Hitachi Accelerated Flash

An Innovative Approach to Solid-State Storage

By Hitachi Data Systems
February 2015
Contents

Executive Summary 3
Introduction 4
The "Catch-22" of Traditional Storage 4
Hitachi Accelerated Flash 5
Hitachi Accelerated Flash Evolution 6
Key Building Blocks of HAF 7
The Flash Module Unit 7
The Flash Module Drive 8
Why Multi-Level Cell (MLC)? 8
How HAF Innovation Extends the Endurance of MLC Flash System 8
Flash Controller ASIC 9
Architectural Highlights 9
Key Benefits of Flash Module Drive Controller Design 10
Innovative Flash Acceleration Software 12
Hitachi Accelerated Flash Portfolio 12
Maximize the Value of Flash Storage With Hitachi Dynamic Tiering 13
Central Management: Hitachi Command Suite 14
Flash Module Drive Performance and Other Metrics 15
Hitachi Storage Performance Comparison 15
Pricing 16
Space and Power Efficiency: Capacity Growth Path 16
Space and Power Efficiency Comparison Chart: Comparing HAF, SSD and HDD 17
Conclusion 17
Appendix A: Contributors 18
Hitachi Accelerated Flash

Executive Summary

Until the advent of solid-state storage that was ready for the data center, the performance gap between storage systems and application server CPUs created troublesome productivity bottlenecks in the data center. Storage architects tried to squeeze performance from data stores in ways that were both inefficient and expensive. But as performance was key in the world of data-intensive applications, cost had to remain a secondary concern.

However, IT departments have finite budgets. Therefore, the mounting costs necessary to achieve robustness have been the primary roadblock to a wider adoption of solid-state storage, particularly in the data center.

With these concerns in mind, Hitachi perceived the following challenges: How can we provide an enterprise-class, high-performance, robust, flash-based storage option that has a 5-year endurance specification? And, how do we do so cost-effectively?

The development team’s answer was a purpose-built flash controller that exploits the attractively priced MLC NAND technology. At the same time, the controller delivers the performance, functionality, robustness and endurance expected in an enterprise-class solution.

The result is a unique storage option, Hitachi Accelerated Flash (HAF), which can deliver nearly 4,000,000 IOPS within a Hitachi storage platform.

HAF is comprised of flash acceleration system software and Hitachi designed flash module drives (FMD). Although FMDs use multilevel cell flash (MLC), they are not solid-state drives (SSD). The FMD is larger than a traditional SSD and delivers 3 to 4 times the IOPS performance. It is more space efficient, with 2x to 4x the terabytes per rack unit density. It is 32% more power efficient, and is one-third lower in cost per terabyte when compared to a similarly configured solution using SSDs.

The brain of the FMD controller is a custom-designed ASIC featuring a quad-core processor. This ASIC has the compute power needed to deliver innovative housekeeping management that eliminates performance bottlenecks. It extends the ECC to deliver robustness, manages wear leveling to heighten endurance, and adaptively refreshes data to secure data integrity.

However, Hitachi looked beyond the flash controller. As a result of a complete refresh of the firmware, the design team optimized Hitachi storage systems for flash, which resulted in the development of flash acceleration software. This holistic design approach ensured that flash seamlessly integrates into existing Hitachi storage systems and is transparent to other functional software.

The purpose of this white paper is to take a close look at Hitachi Accelerated Flash. The discussion focuses on the uniqueness of this solution. It considers why and how HAF can meet the increasing IT challenges of large and small enterprises as they focus on the management challenges of their high-velocity data.
Introduction

The “Catch-22” of Traditional Storage

The storage industry has become very cognizant of the number of conflicting challenges that face the storage architect. However, despite the many issues, performance continues to dominate as a top concern.

There is a multiplier effect facing the storage professional (see Figure 1). It is not simply the increase in performance but the management of the growing disparity between the annual performance gains of client-side initiators (application servers) and traditional data storage (hard disk drive or HDD).

Figure 1. Increasing Performance Gap Between Servers and Storage

With approximate annual performance growth rates of 50% in server CPUs and 10% in storage, there is a widening performance gap. This dichotomy increasingly stresses the ability of storage to serve data at the rate today’s high-performance applications require. Such applications include transaction-heavy database applications, business and financial analytics, data mining, data warehousing, production archive, virtual machine deployment and virtual desktop infrastructure (VDI).

There is little surprise that these data-intensive applications now seem performance-starved. Thus, administrators are looking for a storage solution that can deliver data at the velocity applications need to be fully effective.

To understand the issues that impact storage performance, first consider queue depth, the number of I/O requests waiting to be serviced by the storage system. The more hard disk drive actuators (or spindles) available, the greater the ability of the storage system to respond. However, simply adding more spindles to get I/O performance can be expensive. If the application does not need the capacity that all those hard disk drives bring to bear, it is an enormous waste.

Historically, storage administrators and other storage vendors have applied techniques such as wide striping (reduce latency) or short stroking (reduce seek time) to improve storage response time. Wide striping is simply adding more spindles and distributing the data over a larger number of underutilized spindles. As already noted, this approach is very inefficient if the additional capacity is not required. Short stroking is the practice of only writing data to the disk
sectors located on the outermost portion of the disk platter. Short stroking improves data access time by shortening the actuator seek time. However, it comes with a significant penalty. It uses only the outer cylinders, meaning that only about 25% of the drive is utilized. Short stroking translates to a significant amount of disk space or capacity being wasted (not to mention the power costs of constantly spinning disks with 75% of unutilized capacity). Considering that the candidates for short stroking tend to be the most expensive 15K HDDs, it is another very expensive option.

In high-performance application environments this underutilization may seem a reasonable price to pay in exchange for an approximately 50% decrease in access time (increase in data accessibility). However, many performance-starved applications are often also capacity-intensive. Translation: Short stroking is not the answer. Neither is wide striping.

To achieve a truly efficient data center, performance-starved applications must be addressed with a storage solution that meets both performance and capacity requirements (see Figure 2).

Figure 2. Nonvolatile Memory (NVM) "Solid-State Storage" Filling the Performance Gap

In this type of a solution, high-capacity, solid-state storage in an all-flash storage system or in hybrid storage architectures becomes attractive.

**Hitachi Accelerated Flash**

There is no argument that flash can deliver performance when needed. Companies in the Hitachi customer base have been quick to appreciate the benefits that highly accessible data can deliver. They also appreciate the economic benefits that come with this new tier of storage. The Hitachi portfolio, from the Hitachi Unified Storage 100 family to Hitachi Unified Storage VM (HUS VM) to Hitachi Virtual Storage Platform (VSP), has deployed SSD technology for some time. However, the focus of this paper is to discuss the differentiated Hitachi approach to solid-state storage: Hitachi Accelerated Flash (HAF).
Hitachi has designed a purpose-built flash solution for enterprises. HAF includes system controller software that is optimized for I/Os to flash devices. It also includes Hitachi designed flash module drives (FMD). These components are available options for each of the Hitachi platform families from Hitachi Unified Storage 150 on up to Virtual Storage Platform G1000 (see Figure 3).

Figure 3. Platforms That Support Hitachi Accelerated Flash

An FMD is not an SSD. It is unique flash technology, developed by Hitachi, that builds on the many flexibility, performance, reliability and common management attributes already well known and appreciated in traditional Hitachi storage. FMD controllers are purpose-built to provide consistently high performance. By removing housekeeping tasks from the I/O path, the infamous "write cliff" is eliminated. Also, an extended ECC reduces the refresh rate, buffering and compression eliminate excessive writes, and global wear-leveling management balances NAND cell use patterns. All of these benefits drive increases in the endurance of the FMD. More on these topics later.

The FMD controller can support higher storage densities than traditional SSD configurations. The FMD, the physical equivalent of an SSD, is available in both 1.6TB and 3.2TB capacities compared to the enterprise SSD capacities of no more than 800GB. This 4x increase in supportable capacity translates into significant cost savings.

Hitachi Accelerated Flash Evolution

When the pragmatic value of solid-state storage in the data center was realized, SSDs evolved to become mainstream in the marketplace. The Hitachi approach was to integrate the best available representation of this technology into their existing storage portfolio. However, while implementing commodity solid state in drive form factors (SSD) into traditional storage systems is often the quickest time to market, it also tends to be limited. SSD is restricted in its ability to take full advantage of advanced storage capabilities resident in many storage controllers.

Over time, Hitachi storage systems have evolved to offer a rich portfolio of services and functionality that has established the dominance of HDS as a storage vendor. The challenge, therefore, was for the Hitachi developers to integrate flash technology but without compromising current storage architectures.
The solution to the dilemma was nicely articulated by industry analyst George Crump\(^1\), who stated in one of his blogs: “When a flash system vendor develops their own controller technology it also gives them the freedom to advance their controller functionality so it can provide enterprise class reliability to lower cost MLC flash.”

Hitachi developers could not have agreed more. With HAF, Hitachi set out to design a purpose-built, enterprise-class flash controller that would tightly and transparently integrate with the current Hitachi storage portfolio. The result is a flash-focused controller that synergistically operates with traditional Hitachi storage controllers and exploits the rich portfolio of Hitachi storage innovations. This unique flash controller introduces a significant number of flash-focused innovations that optimize the use of MLC NAND flash as an enterprise-class storage system. The 60-plus patents either granted or applied for are a testament to the advanced innovation accessible in this controller.

Key Building Blocks of HAF
1. Flash module unit (2U chassis that can support up to 12 flash module drives) (see Figure 4).
2. Flash module drive (see Figure 5).
3. Flash controller ASIC (part of the FMD).
4. MLC NAND flash (part of the FMD).
5. Flash acceleration software (licensed feature).

The Flash Module Unit
This chassis consists of a basic 2U enclosure that can be easily integrated into the VSP G1000, VSP, HUS VM or HUS 150 storage systems. The chassis is designed specifically to accommodate 12 FMDs, which are the hot-pluggable unit of capacity. A minimum configuration of 4 FMU basic building block modules will fit into the VSP or VSP G1000 and will support high availability implementation. HUS VM and HUS 150 systems will support the integration of individual flash module units in either a hybrid or an all-flash configuration.

---

The Flash Module Drive

The FMD (see Figure 5) is the basic enterprise storage module for HAF. It integrates the flash controller custom ASIC, MLC NAND flash memory, DDR3 and SAS interfaces to create a complete SAS-interfaced physical LUN device. It can be ordered in capacities of 1.6TB and 3.2TB of flash-based persistent storage capacity.

Figure 5. Flash Module Drive

Why Multi-Level Cell (MLC)?

MLC has a durability number of $10^3$ or at best $10^4$ writes per cell, and single level cell (SLC) can boast a durability number of $10^5$. So, why did Hitachi select MLC for an enterprise-class flash storage system? One reason was cost. The consumer industry has adopted MLC NAND flash and, therefore, it is going to enjoy a much steeper cost decline than SLC flash. However, cost is only relevant in enterprise-class storage systems when reliability, availability and performance can be guaranteed.

The Hitachi development team’s purpose-built approach has enabled them to deliver the cost benefits of MLC in a flash-focused storage. In addition, this storage transparently and synergistically integrates with existing Hitachi storage systems. It delivers the performance, robustness and endurance needed to meet the increasing demands of enterprise workloads.

How HAF Innovation Extends the Endurance of MLC Flash System

The frequency of writes determines the lifespan of flash. Hence, the number of times a NAND cell is subject to a write affects the flash storage’s durability and reliability. This frequency equation is a law of flash physics, and such laws cannot be challenged. However, innovative controller designs can manage and optimize for this reality and extend the endurance of MLC NAND. Such is the case with the Hitachi enterprise-class flash module drive, or FMD.

The following are some of the techniques the development team used to expand endurance in the FMD. Additional discussion on these items will be presented in later sections.

- ECC has been extended to correct 48 bits per 1.4KB. This correction enhances the ability to monitor the degradation of pages and avoids any premature page rewrites.
- The controller reads and recalculates the internal ECC on the complete FMD every 2 days and dynamically optimizes page refresh based on applied error correction.
- Logical-physical address conversion, to enable old data to be erased asynchronously, minimizes housekeeping tasks.
- Buffered write area, to reduce formatting for small writes, efficiently manages formatted page availability.
- Read/modify/write, for writes that are less than a page, minimizes the consumption of formatted pages.
- Data is refreshed at least every 30 days to avoid retention-time degradation.
- Zero data compaction reduces space from unnecessary writes by up to 94%.
- Wear leveling is done locally across the pages in a flash module and globally across all modules in a pool of flash modules. This approach distributes wear and extends the life of a flash module.
- 25% of the flash capacity is overprovisioned.

**Overprovisioning**

Overprovisioning is the practice of including flash memory above the advertised capacity. This overprovisioning increases the write endurance of the specified MLC flash memory capacity and its overall performance. FMD provides 25% overprovisioning.

**Data Reliability**

Each NAND memory block exhibits a different robustness for endurance. To manage this uneven wear characteristic, the FMD controller chip includes a dedicated "engine" that performs a memory-block-analysis function. This continuous scanning and diagnostic assessment of each NAND block effectively manages its optimal time for reclamation and ensures the expectations for enterprise data.

**RAID**

In addition to the NAND block protection, each FMD is protected by RAID. Options are RAID-1, RAID-5 and RAID-6; all options are managed by VSP G1000, VSP, HUS VM or HUS 150.

**Write Endurance Protection**

Organizations using HAF are protected from FMD failure when covered under a maintenance service and support agreement. In addition, FMDs are covered by a 5-year warranty. If an FMD reaches 95% of its specified write endurance, a service information message (SIM) is issued to prompt the replacement of the FMD. Note: A service processor (SVP) or a report from Hitachi Storage Navigator can confirm the write capacity of the FMD.

**Flash Controller ASIC**

At the heart of the FMD is a custom flash controller ASIC engineered by Hitachi (see Figure 6). This ASIC is a quad-core, 1GHz, 32-bit processor complex. It provides unprecedented processing capability for a flash controller, allowing FMDs to avoid the processing pitfalls of other flash controllers.

**Architectural Highlights**

1. Highly parallel architecture.
   a. 8 lanes of PCIe.
   b. PCIe root complex.
   c. Patented flash logic that supports 32 paths to the flash storage.
2. A 4x ARM processor complex that delivers more than sufficient compute power to manage multiple, parallel tasks and service concurrent I/Os.

3. Integrated DDR-3 interface.

4. Support for 128 flash memory chips.

Figure 6. Hitachi Custom Flash Controller Chip

An area of contention within most flash controllers is the number of paths that access the flash storage. Most controllers offer 8, 12 or perhaps 16 paths. Hitachi chose to incorporate 32 parallel paths, combined with the power and flexibility of the multicore processor. This approach enables the parallel processing of multiple tasks. It also allows the removal of housekeeping tasks from the I/O path (wear leveling, ECC, and so forth), eliminating the potential of host I/O blocking.

Eight lanes of PCIe v2.0 are integrated into the chip. The PCIe interface connects to an external SAS target mode controller, providing SAS interconnectivity. This host-side interface includes a full root complex capability. A state-of-the-art DDR3 memory interface completes the FMD controller ASIC.

Key Benefits of Flash Module Drive Controller Design

Write Cliff

The degradation of I/O processing when managing housekeeping tasks and/or when under heavy write workloads is a common occurrence with many flash controllers. Essentially, this degradation is an issue of processor schedule handling, where housekeeping tasks such as garbage collection and wear-leveling demands cause a block in the host I/O. Simply put, it is a symptom of a controller that does not have sufficient processor power to gracefully handle sustained high-performance write environments.

The write cliff is not a concern with the FMD, which removes the housekeeping tasks from the I/O path. This approach is possible thanks to use of a quad-core processor that has more than sufficient processing power.
Wear Leveling

Flash has a limited lifespan. An important characteristic of flash memory is that each write or erase cycle stresses the cell, causing a deterioration of the cell over time. Compounding this fact is the probability of an imbalance or a bias in cell activity, which creates an irregular distribution of these unstable cells.

Resident within the FMD is the necessary intelligence to minimize these challenges and to optimize the life of the NAND flash. The FMD controller monitors the rate of writes, erasures and refreshes to eliminate biased activity by balancing the rate of deterioration in the flash memory blocks. It then manages the physical data location with the best match between the I/O frequency on the data and the NAND flash block usage status.

This durability enhancement is achieved with wear-leveling management and is key in extending the useful lifetime of the NAND flash that populates the FMD.

Data Integrity

To preserve the integrity of data writes, 42 bits of ECC are appended for every 1KB of data written. This action translates to each FMD ECC being able to correct up to 42 bits per 1.4KB, which exceeds the standard MLC spec of 24 bits per 1KB of data. It ensures that even if a bit error is discovered, it can be easily recovered.

Adaptive Data Refresh

To catch bit errors quickly, the FMD will perform what is called a high frequency “adaptive” data refresh. This refresh is where the controller reads and recalculates the internal ECC on the entire FMD every 2 days and dynamically optimizes page refresh based on applied error correction. Note: To secure the integrity of previously written data, all data is rewritten at least every 30 days. This practice is key to not only extending flash cell longevity but also to improving the overall sustained performance.

Periodic Data Diagnosis or Recovery

A uniquely powerful integration of features occurs when HAF is used with one of the supported Hitachi storage systems. The data diagnostic and read retry functions of the FMD are partnered with the periodic data diagnostic and recovery functions resident within Hitachi storage controllers. If the bit errors exceed the ECC correction capability of the FMD, then the data is read out by the read retry function. This function adjusts the parameters of the flash memory and reads the data. The area is then refreshed, meaning that the data is read and copied to a different area before the data becomes unreadable.

High-Speed Formatting

The formatting is done autonomously in the FMD. This highly efficient process is completed in approximately 60 minutes, regardless of the number of drives to format. When compared to the 280 minutes that a similarly configured array of SSDs (22.4TB) would take, it is apparent that systems employing the FMD have a reduced install time.

Block Write Avoidance

Not only does the flash controller chip manage 128 NAND chips, but also it supports a feature, called “block write avoidance,” to enhance write endurance and performance. Any data stream of all "0s" or "1s" is recognized by the algorithm, in real time, which remaps the data with a pointer. In the case of a parity group format, this technique can deliver up to a 94% savings in storage space. A beneficial side effect of this space savings is to effectively increase the overprovisioned area. This extra overprovisioning results in more efficiently run background tasks such as garbage collection and wear leveling and, most importantly, an improvement in the sustained write performance.

This feature not only drastically reduces format times, but also, by eliminating unnecessary write/erase cycles, it extends the effective life of the flash memory.
Note: Unlike traditional commodity SSDs, the FMD integrates with the active data features such as Hitachi Dynamic Tiering page zero and UNMAP.

Innovative Flash Acceleration Software

The Hitachi development team's first task was to dissect the existing storage system firmware and look for opportunities to optimize the code to support flash. An inviolate design principle was that flash optimization was to be achieved without compromising the many storage services and features already available. The result was a complete "flash-optimized" refresh of the firmware that produced significantly improved I/O processing, higher internal thread count support and faster internal data movement. It also reduced response time considerably. These changes combined to double the system performance and, while there were numerous changes to the firmware, they were implemented without the need for any hardware changes.

These microcode changes morphed into "flash acceleration," a feature that is transparent to other functions such as dynamic provisioning, dynamic tiering and local or remote replication. Flash acceleration is enabled or disabled non-disruptively with a license key. This software is included with every Hitachi Virtual Storage Platform G1000, Hitachi Virtual Storage Platform, and Hitachi Unified Storage VM. Flash optimization software provides an additional improvement in random reads through a reduction in processes overhead for flash I/Os. It is available as an optional license with HUS VM.

Hitachi Accelerated Flash Portfolio

HAF is supported on Hitachi Virtual Storage Platform G1000, Hitachi Virtual Storage Platform, Hitachi Unified Storage VM and Hitachi Unified Storage 150.

"All Flash" Storage Systems

AFA configurations are supported on all storage systems that support HAF (see Figure 7).

Hybrid Storage

Hybrid storage is the intelligent integration of multiple media types such a SAS, NL-SAS, SSD and HAF.

Hitachi Dynamic Tiering (HDT) delivers central management of all Hitachi storage platforms. The net end-user benefit of this software-managed combination of disparate media types must be greater than the individual contribution.
Maximize the Value of Flash Storage With Hitachi Dynamic Tiering

With the integration of SSD technology into the data center and the evolution of hybrid storage (SSD + HDD), the initial obligation was to manage the placement of data manually. However, if the promise of solid-state storage was to be realized, this manual task had to be eliminated.

The answer was to enable the storage system to move data between media types autonomously based on the access profile and optional policies a data user had established (see Figure 8).
Hitachi Dynamic Tiering software delivers this functionality. It eliminates the complexities of managing data placement and optimizes the cost-to-performance effectiveness of the storage solution. HDT simplifies the administration of storage by eliminating the need for time-consuming manual intervention. No complex decision criteria are required; data is moved based on simple logic around user policies, media type and speed, and sustained I/O requirements. One to 3 tiers of storage can be defined within a self-managed pool using any storage media types available to the Hitachi platform. Thin provisioned managed pages (32MB to 42MB depending on the platform) are moved between Dynamic Tiering tiers to the most appropriate media based on the workload. For example: Highly active database data migrates to flash storage; aging data that has not been touched for some time is demoted to less performance, and hence less expensive, media.

Central Management: Hitachi Command Suite

Hitachi Command Suite (HCS) delivers central management of all Hitachi storage platforms (see Figure 9).

When flash (SSD or HAF) is introduced into a Hitachi storage system, it integrates seamlessly and without impacting or altering the management scheme. HCS can see and manage flash storage in the same way it manages disk drives.
Flash Module Drive Performance and Other Metrics

Hitachi Storage Performance Comparison

**TABLE 1. COMPARISON OF VSP G1000, VSP, HUS VM AND HUS 150**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Hitachi Virtual Storage Platform G1000</th>
<th>Hitachi Virtual Storage Platform</th>
<th>Hitachi Unified Storage VM</th>
<th>Hitachi Unified Storage 150</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance (IOPS)</td>
<td>3,850,000</td>
<td>1,000,000</td>
<td>500,000</td>
<td>220,000</td>
</tr>
<tr>
<td>Latency</td>
<td>&lt;1.1ms with high consistency</td>
<td>&lt;1ms with high consistency</td>
<td>&lt;1ms with high consistency</td>
<td>&lt;1ms with high consistency</td>
</tr>
<tr>
<td>Capacities Supported</td>
<td>Split 2,026TB</td>
<td>675TB</td>
<td>2,026TB</td>
<td>845TB</td>
</tr>
<tr>
<td>Flash Module Drive (FMD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sequential I/O (Write)</td>
<td>49Gb/sec</td>
<td>13Gb/sec</td>
<td>11Gb/sec</td>
<td>6Gb/sec</td>
</tr>
<tr>
<td>All Flash Configuration</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Hitachi Dynamic Tiering Support</td>
<td>HAF, SSD, SAS, NL-SAS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Management</td>
<td>Hitachi Command Suite</td>
<td>Hitachi Command Suite</td>
<td>Hitachi Command Suite</td>
<td>Hitachi Command Suite</td>
</tr>
</tbody>
</table>

**Table 1 Key Takeaway**

Hitachi Accelerated Flash is a purpose-built flash technology that can deliver nearly 4 million IOPS when deployed in high-performance Hitachi storage systems. At the heart of HAF is a Hitachi flash module drive that the controller integrates transparently and synergistically with Hitachi storage platforms. Translation: All existing storage features and functions, such as management, environmental services, dynamic tiering, dynamic provisioning, local and remote replication, and so forth, are supported in a HAF configuration. They are supported whether hybrid or all flash.
Pricing

Table 2 is based on a single HUS VM capable of delivering 1,000,000 IOPS. The metrics were developed internally using industry sources as references.\(^2\)\(^3\)\(^4\)\(^5\)

**TABLE 2. COST INFORMATION* BASED ON 1,000,000 IOPS HUS VM CONFIGURATIONS**

<table>
<thead>
<tr>
<th>Performance (IOPS)</th>
<th>FMD</th>
<th>SSD</th>
<th>HDD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,000,000</td>
<td>16</td>
<td>80</td>
<td>768</td>
</tr>
<tr>
<td>500,000</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>500,000</td>
<td>13U</td>
<td>69U</td>
<td></td>
</tr>
<tr>
<td>Device Count</td>
<td>1</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Rack Space</td>
<td>13U</td>
<td>69U</td>
<td></td>
</tr>
<tr>
<td>Price per I/O</td>
<td>$1</td>
<td>$3</td>
<td>$7</td>
</tr>
</tbody>
</table>

* The pricing used was US List, current at time of writing.

**Table 2 Key Takeaway**

Based on the configurations that could realize 500,000 IOPS there is a 3x cost advantage when using FMDs instead of traditional SSDs. The HDD numbers have also been included for reference.

Space and Power Efficiency: Capacity Growth Path

**TABLE 3. EFFICIENCY METRICS AT DIFFERING FMD CAPACITIES**

<table>
<thead>
<tr>
<th>FMD Capacity</th>
<th>1.6TB</th>
<th>3.2TB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum FMD per Tray</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Raw Capacity per Tray</td>
<td>19.2TB</td>
<td>38.4TB</td>
</tr>
<tr>
<td>Raw Capacity per Rack Unit</td>
<td>9.6TB</td>
<td>19.2TB</td>
</tr>
<tr>
<td>Power per Tray</td>
<td>369W</td>
<td>381W</td>
</tr>
<tr>
<td>Power Efficiency (Watts per terabytes, per FMD, including chassis)</td>
<td>19.2</td>
<td>9.9</td>
</tr>
</tbody>
</table>

**Table 3 Key Takeaway**

The purpose of Table 3 is to illustrate the capacity growth path for the flash module drive. Currently available FMD capacities are the 1.6TB and 3.2TB.

---

\(^2\) [http://wikibon.org/wiki/Performance_Centric_Data_Storage_TCO_Study_Demonstrates_Cost_Advantages_of_Flash_Arrays](http://wikibon.org/wiki/Performance_Centric_Data_Storage_TCO_Study_Demonstrates_Cost_Advantages_of_Flash_Arrays)

\(^3\) [http://www.thestoragealchemist.com/flash-is-really-about-economics/?sf12791614=1](http://www.thestoragealchemist.com/flash-is-really-about-economics/?sf12791614=1)

\(^4\) [http://wikibon.org/wiki/Cape_Study_The_Hunting_of_the_RARC](http://wikibon.org/wiki/Cape_Study_The_Hunting_of_the_RARC)

Space and Power Efficiency Comparison Chart: Comparing HAF, SSD and HDD

Table 4 is based on a populated HUS VM with HAF, SSD or high-performance HDDs. While based on the HUS VM frame, the comparison highlights the media differences independent of where hosted. The results are assembled to illustrate the efficiency gains that are available with HAF. The HAF numbers are based on the 3.2TB capacity FMD.

**TABLE 4. SPACE AND POWER EFFICIENCY COMPARISON CHART**

<table>
<thead>
<tr>
<th></th>
<th>Flash Module Device</th>
<th>Solid-State Disk</th>
<th>High-Performance 15K HDD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capacity per Drive</td>
<td>3.2TB</td>
<td>400GB</td>
<td>300GB</td>
</tr>
<tr>
<td>Maximum (Max.) Devices Supported</td>
<td>576</td>
<td>128</td>
<td>1152</td>
</tr>
<tr>
<td>Max. Trays Supported</td>
<td>48</td>
<td>6</td>
<td>48</td>
</tr>
<tr>
<td>Max. Devices per Tray</td>
<td>12</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>Rack Units (Minimum/Max.)</td>
<td>2/96</td>
<td>2/12</td>
<td>2/96</td>
</tr>
<tr>
<td>Total Raw Capacity (30% FMD over provisioning included)</td>
<td>2,026TB</td>
<td>51.2TB</td>
<td>345.6TB</td>
</tr>
<tr>
<td>Raw Capacity per Tray</td>
<td>38.4TB</td>
<td>9.6TB</td>
<td>7.2TB</td>
</tr>
<tr>
<td>Raw Capacity per Rack Unit</td>
<td>19.2TB</td>
<td>4.8TB</td>
<td>3.6TB</td>
</tr>
<tr>
<td>Power per 2U Tray</td>
<td>381W</td>
<td>295W</td>
<td>345W</td>
</tr>
<tr>
<td>Power Efficiency per 2U Tray</td>
<td>9.9W/TB</td>
<td>31W/TB</td>
<td>49W/TB</td>
</tr>
</tbody>
</table>

*Table 4 Key Takeaways*

1. **Space Savings:** FMDs have a significant advantage in rack storage density with a 4x advantage over SSDs and a 5.33x advantage over high-performance HDDs.

2. **Power Savings:** FMDs are more power efficient and use 68% less power than SSDs and 80% less power than high-performance HDDs.

**Conclusion**

IT departments have finite budgets. Thus, cost along with robustness have been the primary constraints that have limited a wider adoption of solid-state storage, particularly in the data center.

With these concerns in mind, Hitachi perceived the following challenges: How can we provide an enterprise-class, high-performance, robust, flash-based storage option that has a 5-year endurance specification? And, how do we do so cost-effectively?

The Hitachi development team’s answer was to create a purpose-built flash controller that exploits the attractively priced MLC NAND technology. At the same time, the controller delivers the performance, functionality, robustness and endurance expected in an enterprise-class solution.
Appendix A: Contributors

The information included in this document represents the expertise, feedback and suggestions of a number of contributors from Hitachi Data Systems. The author would like to recognize the following individuals:

- Matt Pujol, product management
- Walter Amsler, product planning
- Ed McElearney, technical operations
- Mike Nalls, product marketing
- Mark Adams, product marketing