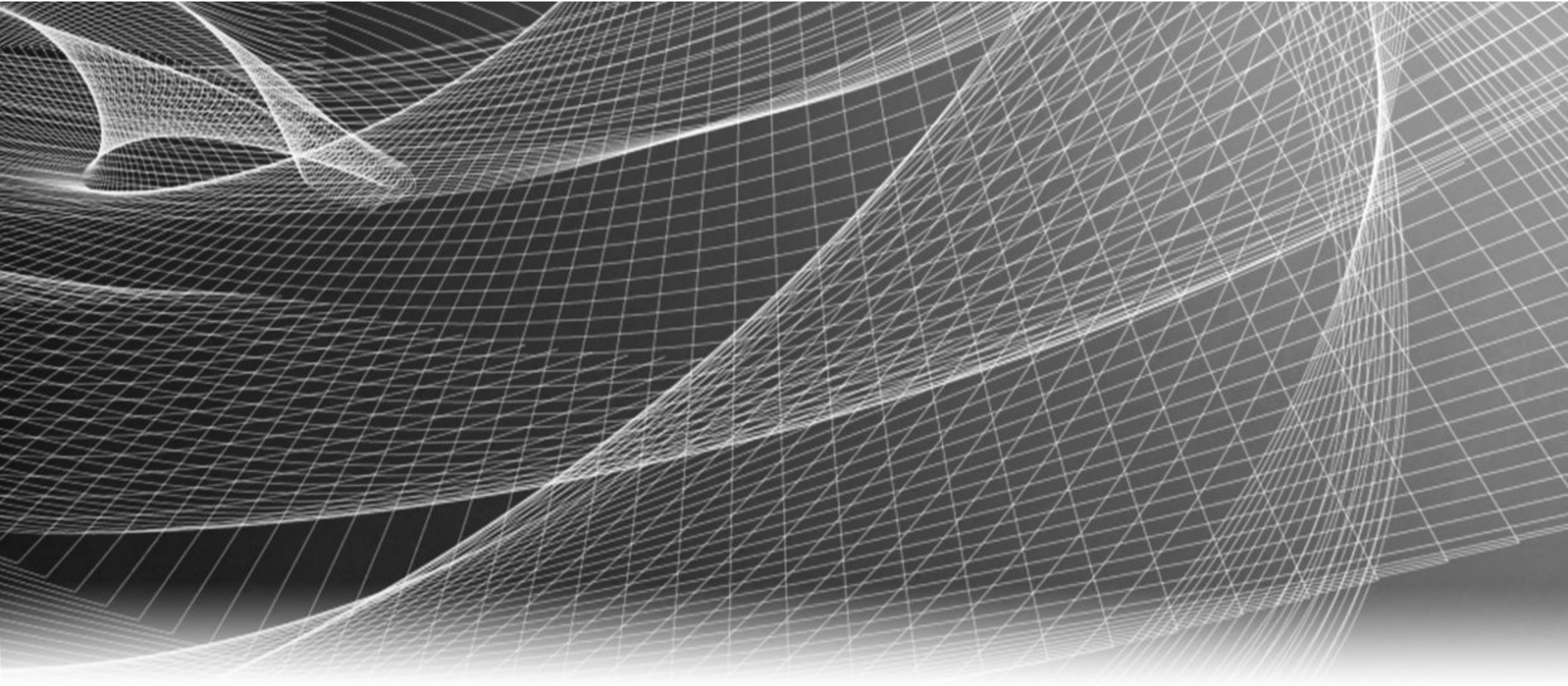

Isilon

OneFS

Version 7.2.0

Web Administration Guide

EMC^2

Published July, 2015

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

The information in this publication is provided as is. EMC Corporation makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose. Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

For the most up-to-date regulatory document for your product line, go to EMC Online Support (https://support.emc.com).

EMC Corporation
Hopkinton, Massachusetts 01748-9103
1-508-435-1000 In North America 1-866-464-7381
www.EMC.com

CONTENTS

Chapter 14 **NDMP backup** 285

CHAPTER 1

Introduction to this guide

This section contains the following topics:

About this guide

This guide describes how the Isilon OneFS web administration interface provides access to cluster configuration, management, and monitoring functionality.

We value your feedback. Please let us know how we can improve this document.

- Take the survey at https://www.research.net/s/isi-docfeedback.
- Send your comments or suggestions to docfeedback@isilon.com.

Isilon scale-out NAS overview

The EMC Isilon scale-out NAS storage platform combines modular hardware with unified software to harness unstructured data. Powered by the OneFS operating system, an EMC Isilon cluster delivers a scalable pool of storage with a global namespace.

The platform's unified software provides centralized web-based and command-line administration to manage the following features:

- A cluster that runs a distributed file system
- Scale-out nodes that add capacity and performance
- Storage options that manage files, block data, and tiering
- Flexible data protection and high availability
- Software modules that control costs and optimize resources

Where to go for support

You can contact EMC Isilon Technical Support for any questions about EMC Isilon products.

Online Support	Live Chat Create a Service Request
Telephone Support	United States: 1-800-SVC-4EMC (800-782-4362) Canada: 800-543-4782 Worldwide: +1-508-497-7901 For local phone numbers in your country, see EMC Customer Support Centers.
Help with online support	For questions specific to EMC Online Support registration or access, email support@emc.com.

CHAPTER 2

Isilon scale-out NAS

This section contains the following topics:

OneFS storage architecture

EMC Isilon takes a scale-out approach to storage by creating a cluster of nodes that runs a distributed file system. OneFS combines the three layers of storage architecture—file system, volume manager, and data protection—into a scale-out NAS cluster.

Each node adds resources to the cluster. Because each node contains globally coherent RAM, as a cluster becomes larger, it becomes faster. Meanwhile, the file system expands dynamically and redistributes content, which eliminates the work of partitioning disks and creating volumes.

Nodes work as peers to spread data across the cluster. Segmenting and distributing data —a process known as striping—not only protects data, but also enables a user connecting to any node to take advantage of the entire cluster's performance.

OneFS uses distributed software to scale data across commodity hardware. Each node helps control data requests, boosts performance, and expands the cluster's capacity. No master device controls the cluster; no slaves invoke dependencies. Instead, each node helps control data requests, boosts performance, and expands the cluster's capacity.

Isilon node components

As a rack-mountable appliance, a storage node includes the following components in a 2U or 4U rack-mountable chassis with an LCD front panel: CPUs, RAM, NVRAM, network interfaces, InfiniBand adapters, disk controllers, and storage media. An Isilon cluster comprises three or more nodes, up to 144.

When you add a node to a cluster, you increase the cluster's aggregate disk, cache, CPU, RAM, and network capacity. OneFS groups RAM into a single coherent cache so that a data request on a node benefits from data that is cached anywhere. NVRAM is grouped to write data with high throughput and to protect write operations from power failures. As the cluster expands, spindles and CPU combine to increase throughput, capacity, and input-output operations per second (IOPS).

EMC Isilon makes several types of nodes, all of which can be added to a cluster to balance capacity and performance with throughput or IOPS:

Node	Use Case
S-Series	IOPS-intensive applications
X-Series	High-concurrency and throughput-driven workflows
NL-Series	Near-primary accessibility, with near-tape value
HD-Series	Maximum capacity

The following EMC Isilon nodes improve performance:

Node	Function
A-Series Performance Accelerator	Independent scaling for high performance
A-Series Backup Accelerator	High-speed and scalable backup-and-restore solution for tape drives over Fibre Channel connections

Internal and external networks

A cluster includes two networks: an internal network to exchange data between nodes and an external network to handle client connections.

Nodes exchange data through the internal network with a proprietary, unicast protocol over InfiniBand. Each node includes redundant InfiniBand ports so you can add a second internal network in case the first one fails.

Clients reach the cluster with 1 GigE or 10 GigE Ethernet. Since every node includes Ethernet ports, the cluster's bandwidth scales with performance and capacity as you add nodes.

Isilon cluster

An Isilon cluster consists of three or more hardware nodes, up to 144. Each node runs the Isilon OneFS operating system, the distributed file-system software that unites the nodes into a cluster. A cluster’s storage capacity ranges from a minimum of 18 TB to a maximum of 15.5 PB.

Cluster administration

OneFS centralizes cluster management through a web administration interface and a command-line interface. Both interfaces provide methods to activate licenses, check the status of nodes, configure the cluster, upgrade the system, generate alerts, view client connections, track performance, and change various settings.

In addition, OneFS simplifies administration by automating maintenance with a job engine. You can schedule jobs that scan for viruses, inspect disks for errors, reclaim disk space, and check the integrity of the file system. The engine manages the jobs to minimize impact on the cluster's performance.

With SNMP versions 2c and 3, you can remotely monitor hardware components, CPU usage, switches, and network interfaces. EMC Isilon supplies management information bases (MIBs) and traps for the OneFS operating system.

OneFS also includes a RESTful application programming interface—known as the Platform API—to automate access, configuration, and monitoring. For example, you can retrieve performance statistics, provision users, and tap the file system. The Platform API integrates with OneFS role-based access control to increase security. See the *Isilon Platform API Reference*.

Quorum

An Isilon cluster must have a quorum to work properly. A quorum prevents data conflicts—for example, conflicting versions of the same file—in case two groups of nodes become unsynchronized. If a cluster loses its quorum for read and write requests, you cannot access the OneFS file system.

For a quorum, more than half the nodes must be available over the internal network. A seven-node cluster, for example, requires a four-node quorum. A 10-node cluster requires a six-node quorum. If a node is unreachable over the internal network, OneFS separates the node from the cluster, an action referred to as splitting. After a cluster is split, cluster operations continue as long as enough nodes remain connected to have a quorum.

In a split cluster, the nodes that remain in the cluster are referred to as the majority group. Nodes that are split from the cluster are referred to as the minority group.

When split nodes can reconnect with the cluster and resynchronize with the other nodes, the nodes rejoin the cluster's majority group, an action referred to as merging.

A OneFS cluster contains two quorum properties:

- read quorum (`efs.gmp.has_quorum`)
- write quorum (`efs.gmp.has_super_block_quorum`)

By connecting to a node with SSH and running the `sysctl` command-line tool as root, you can view the status of both types of quorum. Here is an example for a cluster that has a quorum for both read and write operations, as the command's output indicates with a 1, for true:

```
sysctl efs.gmp.has_quorum
  efs.gmp.has_quorum: 1
sysctl efs.gmp.has_super_block_quorum
  efs.gmp.has_super_block_quorum: 1
```

The degraded states of nodes—such as smartfail, read-only, offline, and so on—affect quorum in different ways. A node in a smartfail or read-only state affects only write quorum. A node in an offline state, however, affects both read and write quorum. In a cluster, the combination of nodes in different degraded states determines whether read requests, write requests, or both work.

A cluster can lose write quorum but keep read quorum. Consider a four-node cluster in which nodes 1 and 2 are working normally. Node 3 is in a read-only state, and node 4 is in a smartfail state. In such a case, read requests to the cluster succeed. Write requests, however, receive an input-output error because the states of nodes 3 and 4 break the write quorum.

A cluster can also lose both its read and write quorum. If nodes 3 and 4 in a four-node cluster are in an offline state, both write requests and read requests receive an input-output error, and you cannot access the file system. When OneFS can reconnect with the nodes, OneFS merges them back into the cluster. Unlike a RAID system, an Isilon node can rejoin the cluster without being rebuilt and reconfigured.

Splitting and merging

Splitting and merging optimize the use of nodes without your intervention.

OneFS monitors every node in a cluster. If a node is unreachable over the internal network, OneFS separates the node from the cluster, an action referred to as splitting. When the cluster can reconnect to the node, OneFS adds the node back into the cluster, an action referred to as merging.

When a node is split from a cluster, it will continue to capture event information locally. You can connect to a split node with SSH and run the `isi events list` command to view the local event log for the node. The local event log can help you troubleshoot the connection issue that resulted in the split. When the split node rejoins the cluster, local events gathered during the split are deleted. You can still view events generated by a split node in the node's event log file located at `/var/log/isi_celog_events.log`.

If a cluster splits during a write operation, OneFS might need to re-allocate blocks for the file on the side with the quorum, which leads allocated blocks on the side without a quorum to become orphans. When the split nodes reconnect with the cluster, the OneFS Collect system job reclaims the orphaned blocks.

Meanwhile, as nodes split and merge with the cluster, the OneFS AutoBalance job redistributes data evenly among the nodes in the cluster, optimizing protection and conserving space.

Storage pools

Storage pools segment nodes and files into logical divisions to simplify the management and storage of data.

A storage pool comprises node pools and tiers. Node pools group equivalent nodes to protect data and ensure reliability. Tiers combine node pools to optimize storage by need, such as a frequently used high-speed tier or a rarely accessed archive.

The SmartPools module groups nodes and files into pools. If you do not activate a SmartPools license, the module provisions node pools and creates one file pool. If you activate the SmartPools license, you receive more features. You can, for example, create multiple file pools and govern them with policies. The policies move files, directories, and file pools among node pools or tiers. You can also define how OneFS handles write operations when a node pool or tier is full. SmartPools reserves a virtual hot spare to reprotect data if a drive fails regardless of whether the SmartPools license is activated.

IP address pools

Within a subnet, you can partition a cluster's external network interfaces into pools of IP address ranges. The pools empower you to customize your storage network to serve different groups of users. Although you must initially configure the default external IP subnet in IPv4 format, you can configure additional subnets in IPv4 or IPv6.

You can associate IP address pools with a node, a group of nodes, or NIC ports. For example, you can set up one subnet for storage nodes and another subnet for accelerator nodes. Similarly, you can allocate ranges of IP addresses on a subnet to different teams, such as engineering and sales. Such options help you create a storage topology that matches the demands of your network.

In addition, network provisioning rules streamline the setup of external connections. After you configure the rules with network settings, you can apply the settings to new nodes.

As a standard feature, the OneFS SmartConnect module balances connections among nodes by using a round-robin policy with static IP addresses and one IP address pool for each subnet. Activating a SmartConnect Advanced license adds features, such as defining IP address pools to support multiple DNS zones.

The OneFS operating system

A distributed operating system based on FreeBSD, OneFS presents an Isilon cluster's file system as a single share or export with a central point of administration.

The OneFS operating system does the following:

- Supports common data-access protocols, such as SMB and NFS.
- Connects to multiple identity management systems, such as Active Directory and LDAP.
- Authenticates users and groups.
- Controls access to directories and files.

Data-access protocols

With the OneFS operating system, you can access data with multiple file-sharing and transfer protocols. As a result, Microsoft Windows, UNIX, Linux, and Mac OS X clients can share the same directories and files.

OneFS supports the following protocols.

SMB
: The Server Message Block (SMB) protocol enables Windows users to access the cluster. OneFS works with SMB 1, SMB 2, and SMB 2.1, as well as SMB 3.0 for Multichannel only. With SMB 2.1, OneFS supports client opportunity locks (oplocks) and large (1 MB) MTU sizes. The default file share is `/ifs`.

NFS
: The Network File System (NFS) protocol enables UNIX, Linux, and Mac OS X systems to remotely mount any subdirectory, including subdirectories created by Windows users. OneFS works with NFS versions 3 and 4. The default export is `/ifs`.

HDFS
: The Hadoop Distributed File System (HDFS) protocol enables a cluster to work with Apache Hadoop, a framework for data-intensive distributed applications. HDFS integration requires you to activate a separate license.

FTP
: FTP allows systems with an FTP client to connect to the cluster and exchange files.

HTTP
: HTTP gives systems browser-based access to resources. OneFS includes limited support for WebDAV.

Identity management and access control

OneFS works with multiple identity management systems to authenticate users and control access to files. In addition, OneFS features access zones that allow users from different directory services to access different resources based on their IP address. Role-based access control, meanwhile, segments administrative access by role.

OneFS authenticates users with the following identity management systems:

- Microsoft Active Directory (AD)
- Lightweight Directory Access Protocol (LDAP)
- Network Information Service (NIS)
- Local users and local groups
- A file provider for accounts in `/etc/spwd.db` and `/etc/group` files. With the file provider, you can add an authoritative third-party source of user and group information.

You can manage users with different identity management systems; OneFS maps the accounts so that Windows and UNIX identities can coexist. A Windows user account managed in Active Directory, for example, is mapped to a corresponding UNIX account in NIS or LDAP.

To control access, an Isilon cluster works with both the access control lists (ACLs) of Windows systems and the POSIX mode bits of UNIX systems. When OneFS must transform a file's permissions from ACLs to mode bits or from mode bits to ACLs, OneFS merges the permissions to maintain consistent security settings.

OneFS presents protocol-specific views of permissions so that NFS exports display mode bits and SMB shares show ACLs. You can, however, manage not only mode bits but also

ACLs with standard UNIX tools, such as the `chmod` and `chown` commands. In addition, ACL policies enable you to configure how OneFS manages permissions for networks that mix Windows and UNIX systems.

Access zones

OneFS includes an access zones feature. Access zones allow users from different authentication providers, such as two untrusted Active Directory domains, to access different OneFS resources based on an incoming IP address. An access zone can contain multiple authentication providers and SMB namespaces.

RBAC for administration

OneFS includes role-based access control (RBAC) for administration. In place of a root or administrator account, RBAC lets you manage administrative access by role. A role limits privileges to an area of administration. For example, you can create separate administrator roles for security, auditing, storage, and backup.

Structure of the file system

OneFS presents all the nodes in a cluster as a global namespace—that is, as the default file share, `/ifs`.

In the file system, directories are inode number links. An inode contains file metadata and an inode number, which identifies a file's location. OneFS dynamically allocates inodes, and there is no limit on the number of inodes.

To distribute data among nodes, OneFS sends messages with a globally routable block address through the cluster's internal network. The block address identifies the node and the drive storing the block of data.

Note

It is recommended that you do not save data to the root `/ifs` file path but in directories below `/ifs`. The design of your data storage structure should be planned carefully. A well-designed directory optimizes cluster performance and cluster administration.

Data layout

OneFS evenly distributes data among a cluster's nodes with layout algorithms that maximize storage efficiency and performance. The system continuously reallocates data to conserve space.

OneFS breaks data down into smaller sections called blocks, and then the system places the blocks in a stripe unit. By referencing either file data or erasure codes, a stripe unit helps safeguard a file from a hardware failure. The size of a stripe unit depends on the file size, the number of nodes, and the protection setting. After OneFS divides the data into stripe units, OneFS allocates, or stripes, the stripe units across nodes in the cluster.

When a client connects to a node, the client's read and write operations take place on multiple nodes. For example, when a client connects to a node and requests a file, the node retrieves the data from multiple nodes and rebuilds the file. You can optimize how OneFS lays out data to match your dominant access pattern—concurrent, streaming, or random.

Writing files

On a node, the input-output operations of the OneFS software stack split into two functional layers: A top layer, or initiator, and a bottom layer, or participant. In read and write operations, the initiator and the participant play different roles.

When a client writes a file to a node, the initiator on the node manages the layout of the file on the cluster. First, the initiator divides the file into blocks of 8 KB each. Second, the initiator places the blocks in one or more stripe units. At 128 KB, a stripe unit consists of 16 blocks. Third, the initiator spreads the stripe units across the cluster until they span a width of the cluster, creating a stripe. The width of the stripe depends on the number of nodes and the protection setting.

After dividing a file into stripe units, the initiator writes the data first to non-volatile random-access memory (NVRAM) and then to disk. NVRAM retains the information when the power is off.

During the write transaction, NVRAM guards against failed nodes with journaling. If a node fails mid-transaction, the transaction restarts without the failed node. When the node returns, it replays the journal from NVRAM to finish the transaction. The node also runs the AutoBalance job to check the file's on-disk striping. Meanwhile, uncommitted writes waiting in the cache are protected with mirroring. As a result, OneFS eliminates multiple points of failure.

Reading files

In a read operation, a node acts as a manager to gather data from the other nodes and present it to the requesting client.

Because an Isilon cluster's coherent cache spans all the nodes, OneFS can store different data in each node's RAM. By using the internal InfiniBand network, a node can retrieve file data from another node's cache faster than from its own local disk. If a read operation requests data that is cached on any node, OneFS pulls the cached data to serve it quickly.

In addition, for files with an access pattern of concurrent or streaming, OneFS pre-fetches in-demand data into a managing node's local cache to further improve sequential-read performance.

Metadata layout

OneFS protects metadata by spreading it across nodes and drives.

Metadata—which includes information about where a file is stored, how it is protected, and who can access it—is stored in inodes and protected with locks in a B+ tree, a standard structure for organizing data blocks in a file system to provide instant lookups. OneFS replicates file metadata across the cluster so that there is no single point of failure.

Working together as peers, all the nodes help manage metadata access and locking. If a node detects an error in metadata, the node looks up the metadata in an alternate location and then corrects the error.

Locks and concurrency

OneFS includes a distributed lock manager that orchestrates locks on data across all the nodes in a cluster.

The lock manager grants locks for the file system, byte ranges, and protocols, including SMB share-mode locks and NFS advisory locks. OneFS also supports SMB opportunistic locks and NFSv4 delegations.

Because OneFS distributes the lock manager across all the nodes, any node can act as a lock coordinator. When a thread from a node requests a lock, the lock manager's hashing algorithm typically assigns the coordinator role to a different node. The coordinator allocates a shared lock or an exclusive lock, depending on the type of request. A shared lock allows users to share a file simultaneously, typically for read operations. An exclusive lock allows only one user to access a file, typically for write operations.

Striping

In a process known as striping, OneFS segments files into units of data and then distributes the units across nodes in a cluster. Striping protects your data and improves cluster performance.

To distribute a file, OneFS reduces it to blocks of data, arranges the blocks into stripe units, and then allocates the stripe units to nodes over the internal network.

At the same time, OneFS distributes erasure codes that protect the file. The erasure codes encode the file's data in a distributed set of symbols, adding space-efficient redundancy. With only a part of the symbol set, OneFS can recover the original file data.

Taken together, the data and its redundancy form a protection group for a region of file data. OneFS places the protection groups on different drives on different nodes—creating data stripes.

Because OneFS stripes data across nodes that work together as peers, a user connecting to any node can take advantage of the entire cluster's performance.

By default, OneFS optimizes striping for concurrent access. If your dominant access pattern is streaming--that is, lower concurrency, higher single-stream workloads, such as with video--you can change how OneFS lays out data to increase sequential-read performance. To better handle streaming access, OneFS stripes data across more drives. Streaming is most effective on clusters or subpools serving large files.

Data protection overview

An Isilon cluster is designed to serve data even when components fail. By default, OneFS protects data with erasure codes, enabling you to retrieve files when a node or disk fails. As an alternative to erasure codes, you can protect data with two to eight mirrors.

When you create a cluster with five or more nodes, erasure codes deliver as much as 80 percent efficiency. On larger clusters, erasure codes provide as much as four levels of redundancy.

In addition to erasure codes and mirroring, OneFS includes the following features to help protect the integrity, availability, and confidentiality of data:

Feature	Description
Antivirus	OneFS can send files to servers running the Internet Content Adaptation Protocol (ICAP) to scan for viruses and other threats.

Feature	Description
Clones	OneFS enables you to create clones that share blocks with other files to save space.
NDMP backup and restore	OneFS can back up data to tape and other devices through the Network Data Management Protocol. Although OneFS supports both NDMP 3-way and 2-way backup, 2-way backup requires an Isilon Backup Accelerator node.
Protection domains	You can apply protection domains to files and directories to prevent changes.

The following software modules also help protect data, but they require you to activate a separate license:

Licensed Feature	Description
SyncIQ	SyncIQ replicates data on another Isilon cluster and automates failover and failback operations between clusters. If a cluster becomes unusable, you can fail over to another Isilon cluster.
SnapshotIQ	You can protect data with a snapshot—a logical copy of data stored on a cluster.
SmartLock	The SmartLock tool prevents users from modifying and deleting files. You can commit files to a write-once, read-many state: The file can never be modified and cannot be deleted until after a set retention period. SmartLock can help you comply with Securities and Exchange Commission Rule 17a-4.

N+M data protection

OneFS supports N+M erasure code levels of N+1, N+2, N+3, and N+4.

In the N+M data model, N represents the number of nodes, and M represents the number of simultaneous failures of nodes or drives that the cluster can handle without losing data. For example, with N+2 the cluster can lose two drives on different nodes or lose two nodes.

To protect drives and nodes separately, OneFS also supports N+M:B. In the N+M:B notation, M is the number of disk failures, and B is the number of node failures. With N +3:1 protection, for example, the cluster can lose three drives or one node without losing data.

The default protection level for clusters larger than 18 TB is N+2:1. The default for clusters smaller than 18 TB is N+1.

The quorum rule dictates the number of nodes required to support a protection level. For example, N+3 requires at least seven nodes so you can maintain a quorum if three nodes fail.

You can, however, set a protection level that is higher than the cluster can support. In a four-node cluster, for example, you can set the protection level at 5x. OneFS protects the data at 4x until a fifth node is added, after which OneFS automatically reprotects the data at 5x.

Data mirroring

You can protect on-disk data with mirroring, which copies data to multiple locations. OneFS supports two to eight mirrors. You can use mirroring instead of erasure codes, or you can combine erasure codes with mirroring.

Mirroring, however, consumes more space than erasure codes. Mirroring data three times, for example, duplicates the data three times, which requires more space than erasure codes. As a result, mirroring suits transactions that require high performance.

You can also mix erasure codes with mirroring. During a write operation, OneFS divides data into redundant protection groups. For files protected by erasure codes, a protection group consists of data blocks and their erasure codes. For mirrored files, a protection group contains all the mirrors of a set of blocks. OneFS can switch the type of protection group as it writes a file to disk. By changing the protection group dynamically, OneFS can continue writing data despite a node failure that prevents the cluster from applying erasure codes. After the node is restored, OneFS automatically converts the mirrored protection groups to erasure codes.

The file system journal

A journal, which records file-system changes in a battery-backed NVRAM card, recovers the file system after failures, such as a power loss. When a node restarts, the journal replays file transactions to restore the file system.

Virtual hot spare

When a drive fails, OneFS uses space reserved in a subpool instead of a hot spare drive. The reserved space is known as a virtual hot spare.

In contrast to a spare drive, a virtual hot spare automatically resolves drive failures and continues writing data. If a drive fails, OneFS migrates data to the virtual hot spare to reprotect it. You can reserve as many as four disk drives as a virtual hot spare.

Balancing protection with storage space

You can set protection levels to balance protection requirements with storage space.

Higher protection levels typically consume more space than lower levels because you lose an amount of disk space to storing erasure codes. The overhead for the erasure codes depends on the protection level, the file size, and the number of nodes in the cluster. Since OneFS stripes both data and erasure codes across nodes, the overhead declines as you add nodes.

VMware integration

OneFS integrates with several VMware products, including vSphere, vCenter, and ESXi.

For example, OneFS works with the VMware vSphere API for Storage Awareness (VASA) so that you can view information about an Isilon cluster in vSphere. OneFS also works with the VMware vSphere API for Array Integration (VAAI) to support the following features for block storage: hardware-assisted locking, full copy, and block zeroing. VAAI for NFS requires an ESXi plug-in.

With the Isilon for vCenter plug-in, you can backup and restore virtual machines on an Isilon cluster. With the Isilon Storage Replication Adapter, OneFS integrates with the

VMware vCenter Site Recovery Manager to recover virtual machines that are replicated between Isilon clusters.

Software modules

You can access advanced features by activating licenses for EMC Isilon software modules.

SmartLock

SmartLock protects critical data from malicious, accidental, or premature alteration or deletion to help you comply with SEC 17a-4 regulations. You can automatically commit data to a tamper-proof state and then retain it with a compliance clock.

SyncIQ automated failover and failback

SyncIQ replicates data on another Isilon cluster and automates failover and failback between clusters. If a cluster becomes unusable, you can fail over to another Isilon cluster. Failback restores the original source data after the primary cluster becomes available again.

File clones

OneFS provides provisioning of full read/write copies of files, LUNs, and other clones. OneFS also provides virtual machine linked cloning through VMware API integration.

SnapshotIQ

SnapshotIQ protects data with a snapshot—a logical copy of data stored on a cluster. A snapshot can be restored to its top-level directory.

SmartPools

SmartPools enable you to create multiple file pools governed by file-pool policies. The policies move files and directories among node pools or tiers. You can also define how OneFS handles write operations when a node pool or tier is full.

SmartConnect

If you activate a SmartConnect Advanced license, you can balance policies to evenly distribute CPU usage, client connections, or throughput. You can also define IP address pools to support multiple DNS zones in a subnet. In addition, SmartConnect supports IP failover, also known as NFS failover.

InsightIQ

The InsightIQ virtual appliance monitors and analyzes the performance of your Isilon cluster to help you optimize storage resources and forecast capacity.

Aspera for Isilon

Aspera moves large files over long distances fast. Aspera for Isilon is a cluster-aware version of Aspera technology for non-disruptive, wide-area content delivery.

HDFS

OneFS works with the Hadoop Distributed File System protocol to help clients running Apache Hadoop, a framework for data-intensive distributed applications, analyze big data.

SmartQuotas

The SmartQuotas module tracks disk usage with reports and enforces storage limits with alerts.

CHAPTER 3

General cluster administration

This section contains the following topics:

General cluster administration overview

You can manage general OneFS settings and module licenses for the EMC Isilon cluster.

General cluster administration covers several areas. You can manage general settings such as cluster name, date and time, and email. You can monitor the cluster status and performance, including hardware components. You can configure how events and notifications are handled, and you can perform cluster maintenance such as adding, removing, and restarting nodes.

Most management tasks are accomplished through both the web administration or command-line interface; however, you will occasionally encounter a task that can only be managed by one or the other.

User interfaces

OneFS provides several interfaces for managing the EMC Isilon cluster.

Interface	Description	Comment
OneFS web administration interface	The browser-based OneFS web administration interface provides secure access with OneFS-supported browsers. Use this interface to view robust graphical monitoring displays and to perform cluster-management tasks.	The OneFS web administration interface uses port 8080 as its default port.
OneFS command-line interface	Run OneFS `isi` commands in the command-line interface to configure, monitor, and manage the cluster. Access to the command-line interface is through a secure shell (SSH) connection to any node in the cluster.	The OneFS command-line interface provides an extended standard UNIX command set for managing the cluster.
OneFS API	The OneFS application programming interface (API) is divided into two functional areas: one area enables cluster configuration, management, and monitoring functionality, and the other area enables operations on files and directories on the cluster. You can send requests to the OneFS API through a Representational State Transfer (REST) interface, which is accessed through resource URIs and standard HTTP methods.	You should have a solid understanding of HTTP/1.1 and experience writing HTTP-based client software before you implement client-based software through the OneFS API.
Node front panel	With the exception of accelerator nodes, the front panel of each node contains an LCD screen with five buttons that you can use to monitor node and cluster details.	Node status, events, cluster details, capacity, IP and MAC addresses, throughput, and drive status are available through the node front panel.

Connecting to the cluster

EMC Isilon cluster access is provided through the web administration interface or through SSH. You can use a serial connection to perform cluster-administration tasks through the command-line interface.

You can also access the cluster through the node front panel to accomplish a subset of cluster-management tasks. For information about connecting to the node front panel, see the installation documentation for your node.

Log in to the web administration interface

You can monitor and manage your EMC Isilon cluster from the browser-based web administration interface.

Procedure

1. Open a browser window and type the URL for your cluster in the address field, replacing *<yourNodeIPaddress>* in the following example with the first IP address you provided when you configured ext-1:

   https://*<yourNodeIPaddress>*:8080

   The system displays a message if your security certificates have not been configured. Resolve any certificate configurations and continue to the web site.

2. Log in to OneFS by typing your OneFS credentials in the **Username** and **Password** fields.

   After you log into the web administration interface, there is a 4-hour login timeout and a 24-hour session inactivity timeout.

Open an SSH connection to a cluster

You can use any SSH client such as OpenSSH or PuTTY to connect to an EMC Isilon cluster.

Before you begin

You must have valid OneFS credentials to log in to a cluster after the connection is open.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster, using the IP address of the node and port number 22.
2. Log in with your OneFS credentials.

   At the OneFS command line prompt, you can use `isi` commands to monitor and manage your cluster.

Licensing

Advanced cluster features are available when you activate licenses for OneFS software modules. Each optional OneFS software module requires you to activate a separate license.

For more information about the following optional software modules, contact your EMC Isilon sales representative.

- HDFS

- InsightIQ
- Isilon for vCenter
- SmartConnect Advanced
- SmartDedupe
- SmartLock
- SmartPools
- SmartQuotas
- SnapshotIQ
- SyncIQ

License status

The status of a OneFS module license indicates whether the functionality provided by a module is available on the cluster.

Licenses exist in one of the following states:

Status	Description
Inactive	The license has not been activated on the cluster. You cannot access the features provided by the corresponding module.
Evaluation	The license has been temporarily activated on the cluster. You can access the features provided by the corresponding module for a limited period of time. After the license expires, the features become unavailable unless the license is reactivated.
Activated	The license has been activated on the cluster. You can access the features provided by the corresponding module.
Expired	The evaluation license has expired on the cluster. You can no longer access the features provided by the corresponding module. The features will remain unavailable unless you reactivate the license.

The following table describes what functionality is available for each license depending on the license's status:

License	Inactive	Evaluation/ Activated	Expired
HDFS	Clients cannot access the cluster through HDFS.	You can configure HDFS settings and clients can access the cluster through HDFS.	You cannot configure HDFS settings. After the HDFS service restarts, clients can no longer access the cluster through HDFS.
InsightIQ	You cannot monitor the cluster with InsightIQ.	You can monitor the cluster with InsightIQ.	InsightIQ stops monitoring the cluster. Data previously collected by InsightIQ is still available on the InsightIQ instance.
Isilon for vCenter	You cannot back up virtual machines that are stored on an	You can back up virtual machines that are stored on an	You cannot create new backups of virtual machines

License	Inactive	Evaluation/ Activated	Expired
	Isilon cluster with Isilon for vCenter.	Isilon cluster with Isilon for vCenter.	that are stored on an Isilon cluster.
SmartPools	All files belong to the default file pool and are governed by the default file pool policy. Virtual hot spare allocation, which reserves space for data repair if a drive fails, is also available.	You can create multiple file pools and file pool policies. You can also manage spillover, which defines how write operations are handled when a storage pool is not writable.	You can no longer manage file pool policies, and the SmartPools job will no longer run. Newly added files will be governed by the default file pool policy, and the SetProtectPlus job will eventually apply the default file pool policy to all files in the cluster. If the SmartPools job is running when the license expires, the job completes before becoming disabled.
SmartConnect Advanced	Client connections are balanced by using a round robin policy. IP address allocation is static. Each external network subnet can be assigned only one IP address pool.	You can access features such as CPU utilization, connection counting, and client connection policies in addition to the round robin policy. You can also configure address pools to support multiple DNS zones within a single subnet, and support IP failover.	You can no longer specify SmartConnect Advanced settings.
SmartDedupe	You cannot deduplicate data with SmartDedupe.	You can deduplicate data with SmartDedupe.	You can no longer deduplicate data. Previously deduplicated data remains deduplicated.
SmartLock	You cannot enforce file retention with SmartLock.	You can enforce file retention with SmartLock.	You cannot create new SmartLock directories or modify SmartLock directory configuration settings for existing directories. You can still commit files to a write once read many (WORM) state, even after the SmartLock license is unconfigured, but you cannot delete WORM-committed files from enterprise directories.

License	Inactive	Evaluation/ Activated	Expired
SnapshotIQ	You can view and manage snapshots generated by OneFS applications. However, you cannot create snapshots or configure SnapshotIQ settings.	You can create, view, and manage snapshots. You can also configure snapshot settings.	You will no longer be able to generate snapshots. Existing snapshot schedules are not deleted; however, the schedules will not generate snapshots. You can still delete snapshots and access snapshot data.
SmartQuotas	You cannot create quotas with SmartQuotas.	You can create quotas with SmartQuotas.	OneFS disables all quotas. Exceeding advisory and soft thresholds does not trigger events. Hard and soft thresholds are not enforced.
SyncIQ	You cannot replicate data with SyncIQ.	You can replicate data with SyncIQ	You will no longer be able to replicate data to remote clusters, and remote clusters will not be able to replicate data to the local cluster. Replication policies will still display a status of enabled; however, future replication jobs created by the policy will fail. If a replication job is in progress when the license expires, the job completes.

License configuration

You can configure or unconfigure some OneFS module licenses.

You can configure a license by performing specific operations through the corresponding module. Not all actions that require you to activate a license will configure the license. Also, not all licenses can be configured. Configuring a license does not add or remove access to any features provided by a module.

You can unconfigure a license only through the `isi license unconfigure` command. You may want to unconfigure a license for a OneFS software module if, for example, you enabled an evaluation version of a module but later decided not to purchase a permanent license. Unconfiguring a module license does not deactivate the license. Unconfiguring a license does not add or remove access to any features provided by a module.

The following table describes both the actions that cause each license to be configured and the results of unconfiguring each license:

License	Cause of configuring	Result of unconfiguring
HDFS	Cannot configure this license.	No system impact.
InsightIQ	Cannot configure this license.	No system impact.

License	Cause of configuring	Result of unconfiguring
Isilon for vCenter	Cannot configure this license.	No system impact.
SmartPools	Create a file pool policy (other than the default file pool policy).	OneFS deletes all file pool policies (except the default file pool policy).
SmartConnect	Configure SmartConnect Advanced settings for at least one IP address pool.	OneFS converts dynamic IP address pools to static IP address pools.
SmartDedupe	Cannot configure this license.	No system impact.
SmartLock	Cannot configure this license.	No system impact.
SnapshotIQ	Create a snapshot schedule.	Deletes all snapshot schedules.
SmartQuotas	Create a quota.	No system impact.
SyncIQ	Create a replication policy.	No system impact.

Activate a license

To access a OneFS module, you must activate a license.

Before you begin

Before you can activate a license, you must obtain a valid license key, and you must have root user privileges on your cluster. To obtain a license key, contact your EMC Isilon sales representative.

Procedure

1. Click **Help** › **About This Cluster**.
2. In the **Licensed Modules** section, click **Activate license**.
3. In the **License key** field, type the license key for the module that you want to enable.
4. Read the end user license agreement, click **I have read and agree**, and then click **Submit**.

View license information

You can view information about the current status of any optional Isilon software modules.

Procedure

1. Click **Help** › **About This Cluster**.
2. In the **Licensed Modules** area, review information about licenses, including status and expiration date.

Unconfigure a license

You can unconfigure a licensed module through the command-line interface.

You must have root user privileges on your Isilon cluster to unconfigure a module license. This procedure is available only through the command-line interface (CLI).

Note

Unconfiguring a license does not deactivate the license.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster.

   You must log in as root.

2. Run the `isi license unconfigure` command.

   The following command unconfigures the license for SmartConnect:

   ```
   isi license unconfigure -m smartconnect
   ```

   If you do not know the module name, run the `isi license` command for a list of OneFS modules and their status.

   OnesFS returns a confirmation message similar to the following text: `The SmartConnect module has been unconfigured`. The license is unconfigured, and any processes enabled for the module are disabled.

Certificates

You can renew the Secure Sockets Layer (SSL) certificate for the Isilon web administration interface or replace it with a third-party SSL certificate.

All Platform API communication, which includes communication through the web administration interface, is over SSL. You can replace or renew the self-signed certificate with a certificate that you generate. To replace or renew an SSL certificate, you must be logged in as root.

Replace or renew the SSL certificate

You can replace or renew the Secure Sockets Layer (SSL) certificate, which is used to access the EMC Isilon cluster through a browser.

Before you begin

When you renew or replace a self-signed SSL certificate, you must provide information for your organization in the format that is described in the Self-signed SSL certificate data example.

The following folders are the default locations for the `server.crt` and `server.key` files in OneFS 6.0 and higher.

- SSL certificate: `/usr/local/apache2/conf/ssl.crt/server.crt`
- SSL certificate key: `/usr/local/apache2/conf/ssl.key/server.key`

Procedure

1. Establish an SSH connection to any node in the cluster.
2. At the command prompt, run the following command to create the appropriate directory.

   ```
   mkdir /ifs/local/
   ```

3. At the command prompt, run the following command to change to the directory.

   ```
   cd /ifs/local/
   ```

4. Choose the type of certificate you want to install.

Option	Description

Third-party (public or private) CA-issued certificate	a. At the command prompt, run the following command to generate a new Certificate Signing Request (CSR) in addition to a new key, where *<common_name>* is the host name, such as isilon.example.com:
	`openssl req -new -nodes -newkey rsa:1024 -keyout <common name>.key \`
	`  -out <common-name>.csr`
	b. Send the contents of the *<common_name>*`.csr` file from the cluster to your Certificate Authority (CA) for signing. When you receive the signed certificate (now a `.crt` file) from the CA, copy the certificate to `/ifs/local/<common-name>.crt`.

Self-signed certificate based on the existing (stock) ssl.key	a. At the command prompt, run the following command to create a two-year certificate. Increase or decrease the value for `-days` to generate a certificate with a different expiration date.
	`cp /usr/local/apache2/conf/ssl.key/server.key ./openssl req -new \/`
	`  -days 730 -nodes -x509 -key server.key -out server.crt`

A renewal certificate is created, based on the existing (stock) `ssl.key` file.

5. (Optional) At the command prompt, run the following command to verify the attributes in an SSL certificate.

```
openssl x509 -text -noout -in <common-name>.crt
```

6. Run the following commands to install the certificate and key:

```
isi services -a isi_webui disable
chmod 640 <common name>.key
isi_for_array -s 'cp /ifs/local/<common-name>.key /usr/local/
apache2/conf/ssl.key/server.key'
isi_for_array -s 'cp /ifs/local/<common-name>.crt /usr/local/
apache2/conf/ssl.crt/server.crt'
isi services -a isi_webui enable
```

7. Run the following command to remove the files in `/ifs/local`.

```
rm /ifs/local/*
```

Verify an SSL certificate update

You can verify the details stored in a Secure Sockets Layer (SSL) certificate.

Procedure

1. Open a web browser window.
2. Browse to `https://<common name>:8080`, where *<common name>* is the host name for the EMC Isilon web administration interface, such as isilon.example.com.
3. In the security details for the web page, verify that the subject line and other details that you provided are correct.

Note

The steps to view security details vary by browser. For example, in some browsers, you can click the padlock icon in the address bar to view the security details for the web page. Follow the steps that are specific to your browser.

Self-signed SSL certificate data example

Self-signed SSL certificate renewal or replacement requires you to provide data such as your fully qualified domain name and a contact email address.

When you renew or replace a self-signed SSL certificate, you are asked to provide data in the format shown in the following example. Some fields in the certificate file contain a default value. If you type `'.'`, the field is left blank when the certificate is generated.

- Country Name (2 letter code) [XX]:**`US`**
- State or Province Name (full name) [Some-State]:**`Washington`**
- Locality Name (for example, city) [default city]:**`Seattle`**
- Organization Name (for example, company) [Internet Widgits Pty Ltd]:**`Isilon`**
- Organizational Unit Name (for example, section) []:**`Support`**
- Common Name (for example, server FQDN or server name) []:**`isilon.example.com`**
- Email Address []:**`support@example.com`**

In addition, you should add the following attributes to be sent with your certificate request:

- Challenge password []:**`Isilon1`**
- Optional company name []:

Cluster identity

You can specify identity attributes for the EMC Isilon cluster.

Cluster name
: The cluster name appears on the login page, and it makes the cluster and its nodes more easily recognizable on your network. Each node in the cluster is identified by the cluster name plus the node number. For example, the first node in a cluster named Images may be named Images-1.

Cluster description
: The cluster description appears below the cluster name on the login page. The cluster description is useful if your environment has multiple clusters.

Login message
: The login message appears as a separate box on the login page. The login message can convey cluster information, login instructions, or warnings that a user should know before logging into the cluster.

Set the cluster name

You can specify a name, description, and login message to your EMC Isilon cluster.

Cluster names must begin with a letter and can contain only numbers, letters, and hyphens. If the cluster is joined to an Active Directory domain, the cluster name must be 11 characters or fewer.

Procedure

1. Click **Cluster Management** › **General Settings** › **Cluster Identity**.
2. (Optional) In the **Cluster Name and Description** area, type a name for the cluster in the **Cluster Name** field and type a description in the **Cluster Description** field.
3. (Optional) In the **Login Message** area, type a title in the **Message Title** field and a message in the **Message Body** field.
4. Click **Submit**.

After you finish

You must add the cluster name to your DNS servers.

Cluster contact information

Isilon Technical Support personnel and event notification recipients will communicate with the specified contacts.

You can specify the following contact information for your EMC Isilon cluster:

- Company name and location
- Primary and secondary contact names
- Phone number and email address for each contact

Specify contact information

You can specify contact information so that Isilon Technical Support personnel and event notification recipients can contact you.

Procedure

1. Click **Cluster Management** › **General Settings** › **Cluster Identity**.
2. In the **Contact Information** area, type the name and contact information in the fields for those details.
3. Click **Submit**.

   To modify this information, you can access the aforementioned screen, or you can go to the **Contact Information** area of **Dashboard** › **Events** › **Notification Settings**.

Cluster date and time

The Network Time Protocol (NTP) service is configurable manually, so you can ensure that all nodes in a cluster are synchronized to the same time source.

The NTP method automatically synchronizes cluster date and time settings through an NTP server. Alternatively, you can set the date and time reported by the cluster by manually configuring the service.

Windows domains provide a mechanism to synchronize members of the domain to a master clock running on the domain controllers, so OneFS adjusts the cluster time to that of Active Directory with a service. If there are no external NTP servers configured, OneFS uses the Windows domain controller as the NTP time server. When the cluster and domain time become out of sync by more than 4 minutes, OneFS generates an event notification.

Note

If the cluster and Active Directory become out of sync by more than 5 minutes, authentication will not work.

Set the cluster date and time

You can set the date, time, and time zone that is used by the EMC Isilon cluster.

Procedure

1. Click **Cluster Management** › **General Settings** › **Date & Time**.

   The **Date and Time** page displays a list of each node's IP address and the date and time settings for each node.
2. From the **Date and time** lists, select the month, date, year, hour, and minute settings.
3. From the **Time zone** list, select a value.

   If the time zone that you want is not in the list, select **Advanced** from the **Time zone** list, and then select the time zone from the **Advanced time zone** list.
4. Click **Submit**.

Specify an NTP time server

You can specify one or more Network Time Protocol (NTP) servers to synchronize the system time on the EMC cluster. The cluster periodically contacts the NTP servers and sets the date and time based on the information it receives.

Procedure

1. Click **Cluster Management** › **General Settings** › **NTP**.
2. (Optional) Add a server.
   a. In the **Server IP or hostname** field, type the host name or IP address of the NTP server, click **Add**, and then click **Submit**.
   b. (Optional) To enable NTP authentication with a keyfile, type the path and file name in the **Keyfile** field, and then click **Submit**.
3. (Optional) Delete a server.
   a. Select the check box next to the server name in the **Server** list for each server that you want to delete.
   b. Click **Delete**.
   c. Click **Submit**.

SMTP email settings

If your network environment requires the use of an SMTP server or if you want to route EMC Isilon cluster event notifications with SMTP through a port, you can configure SMTP email settings.

SMTP settings include the SMTP relay address and port number that email is routed through. You can specify an origination email and subject line for all event notification emails sent from the cluster.

If your SMTP server is configured to support authentication, you can specify a username and password. You can also specify whether to apply encryption to the connection.

Configure SMTP email settings

You can send event notifications through an SMTP mail server. You can also enable SMTP authentication if your SMTP server is configured to support it.

Procedure

1. Click **Cluster Management** › **General Settings** › **Email Settings**.
2. From the **Email Settings** area, in the **SMTP relay address** field, type the fully qualified domain name or the IP address of the SMTP relay..
3. In the **SMTP relay port** field, type the port number.

   The default port number is 25.
4. In the **Send email as** field, type the originating email address that will be displayed in the To line of the email.
5. In the **Subject** field, type the text that will be displayed in the Subject line of the email.
6. (Optional) To require SMTP authentication, from the **Use SMTP AUTH** area, select **Yes**.
   a. In the **Username**, **Password**, and **Confirm password** fields, type the required user credentials.
   b. From the **Connection security** area, select one of the following options:
   - To specify no encryption of the connection, select **No security**.
   - To specify TLS encryption of the connection, select **STARTTLS**.
7. Click **Submit**.

   You can test your configuration by sending a test event notification.

Configuring the cluster join mode

The join mode specifies how a node is added to the EMC Isilon cluster and whether authentication is required. OneFS supports manual and secure join modes for adding nodes to the EMC Isilon cluster.

Mode	Description
Manual	Allows you to manually add a node to the cluster without requiring authorization.
Secure	Requires authorization of every node added to the cluster and the node must be added through the web administration interface or through the `isi devices -a add -d <unconfigured_node_serial_no>` command in the command-line interface. **Note** If you specify a secure join mode, you cannot join a node to the cluster through serial console wizard option `[2] Join an existing cluster`.

Specify the cluster join mode

You can specify a join mode that determines how nodes are added to the EMC Isilon cluster.

Procedure

1. Click **Cluster Management** › **General Settings** › **Join Mode**.
2. In the **Settings** area, select the mode that will determine how nodes can be added to the cluster.
   - **Manual**— joins can be manually initiated.
   - **Secure**—joins can be initiated only by the cluster and require authentication.
3. Click **Submit**.

File system settings

You can configure global file system settings on an EMC Isilon cluster pertaining to access time tracking and character encoding.

You can enable or disable access time tracking, which monitors the time of access on each file. If necessary, you can also change the default character encoding on the cluster.

Enable or disable access time tracking

You can enable access time tracking to support features that require it.

By default, the EMC Isilon cluster does not track the timestamp when files are accessed. You can enable this feature to support OneFS features that use it. For example, access-time tracking must be enabled to configure SyncIQ policy criteria that match files based on when they were last accessed.

Note

Enabling access-time tracking may affect cluster performance.

Procedure

1. Click **File System Management** › **File System Settings** › **Access Time Tracking**.
2. In the **Access Time Tracking** area, select a configuration option.
   - To enable access time tracking, click **Enabled**, and then specify in the **Precision** fields how often to update the last-accessed time by typing a numeric value and by selecting a unit of measure, such as Seconds, Minutes, Hours, Days, Weeks, Months, or Years.

     For example, if you configure a Precision setting of 1 day, the cluster updates the last-accessed time once each day, even if some files were accessed more often than once during the day.
   - To disable access-time tracking, click **Disabled**.
3. Click **Submit**.

Specify the cluster character encoding

You can modify the character encoding set for the EMC Isilon cluster after installation.

Only OneFS-supported character sets are available for selection. UTF-8 is the default character set for OneFS nodes.

Note

If the cluster character encoding is not set to UTF-8, SMB share names are case-sensitive.

You must restart the cluster to apply character encoding changes.

⚠ **CAUTION**

Character encoding is typically established during installation of the cluster. Modifying the character encoding setting after installation may render files unreadable if done incorrectly. Modify settings only if necessary after consultation with Isilon Technical Support.

Procedure

1. Click **File System Management** › **File System Settings** › **Character Encoding**.
2. (Optional) From the **Character encoding** list, select the character-encoding set that you want to use.
3. Click **Submit**, and then click **Yes** to acknowledge that the encoding change becomes effective after the cluster is restarted.
4. Restart the cluster.

Results

After the cluster restarts, the web administration interface reflects your change.

Cluster monitoring

You can monitor the health, performance, and status of your EMC Isilon cluster.

Using the OneFS dashboard from the web administration interface, you can monitor the status and health of the OneFS system. Information is available for individual nodes, including node-specific network traffic, internal and external network interfaces, and details about node pools, tiers, and overall cluster health. You can monitor the following areas of your EMC Isilon cluster health and performance:

Node status
: Health and performance statistics for each node in the cluster, including hard disk drive (HDD) and solid-state drive (SSD) usage.

Client connections
: Number of clients connected per node.

New events
: List of event notifications generated by system events, including the severity, unique instance ID, start time, alert message, and scope of the event.

Cluster size
: **Current** view: Used and available HDD and SSD space and space reserved for the virtual hot spare (VHS). **Historical** view: Total used space and cluster size for a one-year period.

Cluster throughput (file system)
: **Current** view: Average inbound and outbound traffic volume passing through the nodes in the cluster for the past hour. **Historical** view: Average inbound and outbound traffic volume passing through the nodes in the cluster for the past two weeks.

CPU usage

Current view: Average system, user, and total percentages of CPU usage for the past hour. **Historical** view: CPU usage for the past two weeks.

Monitor the cluster

You can monitor the health and performance of an EMC Isilon cluster with charts and tables that show the status and performance of nodes, client connections, events, cluster size, cluster throughput, and CPU usage.

Procedure

1. Click **Dashboard** › **Cluster Overview** › **Cluster Status**.
2. (Optional) View cluster details.
   - Status: To view details about a node, click the ID number of the node.
   - Client connection summary: To view a list of current connections, click **Dashboard** › **Cluster Overview** › **Client Connections Status**.
   - New events: To view more information about an event, click **View details** in the **Actions** column.
   - Cluster size: To switch between current and historical views, click **Historical** or **Current** near the **Monitoring** section heading. In historical view, click **Used** or **Cluster size** to change the display.
   - Cluster throughput (file system): To switch between current and historical views, click **Historical** or **Current** next to the Monitoring section heading. To view throughput statistics for a specific period within the past two weeks, click **Dashboard** › **Cluster Overview** › **Throughput Distribution**.

     Note

     You can hide or show inbound or outbound throughput by clicking **Inbound** or **Outbound** in the chart legend. To view maximum throughput, next to **Show**, select **Maximum**.

   - CPU usage: To switch between current and historical views, click **Historical** or **Current** near the **Monitoring** section heading.

     Note

     You can hide or show a plot by clicking **System**, **User**, or **Total** in the chart legend. To view maximum usage, next to **Show**, select **Maximum**.

View node status

You can view the current and historical status of a node.

Procedure

1. Click **Dashboard** › **Cluster Overview** › **Cluster Status**.
2. (Optional) In the **Status** area, click the ID number for the node that you want to view status for.
3. View node details.
   - Status: To view networks settings for a node interface or subnet or pool, click the link in the **Status** area.

- Client connections: To view current clients connected to this node, review the list in this area.
- Chassis and drive status: To view the state of drives in this node, review this area. To view details about a drive, click the name link of the drive; for example, **Bay1**.
- Node size: To switch between current and historical views, click **Historical** or **Current** next to the **Monitoring** area heading. In historical view, click **Used** or **Cluster size** to change the display accordingly.
- Node throughput (file system): To switch between current and historical views, click **Historical** or **Current** next to the **Monitoring** area heading. To view throughput statistics for a period within the past two weeks, click **Dashboard** › **Cluster Overview** › **Throughput Distribution**.

  Note

  You can hide or show inbound or outbound throughput by clicking **Inbound** or **Outbound** in the chart legend. To view maximum throughput, next to **Show**, select **Maximum**.

- CPU usage: To switch between current and historical views, click **Historical** or **Current** next to the **Monitoring** area heading.

  Note

  You can hide or show a plot by clicking **System**, **User**, or **Total** in the chart legend. To view maximum usage, next to **Show**, select **Maximum**.

Monitoring cluster hardware

You can manually check the status of hardware on the EMC Isilon cluster as well as enable SNMP to remotely monitor components.

View node hardware status

You can view the hardware status of a node.

Procedure

1. Click **Dashboard** › **Cluster Overview** › **Cluster Status**.
2. (Optional) In the **Status** area, click the ID number for a node.
3. In the **Chassis and drive status** area, click **Platform**.

Chassis and drive states

You can view chassis and drive state details.

In a cluster, the combination of nodes in different degraded states determines whether read requests, write requests, or both work. A cluster can lose write quorum but keep read quorum. OneFS provides details about the status of chassis and drives in your cluster. The following table describes all the possible states that you may encounter in your cluster.

State	Description	Interface	Error state
`HEALTHY`	All drives in the node are functioning correctly.	Command-line interface, web administration interface	
`SMARTFAIL` or `Smartfail or restripe in progress`	The drive is in the process of being removed safely from the file system, either because of an I/O error or by user request. Nodes or drives in a smartfail or read-only state affect only write quorum.	Command-line interface, web administration interface	
`NOT AVAILABLE`	A drive is unavailable for a variety of reasons. You can click the bay to view detailed information about this condition. **Note** In the web administration interface, this state includes the `ERASE` and `SED_ERROR` command-line interface states.	Command-line interface, web administration interface	X
`SUSPENDED`	This state indicates that drive activity is temporarily suspended and the drive is not in use. The state is manually initiated and does not occur during normal cluster activity.	Command-line interface, web administration interface	
`NOT IN USE`	A node in an offline state affects both read and write quorum.	Command-line interface, web administration interface	
`REPLACE`	The drive was smartfailed successfully and is ready to be replaced.	Command-line interface only	
`STALLED`	The drive is stalled and undergoing stall evaluation. Stall evaluation is the process of checking drives that are slow or having other issues. Depending on the outcome of the evaluation, the drive may return to service or be smartfailed. This is a transient state.	Command-line interface only	
`NEW`	The drive is new and blank. This is the state that a drive is in when you run the `isi dev` command with the `-a add` option.	Command-line interface only	
`USED`	The drive was added and contained an Isilon GUID but the drive is not from this node. This drive likely will be formatted into the cluster.	Command-line interface only	

State	Description	Interface	Error state
PREPARING	The drive is undergoing a format operation. The drive state changes to HEALTHY when the format is successful.	Command-line interface only	
EMPTY	No drive is in this bay.	Command-line interface only	
WRONG_TYPE	The drive type is wrong for this node. For example, a non-SED drive in a SED node, SAS instead of the expected SATA drive type.	Command-line interface only	
BOOT_DRIVE	Unique to the A100 drive, which has boot drives in its bays.	Command-line interface only	
SED_ERROR	The drive cannot be acknowledged by the OneFS system. **Note** In the web administration interface, this state is included in Not available.	Command-line interface, web administration interface	X
ERASE	The drive is ready for removal but needs your attention because the data has not been erased. You can erase the drive manually to guarantee that data is removed. **Note** In the web administration interface, this state is included in Not available.	Command-line interface only	
INSECURE	Data on the self-encrypted drive is accessible by unauthorized personnel. Self-encrypting drives should never be used for non-encrypted data purposes. **Note** In the web administration interface, this state is labeled Unencrypted SED.	Command-line interface only	X
UNENCRYPTED SED	Data on the self-encrypted drive is accessible by unauthorized personnel. Self-encrypting drives should never be used for non-encrypted data purposes.	Web administration interface only	X

State	Description	Interface	Error state
	Note In the command-line interface, this state is labeled `INSECURE`.		

Check battery status

You can monitor the status of NVRAM batteries and charging systems.

This functionality is available only from the command line on node hardware that supports the command.

Procedure

1. Run the `isi batterystatus` command to view the status of all NVRAM batteries and charging systems on the node.

   The system displays output similar to the following example:

```
battery 1 : Good
battery 2 : Good
```

SNMP monitoring

You can use SNMP to remotely monitor the EMC Isilon cluster hardware components, such as fans, hardware sensors, power supplies, and disks. The default Linux SNMP tools or a GUI-based SNMP tool of your choice can be used for this purpose.

You can enable SNMP monitoring on individual nodes on your cluster, and you can also monitor cluster information from any node. Generated SNMP traps are sent to your SNMP network. You can configure an event notification rule that specifies the network station where you want to send SNMP traps for specific events, so that when an event occurs, the cluster sends the trap to that server. OneFS supports SNMP in read-only mode. OneFS supports SNMP version 2c, which is the default value, and SNMP version 3.

Note

OneFS does not support SNMP v1. Although an option for v1/v2c may be displayed, if you select the v1/v2c pair, OneFS will only monitor through SNMP v2c.

You can configure settings for SNMP v3 alone or for both SNMP v2c and v3.

Note

If you configure SNMP v3, OneFS requires the SNMP-specific security level of AuthNoPriv as the default value when querying the cluster. The security level AuthPriv is not supported.

Elements in an SNMP hierarchy are arranged in a tree structure, similar to a directory tree. As with directories, identifiers move from general to specific as the string progresses from left to right. Unlike a file hierarchy, however, each element is not only named, but also numbered.

For example, the SNMP entity `.iso.org.dod.internet.private.enterprises.isilon.oneFSss.s` sLocalNodeId.0 maps to .1.3.6.1.4.1.12124.3.2.0. The part of the name that refers to the OneFS SNMP namespace is the 12124 element. Anything further to the right of that number is related to OneFS-specific monitoring.

Management Information Base (MIB) documents define human-readable names for managed objects and specify their data types and other properties. You can download MIBs that are created for SNMP-monitoring of an Isilon cluster from the web-administration interface or manage them using the command-line interface. MIBs are stored in /usr/local/share/snmp/mibs/ on a OneFS node. The OneFS ISILON-MIBs serve two purposes:

- Augment the information available in standard MIBs
- Provide OneFS-specific information that is unavailable in standard MIBs

ISILON-MIB is a registered enterprise MIB. Isilon clusters have two separate MIBs:

ISILON-MIB
: Defines a group of SNMP agents that respond to queries from a network monitoring system (NMS) called OneFS Statistics Snapshot agents. As the name implies, these agents snapshot the state of the OneFS file system at the time that it receives a request and reports this information back to the NMS.

ISILON-TRAP-MIB
: Generates SNMP traps to send to an SNMP monitoring station when the circumstances occur that are defined in the trap protocol data units (PDUs).

The OneFS MIB files map the OneFS-specific object IDs with descriptions. Download or copy MIB files to a directory where your SNMP tool can find them, such as /usr/share/snmp/mibs/ or /usr/local/share/snmp/mibs, depending on the tool that you use.

To enable Net-SNMP tools to read the MIBs to provide automatic name-to-OID mapping, add **-m All** to the command, as in the following example:

```
snmpwalk -v2c -c public -m All <node IP> isilon
```

If the MIB files are not in the default Net-SNMP MIB directory, you may need to specify the full path, as in the following example. Note that all three lines are a single command.

```
snmpwalk -m /usr/local/share/snmp/mibs/ISILON-MIB.txt:/usr/local\
/share/snmp/mibs/ISILON-TRAP-MIB.txt:/usr/local/share/snmp/mibs \
/ONEFS-TRAP-MIB.txt -v2c -C c -c public <node IP> enterprises.onefs
```

Note

The previous examples are run from the snmpwalk command on a cluster. Your SNMP version may require different arguments.

Managing SNMP settings

SNMP can be used to monitor cluster hardware and system information. Settings can be configured through either the web administration interface or the command-line interface.

You can enable SNMP monitoring on individual nodes in the cluster, and you can monitor information cluster-wide from any node when you enable SNMP on each node. When using SNMP on an Isilon cluster, you should use a fixed general username. A password for the general user can be configured in the web administration interface.

You should configure a network monitoring system (NMS) to query each node directly through a static IP address. This approach allows you to confirm that all nodes have external IP addresses and therefore respond to SNMP queries. Because the SNMP proxy is enabled by default, the SNMP implementation on each node is configured automatically to proxy for all other nodes in the cluster except itself. This proxy configuration allows the Isilon Management Information Base (MIB) and standard MIBs to be exposed seamlessly through the use of context strings for supported SNMP versions. After you download and save the appropriate MIBs, you can configure SNMP monitoring through either the web administration interface or though the command-line interface.

Configure the cluster for SNMP monitoring

You can configure your EMC Isilon cluster to remotely monitor hardware components using SNMP.

Before you begin

When SNMP v3 is used, OneFS requires the SNMP-specific security level of AuthNoPriv as the default value when querying the cluster. The security level AuthPriv is not supported.

You can enable or disable SNMP monitoring, allow SNMP access by version, and configure other settings, some of which are optional. All SNMP access is read-only.

Note

The Isilon cluster does not generate SNMP traps unless you configure an event notification rule to send events.

Procedure

1. Click **Cluster Management** › **General Settings** › **SNMP Monitoring**.
2. In the Service area of the SNMP Monitoring page, enable or disable SNMP monitoring.
   a. To disable SNMP monitoring, click **Disable**, and then click **Submit**.
   b. To enable SNMP monitoring, click **Enable**, and then continue with the following steps to configure your settings.
3. In the Downloads area, click **Download** for the MIB file that you want to download.

   Follow the download process that is specific to your browser.
4. (Optional) If you are using Internet Explorer as your browser, right-click the **Download** link, select **Save As** from the menu, and save the file to your local drive.

   You can save the text in the file format that is specific to your Net-SNMP tool.
5. Copy MIB files to a directory where your SNMP tool can find them, such as `/usr/share/snmp/mibs/` or `/usr/local/share/snmp/mibs`, depending on the SNMP tool that you use.

   To have Net-SNMP tools read the MIBs to provide automatic name-to-OID mapping, add `-m All` to the command, as in the following example: `snmpwalk -v2c -c public -m All <node IP> isilon`
6. Navigate back to the SNMP Monitoring page and configure General Settings.
   a. In the Settings area, configure protocol access by selecting the version that you want.

      OneFS does not support writable OIDs; therefore, no write-only community string setting is available.
   b. In the **System location** field, type the system name.

      This setting is the value that the node reports when responding to queries. Type a name that helps to identify the location of the node.

   c. Type the contact email address in the **System contact** field.

7. (Optional) If you selected SNMP v1/v2 as your protocol, locate the SNMP v1/v2c Settings section and type the community name in the **Read-only community** field.

   The default community name is `I$ilonpublic`.

   Note

   OneFS no longer supports SNMP v1. Although an option for v1/v2c may be displayed, if you select the v1/v2c pair, OneFS will only monitor through SNMP v2c.

8. Configure SNMP v3 Settings.

   a. In the **Read-only user** field, type the SNMP v3 security name to change the name of the user with read-only privileges.

      The default read-only user is `general`.
      The password must contain at least eight characters and no spaces.

   b. In the **SNMP v3 password** field, type the new password for the read-only user to set a new SNMP v3 authentication password.

      The default password is `password`. We recommend that you change the password to improve security.

   c. Type the new password in the **Confirm password** field to confirm the new password.

9. Click **Submit**.

View SNMP settings

You can review SNMP monitoring settings.

Procedure

1. Click **Cluster Management** › **General Settings** › **SNMP Monitoring**.

Events and notifications

You can monitor the health and performance of your EMC Isilon cluster through OneFS event notifications.

When OneFS identifies an occurrence on your cluster that may require additional attention, an event is generated. OneFS records events related to file system integrity, network connections, hardware, and other vital components of your cluster.

You can select the events that you want to monitor, and you can cancel, quiet, or unquiet events.

In addition, you can configure event notification rules to determine who receives a notification when an event occurs.

Coalesced events

OneFS coalesces related, group events or repeated, duplicate events into a single event.

Coalesced group events

Group events are different types of events that are all related to a single occurrence.

In the following example, a single connection issue might generate the following events:

Event	Description
100010005	A SAS PHY topology problem or change was detected.
100010006	A drive's error log counter indicates there may be a problem.
100010007	A SAS link has exceeded the maximum Bit Error Rate (BER) .
100010008	A SAS link has been disabled for exceeding the maximum Bit Error Rate (BER).

Because the events are triggered by a single occurrence, OneFS creates a group event and combines the related messages under the new group event numbered 24.294. Instead of seeing four events, you will see a single group event alerting you to storage transport issues. You can still view all the grouped events individually if you choose.

To view this coalesced event, run the following command:

```
isi events show 24.924
```

The system displays the following example output of the coalesced group event:

```
          ID: 24.924
        Type: 199990001
    Severity: critical
       Value: 0.0
     Message: Disk Errors detected (Bay 1)
        Node: 21
    Lifetime: Sun Jun 17 23:29:29 2012 - Now
     Quieted: Not quieted
  Specifiers: disk: 35
              val: 0.0
              devid: 24
              drive_serial: 'XXXXXXXXXXXX'
              lba: 1953520064L
              lnn: 21
              drive_type: 'HDD'
              device: 'da1'
              bay: 1
              unit: 805306368
Coalesced by: --
Coalescer Type: Group
Coalesced events:
ID   STARTED   ENDED SEV LNN MESSAGE
24.911 06/17 23:29 --  I   21  Disk stall: Bay 1, Type HDD, LNUM 35.
Disk ...
24.912 06/17 23:29 --  I   21  Sector error: da1 block 1953520064
24.913 06/17 23:29 --  I   21  Sector error: da1 block 2202232
24.914 06/17 23:29 --  I   21  Sector error: da1 block 2202120
24.915 06/17 23:29 --  I   21  Sector error: da1 block 2202104
24.916 06/17 23:29 --  I   21  Sector error: da1 block 2202616
24.917 06/17 23:29 --  I   21  Sector error: da1 block 2202168
24.918 06/17 23:29 --  I   21  Sector error: da1 block 2202106
24.919 06/17 23:29 --  I   21  Sector error: da1 block 2202105
24.920 06/17 23:29 --  I   21  Sector error: da1 block 1048670
```

```
24.921 06/17 23:29 --  I   21  Sector error: da1 block 223
24.922 06/17 23:29 --  C   21  Disk Repair Initiated: Bay 1, Type
HDD, LNUM...
```

Coalesced duplicate events

Duplicate events are the same message repeated in response to an ongoing issue.

In the following example, a SmartQuotas maximum threshold is repeatedly exceeded, and the system coalesces this sequence of identical but discrete occurrences into one event numbered 1.3035.

To view this coalesced event, run the following command:

```
isi events show 1.3035
```

The system displays the following example output of the coalesced duplicate event:

```
          ID: 1.3035
        Type: 500010001
    Severity: info
       Value: 0.0
     Message: SmartQuotas threshold violation on quota violated,
domain direc...
        Node: All
    Lifetime: Thu Jun 14 01:00:00 2012 - Now
     Quieted: Not quieted
  Specifiers: enforcement: 'advisory'
              domain: 'directory /ifs/quotas'
              name: 'violated'
              val: 0.0
              devid: 0
              lnn: 0
Coalesced by: --
Coalescer Type: Duplicate
Coalesced events:
ID  STARTED  ENDED SEV LNN MESSAGE
18.621 06/14 01:00 -- I  All SmartQuotas threshold violation on quota
vio...
18.630 06/15 01:00 -- I  All SmartQuotas threshold violation on quota
vio...
18.638 06/16 01:00 -- I  All SmartQuotas threshold violation on quota
vio...
18.647 06/17 01:00 -- I  All SmartQuotas threshold violation on quota
vio...
18.655 06/18 01:00 -- I  All SmartQuotas threshold violation on quota
vio...
```

Viewing event information

You can view event details, event history, and event logs.

View event details

You can view the details of an event.

Procedure

1. Click **Dashboard** › **Events** › **Summary**.
2. In the **Actions** column of an event whose details you want to view, click **View details**.

View the event history

You can view all events in a chronological list and then select an event to view additional information.

Procedure

1. Click **Dashboard** › **Events** › **Events History**.

   The Events History page displays a list of all events in chronological order—newest to oldest—that have occurred on your Isilon cluster.

View the event log

You can log in to a node through the command-line interface and view the contents of the local event log.

Event logs are typically used for support purposes. You can only view the event log using the command-line interface.

Procedure

1. Establish an SSH connection to any node in the EMC Isilon cluster.
2. View the `/var/log/isi_celog_events.log` file.

   The log file lists all event activity. Each event row contains one of the following event labels:

Event label	Description
COALESCED: FIRST EVENT	An event was tagged as a possible first event in a series of events that can be coalesced. The first event label is a only a placeholder for a potential parent coalescer event.
COALESCER EVENT: ADDED	A parent coalescer event was created.
COALESCED	An event was added as a child beneath a coalescer event.
CREATOR EV COALID UPDATED	A group was created and the placeholder first event label was updated to include actual group information.
DROPPED	An event did not include any new information and was not stored in the master event database.
FORWARDED_TO_MASTER	An event was forwarded to the master node to be stored in the master event database.
DB: STORED	An event was stored in the master event database.
DB: PURGED	An event was removed from the master event database. The database has a limit of 50,000 entries, and old events are purged when that limit is reached.
INVALID EVENT: DROPPED	An event contained invalid information and was not stored in the master event database.

Event label	Description
UPDATE EVENT: DROPPED	A request to update the group information in a parent coalescer event was discontinued.

Responding to events

You can view event details and respond to cluster events.

You can view and manage new events, open events, and recently ended events. You can also view coalesced events and additional, more-detailed information about a specific event. You also can quiet or cancel events.

Quieting, unquieting, and canceling events

You can change an event's state by quieting, unquieting, or canceling an event.

You can select the following actions to change the state of an event:

Quiet

Acknowledges and removes the event from the list of new events and adds the event to a list of quieted events.

Note

If a new event of the same event type is triggered, it is a separate new event and must be quieted.

Unquiet

Returns a quieted event to an unacknowledged state in the list of new events and removes the event from the list of quieted events.

Cancel

Permanently ends an occurrence of an event. The system cancels an event when conditions are met that end its duration, which is bounded by a start time and an end time, or when you cancel the event manually.

Most events are canceled automatically by the system when the event reaches the end of its duration. The event remains in the system until you manually acknowledge or quiet the event. You can acknowledge events through either the web administration interface or the command-line interface.

Change the status for an event

You can change the status of an event by quieting, unquieting, or canceling it.

Procedure

1. Click **Dashboard** › **Events** › **Summary**.
2. Perform one of the following actions:
   - To acknowledge an event, click **Quiet**.
   - To restore an event to an unacknowledged state, click **Unquiet**.
   - To permanently remove an occurrence of an event, click **Cancel**.

Managing event notification settings

You can view and modify event notification settings and configure batch notifications.

Event notification methods

You can define the method by which OneFS delivers notifications.

Email
: You can send email messages to distribution lists and apply email templates to notifications. You can also specify SMTP, authorization, and security settings.

SupportIQ
: You can deliver notifications to Isilon Technical Support over HTTPS, SMTP, or both.

SNMP trap
: You can send SNMP traps to one or more network monitoring stations or trap receivers. Each event can generate one or more SNMP traps. You can download management information base files (MIBs) from the cluster at `/usr/local/share/snmp/mibs/`. The `ISILON-TRAP-MIB.txt` file describes the traps that the cluster can generate, and the `ISILON-MIB.txt` file describes the associated varbinds that accompany the traps.

ESRS
: You can receive alerts from the EMC Secure Remote Support (ESRS) Gateway. The ESRS Gateway is a secure, IP-based customer service support system.

The ESRS Gateway is similar to SupportIQ and performs many of the same functions:

- Send alerts regarding the health of your devices.
- Enable support personnel to run the same scripts used by SupportIQ to gather data from your devices.
- Allow support personnel to establish remote access to troubleshoot your cluster.

Event notification settings

You can specify whether you want to receive event notifications as aggregated batches or as individual notifications for each event. Batch notifications are sent every 10 seconds.

The batch options that are described in this table affect both the content and the subject line of notification emails that are sent in response to system events. You can specify event notification batch options when you configure SMTP email settings.

Setting	Option	Description
Notification batch mode	Batch all	Generates a single email for each event notification.
	Batch by severity	Generates an email that contains aggregated notifications for each event of the same severity, regardless of event category.
	Batch by category	Generates an email that contains aggregated notifications for event of the same category, regardless of severity.
	No batching	Generates one email per event.

Setting	Option	Description
Custom notification template	No custom notification template is set	Sends the email notification in the default OneFS notification template format.
	Set custom notification template	Sends the email notifications in the format that you defined in your custom template file.

View event notification settings

You can view email, SupportIQ, and contact information for event notifications.

Procedure

1. Click **Dashboard** › **Events** › **Notification Settings**.

Modify event notification settings

You can modify email, SupportIQ, and contact settings for event notifications.

Procedure

1. Click **Dashboard** › **Events** › **Notification Settings.**
2. Click the **Modify** link for the setting that you want to change.
3. Click **Submit.**

Specify event-notification batch mode or template settings

You can choose an event-notification batch option to specify whether you want to receive notifications individually or as an aggregate. You also can specify a custom notification template for email notifications.

Before you begin

You must first create a custom notification template and then upload it to a directory at the same level or below `/ifs`; for example, `/ifs/templates`.

Procedure

1. Click **Cluster Management** › **General Settings** › **Email Settings**.
2. In the **Event Notification Settings** area on the General Settings page, select a **Notification batch mode** option.
3. Leave the **Set custom notification template** field blank to use the default notification template.
4. In the **Custom notification template** field, select the custom event notification template.
   - Click **Browse**, navigate to and select the template file that you want to use, and then click **OK.**
   - In the **Set custom notification template** field, type the path and file name of the template file that you want to use.
5. Click **Submit**.

Managing event notification rules

You can create, modify, or delete event notification rules to determine when and how you receive information about specific system events.

Create an event notification rule

You can configure event notification rules based on specified events and event types.

You can configure email notification and SNMP trap generation for a specific event.

Procedure

1. Click **Dashboard** › **Events** › **Event Notification Rules**.
2. In the Notification Rules area on the Cluster Events page, click **Add Rule**.
3. In the **Rule name** field on the Add Notification Rule page, type a name for the rule.
4. In the Recipients area, specify a notification method.
   a. To notify a recipient through email, select **Email**, type the email address to which notifications will be sent, and then click **Add**.
   b. To notify a community through SNMP traps, select **SNMP**, select the community name and the SNMP host, which is the network monitoring station, from the respective lists, and then click **Add**.
   c. To add additional notification recipients or communities, repeat these steps.
5. In the Events area, expand the event types and select the check boxes for the events and event types that you want to trigger this notification.
6. Click **Submit**.

Send a test event notification

You can generate a test event notification to confirm that event notifications are working as you intend.

Procedure

1. Click **Dashboard** › **Events** › **Notification Settings**.
2. In the **Send Test Event** area on the Cluster Events page, click **Send test event**.
3. On the Cluster Events page, click **Summary** to verify whether the test event was successful.

   A corresponding test event notification appears in the New Events list, which appears in the **Message** column as a message similar to `Test event sent from WebUI`.

View event notification rules

You can view a list of event notification rules and details about specific rules.

Procedure

1. Click **Dashboard** › **Events** › **Event Notification Rules**.
2. In the **Actions** column of the rule whose settings you want to view, click **Edit**.
3. When you have finished viewing the rule details, click **Cancel**.

Modify an event notification rule

You can modify event notification rules that you created. System event notification rules cannot be modified.

Procedure

1. Click **Dashboard** › **Events** › **Event Notification Rules**.
2. In the **Actions** column for the rule that you want to modify, click **Edit**.
3. Modify the event notification rule settings as needed.
4. Click **Submit**.

Delete an event notification rule

You can delete event notification rules that you created, but system event notification rules cannot be deleted.

Procedure

1. Click **Dashboard** › **Events** › **Event Notification Rules**.
2. In the **Notification Rules** area, in the **Actions** column for the rule that you want to delete, click **Delete**.
3. Click **Yes** to confirm the deletion.

Cluster maintenance

Trained service personnel can replace or upgrade components in Isilon nodes.

Isilon Technical Support can assist you with replacing node components or upgrading components to increase performance.

Replacing node components

If a node component fails, Isilon Technical Support will work with you to quickly replace the component and return the node to a healthy status.

Trained service personnel can replace the following field replaceable units (FRUs):

- battery
- boot flash drive
- SATA/SAS Drive
- memory (DIMM)
- fan
- front panel
- intrusion switch
- network interface card (NIC)
- InfiniBand card
- NVRAM card
- SAS controller
- power supply

If you configure your cluster to send alerts to Isilon, Isilon Technical Support will contact you if a component needs to be replaced. If you do not configure your cluster to send alerts to Isilon, you must initiate a service request.

Upgrading node components

You can upgrade node components to gain additional capacity or performance.

Trained service personnel can upgrade the following components in the field:

- drive
- memory (DIMM)
- network interface card (NIC)

If you want to upgrade components in your nodes, contact Isilon Technical Support.

Managing drive firmware

If the firmware of any drive in a cluster becomes obsolete, the cluster performance or hardware reliability might get affected. To ensure overall data integrity, you may update the drive firmware to the latest revision by installing the drive support package or the drive firmware package.

You can determine whether the drive firmware on your cluster is of the latest revision by viewing the status of the drive firmware.

Note

We recommend that you contact EMC Isilon Technical Support before updating the drive firmware.

Drive firmware update overview

You can update the drive firmware through drive support packages or drive firmware packages.

Download and install either of these packages from http://support.emc.com depending on the OneFS version running on your cluster and the type of drives on the nodes.

Drive Support Package
For clusters running OneFS 7.1.1 and later, install a drive support package to update the drive firmware. You do not need to reboot the affected nodes to complete the firmware update. A drive support package provides the following additional capabilities:

- Updates the following drive configuration information:
  - List of supported drives
  - Drive firmware metadata
  - SSD wear monitoring data
  - SAS and SATA settings and attributes
- Automatically updates the drive firmware for new and replacement drives to the latest revision before those drives are formatted and used in a cluster. This is applicable only for clusters running OneFS 7.2 and later.

  Note

  Firmware of drives in use cannot be updated automatically.

Drive Firmware Package
For clusters running OneFS versions earlier than 7.1.1, or for clusters with non-bootflash nodes, install a cluster-wide drive firmware package to update the drive firmware. You must reboot the affected nodes to complete the firmware update.

Install a drive support package

For clusters running OneFS 7.1.1 and later, install the drive support package to update your drive firmware to the latest supported revision.

Procedure

1. Go to the EMC Support page that lists all the available versions of the drive support package.
2. Click the latest version of the drive support package and download the file.
3. Open a secure shell (SSH) connection to any node in the cluster and log in.
4. Create or check for the availability of the directory structure `/ifs/data/Isilon_Support/dsp`.
5. Copy the downloaded file to the `dsp` directory through SCP, FTP, SMB, NFS, or any other supported data-access protocols.
6. Unpack the file by running the `tar` command.

   For example, unpack drive support package version 1.4 as follows:

   ```
   tar -zxvf Drive_Support_v1.4.tgz
   ```
7. Install the package by running the `isi_dsp_install` command.

   For example, install drive support package version 1.4 as follows:

   ```
   isi_dsp_install Drive_Support_v1.4.tar
   ```

   Note

   - You must run the **`isi_dsp_install`** command to install the drive support package. Do not use the **`isi pkg`** command.
   - The installation process takes care of installing all the necessary files from the drive support package followed by the uninstallation of the package. You do not need to delete the package after its installation or prior to installing a later version.

View drive firmware status

You can view the status of the drive firmware on the cluster to determine whether you need to update the firmware.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Perform one of the following tasks:
   - To view the drive firmware status of all the nodes, run the following command:

     ```
     isi drivefirmware status
     ```

- To view the drive firmware status of drives on a specific node, run the `isi devices` command with the `-a fwstatus` option. Run the following command to view the drive firmware status of each drive on node 1:

  ```
  isi devices -a fwstatus -d 1
  ```

  The output of the previous command is shown in the following example:

  ```
  Node 1
              Model                      FW             Desired FW
  Bay 1       HGST HUS724030ALA640       MF80AAC0
  Bay 2       HGST HUS724030ALA640       MF80AAC0
  Bay 3       HGST HUS724030ALA640       MF80AAC0
  ```

  Run the following command to view the drive firmware status on node 1 and disk 12:

  ```
  isi devices -a fwstatus -d 1:12
  ```

If a drive firmware update is not required, the `Desired FW` column is empty.

Update the drive firmware

Determine the type of drive in the node, and then execute the update. The node restarts automatically after the loading process is complete.

Before you begin

Install the drive firmware package.

Note

Do not restart or power off the node before the update is complete. When the update process completes successfully, the node restarts automatically.

Procedure

1. Log in to the node through either a serial console port or an internal SSH connection between nodes.

   Note

   External network interfaces are disabled as part of the reboot process. The command in the next step will fail if you run the command from an external network interface.

2. Determine whether a drive firmware update is required by typing the following command:

   ```
   isi_disk_firmware_reboot
   ```

3. Determine whether any of the drives on the node that need to be updated are Western Digital drives by typing the following command:

   ```
   isi_radish -q
   ```

   ⚠ CAUTION

   If you are performing a drive firmware update on Western Digital drives, you must update the drives sequentially to avoid significant drive damage.

4. Type one of the following commands depending on whether the node contains any Western Digital drives that require a drive firmware update:

- To update the drive firmware of a node with any Western Digital drives, type the following command to perform a sequential update:
  `isi_disk_firmware_reboot -sv`
- To update the drive firmware of a node without any Western Digital drives, type the following command to perform a parallel update:
  `isi_disk_firmware_reboot -p`

A drive firmware update takes 20–60 seconds, depending on the drive model. A node containing Western Digital drives takes approximately fifteen minutes to complete the entire process.

After the update process is complete, the node reboots automatically.
If the update is unsuccessful, the LED display on the front panel of the node will indicate an error and the node will not reboot. Wait for a few minutes and then run the `reboot` command to reboot the node manually. If this process is unsuccessful, contact EMC Isilon Technical Support.

Verify a drive firmware update

After you update the drive firmware in a node, confirm that the firmware is updated properly and that the affected drives are operating correctly.

Procedure

1. Ensure that no drive firmware updates are currently in progress by running the following command:

   ```
   isi devices
   ```

   If a drive is currently being updated, `[FW_UPDATE]` appears in the status column.

2. Verify that all drives have been updated by running the following command:

   ```
   isi drivefirmware status
   ```

   If all drives have been updated, the `Desired FW` column is empty.

3. Verify that all affected drives are operating in a healthy state by running the following command:

   ```
   isi devices
   ```

   If a drive is operating in a healthy state, `[HEALTHY]` appears in the status column.

Drive firmware status information

You can view information about the status of the drive firmware through the OneFS command-line interface.

The following example shows the output of the `isi drivefirmware status` command:

```
Model                    FW           Desired FW    Count     Nodes
HGST HUS724030ALA640     MF80AAC0            30        1
```

Where:

Model

Displays the name of the drive model.

FW

Displays the version number of the firmware currently running on the drives.

Desired FW

If the drive firmware should be upgraded, displays the version number of the drive firmware that the firmware should be updated to.

Count

Displays the number of drives of this model that are currently running the specified drive firmware.

Nodes

Displays the LNNs of nodes that the specified drives exist in.

The following example shows the output of the `isi devices` command with the `-a fwstatus` option:

```
Node 1
                 Model                     FW             Desired FW
Bay 1            HGST HUS724030ALA640      MF80AAC0
Bay 2            HGST HUS724030ALA640      MF80AAC0
Bay 3            HGST HUS724030ALA640      MF80AAC0
```

Where:

Drive

Displays the number of the bay that the drive is in.

Note

This column is not labeled in the output. The information appears under the node number.

Model

Displays the name of the drive model.

FW

Displays the version number of the firmware currently running on the drive.

Desired FW

Displays the version number of the drive firmware that the drive should be updated to. If a drive firmware update is not required, the `Desired FW` column is empty.

Automatic update of drive firmware

For clusters running OneFS 7.2 or later, install the latest drive support package on a node to automatically update the firmware for a new or replacement drive.

The information within the drive support package determines whether the firmware of a drive must be updated before the drive is formatted and used. If an update is available, the drive is automatically updated with the latest firmware.

Note

New and replacement drives added to a cluster are formatted regardless of the status of their firmware revision. You can identify a firmware update failure by viewing the firmware status for the drives on a specific node. In case of a failure, run the `isi devices` command with the `fwupdate` action on the node to update the firmware manually. For example, run the following command to manually update the firmware on node 1:

```
isi devices -a fwupdate -d 1
```

Managing cluster nodes

You can add and remove nodes from a cluster. You can also shut down or restart the entire cluster.

Add a node to a cluster

You can add a new node to an existing EMC Isilon cluster.

Before you begin

Before you add a node to a cluster, verify that an internal IP address is available. Add IP addresses as necessary before you add a new node.

If a new node is running a different version of OneFS than a cluster, the system changes the node version of OneFS to match the cluster.

Note

For specific information about version compatibility between OneFS and EMC Isilon hardware, refer to the *Isilon Supportability and Compatibility Guide*.

Procedure

1. Click **Cluster Management** › **Hardware Configuration** › **Add Nodes**.
2. In the **Available Nodes** table, click **Add** for the node that you want to add to the cluster.

Remove a node from the cluster

You can remove a node from an EMC Isilon cluster. When you remove a node, the system smartfails the node to ensure that data on the node is transferred to other nodes in the cluster.

Removing a storage node from a cluster deletes the data from that node. Before the system deletes the data, the FlexProtect job safely redistributes data across the nodes remaining in the cluster.

Procedure

1. Navigate to **Cluster Management** › **Hardware Configuration** › **Remove Nodes**.
2. In the Remove Node area, specify the node you want to remove.
3. Click **Submit**.

   If you remove a storage node, the Cluster Status area displays smartfail progress. If you remove a non-storage accelerator node, it is immediately removed from the cluster.

Modify the LNN of a node

You can modify the logical node number (LNN) of a node. This procedure is available only through the command-line interface (CLI).

The nodes within your cluster can be renamed to any name/integer between 1 and 144. By changing the name of your node, you are resetting the LNN.

Note

Although you can specify any integer as an LNN, we recommend that you do not specify an integer greater than 144. Specifying LNNs above 144 can result in significant performance degradation.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Open the isi config command prompt by running the following command:

   ```
   isi config
   ```

3. Run the `lnnset` command.

   The following command switches the LNN of a node from 12 to 73:

   ```
   lnnset 12 73
   ```

4. Enter `commit` .

Results

You might need to reconnect to your SSH session before the new node name is automatically changed.

Shut down or restart a cluster

You can shut down or restart an entire EMC Isilon cluster.

Procedure

1. Navigate to **Cluster Management** › **Hardware Configuration** › **Shutdown & Reboot Controls.**
2. In the Shut Down or Reboot This Cluster area, specify an action:

Option	Description
Shut down	Shuts down the cluster.
Reboot	Stops then restarts the cluster.

3. Click **Submit.**

Upgrading OneFS

Two options are available for upgrading the OneFS operating system: a rolling upgrade or a simultaneous upgrade. Before upgrading OneFS software, a pre-upgrade check must be performed.

A rolling upgrade individually upgrades and restarts each node in the EMC Isilon cluster sequentially. During a rolling upgrade, the cluster remains online and continues serving clients with no interruption in service, although some connection resets may occur on

SMB clients. Rolling upgrades are performed sequentially by node number, so a rolling upgrade takes longer to complete than a simultaneous upgrade. The final node in the upgrade process is the node that you used to start the upgrade process.

Note

Rolling upgrades are not available for all clusters. For instructions on how to plan an upgrade, prepare the cluster for upgrade, and perform an upgrade of the operating system, see the *OneFS Upgrade Planning and Process Guide.*

A simultaneous upgrade installs the new operating system and restarts all nodes in the cluster at the same time. Simultaneous upgrades are faster than rolling upgrades but require a temporary interruption of service during the upgrade process. Your data is inaccessible during the time that it takes to complete the upgrade process.

Before beginning either a simultaneous or rolling upgrade, OneFS compares the current cluster and operating system with the new version to ensure that the cluster meets certain criteria, such as configuration compatibility (SMB, LDAP, SmartPools), disk availability, and the absence of critical cluster events. If upgrading puts the cluster at risk, OneFS warns you, provides information about the risks, and prompts you to confirm whether to continue the upgrade.

If the cluster does not meet the pre-upgrade criteria, the upgrade does not proceed, and the unsupported statuses are listed.

Remote support

Isilon Technical Support personnel can remotely manage your Isilon cluster to troubleshoot an open support case with your permission.

You can enable remote customer service support through SupportIQ or the EMC Secure Remote Support (ESRS) Gateway.

Remote support using SupportIQ

Isilon Technical Support personnel can remotely manage your Isilon cluster to troubleshoot an open support case with your permission. The Isilon SupportIQ module allows Isilon Technical Support personnel to gather diagnostic data about the cluster.

Isilon Technical Support representatives run scripts that gather data about cluster settings and operations. The SupportIQ agent then uploads the information to a secure Isilon FTP site so it is available for Isilon Technical Support personnel to review. These scripts do not affect cluster services or data availability.

Note

The SupportIQ scripts are based on the Isilon `isi_gather_info` log-gathering tool.

The SupportIQ module is included with the OneFS operating system and does not require you to activate a separate license. You must enable and configure the SupportIQ module before SupportIQ can run scripts to gather data. The feature may have been enabled when the cluster was first set up, but you can enable or disable SupportIQ through the Isilon web administration interface.

In addition to enabling the SupportIQ module to allow the SupportIQ agent to run scripts, you can enable remote access, which allows Isilon Technical Support personnel to monitor cluster events and remotely manage your cluster using SSH or the web administration interface. Remote access helps Isilon Technical Support to quickly identify and troubleshoot cluster issues. Other diagnostic tools are available for you to use in conjunction with Isilon Technical Support to gather and upload information such as packet capture metrics.

Note

If you enable remote access, you must also share cluster login credentials with Isilon Technical Support personnel. Isilon Technical Support personnel remotely access your cluster only in the context of an open support case and only after receiving your permission.

Configuring SupportIQ

OneFS logs contain data that Isilon Technical Support personnel can securely upload, with your permission, and then analyze to troubleshoot cluster problems. The SupportIQ technology must be enabled and configured for this process.

When SupportIQ is enabled, Isilon Technical Support personnel can request logs through scripts that gather cluster data and then upload the data to a secure location. You must enable and configure the SupportIQ module before SupportIQ can run scripts to gather data. The feature may have been enabled when the cluster was first set up.

You can also enable remote access, which allows Isilon Technical Support personnel to troubleshoot your cluster remotely and run additional data-gathering scripts. Remote access is disabled by default. To enable remote SSH access to your cluster, you must provide the cluster password to a Technical Support engineer.

Enable and configure SupportIQ

You can enable and configure SupportIQ to allow the SupportIQ agent to run scripts that gather and upload information about your cluster to Isilon Technical Support personnel. Optionally, you can enable remote access to your EMC Isilon cluster.

Procedure

1. Click **Cluster Management** › **General Settings** › **SupportIQ**.
2. In the **SupportIQ Settings** area, select the **Enable SupportIQ** check box.
3. For **SupportIQ alerts**, select an option.
   - **Send alerts via SupportIQ agent (HTTPS) and by email (SMTP)** – SupportIQ delivers notifications to Isilon through the SupportIQ agent over HTTPS and by email over SMTP.
   - **Send alerts via SupportIQ agent (HTTPS)** – SupportIQ delivers notifications to Isilon only through the SupportIQ agent over HTTPS.
4. (Optional) Enable HTTPS proxy support for SupportIQ.
   a. Select the **HTTPS proxy for SupportIQ** check box.
   b. In the **Proxy host** field, type the IP address or fully qualified domain name (FQDN) of the HTTP proxy server.
   c. In the **Proxy port** field, type the number of the port on which the HTTP proxy server receives requests.
   d. (Optional) In the **Username** field, type the user name for the proxy server.
   e. (Optional) In the **Password** field, type the password for the proxy server.
5. (Optional) Enable remote access to the cluster.
   a. Select the **Enable remote access to cluster via SSH and web interface** check box.

b. Review the remote-access end user license agreement (EULA) and, if you agree to the terms and conditions, select the **I have read and agree to** check box.

6. Click **Submit**.

   A successful configuration is indicated by a message similar to `SupportIQ settings have been updated`.

Disable SupportIQ

You can disable SupportIQ so the SupportIQ agent does not run scripts to gather and upload data about your EMC Isilon cluster.

Procedure

1. Click **Cluster Management** › **General Settings** › **SupportIQ**.
2. Clear the **Enable SupportIQ** check box.
3. Click **Submit**.

SupportIQ scripts

When SupportIQ is enabled, Isilon Technical Support personnel can request logs with scripts that gather cluster data and then upload the data. The SupportIQ scripts are located in the `/usr/local/SupportIQ/Scripts/` directory on each node.

The following table lists the data-gathering activities that SupportIQ scripts perform. These scripts can be run automatically, at the request of an Isilon Technical Support representative, to collect information about your cluster's configuration settings and operations. The SupportIQ agent then uploads the information to a secure Isilon FTP site, so that it is available for Isilon Technical Support personnel to analyze. The SupportIQ scripts do not affect cluster services or the availability of your data.

Action	Description
Clean watch folder	Clears the contents of `/var/crash`.
Get application data	Collects and uploads information about OneFS application programs.
Generate dashboard file daily	Generates daily dashboard information.
Generate dashboard file sequence	Generates dashboard information in the sequence that it occurred.
Get ABR data (as built record)	Collects as-built information about hardware.
Get ATA control and GMirror status	Collects system output and invokes a script when it receives an event that corresponds to a predetermined `eventid`.
Get cluster data	Collects and uploads information about overall cluster configuration and operations.
Get cluster events	Gets the output of existing critical events and uploads the information.
Get cluster status	Collects and uploads cluster status details.
Get contact info	Extracts contact information and uploads a text file that contains it.
Get contents (var/crash)	Uploads the contents of `/var/crash`.

Action	Description
Get job status	Collects and uploads details on a job that is being monitored.
Get domain data	Collects and uploads information about the cluster's Active Directory Services (ADS) domain membership.
Get file system data	Collects and uploads information about the state and health of the OneFS `/ifs/` file system.
Get IB data	Collects and uploads information about the configuration and operation of the InfiniBand back-end network.
Get logs data	Collects and uploads only the most recent cluster log information.
Get messages	Collects and uploads active `/var/log/messages` files.
Get network data	Collects and uploads information about cluster-wide and node-specific network configuration settings and operations.
Get NFS clients	Runs a command to check if nodes are being used as NFS clients.
Get node data	Collects and uploads node-specific configuration, status, and operational information.
Get protocol data	Collects and uploads network status information and configuration settings for the NFS, SMB, FTP, and HTTP protocols.
Get Pcap client stats	Collects and uploads client statistics.
Get readonly status	Warns if the chassis is open and uploads a text file of the event information.
Get usage data	Collects and uploads current and historical information about node performance and resource usage.
`isi_gather_info`	Collects and uploads all recent cluster log information.
`isi_gather_info --incremental`	Collects and uploads changes to cluster log information that have occurred since the most recent full operation.
`isi_gather_info --incremental single node`	Collects and uploads details for a single node. Prompts you for the node number.
`isi_gather_info single node`	Collects and uploads changes to cluster log information that have occurred since the most recent full operation. Prompts you for the node number.
Upload the dashboard file	Uploads dashboard information to the secure Isilon Technical Support FTP site.

Remote support using ESRS Gateway

EMC Isilon clusters support enablement of the ESRS Gateway.

The EMC Secure Remote Support (ESRS) Gateway is a secure, IP-based customer service support system. The EMC ESRS Gateway features include 24x7 remote monitoring and secure authentication with AES 256-bit encryption and RSA digital certificates. You can select monitoring on a node-by node basis, allow or deny remote support sessions, and review remote customer service activities.

The ESRS Gateway is similar to SupportIQ and performs many of the same functions:

- Send alerts regarding the health of your devices.
- Enable support personnel to run the same scripts used by SupportIQ to gather data from your devices.
- Allow support personnel to establish remote access to troubleshoot your cluster.

An important difference between SupportIQ and the ESRS Gateway is that SupportIQ management is cluster-wide; SupportIQ manages all nodes. The ESRS Gateway manages nodes individually; you select which nodes should be managed.

You can only enable one remote support system on your Isilon cluster. The EMC products you use and your type of environment determine which system is most appropriate for your Isilon cluster:

- If your environment comprises one or more EMC products that can be monitored, use the ESRS Gateway.
- If ESRS is currently implemented in your environment, use the ESRS Gateway.
- If your use of ESRS requires the ESRS Client, use SupportIQ. Isilon nodes do not support ESRS Client connectivity.
- If you have a high-security environment, use the ESRS Gateway.
- If the only EMC products in your environment are Isilon nodes, use SupportIQ.

See the most recent version of the document titled *EMC Secure Remote Support Technical Description* for a complete description of EMC Secure Remote Support features and functionality.

Additional documentation on ESRS can be found on the EMC Online Support site.

Configuring ESRS Gateway support

You can configure support for the ESRS Gateway on your Isilon cluster.

Before configuring ESRS Gateway support on your Isilon cluster, at least one ESRS Gateway server must be installed and configured. The server acts as the single point of entry and exit for IP-based remote support activities and monitoring notifications. You can also set up a secondary gateway server as a failover, specify whether to use SMTP if ESRS transmission fails, and specify whether an email should be sent upon transmission failure.

ESRS Gateway support also requires you to designate a subnet as a point for remote access by support personnel. We recommend that you designate a subnet that is dedicated to remote connections through the ESRS Gateway, and that the subnet contains a static IP address pool in the System access zone. If you cannot dedicate a subnet for remote connections, ensure that the first IP address pool in the designated subnet is configured to use static IP addresses and is assigned to the System access zone.

When you enable support for the ESRS Gateway on a cluster, the serial number and IP address of each node is sent to the ESRS Gateway server. Once node information is received, you can:

- Select which nodes you want managed through the ESRS Gateway with the ESRS Configuration Tool.
- Create rules for remote support connection to Isilon nodes with the ESRS Policy Manager.

See the most recent version of the document titled *EMC Secure Remote Site Planning Guide* for a complete description of ESRS Gateway server requirements, installation, and configuration.

See the most recent version of the document titled *EMC Secure Remote Support Gateway for Windows Operations Guide* for a complete description of the ESRS Configuration Tool.

See the most recent version of the document titled *EMC Secure Remote Support Policy Manager Operations Guide* for a complete description of the ESRS Policy Manger.

Additional documentation on ESRS can be found on the EMC Online Support site.

Enable and configure ESRS Gateway support

You can enable support for the ESRS Gateway on an Isilon cluster.

Before you begin

An ESRS Gateway server must be installed and configured before you can enable ESRS Gateway support on an Isilon cluster. SupportIQ must be disabled.

Procedure

1. Run the `isi remotesupport connectemc modify` command to enable and configure ESRS Gateway support.

   The following command enables ESRS Gateway support, specifies a primary gateway, a remote support subnet, and that an SMTP failover should be used.

   ```
   isi remotesupport connectemc modify --enabled yes \
   -- primary-esrs-gateway gw-serv-esrs1 --use-smtp-failover yes \
   --remote-support-subnet subnet0
   ```

Disable ESRS Gateway support

You can disable ESRS Gateway support on the Isilon cluster.

You can disable support for the ESRS Gateway in order to use SupportIQ. Isilon clusters only allow one remote support system to be enabled at a time.

Procedure

1. Disable ESRS Gateway support on an Isilon cluster by running the following command:

   ```
   isi remotesupport connectemc modify --enabled no
   ```

View ESRS Gateway settings

You can view ESRS Gateway settings specified on an EMC Isilon cluster.

Procedure

1. Run the `isi remotesupport connectemc view` command.

   The system displays output similar to the following example:

   ```
                     Enabled: yes
        Primary Esrs Gateway: gw-serv-esrs1
      Secondary Esrs Gateway: gw-serv-esrs2
           Use Smtp Failover: yes
   Email Customer On Failure: no
       Remote Support Subnet: subnet0
   ```

CHAPTER 4

Access zones

This section contains the following topics:

Access zones overview

Although the default view of an EMC Isilon cluster is that of one physical machine, you can partition a cluster into multiple virtual containers called access zones. Access zones allow you to isolate data and control who can access data in each zone.

Access zones support all configuration settings for authentication and identity management services on a cluster, so you can configure authentication providers and provision SMB shares and NFS exports on a zone-by-zone basis. When you create an access zone, a local provider is created automatically, which allows you to configure each access zone with a list of local users and groups. You can also authenticate through a different authentication provider in each access zone.

To control data access, you can direct incoming connections to the access zone through a specific IP address pool. Associating an access zone with an IP address pool restricts authentication to the associated access zone and reduces the number of available and accessible SMB shares and NFS exports. Another advantage to multiple access zones is the ability to configure audit protocol access for individual access zones. You can modify the default list of successful and failed protocol audit events and then generate reports through a third-party tool for an individual access zone.

A cluster includes a built-in access zone named System, where you manage all aspects of a cluster and other access zones. By default, all cluster IP addresses connect to the System zone. Even if you create additional access zones, you configure all access zones in the System zone. Role-based access, which primarily allows configuration actions, is available through only the System zone. All administrators, including those given privileges by a role, must connect to the System zone to configure a cluster.

Configuration management of a non-System access zone is not permitted through SSH, the OneFS Platform API, or the web administration interface. However, you can create and delete SMB shares in an access zone through the Microsoft Management Console (MMC).

Access zone base directory rules

You must assign a base directory to each access zone. A base directory defines the file system tree exposed by an access zone and isolates data contained in the directory to the access zone.

A base directory path is unique for each access zone and cannot overlap or be nested inside base directories of other access zones.

Base directories restrict configuration options for several features such as SMB share paths, NFS exports, the HDFS root directory, and the local provider home directory template. You must observe the following rules when specifying a base directory:

- The base directory cannot be identical to the base directory of any other access zone, except the System zone. For example, you cannot specify `/ifs/data/hr` to both the zone2 and zone3 access zones.
- Cannot overlap with the file system tree of a base directory in any other access zone, except the System zone. For example, if `/ifs/data/hr` is assigned to zone2, you cannot assign `/ifs/data/hr/personnel` to zone3.
- The base directory of the default System access zone is `/ifs` and cannot be modified.

Note

Assigning a base directory that is identical to or overlaps with the System zone is allowed, but only recommended as a temporary base directory when modifying the base directory path and migrating data to the new directory.

Access zones best practices

You can avoid configuration problems on the EMC Isilon cluster when creating access zones by following best practices guidelines.

Best practice	Details
Create unique base directories.	The base directory path of each access zone must be unique and cannot overlap or be nested inside the base directory of another access zone.
Separate the function of the System zone from other access zones.	If you choose to create any additional access zones, do not allow data access in both the System zone and created zones. Reserve the System zone for configuration access, and create additional zones for data access. Move current data out of the System zone and into a new access zone.
Create access zones to isolate data access for different clients or users.	Do not create access zones if a workflow requires data sharing between different classes of clients or users.
Assign only one authentication provider of each type to each access zone.	An access zone is limited to a single Active Directory provider; however, OneFS allows multiple LDAP, NIS, and file authentication providers in each access zone. It is recommended that you assign only one type of each provider per access zone in order to simplify administration.
Avoid overlapping UID or GID ranges for authentication providers in the same access zone.	The potential for zone access conflicts is slight but possible if overlapping UIDs/GIDs are present in the same access zone.
Configure a single DNS server for all access zones.	OneFS does not support one DNS server per access zone. It is recommended that all access zones point to a single DNS server.

Access zone limits

You can follow access zone limits guidelines to help size the workloads on the OneFS system.

If you configure multiple access zones on an EMC Isilon cluster, limits guidelines are recommended for optimal system performance. The limits described in the *EMC Isilon Guidelines for Large Workloads* publication are recommended for heavy enterprise workflows on a cluster, treating each access zone as a separate physical machine.

Quality of service

You can set upper bounds on quality of service by assigning specific physical resources to each access zone.

Quality of service addresses physical hardware performance characteristics that can be measured, improved, and sometimes guaranteed. Characteristics measured for quality of service include but are not limited to throughput rates, CPU usage, and disk capacity. When you share physical hardware in an EMC Isilon cluster across multiple virtual instances, competition exists for the following services:

- CPU
- Memory
- Network bandwidth
- Disk I/O
- Disk capacity

Access zones do not provide logical quality of service guarantees to these resources, but you can partition these resources between access zones on a single cluster. The following table describes a few ways to partition resources to improve quality of service:

Use	Notes
NICs	You can assign specific NICs on specific nodes to an IP address pool that is associated with an access zone. By assigning these NICs, you can determine the nodes and interfaces that are associated with an access zone. This enables the separation of CPU, memory, and network bandwidth.
SmartPools	SmartPools are separated by node hardware equivalence classes, usually into multiple tiers of high, medium, and low performance. The data written to a SmartPool is written only to the disks in the nodes of that pool. Associating an IP address pool with only the nodes of a single SmartPool enables partitioning of disk I/O resources.
SmartQuotas	Through SmartQuotas, you can limit disk capacity by a user or a group or in a directory. By applying a quota to an access zone's base directory, you can limit disk capacity used in that access zone.

Managing access zones

You can create access zones on the EMC Isilon cluster, view and modify access zone settings, and delete access zones.

Create an access zone

You can create an access zone and define a base directory and authentication providers.

Procedure

1. Click **Access** › **Access Zones**.
2. Click **Create an access zone**.
3. In the **Access Zone Name** field, type a name for the access zone.

4. In the **Zone Base Directory** field, type the base directory path for the access zone.

   The path must be unique and not overlap with the path of any other access zone as specified by the base directory rules.

5. From the **Authentication Providers** list, select the method that you want to use to configure an authentication provider for this access zone.

   The available sub-settings differ depending on the method that you select for configuring authentication providers.

   - Select `Use all authentication providers` all authentication providers configured on the EMC Isilon cluster are assigned to the access zone.
   - Select `Manually select authentication providers`, and the **Add an Authentication Provider** button appears.

6. (Optional) If you chose to manually select providers, click **Add an Authentication Provider** in the **New Authentication Provider** area.

   a. From the **Authentication Provider Type** list, select a provider type. A provider type is listed only if an instance of that type exists and is not already in use by the access zone.

   b. From the **Authentication Provider** list, select an available provider instance.

   c. To change the order in which authentication providers are searched during authentication and user lookup, click the title bar of a provider instance and drag it to a new position in the list.

7. Click **Create Access Zone**.

After you finish

Before users can connect to an access zone, you must associate it with an IP address pool.

Associate an IP address pool with an access zone

You can specify which access zone that a user is to connect to by associating an IP address pool with the access zone.

Procedure

1. Click **Cluster Management** › **Network Configuration**.
2. In the **External Network Settings** area, under **Subnets**, click a subnet name—for example, **subnet0**.
3. In the **Basic Settings** area of the **IP Address Pools** settings, click **Edit** for the IP address pool you want to modify.

   The **Configure IP address pool settings** dialog box appears.

4. Select from the **Access Zone** list the access zone to use when connecting through an IP address that belongs to this pool.
5. Click **Submit**.

View a list of access zones

You can view a list of all access zones on the cluster.

Procedure

1. Click **Access** › **Access Zones**.

2. Click **View details** for more information about a specific access zone.

Modify an access zone

You can modify the properties of any access zone with one exception: You cannot change the name of the built-in System zone.

Procedure

1. Click **Access** › **Access Zones**.
2. For the access zone whose settings you want to modify, click **View details**.
3. For each setting that you want to modify, click **Edit**, make the change, and then click **Save**.

Delete an access zone

You can delete any access zone except the built-in System zone. When you delete an access zone, all associated authentication providers remain available to other zones, but IP addresses are not reassigned to other zones. SMB shares, NFS exports, and HDFS data paths are deleted when you delete an access zone; however, the directories and data still exist, and you can map new shares, exports, or paths in another access zone.

Procedure

1. Click **Access** › **Access Zones**.
2. From the table of access zones, click **Delete** next to the access zone that you want to delete.
3. In the **Confirm Delete** dialog box, click **Delete**.

CHAPTER 5

Authentication and access control

This section contains the following topics:

Authentication and access control overview

OneFS supports several methods for ensuring that your cluster remains secure, including UNIX- and Windows-style permissions for data-level access control, access zones for data isolation, and role-based administration control access to system configuration settings.

OneFS is designed for a mixed environment that allows you to configure both Access Control Lists (ACLs) and standard UNIX permissions on the cluster file system.

Note

In most situations, the default settings are sufficient. You can configure additional access zones, custom roles, and permissions policies as necessary for your particular environment.

Role-based access

You can assign role-based access to delegate administrative tasks to selected users.

Role based access control (RBAC) allows the right to perform particular administrative actions to be granted to any user who can authenticate to a cluster. Roles are created by a Security Administrator, assigned privileges, and then assigned members. All administrators, including those given privileges by a role, must connect to the System zone to configure the cluster. When these members log in to the cluster through a configuration interface, they have these privileges. All administrators can configure settings for access zones, and they always have control over all access zones on the cluster.

Roles also give you the ability to assign privileges to member users and groups. By default, only the root user and the admin user can log in to the web administration interface through HTTP or the command-line interface through SSH. Using roles, the root and admin users can assign others to built-in or customer roles that have login and administrative privileges to perform specific administrative tasks.

Note

As a best practice, assign users to roles that contain the minimum set of necessary privileges. For most purposes, the default permission policy settings, system access zone, and built-in roles are sufficient. You can create role-based access management policies as necessary for your particular environment.

Roles and privileges

In addition to controlling access to files and directories through ACLs and POSIX mode bits, OneFS controls configuration-level access through administrator roles. A role is a collection of OneFS privileges that are usually associated with a configuration subsystem. Those privileges are granted to members of that role as they log in to the cluster through the Platform API, command-line interface, or web administration interface

Roles

You can permit and limit access to administrative areas of your EMC Isilon cluster on a per-user basis through roles.

OneFS includes built-in administrator roles with predefined sets of privileges that cannot be modified. The following list describes what you can and cannot do through roles:

- You can assign privileges through role membership.
- You can add any user to a role as long as the user can authenticate to the cluster.
- You can create custom roles and assign privileges to those roles.
- You can add users singly or as groups, including well-known groups.
- You can assign a user as a member of more than one role.
- You can add a group to a role, which grants to all users who are members of that group all of the privileges associated with the role.
- You cannot assign privileges directly to users or groups.

Note

When OneFS is first installed, only users with root- or admin-level can log in and assign users to roles.

Built-in roles

Built-in roles include privileges to perform a set of administrative functions.

The following tables describe each of the built-in roles from most powerful to least powerful. The tables include the privileges and read/write access levels, if applicable, that are assigned to each role. You can assign users and groups to built-in roles and to roles that you create.

Table 1 SecurityAdmin role

Description	Privileges	Read/write access
Administer security configuration on the cluster, including authentication providers, local users and groups, and role membership.	ISI_PRIV_LOGIN_CONSOLE	N/A
	ISI_PRIV_LOGIN_PAPI	N/A
	ISI_PRIV_LOGIN_SSH	N/A
	ISI_PRIV_AUTH	Read/write
	ISI_PRIV_ROLE	Read/write

Table 2 SystemAdmin role

Description	Privileges	Read/write access
Administer all aspects of cluster configuration that are not specifically handled by the SecurityAdmin role.	ISI_PRIV_LOGIN_CONSOLE	N/A
	ISI_PRIV_LOGIN_PAPI	N/A
	ISI_PRIV_LOGIN_SSH	N/A
	ISI_PRIV_SYS_SHUTDOWN	N/A
	ISI_PRIV_SYS_SUPPORT	N/A
	ISI_PRIV_SYS_TIME	N/A
	ISI_PRIV_ANTIVIRUS	Read/write
	ISI_PRIV_AUDIT	Read/write
	ISI_PRIV_CLUSTER	Read/write

Table 2 SystemAdmin role (continued)

Description	Privileges	Read/write access
	ISI_PRIV_DEVICES	Read/write
	ISI_PRIV_EVENT	Read/write
	ISI_PRIV_FTP	Read/write
	ISI_PRIV_HDFS	Read/write
	ISI_PRIV_HTTP	Read/write
	ISI_PRIV_ISCSI	Read/write
	ISI_PRIV_JOB_ENGINE	Read/write
	ISI_PRIV_LICENSE	Read/write
	ISI_PRIV_NDMP	Read/write
	ISI_PRIV_NETWORK	Read/write
	ISI_PRIV_NFS	Read/write
	ISI_PRIV_NTP	Read/write
	ISI_PRIV_QUOTA	Read/write
	ISI_PRIV_REMOTE_SUPPORT	Read/write
	ISI_PRIV_SMARTPOOLS	Read/write
	ISI_PRIV_SMB	Read/write
	ISI_PRIV_SNAPSHOT	Read/write
	ISI_PRIV_STATISTICS	Read/write
	ISI_PRIV_SYNCIQ	Read/write
	ISI_PRIV_VCENTER	Read/write
	ISI_PRIV_WORM	Read/write
	ISI_PRIV_NS_TRAVERSE	N/A
	ISI_PRIV_NS_IFS_ACCESS	N/A

Table 3 AuditAdmin role

Description	Privileges	Read/write access
View all system configuration settings.	ISI_PRIV_LOGIN_CONSOLE	N/A
	ISI_PRIV_LOGIN_PAPI	N/A
	ISI_PRIV_LOGIN_SSH	N/A
	ISI_PRIV_ANTIVIRUS	Read-only
	ISI_PRIV_AUDIT	Read-only
	ISI_PRIV_CLUSTER	Read-only

Table 3 AuditAdmin role (continued)

Description	Privileges	Read/write access
	ISI_PRIV_DEVICES	Read-only
	ISI_PRIV_EVENT	Read-only
	ISI_PRIV_FTP	Read-only
	ISI_PRIV_HDFS	Read-only
	ISI_PRIV_HTTP	Read-only
	ISI_PRIV_ISCSI	Read-only
	ISI_PRIV_JOB_ENGINE	Read-only
	ISI_PRIV_LICENSE	Read-only
	SI_PRIV_NDMP	Read-only
	ISI_PRIV_NETWORK	Read-only
	ISI_PRIV_NFS	Read-only
	ISI_PRIV_NTP	Read-only
	ISI_PRIV_QUOTA	Read-only
	ISI_PRIV_REMOTE_SUPPORT	Read-only
	ISI_PRIV_SMARTPOOLS	Read-only
	ISI_PRIV_SMB	Read-only
	ISI_PRIV_SNAPSHOT	Read-only
	ISI_PRIV_STATISTICS	Read-only
	ISI_PRIV_SYNCIQ	Read-only
	ISI_PRIV_VCENTER	Read-only
	ISI_PRIV_WORM	Read-only

Table 4 VMwareAdmin role

Description	Privileges	Read/write access
Administers remotely all aspects of storage needed by VMware vCenter.	ISI_PRIV_LOGIN_PAPI	N/A
	ISI_PRIV_ISCSI	Read/write
	ISI_PRIV_NETWORK	Read/write
	ISI_PRIV_SMARTPOOLS	Read/write
	ISI_PRIV_SNAPSHOT	Read/write
	ISI_PRIV_SYNCIQ	Read/write
	ISI_PRIV_VCENTER	Read/write
	ISI_PRIV_NS_TRAVERSE	N/A

Table 4 VMwareAdmin role (continued)

Description	Privileges	Read/write access
	ISI_PRIV_NS_IFS_ACCESS	N/A

Table 5 BackupAdmin role

Description	Privileges	Read/write access
Allows backup and restore of files from `/ifs`	ISI_PRIV_IFS_BACKUP	Read-only
	ISI_PRIV_IFS_RESTORE	Read-only

Custom roles

Custom roles supplement built-in roles.

You can create custom roles and assign privileges mapped to administrative areas in your EMC Isilon cluster environment. For example, you can create separate administrator roles for security, auditing, storage provisioning, and backup.

You can designate certain privileges as read-only or read/write when adding the privilege to a role. You can modify this option at any time.

You can add or remove privileges as user responsibilities grow and change.

Privileges

Privileges permit users to complete tasks on an EMC Isilon cluster.

Privileges are associated with an area of cluster administration such as Job Engine, SMB, or statistics.

Privileges have one of two forms:

Action

Allows a user to perform a specific action on a cluster. For example, the ISI_PRIV_LOGIN_SSH privilege allows a user to log in to a cluster through an SSH client.

Read/Write

Allows a user to view or modify a configuration subsystem such as statistics, snapshots, or quotas. For example, the ISI_PRIV_SNAPSHOT privilege allows an administrator to create and delete snapshots and snapshot schedules. A read/write privilege can grant either read-only or read/write access. Read-only access allows a user to view configuration settings; read/write access allows a user to view and modify configuration settings.

Privileges are granted to the user on login to a cluster through the OneFS API, the web administration interface, SSH, or a console session. A token is generated for the user, which includes a list of all privileges granted to the user. Each URI, web-administration interface page, and command requires a specific privilege to view or modify the information available through any of these interfaces.

Note

Privileges are not granted to users that do not connect to the System Zone during login or to users that connect through the deprecated Telnet service, even if they are members of a role.

OneFS privileges

Privileges in OneFS are assigned through role membership; privileges cannot be assigned directly to users and groups.

Table 6 Login privileges

OneFS privilege	User right	Privilege type
ISI_PRIV_LOGIN_CONSOLE	Log in from the console	Action
ISI_PRIV_LOGIN_PAPI	Log in to the Platform API and the web administration interface	Action
ISI_PRIV_LOGIN_SSH	Log in through SSH	Action

Table 7 System privileges

OneFS privilege	User right	Privilege type
ISI_PRIV_SYS_SHUTDOWN	Shut down the system	Action
ISI_PRIV_SYS_SUPPORT	Run cluster diagnostic tools	Action
ISI_PRIV_SYS_TIME	Change the system time	Action

Table 8 Security privileges

OneFS privilege	User right	Privilege type
ISI_PRIV_AUTH	Configure external authentication providers	Read/write
ISI_PRIV_ROLE	Create new roles and assign privileges	Read/write

Table 9 Configuration privileges

OneFS privilege	User right	Privilege type
ISI_PRIV_ANTIVIRUS	Configure antivirus scanning	Read/write
IS_PRIV_AUDIT	Configure audit capabilities	Read/write
ISI_PRIV_CLUSTER	Configure cluster identity and general settings	Read/write

Table 9 Configuration privileges (continued)

OneFS privilege	User right	Privilege type
ISI_PRIV_DEVICES	Create new roles and assign privileges	Read/write
ISI_PRIV_EVENT	View and modify system events	Read/write
ISI_PRIV_FTP	Configure FTP server	Read/write
ISI_PRIV_HDFS	Configure HDFS server	Read/write
ISI_PRIV_HTTP	Configure HTTP server	Read/write
ISI_PRIV_ISCSI	Configure iSCSI server	Read/write
ISI_PRIV_JOB_ENGINE	Schedule cluster-wide jobs	Read/write
ISI_PRIV_LICENSE	Activate OneFS software licenses	Read/write
ISI_PRIV_NDMP	Configure NDMP server	Read/write
ISI_PRIV_NETWORK	Configure network interfaces	Read/write
ISI_PRIV_NFS	Configure the NFS server	Read/write
ISI_PRIV_NTP	Configure NTP	Read/write
ISI_PRIV_QUOTA	Configure file system quotas	Read/write
ISI_PRIV_REMOTE_SUPPORT	Configure remote support	Read/write
ISI_PRIV_SMARTPOOLS	Configure storage pools	Read/write
ISI_PRIV_SMB	Configure the SMB server	Read/write
ISI_PRIV_SNAPSHOT	Schedule, take, and view snapshots	Read/write
ISI_PRIV_SNMP	Configure SNMP server	Read/write
ISI_PRIV_STATISTICS	View file system performance statistics	Read/write
ISI_PRIV_SYNCIQ	Configure SyncIQ	Read/write
ISI_PRIV_VCENTER	Configure VMware for vCenter	Read/write
ISI_PRIV_WORM	Configure SmartLock directories	Read/write

Table 10 Platform API-only privileges

OneFS privilege	User right	Privilege type
ISI_PRIV_EVENT	View and modify system events	Read/write
ISI_PRIV_LICENSE	Activate OneFS software licenses	Read/write
ISI_PRIV_STATISTICS	View file system performance statistics	Read/write

Table 11 File access privileges

OneFS privilege	User right	Privilege type
ISI_PRIV_IFS_BACKUP	Back up files from `/ifs`. **Note** This privilege circumvents traditional file access checks, such as mode bits or NTFS ACLs.	Action
ISI_PRIV_IFS_RESTORE	Restore files from `/ifs`. **Note** This privilege circumvents traditional file access checks, such as mode bits or NTFS ACLs.	Action

Command-line interface privileges

You can perform most tasks granted by a privilege through the command-line interface.

Some OneFS commands require root access, but if you do not have root access, you can perform most of the commands associated with a privilege through the `sudo` program. The system automatically generates a sudoers file of users based on existing roles. Prefixing a command with sudo allows you to run commands that require root access. For example, if you do not have root access, the following command fails:

```
isi alert list
```

If you are on the sudoers list because you are a member of a role that has the ISI_PRIV_EVENT privilege, the following command succeeds:

```
sudo isi alert list
```

The following tables list all One FS commands available, the associated privilege or root-access requirement, and whether `sudo` is required to run the command.

Note

If you are running in compliance mode, additional sudo commands are available.

Table 12 Privileges sorted by CLI command

isi command	Privilege	Requires sudo
isi alert	ISI_PRIV_EVENT	x
isi audit	ISI_PRIV_AUDIT	
isi auth - excluding isi auth role	ISI_PRIV_AUTH	
isi auth role	ISI_PRIV_ROLE	
isi avscan	ISI_PRIV_ANTIVIRUS	x
isi batterystatus	ISI_PRIV_STATISTICS	x
isi config	root	
isi dedupe - excluding isi dedupe stats	ISI_PRIV_JOB_ENGINE	
isi dedupe stats	ISI_PRIV_STATISTICS	
isi devices	ISI_PRIV_DEVICES	x
isi drivefirmware	root	
isi domain	root	
isi email	ISI_PRIV_CLUSTER	x
isi events	ISI_PRIV_EVENT	x
isi exttools	root	
isi fc	root	
isi filepool	ISI_PRIV_SMARTPOOLS	
isi firmware	root	
isi ftp	ISI_PRIV_FTP	x
isi get	root	
isi hdfs	ISI_PRIV_HDFS	
isi iscsi	ISI_PRIV_ISCSI	x
isi job	ISI_PRIV_JOB_ENGINE	
isi license	ISI_PRIV_LICENSE	x
isi lun	ISI_PRIV_ISCSI	x
isi ndmp	ISI_PRIV_NDMP	x
isi networks	ISI_PRIV_NETWORK	x
isi nfs	ISI_PRIV_NFS	
isi perfstat	ISI_PRIV_STATISTICS	x

Table 12 Privileges sorted by CLI command (continued)

isi command	Privilege	Requires sudo
isi pkg	root	
isi quota	ISI_PRIV_QUOTA	
isi readonly	root	
isi remotesupport	ISI_PRIV_REMOTE_SUPPORT	
isi servicelight	ISI_PRIV_DEVICES	x
isi services	root	
isi set	root	
isi smb	ISI_PRIV_SMB	
isi snapshot	ISI_PRIV_SNAPSHOT	
isi snmp	ISI_PRIV_SNMP	x
isi stat	ISI_PRIV_STATISTICS	x
isi statistics	ISI_PRIV_STATISTICS	x
isi status	ISI_PRIV_STATISTICS	x
isi storagepool	ISI_PRIV_SMARTPOOLS	
isi sync	ISI_PRIV_SYNCIQ	
isi tape	ISI_PRIV_NDMP	x
isi target	ISI_PRIV_ISCSI	x
isi update	root	
isi version	ISI_PRIV_CLUSTER	x
isi worm	ISI_PRIV_WORM	
isi zone	ISI_PRIV_AUTH	

Table 13 CLI commands sorted by privilege

Privilege	isi commands	Requires sudo
ISI_PRIV_ANTIVIRUS	• isi avscan	x
ISI_PRIV_AUDIT	• isi audit	
ISI_PRIV_AUTH	• isi auth - excluding isi auth role • isi zone	
ISI_PRIV_IFS_BACKUP	N/A	N/A
ISI_PRIV_CLUSTER	• isi email • isi version	x

Table 13 CLI commands sorted by privilege (continued)

Privilege	isi commands	Requires sudo
ISI_PRIV_DEVICES	• isi devices • isi servicelight	x
ISI_PRIV_EVENT	• isi alert • isi events	x
ISI_PRIV_FTP	• isi ftp	x
ISI_PRIV_HDFS	• isi hdfs	
ISI_PRIV_ISCSI	• isi iscsi • isi lun • isi target	x
ISI_PRIV_JOB_ENGINE	• isi job • isi dedupe - excluding isi dedupe stats	
ISI_PRIV_LICENSE	• isi license	x
ISI_PRIV_NDMP	• isi ndmp • isi tape	x
ISI_PRIV_NETWORK	• isi networks	x
ISI_PRIV_NFS	• isi nfs	
ISI_PRIV_QUOTA	• isi quota	
ISI_PRIV_ROLE	• isi auth role	
ISI_PRIV_REMOTE_SUPPORT	• isi remotesupport	
ISI_PRIV_IFS_RESTORE	N/A	N/A
ISI_PRIV_SMARTPOOLS	• isi filepool • isi storagepool	
ISI_PRIV_SMB	• isi smb	
ISI_PRIV_SNAPSHOT	• isi snapshot	
ISI_PRIV_SNMP	• isi snmp	x
ISI_PRIV_STATISTICS	• isi batterystatus • isi dedupe stats • isi perfstat • isi stat	x

Table 13 CLI commands sorted by privilege (continued)

Privilege	isi commands	Requires sudo
	• isi statistics • isi status	
ISI_PRIV_SYNCIQ	• isi sync	
ISI_PRIV_WORM	• isi worm	
root	• isi config • isi domain • isi drivefirmware • isi exttools • isi fc • isi firmware • isi get • isi pkg • isi readonly • isi services • isi set • isi update	

Data backup and restore privileges

You can assign privileges to a user that are explicitly for cluster data backup and restore actions.

Two privileges allow a user to backup and restore cluster data over supported client-side protocols: ISI_PRIV_IFS_BACKUP and ISI_PRIV_IFS_RESTORE.

⚠ CAUTION

These privileges circumvent traditional file access checks, such as mode bits or NTFS ACLs.

Most cluster privileges allow changes to cluster configuration in some manner. The backup and restore privileges allow access to cluster data from the System zone, the traversing of all directories, and reading of all file data and metadata regardless of file permissions.

Users assigned these privileges use the protocol as a backup protocol to another machine without generating access-denied errors and without connecting as the root user. These two privileges are supported over the following client-side protocols:

- SMB
- RAN API
- FTP
- SSH

Over SMB, the ISI_PRIV_IFS_BACKUP and ISI_PRIV_IFS_RESTORE privileges emulate the Windows privileges SE_BACKUP_NAME and SE_RESTORE_NAME. The emulation means that normal file-open procedures are protected by file system permissions. To enable the backup and restore privileges over the SMB protocol, you must open files with the FILE_OPEN_FOR_BACKUP_INTENT option, which occurs automatically through Windows backup software such as Robocopy. Application of the option is not automatic when files are opened through general file browsing software such as Windows File Explorer.

Both ISI_PRIV_IFS_BACKUP and ISI_PRIV_IFS_RESTORE privileges primarily support Windows backup tools such as Robocopy. A user must be a member of the BackupAdmin built-in role to access all Robocopy features, which includes copying file DACL and SACL metadata.

User permissions utility

You can view expected user or group permissions to a given file or directory with the expected user permissions utility.

The command-line interface expected user permissions utility provides quick discovery of user and group permissions, displaying ACL or mode bits permissions. The utility does not display privileges or SMB share permissions, however.

Note

You must be a member of a role that has ISI_PRIV_LOGIN_SSH and ISI_PRIV_AUTH privileges to run this utility.

For information about viewing group or user permissions, see the View expected user permissions topic.

Authentication

OneFS supports local and remote authentication providers to verify that users attempting to access an EMC Isilon cluster are who they claim to be. Anonymous access, which does not require authentication, is supported for protocols that allow it.

OneFS supports concurrent multiple authentication provider types, which are analogous to directory services. For example, OneFS is often configured to authenticate Windows clients with Active Directory and to authenticate UNIX clients with LDAP. You can also configure NIS, designed by Sun Microsystems, to authenticate users and groups when they access a cluster.

Note

OneFS is RFC 2307-compliant.

Supported authentication providers

You can configure local and remote authentication providers to authenticate or deny user access to an EMC Isilon cluster.

The following table compares features that are available with each of the authentication providers that OneFS supports. In the following table, an x indicates that a feature is fully supported by a provider; an asterisk (*) indicates that additional configuration or support from another provider is required.

Authentication provider	NTLM	Kerberos	User/group management	Netgroups	UNIX properties (RFC 2307)	Windows properties
Active Directory	x	x			*	x
LDAP	*	x		x	x	*
NIS				x	x	
Local	x		x		x	x
File	x			x	x	
MIT Kerberos		x		*	*	*

Authentication provider features

You can configure authentication providers for your environment.

Authentication providers support a mix of the features described in the following table.

Feature	Description
Authentication	All authentication providers support plain-text authentication. You can configure some providers to support NTLM or Kerberos authentication also.
Users and groups	OneFS provides the ability to manage users and groups directly on the cluster.
Netgroups	Specific to NFS, netgroups restrict access to NFS exports.
UNIX-centric user and group properties	Login shell, home directory, UID, and GID. Missing information is supplemented by configuration templates or additional authentication providers.
Windows-centric user and group properties	NetBIOS domain and SID. Missing information is supplemented by configuration templates.

Kerberos authentication

Kerberos is a network authentication provider that negotiates encryption tickets for securing a connection. OneFS supports Active Directory Kerberos and MIT Kerberos authentication providers on an EMC Isilon cluster. If you configure an Active Directory provider, Kerberos authentication is provided automatically. MIT Kerberos works independently of Active Directory.

For MIT Kerberos authentication, you define an administrative domain known as a realm. Within this realm, an authentication server has the authority to authenticate a user, host, or service. You can optionally define a Kerberos domain to allow additional domain extensions to be associated with a realm.

The authentication server in a Kerberos environment is called the Key Distribution Center (KDC) and distributes encrypted tickets. When a user authenticates with an MIT Kerberos provider within a realm, an encrypted ticket with the user's service principal name (SPN) is created and validated to securely pass the user's identification for the requested service.

You can include an MIT Kerberos provider in specific access zones for authentication. Each access zone may include at most one MIT Kerberos provider. You can discontinue authentication through an MIT Kerberos provider by removing the provider from all the referenced access zones.

Keytabs and SPNs overview

A Key Distribution Center (KDC) is an authentication server that stores accounts and keytabs for users connecting to a network service within an EMC Isilon cluster. A keytab is a key table that stores keys to validate and encrypt Kerberos tickets.

One of the fields in a keytab entry is a service principal name (SPN). An SPN identifies a unique service instance within a cluster. Each SPN is associated with a specific key in the KDC. Users can use the SPN and its associated keys to obtain Kerberos tickets that enable access to various services on the cluster. A member of the SecurityAdmin role can create new keys for the SPNs and modify them later as necessary. An SPN for a service typically appears as `<service>/<fqdn>@<realm>`.

Note

SPNs must match the SmartConnect zone name and the FQDN hostname of the cluster. If the SmartConnect zone settings are changed, you must update the SPNs on the cluster to match the changes.

MIT Kerberos protocol support

MIT Kerberos supports certain standard network communication protocols such as HTTP, HDFS, and NFS. MIT Kerberos does not support SMB, SSH, and FTP protocols.

For the NFS protocol support, MIT Kerberos must be enabled for an export and also a Kerberos provider must be included within the access zone.

LDAP

The Lightweight Directory Access Protocol (LDAP) is a networking protocol that enables you to define, query, and modify directory services and resources.

OneFS can authenticate users and groups against an LDAP repository in order to grant them access to the cluster. OneFS supports Kerberos authentication for an LDAP provider.

The LDAP service supports the following features:

- Users, groups, and netgroups.
- Configurable LDAP schemas. For example, the ldapsam schema allows NTLM authentication over the SMB protocol for users with Windows-like attributes.
- Simple bind authentication, with and without SSL.
- Redundancy and load balancing across servers with identical directory data.
- Multiple LDAP provider instances for accessing servers with different user data.
- Encrypted passwords.

Active Directory

The Active Directory directory service is a Microsoft implementation of Lightweight Directory Access Protocol (LDAP), Kerberos, and DNS technologies that can store information about network resources. Active Directory can serve many functions, but the primary reason for joining the cluster to an Active Directory domain is to perform user and group authentication.

When the cluster joins an Active Directory domain, a single Active Directory machine account is created. The machine account establishes a trust relationship with the domain and enables the cluster to authenticate and authorize users in the Active Directory forest. By default, the machine account is named the same as the cluster. If the cluster name is more than 15 characters long, the name is hashed and displayed after joining the domain.

Note

If you configure an Active Directory provider, Kerberos authentication is provided automatically.

Whenever possible, observe the following guidelines when you configure Active Directory providers on a cluster:

- Configure a single Active Directory instance if all domains have a trust relationship.
- Configure multiple Active Directory instances only to grant access to multiple sets of mutually-untrusted domains.

NIS

The Network Information Service (NIS) provides authentication and identity uniformity across local area networks. OneFS includes an NIS authentication provider that enables you to integrate the cluster with your NIS infrastructure.

NIS, designed by Sun Microsystems, can authenticate users and groups when they access the cluster. The NIS provider exposes the passwd, group, and netgroup maps from an NIS server. Hostname lookups are also supported. You can specify multiple servers for redundancy and load balancing.

Note

NIS is different from NIS+, which OneFS does not support.

File provider

A file provider enables you to supply an authoritative third-party source of user and group information to an EMC Isilon cluster. A third-party source is useful in UNIX and Linux environments that synchronize `/etc/passwd`, `/etc/group`, and `etc/netgroup` files across multiple servers.

Standard BSD `/etc/spwd.db` and `/etc/group` database files serve as the file provider backing store on a cluster. You generate the `spwd.db` file by running the `pwd_mkdb` command in the OneFS command-line interface (CLI). You can script updates to the database files.

On an Isilon cluster, a file provider hashes passwords with `libcrypt`. For the best security, we recommend that you use the Modular Crypt Format in the source `/etc/passwd` file to determine the hashing algorithm. OneFS supports the following algorithms for the Modular Crypt Format:

- MD5
- NT-Hash
- SHA-256
- SHA-512

For information about other available password formats, run the `man 3 crypt` command in the CLI to view the crypt man pages.

Note

The built-in System file provider includes services to list, manage, and authenticate against system accounts such as root, admin, and nobody. We recommended that you do not modify the System file provider.

Local provider

The local provider provides authentication and lookup facilities for user accounts added by an administrator.

Local authentication is useful when Active Directory, LDAP, or NIS directory services are not configured or when a specific user or application needs access to the cluster. Local groups can include built-in groups and Active Directory groups as members.

In addition to configuring network-based authentication sources, you can manage local users and groups by configuring a local password policy for each node in the cluster. OneFS settings specify password complexity, password age and re-use, and password-attempt lockout policies.

Data access control

You can configure an EMC Isilon cluster so that both UNIX and Windows users have access to content over NFS and SMB, regardless of the protocol that stored the data.

The OneFS operating system supports multiprotocol data access over Server Message Block (SMB) and Network File System (NFS) with a unified security model. For NFS, the default export on a cluster, `/ifs`, enables Linux and UNIX clients to remotely mount any subdirectory, including subdirectories created by Windows users. Linux and UNIX clients also can mount ACL-protected subdirectories created by a OneFS administrator.

Conversely, for SMB the default file share on a cluster, `/ifs`, provides Windows users access to file system resources over the network that includes resources stored by UNIX and Linux systems. The same access model applies to directories and files.

By default, OneFS maintains the same file permissions regardless of the client’s operating system, the user’s identity management system, or the file sharing protocol. When OneFS must transform a file’s permissions from ACLs to mode bits or vice versa, it merges the permissions into an optimal representation that uniquely balances user expectations and file security.

Authorization

OneFS supports two types of authorization data on a file: Windows-style access control lists (ACLs) and POSIX mode bits (UNIX permissions). Authorization type is based on the ACL policies that are set and on the file-creation method.

Access to a file or directory is governed by either a Windows access control list (ACL) or UNIX mode bits. Regardless of the security model, OneFS enforces access rights consistently across access protocols. A user is granted or denied the same rights to a file when using SMB for Windows file sharing as when using NFS for UNIX file sharing.

An EMC Isilon cluster includes global policy settings that enable you to customize the default ACL and UNIX permissions to best support your environment. Generally, files that are created over SMB or in a directory that has an ACL receive an ACL; otherwise, OneFS relies on the POSIX mode bits that define UNIX permissions. In either case, the owner is represented by a UNIX identifier (UID or GID) or by its Windows identifier (SID). The primary group is represented by a GID or SID. Although mode bits are present when a file has an ACL, the mode bits are provided for only protocol compatibility, not for access checks.

Note

Although you can configure ACL policies to optimize a cluster for UNIX or Windows, you should do so only if you understand how ACL and UNIX permissions interact.

The OneFS file system installs with UNIX permissions as the default. By using Windows Explorer or OneFS administrative tools, you can give a file or directory an ACL. In addition to Windows domain users and groups, ACLs in OneFS can include local, NIS, and LDAP users and groups. After you give a file an ACL, OneFS stops enforcing the file's mode bits, which remain only as an estimate of the effective permissions.

SMB

You can configure SMB shares to provide Windows clients network access to file system resources on the cluster. You can grant permissions to users and groups to carry out operations such as reading, writing, and setting access permissions on SMB shares.

ACLs

In Windows environments, file and directory permissions, referred to as access rights, are defined in access control lists (ACLs). Although ACLs are more complex than mode bits, ACLs can express much more granular sets of access rules. OneFS checks the ACL processing rules commonly associated with Windows ACLs.

A Windows ACL contains zero or more access control entries (ACEs), each of which represents the security identifier (SID) of a user or a group as a trustee. In OneFS, an ACL can contain ACEs with a UID, GID, or SID as the trustee. Each ACE contains a set of rights that allow or deny access to a file or folder. An ACE can optionally contain an inheritance flag to specify whether the ACE should be inherited by child folders and files.

Note

Instead of the standard three permissions available for mode bits, ACLs have 32 bits of fine-grained access rights. Of these, the upper 16 bits are general and apply to all object types. The lower 16 bits vary between files and directories but are defined in a way that allows most applications to apply the same bits for files and directories.

Rights grant or deny access for a given trustee. You can block user access explicitly through a deny ACE or implicitly by ensuring that a user does not directly, or indirectly through a group, appear in an ACE that grants the right.

NFS

You can configure NFS exports to provide UNIX clients network access to file system resources on the cluster.

UNIX permissions

In a UNIX environment, file and directory access is controlled by POSIX mode bits, which grant read, write, or execute permissions to the owning user, the owning group, and everyone else.

OneFS supports the standard UNIX tools for viewing and changing permissions, `ls`, `chmod`, and `chown`. For more information, run the `man ls`, `man chmod`, and `man chown` commands.

All files contain 16 permission bits, which provide information about the file or directory type and the permissions. The lower 9 bits are grouped as three 3-bit sets, called triples, which contain the read, write, and execute (rwx) permissions for each class of users—owner, group, and other. You can set permissions flags to grant permissions to each of these classes.

Unless the user is root, OneFS checks the class to determine whether to grant or deny access to the file. The classes are not cumulative: The first class matched is applied. It is therefore common to grant permissions in decreasing order.

Mixed-permission environments

When a file operation requests an object’s authorization data, for example, with the `ls –l` command over NFS or with the **Security** tab of the **Properties** dialog box in Windows Explorer over SMB, OneFS attempts to provide that data in the requested format. In an environment that mixes UNIX and Windows systems, some translation may be required when performing create file, set security, get security, or access operations.

NFS access of Windows-created files

If a file contains an owning user or group that is a SID, the system attempts to map it to a corresponding UID or GID before returning it to the caller.

In UNIX, authorization data is retrieved by calling `stat(2)` on a file and examining the owner, group, and mode bits. Over NFSv3, the GETATTR command functions similarly. The system approximates the mode bits and sets them on the file whenever its ACL changes. Mode bit approximations need to be retrieved only to service these calls.

Note

SID-to-UID and SID-to-GID mappings are cached in both the OneFS ID mapper and the `stat` cache. If a mapping has recently changed, the file might report inaccurate information until the file is updated or the cache is flushed.

SMB access of UNIX-created files

No UID-to-SID or GID-to-SID mappings are performed when creating an ACL for a file; all UIDs and GIDs are converted to SIDs or principals when the ACL is returned.

OneFS initiates a two-step process for returning a security descriptor, which contains SIDs for the owner and primary group of an object:

1. The current security descriptor is retrieved from the file. If the file does not have a discretionary access control list (DACL), a synthetic ACL is constructed from the file’s lower 9 mode bits, which are separated into three sets of permission triples—one each for owner, group, and everyone. For details about mode bits, see the UNIX permissions topic.
2. Two access control entries (ACEs) are created for each triple: the allow ACE contains the corresponding rights that are granted according to the permissions; the deny ACE

contains the corresponding rights that are denied. In both cases, the trustee of the ACE corresponds to the file owner, group, or everyone. After all of the ACEs are generated, any that are not needed are removed before the synthetic ACL is returned.

Managing roles

You can view, add, or remove members of any role. Except for built-in roles, whose privileges you cannot modify, you can add or remove OneFS privileges on a role-by-role basis.

Note

Roles take both users and groups as members. If a group is added to a role, all users who are members of that group are assigned the privileges associated with the role. Similarly, members of multiple roles are assigned the combined privileges of each role.

View a role

You can view information about built-in and custom roles.

Procedure

1. Click **Access** › **Membership & Roles** › **Roles**.
2. In the **Roles** area, select a role and click **View** / **Edit**.
3. In the **View Role Details** dialog box, view information about the role.
4. Click **Close** to return to the **Membership & Roles** page.

View privileges

You can view user privileges.

This procedure must be performed through the command-line interface (CLI). You can view a list of your privileges or the privileges of another user using the following commands:

Procedure

1. Establish an SSH connection to any node in the cluster.
2. To view privileges, run one of the following commands.
   - To view a list of all privileges, run the following command:

     ```
     isi auth privileges --verbose
     ```

   - To view a list of your privileges, run the following command:

     ```
     isi auth id
     ```

   - To view a list of privileges for another user, run the following command, where *<user>* is a placeholder for another user by name:

     ```
     isi auth mapping token <user>
     ```

Create a custom role

You can create a custom role and add privileges and members to that role.

Procedure

1. Click **Access** › **Membership & Roles** › **Roles**.
2. Click **Create a Role**.
3. In the **Role Name** field, type a name for the role.
4. In the **Description** field, type a description.
5. Click **Add a member to this role** to add a member to the role.
6. Click **Add a privilege to this role** to assign access rights and privileges.
7. Click **Create Role**.

Modify a role

You can modify the description and the user or group membership of any role, including built-in roles. However, you can modify the name and privileges only for custom roles.

Procedure

1. Click **Access** › **Membership & Roles** › **Roles**.
2. In the **Roles** area, select a role and click **View** / **Edit**.

   The **View Role Details** dialog box appears.
3. Click **Edit Role** and modify the settings as needed in the **Edit Role Details** dialog box.
4. Click **Save Changes** to return to the **View Role Details** dialog box.
5. Click **Close**.

Copy a role

You can copy an existing role and add or remove privileges and members for that role as needed.

Procedure

1. Click **Access** › **Membership & Roles** › **Roles**.
2. In the **Roles** area, select a role and click **More** › **Copy**.
3. Modify the role name, description, members, and privileges as needed.
4. Click **Copy Role**.

Add a privilege to a custom role

You can add or remove privileges to a custom role as needed. You can designate certain privileges as read-only or read/write. You cannot modify the privileges assigned to a built-in role. Repeat this procedure for each privilege that you want to add to a custom role.

Procedure

1. Click **Add a privilege to this role** in the dialog box for creating, copying, or editing a role.
2. In the **Add a privilege to this role** dialog box, select an access type for the role.

3. Select a privilege from the list.
4. Click **Add Privilege**.

Add a member to a role

You can add one or more members to a role when creating, copying, or modifying the role. A user or a group can be a member of more than one role. The privileges associated with a role are granted to all members of that role. Repeat this procedure to add more members to the role.

Procedure

1. Click **Add a member to this role** in the dialog box for creating, copying, or editing a role.
2. In the **Select a User** dialog box, select one of following options:
   - **Users**
   - **Groups**
   - **Well-known SIDs**
3. If you selected **User** or **Group**, locate the user or group through one of the following methods:
   - Type the Username or Group Name you want to search for in the text field.
   - Select the authentication provider you want to search for from the **Provider** list. Only providers that are currently configured and enabled on the cluster are listed.
4. Click **Search**.
5. Select a user name, group name, or a well-known SID from the search results to add as members to the role.
6. Click **Select**.

Delete a custom role

Deleting a custom role does not affect the privileges or users that are assigned to it. You cannot delete built-in roles.

Procedure

1. Click **Access** › **Membership & Roles** › **Roles**.
2. In the **Roles** area, select one or more roles, and then perform one of the following actions:
   - To delete a single role, click **More** › **Delete** from the **Actions** column against the selected role.
   - To delete multiple roles, select **Delete Selection** from the **Select a bulk action** list.
3. In the confirmation dialog box, click **Delete**.

Managing authentication providers

You can configure one or more LDAP, Active Directory, NIS, file, and Kerberos providers. A local provider is created automatically when you create an access zone, which allows you to create a configuration for each access zone so it has its own list of local users that can authenticate to it. You also can create a password policy for each local provider to enforce password complexity.

Managing LDAP providers

You can view, configure, modify, and delete LDAP providers. You can discontinue authentication through an LDAP provider by removing it from all access zones that are using it.

Configure an LDAP provider

By default, when you configure an LDAP provider, it is automatically added to the System access zone.

Procedure

1. Click **Access** › **Authentication Providers** › **LDAP**.
2. Click **Add an LDAP provider**.
3. In the **LDAP Provider Name** field, type a name for the provider.
4. In the **Server URIs** field, type one or more valid LDAP server URIs, one per line, in the format ldaps://*<server>*:*<port>* (secure LDAP) or ldap://*<server>*:*<port>* (non-secure LDAP).

   Note

   If you do not specify a port, the default port is used. The default port for non-secure LDAP (ldap://) is 389; for secure LDAP (ldaps://), it is 636. If you specify non-secure LDAP, the bind password is transmitted to the server in clear text.

5. (Optional) Configure the following settings as needed.

   Load balance servers
   Select the check box to connect to a random server, or clear the check box to connect according to the order in which the servers are listed in the **Servers** field.

   Base Distinguished Name
   Type the distinguished name (DN) of the entry at which to start LDAP searches. Base DNs can include `cn` (Common Name), `l` (Locality), `dc` (Domain Component), `ou` (Organizational Unit), or other components. For example, `dc=emc,dc=com` is a base DN for emc.com.

   Bind to
   Type the distinguished name of the entry at which to bind to the LDAP server.

   Password
   Specify the password to use when binding to the LDAP server. Use of this password does not require a secure connection; if the connection is not using Transport Layer Security (TLS), the password is sent in clear text.

6. (Optional) To modify the default settings for user, group, and netgroup queries, click **Default Query Settings**.
7. (Optional) To modify the settings for user queries and home directory provisioning, click **User Query Settings**.
8. (Optional) To modify the settings for group queries, click **Group Query Settings**.
9. (Optional) To modify the settings for netgroup queries, click **Netgroup Query Settings**.
10. (Optional) To modify the default LDAP attributes that contain user information or to modify LDAP security settings, click **Advanced LDAP Settings**.

11.Click **Add LDAP Provider**.

Modify an LDAP provider

You can modify any setting for an LDAP provider except its name. You must specify at least one server for the provider to be enabled.

Procedure

1. Click **Access** › **Authentication Providers** › **LDAP**.
2. In the **LDAP Providers** table, click **View details** for the provider whose settings you want to modify.
3. For each setting that you want to modify, click **Edit**, make the change, and then click **Save**.
4. (Optional) Click **Close**.

Delete an LDAP provider

When you delete an LDAP provider, it is removed from all the access zones. As an alternative, you can stop using an LDAP provider by removing it from each access zone that contains it so that the provider remains available for future use.

Procedure

1. Click **Access** › **Authentication Providers** › **LDAP**.
2. In the **LDAP Providers** table, click **Delete** for the provider you want to delete.
3. In the confirmation dialog box, click **Delete**.

LDAP query settings

You can configure the entry point and depth at which to search for LDAP users, groups, and netgroups. You also can configure the settings for user home directory provisioning.

Note

OneFS is RFC 2307-compliant.

Distinguished Name

Specifies the base distinguished name (base DN) of the entry at which to start LDAP searches for user, group, or netgroup objects. Base DNs can include `cn` (Common Name), `l` (Locality), `dc` (Domain Component), `ou` (Organizational Unit), or other components. For example, `dc=emc,dc=com` is a base DN for emc.com.

Search Scope

Specifies the depth from the base DN at which to perform LDAP searches.
The following values are valid:

`default`

Applies the search scope that is defined in the default query settings. This option is not available for the default query search scope.

`base`

Searches only the entry at the base DN.

`onelevel`

Searches all entries exactly one level below the base DN.

`subtree`

Searches the base DN and all entries below it.

`children`

Searches all entries below the base DN, excluding the base DN itself.

Search Timeout
Specifies the number of seconds after which to stop retrying and fail a search. The default value is `100`. This setting is available only in the default query settings.

Query Filter
Specifies the LDAP filter for user, group, or netgroup objects. This setting is not available in the default query settings.

Authenticate users from this provider
Specifies whether to allow the provider to respond to authentication requests. This setting is available only in the user query settings.

Home Directory Naming
Specifies the path to use as a template for naming home directories. The path must begin with `/ifs` and can contain variables, such as %U, that are expanded to generate the home directory path for the user. This setting is available only in the user query settings.

Create home directories on first login
Specifies whether to create a home directory the first time a user logs in, if a home directory does not already exist for the user. This setting is available only in the user query settings.

UNIX Shell
Specifies the path to the user's login shell, for users who access the file system through SSH. This setting is available only in the user query settings.

LDAP advanced settings

You can configure LDAP security settings and specify the LDAP attributes that contain user information.

Note

OneFS is RFC 2307-compliant.

Name Attribute
Specifies the LDAP attribute that contains UIDs, which are used as login names. The default value is `uid`.

Common Name Attribute
Specifies the LDAP attribute that contains common names (CNs). The default value is `cn`.

Email Attribute
Specifies the LDAP attribute that contains email addresses. The default value is `mail`.

GECOS Field Attribute
Specifies the LDAP attribute that contains GECOS fields. The default value is `gecos`.

UID Attribute
Specifies the LDAP attribute that contains UID numbers. The default value is `uidNumber`.

GID Attribute
Specifies the LDAP attribute that contains GIDs. The default value is `gidNumber`.

Home Directory Attribute
Specifies the LDAP attribute that contains home directories. The default value is `homeDirectory`.

UNIX Shell Attribute
: Specifies the LDAP attribute that contains UNIX login shells. The default value is `loginShell`.

Netgroup Members Attribute
: Specifies the LDAP attribute that contains netgroup members. The default value is `memberNisNetgroup`.

Netgroup Triple Attribute
: Specifies the LDAP attribute that contains netgroup triples. The default value is `nisNetgroupTriple`.

Group Members Attribute
: Specifies the LDAP attribute that contains group members. The default value is `memberUid`.

Unique Group Members Attribute
: Specifies the LDAP attribute that contains unique group members. This attribute is used to determine which groups a user belongs to if the LDAP server is queried by the user's DN instead of the user's name. This setting has no default value.

UNIX Password Attribute
: Specifies the LDAP attribute that contains UNIX passwords. This setting has no default value.

Windows Password Attribute
: Specifies the LDAP attribute that contains Windows passwords. The default value is `ntpasswdhash`.

Certificate Authority File
: Specifies the full path to the root certificates file.

Require secure connection for passwords
: Specifies whether to require a Transport Layer Security (TLS) connection.

Ignore TLS Errors
: Continues over a secure connection even if identity checks fail.

Managing Active Directory providers

You can view, configure, modify, and delete Active Directory providers. OneFS includes a Kerberos configuration file for Active Directory in addition to the global Kerberos configuration file, both of which you can configure through the command-line interface. You can discontinue authentication through an Active Directory provider by removing it from all access zones that are using it.

Configure an Active Directory provider

You can configure one or more Active Directory providers, each of which must be joined to a separate Active Directory domain. By default, when you configure an Active Directory provider, it is automatically added to the System access zone.

Note

Consider the following information when you configure an Active Directory provider:

- When you join Active Directory from OneFS, cluster time is updated from the Active Directory server, as long as an NTP server has not been configured for the cluster.
- If you migrate users to a new or different Active Directory domain, you must re-set the ACL domain information after you configure the new provider. You can use third-party tools such as Microsoft SubInACL.

Procedure

1. Click **Access** › **Authentication Providers** › **Active Directory**.
2. Click **Join a domain**.
3. In the **Domain Name** field, type a fully qualified Active Directory domain name.

   The domain name will also be used as the provider name.
4. In the **User** field, type the username of an account that is authorized to join the Active Directory domain.
5. In the **Password** field, type the password of the user account.
6. (Optional) In the **Organizational Unit** field, type the name of the organizational unit (OU) to connect to on the Active Directory server. Specify the OU in the format *OuName* or *OuName1*/*SubName2*.
7. (Optional) In the **Machine Account** field, type the name of the machine account.

   Note

   If you specified an OU to connect to, the domain join will fail if the machine account does not reside in the OU.
8. (Optional) To enable Active Directory authentication for NFS, select the **Enable Secure NFS** check box.

   Note

   If you specified an OU to connect to, the domain join will fail if the machine account does not reside in the OU.

   If you enable this setting, OneFS registers NFS service principal names (SPNs) during the domain join.
9. (Optional) To configure advanced settings, click **Advanced Active Directory Settings**.
10. Click **Join**.

Modify an Active Directory provider

You can modify the advanced settings for an Active Directory provider.

Procedure

1. Click **Access** › **Authentication Providers** › **Active Directory**.
2. In the **Active Directory Providers** table, click **View details** for the provider whose settings you want to modify.
3. Click **Advanced Active Directory Settings**.
4. For each setting that you want to modify, click **Edit**, make the change, and then click **Save**.
5. (Optional) Click **Close**.

Delete an Active Directory provider

When you delete an Active Directory provider, you disconnect the cluster from the Active Directory domain that is associated with the provider, disrupting service for users who are accessing it. After you leave an Active Directory domain, users can no longer access the domain from the cluster.

Procedure

1. Click **Access** › **Authentication Providers** › **Active Directory**.
2. In the **Active Directory Providers** table, click **Leave** for the domain you want to leave.
3. In the confirmation dialog box, click **Leave**.

Active Directory provider settings

You can view or modify the advanced settings for an Active Directory provider.

Setting	Description
Services For UNIX	Specifies whether to support RFC 2307 attributes for domain controllers. RFC 2307 is required for Windows UNIX Integration and Services For UNIX technologies.
Map to primary domain	Enables the lookup of unqualified user names in the primary domain. If this setting is not enabled, the primary domain must be specified for each authentication operation.
Ignore trusted domains	Ignores all trusted domains.
Trusted Domains	Specifies trusted domains to include if the **Ignore Trusted Domains** setting is enabled.
Domains to Ignore	Specifies trusted domains to ignore even if the **Ignore Trusted Domains** setting is disabled.
Send notification when domain is unreachable	Sends an alert as specified in the global notification rules.
Use enhanced privacy and encryption	Encrypts communication to and from the domain controller.
Home Directory Naming	Specifies the path to use as a template for naming home directories. The path must begin with `/ifs` and can contain variables, such as %U, that are expanded to generate the home directory path for the user.

Setting	Description
Create home directories on first login	Creates a home directory the first time that a user logs in if a home directory does not already exist for the user.
UNIX Shell	Specifies the path to the login shell to use if the Active Directory server does not provide login-shell information. This setting applies only to users who access the file system through SSH.
Query all other providers for UID	If no UID is available in the Active Directory, looks up Active Directory users in all other providers for allocating a UID.
Match users with lowercase	If no UID is available in the Active Directory, normalizes Active Directory user names to lowercase before lookup.
Auto-assign UIDs	If no UID is available in the Active Directory, enables UID allocation for unmapped Active Directory users.
Query all other providers for GID	If no GID is available in the Active Directory, looks up Active Directory groups in all other providers before allocating a GID.
Match groups with lowercase	If no GID is available in the Active Directory, normalizes Active Directory group names to lowercase before lookup.
Auto-assign GIDs	If no GID is available in the Active Directory, enables GID allocation for unmapped Active Directory groups.
Make UID/GID assignments for users and groups in these specific domains	Restricts user and group lookups to the specified domains.

Managing NIS providers

You can view, configure, and modify NIS providers or delete providers that are no longer needed. You can discontinue authentication through an NIS provider by removing it from all access zones that are using it.

Configure an NIS provider

By default, when you configure an NIS provider it is automatically added to the System access zone.

Procedure

1. Click **Access** › **Authentication Providers** › **NIS**.
2. Click **Add a NIS provider**.
3. In the **NIS Provider Name** field, type a name for the provider.
4. In the **Servers** field, type one or more valid NIS server IP addresses, host names, or fully qualified domain names (FQDNs), separated by commas.

   Note

   If the **Load balance servers** option is not selected, servers are accessed in the order in which they are listed.

5. In the **NIS Domain** field, type the domain name.

6. (Optional) Configure the **Load balance servers** setting:
    - To connect to a random server, select the check box.
    - To connect according to the order in which the servers are listed in the **Servers** field, clear the check box.
7. (Optional) Specify the **Default Query Settings**.
    a. In the **Search Timeout** field, specifies the number of seconds after which to stop retrying and fail a search. The default value is 100.
    b. In the **Retry Frequency** field, specify the timeout period in seconds after which a request will be retried. The default value is 5.
8. (Optional) Specify the **User Query Settings**.
    a. Select the **Authenticate users from this provider** check box to allow the provider to respond to authentication requests.
    b. Type a path in the **Home Directory Naming** field to use as a template for naming home directories. The path must begin with `/ifs` and can contain expansion variables, such as %U, which expand to generate the home directory path for the user. For more information, see the Home directories section.
    c. Select the **Create home directories on first login** check box to specify whether to create a home directory the first time a user logs in, if a home directory does not already exist for the user.
    d. Select a path from the **UNIX Shell** list to specify the path to the user's login shell for users who access the file system through SSH.
9. (Optional) Click **Host Name Query Settings** and then configure the **Resolve hosts from this provider** setting:
    - To enable host resolution, select the check box.
    - To disable host resolution, clear the check box.
10. Click **Add NIS provider**.

Modify an NIS provider

You can modify any setting for an NIS provider except its name. You must specify at least one server for the provider to be enabled.

Procedure

1. Click **Access** › **Authentication Providers** › **NIS**.
2. In the **NIS Providers** table, click **View details** for the provider whose settings you want to modify.
3. For each setting that you want to modify, click **Edit**, make the change, and then click **Save**.
4. Click **Close**.

Delete an NIS provider

When you delete an NIS provider, it is removed from all access zones. As an alternative, you can stop using an NIS provider by removing it from each access zone that contains it so that the provider remains available for future use.

Procedure

1. Click **Access** › **Authentication Providers** › **NIS**.

2. In the **NIS Providers** table, click **Delete** for the provider that you want to delete.
3. In the confirmation dialog box, click **Delete**.

Managing file providers

You can configure one or more file providers, each with its own combination of replacement files, for each access zone. Password database files, which are also called user database files, must be in binary format.

Each file provider pulls directly from up to three replacement database files: a group file that has the same format as `/etc/group`; a netgroups file; and a binary password file, `spwd.db`, which provides fast access to the data in a file that has the `/etc/master.passwd` format. You must copy the replacement files to the cluster and reference them by their directory path.

Note

If the replacement files are located outside the `/ifs` directory tree, you must distribute them manually to every node in the cluster. Changes that are made to the system provider's files are automatically distributed across the cluster.

Configure a file provider

You can configure one or more file providers, each with its own combination of replacement files, for each access zone. You can specify replacement files for any combination of users, groups, and netgroups.

Procedure

1. Click **Access** › **Authentication Providers** › **File Provider**.
2. Click **Add a file provider**.
3. In the **File Provider Name** field, type a name for the file provider.
4. In the **File location for authentication information** area, specify one or more of the following files:
   - To specify a user replacement file, in the **Users File** field, type or browse to the location of the `spwd.db` file.
   - To specify a group replacement file, in the **Groups File** field, type or browse to the location of the `group` file.
   - To specify a netgroup replacement file, in the **Netgroups File** field, type or browse to the location of the `netgroup` file.
5. (Optional) Configure the following settings:

Option	Description
Authenticate users from this provider	Specifies whether to allow the provider to respond to authentication requests.
Home Directory Naming	Specifies the path to use as a template for naming home directories. The path must begin with `/ifs` and can contain expansion variables such as %U, which expand to generate the home directory path for the user. For more information, see the Home directories section.

Option	Description
Create home directories on first login	Specifies whether to create a home directory the first time a user logs in, if a home directory does not already exist for the user.
UNIX Shell	Specifies the path to the user's login shell, for users who access the file system through SSH.

6. Click **Add File Provider**.

Generate a password file

Password database files, which are also called user database files, must be in binary format.

This procedure must be performed through the command-line interface (CLI). For command-usage guidelines, run the `man pwd_mkdb` command.

Procedure

1. Establish an SSH connection to any node in the cluster.
2. Run the `pwd_mkdb` *<file>* command, where *<file>* is the location of the source password file.

   Note

   By default, the binary password file, `spwd.db`, is created in the `/etc` directory. You can override the location to store the `spwd.db` file by specifying the `-d` option with a different target directory.

   The following command generates an `spwd.db` file in the `/etc` directory from a password file that is located at `/ifs/test.passwd`:

   ```
   pwd_mkdb /ifs/test.passwd
   ```

   The following command generates an `spwd.db` file in the `/ifs` directory from a password file that is located at `/ifs/test.passwd`:

   ```
   pwd_mkdb -d /ifs /ifs/test.passwd
   ```

Password file format

The file provider uses a binary password database file, `spwd.db`. You can generate a binary password file from a `master.passwd`-formatted file by running the `pwd_mkdb` command.

The `master.passwd` file contains ten colon-separated fields, as shown in the following example:

```
admin:*:10:10::0:0:Web UI Administrator:/ifs/home/admin:/bin/zsh
```

The fields are defined below in the order in which they appear in the file.

Note

UNIX systems often define the passwd format as a subset of these fields, omitting the Class, Change, and Expiry fields. To convert a file from passwd to master.passwd format, add **:0:0:** between the GID field and the Gecos field.

Username
: The user name. This field is case-sensitive. OneFS does not limit the length; many applications truncate the name to 16 characters, however.

Password
: The user's encrypted password. If authentication is not required for the user, you can substitute an asterisk (*) for a password. The asterisk character is guaranteed to not match any password.

UID
: The UNIX user identifier. This value must be a number in the range 0-4294967294 that is not reserved or already assigned to a user. Compatibility issues occur if this value conflicts with an existing account's UID.

GID
: The group identifier of the user's primary group. All users are a member of at least one group, which is used for access checks and can also be used when creating files.

Class
: This field is not supported by OneFS and should be left empty.

Change
: OneFS does not support changing the passwords of users in the file provider. This field is ignored.

Expiry
: OneFS does not support the expiration of user accounts in the file provider. This field is ignored.

Gecos
: This field can store a variety of information but is usually used to store the user's full name.

Home
: The absolute path to the user's home directory, beginning at /ifs.

Shell
: The absolute path to the user's shell. If this field is set to /sbin/nologin, the user is denied command-line access.

Group file format

The file provider uses a group file in the format of the /etc/group file that exists on most UNIX systems.

The group file consists of one or more lines containing four colon-separated fields, as shown in the following example:

```
admin:*:10:root,admin
```

The fields are defined below in the order in which they appear in the file.

Group name
: The name of the group. This field is case-sensitive. Although OneFS does not limit the length of the group name, many applications truncate the name to 16 characters.

Password
: This field is not supported by OneFS and should contain an asterisk (*).

GID
: The UNIX group identifier. Valid values are any number in the range `0-4294967294` that is not reserved or already assigned to a group. Compatibility issues occur if this value conflicts with an existing group's GID.

Group members
: A comma-delimited list of user names.

Netgroup file format

A netgroup file consists of one or more netgroups, each of which can contain members. Hosts, users, or domains, which are members of a netgroup, are specified in a member triple. A netgroup can also contain another netgroup.

Each entry in a `netgroup` file consists of the netgroup name, followed by a space-delimited set of member triples and nested netgroup names. If you specify a nested netgroup, it must be defined on a separate line in the file.
A member triple takes the following form:

```
(<host>, <user>, <domain>)
```

Where *<host>* is a placeholder for a machine name, *<user>* is a placeholder for a user name, and *<domain>* is a placeholder for a domain name. Any combination is valid except an empty triple: `(,,)`.

The following sample file contains two netgroups. The rootgrp netgroup contains four hosts: two hosts are defined in member triples and two hosts are contained in the nested othergrp netgroup, which is defined on the second line.

```
rootgrp (myserver, root, somedomain.com) (otherserver, root,
somedomain.com) othergrp
othergrp (other-win,, somedomain.com) (other-linux,, somedomain.com)
```

Note

A new line signifies a new netgroup. You can continue a long netgroup entry to the next line by typing a backslash character (\) in the right-most position of the first line.

Modify a file provider

You can modify any setting for a file provider, with the exception that you cannot rename the System file provider.

Procedure

1. Click **Access** › **Authentication Providers** › **File Provider**.
2. In the **File Providers** table, click **View details** for the provider whose settings you want to modify.
3. For each setting that you want to modify, click **Edit**, make the change, and then click **Save**.
4. Click **Close**.

Delete a file provider

To stop using a file provider, you can clear all of its replacement file settings or you can permanently delete the provider.

Procedure

1. Click **Access** › **Authentication Providers** › **File Provider**.
2. In the **File Providers** table, select the provider name.
3. Select **Delete** from the **Select an action** list.
4. In the confirmation dialog box, click **Delete**.

Managing local users and groups

When you create an access zone, each zone includes a local provider that allows you to create and manage local users and groups. Although you can view the users and groups of any authentication provider, you can create, modify, and delete users and groups in the local provider only.

View a list of users or groups by provider

You can view the users and groups of any authentication provider.

Procedure

1. Click **Access** › **Membership & Roles**.
2. Click one of the following, depending on what you want to view:

Option	Description
Users	Select this tab to view all users by provider.
Groups	Select this tab to view all groups by provider.

3. From the **Current Access Zone** list, select an access zone.
4. Depending on your selection in step 2, select the local provider for the access zone from the **Users** list or the **Groups** list.

Create a local user

Each access zone includes a local provider that allows you to create and manage local users and groups. When creating a local user account, you can configure its name, password, home directory, UNIX user identifier (UID), UNIX login shell, and group memberships.

Procedure

1. Click **Access** › **Membership & Roles** › **Users**.
2. From the **Current Access Zone** list, select an access zone.
3. From the **Users** list, select the local provider for the zone.
4. Click **Create a user**.
5. In the **Username** field, type a username for the account.
6. In the **Password** field, type a password for the account.
7. (Optional) Configure the following additional settings as needed.

Option	Description
Allow password to expire	Select this check box to specify that the password is allowed to expire.
UID	If this setting is left blank, the system automatically allocates a UID for the account. This is the recommended setting. You cannot assign a UID that is in use by another local user account.
Full Name	Type a full name for the user.
Email Address	Type an email address for the account.
Primary Group	Click **Select group** to specify the owner group using the **Select a Primary Group** dialog box. a. To locate a group under the selected local provider, type a group name or click **Search**. b. Select a group to return to the **Manage Users** window.
Additional Groups	Click **Add group** to specify any additional groups to make this user a member of.
Home Directory	Type the path to the user's home directory. If you do not specify a path, a directory is automatically created at `/ifs/home/`*`<username>`*.
UNIX Shell	This setting applies only to users who access the file system through SSH. From the list, select a shell. By default, the **`/bin/zsh`** shell is selected.
Account Expiration Date	Click the calendar icon to select the expiration date or type the expiration date in the field, and then type the date in the format *<mm>/<dd>/<yyyy>*.
Enable the account	Select this check box to allow the user to authenticate against the local database for SSH, FTP, HTTP, and Windows file sharing through SMB. This setting is not used for UNIX file sharing through NFS.

8. Click **Create**.

Create a local group

In the local provider of an access zone, you can create groups and assign members to them.

Procedure

1. Click **Access** › **Membership & Roles** › **Groups**.
2. From the **Current Access Zone** list, select an access zone.
3. From the **Groups** list, select the local provider for the zone.
4. Click **Create a group**.
5. In the **Group Name** field, type a name for the group.
6. (Optional) To override automatic allocation of the UNIX group identifier (GID), in the **GID** field, type a numerical value.

Note

You cannot assign a GID that is in use by another group. It is recommended that you leave this field blank to allow the system to automatically generate the GID.

7. (Optional) For each member that you want to add to the group, click **Add user** and perform the following tasks in the **Select a User** dialog box:
   a. Search for either **Users** or **Well-known SIDs**.
   b. If you selected **Users**, specify values for the following fields:

      Username

      Type all or part of a user name, or leave the field blank to return all users. Wildcard characters are accepted.

      Access Zone

      Select the access zone that contains the authentication provider that you want to search.

      Provider

      Select an authentication provider.

   c. Click **Search**.
   d. In the **Search Results** table, select a user and then click **Select**.

      The dialog box closes.
8. Click **Create**.

Naming rules for local users and groups

Local user and group names must follow naming rules in order to ensure proper authentication and access to the EMC Isilon cluster.

You must adhere to the following naming rules when creating and modifying local users and groups:

- The maximum name length is 104 characters. It is recommended that names do not exceed 64 characters.
- Names cannot contain the following invalid characters:
  " / \ [ ] : ; | = , + * ? < >
- Names can contain any special character that is not in the list of invalid characters. It is recommend that names do not contain spaces.
- Names are not case sensitive.

Modify a local user

You can modify any setting for a local user account except the user name.

Procedure

1. Click **Access** › **Membership & Roles** › **Users**.
2. From the **Current Access Zone** list, select an access zone.
3. From the **Users** list, select the local provider for the access zone.
4. In the list of users, click **View details** for the local user whose settings you want to modify.
5. For each setting that you want to modify, click **Edit**, make the change, and then click **Save**.

6. Click **Close**.

Modify a local group

You can add or remove members from a local group.

Procedure

1. Click **Access** › **Membership & Roles** › **Groups**.
2. From the **Current Access Zone** list, select an access zone.
3. From the **Groups** list, select the local provider for the access zone.
4. In the list of groups, click **View details** for the local group whose settings you want to modify.
5. For the **Members** setting, click **Edit**.
6. Add or remove the users that you want, and then click **Save**.
7. Click **Close**.

Delete a local user

A deleted user can no longer access the cluster through the command-line interface, web administration interface, or file access protocol. When you delete a local user account, the corresponding home directory remains in place.

Procedure

1. Click **Access** › **Membership & Roles** › **Users**.
2. From the **Current Access Zone** list, select an access zone.
3. From the **Users** list, select the local provider for the access zone.
4. Click **Delete** for the user that you want to delete.
5. In the confirmation dialog box, click **Delete**.

Delete a local group

You can delete a local group even if members are assigned to it; deleting a group does not affect the members of that group.

Procedure

1. Click **Access** › **Membership & Roles** › **Groups**.
2. From the **Current Access Zone** list, select an access zone.
3. From the **Groups** list, select the local provider for the access zone.
4. Click **Delete** for the group that you want to delete.
5. In the confirmation dialog box, click **Delete**.

Naming rules for local users and groups

Local user and group names must follow naming rules in order to ensure proper authentication and access to the EMC Isilon cluster.

You must adhere to the following naming rules when creating and modifying local users and groups:

- The maximum name length is 104 characters. It is recommended that names do not exceed 64 characters.
- Names cannot contain the following invalid characters:
  " / \ [ ] : ; | = , + * ? ‹ ›

- Names can contain any special character that is not in the list of invalid characters. It is recommend that names do not contain spaces.
- Names are not case sensitive.

Managing MIT Kerberos authentication

You can configure an MIT Kerberos provider for authentication without Active Directory. Configuring an MIT Kerberos provider involves creating an MIT Kerberos realm, creating a provider, and joining a predefined realm. Optionally, you can configure an MIT Kerberos domain for the provider. You can also update the encryption keys if there are any configuration changes to the Kerberos provider. You can include the provider in one or more access zones.

Managing MIT Kerberos realms

An MIT Kerberos realm is an administrative domain that defines the boundaries within which an authentication server has the authority to authenticate a user or service. You can create, view, edit, or delete a realm. As a best practice, specify a realm name using uppercase characters.

Create an MIT Kerberos realm

An MIT Kerberos realm is an administrative domain that defines the boundaries within which an authentication server has the authority to authenticate a user or service. You can create a realm by defining a Key Distribution Center (KDC) and an administrative server.

Procedure

1. Click **Access** › **Authentication Providers** › **Kerberos Provider**.
2. Click **Create a Kerberos Realm**.
3. In the **Realm Name** field, type a domain name in uppercase characters. For example, `CLUSTER-NAME.COMPANY.COM`.
4. Select the **Set as the default realm** check box to set the realm as the default.
5. In the **Key Distribution Centers (KDCs)** field add one or more KDCs by specifying the hostnames or IP addresses of the authentication servers.
6. (Optional) In the **Admin Server** field, type the hostname or the IP address of the authentication server to use as the master KDC. If you omit this step, the first KDC that you added previously is used as the default administrative server.
7. (Optional) In the **Default Domain** field, specify the domain name to use for translating the service principal names.
8. Click **Create Realm**.

Modify an MIT Kerberos realm

You can modify an MIT Kerberos realm by modifying the Key Distribution Center (KDC) and the administrative server settings for that realm.

Procedure

1. Click **Access** › **Authentication Providers** › **Kerberos Provider**.
2. In the **Kerberos Realms** table, select a realm and click **View** / **Edit**.
3. In the **View a Kerberos Realm** page, click **Edit Realm**.
4. Select or clear the **Set as the default realm** check box to modify the default realm setting.

5. In the **Key Distribution Centers (KDCs)** field, specify the hostname or the IP address of an alternate authentication server.
6. In the **Admin Server** field, type the hostname or the IP address of an alternate administrative server to use as the master KDC.
7. In the **Default Domain** field, specify an alternate domain name for translating the service principal names (SPNs).
8. Click **Save Changes** to return to the **View a Kerberos Realm** page.
9. Click **Close**.

View an MIT Kerberos realm

You can view details related to the name, Key Distribution Centers (KDCs), and administrative server associated with an MIT Kerberos realm.

Procedure

1. Click **Access** › **Authentication Providers** › **Kerberos Provider**.
2. In the **Kerberos Realms** table, select a realm and click **View** / **Edit** to view the information associated with the realm.

Delete an MIT Kerberos realm

You can delete one or more MIT Kerberos realms and all the associated MIT Kerberos domains. Kerberos realms are referenced by Kerberos providers. Hence before you delete a realm for which you have created a provider, you must first delete that provider.

Procedure

1. Click **Access** › **Authentication Providers** › **Kerberos Provider**.
2. In the **Kerberos Realms** table, select one or more realms and then perform one of the following actions:
   - To delete a single realm, select the realm and click **More** › **Delete** from the **Actions** column.
   - To delete multiple realms, select the realms and then select **Delete Selection** from the **Select a bulk action** list.
3. In the confirmation dialog box, click **Delete**.

Managing MIT Kerberos providers

You can create view, delete, or modify an MIT Kerberos provider. You can also configure the Kerberos provider settings.

Creating an MIT Kerberos provider

You can create an MIT Kerberos provider by obtaining the credentials for accessing a cluster through the Key Distribution Center (KDC) of the Kerberos realm. This process is also known as joining a realm. Thus when you create a Kerberos provider you also join a realm that you have previously created. You must be a member of the SecurityAdmin role to create an MIT Kerberos provider.

Using the web interface, you can perform the following tasks through a single workflow or perform each task individually before creating the provider.

- Defining a realm
- Defining a domain
- Managing a service principal name (SPN)

Create an MIT Kerberos realm, domain, and a provider

You can create an MIT Kerberos realm, domain, and a provider through a single workflow instead of configuring each of these objects individually.

Procedure

1. Click **Access** › **Authentication Providers** › **Kerberos Provider**.
2. Click **Get Started**.
3. In the **Create Realm** section, specify the following parameters:
   a. In the **Realm Name** field, type a domain name preferably in uppercase characters. For example, type `CLUSTER-NAME.COMPANY.COM`.
   b. Select the **Set as the default realm** check box to set the realm as the default.
   c. In the **Key Distribution Centers (KDCs)** field, add one or more KDCs by specifying the hostnames or IP addresses of the authentication servers.
   d. In the **Admin Server** field, type the hostname or IP address of the authentication server to use as the master KDC. If you omit this step, the first KDC that you added previously is used as the default admin server.
   e. (Optional) In the **Default Domain** field, specify the domain name to use for translating the service principal names (SPNs).
4. (Optional) In the **Create Domain(s)** section, specify one or more domain names to associate with the realm.
5. In the **Authenticate to Realm** section, specify the following parameters:
   a. In the **User** field, type a user name who has the permission to create the SPNs in the Kerberos realm.
   b. In the **Password** field, type the password for the user.
6. In the **Create Provider** section, select one of the following options for managing the SPNs:
   - Use the recommended SPNs.
   - Type an SPN in the format `service/principal@realm` to manually associate it with the selected realm. You can add more than one SPNs for association, if necessary.
7. Click **Create Provider and Join Realm**.

Create an MIT Kerberos provider and join a realm

You join a realm automatically as you create an MIT Kerberos provider. A realm defines a domain within which the authentication for a specific user or service takes place.

Before you begin

You must be a member of the SecurityAdmin role to view and access the **Create a Kerberos Provider** button and perform the tasks described in this procedure.

Procedure

1. Click **Access** › **Authentication Providers** › **Kerberos Provider**.
2. Click **Create a Kerberos Provider**.
3. In the **Realm Authentication Information** section, specify the following parameters:
   a. In the **User** field, type a user name who has the permission to create service principal names (SPNs) in the Kerberos realm.
   b. In the **Password** field, type the password for the user.

4. In the **Provider Information** section, specify the following parameters:
   a. Select a pre-existing realm from the list.
   b. Select one of the following options for managing the SPNs:
   - Use the recommended SPNs.
   - Type an SPN in the format `service/principal@realm` to manually associate it with the selected realm. You can add more than one SPN for association, if necessary.
5. Click **Create Provider and Join Realm**.

Modify an MIT Kerberos provider

You can modify the realm authentication information and the service principal name (SPN) information for an MIT Kerberos provider.

Before you begin

You must be a member of the SecurityAdmin role to view and access the **View / Edit** button to modify an MIT Kerberos provider.

Procedure

1. Click **Access** › **Authentication Providers** › **Kerberos Provider**.
2. In the **Kerberos Provider** table, select a domain and click **View / Edit**.
3. In the **View a Kerberos Provider** page, click **Edit Provider**.
4. In the **Realm Authentication Information** section, specify the credentials for a user with permissions to create SPNs in the given Kerberos realm.
5. In the **Provider Information** section, select one of the following options for managing the SPNs:
   - Use the recommended SPNs.
   - Type an SPN in the format `service/principal@realm` to manually associate the SPN with the selected realm. You can add more than one SPN for association, if necessary.
6. Click **Save Changes** to return to the **View a Kerberos Provider** page.
7. Click **Close**.

View an MIT Kerberos provider

You can view information related to MIT Kerberos realms and service principal names (SPNs) associated with an MIT Kerberos provider.

Procedure

1. Click **Access** › **Authentication Providers** › **Kerberos Provider**.
2. In the **Kerberos Providers** table, select a provider and click **View / Edit** to view the provider information including the realm, recommended SPNs, and any other SPNs that are discovered.

Delete an MIT Kerberos provider

You can delete an MIT Kerberos provider and remove it from all the referenced access zones. When you delete a provider, you also leave an MIT Kerberos realm.

Before you begin

You must be a member of the SecurityAdmin role to perform the tasks described in this procedure.

Procedure

1. Click **Access** › **Authentication Providers** › **Kerberos Provider**.
2. In the **Kerberos Providers** table, select one or more providers and then perform one of the following actions:
   - To delete a single provider, select the provider and click **More** › **Delete** from the **Actions** column.
   - To delete multiple providers, select the providers and then select **Delete Selection** from the **Select a bulk action** list.
3. In the confirmation dialog box, click **Delete**.

Configure Kerberos provider settings

You can configure the settings of a Kerberos provider to allow the DNS records to locate the Key Distribution Center (KDC), Kerberos realms, and the authentication servers associated with a Kerberos realm. These settings are global to all the users of Kerberos across all the nodes, services, and access zones. Some settings are applicable only to the client-side Kerberos that is relevant when joining a realm or when communicating with an Active Directory KDC. Typically, you do not need to change the settings after the initial configuration.

Procedure

1. Click **Access** › **Authentication Providers** › **Kerberos Settings**.
2. In the **Default Realm** field, specify the realm to use for the service principal name (SPN). The default realm is the first realm that you create.
3. Select a check box to always send pre-authentication. This is a client-side Kerberos configuration setting.

   Selecting this check box enables the Kerberos ticket requests to include *ENC_TIMESTAMP* as the pre-authentication data even if the authentication server did not request it. This is useful when working with Active Directory servers.
4. Select a check box to specify whether to use the DNS server records to locate the KDCs and other servers for a realm, if that information is not listed for the realm.
5. Select a check box to specify whether to use the DNS text records to determine the Kerberos realm of a host.
6. Click **Save Changes**.

Managing MIT Kerberos domains

You can optionally define MIT Kerberos domains to allow additional domain extensions to be associated with an MIT Kerberos realm. You can always specify a default domain for a realm.

You can create, modify, delete, and view an MIT Kerberos domain. A Kerberos domain name is a DNS suffix that you specify typically using lowercase characters.

Create an MIT Kerberos domain

You optionally create an MIT Kerberos domain to allow additional domain extensions to be associated with an MIT Kerberos realm apart from the default domains.

Before you begin

You must be a member of the SecurityAdmin role to perform the tasks described in this procedure.

Procedure

1. Click **Access** › **Authentication Providers** › **Kerberos Provider**.
2. Click **Create a Kerberos Domain**.
3. In the **Domain** field, specify a domain name which is typically a DNS suffix in lowercase characters.
4. From the **Realm** list, select a realm that you have configured previously.
5. Click **Create Domain**.

Modify an MIT Kerberos domain

You can modify an MIT Kerberos domain by modifying the realm settings.

Before you begin

You must be a member of the SecurityAdmin role to perform the tasks described in this procedure.

Procedure

1. Click **Access** › **Authentication Providers** › **Kerberos Provider**.
2. In the **Kerberos Domains** table, select a domain and click **View** / **Edit**.
3. In the **View a Kerberos Domain** page, click **Edit Domain**.
4. From the **Realm** list, select an alternate realm.
5. Click **Save Changes** to return to the **View a Kerberos Domain** page.
6. Click **Close**.

View an MIT Kerberos domain

You can view the properties of an MIT Kerberos domain mapping.

Procedure

1. Click **Access** › **Authentication Providers** › **Kerberos Provider**.
2. In the **Kerberos Domains** table, select a domain and click **View** / **Edit** to view the properties of the domain mapping.

Delete an MIT Kerberos domain

You can delete one or more MIT Kerberos domain mappings.

Before you begin

You must be a member of the SecurityAdmin role to perform the tasks described in this procedure.

Procedure

1. Click **Access** › **Authentication Providers** › **Kerberos Provider**.
2. In the **Kerberos Domains** table, select one or more domain mappings and then perform one of the following actions:
   - To delete a single domain mapping, select the mapping and click **More** › **Delete** from the **Actions** column.
   - To delete multiple domain mappings, select the mappings and then select **Delete Selection** from the **Select a bulk action** list.

Managing access permissions

The internal representation of identities and permissions can contain information from UNIX sources, Windows sources, or both. Because access protocols can process the information from only one of these sources, the system may need to make approximations to present the information in a format the protocol can process.

View expected user permissions

You can view the expected permissions for user access to a file or directory.

This procedure must be performed through the command-line interface (CLI).

Procedure

1. Establish an SSH connection to any node in the cluster.
2. View expected user permissions by running the `isi auth access` command.

   The following command displays permissions in `/ifs/` for the user that you specify in place of *<username>*:

   ```
   isi auth access <username> /ifs/
   ```

   The system displays output similar to the following example:

   ```
           User
              Name : <username>
               UID : 2018
               SID :
   SID:S-1-5-21-2141457107-1514332578-1691322784-1018
          File
             Owner : user:root
             Group : group:wheel
              Mode : drwxrwxrwx
     Relevant Mode : d---rwx---
   Permissions
          Expected : user:<username> \
           allow
   dir_gen_read,dir_gen_write,dir_gen_execute,delete_child
   ```

3. View mode-bits permissions for a user by running the `isi auth access` command.

   The following command displays verbose-mode file permissions information in `/ifs/` for the user that you specify in place of *<username>*:

   ```
   isi auth access <username> /ifs/ -v
   ```

   The system displays output similar to the following example:

   ```
   User Name : <username> UID \
   : 2018 SID : SID:S-1-5-21-2141457107-1514332578-1691322784-1018
   File Owner : user:root Group : group:wheel Mode : drwxrwxrwx
   Relevant Mode : d---rwx--- Permissions Expected : user:<username>
   allow dir_gen_read,dir_gen_write,dir_gen_execute,delete_child
   ```

4. View expected ACL user permissions on a file for a user by running the `isi auth access` command.

The following command displays verbose-mode ACL file permissions for the file `file_with_acl.tx` in `/ifs/data/` for the user that you specify in place of *<username>*:

```
isi auth access <username> /ifs/data/file_with_acl.tx -v
```

The system displays output similar to the following example:

```
User Name : <username> \
UID : 2097 SID :
SID:S-1-7-21-2141457107-1614332578-1691322789-1018
File Owner : user:<username> Group : group:wheel
Permissions Expected : user:<username>
allow file_gen_read,file_gen_write,std_write_dac
Relevant Acl: group:<group-name> Users allow file_gen_read
user:<username> allow std_write_dac,file_write,
append,file_write_ext_attr,file_write_attr
group:wheel allow file_gen_read,file_gen_write
```

Configure access management settings

Default access settings include whether to send NTLMv2 responses for SMB connections, the identity type to store on disk, the Windows workgroup name for running in local mode, and character substitution for spaces encountered in user and group names.

Procedure

1. Click **Access** › **Settings**.
2. Configure the following settings as needed.

Option	Description
Send NTLMv2	Specifies whether to send only NTLMv2 responses to SMB clients with NTLM-compatible credentials.
On-Disk Identity	Controls the preferred identity to store on disk. If OneFS is unable to convert an identity to the preferred format, it is stored as is. This setting does not affect identities that are currently stored on disk. Select one of the following settings: native Allow OneFS to determine the identity to store on disk. This is the recommended setting. unix Always store incoming UNIX identifiers (UIDs and GIDs) on disk. sid Store incoming Windows security identifiers (SIDs) on disk, unless the SID was generated from a UNIX identifier; in that case, convert it back to the UNIX identifier and store it on disk.
Workgroup	Specifies the NetBIOS workgroup. The default value is `WORKGROUP`.
Space Replacement	For clients that have difficulty parsing spaces in user and group names, specifies a substitute character.

3. Click **Save**.

After you finish

If you changed the on-disk identity selection, it is recommended that you run the PermissionRepair job with the `Convert` repair type to prevent potential permissions errors. For more information, see Update cluster permissions on page 135.

Modify ACL policy settings

You can modify ACL policy settings but the default ACL policy settings are sufficient for most cluster deployments.

⚠ CAUTION

Because ACL policies change the behavior of permissions throughout the system, they should be modified only as necessary by experienced administrators with advanced knowledge of Windows ACLs. This is especially true for the advanced settings, which are applied regardless of the cluster's environment.

For UNIX, Windows, or balanced environments, the optimal permission policy settings are selected and cannot be modified. However, you can choose to manually configure the cluster's default permission settings if necessary to support your particular environment.

Note

You must be logged in to the web administration interface to perform this task.

Procedure

1. Click **Protocols** › **ACLs** › **ACL Policies**.
2. In the **Standard Settings** area, under **Environment**, select the option that best describes your environment, or select **Configure permission policies manually** to configure individual permission policies.
3. If you selected the **Configure permission policies manually** option, configure the settings as needed.

   For more information about these settings, see ACL policy settings options.
4. In the **Advanced Settings** area, configure the settings as needed.

ACL policy settings options

You can configure an ACL policy by choosing from the available settings options.

Environment settings

UNIX only
: Causes cluster permissions to operate with UNIX semantics, as opposed to Windows semantics. Enabling this option prevents ACL creation on the system.

Balanced
: Causes cluster permissions to operate in a mixed UNIX and Windows environment. This setting is recommended for most cluster deployments.

Windows only
: Causes cluster permissions to operate with Windows semantics, as opposed to UNIX semantics. Enabling this option causes the system to return an error on UNIX chmod requests.

Configure permission policies manually

Allows you to configure the individual permissions policy settings available under **Permission Policies**.

Permission policies settings

ACL creation over SMB

Specifies whether to allow or deny creation of ACLs over SMB. Select one of the following options.

Do not allow the creation of ACLs over Windows File Sharing (SMB)

Prevents ACL creation on the cluster.

Allow the creation of ACLs over SMB

Allows ACL creation on the cluster.

Note

Inheritable ACLs on the system take precedence over this setting: If inheritable ACLs are set on a folder, any new files and folders created in that folder will inherit the folder's ACL. Disabling this setting does not remove ACLs currently set on files. If you want to clear an existing ACL, run the `chmod -b` *<mode> <file>* command to remove the ACL and set the correct permissions.

chmod on files with existing ACLs

Controls what happens when a `chmod` operation is initiated on a file with an ACL, either locally or over NFS. This setting controls any elements that set UNIX permissions, including File System Explorer. Enabling this policy setting does not change how `chmod` operations affect files that do not have ACLs. Select one of the following options.

Remove the existing ACL and set UNIX permissions instead

For `chmod` operations, removes any existing ACL and instead sets the `chmod` permissions. Select this option only if you do not need permissions to be set from Windows.

Remove the existing ACL and create an ACL equivalent to the UNIX permissions

Stores the UNIX permissions in a Windows ACL. Select this option only if you want to remove Windows permissions but do not want files to have synthetic ACLs.

Remove the existing ACL and create an ACL equivalent to the UNIX permissions, for all users/groups referenced in old ACL

Stores the UNIX permissions in a Windows ACL. Select this option only if you want to remove Windows permissions but do not want files to have synthetic ACLs.

Merge the new permissions with the existing ACL

Causes Windows and UNIX permissions to operate smoothly in a balanced environment by merging permissions that are applied by `chmod` with existing ACLs. An ACE for each identity (owner, group, and everyone) is either modified or created, but all other ACEs are unmodified. Inheritable ACEs are also left unmodified to enable Windows users to continue to inherit appropriate permissions. UNIX users can set specific permissions for each of those three standard identities, however.

Deny permission to modify the ACL

Prevents users from making NFS and local `chmod` operations. Enable this setting if you do not want to allow permission sets over NFS.

Ignore operation if file has an existing ACL

Prevents an NFS client from making changes to the ACL. Select this option if you defined an inheritable ACL on a directory and want to use that ACL for permissions.

⚠ CAUTION

If you try to run the `chmod` command on the same permissions that are currently set on a file with an ACL, you may cause the operation to silently fail—The operation appears to be successful, but if you were to examine the permissions on the cluster, you would notice that the `chmod` command had no effect. As a workaround, you can run the `chmod` command away from the current permissions and then perform a second `chmod` command to revert to the original permissions. For example, if your file shows 755 UNIX permissions and you want to confirm this number, you could run `chmod 700 file; chmod 755 file`.

ACLs created on directories by UNIX chmod

On Windows systems, the access control entries for directories can define fine-grained rules for inheritance; on UNIX, the mode bits are not inherited. Making ACLs that are created on directories by the chmod command inheritable is more secure for tightly controlled environments but may deny access to some Windows users who would otherwise expect access.
Select one of the following options.

- **Make them inheritable**
- **Do not make them inheritable**

chown/chgrp on files with existing ACLs

Changes a file or folder's owning user or group. Select one of the following options.

Modify the owner and/or group

Causes the chown or chgrp operation to perform as it does in UNIX. Enabling this setting modifies any ACEs in the ACL associated with the old and new owner or group.

Modify the owner and/or group and ACL permissions

Cause the NFS chown or chgrp operation to function as it does in Windows. When a file owner is changed over Windows, no permissions in the ACL are changed.

Ignore operation if file has an existing ACL

Prevents an NFS client from making changes to the owner or group.

Note

Over NFS, the chown or chgrp operation changes the permissions and the owner or owning group. For example, consider a file owned by user Joe with rwx------ (700) permissions, signifying rwx permissions for the owner, but no permissions for anyone else. If you run the chown command to change ownership of the file to user Bob, the owner permissions are still rwx but they now represent the permissions for Bob, rather than for Joe, who lost all of his permissions. This setting does not affect UNIX chown or chgrp operations performed on files with UNIX permissions, and it does not affect Windows chown or chgrp operations, which do not change any permissions.

Access checks (chmod, chown)

In UNIX environments, only the file owner or superuser has the right to run a `chmod` or `chown` operation on a file. In Windows environments, you can implement this policy setting to give users the right to perform `chmod` operations, called the change permissions right, or the right to perform `chown` operations, called the take ownership right.

Note

The take ownership right only gives users the ability to take file ownership, not to give ownership away.

Select one of the following options.

Allow only owners to chmod or chown

Causes `chmod` and `chown` access checks to operate with UNIX-like behavior.

Allow owner and users with 'take ownership' right to chown, and owner and users with 'change permissions' right to chmod

Causes `chmod` and `chown` access checks to operate with Windows-like behavior.

Advanced settings

Treatment of 'rwx' permissions

In UNIX environments, rwx permissions signify two things: A user or group has read, write, and execute permissions; and a user or group has the maximum possible level of permissions.
When you assign UNIX permissions to a file, no ACLs are stored for that file. A Windows system processes only ACLs; Windows does not process UNIX permissions. Therefore, when you view a file's permissions on a Windows system, the cluster must translate the UNIX permissions into an ACL. This type of ACL is called a synthetic ACL. Synthetic ACLs are not stored anywhere; instead, they are dynamically generated as needed and then they are discarded. If a file has UNIX permissions, you may notice synthetic ACLs when you run the `ls` file command on the cluster to view a file’s ACLs.

When you generate a synthetic ACL, the cluster maps UNIX permissions to Windows rights. Windows supports a more granular permissions model than UNIX does, and it specifies rights that cannot easily be mapped from UNIX permissions. If the cluster maps rwx permissions to Windows rights, you must enable one of the following options. The main difference between rwx and Full Control is the broader set of permissions with Full Control.

Select one of the following options.

Retain 'rwx' permissions

Generates an ACE that provides only read, write, and execute permissions.

Treat 'rwx' permissions as Full Control

Generates an ACE that provides the maximum Windows permissions for a user or a group by adding the change permissions right, the take ownership right, and the delete right.

Group owner inheritance

Operating systems tend to work with group ownership and permissions in two different ways: BSD inherits the group owner from the file's parent folder; Windows and Linux inherit the group owner from the file creator's primary group. If you enable a setting that causes the group owner to be inherited from the creator's primary group, you can override it on a per-folder basis by running the `chmod` command to set the set-gid bit. This inheritance applies only when the file is created. For more information, see the manual page for the `chmod` command.

Select one of the following options.

When an ACL exists, use Linux and Windows semantics, otherwise use BSD semantics

Controls file behavior based on whether the new file inherits ACLs from its parent folder. If it does, the file uses the creator's primary group. If it does not, the file inherits from its parent folder.

BSD semantics - Inherit group owner from the parent folder

Causes the group owner to be inherited from the file's parent folder.

Linux and Windows semantics - Inherit group owner from the creator's primary group

Causes the group owner to be inherited from the file creator's primary group.

chmod (007) on files with existing ACLs

Specifies whether to remove ACLs when running the `chmod (007)` command.

Select one of the following options.

chmod(007) does not remove existing ACL

Sets 007 UNIX permissions without removing an existing ACL.

chmod(007) removes existing ACL and sets 007 UNIX permissions

Removes ACLs from files over UNIX file sharing (NFS) and locally on the cluster through the `chmod (007)` command. If you enable this setting, be sure to run the `chmod` command on the file immediately after using `chmod (007)` to clear an ACL. In most cases, you do not want to leave 007 permissions on the file.

Owner permissions

It is impossible to represent the breadth of a Windows ACL's access rules using a set of UNIX permissions. Therefore, when a UNIX client requests UNIX permissions for a file with an ACL over NFS, an action known as a stat, it receives an imperfect approximation of the file's true permissions. By default, executing an `ls -l` command from a UNIX client returns a more open set of permissions than the user expects. This permissiveness compensates for applications that incorrectly inspect the UNIX permissions themselves when determining whether to attempt a file-system operation. The purpose of this policy setting is to ensure that these applications proceed with the operation to allow the file system to properly determine user access through the ACL.

Select one of the following options.

Approximate owner mode bits using all possible group ACEs

Makes the owner permissions appear more permissive than the actual permissions on the file.

Approximate owner mode bits using only the ACE with the owner ID

Makes the owner permissions appear more accurate, in that you see only the permissions for a particular owner and not the more permissive set. This may cause access-denied problems for UNIX clients, however.

Group permissions

Select one of the following options for group permissions:

Approximate group mode bits using all possible group ACEs

Makes the group permissions appear more permissive than the actual permissions on the file.

Approximate group mode bits using only the ACE with the group ID

Makes the group permissions appear more accurate, in that you see only the permissions for a particular group and not the more permissive set. This may cause access-denied problems for UNIX clients, however.

No "deny" ACEs

The Windows ACL user interface cannot display an ACL if any deny ACEs are out of canonical ACL order. To correctly represent UNIX permissions, deny ACEs may be required to be out of canonical ACL order.

Select one of the following options.

Do not modify synthetic ACLs and mode bit approximations

Specifies to not modify synthetic ACL generation; “deny” ACEs will be generated when necessary.

⚠ CAUTION

This option can lead to permissions being reordered, permanently denying access if a Windows user or an application performs an ACL get, an ACL modification, and an ACL set (known as a round trip) to and from Windows.

Remove “deny” ACEs from ACLs. This setting can cause ACLs to be more permissive than the equivalent mode bits

Does not include deny ACEs when generating synthetic ACLs. This setting can cause ACLs to be more permissive than the equivalent mode bits.

Access check (utimes)

You can control who can change utimes, which are the access and modification times of a file, by selecting one of the following options.

Allow only owners to change utimes to client-specific times (POSIX compliant)

Allows only owners to change utimes, which complies with the POSIX standard, an approach that is probably familiar to administrators of UNIX systems.

Allow owners and users with ‘write’ access to change utimes to client-specific times

Allows owners as well as users with write access to modify utimes—a less restrictive approach that is probably familiar to administrators of Windows systems.

Read-only DOS attribute

Deny permission to modify files with DOS read-only attribute over Windows Files Sharing (SMB)

Duplicates DOS-attribute permissions behavior over only the SMB protocol, so that they use the read-only attribute over SMB.

Deny permission to modify files with DOS read-only attribute over both UNIX (NFS) and Windows File Sharing (SMB)

Duplicates DOS-attribute permissions behavior over both NFS and SMB protocols. For example, if permissions are read-only on a file over SMB, permissions are read-only over NFS.

Displayed mode bits

Use ACL to approximate mode bits
Presents the OneFS approximation of the NFS mode bits, based on ACL permissions in the security descriptor.

Always display 777 if ACL exists-
If the approximated NFS permissions are less permissive than those in the security descriptor, you may want to use this setting so the NFS client does not stop with the access check before performing its operation. Use this setting when a third-party application may be blocked if the ACL does not provide the proper access.

Update cluster permissions

You can update file permissions or ownership by running the Repair Permissions job. To prevent permissions issues that can occur after changing the on-disk identity, run this job with the Convert Permissions job to ensure that the changes are fully propagated throughout the cluster.

Procedure

1. Click **Protocols** › **ACLs** › **Repair Permissions Job**.
2. (Optional) From the **Priority** list, select the priority level at which to run the job in relation to other jobs.
3. (Optional) From the **Impact policy** list, select an impact policy for the job to follow.
4. In the **Paths** field, type or browse to the directory in `/ifs` whose permissions you want to repair.
5. (Optional) Click **Add another directory path** and in the added **Paths** field, type or browse for an additional directory in `/ifs` whose permissions you want to repair.

   You can repeat this step to add directory paths as needed.
6. From the **Repair task** list, select one of the following methods for updating permissions:

Option	Description
Clone permissions	Applies the permissions settings for the directory specified by the **Template Directory** setting to the **Path to repair** directory.
Inherit permissions	Recursively applies the ACL of the directory that is specified by the **Template Directory** setting to each file and subdirectory in the specified **Path to repair** directory, according to standard inheritance rules.
Convert permissions	For each file and directory in the specified **Path to repair** directory, converts the owner, group, and access control list (ACL) to the target on-disk identity.

   The remaining settings options differ depending on the selected repair task.
7. In the **Template File or Directory** field, type or browse to the directory in `/ifs` that you want to copy permissions from. This setting applies to only the Clone and Inherit repair types.
8. (Optional) From the **Mapping type** list, select the preferred on-disk identity type to apply. This setting applies to only the Convert permissions repair task.

Option	Description
Global	Applies the system's default identity.
SID (Windows)	Applies the Windows identity.
UNIX	Applies the UNIX identity.
Native	If a user or group does not have an authoritative UNIX identifier (UID or GID), applies the Windows identity (SID)

9. (Optional) From the **Access Zone** list, select an access zone to use for ID mapping. This setting applies to only the Convert permissions repair task.

CHAPTER 6

Identity management

This section contains the following topics:

Identity management overview

In environments with several different types of directory services, OneFS maps the users and groups from the separate services to provide a single unified identity on an EMC Isilon cluster and uniform access control to files and directories, regardless of the incoming protocol. This process is called identity mapping.

Isilon clusters are frequently deployed in multiprotocol environments with multiple types of directory services, such as Active Directory and LDAP. When a user with accounts in multiple directory services logs in to a cluster, OneFS combines the user's identities and privileges from all the directory services into a native access token.

You can configure OneFS settings to include a list of rules for access token manipulation to control user identity and privileges. For example, you can set a user mapping rule to merge an Active Directory identity and an LDAP identity into a single token that works for access to files stored over both SMB and NFS. The token can include groups from Active Directory and LDAP. The mapping rules that you create can solve identity problems by manipulating access tokens in many ways, including the following examples:

- Authenticate a user with Active Directory but give the user a UNIX identity.
- Select a primary group from competing choices in Active Directory or LDAP.
- Disallow login of users that do not exist in both Active Directory and LDAP.

For more information about identity management, see the white paper *Managing identities with the Isilon OneFS user mapping service* at EMC Online Support.

Identity types

OneFS supports three primary identity types, each of which you can store directly on the file system. Identity types are user identifier and group identifier for UNIX, and security identifier for Windows.

When you log on to an EMC Isilon cluster, the user mapper expands your identity to include your other identities from all the directory services, including Active Directory, LDAP, and NIS. After OneFS maps your identities across the directory services, it generates an access token that includes the identity information associated with your accounts. A token includes the following identifiers:

- A UNIX user identifier (UID) and a group identifier (GID). A UID or GID is a 32-bit number with a maximum value of 4,294,967,295.
- A security identifier (SID) for a Windows user account. A SID is a series of authorities and sub-authorities ending with a 32-bit relative identifier (RID). Most SIDs have the form S-1-5-21-*<A>*-*<B>*-*<C>*-*<RID>*, where *<A>*, *<B>*, and *<C>* are specific to a domain or computer and *<RID>* denotes the object in the domain.
- A primary group SID for a Windows group account.
- A list of supplemental identities, including all groups in which the user is a member.

The token also contains privileges that stem from administrative role-based access control.

On an Isilon cluster, a file contains permissions, which appear as an access control list (ACL). The ACL controls access to directories, files, and other securable system objects.

When a user tries to access a file, OneFS compares the identities in the user's access token with the file's ACL. OneFS grants access when the file's ACL includes an access control entry (ACE) that allows the identity in the token to access the file and that does not include an ACE that denies the identity access. OneFS compares the access token of a user with the ACL of a file.

Note

For more information about access control lists, including a description of the permissions and how they correspond to POSIX mode bits, see the white paper titled *EMC Isilon multiprotocol data access with a unified security model* on the EMC Online Support web site.

When a name is provided as an identifier, it is converted into the corresponding user or group object and the correct identity type. You can enter or display a name in various ways:

- UNIX assumes unique case-sensitive namespaces for users and groups. For example, Name and name represent different objects.
- Windows provides a single, case-insensitive namespace for all objects and also specifies a prefix to target an Active Directory domain; for example, domain\name.
- Kerberos and NFSv4 define principals, which require names to be formatted the same way as email addresses; for example, name@domain.com.

Multiple names can reference the same object. For example, given the name support and the domain example.com, support, EXAMPLE\support and support@example.com are all names for a single object in Active Directory.

Access tokens

An access token is created when the user first makes a request for access.

Access tokens represent who a user is when performing actions on the cluster and supply the primary owner and group identities during file creation. Access tokens are also compared against the ACL or mode bits during authorization checks.

During user authorization, OneFS compares the access token, which is generated during the initial connection, with the authorization data on the file. All user and identity mapping occurs during token generation; no mapping takes place during permissions evaluation.

An access token includes all UIDs, GIDs, and SIDs for an identity, in addition to all OneFS privileges. OneFS reads the information in the token to determine whether a user has access to a resource. It is important that the token contains the correct list of UIDs, GIDs, and SIDs. An access token is created from one of the following sources:

Source	Authentication
Username	• SMB impersonate user • Kerberized NFSv3 • Kerberized NFSv4 • NFS export user mapping • HTTP • FTP • HDFS
Privilege Attribute Certificate (PAC)	• SMB NTLM

Source	Authentication
	- Active Directory Kerberos
User identifier (UID)	- NFS AUTH_SYS mapping

Access token generation

For most protocols, the access token is generated from the username or from the authorization data that is retrieved during authentication.

The following steps present a simplified overview of the complex process through which an access token is generated:

Step	Process	Description
1	User identity lookup	Using the initial identity, the user is looked up in all configured authentication providers in the access zone, in the order in which they are listed, until a match is found. The user identity and group list are retrieved from the authenticating provider. Any SIDs, UIDs, or GIDs are added to the initial token. **Note** An exception to this behavior occurs if the AD provider is configured to call other providers, such as LDAP or NIS.
2	ID mapping	The user's identifiers are associated across directory services. All SIDs are converted to their equivalent UID/GID and vice versa. These ID mappings are also added to the access token.
3	User mapping	Access tokens from other directory services are combined. If the username matches any user mapping rules, the rules are processed in order and the token is updated accordingly.
4	On-disk identity calculation	The default on-disk identity is calculated from the final token and the global setting. These identities are used for newly created files.

ID mapping

The Identity (ID) mapping service maintains relationship information between mapped Windows and UNIX identifiers to provide consistent access control across file sharing protocols within an access zone.

Note

ID mapping and user mapping are different services, despite the similarity in names.

During authentication, the authentication daemon requests identity mappings from the ID mapping service in order to create access tokens. Upon request, the ID mapping service returns Windows identifiers mapped to UNIX identifiers or UNIX identifiers mapped to Windows identifiers. When a user authenticates to a cluster over NFS with a UID or GID, the ID mapping service returns the mapped Windows SID, allowing access to files that another user stored over SMB. When a user authenticates to the cluster over

SMB with a SID, the ID mapping service returns the mapped UNIX UID and GID, allowing access to files that a UNIX client stored over NFS.

Mappings between UIDs or GIDs and SIDs are stored according to access zone in a cluster-distributed database called the ID map. Each mapping in the ID map is stored as a one-way relationship from the source to the target identity type. Two-way mappings are stored as complementary one-way mappings.

Mapping Windows IDs to UNIX IDs

When a Windows user authenticates with an SID, the authentication daemon searches the external Active Directory provider to look up the user or group associated with the SID. If the user or group has only an SID in the Active Directory, the authentication daemon requests a mapping from the ID mapping service.

Note

User and group lookups may be disabled or limited, depending on the Active Directory settings. You enable user and group lookup settings through the `isi auth ads modify` command.

If the ID mapping service does not locate and return a mapped UID or GID in the ID map, the authentication daemon searches other external authentication providers configured in the same access zone for a user that matches the same name as the Active Directory user.

If a matching user name is found in another external provider, the authentication daemon adds the matching user's UID or GID to the access token for the Active Directory user, and the ID mapping service creates a mapping between the UID or GID and the Active Directory user's SID in the ID map. This is referred to as an *external mapping*.

Note

When an external mapping is stored in the ID map, the UID is specified as the on-disk identity for that user. When the ID mapping service stores a generated mapping, the SID is specified as the on-disk identity.

If a matching user name is not found in another external provider, the authentication daemon assigns a UID or GID from the ID mapping range to the Active Directory user's SID, and the ID mapping service stores the mapping in the ID map. This is referred to as a *generated mapping*. The ID mapping range is a pool of UIDs and GIDs allocated in the mapping settings.

After a mapping has been created for a user, the authentication daemon retrieves the UID or GID stored in the ID map upon subsequent lookups for the user.

Mapping UNIX IDs to Windows IDs

The ID mapping service creates temporary UID-to-SID and GID-to-SID mappings only if a mapping does not already exist. The UNIX SIDs that result from these mappings are never stored on disk.

UIDs and GIDs have a set of predefined mappings to and from SIDs.

If a UID-to-SID or GID-to-SID mapping is requested during authentication, the ID mapping service generates a temporary UNIX SID in the format S-1-22-1-*<UID>* or S-1-22-2-*<GID>* by applying the following rules:

- For UIDs, the ID mapping service generates a UNIX SID with a domain of S-1-22-1 and a resource ID (RID) matching the UID. For example, the UNIX SID for UID 600 is S-1-22-1-600.

- For GIDs, the ID mapping service generates a UNIX SID with a domain of S-1-22-2 and an RID matching the GID. For example, the UNIX SID for GID 800 is S-1-22-2-800.

ID mapping ranges

In access zones with multiple external authentication providers, such as Active Directory and LDAP, it is important that the UIDs and GIDs from different providers that are configured in the same access zone do not overlap. Overlapping UIDs and GIDs between providers within an access zone might result in some users gaining access to other users' directories and files.

The range of UIDs and GIDs that can be allocated for generated mappings is configurable in each access zone through the `isi auth settings mappings modify` command. The default range for both UIDs and GIDs is 1000000–2000000 in each access zone.

Do not include commonly used UIDs and GIDs in your ID ranges. For example, UIDs and GIDs below 1000 are reserved for system accounts and should not be assigned to users or groups.

User mapping

User mapping provides a way to control permissions by specifying a user's security identifiers, user identifiers, and group identifiers. OneFS uses the identifiers to check file or group ownership.

With the user-mapping feature, you can apply rules to modify which user identity OneFS uses, add supplemental user identities, and modify a user's group membership. The user-mapping service combines a user’s identities from different directory services into a single access token and then modifies it according to the rules that you create.

Note

You can configure mapping rules on a per-zone basis. Mapping rules must be configured separately in each access zone that uses them. OneFS maps users only during login or protocol access.

Default user mappings

Default user mappings determine access if explicit user-mapping rules are not created.

If you do not configure rules, a user who authenticates with one directory service receives the identity information in other directory services when the account names are the same. For example, a user who authenticates with an Active Directory domain as Desktop\jane automatically receives identities in the final access token for the corresponding UNIX user account for jane from LDAP or NIS.

In the most common scenario, OneFS is connected to two directory services, Active Directory and LDAP. In such a case, the default mapping provides a user with the following identity attributes:

- A UID from LDAP
- The user SID from Active Directory
- An SID from the default group in Active Directory

The user's groups come from Active Directory and LDAP, with the LDAP groups and the autogenerated group GID added to the list. To pull groups from LDAP, the mapping service queries the memberUid attribute. The user’s home directory, gecos, and shell come from Active Directory.

Elements of user-mapping rules

You combine operators with user names to create a user-mapping rule.

The following elements affect how the user mapper applies a rule:

- The operator, which determines the operation that a rule performs
- Fields for usernames
- Options
- A parameter
- Wildcards

User-mapping best practices

You can follow best practices to simplify user mapping.

Use Active Directory with RFC 2307 and Windows Services for UNIX

Use Microsoft Active Directory with Windows Services for UNIX and RFC 2307 attributes to manage Linux, UNIX, and Windows systems. Integrating UNIX and Linux systems with Active Directory centralizes identity management and eases interoperability, reducing the need for user-mapping rules. Make sure your domain controllers are running Windows Server 2003 or later.

Employ a consistent username strategy

The simplest configurations name users consistently, so that each UNIX user corresponds to a similarly named Windows user. Such a convention allows rules with wildcard characters to match names and map them without explicitly specifying each pair of accounts.

Do not use overlapping ID ranges

In networks with multiple identity sources, such as LDAP and Active Directory with RFC 2307 attributes, you should ensure that UID and GID ranges do not overlap. It is also important that the range from which OneFS automatically allocates UIDs and GIDs does not overlap with any other ID range. OneFS automatically allocates UIDs and GIDs from the range 1,000,000-2,000,000. If UIDs and GIDs overlap multiple directory services, some users might gain access to other users' directories and files.

Avoid common UIDs and GIDs

Do not include commonly used UIDs and GIDs in your ID ranges. For example, UIDs and GIDs below 1000 are reserved for system accounts; do not assign them to users or groups.

Do not use UPNs in mapping rules

You cannot use a user principal name (UPN) in a user mapping rule. A UPN is an Active Directory domain and username that are combined into an Internet-style name with an @ symbol, such as an email address: jane@example. If you include a UPN in a rule, the mapping service ignores it and may return an error. Instead, specify names in the format DOMAIN\user.com.

Group rules by type and order them

The system processes every mapping rule by default, which can present problems when you apply a deny-all rule—for example, to deny access to all unknown users. In addition, replacement rules might interact with rules that contain wildcard characters. To minimize complexity, it is recommended that you group rules by type and organize them in the following order:

1. Replacement rules: Specify all rules that replace an identity first to ensure that OneFS replaces all instances of the identity.
2. Join, add, and insert rules: After the names are set by any replacement operations, specify join, add, and insert rules to add extra identifiers.
3. Allow and deny rules: Specify rules that allow or deny access last.

Note

Stop all processing before applying a default deny rule. To do so, create a rule that matches allowed users but does nothing, such as an add operator with no field options, and has the break option. After enumerating the allowed users, you can place a catchall deny at the end to replace anybody unmatched with an empty user.

To prevent explicit rules from being skipped, in each group of rules, order explicit rules before rules that contain wildcard characters.

Add the LDAP or NIS primary group to the supplemental groups

When an Isilon cluster is connected to Active Directory and LDAP, a best practice is to add the LDAP primary group to the list of supplemental groups. This lets OneFS honor group permissions on files created over NFS or migrated from other UNIX storage systems. The same practice is advised when an Isilon cluster is connected to both Active Directory and NIS.

On-disk identity

After the user mapper resolves a user's identities, OneFS determines an authoritative identifier for it, which is the preferred on-disk identity.

OnesFS stores either UNIX or Windows identities in file metadata on disk. On-disk identity types are UNIX, SID, and native. Identities are set when a file is created or a file's access control data is modified. Almost all protocols require some level of mapping to operate correctly, so choosing the preferred identity to store on disk is important. You can configure OneFS to store either the UNIX or the Windows identity, or you can allow OneFS to determine the optimal identity to store.

On-disk identity types are UNIX, SID, and native. Although you can change the type of on-disk identity, the native identity is best for a network with UNIX and Windows systems. In native on-disk identity mode, setting the UID as the on-disk identity improves NFS performance.

Note

The SID on-disk identity is for a homogeneous network of Windows systems managed only with Active Directory. When you upgrade from a version earlier than OneFS 6.5, the on-disk identity is set to UNIX. When you upgrade from OneFS 6.5 or later, the on-disk identity setting is preserved. On new installations, the on-disk identity is set to native.

The native on-disk identity type allows the OneFS authentication daemon to select the correct identity to store on disk by checking for the identity mapping types in the following order:

Order	Mapping type	Description
1	Algorithmic mapping	An SID that matches S-1-22-1-UID or S-1-22-2-GID in the internal ID mapping database is converted back to the corresponding UNIX identity, and the UID and GID are set as the on-disk identity.
2	External mapping	A user with an explicit UID and GID defined in a directory service (such as Active Directory with RFC 2307 attributes, LDAP, NIS, or the OneFS file provider or local provider) has the UNIX identity set as the on-disk identity.
3	Persistent mapping	Mappings are stored persistently in the identity mapper database. An identity with a persistent mapping in the identity mapper database uses the destination of that mapping as the on-disk identity, which occurs primarily with manual ID mappings. For example, if there is an ID mapping of GID:10000 to S-1-5-32-545, a request for the on-disk storage of GID:10000 returns S-1-5-32-545.
4	No mapping	If a user lacks a UID or GID even after querying the other directory services and identity databases, its SID is set as the on-disk identity. In addition, to make sure a user can access files over NFS, OneFS allocates a UID and GID from a preset range of 1,000,000 to 2,000,000. In native on-disk identity mode, a UID or GID that OneFS generates is never set as the on-disk identity.

Note

If you change the on-disk identity type, you should run the PermissionRepair job in convert mode to make sure that the disk representation of all files is consistent with the changed setting.

Managing ID mappings

You can create, modify, and delete identity mappings and configure ID mapping settings.

Create an identity mapping

You can create a manual identity mapping between source and target identities or automatically generate a mapping for a source identity.

This procedure is available only through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Run the `isi auth mapping create` command.

   The following command specifies IDs of source and target identities in the zone3 access zone to create a two-way mapping between the identities:

   ```
   isi auth mapping create --2way --source-sid=S-1-5-21-12345 \
   --target-uid=5211 --zone=zone3
   ```

Modify an identity mapping

You can modify the configuration of an identity mapping.

This procedure is available only through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Run the `isi auth mapping modify` command.

   The following command modifies the mapping of the user with UID 4236 in the zone3 access zone to include a reverse, 2-way mapping between the source and target identities:

   ```
   isi auth mapping modify --source-uid=4236 \
   --target-sid=S-1-5-21-12345 --zone=zone3 --2way
   ```

Delete an identity mapping

You can delete one or more identity mappings.

This procedure is available only through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Run the `isi auth mapping delete` command.

   The following command deletes all identity mappings in the zone3 access zone:

   ```
   isi auth mapping delete --all --zone=zone3
   ```

   The following command deletes all identity mappings in the zone3 access zone that were both created automatically and include a UID or GID from an external authentication source:

   ```
   isi auth mapping delete --all --only-external --zone=zone3
   ```

   The following command deletes the identity mapping of the user with UID 4236 in the zone3 access zone:

   ```
   isi auth mapping delete --source-uid=4236 --zone=zone3
   ```

View an identity mapping

You can display mapping information for a specific identity.

This procedure is available only through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Run the `isi auth mapping view` command.

   The following command displays mappings for the user with UID 4236 in the zone3 access zone:

   ```
   isi auth mapping view --uid=4236 --zone=zone3
   ```

The system displays output similar to the following example:

```
Name: user_36
 On-disk: UID: 4236
Unix uid: 4236
Unix gid: -100000
     SMB: S-1-22-1-4236
```

Flush the identity mapping cache

You can flush the ID map cache to remove in-memory copies of all or specific identity mappings.

Modifications to ID mappings may cause the cache to become out-of-sync and users might experience slowness or stalls when authenticating. You can flush the cache to synchronize the mappings.

This procedure is available only through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Run the `isi auth mapping flush` command.

   The following command flushes all identity mappings on the EMC Isilon cluster:

   ```
   isi auth mapping flush --all
   ```

   The following command flushes the mapping of the user with UID 4236 in the zone3 access zone:

   ```
   isi auth mapping flush --source-uid-4236 --zone=zone3
   ```

View a user token

You can view the contents of an access token generated for a user during authentication.

This procedure is available only through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Run the `isi auth mapping token` command.

   The following command displays the access token of a user with UID 4236 in the zone3 access zone:

   ```
   isi auth mapping token --uid=4236 --zone=zone3
   ```

   The system displays output similar to the following example:

   ```
   User
    Name: user_36
    UID: 4236
    SID: S-1-22-1-4236
    On Disk: 4236
   ZID: 3
   Zone: zone3
   Privileges: -
   Primary Group
           Name: user_36
           GID: 4236
   ```

```
        SID: S-1-22-2-4236
  On Disk: 4236
```

Configure identity mapping settings

You can enable or disable automatic allocation of UIDs and GIDS and customize the range of ID values in each access zone. The default range is 1000000–2000000.

This procedure is available only through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Run the `isi auth settings mapping modify` command.

   The following command enables automatic allocation of both UIDs and GIDs in the zone3 access zone and sets their allocation ranges to 25000–50000:

   ```
   isi auth settings mapping modify --gid-range-enabled=yes \
   --gid-range-min=25000 --gid-range-max=50000 --uid-range-
   enabled=yes \
   --uid-range-min=25000 --uid-range-max=50000 --zone=zone3
   ```

View identity mapping settings

You can view the current configuration of identity mapping settings in each zone.

This procedure is available only through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Run the `isi auth settings mapping view` command.

   The following command displays the current settings in the zone3 access zone:

   ```
   isi auth settings mapping view --zone=zone3
   ```

   The system displays output similar to the following example:

   ```
   GID Range Enabled: Yes
       GID Range Min: 25000
       GID Range Max: 50000
   UID Range Enabled: Yes
       UID Range Min: 25000
       UID Range Max: 50000
   ```

Managing user identities

You can manage user identities by creating user-mapping rules.

When you create user-mapping rules, it is important to remember the following information:

- You can only create user-mapping rules if you are connected to the EMC Isilon cluster through the System zone; however, you can apply user-mapping rules to specific access zones. If you create a user-mapping rule for a specific access zone, the rule applies only in the context of its zone.
- When you change user-mapping on one node, OneFS propagates the change to the other nodes.

- After you make a user-mapping change, the OneFS authentication service reloads the configuration.

View user identity

You can view the identities and group membership that a specified user has within the Active Directory and LDAP directory services.

This procedure must be performed through the command-line interface (CLI).

Note

The OneFS user access token contains a combination of identities from Active Directory and LDAP if both directory services are configured. You can run the following commands to discover the identities that are within each specific directory service.

Procedure

1. Establish an SSH connection to any node in the cluster.
2. View a user identity from Active Directory only by running the `isi auth users view` command.

   The following command displays the identity of a user named stand in the Active Directory domain named YORK:

   ```
   isi auth users view --user=YORK\\stand --show-groups
   ```

   The system displays output similar to the following example:

   ```
              Name: YORK\stand
                DN:
   CN=stand,CN=Users,DC=york,DC=hull,DC=example,DC=com
        DNS Domain: york.hull.example.com
            Domain: YORK
          Provider: lsa-activedirectory-provider:YORK.HULL.EXAMPLE.COM
   Sam Account Name: stand
               UID: 4326
               SID: S-1-5-21-1195855716-1269722693-1240286574-591111
     Primary Group
                  ID : GID:1000000
                Name : YORK\york_sh_udg
    Additional Groups: YORK\sd-york space group
                       YORK\york_sh_udg
                       YORK\sd-york-group
                       YORK\sd-group
                       YORK\domain users
   ```

3. Vew a user identity from LDAP only by running the `isi auth users view` command.

   The following command displays the identity of an LDAP user named stand:

   ```
   isi auth user view --user=stand --show-groups
   ```

   The system displays output similar to the following example:

   ```
              Name: stand
           DN:
   uid=stand,ou=People,dc=colorado4,dc=hull,dc=example,dc=com
   DNS Domain: -
       Domain: LDAP_USERS
     Provider: lsa-ldap-provider:Unix LDAP
   Sam Account Name: stand
   ```

```
           UID: 4326
           SID: S-1-22-1-4326
  Primary Group
              ID : GID:7222
            Name : stand
Additional Groups: stand
                   sd-group
                   sd-group2
```

Create a user-mapping rule

You can create a user-mapping rule to manage user identities.

Procedure

1. Click **Access** › **Membership & Roles** › **User Mapping**.
2. From the **Current Access Zone** list, select an access zone that contains the rules you want to manage, and then click **Edit User Mapping Rules**.

   The **Edit User Mapping Rules** dialog box appears.
3. Click **Create a User Mapping Rule.**

   The **Create a User Mapping Rule** dialog box appears.
4. From the **Operation** list, select an operation.

   Depending on your selection, the **Create a User Mapping Rule** displays additional fields.
5. Fill in the fields as needed.
6. Click **Add Rule** to save the rule and return to the **Edit User Mapping Rules** dialog box.
7. In the **User Mapping Rules** area, click the title bar of a rule and drag it to a new position to change the position of a rule in the list.

   Rules are applied in the order they are listed. To ensure that each rule gets processed, list replacement rules first and list allow or deny rules at the end.
8. If the access token is not associated with a default UNIX user or if the default UNIX user does not have a primary UID or GID, select one of the following options for authentication:
   - Generate a primary UID or GID from the reserved range of UIDs and GIDs
   - Deny access to the user
   - Assign another user as the default UNIX user

   Note

   It is recommended that you assign a user from the well-known account that has a read-only access.
9. Click **Save Changes**.

Test a user-mapping rule

After creating a user-mapping rule, you can test it to make sure that the token results for a user are as expected.

Procedure

1. Click **Access** › **Membership & Roles** › **User Mapping**.
2. From the **Current Access Zone** list, select an access zone that contains the rules you want to test.

3. In the **Test User Mapping** section, type or select a user, group, or a well-known SID from the **Select a User** dialog box.
4. Click **Test Mapping**.

   The token results appear in the **Results** section as shown:

```
User
    Name:krb_user_002
    UID:1002
    SID:S-1-22-1-1001
    On disk:1001
    ZID:1
    Zone:System

Privileges:-

Primary Group
    Name:krb_user_001
    GID:1000
    SID:S-1-22-2-1001
    On disk:1000

Supplemental Identities
    Name:Authenticated Users
    GID: -
    SID:S-1-5-11
```

Merge Windows and UNIX tokens

You can use either the join or append operator to merge tokens from different directory services into a single OneFS user token.

When Windows and Unix user names do not match across directory services, you can write user-mapping rules that use either the join or the append operator to merge two user names into a single token. For example, if a user's Windows username is win_bob and the users Unix username is UNIX_bob, you can join or append the user tokens of the two different users.
When you append an account to another account, the append operator adds information from one identity to another: OneFS appends the fields that the options specify from the source identity to the target identity. OneFS appends the identifiers to the additional group list.

Procedure

1. Click **Access** › **Membership & Roles** › **User Mapping**.
2. Select the **Current Access Zone** that contains the rules you want to manage, and then click **Edit User Mapping Rules**.

   The **Edit User Mapping Rules** dialog box appears.
3. Click **Create a User Mapping Rule**.

   The **Create a User Mapping Rule** dialog box appears.
4. From the **Operation** list, select an option:

Option	Description
Join two users together	Inserts the new identity into the token.
Append field from a user	Modifies the access token by adding fields to it.

   Depending on your selection, the **Create a User Mapping Rule** dialog box refreshes to display additional fields.

5. Populate the fields as needed.
6. Click **Add Rule**.

   Note

   Rules are called in the order they are listed. To ensure that each rule gets processed, list replacements first and allow/deny rules last. You can change the order in which a rule is listed by clicking its title bar and dragging it to a new position.

7. Click **Save Changes**.

Retrieve the primary group from LDAP

You can create a user-mapping rule to insert primary group information from LDAP into a user's access token.

By default, the user-mapping service combines information from AD and LDAP but gives precedence to the information from AD. You can create a mapping rule to control how OneFS combines the information, giving precedence to a primary group from LDAP rather than from Active Directory for a user.

Procedure

1. Click **Access** › **Membership & Roles** › **User Mapping**.
2. Select the **Current Access Zone** that contains the rules you want to manage, and then click **Edit User Mapping Rules**.

   The **Edit User Mapping Rules** dialog box appears.
3. Click **Create a User Mapping Rule**.

   The **Create a User Mapping Rule** dialog box appears.
4. From the **Operation** list, select **Insert fields from a user**.

   The **Create a User Mapping Rule** dialog box refreshes to display additional fields.
5. To populate the **Insert Fields into this User** field, perform the following steps:
   a. Click **Browse**.

      The **Select a User** dialog box appears.
   b. Select a user and an Active Directory authentication provider.
   c. Click **Search** to view the search results.
   d. Select a username and click **Select** to return to the **Create a User Mapping Rule** dialog box.

      The primary group of the second user is inserted as the primary group of the first user.
6. Select the **Insert primary group SID and GID** check box.
7. To populate the **Insert Fields from this User** field, perform the following steps:
   a. Click **Browse**.

      The **Select a User** dialog box appears.
   b. Select a user and an LDAP authentication provider.
   c. Click **Search** to view the search results.
   d. Select a username and click **Select** to return to the **Create a User Mapping Rule** dialog box.

8. Click **Add Rule**.

Note

Rules are called in the order they are listed. To ensure that each rule gets processed, list the replacements first and the allow or deny rules at the end. You can change the order in which a rule is listed by clicking its title bar and dragging it to a new position.

9. Click **Save Changes**.

Mapping rule options

Mapping rules can contain options that target the fields of an access token.

A field represents an aspect of a cross-domain access token, such as the primary UID and primary user SID from a user that you select. You can see some of the fields in the OneFS web administration interface. **User** in the web administration interface is the same as username. You can also see fields in an access token by running the command `isi auth mapping token`.

When you create a rule, you can add an option to manipulate how OneFS combines aspects of two identities into a single token. For example, an option can force OneFS to append the supplement groups to a token.

A token includes the following fields that you can manipulate with user mapping rules:

- username
- unix_name
- primary_uid
- primary_user_sid
- primary_gid
- primary_group_sid
- additional_ids (includes supplemental groups)

Options control how a rule combines identity information in a token. The break option is the exception: It stops OneFS from processing additional rules.

Although several options can apply to a rule, not all options apply to all operators. The following table describes the effect of each option and the operators that they work with.

Option	Operator	Description
user	insert, append	Copies the primary UID and primary user SID, if they exist, to the token.
groups	insert, append	Copies the primary GID and primary group SID, if they exist, to the token.
groups	insert, append	Copies all the additional identifiers to the token. The additional identifiers exclude the primary UID, the primary GID, the primary user SID, and the primary group SID.
default_user	all operators except remove groups	If the mapping service fails to find the second user in a rule, the service tries to find the username of the default user. The name of the default user cannot include wildcards. When you set the option for the default user in a rule with the command-line interface, you must set it with an underscore: default_user.

Option	Operator	Description
break	all operators	Stops the mapping service from applying rules that follow the insertion point of the break option. The mapping service generates the final token at the point of the break.

Mapping rule operators

The operator determines what a mapping rule does.

You can create user-mapping rules through either the web-administration interface, where the operators are spelled out in a list, or from the command-line interface.

When you create a mapping rule with the OneFS command-line interface (CLI), you must specify an operator with a symbol. The operator affects the direction in which the mapping service processes a rule. For more information about creating a mapping rule, see the white paper *Managing identities with the Isilon OneFS user mapping service*. The following table describes the operators that you can use in a mapping rule.

A mapping rule can contain only one operator.

Operator	Web interface	CLI	Direction	Description
append	**Append fields from a user**	++	Left-to-right	Modifies an access token by adding fields to it. The mapping service appends the fields that are specified in the list of options (user, group, groups) to the first identity in the rule. The fields are copied from the second identity in the rule. All appended identifiers become members of the additional groups list. An append rule without an option performs only a lookup operation; you must include an option to alter a token.
insert	**Insert fields from a user**	+=	Left-to-right	Modifies an existing access token by adding fields to it. Fields specified in the options list (user, group, groups) are copied from the new identity and inserted into the identity in the token. When the rule inserts a primary user or primary group, it become the new primary user and primary group in the token. The previous primary user and primary group move to the additional identifiers list. Modifying the primary user leaves the token's username unchanged. When inserting the additional groups from an identity, the service adds the new groups to the existing groups.
replace	**Replace one user with a different user**	=>	Left-to-right	Removes the token and replaces it with the new token that is identified by the second username. If the second username is empty, the mapping service removes the first username in the token, leaving no username. If a token contains no username, OneFS denies access with a `no such user` error.

Operator	Web interface	CLI	Direction	Description
remove groups	**Remove supplemental groups from a user**	--	Unary	Modifies a token by removing the supplemental groups.
join	**Join two users together**	&=	Bidirectional	Inserts the new identity into the token. If the new identity is the second user, the mapping service inserts it after the existing identity; otherwise, the service inserts it before the existing identity. The location of the insertion point is relevant when the existing identity is already the first in the list because OneFS uses the first identity to determine the ownership of new file system objects.

CHAPTER 7

Auditing

This section contains the following topics:

Auditing overview

You can audit system configuration changes and SMB and NFS protocol activity on an EMC Isilon cluster. All audit data is stored and protected in the cluster file system and organized by audit topics.

When you enable system configuration auditing, no additional configuration is required; all configuration events that are handled by the application programming interface (API) through the command-line interface (CLI) are tracked and recorded in the `config` audit topic directories.

Auditing can detect many potential sources of data loss, including fraudulent activities, inappropriate entitlements, and unauthorized access attempts. Customers in industries such as financial services, health care, life sciences, and media and entertainment, as well as in governmental agencies, must meet stringent regulatory requirements developed to protect against these sources of data loss.

You can enable and configure protocol auditing for one or more access zones in a cluster. If you enable protocol auditing for an access zone, file-access events through the SMB and NFS protocol are recorded in the protocol audit topic directories. The `protocol` audit log file is consumable by auditing applications that support the EMC Common Event Enabler (CEE). You can specify which events to log in each access zone. For example, you might want to audit the default set of `protocol` events in the System access zone but audit only successful attempts to delete files in a different access zone.

The audit events are logged on the individual nodes where the SMB or NFS client initiated the activity. The events are then stored in a binary file under `/ifs/.ifsvar/audit/logs`. The logs automatically roll over to a new file after the size reaches 1 GB.

Syslog

Syslog is a protocol that is used to convey certain event notification messages. The root user can configure an Isilon cluster to log audit events and forward them to syslog by using the syslog forwarder.

By default, all protocol events that occur on a particular node are forwarded to the `/var/log/audit_protocol.log` file, regardless of the access zone the event originated from.

Syslog is configured with an identity of `audit_protocol`, a facility of `syslog`, and a priority level of `info`.

Enable syslog

By default, audit event forwarding to syslog is not enabled when auditing is enabled. To enable this feature, you must configure audit syslog settings through the command line interface for zones.

Before you begin

Note

To enable audit event forwarding, you must configure audit syslog settings for each access zone. This procedure is available only through the command-line interface (CLI).

Procedure

1. Open a Secure Shell (SSH) connection to any node in the cluster and log in.
2. Run the `isi zone zones modify` command with the `--syslog-forwarding-enabled` option to enable or disable audit syslog.

   The following command enables audit syslog for an access zone named UserZone:

   ```
   isi zone zones modify UserZone --syslog-forwarding-enabled=yes
   ```

   The following command disables audit syslog for the UserZone access zone:

   ```
   isi zone zones modify UserZone --syslog-forwarding-enabled=no
   ```
3. To view the audit settings, run the following command:

   ```
   isi audit settings view
   ```
4. To view the audit configuration settings, run the following command:

   ```
   isi zone zones view <zone>
   ```

   Note

   The syslog forwarder forwards only the zone's audit events to syslog that are set by the `--syslog-audit-events` parameter. This parameter is set to a list of comma-separated audit event types or "all" to set all the audit events. Only audit events that are defined by running the `isi zone zones modify` command with the `--audit-success` and `--audit-failure` options are eligible for forwarding to syslog.

Syslog forwarding

The syslog forwarder is a daemon that, when enabled, retrieves configuration changes and protocol audit events in an access zone and forwards the events to syslog. Only user-defined audit success and failure events are eligible for being forwarded to syslog.

On each node there is an audit syslog forwarder daemon running that will log audit events to the same node's syslog daemon.

Protocol audit events

By default, audited access zones track only certain events on the EMC Isilon cluster, including successful and failed attempts to access files and directories.

The default tracked events are create, close, delete, rename, and set_security.

The names of generated events are loosely based on the Windows I/O request packet (IRP) model in which all operations begin with a create event to obtain a file handle. A create event is required before all I/O operations, including the following: close, create, delete, get_security, read, rename, set_security, and write. A close event marks when the client is finished with the file handle that was produced by a create event.

These internally stored events are translated to events that are forwarded through CEE to the auditing application. The CEE export facilities on OneFS perform this mapping. CEE can be used to connect to any third party application that supports CEE.

Different SMB and NFS clients issue different requests, and one particular version of a platform such as Windows or Mac OS X using SMB might differ from another. Similarly, different versions of an application such as Microsoft Word or Windows Explorer might make different protocol requests. For example, a client with a Windows Explorer window open might generate many events if an automatic or manual refresh of that window occurs. Applications issue requests with the logged-in user's credentials, but you should not assume that all requests are purposeful user actions.

When enabled, OneFS audit will track all changes that are made to the files and directories in SMB shares and NFS exports.

Sample config audit log

You can view both configuration audit and protocol audit logs by running the `isi_audit_viewer` command on any node in the Isilon cluster.

The `isi_audit_viewer -t config` command produces output similar to the following example:

```
[0: Fri Jan 23 16:17:03 2015] {"id":"524e0928-
a35e-11e4-9d0c-005056302134","timestamp":
1422058623106323,"payload":"PAPI config logging started."}

[1: Fri Jan 23 16:17:03 2015] {"id":"5249b99d-
a35e-11e4-9d0c-005056302134","timestamp":1422058623112992,"payload":
{"user":{"token": {"UID":0, "GID":0, "SID": "SID:S-1-22-1-0", "GSID":
"SID:S-1-22-2-0", "GROUPS": ["SID:S-1-5-11", "GID:5", "GID:20", "GID:
70", "GID:10"], "protocol": 17, "zone id": 1, "client":
"10.7.220.97", "local": "10.7.177.176" }},"uri":"/1/protocols/smb/
shares","method":"POST","args":"","body":{"path": "/ifs/data",
"name": "Test"}}}

[2: Fri Jan 23 16:17:05 2015] {"id":"5249b99d-
a35e-11e4-9d0c-005056302134","timestamp":1422058625144567,"payload":
{"status":201,"statusmsg":"Created","body":{"id":"Test"}}}

[3: Fri Jan 23 16:17:39 2015] {"id":"67e7ca62-
a35e-11e4-9d0c-005056302134","timestamp":1422058659345539,"payload":
{"user":{"token": {"UID":0, "GID":0, "SID": "SID:S-1-22-1-0", "GSID":
"SID:S-1-22-2-0", "GROUPS": ["SID:S-1-5-11", "GID:5", "GID:20", "GID:
70", "GID:10"], "protocol": 17, "zone id": 1, "client":
"10.7.220.97", "local": "10.7.177.176" }},"uri":"/1/audit/
settings","method":"PUT","args":"","body":{"config_syslog_enabled":
true}}}

[4: Fri Jan 23 16:17:39 2015] {"id":"67e7ca62-
a35e-11e4-9d0c-005056302134","timestamp":1422058659387928,"payload":
{"status":204,"statusmsg":"No Content","body":{}}}
```

Events come in pairs; a *pre event* is logged before the command is carried out and a *post event* is logged after the event is triggered. These events can be correlated by matching the `id` field. In the above logs, events 1 and 2 are paired, and events 3 and 4 are paired.

The pre event always comes first, and contains user token information, the PAPI path, and whatever arguments were passed to the PAPI call. In event 1, a POST request was made to `/1/protocols/smb/shares` with arguments `path=/ifs/data` and `name=Test`. The post event contains the HTTP return status and any output returned from the server.

Supported event types

You can view or modify the event types that are audited in an access zone.

For the most current list of supported auditing tools, see the Isilon Third-Party Software & Hardware Compatibility Guide.

The following event types are configured by default on each audited access zone:

Event name	Example protocol activity
create	• Create a file or directory • Open a file, directory, or share • Mount a share • Delete a file
close	• Close a directory • Close a modified or unmodified file
rename	Rename a file or directory
delete	Delete a file or directory
set_security	Attempt to modify file or directory permissions

The following event types are available for exporting through CEE:

Event name	Example protocol activity
read	The first read request on an open file handle
write	The first write request on an open file handle
close	The client is finished with an open file handle
get_security	The client reads security information for an open file handle

The following protocol audit events can not be exported through CEE:

Event name	Example protocol activity
logon	SMB session create request by a client
logoff	SMB session logoff
tree_connect	SMB first attempt to access a share

Supported audit tools

You can configure OneFS to send protocol auditing logs to servers that support the EMC Common Event Enabler (CEE).

CEE has been tested and verified to work on several third-party software vendors. For the most current list of supported auditing tools, see the Isilon Third-Party Software & Hardware Compatibility Guide.

Note

We recommend that you install and configure third-party auditing applications before you enable the OneFS auditing feature. Otherwise, the large backlog consumed by this feature may cause results to not be updated for a considerable amount of time.

Enable system configuration auditing

You can enable or disable the auditing of system configuration changes. No additional settings are available.

Note

It is recommended that you install and configure third-party auditing applications before you enable the OneFS auditing feature. Otherwise, the backlog consumed by the tool may be so large that results may be stale for a prolonged time.

Procedure

1. Click **Cluster Management** › **Auditing**.
2. In the **Settings** area, select the **Enable Configuration Change Auditing** checkbox.
3. Click **Save Changes**.

Enable protocol access auditing

You can audit SMB protocol access on a per-access zone basis and optionally forward the generated events to the EMC Common Event Enabler (CEE) for export to third-party products.

Note

It is recommended that you install and configure third-party auditing applications before you enable the OneFS auditing feature. Otherwise, the backlog consumed by the tool may become so large that results might be stale for a prolonged time.

Procedure

1. Click **Cluster Management** › **Auditing**.
2. In the **Settings** area, select the **Enable Protocol Access Auditing** checkbox.
3. In the **Audited Zones** area, click **Add Zones**.
4. In the **Select Zones** dialog box, select the checkbox for one or more access zones, and then click **Add Zones**.
5. (Optional) In the **Event Forwarding** area, specify one or more CEE servers to forward logged events to.
   a. In the **CEE Server URIs** field, type the URI of each CEE server in the CEE server pool.

      The OneFS CEE export service uses round robin load-balancing when exporting events to multiple CEE servers. Valid URIs start with `http://` and include the port number and path to the CEE server if necessary—for example, `http://example.com:12228/cee`.

   b. In the **Storage Cluster Name** field, specify the name of the storage cluster to use when forwarding protocol events.

This name value is typically the SmartConnect zone name, but in cases where SmartConnect is not implemented, the value must match the hostname of the cluster as Varonis recognizes it. If the field is left blank, events from each node are filled with the node name (clustername + lnn). This setting is required only if needed by your third-party audit application.

Note

Although this step is optional, be aware that a backlog of events will accumulate regardless of whether CEE servers have been configured. When configured, CEE forwarding begins with the oldest events in the backlog and moves toward newest events in a first-in-first-out sequence.

6. Click **Save Changes**.

Results

The following protocol events, which are the only events supported by Varonis DatAdvantage, are collected for audited access zones by default: `create`, `close`, `delete`, `rename`, and `set_security`. You can modify the set of events that are audited in an access zone by running the `isi zone zones modify` command in the command-line interface. Because each audited event consumes system resources, it is recommended that you only configure zones for events that are needed by your auditing application.

Auditing settings

You can view or modify basic settings for configuration change auditing and protocol access auditing.

`Enable Configuration Change Auditing`
Audits requests that are made through the API for system configuration changes.

`Enable Protocol Access Auditing`
Audits requests that are made through the SMB protocol to access data.

`Audited Zones`
Specifies one or more access zones to audit. This setting applies only to protocol access auditing.

`CEE Server URIs`
Specifies one or more CEE server URIs where audit events will be forwarded. The OneFS CEE export service uses round robin load-balancing when exporting events to multiple CEE servers. This setting applies only to protocol access auditing.

`Storage Cluster Name`
Specifies the name of the storage cluster to use when forwarding protocol events—typically, the SmartConnect zone name. This setting is required only if needed by your third-party audit application.

CHAPTER 8

File sharing

This section contains the following topics:

File sharing overview

Multi-protocol support in OneFS enables files and directories on the Isilon cluster to be accessed through SMB for Windows file sharing, NFS for UNIX file sharing, secure shell (SSH), FTP, and HTTP. By default, only the SMB and NFS protocols are enabled.

OneFS creates the `/ifs` directory, which is the root directory for all file system data on the cluster. The `/ifs` directory is configured as an SMB share and an NFS export by default. You can create additional shares and exports within the `/ifs` directory tree.

Note

We recommend that you do not save data to the root `/ifs` file path but in directories below `/ifs`. The design of your data storage structure should be planned carefully. A well-designed directory structure optimizes cluster performance and administration.

You can set Windows- and UNIX-based permissions on OneFS files and directories. Users who have the required permissions and administrative privileges can create, modify, and read data on the cluster through one or more of the supported file sharing protocols.

- SMB. Allows Microsoft Windows and Mac OS X clients to access files that are stored on the cluster.
- NFS. Allows Linux and UNIX clients that adhere to the RFC1813 (NFSv3) and RFC3530 (NFSv4) specifications to access files that are stored on the cluster.
- HTTP and HTTPS (with optional DAV). Allows clients to access files that are stored on the cluster through a web browser.
- FTP. Allows any client that is equipped with an FTP client program to access files that are stored on the cluster through the FTP protocol.

Mixed protocol environments

The `/ifs` directory is the root directory for all file system data in the cluster, serving as an SMB share, an NFS export, and a document root directory. You can create additional shares and exports within the `/ifs` directory tree. You can configure your OneFS cluster to use SMB or NFS exclusively. You can also enable HTTP, FTP, and SSH.

Access rights are consistently enforced across access protocols on all security models. A user is granted or denied the same rights to a file whether using SMB or NFS. Clusters running OneFS support a set of global policy settings that enable you to customize the default access control list (ACL) and UNIX permissions settings.

OneFS is configured with standard UNIX permissions on the file tree. Through Windows Explorer or OneFS administrative tools, you can give any file or directory an ACL. In addition to Windows domain users and groups, ACLs in OneFS can include local, NIS, and LDAP users and groups. After a file is given an ACL, the mode bits are no longer enforced and exist only as an estimate of the effective permissions.

Note

We recommend that you configure ACL and UNIX permissions only if you fully understand how they interact with one another.

Write caching with SmartCache

Write caching accelerates the process of writing data to the cluster. OneFS includes a write-caching feature called SmartCache, which is enabled by default for all files and directories.

If write caching is enabled, OneFS writes data to a write-back cache instead of immediately writing the data to disk. OneFS can write the data to disk at a time that is more convenient.

Note

We recommend that you keep write caching enabled. You should also enable write caching for all file pool policies.

OneFS interprets writes to the cluster as either synchronous or asynchronous, depending on a client's specifications. The impacts and risks of write caching depend on what protocols clients use to write to the cluster, and whether the writes are interpreted as synchronous or asynchronous. If you disable write caching, client specifications are ignored and all writes are performed synchronously.

The following table explains how clients' specifications are interpreted, according to the protocol.

Protocol	Synchronous	Asynchronous
NFS	The stable field is set to `data_sync` or `file_sync`.	The stable field is set to `unstable`.
SMB	The `write-through` flag has been applied.	The `write-through` flag has not been applied.

Write caching for asynchronous writes

Writing to the cluster asynchronously with write caching is the fastest method of writing data to your cluster.

Write caching for asynchronous writes requires fewer cluster resources than write caching for synchronous writes, and will improve overall cluster performance for most workflows. However, there is some risk of data loss with asynchronous writes.

The following table describes the risk of data loss for each protocol when write caching for asynchronous writes is enabled:

Protocol	Risk
NFS	If a node fails, no data will be lost except in the unlikely event that a client of that node also crashes before it can reconnect to the cluster. In that situation, asynchronous writes that have not been committed to disk will be lost.
SMB	If a node fails, asynchronous writes that have not been committed to disk will be lost.

We recommend that you do not disable write caching, regardless of the protocol that you are writing with. If you are writing to the cluster with asynchronous writes, and you decide that the risks of data loss are too great, we recommend that you configure your clients to use synchronous writes, rather than disable write caching.

Write caching for synchronous writes

Write caching for synchronous writes costs cluster resources, including a negligible amount of storage space. Although it is not as fast as write caching with asynchronous writes, unless cluster resources are extremely limited, write caching with synchronous writes is faster than writing to the cluster without write caching.

Write caching does not affect the integrity of synchronous writes; if a cluster or a node fails, none of the data in the write-back cache for synchronous writes is lost.

SMB

OneFS includes a configurable SMB service to create and manage SMB shares. SMB shares provide Windows clients network access to file system resources on the cluster. You can grant permissions to users and groups to carry out operations such as reading, writing, and setting access permissions on SMB shares.

The `/ifs` directory is configured as an SMB share and is enabled by default. OneFS supports both user and anonymous security modes. If the user security mode is enabled, users who connect to a share from an SMB client must provide a valid user name with proper credentials.

SMB shares act as checkpoints, and users must have access to a share in order to access objects in a file system on a share. If a user has access granted to a file system, but not to the share on which it resides, that user will not be able to access the file system regardless of privileges. For example, assume a share named `ABCDocs` contains a file named `file1.txt` in a path such as: `/ifs/data/ABCDocs/file1.txt`. If a user attempting to access `file1.txt` does not have share privileges on `ABCDocs`, that user cannot access the file even if originally granted read and/or write privileges to the file.

The SMB protocol uses security identifiers (SIDs) for authorization data. All identities are converted to SIDs during retrieval and are converted back to their on-disk representation before they are stored on the cluster.

When a file or directory is created, OneFS checks the access control list (ACL) of its parent directory. If the ACL contains any inheritable access control entries (ACEs), a new ACL is generated from those ACEs. Otherwise, OneFS creates an ACL from the combined file and directory create mask and create mode settings.

OneFS supports the following SMB clients:

SMB version	Supported operating systems
1.0	Windows 2000 or later Windows XP or later Mac OS X 10.5 or later
2.0	Windows Vista or later Windows Server 2008 or later Mac OS X 10.9 or later
2.1	Windows 7 or later Windows Server 2008 R2 or later
3.0 - Multichannel only	Windows 8 or later Windows Server 2012 or later

SMB shares in access zones

You can create and manage SMB shares within access zones.

You can create access zones that partition storage on the EMC Isilon cluster into multiple virtual containers. Access zones support all configuration settings for authentication and identity management services on the cluster, so you can configure authentication providers and provision SMB shares on a zone-by-zone basis. When you create an access zone, a local provider is created automatically, which allows you to configure each access zone with a list of local users and groups. You can also authenticate through a different Active Directory provider in each access zone, and you can control data access by directing incoming connections to the access zone from a specific IP address in a pool. Associating an access zone with an IP address pool restricts authentication to the associated access zone and reduces the number of available and accessible SMB shares.

Here are a few ways to simplify SMB management with access zones:

- Migrate multiple SMB servers, such as Windows file servers or NetApp filers, to a single Isilon cluster, and then configure a separate access zone for each SMB server.
- Configure each access zone with a unique set of SMB share names that do not conflict with share names in other access zones, and then join each access zone to a different Active Directory domain.
- Reduce the number of available and accessible shares to manage by associating an IP address pool with an access zone to restrict authentication to the zone.
- Configure default SMB share settings that apply to all shares in an access zone.

The Isilon cluster includes a built-in access zone named System, where you manage all aspects of the cluster and other access zones. If you don't specify an access zone when managing SMB shares, OneFS will default to the System zone.

SMB Multichannel

SMB Multichannel supports establishing a single SMB session over multiple network connections.

SMB Multichannel is a feature of the SMB 3.0 protocol that provides the following capabilities:

Increased throughput

OneFS can transmit more data to a client through multiple connections over high speed network adapters or over multiple network adapters.

Connection failure tolerance

When an SMB Multichannel session is established over multiple network connections, the session is not lost if one of the connections has a network fault, which enables the client to continue to work.

Automatic discovery

SMB Multichannel automatically discovers supported hardware configurations on the client that have multiple available network paths and then negotiates and establishes a session over multiple network connections. You are not required to install components, roles, role services, or features.

SMB Multichannel requirements

You must meet software and NIC configuration requirements to support SMB Multichannel on the EMC Isilon cluster.

OneFS can only support SMB Multichannel when the following software requirements are met:

- Windows Server 2012, 2012r2 or Windows 8, 8.1 clients
- SMB Multichannel must be enabled on both the EMC Isilon cluster and the Windows client computer. It is enabled on the Isilon cluster by default.

SMB Multichannel establishes a single SMB session over multiple network connections only on supported network interface card (NIC) configurations. SMB Multichannel requires at least one of the following NIC configurations on the client computer:

- Two or more network interface cards.
- One or more network interface cards that support Receive Side Scaling (RSS).
- One or more network interface cards configured with link aggregation. Link aggregation enables you to combine the bandwidth of multiple NICs on a node into a single logical interface.

Client-side NIC configurations supported by SMB Multichannel

SMB Multichannel automatically discovers supported hardware configurations on the client that have multiple available network paths.

Each node on the EMC Isilon cluster has at least one RSS-capable network interface card (NIC). Your client-side NIC configuration determines how SMB Multichannel establishes simultaneous network connections per SMB session.

Client-side NIC Configuration	Description
Single RSS-capable NIC	SMB Multichannel establishes a maximum of four network connections to the Isilon cluster over the NIC. The connections are more likely to be spread across multiple CPU cores, which reduces the likelihood of performance bottleneck issues and achieves the maximum speed capability of the NIC.
Multiple NICs	If the NICs are RSS-capable, SMB Multichannel establishes a maximum of four network connections to the Isilon cluster over each NIC. If the NICs on the client are not RSS-capable, SMB Multichannel establishes a single network connection to the Isilon cluster over each NIC. Both configurations allow SMB Multichannel to leverage the combined bandwidth of multiple NICs and provides connection fault tolerance if a connection or a NIC fails. **Note** SMB Multichannel cannot establish more than eight simultaneous network connections per session. In a multiple NIC configuration, this might limit the number connections allowed per NIC. For example, if the configuration contains three RSS-capable NICs, SMB Multichannel might establish three connections over the first NIC, three connections over the second NIC and two connections over the third NIC.
Aggregated NICs	SMB Multichannel establishes multiple network connections to the Isilon cluster over aggregated NICs, which results in balanced connections across CPU cores, effective consumption of combined bandwidth, and connection fault tolerance. **Note** The aggregated NIC configuration inherently provides NIC fault tolerance that is not dependent upon SMB.

SMB share management through MMC

OneFS supports the Shared Folders snap-in for the Microsoft Management Console (MMC), which allows SMB shares on the EMC Isilon cluster to be managed using the MMC tool.

Typically, you connect to the global System zone through the web administration interface or the command line interface to manage and configure shares. If you configure access zones, you can connect to a zone through the MMC Shared Folders snap-in to directly manage all shares in that zone.

You can establish a connection through the MMC Shared Folders snap-in to an Isilon node and perform the following SMB share management tasks:

- Create and delete shared folders
- Configure access permission to an SMB share
- View a list of active SMB sessions
- Close open SMB sessions
- View a list of open files
- Close open files

When you connect to a zone through the MMC Shared Folders snap-in, you can view and manage all SMB shares assigned to that zone; however, you can only view active SMB sessions and open files on the specific node that you are connected to in that zone. Changes you make to shares through the MMC Shared Folders snap-in are propagated across the cluster.

MMC connection requirements

You can connect to an EMC Isilon cluster through the MMC Shared Folders snap-in if you meet access requirements.

The following conditions are required to establish a connection through the MMC Shared Folders snap-in:

- You must run the Microsoft Management Console (MMC) from a Windows workstation that is joined to the domain of an Active Directory (AD) provider configured on the cluster.
- You must be a member of the local *<cluster>*\Administrators group.

  Note

  Role-based access control (RBAC) privileges do not apply to the MMC. A role with SMB privileges is not sufficient to gain access.

- You must log in to a Windows workstation as an Active Directory user that is a member of the local *<cluster>*\Administrators group.

Symbolic links and SMB clients

OneFS enables SMB2 clients to access symbolic links in a seamless manner. Many administrators deploy symbolic links to virtually reorder file system hierarchies, especially when crucial files or directories are scattered around an environment.

In an SMB share, a symbolic link (also known as a symlink or a soft link) is a type of file that contains a path to a target file or directory. Symbolic links are transparent to applications running on SMB clients, and they function as typical files and directories.

Support for relative and absolute links is enabled by the SMB client. The specific configuration depends on the client type and version.

A symbolic link that points to a network file or directory that is not in the path of the active SMB session is referred to as an absolute (or remote) link. Absolute links always point to the same location on a file system, regardless of the present working directory, and usually contain the root directory as part of the path. Conversely, a relative link is a symbolic link that points directly to a user's or application's working directory, so you do not have to specify the full absolute path when creating the link.

OneFS exposes symbolic links through the SMB2 protocol, enabling SMB2 clients to resolve the links instead of relying on OneFS to resolve the links on behalf of the clients. To transverse a relative or absolute link, the SMB client must be authenticated to the SMB shares that the link can be followed through. However, if the SMB client does not have permission to access the share, access to the target is denied and Windows will not prompt the user for credentials.

SMB2 and NFS links are interoperable for relative links only. For maximum compatibility, create these links from a POSIX client.

Note

SMB1 clients (such as Windows XP or 2002) may still use relative links, but they are traversed on the server side and referred to as "shortcut files." Absolute links do not work in these environments.

Enabling symbolic links

Before you can fully use symbolic links in an SMB environment, you must enable them.

For Windows SMB clients to traverse each type of symbolic link, you must enable them on the client. Windows supports the following link types:

- local to local
- remote to remote
- local to remote
- remote to local

You must run the following Windows command to enable all four link types:

```
fsutil behavior set SymlinkEvaluation L2L:1 R2R:1 L2R:1 R2L:1
```

For POSIX clients using Samba, you must set the following options in the `[global]` section of your Samba configuration file (`smb.conf`) to enable Samba clients to traverse relative and absolute links:

```
follow symlinks=yes
wide links=yes
```

In this case, "wide links" in the `smb.conf` file refers to absolute links. The default setting in this file is `no`.

Managing symbolic links

After enabling symbolic links, you can create or delete them from the Windows command prompt or a POSIX command line.

Create symbolic links using the Windows `mklink` command on an SMB2 client or the `ln` command from a POSIX command-line interface. For example, an administrator may want to give a user named User1 access to a file named `File1.doc` in the `/ifs/data/` directory without giving specific access to that directory by creating a link named Link1:

```
mklink \ifs\home\users\User1\Link1 \ifs\data\Share1\File1.doc
```

When you create a symbolic link, it is designated as a file link or directory link. Once the link is set, the designation cannot be changed. You can format symbolic link paths as either relative or absolute.

To delete symbolic links, use the `del` command in Windows, or the `rm` command in a POSIX environment.

Keep in mind that when you delete a symbolic link, the target file or directory still exists. However, when you delete a target file or directory, a symbolic link continues to exist and still points to the old target, thus becoming a broken link.

Anonymous access to SMB shares

You can configure anonymous access to SMB shares by enabling the local Guest user and allowing impersonation of the guest user.

For example, if you store files such as browser executables or other data that is public on the internet, anonymous access allows any user to access the SMB share without authenticating.

Managing SMB settings

You can enable or disable the SMB service, configure global settings for the SMB service, and configure default SMB share settings that are specific to each access zone.

Configure SMB server settings

You can enable or disable the SMB server and configure global settings for SMB shares and snapshot directories.

⚠ CAUTION

Modifying the advanced settings could result in operational failures. Be aware of the potential consequences before committing changes to these settings.

Procedure

1. Click **Protocols** › **Windows Sharing (SMB)** › **SMB Server Settings**.
2. From the **SMB service** setting, select **Enabled**.
3. From the **Snapshot Directory Settings** box, choose the system default or a custom configuration for the following settings:
   - Visible at root
   - Accessible at root
   - Visible in subdirectories
   - Accessible in subdirectories
4. Click **Save**.

Configure default SMB share settings

You can configure SMB share settings specific to each access zone.

The default settings are applied to all new shares that are added to the access zone.

⚠ CAUTION

If you modify the default settings, the changes are applied to all existing shares in the access zone unless the setting was configured at the SMB share level.

Procedure

1. Click **Protocols** › **Windows Sharing (SMB)** › **SMB Default Share Settings**.
2. From the **Current Access Zones** drop-down list, select the access zone that the default settings will apply to.
3. From the **File and Directory Permissions Settings** box, choose the system default or a custom configuration for the following settings:
   - Create Permissions
   - Create Mask (Dir)
   - Create Mode (Dir)
   - Create Mask (File)
   - Create Mode (File)
4. From the **Performance Settings** box, choose the system default or a custom configuration for the following settings:
   - Change Notify
   - Oplocks

   Performance settings are advanced and should only be modified if necessary.
5. From the **Security Settings** box, choose the system default or a custom configuration for the following settings:
   - Impersonate Guest
   - Impersonate User
   - NTFS ACL
6. Click **Save**.

Enable or disable SMB Multichannel

SMB Multichannel is required for multiple, concurrent SMB sessions from a Windows client computer to a node in an EMC Isilon cluster. SMB Multichannel is enabled in the Isilon cluster by default.

You can enable or disable SMB Multichannel only through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Run the `isi smb settings global modify` command.

   The following command enables SMB Multichannel on the EMC Isilon cluster:

   ```
   isi smb settings global modify --support-multichannel=yes
   ```

   The following command disables SMB Multichannel on the EMC Isilon cluster:

   ```
   isi smb settings global modify --support-multichannel=no
   ```

Snapshots directory settings

You can view and configure the settings that control the snapshots directories in SMB.

⚠ CAUTION

These settings affect the behavior of the SMB service. Changes to these settings can affect all current and future SMB shares.

Setting	Setting value
Visible at Root	Specifies whether to make the .snapshot directory visible at the root of the share. The default value is `Yes`.
Accessible at Root	Specifies whether to make the .snapshot directory accessible at the root of the share. The default value is `Yes`.
Visible in Subdirectories	Specifies whether to make the .snapshot directory visible in subdirectories of the share root. The default value is `No`.
Accessible in Subdirectories	Specifies whether to make the .snapshot directory accessible in subdirectories of the share root. The default value is `Yes`.

File and directory permission settings

You can view and configure the default source permissions and UNIX create mask/mode bits that are applied when a file or directory is created in an SMB share.

Note

Changes that are made directly to an SMB share override the default settings configured from the **Default SMB Share Settings** tab.

If the mask and mode bits match the default values, a green check mark next to a setting appears, indicating that the specified read (R), write (W), or execute (X) permission is enabled at the user, group, or "other" level. The "other" level includes all users who are not listed as the owner of the share, and are not part of the group level that the file belongs to.

Setting	Setting value
Create Permissions	Sets the default source permissions to apply when a file or directory is created. The default value is `Default ACL`.
Create Mask (Dir)	Specifies UNIX mode bits that are removed when a directory is created, restricting permissions. Mask bits are applied before mode bits are applied.
Create Mode (Dir)	Specifies UNIX mode bits that are added when a directory is created, enabling permissions. Mode bits are applied after mask bits are applied.
Create Mask (File)	Specifies UNIX mode bits that are removed when a file is created, restricting permissions. Mask bits are applied before mode bits are applied.
Create Mode (File)	Specifies UNIX mode bits that are added when a file is created, enabling permissions. Mode bits are applied after mask bits are applied.

SMB performance settings

You can view and configure the change notify and oplocks performance settings of an SMB share.

Note

Changes that are made directly to an SMB share override the default settings configured from the **Default SMB Share Settings** tab.

Setting	Setting value
Change Notify	Configures notification of clients when files or directories change. This helps prevent clients from seeing stale content, but requires server resources. The default value is `Norecurse`.
Oplocks	Indicates whether an opportunistic lock (oplock) request is allowed. An oplock allows clients to provide performance improvements by using locally-cached information. The default value is `Yes`.

SMB security settings

You can view and configure the Impersonate Guest, Impersonate User, and NTFS ACL security settings of an SMB share.

Note

Changes that are made directly to an SMB share override the default settings configured from the **Default SMB Share Settings** tab.

Setting	Setting value
Impersonate Guest	Determines guest access to a share. The default value is `Never`.
Impersonate User	Allows all file access to be performed as a specific user. This must be a fully qualified user name. The default value is `No value`.
NTFS ACL	Allows ACLs to be stored and edited from SMB clients. The default value is `Yes`.

Managing SMB shares

You can configure the rules and other settings that govern the interaction between your Windows network and individual SMB shares on the cluster.

OneFS supports %U, %D, %Z, %L, %0, %1, %2, and %3 variable expansion and automatic provisioning of user home directories.

You can configure the users and groups that are associated with an SMB share, and view or modify their share-level permissions.

Note

We recommend that you configure advanced SMB share settings only if you have a solid understanding of the SMB protocol.

Create an SMB share

When you create an SMB share, you can override the default permissions, performance, and access settings. You can configure SMB home directory provisioning by including expansion variables in the share path to automatically create and redirect users to their own home directories.

Before you begin

You must specify a path to use as the SMB share; create the directory before you create an SMB share. Shares are specific to access zones and the share path must exist under the zone path. Create access zones before you create SMB shares.

Procedure

1. Click **Protocols** › **Windows Sharing (SMB)** › **SMB Shares**.
2. From the **Current Access Zones** drop-down list, select the access zone the share will belong to.
3. Click **Add a share**.
4. In the **Share Name** field, type a name for the share.

   Share names can contain up to 80 characters, and can only contain alphanumeric characters, hyphens, and spaces. Also, if the cluster character encoding is not set to UTF-8, SMB share names are case-sensitive.
5. (Optional) In the **Description** field, type a comment about the share.

   A description is optional, but can be helpful if you are managing multiple shares. This field is limited to 255 characters.
6. In the **Directory to be Shared** field, type the full path of the share, beginning with `/ifs`, or click **Browse** to locate the share.

   Note

   You can specify one or more of the following variables in the directory path but you must select the **Allow Variable Expansion** check box or the string is interpreted literally by the system.

Variable	Expansion
%D	NetBIOS domain name.
%U	User name—for example, **`user_001`**.
%Z	Zone name—for example, **`System`**.
%L	Host name of the cluster, normalized to lowercase.
%0	First character of the user name.
%1	Second character of the user name.
%2	Third character of the user name.

   For example, if a user is in a domain named DOMAIN and has a username of user_1, the path `/ifs/home/%D/%U` expands to `/ifs/home/DOMAIN/user_1`.
7. Apply the initial ACL settings for the directory. You can modify these settings later.

- To apply a default ACL to the shared directory, click **Apply Windows default ACLs**.

  Note

  If the **Auto-Create Directories** setting is enabled, OneFS creates an ACL with the equivalent of UNIX 700 mode bit permissions for any directory that is created automatically.

- To maintain the existing permissions on the shared directory, click **Do not change existing permissions**.

8. (Optional) Configure home directory provisioning settings.
   - To expand path variables such as %U in the share directory path, select **Allow Variable Expansion**.
   - To automatically create home directories when users access the share for the first time, select **Auto-Create Directories**. This option is available only if the **Allow Variable Expansion** option is enabled.
9. (Optional) Apply advanced SMB share settings if needed.
10. Click **Create**.

After you finish

The default permissions configuration is read-only access for the well-known Everyone account. Modify the Users and Groups section settings to allow users to write to the share.

Modify SMB share permissions, performance, or security

You can modify the permissions, performance, and access settings for individual SMB shares.

You can configure SMB home directory provisioning by using directory path, or expansion, variables to automatically create and redirect users to their own home directories.

Note

Any changes made to these settings will only affect the settings for this share. If you need to make changes to default SMB share values, that can be done from the **Default SMB Share Settings** tab.

Procedure

1. Click **Protocols** › **Windows Sharing (SMB)** › **SMB Shares**.
2. From the **Current Access Zones** drop-down list, select the access zone that contains the share you want to modify.
3. From the list of SMB shares, locate the share you want to modify and then click **View details**.
4. For each setting that you want to modify, click **Edit**, make the change, and then click **Save**.
5. To modify the settings for file and directory permissions, performance, or security, click **Advanced SMB Share Settings**.
6. Click **Save**.

Delete an SMB share

You can delete SMB shares that are no longer needed.

Unused SMB shares do not hinder cluster performance. If you delete an SMB share, the share path is deleted but the directory it referenced still exists. If you create a new share with the same path as the share that was deleted, the directory that the previous share referenced will be accessible again through the new share.

Procedure

1. Click **Protocols** › **Windows Sharing (SMB)** › **SMB Shares**.
2. From the **Current Access Zones** drop-down list, select the access zone that contains the share that you want to delete.
3. From the list of SMB shares, select the share that you want to delete.

   Note

   You can delete all of the shares on the cluster by selecting the **Name/Path** option, and then selecting **Delete** from the drop-down menu.

4. Click **Delete**.
5. In the confirmation dialog box, click **Delete** to confirm the deletion.

Limit access to /ifs share for the Everyone account

By default, the `/ifs` root directory is configured as an SMB share in the System access zone. We recommend that you restrict the Everyone account of this share to read-only access.

Procedure

1. Click **Protocols** › **Windows Sharing (SMB)** › **SMB Shares**.
2. From the **Current Access Zone** drop-down list, select `System`.
3. Click **View details** for the `/ifs` share.
4. Click **Edit** next to the Users & Groups option.
5. In the **User/Group Accounts** list, click **Edit** next to the Everyone account.
6. Click **Specify Permission Level** and then select the `Read` option.

   The `Full Control` and `Read-Write` options should not be selected.

7. Click **Save**.

Configure anonymous access to a single SMB share

You can configure anonymous access to data stored on a single share through Guest user impersonation.

Procedure

1. Click **Access** › **Membership & Roles** › **Users**.
2. From the **Current Access Zone** list, select the access zone that contains the share you want to allow anonymous access.
3. From the **Users** list, select **Local**.
   a. Click **View details** for the Guest account.

   b. Next to the **Enable Account** area, click **Edit**.

   c. Select the **Enable Account** check box, and then click **Save**.

4. Click **Protocols** › **Windows Sharing (SMB)** › **SMB Shares**.
5. Click **View details** next to the share you to allow anonymous access.
6. Click **Advanced SMB Share Settings**, and then click **Security Settings**.
7. From the **Impersonate Guest** drop-down list, select **Use Custom**.
8. At the **Confirm Default Override** dialog box, click **Continue**.

   A second drop-down list is displayed.
9. Select **Always** from the drop-down list, and then click **Save**.

Configure anonymous access to all SMB shares in an access zone

You can configure anonymous access to data stored in an access zone through Guest user impersonation.

Procedure

1. Click **Access** › **Membership & Roles** › **Users**.
2. From the **Current Access Zone** list, select the access zone you want to allow anonymous access.
3. From the **Users** list, select **Local**.

   a. Click **View details** for the Guest account.

   b. Next to the **Enable Account** area, click **Edit**.

   c. Select the **Enable Account** check box, and then click **Save**.
4. Click **Protocols** › **Windows Sharing (SMB)** › **Default Share Settings**.
5. From the **Current Access Zone** list, select the access zone you want to allow anonymous access.
6. Click **Show**, and then click **Security Settings**.
7. From the **Impersonate Guest** drop-down list, select **Use Custom**.
8. At the **Confirm Default Override** dialog box, click **Continue**.

   A second drop-down list is displayed.
9. Select **Always** from the drop-down list, and then click **Save**.

Add a user or group to an SMB share

For each SMB share, you can add share-level permissions for specific users and groups.

Procedure

1. Click **Protocols** › **Windows Sharing (SMB)** › **SMB Shares**.
2. From the **Current Access Zones** drop-down list, select the access zone that contains the share you want to add a user or group to.
3. From the list of SMB shares, locate the share you want to modify and then click **View details**.
4. Click **Edit** next to the Users & Groups option.

   The User/Group permission list for the share appears.
5. Click **Add a User or Group**. Then select the option you want to search for.

- Users
- Groups
- Well-known SIDs

6. If you selected User or Group, you can locate the user or group through one of the following methods:
   - Type the Username or Group Name you want to search for in the text field, and then click **Search**.
   - Select the authentication provider you want to search in the text field, and then click **Search**. Only providers that are currently configured and enabled on the cluster are listed.
   - Type the Username or Group Name and select an authentication provider and click **Search**.
7. If you selected Well-known SIDs, click **Search**.
8. In the search results, click the user, group, or SID that you want to add to the SMB share and then click **Select**.
9. By default, the access rights of the new account are set to `Deny All`. To enable a user or group to access the share, follow these additional steps:
   a. Next to the user or group account you added, click **Edit**.
   b. Select **Run as Root** or select **Specify Permission Level** and then select one or more of the following permission levels: `Full Control`, `Read-Write`, and `Read`.
10. Click **Save**.

Configure multi-protocol home directory access

For users who will access this share through FTP or SSH, you can make sure that their home directory path is the same whether they connect through SMB or they log in through FTP or SSH.

This command directs the SMB share to use the home directory template that is specified in the user's authentication provider. This procedure is available only through the command-line interface.

Procedure

1. Establish an SSH connection to any node in the cluster.
2. Run the following command, where *<homedir_share>* is the name of the SMB share:

```
isi smb share modify <homedir_share> --path=""
```

NFS

OneFS provides an NFS server so you can share files on your cluster with NFS clients that adhere to the RFC1813 (NFSv3) and RFC3530 (NFSv4) specifications.

In OneFS, the NFS server is fully optimized as a multi-threaded service running in user space instead of the kernel. This architecture load balances the NFS service across all nodes of the cluster, providing the stability and scalability necessary to manage up to thousands of connections across multiple NFS clients.

NFS mounts execute and refresh quickly, and the server constantly monitors fluctuating demands on NFS services and makes adjustments across all nodes to ensure continuous, reliable performance. Using a built-in process scheduler, OneFS helps ensure fair allocation of node resources so that no client can seize more than its fair share of NFS services.

The NFS server also supports access zones defined in OneFS, so that clients can access only the exports appropriate to their zone. For example, if NFS exports are specified for Zone 2, only clients assigned to Zone 2 can access these exports.

To simplify client connections, especially for exports with large path names, the NFS server also supports aliases, which are shortcuts to mount points that clients can specify directly.

For secure NFS file sharing, OneFS supports NIS and LDAP authentication providers.

NFS exports

You can manage individual NFS export rules that define mount-points (paths) available to NFS clients and how the server should perform with these clients.

In OneFS, you can create, delete, list, view, modify, and reload NFS exports.

NFS export rules are zone-aware. Each export is associated with a zone, can only be mounted by clients on that zone, and can only expose paths below the zone root. By default, any export command applies to the client's current zone.

Each rule must have at least one path (mount-point), and can include additional paths. You can also specify that all subdirectories of the given path or paths are mountable. Otherwise, only the specified paths are exported, and child directories are not mountable.

An export rule can specify a particular set of clients, enabling you to restrict access to certain mount-points or to apply a unique set of options to these clients. If the rule does not specify any clients, then the rule applies to all clients that connect to the server. If the rule does specify clients, then that rule is applied only to those clients.

NFS aliases

You can create and manage aliases as shortcuts for directory path names in OneFS. If those path names are defined as NFS exports, NFS clients can specify the aliases as NFS mount points.

NFS aliases are designed to give functional parity with SMB share names within the context of NFS. Each alias maps a unique name to a path on the file system. NFS clients can then use the alias name in place of the path when mounting.

Aliases must be formed as top-level Unix path names, having a single forward slash followed by name. For example, you could create an alias named `/q4` that maps to `/ifs/data/finance/accounting/winter2015` (a path in OneFS). An NFS client could mount that directory through either of:

```
mount cluster_ip:/q4
```

```
mount cluster_ip:/ifs/data/finance/accounting/winter2015
```

Aliases and exports are completely independent. You can create an alias without associating it with an NFS export. Similarly, an NFS export does not require an alias.

Each alias must point to a valid path on the file system. While this path is absolute, it must point to a location beneath the zone root (`/ifs` on the System zone). If the alias points to a path that does not exist on the file system, any client trying to mount the alias would be denied in the same way as attempting to mount an invalid full pathname.

NFS aliases are zone-aware. By default, an alias applies to the client's current access zone. To change this, you can specify an alternative access zone as part of creating or modifying an alias.

Each alias can only be used by clients on that zone, and can only apply to paths below the zone root. Alias names are unique per zone, but the same name can be used in different zones—for example, `/home`.

When you create an alias in the web administration interface, the alias list displays the status of the alias. Similarly, using the `--check` option of the `isi nfs aliases` command, you can check the status of an NFS alias (status can be: good, illegal path, name conflict, not exported, or path not found).

NFS log files

OneFS writes log messages associated with NFS events to a set of files in `/var/log`.

With the log level option, you can now specify the detail at which log messages are output to log files. The following table describes the log files associated with NFS.

Log file	Description
nfs.log	Primary NFS server functionality (v3, v4, mount)
rpc_lockd.log	NFS v3 locking events through the NLM protocol
rpc_statd.log	NFS v3 reboot detection through the NSM protocol
isi_netgroup_d.log	Netgroup resolution and caching

Managing the NFS service

You can enable or disable the NFS service and specify the NFS versions to support, including NFSv3 and NFSv4. NFS settings are applied across all nodes in the cluster.

Configure NFS file sharing

You can enable or disable the NFS service, and set the lock protection level and security type. These settings are applied across all nodes in the cluster. You can change the settings for individual NFS exports that you define.

Procedure

1. Click **Protocols** › **UNIX Sharing (NFS)** › **Global Settings**.
2. Enable or disable the following settings:
   - NFS Export Service
   - NFSv3
   - NFSv4
3. Select the **Requested Lock Protection Level** setting from the drop-down list.

   Because the NFS service is distributed across all nodes on the cluster, you can select the number of node failures that would be tolerated and still keep the service running. The default setting `[+2]` is the optimal value. We recommend that you not change it.
4. Click the **Reload** button for the **Cached Export Configuration** setting.

   The cached NFS export settings are reloaded to help ensure that changes to DNS or NIS are applied.

5. Click **Save**.

Create a root-squashing rule for the default NFS export

By default, the NFS service implements a root-squashing rule for the default NFS export. This prevents root users on NFS clients from exercising root privileges on the NFS server.

Procedure

1. Click **Protocols** › **UNIX Sharing (NFS)** › **NFS Exports**.
2. For the default export in the NFS Exports list, click **View details**.
3. Verify that **Map Root User** is set to `Use default`. If so, proceed to step 6.
4. If **Map Root User** is set to `Use custom`, restore root user options to these values:

```
User: Map root users to user nobody
Primary Group: No primary group
Secondary Groups: No secondary groups
```

5. Click **Save**.
6. Click **Close**.

Results

With these settings, regardless of the users' credentials on the NFS client, they would not be able to gain root privileges on the NFS server.

NFS global settings

NFS global settings determine how the NFS service operates. You can modify these settings according to your organization's needs.

The following table describes NFS global settings and their default values:

Setting	Description
NFS Export Service	Enables or disables the NFS service. This setting is enabled by default.
NFSv3	Enables or disables support for NFSv3. This setting is enabled by default.
NFSv4	Enables or disables support for NFSv4. This setting is disabled by default.
Requested Lock Protection Level	Determines the number of node failures that can happen before a lock might be lost. The default value is `[+2] Tolerate failure of 2 nodes`.
Cached Export Configuration	Enables you to reload cached NFS exports to help ensure that any domain or network changes take effect immediately.

Managing NFS exports

You can create NFS exports, view and modify export settings, and delete exports that are no longer needed.

The `/ifs` directory is the top-level directory for data storage in OneFS, and is also the path defined in the default export. By default, the `/ifs` export disallows root access, but other enables UNIX clients to mount this directory and any subdirectories beneath it.

Note

We recommend that you modify the default export to limit access only to trusted clients, or to restrict access completely. To help ensure that sensitive data is not compromised, other exports that you create should be lower in the OneFS file hierarchy, and can be protected by access zones or limited to specific clients with either root, read-write, or read-only access, as appropriate.

Create an NFS export

You can create NFS exports to share files in OneFS with UNIX-based clients.

The NFS service runs in user space and distributes the load across all nodes in the cluster. This enables the service to be highly scalable and support thousands of exports. As a best practice, however, you should avoid creating a separate export for each client on your network. It is more efficient to create fewer exports, and to use access zones and user mapping to control access.

Procedure

1. Click **Protocols** › **UNIX Sharing (NFS)** › **NFS Exports**.
2. Click **+ Add an NFS Export**.
3. (Optional) In the **Description** field, type a comment that describes the export.
4. (Optional) Specify the NFS clients that are allowed to access the export.

   You can specify NFS clients in any or all of the client fields, as described in the following table. A client can be identified by host name, IPv4 or IPv6 address, subnet, or netgroup. IPv4 addresses mapped into the IPv6 address space are translated and stored as IPv4 addresses to remove any possible ambiguities.
   You can specify multiple clients in each field by typing one entry per line.

   Note

   If you do not specify any clients, all clients on your network are allowed access to the export. If you specify clients in any of the rule fields, such as **Always Read-Only Clients**, the applicable rule is only applied to those clients.
   If you add the same client to more than one list and the client is entered in the same format for each entry, the client is normalized to a single list in the following order of priority:

   - Root Clients
   - Always Read-Write Clients
   - Always Read-Only Clients
   - Clients

Setting	Description
Clients	Specifies one or more clients to be allowed access to the export. Access level is controlled through export permissions.
Always Read-Write Clients	Specifies one or more clients to be allowed read/write access to the export regardless of the export's access-restriction setting. This is equivalent to adding a client to the **Clients** list with the **Restrict access to read-only** setting cleared.

Setting	Description
Always Read-Only Clients	Specifies one or more clients to be allowed read-only access to the export regardless of the export's access-restriction setting. This is equivalent to adding a client to the **Clients** list with the **Restrict access to read-only** setting selected.
Root Clients	Specifies one or more clients to be mapped as root for the export. This setting is equivalent to adding a client to the **Clients** list and mapping root users to the root user name.

5. For the **Directory Paths** setting, type or browse to the directory that you want to export.

   You can add multiple directory paths by clicking **Add another directory path** for each additional path.

6. Specify export permissions:
   - Restrict actions to read-only.
   - Enable mount access to subdirectories. Allow subdirectories below the path(s) to be mounted.

7. Specify User/Group mapping.

   If you select the **Use custom** option, you can limit access by mapping root users or all users to a specific user and group ID. For root squash, map root users to the user name `nobody`.

8. Specify Security Type(s).

   If you select the **Use custom** option, you can select one or more of the following security types:

   - UNIX (system)
   - Kerberos5
   - Kerberos5 Integrity
   - Kerberos5 Privacy

9. Configure Advanced NFS Export Settings.

   We recommend that you do not change advanced settings unless it is necessary and you fully understand the consequences of these changes.

10. Click **Save**.

Results

The new NFS export is created and shown at the top of the **NFS Exports** list.

Modify an NFS export

You can modify the settings for an existing NFS export. In some cases, modifying an NFS export could invalidate existing NFS client connections.

Procedure

1. Select **Protocols** › **UNIX Sharing (NFS)** › **NFS Exports**.
2. In the **NFS Exports** list, click **View details** for the export that you want to modify.
3. For each setting that you want to modify, click **Edit**, make the change, and then click **Save**.
4. Click **Close**.

5. Click **Hide details** to redisplay the **NFS Exports** list.

Delete an NFS export

You can delete unneeded NFS exports. Any current NFS client connections to these exports become invalid.

Note

You can delete all the exports on a cluster at once. Click the **Export ID/Path** check box at the top of the **NFS Exports** list, and then select **Delete** from the drop-down list to the right.

Procedure

1. Select **Protocols › UNIX Sharing (NFS) › NFS Exports.**
2. In the **NFS Exports** list, click the check box to the left of the export that you want to delete.
3. Click **Delete.**
4. In the confirmation dialog box, click **Delete** to confirm the operation.

Check NFS exports for errors

You can check for errors in NFS exports, such as conflicting export rules, invalid paths, and unresolvable hostnames and netgroups.

This procedure is available only through the CLI.

Procedure

1. Run the `isi nfs exports check` command.

   In the following example output, no errors were found:

   ```
   ID Message
   ----------
   ----------
   Total: 0
   ```

   In the following example output, export 1 contains a directory path that does not currently exist:

   ```
   ID   Message
   -----------------------------------
   1    '/ifs/test' does not exist
   -----------------------------------
   Total: 1
   ```

View and configure default NFS export settings

You can view and configure default NFS export settings. All new exports and any existing exports using default values are affected by changes to the default settings.

Note

Changes to default export settings affect all current and future NFS exports that use default settings, and, if specified incorrectly, could impact the availability of the NFS file sharing service. We recommend that you not make changes to default settings, particularly advanced settings, unless you have experience working with NFS. Instead, you should change settings as needed for individual NFS exports as you create them.

Procedure

1. Select **Protocols** › **UNIX Sharing (NFS)** › **Export Settings**.

   Four common NFS export settings are displayed: **Map Root User**, **Map Non Root User**, **Map Failure**, and **Security Types**. In addition, a link to advanced default export settings is provided.

2. Modify any of default settings that you want to apply to all new NFS exports, or to existing exports that use any of the default values.
3. When you are finished modifying settings, click **Save**.

Basic NFS export settings

The basic NFS export settings are global settings that apply to any new NFS exports that you create.

The basic NFS export settings are described in the following table.

Setting	Default values
Map Root User	User: Map root users to user `nobody` Primary Group: No primary group Secondary Groups: No secondary groups **Note** The default settings result in a root squashing rule whereby no user on the NFS client, even a root user, can gain root privileges on the NFS server.
Map Non Root User	User mapping is disabled by default. We recommend that you specify this setting on a per-export basis, when appropriate.
Map Failure	User mapping is disabled by default. We recommend that you specify this setting on a per-export basis, when appropriate.
Security Types	System default. Available options include `UNIX (sys)`, `Kerberos5`, `Kerberos5 Integrity`, and `Kerberos5 Privacy`.

NFS export performance settings

You can specify settings to control the performance of NFS exports.

The following table describes the performance category of settings for NFS exports:

Setting	Description
Block Size	The block size used to calculate block counts for NFSv3 `FSSTAT` and NFSv4 `GETATTR` requests. The default value is `8192 bytes`.
Commit Asynchronous	If set to yes, allows NFSv3 and NFSv4 COMMIT operations to be asynchronous. The default value is `No`.
Directory Transfer Size	The preferred directory read transfer size reported to NFSv3 and NFSv4 clients. The default value is `131072 bytes`.
Read Transfer Max Size	The maximum read transfer size reported to NFSv3 and NFSv4 clients. The default value is `1048576 bytes`.

Setting	Description
Read Transfer Multiple	The recommended read transfer size multiple reported to NFSv3 and NFSv4 clients. The default value is `512 bytes`.
Read Transfer Size	The preferred read transfer size reported to NFSv3 and NFSv4 clients. The default value is `131072 bytes`.
Setattr Asynchronous	If set to `Yes`, performs set attribute operations asynchronously. The default value is `No`.
Write Datasync Action	The action to perform for DATASYNC writes. The default value is `DATASYNC`.
Write Datasync Reply	The reply to send for DATASYNC writes. The default value is `DATASYNC`.
Write Filesync Action	The action to perform for FILESYNC writes. The default value is `FILESYNC`.
Write Filesync Reply	The reply to send for FILESYNC writes. The default value is `FILESYNC`.
Write Transfer Max Size	The maximum write transfer size reported to NFSv3 and NFSv4 clients. The default value is `1048576 bytes`.
Write Transfer Multiple	The recommended write transfer size reported to NFSv3 and NFSv4 clients. The default value is `512 bytes`.
Write Transfer Size	The preferred write transfer size reported to NFSv3 and NFSv4 clients. The default value is `524288`.
Write Unstable Action	The action to perform for UNSTABLE writes. The default value is `UNSTABLE`.
Write Unstable Reply	The reply to send for UNSTABLE writes. The default value is `UNSTABLE`.

NFS export client compatibility settings

The NFS export client compatibility settings affect the customization of NFS exports.

These settings are described in the following table.

Setting	Setting value
Max File Size	Specifies the maximum file size to allow. This setting is advisory in nature and is returned to the client in a reply to an NFSv3 FSINFO or NFSv4 GETATTR request. The default value is `9223372036854776000 bytes`.
Readdirplus Enable	Enables the use of NFSv3 readdirplus service whereby a client can send a request and received extended information about the directory and files in the export. The default is `Yes`.
Return 32 bit File IDs	Specifies return 32-bit file IDs to the client. The default is `No`.

NFS export behavior settings

The NFS export behavior settings control whether NFS clients can perform certain functions on the NFS server, such as setting the time.

The NFS export behavior settings are described in the following table.

Setting	Description
Can Set Time	When this setting is enabled, OneFS allows the NFS client to set various time attributes on the NFS server. The default value is `Yes`.
Encoding	Overrides the general encoding settings the cluster has for the export. The default value is `DEFAULT`.
Map Lookup UID	Looks up incoming user identifiers (UIDs) in the local authentication database. The default value is `No`.
Symlinks	Informs the NFS client that the file system supports symbolic link file types. The default value is `Yes`.
Time Delta	Sets the server clock granularity. The default value is `1e-9 seconds` (0.000000001 second).

Managing NFS aliases

You can create NFS aliases to simplify exports that clients connect to. An NFS alias maps an absolute directory path to a simple directory path.

For example, suppose you created an NFS export to `/ifs/data/hq/home/archive/first-quarter/finance`. You could create the alias `/finance1` to map to that directory path.

NFS aliases can be created in any access zone, including the System zone.

Create an NFS alias

You can create an NFS alias to map a long directory path to a simple pathname.

Aliases must be formed as a simple Unix-style directory path, for example, `/home`.

Procedure

1. Select **Protocols › UNIX Sharing (NFS)**.

   The **UNIX Sharing (NFS)** page is displayed with four tabs across the top.

2. Click **NFS Aliases**.

   The **NFS Aliases** list appears, displaying at minimum the current zone's NFS root alias. Any other aliases previously created for the current zone are also listed.

3. Click **Create an NFS Alias**.
4. In the **Create an NFS Alias** dialog box, enter a name for the alias.

   The alias name must be formed as a simple UNIX-style path with one element, for example, `/home`.

5. Enter the full path that the alias is to be associated with.

   If you have set up access zones in OneFS, the full path must begin with the root of the current access zone.

6. Click **Create Alias**.

   The **Create an NFS Alias** dialog box closes.

Results

The name, status, and path of the new alias are shown at the top of the **NFS Aliases** list.

Modify an NFS alias

You can modify an NFS alias.

Procedure

1. Select **Protocols › UNIX Sharing (NFS)**.

   The **UNIX Sharing (NFS)** page is displayed with four tabs across the top.

2. Click **NFS Aliases**.

   The **NFS Aliases** list appears, displaying all aliases for the current access zone.
3. Next to the alias that you intend to modify, click **View/Edit**.
4. In the **View NFS Alias Details** dialog box, click **Edit Alias**.
5. Modify either or both the alias name and the path that the alias represents.

   The alias name must be formed as a simple UNIX-style path with one element, for example, `/home`. The path must begin with the root of the current access zone. If you are unsure of the pathname, you can click **Browse** to navigate to the pathname in OneFS.

6. When you are finished modifying the alias, click **Save Changes**.

   The **View NFS Alias Details** dialog box is displayed, and a message at the top of the dialog box indicates the operation was successful.
7. Click **Close**.

   The **View NFS Alias Details** dialog box closes.

Results

The modified alias name, status, and path are shown in the **NFS Aliases** list.

Delete an NFS alias

You can delete an NFS alias.

If an NFS alias is mapped to an NFS export, deleting the alias can disconnect clients that used the alias to mount the export.

Procedure

1. Select **Protocols › UNIX Sharing (NFS)**.

   The **UNIX Sharing (NFS)** page is displayed with four tabs across the top.

2. Click **NFS Aliases**.

   The **NFS Aliases** list appears, displaying all aliases created for the current access zone.
3. Next to the alias that you intend to delete, select **More› Delete Alias**.

   A **Confirm Delete** dialog box appears.
4. Click **Delete**.

   The alias is removed from the **NFS Aliases** list.

List NFS aliases

You can view a list of NFS aliases that have already been defined.

Procedure

1. Select **Protocols** › **UNIX Sharing (NFS)**.

   The **UNIX Sharing (NFS)** page is displayed with four tabs across the top.
2. Click **NFS Aliases**.

   The **NFS Aliases** list appears, displaying all aliases for the current access zone. The names, states, and paths for all aliases are shown.

View an NFS alias

You can view the settings of an NFS alias.

Procedure

1. Select **Protocols** › **UNIX Sharing (NFS)**.

   The **UNIX Sharing (NFS)** page is displayed with four tabs across the top.
2. Click **NFS Aliases**.

   The **NFS Aliases** list appears, displaying all aliases for the current access zone.
3. Next to the alias that you want to view, click **View/Edit**.

   The **View NFS Alias Details** dialog box displays the settings associated with the alias.
4. When you are done viewing the alias, click **Close**.

FTP

OneFS includes a secure FTP service called vsftpd, which stands for Very Secure FTP Daemon, that you can configure for standard FTP and FTPS file transfers.

Enable and configure FTP file sharing

You can set the FTP service to allow any node in the cluster to respond to FTP requests through a standard user account.

You can enable the transfer of files between remote FTP servers and enable anonymous FTP service on the root by creating a local user named anonymous or ftp.

When configuring FTP access, make sure that the specified FTP root is the home directory of the user who logs in. For example, the FTP root for local user jsmith should be `ifs/home/jsmith`.

Procedure

1. Click **Protocols** › **FTP Settings**.
2. Click **Enable**.
3. Select one or more of the following settings:

Option	Description
`Server-to-server transfers`	Enables the transfer of files between two remote FTP servers. This setting is disabled by default.
`Anonymous access`	Enables users with "anonymous" or "ftp" as the user name to access files and directories without requiring authentication. This setting is disabled by default.
`Local access`	Enables local users to access files and directories with their local user name and password, allowing them to upload files directly through the file system. This setting is enabled by default.

4. Click **Submit**.

HTTP and HTTPS

OneFS includes a configurable HTTP service, which is used to request files that are stored on the cluster and to interact with the web administration interface.

OneFS supports both HTTP and its secure variant, HTTPS. Each node in the cluster runs an instance of the Apache HTTP Server to provide HTTP access. You can configure the HTTP service to run in different modes.

Both HTTP and HTTPS are supported for file transfer, but only HTTPS is supported for Platform API calls. The HTTPS-only requirement includes the web administration interface. In addition, OneFS supports a form of the web-based DAV (WebDAV) protocol that enables users to modify and manage files on remote web servers. OneFS performs distributed authoring, but does not support versioning and does not perform security checks. You can enable DAV in the web administration interface.

Enable and configure HTTP

You can configure HTTP and DAV to enable users to edit and manage files collaboratively across remote web servers.

This procedure is available only through the web administration interface.

Procedure

1. Click **Protocols** › **HTTP Settings**.
2. From the **Service** options, select one of the following settings:

Option	Description
Enable HTTP	Allows HTTP access for cluster administration and browsing content on the cluster.
Disable HTTP and redirect to the web interface	Allows only administrative access to the web administration interface. This is the default setting.
Disable HTTP entirely	Closes the HTTP port used for file access. Users can continue to access the web administration interface by specifying the port number in the URL. The default port is 8080.

3. In the **Document root directory** field, type or click **Browse** to navigate to an existing directory in `/ifs`, or click **File System Explorer** to create a new directory and set its permissions.

   Note

   The HTTP server runs as the daemon user and group. To properly enforce access controls, you must grant the daemon user or group read access to all files under the document root, and allow the HTTP server to traverse the document root.

4. In the **Server hostname** field, type the HTTP server name. The server hostname must be a fully-qualified, SmartConnect zone name and valid DNS name. The name must begin with a letter and contain only letters, numbers, and hyphens (-).
5. In the **Administrator email address** field, type an email address to display as the primary contact for issues that occur while serving files.
6. From the **Authentication** list, select an authentication setting:

Option	Description
Off	Disables HTTP authentication.
Basic Authentication Only	Enables HTTP basic authentication. User credentials are sent in plain text.
Integrated Authentication Only	Enables HTTP authentication via NTLM, Kerberos, or both.
Integrated and Basic Authentication	Enables both basic and integrated authentication.
Basic Authentication with Access Controls	Enables HTTP basic authentication and enables the Apache web server to perform access checks.
Integrated Authentication with Access Controls	Enables HTTP integrated authentication via NTLM and Kerberos, and enables the Apache web server to perform access checks.
Integrated and Basic Auth with Access Controls	Enables HTTP basic authentication and integrated authentication, and enables the Apache web server to perform access checks.

7. Click the **Enable DAV** check box. This allows multiple users to manage and modify files collaboratively across remote web servers.
8. Click the **Disable access logging** check box.
9. Click **Submit**.

CHAPTER 9

Home directories

This section contains the following topics:

Home directories overview

When you create a local user, OneFS automatically creates a home directory for the user. OneFS also supports dynamic home directory provisioning for users who access the cluster by connecting to an SMB share or by logging in through FTP or SSH. Regardless of the method by which a home directory was created, you can configure access to the home directory through a combination of SMB, SSH, and FTP.

Home directory permissions

You can set up a user's home directory with a Windows ACL or with POSIX mode bits, which are then converted into a synthetic ACL. The method by which a home directory is created determines the initial permissions that are set on the home directory.

When you create a local user, the user's home directory is created with mode bits by default.

For users who authenticate against external sources, you can specify settings to create home directories dynamically at login time. If a home directory is created during a login through SSH or FTP, it is set up with mode bits; if a home directory is created during an SMB connection, it receives either mode bits or an ACL. For example, if an LDAP user first logs in through SSH or FTP, the user's home directory is created with mode bits. If the same user first connects through an SMB share, the home directory is created with the permissions indicated by the configured SMB settings. If the `--inheritable-path-acl` option is enabled, an ACL is generated; otherwise, mode bits are used.

Authenticating SMB users

You can authenticate SMB users from authentication providers that can handle NT hashes.

SMB sends an NT password hash to authenticate SMB users, so only users from authentication providers that can handle NT hashes can log in over SMB. The following OneFS-supported authentication providers can handle NT hashes:

- Active Directory
- Local
- LDAPSAM (LDAP with Samba extensions enabled)

Home directory creation through SMB

You can create SMB shares by including expansion variables in the share path. Expansion variables give users to access their home directories by connecting to the share. You can also enable dynamic provisioning of home directories that do not exist at SMB connection time.

Note

Share permissions are checked when files are accessed, before the underlying file system permissions are checked. Either of these permissions can prevent access to the file or directory.

Create home directories with expansion variables

You can configure settings with expansion variables to create SMB share home directories.

When users access the EMC Isilon cluster over SMB, home directory access is through SMB shares. You can configure settings with a path that uses a variable expansion syntax, allowing a user to connect to their home directory share.

Note

Home directory share paths must begin with `/ifs/` and must be in the root path of the access zone in which the home directory SMB share is created.

In the following commands, the `--allow-variable-expansion` option is enabled to indicate that %U should be expanded to the user name, which is user411 in this example. The `--auto-create-directory` option is enabled to create the directory if it does not exist:

```
isi smb shares create HOMEDIR --path=/ifs/home/%U \
  --allow-variable-expansion=yes --auto-create-directory=yes
isi smb shares permission modify HOMEDIR --wellknown Everyone \
   --permission-type allow --permission full
isi smb shares view HOMEDIR
```

The system displays output similar to the following example:

```
                                     Share Name: HOMEDIR
                                           Path: /ifs/home/%U
                                    Description:
                     Client-side Caching Policy: manual
Automatically expand user names or domain names: True
Automatically create home directories for users: True
                                      Browsable: True
Permissions:
Account  Account Type Run as Root Permission Type Permission
------------------------------------------------------------
Everyone wellknown    False       allow           full
------------------------------------------------------------
Total: 1
...
```

When user411 connects to the share with the `net use` command, the user's home directory is created at `/ifs/home/user411`. On user411's Windows client, the `net use m:` command connects `/ifs/home/user411` through the HOMEDIR share:

```
net use m: \\cluster.company.com\HOMEDIR /u:user411
```

Procedure

1. Run the following commands on the cluster with the `--allow-variable-expansion` option enabled. The %U expansion variable expands to the user name, and the `--auto-create-directory` option is enabled to create the directory if it does not exist:

   ```
   isi smb shares create HOMEDIR --path=/ifs/home/%U \
     --allow-variable-expansion=yes --auto-create-directory=yes
   isi smb shares permission modify HOMEDIR --wellknown Everyone \
      --permission-type allow --permission full
   ```

2. Run the following command to view the home directory settings:

   ```
   isi smb shares view HOMEDIR
   ```

   The system displays output similar to the following example:

   ```
                                        Share Name: HOMEDIR
                                              Path: /ifs/home/%U
                                       Description:
                        Client-side Caching Policy: manual
   Automatically expand user names or domain names: True
   Automatically create home directories for users: True
                                         Browsable: True
   Permissions:
   Account  Account Type Run as Root Permission Type Permission
   ------------------------------------------------------------
   Everyone wellknown    False       allow           full
   ------------------------------------------------------------
   Total: 1
   ...
   ```

   If user411 connects to the share with the `net use` command, user411's home directory is created at `/ifs/home/user411`. On user411's Windows client, the `net use m:` command connects `/ifs/home/user411` through the HOMEDIR share, mapping the connection similar to the following example:

   ```
   net use m: \\cluster.company.com\HOMEDIR /u:user411
   ```

Create home directories with the --inheritable-path-acl option

You can enable the `--inheritable-path-acl` option on a share to specify that it is to be inherited on the share path if the parent directory has an inheritable ACL.

Before you begin

To perform most configuration tasks, you must log on as a member of the SecurityAdmin role.

By default, an SMB share's directory path is created with a synthetic ACL based on mode bits. You can enable the `--inheritable-path-acl` option to use the inheritable ACL on all directories that are created, either at share creation time or for those dynamically provisioned when connecting to that share.

Procedure

1. Run commands similar to the following examples to enable the `--inheritable-path-acl` option on the cluster to dynamically provision a user home directory at first connection to a share on the cluster:

   ```
   isi smb shares create HOMEDIR_ACL --path=/ifs/home/%U \
      --allow-variable-expansion=yes --auto-create-directory=yes \
      --inheritable-path-acl=yes
   ```

   ```
   isi smb shares permission modify HOMEDIR_ACL \
     --wellknown Everyone \
     --permission-type allow --permission full
   ```

2. Run a `net use` command, similar to the following example, on a Windows client to map the home directory for user411:

   ```
   net use q: \\cluster.company.com\HOMEDIR_ACL /u:user411
   ```

3. Run a command similar to the following example on the cluster to view the inherited ACL permissions for the user411 share:

   ```
   cd /ifs/home/user411
   ls -lde .
   ```

   The system displays output similar to the following example:

   ```
   drwx------ +  2 user411 Isilon Users 0 Oct 19 16:23 ./
    OWNER: user:user411
    GROUP: group:Isilon Users
    CONTROL:dacl_auto_inherited,dacl_protected
    0: user:user411 allow dir_gen_all,object_inherit,container_inherit
   ```

Create special home directories with the SMB share %U variable

The special SMB share name %U enables you to create a home-directory SMB share that appears the same as a user's user name.

You typically set up a %U SMB share with a share path that includes the %U expansion variable. If a user attempts to connect to a share matching the login name and it does not exist, the user connects to the %U share instead and is directed to the expanded path for the %U share.

Note

If another SMB share exists that matches the user's name, the user connects to the explicitly named share rather than to the %U share.

Procedure

1. Run the following command to create a share that matches the authenticated user login name when the user connects to the share:

   ```
   isi smb share create %U /ifs/home/%U \
     --allow-variable-expansion=yes --auto-create-directory=yes \
     --zone=System
   ```

   After running this command, user Zachary will see a share named 'zachary' rather than '%U', and when Zachary tries to connect to the share named 'zachary', he will be directed to `/ifs/home/zachary`. On a Windows client, if Zachary runs the following commands, he sees the contents of his `/ifs/home/zachary` directory:

   ```
   net use m: \\cluster.ip\zachary /u:zachary
   cd m:
   dir
   ```

   Similarly, if user Claudia runs the following commands on a Windows client, she sees the directory contents of `/ifs/home/claudia`:

   ```
   net use m: \\cluster.ip\claudia /u:claudia
   cd m:
   dir
   ```

   Zachary and Claudia cannot access one another's home directory because only the share 'zachary' exists for Zachary and only the share 'claudia' exists for Claudia.

Home directory creation through SSH and FTP

You can configure home directory support for users who access the cluster through SSH or FTP by modifying authentication provider settings.

Set the SSH or FTP login shell

You can use the `--login-shell` option to set the default login shell for the user.

By default, the `--login-shell` option, if specified, overrides any login-shell information provided by the authentication provider, except with Active Directory. If the `--login-shell` option is specified with Active Directory, it simply represents the default login shell if the Active Directory server does not provide login-shell information.

Procedure

1. Run the following command to set the login shell for all local users to `/bin/bash`:

   ```
   isi auth local modify System --login-shell /bin/bash
   ```

2. Run the following command to set the default login shell for all Active Directory users in your domain to `/bin/bash`:

   ```
   isi auth ads modify YOUR.DOMAIN.NAME.COM --login-shell /bin/bash
   ```

Set SSH/FTP home directory permissions

You can specify home directory permissions for a home directory that is accessed through SSH or FTP by setting a umask value.

Before you begin

To perform most configuration tasks, you must log on as a member of the SecurityAdmin role.

When a user's home directory is created at login through SSH or FTP, it is created using POSIX mode bits. The permissions setting on a user's home directory is set to 0755, then masked according to the umask setting of the user's access zone to further limit permissions. You can modify the umask setting for a zone with the `--home-directory-umask` option, specifying an octal number as the umask value.

Procedure

1. Run the following command to view umask setting:

   ```
   isi zone zones view System
   ```

   The system displays output similar to the following example:

   ```
                     Name: System
                     Path: /ifs
               Cache Size: 4.77M
            Map Untrusted:
           Auth Providers: -
             NetBIOS Name:
       All Auth Providers: Yes
   ```

```
    User Mapping Rules: -
  Home Directory Umask: 0077
    Skeleton Directory: /usr/share/skel
         Audit Success: create, delete, rename, set_security, close
         Audit Failure: create, delete, rename, set_security, close
   HDFS Authentication: all
           HDFS Keytab: /etc/hdfs.keytab
   HDFS Root Directory: /ifs
       WebHDFS Enabled: Yes
Syslog Forwarding Enabled: No
   Syslog Audit Events: create, delete, rename, set_security
               Zone ID: 1
```

In the command result, you can see the default setting for `Home Directory Umask` for the created home directory is `0700`, which is equivalent to (`0755` & ~(`077`)). You can modify the `Home Directory Umask` setting for a zone with the `--home-directory-umask` option, specifying an octal number as the umask value. This value indicates the permissions that are to be disabled, so larger mask values indicate fewer permissions. For example, a umask value of `000` or `022` yields created home directory permissions of `0755`, whereas a umask value of `077` yields created home directory permissions of `0700`.

2. Run a command similar to the following example to allow a group/others write/ execute permission in a home directory:

```
isi zone zones modify System --home-directory-umask=022
```

In this example, user home directories will be created with mode bits `0755` masked by the umask field, set to the value of `022`. Therefore, user home directories will be created with mode bits `0755`, which is equivalent to (`0755` & ~(`022`)).

Set SSH/FTP home directory creation options

You can configure home directory support for a user who accesses the cluster through SSH or FTP by specifying authentication provider options.

Procedure

1. Run the following command to view settings for an Active Directory authentication provider on the cluster:

```
isi auth ads list
```

The system displays output similar to the following example:

```
Name                 Authentication Status DC Name Site
---------------------------------------------------------
YOUR.DOMAIN.NAME.COM Yes            online -       SEA
---------------------------------------------------------
Total: 1
```

2. Run the `isi auth ads modify` command with the `--home-directory-template` and `--create-home-directory` options.

```
isi auth ads modify YOUR.DOMAIN.NAME.COM \
--home-directory-template=/ifs/home/ADS/%D/%U \
--create-home-directory=yes
```

3. Run the `isi auth ads view` command with the `--verbose` option.

The system displays output similar to the following example:

```
                    Name: YOUR.DOMAIN.NAME.COM
          NetBIOS Domain: YOUR
          ...
   Create Home Directory: Yes
 Home Directory Template: /ifs/home/ADS/%D/%U
             Login Shell: /bin/sh
```

4. Run the `id` command.

   The system displays output similar to the following example:

   ```
   uid=1000008(<your-domain>\user_100) gid=1000000(<your-domain>
   \domain users)
   groups=1000000(<your-domain>\domain users),1000024(<your-domain>
   \c1t),1545(Users)
   ```

5. (Optional) To verify this information from an external UNIX node, run the `ssh` command from an external UNIX node.

   For example, the following command would create `/ifs/home/ADS/`*`<your-domain>`*`/user_100` if it did not previously exist:

   ```
   ssh <your-domain>\\user_100@cluster.isilon.com
   ```

Provision home directories with dot files

You can provision home directories with dot files.

Before you begin

To perform most configuration tasks, you must log on as a member of the SecurityAdmin role.

The skeleton directory, which is located at `/usr/share/skel` by default, contains a set of files that are copied to the user's home directory when a local user is created or when a user home directory is dynamically created during login. Files in the skeleton directory that begin with `dot.` are renamed to remove the `dot` prefix when they are copied to the user's home directory. For example, `dot.cshrc` is copied to the user's home directory as `.cshrc`. This format enables dot files in the skeleton directory to be viewable through the command-line interface without requiring the `ls -a` command.
For SMB shares that might use home directories that were provisioned with dot files, you can set an option to prevent users who connect to the share through SMB from viewing the dot files.

Procedure

1. Run the following command to display the default skeleton directory in the System access zone:

   ```
   isi zone zones view System
   ```

   The system displays output similar to the following example:

   ```
                  Name: System
   ...
    Skeleton Directory: /usr/share/skel
   ```

2. Run the `isi zone zones modify` command to modify the default skeleton directory.

The following command modifies the default skeleton directory, `/usr/share/skel`, in an access zone, where System is the value for the *<zone>* option and `/usr/share/skel2` is the value for the *<path>* option:

```
isi zone zones modify System --skeleton-directory=/usr/share/skel2
```

Home directory creation in a mixed environment

If a user logs in through both SMB and SSH, it is recommended that you configure home directory settings so the path template is the same for the SMB share and each authentication provider against which the user is authenticating through SSH.

Interactions between ACLs and mode bits

Home directory setup is determined by several factors, including how users authenticate and the options that specify home directory creation.

A user's home directory may be set up with either ACLs or POSIX mode bits, which are converted into a synthetic ACL. The directory of a local user is created when the local user is created, and the directory is set up with POSIX mode bits by default. Directories can be dynamically provisioned at log in for users who authenticate against external sources, and in some cases for users who authenticate against the File provider. In this situation, the user home directory is created according to how the user first logs in.

For example, if an LDAP user first logs in through SSH or FTP and the user home directory is created, it is created with POSIX mode bits. If that same user first connects through an SMB home directory share, the home directory is created as specified by the SMB option settings. If the `--inherited-path-acl` option is enabled, ACLs are generated. Otherwise, POSIX mode bits are used.

Default home directory settings in authentication providers

The default settings that affect how home directories are set up differ, based on the authentication provider that the user authenticates against.

Authentication provider	Home directory	Home directory creation	UNIX login shell
Local	• `--home-directory-template=/ifs/home/%U` • `--create-home-directory=yes` • `--login-shell=/bin/sh`	Enabled	`/bin/sh`
File	• `--home-directory-template=""`	Disabled	None

Authentication provider	Home directory	Home directory creation	UNIX login shell
	• `--create-home-directory=no`		
Active Directory	• `--home-directory-template=/ifs/home/%D/%U` • `--create-home-directory=no` • `--login-shell=/bin/sh` **Note** If available, provider information overrides this value.	Disabled	`/bin/sh`
LDAP	• `--home-directory-template=""` • `--create-home-directory=no`	Disabled	None
NIS	• `--home-directory-template=""` • `--create-home-directory=no`	Disabled	None

Supported expansion variables

You can include expansion variables in an SMB share path or in an authentication provider's home directory template.

OneFS supports the following expansion variables. You can improve performance and reduce the number of shares to be managed when you configure shares with expansion variables. For example, you can include the %U variable for a share rather than create a share for each user. When a %U is included in the name so that each user's path is different, security is still ensured because each user can view and access only his or her home directory.

Note

When you create an SMB share through the web administration interface, you must select the **Allow Variable Expansion** check box or the string is interpreted literally by the system.

Variable	Value	Description
%U	User name (for example, user_001)	Expands to the user name to allow different users to use different home directories. This variable is typically included at the end of the path. For example, for a user named user1, the path `/ifs/home/%U` is mapped to `/ifs/home/user1`.
%D	NetBIOS domain name (for example, YORK for YORK.EAST.EXAMPLE.COM)	Expands to the user's domain name, based on the authentication provider: • For Active Directory users, %D expands to the Active Directory NetBIOS name. • For local users, %D expands to the cluster name in uppercase characters. For example, for a cluster named cluster1, %D expands to CLUSTER1. • For users in the System file provider, %D expands to UNIX_USERS. • For users in other file providers, %D expands to FILE_USERS. • For LDAP users, %D expands to LDAP_USERS. • For NIS users, %D expands to NIS_USERS.
%Z	Zone name (for example, ZoneABC)	Expands to the access zone name. If multiple zones are activated, this variable is useful for differentiating users in separate zones. For example, for a user named user1 in the System zone, the path `/ifs/home/%Z/%U` is mapped to `/ifs/home/System/user1`.
%L	Host name (cluster host name in lowercase)	Expands to the host name of the cluster, normalized to lowercase. Limited use.
%0	First character of the user name	Expands to the first character of the user name.
%1	Second character of the user name	Expands to the second character of the user name.
%2	Third character of the user name	Expands to the third character of the user name.

Note

If the user name includes fewer than three characters, the %0, %1, and %2 variables wrap around. For example, for a user named ab, the variables maps to a, b, and a, respectively. For a user named a, all three variables map to a.

Domain variables in home directory provisioning

You can use domain variables to specify authentication providers when provisioning home directories.

The domain variable (%D) is typically used for Active Directory users, but it has a value set that can be used for other authentication providers. %D expands as described in the following table for the various authentication providers.

Authenticated user	%D expansion
Active Directory user	Active Directory NetBIOS name—for example, YORK for provider YORK.EAST.EXAMPLE.COM.
Local user	The cluster name in all-uppercase characters—for example, if the cluster is named MyCluster, %D expands to MYCLUSTER.
File user	• UNIX_USERS (for System file provider) • FILE_USERS (for all other file providers)
LDAP user	LDAP_USERS (for all LDAP authentication providers)
NIS user	NIS_USERS (for all NIS authentication providers)

CHAPTER 10

Snapshots

This section contains the following topics:

Snapshots overview

A OneFS snapshot is a logical pointer to data that is stored on a cluster at a specific point in time.

A snapshot references a directory on a cluster, including all data stored in the directory and its subdirectories. If the data referenced by a snapshot is modified, the snapshot stores a physical copy of the data that was modified. Snapshots are created according to user specifications or are automatically generated by OneFS to facilitate system operations.

To create and manage snapshots, you must activate a SnapshotIQ license on the cluster. Some applications must generate snapshots to function but do not require you to activate a SnapshotIQ license; by default, these snapshots are automatically deleted when OneFS no longer needs them. However, if you activate a SnapshotIQ license, you can retain these snapshots. You can view snapshots generated by other modules without activating a SnapshotIQ license.

You can identify and locate snapshots by name or ID. A snapshot name is specified by a user and assigned to the virtual directory that contains the snapshot. A snapshot ID is a numerical identifier that OneFS automatically assigns to a snapshot.

Data protection with SnapshotIQ

You can create snapshots to protect data with the SnapshotIQ software module. Snapshots protect data against accidental deletion and modification by enabling you to restore deleted and modified files. To use SnapshotIQ, you must activate a SnapshotIQ license on the cluster.

Snapshots are less costly than backing up your data on a separate physical storage device in terms of both time and storage consumption. The time required to move data to another physical device depends on the amount of data being moved, whereas snapshots are always created almost instantaneously regardless of the amount of data referenced by the snapshot. Also, because snapshots are available locally, end-users can often restore their data without requiring assistance from a system administrator. Snapshots require less space than a remote backup because unaltered data is referenced rather than recreated.

Snapshots do not protect against hardware or file-system issues. Snapshots reference data that is stored on a cluster, so if the data on the cluster becomes unavailable, the snapshots will also be unavailable. Because of this, it is recommended that you back up your data to separate physical devices in addition to creating snapshots.

Snapshot disk-space usage

The amount of disk space that a snapshot consumes depends on both the amount of data stored by the snapshot and the amount of data the snapshot references from other snapshots.

Immediately after OneFS creates a snapshot, the snapshot consumes a negligible amount of disk space. The snapshot does not consume additional disk space unless the data referenced by the snapshot is modified. If the data that a snapshot references is modified, the snapshot stores read-only copies of the original data. A snapshot consumes only the space that is necessary to restore the contents a directory to the state it was in when the snapshot was taken.

To reduce disk-space usage, snapshots that reference the same directory reference each other, with older snapshots referencing newer snapshots. If a file is deleted, and several snapshots reference the file, a single snapshot stores a copy the file, and the other snapshots reference the file from the snapshot that stored the copy. The reported size of a snapshot reflects only the amount of data stored by the snapshot and does not include the amount of data referenced by the snapshot.

Because snapshots do not consume a set amount of storage space, there is no available-space requirement for creating a snapshot. The size of a snapshot grows according to how the data referenced by the snapshot is modified. A cluster cannot contain more than 20,000 snapshots.

Snapshot schedules

You can automatically generate snapshots according to a snapshot schedule.

With snapshot schedules, you can periodically generate snapshots of a directory without having to manually create a snapshot every time. You can also assign an expiration period that determines when SnapshotIQ deletes each automatically generated snapshot.

Snapshot aliases

A snapshot alias is a logical pointer to a snapshot. If you specify an alias for a snapshot schedule, the alias will always point to the most recent snapshot generated by that schedule. Assigning a snapshot alias allows you to quickly identify and access the most recent snapshot generated according to a snapshot schedule.

If you allow clients to access snapshots through an alias, you can reassign the alias to redirect clients to other snapshots. In addition to assigning snapshot aliases to snapshots, you can also assign snapshot aliases to the live version of the file system. This can be useful if clients are accessing snapshots through a snapshot alias, and you want to redirect the clients to the live version of the file system.

File and directory restoration

You can restore the files and directories that are referenced by a snapshot by copying data from the snapshot, cloning a file from the snapshot, or reverting the entire snapshot.

Copying a file from a snapshot duplicates the file, which roughly doubles the amount of storage space consumed. Even if you delete the original file from the non-snapshot directory, the copy of the file remains in the snapshot.

Cloning a file from a snapshot also duplicates the file. However, unlike a copy, which immediately consumes additional space on the cluster, a clone does not consume any additional space on the cluster unless the clone or cloned file is modified.

Reverting a snapshot replaces the contents of a directory with the data stored in the snapshot. Before a snapshot is reverted, SnapshotIQ creates a snapshot of the directory that is being replaced, which enables you to undo the snapshot revert later. Reverting a snapshot can be useful if you want to undo a large number of changes that you made to files and directories. If new files or directories have been created in a directory since a snapshot of the directory was created, those files and directories are deleted when the snapshot is reverted.

Note

If you move a directory, you cannot revert snapshots of the directory that were taken before the directory was moved.

Best practices for creating snapshots

Consider the following snapshot best practices when working with a large number of snapshots.

It is recommended that you do not create more than 1,000 snapshots of a single directory to avoid performance degradation. If you create a snapshot of a root directory, that snapshot counts towards the total number of snapshots for any subdirectories of the root directory. For example, if you create 500 snapshots of `/ifs/data` and 500 snapshots of `/ifs/data/media`, you have created 1000 snapshots of `/ifs/data/media`. Avoid creating snapshots of directories that are already referenced by other snapshots.

It is recommended that you do not create more than 1000 hard links per file in a snapshot to avoid performance degradation. Always attempt to keep directory paths as shallow as possible. The deeper the depth of directories referenced by snapshots, the greater the performance degradation.

Creating snapshots of directories higher on a directory tree will increase the amount of time it takes to modify the data referenced by the snapshot and require more cluster resources to manage the snapshot and the directory. However, creating snapshots of directories lower on directories trees will require more snapshot schedules, which can be difficult to manage. It is recommended that you do not create snapshots of `/ifs` or `/ifs/data`.

You can create up to 20,000 snapshots on a cluster at a time. If you create a large number of snapshots, you might not be able to manage snapshots through the OneFS web administration interface. However, you can manage any number of snapshots through the OneFS command-line interface.

Note

It is recommended that you do not disable the snapshot delete job. Disabling the snapshot delete job prevents unused disk space from being freed and can also cause performance degradation.

If the system clock is set to a time zone other than Coordinated Universal Time (UTC), SnapshotIQ modifies snapshot duration periods to match Daylight Savings Time (DST). Upon entering DST, snapshot durations are increased by an hour to adhere to DST; when exiting DST, snapshot durations are decreased by an hour to adhere to standard time.

Best practices for creating snapshot schedules

Snapshot schedule configurations can be categorized by how they delete snapshots: ordered deletions and unordered deletions.

An ordered deletion is the deletion of the oldest snapshot of a directory. An unordered deletion is the deletion of a snapshot that is not the oldest snapshot of a directory. Unordered deletions take approximately twice as long to complete and consume more cluster resources than ordered deletions. However, unordered deletions can save space by retaining a smaller total number of snapshots.

The benefits of unordered deletions versus ordered deletions depend on how often the data referenced by the snapshots is modified. If the data is modified frequently, unordered deletions will save space. However, if data remains unmodified, unordered deletions will most likely not save space, and it is recommended that you perform ordered deletions to free cluster resources.

To implement ordered deletions, assign the same duration period for all snapshots of a directory. The snapshots can be created by one or multiple snapshot schedules. Always ensure that no more than 1000 snapshots of a directory are created.

To implement unordered snapshot deletions, create several snapshot schedules for a single directory, and then assign different snapshot duration periods for each schedule. Ensure that all snapshots are created at the same time when possible.

The following table describes snapshot schedules that follow snapshot best practices:

Table 14 Snapshot schedule configurations

Deletion type	Snapshot frequency	Snapshot time	Snapshot expiration	Max snapshots retained
Ordered deletion (for mostly static data)	Every hour	Beginning at 12:00 AM Ending at 11:59 AM	1 month	720
Unordered deletion (for frequently modified data)	Every other hour	Beginning at 12:00 AM Ending at 11:59 PM	1 day	27
	Every day	At 12:00 AM	1 week	
	Every week	Saturday at 12:00 AM	1 month	
	Every month	The first Saturday of the month at 12:00 AM	3 months	

File clones

SnapshotIQ enables you to create file clones that share blocks with existing files in order to save space on the cluster. A file clone usually consumes less space and takes less time to create than a file copy. Although you can clone files from snapshots, clones are primarily used internally by OneFS.

The blocks that are shared between a clone and cloned file are contained in a hidden file called a shadow store. Immediately after a clone is created, all data originally contained in the cloned file is transferred to a shadow store. Because both files reference all blocks from the shadow store, the two files consume no more space than the original file; the clone does not take up any additional space on the cluster. However, if the cloned file or clone is modified, the file and clone will share only blocks that are common to both of them, and the modified, unshared blocks will occupy additional space on the cluster.

Over time, the shared blocks contained in the shadow store might become useless if neither the file nor clone references the blocks. The cluster routinely deletes blocks that are no longer needed. You can force the cluster to delete unused blocks at any time by running the ShadowStoreDelete job.

Clones cannot contain alternate data streams (ADS). If you clone a file that contains alternate data streams, the clone will not contain the alternate data streams.

Shadow-store considerations

Shadow stores are hidden files that are referenced by cloned and deduplicated files. Files that reference shadow stores behave differently than other files.

- Reading shadow-store references might be slower than reading data directly. Specifically, reading non-cached shadow-store references is slower than reading non-cached data. Reading cached shadow-store references takes no more time than reading cached data.
- When files that reference shadow stores are replicated to another Isilon cluster or backed up to a Network Data Management Protocol (NDMP) backup device, the shadow stores are not transferred to the target Isilon cluster or backup device. The files are transferred as if they contained the data that they reference from shadow stores. On the target Isilon cluster or backup device, the files consume the same amount of space as if they had not referenced shadow stores.
- When OneFS creates a shadow store, OneFS assigns the shadow store to a storage pool of a file that references the shadow store. If you delete the storage pool that a shadow store resides on, the shadow store is moved to a pool occupied by another file that references the shadow store.
- OneFS does not delete a shadow-store block immediately after the last reference to the block is deleted. Instead, OneFS waits until the ShadowStoreDelete job is run to delete the unreferenced block. If a large number of unreferenced blocks exist on the cluster, OneFS might report a negative deduplication savings until the ShadowStoreDelete job is run.
- Shadow stores are protected at least as much as the most protected file that references it. For example, if one file that references a shadow store resides in a storage pool with +2 protection and another file that references the shadow store resides in a storage pool with +3 protection, the shadow store is protected at +3.
- Quotas account for files that reference shadow stores as if the files contained the data referenced from shadow stores; from the perspective of a quota, shadow-store references do not exist. However, if a quota includes data protection overhead, the quota does not account for the data protection overhead of shadow stores.

Snapshot locks

A snapshot lock prevents a snapshot from being deleted. If a snapshot has one or more locks applied to it, the snapshot cannot be deleted and is referred to as a locked snapshot. If the duration period of a locked snapshot expires, OneFS will not delete the snapshot until all locks on the snapshot have been deleted.

OneFS applies snapshot locks to ensure that snapshots generated by OneFS applications are not deleted prematurely. For this reason, it is recommended that you do not delete snapshot locks or modify the duration period of snapshot locks.

A limited number of locks can be applied to a snapshot at a time. If you create snapshot locks, the limit for a snapshot might be reached, and OneFS could be unable to apply a snapshot lock when necessary. For this reason, it is recommended that you do not create snapshot locks.

Snapshot reserve

The snapshot reserve enables you to set aside a minimum percentage of the cluster storage capacity specifically for snapshots. If specified, all other OneFS operations are unable to access the percentage of cluster capacity that is reserved for snapshots.

Note

The snapshot reserve does not limit the amount of space that snapshots can consume on the cluster. Snapshots can consume a greater percentage of storage capacity specified by the snapshot reserve. It is recommended that you do not specify a snapshot reserve.

SnapshotIQ license functionality

You can create snapshots only if you activate a SnapshotIQ license on a cluster. However, you can view snapshots and snapshot locks that are created for internal use by OneFS without activating a SnapshotIQ license.

The following table describes what snapshot functionality is available depending on whether the SnapshotIQ license is active:

	Inactive	Active
Create snapshots and snapshot schedules	No	Yes
Configure SnapshotIQ settings	No	Yes
View snapshot schedules	Yes	Yes
Delete snapshots	Yes	Yes
Access snapshot data	Yes	Yes
View snapshots	Yes	Yes

If you a SnapshotIQ license becomes inactive, you will no longer be able to create new snapshots, all snapshot schedules will be disabled, and you will not be able to modify snapshots or snapshot settings. However, you will still be able to delete snapshots and access data contained in snapshots.

Creating snapshots with SnapshotIQ

To create snapshots, you must configure the SnapshotIQ licence on the cluster. You can create snapshots either by creating a snapshot schedule or manually generating an individual snapshot.

Manual snapshots are useful if you want to create a snapshot immediately, or at a time that is not specified in a snapshot schedule. For example, if you plan to make changes to your file system, but are unsure of the consequences, you can capture the current state of the file system in a snapshot before you make the change.

Before creating snapshots, consider that reverting a snapshot requires that a SnapRevert domain exist for the directory that is being reverted. If you intend on reverting snapshots for a directory, it is recommended that you create SnapRevert domains for those directories while the directories are empty. Creating a domain for a directory that contains less data takes less time.

Create a SnapRevert domain

Before you can revert a snapshot that contains a directory, you must create a SnapRevert domain for the directory. It is recommended that you create SnapRevert domains for a directory while the directory is empty.

The root path of the SnapRevert domain must be the same root path of the snapshot. For example, a domain with a root path of `/ifs/data/media` cannot be used to revert a snapshot with a root path of `/ifs/data/media/archive`. To revert `/ifs/data/media/archive`, you must create a SnapRevert domain with a root path of `/ifs/data/media/archive`.

Procedure

1. Click **Cluster Management** › **Job Operations** › **Job Types**.
2. In the **Job Types** area, in the **DomainMark** row, from the **Actions** column, select **Start Job**.
3. In the **Domain Root Path** field, type the path of a snapshot root directory.
4. From the **Type of domain** list, select **SnapRevert**.
5. Ensure that the **Delete this domain** check box is cleared.
6. Click **Start Job**.

Create a snapshot schedule

You can create a snapshot schedule to continuously generate snapshots of directories.

Procedure

1. Click **Data Protection** › **SnapshotIQ** › **Snapshot Schedules**.
2. Click **Create a snapshot schedule**.
3. (Optional) In the **Create a Snapshot Schedule** area, in the **Schedule Name** field, type a name for the snapshot schedule.
4. (Optional) In the **Naming pattern for Generated Snapshots** field, type a naming pattern. Each snapshot generated according to this schedule is assigned a name based on the pattern.

   For example, the following naming pattern is valid:

   ```
   WeeklyBackup_%m-%d-%Y_%H:%M
   ```

   The example produces names similar to the following:

   ```
   WeeklyBackup_07-13-2014_14:21
   ```
5. In the **Directory Path** field, specify the directory that you want to be contained in snapshots that are generated according to this schedule.
6. From the **Snapshot Frequency** list, select how often you want to generate snapshots according to the schedule.

Option	Description
Generate snapshots every day, or skip generating snapshots for a specified number of days.	Select **Daily**, and specify how often you want to generate snapshots.
Generate snapshots on specific days of the week, and optionally skip generating snapshots for a specified number of weeks.	Select **Weekly**, and specify how often you want to generate snapshots.
Generate snapshots on specific days of the month, and optionally skip generating snapshots for a specified number of months.	Select **Monthly**, and specify how often you want to generate snapshots.
Generate snapshots on specific days of the year.	Select **Yearly**, and specify how often you want to generate snapshots.

Note

A snapshot schedule cannot span multiple days. For example, you cannot specify to begin generating snapshots at 5:00 PM Monday and end at 5:00 AM Tuesday. To continuously generate snapshots for a period greater than a day, you must create two snapshot schedules. For example, to generate snapshots from 5:00 PM Monday to 5:00 AM Tuesday, create one schedule that generates snapshots from 5:00 PM to 11:59 PM on Monday, and another schedule that generates snapshots from 12:00 AM to 5:00 AM on Tuesday.

7. (Optional) To assign an alternative name to the most recent snapshot generated by the schedule, specify a snapshot alias.
   a. Next to **Create an Alias**, click **Yes**.
   b. To modify the default snapshot alias name, in the **Alias Name** field, type an alternative name for the snapshot.
8. (Optional) To specify a length of time that snapshots generated according to the schedule exist on the cluster before they are automatically deleted by OneFS, specify an expiration period.
   a. Next to **Snapshot Expiration**, click **Snapshots expire**.
   b. Next to **Snapshots expire**, specify how long you want to retain the snapshots generated according to the schedule.
9. Click **Create**.

Create a snapshot

You can create a snapshot of a directory.

Procedure

1. Click **Data Protection › SnapshotIQ › Summary**.
2. Click **Capture a new snapshot**.
3. (Optional) In the **Capture a Snapshot** area, in the **Snapshot Name** field, type a name.
4. In the **Directory Path** field, specify the directory that you want the snapshot to contain.
5. (Optional) To create an alternative name for the snapshot, specify a snapshot alias.

   a. Next to **Create an Alias**, click **Yes**.

   b. To modify the default snapshot alias name, in the **Alias Name** field, type an alternative name for the snapshot.

6. (Optional) To assign a time that OneFS will automatically delete the snapshot, specify an expiration period.

   a. Next to **Snapshot Expiration**, click **Snapshot Expires on**.

   b. In the calendar, specify the day that you want the snapshot to be automatically deleted.

7. Click **Capture**.

Snapshot naming patterns

If you schedule snapshots to be automatically generated, either according to a snapshot schedule or a replication policy, you must assign a snapshot naming pattern that determines how the snapshots are named. Snapshot naming patterns contain variables that include information about how and when the snapshot was created.

The following variables can be included in a snapshot naming pattern:

Variable	Description
`%A`	The day of the week.
`%a`	The abbreviated day of the week. For example, if the snapshot is generated on a Sunday, `%a` is replaced with `Sun`.
`%B`	The name of the month.
`%b`	The abbreviated name of the month. For example, if the snapshot is generated in September, `%b` is replaced with `Sep`.
`%C`	The first two digits of the year. For example, if the snapshot is created in 2014, `%C` is replaced with `20`.
`%c`	The time and day. This variable is equivalent to specifying **`%a %b %e %T %Y`**.
`%d`	The two digit day of the month.
`%e`	The day of the month. A single-digit day is preceded by a blank space.
`%F`	The date. This variable is equivalent to specifying **`%Y-%m-%d`**
`%G`	The year. This variable is equivalent to specifying **`%Y`**. However, if the snapshot is created in a week that has less than four days in the current year, the year that contains the majority of the days of the week is displayed. The first day of the week is calculated as Monday. For example, if a snapshot is created on Sunday, January 1, 2017, `%G` is replaced with `2016`, because only one day of that week is in 2017.
`%g`	The abbreviated year. This variable is equivalent to specifying **`%y`**. However, if the snapshot was created in a week that has less than four days in the current year, the year that contains the majority of the days of the week is displayed. The first day of the week is calculated as Monday. For example, if a

Variable	Description
	snapshot is created on Sunday, January 1, 2017, %g is replaced with 16, because only one day of that week is in 2017.
%H	The hour. The hour is represented on the 24-hour clock. Single-digit hours are preceded by a zero. For example, if a snapshot is created at 1:45 AM, %H is replaced with 01.
%h	The abbreviated name of the month. This variable is equivalent to specifying **%b**.
%I	The hour represented on the 12-hour clock. Single-digit hours are preceded by a zero. For example, if a snapshot is created at 1:45 PM, %I is replaced with 01.
%j	The numeric day of the year. For example, if a snapshot is created on February 1, %j is replaced with 32.
%k	The hour represented on the 24-hour clock. Single-digit hours are preceded by a blank space.
%l	The hour represented on the 12-hour clock. Single-digit hours are preceded by a blank space. For example, if a snapshot is created at 1:45 AM, %I is replaced with 1.
%M	The two-digit minute.
%m	The two-digit month.
%p	AM or PM.
%{PolicyName}	The name of the replication policy that the snapshot was created for. This variable is valid only if you are specifying a snapshot naming pattern for a replication policy.
%R	The time. This variable is equivalent to specifying **%H:%M**.
%r	The time. This variable is equivalent to specifying **%I:%M:%S %p**.
%S	The two-digit second.
%s	The second represented in UNIX or POSIX time.
%{SrcCluster}	The name of the source cluster of the replication policy that the snapshot was created for. This variable is valid only if you are specifying a snapshot naming pattern for a replication policy.
%T	The time. This variable is equivalent to specifying **%H:%M:%S**
%U	The two-digit numerical week of the year. Numbers range from 00 to 53. The first day of the week is calculated as Sunday.
%u	The numerical day of the week. Numbers range from 1 to 7. The first day of the week is calculated as Monday. For example, if a snapshot is created on Sunday, %u is replaced with 7.

Variable	Description
`%V`	The two-digit numerical week of the year that the snapshot was created in. Numbers range from `01` to `53`. The first day of the week is calculated as Monday. If the week of January 1 is four or more days in length, then that week is counted as the first week of the year.
`%v`	The day that the snapshot was created. This variable is equivalent to specifying **`%e-%b-%Y`**.
`%W`	The two-digit numerical week of the year that the snapshot was created in. Numbers range from `00` to `53`. The first day of the week is calculated as Monday.
`%w`	The numerical day of the week that the snapshot was created on. Numbers range from `0` to `6`. The first day of the week is calculated as Sunday. For example, if the snapshot was created on Sunday, `%w` is replaced with `0`.
`%X`	The time that the snapshot was created. This variable is equivalent to specifying **`%H:%M:%S`**.
`%Y`	The year that the snapshot was created in.
`%y`	The last two digits of the year that the snapshot was created in. For example, if the snapshot was created in 2014, `%y` is replaced with `14`.
`%Z`	The time zone that the snapshot was created in.
`%z`	The offset from coordinated universal time (UTC) of the time zone that the snapshot was created in. If preceded by a plus sign, the time zone is east of UTC. If preceded by a minus sign, the time zone is west of UTC.
`%+`	The time and date that the snapshot was created. This variable is equivalent to specifying **`%a %b %e %X %Z %Y`**.
`%%`	Escapes a percent sign. "`100%%`" is replaced with `100%`.

Managing snapshots

You can delete and view snapshots. You can also modify the name, duration period, and alias of an existing snapshot. However, you cannot modify the data contained in a snapshot; the data contained in a snapshot is read-only.

Reducing snapshot disk-space usage

If multiple snapshots contain the same directories, deleting one of the snapshots might not free the entire amount of space that the system reports as the size of the snapshot. The size of a snapshot is the maximum amount of data that might be freed if the snapshot is deleted.

Deleting a snapshot frees only the space that is taken up exclusively by that snapshot. If two snapshots reference the same stored data, that data is not freed until both snapshots are deleted. Remember that snapshots store data contained in all subdirectories of the root directory; if snapshot_one contains `/ifs/data/`, and snapshot_two contains `/ifs/data/dir`, the two snapshots most likely share data.

If you delete a directory, and then re-create it, a snapshot containing the directory stores the entire re-created directory, even if the files in that directory are never modified.

Deleting multiple snapshots that contain the same directories is more likely to free data than deleting multiple snapshots that contain different directories.

If multiple snapshots contain the same directories, deleting older snapshots is more likely to free disk-space than deleting newer snapshots.

Snapshots that are assigned expiration dates are automatically marked for deletion by the snapshot daemon. If the daemon is disabled, snapshots will not be automatically deleted by the system. It is recommended that you do not disable the snapshot daemon.

Delete snapshots

You can delete a snapshot if you no longer want to access the data contained in the snapshot.

OneFS frees disk space occupied by deleted snapshots when the SnapshotDelete job is run. Also, if you delete a snapshot that contains clones or cloned files, data in a shadow store might no longer be referenced by files on the cluster; OneFS deletes unreferenced data in a shadow store when the ShadowStoreDelete job is run. OneFS routinely runs both the shadow store delete and SnapshotDelete jobs. However, you can also manually run the jobs at any time.

Procedure

1. Click **Data Protection** › **SnapshotIQ** › **Snapshots**.
2. Specify the snapshots that you want to delete.
   a. For each snapshot you want to delete, in the **Saved File System Snapshots** table, in the row of a snapshot, select the check box.
   b. From the **Select an action** list, select **Delete**.
   c. In the confirmation dialog box, click **Delete**.
3. (Optional) To increase the speed at which deleted snapshot data is freed on the cluster, run the SnapshotDelete job.
   a. Click **Cluster Management** › **Job Operations** › **Job Types**.
   b. In the **Job Types** area, in the **SnapshotDelete** row, from the **Actions** column, select **Start Job**.
   c. Click **Start**.
4. (Optional) To increase the speed at which deleted data shared between deduplicated and cloned files is freed on the cluster, run the ShadowStoreDelete job.

   Run the ShadowStoreDelete job only after you run the SnapshotDelete job.

   a. Click **Cluster Management** › **Job Operations** › **Job Types**.
   b. In the **Job Types** area, in the **ShadowStoreDelete** row, from the **Actions** column, select **Start Job**.
   c. Click **Start**.

Modify snapshot attributes

You can modify the name and expiration date of a snapshot.

Procedure

1. Click **File System Management** › **SnapshotIQ** › **Snapshots**.
2. In the **Saved File System Snapshots** table, in the row of a snapshot, click **View Details**.
3. In the **Snapshot Details** area, modify snapshot attributes.
4. Next to each snapshot attribute that you modified, click **Save**.

Assign a snapshot alias to a snapshot

You can assign a snapshot alias to a snapshot.

Procedure

1. Click **Data Protection** › **SnapshotIQ** › **Snapshots**.
2. Above the **Saved File System Snapshots** table, click **View snapshot aliases**.
3. In the **Snapshot Aliases** table, in the row of an alias, click **View details**.
4. In the **Snapshot Alias Details** pane, in the **Alias Name** area, click **Edit**.
5. In the **Alias Name** field, type a new alias name.
6. Click **Save**.

View snapshots

You can view snapshots.

Procedure

1. Click **Data Protection** › **SnapshotIQ** › **Snapshots**.

   The 50 most recently generated snapshots appear in the **Saved File System Snapshots** table.
2. (Optional) To view additional snapshots, at the bottom of the **Saved File System Snapshots** table, click **Show 50 more**.

Snapshot information

You can view information about snapshots, including the total amount of space consumed by all snapshots.

The following information is displayed in the **Saved Snapshots** area:

SnapshotIQ Status
: Indicates whether a SnapshotIQ license has been activated on the cluster.

Total Number of Saved Snapshots
: Indicates the total number of snapshots that exist on the cluster.

Total Number of Snapshots Pending Deletion
: Indicates the total number of snapshots that were deleted on the cluster since the last snapshot delete job was run. The space consumed by the deleted snapshots is not freed until the snapshot delete job is run again.

Total Number of Snapshot Aliases

Indicates the total number of snapshot aliases that exist on the cluster.

Capacity Used by Saved Snapshots

Indicates the total amount of space consumed by all snapshots.

Restoring snapshot data

You can restore snapshot data through various methods. You can revert a snapshot or access snapshot data through the snapshots directory.

From the snapshots directory, you can either clone a file or copy a directory or a file. The snapshots directory can be accessed through Windows Explorer or a UNIX command line. You can disable and enable access to the snapshots directory for any of these methods through snapshots settings.

Revert a snapshot

You can revert a directory back to the state it was in when a snapshot was taken. Before OneFS reverts a snapshot, OneFS generates a snapshot of the directory being reverted, so that data stored in the directory is not lost. OneFS does not delete a snapshot after reverting it.

Before you begin

- Create a SnapRevert domain for the directory.
- Create a snapshot of a directory.

Procedure

1. Click **Cluster Management › Job Operations › Job Types**.
2. In the **Job Types** area, in the **SnapRevert** row, from the **Actions** column, select **Start Job**.
3. (Optional) To specify a priority for the job, from the **Priority** list, select a priority.

   Lower values indicate a higher priority. If you do not specify a priority, the job is assigned the default snapshot revert priority.

4. (Optional) To specify the amount of cluster resources the job is allowed to consume, from the **Impact policy** list, select an impact policy.

   If you do not specify a policy, the job is assigned the default snapshot revert policy.

5. In the **Snapshot ID to revert** field, type the name or ID of the snapshot that you want to revert, and then click **Start**.

Restore a file or directory using Windows Explorer

If the Microsoft Shadow Copy Client is installed on your computer, you can use it to restore files and directories that are stored in snapshots.

This method of restoring files and directories does not preserve the original permissions. Instead, this method assigns the file or directory the same permissions as the directory you are copying that file or directory into. To preserve permissions while restoring data from a snapshot, run the `cp` command with the `-a` option on a UNIX command line.

Note

You can access up to 64 snapshots of a directory through Windows explorer, starting with the most recent snapshot. To access more than 64 snapshots for a directory, access the cluster through a UNIX command line.

Procedure

1. In Windows Explorer, navigate to the directory that you want to restore or the directory that contains the file that you want to restore.

   If the directory has been deleted, you must recreate the directory.

2. Right-click the folder, and then click **Properties**.
3. In the **Properties** window, click the **Previous Versions** tab.
4. Select the version of the folder that you want to restore or the version of the folder that contains the version of the file that you want to restore.
5. Restore the version of the file or directory.
   - To restore all files in the selected directory, click **Restore**.
   - To copy the selected directory to another location, click **Copy** and then specify a location to copy the directory to.
   - To restore a specific file, click **Open**, and then copy the file into the original directory, replacing the existing copy with the snapshot version.

Restore a file or directory through a UNIX command line

You can restore a file or directory through a UNIX command line.

Procedure

1. Open a connection to the cluster through a UNIX command line.
2. (Optional) To view the contents of the snapshot you want to restore a file or directory from, run the `ls` command for a directory contained in the snapshots root directory.

   For example, the following command displays the contents of the `/archive` directory contained in Snapshot2012Jun04:

   ```
   ls /ifs/.snapshot/Snapshot2014Jun04/archive
   ```

3. Copy the file or directory by using the `cp` command.

   For example, the following command creates a copy of file1:

   ```
   cp -a /ifs/.snapshot/Snapshot2014Jun04/archive/file1 \
   /ifs/archive/file1_copy
   ```

Clone a file from a snapshot

You can clone a file from a snapshot. This procedure is available only through the command-line interface (CLI).

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. To view the contents of the snapshot you want to restore a file or directory from, run the `ls` command for a subdirectory of the snapshots root directory.

For example, the following command displays the contents of the `/archive` directory contained in Snapshot2014Jun04:

```
ls /ifs/.snapshot/Snapshot2014Jun04/archive
```

3. Clone a file from the snapshot by running the `cp` command with the `-c` option.

   For example, the following command clones test.txt from Snapshot2014Jun04:

```
cp -c /ifs/.snapshot/Snapshot2014Jun04/archive/test.txt \
/ifs/archive/test_clone.text
```

Managing snapshot schedules

You can modify, delete, and view snapshot schedules.

Modify a snapshot schedule

You can modify a snapshot schedule. Any changes to a snapshot schedule are applied only to snapshots generated after the modifications are made. Existing snapshots are not affected by schedule modifications.

If you modify the alias of a snapshot schedule, the alias is assigned to the next snapshot generated based on the schedule. However, if you do this, the old alias is not removed from the last snapshot that it was assigned to. Unless you manually remove the old alias, the alias will remain attached to the last snapshot that it was assigned to.

Procedure

1. Click **Data Protection** › **SnapshotIQ** › **Snapshot Schedules**.
2. In the **Snapshot Schedules** table, in the row of the snapshot schedule you want to modify, click **View details**.
3. In the **Snapshot Schedule Details** area, modify snapshot schedule attributes.
4. Next to each snapshot schedule attribute that you modified, click **Save**.

Delete a snapshot schedule

You can delete a snapshot schedule. Deleting a snapshot schedule will not delete snapshots that were previously generated according to the schedule.

Procedure

1. Click **Data Protection** › **SnapshotIQ** › **Snapshot Schedules**.
2. In the **Snapshot Schedules** table, in the row of the snapshot schedule you want to delete, click **Delete**.
3. In the **Confirm Delete** dialog box, click **Delete**.

View snapshot schedules

You can view snapshot schedules.

Procedure

1. Click **Data Protection** › **SnapshotIQ** › **Snapshot Schedules**.
2. In the **Snapshot Schedules** table, view snapshot schedules.
3. (Optional) To view detailed information about a snapshot schedule, in the **Snapshot Schedules** table, in the row of the snapshot schedule that you want to view, click **View details**.

Snapshot schedule settings are displayed in the **Snapshot Schedule Details** area. Snapshots that are scheduled to be generated according to the schedule are displayed in the **Snapshot Calendar** area.

Managing snapshot aliases

You can configure snapshot schedules to assign a snapshot alias to the most recent snapshot created by a snapshot schedule. You can also manually assign snapshot aliases to specific snapshots or the live version of the file system.

Configure a snapshot alias for a snapshot schedule

You can configure a snapshot schedule to assign a snapshot alias to the most recent snapshot created by the schedule.

Procedure

1. Click **Data Protection** › **SnapshotIQ** › **Snapshot Schedules**
2. In the **Snapshot Schedules** table, in the row of the snapshot schedule you want to configure, click **View details**.
3. Next to **Snapshot Alias**, click **Create an Alias**.
4. In the **Snapshot Alias** area, select **Yes**.
5. In the **Alias Name** field, type the name of the snapshot alias.
6. In the **Snapshot Alias** area, click **Save**.

Assign a snapshot alias to a snapshot

You can assign a snapshot alias to a snapshot.

Procedure

1. Click **Data Protection** › **SnapshotIQ** › **Snapshots**.
2. Above the **Saved File System Snapshots** table, click **View snapshot aliases**.
3. In the **Snapshot Aliases** table, in the row of an alias, click **View details**.
4. In the **Snapshot Alias Details** pane, in the **Alias Name** area, click **Edit**.
5. In the **Alias Name** field, type a new alias name.
6. Click **Save**.

Reassign a snapshot alias to the live file system

You can reassign a snapshot alias to redirect clients from a snapshot to the live file system. This procedure is available only through the command-line interface (CLI).

Procedure

1. Run the `isi snapshot aliases modify` command.

   The following command reassigns the latestWeekly alias to the live file system:

   ```
   isi snapshot aliases modify latestWeekly --target LIVE
   ```

View snapshot aliases

You can view a list of all snapshot aliases. This procedure is available only through the command-line interface (CLI).

Procedure

1. View a list of all snapshot aliases by running the following command:

   ```
   isi snapshot aliases list
   ```

   If a snapshot alias references the live version of the file system, the `Target ID` is `-1`.

2. (Optional) View information about a specific snapshot by running the `isi snapshot aliases view` command.

   The following command displays information about latestWeekly:

   ```
   isi snapshot aliases view latestWeekly
   ```

Snapshot alias information

You can view information about snapshot aliases through the output of the `isi snapshot aliases view` command.

ID
The numerical ID of the snapshot alias.

Name
The name of the snapshot alias.

Target ID
The numerical ID of the snapshot that is referenced by the alias.

Target Name
The name of the snapshot that is referenced by the alias.

Created
The date that the snapshot alias was created.

Managing with snapshot locks

You can delete, create, and modify the expiration date of snapshot locks.

⚠ CAUTION

It is recommended that you do not create, delete, or modify snapshots locks unless you are instructed to do so by Isilon Technical Support.

Deleting a snapshot lock that was created by OneFS might result in data loss. If you delete a snapshot lock that was created by OneFS, it is possible that the corresponding snapshot might be deleted while it is still in use by OneFS. If OneFS cannot access a snapshot that is necessary for an operation, the operation will malfunction and data loss might result. Modifying the expiration date of a snapshot lock created by OneFS can also result in data loss because the corresponding snapshot can be deleted prematurely.

Create a snapshot lock

You can create snapshot locks that prevent snapshots from being deleted. This procedure is available only through the command-line interface (CLI).

Although you can prevent a snapshot from being automatically deleted by creating a snapshot lock, it is recommended that you do not create snapshot locks. To prevent a snapshot from being automatically deleted, it is recommended that you extend the duration period of the snapshot.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Create a snapshot lock by running the `isi snapshot locks create` command.

   For example, the following command applies a snapshot lock to "SnapshotApril2012", sets the lock to expire in one month, and adds a description of "Maintenance Lock":

   ```
   isi snapshot locks create SnapshotApril2012 --expires 1M \
   --comment "Maintenance Lock"
   ```

Modify a snapshot lock expiration date

You can modify the expiration date of a snapshot lock. This procedure is available only through the command-line interface (CLI).

⚠ CAUTION

It is recommended that you do not modify the expiration dates of snapshot locks.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Run the `isi snapshot locks modify` command.

   The following command sets an expiration date two days from the present date for a snapshot lock with an ID of 1 that is applied to a snapshot named SnapshotApril2014:

   ```
   isi snapshot locks modify SnapshotApril2014 1 --expires 2D
   ```

Delete a snapshot lock

You can delete a snapshot lock. This procedure is available only through the command-line interface (CLI).

⚠ CAUTION

It is recommended that you do not delete snapshot locks.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Delete a snapshot lock by running the `isi snapshot locks delete` command.

For example, the following command deletes a snapshot lock that is applied to SnapshotApril2014 and has a lock ID of 1:

```
isi snapshot locks delete Snapshot2014Apr16 1
```

The system prompts you to confirm that you want to delete the snapshot lock.

3. Type **yes** and then press ENTER.

Snapshot lock information

You can view snapshot lock information through the `isi snapshot locks view` and `isi snapshot locks list` commands.

ID

Numerical identification number of the snapshot lock.

Comment

Description of the snapshot lock. This can be any string specified by a user.

Expires

The date that the snapshot lock will be automatically deleted by OneFS.

Count

The number of times the snapshot lock is held.
The file clone operation can hold a single snapshot lock multiple times. If multiple file clones are created simultaneously, the file clone operation holds the same lock multiple times, rather than creating multiple locks. If you delete a snapshot lock that is held more than once, you will delete only one of the instances that the lock is held. In order to delete a snapshot lock that is held multiple times, you must delete the snapshot lock the same number of times as displayed in the count field.

Configure SnapshotIQ settings

You can configure SnapshotIQ settings that determine how snapshots can be created and the methods that users can access snapshot data.

Procedure

1. Click **Data Protection** › **SnapshotIQ** › **Settings**.
2. Modify SnapshotIQ settings, and then click **Save**.

SnapshotIQ settings

SnapshotIQ settings determine how snapshots behave and can be accessed.

The following SnapshotIQ settings can be configured:

Snapshot Scheduling

Determines whether snapshots can be generated.

Note

Disabling snapshot generation might cause some OneFS operations to fail. It is recommended that you do not disable this setting.

Auto-create Snapshots

Determines whether snapshots are automatically generated according to snapshot schedules.

Auto-delete Snapshots

Determines whether snapshots are automatically deleted according to their expiration dates.

NFS Visibility & Accessibility

Root Directory Accessible

Determines whether snapshot directories are accessible through NFS.

Root Directory Visible

Determines whether snapshot directories are visible through NFS.

Sub-directories Accessible

Determines whether snapshot subdirectories are accessible through NFS.

SMB Visibility & Accessible

Root Directory Accessible

Determines whether snapshot directories are accessible through SMB.

Root Directory Visible

Determines whether snapshot directories are visible through SMB.

Sub-directories Accessible

Determines whether snapshot subdirectories are accessible through SMB.

Local Visibility & Accessibility

Root Directory Accessible

Determines whether snapshot directories are accessible through the local file system. You can access the local file system through an SSH connection or the local console.

Root Directory Visible

Determines whether snapshot directories are visible through the local file system. You can access the local file system through an SSH connection or the local console.

Sub-directories Accessible

Determines whether snapshot subdirectories are accessible through the local file system. You can access the local file system through an SSH connection or the local console.

Set the snapshot reserve

You can specify a minimum percentage of cluster-storage capacity that you want to reserve for snapshots. This procedure is available only through the command-line interface (CLI).

The snapshot reserve does not limit the amount of space that snapshots are allowed to consume on the cluster. Snapshots can consume more than the percentage of capacity specified by the snapshot reserve. It is recommended that you do not specify a snapshot reserve.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Set the snapshot reserve by running the `isi snapshot settings modify` command with the `--reserve` option.

   For example, the following command sets the snapshot reserve to 20%:

   ```
   isi snapshot settings modify --reserve 20
   ```

CHAPTER 11

Deduplication with SmartDedupe

This section contains the following topics:

Deduplication overview

The SmartDedupe software module enables you to save storage space on your cluster by reducing redundant data. Deduplication maximizes the efficiency of your cluster by decreasing the amount of storage required to store multiple files with similar blocks.

SmartDedupe deduplicates data by scanning an Isilon cluster for identical data blocks. Each block is 8 KB. If SmartDedupe finds duplicate blocks, SmartDedupe moves a single copy of the blocks to a hidden file called a shadow store. SmartDedupe then deletes the duplicate blocks from the original files and replaces the blocks with pointers to the shadow store.

Deduplication is applied at the directory level, targeting all files and directories underneath one or more root directories. You can first assess a directory for deduplication and determine the estimated amount of space you can expect to save. You can then decide whether to deduplicate the directory. After you begin deduplicating a directory, you can monitor how much space is saved by deduplication in real time.

SmartDedupe does not deduplicate files that are 32 KB and smaller, because doing so would consume more cluster resources than the storage savings are worth. Each shadow store can contain up to 255 blocks. Each block in a shadow store can be referenced 32000 times.

Deduplication jobs

Deduplication is performed by maintenance jobs referred to as deduplication jobs. You can monitor and control deduplication jobs as you would any other maintenance job on the cluster. Although the overall performance impact of deduplication is minimal, the deduplication job consumes 256 MB of memory per node.

When a deduplication job is first run on a cluster, SmartDedupe samples blocks from each file and creates index entries for those blocks. If the index entries of two blocks match, SmartDedupe scans the blocks adjacent to the matching pair and then deduplicates all duplicate blocks. After a deduplication job samples a file once, new deduplication jobs will not sample the file again until the file is modified.

The first deduplication job you run might take significantly longer to complete than subsequent deduplication jobs. The first deduplication job must scan all files under the specified directories to generate the initial index. If subsequent deduplication jobs take a long time to complete, this most likely indicates that a large amount of data is being deduplicated. However, it can also indicate that clients are creating a large amount of new data on the cluster. If a deduplication job is interrupted during the deduplication process, the job will automatically restart the scanning process from where the job was interrupted.

It is recommended that you run deduplication jobs when clients are not modifying data on the cluster. If clients are continually modifying files on the cluster, the amount of space saved by deduplication is minimal because the deduplicated blocks are constantly removed from the shadow store. For most clusters, it is recommended that you start a deduplication job every ten days.

The permissions required to modify deduplication settings are not the same as those needed to run a deduplication job. Although a user must have the maintenance job permission to run a deduplication job, the user must have the deduplication permission to modify deduplication settings. By default, the deduplication job is configured to run at a low priority.

Data replication and backup with deduplication

When deduplicated files are replicated to another Isilon cluster or backed up to a tape device, the deduplicated files no longer share blocks on the target Isilon cluster or backup device. However, although you can deduplicate data on a target Isilon cluster, you cannot deduplicate data on an NDMP backup device.

Shadows stores are not transferred to target clusters or backup devices. Because of this, deduplicated files do not consume less space than non-deduplicated files when they are replicated or backed up. To avoid running out of space, you must ensure that target clusters and tape devices have enough free space to store deduplicated data as if the data had not been deduplicated. To reduce the amount of storage space consumed on a target Isilon cluster, you can configure deduplication for the target directories of your replication policies. Although this will deduplicate data on the target directory, it will not allow SyncIQ to transfer shadow stores. Deduplication is still performed by deduplication jobs running on the target cluster.

The amount of cluster resources required to backup and replicate deduplicated data is the same as for non-deduplicated data. You can deduplicate data while the data is being replicated or backed up.

Snapshots with deduplication

You cannot deduplicate the data stored in a snapshot. However, you can create snapshots of deduplicated data.

If you create a snapshot for a deduplicated directory, and then modify the contents of that directory, the references to shadow stores will be transferred to the snapshot over time. Therefore, if you enable deduplication before you create snapshots, you will save more space on your cluster. If you implement deduplication on a cluster that already has a significant amount of data stored in snapshots, it will take time before the snapshot data is affected by deduplication. Newly created snapshots can contain deduplicated data, but snapshots created before deduplication was implemented cannot.

If you plan on reverting a snapshot, it is best to revert the snapshot before running a deduplication job. Restoring a snapshot can overwrite many of the files on the cluster. Any deduplicated files are reverted back to normal files if they are overwritten by a snapshot revert. However, after the snapshot revert is complete, you can deduplicate the directory and the space savings persist on the cluster.

Deduplication considerations

Deduplication can significantly increase the efficiency at which you store data. However, the effect of deduplication varies depending on the cluster.

You can reduce redundancy on a cluster by running SmartDedupe. Deduplication creates links that can impact the speed at which you can read from and write to files. In particular, sequentially reading chunks smaller than 512 KB of a deduplicated file can be significantly slower than reading the same small, sequential chunks of a non-deduplicated file. This performance degradation applies only if you are reading non-cached data. For cached data, the performance for deduplicated files is potentially better than non-deduplicated files. If you stream chunks larger than 512 KB, deduplication does not significantly impact the read performance of the file. If you intend on streaming 8 KB or less of each file at a time, and you do not plan on concurrently streaming the files, it is recommended that you do not deduplicate the files.

Deduplication is most effective when applied to static or archived files and directories. The less files are modified, the less negative effect deduplication has on the cluster. For example, virtual machines often contain several copies of identical files that are rarely modified. Deduplicating a large number of virtual machines can greatly reduce consumed storage space.

Shadow-store considerations

Shadow stores are hidden files that are referenced by cloned and deduplicated files. Files that reference shadow stores behave differently than other files.

- Reading shadow-store references might be slower than reading data directly. Specifically, reading non-cached shadow-store references is slower than reading non-cached data. Reading cached shadow-store references takes no more time than reading cached data.
- When files that reference shadow stores are replicated to another Isilon cluster or backed up to a Network Data Management Protocol (NDMP) backup device, the shadow stores are not transferred to the target Isilon cluster or backup device. The files are transferred as if they contained the data that they reference from shadow stores. On the target Isilon cluster or backup device, the files consume the same amount of space as if they had not referenced shadow stores.
- When OneFS creates a shadow store, OneFS assigns the shadow store to a storage pool of a file that references the shadow store. If you delete the storage pool that a shadow store resides on, the shadow store is moved to a pool occupied by another file that references the shadow store.
- OneFS does not delete a shadow-store block immediately after the last reference to the block is deleted. Instead, OneFS waits until the ShadowStoreDelete job is run to delete the unreferenced block. If a large number of unreferenced blocks exist on the cluster, OneFS might report a negative deduplication savings until the ShadowStoreDelete job is run.
- Shadow stores are protected at least as much as the most protected file that references it. For example, if one file that references a shadow store resides in a storage pool with +2 protection and another file that references the shadow store resides in a storage pool with +3 protection, the shadow store is protected at +3.
- Quotas account for files that reference shadow stores as if the files contained the data referenced from shadow stores; from the perspective of a quota, shadow-store references do not exist. However, if a quota includes data protection overhead, the quota does not account for the data protection overhead of shadow stores.

SmartDedupe license functionality

You can deduplicate data only if you activate a SmartDedupe license on a cluster. However, you can assess deduplication savings without activating a SmartDedupe license.

If you activate a SmartDedupe license, and then deduplicate data, the space savings are not lost if the license becomes inactive. You can also still view deduplication savings while the license is inactive. However, you will not be able to deduplicate additional data until you re-activate the SmartDedupe license.

Managing deduplication

You can manage deduplication on a cluster by first assessing how much space you can save by deduplicating individual directories. After you determine which directories are worth deduplicating, you can configure SmartDedupe to deduplicate those directories specifically. You can then monitor the actual amount of disk space you are saving.

Assess deduplication space savings

You can assess the amount of disk space you will save by deduplicating a directory.

Procedure

1. Click **File System Management** › **Deduplication** › **Settings**.
2. In the **Assess Deduplication** area, click **Browse** and select a directory that you want to deduplicate.

   If you assess multiple directories, disk savings are not differentiated by directory in the deduplication report.
3. Click **Cluster Management** › **Job Operations** › **Job Types**.
4. In the **Job Types** table, in the row of the **DedupeAssessment** job, from the **Actions** column, select **Start Job**.
5. Click **Cluster Management** › **Job Operations** › **Job Summary**.
6. Wait for the assessment job to complete.

   When the DedupeAssessment job is complete, the job is removed from the **Active Jobs** table.
7. Click **File System Management** › **Deduplication** › **Summary**.

   In the **Deduplication Assessment Reports** table, in the row of the most recent assessment job, click **View Details**.
8. View the amount of disk space that will be saved if you deduplicate the directory.

   The number of blocks that will be deduplicated is displayed in the **Deduped blocks** field.

Specify deduplication settings

You can specify which directories you want to deduplicate.

Procedure

1. Click **File System Management** › **Deduplication** › **Settings**.
2. In the **Deduplication Settings** area, click **Browse** and select a directory that you want to deduplicate.
3. (Optional) Specify additional directories.
   a. Click **Add another directory path**.
   b. Click **Browse** and select a directory that you want to deduplicate.
4. Click **Cluster Management** › **Job Operations** › **Jobs Types**.
5. In the **Jobs** table, in the row of the Dedupe job, click **View/Edit**.
6. Click **Edit Job Type**.
7. Modify the settings of the deduplication job, and then click **Save Changes**.

View deduplication space savings

You can view the amount of disk space that you are currently saving with deduplication.

Procedure

1. Click **File System Management** › **Deduplication** › **Summary**.
2. In the **Deduplication Savings** area, view the amount of disk space saved.

View a deduplication report

After a deduplication job completes, you can view information about the job in a deduplication report.

Procedure

1. Click **File System Management** › **Deduplication** › **Summary**.
2. Select a deduplication report.
   - To view a report about a deduplication job, in the **Deduplication Reports** table, click **View Report**.
   - To view a report about a deduplication assessment job, in the **Deduplication Assessment Reports** table, click **View Report**.

Deduplication job report information

You can view the following deduplication specific information in deduplication job reports:

Start time
: The time the deduplication job started.

End time
: The time the deduplication job ended.

Iteration Count
: The number of times that SmartDedupe interrupted the sampling process. If SmartDedupe is sampling a large amount of data, SmartDedupe might interrupt sampling in order to start deduplicating the data. After SmartDedupe finishes deduplicating the sampled data, SmartDedupe will continue sampling the remaining data.

Scanned blocks
: The total number of blocks located underneath the specified deduplicated directories.

Sampled blocks
: The number of blocks that SmartDedupe created index entries for.

Deduped blocks
: The number of blocks that were deduplicated.

Dedupe percent
: The percentage of scanned blocks that were deduplicated.

Created dedupe requests
: The total number of deduplication requests created. A deduplication request is created for each matching pair of data blocks. For example, if you have 3 data blocks that all match, SmartDedupe creates 2 requests. One of the requests could pair file1 and file2 together and the other request could pair file2 and file3 together.

Successful dedupe requests
The number of deduplication requests that completed successfully.

Failed dedupe requests
The number of deduplication requests that failed. If a deduplication request fails, it doesn't mean that the job failed too. A deduplication request can fail for any number of reasons. For example, the file might have been modified since it was sampled.

Skipped files
The number of files that were not scanned by the deduplication job. SmartDedupe skips files for a number of reasons. For example, SmartDedupe skips files that have already been scanned and haven't been modified since. SmartDedupe also skips all files that are smaller than 4 KB.

Index entries
The number of entries that currently exist in the index.

Index lookup attempts
The total number of lookups that have been done by earlier deduplication jobs plus the number of lookups done by this deduplication job. A lookup is when the deduplication job attempts to match a block that was indexed with a block that hasn't been indexed.

Index lookup hits
The number of blocks that matched index entries.

Deduplication information

You can view the amount of disk space saved by deduplication in the **Deduplication Savings** area:

Space Savings
The total amount of physical disk space saved by deduplication, including protection overhead and metadata. For example, if you have three identical files that are all 5 GB, the estimated physical saving would be greater than 10 GB, because deduplication saved space that would have been occupied by file metadata and protection overhead.

Deduplicated data
The amount of space on the cluster occupied by directories that were deduplicated.

Other data
The amount of space on the cluster occupied by directories that were not deduplicated.

CHAPTER 12

Data replication with SyncIQ

This section contains the following topics:

SyncIQ backup and recovery overview

OneFS enables you to replicate data from one Isilon cluster to another through the SyncIQ software module. You must activate a SyncIQ license on both Isilon clusters before you can replicate data between them.

You can replicate data at the directory level while optionally excluding specific files and sub-directories from being replicated. SyncIQ creates and references snapshots to replicate a consistent point-in-time image of a root directory. Metadata such as access control lists (ACLs) and alternate data streams (ADS) are replicated along with data.

SyncIQ enables you to maintain a consistent backup copy of your data on another Isilon cluster. SyncIQ offers automated failover and failback capabilities that enable you to continue operations on another Isilon cluster if a primary cluster becomes unavailable.

Replication policies and jobs

Data replication is coordinated according to replication policies and jobs. Replication policies specify what data is replicated, where the data is replicated to, and how often the data is replicated. Replication jobs are the operations that replicate data from one Isilon cluster to another. SyncIQ generates replication jobs according to replication policies.

A replication policy specifies two clusters: the source and the target. The cluster on which the replication policy exists is the source cluster. The cluster that data is being replicated to is the target cluster. When a replication policy starts, SyncIQ generates a replication job for the policy. When a replication job runs, files from a directory on the source cluster are replicated to a directory on the target cluster; these directories are known as source and target directories.

After the first replication job created by a replication policy finishes, the target directory and all files contained in the target directory are set to a read-only state, and can be modified only by other replication jobs belonging to the same replication policy. There is no limit to the number of replication policies that can exist on a cluster.

Note

To prevent permissions errors, make sure that ACL policy settings are the same across source and target clusters.

You can create two types of replication policies: synchronization policies and copy policies. A synchronization policy maintains an exact replica of the source directory on the target cluster. If a file or sub-directory is deleted from the source directory, the file or directory is deleted from the target cluster when the policy is run again.

You can use synchronization policies to fail over and fail back data between source and target clusters. When a source cluster becomes unavailable, you can fail over data on a target cluster and make the data available to clients. When the source cluster becomes available again, you can fail back the data to the source cluster.

A copy policy maintains recent versions of the files that are stored on the source cluster. However, files that are deleted on the source cluster are not deleted from the target cluster. Failback is not supported for copy policies. Copy policies are most commonly used for archival purposes.

Copy policies enable you to remove files from the source cluster without losing those files on the target cluster. Deleting files on the source cluster improves performance on the source cluster while maintaining the deleted files on the target cluster. This can be useful if, for example, your source cluster is being used for production purposes and your target cluster is being used only for archiving.

After creating a job for a replication policy, SyncIQ must wait until the job completes before it can create another job for the policy. Any number of replication jobs can exist on a cluster at a given time; however, only five replication jobs can run on a source cluster at the same time. If more than five replication jobs exist on a cluster, the first five jobs run while the others are queued to run. The number of replication jobs that a single target cluster can support concurrently is dependent on the number of workers available on the target cluster.

You can replicate any number of files and directories with a single replication job. You can prevent a large replication job from overwhelming the system by limiting the amount of cluster resources and network bandwidth that data synchronization is allowed to consume. Because each node in a cluster is able to send and receive data, the speed at which data is replicated increases for larger clusters.

Automated replication policies

You can manually start a replication policy at any time, but you can also configure replication policies to start automatically based on source directory modifications or a schedule.

You can configure a replication policy to run according to a schedule, so that you can control when replication is performed. You can also configure a replication policy to start when SyncIQ detects a modification to the source directory, so that SyncIQ maintains a more current version of your data on the target cluster.

Scheduling a policy can be useful under the following conditions:

- You want to replicate data when user activity is minimal
- You can accurately predict when modifications will be made to the data

Configuring a policy to start when changes are made to the source directory can be useful under the following conditions:

- You want retain a consistent copy of your data at all times
- You are expecting a large number of changes at unpredictable intervals

For policies that are configured to start whenever changes are made to the source directory, SyncIQ checks the source directories every ten seconds. SyncIQ does not account for excluded files or directories when detecting changes, so policies that exclude files or directories from replication might be run unnecessarily. For example, assume that newPolicy replicates `/ifs/data/media` but excludes `/ifs/data/media/temp`. If a modification is made to `/ifs/data/media/temp/file.txt`, SyncIQ will run newPolicy, but will not replicate `/ifs/data/media/temp/file.txt`.

If a policy is configured to start whenever changes are made to its source directory, and a replication job fails, SyncIQ will wait one minute before attempting to run the policy again. SyncIQ will increase this delay exponentially for each failure up to a maximum delay of eight hours. You can override the delay by running the policy manually at any time. After a job for the policy completes successfully, SyncIQ will resume checking the source directory every ten seconds.

Source and target cluster association

SyncIQ associates a replication policy with a target cluster by marking the target cluster when the job runs for the first time. Even if you modify the name or IP address of the target cluster, the mark persists on the target cluster. When a replication policy is run, SyncIQ checks the mark to ensure that data is being replicated to the correct location.

On the target cluster, you can manually break an association between a replication policy and target directory. Breaking the association between a source and target cluster causes the mark on the target cluster to be deleted. You might want to manually break a target association if an association is obsolete. If you break the association of a policy, the policy is disabled on the source cluster and you cannot run the policy. If you want to run the disabled policy again, you must reset the replication policy.

Note

Breaking a policy association causes either a full or differential replication to occur the next time you run the replication policy. During a full or differential replication, SyncIQ creates a new association between the source and target clusters. Depending on the amount of data being replicated, a full or differential replication can take a very long time to complete.

Full and differential replication

If a replication policy encounters an issue that cannot be fixed (for example, if the association was broken on the target cluster), you might need to reset the replication policy. If you reset a replication policy, SyncIQ performs either a full or differential replication the next time the policy is run. You can specify the type of replication that SyncIQ performs.

During a full replication, SyncIQ transfers all data from the source cluster regardless of what data exists on the target cluster. A full replication consumes large amounts of network bandwidth and can take a very long time to complete. However, a full replication is less strenuous on CPU usage than a differential replication.

During a differential replication, SyncIQ first checks whether a file already exists on the target cluster and then transfers only data that does not already exist on the target cluster. A differential replication consumes less network bandwidth than a full replication; however, differential replications consume more CPU. Differential replication can be much faster than a full replication if there is an adequate amount of available CPU for the differential replication job to consume.

Controlling replication job resource consumption

You can create rules that limit the network traffic created and the rate at which files are sent by replication jobs. You can also specify the number of workers that are spawned by a replication policy to limit the amount of cluster resources that are consumed. Also, you can restrict a replication policy to connect only to a specific storage pool.

You can create network-traffic rules that control the amount of network traffic generated by replication jobs during specified time periods. These rules can be useful if, for example, you want to limit the amount of network traffic created during other resource-intensive operations.

You can create multiple network traffic rules to enforce different limitations at different times. For example, you might allocate a small amount of network bandwidth during peak business hours, but allow unlimited network bandwidth during non-peak hours.

When a replication job runs, OneFS generates workers on the source and target cluster. Workers on the source cluster send data while workers on the target cluster write data. OneFS generates no more than 40 workers for a replication job. You can modify the maximum number of workers generated per node to control the amount of resources that a replication job is allowed to consume. For example, you can increase the maximum number of workers per node to increase the speed at which data is replicated to the target cluster.

You can also reduce resource consumption through file-operation rules that limit the rate at which replication policies are allowed to send files. However, it is recommended that you only create file-operation rules if the files you intend to replicate are predictably similar in size and not especially large.

Replication reports

After a replication job completes, SyncIQ generates a report that contains detailed information about the job, including how long the job ran, how much data was transferred, and what errors occurred.

If a replication report is interrupted, SyncIQ might create a subreport about the progress of the job so far. If the job is then restarted, SyncIQ creates another subreport about the progress of the job until the job either completes or is interrupted again. SyncIQ creates a subreport each time the job is interrupted until the job completes successfully. If multiple subreports are created for a job, SyncIQ combines the information from the subreports into a single report.

SyncIQ routinely deletes replication reports. You can specify the maximum number of replication reports that SyncIQ retains and the length of time that SyncIQ retains replication reports. If the maximum number of replication reports is exceeded on a cluster, SyncIQ deletes the oldest report each time a new report is created.

You cannot customize the content of a replication report.

Note

If you delete a replication policy, SyncIQ automatically deletes any reports that were generated for that policy.

Replication snapshots

SyncIQ generates snapshots to facilitate replication, failover, and failback between Isilon clusters. Snapshots generated by SyncIQ can also be used for archival purposes on the target cluster.

Source cluster snapshots

SyncIQ generates snapshots on the source cluster to ensure that a consistent point-in-time image is replicated and that unaltered data is not sent to the target cluster.

Before running a replication job, SyncIQ creates a snapshot of the source directory. SyncIQ then replicates data according to the snapshot rather than the current state of the cluster, allowing users to modify source-directory files while ensuring that an exact point-in-time image of the source directory is replicated.

For example, if a replication job of `/ifs/data/dir/` starts at 1:00 PM and finishes at 1:20 PM, and `/ifs/data/dir/file` is modified at 1:10 PM, the modifications are not reflected on the target cluster, even if `/ifs/data/dir/file` is not replicated until 1:15 PM.

You can replicate data according to a snapshot generated with the SnapshotIQ tool. If you replicate data according to a SnapshotIQ snapshot, SyncIQ does not generate another snapshot of the source directory. This method can be useful if you want to replicate identical copies of data to multiple Isilon clusters.

SyncIQ generates source snapshots to ensure that replication jobs do not transfer unmodified data. When a job is created for a replication policy, SyncIQ checks whether it is the first job created for the policy. If it is not the first job created for the policy, SyncIQ compares the snapshot generated for the earlier job with the snapshot generated for the new job.

SyncIQ replicates only data that has changed since the last time a snapshot was generated for the replication policy. When a replication job is completed, SyncIQ deletes the previous source-cluster snapshot and retains the most recent snapshot until the next job is run.

Target cluster snapshots

When a replication job is run, SyncIQ generates a snapshot on the target cluster to facilitate failover operations. When the next replication job is created for the replication policy, the job creates a new snapshot and deletes the old one.

If a SnapshotIQ license has been activated on the target cluster, you can configure a replication policy to generate additional snapshots that remain on the target cluster even as subsequent replication jobs run.

SyncIQ generates target snapshots to facilitate failover on the target cluster regardless of whether a SnapshotIQ license has been configured on the target cluster. Failover snapshots are generated when a replication job completes. SyncIQ retains only one failover snapshot per replication policy, and deletes the old snapshot after the new snapshot is created.

If a SnapshotIQ license has been activated on the target cluster, you can configure SyncIQ to generate archival snapshots on the target cluster that are not automatically deleted when subsequent replication jobs run. Archival snapshots contain the same data as the snapshots that are generated for failover purposes. However, you can configure how long archival snapshots are retained on the target cluster. You can access archival snapshots the same way that you access other snapshots generated on a cluster.

Data failover and failback with SyncIQ

SyncIQ enables you to perform automated data failover and failback operations between Isilon clusters. If a cluster is rendered unusable, you can fail over to another Isilon cluster, enabling clients to access their data on the other cluster. If the unusable cluster becomes accessible again, you can fail back to the original Isilon cluster.

For the purposes of explaining failover and failback procedures, the cluster originally accessed by clients is referred to as the primary cluster, and the cluster that client data is originally replicated to is referred to as the secondary cluster. Failover is the process that allows clients to modify data on a secondary cluster. Failback is the process that allows clients to access data on the primary cluster again and begins to replicate data back to the secondary cluster.

Failover and failback can be useful in disaster recovery procedures. For example, if a primary cluster is damaged by a natural disaster, you can migrate clients to a secondary cluster until the primary cluster is repaired and then migrate the clients back to the primary cluster.

You can fail over and fail back to facilitate scheduled cluster maintenance. For example, if you are upgrading the primary cluster, you might want to migrate clients to a secondary cluster until the upgrade is complete and then migrate clients back to the primary cluster.

Note

Data failover and failback is not supported for SmartLock directories.

Data failover

Data failover is the process of preparing data on a secondary cluster to be modified by clients. After you fail over to a secondary cluster, you can redirect clients to modify their data on the secondary cluster.

Before failover is performed, you must create and run a replication policy on the primary cluster. You initiate the failover process on the secondary cluster. Failover is performed per replication policy; to migrate data that is spread across multiple replication policies, you must initiate failover for each replication policy.

You can use any replication policy to fail over. However, if the action of the replication policy is set to copy, any file that was deleted on the primary cluster will be present on the secondary cluster. When the client connects to the secondary cluster, all files that were deleted on the primary cluster will be available to the client.

If you initiate failover for a replication policy while an associated replication job is running, the failover operation completes but the replication job fails. Because data might be in an inconsistent state, SyncIQ uses the snapshot generated by the last successful replication job to revert data on the secondary cluster to the last recovery point.

If a disaster occurs on the primary cluster, any modifications to data that were made after the last successful replication job started are not reflected on the secondary cluster. When a client connects to the secondary cluster, their data appears as it was when the last successful replication job was started.

Data failback

Data failback is the process of restoring clusters to the roles they occupied before a failover operation. After data failback is complete, the primary cluster hosts clients and replicates data to the secondary cluster for backup.

The first step in the failback process is updating the primary cluster with all of the modifications that were made to the data on the secondary cluster. The next step in the failback process is preparing the primary cluster to be accessed by clients. The final step in the failback process is resuming data replication from the primary to the secondary cluster. At the end of the failback process, you can redirect users to resume accessing their data on the primary cluster.

You can fail back data with any replication policy that meets all of the following criteria:

- The source directory is not a SmartLock directory.
- The policy has been failed over.
- The policy is a synchronization policy.
- The policy does not exclude any files or directories from replication.

Recovery times and objectives for SyncIQ

The Recovery Point Objective (RPO) and the Recovery Time Objective (RTO) are measurements of the impacts that a disaster can have on business operations. You can calculate your RPO and RTO for a disaster recovery with replication policies.

RPO is the maximum amount of time for which data is lost if a cluster suddenly becomes unavailable. For an Isilon cluster, the RPO is the amount of time that has passed since the last completed replication job started. The RPO is never greater than the time it takes for two consecutive replication jobs to run and complete.

If a disaster occurs while a replication job is running, the data on the secondary cluster is reverted to the state it was in when the last replication job completed. For example, consider an environment in which a replication policy is scheduled to run every three hours, and replication jobs take two hours to complete. If a disaster occurs an hour after a replication job begins, the RPO is four hours, because it has been four hours since a completed job began replicating data.

RTO is the maximum amount of time required to make backup data available to clients after a disaster. The RTO is always less than or approximately equal to the RPO, depending on the rate at which replication jobs are created for a given policy.

If replication jobs run continuously, meaning that another replication job is created for the policy before the previous replication job completes, the RTO is approximately equal to the RPO. When the secondary cluster is failed over, the data on the cluster is reset to the state it was in when the last job completed; resetting the data takes an amount of time proportional to the time it took users to modify the data.

If replication jobs run on an interval, meaning that there is a period of time after a replication job completes before the next replication job for the policy starts, the relationship between RTO and RPO depends on whether a replication job is running when the disaster occurs. If a job is in progress when a disaster occurs, the RTO is roughly equal to the RPO. However, if a job is not running when a disaster occurs, the RTO is negligible because the secondary cluster was not modified since the last replication job ran, and the failover process is almost instantaneous.

SyncIQ license functionality

You can replicate data to another Isilon cluster only if you activate a SyncIQ license on both the local cluster and the target cluster.

If a SyncIQ license becomes inactive, you cannot create, run, or manage replication policies. Also, all previously created replication policies are disabled. Replication policies that target the local cluster are also disabled. However, data that was previously replicated to the local cluster is still available.

Creating replication policies

You can create replication policies that determine when data is replicated with SyncIQ.

Excluding directories in replication

You can exclude directories from being replicated by replication policies even if the directories exist under the specified source directory.

Note

Failback is not supported for replication policies that exclude directories.

By default, all files and directories under the source directory of a replication policy are replicated to the target cluster. However, you can prevent directories under the source directory from being replicated.

If you specify a directory to exclude, files and directories under the excluded directory are not replicated to the target cluster. If you specify a directory to include, only the files and directories under the included directory are replicated to the target cluster; any directories that are not contained in an included directory are excluded.

If you both include and exclude directories, any excluded directories must be contained in one of the included directories; otherwise, the excluded-directory setting has no effect. For example, consider a policy with the following settings:

- The root directory is `/ifs/data`
- The included directories are `/ifs/data/media/music` and `/ifs/data/media/movies`
- The excluded directories are `/ifs/data/archive` and `/ifs/data/media/music/working`

In this example, the setting that excludes the `/ifs/data/archive` directory has no effect because the `/ifs/data/archive` directory is not under either of the included directories. The `/ifs/data/archive` directory is not replicated regardless of whether the directory is explicitly excluded. However, the setting that excludes the `/ifs/data/media/music/working` directory does have an effect, because the directory would be replicated if the setting was not specified.

In addition, if you exclude a directory that contains the source directory, the exclude-directory setting has no effect. For example, if the root directory of a policy is `/ifs/data`, explicitly excluding the `/ifs` directory does not prevent `/ifs/data` from being replicated.

Any directories that you explicitly include or exclude must be contained in or under the specified root directory. For example, consider a policy in which the specified root directory is `/ifs/data`. In this example, you could include both the `/ifs/data/media` and the `/ifs/data/users/` directories because they are under `/ifs/data`.

Excluding directories from a synchronization policy does not cause the directories to be deleted on the target cluster. For example, consider a replication policy that synchronizes `/ifs/data` on the source cluster to `/ifs/data` on the target cluster. If the policy excludes `/ifs/data/media` from replication, and `/ifs/data/media/file` exists on the target cluster, running the policy does not cause `/ifs/data/media/file` to be deleted from the target cluster.

Excluding files in replication

If you do not want specific files to be replicated by a replication policy, you can exclude them from the replication process through file-matching criteria statements. You can configure file-matching criteria statements during the replication-policy creation process.

Note

You cannot fail back replication policies that exclude files.

A file-criteria statement can include one or more elements. Each file-criteria element contains a file attribute, a comparison operator, and a comparison value. You can combine multiple criteria elements in a criteria statement with Boolean "AND" and "OR" operators. You can configure any number of file-criteria definitions.

Configuring file-criteria statements can cause the associated jobs to run slowly. It is recommended that you specify file-criteria statements in a replication policy only if necessary.

Modifying a file-criteria statement will cause a full replication to occur the next time that a replication policy is started. Depending on the amount of data being replicated, a full replication can take a very long time to complete.

For synchronization policies, if you modify the comparison operators or comparison values of a file attribute, and a file no longer matches the specified file-matching criteria, the file is deleted from the target the next time the job is run. This rule does not apply to copy policies.

File criteria options

You can configure a replication policy to exclude files that meet or do not meet specific criteria.

You can specify file criteria based on the following file attributes:

Date created

Includes or excludes files based on when the file was created. This option is available for copy policies only.
You can specify a relative date and time, such as "two weeks ago", or specific date and time, such as "January 1, 2012." Time settings are based on a 24-hour clock.

Date accessed

Includes or excludes files based on when the file was last accessed. This option is available for copy policies only, and only if the global access-time-tracking option of the cluster is enabled.
You can specify a relative date and time, such as "two weeks ago", or specific date and time, such as "January 1, 2012." Time settings are based on a 24-hour clock.

Date modified

Includes or excludes files based on when the file was last modified. This option is available for copy policies only.
You can specify a relative date and time, such as "two weeks ago", or specific date and time, such as "January 1, 2012." Time settings are based on a 24-hour clock.

File name

Includes or excludes files based on the file name. You can specify to include or exclude full or partial names that contain specific text.
The following wildcard characters are accepted:

Note

Alternatively, you can filter file names by using POSIX regular-expression (regex) text. Isilon clusters support IEEE Std 1003.2 (POSIX.2) regular expressions. For more information about POSIX regular expressions, see the BSD man pages.

Table 15 Replication file matching wildcards

Wildcard character	Description
`*`	Matches any string in place of the asterisk. For example, `m*` matches `movies` and `m123`.
`[ ]`	Matches any characters contained in the brackets, or a range of characters separated by a dash. For example, `b[aei]t` matches `bat`, `bet`, and `bit`. For example, `1[4-7]2` matches `142`, `152`, `162`, and `172`. You can exclude characters within brackets by following the first bracket with an exclamation mark. For example, `b[!ie]` matches `bat` but not `bit` or `bet`. You can match a bracket within a bracket if it is either the first or last character. For example, `[[c]at` matches `cat` and `[at`. You can match a dash within a bracket if it is either the first or last character. For example, `car[-s]` matches `cars` and `car-`.
`?`	Matches any character in place of the question mark. For example, `t?p` matches `tap`, `tip`, and `top`.

Path

Includes or excludes files based on the file path. This option is available for copy policies only.
You can specify to include or exclude full or partial paths that contain specified text. You can also include the wildcard characters `*`, `?`, and `[ ]`.

Size

Includes or excludes files based on their size.

Note

File sizes are represented in multiples of 1024, not 1000.

Type

Includes or excludes files based on one of the following file-system object types:

- Soft link
- Regular file
- Directory

Configure default replication policy settings

You can configure default settings for replication policies. If you do not modify these settings when creating a replication policy, the specified default settings are applied.

Procedure

1. Click **Data Protection** › **SyncIQ** › **Settings**.
2. In the **Default Policy Settings** section, if you want policies to connect only to nodes in a specified SmartConnect zone, select **Connect only to the nodes within the target cluster SmartConnect zone.**

   Note

   This option will affect only policies that specify the target cluster as a SmartConnect zone.

3. Specify which nodes you want replication policies to connect to when a policy is run.

Option	Description
Connect policies to all nodes on a source cluster.	Click **Run the policy on all nodes in this cluster.**
Connect policies only to nodes contained in a specified subnet and pool.	a. Click **Run the policy only on nodes in the specified subnet and pool.** b. From the **Subnet and pool** list, select the subnet and pool .

   Note

   SyncIQ does not support dynamically allocated IP address pools. If a replication job connects to a dynamically allocated IP address, SmartConnect might reassign the address while a replication job is running, which would disconnect the job and cause it to fail.

4. Click **Submit**.

Create a replication policy

You can create a replication policy with SyncIQ that defines how and when data is replicated to another Isilon cluster. Configuring a replication policy is a five-step process.

Configure replication policies carefully. If you modify any of the following policy settings after the policy is run, OneFS performs either a full or differential replication the next time the policy is run:

- Source directory
- Included or excluded directories
- File-criteria statement
- Target cluster name or address

  This applies only if you target a different cluster. If you modify the IP or domain name of a target cluster, and then modify the replication policy on the source cluster to match the new IP or domain name, a full replication is not performed.
- Target directory

Configure basic policy settings

You must configure basic settings for a replication policy.

Procedure

1. Click **Data Protection** › **SyncIQ** › **Policies**.
2. Click **Create a SyncIQ policy**.
3. In the **Settings** area, in the **Policy name** field, type a name for the replication policy.
4. (Optional) In the **Description** field, type a description for the replication policy.
5. In the **Action** area, specify the type of replication policy.
   - To copy all files from the source directory to the target directory, click **Copy**.

     Note

     Failback is not supported for copy policies.
   - To copy all files from the source directory to the target directory and delete any files on the target directory that are not in the source directory, click **Synchronize**.
6. In the **Run job** area, specify whether replication jobs will be run.

Option	Description
Run jobs only when manually initiated by a user.	Click **Only manually**.
Run jobs automatically according to a schedule.	a. Click **On a schedule**. b. Specify a schedule. If you configure a replication policy to run more than once a day, you cannot configure the interval to span across two calendar days. For example, you cannot configure a replication policy to run every hour starting at 7:00 PM and ending at 1:00 AM.
Run jobs automatically every time a change is made to the source directory.	Click **Whenever the source is modified**.

After you finish

The next step in the process of creating a replication policy is specifying source directories and files.

Specify source directories and files

You must specify the directories and files you want to replicate.

Procedure

1. In the **Source Cluster** area, in the **Source Root Directory** field, type the full path of the source directory that you want to replicate to the target cluster.

   You must specify a directory contained in `/ifs`. You cannot specify the `/ifs/.snapshot` directory or subdirectory of it.

2. (Optional) Prevent specific subdirectories of the root directory from being replicated.
   - To include a directory, in the **Included Directories** area, click **Add a directory path**.
   - To exclude a directory, in the **Excluded Directories** area, click **Add a directory path**.
3. (Optional) Prevent specific files from being replicated by specifying file matching criteria.
   1. In the **File Matching Criteria** area, select a filter type.
   2. Select an operator.
   3. Type a value.

   Files that do not meet the specified criteria will not be replicated to the target cluster. For example, if you specify `File Type doesn't match .txt`, SyncIQ will not replicate any files with the .txt file extension. If you specify `Created after 08/14/2013`, SyncIQ will not replicate any files created before August 14th, 2013. If you want to specify more than one file matching criterion, you can control how the criteria relate to each other by clicking either **Add an "Or" condition** or **Add an "And" condition**.

4. Specify which nodes you want the replication policy to connect to when the policy is run.

Option	Description
Connect the policy to all nodes in the source cluster.	Click **Run the policy on all nodes in this cluster.**
Connect the policy only to nodes contained in a specified subnet and pool.	a. Click **Run the policy only on nodes in the specified subnet and pool.** b. From the **Subnet and pool** list, select the subnet and pool .

Note

SyncIQ does not support dynamically allocated IP address pools. If a replication job connects to a dynamically allocated IP address, SmartConnect might reassign the address while a replication job is running, which would disconnect the job and cause it to fail.

After you finish

The next step in the process of creating a replication policy is specifying the target directory.

Specify the policy target directory

You must specify a target cluster and directory to replicate data to.

Procedure

1. In the **Target Cluster** area, in the **Target Host** field, type one of the following:
   - The fully qualified domain name of any node in the target cluster.
   - The host name of any node in the target cluster.
   - The name of a SmartConnect zone in the target cluster.
   - The IPv4 or IPv6 address of any node in the target cluster.
   - `localhost`

     This will replicate data to another directory on the local cluster.

   Note

   SyncIQ does not support dynamically allocated IP address pools. If a replication job connects to a dynamically allocated IP address, SmartConnect might reassign the address while a replication job is running, which would disconnect the job and cause it to fail.

2. In the **Target Directory** field, type the absolute path of the directory on the target cluster that you want to replicate data to.

   ⚠ CAUTION

   If you specify an existing directory on the target cluster, ensure that the directory is not the target of another replication policy. If this is a synchronization policy, ensure that the directory is empty. All files are deleted from the target of a synchronization policy the first time the policy is run.

   If the specified target directory does not already exist on the target cluster, the directory is created the first time the job is run. It is recommended that you do not specify the `/ifs` directory. If you specify the `/ifs` directory, the entire target cluster is set to a read-only state, preventing you from storing any other data on the cluster.

   If this is a copy policy, and files in the target directory share the same name as files in the source directory, the target directory files are overwritten when the job is run.

3. If you want replication jobs to connect only to the nodes included in the SmartConnect zone specified by the target cluster, click **Connect only to the nodes within the target cluster SmartConnect Zone**.

After you finish

The next step in the process of creating a replication policy is specifying policy target snapshot settings.

Configure policy target snapshot settings

You can optionally specify how archival snapshots are generated on the target cluster. You can access archival snapshots the same way that you access SnapshotIQ snapshots.

SyncIQ always retains one snapshot on the target cluster to facilitate failover, regardless of these settings.

Procedure

1. To create archival snapshots on the target cluster, in the **Target Snapshots** area, click **Capture snapshots on the target cluster**.
2. (Optional) To modify the default alias of the last snapshot created according to the replication policy, in the **Snapshot Alias Name** field, type a new alias.

   You can specify the alias name as a snapshot naming pattern. For example, the following naming pattern is valid:

   ```
   %{PolicyName}-on-%{SrcCluster}-latest
   ```

   The previous example produces names similar to the following:

   ```
   newPolicy-on-Cluster1-latest
   ```
3. (Optional) To modify the snapshot naming pattern for snapshots created according to the replication policy, in the **Snapshot Naming Pattern** field, type a naming pattern. Each snapshot generated for this replication policy is assigned a name based on this pattern.

   For example, the following naming pattern is valid:

   ```
   %{PolicyName}-from-%{SrcCluster}-at-%H:%M-on-%m-%d-%Y
   ```

   The example produces names similar to the following:

   ```
   newPolicy-from-Cluster1-at-10:30-on-7-12-2012
   ```
4. Select one of the following options:
   - Click **Snapshots do not expire**.
   - Click **Snapshots expire after...** and specify an expiration period.

After you finish

The next step in the process of creating a replication policy is configuring advanced policy settings.

Configure advanced policy settings

You can optionally configure advanced settings for a replication policy.

Procedure

1. (Optional) In the **Worker Threads Per Node** field, specify the maximum number of concurrent processes per node that will perform replication operations.

   Note

   Do not modify the default setting without consulting Isilon Technical Support.
2. (Optional) From the **Log Level** list, select the level of logging you want SyncIQ to perform for replication jobs.

   The following log levels are valid, listed from least to most verbose:
   - **Fatal**
   - **Error**
   - **Info**
   - **Copy**

- **Debug**
- **Trace**

Replication logs are typically used for debugging purposes. If necessary, you can log in to a node through the command-line interface and view the contents of the `/var/log/isi_migrate.log` file on the node.

Note

Notice is the recommended log level.

3. (Optional) If you want SyncIQ to perform a checksum on each file data packet that is affected by the replication policy, select the **Validate File Integrity** check box.

   If you enable this option, and the checksum values for a file data packet do not match, SyncIQ retransmits the affected packet.

4. (Optional) To modify the length of time SyncIQ retains replication reports for the policy, in the **Keep Reports For** area, specify a length of time.

   After the specified expiration period has passed for a report, SyncIQ automatically deletes the report.

   Some units of time are displayed differently when you view a report than how they were originally entered. Entering a number of days that is equal to a corresponding value in weeks, months, or years results in the larger unit of time being displayed. For example, if you enter a value of `7 days`, 1 week appears for that report after it is created. This change occurs because SyncIQ internally records report retention times in seconds and then converts them into days, weeks, months, or years.

5. (Optional) Specify whether to record information about files that are deleted by replication jobs by selecting one of the following options:

   - Click **Record when a synchronization deletes files or directories**.
   - Click **Do not record when a synchronization deletes files or directories**.

   This option is applicable for synchronization policies only.

After you finish

The next step in the process of creating a replication policy is saving the replication policy settings.

Save replication policy settings

SyncIQ does not create replication jobs for a replication policy until you save the policy.

Before you begin

Review the current settings of the replication policy. If necessary, modify the policy settings.

Procedure

1. Click **Create Policy**.

After you finish

You can increase the speed at which you can failback a replication policy by creating a SyncIQ domain for the source directory of the policy.

Create a SyncIQ domain

You can create a SyncIQ domain to increase the speed at which failback is performed for a replication policy. Because you can fail back only synchronization policies, it is not necessary to create SyncIQ domains for copy policies.

Failing back a replication policy requires that a SyncIQ domain be created for the source directory. OneFS automatically creates a SyncIQ domain during the failback process. However, if you intend on failing back a replication policy, it is recommended that you create a SyncIQ domain for the source directory of the replication policy while the directory is empty. Creating a domain for a directory that contains less data takes less time.

Procedure

1. Click **Cluster Management** › **Job Operations** › **Job Types**.
2. In the **Job Types** area, in the **DomainMark** row, from the **Actions** column, select **Start Job**.
3. In the **Domain Root Path** field, type the path of a source directory of a replication policy.
4. From the **Type of domain** list, select **SyncIQ**.
5. Ensure that the **Delete domain** check box is cleared.
6. Click **Start Job**.

Assess a replication policy

Before running a replication policy for the first time, you can view statistics on the files that would be affected by the replication without transferring any files. This can be useful if you want to preview the size of the data set that will be transferred if you run the policy.

Note

You can assess only replication policies that have never been run before.

Procedure

1. Click **Data Protection** › **SyncIQ** › **Policies**.
2. In the **SyncIQ Policies** table, in the row of a replication policy, from the **Actions** column, select **Assess Sync**.
3. Click **Data Protection** › **SyncIQ** › **Summary**.
4. After the job completes, in the **SyncIQ Recent Reports** table, in the row of the replication job, click **View Details**.

   The report displays the total amount of data that would have been transferred in the **Total Data** field.

Managing replication to remote clusters

You can manually run, view, assess, pause, resume, cancel, resolve, and reset replication jobs that target other clusters.

After a policy job starts, you can pause the job to suspend replication activities. Afterwards, you can resume the job, continuing replication from the point where the job was interrupted. You can also cancel a running or paused replication job if you want to free the cluster resources allocated for the job. A paused job reserves cluster resources whether or not the resources are in use. A cancelled job releases its cluster resources and allows another replication job to consume those resources. No more than five running and paused replication jobs can exist on a cluster at a time. However, an unlimited number of canceled replication jobs can exist on a cluster. If a replication job remains paused for more than a week, SyncIQ automatically cancels the job.

Start a replication job

You can manually start a replication job for a replication policy at any time.

If you want to replicate data according to an existing snapshot, at the OneFS command prompt, run the `isi sync jobs start` command with the `--source-snapshot` option. You cannot replicate data according to snapshots generated by SyncIQ.

Procedure

1. Click **Data Protection** › **SyncIQ** › **Policies**.
2. In the **SyncIQ Policies** table, in the **Actions** column for a job, select **Start Job**.

Pause a replication job

You can pause a running replication job and then resume the job later. Pausing a replication job temporarily stops data from being replicated, but does not free the cluster resources replicating the data.

Procedure

1. Click **Data Protection** › **SyncIQ** › **Summary**.
2. In the **Active Jobs** table, in the **Actions** column for a job, click **Pause Running Job**.

Resume a replication job

You can resume a paused replication job.

Procedure

1. Click **Data Protection** › **SyncIQ** › **Summary**.
2. In the **Currently Running** table, in the **Actions** column for a job, click **Resume Running Job**.

Cancel a replication job

You can cancel a running or paused replication job. Cancelling a replication job stops data from being replicated and frees the cluster resources that were replicating data. You cannot resume a cancelled replication job. To restart replication, you must start the replication policy again.

Procedure

1. Click **Data Protection** › **SyncIQ** › **Summary**.
2. In the **Active Jobs** table, in the **Actions** column for a job, click **Cancel Running Job**.

View active replication jobs

You can view information about replication jobs that are currently running or paused.

Procedure

1. Click **Data Protection** › **SyncIQ** › **Policies**.

2. In the **Active Jobs** table, review information about active replication jobs.

Replication job information

You can view information about replication jobs through the **Active Jobs** table.

Status
: The status of the job. The following job statuses are possible:

Running
: The job is currently running without error.

Paused
: The job has been temporarily paused.

Policy Name
: The name of the associated replication policy.

Started
: The time the job started.

Elapsed
: How much time has elapsed since the job started.

Transferred
: The number of files that have been transferred, and the total size of all transferred files.

Source Directory
: The path of the source directory on the source cluster.

Target Host
: The target directory on the target cluster.

Actions
: Displays any job-related actions that you can perform.

Initiating data failover and failback with SyncIQ

You can fail over from one Isilon cluster to another if, for example, a cluster becomes unavailable. You can then fail back to a primary cluster if the primary cluster becomes available again. You can revert failover if you decide that the failover was unnecessary, or if you failed over for testing purposes.

Note

Although you cannot fail over or fail back SmartLock directories, you can recover SmartLock directories on a target cluster. After you recover SmartLock directories, you can migrate them back to the source cluster.

Fail over data to a secondary cluster

You can fail over to a secondary Isilon cluster if, for example, a cluster becomes unavailable.

Before you begin

Create and successfully run a replication policy.

Complete the following procedure for each replication policy that you want to fail over.

Procedure

1. On the secondary Isilon cluster, click **Data Protection › SyncIQ › Local Targets**.
2. In the **SyncIQ Local Targets** table, in the row for a replication policy, from the **Actions** column, select **Allow Writes**.
3. On the primary cluster, modify the replication policy so that it is set to run only manually.

   This step will prevent the policy on the primary cluster from automatically running a replication job. If the policy on the primary cluster runs a replication job while writes are allowed to the target directory, the job will fail and the policy will be set to an unrunnable state. If this happens, modify the replication policy so that it is set to run only manually, resolve the policy, and complete the failback process. After you complete the failback process, you can modify the policy to run according to a schedule again.

After you finish

Direct clients to begin accessing the secondary cluster.

Revert a failover operation

Failover reversion undoes a failover operation on a secondary cluster, enabling you to replicate data from the primary cluster to the secondary cluster again. Failover reversion is useful if the primary cluster becomes available before data is modified on the secondary cluster or if you failed over to a secondary cluster for testing purposes.

Before you begin

Fail over a replication policy.

Reverting a failover operation does not migrate modified data back to the primary cluster. To migrate data that clients have modified on the secondary cluster, you must fail back to the primary cluster.

Complete the following procedure for each replication policy that you want to fail over:

Procedure

1. Click **Data Protection › SyncIQ › Local Targets**.
2. In the **SyncIQ Local Targets** table, in the row for a replication policy, from the **Actions** column, select **Disallow Writes**.

Fail back data to a primary cluster

After you fail over to a secondary cluster, you can fail back to the primary cluster.

Before you begin

Fail over a replication policy.

Procedure

1. On the primary cluster, click **Data Protection › SyncIQ › Policies**.
2. In the **SyncIQ Policies** table, in the row for a replication policy, from the **Actions** column, select **Resync-prep**.

   SyncIQ creates a mirror policy for each replication policy on the secondary cluster.

SyncIQ names mirror policies according to the following pattern:

```
<replication-policy-name>_mirror
```

3. On the secondary cluster, replicate data to the primary cluster by using the mirror policies.

   You can replicate data either by manually starting the mirror policies or by modifying the mirror policies and specifying a schedule.

4. Prevent clients from accessing the secondary cluster and then run each mirror policy again.

   To minimize impact to clients, it is recommended that you wait until client access is low before preventing client access to the cluster.

5. On the primary cluster, click **Data Protection** › **SyncIQ** › **Local Targets**.
6. In the **SyncIQ Local Targets** table, from the **Actions** column, select **Allow Writes** for each mirror policy.
7. On the secondary cluster, click **Data Protection** › **SyncIQ** › **Policies**.
8. In the **SyncIQ Policies** table, from the **Actions** column, select **Resync-prep** for each mirror policy.

After you finish

Redirect clients to begin accessing the primary cluster.

Performing disaster recovery for SmartLock directories

Although you cannot fail over or fail back SmartLock directories, you can recover SmartLock directories on a target cluster. After you recover SmartLock directories, you can migrate them back to the source cluster.

Recover SmartLock directories on a target cluster

You can recover SmartLock directories that you have replicated to a target cluster.

Before you begin

Create and successfully run a replication policy.

Complete the following procedure for each SmartLock directory that you want to recover.

Procedure

1. On the target cluster, click **Data Protection** › **SyncIQ** › **Local Targets**.
2. In the **SyncIQ Local Targets** table, in the row of the replication policy, enable writes to the target directory of the policy.
   - If the last replication job completed successfully and a replication job is not currently running, select **Allow Writes**.
   - If a replication job is currently running, wait until the replication job completes, and then select **Allow Writes**.
   - If the primary cluster became unavailable while a replication job was running, select **Break Association**.
3. If you clicked **Break Association**, restore any files that are left in an inconsistent state.

a. Delete all files that are not committed to a WORM state from the target directory.

b. Copy all files from the failover snapshot to the target directory.

   Failover snapshots are named according to the following naming pattern:

   ```
   SIQ-Failover-<policy-name>-<year>-<month>-<day>_<hour>-<minute>-
   <second>
   ```

   Snapshots are stored in the `/ifs/.snapshot` directory.

4. If any SmartLock directory configuration settings, such as an autocommit time period, were specified for the source directory of the replication policy, apply those settings to the target directory.

   Because autocommit information is not transferred to the target cluster, files that were scheduled to be committed to a WORM state on the source cluster will not be scheduled to be committed at the same time on the target cluster. To ensure that all files are retained for the appropriate time period, you can commit all files in target SmartLock directories to a WORM state. For example, the following command automatically commits all files in `/ifs/data/smartlock` to a WORM state after one minute.

   ```
   isi smartlock modify --path /ifs/data/smartlock --autocommit 1n
   ```

After you finish

Redirect clients to begin accessing the target cluster.

Migrate SmartLock directories

You might want to migrate SmartLock directories if you restored the directories on a target cluster, and want to transfer those directories either back to the source cluster or to a new cluster.

Procedure

1. On a cluster, create a replication policy for each directory that you want to migrate.

   The policies must meet the following requirements:

   - The source directory is the SmartLock directory that you are migrating.
   - The target directory is an empty SmartLock directory. The source and target directories must be of the same SmartLock type. For example, if the target directory is a compliance directory, the source must also be a compliance directory.

2. Replicate data to the target cluster by running the policies you created.

   You can replicate data either by manually starting the policies or by specifying a policy schedule.

3. (Optional) To ensure that SmartLock protection is enforced for all files, commit all files in the SmartLock source directory to a WORM state.

   Because autocommit information is not transferred to the target cluster, files that were scheduled to be committed to a WORM state on the source cluster will not be scheduled to be committed at the same time on the target cluster. To ensure that all files are retained for the appropriate time period, you can commit all files in target SmartLock directories to a WORM state.

For example, the following command automatically commits all files in `/ifs/data/smartlock` to a WORM state after one minute:

```
isi smartlock modify --path /ifs/data/smartlock --autocommit 1n
```

This step is unnecessary if you have not configured an autocommit time period for the SmartLock directory being replicated.

4. Prevent clients from accessing the source cluster and run the policy that you created.

   To minimize impact to clients, it is recommended that you wait until client access is low before preventing client access to the cluster.

5. On the target cluster, click **Data Protection** › **SyncIQ** › **Local Targets**.
6. In the **SyncIQ Local Targets** table, in the row of each replication policy, from the **Actions** column, select **Allow Writes**.
7. (Optional) If any SmartLock directory configuration settings, such as an autocommit time period, were specified for the source directories of the replication policies, apply those settings to the target directories.
8. (Optional) Delete the copy of your SmartLock data on the source cluster.

   If the SmartLock directories are compliance directories or enterprise directories with the privileged delete functionality permanently disabled, you cannot recover the space consumed by the source SmartLock directories until all files are released from a WORM state. If you want to free the space before files are released from a WORM state, contact Isilon Technical Support for information about reformatting your cluster.

Managing replication policies

You can modify, view, enable and disable replication policies.

Modify a replication policy

You can modify the settings of a replication policy.

If you modify any of the following policy settings after a policy runs, OneFS performs either a full or differential replication the next time the policy runs:

- Source directory
- Included or excluded directories
- File-criteria statement
- Target cluster
  This applies only if you target a different cluster. If you modify the IP or domain name of a target cluster, and then modify the replication policy on the source cluster to match the new IP or domain name, a full replication is not performed.
- Target directory

Procedure

1. Click **Data Protection** › **SyncIQ** › **Policies**.
2. In the **SyncIQ Policies** table, in the row for a policy, click **View/Edit**.
3. In the **View SyncIQ Policy Details** dialog box, click **Edit Policy**.
4. Modify the settings of the replication policy, and then click **Save Changes**.

Delete a replication policy

You can delete a replication policy. Once a policy is deleted, SyncIQ no longer creates replication jobs for the policy. Deleting a replication policy breaks the target association on the target cluster, and allows writes to the target directory.

If you want to temporarily suspend a replication policy from creating replication jobs, you can disable the policy, and then enable the policy again later.

Procedure

1. Click **Data Protection** › **SyncIQ** › **Policies**.
2. In the **SyncIQ Policies** table, in the row for a policy, select **Delete Policy**.
3. In the confirmation dialog box, click **Delete**.

   Note

   The operation will not succeed until SyncIQ can communicate with the target cluster; until then, the policy will not be removed from the **SyncIQ Policies** table. After the connection between the source cluster and target cluster is reestablished, SyncIQ will delete the policy the next time that the job is scheduled to run; if the policy is configured to run only manually, you must manually run the policy again. If SyncIQ is permanently unable to communicate with the target cluster, run the `isi sync policies delete` command with the `--local-only` option. This will delete the policy from the local cluster only and not break the target association on the target cluster. For more information, see the *OneFS CLI Administration Guide*.

Enable or disable a replication policy

You can temporarily suspend a replication policy from creating replication jobs, and then enable it again later.

Note

If you disable a replication policy while an associated replication job is running, the running job is not interrupted. However, the policy will not create another job until the policy is enabled.

Procedure

1. Click **Data Protection** › **SyncIQ** › **Policies**.
2. In the **SyncIQ Policies** table, in the row for a replication policy, select either **Enable Policy** or **Disable Policy**.

   If neither **Enable Policy** nor **Disable Policy** appears, verify that a replication job is not running for the policy. If an associated replication job is not running, ensure that the SyncIQ license is active on the cluster.

View replication policies

You can view information about replication policies.

Procedure

1. Click **Data Protection** › **SyncIQ** › **Policies**.

2. In the **SyncIQ Policies** table, review information about replication policies.

Replication policy information

You can view information about replication policies through the **SyncIQ Policies** table.

Policy Name
: The name of the policy.

State
: Whether the policy is enabled or disabled.

Last Known Good
: When the last successful job ran.

Schedule
: When the next job is scheduled to run. A value of **Manual** indicates that the job can be run only manually. A value of **When source is modified** indicates that the job will be run whenever changes are made to the source directory.

Source Directory
: The path of the source directory on the source cluster.

Target Host : Directory
: The IP address or fully qualified domain name of the target cluster and the full path of the target directory.

Actions
: Any policy-related actions that you can perform.

Replication policy settings

You configure replication policies to run according to replication policy settings.

Policy name
: The name of the policy.

Description
: Describes the policy. For example, the description might explain the purpose or function of the policy.

Enabled
: Determines whether the policy is enabled.

Action
: Determines the how the policy replicates data. All policies copy files from the source directory to the target directory and update files in the target directory to match files on the source directory. The action determines how deleting a file on the source directory affects the target. The following values are valid:

  Copy
  : If a file is deleted in the source directory, the file is not deleted in the target directory.

  Synchronize
  : Deletes files in the target directory if they are no longer present on the source. This ensures that an exact replica of the source directory is maintained on the target cluster.

Run job
: Determines whether jobs are run automatically according to a schedule or only when manually specified by a user.

Last Successful Run
Displays the last time that a replication job for the policy completed successfully.

Last Started
Displays the last time that the policy was run.

Source Root Directory
The full path of the source directory. Data is replicated from the source directory to the target directory.

Included Directories
Determines which directories are included in replication. If one or more directories are specified by this setting, any directories that are not specified are not replicated.

Excluded Directories
Determines which directories are excluded from replication. Any directories specified by this setting are not replicated.

File Matching Criteria
Determines which files are excluded from replication. Any files that do not meet the specified criteria are not replicated.

Restrict Source Nodes
Determines whether the policy can run on all nodes on the source cluster or run only on specific nodes.

Target Host
The IP address or fully qualified domain name of the target cluster.

Target Directory
The full path of the target directory. Data is replicated to the target directory from the source directory.

Restrict Target Nodes
Determines whether the policy can connect to all nodes on the target cluster or can connect only to specific nodes.

Capture Snapshots
Determines whether archival snapshots are generated on the target cluster.

Snapshot Alias Name
Specifies an alias for the latest archival snapshot taken on the target cluster.

Snapshot Naming Pattern
Specifies how archival snapshots are named on the target cluster.

Snapshot Expiration
Specifies how long archival snapshots are retained on the target cluster before they are automatically deleted by the system.

Workers Threads Per Node
Specifies the number of workers per node that are generated by OneFS to perform each replication job for the policy.

Log Level

Specifies the amount of information that is recorded for replication jobs.
More verbose options include all information from less verbose options. The following list describes the log levels from least to most verbose:

- **Fatal**
- **Error**
- **Notice**
- **Info**
- **Copy**
- **Debug**
- **Trace**

Replication logs are typically used for debugging purposes. If necessary, you can log in to a node through the command-line interface and view the contents of the `/var/log/isi_migrate.log` file on the node.

Note

Notice is the recommended log level.

Validate File Integrity

Determines whether OneFS performs a checksum on each file data packet that is affected by a replication job. If a checksum value does not match, OneFS retransmits the affected file data packet.

Keep Reports For

Specifies how long replication reports are kept before they are automatically deleted by OneFS.

Log Deletions on Synchronization

Determines whether OneFS records when a synchronization job deletes files or directories on the target cluster.

The following replication policy fields are available only through the OneFS command-line interface.

Source Subnet

Specifies whether replication jobs connect to any nodes in the cluster or if jobs can connect only to nodes in a specified subnet.

Source Pool

Specifies whether replication jobs connect to any nodes in the cluster or if jobs can connect only to nodes in a specified pool.

Password Set

Specifies a password to access the target cluster.

Report Max Count

Specifies the maximum number of replication reports that are retained for this policy.

Target Compare Initial Sync

Determines whether full or differential replications are performed for this policy. Full or differential replications are performed the first time a policy is run and after a policy is reset.

Source Snapshot Archive

Determines whether snapshots generated for the replication policy on the source cluster are deleted when the next replication policy is run. Enabling archival source snapshots does not require you to activate the SnapshotIQ license on the cluster.

Source Snapshot Pattern

If snapshots generated for the replication policy on the source cluster are retained, renames snapshots according to the specified rename pattern.

Source Snapshot Expiration

If snapshots generated for the replication policy on the source cluster are retained, specifies an expiration period for the snapshots.

Restrict Target Network

Determines whether replication jobs connect only to nodes in a given SmartConnect zone. This setting applies only if the Target Host is specified as a SmartConnect zone.

Target Detect Modifications

Determines whether SyncIQ checks the target directory for modifications before replicating files. By default, SyncIQ always checks for modifications.

Note

Disabling this option could result in data loss. It is recommended that you consult Isilon Technical Support before disabling this option.

Resolve

Determines whether you can manually resolve the policy if a replication job encounters an error.

Managing replication to the local cluster

You can interrupt replication jobs that target the local cluster.

You can cancel a currently running job that targets the local cluster, or you can break the association between a policy and its specified target. Breaking a source and target cluster association causes SyncIQ to perform a full replication the next time the policy is run.

Cancel replication to the local cluster

You can cancel a replication job that is targeting the local cluster.

Procedure

1. Click **Data Protection** › **SyncIQ** › **Local Targets**.
2. In the **SyncIQ Local Targets** table, specify whether to cancel a specific replication job or all replication jobs targeting the local cluster.
   - To cancel a specific job, in the row for a replication job, select **Cancel Running Job**.
   - To cancel all jobs targeting the local cluster, select the check box to the left of **Policy Name** and then select **Cancel Selection** from the **Select a bulk action** list.

Break local target association

You can break the association between a replication policy and the local cluster. Breaking the target association will allow writes to the target directory but will also require you to reset the replication policy before you can run the policy again.

⚠ CAUTION

After a replication policy is reset, SyncIQ performs a full or differential replication the next time the policy is run. Depending on the amount of data being replicated, a full or differential replication can take a very long time to complete.

Procedure

1. Click **Data Protection** › **SyncIQ** › **Local Targets**.
2. In the **SyncIQ Local Targets** table, in the row for a replication policy, select **Break Association**.
3. In the **Confirm** dialog box, click **Yes**.

View replication policies targeting the local cluster

You can view information about replication policies that are currently replicating data to the local cluster.

Procedure

1. Click **Data Protection** › **SyncIQ** › **Local Targets**.
2. In the **SyncIQ Local Targets** table, view information about replication policies.

Remote replication policy information

You can view information about replication policies that are currently targeting the local cluster.

The following information is displayed in the **SyncIQ Local Targets** table:

ID
: The ID of the replication policy.

Policy Name
: The name of the replication policy.

Source Host
: The name of the source cluster.

Source Cluster GUID
: The GUID of the source cluster.

Coordinator IP
: The IP address of the node on the source cluster that is acting as the job coordinator.

Updated
: The time when data about the policy or job was last collected from the source cluster.

Target Path
: The path of the target directory on the target cluster.

Status
: The current status of the replication job.

Actions

Displays any job-related actions that you can perform.

Managing replication performance rules

You can manage the impact of replication on cluster performance by creating rules that limit the network traffic created and the rate at which files are sent by replication jobs.

Create a network traffic rule

You can create a network traffic rule that limits the amount of network traffic that replication policies are allowed to generate during a specified time period.

Procedure

1. Click **Data Protection › SyncIQ › Performance Rules**.
2. Click **Create a SyncIQ Performance Rule**.
3. From the **Rule Type** list, select **Bandwidth**.
4. In the **Limit** field, specify the maximum number of bytes per second that replication policies are allowed to send.
5. In the **Schedule** area, specify the time and days of the week that you want to apply the rule.
6. Click **Create Performance Rule**.

Create a file operations rule

You can create a file-operations rule that limits the number of files that replication jobs can send per second.

Procedure

1. Click **Data Protection › SyncIQ › Performance Rules**.
2. Click **Create a SyncIQ Performance Rule**.
3. From the **Rule Type** list, select **Bandwidth**.
4. In the **Limit** field, specify the maximum number of files per second that replication policies are allowed to send.
5. In the **Schedule** area, specify the time and days of the week that you want to apply the rule.
6. Click **Create Performance Rule**.

Modify a performance rule

You can modify a performance rule.

Procedure

1. Click **Data Protection › SyncIQ › Performance Rules**.
2. In the **SyncIQ Performance Rules**, in the row for the rule you want to modify, click **View/Edit**.
3. Click **Edit Performance Rule**.
4. Modify rule settings, and then click **Save Changes**.

Delete a performance rule

You can delete a performance rule.

Procedure

1. Click **Data Protection** › **SyncIQ** › **Performance Rules**.
2. In the **SyncIQ Performance Rules** table, in the row for the rule you want to delete, select **Delete Rule**.
3. In the **Confirm Delete** dialog box, click **Delete**.

Enable or disable a performance rule

You can disable a performance rule to temporarily prevent the rule from being enforced. You can also enable a performance rule after it has been disabled.

Procedure

1. Click **Data Protection** › **SyncIQ** › **Performance Rules**.
2. In the **SyncIQ Performance Rules** table, in the row for a rule you want to enable or disable, select either **Enable Rule** or **Disable Rule**.

View performance rules

You can view information about replication performance rules.

Procedure

1. Click **Data Protection** › **SyncIQ** › **Performance Rules**.
2. In the **SyncIQ Performance Rules** table, view information about performance rules.

Managing replication reports

In addition to viewing replication reports, you can configure how long reports are retained on the cluster. You can also delete any reports that have passed their expiration period.

Configure default replication report settings

You can configure the default amount of time that SyncIQ retains replication reports for. You can also configure the maximum number of reports that SyncIQ retains for each replication policy.

Procedure

1. Click **Data Protection** › **SyncIQ** › **Settings**.
2. In the **Report Settings** area, in the **Keep Reports For** area, specify how long you want to retain replication reports for.

   After the specified expiration period has passed for a report, SyncIQ automatically deletes the report.

   Some units of time are displayed differently when you view a report than how you originally enter them. Entering a number of days that is equal to a corresponding value in weeks, months, or years results in the larger unit of time being displayed. For example, if you enter a value of 7 days, 1 week appears for that report after it is created. This change occurs because SyncIQ internally records report retention times in seconds and then converts them into days, weeks, months, or years for display.

3. In the **Number of Reports to Keep Per Policy** field, type the maximum number of reports you want to retain at a time for a replication policy.
4. Click **Submit**.

Delete replication reports

Replication reports are routinely deleted by SyncIQ after the expiration date for the reports has passed. SyncIQ also deletes reports after the number of reports exceeds the specified limit. Excess reports are periodically deleted by SyncIQ; however, you can manually delete all excess replication reports at any time. This procedure is available only through the command-line interface (CLI).

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Delete excess replication reports by running the following command:

```
isi sync reports rotate
```

View replication reports

You can view replication reports and subreports.

Procedure

1. Click **Data Protection** › **SyncIQ** › **Reports**.
2. In the **SyncIQ Reports** table, in the row for a report, click **View Details**.

   If a report is composed of subreports, the report is displayed as a folder. Subreports are displayed as files within report folders.

Replication report information

You can view information about replication jobs through the **Reports** table.

Policy Name

The name of the associated policy for the job. You can view or edit settings for the policy by clicking the policy name.

Status

Displays the status of the job. The following job statuses are possible:

Running

The job is currently running without error.

Paused

The job has been temporarily paused.

Finished

The job completed successfully.

Failed

The job failed to complete.

Started

Indicates when the job started.

Ended

Indicates when the job ended.

Duration
: Indicates how long the job took to complete.

Transferred
: The total number of files that were transferred during the job run, and the total size of all transferred files. For assessed policies, `Assessment` appears.

Source Directory
: The path of the source directory on the source cluster.

Target Host
: The IP address or fully qualified domain name of the target cluster.

Action
: Displays any report-related actions that you can perform.

Managing failed replication jobs

If a replication job fails due to an error, SyncIQ might disable the corresponding replication policy. For example SyncIQ might disable a replication policy if the IP or hostname of the target cluster is modified. If a replication policy is disabled, the policy cannot be run.

To resume replication for a disabled policy, you must either fix the error that caused the policy to be disabled, or reset the replication policy. It is recommended that you attempt to fix the issue rather than reset the policy. If you believe you have fixed the error, you can return the replication policy to an enabled state by resolving the policy. You can then run the policy again to test whether the issue was fixed. If you are unable to fix the issue, you can reset the replication policy. However, resetting the policy causes a full or differential replication to be performed the next time the policy is run.

Note

Depending on the amount of data being synchronized or copied, a full and differential replications can take a very long time to complete.

Resolve a replication policy

If SyncIQ disables a replication policy due to a replication error, and you fix the issue that caused the error, you can resolve the replication policy. Resolving a replication policy enables you to run the policy again. If you cannot resolve the issue that caused the error, you can reset the replication policy.

Procedure

1. Click **Data Protection** › **SyncIQ** › **Policies**.
2. In the **Policies** table, in the row for a policy, select **Resolve**.

Reset a replication policy

If a replication job encounters an error that you cannot resolve, you can reset the corresponding replication policy. Resetting a policy causes OneFS to perform a full or differential replication the next time the policy is run. Resetting a replication policy deletes the latest snapshot generated for the policy on the source cluster.

⚠ CAUTION

Depending on the amount of data being replicated, a full or differential replication can take a very long time to complete. Reset a replication policy only if you cannot fix the issue that caused the replication error. If you fix the issue that caused the error, resolve the policy instead of resetting the policy.

Procedure

1. Click **Data Protection** › **SyncIQ** › **Policies**.
2. In the **SyncIQ Policies** table, in the row for a policy, select **Reset Sync State**.

Perform a full or differential replication

After you reset a replication policy, you must perform either a full or differential replication.

Before you begin

Reset a replication policy.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in through the root or compliance administrator account.
2. Specify the type of replication you want to perform by running the `isi sync policies modify` command.
   - To perform a full replication, disable the `--target-compare-initial-sync` option.

     For example, the following command disables differential synchronization for newPolicy:

     ```
     isi sync policies modify newPolicy \
     --target-compare-initial-sync false
     ```
   - To perform a differential replication, enable the `--target-compare-initial-sync` option.

     For example, the following command enables differential synchronization for newPolicy:

     ```
     isi sync policies modify newPolicy \
     --target-compare-initial-sync true
     ```
3. Run the policy by running the `isi sync jobs start` command.

   For example, the following command runs newPolicy:

   ```
   isi sync jobs start newPolicy
   ```

Managing changelists

You can create and view changelists that describe what data was modified by a replication job. Changelists are most commonly accessed by applications through the OneFS Platform API.

To create a changelist, you must enable changelists for a replication policy. If changelists are enabled for a policy, SyncIQ does not automatically delete the repstate files generated by the policy; if changelists are not enabled for a policy, SyncIQ automatically deletes the repstate files after the corresponding replication jobs complete. SyncIQ generates one repstate file for each replication job. Because a large amount of repstate files can consume a large amount of disk space, it is recommended that you do not enable changelists for a policy unless it is necessary for your workflow.

If changelists are enabled for a policy, SyncIQ does not automatically delete source cluster snapshots for the policy. To create a changelist, you must have access to two consecutive snapshots and an associated repstate generated by a replication policy.

Create a changelist

You can create a changelist to view what data was modified by a replication job.

Before you begin

Through the OneFS command line, enable changelists for a replication policy, and then run the policy at least twice. The following command enables changelists for newPolicy:

```
isi sync policies modify newPolicy --changelist true
```

Note

You can enable changelists only through the command-line interface (CLI).

Procedure

1. (Optional) Record the IDs of the snapshots generated by the replication policy.
   a. Click **File System Management** › **SnapshotIQ** › **Snapshots**.
   b. In the row of the snapshots that were created for the replication policies that you want to create a change list for, click, **View Details** and record the IDs of the snapshots

      The snapshots must have been generated sequentially for the same replication policy. Changelist snapshots are generated according to the following snapshot naming convention:

      ```
      SIQ-Changelist-<policy-name>-<date>
      ```

      If source-archival snapshots are enabled for the policy, the change-list snapshots are named according to the naming convention specified by the policy.
2. Click **Cluster Management** › **Job Operations** › **Job Types**.
3. In the **Job Types** area, in the **ChangelistCreate** row, from the **Actions** column, select **Start Job**.
4. In the **Older Snapshot ID** field, type the ID of the snapshot generated when the replication policy started.
5. In the **Newer Snapshot ID** field, type the ID of the snapshot generated when the replication policy ended.
6. (Optional) If you want to recreate the changelist later, select **Retain the replication record after a changelist is created**. If you do not select this option, the repstate file used to generate the changelist is deleted after the changelist is created.
7. Click **Start Job**.

View a changelist

You can view a changelist that describes what data was modified by a replication job. This procedure is available only through the command-line interface (CLI).

Procedure

1. View the IDs of changelists by running the following command:

   ```
   isi_changelist_mod -l
   ```

   Changelist IDs include the IDs of both snapshots used to create the changelist. If OneFS is still in the process of creating a changelist, `inprog` is appended to the changelist ID.

2. (Optional) View all contents of a changelist by running the `isi_changelist_mod` command with the `-a` option.

   The following command displays the contents of a changelist named 2_6:

   ```
   isi_changelist_mod -a 2_6
   ```

3. View a specific changelist entry by running the `isi_changelist_mod` command with the `-g` option.

   The following command displays an entry with a LIN of 1003402c3 from a changelist named 2_6:

   ```
   isi_changelist_mod -g 2_6 1003402c3
   ```

Changelist information

You can view the information contained in changelists.

Note

The information contained in changelists is meant to be consumed by applications through the OneFS Platform API. The information might be less useful when consumed through the command-line interface (CLI).

The following information is displayed in the output of the `isi_changelist_mod` command:

lin
: The LIN of the changelist entry. Metadata entries are assigned a LIN of 1.

entry_type
: The type of changelist entry. The field is set to either `metadata` or `file`.

size
: The total size of the changelist entry, in bytes.

reserved
: This field is not currently used by changelists.

root_path
: The root path of the snapshots used to create the changelist.

owning_job_id
: The ID of the ChangelistCreate job that created the changelist.

num_cl_entries
: The number of changelist entries in the changelist.

root_path_size
The total size of the null-terminated UTF-8 string that contains the root path of the snapshots, in bytes.

root_path_offset
The number of bytes between the start of the changelist entry structure and the null-terminated UTF-8 string that contains the root path of the snapshots.

path
The path, relative to the root path, of the file or directory that was modified or removed.

type
If an item was modified, describes the type of item that was modified. The following types of items might have been modified:

regular
A regular file was modified.

directory
A directory was modified.

symlink
A symbolic link was modified.

fifo
A first-in-first-out (FIFO) queue was modified.

socket
A Unix domain socket was modified.

char device
A character device was modified.

block device
A block device was modified.

unknown
An unknown type of file was modified.

If any type of item was removed, this field is set to `(REMOVED)`.

size
The size of the item that was modified, in bytes. If an item was removed, this field is set to `0`.

path_size
The total size of the null-terminated UTF-8 string that contains the path, relative to the root path, of the file or directory that was modified or removed, in bytes.

path_offset
The number of bytes between the start of the changelist entry structure and the path, relative to the root path, of the file or directory that was modified or removed.

atime
The POSIX timestamp of when the item was last accessed.

atimensec
The number of nanoseconds past the atime that the item was last accessed.

ctime
The POSIX timestamp of when the item was last changed.

ctimensec
The number of nanoseconds past the ctime that the item was last changed.

mtime

The POSIX timestamp of when the item was last modified.

mtimensec

The number of nanoseconds past the mtime that the item was last modified.

CHAPTER 13

Data layout with FlexProtect

This section contains the following topics:

FlexProtect overview

An Isilon cluster is designed to continuously serve data, even when one or more components simultaneously fail. OneFS ensures data availability by striping or mirroring data across the cluster. If a cluster component fails, data stored on the failed component is available on another component. After a component failure, lost data is restored on healthy components by the FlexProtect proprietary system.

Data protection is specified at the file level, not the block level, enabling the system to recover data quickly. Because all data, metadata, and parity information is distributed across all nodes, the cluster does not require a dedicated parity node or drive. This ensures that no single node limits the speed of the rebuild process.

File striping

OneFS uses the internal network to automatically allocate and stripe data across nodes and disks in the cluster. OneFS protects data as the data is being written. No separate action is necessary to stripe data.

OneFS breaks files into smaller logical chunks called stripes before writing the files to disk; the size of each file chunk is referred to as the stripe unit size. Each OneFS block is 8 KB, and a stripe unit consists of 16 blocks, for a total of 128 KB per stripe unit. During a write, OneFS breaks data into stripes and then logically places the data in a stripe unit. As OneFS stripes data across the cluster, OneFS fills the stripe unit according to the number of nodes and protection level.

OneFS can continuously reallocate data and make storage space more usable and efficient. As the cluster size increases, OneFS stores large files more efficiently.

Requested data protection

The requested protection of data determines the amount of redundant data created on the cluster to ensure that data is protected against component failures. OneFS enables you to modify the requested protection in real time while clients are reading and writing data on the cluster.

OneFS provides several data protection settings. You can modify these protection settings at any time without rebooting or taking the cluster or file system offline. When planning your storage solution, keep in mind that increasing the requested protection reduces write performance and requires additional storage space for the increased number of nodes.

OneFS uses the Reed Solomon algorithm for N+M protection. In the N+M data protection model, N represents the number of data-stripe units, and M represents the number of simultaneous node or drive failures—or a combination of node and drive failures—that the cluster can withstand without incurring data loss. N must be larger than M.

In addition to N+M data protection, OneFS also supports data mirroring from 2x to 8x, allowing from two to eight mirrors of data. In terms of overall cluster performance and resource consumption, N+M protection is often more efficient than mirrored protection. However, because read and write performance is reduced for N+M protection, data mirroring might be faster for data that is updated often and is small in size. Data mirroring requires significant overhead and might not always be the best data-protection method. For example, if you enable 3x mirroring, the specified content is duplicated three times on the cluster; depending on the amount of content mirrored, this can consume a significant amount of storage space.

FlexProtect data recovery

OneFS uses the FlexProtect proprietary system to detect and repair files and directories that are in a degraded state due to node or drive failures.

OneFS protects data in the cluster based on the configured protection policy. OneFS rebuilds failed disks, uses free storage space across the entire cluster to further prevent data loss, monitors data, and migrates data off of at-risk components.

OneFS distributes all data and error-correction information across the cluster and ensures that all data remains intact and accessible even in the event of simultaneous component failures. Under normal operating conditions, all data on the cluster is protected against one or more failures of a node or drive. However, if a node or drive fails, the cluster protection status is considered to be in a degraded state until the data is protected by OneFS again. OneFS reprotects data by rebuilding data in the free space of the cluster. While the protection status is in a degraded state, data is more vulnerable to data loss.

Because data is rebuilt in the free space of the cluster, the cluster does not require a dedicated hot-spare node or drive in order to recover from a component failure. Because a certain amount of free space is required to rebuild data, it is recommended that you reserve adequate free space through the virtual hot spare feature.

As you add more nodes, the cluster gains more CPU, memory, and disks to use during recovery operations. As a cluster grows larger, data restriping operations become faster.

Smartfail

OneFS protects data stored on failing nodes or drives through a process called smartfailing.

During the smartfail process, OneFS places a device into quarantine. Data stored on quarantined devices is read only. While a device is quarantined, OneFS reprotects the data on the device by distributing the data to other devices. After all data migration is complete, OneFS logically removes the device from the cluster, the cluster logically changes its width to the new configuration, and the node or drive can be physically replaced.

OneFS smartfails devices only as a last resort. Although you can manually smartfail nodes or drives, it is recommended that you first consult Isilon Technical Support.

Occasionally a device might fail before OneFS detects a problem. If a drive fails without being smartfailed, OneFS automatically starts rebuilding the data to available free space on the cluster. However, because a node might recover from a failure, if a node fails, OneFS does not start rebuilding data unless the node is logically removed from the cluster.

Node failures

Because node loss is often a temporary issue, OneFS does not automatically start reprotecting data when a node fails or goes offline. If a node reboots, the file system does not need to be rebuilt because it remains intact during the temporary failure.

If you configure N+1 data protection on a cluster, and one node fails, all of the data is still accessible from every other node in the cluster. If the node comes back online, the node rejoins the cluster automatically without requiring a full rebuild.

To ensure that data remains protected, if you physically remove a node from the cluster, you must also logically remove the node from the cluster. After you logically remove a node, the node automatically reformats its own drives, and resets itself to the factory default settings. The reset occurs only after OneFS has confirmed that all data has been reprotected. You can logically remove a node using the smartfail process. It is important that you smartfail nodes only when you want to permanently remove a node from the cluster.

If you remove a failed node before adding a new node, data stored on the failed node must be rebuilt in the free space in the cluster. After the new node is added, OneFS distributes the data to the new node. It is more efficient to add a replacement node to the cluster before failing the old node because OneFS can immediately use the replacement node to rebuild the data stored on the failed node.

Requesting data protection

You can request the protection of a file or directory by setting its requested protection. This flexibility enables you to protect distinct sets of data at different levels.

The default requested protection of node pools is N+2:1, which means that two drives or one node can fail without causing any data loss. For clusters or node pools containing less than two petabytes or fewer than 16 nodes, N+2:1 is the recommended requested protection. However, if the cluster or node pool is larger, you might consider higher requested protection.

OneFS allows you to request protection that the cluster is currently incapable of matching. If you request an unmatchable protection, the cluster will continue trying to match the requested protection until a match is possible. For example, in a four-node cluster, you might request a protection of 5x. In this example, OneFS would protect the data at 4x until you added a fifth node to the cluster, at which point OneFS would reprotect the data at the 5x.

Note

For 4U Isilon IQ X-Series and NL-Series nodes, and IQ 12000X/EX 12000 combination platforms, the minimum cluster size of three nodes requires a minimum of N+2:1.

Requested protection settings

Requested protection settings determine the level of hardware failure that a cluster can recover from without suffering data loss.

Requested protection setting	Minimum number of nodes required	Definition
[+1n]	3	The cluster can recover from one drive or node failure without sustaining any data loss.
[+2d:1n]	3	The cluster can recover from two simultaneous drive failures or one node failure without sustaining any data loss.
[+2n]	4	The cluster can recover from two simultaneous drive or node failures without sustaining any data loss.
[+3d:1n]	3	The cluster can recover from three simultaneous drive failures or one

Requested protection setting	Minimum number of nodes required	Definition
		node failure without sustaining any data loss.
[+3d:1n1d]	3	The cluster can recover from three simultaneous drive failures or simultaneous failures of one node and one drive without sustaining any data loss.
[+3n]	6	The cluster can recover from three simultaneous drive or node failures without sustaining any data loss.
[+4d:1n]	3	The cluster can recover from four simultaneous drive failures or one node failure without sustaining any data loss.
[+4d:2n]	4	The cluster can recover from four simultaneous drive failures or two node failures without sustaining any data loss.
[+4n]	8	The cluster can recover from four simultaneous drive or node failures without sustaining any data loss.
Nx (Data mirroring)	N For example, 5x requires a minimum of five nodes.	The cluster can recover from N - 1 drive or node failures without sustaining data loss. For example, 5x protection means that the cluster can recover from four drive or node failures.

Requested protection disk space usage

Increasing the requested protection of data also increases the amount of space consumed by the data on the cluster.

The parity overhead for N + M protection depends on the file size and the number of nodes in the cluster. The percentage of parity overhead declines as the cluster gets larger.

The following table describes the estimated amount of overhead depending on the requested protection and the size of the cluster or node pool. The table does not show recommended protection levels based on cluster size.

Number of nodes	[+1n]	[+2d:1n]	[+2n]	[+3d: 1n]	[+3d: 1n1d]	[+3n]	[+4d: 1n]	[+4d: 2n]	[+4n]
3	2 +1 (33%)	4 + 2 (33%)	—	6 + 3 (33%)	3 + 3 (50%)	—	8 + 4 (33%)	—	—

Number of nodes	[+1n]	[+2d:1n]	[+2n]	[+3d: 1n]	[+3d: 1n1d]	[+3n]	[+4d: 1n]	[+4d: 2n]	[+4n]
4	3 +1 (25%)	6 + 2 (25%)	—	9 + 3 (25%)	5 + 3 (38%)	—	12 + 4 (25%)	4 + 4 (50%)	—
5	4 +1 (20%)	8 + 2 (20%)	3 + 2 (40%)	12 + 3 (20%)	7 + 3 (30%)	—	16 + 4 (20%)	6 + 4 (40%)	—
6	5 +1 (17%)	10 + 2 (17%)	4 + 2 (33%)	15 + 3 (17%)	9 + 3 (25%)	—	16 + 4 (20%)	8 + 4 (33%)	—
7	6 +1 (14%)	12 + 2 (14%)	5 + 2 (29%)	15 + 3 (17%)	11 + 3 (21%)	4 + 3 (43%)	16 + 4 (20%)	10 + 4 (29%)	—
8	7 +1 (13%)	14 + 2 (12.5%)	6 + 2 (25%)	15 + 3 (17%)	13 + 3 (19%)	5 + 3 (38%)	16 + 4 (20%)	12 + 4 (25%)	—
9	8 +1 (11%)	16 + 2 (11%)	7 + 2 (22%)	15 + 3 (17%)	15+3 (17%)	6 + 3 (33%)	16 + 4 (20%)	14 + 4 (22%)	5 + 4 (44%)
10	9 +1 (10%)	16 + 2 (11%)	8 + 2 (20%)	15 + 3 (17%)	15+3 (17%)	7 + 3 (30%)	16 + 4 (20%)	16 + 4 (20%)	6 + 4 (40%)
12	11 +1 (8%)	16 + 2 (11%)	10 + 2 (17%)	15 + 3 (17%)	15+3 (17%)	9 + 3 (25%)	16 + 4 (20%)	16 + 4 (20%)	8 + 4 (33%)
14	13 + 1 (7%)	16 + 2 (11%)	12 + 2 (14%)	15 + 3 (17%)	15+3 (17%)	11 + 3 (21%)	16 + 4 (20%)	16 + 4 (20%)	10 + 4 (29%)
16	15 + 1 (6%)	16 + 2 (11%)	14 + 2 (13%)	15 + 3 (17%)	15+3 (17%)	13 + 3 (19%)	16 + 4 (20%)	16 + 4 (20%)	12 + 4 (25%)
18	16 + 1 (6%)	16 + 2 (11%)	16 + 2 (11%)	15 + 3 (17%)	15+3 (17%)	15 + 3 (17%)	16 + 4 (20%)	16 + 4 (20%)	14 + 4 (22%)
20	16 + 1 (6%)	16 + 2 (11%)	16 + 2 (11%)	16 + 3 (16%)	16 + 3 (16%)	16 + 3 (16%)	16 + 4 (20%)	16 + 4 (20%)	16 + 4 (20%)
30	16 + 1 (6%)	16 + 2 (11%)	16 + 2 (11%)	16 + 3 (16%)	16 + 3 (16%)	16 + 3 (16%)	16 + 4 (20%)	16 + 4 (20%)	16 + 4 (20%)

The parity overhead for mirrored data protection is not affected by the number of nodes in the cluster. The following table describes the parity overhead for requested mirrored protection.

2x	3x	4x	5x	6x	7x	8x
50%	67%	75%	80%	83%	86%	88%

CHAPTER 14

NDMP backup

This section contains the following topics:

NDMP backup and recovery overview

In OneFS, you can back up and restore file-system data through the Network Data Management Protocol (NDMP). From a backup server, you can direct backup and recovery processes between an Isilon cluster and backup devices such as tape devices, media servers, and virtual tape libraries (VTLs).

OneFS supports both three-way and two-way NDMP backup models. Three-way NDMP backup is also known as the remote NDMP backup and the two-way NDMP backup is known as the local or direct NDMP backup. During a three-way NDMP backup operation, a data management application (DMA) on a backup server instructs the cluster to start backing up data to a tape media server that is either attached to the LAN or directly attached to the DMA.

During a two-way NDMP backup operation, a DMA on a backup server instructs a Backup Accelerator node on the cluster to start backing up data to a tape media server that is attached to the Backup Accelerator node.

Two-way NDMP backup is significantly faster than the three-way NDMP backup. It is also the most efficient method in terms of cluster resource consumption. However, a two-way NDMP backup requires that you attach one or more Backup Accelerator nodes to the cluster.

In both the two-way and three-way NDMP backup models, file history data is transferred from the cluster to the backup server. Before a backup begins, OneFS creates a snapshot of the targeted directory, then backs up the snapshot, which ensures that the backup image represents a specific point in time.

You do not need to activate a SnapshotIQ license on the cluster to perform NDMP backups. If you have activated a SnapshotIQ license on the cluster, you can generate a snapshot through the SnapshotIQ tool, and then back up the same snapshot to multiple tape devices. If you back up a SnapshotIQ snapshot, OneFS does not create another snapshot for the backup.

Note

If you are backing up SmartLock directories for compliance purposes, it is recommended that you do not specify autocommit time periods for the SmartLock directories. This is because, depending on the autocommit period, files in the SmartLock directories may still be subject to change.

NDMP two-way backup

The NDMP two-way backup is also known as the local or direct NDMP backup. To perform NDMP two-way backups, you must attach a Backup Accelerator node to your Isilon cluster and attach a tape device to the Backup Accelerator node. You must then use OneFS to detect the tape device before you can back up to that device.

You can connect supported tape devices directly to the Fibre Channel ports of a Backup Accelerator node. Alternatively, you can connect Fibre Channel switches to the Fibre Channel ports on the Backup Accelerator node, and connect tape and media changer devices to the Fibre Channel switches. For more information, see your Fibre Channel switch documentation about zoning the switch to allow communication between the Backup Accelerator node and the connected tape and media changer devices.

If you attach tape devices to a Backup Accelerator node, the cluster detects the devices when you start or restart the node or when you re-scan the Fibre Channel ports to discover devices. If a cluster detects tape devices, the cluster creates an entry for the path to each detected device.

If you connect a device through a Fibre Channel switch, multiple paths can exist for a single device. For example, if you connect a tape device to a Fibre Channel switch, and then connect the Fibre Channel switch to two Fibre Channel ports, OneFS creates two entries for the device, one for each path.

Note

If you perform an NDMP two-way backup operation, you must assign static IP addresses to the Backup Accelerator node. If you connect to the cluster through a data management application (DMA), you must connect to the IP address of a Backup Accelerator node. If you perform an NDMP three-way backup, you can connect to any node in the cluster.

Snapshot-based incremental backups

You can implement snapshot-based incremental backups to increase the speed at which these backups are performed.

During a snapshot-based incremental backup, OneFS checks the snapshot taken for the previous NDMP backup operation and compares it to a new snapshot. OneFS then backs up all data that was modified since the last snapshot was made.

If the incremental backup does not involve snapshots, OneFS must scan the directory to discover which files were modified. OneFS can perform incremental backups significantly faster if snapshots are referenced.

You can perform incremental backups without activating a SnapshotIQ license on the cluster. Although SnapshotIQ offers a number of useful features, it does not enhance snapshot capabilities in NDMP backup and recovery.

Note

If you run an NDMP backup on a cluster with a SnapshotIQ license, the snapshot visibility must be turned on for SMB, NFS, and local clients for a successful completion of the operation.

Set the `BACKUP_MODE` environment variable to `SNAPSHOT` to enable snapshot-based incremental backups. If you enable snapshot-based incremental backups, OneFS retains each snapshot taken for NDMP backups until a new backup of the same or lower level is performed. However, if you do not enable snapshot-based incremental backups, OneFS automatically deletes each snapshot generated after the corresponding backup is completed or canceled.

After setting the `BACKUP_MODE` environment variable, snapshot-based incremental backup works with certain data management applications (DMAs) as listed in the next table.

Table 16 DMA support for snapshot-based incremental backups

DMA	Supported
Symantec NetBackup	Yes
EMC Networker	Yes
EMC Avamar	Yes

Table 16 DMA support for snapshot-based incremental backups (continued)

DMA	Supported
Commvault Simpana	No
IBM Tivoli Storage Manager	No
Symantec Backup Exec	Yes
Dell NetVault	No
ASG-Time Navigator	No

NDMP protocol support

You can back up cluster data through version 3 or 4 of the NDMP protocol.

OneFS supports the following features of NDMP versions 3 and 4:

- Full (level 0) NDMP backups
- Incremental (levels 1-10) NDMP backups

  Note

  In a level 10 NDMP backup, only data changed since the most recent incremental (level 1-9) backup or the last level 10 backup is copied. By repeating level 10 backups, you can be assured that the latest versions of files in your data set are backed up without having to run a full backup.

- Token-based NDMP backups
- NDMP TAR backup type
- Path-based and dir/node file history format
- Direct Access Restore (DAR)
- Directory DAR (DDAR)
- Including and excluding specific files and directories from backup
- Backup of file attributes
- Backup of Access Control Lists (ACLs)
- Backup of Alternate Data Streams (ADSs)
- Backup Restartable Extension (BRE)

OneFS supports connecting to clusters through IPv4 or IPv6.

Supported DMAs

NDMP backups are coordinated by a data management application (DMA) that runs on a backup server.

OneFS supports the following DMAs:

- Symantec NetBackup
- EMC NetWorker

- EMC Avamar
- Symantec Backup Exec
- IBM Tivoli Storage Manager
- Dell NetVault
- CommVault Simpana
- ASG-Time Navigator

Note

All supported DMAs can connect to an Isilon cluster through IPv4. CommVault Simpana is currently the only DMA that also supports connecting to an Isilon cluster through IPv6.

See the Isilon Third-Party Software and Hardware Compatibility Guide for the latest information about supported DMAs.

NDMP hardware support

OneFS can back up data to and restore data from tape devices and virtual tape libraries (VTLs).

Supported tape devices
OneFS supports the following types of emulated and physical tape devices for two-way NDMP backups:

- LTO-3
- LTO-4
- LTO-5
- LTO-6

Note

OneFS supports only the LTO-3 tape device for a FalconStor VTL. It supports only the LTO-3 and LTO-4 tape devices for a Data Domain VTL.

For three-way NDMP backups, the data management application (DMA) determines the tape devices that are supported.

Supported tape libraries
For both two-way and three-way NDMP backups, OneFS supports all the tape libraries that are supported by the DMA.

Supported virtual tape libraries
For two-way NDMP backups, OneFS supports FalconStor and Data Domain VTLs.

For three-way NDMP backups, the DMA determines the virtual tape libraries that will be supported.

NDMP backup limitations

OneFS NDMP backups have the following limitations:

- Does not support more than 4 KB path length.
- Does not back up file system configuration data, such as file protection level policies and quotas.

- Cannot back up tape blocks larger than 512 KB.
- Does not support multiplexing across multiple streams.
- Does not support multiple concurrent backups onto the same tape.
- Does not support restoring data from a file system other than OneFS. However, you can migrate data via the NDMP protocol from a NetApp or EMC VNX storage system to OneFS.
- Backup Accelerator nodes cannot interact with more than 1024 device paths, including the paths of tape and media changer devices. For example, if each device has four paths, you can connect 256 devices to a Backup Accelerator node. If each device has two paths, you can connect 512 devices.

NDMP performance recommendations

Consider the following recommendations to optimize OneFS NDMP backups.

General performance recommendations

- Install the latest patches for OneFS and your data management application (DMA).
- Run a maximum of eight NDMP concurrent sessions per A100 Backup Accelerator node and four NDMP concurrent sessions per Annapurna Backup Accelerator node to obtain optimal throughput per session.
- NDMP backups result in very high Recovery Point Objectives (RPOs) and Recovery Time Objectives (RTOs). You can reduce your RPO and RTO by attaching one or more Backup Accelerator nodes to the cluster and then running two-way NDMP backups.
- The throughput for an Isilon cluster during the backup and restore operations is dependent on the dataset and is considerably reduced for small files.
- If you are backing up multiple directories that contain small files, set up a separate schedule for each directory.
- If you are performing three-way NDMP backups, run multiple NDMP sessions on multiple nodes in your Isilon cluster.
- Restore files through Direct Access Restore (DAR), especially if you restore files frequently. However, it is recommended that you do not use DAR to restore a full backup or a large number of files, as DAR is better suited to restoring smaller numbers of files.
- Restore files through Directory DAR (DDAR) if you restore large numbers of files frequently.
- Use the largest tape record size available for your version of OneFS to increase throughput.
- If possible, do not include or exclude files from backup. Including or excluding files can affect backup performance, due to filtering overhead.
- Limit the depth of nested subdirectories in your file system.
- Limit the number of files in a directory. Distribute files across multiple directories instead of including a large number of files in a single directory.
- Use path-based file history format.
- Configure multiple policies when scheduling backup operations, with each policy capturing a portion of the file system. Do not attempt to back up the entire file system through a single policy.

- If you are backing up a large number of small files, it is recommended that you distribute the files across multiple NDMP jobs. Spreading backup operations across multiple NDMP jobs can significantly improve the performance of clusters.

SmartConnect recommendations

- A two-way NDMP backup session with SmartConnect requires backup accelerators for backup and restore operations. However, a three-way NDMP session with SmartConnect does not require backup accelerators for these operations.
- For a two-way NDMP backup session with SmartConnect, connect to the NDMP session through a dedicated SmartConnect zone consisting of a pool of Network Interface Cards (NICs) on the backup accelerator nodes.
- For a two-way NDMP backup session without SmartConnect, initiate the backup session through a static IP address or fully qualified domain name of the backup accelerator node.
- For a three-way NDMP backup operation, the front-end Ethernet network or the interfaces of the nodes are used to serve the backup traffic. Therefore, it is recommended that you configure a DMA to initiate an NDMP session only using the nodes that are not already overburdened serving other workloads or connections.
- For a three-way NDMP backup operation with or without SmartConnect, initiate the backup session using the IP addresses of the nodes that are identified for running the NDMP sessions.
- For a three-way NDMP backup operation, when selecting nodes to include in the SmartConnect zone, make sure to include nodes of the same type in the zone. A SmartConnect zone with mixed node types, for example, S210 and NL400 nodes, might degrade the performance of the backup operation.

Backup Accelerator recommendations

- Assign static IP addresses to Backup Accelerator nodes.
- Attach more Backup Accelerator nodes to larger clusters. The recommended number of Backup Accelerator nodes is listed in the following table.

Table 17 Nodes per Backup Accelerator node

Node type	Recommended number of nodes per Backup Accelerator node
X-Series	3
NL-Series	3
S-Series	3
HD-Series	3

- Attach more Backup Accelerator nodes if you are backing up to more tape devices.

DMA-specific recommendations

- Enable parallelism for the DMA if the DMA supports this option. This allows OneFS to back up data to multiple tape devices at the same time.

Excluding files and directories from NDMP backups

You can exclude files and directories from NDMP backup operations by specifying NDMP environment variables through a data management application (DMA). If you include a file or directory, all other files and directories are automatically excluded from backup operations. If you exclude a file or directory, all files and directories except the excluded one are backed up.

You can include or exclude files and directories by specifying the following character patterns:

Table 18 NDMP file and directory matching wildcards

Character	Description	Example	Includes or excludes the following directories
*	Takes the place of any character or characters	archive*	`/ifs/data/archive1` `/ifs/data/archive42_a/media`

[]	Takes the place of a range of letters or numbers	data_store_[a-f]
		data_store_[0-9]
			`/ifs/data/data_store_a`
			`/ifs/data/data_store_c`
			`/ifs/data/data_store_8`

?	Takes the place of any single character	user_?	`/ifs/data/user_1` `/ifs/data/user_2`
\	Includes a blank space	user\ 1	`/ifs/data/user 1`

Note

" " are required for Symantec NetBackup when multiple patterns are specified. The patterns are not limited to directories.

Unanchored patterns such as `home` or `user1` target a string of text that might belong to many files or directories. Anchored patterns target specific file pathnames, such as `ifs/data/home`. You can include or exclude either type of pattern.

For example, suppose you want to back up the `/ifs/data/home` directory, which contains the following files and directories:

- `/ifs/data/home/user1/file.txt`
- `/ifs/data/home/user2/user1/file.txt`
- `/ifs/data/home/user3/other/file.txt`
- `/ifs/data/home/user4/emptydirectory`

If you include the `/ifs/data/home` directory, all files and directories, including `emptydirectory` would be backed up.

If you specify both the include and exclude patterns, the include pattern is first processed followed by the exclude pattern.

If you specify both the include and exclude patterns, any excluded files or directories under the included directories would not be backed up. If the excluded directories are not found in any of the included directories, the exclude specification would have no effect.

Note

Specifying unanchored patterns can degrade the performance of backups. It is recommended that you avoid unanchored patterns whenever possible.

Configuring basic NDMP backup settings

You can configure NDMP backup settings to control how these backups are performed for the cluster. You can also configure OneFS to interact with a specific data management application (DMA) for NDMP backups.

Configure and enable NDMP backup

OneFS prevents NDMP backups by default. Before you can perform NDMP backups, you must enable NDMP backups and configure NDMP settings.

Procedure

1. Click **Data Protection** › **Backup** › **NDMP Settings**.
2. In the **Service** area, click **Enable**.
3. (Optional) To specify a port through which data management applications (DMAs) access the cluster, or the DMA vendor that OneFS is to interact with, in the **Settings** area, click **Edit settings**.
   - In the **Port number** field, type a port number.
   - From the **DMA vendor** list, select the name of the DMA vendor to manage backup operations.

     If your DMA vendor is not included in the list, select **generic**. However, note that any vendors not included on the list are not officially supported and might not function as expected.
4. Add an NDMP user account through which your DMA can access the cluster.

Disable NDMP backup

You can disable NDMP backup if you no longer want to use this backup method.

Procedure

1. Click **Data Protection** › **Backup** › **NDMP Settings**.
2. In the **Service** area, click **Disable**.

View NDMP backup settings

You can view current NDMP backup settings. These settings define whether NDMP backup is enabled, the port through which your data management application (DMA) connects to the cluster, and the DMA vendor that OneFS is configured to interact with.

Procedure

1. Click **Data Protection** › **Backup** › **NDMP Settings** and view NDMP backup settings.
2. In the **Settings** area, review NDMP backup settings.

NDMP backup settings

You can configure the following settings to control how NDMP backups are performed on the cluster.

Port number

The number of the port through which the data management application (DMA) can connect to the cluster.

DMA vendor
The DMA vendor that the cluster is configured to interact with.

Managing NDMP user accounts

You can create, delete, and modify the passwords of NDMP user accounts.

Create an NDMP user account

Before you can perform NDMP backups, you must create an NDMP user account through which your data management application (DMA) can access the Isilon cluster.

Procedure

1. Click **Data Protection** › **Backup** › **NDMP Settings**.
2. In the **NDMP Administrators** area, click **Add administrator**.
3. In the **Add Administrator** dialog box, in the **Name** field, type a name for the account.
4. In the **Password** and **Confirm password** fields, type a password for the account.
5. Click **Submit**.

Modify the password of an NDMP user account

You can modify the password for an NDMP user account.

Procedure

1. Click **Data Protection** › **Backup** › **NDMP Settings**.
2. In the **NDMP Administrator** table, in the row for an NDMP user account, click **Change password**.
3. In the **Password** and **Confirm password** fields, type a new password for the account.
4. Click **Submit**.

Delete an NDMP user account

You can delete an NDMP user account.

Procedure

1. Click **Data Protection** › **Backup** › **NDMP Settings**.
2. In the **NDMP Administrators** table, in the row for an NDMP user account, click **Delete**.
3. In the **Confirm** dialog box, click **Yes**.

View NDMP user accounts

You can view information about NDMP user accounts.

Procedure

1. Click **Data Protection** › **Backup** › **NDMP Settings**.
2. In the **NDMP administrators** area, review information about NDMP user accounts.

Managing NDMP backup devices

After you attach a tape or media changer device to a Backup Accelerator node, you must configure OneFS to detect and establish a connection to the device. After the connection between the cluster and the backup device is established, you can modify the name that the cluster has assigned to the device, or disconnect the device from the cluster.

Detect NDMP backup devices

If you connect a tape device or media changer to a Backup Accelerator node, you must configure OneFS to detect the device. Only then can OneFS back up data to and restore data from the device. In OneFS, you can scan a specific node, a specific port, or all ports on all nodes.

Procedure

1. Click **Data Protection** › **Backup** › **Devices**.
2. Click **Discover devices**.
3. (Optional) To scan only a specific node for NDMP devices, from the **Nodes** list, select a node.
4. (Optional) To scan only a specific port for NDMP devices, from the **Ports** list, select a port.

   If you specify a port and a node, only the specified port on the node is scanned. However, if you specify only a port, the specified port will be scanned on all nodes.
5. (Optional) To remove entries for devices or paths that have become inaccessible, select the **Delete inaccessible paths or devices** check box.
6. Click **Submit**.

Results

For each device that is detected, an entry is added to either the **Tape Devices** or **Media Changers** tables.

Modify the name of an NDMP backup device

You can modify the name of an NDMP backup device in OneFS.

Procedure

1. Click **Data Protection** › **Backup** › **Devices**.
2. In the **Tape Devices** table, click the name of a backup device entry.
3. In the **Rename Device** dialog box, in the **Device Name** field, type a new name for the backup device.
4. Click **Submit**.

Delete an entry for an NDMP backup device

If you physically remove an NDMP device from a cluster, OneFS retains the entry for the device. You can delete a device entry for a removed device. You can also remove the device entry for a device that is still physically attached to the cluster; this causes OneFS to disconnect from the device.

If you remove a device entry for a device that is connected to the cluster, and you do not physically disconnect the device, OneFS will detect the device the next time it scans the ports. You cannot remove a device entry for a device that is currently in use.

Procedure

1. Click **Data Protection** › **Backup** › **Devices**.

2. In the **Tape Devices** table, in the row for the target device, click **Delete device**.
3. In the **Confirm** dialog box, click **Yes**.

View NDMP backup devices

You can view information about tape and media changer devices that are currently attached to your Isilon cluster.

Procedure

1. Click **Data Protection** › **Backup** › **Devices**.
2. In the **Tape Devices** and **Media Changers** tables, review information about NDMP backup devices.

NDMP backup device settings

OneFS creates a device entry for each device you attach to the cluster through a Backup Accelerator node.

The following table describes the settings available in the **Tape Devices** and **Media Changers** tables:

Table 19 NDMP backup device settings

Setting	Description
Name	A device name assigned by OneFS.
State	Indicates whether the device is in use. If data is currently being backed up to or restored from the device, `Read/Write` appears. If the device is not in use, `Closed` appears.
WWN	The world wide node name (WWNN) of the device.
Product	The name of the device vendor and the model name or number of the device.
Serial Number	The serial number of the device.
Paths	The name of the Backup Accelerator node that the device is attached to and the port number or numbers to which the device is connected.
LUN	The logical unit number (LUN) of the device.
Port ID	The port ID of the device that binds the logical device to the physical device.
WWPN	The world wide port name (WWPN) of the port on the tape or media changer device.

Managing NDMP backup ports

You can manage the Fibre Channel ports that connect tape and media changer devices to a Backup Accelerator node. You can also enable, disable, or modify the settings of an NDMP backup port.

Modify NDMP backup port settings

You can modify the settings of an NDMP backup port.

Procedure

1. Click **Data Protection** › **Backup** › **Ports**.
2. In the **Sessions** table, click the name of a port.
3. In the **Edit Port** dialog box, modify port settings as needed, and then click **Submit**.

Enable or disable an NDMP backup port

You can enable or disable an NDMP backup port.

Procedure

1. Click **Data Protection** › **Backup** › **Ports**.
2. In the **Ports** table, in the row of a port, click **Enable** or **Disable**.

View NDMP backup ports

You can view information about Fibre Channel ports of Backup Accelerator nodes attached to a cluster.

Procedure

1. Click **Data Protection** › **Backup** › **Ports**.
2. In the **Ports** table, review information about NDMP backup ports.

NDMP backup port settings

OneFS assigns default settings to each port on each Backup Accelerator node attached to the cluster. These settings identify each port and specify how the port interacts with NDMP backup devices.

The settings that appear in the **Ports** table are as follows:

Table 20 NDMP backup port settings

Setting	Description
Port	The name of the Backup Accelerator node, and the number of the port.
Topology	The type of Fibre Channel topology that the port is configured to support.. Options are: **Point to Point** A single backup device or Fibre Channel switch directly connected to the port. **Loop** Multiple backup devices connected to a single port in a circular formation. **Auto** Automatically detects the topology of the connected device. This is the recommended setting, and is required if you are using a switched-fabric topology.

Table 20 NDMP backup port settings (continued)

Setting	Description
WWNN	The world wide node name (WWNN) of the port. This name is the same for each port on a given node.
WWPN	The world wide port name (WWPN) of the port. This name is unique to the port.
Rate	The rate at which data is sent through the port. The rate can be set to `1 Gb/s`, `2 Gb/s`, `4 Gb/s`, `8 Gb/s`, and `Auto`. `8 Gb/s` is available for A100 nodes only. If set to `Auto`, OneFS automatically negotiates with the DMA to determine the rate. `Auto` is the recommended setting.

Managing NDMP backup sessions

You can view the status of NDMP backup sessions or terminate a session that is in progress.

End an NDMP session

You can end an NDMP backup or restore session at any time.

Procedure

1. Click **Data Protection** › **Backup** › **Sessions**.
2. In the **Sessions** table, in the row of the NDMP session that you want to end, click **Kill**.
3. In the **Confirm** dialog box, click **Yes**.

View NDMP sessions

You can view information about active NDMP sessions.

Procedure

1. Click **Data Protection** › **Backup** › **Sessions**.
2. In the **Sessions** table, review information about NDMP sessions.

NDMP session information

You can view information about active NDMP sessions.

The following information is included in the **Sessions** table, as follows:

Table 21 NDMP session information

Item	Description
Session	The unique identification number that OneFS assigned to the session.
Elapsed	How much time has elapsed since the session started.
Transferred	The amount of data that was transferred during the session.

Table 21 NDMP session information (continued)

Item	Description
Throughput	The average throughput of the session over the past five minutes.
Client/Remote	The IP address of the backup server that the data management application (DMA) is running on. If a three-way NDMP backup or restore operation is currently running, the IP address of the remote tape media server also appears.
Mover/Data	The current state of the data mover and the data server. The first word describes the activity of the data mover. The second word describes the activity of the data server. The data mover and data server send data to and receive data from each other during backup and restore operations. The data mover is a component of the backup server that receives data during backups and sends data during restore operations. The data server is a component of OneFS that sends data during backups and receives information during restore operations. The following states might appear: Active The data mover or data server is currently sending or receiving data. Paused The data mover is temporarily unable to receive data. While the data mover is paused, the data server cannot send data to the data mover. The data server cannot be paused. Idle The data mover or data server is not sending or receiving data. Listen The data mover or data server is waiting to connect to the data server or data mover.
Operation	The type of operation (backup or restore) that is currently in progress. If no operation is in progress, this field is blank. Backup (0-10) Indicates that data is currently being backed up to a media server. The number indicates the level of NDMP backup. Restore Indicates that data is currently being restored from a media server.
Source/Destination	If an operation is currently in progress, specifies the `/ifs` directories that are affected by the operation. If a backup is in progress, displays the path of the source directory that is being backed up. If a restore operation is in progress, displays the path of the directory that is being restored along with the destination directory to which the tape media server is restoring data. If you are restoring data to the same location that you backed up your data from, the same path appears twice.

Table 21 NDMP session information (continued)

Item	Description
Device	The name of the tape or media changer device that is communicating with the cluster.
Mode	How OneFS is interacting with data on the backup media server, as follows: Read/Write OneFS is reading and writing data during a backup operation. Read OneFS is reading data during a restore operation. Raw The DMA has access to tape drives, but the drives do not contain writable tape media.

Managing restartable backups

A restartable backup is a type of NDMP backup that you can enable in your data management application (DMA). If a restartable backup fails, for example, because of a power outage, you can restart the backup from a checkpoint close to the point of failure. In contrast, when a non-restartable backup fails, you must back up all data from the beginning, regardless of what was transferred during the initial backup process.

After you enable restartable backups from your DMA, you can manage restartable backup contexts from OneFS. These contexts are the data that OneFS stores to facilitate restartable backups. Each context represents a checkpoint that the restartable backup process can return to if a backup fails.

Restartable backups are supported only for EMC NetWorker 8.1 and later.

Configure restartable backups for EMC NetWorker

You must configure EMC NetWorker to enable restartable backups and, optionally, define the checkpoint interval.

If you do not specify a checkpoint interval, NetWorker uses the default interval of 5 GB.

Procedure

1. Configure the client and the directory path that you want to back up as you would normally.
2. In the **Client Properties** dialog box, enable restartable backups.
   a. On the **General** page, click the **Checkpoint enabled** checkbox.
   b. In the **Checkpoint granularity** drop-down list, select `File`.
3. In the **Application information** field, type any NDMP variables that you want to specify.

   The following variable setting specifies a checkpoint interval of 1 GB: **`CHECKPOINT_INTERVAL_IN_BYTES=1GB`**
4. Finish configuration and click **OK** in the **Client Properties** dialog box.
5. Start the backup.

6. If the backup is interrupted—for example, because of a power failure—restart it.
   a. On the **Monitoring** page, locate the backup process in the **Groups** list.
   b. Right-click the backup process and then, in the context menu, click **Restart**.

   NetWorker automatically restarts the backup from the last checkpoint.

Delete a restartable backup context

After a restartable backup context is no longer needed, your data management application (DMA) automatically requests that OneFS delete the context. You can manually delete a restartable backup context before the DMA requests it.

Note

It is recommended that you do not manually delete restartable backup contexts. Manually deleting a restartable backup context requires you to restart the corresponding NDMP backup from the beginning.

Procedure

1. Run the `isi ndmp extensions contexts delete` command.

   The following command deletes a restartable backup context with an ID of 792eeb8a-8784-11e2-aa70-0025904e91a4:

   ```
   isi ndmp extensions contexts delete 792eeb8a-8784-11e2-
   aa70-0025904e91a4
   ```

View restartable backup contexts

You can view restartable backup contexts that have been configured.

Procedure

1. View all backup contexts by running the following command:

   ```
   isi ndmp extensions contexts list
   ```

2. To view detailed information about a specific backup context, run the `isi ndmp extensions contexts view` command.

   The following command displays detailed information about a backup context with an ID of 792eeb8a-8784-11e2-aa70-0025904e91a4:

   ```
   isi ndmp extensions contexts view 792eeb8a-8784-11e2-
   aa70-0025904e91a4
   ```

Configure restartable backup settings

You can specify the number of restartable backup contexts that OneFS retains at a time, up to a maximum of 1024 contexts.

Procedure

1. Run the `isi ndmp extensions settings modify` command.

   The following command sets the maximum number of restartable backup contexts to 128:

   ```
   isi ndmp extensions settings modify --bre_max_contexts 128
   ```

View restartable backup settings

You can view the current limit of restartable backup contexts that OneFS retains at one time.

Procedure

1. Run the following command:

   ```
   isi ndmp extensions settings view
   ```

Managing file list backups

If your data management application (DMA) can pass environment variables to OneFS, you can control backups by specifying a file list.

Currently, EMC Networker and Symantec NetBackup can pass environment variables to OneFS.

With a normal NDMP level 0 (full) backup, your DMA backs up an entire source directory. With an NDMP incremental (level 1-10) backup, your DMA backs up only those files that have been created or changed since the previous incremental backup of the same level.

When you specify a file list backup, only the listed files and subdirectories in the source directory are backed up. With a level 0 file list backup, all listed files and directories in the source directory are backed up.

A backup level other than 0 triggers an incremental file list backup. In an incremental file list backup, only the listed files that were created or changed in the source directory since the last incremental backup of the same level are backed up.

To configure a file list backup, you must complete the following tasks:

- Create the file list and place it in OneFS
- Specify the path of the source directory
- Specify the file list location

The file list is an ASCII text file that lists the pathnames of files to be backed up. The pathnames must be relative to the path specified in the `FILESYSTEM` environment variable. Absolute file paths in the file list are not supported. The pathnames of all files must be included, or they are not backed up. For example, if you include the pathname of a subdirectory, only the subdirectory, not the files it contains, is backed up.

To specify the full path of the source directory to be backed up, you must specify the `FILESYSTEM` environment variable in your DMA. For example:

```
FILESYSTEM=/ifs/data/projects
```

To specify the pathname of the file list, you must specify the environment variable, `BACKUP_FILE_LIST` in your DMA. The file list must be accessible from the node performing the backup. For example:

```
BACKUP_FILE_LIST=/ifs/data/proj_list.txt
```

Format of a backup file list

You must create a file list to enable a file list backup.

A file list backup requires an ASCII text file in a particular format to identify the pathnames of files to be backed up. Following is an example of a file list with pathnames relative to `/ifs/data/projects`:

```
project-summary_rev0.xls
project-summary_rev3.xls
proj001/plan/plan001.doc
proj001/plan/plan001.pdf
proj001/plan/proj-logo.png
proj001/plan/schedule001.xls
proj001/plan/stakeholders_list.txt
proj001/plan/team_pic.png
proj002/plan/logo.png
proj002/plan/projplan002.doc
proj002/plan/projplan002.pdf
proj002/plan/sched-002.xls
proj002/plan/stakeholders.txt
proj002/plan/team.png
proj005/plan/projectlogo.png
proj005/plan/projectteam.png
proj005/plan/proj-plan005.doc
proj005/plan/proj-plan005.pdf
proj005/plan/schedule.xls
proj005/plan/stakeholders.txt
```

As shown in the example, the pathnames are relative to the full path of the source directory, which you specify in the FILESYSTEM environment variable. Absolute file paths are not supported in the file list.

Also as shown, the directories and files must be in sorted order for the backup to be successful.

The pathnames of all files must be included in the file list, or they are not backed up. For example, if you only include the pathname of a subdirectory, the subdirectory is backed up, but not the files the subdirectory contains. The exception is ADS (alternate data streams). All ADS associated with a file to be backed up are automatically backed up.

Placement of the file list

Before you can perform a file list backup, you must place the file list in OneFS.

For example, suppose the FILESYSTEM environment variable specifies the full path of the directory to be backed up as `/ifs/data/projects`. You can place the text file containing the file list anywhere within the `/ifs` path.

Start a file list backup

You can configure and start a file list backup from your data management application (DMA).

Before you begin

You should have already specified and saved the list of files to be backed up in an ASCII text file.

Configure a file list backup from your DMA as you would any backup, but with a few additional steps as described in the following procedure.

Procedure

1. Copy the file list to the OneFS file system on the cluster containing the files to be backed up.

   For example, if the directory that you specify in the `FILESYSTEM` environment variable is `/ifs/data/projects`, you could place your file list at `/ifs/data`.

2. In your DMA, specify the `BACKUP_FILE_LIST` environment variable to be the full pathname of the file list.

   For example, if the file list was named `proj_backup.txt`, and you placed it at `/ifs/data`, specify `/ifs/data/proj_backup.txt` as the full pathname of the file list.

3. Start your backup as you normally would.

Results

The files in your file list are backed up as specified.

NDMP restore operations

NDMP supports the following types of restore operations:

- Parallel restore (multi-threaded process)
- Serial restore (single-threaded process)

Parallel restore operation

In OneFS, the default NDMP restore for any path is the parallel restore operation. Parallel (multi-threaded) restore enables faster full or partial restore operations by writing data to the cluster as fast as the data can be read from the tape. However, if you specify DAR (direct access restore), the operation reverts to serial processing.

Specify a serial restore operation

You can use the `RESTORE_OPTIONS` environment variable to specify a serial (single-threaded) restore operation.

Procedure

1. In your data management application, configure a restore operation as you normally would.

2. Make sure that the `RESTORE_OPTIONS` environment variable is set to **`1`** on your data management application.

   If the `RESTORE_OPTIONS` environment variable is not already set to **`1`**, specify the `isi ndmp settings variables modify` command from the OneFS command line. The following command specifies serial restore for the `/ifs/data/projects` directory:

   ```
   isi ndmp settings variables modify --path /ifs/data/projects --
   name restore_options --value 1
   ```

   The value of the `path` option is the `FILESYSTEM` environment variable set during the backup operation. The value that you specify for the `name` option is case sensitive.

3. Start the restore operation.

Specify a serial restore operation

You can use the RESTORE_OPTIONS environment variable to specify a serial (single-threaded) restore operation.

Procedure

1. In your data management application, configure a restore operation as you normally would.
2. Make sure that the RESTORE_OPTIONS environment variable is set to **1** on your data management application.

   If the RESTORE_OPTIONS environment variable is not already set to **1**, specify the isi ndmp settings variables modify command from the OneFS command line. The following command specifies serial restore for the /ifs/data/projects directory:

   ```
   isi ndmp settings variables modify --path /ifs/data/projects --
   name restore_options --value 1
   ```

   The value of the path option is the FILESYSTEM environment variable set during the backup operation. The value that you specify for the name option is case sensitive.
3. Start the restore operation.

Sharing tape drives between clusters

Multiple Isilon clusters, or an Isilon cluster and a third-party NAS system, can be configured to share a single tape drive. This helps to maximize the use of the tape infrastructure in your data center.

In your data management application (DMA), you must configure NDMP to control the tape drive and ensure that it is shared properly. The following configurations are supported.

OneFS Versions	Supported DMAs	Tested configurations
• 7.1.1 • 7.1.0.1 (and later)* • 7.0.2.5 • 6.6.5.26	• EMC NetWorker 8.0 and later • Symantec NetBackup 7.5 and later	• Isilon Backup Accelerator with a second Backup Accelerator • Isilon Backup Accelerator with a NetApp storage system
* The tape drive sharing function is not supported in the OneFS 7.0.1 release.		

EMC NetWorker refers to the tape drive sharing capability as DDS (dynamic drive sharing). Symantec NetBackup uses the term SSO (shared storage option). Consult your DMA vendor documentation for configuration instructions.

Managing default NDMP settings

In OneFS, you can manage NDMP backup and restore operations by specifying default NDMP environment variables. You can also override default NDMP environment variables through your data management application (DMA). For more information about specifying NDMP environment variables through your DMA, see your DMA documentation.

Set default NDMP settings for a directory

You can set default NDMP settings for a directory.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Set default NDMP settings by running the `isi ndmp settings variables create` command.

   For example, the following command enables snapshot-based incremental backups for `/ifs/data/media`:

   ```
   isi ndmp settings variables create /ifs/data/media BACKUP_MODE
   SNAPSHOT
   ```

Modify default NDMP settings for a directory

You can modify the default NDMP settings for a directory.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Modify default NDMP settings by running the `isi ndmp settings variables modify` command.

   For example, the following command sets the default file history format to path-based format for `/ifs/data/media`:

   ```
   isi ndmp settings variables modify /ifs/data/media HIST F
   ```

3. (Optional) To remove a default NDMP setting for a directory, run the `isi ndmp settings variables delete` command:

   For example, the following command removes the default file history format for `/ifs/data/media`:

   ```
   isi ndmp settings variables delete /ifs/data/media --name HIST
   ```

View default NDMP settings for directories

You can view the default NDMP settings for directories.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. View default NDMP settings by running the following command:

   ```
   isi ndmp settings variables list
   ```

NDMP environment variables

You can specify default settings of NDMP backup and restore operations through NDMP environment variables. You can also specify NDMP environment variables through your data management application (DMA).

Table 22 NDMP environment variables

Environment variable	Valid values	Default	Description
`BACKUP_MODE=`	`TIMESTAMP` `SNAPSHOT`	`TIMESTAMP`	Enables or disables snapshot-based incremental backups. To enable snapshot-based incremental backups, specify `SNAPSHOT`. To disable snapshot-based incremental backups, specify `TIMESTAMP`.
`FILESYSTEM=`	*<file-path>*	None	Specifies the full path of the directory you want to back up. Must be specified by the DMA before starting the backup, or an error is generated.
`LEVEL=`	*<integer>*	`0`	Specifies the level of NDMP backup to perform. The following values are valid: **0** Performs a full NDMP backup. **1 - 9** Performs an incremental backup at the specified level. **10** Performs unlimited incremental backups.
`UPDATE=`	`Y` `N`	`Y`	Determines whether OneFS updates the dump dates file. **Y** OneFS updates the dump dates file.

Table 22 NDMP environment variables (continued)

Environment variable	Valid values	Default	Description
			N OneFS does not update the dump dates file.
`HIST=`	*‹file-history-format›*	`Y`	Specifies the file history format. The following values are valid: **D** Specifies dir/node file history. **F** Specifies path-based file history. **Y** Specifies the default file history format determined by your NDMP backup settings. **N** Disables file history.
`DIRECT=`	`Y` `N`	`N`	Enables or disables Direct Access Restore (DAR) and Directory DAR (DDAR). The following values are valid: **Y** Enables DAR and DDAR. **N** Disables DAR and DDAR.
`FILES=`	*‹file-matching-pattern›*	None	If you specify this option, OneFS backs up only files and directories that meet the specified pattern. Separate multiple patterns with a space.
`EXCLUDE=`	*‹file-matching-pattern›*	None	If you specify this option, OneFS does not back up files and directories that meet the specified pattern.

Table 22 NDMP environment variables (continued)

Environment variable	Valid values	Default	Description
			Separate multiple patterns with a space.
`RESTORE_HARDLINK _BY_TABLE=`	`Y` `N`	`N`	Determines whether OneFS recovers hard links by building a hard-link table during restore operations. Specify this option if hard links were incorrectly backed up, and restore operations are failing. If a restore operation fails because hard links were incorrectly backed up, the following message appears in the NDMP backup logs: `Bad hardlink path for <path>`
`BACKUP_FILE_LIST=`	*<file-path>*	None	Specifies the pathname in OneFS of the file list to control the backup. This variable must be passed from the DMA initiating the backup. Currently, only EMC Networker and Symantec NetBackup can pass environment variables to OneFS.

`RESTORE_OPTIONS=`	`0`
	`1`
		`0`	The restore operation, by default, is multi-threaded to improve performance. To change the restore operation to single-threaded, specify `RESTORE_OPTIONS=1`

Managing snapshot based incremental backups

After you enable snapshot-based incremental backups, you can view and delete the snapshots created for these backups.

Enable snapshot-based incremental backups for a directory

You can configure OneFS to perform snapshot-based incremental backups for a directory by default. You can also override the default setting in your data management application (DMA).

Procedure

1. Run the `isi ndmp settings variable create` command.

   The following command enables snapshot-based incremental backups for `/ifs/data/media`:

   ```
   isi ndmp settings variables create /ifs/data/media BACKUP_MODE
   SNAPSHOT
   ```

View snapshots for snapshot-based incremental backups

You can view snapshots generated for snapshot-based incremental backups.

Procedure

1. Run the following command:

   ```
   isi ndmp dumpdates list
   ```

Delete snapshots for snapshot-based incremental backups

You can delete snapshots created for snapshot-based incremental backups.

Note

It is recommended that you do not delete snapshots created for snapshot-based incremental backups. If all snapshots are deleted for a path, the next backup performed for the path is a full backup.

Procedure

1. Run the `isi ndmp dumpdates delete` command.

   The following command deletes all snapshots created for backing up `/ifs/data/media`:

   ```
   isi ndmp dumpdates delete /ifs/data/media
   ```

View NDMP backup logs

You can view information about NDMP backup and restore operations through NDMP backup logs.

Procedure

1. Click **Data Protection** › **Backup** › **Logs**.
2. In the **Log Location** area, from the **Node** list, select a node.
3. In the **Log Contents** area, review information about NDMP backup and restore operations.

CHAPTER 15

File retention with SmartLock

This section contains the following topics:

SmartLock overview

You can prevent users from modifying and deleting files on an EMC Isilon cluster with the SmartLock software module. You must activate a SmartLock license on a cluster to protect data with SmartLock.

With the SmartLock software module, you can create SmartLock directories and commit files within those directories to a write once read many (WORM) state. You cannot erase or re-write a file committed to a WORM state. After a file is removed from a WORM state, you can delete the file. However, you can never modify a file that has been committed to a WORM state, even after it is removed from a WORM state.

Compliance mode

SmartLock compliance mode enables you to protect your data in compliance with the regulations defined by U.S. Securities and Exchange Commission rule 17a-4. You can upgrade a cluster to compliance mode during the initial cluster configuration process, before you activate the SmartLock license. To upgrade a cluster to SmartLock compliance mode after the initial cluster configuration process, contact Isilon Technical Support.

If you upgrade a cluster to compliance mode, you will not be able to log in to that cluster through the root user account. Instead, you can log in to the cluster through the compliance administrator account that is configured either during initial cluster configuration or when the cluster is upgraded to compliance mode. If you are logged in through the compliance administrator account, you can perform administrative tasks through the `sudo` command.

SmartLock directories

In a SmartLock directory, you can commit a file to a WORM state manually or you can configure SmartLock to automatically commit the file. You can create two types of SmartLock directories: enterprise and compliance. However, you can create compliance directories only if the cluster has been upgraded to SmartLock compliance mode. Before you can create SmartLock directories, you must activate a SmartLock license on the cluster.

Enterprise directories enable you to protect your data without restricting your cluster to comply with regulations defined by U.S. Securities and Exchange Commission rule 17a-4. If you commit a file to a WORM state in an enterprise directory, the file can never be modified and cannot be deleted until the retention period passes. However, if you are logged in through the root user account, you can delete the file before the retention period passes through the privileged delete feature. The privileged delete feature is not available for compliance directories. Enterprise directories reference the system clock to facilitate time-dependent operations, including file retention.

Compliance directories enable you to protect your data in compliance with the regulations defined by U.S. Securities and Exchange Commission rule 17a-4. If you commit a file to a WORM state in a compliance directory, the file cannot be modified or deleted before the specified retention period has expired. You cannot delete committed files, even if you are logged in to the compliance administrator account. Compliance directories reference the compliance clock to facilitate time-dependent operations, including file retention.

You must set the compliance clock before you can create compliance directories. You can set the compliance clock only once. After you set the compliance clock, you cannot modify the compliance clock time. The compliance clock is controlled by the compliance clock daemon. Because root and compliance administrator users can disable the compliance clock daemon, it is possible for those users to increase the retention period of WORM committed files in compliance mode. However, it is not possible to decrease the retention period of a WORM committed file.

Replication and backup with SmartLock

You must ensure that SmartLock directories remain protected during replication and backup operations.

If you are replicating SmartLock directories with SyncIQ, it is recommended that you configure all nodes on the source and target clusters into Network Time Protocol (NTP) peer mode to ensure that the node clocks are synchronized. For compliance clusters, it is recommended that you configure all nodes on the source and target clusters into NTP peer mode before you set the compliance clock to ensure that the compliance clocks are initially set to the same time.

Note

If you replicate data to a SmartLock directory, do not configure SmartLock settings for that directory until you are no longer replicating data to the directory. Configuring an autocommit time period for a SmartLock directory that you are replicating to can cause replication jobs to fail. If the target directory commits a file to a WORM state, and the file is modified on the source cluster, the next replication job will fail because it cannot update the file.

SmartLock replication and backup limitations

Be aware of the limitations of replicating and backing up SmartLock directories with SyncIQ and NDMP.

If the source or target directory of a SyncIQ policy is a SmartLock directory, replication might not be allowed. For more information, see the following table:

Source directory type	Target directory type	Allowed
Non-SmartLock	Non-SmartLock	Yes
Non-SmartLock	SmartLock enterprise	Yes
Non-SmartLock	SmartLock compliance	No
SmartLock enterprise	Non-SmartLock	Yes; however, retention dates and commit status of files will be lost.
SmartLock enterprise	SmartLock enterprise	Yes
SmartLock enterprise	SmartLock compliance	No
SmartLock compliance	Non-SmartLock	No
SmartLock compliance	SmartLock enterprise	No
SmartLock compliance	SmartLock compliance	Yes

If you are replicating a SmartLock directory to another SmartLock directory, you must create the target SmartLock directory prior to running the replication policy. Although OneFS will create a target directory automatically if a target directory does not already exist, OneFS will not create a target SmartLock directory automatically. If you attempt to replicate an enterprise directory before the target directory has been created, OneFS will create a non-SmartLock target directory and the replication job will succeed. If you replicate a compliance directory before the target directory has been created, the replication job will fail.

If you replicate SmartLock directories to another cluster with SyncIQ, the WORM state of files is replicated. However, SmartLock directory configuration settings are not transferred to the target directory.

For example, if you replicate a directory that contains a committed file that is set to expire on March 4th, the file is still set to expire on March 4th on the target cluster. However, if the directory on the source cluster is set to prevent files from being committed for more than a year, the target directory is not automatically set to the same restriction.

If you back up data to an NDMP device, all SmartLock metadata relating to the retention date and commit status is transferred to the NDMP device. If you restore data to a SmartLock directory on the cluster, the metadata persists on the cluster. However, if the directory that you restore to is not a SmartLock directory, the metadata is lost. You can restore to a SmartLock directory only if the directory is empty.

SmartLock license functionality

You must activate a SmartLock license on a cluster before you can create SmartLock directories and commit files to a WORM state.

If a SmartLock license becomes inactive, you will not be able to create new SmartLock directories on the cluster, modify SmartLock directory configuration settings, or delete files committed to a WORM state in enterprise directories before their expiration dates. However, you can still commit files within existing SmartLock directories to a WORM state.

If a SmartLock license becomes inactive on a cluster that is running in SmartLock compliance mode, root access to the cluster is not restored.

SmartLock considerations

- If a file is owned exclusively by the root user, and the file exists on a cluster that is in SmartLock compliance mode, the file will be inaccessible, because the root user account is disabled in compliance mode. For example, this can happen if a file is assigned root ownership on a cluster that has not been upgraded to compliance mode, and then the file is replicated to a cluster in compliance mode. This can also occur if a file is assigned root ownership before a cluster is upgraded to SmartLock compliance mode or if a root-owned file is restored on a compliance cluster after being backed up.
- It is recommended that you create files outside of SmartLock directories and then transfer them into a SmartLock directory after you are finished working with the files. If you are uploading files to a cluster, it is recommended that you upload the files to a non-SmartLock directory, and then later transfer the files to a SmartLock directory. If a file is committed to a WORM state while the file is being uploaded, the file will become trapped in an inconsistent state.
  Files can be committed to a WORM state while they are still open. If you specify an autocommit time period for a directory, the autocommit time period is calculated according to the length of time since the file was last modified, not when the file was closed. If you delay writing to an open file for more than the autocommit time period, the file will be automatically committed to a WORM state, and you will not be able to write to the file.

- In a Microsoft Windows environment, if you commit a file to a WORM state, you can no longer modify the hidden or archive attributes of the file. Any attempt to modify the hidden or archive attributes of a WORM committed file will generate an error. This can prevent third-party applications from modifying the hidden or archive attributes.

Set the compliance clock

Before you can create SmartLock compliance directories, you must set the compliance clock. This procedure is available only through the command-line interface (CLI).

Setting the compliance clock configures the clock to the same time as the cluster system clock. Before you set the compliance clock, ensure that the system clock is set to the correct time. If the compliance clock later becomes unsynchronized with the system clock, the compliance clock will slowly correct itself to match the system clock. The compliance clock corrects itself at a rate of approximately one week per year.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in through the compliance administrator account.
2. Set the compliance clock by running the following command.

```
isi worm cdate set
```

View the compliance clock

You can view the current time of the compliance clock. This procedure is available only through the command-line interface (CLI).

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in through the compliance administrator account.
2. View the compliance clock by running the following command:

```
isi worm cdate view
```

Creating a SmartLock directory

You can create a SmartLock directory and configure settings that control how long files are retained in a WORM state and when files are automatically committed to a WORM state. You cannot move or rename a directory that contains a SmartLock directory.

It is recommended that you set SmartLock configuration settings only once and do not modify the settings after files have been added to the SmartLock directory.

Retention periods

A retention period is the length of time that a file remains in a WORM state before being released from a WORM state. You can configure SmartLock directory settings that enforce default, maximum, and minimum retention periods for the directory.

If you manually commit a file, you can optionally specify the date that the file is released from a WORM state. You can configure a minimum and a maximum retention period for a SmartLock directory to prevent files from being retained for too long or too short a time period. It is recommended that you specify a minimum retention period for all SmartLock directories.

For example, assume that you have a SmartLock directory with a minimum retention period of two days. At 1:00 PM on Monday, you commit a file to a WORM state, and specify the file to be released from a WORM state on Tuesday at 3:00 PM. The file will be released from a WORM state two days later on Wednesday at 1:00 PM, because releasing the file earlier would violate the minimum retention period.

You can also configure a default retention period that is assigned when you commit a file without specifying a date to release the file from a WORM state.

Autocommit time periods

You can configure an autocommit time period for SmartLock directories. An autocommit time period causes files that have been in a SmartLock directory for a period of time without being modified to be automatically committed to a WORM state.

If you modify the autocommit time period of a SmartLock directory that contains uncommitted files, the new autocommit time period is immediately applied to the files that existed before the modification. For example, consider a SmartLock directory with an autocommit time period of 2 hours. If you modify a file in the SmartLock directory at 1:00 PM, and you decrease the autocommit time period to 1 hour at 2:15 PM, the file is instantly committed to a WORM state.

If a file is manually committed to a WORM state, the read-write permissions of the file are modified. However, if a file is automatically committed to a WORM state, the read-write permissions of the file are not modified.

Create a SmartLock directory

You can create a SmartLock directory and commit files in that directory to a WORM state. This procedure is available only through the command-line interface (CLI).

Before creating a SmartLock directory, be aware of the following conditions and requirements:

- You cannot create a SmartLock directory as a subdirectory of an existing SmartLock directory.
- Hard links cannot cross SmartLock directory boundaries.
- Creating a SmartLock directory causes a corresponding SmartLock domain to be created for that directory.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Run the `isi worm domains create` command.

   If you specify the path of an existing directory, the directory must be empty.

   The following command creates a compliance directory with a default retention period of four years, a minimum retention period of three years, and an maximum retention period of five years:

   ```
   sudo isi worm domains create /ifs/data/SmartLock/directory1 \
   --compliance --default-retention 4Y --min-retention 3Y \
   --max-retention 5Y --mkdir
   ```

   The following command creates an enterprise directory with an autocommit time period of thirty minutes and a minimum retention period of three months:

   ```
   isi worm domains create /ifs/data/SmartLock/directory2 \
   --autocommit-offset 30m --min-retention 3M --mkdir
   ```

Managing SmartLock directories

You can modify SmartLock directory settings only 32 times per directory. SmartLock directory settings include the default, minimum, and maximum retention period and the autocommit time period.

A SmartLock directory can be renamed only if the directory is empty.

Modify a SmartLock directory

You can modify the SmartLock configuration settings for a SmartLock directory. This procedure is available only through the command-line interface (CLI).

Note

You can modify SmartLock directory settings only 32 times per directory. It is recommended that you set SmartLock configuration settings only once and do not modify the settings after files are added to the SmartLock directory.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Modify SmartLock configuration settings by running the `isi worm modify` command.

   The following command sets the default retention period to one year:

   ```
   isi worm domains modify /ifs/data/SmartLock/directory1 \
   --default-retention 1Y
   ```

View SmartLock directory settings

You can view the SmartLock directory settings for SmartLock directories. This procedure is available only through the command-line interface (CLI).

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. View all SmartLock domains by running the following command:

   ```
   isi worm domains list
   ```

   The system displays output similar to the following example:

   ```
   ID    Path                           Type
   -----------------------------------------------
   65536 /ifs/data/SmartLock/directory1 enterprise
   65537 /ifs/data/SmartLock/directory2 enterprise
   65538 /ifs/data/SmartLock/directory3 enterprise
   -----------------------------------------------
   ```

3. (Optional) To view detailed information about a specific SmartLock directory, run the `isi worm domains view` command.

   The following command displays detailed information about `/ifs/data/SmartLock/directory2`:

   ```
   isi worm domains view /ifs/data/SmartLock/directory2
   ```

The system displays output similar to the following example:

```
                ID: 65537
              Path: /ifs/data/SmartLock/directory2
              Type: enterprise
               LIN: 4295426060
Autocommit Offset: 30m
    Override Date: -
Privileged Delete: off
Default Retention: 1Y
    Min Retention: 3M
    Max Retention: -
   Total Modifies: 3/32 Max
```

SmartLock directory configuration settings

You can configure SmartLock directory settings that determine when files are committed to and how long files are retained in a WORM state.

ID

The numerical ID of the corresponding SmartLock domain.

Path

The path of the directory.

Type

The type of SmartLock directory.

LIN

The inode number of the directory.

Autocommit offset

The autocommit time period for the directory. After a file exists in this SmartLock directory without being modified for the specified time period, the file is automatically committed to a WORM state.

Times are expressed in the format "*<integer> <time>*", where *<time>* is one of the following values:

Y

Specifies years

M

Specifies months

W

Specifies weeks

D

Specifies days

H

Specifies hours

m

Specifies minutes

s

Specifies seconds

Override date

The override retention date for the directory. Files committed to a WORM state are not released from a WORM state until after the specified date, regardless of the maximum retention period for the directory or whether a user specifies an earlier date to release a file from a WORM state.

Privileged delete

Indicates whether files in the directory can be deleted through the privileged delete functionality.

on

The root user can delete files committed to a WORM state by running the `isi worm files delete` command.

off

WORM committed files cannot be deleted, even through the `isi worm files delete` command.

disabled

WORM committed files cannot be deleted, even through the `isi worm files delete` command. After this setting is applied, it cannot be modified.

Default retention period

The default retention period for the directory. If a user does not specify a date to release a file from a WORM state, the default retention period is assigned. Times are expressed in the format "*<integer> <time>*", where *<time>* is one of the following values:

Y

Specifies years

M

Specifies months

W

Specifies weeks

D

Specifies days

H

Specifies hours

m

Specifies minutes

s

Specifies seconds

`Forever` indicates that WORM committed files are retained permanently by default. `Use Min` indicates that the default retention period is equal to the minimum retention date. `Use Max` indicates that the default retention period is equal to the maximum retention date.

Minimum retention period

The minimum retention period for the directory. Files are retained in a WORM state for at least the specified amount of time, even if a user specifies an expiration date that results in a shorter retention period.
Times are expressed in the format "*<integer> <time>*", where *<time>* is one of the following values:

Y

Specifies years

M

Specifies months

W

Specifies weeks

D

Specifies days

H

Specifies hours

m

Specifies minutes

s

Specifies seconds

`Forever` indicates that all WORM committed files are retained permanently.

Maximum retention period

The maximum retention period for the directory. Files cannot be retained in a WORM state for more than the specified amount of time, even if a user specifies an expiration date that results in a longer retention period.
Times are expressed in the format "*<integer> <time>*", where *<time>* is one of the following values:

Y

Specifies years

M

Specifies months

W

Specifies weeks

D

Specifies days

H

Specifies hours

m

Specifies minutes

s

Specifies seconds

`Forever` indicates that there is no maximum retention period.

Total modifies

The total number of times that SmartLock settings have been modified for the directory. You can modify SmartLock settings only 32 times per directory.

Managing files in SmartLock directories

You can commit files in SmartLock directories to a WORM state by removing the read-write privileges of the file. You can also set a specific date at which the retention period of the file expires. Once a file is committed to a WORM state, you can increase the retention period of the file, but you cannot decrease the retention period of the file. You cannot move a file that has been committed to a WORM state, even after the retention period for the file has expired.

The retention period expiration date is set by modifying the access time of a file. In a UNIX command line, the access time can be modified through the `touch` command. Although there is no method of modifying the access time through Windows Explorer, you can modify the access time through Windows Powershell. Accessing a file does not set the retention period expiration date.

If you run the `touch` command on a file in a SmartLock directory without specifying a date on which to release the file from a SmartLock state, and you commit the file, the retention period is automatically set to the minimum retention period specified for the SmartLock directory. If you have not specified a minimum retention period for the SmartLock directory, the file is assigned a retention period of zero seconds. It is recommended that you specify a minimum retention period for all SmartLock directories.

Set a retention period through a UNIX command line

You can specify when a file will be released from a WORM state through a UNIX command line.

Procedure

1. Open a connection to any node in the cluster through a UNIX command line and log in.
2. Set the retention period by modifying the access time of the file through the `touch` command.

   The following command sets an expiration date of June 1, 2015 for `/ifs/data/test.txt`:

   ```
   touch -at 201506010000 /ifs/data/test.txt
   ```

Set a retention period through Windows Powershell

You can specify when a file will be released from a WORM state through Microsoft Windows Powershell.

Procedure

1. Open the Windows PowerShell command prompt.
2. (Optional) Establish a connection to the cluster by running the `net use` command.

   The following command establishes a connection to the `/ifs` directory on cluster.ip.address.com:

   ```
   net use "\\cluster.ip.address.com\ifs" /user:root password
   ```

3. Specify the name of the file you want to set a retention period for by creating an object.

The file must exist in a SmartLock directory.

The following command creates an object for `/smartlock/file.txt`:

```
$file = Get-Item "\\cluster.ip.address.com\ifs\smartlock\file.txt"
```

4. Specify the retention period by setting the last access time for the file.

   The following command sets an expiration date of July 1, 2015 at 1:00 PM:

   ```
   $file.LastAccessTime = Get-Date "2015/7/1 1:00 pm"
   ```

Commit a file to a WORM state through a UNIX command line

You can commit a file to a WORM state through a UNIX command line.

To commit a file to a WORM state, you must remove all write privileges from the file. If a file is already set to a read-only state, you must first add write privileges to the file, and then return the file to a read-only state.

Procedure

1. Open a connection to the cluster through a UNIX command line interface and log in.
2. Remove write privileges from a file by running the `chmod` command.

   The following command removes write privileges of `/ifs/data/smartlock/file.txt`:

   ```
   chmod ugo-w /ifs/data/smartlock/file.txt
   ```

Commit a file to a WORM state through Windows Explorer

You can commit a file to a WORM state through Microsoft Windows Explorer. This procedure describes how to commit a file through Windows 7.

To commit a file to a WORM state, you must apply the read-only setting. If a file is already set to a read-only state, you must first remove the file from a read-only state and then return it to a read-only state.

Procedure

1. In Windows Explorer, navigate to the file you want to commit to a WORM state.
2. Right-click the folder and then click **Properties**.
3. In the **Properties** window, click the **General** tab.
4. Select the **Read-only** check box, and then click **OK**.

Override the retention period for all files in a SmartLock directory

You can override the retention period for files in a SmartLock directory. All files committed to a WORM state within the directory will remain in a WORM state until after the specified day. This procedure is available only through the command-line interface (CLI).

If files are committed to a WORM state after the retention period is overridden, the override date functions as a minimum retention date. All files committed to a WORM state do not expire until at least the given day, regardless of user specifications.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.

2. Override the retention period expiration date for all WORM committed files in a SmartLock directory by running the `isi worm modify` command.

   For example, the following command overrides the retention period expiration date of `/ifs/data/SmartLock/directory1` to June 1, 2014:

   ```
   isi worm domains modify /ifs/data/SmartLock/directory1 \
   --override-date 2014-06-01
   ```

Delete a file committed to a WORM state

You can delete a WORM committed file before the expiration date only if you are logged in as the root user or compliance administrator. This procedure is available only through the command-line interface (CLI).

Before you begin

Privileged delete functionality must not be permanently disabled for the SmartLock directory that contains the file.

Procedure

1. Open a connection to the cluster through a UNIX command line and log in through either the root user or compliance administrator account.
2. If privileged delete functionality was disabled for the SmartLock directory, modify the directory by running the `isi worm domains modify` command with the `--privileged-delete` option.

   The following command enables privileged delete for `/ifs/data/SmartLock/directory1`:

   ```
   isi worm domains modify /ifs/data/SmartLock/directory1 \
   --privileged-delete true
   ```
3. Delete the WORM committed file by running the `isi worm files delete` command.

   The following command deletes `/ifs/data/SmartLock/directory1/file`:

   ```
   isi worm files delete /ifs/data/SmartLock/directory1/file
   ```

   The system displays output similar to the following:

   ```
   Are you sure? (yes, [no]):
   ```
4. Type **yes** and then press ENTER.

View WORM status of a file

You can view the WORM status of an individual file. This procedure is available only through the command-line interface (CLI).

Procedure

1. Open a connection to the cluster through a UNIX command line.
2. View the WORM status of a file by running the `isi worm files view` command.

   For example, the following command displays the WORM status of a file:

   ```
   isi worm files view /ifs/data/SmartLock/directory1/file
   ```

The system displays output similar to the following:

```
WORM Domains
ID    Root Path
------------------------------------
65539 /ifs/data/SmartLock/directory1

WORM State: COMMITTED
   Expires: 2015-06-01T00:00:00
```

CHAPTER 16

Protection domains

This section contains the following topics:

Protection domains overview

Protection domains are markers that prevent modifications to files and directories. If a domain is applied to a directory, the domain is also applied to all of the files and subdirectories under the directory. You can specify domains manually; however, OneFS usually creats domains automatically.

There are three types of domains: SyncIQ, SmartLock, and SnapRevert. SyncIQ domains can be assigned to source and target directories of replication policies. OneFS automatically creates a SyncIQ domain for the target directory of a replication policy the first time that the policy is run. OneFS also automatically creates a SyncIQ domain for the source directory of a replication policy during the failback process. You can manually create a SyncIQ domain for a source directory before you initiate the failback process, but you cannot delete a SyncIQ domain that mark the target directory of a replication policy.

SmartLock domains are assigned to SmartLock directories to prevent committed files from being modified or deleted. OneFS automatically creates a SmartLock domain when a SmartLock directory is created. You cannot delete a SmartLock domain. However, if you delete a SmartLock directory, OneFS automatically deletes the SmartLock domain associated with the directory.

SnapRevert domains are assigned to directories that are contained in snapshots to prevent files and directories from being modified while a snapshot is being reverted. OneFS does not automatically create SnapRevert domains. You cannot revert a snapshot until you create a SnapRevert domain for the directory that the snapshot contains. You can create SnapRevert domains for subdirectories of directories that already have SnapRevert domains. For example, you could create SnapRevert domains for both `/ifs/data` and `/ifs/data/archive`. You can delete a SnapRevert domain if you no longer want to revert snapshots of a directory.

Protection domain considerations

You can manually create protection domains before they are required by OneFS to perform certain actions. However, manually creating protection domains can limit your ability to interact with the data marked by the domain.

- Copying a large number of files into a protection domain might take a very long time because each file must be marked individually as belonging to the protection domain.
- You cannot move directories in or out of protection domains. However, you can move a directory contained in a protection domain to another location within the same protection domain.
- Creating a protection domain for a directory that contains a large number of files will take more time than creating a protection domain for a directory with fewer files. Because of this, it is recommended that you create protection domains for directories while the directories are empty, and then add files to the directory.
- If a domain is currently preventing the modification or deletion of a file, you cannot create a protection domain for a directory that contains that file. For example, if `/ifs/data/smartlock/file.txt` is set to a WORM state by a SmartLock domain, you cannot create a SnapRevert domain for `/ifs/data/`.

Create a protection domain

You can create replication or snapshot revert domains to facilitate snapshot revert and failover operations. You cannot create a SmartLock domain. OneFS automatically creates a SmartLock domain when you create a SmartLock directory.

Procedure

1. Click **Cluster Management** › **Job Operations** › **Job Types**.
2. In the **Job Types** area, in the **DomainMark** row, from the **Actions** column, select **Start Job**.
3. In the **Domain Root Path** field, type the path of the directory you want to create a protection domain for.
4. From the **Type of domain** list, specify the type of domain you want to create.
5. Ensure that the **Delete this domain** check box is cleared.
6. Click **Start Job**.

Delete a protection domain

You can delete a replication or snapshot revert domain if you want to move directories out of the domain. You cannot delete a SmartLock domain. OneFS automatically deletes a SmartLock domain when you delete a SmartLock directory.

Procedure

1. Click **Cluster Management** › **Job Operations** › **Job Types**.
2. In the **Job Types** area, in the **DomainMark** row, from the **Actions** column, select **Start Job**.
3. In the **Domain Root Path** field, type the path of the directory you want to delete a protection domain for.
4. From the **Type of domain** list, specify the type of domain you want to delete.
5. Select **Delete this domain**.
6. Click **Start Job**.

CHAPTER 17

Data-at-rest-encryption

This section contains the following topics:

Data-at-rest encryption overview

You can enhance data security with a EMC Isilon cluster that contains only self-encrypting-drive nodes, providing data-at-rest protection.

The OneFS system is available as a cluster that is composed of Isilon OneFS nodes that contain only self-encrypting drives (SEDs). The system requirements and management of data at rest on self-encrypting nodes are identical to that of nodes that do not contain self-encrypting drives. Clusters of mixed node types are not supported.

Self-encrypting drives

Self-encrypting drives store data on a EMC Isilon cluster that is specially designed for data-at-rest encryption.

Data-at-rest- encryption on self-encrypted drives occurs when data that is stored on a device is encrypted to prevent unauthorized data access. All data written to the storage device is encrypted when it is stored, and all data read from the storage device is decrypted when it is read. The stored data is encrypted with a 256-bit data AES encryption key and decrypted in the same manner. OneFS controls data access by combining the drive authentication key with on-disk data-encryption keys.

Note

All nodes in a cluster must be of the self-encrypting drive type. Mixed nodes are not supported.

Data security on self-encrypted drives

Smartfailing self-encrypted drives guarantees data security after removal.

Data on self-encrypted drives is protected from unauthorized access by authenticating encryption keys. Encryption keys never leave the drive. When a drive is locked, successful authentication unlocks the drive for data access.

The data on self-encrypted drives is rendered inaccessible in the following conditions:

- When a self-encrypting drive is smartfailed, drive authentication keys are deleted from the node. The data on the drive cannot be decrypted and is therefore unreadable, which secures the drive.
- When a drive is smartfailed and removed from a node, the encryption key on the drive is removed. Because the encryption key for reading data from the drive must be the same key that was used when the data was written, it is impossible to decrypt data that was previously written to the drive. When you smartfail and then remove a drive, it is cryptographically erased.

  Note

  Smartfailing a drive is the preferred method for removing a self-encrypted drive. Removing a node that has been smartfailed guarantees that data is inaccessible.
- When a self-encrypting drive loses power, the drive locks to prevent unauthorized access. When power is restored, data is again accessible when the appropriate drive authentication key is provided.

Data migration to a self-encrypted-drives cluster

You can migrate data from your existing cluster to a cluster of self-encrypted-drive nodes.

The Isilon cluster does not support the coexistence of regular and self-encrypted nodes. However, if you have data on an existing Isilon cluster that you want to migrate to a cluster of self-encrypted nodes, you can add self-encrypted nodes to your existing cluster one time only to migrate your data.

Note

Before you begin the data-migration process, both clusters must be upgraded to the same OneFS version.

During data migration, an error is generated that indicates you are running in mixed mode, which is not supported and is not secure. The data migrated to the self-encrypted drives is not secure until the smartfail process is completed for the non-encrypted drives.

⚠ CAUTION

Data migration to a cluster of self-encrypted-drive nodes must be performed by Isilon Professional Services. For more information, contact your EMC Isilon representative.

Chassis and drive states

You can view chassis and drive state details.

In a cluster, the combination of nodes in different degraded states determines whether read requests, write requests, or both work. A cluster can lose write quorum but keep read quorum. OneFS provides details about the status of chassis and drives in your cluster. The following table describes all the possible states that you may encounter in your cluster.

State	Description	Interface	Error state
`HEALTHY`	All drives in the node are functioning correctly.	Command-line interface, web administration interface	
`SMARTFAIL` or `Smartfail or restripe in progress`	The drive is in the process of being removed safely from the file system, either because of an I/O error or by user request. Nodes or drives in a smartfail or read-only state affect only write quorum.	Command-line interface, web administration interface	
`NOT AVAILABLE`	A drive is unavailable for a variety of reasons. You can click the bay to view detailed information about this condition.	Command-line interface, web administration interface	X

State	Description	Interface	Error state
	Note In the web administration interface, this state includes the `ERASE` and `SED_ERROR` command-line interface states.		
`SUSPENDED`	This state indicates that drive activity is temporarily suspended and the drive is not in use. The state is manually initiated and does not occur during normal cluster activity.	Command-line interface, web administration interface	
`NOT IN USE`	A node in an offline state affects both read and write quorum.	Command-line interface, web administration interface	
`REPLACE`	The drive was smartfailed successfully and is ready to be replaced.	Command-line interface only	
`STALLED`	The drive is stalled and undergoing stall evaluation. Stall evaluation is the process of checking drives that are slow or having other issues. Depending on the outcome of the evaluation, the drive may return to service or be smartfailed. This is a transient state.	Command-line interface only	
`NEW`	The drive is new and blank. This is the state that a drive is in when you run the `isi dev` command with the `-a add` option.	Command-line interface only	
`USED`	The drive was added and contained an Isilon GUID but the drive is not from this node. This drive likely will be formatted into the cluster.	Command-line interface only	
`PREPARING`	The drive is undergoing a format operation. The drive state changes to HEALTHY when the format is successful.	Command-line interface only	
`EMPTY`	No drive is in this bay.	Command-line interface only	
`WRONG_TYPE`	The drive type is wrong for this node. For example, a non-SED drive in a SED node, SAS instead of the expected SATA drive type.	Command-line interface only	
`BOOT_DRIVE`	Unique to the A100 drive, which has boot drives in its bays.	Command-line interface only	

State	Description	Interface	Error state
SED_ERROR	The drive cannot be acknowledged by the OneFS system. **Note** In the web administration interface, this state is included in Not available.	Command-line interface, web administration interface	X
ERASE	The drive is ready for removal but needs your attention because the data has not been erased. You can erase the drive manually to guarantee that data is removed. **Note** In the web administration interface, this state is included in Not available.	Command-line interface only	
INSECURE	Data on the self-encrypted drive is accessible by unauthorized personnel. Self-encrypting drives should never be used for non-encrypted data purposes. **Note** In the web administration interface, this state is labeled Unencrypted SED.	Command-line interface only	X
UNENCRYPTED SED	Data on the self-encrypted drive is accessible by unauthorized personnel. Self-encrypting drives should never be used for non-encrypted data purposes. **Note** In the command-line interface, this state is labeled INSECURE.	Web administration interface only	X

Smartfailed drive REPLACE state

You can see different drive states during the smartfail process.

If you run the `isi dev` command while the drive in bay 1 is being smartfailed, the system displays output similar to the following example:

```
Node 1, [ATTN]
  Bay 1        Lnum 11      [SMARTFAIL]    SN:Z296M8HK
000093172YE04  /dev/da1
  Bay 2        Lnum 10      [HEALTHY]      SN:Z296M8N5
00009330EYE03  /dev/da2
  Bay 3        Lnum 9       [HEALTHY]      SN:Z296LBP4
00009330EYE03  /dev/da3
  Bay 4        Lnum 8       [HEALTHY]      SN:Z296LCJW
00009327BYE03  /dev/da4
  Bay 5        Lnum 7       [HEALTHY]      SN:Z296M8XB
00009330KYE03  /dev/da5
  Bay 6        Lnum 6       [HEALTHY]      SN:Z295LXT7
000093172YE03  /dev/da6
  Bay 7        Lnum 5       [HEALTHY]      SN:Z296M8ZF
00009330KYE03  /dev/da7
  Bay 8        Lnum 4       [HEALTHY]      SN:Z296M8SD
00009330EYE03  /dev/da8
  Bay 9        Lnum 3       [HEALTHY]      SN:Z296M8QA
00009330EYE03  /dev/da9
  Bay 10       Lnum 2       [HEALTHY]      SN:Z296M8Q7
00009330EYE03  /dev/da10
  Bay 11       Lnum 1       [HEALTHY]      SN:Z296M8SP
00009330EYE04  /dev/da11
  Bay 12       Lnum 0       [HEALTHY]      SN:Z296M8QZ
00009330JYE03  /dev/da12
```

If you run the `isi dev` command after the smartfail completes successfully, the system displays output similar to the following example, showing the drive state as `REPLACE`:

```
Node 1, [ATTN]
  Bay 1        Lnum 11      [REPLACE]      SN:Z296M8HK
000093172YE04  /dev/da1
  Bay 2        Lnum 10      [HEALTHY]      SN:Z296M8N5
00009330EYE03  /dev/da2
  Bay 3        Lnum 9       [HEALTHY]      SN:Z296LBP4
00009330EYE03  /dev/da3
  Bay 4        Lnum 8       [HEALTHY]      SN:Z296LCJW
00009327BYE03  /dev/da4
  Bay 5        Lnum 7       [HEALTHY]      SN:Z296M8XB
00009330KYE03  /dev/da5
  Bay 6        Lnum 6       [HEALTHY]      SN:Z295LXT7
000093172YE03  /dev/da6
  Bay 7        Lnum 5       [HEALTHY]      SN:Z296M8ZF
00009330KYE03  /dev/da7
  Bay 8        Lnum 4       [HEALTHY]      SN:Z296M8SD
00009330EYE03  /dev/da8
  Bay 9        Lnum 3       [HEALTHY]      SN:Z296M8QA
00009330EYE03  /dev/da9
  Bay 10       Lnum 2       [HEALTHY]      SN:Z296M8Q7
00009330EYE03  /dev/da10
  Bay 11       Lnum 1       [HEALTHY]      SN:Z296M8SP
00009330EYE04  /dev/da11
  Bay 12       Lnum 0       [HEALTHY]      SN:Z296M8QZ
00009330JYE03  /dev/da12
```

If you run the `isi dev` command while the drive in bay 3 is being smartfailed, the system displays output similar to the following example:

```
Node 1, [ATTN]
  Bay 1        Lnum 11      [REPLACE]      SN:Z296M8HK
000093172YE04  /dev/da1
  Bay 2        Lnum 10      [HEALTHY]      SN:Z296M8N5
00009330EYE03  /dev/da2
  Bay 3        Lnum 9       [SMARTFAIL]    SN:Z296LBP4
00009330EYE03  N/A
  Bay 4        Lnum 8       [HEALTHY]      SN:Z296LCJW
00009327BYE03  /dev/da4
  Bay 5        Lnum 7       [HEALTHY]      SN:Z296M8XB
00009330KYE03  /dev/da5
  Bay 6        Lnum 6       [HEALTHY]      SN:Z295LXT7
000093172YE03  /dev/da6
  Bay 7        Lnum 5       [HEALTHY]      SN:Z296M8ZF
00009330KYE03  /dev/da7
  Bay 8        Lnum 4       [HEALTHY]      SN:Z296M8SD
00009330EYE03  /dev/da8
  Bay 9        Lnum 3       [HEALTHY]      SN:Z296M8QA
00009330EYE03  /dev/da9
  Bay 10       Lnum 2       [HEALTHY]      SN:Z296M8Q7
00009330EYE03  /dev/da10
  Bay 11       Lnum 1       [HEALTHY]      SN:Z296M8SP
00009330EYE04  /dev/da11
  Bay 12       Lnum 0       [HEALTHY]      SN:Z296M8QZ
00009330JYE03  /dev/da12
```

Smartfailed drive ERASE state

At the end of a smartfail process, OneFS attempts to delete the authentication key on a drive if it is unable to reset the key.

Note

- To securely delete the authentication key on a single drive, smartfail the individual drive.
- To securely delete the authentication key on a single node, smartfail the node.
- To securely delete the authentication keys on an entire cluster, smartfail each node and run the `isi_reformat_node` command on the last node.

Upon running the `isi dev` command, the system displays output similar to the following example, showing the drive state as ERASE:

```
Node 1, [ATTN]
  Bay 1        Lnum 11      [REPLACE]      SN:Z296M8HK
000093172YE04  /dev/da1
  Bay 2        Lnum 10      [HEALTHY]      SN:Z296M8N5
00009330EYE03  /dev/da2
  Bay 3        Lnum 9       [ERASE]        SN:Z296LBP4
00009330EYE03  /dev/da3
```

Drives showing the ERASE state can be safely retired, reused, or returned.

Any further access to a drive showing the ERASE state requires the authentication key of the drive to be set to its default manufactured security ID (MSID). This action erases the data encryption key (DEK) on the drive and renders any existing data on the drive permanently unreadable.

CHAPTER 18

SmartQuotas

This section contains the following topics:

SmartQuotas overview

The SmartQuotas module is an optional quota-management tool that monitors and enforces administrator-defined storage limits. Using accounting and enforcement quota limits, reporting capabilities, and automated notifications, SmartQuotas manages storage use, monitors disk storage, and issues alerts when disk-storage limits are exceeded.

Quotas help you manage storage usage according to criteria that you define. Quotas are used as a method of tracking—and sometimes limiting—the amount of storage that a user, group, or project consumes. Quotas are a useful way of ensuring that a user or department does not infringe on the storage that is allocated to other users or departments. In some quota implementations, writes beyond the defined space are denied, and in other cases, a simple notification is sent.

The SmartQuotas module requires a separate license. For additional information about the SmartQuotas module or to activate the module, contact your EMC Isilon sales representative.

Quota types

OneFS uses the concept of quota types as the fundamental organizational unit of storage quotas. Storage quotas comprise a set of resources and an accounting of each resource type for that set. Storage quotas are also called storage domains.

Storage quotas creation requires three identifiers:

- The directory to monitor.
- Whether snapshots are to be tracked against the quota limit.
- The quota type (directory, user, or group).

You can choose a quota type from the following entities:

Directory
: A specific directory and its subdirectories.

User
: Either a specific user or default user (every user). Specific-user quotas that you configure take precedence over a default user quota.

Group
: All members of a specific group or all members of a default group (every group). Any specific-group quotas that you configure take precedence over a default group quota. Associating a group quota with a default group quota creates a linked quota.

You can create multiple quota types on the same directory, but they must be of a different type or have a different snapshot option. You can specify quota types for any directory in OneFS and nest them within each other to create a hierarchy of complex storage-use policies.

Nested storage quotas can overlap. For example, the following quota settings ensure that the finance directory never exceeds 5 TB, while limiting the users in the finance department to 1 TB each:

- Set a 5 TB hard quota on `/ifs/data/finance`.
- Set 1 TB soft quotas on each user in the finance department.

Note

You should not create quotas of any type on the OneFS root (`/ifs`). A root-level quota may significantly degrade performance.

Default quota type

Default quotas automatically create other quotas for users or groups in a specified directory.

A default quota specifies a policy for new entities that match a trigger. The default-user@/ifs/cs becomes specific-user@/ifs/cs for each specific-user that is not otherwise defined.

For example, you can create a default-user quota on the `/ifs/dir-1` directory, where that directory is owned by the root user. The default-user type automatically creates a new domain on that directory for root and adds the usage there:

```
my-OneFS-1# mkdir /ifs/dir-1
my-OneFS-1# isi quota quotas create /ifs/dir-1 default-user
my-OneFS-1# isi quota quotas ls --path=/ifs/dir-1
Type         AppliesTo  Path       Snap  Hard  Soft  Adv  Used
---------------------------------------------------------------
default-user DEFAULT    /ifs/dir-1 No    -     -     -    0b
user         root       /ifs/dir-1 No    -     -     -    0b
---------------------------------------------------------------
```

Now add a file that is owned by a different user (admin).

```
my-OneFS-1# touch /ifs/dir-1/somefile
my-OneFS-1# chown admin /ifs/dir-1/somefile
my-OneFS-1# isi quota quotas ls --path=/ifs/dir-1
Type         AppliesTo  Path       Snap  Hard  Soft  Adv  Used
---------------------------------------------------------------
default-user DEFAULT    /ifs/dir-1 No    -     -     -    0b
user         root       /ifs/dir-1 No    -     -     -    26b
user         admin      /ifs/dir-1 No    -     -     -    0b
---------------------------------------------------------------
Total: 3
```

In this example, the default-user type created a new specific-user type automatically (user:admin) and added the new usage to it. Default-user does not have any usage because it is used only to generate new quotas automatically. Default-user enforcement is copied to a specific-user (user:admin), and the inherited quota is called a linked quota. In this way, each user account gets its own usage accounting.

Defaults can overlap. For example, default-user@/ifs/dir-1 and default-user@/ifs/cs both may be defined. If the default enforcement changes, OneFS storage quotas propagate the changes to the linked quotas asynchronously. Because the update is asynchronous, there is some delay before updates are in effect. If a default type, such as every user or every group, is deleted, OneFS deletes all children that are marked as inherited. As an option, you can delete the default without deleting the children, but it is important to note that this action breaks inheritance on all inherited children.

Continuing with the example, add another file that is owned by the root user. Because the root type exists, the new usage is added to it.

```
my-OneFS-1# touch /ifs/dir-1/anotherfile
my-OneFS-1# isi quota ls -v --path=/ifs/dir-1 --format=list
      Type: default-user
```

```
 AppliesTo: DEFAULT
      Path: /ifs/dir-1
      Snap: No
Thresholds
          Hard : -
          Soft : -
           Adv : -
         Grace : -
     Usage
            Files : 0
    With Overhead : 0.00b
     W/O Overhead : 0.00b
      Over: -
  Enforced: No
 Container: No
    Linked: -
----------------------------------------------------------------------
      Type: user
 AppliesTo: root
      Path: /ifs/dir-1
      Snap: No
Thresholds
          Hard : -
          Soft : -
           Adv : -
         Grace : -
     Usage
            Files : 2
    With Overhead : 3.50K
     W/O Overhead : 55.00b
      Over: -
  Enforced: No
 Container: No
    Linked: Yes
----------------------------------------------------------------------
-
      Type: user
 AppliesTo: admin
      Path: /ifs/dir-1
      Snap: No
Thresholds
          Hard : -
          Soft : -
           Adv : -
         Grace : -
     Usage
            Files : 1
    With Overhead : 1.50K
     W/O Overhead : 0.00b
      Over: -
  Enforced: No
 Container: No
    Linked: Yes
```

The enforcement on default-user is copied to the specific-user when the specific-user allocates within the type, and the new inherited quota type is also a linked quota.

Note

Configuration changes for linked quotas must be made on the parent quota that the linked quota is inheriting from. Changes to the parent quota are propagated to all children. To override configuration from the parent quota, you must unlink the quota first.

Usage accounting and limits

Storage quotas support two usage types that you can create to manage storage space. The usage types are accounting and enforcement limits.

You can configure OneFS quotas by usage type to track or limit storage use. The accounting option, which monitors disk-storage use, is useful for auditing, planning, and billing. Enforcement limits set storage limits for users, groups, or directories.

Accounting

The accounting option tracks but does not limit disk-storage use. Using the accounting option for a quota, you can monitor inode count and physical and logical space resources. Physical space refers to all of the space used to store files and directories, including data and metadata in the domain. Logical space refers to the sum of all files sizes, excluding file metadata and sparse regions. User data storage is tracked using logical-space calculations, which do not include protection overhead. As an example, by using the accounting option, you can do the following:

- Track the amount of disk space used by various users or groups to bill each user, group, or directory for only the disk space used.
- Review and analyze reports that help you identify storage usage patterns and define storage policies.
- Plan for capacity and other storage needs.

Enforcement limits

Enforcement limits include all of the functionality of the accounting option, plus the ability to limit disk storage and send notifications. Using enforcement limits, you can logically partition a cluster to control or restrict how much storage that a user, group, or directory can use. For example, you can set hard- or soft-capacity limits to ensure that adequate space is always available for key projects and critical applications and to ensure that users of the cluster do not exceed their allotted storage capacity. Optionally, you can deliver real-time email quota notifications to users, group managers, or administrators when they are approaching or have exceeded a quota limit.

Note

If a quota type uses the accounting-only option, enforcement limits cannot be used for that quota.

The actions of an administrator logged in as root may push a domain over a quota threshold. For example, changing the protection level or taking a snapshot has the potential to exceed quota parameters. System actions such as repairs also may push a quota domain over the limit.

The system provides three types of administrator-defined enforcement thresholds.

Threshold type	Description
Hard	Limits disk usage to a size that cannot be exceeded. If an operation, such as a file write, causes a quota target to exceed a hard quota, the following events occur: • the operation fails • an alert is logged to the cluster

Threshold type	Description
	- a notification is issued to specified recipients. Writes resume when the usage falls below the threshold.
Soft	Allows a limit with a grace period that can be exceeded until the grace period expires. When a soft quota is exceeded, an alert is logged to the cluster and a notification is issued to specified recipients; however, data writes are permitted during the grace period. If the soft threshold is still exceeded when the grace period expires, data writes fail, and a hard-limit notification is issued to the recipients you have specified. Writes resume when the usage falls below the threshold.
Advisory	An informational limit that can be exceeded. When an advisory quota threshold is exceeded, an alert is logged to the cluster and a notification is issued to specified recipients. Advisory thresholds do not prevent data writes.

Disk-usage calculations

For each quota that you configure, you can specify whether data-protection overhead is included in future disk-usage calculations.

Most quota configurations do not need to include overhead calculations. If you do not include data-protection overhead in usage calculations for a quota, future disk-usage calculations for the quota include only the space that is required to store files and directories. Space that is required for the data-protection setting of the cluster is not included.

Consider the same example user, who is now restricted by a 40 GB quota that does not include data-protection overhead in its disk-usage calculations. If your cluster is configured with a 2x data-protection level and the user writes a 10 GB file to the cluster, that file consumes 20 GB of space but the 10GB for the data-protection overhead is not counted in the quota calculation. In this example, the user has reached 25 percent of the 40 GB quota by writing a 10 GB file to the cluster. This method of disk-usage calculation is recommended for most quota configurations.

If you include data-protection overhead in usage calculations for a quota, future disk-usage calculations for the quota include the total amount of space that is required to store files and directories, in addition to any space that is required to accommodate your data-protection settings, such as parity or mirroring. For example, consider a user who is restricted by a 40 GB quota that includes data-protection overhead in its disk-usage calculations. If your cluster is configured with a 2x data-protection level (mirrored) and the user writes a 10 GB file to the cluster, that file actually consumes 20 GB of space: 10 GB for the file and 10 GB for the data-protection overhead. In this example, the user has reached 50 percent of the 40 GB quota by writing a 10 GB file to the cluster.

Note

Cloned and deduplicated files are treated as ordinary files by quotas. If the quota includes data protection overhead, the data protection overhead for shared data is not included in the usage calculation.

You can configure quotas to include the space that is consumed by snapshots. A single path can have two quotas applied to it: one without snapshot usage, which is the default, and one with snapshot usage. If you include snapshots in the quota, more files are included in the calculation than are in the current directory. The actual disk usage is the sum of the current directory and any snapshots of that directory. You can see which snapshots are included in the calculation by examining the `.snapshot` directory for the quota path.

Note

Only snapshots created after the QuotaScan job finishes are included in the calculation.

Quota notifications

Quota notifications are generated for enforcement quotas, providing users with information when a quota violation occurs. Reminders are sent periodically while the condition persists.

Each notification rule defines the condition that is to be enforced and the action that is to be executed when the condition is true. An enforcement quota can define multiple notification rules. When thresholds are exceeded, automatic email notifications can be sent to specified users, or you can monitor notifications as system alerts or receive emails for these events.

Notifications can be configured globally, to apply to all quota domains, or be configured for specific quota domains.

Enforcement quotas support the following notification settings. A given quota can use only one of these settings.

Limit notification settings	Description
Turn Off Notifications for this Quota	Disables all notifications for the quota.
Use Default Notification Rules	Uses the global default notification for the specified type of quota.
Use Custom Notification Rules	Enables the creation of advanced, custom notifications that apply to the specific quota. Custom notifications can be configured for any or all of the threshold types (hard, soft, or advisory) for the specified quota.

Quota notification rules

You can write quota notification rules to generate alerts that are triggered by event thresholds.

When an event occurs, a notification is triggered according to your notification rule. For example, you can create a notification rule that sends an email when a disk-space allocation threshold is exceeded by a group.

You can configure notification rules to trigger an action according to event thresholds (a notification condition). A rule can specify a schedule, such as "every day at 1:00 AM," for executing an action or immediate notification of certain state transitions. When an event occurs, a notification trigger may execute one or more actions, such as sending an email or sending a cluster alert to the interface. The following examples demonstrate the types of criteria that you can use to configure notification rules.

- Notify when a threshold is exceeded; at most, once every 5 minutes
- Notify when allocation is denied; at most, once an hour
- Notify while over threshold, daily at 2 AM
- Notify while grace period expired weekly, on Sundays at 2 AM

Notifications are triggered for events grouped by the following categories:

Instant notifications

Includes the write-denied notification, triggered when a hard threshold denies a write, and the threshold-exceeded notification, triggered at the moment a hard, soft, or advisory threshold is exceeded. These are one-time notifications because they represent a discrete event in time.

Ongoing notifications

Generated on a scheduled basis to indicate a persisting condition, such as a hard, soft, or advisory threshold being over a limit or a soft threshold's grace period being expired for a prolonged period.

Quota reports

The OneFS SmartQuotas module provides reporting options that enable administrators to manage cluster resources and analyze usage statistics.

Storage quota reports provide a summarized view of the past or present state of the quota domains. After raw reporting data is collected by OneFS, you can produce data summaries by using a set of filtering parameters and sort types. Storage-quota reports include information about violators, grouped by threshold types. You can generate reports from a historical data sample or from current data. In either case, the reports are views of usage data at a given time. OneFS does not provide reports on data aggregated over time, such as trending reports, but you can use raw data to analyze trends. There is no configuration limit on the number of reports other than the space needed to store them.

OneFS provides the following data-collection and reporting methods:

- Scheduled reports are generated and saved on a regular interval.
- Ad hoc reports are generated and saved at the request of the user.
- Live reports are generated for immediate and temporary viewing.

Scheduled reports are placed by default in the `/ifs/.isilon/smartquotas/reports` directory, but the location is configurable to any directory under `/ifs`. Each generated report includes quota domain definition, state, usage, and global configuration settings. By default, ten reports are kept at a time, and older reports are purged. You can create ad hoc reports at any time to view the current state of the storage quotas system. These live reports can be saved manually. Ad hoc reports are saved to a location that is separate from scheduled reports to avoid skewing the timed-report sets.

Creating quotas

You can create two types of storage quotas to monitor data: accounting quotas and enforcement quotas. Storage quota limits and restrictions can apply to specific users, groups, or directories.

The type of quota that you create depends on your goal.

- Enforcement quotas monitor and limit disk usage. You can create enforcement quotas that use any combination of hard limits, soft limits, and advisory limits.

> **Note**
>
> Enforcement quotas are not recommended for snapshot-tracking quota domains.

- Accounting quotas monitor, but do not limit, disk usage.

Note

After you create a new quota, it begins to report data almost immediately, but the data is not valid until the QuotaScan job completes. Before using quota data for analysis or other purposes, verify that the QuotaScan job has finished.

Create an accounting quota

You can create an accounting quota to monitor but not limit disk usage.

Optionally, you can include snapshot data, data-protection overhead, or both in the accounting quota.

Procedure

1. Click **File System Management** › **SmartQuotas** › **Quotas & Usage**.
2. On the **Storage Quotas & Usage** page, click **Create a storage quota**.
3. From the **Quota Type** list, select the target for this quota: a **directory**, **user**, or **group**.
4. Depending on the target that you selected, select the entity that you want to apply the quota to. For example, if you selected **User** from the **Quota Type** list, you can target either all users or a specific user.
5. In the **Directory path** field, type the path and directory for the quota, or click **Browse**, and then select a directory.
6. (Optional) In the **Usage Accounting** area, select the options that you want.
   - To include snapshot data in the accounting quota, select the **Include Snapshot Data** check box.
   - To include the data-protection overhead in the accounting quota, select the **Include Data-Protection Overhead** check box.
   - To include snapshot data in the accounting quota, select the **Include Snapshot Data** check box.
7. In the **Usage Limits** area, click **No Usage Limit (Accounting Only)**.
8. Click **Create Quota**.

After you finish

After you create a quota, it begins to report data almost immediately, but the data is not valid until the QuotaScan job completes. Before using quota data for analysis or other purposes, verify that the QuotaScan job has finished.

Create an enforcement quota

You can create an enforcement quota to monitor and limit disk usage.

You can create enforcement quotas that set hard, soft, and advisory limits.

Procedure

1. Click **File System Management** › **SmartQuotas** › **Quotas & Usage**.
2. On the **Storage Quotas & Usage** page, click **Create a storage quota**.

3. From the **Quota Type** list, select the target for this quota: a directory, user, or group.
4. Depending on the target that you selected, select the entity that you want to apply the quota to. For example, if you selected User from the **Quota Type** list, you can target all users or a specific user.
5. In the **Directory path** field, type the path and directory for the quota, or click **Browse**, and then select a directory.
6. (Optional) In the **Usage Accounting** area, click the **Include Snapshot Data** check box, the **Include Data-Protection Overhead** check box, or both to include them in the quota.
7. In the **Usage Limits** area, click **Specify Usage Limits**.
8. Click the check box next to the option for each type of limit that you want to enforce.
9. Type numerals in the fields and select from the lists the values that you want to use for the quota.
10. In the **Limit Notations** area, click the notification option that you want to apply to the quota.
11. To generate an event notification, select the **Create cluster event** check box.
12. (Optional) If you selected the option to use custom notification rules, click the link to expand the custom notification type that applies to the usage-limit selections.
13. Click **Create Quota**.

After you finish

After you create a quota, it begins to report data almost immediately but the data is not valid until the QuotaScan job completes. Before using quota data for analysis or other purposes, verify that the QuotaScan job has finished.

Managing quotas

You can modify the configured values of a storage quota, and you can enable or disable a quota. You can also create quota limits and restrictions that apply to specific users, groups, or directories.

Quota management in OneFS is simplified by the quota search feature, which helps you to locate a quota or quotas by using filters. You can unlink quotas that are associated with a parent quota, and configure custom notifications for quotas. You can also disable a quota temporarily and then enable it when needed.

Note

Moving quota directories across quota domains is not supported.

Search for quotas

You can search for a quota using a variety of search criteria.

By default, all storage quotas and display options are listed on this page before you apply report or search filters. If the Quotas & Storage section is collapsed, click **Define quota display**.

Procedure

1. Click **File System Management** › **SmartQuotas** › **Quotas & Usage**.
2. In the **Quotas & Usage** area, for **Report Filters**, select **Search for specific quotas within this report**.

3. In the **Quota Type** list, select the quota type that you want to find.
4. If you selected **User Quota** or **Group quota** for a quota type, type a full or partial user or group name in the **User** or **Group** field.

   You can use the wildcard character (*) in the **User** or **Group** field.

   - To search for only default users, select the **Only show default users** checkbox.
   - To search for only default groups, select the **Only show default groups** check box.
5. In the **Directory Path** field, type a full or partial path.

   You can use the wildcard character (*) in the **Directory Path** field.

   - To search subdirectories, select the **Include subdirectories** check box.
   - To search for only quotas that are in violations, select the **Only show quotas for which usage limits are currently in violation** check box.
6. (Optional) Click **Update Display**.

   Quotas that match the search criteria appear in the sections where quotas are listed.

Results

An accounting or enforcement quota with a threshold value of zero is indicated by a dash (–). You can click the column headings to sort the result set.

Note

To clear the result set and display all storage quotas, in the Quotas & Usage area, select **Show all quotas and usage for this report** for Report Filters, and then click **Update Display**.

Manage quotas

Quotas help you monitor and analyze the current or historical use of disk storage. You can search for quotas, and you can view, modify, delete, and unlink a quota.

An initial QuotaScan job must run for the default or scheduled quotas, or the data displayed may be incomplete.

Before you modify a quota, consider how the changes will affect the file system and end users.

Note

- The options to edit or delete a quota appear only when the quota is not linked to a default quota.
- The option to unlink a quota is available only when the quota is linked to a default quota.

Procedure

1. Click **File System Management › SmartQuotas › Quotas & Usage**.
2. From the **Quota Report** options, select the type of quota report that you want to view or manage.
   - To monitor and analyze current disk storage use, click **Show current quotas and usage (Live Report)**.

- To monitor and analyze historical disk storage use, click **Show archived quota report** to select from the list of archived scheduled and manually generated quota reports.

3. For **Report Filters**, select the filters to be used for this quota report.
   - To view all information in the quota report, click **Show all quotas and usage for this report**.
   - To filter the quota report, click **Search for specific quotas within this report**, and then select the filters that you want to apply.
4. Click **Update Display**.
5. (Optional) Select a quota to view its settings or to perform the following management actions.
   - To review or edit this quota, click **View details**.
   - To delete this quota, click **Delete**.
   - To unlink a linked quota, click **Unlink**.

   Note

   Configuration changes for linked quotas must be made on the parent (default) quota that the linked quota is inheriting from. Changes to the parent quota are propagated to all children. If you want to override configuration from the parent quota, you must first unlink the quota.

Export a quota configuration file

You can export quota settings as a configuration file, which can then be imported for reuse to another Isilon cluster. You can also store the exported quota configurations in a location outside of the cluster. This task may only be performed from the OneFS command line interface.

You can pipe the XML report to a file or directory. The file can then be imported to another cluster.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. At the command prompt, run the following command:

```
isi_classic quota list --export
```

   The quota configuration file displays as raw XML.

Import a quota configuration file

You can import quota settings in the form of a configuration file that has been exported from another Isilon cluster. This task can only be performed from the OneFS command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Navigate to the location of the exported quota configuration file.

3. At the command prompt, run the following command, where *<filename>* is the name of an exported configuration file:

   ```
   isi_classic quota import --from-file=<filename>
   ```

   The system parses the file and imports the quota settings from the configuration file. Quota settings that you configured before importing the quota configuration file are retained, and the imported quota settings are effective immediately.

Managing quota notifications

Quota notifications can be enabled or disabled, modified, and deleted.

By default, a global quota notification is already configured and applied to all quotas. You can continue to use the global quota notification settings, modify the global notification settings, or disable or set a custom notification for a quota.

Enforcement quotas support four types of notifications and reminders:

- Threshold exceeded
- Over-quota reminder
- Grace period expired
- Write access denied

If a directory service is used to authenticate users, you can configure notification mappings that control how email addresses are resolved when the cluster sends a quota notification. If necessary, you can remap the domain that is used for quota email notifications and you can remap Active Directory domains, local UNIX domains, or both.

Configure default quota notification settings

You can configure default global quota notification settings that apply to all quotas of a specified threshold type.

The custom notification settings that you configure for a quota take precedence over the default global notification settings.

Procedure

1. Click **File System Management** › **SmartQuotas** › **Settings**.
2. (Optional) On the Quota Settings page, for **Scheduled Reporting**, select **On**.
3. Click **Change Schedule**, and then select a report frequency from the list.
4. Select the reporting schedule options that you want, and then click **Select**.
5. In the **Scheduled Report Archiving** area, you can configure the following size and directory options:
   - To configure the number of live reports that you want to archive, type the number of reports in the **Limit archive size** field.
   - To specify an archive directory that is different from the default, in the **Archive Directory** field, type the path or click **Browse** to select the path.
6. In the **Manual Report Archiving** area, you can configure the following size and directory options:
   - To configure the number of live reports that you want to archive, type the number of reports in the **Limit archive size** field.

- To specify an archive directory that is different from the default, in the **Archive Directory** field, type the path or click **Browse** to select the path.

7. In the **Email Mapping Rules** area, choose each mapping rule that you want to use by selecting the check box in the **Provider Type** column.
8. In the Notification Rules area, define default notification rules for each rule type.
   - To expand the list of limit notifications rules types, click **Default Notifications Settings**.
   - To display default settings options for advisory-limit notification rules, click **Advisory Limit Notification Rules**.
   - To set the advisory-limit options that you want, click **Event: Advisory Limit Value Exceeded** and **Event: While Advisory Limit Remains Exceeded**.
   - To display default settings for soft-limit notifications, click **Soft Limit Notification Rules**.
   - To set the soft-limit options that you want, click **Event: Soft Limit Value Exceeded**, **Event: While Soft Limit Remains Exceeded**, **Event: Soft Limit Grace Period Expired**, and **Event: Soft Limit Write Access Denied**.
   - To display the options for a hard-limit notification rule, click **Hard Limit Notification Rules** .
   - To set the hard-limit options that you want, click **Event: Hard Limit Write Access Denied** and **Event: While Hard Limit Remains Exceeded**.
9. Click **Save**.

After you finish

After you create a new quota, it begins to report data almost immediately, but the data is not valid until the QuotaScan job completes. Before using quota data for analysis or other purposes, verify that the QuotaScan job has finished.

Configure custom quota notification rules

You can configure custom quota notification rules that apply only to a specified quota.

Before you begin

To configure a custom notification rule, an enforcement quota must exist or be in the process of being created. To configure notifications for an existing enforcement quota, follow the procedure to modify a quota and then use these steps from the Quota Details pane of the specific quota.

Quota-specific custom notification rules must be configured for that quota. If notification rules are not configured for a quota, the default event notification configuration is used. For more information about configuring default notification rules, see Create an event notification rule.

Procedure

1. In the **Limit Notifications** area, click **Use Custom Notification Rules**.

   The links display for the rules options that are available for the type of notification that you have selected for this quota.
2. Click the **View details**, and then click **Edit limit notifications**.
3. Click the link for the limit notification type that you want to configure for this quota. From the list, select the target for this quota: a directory, user, or group.

   The **Limit Notification Rules** options display for the selection type.

4. Select or type the values to configure the custom notification rule for this quota.
5. Click **Create quota** when you have completed configuring the settings for this notification rule.

Results

The quota appears in the **Quotas & Usage** list.

After you finish

After you create a new quota, it begins to report data almost immediately, but the data is not valid until the QuotaScan job completes. Before using quota data for analysis or other purposes, verify that the QuotaScan job has finished.

Map an email notification rule for a quota

Email notification mapping rules control how email addresses are resolved when the cluster sends a quota notification.

If necessary, you can remap the domain used for SmartQuotas email notifications. You can remap Active Directory Windows domains, local UNIX domains, or NIS domains.

Note

You must be logged in to the web administration interface to perform this task.

Procedure

1. Click **File System Management** › **SmartQuotas** › **Settings**.
2. (Optional) In the **Email Mapping** area, click **Create an email mapping rule**.
3. From the **Provider Type** list, select the authentication provider type for this notification rule. The default is `Local`. To determine which authentication providers are available on your cluster, navigate to **Access** › **Authentication Providers**.
4. From the **Current Domain** list, select the domain that you want to use for the mapping rule. If the list is blank, navigate to **Cluster Management** › **Network Configuration**, and then click **Edit** in the **DNS Settings** area to specify the domains that you want to use for mapping.
5. In the **Map-to-Domain** field, type the name of the domain that you want to map email notifications to. This can be the same domain name you selected from the **Current Domain** list. To specify multiple domains, separate the domain names with commas.
6. Click **Save Rule**.

Email quota notification messages

If email notifications for exceeded quotas are enabled, you can customize Isilon-provided templates for email notifications or create your own.

There are three email notification templates provided with OneFS. The templates are located in `/etc/ifs` and are described in the following table:

Template	Description
`quota_email_template.txt`	A notification that disk quota has been exceeded.
`quota_email_grace_template.txt`	A notification that disk quota has been exceeded (also includes a parameter to define a grace period in number of days).

Template	Description
`quota_email_test_template.txt`	A notification test message you can use to verify that a user is receiving email notifications.

If the default email notification templates do not meet your needs, you can configure your own custom email notification templates using a combination of text and SmartQuotas variables. Whether you choose to create your own templates or modify the existing ones, make sure that the first line of the template file is a `Subject:` line. For example:

```
Subject: Disk quota exceeded
```

If you want to include information about the message sender, include a `From:` line immediately under the subject line. If you use an email address, include the full domain name for the address. For example:

```
From:  administrator@abcd.com
```

In this example of the `quota_email_template.txt` file, a `From:` line is included. Additionally, the default text "Contact your system administrator for details" at the end of the template is changed to name the administrator:

```
Subject: Disk quota exceeded
From: administrator@abcd.com

The <ISI_QUOTA_TYPE> disk quota on directory <ISI_QUOTA_PATH>
owned by <ISI_QUOTA_OWNER> on <ISI_QUOTA_NODE> was exceeded.

The quota limit is <ISI_QUOTA_THRESHOLD>, and <ISI_QUOTA_USAGE>
is currently in use. You may be able to free some disk space by
deleting unnecessary files. If your quota includes snapshot usage,
your administrator may be able to free some disk space by deleting
one or more snapshots. Contact Jane Anderson (janderson@abcd.com)
for details.
```

This is an example of a what a user will see as an emailed notification (note the SmartQuotas variables are resolved):

```
Subject: Disk quota exceeded
From: administrator@abcd.com

The advisory disk quota on directory /ifs/data/sales_tools/collateral
owned by jsmith on production-Boris was exceeded.

The quota limit is 10 GB, and 11 GB is in use. You may be able
to free some disk space by deleting unnecessary files. If your
quota includes snapshot usage, your administrator may be able
to free some disk space by deleting one or more snapshots.
Contact Jane Anderson (janderson@abcd.com) for details.
```

Custom email notification template variable descriptions

An email template contains text and, optionally, variables that represent values. You can use any of the SmartQuotas variables in your templates.

Variable	Description	Example
ISI_QUOTA_PATH	Path of quota domain	`/ifs/data`
ISI_QUOTA_THRESHOLD	Threshold value	20 GB

Variable	Description	Example
ISI_QUOTA_USAGE	Disk space in use	10.5 GB
ISI_QUOTA_OWNER	Name of quota domain owner	jsmith
ISI_QUOTA_TYPE	Threshold type	Advisory
ISI_QUOTA_GRACE	Grace period, in days	5 days
ISI_QUOTA_EXPIRATION	Expiration date of grace period	Fri May 22 14:23:19 PST 2015
ISI_QUOTA_NODE	Hostname of the node on which the quota event occurred	someHost-prod-wf-1

Customize email quota notification templates

You can customize Isilon-provided templates for email notifications. This task may only be performed from the OneFS command line interface.

This procedure assumes you are using the provided templates, which are located in the `/etc/ifs` directory.

Note

We recommend that you do not edit the templates directly. Instead, copy them to another directory to edit and deploy them.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Copy one of the default templates to a directory in which you will edit the file and later access it through the OneFS web administration interface. For example:

```
cp /etc/ifs/quota_email_template.txt /ifs/data/quotanotifiers/
quota_email_template_copy.txt
```

3. Open the desired template file in a text editor. For example:

```
edit /ifs/data/quotanotifiers/quota_email_template_copy.txt
```

   The template displays in the editor.
4. Edit the template as desired. If you are using or creating a customized template, make sure the template has a `Subject:` line.
5. Save your changes. Template files must be saved as .txt files.
6. In the web administration interface, navigate to **File System** › **SmartQuotas** › **Quotas & Usage**.
7. Select the desired quota for which you wish to set a notification rule.
8. Click the **Settings** tab.
9. Select the notification rule you wish to use with the template you created (for example, **Advisory Limit Notification Rules**). For expanded information about setting notification rules, refer to the instructions for configuring default quota notification settings and configuring custom quota notification rules in this chapter.
10. Select the desired event for the template (for example, **Event: Advisory Limit Value Exceeded**).

11. In the **Send Email** area, select one of the owner notification type check boxes.
12. In the **Message Template** field, enter or browse to find the template you copied or customized.
13. (Optional) In the **Event** area, select **Create Cluster Event** to generate an event notification in addition to the email notification.
14. (Optional) In the **Delay** area, select the desired amount of time to wait before generating a notification. The default is zero minutes.

    Repeat steps 9 through 14 to specify an email notification template for each notification rule you wish to create for the quota.
15. Click **Save**.

Managing quota reports

You can configure and schedule reports to help you monitor, track, and analyze storage use on an Isilon cluster.

You can view and schedule reports and customize report settings to track, monitor, and analyze disk storage use. Quota reports are managed by configuring settings that give you control over when reports are scheduled, how they are generated, where and how many are stored, and how they are viewed. The maximum number of scheduled reports that are available for viewing in the web-administration interface can be configured for each report type. When the maximum number of reports are stored, the system deletes the oldest reports to make space for new reports as they are generated.

Create a quota report schedule

You can configure quota report settings to generate the quota report on a specified schedule.

These settings determine whether and when scheduled reports are generated, and where and how the reports are stored. If you disable a scheduled report, you can still run unscheduled reports at any time.

Procedure

1. Click **File System Management** › **SmartQuotas** › **Settings**.
2. (Optional) On the Quota settings page, for **Scheduled Reporting**, click **On**.

   The **Report Frequency** option appears.
3. Click **Change schedule**, and select the report frequency that you want to set from the list.
4. Select the reporting schedule options that you want.
5. Click **Save**.

Results

Reports are generated according to your criteria and can be viewed in the Generated Reports Archive.

Generate a quota report

In addition to scheduled quota reports, you can generate a report to capture usage statistics at a point in time.

Before you begin

Quotas must exist and the initial QuotaScan job must run before you can generate a quota report.

Procedure

1. Click **File System Management › SmartQuotas › Generated Reports Archive**.
2. In the **Generated Quota Reports Archive** area, click **Generate a quota report**.
3. Click **Generate Report**.

Results

The new report appears in the Quota Reports list.

Locate a quota report

You can locate quota reports, which are stored as XML files, and use your own tools and transforms to view them. This task can only be performed from the OneFS command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Navigate to the directory where quota reports are stored. The following path is the default quota report location:

   `/ifs/.isilon/smartquotas/reports`

   Note

   If quota reports are not in the default directory, you can run the `isi quota settings` command to find the directory where they are stored.

3. At the command prompt, run the `ls` command.
   - To view a list of all quota reports in the directory, run the following command:

     ```
     ls -a *.xml
     ```
   - To view a specific quota report in the directory, run the following command:

     ```
     ls <filename>.xml
     ```

Basic quota settings

When you create a storage quota, the following attributes must be defined, at a minimum. When you specify usage limits, additional options are available for defining your quota.

Option	Description
Directory Path	The directory that the quota is on.

Option	Description
User Quota	Select to automatically create a quota for every current or future user that stores data in the specified directory.
Group Quota	Select to automatically create a quota for every current or future group that stores data in the specified directory.
Include Snapshot Data	Select to count all snapshot data in usage limits; cannot be changed after the quota is created.
Include Data-Protection Overhead	Select to count protection overhead in usage limits.
No Usage Limit	Select to account for usage only.
Specify Usage Limits	Select to enforce advisory, soft, or absolute limits.

Advisory limit quota notification rules settings

You can configure custom quota notification rules for advisory limits for a quota. These settings are available when you select the option to use custom notification rules.

Option	Description	Exceeded	Remains exceeded
Send email	Specify the type of email to use.	Yes	Yes
Notify owner	Select to send an email notification to the owner of the entity.	Yes	Yes
Notify another	Select to send an email notification to another recipient and type the recipient's email address.	Yes	Yes
Message template	Select from the following template types for use in formatting email notifications: • Default (leave **Message Template** field blank to use default) • Custom	Yes	Yes
Create cluster event	Select to generate an event notification for the quota when exceeded.	Yes	Yes
Delay	Specify the length of time (hours, days, weeks) to delay before generating a notification.	Yes	No
Frequency	Specify the notification and alert frequency: daily, weekly, monthly, yearly; depending on selection, specify intervals, day to send, time of day, multiple emails per rule.	No	Yes

Soft limit quota notification rules settings

You can configure custom soft limit notification rules for a quota. These settings are available when you select the option to use custom notification rules.

Option	Description	Exceeded	Remains exceeded	Grace period expired	Write access denied
Send email	Specify the recipient of the email notification.	Yes	Yes	Yes	Yes
Notify owner	Select to send an email notification to the owner of the entity.	Yes	Yes	Yes	Yes
Notify another	Select to send an email notification to another recipient and type the recipient's email address.	Yes	Yes	Yes	Yes
Message template	Select from the following template types for use in formatting email notifications: • Default (leave **Message Template** field blank to use default) • Custom	Yes	Yes	Yes	Yes
Create cluster event	Select to generate an event notification for the quota.	Yes	Yes	Yes	Yes
Delay	Specify the length of time (hours, days, weeks) to delay before generating a notification.	Yes	No	No	Yes
Frequency	Specify the notification and alert frequency: daily, weekly, monthly, yearly; depending on selection, specify intervals, day to send, time of day, multiple emails per rule.	No	Yes	Yes	No

Hard limit quota notification rules settings

You can configure custom quota notification rules for hard limits for a quota. These settings are available when you select the option to use custom notification rules.

Option	Description	Write access denied	Exceeded
Send email	Specify the recipient of the email notification.	Yes	Yes
Notify owner	Select to send an email notification to the owner of the entity.	Yes	Yes
Notify another	Select to send an email notification to another recipient and type the recipient's email address.	Yes	Yes
Message template	Select from the following template types for use in formatting email notifications: • Default (leave **Message Template** field blank to use default) • Custom	Yes	Yes

Option	Description	Write access denied	Exceeded
Create cluster event	Select to generate an event notification for the quota when exceeded.	Yes	Yes
Delay	Specify the length of time (hours, days, weeks) to delay before generating a notification.	Yes	No
Frequency	Specify the notification and alert frequency: daily, weekly, monthly, yearly; depending on selection, specify intervals, day to send, time of day, multiple emails per rule.	No	Yes

Limit notification settings

You have three notification options when you create an enforcement quota: use default notification rules, turn off notifications, or use custom notification rules. Enforcement quotas support the following notification settings for each threshold type. A quota can use only one of these settings.

Notification setting	Description
Use Default Notification Rules	Uses the default notification rules that you configured for the specified threshold type.
Turn Off Notifications for this Quota	Disables all notifications for the quota.
Use Custom Notification Rules	Provides settings to create basic custom notifications that apply to only this quota.

Quota report settings

You can configure quota report settings that track disk usage. These settings determine whether and when scheduled reports are generated, and where and how reports are stored. When the maximum number of reports are stored, the system deletes the oldest reports to make space for new reports as they are generated.

Setting	Description
Scheduled reporting	Enables or disables the scheduled reporting feature. • **Off**. Manually generated on-demand reports can be run at any time. • **On**. Reports run automatically according to the schedule that you specify.
Report frequency	Specifies the interval for this report to run: daily, weekly, monthly, or yearly. You can use the following options to further refine the report schedule. **Generate report every**. Specify the numeric value for the selected report frequency; for example, every 2 months. **Generate reports on**. Select the day or multiple days to generate reports. **Select report day by**. Specify date or day of the week to generate the report. **Generate one report per specified by**. Set the time of day to generate this report.

Setting	Description
	Generate multiple reports per specified day. Set the intervals and times of day to generate the report for that day.
Scheduled report archiving	Determines the maximum number of scheduled reports that are available for viewing on the SmartQuotas **Reports** page. **Limit archive size** for scheduled reports to a specified number of reports. Type the integer to specify the maximum number of reports to keep. **Archive Directory**. Browse to the directory where you want to store quota reports for archiving.
Manual report archiving	Determines the maximum number of manually generated (on-demand) reports that are available for viewing on the SmartQuotas **Reports** page. **Limit archive size** for live reports to a specified number of reports. Type the integer to specify the maximum number of reports to keep. **Archive Directory**. Browse to the directory where you want to store quota reports for archiving.

CHAPTER 19

Storage Pools

This section contains the following topics:

Storage pools overview

OneFS organizes different node types into separate node pools, and you can further organize these node pools into logical tiers of storage. By activating a SmartPools license, you can create file pool policies that store files in these tiers automatically, based on file matching criteria that you specify.

Without an active SmartPools license, OneFS manages all node pools as a single pool of storage and stores file data and metadata across the entire cluster, ensuring that data is protected, secure, and readily accessible. All files belong to the default file pool and are governed by the default file pool policy. In this mode, OneFS provides a number of functions that you can deploy, such as autoprovisioning, compatibilities, virtual hot spare, SSD strategies, global namespace acceleration (GNA), L3 cache, and storage tiers.

When you activate a SmartPools license, you can take further control of your data set to optimize the performance of your cluster. Additional functions become available, including custom file pool policies that take precedence over the default policy, and spillover management.

The following table summarizes storage pool functions based on whether a SmartPools license is inactive or active.

Function	Without active SmartPools license	With active SmartPools license
Automatic storage pool provisioning	Yes	Yes
Compatibilities (node equivalence)	Yes	Yes
Virtual hot spare	Yes	Yes
SSD strategies	Yes	Yes
L3 cache	Yes	Yes
Tiers	Yes	Yes
GNA	Yes	Yes
File pool policies	No	Yes
Spillover management	No	Yes

Storage pool functions

OneFS automatically groups equivalence-class nodes into node pools when the system is installed and whenever nodes are added to the cluster. This autoprovisioning of nodes into node pools enables OneFS to optimize reliability and data protection on the cluster.

Without an active SmartPools license, OneFS uses a default file pool policy to organize all data into a single file pool. In this mode, OneFS distributes data across the entire cluster so that data is protected and readily accessible.

OneFS includes the following functions with or without an active SmartPools license:

Autoprovisioning of node pools

OneFS automatically groups equivalence-class nodes into node pools for optimal storage efficiency and protection.

Compatibilities (node equivalence)

Enable certain nodes that are not equivalence-class to join existing node pools. OneFS supports compatibilities between Isilon S200 and S210 nodes, and between Isilon X400 and X410 nodes.

Tiers

Group node pools into logical tiers of storage. It is recommended that you activate a SmartPools license for this feature. An active SmartPools license enables you to create custom file pool policies and direct different file pools to appropriate storage tiers.

Default file pool policy

Governs all file types and can store files anywhere on the cluster. Custom file pool policies, which require an SmartPools license, take precedence over the default file pool policy.

Requested protection

Specify a requested protection setting for the default file pool, per node pool, or even on individual files. You can leave the default setting in place, or choose the suggested protection calculated by OneFS to ensure optimal data protection.

Virtual hot spare

Reserve a percentage of available storage for data repair in the event of a disk failure.

SSD strategies

Define the type of data that is stored on SSDs in the cluster, for example, storing metadata for read/write acceleration.

L3 cache

Specify that SSDs in nodes exclusively cache data and metadata, thus enabling faster access to the most frequently requested files.

Global namespace acceleration

Activate global namespace acceleration (GNA), which allows data stored on node pools without SSDs to access SSDs elsewhere in the cluster to store extra metadata mirrors. Extra metadata mirrors accelerate metadata read operations.

When you activate a SmartPools license, OneFS provides access to the following additional functions:

Custom file pool policies

Create custom file pool policies to identify different classes of files, and store these file pools in logical storage tiers. For example, you could define a high-performance tier of Isilon S-series node pools, and an archival tier of high-capacity Isilon NL400 and HD400 node pools. Then, with custom file pool policies, you could identify file pools based on matching criteria, and define actions to perform on these pools. For example, one file pool policy could identify all JPEG files older than a year and store them in an archival tier. Another policy could move all files that have been created or modified within the last three months to a performance tier.

Storage pool spillover

Enable automated capacity overflow management for storage pools. Spillover defines how to handle write operations when a storage pool is not writable for some reason. If spillover is enabled, data is redirected to a specified storage pool. If spillover is disabled, new data writes fail and an error message is sent to the client attempting the write operation.

Autoprovisioning

When you add a node to your cluster, OneFS automatically assigns the node to a node pool. With node pools, OneFS can ensure optimal performance, load balancing, and reliability of the file system. Autoprovisioning reduces the time required for the manual management tasks associated with resource planning and configuration.

Nodes are not provisioned, meaning they are not associated with each other and are not writable, until at least three nodes of an equivalence class are added to the cluster. If you have added only two nodes of an equivalence class to your cluster, no data is stored on those nodes until you add a third node of the same equivalence class.

Similarly, if a node goes down or is removed from the cluster so that fewer than three equivalence-class nodes remain, the node pool becomes under-provisioned. However, the two remaining nodes are still writable. If only one node remains, that node is not writable, but remains readable.

OneFS offers a compatibility function, also referred to as node equivalency, that enables certain node types to be provisioned to existing node pools, even when there are fewer than three equivalence-class nodes. For example, an S210 node could be provisioned and added to a node pool of S200 nodes. Similarly, an X410 node could be added to a node pool of X400 nodes. The compatibility function ensures that you can add nodes one at a time to your cluster and still have them be fully functional peers within a node pool.

Node pools

A node pool is a collection of three or more nodes. As you add nodes to an Isilon cluster, OneFS automatically provisions them into node pools based on characteristics such as series, drive size, RAM, and SSD-per-node ratio. Nodes with identical characteristics are called equivalence-class nodes.

If you add fewer than three nodes of a node type, OneFS cannot autoprovision the nodes to your cluster. In these cases, you can create compatibilities. Compatibilities enable OneFS to provision nodes that are not equivalence-class to a compatible node pool.

After provisioning, each node in the OneFS cluster is a peer, and any node can handle a data request. Each provisioned node increases the aggregate disk, cache, CPU, and network capacity on the cluster.

You can move nodes from an automatically managed node pool into one that you define manually. This capability is available only through the OneFS command-line interface. If you attempt to remove nodes from a node pool such that the removal would leave fewer than three nodes in the pool, the removal fails. When you remove a node from a manually defined node pool, OneFS attempts to move the node into a node pool of the same equivalence class, or into a compatible node pool.

Node compatibilities

OneFS requires that a node pool contain at least three nodes so that the operating system can write data and perform the necessary load balancing and data protection operations. You can enable certain nodes to be provisioned to an existing node pool by defining a compatibility.

Note

The compatibility function is also referred to as node equivalency.

If you add fewer than three Isilon S210 or X410 nodes to your cluster, and you have existing S200 or X400 node pools, you can create compatibilities to provision the new nodes and make them functional within the cluster. Only S210 and X410 nodes are eligible for compatibility.

To be provisioned, the S210 or X410 nodes must have the same drive configurations as their S200 and X400 counterparts, and must have compatible RAM amounts, as shown in the following table:

S200/S210 Compatibility		X400/X410 Compatibility	
S200 RAM	**S210 RAM**	**X400 RAM**	**X410 RAM**
24 GB	32 GB	24 GB	32 GB
48 GB	64 GB	48 GB	64 GB
96 GB	128 GB	96 GB	128 GB
	256 GB	192 GB	256 GB

Note

After you have added three or more S210 or X410 nodes to your cluster, you should consider removing the compatibilities that you have created. This step enables OneFS to autoprovision new S210 or X410 node pools and take advantage of the performance specifications of the newer node types.

Manual node pools

If the node pools automatically provisioned by OneFS do not meet your needs, you can configure node pools manually. You do this by moving nodes from an existing node pool into the manual node pool.

This capability enables you to store data on specific nodes according to your purposes, and is available only through the OneFS command-line interface.

⚠ CAUTION

It is recommended that you enable OneFS to provision nodes automatically. Manually created node pools might not provide the same performance and efficiency as automatically managed node pools, particularly if your changes result in fewer than 20 nodes in the manual node pool.

Virtual hot spare

Virtual hot spare (VHS) settings enable you to reserve disk space to rebuild the data in the event that a drive fails.

You can specify both a number of virtual drives to reserve and a percentage of total storage space. For example, if you specify two virtual drives and 15 percent, each node pool reserves virtual drive space equivalent to two drives or 15 percent of their total capacity (whichever is larger).

You can reserve space in node pools across the cluster for this purpose by specifying the following options:

- At least 1–4 virtual drives.

- At least 0–20% of total storage.

OneFS calculates the larger number of the two factors to determine the space that is allocated. When configuring VHS settings, be sure to consider the following information:

- If you deselect the option to **Ignore reserved space when calculating available free space** (the default), free-space calculations include the space reserved for VHS.
- If you deselect the option to **Deny data writes to reserved disk space** (the default), OneFS can use VHS for normal data writes. We recommend that you leave this option selected, or data repair can be compromised.
- If **Ignore reserved space when calculating available free space** is enabled while **Deny data writes to reserved disk space** is disabled, it is possible for the file system to report utilization as more than 100 percent.

Note

VHS settings affect spillover. If the VHS option **Deny data writes to reserved disk space** is enabled while **Ignore reserved space when calculating available free space** is disabled, spillover occurs before the file system reports 100% utilization.

Spillover

When you activate a SmartPools license, you can designate a node pool or tier to receive spillover data when the hardware specified by a file pool policy is full or otherwise not writable.

If you do not want data to spill over to a different location because the specified node pool or tier is full or not writable, you can disable this feature.

Note

Virtual hot spare reservations affect spillover. If the setting **Deny data writes to reserved disk space** is enabled, while **Ignore reserved space when calculating available free space** is disabled, spillover occurs before the file system reports 100% utilization.

Suggested protection

Based on the configuration of your Isilon cluster, OneFS automatically calculates the amount of protection that is recommended to maintain EMC Isilon's stringent data protection requirements.

OneFS includes a function to calculate the suggested protection for data to maintain a theoretical mean-time to data loss (MTTDL) of 5000 years. Suggested protection provides the optimal balance between data protection and storage efficiency on your cluster.

By configuring file pool policies, you can specify one of multiple requested protection settings for a single file, for subsets of files called file pools, or for all files on the cluster.

It is recommended that you do not specify a setting below suggested protection. OneFS periodically checks the protection level on the cluster, and alerts you if data falls below the recommended protection.

Protection policies

OneFS provides a number of protection policies to choose from when protecting a file or specifying a file pool policy.

The more nodes you have in your cluster, up to 20 nodes, the more efficiently OneFS can store and protect data, and the higher levels of requested protection the operating system can achieve. Depending on the configuration of your cluster and how much data is stored, OneFS might not be able to achieve the level of protection that you request. For example, if you have a three-node cluster that is approaching capacity, and you request +2n protection, OneFS might not be able to deliver the requested protection.

The following table describes the available protection policies in OneFS.

Protection policy	Summary
+1n	Tolerate the failure of 1 drive or the failure of 1 node
+2d:1n	Tolerate the failure of 2 drives or the failure of 1 node
+2n	Tolerate the failure of 2 drives or the failure of 2 nodes
+3d:1n	Tolerate the failure of 3 drives or the failure of 1 node
+3d:1n1d	Tolerate the failure of 3 drives or the failure of 1 node and 1 drive
+3n	Tolerate the failure of 3 drives or the failure of 3 nodes
+4d:1n	Tolerate the failure of 4 drives or the failure of 1 node
+4d:2n	Tolerate the failure of 4 drives or the failure of 2 nodes
+4n	Tolerate the failure of 4 drives or the failure of 4 nodes
Mirrors: 2x 3x 4x 5x 6x 7x 8x	Duplicates, or mirrors, data over the specified number of nodes. For example, 2x results in two copies of each data block. **Note** Mirrors can use more data than the other protection policies, but might be an effective way to protect files that are written non-sequentially or to provide faster access to important files.

SSD strategies

OneFS clusters can contain nodes that include solid-state drives (SSD). OneFS autoprovisions equivalence-class nodes with SSDs into one or more node pools. The SSD strategy defined in the default file pool policy determines how SSDs are used within the cluster, and can be set to increase performance across a wide range of workflows.

You can configure file pool policies to apply specific SSD strategies as needed. When you select SSD options during the creation of a file pool policy, you can identify the files in the OneFS cluster that require faster or slower performance. When the SmartPools job runs, OneFS uses file pool policies to move this data to the appropriate storage pool and drive type.

The following SSD strategy options that you can set in a file pool policy are listed in order of slowest to fastest choices:

Avoid SSDs
: Writes all associated file data and metadata to HDDs only.

⚠ CAUTION

Use this option to free SSD space only after consulting with Isilon Technical Support personnel. Using this strategy can negatively affect performance.

Metadata read acceleration
: Writes both file data and metadata to HDDs. This is the default setting. An extra mirror of the file metadata is written to SSDs, if available. The SSD mirror is in addition to the number of mirrors, if any, required to satisfy the requested protection.

Metadata read/write acceleration
: Writes file data to HDDs and metadata to SSDs, when available. This strategy accelerates metadata writes in addition to reads but requires about four to five times more SSD storage than the **Metadata read acceleration** setting. Enabling GNA does not affect read/write acceleration.

Data on SSDs
: Uses SSD node pools for both data and metadata, regardless of whether global namespace acceleration is enabled. This SSD strategy does not result in the creation of additional mirrors beyond the normal requested protection but requires significantly increased storage requirements compared with the other SSD strategy options.

Global namespace acceleration

Global namespace acceleration (GNA) allows data stored on node pools without SSDs to use SSDs elsewhere in the cluster to store extra metadata mirrors. Extra metadata mirrors can improve file system performance by accelerating metadata read operations.

You can only enable GNA if 20% or more of the nodes in the cluster contain at least one SSD and 1.5% or more of the total cluster storage is SSD-based. For best results, ensure that at least 2.0% of the total cluster storage is SSD-based before enabling global namespace acceleration.

If the ratio of SSDs to non-SSDs on the cluster falls below the 1.5% threshold, GNA becomes inactive even if enabled. GNA is reactivated when the ratio is corrected. When GNA is inactive, existing SSD mirrors are readable but newly written metadata does not include the extra SSD mirror.

Note

If GNA is enabled for the cluster, file pool policies that direct data to node pools with L3 cache enabled should also set the SSD strategy to `Avoid SSDs`. Otherwise, additional SSD mirrors would be created for data that is already accelerated by L3 cache. This is an inefficient use of SSD storage space and is not recommended.

L3 cache overview

You can configure nodes with solid-state drives (SSDs) to increase cache memory and speed up file system performance across larger working file sets.

OneFS caches file data and metadata at multiple levels. The following table describes the types of file system cache available on an Isilon cluster.

Name	Type	Profile	Scope	Description
L1 cache	RAM	Volatile	Local node	Also known as front-end cache, holds copies of file system metadata and data requested by the front-end network through NFS, SMB, HTTP, and so on.
L2 cache	RAM	Volatile	Global	Also known as back-end cache, holds copies of file system metadata and data on the node that owns the data.
SmartCache	Variable	Non-volatile	Local node	Holds any pending changes to front-end files waiting to be written to storage. This type of cache protects write-back data through a combination of RAM and stable storage.
L3 cache	SSD	Non-volatile	Global	Holds file data and metadata released from L2 cache, effectively increasing L2 cache capacity.

OneFS caches frequently accessed file and metadata in available random access memory (RAM). Caching enables OneFS to optimize data protection and file system performance. When RAM cache reaches capacity, OneFS normally discards the oldest cached data and processes new data requests by accessing the storage drives. This cycle is repeated each time RAM cache fills up.

You can deploy SSDs as L3 cache to reduce the cache cycling issue and further improve file system performance. L3 cache adds significantly to the available cache memory and provides faster access to data than hard disk drives (HDD).

As L2 cache reaches capacity, OneFS evaluates data to be released and, depending on your workflow, moves the data to L3 cache. In this way, much more of the most frequently accessed data is held in cache, and overall file system performance is improved.

For example, consider a cluster with 128GB of RAM. Typically the amount of RAM available for cache fluctuates, depending on other active processes. If 50 percent of RAM is available for cache, the cache size would be approximately 64GB. If this same cluster had three nodes, each with two 200GB SSDs, the amount of L3 cache would be 1.2TB, approximately 18 times the amount of available L2 cache.

L3 cache is enabled by default for new node pools. A node pool is a collection of nodes that are all of the same equivalence class, or for which compatibilities have been created. L3 cache applies only to the nodes where the SSDs reside. For the HD400 node, which is primarily for archival purposes, L3 cache is on by default and cannot be turned off. On the HD400, L3 cache is used only for metadata.

If you enable L3 cache on a node pool, OneFS manages all cache levels to provide optimal data protection, availability, and performance. In addition, in case of a power failure, the data on L3 cache is retained and still available after power is restored.

Note

Although some benefit from L3 cache is found in workflows with streaming and concurrent file access, L3 cache provides the most benefit in workflows that involve random file access.

Migration to L3 cache

L3 cache is enabled by default on new nodes. If you are upgrading your cluster from an older release (pre-OneFS 7.1.1), you must enable L3 cache manually on node pools that have SSDs. When you enable L3 cache, OneFS activates a process that migrates SSDs from storage disks to cache. File data currently on SSDs is moved elsewhere in the cluster.

You can enable L3 cache as the default for all new node pools or manually for a specific node pool, either through the command line or from the web administration interface. You can enable L3 cache only on node pools whose nodes have SSDs.

Depending on the amount of data stored in your SSDs, the migration process can take some time. OneFS displays a message informing you that the migration is about to begin:

```
WARNING: Changes to L3 cache configuration can have a long completion
time. If this is a concern, please contact EMC Isilon Support for
more information.
```

You must confirm whether OneFS should proceed with the migration. After you do, OneFS handles the migration intelligently as a background process. You can continue to administer your cluster during the migration.

If you choose to disable L3 cache on a node pool, the migration process is very fast.

L3 cache on HD400 node pools

The HD400 is a high-capacity node designed primarily for archival workflows. L3 cache is turned on by default on HD400 node pools and cannot be turned off.

Archival workflows feature a higher percentage of data writes compared to data reads. Consequently, L3 cache on HD400 node pools holds only metadata, which improves the speed of file system traversal activities such as directory lookup. L3 cache on HD400 nodes does not contain file data.

Tiers

A tier is a user-defined collection of node pools that you can specify as a storage pool for files. A node pool can belong to only one tier.

You can create tiers to assign your data to any of the node pools in the tier. For example, you can assign a collection of node pools to a tier specifically created to store data that requires high availability and fast access. In a three-tier system, this classification may be Tier 1. You can classify data that is used less frequently or that is accessed by fewer users as Tier-2 data. Tier 3 usually comprises data that is seldom used and can be archived for historical or regulatory purposes.

File pool policies

File pool policies define sets of files—file pools—and where and how they are stored on your cluster. You can configure multiple file pool policies with filtering rules that identify specific file pools and the requested protection and I/O optimization settings for these file pools. Creating custom file pool policies requires an active SmartPools license.

The initial installation of OneFS places all files into a single file pool, which is subject to the default file pool policy. Without an active SmartPools license, you can configure only the default file pool policy, which controls all files and stores them anywhere on the cluster.

With an active SmartPools license, OneFS augments basic storage functions by enabling you to create custom file pool policies that identify, protect, and control multiple file pools. With a custom file pool policy, for example, you can define and store a file pool on a specific node pool or tier for fast access or archival purposes.

When you create a file pool policy, flexible filtering criteria enable you to specify time-based attributes for the dates that files were last accessed, modified, or created. You can also define relative time attributes, such as 30 days before the current date. Other filtering criteria include file type, name, size, and custom attributes. The following examples demonstrate a few ways you can configure file pool policies:

- A file pool policy to set stronger protection on a specific set of important files.
- A file pool policy to store frequently accessed files in a node pool that provides the fastest reads or read/writes.
- A file pool policy to evaluate the last time files were accessed, so that older files are stored in a node pool best suited for regulatory archival purposes.

When the SmartPools job runs, typically once a day, it processes file pool policies in priority order. You can edit, reorder, or remove custom file pool policies at any time. The default file pool policy, however, is always last in priority order. Although you can edit the default file pool policy, you cannot reorder or remove it. When custom file pool policies are in place, the settings in the default file pool policy apply only to files that are not covered by another file pool policy.

When new files are created, OneFS temporarily chooses a storage pool based on file pool policies in place when the last SmartPools job ran. OneFS might apply new storage settings and move these new files when the next SmartPools job runs, based on a new matching file pool policy.

Managing node pools in the web administration interface

You can manage node pools through the OneFS web administration interface. You must have the SmartPools or higher administrative privilege.

Add node pools to a tier

You can group available node pools into tiers.

A node pool can only be added to one tier at a time. If no node pools are listed as available, they already belong to other tiers.

Procedure

1. Click **File System › Storage Pools › SmartPools**.

   The **SmartPools** page displays two groups: **Tiers & Node Pools** and **Compatibilities**
2. In the **Tiers & Node Pools** area, click **View/Edit** next to the tier.
3. In the **View Tier Details** page, click **Edit Tier**.

   The **Edit Tier Details** page is displayed.

4. In the **Available Node Pools** list, select a node pool and click **Add**.

   The node pool moves to the **Selected Node Pools for this Tier** list.

5. Repeat step 4 for each node pool you intend to add. When all node pools have been added, click **Save Changes**.

   A message informs you that the operation was successful. The **View Tier Details** page remains open.

6. Click **Close**.

   The **Tiers & Node Pools** group now shows that the node pools are part of the tier.

Change the name or requested protection of a node pool

You can change the name or the requested protection of a node pool.

Procedure

1. Click **File System › Storage Pools › SmartPools**.
2. In the **Tiers & Node Pools** group, in the row of the node pool that you want to modify, click **View/Edit**.

   The **View Node Pools Details** page appears.

3. Click **Edit.**

   The **Edit Node Pools Details** page appears.

4. Enter a new name for the node pool, or specify a new requested protection level from the list, or do both.

   A node pool name can start only with a letter or underscore character, and otherwise can contain only letters, numbers, hyphens, underscores, or periods.

5. Click **Save Changes** in the **Edit Node Pools Details** page.
6. Click **Close** in the **View Node Pools Details** page.

Add a compatible node to a node pool

OneFS automatically adds a new equivalence-class node to an existing node pool. For a new node that is not equivalence-class, you can create a compatibility to add the node to an existing node pool.

Compatibilities work only between S200 and S210 nodes, and X400 and X410 nodes. For example, if you have a node pool made up of three or more S200 nodes, you can create a compatibility so that new S210 nodes are automatically added to the S200 node pool. Similarly, if you have a node pool made up of three or more X400 nodes, you can create a compatibility so that new X410 nodes are automatically added to the X400 node pool.

Note

The S210 and X410 nodes must have compatible RAM and the same drive configurations as their S200 and X400 counterparts to be provisioned into those node pools.

Procedure

1. Click **File System › Storage Pools › SmartPools**.

   The **SmartPools** tab displays two lists: **Tiers & Node Pools** and **Compatibilities.**

2. Click **Create a Compatibility**.

The **Create a Compatibility** dialog box displays drop-down lists of compatibility options.

3. In the **Make** list, select either `S200` or `X400`, as appropriate.
4. In the **compatible with** list, select either `S210` or `X410`, as appropriate.
5. Click **Create a Compatibility**.

   A **Confirm Changes to this Cluster** dialog box appears with two check boxes that you must select before proceeding. The check boxes describe the result of the operation.
6. Select all check boxes, then click **Confirm**.

Results

The compatibility is created and described in the **Compatibilities** list. The result of the compatibility appears in the **Tiers & Node Pools** list.

Merge compatible node pools

You can merge multiple compatible node pools to optimize efficiency across the cluster.

For example, if you had six S200 nodes in one node pool and three S210 nodes in a second node pool, you could create a compatibility to merge the two node pools into one nine-node pool. Larger node pools, up to 20 nodes, enable OneFS to protect data more efficiently, thus preserving more storage space for new data.

Note

Newer node types typically have better performance specifications than older node types, so merging them with older node types could reduce performance. Also, when two node pools are merged, OneFS restripes the data to take advantage of the larger node pool, which could take considerable time, depending on the size of your data set.

Procedure

1. Click **File System** › **Storage Pools** › **SmartPools**.

   The **SmartPools** tab displays two lists: **Tiers & Node Pools** and **Compatibilities**.
2. Click **Create a Compatibility**.

   The **Create a Compatibility** dialog box displays drop-down lists of compatibility options.
3. In the **Make** list, select either `S200` or `X400`, as appropriate.
4. In the **compatible with** list, select either `S210` or `X410`, as appropriate.
5. Click **Create a Compatibility**.

   A **Confirm Changes to this Cluster** dialog box appears with a two check boxes that you have to select before proceeding. The check boxes describe the result of the operation.
6. Click **Confirm**.

   If the two node pools being merged have different settings, for example L3 cache, OneFS prompts you to resolve the differences before merging the node pools.

Results

The compatibility appears in the **Compatibilities** list. In the **Tiers & Node Pools** list, the two former node pools are joined as one.

Delete a compatibility

You can delete a compatibility, and any nodes that are part of a node pool because of this compatibility are removed from the node pool.

⚠ CAUTION

Deleting a compatibility could result in unintended consequences. For example, if you delete a compatibility, and fewer than three compatible nodes are removed from the node pool, those nodes would be removed from your cluster's available pool of storage. The next time the SmartPools job runs, data on those nodes would be restriped elsewhere on the cluster, which could be a lengthy process. If three or more compatible nodes are removed from the node pool, these nodes would form their own node pool, but data could still be restriped. Any file pool policy pointing to the original node pool would now point to the node pool's tier if one existed, or otherwise to a new tier created by OneFS.

Procedure

1. Click **File System** › **Storage Pools** › **SmartPools**.

   The **SmartPools** tab displays two groups: **Tiers & Node Pools** and **Compatibilities**.

2. In the **Compatibilities** list, next to the compatibility that you want to delete, click **Delete**.

   The **Delete a Compatibility** dialog box appears, describing the compatibility and what will happen when you delete it.

3. Click **Delete Compatibility**.

   A **Confirm Changes to this Cluster** dialog box appears with a list of four check boxes that you must select before proceeding.

4. Select all check boxes, and click **Confirm**.

Results

The compatibility is deleted, and the new state of the affected nodes appears in the **Tiers & Node Pools** list.

Managing L3 cache from the web administration interface

You can manage L3 cache globally or on specific node pools from the web administration interface. You must have the SmartPools or higher administrative privilege. On HD400 nodes, L3 cache is turned on by default and cannot be turned off.

Set L3 cache as the default for node pools

You can set L3 cache as the default, so that when new node pools are created, L3 cache is enabled automatically.

Before you begin

L3 cache is only effective on nodes that include SSDs. If none of your clusters have SSD storage, there is no need to enable L3 cache as the default.

Procedure

1. Click **File System** › **Storage Pools** › **SmartPools Settings**.

   The **Edit SmartPools Settings** page appears.

2. Under **Local Storage Settings**, click **Use SSDs as L3 Cache by default for new node pools**.
3. Click **Save Changes**.

Results

As you add new nodes with SSDs to your cluster, and OneFS designates new node pools, these node pools automatically have L3 cache enabled. New node pools without SSDs do not have L3 cache enabled by default.

Set L3 cache on a specific node pool

You can turn on L3 cache for a specific node pool.

Procedure

1. Click **File System** › **Storage Pools** › **SmartPools**.

   The **SmartPools** page, showing a list of tiers and node pools, appears.
2. In the **Tiers & Node Pools** list, click **View/Edit** next to the target node pool.

   The **View Node Pool Details** dialog box appears, showing the current settings of the node pool.
3. Click **Edit**.

   The **Edit Node Pool Details** dialog box appears.
4. Click the **Enable L3 cache** check box.

   The check box is grayed out for node pools that do not have SSDs, or for which the setting cannot be changed.
5. Click **Save Changes**.

   The **Confirm Change to L3 Cache Setting** message box appears.
6. Click the **Continue** button.

   The migration process to L3 cache begins and can take awhile, depending on the number and size of the SSDs in the node pool. When the migration process is complete, the **View Node Pool Details** dialog box appears.
7. Click **Close**.

Restore SSDs to storage drives for a node pool

You can disable L3 cache for SSDs on a node pool and restore those SSDs to storage drives.

Note

On HD400 node pools, SSDs are only used for L3 cache, which is turned on by default and cannot be turned off. All other node pools with SSDs for L3 cache can have their SSDs migrated back to storage drives.

Procedure

1. Click **File System** › **Storage Pools** › **SmartPools**.
2. In the **Tiers & Node Pools** area of the **SmartPools** tab, select **View/Edit** next to the target node pool.

   The **View Node Pool Details** dialog box appears, showing the current settings of the node pool.

3. Click **Edit**.

   The **Edit Node Pool Details** dialog box appears.

4. Clear the **Enable L3 cache** check box.

   The setting is grayed out for node pools without SSDs, or for which the setting cannot be changed.

5. Click **Save Changes**.

   The **Confirm Change to L3 Cache Setting** message box appears.

6. Click **Continue**.

   The migration process to disable L3 cache begins and can take awhile, depending on the number and size of the SSDs in the node pool. When the migration process is complete, the **View Node Pool Details** dialog box appears.

7. Click **Close**.

Managing tiers

You can move node pools into tiers to optimize file and storage management. Managing tiers requires the SmartPools or higher administrative privilege.

Create a tier

You can group create a tier that contains one or more node pools. You can use the tier to store specific categories of files.

Procedure

1. Click **File System** › **Storage Pools** › **SmartPools**.

   The **SmartPools** tab appears with two sections: **Tiers & Node Pools** and **Compatibilities**.

2. In the **Tiers & Node Pools** section, click **Create a Tier**.
3. In the **Create a Tier** page that appears, enter a name for the tier.
4. For each node pool that you want to add to the tier, select a node pool from the **Available Node Pools** list, and click **Add**.

   The node pool is moved into the **Selected Node Pools for this Tier** list.

5. Click **Create Tier**.

   The **Create a Tier** page closes, and the new tier is added to the **Tiers & Node Pools** area. The node pools that you added are shown below the tier name.

Edit a tier

You can modify the name and change the node pools that are assigned to a tier.

A tier name can contain alphanumeric characters and underscores but cannot begin with a number.

Procedure

1. Click **File System** › **Storage Pools** › **SmartPools**.

   The **SmartPools** tab displays two groups: **Tiers & Node Pools** and **Compatibilities**.

2. In the **Tiers & Node Pools** area, next to the tier you want to edit, click **View/Edit**.

3. In the **View Tier Details** dialog box, click **Edit Tier**.
4. In the **Edit Tier Details** dialog box, modify the following settings as needed:

Option	Description
`Tier Name`	To change the name of the tier, select and type over the existing name.
`Node Pool Selection`	To change the node pool selection, select a node pool, and click either **Add** or **Remove.**

5. When you have finished editing tier settings, click **Save Changes**.
6. In the **View Tier Details** dialog box, click **Close**.

Delete a tier

You can delete a tier that has no assigned node pools.

Before you begin

If you want to delete a tier that does have assigned node pools, you must first remove the node pools from the tier.

Procedure

1. Click **File System** › **Storage Pools** › **SmartPools**.

   The **SmartPools** tab displays two lists: **Tiers & Node Pools** and **Compatibilities**.
2. In the **Tiers & Node Pools** list, next to the tier that you want to delete, click **More** › **Delete Tier**.

   A message box asks you to confirm or cancel the operation.
3. Click **Delete Tier** to confirm the operation.

Results

The tier is removed from the **Tiers & Node Pools** list.

Creating file pool policies

You can configure file pool policies to identify logical groups of files called file pools, and you can specify storage operations for these files.

Before you can create file pool policies, you must activate a SmartPools license, and you must have the SmartPools or higher administrative privilege.

File pool policies have two parts: file-matching criteria that define a file pool, and the actions to be applied to the file pool. You can define file pools based on characteristics, such as file type, size, path, birth, change, and access timestamps, and combine these criteria with Boolean operators (AND, OR).

In addition to file-matching criteria, you can identify a variety of actions to apply to the file pool. These actions include:

- Setting requested protection and data-access optimization parameters
- Identifying data and snapshot storage targets
- Defining data and snapshot SSD strategies
- Enabling or disabling SmartCache

For example, to free up disk space on your performance tier (S-series node pools), you could create a file pool policy to match all files greater than 25 MB in size, which have not been accessed or modified for more than a month, and move them to your archive tier (NL-series node pools).

You can configure and prioritize multiple file pool policies to optimize file storage for your particular work flows and cluster configuration. When the SmartPools job runs, by default once a day, it applies file pool policies in priority order. When a file pool matches the criteria defined in a policy, the actions in that policy are applied, and lower-priority custom policies are ignored for the file pool.

After the list of custom file pool policies is traversed, if any of the actions are not applied to a file, the actions in the default file pool policy are applied. In this way, the default file pool policy ensures that all actions apply to every file.

Note

You can reorder the file pool policy list at any time, but the default file pool policy is always last in the list of file pool policies.

OneFS also provides customizable template policies that you can copy to make your own policies. These templates, however, are only available from the OneFS web administration interface.

Create a file pool policy

You can create a file pool policy to define a specific file set and specify SmartPools actions to be applied to the matched files. These SmartPools actions include moving files to certain tiers or node pools, changing the requested protection levels, and optimizing write performance and data access.

⚠ CAUTION

If existing file pool policies direct data to a specific storage pool, do not configure other file pool policies with `anywhere` for the Data storage target option. Because the specified storage pool is included when you use `anywhere`, target specific storage pools to avoid unexpected results.

Procedure

1. Click **File System** › **Storage Pools** › **File Pool Policies**.
2. Click **Create a File Pool Policy**.
3. In the **Create a File Pool Policy** dialog box, enter a policy name and, optionally, a description.
4. Specify the files to be managed by the file pool policy.

   To define the file pool, you can specify file matching criteria by combining IF, AND, and OR conditions. You can define these conditions with a number of file attributes, such as name, path, type, size, and timestamp information.
5. Specify SmartPools actions to be applied to the selected file pool.

   You can specify storage and I/O optimization settings to be applied.
6. Click **Create Policy**.

Results

The file pool policy is created and applied when the next scheduled SmartPools system job runs. By default, this job runs once a day, but you also have the option to start the job immediately.

File-matching options for file pool policies

You can configure a file pool policy for files that match specific criteria.

The following file-matching options can be specified when you create or edit a file pool policy.

Note

OneFS supports UNIX shell-style (glob) pattern matching for file name attributes and paths.

The following table lists the file attributes that you can use to define a file pool policy.

File attribute	Specifies
Name	Includes or excludes files based on the file name. You can specify whether to include or exclude full or partial names that contain specific text. Wildcard characters are allowed.
Path	Includes or excludes files based on the file path. You can specify whether to include or exclude full or partial paths that contain specified text. You can also include the wildcard characters *, ?, and [ ].
File type	Includes or excludes files based on one of the following file-system object types: • File • Directory • Other
Size	Includes or excludes files based on their size. **Note** File sizes are represented in multiples of 1024, not 1000.
Modified	Includes or excludes files based on when the file was last modified. In the web administration interface, you can specify a relative date and time, such as "older than 2 weeks," or a specific date and time, such as "before January 1, 2012." Time settings are based on a 24-hour clock.
Created	Includes or excludes files based on when the file was created. In the web administration interface, you can specify a relative date and time, such as "older than 2 weeks," or a specific date and time, such as "before January 1, 2012." Time settings are based on a 24-hour clock.
Metadata changed	Includes or excludes files based on when the file metadata was last modified. This option is available only if the global access-time-tracking option of the cluster is enabled.

File attribute	Specifies
	In the web administration interface, you can specify a relative date and time, such as "older than 2 weeks," or a specific date and time, such as "before January 1, 2012." Time settings are based on a 24-hour clock.
Accessed	Includes or excludes files based on when the file was last accessed based on the following units of time: In the web administration interface, you can specify a relative date and time, such as "older than 2 weeks," or a specific date and time, such as "before January 1, 2012." Time settings are based on a 24-hour clock. **Note** Because it affects performance, access time tracking as a file pool policy criterion is disabled by default.
File attribute	Includes or excludes files based on a custom user-defined attribute.

Valid wildcard characters

You can combine wildcard characters with file-matching options to define a file pool policy.

OneFS supports UNIX shell-style (glob) pattern matching for file name attributes and paths.

The following table lists the valid wildcard characters that you can combine with file-matching options to define a file pool policy.

Wildcard	Description
`*`	Matches any string in place of the asterisk. For example, `m*` matches `movies` and `m123`.
`[a-z]`	Matches any characters contained in the brackets, or a range of characters separated by a hyphen. For example, `b[aei]t` matches `bat`, `bet`, and `bit`, and `1[4-7]2` matches `142`, `152`, `162`, and `172`. You can exclude characters within brackets by following the first bracket with an exclamation mark. For example, `b[!ie]` matches `bat` but not `bit` or `bet`. You can match a bracket within a bracket if it is either the first or last character. For example, `[[c]at` matches `cat` and `[at`. You can match a hyphen within a bracket if it is either the first or last character. For example, `car[-s]` matches `cars` and `car-`.
`?`	Matches any character in place of the question mark. For example, `t?p` matches `tap`, `tip`, and `top`.

SmartPools settings

SmartPools settings include directory protection, global namespace acceleration, L3 cache, virtual hot spare, spillover, requested protection management, and I/O optimization management.

Settings in Web Admin	Settings in CLI	Description	Notes
Increase directory protection to a higher level than its contents	--protect-directories-one-level-higher	Increases the amount of protection for directories at a higher level than the directories and files that they contain, so that data that is not lost can still be accessed. When device failures result in data loss (for example, three drives or two nodes in a +2:1 policy), enabling this setting ensures that intact data is still accessible.	This setting should be enabled (the default). When this setting is disabled, the directory that contains a file pool is protected according to your protection-level settings, but the devices used to store the directory and the file may not be the same. There is potential to lose nodes with file data intact but not be able to access the data because those nodes contained the directory. As an example, consider a cluster that has a +2 default file pool protection setting and no additional file pool policies. OneFS directories are always mirrored, so they are stored at 3x, which is the mirrored equivalent of the +2 default. This configuration can sustain a failure of two nodes before data loss or inaccessibility. If this setting is enabled, all directories are protected at 4x. If the cluster experiences three node failures, although individual files may be inaccessible, the directory tree is available and provides access to files that are still accessible. In addition, if another file pool policy protects some files at a higher level, these too are accessible in the event of a three-node failure.
Enable global namespace acceleration	--global-namespace-acceleration-enabled	Specifies whether to allow per-file metadata to use SSDs in the node pool. • When disabled, restricts per-file metadata to the storage pool policy of the file, except in the case of spillover. This is the default setting.	This setting is available only if 20 percent or more of the nodes in the cluster contain SSDs and at least 1.5 percent of the total cluster storage is SSD-based. If nodes are added to or removed from a cluster, and the SSD thresholds are no longer satisfied, GNA becomes inactive. GNA remains enabled, so that when the

Settings in Web Admin	Settings in CLI	Description	Notes
		- When enabled, allows per-file metadata to use the SSDs in any node pool.	SSD thresholds are met again, GNA is reactivated. **Note** Node pools with L3 cache enabled are effectively invisible for GNA purposes. All ratio calculations for GNA are done exclusively for node pools without L3 cache enabled.
Use SSDs as L3 Cache by default for new node pools	--ssd-l3-cache-default-enabled	For node pools that include solid-state drives, deploy the SSDs as L3 cache. L3 cache extends L2 cache and speeds up file system performance across larger working file sets.	L3 cache is enabled by default on new node pools. When you enable L3 cache on an existing node pool, OneFS performs a migration, moving any existing data on the SSDs to other locations on the cluster. OneFS manages all cache levels to provide optimal data protection, availability, and performance. In case of a power failure, the data on L3 cache is retained and still available after power is restored.
Virtual Hot Spare	--virtual-hot-spare-deny-writes --virtual-hot-spare-hide-spare --virtual-hot-spare-limit-drives --virtual-hot-spare-limit-percent	Reserves a minimum amount of space in the node pool that can be used for data repair in the event of a drive failure. To reserve disk space for use as a virtual hot spare, select from the following options: - **Ignore reserved disk space when calculating available free space**. Subtracts the space reserved for the virtual hot spare when calculating available free space. - **Deny data writes to reserved disk space**. Prevents write operations from using reserved disk space. - **VHS Space Reserved**. You can reserve a minimum number of virtual drives (1-4), as well as a minimum percentage of total disk space (0-20%).	If you configure both the minimum number of virtual drives and a minimum percentage of total disk space when you configure reserved VHS space, the enforced minimum value satisfies both requirements. If this setting is enabled and **Deny new data writes** is disabled, it is possible for the file system utilization to be reported at more than 100%.

Settings in Web Admin	Settings in CLI	Description	Notes
Enable global spillover	--no-spillover	Specifies how to handle write operations to a node pool that is not writable.	- When enabled, redirects write operations from a node pool that is not writable either to another node pool or anywhere on the cluster (the default). - When disabled, returns a disk space error for write operations to a node pool that is not writable.
Spillover Data Target	--spillover-target --spillover-anywhere	Specifies another storage pool to target when a storage pool is not writable.	When spillover is enabled, but it is important that data writes do not fail, select **anywhere** for the **Spillover Data Target** setting, even if file pool policies send data to specific pools.
Manage protection settings	--automatically-manage-protection	When this setting is enabled, SmartPools manages requested protection levels automatically.	When **Apply to files with manually-managed protection** is enabled, overwrites any protection settings that were configured through File System Explorer or the command-line interface.
Manage I/O optimization settings	--automatically-manage-io-optimization	When enabled, uses SmartPools technology to manage I/O optimization.	When **Apply to files with manually-managed I/O optimization settings** is enabled, overwrites any I/O optimization settings that were configured through File System Explorer or the command-line interface

Managing file pool policies

You can modify, reorder, copy, and remove custom file pool policies. Although you can modify the default file pool policy, you cannot reorder or remove it.

To manage file pool policies, you can perform the following tasks:

- Modify file pool policies
- Modify the default file pool policy
- Copy file pool policies
- Use a file pool policy template
- Reorder file pool policies
- Delete file pool policies

Configure default file pool protection settings

You can configure default file pool protection settings. The default settings are applied to any file that is not covered by another file pool policy.

⚠ CAUTION

If existing file pool policies direct data to a specific storage pool, do not add or modify a file pool policy to the `anywhere` option for the Data storage target option. Target a specific file pool instead.

Procedure

1. Click **File System** › **Storage Pools** › **File Pool Policies**.
2. In the **File Pool Policies** tab, next to Default Policy in the list, click **View/Edit**.

   The **View Default Policy Details** dialog box is displayed.
3. Click **Edit Policy**.

   The **Edit Default Policy Details** dialog box is displayed.
4. In the **Apply SmartPools Actions to Selected Files** section, choose the storage settings that you want to apply as the default for **Storage Target**, **Snapshot Storage Target**, and **Requested Protection Level**.
5. Click **Save Changes**, and then click **Close**.

Results

The next time the SmartPools job runs, the settings that you selected are applied to any file that is not covered by another file pool policy.

Default file pool requested protection settings

Default protection settings include specifying the data storage target, snapshot storage target, requested protection, and SSD strategy for files that are filtered by the default file pool policy.

Settings (Web Admin)	Settings (CLI)	Description	Notes
Storage Target	--data-storage-target --data-ssd-strategy	Specifies the storage pool (node pool or tier) that you want to target with this file pool policy. ⚠ CAUTION **If existing file pool policies direct data to a specific storage pool, do not configure other file pool policies with `anywhere` for the Data storage target option. Because the specified storage pool is included when you use `anywhere`, target specific storage pools to avoid unintentional file storage locations.** Select one of the following options to define your SSD strategy:	Note If GNA is not enabled and the storage pool that you choose to target does not contain SSDs, you cannot define an SSD strategy. **Use SSDs for metadata read acceleration** writes both file data and metadata to HDD storage pools but adds an additional SSD mirror if possible to accelerate read performance. Uses HDDs to provide reliability and an extra

Settings (Web Admin)	Settings (CLI)	Description	Notes
		Use SSDs for metadata read acceleration Default. Write both file data and metadata to HDDs and metadata to SSDs. Accelerates metadata reads only. Uses less SSD space than the **Metadata read/write acceleration** setting. Use SSDs for metadata read/write acceleration Write metadata to SSD pools. Uses significantly more SSD space than **Metadata read acceleration**, but accelerates metadata reads and writes. Use SSDs for data & metadata Use SSDs for both data and metadata. Regardless of whether global namespace acceleration is enabled, any SSD blocks reside on the storage target if there is room. Avoid SSDs Write all associated file data and metadata to HDDs only. ⚠ CAUTION **Use this to free SSD space only after consulting with Isilon Technical Support personnel; the setting can negatively affect performance.**	metadata mirror to SSDs, if available, to improve read performance. Recommended for most uses. When you select **Use SSDs for metadata read/write acceleration** , the strategy uses SSDs, if available in the storage target, for performance and reliability. The extra mirror can be from a different storage pool using GNA enabled or from the same node pool. Neither the **Use SSDs for data & metadata** strategy nor the **Use SSDs for data & metadata** strategy result in the creation of additional mirrors beyond the normal requested protection. Both file data and metadata are stored on SSDs if available within the file pool policy. This option requires a significant amount of SSD storage.
Snapshot storage target	--snapshot-storage-target --snapshot-ssd-strategy	Specifies the storage pool that you want to target for snapshot storage with this file pool policy. The settings are the same as those for data storage target, but apply to snapshot data.	Notes for data storage target apply to snapshot storage target
Requested protection	--set-requested-protection	**Default of storage pool**. Assign the default requested protection of the storage pool to the filtered files. **Specific level**. Assign a specified requested protection to the filtered files.	To change the requested protection , select a new value from the list.

Configure default I/O optimization settings

You can configure default I/O optimization settings.

Procedure

1. Click **File System** › **Storage Pools** › **File Pool Policies**.
2. In the **File Pool Policies** tab, next to Default Policy in the list, click **View/Edit**.

   The **View Default Policy Details** dialog box is displayed.
3. Click **Edit Policy**.

The **Edit Default Policy Details** dialog box is displayed.

4. In the **Apply SmartPools Actions to Selected Files** section, under I/O Optimization Settings,choose the settings that you want to apply as the default for **Write Performance** and **Data Access Pattern**.
5. Click **Save Changes**, and then click **Close**.

Results

The next time the SmartPools job runs, the settings that you selected are applied to any file that is not covered by another file pool policy.

Default file pool I/O optimization settings

You can manage the I/O optimization settings that are used in the default file pool policy, which can include files with manually managed attributes.

To allow SmartPools to overwrite optimization settings that were configured using File System Explorer or the `isi set` command, select the **Including files with manually-managed I/O optimization settings** option in the **Default Protection Settings** group. In the CLI, use the `--automatically-manage-io-optimization` option with the `isi storagepool settings modify` command.

Setting (Web Admin)	Setting (CLI)	Description	Notes
Write Performance	--enable-coalescer	Enables or disables SmartCache (also referred to as the coalescer).	**Enable SmartCache** is the recommended setting for optimal write performance. With asynchronous writes, the Isilon server buffers writes in memory. However, if you want to disable this buffering, we recommend that you configure your applications to use synchronous writes. If that is not possible, disable SmartCache.
Data Access Pattern	--data-access-pattern	Defines the optimization settings for accessing concurrent, streaming, or random data types.	Files and directories use a concurrent access pattern by default. To optimize performance, select the pattern dictated by your workflow. For example, a workflow heavy in video editing should be set to **Optimize for streaming access**. That workflow would suffer if the data access pattern was set to **Optimize for random access**.

Modify a file pool policy

You can modify a file pool policy.

⚠ CAUTION

If existing file pool policies direct data to a specific storage pool, do not configure other file pool policies with `anywhere` for the Data storage target option. Because the specified storage pool is included when you use `anywhere`, target specific storage pools to avoid unintentional file storage locations.

Procedure

1. Click **File System** › **Storage Pools** › **File Pool Policies**.
2. In the **File Pool Policies** list, next to the policy you want to modify, click **View/Edit**.

   The **View File Pool Policy Details** dialog box is displayed.

3. Click **Edit Policy**.

   The **Edit File Pool Policy Details** dialog box is displayed.

4. Modify the policy settings, and then click **Save Changes**.
5. Click **Close** in the **View File Pool Policy Details** dialog box.

Results

Changes to the file pool policy are applied when the next SmartPools job runs. You can also start the SmartPools job manually to execute the policy immediately.

Prioritize a file pool policy

You can change the priority of custom file pool policies. File pool policies are evaluated in descending order according to their position in the file pool policies list.

By default, new policies are inserted immediately above the default file pool policy, which is always last in the list and therefore lowest in priority. You can give a custom policy higher or lower priority by moving it up or down in the list.

Procedure

1. Click **File System** › **Storage Pools** › **File Pool Policies**.

   The **File Pool Policies** tab displays two lists: **File Pool Policies** and **Policy Templates**.

2. In the **File Pool Policies** list, in the **Order** column, click an arrow icon next to a policy to move it up or down in the priority order.
3. Repeat the above step for each policy whose priority you want to change.

Results

When the SmartPools system job runs, it processes the file pool policies in priority order. The default file pool policy is applied to all files that are not matched by any other file pool policy.

Create a file pool policy from a template

You can create a new file pool policy from a policy template. The templates are pre-configured for typical work flows, such as archiving older files or managing virtual machines (.vmdk files).

Procedure

1. Click **File System** › **Storage Pools** › **File Pool Policies**.

   The **File Pool Policies** tab provides two lists: **File Pool Policies** and **Policy Templates**.

2. In the **Policy Templates** list, next to the template name that you want to use, click **View/Use Template**.

   The **View File Pool Policy Template Details** dialog box opens.

3. Click **Use Template**.

   The **Create a File Pool Policy** dialog box opens.

4. (Required) Specify a policy name and description, and modify any of the policy settings.
5. Click **Create Policy**.

Results

The new custom policy is added to the **File Pool Policies** list directly above the default policy.

Delete a file pool policy

You can delete any file pool policy except the default policy.

When you delete a file pool policy, its file pool is controlled either by another file pool policy or by the default policy the next time the SmartPools job runs.

Procedure

1. Click **File System** › **Storage Pools** › **File Pool Policies**.

   The **File Pool Policies** tab displays two lists: **File Pool Policies** and **Policy Templates**.
2. In the **File Pool Policies** list, next to the policy that you want to delete, click **More** › **Delete Policy**.
3. In the **Confirm Delete** dialog box, click **Delete**.

Results

The file pool policy is removed from the **File Pool Policies** list.

Monitoring storage pools

You can access information on storage pool health and usage.

The following information is available:

- File pool policy health
- SmartPools health, including tiers, node pools, and subpools
- For each storage pool, percentage of HDD and SSD disk space usage
- SmartPools job status

Monitor storage pools

You can view the status of file pool policies, SmartPools, and settings.

Procedure

1. Click **File System** › **Storage Pools** › **Summary**.

   The **Summary** tab displays two areas: the **Status** list and the **Local Storage Usage** graph.
2. In the **Status** list, check the status of policies, SmartPools, and SmartPools settings.
3. (Optional) If the status of an item is other than `Good`, you can click **View Details** to view and fix any issues.
4. In the **Local Storage Usage** area, view the statistics associated with each node pool.

   If node pool usage is unbalanced, for example, you might want to consider whether to modify your file pool policies.

View subpools health

OneFS exposes unhealthy subpools in a list so that you can correct any issues.

A subpool is otherwise known as a disk pool, a collection of disks that is part of a node pool.

Procedure

1. Click **File System** › **Storage Pools** › **SmartPools**.

   The **SmartPools** tab displays three groupings: **Tiers & Node Pools**, **Compatibilities**, and **Subpools Health**.
2. In the **Subpools Health** area, review details of, and mitigate, any unhealthy subpools.

View the results of a SmartPools job

You can review detailed results from the last time the SmartPools job ran.

Procedure

1. Click **Cluster Management** › **Job Operations** › **Job Reports**.

   The **Jobs Reports** tab displays a list of job reports.
2. In the **Job Reports** list, in the **Type** column, find the latest SmartPools job, and click **View Details**.

   The **View Job Report Details** dialog box opens, displaying the job report.
3. Scroll through the report to see the results of each file pool policy.
4. Click **Close** in the **View Job Report Details** dialog box when you are finished.

CHAPTER 20

System jobs

This section contains the following topics:

System jobs overview

The most critical function of OneFS is maintaining the integrity of data on your Isilon cluster. Other important system maintenance functions include monitoring and optimizing performance, detecting and mitigating drive and node failures, and freeing up available space.

Because maintenance functions use system resources and can take hours to run, OneFS performs them as jobs that run in the background through a service called Job Engine. The time it takes for a job to run can vary significantly depending on a number of factors. These include other system jobs that are running at the same time; other processes that are taking up CPU and I/O cycles while the job is running; the configuration of your cluster; the size of your data set; and how long since the last iteration of the job was run.

Up to three jobs can run simultaneously. To ensure that maintenance jobs do not hinder your productivity or conflict with each other, Job Engine categorizes them, runs them at different priority and impact levels, and can temporarily suspend them (with no loss of progress) to enable higher priority jobs and administrator tasks to proceed.

In the case of a power failure, Job Engine uses a checkpoint system to resume jobs as close as possible to the point at which they were interrupted. The checkpoint system helps Job Engine keep track of job phases and tasks that have already been completed. When the cluster is back up and running, Job Engine restarts the job at the beginning of the phase or task that was in process when the power failure occurred.

As system administrator, through the Job Engine service, you can monitor, schedule, run, terminate, and apply other controls to system maintenance jobs. The Job Engine provides statistics and reporting tools that you can use to determine how long different system jobs take to run in your OneFS environment.

Note

To initiate any Job Engine tasks, you must have the role of SystemAdmin in the OneFS system.

System jobs library

OneFS contains a library of jobs that runs in the background to maintain your Isilon cluster. Some jobs are automatically started by OneFS when particular conditions arise, and some jobs have a default schedule. However, you can run all jobs manually or schedule them according to your workflow.

Job name	Description	Exclusion Set	Impact Policy	Priority	Operation
AutoBalance	Balances free space in a cluster, and is most efficient in clusters that contain only hard disk drives (HDDs). Run as part of MultiScan, or automatically by the system if MultiScan is disabled.	Restripe	Low	4	Auto

Job name	Description	Exclusion Set	Impact Policy	Priority	Operation
AutoBalanceLin	Balances free space in a cluster, and is most efficient in clusters when file system metadata is stored on solid state drives (SSDs). Run as part of MultiScan, or automatically by the system if MultiScan is disabled.	Restripe	Low	4	Auto
AVScan	Performs an antivirus scan on all files.	None	Low	6	Manual
Collect	Reclaims free space that previously could not be freed because the node or drive was unavailable. Run as part of MultiScan, or automatically by the system if MultiScan is disabled.	Mark	Low	4	Auto
Dedupe*	Scans a directory for redundant data blocks and deduplicates all redundant data stored in the directory. Available only if you activate a SmartDedupe license.	None	Low	4	Manual
DedupeAssessment	Scans a directory for redundant data blocks and reports an estimate of the amount of space that could be saved by deduplicating the directory.	None	Low	6	Manual
DomainMark	Associates a path, and the contents of that path, with a domain.	None	Low	5	Manual
FlexProtect	Scans the file system after a device failure to ensure that all files remain protected. FlexProtect is most efficient in clusters that contain only HDDs.	Restripe	Medium	1	Auto

Job name	Description	Exclusion Set	Impact Policy	Priority	Operation
	Note Unlike HDDs and SSDs that are used for storage, when an SSD used for L3 cache fails, the drive state should immediately change to REPLACE without a FlexProtect job running. An SSD drive used for L3 cache contains only cache data that does not have to be protected by FlexProtect. After the drive state changes to REPLACE, you can pull and replace the failed SSD.				
FlexProtectLin	Scans the file system after a node failure to ensure that all files remain protected. Most efficient when file system metadata is stored on SSDs.	Restripe	Medium	1	Auto
FSAnalyze	Gathers information about the file system.	None	Low	1	Scheduled
IntegrityScan	Verifies file system integrity.	Mark	Medium	1	Manual
MediaScan	Locates and clears media-level errors from disks.	Restripe	Low	8	Scheduled
MultiScan	Performs the work of the AutoBalance and Collect jobs simultaneously.	Restripe Mark	Low	4	Auto
PermissionRepair	Corrects file and directory permissions in the `/ifs` directory.	None	Low	5	Manual
QuotaScan*	Updates quota accounting for domains created on an existing file tree. Available only if you activate a SmartQuotas license.	None	Low	6	Auto

Job name	Description	Exclusion Set	Impact Policy	Priority	Operation
SetProtectPlus	Applies a default file policy across the cluster. Runs only if a SmartPools license is not active.	Restripe	Low	6	Manual
ShadowStoreDelete	Frees space that is associated with a shadow store.	None	Low	2	Scheduled
SmartPools*	Enforces SmartPools file policies. Available only if you activate a SmartPools license.	Restripe	Low	6	Scheduled
SnapRevert	Reverts an entire snapshot back to head.	None	Low	5	Manual
SnapshotDelete	Creates free space associated with deleted snapshots.	None	Medium	2	Auto
TreeDelete	Deletes a specified file path in the `/ifs` directory.	None	Medium	4	Manual
Upgrade	Upgrades the file system after a software version upgrade. **Note** The Upgrade job should be run only when you are updating your cluster with a major software version. For complete information, see the *Isilon OneFS Upgrade Planning and Process Guide*.	Restripe	Medium	3	Manual
* Available only if you activate an additional license					

Job operation

OneFS includes system maintenance jobs that run to ensure that your Isilon cluster performs at peak health. Through the Job Engine, OneFS runs a subset of these jobs automatically, as needed, to ensure file and data integrity, check for and mitigate drive and node failures, and optimize free space. For other jobs, for example, Dedupe, you can use Job Engine to start them manually or schedule them to run automatically at regular intervals.

The Job Engine runs system maintenance jobs in the background and prevents jobs within the same classification (exclusion set) from running simultaneously. Two exclusion sets are enforced: restripe and mark.

Restripe job types are:

- AutoBalance
- AutoBalanceLin
- FlexProtect
- FlexProtectLin
- MediaScan
- MultiScan
- SetProtectPlus
- SmartPools

Mark job types are:

- Collect
- IntegrityScan
- MultiScan

Note that MultiScan is a member of both the restripe and mark exclusion sets. You cannot change the exclusion set parameter for a job type.

The Job Engine is also sensitive to job priority, and can run up to three jobs, of any priority, simultaneously. Job priority is denoted as 1–10, with 1 being the highest and 10 being the lowest. The system uses job priority when a conflict among running or queued jobs arises. For example, if you manually start a job that has a higher priority than three other jobs that are already running, Job Engine pauses the lowest-priority active job, runs the new job, then restarts the older job at the point at which it was paused. Similarly, if you start a job within the restripe exclusion set, and another restripe job is already running, the system uses priority to determine which job should run (or remain running) and which job should be paused (or remain paused).

Other job parameters determine whether jobs are enabled, their performance impact, and schedule. As system administrator, you can accept the job defaults or adjust these parameters (except for exclusion set) based on your requirements.

When a job starts, the Job Engine distributes job segments—phases and tasks—across the nodes of your cluster. One node acts as job coordinator and continually works with the other nodes to load-balance the work. In this way, no one node is overburdened, and system resources remain available for other administrator and system I/O activities not originated from the Job Engine.

After completing a task, each node reports task status to the job coordinator. The node acting as job coordinator saves this task status information to a checkpoint file. Consequently, in the case of a power outage, or when paused, a job can always be restarted from the point at which it was interrupted. This is important because some jobs can take hours to run and can use considerable system resources.

Job performance impact

The Job Engine service monitors system performance to ensure that maintenance jobs do not significantly interfere with regular cluster I/O activity and other system administration tasks. Job Engine uses impact policies that you can manage to control when a job can run and the system resources that it consumes.

Job Engine has four default impact policies that you can use but not modify. The default impact policies are:

Impact policy	Allowed to run	Resource consumption
LOW	Any time of day.	Low
MEDIUM	Any time of day.	Medium
HIGH	Any time of day.	High
OFF_HOURS	Outside of business hours. Business hours are defined as 9AM to 5pm, Monday through Friday. OFF_HOURS is paused during business hours.	Low

If you want to specify other than a default impact policy for a job, you can create a custom policy with new settings.

Jobs with a low impact policy have the least impact on available CPU and disk I/O resources. Jobs with a high impact policy have a significantly higher impact. In all cases, however, the Job Engine uses CPU and disk throttling algorithms to ensure that tasks that you initiate manually, and other I/O tasks not related to the Job Engine, receive a higher priority.

Job priorities

Job priorities determine which job takes precedence when more than three jobs of different exclusion sets attempt to run simultaneously. The Job Engine assigns a priority value between 1 and 10 to every job, with 1 being the most important and 10 being the least important.

The maximum number of jobs that can run simultaneously is three. If a fourth job with a higher priority is started, either manually or through a system event, the Job Engine pauses one of the lower-priority jobs that is currently running. The Job Engine places the paused job into a priority queue, and automatically resumes the paused job when one of the other jobs is completed.

If two jobs of the same priority level are scheduled to run simultaneously, and two other higher priority jobs are already running, the job that is placed into the queue first is run first.

Managing system jobs

The Job Engine enables you to control periodic system maintenance tasks that ensure OneFS file system stability and integrity. As maintenance jobs run, the Job Engine constantly monitors and mitigates their impact on the overall performance of the cluster.

As system administrator, you can tailor these jobs to the specific workflow of your Isilon cluster. You can view active jobs and job history, modify job settings, and start, pause, resume, cancel, and update job instances.

View active jobs

If you are noticing slower system response while performing administrative tasks, you can view jobs that are currently running on your Isilon cluster.

Procedure

1. Click **Cluster Management** › **Job Operations** › **Job Summary**.
2. In the **Active Jobs** table, view status information about all currently running jobs, job settings, and progress details.
   a. You can perform bulk actions on the active jobs by selecting the **Status** check box, then selecting an action from the **Select a bulk action** drop-down list.

View job history

If you want to check the last time a critical job ran, you can view recent activity for a specific job, or for all jobs.

Procedure

1. Click **Cluster Management** › **Job Operations** › **Job Reports**.

   The **Job Reports** table displays a chronological list of the last ten job events that have occurred on the cluster. Event information includes the time the event occurred, the job responsible for the event, and event results.
2. Filter reports by job type by selecting the job from the **Filter by Job Type** drop-down list and clicking **Reset**.
3. Click on **View Details** next to a job name to view recent events for only that job.

   Recent events for the job appear in the **View Job Report Details** window, and include information such as start time, duration, and whether or not the job was successful.

Start a job

By default, only some system maintenance jobs are scheduled to run automatically. However, you can start any of the jobs manually at any time.

Procedure

1. Click **Cluster Management** › **Job Operations** › **Job Summary**.
2. In the **Active Jobs** window, select the job you want to start and click **More**.
3. Click **Start Running Job**.

Pause a job

You can pause a job temporarily to free up system resources.

Procedure

1. Click **Cluster Management** › **Job Operations** › **Job Summary**.
2. In the **Active Jobs** table, click **More** for the job that you want to pause.
3. Click **Pause Running Job** in the menu that appears.

   The job remains paused until you resume it.

Resume a job

You can resume a paused job.

Procedure

1. Click **Cluster Management** › **Job Operations** › **Job Summary**.
2. In the **Active Jobs** table, click **More** for the job that you want to pause.
3. Click **Resume Running Job** in the menu that appears.

Results

The job continues from the phase or task at which it was paused.

Cancel a job

If you want to free up system resources, or for any reason, you can permanently discontinue a running, paused, or waiting job.

Procedure

1. Click **Cluster Management** › **Job Operations** › **Job Summary**.
2. In the **Active Jobs** table, click **More** for the job that you want to cancel.
3. Click **Cancel Running Job** in the menu that appears.

Update a job

You can change the priority and impact policy of a running, waiting, or paused job.

When you update a job, only the current instance of the job runs with the updated settings. The next instance of the job returns to the default settings for that job.

Note

To change job settings permanently, see "Modify job type settings."

Procedure

1. Click **Cluster Management** › **Job Operations** › **Job Summary**.
2. In the **Active Jobs** table, click **View/Edit** for the job that you want to update.
3. (Required) In the **View Active Job Details** window, click **Edit Job**.
   a. Select a new priority level from the **Priority** drop-down list.
   b. Select an impact policy level from the **Impact Policy** drop-down list.
4. Click **Save Changes**.

   When you update a running job, the job automatically resumes. When you update a paused or idle job, the job remains in that state until you restart it.

Modify job type settings

You can customize system maintenance jobs for your administrative workflow by modifying the default priority level, impact level, and schedule for a job type.

Procedure

1. Click **Cluster Management** › **Job Operations** › **Job Types**.
2. In the **Job Types** table, locate the row for the policy you want to modify and click **View / Edit**.

The **View Job Type Details** window appears, displaying current default settings, schedule, current state, and recent activity.

3. Click **Edit Job Type**. The **Edit Job Type Details** window appears.
4. Modify the details you want to change. You can modify the default priority, the default impact policy, whether the job is enabled, and whether the job runs manually or on a schedule.
5. Click **Scheduled** to modify a job schedule, then select the schedule option from the drop-down list.
6. Click **Save Changes**.

   The modifications are saved and applied to all instances of that job type. The results are shown in the **View Job Type Details** window.
7. Click **Close**.

Managing impact policies

For system maintenance jobs that run through the Job Engine service, you can create and assign policies that help control how jobs affect system performance.

As system administrator, you can create, copy, modify, and delete impact policies, and view their settings.

Create an impact policy

The Job Engine includes four impact policies, which you cannot modify or delete. However, you can create and configure new impact policies.

Procedure

1. Click **Cluster Management › Job Operations › Impact Policies**.
2. Click **Add an Impact Policy**.

   The **Create Impact Policy** window appears.
3. In the **Name** text field, type a name for the policy. This field is required.
4. (Required) In the **Description** text field, type a comment about the impact policy.

   Include information specific to the impact policy such as unique schedule parameters or logistical requirements that make the impact policy necessary.
5. Click **Add an Impact Policy Interval**.
   a. In the **Add an Impact Policy Interval** window, select the impact level and start and end times from the drop-down lists.
   b. Click **Add Impact Policy Interval**.

   The **Add an Impact Policy Interval** window disappears, and the settings you selected appear in the **Impact Schedule** table.
6. Click **Create Impact Policy**.

   Your copy of the impact policy is saved and is listed in alphabetical order in the **Impact Policies** table.

Copy an impact policy

You can use a default impact policy as the template for a new policy by making and modifying a copy.

Procedure

1. Click **Cluster Management** › **Job Operations** › **Impact Policies**.
2. In the **Impact Policies** table, locate the row for the policy you want to copy and click **More**. The **Copy Impact Policy** window appears.
3. In the **Name** field, type a name for the new policy.
4. In the **Description** text field, enter a description for the new policy.

   Include information specific to the impact policy such as unique schedule parameters or logistical requirements that make the impact policy necessary.

5. Click **Add an Impact Policy Interval**.
   a. In the **Add an Impact Policy Interval** window, select the impact level and start and end times from the drop-down lists.
   b. Click **Add Impact Policy Interval**.

   The **Add an Impact Policy Interval** window closes, and the settings you selected appear in the **Impact Schedule** table.
6. Click **Copy Impact Policy**.

   Your copy of the impact policy is saved and is listed in alphabetical order in the **Impact Policies** table.

Modify an impact policy

You can change the name, description, and impact intervals of a custom impact policy.

Before you begin

You cannot modify the default impact policies, HIGH, MEDIUM, LOW, and OFF_HOURS. If you want to modify a policy, create and modify a copy of a default policy.

Procedure

1. Navigate to **Cluster Management** › **Job Operations** › **Impact Policies**.
2. In the **Impact Policies** table, click **View / Edit** for the policy you want to modify.

   The **Edit Impact Policy** window appears.
3. Click **Edit Impact Policy**, and modify one or all of the following:

Option	Description
Policy description	a. In the **Description** field, type a new overview for the impact policy. b. Click **Submit**.
Impact schedule	a. In the **Impact Schedule** area, modify the schedule of the impact policy by adding, editing, or deleting impact intervals. b. Click **Save Changes**.

The modified impact policy is saved and listed in alphabetical order in the **Impact Policies** table.

Delete an impact policy

You can delete impact policies that you have created.

You cannot delete default impact policies, `HIGH`, `MEDIUM`, `LOW`, and `OFF_HOURS`.

Procedure

1. Click **Cluster Management** › **Job Operations** › **Impact Policies**.
2. In the **Impact Policies** table, click **More** next to the custom impact policy that you want to delete.
3. Click **Delete**.

   A confirmation dialog box appears.
4. In the confirmation dialog box, click **Delete**.

View impact policy settings

You can view the impact policy settings for any job.

Procedure

1. Click **Cluster Management** › **Job Operations** › **Job Types**.

   The **Job Types** table is displayed.
2. If necessary, scroll through the **Job Types** table to find a specific job.

   The impact policy settings for the job are shown in the **Job Types** table.

Viewing job reports and statistics

You can generate reports for system jobs and view statistics to better determine the amounts of system resources being used.

Most system jobs controlled by the Job Engine run at a low priority and with a low impact policy, and generally do not have a noticeable impact on cluster performance.

A few jobs, because of the critical functions they perform, run at a higher priority and with a medium impact policy. These jobs include FlexProtect and FlexProtect Lin, FSAnalyze, SnapshotDelete, and TreeDelete.

As a system administrator, if you are concerned about the impact a system job might have on cluster performance, you can view job statistics and reports. These tools enable you to view detailed information about job load, including CPU and memory usage and I/O operations.

View statistics for a job in progress

You can view statistics for a job in progress.

Procedure

1. Click **Cluster Management** › **Job Operations** › **Job Summary**.

   You can view jobs that are running in the **Active Jobs** area.
2. Click the **View/Edit** option to the right of the job entry.

Results

The **View Active Jobs Details** screen opens, where you can view statistics such as processed data, elapsed time, phase, and progress, including an estimate of the time remaining for the job to complete.

View a report for a completed job

After a job finishes, you can view a report about the job.

Before you begin

A report for a job is not available until after the job is completed.

Procedure

1. Click **Cluster Management › Job Operations › Job Reports**.

   The **Job Reports** screen appears with a list of the last 10 completed jobs.

2. Locate the job whose report you want to view.

   If the job is not on the first page of the **Job Reports** screen, click the right-arrow icon to page through the list until you locate your job.

3. Click **View Details**.

   The **View Job Report Details** screen appears, listing job statistics such as elapsed time, CPU and memory usage, and total I/O operations.

4. When you are finished viewing the report, click **Close**.

CHAPTER 21

Networking

This section contains the following topics:

Networking overview

After you determine the topology of your network, you can set up and manage your internal and external networks.

There are two types of networks associated with an EMC Isilon cluster:

Internal
: Nodes communicate with each other using a high speed low latency InfiniBand network. You can optionally configure a second InfiniBand network as a failover for redundancy.

External
: Clients connect to the cluster through the external network with Ethernet. The Isilon cluster supports standard network communication protocols, including NFS, SMB, HTTP, and FTP. The cluster includes various external Ethernet connections, providing flexibility for a wide variety of network configurations. External network speeds vary by product.

Internal network overview

The EMC Isilon cluster must connect to at least one high-speed, low-latency InfiniBand switch for internal communications and data transfer. The connection to the InfiniBand switch is also referred to as an internal network. The internal network is separate from the external network (Ethernet) by which users access the cluster.

Upon initial configuration of your cluster, OneFS creates a default internal network for the InfiniBand switch. The interface to the default internal network is int-a. An internal network for a second InfiniBand switch can be added for redundancy and failover. Failover allows continuous connectivity during path failures. The interface to the default internal network is int-b, which is referred to as int-b/failover in the web administration interface.

Internal IP address ranges

The number of IP addresses assigned to the internal network determines how many nodes can be joined to the EMC Isilon cluster.

When you initially configure the cluster, you specify one or more IP address ranges for the internal InfiniBand network. This range of addresses is used by the nodes to communicate with each other. It is recommended that you create a range of addresses large enough to accommodate adding additional nodes to your cluster. If the IP address range defined during the initial configuration is too restrictive for the size of the internal network, you can add ranges to the int-a network and int-b network. For certain configuration changes, such as deleting an IP address assigned to a node, the cluster must be restarted.

While all clusters will have, at minimum, one internal InfiniBand network (int-a), to enable a second internal network (int-b) you must assign another IP address range to it. To enable internal network failover, assign an IP address range to the failover network. This range is used to refer to the actual IP addresses in use to provide seamless internal IP address failover.

Internal network failover

You can configure an internal switch as a failover network to provide redundancy for intra-cluster communications.

Enable an internal failover network by connecting the int-a interfaces of each node in the cluster to one switch, connecting the int-b ports on each node to another switch, and then restarting the EMC Isilon cluster.

In addition to the IP address range assigned to the int-a internal network, if you enable failover on a second InfiniBand switch, you must assign an IP address range that points to actual IP addresses used by the cluster. These addresses enable seamless failover in the event that either the int-a or int-b switches fail.

External client network overview

You connect a client computer to the EMC Isilon cluster through the external network. OneFS supports network subnets, IP address pools, and features network provisioning rules to simplify configuration.

Subnets simplify external (front-end) network management and provide flexibility in implementing and maintaining the cluster network. You can create IP address pools within subnets to partition your network interfaces according to workflow or node type. You can configure external network settings through provisioning rules and then those rules are applied to nodes that are added to the cluster.

You must initially configure the default external IP subnet in IPv4 format. After configuration is complete, you can configure additional subnets using IPv4 or IPv6.

IP address pools can be associated with a node or a group of nodes as well as with the NIC ports on the nodes. For example, based on the network traffic that you expect, you might decide to establish one subnet for storage nodes and another subnet for accelerator nodes.

How you set up your external network subnets depends on your network topology. In a basic network topology where all client-node communication occurs through a single gateway, only a single external subnet is required. If clients connect through multiple subnets or internal connections, you must configure multiple external network subnets.

External network settings

A default external network subnet is created during the initial set up of your EMC Isilon cluster. You can make modifications to this subnet, create new subnets, and make additional configuration changes to the external network.

During initial cluster setup, OneFS performs the following actions:

- Creates a default external network subnet called subnet0, with the specified netmask, gateway, and SmartConnect service address.
- Creates a default IP address pool called pool0 with the specified IP address range, the SmartConnect zone name, and the external interface of the first node in the cluster as the only member.
- Creates a default network provisioning rule called rule0, which automatically assigns the first external interface for all newly added nodes to pool0.
- Adds pool0 to subnet0 and configures pool0 to use the virtual IP of subnet0 as its SmartConnect service address.
- Sets the global, outbound DNS settings to the domain name server list and DNS search list, if provided.

Once the initial external network has been established, you can configure the following information about your external network:

- Netmask
- IP address range
- Gateway
- Domain name server list (optional)
- DNS search list (optional)
- SmartConnect zone name (optional)
- SmartConnect service address (optional)

You can make modifications to the external network through the web administration interface and the command-line interface.

IP address pools

You can partition EMC Isilon cluster nodes and external network interfaces into logical IP address pools. IP address pools are also utilized when configuring SmartConnect zones and IP failover support for protocols such as NFS. Multiple pools for a single subnet are available only if you activate a SmartConnect Advanced license.

IP address pools:

- Map available addresses to configured interfaces.
- Belong to external network subnets.
- Partition network interfaces on your cluster into pools.
- Can be to assigned to groups in your organization.

The IP address pool of a subnet consists of one or more IP address ranges and a set of cluster interfaces. All IP address ranges in a pool must be unique.

A default IP address pool is configured during the initial cluster setup through the command-line configuration wizard. You can modify the default IP address pool at any time. You can also add, remove, or modify additional IP address pools.

If you add external network subnets to your cluster through the subnet wizard, you must specify the IP address pools that belong to the subnet.

IP address pools are allocated to external network interfaces either dynamically or statically. The static allocation method assigns one IP address per pool interface. The IP addresses remain assigned, regardless of that interface's status, but the method does not guarantee that all IP addresses are assigned. The dynamic allocation method distributes all pool IP addresses, and the IP address can be moved depending on the interface's status and connection policy settings.

IPv6 support

You can configure dual stack support for IPv6.

With dual-stack support in OneFS, you can configure both IPv4 and IPv6 addresses. However, configuring an EMC Isilon cluster to use IPv6 exclusively is not supported. When you set up the cluster, the initial subnet must consist of IPv4 addresses.

The following table describes important distinctions between IPv4 and IPv6.

IPv4	IPv6
32-bit addresses	128-bit addresses

IPv4	IPv6
Subnet mask	Prefix length
Address Resolution Protocol (ARP)	Neighbor Discovery Protocol (NDP)

SmartConnect module

SmartConnect is a module that specifies how the DNS server on the EMC Isilon cluster handles connection requests from clients and the methods used to assign IP addresses to network interfaces.

Settings and policies configured for SmartConnect are applied per IP address pool. You can configure basic and advanced SmartConnect settings.

SmartConnect Basic

SmartConnect Basic is included with OneFS as a standard feature and does not require a license.

SmartConnect Basic supports the following settings:

- Specification of the DNS zone
- Round robin connection balancing method
- Service subnet to answer DNS requests

SmartConnect Basic has the following limitations to IP address pool configuration:

- You may only specify a static IP address allocation policy.
- You cannot specify an IP address failover policy.
- You cannot specify an IP address rebalance policy.
- You may only assign one IP address pool per external network subnet.

SmartConnect Advanced

SmartConnect Advanced extends the settings available from SmartConnect Basic. It requires an active license.

SmartConnect Advanced supports the following settings:

- Round robin, CPU utilization, connection counting, and throughput balancing methods.
- Static and dynamic IP address allocation.

SmartConnect Advanced allows you to specify the following IP address pool configuration options:

- You can define an IP address failover policy for the IP address pool.
- You can define an IP address rebalance policy for the IP address pool.
- SmartConnect Advanced supports multiple IP address pools per external subnet to allow multiple DNS zones within a single subnet.

Connection balancing

The connection balancing policy determines how the DNS server handles client connections to the EMC Isilon cluster.

You can specify one of the following balancing methods:

Round robin

Selects the next available node on a rotating basis. This is the default method. Without a SmartConnect license for advanced settings, this is the only method available for load balancing.

Connection count

Determines the number of open TCP connections on each available node and selects the node with the fewest client connections.

Network throughput

Determines the average throughput on each available node and selects the node with the lowest network interface load.

CPU usage

Determines the average CPU utilization on each available node and selects the node with lightest processor usage.

IP address allocation

The IP address allocation policy ensures that all of the IP addresses in the pool are assigned to an available network interface.

You can specify whether to use static or dynamic allocation.

Static

Assigns one IP address to each network interface added to the IP address pool, but does not guarantee that all IP addresses are assigned.

Once assigned, the network interface keeps the IP address indefinitely, even if the network interface becomes unavailable. To release the IP address, remove the network interface from the pool or remove it from the cluster.

Without a license for SmartConnect Advanced, static is the only method available for IP address allocation.

Dynamic

Assigns IP addresses to each network interface added to the IP address pool until all IP addresses are assigned. This guarantees a response when clients connect to any IP address in the pool.

If a network interface becomes unavailable, its IP addresses are automatically moved to other available network interfaces in the pool as determined by the IP address failover policy.

This method is only available with a license for SmartConnect Advanced.

Allocation recommendations based on file sharing protocols

It is recommended that you select a static allocation method if your clients connect through stateful protocols and a dynamic allocation method with stateless protocols.

The following table displays several common protocols and the recommended allocation method:

File sharing protocol	Recommended allocation method
• SMB • NFSv4 • HTTP • FTP	Static

File sharing protocol	Recommended allocation method
• sFTP • FTPS • HDFS • SyncIQ	
• NFSv2 • NFSv3	Dynamic

IP address failover

The IP address failover policy specifies how to handle the IP addresses of network interfaces that become unavailable

To define an IP address failover policy, you must have a license for SmartConnect Advanced, and the IP address allocation policy must be set to dynamic. Dynamic IP allocation ensures that all of the IP addresses in the pool are assigned to available network interfaces.

When a network interface becomes unavailable, the IP addresses that were assigned to it are redistributed to available network interfaces according to the IP address failover policy. Subsequent client connections are directed to the new network interfaces.

You can select one of the following the connection balancing methods to determine how the IP address failover policy selects which network interface receives a redistributed IP address:

- Round robin
- Connection count
- Network throughput
- CPU usage

IP address rebalancing

The IP address rebalance policy specifies when to redistribute IP addresses if one or more previously unavailable network interfaces becomes available again.

To define an IP address rebalance policy, you must have a license for SmartConnect Advanced, and the IP address allocation policy must be set to dynamic. Dynamic IP addresses allocation ensures that all of the IP addresses in the pool are assigned to available network interfaces.

You can set rebalancing to occur manually or automatically:

Manual

Does not redistribute IP addresses until you manually issue a rebalance command through the command-line interface.

Upon rebalancing, IP addresses will be redistributed according to the connection balancing method specified by the IP address failover policy defined for the IP address pool.

Automatic

Automatically redistributes IP addresses according to the connection balancing method specified by the IP address failover policy defined for the IP address pool.

Automatic rebalance may also be triggered by changes to cluster nodes, network interfaces, or the configuration of the external network.

Note

Rebalancing can disrupt client connections. Ensure the client workflow on the IP address pool is appropriate for automatic rebalancing.

SmartConnect DNS service

The SmartConnect service IP address handles client DNS requests and is configured as a subnet setting.

You must have at least one subnet configured with a SmartConnect service IP address in order to handle client DNS requests.

Do not designate an IP address from a pool as the SmartConnect service IP. The SmartConnect service IP should only answer DNS requests; client connections through the SmartConnect service IP result in unexpected behavior or disconnection.

When configuring IP address pool settings, you can designate any subnet with a service IP address to act as the SmartConnect DNS service for the pool. You can assign a SmartConnect DNS service to multiple pools.

The SmartConnect DNS service handles all incoming DNS requests on behalf of each associated pool's SmartConnect zone and it distributes the requests according to each pool's connection balancing policy.

Any pool that does not specify a SmartConnect DNS service is excluded when answering incoming DNS requests.

Note

SmartConnect requires that you add a new name server (NS) record to the existing authoritative DNS zone that contains the cluster, and you must provide the fully qualified domain name (FQDN) of the SmartConnect zone.

DNS name resolution

You can designate up to three DNS servers and up to six search domains for your external network.

You can configure the DNS server settings during initial cluster configuration with the command-line Configuration wizard. After the initial configuration, you can modify the DNS server settings through the web administration interface or through the `isi networks` command.

NIC aggregation

Network interface card (NIC) aggregation, also known as link aggregation, is optional, and enables you to combine the bandwidth of a node's physical network interface cards into a single logical connection. NIC aggregation provides improved network throughput.

Note

Configuring link aggregation is an advanced function of network switches. Consult your network switch documentation before configuring your EMC Isilon cluster for link aggregation.

NIC aggregation can be configured during the creation of a new external network subnet. Alternatively, you can configure NIC aggregation on the existing IP address pool of a subnet.

- OneFS provides support for the following link aggregation methods:

  Link Aggregation Control Protocol (LACP)
  : This method supports the IEEE 802.3ad Link Aggregation Control Protocol (LACP). Balances outgoing traffic across the interfaces based on hashed protocol header information that includes the source and destination address and the VLAN tag, if available. Also assembles interfaces of the same speed into groups called Link Aggregated Groups (LAGs) and balances traffic across the fastest LAGs. This option is the default mode for new pools.

  Legacy Fast EtherChannel (FEC) mode
  : This method aggregates network interfaces through an older FEC driver recommended for OneFS 6.5 and earlier.

  Etherchannel (FEC)
  : This method provides static balancing on aggregated interfaces through the Cisco Fast EtherChannel (FEC) driver, which is found on older Cisco switches. Capable of load balancing traffic across Fast Ethernet links. Allows multiple physical Fast Ethernet links to combine into one logical channel.

  Active / Passive Failover
  : This method switches to the next active interface when the primary interface becomes unavailable. Manages traffic only through a primary interface. The second interface takes over the work of the first as soon as it detects an interruption in communication.

  Round-Robin
  : This method rotates connections through the nodes in a first-in, first-out sequence, handling all processes without priority. Balances outbound traffic across all active ports in the aggregated link and accepts inbound traffic on any port.

- Some NICs may allow aggregation of ports only on the same network card.
- For LACP and FEC aggregation modes, the switch must support IEEE 802.3ad link aggregation. Since the trunks on the network switch must also be configured, the node must be connected with the correct ports on the switch.

Routing options

By default, outgoing client traffic on the EMC Isilon cluster is destination-based; traffic is routed to a particular gateway based on where the traffic is going. OnefS supports source-based routing and static routes; these options allow for more granular control of the direction of outgoing client traffic.

Source-based routing

Source-based routing selects which gateway to direct outgoing client traffic through based on the source IP address in each packet header.

In the following example, you enable source-based routing on an Isilon cluster that is connected to SubnetA and SubnetB. Each subnet is configured with a SmartConnect zone and a gateway, also labeled A and B. When a client on SubnetA makes a request to SmartConnect ZoneB, the response originates from ZoneB. This results in a ZoneB address as the source IP in the packet header, and the response is routed through GatewayB. Without source-based routing, the default route is destination-based, so the response is routed through GatewayA.

In another example, a client on SubnetC, which is not connected to the Isilon cluster, makes a request to SmartConnect ZoneA and ZoneB. The response from ZoneA is routed through GatewayA, and the response from ZoneB is routed through GatewayB. In other words, the traffic is split between gateways. Without source-based routing, both responses are routed through the same gateway.

When enabled, source-based routing automatically scans your network configuration to create client traffic rules. If you make modifications to your network configuration, such as changing the IP address of a gateway server, source-based routing adjusts the rules. Source-based routing is applied across the entire cluster and only supports the IPv4 protocol.

Enabling or disabling source-based routing goes into effect immediately. Packets in transit continue on their original courses, and subsequent traffic is routed based on the status change. Transactions composed of multiple packets might be disrupted or delayed if the status of source-based routing changes during transmission.

Source-based routing can conflict with static routes. If a routing conflict occurs, source-based routing rules are prioritized over the static route.

You might enable source-based routing if you have a large network with a complex topology. For example, if your network is a multi-tenant environment with several gateways, traffic is more efficiently distributed with source-based routing.

Static routing

A static route directs outgoing client traffic to a specified gateway based on the IP address the client is connected through.

You configure static routes by IP address pool, and each route applies to all network interfaces that are members of the IP address pool. Static routes only support the IPv4 protocol.

You might configure static routing in order to connect to networks that are unavailable through the default routes or if you have a small network that only requires one or two routes.

Note

If you have upgraded from a version earlier than OneFS 7.0, existing static routes that were added through rc scripts will no longer work and must be re-created by running the `isi networks modify pool` command with the `--add-static-routes` option.

VLANs

Virtual LAN (VLAN) tagging is an optional setting that enables an EMC Isilon cluster to participate in multiple virtual networks.

You can partition a physical network into multiple broadcast domains, or virtual local area networks (VLANs). You can enable a cluster to participate in a VLAN which allows multiple cluster subnet support without multiple network switches; one physical switch enables multiple virtual subnets.

VLAN tagging inserts an ID into packet headers. The switch refers to the ID to identify from which VLAN the packet originated and to which network interface a packet should be sent.

Configuring the internal network

You can modify the internal network settings of your EMC Isilon cluster.

The following actions are available:

- Modify the IP address ranges of the internal network and the int-b/failover network
- Modify the internal network netmask
- Configure and enable an internal failover network
- Disable internal network failover

You can configure the int-b/failover network to provide backup in the event of an int-a network failure. Configuration involves specifying a valid netmask and IP address range for the failover network.

Modify the internal IP address range

Each internal InfiniBand network requires an IP address range. The ranges should have a sufficient number of IP addresses for present operating conditions as well as future expansion and addition of nodes. You can add, remove, or migrate IP addresses for both the internal (int-a) and failback (int-b) networks.

Procedure

1. Click **Cluster Management › Network Configuration**.
2. In the **Internal Networks Settings** area, select the network that you want to add IP addresses for.
   - To select the int-a network, click **int-a**.
   - To select the int-b/failover network, click **int-b/Failover**.
3. In the **IP Ranges** area, you can add, delete, or migrate your IP address ranges.

   Ideally, the new range is contiguous with the previous one. For example, if your current IP address range is 192.168.160.60 - 192.168.160.162, the new range should start with 192.168.160.163.
4. Click **Submit**.
5. Restart the cluster, if needed.
   - If the IP address changes are within the internal network netmask, you do not need to restart the cluster.
   - If the IP address changes reduce the current range, you must restart the cluster.

- If you migrate the IP address ranges, you must restart the cluster.

Modify the internal network netmask

You can modify the netmask value for the internal network.

If the netmask is too restrictive for the size of the internal network, you must modify the netmask settings. It is recommended that you specify a class C netmask, such as `255.255.255.0`, for the internal netmask. This netmask is large enough to accommodate future nodes.

Note

For the changes in netmask value to take effect, you must reboot the cluster.

Procedure

1. Click **Cluster Configuration** › **Network Configuration**.
2. In the **Internal Network Settings** area, select the network that you want to configure the netmask for.
    - To select the int-a network, click **int-a**.
    - To select the int-b/Failover network, click **int-b/Failover**.

    It is recommended that the netmask values you specify for int-a and int-b/failover are the same. If you modify the netmask value of one, modify the other.
3. In the **Netmask** field, type a netmask value.

    You cannot modify the netmask value if the change invalidates any node addresses.
4. Click **Submit**.

    A dialog box prompts you to reboot the cluster.
5. Specify when you want to reboot the cluster.
    - To immediately reboot the cluster, click **Yes**. When the cluster finishes rebooting, the login page appears.
    - Click **No** to return to the **Edit Internal Network** page without changing the settings or rebooting the cluster.

Configure and enable internal failover

You can enable an internal failover on your EMC Isilon cluster.

Procedure

1. Click **Cluster Management** › **Network Configuration**.
2. In the **Internal Network Settings** area, click **int-b/Failover**.
3. In the **IP Ranges** area, for the **int-b** network, click **Add range**.
4. On the **Add IP Range** dialog box, enter the IP address at the low end of the range in the first **IP range** field.
5. In the second **IP range** field, type the IP address at the high end of the range.

    Ensure that there is no overlap of IP addresses between the int-a and int-b/failover network ranges. For example, if the IP address range for the int-a network is 192.168.1.1 - 192.168.1.100, specify a range of 192.168.2.1 - 192.168.2.100 for the int-b network.

6. Click **Submit**.
7. In the **IP Ranges** area for the **Failover** network, click **Add range**.

   Add an IP address range for the failover network, ensuring there is no overlap with the int-a network or the int-b network.

   The **Edit Internal Network** page appears, and the new IP address range appears in the **IP Ranges** list.
8. In the **Settings** area, specify a valid netmask. Ensure that there is no overlap between the IP address range for the int-b network or for the failover network.

   It is recommended that the netmask values you specify for int-a and int-b/failover are the same.
9. In the **Settings** area, for **State**, click **Enable** to enable the int-b and failover networks.
10. Click **Submit**.

    The **Confirm Cluster Reboot** dialog box appears.
11. Restart the cluster by clicking **Yes**.

Disable internal network failover

You can disable the int-b and failover internal networks.

Procedure

1. Click **Cluster Management** › **Network Configuration**.
2. In the **State** area, click **Disable**.
3. Click **Submit**.

   The **Confirm Cluster Reboot** dialog box appears.
4. Restart the cluster by clicking **Yes**.

Configuring an external network

You can configure all external network connections between the EMC Isilon cluster and client computers.

Adding a subnet

You can add and configure an external subnet.

Adding a subnet to the external network encompasses these tasks:

Procedure

1. Configuring subnet settings.
2. Adding an IP address to a new subnet.
3. (Optional) Configuring SmartConnect settings for a new subnet.
4. Selecting interface members for a new subnet.

Configure subnet settings

You can add a subnet to the external network of a cluster using the web administration interface or the Isilon command line. This procedure describes using the web administration interface to add a subnet.

Procedure

1. Click **Cluster Management** › **Network Configuration**.
2. In the **External Network Settings** area, click **Add subnet**.
3. In the **Basic** section, in the **Name** field, type a unique name for the subnet.

   The name can be up to 32 alphanumeric characters long and can include underscores or hyphens, but not spaces or other punctuation.
4. (Optional) In the **Description** field, type a descriptive comment about the subnet.

   The comment can be no more than 128 characters.
5. Specify the IP address format for the subnet and configure an associated netmask or prefix length setting:
   - For an IPv4 subnet, click **IPv4** in the **IP Family** list. In the **Netmask** field, type a dotted decimal octet (x.x.x.x) that represents the subnet mask.
   - For an IPv6 subnet, click **IPv6** in the **IP family** list. In the **Prefix length** field, type an integer (ranging from `1` to `128`) that represents the network prefix length.
6. In the **MTU** list, type or select the size of the maximum transmission units the cluster uses in network communication. Any numerical value is allowed, but must be compatible with your network and the configuration of all devices in the network path. Common settings are `1500` (standard frames) and `9000` (jumbo frames).

   Although OneFS supports both 1500 MTU and 9000 MTU, using a larger frame size for network traffic permits more efficient communication on the external network between clients and cluster nodes. For example, if a subnet is connected through a 10 GbE interface and NIC aggregation is configured for IP address pools in the subnet, it is recommended you set the MTU to 9000. To benefit from using jumbo frames, all devices in the network path must be configured to use jumbo frames.
7. In the **Gateway address** field, type the IP address of the gateway server device through which the cluster communicates with systems outside of the subnet.
8. In the **Gateway priority** field, type an integer for the priority of the subnet gateway for nodes assigned to more than one subnet.

   You can configure only one default gateway per node, but each subnet can be assigned a gateway. When a node belongs to more than one subnet, this option enables you to define the preferred default gateway. A value of `1` represents the highest priority, and `10` represents the lowest priority.
9. If you plan to use SmartConnect for connection balancing, in the **SmartConnect service IP** field, type the IP address that will receive all incoming DNS requests for each IP address pool according to the client connection policy. You must have at least one subnet configured with a SmartConnect service IP in order to use connection balancing.
10. (Optional) In the **Advanced** section, you can enable VLAN tagging if you want to enable the cluster to participate in virtual networks.

Note

Configuring a VLAN requires advanced knowledge of network switches. Consult your network switch documentation before configuring your cluster for a VLAN.

11. If you enable VLAN tagging, you must also type a `VLAN ID` that corresponds to the ID number for the VLAN set on the switch, with a value from 2 to 4094.
12. (Optional) In the **Hardware load balancing** field, type the IP address for a hardware load balancing switch using Direct Server Return (DSR). This routes all client traffic to the cluster through the switch. The switch determines which node handles the traffic for the client, and passes the traffic to that node.
13. Click **Next**.

    The **Step 2 of 4 -- IP Address Pool Settings** dialog box appears.

After you finish

The next step in the process of adding a new subnet is adding an IP address pool.

Add an IP address pool to a new subnet

You can partition the external network interface of your cluster into groups, or pools, of unique IP address ranges in a subnet.

Before you begin

You must specify basic subnet settings by completing the previous subnet wizard page.

Note

If your cluster is running SmartConnect Basic for connection balancing, you can configure only one IP address pool per subnet. If you activate a SmartConnect Advanced license, you can configure unlimited IP address pools per subnet.

Procedure

1. In the **Step 2 of 4 — IP Address Pool Settings** dialog box, type a unique **Name** for the IP address pool. The name can be up to 32 alphanumeric characters long and can include underscores or hyphens, but no spaces or other punctuation.
2. Type a `Description` for the IP address pool. The description can contain up to 128 characters.
3. In the **Access zone** list, click to select an access zone for the pool. OneFS includes a default system access zone.
4. In the **IP range (low-high)** area, click **New**.

   OneFS adds an IP address range with default **Low IP** and **High IP** values.
5. Click to select the default **Low IP** value. Replace the default value with the starting IP address of the subnet's IP address pool.
6. Click to select the default **High IP** value. Replace the default value with the ending IP address of the subnet's IP address pool.
7. (Optional) Add IP address ranges to the IP address pool by repeating steps 3 through 6 as needed.
8. Click **Next**.

   The **Step 3 of 4 — SmartConnect Settings** dialog box appears.

After you finish

The next step in the process of adding a new subnet is configuring SmartConnect settings, which is optional. If you do not wish to configure SmartConnect settings, the next step is adding network interface members to the new subnet.

Configure SmartConnect settings for a new subnet

You can configure subnet connection balancing for a cluster's external network with the default SmartConnect Basic feature of OneFS. You can configure advanced settings if you activate a SmartConnect Advanced license.

Before you begin

You must specify basic subnet settings and add at least one IP address pool range to the new subnet by completing the previous subnet wizard pages.

You can configure SmartConnect as an optional module to balance client connections on the external network of your cluster. A SmartConnect Advanced license must be active for certain options. An active SmartConnect Advanced license adds additional advanced balancing policies to evenly distribute CPU usage, client connections, or throughput. An active license also lets you define IP address pools to support multiple DNS zones in a subnet. In addition, SmartConnect supports IP failover, also known as NFS failover. In contrast, with SmartConnect Basic you can only set a round robin balancing policy.

Note

SmartConnect requires that you add a new name server (NS) record to the existing authoritative DNS zone that contains the cluster and that you delegate the SmartConnect zone as a fully qualified domain name (FQDN).

Procedure

1. In the **Step 3 of 4 — SmartConnect Settings** dialog box, type a **Zone name** for the SmartConnect zone that this IP address pool represents. The zone name must be unique among the pools served by the SmartConnect service subnet specified in Step 3 below.
2. In the **Connection policy** list, select the type of connection balancing policy set by the IP address pool of this subnet. The connection balancing policy determines how SmartConnect distributes incoming DNS requests across the members of an IP address pool.

Option	Description
Round Robin	Selects the next available node on a rotating basis, and is the default policy if no other policy is selected.
Connection Count	Determines the number of open TCP connections on each available node to optimize the cluster usage.
Network Throughput	Sets the overall average throughput volume on each available node to optimize the cluster usage.
CPU Usage	Examines average CPU utilization on each available node to optimize the cluster usage.

3. In the **SmartConnect service subnet** list, select the name of the external network subnet whose SmartConnect service will answer DNS requests on behalf of the IP address pool. A pool can have only one SmartConnect service answering DNS requests. If this option is left blank, the IP address pool the subnet belongs to is excluded when SmartConnect answers incoming DNS requests for the cluster.

Note

If you have activated a SmartConnect Advanced license, complete the following steps for the options in the **SmartConnect Advanced** section of this wizard page.

4. In the **IP allocation method** list, select the method by which IP addresses are assigned to the member interfaces for this IP address pool:

Option	Description
Static	Select this IP allocation method to assign IP addresses when member interfaces are added to the IP pool. As members are added to the pool, this method allocates the next unused IP address from the pool to each new member. After an IP address is allocated, the pool member keeps the address indefinitely unless one of the following items is true: • The member interface is removed from the network pool. • The member node is removed from the cluster. • The member interface is moved to another IP address pool.
Dynamic	Select this IP allocation method to ensure that all IP addresses in the IP address pool are assigned to member interfaces, which allows clients to connect to any IP addresses in the pool and be guaranteed a response. If a node or an interface becomes unavailable, their IP addresses are automatically moved to other available member interfaces in the pool.

If you select the dynamic IP allocation method, you can specify the SmartConnect **Rebalance policy** and the **IP failover policy in the next two steps.**

5. Select the type of SmartConnect **Rebalance policy** to redistribute IP addresses. IP address redistribution occurs when node interface members in an IP address pool become available. These options can only be selected if the IP allocation method is set to **Dynamic**.

Option	Description
Automatic Failback (default)	Automatically redistributes IP addresses. The automatic rebalance is triggered by a change to one of the following items. • Cluster membership. • Cluster external network configuration. • A member network interface.
Manual Failback	Does not redistribute IP addresses until you manually issue a rebalance command through the command-line interface.

6. The **IP failover policy**—also known as NFS failover—determines how to redistribute the IP addresses among remaining members of an IP address pool when one or more members are unavailable. In order to enable IP failover, set the **IP allocation method** to **Dynamic**, and then select an **IP failover policy**:

Option	Description
Round Robin	Selects the next available node on a rotating basis, and is the default policy if no other policy is selected.
Connection Count	Determines the number of open TCP connections on each available node to optimize the cluster usage.
Network Throughput	Sets the overall average throughput volume on each available node to optimize the cluster usage.
CPU Usage	Examines average CPU utilization on each available node to optimize the cluster usage.

7. Click **Next** to store the changes that you made to this wizard page.

   The **Step 4 of 4 — IP Address Pool members** dialog box appears.

After you finish

The next step in the process of adding a new subnet is adding network interface members.

Select interface members for a new subnet

You can select which network interfaces are in the IP address pool that belongs to the external network subnet.

Before you begin

You must specify basic subnet settings and add at least one IP address pool range to the new subnet by completing previous subnet wizard pages.

Procedure

1. In the **Step 4 of 4 — IP Address Pool Members** dialog box, select which **Available interfaces** on which nodes you want to assign to the current IP address pool, and then click the **right arrow** button to move them to the **Interfaces in current pool**.

   Alternatively, drag and drop the selected interfaces between the **Available interfaces** table and the **Interfaces in current pool** table.

   Selecting an available interface for a node that has a **Type** designated **Aggregation** bonds together the external interfaces for the selected node.

   In the case of aggregated links, choose the aggregation mode that corresponds to the switch settings from the **Aggregation mode** drop-down.

   Note

   Configuring link aggregation requires advanced knowledge of how to configure network switches. Consult your network switch documentation before configuring your cluster for link aggregation.

2. When you have finished assigning external network interfaces to the IP address pool, click **Submit**.

   The external subnet settings you configured by using the Subnet wizard appear on the **Edit Subnet** page.

Managing external network subnets

You can configure subnets on an external network to manage connections between the EMC Isilon cluster and client computers.

Modify external subnet settings

You can modify the subnet for the external network.

Note

Modifying an external network subnet that is in use can disable access to the cluster and the web administration interface. OneFS displays a warning if deleting a subnet will terminate communication between the cluster and the web administration interface.

Procedure

1. Click **Cluster Management** › **Network Configuration**.
2. In the **External Network Settings** area, click the name of the subnet you want to modify.
3. In the **Settings** area, click **Edit.**
4. Modify the **Basic** subnet settings as needed.

Option	Description
Description	A descriptive comment that can be up to 128 characters.
Netmask	The subnet mask for the network interface. This field appears only for IPv4 subnets.
MTU	The maximum size of the transmission units the cluster uses in network communication. Any numerical value is allowed, but might not be compatible with your network. Common settings are 1500 (standard frames) and 9000 (jumbo frames).
Gateway address	The IP address of the gateway server through which the cluster communicates with systems outside of the subnet.
Gateway priority	The priority of the subnet's gateway for nodes that are assigned to more than one subnet. Only one default gateway can be configured on each Isilon node, but each subnet can have its own gateway. If a node belongs to more than one subnet, this option enables you to define the preferred default gateway. A value of 1 is the highest priority, with 10 being the lowest priority.
SmartConnect service IP	The IP address that receives incoming DNS requests from outside the cluster. SmartConnect responds to these DNS requests for each IP address pool according to the pool's client-connection policy. To use connection balance, at least one subnet must be configured with a SmartConnect service IP address.

5. (Optional) Modify the **Advanced** settings as needed.

Note

Configuring a virtual LAN requires advanced knowledge of network switches. Consult your network switch documentation before configuring your cluster for a VLAN. If you are not using a virtual LAN, leave the VLAN options disabled.

Option	Description
VLAN tagging	You can enable VLAN tagging. VLAN tagging allows a cluster to participate in multiple virtual networks. VLAN support provides security across subnets that is otherwise available only by purchasing additional network switches.
VLAN ID	If you enabled VLAN tagging, type a VLAN ID that corresponds to the ID number for the VLAN that is set on the switch, with a value from 1 to 4094.
Hardware load balancing IPs	You can enter the IP address for a hardware load balancing switch that uses Direct Server Return (DSR).

6. Click **Submit**.

Remove an external subnet

You can delete an external network subnet that you no longer need.

Deleting an external network subnet that is in use can prevent access to the cluster and the web administration interface. OneFS displays a warning if deleting a subnet will terminate communication between the cluster and the web administration interface.

Procedure

1. Click **Cluster Management** › **Network Configuration**.
2. In the **External Network Settings** area, click the name of the subnet you want to delete.

   The **Edit Subnet** page appears for the subnet you specified.
3. Click **Delete subnet**.

   A confirmation dialogue box appears.
4. Click **Yes** to delete the subnet.

   If the subnet you are deleting is used to communicate with the web administration interface, the confirmation message will contain an additional warning.

Create a static route

You can create a static route to connect to networks that are unavailable through the default routes.

You configure a static route on a per-pool basis. A static route can be configured only with the command-line interface and only with the IPv4 protocol.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Create a static route by running the following command: `isi networks modify pool --name` *<subnetname>:<poolname>* `--add-static-routes` *<subnet>/<netmask>-<gateway>*

The system displays output similar to the following example:

```
Modifying pool 'subnet0:pool0': Saving:
OK
        OK
```

3. To verify that the static route was created, run the following command: **`isi networks ls pools -v`**.

Remove a static route

You can remove static routes.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Remove a static route by running the following command: **`isi networks modify pool`** `--name` *<subnetname>*:*<poolname>* `--remove-static-routes` *<subnet>*/*<netmask>*-*<gateway>*

   The system displays output similar to the following example:

   ```
   Modifying pool 'subnet0:pool0':

   Saving:
                 OK
   ```

3. To ensure that the static route was created, run the following command: **`isi networks ls pools -v`**.

Enable or disable VLAN tagging

You can configure a cluster to participate in multiple virtual private networks, also known as virtual LANs or VLANs. You can also configure a VLAN when creating a subnet using the Subnet wizard.

Procedure

1. Click **Cluster Management** › **Network Configuration**.
2. In the **External Network Settings** area, click the name of the subnet that contains the IP address pool that you want to add interface members to.
3. In the **Settings** area for the subnet, click **Edit**.
4. In the **VLAN tagging** list, select **Enabled** or **Disabled**.

   If you select **Enabled**, proceed to the next step. If you select **Disabled**, proceed to Step 6.
5. In the **VLAN ID** field, type a number between 2 and 4094 that corresponds to the VLAN ID number set on the switch.
6. Click **Submit**.

Managing IP address pools

IP address pools allow you to manage the IP addresses clients use to connect to an EMC Isilon cluster. You can also add network interfaces to IP address pools. Each IP address pool is associated with a subnet.

Add an IP address pool

You can add an IP address pool to a external network subnet.

Procedure

1. Click **Cluster Management** › **Network Configuration**.
2. Click the name of the subnet to which you are adding an IP address pool.
3. In the **IP Address Pools** area, click **Add pool**.
4. Specify basic settings for the new IP address pool and click **Next**.
5. (Optional) Specify SmartConnect settings and click **Next**.
6. (Optional) Add network interface members and click **Submit**.

Modify an IP address pool

You can use the web interface to modify IP address pool settings.

Procedure

1. Click **Cluster Management** › **Network Configuration**.
2. Click the name of the subnet containing the pool you want to modify.
3. In the **Basic Settings** area, click **Edit** for the IP address pool you want to modify.
4. Modify the address pool settings and click **Submit**.

Delete an IP address pool

You can use the web interface to delete IP address pool settings.

Procedure

1. Click **Cluster Management** › **Network Configuration**.
2. Click the name of the subnet containing the pool you want to delete.
3. Click **Delete pool** by the pool you want to delete.

Modify a SmartConnect zone

You can modify the settings of a SmartConnect zone that you created for an external network subnet using the Subnet wizard.

Procedure

1. Click **Cluster Management** › **Network Configuration**.
2. Click the name of the external network subnet that contains the SmartConnect zone you want to modify.
3. In the **SmartConnect settings** area of the pool containing the SmartConnect settings you want to modify, click **Edit**.
4. Modify the **Zone name,** and then click **Submit**.

   The SmartConnect zone should be a fully-qualified domain name (FQDN).

Disable a SmartConnect zone

You can remove a SmartConnect zone from an external network subnet.

Procedure

1. Click **Cluster Management › Network Configuration**.
2. Click the name of the external network subnet that contains the SmartConnect zone you want to disable.
3. In the **SmartConnect settings** area for the pool containing the SmartConnect zone you want to delete, click **Edit**.
4. To disable the SmartConnect zone, delete the name of the SmartConnect zone from the **Zone name** field and leave the field blank.
5. Click **Submit** to disable the SmartConnect zone.

Configure IP failover

You can configure IP failover to reassign an IP address from an unavailable node to a functional node, which enables clients to continue communicating with the cluster, even after a node becomes unavailable.

Procedure

1. Click **Cluster Management › Network Configuration**
2. In the **External Network Settings** area, click the name of the subnet for which you want to set up IP failover.
3. Expand the area of the pool you want to modify and click **Edit** in the **SmartConnect Settings** area.
4. (Optional) In the **Zone name** field, enter a name for the zone, using no more than 128 characters.
5. In the **Connection Policy** list, select a balancing policy:

Option	Description
Round Robin	Selects the next available node on a rotating basis, and is the default state if no other policy is selected.
Connection Count	Determines the number of open TCP connections on each available node to optimize the cluster usage.
Network Throughput	Uses the overall average throughput volume on each available node to optimize the cluster usage.
CPU Usage	Examines average CPU utilization on each available node to optimize the cluster usage.

6. If you purchased a license for SmartConnect Advanced, you will also have access to the following lists:

   IP allocation method

   This setting determines how IP addresses are assigned to clients. Select either `Dynamic` or `Static`.

   Rebalance Policy

   This setting defines the client redirection policy for when a node becomes unavailable. The **IP allocation** list must be set to `Dynamic` in order for rebalance policy options to be selected.

IP failover policy

This setting defines the client redirection policy when an IP address becomes unavailable.

Allocate IP addresses to accommodate new nodes

You can expand capacity by adding new nodes to your Isilon cluster.

After the hardware installation is complete, you can allocate IP addresses for a new node on one of the cluster's existing external network subnets, and then add the node's external interfaces to the subnet's IP address pool.

You can also use network provisioning rules to automate the process of configuring the external network interfaces for new nodes when they are added to a cluster, although you may still need to allocate more IP addresses for the new nodes, depending on how many are already configured.

Procedure

1. Click **Cluster Management** › **Network Configuration**.
2. In the **External Network Settings** area, click the name of the subnet that contains the IP address pool that you want to allocate more IP addresses to in order to accommodate the new nodes.
3. In the **Basic settings** area, click **Edit**.
4. Click **New** to add a new IP address range using the **Low IP** and **High IP** fields. or click the respective value in either the **Low IP** or **High IP** columns and type a new beginning or ending IP address.
5. Click **Submit**.
6. In the **Pool members** area, click **Edit**.
7. In the **Available Interfaces** table, select one or more interfaces for the newly added node, and then click the **right arrow** button to move the interfaces into the **Interfaces in current pool** table.
8. Click **Submit** to assign the new node interfaces to the IP address pool.

Managing network interface members

You can assign nodes and network interfaces to specific IP address pools.

You can also aggregate network interfaces and specify the aggregation method.

Modify the interface members of a subnet

You can use the web interface to modify interface member settings.

Procedure

1. Click **Cluster Management** › **Network Configuration**.
2. In the **External Network Settings** area, click the subnet containing the interface members you want to modify.
3. Click **Edit** next to the **Pool members** area.

Remove interface members from an IP address pool

You can use the web interface to remove interface members from an IP address pool.

Procedure

1. Click **Cluster Management** › **Network Configuration**.
2. In the **External Network Settings** area, click the name of the subnet containing the IP address pool that you want to remove the interface member(s) for.
3. In the **Pool members** area, click **Edit**.
4. Remove a node's interface from the IP address pool by clicking the node's name in the **Interfaces in current pool** column, and then click the **left arrow** button. You can also drag and drop node interfaces between the **Available Interfaces** list and the **Interfaces in current pool**column.
5. When you have finished removing node interfaces from the IP address pool, click **Submit**.

Configure NIC aggregation

You can configure cluster IP address pools to use NIC aggregation.

Before you begin

You must enable NIC aggregation on the cluster before you can enable NIC aggregation on the switch. If the cluster is configured but the switch is not configured, then the cluster can continue to communicate. If the switch is configured, but the cluster is not configured, the cluster cannot communicate, and you are unable to configure the cluster for NIC aggregation.

This procedure describes how to configure network interface card (NIC) aggregation for an IP address pool belonging to an existing subnet. You can also configure NIC aggregation while configuring an external network subnet using the Subnet wizard.

Configuring NIC aggregation means that multiple, physical external network interfaces on a node are combined into a single logical interface. If a node has two external Gigabit Ethernet interfaces, both will be aggregated. On a node with both Gigabit and 10 Gigabit Ethernet interfaces, both types of interfaces can be aggregated, but only with interfaces of the same type. NIC aggregation cannot be used with mixed interface types.

An external interface for a node cannot be used by an IP address pool in both an aggregated configuration and an individual interface. You must remove the individual interface for a node from the **Interfaces in current pool** table before configuring an aggregated NIC. Otherwise, the web administration interface displays an error message when you click **Submit**.

Note

Configuring link aggregation requires advanced knowledge of network switches. Consult your network switch documentation before configuring your cluster for NIC aggregation.

Procedure

1. Click **Cluster Management** › **Network Configuration**.
2. In the **External Network Settings** area, click the name of the subnet that contains the IP address pool that you want to add aggregated interface members to.
3. In the **Pool members** area, click **Edit**.

   In the case of multiple IP address pools, expand the pool that you want to add the aggregated interfaces to, and then click **Edit** in the **Pool members** area.

4. In the **Available interfaces** table, click the aggregated interface for the node, which is indicated by a listing of AGGREGATION in the **Type** column.

   For example, if you want to aggregate the network interface card for Node 2 of the cluster, click the interface named **ext-agg, Node 2** under **Available interfaces**, and then click the right-arrow button to move the aggregated interface to the **Interfaces in current pool** table.

5. From the **Aggregation mode** drop-down, select the appropriate aggregation mode that corresponds to the network switch settings.

   Note

   Consult your network switch documentation for supported NIC aggregation modes.

   OneFS supports the following NIC aggregation modes:

Option	Description
Link Aggregation Control Protocol (LACP)	This method supports the IEEE 802.3ad Link Aggregation Control Protocol (LACP). Balances outgoing traffic across the interfaces based on hashed protocol header information that includes the source and destination address and the VLAN tag, if available. Also assembles interfaces of the same speed into groups called Link Aggregated Groups (LAGs) and balances traffic across the fastest LAGs. This option is the default mode for new pools.
Legacy Fast EtherChannel (FEC) mode	This method aggregates network interfaces through an older FEC driver recommended for OneFS 6.5 and earlier.
Etherchannel (FEC)	This method provides static balancing on aggregated interfaces through the Cisco Fast EtherChannel (FEC) driver, which is found on older Cisco switches. Capable of load balancing traffic across Fast Ethernet links. Allows multiple physical Fast Ethernet links to combine into one logical channel.
Active / Passive Failover	This method switches to the next active interface when the primary interface becomes unavailable. Manages traffic only through a primary interface. The second interface takes over the work of the first as soon as it detects an interruption in communication.
Round-Robin Tx	This method rotates connections through the nodes in a first-in, first-out sequence, handling all processes without priority. Balances outbound traffic across all active ports in the aggregated link and accepts inbound traffic on any port.

6. Click **Submit**.

NIC and LNI aggregation options

Network interface card (NIC) and logical network interface (LNI) mapping options can be configured for aggregation.

The following list provides guidelines for interpreting the aggregation options.

- Nodes support multiple network card configurations.
- LNI numbering corresponds to the physical positioning of the NIC ports as found on the back of the node. LNI mappings are numbered from left to right.
- Aggregated LNIs are listed in the order in which they are aggregated at the time they are created.
- NIC names correspond to the network interface name as shown in command-line interface tools such as `ifconfig` and `netstat`.

LNI	NIC	Aggregated LNI	Aggregated NIC	Aggregated NIC (Legacy FEC mode
ext-1 ext-2	em0 em1	ext-agg = ext-1 + ext-2	lagg0	fec0

ext-1
ext-2
ext-3
ext-4
	em2
	em3
	em0
	em1
		ext-agg = ext-1 + ext-2
		ext-agg-2 = ext-3 + ext-4
		ext-agg-3 = ext-3 + ext-4 + ext-1 + ext-2
			lagg0
			lagg1
			lagg2
				fec0
				fec1
				fec2

ext-1 ext-2 10gige-1 10gige-1	em0 em1 cxgb0 cxgb1	ext-agg = ext-1 + ext-2 10gige-agg-1 = 10gige-1 + 10gige-2	lagg0 lagg1	fec0 fec1

Remove an aggregated NIC from an IP address pool

You can remove an aggregated NIC configuration from an IP address pool if your network environment has changed. However, you must first replace the aggregated setting with single-NIC settings in order for the node to continue supporting network traffic.

Procedure

1. Click **Cluster Management** › **Network Configuration**.
2. In the **External Network Settings** area, click the name of the subnet that contains the IP address pol with the NIC aggregation settings you want to remove.
3. In the **Pool members** area, click **Edit**.
4. Select the name of the aggregated NIC for the node that you want to remove in the **Interfaces in current pool** table, and then click the **left arrow** button to move the name into the **Available interfaces** table.
5. Select one or more individual interfaces for the node in the **Available interfaces** table, and then click the **right arrow** button to move the interfaces into the **Interfaces in current pool** table.
6. When you have completed modifying the node interface settings, click **Submit**.

Move nodes between IP address pools

You can move nodes between IP address pools in the event of a network reconfiguration or installation of a new network switch.

The process of moving nodes between IP address pools involves creating a new IP address pool and then assigning it to the nodes so that they are temporarily servicing multiple subnets. After testing that the new IP address pool is working correctly, the old IP address pool can safely be deleted.

Procedure

1. Create a new IP address pool with the interfaces belonging to the nodes you want to move.
2. Verify that the new IP address pool functions properly by connecting to the nodes you want to move with IP addresses from the new pool.
3. Delete the old IP address pool.

Reassign a node to another external subnet

You can move a node interface to a different subnet.

Nodes can be reassigned to other subnets.

Procedure

1. Click **Cluster Management** › **Network Configuration**.
2. In the **External Settings** area, click the subnet containing the node that you want to modify.
3. In the **IP Address Pools** area, click **Edit** next to the **Pool members** area.
4. Reassign the interface members that you want to move by dragging and dropping them from one column to other, or by clicking on an interface member and using the **left arrow** and **right arrow** buttons.

Configure DNS settings

You can configure the domain name servers (DNS) and DNS search list to resolve host names for the EMC Isilon cluster.

Procedure

1. Click **Cluster Management** › **Networking Configuration**.
2. In the **DNS Settings**area, click **Edit**.
3. In the **Domain name server(s)** field, enter up to three domain name server IP addresses. You can specify domain name server addresses in IPv4 or IPv6 format.
4. In the **DNS search list** field, enter up to six DNS search suffixes. DNS search suffixes are appended to unqualified host names .
5. (Optional) In the **DNS resolver options** field, enter advanced DNS configuration options.

   Note

   DNS resolver options are described in the `/etc/resolv.conf` file. Setting DNS resolver options is not recommended. Most clusters do not need DNS resolver options and setting them may change how OneFS performs DNS lookups.

6. Click **Submit**.

Managing external client connections with SmartConnect

You can manage settings that determine how IP addresses are allocated to client connection requests.

With the Basic SmartConnect module, you can specify connection balancing and IP allocation policies. If you have activated a Advanced SmartConnect license, you can also specify settings that manage IP failover and rebalancing.

Configure client connection balancing

You can configure connection balancing for your cluster's external network connections with SmartConnect.

Before you begin

You must first enable SmartConnect by setting up a SmartConnect service address on the external network subnet that answers incoming DNS requests.

You might have already configured SmartConnect while setting up an external network subnet using the Subnet wizard. However, you can configure or modify connection balancing settings at any time as your networking requirements change.

Procedure

1. Click **Cluster Management** › **Network Configuration**.
2. In the **External Network Settings** area, click the link for the subnet that you want to configure for connection balancing.
3. In the **Settings** area, verify that the SmartConnect service IP was configured.

   If the **SmartConnect service IP** field reads `Not set`, click **Edit**, and then specify the IP address that DNS requests are directed to.
4. In the **SmartConnect settings** area , click **Edit**.
5. In the **Zone name** field, type a name for the SmartConnect zone that this IP address pool represents. The zone name must be unique among the pools served by the SmartConnect service subnet that is specified in Step 7 below.
6. In the **Connection policy** drop-down list, select the type of connection balancing policy that is configured for the IP address pool for this zone. The policy determines how SmartConnect distributes incoming DNS requests across the members of an IP address pool.

   Note

   Round robin is the only connection policy available if you have not activated a SmartConnect Advanced license.

Option	Description
Round Robin	Selects the next available node on a rotating basis, and is the default policy if no other policy is selected.
Connection Count	Determines the number of open TCP connections on each available node to optimize the cluster usage.

Option	Description
Network Throughput	Sets the overall average throughput volume on each available node to optimize the cluster usage.
CPU Usage	Examines average CPU utilization on each available node to optimize the cluster usage.

7. In the **SmartConnect service subnet** list, select the name of the external network subnet whose SmartConnect service answers DNS requests on behalf of the IP address pool. A pool can have only one SmartConnect service answering DNS requests. If this option is left blank, the IP address pool that the SmartConnect service belongs to is excluded when SmartConnect answers incoming DNS requests for the cluster.

   If you have activated a SmartConnect Advanced license, complete the following steps in the **SmartConnect Advanced** area.

8. In the **IP allocation method** list, select the method by which IP addresses are assigned to the member interfaces for this IP address pool.

Managing network interface provisioning rules

You can configure provisioning rules to automate the configuration of external network interfaces.

Provisioning rules specify how new nodes are configured when they are added to an EMC Isilon cluster

If the new node type matches the type defined in a rule, the new node's interface name is added to the subnet and the IP address pool specified in the rule.

For example, you can create a provisioning rule that configures new Isilon storage nodes, and another rule that configures new accelerator nodes.

OneFS automatically checks for multiple provisioning rules when new rules are added to ensure there are no conflicts.

Create a node provisioning rule

Configure one or more provisioning rules to automate the process of adding new nodes to your Isilon cluster. All Isilon nodes support provisioning rules.

Before you begin

External network subnets and IP address pools must be configured before creating node provisioning rules. You must also verify that the IP address pool included in the provisioning rule has sufficient IP addresses to accommodate the new node's client connections.

Procedure

1. Click **Cluster Management** › **Network Configuration**.
2. In the **Provisioning Rules** area, click **Add rule**.
3. In the **Name** field, type a unique name for the provisioning rule. The rule name can be a maximum of 32 characters and can include spaces or other punctuation.
4. (Optional) In the **Description** field, type a descriptive comment about the provisioning rule.

5. In the **If node type is** list, select the type of node to which you want to apply the rule:

Option	Description
Any	Apply the provisioning rule to all types of Isilon nodes that join the cluster.
Storage-i	Apply the provisioning rule only to Isilon i-Series storage nodes that join the cluster.
Accelerator-i	Apply the provisioning rule only to Isilon i-Series performance accelerator nodes that join the cluster.
Storage	Apply the provisioning rule only to Isilon storage nodes that join the cluster.
Accelerator	Apply the provisioning rule only to performance-accelerator nodes that join the cluster.
Backup-Accelerator	Apply the provisioning rule only to Isilon backup-accelerator nodes that join the cluster.

6. In the **then assign interface** list, assign one of the following interfaces to the external network subnet and IP address pool for the node specified in the rule:

Option	Description
ext-1	The first external Gigabit Ethernet interface on the cluster.
ext-2	The second external Gigabit Ethernet interface on the cluster.
ext-3	The third external Gigabit Ethernet interface on the cluster.
ext-4	The fourth external Gigabit Ethernet interface on the cluster.
ext-agg	The first and second external Gigabit Ethernet interfaces aggregated together.
ext-agg-2	The third and fourth external Gigabit Ethernet interfaces aggregated together.
ext-agg-3	The first four external Gigabit Ethernet interfaces aggregated together.
ext-agg-4	All six Gigabit Ethernet interfaces aggregated together.
10gige-1	The first external 10 Gigabit Ethernet interface on the cluster.
10gige-2	The second external 10 Gigabit Ethernet interface on the cluster.
10gige-agg-1	The first and second external 10 Gigabit Ethernet interfaces aggregated together.

7. In the **Subnet** list, select the external subnet that the new node will join.
8. In the **Pool** list, select the IP address pool of the subnet that should be used by the new node.
9. Click **Submit**.

Modify a node provisioning rule

You can modify node provisioning rules.

Procedure

1. Click **Cluster Configuration** › **Network Configuration**.
2. In the **Provisioning Rules** area, click the name of the rule you want to modify.
3. Modify the provisioning rule settings as needed.
4. Click **Submit**.

Delete a node provisioning rule

You can delete a provisioning rule that is no longer necessary.

Procedure

1. Click **Cluster Management** › **Network Configuration.**
2. In the **Provisioning Rules** area, click **Delete** next to the rule you want to delete.

   A confirmation dialog box appears.
3. Click **Yes** to delete the rule, or click **No** to keep the rule.

Managing routing options

You can control the direction of outgoing client traffic through source-based routing or static route configuration.

Enable or disable source-based routing

You can enable source-based routing to ensure that outgoing client traffic is routed to the gateway of the source IP address in the packet header. If you disable source-based routing, outgoing traffic is destination-based or it follows static routes.

Before you begin

Source-based routing rules are prioritized over static routes. You can check if there are static routes configured in any IP address pools by running the following command:

```
isi networks list pools -v
```

Source-based routing is enabled or disabled on the entire EMC Isilon cluster and supports only the IPv4 protocol.
Enabling and disabling source-based routing is only supported through the command-line interface.

Procedure

1. Enable source-based routing by running the following command:

   ```
   isi networks sbr enable
   ```

2. Disable source-based routing by running the following command:

   ```
   isi networks sbr disable
   ```

Add or remove a static route

You can configure static routes to direct outgoing client traffic through a specific gateway.

Before you begin

Source-based routing rules are prioritized over static routes. You can check if source-based routing is enabled on the cluster by running the following command:

```
isi networks
```

Configure a static route on a per-pool basis. Static routes support only the IPv4 protocol. You can only add or remove a static route through the command-line interface.

Procedure

1. (Optional) Identify the name of the IP address pool you want to modify for static routes by running the following command:

   ```
   isi networks list pools
   ```

2. Configure static routes on a pool by running the `isi networks modify pool` command.

   Specify the route in classless inter-domain routing (CIDR) notation format.

   The following command adds a static route to pool5 and assigns the route to all network interfaces that are members of the pool:

   ```
   isi networks modify pool subnet10:pool5 --add-static-
   routes=192.168.205.128/24-192.168.205.2
   ```

   The following command removes a static route from pool5:

   ```
   isi networks modify pool subnet10:pool5 --remove-static-
   routes=192.168.205.128/24-192.168.205.2
   ```

CHAPTER 22

Hadoop

This section contains the following topics:

Hadoop overview

Hadoop is an open-source platform that runs analytics on large sets of data across a distributed file system.

In a Hadoop implementation on an EMC Isilon cluster, OneFS acts as the distributed file system and HDFS is supported as a native protocol. Clients from a Hadoop cluster connect to the Isilon cluster through the HDFS protocol to manage and process data.

Hadoop support on the cluster requires you to activate an HDFS license. To obtain a license, contact your EMC Isilon sales representative.

Hadoop architecture

Hadoop consists of a compute layer and a storage layer.

In a typical Hadoop implementation, both layers exist on the same cluster.

Hadoop compute layer

MapReduce 2.0 (YARN) is the task processing engine of the Hadoop compute layer.

MapReduce runs a variety of jobs (also known as applications), or queries, on large sets of data and pulls information out. MapReduce relies on the following key components:

ResourceManager

Global authority that allocates resources (such as CPU, memory, disk, network) to NodeManagers, and schedules jobs based on their resource requirements.

ApplicationMaster

Per-job component that negotiates job resources from the ResourceManager and tracks job status.

NodeManager

Per-node component that launches jobs and monitors job resource consumption.

HDFS storage layer

The storage layer is known as the Hadoop distributed file system (HDFS).

The storage layer contains the data accessed and processed by the compute layer. HDFS relies on two key components:

DataNode

Node that stores data and serves read and write requests as directed by the NameNode component.

NameNode

Node that stores in-memory maps of every file, including information about which DataNode the file resides on and the location of the file on the DataNode.

A typical Hadoop implementation contains one NameNode that acts as a master and routes requests for data access to the proper DataNode.

How Hadoop is implemented on OneFS

In a Hadoop implementation on the EMC Isilon cluster, data is stored on OneFS. HDFS is supported as a protocol, which is used by Hadoop compute clients to access data.

A Hadoop implementation with OneFS differs from a typical Hadoop implementation in the following ways:

- The compute and storage layers are on separate clusters instead of the same cluster.
- Instead of storing data within a Hadoop distributed file system, the storage layer functionality is fulfilled by OneFS on an EMC Isilon cluster. Nodes on the Isilon cluster function as both a NameNode and a DataNode.
- The compute layer is established on a Hadoop compute cluster that is separate from the Isilon cluster. MapReduce and its components are installed on the Hadoop compute cluster only.
- Rather than a storage layer, HDFS is implemented on OneFS as a native, lightweight protocol layer between the Isilon cluster and the Hadoop compute cluster. Clients from the Hadoop compute cluster authenticate over HDFS to access data on the Isilon cluster.
- In addition to HDFS, clients from the Hadoop compute cluster can connect to the Isilon cluster over any protocol that OneFS supports such as NFS, SMB, FTP, and HTTP.
- Hadoop compute clients can connect to any node on the Isilon cluster that functions as a NameNode instead of being routed by a single NameNode.

Hadoop distributions supported by OneFS

You can run most of the common Hadoop distributions with the EMC Isilon cluster.

OneFS supports the following Hadoop distributions:

Hadoop distribution	Versions supported
Cloudera CDH	• 3 (Updates 2–5) • 4.2 • 5.0 • 5.1 • 5.2
Cloudera Manager	• 4.0 • 5.2
Greenplum GPHD	• 1.1 • 1.2
HAWQ	• 1.1.0.1
Hortonworks Data Platform	• 1.1.1–1.3.3 (non-GUI) • 2.1

Hadoop distribution	Versions supported
Pivotal HD	• 1.0.1 • 2.0
Apache Hadoop	• 0.20.203 • 0.20.205 • 1.0.0–1.0.3 • 1.2.1 • 2.0.x • 2.2–2.4

WebHDFS

OneFS supports access to HDFS data through WebHDFS client applications.

WebHDFS is a RESTful programming interface based on HTTP operations such as GET, PUT, POST, and DELETE that is available for creating client applications. WebHDFS client applications allow you to access HDFS data and perform HDFS operations through HTTP and HTTPS.

WebHDFS is supported by OneFS on a per-access zone basis and is enabled by default.

WebHDFS supports simple authentication and Kerberos authentication. If the HDFS authentication method for an access zone is set to `All`, OneFS uses simple authentication by default.

Note

To prevent unauthorized client access through simple authentication, disable WebHDFS in each access zone that should not support it.

Secure impersonation

Secure impersonation enables you to create proxy users that can impersonate other users to run Hadoop jobs.

You might configure secure impersonation if you use applications, such as Apache Oozie, to automatically schedule, manage, and run Hadoop jobs. For example, you can create an Oozie proxy user that securely impersonates a user called HadoopAdmin, which allows the Oozie user to request that Hadoop jobs be performed by the HadoopAdmin user.

You configure proxy users for secure impersonation on a per–zone basis, and users or groups of users that you assign as members to the proxy user must be from the same access zone. A member can be one or more of the following identity types:

- User specified by user name or UID
- Group of users specified by group name or GID
- User, group, machine, or account specified by SID
- Well-known user specified by name

If the proxy user does not present valid credentials or if a proxy user member does not exist on the cluster, access is denied. The proxy user can only access files and subdirectories located in the HDFS root directory of the access zone. It is recommended that you limit the members that the proxy user can impersonate to users that have access only to the data the proxy user needs.

Ambari agent

The Ambari client/server framework is a third-party tool that enables you to configure, manage, and monitor a Hadoop cluster through a browser-based interface. The OneFS Ambari agent allows you to monitor the status of HDFS services on the EMC Isilon cluster through the Ambari interface.

The Ambari agent is configured per access zone; you can configure the OneFS Ambari agent in any access zone that contains HDFS data. To start an Ambari agent in an access zone, you must specify the address of the external Ambari server and the address of a NameNode that acts as the point of contact for the access zone.

The external Ambari server receives communications from the OneFS Ambari agent. Once the Ambari agent assigned to the access zone registers with the Ambari server, the agent provides a heartbeat status at regular intervals. The OneFS Ambari agent does not provide metrics or alerts to the Ambari server. The external Ambari server must be specified by a resolvable hostname, FQDN, or IP address and must be assigned to an access zone.

The NameNode is the designated point of contact in an access zone that Hadoop services managed through the Ambari interface should connect through. For example, if you manage services such as YARN or Oozie through the Ambari interface, the services will connect to the access zone through the specified NameNode. The Ambari agent communicates the location of the designated NameNode to the Ambari server, and to the Ambari interface, the NameNode represents the access zone. If you change the designated NameNode address, the Ambari agent will inform the Ambari server. The NameNode must be a resolvable SmartConnect zone name or an IP address from the IP address pool associated with the access zone.

Note

The specified NameNode value maps to the NameNode, secondary NameNode, and DataNode components on the Ambari interface.

The OneFS Ambari agent is based on the Apache Ambari framework and is compatible with Ambari server versions 1.5.1.110 and 1.6.0.

Virtual HDFS racks

You can create a virtual HDFS rack of nodes on the EMC Isilon cluster to optimize performance and reduce latency when accessing HDFS data.

A virtual HDFS rack enables you to specify a pool of preferred HDFS nodes on the EMC Isilon cluster and specify an associated pool of Hadoop compute clients.

When a Hadoop compute client from the defined pool connects to the cluster, OneFS returns at least two IP addresses from the pool of preferred HDFS nodes.

Virtual HDFS racks allow you to fine-tune client connectivity by directing Hadoop compute clients to go through quicker, less-busy switches or to faster nodes, depending on your network topology.

HDFS implementation considerations

Implementing HDFS requires you to take other areas of OneFS into consideration to ensure successful connections and access to HDFS data.

HDFS directories and Hadoop user accounts

Before implementing Hadoop, ensure that the directories and user accounts that you will need for Hadoop are configured on the EMC Isilon cluster.

When you set up directories, files, accounts, and permissions, ensure that they have the correct permissions so that Hadoop clients and applications can access the directories and files. Directories and permissions will vary by Hadoop distribution, environment, requirements, and security policies.

You must also ensure that the user accounts that your Hadoop distribution requires are configured on the Isilon cluster on a per-zone basis. The user accounts that you need and the associated owner and group settings vary by distribution, requirements, and security policies. The profiles of the accounts, including UIDs and GIDS, on the Isilon cluster should match those of the accounts on your Hadoop compute clients.

OneFS must be able to look up a local Hadoop user by name. If there are no directory services, such as Active Directory or LDAP, that can perform a user lookup, you must create a local Hadoop user. If directory services are available, a local user account is not required.

HDFS settings in access zones

Some HDFS attributes must be configured for each access zone on the EMC Isilon cluster.

You configure one HDFS root directory for each access zone. When a Hadoop compute client connects to the cluster, the user can access all files and sub-directories in the specified root directory. Unlike NFS mounts or SMB shares, clients connecting to the cluster through HDFS cannot be given access to individual folders within the root directory. The default HDFS directory is `/ifs`.

If you have multiple Hadoop workflows that require separate sets of data, you can create multiple access zones and configure a unique HDFS root directory for each zone.

You configure the authentication method for each access zone. HDFS supports simple authentication, Kerberos authentication, or both. When a Hadoop compute client connects to an access zone on the Isilon cluster, the client must authenticate with the method specified for that access zone. By default, HDFS accepts both simple and Kerberos authentication.

Proxy users are configured on a per-zone basis. Members assigned to the proxy user must belong to the same access zone.

When HDFS is licensed, the HDFS service is enabled on the entire cluster. You cannot disable HDFS on a per-access zone basis. If you create multiple access zones, you must configure HDFS settings for each zone that you want Hadoop compute clients to access through HDFS.

HDFS and SmartConnect

You can set up a SmartConnect zone for connections from Hadoop compute clients.

SmartConnect is a module that specifies how the DNS server on the EMC Isilon cluster handles connection requests from clients.

Each SmartConnect zone represents a specific pool of IP addresses. When you associate a SmartConnect zone with an access zone, OneFS only allows Hadoop clients connecting through the IP addresses in the SmartConnect zone to reach the HDFS data in the access zone. A root HDFS directory is specified for each access zone. This configuration isolates data within access zones and allows you to restrict client access to the data.

A SmartConnect zone evenly distributes NameNode requests from Hadoop compute clients across the access zone. When a Hadoop compute client makes an initial DNS request to connect to the SmartConnect zone, the Hadoop client is routed to an Isilon node that serves as a NameNode. Subsequent requests from the Hadoop compute client go the same node. When a second Hadoop client makes a DNS request for the SmartConnect zone, SmartConnect balances the traffic and routes the client connection to a different node than that used by the previous Hadoop compute client.

If you create a SmartConnect zone, you must add a new name server (NS) record as a delegated domain to the authoritative DNS zone that contains the Isilon cluster. On the Hadoop compute cluster, you must add the name of the DNS entry of the SmartConnect zone to the `core-site.xml` file so that your Hadoop compute clients connect to a NameNode with the DNS name of the zone.

SmartConnect is discussed in further detail in the *Networking* section of this guide.

Implementing Hadoop with OneFS

To support Hadoop on the EMC Isilon cluster, you must configure HDFS on the Isilon cluster to communicate with a Hadoop cluster.

The process for configuring HDFS on the Isilon cluster is summarized in the following list:

- Activate a license for HDFS. When a license is activated, the HDFS service is enabled by default.
- Create directories on the cluster that will be set as HDFS root directories.
- Create a SmartConnect zone for balancing connections from Hadoop compute clients.
- Create local Hadoop users in access zones that do not have directory services such as Active Directory or LDAP.
- Set the HDFS root directory in each access zone that supports HDFS connections.
- Enable or disable WebHDFS in each access zone.
- Set an authentication method in each access zone that supports HDFS connections.
- Configure HDFS service settings on the cluster.
- Configure proxy users for secure impersonation.
- Configure virtual HDFS racks.

Managing the HDFS service

You can configure HDFS service settings on the EMC Isilon cluster to improve performance for HDFS workflows.

Configure HDFS service settings

You can configure HDFS service settings to improve performance for HDFS workflows.

HDFS service settings can be configured only through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and then log in.
2. Run the `isi hdfs settings modify` command.

   The following example command sets the block size to 1 GB:

   ```
   isi hdfs settings modify --default-block-size=1G
   ```

   You must specify the block size in bytes. Suffixes K, M, and G are allowed.

   The following example command sets the checksum type to crc32:

   ```
   isi hdfs settings modify --default-checksum-type=crc32
   ```

   The following example command sets the log level to CRIT:

   ```
   isi hdfs settings modify --server-log-level=CRIT
   ```

   The following example command sets the number of server threads to 32:

   ```
   isi hdfs settings modify --server-threads=32
   ```

HDFS service settings

HDFS service settings affect the performance of HDFS workflows.

You can configure the following HDFS service settings:

Setting	Description
Block size	The HDFS block size setting on the EMC cluster determines how the HDFS service returns data upon read requests from Hadoop compute client. You can modify the HDFS block size on the cluster to increase the block size from the default of 64 MB up to 128 MB. Increasing the block size enables the Isilon cluster nodes to read and write HDFS data in larger blocks and optimize performance for most use cases. The Hadoop cluster maintains a different block size that determines how a Hadoop compute client writes a block of file data to the Isilon cluster. The optimal block size depends on your data, how you process your data, and other factors. You can configure the block size on the Hadoop cluster in the `hdfs-site.xml` configuration file in the dfs.block.size property.
Checksum type	The HDFS service sends the checksum type to Hadoop compute clients, but it does not send any checksum data, regardless of the checksum type. The default checksum type is set to `None`. If you Hadoop distribution requires a checksum type other than `None` to the client, you can set the checksum type to `CRC32` or `CRC32C`.
Service threads	The HDFS service generates multiple threads to handle HDFS traffic from EMC Isilon nodes. By default, the service thread value is set to `auto`, which calculates the thread count by multiplying the number of cores on a node by eight and adding a minimum threshold of thirteen. It generates a maximum of 96 threads on a node. To support a large system of Hadoop compute clients, you might need to increase the number of threads. If you are distributing HDFS traffic across all of

Setting	Description
	the nodes in an Isilon cluster through a SmartConnect zone, the total number of HDFS service threads should equal at least half of the total number of maps and reduces on the Hadoop compute cluster. The maximum thread count is 256 per node.
Logging level	The HDFS service supports the following logging levels: • Emergency—panic conditions broadcast to all users • Alert—conditions that must be corrected immediately, such as a corrupt system database • Critical—critical conditions, such as a hard device error • Error—general errors • Notice—conditions that are not errors, but might require special handling • Information—information messages that do not require action • Debug—information typically useful only when debugging a program

View HDFS service settings

You can view configuration details for the HDFS service.

HDFS service settings can be viewed only through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and then log in.
2. Run the `isi hdfs settings view` command.

   The system displays output similar to the following example:

   ```
   Default Block Size:     64M
   Default Checksum Type:  none
   Server Log Level:       crit
   Server Threads:         auto
   ```

Enable or disable the HDFS service

The HDFS service, which is enabled by default after you activate an HDFS license, can be enabled or disabled by running the `isi services` command.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in by using the root user account.
2. At the command prompt, run the `isi service` command to enable or disable the HDFS service, isi_hdfs_d.
   - To enable the HDFS service, run the following command:

     `isi services isi_hdfs_d enable`
   - To disable the HDFS service, run the following command:

     `isi services isi_hdfs_d disable`

Managing HDFS access zone settings

Some HDFS attributes must be configured for each access zone on the EMC Isilon cluster.

You can specify values for the following HDFS attributes within each access zone:

- HDFS root directory
- Authentication method
- WebHDFS support
- Ambari agent settings

Supported HDFS authentication methods

The authentication method determines what credentials are required by OneFS to establish a Hadoop compute client connection.

An HDFS authentication method is specified for each access zone. OneFS supports the following authentication methods for HDFS:

Authentication method	Description
Simple only	Requires only a user name to establish client connections.
Kerberos only	Requires Kerberos credentials to establish client connections. **Note** You must configure Kerberos as an authentication provider on the EMC Isilon cluster, and you must modify the `core-site.xml` file on clients running Hadoop 2.2 and later.
All (default value)	Accepts both simple authentication and Kerberos credentials. If Kerberos settings and file modifications are not completed, client connections default to simple authentication. ⚠ CAUTION **To prevent unintended access through simple authentication, set the authentication method to `Kerberos only` to enforce client access through Kerberos.**

Set the HDFS authentication method in an access zone

You can configure the HDFS authentication method within each access zone on the EMC Isilon cluster.

Before you begin

If you want to Hadoop clients to connect to an access zone through Kerberos, a Kerberos authentication provider must be configured on the cluster.

HDFS access zone settings can only be configured through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.

2. (Optional) To identify the name of the access zone that you want to modify for HDFS, run the following command:

   ```
   isi zone zones list
   ```

3. Set the HDFS authentication method for the access zone by running the `isi zone zones modify <zone>` command, where *<zone>* is the name of the zone.

   The following command specifies that Hadoop compute clients connecting to zone3 must be identified through the simple authentication method:

   ```
   isi zone zones modify zone3 --hdfs-authentication=simple_only
   ```

   The following example command specifies that Hadoop compute clients connecting to zone3 must be identified through the Kerberos authentication method:

   ```
   isi zone zones modify zone3 --hdfs-authentication=kerberos_only
   ```

After you finish

To ensure that users can authenticate through Kerberos, you must modify the `core-site.xml` file on clients running Hadoop 2.2 and later.

Configure HDFS authentication properties on the Hadoop client

If you want clients running Hadoop 2.2 and later to connect to an access zone through Kerberos, you must make some modifications to the `core-site.xml` and `hdfs-site.xml` files on the Hadoop clients.

Before you begin

Kerberos must be set as the HDFS authentication method and a Kerberos authentication provider must be configured on the cluster.

Procedure

1. Go to the `$HADOOP_CONF` directory on your Hadoop client.
2. Open the `core-site.xml` file in a text editor.
3. Set the value of the hadoop.security.token.service.use_ip property to **`false`** as shown in the following example:

   ```
   <property>
     <name>hadoop.security.token.service.use_ip</name>
     <value>false</value>
   </property>
   ```

4. Save and close the `core-site.xml` file.
5. Open the `hdfs-site.xml` file in a text editor.
6. Set the value of the dfs.namenode.kerberos.principal.pattern property to the Kerberos realm as shown in the following example:

   ```
   <property>
     <name>dfs.namenode.kerberos.principal.pattern</name>
     <value>hdfs/*@storage.company.com</value>
   </property>
   ```

7. Save and close the `hdfs-site.xml` file.

Create a local Hadoop user

OneFS must be able to look up a local Hadoop user by name. If there are no directory services in an access zone that can perform a user lookup, you must create a local Hadoop user that maps to a user on a Hadoop compute client for that access zone. If directory services are available, a local user account is not required.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster.
2. Run the `isi auth users create` command.

   The following example command creates a user named hadoop-user1 assigned to a local authentication provider within the zone3 access zone:

   ```
   isi auth users create --name=hadoop-user1 --provider=local --
   zone=zone3
   ```

Set the HDFS root directory in an access zone

You configure one HDFS root directory for each access zone.

Before you begin

The directory structure you want to set as the root path should already exist on the OneFS file system.

HDFS access zone settings can only be configured through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the EMC Isilon cluster and log in.
2. (Optional) To identify the name of the zone you want to modify for HDFS, run the following command:

   ```
   isi zone zones list
   ```

3. Configure the HDFS root directory for this access zone by running the `isi zone zones modify <zone>` command, where *<zone>* is the name of the zone.

   The following command specifies that Hadoop compute clients connecting to zone3 are given access to the `/ifs/hadoop/` directory:

   ```
   isi zone zones modify zone3 --hdfs-root-directory=/ifs/hadoop
   ```

Enable or disable WebHDFS within an access zone

You can specify whether WebHDFS is supported per access zone.

HDFS access-zone settings can only be configured through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the EMC Isilon cluster and log in.
2. (Optional) To identify the name of the zone you want to modify for HDFS, run the following command:

   ```
   isi zone zones list
   ```

3. Enable or disable WebHDFS in the access zone by running the `isi zone zones modify` command.

   The following command enables WebHDFS in zone3:

   ```
   isi zone zones modify zone3 --webhdfs-enabled=yes
   ```

   The following command disables WebHDFS in zone3:

   ```
   isi zone zones modify zone3 --webhdfs-enabled=no
   ```

Configure Ambari agent settings

You can configure Ambari agent support in each access zone that contains HDFS data.

Ambari agent settings can only be configured through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. (Optional) To identify the name of the access zone that you want to configure for Ambari agent settings, run the following command:

   ```
   isi zone zones list
   ```

3. Run the `isi zone zones modify` command.

   The following example specifies company.ambari.server.com as the external Ambari server that receives communication from the Ambari agent running in the zone3 access zone:

   ```
   isi zone zones modify zone3 \
   --hdfs-ambari-server=company.ambari.server.com
   ```

   The following example designates the IP address 192.168.205.5 as the point of contact in the zone3 access zone for Hadoop services managed through the Ambari interface:

   ```
   isi zone zones modify zone3 --hdfs-ambari-namenode=192.168.205.5
   ```

Configuring secure impersonation

Configure and manage proxy users that can securely impersonate other users and groups.

Create a proxy user

You can create a proxy user that securely impersonates another user.

This procedure is available only through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Run the `isi hdfs proxyusers create` command.

   The following command designates hadoop-user23 in zone1 as a new proxy user:

   ```
   isi hdfs proxyusers create hadoop-user23 --zone=zone1
   ```

The following command designates hadoop-user23 in zone1 as a new proxy user and adds the group hadoop-users to the list of members that the proxy user can impersonate:

```
isi hdfs proxyusers create hadoop-user23 --zone=zone1 --add-
group=hadoop-users
```

The following command designates hadoop-user23 in zone1 as a new proxy user and adds UID 2155 to the list of members that the proxy user can impersonate:

```
isi hdfs proxyusers create hadoop-user23 --zone=zone1 --add-
UID=2155
```

Modify a proxy user

You can modify a proxy user that securely impersonates another user.

This procedure is available only through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Run the `isi hdfs proxyusers modify` command.

   The following command adds the well-known user local to, and removes the user whose UID is 2155 from, the list of members for proxy user hadoop-user23 in zone1:

   ```
   isi hdfs proxyusers modify hadoop-user23 --zone=zone1 --add-
   wellknown=local --remove-uid=2155
   ```

Delete a proxy user

You can delete a proxy user that securely impersonates another user.

This procedure is available only through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Run the `isi hdfs proxyusers delete` command.

   The following command deletes hadoop-user23 in zone1 from the list of proxy users:

   ```
   isi hdfs proxyusers delete hadoop-user23 --zone=zone1
   ```

List the members of a proxy user

You can display all groups of users and individual users, known as members, that can be impersonated by a specific proxy user.

This procedure is available only through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Run the `isi hdfs proxyusers members list` command.

   The following command displays a detailed list of the users and groups of users that are members of proxy user hadoop-user23 in zone1:

   ```
   isi hdfs proxyusers members list hadoop-user23 --zone=zone1 -v
   ```

The system displays output similar to the following example:

```
Type: user
Name: krb_user_005
  ID: UID:1004
-----------------------
Type: group
Name: krb_users
  ID: SID:S-1-22-2-1003
-----------------------
Type: wellknown
Name: LOCAL
  ID: SID:S-1-2-0
```

View proxy users

You can view all proxy users in a specific zone or you can view information for a specific proxy user.

This procedure is available only through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. To view a list of all proxy users configure in a specific access zone, run the `isi hdfs proxyusers list` command.

   The following command displays a list of all proxy users configured in zone1:

   ```
   isi hdfs proxyusers list --zone=zone1
   ```

   The system displays output similar to the following example:

   ```
   Name
   -------------
   hadoop-user23
   hadoop-user25
   hadoop-user28
   -------------
   Total: 3
   ```

3. To view the configuration details for a specific proxy user, run the `isi hdfs proxyusers view`command.

   The following command displays the configuration details for the hadoop-user23 proxy user in zone1:

   ```
   isi hdfs proxyusers view hadoop-user23 --zone=zone1
   ```

   The system displays output similar to the following example:

   ```
   Name: hadoop-user23
   Members: krb_users
            LOCAL
            krb_user_004
   ```

Managing virtual HDFS racks

You can manage virtual HDFS racks of nodes on the EMC Isilon cluster.

A virtual HDFS rack is a pool of nodes on the Isilon cluster associated with a pool of Hadoop compute clients. You can create, modify, and delete virtual racks.

Create a virtual HDFS rack

You can create a virtual HDFS rack on the EMC Isilon cluster.

Virtual HDFS racks can only be created through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Create a virtual HDFS rack by running the `isi hdfs racks create` command.

   A rack name begins with a forward slash—for example, /hdfs-rack2.

   The following command creates a rack named /hdfs-rack2:

   ```
   isi hdfs racks create /hdfs-rack2
   ```

   The following command creates a rack named hdfs-rack2, specifies 120.135.26.10-120.135.26.20 as the IP address range of Hadoop compute client associated with the rack, and specifies subnet0:pool0 as the pool of Isilon nodes assigned to the rack:

   ```
   isi hdfs racks create /hdfs-rack2 --client-ip-
   ranges=120.135.26.10-120.135.26.20 --ip-pools=subnet0:pool0
   ```

Modify a virtual HDFS rack

You can modify the settings of a virtual HDFS rack.

Virtual HDFS racks can only be modified through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. (Optional) To identify the name of the virtual HDFS rack you want to modify, run the following command:

   ```
   isi hdfs racks list
   ```

3. Modify the virtual HDFS rack by running the `isi hdfs racks modify` command.

   A rack name begins with a forward slash—for example, /hdfs-rack2.

   The following example command renames a rack named /hdfs-rack2 to /hdfs-rack5:

   ```
   isi hdfs racks modify /hdfs-rack2 --new-name=/hdfs-rack5
   ```

   The following example command adds 120.135.26.30-120.135.26.40 to the list of existing Hadoop compute client IP addresses on the rack named /hdfs-rack2:

   ```
   isi hdfs racks modify /hdfs-rack2 --add-client-ip-
   ranges=120.135.26.30-120.135.26.40
   ```

   In addition to adding a new range to the list of existing ranges, you can modify the client IP address ranges by replacing the current ranges, deleting a specific range or deleting all ranges.

The following example command replaces any existing IP pools with subnet1:pool1 and subnet2:pool2 on the rack named /hdfs-rack2:

```
isi hdfs racks modify /hdfs-rack2 --ip-
pools=subnet1:pool1,subnet2:pool2
```

In addition to replacing the list of existing pools with new pools, you can modify the IP pools by adding pools to list of current pools, deleting a specific pool or deleting all pools.

Delete a virtual HDFS rack

You can delete a virtual HDFS rack from an EMC Isilon cluster.

Virtual HDFS racks can only be deleted through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. (Optional) To identify the name of the virtual HDFS rack you want to delete, run the following command:

   ```
   isi hdfs racks list
   ```

3. Delete a virtual HDFS rack by running the `isi hdfs racks delete` command.

   A rack name begins with a forward slash—for example, /hdfs-rack2.

   The following command deletes the virtual HDFS rack named /hdfs-rack2:

   ```
   isi hdfs racks delete /hdfs-rack2
   ```

4. At the prompt, type **`yes`**.

View virtual HDFS racks

You can view information for all virtual HDFS racks or for a specific rack.

Virtual HDFS rack settings can only be viewed through the command-line interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. To view a list of all virtual HDFS racks configured on the cluster, run the `isi hdfs racks list` command.

   The system displays output similar to the following example:

   ```
   Name        Client IP Ranges        IP Pools
   -----------------------------------------------
   /hdfs-rack1 10.20.30.40-10.20.30.50 subnet0:pool0
   /hdfs-rack2 20.10.30.10-20.10.30.20 subnet1:pool1
   -----------------------------------------------
   Total: 2
   ```

   The following example command displays setting details for all virtual HDFS racks configured on the cluster:

   ```
   isi hdfs racks list -v
   ```

The system displays output similar to the following example:

```
Name: /hdfs-rack1
Client IP Ranges: 10.20.30.40-10.20.30.50
IP Pools: subnet0:pool0
-----------------------------------------
Name: /hdfs-rack2
Client IP Ranges: 20.10.30.10-20.10.30.20
IP Pools: subnet1:pool1
```

3. To view the setting details for a specific virtual HDFS rack, run the `isi hdfs racks view` command:

   Each rack name begins with a forward slash—for example **`/hdfs-rack2`**.

   The following example command displays setting details for the virtual HDFS rack named /hdfs-rack2:

   ```
   isi hdfs racks view /hdfs-rack2
   ```

   The system displays output similar to the following example:

   ```
   Name: /hdfs-rack2
   Client IP Ranges: 20.10.30.10-20.10.30.20
   IP Pools: subnet1:pool1
   ```

CHAPTER 23

Antivirus

This section contains the following topics:

Antivirus overview

You can scan the files you store on an Isilon cluster for computer viruses and other security threats by integrating with third-party scanning services through the Internet Content Adaptation Protocol (ICAP). OneFS sends files through ICAP to a server running third-party antivirus scanning software. These servers are referred to as ICAP servers. ICAP servers scan files for viruses.

After an ICAP server scans a file, it informs OneFS of whether the file is a threat. If a threat is detected, OneFS informs system administrators by creating an event, displaying near real-time summary information, and documenting the threat in an antivirus scan report. You can configure OneFS to request that ICAP servers attempt to repair infected files. You can also configure OneFS to protect users against potentially dangerous files by truncating or quarantining infected files.

Before OneFS sends a file to be scanned, it ensures that the scan is not redundant. If a file has already been scanned and has not been modified, OneFS will not send the file to be scanned unless the virus database on the ICAP server has been updated since the last scan.

Note

Antivirus scanning is available only if all nodes in the cluster are connected to the external network.

On-access scanning

You can configure OneFS to send files to be scanned before they are opened, after they are closed, or both. Sending files to be scanned after they are closed is faster but less secure. Sending files to be scanned before they are opened is slower but more secure.

If OneFS is configured to ensure that files are scanned after they are closed, when a user creates or modifies a file on the cluster, OneFS queues the file to be scanned. OneFS then sends the file to an ICAP server to be scanned when convenient. In this configuration, users can always access files without any delay. However, it is possible that after a user modifies or creates a file, a second user might access the file before the file is scanned. If a virus was introduced to the file from the first user, the second user will be able to access the infected file. Also, if an ICAP server is unable to scan a file, the file will still be accessible to users.

If OneFS ensures that files are scanned before they are opened, when a user attempts to download a file from the cluster, OneFS first sends the file to an ICAP server to be scanned. The file is not sent to the user until the scan is complete. Scanning files before they are opened is more secure than scanning files after they are closed, because users can access only scanned files. However, scanning files before they are opened requires users to wait for files to be scanned. You can also configure OneFS to deny access to files that cannot be scanned by an ICAP server, which can increase the delay. For example, if no ICAP servers are available, users will not be able to access any files until the ICAP servers become available again.

If you configure OneFS to ensure that files are scanned before they are opened, it is recommended that you also configure OneFS to ensure that files are scanned after they are closed. Scanning files as they are both opened and closed will not necessarily improve security, but it will usually improve data availability when compared to scanning files only when they are opened. If a user wants to access a file, the file may have already been scanned after the file was last modified, and will not need to be scanned again if the ICAP server database has not been updated since the last scan.

Antivirus policy scanning

You can create antivirus scanning policies that send files from a specified directory to be scanned. Antivirus policies can be run manually at any time, or configured to run according to a schedule.

Antivirus policies target a specific directory on the cluster. You can prevent an antivirus policy from sending certain files within the specified root directory based on the size, name, or extension of the file. Antivirus policies do not target snapshots. Only on-access scans include snapshots. Antivirus scans are handled by the OneFS job engine, and function the same as any system job.

Individual file scanning

You can send a specific file to an ICAP server to be scanned at any time.

If a virus is detected in a file but the ICAP server is unable to repair it, you can send the file to the ICAP server after the virus database had been updated, and the ICAP server might be able to repair the file. You can also scan individual files to test the connection between the cluster and ICAP servers.

Antivirus scan reports

OneFS generates reports about antivirus scans. Each time that an antivirus policy is run, OneFS generates a report for that policy. OneFS also generates a report every 24 hours that includes all on-access scans that occurred during the day.

Antivirus scan reports contain the following information:

- The time that the scan started.
- The time that the scan ended.
- The total number of files scanned.
- The total size of the files scanned.
- The total network traffic sent.
- The network throughput that was consumed by virus scanning.
- Whether the scan succeeded.
- The total number of infected files detected.
- The names of infected files.
- The threats associated with infected files.
- How OneFS responded to detected threats.

ICAP servers

The number of ICAP servers that are required to support an Isilon cluster depends on how virus scanning is configured, the amount of data a cluster processes, and the processing power of the ICAP servers.

If you intend to scan files exclusively through antivirus scan policies, it is recommended that you have a minimum of two ICAP servers per cluster. If you intend to scan files on access, it is recommended that you have at least one ICAP server for each node in the cluster.

If you configure more than one ICAP server for a cluster, it is important to ensure that the processing power of each ICAP server is relatively equal. OneFS distributes files to the ICAP servers on a rotating basis, regardless of the processing power of the ICAP servers. If one server is significantly more powerful than another, OneFS does not send more files to the more powerful server.

Supported ICAP servers

OneFS supports ICAP servers running the following antivirus scanning software:

- Symantec Scan Engine 5.2 and later.
- Trend Micro Interscan Web Security Suite 3.1 and later.
- Kaspersky Anti-Virus for Proxy Server 5.5 and later.
- McAfee VirusScan Enterprise 8.7 and later with VirusScan Enterprise for Storage 1.0 and later.

Anitvirus threat responses

You can configure the system to repair, quarantine, or truncate any files that the ICAP server detects viruses in.

OneFS and ICAP servers react in one or more of the following ways when threats are detected:

Alert

All threats that are detected cause an event to be generated in OneFS at the warning level, regardless of the threat response configuration.

Repair

The ICAP server attempts to repair the infected file before returning the file to OneFS.

Quarantine

OneFS quarantines the infected file. A quarantined file cannot be accessed by any user. However, a quarantined file can be removed from quarantine by the root user if the root user is connected to the cluster through secure shell (SSH).

If you backup your cluster through NDMP backup, quarantined files will remain quarantined when the files are restored. If you replicate quarantined files to another Isilon cluster, the quarantined files will continue to be quarantined on the target cluster. Quarantines operate independently of access control lists (ACLs).

Truncate

OneFS truncates the infected file. When a file is truncated, OneFS reduces the size of the file to zero bytes to render the file harmless.

You can configure OneFS and ICAP servers to react in one of the following ways when threats are detected:

Repair or quarantine

Attempts to repair infected files. If an ICAP server fails to repair a file, OneFS quarantines the file. If the ICAP server repairs the file successfully, OneFS sends the file to the user. Repair or quarantine can be useful if you want to protect users from accessing infected files while retaining all data on a cluster.

Repair or truncate

Attempts to repair infected files. If an ICAP server fails to repair a file, OneFS truncates the file. If the ICAP server repairs the file successfully, OneFS sends the file to the user. Repair or truncate can be useful if you do not care about retaining all data on your cluster, and you want to free storage space. However, data in infected files will be lost.

Alert only

Only generates an event for each infected file. It is recommended that you do not apply this setting.

Repair only

Attempts to repair infected files. Afterwards, OneFS sends the files to the user, whether or not the ICAP server repaired the files successfully. It is recommended that you do not apply this setting. If you only attempt to repair files, users will still be able to access infected files that cannot be repaired.

Quarantine

Quarantines all infected files. It is recommended that you do not apply this setting. If you quarantine files without attempting to repair them, you might deny access to infected files that could have been repaired.

Truncate

Truncates all infected files. It is recommended that you do not apply this setting. If you truncate files without attempting to repair them, you might delete data unnecessarily.

Configuring global antivirus settings

You can configure global antivirus settings that are applied to all antivirus scans by default.

Exclude files from antivirus scans

You can prevent files from being scanned by antivirus policies.

Procedure

1. Click **Data Protection › Antivirus › Settings**.
2. In the **File size restriction** area, specify whether to exclude files from being scanned based on size.
   - Click **Scan all files regardless of size**.
   - Click **Only scan files smaller than the maximum file size** and specify a maximum file size.
3. In the **Filename restrictions** area, specify whether to exclude files from being scanned based on file names and extensions.
   - Click **Scan all files**.
   - Click **Only scan files with the following extensions or filenames**.
   - Click **Scan all files except those with the following extensions or filenames**.
4. (Optional) If you chose to exclude files based on file names and extensions, specify which files will be selected.
   a. In the **Extensions** area, click **Edit list**, and specify extensions.

b. In the **Filenames** area, click **Edit list**, and specify filenames.

You can specify the following wild cards:

Wildcard	Description
`*`	Matches any string in place of the asterisk. For example, specifying `m*` would match movies and m123.
`[ ]`	Matches any characters contained in the brackets, or a range of characters separated by a dash. For example, specifying `b[aei]t` would match bat, bet, and bit. For example, specifying `1[4-7]2` would match 142, 152, 162, and 172. You can exclude characters within brackets by following the first bracket with an exclamation mark. For example, specifying `b[!ie]` would match bat but not bit or bet. You can match a bracket within a bracket if it is either the first or last character. For example, specifying `[[c]at` would match cat, and [at. You can match a dash within a bracket if it is either the first or last character. For example, specifying `car[-s]` would match cars, and car-.
`?`	Matches any character in place of the question mark. For example, specifying `t?p` would match tap, tip, and top.

5. Click **Submit**.

Configure on-access scanning settings

You can configure OneFS to automatically scan files as they are accessed by users. On-access scans operate independently of antivirus policies.

Procedure

1. Click **Data Protection** › **Antivirus** › **Settings**.
2. In the **On Access Scans** area, specify whether you want files to be scanned as they are accessed.
   - To require that all files be scanned before they are opened by a user, select **Scan files when they are opened**, and then specify whether you want to allow access to files that cannot be scanned.
   - To scan files after they are closed, select **Scan files when they are closed**.
3. In the **Directories to be scanned** area, specify the directories that you want to apply on-access settings to.

   If no directories are specified, on-access scanning settings are applied to all files. If you specify a directory, only files from the specified directories will be scanned as they are accessed.
4. Click **Submit**.

Configure antivirus threat response settings

You can configure how OneFS responds to detected threats.

Procedure

1. Click **Data Protection** › **Antivirus** › **Settings**.
2. In the **Action on detection** area, specify how you want OneFS to react to potentially infected files.

Configure antivirus report retention settings

You can configure how long OneFS retains antivirus reports before automatically deleting them.

Procedure

1. Click **Data Protection** › **Antivirus** › **Settings**.
2. In the **Reports** area, specify how long you want OneFS to keep reports for.

Enable or disable antivirus scanning

You can enable or disable all antivirus scanning. This procedure is available only through the web administration interface.

Procedure

1. Click **Data Protection** › **Antivirus** › **Summary**.
2. In the **Service** area, click **Enable** or **Disable**.

Managing ICAP servers

Before you can send files to be scanned on an ICAP server, you must configure OneFS to connect to the server. You can test, modify, and remove an ICAP server connection. You can also temporarily disconnect and reconnect to an ICAP server.

Add and connect to an ICAP server

You can add and connect to an ICAP server. After a server is added, OneFS can send files to the server to be scanned for viruses.

Procedure

1. Click **Data Protection** › **Antivirus** › **Summary**.
2. In the **ICAP Servers** area, click **Add server**.
3. In the **Add ICAP Server** dialog box, in the **ICAP URL** field, type the IP address of an ICAP server.
4. (Optional) In the **Description** field, type a description of this ICAP server.
5. Click **Submit**.

   The ICAP server is displayed in the **ICAP Servers** table.

Test an ICAP server connection

You can test the connection between the cluster and an ICAP server. This procedure is available only through the web administration interface.

Procedure

1. Click **Data Protection** › **Antivirus** › **Summary**.
2. In the **ICAP Servers** table, in the row for the ICAP server, click **Test connection**.

   If the connection test succeeds, the **Status** column displays a green icon. If the connection test fails, the **Status** column displays a red icon.

Modify ICAP connection settings

You can modify the IP address and optional description of ICAP server connections.

Procedure

1. Click **Data Protection** › **Antivirus** › **Summary**.
2. In the **ICAP Servers** table, in the row for an ICAP server, click **Edit**.
3. Modify settings, and then click **Submit**.

Temporarily disconnect from an ICAP server

If you want to prevent OneFS from sending files to an ICAP server, but want to retain the ICAP server connection settings, you can temporarily disconnect from the ICAP server.

Procedure

1. Click **Data Protection** › **Antivirus** › **Summary**.
2. In the **ICAP Servers** table, in the row for an ICAP server, click **Disable**.

Reconnect to an ICAP server

You can reconnect to an ICAP server that you have temporarily disconnected from.

Procedure

1. Click **Data Protection** › **Antivirus** › **Summary**.
2. In the **ICAP Servers** table, in the row for an ICAP server, click **Enable**.

Remove an ICAP server

You can permanently disconnect from the ICAP server.

Procedure

1. Click **Data Protection** › **Antivirus** › **Summary**.
2. In the **ICAP Servers** table, in the row for an ICAP server, click **Delete**.

Create an antivirus policy

You can create an antivirus policy that causes specific files to be scanned for viruses each time the policy is run.

Procedure

1. Click **Data Protection** › **Antivirus** › **Policies**.
2. Click **Add policy**.
3. In the **Name** field, type a name for the antivirus policy.
4. Click **Add directory** and select a directory that you want to scan.

   Optionally, repeat this step to specify multiple directories.

5. In the **Restrictions** area, specify whether you want to enforce the file size and name restrictions specified by the global antivirus settings.
   - Click **Enforce file size and filename restrictions**.
   - Click **Scan all files within the root directories**.
6. In the **Run policy** area, specify whether you want to run the policy according to a schedule or manually.

   Scheduled policies can also be run manually at any time.

Option	Description
Run the policy only manually.	Click **Manually**
Run the policy according to a schedule.	a. Click **Scheduled**. b. In the **Interval** area, specify on what days you want the policy to run. c. In the **Frequency** area, specify how often you want the policy to run on the specified days.

7. Click **Submit**.

Managing antivirus policies

You can modify and delete antivirus policies. You can also temporarily disable antivirus policies if you want to retain the policy but do not want to scan files.

Modify an antivirus policy

You can modify an antivirus policy.

Procedure

1. Click **Data Protection** › **Antivirus** › **Policy**.
2. In the **Policies** table, click the name of the antivirus policy that you want to modify.
3. Modify settings, and then click **Submit**.

Delete an antivirus policy

You can delete an antivirus policy.

Procedure

1. Click **Data Protection** › **Antivirus** › **Policies**.
2. In the **Policies** table, in the row for an antivirus policy, click **Delete**.

Enable or disable an antivirus policy

You can temporarily disable antivirus policies if you want to retain the policy but do not want to scan files.

Procedure

1. Click **Data Protection** › **Antivirus** › **Policies**.
2. In the **Policies** table, in the row for an antivirus policy, click **Enable** or **Disable**.

View antivirus policies

You can view antivirus policies.

Procedure

1. Click **Data Protection** › **Antivirus** › **Policies**.
2. In the **Policies** table, view antivirus policies.

Managing antivirus scans

You can scan multiple files for viruses by manually running an antivirus policy, or scan an individual file without an antivirus policy. You can also stop antivirus scans.

Scan a file

You can manually scan an individual file for viruses. This procedure is available only through the command-line interface (CLI).

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Run the `isi avscan manual` command.

   For example, the following command scans `/ifs/data/virus_file`:

   ```
   isi avscan manual /ifs/data/virus_file
   ```

Manually run an antivirus policy

You can manually run an antivirus policy at any time. This procedure is available only through the web administration interface.

Procedure

1. Click **Data Protection** › **Antivirus** › **Policies**.
2. In the **Policies** table, in the row for a policy, click **Start**.

Stop a running antivirus scan

You can stop a running antivirus scan. This procedure is available only through the web administration interface.

Procedure

1. Click **Data Protection** › **Antivirus** › **Summary**.
2. In the **Currently Running** table, in the row for an antivirus scan, click **Cancel**.

Managing antivirus threats

You can repair, quarantine, or truncate files in which threats are detected. If you think that a quarantined file is no longer a threat, you can rescan the file or remove the file from quarantine.

Manually quarantine a file

You can quarantine a file to prevent the file from being accessed by users.

Procedure

1. Click **Data Protection** › **Antivirus** › **Detected Threats**.
2. In the **Detected Threats** table, in the row of a file, click **Quarantine**.

Rescan a file

You can rescan the file for viruses if, for example, you believe that a file is no longer a threat.

Procedure

1. Click **Data Protection** › **Antivirus** › **Detected Threats**.
2. In the **Detected Threats** table, in the row of a file, click **Rescan**.

Remove a file from quarantine

You can remove a file from quarantine if, for example, you believe that the file is no longer a threat.

Procedure

1. Click **Data Protection** › **Antivirus** › **Detected Threats**.
2. In the **Detected Threats** table, in the row of a file, click **Restore**.

Manually truncate a file

If a threat is detected in a file, and the file is irreparable and no longer needed, you can truncate the file.

Procedure

1. Click **Data Protection** › **Antivirus** › **Detected Threats**.
2. In the **Detected Threats** table, in the row for a file, click **Truncate**.

   The **Confirm** dialog box appears.
3. Click **Yes**.

View threats

You can view files that have been identified as threats by an ICAP server.

Procedure

1. Click **Data Protection** › **Antivirus** › **Detected Threats**.
2. In the **Detected Threats** table, view potentially infected files.

Antivirus threat information

You can view information about the antivirus threats that are reported by an ICAP server.

Status
: The color of the icon indicates the status of the potentially infected file.

  Red
  : OneFS did not take any action on the file.

  Orange
  : OneFS truncated the file.

  Yellow
  : OneFS quarantined the file.

Threat
: Displays the name of the detected threat as it is recognized by the ICAP server.

Filename
: Displays the name of the potentially infected file.

Directory
: Displays the directory in which the file is located.

Remediation
: Indicates how OneFS responded to the file when the threat was detected. If OneFS did not quarantine or truncate the file, `Infected` appears.

Detected
: Displays the time that the threat was detected.

Policy
: Displays the name of the antivirus policy that caused the threat to be detected. If the threat was detected as a result of a manual antivirus scan of an individual file, `Manual scan` appears.

Currently
: Displays the current state of the file.

File size
: Displays the size of the file in bytes. Truncated files display a size of zero bytes.

Managing antivirus reports

In addition to viewing antivirus reports through the web administration interface, you can export reports to a comma-separated values (CSV) file. You can also view events that are related to antivirus activity.

Export an antivirus report

You can export an antivirus report to a comma separated values (CSV) file.

Procedure

1. Click **Data Protection** › **Antivirus** › **Reports**.
2. In the **Reports** table, in the row for a report, click **Export**.
3. Save the CSV file.

View antivirus reports

You can view antivirus reports.

Procedure

1. Click **Data Protection** › **Antivirus** › **Reports**.
2. In the **Reports** table, in the row for a report, click **View Details**.

View antivirus events

You can view events that relate to antivirus activity.

Procedure

1. Click **Dashboard** › **Events** › **Event History**.
2. In the **Event History** table, view all events.

   All events related to antivirus scans are classified as warnings. The following events are related to antivirus activities:

   `Anti-Virus scan found threats`
   A threat was detected by an antivirus scan. These events refer to specific reports on the Antivirus Reports page but do not provide threat details.

   `No ICAP Servers available`
   OneFS is unable to communicate with any ICAP servers.

   `ICAP Server Unresponsive or Invalid`
   OneFS is unable to communicate with an ICAP server.

CHAPTER 24

VMware integration

This section contains the following topics:

VMware integration overview

OneFS integrates with VMware infrastructures, including vSphere, vCenter, and ESXi. VMware integration enables you to view information about and interact with Isilon clusters through VMware applications.

OneFS interacts with VMware infrastructures through VMware vSphere API for Storage Awareness (VASA) and VMware vSphere API for Array Integration (VAAI).

OneFS integrates with VMware vCenter through the Isilon for vCenter plug-in. The Isilon for vCenter plug-in enables you to locally backup and restore virtual machines on an Isilon cluster. For more information about Isilon for vCenter, see the following documents:

- *Isilon for vCenter Release Notes*
- *Isilon for vCenter Installation Guide*
- *Isilon for vCenter User Guide*

VAAI

OneFS uses VMware vSphere API for Array Integration (VAAI) to support offloading specific virtual machine storage and management operations from VMware ESXi hypervisors to an Isilon cluster.

VAAI support enables you to accelerate the process of creating virtual machines and virtual disks. For OneFS to interact with your vSphere environment through VAAI, your VMware environment must include ESXi 5.0 or later hypervisors.

If you enable VAAI capabilities for an Isilon cluster, when you clone a virtual machine residing on the cluster through VMware, OneFS clones the files related to that virtual machine.

To enable OneFS to use VMware vSphere API for Array Integration (VAAI), you must install the VAAI NAS plug-in for Isilon on the ESXi server. For more information on the VAAI NAS plug-in for Isilon, see the *VAAI NAS plug-in for Isilon Release Notes*.

VASA

OneFS communicates with VMware vSphere through VMware vSphere API for Storage Awareness (VASA).

VASA support enables you to view information about Isilon clusters through vSphere, including Isilon-specific alarms in vCenter. VASA support also enables you to integrate with VMware profile driven storage by providing storage capabilities for Isilon clusters in vCenter. For OneFS to communicate with vSphere through VASA, your VMware environment must include ESXi 5.0 or later hypervisors.

Isilon VASA alarms

If the VASA service is enabled on an Isilon cluster and the cluster is added as a VMware vSphere API for Storage Awareness (VASA) vendor provider in vCenter, OneFS generates alarms in vSphere.

The following table describes the alarm that OneFS generates:

Alarm name	Description
Thin-provisioned LUN capacity exceeded	There is not enough available space on the cluster to allocate space for writing data to thinly provisioned LUNs. If this condition persists, you will not be able to write to the virtual machine on this cluster. To resolve this issue, you must free storage space on the cluster.

VASA storage capabilities

OneFS integrates with VMware vCenter through VMware vSphere API for Storage Awareness (VASA) to display storage capabilities of Isilon clusters in vCenter.

The following storage capabilities are displayed through vCenter:

Archive
The Isilon cluster is composed of Isilon NL-Series nodes. The cluster is configured for maximum capacity.

Performance
The Isilon cluster is composed of Isilon i-Series, Isilon X-Series, or Isilon S-Series nodes. The cluster is configured for maximum performance.

Note

If a node type supports SSDs but none are installed, the cluster is recognized as a capacity cluster.

Capacity
The Isilon cluster is composed of Isilon X-Series nodes that do not contain SSDs. The cluster is configured for a balance between performance and capacity.

Hybrid
The Isilon cluster is composed of nodes associated with two or more storage capabilities. For example, if the cluster contained both Isilon S-Series and NL-Series nodes, the storage capability of the cluster is displayed as `Hybrid`.

Configuring VASA support

To enable VMware vSphere API for Storage Awareness (VASA) support for a cluster, you must enable the VASA daemon on the cluster, download the Isilon vendor provider certificate and add the Isilon vendor provider in vCenter.

Enable VASA

You must enable an Isilon cluster to communicate with VMware vSphere API for Storage Awareness (VASA) by enabling the VASA daemon.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Enable VASA by running the following command:

```
isi services isi_vasa_d enable
```

Download the Isilon vendor provider certificate

To add an Isilon cluster VASA vendor provider in VMware vCenter, you must use a vendor provider certificate.

Procedure

1. In a supported web browser, connect to an Isilon cluster at `https://<IPAddress>`, where *<IPAddress>* is the IP address of the Isilon cluster.
2. Add a security exception and view the security certificate to make sure that it is current.
3. Download the security certificate and save it to a location on your machine.

   For more information about exporting a security certificate, see the documentation of your browser.

   Note

   Record the location of where you saved the certificate. You will need this file path when adding the vendor provider in vCenter.

Add the Isilon vendor provider

You must add an Isilon cluster as a vendor provider in VMware vCenter before you can view information about the storage capabilities of the cluster through vCenter.

Before you begin

Download a vendor provider certificate.

Procedure

1. In vCenter, navigate to the **Add Vendor Provider** window.
2. Fill out the following fields in the **Add Vendor Provider** window:

   Name
   : Type a name for this VASA provider. Specify as any string. For example, type **`EMC Isilon Systems`**.

   URL
   : Type **`https://<IPAddress>:8081/vasaprovider`**, where *<IPAddress>* is the IP address of a node in the Isilon cluster.

   Login
   : Type **`root`**.

   Password
   : Type the password of the root user.

   Certificate location
   : Type the file path of the vendor provider certificate for this cluster.

3. Select the **Use Vendor Provider Certificate** box.
4. Click **OK**.

Disable or re-enable VASA

You can disable or re-enable an Isilon cluster to communicate with VMware vSphere through VMware vSphere API for Storage Awareness (VASA).

To disable support for VASA, you must disable both the VASA daemon and the Isilon web administration interface. You will not be able to administer the cluster through an internet browser while the web interface is disabled. To re-enable support for VASA, you must enable both the VASA daemon and the web interface.

Procedure

1. Open a secure shell (SSH) connection to any node in the cluster and log in.
2. Disable or enable the web interface by running one of the following commands:
   - **`isi services apache2 disable`**
   - **`isi services apache2 enable`**
3. Disable or enable the VASA daemon by running one of the following commands:
   - **`isi services isi_vasa_d disable`**
   - **`isi services isi_vasa_d enable`**

Troubleshooting VASA storage display failures

If you are unable to view information about Isilon clusters through vSphere, follow the troubleshooting tips given below to fix the issue.

- Verify that the vendor provider certificate is current and has not expired.
- Verify that the Isilon cluster is able to communicate with VASA through the VASA daemon. If the VASA daemon is disabled, run the following command to enable it:

  ```
  isi services isi_vasa_d enable
  ```
- Verify that the date and time on the cluster is set correctly.
- Verify that data has been synchronized properly from vCenter.

CHAPTER 25

File System Explorer

This section contains the following topics:

File System Explorer overview

The File System Explorer is a web-based interface that enables you to manage the content stored on the cluster. You can use the File System Explorer to navigate the Isilon file system (`/ifs`), add directories, and manage file and directory properties including data protection, I/O optimization, and UNIX permissions. The File System Explorer is available only if you are logged in through the root user account.

Isilon file system directory permissions are initially set to allow full access for all users. Any user can delete any file, regardless of the permissions on the individual file. Depending on your environment, you want to establish permission restrictions through the File System Explorer.

You can view and configure file and directory properties from within Windows clients that are connected to the cluster. However, because Windows and UNIX permissions differ from one another, you must be careful not to make any unwanted changes that affect file and directory access.

Note

The File System Explorer displays up to 1000 files in a directory. If more than 1000 files exist within a directory, the files are displayed without additional information, such as file size and last modified date.

Browse the file system

You can browse the Isilon file system (`/ifs`) through the File System Explorer.

Procedure

1. Navigate to **File System Management › File System Explorer**.
2. View files and directories.
   - You can expand and collapse directories in the **Directories** pane.
   - The contents of the selected directory are displayed in the right pane. You can view the contents of another directory by clicking the directory in the **Directories** pane.

File System Explorer icons

The following icons appear in the File System Explorer:

Folder
: Represents a directory.

Folder with a red lock
: Represents a directory that contains a SmartLock directory.

Red lock
: Represents a SmartLock directory.

Yellow lock
: Represents a directory that is contained within a SmartLock directory.

Sheet of paper
: Represents a file

Sheet of paper with a red lock
Represents a file that has been committed to a WORM state.

Create a directory

You can create a directory under `/ifs` through the File System Explorer.

Procedure

1. Navigate to **File System Management** › **File System Explorer**.
2. In the **Directories** pane, specify where you want to create the directory.
3. Click **Add Directory**.
4. In the **New Directory Properties** dialog box, in the **Directory name** field, type a name for the directory.
5. From the **User** list, select the owner of the directory.
6. From the **Group** list, select the group for the directory.
7. From the **Permissions** table, specify the basic permissions for the directory.
8. Click **Submit**.

Modify file and directory properties

You can modify the data protection, I/O optimization, and UNIX permission properties of files and directories through the File System Explorer.

Procedure

1. Navigate to **File System Management** › **File System Explorer**.
2. In the **Directories** pane, click the directory that contains the file or directory that you want to modify permissions for.
3. In the right pane, in the row of the file or directory you want to modify permissions for, click **Properties**.
4. In the **Properties** dialog box, specify the properties of the file or directory.
5. Click **Submit**.

View file and directory properties

You can view the data protection, I/O optimization, and UNIX permission properties of files and directories through the File System Explorer.

Procedure

1. Navigate to **File System Management** › **File System Explorer**.
2. In the **Directories** pane, click the directory that contains the file or directory that you want to view permissions for.
3. In the right pane, in the row of the file or directory you want to view permissions for, click **Properties**.
4. In the **Properties** dialog box, view the properties of the file or directory.

File and directory properties

Each file and directory is assigned specific data protection, I/O optimization, and UNIX permission properties that you can view through the File System Explorer.

The following properties are displayed in the **Properties** dialog box of the File System Explorer:

Protection Settings

Settings management
: Specifies whether protection settings are managed manually or by SmartPools. If you modify either or both protection settings, this property automatically refreshes to `Manually managed`. If you specify **Managed by SmartPools**, the protection settings will automatically refresh to match the SmartPools specifications the next time the SmartPools job is run.

Disk pool
: The disk pool whose requested protection is applied if SmartPools is configured to manage protection settings. This property is available only if SmartPools is licensed and enabled on the cluster.

SSD
: The SSD strategy that will be used for user data and metadata if solid-state drives (SSDs) are available. The following SSD strategies are available:

Metadata acceleration
: OneFS creates a mirror backup of file metadata on an SSD and writes the rest of the metadata plus all user data to hard disk drives (HDDs). Depending on the global namespace acceleration setting, the SSD mirror might be an extra mirror in addition to the number required to satisfy the protection level.

Avoid SSDs
: OneFS does not write data or metadata to SSDs. OneFS writes all data and metadata to HDDs only.

Data on SSDs
: Similar to metadata acceleration, OneFS creates a mirror backup of file metadata on an SSD and writes the rest of the metadata plus all user data to hard disk drives. However, OneFS also writes one copy of the file user data (if mirrored) or all of the data (if not mirrored) to SSDs. All SSD blocks reside on the file target pool if there is adequate space available, regardless of whether global namespace acceleration is enabled. OneFS does not create additional mirrors beyond the normal protection level.

Actual protection
: The FlexProtect or data-mirroring requested protection for this file or directory. If SmartPools is licensed and enabled on the cluster, the default requested protection for files and directories is inherited from the specified disk pool.

I/O Optimization Settings

Settings Management

Specifies whether I/O Optimization Settings are managed manually or by SmartPools. If you modify either or both I/O optimization settings, this property automatically refreshes to `Manually managed`. If you specify **Managed by SmartPools,** the I/O optimization settings values will automatically refresh to match the SmartPools specifications the next time the SmartPools job is run.

SmartCache

Specifies whether write caching with SmartCache is enabled for this file or directory.

Data access pattern

The optimization settings for accessing data. The following data access patterns are available:

Concurrency

File or directory is optimized to support many clients simultaneously.

Streaming

File or directory is optimized for high-speed streaming of a single file. For example, this pattern can be useful if a single client needs to read very quickly from a single file.

Random

File or directory is optimized for unpredictable access.

The default data access pattern of iSCSI LUNs is the random access pattern. The default data access pattern of other files and directories is the concurrent access pattern.

UNIX Permissions

User

The owner of the file or directory.

Group

The group of the file or directory.

Permissions

The basic permissions for the file or directory.