All-flash storage arrays have pitted longtime storage vendors against pesky startups. This guide will help you compare flash storage products and decide if an all-flash array is right for you.
Contents

All-flash arrays: comprehensive chart
Page 2

IBM FlashSystem 840 focuses on performance, V840 is more feature-rich
Page 11

Pure Storage bolsters FlashArray line with more capacity choices
Page 18

Experts weigh In On all-flash arrays
Page 28

When to consider the purchase of an all-flash array
Page 31

Table of Contents

Everything you wanted to know about all-flash arrays in one big chart
  • All-flash array comparison chart..........................................................Page 2

Devil in the details? Snapshot profiles of all-flash array vendors
  • Dell flash strategy centers on updated Compellent, EqualLogic arrays..........................................................Page 2
  • EMC touts performance in scale-out XtremIO, aims VNX-F at price-conscious...Page 4
  • HP makes case that all-flash 3PAR StoreServ 7450 is designed for flash........Page 6
  • Hitachi boosts all-flash HUS virtual machine with custom drives, optimized software..........................................................Page 9
  • IBM FlashSystem 840 focuses on performance, V840 is more feature-rich.....Page 11
  • Kaminario beefs up all-flash array with enterprise storage features.............Page 13
  • NetApp offers all-flash EF-Series, FAS systems prepping new FlashRay.......Page 14
  • Nimbus Gemini all-flash array line set to add higher density option.............Page 16
  • Pure Storage bolsters FlashArray line with more capacity choices.............Page 18
  • Skyera turns heads on price with skyHawk, readies larger skyEagle..........Page 19
  • SolidFire seeks to grow customer base beyond cloud to enterprise IT........Page 21
  • Tegile breaks hybrid focus with release of new T3800 all-flash array..........Page 23
  • Violin extends focus beyond speed, plans enterprise storage features........Page 25

Experts weigh in on all-flash arrays
  • All-flash arrays: Performance vs. function.............................................Page 28
  • When to consider the purchase of an all-flash array.............................Page 31
  • Making sense of all-flash array vendors’ wild performance claims............Page 32
EVERYTHING YOU WANTED TO KNOW ABOUT ALL-FLASH ARRAYS IN ONE BIG CHART

TechTarget's all-flash array guide profiles 15 product lines from 13 vendors, large and small. Click the link below to check out our all-flash array comparison chart.

All-flash array comparison chart

All-flash arrays may differ on capacity, drive type, networking options and storage-savings features, but they all improve performance. However, comparing products to each other can be difficult, because vendors use differing configurations, workloads and block sizes to test their products.

Price is also tough to compare. Prices per GB vary among vendors based on their products' configurations and whether or not they factor in data reduction technologies such as deduplication and compression. Prices per IOPS are no easier.

Our all-flash array comparison chart provides an overview of all-flash arrays on a variety of important criteria that will help you decide on the right array for your needs. You will find information on storage architecture, flash type, array capacity, networking technology, warranty/support, and storage-saving and management features.

The chart can be printed in landscape mode, scaled to fit the page, if desired. Download our all-flash array comparison chart now.

DEVIL IN THE DETAILS? SNAPSHOT PROFILE OF ALL-FLASH ARRAY VENDORS

We profiled 13 vendors of all-flash arrays and spotlighted their products’ areas of distinction and deficiency compared to the others.

Dell flash strategy centers on updated Compellent, EqualLogic arrays

Dell Inc.’s Compellent and EqualLogic all-flash arrays trace their origins to hard disk drives, but they’ve had plenty of time to pick up new features and capabilities geared for solid-state storage.
Compellent customers technically could create an all-flash configuration as far back as October 2009 by loading up a standard array with solid-state drives (SSDs). But with a major platform upgrade in August 2013, Compellent arrays gained such cutting-edge abilities as automated tiering of sub-LUN data across different types of SSDs. Users can tier between write-intensive single-level cell (SLC) drives and read-intensive multilevel cell (MLC) drives and tack on a layer of hard disk drives (HDDs) if they choose.

Future Compellent SC8000 releases are due to include such capabilities as inline deduplication and compression, and quality of service (QoS), as well as to add support for RESTful application programming interfaces and OpenStack cloud technology drivers. Those enhancements would bring the product into line with some of the more feature-rich all-flash arrays.

The Compellent all-flash array is among the highest-capacity options in the market, with a maximum raw capacity of 1.6 PB in a dual-controller setup, and the product stands out with its support of 16 Gbps Fibre Channel and Fibre Channel over Ethernet storage networking.

Compellent is targeted to medium-sized to large organizations, while Dell's line of iSCSI-only EqualLogic arrays typically takes aim at the use cases of small- to medium-sized businesses.

Compellent's list price is $130,000 with six SLC SSDs and six MLC SSDs, and Dell claims the Compellent all-flash arrays are priced from $5 per gigabyte to $10 per GB with 48 SLC SSDs and 240 MLC SSDs.

Like Compellent, EqualLogic launched its first all-SSD-capable model in 2009. But the December PS6210S release takes advantage of controller and firmware technology redesigned specifically with flash in mind. The end result is a performance claim by Dell of 1.2 million IOPS in a configuration with a virtual pool of eight all-flash arrays.

The PS6210S all-flash EqualLogic array supports only the high-performance but more expensive SLC flash, as the industry increasingly moves to cheaper MLC technology. Dell supplied a starting price of $8 per GB for the PS6210S, based on a configuration with two dozen 800-GB SSDs, placing it roughly in the middle of the pack of the lines of all-flash arrays.

The most aggressive all-flash array vendors claim their prices are less than $5 per GB, but those products often employ data reduction technology to hit those numbers. Neither the all-flash EqualLogic arrays nor Compellent arrays support inline deduplication and compression, two of the most important storage-saving features. They support only post-process deduplication for file data within a unified storage system via the Dell Fluid File System, and...
Compellent this month added support for post-process compression, according to a Dell spokesperson.

Dell's EqualLogic division has credited the PS6210S array's redesigned controller and software for allowing it to stake a claim of flash performance at the price of disk. The company declined to supply an estimated price per IOPS.

EqualLogic supports such enterprise storage features as thin provisioning, snapshots, replication and data encryption, and unlike Compellent, it doesn't charge an extra fee for any of them. All software is included with the purchase of the array.

Like most all-flash array vendors, Dell said it would replace flash drives in EqualLogic arrays if they wear out before the warranty expires, with no limitation on the number of writes. By contrast, Compellent replaces flash drives only if they are within the SSD's "rated life" -- the maximum amount of bytes that can be written to the device. Compellent's warranty covers defects in workmanship and/or materials, but it does not cover problems related to drives reaching their maximum-rated life, according to a Dell representative.

Dell claims the Compellent SC8000 has no performance degradation during controller failure in typical cases. There could be as much as 50% degradation if the controllers are running at peak performance and maximizing I/O and bandwidth, but Dell claims that is rare if a system is properly sized. Compellent uses redundant active-active controllers.

EqualLogic arrays avoid performance degradation with active-standby controllers. Only one controller is used in normal operating, while the second controller mirrors the cache of the active controller.

**EMC touts performance in scale-out XtremIO, aims VNX-F at price-conscious**

EMC Corp.'s chief all-flash array options are the scale-out XtremIO that was purpose built for flash and the scale-up VNX-F that is an offshoot of its traditional disk-based VNX product line.

The Hopkinton, Mass.-based storage giant positions its EMC XtremIO product for customers looking for predictable low-latency performance and scalability across a broad set of workloads. EMC acquired Israel-based XtremIO in May 2012, and the first XtremIO all-flash array became generally available in November 2013.
VNX-F -- the all-flash version of the VNX midrange array family -- also shipped in November, featuring a new flash-optimized VNX2 multicore software architecture. VNX-F targets customers who want the performance characteristics of an all-flash system yet prioritize price per GB and price per IOPS, according to Aaron Chaisson, director of cross business product marketing at EMC.

"EMC offers all-flash that address a number of different price points, use cases, applications, deployment scenarios and workloads," Chaisson wrote in an email.

EMC listed the VNX-F price per GB as less than $5 for a system with 46.2 TB raw capacity but declined to provide the price per GB of the XtremIO array. EMC also declined to provide information on price per IOPS for either product. Company representatives said pricing for the least expensive configuration of the XtremIO array starts at $200,000, and the least expensive configuration of the VNX-F with 10.8 TB raw capacity has a list price of less than $240,000. A new XtremIO entry configuration option starting at 5 TB raw capacity (3.3 TB usable) and 70 TB usable logical capacity will be available in July, according to Fenselau.

The XtremIO all-flash array had limited availability in April 2013, although its general availability date was November. The product has 10 TB and 20 TB scaling units called X-Bricks that use enterprise-grade multilevel cell flash (eMLC) and can scale to 80 TB clusters.

"The advantage of the scale-out architecture is you can scale further as we qualify 8-brick and 16-brick clusters, and the performance increases linearly," said Andy Fenselau, senior director of product and solutions marketing at EMC's XtremIO business unit, via an email, noting that capacity wasn't the only metric that would expand. He said the eight-brick qualification is expected by year's end, with plans for the 16-brick to follow in 2015.

Features added with the XtremIO generally available release included thin provisioning, inline data deduplication and support for 8 Gbps Fibre Channel and 10 GbE iSCSI connectivity. Updates added this month include snapshots, data-at-rest encryption and support for REST-based APIs and an OpenStack driver, according to EMC.

Those additional features will help XtremIO to catch up to many other all-flash arrays that beat EMC to market. EMC's VNX-F notably lacks data encryption and in line deduplication.

Replication support in the VNX-F and XtremIO products is available through for-fee add-ons. VNX offers replication and three data center replication through a Remote Protection Suite. Another option for replication is the for-
fee Local Protection Suite, which also supplies snapshots. XtremIO supports replication and three data center replication through EMC’s VPLEX and by running VPLEX in combination with RecoverPoint. Later this year, XtremIO plans to add support for replication through RecoverPoint alone, according to Fenselau.

Fenselau said discounted “frame” pricing is available for VPLEX. He added that many EMC customers have standardized on VPLEX and RecoverPoint. He said he views replication support via additional for-fee products in the same way as vendors that provide array-based replication or any data service for an additional charge.

The XtremIO array stands out in performance, with maximum read IOPS of 1,200,000 IOPS for a random workload at 4K block size, and the VNX-F was notable in throughput at 17.68 GBps for reads with a large-block, sequential workload.

The XtremIO product is among the minority of all-flash arrays to use custom-built RAID technology. Fenselau said the XtremIO Data Protection (XDP) was built from the ground up for flash and performs better than RAID 10 while minimizing flash writes with an 8% capacity overhead.

Fenselau said the XtremIO all-flash array has more than 200 customers. EMC declined to supply customer information for VNX-F other than to say that current product use cases include virtual servers, virtualized databases, virtual desktop infrastructure (VDI), test and development environments, online transaction processing (OLTP) applications, business intelligence and billing systems.

“EMC has never been a company to shy away from having some overlap in its portfolio,” said Mark Peters, a senior analyst at Milford, Mass.-based Enterprise Strategy Group Inc.

EMC said there is an average performance degradation of 20% during XtremIO controller failure. The vendor did not disclose the degradation percentage for the VNX-F.

**HP makes case that all-flash 3PAR StoreServ 7450 is designed for flash**

Hewlett-Packard Co.’s 3PAR StoreServ 7450 uses the same core architecture as the entire StoreServ family of products, but the company makes a case that the array was designed strictly for flash.
Priyadarshi Prasad, a senior product manager for HP 3PAR, said the 7450 has the same SAS interconnects and enclosure "plumbing" as the disk-based StoreServ products, but the product uses different CPU and memory to hit performance numbers of 900,000 random read IOPS and latency of less than 0.7 milliseconds with a quad-controller system.

Also helping to boost performance is the StoreServ 7450’s flash-optimized software layer that incorporates HP’s Adaptive technologies for read, write, I/O processing, cache and other functions, according to Prasad. The benefits of the flash-focused Adaptive technologies extend to the other StoreServ arrays, he added.

"Obviously HP is working very much towards its whole converged infrastructure," using the same operating platform and user interface across its 3PAR platforms, said Mark Peters, a senior analyst at Milford, Mass.-based Enterprise Strategy Group Inc.

"The interesting aspect is how they can potentially integrate with the rest of the 3PAR lineup," said Henry Baltazar, a senior analyst at Cambridge, Mass.-based Forrester Research Inc. "They are doing some interesting things with federation capabilities to move data across these systems. That could be interesting for people that are trying to optimize workloads and make sure that, when workloads are cold, they're not consuming a lot of flash."

HP’s peer-to-peer federation allows users to nondisruptively move applications from the StoreServ 7450 all-flash array to spinning media and vice versa, according to Prasad.

Prasad also highlighted the 3PAR StoreServ 7450’s advanced capabilities such as three-data-center replication, automatic failover of a front-end controller node port that experiences laser loss (helpful if a cable is mistakenly pulled out), and granular quality of service (QoS) that allows users to assign minimum thresholds for IOPS, bandwidth and latency.

According to Prasad, support for QoS varies among all-flash array vendors. He said HP and SolidFire support the feature in the "truest sense of the word," with the ability to prioritize one application over another in software, as opposed to basic hardware partitioning that can require administrative effort and knowledge. The HP-SolidFire approach allows users to "change it every second, every minute of the day just by switching a knob," he said.

HP touts the StoreServ 7450’s full-feature set as one of the product’s differentiators. Other capabilities include snapshots, inline zero-detect deduplication, asynchronous and synchronous replication and data-at-rest...
encryption. One of the few features the product does not offer is inline data compression.

The HP 3PAR StoreServ 7450 is the only major all-flash array to support all three major types of flash drives, from the most expensive and highest endurance single-level cell (SLC) to midgrade enterprise multilevel cell (eMLC) and the least expensive and lower endurance MLC drives.

HP recently changed the warranty on its MLC drives from three years to five years, but the warranty on the SLC and eMLC remains three years.

"Customers have not been asking for any other warranty" on the SLC and eMLC drives, explained Prasad. "It's only on the MLC drives that customers have some concerns on whether these types of drives will wear out after three years. We went back and told them, 'Look, if you are concerned, we'll offer you a five-year warranty on these MLC SSDs."

Equipped with 920 GB MLC SSDs, the four-node 3PAR StoreServ 7450 is among the higher capacity options available in all-flash arrays with a maximum of 220 TB of raw capacity. Most 7450 sales are four-node systems, according to Prasad.

HP 3PAR StoreServ 7450 customers include the research and development group at Nuance Communications, which makes voice recognition and language software, and ExactTarget, a Salesforce.com company specializing in marketing automation and campaign management.

HP said there could be performance degradation of up to 25% during controller failure in a two-node configuration and up to 15% in a four-node configuration. There would be no impact if CPUs and controllers operate at less than 75% utilization, according to HP. Data is accessible during controller failure.

Estimated street pricing for the StoreServ 7450 starts at $38,000 for a two-node configuration with 800 GB raw capacity. HP claims the array costs $6 per GB based on raw capacity (without factoring in deduplication) for a two-node system with 48 920 GB SSDs and $0.18 per IOPS for two nodes with 24 100 GB SSDs.

Hitachi boosts all-flash HUS virtual machine with custom drives, optimized software

Hitachi Data Systems Corp. (HDS) may not sell an array that was designed at the outset exclusively for flash, but the company has expended
considerable effort to optimize its Hitachi Unified Storage (HUS) VM for the ultra-fast storage media.

HDS took the unusual step of building its own flash module drives (FMDs) and made flash-specific modifications to its Basic Operating System (BOS) to help the HUS VM crack the million read IOPS threshold that growing numbers of all-flash array vendors claim.

The HUS VM is among a handful of all-flash arrays to use custom-built flash drives. Hitachi's FMDs can store 3.2 TB and boost the array's maximum flash storage capacity to 302 TB, making it one of the higher capacity all-flash arrays on the market.

FMD includes multilevel cell (MLC) NAND media, battery backup and a custom-designed embedded controller. The company claims that it has either received or applied for more than 60 patents for drive technologies such as custom error-correcting code (ECC) to monitor page degradation and avoid premature page rewrites.

With its initial release in September 2012, the first all-SSD-capable HUS VM achieved about 250,000 read IOPS. A June 2013 release boosted performance for a single system to about 500,000 read IOPS, and an October update doubled the number to 1 million, according to Bob Madaio, senior director of infrastructure platforms marketing at HDS.

Madaio said the performance bump to 500,000 IOPS stemmed from a system software upgrade branded as Hitachi Accelerated Flash (HAF), and the leap to 1 million IOPS resulted from further flash optimizations for which the HDS adds a "small upcharge" relative to the overall system cost. HDS field representatives can advise customers if they will be able to achieve 1 million IOPS based on how their applications work.

"We all market these high IOPS numbers, but what really matters is low and consistent response time," he said. "A million IOPS performance takes a very special high-intensity read environment to really make sense."

Madaio said some users may need performance at all costs, but most scenarios don't require a million IOPS. Likening all-flash arrays to rocket ships and disk-based arrays to cars, he said, "Yes, there are different models of rocket ships, and I guess one is faster than the other. But, everyone was using a car, so the difference is so massive that the rest of it becomes almost insignificant for many customers."

HDS doesn't sell specifically designated all-flash arrays, although all of the company's products can be purchased and deployed in flash-only mode.
Madaio said the company has shipped all-flash high-end Virtual Storage Platform (VSP) arrays, but the entry-level enterprise HUS VM is the most logical product for broad all-flash use cases.

"We can ship a 600 TB flash system today with VSP with no issue, but the use cases for 600 TB of flash in one system are relatively small," Madaio said. "We don't find most customers looking and thinking of big enterprise systems in all-flash perspective. That's the reason we didn't really go there."

But one potential advantage HDS has is integration across its product lines. For instance, users can replicate between HUS VM and VSP through native replication in both products.

Another differentiator is the Hitachi Universal Volume Manager external storage virtualization software, which enables a user to view and manage the storage environment as a single pool. The software can also assist with data migration between Hitachi arrays or between Hitachi arrays and third-party storage.

Customers have the option to upgrade HUS VM from an all-flash to a tiered hybrid system with a combination of hard-disk drives (HDDs), solid-state drives (SSDs) and Hitachi's FMDs. They can also upgrade to a unified system with the addition of Hitachi NAS for network file services.

HUS VM supports block, file and object storage. The product is also among the most feature-rich all-flash options with support for capabilities such as thin provisioning, inline deduplication, snapshots, clones, data encryption and replication. Hitachi recently added three-data-center replication through a fee-based Remote Replication Extended software package.

List pricing for an all-flash HUS VM starts at $304,570 for 64 GB of cache and four 1.6 TB flash modules. HDS claims a $20 per GB list price for an array with 96 1.6 TB flash modules and $0.73 per IOPS for an array with 16 1.6 TB flash modules.

HUS VM's all-flash customers include Infosys, a consulting, technology and outsourcing firm, and Matiq, a major food supplier in Norway.

IBM targets two market segments with FlashSystem platform

The FlashSystem 840 and V840 models that IBM unveiled this year include an all-flash array built for performance and another with enterprise storage features.
IBM recently joined the growing list of vendors claiming to surpass a million IOPs with its announcement that its new FlashSystem 840 doubled the performance and bandwidth of its predecessor, the FlashSystem 820, and dropped the read and write latency down to barely detectable microseconds.

How well IBM's purported 1.1 million IOPS stacks up against the millions of IOPS boasted by seven other all-flash array vendors will remain a mystery unless all vendors agree to use standard testing methodology. Until that day, suffice it to say, they're all faster than disk-based arrays.

We do know the FlashSystem 840 appliance lacks many of the storage-saving and management features increasingly showing up in performance-focused enterprise all-flash arrays. That's where IBM passes the baton to its FlashSystem V840 Enterprise Performance Solution, which bundles the FlashSystem 840 and IBM System Storage SAN Volume Controller, or SVC, software stack.

The FlashSystem V840 array supports such capabilities as thin provisioning, inline compression, snapshots, replication, data encryption and quality of service, whereas the 840 checks off only encryption and QoS from that list. Neither product supports inline deduplication.

"The reason we have two products is there's a segment of the market that needs extreme performance. They are truly doing application acceleration. And they don't want any bells and whistles getting in the way," said Kevin Powell, business line program manager for IBM FlashSystem.

Powell said the other half of the market uses flash as an alternative to hard-disk drives (HDDs). Those organizations want high performance and low latency but also need the enterprise features they're accustomed to using with their traditional disk-based storage systems.

Customers do have the option to completely configure IBM's disk-based systems, such as the DS8000 and Storwize arrays, with solid-state drives, But Powell said IBM doesn't count those products in the same camp as the arrays designed to take advantage of flash technology.

IBM got a running start in the purpose-built flash array market with its 2012 acquisition of Texas Memory Systems (TMS). By April, IBM was putting out its own branded versions of TMS products -- the FlashSystem 710, 720, 810 and 820 -- with such minor updates as an improved graphical user interface. This year's FlashSystem 840 and V840 mark "the first real IBM products," according to Powell.
"We built the flash modules, the controllers, the interface cards, everything, so that we minimize the latency," he said. "It's not just flash running through a disk system."

Few all-flash arrays use flash modules custom-built by the vendor. IBM sources its **NAND flash** directly from the manufacturers, and its proprietary **enterprise multilevel cell (eMLC)** modules are lengthier and about 50% wider than a standard 2.5-inch disk, according to Powell. They also buck the industry trend toward cheaper **MLC flash**.

"We purposely chose eMLC technology over lesser-grade MLC because the IBM customer set is more of an enterprise customer set," Powell said. "We are focused on reliability."

Powell said IBM has worked on write protection, **wear leveling** and other features to improve the endurance of its flash modules. He noted that IBM designed the system targeting "five-plus years of endurance" for a typical to heavy workload, although extreme use cases could potentially wear out the flash more quickly.

"All flash technology, from all vendors, will wear out over time based on the number of erase cycles," Powell said, adding that IBM does not state or publish a guaranteed number of reads or writes.

IBM places caveats on its flash warranty. The overwhelming majority of all-flash-array vendors say they would replace flash drives even if they wear out before the product warranty expires. Powell said IBM replaces failed products due to bad components or other malfunctions, but flash read/write wear-out is not covered under the warranty.

"Our service and support organization is here, and we can construct a contract as necessary for any customer requirements we need to," Powell said. "IBM has a long history of supporting customers for very long durations with our mainframe and our enterprise storage systems, so we know how to take care of enterprise customers."

IBM continues to sell and support FlashSystem 710, 720, 810 and 820. But Powell said those products will "phase out before the end of this year."

List prices begin at $50,000 for the FlashSystem 840 and $100,000 for the V840. IBM claims a price per gigabyte (without compression) of $14.08 for the 840 and $14.70 for the V840, and a per-IOPS price of 33 cents for the 840 and 60 cents for the V840.
IBM said there is 50% performance degradation during controller failure. Data is accessible during controller failure.

**Kaminario beeps up all-flash array with enterprise storage features**

Early versions of Kaminario Inc.'s K2 all-flash arrays stood out with their performance claims but lacked many of the capacity-saving and data management features typically found in traditional enterprise storage arrays.

That changed with Kaminario K2 version 5 (v5) released in May.

Ritu Jyoti, chief product officer at the Newton, Massachusetts-headquartered company, said Kaminario K2 version 5 (v5) enhanced the product's enterprise readiness with the addition of features such as thin provisioning, inline deduplication and inline compression in May and replication and data encryption planned for later in the year.

"V5 is a huge, huge leap forward for us," Jyoti said.

She said interest was building early this year for the new v5 product, and some customers purchased K2 v5 even before its release. Kaminario-supplied numbers claim the price per GB for the least expensive configuration dropped from $8.69 for v4 to $2.55 for v5 with deduplication and compression factored in. Also with v5, the Kaminario K2 product shifted from more expensive 1.6 TB enterprise multilevel cell (eMLC) SSDs to less expensive 800 GB MLC drives, according to the company.

Kaminario's K2 offers a seven-year endurance guarantee on its flash drives - a guarantee that is written into the warranty contract, according to Jyoti.

Last year, Kaminario announced a performance consistency guarantee promising no more than 25% performance degradation during a system failure. The company claimed that if the degradation failed to meet the guarantee, it would give customers the new hardware necessary to allow the product to do so.

Kaminario's all-flash arrays stand out for their performance claims, with maximum read and write IOPS in excess of 1 million and double-digit read throughput. Kaminario-supplied figures for K2 v4 (maximum read IOPS of 4 million and maximum read throughput of 64 GB/s) were actually greater than the numbers provided for K2 v5 (maximum read IOPS of 2 million and maximum read throughput of 25.6 GB/s) because the company qualified v4 for higher performance than it has done so far for v5, Jyoti said.
Jyoti said Kaminario’s K2 can scale capacity and performance to a greater degree than it does now, and the company can undertake further qualifications if it chooses. But, she said Kaminario has not seen customer demand for higher capacity and performance.

"We [have] a scale-out architecture, so we can add more nodes and scale both performance and capacity," Jyoti wrote in an email. "With v5, we are adding scale-up as well -- meaning we can scale capacity without adding performance. In other words, the customers can dynamically scale capacity and performance, all fully automated. In theory, these numbers can scale infinitely, but we do thorough qualification and state the numbers that we have qualified."

Kaminario’s scale-out K2 v4 had a raw maximum capacity of 240 TB and a usable maximum of 120 TB. The new K2 v5 boosted the maximum raw capacity to 307.2 TB and the maximum usable to 236.8 TB, with the option to scale further, according to the company.

NetApp offers all-flash EF-Series, FAS systems while prepping new FlashRay

NetApp Inc. offers all-flash array options with its EF-Series and FAS storage systems, but its upcoming FlashRay product could make it a more serious player in the solid-state marketplace.

The Sunnyvale, California-based company started from scratch on the scale-out FlashRay design to tailor the product for ultra-fast flash storage. NetApp said the multi-protocol array will use a new internally developed operating system and include enterprise storage efficiency features such as inline deduplication and compression. FlashRay is currently in beta testing and scheduled for general release this year.

"NetApp is still flushing out their flash story," said Henry Baltazar, a senior analyst at Cambridge, Massachusetts-based Forrester Inc. "A lot is riding on whatever happens with FlashRay. I don't think their impressive stuff happens until FlashRay comes out."

In the meantime, NetApp has customers deploying its EF-Series and its FAS arrays in all-SSD configurations, according to Rip Wilson, senior product marketing manager for all-flash arrays at NetApp. The EF-Series is based on technology that NetApp acquired from LSI in 2011.

"We believe there is no one-size-fits-all approach to flash," said Wilson via an email. "Enterprises need different features -- performance, efficiency,
All-flash arrays: comprehensive chart
Page 2

IBM FlashSystem 840 focuses on performance, V840 is more feature-rich
Page 11

Pure Storage bolsters FlashArray line with more capacity choices
Page 18

Experts weigh In On all-flash arrays
Page 28

When to consider the purchase of an all-flash array
Page 31

reliability, flexibility and choice -- to maximize the true value of flash. Every workload is different.”

The EF540 ships with enterprise-grade single-level cell (SLC) solid-state drives (SSDs) and the EF550 features enterprise multilevel cell (eMLC) flash drives. SLC drives carry the highest endurance, but eMLC drives are less expensive.

Many competitors have been moving to even cheaper multilevel cell (MLC) flash drives. FlashRay is expected to use MLC flash.

NetApp’s EF550 stands out with its double-digit throughput and microsecond latency figures as well as its networking support for 16 Gbps Fibre Channel, 10 Gb Ethernet, 40 Gb InfiniBand and 6 Gbps SAS. The product also offers plentiful RAID options, including NetApp’s proprietary Dynamic Disk Pools (DDP).

Wilson said DDP evenly distributes data, dual parity and spare capacity across the array’s entire pool of drives to simplify setup and maximize utilization because there are no RAID sets to configure and no idle hot spare drives. NetApp also claims DDP is able to maintain consistent performance by eliminating hot spots, minimizing the performance impact of a drive failure and returning the system to optimal condition more quickly than traditional RAID.

From a storage feature standpoint, the EF-Series offers capabilities such as thin provisioning, snapshots and replication, and unlike many vendors, NetApp doesn’t charge an extra fee for them. But, the product lacks support for inline deduplication, compression, and quality of service.

The EF550 supports SSDs of up to 1.6 TB while the EF540 supports 800 GB. The EF550 scales to 192 TB.

List price for the EF540 starts at $175,000 for 9.6 TB raw capacity, the EF550 starts at $85,000 for 2.4 TB. NetApp claims its price per GB is $18.25 and price per IOPS is $0.63 for both systems, based on 9.6 TB with 12 800 GB SSDs.

There is performance degradation of 10 percent to 50 percent during controller failure for the EF550. Data is accessible during controller failure.
Nimbus Gemini all-flash array line set to add higher density option

One of the early movers in the all-flash array market, Nimbus Data Systems Inc. sells multiprotocol systems that promise high performance and a broad range of enterprise storage features. But the maximum raw capacity was limited to 48 TB in the Gemini F400 and F600 lines released last year.

The South San Francisco-based company’s Gemini X-series is poised to expand the maximum raw capacity to 96 TB in a 2U rack that will use its latest custom-built 4 TB flash drives. The X-series array can scale out to close to 1 PB in a clustered configuration, and with data reduction, it holds the potential to store several petabytes, depending on the data set.

The new Nimbus Gemini X-series uses Flash Directors and Flash Nodes and combines a nonblocking 320 Gbps RDMA switch, up to 40 configurable ports for connectivity and centralized cluster management software. The Flash Director virtualizes the storage capacity into a single namespace and load balances I/O across the nodes.

Nimbus CEO Thomas Isakovich said customers have discussed the prospects of large all-flash deployments for their entire tier-one environments. They can start evaluating the Gemini X-series this month. Nimbus expects to make the product generally available in the summer.

The current Nimbus Gemini F400 and F600 systems are among few all-flash arrays to support block, file and object storage. They also use custom-built multilevel cell (MLC) flash drives, and their 2 TB drive capacity is among the highest of arrays on the market. Also, the Nimbus drives are the only SSDs to use 12 Gbps SAS as opposed to 6 Gbps SAS.

For storage networking, the Nimbus F600 is the only all-flash array to support 40 Gbps Ethernet and 56 Gbps InfiniBand. It also supports 16 Gbps Fibre Channel.

Nimbus arrays support thin provisioning, snapshots, clones, replication, data encryption and quality of service. Nimbus provides all software in the base price of the system.

Performance figures are difficult to compare because all vendors use different configurations and conditions, but Nimbus claims a maximum read IOPS of 2 million with a 4K block size, a maximum read throughput of 12 GBps with a 64K block and read/write latency of microseconds are among the most impressive of flash arrays.
Nimbus guarantees its all-flash arrays, including the SSDs, for up to 10 years so long as the customer has an active support contract. The company claims its flash lifecycle management technology increases the native endurance of MLC flash by more than 50 times. Isakovich has said that customers would not be able to wear out the arrays even if they did nothing but writes during the 10 years.

The company began selling all-flash arrays in 2010 and claims to have collected more than 500 customers in industries such as finance, health care, media, government, education, technology and cloud services. Customers have included eBay, Cloudmark, Mitsubishi Power Systems, Loxogon, Tibco Software and Bangor, Maine-based St. Joseph Healthcare. Primary use cases have been database applications, virtualization (both desktops and servers) and cloud computing.

Nimbus’ warranty is for up to 10 years, including all hardware including flash drives and a full software subscription.

Pricing for the Gemini F series begins at $40,000 for a one-controller and 3 TB raw capacity configuration, $55,000 for two controllers and 3 TB, and $280,000 for two controllers and 48 TB. Nimbus claims its price per GB is $4.88 based on redundancy, deduplication and workload type with an $0.02 per IOPS price for a 3 TB entry level configuration.

Isakovich said Nimbus’ asymmetric active-active design eliminates performance degradation during controller failure.

Pure Storage bolsters FlashArray line with more capacity choices

Pure Storage Inc.’s FlashArray line reached its fourth generation this year, adding replication and new entry-level and higher capacity models as it girded for a battle in an increasingly crowded all-flash array market space.

Flash array startup specialists such as Pure Storage now find themselves in competition with long-standing storage vendors.

The new FA-450 scales up to 70 TB of raw capacity and supports 16 Gbps Fibre Channel to give large enterprises an additional option, and the FA-405 provides up to 11 TB of raw capacity with smaller and distributed businesses in mind. The new models have the same architecture as the existing FA-420, which offers up to 35 TB of raw capacity. FA-420 customers can upgrade non-disruptively to the FA-450 if they want to increase capacity or performance.
Highlights in new product

"Pure Storage runs on an x86 box. They use commodity SSDs. Their innovation is in the software," said Henry Baltazar, a senior analyst at Cambridge, Massachusetts-based Forrester Inc. "They've got a ton of funding [$470 million]. There's a lot of recognition around them. They've got a lot of feet on the street."

Pure Storage claims to deliver all-flash storage at a cost that is competitive with, if not lower than, systems with hard-disk drives. Its FlashArray FA-450's purported price per GB of less than $3 with data reduction factored in placed it among the lowest among all-flash arrays.

Pure Storage is among the few all-flash vendors to support both inline deduplication and compression. Another area where Pure Storage's FlashArray stands out from the pack includes its use of proprietary RAID technology.

Lowlights

On the flip side, the product lacks support for scale-out architecture and quality of service. Most all-flash arrays support those capabilities.

The company claims FlashArray can store in excess of 120 TB through data reduction. Deduplication and compression, however, can vary by data set. Pure Storage's website includes a ticker that it claims shows the average data reduction rate, factoring in deduplication and compression, across all deployments.

What can you do with it?

Typical use cases for Pure Storage's FlashArray include heavily virtualized server environments, database acceleration and virtual desktop infrastructure (VDI) installations. A list of case studies on the company's website showed data reduction ratios ranging from 5 to 1 to 17 to 1 for VDI deployments and from 3.5 to 1 to 10 to 1 for databases and/or VMware deployments.

More than a dozen of the case studies on Pure Storage's website involved VDI deployments. Organizations using FlashArray for VDI included the San Jose Sharks' NHL hockey team, the Waxahachie (Texas) independent school district and Riverview Hospital in Noblesville, Indiana.

FA-405 pricing starts at less than $100,000 for 10 TB of usable capacity.
There is no performance degradation during controller failure, and data is accessible. "We cap the performance in what a single system can provide, just for this purpose -- explicitly to provide the same performance during upgrades and/or failures," Jim Sangster, Pure's director of product and solutions marketing, wrote in an email.

"This is intentionally how FlashArray is architected and delivered through the two systems with active-active front end and active-passive back end and the notion that when one fails, it is the other that is running full tilt. When both are in normal operation, the I/O ports from both are used, but we are never using performance of both."

Skyera turns heads on price with skyHawk, readies larger skyEagle

On paper, the all-flash arrays from Skyera Inc. have turned plenty of heads for their pricing.

San Jose, California-based Skyera claims its first product, skyHawk, has a price per gigabyte of $2.99 based on a configuration of 44 TB usable capacity. With 3-to-1 data reduction factored in, the price per gigabyte drops to 99 cents, according to Raymond Pang, director of product marketing at Skyera.

"They're incredibly impressive on the pricing front," said Mark Peters, a senior analyst at Milford, Mass.-based Enterprise Strategy Group Inc. "Many of the all-flash vendors will tell you that they can compete with spinning drive-[based arrays] on price. Usually that's as a result of compression and deduplication. Skyera and Nimbus will say, 'Now we can get down to comparison prices even without using those other tools.'"

But while Skyera first announced skyHawk in August 2012, it didn't ship in volume until April 2014. Pang said skyHawk has shipped in limited availability to an unspecified number of large enterprises and government agencies since the third quarter of 2013. The follow-on skyEagle is not expected until the second half of 2014, according to Pang.

While skyHawk targets smaller-scale, lighter-weight, lower-power use cases and enterprise test and development environments, the upcoming higher-density skyEagle will aim to expand into enterprise mission-critical applications and the workloads of cloud service providers, Pang said.

If skyEagle lives up to the company's claims, the product's capacity will max out at 650 TB raw (500 TB usable). The post-deduplication capacity can vary based on the data set, but whatever the case, the numbers would dwarf
skyHawk's capacity limits of 44 TB usable and an estimated 88 TB post-data reduction.

"Skyera is promising to deliver a box that's 1U and almost a petabyte of storage, so that's significant. They have custom-built hardware, but they're also trying to design their own software," said George Crump, president and founder of Storage Switzerland LLC. He added that Skyera is "somebody to keep an eye on, but that's about it right now."

The skyHawk arrays use custom-built storage blades rather than traditional solid-state drives, proprietary flash-optimized RAID-SE technology, multi-level cell (MLC) flash and inline compression, and support file-based storage, in addition to block.

Henry Baltazar, a senior analyst at Cambridge, Massachusetts-based Forrester Inc., said Skyera's intellectual value and expertise lie in its hardware controllers and getting the most endurance out of consumer-grade MLC flash. "That's how they're able to get the really aggressive price points," he said.

The performance numbers that Skyera is already claiming for the upcoming skyEagle would place the product among the upper tier of all-flash arrays, if they hold up. The company listed the maximum read IOPS at 5 million, and the maximum write IOPS at 1.5 million.

Skyera's plans for the future skyEagle release also call for the addition of scale-out capabilities to the scale-up architecture, inline deduplication and support for Fibre Channel connectivity.

Skyera is one of many vendors to charge fees for select features in its all-flash array. The company requires a for-fee license for remote replication, whether synchronous or asynchronous, and charges for inline compression and data encryption. Inline deduplication will also carry a fee in the future.

Skyera claims less than 50% performance degradation if one node is lost in a two-controller configuration for skyHawk and skyEagle.

**SolidFire seeks to grow customer base beyond cloud to enterprise IT**

With upcoming updates to its all-flash arrays, SolidFire Inc. will try to broaden its reach beyond the cloud service providers that it initially targeted and find a more general enterprise IT audience.
New features that the Boulder, Colorado-based startup plans to add this year include 16 Gbps Fibre Channel connectivity; real-time replication; support for a single cluster that combines storage nodes of differing capacities, performance and protocols; and native snapshot-based backup and restore with any object store that uses an Amazon S3 or OpenStack Swift APIs.

So far, SolidFire's all-flash arrays have largely held their appeal for cloud providers that need to scale storage and performance. Customers have a choice of three different all-flash storage nodes, each in a 1U form factor with 10 multi-level cell (MLC) flash drives.

The smallest SF3010 node supplies 3 TB of raw capacity, 12 TB of "effective" capacity with data reduction factored in, and 50,000 IOPS for an 80/20 mixed read/write workload. The high-end SF9010 node offers 9.6 TB of raw capacity, 34 TB of "effective" capacity, and 75,000 IOPS for an 80/20 mixed read/write workload, according to Jay Prassl, vice president of marketing at SolidFire.

But, Prassl said, customers can mix and match any of SolidFire's products within the same cluster. Storage footprints commonly start at a five-node minimum configuration and have the potential to expand to as many as 100 nodes, he said.

That means a cloud provider with extreme needs could potentially buy a 100-node SF9010 all-flash system with a raw capacity of 960 TB, a post-data-reduction capacity as high as 3.4 PB, depending on the data type, and IOPS well into the millions, according to Prassl's numbers.

Beyond the scale-out capabilities, another SolidFire feature that has drawn great interest from cloud providers is quality of service (QoS). With SolidFire's fine-grained QoS settings, providers can assign minimum, maximum and burst IOPS on a per-volume basis.

"They're one of the first to really push QoS," said Henry Baltazar, a senior analyst at Cambridge, Massachusetts-based Forrester Research Inc.

Baltazar noted that SolidFire's innovation lies in software rather than hardware engineering. The company uses commodity x86 servers and standard MLC SSDs fortified for enterprise use, he said.

SolidFire's all-flash arrays stand out among competing all-flash arrays in price and performance, but those two criteria can be difficult to compare when vendors use varying configurations and test conditions.
SolidFire claimed its price per gigabyte is less than $3 with deduplication, compression and thin provisioning taken into consideration, placing it among the lowest on the market. On the other hand, the street price of "less than $200,000" for the least expensive configuration (the five-node SF3010) isn't quite as low -- several vendors list a price less than $100,000 as their starting price for low-capacity models. SolidFire also claims a $1 per IOPS price for all its systems.

But SolidFire is in the minority of all-flash array vendors with its support for both inline deduplication and inline compression, so its arrays have the potential to store more data than the raw capacity figures indicate, contingent on the data type.

Another area where SolidFire stands out is its custom data protection technology. The company's proprietary Helix distributed replication algorithm spreads at least two redundant copies of data across all drives within the system. The company claims the "RAID-less" approach permits the system to sustain multiple, concurrent failures and isolate them to avoid a performance impact.

SolidFire's customers includes cloud and managed hosting providers such as ViaWest, Calligo, Colt, Datapipe, Clearview, CloudSigma, Internap and ServInt. Other large customers have included eBay, PayPal and SunGard.

Because of its controller-less clustered architecture, SolidFire claims the loss of nodes has no performance impact on volumes.

"SolidFire isolates failures, minimizing the impact on systemwide performance," Prassl said. "Within a SolidFire system, even though there is a decrease in the total available performance of the system, there is no impact to the individual performance level set for each individual volume. This capability is enabled through SolidFire's guaranteed quality-of-service capabilities, and applies to every volume within the system."

**IT Tegile breaks hybrid focus with release of new T3800 all-flash array**

Tegile Systems Inc.'s primary focus is selling hybrid arrays equipped with hard-disk and solid-state drives, but the company pushed into new territory this week with the release of its T3800 all-flash model.

The T3800 targets high-performance workloads such as big-data analysis and online transaction processing (OLTP) applications and affords a maximum raw capacity of 332 TB, or a possible 1.3 PB TB usable with data reduction factored in, according to Tegile.
"As we move into these types of applications, latency predictability becomes paramount," said Rob Commins, vice president of marketing at Tegile Systems. He said the company added the T3800 in response to customer demand for a more latency-focused product.

Prior to the T3800, Tegile's only all-flash option was the HA2800. Most customers purchased the HA2800 with hard disk drives (HDDs) or ended up adding disks later, according to Commins.

Commins said the HA 2800 all-flash array attracted about 75 of Tegile's 500 customers. They could "pin" a volume in flash with a HA2800 array used in hybrid mode, he said.

In all-flash mode, Tegile's HA2800 has a raw capacity of 4.4 TB. With deduplication, Tegile claims that flash capacity jumps to 22 TB. But the real attraction for many Tegile customers is the ability to add hard-disk drives to expand the HA2800's capacity to 148.4 TB, or 742 TB after deduplication.

"We have found that most customers looking at all-flash solutions also have capacity-based applications, or that their performance-hungry apps only require that 3% to 10% of their data be in flash at a given time," Commins said. He cited the most common use cases for hybrid arrays as database optimization and large virtual server and desktop environments.

Features that make Tegile's T3800 and HA2800 stand out from other all-flash arrays include their support of block and file storage and both inline data deduplication and inline compression. On the product roadmap for 2014 is the addition of scale-out capabilities to augment the current product's scale-up architecture, Commins said.

Performance statistics are tough to compare because all-flash array vendors test their products using differing configurations and workloads, but the maximum read IOPS of 350,000 for the T3800 and the 330,000 for the HA2800 (at 4K block size with a random workload) were not among the highest. The read latency for each array is 1 millisecond.

The product's exclusive use of eMLC flash drives -- which offer greater endurance than less-expensive MLC drives -- is not the norm and perhaps factors into the product's higher price compared to some all-flash arrays.

"Using eMLC drives, we see very few problems or returns," Commins said. "Do they cost us more in our [bill of materials] cost? Yes. Do they offer superior customer experience? Definitely."
The list price for the least expensive configuration of Tegile's T3800 is $550,000 (with 44 TB raw capacity), and the estimated street price per GB is $1.10 with data reduction factored in. The list price for Tegile's HA2800 is $235,000 for 4.4 TB raw capacity, and the estimated street price per GB is $8 with data reduction factored in. Unlike many all-flash array vendors, Tegile Systems does not charge an extra fee for its storage-savings and data management features.

The T3800 and HA2800 are full-featured storage systems that support a wide range of efficiency and management technologies such as thin provisioning, snapshots, replication and VM-aware metrics. Among their few areas of deficiency in comparison to other arrays is a lack of support for data encryption and a driver for the open source OpenStack cloud technology platform.

Tegile said there is performance degradation up to 50% during a controller failure. Data can be accessed during a controller failure.

**Violin extends focus beyond speed, plans enterprise storage features**

One of the earliest vendors in the all-flash array market, Violin Memory Inc., made performance and low latency its top concerns and concentrated on its proprietary memory module, memory fabric architecture and flash-optimized RAID technology.

The Santa Clara, California-based company has never built storage systems with spinning disks. Its laser focus on solid-state storage has produced technologies to manage the array instead of managing individual solid-state drives (SSDs) and multiple layers of controllers, according to Erik Ottem, the director of product marketing at Violin Memory.

"This provides a more efficient approach, which leads to higher performance, consistently low latency, at a lower cost," Ottem said via email.

VRAID, custom flash modules

Violin is among the minority of flash vendors with proprietary RAID technology, known as vRAID, and custom-built flash modules. The Violin Intelligence Memory Modules (VIMMs) can store up to 1 TB, and they connect to the memory fabric via PCI Express (PCIe).

According to the Violin website, Violin 6000 Series array models range in capacity from 6.5 TB raw (3.5 TB usable) to 70 TB (44 TB usable) in
performance from 200,000 IOPS to 1 million IOPS and in bandwidth from 1.5 GBps to 4 GBps. Latency is under 500 microseconds or 250 microseconds, depending on the model. Some products feature higher-endurance single-level cell (SLC) flash, and others have lower-cost multi-level cell (MLC) flash.

The Violin 6264, the Violin 6232, and the Violin 6224 have raw capacities of 70 TB, 35 TB and 26 TB, respectively, and they all use MLC VIMMs. The Violin 6000 Series supports high-performance InfiniBand for storage networking.

But, the Violin 6224, 6232 and 6264 arrays lack some features and capabilities that many all-flash arrays have. Unlike most flash arrays, the Violin product line does not support nondisruptive hardware upgrades and zero block reclaim.

Also, Violin does not support inline deduplication, inline compression, quality of service and replication. However, inline deduplication, inline compression, zero block reclaim, replication, three data center replication and VMware vSphere Metro Storage Cluster (vMSC) certification are in the product plans for this year, according to Ottem.

"Because they were early into the market, they were aimed at the all-out speed category. They didn't have any feature and functionality, nor indeed did they think it was necessary," said Mark Peters, a senior analyst at Milford, Mass.-based Enterprise Strategy Group Inc. "As the market is maturing, that need to have functionality is increasing."

Functionality plus storage management

Violin charges extra for several storage management features. The company’s vMOS software package, which incorporates features from Symantec Corp.’s Storage Foundation, includes thin provisioning, snapshots and clones, according to Ottem. Violin also charges a fee for encryption of data at rest, he said.

Violin recently launched a Windows Flash Array (WFA) that it jointly developed with Microsoft. The scale-out, file-based WFA uses Violin’s 6264 array with the Windows Server 2012 R2 operating system embedded. The new product is designed to improve the performance of enterprise and cloud workloads that use Microsoft products such as SQL Server databases, SharePoint collaboration software and Hyper-V virtualization applications. Users also gain access to Windows Storage Server 2012 R2 storage features, including thin provisioning, data deduplication, compression, snapshots, mirroring and data-in-flight encryption.
Violin reports less than 10% performance degradation if a controller fails.

The estimated street price for the Violin 6000 Series is around $100,000 for 13 TB of raw (7 TB usable) capacity in the 6212. Violin claims a street price of $4 per GB to $4.50 per GB for its 6264.

Chief use cases for Violin’s all-flash arrays include customer relationship management, enterprise resource planning and transaction-oriented applications, databases, online transaction processing, real-time analytics and virtual environments. Violin customers profiled in case studies include Collier County (Florida) Public Schools, Pella Corp., Gilead Sciences Inc. and the University of California at Davis.

EXPERTS WEIGH IN ON ALL-FLASH ARRAYS

We asked industry analysts and consultants to offer advice and delve into some of the fine points about all-flash arrays to provide a deeper level of insight into the product market.

All-flash arrays: Performance vs. function

Nearly every storage vendor now offers all-flash storage arrays, and IT professionals are beginning to recognize the need for these high-performance storage systems. But how does an IT pro decide which of the many all-flash arrays are best suited for their organization and performance demands?

Performance vs. function

As the all-flash storage array market begins to mature, there are two categories of arrays emerging.

The first are all-flash arrays that were designed from the ground up to be all-flash arrays. They typically have optimized hardware designs that focus on extracting the maximum possible performance from the flash within the array. The vendors in this space are almost all emerging technology companies or startups. In most cases, their focus on hardware and performance is at the expense of storage software services. These are the features that many storage administrators now count on to do their jobs, providing capabilities like snapshots, replication and cloning.

These arrays are known for generating millions of IOPS per system. However, there really is no established method for how those high IOPS
numbers are obtained. They can be generated from a single workload or multiple workloads accessing the system at the same time.

The other category is made up of all-flash arrays that are more feature-oriented. These are typically systems from established vendors, as well as a few startups, that choose to focus on the software functionality (providing a feature-rich’ experience), often at the expense of maximum performance. Typically, these systems either use legacy hardware from the established vendor and retrofit their old arrays with solid-state drives (SSD) or, in the case of a startup, use off-the-shelf hardware to keep costs down.

These systems can often generate 200-400k IOPS per system. Some scale-out, software-rich systems will claim an aggregate performance of millions of IOPS as well but, as mentioned above, the devil is in the details. They typically have a performance limit per volume or per node within the scale-out cluster. This means they can scale to millions of IOPS like the performance-focused systems described above, but it takes many nodes to get there and to see that extreme performance requires multiple workloads all running concurrently. A scale-out system cannot deliver millions of IOPS to a single workload or thread.

Which is best?

We are often asked which method is best. The answer, as usual, depends on the needs of the data center and the specific applications that are running. Most data centers, while performance-constrained, are not constrained to the point that they will typically exceed the baseline performance of a feature-rich all-flash array. Also, most organizations will take great comfort in the availability of the feature sets they have become accustomed to from legacy hard disk arrays.

There are environments with a need for more than a half million IOPS, but it's how those IOPS are needed that will help determine the best system for a particular data center. If the need for performance is distributed across more than a few workloads, the all-flash systems that can provide scale-out linear performance growth are ideal.

If the environment has a single workload that needs more than half a million IOPS, then the performance-focused systems are needed. As stated above, these systems can provide millions of IOPS to a single workload.

Middle ground?

Is there room in the middle? Does a storage system exist that can meet the needs of a performance-demanding workload, yet still provide the feature-
rich environment that more traditional applications require? There are several vendors that provide this class of solution. This type of system must be designed first as a performance-focused system, then have software added to it. While the addition of that software will add some latency, it will not impact most applications. These systems typically have performance to spare.

This software can be added in several ways. Some vendors provide an appliance that the performance-focused system can be connected into, allowing it to take advantage of all the features that the appliance can provide. This storage virtualization approach also allows the all-flash array to be somewhat integrated from a software services perspective.

Other vendors have the ability to load storage software onto a co-processor within the flash array itself. This provides a tighter integration experience and saves the cost of an external appliance.

Finally, all of these hardware-focused systems could work with any of the software-defined storage solutions that are on the market today, including those converged solutions that run within the hypervisor architecture. The key, though, is to make sure that that software-defined solution can support external, shared storage (not all do).

While combining a hardware-focused solution with either an appliance or hypervisor that delivers the storage services, it's key to remember there remains one big challenge. That hardware-focused flash solution must be delivered at a price point (including software) that is in the same range as the feature-rich solutions described above. In most cases, the feature-rich solutions are still the most cost-effective, and again, 400k IOPS is more than enough for most data centers.

All-flash arrays are becoming mainstream. Many vendors in the space claim price parity with "performance-focused" hard drive arrays. These would be arrays from name-brand vendors that are using 15K RPM drives. This claim is generally true, so any data center looking to buy a performance-focused disk array should be seriously considering an all-flash array.

The choice within the all-flash segment is largely dependent on what the needs of the data center are. For most data centers, the feature-rich solutions will be all they need. But it may be worth the investigative step to confirm that and to then determine if they need a scale-up or scale-out system.
When to consider the purchase of an all-flash array

Under what circumstances should an IT organization consider the purchase of an all-flash storage array?

This is a little glib, but I would always buy an all-flash array for everything if I had more money than God, and therefore I could always get the latest, greatest thing because -- well, why not? The only reason we put data anywhere outside of memory is because we can't afford to put it all in memory. Or, maybe if you are in the sort of business where you want to impress your customers and clients that you have your entire data center in one rack, and it goes faster than the speed of light, then absolutely do that.

For more normal people and the vast majority of users, I don't think they should buy an all-flash array because it's an all-flash array. In other words, don't get attracted to new, shiny things just because they're new, shiny things. The first and foremost thing, as with the purchase of anything in life -- but particularly in IT -- is that you need to start with what you are trying to do and determine whether the devices that you're choosing between do at least what you need.

Next are a couple of things to consider especially. One is functionality. Clearly if you are expecting fairly rich functionality with snapshots and thin provisioning and yadda yadda, then you need to ensure that that is within the all-flash device or can be used with the device that you're looking at. The other thing, from the shiny object side, is that it's all too easy to get carried away with chasing specifications and end up buying more than you need.

You certainly need to be looking at the capacity, scalability and the ease of use of the system. There's no point buying something that can't grow with you. And there's no point buying something that is so difficult to use that it actually wastes more manpower than you gain somewhere else.

And perhaps most important of all is the financial side. I hold to the assertion that every purchase in storage is a financial purchase. How are you going to use this device? Are you going to use it for persistent data with certain "point" applications, or maybe in some situations -- where you can both afford and need to do it -- put your entire infrastructure on this flash unit? Or are you going to use it as a caching or tiering device? That will obviously impact both the amount you need to buy, and more importantly, it will impact the financial implications.

The judicious use of some amount of flash can preclude a lot of wasted spinning-disk capacity that otherwise has traditionally been bought purely to
get the IOPS and not the terabytes. The overall economics of some amount of flash can actually be very positive on your overall cost per gigabyte.

I don't think there are particular applications that are silver bullets for all-flash arrays. Clearly the sorts of features that have higher I/O requirements and high read requirements, such as databases and virtual desktop infrastructure, are good examples, but they are by no means the only ones.

Making sense of all-flash array vendors' wild performance claims

The performance numbers of all-flash arrays are soaring into the stratosphere, but it's often difficult to interpret vendors' claims of millions of IOPS, tens of megabytes per second and hundreds of barely negligible microseconds. Not only do the manufacturers use differing conditions to test flash performance, but workloads also rarely, if ever, mimic the real-world environments of their customers. As a result, it can be difficult to compare array performance accurately.

Dennis Martin, president and founder of Demartek LLC, an analyst organization which operates its own on-site test lab in Golden, Colo., has done extensive flash performance testing. In this podcast interview with SearchSolidStateStorage, Martin explains the importance of I/O per second (IOPS), throughput and latency for varying types of workloads, the significance of block size and random or sequential workloads and the chances that a million IOPS or microsecond latency will matter to the average enterprise data center.

Download this podcast

Of IOPS, throughput and latency, which figure do you think is the most important one for an enterprise end user trying to sort out the performance claims of all-flash array vendors?

Dennis Martin: We look at all three of those -- IOPS, throughput, sometimes called megabytes per second, and latency, sometimes called response time. All three of them are equally important, just generally speaking, because they all give you a different dimension of the flash performance. Some applications tend to be more sensitive to, or tend to drive, one of those metrics, either IOPS or throughput or latency, more than others. And some applications might be interested in two out of three. For example, they might be interested in IOPS and latency, or they might be interested in throughput and latency. Generally you don't see a single application that wants to get both high IOPS and high throughput. And then latency is a second figure that
is frequently needed, typically more so for IOPS-oriented workloads than throughput.

Can you give me examples of real-world applications that would need high IOPS versus high throughput versus low latency?

Martin: A high IOPS workload would be any kind of transaction workload, like a transactional database where you have lots of customers buying things and accessing the database at the same time. You might also think of it as not just one application but multiple applications. Before, you might not have put too many applications on one storage system because of the contention for the hard drives. In an all-flash system, some of those limitations are lifted, so you can put multiple applications on there. You might have a transaction database application that's very critical. You might put an email system on there. You might put some sort of a collection of virtual servers, or you might even do virtual desktops. There are all kinds of different applications that tend to be more IOPS-oriented.

If you're looking for throughput, then those are typically video streaming or backups or data warehousing or something where you're scanning over or looking at lots of data at once.

And then for latencies, sometimes it's a combination of the ones I just mentioned. You'll see some of the IOPS workloads also are interested in very low latencies. Sometimes they want to push latencies a little bit more because they're just looking for the fastest turnaround.

The performance claims of the major vendors of all-flash arrays are getting higher and higher, and cracking the million-read IOPS threshold is becoming commonplace. How important is a million IOPS for the average enterprise data center?

Martin: Generally, you're talking about a transaction-oriented workload where every individual I/O counts as far as how fast it happens. So, a million read IOPS is becoming commonplace. That's an important number because that just tells you where the high end is of this storage system. It also might tell you that you can run multiple applications against this same storage array because now it has the headroom to handle it.

Vendors often use a block size of 4K to get their maximum IOPS, and they use larger block sizes of 64, 128 or even 256K to get their throughput figures. Can you explain how the block size affects the performance figures?

Martin: One of the reasons you see 4K for all-flash arrays is because the minimum page size on flash devices is typically 4K. So, that's the smallest
you would go because that's the smallest the actual flash media itself would accept as an I/O. Some of them are starting to get a little larger now, like 8K, but 4K is a good place to start.

If you're looking at IOPS numbers, the smaller block sizes are going to have larger IOPS numbers. So, a 4K performance is going to be higher than an 8K block size performance and so on. As the block sizes get larger in an IOPS-oriented workload, the number of IOPS will decrease. What that means is, even though it's fewer IOPS, you're getting more data because the block sizes are bigger. You're getting bigger chunks of data.

If the application is an IOPS-oriented workload -- that means you're interested in how quickly individual transactions can occur -- the workloads that generally do that tend to have the smaller block sizes. And so you're going to see 4K or 8K, something in that range, often for database workloads but not limited to database workloads.

When you get to the larger block sizes, like 64K, 128K, 256K or even higher, then you're typically talking about a throughput-oriented workload, which is not so much transaction sensitive, but it's more how much bandwidth can you get out of it or how big of a boat can you push through this. So, with the smaller block sizes, the bandwidths will be smaller. And as you push the block size up, then the bandwidth will go up and the throughput will go up.

So, IOPS will be high for small block, and throughput is going to see better numbers at the large block.

How does the random or sequential nature of the workload used in the test environment affect the performance numbers?

Martin: If you're doing sequential workloads, that means you're doing a backup job or you're streaming a video or something where you start at one point in the data and you just keep accessing contiguous blocks in sequential fashion. Sequential workloads tend to be faster generally because they are doing things contiguously, and typically the sequential workloads are also using larger blocks.

If you're doing random, you're doing something like a transactional workload or a file directory workload where you've got lots of different users accessing different things all at the same time. There's no real way to predict what the next block will be that is requested because you've got different users doing different things. The more users you have, the more applications you have that are accessing the same storage, the more and more it's going to shift over toward a random workload by the time it hits the storage system. Random workloads tend to be harder on storage systems, and this is true of
We're increasingly seeing latencies for all-flash arrays drop from milliseconds to microseconds. How important is the distinction between microseconds and milliseconds for the average enterprise data center?

Martin: For some workloads, latencies are actually more important than either IOPS or throughput. So where the round trip time or the latency is extremely important, you want it to of course be as low as possible. And microseconds are the next order of magnitude smaller than milliseconds. Just for definition purposes, 1,000 microseconds is equal to one millisecond.

Typically, with hard-drive arrays, you’re going to have millisecond range latencies [of] 2 or 3 or 4 or 5. If it’s not so good, it’ll be of course higher than that -- into double digits.

For some applications, that's just not good enough, and that's why people are looking at all-flash arrays because now you can get latencies down into the very low milliseconds or down into the microseconds, which just means you’ve got very fast turnaround time.

In the final analysis, what recommendations do you have for IT pros trying to sift through the performance claims of all-flash array vendors?

Martin: Look for which of these tests are similar to the workloads that you have. If they’re running, for example, a 100% random read workload test that has 4K block size, then you have to ask yourself, ‘Is my workload like this?’ The tests you want to pay attention to are the ones that are closest to your workloads in your environment. So, that presupposes that you know what those workloads are and that you’ve been measuring them and you at least have some ideas.

I would go through the tests and say, ‘All right, is it read versus write? What’s the percentage? What's the read-write ratio? Is it 100% read? Or is it 50% read? That kind of thing. Again, look at the random vs. sequential. Is it 100% random? Is it 50% random? I would look at the block size. Make sure the block sizes are the same. I would also look at protocol to see if this is a file workload versus a block workload. If they go into it, then I would look for a fifth item in my list of I/O characteristics, and that would be queue depth. That's the number of I/Os that you should have at the same time. So, you really have all of those five I/O characteristics or parameters that you want to look at and compare.
Free resources for technology professionals

TechTarget publishes targeted technology media that address your need for information and resources for researching products, developing strategy and making cost-effective purchase decisions. Our network of technology-specific Web sites gives you access to industry experts, independent content and analysis and the Web's largest library of vendor-provided white papers, webcasts, podcasts, videos, virtual trade shows, research reports and more—drawing on the rich R&D resources of technology providers to address market trends, challenges and solutions. Our live events and virtual seminars give you access to vendor neutral, expert commentary and advice on the issues and challenges you face daily. Our social community IT Knowledge Exchange allows you to share real world information in real time with peers and experts.

What makes TechTarget unique?

TechTarget is squarely focused on the enterprise IT space. Our team of editors and network of industry experts provide the richest, most relevant content to IT professionals and management. We leverage the immediacy of the Web, the networking and face-to-face opportunities of events and virtual events, and the ability to interact with peers—all to create compelling and actionable information for enterprise IT professionals across all industries and markets.

Related TechTarget Websites

- SearchDataBackup
- SearchCloudStorage
- SearchDisasterRecovery
- SearchVirtualStorage
- SearchStorage
- SearchSMBStorage