EMC® ScaleIO®

Version 2.0.x

Deployment Guide

P/N 302-003-287

REV 03

EMC²

Copyright © 2016-2017 Dell Inc. or its subsidiaries. All rights reserved.

Published March 2017

Dell 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." DELL 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 DELL SOFTWARE DESCRIBED IN THIS PUBLICATION REQUIRES AN APPLICABLE SOFTWARE LICENSE.

Dell, EMC, and other trademarks are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be the property of their respective owners. Published in the USA.

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

CONTENTS

Figures 9

Tables 11

Preface 13

Part 1 Introduction 15

Chapter 1 Introduction to EMC ScaleIO 17
What is ScaleIO? 18
System requirements 18
ScaleIO cluster components 18
Physical server requirements 19
Supported operating systems 20
ScaleIO Gateway server requirements 22
Other requirements 23
What's new in this version? 23
Product limits 24

Chapter 2 Architecture 27
ScaleIO Architecture Overview 28
System 28
Hardware 28
Software 28
The MDM cluster 29
Storage definitions 31
Protection Domains 31
Storage Pools 31
Fault Sets 33
Naming 34
Protection and load balancing 34
Rebuild 34
Rebuild throttling 35
Rebalance 35
Rebalance throttling 36
Checksum protection 37
Caching 37
Networking 40
VMware limitation: 43
Virtual IP Address 43
Implementing ScaleIO 44
Physical layer 44
SAN virtualization layer 46
Other functions 47
Maintenance 50
Maintaining the physical layer 50
Instant maintenance mode 51

Maintaining the virtualization layer....51
S.M.A.R.T. hardware monitoring.... 52
List of approved RAID controllers.... 53
Snapshots.... 54
Snapshot operations....55
Management tools....55
Configuring direct attached storage (DAS).... 56
Implementing ScaleIO over a virtual system.... 56
VMware....56
Xen implementation.... 57

Part 2 Deploying ScaleIO Systems 59

Chapter 3 Deploying ScaleIO on Physical Servers 61
Overview of ScaleIO deployment on physical servers....62
Linux package names....63
Preparing the Installation Manager and the Gateway.... 63
Preparing the IM on a Linux server.... 64
Preparing the IM on a Windows server.... 66
Which installation mode is best for you?....67
Installing with the full Installation Manager....67
Preparing the CSV topology file.... 68
Installing with the Installation Manager....73
Installing with the Installation Manager wizard.... 83
Setting up the installation....84
Running the installation.... 87
Enabling your storage....89
Deploying on Ubuntu, Oracle Linux, or CoreOS servers....90
Ensuring the kernel version is correct.... 91
Creating the configuration file via the ScaleIO Gateway....92
Creating the configuration file manually.... 94
Running the ScaleIO Gateway on CoreOS container.... 95
Creating a mirror repository.... 95
Deploying on OpenStack.... 96
ScaleIO with Mirantis OpenStack.... 97
ScaleIO with Canonical (Ubuntu) OpenStack.... 97
ScaleIO with OpenStack Liberty....97
ScaleIO with OpenStack Mitaka.... 98
Environment and system requirements....99
Installing the ScaleIO GUI.... 100
Configuring Installation Manager properties.... 100

Chapter 4 Deploying ScaleIO on ESX Servers 103
Overview of ScaleIO deployment on ESX servers.... 104
Before you begin - ESX....104
Deployment prerequisites....104
Other deployment considerations....107
Deploying ScaleIO....108
Registering the ScaleIO plug-in and uploading the OVA template.... 108
Accessing the plug-in.... 111
Installing the SDC on ESX hosts....111
Deploying ScaleIO....113
Installing the ScaleIO GUI....130

Chapter 5 **Post-Deployment Activities** **133**
Post-deployment checklist....134
Getting started post-deployment....135
Creating a LockBox....135
Configuring ESRS....136
Configuring native users....137
Configuring LDAP users....138
Configuring SNMP after deployment....139
Create and map volumes....139

Part 3 **Reference** **143**

Chapter 6 **Manual Deployment** **145**
Manual deployment on physical servers....146
MDM cluster setup....148
Deploying a 3-node cluster....149
Deploying a 5-node cluster....154
Post-installation considerations....160
Provisioning storage example....166
Deploying ScaleIO on ESX....168
Deploying the ScaleIO virtual machine (SVM)....169
Configuring the UUID on virtual machines....170
Installing the SDC directly on an ESX host....171
Optimizing the guest operating system performance....172
Manually removing the VIB file from the ESX host....173

Chapter 7 **Advanced Topics** **175**
Advanced Gateway topics....176
Enabling and disabling Gateway components....176
Using a custom Java configuration for the Gateway....177
Installing the Gateway without assigning an admin password....177
Certificate management for ScaleIO Gateway....177
OpenStack interoperation with the ScaleIO Gateway....181
Generating a self-signed certificate using the keytool utility....181
Advanced Installation Manager topics....182
Adding devices to SDS nodes on Windows servers....182
Installing without validating Linux devices....183
Using SSH authentication on the ScaleIO Gateway....183
Advanced VMware plug-in topics....184
Advanced settings options....184
Registering the ScaleIO plug-in manually....185
Troubleshooting plug-in registration issues....187
Deploy ScaleIO in a 2-layer environment....188
Deploying ScaleIO in a 2-layer environment using the Installation Manager and plug-in....188
Manually deploying ScaleIO in a 2-layer environment....189
ESXi deployment prerequisites....190

Chapter 8 **Maintaining a ScaleIO System** **195**
Maintaining a ScaleIO system....196
Installing RFcache on servers in an existing ScaleIO system....196
Extending the MDM cluster from 3 to 5-node....197
Creating a LockBox for SNMP/ESRS....200

Working with Dynamic Host Name resolution for SNMP in ScaleIO... 200
Switching to secured authentication mode..... 201
Using SCLI in non-secure mode.....205
Extending an existing ScaleIO system..... 205
Configuring virtual IP addresses using Installation Manager..... 206
Removing components..... 207

Chapter 9 System Analysis 211
System analysis overview.....212
Creating the system analysis report..... 213
System analysis report description.....215

Chapter 10 Configuring ESRS connection properties 221
Before configuring ESRS.....222
Registering the system with the ESRS Gateway..... 222
Performing other ESRS configuration activities..... 224
Validating ConnectEMC Dial-Home..... 225

Chapter 11 Troubleshooting 227
Troubleshooting ScaleIO.....228
Adding a cache device to RFcache Storage Pool.....228
After SDC installation, the ScaleIO SVM does not start automatically.....230
Application server does not see a ScaleIO volume.....231
Cannot log in to the Installation Manager after upgrade.....231
Certificate error when installing SDC on Windows servers..... 232
Enabling acceleration in a new node..... 232
S.M.A.R.T. hardware alerts are not displayed.....235
Gateway server recovery.....236
Recovering Gateway on a new server during upgrade..... 236
Recovering IM with LIA trusted IP feature enabled.....238
SNMP configuration recovery post gateway crash during upgrade... 238
Installation Manager returns an error.....239
Installation with the Installation Manager fails..... 239
LIA upgrade fails in Windows 2008.....240
Older SDCs cannot communicate with newer SDSs after upgrade to v2.x 240
Removing RFcache (the xcache package) from Windows servers..... 241
Removing RFcache leftovers from the Windows OS registry..... 241
Replacing a faulty caching device in ScaleIO.....241
Replacing a faulty cached storage device in ScaleIO..... 243
ScaleIO CLI or GUI cannot connect to an MDM.....244
SCLI add_sds command fails due to communication error or MDM crash.... 245
SVM manual memory allocation.....245
Viewing the status of volumes by drv_cfg --rescan command.....247
Virtual IP feature is not functional..... 247
Enabling acceleration in a new node..... 248
The VMware plug-in responds slowly..... 251
Solving ScaleIO performance issues.....251
Mismatch in IO counters.....252
Deploying SVM on a node with a management IP address from a different subnet than the node with the SVM template..... 252

Speeding up rebuild and rebalance processes....252
RFcache is not installed on an upgraded Windows server....253
SSD devices are not recognized in the operating system.... 254
ESX competing thread setting resets on server reboot.... 255

Chapter 12 Frequently Asked Questions 257
Associating ScaleIO volumes with physical disks.... 258
Volume information - Linux....258
Volume information - Windows....259
Changing default ports....260
Adding an external SDC to an existing ScaleIO system.... 261
Installing SDC on an ESX server and connecting it to ScaleIO.....261
Installing SDC on a Linux server and connecting it to ScaleIO..... 262
Installing SDC on a Windows server and connecting it to ScaleIO.... 263
Changing the LIA configuration file....263
Cleaning the ScaleIO VMware environment and performing a clean install.... 264
Configuring ScaleIO devices in Linux LVM....266
Configuring session timeout parameters.... 266
Fixing keytool errors.... 267
Error during rpm installation command.... 267
Error during rpm upgrade command.... 267
Installing Java on SUSE 12 servers.... 267
Mounting ScaleIO....268
The ScaleIO Gateway web server isn’t responding....269
The ScaleIO Gateway (REST service, Installation Manager) may be disabled:.... 269
The ScaleIO Gateway web server isn't responsive and the following error appears in the catalina log file:.... 270
Upgrading the Gateway when a custom certificate is used.... 271
Uploading a new OVA....271
Using the same data network for different NICs.... 271
What to do when the default self-signed certificate expires....271
Xen v6.5 High Availability....271

Glossary 273

FIGURES

1 5-node MDM cluster 30
2 Protection Domains and Storage Pools 32
3 Protection Domains, Storage Pools, and Fault Sets 33
4 ScaleIO system deployed on a single network (Windows) 42
5 ScaleIO system deployed on separate networks (Windows) 43
6 Physical layout example—3-node cluster 45
7 Snapshot operations 54
8 ScaleIO implementation on ESX 57
9 ScaleIO Xen virtual machine architecture 58
10 CSV—complete 68
11 IM Installation Configuration screen 76
12 OpenStack Liberty integration 98
13 OpenStack Mitaka integration 99
14 EMC ScaleIO screen 111
15 Before extending 198
16 After extending 198
17 Set Virtual IPs for ScaleIO system screen 207

TABLES

1 Server physical requirements....19
2 Supported operating systems - ScaleIO components....20
3 Product limits....24
4 MDM cluster modes....30
5 Caching modes....37
6 Caching support matrix....39
7 IP address configurations in ScaleIO (based on CSV file)....40
8 Linux package formats....63
9 CSV topology spreadsheets....69
10 driver_sync.conf parameters....94
11 Sample manual installation topology....149
12 Sample manual installation topology....154
13 Provisioning storage example....166
14 Default ports....260

Preface

As part of an effort to improve its product lines, EMC periodically releases revisions of its software and hardware. Therefore, some functions described in this document might not be supported by all versions of the software or hardware currently in use. The product release notes provide the most up-to-date information on product features.

Contact your EMC technical support professional if a product does not function properly or does not function as described in this document.

Note

This document was accurate at publication time. Go to EMC Online Support (https://support.emc.com) to ensure that you are using the latest version of this document.

Related documentation

The release notes for your version includes the latest information for your product.

The following EMC publication sets provide information about your ScaleIO or ScaleIO Ready Node product:

- ScaleIO software (downloadable as ScaleIO Software <version> Documentation set)
- ScaleIO Ready Node with AMS (downloadable as ScaleIO Ready Node with AMS Documentation set)
- ScaleIO Ready Node no AMS (downloadable as ScaleIO Ready Node no AMS Documentation set)
- VxRack Node 100 Series (downloadable as VxRack Node 100 Series Documentation set)

You can download the release notes, the document sets, and other related documentation from EMC Online Support.

Typographical conventions

EMC uses the following type style conventions in this document:

Bold	Used for names of interface elements, such as names of windows, dialog boxes, buttons, fields, tab names, key names, and menu paths (what the user specifically selects or clicks)
Italic	Used for full titles of publications referenced in text
`Monospace`	Used for: • System code • System output, such as an error message or script • Pathnames, filenames, prompts, and syntax • Commands and options
Monospace italic	Used for variables
`Monospace bold`	Used for user input
[ ]	Square brackets enclose optional values

|	Vertical bar indicates alternate selections - the bar means “or”
{ }	Braces enclose content that the user must specify, such as x or y or z
...	Ellipses indicate nonessential information omitted from the example

Where to get help

EMC support, product, and licensing information can be obtained as follows:

Product information

For documentation, release notes, software updates, or information about EMC products, go to EMC Online Support at https://support.emc.com.

Technical support

Go to EMC Online Support and click Service Center. You will see several options for contacting EMC Technical Support. Note that to open a service request, you must have a valid support agreement. Contact your EMC sales representative for details about obtaining a valid support agreement or with questions about your account.

Your comments

Your suggestions will help us continue to improve the accuracy, organization, and overall quality of the user publications. Send your opinions of this document to techpubcomments@emc.com.

PART 1

Introduction

This part gives an overview of ScaleIO benefits and architecture. Chapters include:

Chapter 1, "Introduction to EMC ScaleIO"

Chapter 2, "Architecture"

CHAPTER 1

Introduction to EMC ScaleIO

This chapter introduces EMC® ScaleIO®. Topics include:

- What is ScaleIO? 18
- System requirements 18
- What's new in this version? 23
- Product limits 24

What is ScaleIO?

ScaleIO is a software-only solution that uses existing servers' local disks and LAN to create a virtual SAN that has all the benefits of external storage—but at a fraction of cost and complexity. ScaleIO utilizes the existing local storage devices and turns them into shared block storage. For many workloads, ScaleIO storage is comparable to, or better than external shared block storage.

The lightweight ScaleIO software components are installed on the application servers and communicate via a standard LAN to handle the application I/O requests sent to ScaleIO block volumes. An extremely efficient decentralized block I/O flow, combined with a distributed, sliced volume layout, results in a massively parallel I/O system that can scale up to thousands of nodes.

ScaleIO is designed and implemented with enterprise-grade resilience. Furthermore, the software features an efficient distributed self-healing process that overcomes media and server failures, without requiring administrator involvement.

Dynamic and elastic, ScaleIO enables administrators to add or remove servers and capacity on-the-fly. The software immediately responds to the changes, rebalancing the storage distribution and achieving a layout that optimally suits the new configuration.

Because ScaleIO is hardware agnostic, the software works efficiently with various types of disks, including: magnetic (HDD) and solid-state disks (SSD), flash PCI Express (PCIe) cards, networks, and hosts.

ScaleIO can easily be installed in an existing infrastructure as well as in green field configurations.

System requirements

This section lists the requirements for system components.

ScaleIO cluster components

List of required ScaleIO servers.

- ScaleIO component servers:
  - 3-node cluster
    - One Master MDM
    - One Slave MDM
    - One TieBreaker
    - Minimum of three SDSs (on the same servers as the above components, or on three different servers)
    - Any number of SDCs (on the same servers as the above components, or on different servers)
  - 5-node cluster
    - One Master MDM
    - Two Slave MDMs

  - Two TieBreakers
  - Minimum of three SDSs (on the same servers as the above components, or on three different servers)
  - Any number of SDCs (on the same servers as the above components, or on different servers)
- ScaleIO Gateway server, on the same server as any of the ScaleIO components, or on a separate server. Do not install the Gateway on a server on which RFcache (the `xcache` package) or SDC will be installed.

Physical server requirements

Table 1 Server physical requirements

Component	Requirement
Processor	One of the following: • Intel or AMD x86 64-bit (recommended) • Intel or AMD x86 32-bit (for Xen only)
Physical memory	ScaleIO component requirements: • 500 MB RAM for the Meta Data Manager (MDM) • 500 MB RAM for each ScaleIO Data Server (SDS) • 50 MB RAM for each ScaleIO Data Client (SDC) DAS Cache memory requirements (ScaleIO Ready Node). Add to every SVM/node that will be using DAS Cache: • 1U1N servers—500 MB RAM • 2U1N servers—1 GB RAM For optimal performance, remove the swap space (hard disk virtual memory) from the operating system disk. If other applications are running, consult with the application vendor to determine the best practice for swap space and size allocation.
Disk space	• 1 GB for each physical node or Xen hypervisor • 10 GB for VMware topologies
Connectivity	One of the following: • 1 GbE or 10 GbE (recommended) network • IP-over-InfiniBand network Dual-port network interface cards (recommended) Ensure the following: • There is network connectivity between all components. • Network bandwidth and latency between all nodes is acceptable, according to application demands. • Ethernet switch supports the bandwidth between network nodes. • MTU settings are consistent across all servers and switches. • The following TCP ports are not used by any other application, and are open in the local firewall of the server:

Table 1 Server physical requirements (continued)

Component	Requirement
	▪ MDM: 6611 and 9011 ▪ SDS: 7072 (for multiple SDS, ports 7073-7076) ▪ ScaleIO Gateway (includes REST Gateway, Installation Manager, and SNMP trap sender): 80 and 443 ▪ Light Installation Agent (LIA): 9099 ▪ SDBG ports: MDM 25620, SDS 25640, Multiple SDS 25641-25644 (not 25640) • The following UDP port is open in the local firewall of the server: ▪ SNMP traps: 162 **Note** You can change the default ports. For more information, see "Changing default ports" in the user documentation.

Supported operating systems

For the most updated list, see the EMC Simple Support Matrix (ESSM) at https://elabnavigator.emc.com.

Table 2 Supported operating systems - ScaleIO components

Operating system	Requirement
Linux	Supported versions: • CentOS 6.x-7.x, Oracle Linux 6.5/7.x • Red Hat 6.x-7.x • SLES 11.3, SLES 12, SLES 12.1 • Ubuntu 14.04, Ubuntu 16.04 **Note** Before deploying on Ubuntu servers, you must prepare the environment, as described in the *EMC ScaleIO Deployment Guide*. Packages required for all components, all Linux flavors: • numactl • libaio Additional packages required for MDM components: • bash-completion (for scli completion) • Latest version of Python 2.X When installing the MDM component on Linux CentOS 6 or RHEL 6 hosts (for software-only systems), set the shared memory parameter in the /etc/sysctl.conf file to at least the

Table 2 Supported operating systems - ScaleIO components (continued)

Operating system	Requirement

	following value: kernel.shmmax=209715200. To use this value, type the `sysctl -p` command.
	Requirements for running the GUI:
	• Java 1.8 64-bit, or later (For SUE 12, see "Installing Java on SUE 12 servers" in the *ScaleIO Deployment Guide*.)
	• Screen resolution: 1366 x 768 minimum
	• Additional supported operating systems:
	  ▪ Windows 7
	  ▪ Windows 2010
	To use the secure authentication mode, ensure that OpenSSL 64-bit v1.0.1 or later is installed on all servers in the system.
	To use the secure authentication mode on SLES 11.3 servers, ensure that the OpenSSL on the server is v1.0.1 or greater, or install these packages (from the ISO in the Complete VMware SW download container) on the server:
	• `libopenssl1_0_0-1.0.1g-0.40.1.x86_64.rpm`
	• `openssl1-1.0.1g-0.40.1.x86_64.rpm`
	To use LDAP, ensure that OpenLDAP 2.4 is installed on all servers.

Windows	Supported versions:
	• 2008 R2, 2012, 2012 R2, or 2016 (in v2.0.1.1 and later)
	• For VxRack Node 100 Series, only 2012 R2 is supported.
	• On all MDM servers, install the EMC-provided `PythonModulesInstall.exe` on all MDM nodes. The file is supplied on the ISO, or download from the EMC Online Support site (search for ScaleIO Python Installation Modules) on https://support.emc.com.
	• To install SDC or RFcache on 2008 R2, ensure that Microsoft Security Update KB3033929 is installed.
	Requirements for running the GUI:
	• Java 1.8 64-bit or later
	• Screen resolution: 1366 x 768 minimum
	• Additional supported operating systems:
	  ▪ Windows 7
	  ▪ Windows 2010
	To use the secure authentication mode, ensure that these are installed on all servers in the system:
	• OpenSSL 64-bit v1.0.1 or later
	• Visual C++ redistributable 2010 package, 64-bit
	To use RFcache, ensure that Visual C++ redistributable 2010 package, 64-bit is installed on all servers in the MDM cluster and on all SDSs.

Hypervisors	• VMware ESXi OS: 5.5 or 6.0, managed by vCenter 5.5 (update 2e or later) or 6.0

Table 2 Supported operating systems - ScaleIO components (continued)

Operating system	Requirement
	• Hyper-V • Xenserver 6.5 or 7.0 **Note** OpenSSL 64-bit v1.0.1 is supported on XEN6.5 SP1 (or later) • RedHat KVM
OpenStack	Supported versions: • Liberty • Mitaka • Newton • Redhat OpenStack Platform • Mirantis OpenStack • Canonical Ubuntu OpenStack.

ScaleIO Gateway server requirements

Component	Requirement
Supported operating systems	• Windows 2008 R2, 2012 R2, or 2016, including the Visual C++ redistributable 2010 package, 64-bit • Linux: ▪ CentOS 6.x-7.x ▪ Oracle Linux 6.5/7.x ▪ Red Hat 6.x-7.x ▪ SLES 11.3, SLES 12, SLES 12.1 ▪ Ubuntu 14.04, Ubuntu 16.04 Every server requires 2 cores and a minimum of 3 GB available RAM.
Supported web browsers	• Internet Explorer 10, or later • Firefox, version 42, or later • Chrome, version 45, or later
Java requirements	• 1.8 or later, 64-bit
Other	• For a Windows Installation Manager (IM): ▪ The WMI service must be enabled on the IM server and on all Windows ScaleIO nodes. ▪ Don't install the Gateway on a server on which RFcache (the `xcache` package) or SDC will be installed.

Component	Requirement
	• The Gateway server must have connectivity to all the nodes that are being installed. If you are using separate networks for management and data, the server must be able to communicate with both networks.

Other requirements

ScaleIO requires that you use a minimum of three SDS servers, with a combined free capacity of at least 300 GB. These minimum values are true per system and per Storage Pool.

NOTICE

ScaleIO installation enables unlimited use of the product, in non-production environments . To obtain a license for production use, and to receive technical support, open a service ticket with EMC Support at https://support.emc.com.

For complete information on licensing, see the *ScaleIO User Guide*.

What's new in this version?

This version introduces the following new features:

- The MDM cluster can be assigned a virtual IP address, that will be used for communications between the MDM cluster and SDCs. MDMs are sometimes switched during normal operation of the cluster, and the virtual IP address will always be mapped to the active MDM. The virtual IP address on the MDM cluster ensures that all SDCs remain connected to the MDM, even after the physical server is replaced, thus simplifying maintenance procedures and preventing the need to reconfigure all SDCs in a system.
- The Get Info operation has been enhanced to include an offline GUI, a snapshot of the ScaleIO GUI display at a given point in time. This enables ScaleIO support personnel to get a clearer view of a system, so they can provide better assistance.
- vSphere rollback improvement: you can roll back a failed deployment and the deployment settings are saved for the redeployment. The roll-back option in the vSphere Web plug-in enables you to roll back a failed deployment. The roll-back feature now retains all of the current settings in the wizard for redeployment. This allows you to roll back the failed deployment, make changes to specific settings without needing to re-enter all other data, and then restart the deployment process. Rolling back is useful when you have entered incorrect information in one of the deployment settings or when one of the devices used in the deployment has failed.
  The roll-back option can only save all settings when you are using the same computer and the same browser from which you initially deployed. Similarly, if you delete your browser's cookies or history, your settings will not be saved. Once you exit the deployment wizard, the deployment settings are not retained.
- The GUI now includes a filter in the **Frontend > Snapshot** view, similar to the filter in the **Volumes** and **SDC** views. The filter lets you filter out the information shown in the table, so that only the objects related to the items selected in the filter are visible.

- A login banner can be displayed upon login to the system. It can be used to communicate messages or to obtain user consent to real-time monitoring of information and retrieval of stored files. When the login banner is set up, it appears during the system login process before the login credential prompts.
- The ScaleIO bare-metal solution now provides monitoring capabilities for RAID controllers and storage devices compatible with S.M.A.R.T. (Self-Monitoring, Analysis and Reporting Technology) protocols. In Linux-based environments, S.M.A.R.T. compatible HDDs, SSDs and RAID storage controllers can be monitored for S.M.A.R.T. attributes such as temperature, SSD wear level, and error counters. Physical device information such as model, firmware, serial number, etc. are also shown. These can help the user identify and maintain the devices. LEDs can also be lit on these hardware devices, to simplify physical identification for maintenance purposes.
- ScaleIO has implemented VMware's vStorage API for Storage Awareness (VASA). ScaleIO's VASA enables control of ScaleIO storage to be handled by the vCenter administrator. ScaleIO's VASA uses Virtual Volumes (VVols) to enable VMs hosted on ESXs in the vCenter to have their storage mapped directly to storage in ScaleIO. For more information, see the *Using ScaleIO with VVols* technical notes.

Note

The implementation of VVols has not yet been certified by VMware and therefore is not recommended for production environments.

Product limits

The following table lists product capabilities:

Table 3 Product limits

Item	Limit
ScaleIO System raw capacity	300 GB—16 PB
Device size	100 GB—8 TB
Minimum Storage Pool capacity	300 GB
Volume size	8 GB—1 PB
Maximum number of volumes/snapshots in system	32,768[a]
Maximum number of volumes/snapshots in Protection Domain	32,768
Maximum number of volumes + snapshots in single VTree	32
Maximum capacity per SDS	96 TB
SDSs per system	1024
SDSs per Protection Domain	128[a]
Maximum devices (disks) per SDS server	64[b]
Maximum devices (disks) per Storage Pool	300[a]
Minimum devices (disks) per Storage Pool	3, on different SDSs

Table 3 Product limits (continued)

Item	Limit
Maximum SDCs per system	1024 When using replication with RecoverPoint, the maximum number of SDCs is reduced by the number of RPAs in the system.[c]
Maximum volumes that can be mapped to a single SDC	8192
Maximum Protection Domains per system	256
Maximum Storage Pools	1024
Maximum Storage Pools per Protection Domain	64
Maximum Fault Sets per Protection Domain	64
Maximum IP addresses per server (MDM and SDS)	8
RAM Cache	128 MB—300 GB

a. If more are needed, contact EMC Support.
b. On VMware servers, the maximum devices per SDS is 59.
c. Replication support is version-specific. For information, see the ESSM.

CHAPTER 2

Architecture

This chapter describes the ScaleIO architecture. Topics include:

- ScaleIO Architecture Overview 28
- System 28
- The MDM cluster 29
- Storage definitions 31
- Protection and load balancing 34
- Caching 37
- Networking 40
- Virtual IP Address 43
- Implementing ScaleIO 44
- Other functions 47
- Maintenance 50
- S.M.A.R.T. hardware monitoring 52
- Snapshots 54
- Management tools 55
- Configuring direct attached storage (DAS) 56
- Implementing ScaleIO over a virtual system 56

ScaleIO Architecture Overview

This chapter describes the ScaleIO architecture overview.

ScaleIO is a software-only solution. ScaleIO components are lightweight, highly available software components, installed on new or existing servers alongside your production applications (hypervisors, databases, web applications, etc.). The system can be installed directly on the servers, or over a virtual server system (hypervisor or virtual machines).

System

The ScaleIO system is based on a hardware and a software component.

Hardware

In general, hardware can be the existing application servers used by the datacenter, or a new set of nodes (if, for example, you want to dedicate all nodes solely for the purpose of running the ScaleIO SAN storage system).

- *Nodes*
  Nodes, or servers, are the basic computer unit used to install and run the ScaleIO system. They can be the same servers used for the applications (server convergence), or a dedicated cluster. In any case, ScaleIO is hardware-agnostic, and therefore, aside from performance considerations, the type of server is inconsequential.
- *Storage Media*
  The storage media can be any storage media, in terms of the type (HDD, SSD, or PCIe flash cards) and anywhere (DAS, or external).

Software

The ScaleIO virtual SAN consists of the following software components:

- Meta Data Manager—MDM: Configures and monitors the ScaleIO system. The MDM can be configured in redundant Cluster Mode, with three members on three servers or five members on five servers, or in Single Mode on a single server.

  NOTICE

  It is not recommended to use Single Mode in production systems, except in temporary situations. The MDMs contains all the metadata required for system operation. Single Mode has no protection, and exposes the system to a single point of failure.

- ScaleIO Data Server—SDS: Manages the capacity of a single server and acts as a back-end for data access. The SDS is installed on all servers contributing storage devices to the ScaleIO system. These devices are accessed through the SDS.
- ScaleIO Data Client—SDC: A lightweight device driver that exposes ScaleIO volumes as block devices to the application that resides on the same server on which the SDC is installed.

Depending on the desired configuration (described later), the software components are installed on the server and give rise to a virtual SAN layer exposed to the applications residing on the servers.

The MDM cluster

The MDM serves as the monitoring and configuration agent of the ScaleIO system. The MDM is mainly used for management which consists of migration, rebuilds, and all system-related functions. No I/O run through the MDM.

To support high availability, three or more instances of MDM run on different servers. In a multi-MDM environment, one MDM is given the Master role, and the others act as Slave or TieBreaker MDMs.

The MDM cluster comprises a combination of Master MDM, Slave MDMs, and TieBreaker MDMs.

The following terms are relevant to the MDM, the building blocks of the MDM cluster:

- MDM

  Any server with the MDM package installed on it. An MDM can be given a Manager or a TieBreaker (default) role, during installation. MDMs have a unique MDM ID, and can be given unique names.

  Before the MDM can be part of the cluster, it must be promoted to a Standby MDM.

- Standby MDM and TieBreaker

  An MDM and a TieBreaker can be added to a system as a standby. Once added, the standby MDM or TieBreaker is attached, or locked, to that specific system.

  A standby MDM can be called on to assume the position of a Manager MDM or TieBreaker MDM when it is promoted to be a cluster member.

- Manager MDM

  An MDM that can act as a Master or a Slave in the cluster. Manager MDMs have a unique system ID, and can be given unique names. A manager can be a standby or a member of the cluster.

  In ScaleIO documentation, “MDM” refers to a manager, unless specified otherwise.

- TieBreaker MDM

  An MDM whose sole role is to help determine which MDM is the master. A TieBreaker can be a standby or a member of the cluster. A TieBreaker MDM is not a manager.

  In a 3-node cluster, there is one TB; in a 5-node cluster, there are two TBs. This ensures that there are always an odd number of MDMs in a cluster, which guarantees that there is always a majority in electing the master.

The following terms are relevant to the MDM cluster, specifically:

- Master MDM (used to be called Primary MDM)

  The MDM in the cluster that controls the SDSs and SDCs. The Master MDM contains and updates the MDM repository, the database that stores the SDS configuration, and how data is distributed between the SDSs in the system. This repository is constantly replicated to the Slave MDMs, so they can take over with no delay.

  Every MDM cluster has one Master MDM.

- Slave MDM (used to be called Secondary MDM)

  An MDM in the cluster that is ready to take over the Master MDM role if ever necessary.

In a 3-node cluster, there is one Slave MDM, thus allowing for a single point of failure. In a 5-node cluster, there are two Slave MDMs, thus allowing for two points of failure. This increased resiliency is a major benefit to enabling the 5-node cluster.

- Replica

  An MDM that contains a replica of the MDM repository. This includes the Master MDM and any Slave MDMs in the MDM cluster.

The following table describes the available cluster modes:

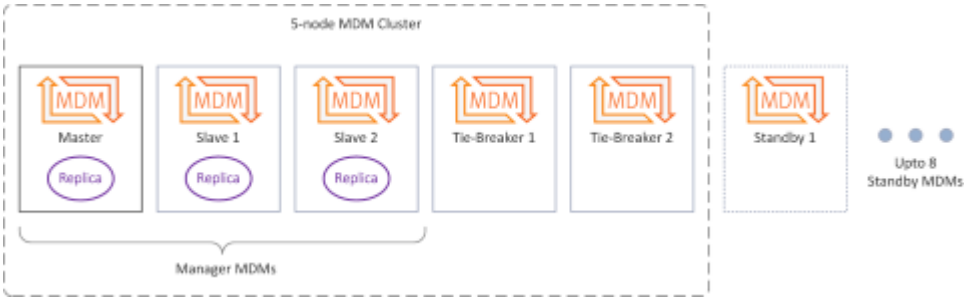
Table 4 MDM cluster modes

Cluster mode	Members	Description
3-node (default)	• Master MDM • Slave MDM • TieBreaker	3-node cluster has two copies of the repository, thus can withstand one MDM cluster failure.
5-node	• Master MDM • Two Slave MDM • Two TieBreaker	5-node cluster has three copies of the repository, thus can withstand two MDM cluster failure.
Single-node	• Master MDM	Single-node cluster has only one copy of the repository, thus it cannot withstand failure. It is not recommended to use Single Mode in production systems, except in temporary situations.

In addition to the cluster members, you can prepare standby Managers and TieBreaker nodes, for a total of thirteen cluster and standby MDMs.

The following figure illustrates a 5-node MDM cluster:

Figure 1 5-node MDM cluster



All members of the MDM cluster have the same MDM package installed on them.

Before a server makes its way into the MDM cluster, it must follow the following path:

1. Install the MDM package on the server.
   During the installation, you determine if the server will be a Manager or a TieBreaker (default).
2. Promote the server to Standby status, either as a Manager or as a TieBreaker.

3. Add the standby server to the MDM cluster. A Manager, once entered into the cluster can take on the Master or Slave state.

MDM cluster creation is done automatically when deploying a system with any of the automated deployment tools.

Storage definitions

When configuring a ScaleIO system, you should take the following concepts into account: Protection Domains, Storage Pools, and Fault Sets. Together, these elements link the physical layer with the virtualization layer.

Protection Domains

A Protection Domain is a logical entity that contains a group of SDSs that provide backup for each other. Each SDS belongs to one (and only one) Protection Domain. Thus, by definition, each Protection Domain is a unique set of SDSs. In Figure 2 there are three Protection Domains. The one in the middle (fully depicted) consists of seven SDSs, each with two storage devices.

The maximum recommended number of nodes in a Protection Domain is 100. This enables the following:

- optimal performance
- reduction of theoretical *mean time between failure* issues
- ability to sustain multiple failures in different Protection Domains

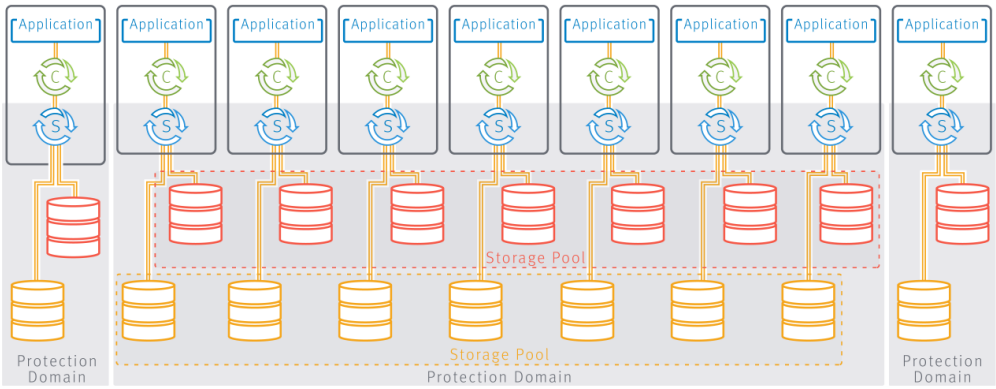
You can add Protection Domains during installation. In addition, you can modify Protection Domains post-installation with all the management clients (except for OpenStack) .

Storage Pools

Storage Pools allow the generation of different storage tiers in the ScaleIO system. A Storage Pool is a set of physical storage devices in a Protection Domain. Each storage device belongs to one (and only one) Storage Pool. In Figure 2, there are 2 Storage Pools depicted.

When a volume is configured over the virtualization layer (see “SAN virtualization layer”), it is distributed over all devices residing in the same Storage Pool. Each volume block has two copies located on two different SDSs. This allows the system to maintain data availably following a single-point failure. The data will still be available following multiple failures, as long as each failure took place in a different storage pool.

Figure 2 Protection Domains and Storage Pools



To provide consistent performance it is recommended that all devices in the Storage Pool will have similar storage properties.

For example, consider Figure 2. If all SDSs in a Protection Domain have two physical drives associated with them—one HDD and the other SSD— then you should define two Storage Pools:

- Capacity Storage Pool

  Consists of all HDDs in the Protection Domain

- Performance Pool

  Consists of all SSDs in the Protection Domain

Note

Mixing different types of media in the same pool is allowed, but be aware that due to the distribution of the data, performance will be limited to the least-performing member of the Storage Pool.

ScaleIO might not perform optimally if there are large differences between the sizes of the devices in the Storage Pool, for example, if one device is as big as the rest of the devices. If in doubt, contact ScaleIO support.

Each Storage Pool can work in one of the following modes:

- Zero padding enabled

  Ensures that every read from an area previously not written to returns zeros. Some applications might depend on this behavior. Furthermore, zero padding ensures that reading from a volume will not return information that was previously deleted from the volume.

  This behavior incurs some performance overhead on the first write to every area of the volume.

- Zero padding disabled (default)

  A read from an area previously not written to will return unknown content. This content might change on subsequent reads.

Zero padding must be enabled if you plan to use any other application that assumes that when reading from areas not written to before, the storage will return zeros or consistent data.

Note

The zero padding policy cannot be changed after the addition of the first device to a specific Storage Pool.

You can add Storage Pools during installation. In addition, you can modify Storage Pools post-installation with most of the management clients.

Fault Sets

A Fault Set is a logical entity that contains a group of SDSs within a Protection Domain, that have a higher chance of going down together, for example if they are all powered in the same rack. By grouping them into a Fault Set, you are telling ScaleIO that the data mirroring for all devices in this Fault Set, should take place on SDSs that are outside of this Fault Set.

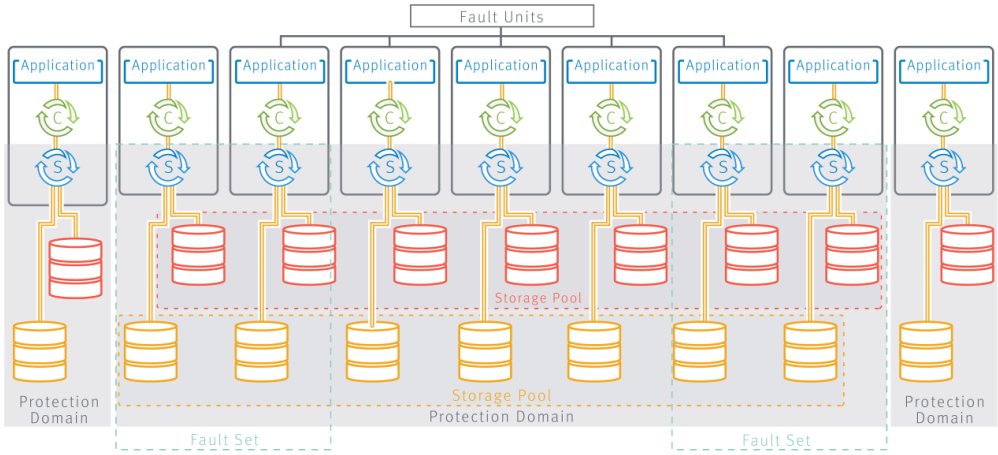
When defining Fault Sets, we refer to the term fault units, where a fault unit can be either a Fault Set, or an SDS not associated with a Fault Set (you may think of it as a Fault Set of a single SDS).

There must be enough capacity within at least 3 fault units to enable mirroring.

If Fault Sets are defined, you can use any combination of fault units, for example:

- SDS1, SDS2, SDS3
- FS1, SDS1, SDS2
- FS1, FS2, SDS1
- FS1, FS2, FS3
  Figure 3 on page 33 illustrates the same configuration as Figure 2, with the addition of Fault Sets.

  Figure 3 Protection Domains, Storage Pools, and Fault Sets



To use Fault Sets, you must work in the following order:

1. Ensure that a Protection Domain exists, or add a new one.
2. Ensure that a Storage Pool and Fault Sets (minimum of 3 fault units) exist, or add new ones.
3. Add the SDS, designating the PD and FS, and at the same time, adding the SDS devices into a Storage Pool.
   The automated deployment and installation tools follow this order automatically.

You can only create and configure Fault Sets before adding SDSs to the system, and configuring them incorrectly may prevent the creation of volumes. An SDS can only be added to a Fault Set during the creation of the SDS.

You define Fault Sets and add SDSs to them during installation, using the following management tools:

- Installation manager
- CLI
- REST
- VMware plug-in

In addition, you can also add Fault Sets when adding SDS nodes after initial installation.

Naming

It is recommended to name all ScaleIO objects with meaningful names. This will make it easier when defining volumes, associating them with applications, etc.

From the previous example, the Storage Pools can be named "Capacity_Storage" and "Performance_Storage," which allows you to identify the different tiers.

As for Protection Domains, one example would be separating the SDSs used by the finance department from those used by the engineering department. This segregation of different departments is very beneficial in many aspects (security being one of them). Thus, one might name the domains "Financial-PD" and "Engineering-PD."

The Fault Sets could be called "FS_Rack01" and "FS_Rack02."

Protection and load balancing

ScaleIO maintains the user data in a RAID-1 mesh mirrored layout. Each piece of data is stored on two different servers. The copies are randomly distributed over the storage devices. Rebuild and rebalance processes are fully automated, but are configurable.

Rebuild

When a failure occurs, such as on a server, device or network failure, ScaleIO immediately initiates a process of protecting the data. This process is called Rebuild, and comes in two flavors:

- Forward rebuild is the process of creating another copy of the data on a new server. In this process, all the devices in the Storage Pool work together, in a many-to-many fashion, to create new copies of all the failed storage blocks. This method ensures an extremely fast rebuild.
- Backward rebuild is the process of re-synchronization of one of the copies. This is done by passing to the copy only changes made to the data while this copy was inaccessible. This process minimizes the amount of data transferred over the network during recovery.

ScaleIO automatically selects the type of rebuild to perform. This implies that in some cases, more data will be transferred to minimize the time that the user data is not fully protected.

Rebuild throttling

Rebuild throttling sets the rebuild priority policy for a Storage Pool. The policy determines the priority between the rebuild I/O and the application IO when accessing SDS devices. Please note that application I/Os are continuously served.

Applying rebuild throttling will on one hand increase the time the system is exposed with a single copy of some of data, but on the other hand, will reduce the impact on the application. One has to make a decision and choose the right balance between the two.

The following possible priority policies may be applied:

- No Limit: No limit on rebuild I/Os.

  Any rebuild I/O is submitted to the device immediately, without further queuing. Please note that rebuild I/Os are relatively large and hence setting this policy will speed up the rebuild, but will have the maximal effect on the application I/O.
- Limit Concurrent I/O: Limit the number of concurrent rebuild I/Os per SDS device (default).

  The rebuild I/Os are limited to a predefined number of concurrent I/Os. Once the limit is reached, the next incoming rebuild I/O waits until the completion of a currently executed rebuild I/O. This will complete the Rebuild quickly for best reliability, however, there is a risk of host application impact.
- Favor Application I/O: Limit rebuild in both bandwidth and concurrent I/Os.

  The rebuild I/Os are limited both in bandwidth and in the amount of concurrent I/Os. As long as the number of concurrent rebuild I/Os, and the bandwidth they consume, do not exceed the predefined limits, rebuild I/Os will be served. Once either threshold is reached, the rebuild I/Os wait until both I/O and bandwidth are below their thresholds. For example, setting the value to "1" will guarantee the device will only have one concurrent rebuild IO at any given moment, which will ensure the application IOs only wait for 1 rebuild IO at worst case.

  This imposes bandwidth on top of the Limit Concurrent I/Os option, which is a prerequisite to using this policy.
- Dynamic Bandwidth Throttling: This policy is similar to Favor Application I/O, but extends the interval in which application I/Os are considered to be flowing by defining a minimal quiet period. This quiet period is defined as a certain interval in which no application I/Os occurred. Note that the limits on the rebuild bandwidth and concurrent I/Os are still imposed.
- Default Values:
  - The default policy for rebuild is: Limit Concurrent I/O
  - Rebuild concurrent I/O Limit: 1 concurrent I/O

  Note

  Rebuild throttling affects the system's performance and should only be used by advanced users.

Rebalance

Rebalance is the process of moving one of the data copies to a different server. It occurs when ScaleIO detects that the user data is not evenly balanced across the devices in a Storage Pool. This can occur as a result of several conditions such as: SDS addition/removal, device addition/removal, or following a recovery from a failure.

ScaleIO will move copies of the data from the most utilized devices to the least utilized ones.

Both Rebuild and Rebalance compete with the application IO for the system resources. This includes network, CPU and disks. ScaleIO provides a very rich set of parameters that can control this resource consumption. While the system is factory-tuned for balancing between speedy rebuild/rebalance and minimization of the effect on the application IO, the user has very fine-grain control over the rebuild and rebalance behavior.

Rebalance throttling

Rebalance throttling sets the rebalance priority policy for a Storage Pool. The policy determines the priority between the rebalance I/O and the application IO when accessing SDS devices. Please note that application I/Os are continuously served. Rebalance, unlike rebuild, does not impact the system's reliability and therefore reducing its impact is not risky.

The following possible priority policies may be applied:

- No Limit: No limit on rebalance I/Os.

  Any rebalance I/O is submitted to the device immediately, without further queuing. Please note that rebalance I/Os are relatively large and hence setting this policy will speed up the rebalance, but will have the maximal effect on the application I/O.
- Limit Concurrent I/O: Limit the number of concurrent rebalance I/Os per SDS device.

  The rebalance I/Os are limited to a predefined number of concurrent I/Os. Once the limit is reached, the next incoming rebalance I/O waits until the completion of a currently executed rebalance I/O. For example, setting the value to "1" will guarantee that the device will only have one rebalance IO at any given moment, which will ensure that the application IOs only wait for 1 rebalance IO in the worst case.
- Favor Application I/O: Limit rebalance in both bandwidth and concurrent I/Os.

  The rebalance I/Os are limited both in bandwidth and in the amount of concurrent I/Os. As long as the number of concurrent rebalance I/Os, and the bandwidth they consume, do not exceed the predefined limits, rebalance I/Os will be served. Once either limiter is reached, the rebalance I/Os wait until such time that the limits are not met again.

  This imposes a bandwidth limit on top of the Limit Concurrent I/Os option.
- Dynamic Bandwidth Throttling: This policy is similar to Favor Application I/O, but extends the interval in which application I/Os are considered to be flowing by defining a minimal quiet period. This quiet period is defined as a certain interval in which no application I/Os occurred. Note that the limits on the rebalance bandwidth and concurrent I/Os are still imposed.
- Default Values:
  - The default policy for rebalance: Favor Application I/O
  - Rebalance concurrent I/O Limit: 1 concurrent I/O per SDS device
  - Rebalance bandwidth limit: 10240 KB/s

Note

Rebalance throttling affects the system's performance and should only be used by advanced users.

Checksum protection

This feature addresses errors that change the payload during the transit through the ScaleIO system. ScaleIO protects data in-flight by calculating and validating the checksum value for the payload at both ends.

- During write operations, the checksum is calculated when the SDC receives the write request from the application. This checksum is validated just before each SDS writes the data on the storage device.
- During read operations, the checksum is calculated when the data is read from the SDS device, and is validated by the SDC before the data returns to the application. If the validating end detects a discrepancy, it will initiate a retry. The checksum will be done in the granularity of a sector (1/2KB).

This feature applies to all IOs: Application, Rebuild, Rebalance, and Migrate. The checksum is also kept in RMcache (Read Memory Cache), protecting every block that is maintained in SDS memory against memory corruption. The checksum feature can be enabled at the Protection Domain level, and defined at the Storage Pool level. The feature is T10/DIF-ready.

Note

The checksum feature may have a major impact on performance and availability. Contact EMC customer support to verify if your use case is relevant.

Caching

ScaleIO offers a number of caching options, for the purpose of enhancing system performance.

The following caching options are supported by ScaleIO:

- RAM Read Cache (using DRAM server memory)
- Read Flash Cache (using SSD devices)

In addition, the following caching solutions are available:

- CacheCade (using SSD devices) - available in VxRack Node 100 Series systems
- DAS Cache (using SSD devices) - available in ScaleIO Ready Node systems

SSDs used for caching cannot be used for storage purposes.

The following table summarizes information about the caching modes provided by the system.

Table 5 Caching modes

Mode	Description	Considerations	Default Setting
RAM Read Cache (RMcache)	Read-only caching performed by server RAM.	RAM Read cache, the fastest type of caching, uses RAM that is allocated for caching. Its size is limited to the amount of allocated RAM.	Disabled, except when storage-only nodes are deployed.

Table 5 Caching modes (continued)

Mode	Description	Considerations	Default Setting
		Note The amount that may be allocated is limited, and can never be the maximum available RAM.	
RFcache	Read-only caching performed by one or more dedicated SSD devices in the server.	RFcache uses the full capacity of SSD devices (up to eight) to provide a larger footprint of read-only LRU (Least Recently Used) based-caching resources for the SDS. This type of caching reacts quickly to workload changes to speed up HDD Read performance. Several SSD devices can be allocated to a shared cache pool, and therefore the cache size is limited in size only by the amount of SSDs allocated for this purpose. The RFcache driver must be installed during deployment. Caching devices can be defined either during the installation process or after deployment. Limitations: RFcache does not support partitions on devices installed on Windows nodes. Support matrix: • An RFcache device (flash device) can be partitioned only on Linux. • An SDS storage/source device cannot be partitioned if it needs to be accelerated by RFcache. • An SDS storage/source device as a file (over file system), cannot be accelerated by RFcache.	
CacheCade	Read and write-back caching performed by one or more dedicated SSD devices in the server.	CacheCade uses the full capacity of one or more SSD devices to provide a large footprint of both read and write-back caching resources to the SDS. This caching mode moves "hot" (active) chunks of data from HDDs to cache, for Read and Write buffering. For write-back caching, the write is temporarily written to the SSD, which is much faster than an HDD, allowing faster response of the SDS to write acknowledgment. Two SSD devices can be allocated to a shared cache pool, up to a maximum size of 512 GB in total.	Disabled

Table 5 Caching modes (continued)

Mode	Description	Considerations	Default Setting
		Note If a fault occurs in the caching device before the writes have been offloaded, all the HDD devices cached by CacheCade acquire failed status, and a rebuild process commences in VxRack Node 100 Series. Once the rebuild is over, the caching disk can be replaced, all caching has stopped in the storage pool, and the HDD members in the storage pool can be cleared of errors.	
DAS Cache	Read and write-back caching performed by one or more dedicated SSD devices in the server	DAS Cache uses the full capacity of one or more SSD devices to provide a large footprint of both read and write-back caching resources to the SDS. This caching mode moves "hot" (active) chunks of data from HDDs to cache, for Read and Write buffering. For write-back caching, the write is temporarily written to the SSD, which is much faster than an HDD, allowing faster response of the SDS to write acknowledgment. One SSD device can accelerate several HDDs (in DAS Cache they are called "Volumes"). Striping the Cache on two devices is not supported in the ScaleIO Ready Node solution. **Note** If a fault occurs in the caching device before the writes have been offloaded, all the HDD devices cached by DAS Cache acquire failed status, and a rebuild process commences in ScaleIO. Once the rebuild is over, the caching disk can be replaced, all caching has stopped in the storage pool, and the HDD members in the storage pool can be cleared of errors.	Disabled

The following table illustrates the caching support matrix:

Table 6 Caching support matrix

System	RFcache	RMcache	DAS Cache	CacheCade
ScaleIO	Yes	Yes		
ScaleIO Ready Node	Yes	Yes	Yes	
VxRack Node 100 Series	Yes	Yes		Yes

Networking

In ScaleIO, inter-node communication (for the purposes of managing data locations, rebuild and rebalance, and for application access to stored data) can be done on one IP network, or on separate IP networks. Management (via any of the management interfaces) can be done in the following ways:

- Via a separate network with access to the other ScaleIO components
- On the same network

These options can be configured a) during deployment in the full Installation Manager (via the CSV topology file) and using the VMware plug-in, as well as b) after deployment with the CLI.

This section describes how to choose from these options, depending on your organization's requirements, security considerations, performance needs, and IT environment.

ScaleIO networking considerations:

- Single IP network: All communications and IOs used for management and for data storage are performed on the same IP network. This setup offers the following benefits:
  - Ease of use
  - Fewer IP addresses required
- Multiple separate IP networks: Separate networks are used for management and for data storage, or separate networks are used within the data storage part of the system. This setup offers the following benefits:
  - Security
  - Redundancy
  - Performance
  - Separate IP roles in order to separate between customer data and internal management

Note

Network high availability can be implemented by using NIC-bonding (refer to relevant operating system vendor guidelines for best practices) or by using several data networks in ScaleIO.

For more information about MTU performance considerations and best practices, see the *EMC Fine-Tuning ScaleIO Performance Technical Notes*.

The following table describes the range of potential IP address configurations:

Table 7 IP address configurations in ScaleIO (based on CSV file)

Column in CSV file	MDM Mgmt IP	MDM IPs	SDS All IPs	SDS-SDS Only IPs	SDS-SDC Only IPs
Comments	Management Access	Control Network	Rebuild and Data Path Network	Rebuild Network	Data Path Network

Table 7 IP address configurations in ScaleIO (based on CSV file) (continued)

Column in CSV file	MDM Mgmt IP	MDM IPs	SDS All IPs	SDS-SDS Only IPs	SDS-SDC Only IPs
	Optional, but recommended; not applicable for Tie-Breaker IP addresses that can be used to provide access to ScaleIO management applications, such as CLI, GUI, REST API, OpenStack. This IP address must be externally accessible.	Mandatory IP addresses used for MDM control communications with SDSs and SDCs, used to convey data migration decisions, but no user data passes through the MDM. Must be on the same network as the data network. Must be externally accessible if no MDM Management IP addresses are used.	IP addresses used for both SDS-SDS and SDS-SDC communications. These IP addresses will also be used to communicate with the MDM	IP addresses used for SDS-SDS communication only. These addresses are used for rebuild & rebalance operations. These IP addresses will also be used to communicate with the MDM.	IP addresses used for SDS-SDC communication. These addresses are only used for read-write user data operations.

The following combinations can be used for SDS/SDC:

- Only *SDS All IPs*
- Only *SDS-SDS Only IPs* + *SDS-SDC Only IPs*
- *SDS All IPs* + either *SDS-SDS Only IPs* or *SDS-SDC Only IPs* (can be used in cases of multiple networks; ensure that you do not use the same IP address more than once in the networks).
- *SDS All IPs* + both *SDS-SDS Only IPs* and *SDS-SDC Only IPs* (can be used in cases of multiple networks; ensure that you do not use the same IP address more than once in the networks).

Note

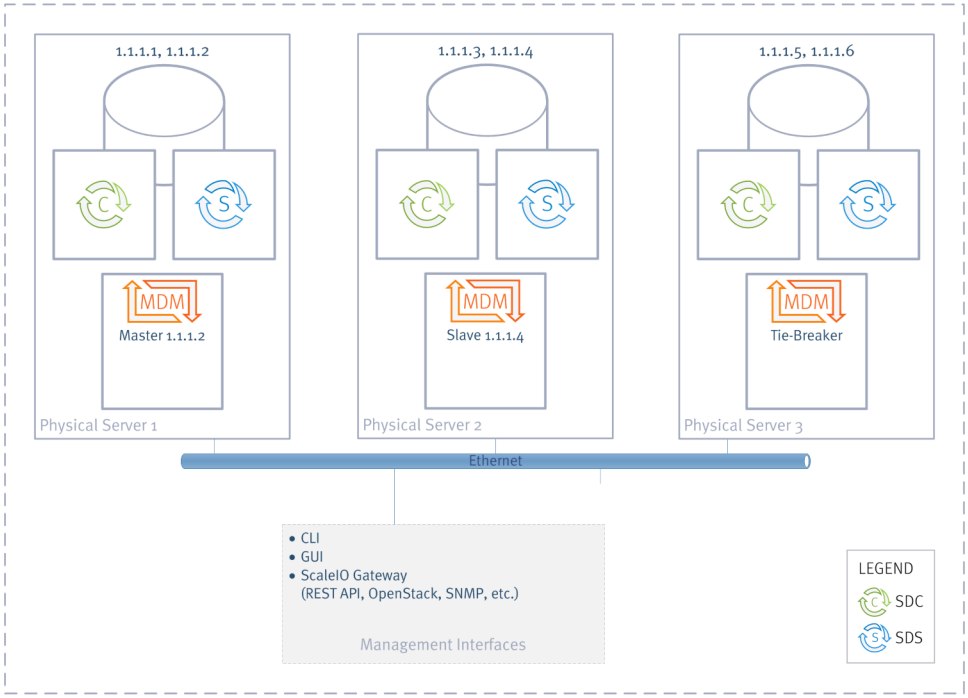
On Windows and Linux, only the MDM needs a management IP address. On VMware, all ScaleIO VMs need to have a management IP address as well as another address for the data network, the network on which traffic flows between SDSs and SDCs for read/writes, rebuild, and rebalance.

In the following example drawing for separate networks, a very simple example is shown, where the management and storage parts of the system are on different networks. In more complex configurations, MDMs, SDCs and SDSs can be on separate networks. Up to 8 separate networks per ScaleIO system are supported.

The following figures show example configurations and the corresponding fields in a CSV configuration file:

Figure 4 ScaleIO system deployed on a single network (Windows)

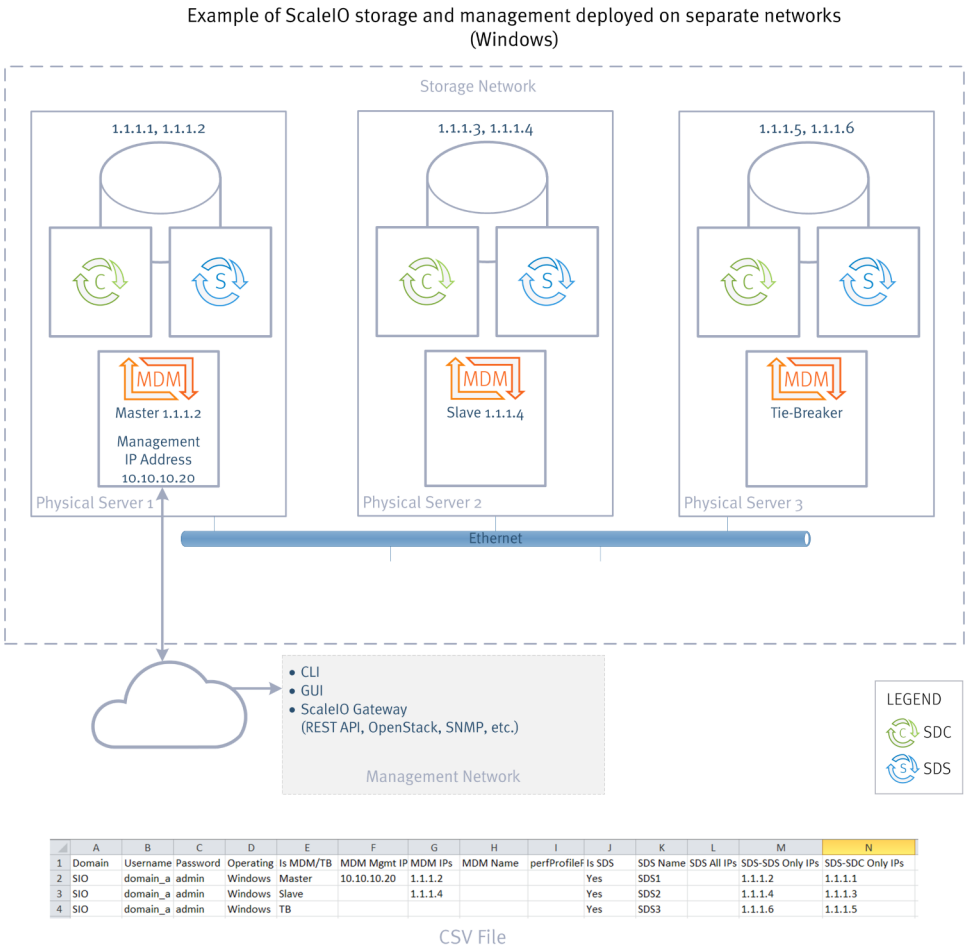
Example of a ScaleIO system deployed on a single network (Windows)



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Domain	Username	Password	Operating	Is MDM/TB	MDM Mgmt IP	MDM IPs	MDM Name	perfProfileF	Is SDS	SDS Name	SDS All IPs	SDS-SDS Only IPs	SDS-SDC Only IPs
2	SIO	domain_a	admin	Windows	Master		1.1.1.2			Yes	SDS1		1.1.1.2	1.1.1.1
3	SIO	domain_a	admin	Windows	Slave		1.1.1.4			Yes	SDS2		1.1.1.4	1.1.1.3
4	SIO	domain_a	admin	Windows	TB					Yes	SDS3		1.1.1.6	1.1.1.5

CSV File

Figure 5 ScaleIO system deployed on separate networks (Windows)



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Domain	Username	Password	Operating	Is MDM/TB	MDM Mgmt IP	MDM IPs	MDM Name	perfProfileF	Is SDS	SDS Name	SDS All IPs	SDS-SDS Only IPs	SDS-SDC Only IPs
2	SIO	domain_a	admin	Windows	Master	10.10.10.20	1.1.1.2			Yes	SDS1		1.1.1.2	1.1.1.1
3	SIO	domain_a	admin	Windows	Slave		1.1.1.4			Yes	SDS2		1.1.1.4	1.1.1.3
4	SIO	domain_a	admin	Windows	TB					Yes	SDS3		1.1.1.6	1.1.1.5

CSV File

VMware limitation:

Multiple IP subnets used for the ScaleIO Data network cannot be on the same subnet in a VMware setup.

For more information, see the VMware limitation in the following link:

http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=2010877

ScaleIO only supports the following network configurations when deployed on VMware:

- A single data storage network
- Two or more data networks, each on separate IP subnets
- A single IP data network using several NIC-bonding configurations, or vSwitch load balancing

Virtual IP Address

Up to four virtual IP addresses can be defined for the MDM cluster. SDCs are then mapped to the MDM cluster's virtual IP addresses, instead of to static MDM IP addresses. MDMs are sometimes switched during normal operation of the cluster, and the virtual IP address will always be mapped to the active MDM. The use of virtual IP

addresses simplifies maintenance procedures on the MDM cluster, because system components communicate via the virtual IP addresses. Therefore, SDCs do not need to be reconfigured when a server hosting an MDM is replaced.

Note

Virtual IP addresses are not currently supported on Windows-based systems.

In new installations in Linux environments, the MDM cluster's virtual IP address can be added and mapped using the Installation Manager CSV file. In VMware environments, virtual IP addresses are mandatory, and configuration is performed using the ScaleIO VMware Installation Wizard, in the Configure SVM stage. The REST API can also be used to add virtual IP addresses to the cluster. In all cases, a virtual IP NIC placeholder must be mapped to each virtual IP address. Ensure that there are NICs available for this purpose.

Existing systems may be extended to include additional MDMs to a cluster. The new MDMs should be mapped to the existing virtual IP addresses.

If virtual IP addresses need to be modified, you must use the CLI or the REST API (not the IM or the vSphere plug-in), and it must be done with extreme caution.

All SDCs will require reconfiguration, to reflect the changes made to the MDM cluster. Otherwise, the SDCs will not be able to communicate with the MDM cluster, and volumes will not be accessible.

Implementing ScaleIO

Implementing a ScaleIO system is, in general, a two-step process: first build the physical storage layer, then configure the virtual SAN layer on top of it.

Physical layer

The physical layer consists of the hardware (servers with storage devices and the network between them) and the ScaleIO software installed on them.

Typically, each SDS is physically located on a separate server, but ScaleIO also supports the installation of multiple SDSs per server.

The Multiple SDS feature allows you to take fuller advantage of the server's computing resources, particularly when no applications are running along-side the SDS. Each SDS on the server is unique, with its own name (for example, sds1, sds2), path, and ports used. This uniqueness provides better control over the SDSs installed. Each SDS can be installed, removed, or upgraded independently. It is up to the user to control the scope of the SDS object. For example, if the SDSs are placed in the same Protection Domain, the user must put them in a single Fault Set.

ScaleIO currently supports four SDSs running on each server. Each SDS is installed using a different RPM. The user can start with a single SDS, and add more SDSs later.

Currently, there is no Windows support for the Multiple SDS feature.

To implement the physical layer, perform the following steps:

1. Install the MDM component on the MDM nodes in one of the following configurations:
   - Three-node redundant cluster (one Master MDM, one Slave MDM, and one Tie-Breaker).
   - Five-node redundant cluster (one Master MDM, two Slave MDMs, and two Tie-Breakers).

- Single node (one master MDM).

  NOTICE

  It is not recommended to use Single Mode in production systems, except in temporary situations. The MDM contains all the metadata required for system operation. Single Mode has no protection, and exposes the system to a single point of failure.

  Note

  MDMs do not require dedicated nodes. They can be installed on nodes hosting other ScaleIO components.

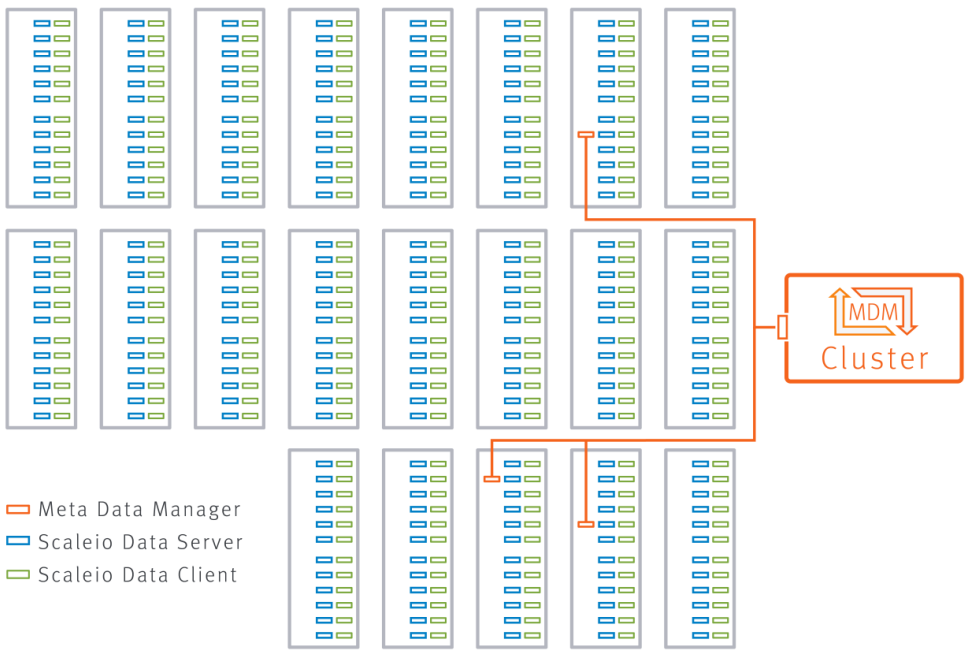
2. Install the SDS component on all nodes that will contribute some, or all, of their physical storage.

   Divide the SDS nodes into Protection Domains. Each SDS can be a member of only one Protection Domain.

   Per Protection Domain, divide the physical storage units into Storage Pools, and optionally, into Fault Sets.

3. Install the SDC component on all nodes on which the application will access the data exposed by the ScaleIO volumes.

   Figure 6 Physical layout example—3-node cluster



Communication is done over the existing LAN using standard TCP/IP. The MDM and SDS nodes can be assigned up to eight IP addresses, enabling wider bandwidth and better I/O performance and redundancy.

You can perform physical layer setup using the following methods:

- ScaleIO Installation Manager a web-client based tool

- ScaleIO VMware plug-in a VMware plug-in
- Manual installation manual installation procedures

  After completing this installation, the physical layer is ready, and can expose a virtual storage layer.

SAN virtualization layer

The MDM cluster manages the entire system. It aggregates the entire storage exposed to it by all the SDSs to generate a virtual layer - virtual SAN storage. Volumes can now be defined over the Storage Pools and can be exposed to the applications as a local storage device using the SDCs.

To expose the virtual SAN devices to your servers (the ones on which you installed and configured SDCs), perform the following:

- Define volumes. Each volume defined over a Storage Pool is evenly distributed over all members using a RAID protection scheme. By having all SDS members of the Storage Pool participate, ScaleIO ensures:
  - Highest and most stable and consistent performance possible
  - Rapid recovery and redistribution of data
  - Massive IOPS and throughput

    You can define volumes as follows:
  - Thick

    Capacity is allocated immediately, even if not actually used. This can cause capacity to be allocated, but never used, leading to wasted capacity.

    Thick capacity provisioning is limited to available capacity.
  - Thin

    Capacity is “on reserve,” but not allocated until actually used. This policy enables more flexibility in provisioning.

    Whereas thick capacity is limited to available capacity, thin capacity provisioning can be oversubscribed, as follows:

    Maximum thin capacity provisioning = 5 * (gross capacity - used capacity)

    When capacity usage reaches the level where it may cause IO errors, alerts are generated. At certain higher capacity levels, volumes (even thin volumes) can no longer be created.

    Example:

    In a system with 3 SDSs, each with 10 TB, there are 30 TB of storage.

    In the system, there is already a thick-provisioned volume that takes up 15 TB of the gross capacity (created by adding a 7.5 TB volume).

    MDM will allow a total of 300 TB gross to be provisioned, and since 15 TB are already allocated, you can add a thin-provisioned volume of 285 TB gross (by adding a 142.5 TB volume) or a thick-provisioned volume of 15 TB gross.
- Map volumes. Designate which SDCs can access the given volumes. This gives rise to the following:
  - Access control per volume exposed
  - Shared nothing or shared everything volumes

Once an SDC is mapped to a volume, it immediately gets access to the volume and exposes it locally to the applications as a standard block device. These block devices appear as `/dev/sciniX` where `X` is a letter, starting from “a.”

For example:

- `/dev/scinia`
- `/dev/scinib`

- The maximum amount of partitions for the scini disk is 15.

- In a Windows environment, the device looks like any other local disk device, as shown in the Device Manager.

  The maximum amount of volumes that can be mapped to an SDC is listed in Table 2, “Product limits.”

Note

SDC mapping is similar to LUN mapping, in the sense that it only allows volume access to clients that were explicitly mapped to the volume.

This is the end of the system setup.

Other functions

ScaleIO includes the following functions:

- EMC Secure Remote Support (ESRS)

  ESRS support enables secure, high-speed, 24x7, remote connection between EMC and customer installations, including:

  - Remote monitoring
  - Remote diagnosis and repair
  - Daily sending of logs, alerts, and ScaleIO topology

- Syslog

  The MDM syslog service can send events, via TCP/IP, to RFC 6587-compliant remote (or local) Syslog servers. Messages are sent with facility local0, by default. Once the syslog service is started, all events will be sent until the service is stopped.

- Get Info

  Get Info assembles a ZIP file of system logs for troubleshooting. You can run this function from a local node for its own logs, using the CLI, or by using the Installation Manager to assemble logs from all MDM and SDS nodes in the system. In addition to the log files, a visual snapshot of the ScaleIO GUI, from the time you perform the operation, can be saved, to better enable support options.

  The Get Info function is described in the *Log Collection Guide*.

- Quality of Service (QoS)

  You can adjust the amount of bandwidth and storage that any given SDC can use. You can configure this with the CLI and the REST interface, on a per client/per volume basis.

- Background Device Scanner

  The Background Device Scanner ("scanner") enhances the resilience of your ScaleIO system by constantly searching for, and fixing, device errors before they

can affect your system. This provides increased data reliability than the media's checksum scheme provides. The scanner seeks out corrupted sectors of the devices in that pool, provides SNMP reporting about errors found, and keeps statistics about its operation.

When a scan is completed, the process starts again, thus adding constant protection to your system.

You can set the scan rate (default: 1 MB/second per device), which limits the bandwidth allowed for scanning, and choose from the following scan modes:

- Device only mode

  The scanner uses the device's internal checksum mechanism to validate the primary and secondary data. If a read succeeds in both devices, no action is taken. If a faulty area is read, an error will be generated.

  If a read fails on one device, the scanner attempts to correct the faulty device with the data from the good device. If the fix succeeds, the error-fixes counter is increased. If the fix fails, a device error is issued.

  Note

  A similar algorithm is performed every time an application read fails on the primary device.

  If the read fails on both devices, the scanner skips to the next storage block.

- Data comparison mode (only available if zero padding is enabled)

  The scanner performs the same algorithm as above, with the following additions:

  After successful reads of primary and secondary, the scanner calculates and compares their checksums. If this comparison fails, the compare errors counter is increased, and the scanner attempts to overwrite the secondary device with the data from the primary device. If this fails, a device error is issued.

  The scanning function is enabled and disabled (default) at the Storage Pool level, and this setting affects all devices in the Storage Pool. You can make these changes at anytime, and you can add/remove volumes and devices while the scanner is enabled.

  When adding a device to a Storage Pool in which the scanner is enabled, the scanning will start about 30 seconds after the device is added.

- AD over LDAP or LDAPS authentication

  User authentication may be done using AD (Active Directory) over LDAP (Lightweight Directory Access Protocol) or LDAPS (Secure LDAP). ScaleIO can support both AD users that are fully controlled through the customer’s existing centralized location, and local users (as has been supported in earlier ScaleIO versions). You can associate groups from the AD with the existing ScaleIO roles in order to ensure the Role-Based Access (RBAC) model. When a user logs on to the ScaleIO system, the MDM identifies that the user belongs to the AD domain, and authenticates the user against the AD server over secured communications. Once the user is authenticated, ScaleIO accepts the group to which the user belongs according to the AD, and associates the appropriate role and its user permissions to that user. The AD implementation is fully redundant.

Note

The authorization permissions of each role are defined differently for local authentication, and for LDAP/LDAPS authentication.

The benefits of using AD over LDAP/LDAPS include:

- Full control of ScaleIO users through the main user repository
- No need to specify a local user for each customer

  If the AD directory is down, the administrator can always use local users to maintain the ScaleIO system.

- Oscillating failure handling

  The Oscillating Failures feature detects and reports various oscillating failures, in cases when components fail repeatedly and cause unnecessary failovers, and therefore disruptions to normal system operation. Typical examples of oscillating failures include:

  - A disk that accepts some I/Os and rejects others
  - A node with interrupted connectivity
  - A node that is constantly busy and therefore handles some I/Os too slowly
  - A disk that is sometimes slow to respond
  - A network that is experiencing disruptions

    The smart detection of such failures provides the ability to handle error situations, and to reduce their impact on normal system operation. Oscillating failure handling can be set for MDMs, SDSs and for SDCs. For SDSs, failure handling can be defined per Protection Domain or per Storage Pool.

  - Oscillating failure counters

    The following table describes the oscillating failure counters:

Oscillating failure counters	Description
(sds_sds/sdc_mdm/sdc_sds/mdm_sds) network_disconnections	Measures the number of network disconnections (socket closed) between two components per IP address
sds_decoupled	Measures the number of times an SDS process is down, as detected by the MDM
sds_configuration_failures	Measures the number of times the MDM fails to configure an SDS, when connecting to an SDS (failures occur during the reconfiguration phase)
sds_receive_buffer_allocation_failures	Measures the number of times an SDS fails to allocate buffer for receiving messages
sdc_long_operations	Measures the number of SDC RPC operations that take longer

Oscillating failure counters	Description
	than the predefined threshold (default threshold is 5 seconds)
sdc_memory_allocation_failures	Measures the number of memory allocation failures in each SDC
sdc_socket_allocation_failures	Measures the number of socket allocation failures in each SDC
sds_device_long_successful_ios	Measures the number of successful IOs to an SDS device, which take longer than the predefined threshold (default threshold is 250 milliseconds)

- Secure connectivity with external components

  This feature allows external components to authenticate the MDM. After authentication, communication between the MDM and external components is performed using TLS (Transport Layer Security) protocols.

  Secure communication with the MDM is authenticated by the following components:

  - CLI client

    Note

    If the secure mode is not enabled, modifications are necessary to run SCLI commands.

  - ScaleIO Gateway
  - GUI client
  - IM client

    Once added in the trust point, all communications will require authentication, followed by communications over TLS. The same method is employed between the IM and all LIAs.

Maintenance

Maintenance of ScaleIO is primarily limited to configuration changes of the physical and virtual layers. It requires minimal user attention. When maintenance or planned restart of an SDS is required, the Maintenance Mode feature can be used to streamline system operation.

Maintaining the physical layer

In the physical layer, maintenance is limited to adding and removing hardware units and configuring them into the ScaleIO system. These operations are usually a result of:

- Scaling out when there is a need for additional capacity. This usually results in adding more storage media to the existing servers, or adding additional servers.
- Hardware failure. In cases where there is a hardware (storage media or server) failure and it needs to be replaced.

In all of the above cases, the operation will require adding or removing storage capacity from the system. In some cases, it may include adding or removing an entire server, and its associated storage media, from the configuration. As far as ScaleIO is concerned, all of these activities translate to SDS reconfigurations.

If the removed server is an SDC node, or the server to be added requires exposing storage locally, SDC reconfiguration will happen as well.

- Adding or removing storage media. Add or remove the media from the SDS with which it is associated. ScaleIO will redistribute the data accordingly and seamlessly.
- Adding or removing a node. Add or remove the SDC and SDS residing on the node. ScaleIO will redistribute the data accordingly and seamlessly.

Instant maintenance mode

Instant Maintenance Mode enables you to restart a server that hosts an SDS, with minimal impact on the ScaleIO system, thus bypassing the disruption and effort caused by disorderly shutdown, Protection Domain shutdown, and orderly shutdown.

Whereas ScaleIO Ready Node always uses two copies of user data, invoking Maintenance Mode introduces an additional copy that stores all writes created during maintenance to an SDS or Fault Set (created during maintenance) in both a primary location and a new location. This copy prevents data loss if a single failure occurs.

When the SDS or Fault Set is returned from Maintenance Mode, only the new writes are required to be resynchronized, thus minimizing data transfer during and after the update.

Instant Maintenance Mode does not interrupt application IOs; it can be run on any amount of members of a Fault Set; and it can run in parallel on different Protection Domains. While an SDS is in Maintenance Mode, most ScaleIO operations (like adding a volume) cannot be performed in the Fault Set, Protection Domain, or Storage Pool in which the SDS and its devices reside.

To invoke maintenance mode, the following conditions are required:

- Only one Fault Unit (or standalone SDS) can be in Maintenance Mode at any given time.
- No other SDSs can be in degraded or failed state (force override can be used).
- There must be adequate space on other SDSs for the additional backup (force override can be used).

Note

Use of force override options when entering Maintenance Mode can lead to data unavailability while Maintenance Mode is activated.

While an SDS is in Maintenance Mode, it can be shut down with no danger to data.

Maintaining the virtualization layer

The following operations may be performed on volumes that are exposed by the ScaleIO virtual SAN:

- Add or remove a volume:

  Create or delete a volume in the system.
- Increase volume size:

Add capacity to a given volume, as needed. The change in volume size occurs seamlessly without interrupting I/O.

- Map and unmap volumes to an SDC:

  This enables or disables access to a volume by an SDC, and thus by an application residing on the same node.

S.M.A.R.T. hardware monitoring

The ScaleIO bare-metal solution now provides monitoring capabilities for RAID controllers and storage devices compatible with S.M.A.R.T. (Self-Monitoring, Analysis and Reporting Technology) protocols.

In Linux-based environments, S.M.A.R.T.-compatible HDDs, SSDs and RAID storage controllers can be monitored for S.M.A.R.T. attributes such as temperature, SSD wear level, and error counters. LEDs can also be lit on these hardware devices, to simplify physical identification for maintenance purposes.

Each hardware vendor defines specific thresholds for the S.M.A.R.T. attributes. This feature currently supports storage devices controlled by LSI, HP and Dell RAID controllers, and stand-alone devices. During system deployment, an external monitoring tool is installed as part of the LIA on each node. Additional RAID controller tools must be installed manually after system deployment: storcli for LSI RAID controllers, hpssacli for HP RAID controllers, or perccli for DELL RAID controllers. These tools are used by the system to collect the counters that are returned to the MDM.

Note

In some cases, LSI RAID controllers may report vendor information as "AVAGO" instead of LSI.

The MDM queries the SDSs at set intervals, and stores the returned information. This information can be viewed using CLI queries. In addition, when thresholds are crossed for S.M.A.R.T. attributes, alerts are generated by the system.

When the CLI is used to query device information, physical device information, such as serial number, model name, vendor etc., temperature, and wear level information (for SSDs only) is included in the returned response.

The following CLI commands can be used for hardware monitoring:

- `scli --set_sds_device_led` - turns the device LED on or off, to allow you to physically identify a device in the chassis
- `scli --query_sds_device_info` - displays device data

For information about the use of CLI commands, see the *ScaleIO CLI Reference Guide*.

The following URIs can be used to query devices using the REST API, in order to view S.M.A.R.T.-related information:

- `api/instances/Device::{id}`
- `/api/instances/Sds::{id}/relationships/Device`

For information about the use of REST API URIs, see the *ScaleIO User Guide*.

You can use the GUI to monitor S.M.A.R.T.-related alerts in the **Alerts** view.

In addition, SNMP traps and ESRS alert codes can be used to monitor alerts triggered by devices compatible with S.M.A.R.T.

List of approved RAID controllers

Provides high-level specifications of RAID controllers, which are tested and certified by ScaleIO.

ScaleIO-certified RAID controllers

The following table describes the ScaleIO-certified RAID controllers:

Manufacturer	Specifications
HP	• Model Name: Smart Array P440ar • Vendor Name: HP • Firmware Version: 3.56 • Driver Version: 3.4.10 • Driver Name: hpsa • PCI Address: 0000:03:00.0
DELL	• Model Name: PERC H730 Mini • Vendor Name: Dell • Firmware Version: 25.3.0.0016 • Driver Version: 06.807.10.00-rh • Driver Name: megaraid_sas • PCI Address: 00:02:00:00
LSI	• Model Name: LSI MegaRAID SAS 9271-8i • Vendor Name: LSI • Firmware Version: 23.12.0-0021 • Driver Version: 06.810.09.00-rh • Driver Name: megaraid_sas • PCI Address: 00:82:00:00
	• Model Name: LSI MegaRAID SAS 9271-8i • Vendor Name: LSI • Firmware Version: 23.12.0-0018 • Driver Version: 06.805.06.01-rc • Driver Name: megaraid_sas • PCI Address: 00:82:00:00
	• Model Name: LSI MegaRAID SAS 9271-8i • Vendor Name: LSI • Firmware Version: 23.12.0-0021 • Driver Version: 06.805.06.01-rc • Driver Name: megaraid_sas • PCI Address: 00:82:00:00

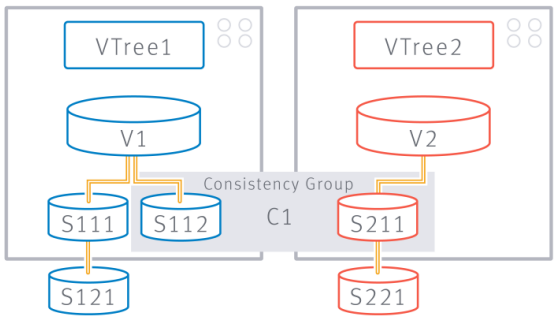
Manufacturer	Specifications
	• Model Name: LSI MegaRAID SAS 9271-8i • Vendor Name: LSI • Firmware Version: 23.34.0-0005 • Driver Version: 06.810.08.00 • Driver Name: megaraid_sas • PCI Address: 00:82:00:00

Snapshots

The ScaleIO storage system enables you to take snapshots of existing volumes, up to 31 per volume. The snapshots are thin provisioned and are extremely quick. For more information about thin provisioning, see "SAN virtualization layer".

Once a snapshot is generated, it becomes a new, unmapped volume in the system. You can manipulate it in the same manner as any other volume exposed by the ScaleIO storage system.

Figure 7 Snapshot operations



The structure related to all the snapshots resulting from one volume is referred to as a VTree (short for Volume Tree). It is a tree spanning from the source volume as the root, whose siblings are either snapshots of the volume itself or descendants of it. Thus, some snapshot operations are related to the VTree, and may affect parts of it. In Figure 8 in BLUE, S111 and S_{112} are snapshots of V1. S121 is a snapshot of snapshot S111. Together, V1 and S1xy are the VTree of V1.

When taking a snapshot in the system, you can specify more than one volume. All snapshots taken together form a consistency group. They are consistent in the sense that they were all taken at the same time. So if there is a contextual relationship between the data contained by all the snapshot members, then that set is meaningful. The consistency group allows manipulation of the entire set.

If you remove an entire consistency group, all of the snapshots that were taken together will be removed. Back to Figure 8, in RED, S211 is a snapshot of V2. Since S112 and S211 were taken together, they compose a consistency group (in ORANGE) designated as C1.

Note

The consistency group is only for convenience purposes. There are no protection measures done by ScaleIO to conserve the consistency group. For example, you can remove a snapshot that is a member of a consistency group.

Snapshot operations

The following operations are snapshot related. Since all snapshots are volumes in the system, all volume operations are applicable to snapshots.

- Take a snapshot of a volume

  The `snapshot_volume` command allows a snapshot of a given volume. When specifying more than one volume (a list), a consistency group is generated.
- Map volume to an SDC
  The `map_volume_to_sdc` command maps the snapshot volume to an SDC and thus exposes it to the applications. This is no different than any other volume mapping operation.
- Remove a volume

  The `remove_volume` command affects snapshots in several ways. Since snapshots are volumes, the way to delete a snapshot is exactly the same as removing a volume from the system.

Management tools

You can provision, maintain, and monitor ScaleIO with the following management clients:

- Command Line Interface (CLI)

  The CLI enables you to perform the entire set of configure, maintain, and monitor activities in a ScaleIO system.
- Graphical User Interface (GUI)

  The GUI enables you to perform standard configure and maintain activities, as well as to monitor the storage system's health and performance. You can use the GUI to view the entire system, and then drill down to different elements.
- VMware plug-in (plug-in)

  The plug-in enables you to perform basic provision and maintain activities in the VMware environment. In addition, the plug-in provides a wizard to deploy ScaleIO in the VMware environment.
- OpenStack

  ScaleIO provides Cinder and Nova drivers, which enable interoperation between a ScaleIO system and an OpenStack cloud operating system.
- REST Gateway

  A REST API can be used to expose monitoring and provisioning via the REST interface. The REST server is installed as part of the ScaleIO Gateway.

  Many ScaleIO activities can be performed in more than one management tool.

  The following tool is also provided:
- Installation Manager (IM)

The IM is used for installing ScaleIO, upgrading and uninstalling components, as well as running the get-info operation. The IM is installed as part of the ScaleIO Gateway.

Configuring direct attached storage (DAS)

ScaleIO works with any free capacity—internal or direct-attached devices, either magnetic hard disk drives (HDD) or flash-based devices such as solid state drive (SSD) and PCIe cards. Although ScaleIO can work with any device topology, it is recommended to configure the raw devices as stand-alone devices.

NOTICE

SDS devices are not write-protected; their contents can be erased or formatted. Exercise caution when using disk utilities.

If the server has a RAID controller, ScaleIO prefers to use the controller's caching abilities for better performance, but is better utilized when all devices are configured as stand-alone (i.e. setting each of the devices to RAID-0 separately). For HDD devices, it is recommended to enable RAID-controller caching. As for flash devices, it depends on the device behavior.

For Windows, when using a physical disk drive, it is recommended to generate a single, unformatted partition over the entire disk.

For more information about preparing Windows devices, see the "Adding devices to SDS nodes on Windows servers" section in the *EMC ScaleIO Deployment Guide*.

Note

For HDDs: It is recommended to use RAID-controller caching when available as follows:

- READ/WRITE: if cache is battery-backed
- READ ONLY: if cache is NOT battery-backed

For flash devices (e.g. SSD): Depends on the device

Implementing ScaleIO over a virtual system

This section provides an overview of how ScaleIO is implemented in a virtualized environment.

VMware

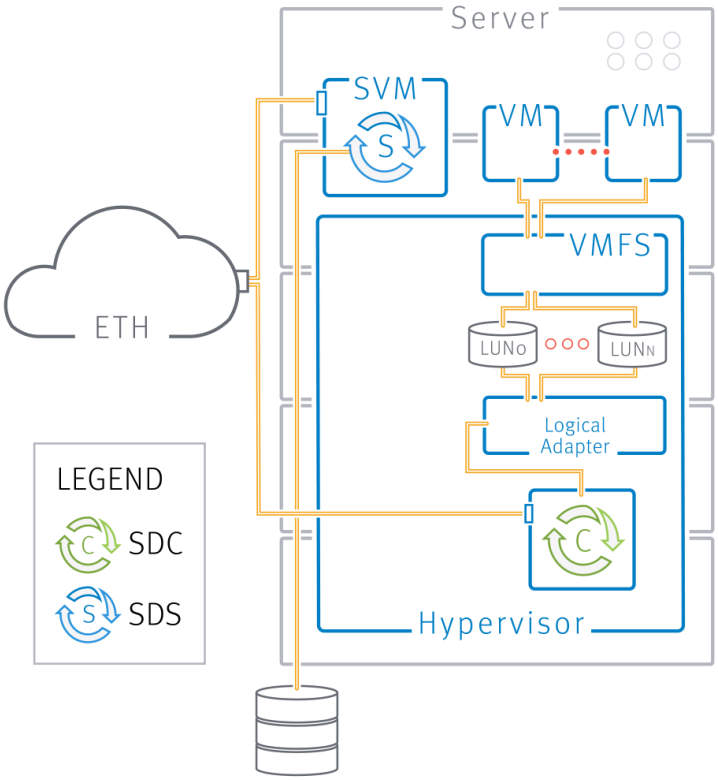
In the VMware environment, the MDM and SDS components are installed on a dedicated SVM, whereas the SDC is installed directly on the ESX host.

Note

Installing the SDC on the ESX host requires a restart of the ESX host.

This implementation is illustrated in the following figure:

Figure 8 ScaleIO implementation on ESX



Note

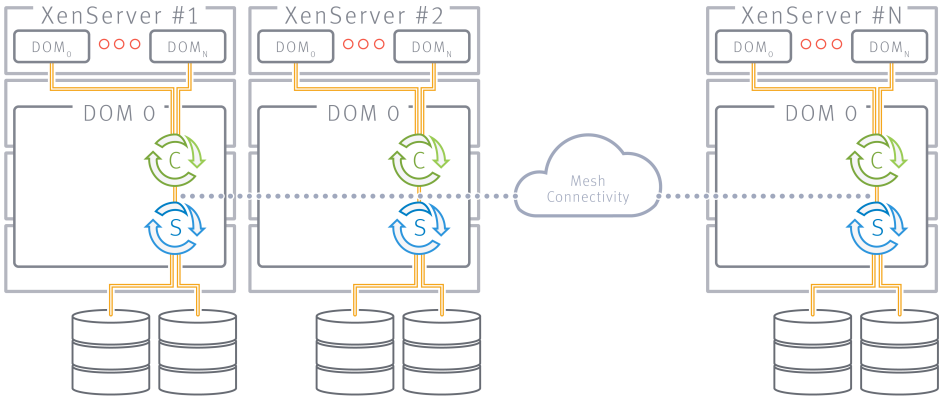
The LUNs in the previous figure can be formatted with VMFS, and then exposed using the ESXi host to the virtual machine, or can be used as RDM devices. When the LUNs are used as RDM devices, the VMFS layer is omitted.

Installation in a VMware environment is enabled via the VMware plug-in.

Xen implementation

In a Xen environment, both the SDC and SDS are installed in Dom0 as would be on a physical node. Dom0 accesses the storage media through the SDS and exposes volumes based on ScaleIO through the SDC.

Figure 9 ScaleIO Xen virtual machine architecture



Information on provisioning in a Xen environment is described in this guide.

PART 2

Deploying ScaleIO Systems

This part describes how to deployScaleIO in your environment. Chapters include:

Chapter 3, "Deploying ScaleIO on Physical Servers"

Chapter 4, "Deploying ScaleIO on ESX Servers"

Chapter 5, "Post-Deployment Activities"

CHAPTER 3

Deploying ScaleIO on Physical Servers

This chapter describes how to install, deploy, and perform initial configuration of ScaleIO software components on physical servers. Topics include:

- Overview of ScaleIO deployment on physical servers 62
- Preparing the Installation Manager and the Gateway 63
- Installing with the full Installation Manager 67
- Installing with the Installation Manager wizard 83
- Deploying on Ubuntu, Oracle Linux, or CoreOS servers 90
- Deploying on OpenStack 96
- Installing the ScaleIO GUI 100
- Configuring Installation Manager properties 100

Overview of ScaleIO deployment on physical servers

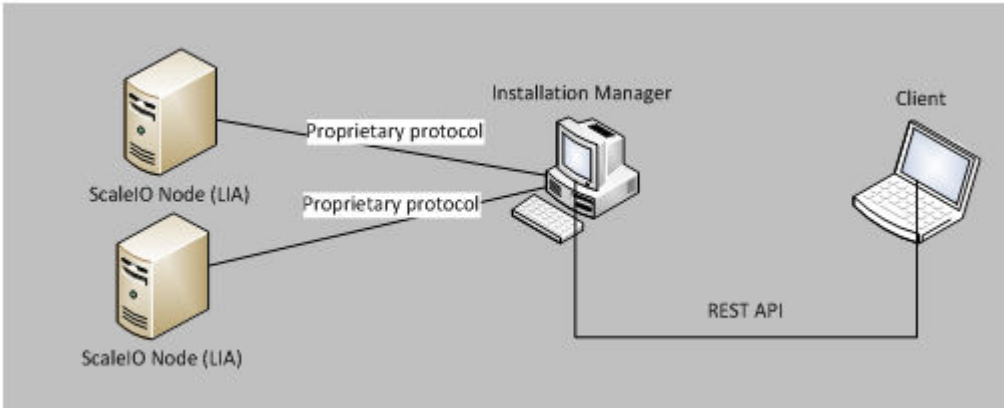
This section describes an overview of ScaleIO deployment on physical servers.

Note

The ScaleIO installation enables unlimited use of the product in non-production environments. To obtain a license for production use, and to receive technical support, open a service ticket with EMC Support at https://support.emc.com.

For complete information on licensing, see the *EMC ScaleIO ScaleIO User Guide.*

In physical environments, and in non-VMware virtual environments (Xen, Hyper-V, and KVM), you use the ScaleIO Installation Manager (IM) to install and configure the ScaleIO components on multiple nodes from one central server, via a web client.



The Installation Manager can be used to install numerous ScaleIO systems, from one central workstation.

Note

The IM can be disabled and configured, as described in "Configuring the Installation Manager".

As part of the installation process, the ScaleIO Lightweight Installation Agent (LIA) is installed. The LIA is a required component for many maintenance operations. During initial installation, the IM establishes trust with the LIA, which facilitates its operation.

Note

To make changes to the LIA configuration file, see "Changing the LIA configuration file".

You can limit the operations that LIA can perform by making changes to its configuration file. In Windows environments, LIA will only install and upgrade packages that are signed by ScaleIO and associated with the ScaleIO system.

The Installation Manager has a REST API that enables install, extend, and uninstall functionality. These features are described in the *EMC ScaleIO User Guide*.

Linux package names

ScaleIO Linux installation package name formats.

Package formats

Throughout this section, Linux packages are displayed in the following format:

`EMC-ScaleIO-<component>-2.0-13000.X.<flavor>.x86_64.rpm`

where:

- *<component>* is the ScaleIO component: mdm, lia, sds, etc.
- *<flavor>* is the OS for your environment, according to the following table:

Table 8 Linux package formats

Linux flavor	Package format	Example
CoreOS	CoreOS	`EMC-ScaleIO-mdm-2.0-13000.X.CoreOS.x86_64.tar` CoreOS packages may need to be extracted before use. This is described where relevant.
RHEL6/OL6/CentOS	el6	`EMC-ScaleIO-mdm-2.0-13000.X.el6.x86_64.rpm`
RHEL7/OL7/CentOS7	el7	`EMC-ScaleIO-mdm-2.0-13000.X.el7.x86_64.rpm`
SLES11 sp3	sles11.3	`EMC-ScaleIO-mdm-2.0-13000.X.sles11.3.x86_64.rpm`
SLES12	sles12.0	`EMC-ScaleIO-mdm-2.0-13000.X.sles12.0.x86_64.rpm`
SLES12 sp1	sles12.1	`EMC-ScaleIO-mdm-2.0-13000.X.sles12.1.x86_64.rpm`
Ubuntu14.04	Ubuntu.14.04	`EMC-ScaleIO-mdm-2.0-13000.X.Ubuntu.14.04.x86_64.tar` Ubuntu packages may need to be extracted before use. This is described where relevant.
Ubuntu16.04	Ubuntu.16.04	`EMC-ScaleIO-mdm-2.0-13000.X.Ubuntu.16.04.x86_64.tar` Ubuntu packages may need to be extracted before use. This is described where relevant.
Xen 6.5/Xen 6.5 SP1	xs6.5.0	`EMC-ScaleIO-mdm-2.0-13000.X.xs6.5.0.x86_64.rpm`
Xen 7.0	xs7.0.0	`EMC-ScaleIO-mdm-2.0-13000.X.xs7.0.0.x86_64.rpm`

Use the packages and the installation commands that match your Linux operating system environment.

On Xen servers, the syntax for ScaleIO CLI commands is `siocli`, as opposed to `scli`.

Note

Before upgrading CoreOs, Ubuntu, or Oracle Linux (OL) systems, ensure that the ScaleIO environment is prepared, as described in Deploying on Ubuntu, Oracle Linux, or CoreOS servers on page 90.

Preparing the Installation Manager and the Gateway

Before installing ScaleIO, you must prepare the Installation Manager (IM). The Installation Manager is installed as part of the ScaleIO Gateway. The Gateway also

includes the REST Gateway and SNMP trap sender functionality. In this document, the terms IM server and Gateway server are synonymous.

You can enable and disable Gateway components, as described in "Enabling and disabling Gateway components".

The Gateway can be installed on the same node as other ScaleIO components. Ensure that the node has adequate memory to run the IM (minimum of 3 GB) and any other applications.

To install ScaleIO on Linux nodes, the IM server can be a Windows or Linux server. To install ScaleIO on Windows nodes, or to install in a mixed-OS environment, your IM server must be a Windows server. The IM server must have connectivity to both the data and management ScaleIO networks.

Ensure that your IM server meets the system requirements described in System requirements on page 18.

The Gateway cannot be installed on a CoreOS server. In CoreOS environment, deploy the Gateway on a supported Windows or Linux server.

Preparing the IM on a Linux server

This section describes how to prepare the Installation Manager (IM) on a Linux server. You can use a Linux IM to deploy to Linux servers only.

Procedure

1. Download and extract the installation files needed for your operating system.

   This includes at a minimum the files for the ScaleIO Gateway and components.

   Note

   Alternatively, you can download the Complete Software Download, which contains all files for all operating systems.

   You may also want the files to install the GUI. You need the Components ZIP file for your Linux operating system as well as the Gateway for Linux. The GUI for Linux is a separate file.

   You can download all files from EMC Online Support (https://support.emc.com).

   Note

   To install with a customized Java configuration, see "Using a custom Java configuration for the Gateway".

2. From the extracted download file, install the ScaleIO Gateway on the Linux IM server, by running the following command (all on one line):

   - RHEL/Centos/Oracle Linux/SLES

```
GATEWAY_ADMIN_PASSWORD=<new_GW_admin_password> rpm -i /tmp/
EMC-ScaleIO-gateway-2.0-13000.X.x86_64.rpm (for SLES 11.3
add --nodeps)
```

- Ubuntu

```
GATEWAY_ADMIN_PASSWORD=<new_GW_admin_password> dpkg -i /tmp/
EMC-ScaleIO-gateway-2.0-13000.X.amd64.deb
```

where *<new_GW_admin_password>* is a password that you define to access the IM.

To install without assigning a Gateway password, see "Installing the Gateway without assigning an admin password".

Note

To include special characters in the password, such as ! @ $, when assigning the password, add the "\" character before the special character. For example, to assign the password Scaleio$$, enter it as Scaleio\$\$. When logging in, the "\" character is not used.

Note

For installation on Ubuntu, Oracle Linux, or CoreOS servers, read this section before continuing with the deployment: "Deploying on Ubuntu servers".

3. When installing on XEN servers, add the following changes to the `/opt/emc/scaleio/gateway/conf/server.xml` file:
   a. Edit the `Connector port="80" protocol="HTTP/1.1` class, as follows:

      Change `redirectPort` from **443** to **8443**.

   b. Edit the `Connector port="443" protocol="org.apache.coyote.http11.Http11NioProtocol"` header, as follows:

      Change the `Connector port` from **443** to **8443**.

   c. Restart the scaleio-gateway.

After you finish

Continue with your choice of IM methods:

- To use the IM wizard, skip to "Installing with the Installation Manager wizard".
- To use the full IM to install ScaleIO and the replication splitter for RecoverPoint, continue with "Installing the replication splitter for RecoverPoint".

  Note

  Replication support is version-specific; see the ESSM for full details.

- To use the full IM to install ScaleIO only, continue with "Installing with the full Installation Manager".

Note

You can configure the Installation Manager, as described in "Configuring the Installation Manager"

Installing the replication splitter for RecoverPoint

Note

If you are not installing the replication splitter for RecoverPoint, skip to "Installing with the full Installation Manager".

This topic describes steps that are required when using the IM to install the replication splitter for RecoverPoint.

Note

Replication support is version-specific; see the ESSM for full details.

There are additional post-installation configurations steps. For full instructions, refer to the EMC ScaleIO Write Splitter for RecoverPoint Technical Notes before beginning the installation.

To install the replication splitter for RecoverPoint, you must do the following:

1. Enable zero-padding on the Storage Pools, as described in "Configuring the Installation Manager".
2. In the installation CSV file:
   - Add the **Splitter RPA IP** column, and enter the information, as described in "Preparing the CSV topology file".
   - In the **Optimize IOPS** column, set the value to `Yes`.
3. Upload the splitter package to the IM server.

   Continue with "Installing with the full Installation Manager".

Preparing the IM on a Windows server

This section describes how to prepare the Installation Manager (IM) on a Windows server. You can use a Windows IM to deploy to Windows and Linux servers.

Don't install the Gateway on a server on which RFcache (the xcache package) will be installed.

Procedure

1. Download and extract the installation files needed for your operating system.

   This includes at minimum the files for the ScaleIO Gateway and components. You may also want the files to install the GUI.

   All ScaleIO files are located in the `Complete Windows Software Download` ZIP file.

   Note

   To use this Gateway to deploy ScaleIO on Linux nodes, too, you should also download the ScaleIO components for the relevant Linux OS.

   You can download all files from EMC Online Support (https://support.emc.com).

2. From the extracted download file, copy the ScaleIO Gateway MSI to the IM server:

   `EMC-ScaleIO-gateway-2.0-13000.X-x64.msi`

3. Run the file, and enter a newGateway Admin password that will be used to access the IM.

Note

To add more devices than the amount of available drive letters, and for other options, see "Adding devices to SDS nodes on Windows servers".

---

4. Prepare disks for storage:
   a. Ensure that devices to be used are on-line, and initialized as MBR or GPT.
   b. On each disk, use the Windows Disk Management to create a new, simple volume.

      Assign a drive letter, and select **Do not format this volume**.
5. Choose which IM mode to use:
   - To use the full IM, continue with "Installing with the full Installation Manager".
   - To use the IM wizard, skip to "Installing with the Installation Manager wizard".

Note

You can configure the Installation Manager, as described in "Configuring the Installation Manager"

Which installation mode is best for you?

The Installation Manager can be used in the following modes:

- Full IM

  This mode, which enables the highest level of customization, uses a user-prepared CSV topology file to install ScaleIO and configure nodes. The CSV file is used by the IM to set up and configure all the nodes. To add additional servers after the initial installation, you will also use the combination of CSV topology file and IM.
- IM wizard

  This abbreviated mode enables you to get a ScaleIO system up and running in the simplest manner, with preset node configuration, where all management and data communication are on the same network. This mode is perfect for a single Protection Domain, fully-converged system. No CSV file is required.

  After preparing the IM, you can choose which mode to use.

  To prepare the ScaleIO Installation Manager (IM), proceed to the section that matches the operating system of your IM server:
- "Preparing the IM on a Linux server"
- "Preparing the IM on a Windows server"

Installing with the full Installation Manager

This section describes how to use the full Installation Manager to install ScaleIO components. To use the IM wizard, go to "Installing with the Installation Manager wizard".

Note

ScaleIO installation enables unlimited use of the product, in non-production environments. To obtain a license for production use, and to receive technical support, open a service ticket with EMC Support at https://support.emc.com.

For complete information on licensing, see the *EMC ScaleIO User Guide.*

To use the full Installation Manager, you must first prepare a CSV topology file.

Preparing the CSV topology file

You can edit a CSV file with Excel or file-editing software. In this document, we will refer to and illustrate the CSV file as a spreadsheet.

The following CSV templates are provided as part of the software download, in the Gateway software packages:

- Complete

  This spreadsheet template contains all available fields, both required and optional.
- Minimal

  This spreadsheet template contains only the required fields. The optional fields (those that are in the complete spreadsheet, but not in the minimal one) will be assigned default values.

Note

The CSV can also be used for removing installed components.

Procedure

1. Fill in your site-specific information in the appropriate places, overwriting the default information provided in the file.

   You only need to use one spreadsheet for the installation, as follows:

   - To manually enter all configuration details, use the Complete spreadsheet.
   - To use default values for the non-mandatory fields, use the Minimal spreadsheet.
   - To configure non-default values for columns that are not in the Minimal spreadsheet, either use the Complete spreadsheet, or copy the column heading from there into the minimal spreadsheet and enter your custom values into the minimal spreadsheet.

   The following figure illustrates part of the Complete CSV file, to add Linux and Windows nodes:

Figure 10 CSV—complete

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Domain	Username	Password	Operating System	Is MDM/TB	MDM Mgmt IP	MDM IPs	MDM Name	perfProfile ForMDM	Virtual Ips	Virtual IP NICs	Is SDS	SDS Name	SDS All IPs	SDS-SDS C	SDS-SDC C	Protectior	Fault Set
2		root	Password1	linux	Master	10.76.1.1	10.76.60.195	MDM1	Default	10.76.60.203	eth1	Yes		10.76.60.195			domain1	fs1
3		root	Password1	linux	Slave	10.76.1.2	10.76.60.17	MDM2	Default	10.76.60.203	eth1	Yes		10.76.60.17			domain1	fs1
4		root	Password1	linux	Slave	10.76.1.3	10.76.60.18	MDM3	Default	10.76.60.203	eth1	Yes		10.76.60.18			domain1	fs1
5		root	Password1	linux	TB		10.76.60.33	TB1	Default			Yes	SDS1	10.76.60.33	10.76.2.2,:	10.76.3.3	domain1	fs2
6	localhost	administrator	Password1	windows	TB		10.76.60.19	TB2	Default			No						
7		root	Password	linux								Yes-1	MultipleSds_1	10.76.60.55			domain1	fs3
8		root	Password1	linux	Standby-Slave	10.76.1.4	10.76.60.20	MDM4	Default	10.76.60.203	eth1	No						
9	localhost	administrator	Password1	windows	Standby-TB		10.76.60.21	TB3	Default			No						
10		root	Password	linux								Yes-2	MultipleSds_2	10.76.60.55			domain1	fs3
11		root	Password	linux								Yes-3	MultipleSds_3	10.76.60.55			domain1	fs3
12		root	Password	linux								Yes-4	MultipleSds_4	10.76.60.55			domain1	fs3

2. The following are the differences between Linux and Windows:

a. For Linux nodes, the **Domain** column is not relevant.

b. Multiple SDS is supported on Linux nodes only.

The following table describes the fields in the spreadsheets. The required fields appear in both spreadsheets. All field names are case-sensitive; the order of the columns is not significant.

Table 9 CSV topology spreadsheets

Field	Description	Required
Domain	If using a domain user, the name of the domain (not relevant for Linux)	
Username	The name of the user. In Linux, this value is always `root`; in Windows, a user with administrator rights (default: `administrator`).	
IPs	IP address of the node. **Note** Entering IP addresses in this column will use the same IP addresses for management and data IP addresses for all nodes. To separate the IP roles, leave this column blank, and use the **MDM Mgmt IP** and **MDM IPs** columns instead.	Yes[a]
Password	Password used to log in to the node. In Linux, the root password. In Windows, the password of a user with administrator rights. **Note** To authenticate with SSH instead of node passwords, see "Using SSH authentication on the ScaleIO Gateway".	Yes
Operating System	The server's OS: `linux` or `windows`	Yes
Is MDM/TB	`Master, Slave, TB, Standby-Slave, Standby-TB`, or `blank` (if not an MDM)	Yes
MDM Mgmt IP	The IP address for the management-only network	
MDM IPs	MDM IP addresses used to communicate with other ScaleIO components in the storage network. Comma-separated, no spaces.	
MDM Name	The name to be assigned to an MDM node. MDM names must be unique. Ensure that each server on which MDM is installed has a unique hostname.	
perfProfileForMDM	Optional performance profile to set for MDMs. `High` or `Default` (same as leaving this field empty).	
Virtual IPs	A virtual IP address for each possible manager MDM, to be used for communication between the MDM cluster and SDCs. Only one virtual IP address can be mapped to each NIC, with a maximum of four virtual IP addresses per system. Virtual IP addresses are not supported on nodes using Windows operating system.	
Virtual IP NICs	The NIC to which the virtual IP addresses are mapped.	
Is SDS	• `Yes` or `No` • To configure multiple SDSs on a server, see "Configuring multiple SDSs on one server".	Yes
SDS Name	The name for the SDS node	

Table 9 CSV topology spreadsheets (continued)

Field	Description	Required
SDS All IPs	SDS IP addresses to be used for communication among all nodes. Comma-separated, no spaces.	
SDS-SDS Only IPs	SDS IP addresses to be used for communication among SDS nodes only. Comma-separated, no spaces.	
SDS-SDC Only IPs	SDS IP addresses to be used for communication among SDS and SDC nodes only. Comma-separated, no spaces. **Note** For SDC-only nodes, enter the IP in this column.	
Protection Domain	The Protection Domain to which to assign this SDS.	
Fault Set	The Fault Set to which to assign this SDS. **Note** When using Fault Sets, you must assign a minimum of three, and follow the guidelines described in "Fault Sets".	

SDS Device List	• The devices to add to the SDS, for example: Linux `/dev/sdb,/dev/sdc` or Windows `d,e,f`. Comma-separated, no spaces.[b]
	• Balance the total device capacity over all SDSs.
	• ScaleIO might not perform optimally if there are large differences between the sizes of the devices in the Storage Pool, for example, if one device is as big as the rest of the devices. If in doubt, contact ScaleIO support.
	• When adding devices that were used in a previous ScaleIO system, follow the instruction in "Adding SDS devices that were used in a previous ScaleIO system".
	Note
	Before adding Windows SDS devices see "Adding devices to SDS nodes on Windows servers".
	• When specifying the SDS device path on a Linux node, use the path according to how it is listed in `cat/proc/partitions` (and not according to the output of `fdisk -l`)
	For example:
	`fdisk` output: `/dev/mapper/samah-lv1, /dev/sdb`
	`cat /proc/partitions` output: `dm-3, sdb`
	• Use these values in the CSV: `/dev/dm-3,/dev/sdb`	

StoragePool List	Sets Storage Pool names. When using this field, the Storage Pools will be appropriated to the devices in the device list, respectively. The amount of Storage Pools must equal the amount of devices. Comma-separated, no spaces.	

Table 9 CSV topology spreadsheets (continued)

Field	Description	Required
	Note If no Storage Pool is set, one will be automatically created during installation, named `default`.	
SDS Device Names	Sets user-defined device names. When using this field, the device names will be appropriated to the devices in the device list, respectively. The amount of names must equal the amount of devices. Comma-separated, no spaces.	
RFCache	Install the xcache package on an SDS: `Yes` or `No` (default). If you are installing xcache on a Windows server, a server restart is necessary. **Note** If you wish to configure RFCache devices as part of the deployment, set this column to `Yes`, **RFCache Device List** column to a list of comma-separated devices, and the **RFCache Pool List** to a corresponding comma-separated list of Storage Pools.	
RFCache Device List	List of devices to provide RFcache acceleration on an SDS. Up to eight devices, comma-separated, with no spaces. If RFcache is `Yes`, and this field is left blank, you can add RFcache devices after installation.	
RFCache Pool List	The name of the Storage Pool for which RFcache acceleration is to be provided. When using this field, the RFcache pools will be appropriated to the devices in the RFcache device list, respectively. The amount of names must equal the amount of devices. Comma-separated, no spaces. If there are values for the RFcache Device List, there must be values for this field, too.	
perfProfileForSDS	Optional performance profile to set for SDSs. `High` or `Default` (same as leaving this field empty).	
Is SDC	`Yes` or `No`	Yes
perfProfileForSDC	Optional performance profile to set for SDCs. `High` or `Default` (same as leaving this field empty).	
SDC Name	Name for the SDC.	

a. This field is required, unless you specify separate data and management IP addresses.
b. To enable volume creation, you must add (at least) one device to (at least) 3 SDSs, where each SDS is in a separate fault unit, and each device has a minimum of 100 GB free storage capacity. You can do that via the CSV, or at a later stage.

Adding SDS devices that were used in a previous ScaleIO system

Devices that were used in previous ScaleIO systems may not have been removed properly. When trying to add these device, the deployment will fail. This is most likely to occur in the following cases:

- When installing ScaleIO on SDS servers whose devices were previously used as SDS devices.
- When extending a system, with an SDS whose devices were previously used as SDS devices.

- When extending a system, by adding an additional SDS (multiple SDS) to a node that already has an SDS installed on it, regardless of the devices that are being used.

  To prevent this issue, perform the following, before running the deployment:

Procedure

1. In the `gatewayUser.properties` file, set the value of `add.sds.with.force.on.next.run = true`.

   For more information, see "Configuring the Installation Manager".

2. Restart the scaleio-gateway service.

   After finishing and marking the deployment complete, the flag will revert to `false`, so you will need to set it again, as necessary.

Configuring multiple SDSs on one server

To configure multiple SDSs on one server, follow these guidelines:

- Multiple SDS is supported with the full installation only, not with the installation wizard.
- Multiple SDS is supported on Linux only.
- Use a separate line for each SDS. This will result in multiple lines for the same server.
- On any server, you cannot mix one SDS and multiple SDSs.
- In the **Is SDS** column, enter **`Yes-`*`X`*** on each line, where *X* is the number of the SDS. For example: `Yes-1, Yes-2`

  Note

  On any SDS IP address, you should enter either one line with the Yes value, or a minimum of two lines with Yes-X options. You cannot mix the Yes value (which means single SDS) with any Yes-X value (which means multiple SDS) on the same IP address.

- When assigning multiple SDSs to a server, the SDS numbers must be consecutive, beginning with 1. For example, to assign three SDSs to one server:
  - Valid—Yes-1, Yes-2, and Yes-3
  - Invalid—Yes-1, Yes-2, and Yes-4
- When extending a node on which there are multiple SDS, you must first set add.sds.with.force.on.next.run = true, as described in "Adding SDS devices that were used in a previous ScaleIO system".

In the following example, the first four lines (under the field names) represent one server with four SDSs. The last line is a different server, with only one SDS.

D	E	F
Is SDS	SDS Name	SDS Device List
Yes-1	sds_76-1	/dev/sdb,/dev/sdc
Yes-2	sds_76-2	/dev/sdd,/dev/sde
Yes-3	sds_76-3	/dev/sdf,/dev/sdg
Yes-4	sds_76-4	/dev/sdh,/dev/sdi
Yes	sds_244	/dev/sdb,/dev/sdc,/de

Installing with the Installation Manager

This section describes how to install and configure ScaleIO components using the Installation Manager.

You need to do the following steps:

- "Log in to the IM server"
- "Upload installation packages"
- "Upload CSV file"
- "Initiate the installation"
- "Complete the installation"

Log in to the IM server

Procedure

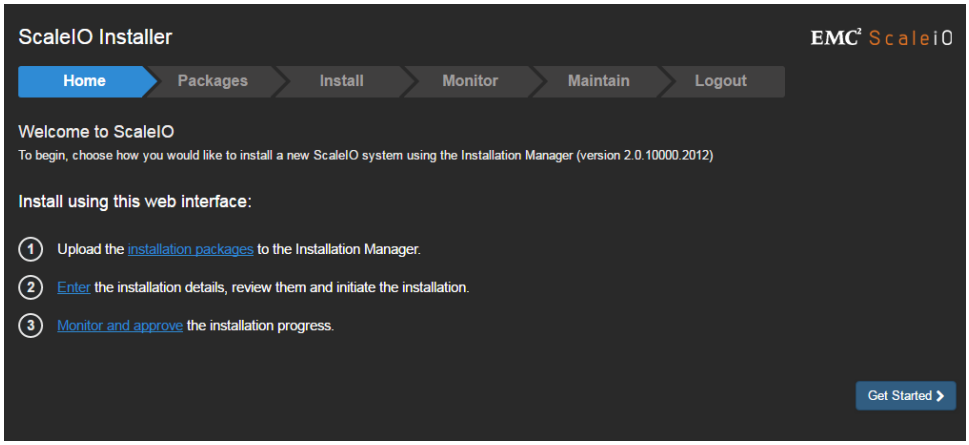
1. Log in to: https://*<IM_Server_IP>*

   where *<IM_Server_IP>* is the IP address of the server where you installed the IM package.

2. Accept the certificate warning; alternatively, install your own certificate for the Tomcat server.

3. Enter the default user name, `admin`, and the password defined when the IM was prepared, then click **Login**.

   The **Welcome** screen appears:

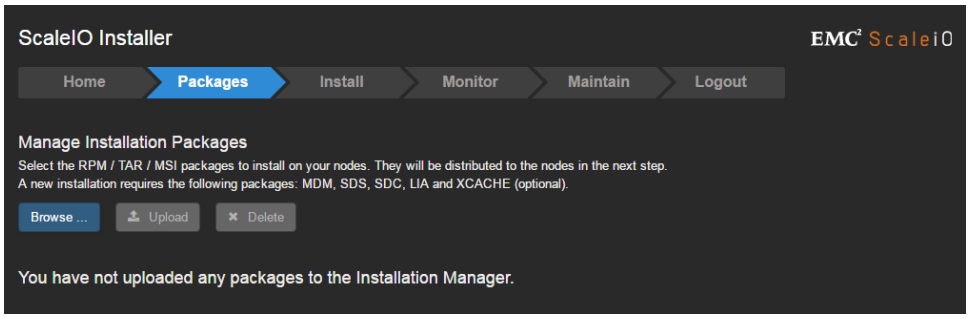


Upload installation packages

Procedure

1. Click **Packages**.

   You may need to re-authenticate with your login credentials. The **Manage Installation Packages** window appears:

   Note

   To use this Gateway to install packages for Windows and Linux, you can upload all the files at once.

2. Browse to where the ScaleIO packages are located for the OS of the servers you are deploying.
3. Select all the relevant files for the operating systems you want to deploy on, then click **Open**.

   File formats can be RPM, TAR, or MSI. Minimally, you must select these packages:

   a. MDM

   b. SDS (To enable multiple SDSs on one server, select all the SDS packages.)

   c. SDC

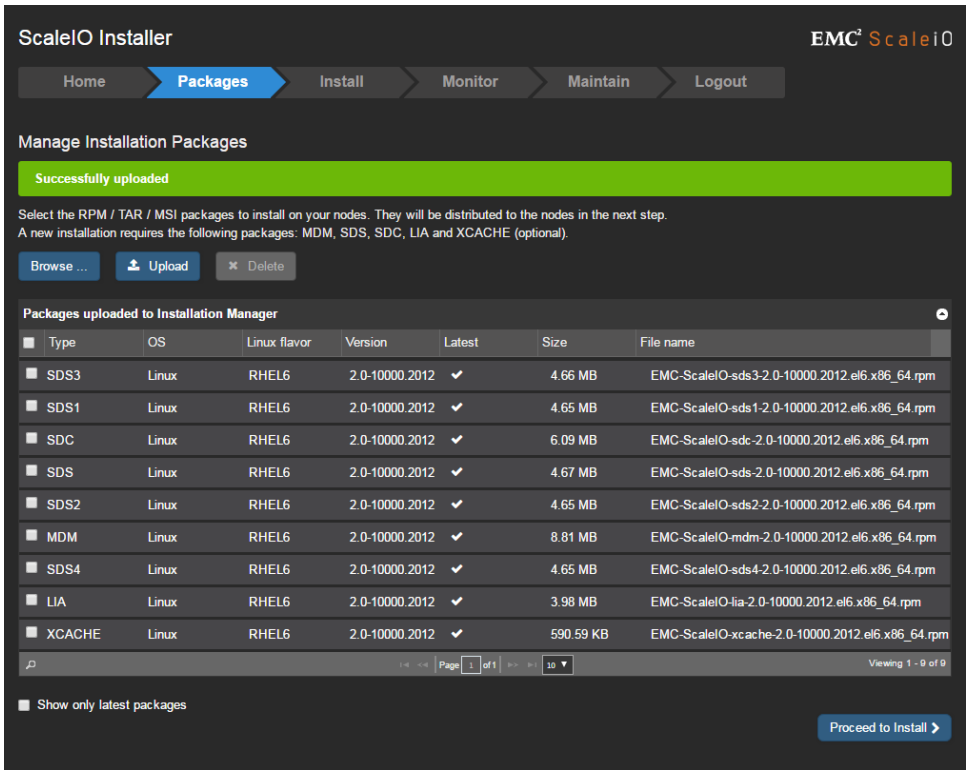
   d. XCache (RFcache)

   e. LIA

   The **Browse** button changes its appearance to reflect the selected files.

4. Click **Upload**.

   The uploaded installation packages (RHEL, in this example) appear in the file table:

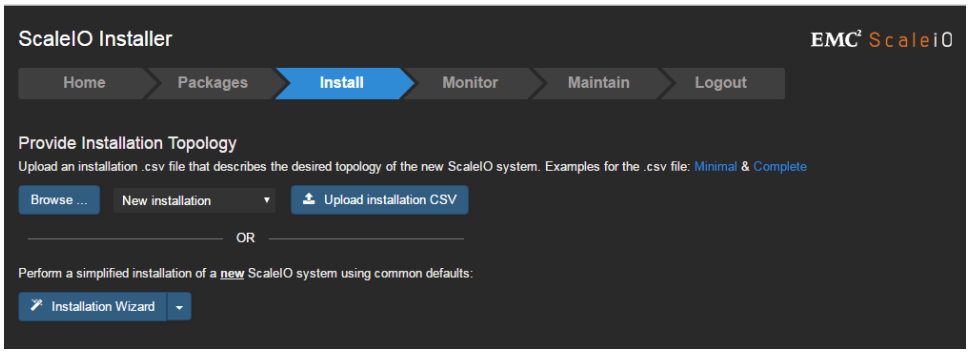


5. Click **Proceed to Install** to proceed to the next step.

Note

To install the replication splitter for RecoverPoint, be sure to upload the splitter package. Replication support is version-specific; see the ESSM for full details.

The **Provide Installation Topology** screen appears:



Upload CSV file

If you have not created the CSV file yet, you can download a template by clicking **Minimal** or **Complete** on this screen.

Procedure

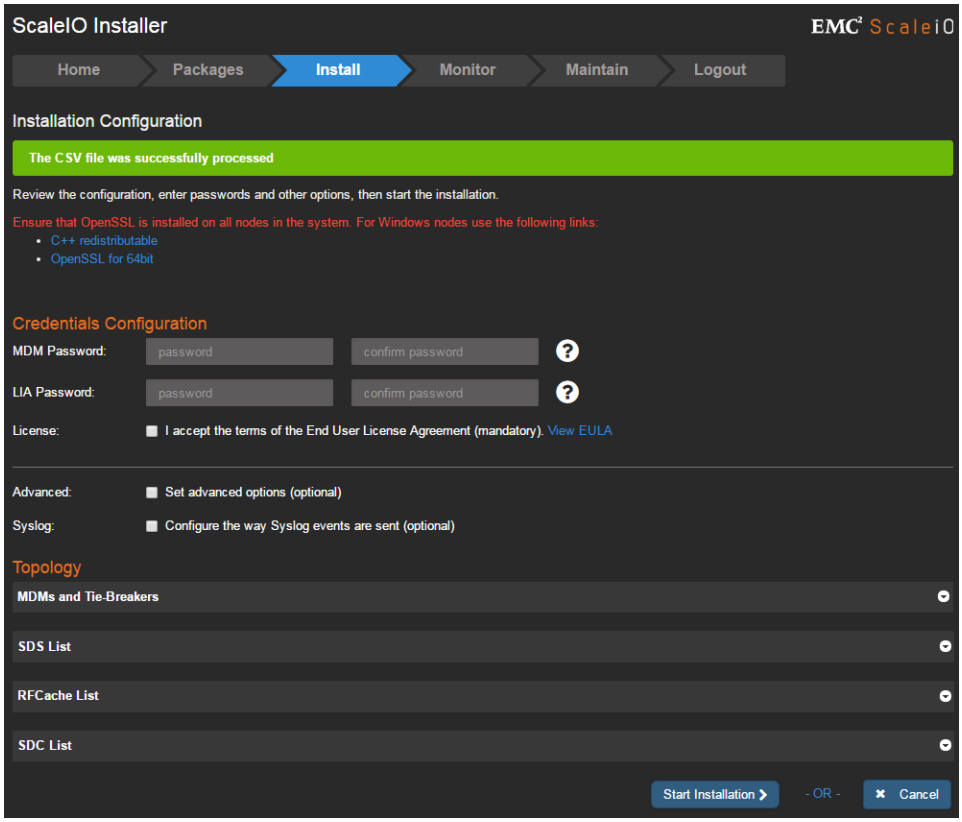
1. Click **Browse**, browse to where the installation CSV file is located, select the file, and click **Open**.
2. For a new installation, select **New installation**.

   To update an existing installation, click the down-arrow and select **Add to existing sys**.

3. Click **Upload Installation CSV**.

   After successfully uploading the file, the **Installation Configuration** screen appears:

Figure 11 IM Installation Configuration screen



   You can expand or contract the lists in the **Topology** section.

Initiate the full installation

Procedure

1. Enter a new MDM password.

   The MDM password is a password that is configured for the MDM during the installation. It must meet the following criteria:

   - Between 6 and 31 ASCII-printable characters
   - Includes at least 3 of the following groups: [a-z], [A-Z], [0-9], special chars (!@#$ ...)
   - Contains no white spaces

   Note

   When using the Installation Manager to extend (rather than to install) a system, enter the MDM credentials that you entered during the initial installation.

2. Enter the LIA password.

The LIA password is a new password that will be used to authenticate communication between the IM and the LIA. It must meet the same criteria as the MDM password, as listed in the previous step.

Note

When using the Installation Manager to extend (rather than to install) a system, enter the LIA credentials that you entered during the initial installation.

3. Review and accept the end user license agreement.

Note

ScaleIO installation enables unlimited use of the product in non-production environments. To obtain a license for production use, and to receive technical support, open a service ticket with EMC Support at https://support.emc.com.

For complete information on licensing, see the *EMC ScaleIO User Guide.*

4. To use any of the following advanced options, select **Set advanced options**:

   You can use the Installation Manager to configure Syslog event reporting. You can also configure these features after installation, via the CLI.

Option type	Option	Description
Skip options	Skip upload	Don't upload packages. You can use this if the packages were already uploaded.
	Skip installation	Don't install packages. You can use this if the packages were already installed.
	Skip configuration	Don't configure ScaleIO components. You can use this if you only want to upload and install packages.
	Skip Linux devices validation	Don't validate Linux device names. For more information, see "Installing without validating Linux devices".
Configuration options	Enable zero-padding	Enable zero-padding of new Storage Pools. Zero padding is required for using the background scanner in data comparison mode and for use with RecoverPoint splitter replication. **Note** Replication support is version-specific; see the ESSM for full details.

Option type	Option	Description
	Enable alert service	Enable the alert service required for SNMP and ESRS reporting. This also creates and configures the lockbox. If you select this option, the **Traps Receiver IP/ Hostname** field is displayed. Enter up to two (comma-separated) IP addresses or hostnames of the SNMP trap receiver servers.
Security options	Disable secure communication with MDM	Disable the need for secure communication mode between management clients and the MDM. **Note** Disabling secure communication has security implications, described in "Using SCLI in non-secure mode".
	Disable secure communication with LIA	Disable the need for secure communication mode between management clients and the LIA. **Note** Disabling secure communication has security implications, described in "Using SCLI in non-secure mode".
	Disable authentication in internal components	Disable the need for authentication between SDSs and MDMs.
	Use trusted IP addresses	Limit the ScaleIO Gateways that can communicate with this LIA. If you select this option, enter the IP addresses to allow, including the IP address of this Gateway. If the Gateway uses multiple IP address (for example, one for management and another for data), enter all the addresses.

5. To configure Syslog reporting, select **Configure MDM to send messages to syslog server**, and enter the following parameters:

   a. **Syslog Server**

      The host name or IP address of the syslog server to where the messages are to be sent. Enter up to two servers, comma-separated.

   b. **Port**

      The port of the syslog server (default 1468)

   c. **Syslog Facility**

      The facility level (default: Local0)

6. Review the displayed information, then click **Start Installation**.

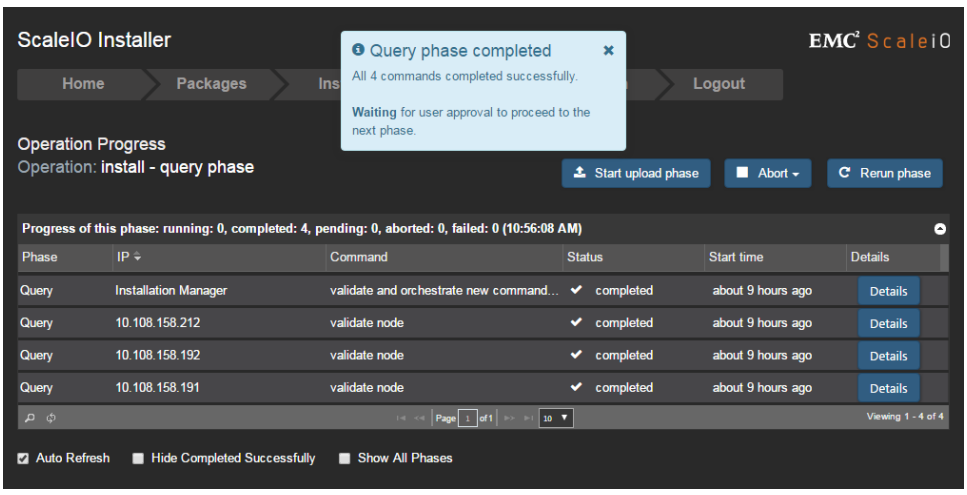
   Note

   If you are installing RFcache (the `xcache` package) on a Windows server, a server restart is necessary. To continue with the installation, click **Allow restart**.

7. Click the **Monitor** tab.

Complete the installation

The installation performs the following phases: query, upload, install, and configure. The monitor **Install - query** screen appears:



Note

When adding nodes to an existing system, the query phase does expect to find currently-installed nodes.

The IM validates that there are no previously installed ScaleIO components on any of the requested installation nodes.

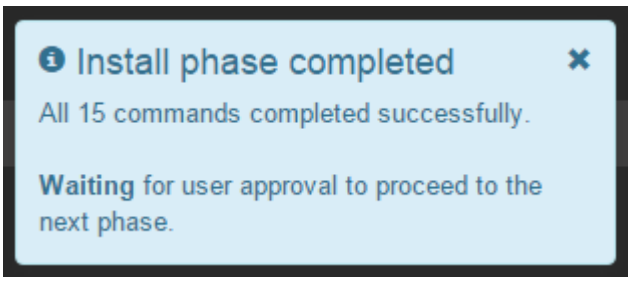
Note

When using the IM to install the replication splitter for RecoverPoint, this step is not performed on the SDC nodes. Replication support is version-specific; see the ESSM for full details.

You can change the display by selecting the following:

- Auto Refresh (default)
- Hide Completed Successfully
- Show All Phases

At the end of each phase, a status message is displayed, and you need to approve moving to the next phase.



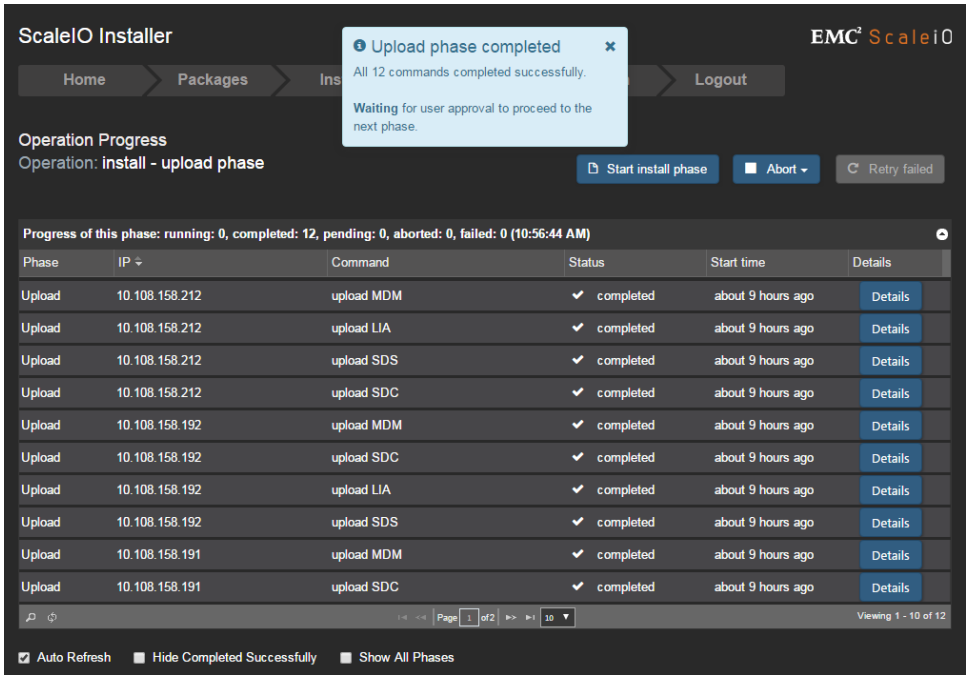
If an error message is displayed during the process, you can retry the failed operations (with the **Retry failed** button) continue to the next phase, or abort the installation.

Note

If you proceed to the next stage, you will not be able to use the IM to retry failed tasks.

Procedure

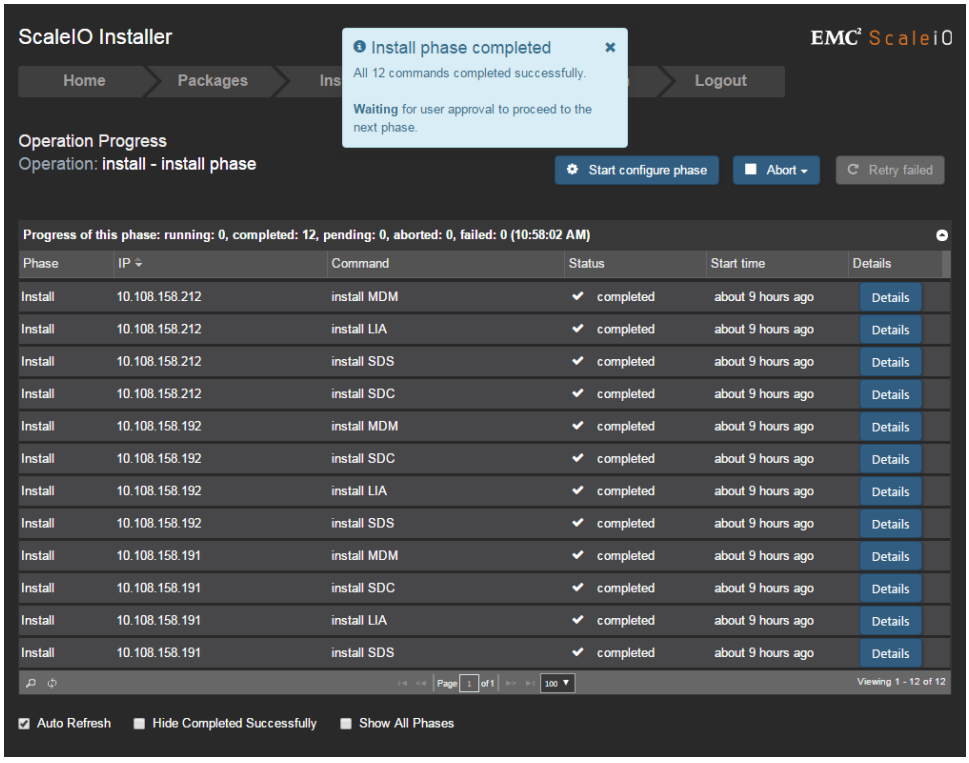
1. When this phase is completed, click **Start upload phase**.
   The **Install - upload** screen appears:



This screen displays which ScaleIO packages are being uploaded to each server, and the status of each command.

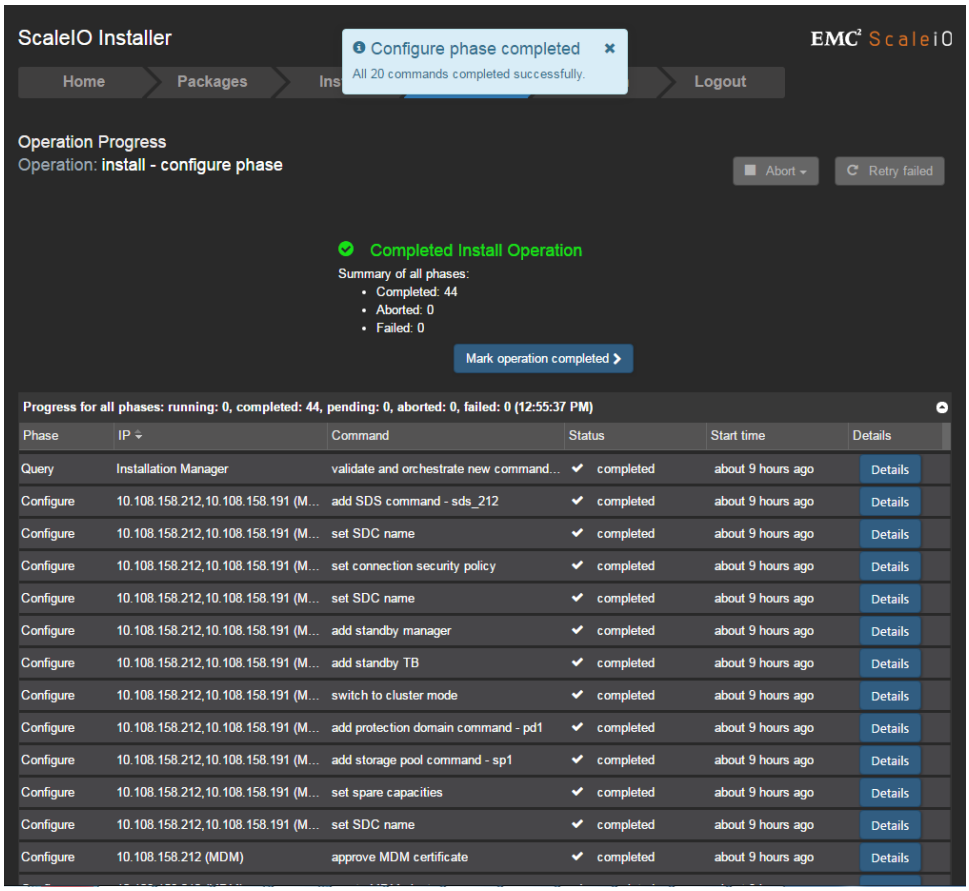
2. Click **Start install phase** to proceed to the next stage.

   The **Install - install** screen appears:

3. 

   When all Install command rows are in Completed status, click **Start configure phase**.

   The **Install - configure** screen appears:



Note

Marking the operation completed signals to the IM that it can now be used for other installations, upgrades, and so on.

4. When all processes are finished, click **Mark operation completed**.

   The ScaleIO system installation is complete!
   A post-operation notice is displayed. The steps in this notice are described in the post-deployment checklist in the *EMC ScaleIO Deployment Guide*, as well as in the *EMC ScaleIO User Guide*.

5. Install the GUI, as described in Installing the ScaleIO GUI on page 100.

   To log in and use the GUI, see the *EMC ScaleIO User Guide*.

After you finish

It is highly recommended to run the ScaleIO system analysis to analyze the ScaleIO system immediately after deployment, before provisioning volumes, and before using the system in production. For more information, see System analysis overview on page 212.

Best practice suggestion

For optimal performance in environments with more than 60,000 IOPS, see the *Fine-Tuning ScaleIO Performance Technical Notes* .

ScaleIO support on Linux servers running Veritas Cluster Server (VCS) software

When a machine running VCS software on top of ScaleIO volumes is rebooted, the VCS virtual devices may enter failed state. This is due to a change introduced in ScaleIO v2.0, that can cause the SDC driver to temporarily hang, and the ScaleIO volumes to appear only after a 45-second delay during machine boot. When the VCS devices enter the failed state, manual intervention is required to refresh them, using VCS command line utilities.
To bypass this issue, perform the following steps on every server running the SDC, after installing the SDC:

Procedure

1. Open `/etc/init.d/scini` for editing.
2. In the line that calls drv_cfg --boot_wait, add this text: `--seconds 2`

   The corrected line should read:

   ```
   /bin/emc/scaleio/drv_cfg --boot_wait --seconds 2
   ```

3. Save the edited file.

Results

This change shortens the SDC driver's wait time on the MDM during machine reboot and (in most of the cases) will avoid the hang. In the very rare case where the SDC didn't receive its first MDM update in 2 seconds, the ScaleIO volumes will appear only after 45 seconds, causing the VCS devices to fail as described above. If this happens, use VCS command line utilities to rediscover the SDC volumes.

Where to go from here

Now that your ScaleIO system is up and running, the next task is to create and map volumes. The SDCs expose these volumes as local storage devices to the applications servers. For more information, see "SAN virtualization layer".

You can create and map volumes using any of the ScaleIO CLI management utilities, as described in the *EMC ScaleIO User Guide*.

Installing with the Installation Manager wizard

This section describes how to use the installation wizard, the quickest way to get a ScaleIO system up and running.

Note

ScaleIO installation enables unlimited use of the product, in non-production environments. To obtain a license for production use, and to receive technical support, open a service ticket with EMC Support at https://support.emc.com.

For complete information on licensing, see the *EMC ScaleIO User Guide.*

The wizard is most suitable for environments when all servers are converged and in one Protection Domain.

Note

To use the wizard, you must prepare the Installation Manager, as described in "Preparing the Installation Manager and the Gateway".

The wizard prepares all ScaleIO management and data communication on a single network. The IM must have access to that network to succeed. You can add additional data network IP addresses for SDS-SDS and SDS-SDC communication via the CLI or the GUI.

The wizard guides you through the following tasks:

- "Task 1: Setting up the installation"
- "Task 2: Running the installation"

Setting up the installation

This section describes how to set up the wizard installation:

Procedure

1. Log in to the Installation Manager (IM) server:
   a. Point your browser to this URL: https://*<IM_Server_IP>*

      where *<IM_Server_IP>* is the IP address of the server where you installed the IM package

   b. Accept the certificate warning; alternatively, install your own certificate for the Tomcat server.

   c. Enter the default user name, `admin`, and the password defined when the IM was prepared, then click **Login**.
2. Upload installation packages:
   a. From the **Welcome** screen, click **Packages**.

      You may need to re-authenticate.

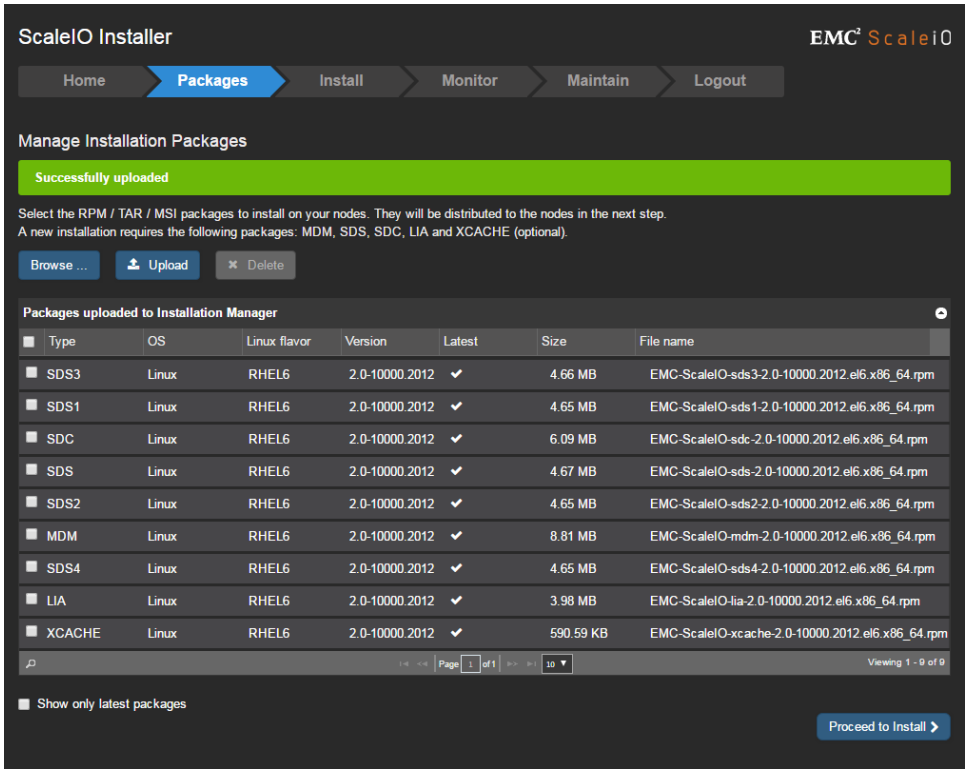
   b. From the **Manage Installation Packages** screen, browse to the extracted ScaleIO component packages, select the files, and click **Open**.

      File formats can be RPM, TAR, or MSI.
      Minimally, you must select these packages:

      - MDM
      - SDS (The SDS-X packages are not needed.)
      - SDC
      - xcache
      - LIA

   c. Click **Upload**.

The uploaded installation packages appear in the file table.



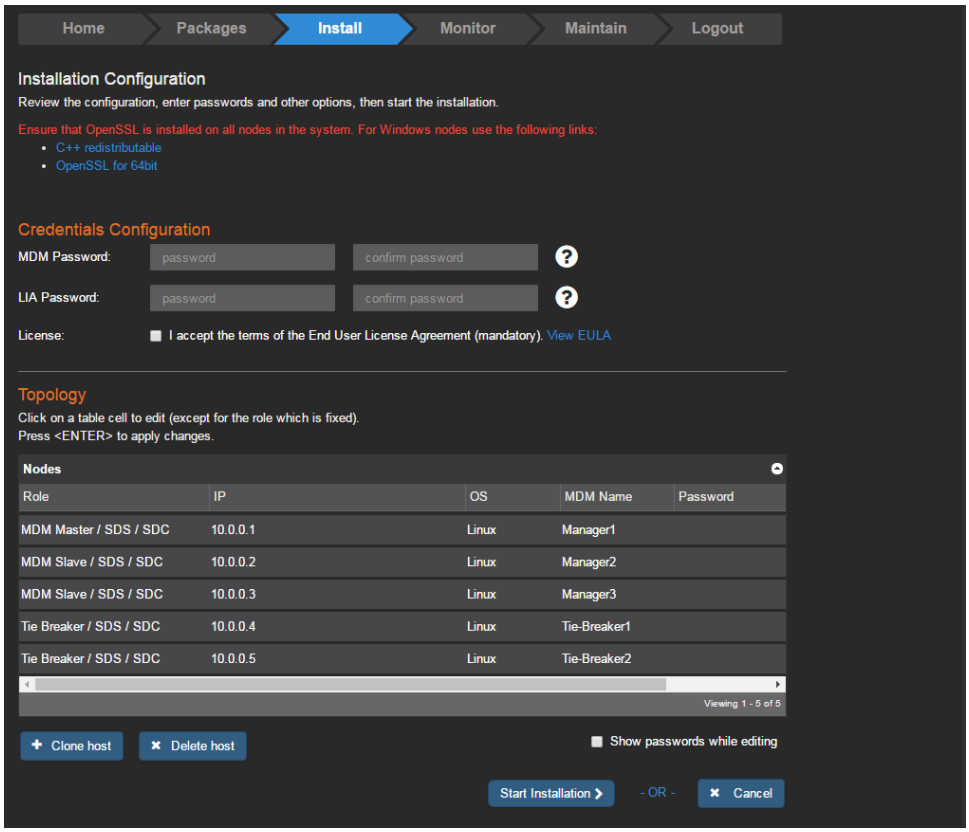
d. Click **Proceed to Install**.

3. In the **Provide Installation Topology** screen, click **Installation wizard**.
4. Select the format of the MDM cluster:
   - 3-node cluster (default)

     Configures a Master MDM, a Slave MDM, and a Tie Breaker MDM.
   - 5-node cluster

     Configures a Master MDM, two Slave MDMs, and two Tie Breaker MDMs.

   The **Installation Configuration** window appears:



Initiate the wizard installation

Procedure

1. Enter a new MDM Password and LIA Password.

   The MDM and LIA passwords must meet the following criteria:

   - Between 6 and 31 ASCII-printable characters
   - Include at least 3 of the following groups: [a-z], [A-Z], [0-9], special chars (! @#$ ...)
   - No white spaces

   The **Topology** section is populated with nodes, according to the MDM cluster chosen in the previous step. Each node is also defined as an SDS and an SDC.

2. Review and accept the end user license agreement.

   NOTICE

   ScaleIO installation enables unlimited use of the product in non-production environments. To obtain a license for production use, and to receive technical support, open a service ticket with EMC Support at https://support.emc.com.

   For complete information on licensing, see the *EMC ScaleIO User Guide.*

3. In the **Topology** section, enter server information:

Note

To see the passwords, press **Enter**, then select **Show passwords while editing**.

a. For each node, change the IP address, select the host operating system, and enter the host password (for Linux, the password of the root user; for Windows, the password of the administrator user).

b. To add more hosts, click **Clone host**.

The assigned IP is derived from the IP of the previous host. Each cloned host is enabled to act as an SDS and an SDC.

4. Click **Start Installation**.

Note

A post-installation notice appears, that outlines the steps that are required after the wizard completes. The content of this notice will be repeated at the end of the process, or you can view it by clicking **Post installation instructions** from the **Monitor** screen.

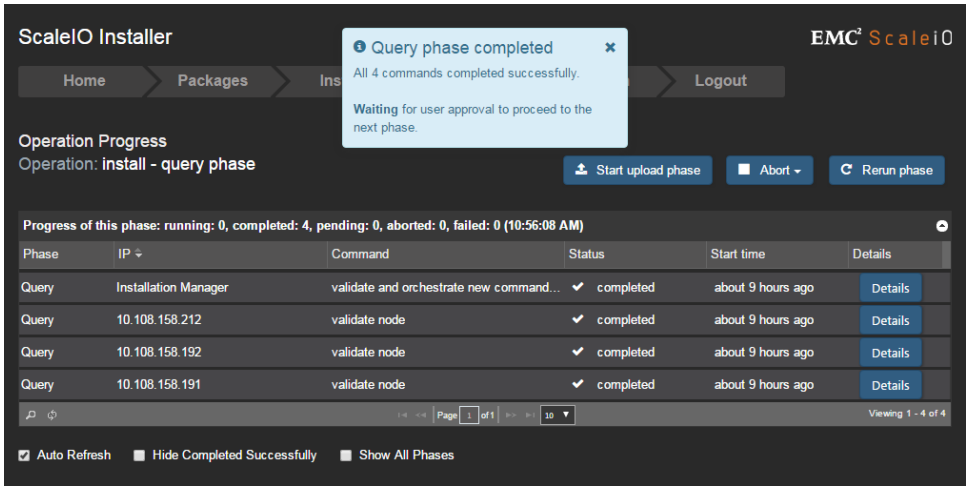
Running the installation

The installation performs the following phases: query, upload, install, and configure. At the end of each phase, you need to approve moving to the next phase.

Procedure

1. Click the **Monitor** tab.

The **Install - query** screen appears.

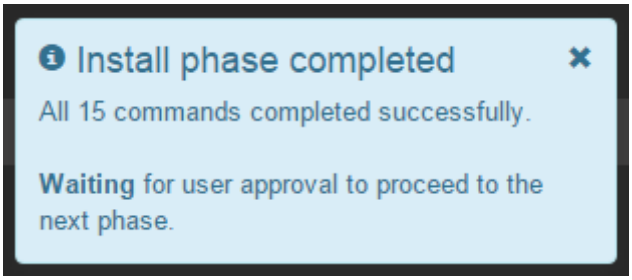


2. When this phase completes successfully, click **Start upload phase**.

Note

You can proceed even if some of the tasks fail, then return to complete the tasks later.

The**Install - upload** screen shows which ScaleIO packages are being uploaded to each server. When all commands in a phase are complete, a message, similar to the following, appears:



The installation process waits for your approval before moving on to the next stage.

If an error message is displayed during the process, you can click the **Retry failed** button to retry the failed operations. Or you can continue to the next phase, or abort the installation.

Note

If you proceed to the next stage, you will not be able to use the IM to retry failed tasks.

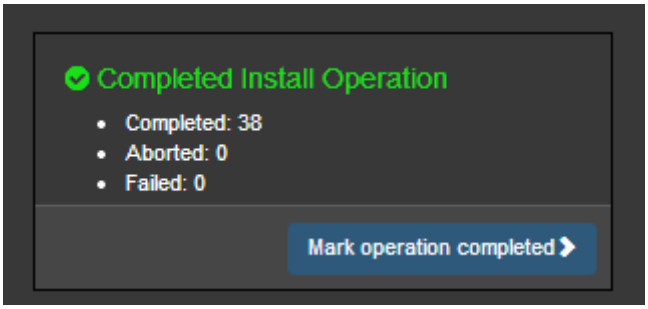
---

3. Click **Start install phase** to proceed to the next stage.

   The **Install - install** screen shows installation progress.

4. When all Install command rows are in Completed status, click **Start configure phase**.

   The **Install - configure** screen displays configuration progress.
   When all processes are completed, the following message appears:



5. Click **Mark operation completed**.

   The ScaleIO system installation is complete! The wizard installation creates one Protection Domain and one Storage Pool, both named "default". These are used in the following section.
   The post-installation notes appear, directing you to the steps necessary to start using your storage. These steps are described in the next section.It is highly recommended to run the ScaleIO system analysis to analyze the ScaleIO system immediately after deployment, before provisioning volumes, and before using the system in production. For more information, see System analysis overview on page 212.

After you finish

If you are using Linux servers that are running Veritas Cluster Server (VCS) software, see ScaleIO support on Linux servers running Veritas Cluster Server (VCS) software on page 83 for post-installation instructions.

Enabling your storage

After the ScaleIO system is installed, follow these steps to enable using the storage. You must issue these commands from the Master MDM, either directly, or via SSH/RDP.

Note

You can also use the GUI to add SDS devices.

Procedure

1. Add SDS devices, according to the following guidelines:
   - You must add at least one device per SDS, to at least 3 SDSs, with a minimum of 100 GB free storage capacity per device.
   - Balance the total device capacity among all SDSs.

   a. Log in, by typing the following command:

   ```
   scli --login --username
   <mdm_username> --password <mdm_password>
   ```

   On Windows-based MDMs, preface commands with `cli`, instead of `scli`.

   Note

   If bash completion isn't enabled, run `./etc/bash_completion/scli`.

   b. Add devices, by typing the following command:

   ```
   scli --add_sds_device --sds_ip <IP> --storage_pool_name
   default --device_path <device_path>
   ```

   Linux example:

   ```
   scli --add_sds_device --sds_ip 192.168.212.10 --
   storage_pool_name default --device_path /dev/sdX
   ```

   Note

   Mapped volumes appear to the SDC as `/dev/sciniX`, where X is a letter, starting from "a". For more information, see "Mounting ScaleIO".

Windows example:

```
cli --add_sds_device --sds_ip 192.168.212.10 --
storage_pool_name default --device_path d
```

Note

For other options regarding adding Windows devices, see "Adding devices to SDS nodes on Windows servers".

2. Add and map a volume:

   a. Add a volume, by typing the following command:

```
scli --add_volume --storage_pool_name default --size_gb
<SIZE> --volume_name <NAME> --protection_domain <NAME>
```

   Example:

```
scli --add_volume --storage_pool_name default --size_gb 16
--volume_name vol01 --protection_domain pd01
```

   b. Map a volume to an SDC, by typing the following command:

```
scli --map_volume_to_sdc --volume_name <NAME> --sdc_ip <IP>
```

   Example:

```
scli --map_volume_to_sdc --volume_name vol01 --sdc_ip
192.168.212.19
```

   Your storage is ready to use.

   To view the installed nodes and storage, use the GUI or the CLI `--query_all` command .

Deploying on Ubuntu, Oracle Linux, or CoreOS servers

This section describes steps that must be taken when deploying SDC or RFcache on Ubuntu, Oracle Linux (OL), or CoreOS servers. You must ensure that the ScaleIO SDC and RFcache (XC) drivers are compatible with the Ubuntu/OL/CoreOS kernel.

Note

All references to Ubuntu in this section apply equally to supported versions of OL and CoreOS.

To ensure that a client node always has an SDC/XC kernel driver that matches the client kernel, the SDC/XC startup script uses `driver_sync.sh`, a synchronization utility that fetches SDC/XC kernel drivers from a configured repository. This repository can be either EMC's repository or a local repository in your private network.

For Ubuntu nodes with ScaleIO SDC/XC drivers:

- The `driver_sync.sh` script is located, by default, at `/bin/emc/scaleio/`.
- In the repository, a driver package tarball is located at `<base_dir>/Ubuntu/<scaleio_version>/<kernel_version>/scini.tar`.

For CoreOS nodes with ScaleIO SDC/XC drivers:

- The `driver_sync.sh` script is located, by default, at the location of the script in the OEM folder.
- In the repository, a driver package tarball is located at `<base_dir>/CoreOS/<scaleio_version>/<kernel_version>/scini.tar`.

It may be necessary to manually restart the driver/s and services on SDCs.

In a normal work flow, an sdc/xcache driver will be loaded on boot. If there is a mismatch between the version of the running kernel and the kernel against which the driver was compiled, the `driver_sync.sh` script will be run automatically to fetch the appropriate driver. If the script fails for any reason (incorrect configuration etc.) after fixing the issue with the `driver_sync.sh` script, the driver services must be restarted manually, as described below. Root permission is required to run the commands.

To manually restart driver/s and services, run these commands:

```
service scini restart
```

```
service xcache restart
```

Ensuring the kernel version is correct

Before you begin, perform the following:

Procedure

1. Ensure that your kernel version matches a version in the repository, by performing the following:

   Note

   If all Ubuntu servers have the same kernel, you can run this command on a single server.

   a. Run this command on every Ubuntu server:

   ```
   uname -r
   ```

   Output similar to the following is displayed:

   ```
   3.16.0-62-generic
   ```

   b. Compare the output to the Ubuntu kernel version in the EMC repository:

   ftp://QNzgdxXix:Aw3wFAwAq3@ftp.emc.com

2. To use a mirror repository, you must ensure that the SSH public and SSH private keys are located in all system nodes in the same path.

The private key should have directory-level permission only (chmod 700 <private_key_path>). The public key can have all permissions. This is not necessary when using the EMC repository.

3. The GPG key must be located in all system nodes in the same path. This key is required for using a mirror or EMC repository.

After you finish

The configuration of `driver_sync.sh` is specified in `driver_sync.conf` file, which can be created in the following methods:

- "Creating the configuration file via the ScaleIO Gateway"
- "Creating the configuration file manually"

Creating the configuration file via the ScaleIO Gateway

This section describes how to create the `driver_sync.conf` file during the Gateway deployment process. Before deploying, you set parameters that will be used by the Gateway, during deployment, to create the file for your environment.

To edit the properties, perform the following:

Procedure

1. Use a text editor to open the `gatewayUser.properties` file, located in the following directory on the Gateway server:

   `/opt/emc/scaleio/gateway/webapps/ROOT/WEB-INF/classes`

2. Edit the file by adding the parameters described below.
3. Save and close the file.
4. Import the GPG key by running the following command, on every Ubuntu node that will contain SDC or RFcache:

```
apt-key add <new_public_key>
```

5. Restart the scaleio-gateway service by running the following command:

```
service scaleio-gateway restart
```

After the changes are made, use the Installation Manager to deploy the system, using the installation files for Ubuntu.

Parameters to add

- `sdc.kernel.repo.SCINI_REPO_ADDRESS`

  The value of this variable is encoded as *<protocol>*://[*<address>*]/*<path>*

  Where *<protocol>*:= `ftp` | `sftp` | `file`

  The following example uses the address of the EMC repository:

```
sdc.kernel.repo.SCINI_REPO_ADDRESS=ftp://
QNzgdxXix:Aw3wFAwAq3@ftp.emc.com
```

- `sdc.kernel.repo.SCINI_REPO_USER`

Contains the user name that is used as login name. Needed when the protocol is FTP or SFTP.

Example:

```
sdc.kernel.repo.SCINI_REPO_USER=admin
```

- `sdc.kernel.repo.SCINI_REPO_PASSWORD`

  Represents the password. Needed only for FTP protocol.

  Example:

```
sdc.kernel.repo.SCINI_REPO_PASSWORD=password
```

- `sdc.kernel.repo.SCINI_REPO_USER_KEY=`*<path_to_private_key>*

  Contains a path to the private RSA key used for SSH connections (user identity file). Needed only for SFTP.

  Example:

```
sdc.kernel.repo.SCINI_REPO_USER_KEY=<path_to_private_key>
```

- `sdc.kernel.repo.SCINI_REPO_HOST_KEY=`*<path_to_public_key>*

  Contains a path to the repository server's public host key, used for SSH connections. Needed only for SFTP.

  Example:

```
sdc.kernel.repo.SCINI_REPO_HOST_KEY=<path_to_public_key>
```

- `sdc.kernel.repo.SCINI_REPO_MODULE_SIGCHECK`

  Set to `0` or `1`, to determine whether the fetched kernel modules must be verified.

  Example:

```
sdc.kernel.repo.SCINI_REPO_MODULE_SIGCHECK=1
```

- `sdc.kernel.repo.SCINI_REPO_EMC_KEY`

  Contains a path to the EMC GPG public key file that is needed for signature verification. You can retrieve the GPG key (`RPM-GPG-KEY-ScaleIO_2.0.13000.X.0`) from the folder in the root of the ISO: `ScaleIO_2.0.0.2_GPG-RPM-KEY_Download.zip`

  This is the same file used by LIA to verify RPMs. It needs to be set only if `SCINI_REPO_MODULE_SIGCHECK` was set to `1`.

  Example:

```
sdc.kernel.repo.SCINI_REPO_EMC_KEY=<path_to_key>/RPM-GPG-KEY-
ScaleIO_2.0.5014.0
```

You can now deploy the system as described in "Installing with the full Installation Manager" in the *ScaleIO Deployment Guide*.

During the deployment, a `driver_sync.conf` file is generated for the SDC and for XC in their respective `bin` folders:

- SDC: `/bin/emc/scaleio/scini_sync`
- XC: `/bin/emc/scaleio/xcache_sync`

Creating the configuration file manually

This topic describes how to create the `driver_sync.conf` file manually. Actually, if you already deployed the system with the Gateway, the file already exists in the `bin` folder listed above, with dummy values, and you will need to edit that file, and then copy it to the `bin` folder on every SDC/XC server in the system.

The configuration of the `driver_sync` script is read from `driver_sync.conf` (an example appears below), which contains the following parameters:

Table 10 driver_sync.conf parameters

Parameter	Description
repo_address	Encodes the chosen method and address.
repo_user	Identifies the user name used when logging in to the drivers server. Required for SFTP and FTP methods.
repo_password	The FTP server login password (for user repo_user). Required for FTP method
user_private_rsa_key	Contains the path of the private RSA key file used to SFTP connection. This key file (provided by ScaleIO on installation) enables SFTP connection without a password.
local_dir	The local directory in which the downloaded SDC/XC drivers are stored. For SDC/XC clients, this value is set up during installation and doesn't normally require a change.
repo_public_rsa_key	The path of the public RSA key of the repository server machine (also known as the host public key).
module_sigcheck	Determines whether to verify (1) or not to verify (0) the downloaded kernel drivers.
emc_public_gpg_key	A key that is used to verify signatures of downloaded drivers. Enter the path to the public GPG signing key of ScaleIO (just like the one that comes with ScaleIO LIA).
The following parameters are used when driver_sync.sh is used to synchronize a local repository against a remote one.	
sync_pattern	A regular expression that can be used to fetch several drivers at once. This parameter is not normally needed, and becomes relevant only when driver_sync.sh is used by your own private repository to fetch from a remote (e.g. EMC) repository.
is_local_repo	Set to 0 (default) or omit, for SDC/XC clients. Set to 1 when driver_sync.sh is used to synchronize local repository against a remote repository. Setting to 1 will cause the fetched tar files not to be extracted.

Example:

```
##############################################
#driver_sync Configuration file
#Everything after a '#' until the end of the line is ignored
##############################################
#Repository address, prefixed by protocol
repo_address = sftp://localhost/path/to/repo/dir
#repo_address = ftp://localhost/path/to/repo/dir
#repo_address = file://local/path/to/repo/dir
# Repository user (valid for ftp/sftp protocol)
repo_user = scini
# Repository password (valid for ftp protocol)
repo_password = scini
# Local directory for modules
local_dir = /bin/emc/scaleio/scini_cache/
# User's RSA private key file (sftp protocol)
user_private_rsa_key = /bin/emc/scaleio/scini_key
# Repository host public key (sftp protocol)
repo_public_rsa_key = /bin/emc/scaleio/scini_repo_key.pub
# Should the fetched modules' signatures be checked [0, 1]
module_sigcheck = 1
# EMC public signature key (needed when module_sigcheck is 1)
emc_public_gpg_key = /bin/emc/scaleio/emc_key.pub
# Sync pattern (regular expression) for massive retrieve
sync_pattern = .*
```

Running the ScaleIO Gateway on CoreOS container

Run the ScaleIO Gateway on the docker container

This step is necessary when you do not have another Gateway available to deploy ScaleIO.
When building the Dockerfile, keep the following in mind:

- Set the base image to CentOS 6.x.
- Install Java 1.8.
- Expose the ports necessary for the Gateway (port 80 and port 443).

Procedure

1. Use a Dockerfile to build a container with the following RPM:

   `EMC-ScaleIO-gateway-2.0-13000.X.x86_64.rpm`

2. Run the container.

Creating a mirror repository

This section describes how to set up a local repository, necessary when connecting all SDC/XC nodes to the ScaleIO repository is not possible, perhaps due to security considerations or other limitations.

A common workflow would be to configure your mirror repository to synchronize with the ScaleIO repository, while your SDC/XC nodes will fetch drivers (synchronize) from the mirror repository.

Driver repository directory hierarchy

The directory hierarchy in the repository consists of the following levels:

<Distro>/<ScaleIOVersion>/<Kernel Version>/

where:

- *<Distro>* is either `CoreOS` or `Ubuntu` or `OEL` (for Oracle Linux)
- *< ScaleIOVersion>* is encoded as `x.x.x.x`
- *<Kernel Version>* is the output of `uname -r`.

The files themselves are tar files named `scini.tar` and `xcache.tar`.

For example:

<Repo_base>`/CoreOS/2.0.30.0/4.1.7-coreos/scini.tar` is a tar file of the SDC/XC driver for ScaleIO version 2.0.30.0 for CoreOS with kernel version 4.1.7-coreos.

Synchronizing the repository

In normal workflow, you do not have to (and should not) modify the directory structure manually. Instead, you should use `driver_sync.sh`, which provides a proper `driver_sync.conf`, as described in the previous section. With the configuration file properly written, synchronizing is done by simply calling `driver_sync.sh sync`.

In the `driver_sync.conf` file, you must set the `is_local_repo` parameter to `1`. This will instruct `driver_sync.sh` not to untar the downloaded files.

It is also recommended that you change the `sync_pattern` parameter, a regular expression, to something more specific than `".*"`. Otherwise the sync operation will mirror the entire EMC driver repository every time.

Setting up a new SFTP-based repository

To save time in setting up a new SFTP-based driver repository, use the `driver_repo_wizard.sh` script. This script can be found in the following locations: under the `scini_sync` or `xcache_sync` directory in `/bin/emc/scaleio` (Ubuntu) or `/usr/share/oem/bin/emc/scaleio` (CoreOS).

- Ubuntu or Oracle Linux:
  - `/bin/emc/scaleio/scini_sync`
  - `/bin/emc/scaleio/xcache_sync/`
- CoreOS:
  - `/usr/share/oem/bin/emc/scaleio/scini_sync`
  - `/usr/share/oem/bin/emc/scaleio/xcache_sync`

The script does the following:

1. Creates a new user with a name you provide. The user will be created shell-less.
2. Creates a pair of RSA keys that will be used for passwordless SSH connection. You must provide the created private key to all your client nodes, while the public key will be added to the `authorized_keys` file of the created user.
3. Updates the SSH service configuration file (`/etc/ssh/ssh_config`) for the new user, instructing to `chroot` the user to its home directory, which should also be the repository base directory.

Deploying on OpenStack

This section describes how to deploy ScaleIO in the following OpenStack environments:

- "ScaleIO with Mirantis OpenStack"
- "ScaleIO with Canonical (Ubuntu) OpenStack"
- "ScaleIO with OpenStack Liberty"

ScaleIO with Mirantis OpenStack

EMC has developed open-sourced Fuel Plug-ins to simplify ScaleIO deployment as well as integrate your Mirantis OpenStack deployment with existing ScaleIO storage systems. Searching for ScaleIO on the Mirantis Fuel Plugin site provides links to software and documentation.

You can install ScaleIO manually, as well, or in an automated fashion with other automation frameworks. Visit http://emccode.com to see what open-sourced code exists to assist in automating ScaleIO.

ScaleIO with Canonical (Ubuntu) OpenStack

EMC has developed open-sourced Juju Charms to simplify deployments of ScaleIO as well as integrate your Canonical (Ubuntu) OpenStack deployment with existing ScaleIO storage systems. Searching for ScaleIO on the Canonical Juju Charm site provides links to software and documentation.

You can install ScaleIO manually, as well, or in an automated fashion with other automation frameworks. Visit http://emccode.com to see what open-sourced code exists to assist in automating ScaleIO.

ScaleIO with OpenStack Liberty

The ScaleIO elastic storage solution for OpenStack Liberty does not contain an OpenStack installation package, because the ScaleIO drivers are already included in the Liberty version. The following link supplies all the information required for configuring the Cinder driver:

http://docs.openstack.org/liberty/config-reference/content/ScaleIO-cinder-driver.html

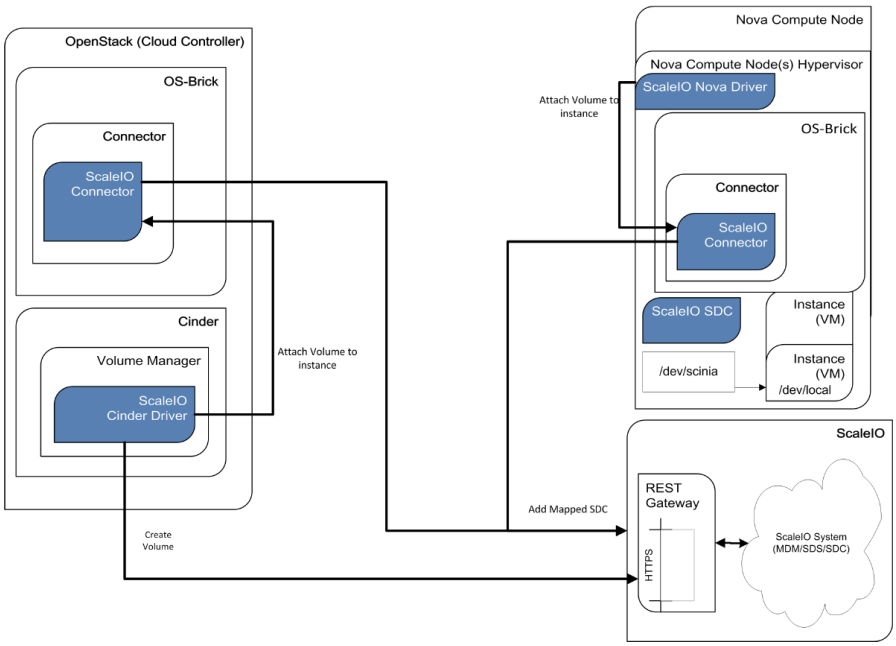
The figure below illustrates the implementation of ScaleIO with OpenStack Liberty.

- Communication between the Cinder driver and the ScaleIO MDM is implemented using the REST API. This happens usually in volume-only related operations such as create\remove volume.
- The OpenStack OS-Brick component discovers volumes that are attached to hosts, and also controls attach\detach volume operations to\from hosts. When attach\detach volume is being called by either Nova or Cinder, it uses the common code in OS-Brick to perform the action with the ScaleIO driver, which calls add \remove mapped SDC commands.

  The OpenStack flow is as follows:

1. Create volume (Cinder operation): Calls the ScaleIO `add_volume` command
2. Attach volume to instance: Initiated by Nova or Cinder:
   a. The Cinder\Nova driver calls OS-Brick to perform the action.
   b. The OS-Brick ScaleIO driver detects the volume and then calls the ScaleIO REST command `addMappedSdc`.
3. After the volume is attached to the compute node, both OpenStack and ScaleIO will recognize it as mapped.

Figure 12 OpenStack Liberty integration



ScaleIO with OpenStack Mitaka

The ScaleIO elastic storage solution for OpenStack Mitaka does not contain an OpenStack installation package, because the ScaleIO drivers are already included in the Mitaka version. The following link supplies all the information required for configuring the Cinder driver:

http://docs.openstack.org/mitaka/config-reference/block-storage/drivers/emc-scaleio-driver.html

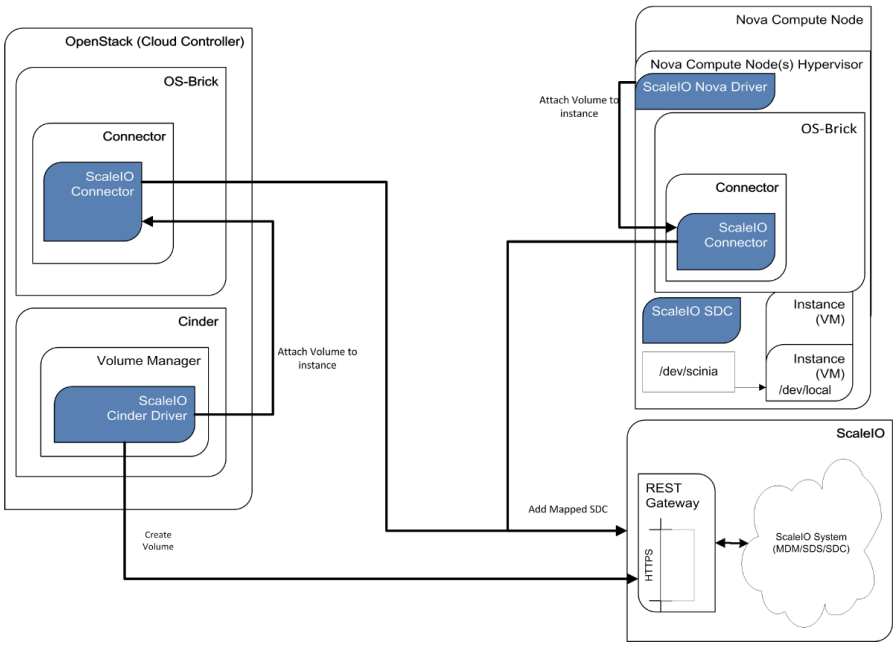
The figure below illustrates the implementation of ScaleIO with OpenStack Mitaka.

- Communication between the Cinder driver and the ScaleIO MDM is implemented using the REST API. This happens usually in volume-only related operations such as create\remove volume.
- The OpenStack OS-Brick component discovers volumes that are attached to hosts, and also controls attach\detach volume operations to\from hosts. When attach\detach volume is being called by either Nova or Cinder, it uses the common code in OS-Brick to perform the action with the ScaleIO driver, which calls add \remove mapped SDC commands.

  The OpenStack flow is as follows:

1. Create volume (Cinder operation): Calls the ScaleIO `add_volume` command
2. Attach volume to instance: Initiated by Nova or Cinder:
   a. The Cinder\Nova driver calls OS-Brick to perform the action.
   b. The OS-Brick ScaleIO driver detects the volume and then calls the ScaleIO REST command `addMappedSdc`.
3. After the volume is attached to the compute node, both OpenStack and ScaleIO will recognize it as mapped.

Figure 13 OpenStack Mitaka integration



Environment and system requirements

The following environment is supported:

- Liberty and later releases, with either all-in-one OpenStack installation, or multi-node (Compute nodes) installation.
- RHEL 6.x or 7.x (or their CentOS equivalents), or Ubuntu 14.04 or 16.04, using KVM as the hypervisor engine.
- The ScaleIO SDC component must be installed and configured on the OpenStack compute node.
- The ScaleIO gateway (which includes the REST gateway) must be installed as part of the ScaleIO deployment. This gateway is typically installed at the beginning of ScaleIO installation. For more information, see the *ScaleIO Deployment Guide*. The ScaleIO gateway must be accessible from the OpenStack cloud controller. Ensure that you know the IP address of the ScaleIO gateway, and its user name and password, because they are required for the Cinder driver's installation and configuration procedures.
- ScaleIO REST Gateway certificate validation—the OpenStack ScaleIO driver communicates with the REST Gateway through https, (over SSL). By default, the driver ignores verification of the ScaleIO REST Gateway's SSL certificate, but it can verify the certificate if the following configuration parameters are defined:
  - `verify_server_certificate`—set to `True`, if the server's certificate must be verified, and to `False` if no verification is required.
  - `server_certificate_path`—If the parameter `verify_server_certificate` is set to `True`, specify the location of the `.pem` file containing the server's certificate.

For instructions for generating a self-signed certificate using the keytool utility, see the *EMC ScaleIO Deployment Guide*.

- The OpenStack driver does not support Storage Pool and domain names that contain special characters, including URL escape characters such as `#`, `/`, and so on. If you would like to create a volume on such a Storage Pool or domain in OpenStack, ensure that you specify the Storage Pool or domain using its ID, and not its name.

Installing the ScaleIO GUI

This topic describes how to install the ScaleIO GUI. You can install the GUI on a Windows or Linux workstation.

Ensure that the workstation satisfies the minimum requirements, as described in System requirements on page 18.

Procedure

1. Install the GUI:

OS	Command
Windows	`EMC-ScaleIO-gui-2.0-13000.X.msi`
RHEL	`rpm -i EMC-ScaleIO-gui-2.0-13000.X.noarch.rpm`
Debian	Run with administrator privileges: `sudo dpkg -i EMC-ScaleIO-gui-2.0-13000.X.deb`

Results

The GUI is installed.

After you finish

To log in and use the GUI, see the *EMC ScaleIO User Guide*.

Configuring Installation Manager properties

Configure the following Installation Manager properties by editing the `gatewayUser.properties` file:

- Enable the Installation Manager (default: true)

  To disable, set `features.enable_IM=false`.

  You can completely disable the use of the IM's default port, 443, by setting both this property and the `features.enable_gateway` property to `false`.

- Enable the reuse of previously used devices, and extend a multi-node SDS.

  To enable, set `add.sds.with.force.on.next.run=true` .

  After finishing and marking the deployment complete, the flag will revert to `false`, so you will need to set it again, as necessary.

- Exclude RecoverPoint RPA nodes from being upgraded with the Installation Manager.

  To exclude RPA nodes while upgrading with the Installation Manager, list the IP addresses to ignore on the `im.ip.ignore.list` line.

  Example:

  `im.ip.ignore.list=10.0.0.1,10.0.0.2,...`

To edit the properties, perform the following:

Procedure

1. Use a text editor to open the `gatewayUser.properties` file, located in the following directory on the Installation Manager/Gateway server:

Gateway installed on	Location of gatewayUser.properties file
Windows	C:\Program Files\EMC\ScaleIO\Gateway\webapps\ROOT\WEB-INF\classes\
Linux	/opt/emc/scaleio/gateway/webapps/ROOT/WEB-INF/classes

2. Edit the file with the desired changes.
3. Save and close the file.
4. Restart the scaleio-gateway service:
   - Windows: Restart the EMC ScaleIO Gateway service
   - Linux: Type the following command:

```
service scaleio-gateway restart
```

Results

Configuration is complete.

CHAPTER 4

Deploying ScaleIO on ESX Servers

This chapter describes how to install, deploy, and perform initial configuration of ScaleIO software components. Topics include:

- Overview of ScaleIO deployment on ESX servers 104
- Before you begin - ESX 104
- Deploying ScaleIO 108

Overview of ScaleIO deployment on ESX servers

This chapter describes an overview of ScaleIO deployment on ESX servers.

ScaleIO provides you with automated tools to install and configure all ScaleIO components on ESX servers from one workstation, in both physical and virtual environments.

Note

ScaleIO installation enables unlimited use of the product, in non-production environments. To obtain a license for production use, and to receive technical support, open a service ticket with EMC Support at https://support.emc.com.

For complete information on licensing, see the *EMC ScaleIO User Guide.*

In VMware, use the ScaleIO vSphere plug-in deployment wizard to install the MDM and SDS components on a dedicated ScaleIO virtual machine (SVM). The SDC is installed directly on the ESX host.

The ScaleIO vSphere VMware deployment wizard enables you to perform all of these activities in a simple, efficient manner, over all the machines in a vCenter. The plug-in can also be used for provisioning ScaleIO nodes.

You can also use the deployment wizard with existing ScaleIO systems, to register them to vSphere, and to add SDS server nodes, or extend the existing system.

Note

It is highly recommended to use the deployment wizard for deployment and provisioning. If the need arises, you can perform these tasks manually, or in a combination with the wizard.

Before you begin - ESX

This section describes deployment prerequisites and other issues to consider before deploying ScaleIO.

Deployment prerequisites

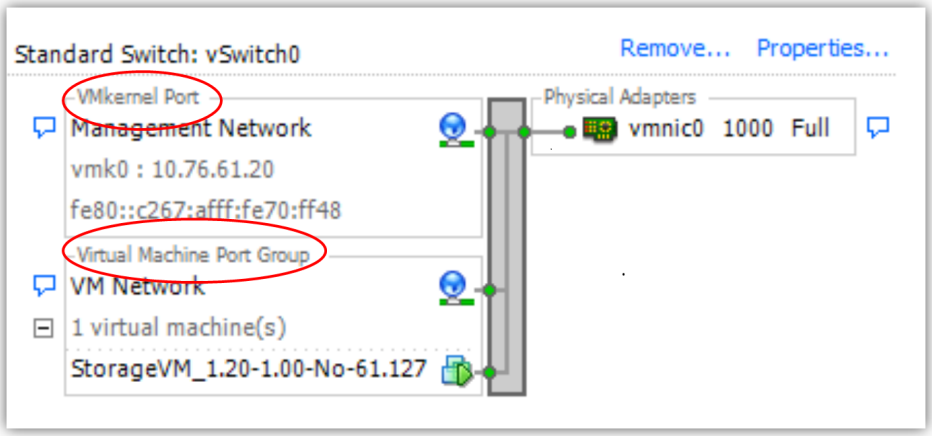
Before you deploy ScaleIO, ensure compliance with the following prerequisites:

- The ESXi hosts must meet the hardware requirements listed in System requirements on page 18.
- For deployment in the ScaleIO Ready Node environment, prepare the ESXi hosts before deployment, as described in ESXi deployment prerequisites on page 190.
- All ESXi hosts selected to have either an MDM, Tie-Breaker, or SDS component installed on them, must have a defined local datastore, with a minimum of 10 GB free space (to be used for the SVM). If the ESX is only being used as an SDC, there is no need for this datastore.
- A minimum of three devices to be added to SDS, that all meet the following prerequisites:
  - A minimum of 100 GB available storage capacity.

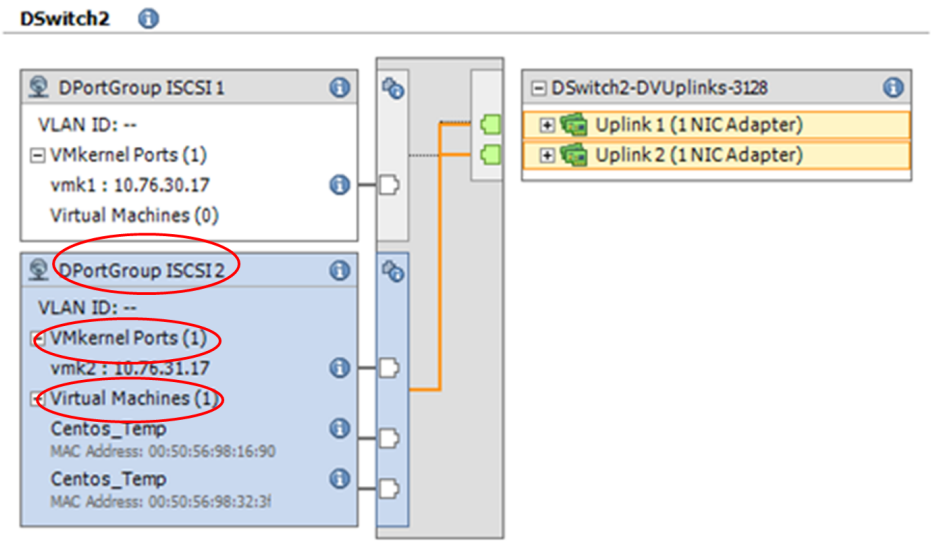
  - The devices must be free of partitions.
  - If a device has a datastore on it, before adding the device, you must either remove the datastore, or use the plug-in **Advanced settings** option to enable VMDK creation.
  - If the device has the ESXi operating system on it, you must use the plug-in **Advanced settings** option to enable VMDK creation.
- The host from which you run the PowerShell (.ps1) script must have the following prerequisites:
  - Runs on Windows, with 64-bit Java installed
  - PowerCLI from VMware (not Windows PowerShell) is installed. PowerCLI version should match the version of the vCenter (for example, both at 6.0).
  - Has incoming and outgoing communication access to the vCenter

    Note

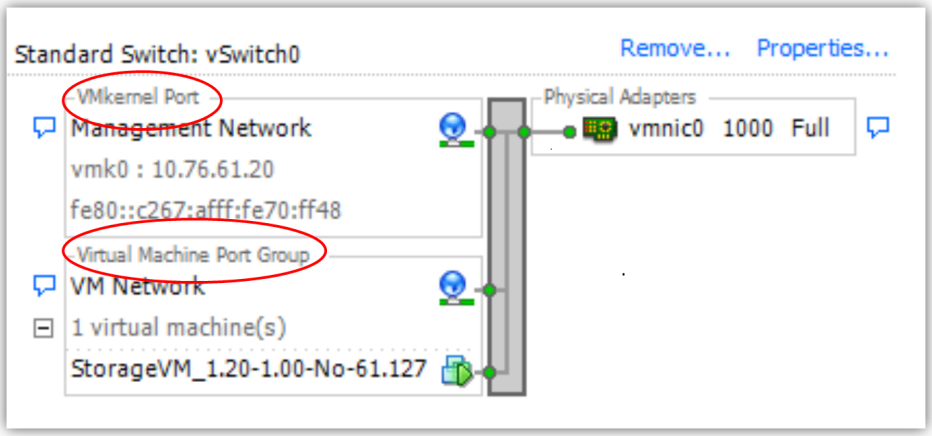
    If your vCenter runs on Windows OS, it is recommended to use it.
- The vSphere web client (Virgo) server must have access to the host on which the PowerShell script will be used.
- The management network on all of the ESXs that are part of the ScaleIO system must have the following items configured:
  - Virtual Machine Port Group. (The name must be the same on all of the ESX hosts.)



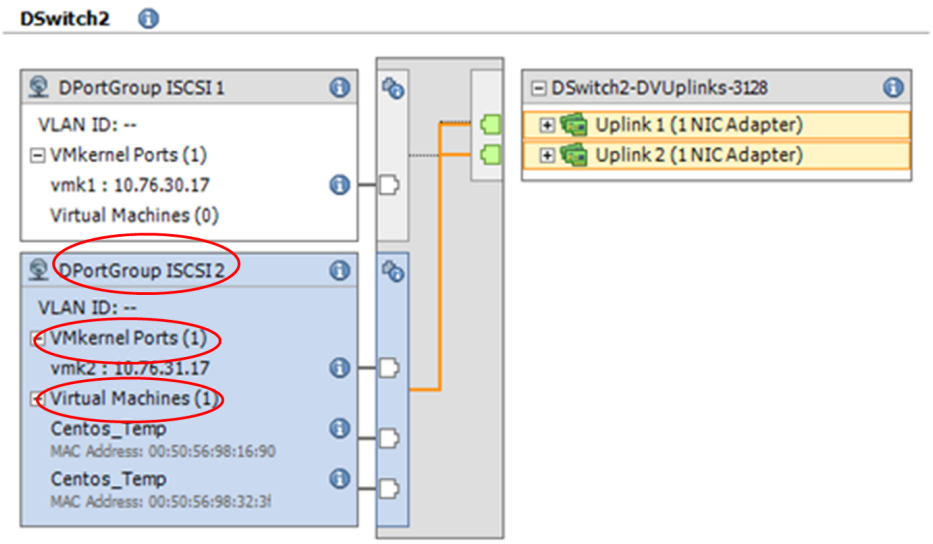
  - When using distributed switches, the vDS must have the following items configured:
    - VMKernel port (necessary only if using a single network)
    - dvPortGroup for virtual machines



- Virtual Machine Port Group. (The name must be the same on all of the ESX hosts.)



- When using distributed switches, the vDS must have the following items configured:
  - VMKernel port (necessary only if using a single network)
  - dvPortGroup for virtual machines



- If using only a single network (management), you must manually configure the following:
  - vSwitch
  - VMKernel Port
  - Virtual Machine Port Group
  - VMKernel Port Binding

Other deployment considerations

During deployment, you may need to address the following items:

- How will the ESX devices be added to the SVM?
  - RDM (highly recommended)

    Using RDM mapping, a device is created on the SVM that points to the physical disk on the ESX.

    You can add RDM devices that are connected via RAID. If a local RDM is not connected via RAID, it may not be supported. To ensure the compatibility of these devices, you can add them as VMDK, or you can select **Enable RDMs on non parallel SCSI controllers**, as described in "Advanced settings options". Enable this option before beginning the deployment.

    Before enabling this feature, contact EMC Support.
  - VMDK

    A new datastore is created, with a VMDK, and the VMDK is added to the SVM. ScaleIO requires thick provisioning, so this process can take a long time.

    In almost all cases, RDM is the preferred method to add physical devices. Use the VMDK method only in the following scenarios:
    - The physical device does not support RDM.
    - The device already has a datastore, and the device isn't being completely used. The excess area that is not already being used will be added as a ScaleIO device.

Note

To use VMDK, select **Enable VMDK creation**, as described in "Advanced settings options".

- Are you deploying a very large system?

  If you are deploying a very large ScaleIO system (several hundred nodes), you can increase the parallelism limit (default: 100), thus speeding up the deployment. This is dependent on the processing power of the vCenter.

  To increase the parallelism limit, use the plug-in **Advanced settings**, as described in "Advanced settings options".
- Do you want to use separate networks for data and management (recommended), and which IP addresses will you use for the SVMs and VMkernels?
- Are there flash devices that will be added to the SDS?
- Do you want to create Fault Sets?

Deploying ScaleIO

This section describes how to deploy ScaleIO in the VMware environment.

To deploy in a 2-layer environment, where only the SDCs are on ESXi servers, see Deploying ScaleIO in a 2-layer environment using the Installation Manager and plug-in on page 188.

Deployment entails the following tasks:

- "Task 1: Registering the ScaleIO plug-in and uploading the OVA template"
- "Task 2: Accessing the plug-in"
- "Task 3: Installing the SDC on ESX hosts"
- "Task 4: Deploying ScaleIO"

Note

After deployment is complete, set all SVMs to start automatically with the system.

Registering the ScaleIO plug-in and uploading the OVA template

This topic describes how to use PowerCLI to register the ScaleIO plug-in on a vCenter and to upload the OVA template.

Note

To use your own web server, see "Manual registration of the ScaleIO plug-in".

Before you begin, ensure that there is communication between the vSphere web client server (usually installed on the vCenter) and the web server storing the plug-in.

Procedure

1. Copy the following files to your designated host (preferably your vCenter), that has PowerCLI installed:

   `ScaleIOVM_2nics_2.0.13000.X.ova`

   `EMC-ScaleIO-vSphere-plugin-installer-2.0-13000.X.zip`

2. Extract the contents of the ZIP file.
3. Using PowerCLI for VMware, set to `Run as administrator`, run the following script:

```
ScaleIOPluginSetup-2.0-13000.X.ps1
```

   a. Enter the vCenter name or IP address, user name, and password.

   Note

   You can also use this script to unregister a plug-in or to create an SVM template.

   b. For `Choose Mode`, choose option `1, Register ScaleIO plug-in`.
   c. Read the upgrade notice, and enter `y` to continue.
   d. For `Select Registration Mode`, choose `Standard` (simplified, using embedded Tomcat).

   This step may take a few minutes.

   Note

   You can use the `Advanced` option to install the plug-in using a ScaleIO Gateway from a previous installation or using your own web service. In either case, you must place this version's `plugin.zip` file (`EMC-ScaleIO-vSphere-web-plugin-2.0-13000.X.zip`) in your resources folder before running the installation. To use a previous ScaleIO Gateway, the resource folder is `ScaleIO Gateway installation folder\webapps\root\resources`.

   e. If necessary, accept the thumbprint.
4. Log in to the vSphere web client.

   If you are already logged in, log out, then log in again.
5. From the vSphere **Home** tab, verify that the ScaleIO icon

is visible in the **Inventories** section.

Note

If the SIO icon is missing, the vSphere web client (Virgo) server failed to download/register the plug-in, due to one of the following reasons:

- A connectivity problem (for example, network/firewall etc.) between the vSphere web client server and the web server storing the plug-in. To resolve the issue, verify that there is communication between the vSphere web client server and the web server that is storing the plug-in.
- URL problem. To resolve the issue, verify that the URL is `https://` and is pointing to the correct web server IP address (i.e. ScaleIO Gateway).

For information on how to use the log to troubleshoot problems that may arise, see "Troubleshooting plug-in registration issues".

6. In the PowerCLI window, press **Enter** to finish the plug-in download and return to the menu.
7. To upload the OVA template, perform the following:
   a. From the PowerCLI script, select 3 to `Create SVM template`.
   b. Enter the parameters described in the following table. (If you are already logged in, some of these parameters may not be necessary.)

Parameter	Description
vcenter	vCenter name or IP address
user name	vCenter user name
password	vCenter password
datacenter	The name of the data center where the datastore that will store the template resides
ova_path	The path of the SVM's OVA
datastores	A list of datastores, up to eight, on which the templates will be created. Press ENTER to stop specifying datastores.

Note

For best results, enter a local (not shared) datastore for each ESX server.

For faster, parallel, deployment in large-scale environments, you can use the OVA to create SVM templates on as many as eight datastores. To do so, enter the datastore names, and when you are done, leave the next line blank. The following example shows how to enter two datastores:

```
datastores[0]: datastore1
datastores[1]: datastore1 (1)
datastores[2]:
```

The upload procedure can take several minutes, during which time a temporary SVM is created, the templates are created, then the temporary SVM is deleted.

When each template is created, a message similar to the following appears:

```
The template EMC ScaleIO SVM Template (v2.0.13000.X) was
successfully created
```

c. When the process is complete, enter **4** to exit the plug-in script.

Accessing the plug-in

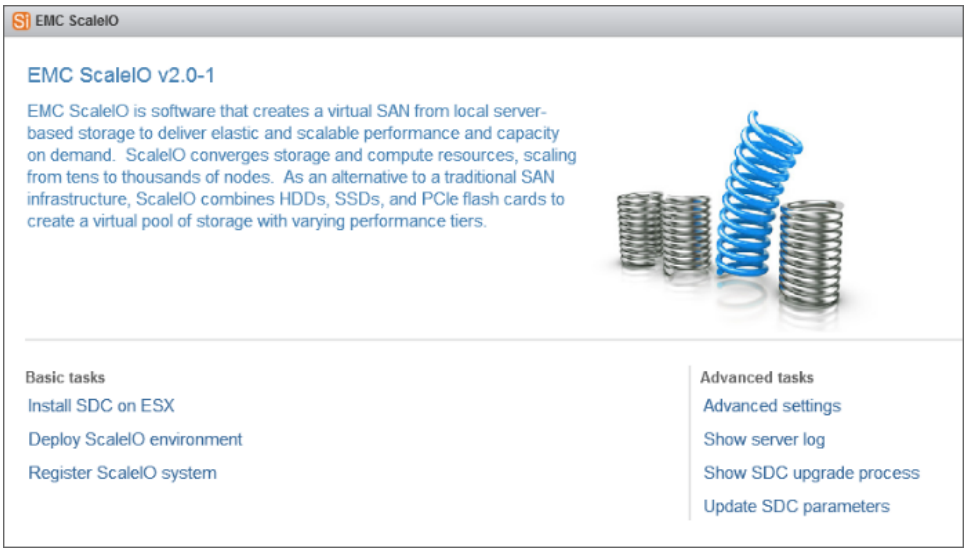
This topic describes how to access the plug-in from the vSphere Web Client home screen.

From the vSphere Web Client home tab, click

.

The **EMC ScaleIO** screen appears:

Figure 14 EMC ScaleIO screen



Installing the SDC on ESX hosts

This topic describes how to install the SDC component on ESX hosts. This is a prerequisite for deploying a new ScaleIO system, and it must be applied to every ESX host that will be in the ScaleIO system, regardless of what role it is playing.

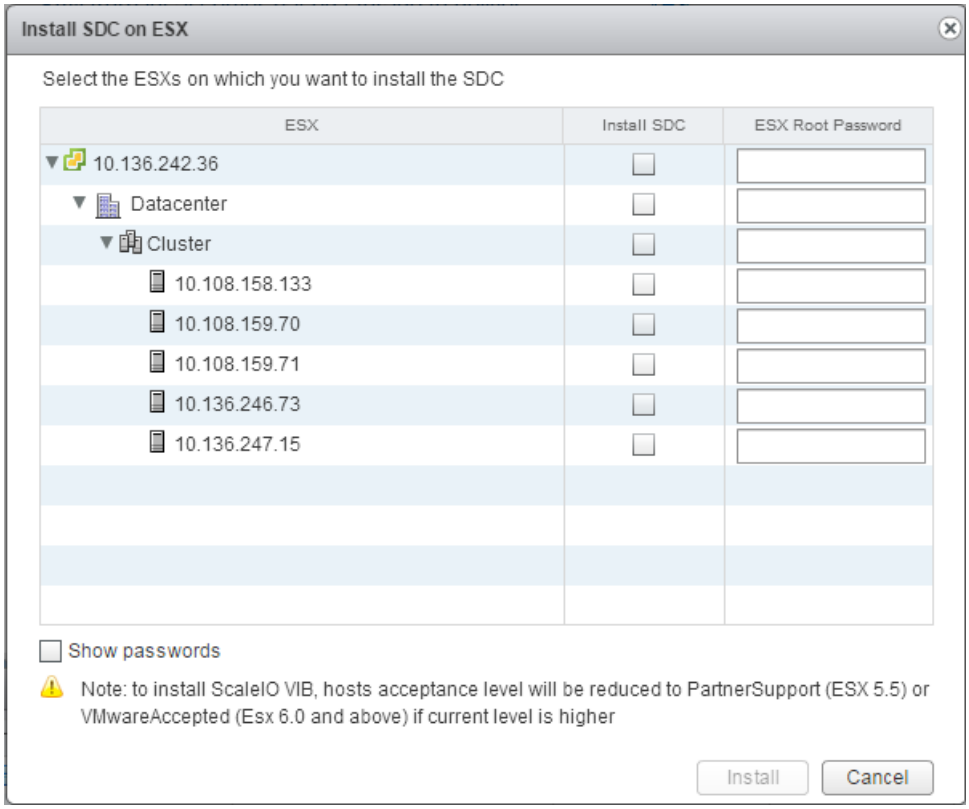
This process installs the SDC on the ESX host, after which you must restart the ESX host.

To install the SDC component, perform the following:

Procedure

1. From the **Basic tasks** section of the **EMC** ScaleIO screen, click **Install SDC on ESX**.

The **Install SDC on ESX** screen appears, showing all the ESX hosts on all the connected vCenters:



Note

Entering the cluster password will automatically enter the same password for all servers in the cluster.

2. Select the ESX hosts, and enter the root password for each host.

Note

It is highly recommended to select all ESX hosts that may be included in an ESX system, even if only in the future.

3. Click **Install**.

   The status appears in the dialog.

4. When finished, click **Finish**.
5. Restart each ESX host.

Note

You must restart the ESX hosts before proceeding.

6. When using all-SSD chassis, perform this step.

   Sometimes the ESX nodes in an all-SSD chassis do not identify the node disks as Standard Parallel SCSI devices. If so, you must enable the **RDMs on non Parallel SCSI Controllers** option.

a. In the **EMC ScaleIO** screen, click **Advanced Settings** to display the settings options.

b. Select the **Enable RDMs on non Parallel SCSI Controller** option and click **OK**.

After you finish

Proceed with the ScaleIO deployment.

Deploying ScaleIO

This section describes how to use the deployment wizard to deploy ScaleIO in the VMware environment.

You can use the wizard to deploy ScaleIO from the beginning, or to register a ScaleIO system that has already been deployed, but not registered with vSphere.

To register an existing system, click **Register ScaleIO system**, and enter the credentials.

Before you begin, you may want to configure advanced installation options:

- Enable creation of VMDK
- Enable RDMs on non-parallel SCSI controllers
- Allow taking over devices that were used in other ScaleIO systems
- Allow the use of non-local datastores for the Gateway
- Increase parallelism limit

To access these settings, click **Advanced settings** on the **EMC ScaleIO** screen. For more information, see "Advanced settings options".

To deploy a new system, perform the following:

Procedure

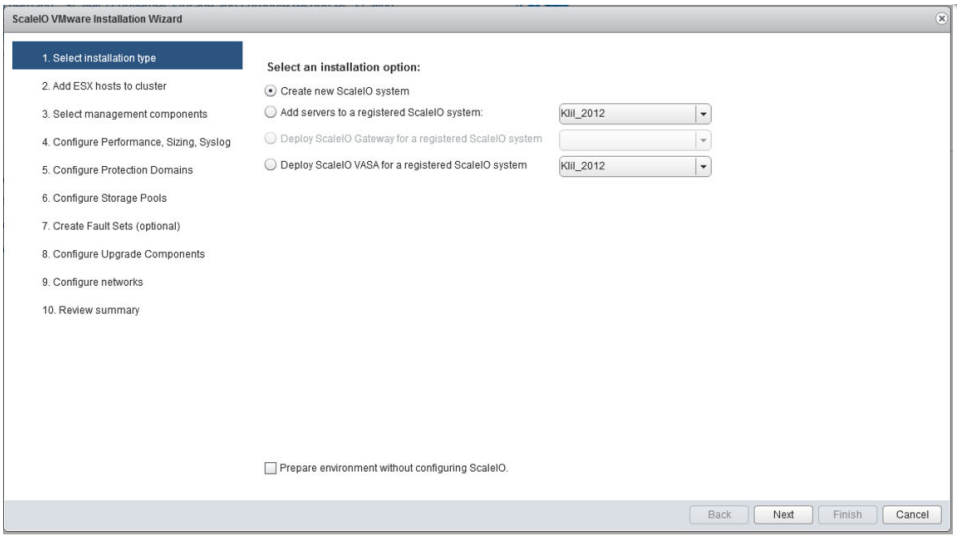
1. From the **Basic tasks** section of the screen, click **Deploy ScaleIO environment**.

   The ScaleIO VMware Deployment Wizard begins. If you exited the previous deployment before completion, you will be able to return from where you left off.

   NOTICE

   The deployment wizard assumes that you are using the provided ScaleIO OVA template to create the ScaleIO virtual machines.

   The **Select Installation** type screen appears:

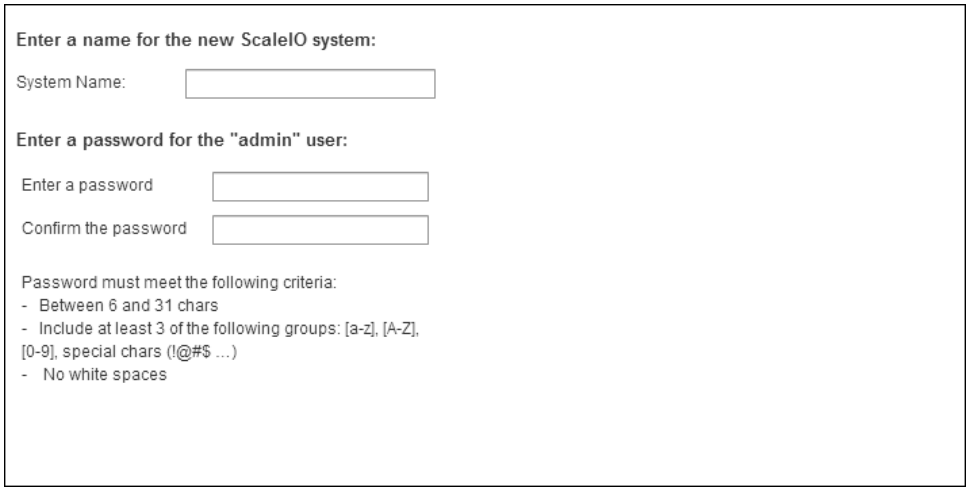


2. In the **Select Installation** screen, select one of the following, then click **Next**:
   - To create a new system, select **Create new ScaleIO system**.
   - To extend a system that was created previously, select **Add servers to a registered ScaleIO system**, then select the existing system from the drop-down list.

     Note

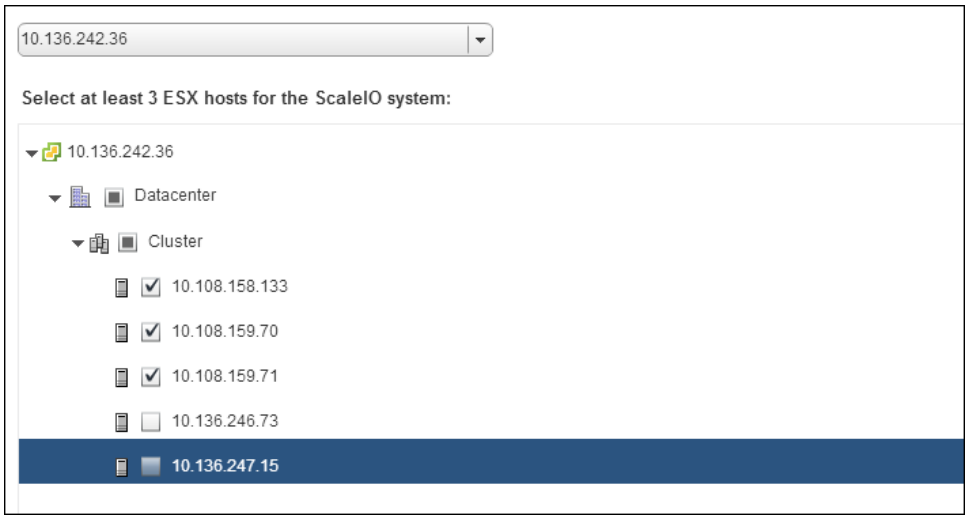
     Before extending an existing SVM with a new role or component, you must perform manual memory allocation on the SVM, as described in "SVM manual memory allocation".
   - To deploy a Gateway on a system that was created previously, select **Deploy ScaleIO Gateway for a registered ScaleIO system**, then select the existing system from the drop-down list.
   - To use the wizard to prepare the environment without performing ScaleIO configuration (Protection Domains, Storage Pools, and so on), select **Prepare environment without configuring ScaleIO**.

3. If you selected **Create a new ScaleIO system**, review and agree to the license terms, then click **Next**.
   If you selected **Create a new ScaleIO system**, the **Create New System** screen appears:



4. In the **Create New System** screen, enter the following, then click **Next**:
   a. System Name: Enter a unique name for this system.
   b. Admin Password: Enter and confirm a password for the ScaleIO admin user. The password must meet the listed criteria.

   The **Add ESX Hosts to Cluster** screen appears:



5. Select the vCenter on which to deploy the ScaleIO system.

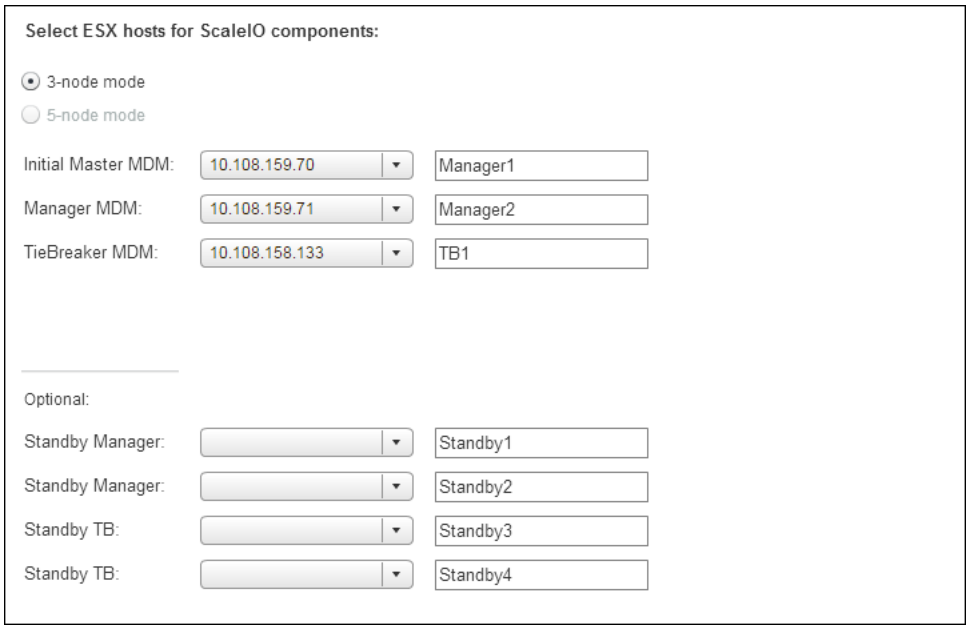
   The vCenter information is populated in the lower part of the screen.

6. Open the vCenter, select the ESX hosts to add to the ScaleIO system, then click **Next**.

Note

To configure ScaleIO, you must select a minimum of three ESX hosts. ESX hosts that do not have the SDC installed will not be available.

The **Select Management Components** screen appears:



7. In the **Select Management Components** screen, perform the following:
   a. Select to deploy a 3-node or 5-node cluster.

      The fields on this screen will change, depending on your choice.

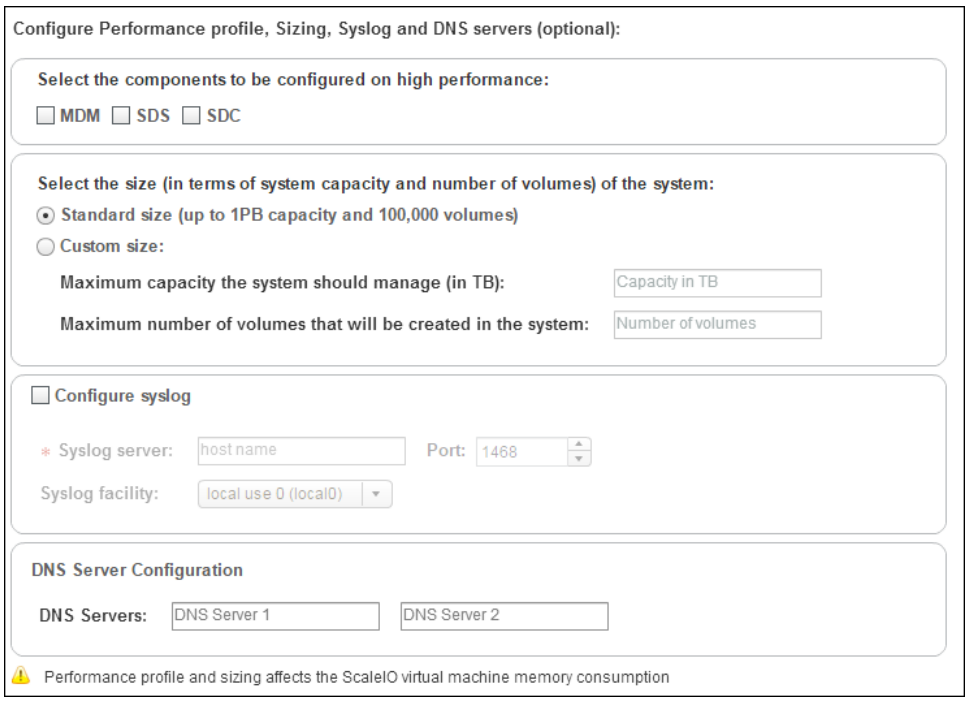
   b. Select an ESX server to serve for each of the MDM cluster roles.

      You can give a name to the MDM servers, such as Manager1, etc.

   c. Select ESX servers to serve as Standby manager and tiebreaker roles (optional).

   d. Click **Next**.

      The **Configure Performance, Sizing, and Syslog** screen appears:



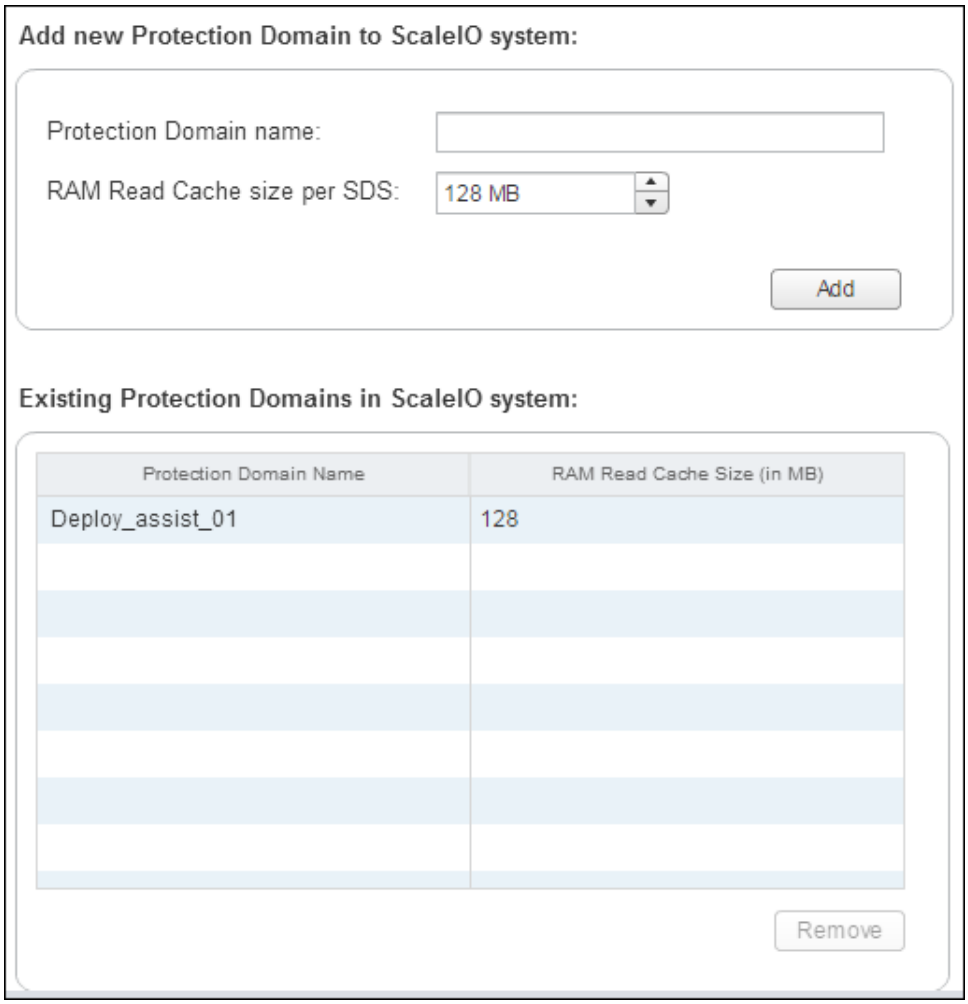
8. Configure the following settings (optional), then click **Next**:
   a. To select components to be configured for high-performance, choose those components.

   b. To configure the allocation of SVM RAM, choose from the following:
      - To use default RAM allocation, select **Standard size**.
      - To use custom settings, select **Custom size**, and enter the maximum capacity and maximum number of volumes.

   c. To configure syslog reporting, select **Configure syslog**, and enter the syslog server, port (default: 1468), and facility (default: 0).

   d. To configure DNS servers, enter their details.

      The **Configure Protection Domains** screen appears:



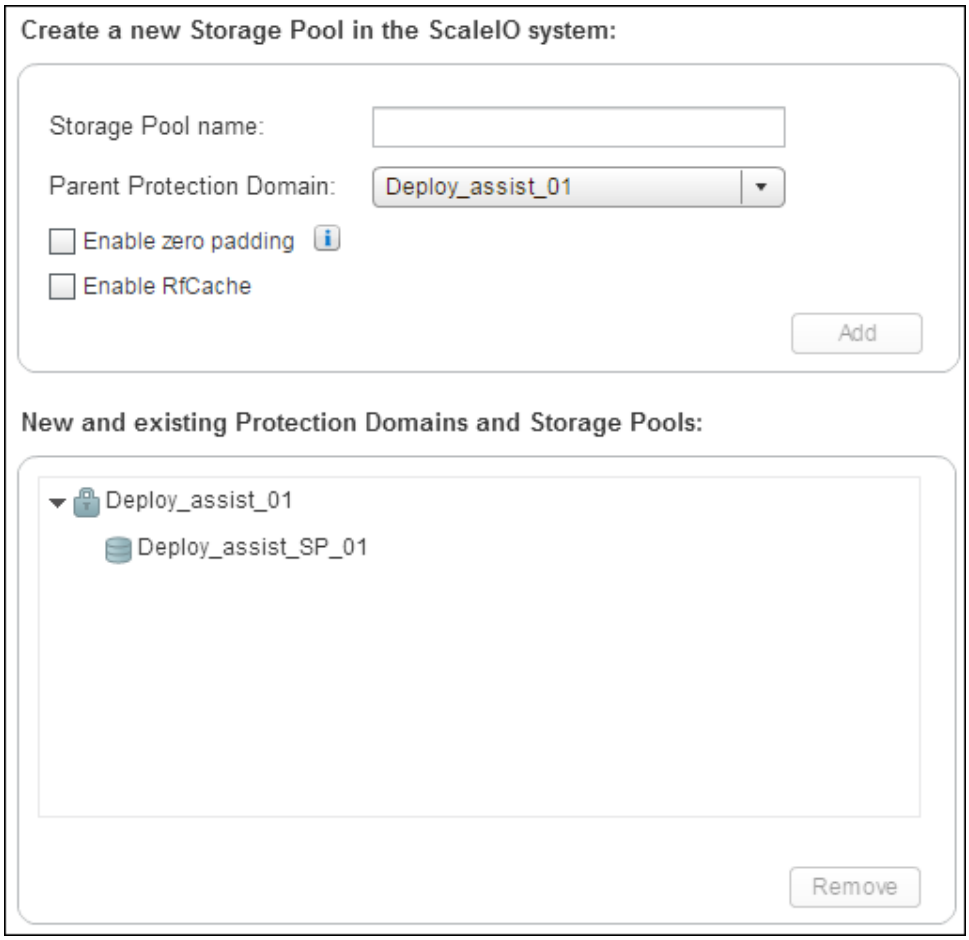
In the **Configure Protection Domains** screen, you can create (or remove) Protection Domains (PD). You must create at least one PD.

9. To create a PD, enter the following, then click **Add**:
   a. Protection Domain name: It is recommended to use a meaningful name.
   b. RAM Read Cache size per SDS: Minimum 128 MB (You can increase this for your environment needs.)

   The added PDs appear in the lower section of the screen, together with the existing PDs. To remove a newly created PD, select it and click **Remove**.

10. Repeat to create additional PDs, then click **Next**.

    The **Configure Storage Pools** screen appears:



In the **Configure Storage Pools** screen, you can create (or remove) Storage Pools (SP). You must create at least one SP.

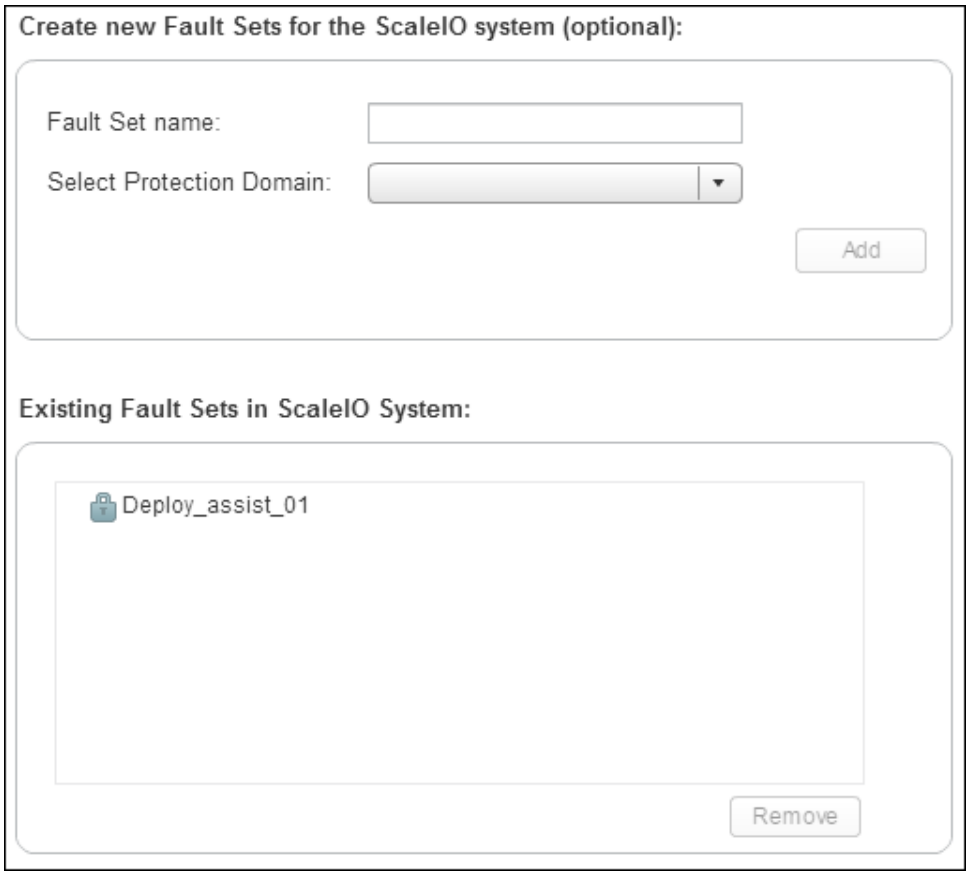
11. To create a SP, configure the following, then click **Add**:
    a. Storage Pool name: It is recommended to use meaningful names.
    b. Select to which PD to add the SP.
    c. To enable zero padding, select **Enable zero padding**.

       Zero padding must be enabled for use with RecoverPoint replication. Zero padding is required for using the background scanner in data comparison mode.

    d. To enable Read Flash cache, select **Enable RFcache**.

       Added SPs appear in the lower section of the screen, inside the folder of the parent PD. You can remove a newly created SP by selecting it and clicking **Remove**.
12. Repeat to create additional SPs, then click **Next**.

    The **Create Fault Sets** screen appears:



You can use this screen to create Fault Sets (optional).

Note

When defining Fault Sets, you must follow the guidelines described in "Fault Sets". Failure to do so may prevent creation of volumes.

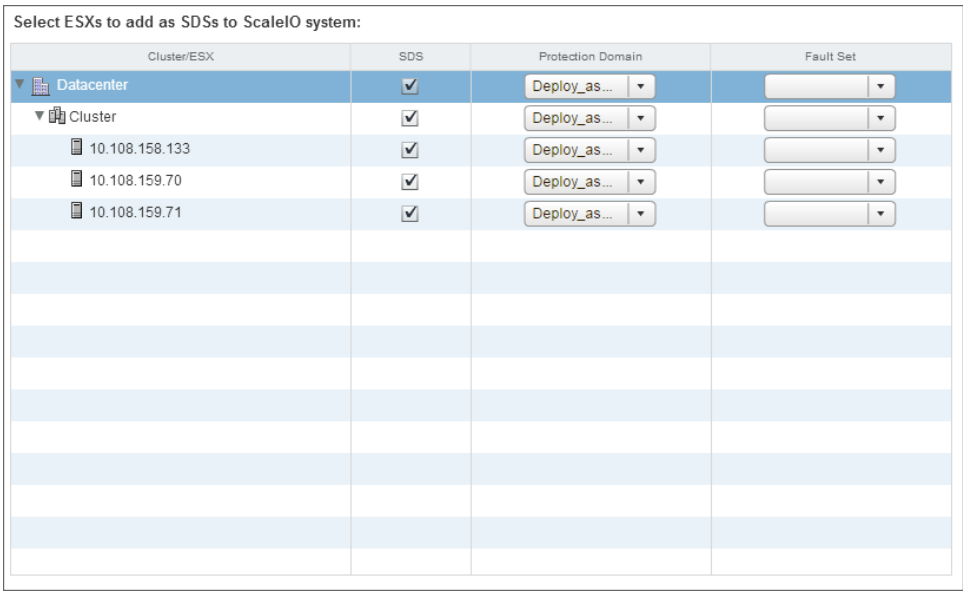
---

13. To create a Fault Set, enter the name, select to which PD to add the Fault Set, then click **Add**. It is recommended to use meaningful names.

    Added Fault Sets appear in the lower section of the screen, inside the folder of the parent PD. You can remove a newly created Fault Set by selecting it and clicking **Remove**.

14. Repeat to create additional Fault Sets (minimum of three), then click **Next**.

    The **Add SDSs** screen appears:

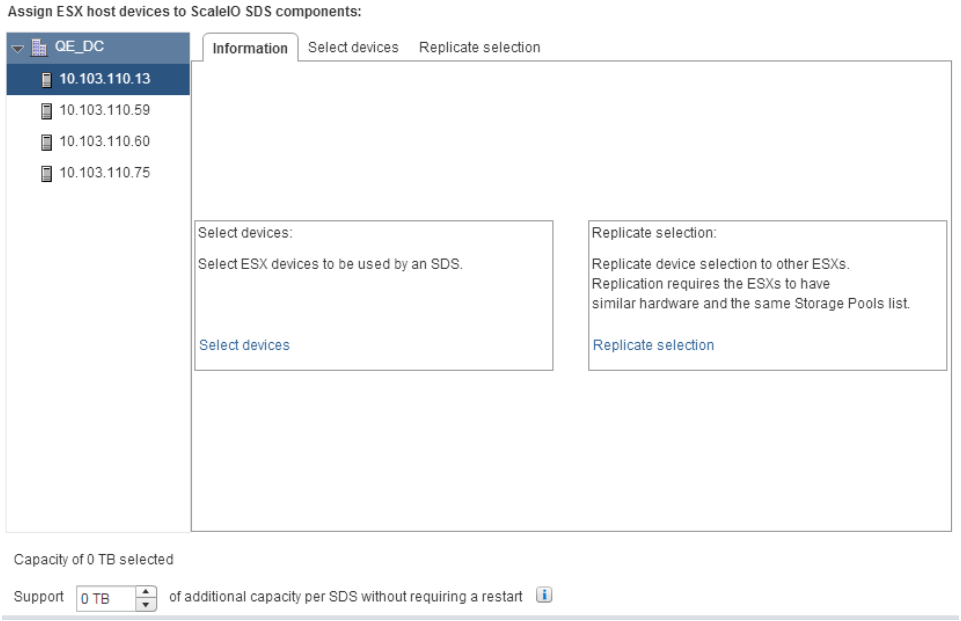


Note

To make the same selections for every ESX in a cluster, you can make your selections per cluster or datacenter.

15. Configure the following for every ESX host or SVM, then click **Next**:
    - Whether it is acting as an SDS.
    - If the SVM is an SDS, select a Protection Domain (required) and Fault Set (optional).

    The **Add devices to SDSs—Information tab** appears:



When this screen appears, only the clusters are displayed.

Open the clusters, and select an ESX host.

From this screen, you can perform the following:

a. Select devices (from the **Select devices** tab)

Select storage devices to add to a single SDS.

b. **Support ... additional capacity per SDS without requiring a restart**

Allocate additional RAM for this future need

c. Replicate selections (from the **Replicate selections** tab)

Select devices for other SDSs, by replicating selections made in the **Select devices** tab. This can be very useful if your ESXs have identical attached devices.

Once you teach the deployment wizard what kind of selections you made for one SDS, the source, in the **Select devices**tab, you can use this option to easily select all of the same type of devices for other target SDSs. For example, if you select an SSD device for the source ESX, and then replicate this selection to the targets, the deployment wizard can automatically select all other SSD devices on the target SDSs.

Note

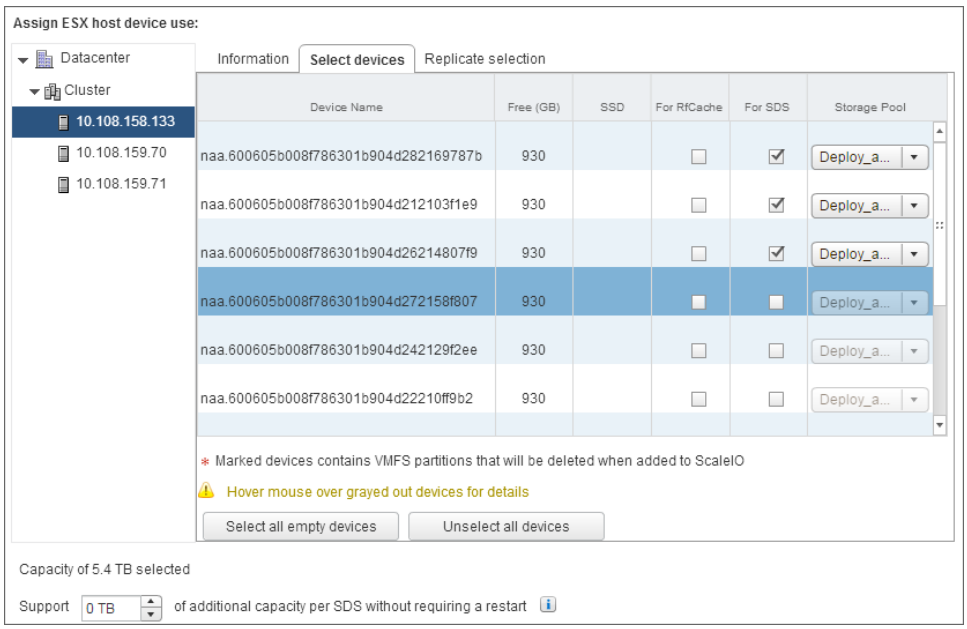
Device matching is performed based on the device runtime name.

To replicate device selections, all of the following conditions must be met:

- The number of devices on each ESX must be identical.
- Source and target devices must be identical in the following ways: a) both are SSD or non-SSD, b) both have datastores on them or do not, c) both are roughly the same size (within 20%), and d) both are connected via a RAID controller or directly attached.
- At least one of the following conditions must be met: a) both SDSs are in the same Protection Domain, b) both SDSs are in different Protection Domains, but with the same list of Storage Pools, or c) the target SDS is in a Protection Domain with only one Storage Pool.

16. Click **Select devices**.

The **Select devices** tab appears:



This screen shows the devices whose free space can be added to the selected ESX host/SDS. You should balance the capacity over the selected SDS.

If the device is an SSD device and has an HBA330 controller, a check is displayed in the SSD column.

17. To add a device's space to an SDS, perform the following:
    a. To use the SSD capacity to provide RFcache acceleration, select **For RFcache**.
    b. To use a device's capacity (SSD or HDD) for storage, select **For SDS**.
    c. Choose the Storage Pool to which to assign the device.

    Note

    If the selected SP has RFcache enabled, you must select at least one RFcache device for that SDS node.

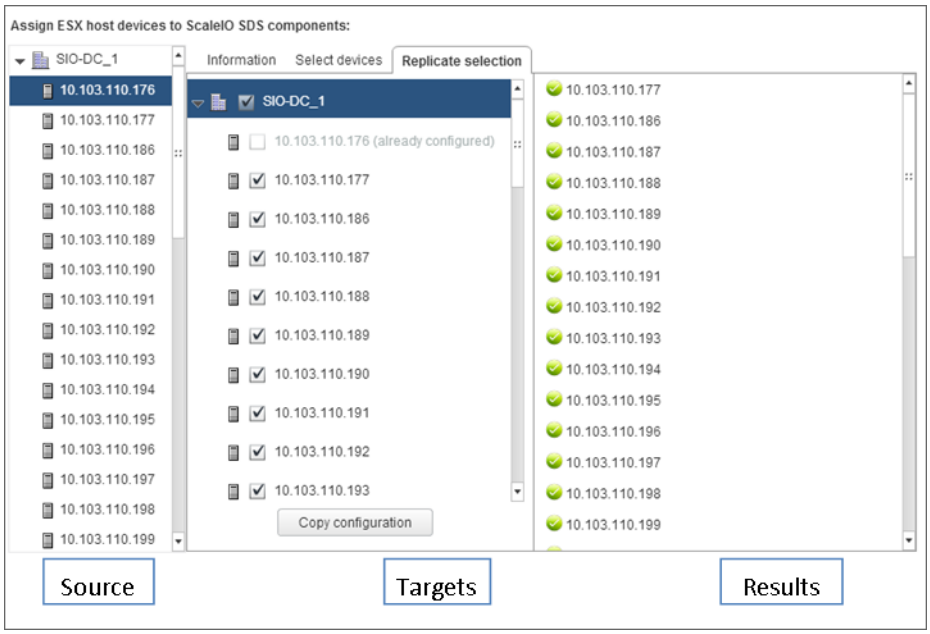
    It is not recommended to mix device types in the same Storage Pool.

    Note

    If you selected to create VMDK (before the deployment), the following options appear:

    - **Create VMDK**. Select this for all relevant devices.
    - **Select all available devices**. Click this to select all devices with a VMFS, and with unused capacity that can be added to the ScaleIO system.

18. To replicate selections to other SDSs, perform the following:
    a. Select the **Replicate selection** tab.



    b. Select the ESX whose device selection you wish to replicate.

    This is the source ESX.

c. Select the target ESXs to which to replicate the selection of the source ESX.

d. Click **Copy configuration**.

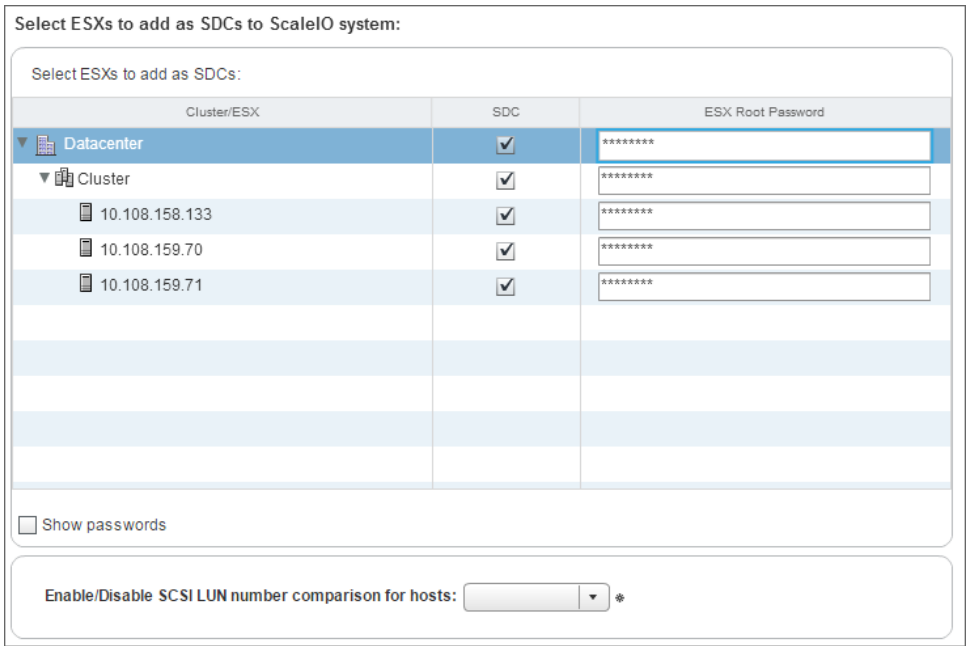
The results are displayed in the right pane of the screen.

19. When you have selected devices for all SDSs, click **Next**.

Note

You must select at least one device for each SDS.

The **Add SDCs** screen appears:



20. Enter the SDC settings, as follows:

a. For each ESX to be added as an SDC, perform the following:

a. Select the **SDC** check box.

b. Type the ESX root password.

Note

To show the entered ESX passwords, select **Show passwords**.

b. Choose whether to enable or disable the LUN comparison for ESX hosts.

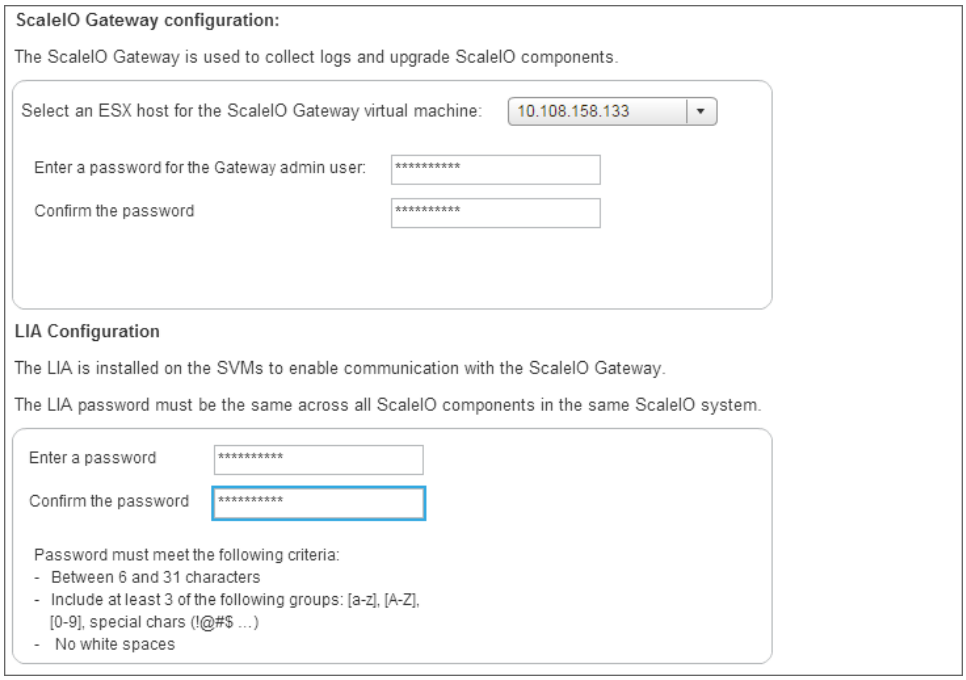
In general, in environments where the SDC is installed on ESX and also on Windows/Linux hosts, you should set this to **Disable**.

Note

Before enabling LUN comparison, consult your environment administrator.

c. Click **Next** to continue.

The **Configure Upgrade Components** dialog appears:



21. Configure the ScaleIO Gateway, by performing the following:
    a. Select an ESX to host the Gateway virtual machine.

       A unique SVM will be created for the Gateway.
       If the previously-selected ESX servers do not have sufficient free space (on any datastore) to contain the ScaleIO SVM template, an SVM, and the Gateway SVM, you will not have an option to select an ESX in this step. It will be done automatically.

    b. Enter and confirm a password for the Gateway administrative user.
22. Configure the Lightweight Installation Agent (LIA) on the SVMs:
    a. Enter and confirm a password for the LIA.

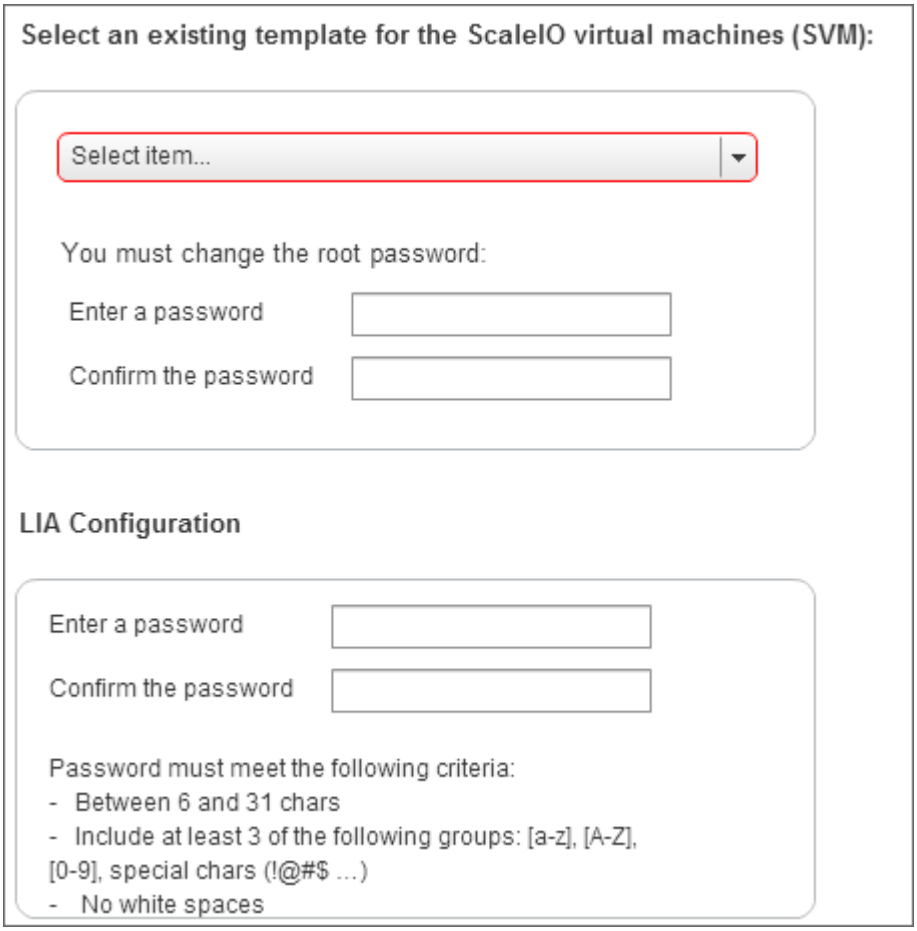
       The password must be the same across all SVMs in the system.

    b. Click **Next**.

       Note

       You can only move forward if the passwords meet the listed criteria, and if the confirmation passwords match the entered passwords.

       The **Select OVA Template** screen appears:



23. In the **Select OVA Template** screen, perform the following, then click **Next**:

    a. Select the template to use to create the ScaleIO virtual machines (SVM).

    The default is **EMC ScaleIOSVM Template**. If you uploaded a template to multiple datastores, you can select them all, for faster deployment.

    If the Gateway selection was performed automatically in the previous step (indicating insufficient space), you must choose at least two templates in this step, one of which will be converted to the Gateway SVM.

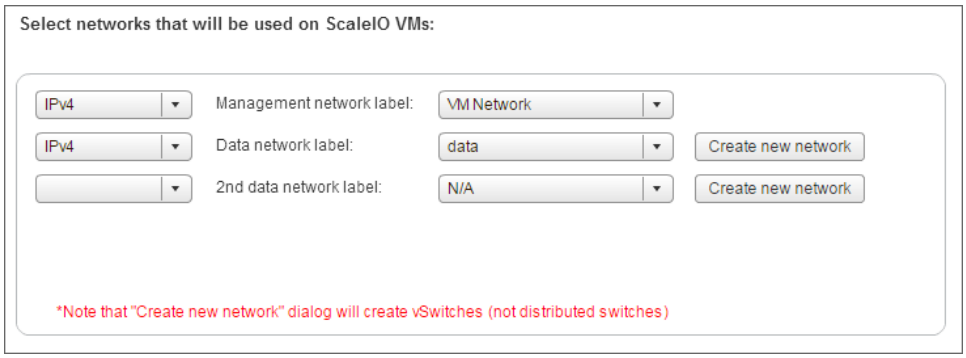
    After selecting the templates, the deployment wizard will automatically select one of the ESXs with the templates to host the Gateway and during deployment will convert the template to a VM for the Gateway (instead of cloning the template).

    Note

    If you select a custom template, ensure that it is compatible with the VMware plug-in and the ScaleIO MDM.

    b. Enter and confirm a new root password, that will be used for all SVMs to be created.

    The **Configure Networks** screen appears:



In the **Configure Networks** screen, you choose to use either a single network for management and data transfer, or to use separate networks. Separating the networks is recommended for security and increased efficiency. You can select one data network or two.

The management network, used to connect and manage the SVMs, is normally connected to the client management network, a 1 GB network.

The data network is internal, enabling communication between the ScaleIO components, and is recommended to be at least a 10GB network.

Note

The selected networks must have communication with all of the system nodes. In some cases, while the wizard does verify that the network names match, this does not guarantee communication, as the VLAN IDs may have been manually altered.

---

24. Enter the following information:

    a. To use one network, select a protocol (IPv4 or IPv6), and a management network, then proceed to step 25 .

    NOTICE

    When the MDM and SDS are on physical servers, IPv6 is fully supported. Manually configure the IPv6 on the ESX (VMKernel). IPv6 is not supported in VMware converged solutions. Contact EMC support for more information.

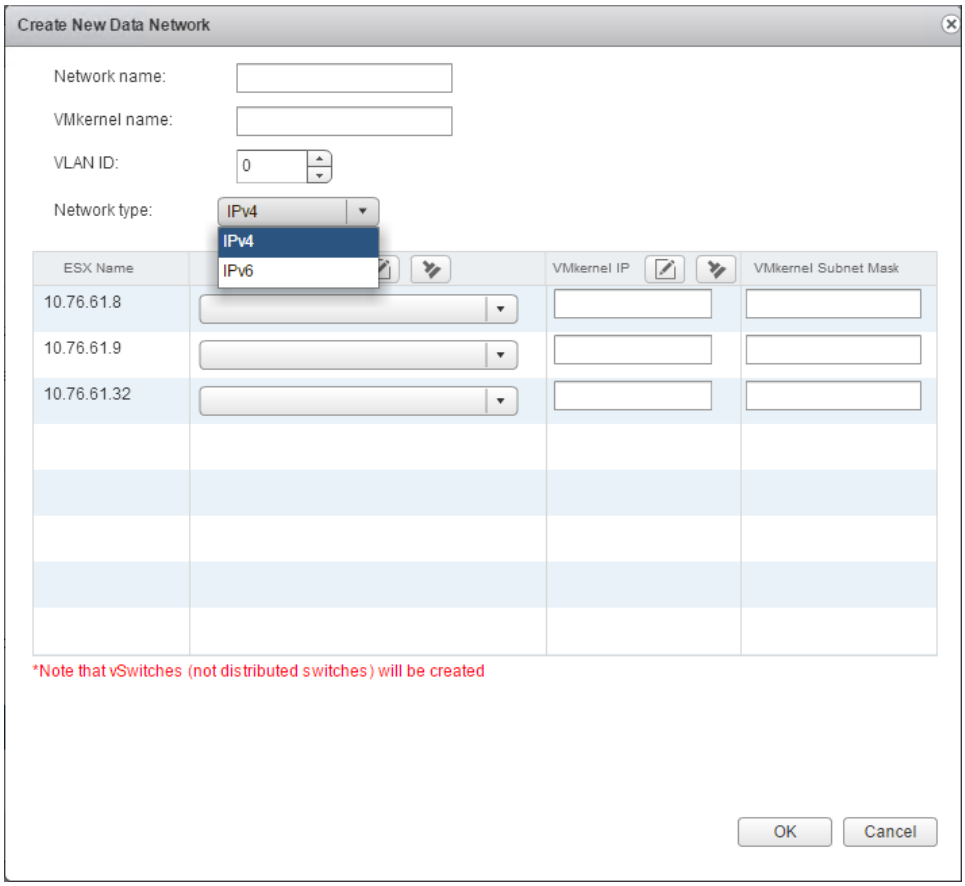
    Note

    For best results, it is highly recommended to use the plug-in to create the data networks, as opposed to creating them manually.

    ---

    b. To use separate networks, select a protocol (IPv4 or IPv6) for the management network label and one or two data network labels. If the data network already exists, select it from the drop-down box. Otherwise, configure the data network by clicking **Create new network**.

    The **Create New Data Network** screen appears:



In the **Create New Data Network** screen, enter the following information:

- **Network name**: The name of the VMware network
- **VMkernel name**: The name of the VMkernel (used to support multipathing)
- **VLAN ID**: the network ID
- **Network type**: IPv4 or IPv6
- For each listed ESX, select a **Data NIC**, a **VMkernel IP**, and a **VmKernel Subnet Mask**.

Note

You can click to auto-fill the values for **Data NIC** and **VMkernel IP**.

---

c. Click **OK**.

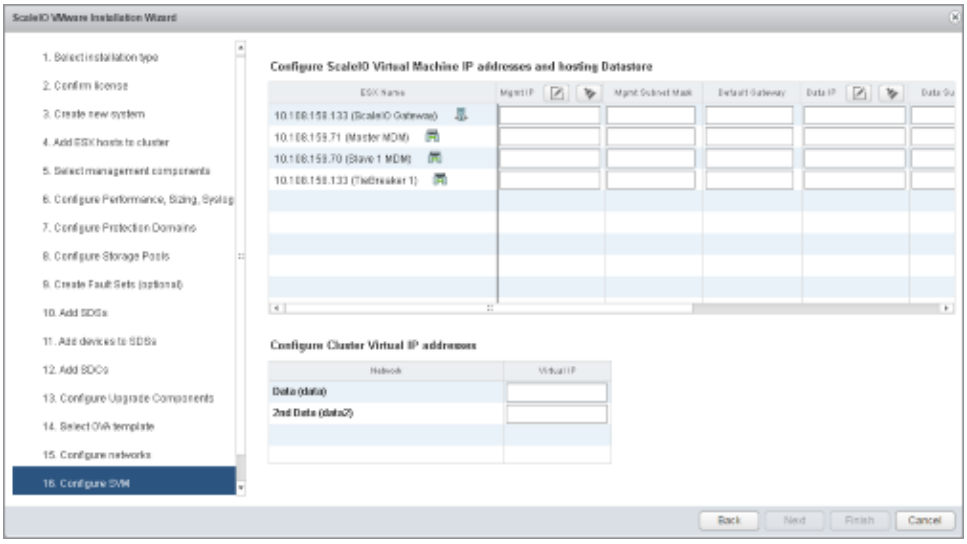
The data network is created.

The wizard will automatically configure the following for the data network:

- vSwitch
- VMkernel Port
- Virtual Machine Port Group
- VMkernel Port Binding

25. Click **Next**.

The **Configure SVM** screen appears.



26. For each SVM, perform the following, then click **Next**.

    a. Enter the IP address, Subnet mask, and Default Gateway for the management network, then the data network.

    b. Enter the Cluster Virtual IP for each network interface.

    c. You can select a datastore, or allow automatic selection.

    d. Configure the cluster's virtual IP addresses by entering the virtual IP address for each data network.

    e. Click **Next**.

    Icons indicate the role that the server plays in the ScaleIO system. You can select to auto-fill the values for IP.

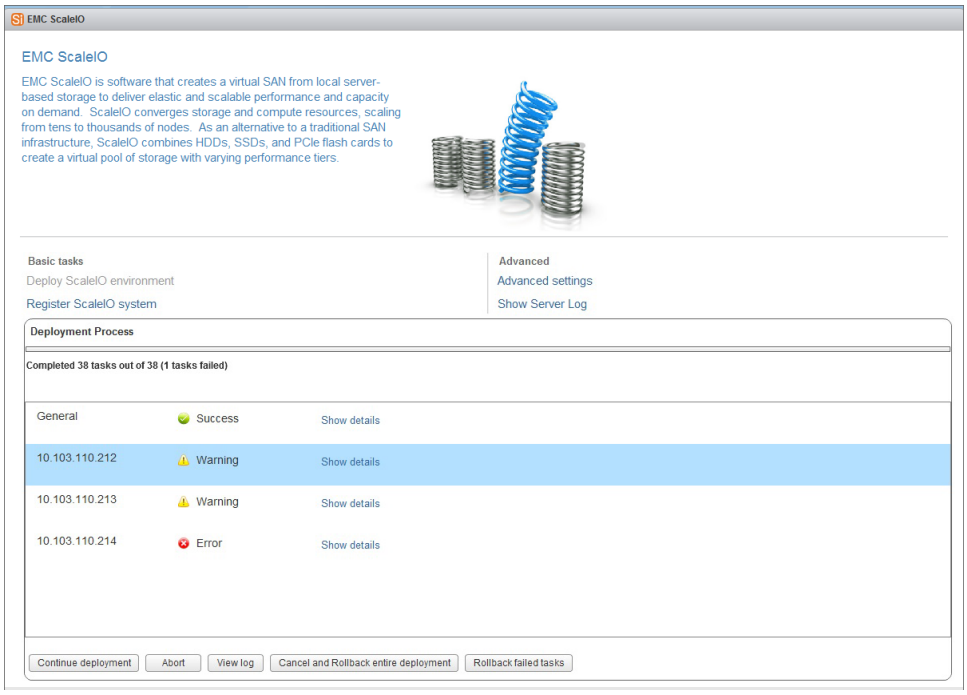
    The **Review Summary** screen appears.

27. Review the configuration.

    Click **Finish** to begin deployment or **Back** to make changes.

28. Enter the vCenter user name and password, then click **OK** to begin the deployment.

    The **Deployment Progress** screen appears:



During the deployment process you can view progress, pause the deployment, and view logs.

To pause the deployment, click **Pause**. Steps that are already in progress will pause after they complete.

After pausing, select one of the following options:

- **Continue deployment** to continue.
- **Abort** to abort the deployment process.
- **Cancel and Rollback entire deployment** to roll back all deployment activities.

  Note

  Rollback cannot be canceled once started.

  For more details on rolling back deployment, see Rolling back the deployment wizard in the vSphere Web plug-in on page 129
- **Rollback failed tasks** to roll back only the tasks that failed.

  Note

  Rollback cannot be canceled once started.

29. When the deployment is complete, click **Finish**.

    If a task failed, click **Continue deployment** to try again.

Rolling back the deployment wizard in the vSphere Web plug-in

When deployment fails due to incorrect input or failed devices, you can roll back the deployment wizard, fix or change the specific deployment setting, and restart deployment.

Procedure

1. When deployment fails, click **Cancel and roll back entire deployment**.
2. In the **Rollback settings** dialog box, select whether you want to remove the SDC configuration.
   - If you select **Remove SDCs configuration (requires a manual restart of the ESXs)**, the ScaleIO VMs are deleted, the MDM IPs are removed from the ESXs, and you must manually restart the ESXs.
   - If you do not select **Remove SDCs configuration (requires a manual restart of the ESXs)**, the ScaleIO VMs are deleted, the MDM IPs are not removed from the ESXs, and you do not need to manually restart the ESXs.
3. When the deployment roll back is complete, select **Finish deployment and return to wizard**.

   If you select **Finish**, you will not be returned to the wizard and will have to start the deployment from the beginning; your deployment details will not be saved. After confirming that you want to finish the deployment and return to the wizard, the last screen in the deployment wizard is displayed.
4. Click **Back** or **Next** to navigate backwards or forwards through the deployment wizard.

   Change the settings that caused the initial deployment to fail.
5. Click **Finish** to complete the wizard and begin deployment.

Where to go from here

Now that your ScaleIO system is up and running, the next task is to create volumes. For more information, see the *EMC ScaleIO User Guide*.

Installing the ScaleIO GUI

This topic describes how to install the ScaleIO GUI. You can install the GUI on a Windows or Linux workstation.

Ensure that the workstation satisfies the minimum requirements, as described in System requirements on page 18.

Procedure

1. Install the GUI:

OS	Command
Windows	`EMC-ScaleIO-gui-2.0-13000.X.msi`
RHEL	`rpm -i EMC-ScaleIO-gui-2.0-13000.X.noarch.rpm`
Debian	Run with administrator privileges: `sudo dpkg -i EMC-ScaleIO-gui-2.0-13000.X.deb`

Results

The GUI is installed.

After you finish

To log in and use the GUI, see the *EMC ScaleIO User Guide*.

CHAPTER 5

Post-Deployment Activities

This chapter describes activities that are performed after deployment of a ScaleIO system. Topics include:

- Post-deployment checklist 134
- Getting started post-deployment 135

Post-deployment checklist

Use the following checklist to verify that you complete the required and optional post-deployment procedures.

Post-deployment procedures

Procedure	Notes	Mandatory?	Done
Install ScaleIO license	A ScaleIO license is necessary when deploying ScaleIO in a production environment.	Optional	
Review ScaleIO performance best practices		Recommended	
Review ScaleIO security best practice configurations		Recommended	
Run the system analysis tool from the Installation Manager (IM), prior to provisioning.	It is highly recommended to run the system analysis tool immediately after deployment, before provisioning volumes, and before using the system in production. You can also use it to check the health of a system that is already operational. The system analysis tool supports only RHEL 6.x servers and IPv4 network configuration. For more information, see System analysis overview on page 212.	Highly recommended	
Add and map volumes	Adding and mapping volumes can be performed using several management tools. For more information, see Create and map volumes on page 139	Mandatory	
Create a LockBox and add the MDM credentials to it	A LockBox can be created during installation with the Installation Manager (IM) when selecting **Set advanced options**. Only when a LockBox has not been created during installation, manually create a LockBox. For more information, see Creating a LockBox on page 135. If you want to use SNMP, ESRS, or LDAP, you must have a LockBox.	Optional	
Configure EMC Secure Remote Support (ESRS)	For more information, see Configuring ESRS connection properties on page 221.	Optional	
Define native users in the system and change the default admin password.	For instructions, see Configuring native users on page 137.	Optional	
Define LDAP users (if using ScaleIO with LDAP).	For instructions, see Configuring LDAP users on page 138. For detailed information, see *EMC ScaleIO User Roles and LDAP Usage Technical Notes*.	Optional	
Enable SNMP and configure the SNMP trap receiver	If SNMP was not configured during installation using the advanced settings in the Installation Manager, it can be configured post-deployment. For more information, see Configuring SNMP after deployment on page 139.	Optional	

Procedure	Notes	Mandatory?	Done
Change monitor user password		Optional	
Back up all SSH keys from nodes and ScaleIO VMs		Optional	

Getting started post-deployment

Perform the following procedures to get started quickly post-deployment:

- Creating a LockBox on page 135
- Configuring ESRS on page 136
- Configuring native users on page 137
- Configuring LDAP users on page 138 (optional)
- Configuring SNMP after deployment on page 139
- Create and map volumes on page 139

Creating a LockBox

Create a LockBox and add the MDM credentials.

Before you begin

- ESRS Gateway v3 version 3.08 or higher must be installed and configured. It is recommended to create at least two ESRS Gateways and define them as cluster via the backend server.
- ESRS Gateway must be reachable from ScaleIO on port 9443.
- The ScaleIO license must be installed.

Ensure you have:

- One or more IP addresses of the ESRS gateway servers. Note that ESRS does not currently support IPv6.
- ESRS username and password.
- ScaleIO Gateway IP address, username, and password.
- MDM username and password.
- The ScaleIO Management IP address to be used as the Connect-In IP address. It must be an IP address that is accessible from the ESRS Gateway (for example, in case of NAT).

Use SioGWTool to configure a LockBox. SioGWTool should be used to create a LockBox only when a Lockbox has not yet been created.

A LockBox can be created during installation with the Installation Manager (IM). For more information on creating a LockBox during installation, see the *Deployment Guide*.

To use SioGWTool, input the appropriate path, based on your operating system, and append the commands to the end of the filepath:

- Linux SioGWTool filepath: `/opt/emc/scaleio/gateway/bin/SioGWTool.sh`
- Windows SioGWTool filepath: `C:\Program Files\EMC\ScaleIO\Gateway\bin\SioGWTool.bat`

Procedure

1. Create a LockBox:

```
<SioGWTool_PATH> --change_lb_passphrase --new_passphrase
<NEW_PASSPHRASE>
```

Windows example:

```
 C:\Program Files\EMC\ScaleIO\Gateway\bin\SioGWTool.bat --
set_mdm_credentials --mdm_user admin --mdm_password Scaleio123
```

2. Add MDM credentials to the LockBox:

```
<SioGWTool_PATH> --set_mdm_credentials --mdm_user
<MDM_USERNAME> --mdm_password <MDM_PASSWORD>
```

Configuring ESRS

Enable EMC Secure Remote Support (ESRS) for remote support.

Before you begin

Ensure that a LockBox has already been created and the MDM credentials have been added to it.

Enable the ESRS feature in the `gatewayUser.properties` file.

Note

Detailed information about configuring ESRS is provided in Before configuring ESRS on page 222.

Procedure

1. Use a text editor to open the `gatewayUser.properties` file, located in the following directory on the Installation Manager/Gateway server:
   - Linux: `/opt/emc/scaleio/gateway/webapps/ROOT/WEB-INF/classes`
   - Windows: `C:\Program Files\EMC\ScaleIO\Gateway\webapps\ROOT\WEB-INF\classes\`
2. Locate the parameter `features.enable_esrs` and edit it as follows:

```
features.enable_esrs=true
```

3. Save and close the file.
4. Add the ESRS Gateway's certificate to the truststore.
   a. In a web browser, browse to `<ESRS_Gateway_IP_address:9443>`.
   b. Download the certificate that is displayed, and save it as a file.

c. Log in to REST and get a token. Make a note of the token.

```
curl -k -v --basic --user
admin:<mdm_admin_password>  https://
<ScaleIO_Gateway_IP_address>/api/login
```

d. Using REST, add the certificate to truststore.

```
curl -k -v --basic -
uadmin:<token_received_from_previous_command> --form
"file=@<path_to_certificate_file>" https://
<ScaleIO_Gateway_IP_address>/api/trustHostCertificate/Mdm
```

5. Restart the scaleio-gateway service:
   - Linux: Run the command `service scaleio-gateway restart`
   - Windows: Restart the EMC Gateway service.
6. Register an ESRS gateway:

```
<SioGWTool_PATH> --register_esrs_gateway --scaleio_gateway_ip
<IP_ADDRESS:PORT> --scaleio_gateway_user <USER> --
scaleio_gateway_password <PASSWORD> --esrs_gateway_ip
<IP_ADDRESS> --esrs_gateway_user <USER> --
esrs_gateway_password <PASSWORD> --connect_in_ip <IP_ADDRESS>
```

Configuring native users

Configure native users using the CLI.

Use the following SCLI commands to create and modify native ScaleIO users. To run CLI commands, you must be logged in. Actual command syntax is operating-system dependent. For more information, see the *EMC ScaleIO CLI Reference Guide*.
SCLI is installed as part of the MDM component and can be found in the following path:

- Linux and VMware: `scli`
- Xen: `siocli`
- Windows: `C:\Program Files\emc\scaleio\MDM\bin`

Procedure

1. Log in as admin user to the CLI:

```
scli --login --username admin --password <PASSWORD>
```

   where *<PASSWORD>* is the password used for the first login.

2. (Optional) Change the default admin password:

```
scli --set_password --old_password <PASSWORD> --new_password
<NEW_PASSWORD>
```

3. Add users:

```
scli --add_user --username <NAME> --user_role <Monitor|
Configure|BackEndConfigure|FrontEndConfigure|Security|
Administrator>
```

A randomly generated password for the created user is returned.

4. (Optional) Log in with the new user and then change its password:

```
scli --login --username <NAME> --password <DEFAULT_PASSWORD>
```

```
scli --set_password --old_password <OLD_PASSWORD> --
new_password <NEW_PASSWORD>
```

This is optional because the GUI, CLI, or REST will enforce a password change if user logs in with the original password.

Configuring LDAP users

Configure LDAP users if you are using LDAP with ScaleIO.

For more information on using LDAP with ScaleIO, see the *EMC ScaleIO User Roles and LDAP Usage Technical Notes*. To run SCLI commands, you must be logged in. Actual command syntax is operating-system dependent. For more information, see the *EMC ScaleIO CLI Reference Guide*.

Procedure

1. Create Active Directory (AD) groups that correspond to the user roles offered by ScaleIO.
2. Use the CLI to add LDAP service to the MDM:

```
scli --add_ldap_service --ldap_service_uri <LDAP_URI> --
ldap_base_dn <LDAP_DN>
```

3. Use the CLI to assign an LDAP group to the user role:

```
scli --assign_ldap_groups_to_roles (--ldap_service_id
<LDAP_SERVICE_ID> | --ldap_service_name <LDAP_SERVICE_NAME>)
[--administrator_role_dn <LDAP_GROUP_DN>] [--security_role_dn
<LDAP_GROUP_DN>] [--backend_config_role_dn <LDAP_GROUP_DN>]
[--frontend_config_role_dn <LDAP_GROUP_DN>] [--
monitor_role_dn <LDAP_GROUP_DN>]
```

4. Use the CLI to set the system to mixed authentication method, LDAP and native:

```
scli --set_user_authentication_method --
native_and_ldap_authentication
```

5. Log in again to apply these changes.

Configuring SNMP after deployment

Configure Simple Network Management Protocol (SNMP) for error reporting, if it was not configured during installation.

Before you begin

Ensure that a LockBox has already been created and the MDM credentials have been added to it. For detailed instructions on creating a LockBox, see Creating a LockBox on page 135.

Enable the SNMP feature in the `gatewayUser.properties` file.

Procedure

1. Use a text editor to open the `gatewayUser.properties` file, located in the following directory on the Installation Manager/Gateway server:
   - Linux: `/opt/emc/scaleio/gateway/webapps/ROOT/WEB-INF/classes`
   - Windows: `C:\Program Files\EMC\ScaleIO\Gateway\webapps\ROOT\WEB-INF\classes\`
2. Locate the parameter `features.enable_snmp` and edit it as follows:

   ```
   features.enable_snmp=true
   ```

3. To add the trap receiver IP address, edit the parameter `snmp.traps_receiver_ip` as follows:

   ```
   snmp.traps_receiver_ip <TRAP_IP_1>, <TRAP_IP_2>
   ```

   The SNMP trap receivers' IP address parameter supports up to two comma-separated or semi-colon-separated hostnames or IP addresses.
4. You can optionally change the following parameters:

Option	Description
snmp.sampling_frequency	The MDM sampling period. The default is 30.
snmp.resend_frequency	The frequency of resending existing traps. The default is 0, which means that traps are sent all the time.

5. Save and close the file.
6. Restart the scaleio-gateway service:
   - Linux: Run the command `service scaleio-gateway restart`
   - Windows: Restart the EMC Gateway service.

Create and map volumes

Create volumes from devices added to SDS nodes, and then map the volumes to SDC nodes.

You can create and map volumes using various management tools, including the CLI, GUI, vSphere web plug-in, and REST API.

Creating and mapping volumes is necessary before applications can access the volumes. In addition, you may create additional volumes and map them as part of the maintenance of the virtualization layer. For additional details on creating and mapping volumes, see the user documentation.

Creating and mapping volumes using the CLI

Create a volume and map a volume to an SDC using the CLI.

Use the following SCLI commands to create and map volumes. For more details on SCLI command usage, see the *EMC ScaleIO CLI Reference Guide*.

Procedure

1. Create a volume:

```
scli --add_volume --size_gb <SIZE> --volume_name <VOL_NAME> --
protection_domain_name <PD_NAME> --storage_pool_name <SP_NAME>
```

2. Map a volume to an SDC:

```
scli --map_volume_to_sdc --volume_name <VOL_NAME> --sdc_ip
<SDC_IP>
```

Creating and mapping volumes using the GUI

Create a volume and map a volume to an SDC using the GUI.

For more information on creating and mapping volumes using the GUI, see the user documentation.

Procedure

1. Open the GUI and log in with credentials:
   - Windows: Select **Windows Start** > **All Programs** > **EMC** > **EMC ScaleIO GUI**
   - Linux: run the script `/opt/emc/scaleio/gui/run.sh`
2. Creating volumes:
   a. In the **Frontend** > **Volumes** view, select the Storage Pool to which you want to add the volume.
   b. From the **Command** menu or context-sensitive menu, select **Add Volume**.

      The **Add Volume** window is displayed.
   c. To create more than one volume, select **Create multiple volumes** and type the number of volumes in the **Copies** box.
   d. Type a name in the **Name** box.
   e. To start from a specific number other than 1, type it in the **Start numbering at** box.

      This number will be the first number in the series that will be appended to the volume name.
   f. Type a number in the **Size** box, representing the volume size in GB (allocation granularity is 8 GB).

g. Select **Thick** (default) or **Thin** provisioning options.

h. Enable or disable the RAM Read Cache feature by selecting or clearing **Use RAM Read Cache**.

i. Click **OK**.

3. Mapping volumes to an SDC:

   a. In the **Frontend** > **Volumes** view, select the volumes.

   b. From the **Command** menu or context-sensitive menu, select **Map Volumes**.

      The **Map Volumes** window is displayed, showing a list of the volumes that will be mapped.

   c. In the **Select Nodes** panel, select one or more SDCs to which you want to map the volumes.

   d. Click **Map Volumes**.

      The progress of the operation is displayed at the bottom of the window. It is recommended to keep the window open until the operation is completed, and until you can see the result of the operation.

Creating and mapping volumes using the vSphere plug-in

Create a volume and map a volume to an SDC using the vSphere plug-in.

Procedure

1. To open the plug-in, from the vSphere Web Client home tab, click:

2. From the **Storage Pools** screen, click **Actions** > **Create volumes**.
3. In the **Create Volume** dialog, enter the volume information.
4. To map the volume to ESXs, select **Map volume to ESXs**.
5. In the **Select ESXs** area, select the clusters or ESXs to which this volume should be mapped.

PART 3

Reference

This part supplies additional information on advanced reference topics. Chapters include:

Chapter 6, "Manual Deployment"

Chapter 7, "Advanced Topics"

Chapter 8, "Maintaining a ScaleIO System"

Chapter 9, "System Analysis"

Chapter 10, "Configuring ESRS connection properties"

Chapter 11, "Troubleshooting"

Chapter 12, "Frequently Asked Questions"

CHAPTER 6

Manual Deployment

This chapter describes how to deploy ScaleIO components in situations when the Installation Manager and VMware deployment wizard are not suitable.

NOTICE

This chapter is for reference purposes only. It is highly recommended to install ScaleIO with the Installation Manager or the VMware deployment wizard. If that is not possible, contact EMC Support.

Topics include:

- Manual deployment on physical servers 146
- Deploying ScaleIO on ESX 168
- Manually removing the VIB file from the ESX host 173

Manual deployment on physical servers

This section describes how to deploy ScaleIO manually on Linux and Windows physical servers.

Note

ScaleIO installation enables unlimited use of the product, in non-production environments. To obtain a license for production use, and to receive technical support, open a service ticket with EMC Support at https://support.emc.com.

For complete information on licensing, see the *EMC ScaleIO User Guide*.

Before deploying SDC or RFcache on Ubuntu, Oracle Linux (OL), or CoreOS servers, see the requirements described in Deploying on Ubuntu, Oracle Linux, or CoreOS servers on page 90.

To deploy ScaleIO manually, you must install the following components:

- MDM

  This package is for the MDM and TieBreaker (TB) nodes. You can install MDM in Cluster mode (either 3-node or 5-node) or in Single mode (a single MDM on one node).

  NOTICE

  It is not recommended to use Single Mode in production systems, except in temporary situations. The MDM contains all the metadata required for system operation. Single Mode has no protection, and exposes the system to a single point of failure.

- SDS

  The SDS is installed on all servers that will contribute storage devices to the ScaleIO system.
- SDC

  The SDC is installed on every server that will expose ScaleIO volumes to the application running on it.
- LIA

  Install the Lightweight Installation Agent (LIA) on every node in the ScaleIO system. This enables the performance of various maintenance operations (upgrade, uninstall, and others).
- Gateway (optional)

  The Gateway can be installed on a Windows or Linux server, within, or external to, the ScaleIO system. To use the Gateway, you must also install the LIA on every server.

This section describes how to perform a manual installation of the ScaleIO components on Linux or Windows servers. It is highly recommended to use the automated Installation Manager as described in Deploying ScaleIO on Physical Servers on page 61.

Before installing, ensure that your servers meet the requirements listed in System requirements on page 18.

If needed, install the v2.0.1 `GPG-RPM-KEY` file before beginning the deployment. This is used to used to authenticate the RPM packages.

Manual installation may require use of SCLI commands. These commands are described, in full, in the *ScaleIO CLI Reference Guide*.

You can deploy as a 3-node or 5-node cluster.

The procedures are described with Linux RHEL/CentOS steps.

The following table describes modifications for deploying on other operating systems:

OS	Modifications required
Ubuntu/OL	Before installing ScaleIO, ensure that you have followed the procedures described in Deploying on Ubuntu, Oracle Linux, or CoreOS servers on page 90. • ScaleIO component Ubuntu packages are delivered as TAR files. Before installing, perform the following: 1. Untar all the packages: `tar -xvf <tar_file>` This yields SIOB files. 2. Extract the DEB from the SIOB files: `siob_extract <siob_file>` 3. The DEB file is in the following format: `EMC-ScaleIO-mdm-2.0-13000.X.Ubuntu.X.X.x86_64.deb` You will use the extracted DEB files for the installation. • Some commands are a bit different, noted where applicable.
CoreOS	Before installing ScaleIO, ensure that you have followed the procedures described in Deploying on Ubuntu, Oracle Linux, or CoreOS servers on page 90. • ScaleIO component CoreOS packages are delivered as TAR files. Before installing, perform the following: 1. Untar all the packages: `tar -xvf <tar_file>` This yields SIOB files. 2. Extract the BSX from the SIOB files: `siob_extract <siob_file>` 3. The BSX file is in the following format: `EMC-ScaleIO-mdm-2.0-13000.X.CoreOS.x86_64.deb` You will use the extracted BSX files for the installation. • Some commands are a bit different, noted where applicable.

OS	Modifications required
Windows	• Windows package names are in the following format: `EMC-ScaleIO-mdm-2.0-13000.X.msi` • To restart services, use **Services** > **Restart**. • Some commands are a bit different, as noted where applicable.

Proceed to the section that describes the cluster mode to install:

"Deploying a 3-node cluster"

"Deploying a 5-node cluster"

MDM cluster setup

You can set up the MDM cluster in the following modes:

- 3-node cluster

  One Master MDM, one Slave MDM, and one Tie-Breaker. This cluster can sustain a single point of failure.

- 5-node cluster

  One Master MDM, two Slave MDMs, and two Tie-Breakers. This extended cluster can sustain two points of failure.

- Single-node cluster

  One Master MDM. Single-node cluster has only one copy of the repository, thus it cannot sustain a failure. It is not recommended to use Single Mode in production systems, except in temporary situations.

  An MDM can be installed on any server. For example, one MDM can run on an application server, while the second MDM can run on any other server, for example one of the SDS nodes.

  Note

  For CLI commands, in cluster mode, use the IP address of the Master MDM or the virtual IP address.

  The following is the proper order for creating an MDM cluster:

Order of creating a cluster:

1. Install the MDMs with Manager role (2 in a 3-node, or 3 in a 5-node cluster).
2. Install the MDMs with TB role (1 in a 3-node, or 2 in a 5-node).
3. Create a single-node cluster from one of the Manager MDMs.
4. Add Slave MDMs to the cluster, as Standbys.
5. Add TBs to the cluster, as Standbys.
6. Switch the cluster to 3 or 5-node mode, while designating the standby Slave and standby TB names in the cluster.

Deploying a 3-node cluster

For this example for a 3-node cluster, we will use the following sample topology:

Table 11 Sample manual installation topology

Role	Management IP	MDM IP	Virtual IP address	Virtual IP NIC	MDM name	Cluster role
MDM 1	10.103.110.152	192.168.1.152	192.168.111.152	eth4	mdm152	Master
MDM 2	10.103.110.153	192.168.1.153	192.168.111.152	eth4	mdm153	Slave
TB 1	10.103.110.154	192.168.1.154			tb154	TB
SDS and SDC will be installed on all nodes.						

Installing the packages

In this section, be sure to use the installation packages and commands that are appropriate for your operating system environment.

Note

If you intend to use the replication feature via RecoverPoint on RHEL 6.x servers, refer to the *EMC ScaleIO Write Splitter for RecoverPoint Technical Notes* before beginning the installation. Replication support is version-specific; see the ESSM for full details.

To install ScaleIO components, perform the following:

Procedure

1. Copy the following file to all MDM and TB servers (MDM 1, MDM 2, and TB 1):
   a. `EMC-ScaleIO-mdm-2.0-13000.X.elX.x86_64.rpm`
2. Copy the following files to all servers that will be SDC and/or SDS servers (in our example, all the servers):
   - `EMC-ScaleIO-sds-2.0-13000.X.elX.x86_64.rpm`
   - `EMC-ScaleIO-sdc-2.0-13000.X.elX.x86_64.rpm`
3. Install the MDM and configure the Manager role, by running the following command on MDM 1 and MDM 2:

OS	Command
RHEL/ CentOS /OL	`MDM_ROLE_IS_MANAGER=1 rpm -i <mdm_path.rpm>`

OS	Command
	Example `MDM_ROLE_IS_MANAGER=1 rpm -i EMC-ScaleIO-mdm-2.0-13000.X.el6.x86_64.rpm`
Ubuntu	`MDM_ROLE_IS_MANAGER=1 dpkg -i <mdm_path.deb>`
CoreOS	`MDM_ROLE_IS_MANAGER=1 /.<mdm_path.bsx>`

Windows	`msiexec /i <mdm_path.msi> MDM_ROLE_IS_MANAGER=1`
	Example
	`msiexec /i EMC-ScaleIO-mdm-2.0-13000.X.msi MDM_ROLE_IS_MANAGER=1`

Note

The default MDM credentials are:

- user name: admin
- password: admin

4. Install the MDM on TB 1:

OS	Command

RHEL/ CentOS /OL	`MDM_ROLE_IS_MANAGER=0 rpm -i <mdm_path.rpm>`
	Example:
	`MDM_ROLE_IS_MANAGER=0 rpm -i EMC-ScaleIO-mdm-2.0-13000.X.el6.x86_64.rpm`

Ubuntu	`MDM_ROLE_IS_MANAGER=0 dpkg -i <mdm_path.deb>`
CoreOS	`MDM_ROLE_IS_MANAGER=0 ./<mdm_path.bsx>`

OS	Command

Windows	`msiexec /i <mdm_path.msi> MDM_ROLE_IS_MANAGER=0`
	Example
	`msiexec /i EMC-ScaleIO-mdm-2.0-13000.X.msi MDM_ROLE_IS_MANAGER=0`

5. Install the SDS on every server that will contribute storage drives to the ScaleIO system:

```
rpm -i <sds_path.rpm>
```

Example

```
rpm -i EMC-ScaleIO-sds-2.0-13000.X.el6.x86_64.rpm
```

Note

Including the IP addresses of as many potential MDMs at this time (up to 8), or of the cluster virtual IP addresses, will make it easier to switch MDMs in the future.

6. Install the SDC on every server that will expose ScaleIO volumes to the application running on it. You can add all the IP addresses of each MDM, or the virtual IP addresses for the MDMs.

Note

Before installing SDC or RFcache on Ubuntu, Oracle Linux (OL), or CoreOS servers, perform the steps described in Deploying on Ubuntu, Oracle Linux, or CoreOS servers on page 90.

Option	Description

RHEL/ CentOS / OL	`MDM_IP=<list of virtual IP or MDM IP addresses> rpm -i <sdc_path.rpm>`
	Example
	`MDM_IP=192.168.1.152,192.168.1.153,192.168.1.154 rpm -i EMC-ScaleIO-sdc-2.0-13000.X.el6.x86_64.rpm`

Ubuntu	`MDM_IP=<list of virtual IP or MDM IP addresses> dpkg -i <sdc_path.deb>`

Option	Description
CoreOS	`MDM_IP=<list of virtual IP or MDM IP addresses> ./ <sdc_path.bsx>`

Windows	`msiexec /i <sdc_path.msi> MDM_IP="<list of virtual IP or MDM IP addresses>"`
	Example
	`msiexec /i EMC-ScaleIO-sdc-2.0-13000.X.msi MDM_IP="192.168.1.152,192.168.1.153,192.168.1.154"`

All of the components are installed. The next step is to form the MDM management cluster.

Creating the MDM cluster

Procedure

1. Create the MDM single-node cluster on MDM 1, by running the following command on MDM 1:

```
scli --create_mdm_cluster --master_mdm_ip <IP of MDM 1> --
master_mdm_management_ip <Management IP of MDM 1> --
master_mdm_name <Name to give MDM 1> [--cluster_virtual_ip
<IP>]  [--master_mdm_virtual_ip_interface <INTF>] --
accept_license
```

Example:

```
scli --create_mdm_cluster --master_mdm_ip 192.168.1.152 --
master_mdm_management_ip 10.103.110.152 --master_mdm_name
mdm152 --cluster_virtual_ip 192.168.111.152 --
master_mdm_virtual_ip_interface eth4 --accept_license
```

Note

Each MDM name must be unique. Ensure that each server on which MDM is installed has a unique hostname.

2. Log in and change the MDM password, by running the following commands on MDM 1:

```
scli --login --username admin --password admin
```

```
scli --set_password --old_password admin --new_password
Scaleio123
```

The login password is changed from admin to Scaleio123.

3. Log in with the new credentials:

```
scli --login --username admin --password Scaleio123
```

4. Add MDM 2 to the MDM cluster as a Standby MDM, by running the following command on MDM 1:

```
scli --add_standby_mdm --new_mdm_ip <IP of MDM 2> --mdm_role
manager --new_mdm_management_ip <Management IP of MDM 2> --
new_mdm_virtual_ip_interface <INTF> --new_mdm_name <Name to
give MDM 2>
```

Example:

```
scli --add_standby_mdm --new_mdm_ip 192.168.1.153 --mdm_role
manager --new_mdm_management_ip 10.103.110.153 --
new_mdm_virtual_ip_interface eth4 --new_mdm_name mdm153
```

5. Add TB 1 to the MDM cluster as a TieBreaker, by running the following command on MDM 1:

```
scli --add_standby_mdm --new_mdm_ip <IP of TB 1> --mdm_role
tb --new_mdm_name <Name to give TB 1>
```

Example:

```
scli --add_standby_mdm --new_mdm_ip 192.168.1.154 --mdm_role
tb --new_mdm_name tb154
```

6. (Optional) Query the cluster to verify that it's running in single mode, by running the following command on MDM 1:

```
scli --query_cluster
```

Output, similar to the following, is displayed:

```
Cluster: Mode: 1_node
```

7. Switch to 3-node cluster mode, by running the following command on MDM 1:

```
scli --switch_cluster_mode --cluster_mode 3_node --
add_slave_mdm_name <MDM 2 name> --add_tb_name <TB 1 name>
```

Example:

```
scli --switch_cluster_mode --cluster_mode 3_node --
add_slave_mdm_name mdm153 --add_tb_name tb154
```

8. (Optional) Query the cluster to verify that it's running in cluster mode, by running the following command on MDM 1:

```
scli --query_cluster
```

Output, similar to the following, is displayed:

```
# scli --query_cluster
Cluster:
    Mode: 3_node, State: Normal, Active: 3/3, Replicas: 2/2
    Virtual IPs: 192.168.111.152
Master MDM:
    Name: MDM152, ID: 0x7d76db0d2ad840b0
        IPs: 192.168.1.152, Management IPs: 10.103.110.152,
Port: 9011, Virtual IP interfaces: eth4
        Version: 2.0.13000
Slave MDMs:
    Name: MDM153, ID: 0x41788ea96626fe21
        IPs: 192.168.1.153, Management IPs: 10.103.110.153,
Port: 9011, Virtual IP interfaces: eth4
        Status: Normal, Version: 2.0.13000
Tie-Breakers:
    Name: MDM154, ID: 0x3d8263b64fcf2642
        IPs: 192.168.1.154, Port: 9011
        Status: Normal, Version: 2.0.13000
```

After completion, see "Post-installation considerations".

Deploying a 5-node cluster

For this example for a 5-node cluster, we will use the following sample topology:

Table 12 Sample manual installation topology

Role	Management IP	MDM IP	Virtual IP address	Virtual IP NIC	MDM name	Cluster role
MDM 1	10.103.110.152	192.168.1.152	192.168.111.152	eth4	mdm152	Master
MDM 2	10.103.110.153	192.168.1.153	192.168.111.152	eth4	mdm153	Slave 1
MDM 3	10.103.110.154	192.168.1.154	192.168.111.152	eth4	mdm154	Slave 2
TB 1	10.103.110.155	192.168.1.155			tb155	TB 1
TB 2	10.103.110.156	192.168.1.156			tb156	TB 2
SDS and SDC will be installed on all nodes.						

Installing the packages

In this section, be sure to use the installation packages and commands that are appropriate for your operating system environment.

Note

If you intend to use the replication feature via RecoverPoint on RHEL 6.x servers, refer to the *EMC ScaleIO Write Splitter for RecoverPoint Technical Notes* before beginning the installation. Replication support is version-specific; see the ESSM for full details.

To install ScaleIO components, perform the following:

Procedure

1. Copy the following file to all MDM and TB servers (MDM 1, 2, 3 and TB 1 and 2):
   a. `EMC-ScaleIO-mdm-2.0-13000.X.elX.x86_64.rpm`
2. Copy the following files to all servers that will be SDC and/or SDS servers (in our example, all the servers):
   - `EMC-ScaleIO-sds-2.0-13000.X.elX.x86_64.rpm`
   - `EMC-ScaleIO-sdc-2.0-13000.X.elX.x86_64.rpm`
3. Install the MDM and configure the Manager role, by running the following commands on the Master (MDM 1) and Slave MDMs (MDM 2 and MDM 3):

OS	Command

RHEL/ CentOS /OL	`MDM_ROLE_IS_MANAGER=1 rpm -i <mdm_path.rpm>`
	Example
	`MDM_ROLE_IS_MANAGER=1 rpm -i EMC-ScaleIO-mdm-2.0-13000.X.el6.x86_64.rpm`

Ubuntu	`MDM_ROLE_IS_MANAGER=1 dpkg -i <mdm_path.deb>`
CoreOS	`MDM_ROLE_IS_MANAGER=1 /.<mdm_path.bsx>`

Windows	`msiexec /i <mdm_path.msi> MDM_ROLE_IS_MANAGER=1`
	Example
	`msiexec /i EMC-ScaleIO-mdm-2.0-13000.X.msi MDM_ROLE_IS_MANAGER=1`

Note

The default MDM credentials are:

- user name: admin
- password: admin

4. Install the MDM on the TieBreaker MDMs (TB 1 and TB 2):

OS	Command

RHEL/ CentOS /OL	`MDM_ROLE_IS_MANAGER=0 rpm -i <mdm_path.rpm>`
	Example:
	`MDM_ROLE_IS_MANAGER=0 rpm -i EMC-ScaleIO-mdm-2.0-13000.X.el6.x86_64.rpm`

Ubuntu	`MDM_ROLE_IS_MANAGER=0 dpkg -i <mdm_path.deb>`
CoreOS	`MDM_ROLE_IS_MANAGER=0 ./<mdm_path.bsx>`

Windows	`msiexec /i <mdm_path.msi> MDM_ROLE_IS_MANAGER=0`
	Example
	`msiexec /i EMC-ScaleIO-mdm-2.0-13000.X.msi MDM_ROLE_IS_MANAGER=0`

5. Install the SDS on every server that will contribute storage drives to the ScaleIO system:

```
rpm -i <sds_path.rpm>
```

Example

```
rpm -i EMC-ScaleIO-sds-2.0-13000.X.el6.x86_64.rpm
```

Note

Including the IP addresses of as many potential MDMs at this time (up to 8), or of the cluster virtual IP addresses, will make it easier to switch MDMs in the future.

6. Install the SDC on every server that will expose ScaleIO volumes to the application running on it. You can add all the IP addresses of each MDM, or the virtual IP addresses for the MDMs.

Note

Before installing SDC or RFcache on Ubuntu, Oracle Linux (OL), or CoreOS servers, perform the steps described in Deploying on Ubuntu, Oracle Linux, or CoreOS servers on page 90.

Option	Description

RHEL/ CentOS / OL	`MDM_IP=<list of virtual IP or MDM IP addresses> rpm -i <sdc_path.rpm>`
	Example
	`MDM_IP=192.168.1.152,192.168.1.153,192.168.1.154 rpm -i EMC-ScaleIO-sdc-2.0-13000.X.el6.x86_64.rpm`

Ubuntu	`MDM_IP=<list of virtual IP or MDM IP addresses> dpkg -i <sdc_path.deb>`
CoreOS	`MDM_IP=<list of virtual IP or MDM IP addresses> ./ <sdc_path.bsx>`

Windows	`msiexec /i <sdc_path.msi> MDM_IP="<list of virtual IP or MDM IP addresses>"`
	Example
	`msiexec /i EMC-ScaleIO-sdc-2.0-13000.X.msi MDM_IP="192.168.1.152,192.168.1.153,192.168.1.154"`

All of the components are installed. The next step is to form the MDM management cluster.

Creating the MDM cluster

Procedure

1. Create the MDM single-node cluster on MDM 1, by running the following command on MDM 1:

```
scli --create_mdm_cluster --master_mdm_ip <IP of MDM 1> --
master_mdm_management_ip <Management IP of MDM 1> --
master_mdm_name <Name to give MDM 1> [--cluster_virtual_ip
```

```
<IP>]  [--master_mdm_virtual_ip_interface <INTF>] --
accept_license
```

Example:

```
scli --create_mdm_cluster --master_mdm_ip 192.168.1.152 --
master_mdm_management_ip 10.103.110.152 --master_mdm_name
mdm152 --cluster_virtual_ip 192.168.111.152 --
master_mdm_virtual_ip_interface eth4 --accept_license
```

Note

Each MDM name must be unique. Ensure that each server on which MDM is installed has a unique hostname.

2. Log in and change the MDM password, by running the following commands on MDM 1:

```
scli --login --username admin --password admin
```

```
scli --set_password --old_password admin --new_password
Scaleio123
```

The login password is changed from admin to Scaleio123.

3. Log in with the new credentials:

```
scli --login --username admin --password Scaleio123
```

4. Add the Slave MDMs (MDM 2 and MDM 3) to the MDM cluster as Standby MDMs, by running the following commands on MDM 1:

```
scli --add_standby_mdm --new_mdm_ip <IP of MDM X> --mdm_role
manager --new_mdm_management_ip <Management IP of MDM X> --
new_mdm_virtual_ip_interface <INTF> --new_mdm_name <Name to
give MDM X>
```

Examples:

```
scli --add_standby_mdm --new_mdm_ip 192.168.1.153 --mdm_role
manager --new_mdm_management_ip 10.103.110.153 --
new_mdm_virtual_ip_interface eth4 --new_mdm_name mdm153
```

```
scli --add_standby_mdm --new_mdm_ip 192.168.1.154 --mdm_role
manager --new_mdm_management_ip 10.103.110.154 --new_mdm_name
mdm154
```

5. Add TB 1 and TB 2 to the MDM cluster as a TieBreaker, by running the following commands on MDM 1:

```
scli --add_standby_mdm --new_mdm_ip <IP of TB X> --mdm_role
tb --new_mdm_name <Name to give TB X>
```

Examples:

```
scli --add_standby_mdm --new_mdm_ip 192.168.1.155 --mdm_role
tb --new_mdm_name tb155
```

```
scli --add_standby_mdm --new_mdm_ip 192.168.1.156 --mdm_role
tb --new_mdm_name tb156
```

6. (Optional) Query the cluster to verify that it's running in single mode, by running the following command on MDM 1:

```
scli --query_cluster
```

Output, similar to the following, is displayed:

```
Cluster: Mode: 1_node
```

7. Switch to 5-node cluster mode, by running the following command on MDM 1:

```
scli --switch_cluster_mode --cluster_mode 5_node --
add_slave_mdm_name <MDM 2 name>,<MDM 3 name> --add_tb_name
<TB 1 name>,<TB 2 name>
```

Example:

```
scli --switch_cluster_mode --cluster_mode 5_node --
add_slave_mdm_name mdm153,mdm154 --add_tb_name tb155,tb156
```

8. (Optional) Query the cluster to verify that it's running in cluster mode, by running the following command on MDM 1:

```
scli --query_cluster
```

Output, similar to the following, is displayed:

```
# scli --query_cluster
Cluster:
Mode: 5_node, State: Normal, Active: 5/5, Replicas: 3/3
Virtual IPs: 192.168.111.152
Master MDM:
Name: mdm6, ID: 0x1cb79cb46ed13200
IPs: 192.168.247.6, 172.16.247.6, Management IPs: 10.136.247.6,
Port: 9011, Virtual IP interfaces: eth4
Version: 2.0.13000
Slave MDMs:
Name: mdm8, ID: 0x67d759a83ec90d42
```

```
IPs: 172.16.247.8, 192.168.247.8, Management IPs: 10.136.247.8,
Port: 9011, Virtual IP interfaces: eth4
Status: Normal, Version: 2.0.13000
Name: mdm7, ID: 0x4d9cb20c720f7481
IPs: 192.168.247.7, 172.16.247.7, Management IPs: 10.136.247.7,
Port: 9011, Virtual IP interfaces: eth4
Status: Normal, Version: 2.0.13000
Tie-Breakers:
Name: mdm9, ID: 0x64bd85740ae9b2a4
IPs: 172.16.247.9, 192.168.247.9, Port: 9011
Status: Normal, Version: 2.0.13000
Name: mdm10, ID: 0x593656bc4ef59083
IPs: 172.16.247.10, 192.168.247.10, Port: 9011
Status: Normal, Version: 2.0.13000
```

Post-installation considerations

This section describes various post-installation considerations. Proceed to the section that matches your operating system environment:

- "Linux"
- "Windows"

Linux

Each installed component creates subdirectories for its use in the following directory:

`/opt/emc/scaleio/<component>/`

For example: `/opt/emc/scaleio/mdm/`

Note

If you are running Veritas Cluster Server (VCS) software, see ScaleIO support on Linux servers running Veritas Cluster Server (VCS) software on page 83 for post-installation instructions.

Installing RFcache - Linux

To enable RFcache, where the RFcache devices provide acceleration, you need to install it, then enable it.

Every SDS that provides acceleration must have at least one RFcache device on it. For a SP with RFcache enabled, you must select at least one RFcache device for each SDS whose devices are part of that SP.

Before installing RFcache on Ubuntu, Oracle Linux, or CoreOS servers, see the requirements described in Ensuring the kernel version is correct on page 91.

Procedure

1. Install RFcache on every SDS server that is to provide acceleration.

   RFcache is enabled with the xcache package:

   ```
   rpm -i EMC-ScaleIO-xcache-2.0-13000.X.elX.x86_64.rpm
   ```

2. Add the RFcache device, by running the following command on the Master MDM:

```
scli --add_sds_rfcache_device --sds_name <SDS_NAME> --
rfcache_device_path <PATH> --rfcache_device_name <DEVICE_NAME>
```

Example:

```
scli --add_sds_rfcache_device --sds_name sds_192 --
rfcache_device_path /dev/sdh --rfcache_device_name sdh
```

3. Enable RFcache on the SP, by running the following command on the Master MDM:

```
scli --set_rfcache_usage --use_rfcache --
protection_domain_name <PD_NAME> --storage_pool_name <SP_NAME>
```

Example:

```
scli scli --set_rfcache_usage --use_rfcache --
protection_domain_name pd_1 --storage_pool_name sp_1
```

Installing the ScaleIO GUI

This topic describes how to install the ScaleIO GUI. You can install the GUI on a Windows or Linux workstation.

Ensure that the workstation satisfies the minimum requirements, as described in System requirements on page 18.

Procedure

1. Install the GUI:

OS	Command
Windows	`EMC-ScaleIO-gui-2.0-13000.X.msi`
RHEL	`rpm -i EMC-ScaleIO-gui-2.0-13000.X.noarch.rpm`
Debian	Run with administrator privileges: `sudo dpkg -i EMC-ScaleIO-gui-2.0-13000.X.deb`

Results

The GUI is installed.

After you finish

To log in and use the GUI, see the *EMC ScaleIO User Guide*.

Installing the Gateway - Linux

To install the ScaleIO Gateway, follow the instructions described in "Preparing the Installation Manager and the Gateway".

CoreOS servers do not support Gateway installation. Deploy a Gateway on another Windows or Linux server.

The ScaleIO Gateway includes the REST Gateway, the Installation Manager (IM), and the SNMP trap sender and ESRS functionality.

The Gateway relies on the presence of the LIA on every server. To install the LIA, see the following section.

Installing the LIA - Linux

This section describes how to install the LIA.

Procedure

1. Copy the LIA package to every server:

```
EMC-ScaleIO-lia-2.0-13000.X.elX.x86_64.rpm
```

2. Install the LIA component, by running the following command on every server:

```
TOKEN=<password> rpm -i <Full RPM file path>
```

   The password must meet the following criteria:

   a. Between 6 and 31, ASCII-printable characters
   b. No blank spaces
   c. Include at least 3 of the following groups: [a-z], [A-Z], [0-9], special chars (! @#$ ...)

   For example:

```
TOKEN=Scaleio123 rpm -i EMC-ScaleIO-
lia-2.0-13000.X.elX.x86_64.rpm
```

3. Import the system installation ID into the LIA:

   a. Create the following file:

```
/opt/emc/scaleio/lia/cfg/installation_id.txt
```

   b. Query the MDM for the installation ID by running the following command:

```
scli --query_all|grep "Installation ID"
```

c. Copy the hexadecimal number that is displayed into the file created in the previous step.

d. Restart the LIA service:

```
pkill lia
```

The LIA component requires security configuration, as described in the *EMC ScaleIO Security Configuration Guide*.

After installing, you need to provision storage. See "Provisioning storage example".

After you finish

It is highly recommended to run the ScaleIO system analysis to analyze the ScaleIO system immediately after deployment, before provisioning volumes, and before using the system in production. For more information, see System analysis overview on page 212.

After installing, you need to provision storage. See "Provisioning storage example".

Note

For optimal performance in environments with more than 60,000 IOPS, see the *Fine-Tuning ScaleIO Performance Technical Notes*.

Windows

Each installed component creates subdirectories for its use in the following directory:

`c:\Program files\emc\scaleio\<component>\`

For example: `c:\Program files\emc\scaleio\mdm\`

Installing RFcache - Windows

To enable RFcache, where the RFcache devices provide acceleration, you need to install it (the xcache package), then enable it.

When installing xcache, a server restart is necessary.

Every SDS that provides acceleration must have at least one RFcache device on it. For a SP with RFcache enabled, you must select at least one RFcache device for each SDS whose devices are part of that SP.

Procedure

1. Install RFcache on every SDS server that is to provide acceleration.

   RFcache is enabled with the xcache package:

   ```
   EMC-ScaleIO-xcache-2.0-13000.X.msi
   ```

   Note

   After installation completes, you must restart the server.

   ---

2. Add the RFcache device, by running the following command on the Master MDM:

```
scli --add_sds_rfcache_device --sds_name <SDS_NAME> --
rfcache_device_path <PATH>
```

Example

```
scli --add_sds_rfcache_device --sds_name sds_192 --
rfcache_device_path j
```

Note

The RFcache device path is represented by a drive letter.

3. Enable RFcache on the SP, by running the following command on the Master MDM:

```
scli --set_rfcache_usage --use_rfcache --
protection_domain_name <PD_NAME> --storage_pool_name <SP_NAME>
```

Example

```
scli scli --set_rfcache_usage --use_rfcache --
protection_domain_name pd_1 --storage_pool_name sp_1
```

Installing the ScaleIO GUI

This topic describes how to install the ScaleIO GUI. You can install the GUI on a Windows or Linux workstation.

Ensure that the workstation satisfies the minimum requirements, as described in System requirements on page 18.

Procedure

1. Install the GUI:

OS	Command
Windows	`EMC-ScaleIO-gui-2.0-13000.X.msi`
RHEL	`rpm -i EMC-ScaleIO-gui-2.0-13000.X.noarch.rpm`
Debian	Run with administrator privileges: `sudo dpkg -i EMC-ScaleIO-gui-2.0-13000.X.deb`

Results

The GUI is installed.

After you finish

To log in and use the GUI, see the *EMC ScaleIO User Guide*.

Installing the Gateway - Windows

To install the ScaleIO Gateway, follow the instructions described in "Preparing the Installation Manager and the Gateway".

The ScaleIO Gateway includes the REST Gateway, the Installation Manager (IM), and the SNMP trap sender functionality.

Installing the LIA - Windows

Procedure

1. Install the LIA component on every node (optional), by running the following command:

   ```
   msiexec /i EMC-ScaleIO-lia-2.0-13000.X.msi TOKEN=<password>
   ```

   The password must meet the following criteria:

   a. Between 6 and 31, ASCII-printable characters

   b. No blank spaces

   c. Include at least 3 of the following groups: [a-z], [A-Z], [0-9], special chars (!@#$ ...)

2. Import the system installation ID into the LIA:

   a. Create the following file:

   `C:\Program Files\EMC\scaleio\lia\cfg\installation_id.txt`

   b. Log in, then query the MDM for the installation ID by running the following command:

   ```
   scli --query_all
   ```

   As part of the output, the installation ID is displayed.

   c. Copy the installation ID into the new file, and add a blank line under it.

   d. Restart the LIA service:

   ```
   EMC scaleio LIA service
   ```

   The LIA component requires security configuration, as described in the *EMC ScaleIO Security Configuration Guide*.

Note

For optimal performance in environments with more than 60,000 IOPS, see the *Fine-Tuning ScaleIO Performance Technical Notes*.

Provisioning storage example

This section shows a simple illustration of how to use the CLI to add capacity and create and map volumes. You can also use the GUI, REST API, or the VMware plug-in.

The IP addresses in the example are taken from the 3-node cluster manual installation example, described in "Deploying a 3-node cluster", but the concepts apply to 5-node clusters, too.

Adding capacity

This section describes how to add capacity. For our example, we will use the following names:

Table 13 Provisioning storage example

Item to add	Name
Protection Domain	pd1
Storage Pool	sp1
Devices	/dev/sdb, /dev/sdc, and /dev/sdd

All scli commands are run from the Master MDM (MDM 1), unless otherwise noted.

Procedure

1. Define a Protection Domain, `pd1`, by running the following command, from the Master MDM (MDM 1):

```
scli --add_protection_domain --protection_domain_name pd1
```

Note

To run scli commands from another server, add `--mdm_ip <IP MDM 1>,<IP MDM 2>` to each command line.

2. Add a Storage Pool, `sp1`, by running the following command:

```
scli --add_storage_pool --protection_domain_name pd1 --
storage_pool_name sp1
```

3. Add an SDS on MDM 1, and its devices to `pd1`, and name it **`sds152`**, by running the following command:

```
scli --add_sds --sds_ip <sds IP>
--device_path <sds devices>
--protection_domain_name <protection domain name>
```

```
--storage_pool_name <storage pool name>
--sds_name <sds name>
```

Example:

```
scli --add_sds --sds_ip 192.168.1.152 --device_path /dev/
sdb,/dev/sdc,/dev/sdd --storage_pool_name sp1 --
protection_domain_name pd1 --sds_name sds152
```

Note

ScaleIO might not perform optimally if there are large differences between the sizes of the devices in the Storage Pool, for example, if one device is as big as the rest of the devices. If in doubt, contact ScaleIO support.

To add devices on a Windows-based SDS, you can use either a drive letter or a file for the SDS. For more information, see "Adding devices to SDS nodes on Windows servers".

---

4. Add an SDS on MDM 2, and its devices to `pd1`, and name it **`sds153`**, by running the following command:

```
scli --add_sds --sds_ip 192.168.1.153 --device_path /dev/
sdb,/dev/sdc,/dev/sdd --storage_pool_name sp1 --
protection_domain_name pd1 --sds_name sds153
```

5. Add an SDS on TB 1, and its devices to `pd1`, add meaningful names to the devices (optional), and name the SDS **`sds154`**, by running the following command:

```
scli --add_sds --sds_ip 192.168.1.154 --device_path /dev/
sdb,/dev/sdc,/dev/sdd --device_name
node_12_dev1,node_12_dev5,node_12_dev7 --storage_pool_name
sp1 --protection_domain_name pd1 --sds_name sds154
```

Creating and mapping volumes

This section describes how to create and map volumes.

Procedure

1. Define a 2500 GB volume, and name it **`vol1`**, by running the following command:

```
scli --add_volume --volume_name <volume name>
--protection_domain_name <protection domain name>
--storage_pool_name <storage pool name> --size_gb <size>
```

Example:

```
scli --add_volume --volume_name vol1 --protection_domain_name
pd1 --storage_pool_name sp1 --size_gb 2500
```

2. Map vol1 to the SDC on MDM 1, by running the following command:

```
scli --map_volume_to_sdc --volume_name vol1 --sdc_ip <IP MDM
1>
```

Example:

```
scli --map_volume_to_sdc --volume_name vol1 --sdc_ip
192.168.1.152
```

Configuring spare capacity

This section describes how to configure spare capacity for your system. This step is highly recommended in any system with less than 10 SDSs or Fault Sets.

Modify the spare capacity, by running the following command on MDM 1:

```
scli --modify_spare_policy --protection_domain_name pd1 --
storage_pool_name sp1 --spare_percentage 34 [--i_am_sure]
```

The number 34 represents the percentage of total capacity set aside to ensure data integrity during server failures. The percentage is derived by 1/(number of SDS), which yields the recommended percentage for less than 10 balanced servers.
For more information, see “Modifying spare policy” in the EMC ScaleIO User Guide.

Deploying ScaleIO on ESX

This section describes topics related to installing ScaleIO in VMware environments.

It is highly recommended to deploy using the VMware plug-in deployment wizard. This appendix describes tasks that are necessary in cases where this is not possible or desirable, such as for scripting. Topics include:

- Deploying the ScaleIO virtual machine (SVM)
- Configuring the UUID on virtual machines
- Installing the SDC directly on an ESX host
- Optimizing the guest operating system performance

After manual deployment on ESX, you must perform manual memory allocation on the SVM, as described in “SVM manual memory allocation”.

Before starting, obtain the IP addresses for all of the nodes to be installed, and an additional IP address for the MDM cluster, as described in the following table:

Number of networks	IP addresses required
1 network (management network)	• 1 for vmnic • 1 for SVM
2 networks	• 2 for vmnic (1 for mgmt, 1 for data) • 2 for SVM (1 for mgmt, 1 for data)
3 networks	• 3 for vmnic (1 for mgmt, 2 for data)

Number of networks	IP addresses required
	• 3 for SVM (1 for mgmt, 2 for data)

Deploying the ScaleIO virtual machine (SVM)

This topic describes how to deploy the SVM and prepare it for installation of ScaleIO. This is necessary when the SDC installed on ESX, but all other components are installed on ScaleIO virtual machines (SVM).

Note

Before deploying, ensure that networks for each ESX to be part of the system have a properly defined Port Group and VMkernel.

To deploy the SVM on the ESX, perform the following:

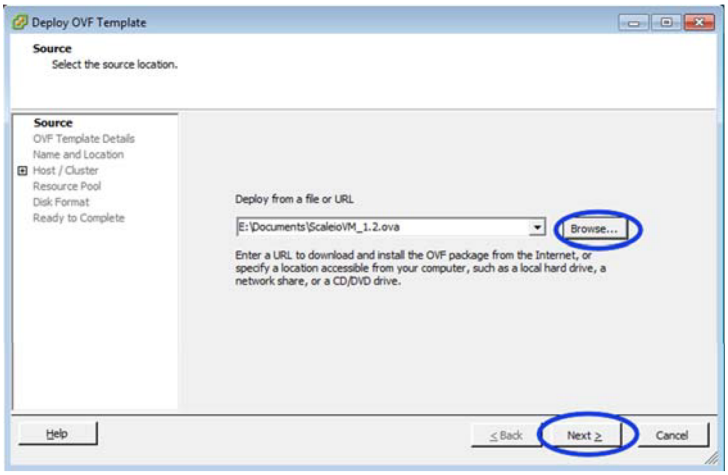
Procedure

1. Download the installation software from the EMC Support site (https://support.emc.com).

   If necessary, extract the files to an installation folder.

2. Download the following document: http://www.vmware.com/files/pdf/techpaper/vmware-multipathing-configuration-software-iSCSI-port-binding.pdf.
3. Use the vSphere client to connect to a vCenter.
4. Select **File** > **Deploy OVF Template**.

   The Deploy OVF Template wizard appears.



5. Enter the full path to the OVA, and click **Next**.
6. Assign a unique name to the VM, accept all remaining default values, and click **Finish**.
7. Clone the SVM to each ESX host, making sure to assign a unique name to each VM.
8. On each ESX physical host, configure the network and NTP (and DNS, if necessary).

a. Using the console, start the SVM, and login.

The default user name is `root`, and the default password is `admin`.

b. Configure the network:

- IP address
- Gateway address
- DNS
- NTP server

Note

When using the OVA, the installation packages can be found in /root/install/.

9. On each SVM in the system, add the local disk devices as RDM devices.

   Perform this task by changing the SVM settings.

10. Configure each ESX host to enable *Start and Stop Virtual Machines with the system*, and set the SVM to do *Automatic Startup*.

11. On each SVM in the system, install the relevant ScaleIO components, as described in "Manual deployment on physical servers".

    The packages are located under `/root/install`

Note

After manual deployment on ESX, you must perform manual memory allocation on the SVM, as described in "SVM manual memory allocation".

Configuring the UUID on virtual machines

To configure the virtual machine, you must enable the UUID attribute on every virtual machine, by performing the following:

Procedure

1. Start the vSphere Client, and log in to a vCenter Server.
2. Select **Virtual Machines and Templates**.
3. From the **Virtual Machines** tab, right-click the virtual machine, and choose **Power** > **Power Off**.

   The virtual machine powers off.

4. Right-click the virtual machine, and click **Edit Settings**.
5. From the **Options** tab, select **Advanced** > **General** in the settings column.
6. Click **Configuration Parameters**.

   The **Configuration Parameters** window appears.

7. Click **Add Row**, and enter the following:

   a. In the **Name** column, enter `disk.EnableUUID`.

   b. In the **Value** column, enter `TRUE`.

8. Click **OK**.
9. Power on the virtual machine.
10. Repeat this procedure for every SVM in the ScaleIO system.

Installing the SDC directly on an ESX host

To install the SDC component directly on an ESXi host, perform the following:

Procedure

1. Set the acceptance level of your host to `PartnerSupported`, by entering:

```
esxcli software acceptance set --level=PartnerSupported
```

2. Install the SDC VIB, by entering the following:

   `esxcli software vib install -d <full_path_to_VIB>>`
   where *<full_path_to_VIB>* is the full path to the `sdc-2.0-13000.X-esx5.5.zip` file.

   Note

   To install without changing the acceptance level, use this command instead:
   `esxcli software vib install -f -d <VIB location>`

3. Reboot the ESX host.

   The SDC will not automatically boot up at this point.
   Update the SDC GUID and MDM IP address parameters:

   a. GUID, for example: 12345678-90AB-CDEF-1234-567890ABCDEF

   Note

   Including the IP addresses of as many potential manager MDMs as possible at this time will make it easier to switch MDMs in the future. You can add a total of eight IP addresses.

   b. MDM IP addresses.

   You can define multiple MDM clusters, each with multiple IP addresses.

   - Separate the IP addresses of the same MDM cluster with a “,” symbol.
   - Separate multiple MDM clusters with the “+” symbol.

     In the following example, there are two MDM clusters. The first has two IP addresses and the second has only one:

     `10.20.30.40,50.60.70.80+11.22.33.44`

4. To update the SDC parameters, enter the following command:

```
esxcli system module parameters set -m scini -p
"IoctlIniGuidStr=<GUID> IoctlMdmIPStr=<MDM_IPs>"
```

where:

a. `GUID` is the GUID (maximum of 1024 digits)

b. `MDM_IPs` is the IP addresses, as described earlier

5. Back up the parameters, by entering the following:

```
/bin/auto-backup.sh
```

6. Load the SDC module, by entering the following:

```
esxcli system module load -m scini
```

Optimizing the guest operating system performance

The configuration of the guest operating system affects the performance of the system. The ScaleIO VM (SVM) is optimized. If the storage VM that is hosting the SDS and MDM roles is not supplied by ScaleIO, refer to the following guideline and procedures.

Modify the device scheduler (Linux guests only)

The following command modifies the I/O scheduler of the devices. It should be run for each device that exposes a ScaleIO volume inside the guest virtual machine.

Command

```
echo noop > /sys/block/<device_name>/queue/scheduler
```

Example

```
echo noop > /sys/block/sds/queue/scheduler
```

Modifying the SCSI controller type of the guest virtual machine

To configure the SCSI controller type to use a Paravirtual SCSI (PVSCSI) adapter, perform the following steps for each guest using ScaleIO storage.

Note

Before performing the following steps, consult VMware documentation to verify that your guest operating system supports Paravirtual SCSI adapters.

Procedure

1. Power off the guest virtual machine.
2. Open the vSphere client, and right click on the guest.
3. Select **Edit Settings**.
4. Select **SCSI Controller 0**.
5. Click **Change Type**.
6. Select VMware **Paravirtual**.
7. Click **OK**.

Manually removing the VIB file from the ESX host

Procedure

1. In an SSH session, check the version of VMware ESXi installed on the host by typing this command:

   ```
   esxcli system version get
   ```

   Example of output:

   ```
   Product: VMware ESXi
   Version: 5.5.0
   Build: Releasebuild-2068190
   Update: 2
   ```

   In the above example, the version is `5.5.0`.

2. Type the command that corresponds to the version installed on the host. For example:

   a. For ESXi 5.5:

   ```
   esxcli software vib remove --vibname scaleio-sdc-esx5.5
   ```

   b. For ESXi 6.0:

   ```
   esxcli software vib remove --vibname scaleio-sdc-esx6.0
   ```

   Example of output:

   ```
   Removal Result
   Message: The update completed successfully, but the system
   needs to be rebooted for the changes to be effective.
   Reboot Required: true
   VIBs Installed:
   VIBs Removed: VMware_bootbank_scaleio-sdc-esx5.5_1.32-201.0
   VIBs Skipped:
   ```

3. Open the `esx.conf` file for editing by typing this command:

   ```
   vi /etc/vmware/esx.conf
   ```

   a. Search for the line containing the scini text (use key sequence /scini).

   b. Delete the line using either the <X> key or <D> to delete to full line.

   c. Save and exit using the key sequence: <ESC>+<:wq!> and <Enter>.

CHAPTER 7

Advanced Topics

This chapter describes advanced ScaleIO installation and upgrade topics. Topics include:

- Advanced Gateway topics....176
- Advanced Installation Manager topics....182
- Advanced VMware plug-in topics....184
- Deploy ScaleIO in a 2-layer environment....188
- ESXi deployment prerequisites....190

Advanced Gateway topics

This section describes the following advanced issues for Gateway installation:

- "Enabling and disabling Gateway components"
- "Using a custom Java configuration for the Gateway"
- "Installing the Gateway without assigning an admin password"
- "Certificate management for ScaleIO Gateway"
- "OpenStack interoperation with the ScaleIO Gateway"
- "Generating a self-signed certificate using the keytool utility"

Enabling and disabling Gateway components

You can use the Features Enabler to enable and disable each Gateway feature. The Features Enabler is activated by setting the values in the Features Enabler section of the `gatewayUser.properties` file.

- Enable SNMP (default: false)

  To enable, set `features.enable_snmp= true`

- Enable Gateway (default: true)

  To disable, set `features.enable_gateway=false`

  You can disable the use of the default port, 443, by setting both this property and the `features.enable_IM` to `false`.

- Enable Installation Manager (default: true)

  To disable, set `features.enable_IM=false`

  You can disable the use of the default port, 443, by setting both this property and the `features.enable_gateway` to `false`.

- Enable ESRS (default: false)

  To enable, set `features.enable_esrs=false true`

- Disable local user login (default: false)

  To restrict login to LDAP users only, set

  ```
  gateway-admin.disable.local.login=true
  ```

- Enable adding SDS devices that have a signature from a previous ScaleIO system (default: false). This can happen in the following cases:
  - Add SDS devices that were used in a previous ScaleIO system.
  - Extending a system with an SDS whose devices were used in a previous system.
  - Extending a system by adding SDSs to a node that already has an SDS installed.

To enable this feature, set `add.sds.with.force.on.next.run=true`

This setting is valid for one deployment only. After being marked completed, it reverts to false.

To configure the properties, perform the following:

Procedure

1. Use a text editor to open the `gatewayUser.properties` file, located in the following directory on the Installation Manager/Gateway server:

Gateway installed on	Location of gatewayUser.properties file
Windows	C:\Program Files\EMC\ScaleIO\Gateway\webapps\ROOT\WEB-INF\classes\
Linux	/opt/emc/scaleio/gateway/webapps/ROOT/WEB-INF/classes

2. Edit the file with the desired changes.
3. Save and close the file.
4. Restart the scaleio-gateway service:
   a. Windows:
      Restart the EMC ScaleIO Gateway service.
   b. Linux
      Type the command `service scaleio-gateway restart`.

   Configuration is complete.

Using a custom Java configuration for the Gateway

To use a custom Java configuration, use the following command to install the Gateway, in place of the normal command:

```
GATEWAY_ADMIN_PASSWORD=<new_GW_admin_password>
SIO_GW_JAVA=<path_to_java> rpm -U /tmp/EMC-ScaleIO-
gateway-2.0-13000.X.x86_64.rpm
```

where *<new_GW_admin_password>* is a password that you define to access the Installation Manager administration commands and web interface, and *<path_to_java>* is the path to the desired Java files.

Installing the Gateway without assigning an admin password

You can install the ScaleIO Gateway without adding the `GATEWAY_ADMIN_PASSWORD` variable.

To add this variable later, perform one of the following:

Procedure

1. Use a REST URI command, as described in the *EMC ScaleIO User Guide*.
2. Edit the user properties file, and then restart the scaleio-gateway service, as described in the *EMC ScaleIO User Guide*.

Certificate management for ScaleIO Gateway

This section explains how to replace the ScaleIO gateway's self-signed security certificate with your organization's "trusted" certificate, and how to create a new

"trusted" certificate. The ScaleIO gateway automatically creates its own self-signed security certificate when it is installed or upgraded. If your organization has no special security certificate requirements, you can keep working with the default certificate.

This section includes the following topics:

- "Replacing the default self-signed security certificate with your own trusted certificate"
- "Replacing the default self-signed security certificate with your own self-signed certificate"
- "Upgrading the Gateway when a non-default certificate is used"

For a description of how to fix a keytool error, see "Fixing keytool errors".

Replacing the default self-signed security certificate with your own trusted certificate

You can create your own trusted certificate, and then replace the default certificate with the one that you created.

Procedure

1. Find the location of `keytool` on your server, and open it.

   It is a part of the Java (JRE or JDK) installation on your server, in the `bin` directory. For example:

   a. `C:\Program Files\Java\jdk1.7.0_25\bin\keytool.exe`

   b. `/usr/bin/keytool`

2. Generate your RSA private key, by typing this command:

   ```
   keytool -genkey -alias <YOUR_ALIAS> -keyalg RSA -keystore
   <PATH_TO_NEW_KEYSTORE_FILE>
   ```

   a. If you want to define a password, add the following parameters to the command. Use the same password for both parameters.

   ```
   -storepass <KEYSTORE_PASSWORD> -keypass <KEYSTORE_PASSWORD>
   ```

   Note

   Specify a directory outside the ScaleIO Gateway installation directory for the newly created keystore file. This will prevent it from being overwritten when the ScaleIO Gateway is upgraded or reinstalled.

3. If you already have a Certificate Signing Request (CSR), skip this step.

   If you need a CSR, generate one by typing the following command. (If you did not define a keystore password in the previous step, omit the password flags.)

   ```
   keytool -certreq -keyalg RSA -alias <YOUR_ALIAS> -file
   certreq.txt -keystore <PATH_TO_NEW_KEYSTORE_FILE> -storepass
   <KEYSTORE_PASSWORD> -keypass <KEYSTORE_PASSWORD>
   ```

4. If you already have an SSL certificate, skip this step.

   If you need an SSL certificate, use your CSR to obtain a new certificate from a third-party trusted SSL certificate provider. Save the certificate file on your server, outside the ScaleIO Gateway installation directory.

5. Import the Trusted Root, by typing this command. (If you did not define a keystore password, omit the password flags.)

```
keytool -import -alias root -keystore
<PATH_TO_NEW_KEYSTORE_FILE> -trustcacerts -file <LOCATION
OF_YOUR_root.cer_FILE> -storepass <KEYSTORE_PASSWORD> -
keypass <KEYSTORE_PASSWORD>
```

Note

.
The certificate must be in x.509 format.

If a message appears saying that the root is already in the system-wide store, import it anyway.

6. Import the intermediate certificates, by typing the command. (If you did not define a keystore password, omit the password flags.)

```
keytool -import -alias intermediateCA -keystore
<PATH_TO_NEW_KEYSTORE_FILE> -trustcacerts -file
<LOCATION_OF_YOUR_intermediate.cer_FILE> -storepass <keystore
password> -keypass <keystore password>
```

You must provide a unique alias name for every intermediate certificate that you upload with this step.

7. Install the SSL Certificate under the same alias that the CSR was created from (<YOUR_ALIAS> in previous steps), by typing the command (if you did not define a keystore password, omit the password flags):

```
keytool -import -alias <YOUR_ALIAS> -keystore
<PATH_TO_NEW_KEYSTORE_FILE> -trustcacerts -file
<LOCATION_OF_SSL_CERTIFICATE> -storepass <keystore password> -
keypass <keystore password>
```

8. Edit the following items in the file *<ScaleIO_GATEWAY_INSTALLATION DIRECTORY>*`\conf\catalina.properties`:
   a. `keystore.file=`*<PATH_TO_NEW_KEYSTORE_FILE>*
   b. keystore.password=*<PASSWORD_DEFINED_DURING_KEYSTORE_CREATION>*

      If you did not define a password, the default password is `changeit`.
9. Restart the scaleio-gateway service:
   - Windows
     From the Windows **Services** management window, restart the EMC ScaleIO Gateway service.
   - Linux
     Type the following command:

```
service scaleio-gateway restart
```

Replacement of the security certificate is complete.

Replacing the default self-signed security certificate with your own self-signed certificate

Procedure

1. Find the location of `keytool` on your server, and open it.

   It is usually a part of the Java (JRE or JDK) installation on your server, in the `path` or `bin` directory. For example:

   a. `C:\Program Files\Java\jdk1.7.0_25\bin\keytool.exe`

   b. `/usr/bin/keytool`

2. Type the following command.

   If you want to use the password generated by default (`changeit`), omit the `-keypass` and `-keystore` parameters from the command.

```
keytool -genkey -alias <YOUR_ALIAS> -keyalg RSA -validity 360
-keysize 2048 -storepass <KEYSTORE_PASSWORD> -keypass
<KEYSTORE_PASSWORD> -keystore <PATH_TO
_THE_CREATED_KEYSTORE_FILE>
```

Note

Store your keystore file in a directory outside the ScaleIO Gateway installation directory. This will prevent it from being overwritten when the ScaleIO Gateway is upgraded or reinstalled.

---

3. Edit the following items in the file *<ScaleIO_GATEWAY_INSTALLATION DIRECTORY>*`\conf\catalina.properties`:

   a. `keystore.file=`*<PATH_TO_NEW_KEYSTORE_FILE>*

   b. `keystore.password=`*<PASSWORD_DEFINED_DURING_KEYSTORE_CREATION>*

   If you did not define a password, the default password is `changeit`.

4. Restart the scaleio-gateway service:

   - Windows
     From the Windows **Services** management window, restart the EMC ScaleIO Gateway service.
   - Linux
     Type the following command:

```
service scaleio-gateway restart
```

Replacement of the security certificate is complete.

Upgrading the Gateway when a non-default certificate is used

If a non-default security certificate is used on the ScaleIO Gateway in the Windows and Linux environment (for example, if the certificate is signed by the user organization CA), you must save a copy of the certificate (`*.keystore` file) and the `catalina.properties` file before you upgrade the gateway. After the upgrade is complete, you must copy these files back to their original location.

The default file locations, per operating system, are:

Linux:

`/opt/emc/scaleio/gateway/conf/catalina.properties`

`/opt/emc/scaleio/gateway/conf/certificates/.keystore`

Windows (64 bit):

`C:\Program Files\EMC\ScaleIO\Gateway\conf\catalina.properties`

`C:\Program Files\EMC\ScaleIO\Gateway\conf\certificates\.keystore`

OpenStack interoperation with the ScaleIO Gateway

The OpenStack ScaleIO Cinder driver communicates with the ScaleIO Gateway through HTTPS (in other words, over SSL). By default, the driver ignores the gateway SSL certificate verification. However, the ScaleIO Cinder driver can be configured to verify the certificate.

Note

You can generate a self-signed certificate (.PEM file), using the keytool utility. For more information, see "Generating a self-signed certificate using the keytool utility".

To enable certificate verification, add the following parameters to the file `/etc/cinder/cinder_scaleio.config` on the Cinder node:

`verify_server_certificate=true`

`server_certificate_path=<PATH_TO_PEM_FILE>`

Generating a self-signed certificate using the keytool utility

This section describes how to generate self-signed certificates, using the keytool utility. The certificates can by used by the OpenStack ScaleIO driver, to communicate with the ScaleIO Gateway. For more information about configuring the driver, see "OpenStack interoperation with the ScaleIO Gateway".

To generate a self-signed certificate using the keytool utility, perform the following steps:

Procedure

1. Create a keystore file (.JKS) by typing the command:

```
keytool -genkeypair -keysize 1024 -alias herong_key -keypass
keypass -keystore herong.jks -storepass jkspass
```

2. Export the keystore file to a .PEM file, by typing the command:

```
keytool -exportcert -alias herong_key -keypass keypass -
keystore herong.jks -storepass jkspass -rfc -file
keytool_crt.pem
```

The certificate is stored in the file `<keytool_crt.pem>`. During configuration of the Cinder driver, the path to this .PEM file is required.

Advanced Installation Manager topics

This section describes the following advanced issues for Installation Manager installations:

- "Adding devices to SDS nodes on Windows servers"
- "Installing without validating Linux devices"
- "Using SSH authentication on the ScaleIO Gateway"

Adding devices to SDS nodes on Windows servers

You can add devices to SDS nodes on Windows server by the following methods:

- Using a drive letter
- Using a file for the SDS with the `prepare_disk.exe` utility. This utility creates a mount point, formats it, and creates a file on it. To add the SDS device, use the full path of this file for the device path, as described in the following steps.

Using a physical disk or using a mount point on a raw partition are both not supported.

Note

As an alternative to manually running the `prepare_disk.exe` utility, you can use the CSV file during installation or extending of ScaleIO. To do so, in the **SDS Device List** column, add both drive letters and PhysicalDrive*X* (where *X* is the disk number from the Windows Disk Manager). You can also use this method with the CSV in situations where you need to add more devices than the amount of available drive letters for raw devices (d-z, excluding previously-allocated drives).

The following formats are acceptable in the **SDS Device List** column of the CSV file and are not case-sensitive:

- //./physicaldrive*X*
- \\.\physicaldrive*X*
- //?/physicaldrive*X*
- \\?\physicaldrive*X*
- .physicaldrive*X*
- ?physicaldrive*X*
- physicaldrive*X*

Procedure

1. On the SDS, open a command line.
2. Verify that the file `prepare_disk.exe` exists on the SDS (folder location `C:\Program Files\EMC\scaleio\sds\bin`).
3. Run the command `prepare_disk.exe \\?\PhysicalDrive`*X*

   where *X* is the disk number from the Windows Disk Manager.
   Example:

```
C:\Program Files\EMC\scaleio> "C:\Program Files\EMC\scaleio
\sds\bin\prepare_disk.exe" \\?\PhysicalDrive8
```

Output similar to the following is displayed:

```
mosShmClient_Thrd:00481: Successfully connected to shared
memory server.
mosIO_close:00159: Closing FD(576)
```

4. Verify that the disk drive exists in the location `SYSTEMDRIVE\scaleio_devices`.

   Example:

   ```
   C:\scaleio_devices\PhysicalDrive8\PhysicalDrive8.bin
   ```

5. Add the new SDS device (one device per `scli --add_sds_device` command) using SCLI, or using the GUI:

   Example:

   ```
   scli --add_sds_device --sds_ip 10.76.61.6 --device_path C:
   \scaleio_devices\PhysicalDrive8\PhysicalDrive8.bin
   ```

   ```
   C:\Program Files\EMC\scaleio\mdm\bin>scli --add_sds_device --
   sds_ip 10.76.61.6 --device_path C:\scaleio_devices
   \PhysicalDrive8\PhysicalDrive8.bin
   ```

   Output similar to the following is displayed:

   ```
   Successfully added device C:\scaleio_devices
   \PhysicalDrive8\PhysicalDrive8.bin to SDS. New device ID:
   5b9f159300010002
   ```

Installing without validating Linux devices

During the Installation Manager validate stage, errors in the entering of Linux device names (such as `/dev/sdbb` instead of `/dev/sdb`) will prevent the continuation of the installation. This can be a problem if you want to add valid devices that do not meet the `/proc/partitions` standard (such as aliases and logical names).

To prevent the installation from getting stuck, you can select **Skip Linux devices validation** from the **Advanced settings** in the IM, as described in step 4.

When this option has been selected, the validate stage will complete, but the affected node will show a completed (warning) status. You can click **Details** to see the error, and you can choose to proceed with the installation.

Using SSH authentication on the ScaleIO Gateway

A manually generated public-private key pair can be used to perform SSH key authentication, instead of passwords, between the ScaleIO Gateway and ScaleIO system servers. The public key is placed on all servers that will allow access to the ScaleIO Gateway. The ScaleIO Gateway is the owner of the matching private key.

SioGWTool is used to upload the private key to the ScaleIO Gateway. SioGWTool is located in:

- Linux: `/opt/emc/scaleio/gateway/bin/SioGWTool.sh`

- Windows: `C:\Program Files\EMC\ScaleIO\Gateway\bin\SioGWTool.bat`

Note

In a new installation, if SSH key authentication will be used, there is no need to add node passwords to the CSV file used for system deployment.

Procedure

1. Create the private and public key pair:
   a. Type the following command in the command line: `cd ~/.ssh/`
   b. Type the following command: `ssh-keygen -t rsa`

      This command generates a private and public key. When performing this command, you can add a passphrase, or generate the key without a passphrase. Two SSH keys are produced: `id_rsa` and `id_rsa.pub`.
2. On each ScaleIO server, type this command to store the public key in the `authorized_keys` file: `cat ~/.ssh/id_rsa.pub | cat >> ~/.ssh/authorized_keys`
3. Use SioGWTool in command line to upload the private key to the gateway, by typing the following command:

```
SioGWTool --update_ssh_key --ssh_key <PATH_TO_SSH_KEY_FILES>
[--key_passphrase <KEY_PASSPHRASE>]
```

   The key passphrase is optional. Add this if you generated it when you created the key pair. During deployment, the gateway will use the private key for authentication.

Advanced VMware plug-in topics

This section describes the following advanced issues for plug-in installation:

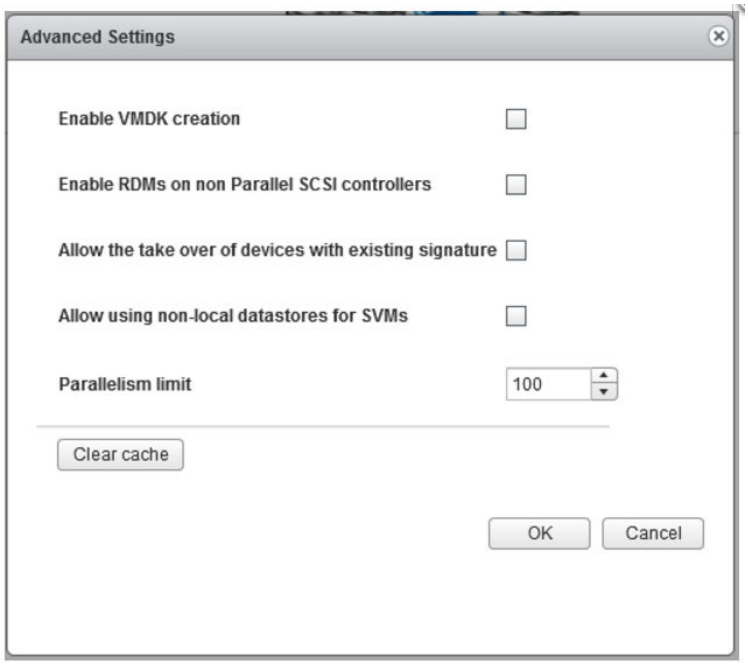
- "Advanced settings options"
- "Manual registration of the ScaleIO plug-in"
- "Troubleshooting plug-in registration issues"

Advanced settings options

This topic describes the advanced settings options that you can use to configure the VMware plug-in deployment settings.

To access these settings, click from the **EMC ScaleIO** screen, click **Advanced settings**.

The **Advanced Settings** dialog appears:



Enable the features you want before starting the ScaleIO deployment, or, if you are in the midst of a deployment, you can click **Cancel**, then make the changes. When you restart the deployment, you can pick up from where you left off.

- **Enable VMDK creation**

  Enables the addition of devices as VMDK, as opposed to RDM.

- **Enable RDMs on non Parallel SCSI controllers**

  Enables non-SCSI controEnable RDMs on non Parallel SCSI controllersller devices to be added as RDM. Note: Do not enable this option if the device does not support SCSI Inquiry Vital Data Product (VPD) page code 0x83.

- **Allow the take over of devices with existing signature**

  Enables the reuse of devices that were part of a previous ScaleIO system, and were not removed from that system properly. This will revert to not selected after use, and must be reactivated.

- **Allow using non-local datastores for ScaleIO gateway**

  Enable deployment of the Gateway SVM on non-local datastores (such as remote storage), in case of insufficient space on the local DS, or if no DS can be found.

- **Parallelism limit**

  Enables the increase of the parallelism limit (default: 100), thus speeding up the deployment, which can be useful in deployment of a very large ScaleIO system (several hundred nodes). This is dependent on the processing power of the vCenter.

Registering the ScaleIO plug-in manually

This topic describes an advanced way to use PowerCLI to register the ScaleIO plug-in on a vCenter.

The plug-in is provided as a ZIP file that can be downloaded by the vSphere web client servers in your environment. The ZIP file can be downloaded directly from the EMC Online Support site, or, if the web servers may not have internet access, from a file server.

Before you begin, ensure that there is communication between the vSphere web client server and the web server storing the plug-in.

Procedure

1. You can upload the ZIP file to an HTTP or an HTTPS server.

   If you are uploading the ZIP file to an HTTP server, perform the following:

   a. On the computer where the vSphere Web client is installed, locate the `webclient.properties` file.

vCenter	Operating system	Path to file
5.x	Windows 2003	%ALLUSERSPROFILE%\Application Data\VMware\vSphere Web Client
	Windows 2008/2012	%ALLUSERSPROFILE%\VMware\vSphere Web Client
	Linux	/var/lib/vmware/vsphere-client
6.x	Windows 2008/2012	C:\ProgramData\VMware\vCenterServer\cfg\vsphere-client
	Linux	/etc/vmware/vsphere-client/

   b. Add the following line to the file:

   `allowHttp=true`

   c. Restart the VMware vSphere Web Client service.

2. Using PowerCLI for VMware, and set to **Run as administrator**, run `Set-ExecutionPolicy AllSigned` (SCI-638).
3. Close PowerCLI, and reopen it, set as **Run as administrator**.
4. Extract `EMC-ScaleIO-vSphere-plugin-installer-2.0-13000.X.zip`
5. Upload `EMC-ScaleIO-vSphere-web-plugin-2.0-13000.X.zip` to the HTTP/HTTPS server.
6. From PowerCLI, run the following script: `ScaleIOPluginSetup-2.0-130000.X.ps1`

   a. Enter the vCenter name or IP address, user name, and password.

   b. For **Select Mode**, choose option `1, Register SIO plugin`.

   c. For **Select Registration Mode**, choose `Advanced`.

   d. Enter the full URL of the plug-in ZIP file.

   If the ZIP file is located on a ScaleIO Gateway, enter the following URL: https:*//<SCALEIO_GATEWAY_IP_ADDRESS>*.
   If you are manually placing the zip file on the server, place it in the `\webapps\root\resources` folder.

   e. If necessary, accept the thumbprint.

7. Log out, and log back in to the vSphere web client.

   Example for HTTP server:

```
.\registerScaleIOPlugin.ps1 -vcenter 10.103.109.16
    -userUrl "http://10.76.60.14/sample/ScaleIO-vSphere-web-
plugin-1.30.0.160.zip"
```

Example for HTTPS server:

.\registerScaleIOPlugin.ps1 -vcenter 10.103.109.16

```
-userUrl
"https://10.76.61.139/sample/ScaleIO-vSphere-web-
plugin-1.30.0.160.zip"
-thumbprint CA:66:49:D0:CE:D9:8C:A0:D0:93:E3:83:DE:59:25:5F:
79:E1:53:B6
-adminEmail test.email@emc.com
```

The script registers the plug-in and the following message appears:

```
Registering ScaleIO extension...
```

```
The extension was registered successfully
```

Troubleshooting plug-in registration issues

You can use the following logs to assist in troubleshooting problems that may occur during registration of the VMware plug-in. To find relevant log entries, search for scaleio in the log file.

The vSphere web client (Virgo) logs are located in the following directories:

vCenter	Operating system	Path to file
5.x	Windows	C:\ProgramData\VMware\vSphere Web Client\serviceability\logs
	Linux	/var/log/vmware/vsphere-client/
6.x	Windows	C:\ProgramData\VMWare\vCenterServer\logs\vsphere-client\logs
	Linux	/var/log/vmware/vsphere-client/logs

Other relevant logs:

vCenter	Operating system	Path to file
5.x	Windows	C:\Windows\System32\config \systemprofile\AppData\Roaming \VMware\scaleio
	Linux	/opt/.vmware/scaleio
6.x	Windows	C:\Users\vspherewebclientsvc \AppData\Roaming\VMware\scaleio
	Linux	/etc/vmware/vsphere-client/vc-packages/scaleio

Deploy ScaleIO in a 2-layer environment

Deploy the ScaleIO MDM and SDS components on physical Linux servers, and the SDCs on ESXi servers.

You can deploy ScaleIO in a 2-layer environment using one of the following methods:

- Deploying ScaleIO in a 2-layer environment using the Installation Manager and plug-in on page 188
- Manually deploying ScaleIO in a 2-layer environment on page 189

Deploying ScaleIO in a 2-layer environment using the Installation Manager and plug-in

Deploy the ScaleIO MDM and SDS components on physical Linux servers, and the SDCs on ESXi servers.

The steps in this task are described fully in the *ScaleIO Deployment Guide*.

Procedure

1. Deploy the physical Linux environment:
   a. Install the ScaleIO Gateway, as described in Preparing the Installation Manager and the Gateway on page 63.
   b. Deploy ScaleIO, as described in Installing with the full Installation Manager on page 67.

   Note

   When preparing the CSV, you must enter No in the Is SDC column for all nodes.

   c. Install the ScaleIO GUI, as described in Installing the ScaleIO GUI on page 100.
2. Register the Linux system:
   a. Register the ScaleIO plug-in, as described in Registering the ScaleIO plug-in and uploading the OVA template on page 108.

      Do not upload the OVA template.
   b. Log in to the vSphere client.
   c. From the ScaleIO **Basic tasks** list, click **Register ScaleIO system**.
   d. Follow the steps to register the Linux system.
3. Install the SDC on ESX, as described in Installing the SDC on ESX hosts on page 111.
4. Deploy ScaleIO:
   a. Start the deployment, as described in Deploying ScaleIO on page 113.
   b. In the **Select Installation** screen, select **Add servers to a registered ScaleIO system**.
   c. In the **Add ESX hosts to cluster** screen, select the relevant datacenter.

d. Click **Next** through the following screens until you reach the **Add SDCs** screen.

e. In the **Add SDCs** screen:

   a. Type the ESX root password.

   b. Select **Disable** for the SCSI comparison.

f. Click **Next** through the following screens until you reach the **Configure Upgrade Components** screen.

g. Clear the **Install ScaleIO Gateway** option.

h. In the **Configure Upgrade Components screen**, select the management and data networks.

i. Review the deployment summary and click **Finish**.

j. Enter the vCenter password.

   The deployment process (adding the SDC) begins.

Manually deploying ScaleIO in a 2-layer environment

Manually deploy the ScaleIO MDM and SDS components on physical Linux servers, and the SDCs on ESXi servers.

The steps in this task are described fully in the *ScaleIO Deployment Guide*.

Procedure

1. Manually deploy the physical Linux environment:

   a. Install the MDM, as described in MDM cluster setup on page 148.

      You can set up the MDM cluster in 3-node, 5-node, or single-node mode.

   b. Install and enable RFCache, if using, as described in Installing RFcache - Linux on page 160.

   c. Install the ScaleIO GUI, as described in Installing the ScaleIO GUI on page 100.

   d. Install the Gateway, as described in Installing the Gateway - Linux on page 162.

   e. Install the LIA, as described in Installing the LIA - Linux on page 162.

2. Register the Linux system:

   a. Register the ScaleIO plug-in, as described in Registering the ScaleIO plug-in and uploading the OVA template on page 108.

      Do not upload the OVA template.

   b. Log in to the vSphere client.

   c. From the ScaleIO **Basic tasks** list, select **Register ScaleIO system**.

   d. Follow the steps to register the Linux system.

3. Install the SDC on the ESX, as described in Installing the SDC directly on an ESX host on page 171.

   You will need to reboot the ESX host during the installation.

ESXi deployment prerequisites

Before deploying ScaleIO on ESXi servers in the ScaleIO Ready Node environment, you must prepare the servers.

Before you begin

Ensure that you have:

- IP address of VMware vCenter server
- User name and password for accessing the vCenter with the vSphere Web Client

Before deploying ScaleIO systems, you must configure VMware security features. This procedure describes how to do so with the vSphere web client.

Note

Ensure that the remote connection is established within 1-2 minutes, otherwise, the lockdown mode is enabled on the ESXi servers, preventing you from subsequent remote connectivity. When in lockdown mode, you can connect to the ESXi servers locally.

Procedure

1. Log in to the vSphere web client.

   The **vSphere Web Client** screen is displayed with **Navigator** and **Home** panes; the **Navigator** pane with network entities as nodes and the **Home** pane with network entities as icons.

2. In the **Navigator** pane, select **Home** > **Hosts and Clusters**.

   Preconfigured hosts, clusters, and datacenters are displayed as navigation tree nodes, in the **Navigator** pane.

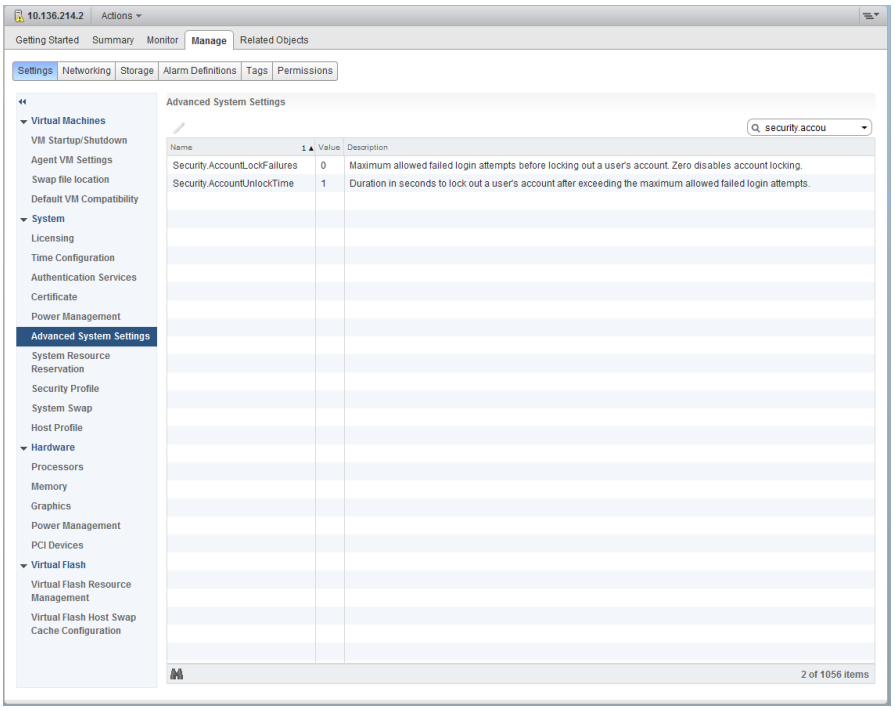
3. Configure the security-related parameters:

   a. Select an ESX server, under a host/cluster/datacenter node.

      The ESX servers are built under a host, cluster, or datacenter node, in the **Navigator** pane.

   b. Under **Manage** menu, select **Settings** > **System** > **Advanced System Settings**.

A list of system-defined parameters, with predefined parameter values, is displayed in the work area.



c. Using the **Search** ( ) field, filter the following security parameters:

- `Set Security.AccountLockFailures` =*0*
  The parameter value defines the maximum number of failed login attempts allowed, before locking out the user's account. The zero (0) value disables account locking.

  Note

  To disable the locking out of user accounts, the parameter value must be set as zero (0).

- `Set Security.AccountUnlockTime` = *1* (in seconds)
  The parameter value defines the duration in seconds to lock out a user's account, after exceeding the maximum number of allowed failed login attempts.

  Note

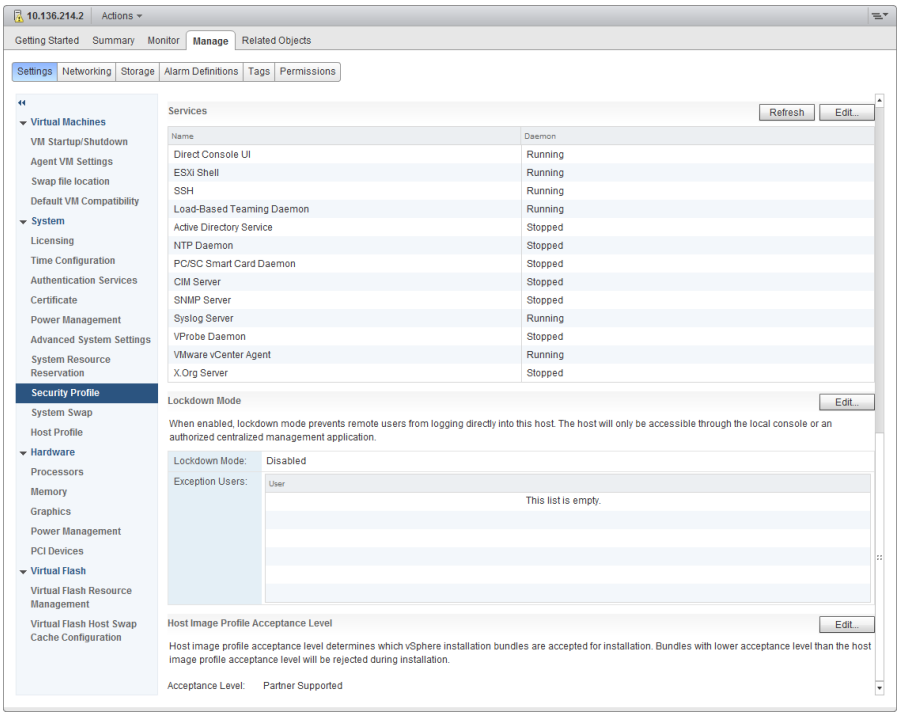
  For ease of operation, it is recommended that the parameter value be set as 1 or 2 seconds.

d. If required, update the security parameter values by clicking the **Edit** ( ) icon.

4. Configure the `Lockdown Mode` parameter:

   a. In the **vSphere Web Client** screen, under **Manage** menu, select **Settings** > **System** > **Security Profile**.

System-defined parameter panes such as **Services** and **Lockdown Mode**, with predefined parameter options, are displayed in the work area.



b. Scroll down to access the **Lockdown Mode** pane and verify the value of the `Lockdown Mode` parameter.

Note

The parameter option of `Lockdown Mode` must be Disabled. If it is not, change it. To change the parameter option from *Enable* to *Disable*, click **Edit** in the **Lockdown Mode** pane and change the parameter value to *Disable*.

---

5. Configure the SSH parameter:

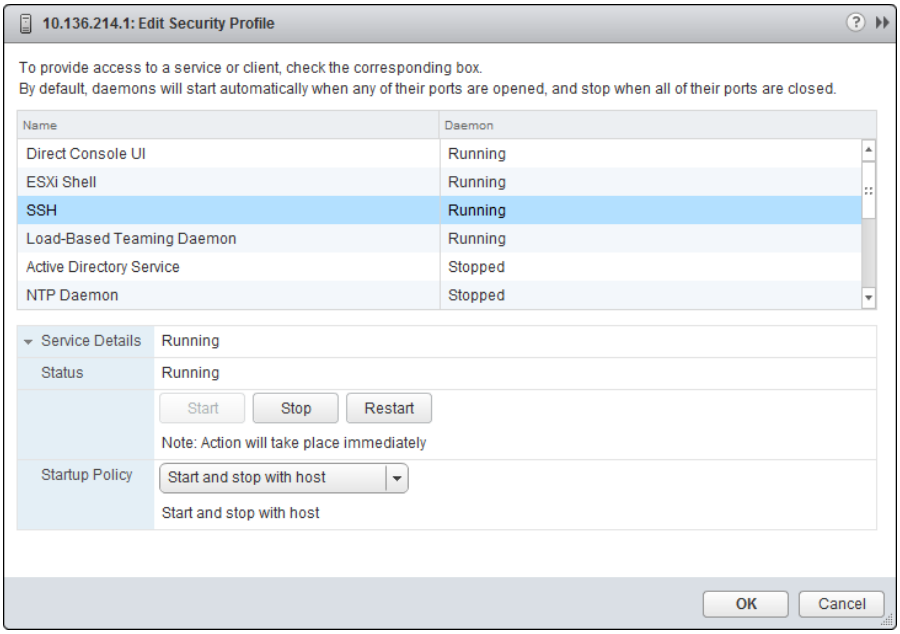
   a. In the **vSphere Web Client** screen, under **Manage** menu, select **Settings** > **System** > **Security Profile**.

   System-defined parameter panes such as **Services** and **Lockdown Mode**, with predefined parameter options, are displayed in the work area.

   b. Scroll down to access the **Services** pane and verify the SSH option.

Note

SSH must be Running (that is, the SSH connection is established). If SSH is Stopped, you must change it to Running.



c. To change SSH option to Running:

   a. Click **Edit** in the **Services** pane.

   b. In the **Edit Security Profile** screen, select the SSH parameter.
   The SSH status is *Stopped*.

   c. Click **Start**.
   The SSH status changes to *Running*.

   For SSH status as *Stopped*, **Start** is active and **Stop** is grayed out.

d. Click **OK**.

   The SSH status changes to *Running*.

6. Restart the management service or restart the server.

CHAPTER 8

Maintaining a ScaleIO System

This chapter describes how to maintain ScaleIO systems. Topics include:

The upgrade procedures are described in the *ScaleIO Software Upgrade Guide*.

- Maintaining a ScaleIO system 196

Maintaining a ScaleIO system

This section describes how to perform the following maintenance activities to an existing ScaleIO system:

- "Installing RFcache on servers in an existing ScaleIO system"
- "Extending the MDM cluster from 3 to 5-node"
- "Creating a LockBox for SNMP/ESRS"
- "Switching to secured authentication mode"
- "Using SCLI in non-secure mode"
- "Extending an existing ScaleIO system"
- "Removing components"

Log collection is described in the *EMC ScaleIO Log Retrieval Technical Notes*, which can be downloaded from EMC Online Support.

Installing RFcache on servers in an existing ScaleIO system

This section describes how to install RFcache to servers in an existing ScaleIO system.

Proceed to the section that matches your operating system environment:

- "Installing RFcache in physical servers"
- "Installing RFcache in VMware servers"

Installing RFcache in physical servers

Installing RFcache in physical Linux or Windows servers is performed with the Installation Manager and the CSV topology file, as described in "Extending an existing ScaleIO system".

The only difference between Linux and Windows installation is that in a Windows installation, the servers need to be restarted after installing.

To install RFcache, perform the following:

Procedure

1. Update the CSV file with the RFcache fields, as described in "Preparing the CSV topology file".
2. Follow the instructions in "Extending an existing ScaleIO system".

Installing RFcache in VMware servers

Installing RFcache in VMware servers is performed manually and with the VMware plug-in. First, you copy the package to the SVMs, and then use the plug-in to configure its use.

To install RFcache on SVMs, perform the following:

Procedure

1. Copy the RFcache package to the SVMs.
2. Install RFcache on every SDS server that is to provide acceleration. RFcache is enabled with the xcache package:

- If the server was upgraded from v1.32.x.x, copy the following file to all SVMs running an SDS, then install it with the following command:

```
rpm -i EMC-ScaleIO-xcache-2.0-xxx.
0.sles11.3.*SVM*.x86_64.rpm
```

- If the server was installed as v2.0 or later, copy the following file to all v2.0 SVMs running an SDS, then install it with the following command:

```
rpm -i EMC-ScaleIO-xcache-2.0-xxx.0.sles11.3.x86_64
```

3. Use the VMware plug-in to enable RFcache on the SDS:
4. Click **SDSs**.
5. Right-click an SDS, and select **Add SDS device**.
6. Select **RFcache**, then click **OK**.
7. Repeat the previous step for every SDS on which you want to enable RFcache.

Extending the MDM cluster from 3 to 5-node

This section describes how to extend the MDM cluster from 3-node to 5-node.

Proceed to the section that matches your operating system environment:

- "Extending the MDM cluster in physical servers"
- "Extending the MDM cluster in VMware servers"

Extending the MDM cluster in physical servers

Before you begin

If you have upgraded your ScaleIO environment from v1.32.x, you must modify the CSV topology file prior to extending the MDM cluster. For more information, see Converting the CSV topology file after upgrade from ScaleIO v1.32.x on page 198.

You can extend the MDM cluster in physical Linux or Windows servers using the Installation Manager and the CSV topology file.

To extend the MDM cluster, perform the following:

Procedure

1. Get one of the following CSV topology files:
   - The converted CSV used to deploy a v1.32.x system that has been upgraded to v2.x.
   - The CSV used to deploy the v2.0 system in its current 3-node cluster mode.
   - Download the complete or minimal CSV (provided in the ISO, or can be downloaded from the Installation Manager) and fill-in the current system topology fields in the CSV.
2. Edit, and save the CSV with one of these options:
   a. Add two new hosts (two new lines) to the System topology, and in the **Is MDM/TB** column for those lines, designate one as a Slave and one as a TieBreaker (TB) role.
   b. In the **Is MDM/TB** column of two existing hosts that were not part of the MDM cluster (Figure 13), add a Slave and a TieBreaker (TB) role.

Figure 15 Before extending

	A	B	C	D	E	F	G
1	IPs	Password	Operating System	Is MDM/TB	Is SDS	SDS Device List	Is SDC
2	10.76.60.1	Password1	linux	Primary	Yes	/dev/sdb	Yes
3	10.76.60.2	Password1	linux	Secondary	Yes	/dev/sdb	Yes
4	10.76.60.3	Password1	windows	TB	Yes	g	Yes
5	10.76.60.4	Password1	linux		Yes	/dev/sdb	Yes
6	10.76.60.5	Password1	windows		Yes	g	Yes

Figure 16 After extending

	A	B	C	D	E	F	G
1	IPs	Password	Operating System	Is MDM/TB	Is SDS	SDS Device List	Is SDC
2	10.76.60.1	Password1	linux	Master	Yes	/dev/sdb	Yes
3	10.76.60.2	Password1	linux	Slave	Yes	/dev/sdb	Yes
4	10.76.60.3	Password1	windows	TB	Yes	g	Yes
5	10.76.60.4	Password1	linux	Slave	Yes	/dev/sdb	Yes
6	10.76.60.5	Password1	windows	TB	Yes	g	Yes

3. From the **Packages** tab, upload all ScaleIO packages, per the host OS.
4. From the **Install** tab, select the edited CSV file, and select **Add to existing system** from the drop-down menu.
5. Click **Upload installation CSV**.
6. Start the installation, and monitor as normal.

Converting the CSV topology file after upgrade from ScaleIO v1.32.x

Convert the v1.32.x CSV topology file prior to extending the MDM cluster from 3 nodes to 5 nodes.

Procedure

1. Open the CSV file used to deploy the v1.32.x system.
2. Change the column header from **SDS Pool List** to **StoragePool List**.
3. Replace references to v1.32.x roles with their corresponding v2.x roles.

   Instances of "Primary" should be changed to "Master." Instances of "Secondary" should be changed to "Slave."

After you finish

Proceed with Extending the MDM cluster in physical servers on page 197.

Extending the MDM cluster in VMware servers

This section describes how to extend the MDM cluster in VMware servers. This task is performed by adding two additional ESX servers, and assigning the new roles to them, described in "Task 1: Extending the MDM cluster".

When you add new manager MDMs, to ensure continued SDC-MDM communication, you should update the SDCs in the system with the new MDM IP addresses. You can do this simply with the VMware plug-in, described in "Task 2: Updating the SDC parameters".

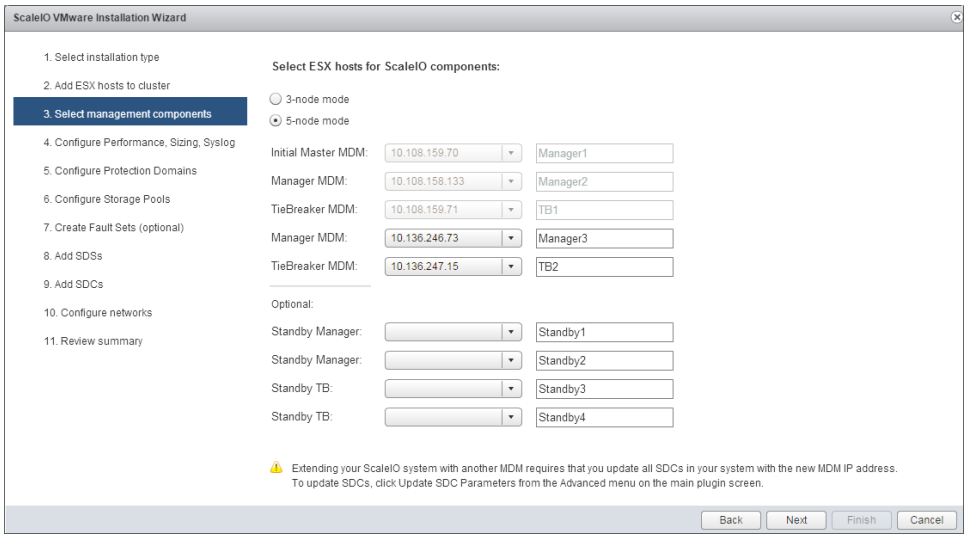
Note

If an SDC in your system is from ScaleIO pre-v1.32.4, you must contact EMC Support to get an update `drv_cfg` file for your environment.

Extending the MDM cluster

Procedure

1. If the ESX servers to be added do not have the SDC component on them, install the SDC on each of the ESX servers, as described in "Task 3: Installing the SDC on ESX hosts".
2. Begin a new deployment, starting with step 1.
3. In the **Select Installation** screen, select **Add servers to a registered ScaleIO system**, and select the system you want to extend.
4. In the Select Management Components screen, perform the following:
5. Select **5-node mode**.
6. In the **Manager MDM** and **TieBreaker MDM** fields that now appear, select the nodes to add to the cluster.



7. Click **Next**, and continue the deployment.

   You can skip steps that do not need to be changed.

   Note

   When adding components, the wizard adjusts the displayed screens to options that are relevant to the current ScaleIO system.

8. Complete the deployment.

Updating the SDC parameters

This section describes how to use the VMware plug-in to update the SDCs with system parameters that are needed to maintain SDC-MDM communication.

Procedure

1. From the plug-in **Advanced tasks** menu, click **Update SDC parameters**, and follow instructions to complete that process.
2. Ensure that SDC parameters were updated by running this command on each ESX:

```
cat /etc/vmware/esx.conf |grep scini|grep -i mdm
```

Creating a LockBox for SNMP/ESRS

A LockBox file must exist on the ScaleIO Gateway for use by the SNMP and ESRS features. The LockBox is used to securely store MDM authentication credentials. SioGWTool is used to create the LockBox, and to add MDM authentication credentials.

For more information about configuring a LockBox for SNMP, and other SNMP configuration activities, see "Configuring SNMP properties after deployment" in the *EMC ScaleIO User Guide*.

For more information about configuring a LockBox for ESRS, and other ESRS configuration activities, see Configuring ESRS connection properties on page 221.

Working with Dynamic Host Name resolution for SNMP in ScaleIO

When working with Dynamic Host Name resolution and SNMP, specific configurations are required in the ScaleIO Gateway and on the DNS server.

Before you begin

Ensure that SNMP is already enabled and that SNMP traps are being received by at least one SNMP trap receiver.

The following procedures are required in order to set up Dynamic Host Name resolution:

- On the ScaleIO Gateway, add the host name of the SNMP trap receiver to the appropriate parameter in the `gatewayUser.properties` file
- On the DNS server, configure the SNMP trap receiver properties, in order to support dynamic host name resolution
- On the DNS server, reduce the `"Time To Live" (TTL)` setting for the SNMP trap receiver

Procedure

1. On the ScaleIO Gateway:
   a. Open the `gatewayUser.properties` file in a text editor. The file location is:

      On Linux: `/opt/emc/scaleio/gateway/webapps/ROOT/WEB-INF/classes/gatewayUser.properties`

      On Windows: `C:\Program Files\EMC\ScaleIO\Gateway\webapps\ROOT\WEB-INF\classes\gatewayUser.properties`

   b. Add the host name of the SNMP trap receiver to the property `snmp.traps_receiver_ip=`. If there is more than one IP address or host name, use a comma-separated list.

c. Save the file.

d. Restart the scaleio-gateway service.

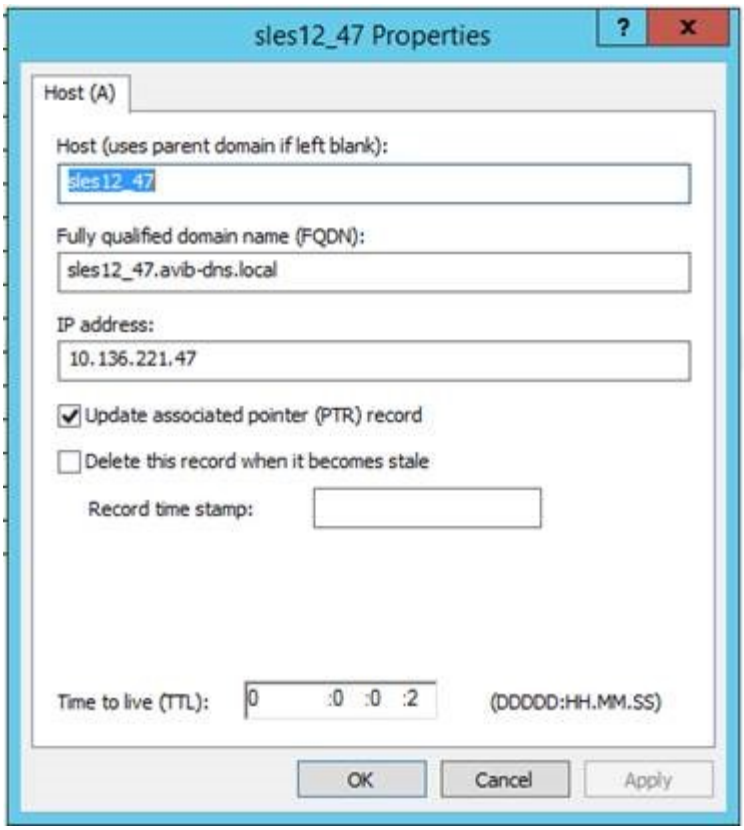
e. Verify that traps are being received at all configured trap receiver hosts.

2. On the DNS server, configure dynamic host name resolution support for the trap receiver.
3. On the DNS server, reduce the TTL setting for the trap receiver. For example, on Windows, perform the following:

   a. Open the **DNS manager** window.

   b. Click **View**, and select the **Advanced** option.

   c. Right-click the trap receiver, and select **Properties**. A window similar to the following is displayed:



   d. At the bottom of the window, in the **Time to live (TTL)** field, change the value from one hour to one or two seconds.

   e. Click **OK**.

Results

Dynamic Host Name resolution configuration is complete.

Switching to secured authentication mode

This section describes how to switch the authentication mode of a ScaleIO system from non-secured to secured mode.

This can be relevant when upgrading to 2.0, where the upgraded system is always in non-secure mode, or after performing a clean install with the Installation Manager and choosing non-secure authentication, via the advanced installation options.

Secure authentication can only be enabled on operating systems that support Open SSL v1.0.1 or later.

You need to update the following areas:

- Internal secure mode (between the MDM and SDS)
- External secure mode (between the management clients and MDM)
- LIA authentication (between the Installation Manager and the LIA)

Proceed to the section that describes your environment:

- "Physical Linux servers"
- "Physical Windows servers"
- "VMware servers"

Physical Linux servers

This topic describes how to change the authentication mode from non-secure to secure on physical Linux servers.

Before beginning, verify the version of the OpenSSL on the server, by running this command:

```
rpm -qa | grep -i openssl
```

If the version is 1.0.1 or greater, you can continue. If it is not, secure authentication cannot be enabled.

Procedure

1. Log in to the SCLI, in a non-secure fashion:

   ```
   scli --login --username <USER> --password <PASSWORD> --
   use_nonsecure_communication
   ```

2. Change to internal secure mode, by running the following command:

   ```
   scli --set_component_authentication_properties --
   use_authentication --use_nonsecure_communication
   ```

3. Change to external secure mode, by running the following command:

   ```
   scli --set_management_client_communication --
   enable_client_secure_communication --
   use_nonsecure_communication
   ```

4. Enable secure LIA authentication, by performing the following:

   Note

   This procedure can take quite some time, and there is no roll-back.

   a. Log in to the IM server (https://*<IM_Server_IP>*).

   b. Click the **Maintain** tab.

   c. From the **Maintenance operation** screen, perform the following:

      a. Type the Master (or Primary) MDM credentials.

      b. Type the LIA password.

c. In **Advanced Options**, select **Allow non-secure communication with LIAs**.

d. Click **Retrieve system topology**.

e. If a certificate approval is required, approve it.

f. Type the authentication credentials again.

g. Click **Retrieve system topology**.

h. Click **Security settings** and select **Enable LIAs security**.

i. Type the authentication credentials again.

j. Confirm the request, and click **Enable LIAs security**.

k. If a certificate approval is required, approve it.

l. When complete, click **Mark operation completed**.

Physical Windows servers

This topic describes how to change the authentication mode from non-secure to secure on physical Windows servers.

Before beginning, verify (in **Programs and Features**) that OpenSSL version 1.0.1 or greater is installed. If it is not, secure authentication cannot be enabled.

Procedure

1. Log in to the SCLI, in a non-secure fashion:

   ```
   scli --login --username <USER> --password <PASSWORD> --
   use_nonsecure_communication
   ```

2. Change to internal secure mode, by running the following command:

   ```
   scli --set_component_authentication_properties --
   use_authentication --use_nonsecure_communication
   ```

3. Change to external secure mode, by running the following command:

   ```
   scli --set_management_client_communication --
   enable_client_secure_communication --
   use_nonsecure_communication
   ```

4. Enable secure LIA authentication, by performing the following:

   Note

   This procedure can take quite some time, and there is no roll-back.

   a. Log in to the IM server (https://*<IM_Server_IP>*).

   b. Click the **Maintain** tab.

   c. From the **Maintenance operation** screen, perform the following:

      a. Type the Master (or Primary) MDM credentials.

      b. Type the LIA password.

      c. In **Advanced Options**, select **Allow non-secure communication with LIAs**.

d. Click **Retrieve system topology**.
e. If a certificate approval is required, approve it.
f. Type the authentication credentials again.
g. Click **Retrieve cluster topology**.
h. Click **Security settings** and select **Enable LIAs security**.
i. Type the authentication credentials again.
j. Confirm the request, and click **Enable LIAs security**.
k. If a certificate approval is required, approve it.
l. When complete, click `Mark operation completed`.

VMware servers

This topic describes how to change the authentication mode from non-secure to secure on VMware servers.

To enable secure authentication mode, perform the following:

Procedure

1. Log in to the SCLI, in a non-secure fashion:

   ```
   scli --login --username <USER> --password <PASSWORD> --
   use_nonsecure_communication
   ```

2. Change the internal secure mode, by running the following command:

   ```
   scli --set_component_authentication_properties --
   use_authentication --use_nonsecure_communication
   ```

3. Change the external secure mode, by running the following command:

   ```
   scli --set_management_client_communication --
   enable_client_secure_communication --
   use_nonsecure_communication
   ```

4. Enable secure LIA authentication, by performing the following:

   a. Open the ScaleIO Gateway Installation Manager:

      a. From the plug-in **ScaleIO Systems** screen, right-click the system for which you want to perform this operation.
      b. Select **Open ScaleIO Gateway**.

      Note

      Opening the Gateway can take some time.

      c. From the **IM welcome** screen, enter the default IM credentials.

   b. From the **IM web client** main menu, click **Maintain**.
   c. In **Advanced Options**, select **Allow non-secure communication with LIAs**.

      The **Maintenance operation** screen appears.

      a. Click **Security settings** and select **Enable LIAs security**.
      b. Confirm the request, and click **Enable LIAs security**.

c. If a certificate approval is required, approve the certificates.

5. From the main plug-in window, click **ScaleIO systems**.
6. Right-click the system to register, and select **Reregister ScaleIO system**.
7. Enter the user name and password of the system, and click **OK**.
8. Approve the Master MDM certificate by clicking **Accept**.
9. From the plug-in **ScaleIO Systems** screen, right-click the system and accept the Slave MDM certificate.

Note

Before performing maintenance operations in a system where replication is enabled on SDC nodes, you should exclude the RPA nodes from the detection list , as described in "Configuring the Installation Manager".

Using SCLI in non-secure mode

If ScaleIO is running in non-secure mode, you must make the following change to enable running commands.

On every MDM server, disable secure communication:

Procedure

1. Windows

   In the SCLI `conf.txt` file, add `cli_use_secure_communication=0`

2. Linux

   Run this command: `echo cli_use_secure_communication=0 >> ~/.scli/conf.txt`

Extending an existing ScaleIO system

This section describes how to add components to an existing ScaleIO installation.

In physical environments, you add components with the Installation Manager. In VMware environments, you add components with the VMware deployment wizard.

Adding components with the Installation Manager

This section describes how to add components with the Installation Manager.

To add components, you first need to update the CSV topology file with the new components, then you can use the web client to add them.

The secure communication mode of components must match the mode of the system to which the components are being added. If they do not match, you must either change the mode of the components to be added or change the mode of the system so that they match.

Procedure

1. Follow the procedure described in "Installing with the Installation Manager".

Note

Use the same LIA password that was configured during initial installation.

2. In the Upload CSV stage, browse to the updated CSV file, and select **Add to existing sys**.
3. Upload the CSV, and continue as normal.

Adding components with the VMware deployment wizard

This section describes how to add ScaleIO components to an existing system with the VMware deployment wizard.

Note

The following procedure cannot be used to add an SDS component to an existing SVM. To do so, contact EMC Support.

Procedure

1. To add components, begin the deployment, starting with step 1.
2. In the **Select Installation** screen, select **Add servers to a registered ScaleIO system**, and select the system you want to extend.
3. Continue with the deployment steps, adding the new nodes.

   You can skip steps that do not need to be changed.

   Note

   When adding components, the wizard adjusts the displayed screens to options that are relevant to the current ScaleIO system.

4. Complete the deployment.

   Note

   After extending an existing SVM with a new ScaleIO role/component, you must perform manual memory allocation on the SVM, as described in "SVM manual memory allocation".

Configuring virtual IP addresses using Installation Manager

Configure virtual IP addresses using the **Maintain** menu in the Installation Manager (IM).

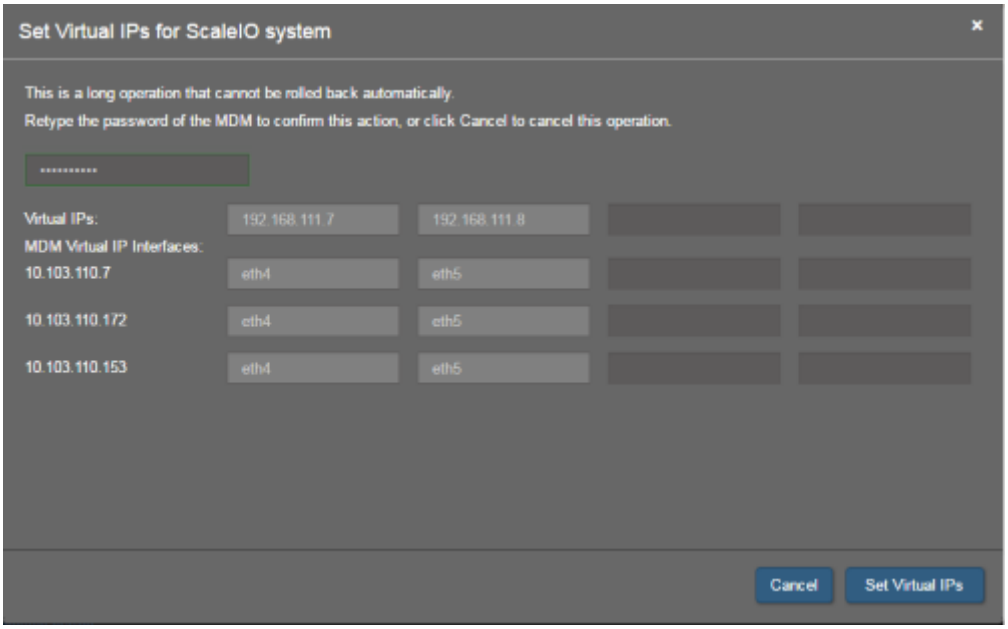
You can assign a virtual IP address for each possible manager MDM, which will be used for communications between the MDM cluster and SDCs. Only one virtual IP address can be mapped to each NIC, with a maximum of four virtual IP addresses per system. Virtual IP addresses are not supported on nodes using Windows operating system. The IM can be used only to assign new virtual IP addresses; to change or remove existing virtual IP addresses, use the appropriate CLI commands.

Procedure

1. In the Installation Manager, select **Maintain**.

2. Select **Set Virtual IPs**.
3. In the **Set Virtual IPs for ScaleIO system** screen, enter the MDM password.

Figure 17 Set Virtual IPs for ScaleIO system screen



4. For each MDM that you wish to set a virtual IP address, enter a virtual IP address and the NIC to which it will be mapped. For each new virtual IP address, enter the virtual IP address and NIC name for each MDM to which it will be mapped.
5. Click **Set Virtual IPs**.

Results

The virtual IP address is configured and all of the SDCs are updated with the new virtual IP address.

Removing components

This section describes how to remove ScaleIO components.

To uninstall ScaleIO components, use the Installation Manager. This requires that the LIA be installed in all nodes to be changed.

When removing RFcache (the xcache package) on a Windows server, a server restart is necessary after the removal.

Proceed to the appropriate section:

- "Removing ScaleIO components using the IM"

To unregister the vSphere plug-in, refer to the *ScaleIO Software Upgrade Guide*.

Removing ScaleIO components using the IM

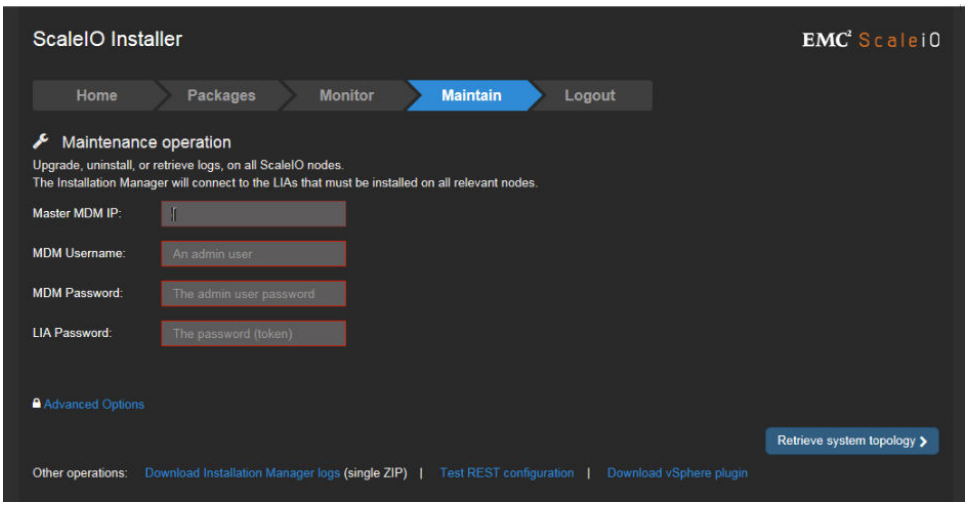
This section describes how to use the Installation Manager to remove installed ScaleIO components. All components in the ScaleIO system that is being accessed will be removed. This information is attained from the LIA that is installed on every node.

To uninstall:

Procedure

1. Log in to the web client, as described in "Installing with the Installation Manager".
2. From the IM web client main menu, click **Maintain**.

   The **Maintenance operation** screen appears:



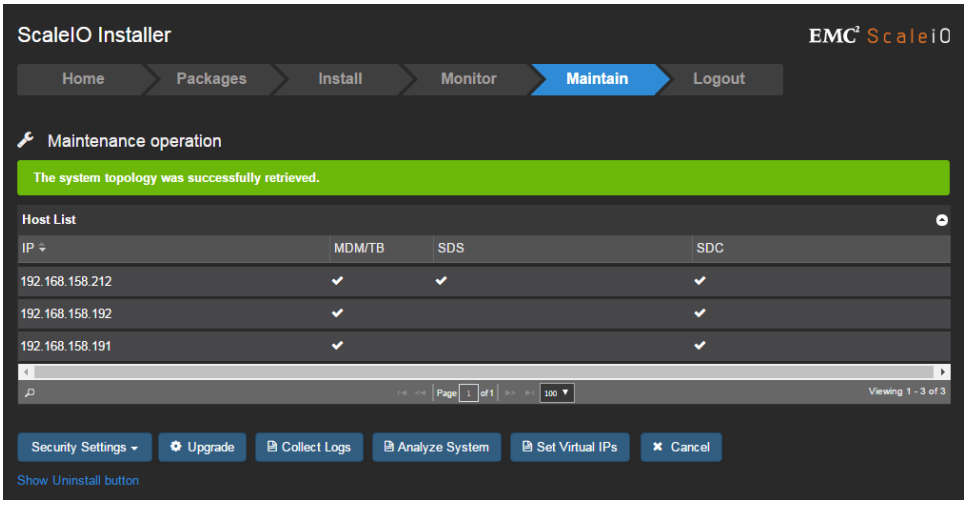
Note

You can download the Installation Manager logs by clicking **Installation Manager logs**.

---

3. Enter the Primary MDM IP address and password.
4. In the **Connection** section, choose **Use LIA**, and type the LIA password.
5. Click **Retrieve system topology**.

   The system topology appears:

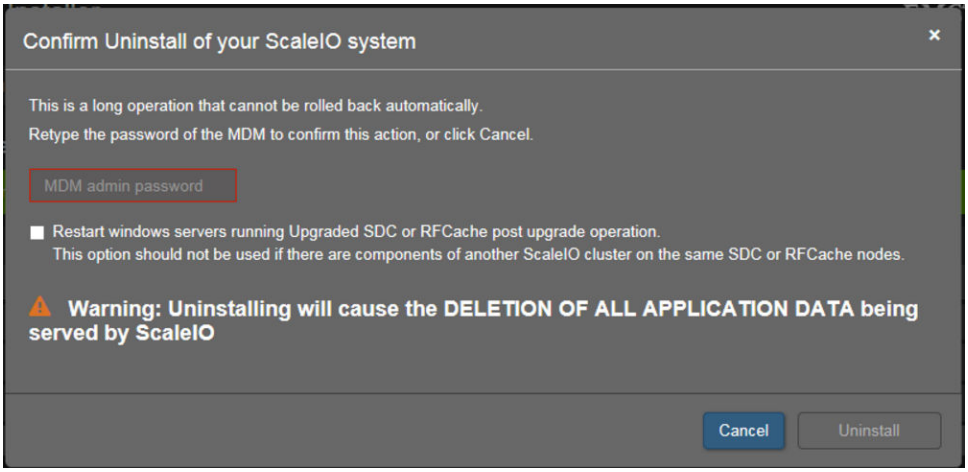


6. Click the **Show Uninstall button** link, and confirm enabling this option.

   The uninstall operation may take some time, depending on your system topology. This operation cannot be rolled back.

7. Click **Uninstall**.

   A confirmation dialog appears:



   NOTICE

   Uninstalling an SDC component requires a machine restart. If you are uninstalling SDC components on Windows servers, select to enable automatic restart (on those servers only). Alternatively, you can manually restart these servers after removing the SDC.

   On Linux servers, if the kernel module is busy, perform a manual restart.

8. Enter the MDM password, select to reboot servers (optional), and click **Uninstall**.
9. To monitor the uninstallation progress, click **Monitor**.
10. When the uninstallation is complete, click **Mark operation completed**.

Unregistering the ScaleIO plug-in

To remove the plug-in, perform the following:

Procedure

1. Run the script to remove the plug-in:
   a. From the folder where you extracted the `EMC-ScaleIO-vSphere-web-plugin-package-1.32-XXX.X.zip` file, use PowerCLI to run the `ScaleIOPluginSetup-1.32-XXX.X.ps1` script.
   b. Select option `2`, `Unregister ScaleIO plugin`.
2. Enter the vCenter credentials and confirm the script actions.
3. Log out, then log back in to the vSphere web client.
4. The plug-in is no longer registered.

CHAPTER 9

System Analysis

This chapter describes the ScaleIO system analysis tool. Topics include:

- System analysis overview 212

System analysis overview

This topic describes ScaleIO system analysis and the environment required to use it.

ScaleIO system analysis enables you to identify potential issues with your ScaleIO system which may prevent best performance. It is highly recommended to analyze the ScaleIO system immediately after deployment, before provisioning volumes, and before using the system in production. You can also use it to check the health of a system that is already operational.

The system analysis is invoked from the ScaleIO Installation Manager (IM). The analysis checks the following:

- ScaleIO components are up and running
- Ping between two relevant nodes in the system
- Connectivity within the ScaleIO configuration (for example, connectivity between SDSs within a Protection Domain, connection of SDCs with the cluster virtual IP address)
- Network configuration
- RAID controller and device configuration

Using the system analysis, you can detect any potential issues in the system, then rectify them, before provisioning and using the system in a production environment.

Environment requirements and prerequisites:

- Supports only RHEL 6.x servers, with the following 3rd-party tools:
  - Netcat
  - StorCLI
  - PERCCLI
  - smartctl
- Requires a ScaleIO Gateway server:
  - On a Linux or Windows server. ScaleIO is tested on RHEL 6.x and RHEL 7.x as well as Windows2K12 servers.
  - With at least 1 GB of free disk space per node in the system
- A web browser, that is supported by the IM.
- Supports LSI RAID controller cards.
- Supports IPv4 network configuration.

Best-practice recommendation:

1. Deploy ScaleIO.
2. Analyze system to identify issues that should be fixed.
3. Fix issues.
4. Analyze system to verify that the issues have been fixed.
5. When the system meets your satisfaction, you can move it into production.

Limitations and compatibility:

- On servers with IPv4 and IPv6, the IPv6 (only) analysis is not conclusive; IPv6 is not supported.

- On servers with MegaRAID, the RAID function analysis is not supported; only StorCLI is supported.
- On servers with multiple SDSs, analysis is performed but it is not conclusive.
- RFCache analysis is not performed.
- Due to default Internet Explorer settings, to expand a report, you may need to grant permission for IE to run scripts. For more information, see the version release notes.

Creating the system analysis report

The topic describes how to create and display the ScaleIO system analysis report.

Before you begin

Ensure that you have access to the following:

- A web browser that is supported by the ScaleIO Installation Manager (IM)
- IP address of the ScaleIO Gateway server
- The Gateway admin username (default: admin) and password (defined during deployment)
- Master MDM IP address, username, and password (defined during deployment)
- LIA password

Procedure

1. Log in to the IM server:

   a. Point your browser to **https://*<IM_Server_IP_address>***

   where **<IM_Server_IP_address>** is the IP address of the server on which the Gateway package is installed.

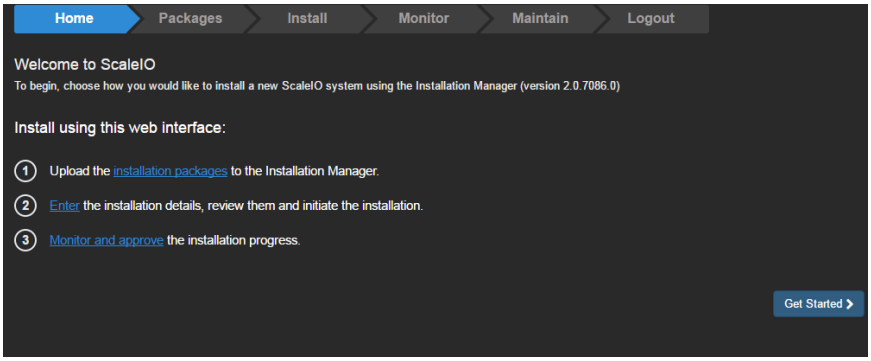
   The **ScaleIO Installer** login screen is displayed:

   b. If a login banner is displayed, accept it to continue.

   c. Type the Gateway username (default is admin) and password.

   d. Click **Login**.

   The **ScaleIO Installer Home** screen is displayed.



2. Generate the system analysis report:

   a. Click the **Maintain** tab.

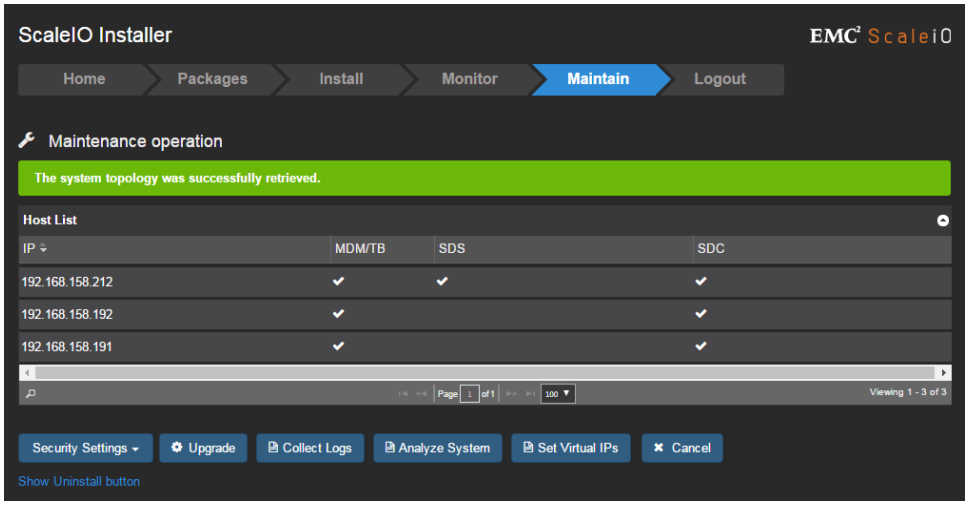
      The **Maintenance operation** screen is displayed.
      If a warning is displayed that indicates that a previous operation is not completed, it means that the last IM operation was not marked as complete. To continue, click the **Monitor** tab, mark the operation complete, then return to the **Maintain** tab.

   b. Enter the authentication credentials:

      - Master MDM IP address, user name, and password
      - LIA password

   c. Click **Retrieve system topology**.

      The system topology is displayed.



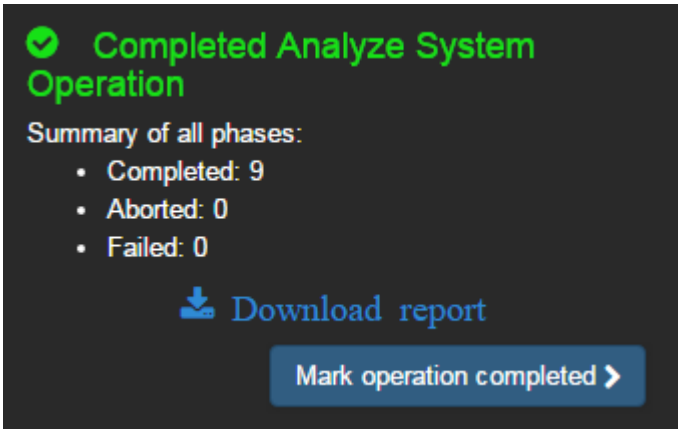
   d. Click **Analyze System**.

      The system analysis begins.

   e. Monitor the progress of the system analysis by clicking the **Monitor** tab.

      When the process completes (which could take several minutes), a **Download report** link is displayed.

   f. Click the **Download report** link.



The report is saved, in ZIP format, in the default download location of the server on which the report was run.

g. To enable the IM to be available for subsequent operations, click **Mark operation completed**.

3. Display the analysis report:

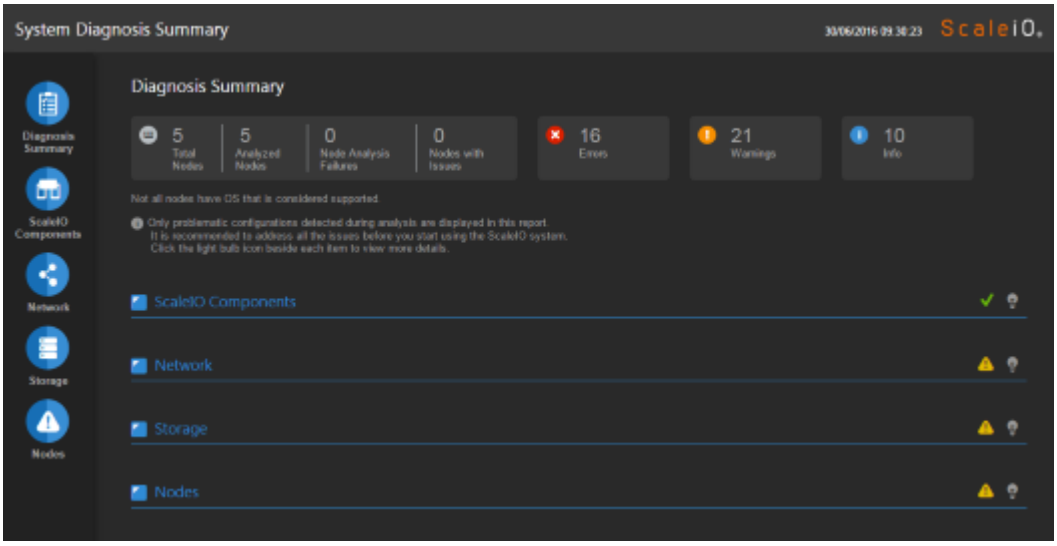
   a. Unzip the downloaded file.

   The file includes the `ScaleioSystemDiagnosisReport.html` analysis file, and several TGZ files (one for each node, in the `dumps` folder).

   b. Double-click `ScaleioSystemDiagnosisReport.html`.

Results

The **System Diagnosis Summary** report is displayed in the default web browser on your computer.

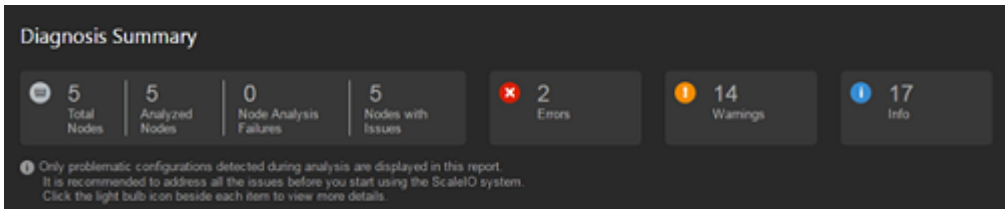


System analysis report description

This topic describes how to get the most benefit from the system analysis report.

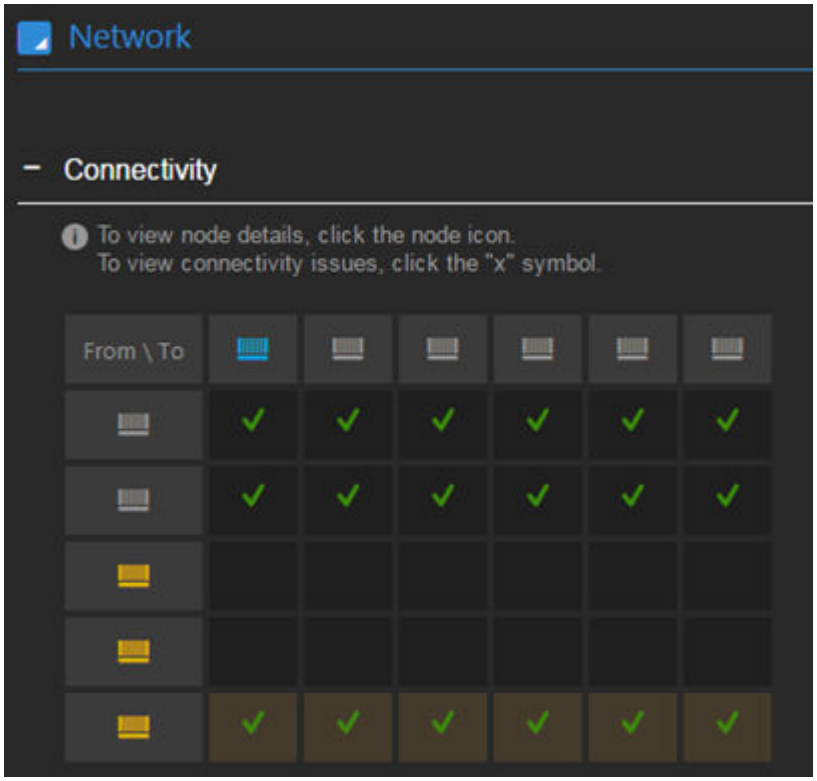
At the top of the report, the Diagnosis Summary shows the number of the issues that have been detected system-wide. The summary shows the following categories:

- Node analysis
  - How many nodes are analyzed
  - How many nodes could not be analyzed
  - How many nodes have issues
- Severity analysis
  - How many issues of each severity level were found



When the analysis first opens, the major sections are shown in summary form. You can expand them to show detailed diagnostic reports, as follows:

- ScaleIO components:
  This section of the report shows the non-running ScaleIO server components, that is, SDS, SDC, and MDM. Failure of these components may affect system performance and data availability. Each of the ScaleIO server components supports the following functionality:
  - SDS server
    The SDS (ScaleIO Data Server) manages the capacity of a single server and acts as the back-end for data access. The SDS is installed on all the servers contributing storage devices to the ScaleIO system. Failure of an SDS may affect the cluster performance and data availability.
  - SDC server
    The SDC (ScaleIO Data Client) server is installed on each server that needs access the ScaleIO storage and it is the gateway to the ScaleIO storage. Failure of an SDC server denies its application access to the ScaleIO storage.
  - MDM server
    The MDM (Meta Data Manager) server controls and monitors the ScaleIO system. Failure of an MDM may affect the cluster performance and data availability.
- Network:
  This section of the report checks the connectivity between various ScaleIO components, as well as the NIC configuration and performance. The network issues may impact the system performance and data availability.



- Connectivity
  Performing pings between ScaleIO components leads to detecting and resolving connectivity-related issues in the system. If the regular pings succeed, then the MTU pings, followed by the Ncat pings are performed to the ports used by the ScaleIO application.

  If virtual IP addresses are assigned to the MDMs in the cluster, a logical host, called *MDM cluster* is displayed (represented by a blue host icon) in the analysis report. The following issues are tested:

  - SDC connectivity with the virtual IP addresses.
  - All virtual IP addresses are configured.
  - Only one physical node is configured to use the virtual IP address.

  To view specific information:

  - To view the node details, click the node icon.
  - To view connectivity issues, click the X symbol
- NIC Configuration and Performance
  This section of the report displays the configurations that do not meet the best practice recommendations described in the user documentation.
- MDM cluster
  Shows status of virtual IP addresses assigned to the MDMs in the cluster.

- Storage:
  The Storage section of the report describes the issues associated with the RAID controllers, storage devices, and storage pool configurations.

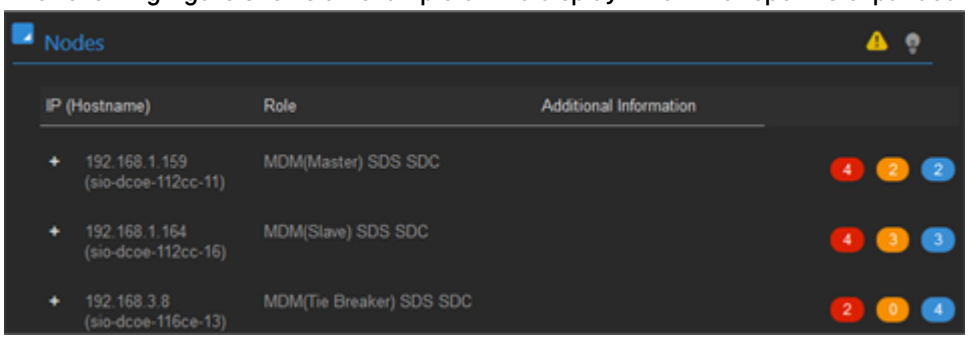
  - RAID Controller
    Describes the issues related to the physical disks

  - Devices
    Describes the list of problematic or potentially problematic storage devices
  - Storage Pool Uniformity
    Indicates Storage Pools with non-homogeneous disk performance
- Nodes (groups all of the reported issues per node): Describes the detected issues, as listed above, grouped by node

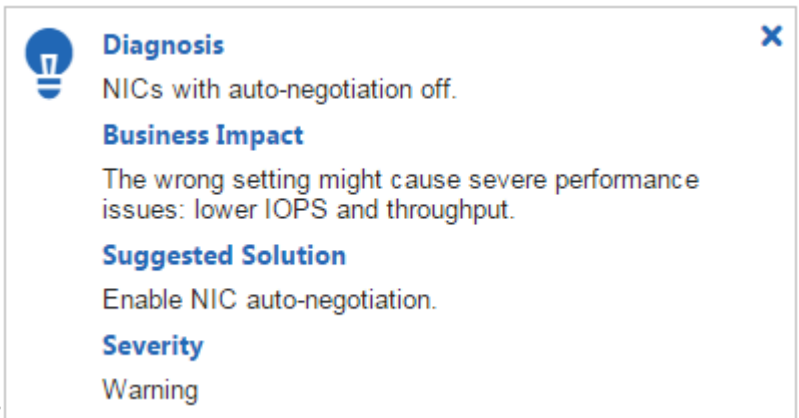
You can show more (or less) information, as follows:

- Use the Expand ( ) button and + symbol to drill down.
- Use the Collapse ( ) button and – symbol to hide information.

The following figure shows an example of the display when the report is expanded.



You can show additional details of an error by clicking the ( ) icon. A pop-up window similar to the following is displayed.

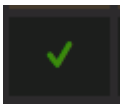
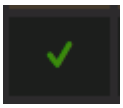
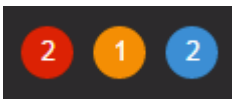


The following table describes how to navigate the analysis information:

To display this...	Click this symbol...
Diagnosis Summary	
ScaleIO component issues, sorted according to SDS, SDC, and MDM.	

To display this...	Click this symbol...
Network connectivity (including virtual IP address configuration) and NICs	
Storage, including RAID controller, devices, and Storage Pool uniformity	
Nodes, displayed according to IP address and system role	
Collapse a list	
Expand a list	
Open a pop-up containing Diagnosis, Business Impact, and Suggested Solution information.	

The following table describes other symbols and interface elements used in the analysis report display:

Symbol / Interface Elements	Description
	All connectivity tests passed
or	Connectivity issues were detected (for more information, click the symbol)
 (black background)	Connectivity matrix background: all pertinent connectivity tools are available
(brown background)	Connectivity matrix background: connectivity tests could not be performed, due to a missing tool
	Error counters: • Red - Error • Orange - Warning • Blue - Info
	All connectivity test tools are available for use on the node
	Some connectivity tests could not be performed on this node, due to missing tools (for more information, click the symbol)

Symbol / Interface Elements	Description
	A logical host, that represents the MDM cluster, when virtual IP addresses are configured

Note

An empty cell in the connectivity matrix indicates that no connectivity check was performed. In such cases, no connectivity is expected.

CHAPTER 10

Configuring ESRS connection properties

This chapter explains how to configure connection properties, to allow communication between the ScaleIO system and ESRS (EMC Secure Remote Support) servers. ESRS can be used by ASP Partners for remote hardware support. For more information about using ESRS, refer to ESRS documentation.

Use SioGWTool in command line to configure the connection properties.

Topics in this chapter include:

- Before configuring ESRS 222
- Registering the system with the ESRS Gateway 222
- Performing other ESRS configuration activities 224
- Validating ConnectEMC Dial-Home 225

Before configuring ESRS

Before configuring ESRS, ensure that your system meets the following requirements, and that you have the following information:

- Requirements:
  - ESRS Gateway v3 version 3.08 or higher must be installed and configured.
  - ESRS Gateway must be reachable from the ScaleIO-Gateway node on port 9443.
  - The ScaleIO license must be installed.
  - The ScaleIO Master MDM Management IP address to be used as the Connect-In IP address must be an IP address that is accessible from the ESRS Gateway (for example, in case of NAT). Ensure that you know this IP address, because you will need it when you perform the registration procedure.
  - On the ScaleIO Gateway, in the `gatewayUser.properties` file, set the property `features.enable_esrs=true`
- Information:
  - One or more IP addresses of the ESRS servers.
  - ESRS user and password credentials for the ScaleIO command.
  - MDM user name and password.

Registering the system with the ESRS Gateway

Enable use of the ESRS feature on your ScaleIO system, add the ESRS Gateway's certificate to the truststore, and register your system on the ESRS Gateway.

Before you begin

See Before configuring ESRS on page 222.

The REST procedures described below are performed using cURL. You may use other similar tools and their corresponding commands to perform those steps.

A web browser is required for this procedure.

In order to work with ESRS, the LockBox must be configured, and the MDM credentials must be added to it. The LockBox is required for both SNMP and ESRS. The same LockBox is used for both functionalities. Use SioGWTool in command line to create a LockBox, to register the ESRS gateway address, and to set the connect-in IP address. SioGWTool is located in:

- Linux:

  `/opt/emc/scaleio/gateway/bin/SioGWTool.sh`
- Windows:

  `C:\Program Files\EMC\ScaleIO\Gateway\bin\SioGWTool.bat`

Procedure

1. If no LockBox exists on the ScaleIO system (one may already have been created for SNMP, for example), create a LockBox. In command line, use the following SioGWTool commands:

```
<SioGWTool_PATH> --change_lb_passphrase --new_passphrase
<new_passphrase>
```

```
SioGWTool --set_mdm_credentials --mdm_user <mdm_user_name>
--mdm_password <mdm_password>
```

2. Enable the ESRS feature in the `gatewayUser.properties` file.
   a. Use a text editor to open the `gatewayUser.properties` file, located in the following directory on the Installation Manager/ScaleIO Gateway server:

Gateway installed on	Location of gatewayUser.properties file
Windows	`C:\Program Files\EMC\ScaleIO\Gateway\webapps\ROOT\WEB-INF\classes\`
Linux	`/opt/emc/scaleio/gateway/webapps/ROOT/WEB-INF/classes`

   b. Change the enable_esrs feature to "true":

      features.enable_esrs=true

   c. Save the file.
3. Add the ESRS Gateway's certificate to the truststore.
   a. In a web browser, browse to `<ESRS_Gateway_IP_address:9443>`.
   b. Download the certificate that is displayed, and save it as a file.
   c. Log in to REST and get a token. Make a note of the token.

```
curl -k -v --basic --user
admin:<mdm_admin_password>  https://
<ScaleIO_Gateway_IP_address>/api/login
```

   d. Using REST, add the certificate to truststore.

```
curl -k -v --basic -
uadmin:<token_received_from_previous_command> --form
"file=@<path_to_certificate_file>" https://
<ScaleIO_Gateway_IP_address>/api/trustHostCertificate/Mdm
```

4. Restart the scaleio-gateway service.

Option	Description
Windows	Restart the EMC ScaleIO Gateway service
Linux	Type the command `service scaleio-gateway restart`

5. Register the ScaleIO system on the ESRS Gateway. In command line, use the following SioGWTool command:

```
<SioGWTool_PATH> --register_esrs_gateway --scaleio_gateway_ip
<IP_ADDRESS:PORT> --scaleio_gateway_user <USER> --
```

```
scaleio_gateway_password <PASSWORD> --esrs_gateway_ip
<IP_ADDRESS> --esrs_gateway_user <USER>
--esrs_gateway_password <PASSWORD> --connect_in_ip <IP_ADDRESS>
```

where `--esrs_gateway_user` is the user name used for EMC support (typically, an email address), and `--esrs_gateway_password` is its corresponding password. `--connect_in_ip` is the MGMT IP address of the current Master MDM.

Results

ESRS registration is complete.

Performing other ESRS configuration activities

Use SioGWTool in command line for ESRS configuration activities.

SioGWTool is located in:

- Linux: `/opt/emc/scaleio/gateway/bin/SioGWTool.sh`
- Windows: `C:\Program Files\EMC\ScaleIO\Gateway\bin\SioGWTool.bat`

Procedure

1. The following activities can be performed with SioGWTool:

Activity	Command
To unregister from all ESRS Gateways, type the command:	`<SioGWTool_PATH> --unregister_esrs_gateway --scaleio_gateway_ip <ip address:port> --scaleio_gateway_user <user> --scaleio_gateway_password <password>`
To remove a specific ESRS Gateway from the ScaleIO Gateway configuration, type the command:	`<SioGWTool_PATH> --remove_esrs_gateway --scaleio_gateway_ip <scaleio_gateway_ip_address> --scaleio_gateway_user <scaleio_gateway_user> --scaleio_gateway_password <scaleio_gateway_password> --esrs_gateway_ip <esrs_gateway_ip_address>`
To reset the ScaleIO Gateway admin password, type the command:	`<SioGWTool_PATH> --reset_password`
To start ESRS logic on the ScaleIO Gateway (alerts will be sent ot the ESRS server), type the command:	`<SioGWTool_PATH> --start_esrs`
To stop ESRS logic from running on the ScaleIO Gateway (alerts will not be sent to the ESRS server), type the command:	`<SioGWTool_PATH> --stop_esrs`

Activity	Command
To check the connection to a registered ESRS Gateway (probing will be done throughout the servers to find one that can be connected), type the command:	`<SioGWTool_PATH> --check_esrs_connectivity`
To show a list of the available SioGWTool commands, type the command:	`<SioGWTool_PATH> --help`

Validating ConnectEMC Dial-Home

Verify that ConnectEMC is correctly configured and is communicating with EMC Global Services, by performing a dial-home test.

This procedure describes how to quickly verify the connectivity between the ESRS-GW and ESRS servers, either after installation or after registering to ESRS. The ESRS GW is located at the customer's site and has a verified connection to ESRS servers as part of its installation.

Procedure

1. While logged in to the ESRS- gateway as the gpadmin Operating System user, remove the contents of the directory `/opt/connectemc/archive`.
2. Create a test event for ConnectEMC using the following command:

```
touch /opt/connectemc/poll/testerror
```

Results

Within up to 10 minutes (depending on your polling interval setting), ConnectEMC will generate a file in one of the following directories, depending on the test result:
If the connection was successful, a file will be saved in `/opt/connectemc/archive`.

If the service fails, a file will be saved in `/opt/connectemc/archive/failed`.

CHAPTER 11

Troubleshooting

This chapter describes ScaleIO troubleshooting. Topics include:

- Troubleshooting ScaleIO 228
- Adding a cache device to RFcache Storage Pool 228
- After SDC installation, the ScaleIO SVM does not start automatically 230
- Application server does not see a ScaleIO volume 231
- Cannot log in to the Installation Manager after upgrade 231
- Certificate error when installing SDC on Windows servers 232
- Enabling acceleration in a new node 232
- S.M.A.R.T. hardware alerts are not displayed 235
- Gateway server recovery 236
- Installation Manager returns an error 239
- Installation with the Installation Manager fails 239
- LIA upgrade fails in Windows 2008 240
- Older SDCs cannot communicate with newer SDSs after upgrade to v2.x 240
- Removing RFcache (the xcache package) from Windows servers 241
- Removing RFcache leftovers from the Windows OS registry 241
- Replacing a faulty caching device in ScaleIO 241
- Replacing a faulty cached storage device in ScaleIO 243
- ScaleIO CLI or GUI cannot connect to an MDM 244
- SCLI add_sds command fails due to communication error or MDM crash 245
- SVM manual memory allocation 245
- Viewing the status of volumes by drv_cfg --rescan command 247
- Virtual IP feature is not functional 247
- The VMware plug-in responds slowly 251
- Solving ScaleIO performance issues 251
- Mismatch in IO counters 252
- Deploying SVM on a node with a management IP address from a different subnet than the node with the SVM template 252
- Speeding up rebuild and rebalance processes 252
- RFcache is not installed on an upgraded Windows server 253
- SSD devices are not recognized in the operating system 254
- ESX competing thread setting resets on server reboot 255

Troubleshooting ScaleIO

This chapter describes solutions to issues that may arise. Most of the issues requiring troubleshooting are related to events and CLI return messages:

Event notifications
Upon changes or events in the system, ScaleIO may generate a system event that will be logged in a file. When addressing an event, look for the proper entry in the System Events appendix of the *EMC ScaleIO User Guide*, according to the name of the event received. For each event that requires attention, the entry will include a description of a possible action to take. Follow the instructions before contacting EMC Support.

CLI messages
When issuing a CLI command, ScaleIO generates a textual response describing the outcome of the command. In some cases, the result may be a failure. Upon receiving a failure message, you can address it by looking for the proper entry in the Return Messages appendix of the *EMC ScaleIO User Guide*, according to the text of the message received.

For each return message that requires attention, the entry will include a description of a possible action to take. Follow the instructions before contacting EMC Support.

Adding a cache device to RFcache Storage Pool

Add a cache device to an RFcache Storage Pool in a ScaleIO system.

Before you begin

Ensure that you have admin-level username and password for accessing ScaleIO.

Each RFcache pool on an SDS supports up to eight SDS devices as members in the RFcache Storage Pool. You can add additional devices while caching is in progress.

Procedure

1. Log in to an SDS.
2. Identify the SSD and HDDs to be accelerated and used as the cache devices, by running the following system commands:

SDS Operating System	Description
ESXi	Run the esxcli commands or open the vSphere View GUI to check the device properties.
Linux	Run the following commands: a. Run the command `fdisk -l` and check the size of the device in the output. b. Run the command `smartctl` and identify the type of device in the output.
Windows	Run any of the following commands: • `wmic diskdrive list brief /format:list` and check the device properties in the output.

SDS Operating System	Description
	• `wmic diskdrive get name,size,model` and check the device name, size, and model number in the output.

3. Log in to the Master MDM.

4. Add an SSD device to the default cache pool, by running the following command:

```
scli --add_sds_rfcache_device (--sds_id <ID> |--sds_name
<NAME> | --sds_ip <IP> [--sds_port <PORT>])
--rfcache_device_path <PATH> [--rfcache_device_name <NAME>]
```

Example:

```
scli --add_sds_rfcache_device --sds_name SDS1 --
rfcache_device_path /dev/sdb --rfcache_device_name
rfcache1_SDS1_sdb
```

5. Repeat the previous steps, from the beginning of the procedure, for each new SDS device to be configured in the RFcache Storage Pool.

6. Verify the status of acceleration per SDS, by running the following command:

```
scli --query_sds (--sds_id <ID> | --sds_name <NAME> |--sds_ip
<IP>)
```

Example:

```
scli --query_sds --sds_ip 192.168.1.6
```

Output, similar to the following, appears, for different command types:

- Command to verify the RFcache status:

```
scli --query_sds (--sds_id <ID> | --sds_name <NAME> |--
sds_ip <IP>)
```

Example:

```
scli --query_sds --sds_ip 9.10.21.141
```

Output:

```
SDS 4eecc17300000002 Name: SDS_10.136.215.203 Version:
2.0.13000
```

```
Protection Domain: 928cd6e100000000, Name:
Default_Protection_Domain
DRL mode: Volatile
Authentication error: None
IP information (total 2 IPs):
         1: 9.10.21.141      Role: All (SDS and SDC)
         2: 9.10.121.141     Role: All (SDS and SDC)
        Port: 7072
RAM Read Cache information:
        128.0 MB (131072 KB) total size
        Cache is enabled
        RAM Read Cache memory allocation state is PENDING
Rfcache enabled
.
.
.

Rfcache device information (total 1 devices):
         1: Name: 50000396CC883494  Path: /dev/sdc  Original-
path: /dev/sdc  Size 762495MB  ID: e177c11700020000
```

- Command to verify storage pool status:

```
scli --query_storage_pool (((--protection_domain_id <ID> |
--protection_domain_name <NAME>) --storage_pool_name
<NAME>) | --storage_pool_id <ID>)
```

Example:

```
scli --query_storage_pool --storage_pool_name
Cached_Hdd_Pool_1 --protection_domain_name
Default_Protection_Domain
```

Output:

```
Storage Pool Cached_Hdd_Pool_1 (Id: ada1bb8500000000) has 0
volumes and 23.0 TB (23536 GB) available for volume allocation
        The number of parallel rebuild/rebalance jobs: 2
        Rebuild is enabled and using Limit-Concurrent-IO
policy with the following parameters:
        Number of concurrent IOs per device: 1
        Rebalance is enabled and using Favor-Application-IO
policy with the following parameters:
        Number of concurrent IOs per device: 1, Bandwidth
limit per device: 10240 KB per second
        Background device scanner: Disabled
        Zero padding is enabled
        Spare policy: 36% out of total
        Checksum mode: disabled
        Doesn't use RAM Read Cache
        Uses Flash Read Cache
```

Results

The new cache devices are configured in the RFcache Storage Pool in ScaleIO.

After SDC installation, the ScaleIO SVM does not start automatically

After deployment is completed, set all ScaleIO VMs to start automatically:

Procedure

1. Click the ESX **Configuration** tab.
2. From the **Software** section, click **Virtual Machine Startup/Shutdown**.
3. Click **Properties**.
4. In the dialog box, select **Allow virtual machines to start and stop automatically with the system**.
5. Select the SVM and move it to the **Automatic Startup** list.
6. Click **OK**.
7. Repeat this process for all SVMs.

Application server does not see a ScaleIO volume

Perform the following steps:

Procedure

1. Check if the ScaleIO system is operational by typing the following command:

   ```
   scli --mdm_ip <mdm IP> --query_all
   ```

2. Check if the volume is mapped to any of the SDC servers by running the `--query_all_volumes` command.
3. Determine if the SDC is installed on the server by typing the following command:

   ```
   rpm -qa | grep sdc
   ```

4. Determine if the SDC is installed on the server by typing the following command:

   ```
   esxcli software vib list|grep sdc
   ```

5. Determine if the SDC is connected to an MDM by typing the following command:

   ```
   scli --mdm_ip <mdm IP> --query_all_sdc
   ```

6. Ensure that the MDM management IP address is up and running.
7. On an application server, rescan for new volumes by performing the following:
   - Linux Type:

     ```
     /opt/emc/scaleio/sdc/bin/drv_cfg --rescan
     ```

   - ESX

     Rescan for new devices

Cannot log in to the Installation Manager after upgrade

If you upgrade the IM, and are then unable to log in, restart the EMC ScaleIO Gateway service.

Certificate error when installing SDC on Windows servers

Due to a Microsoft Windows limitation, when installing an SDC component on a Windows 2008 R2 server, a certificate error, similar to the following may be displayed:

To solve this issue, ensure that Microsoft Security Update KB3033929 is installed.

Enabling acceleration in a new node

Enable acceleration in a new SDS node in a ScaleIO system.

Before you begin

Ensure that you have:

- Admin-level username and password for accessing ScaleIO
- The RFcache package (`xcache`) for installation on a new SDS
- This task requires other ScaleIO procedures, that are described in other documents:
  - Installing an SDS in ScaleIO
  - Adding an SDS to the MDM

- Installing the RFcache package, using Installation Manager (IM)
- Accessing the SVM

Procedure

1. Install an SDS in the ScaleIO system and add it to the MDM.
2. Install the RFcache package on the SDS:

SDS OS	Description
ESXi	Access the SVM and then run the following command: `rpm -i EMC-ScaleIO-xcache-2.0-xxx.0.sles11.3.*SVM*.x86_64.rpm`
Linux	Install the RFcache package, using Installation Manager.
Windows	Install the RFcache package, using Installation Manager.

3. Identify the SSD and HDDs to be accelerated and used as cache devices, by running the following commands:

SDS OS	Description
ESXi	Run the esxcli commands or open the vSphere View GUI to check the device properties.
Linux	Run the following commands: a. Run the command `fdisk -l` and check the size of the device in the output. b. Run the command `smartctl` and identify the type of device in the output.
Windows	Run any of the following commands: • `wmic diskdrive list brief /format:list` and check the device properties in the output. • `wmic diskdrive get name,size,model` and check the device name, size, and model number in the output.

4. Log in to the Master MDM.
5. Add an SSD device to the default cache pool, by running the following command:

```
scli --add_sds_rfcache_device (--sds_id <ID> |--sds_name
<NAME> | --sds_ip <IP> [--sds_port <PORT>])
--rfcache_device_path <PATH> [--rfcache_device_name <NAME>]
```

Example:

```
scli --add_sds_rfcache_device --sds_name SDS1 --
rfcache_device_path /dev/sdb --rfcache_device_name
rfcache1_SDS1_sdb
```

6. Repeat the previous steps, from the beginning of the procedure, for each new SDS device to be configured in the RFcache Storage Pool.

7. Enable caching on the Storage Pool in the Protection Domain, by running the following command:

```
scli --set_rfcache_usage (((--protection_domain_id <ID> | --
protection_domain_name <NAME>) --storage_pool_name <NAME>) |
--storage_pool_id <ID>) [--use_rfcache | --dont_use_rfcache]
```

Example:

```
scli --set_rfcache_usage --protection_domain_name PD1 --
storage_pool_name SP1 --use_rfcache
```

8. Verify the status of acceleration per SDS, by running the following command:

```
scli --query_sds (--sds_id <ID> | --sds_name <NAME> |--sds_ip
<IP>)
```

Example:

```
scli --query_sds --sds_ip 192.168.1.6
```

Output, similar to the following, appears, for different command types:

- Command to verify the RFcache status:

```
scli --query_sds (--sds_id <ID> | --sds_name <NAME> |--
sds_ip <IP>)
```

Example:

```
scli --query_sds --sds_ip 9.10.21.141
```

Output:

```
SDS 4eecc17300000002 Name: SDS_10.136.215.203 Version:
2.0.13000
Protection Domain: 928cd6e100000000, Name:
Default_Protection_Domain
DRL mode: Volatile
Authentication error: None
IP information (total 2 IPs):
         1: 9.10.21.141      Role: All (SDS and SDC)
         2: 9.10.121.141     Role: All (SDS and SDC)
        Port: 7072
RAM Read Cache information:
        128.0 MB (131072 KB) total size
        Cache is enabled
        RAM Read Cache memory allocation state is PENDING
Rfcache enabled
.
```

```
.
.

Rfcache device information (total 1 devices):
         1: Name: 50000396CC883494  Path: /dev/sdc  Original-
path: /dev/sdc  Size 762495MB  ID: e177c11700020000
```

- Command to verify storage pool status:

```
scli --query_storage_pool (((--protection_domain_id <ID> |
--protection_domain_name <NAME>) --storage_pool_name
<NAME>) | --storage_pool_id <ID>)
```

Example:

```
scli --query_storage_pool --storage_pool_name
Cached_Hdd_Pool_1 --protection_domain_name
Default_Protection_Domain
```

Output:

```
Storage Pool Cached_Hdd_Pool_1 (Id: ada1bb8500000000) has 0
volumes and 23.0 TB (23536 GB) available for volume allocation
        The number of parallel rebuild/rebalance jobs: 2
        Rebuild is enabled and using Limit-Concurrent-IO
policy with the following parameters:
        Number of concurrent IOs per device: 1
        Rebalance is enabled and using Favor-Application-IO
policy with the following parameters:
        Number of concurrent IOs per device: 1, Bandwidth
limit per device: 10240 KB per second
        Background device scanner: Disabled
        Zero padding is enabled
        Spare policy: 36% out of total
        Checksum mode: disabled
        Doesn't use RAM Read Cache
        Uses Flash Read Cache
```

9. Repeat this entire procedure on the rest of the SDS nodes.

Results

The new nodes are configured in the RFcache Storage Pool in ScaleIO.

S.M.A.R.T. hardware alerts are not displayed

Hardware alerts are not displayed in the ScaleIO GUI.

1. Ensure that the proper driver/disk utility is installed on the servers:
   - LSI: storcli
   - HP servers: hpssacli
   - Dell servers: perccli

   The CLI utility enables you to display the hardware alerts for the RAID controller and disk status.
2. Ensure that the LIA process is running on the server.
   LIA is used for alert display.
3. Ensure that the specific RAID controller is supported on the server.

Gateway server recovery

Describes the procedures for recovering a ScaleIO gateway when the physical host server fails:

- Recovering Gateway on a new server during upgrade on page 236
- Recovering IM with LIA trusted IP feature enabled on page 238
- SNMP configuration recovery post gateway crash during upgrade on page 238

Recovering Gateway on a new server during upgrade

Recover the ScaleIO Gateway (GW) on a new server and continue with the process of ScaleIO upgrade, when the GW server crashes during ScaleIO upgrade.

If the GW server fails during an upgrade, the Installation Manager (IM) will also go down. You must recover the GW first, and then the IM. This procedure will cause the MDM to use the new GW in place of the failed one.
To recover the Gateway on a new server, perform the following steps:

Procedure

1. Install the Gateway on a new server (identical to the one that crashed), using any of the following options:
   - Linux OS: Use the GW RPM file.
   - Windows OS: Use the GW MSI file.
2. If the LIA-trusted IPs feature was enabled during installation, you must perform the steps described in Recovering IM with LIA trusted IP feature enabled on page 238 before continuing with this procedure.
3. Open the `GatewayUser.properties` file, located at `/opt/emc/scaleio/gateway/webapps/ROOT/WEB-INF/classes/GatewayUser.properties`, and edit the IM parameters.

Parameter	Description	Action
`upgrade.mdm.data.ips`	Configures MDM data IPs	Enter comma-separated IP addresses for different cluster nodes. • If a single node has multiple IP addresses, then enter plus-sign-separated IPs. • Ensure that the sequence of items in the list matches the order of data, management (mgmt), and role for the same node. • Reserve the empty values for TieBreaker management (TB mgmt) IP addresses. Example: • `upgrade.mdm.data.ips` = 192.100.1.10,192.100.1.49,192.100.1.9,192.100.1.17,192.100.1.11 • `upgrade.mdm.mgmt.ips` = 10.76.60.41,10.76.60.48,10.76.60.41,,
`upgrade.mdm.mgmt.ips`	Configures MDM management IPs	Enter comma-separated IP addresses for different cluster nodes. • If a single node has multiple IP addresses, then enter plus-sign-separated IP addresses

Parameter	Description	Action
		• Ensure that the sequence of items in the list matches the order of data, management (mgmt), and role for the same node • Reserve the empty values for TieBreaker management (TB mgmt) IP addresses. Example: • `upgrade.mdm.data.ips` = 192.100.1.10,192.100.1.49,192.100.1.9,192.100.1.17,192.100.1.11 • `upgrade.mdm.mgmt.ips` = 10.76.60.41,10.76.60.48,10.76.60.41,,
`upgrade.mdm.role`	Configures MDM roles	Enter the role of the MDM server; mdm, slave, or tb (TieBreaker). The line of the MDM role should correlate to the MDM data IP and MDM management IP addresses. Example: `upgrade.mdm.role` = mdm,slave,slave,tb,tb
`upgrade.mdm.version.target`	Configures the target version of MDM	Enter the MDM target version to which you want to upgrade. The first item in the version.target must represent the first cluster node. Example: `upgrade.mdm.version.target` = 2.0-828.0,2.0-828.0,2.0-828.0
`upgrade.mdm.version.orig`	Configures the original version of MDMs	Enter the MDM's original version, from which you are upgrading. The first item in the version.orig must represent the first cluster node. Example: `upgrade.mdm.version.orig` = 1.32-255.0,1.32-255.0,1.32-255.0
`upgrade.mdm.actor.port`	Configures MDM actor ports	Enter comma-separated MDM actor ports, as an ordered list. Each entry is a relevant option, rollback or upgrade. Option selection relates to the operation, version upgrade or rollback that was being performed before the GW crashed. Example: `upgrade.mdm.status` = upgrade
`upgrade.sds.mm_list`	Removes the SDS from Maintenance mode	Change the status of the SDS, which was in maintenance mode before the GW server crash using the `upgrade.sds.mm_list` parameter. Example: `upgrade.sds.mm_list` = 10.76.60.32,10.76.60.33,10.76.60.34,10.76.60.35,10.76.60.36

Results

The IM parameters in the `GatewayUser.properties` file match with the scenario that existed before the GW server crash.

Recovering IM with LIA trusted IP feature enabled

Describes how to recover the Installation Manager (IM) when the LIA-trusted IP feature was enabled in the crashed IM.

When the IM crashes during system upgrade and the LIA trusted IPs feature was enabled with a faulty IM, you are required to update the LIA's white list with the new IM IPs.

Note

When this procedure is required and

To update the LIA's white list, perform the following steps on all nodes:

Procedure

1. Access the LIA node in the system.
2. For a server with a Linux-based OS environment, perform the following steps:
   a. Go to: `/opt/emc/scaleio/lia/bin/`
   b. Run the following one-line command:

```
./update_conf_bin 7 lia_trusted_ips
<Comma_separated_Scaleio_gw_IPs> 1
```

   Ensure that you enter all of the ScaleIO GW (IM) hosts' available IP addresses, even if those addresses are not used by ScaleIO.
   Example:

```
./update_conf_bin 7 lia_trusted_ips 1.2.3.4,5.6.7.8 1
```

3. For a server with a Windows-based OS environment, perform the following steps:
   a. Go to: `C:\Program Files\EMC\scaleio\sds\bin`
   b. Run the following one-line command:

```
update_conf.exe 7 lia_trusted_ips
<Comma_separated_Scaleio_gw_IPs> 1
```

   Ensure that you enter all of the ScaleIO GW (IM) hosts' available IPs, even if the IPs are not used by ScaleIO.
   Example:

```
update_conf.exe 7 lia_trusted_ips 1.2.3.4,5.6.7.8 1
```

SNMP configuration recovery post gateway crash during upgrade

Recover SNMP configuration when the gateway crashes during ScaleIO upgrade.

There are two methods of recovering SNMP configuration:

- Recovering SNMP configuration using Installation Manager on page 239
- Manually recovering SNMP configuration on page 239

Recovering SNMP configuration using Installation Manager

Recover SNMP configuration, when the gateway crashes during ScaleIO upgrade, using ScaleIO Installation Manager.

Procedure

1. Delete the alertservice user from the MDM.
2. Open the IM GUI and extend the ScaleIO system, using an updated CSV file.
3. Under **Advanced Options**, select **Enable alertservice** and enter the **TRAP receiver** parameters.
4. Complete the extend operation.

Manually recovering SNMP configuration

Recover SNMP configuration manually, when the gateway crashes during ScaleIO upgrade.

Procedure

1. Create a LockBox.
2. On the Gateway, edit the `gatewayUser.properties` file:

   For more information, see the *EMC ScaleIO User Guide*

   a. Change the value of `features.enable_snmp` to *true*.

   b. Enter IP addresses in `snmp.traps_receiver_ip`.

      ScaleIO supports up to two comma-separated TRAP receivers.

   c. In `snmp.sampling_frequency`, update the MDM sampling period (optional step).

   d. In `snmp.resend_frequency`, update the time period to re-send the already existing TRAPs (optional step).
3. Restart the scaleio-gateway service.

Installation Manager returns an error

In the Installation Manager, if an `HTTP 404 Webpage not available` error (or similar) is displayed, the IM may have been disabled.

For more information, refer to the “Configuring the Installation Manager” section in the *Deployment Guide*.

Installation with the Installation Manager fails

If the following error message appears when installing with the Installation Manager,

```
Command failed: Could not connect to <IP_address>
```

It could be that this node is not in an accessible network.

LIA upgrade fails in Windows 2008

When upgrading or installing the ScaleIO system on Windows 2008R2-based servers, the upgrade or installation may fail and generate this error message: `<package X> for OS windows is not signed`.

Due to an issue in Windows 2008, the certificate issuer might not be recognized. This will result in failure to upgrade or install the LIA MSI in Windows 2008.

To resolve the problem, use the following workaround:

Procedure

1. Download and extract the `roots.zip` root certificates package, located at https://www.symantec.com/theme/roots.
2. Manually install all the certificates in the VeriSign Root Certificates subfolders:
   a. Open a subfolder.
   b. Double-click a certificate, and select **Install Certificate**.
   c. Select **Place all the certificates in the following store**, and then browse to and select the **Trusted Root Certification Authorities** store.
   d. Click **OK**, then **Next**, and then **Finish** .
   e. Review and approve the security warning.

      A message appears, informing you that the import was successful.

   f. Repeat for each certificate.
3. Retry the upgrade or installation.

   The upgrade or installation should now succeed.

Older SDCs cannot communicate with newer SDSs after upgrade to v2.x

After upgrading to v2.x, older SDCs may not be able to communicate with newer SDSs.

If the IP addresses of a v2.x SDS are defined in IPv6 only, older-generation SDCs using IPv4 will not be able to communicate with the SDS. This can cause I/O errors. This can occur when v1.32 SDCs are added to a v2.0 system, as v2.0 supports IPv6, whereas earlier versions do not.

To resolve the problem, if the SDCs are to remain pre-v2.0:

- Retain at least one IPv4 address on the SDS.
- Before removing all the IPv4 IP addresses from the SDSs, ensure that all SDCs are upgraded to v2.x and then rebooted .

Removing RFcache (the xcache package) from Windows servers

When using the Control Panel Add/Remove to remove xcache on a Windows server, if SDC is also installed on that server, you must remove SDC first, restart the machine, and then remove xcache.

Removing RFcache leftovers from the Windows OS registry

Run a batch file to remove RFcache leftovers from the Windows registry.

Uninstalling RFcache leaves some leftover entries in the Windows registry. This will cause errors the next time RFcache is installed. To remove the leftover entries in the registry, use the `Clean_XC_registry.bat` file included in the ScaleIO Windows download.

Procedure

1. Locate the `Clean_XC_registry.bat` file that is included in the ScaleIO Windows download.
2. From the Windows command prompt, run the `Clean_XC_registry.bat` file.

Replacing a faulty caching device in ScaleIO

Replace a faulty caching device in a ScaleIO system.

Before you begin

Ensure that you have:

- Administrator authentication credentials to run ScaleIO SCLI commands
- Access to the ScaleIO nodes
- Any of the following parameters:
  - SDS name
  - SDS IP address
  - SDS ID

This procedure applies to ScaleIO software-only systems. It does not apply to AMS-supported ScaleIO systems.

Procedure

1. Log in to the scli interface.
2. Identify the failed RFcache device:

```
scli --query_sds --sds_name <NAME> | --sds_ip <IP> |--sds_id
<ID>
```

Output, similar to the following, appears for the RFcache devices:

```
Rfcache device information (total 1 devices):
         1: Name: N/A  Path: /dev/sdi  Original-path: /dev/sdi
Size 952719MB  ID: dbfb70fc00000000
```

For a faulty device, the command output displays the fault (error) at the end of the output, while the output for a healthy device does not.

3. Remove the faulty RFcache device from the Storage Pool:

```
scli --remove_sds_rfcache_device (--rfcache_device_id <ID> |
((--sds_id <ID> | --sds_name <NAME> | --sds_ip <IP> [--
sds_port <PORT>]) (--rfcache_device_name <NAME> | --
rfcache_device_path <PATH>)))
```

Example:

```
scli --remove_sds_rfcache_device --rfcache_device_id
7cf1ce1f00000000
```

4. Physically replace the faulty device, using the relevant system and vendor guidelines.
5. Add a new device:

```
scli --add_sds_rfcache_device (--sds_id <ID> |--sds_name
<NAME> | --sds_ip <IP> [--sds_port <PORT>])
--rfcache_device_path <PATH> [--rfcache_device_name <NAME>]
```

Example:

```
scli --add_sds_rfcache_device --sds_name SDS1 --
rfcache_device_path /dev/sdb --rfcache_device_name
rfcache1_SDS1_sdb
```

6. Reset the device error counter:

```
scli --clear_sds_rfcache_error (--sds_id <ID> |--sds_name
<NAME> | --sds_ip <IP> [--sds_port <PORT>])
```

7. Verify the status of acceleration per SDS:

```
scli --query_sds --sds_name <NAME> | --sds_ip <IP> |--sds_id
<ID>
```

Example:

```
scli --query_sds --sds_name SDS1
```

The command output should display the device state without errors (that is, the replaced device in healthy state); otherwise, repeat the procedure.

Results

The faulty caching device is replaced in the ScaleIO system.

Replacing a faulty cached storage device in ScaleIO

Replace a faulty cached storage device in a ScaleIO system.

Before you begin

Ensure that you have:

- Administrator authentication credentials to run ScaleIO SCLI commands
- Access to the ScaleIO nodes
- Any one of the following parameters:
  - SDS name
  - SDS IP address
  - SDS ID

This procedure applies to ScaleIO software-only systems. It does not apply to AMS-supported ScaleIO systems.

Procedure

1. Log in to the scli interface.
2. Identify the failed cached storage HDD device:

```
scli --query_sds --sds_name <NAME> | --sds_ip <IP> |--sds_id
<ID>
```

Output, similar to the following, appears:

```
Device information (total 7 devices):
         1: Name: N/A  Path: /dev/sdb  Original-path: /dev/sdb
ID: ffcfa55700000000
                Storage Pool: sp1, Capacity: 929 GB Error-
fixes: 0 scanned 0 MB, Compare errors: 0 State: Normal
         2: Name: N/A  Path: /dev/sdc  Original-path: /dev/sdc
ID: ffcfa55800000001
                Storage Pool: sp1, Capacity: 929 GB Error-
fixes: 0 scanned 0 MB, Compare errors: 0 State: Normal
         3: Name: N/A  Path: /dev/sdd  Original-path: /dev/sdd
ID: ffcfa55900000002
                Storage Pool: sp1, Capacity: 929 GB Error-
fixes: 0 scanned 0 MB, Compare errors: 0 State: Normal
         4: Name: N/A  Path: /dev/sde  Original-path: /dev/sde
ID: ffcfa55a00000003
                Storage Pool: sp2, Capacity: 929 GB Error-
fixes: 0 scanned 0 MB, Compare errors: 0 State: Normal
         5: Name: N/A  Path: /dev/sdf  Original-path: /dev/sdf
ID: ffcfa55b00000004
                Storage Pool: sp2, Capacity: 929 GB Error-
fixes: 0 scanned 0 MB, Compare errors: 0 State: Normal
         6: Name: N/A  Path: /dev/sdg  Original-path: /dev/sdg
ID: ffcfa55c00000005
                Storage Pool: sp2, Capacity: 929 GB Error-
fixes: 0 scanned 0 MB, Compare errors: 0 State: Normal
         7: Name: N/A  Path: /dev/sdh  Original-path: /dev/sdh
ID: ffcfa55d00000006
                Storage Pool: sp3, Capacity: 929 GB Error-
fixes: 0 scanned 0 MB, Compare errors: 0 State: Normal
```

For a faulty device, the command output displays the fault (error) at the end of the output, while the output for a healthy device does not.

3. Remove the faulty device from the Storage Pool:

```
scli --remove_sds_device (--device_id <ID> |((--sds_id <ID> |
--sds_name <NAME> | --sds_ip <IP> [--sds_port <PORT>]) (--
device_name <NAME> | --device_path <PATH>)))
```

4. Physically replace the faulty device, using the relevant system and vendor guidelines.
5. Add a new device, as a cached device, to the Storage Pool:

```
scli --add_sds_device (--sds_id <ID> | --sds_name <NAME> | --
sds_ip <IP> [--sds_port <PORT>])
--device_path <PATH> [--device_name <NAME>] (--
storage_pool_name <NAME>) | --storage_pool_id <ID>)
```

When the device is added to the Storage Pool, the device starts being cached.

6. Verify the status of the replaced device:

```
scli --query_sds --sds_name <NAME> | --sds_ip <IP> |--sds_id
<ID>
```

Example:

```
scli --query_sds --sds_name SDS1
```

The command output should display the device state without errors (that is, the replaced device in healthy state); otherwise, repeat the procedure.

Results

The faulty cached device is replaced in the ScaleIO system.

ScaleIO CLI or GUI cannot connect to an MDM

Perform the following steps:

Procedure

1. Ping the MDM IP address to ensure you have connectivity.
2. Ensure that you are connecting to the IP address of the Master MDM.

   If the MDM ownership has changed, try to connect to the IP address of the Slave MDM.
3. Check if the MDM is running, by typing the following command:

   ```
   ps -ef| grep mdm
   ```
4. Ensure that the management IP address is up and running.

SCLI add_sds command fails due to communication error or MDM crash

When you run the SCLI `add_sds` command, the command may fail, either due to communication error or the primary MDM going offline.

Normally, when you run `add_sds`, the SDS receives a request and returns acknowledgment (ACK) to the MDM. However, after you initiate an `add_sds` command, if the MDM goes down for any reason, the MDM may not receive the ACK message (because the MDM is offline), but the SDS presumes that it is attached to the MDM.

Thus, when the MDM does not receive the ACK message, retrying the `add_sds` command fails with status as `SDS is already attached to this MDM`.

To resolve this issue, it is recommended that you clean the SDS configuration by running the `add_sds` command with `force_clean` flag and then re-run `add_sds`.

SVM manual memory allocation

When using the plug-in for a clean deployment, SVM memory allocation is performed automatically. In the following cases, SVM memory allocation must be performed manually:

- Manual deployment on VMware.
- Extending an existing SVM with a new ScaleIO role/component, whether this is being done with the plug-in or manually.

  Workaround: Perform all the parts of step 1 and step 2 before extending the additional role/component on the SVM. Perform steps one SVM at a time.
- Changing the SDS performance profile, post deployment, or after an upgrade from v1.32.x to v2.0.

  Workaround: Perform all the parts of step 1 , one SVM at a time.
- Post backend upgrade from v1.32.x to v2.0.

  Workaround: Perform all the parts of step 1 and step 2 on all SVMs that were upgraded, in the following order, one SVM at a time.

Procedure

1. For SVMs that are SDS-only, perform the following:
   a. Move the SDS to maintenance mode (MM).
   b. Shut down the SVM.
   c. Increase SVM memory, according to the formula below.
   d. Power up the SVM.
   e. Exit MM.
2. For SVMs that are MDM (Master, Slave, or TB, may contain SDS, also):
   a. Start with Slaves and TBs:
      a. Move the SDS to maintenance mode (MM).
      b. Shut down the SVM.

c. Increase SVM memory, according to the formula below.
d. Power up the SVM.
e. Exit MM.

b. Proceed with the Master MDM:
a. Switch ownership, so the Master MDM is now a Slave MDM.
b. Move the SDS to maintenance mode (MM).
c. Shut down the SVM.
d. Increase SVM memory, according to the formula below.
e. Power up the SVM.
f. Exit MM.

Results

The memory allocation formula:

Component	Memory allocation rules		
Base SVM	• 350 MB		
MDM (Master/Slave)	• 470 MB + (500 KB * 8 TB of volume capacity) + (1.44 KB * number of volumes) + (4 KB * number of SDS devices) • Maximum supported volumes: 256 K		
Tie-Breaker MDM	• 50 MB		
SDS	• (Base) 536 MB + (RmCache Size) * 1.15 + (Storage capacity in TB) * 53 MB • For SDS high performance profile, we add 195 MB.		
SDC	• 132 KB + 23 MB * (number of MDMs) + 25 KB * (number of SDSs) + 1.5 KB * (number of volumes) + 16 B * (number of volume blocks) + 24 KB * (8 TB of volume capacity) • Volume blocks: 1 GB storage = 8 volume blocks		
RFcache	• 16 * (cache_size/page_size) • Commonly-used sizes:		
	RFcache page size	RFcache memory requirement, if the cache device is 800 GB	RFcache memory requirement, if the cache device is 1.6 TB
	64 K	200 MB	400 MB
	32 K	400 MB	800 MB
	16K	800 MB	1.6 GB
	8 K	1.6 GB	3.2 GB
	4 K	3.2 GB	6.4 GB

Viewing the status of volumes by drv_cfg --rescan command

View the status of volumes configured on a SDC node by running the `drv_cfg --rescan` command.

Before you begin

Ensure that the volumes are created and mapped to the SDC nodes in the ScaleIO system.

In a specific scenario in ScaleIO, volumes may not be visible in the ESX, Linux, or Windows OS, or you may want to view the status of volumes immediately, for any of the following reasons:

- Confirming changes in volume-related configuration
- Viewing volumes immediately
- Confirming inaccuracies viewed in the volume display (for reasons such as temporary communication-related issues between system components)

You may troubleshoot and view the status of volumes immediately by running the `drv_cfg --rescan` command.

Perform the following steps:

Procedure

1. Check if the volume is mapped to an SDC server.
2. Determine if the SDC is installed on the server and connected to an MDM.
3. Ensure that the MDM management IP address is up and running.
4. On an application server, re-scan the volumes, by running the following command:

   ```
   drv_cfg --rescan
   ```

   For additional information, see Application server does not see a ScaleIO volume on page 231.

Results

The up-to-date status of volumes is displayed.

Virtual IP feature is not functional

Configuring virtual IP addresses on an MDM cluster via vSphere plug-in may disable the virtual IP feature in the system.

It is recommended that you configure the virtual IP addresses on a physical system only via Installation Manager (IM). The virtual IP configuration using vSphere plug-in may lead to communication issues with SDCs, which may trigger IO errors and cause the virtual IP feature to fail.
However, if you have already used the vSphere plug-in instead, you must perform the following steps to resolve the issue:

Procedure

1. Remove the virtual IP addresses by running the following command:

```
scli --modify_cluster_virtual_ips
```

2. Remove the virtual IP interfaces by running the following command:

```
scli --modify_virtual_ip_interfaces
```

3. Edit the IP addresses in the `drv_cfg` file of each SDC in the system.
4. Reconfigure the virtual IP addresses, via IM.

Results

The virtual IP addresses are configured on the MDM cluster in the system, via IM, preventing the failure of the virtual IP feature.

Enabling acceleration in a new node

Enable acceleration in a new SDS node in a ScaleIO system.

Before you begin

Ensure that you have:

- Admin-level username and password for accessing ScaleIO
- The RFcache package (`xcache`) for installation on a new SDS
- This task requires other ScaleIO procedures, that are described in other documents:
  - Installing an SDS in ScaleIO
  - Adding an SDS to the MDM
  - Installing the RFcache package, using Installation Manager (IM)
  - Accessing the SVM

Procedure

1. Install an SDS in the ScaleIO system and add it to the MDM.
2. Install the RFcache package on the SDS:

SDS OS	Description
ESXi	Access the SVM and then run the following command: `rpm -i EMC-ScaleIO-xcache-2.0-xxx.0.sles11.3.*SVM*.x86_64.rpm`
Linux	Install the RFcache package, using Installation Manager.
Windows	Install the RFcache package, using Installation Manager.

3. Identify the SSD and HDDs to be accelerated and used as cache devices, by running the following commands:

SDS OS	Description
ESXi	Run the esxcli commands or open the vSphere View GUI to check the device properties.
Linux	Run the following commands: a. Run the command `fdisk -l` and check the size of the device in the output. b. Run the command `smartctl` and identify the type of device in the output.
Windows	Run any of the following commands: • `wmic diskdrive list brief /format:list` and check the device properties in the output. • `wmic diskdrive get name,size,model` and check the device name, size, and model number in the output.

4. Log in to the Master MDM.

5. Add an SSD device to the default cache pool, by running the following command:

```
scli --add_sds_rfcache_device (--sds_id <ID> |--sds_name
<NAME> | --sds_ip <IP> [--sds_port <PORT>])
--rfcache_device_path <PATH> [--rfcache_device_name <NAME>]
```

Example:

```
scli --add_sds_rfcache_device --sds_name SDS1 --
rfcache_device_path /dev/sdb --rfcache_device_name
rfcache1_SDS1_sdb
```

6. Repeat the previous steps, from the beginning of the procedure, for each new SDS device to be configured in the RFcache Storage Pool.

7. Enable caching on the Storage Pool in the Protection Domain, by running the following command:

```
scli --set_rfcache_usage (((--protection_domain_id <ID> | --
protection_domain_name <NAME>) --storage_pool_name <NAME>) |
--storage_pool_id <ID>) [--use_rfcache | --dont_use_rfcache]
```

Example:

```
scli --set_rfcache_usage --protection_domain_name PD1 --
storage_pool_name SP1 --use_rfcache
```

8. Verify the status of acceleration per SDS, by running the following command:

```
scli --query_sds (--sds_id <ID> | --sds_name <NAME> |--sds_ip
<IP>)
```

Example:

```
scli --query_sds --sds_ip 192.168.1.6
```

Output, similar to the following, appears, for different command types:

- Command to verify the RFcache status:

```
scli --query_sds (--sds_id <ID> | --sds_name <NAME> |--
sds_ip <IP>)
```

Example:

```
scli --query_sds --sds_ip 9.10.21.141
```

Output:

```
SDS 4eecc17300000002 Name: SDS_10.136.215.203 Version:
2.0.13000
Protection Domain: 928cd6e100000000, Name:
Default_Protection_Domain
DRL mode: Volatile
Authentication error: None
IP information (total 2 IPs):
         1: 9.10.21.141      Role: All (SDS and SDC)
         2: 9.10.121.141     Role: All (SDS and SDC)
        Port: 7072
RAM Read Cache information:
        128.0 MB (131072 KB) total size
        Cache is enabled
        RAM Read Cache memory allocation state is PENDING
Rfcache enabled
.
.
.

Rfcache device information (total 1 devices):
         1: Name: 50000396CC883494  Path: /dev/sdc  Original-
path: /dev/sdc  Size 762495MB  ID: e177c11700020000
```

- Command to verify storage pool status:

```
scli --query_storage_pool (((--protection_domain_id <ID> |
--protection_domain_name <NAME>) --storage_pool_name
<NAME>) | --storage_pool_id <ID>)
```

Example:

```
scli --query_storage_pool --storage_pool_name
Cached_Hdd_Pool_1 --protection_domain_name
Default_Protection_Domain
```

Output:

```
Storage Pool Cached_Hdd_Pool_1 (Id: ada1bb8500000000) has 0
volumes and 23.0 TB (23536 GB) available for volume allocation
        The number of parallel rebuild/rebalance jobs: 2
        Rebuild is enabled and using Limit-Concurrent-IO
policy with the following parameters:
        Number of concurrent IOs per device: 1
        Rebalance is enabled and using Favor-Application-IO
policy with the following parameters:
        Number of concurrent IOs per device: 1, Bandwidth
limit per device: 10240 KB per second
        Background device scanner: Disabled
        Zero padding is enabled
        Spare policy: 36% out of total
        Checksum mode: disabled
        Doesn't use RAM Read Cache
        Uses Flash Read Cache
```

9. Repeat this entire procedure on the rest of the SDS nodes.

Results

The new nodes are configured in the RFcache Storage Pool in ScaleIO.

The VMware plug-in responds slowly

In an environment with a large-scale of volumes, the plug-in response may slow down. To solve this, increase the maximum memory size to 2GB and then restart the vSphere web client service.

- Windows

  In the configuration file, `C:\Program Files\VMware\Infrastructure\vSphereWebClient\server\bin\service\conf\wrapper.conf`, look for a string similar to `= -Xmx` (the line should also start with `wrapper.java.additional`), and change the value to 2048M.

- Linux

  In the configuration file, `usr/lib/vmware-vsphere-client/server/wrapper/conf/wrapper.conf`, look for `wrapper.java.maxmemory=` and change the value to 2048.

  The `wrapper.java.maxmemory` parameter may not exist in vCenter 6.0 `wrapper.conf`, If this is the case, add it manually:

  `wrapper.java.maxmemory=2048`

Solving ScaleIO performance issues

ScaleIO is designed to generate the best performance possible from any given system configuration, by using all possible nodes and distributing the data evenly among them.

When the system performance does not meet your expectations, verify if:

- The relevant volume is allocated to a high-performance Storage Pool.
  For example, is it allocated to a pool consisting of SSDs only? If not, using such a Storage Pool will generate better performance.
- The relevant volume resides in a Storage Pool that consists of different storage drives, with different performances?
  A low performance drive in the pool will slow down all the members (waiting for it to respond). If possible, avoid mixing different types of drives.

- The network in use provides maximal network bandwidth to all the ports in use by ScaleIO.

For a full performance review, use the following resources:

- For all ScaleIO-related products, see *ScaleIO Performance Tuning Best Practice Guide*.
- For ScaleIO software only, see *System Analysis Best Practice Guide*.

After following these suggestions, you may contact EMC Support for professional analysis and assistance.

Mismatch in IO counters

Resolving mismatch in the IO counters of ScaleIO Dashboard and customer application.

The IO counter in ScaleIO Dashboard might display a read/write IO value which is larger than the one displayed in the customer application, although the total bandwidth in both the cases is the same.

The mismatch between the IO counters occurs when the customer operating system, which triggers the IO requests to the SDS, splits them into two or more smaller-sized IO requests; the total bandwidth being the same as the bandwidth of the original IO request. Thus the IOs are split at the level of the customer operating system (which causes ScaleIO to see a larger number of IOs), and are integrated as well at the same level (the level that splits them).

Currently, ScaleIO does not support any workaround for such scenarios.

Deploying SVM on a node with a management IP address from a different subnet than the node with the SVM template

Due to a VMware limitation, it is not possible to deploy an SVM on a node with a management IP address that is from a different subnet than the node that hosts the SVM template.

When it is necessary to deploy an SVM on a node with a management IP address that is from a different subnet than the node hosting the SVM template, use one of the following workarounds:

- Add the management IP address from the same SVM subnet to the node that hosts the SVM template, and add the new SVM to an existing system.
- Upload the SVM template to the node that will host the new SVM, and add that node to an existing system. Use the template you uploaded to deploy the SVM.

Speeding up rebuild and rebalance processes

You can set concurrent activity limits to speed up rebuild and rebalance processes.

You can speed up rebuild and rebalance processes by using the `scli --set_rebuild_rebalance_parallelism` command to increase the amount of concurrent activities. This can be very useful If you are planning maintenance, have a DR situation, or must have faster rebuild/rebalance times.

Check the service-level agreement (SLA) of your environment/application to ensure that the selected setting does not adversely affect your applications/clients. Test your environment to determine what the optimal "limit" is.

The following example sets concurrent activity to 10:

```
scli --set_rebuild_rebalance_parallelism
--protection_domain_name default --storage_pool_name default
--limit 10 --mdm_ip 10.13.168.138
```

For more information, see the *CLI Reference Guide*.

RFcache is not installed on an upgraded Windows server

When using the Installation Manager (IM) to upgrade ScaleIO from v2.x to v2.0.1 on a Windows server, the LIA removes the old XC file (`rfcache`) and then reboots the host. After the reboot, the LIA fails to install the new XC. This error is displayed in the IM.
To avoid this, before starting the upgrade process from the Installation Manager, perform the following steps:

Note

If the SDC is also installed on the Windows server, you must first remove the SDC, restart the machine, and then remove xcache.

Procedure

1. Log in to the Master MDM, either directly, or via SSH/RDP.
2. On each relevant Storage Pool, disable RFcache:

   ```
   scli --set_rfcache_usage --protection_domain_name <PD_NAME> --
   storage_pool_name <SP_NAME> --dont_use_rfcache
   ```

3. Remove each RFcache device:

   ```
   scli --remove_sds_rfcache_device --sds_ip <SDS_IP> --
   rfcache_device_path <RFCACHE_PATH>
   ```

4. On all Windows hosts with XC, perform the following steps, one host at a time:
   a. From **Control Panel** > **Add/Remove Programs**, uninstall ScaleIO XCACHE.
   b. When prompted, reboot the host.
   c. When the host has finished rebooting, clean the registry by running the `Clean_XC_registry.bat` script.

      The script is included in `Complete_Windows_SW_Download.xip`.
5. Upgrade the Gateway.

   From the extracted download file, copy the ScaleIO Gateway MSI to the IM server and run the file: `EMC-ScaleIO-gateway-2.0-13000.X-x64.msi`.
6. After the Gateway is upgraded, when uploading packages to the Gateway, do not upload the XC package.

7. Start the upgrade, as described in the relevant section of the upgrade documentation.
8. Once the upgrade is complete, perform an extend operation to add back the XC (rfcache).

SSD devices are not recognized in the operating system

SSD devices are not recognized correctly by the ESXi host and are not displayed as SSDs in the vSphere plug-in.

Procedure

1. Select the appropriate method to identify the devices' type and to correspond it to the device ID.

Option	Description

DirectPath is not configured	Perform the following procedures on the ESXi host:
	a. List the devices and their types:
	`cd /opt/lsi/perccli/`
	`./perccli /c0/eall/sall show`
	The devices are displayed with their DG number. The devices' type is displayed in the **Med** column.
	b. Display the DG number and corresponding VD number:
	`./perccli /c0/vall show`
	c. Run the following command:
	`esxcli storage core path list`
	The VD number is listed in the **Target** column. Identify the row with the value in the **naa** column that matches the Target number.

DirectPath is configured	Perform the following procedures on the SVM:
	a. List the devices and their types:
	`/opt/MegaRAID/perccli/perccli64 /c0/eall/sall show`
	The devices are displayed with their DG number. The devices' type is displayed in the **Med** column.
	b. Display the DG number and corresponding VD number:
	`/opt/MegaRAID/perccli/perccli64 /c0/vall show`

Option	Description
	c. Run the following command: `ls -l /dev/disk/by-path/` Search for "pci-0000:02:00.0-scsi-0:2:*X*:0", where *X* is the VD number from the previous step.

ESX competing thread setting resets on server reboot

The setting for ESXi outstanding IO for competing worlds (threads) on a ScaleIO volume is not persistent, and needs to be applied again after every ESXi host reboot.

You can make the setting persistent by adding a script to the ESXi cron job.

For more information on using a cron job, see https://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1033346

Procedure

1. Add the following script on the ESXi host crontab:

```
/etc/rc.local.d/local.sh
#!/bin/sh

line_item=$(esxcli storage core device list|egrep '^((\w+\.)|
( )((Display Name:)|(Devfs Path:)|(Vendor:)|(Model:)|
(Revision:)|(SCSI Level:)|(No of outstanding IOs with
competing worlds:)))(.?)'|sed 's/: $/: Null/g;s/ //g;s/(.)
(worlds\: [0-9]*$)/\1\2\n\n/'|sed -r ':a;N;$!ba;s/\n
{1}

/,/g;s/(\,\,\,)/\n/g;s/\,\,//g;s/ /-/g')

for item in $
{line_item}

; do
new_item=$(echo $
{item}

|sed -r 's/-/ /g;s/(\w+(\s)?\w+)+\:\s//g')
device=$(echo $
{new_item}|cut -d"," -f1);
vendor=$(echo ${new_item}

|cut -d"," -f4);
req_outstanding=$(echo $
{new_item}

|cut -d"," -f8);
if ! [[ "$
{req_outstanding}

" == "256" ]]; then
if [[ "$
{vendor}

" == "EMC" ]]; then
esxcli storage core device set -d $
```

```
{device}

-O 256;
fi
fi
done
```

Results

The settings will be persistent.

CHAPTER 12

Frequently Asked Questions

This chapter describes frequently asked questions regarding ScaleIO. Topics include:

- Associating ScaleIO volumes with physical disks....258
- Changing default ports....260
- Adding an external SDC to an existing ScaleIO system....261
- Changing the LIA configuration file....263
- Cleaning the ScaleIO VMware environment and performing a clean install....264
- Configuring ScaleIO devices in Linux LVM....266
- Configuring session timeout parameters....266
- Fixing keytool errors....267
- Installing Java on SUSE 12 servers....267
- Mounting ScaleIO....268
- The ScaleIO Gateway web server isn't responding....269
- Upgrading the Gateway when a custom certificate is used....271
- Uploading a new OVA....271
- Using the same data network for different NICs....271
- What to do when the default self-signed certificate expires....271
- Xen v6.5 High Availability....271

Associating ScaleIO volumes with physical disks

This section describes how to associate volumes with physical disks.

Contact ScaleIO Customer Support for access to the troubleshooting utility.

To get ScaleIO volume information, run the `scli --query_all_volumes` (or `--query_all` or `--query_volume`) command.

Output similar to the following appears:

```
Query-all-volumes returned 10 volumes
Protection Domain 0728185d00000000 Name: pd1
Storage Pool ad99eaab00000000 Name: default
<No volumes defined>

Storage Pool ad99eaac00000001 Name: sp1
Volume ID: fac22a6300000000 Name: vol0 Size: 152.0 GB (155648 MB) Mapped to 1 SDC Thick-provisioned
Volume ID: fac22a6400000001 Name: vol1 Size: 400.0 GB (409600 MB) Mapped to 1 SDC Thin-provisioned
Volume ID: fac22a6500000002 Name: vol2 Size: 80.0 GB (81920 MB) Mapped to 1 SDC Thick-provisioned
Volume ID: fac22a6600000003 Name: vol3 Size: 392.0 GB (401408 MB) Mapped to 1 SDC Thick-provisioned
Volume ID: fac22a6700000004 Name: vol4 Size: 96.0 GB (98304 MB) Mapped to 1 SDC Thin-provisioned
Volume ID: fac22a6800000005 Name: vol5 Size: 112.0 GB (114688 MB) Mapped to 1 SDC Thick-provisioned
Volume ID: fac22a6900000006 Name: vol6 Size: 96.0 GB (98304 MB) Mapped to 1 SDC Thin-provisioned
Volume ID: fac22a6a00000007 Name: vol7 Size: 176.0 GB (180224 MB) Mapped to 1 SDC Thick-provisioned
Volume ID: fac22a6b00000008 Name: vol8 Size: 272.0 GB (278528 MB) Mapped to 1 SDC Thick-provisioned
Volume ID: fac22a6c00000009 Name: vol9 Size: 360.0 GB (368640 MB) Mapped to 1 SDC Thin-provisioned
```

This output shows the Volume ID and name, as well as other volume information.

Volume information - Linux

On the SDC host, run the following command to get the operating system volume information that correlates to the ScaleIO scini device name:

```
ls -l /dev/disk/by-id/ |grep scini
```

Output, similar to the following appears:

```
lrwxrwxrwx 1 root root 12 Aug 25 19:40 emc-vol-62c093a52d14aec7-fac22a6300000000 -> ../../scinia
lrwxrwxrwx 1 root root 12 Aug 25 19:40 emc-vol-62c093a52d14aec7-fac22a6400000001 -> ../../scinic
lrwxrwxrwx 1 root root 12 Aug 25 19:40 emc-vol-62c093a52d14aec7-fac22a6500000002 -> ../../scinib
lrwxrwxrwx 1 root root 12 Aug 25 19:41 emc-vol-62c093a52d14aec7-fac22a6600000003 -> ../../scinie
lrwxrwxrwx 1 root root 12 Aug 25 19:41 emc-vol-62c093a52d14aec7-fac22a6700000004 -> ../../scinid
lrwxrwxrwx 1 root root 12 Aug 25 19:42 emc-vol-62c093a52d14aec7-fac22a6800000005 -> ../../scinif
lrwxrwxrwx 1 root root 12 Aug 25 19:42 emc-vol-62c093a52d14aec7-fac22a6900000006 -> ../../scinig
lrwxrwxrwx 1 root root 12 Aug 25 19:42 emc-vol-62c093a52d14aec7-fac22a6a00000007 -> ../../scinii
lrwxrwxrwx 1 root root 12 Aug 25 19:42 emc-vol-62c093a52d14aec7-fac22a6b00000008 -> ../../scinih
lrwxrwxrwx 1 root root 12 Aug 25 19:43 emc-vol-62c093a52d14aec7-fac22a6c00000009 -> ../../scinij
```

This output shows the scini volume name and the volume ID.

By matching the volume ID in both outputs, you can match the operating system names, sciniX, with the ScaleIO volume name.

For example:

- scinia = fac22a6300000000 = vol0
- scinic = fac22a6400000001 = vol1

Alternatively, run the `sg_inq /dev/sciniX` SCSI query command. The result of this command includes the EMC volume ID at the bottom of the output, as illustrated in the following figure:

```
Vendor identification: EMC
Product identification: ScaleIO
Product revision level: 1.3
Unit serial number: EMC-62c093a52d14aec7-fac22a6300000000
```

Note

The sg3_utils must be installed on the Linux host in order to run this command.

Volume information - Windows

The `sg_inq.exe` file was added to the MSI installation and can be found at `C:\Program Files\EMC\ScaleIO\SDC\diag\`.

1. Run the `sg_inq HardiskX` SCSI query command. The result of this command includes the EMC volume ID at the bottom of the output.
2. On the MDM, get the ScaleIO volume information:

   ```
   drv_cfg --query_vol
   ```

   Output similar to the following is displayed:

   ```
   Retrieved 5 volume(s)
   VOL-ID 6acb988100000000 MDM-ID 0b246c9a755ca3dd
   VOL-ID 6acb988200000001 MDM-ID 0b246c9a755ca3dd
   VOL-ID 6acb988300000002 MDM-ID 0b246c9a755ca3dd
   VOL-ID 6acb988400000003 MDM-ID 0b246c9a755ca3dd
   VOL-ID 6acb988500000004 MDM-ID 0b246c9a755ca3dd
   ```

3. On the SDC host, get the operating system volume information that correlates to the ScaleIO scini device name:

   ```
   #esxcli storage core device list|grep eui
   ```

   Output similar to the following appears:

   ```
   eui.0b246c9a755ca3dd6acb988400000003
     Display Name: EMC Fibre Channel Disk (eui.0b246c9a755ca3dd6acb988400000003)
     Devfs Path: /vmfs/devices/disks/eui.0b246c9a755ca3dd6acb988400000003
   eui.0b246c9a755ca3dd6acb988300000002
     Display Name: EMC Fibre Channel Disk (eui.0b246c9a755ca3dd6acb988300000002)
     Devfs Path: /vmfs/devices/disks/eui.0b246c9a755ca3dd6acb988300000002
   eui.0b246c9a755ca3dd6acb988200000001
     Display Name: EMC Fibre Channel Disk (eui.0b246c9a755ca3dd6acb988200000001)
     Devfs Path: /vmfs/devices/disks/eui.0b246c9a755ca3dd6acb988200000001
   eui.0b246c9a755ca3dd6acb988100000000
     Display Name: EMC Fibre Channel Disk (eui.0b246c9a755ca3dd6acb988100000000)
     Devfs Path: /vmfs/devices/disks/eui.0b246c9a755ca3dd6acb988100000000
   eui.0b246c9a755ca3dd6acb988500000004
     Display Name: EMC Fibre Channel Disk (eui.0b246c9a755ca3dd6acb988500000004)
     Devfs Path: /vmfs/devices/disks/eui.0b246c9a755ca3dd6acb988500000004
   ```

4. Match the last 16 characters of the EUI to the ScaleIO volume ID, and the first 16 characters to the MDM-ID.

Changing default ports

The following table lists the TCP ports that are used by ScaleIO. Prior to installing or upgrading a system, ensure that these ports are not in use by other processes.

If they are, either free these ports or change the ScaleIO ports to another available port.

Table 14 Default ports

Port used by	Port #	File to change	Field to modify (or to add, if it does not exist)	Notes
MDM listener	6611	**Note** Cannot be modified, and must be available		
MDM Cluster member	9011	`/opt/emc/scaleio/mdm/cfg/conf.txt`	`actor_cluster_port=`*<NEW_PORT>*	
SDS listener	7072*x*	`/opt/emc/scaleio/sds/cfg/conf.txt`	`tgt_port=`*<NEW_PORT>*	For cases where multiple SDSs are installed on the same physical server, use ports 7072+*x*, where *x* is the index of the SDS.
LIA listener	9099	`/opt/emc/scaleio/lia/cfg/conf.txt`	`lia_port=`*<NEW_PORT>*	
Gateway (secure)	443	{gateway installation directory} `/conf/catalina.properties`	`http.port=443`	
Gateway (not secure)	80	{gateway installation directory} `/conf/catalina.properties`	`ssl.port=80`	After changing either of the ports, you must restart the service/daemon, as described in The ScaleIO Gateway web server isn't responding on page 269.
SDBG for MDM (Manager)	25620			These ports are used by ScaleIO internal debugging tools to extract live information from the system for debugging purposes.
SDBG for MDM (Tie-Breaker)	25600			
SDBG for SDS	25640			These ports are used by ScaleIO internal debugging tools to extract live

Table 14 Default ports (continued)

Port used by	Port #	File to change	Field to modify (or to add, if it does not exist)	Notes
				information from the system for debugging purposes.
SDBG for Multi-SDS	2564*x*			These ports are used by ScaleIO internal debugging tools to extract live information from the system for debugging purposes. For cases where multiple SDSs are installed on the same physical server, use ports 2564+*x*, where *x* is the index of the SDS.

Adding an external SDC to an existing ScaleIO system

During manual installation, you can install the SDC according to the OS-specific instructions below, and it will be connected to the existing ScaleIO system:

- Installing SDC on an ESX server and connecting it to ScaleIO on page 261
- Installing SDC on a Linux server and connecting it to ScaleIO on page 262
- Installing SDC on a Windows server and connecting it to ScaleIO on page 263

Installing SDC on an ESX server and connecting it to ScaleIO

Install the SDC with the appropriate parameters to connect it to an existing ScaleIO system.

Before you begin

Ensure that you have:

- The virtual IP address or MDM IP address of the existing system
- Login credentials for the SDC
- The appropriate installation packages for the SDC
- Access to the drv_cfg tool. Contact EMC support for access to this tool on ESX.

The following procedure describes installing an external SDC on an ESX server using the esxcli. Alternatively, you can install the external SDC using the vSphere plug-in. For more information, see "Installing the SDC on ESX hosts" in the *ScaleIO Deployment Guide*.

Procedure

1. On the ESX on which you are installing the SDC, set the acceptance level:

```
esxcli --server=<SERVER_NAME> software acceptance set --
level=PartnerSupported
```

where *<SERVER_NAME>* is the ESX on which you are installing the SDC.

2. Install the SDC:

```
esxcli software vib update -d "Full Path"
```

3. Set the IP address of the MDM:

```
esxcli system module parameters set -m scini -p
"IoctlIniGuidStr=<XXXXXX> IoctlMdmIPStr=<LIST_VIP_MDM_IPS>"
```

where

- *<LIST_VIP_MDM_IPS>* is a comma-separated list of the MDM IP addresses or the virtual IP address of the MDM
- *<XXXXXX>* is the version

Results

The SDC is installed on the ESX server and is connected to the ScaleIO system.

Installing SDC on a Linux server and connecting it to ScaleIO

Install the SDC with the appropriate parameters to connect it to an existing ScaleIO system.

Before you begin

Ensure that you have:

- The virtual IP address or MDM IP address of the existing system
- Login credentials for the SDC
- The appropriate installation packages for the SDC

The following procedure describes manually installing an external SDC on a Linux server. Alternatively, you can install the external SDC using the Installation Manager. For more information, see "Extending an existing ScaleIO system" in the *ScaleIO Deployment Guide*.

Procedure

1. Install the SDC:
   - RHEL/CentOS /Oracle Linux

     ```
     MDM_IP=<LIST_VIP_MDM_IPS> rpm -i <SDC_PATH>.rpm
     ```

   - Ubuntu

     ```
     MDM_IP=<LIST_VIP_MDM_IPS> dpkg -i <SDC_PATH>.deb
     ```

   - CoreOS

     ```
     MDM_IP=<LIST_VIP_MDM_IPS> ./<LIST_VIP_MDM_IPS>.bsx
     ```

   where

- *<LIST_VIP_MDM_IPS>* is a comma-separated list of the MDM IP addresses or the virtual IP address of the MDM
- *<SDC_PATH>* is the path where the SDC installation package is located

Results

The SDC is installed on the Linux server and is connected to the ScaleIO system.

Installing SDC on a Windows server and connecting it to ScaleIO

Install the SDC with the appropriate parameters to connect it to an existing ScaleIO system.

Before you begin

Ensure that you have:

- The virtual IP address or MDM IP address of the existing system
- Login credentials for the SDC
- The appropriate installation packages for the SDC

The following procedure describes manually installing an external SDC on a Windows server. Alternatively, you can install the external SDC using the Installation Manager. For more information, see "Extending an existing ScaleIO system" in the *ScaleIO Deployment Guide*.

Procedure

1. On the Windows server on which you are installing the SDC, run:

   ```
   msiexec /i <SDC_PATH>.msi MDM_IP=<LIST_VIP_MDM_IPS>
   ```

   where

   - *<SDC_PATH>* is the path where the SDC installation package is located
   - *<LIST_VIP_MDM_IPS>* is a comma-separated list of the MDM IP addresses or the virtual IP address of the MDM

Results

The SDC is installed on the Windows server and is connected to the ScaleIO system.

Changing the LIA configuration file

You can change the default behavior of the LIA by editing the configuration file:

- Windows: `C:\Program Files\emc\scaleio\LIA\cfg\conf.txt`
- Linux: `/opt/emc/scaleio/lia/cfg/conf.txt`

  The following are some values relevant to LIA behavior:

  ```
  lia_token=5
  lia_enable_install=1
  lia_enable_uninstall=1
  lia_enable_configure_fetch_logs=1
  ```

For example, to restrict which Gateway IP addresses can access the LIA, add those IP addresses to this line in the `conf.txt` file:

```
lia_trusted_ips=<IP_ADDRESS_1>,<IP_ADDRESS_2>
```

To set this during LIA installation, set the TRUSTED_IPS environment variable. For example:

```
TRUSTED_IPS=1.2.3.4,5.6.7.8 rpm -i lia.rpm
```

Cleaning the ScaleIO VMware environment and performing a clean install

This topic explains how to clean the ScaleIO VMware environment and perform a clean install while using previously defined networks.

Before you begin

Before you begin, unmap and delete any ScaleIO volumes in your system.

If necessary, unregister your ScaleIO system from within the plugin and delete all the ScaleIO SVMs.

Procedure

1. Set to **Run as administrator**, close the existing PowerCLI sessions and open a new one.
2. Using the PS1 script, unregister the plugin.
3. Stop the vSphere web client service:

   VC Linux: `service vsphere-client stop`
4. Delete the contents of the plug-in folder.

   The vSphere web client (Virgo) plug-in folders are located at:

vCenter	Operating system	Path to file
5.x	Windows	C:\ProgramData\VMware\vSphere Web Client\vc-packages\vsphere-client-serenity
	Linux	/var/lib/vmware/vsphere-client/vc-packages/vsphere-client-serenity/
6.x	Windows	C:\ProgramData\VMware\vCenterServer\cfg\vsphere-client\vc-packages\vsphere-client-serenity
	Linux	/etc/vmware/vsphere-client/vc-packages/vsphere-client-serenity

5. Delete the `scaleio` folder or its contents.

   The `scaleio` folders are located at:

vCenter	Operating system	Path to file
5.x	Windows	C:\Windows\System32\config\systemprofile\AppData\Roaming\VMware\scaleio
	Linux	/opt/.vmware/scaleio
6.x	Windows	C:\Users\vspherewebclientsvc\AppData\Roaming\VMware\scaleio
	Linux	/etc/vmware/vsphere-client/vc-packages/scaleio

6. Clean the Virgo logs folder.

   The Virgo log folders are located at:

vCenter	Operating system	Path to file
5.x	Windows	C:\ProgramData\VMware\vSphere Web Client\serviceability\logs
	Linux	/var/log/vmware/vsphere-client/
6.x	Windows	C:\ProgramData\VMWare\vCenterServer\logs\vsphere-client\logs
	Linux	/var/log/vmware/vsphere-client/logs

7. Start the vSphere web client service:

   VC Linux: `service vsphere-client start`

8. Clear your web browser's cache and cookies, or else open a different web browser.
9. Using the PS1 script, register the plugin via PowerCLI.

   Note

   Do not press ENTER at this point.

10. After you have logged in to the vSphere web client to complete the registration and you see the ScaleIO icon, press ENTER in the PowerCLI session.

    This stops the embedded Tomcat server.

11. If necessary, remove the SDC module parameters and VIB from the ESXs:

    a. Connect via SSH to each ESX.

    b. Run:

```
~ # esxcli system module parameters set -m scini -p ""
~ # esxcli software vib remove -n scaleio-sdc-esx5.5 / 6.0
```

    c. Reboot each ESX.

Configuring ScaleIO devices in Linux LVM

To configure devices, perform the following:

Procedure

1. Edit the `/etc/lvm/lvm.conf` file by adding the following line:

```
types = [ "scini", 16 ]
```

2. If only scini devices are to be used, you can add the following filter:

```
filter = [ "a|/dev/scini*|", "r/.*/" ]
```

3. Once configured, the `lvmdiskscan` command should yield results similar to the following:

```
/dev/scinia  [      96.00 GiB] LVM physical volume
/dev/scinib  [     320.00 GiB] LVM physical volume
/dev/scinic1 [      56.00 GiB]
/dev/scinid  [      32.00 GiB]
1 disk
1 partition
2 LVM physical volume whole disks
0 LVM physical volumes
```

4. Continue with normal LVM steps.

Configuring session timeout parameters

When a user is authenticated by the system, all commands are performed with the user's respective role until a logout is performed, or until the session expires by reaching one of the following timeouts:

- Maximum session length (default: 8 hours)
- Session idle time (default: 10 minutes)

You can modify these parameters, by editing the MDM `conf.txt` file:

- Linux: `/opt/emc/scaleio/mdm/cfg/conf.txt`
- Windows: `C:\Program Files\emc\scaleio\mdm\cfg\conf.txt`

1. To configure maximum session length, edit the value of the `user_session_hard_timeout_secs` parameter. The minimum is 10 seconds, maximum 10 years, and default 8 hours.
2. To configure session idle time, edit the value of the `user_session_timeout_secs` parameter. The minimum is 10 seconds, maximum 3 months, default 10 minutes.
3. After changing the parameters, restart the MDM service (delete and create service) for the changes to take effect.
4. To ensure persistence after MDM restart, make these changes on every MDM.

Fixing keytool errors

Error during rpm installation command

Error message:

```
No keytool path was found. Please pass SIO_GW_KEYTOOL as an argument
to the rpm installation command.
```

If a message similar to this is displayed after executing the rpm command to install the Gateway, add the location of the `/bin/keytool` file on your server to the command.

Example:

```
SIO_GW_KEYTOOL=/usr/lib/jvm/java-1.6.0-openjdk-1.6.0.0.x86_64/jre
rpm -U <gateway_installation_file_name>.rpm
```

Error during rpm upgrade command

Error message:

```
No keytool path was found. Set the environment variable SIO_GW_KEYTOOL
```

If a message similar to this is displayed after executing the rpm command to upgrade the Gateway, add the location of the `/bin/keytool` file on your server to the command.

Example:

```
SIO_GW_KEYTOOL=/usr/java/default/bin/ rpm -U /tmp/EMC-ScaleIO-
gateway-1.32-363.0.x86_64.rpm
```

Installing Java on SUSE 12 servers

Installation of Java is different in SLES-based distributions because SLES uses update-alternatives commands. For SUSE, we use a TGZ file in place of RPM.

To install Java on SUSE 12 servers:

Procedure

1. Untar the TGZ (for example, `jre-8u60-linux-x64.tar.gz`) to `/usr/java`.

   This creates a directory of `/usr/java/jre1.8.0_60/`.

2. Apply the std update-alternatives procedure:

   ```
   /usr/sbin/update-alternatives --install "/usr/bin/java"
   "java" "/usr/java/jre1.8.0_60/bin/java" 40
   /usr/sbin/update-alternatives --config java
   /usr/sbin/update-alternatives --install "/usr/bin/keytool”
   "keytool" "/usr/java/jre1.8.0_60/bin/keytool" 40
   /usr/sbin/update-alternatives --config keytool
   ```

Mounting ScaleIO

The exposed volumes are connected to the servers via the network. To configure mounting options of devices, follow the instructions for your OS.

Use persistent device names, described in full in Associating ScaleIO volumes with physical disks on page 258.

To mount :

Procedure

1. Determine the `/dev/disk/by-id` correlation to `/dev/sciniX`:

```
ls -l /dev/disk/by-id/ |grep scini
```

Output similar to the following appears:

```
lrwxrwxrwx 1 root root 12 Mar 2 05:35 emc-
vol-7ec27ef55b8f2108-85a0f0330000000a -> ../../scinia
lrwxrwxrwx 1 root root 12 Mar 2 05:35 emc-
vol-7ec27ef55b8f2108-85a0f03200000009 -> ../../scinib
lrwxrwxrwx 1 root root 12 Mar 2 05:35 emc-
vol-7ec27ef55b8f2108-85a0f02c00000003 -> ../../scinic
```

2. Run the mount command:

```
mount /dev/disk/by-id/<EMC-vol-id>
```

Example:

```
mount /dev/disk/by-id/emc-
vol-7ec27ef55b8f2108-85a0f0330000000a /mnt_scinia
```

3. To make the mount command persistent, edit the `/etc/fstab` file according to the instructions for your operating system:
   - RHEL 6.x:
     a. In `/etc/fstab`, use a text editor to add the ScaleIO mount lines:

```
/dev/disk/by-id/emc-
vol-7ec27ef55b8f2108-85a0f0330000000a /mnt_scinia ext4
defaults 0 0
```

     b. In `/etc/rc.local`, use a text editor to add the mount commands:

```
mount /mnt_scinia
```

   - RHEL 7.x:
     In `/etc/fstab`, use a text editor to add `_netdev` to the mount lines.

Example:

```
/dev/disk/by-id/emc-vol-7ec27ef55b8f2108-85a0f0330000000a /
mnt_scinia ext4 defaults,_netdev 0 0
```

Ensure that you comply with the `netdev` and syntax rules for your file system, as described in the `man` page.

- SLES:
  In `/etc/fstab`, use a text editor to add `nofail` to the ScaleIO Ready Node mount lines.

  Example:

```
/dev/disk/by-id/emc-vol-7ec27ef55b8f2108-85a0f0330000000a /
mnt_scinia ext3 nofail 0 0
```

  Ensure that you comply with the `nofail` and syntax rules for your file system, as described in the `man` page.

The ScaleIO Gateway web server isn't responding

The ScaleIO Gateway (REST service, Installation Manager) may be disabled:

The ScaleIO Gateway seems to be locked or disabled, and returns the HTTP status code 401 or 403.

Solution

- Ensure that the Gateway is enabled, as described in the documentation.
- In the `gatewayUser.properties` file, ensure that the `gateway-admin.password` property has a non-blank password. If the password is blank, the gateway has been locked.

  The following table shows the location of the gatewayUser.properties file:

Gateway installed on	Location of gatewayUser.properties file
Windows, 64-bit	C:\Program Files\EMC\ScaleIO\Gateway\webapps\ROOT\WEB-INF\classes\
Linux	/opt/emc/scaleio/gateway/webapps/ROOT/WEB-INF/classes

  To reset the Scaleio-Gateway password, perform the following steps:

Procedure

1. Use SioGWTool to reset the password by typing the following command:

```
SioGWTool --reset_password --password <new_scaleio-
gateway_password> --config_file
<path_to_file_gatewayUser.properties>
```

Note

The path to SioGWTool is:

Linux: `/opt/emc/scaleio/gateway/bin/SioGWTool.sh`

Windows: `C:\Program Files\EMC\ScaleIO\Gateway\bin\SioGWTool.bat`

2. Restart the scaleio-gateway service

The ScaleIO Gateway web server isn't responsive and the following error appears in the catalina log file:

- Windows:

  `C:\Program Files\EMC\ScaleIO\Gateway\logs\catalina.<date>.log`

- Linux:

  `/opt/emc/scaleio/gateway/logs`

```
2014-06-21 22:50:57,113 [main] ERROR
o.a.coyote.http11.Http11NioProtocol - Failed to initialize end
point associated with ProtocolHandler ["http-nio-443"]
java.net.BindException: Address already in use: bind
```

Solution

Perform one of the following:

Procedure

1. Find the service/daemon that is currently occupying that port and stop it:
   - Windows

     Run: netstat -anb
   - Linux

     Run: `netstat -alp`

   On Windows, one of the common applications that occupies this port is the VMware workstation, which uses this port for the shared VM feature. You can configure VMware workstation to use a different port via the Settings dialog, or you can disable the shared VM feature.

   Once the port is free, restart the scaleio-gateway service:

   - Windows

     Restart the EMC ScaleIO Gateway service.
   - Linux

     Type the command `service scaleio-gateway restart`
2. Change the ScaleIO Gateway web server to run on a different port, as described in "Changing default ports".

After doing so, restart the ScaleIO Gateway service/daemon, as described above. Access the Gateway with the new port. For example: https://<host>:<port>

Upgrading the Gateway when a custom certificate is used

If a custom security certificate is used on the Gateway (Windows and Linux environments), you must save a copy of the certificate (`*.keystore` file) and the `catalina.properties` file before you upgrade the gateway. After the upgrade is complete, you must copy these files back to their original location.

The default file locations, per operating system, are:

Linux:

`/opt/emc/scaleio/gateway/conf/catalina.properties`

`/opt/emc/scaleio/gateway/conf/certificates/.keystore`

Windows (64 bit):

`C:\Program Files\EMC\ScaleIO\Gateway\conf\catalina.properties`

`C:\Program Files\EMC\ScaleIO\Gateway\conf\certificates\.keystore`

Uploading a new OVA

If you have already used the OVA to create a template, you cannot create another template with the same name in the same datacenter.

Either remove the original template first, or use the `ScaleIOPluginSetup-2.0-13000.X.ps1` script, option #3, to assign a different name to the new template.

You can also upload the OVA manually using the VMware OVA upload tools. Configure the networks manually, after deployment or during the wizard menus. For more information, see the VMware user guides.

Using the same data network for different NICs

This configuration is supported, but it could reduce efficiency of outgoing communication and deny you the benefits of high availability of the multiple networks.

What to do when the default self-signed certificate expires

If the default self-signed security certificate is used on the Gateway, it expires after approximately one year. When you upgrade the gateway, the self-signed certificate is automatically replaced with a new one. If your self-signed security certificate expires, you can create a new one using the Java keytool utility.

Xen v6.5 High Availability

Enable use of volumes in HA, by enabling the volumes to be recognized as HBA.

Procedure

1. Install ScaleIO as described in the *EMC ScaleIO Deployment Guide*.
2. Modify the `lvm.conf` file, as described in the "Adding a volume" section of the *ScaleIO User Guide*, Xen appendix.
3. Create and map a ScaleIO volume to your Xen hosts.
4. From the Xen pool master, list the available storage:

   ```
   #ls -l /dev/disk/by-id/
   ```

   This will return an overview of the attached ScaleIO volumes, similar to the following:

   ```
   # scsi-"volume id" -> ../../scinia
   ```

5. From the Xen pool master, issue the Xen Storage Repository create command:

   ```
   # xe sr-create name-label="Any_name" content-type="ScaleIO"
   shared=true device-config:SCSIid="volume id" type=lvmohba
   ```

   where `volume id` is the value of volume id identified in the `scsi-"volume id"` output in the previous step.

Results

Your ScaleIO volume should now appear on your Xen hosts as a Storage Repository (SR).

GLOSSARY

A

Active Directory Active Directory (AD) provides directory-based identity-related services. It maintains a directory that is used to centrally store identity information and security principles, and uses them to authenticate and authorize users and devices.

Active Forward Rebuild A copy of stored data is currently being rebuilt on another server, due to planned or unplanned shutdown of a server.

B

Backward Rebuild Data is rebuilt on servers that went offline and became active again. Forward rebuilds can take a long time, and therefore, it can be quicker to restore and update the data on a server which has come back online, than it is to do an entire rebuild on a different server.

BWC Bandwidth counters.

C

Cache Random access electronic storage used to retain frequently used data for faster access by the channel.

ScaleIO can be configured to use server RAM cache, SSD devices, or PCIe devices in the system, to improve system performance.

CacheCade Read and Write caching of storage devices performed by one or more designated SSD devices in the system.

Cache Hit Rate The percentage of I/Os from cache.

Cache Skip Data is written directly to storage, bypassing the cache. Reasons for cache skips include: I/Os were too large, the cache device was busy, or I/Os were unaligned. The cache can also be configured to always work in passthrough mode.

Cache Writes Handling Mode The caching write-mode used by the system: passthrough mode (writes to storage only), or cached mode (by default, writes both to cache and to storage).

Cluster Mode is controlled by a cluster of MDM nodes, minimally consisting of a Master MDM, Slave MDM, and a Tie-Breaker node. 5-node clusters consist of one Master MDM, two Slave MDMs, and two Tie-Breakers.

D

Degraded Capacity The capacity is available, but is not protected in case of another failure

Device Physical storage device, such as a flash drive, or magnetic disk

DRL Dirty Region Logging: DRL bits indicate if data is in-writing to a certain location. Once the data is written in both primary and secondary locations, the DRL bit associated with the written location is cleared. These bits can be either stored in DRAM only (memory_only) or also backed up in non-volatile memory (hardened). The former delivers better I/O performance; the latter reduces data movement following a power-cycle giving rise to a faster rebuild.

F

Failed Capacity The capacity is inaccessible due to a failure, and data integrity is at risk

Fault Sets A logical entity that ensures that SDS data is backed up on SDSs that belong to other Fault Sets, thus preventing double-point-of-failure scenarios if rack power outages occur.

Forward Rebuild Data in storage will be rebuilt on another server, due to planned or unplanned shutdown of a server.

I

ID Identifier, a unique sequence of characters that identifies an object in the system. In some CLI commands, an ID can be used to specify a system component.

IP Role The role of the IP address configured for an SDS. Each SDS can have several IP addresses associated with it. Each IP address can serve a different purpose, or role. IP roles include: SDS, SDC, or both SDS and SDC.

L

LDAP The Lightweight Directory Access Protocol (LDAP) is a directory service protocol that runs on a layer above the TCP/IP stack. It provides a mechanism used to connect to, search, and modify Internet directories using Client-Server architecture. In ScaleIO, LDAP is the protocol used by the MDM to communicate with Active Directory (AD) for authentication purposes.

M

Management IPs The IP addresses of the MDMs defined in the system that can be used to access the MDM from CLI, GUI and REST.

Management Port The Port number used by the MDM for purposes of communicating with the nodes in the network.

Manager MDM An MDM that can act as a Master or a Slave in the cluster. Manager MDMs have a unique system ID, and can be given unique names. A manager can be a standby or a member of the cluster.

Master MDM The MDM in the cluster that controls the SDSs and SDCs.

MDM Any server with the MDM package installed on it. An MDM can be given a Manager or a TieBreaker (default) role, during installation. MDMs have a unique MDM ID, and can be given unique names.

P

Page Size The page size, typically in KB, used for caching purposes by Read Flash Cache.

Pass-Through Mode Data is passed through to or from storage devices without being cached by Read Flash Cache.

Pending Backward Rebuild A backward rebuild is waiting in a queue, and will be performed when possible, according to rebuild throttling policy.

Primary MDM See Master MDM.

Protected Capacity Capacity that has an accessible copy in the system, in case of failure.

Protection Domain A unique set of SDSs grouped together for reliability and tenancy separation.

R

RAM Read Cache (RMcache) Server RAM that is reserved for caching storage devices in a Storage Pool.

Read Flash Cache (RFcache) Read-only caching of storage devices performed by one or more designated SSD devices and PCIe flash devices in a system.

Rebalance When detects lopsided use of storage capacity, or when new nodes are added, it redistributes data across the nodes, in order to improve performance.

Rebuild When detects a failure in the network, it creates a new copy of the data from the failed component, in a new location, to ensure data integrity.

Restricted MDM Mode A mode set in which commands can only be performed from an MDM machine.

Restricted SDC Mode Only approved SDCs can access the MDM. When this mode is enabled, volumes can only be added to approved SDCs.

S

SDC Data Client, a lightweight device driver that exposes volumes as block devices to the application residing on the same server on which the SDC is installed.

SDS Data Server, which manages the capacity of a single server and acts as a back-end for data access. The SDS is installed on all servers contributing storage devices to the system.

Secondary MDM See See Slave MDM.

Single Mode A single MDM manages the network. This mode has no backup protection, and should not be used in production environments.

Slave MDM An MDM in the cluster that is ready to take over the Master MDM role if ever necessary.

Snapshot Capacity The amount of capacity occupied by snapshots of volumes.

Spare Capacity Capacity that is reserved for system use, when recovery from failure is required. This capacity cannot be used for storage purposes.

Spare Percentage Policy This policy determines the amount of capacity that must always be reserved as free space.

Standby MDM An MDM that has been locked to a specific ScaleIO system.

Storage Pool A sub-set of physical storage devices in a Protection Domain. Each storage device can only belong to one Storage Pool. User volumes will always use the storage of a single Storage Pool.

T

Thick Capacity Capacity allocated for thick volumes.

Thick Provisioned Volume Volume that has all its capacity pre-allocated on creation.

Thin Capacity Capacity allocated for thin volumes.

Thin Provisioned Volume Volume for which capacity is allocated on demand (by writing to the volume).

Throttling Throttling controls resource prioritization for rebuild and rebalance processes. Throttling can be controlled per Protection Domain or per Storage Pool (by configuring rebuild and rebalance policies).

Tie-Breaker The Tie-Breaker is used to determine which of the MDM nodes will take control over the ScaleIO system. As the number of MDMs is even, the Tie-Breaker ensures that there will always be one Master MDM achieving cluster quorum.

U

Unavailable Capacity Capacity that is not being used, but is also unavailable (due to server outage).

Unused Capacity Capacity that is not currently being used for any purpose in the system.

V

Volume A general term referring to a storage device. In the system, a volume consists of multiple blocks spread evenly on a storage pool devices.

W

Widget The full screen view can be minimized into a widget, which is a small window that floats on your screen, over other applications. Property sheets can also be minimized into widgets.

Write Misses Write requests that were not found in cache