White Paper

EMC VNX2 FAST VP

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

A Detailed Review

Abstract

This white paper discusses EMC® Fully Automated Storage Tiering for Virtual Pools (FAST VP™) technology and describes its features and implementation. Details on how to use the product in Unisphere™ are discussed, and usage guidance and major customer benefits are also included.

March 2016

Part Number H12102.3

Table of Contents

Executive summary

Fully Automated Storage Tiering for Virtual Pools (FAST VP™) can lower the Total Cost of Ownership (TCO) and increase performance by intelligently managing data placement. When FAST VP is implemented, the storage system measures, analyzes, and implements a dynamic storage-tiering policy in a faster and more efficient way than a human analyst.

Storage provisioning can be repetitive and time consuming, and estimates calculated incorrectly can produce uncertain results. It is not always obvious how to match capacity to the performance requirements of a workload's data. Even when a match is achieved, requirements change, and a storage system's provisioning will require constant adjustments.

Storage tiering allows a storage pool to use varying levels of drives. The LUNs within the storage pool can use the necessary storage capacity for various applications while meeting some of the required performance goals. FAST VP uses IO statistics at a 256MB slice granularity (known as *sub-LUN tiering*). It uses the relative activity level of each slice to determine the need to promote to higher tiers of storage. For arrays containing Flash drives, this creates a Flash-optimized array. Relocation is initiated at your discretion through either manual initiation or an automated scheduler. FAST VP removes the need for manual, resource intensive, LUN migrations, while still providing the performance levels required by the most active dataset.

FAST VP is a licensed feature available on the VNX®2 series. FAST VP licenses are available for both individual platforms or as part of the FAST Suite, which includes licenses for both FAST VP and Multicore FAST Cache.

This white paper discusses the EMC FAST VP technology and describes its features, functions, and management. It assumes that the FAST VP enabler is installed on the array.

Audience

This white paper is intended for EMC customers, partners, and employees who are considering using the FAST VP product. Some familiarity with EMC midrange storage systems is assumed. You should be familiar with the material discussed in the white papers *Introduction to the EMC VNX2 Series* and *Virtual Provisioning for the VNX2 Series.*

Terminology

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.

Command line interface (CLI) - Interface for typing commands through the Control Station to perform tasks that include the management and configuration of the database and Data Movers, and the monitoring of statistics for VNX for File cabinet components.

Diskmark - A disk identification number on VNX for File.

File System - Method of cataloging and managing the files and directories on a system.

Fully Automated Storage Tiering Virtual Pools (FAST VP) - Automatically assigns different categories of data to different types of storage media within a tiered pool. Data categories can be based on performance requirements, frequency of use, cost, and other considerations. This is done by retaining the most frequently accessed or important data on fast, high-performance (more expensive) drives, and moving the less frequently accessed and less important data to lower-performance (less expensive) drives.

Graphical User Interface (GUI) - Software that uses graphical objects such as pull-down menus and operations such as drag-and-drop to allow the user to enter commands and execute functions.

Load Balance - Relocation of slices across and within tiers based on relative slice temperature.

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

Redundant Array of Independent Disks (RAID) - Method for storing information where the data is stored on multiple disk drives to increase performance and storage capacities and to provide redundancy and fault tolerance.

RAID Group - Group of disks that take on a RAID configuration

Rebalance - Automated data relocation to improve the data imbalance when new, empty drives are combined with relatively full, existing drives in a pool.

Skew - Small percentage of capacity that is responsible for most of the IO activity.

Slice - Minimum increment of capacity that can be allocated to a pool LUN. Pool LUNs are comprised of slices.

Slice Relocation - Movement of 256MB blocks of data within or between tiers, based on the relative activity level (temperature) of each block.

Storage Pool for Block - 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.

Temperature - Moving average of a slice's activity level over time.

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

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

Tier - Different categories of media in a storage pool. These categories provide various levels of performance and capacity through several drive types. Available tiers are Extreme Performance, Performance, and Capacity.

Introduction

Data has a lifecycle. As data progresses through its lifecycle, it experiences varying levels of activity. When data is created, it is typically heavily used. As data ages, it is accessed less often. Fully Automated Storage Tiering for Virtual Pools (FAST VP) is a simple and elegant solution for dynamically matching storage requirements with changes in the frequency of data access. FAST VP segregates disk drives into three categories, called tiers:

- **Extreme Performance Tier** – Flash drives
- **Performance Tier** – Serial Attach SCSI (SAS) drives
- **Capacity Tier** – Near-Line SAS (NL-SAS) drives

FAST VP is a feature designed to aggressively reduce Total Cost of Ownership (TCO) and maintain performance. Rather than using only one type of drive, mixing Flash, SAS, and NL-SAS drives allows you to maintain performance requirements while reducing the drive count. In some cases, a nearly two-thirds reduction in drive count can be achieved, while in other cases, performance throughput can double by simply adding less than 10 percent of a pool's total capacity to Flash drives.

FAST VP has proven highly effective for a number of applications. Tests in OLTP environments with Oracle[1] and Microsoft SQL Server[2] show that using FAST VP with a heterogeneous pool containing Flash drives instead of a homogenous pool can lower capital expenditure by 15 to 38 percent, reduce power and cooling costs by over 40 percent, and still increase performance. For more information about these benefits, see the "What drive mix is right for my IO profile?" section of this paper.

The VNX series of storage systems delivers high value by providing a unified approach to auto-tiering for file and block data. Both block and file data can use virtual pools and FAST VP. This provides compelling value for users who want to optimize the use of high-performance drives across their environment.

You can use FAST VP in combination with other performance optimization software, such as Multicore FAST Cache. A common strategy is to use FAST VP to gain TCO benefits while using Multicore FAST Cache to boost overall system performance. There are also several scenarios in which it makes sense to use FAST VP for both purposes. This paper discusses considerations for the best deployment of FAST VP.

[1] *Leveraging Fully Automated Storage Tiering (FAST) with Oracle Database Applications* EMC white paper

[2] *EMC Tiered Storage for Microsoft SQL Server 2008 – Enabled by EMC Unified Storage and EMC Fully Automated Tiering (FAST)* EMC white paper

Using FAST VP for Block

Storage Pools

Storage pools are the framework that allows FAST VP to take full advantage of different drive types. A pool is somewhat analogous to a RAID group, which is a physical collection of drives on which logical units (LUNs) are created. Pools 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 more resources and requires minimal planning and management efforts.

Pools can be homogeneous or heterogeneous. Homogeneous pools have a single drive type (Flash, SAS, or NL-SAS), whereas heterogeneous pools have different drive types.

Homogeneous pools

Homogeneous pools are recommended for applications with limited skew, such that their access profiles can be very random across a large address range. Multiple LUNs with similar profiles can share the same pool resources. These LUNs provide more predictable performance based on the drive type employed. In a homogeneous pool, only one drive type (Flash, SAS, or NL-SAS) is selected during pool creation.

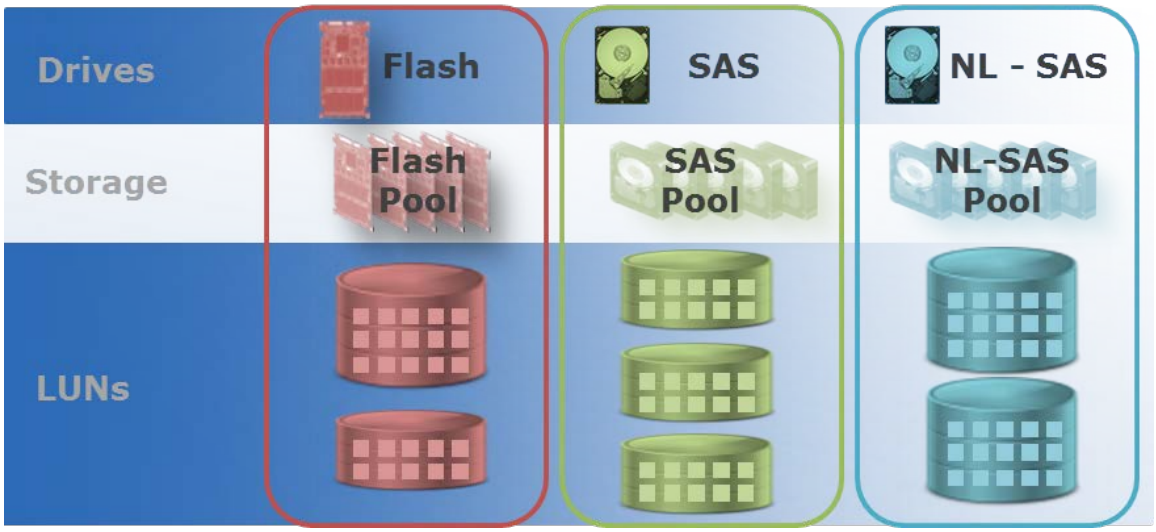


Figure 1. Homogeneous Storage Pools

As seen in Figure 1, each pool contains various sized LUNs that consist of one particular drive type.

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.

Heterogeneous pools

Heterogeneous pools consist of LUNs with different drive types. VNX supports Flash, SAS, and NL-SAS drives in one pool. As shown in Figure 2, there can be a maximum of three drive types in a heterogeneous pool. Data in a particular LUN can reside on some or all of the different drive types.

Figure 2. Heterogeneous Storage Pools

Storage Tiers

When creating storage pools on a VNX, data can use three different categories of media in a single pool. These categories, referred to as storage tiers, provide various levels of performance and capacity through several drive types. The available tiers when creating a storage pool are:

- Extreme Performance
- Performance
- Capacity

FAST VP differentiates each of these tiers by drive type, but it does not take rotational speed into consideration. EMC strongly encourages that you avoid mixing rotational speeds per drive type in a given pool. If multiple rotational-speed drives exist in the array, you should implement multiple pools as well.

FAST VP can leverage one, two, or all three storage tiers in a single pool. Each tier offers unique advantages in terms of performance and cost.

Per-Tier RAID Configuration

During storage pool creation, you can select RAID protection on a per-tier basis. Each tier has a single RAID type, and once the RAID configuration is set for that tier in the pool, it cannot be changed. Table 1 shows the RAID configurations that are supported for each tier.

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

Because RAID 5 (8+1) and RAID 6 (14+2) have high data to parity ratios, they provide 50% savings compared to RAID 5 (4+1) and RAID 6 (6+2), respectively. The tradeoff for higher data to parity ratios are larger fault domains and potentially longer rebuild times. This is especially true for RAID 5, which has only a single parity drive. EMC advises you to choose carefully between RAID 5 (4+1) and RAID 5 (8+1) and decide whether robustness or efficiency is a higher priority. Robustness is less likely to be an issue for RAID 6, because it has two parity drives.

When creating a storage pool with any of the available tiers, EMC recommends that you add drives in multiples of the preferred RAID configuration. If you select a drive count that is not a multiple of the preferred RAID configuration, the system creates non-preferred drive groups, which can affect performance.

For best practice recommendations, refer to the *EMC Unified Storage Best Practices for Performance and Availability – Common Platform and Block Storage* white paper on EMC Online Support.

Extreme Performance Tier

Use the Extreme Performance tier when response times and performance are the most important criteria for storage. The Extreme Performance tier uses Flash technology – solid-state drives (SSDs) that contain no moving parts. This revolutionizing technology eliminates rotational latencies and can lead to a performance boost and significant energy savings. The VNX2 series also supports SAS Flash 2 Flash drives. Compared to SAS Flash drives, SAS Flash 2 Flash drives are more cost effective and appropriate to use when data change rates are moderate.

Tests show that adding a small percentage of Flash capacity to storage and using intelligent tiering products, such as FAST VP, can deliver double-digit percentage gains in throughput and response time performance in some applications.

Flash drives have a higher per-GB cost but a lower per IO cost compared to traditional spinning drives. To receive the best return, use Flash drives for data that requires fast response times and high IOPS. FAST VP enables you to optimize the use of these high-performance resources, because it automatically relocates “hot” data to the Flash drives at a sub-LUN level.

Performance Tier

Use the Performance tier to achieve a combination of performance and capacity. This tier, composed of Serial Attach SCSI (SAS) drives, offers high levels of performance, reliability, and capacity. SAS drives are based on industry-standardized, enterprise-level, mechanical hard-drive technology that stores digital data on a series of rapidly rotating magnetic platters.

The Performance tier includes both 10K and 15K RPM spinning drives, which are available on all EMC midrange storage systems. When adding these drives to a storage pool, EMC recommends using a single rotational speed per tier. The Performance tier is valuable, because it offers high, all-around performance with consistent response times, high throughput, and good bandwidth at a mid-level price point.

Capacity Tier

Use the Capacity tier to decrease the cost per GB of data. This tier, consisting of 7.2K RPM Near-Line SAS (NL-SAS) drives, is designed for maximum capacity at a modest performance level. Although NL-SAS drives have a slower rotational speed compared to drives in the Performance tier, NL-SAS drives can significantly reduce energy use and free up capacity in the more expensive and higher performing storage tiers.

In a typical system, up to 95% of application data has little IO activity. Since NL-SAS drives cost less than performance drives on a per-GB basis, and their cost is a small fraction of the cost of Flash drives, they are the most appropriate type of media for this "cold" data. NL-SAS drives consume 96% less power per terabyte than performance drives and offer a compelling opportunity for TCO improvement that considers both purchase cost and operational efficiency.

Note: For more information about the latest supported drives, refer to the *Introduction to the EMC VNX2 Series – VNX5200, VNX5400, VNX5600, VNX5800, VNX7600, & VNX8000* white paper found on EMC Online Support.

Considerations for Using Tiers

When all three tiers are available, there are considerations for using one tier or another. For example, response times are lower with Flash drives compared to NL-SAS drives, but capacity is greater in the Capacity tier as compared to the Extreme Performance tier. Table 2 compares the three tiers.

Table 2. Comparison of the Extreme Performance, Performance, & Capacity Tiers

	Extreme Performance (Flash)	Performance (SAS)	Capacity (NL-SAS)
Performance	User Reponse Time		
	1-5 ms	≈5 ms	7-10 ms
	Multi-Access Reponse Time		
	‹ 10 ms	10-50 ms	≤ 100ms
	High IOPS/GB and Low Latency	High bandwidth with contending workloads	Low IOPS/GB
			Leverages storage array SP cache for sequential and large block access
Strengths	Provides extremely fast access for reads	Sequential reads leverage read-ahead	
		Sequential writes leverage system optimizations favoring disks	
	Executes multiple sequential streams better than SAS	Read/write mixes provide predictable performance	Large IO is serviced efficiently
Observations	Writes slower than reads	Uncached writes are slower than reads	Long response times for heavy-write loads
	Heavy concurrent writes affect read rates		
	Single-threaded, large, sequential IO is equivalent to SAS		

FAST VP LUN Management

Although FAST VP is a feature that is implemented for the entire storage pool, you can modify settings at the LUN level. During the creation of a LUN via the GUI or Navisphere Secure Command Line Utility (*naviseccli*), you can define a tiering policy. A tiering policy specifies where the initial placement of data will be and how that data will be relocated within the pool during scheduled and manually invoked relocation periods. FAST VP bases decisions for how data relocation occurs on performance statistics collected every hour. For more information, see the "FAST VP Algorithm" section in this paper.

Tiering Policies

FAST VP is a completely automated feature that implements a set of user-defined tiering polices to ensure the best performance for various business needs. These requirements are met by using FAST VP tiering policies, which determine how new allocations and ongoing relocations should apply within a storage pool. The relocations are done for each set of logically ordered blocks, called slices. FAST VP uses an algorithm to make data relocation decisions based on the activity level of each slice. It ranks the order of data relocation across all LUNs within each separate pool. The system uses this information in combination with the per LUN tiering policy to create a candidate list for data movement.

The following LUN level policies are available:

- Highest Available Tier
- Auto-Tier
- Start High then Auto-Tier
- Lowest Available Tier
- No Data Movement

Users can set all LUN level policies except the "No Data Movement" policy both during and after LUN creation. The "No Data Movement" policy is only available after LUN creation.

Highest Available Tier

Use the "Highest Available Tier" policy when quick response times are a priority. This tier is effective for LUNs which, although not always the most active, require high levels of performance whenever they are accessed. The "Highest Available Tier" policy starts with the "hottest" slices first and places them in the highest available tier until the tier's capacity or performance capability limit is hit. Then it places the slices into the next highest tier.

Auto-Tier

A small portion of a large set of data may be responsible for most of the IO activity in a system. FAST VP allows for moving a small percentage of the "hot" data to higher tiers while maintaining the rest of the data in the lower tiers. The "Auto-Tier" policy automatically relocates data to the most appropriate tier based on the activity level of each data slice. Slices provisioned to a LUN are relocated based on the highest performance disk drives available and the LUN's slice temperature. Although this setting relocates data based on the LUN's performance statistics, LUNs set with "Highest available Tier" take precedence.

Start High then Auto-Tier (Recommended)

"Start High then Auto-Tier" is the default policy for each newly created pool LUN. This is the recommended policy, because it takes advantage of the "Highest Available Tier" and "Auto-Tier" policies. "Start High then Auto-Tier" sets the preferred tier for initial data allocation to the highest performing disk drives with available space, and then it relocates the LUN's data based on the performance statistics and the auto-tiering algorithm. This tiering policy maximizes initial performance and takes full advantage of the highest performing drives while improving TCO. With this tiering policy, less active data is moved to lower tiers, making room for more active data in the higher tiers. This tiering policy is especially good for LUNs that exhibit skew.

Lowest Available Tier

Use the "Lowest Available Tier" policy when cost effectiveness is the highest priority. With this policy, data is initially placed on the lowest available tier with capacity. Select this policy for LUNs that are not performance sensitive or response-time sensitive. Regardless of their activity level, all slices of these LUNs will remain on the lowest storage tier available in their pool. Slices with "colder" temperatures in the LUN have higher priority to be relocated to the lowest tier.

No Data Movement

Select the "No Data Movement" policy after a LUN is created. If a LUN is configured with this policy, no slices provisioned to the LUN are relocated across tiers. Data remains in its current position, but can still be relocated within the tier. The system still collects statistics on these slices after the tiering policy is changed.

Initial Slice Placement within a Tier

The tiering policies listed above affect the initial placement of a LUN's slices within the available tier. For example, initial placement with a pool set to "Auto-Tier" results in the data being distributed across all storage tiers available in the tier. The distribution of the data is based on the available capacity of the pool. If 70% of a pool's free capacity resides in the lowest available tier, then 70% of the new slices are placed in that tier. LUNs set to "Highest Available Tier" and "Lowest Available Tier" have their slices placed in the highest or lowest tiers, respectively, that have capacity available. Since the "No Data Movement" policy can only be applied after a LUN is created, initial placement is set to the policy that was used prior to the change to no movement.

Summary of Tiering Policies

Table 3 summarizes the tiering policies.

Table 3. Tiering Policies

Tiering Policy	Corresponding Initial Tier Placement	Description
Highest Available Tier	Highest Available Tier	Sets the preferred tier for the initial data placement and subsequent data relocation (if applicable) to the highest-performing disk drives with available space.
Auto-Tier	Optimize for Pool Performance	Sets the initial data placement to Optimized for Pool and then relocates the LUN's data based on the LUN's performance statistics such that data is relocated among tiers according to I/O activity.
Start High then Auto-Tier (Default)	Highest Available Tier	First sets the preferred tier for the initial data placement to the highest-performing disk drives with available space, then relocates the LUN's data based on the LUN's performance statistics and the auto-tiering algorithm.
Lowest Available Tier	Lowest Available Tier	Sets the preferred tier for the initial data placement and subsequent data relocation (if applicable) to the most cost-effective disk drives with available space.
No Data Movement	Based on previous FAST VP policy used	Prevents any ongoing data movement for this LUN. The current placement of data is retained. Only available as a LUN property after the LUN is created.

Managing FAST VP at the storage pool level

Once you have created LUNs with the most appropriate tiering policy, you can view and manage FAST VP properties at the storage pool level. When viewing the properties of a specific pool, the Tiering tab provides summary information about the amount of data targeted to be moved up, down, and within tiers. This tab also provides the ability to choose to use manual or scheduled data relocation for the pool.

Figure 3 shows these properties.

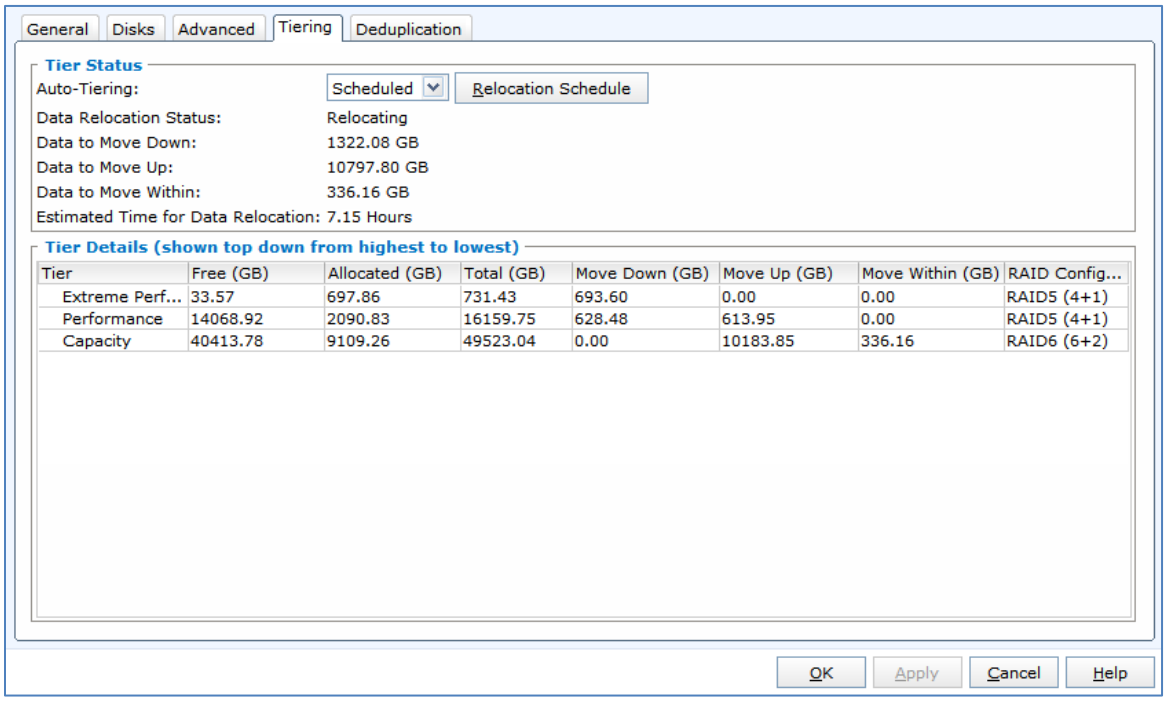


Figure 3. Storage Pool Tiering Properties

In the storage pool tiering properties window, users can view both the tier status and the tier details for the pool. The Tier Status section enables you to define the method for starting data relocation for the pool. Choose "Manual" to manually start the data relocation on a selected storage pool at any time. Choose "Scheduled" to have the system perform data relocation based on the settings defined in the relocation schedule. Click the **Relocation Schedule** button to bring up the **Manage Auto Tiering** window and view this schedule.

The Tier Status section displays the following information:

- Data Relocation Status, which can be:
  - Ready (no active data relocations)
  - Relocating (data relocations are in progress)
  - Paused (all data relocations for this system are paused)
- Quantity of pool data, in GB, that is ready to move up, down, or be redistributed within existing tiers for load balancing.

  This is the amount of data to move during the next scheduled data relocation. This number will change, since a recalculation is done right before the relocation window starts. If the Data Relocation Status is *Relocating*, this quantity is the amount of data to be moved during the current data relocation.
- Estimated time for the data relocation based on the data gathered by the software.

The Tier Details section displays the following information:

- Data distribution per tier, including the quantity of free capacity, allocated capacity per tier, and total capacity.
- Amount of data that is ready to be moved up, down, or redistributed within the tier.
- RAID type for each tier.

Relocation Schedule

The FAST VP feature allows for automatic data relocation based on a user-defined relocation schedule. This schedule defines when and how frequently the array starts data relocation on the participating storage pools. These scheduled relocations, which take place on all pools concurrently, can be arranged to take place during off-hours to minimize any adverse performance impact. The **Manage Auto-Tiering** window accepts changes to be made to the schedule. You can launch this window from the storage pool tiering properties or from the link in the task pane located in the Storage section of Unisphere.

Figure 4 depicts the **Manage Auto-Tiering** window.

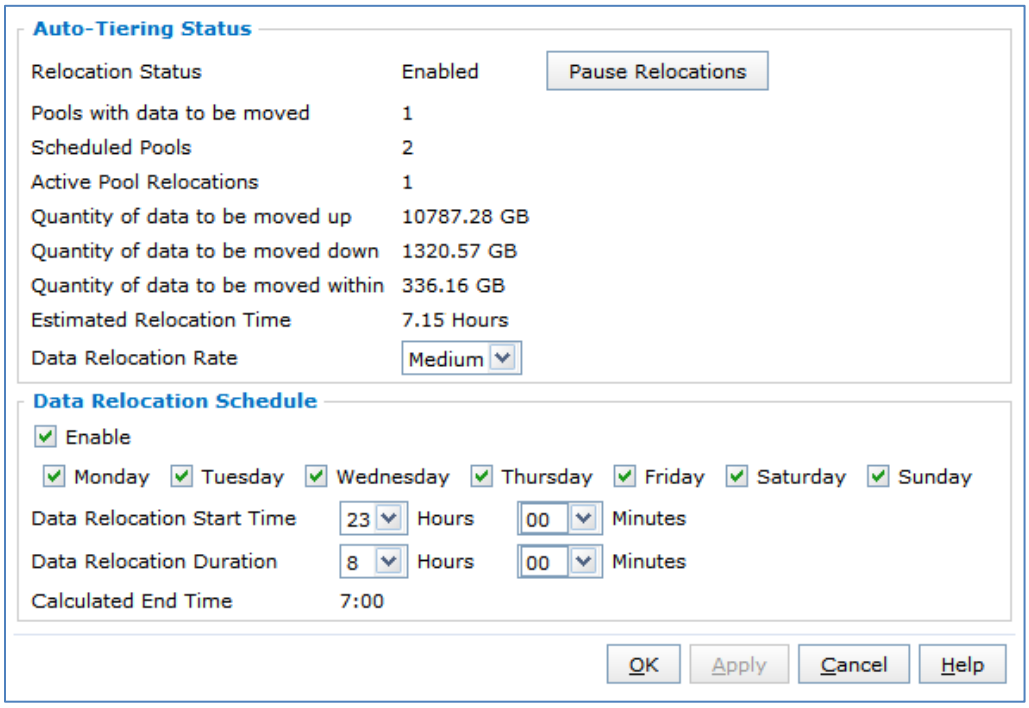


Figure 4. Manage Auto-Tiering Window

In the **Manage Auto-Tiering** window, you can view both the current auto-tiering status and the data relocation schedule. The Auto-Tiering Status section indicates whether data relocation is enabled or paused. When data relocation is paused, all active relocation tasks for all storage pools pause, meaning no future relocation requests

will be issued. Use the button to the right of the **Relocation Status** field to pause or resume data relocations.

This section also displays the "Pools with data to be moved" field. This is the number of tiered pools that have data queued for relocation. "Scheduled Pools" refers to the number of tiered pools associated with the data relocation schedule. "Active Pool Relocations" is the number of tiered pools currently participating in manual or scheduled data relocation operations. The Auto-Tiering Status section also shows the total amount of data in all tiered storage pools that is ready to move up to the next higher tier, down to the next lower tier, and/or be redistributed on the disks for load balancing, along with the estimated time (in hours) for completing all data relocation operations.

The Data Relocation Schedule section enables users to define the operational rate for the data relocation operation. Valid values are **Low**, **Medium** (default), or **High**, where **Low** has the least impact on system performance, and **High** has the most impact.

This section also gives users the ability to disable/enable the schedule and change the relocation start time, relocation duration, and days on which relocations can take place. Multiple relocation operations can occur during this timeframe, and the software assigns this schedule to all participating pools. In the example shown in Figure 4, data relocations are allowed to run each day of the week for an eight hour window, starting at 23:00 (11:00 PM) and concluding at 7:00 AM.

By default, the data relocation state for a created pool is set to **Scheduled**. As Figure 5 shows, users can right-click on a pool to change the data relocation state to **Manual**, and can stop or start data relocation for the pool.

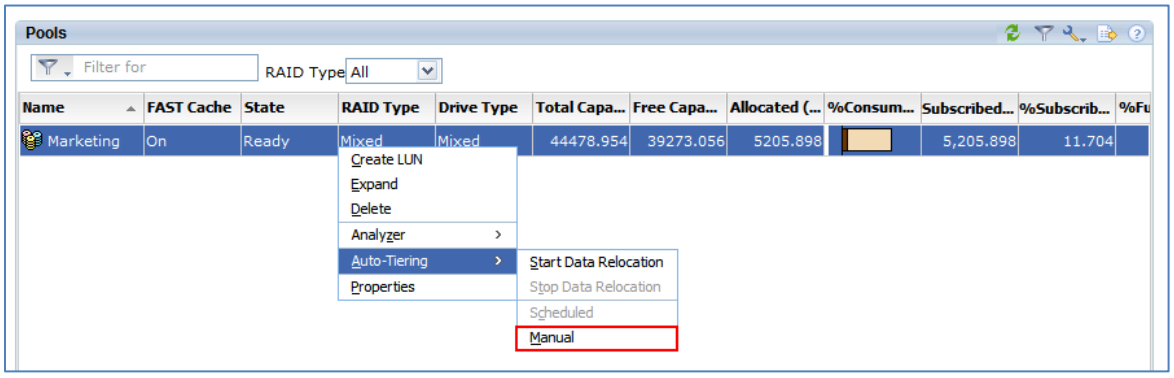


Figure 5. Scheduled Auto-Tiering

Manual Relocation

Users can initiate manual data relocation for selected pools, whether or not they are included in the auto-tiering schedule, at any time using the GUI or *naviseccli*. FAST VP analyzes all gathered statistics prior to beginning the relocation to ensure that up-to-date statistics and settings are properly accounted for prior to relocation. If the relocation completes within the defined timeframe, FAST VP recollects statistics and continues relocation until the timeframe closes.

When users select **Start Data Relocation** in the GUI, they can set the data relocation rate and duration. The data relocation rate can be set to **Low**, **Medium**, or **High**. The default is the relocation rate set in the **Manage Auto-Tiering** window. (See Figure 4.) Users can also set the relocation duration, which specifies the time, in hours and minutes, at which data relocations can occur.

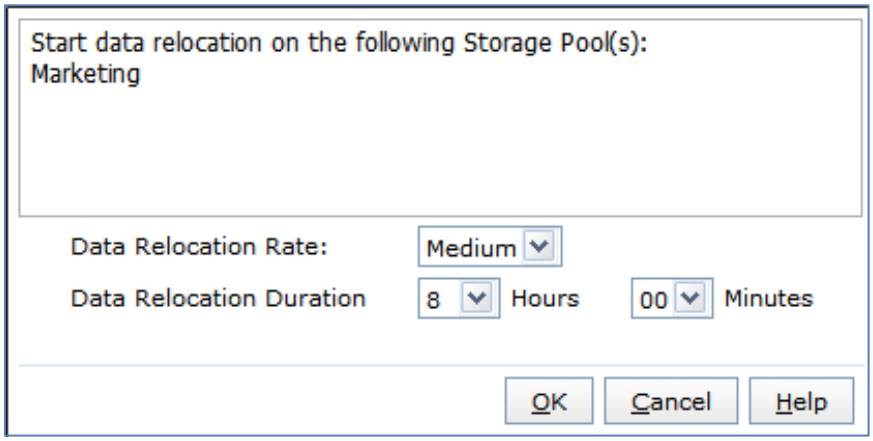


Figure 6. Start Data Relocation

Although the automatic scheduler is an array-wide setting, the manual relocation operation allows for immediate relocation to take place at the pool level. Common situations for initiating a manual relocation include:

- When reconfiguring a pool, if LUN properties, such as the tiering policy, change, and the new priority structure must be realized immediately
- As part of a script for a finer-grained relocation schedule

Using FAST VP for File

Management

The process of implementing FAST VP for file storage begins by provisioning LUNs from a block storage pool with mixed tiers. The provisioned LUNs are then added to the ~filestorage Storage Group. After performing a rescan of the storage systems via the link on the Storage page in Unisphere, a diskmark operation is started that presents the newly created storage to a new file storage pool. The rescan automatically creates and maps the storage pool for File to the block storage pool using the same pool name. Additionally, it creates disk volumes in a 1:1 mapping for each LUN that was added to the ~filestorage Storage Group. At this point, you can create file systems from the disk volumes contained in the new storage pool for File.

The FAST VP policy that was applied to the LUNs presented to the storage pool for File operates the same as with other LUNs in the system, dynamically migrating data between storage tiers in the pool.

Figure 7 illustrates the FAST VP for File workflow.

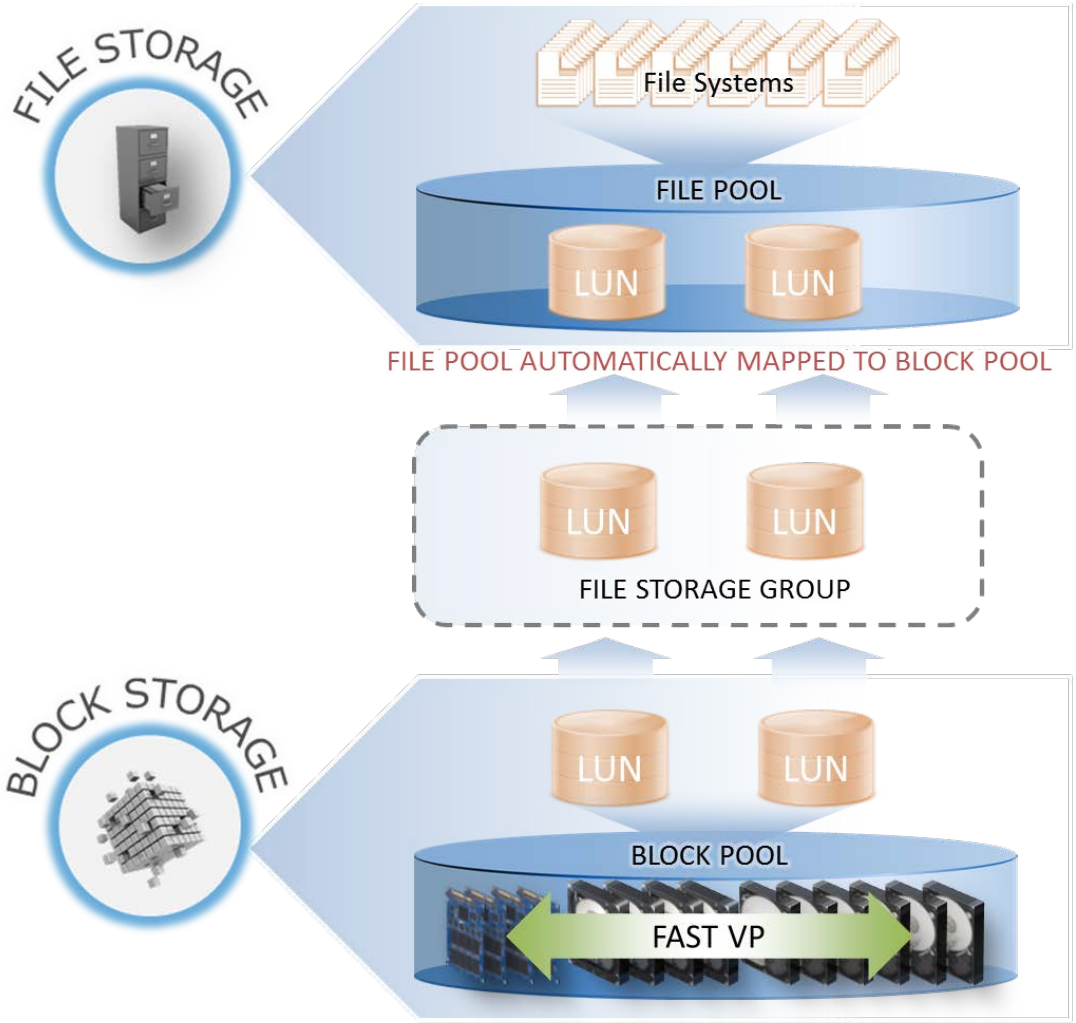


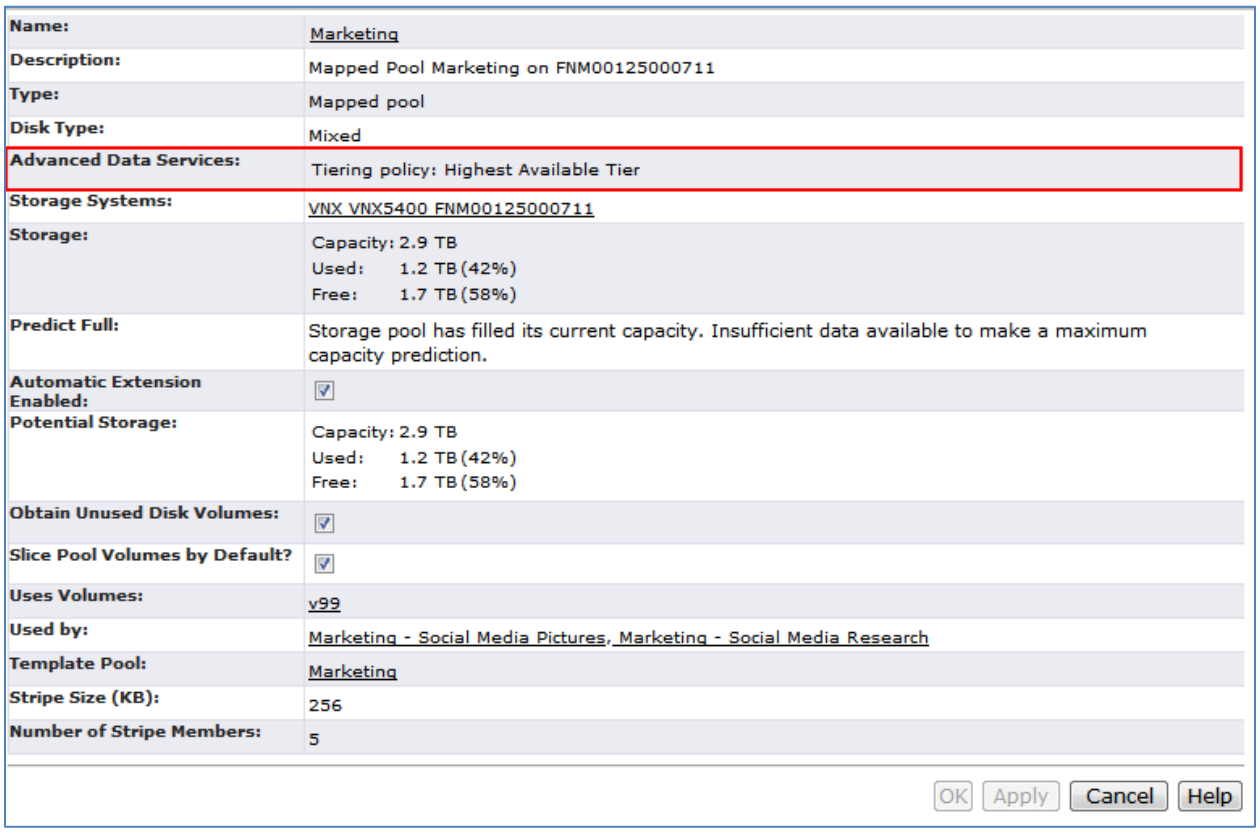
Figure 7. FAST VP for File Workflow

FAST VP for File is supported for both Unisphere and the CLI. Its properties can be seen in various areas throughout Unisphere, including on the properties page of pools for File and the properties page for both volumes and file systems. The Advanced Data Services section lists the tiering policy for the LUN along with other enabled data services, including values for whether the LUN is thin and/or mirrored, and whether its data is compressed. This section displays only if the storage pool on which the file system is created is a tiered pool, or if the file system is created directly from pool-based LUNs using manual volume management. If Advanced Data Services are not enabled, the tiering policy does not appear.

For more information about other Advanced Data services, refer to the *Introduction to EMC VNX Storage Efficiency Technologies* white paper found on EMC's Online Support.

Creating a File System using FAST VP LUNs

As outlined in the diagram above, you can create file systems once LUNs are presented to the ~filestorage Storage Group, and a storage pool for File is automatically mapped to the block storage. After the new storage pool for File is created, you can view the corresponding tiering policy in the properties window for that particular pool, as shown in Figure 8.



Name:	Marketing
Description:	Mapped Pool Marketing on FNM00125000711
Type:	Mapped pool
Disk Type:	Mixed
Advanced Data Services:	Tiering policy: Highest Available Tier
Storage Systems:	VNX VNX5400 FNM00125000711
Storage:	Capacity: 2.9 TB Used: 1.2 TB (42%) Free: 1.7 TB (58%)
Predict Full:	Storage pool has filled its current capacity. Insufficient data available to make a maximum capacity prediction.
Automatic Extension Enabled:	☑
Potential Storage:	Capacity: 2.9 TB Used: 1.2 TB (42%) Free: 1.7 TB (58%)
Obtain Unused Disk Volumes:	☑
Slice Pool Volumes by Default?	☑
Uses Volumes:	v99
Used by:	Marketing - Social Media Pictures, Marketing - Social Media Research
Template Pool:	Marketing
Stripe Size (KB):	256
Number of Stripe Members:	5

OK Apply Cancel Help

Figure 8. Properties Window for a Storage Pool for File

As seen Figure 8, the tiering policy is displayed in the Advanced Data Services section. If multiple LUNs with different tiering policies are present in the pool, a tiering policy of "Mixed" is displayed.

Similarly, once you create a file system, the tiering policy is also displayed in the properties window of that particular file system, as shown in Figure 9.

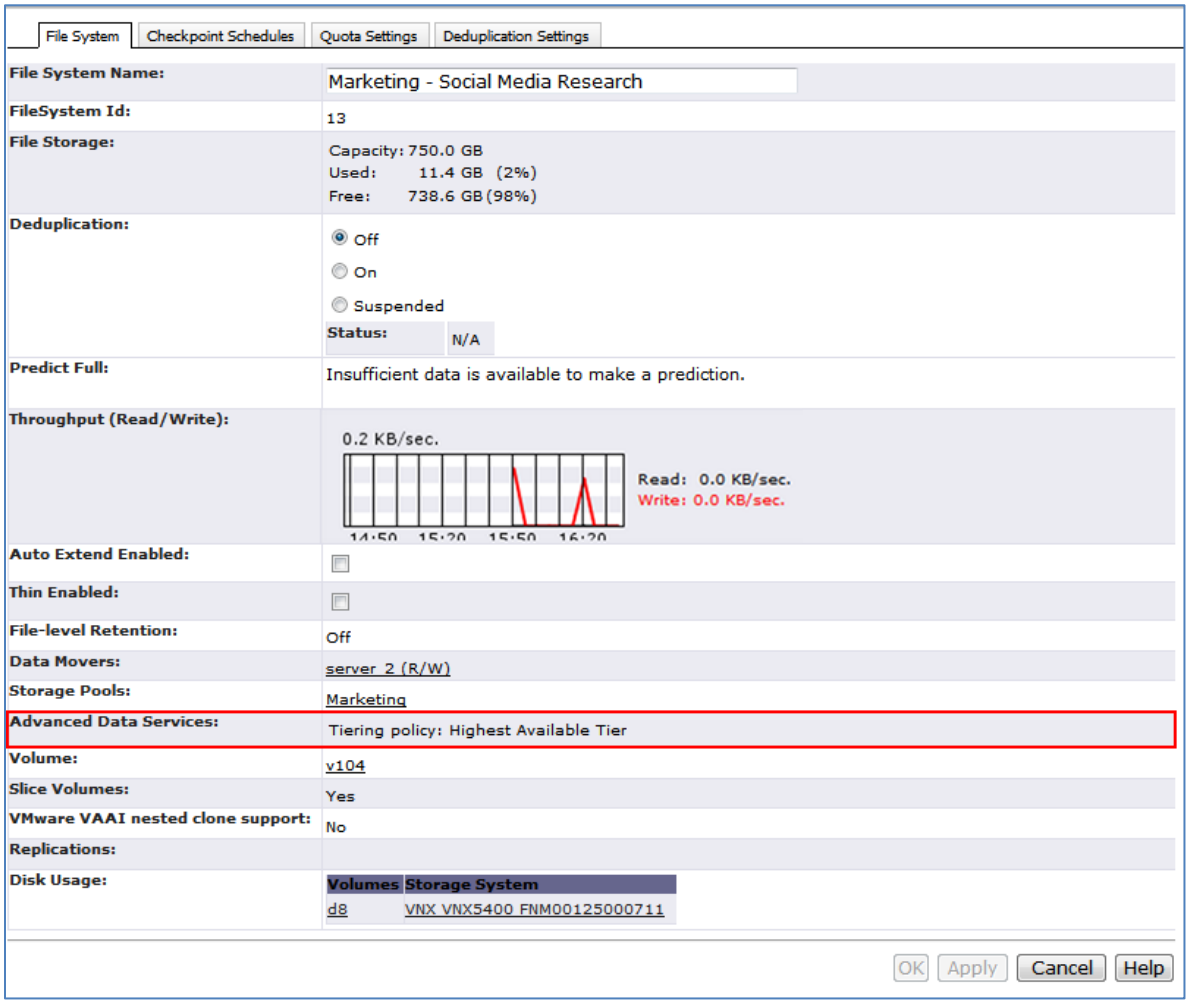


Figure 9. File System Properties Window

As seen in Figure 9, you can view all of the properties of the file system along with the tiering policy.

For applications requiring consistent and predictable performance, EMC **strongly** recommends that you create File Systems using thick LUNs with Block Deduplication disabled. In VNX OE for Block Release 33, the default LUN type for a new LUN is thin; this can be accomplished by unchecking the 'Thin' checkbox when creating the LUN. Block Deduplication will be disabled by default. Deduplication and compression optimization can be achieved using the thin and File deduplication (which includes compression) attributes at the File System level. Also, when creating a storage pool, the entire pool should be allocated to file systems, and it is recommended to:

- Create approximately 1 LUN for every 4 drives in the storage pool.
- Create LUNs in even multiples of 10.

- Number of LUNs = (number of drives in pool divided by 4), rounded up to nearest multiple of 10.

- Make all LUNs the same size.
- Balance LUN ownership across SPA and SPB.
- Leave 5% available in the pool for relocations and recovery operations.
- Have the same FAST VP tiering policy for all LUNs in a given VNX OE for File storage pool.
  - Create a user-defined storage pool to separate VNX OE for File LUNs from the same Block storage pool that have different tiering policies.
- Use Thin Enabled File Systems, when using FAST VP with VNX OE for File, for increased benefit from FAST VP multi-tiering.

When using MirrorView for file replication, it is important to remember that VNX file configurations will not allow mixing of mirrored and non-mirrored types in a pool. If you try to do this, the disk mark will fail.

Note: For more information about best practices for VNX for File, refer to the *Applied Best Practices Guide: EMC VNX Unified Best Practices* found on EMC's Online Support.

FAST VP Operations

FAST VP operates by relocating the most active data up to the highest available tier (either the Extreme Performance or Performance Tier). To ensure sufficient space in the higher tiers, relocations attempt to reclaim 10% free space in the tiers to allow for new slice allocation that occurs when new LUNs are created or when thin LUNs consume additional capacity. By reclaiming this 10% headroom, the least active slices within each tier move to lower tiers (Performance or Capacity).

FAST VP works at a granularity of 256 MB and relocates data by moving the entire slice to the appropriate tier, depending on the tiering policy selected for that particular LUN.

FAST VP Algorithm

FAST VP uses three different strategies to improve performance, capacity, and TCO. These techniques help identify and relocate slices to the most appropriate tiers by collecting statistics on each slice, analyzing the data, and relocating each slice based on its activity level.

Statistics Collection

As previously noted, a slice of data is considered "hotter" (more active) or "colder" (less active) than another slice of data based on the relative activity level of those slices. The activity level of a particular slice is determined by counting the number of IOs, reads and writes, bound for each slice. FAST VP maintains a cumulative IO count

and “weighs” each IO according to how recently it arrived. This weight deteriorates over time, and newer IOs are given a higher weight. After approximately 24 hours, the weights of new IOs are nearly cut in half and continue to decrease. This statistics collection occurs continuously in the background for all pool LUNs.

Analysis

FAST VP analyzes the collected data once per hour. This analysis process produces a ranking order from “hottest” to “coldest” for each slice within the pool. Before relocation is invoked with automatic or manual relocation, FAST VP performs a final calculation and creates a candidate list of slices to move up, down, and within a pool. The ranking of a LUN and its slices can be influenced by changing the tiering policy, in which case the tiering policy takes precedence over activity level.

Relocation

During the user-defined relocation window, FAST VP promotes slices according to the candidate list that it created in the analysis stage. During relocation, FAST VP prioritizes relocating slices to higher tiers. Slices are only relocated to lower tiers if the space they occupy is required for a higher priority slice. This way, FAST VP ensures that the higher performing drives are always used.

After data is added to a pool, FAST VP attempts to move it to the higher tiers if space is available and the tiering policy allows for it. The relocation process aims to keep 10% of the space in all tiers free. This creates space for any new slice allocations of higher priority LUNs before the next relocation. Lower tiers are used for capacity as the need arises. This entire relocation process is done automatically based on the user-defined relocation schedule, or manually, if user-initiated. Figure 10 provides an illustration of how FAST VP can improve slice placement for a pool.

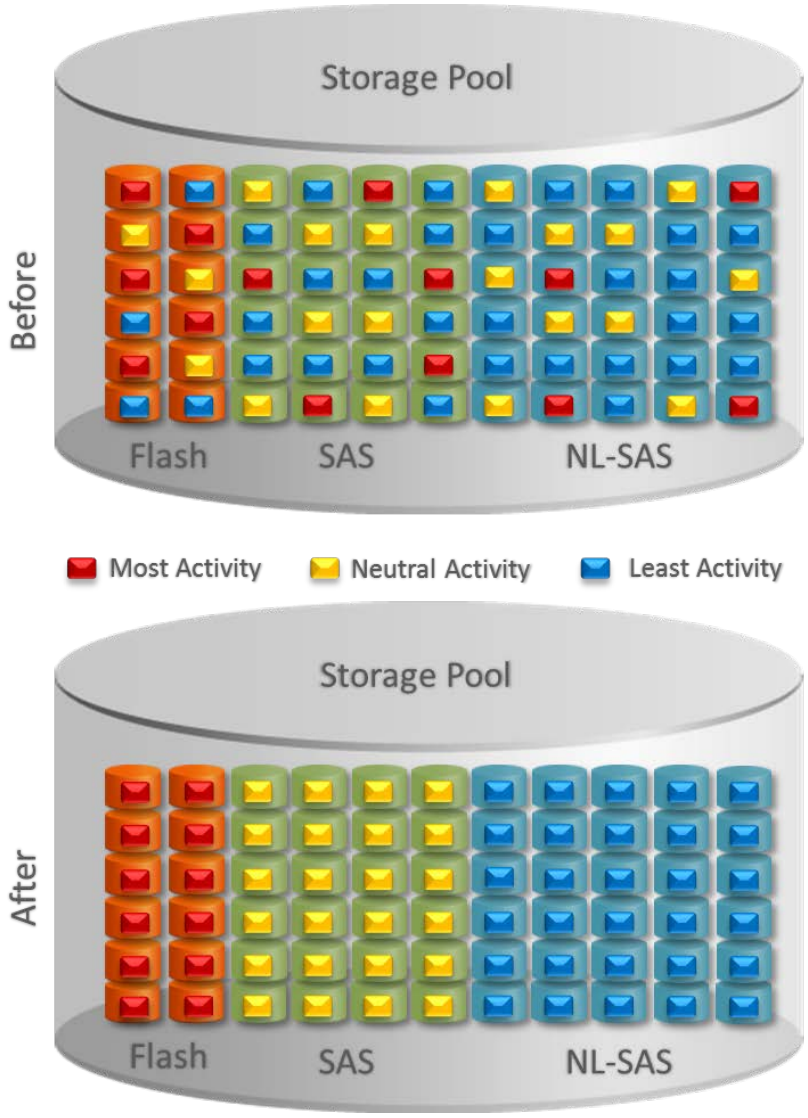


Figure 10. FAST VP Slice Relocation

As seen in Figure 10, before FAST VP is installed on the array, data ranging from very active to not active at all resides in a pool across each tier. The more expensive and higher performing drives may contain non-active data that would be more appropriate in the Capacity tier, for which the cost per GB is less expensive. Once FAST VP is installed, the array automatically collects and analyzes statistics and relocates data across tiers to better utilize the more expensive drives. FAST VP moves the most active data to a higher tier and the least active data to a lower tier.

Rebalance

Upon the expansion of a storage pool, the system recognizes the newly added space and initiates an auto-tiering data relocation operation called rebalancing. A rebalance relocates slices within the whole pool to achieve the best performance. If the pool is heterogeneous, relocation happens not only within each tier, but also across tiers.

When a rebalance is initiated and running, Unisphere displays a percentage number to indicate the progress. Figure 11 shows this progress report, which is displayed in the pool properties.

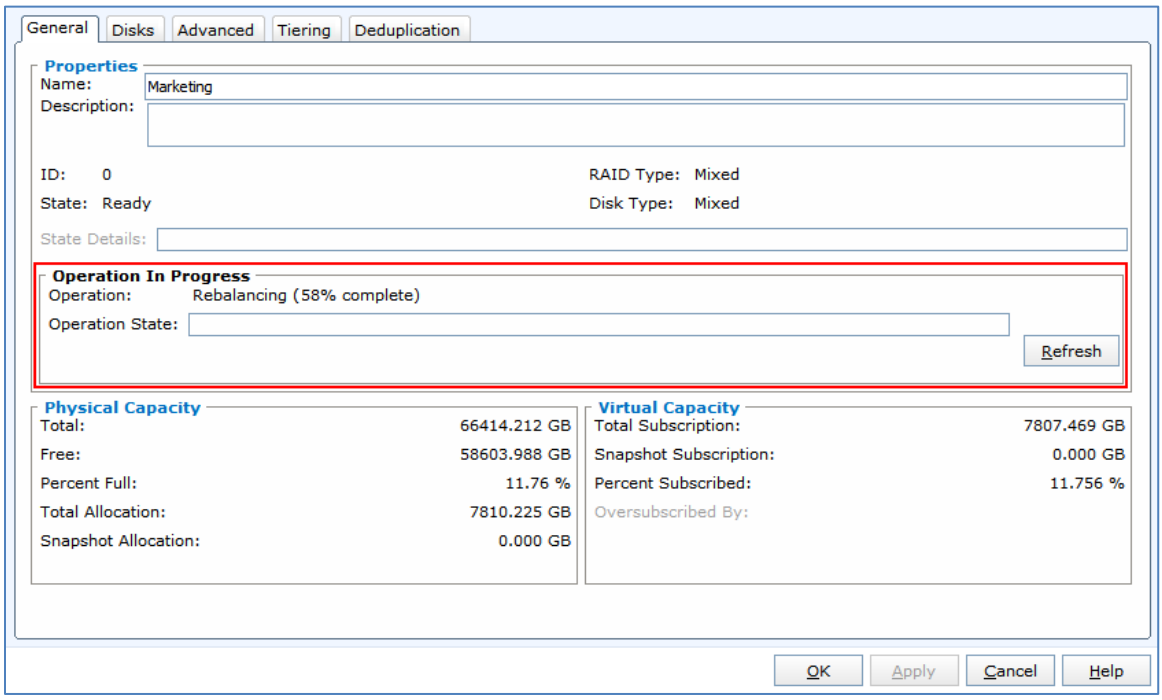


Figure 11. Rebalance in Progress

Figure 11 shows a rebalance operation that is 58% complete. When the pool is expanding, the progress report also shows the operation state, which can be **New**, **Running**, or **Failed**. There is no relocation window for a rebalance operation, which runs until all slices in the relocation candidate list have been processed.

If you do not want an automatic rebalance to occur, because the newly added drives will be used for a special purposed LUN, you can issue a rebalance skip using *naviseccli*. For example:

```
storagepool -expand -id 0 -disks 0_1_5-0_1_10 –skipRules
```

Note: For a more detailed synopsis of this command and others, refer to the *VNX Command Line Interface Reference for File* documentation, located on EMC Online

Load-Balance

In addition to relocating slices across tiers based on relative slice temperature, FAST VP can also relocate slices within a tier to achieve maximum performance across drives in that tier. Some drives within a tier may be more heavily used than others. To improve performance, data slices are relocated within the tier to balance the load, if the tier is not 100% full. Maintaining 5% or more of free space allows for load-balancing to work efficiently. This is accomplished by extending auto-tiering data relocation to also analyze and move data within a tier.

General Guidance and Recommendations for FAST VP and Multicore FAST Cache

As noted previously, the FAST Suite consists of FAST VP and Multicore FAST Cache. These two products are packaged together, work together, and complement each other. Multicore FAST Cache allows the storage system to provide Flash drive class performance to the most heavily accessed chunks of data across the entire system. Multicore FAST Cache absorbs IO bursts from applications, thereby reducing the load on back-end hard drives. This improves the performance of the storage solution. For more details on Multicore FAST Cache, refer to the *EMC VNX Multi-Core FAST Cache* white paper found on EMC Online support. Table 4 compares FAST VP and Multicore FAST Cache.

Table 4. Comparison between FAST VP and Multicore FAST Cache Features

Multicore FAST Cache	FAST VP
Allows Flash drives to be used to extend the existing caching capacity of the storage system.	Allows a single LUN to leverage the advantages of multiple drive types through the use of storage pools.
Granularity is 64 KB.	Granularity is 256 MB.
Data that is accessed frequently is copied from HDDs to Flash drives.	Data is moved between different storage tiers based on weighted-average-of-access statistics collected over a period of time.
Used when workload changes are unpredictable and very dynamic, and require a quick response time.	Used when workload pattern changes are predictable and relatively low.
Constantly promotes frequently accessed HDD data to Multicore FAST Cache. There are no relocation cycles.	Data movement occurs in scheduled or manually invoked relocation windows.
Real-time monitoring decides which data needs to be promoted to Multicore FAST Cache.	Hourly analysis decides which portion of data needs to be moved.

Multicore FAST Cache and FAST VP can be used together for sub-LUN tiering features to yield high performance and improved TCO for the storage system. For example, in scenarios where limited Flash drives are available, the Flash drives can be used to create the Multicore FAST Cache, and FAST VP can be used on a two-tier pool

(Performance and Capacity). From a performance point of view, Multicore FAST Cache dynamically provides performance benefits to bursts of data, while FAST VP moves "hotter" data to higher tiers. From a TCO perspective, Multicore FAST Cache, with a small number of Flash drives, serves the data that is accessed most frequently, while FAST VP optimizes disk utilization and efficiency.

As a general rule, use Multicore FAST Cache when storage system performance needs to be improved immediately for burst-prone data. Multicore FAST Cache complements FAST VP by focusing on improving performance, while FAST VP focuses on improving TCO.

The Multicore FAST Cache feature is storage-tier-aware and works with FAST VP to make sure that storage system resources are not wasted by unnecessarily copying data to Multicore FAST Cache, if that data is already on a Flash drive. If FAST VP moves a chunk of data to the Extreme Performance Tier, Multicore FAST Cache will not promote that chunk of data into Multicore FAST Cache, even if Multicore FAST Cache promotion criteria are met.

A general recommendation for the initial deployment of Flash drives in a storage system is to use them for Multicore FAST Cache. In almost all cases, Multicore FAST Cache, with a 64KB granularity, offers the industry's best optimization of Flash technology.

Large Scale Migrations

EMC recommends that large scale migrations turn off tiering until the migration completes. If critical data is being tiered during or immediately following the migration, EMC recommends using the "Highest Available" tiering policy.

When the migration process starts, it is best to fill the highest tiers of the pool first. This is especially important for live migrations. Using the Auto-Tier setting would place some data in the Capacity tier. At this point, FAST VP has not yet run an analysis on the new data, so it cannot distinguish between "hot" and "cold" data. Therefore, with the Auto-Tier setting, some of the busiest data may be placed in the Capacity tier.

In these cases, you should set the target pool LUNs to the highest tier. This way, all data is initially allocated to the highest tiers in the pool. As the higher tiers fill and capacity from the Capacity (NL-SAS) tier starts to be allocated, you can stop the migration and start manual FAST VP relocation.

Assuming the analysis process has run for a sufficient amount of time, relocation orders the slices by rank and moves data appropriately. In addition, since the relocation attempts to free 10% of the highest tiers, these tiers will have more capacity for new slice allocations.

This iterative process can be used while more data is migrated into the pool. You can start FAST VP to begin relocation when most of the new data is allocated to the Capacity tier. Once all of the data is migrated into the pool, you can select the best tiering policy.

What drive mix is right for my IO profile?

It is common for a small percentage of overall capacity to be responsible for most of the IO activity. Analysis of an IO profile may indicate that 85 percent of the IO activity to a volume only involves 15 percent of the capacity. The resulting active capacity is called the working set. Software like FAST VP and Multicore FAST Cache keeps the working set on the highest-performing drives.

It is common for OLTP environments to yield working sets of 20 percent or less of their total capacity. These profiles hit the sweet spot for FAST VP and Multicore FAST Cache. Other IO profiles, such as Decision Support Systems (DSS), may have much larger working sets. In these cases, you can use FAST VP to deploy Flash drives, because DSS workloads are not typically Multicore-FAST-Cache-friendly. You can use Capacity Tier drives to lower TCO.

At a minimum, the capacity across the Performance Tier and Extreme Performance Tier (and/or Multicore FAST Cache) should accommodate the working set. However, capacity is not the only consideration. The spindle count of these tiers needs to be sized to handle the IO load of the working set.

Drives in the Performance Tier are versatile in handling a wide spectrum of IO profiles. Therefore, EMC highly recommends including Performance Tier drives in each pool. The Multicore FAST Cache can be an effective tool for handling a large percentage of activity, but inevitably, there will be IOs that are not promoted or are cache misses. The Performance Tier drives offer good performance for these IOs.

The Performance Tier drives also facilitate faster promotion of data into the Multicore FAST Cache by quickly providing promoted 64 KB chunks to the Multicore FAST Cache. This minimizes Multicore FAST Cache warm-up time as some data gets “hot” and other data goes “cold”. If the IO profile has a significant component of random writes, these are best served from Performance Tier drives as opposed to Capacity drives.

You can use Capacity drives to optimize TCO. This often equates to comprising 60% to 80% of the pool’s capacity. Of course, there are also profiles with low IOPS/GB and or sequential workloads that may result in the use of a higher percentage of Capacity Tier drives.

You can engage EMC Professional Services and qualified partners for detailed analysis and recommendations for future configurations. They have the tools and expertise to make very specific recommendations for tier composition based on an existing IO profile.

Conclusion

With the use of FAST VP, users can remove complexity and management overhead from their environments. FAST VP utilizes Flash, SAS, and NL-SAS drives (or any combination thereof) within a single pool. LUNs within the pool can then leverage the advantages of each drive type at the 256 MB slice granularity. This sub-LUN-level tiering ensures that the most active dataset resides on the best-performing drive tier

available, while maintaining infrequently used data on lower-cost, high-capacity drives.

Relocations can occur without user interaction on a predetermined schedule, making FAST VP a truly automated offering. In the event that relocation is required on-demand, you can invoke FAST VP relocation on an individual pool using the Unisphere GUI or CLI.

Both FAST VP and Multicore FAST Cache work by placing data segments on the most appropriate storage tier based on their usage pattern. These two solutions are complementary, because they work on different granularity levels and frequency. Implementing both FAST VP and Multicore FAST Cache can significantly improve performance and reduce cost in the storage environment.

References

The following white papers are available on the EMC Online Support website:

- *Introduction to the EMC VNX2 Series – VNX5200 VNX5400, VNX5600, VNX5800, VNX7600, & VNX8000*
- *EMC Tiered Storage for Microsoft SQL Server 2008—Enabled by EMC Unified Storage and EMC Fully Automated Storage Tiering (FAST)*
- *EMC Unified Storage Best Practices for Performance and Availability – Common Platform and Block — Applied Best Practices*
- *EMC Unisphere Unified Storage Management Solution for the VNX2 Series*
- *EMC Virtual Provisioning for the VNX2 Series – Applied Technology*
- *EMC VNX2 Multicore FAST Cache*
- *Introduction to EMC VNX2 Storage Efficiency Technologies*
- *Leveraging EMC Unified Storage System Dynamic LUNs for Data Warehouse Deployments*
- *Leveraging Fully Automated Storage Tiering (FAST) with Oracle Database Applications*