WHITE PAPER

Saving Time and Money with Savvy Use of Flash in Automated Storage Tiering

Sponsored by: EMC

Ashish Nadkarni

September 2013

EXECUTIVE SUMMARY

In a sluggish economy, getting the best ROI on every IT dollar spent is the top priority for almost every business. Storage budgets in most IT environments continue to remain flat or constrained, while data storage requirements continue to grow at an unsustainable pace. For example, even with a conservative estimate of 60% year-over-year growth, a 100 terabyte (TB) datacenter in 2013 will be an 11 petabyte (PB) datacenter by 2023. The customer movement to implement Big Data and decision support applications translates into retaining increased amounts of historical data and further growing the total storage environment. Performance requirements for the storage infrastructure are increasing too. Thanks to the move toward virtualized datacenters, ever-growing number of applications, the digitization of data, and other trends in IT, the demand placed on an infrastructure to deliver an increased number of transactions will grow exponentially. Businesses are looking to storage vendors for relief. The status quo of buying more and more poorly utilized disks to store more data is no longer acceptable.

Designing a storage environment to scale for both capacity and performance used to be somewhat mutually exclusive. Armed with new discoveries on data life cycle, including a phenomenon known as "data decay," storage vendors such as EMC are tackling this problem head-on. At their disposal are newer generations of optimization and automated tiering technologies that leverage flash/SSDs as:

- A caching layer that serves the hottest but often smaller data sets from secondary cache
- A performance tier in a pooled and virtualized multi-tiered environment that caters only to performance-intensive data sets

By using SSDs in a creative manner, one that is markedly different from how they were traditionally used, storage vendors aim to make their storage solutions more physically and operationally efficient. Such technologies also enable businesses to reduce acquisition and maintenance costs by adopting a pay-as-you-need model for storage purchases. SSDs have now become a mainstay in the enterprise.

Global Headquarters: 5 Speen Street Framingham, MA 01701 USA P.508.872.8200 F.508.935.4015 www.idc.com

Vendors such as EMC are leading this charge by making it easier for their clients to benefit from this technology. By way of EMC's FLASH 1st strategy, EMC solutions such as the next-generation VNX systems with MCx — with multi-core-enabled software and hardware optimized to deliver more performance than ever — are enabling customers to shed antiquated ways of configuring storage with data statically placed on preconfigured tiers. Now, this model can be replaced with the FAST Suite — a set of software that includes an automated tiering solution that dynamically moves data between tiers based on its performance and decay life and an extendable cache to service workloads that exhibit IOPS bursts of observed activity: Highly active data is automatically placed on the highest-performing tier, whereas low activity data is moved to the most cost-effective tier. By leveraging SSDs in this model, businesses can reduce their overall footprint without compromising performance. This can result in performance levels of up to 1 million IOPS from a single VNX.

IN THIS WHITE PAPER

In this white paper, IDC examines EMC's FLASH 1st strategy — which pairs flash drives with EMC's FAST VP (fully automated tiering for virtual pools) solution and FAST Cache, effectively an extension of the DRAM Cache, in its MCx-enabled next-generation VNX. In a VNX configured with FAST VP, storage is managed in pools that consist of multiple drive types, including flash, high-performance SAS, and/or high-capacity nearline SAS (NL-SAS). An additional flash layer in the form of FAST Cache is made available to accelerate high I/O bursts. During the normal course of operations, FAST VP user-defined policies move data intelligently and automatically from tier to tier depending on its on-demand performance characteristics. The more superior the policy, the more time that data set spends on the faster tier. For example, the FAST default policy of "FLASH 1st + Auto-Tier" seeks to place newly created, highly active data on the low-latency SSD tier.

The granularity of data movement ensures speed and efficiency while performance requirements of the entire data set dictate the overall composition of the various tiers. In most cases, only 5% of overall capacity is SSD, with 10% being high-performance SAS HDDs and the remaining 85% pool capacity being high-capacity nearline HDDs. With FAST VP moving highly active data to the high-performance SSD tier, it is not uncommon for this tier to be responsible for up to 95% of all transactions. This results in a smaller footprint and a reduction in power and cooling costs compared with a traditionally configured storage system.

The nature of this technology makes it suitable for most types of workloads. This is so because most workloads exhibit data activity skew where a small subset of data is responsible for most of the data set's overall activity. Being able to have the storage system take advantage of such data skew by dynamically placing data according to its activity provides quantifiable benefits to businesses in various industries, including healthcare and media, as well as cloud service providers. IDC research found that, on average, such organizations were able to reduce the average physical footprint by up to 66% while delivering more transactions (IOPS) than similarly configured environment with traditional HDD-based storage media. The new dynamic environment using VNX with auto-tiering required 3X less floor space and 4X fewer drives to achieve the same usable capacity.

IDC also noted that such a conversion helped organizations increase the productivity of their storage resources. While such benefits are not directly quantifiable, each business IDC interviewed for this white paper was pleased with the fact that this conversion shifted the focus of its resources from day-to-day storage optimization tasks to strategic initiatives, which were much better aligned with business objectives.

INTRODUCTION AND MARKET BACKGROUND

Demands for higher performance and the economics of flash are fueling the growth of flash drives as IDC forecasts the I/O-intensive segment (SSD and flash) will grow the fastest by capacity at a 74.3% CAGR through 2017. The market demand for performance with flash, as forecast by IDC in the I/O-intensive segment, will begin to strain storage architectures designed for performance- and capacity-optimized HDDs.

However, these numbers do not tell the pain points experienced by IT managers who support this growth. Managers have to deal with growing data, more performance-hungry applications, and acquiring and maintaining more storage with flat or declining storage budgets. For the longest time, storage vendors have focused on enabling businesses to rightsize the storage infrastructure by providing tools that could report on how their applications accessed storage and how data was served.

The earliest storage tiering solutions focused on placing data statically on the appropriate tiers. IT organizations used to configure storage environments to always meet the aggregate peak requirements of all their workloads combined. They also had to configure such storage systems with capacity to meet the data growth requirements of their businesses. This proposition almost always resulted in overpurchasing of assets that could perform and scale, making it very expensive and inefficient as most of these assets were underutilized most of the time.

These solutions were followed by versions of analytical tools that provided a great deal of information on data types and workloads. This gave birth to a new generation of storage tiering solutions that were focused on data life cycle and could purportedly move data dynamically from one location to another. The promise here was that such solutions would allow the storage environments to comprise tiers made from the fastest drives to the slowest drives, thereby resulting in lower cost. Unfortunately, very few businesses were able to realize that goal primarily for the following reasons:

- Such tools moved data at the volume level, and even if there were only a few "hotspots," the entire volume would end up being moved, resulting in businesses having to reserve a lot of excess capacity that would be used only as a transient tier.

- Such solutions were available only in enterprise systems, making them off limits for businesses that could not afford them. Even when these tools were available in midrange systems, they were somewhat performance intensive and resulted in controller overhead.

Somewhere along the way, the birth of SSDs influenced storage tiering solutions by offering a great way to provide extreme performance at the top tier in the statically tiered storage environments. They also found a good deal of adoption in the earlier generations of automated tiering solutions.

SSDs were introduced in the market originally as a replacement for spinning media. The promise was that SSDs did not have the wear and tear associated with traditional rotational media, and they simultaneously provided a lot of IOPS in the same footprint. Because of this dual benefit, vendors were quick to come up with solutions that allowed businesses to substitute flash in place of traditional drive types during the purchase.

In 2008, EMC became the first storage vendor to introduce flash as a drive type in its storage offerings. Given the price point of flash at that time, the drives were available only in EMC's top-of-the-line DMX platform, making their adoption slow. At the time, EMC positioned this offering specifically for businesses that wanted and could afford extreme performance out of a shared networked storage system for mission-critical applications (or components thereof) that could leverage the SSD tier.

However, the biggest challenge to the adoption of this offering, other than a very high price point, was the inability of automated tiering solutions that could move data quickly in and out of this tier. Businesses that deployed SSDs in their storage systems therefore had to either use it as a static tier — which meant that the data lived there in perpetuity unless it was manually moved out — or figure out a scripted approach to quickly moving parts of data in and out of this tier. Neither approach was economical.

The other lesser but nonetheless significant problem was capacity. Given the size of flash drives at that time, a traditional RAID group made up of flash did not yield much capacity. For example, a traditional RAID 5 7+1 RAID group consisting of 128GB SSDs yielded only around 750GB of usable capacity. This was at a time when a similar configuration with a typical 15,000rpm 300GB Fibre Channel drive could yield around 1.8TB for a price that was four to five times lower. With huge amounts of cache in the enterprise system, and with arguably reasonable performance from a Fibre Channel drive-based RAID group, the flash tier found very few use cases.

This initially resulted in a bit of a slump when it came to flash adoption in the enterprise.

Static Data, Data Decay, and the New Approach to Storage Tiering

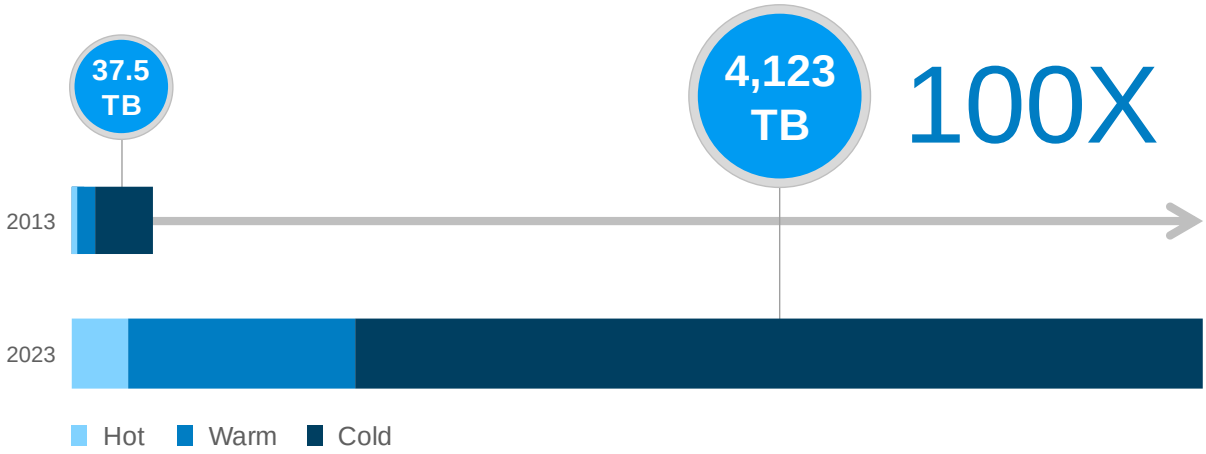
Fast-forward five years and this situation is changing. By using newer analytics tools, storage vendors are discovering that the data stored on their systems undergoes a silent transformation as it grows. On one hand, while this data may be growing at a phenomenal pace, analytical tools demonstrate that the ratio of cold to warm to hot data (i.e., data that is accessed only occasionally, frequently, and constantly) changes over time.

For example, EMC estimates that over a 10-year period, static data in an average IT environment will see a 100-fold increase (see Figure 1). While the relative size of each tier remains constant, the nominal size of the coldest tier is the largest. Data itself starts on the highest tier and then moves down and accumulates on the

coldest tier. Traditional static tiering approaches would not have been able to adapt to this phenomenon because they either lacked or had insufficient capabilities to quickly react to the changes in workloads.

FIGURE 1

Anticipated Evolution of 50TB in 10 Years



Source: EMC, 2013

The other phenomenon being brought to light by such tools is data decay. These tools have been able to model the "decay" that data in any environment undergoes as a function of time. In most environments, data, when it is first created, is highly active (hot). Workloads accessing it are performance hungry, and even the slightest latency is felt everywhere. Over time, however, the "temperature" of this data drops as activity declines, and it keeps falling until such time that it eventually "freezes" (i.e., is never accessed because these workloads now attend to newer data). Data decay allows faster tiers such as the flash to be leveraged to service the performance requirements of a much larger data set.

These factors have forced vendors to leverage newer technologies such as storage optimization and newer versions of automated tiering to make their storage systems more cost effective. Storage frames leveraging such technologies are now able to store much more data much more efficiently and offer far greater performance in terms of IOPS and throughput, and they are able to do so with a smaller physical footprint, at lower capex and opex costs, and using fewer resources. Flash drives (SSDs) are a huge contributor to the success of this storage efficiency formula that puts businesses in control of what previously were runaway storage costs fueled by unsustainable data growth:

- Storage optimization technologies target the reclamation of underutilized, redundant, or wasted storage. Thin provisioning, for example, reduces allocated wastage, whereas deduplication can eliminate redundant data blocks. Finally, compression targets the reduction of data at rest. By deploying storage optimization technologies for highly optimized data sets, businesses can save up to 75% of allocated storage.

- Automated tiering solutions function on the basic premise that of the entire data set in an environment, only a part of it is dynamic and only a fraction of dynamic data is performance critical. Oversubscribing the entire storage environment for a fraction of the performance-critical data causes wastage and is unsustainable in the long run. Using intelligent data movement technologies enables performance-sensitive data blocks to be placed on the appropriate tier in a timely manner while the static counterparts continue to rest on slower media.

The Impact of Storage Tiering on the SSD Market

There is no doubt that one of the principal reasons behind the growth of the SSD market is the large-scale adoption of SSDs in the enterprise. The SSDs of today are much more reliable, have more capacity, and offer a fair bit of value for the price. IDC estimates that in CY 1Q13, worldwide SSD shipments increased 90% year over year to 13 million units on revenue that increased 44% to $2 billion. The revenue growth for the worldwide SSD market was a result mainly of strong SSD shipment growth in the enterprise segment and client segment from PC applications. In the enterprise storage system and server market, higher SSD shipments helped boost revenue in the enterprise segment.

However, SSDs are still expensive, and they are not yet at a point where they can be considered as a wholesale replacement for spinning media in storage environments. But in spite of that, the adoption of flash drives in storage environments of all sizes continues to increase.

Arguably then, the next big reason for growth of SSDs is coming from their increasing presence in external disk storage systems. IDC attributes this primarily to the growing adoption of next-generation storage optimization and tiering solutions such as EMC FAST VP and FAST Cache. The focus of using SSDs in these systems is shifting away from low capacity, higher price, and higher performance to being a transient tier or cache that optimizes performance and capacity at an acceptable price point. More importantly, such solutions are now available in midrange systems, making it easier for businesses of all sizes to avail of their benefits. No longer are SSDs and automated tiering solutions an "enterprise only" feature that is the privilege of a chosen few.

THE EMC FLASH 1ST STRATEGY: MAKING A LITTLE FLASH GO A LONG WAY

EMC has been one of the early adopters of storage optimization and tiering technologies that leverage SSDs. Its FLASH 1st strategy brings together flash drives with the FAST Suite to provide out-of-the-box storage efficiency and value without compromising performance. Today, the FAST Suite is made up of FAST Cache and FAST VP and is targeted specifically at the VNX platform. In the future, it may include other technologies that augment the performance of the storage system itself.

Data Decay: The Core of the FLASH 1st Strategy

In creating the FAST Suite, EMC has paid closer attention to the data decay model and the resulting success of the storage optimization solutions that target data decay first (as opposed to those that provide other types of optimization without examining the data shelf life).

The incredibly short shelf life of data is becoming apparent thanks to analytical tools that can measure data in most environments. In a lot of ways, most data generated has the "shelf life of a banana." In other words, in most cases, freshly created data has read and write sweet spots that make it fairly "hot." Subsequently, as new data is created, it is put to the side, thereby reducing its temperature or value — at which time the numbers of write operations decrease. Eventually the data is accessed very infrequently or not at all and hardly ever modified. EMC, in its analysis, has concluded that in most cases, when data is older than 90 days, the probability of anyone ever requesting that data falls to 10%. After 180 days, that probability falls to 1%, and when this data is a year old, there is only 0.01% probability that anyone would request it.

By moving data across tiers according to its actual decay in activity, FAST optimizes the ROI of the IT storage budget. The FAST Suite ensures that highly active data is optimized for the lowest dollar per IOPS and for the lowest dollar per GB when that very data becomes cold. In other words, FAST delivers optimum economic efficiency over the data's life cycle.

A key benefit of constant and dynamic data placement is that the most expensive SSD tier is replenished constantly with only the highest activity data. As data activity falls for a given data set, it is evicted from the highest $/GB tier and moved down the cost hierarchy. As data moves down, the capacity on the SSD tier is thus freed up for newer and "hotter" data. EMC estimates that most configurations will encounter a 100% churn of data in the SSD tier roughly six times a year. Therefore, the dynamic data movement of FAST in most cases will yield a 600% ROI for the SSD tier. This makes deployment of SSDs in the enterprise a much more attractive option. By constantly reusing the expensive flash tier in this fashion, FAST makes a little flash go a long way by essentially delivering the service equivalent of an SSD tier six times larger.

EMC's FAST Suite is based on the following key principles:

- Simply deploying high-performance flash drives and high-capacity disk drives in a storage system does not guarantee efficiency. To gain efficiency, companies need a tiering solution that automatically monitors data as it ages and moves it to lower-cost, high-capacity tiers once they have determined that data activity has declined.
- This system needs to be adaptable to multiple and diverse workloads by altering the ratio of the various tiers and the frequency with which the movement occurs. Different data sets may exhibit varying degrees of "popularity decay," but the older and less active the data becomes, the less performance the storage system will be required to dedicate to it. The system should "learn" such patterns and/or allow storage administrators to quickly tune it. Once this scheme is automated, the system should optimize data storage for both best performance and lowest capacity cost.

- Directly proportional to this tiering blend should be the economy: As data declines in popularity, it is moved to lower-performing and less expensive storage tiers. The more granular the move, the lower the overhead and the higher the optimization and savings. For example: When 10GB of data in a 10TB data set is "hot," only that 10GB "chunk" need occupy the flash tier (as opposed to the entire 10TB) and obtain necessary performance with fewer flash drives (the IOPS of a flash drive is roughly 30 times that of a high-performance HDD and 65 times that of a nearline SAS). When this 10GB chunk gets older, it is moved from flash to the nearline SAS tier, thereby reducing its occupied cost to less than one one-hundredth of the original cost.

The system must be able to cater to frequently requested data or data that is always "hot" by keeping it on the fastest tier as long as it is being requested frequently.

FAST VP with Flash as a Tier

EMC first introduced FAST or Fully Automated Storage Tiering on its DMX platform in 2009 and later on the VNX platform. FAST laid the foundation for changing the way storage is viewed and provisioned, moving away from traditional RAID groups and LUNs (logical unit numbers or volumes) to the concept of wide striped pooled storage where data is looked at in "slices." When FAST was introduced, it was already well established in the market that the firmware in most systems could interpret, guess, and adapt to data access and I/O patterns. However, other than tuning its cache and doing prefetches and some minor tweaks, the firmware could not do much because data was really stored in LUNs that existed on traditional RAID groups. Enter FAST. With pooled storage comprising various tiers of capacity and performance disks (including SSDs), the system could actually respond to changes to incoming I/O and access by placing the data on the appropriate tier that was best suited to service it. Today, recent enhancements to FAST VP incorporate sub-LUN tiering in 256MB "slices." This granular movement makes FAST VP more efficient and agile: FAST VP optimizes storage pools automatically, ensuring that active data is being served from the highest-performing tier, usually flash, while cold data is moved to lower-cost, high-capacity disk tiers. FAST allows flash to be used for high-performance workloads.

FAST Cache

In addition to servicing routinely active, "hot" data from the flash tier in FAST VP pools, some of the SSDs installed in the system can be configured as FAST Cache, specifically to provide a real-time performance boost to data blocks accessed by the "hottest" workloads. FAST Cache is essentially an extension of the system cache, is read/write capable, is nondisruptive, and can be configured to hold a usable capacity of up to 4.2TB on the MCx-enabled VNX.

FAST Cache is like a revolving FIFO buffer. The "garbage collection" mechanisms in the system automatically identify and evict data that is not worthy of being serviced from cache.

BENEFITS OF AUTOMATED STORAGE TIERING AND OPTIMIZATION SOLUTIONS

The primary benefit of the storage optimization and tiering solutions in the EMC FAST Suite is that they allow businesses to rightsize the storage environment. The derived benefit of this is a storage framework that costs less to build out and costs less to operate but offers performance that is equal to or better than that of a similarly configured framework with traditional storage elements only.

IDC recently conducted a study to quantify the business benefits of storage optimization and automated tiering solutions using EMC's FLASH 1st strategy. During this study, IDC surveyed businesses that had deployed VNX systems with FLASH 1st. Specific figures for savings and costs came from interviewing IT managers at companies of different sizes and in different industries such as media, healthcare, and telecom. For the survey, IDC asked the IT managers a series of questions regarding their storage strategy and how deploying VNX systems with FLASH 1st impacted their organizations. These organizations have very complex and extensive storage needs but limited budgets, requiring them in all cases to think outside the box.

Capex Savings

The traditional approach during the storage requirements gathering process is to use standard sizing metrics such as IOPS per drive configuration or GB per drive configuration. Unfortunately, these two metrics are mutually exclusive. In almost every case, the resulting configuration either is oversubscribed from a capacity perspective or remains underutilized from a performance perspective. The net result is businesses often overspend on storage.

As the FLASH 1st strategy changes that line of thinking, IT managers feel that the burden on their IT "checkbooks" has been relieved. When a VNX system is configured with FAST, the usual metrics are tossed out and replaced by metrics that bring into focus the "hotness" of the data, the overall size of the data, and the data life cycle. The resulting configuration, where the highest quartile of IOPS is serviced by the smallest quartile of (high-performance) storage tier, thus has a smaller footprint than in the former situation, and the storage system performance is equal to or better than the performance in the previous situation. In fact, one could argue that in the latter situation, the efficiency of the storage system is much higher because all components are utilized optimally. This has huge implications on acquisition costs.

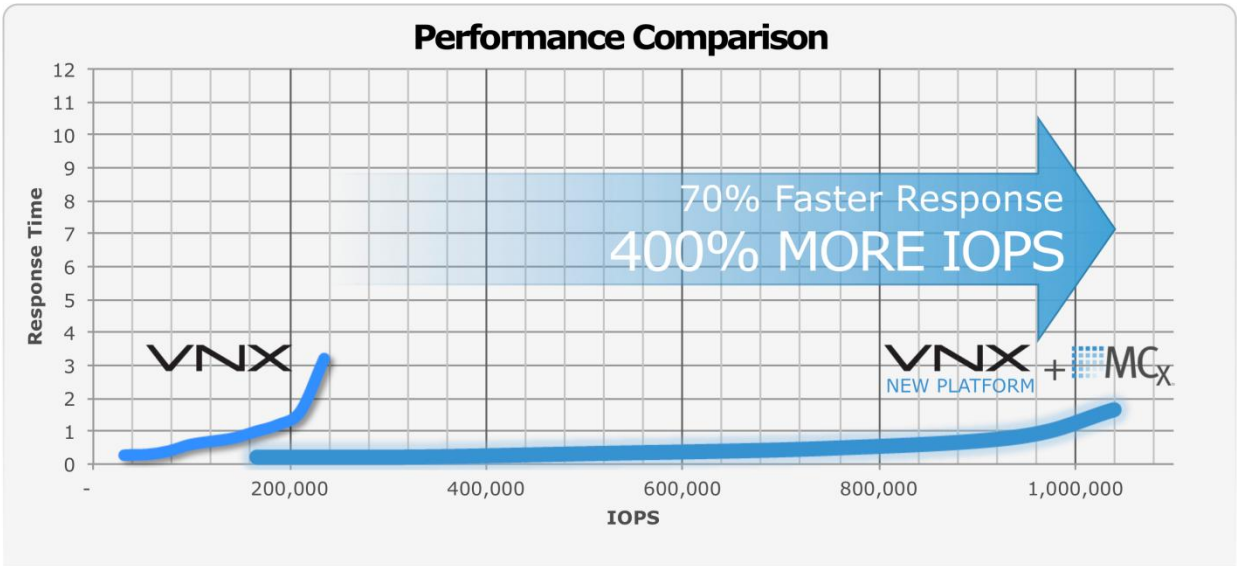
For example, a leading healthcare provider in the United States put out an RFP to storage vendors to quote a storage configuration that could support a virtual server infrastructure running an electronic health records (EHR) system from MEDITECH. Two of the three vendors that responded proposed storage that was configured in a traditional manner. To support IOPS of over 20,000, one vendor stated that the company would need 120 drives. The third vendor, EMC, responded with a VNX and FLASH 1st configuration that could accomplish the same result with less than half the number of drives as its competitors. The resulting configuration was more agile and performed as well as its competition, but more importantly, it was cheaper to acquire and maintain.

The next-generation VNX with MCx means that EMC can deliver a midrange series with the performance of an all-flash array and the efficiency of tiering (see Figure 2).

- The modular scalability of the VNX models can address the growing customer adoption of virtualized infrastructures and cloud computing environments.
- The redesign of the VNX family makes it possible to take better advantage of the low latency of flash and deliver highly optimized data services.
- This redesign also enables more efficient capacity management with tiered SSD and HDD storage for a superior customer experience.

FIGURE 2

VNX Performance Comparison



Source: EMC, 2013

Opex Savings

It should come as no surprise that datacenter power consumption is slowly rising to the top of every CIO's list of worries. Datacenter power consumption will only increase as data grows and the infrastructure necessary to support that growth is upgraded. Therefore, businesses are always looking for ways to cut back on power and cooling. Rotating hard drives in servers and storage systems are one of the worst offenders when it comes to heat generation. At an average of 18W per 15,000rpm Fibre Channel or SAS drive, a system configured with 100 drives consumes around 1.8KW in drives alone. This does not include the power needed to cool the drives.

Compared with rotating hard drives in the manner described above, SSDs consume an average of 5W, while nearline SAS drives consume an average of 12W. By utilizing VNX with FLASH 1st, for example, customers can slash power consumption

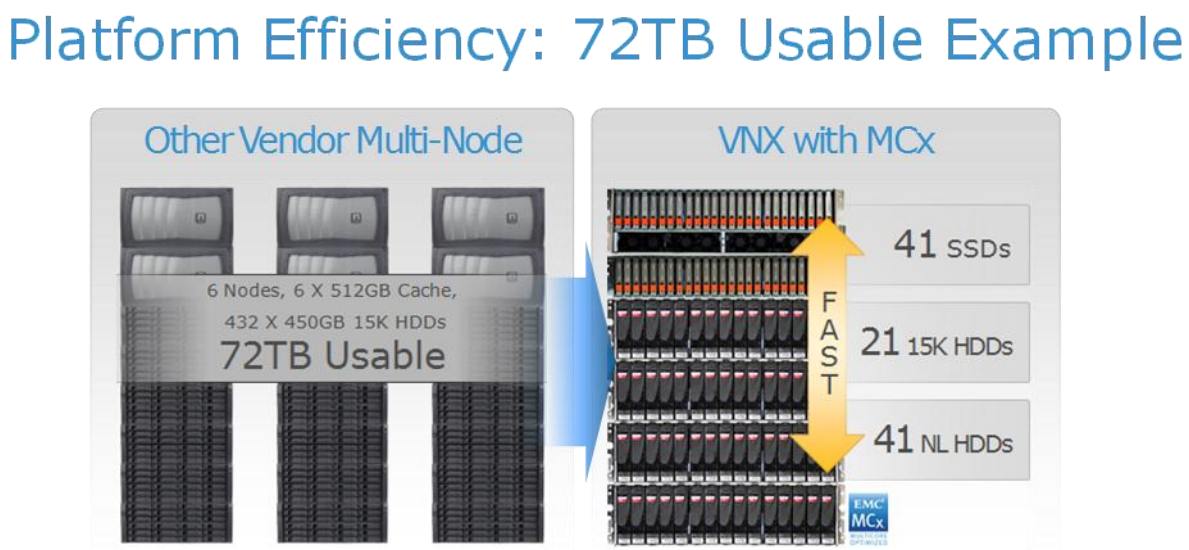
by around 50% simply by altering the mix of flash, 15,000rpm, and nearline drives. This is because flash drives not only consume less power but also offer nearly 30 times the IOPS of the 15,000rpm drives.

This was one of the criteria used by a telecommunications company in Morocco when it was selecting storage for a private cloud virtual desktop deployment. To sustain a deployment that would eventually support 7,600 users, the company required a storage configuration that could sustain 18,300 IOPS for short bursts. Traditionally, configured storage would have required 102 15,000rpm drives in a single-tier configuration to reach this IOPS number. Unfortunately, because each drive is a massive 600GB, this older strategy would require the company to acquire 60TB of unneeded capacity. This would have resulted in an underutilized configuration that occupied the equivalent of an entire rack and left little room for expansion should the number of users (hence IOPS) increase. Furthermore, the drives in this configuration would have consumed 1.8KW in energy alone. By going with a VNX in a three-tier configuration, the company achieved the same number of IOPS in 22 drives. This storage configuration now occupies the equivalent of half a rack with energy savings of over 1.2KW just by altering the number and type of drives. "We don't have a space issue in our datacenter, but even then we could not justify not going with a configuration that was cheaper, occupied only half a rack, and consumed 1.7KW less energy," said the IT manager.

MCx-enabled VNX with FLASH 1st takes advantage of multi-core CPUs to provide superior ROI advantages. As illustrated in Figure 3, an array without efficiency features of VNX with auto-tiering would require 3X the floor space and 4X the drive count to achieve the same usable capacity.

FIGURE 3

VNX with MCx Configuration Comparison



Source: EMC, 2013

Generally, the next-generation VNX series with auto-tiering can be expected to deliver significantly greater performance than other systems, when configured for similar capacity and costs.

Increase in IT Productivity

In discussions about tangible savings realized by deploying storage tiering and optimization solutions, it is often easy to overlook the positive impact these solutions have on storage resources. Earlier versions of storage tiering and mobility solutions were slow, clunky to use, and very resource intensive. Businesses quickly realized that the direct savings in acquisition costs were soon offset by an increased burden on their storage staff. Besides, any manual effort is always fraught with risks of inefficiency, human error, and, worst of all, data loss. With a storage platform that automates data movement, IT resources no longer have to micromanage the storage environment and thus are freed up to work on new or strategic initiatives more rapidly. In similar research related to intelligent storage automation, IDC found that IT productivity gains reduced the operating cost per TB by 3.5 FTE hours or 48%.

As one of the core tiers of the infrastructure, when the storage system meets or exceeds performance expectations, in most cases the higher tiers, such as the compute and application stacks, function in a similar manner, like a well-oiled machine. This often produces the benefit of improved end-user functionality. While this benefit is not measurable in all cases, it is worth noting that all clients in IDC's study acknowledged that deploying storage automation solutions improved the overall well-being of other IT resources.

For example, in the case of a United States–based media company, nightly data load times for primary databases fell from over 14 hours to less than 3 hours — a reduction of 79%. "The immediate impact of deploying the solution was that everyone from the application management side, including the systems management teams, was freed from the daily struggle of having to allocate time to troubleshooting failures or slowdowns in data loads," said the IT manager at this business.

DEPLOYMENT SCENARIOS AND DRIVERS FOR THE EMC FAST SUITE

During the interviews, businesses provided IDC with a variety of reasons for deploying EMC VNX systems with FLASH 1st:

- On-demand performance without having to prepopulate the system was one of the primary themes that resonated among businesses.
- Not having to purchase two rack spaces' worth of capacity that would be used only once a day, a few times a week, or, in some cases, only a few times a year was a huge benefit of EMC's solution versus its competition.

An added benefit of this solution is that the configuration is flexible and somewhat future proof — allowing businesses to add capacity in IOPS or in GB as needed without costly forklift upgrades. The storage systems scale linearly, and storage administrators can provide added boost to performance-intensive workloads by dynamically adding more

flash drives to FAST Cache and FAST VP pools. Additionally, for workloads that are always data intensive, administrators can "pin" this data onto the flash tier (i.e., so that it is permanently served from that tier and excluded from any automated movement that could move it to slower tiers when the peak workload subsides). Similarly, administrators can add more storage capacity by adding SAS or nearline SAS drives to the FAST VP pools. This characteristic nature of FAST makes it well suited for service provider or private cloud models where there is a tight coupling between acquired and provisioned capacity to keep costs under control.

Acquisition and ongoing maintenance costs are another key reason why businesses consider this solution for new deployments. Budget-minded businesses can now purchase a smaller platform and get into the FAST technology and then scale as needed — often minimizing the need to sink an investment in a more expensive platform that will remain underutilized for much of its lifetime.

Ongoing maintenance costs were cheaper too. Because SSDs and NL-SAS drives consume much less power than 15,000rpm SAS or Fibre Channel drives, a blended configuration consisting of 10% flash, 30–40% 15,000rpm drives, and the rest with 7,200rpm NL-SAS drives comes in at 60% of the total drive count as a traditionally configured storage system. In this situation, downsizing the VNX model resulted in the use of 75% less power and 65% less space.

The healthcare provider that purchased this solution to service its MEDITECH EHR system was thrilled to be able to kill two birds with one stone: It acquired a system that was fully certified by the EHR vendor and leveraged it for its heavily virtualized application environment, including email and productivity applications. Furthermore, this solution was much more economical to maintain than its counterparts, coming in at slightly higher than half the number of drives. Initially, it had "pinned" its EHR data to FAST Cache, but after seeing the robustness of the system, it removed this pinning and is now planning to further expand the FAST VP pool with nearline SAS drives and an additional layer of flash. "Investment in EMC has brought down our hard and soft costs. We now have happier end users — which means fewer calls to the storage team and fewer trouble tickets. The VNX solution has cut down a lot of stress in the environment," remarked the storage manager at this provider.

"Investment in EMC has brought down our hard and soft costs. We now have happier end users — which means fewer calls to the storage team and fewer trouble tickets. The VNX solution has cut down a lot of stress in the environment."

Even though a casual glance at this solution makes it appear very complex and resource intensive, it is anything but that. All interviewees agreed that while they initially had some reservations about how much additional workload it would add to their already overworked staff, they were pleasantly surprised by how easy the FAST Suite was to maintain. "On the contrary," one IT manager remarked, "the solution has liberated my staff from the constant storage micromanagement that consumed most of their time." Businesses that have to deal with unsustainable data growth with skeletal storage staff thus will find storage optimization technologies such as FAST VP indispensable.

"The solution has liberated my staff from the constant storage micromanagement that consumed most of their time."

CHALLENGES FOR EMC'S FLASH 1ST STRATEGY

As with any new innovation, EMC has to handle its share of challenges with its FLASH 1st strategy. One of the biggest challenges for EMC will be to convince businesses that their data will remain safe inside a storage frame that now has an additional abstraction layer above the "tried and tested" RAID elements. This data not only is constantly in flux (i.e., being moved from tier to tier and in some cases being copied into FAST Cache) but also is being sliced and diced at a sub-LUN level so that, at any given moment, part of it could reside on one tier while part of it could be somewhere else. Storage managers who subscribe to a more traditional approach to data placement could push back, demanding instead that this technology mature before they place mission-critical data on it. However, EMC could showcase the capabilities of this technology by focusing on the following key benefits:

- The solution provides improved performance in a much smaller physical footprint.
- The solution examines data decay in the environment and accordingly optimizes $/IOPS for performance without compromising on $/GB vis-à-vis a traditionally configured storage system with no automated tiering solutions.
- The ability to scale up the storage subsystem on demand without having to prepurchase the components makes it more cost effective.
- Organizations can benefit from tangible savings that impact both capex and opex budgets. This includes savings such as acquisition costs, ongoing maintenance costs, and facility costs such as power and cooling.
- Indirectly measurable savings, such as liberating resources from day-to-day management tasks, allow staff to focus on longer-term and more strategic initiatives.

As far as the risks go, EMC is in a strong position to convince its customers that both its VNX systems and its FAST solution are based on tried and tested components. While no technology is risk free, businesses could find it reassuring that the platform in which they are investing is stable.

Additionally, EMC could focus on FAST as an ideal solution for businesses seeking to consolidate legacy environments residing on traditionally configured but often inefficiently utilized storage systems. The primary benefit offered by a FLASH 1st strategy during such conversions includes the overall reduction in storage footprint while meeting the demand for data growth and performance with IOPS-intensive applications.

CONCLUSION AND ESSENTIAL GUIDANCE

There is no doubt that both storage optimization and flash technologies are here to stay. The initial charters for both technologies were very different: One was focused on saving costs, whereas the other focused on providing extreme performance. However, as EMC is demonstrating via its FLASH 1st strategy, a case is being made for both technologies to coexist in the same environment in order to provide the most cost-effective solution.

There is also no doubt that unprecedented data growth will continue to strain organizations everywhere. Vendors that make the most of flash-based storage optimization technologies that are based on validating the shelf life of data to address not just the thirst for more performance but also ongoing pain points such as burgeoning storage infrastructure costs, lack of infrastructure agility, and resource constraints will come out winners. As with other mainstay storage technologies, solutions such as FAST will ultimately be looked upon as enablers for:

1. Increased storage and data management efficiency
2. Increased application performance
3. Reduced storage infrastructure costs
4. Increased IT productivity

Organizations that leverage such solutions will benefit from an efficient, agile, easy-to-manage, and scalable storage infrastructure.