

Virtual Provisioning for the EMC VNX2 Series

VNX5200, VNX5400, VNX5600, VNX5800, VNX7600, & VNX8000
Applied Technology

Abstract

This white paper discusses the benefits of Virtual Provisioning™ on the EMC® VNX2 series storage systems. It provides an overview of this technology and describes how Virtual Provisioning is implemented on the VNX2.

March 2016

Copyright © 2016 EMC Corporation. All Rights Reserved.

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

The information in this publication is provided “as is.” EMC Corporation makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on EMC.com.

Part Number H12204.6

Table of Contents

Executive summary	5
Audience.....	5
Terminology	5
Introduction	7
Business requirements	7
What is Virtual Provisioning?	8
Storage Pools	10
Pool Attributes	11
Oversubscribing a pool.....	12
Monitoring pool capacity.....	12
Expanding Pools.....	13
SAS Flash 3 drives in all-Flash Storage Pools	15
Pool LUNs	15
Thin LUNs.....	15
Thick LUNs	15
Thin LUNs vs. Thick LUNs.....	16
Creating pool LUNs.....	16
Expanding pool LUNs	18
When to use Classic, Thick, and Thin LUNs	18
Using Virtual Provisioning for VNX File	20
Thin LUNs for VNX File	20
VNX Block Deduplication for VNX File	20
File System and Checkpoint Storage Space Reclaim	20
Low Space Handling for VNX File	23
Out of Space Handling for VNX File	29
Thick and Thin-enabled File Systems	34
Creating thin-enabled File Systems.....	34
Monitoring thin-enabled File Systems.....	37
Conclusion	39
References	39
Appendix A: Thin LUNs	40
Thin LUN space reclamation via Migration	40
Using thin LUNs with applications	41
Host-based File Systems	41
VMware	42

Hyper-V	42
Appendix B: Low Space and Out of Space Handling Interoperability.....	43
Low Space Handling.....	43
Out of Space Handling.....	45

Executive summary

EMC® VNX Virtual Provisioning provides pool-based storage provisioning by implementing pool LUNs that can be either thin or thick. Thin LUNs provide on-demand storage that maximizes the utilization of your storage by allocating storage as it is needed. Thick LUNs provide high and predictable performance for your applications mainly because all of the user capacity is reserved and allocated upon creation. Both types of LUNs benefit from the ease-of-use features of pool-based provisioning. Pools and pool LUNs are also the building blocks for advanced data services such as Fully Automated Storage Tiering for Virtual Pools (FAST™ VP), Compression, and Deduplication.

Virtual Provisioning enables organizations to reduce storage costs by increasing capacity utilization, simplifying storage management, and reducing application downtime. Virtual Provisioning also helps companies to reduce power and cooling requirements and reduce capital expenditures.

Audience

This white paper is intended for IT planners, storage architects, administrators, and others involved in evaluating, managing, operating, or designing VNX storage systems.

Terminology

The following terminology appears in this white paper:

Allocated capacity — For a pool, this is the space currently used by all LUNs in the pool. For a thin LUN, this is the physical space used by the LUN. For a thick LUN, this is the host-visible capacity used by the LUN. Allocated capacity is slightly larger than the capacity used by the host because metadata exists at the pool LUN level. This is also known as “total allocation.”

Available capacity — The amount of actual physical pool space that is currently not allocated for pool LUNs.

Automatic Volume Management (AVM) — Feature of VNX that creates and manages File volumes automatically. AVM organizes volumes into Storage Pools for File that can be allocated to File Systems.

Classic LUN — A logical unit of storage created on a user-defined RAID group. The amount of physical space allocated is the same as the user capacity seen by the host server. Classic LUNs cannot be created on a pool; they are always created on a RAID group.

High water mark — Trigger point at which VNX performs one or more actions, such as sending a warning message or extending a File System, as directed by the related feature’s software/parameter settings.

Low space threshold — A value in GBs which defines when VNX File should take action on the VNX File resources to protect them from the underlying storage becoming full. The Low Space Threshold only applies to Block Storage Pools providing thin LUNs to VNX File.

LUN migration — A VNX feature that migrates data to another LUN with minimal disruption to running applications.

Mapped pool — A Storage Pool for File that is created during the storage discovery process for use on VNX for File. It is a one-to-one mapping with a VNX Storage Pool for Block.

Oversubscribed capacity — The amount of user capacity configured for pool LUNs that exceeds the physical capacity in a pool. Oversubscribing capacity is supported via thin LUNs.

Pool LUN — A logical unit of storage created in a pool. A pool LUN can be either a thin LUN or a thick LUN.

Pool Threshold alert — An alert issued when the % Full Threshold has been exceeded.

Slice — The minimum increment of capacity that can be allocated to a pool LUN. Pool LUNs are comprised of slices.

Storage Pool for Block — A group of drives for configuring pool LUNs (thick and thin).

Storage Pool for File — Groups of available File disk volumes organized by AVM that are used to allocate available storage to File Systems. They can be created automatically when using AVM or manually by the user.

Subscribed capacity — The total amount of capacity configured for LUNs in the pool. This value can be greater than the available user capacity. The available user capacity can be expanded by adding drives to the pool.

Thick LUN — A type of pool LUN in which allocated physical space is equal to the user capacity seen by the host server.

Thin-enabled File System — A File System that lets you allocate storage based on long-term projections, while you consume only the File System resources that you currently need. NFS or CIFS clients and applications see the virtual maximum size of the File System of which only a portion is physically allocated.

Thin-friendly — A term that is frequently used for File Systems and applications that do not pre-allocate all of the storage space during initialization. This term is also used for File Systems that reuse deleted space before consuming additional storage. Both of these features improve capacity utilization in thin provisioning.

Thin LUN — A type of pool LUN where physical space allocated can be less than the user capacity seen by the host server.

Total Capacity — The total amount of physical storage capacity in the pool that is available for pool LUNs. It is measured as raw disk capacity minus RAID and other overhead. This is also referred to as “usable capacity” or “host visible capacity.”

Volume — On VNX for File, a virtual disk into which a file system, database management system, or other application places data. A volume can be a single disk partition or multiple partitions on one or more physical drives.

% Full — The percentage of pool capacity that is currently consumed. It is calculated using this formula: consumed capacity / user capacity = % Full.

% Full Threshold — A user-configurable threshold that triggers the system to generate an alert when it is exceeded.

Introduction

One of the biggest challenges facing storage administrators is balancing the storage requirements for various competing applications in their data centers. Administrators are typically forced to allocate space, initially, based on anticipated storage growth. They do this to reduce the management expense and application downtime incurred when they need to add more storage as their business grows. This generally results in the over-provisioning of storage capacity, which then leads to higher costs; increased power, cooling, and floor space requirements; and lower capacity utilization rates. Even with careful planning, it may be necessary to provision additional storage in the future. This may require application downtime depending on the operating systems involved.

VNX Virtual Provisioning technology is designed to address these concerns. Thin LUNs and thin-enabled File Systems can present more storage to an application than is physically available. Storage managers are freed from the time-consuming administrative work of deciding how to allocate drive capacity. Instead, an array-based mapping service builds and maintains all of the storage structures based on a few high-level user inputs. Drives are grouped into storage pools that form the basis for provisioning actions and advanced data services. Physical storage from the pool can be automatically allocated only when required for maximum efficiency.

Business requirements

Organizations, both large and small, need to reduce the cost of managing their storage infrastructure while meeting rigorous service level requirements and accommodating explosive storage capacity growth.

Thin provisioning addresses several business objectives that have drawn increasing focus:

- **Reducing capital expenditures and ongoing costs**
Thin provisioning reduces capital costs by delivering storage capacity on actual demand instead of allocating storage capacity on anticipated demand. Ongoing costs are reduced because fewer drives consume less power and cooling, and less floor space.
- **Maximizing the utilization of storage assets**
Organizations need to accommodate growth by drawing more value from the same

or fewer storage resources. Operational efficiency remains an ongoing challenge, as organizations often over-allocate storage to applications to reduce the risk of outage and the need to re-provision later on.

- **Reducing the cost of storage administration**
“Ease-of-use” initiatives span multiple aspects of storage processes, including staff training, initial storage provisioning, the addition of new storage, and the management and monitoring of storage systems. Virtual provisioning simplifies the process of adding storage.

What is Virtual Provisioning?

Storage provisioning is the process of assigning storage resources to meet the capacity, availability, and performance needs of applications.

First, let’s take a look at how traditional storage provisioning works. With traditional block storage provisioning, you create a RAID group with a particular RAID protection level and a certain number of drives. RAID groups are restricted to a single drive type and to a maximum of 16 drives. When LUNs are bound on the RAID group, the host-reported capacity of the LUNs is equal to the amount of physical storage capacity allocated. The entire amount of physical storage capacity must be present on day one, resulting in low levels of utilization, and recovering underutilized space remains a challenge. **Figure 1** shows traditional storage provisioning where the physically allocated space is equal to the space that is reported to the application.

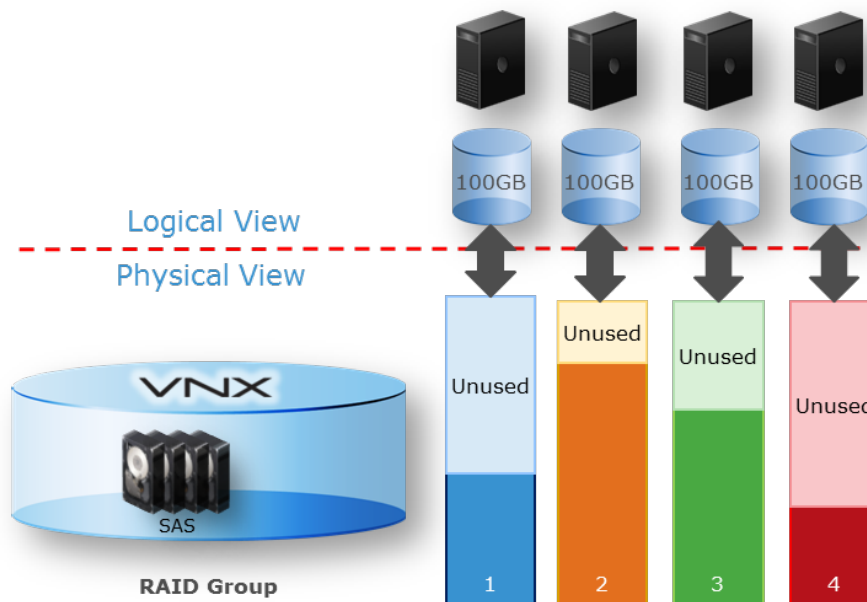


Figure 1 Traditional storage provisioning

With traditional provisioning, the storage administrator needs to carefully carve out the storage for an application based on the amount forecasted by the application administrator. There is a tendency for these forecasts to be inflated. In some companies, an application administrator may monitor storage space and ask the

storage administrator to provision additional storage. The storage administrator must rely on timely and accurate communications from various IT teams to effectively manage storage space utilization.

VNX Virtual Provisioning utilizes storage pool-based provisioning technology which is designed to save time, increase efficiency, reduce costs, and improve ease of use. Storage pools can be used to create thick and thin LUNs. Thick LUNs provide high and predictable performance for your applications mainly because all of the user capacity is reserved and allocated upon creation. **Figure 2** shows thick provisioning from a storage pool.

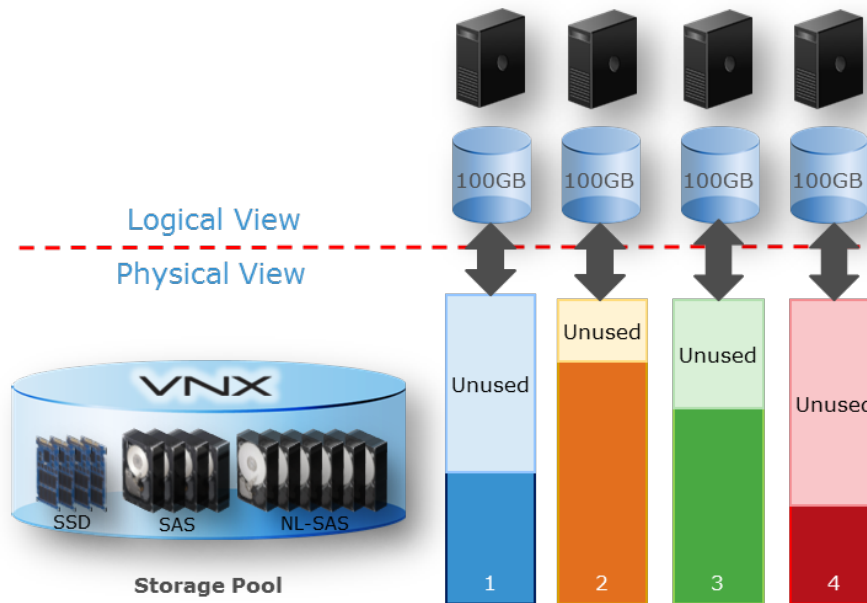


Figure 2 Thick Provisioning

With thin provisioning, the user capacity (storage perceived by the host) can be larger than the available capacity on the storage system. Thin LUNs can be sized to accommodate growth without regard for currently available assets. Physical storage is assigned to the server in a capacity-on-demand fashion from the shared pool. **Figure 3** shows thin provisioning where a certain capacity is reported to the application but only the consumed capacity is allocated from the Storage Pool.

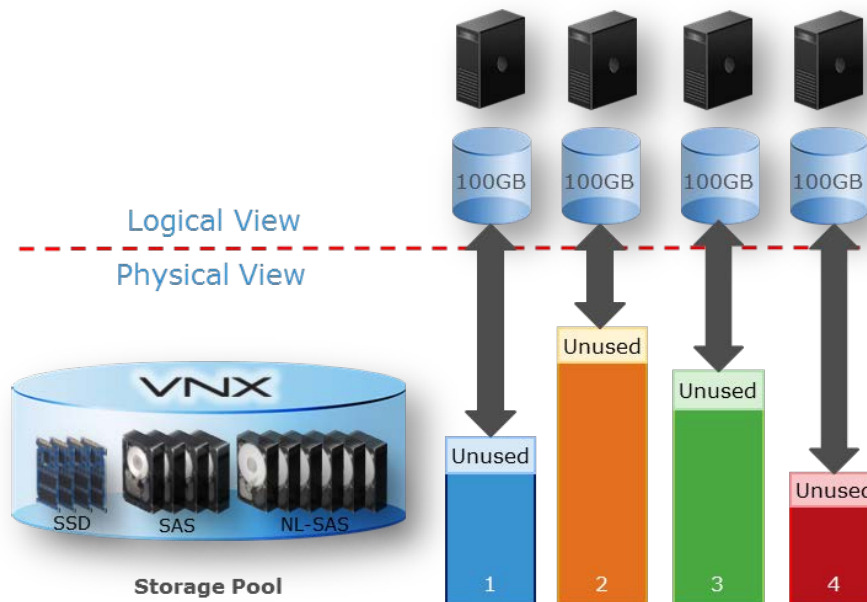


Figure 3 Thin Provisioning

Virtual Provisioning technology provides the flexibility of choosing between thick or thin LUNs based on the application requirements. It also supports existing VNX features such as hot sparing, proactive sparing, and the ability to migrate data between thin LUNs, thick LUNs, or classic LUNs without incurring application downtime. The ability to non-disruptively migrate data to different LUN and disk types provides the best solution for meeting your changing application and business requirements without incurring downtime.

Storage Pools

Virtual Provisioning utilizes Storage Pool technology. A pool is somewhat analogous to a RAID Group, which is a physical collection of drives on LUNs are created. But pools have several advantages over RAID Groups:

- Pools allow you to take advantage of advanced data services like FAST VP, Compression, and Deduplication
- Multiple drive types can be mixed into a pool to create multiple tiers with each tier having its own RAID configuration
- They can contain a few drives or hundreds of drives whereas RAID Groups are limited to 16 drives
- Because of the large number of drives supported in a pool, pool-based provisioning spreads workloads over many resources requiring minimal planning and management effort
- When selecting a number of drives that result in multiple RAID Groups, the system will automatically create the private RAID Groups. See [Table 1](#) for the preferred drive count options for each RAID type

Table 1 RAID Configuration Options

RAID Type	Preferred Drive Count Options
RAID 1/0	4+4
RAID 5	4+1, 8+1
RAID 6	6+2, 14+2

Pool Attributes

Pools are simple to create because they require few user inputs:

- Pool Name: For example, “Pool 0”
- Drives: Number and type
- RAID Protection level

Figure 4 shows an example of how to create a Storage Pool named Pool 1 with RAID 5 (4+1) protection for the SAS Flash 2 drives, RAID 5 (4+1) protection for the SAS drives, and RAID 6 (6+2) for the NL-SAS drives.

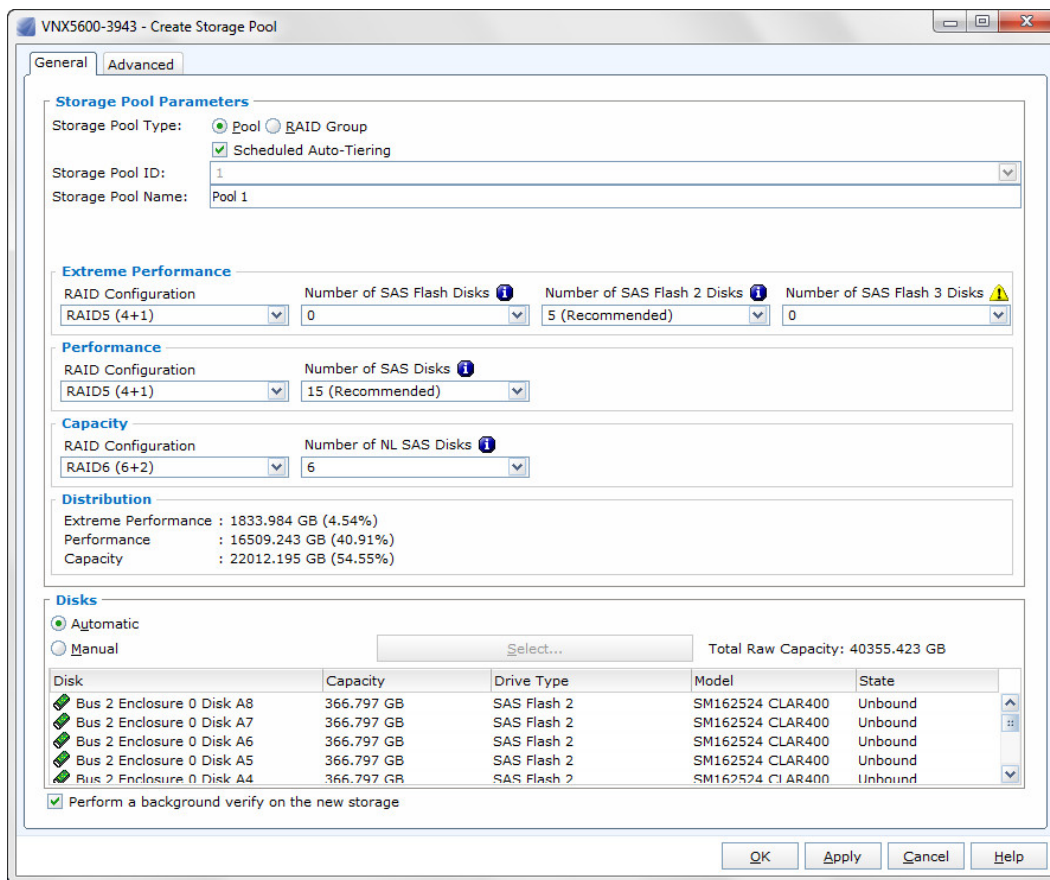


Figure 4 Create Storage Pool page

Users are advised to make every effort to ensure that pools are created with common drive capacities. Unisphere will display a warning when you attempt to create a pool with mixed-size disk types (Flash, SAS, or NL-SAS). However, different types of drives can have different sizes. To maximize space utilization, however, all drives of a

particular type should be the same size in each pool because it is possible that the larger drives of the same type will be truncated. If it becomes absolutely necessary to use different capacity drives to create a pool, create the pool in stages to avoid truncation. For example, if you have ten 600 GB SAS drives and five 300 GB SAS drives, first create the pool by selecting only the ten 600 GB drives, and then expand the pool by adding the five 300 GB drives.

Oversubscribing a pool

Thin provisioning allows you to oversubscribe a pool where capacity presented to the hosts exceeds the physical capacity in a pool. [Figure 5](#) shows an example of an oversubscribed pool.

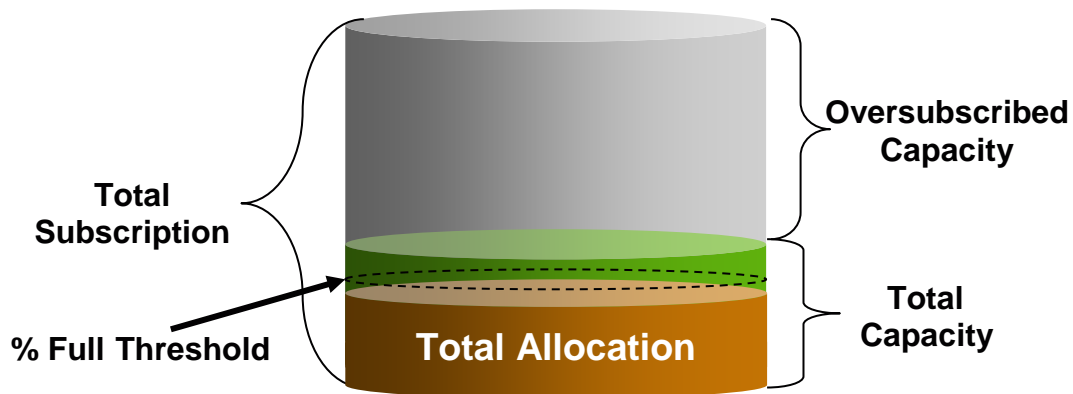


Figure 5 Pool capacity diagram

- Total Capacity is the amount of physical capacity available to all LUNs in the pool
- Total Allocation is the amount of physical capacity that is currently assigned to LUNs
- Total Subscription is the total capacity reported to the host
- Oversubscribed Capacity is the amount of capacity that exceeds the capacity in a pool

Monitoring pool capacity

Managing pool capacity is crucial when using thin provisioning to oversubscribe a pool. This ensures oversubscribed pools do not run out of space which may lead to write failures.

Pools are monitored using the % Full Threshold Alert. This alert is only active if there are one or more thin LUNs in a pool, since thin LUNs are the only way to oversubscribe a pool. If the pool only contains thick LUNs, the alert is not active as there is no risk of running out of space due to oversubscription. You can specify the value for % Full Threshold ($\text{Total Allocation} / \text{Total Capacity}$) when a pool is created or in the Advanced tab of the pool properties page.

Alerts can be monitored by using the Alerts tab in Unisphere. Using Unisphere's Event Monitor wizard, you can also select the option of receiving alerts through email, a paging service, or an SNMP trap.

Table 2 displays information about the user-configurable and built-in threshold alerts and their settings.

Table 2 Threshold alerts

Threshold Type	Threshold Type	Threshold Default	Alert Severity
User Configurable	50%-80%	70%	Warning
Built-in	N/A	85%	Critical

You can also monitor the Storage Pool capacity information in the Storage Pool Properties page in Unisphere as shown in Figure 6. This displays Physical and Virtual Capacity information such as:

- Total Capacity
- Free Capacity
- Percent Full
- Total Allocation
- Snapshot Allocation
- Total Subscription
- Snapshot Subscription
- Percent Subscribed
- Oversubscribed By (if applicable)

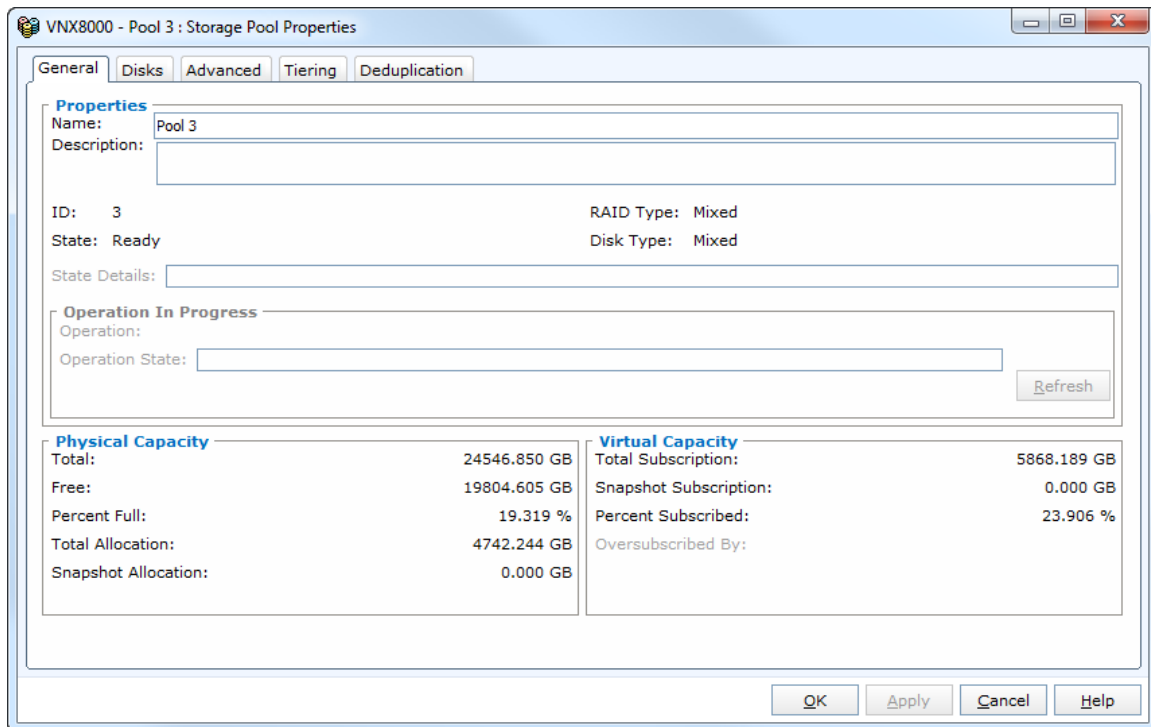


Figure 6 Storage Pool Properties page

Expanding Pools

Since pools can run out of space, it is a best practice to ensure that a monitoring strategy is in place and you have the appropriate resources to expand the pool when

necessary. Adding drives to a pool is a non-disruptive operation and the increased capacity can be immediately used by LUNs in the pool.

When a Storage Pool is expanded, the sudden introduction of new empty drives combined with relatively full, existing drives may cause a data imbalance. This imbalance is resolved by automatic one-time data relocation, referred to as a rebalance. This rebalance relocates a portion of the data to the new drives, based on capacity, to utilize the new spindles.

With Fully Automated Storage Tiering for Virtual Pools (FAST VP) enabled, this rebalance will take performance data into account and load balance across the new spindles. Refer to the *EMC VNX2 FAST VP – A Detailed Review* White Paper for more information on FAST VP.

Pools can be as large as the maximum number of drives (excluding vault drives and hot spares) allowed per system type. For example, a VNX7600 can contain 996 drives in a single pool or between all pools. Vault drives (the first four drives in a storage system) cannot be part of a pool, so Unisphere dialog boxes and wizards do not allow you to select these drives. Large pools must be created by using multiple operations. Depending on the system type, pools can be created by using the maximum allowed drive increment and then expanded until you reach the desired number of drives in a pool. Once the pool is fully initialized, you can create LUNs on it. For example, to create a 240-drive pool on a VNX5600, you need to create a pool with 120 drives and then expand the pool with another 120 drives. You can also expand pools at a later time if more storage is needed. The maximum allowed drive increments for different system types are shown in [Table 3](#).

Table 3 Drive increments for VNX models

VNX model	Maximum Drive Increments
VNX5200	80
VNX5400	80
VNX5600	120
VNX5800	120
VNX7600	120
VNX8000	180

Users need to be aware of fault domains when using large pools. A fault domain refers to data availability. A Virtual provisioning pool is made up of one or more private RAID groups. A pool's fault domain is a single-pool, private RAID group. That is, the availability of a pool is the availability of any single private RAID group. Unless RAID 6 is the pool's level of protection, avoid creating pools with very large numbers of RAID groups. For more information regarding the benefits of smaller pools, refer to the *EMC VNX Unified Best Practices for Performance – Applied Best Practices* white paper on EMC Online Support. The maximum pool and LUN limits for the available models are shown in [Table 4](#).

Table 4 Pool and LUN limits

Model	Max pools	Max disks per pool	Max disks per array	Max LUNs per pool/array
VNX5200	15	121	125	1000
VNX5400	20	246	250	1000
VNX5600	40	496	500	1100
VNX5800	40	746	750	2100
VNX7600	60	996	1000	3000
VNX8000	60	1496	1500	4000

SAS Flash 3 drives in all-Flash Storage Pools

In VNX OE for Block version 05.33.009.5.155 multiple sizes of SAS Flash 3 drives are introduced. SAS Flash 3 drives provide Flash performance at a reduce cost when compared to SAS Flash and SAS Flash 2 Flash drives. The SAS Flash 3 drives do not support FAST VP and are only supported in homogeneous all Flash Storage Pools with drives of the same Flash type.

Pool LUNs

A VNX pool LUN is similar to a classic LUN in many ways. Many of the same Unisphere operations and CLI commands can be used on pool LUNs and classic LUNs. Most user-oriented functions work the same way, including underlying data integrity features, LUN migration, local and remote protection, and LUN properties information.

Pool LUNs are comprised of a collection of slices and have the option to be thin or thick. A slice is a unit of capacity which is allocated from the private RAID Groups to the pool LUN when it needs additional storage. Starting with VNX Operating Environment (OE) for Block Release 33, the slice size has been reduced from 1 GB to 256 MB.

Thin LUNs

The primary difference between thin LUNs compared to Classic and thick LUNs is that thin LUNs have the ability to present more storage to an application than what is physically allocated. Presenting storage that is not physically available avoids underutilizing the storage system’s capacity.

Data and LUN metadata is written to thin LUNs in 8 KB chunks. Thin LUNs consume storage on an as-needed basis from the underlying pool. As new writes come into a thin LUN, more physical space is allocated in 256 MB slices.

Thick LUNs

Thick LUNs are also available in VNX. Unlike a thin LUN, a thick LUN’s capacity is fully reserved and allocated on creation so it will never run out of capacity. Users can also better control which tier the slices are initially written to. For example, as pools are initially being created and there is still sufficient space in the highest tier, users can

be assured that when they create a LUN with either “Highest Available Tier” or “Start High, then Auto-Tier”, data will be written to the highest tier – because the LUN is allocated immediately.

Thin LUNs vs. Thick LUNs

Thin LUNs typically have lower performance than thick LUNs because of the indirect addressing. Thus, the mapping overhead for a thick LUN is much less when compared to a thin LUN.

Thick LUNs have more predictable performance than thin LUNs because the slice allocation is assigned at creation. However, thick LUNs do not provide the flexibility of oversubscribing like a thin LUN does so they should be used for applications where performance is more important than space savings.

Thick and thin LUNs can share the same pool, allowing them to have the same ease-of-use and benefits of pool-based provisioning.

Creating pool LUNs

Starting with VNX Operating Environment (OE) for Block Release 33, thin is the default option when creating a new LUN in Unisphere.

As shown in [Figure 7](#), a pool LUN can be created by providing the following information:

- Storage Pool – The pool to create the LUN from
- Thin, not Thin, or Block Deduplication enabled
- User Capacity - Amount of host visible user capacity

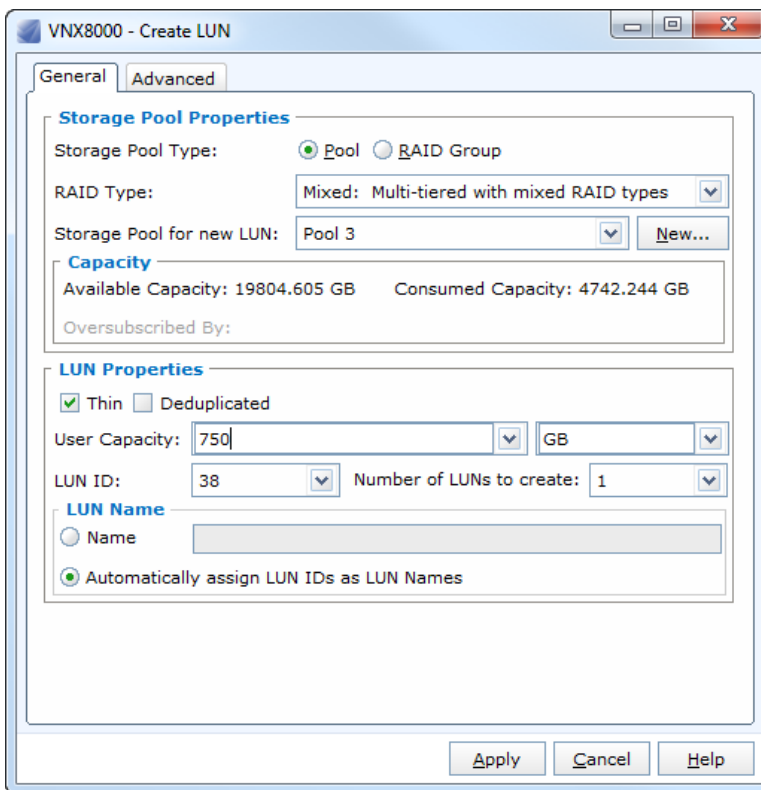


Figure 7 Create LUN page

The minimum LUN size is 1 MB and the maximum LUN size is 256 TB. However, the minimum consumed capacity for any LUN is 1.75 GB because of the 1.5 GB overhead and .25 GB of initial capacity allocation for incoming writes. [Figure 8](#) shows the 1.75 GB initial consumed capacity on a new 500 GB LUN.

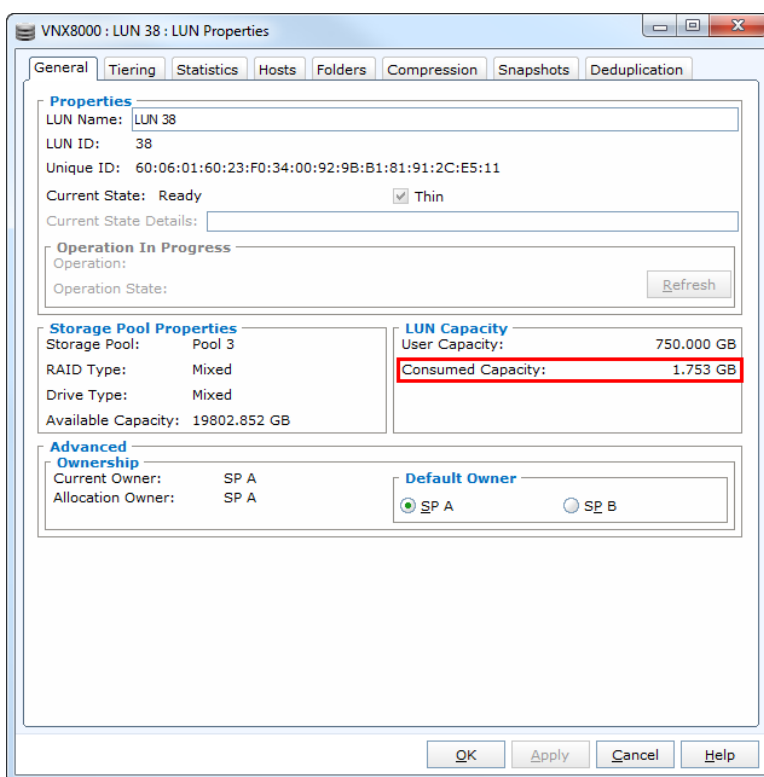


Figure 8 Thin LUN Properties page

Expanding pool LUNs

Virtual Provisioning also simplifies storage management tasks by providing the ability to expand pool LUNs with a few simple clicks. The underlying pools can also be expanded by adding drives non-disruptively when additional physical storage space is required. Using thin provisioning reduces the time and effort required to provision additional storage, and avoids provisioning storage that may not be needed.

Note: LUNs that are assigned to VNX for File cannot be expanded. If additional capacity is required on VNX for File, new LUNs should be created and assigned for AVM to manage.

For a thick LUN, the pool must have enough storage for the LUN expansion to succeed, whereas for a thin LUN the storage does not need to be available. It is important to note that you cannot expand a pool LUN if it is part of a LUN-migration or data protection operation such as MirrorView.

When to use Classic, Thick, and Thin LUNs

It is important to understand your application requirements and select the approach that meets your needs. If conditions change, you can use VNX LUN migration to migrate among thin, thick, and classic LUNs.

Use pool-based thin LUNs for:

- Applications with moderate performance requirements

- Taking advantage of advanced data services like FAST VP, VNX Snapshots, Compression, and Deduplication
- Ease of setup and management
- Best storage efficiency
- Energy and capital savings
- Applications where space consumption is difficult to forecast

Use pool-based thick LUNs for:

- Applications that require good performance
- Taking advantage of advanced data services like FAST VP and VNX Snapshots
- Storage assigned to VNX for File
- Ease of setup and management

Use classic LUNs for:

- Applications that require extreme performance (for example, when milliseconds of performance are critical)
- The most predictable performance
- Precise data placement on physical drives and logical data objects
- Physical separation of data

For more information on thin LUNs, please refer to Appendix A: Thin LUNs.

Using Virtual Provisioning for VNX File

VNX also supports Virtual Provisioning for VNX File.

Thin LUNs for VNX File

In VNX OE for Block version 05.33.008.5.119 and VNX OE for File 8.1.8.119 new space efficiency features such as VNX Block Deduplication for File and File System and Checkpoint Storage Space Reclaim are introduced for VNX File. Each of these features utilize thin LUNs as the underlying VNX File storage. When utilizing thin LUNs, the user has the option to oversubscribe the Block Storage Pool, which can be a concern. This release also includes Low Space and Out of Space Handling for VNX File which protects VNX File from a Block Storage Pool out of space condition. EMC **strongly** recommends creating File Systems with a Log Type of Split when using thin LUNs as the underlying storage for VNX File. In VNX OE for File 8.1.6.96 and later, Split is the default option for Log Type when creating a File System.

VNX Block Deduplication for VNX File

In VNX OE for Block version 05.33.008.5.119 and VNX OE for File 8.1.8.119 VNX Block Deduplication can be used with VNX File. Enabling VNX Block Deduplication for a VNX File mapped storage pool or meta volume causes all underlying LUNs to become deduplication enabled thin LUNs. VNX Block Deduplication utilizes an 8 KB Fixed-Block Deduplication method of removing redundant data from all deduplication enabled LUNs within a Block Storage Pool. For more information on VNX Block Deduplication for File, management, and differences between VNX Block Deduplication for File and VNX File Deduplication and Compression refer to the *EMC VNX2 Deduplication and Compression* white paper on EMC Online Support.

File System and Checkpoint Storage Space Reclaim

File System and Checkpoint Storage Space Reclaim are new features in VNX OE for Block version 05.33.008.5.119 and VNX OE for File 8.1.8.119. These features are only available when the underlying storage for the File System or Checkpoint Storage (SnapSure SavVol) is comprised of thin LUNs. The File System itself can be Thick or Thin. With these features, the user is able reclaim space no longer in use by a File System or the Checkpoint Storage to the underlying Block thin LUNs. The freed space within the thin LUNs can then be used to store new data within the File System or Checkpoint Storage. If enough space is reclaimed, the freed space in the form of LUN slices will be released back to the Block Storage Pool. The File System must be mounted to issue reclaim, and issuing reclaim does not shrink the File System or Checkpoint Storage.

When an application writes data to a File System built on a Mapped Storage Pool, space is allocated from the Storage Pool and the File System grows. When data is deleted, space previously allocated stays with the File System. This space is marked for overwrite and only writing new data to the File System will reuse this space. File System Space Reclaim specifically targets this previously used space within the File System and attempts to return the free blocks to the Block Storage Pool. When

issued, an internal process reviews which blocks contain data and which are not currently in use within the File System. Reclaim operations, similar to SCSI UNMAP commands, are then sent to the underlying thin LUNs for the blocks that can be reclaimed.

Checkpoint Storage Space Reclaim has a similar process of freeing space that is no longer needed to the Block Storage Pool. When a SnapSure Checkpoint is deleted or refreshed, the Checkpoint Storage may contain blocks of data that are no longer needed. As with deleted data in a File System, these blocks are marked for overwrite. When Checkpoint Storage Space Reclaim is issued, the process issues reclaim operations for these blocks to the underlying thin LUNs.

When File System or Checkpoint Storage Space Reclaim is started, blocks to be freed are sent reclaim commands. Once all reclaim commands are sent, the Space Reclaim operation finishes. While the Space Reclaim process is complete the block operations on the Storage Pool, such as slice evacuations, may continue. It is suggested that Space Reclaim be started off peak hours, so that reclaim activity does not contend with the primary workload.

Management

You can easily manage File System and Checkpoint Storage Space Reclaim by using either the Unisphere software or Control Station CLI. When a File System is created on thin LUNs, a new tab named Space Reclaim appears in the File System Properties Window. Similarly, when the SnapSure Checkpoint Storage is on thin LUNs, a new tab named Checkpoint Storage Space Reclaim appears in the File System Properties Window. **Figure 9** shows an example of the two new reclaim tabs. While File System and Checkpoint Storage Space Reclaim are two entirely separate features, the controlling of Space Reclaim is very similar.

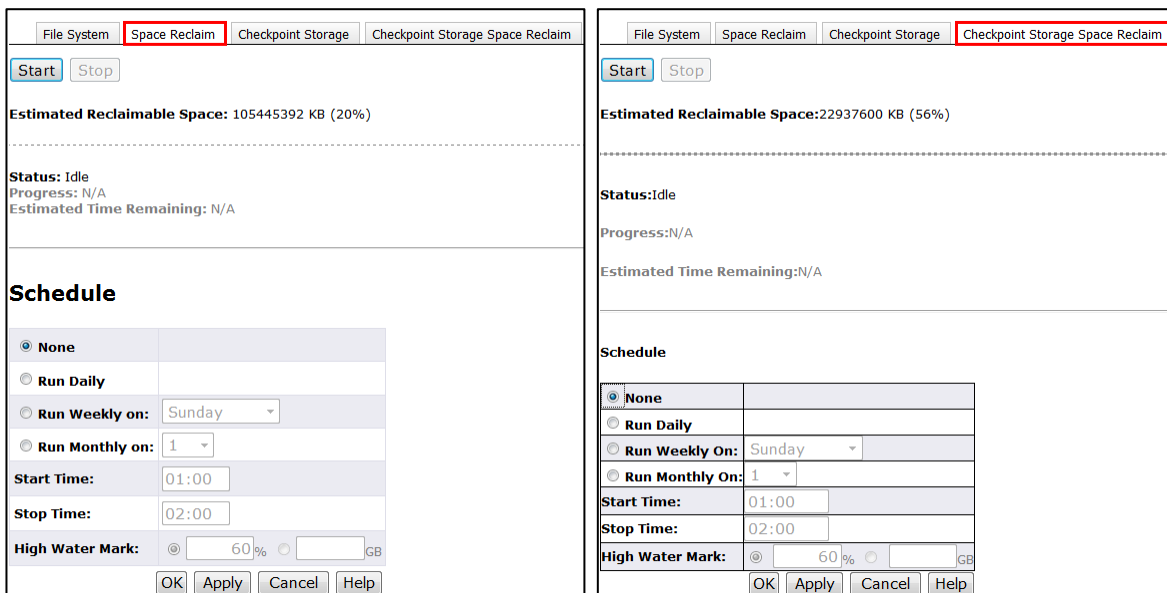


Figure 9 Reclaim Tabs within the File System Properties Window

At the very top of each tab is a **Start** and **Stop** control. These are manual controls for each of the reclaim operations. The user may start a reclaim process or stop a running reclaim process at any point in time. Stopping a running process will only stop the remaining reclaim commands from running, and will not stop any background Storage Pool operations that may have started. Also shown is the **Estimated Reclaimable Space**, which shows the estimated amount of space in KB and % of the File System space which can be reclaimed.

In the middle of each tab is the **Status**, which is the status of reclaim for the tab. The status can either be:

- **Idle** – No Space Reclaim operations are currently running
- **Pending** – A reclaim operation is waiting for other reclaims to finish before running
- **Running** – A reclaim operation is currently running

Also shown is the **Progress**, which is shown as a percentage complete of the overall reclaim operation. Lastly the **Estimated Time Remaining**, which is a time estimate on when the reclaim operations will finish. Neither the Progress nor the Estimated Time Remaining includes an estimate for when the background operations, such as slice evacuations, on the Storage Pool will finish.

In the bottom of each reclaim tab in [Figure 9](#) is a **Schedule**, which the user can customize. The user has the following choices for a schedule:

- **None** – No schedule is currently specified for reclaim (Default)
- **Run Daily** – The reclaim operation is scheduled to Run Daily
- **Run Weekly on** – The day of the week the user would like reclaim operations to run
- **Run Monthly on** – The date within a month the user would like the reclaim operations to run

Following the above options is the **Start Time** and **Stop Time**. With these the user can schedule the time window for reclaim operations to start and stop. If the **Stop Time** is reached and the reclaim operations have not finished running, they will need to be issued if needed in the next reclaim window. The **Start Time** is required, but the **Stop Time** is not. If the **Stop Time** is left blank, the reclaim operations will run to completion no matter how long they take.

Lastly in the reclaim tabs in [Figure 9](#) is the **High Water Mark**. The **High Water Mark** is the minimum threshold for savings in which a reclaim schedule will be allowed to run. The **High Water Mark** can be one of the following:

- **Percentage** – This specifies the minimum percentage of space able to be reclaimed for a scheduled reclaim operation to start in relation to the size of the File System or Checkpoint Storage. For instance, if the total space used for existing and deleted data within the File System or Checkpoint Storage is 100 GBs, and the **High Water Mark** is 10%, a scheduled reclaim operation will only start if over 10% of the space can be reclaimed at that time. In this example, reclaim will only run if at least 10 GBs can be reclaimed.

- **GB** – This specifies the minimum amount of reclaimable space needed, in GBs, to allow the scheduled reclaim operations to start.

File System Space Reclaim can also be issued when deleting a File System. **Figure 10** shows the **Perform Space Reclaim** option found within the File System Confirm Delete window. By default the **Perform Space Reclaim** box is not selected, and a reclaim operation cannot be issued once a File System is deleted. Checkpoint Storage Space Reclaim will start automatically when the last SnapSure Checkpoint is deleted for a File System.

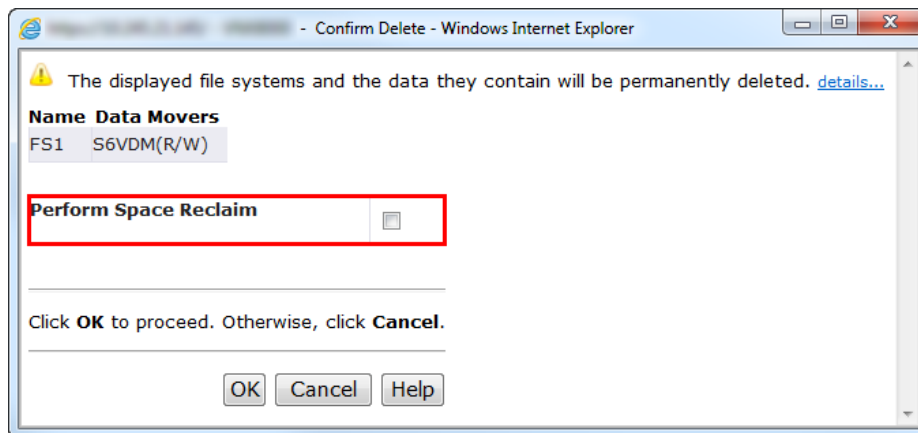


Figure 10 File System Delete Window

File System and Checkpoint Storage Space Reclaim can also be controlled via Control Station CLI. The `server_reclaim` command will show the status of reclaim, the Estimated Reclaimable Space, and has Start and Stop controls for File System and Checkpoint Storage Space Reclaim. To list the reclaim schedules on a system, view more information about a reclaim schedule, and to create or modify an existing schedule use the `nas_reclaim_schedule` command. When deleting a File System via the `nas_fs` CLI command, you can also specify the `-reclaim` option to issue a reclaim upon delete. For detailed information on each of these commands and others, please consult the *VNX Command Line Interface Reference for File* documentation, located on EMC Online Support.

Low Space Handling for VNX File

Low Space Handling for VNX File is a new feature in VNX OE for Block version 05.33.008.5.119 and VNX OE for File 8.1.8.119. With thin LUNs, the user has the ability to create LUNs with sizes which exceed the physical capacity within a Storage Pool. If the Storage Pool is oversubscribed, there is potential for the Storage Pool to become full. While many alerts warn the user a Storage Pool is becoming full, Low Space Handling will take action on VNX File resources utilizing thin LUNs from a Block Storage Pool low on free space. While under Low Space Handling, affected Read/Write File Systems and SnapSure Checkpoints are remounted Read-Only, and any Replication Sessions are stopped. No block-only resources are affected by Low Space Handling.

One component to Low Space Handling is the Low Space Threshold. The Low Space Threshold is a value in GBs which defines when VNX File should take action on the VNX File resources to protect them from the underlying storage becoming full. This value is set on each Block Storage Pool, but is only applicable to Pools which provide thin LUNs to VNX File. The Low Space Threshold is configuration specific, and is directly related to the total amount memory across all Data Movers on the system.

When a Block Storage Pool's Free Capacity is less than the Low Space Threshold, and thin LUNs within the Pool are provided to VNX File, an error message is logged on the system. VNX File receives this alert and determines if action needs to be taken. All Read/Write File Systems created on the affected storage will be remounted Read-Only along with any Read/Write SnapSure Checkpoints relating to the File System. If these File Systems also have replication sessions, the sessions are stopped and left in a restartable state. For any Read/Write SnapSure Checkpoints whose Checkpoint Storage (SavVol) is located on the affected storage but their Production File Systems do not, they will also be remounted Read-Only along with their Production File Systems. Virtual Data Mover root File Systems are left unaffected by Low Space Handling. A table for allowed and restricted operations while under Low Space Handling can be found in [Appendix B: Low Space and Out of Space Handling Interoperability](#).

To clear a low space condition, first the user must increase the Free Capacity within the Block Storage Pool. The Free Capacity of the Storage Pool must exceed the Low Space Threshold. Adding additional disks to the pool is the most efficient method to increase the Free Capacity within a Storage Pool. Once there is enough Free Capacity within the Block Storage Pool, the low space condition is cleared and the user can restore access to the affected File Systems, SnapSure Checkpoints, and Replication Sessions.

After restoring access to an impacted File System, the File System remounts as Read/Write, and any affected SnapSure Checkpoints associated with the File System are automatically remounted as Read/Write. If the File System is the source to a Replication Session, the session is automatically restarted. If the File System is a destination image of a Replication Session, the user must manually start the Replication Session from the source system.

To increase the Free Capacity of the Block Storage Pool the user may also attempt to utilize File System or Checkpoint Storage Space Reclaim. Only space consumed by deleted data before Low Space Handling occurred can be reclaimed as the affected File Systems will be Read-Only. There is no guarantee Space Reclaim will free enough storage to clear the low space condition.

Calculating the Low Space Threshold

When utilizing thin LUNs on VNX File, it is best to understand the size of the Low Space Threshold for the system. To determine the Low Space Threshold you must first determine the number of Data Movers installed on the system. To do so, issue the `nas_server -list` command via the Control Station:


```
[nasadmin@VNX8000 ~]$ nas_server -list
id      type  acl  slot groupID  state  name
1       1    0   2      0         0  server_2
2       1    0   3      0         0  server_3
3       1    0   4      0         0  server_4
4       1    0   5      0         0  server_5
5       1    0   6      0         0  server_6
6       1    0   7      0         0  server_7
7       4    0   8      0         0  server_8
```

In this example, the VNX8000 has 7 Data Movers installed and configured. Next, you need to determine the Total main memory (MB) installed for a Data Mover on the system. All of the Data Movers will have the same memory size within a system so only checking one is necessary. To check the memory size, issue the `server_sysconfig -Platform` command via the Control Station to a Data Mover installed on the system:

```
[nasadmin@VNX8000 ~]$ server_sysconfig server_2 -Platform
server_2 :
Processor           = Intel Six Core Westmere
Processor speed (MHz) = 2800
Total main memory (MB) = 24576
Model               = VNX8000
QPI speed (MHz)     = 6400
Bios Version        = 08.10
Post Version         = Rev. 43.20
Firmware Version    = Rev. 12.50
Family ID           = 0x0012
FRU ID              = 0x000c
```

To finish calculating the Low Space Threshold you multiply the number of Data Movers installed in the system by the Total main memory (MB) of one of the Data Movers. In this example, the Low Space Threshold is $7 * 24576 \text{ MBs} = 172032 \text{ MBs}$. When the Free Capacity of the Block Storage Pool is below the Low Space Threshold, Low Space Handling will take effect.

Restoring from Low Space Handling

When logging into Unisphere after encountering Low Space Handling, the user will see alerts similar to those shown in [Figure 11](#). The alerts shown include Storage Pool Space Usage Alerts along with the *Low space left on storage pool (StoragePoolName)* alert, which is part of the Low Space Handling feature. This alert indicates that the Storage Pool listed has crossed the Low Space Threshold.

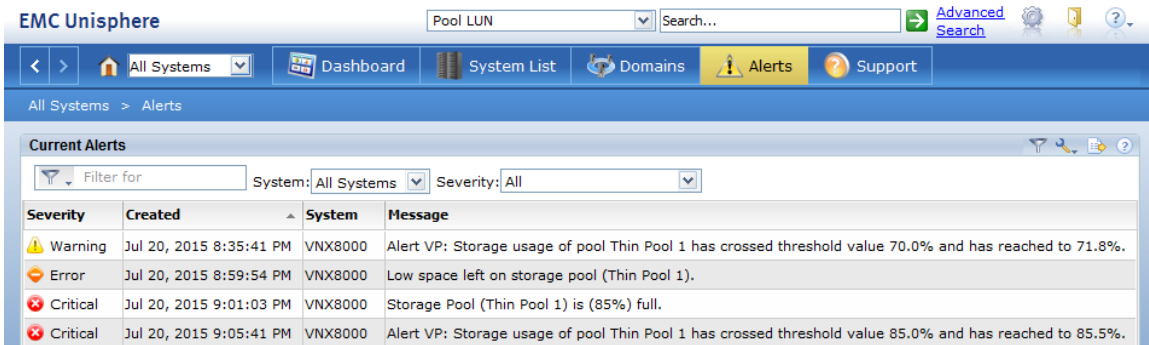


Figure 11 Unisphere Alerts - Low Space Handling

When reviewing Unisphere to determine which File Systems are affected by Low Space Handling, you must view the Mounts tab under Storage > Storage Configuration > File Systems. As shown in Figure 12, the **State** column labels the affected File Systems with a state of **Read Only (Low Space)**. This indicates these File Systems were directly affected by Low Space Handling.

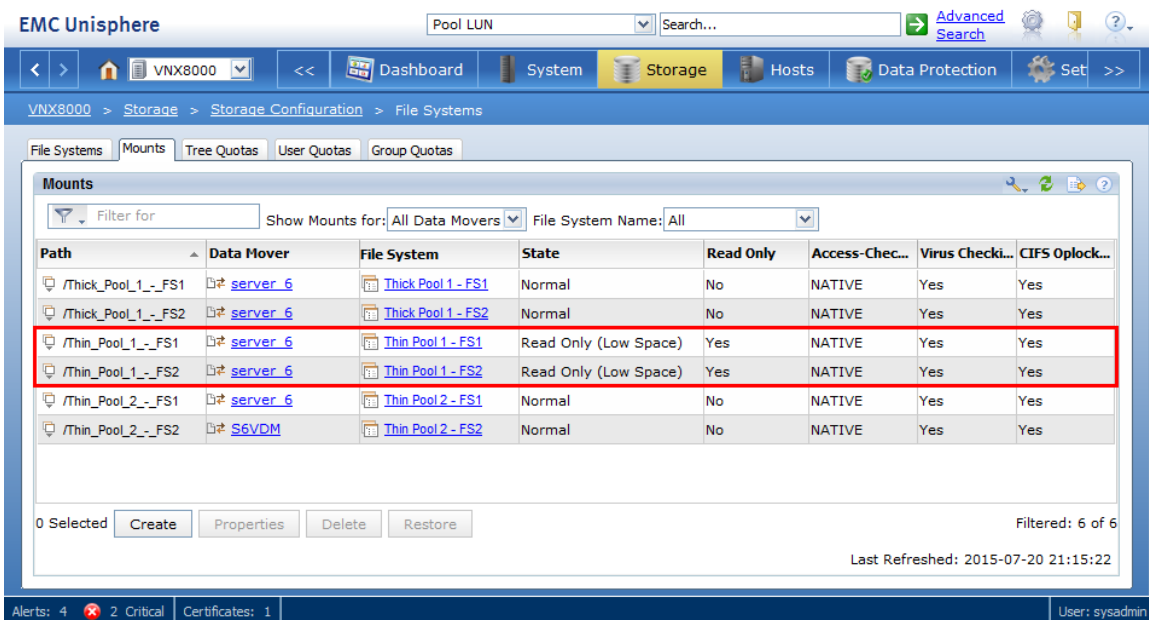


Figure 12 File Systems Mounts tab – Read Only (Low Space)

After reviewing the affected File Systems, you may need to review which SnapSure Checkpoints and Replication Sessions are currently impacted. For SnapSure Checkpoints, navigate to the Mounts tab under Data Protection > Snapshots > File System Checkpoints. As shown in Figure 13, the **State** column shows **Read Only (Low Space)** to indicate the checkpoints have been remounted Read-Only due to Low Space Handling.

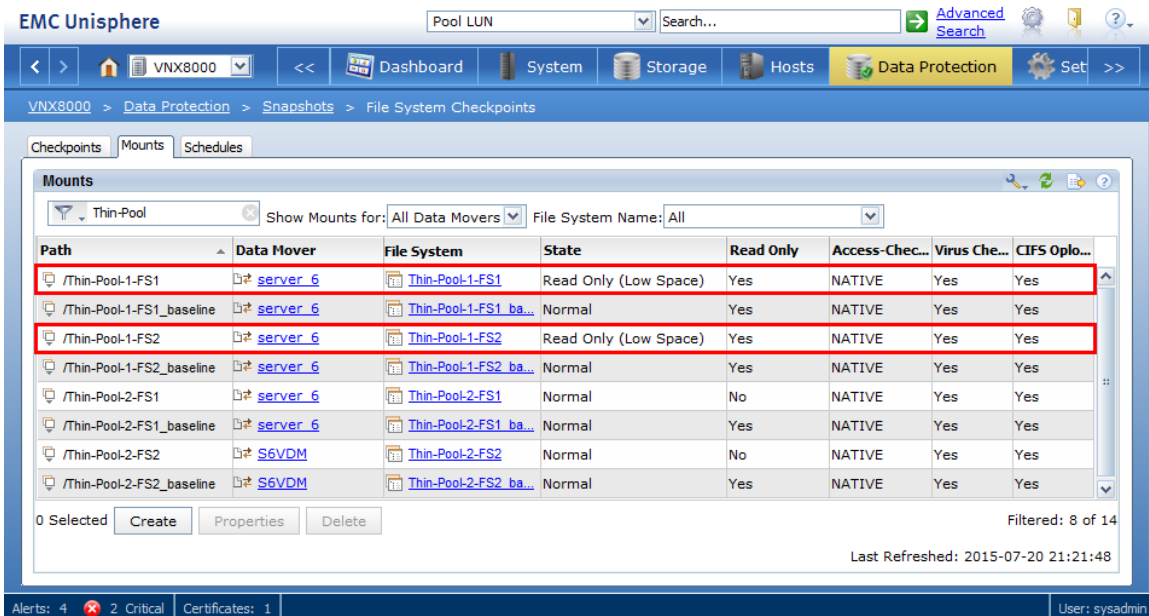


Figure 13 SnapSure Checkpoints – Read Only (Low Space)

For impacted Replication Sessions, you will need to review the Replications tab under Data Protection > Mirrors and Replications > Replications for File. As shown in Figure 14, stopped replication sessions exist in this system. Only confirming that the File System was affected will confirm if the session was stopped due to Low Space Handling.

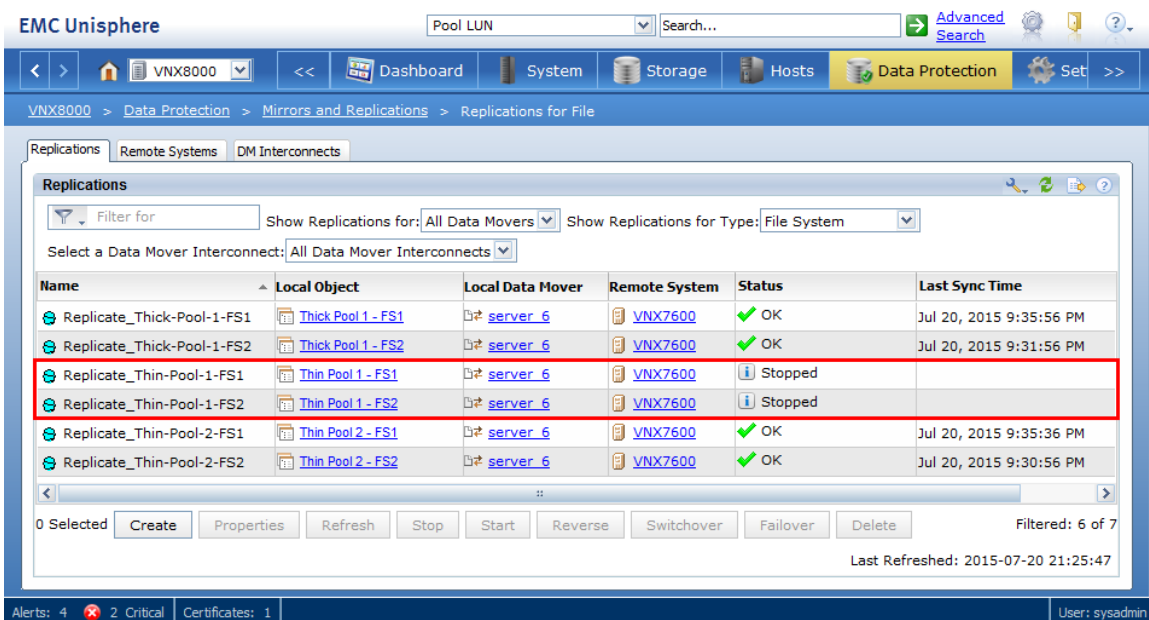


Figure 14 Replications for File – Stopped Replication Sessions

After increasing the Free Capacity within the affected Block Storage Pool, the user needs to restore access to the affected File Systems, SnapSure Checkpoints, and resume any replication session stopped by Low Space Handling. In Unisphere,

navigate to the Mounts tab under Storage > Storage Configuration > File Systems. Select one or multiple affected File Systems to restore and click the **Restore** button as shown in [Figure 15](#). After following the dialog boxes, access to the affected File Systems will be restored, along with their affected SnapSure Checkpoints. If the File System is replicated, replication will automatically resume once the File System is restored. If the File System is a destination image of a Replication Session, the user must manually start the Replication Session from the source system.

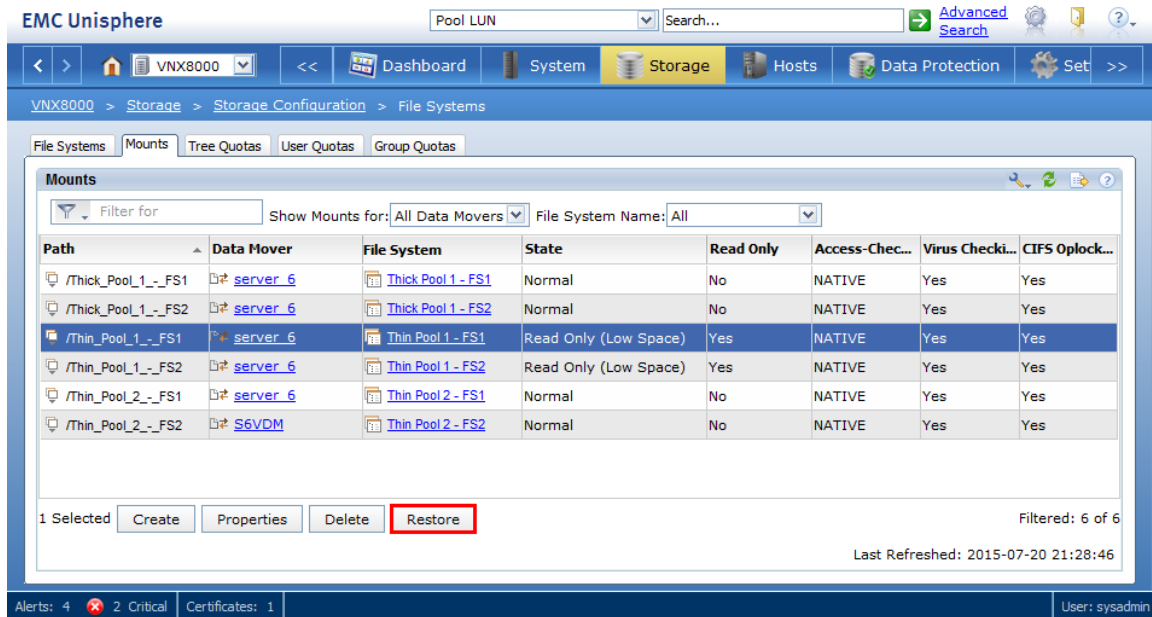


Figure 15 File Systems Mounts tab – Restore Button

Via Control Station CLI the user can also view the File System and SnapSure Checkpoint mounts to determine if Low Space Handling has occurred. In the output of the `server_mount` command, File Systems and SnapSure Checkpoints marked with **state=ls** have been affected by Low Space Handling. An example of the output of the `server_mount` command can be found below:

```
[nasadmin@VNX8000 ~]$ server_mount server_6
server_6 :
root_fs_6 on / uxfs,perm,rw
root_fs_common on /.etc_common uxfs,perm,ro
root_fs_vdm_S6VDM on /root_vdm_2/.etc uxfs,perm,rw
Thick Pool 1 - FS1 on /Thick_Pool_1_-_FS1 uxfs,perm,rw
Thin Pool 1 - FS1 on /Thin_Pool_1_-_FS1 uxfs,perm,ro,lsrw,state=ls
Thin Pool 2 - FS1 on /Thin_Pool_2_-_FS1 uxfs,perm,rw
Thick-Pool-1-FS1_baseline on /Thick-Pool-1-FS1_baseline ckpt,perm,ro
Thick-Pool-1-FS1 on /Thick-Pool-1-FS1 ckpt,perm,rw
Thin-Pool-1-FS1_baseline on /Thin-Pool-1-FS1_baseline ckpt,perm,ro
Thin-Pool-1-FS1 on /Thin-Pool-1-FS1 ckpt,perm,ro,lsrw,state=ls
Thin-Pool-2-FS1_baseline on /Thin-Pool-2-FS1_baseline ckpt,perm,ro
Thin-Pool-2-FS1 on /Thin-Pool-2-FS1 ckpt,perm,rw
```

After increasing the Free Capacity within the Block Storage Pool, the user can use the `server_mount -restore` command via Control Station CLI to restore access to

the affected File Systems and SnapSure Checkpoints. For detailed information on the usage of each of these commands and others, please consult the *VNX Command Line Interface Reference for File* documentation, located on EMC Online Support.

Out of Space Handling for VNX File

With thin LUNs, the user has the ability to create LUNs with sizes which exceed the physical capacity within a Storage Pool. If the Storage Pool is oversubscribed, there is potential for the Storage Pool to become full. Out of Space Handling, as with Low Space Handling, is a protection feature for VNX File when using thin LUNs as the underlying storage. Out of Space Handling for VNX File is a new feature in VNX OE for Block version 05.33.008.5.119 and VNX OE for File 8.1.8.119.

The purpose of Out of Space Handling is to limit the impact when an out of space condition exists on the system. Out of Space Handling only occurs when the Block Storage Pool providing thin LUNs to VNX File becomes full and Low Space Handling has not occurred on all affected File Systems. In the instance where the Block Storage Pool receives a write from a Data Mover but there is no Free Capacity within the pool to save it, a failed write occurs and the Data Mover will panic. At this point, Out of Space Handling occurs, the Data Mover fails over to the standby, and access is restored to File Systems and other resources not affected by the out of space condition.

When Out of Space Handling occurs, all File Systems and SnapSure Checkpoints created on the affected Block Storage Pool will be unmounted from the Data Mover. Any replication sessions for affected File Systems or sessions with Replication Checkpoints utilizing space from the affected storage will also be stopped. If a Virtual Data Mover has been created on the affected storage, any File Systems mounted on the Virtual Data Mover will also be unmounted regardless if the File Systems are created on the affected storage or not. After all affected File Systems and SnapSure Checkpoints are unmounted and replication is stopped where necessary, the Data Mover is allowed to boot normally and restore access to unaffected File Systems, SnapSure Checkpoints, and replication sessions.

As with Low Space Handling, the user must increase the Free Capacity within the Block Storage Pool before restoring access to the affected resources. The Free Capacity of the Block Storage Pool must exceed the Low Space Threshold for the restore to occur. Once enough Free Capacity exists within the Block Storage Pool, the user can utilize the Restore operation to restore access to the affected resources.

During the restore operation, access to affected File Systems, SnapSure Checkpoints, and Virtual Data Movers is restored. However, the Data Mover must be rebooted before Replication Sessions that were affected by the out of space condition are able to be restarted. Issuing a failback from the standby Data Mover will cause a reboot to occur and Replication Sessions will be able to be restarted. If an affected File System is a destination image of a Replication Session, the user must manually start the Replication Session from the source system.

A table for allowed and restricted operations while under Out of Space Handling can be found in [Appendix B: Low Space and Out of Space Handling Interoperability](#).

Restoring from Out of Space Handling

When logging into Unisphere after encountering Out of Space Handling, the user will see alerts similar to those shown in [Figure 16](#). The alerts shown include Storage Pool Space Usage Alerts, information about the Data Mover panic, and the *Storage pool (StoragePoolName) is out of space* alert, which is part of the Out of Space Handling feature. This alert indicates that the Storage Pool listed has encountered an Out of Space Condition, and that Out of Space Handling may have occurred.

The screenshot shows the EMC Unisphere interface with the Alerts section active. The 'Current Alerts' table displays the following data:

Severity	Created	System	Message
Critical	Jul 20, 2015 10:08:50 PM	VNX8000	Slot 6 has panicked.
Error	Jul 20, 2015 10:09:56 PM	VNX8000	Low space left on storage pool (Thin Pool 1).
Critical	Jul 20, 2015 10:10:07 PM	VNX8000	Storage pool (Thin Pool 1) is out of space.
Critical	Jul 20, 2015 10:10:26 PM	VNX8000	Storage Pool (Thin Pool 1) is (100%) full.
Critical	Jul 20, 2015 10:15:44 PM	VNX8000	Alert VP: Storage usage of pool Thin Pool 1 has crossed threshold value 85.0% and has reached to 100.0%.
Warning	Jul 21, 2015 1:00:20 AM	VNX8000	DME 2 Data Mover 6: out of service
Warning	Jul 21, 2015 1:00:20 AM	VNX8000	VNX VNX8000 APM001410615432007: Warning

At the bottom of the alert list, it shows '0 Selected', 'Details', 'Delete', and '7 items'. A summary bar at the bottom indicates 'Alerts: 7', '4 Critical', and 'Certificates: 1'. The user is identified as 'sysadmin'.

Figure 16 Unisphere Alerts – Out of Space Handling

When reviewing Unisphere to determine which File Systems are affected by Out of Space Handling, you must view the Mounts tab under Storage > Storage Configuration > File Systems. As shown in [Figure 17](#), the **State** column labels the affected File Systems with a state of **Unmounted (Out of Space)**. This indicates these File Systems were affected by Out of Space Handling. In this example, the naming of the File Systems indicates the pool the File System was created on and the File System number. File Systems with “Thin Pool 1” in the name were created on the affected Storage Pool, and are directly affected by the out of space condition. “Thin Pool 2 – FS2” is affected as it is mounted to a Virtual Data Mover which was created on the affected storage.

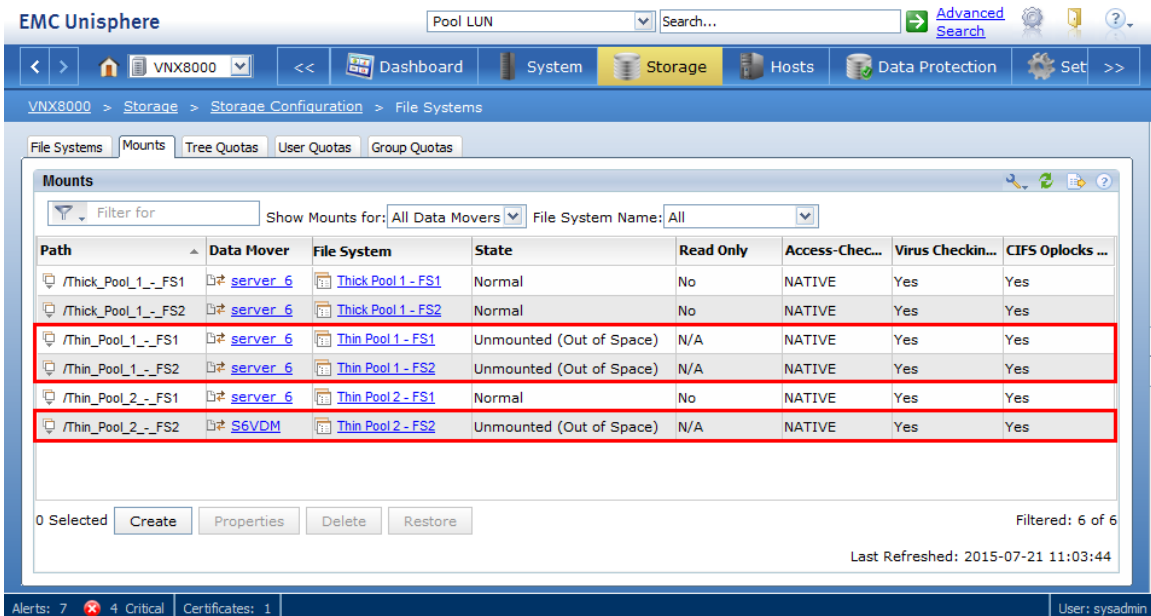


Figure 17 File Systems Mounts tab – Unmounted (Out of Space)

After reviewing the affected File Systems, you may need to review which SnapSure Checkpoints and Replication Sessions are currently impacted. For SnapSure Checkpoints, navigate to Checkpoints under Data Protection > Snapshots > File System Checkpoints. As shown in Figure 18, the **State** column shows **Unmounted (Out of Space)** to indicate the checkpoints have been unmounted due to Out of Space Handling.

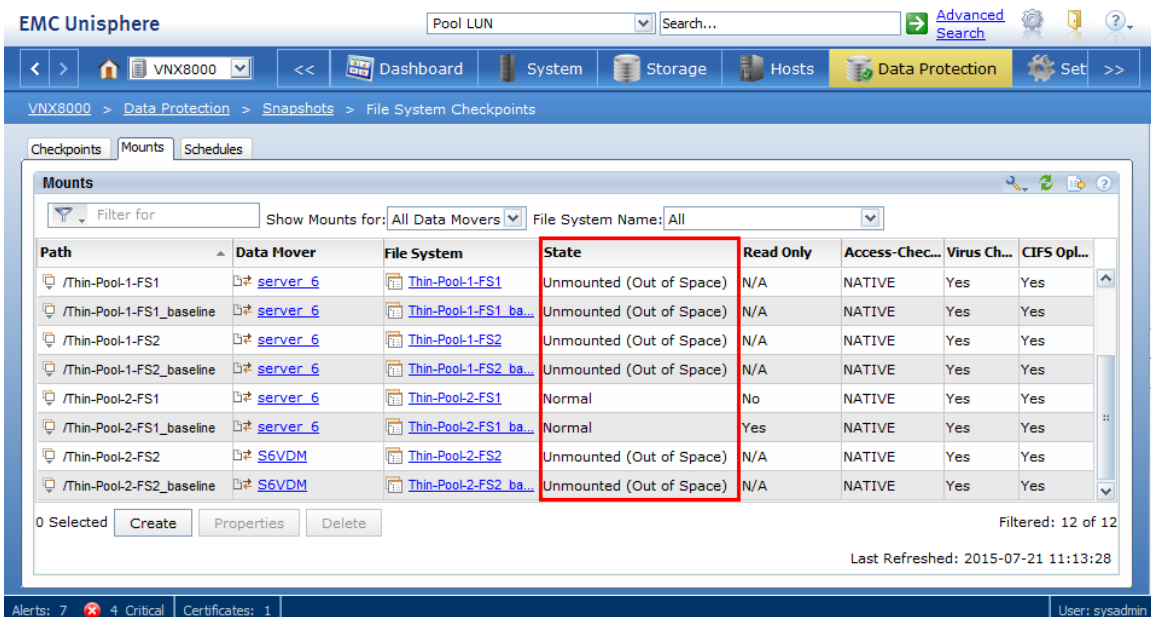


Figure 18 Snapsure Checkpoints –Unmounted (Out of Space)

Via Control Station CLI the user can view the File System and SnapSure Checkpoint mounts to determine if Out of Space Handling has occurred. In the output of the

server_mount command, File Systems and SnapSure Checkpoints marked with **state=oos** have been affected by Out of Space Handling. An example of the output of the server_mount command can be found below:

```
[nasadmin@VNX8000 ~]$ server_mount server_6
server_6 :
root_fs_6 on / udfs,perm,rw
root_fs_common on /.etc_common udfs,perm,ro
root_fs_vdm_S6VDM on /root_vdm_2/.etc udfs,perm,rw,state=oos,<unmounted>
Thick Pool 1 - FS1 on /Thick_Pool_1_-_FS1 udfs,perm,rw
Thin Pool 1 - FS1 on /Thin_Pool_1_-_FS1 udfs,perm,ro,lsrw,state=oos,<unmounted>
Thin Pool 1 - FS2 on /Thin_Pool_1_-_FS2 udfs,perm,rw,state=oos,<unmounted>
Thin Pool 2 - FS1 on /Thin_Pool_2_-_FS1 udfs,perm,rw
Thin Pool 2 - FS2 on /root_vdm_2/Thin_Pool_2_-_FS2
udfs,perm,rw,state=oos,<unmounted>
Thin-Pool-1-FS1_baseline on /Thin-Pool-1-FS1_baseline
ckpt,perm,ro,state=oos,<unmounted>
Thin-Pool-1-FS1 on /Thin-Pool-1-FS1 ckpt,perm,rw,state=oos,<unmounted>
Thin-Pool-1-FS2_baseline on /Thin-Pool-1-FS2_baseline
ckpt,perm,ro,state=oos,<unmounted>
Thin-Pool-1-FS2 on /Thin-Pool-1-FS2 ckpt,perm,rw,state=oos,<unmounted>
Thin-Pool-2-FS1_baseline on /Thin-Pool-2-FS1_baseline ckpt,perm,ro
Thin-Pool-2-FS1 on /Thin-Pool-2-FS1 ckpt,perm,rw
Thin-Pool-2-FS2_baseline on /root_vdm_2/Thin-Pool-2-FS2_baseline
ckpt,perm,ro,state=oos,<unmounted>
Thin-Pool-2-FS2 on /root_vdm_2/Thin-Pool-2-FS2
ckpt,perm,rw,state=oos,<unmounted>
```

After navigating to the Replications tab under Data Protection > Mirrors and Replications > Replications for File, you will see that Replication Sessions for the impacted File Systems are not shown. This is depicted in [Figure 19](#). As the affected File Systems are unmounted, and the replication services for these File Systems are stopped, the Replication Sessions will not be shown until the File System is restored and the Data Mover is rebooted.

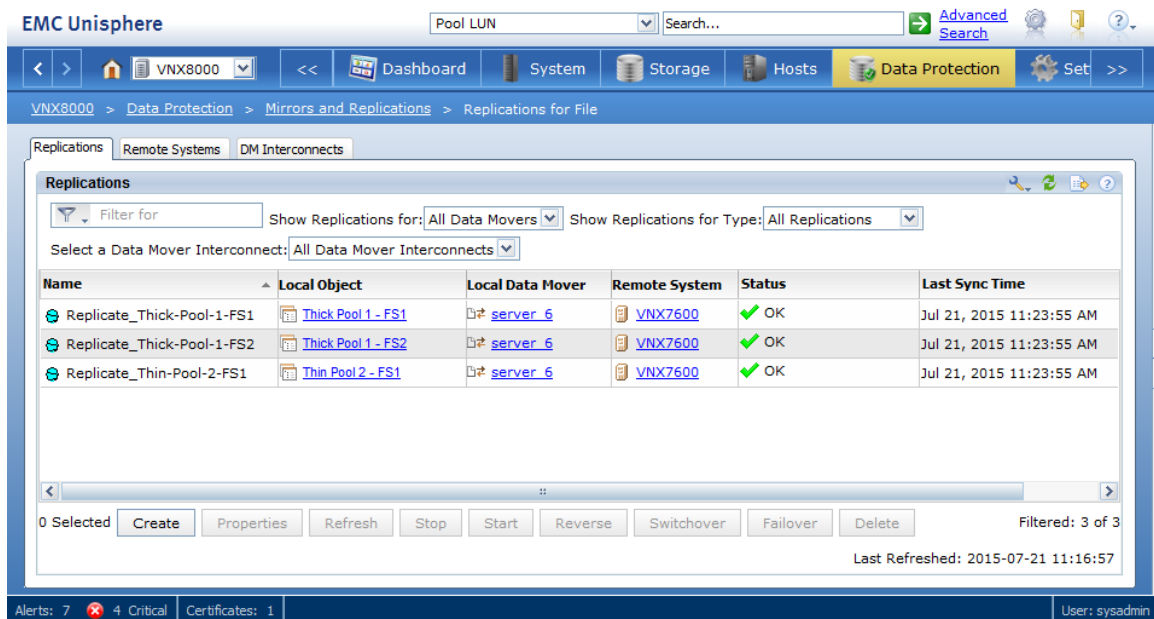


Figure 19 Replications for File

After increasing the Free Capacity within the affected Block Storage Pool, the user needs to restore access to the affected File Systems, SnapSure Checkpoints, and Virtual Data Movers. In Unisphere, navigate to the Mounts tab under Storage > Storage Configuration > File Systems. Select one or multiple affected File Systems to restore and click the **Restore** button as shown in [Figure 20](#). After following the dialog boxes, access to the affected File Systems will be restored, along with their affected SnapSure Checkpoints.

Note: If the Restore operation is issued from Unisphere, the Replication Sessions will not automatically be restarted after the Data Mover is rebooted. The user will need to manually Stop and Start the Replication Sessions once the Data Mover has been rebooted.

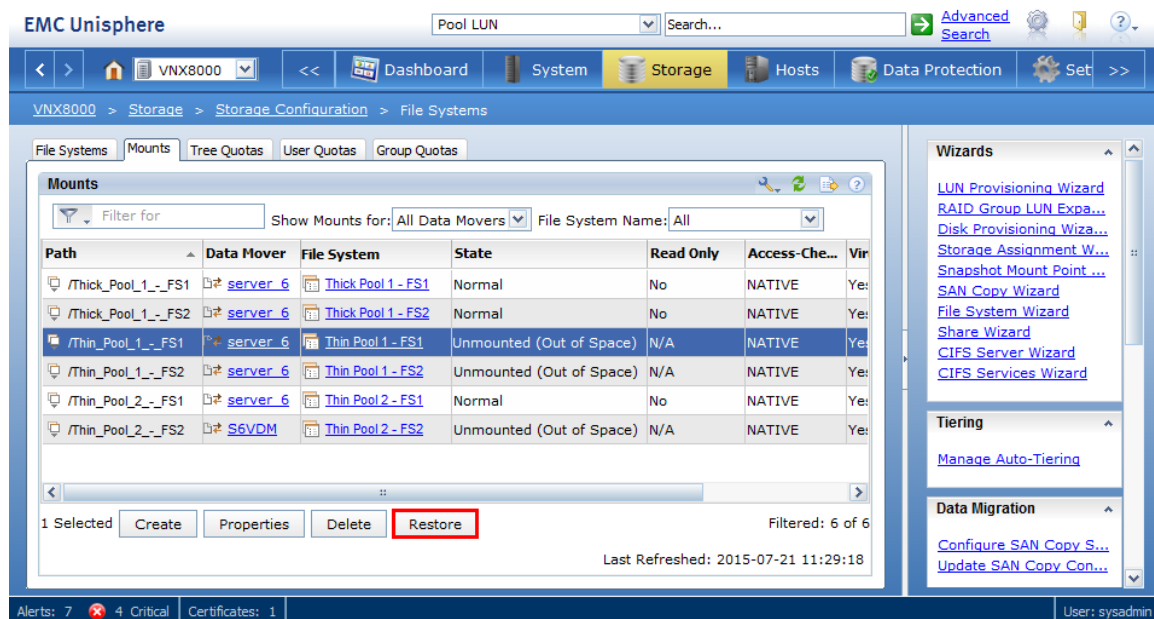


Figure 20 File Systems Mounts tab – Restore Button

To restore Replication Sessions, you must first reboot the Data Mover. As the Data Mover is currently failed over to the standby due to Out of Space Handling, fail back the Data Mover using the `server_standby -restore` command via the Control Station. Issuing a fail back operation will cause a warm reboot of the Data mover to occur. After failing back the Data Mover, the user must Stop and Start each Replication Session to resume replication.

After increasing the Free Capacity within the Block Storage Pool, the user can also use the `server_mount -restore` command via Control Station CLI to restore access to the affected File Systems and Snapsure Checkpoints. If utilizing the `server_mount -restore -all` command, the Data Mover is automatically failed back to the primary and access will be restored to the File Systems, SnapSure Checkpoints, Virtual Data Movers, and Replication Sessions. If an affected File System is a destination image of a Replication Session, the user must manually start the Replication Session from the source system. For detailed information on the

usage of each of these commands and others, please consult the *VNX Command Line Interface Reference for File* documentation, located on EMC Online Support.

Thick and Thin-enabled File Systems

First, let's take a look at how thick File Systems work. Thick file systems are fully allocated on creation and the size of the File System is equal to the size that's reported to the host. This can lead to a large amount of allocated storage that is unused.

Thick File Systems can be created with a smaller initial size and be extended when necessary. You also have the option of taking advantage of the auto-extension feature which will automatically extend the File System when it's nearly fully. To use this, you must provide the High Water Mark (default of 90%) and the maximum size the File System should grow to.

Like a thick File System, a thin-enabled File System (as opposed to a thick File System on a thin LUN) is also configured with an initial size that is fully allocated. However, the thin-enabled File System has defined attributes that determine when it should be extended and what its maximum size should be. The maximum File System size is what the end user sees as the available capacity in the File System.

It is important to note that thin and thick File Systems have different usage characteristics. Some applications, I/O workloads, and storage deployment scenarios may see performance improvements from using thin-enabled File Systems. However, it is important to note that these improvements may change over time as the thin-enabled File System expands and as the data is used, deleted, or modified.

With a thin-enabled File System, performance improves mostly with random and mixed read I/O workloads. Because the thin file system initially occupies less space on disk than a fully provisioned file system, there are smaller disk seeks required for random reads. Disk seeks impact I/O latency, so minimizing seeks can improve performance.

With sequential read I/O, disk seeks are already infrequent, and therefore a performance improvement would not be noticed. Write I/O will also not see much performance improvement as disk seeks are usually not necessary, or only minimal (except for random overwriting), and will mostly be handled by cache anyway. It should be emphasized that this performance improvement may decrease over time as the File System is further used, extended, and fragmented which increases the size of disk seeks and the corresponding latency.

Creating thin-enabled File Systems

For applications requiring consistent and predictable performance, EMC **strongly** recommends that you create File Systems that use thick, non-compressed LUNs with Block Deduplication disabled. In VNX OE for Block, Release 33, the default LUN type for a new LUN is thin; therefore, uncheck the Thin checkbox when you create the LUN. Block Deduplication and Compression will be disabled by default. You can achieve

thin, deduplication, and compression optimizations by using the thin and File deduplication (which includes compression) attributes at the File System level.

Prior to creating any File System, you must first assign storage to VNX File. To do this:

1. Provision LUNs (thick LUNs recommended) from a pool.
2. Assign the LUNs to the protected ~filestorage Storage Group.
3. Initiate a rescan of the storage systems (under the System tab in Unisphere). This will initiate the following operations:
 - a. Run a diskmark which makes the LUNs available to VNX for File components for use as File storage space.
 - b. Create a Storage Pool for File that is the same name as the corresponding Storage Pool.
 - c. Create disk volumes in a 1:1 mapping for each LUN that was added to the Storage Group.

Once that is complete, you can view information about the Storage Pool for File in the properties page as shown in [Figure 21](#).

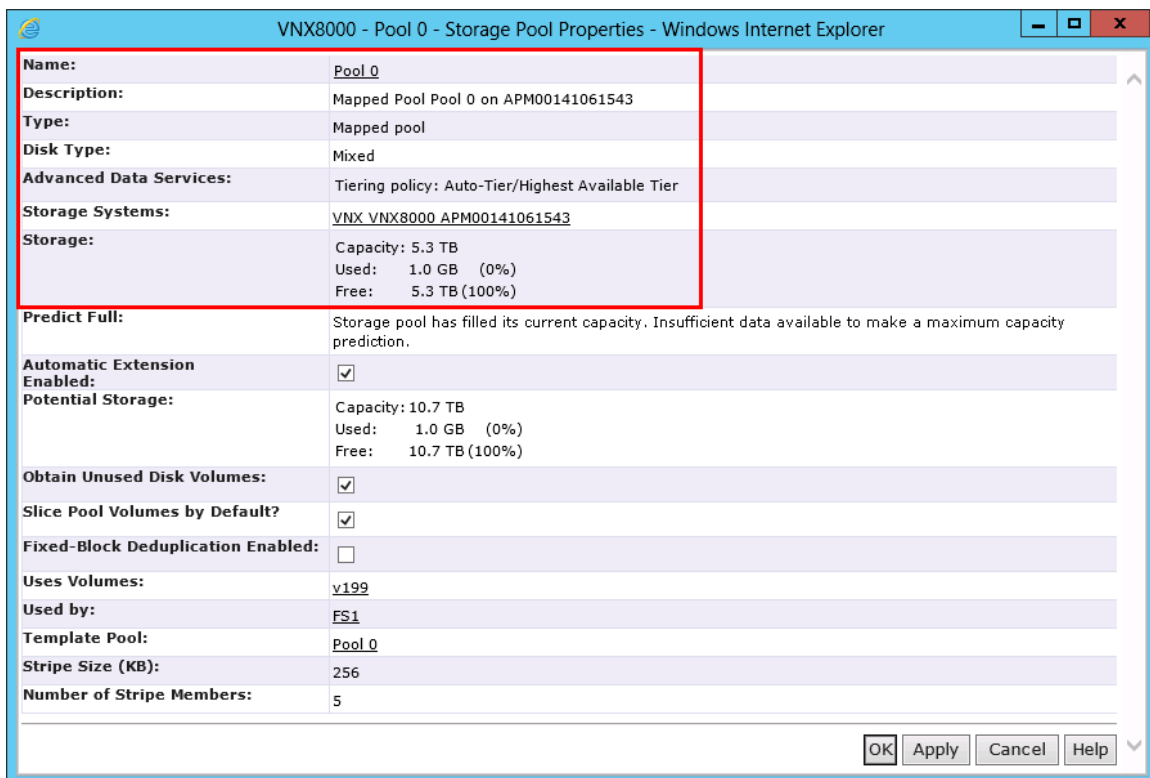


Figure 21 Storage Pool for File properties page

Once the Storage Pool for File is available, you can use it to start creating File Systems. Thin-enabled File Systems can be created with both pool-based and classic LUNs as the underlying storage. Creating a thin-enabled File System is similar to creating a thick File System in Unisphere. [Figure 22](#) shows the Create File System page which can be used to create thin-enabled File Systems.

The screenshot shows a web-based configuration interface for creating a file system. The 'Storage Capacity' field is set to 1 GB. The 'Auto Extend Enabled' and 'Thin Enabled' checkboxes are checked. The 'High Water Mark' is set to 90%. The 'Maximum Capacity (MB)' is set to 2097152. The 'File System Name' is 'FS1'. The 'Storage Pool' is 'Pool 0 10.7 TB (11263990 MB)'. The 'Data Mover (R/W)' is 'server_6'. The 'Mount Point' is 'Default'.

Figure 22 Create File System page

First, you enter the Storage Capacity value as you would for any new File System. Once you check the Thin Enabled checkbox, the Auto-Extend Enabled checkbox will also automatically be checked. This allows you to control when the File System extends and the maximum size it will grow to. You must enter the High Water Mark for auto-extension and the Maximum Capacity, which specifies the File System capacity that is visible to the client.

After creating the File System, 2 TB will be visible to the host as shown in [Figure 23](#), even though VNX has actually only allocated 1 GB of storage. The File System will auto-extend when it reaches the 90% full threshold until it reaches the Maximum Capacity of 2 TB.

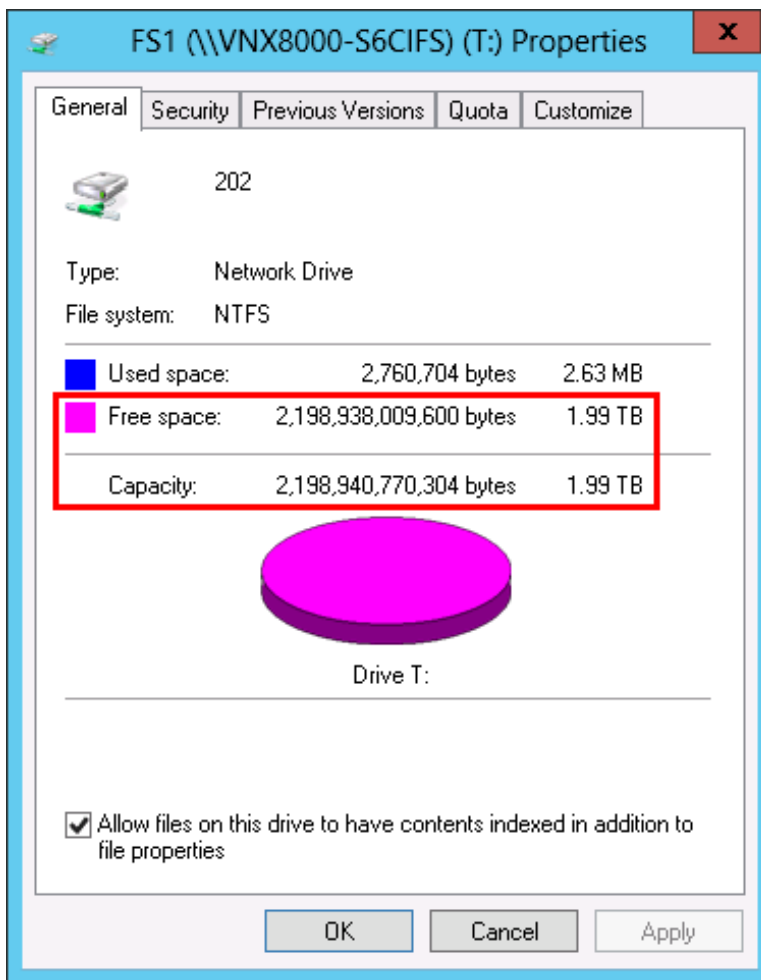


Figure 23 10 TB thin-enabled File System properties

Monitoring thin-enabled File Systems

Just as with thin LUNs, it is crucial to track the utilization of thin-enabled File Systems to avoid running out of space. Enabling thin with automatic File System extension does not automatically reserve space from the pool for that File System.

Administrators must ensure that adequate storage space exists so that the automatic extension operation can succeed. If the available storage is less than the amount of storage required to extend the file system, automatic extension fails. In this case, the user receives an error message when the file system runs out of space. Note that the user may get this message even though it appears that there is free space in their file system.

There are several methods to proactively monitor the utilization of thin-enabled File Systems and the Storage Pool for Files on which they were created. You can use Unisphere to configure proactive alerts when a thin-enabled File System or Storage Pool for File is nearly full. You can configure two types of storage used notifications:

- Current size – How much of the currently allocated File System or Storage Pool for File capacity is used

- Maximum size – How much of the configured maximum File System or Storage Pool for File capacity is used

Alert notifications can be configured to log an event to the event log, send an email, or generate a Simple Network Management Protocol (SNMP) trap. Refer to the *Configuring Events and Notifications on VNX for File* document for more information on setting up event notifications.

Prediction information on when a thin file system is expected to be full is also available. **Figure 24** shows the information that is provided on the properties page of a thin-enabled File System.

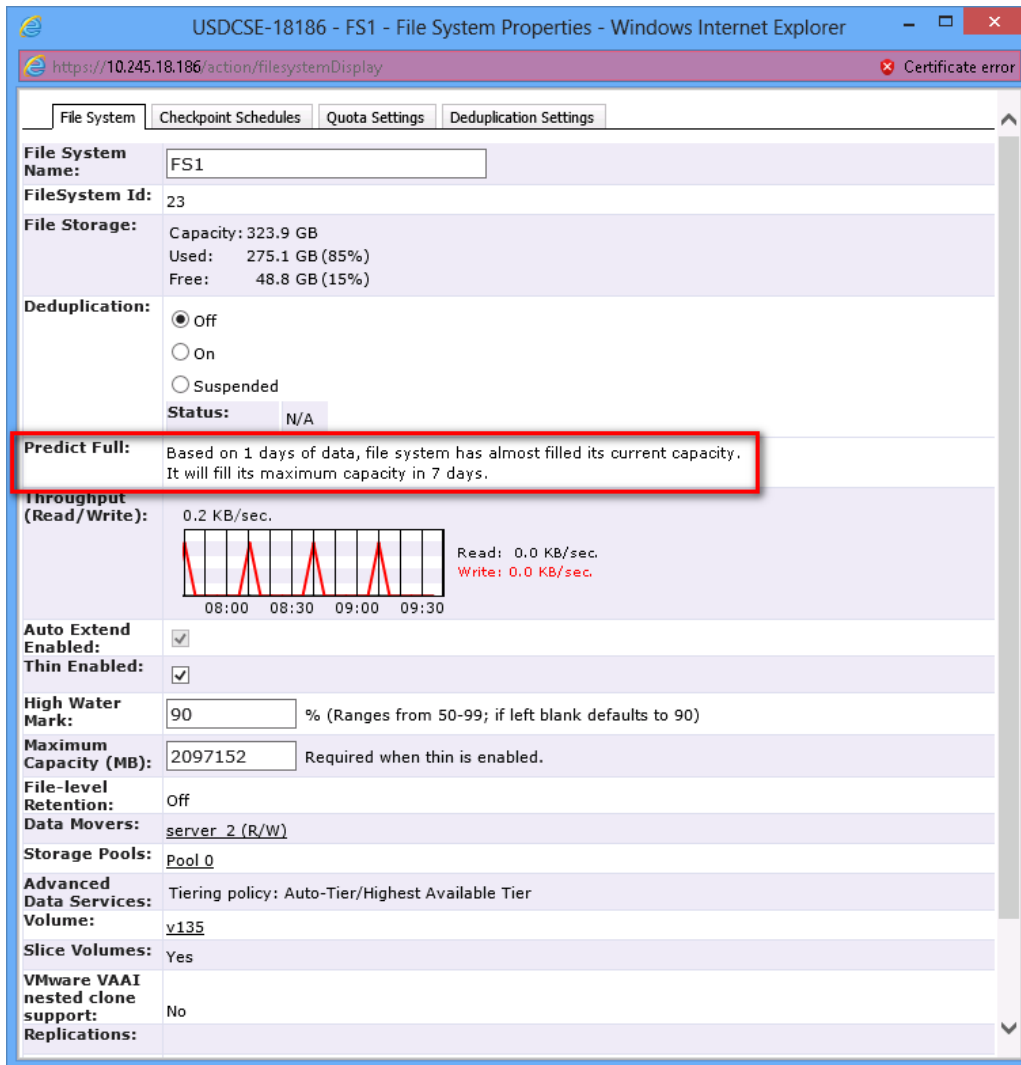


Figure 24 Thin-enabled File System Properties

Proper tracking of storage usage allows you to provision more storage when you need it and helps avoid shortages that could impact users and applications.

Conclusion

When implemented appropriately, Virtual Provisioning can be a powerful complement to an organization's processes and technologies for improving ease of use and utilizing storage capacity more efficiently. VNX Virtual Provisioning integrates well with existing management and business continuity technologies, and is an important advancement in capabilities for VNX customers.

VNX Virtual Provisioning:

- Saves time
 - Easy to create and expand pool LUNs and File Systems
 - Easy to monitor and manage
- Reduces provisioning uncertainty
 - Decisions are easy to modify
 - No impact on host servers
 - Load balances across the pool
- Reduces initial investment and saves energy
 - Highly space-efficient
 - Multiple applications share resources
 - Physical storage can be added as required
- Supports existing VNX features
 - Migration is supported among all types of LUNs
 - VNX Snapshots, SnapView snapshots and clones, SnapSure checkpoints, Replicator, MirrorView/S, MirrorView/A, SAN Copy
 - Unisphere Analyzer

References

- EMC Online Support
 - EMC Virtual Provisioning Release Notes
 - EMC VNX Unified Best Practices for Performance — Applied Best Practices Guide
 - EMC VNX2 FAST VP — A Detailed Review
 - EMC VNX2 Deduplication and Compression
 - Managing Volumes and File Systems with VNX AVM

Appendix A: Thin LUNs

Thin LUN space reclamation via Migration

Unused storage is often locked by applications that do not need the storage that was originally allocated to them. Space reclamation allows the system to regain allocated storage that is not used. This feature works with LUN migration, SAN Copy™, or PowerPath® Migration Enabler (PPME) to migrate from thick to thin LUNs. Since thin LUNs only consume storage to which data is written, all the allocated but unused storage is returned to the pool so that it can be used by other LUNs in the pool. The process is completely transparent and allows you to move your applications without requiring any downtime.

Space reclamation occurs automatically when performing a SAN Copy pull operation on source volumes from other VNX, CLARiiON®, Symmetrix®, and supported third-party systems to a thin destination LUN on VNX.

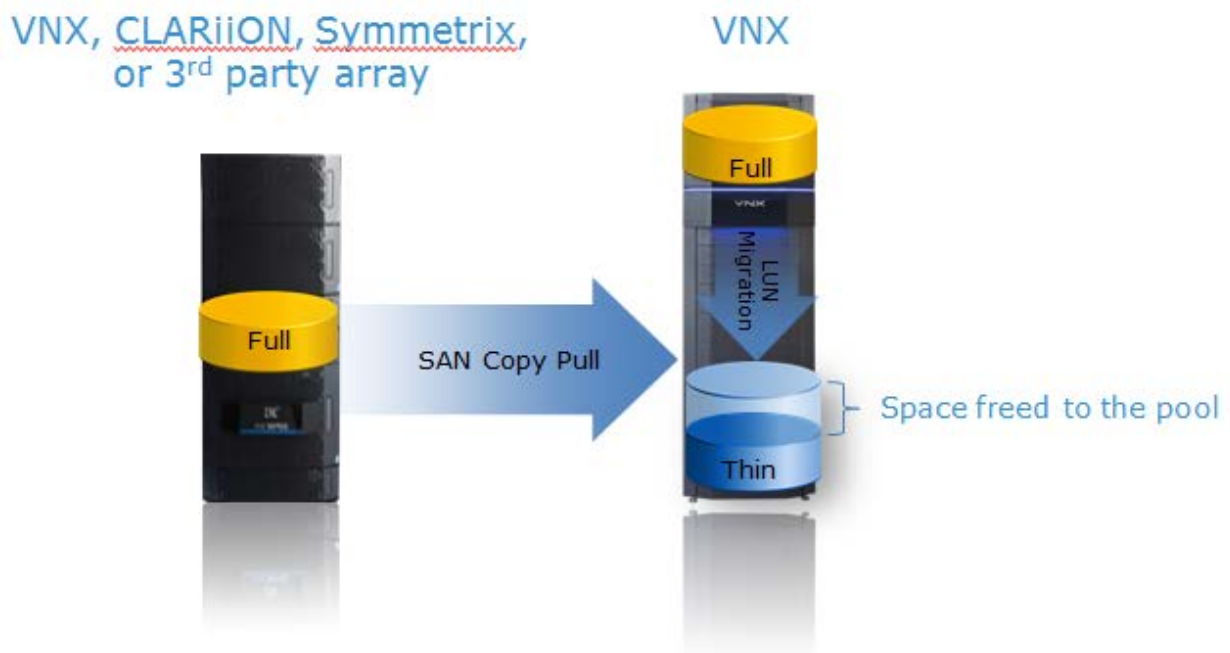


Figure 25 Space reclamation

Space reclamation also occurs when you perform a LUN migration to move an existing classic LUN or thick LUN to a thin LUN within the same array. The software detects zeros at 8 KB chunk granularity. For example, it will only migrate 8 KB chunks with data in them. If the 8 KB chunk is filled with zeros, it is not migrated to a thin LUN and the space is freed.

Another option is to use EMC PowerPath® Migration Enabler (PPME). This is a host-based migration tool that enables non-disruptive or minimally disruptive data migration between storage systems or between logical units within a single storage system. The Host Copy technology in PPME works with the host operating system to

migrate data from the source LUN to the thin target LUN. Host Copy migrations for thin LUNs are supported on Windows, Linux, AIX, and Solaris hosts connected to VNX.

When using space reclamation with NTFS, it is important to note that NTFS doesn't actually remove data from disk after files are deleted. Instead, the data is overwritten when the space is required. Since the migration software processes data at the bit level, it has no notion of a file. Therefore, the data of deleted files is processed the same way as active, non-deleted files. This can significantly reduce the benefits received from space reclamation. The Microsoft SDelete utility for NTFS offers an easy way to permanently delete files prior to performing space reclamation. Running the utility with the -c option replaces deleted files with zeros thereby allowing the space reclamation processes to work effectively.

Using thin LUNs with applications

Due to the storage-on-demand feature of thin LUNs, not all application environments are well suited to thin LUNs. In general, it is a best practice to use thin-friendly applications that do not pre-allocate all of the storage space during initialization. Thin-friendly applications also reuse deleted space before consuming additional storage. The following are guidelines for using thin LUNs with applications most commonly used with VNX.

Host-based File Systems

When creating a file system on a thin LUN, you need to consider how much metadata is written to the thin LUN. An inefficient file system, which writes a lot of metadata to the LUN, causes thin devices to become fully allocated more quickly than with an efficient file system.

Another thin-friendly feature is the ability to effectively reuse deleted space. When you create a new file on the host file system, depending on the file system, it may or may not use the space freed up by a deleted file. If it writes the new file in the previously reclaimed area, then the thin LUN does not consume more space from the pool. However, if it writes in a previously unwritten area, then the thin LUN consumes more space from the pool.

With NTFS, some operations are thin-friendly. For example, an NTFS format does not pre-allocate the physical space. Instead, it creates file system metadata that only consumes a few gigabytes of metadata on the thin LUN. NTFS writes new data on the LUN and updates its own metadata accordingly. However, NTFS is not very effective when it comes to reusing deleted space.

Other file systems like Linux ext2 and ext3, Solaris ZFS and UFS, and Symantec VxFS do not pre-allocate all the storage and also effectively reuse deleted space before allocating new space and thus work nicely with thin LUNs. For the latest list of file systems supported with VNX Virtual Provisioning, refer to the *EMC Support Matrix* at on EMC Online Support.

VMware

For VMware environments, the Virtual Machine File System (VMFS) has many characteristics that are thin-friendly. First, a minimal number of thin extents are allocated from the pool when a VMware file system is created on thin LUNs. Also, a VMFS Datastore reuses previously allocated blocks that are beneficial to thin LUNs. When using RDM volumes, the file system or device created on the guest OS dictates whether the RDM volume is thin-friendly.

When creating a VMware virtual disk, LUNs can be provisioned as:

- Thick Provision Lazy Zeroed
- Thick Provision Eager Zeroed
- Thin Provision

Thick Provision Lazy Zeroed is the default and recommended virtual disk type for thin LUNs. When using this method, the storage required for the virtual disk is reserved in the Datastore, but the VMware kernel does not initialize all the blocks at creation.

The VMware kernel also provides other mechanisms for creating virtual drives that are not thin-friendly. The Thick Provision Eager Zeroed format is not recommended for thin LUNs because it performs a write to every block of the virtual disk at creation. This results in equivalent storage use in the thin pool.

When using Thin Provision, space required for the virtual disk is not allocated at creation. Instead, it is allocated and zeroed out on demand.

As of vSphere 5, there is also the ability to perform thin LUN space reclamation at the storage system level. VMFS 5 uses the SCSI UNMAP command to return space to the storage pool when created on thin LUNs. SCSI UNMAP is used any time VMFS 5 deletes a file, such as Storage vMotion, delete VM, delete snapshot, etc. Earlier versions of VMFS would only return the capacity at the file system level. vSphere 5 greatly simplifies the process by conducting space reclaim automatically.

In addition, features such as VMware DRS, Converter, VM Clones, Storage vMotion, Cold Migration, Templates, and vMotion are thin-friendly.

Hyper-V

In Hyper-V environments, EMC recommends that you select the dynamically expanding option for the Virtual Hard Disk (VHD) when used with VNX thin LUNs as this preserves disk resources. When using pass-through volumes, the file system or guest OS dictates whether the volume will be thin-friendly. For more information on using Hyper-V Server, see the *Using EMC CLARiiON with Microsoft Hyper-V Server* white paper on EMC Online Support.

Appendix B: Low Space and Out of Space Handling Interoperability

The following charts outline which operations are allowed or blocked while under Low Space Handling and Out of Space Handling.

Low Space Handling

Operations	Allowed?	Notes:
File System	Create/Mount/Extend	No If issued before Low Space Handling occurred, the operations will run. Operation may cause Out of Space.
	Delete/Unmount	Yes
	Read Operation to the FS	Yes
	Write Operation to the FS	No
Virtual Data Movers	Create	No
	Delete	Yes If no File Systems exist on the Virtual Data Mover.
RepV2	Create/Start/Failover/Switchover/Reverse/Modify/Refresh/Delete	No For operations started but not completed before Low Space was encountered, the operations will be aborted.
	Stop	Yes
SnapSure	Create/Modify	No If issued before Low Space Handling occurred, the operations will run. Operation may cause Out of Space.
	Delete	Yes Deleting of the oldest check point is allowed. If issued before Low Space Handling occurred, the operations will run.
	Restore/Refresh	No If issued before Low Space Handling occurred, the operations will run.
FSCK	Start	No If issued before Low Space Handling occurred, the operations will run.
Quotas	On/Off/Report/Edit/Check/Tree List	No If issued before Low Space Handling occurred, the operations will run.
VNX Block	Enable/Disable	Yes

Deduplication for VNX File			
File Deduplication and Compression	Enabled/Disable	No	If issued before Low Space Handling occurred, it will be suspended until after the File System is mounted Read Only.
	Suspend	No	If issued before Low Space Handling occurred, it will be cancelled.
DHSM	Offline/Enable/Disable	No	If issued before Low Space Handling occurred, the operations will run.
	Recall	No	If issued before Low Space Handling occurred, the operations will fail.
NDMP Backup	File System (Source)	Yes	
	SnapSure Checkpoint (Source)	No	If issued before Low Space Handling occurred, the operations will run.
	Restore	No	If issued before Low Space Handling occurred, the operations will run
VAAI VMDK Clone/Instant Provision/Snap		No	If issued before Low Space Handling occurred, the operations will run.
CAVA		No	If issued before Low Space Handling occurred, the operations will fail.

Out of Space Handling

While under an Out of Space condition, none of the following operations are allowed on the affected storage. The following operations will be allowed once the Out of Space condition is cleared and access has been restored.

- File System
 - Create/Mount/Delete/Unmount/Extend/Read or Write to FS
- Virtual Data Movers
 - Create/Delete
- RepV2
 - Create/Start/Stop/Failover/Switchover/Reverse/Modify/Refresh/Delete
- SnapSure
 - Create/Modify/Delete/Restore/Refresh
- FSCK
 - Start
- Quota
 - On/Off/Report/Edit/Check/Tree List
- File Deduplication and Compression
 - Enable/Disable/Suspend
- DHSM
 - Offline/Enable/Disable/Recall
- NDMP
 - File System (Source)/SnapSure Checkpoint (Source)/Restore
- VAAI VMDK Clone/Instant Provision/Snap
- CAVA