White Paper

EMC FAST for CLARiiON

A Detailed Review

Abstract

This white paper introduces the EMC® CLARiiON® Fully Automated Storage Tiering (FAST) technology and describes its features and implementation. Details on how to work with the product in the Unisphere™ operating environment are discussed, and usage guidance and major customer benefits are also included.

September 2010

Part Number h8058

Table of Contents

Executive summary

Storage provisioning can be repetitive and time-consuming and 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 may require constant adjustment. Storage tiering is one solution. Storage tiering puts several different types of storage devices into an automatically managed storage "pool." LUNs use the storage capacity they need from the pool, on the devices with the performance they need. Fully Automated Storage Tiering (FAST) is the EMC® CLARiiON® feature that allows a single LUN to leverage the advantages of Flash, Fibre Channel, and SATA drives through the use of pools.

FAST solves theses issues by providing automated sub-LUN-level tiering. FAST collects I/O activity statistics at the 1 GB granularity level (known as a slice). The relative activity level of each slice is used to determine which slices should be promoted to higher tiers of storage. Relocation is initiated at the user's discretion through either manual initiation or an automated scheduler.

Through the frequent relocation of 1 GB slices, FAST continuously adjusts to the dynamic nature of modern storage environments. This removes the need for manual, resource-intensive LUN Migrations while still providing the performance levels required by the most active dataset.

FAST functionality

FAST is a powerful new product from EMC available for the CLARiiON CX4 product line. FAST has the unique ability to relocate data from one tier of storage to another within the same pool. This allows sub-LUN slices of data to leverage the advantages of different drive types at different points in their lifecycles.

Pre-provisioning tasks are significantly reduced by adding FAST to a storage environment. LUNs placed into pools have their slices automatically allocated. As performance needs change, FAST continuously relocates data to the appropriate tier.

I/O activity statistics are collected and analyzed by FAST behind the scenes. This analysis is then combined with user-defined tiering preferences to produce a list of slice relocations. Relocations take place either automatically, according to a user-defined schedule, or manually, at the user's discretion.

FAST functionality can be leveraged in combination with FAST Cache technology to produce a high-performing, agile system at a reduced total cost of ownership compared to what traditionally tiered storage systems can offer.

Introduction

This white paper introduces the EMC CLARiiON Fully Automated Storage Tiering (FAST) technology and describes its features and implementation. Details on how to work with the product in the Unisphere™ operating environment are discussed, and usage guidance and major customer benefits are also included.

Audience

This white paper is intended for EMC customers, partners, and employees who are considering the use of the FAST product in CLARiiON storage systems. It is assumed that the reader is familiar with CLARiiON storage systems and their management software. A good introductory paper to CLARiiON-based storage is the *EMC CLARiiON Storage System Fundamentals for Performance and Availability* white paper available on Powerlink®.

Storage tiering

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 it ages, it is accessed less often. This usage pattern, and the efficient storage of data according to its use, has been difficult to address with previous generations of storage technology. FAST is a simple and elegant approach to the changing use of data over time. FAST has the ability to move sub-LUN-level slices of data from one storage tier to another. This occurs automatically and transparently to the host environment. Each CLARiiON supports three storage tiers each using a different physical storage device technology — Flash, Fibre Channel (FC), and SATA. Each tier offers unique advantages. FAST can leverage all three of these tiers at once, or any two at a time.

Advantages of Flash, Fibre Channel, and SATA storage tiers

Different storage tiers offer different benefits. FC and SATA drives are mature storage technologies offering high levels of performance, reliability, and high capacities. These drives are based on industry-standardized, enterprise, mechanical hard-drive technology that stores digital data on a series of rapidly rotating magnetic platters.

Fibre Channel storage devices

10k and 15k rpm FC drives are available on all CLARiiON and Celerra® storage systems. FC drives have been the performance medium of choice for many years. They also have the highest availability of any mechanical storage device. FC drives continue to serve as a valuable storage tier, offering high all-around performance including consistent response times, high throughput, and good bandwidth at a midtier price point.

SATA storage devices

SATA drives are designed for maximum capacity at a modest performance level. SATA drives are slower than FC drives; 7.2k and 5.4k rpm SATA drives are available on CLARiiON and Celerra arrays. The reduced rotational speed is a trade-off for SATA's significantly larger capacity. For example, the largest SATA drives available for CLARiiON systems are 2 TB, compared to the 600 GB FC drives and 400 GB Flash drives. Data should be migrated from FC drives to SATA drives only if it requires modest performance. Migrating to SATA drives can significantly reduce energy use and free capacity in higher storage tiers.

Flash storage devices

Enterprise Flash drives are relatively new to the enterprise storage marketplace and are quickly gaining popularity. They deliver extremely high performance, and are recommended for data that requires fast response times and/or very high throughput (IOPS). Bandwidth-intensive applications perform only slightly better on Flash drives than FC drives. Flash drives are built on solid-state drive (SSD) technology that has no moving parts. The absence of moving parts makes these drives highly energy-efficient, and eliminates rotational latencies. Therefore, migrating data from FC to Flash can both boost performance and create significant energy savings. Enterprise Flash drives have a higher per-GB cost than FC or SATA drives. To receive the best return, you should use Flash drives for data that requires fast response times and/or high IOPS.

Table 1. Feature tradeoffs for FC, Flash, and SATA II drives

	Fibre Channel Drives	Enterprise FLASH Drives	SATA II Drives
Performance	• High bandwidth with contending workloads • Sole use response time ~5 ms. • Multi-access response time 10 -50 ms.	• For high IOPS/GB and low latency • Sole use response time <1 - 5 ms. • Multi-access response time < 10 ms.	• For low IOPS/GB • Sole use response time 7 - 10 ms. • Multi-access response time up to 100 ms. • Leverage storage array cache for sequential and large block access
Abilities	• Sequential reads can leverage read-ahead • Sequential writes can leverage system optimizations favoring disks • Read/Write mixes give predictable performance	• Provides extremely fast access for reads • Can execute multiple sequential streams better than Fibre Channel	• Large I/O is serviced fairly efficiently • Sequential reads can leverage read-ahead • Sequential writes can leverage system optimizations favoring disks
Limitations	• Slower than SP Cache • Uncached writes are slower than reads	• Writes slower than reads • Heavy concurrent writes affect read rates • Single thread Sequential/large I/O only equivalent to Fibre Channel	• Long response times when under heavy write loads • Not as good as FC at handling multiple streams

FAST operation

FAST operates by continuously relocating the most active data up to the highest available tier (typically Flash or FC). To ensure sufficient space in the higher tiers FAST relocates less active data to lower tiers (FC or SATA). FAST works at a granularity of 1 GB. Each 1 GB block of data is referred to as a "slice." When FAST relocates data, it will move the entire slice to a different storage tier.

Automated scheduler

As its name implies, FAST is a completely automated tool. To this end, relocations can be scheduled to occur automatically. It is recommended that relocations be scheduled during off-hours to minimize any potential performance impact the relocations may cause. Figure 1 shows the Manage Auto-Tiering window.

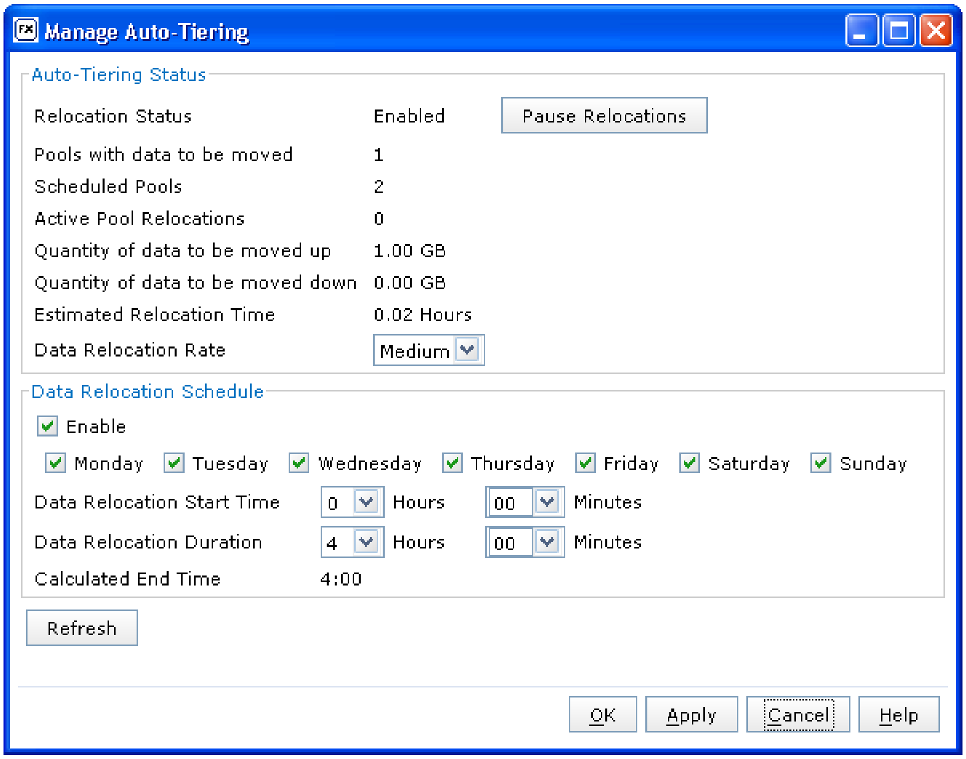

Figure 1. Manage Auto-Tiering window

From this status window, users can control the data relocation rate. The default rate is set to Medium so as not to significantly impact host I/O. This rate will relocate approximately 500 GB per hour[1].

Manual relocation

Manual relocation is initiated by the user through either the Unisphere GUI or the CLI. It can be initiated at any time. When a manual relocation is initiated, FAST performs analysis on all statistics gathered, independent of its regularly scheduled hourly analysis, prior to beginning the relocation. This ensures that up-to-date statistics and settings are properly accounted for prior to relocation.

Although the automatic scheduler is an array-wide setting, manual relocation is enacted at the pool level only. Common situations when users may want to initiate a manual relocation on a specific pool include the following:

- When reconfiguring the pool (for example, adding a new tier of drives)
- When LUN properties have been changed and the new priority structure needs to be realized immediately
- As part of a script

Pools

FLARE® release 30 leverages a new object, heterogeneous storage pools, which allows FAST to fully utilize each of the storage tiers discussed above. LUNs can then be created at the pool level. These pool LUNs are no longer bound to a single storage tier; instead, they can be spread across different storage tiers within the same pool.

[1] This rate depends on array utilization and other tasks competing for array resources. High utilization rates may reduce this relocation rate.

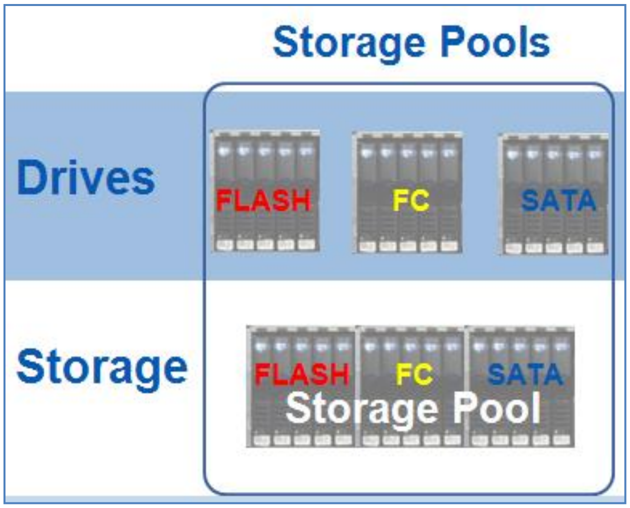

Figure 2. Heterogeneous storage pool concept

LUNs must reside in a pool to be eligible for FAST relocation. RAID groups are by definition homogeneous and therefore are not eligible for tiering. LUNs in RAID groups can be migrated to pools using LUN Migration. For a more in-depth discussion of pools, please see the white paper *EMC CLARiiON Virtual Provisioning - Applied Technology*.

FAST pool management

FAST properties can be viewed and managed at the pool level. Figure 3 shows the tiering information for a specific FAST pool.

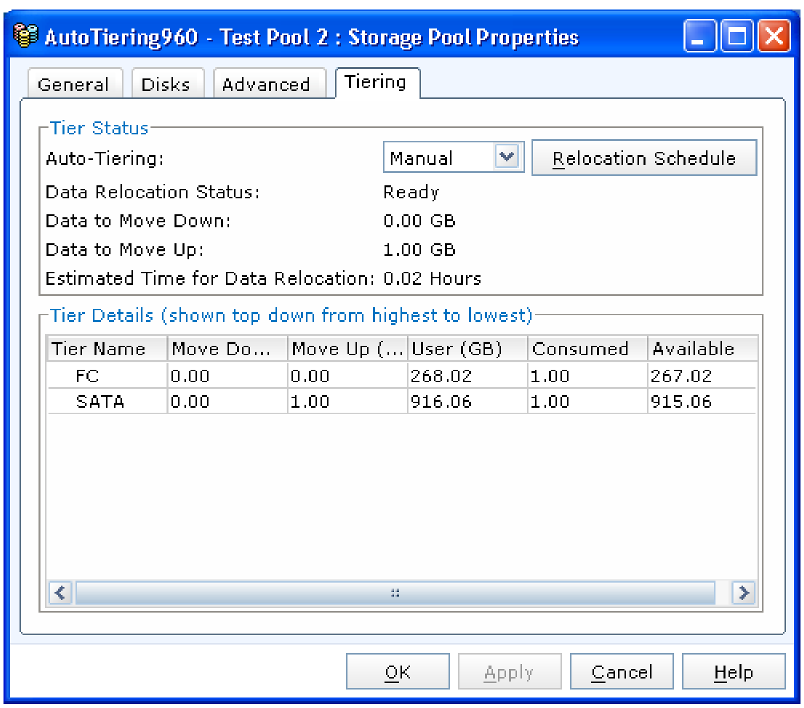

Figure 3. Storage Pool Properties window

The Tier Status section of the window shows FAST relocation information specific to the pool selected. Scheduled relocation can be selected at the pool level from the drop-down menu labeled Auto-Tiering. This can be set to either Automatic or Manual. Users can also connect to the array-wide relocation schedule using the button located in the top right corner. Scheduling policies are covered in the "Relocation" section on page 10. Data Relocation Status displays what state the pool is in with regards to FAST. The Ready state indicates that relocation can begin on this pool at any time. The amount of data bound for a slower tier is shown next to Data to Move Down and that amount of data bound for a faster tier is listed next to Data to Move Up. Below that is the estimated time required to migrate all data within the pool to the appropriate tier.

In the Tier Details section, users can see the exact distribution of their data. This panel shows all tiers of storage residing in the pool. Each tier then displays the amount of data to be moved up and down, the total capacity allocated, the consumed capacity, and the capacity available for new data.

FAST LUN management

Some FAST properties are managed at the LUN level. Figure 4 shows the tiering information for a single LUN.

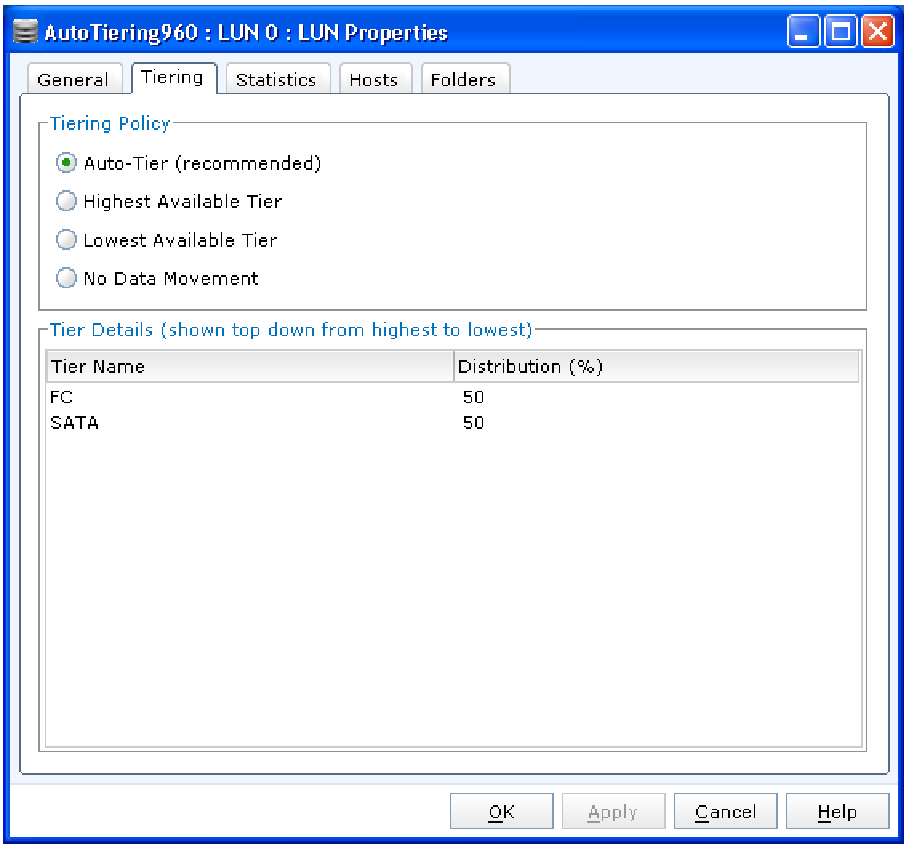

Figure 4. LUN Properties window

The Tier Details section displays the current distribution of 1 GB slices within the LUN. The Tiering Policy section displays the available options for tiering policy.

Tiering policies

There are four tiering policies available within FAST:

- Auto-tier (recommended)
- Highest available tier
- Lowest available tier
- No data movement

Auto-tier
Auto-tier is the default setting for all pool LUNs upon their creation. FAST will relocate slices of these LUNs based solely on their activity level after all slices with the *highest/lowest available tier* have been relocated.

Highest available tier
Highest available tier should be selected for those LUNs which, although not always the most active, require high levels of performance whenever they are accessed. FAST will prioritize slices of a LUN with *highest available tier* selected above all other settings.

Lowest available tier
Lowest available tier should be selected for LUNs that are not performance- or response-time-sensitive. FAST will maintain slices of these LUNs on the lowest storage tier available regardless of activity level.

No data movement
No data movement may only be selected after a LUN has been created. FAST will not move slices from their current positions once the *no data movement* selection has been made.

Initial placement

The tiering policy chosen also affects the initial placement of a LUN's slices within the available tiers. Initial placement with the pool set to *auto-tier* will result in the data being distributed across all storage tiers available within the pool. LUNs set to *highest available tier* will have their component slices placed on the highest tier that has capacity available. LUNs set to *lowest available tier* will have their component slices placed on the lowest tier that has capacity available.

FAST algorithm

To successfully identify and move the correct slices, FAST relies on three main components for its algorithm: statistics collection, analysis, and relocation.

Statistics collection

One slice of data is deemed "hotter" (more activity) or "colder" (less activity) than another based on the relative activity level of those slices. Activity level is determined simply by counting the number of I/Os, reads, and writes bound for each slice. FAST maintains a cumulative I/O count and weights each I/O by how recently it arrived. This weight deteriorates over time. New I/O is given full weight. After approximately 24 hours, the same I/O will carry only about half-weight. After a week, the same I/O carries nearly no weight. Statistics collection happens continuously in the background on all pool LUNs.

Analysis

Once per hour, the collected data is analyzed. This analysis produces a rank ordering of each slice within the pool. The ranking progresses from the "hottest" slices to the "coldest." This ranking is relative to the pool. A "hot" slice in one pool may be "cold" by another pool's ranking. There is no system-level threshold for activity level.

The user can manually impact the ranking of a LUN and its component slices. Each LUN has a setting for tiering policy that is controlled through Unisphere[2]. Tiering policy takes precedence over activity level.

Relocation

During user-defined relocation windows, 1 GB slices are promoted according to the rank ordering performed in the analysis stage. During relocation, FAST will prioritize relocating slices to higher tiers.

[2] See the "Tiering policies" section on page 9 for more details on this.

Slices are only relocated to lower tiers if the space they occupy is required for a higher priority slice. In this way, FAST attempts to fully utilize the highest performing spindles first. Lower tier spindles are utilized as capacity demand grows. Relocation can be initiated either manually or by a user-configurable, automated scheduler.

Guidance and recommendations

The following displays the total number of LUNs that can be set to leverage FAST based on the array model. These limits are the same as the total number of pool LUNs per system. Therefore, all pool LUNs in any given system can leverage FAST.

Table 2. FAST LUN limits

Array model	**CX4-120**	**CX4-240**	**CX4-480**	**CX4-960**
Maximum number of FAST LUNs	512	1024	2048	2048

FAST differentiates tiers by drive type. It does not take rotational speed into consideration. Although multiple rotational speeds can be implemented within a tier, it is not recommended. Users are strongly encouraged to implement only one rotational speed per tier within the pool. If multiple rotational speed drives exist in the array, multiple pools should be implemented as well.

EMC Professional Services can be engaged to assist with properly sizing tiers and pools to maximize investment.

FAST and FAST Cache

FAST Cache is a new feature being introduced in FLARE release 30 that enables the system cache to be expanded by using Flash drives as a second level of cache. This allows the storage system to provide “Flash drive” class performance to the most heavily accessed chunks of data. FAST Cache absorbs I/O bursts from applications, thereby reducing the load on back-end hard disks. This improves the performance of the storage solution. For more details on this feature, refer to the *EMC CLARiiON and Celerra Unified FAST Cache* white paper available on Powerlink.

The following table compares the FAST and FAST Cache features.

Table 3. Comparison between the FAST and FAST Cache features

FAST Cache	**FAST**
Enables Flash drives to be used to extend the existing caching capacity of the storage system.	Leverages pools to provide sub-LUN tiering, enabling the utilization of multiple tiers of storage simultaneously
Has finer granularity – 64 KB	Less granular compared to FAST Cache – 1 GB
Copies data from HDDs to Flash drives when they get accessed frequently	*Moves* data between different storage tiers based on a weighted average of access statistics collected over a period of time
Is designed primarily to improve performance	While it can improve performance, it is primarily designed to improve ease of use and reduce TCO

FAST Cache and the FAST sub-LUN tiering features can be used together to yield high performance and TCO from the storage system. As an example, Flash drives can be used to create FAST Cache and the FAST sub-LUN tiering feature can be used on a pool consisting of Fibre Channel and SATA disk drives. From a performance point of view, FAST Cache will provide immediate performance benefit to any bursty data while FAST will move warmer data to Fibre Channel drives and colder data to SATA drives. From a

TCO perspective, FAST Cache with a small number of Flash drives serves the data that is accessed most frequently, while FAST sub-LUN tiering with Fibre Channel and SATA drives can optimize disk utilization and efficiency.

As a general rule, FAST Cache should be used in cases where storage system performance needs to be improved immediately for burst-prone data. On the other hand, FAST optimizes the storage system TCO as it moves data to the *appropriate* storage tier based on sustained data access and demands over time. FAST Cache focuses on improving performance while FAST focuses on improving TCO. Both features are complementary to each other and help in improving performance *and* TCO.

The FAST Cache feature is storage-tier-aware and works with the FAST sub-LUN tiering feature to make sure that the storage system resources are not wasted on any unnecessary tasks. The following scenarios explain some use cases:

- If FAST technology moves a chunk of data to the Flash drive storage tier in the pool based on its access profile, then FAST Cache will not promote that chunk of data into FAST Cache, even when FAST Cache criteria is met for promotion. This ensures that the storage system resources are not wasted in copying data from one Flash drive to another.
- If a bursty workload starts accessing a particular chunk of a FAST Cache enabled LUN, FAST will not move that data chunk immediately, but FAST Cache will copy those chunks of data into the cache. Once the data gets promoted, a majority of I/O operations would be served from the FAST Cache. This can result in a lower activity level on the back-end LUNs and FAST algorithms might not reach the threshold to start moving those data blocks to a higher storage tier. A FAST-initiated data move is therefore avoided in cases where there is a temporary burst in application workload.
- In contrast to the previous scenario, if the application workload has increased on a sustained basis, FAST Cache will need to write back data into the LUNs, which may reside on FC or SATA, to make space for new promotions. This will register as back-end activity and FAST sub-LUN tiering will eventually schedule a move of the data chunks to a higher storage tier – which may be Flash drives. When this move is completed, FAST Cache will not copy any data that is already in the Flash drive storage tier.

Conclusion

Through the use of FAST, users can remove complexity and management overhead from their environments. FAST utilizes Flash, Fibre Channel, and SATA drives (or any combination thereof) within a single pool. LUNs within the pool can then leverage the advantages of each drive type at the 1 GB 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 occur without user interaction on a predetermined schedule, making FAST a truly automated offering.   In the event that relocation is required on-demand, FAST relocation can be invoked through Unisphere on an individual pool.

Both FAST and 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 time tables. Implementing both FAST and Fast Cache can significantly improve performance and reduce cost in the environment.

References

The following white papers are available on Powerlink:

- *EMC CLARiiON Storage System Fundamentals for Performance and Availability*
- *EMC CLARiiON Best Practices for Performance and Availability*
- *EMC CLARiiON and Celerra Unified FAST Cache*
- *Introducing EMC Unisphere: A Common Midrange Element Manager*
- *EMC CLARiiON Virtual Provisioning*
- *An Introduction to EMC CLARiiON and Celerra Unified Platform Storage Device Technology*