All-Flash: The Essential Guide

Your guide to all-flash storage, how it stacks up against hybrid and PCIe, and how to measure the benefits
In this e-guide:
Computer Weekly keeps its finger on the pulse in the world of data storage, with regular news, features and practical articles. This guide offers a comprehensive survey of the all-flash array market. We look at all-flash products from the big six storage vendors and the startups and specialists. Plus, we give the lowdown on all-flash vs hybrid flash arrays and all-flash vs server-side PCIe SSD. Finally, we look at the shortcomings of measuring the value of flash by metrics designed for spinning disk and suggest practical alternatives.

Antony Adshead, Storage editor
In all-flash storage 2015 was year of consolidation, incremental improvement and price reduction from the big six storage suppliers.

In a short space of time, the big six storage providers have built out their flash product offerings into mature and scalable platforms. The past few months has seen a focus on price, with many suppliers – including the startups – aiming for the magical $1/GB price point.

The all-flash market has become super-competitive, with only Hewlett Packard Enterprise (HPE) seeing revenue growth over the past 12 months (at least according to IDC), and an obvious race to the bottom on price ensuing as 3D and TLC NAND starts to be adopted.

All the incumbent suppliers (HDS, EMC, HPE, NetApp, IBM and Dell) started the year with existing platforms that underwent upgrades to features, capacity and performance. The big six – except for NetApp – had settled on a flash product strategy, with a mix of acquisition or in-house development.

The question we have to ask is where the market will choose to compete next. Indications are that we’ll see a focus on automation and systems, and
matching of all-flash platforms to specific workloads such as containers, virtualisation and databases.

**Hitachi Data Systems**

Hitachi Data Systems (HDS) announced its VSP F-Series in November 2015. Up to that point, VSP G-series arrays supported all-flash configurations, but they weren’t eligible for performance measurement by analyst firms that insist all-flash arrays must be incapable of supporting traditional spinning media.

HDS felt its leadership in the market, which is based on proprietary flash drive technology, wasn’t being accurately reported. The F-Series – available in models F400, F600 and F800 – seeks to redress the balance with system capacities up to 448TB (after data reduction).

In November 2015, HDS also released its second generation of Flash Module Devices (FMDs). These are a bespoke technology that integrates NAND flash and HDS-developed hardware into a custom flash module with capacities up to 6.4TB. The module also introduces inline data compression that it claims has no effect on performance.

In January 2016, HDS added another key piece to its all-flash offerings, with the HFS A Series all-flash arrays. These are data deduplication-enabled arrays that allow HDS to target high-performance, all-flash workloads such as VDI.
The A series packs up to 60 standard SSDs into one 3.5 inch 2U chassis and uses off-the-shelf 16TB multi-level cell (MLC) SSDs, rather than the 6TB-plus custom-built FMDs that HDS offers with its other all-flash options.

A Series models available are the A220, with 10 drives and a raw capacity of 16 TB (effective capacity 64TB with data reduction); the A250 with 30 and capacity of 48 TB (192TB effective); and the A270 with 60 SSDs for a raw capacity of 96TB effective capacity of 384TB.

**EMC**

At EMC World in May 2015, EMC announced the release of XtremIO 4.0 and the parallel availability of XIOS 4.0, the operating system that drives the XtremIO platform. The hardware allows systems to scale up to eight X-Brick nodes in two-brick increments, each at a capacity of 40TB using 16GB SSDs.

Maximum XtremIO system capacity is now 320TB (269TB usable) or 1,612TB effective capacity with data reduction. Availability features in XtremIO were improved, with the ability to non-disruptively add nodes to an existing cluster. The platform was also upgraded with native RecoverPoint integration to provide data replication between appliances in geographically dispersed locations.
Eighteen months down the line, there is still no release date in sight for EMC’s other all-flash platform, DSSD. EMC acquired the company in 2014, making a big announcement at EMC World that year.

DSSD is slated to be a direct-connect all-flash appliance that will deliver significantly higher levels of performance and lower latency compared with even that seen with XtremIO. The product will be targeted at niche applications that demand ultra-high levels of throughput, but which still need the benefits of externally connected storage.

**Hewlett Packard Enterprise**

HPE upgraded and replaced its 7000 and 10000 platforms with the **8000** and **20000 series**. The 8450 all-flash system replaces the previous 7450 model and uses the latest Generation 5 custom **ASIC** that 3PAR systems have become well known for.

At the top end of the range, the HPE 3PAR 20850 scales up to 4PB of raw capacity and an effective capacity – after data reduction – of 15PB. HPE plans to improve performance by use of 3.84GB SSDs based on **3D-NAND** technology later this year.

The 3PAR platforms introduce functionality that now includes **Online Import** for HDS and XIV systems, bi-directional Peer Motion, **asynchronous** replication and improved end-to-end data integrity with **Persistent Checksum**.
Since acquisition in 2010, HPE has invested heavily in the 3PAR platform, and has continually brought out new features and products.

**NetApp**

NetApp has seen a year of change in terms of strategy with the death of FlashRay and the proposed acquisition of all-flash array start-up Solidfire. The FlashRay product (codenamed Mars) was intended to meet demand for bespoke all-flash systems from competitors like EMC with its XtremIO. However, the project seemed to be continually delayed and suffered the final ignominy when project head Brian Pawlowski left for Pure Storage in May 2015.

NetApp says the IP developed for FlashRay has been integrated into the all-flash version of Clustered ONTAP, otherwise known as the **AFF or All-Flash FAS**, released in 2014.

NetApp worked on improvements that optimise read/write performance and most recently implemented inline data deduplication, a feature addition that was long overdue. The top of the range system now scales up to 2880 (NAS) and 960 (SAN) drives, with maximum raw capacities of 10.94PB and 3PB respectively, or 45PB and 15PB after data optimisation.

The proposed acquisition of Solidfire, announced in December 2015, provides NetApp with the ability to target cost-conscious and automation-
In this e-guide

- All-flash array roundup 2016: The big six
- All-flash storage roundup 2016: The startups
- Hybrid flash vs all-flash storage: When is some flash not enough?
- PCIe SSD vs all-flash for enterprise storage
- How to measure flash storage’s true value

- obsessed service providers with a system that is more scale-out than FAS can offer.

**IBM**

IBM has continued to focus on its FlashSystem product line, the successor to technology acquired with the purchase of Texas Memory Systems in 2012.

In February 2015, IBM announced the FlashSystem 900 and FlashSystem V9000. The former is a flash appliance that scales to a maximum 105TB of capacity, with write latency as low as 90 microseconds. The latter is a hybrid appliance built from FlashSystem and IBM’s SVC (SAN Volume Controller) technology. This pairing allows IBM to introduce features, such as replication and real-time compression, that would otherwise not exist in the 900 platform.

The positioning for 900 and V9000 models is one of performance over features. The 900 offers low latency and high throughput, while the V9000 sacrifices some latency to implement the benefits of flexible configuration, data reduction and high availability.

**Dell**

Dell continues to focus on all-flash implementations of its SC series arrays, developed from technology acquired when it purchased Compellent in 2011.
The Dell SC4020 was one of the first all-flash systems to use **TLC** drive technology, a lower-cost evolution of the MLC storage used in most all-flash systems today. A fully specified SC4020 provides up to 1PB of storage in a four-controller clustered configuration, with the capability to use any of **SLC, MLC or TLC** drive types.
All-flash storage roundup 2016: The startups

Chris Evans, Contributor

2015 was a mixed one for the startups in the all-flash array market. There have been acquisitions, initial public offerings, new entrants and some all-flash pioneers have not fared as well as they might like.

There have, however, been a number of developments among the startups in the flash storage market.

The first is in feature sets. All suppliers have started to produce feature-rich products that have data protection (RAID, replication) and optimisation (thin provisioning, compression, data deduplication) features as standard.

At the same time, price has become a battleground, with $1/GB being the magic target each supplier wants to achieve, in much the same way the 1ms response time was seen as the magic performance number. This low price point is being achieved through the use of the latest technology such as TLC NAND, as well as data optimisation techniques, which – of course – use supplier-quoted reduction ratios.
All-flash has become extremely competitive and hybrid flash array suppliers have entered the all-flash market with new variants of their products. A notably absent supplier in this regard is Nimble Storage, which currently only has hybrid offerings.

With the amount of competition in place, 2016 could see an interesting shakedown in the market, with more startups acquired or going out of business.

For example, this year's roundup excludes Nimbus Data, which had no product news or announcements. Skyera, which featured in previous roundups, was acquired by HGST and has not resurfaced as a product under the new company.

**Coho Data**

Coho Data raised a further $30m in investment in May 2015 and released its first all-flash product, the **2000f**. Coho's architecture is built of multiple MicroArrays, 2U storage nodes that are connected by 10Gbps Ethernet SDN-based networking infrastructure.

Each 2000f system consists of two controllers, four 16TB PCIe SSD flash devices and up to 24 SSDs in multiple capacity configurations. This provides scalability from 7TB to 47TB of raw capacity or 11TB to 93TB of usable capacity after data optimisation.
Each 2000f node is capable of up to 320,000 read IOPS or 220,000 read/write IOPS. Clusters can be built by scaling out multiple MicroArrays, which – in this case – can be all-flash-based or hybrid as required.

**Kaminario**

Kaminario announced v5.5 of its K2 platform in August 2015.

The latest systems make use of 3D TLC NAND technology, making Kaminario one of only a few current suppliers to use this (others include Dell, Hewlett Packard Enterprise and SolidFire).

K2 systems can now scale out and scale up to a maximum raw capacity of around 740TB and a projected 1.44PB after data reduction technologies are applied. Performance is rated at 250,000 IOPS for a single K-block system, and up to 1 million IOPS and 12.8GBps of bandwidth for a fully scaled K2.

Version 5.5 of K2 also introduced native array replication and improved financial terms called Perpetual Array, which allows customers to mix and match hardware generations while extending maintenance across an entire K2 cluster.

**NexGen Storage**

NexGen Storage came to the all-flash market with two arrays in November 2015. The N5-1500 and N5-3000 models both use a mix of PCIe flash SSDs (2.6TB in each), combined with either 15TB to 60TB or 30TB to 60TB of...
solid state drives. The two models effectively set two possible entry points for customers.

NexGen is one of only a few suppliers looking to tier flash by deploying PCIe and SSD. This points to a trend we may see more of in 2016. Its systems also offer quality of service to enable scaling in a multi-tenant environment. Both models scale up to 450,000 IOPS and up to 6GBps of throughput.

**Pure Storage**

Pure Storage moved from startup to established business with its Initial Public Offering in October 2015. Prior to that, in June 2015, the company launched a revamp of its FlashArray platform based on new hardware, with three new models under the FlashArray//m brand. The //m20 (5TB to 40TB raw), //m50 (30TB to 88TB raw) and //m70 (44TB to 136TB) models scale to in excess of 120TB, 250TB and 400TB respectively after data optimisation.

Performance scales from 150,000 IOPS at the low end with the //m20 and up to 300,000 IOPS for the //m70, all with less than 1ms latency. Pure has also introduced Evergreen Storage, a buying model to reduce the impact of forklift upgrades.
SolidFire

SolidFire released version 8 of its Element OS, code named Oxygen, in June 2015. The release brings in incremental improvements in multi-tenancy through extended VLAN tagging, new data protection features (synchronous replication, snapshot replication) and additional management features with LDAP authentication.

In February 2015, SolidFire released the SF9605, the latest option in its scale-out node platform. Nodes now scale from 2.4TB to 9.6TB in capacity (raw) or from 10TB to 40TB usable, after data optimisation. Each node is capable of around 50,000 IOPS (SF9010 at 75,000 IOPS).

At the time of writing, NetApp was in the process of acquiring SolidFire with the deal set to close in the first quarter of 2016.

Tintri

Tintri, an existing hybrid array startup, moved into the all-flash market in August 2015 with the release of two all-flash VMstore systems (T5060 and T5080). This was followed in December 2015 with a new entry-level model, the T5040.

The three models scale from 5.76TB to 23TB of raw capacity or 18TB to 73TB after data optimisation. Tintri claims the systems will support between 1,500 and 5,000 virtual machines, with NFS and SMB3 support covering
In this e-guide

- All-flash array roundup 2016: The big six
- All-flash storage roundup 2016: The startups
- Hybrid flash vs all-flash storage: When is some flash not enough?
- PCIe SSD vs all-flash for enterprise storage
- How to measure flash storage’s true value

All-Flash: The Essential Guide

VMware vSphere, Red Hat Enterprise Virtualisation, Microsoft Hyper-V and XenServer hypervisors. VMstore also supports OpenStack (Juno release) and later.

**Tegile**

Tegile, also an existing hybrid storage startup, first came to the all-flash market with the release of the IntelliFlash T3800 in June 2014. The product family was extended in November 2014 with the T3600 and T3700.

The three systems, based on ZFS, all come with 192GB of system memory and offer varying levels of flash capacity from 12TB to 48TB (raw), scaling to 314TB to 336TB (raw) across the range.

In August 2015, Tegile announced the IntelliFlash HD platform, a system that combines the IntelliFlash controller and high density InfiniFlash enclosures from SanDisk. This system provides up to 512TB of flash in 3U, with sub-millisecond latency and a quoted 5 million IOPS per rack. High density flash is another trend to look out for in 2016.

**Violin Memory**

Violin Memory released the Flash Storage Platform (FSP) in February 2015, with additional models in December 2015.

The FSP 7000 series now has five variations (7700, 7600, 7300, 7300E and 7250) that scale from the entry-point FSP 7250 with 8.8TB to 26.6TB and
250,000 IOPS at 1ms latency to the FSP 7700 with 700TB (raw), 2.1PB usable and up to 2 million IOPS, 10GBps throughput and a minimum of 0.18ms latency.

Violin still offers the previous 6000 generation systems and Windows Flash Array, an SMB-based platform for use with Microsoft Windows systems and applications.

X-IO Technologies

X-IO Technologies debuted an all-flash version of its Intelligent Storage Element (ISE) to the market in March 2015. The three ISE 800 series models (ISE 820 G3, ISE 850 G3 and ISE 860G3) provide between 6.4TB and 512TB of capacity at up to 400,000 IOPS and 5GBps of throughput.

At release, the ISE 800 held the top spot for the Storage Performance Council’s SPC-1 price/performance benchmark.

X-IO added features to the ISE series, including synchronous mirroring, thin provisioning, quality of service and management through a Rest-based application programming interface.
Hybrid flash vs all-flash storage: When is some flash not enough?

Bryan Betts, Guest Contributor

Let's start by taking a brief look at the available options. Much has been written in these pages about the growing acceptance and usefulness of flash storage within enterprise-grade arrays. Initially, flash-based solid-state drives (SSDs) were expensive, so the fastest and least costly way to take advantage of them was to retrofit them to an existing disk-based array — as a tier zero, for example.

As prices fell, tier one flash also became an option, especially for those with auto-tiering technology that could automatically move the hottest data onto the fastest tier. However, this still left the array optimised for spinning disks — although this is changing as suppliers update their array software — so other developers realised they could instead design an array optimised for flash, with secondary tiers of cheaper spinning disks for longer-term data.

Remember here that flash is not just a faster version of spinning disk — it is a fundamentally different medium. Yes, it can perform the same tasks as a disk drive, but it works differently, so if your array firmware and your applications continue to address it as a disk you will be wasting much of the advantage available to you. For instance, applications may buffer writes to
cope with disk latency, but your developers now need not factor that into things.

Lastly, there are the all-flash arrays with no spinning disks at all. Initially these – just like their DRAM-based forerunners from Digital Equipment, Texas Memory Systems and others – targeted the most performance-hungry and latency-sensitive applications where cost was much less of an issue, but as time went by and flash costs fell further, the all-flash approach started to make sense for a much broader spread of enterprise applications.

This last process was greatly assisted by the adoption of denser and therefore cheaper flash technologies such as multi-level cell (MLC) and triple-level cell (TLC), which store two and three bits per cell respectively, and have enabled the creation of capacity and performance-optimised variants of flash. The denser chips are less reliable, but as in several other areas of technology, we can use software to more than compensate for this.

Another factor that made all-flash arrays easier to adopt was that while they initially lacked the sophisticated storage management capabilities of the established disk arrays, this is no longer the case. Some may even exceed the capabilities of disk arrays, especially in areas where flash excels, such as continuous data protection (CDP), guaranteed quality of service and the creation of capacity-free snapshots.
Analyse the workload

So why would you still want a hybrid array, and why would you choose a purpose-designed hybrid over a retrofitted one? To start answering those questions you must analyse the workload. The more random it is — and virtual desktop infrastructure (VDI) is highly random, for example — then the more appropriate all-flash will be.

Customer-facing online transaction systems and databases are also sweetspots for all-flash, as is virtualisation in general. Sequential workloads and those with large proportions of cold data are a different matter. Essentially though, what we used to call cache hits and misses still matter, and with all-flash everything can be a cache hit.

With raw flash capacity costing an average of perhaps $5 per GB, hybrid arrays also tend to be cheaper, but cost comparisons between disk and flash are nuanced. For instance, flash is more compact and less power-hungry, so what you win on the purchase price of a hybrid you might lose on its operating costs and the space it takes up.

Be aware too that while some all-flash suppliers are already claiming price parity, this is hard to assess because of the effects of the data reduction technologies used to get more data into a given volume of flash. These technologies — the primary ones being data deduplication and compression — can reduce the effective price per GB by a factor of 5:1 or more (to $1 per
In this e-guide

- All-flash array roundup 2016: The big six
- All-flash storage roundup 2016: The startups
- Hybrid flash vs all-flash storage: When is some flash not enough?
- PCIe SSD vs all-flash for enterprise storage
- How to measure flash storage's true value

GB or less), but their performance can vary considerably depending on the type of data.

As an aside, this is why any decent flash-enabled array needs to include a variety of data reduction tools. For example, data deduplication will work much better on virtual machine images or some file-based tasks than on databases, while compression will work better on databases than for photos or videos. In addition, some arrays allow you to turn these features off on a per-LUN basis – look out for this if you plan to consolidate multiple applications onto a flash array, as your most latency-sensitive applications may need data reduction turned off.

The advantage of purpose-designed hybrids over retrofits is primarily that the former should accelerate all I/O, not just the data placed on the flash tier. However, this is increasingly true of all hybrid arrays, thanks to improved data management software which automatically promotes the hottest data to flash. Indeed, with some systems you must actively pin data to the flash tier if you want it to stay there regardless of its "temperature".

Many hybrids use flash as a read-only cache, while others use it as a write cache too, in part to speed up access to slower hard drives. Writes may be cached in RAM until they can be written as an entire flash page, or as sequential input/output (I/O).
Interoperability with existing infrastructure

The question is complicated further by whether or not the new all-flash or hybrid array must interoperate with an existing storage infrastructure. For example, if your current disk-based NetApp filers are coming off maintenance then you could consider going all-flash, otherwise you will want at least some interoperability. Adding flash to create a hybrid may be preferred; alternatively if your filers are clustered, check whether you can add an all-flash unit to the cluster, in effect turning the whole thing into a hybrid.

Scalability and the upgrade path is also important. How granular are the all-flash options and will you end up paying for a bigger system than you need, and will it need a forklift upgrade when it reaches capacity?

In summary, there is little argument now that the first tier should be flash. Major storage suppliers already report that shipments of 15,000rpm disks have almost entirely been replaced by SSDs. The question is what do we choose for the second (and perhaps third and fourth) tiers – disk, or more flash?

Hard disk could be better for some tasks, such as media streaming, and hard disk technology is still evolving, but its properties change as it does so. For example, the latest high-capacity shingled media is good for reading, but difficult to write to and delete from, making it more of an archive medium.
Flash too is a read-optimised technology that presents additional challenges when you need to write to it, such as the requirement to erase used space before rewriting to it. And while flash is currently moving to denser yet more reliable 3D cell structures, current NAND flash technologies may only have a couple of generations left. There are several more non-volatile technologies evolving in the wings though, such as magnetoresistive RAM, ferroelectric RAM and phase-change RAM.

Flash and its successors will therefore migrate to more and more tiers. Disk will still have its uses and advantages, but like tape before it, they will become ever more narrowly defined.
Solid-state storage, like rotating media, can slot into an enterprise computing architecture in a variety of ways.

And while much recent interest has focused on all-flash arrays, there are other key locations where flash-based storage can play a big role in accelerating application performance -- and especially hyperscale computing performance.

Key among these is server-side flash, installed in a PCI Express (PCIe) slot where it can work as local disk storage or as cache for a server. The former replaces at least some of the requirement for all-flash array-style shared storage, while the latter can augment network storage or the server’s main memory.

 PCIe SSD has several advantages when used as storage local to the server. Not only is it rather cheaper than the same volume of all-flash array capacity, but PCIe eliminates the host bus adapter (HBA) or drive controller and its latency, plus the overhead from the network connection.
This can bring latency down from milliseconds to microseconds. At the same time, all the other benefits solid-state storage has over spinning disk are there, such as lower power consumption, heat generation and vibration, greater robustness and higher density.

Most SSDs today are based on flash but, to be accurate, we should speak of non-volatile memory (NVM) in general. That is because there are other memory types besides flash in use and in development that promise even better performance.

The problem with putting SSD in the server is the same problem you have with any direct-attached storage (DAS).

Software-defined storage

That is, unless you are going to run specialist disk-sharing software -- for example, using software-defined storage tools to abstract local storage and pool it network-wide -- then local storage is available only to its host server. But, as we will see, there are emerging use-cases such as hyperscale computing, where local SSD can actually be more useful and cost-effective than shared storage in the network.

So, the big advantage of all-flash arrays and other forms of shared storage is that it can be available to any connected system. Unallocated space can be shared, as can data on shared volumes. The former is a win because DAS is typically over-provisioned to allow for future growth and, of course, we
In this e-guide

- All-flash array roundup 2016: The big six
- All-flash storage roundup 2016: The startups
- Hybrid flash vs all-flash storage: When is some flash not enough?
- PCIe SSD vs all-flash for enterprise storage
- How to measure flash storage's true value

typically over-specify by a considerable margin to avoid the pain of having to migrate or rebuild an overloaded server in the future.

Combined with technologies such as thin provisioning, where a logical volume takes up only the physical space on disk that its data requires and not its full provisioned space, shared storage has allowed considerable economies to be achieved. Also, once you add appropriate locking mechanisms, servers and applications can share access to the same data, yielding even more advantage.

Server-side SSD on a PCIe board can still be useful in this environment as cache, of course. Caching pre-loads the most frequently accessed data, with the master copy still stored on, and synchronised back to, shared storage in the network. This can be a good way to boost the performance of specific applications.

However, the possibility of a cache miss -- where the required data is not cached and must be retrieved normally over the network -- means that the resulting latency improvement is not consistent or guaranteed.

This is made worse by the fact that some applications have data access patterns that are difficult to identify and cache accurately, making cache misses more likely. There are also issues involved in maintaining coherency between local cache and the remote master data, and questions over the suitability for caching of the large files that store entire virtual machines.
More and more interest is therefore turning towards the idea of using server-side PCIe SSD as local working storage, helped both by the falling cost of flash in particular and by two key developments in IT.

The first is the creation of specifications to implement enterprise-grade server-side SSD, and second is the evolution of new computing paradigms that make shared working storage more of a liability than an advantage.

**Relevant standard**

One relevant standard is obviously PCIe itself, which is currently at generation 3.0, with generation 3.1 coming and 4.0 under development. Architecturally, PCIe is serial, with the ability to consolidate or bond multiple lanes into a single high-speed point-to-point connection.

As well as PC expansion boards, there are 2.5-inch SSDs that use four-lane PCIe over a U.2 (formerly SFF-8639) connector, while two-lane PCIe is also part of the SATA Express (SATAe) specification for external storage that is intended to succeed SATA 3.0.

In addition, the Intel/Apple Thunderbolt interface incorporates four-lane PCIe, while there are specifications to extend PCIe outside the system box, turning it into a kind of short-range storage network but without the latency of a SAN.
Perhaps more important, in some ways, is NVM Express (NVMe), which is a specification to access solid-state storage on a PCIe device (whether PCIe, U.2 or SATAe). It allows the host to take full advantage of the parallelism and low latency of a PCIe-connected SSD, where formerly these devices required proprietary software drivers.

In effect, where PCIe cuts out the storage controller latency, NVMe removes most of the remaining software latency. NVMe means the operating system needs only one standard driver to support any NVMe SSD -- and SSD developers do not need to create their own drivers, thus removing the scope for compatibility issues they can bring. It is also optimised for solid-state storage, unlike superficially similar technologies developed in simpler and slower days, such as AHCI.

(Not all server chipsets will support PCIe 3.0, though, so you need to check. You also need to ensure your server and your storage product support NVMe and the right level of cabling.)

**Hyperscale computing**

Then, building on that underlying evolution in the server platform comes the paradigm shift toward hyperscale computing.

Legacy applications were built on the assumption that high reliability would be baked into the infrastructure, but the same is not true of the modern scale-out design models popularised by the likes of Google and Facebook.
Here, the hyperscale concept comprises independent yet connected nodes where redundancy is at the level of entire server/storage nodes rather than components in the server and shared storage architecture.

So where legacy applications require shared storage in case they need to failover, and are therefore better suited to an all-flash array-type deployment, modern scale-out applications can have thousands of clustered nodes. Each node is typically based on off-the-shelf hardware and provides compute, storage and networking resources.

The nodes are then clustered together and managed as a single entity, replicating among themselves for availability via distributed storage platforms such as Apache Cassandra, Ceph and Hadoop -- you could even call the result a redundant array of inexpensive servers.

All of a sudden, the advantage of having local storage with super-low latency becomes greater than the disadvantage of that storage not being shared.
Flash storage, or using the broader term, solid-state storage, suffers from an inadequate measure of value. Flash storage provides a step-function improvement in the ability to store and retrieve information. The value of processing with flash storage compared to access from electro-mechanical devices is not easy to express.

Many in the industry still use a “data at rest” measure, which is the cost of storing data. That fails to represent more valuable characteristics such as access time and longevity. The data at rest measure, given as dollars per GB, can be misleading and does not convey real economic value. If that is the only measure to use for information storage, then you should use magnetic tape for all operations because it is the least expensive media.

Some vendors also use a dollars per IOPs measure for all-flash storage systems. This measure does not represent the value of what flash can accomplish because it is an aggregate number. This means it represents the total number of I/Os a system can do, which could be from thousands of short-stroked disk drives. It does not directly reflect the response time increase, which is the most meaningful measure in accelerating applications and getting more work done.
So if these measures are inadequate, what is the best way to gauge the value of flash storage? It actually varies depending on the case. Flash can provide key improvements, including consolidation, acceleration, reduction in physical space/power/cooling, longevity, and reduced tuning. Let’s look at these:

- Consolidation – The greater performance levels of flash storage allow for the deployment of more diverse workloads on a single system. With larger capacity flash storage systems, workloads running on multiple spinning disk systems can be consolidated to a single flash storage system. The value of consolidation includes a reduction of the number of systems to manage and the physical space required.

- Acceleration – The first deployments of flash systems focused on accelerating applications (mostly databases), and virtual machine or desktop environments. Acceleration enabled more transactions and improvements in the number of VMs and desktops supported. The successes here drove the shift to more widespread use of solid-state storage technology.

- Physical space – Flash technology increases the capacity per chip and results in less physical space required. Even flash packaged in solid-state drives have eclipsed the capacity points of hard disk drives. With flash storage, more information can already be contained within physical space than was previously possible and technology
gains are still improving in this area. This is important for most organizations where information storage represents a large physical presence.

- Power and cooling – Storage devices using flash technology consume less power and generate less heat (requiring less cooling) than devices with motors and actuators. There is an obvious reduction in cost from this improvement. But this becomes more important when physical plant limitations prevent bringing in more power and cooling to the data center.

- Longevity – Probably the least understood added value from flash storage is the greater longevity in usage for the flash devices and the economic impact that brings. The reliability and wear characteristics are different from electro-mechanical devices, and have reached a point where vendors are giving seven- and 10-year guarantees and even some lifetime warranties with ongoing support contracts. This dramatically changes the economics from the standpoint of total cost of ownership over the long lifespan. The key driver of this is the disaggregation of the storage controller or server from the flash storage enclosures that allows controllers to be updated independently. This has led to some “evergreen” offerings by vendors, which actualizes the economic value in this area.
In this e-guide

- All-flash array roundup 2016: The big six
- All-flash storage roundup 2016: The startups
- Hybrid flash vs all-flash storage: When is some flash not enough?
- PCIe SSD vs all-flash for enterprise storage
- How to measure flash storage's true value

- Reduction in tuning – One of the most often reported benefits (which can be translated to economic value) from deployment of flash storage is the reduction in performance tuning required. This means there is no longer a need to chase performance problems and move data to balance workloads with actuator arms.

It is clear that a data at rest measure is inadequate. Nevertheless, price is always an issue and the cost for flash storage continues to decline at a steep rate because of the investment in technology. Data reduction in the form of compression and deduplication also is a given for the most part in flash storage, multiplying the capacity stored per unit by 4-1 or 5-1 in most cases. The continued technology advances will improve costs even more.

The $/GB data at rest measure is difficult for many to stop using, even though it misrepresents true value. People do it because it is easy and it is a habit after years of measuring value that way. However, it is wrong. There needs to be another relatively simple measure to encompass all the values noted earlier. It may take a while for that to come about. In the meantime, we will continue to look at economic value, do TCO economic models, and explain real value as part of evaluating solutions to handling information.

Next article
In this e-guide

- All-flash array roundup 2016: The big six
- All-flash storage roundup 2016: The startups
- Hybrid flash vs all-flash storage: When is some flash not enough?
- PCIe SSD vs all-flash for enterprise storage
- How to measure flash storage's true value

---

Getting more CW+ exclusive content

As a CW+ member, you have access to TechTarget’s entire portfolio of 120+ websites. CW+ access directs you to previously unavailable “platinum members-only resources” that are guaranteed to save you the time and effort of having to track such premium content down on your own, ultimately helping you to solve your toughest IT challenges more effectively—and faster—than ever before.

Take full advantage of your membership by visiting www.computerweekly.com/eproducts

Images: Fotalia

© 2016 TechTarget. No part of this publication may be transmitted or reproduced in any form or by any means without written permission from the publisher.