Vendor Profile

Violin Memory: Laying the Foundation for the All-Silicon Data Center

Eric Burgener

IDC OPINION

Enterprises that think that flash is still too expensive to deploy in their data centers are mistaken. Flash optimized storage architectures are delivering more aggressive $/IOPS, IOPS/TB, and $/transaction than ever. When it comes to dedicated workloads hosted on 3rd computing platforms that leverage virtual infrastructure and require high performance, HDD can no longer cost-effectively compete. But this alone will not usher in the era of the all-silicon data center.

Being able to reliably consolidate mixed workloads has been the strength of traditional storage architectures. With flash technology now widely available in shared storage arrays, the industry is poised to undergo a transformation that very likely will bring the all-silicon data center about for more than just specialized environments. What has been widely lacking in flash optimized architectures, however, has been the enterprise-class reliability, availability, scalability and management functionality (i.e. mature data services) necessary to host mixed data center workloads. Flash array vendors understand this and are working to provide that functionality in their flash optimized storage architectures, intending to position their products as a viable replacement for enterprise-class storage arrays during the next refresh cycle.

Violin Memory, an early entrant into the flash optimized storage architecture game, has already been selling solutions that allow their customers to consolidate mixed workloads onto their platforms and, in the process, save hundreds of thousands to millions of dollars. With a new management team and a more finely tuned focus and strategy around leveraging backplane memory-based architectures in shared storage solutions, Violin is clearly positioning their products for dense, mixed workload consolidation with the functionality and bellwether references to back it up. With their latest release in particular, the Concerto 7000, they are offering a compelling argument for the all-silicon data center.

IN THIS VENDOR PROFILE

This IDC Vendor Spotlight discusses the emergence of a data center, flash optimized storage architecture that IDC believes will ultimately replace traditional storage architectures as the enterprise storage workhorse of the future, making the all-silicon data center (for primary storage) a very viable proposition. One of the early entrants in this space, Violin Memory, is then profiled.

SITUATION OVERVIEW

Flash as a storage medium has quickly penetrated the data center. With the 3rd computing platform build out, driven by mobile, social media, big data/analytics and cloud, in full swing, spinning disk just cannot cost-effectively deliver on the performance requirements of a very dynamic, increasingly agile virtual infrastructure. Flash performance, characterized by extremely high throughput and low latencies, has driven its rapid acceptance in this new era.

Flash packaging in arrays has gone through several iterations in short order. Flash was initially packaged as a cache which could front end existing spinning disk arrays. It could be either added into the array in a special "flash tray", or the entire array could be front-ended by a flash appliance. Capacities were usually quite small, but efficient caching algorithms could maximize cache hit rates to deliver flash performance a high percentage of the time for most applications. This was an intelligent approach for vendors, since enterprises could easily add a few flash devices to existing configurations, preserving existing investments while familiarizing themselves with the benefits and characteristics of this new technology.

In the next phase, vendors began to architect small storage arrays that had been built specifically with flash in mind, and sell them for use with a single application environment that required extremely high performance. Although vendors did not use the term, this was the dawn of the concept of the "flash optimized storage architecture". These systems included features such as expanded controller and internal system architecture bandwidth, write minimization, wear-leveling, and free space management (referred colloquially to as "garbage collection") that specifically took into account the different way that flash media handled read and write I/O. Given that flash delivered three orders of magnitude better throughput and latency than spinning disk, the bandwidth limitations of legacy storage architectures were a significant handicap in using flash optimally. Greater internal bandwidth as well as other targeted enhancements began to resolve early data center concerns with flash media usage, which included issues like peak performance capabilities, endurance, reliability and predictability.

Now, the industry is on the cusp of a new horizon in flash-based arrays: the flash optimized, enterprise class storage solution. Flash as a technology in the data center is here to stay, and IDC believes that all enterprises should already have at least some flash deployed. Flash will ultimately transform data center storage solutions, but to do so it must be intelligently packaged in flash optimized storage architectures that are designed not only for performance and durability, but also to meet high bars for reliability, availability, scalability and ability to integrate into existing data center workflows. The signature of the enterprise storage workhorse is its ability to reliably host mixed application workloads, predictably delivering needed performance as configurations scale, while meeting manageability and availability requirements.

The Next Wave: Flash Optimized Enterprise Arrays

For flash optimized arrays to start to become viable replacements for legacy enterprise arrays, several key areas have to be addressed. First, they must provide reliability on par with those arrays, leveraging write minimization, wear leveling, and garbage collection to deliver endurance that at least meets a standard 3 to 5 year depreciation life cycle. Data protection schemes should be configurable so that RAID overhead can be traded off against required resiliency to meet specific requirements.

Data protection must include replication capabilities so that an effective disaster recovery strategy can be implemented for applications hosted on the array. The array should support shared access to storage in clustered environments to support failover and live migration operations. [Note that failover and live migration are two very different operations used for different purposes.]

To be viable as a platform for mixed workloads, the array must support sufficient capacity. Arrays that support only up to 32TB or 64TB are not large enough for the secondary economic benefits of flash deployment – far fewer devices to meet performance requirements and much lower energy and floor space consumption relative to spinning disk solutions – to really start to impact the total cost of ownership (TCO) of storage solutions. Platforms need to support at least hundreds of terabytes, with a well defined growth path to petabyte support under a single management umbrella. As workload consolidation density and overall flash capacity increases, these secondary benefits will start to have a real impact.

Scalable, enterprise-class data services must also be available. In-line optional compression and deduplication, designed to work in memory to minimize write I/O to flash, allow customers to optimize data reduction strategies based on application requirements. All storage should be thin provisioned all the time, and this feature will work in conjunction with data reduction to make the most of available flash capacity. Scalable snapshot and clone implementations that do not impact performance as more of them are created and/or retained are critical as inputs to administrative and maintenance operations, test and development environments, or data protection regimens.

One of the most pressing concerns of IT management, based on recent IDC survey data, is access to sufficiently skilled IT expertise. As budget considerations force IT headcount reductions, more and more administrative tasks are moving to IT generalists. These personnel generally have strong virtual server management skills, but are less familiar with storage operations. Automation is becoming a key mantra among CIOs as a way to reliably increase administrative span of control as staffing levels shrink. For an array to be considered a viable platform for dense workload consolidation, it must fit in with this strategy and be easy to integrate into pre-existing data center workflows. Support for snapshot APIs such as Microsoft Volume Shadowcopy Services (VSS), VMware APIs for Data Protection (VADP), Oracle Recovery Manager (RMAN) and others are required. APIs like Microsoft ODX and VMware VAAI that support host off-load for various data services are important. And management APIs like SNMP, SMI-S and others that enable monitoring and management through centralized consoles can help to increase the administrative span of control, making the storage solution easier to manage to business objectives.

The Importance of Quality of Service (QoS) To Delivering Predictable Performance

Garbage collection was referred to briefly earlier. There is the potential for garbage collection to interfere with real time I/O performance, thereby unpredictably inserting added latencies into I/O operations (in particular write I/O). In enterprise environments, the requirement to deliver predictable performance as a storage configuration scales is paramount for many production applications. Flash systems targeted for use in these environments therefore have to consciously address the possibility that garbage collection could interfere with the ability to deliver predictable performance.

There are two basic designs in use for all flash arrays: those that use backplane memory and those that use solid state disks (SSDs). With backplane memory-based architectures, the system can be

built to have absolute awareness of what is going on at the individual flash cell level, using this information to definitively see when and where garbage collection operations are going on. This information can then be used to manage real time I/O and garbage collection operations collectively *at a systems level* to ensure that the two do not interfere with each other. Managing quality of service is therefore integrated into the other flash management operations that are performed at the system level, and these systems do not require a separate set of QoS controls to deliver predictable performance as configurations scale. QoS is in effect "built in". This design will offer the most predictable performance, particularly as the storage system becomes more heavily loaded.

In SSD-based systems, device-level garbage collection is performed by the SSD devices themselves, and the flash system does not necessarily have full visibility regarding what is happening inside these devices. If the system attempts real time I/O on a device that is performing garbage collection, latencies could be unpredictably impacted. For that reason, a separate set of QoS controls, operating at the system level, are required to collect that information at the device level and manage performance. This is required in addition to using separate processors to handle real time I/O and garbage collection, because ultimately the I/O for both operations comes down to individual flash cells. This design can deliver very predictable performance, but as systems become more heavily loaded does have the potential to not deliver the same level of predictable performance as backplane memory-based flash systems.

Enterprise storage systems are expected to deliver predictable performance, and flash-based systems must be architected to deliver that, but that does not mean that they have to have separate QoS controls to do so. It depends on the specific system architecture implemented.

A Note About Data Reduction

Compression and deduplication are the two popular forms of data reduction in use, but because they operate very differently they deliver differing data reduction performance on different workloads. Compression performs better than deduplication on file, print, database, mail, and mixed virtual workloads, delivering data reduction ratios of from 2:1 to 4:1. Deduplication, if performed, would provide minimal additional reductions and for that reason is likely not worth the additional processing overhead and latencies it may introduce. For virtual desktop infrastructure (VDI) environments, however, where hundreds or thousands of desktops may be running the same exact version of Windows and the same applications (e.g. Microsoft Office, etc.), deduplication can provide data reduction ratios by itself in the range of 7:1 to 9:1, making it extremely valuable. When data reduction is performed in-line, it not only increases usable capacity but also contributes to write minimization, a valuable benefit in flash-based systems, because less data has to be written through to flash.

For optimal efficiency, the data reduction technology should be matched to the workload type to optimally balance the additional processing overhead and latency with the space savings benefits. Systems that offer optional compression and deduplication offer good flexibility, and those that would allow them to enabled or disabled at more granular levels, such as for individual applications, would provide the most flexibility. For workloads that require latencies in the 100 – 200 millisecond range or where it provides minimal space savings, it can be entirely turned off. For workloads that require latencies in the 500 – 700 millisecond range compression, deduplication or both can be turned on or off as appropriate based on the application.

Because of initial space savings and write minimization issues, data reduction that is done in a post-processing manner is less valuable than data reduction done using in-line methods, but the actual latency impacts must also be evaluated to ensure that service level agreements (SLAs) can be met.

The All-Silicon Data Center

As flash component prices continue to drop, flash systems become more price competitive with traditional HDD-based systems. As enterprises start to deploy flash in larger volumes and with mixed workloads, the secondary economic benefits make flash systems aggressively competitive with HDD. Metrics like $/IOPS, IOPS/TB, $/transaction and especially TCO are much better ways to evaluate the value that flash systems bring to the table. As these systems host denser, consolidated workloads, this will become even more true. It is not unreasonable to expect in the future that all primary storage could be cost-effectively hosted on flash, particularly if that data center is heavily leveraging the 3rd computing platform. Once flash optimized storage solutions that also meet the enterprise requirements discussed above become widely available, the all-silicon data center becomes a very real future.

Company Overview

Violin Memory was a pioneer in the flash space, shipping their first product, a backplane memory based caching appliance, in 2009. The original founders of the company were primarily chip architects focused on designing what at the time may have been the first "flash optimized storage architecture" to get the most I/O performance, in terms of high throughput and low latency, out of a storage system. The choice of a backplane memory-based architecture was critical in meeting this objective. Violin's solutions portfolio includes the Violin Maestro All Flash Caching Appliance, Violin 6000 All Flash Array, the Violin Windows Flash Array, and the Violin Concerto 7000 – all leveraging their Flash Fabric Architecture for high storage performance, resilience and density. Violin sells flash-based storage solutions through a direct sales force and a worldwide network of resellers and OEMs, including Dell, NEC, Fujitsu and others.

Their customer base includes close to 400 mostly Fortune 2000 customers worldwide in financial services, manufacturing, media and entertainment, healthcare, government, internet, retail and other markets. They have had considerable success selling to market leaders in their industries, with 3 of the top 10 largest corporations, 6 of the top 10 largest telcos, 3 of the top 10 largest retailers, the top 5 largest software companies, and 9 of the top 15 largest IT companies in the world as their customers.

Company Strategy

Taking a cue from past experience, Violin has recently diversified their revenue stream across a larger number of direct, channel and OEM sources. They have divested themselves of the PCIe-based flash card business, and going forward will concentrate their efforts on the shared all-flash array solutions model, building on the success they've already had consolidating mission-critical and other enterprise application workloads at Fortune 1000 companies in the financial services, telco, media and entertainment, healthcare, retail, technology and other sectors worldwide. It is clear that they are throwing their hat into the ring to compete as the next generation enterprise storage workhorse, driving the price points and functionality that make the all-silicon data center a viable option in the future.

It is interesting to note that Violin products are already being used as consolidation platforms for mixed workloads, with storage savings spanning hundreds of thousands to millions of dollars at individual Fortune 1000 customers. Violin is farther down this path, particularly with the release of the Concerto 7000, than most other flash array vendors.

Violin's strategy to go after the Windows and Hyper-V markets with a solution that leverages the Microsoft storage stack is also interesting, and a first for a flash array vendor. The Windows branding and functionality, based on joint development work done by both Microsoft and Violin and embedded in the Windows distribution (but only accessible by Violin), could be very attractive to Microsoft-only accounts. When enterprises are looking to consolidate Windows, Unix and/or Linux workloads together on the same storage platform, the Concerto 7000 will likely be more attractive.

FUTURE OUTLOOK

The company grew revenues rapidly to $74M in FY13, but was struggling in the wake of their IPO in September 2013. The board brought in a new CEO, Kevin DeNuccio, in February 2014, and in four short months he has brought in an almost entirely new executive management team and renewed the company's focus in leveraging backplane memory-based architectures for the shared flash array market. Today, Violin is a newly energized company poised to succeed.

Although DeNuccio has remade the company, he has identified and brought forward key competitive assets that will be critical to their future success. Violin's maturity with flash optimized storage architectures gives them an advantage over later market entrants with less data center workload consolidation experience. Enterprise-class reliability, availability, and scalability are highly valued by CIOs for good reason: they are necessary capabilities that take time to prove out.

Two key product introductions this year, the Windows Flash Array in April and the Concerto 7000 in June, signify the company's serious intent to pursue the next horizon in flash arrays – more scalable storage solutions with enterprise-class reliability, availability and data services whose intended use will be mixed workload consolidation. The Windows Flash Array, targeted obviously at the Windows market, leverages data services in the Microsoft storage stack to offer a powerful yet familiar platform for Windows administrators, while the Concerto 7000 is targeted at mixed enterprise workloads across Windows, Linux and Unix environments, providing in-house developed, enterprise-class data services that leverage hardware assist in many areas. The strategy with the Windows Flash Array leverages Microsoft's own experience delivering storage-based data services in Windows environments, while the Concerto 7000 leverages Violin's own experience, honed through the production use of many of these capabilities over several years with the Violin 6000 All Flash Array in large enterprise accounts.

Will the industry get to an all-silicon data center, and if so, how quickly? A key driver for this will be the TCO of all flash solutions, and that will be driven in large part by flash vendors' ability to reliably consolidate mixed enterprise workloads on their platforms. With the reboot of Violin, they have made clear they understand not only that this is the next battleground, but what it takes to win.

Challenges and Opportunities

Despite Violin's new direction, they have to overcome the market's consternation with their post IPO performance in 2013. Clear and consistent marketing messages around their new focus and their past success with mixed workload consolidation are the order of the day. Competitors still attempt to pigeonhole Violin as a dedicated, single application flash-based solution. Until recently, their largest capacity system supported only 64TB of flash, but with the introduction of the Window Flash Array and the Concerto 7000 they can cluster four nodes of the same type together under a single namespace for 256TB of raw flash capacity. Given their in-line compression and deduplication, customers can easily double or triple this for usable capacities approaching a petabyte. Replication capabilities that make it easier to craft disaster recovery solutions for Violin configurations are relatively new, and Violin needs to prove that they, along with the other data services they support, are reliable and scalable under the demands of enterprise environments.

With a mature flash optimized architecture, significantly increased scalability, and a strong group of referenceable bellwether accounts who have already used their solutions for workload consolidation, Violin has good opportunities open to it. Two strong new products announced this year are clearly intended to take them down what IDC thinks is the path to success in the enterprise for flash arrays – providing a reliable, scalable workload consolidation platform for enterprise applications. Violin needs to create sales momentum around these new products and make sure there is strong awareness about their successes in the market. The new chief marketing officer, Eric Herzog, is already getting the message out to prospects about exactly how much money their customers have saved by moving mixed workloads off of legacy storage onto Violin platforms. The magnitude of these savings will turn CIO heads.

ESSENTIAL GUIDANCE

Flash technology provides undeniable performance benefits to the data center when deployed in small capacities, and offers significant economic benefits as well when deployed in larger capacities. The industry is quickly passing through the adolescence of flash, and with the next goal for flash array vendors clearly defined entering adulthood. That goal is to provide the replacement for HDD-based enterprise storage workhorse arrays. All enterprises should already be deploying flash in at least some capacity, and should closely evaluate flash optimized, enterprise-class storage architectures during their next technology refresh cycle.

Despite a bit of trouble in the recent past, Violin has much to recommend it to customers looking for a flash-based mixed workload consolidation platform: a proven flash optimized architecture, enterprise-class data services that enable mixed workload consolidation, and a track record in doing exactly that with a number of bellwether Fortune 1000 accounts. This is a compelling message for enterprises looking to upgrade their storage platforms to better accommodate the needs of the 3rd computing platform. Now that the company is focused on the use of backplane memory-based architectures for shared storage arrays, Violin needs to ensure that the market is aware of what this remade company has to offer.

LEARN MORE

Related Research

Worldwide Storage Predictions 2014: Storage Disruption – Flash, Cloud and Software-Based Storage (IDC #WC20140109, January 2014)

Worldwide Hybrid External Enterprise Storage Systems 2013 – 2017 Forecast (IDC #245117, December 2013)

All-Flash Array Performance Testing Framework (IDC #241856, June 2013)

Worldwide Enterprise All-Solid State Storage Array 2013 – 2016 Forecast (IDC #240424, April 2013)

IDC's Worldwide SSD and Storage Tiering Taxonomy, 2012 (IDC #237759, November 2012)

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

Global Headquarters

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-insights-community.com
www.idc.com