

# The Early Implications of NVMe/TCP on Ethernet Network Designs

The ratification in November 2018 of the NVMe/TCP standard officially opened the doors for NVMe/TCP to begin to find its way into corporate IT environments. Earlier this week I had the opportunity to listen in on a [webinar](#) that SNIA hosted which provided an update on NVMe/TCP's latest developments and its implications for enterprise IT. Here are four key takeaways from that presentation and how these changes will impact corporate data center Ethernet network designs.

***First, NVMe/TCP will accelerate the deployment of NVMe in enterprises.***

NVMe is already available in networked storage environments using competing protocols such as RDMA which ships as [RoCE](#) (RDMA over Converged Ethernet). The challenge is no one (well, very few anyway) use RDMA in any meaningful way in their environment so using RoCE to run NVMe never gained and will likely never gain any momentum.

The availability of NVMe over TCP changes that. Companies already understand TCP, deploy it everywhere, and know how to scale and run it over their existing Ethernet networks. NVMe/TCP will build on this legacy infrastructure and knowledge.

***Second, any latency that NVMe/TCP introduces still pales in comparison to existing storage networking protocols.***

Running NVMe over TCP does introduces latency versus using

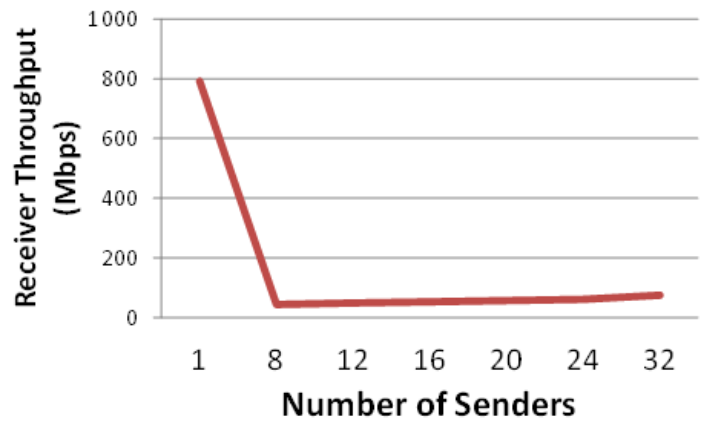
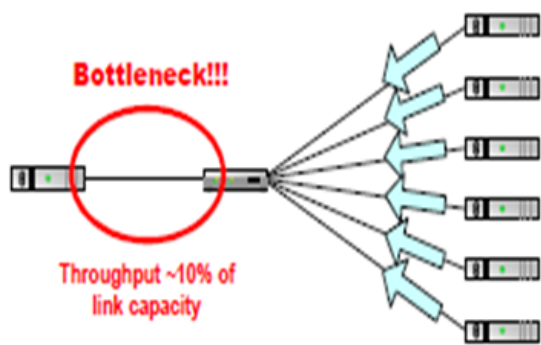
RoCE. However, the latency that TCP introduces is nominal and will likely be measured in microseconds in most circumstances. Most applications will not even detect this level of latency due to the substantial jump in performance that natively running NVMe over TCP will provide versus using existing storage protocols such as iSCSI and FC.

***Third, the introduction of NVMe/TCP will require companies implement Ethernet network designs that minimize latency.***

Ethernet networks may implement buffering in Ethernet switches to handle periods of peak workloads. Companies will need to modify that network design technique when deploying NVMe/TCP as buffering introduces latency into the network and NVMe is highly latency sensitive. Companies will need to more carefully balance how much buffering they introduce on Ethernet switches.

***Fourth, get familiar with the term “incast collapse” on Ethernet networks and how to mitigate it.***

NVMe can support up to 64,000 queues. Every queue that NVMe opens up initiates a TCP session. Here is where challenges may eventually surface. Simultaneously opening up multiple queues will result in multiple TCP sessions initiating at the same time. This could, in turn, have all these sessions arrive at a common congestion point in the Ethernet network at the same time. The network remedies this by having all TCP sessions backing off at the same time, or an [incast collapse](#), creating latency in the network.



*[Source: University of California-Berkeley](#)*

Historically this has been a very specialized and rare occurrence in networking due to the low probability that such an event would ever take place. But the introduction of NVMe/TCP into the network makes the possibility of such an event much more likely to occur, especially as more companies deploy NVMe/TCP into their environment.

## The Ratification of the NVMe/TCP

Ratification of the NVMe/TCP standard potentially makes every enterprise data center a candidate for storage systems that can deliver dramatically better performance to their work loads. Until the performance demands of every workload in a data center are met instantaneously, some workload requests will queue up behind a bottleneck in the data center infrastructure.

Just as introducing flash memory into enterprise storage systems revealed bottlenecks in storage operating system software and storage protocols, NVMe/TCP-based storage systems will reveal bottlenecks in data center networks. Enterprises seeking to accelerate their applications by implementing NVMe/TCP-based storage systems may discover bottlenecks in their networks that need to be addressed in order to see the full benefits that NVMe/TCP-based storage.

To view this presentation in its entirety, follow this [link](#).

---

# DCIG Quick Look: iXsystems TrueNAS X10 Offers an Affordable Offramp from Public Cloud Storage

For many of us, commuting in rush hour with its traffic jams is an unpleasant fact of life. But I once had a job on the outer edge of a metropolitan area. I was westbound when most were eastbound. I often felt a little sorry for the mass of people stuck in traffic as I zoomed—with a smile on my face—in the opposite direction. Today there is a massive flow of workloads and their associated storage to the public cloud. But there are also a lot of companies moving workloads off the public cloud, and their reason is cloud economics.

## Cloud Economics Are Not Always Economical

In a recent conversation with [iXsystems](#), it indicated that many of its new customers are coming to it in search of lower-than-public-cloud costs. Gary Archer, Director of Storage Marketing at iXsystems met with DCIG earlier this month to brief us on a forthcoming product. It turns out the product was not the rumored hyperconverged infrastructure appliance. Instead, he told us iXsystems was about to reach a new low as in a new low starting price and cost per gigabyte for enterprise-grade storage.

A lot of companies look at iXsystems because they want to reduce costs by migrating workloads off the public cloud. These customers find the Z-Series enterprise-grade open source storage attractive, but asked for a lower entry price and lower cost per GB.

## iXsystems TrueNAS X10 is Economical by Design

To meet this demand, iXsystems chose current enterprise-grade, but not the highest-end, hardware for its new TrueNAS X10. For example, each controller features a single 6-core Intel Broadwell Xeon CPU. In an era of ever-larger DRAM caches, each X10 controller has just 32GB of ECC DRAM. Dual one-gigabit Ethernet is built in. 10 GbE is optional. Storage capacity is provided exclusively by SAS-attached hard drives. Flash memory is used, but only as cache.

The TrueNAS X10 retains all the redundancy and reliability features of the Z-Series, but at a starting price of just \$5,500. A 20 TB system costs less than \$10,000, and a 120 TB system costs less than \$18,000 street. So, the X10 starts at \$0.50/GB and ranges down to \$0.15/GB. Expansion via disk shelves should drive the \$/GB even lower.

iXsystems positions the TrueNAS X10 as entry-level enterprise-grade unified storage. As such, the TrueNAS X10 will make a cost-effective storage target for backups, video surveillance and file sharing workloads; but not for workloads characterized by random writes. Although iXsystems lists in-line deduplication and compression on its spec sheet, the relatively limited DRAM cache and CPU performance mean you should probably only implement deduplication with caution. By way of example, the default setting for deduplication is off.

In the TrueNAS X10, iXsystems delivers enterprise-grade storage for companies that want to save money by moving off the public cloud. The X10 will also be attractive to companies that have outgrown the performance, capacity or limited data services offered by SMB-focused NAS boxes.

The TrueNAS X10 is not for every workload. But companies with monthly public cloud bills that have climbed into the tens of thousands may find that “cloud economics” are driving them to seek out affordable on premise alternatives. Seek and ye shall

find.

---

# **DCIG 2016-17 Midmarket Enterprise Storage Array Buyer's Guide Now Available**

Since 2010, DCIG Buyer's Guides have been helping organizations make better technology purchasing decisions, faster. DCIG Buyer's Guides drive time, and therefore cost, out of the technology selection process by helping enterprise technology purchasers understand key product considerations and by giving them access to normalized comparative feature data. This enables technology purchasers to quickly identify a short list of products that possess the features required by the organization and then focus their evaluation efforts on the short-listed products.

DCIG's analysts do the leg work for enterprise technology purchasers by:

- Identifying a common technology need with many competing solutions but with little comparative data available to technology purchasers
- Scanning the environment to identify available products in the marketplace
- Gathering normalized data about the features each product supports
- Providing an objective, third-party evaluation of those features from an end-user perspective
- Describing key product considerations and important changes in the marketplace
- Presenting DCIG's opinions and product feature data in a

way that facilitates rapid feature-based comparisons

DCIG recently adopted a Body of Research approach that enables DCIG to be much more responsive to changes in the marketplace. In DCIG's original approach, developing a particular buyer's guide frequently required 9 months from identification of the need to publication of the buyer's guide. Using DCIG's updated body of research methodology, DCIG can produce a specific Buyer's Guide Edition within two months of identifying the need.

By researching an extensive range of products and consolidating the collected data into a single topic-based data repository, DCIG has the flexibility to quickly and effectively analyze that data based on a wide variety of use cases. These use cases may be the traditional classifications based on features such as protocol support and scalability, or any set of features that define an emerging marketplace such as a specific application certification.

DCIG recently completed the first phase of its Enterprise Storage Array Body of Research. This body of research includes more than 130 products from more than 20 vendors. In order to keep this first phase of the research to a manageable scope, DCIG used the following inclusion and exclusion criteria:

- Must be available as an appliance that includes its own hardware and software
- Must be a traditional or hybrid array (the vendor must support a configuration that includes hard disk drives)
- Must support a dual, redundant controller configuration
- Must not be marketed as scale-out NAS
- Must not be marketed as an all-flash array

Scale-out NAS and all-flash arrays are certainly enterprise storage arrays, and overlap in many ways with products in this body of research. The second phase of DCIG's enterprise storage array research will add these products to the DCIG

Enterprise Storage Array Body of Research by the end of 2016.

DCIG is pleased to announce the availability of the *DCIG 2016-17 Midmarket Enterprise Storage Array Buyer's Guide* as the first Buyer's Guide Edition developed from this body of research. Other Buyer's Guides based on this body of research will be published in the coming weeks and months, including the *2016-17 Midrange Unified Storage Array Buyer's Guide* and the *2016-17 High End Storage Array Buyer's Guide*.



The *DCIG 2016-17 Midmarket Enterprise Storage Array Buyer's Guide* includes seventeen (17) storage arrays from the following ten (10) storage providers (in alphabetical order): AMI, Dell, EMC, Fujitsu, HPE, iXsystems, Nimble Storage, Pivot3, Seagate and Tegile. To identify products likely to be of greatest interest to midmarket organizations, DCIG evaluated arrays with a maximum raw storage of 500TB.

The arrays that met DCIG's inclusion requirements and achieved a ranking of *Good*, *Excellent* or *Recommended* are included in this Buyer's Guide. Most of the arrays are the "Lite" version in a series of products. The included products generally provide all of the features of other products in the series, but at a smaller scale and lower cost.

Like all prior DCIG Buyer's Guides, the *DCIG 2016-17 Midmarket Enterprise Storage Array Buyer's Guide* does the heavy lifting for organizations as they look to purchase a midmarket enterprise storage array by:



- Delineating the storage array features that are supported
- Weighting these features according to what end users consider most important
- Ranking each product
- Creating a standardized one-page data sheet for each product

The end result is that the *DCIG 2016-17 Midmarket Enterprise Storage Array Buyer's Guide* drives time and cost out of the product selection process by enabling prospective buyers to do "at-a-glance" comparisons between many different arrays. The standardized one-page data sheets make it easy to do quick, side-by-side comparisons of midmarket storage arrays so organizations may quickly arrive at a short list of products that may meet their requirements.

The *DCIG 2016-17 Midmarket Enterprise Storage Array Buyer's Guide* is available immediately through the DCIG Analysis Portal for subscribing users at <https://portal.dcig.com>. End users new to the DCIG Analysis Portal may register using this [link](#) to access this Buyer's Guide.

End users registering to access this report via the DCIG Analysis Portal also gain access to the DCIG Interactive Buyer's Guide (IBG). The IBG enables organizations take the next step in the product selection process by generating custom reports, including comprehensive side-by-side feature comparisons of the arrays in which the organization is most interested.

---

# Server-based Storage Makes Accelerating Application Performance Insanely Easy

In today's enterprise data centers, when one thinks performance, one thinks flash. That's great. But that thought process can lead organizations to think that "all-flash arrays" are the only option they have to get high levels of performance for their applications. That thinking is now so outdated. The latest server-based storage solution from [Datrium](#) illustrates how accelerating application performance just became insanely easy by simply clicking a button versus resorting to upgrading some hardware in their environment.

As flash transforms the demands of application owners, organizations want more options to cost-effectively deploy and manage it. These include:

- Putting lower cost flash on servers as it performs better on servers than across a SAN.
- Hyper-converged solutions have become an interesting approach to server-based storage. However, concerns remain about fixed compute/capacity scaling requirements and server hardware lock-in.
- Array-based arrays have taken off in large part because they provide a pool of shared flash storage accessible to multiple servers.

Now a fourth, viable flash option has appeared on the market. While I have always had some doubts about server-based storage solutions that employ server-side software, today I changed my viewpoint after reviewing Datrium's DVX [Server-powered Storage System](#).

[Datrium](#) has the obvious advantages over arrays as it leverages the vast, affordable and often under-utilized server

resources. But unlike hyper-converged systems, it scales flexibly and does not require a material change in server sourcing.

To achieve this ends, [Datrium](#) has taken a very different approach with its “server-powered” storage system design. In effect, Datrium split speed from durable capacity in a single end-to-end system. Storage performance and data services tap host compute and flash cache, driven by Datrium software that is uploaded to the virtual host. It then employs its DVX appliance, an integrated external storage appliance, that permanently holds data and orchestrates the DVX system protects application data in the event of server or flash failure.

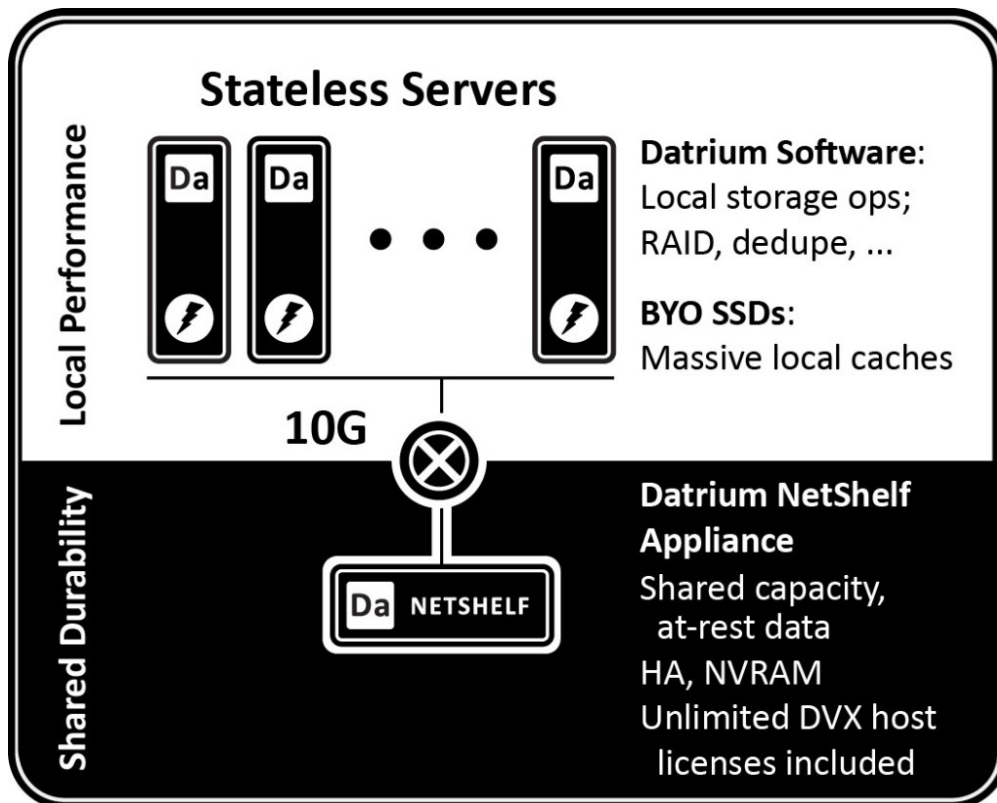
This approach has a couple meaningful takeaways versus traditional arrays:

- Faster flash-based performance given it is local to the server versus accessed across a SAN
- Lower cost since server flash drives cost far less than flash drives found on an all-flash array.

But it also addresses some concerns that have been raised about hyper-converged systems:

- Organizations may independently scale compute and capacity
- Plugs into an organization’s existing infrastructure.

## **Datrium Offers a New Server-based Storage Paradigm**



Source: Datrium

Datrium [DVX](#) provides the different approach needed to create a new storage paradigm. It opens new doors for organizations to:

1. **Leverage excess CPU cycles and flash capacity on ESX servers.** ESX servers now exhibit the same characteristics that the physical servers they replaced once did: they have excess, idle CPU. By deploying server-based storage software at the hypervisor level, organizations can harness this excess, idle CPU to improve application performance.
2. **Capitalize on lower-cost server-based flash drives.** Regardless of where flash drives reside (*server-based or array-based,*) they deliver high levels of performance. However, server-based flash costs much less than array-based flash while providing greater flexibility to add more capacity going forward.

**Accelerating Application Performance Acceleration Just Became Insanely Easy**

Access to excess server-based memory, CPU and flash combine to offer another feature that array-based flash can never deliver: push button application performance. By default, when the Datrium storage software installs on ESX hypervisor, it limits itself to 20 percent of the available vCPU available to each VM. However, not every VM uses all of its available vCPU with many VMs only using only 10-40 percent of their available resources.

Using Datrium's [DIESL Hyperdriver Software](#) version 1.0.6.1, VM administrators can non-disruptively tap into these latent vCPU cycles. Using Datrium's new Insane Mode, they may increase the available vCPU cycles a VM can access from 20 to 40 percent with a click of a button. While the host VM must have latent vCPU cycles available to accomplish this task, this is a feature that array-based flash would be hard-pressed to ever offer and unlikely could ever do with the click of a button.

Server-based storage designs have shown a lot of promise over the years but have not really had the infrastructure available to them to build a runway to success. That has essentially changed and Datrium is one of the first solutions to come to market that recognizes this fundamental change in the infrastructure of data centers and has brought a product to market to capitalize on it. As evidenced by the Insane Mode in its latest software release, organizations may now harness next generation server-based storage designs and accelerate application performance while dramatically lowering complexity and costs in their environment.

---

# Fibre Channel (FC) HBAs Will Not Be Embedded on Server Motherboards Anytime Soon; Interview with QLogic's Vikram Karvat, Part 2

Ethernet adapters began migrating to LAN on motherboard solutions in the late 1990s. Yet this practice never took hold for other technologies like Fibre Channel. The Fibre Channel (FC) market even today, as Gen 6 (32Gb) is being introduced, is dominated by host bus adapters (HBAs). In this second installment in my interview with QLogic's Vice President of Products, Marketing and Planning, Vikram Karvat, he explains why 32Gb FC HBAs are still installed separately in servers, as well provides insight into what new features may be released in the Gen 7 FC protocol.



[QLogic 2700](#) HBA; Source: QLogic

**Jerome:** *Are the new QLogic 32Gb FC HBAs embedded in the server and/or storage array mother boards? If not, are there any plans to do so?*

**Vikram:** The HBAs being discussed here are pretty much entirely add in cards on the server side. There are no embedded FC HBAs on servers. As a result, the FC HBA port counts analysts report represent not only the ports that are shipped from the vendors, but ports that are actually being deployed for use on

an annual basis. This represents as close to a natural demand in any market as you could hope to measure.

**Jerome:** *Why haven't FC HBAs gone to being embedded?*

**Vikram:** A network card typically goes embedded when it hits north of 50 percent connectivity. To get to north of 50 percent for FC, you would probably have to quintuple its volume. It's a different set of economics.

We previously talked a little bit about the increased use of FC on the all-flash array (AFA) side, but [QLogic](#) is also actually seeing an increase in use and deployment of FC SANs in emerging markets like China. FC SAN deployments in China grew by 15 percent last year. That is huge and the growth rate has been like that for probably the last two to three years. Obviously, in the early years, growing even faster than that, but from a relatively modest base.

But it's no longer just a modest base anymore. It's significant at a global scale in terms of how many SANs are being deployed. Again, not to the same scale as in North America. But nonetheless, it's measurable and is making an impact towards keeping the market relatively stable.

From a use case perspective, it's interesting because it's a market that tends not to want to spend money on something unless it's absolutely necessary. It's an indicator of the stability of the FC market. And FC remains the predominant storage interconnect for storage arrays, as well as, servers. There are areas of growth like AFAs and emerging markets. All in all, FC is not a bad story. FC offers the availability, reliability, security, and lossless fabric that enterprises want.

Further, there is a lot of discussion about Remote Direct Memory Access (RDMA) and storage options with very low CPU utilization, etc. But FC has always been a fully offloaded architecture with ultra-low CPU utilization (*in the single*

*digits*) which is why it is used for Online Transaction Processing (OLTP) types of infrastructures and it has always been [zero copy](#) (*i.e. – does not require the CPU to perform the task of copying data from one memory area to another.*)

The notion that there are new storage networking implementations out there that are more efficient is potentially a bit of a fallacy there. FC, as an industry, has not made a big deal out of its strengths because the industry just assumed everybody understood these concepts. We are having to remind people now.

**Jerome:** *As FC is so mature and stable, what innovation is occurring?*

**Vikram:** There a number of areas of innovation where the industry is investing. Obviously [Gen 6](#) FC is good. Moving forward, the FC industry is actually in the process of defining Gen 7 FC as a next step up. Layering on to that, we are innovating in the flash space with Fibre Channel over Non-Volatile Memory Express (FC-NVMe).

FC-NVMe is an industry initiative to directly map the NVMe drive over a fabric. Why you would ask? The normal reasons why you map something over a fabric is for the ability to share, create pools, provision, and manage storage more effectively when it is connected, as opposed to having islands of flash floating around in servers.

The unique thing about FC-NVMe is that instead of using the standard SCSI stack, it actually bypasses the SCSI stack and uses native NVMe semantics to reduce both the latency of the access and the CPU associated with the SCSI infrastructure on both the storage and server sides.

You are effectively taking a technology that was initially very focused on driving latency and performance within a server and extending it out of the box to get some of these additional benefits. We were recently demonstrated the ability



to run FC-NVMe, as well as traditional FCP traffic, simultaneously, on existing fabrics.

When we talk about developing a new technology, it's like, "Hey, here's my new thing. Oh by the way to get this, you have to go buy a whole bunch of new stuff."

What QLogic is doing is layering this functionality onto the infrastructure that's already in place. It effectively comes for free.

We are pretty excited about that. We got a lot of interest from our OEM customers. I suspect that over the course of the next year, as this technology starts getting in front of end users via our OEM customers that our OEM customers will find it even more attractive. Again there's everything to gain and nothing to lose.

*In [Part I](#) of this series, we took a look at why all-flash arrays are driving the need for 32Gb fibre channel.*

*In the third and final of this interview series, Vikram reveals what new FC HBA features service providers are most eager to see and use.*

---

## **All-Flash Arrays Driving Need for 32Gb Fibre Channel; Interview with QLogic's Vikram Karvat, Part I**

All-flash arrays, cloud computing, cloud storage, and converged and hyper-converged infrastructures may grab many of

today's headlines. But the decades old Fibre Channel protocol is still a foundational technology present in many data centers with it holding steady in the U.S. and even gaining increased traction in countries such as China. In this first installment, QLogic's Vice President of Products, Marketing and Planning, Vikram Karvat, provides some background as to why fibre channel (FC) remains relevant and how all-flash arrays are one of the forces driving the need for 32Gb FC.



**Jerome:** *Vik, thanks for taking time out of your schedule to share a bit about 32Gb fibre channel. Before we begin, for the benefit of DCIG's readers, can you share a bit about QLogic and what has been going on over there for the past few years?*

**Vikram:** Thanks, Jerome. Many of your readers are probably familiar with [QLogic](#) from the fibre channel side as it has continued to be a preeminent player in that space. However, QLogic has had a few changes in the last few years.

Mostly notably, QLogic [acquired](#) Brocade's Fibre Channel HBA assets about two years ago. As a result of concluding that transaction in early 2014, QLogic was able to move that relationship to a new level in terms of technical cooperation,

alignment on road maps and technologies, etc.

The other significant change was that QLogic [acquired](#) Broadcom's Ethernet controller assets. QLogic already had its own portfolio of Ethernet controllers with which it had been relatively successful on the host side, and very, very successful on the storage side; but the Broadcom assets brought a different level of scale to our overall Ethernet portfolio and immediately put QLogic in a very, very strong number two position in Ethernet.

The net net is that today QLogic has the number one position in Fibre Channel and the number two position in the 10Gb Ethernet on the host/server side of the business. This is important because it allows QLogic to look at certain types of technology that would benefit from end-to-end integration. It also has some interesting benefits as QLogic moves forward.

**Jerome:** *Tell me about 32Gb Fibre Channel. What is happening on that front?*

**Vikram:** The next instantiation of the Fibre Channel roadmap is Gen 6 (32Gb) Fibre Channel (FC) which QLogic is releasing today. A lot of people ask me, *"Why do you need Gen 6 FC? Do we need more performance?"*

There is always some of that. You do need more performance to support today's latest technologies, such as multi-core processors and multichannel memory on servers, but then you also have the move towards non-volatile storage in servers, as well as in the storage arrays. Further, databases just keep getting bigger and bigger and the response time requirements for accessing these content repositories keeps getting shorter. Gen 6 FC performance advantages play directly into all of these demands from both a bandwidth and an IOPS perspective.

But there's more than just performance advantages with the shift to Gen 6 FC. IT organizations are under tremendous OPEX

pressure. They need to maintain service-level agreements (SLAs) but with fewer people so they have to find ways to work more efficiently. Further, they are under pressure to increase scalability and deliver faster provisioning of new storage on demand.

This is where some of the features and functions that QLogic offers with its new Gen 6 FC adapters deliver as much value and, in some cases, maybe even more value than the performance benefits of Gen 6 FC.

***Jerome:** Isn't QLogic introducing new technology and innovating in a market that is in decline?*

***Vikram:*** There has been a general sense in the industry that Fibre Channel is on a steep decline. I would propose to you today that that may not be entirely true. It's certainly not the growing market that it was a decade ago, but it's not ending any time soon.

The data points here just serve to underscore that. On the external block-based storage side, a lot of Fibre Channel connectivity has actually gone up in terms of a mix of total ports. Some of that is driven by the still significant need for Fibre Channel in traditional arrays.

Some of this demand is also being driven by all-flash arrays. Almost 80 percent of these are connected via Fibre Channel. Then, if you look at Fibre Channel just in raw terms of how many Fibre Channel ports there are per unit of storage capacity, it's actually higher on all-flash arrays than it is on traditional storage arrays, just because of the performance levels associated with flash.

The result is that we've actually seen a slight uptick over the last three years in overall mix of Fibre Channel connectivity on external storage controllers. The actual number of port shipments has been holding steady for the last couple of years. We expect the same to hold true for 2016,

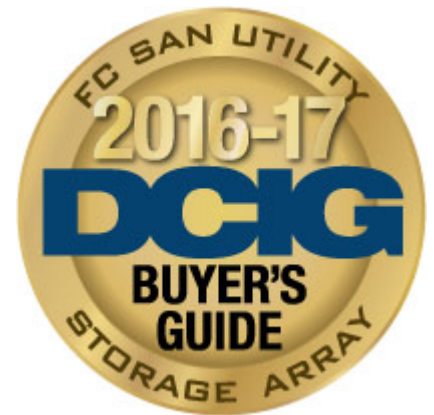
with just slightly north of two million ports of server side HBA connectivity. Again, this might take some people by surprise because there's been the general sense that the market has been in decline, but the numbers actually show that from a standard HBA perspective, it's pretty stable.

*In Part II of this interview series, Vikram shares his thoughts about industry initiatives to directly map the NVMe drive over Fibre Channel fabric.*

---

## **DCIG 2016-17 FC SAN Utility Storage Array and Utility SAN Storage Array Buyer's Guides Now Available**

DCIG is pleased to announce the availability of its **2016-17 FC SAN Utility Storage Array Buyer's Guide** and **2016-17 Utility SAN Storage Array Buyer's Guide** that each weight more than 100 features and rank 62 arrays from thirteen (13) different storage providers. These Buyer's Guide Editions are products of DCIG's updated research methodology where DCIG creates specific Buyer's Guide Editions based upon a larger, general body of research on a topic. As past Buyer's Guides have done, it continues to rank products as Recommended, Excellent, Good and Basic as well as offer the product information that organizations need to make informed buying decisions on FC SAN Utility and multiprotocol Utility SAN storage arrays.



Over the years organizations have taken a number of steps to better manage the data that they already possess as well as prepare themselves for the growth they expect to experience in the future. These steps usually involve either deleting data that they have determined they do not need or should not keep while archiving the rest of it on a low cost media such as optical, tape or even with public cloud storage providers.

Fibre Channel (FC) and multiprotocol SAN storage arrays configured as utility storage arrays represent a maturation of the storage array market. Storage arrays using hard disk drives (HDDs) are still the predominant media used to host and service high performance applications. But with the advent of flash and solid state drives (SSDs), this reality is rapidly changing. Flash-based arrays are rapidly supplanting all-HDD storage arrays to host business-critical, performance sensitive applications as flash-based arrays can typically provide sub-two millisecond read and write response times.

However, the high levels of performance these flash-based arrays offer comes with a price – up to 10x more than all HDD-based utility storage arrays. This is where HDD-based arrays in general, and SAN utility storage arrays in particular, find a new home. These array may host and service applications with infrequently accessed or inactive data such as archived, backup and file data.

Many if not most organizations still adhere to a “*keep it all forever*” mentality when it comes to managing data for various

reasons. These factors have led organizations to adopt a “*delete nothing*” approach to managing their data as this is often their most affordable and prudent option. The challenge with this technique is that as data volumes continue to grow and retention periods remain non-existent, organizations need to identify solutions on which they can affordably store all of this data.

Thanks to the continuing drop per GB in disk’s cost that day has essentially arrived. The emergence of highly available and reliable utility storage arrays that scale into the petabytes at a cost of well below \$1/GB opens the doors for organizations to confidently and cost-effectively keep almost any amount of data online and accessible for their business needs.

Utility storage arrays also offer low millisecond response times (8 – 10 ms) for application reads and writes. This is more than adequate performance for most archival or infrequently accessed data. These arrays deliver millisecond response times while supporting hundreds of terabytes if not petabytes of storage capacity at under a dollar per gigabyte.

The 2016-17 FC SAN Utility Storage Array Buyer’s Guide specifically covers those storage arrays that support the Fibre Channel storage networking protocol. The 2016-17 Utility SAN Storage Array Buyer’s Guide scores the arrays for their support for both FC and iSCSI storage networking protocols. All of the included utility storage arrays are available in highly available, reliable configurations and list for \$1/GB or less. While the arrays in this Guide may support other storage networking protocols, other specific protocols were not weighted in arriving in the conclusions in these Buyer’s Guide Editions.

DCIG’s succinct analysis provides insight into the state of the SAN utility storage array marketplace. It identifies the significant benefits organizations can expect to realize by

implementing a utility storage array, key features that organizations should evaluate on these arrays and includes brief observations about the distinctive features of each array. The storage array rankings provide organizations with an “at-a-glance” overview of this marketplace. DCIG complements these rankings with standardized, one-page data sheets that facilitate side-by-side product comparisons so organizations may quickly get to a short list of products that may meet their requirements.

Registration to access these Buyer’s Guides may be done via the DCIG Analysis Portal which includes access to DCIG Buyer’s Guides in PDF format as well as the DCIG Interactive Buyer’s Guide (IBG). Using the IBG, organizations may dynamically drill down and compare and contrast FC SAN and Utility SAN arrays by generating custom reports, including comprehensive strengths and weaknesses reports that evaluate a much broader base of features than what is found in the published Guide. Both the IBG and this Buyer’s Guide may be accessed after [registering](#) for the [DCIG Analysis Portal](#).

---

## **DCIG 2016-17 iSCSI SAN Utility Storage Array Buyer’s Guide Now Available**

DCIG is pleased to announce the availability of its **2016-17 iSCSI SAN Utility Storage Array Buyer’s Guide** that weights more than 100 features and ranks 67 arrays from fourteen (14) different storage providers. This Buyer’s Guide Edition reflects the first use of DCIG’s updated research methodology



where DCIG creates specific Buyer's Guide Editions based upon a larger, general body of research on a topic. As past Buyer's Guides have done, it continues to rank products as Recommended, Excellent, Good and Basic as well as offer the product information that organizations need to make informed buying decisions on iSCSI SAN utility storage arrays.



Over the years organizations have taken a number of steps to better manage the data that they already possess as well as prepare themselves for the growth they expect to experience in the future. These steps usually involve either deleting data that they have determined they do not need or should not keep while archiving the rest of it on a low cost media such as optical, tape or even with public cloud storage providers.

iSCSI SAN storage arrays configured as utility storage arrays represent a maturation of the storage array market. Storage arrays using hard disk drives (HDDs) are still the predominant media used to host and service high performance applications. But with the advent of flash and solid state drives (SSDs), this reality is rapidly changing. Flash-based arrays are rapidly supplanting all-HDD storage arrays to host business-critical, performance sensitive applications as flash-based arrays can typically provide sub-two millisecond read and write response times.

However the high levels of performance these flash-based arrays offer comes with a price – up to 10x more than all HDD-based utility storage arrays. This is where HDD-based arrays

in general, and iSCSI SAN utility storage arrays in particular, find a new home. These array may host and service applications with infrequently accessed or inactive data such as archived, backup and file data.

Many if not most organizations still adhere to a “*keep it all forever*” mentality when it comes to managing data for various reasons. These factors have led organizations to adopt a “*delete nothing*” approach to managing their data as this is often their most affordable and prudent option. The challenge with this technique is that as data volumes continue to grow and retention periods remain non-existent, organizations need to identify solutions on which they can affordably store all of this data.

Thanks to the continuing drop per GB in disk’s cost that day has essentially arrived. The emergence of highly available and reliable iSCSI SAN utility storage arrays that scale into the petabytes at a cost of well below \$1/GB opens the doors for organizations to confidently and cost-effectively keep almost any amount of data online and accessible for their business needs.

iSCSI SAN utility storage arrays also offer low millisecond response times (8 – 10 ms) for application reads and writes. This is more than adequate performance for most archival or infrequently accessed data. These arrays deliver millisecond response times while supporting hundreds of terabytes if not petabytes of storage capacity at under a dollar per gigabyte.

This Buyer’s Guide edition specifically covers those storage arrays that support the iSCSI storage networking protocol, are available in highly available, reliable configurations and list for \$1/GB or less. While the arrays in this Guide may and often do support other storage networking protocols, those specific protocols were not weighted in arriving in the conclusions in this Buyer’s Guide Edition

DCIG's succinct analysis provides insight into the state of the iSCSI SAN utility storage array marketplace. It identifies the significant benefits organizations can expect to realize by implementing an iSCSI SAN utility storage array, key features that organizations should evaluate on these arrays and includes brief observations about the distinctive features of each array. The iSCSI SAN utility storage array ranking provide organizations with an "at-a-glance" overview of this marketplace. DCIG complements these rankings with standardized, one-page data sheets that facilitate side-by-side product comparisons so organizations may quickly get to a short list of products that may meet their requirements.

Registration to access this Buyer's Guide may be done via the DCIG Analysis Portal which includes access to DCIG Buyer's Guides in PDF format as well as the DCIG Interactive Buyer's Guide (IBG). Using the IBG, organizations may dynamically drill down and compare and contrast iSCSI SAN arrays by generating custom reports, including comprehensive strengths and weaknesses reports that evaluate a much broader base of features that what is found in the published Guide. Both the IBG and this Buyer's Guide may be accessed after [registering](#) for the [DCIG Analysis Portal](#).

---

## **HP 3PAR StoreServ 8000 Series Lays Foundation for Flash Lift-off**

Almost any hybrid or all-flash storage array will accelerate performance for the applications it hosts. Yet many organizations need a storage array that scales beyond just

accelerating the performance of a few hosts. They want a solution that both solves their immediate performance challenges and serves as a launch pad to using flash more broadly in their environment.

Yet putting flash in legacy storage arrays is not the right approach to accomplish this objective. Enterprise-wide flash deployments require purpose-built hardware backed by Tier-1 data services. The HP 3PAR [StoreServ 8000 series](#) provides a fundamentally different hardware architecture and complements this architecture with mature software services. Together these features provide organizations the foundation they need to realize flash's performance benefits while positioning them to expand their use of flash going forward.

## **A Hardware Foundation for Flash Success**

Organizations almost always want to immediately realize the performance benefits of flash and the HP 3PAR [StoreServ 8000 series](#) delivers on this expectation. While flash-based storage arrays use various hardware options for flash acceleration, the 8000 series complements the enterprise-class flash HP 3PAR StoreServ 20000 series while separating itself from competitive flash arrays in the following key ways:

- **Scalable, Mesh-Active architecture.** An Active-Active controller configuration and a scale-out architecture are considered the best of traditional and next-generation array architectures. The HP 3PAR StoreServ 8000 series brings these options together with its Mesh-Active architecture which provides high-speed, synchronized communication between the up-to-four controllers within the 8000 series.
- **No internal performance bottlenecks.** One of the secrets to the 8000's ability to successfully transition from managing HDDs to SSDs and still deliver on flash's performance benefits is its programmable ASIC. The HP 3PAR ASIC, now it's 5<sup>th</sup> generation, is programmed to

manage flash and optimize its performance, enabling the 8000 series to achieve **over 1 million IOPs**.

- **Lower costs without compromise.** Organizations may use lower-cost commercial MLC SSDs (cMLC SSDs) in any 8000 series array. Then leveraging its Adaptive Sparing technology and Gen5 ASIC, it optimizes capacity utilization within cMLC SSDs to achieve high levels of performance, extends media lifespan which are backed by a 5-year warranty, and increases usable drive capacity by up to 20 percent.
- **Designed for enterprise consolidation.** The 8000 series offers both 16Gb FC and 10Gb Ethernet host-facing ports. These give organizations the flexibility to connect performance-intensive applications using Fibre Channel or cost-sensitive applications via either iSCSI or NAS using the 8000 series' File Persona feature. Using the 8000 Series, organizations can start with configurations as small as 3TB of usable flash capacity and scale to 7.3TB of usable flash capacity.

## **A Flash Launch Pad**

As important as hardware is to experiencing success with flash on the 8000 series, HP made a strategic decision to ensure its converged flash and all-flash 8000 series models deliver the same mature set of data services that it has offered on its all-HDD HP 3PAR StoreServ systems. This frees organizations to move forward in their consolidation initiatives knowing that they can meet enterprise resiliency, performance, and high availability expectations even as the 8000 series scales over time to meet future requirements.

For instance, as organizations consolidate applications and their data on the 8000 series, they will typically consume less storage capacity using the 8000 series' native thin provisioning and deduplication features. While storage savings vary, HP finds these features usually result in about 4:1 data reduction ratio which helps to drive down the effective price

of flash on an 8000 series array to as low as \$1.50/GB.

Maybe more importantly, organizations will see minimal to no slowdown in application performance even as they implement these features, as they may be turned on even when running mixed production workloads. The 8000 series compacts data and accelerates application performance by again leveraging its Gen5 ASICs to do system-wide striping and optimize flash media for performance.

Having addressed these initial business concerns around cost and performance, the 8000 series also brings along the HP 3PAR StoreServ's existing data management services that enable organizations to effectively manage and protect mission-critical applications and data. Some of these options include:

- ***Accelerated data protection and recovery.*** Using HP's Recovery Manager Central ([RMC](#)), organizations may accelerate and centralize application data protection and recovery. RMC can schedule and manage snapshots on the 8000 series and then directly copy those snapshots to and from HP StoreOnce without the use of a third-party backup application.
- ***Continuous application availability.*** The HP 3PAR [Remote Copy](#) software either asynchronously or synchronously replicates data to another location. This provides recovery point objectives (RPOs) of minutes, seconds, or even non-disruptive application failover.
- ***Delivering on service level agreements (SLAs).*** The 8000 series' Quality of Service ([QoS](#)) feature ensures high priority applications get access to the resources they need over lower priority ones to include setting sub-millisecond response times for these applications. However QoS also ensures lower priority applications are serviced and not crowded out by higher priority applications.
- ***Data mobility.*** HP 3PAR StoreServ creates a federated storage pool to facilitate non-disruptive, bi-

directional data movement between any of up to four (4) midrange or high end HP 3PAR arrays.

## **Onboarding Made Fast and Easy**

Despite the benefits that flash technology offers and the various hardware and software features that the 8000 series provides to deliver on flash's promise, migrating data to the 8000 series is sometimes viewed as the biggest obstacle to its adoption. As organizations may already have a storage array in their environment, moving its data to the 8000 series can be both complicated and time-consuming. To deal with these concerns, HP provides a relatively fast and easy process for organizations to migrate data to the 8000 series.

In as few as five steps, existing hosts may discover the 8000 series and then access their existing data on their old array through the 8000 series without requiring the use of any external appliance. As hosts switch to using the 8000 series as their primary array, Online Import non-disruptively copies data from the old array to the 8000 series in the background. As it migrates the data, the 8000 series also reduces the storage footprint by as much as 75 percent using its thin-aware functionality which only copies blocks which contain data as opposed to copying all blocks in a particular volume.

Maybe most importantly, data migrations from EMC, HDS or HP EVA arrays (*and others to come*) to the 8000 series may occur in real time. Hosts read data from volumes on either the old array or the new 8000 series with hosts only writing to the 8000 series. Once all data is migrated, access to volumes on the old array is discontinued.

## **Achieve Flash Lift-off Using the HP 3PAR StoreServ 8000 Series**

Organizations want to introduce flash into their environment but they want to do so in a manner that lays a foundation for their broader use of flash going forward without creating a new storage silo that they need to manage in the near term.

The HP 3PAR StoreServ 8000 series delivers on these competing requirements. Its robust hardware and mature data services work hand-in-hand to provide both the high levels of performance and Tier-1 resiliency that organizations need to reliably and confidently use flash now and then expand its use in the future. Further, they can achieve lift-off with flash as they can proceed without worrying about how they will either keep their mission-critical apps online or cost-effectively migrate, protect or manage their data once it is hosted on flash.

---

## **My Top 3 Reasons as to What Went Wrong with Appliance and Storage Array-based Storage Virtualization Solutions**

In the early 2000's I was a big believer in appliance and/or array-based storage virtualization technology. To me, it seemed like the most logical choice to solve some of the most pressing problems such as data migrations, storage optimization and reducing storage networking's overall management complexity that were confronting the deployment of storage networks in enterprise data centers. Yet here we find ourselves in 2015 and, while appliance and storage array-based storage virtualization still exists, it certainly never became the runaway success that many envisioned at the time. Here are my top 3 reasons as to what went wrong with this technology and why it has yet to fully realize its promise.

- 1. It did not and still does not sufficiently scale to meet***



**enterprise requirements.** The big appeal to me of storage virtualization appliances and/or array controllers was that they could aggregate all of an infrastructure's storage arrays and their capacity into one giant pool of storage which could then be centrally managed. As I came to learn, the problem with this philosophy was that none of the solutions could fully scale to manage all of the storage capacity in one's data center and certainly not in my data center.

In the early 2000's I was managing what seemed like an unimaginably large amount of storage capacity (*four storage arrays with over 11TBs of storage capacity.*) Even in that environment (*considered small by any of today's standards,*) the storage virtualization solution I brought in-house only scaled to manage 1TB of capacity. So instead of simplifying my environment and presenting me with only one storage console to manage, it become just another one to manage which increased my complexity rather than reducing it.

**2. Storage virtualization vendors failed to represent the capabilities of high end arrays.** One of the big claims that vendors of storage virtualization appliances made was that you could virtualize high end arrays such as the EMC Symmetrix, the IBM Shark or the HDS Tagmastore. This would eliminate or minimize the need to license their management software, increase application performance and simplify your environment even as you lowered storage costs.

While there was some merit in these claims, they failed to mention that by putting their virtualization appliance in front of these high end arrays you also lost some of the functionality of these high end arrays. For instance, if you had an Oracle Database that communicated directly with the front end controllers on an EMC Symmetrix for data management or performance reasons, that functionality largely went away once a storage virtualization appliance

was put in front of it. While some of those issues have been addressed in modern storage virtualization appliances, they have not and will likely never be fully addressed.

3. ***Heterogeneous environments were way more (and remain way more) complex than anyone likes to admit.*** Another claim made by storage virtualization appliances and/or controllers was the idea that organizations could connect any operating system to any storage array using any network interface and use these storage virtualization appliances to manage and move data between them. This would give organizations a great deal of flexibility to introduce any server and/or storage array onto their storage network and give them more negotiating power to boot.

What I came to learn and now fully understand is that this ideal is simply not an option in enterprise storage networking environments. These environments want tested and proven end-to-end configurations. In this way, if anything went wrong, there was a vendor that they could hold accountable to come in and fix the problem, end-to-end. When presented with this requirement, most storage virtualization vendors were typically unwilling to produce the certifications that illustrated such end-to-end interoperability or provide any guarantees that would resolve the issues that arose. In some cases, even if they did, no one believed they could deliver on them. As such, few enterprises were willing to bet on this type of technology.

Having provided these reasons as to why appliance and/or storage array-based storage virtualization failed to fully gain the wide adoption that many expected them to gain, one should not assume this technology has died off. If anything, storage vendors have learned these lessons with successful deployments of this technology now in place and I even see the adoption of this technology gaining some momentum. In an

upcoming blog entry, I will share some thoughts and tips as to why it is seeing a rebirth and ways in which to successfully deploy these solutions in today's data centers.

---

## My Top 3 Reasons as to What Went Wrong with Appliance and Storage Array-based Storage Virtualization Solutions

In the early 2000's I was a big believer in appliance and/or array-based storage virtualization technology. To me, it seemed like the most logical choice to solve some of the most pressing problems such as data migrations, storage optimization and reducing storage networking's overall management complexity that were confronting the deployment of storage networks in enterprise data centers. Yet here we find ourselves in 2015 and, while appliance and storage array-based storage virtualization still exists, it certainly never became the runaway success that many envisioned at the time. Here are my top 3 reasons as to what went wrong with this technology and why it has yet to fully realize its promise.

1. ***It did not and still does not sufficiently scale to meet enterprise requirements.*** The big appeal to me of storage virtualization appliances and/or array controllers was that they could aggregate all of an infrastructure's storage arrays and their capacity into one giant pool of storage which could then be centrally managed. As I came to learn, the problem with this philosophy was that none of the solutions could fully scale to manage all of the

storage capacity in one's data center and certainly not in my data center.

In the early 2000's I was managing what seemed like an unimaginably large amount of storage capacity (*four storage arrays with over 11TBs of storage capacity.*) Even in that environment (*considered small by any of today's standards,*) the storage virtualization solution I brought in-house only scaled to manage 1TB of capacity. So instead of simplifying my environment and presenting me with only one storage console to manage, it become just another one to manage which increased my complexity rather than reducing it.

**2. *Storage virtualization vendors failed to represent the capabilities of high end arrays.*** One of the big claims that vendors of storage virtualization appliances made was that you could virtualize high end arrays such as the EMC Symmetrix, the IBM Shark or the HDS Tagmastore. This would eliminate or minimize the need to license their management software, increase application performance and simplify your environment even as you lowered storage costs.

While there was some merit in these claims, they failed to mention that by putting their virtualization appliance in front of these high end arrays you also lost some of the functionality of these high end arrays. For instance, if you had an Oracle Database that communicated directly with the front end controllers on an EMC Symmetrix for data management or performance reasons, that functionality largely went away once a storage virtualization appliance was put in front of it. While some of those issues have been addressed in modern storage virtualization appliances, they have not and will likely never be fully addressed.

**3. *Heterogeneous environments were way more (and remain way more) complex than anyone likes to admit.*** Another claim made by storage virtualization appliances and/or

controllers was the idea that organizations could connect any operating system to any storage array using any network interface and use these storage virtualization appliances to manage and move data between them. This would give organizations a great deal of flexibility to introduce any server and/or storage array onto their storage network and give them more negotiating power to boot.

What I came to learn and now fully understand is that this ideal is simply not an option in enterprise storage networking environments. These environments want tested and proven end-to-end configurations. In this way, if anything went wrong, there was a vendor that they could hold accountable to come in and fix the problem, end-to-end. When presented with this requirement, most storage virtualization vendors were typically unwilling to produce the certifications that illustrated such end-to-end interoperability or provide any guarantees that would resolve the issues that arose. In some cases, even if they did, no one believed they could deliver on them. As such, few enterprises were willing to bet on this type of technology.

Having provided these reasons as to why appliance and/or storage array-based storage virtualization failed to fully gain the wide adoption that many expected them to gain, one should not assume this technology has died off. If anything, storage vendors have learned these lessons with successful deployments of this technology now in place and I even see the adoption of this technology gaining some momentum. In an upcoming blog entry, I will share some thoughts and tips as to why it is seeing a rebirth and ways in which to successfully deploy these solutions in today's data centers.

---

# **Four Early Insights from the Forthcoming DCIG 2015-16 Enterprise Midrange Array Buyer's Guide**

DCIG is preparing to release the DCIG 2015-16 Enterprise Midrange Array Buyer's Guide. The Buyer's Guide will include data on 33 arrays or array series from 16 storage providers. The term "Enterprise" in the name Enterprise Midrange Array, reflects a class of storage system that has emerged offering key enterprise-class features at prices suitable for mid-sized budgets.

In many businesses, there is an expectation that applications and their rapidly growing data will be available 24x7x365. Consequently, their storage systems must go beyond traditional expectations for scalable capacity, performance, reliability and availability. For example, not only must the storage system scale, it must scale without application downtime.

These expectations are not new to large enterprises and the high end storage systems that serve them. What is new is that these expectations are now held by many mid-sized organizations—the kind of organizations for which the products in this guide are intended.

While doing our research for the upcoming Buyer's Guide, DCIG has made the following observations regarding the fit between the expectations of mid-sized organizations and the features of the enterprise midrange arrays that will be included in the Buyer's Guide:

***Non-disruptive upgrades.*** In order to meet enterprises' expectations, storage systems must go beyond the old standard availability features like hot swap drives and redundant controllers to provide for uninterrupted operations even during storage system software and hardware upgrades. Consequently, this year's guide evaluates multiple NDU features and puts them literally at the top of the list on our data sheets. Over one third of the Enterprise Midrange Arrays support non-disruptive upgrade features.

***Self-healing technologies.*** While self-healing features are relatively new to midrange storage arrays, these technologies help an array achieve higher levels of availability by enabling the array to detect and resolve certain problems quickly, and with no or minimal human intervention.

Self-healing technologies have been implemented by some storage vendors, but these are seldom mentioned on product specification sheets. DCIG attempted to discover which arrays have implemented self-healing technologies such as bad block repair, failed disk isolation, low-level formatting and power cycling of individual drives; but we suspect (and hope) that more arrays have implemented self-healing capabilities than we were able to confirm through our research.

***Automation.*** Data center automation is an area of growing emphasis for many organizations because it promises to reduce the cost of data center management and enable IT to be more agile in responding to changing business requirements. Ultimately, automation means more staff time can be spent addressing business requirements rather than performing routine storage management tasks.

Organizations can implement automation in their environment through management interfaces that are scriptable or through APIs and SDKs provided by storage vendors. Last year's Enterprise Midrange Array Buyer's Guide prediction that 'support for automated provisioning would improve in the near

future' was correct. While less than 20% of midrange arrays in last year's Buyer's Guide exposed an API for third-party automation tools, the percentage has more than doubled to 50% in this year's guide. Provision of an SDK for integration with management platforms saw a similar increase, rising from 11% to 25%.

***Multi-vendor virtualization.*** A growing number of organizations are embracing a multi-vendor approach to virtualization. Reflecting this trend, support for Microsoft virtualization technologies is gaining ground on VMware among enterprise midrange arrays.

The percentage of arrays that can be managed from within Microsoft's System Center Virtual Machine Manager (SCVMM) now matches vSphere/vCenter support at 33%. Support for Microsoft Windows Offloaded Data Transfer (ODX), a Windows Server 2012 technology that enhances array throughput, is now at 19%.

Although the gap between Microsoft and VMware support is narrowing, support for VMware storage integrations also continues to grow. VAAI 4.1 is supported by 90% of the arrays, while SIOC, VASA and VASRM are now supported by over 50% of the arrays.

The *DCIG 2015-16 Enterprise Midrange Array Buyer's Guide* will provide organizations with a valuable tool to cut time and cost from the product research and purchase process. DCIG looks forward to providing prospective storage purchasers and others with an interest in the storage marketplace with this tool in the very near future.

---



# 10 Characteristics That Help to Define Today's High End Storage Arrays

It has been said that everyone knows what "*normal*" is but that it is often easier to define "*abnormal*" than it is to define "*normal*." To a certain degree that axiom also applies to defining "high end storage arrays." Everyone just seems to automatically assume that a certain set of storage arrays are in the "*high end*" category but when push comes to shove, people can be hard-pressed to provide a working definition as to what constitutes a high end storage array in today's crowded storage space.

Over the last few weeks the analysts at DCIG have certainly wrestled with some of those same issues regarding the definition of a high end storage array. Whereas the highest levels of availability, capacity and performance were once the defining attributes of these arrays, the providers of these arrays can no longer claim that they exclusively deliver these features. Many storage arrays classified as "*enterprise midrange*" or "*midrange*" offer similar or even higher levels of availability, capacity and performance than the storage arrays typically classified as "*high end*."

This is not to imply that a high end class of arrays does not exist. Such arrays do exist and it is important that organizations and enterprises recognize these arrays for what they are. However the features or characteristics that make them "*high end*" may, in some cases, differ from even a few years ago. To shed some light on what makes these storage arrays "*high end*," DCIG has come up with 10 characteristics that organizations should look for to distinguish between an array that is "*high end*" and one that is "*midrange*."

1. **FICON connectivity to an IBM mainframe.** In talking to a number of end users, VARs and vendors, FICON connectivity to IBM mainframes running z/OS is often where the difference between mainframe and midrange begins and ends. In short, if it does not offer FICON connectivity to a mainframe, it is not a high end storage array.
2. **Fibre Channel (FC) block-based storage connectivity.** Absent FICON connectivity, the storage array must minimally offer block-based FC connectivity to even have a shot at being considered a high end storage array. While a number of storage arrays considered high end may support Ethernet block-based protocols such as iSCSI or FCoE (Fibre Channel over Ethernet,) support for these protocols alone are not enough to bridge the midrange to high end gulf.
3. **Multiple Active-Active controller/blade/processor pairs.** A number of midrange arrays offer an "Active-Active" controller configuration where a pair of controllers permits concurrent access to data on the same backend disk. What differentiates a high end array from a midrange array is the availability of multiple pairs of these Active-Active controllers (also called "*blade pairs*" or "*processor pairs*" on some arrays) on the same physical array that are all part of the same logical array configuration.
4. **High levels of cache and capacity.** Despite the encroachment on this territory by multiple midrange arrays, high end storage arrays as a group still generally support far higher levels of cache and storage capacity than most midrange arrays. One should generally expect the amount of cache available on a high end storage array to scale into the hundreds if not thousands of GBs and provide support for PBs of storage capacity.
5. **Large number of multi-core processors.** The multiple blade/controller/processor pairs in a high end storage

array deliver much more than high availability. They also provide access to much higher levels of performance. This becomes critically important in environments that are handling mixed workloads that may include sequential reads, sequential writes and random access, small block transactions.

6. ***Scale-out and scale-up configurations.*** Midrange array providers often tout the scale-out or scale-up capabilities of their arrays like they are the best thing since sliced bread. High end storage providers tend to yawn, stretch and say, *"It is about time you offer those features on your array."* In other words, scale-out and scale-up are part and parcel to the configuration of every high end storage array.
7. ***Detailed system analysis, performance monitoring and troubleshooting.*** High end storage arrays give organizations unparalleled flexibility to gather and analyst system data. This may then be used to quickly, accurately and confidently pinpoint where a performance bottleneck is occurring or what piece of hardware inside of the storage array is malfunctioning. Most midrange storage arrays do not offer this level of diagnostics or capabilities to troubleshoot a performance or system issue.
8. ***Tested, certified configurations.*** While midrange array also "certify" their arrays with certain OSes and applications, the certification process in my mind for midrange arrays has always been a little suspect. This concern stems from the large number of applications and operating systems for which midrange arrays must be certified and the diverse environments into which they are deployed. Due to the smaller number of application- and OS-specific environments into which high end storage arrays are deployed, the level of confidence that enterprises may have about the quality and thoroughness of the interoperability testing and the quality of the features available can be higher.

9. **Starting list price of \$250,000 or higher.** All of these features, high levels of capacity and performance and certifications come at a price. While these high end storage arrays may actually be price competitive on a per GB basis with some midrange arrays, you first need an environment that justifies the scale that these high end arrays bring to the table.
  10. **Non-disruptive operations across two or more data centers.** Many storage arrays offer one or more forms of replications. But what is arguably becoming a defining feature on high end arrays is their ability to deliver synchronous replication to at least two storage arrays and then sync the applications (think VMs) with the underlying replication activities so as to guarantee non-disruptive operation of applications. While this feature was initially designed to deliver disaster recovery, more enterprises are looking to leverage this capability for load balancing, non-disruptive failovers and failbacks and even to lower their data center operating costs.
- 

## **The Challenges of Delivering Inline Deduplication on a High Performance Production Storage Array**

The use of data reduction technologies such as compression and deduplication to reduce storage costs are nothing new. Tape drives have used compression for decades to increase backup data densities on tape while many modern deduplicating backup

appliances use compression and deduplication to also reduce backup data stores. Even a select number of existing HDD-based storage arrays use data compression and deduplication to minimize data stores for large amounts of file data stored in archives or on networked attached file servers.

The challenges of using these two technologies change when they are implemented in high performance environments. The more predictable data access patterns with lots of redundant data that exist in archive, backup, and, to some extent, file serving environments are replaced in high performance environments with applications that potentially have highly random data access patterns where data does not deduplicate as well. Capacity reductions of production data are not as significant (maybe in the 2-3x range) as in backup which can achieve deduplication ratios of up to 8-20X or even higher.

Aggravating the situation, there is little to no tolerance for performance interruptions in the processing of production data – raw or deduplicated. While organizations may tolerate the occasional slow periods of deduplication performance for archive, backup and file servers data stores, consistently high levels of application performance with no interruptions are the expectations here.

Yet when it comes to deduplicating data, there is a large potential for a performance hit. In high performance production environments with high data change rates and few or no periods of application inactivity, all deduplication must be done inline. This requires the analysis of incoming data by breaking packets of data apart into smaller chunks, creating a hash and comparing that hash to existing hashes in the deduplication metadata database to determine if that chunk of data is unique or a duplicate.

If the array determines a chunk of data is a duplicate, there is also a very small chance that a hash collision could occur. Should the all-flash array fail to detect and appropriately

handle this collision, data may be compromised.

These expectations for high levels of data integrity and performance requires large amount of cache or DRAM to host the deduplication metadata. Yet all-flash storage arrays only contain fixed amounts of DRAM. This may limit the maximum amount of flash storage capacity on the array as it makes no sense for the array to offer flash storage capacity beyond the amount of data that it can effectively deduplicate.

These all-flash array capacity limits are reflected in the results of the most recent DCIG 2014-15 Flash Memory Storage Array [Buyer's Guide](#). Of the 36 all-flash array models evaluated, only 42 percent of them could scale to 100 TB or more of flash capacity. Of these models that could scale to more than 100 TB, they:

- Did not support the use of data deduplication at the time the Guide was published
- Did not publicly publish any performance data with deduplication turned “*On*” implying that they recommend turning deduplication “*Off*” when hosting performance sensitive applications
- Use scale-out architectures with high node counts (up to 100) that are unlikely to be used in most production environments

The need to scale to 100 TBs or more of flash storage capacity is quickly becoming a priority. HP reports that already 25% of its HP 3PAR StoreServ 7450 all-flash arrays ship with 80TBs or more of raw capacity as its customers want to move more than just their high performance production data from HDDs to flash. They want to store all of their production data on flash. Further, turning deduplication off for any reason when hosting high performance application on these arrays is counter intuitive since these arrays are specifically designed and intended to host high performance applications. This is why, as organizations look to acquire all-flash storage arrays

to host multiple applications in their environment, they need to look at how well they optimize both capacity and performance to keep their costs under control.

---

## Five Technologies that Companies Should Prioritize in 2014

One of the more difficult tasks for anyone deeply involved in technology is the ability to see the forest from the trees. Often responsible for supporting the technical components that make up today's enterprise infrastructures, to step back and recommend which technologies are the right choices for their organization going forward is a more difficult feat. While there is no one right answer that applies to all organizations, five (5) technologies – some new as well as some old technologies that are getting a refresh – merit that organizations prioritize them in the coming months and years.

Already in 2014 DCIG has released three [Buyer's Guides](#) and has many more planned for release in the coming weeks and months. While working on those Guides, DCIG has also engaged with multiple end-users to discuss their experiences with various technologies and how they prioritize technology buying decisions. This combination of sources – a careful examination of included features on products coupled with input from end-users and vendors – is painting a new picture of five specific technologies that companies should examine and prioritize in their purchasing plans going forward.

- **Backup software with a recovery focus.** Survey after survey shows that backup remains a big issue in many

organizations. However I am not sure who is conducting these surveys or who they are surveying because I now regularly talk to organizations that have backup under control. They have largely solved their ongoing backup issues by using new or updated backup software that is better equipped to use disk as the primary backup target.

It is as they adopt this new backup software and eliminate their backup problems, their focus is turning to recovery. A good example is an individual with whom I spoke this past week. He switched to a new backup software solution that solved his organization's long standing backup issues while enabling it to lay a foundation for application recovery to the cloud.

- ***Converged infrastructures.*** Converged infrastructure solutions are currently generating a great deal of interest as they eliminate much of the time and effort that organizations have to internally exert to configure, deploy and support a solution. However in conversations I have had over the last few weeks and months, it is large organizations that appear to be the most apt to deploy them.
- ***Heterogeneous infrastructures.*** Heterogeneous infrastructures were all the rage for many years among all size organizations as they got IT vendors to compete on price. But having too many components from too many providers created too much complexity and resulting administrative costs— especially in large organizations.

That said, small and mid-sized businesses (SMBs) with smaller IT infrastructures still have the luxury of acquiring IT gear from multiple providers without resulting in their environments becoming too complex to manage. Further, SMBs remain price conscious. As such, they are more willing to sacrifice the notion of “proven” end-to-end configurations to get the cutting edge features and/or the lower prices



that heterogeneous infrastructures are more apt to offer.

- **Flash primed to displace more HDDs.** Those close to the storage industry recognize flash for the revolutionary technology that it is. However I just spoke to an individual this past week that is very technical but who has a web design and programming focus so he and his company were not that familiar with flash. He said that as they have learned more about it, they are re-examining their storage infrastructure and how and where they can best deploy it to accelerate the performance of their applications.

Conversations such as these hint that while flash has already gained acceptance among techies, its broader market adoption and acceptance is yet to come. To date, its cost has been relatively high. However more products offer flash as a cache (*such as occurs in hybrid storage arrays*) and offer technologies such as deduplication and compression. This will further drive down its effective cost per TB. By way of example, I was talking to one individual yesterday who already offers a flash-based solution for under \$300/TB (less than 30 cents/GB.)

- **Tape poised to become the cloud archive medium of choice.** When organizations currently look at how to best utilize the cloud, they typically view it as the ideal place to store their archival data for long term retention. This sets the cloud up as an ideal place in which to deploy tape as preferred medium to store this data largely due to its low operational costs, long media life and infrequent data access.

To accommodate this shift in how organizations are using tape libraries as well as make them more appealing to cloud services providers, tape library providers are adding REST APIs to their tape library interfaces so they appear as a storage target. While most organizations may not know (or

care) that the data they send to the cloud is stored to tape, they do care about its cost. By cloud providers storing data on tape, they can drive down these costs to a penny or less per GB per month.

---

## Drawing the Line between Open Source and Proprietary Code; Interview with iXsystems CTO Jordan Hubbard, Part V

Establishing a standard as to how an organization uses proprietary and open source code is at best difficult for most organizations. But iXsystems has essentially bet its future on the continued use of open source code in its product line. This makes it an imperative that it get this decision right to continue fostering support for its product in the open source community. This fifth entry in my interview series with [iXsystems](#)' CTO Jordan Hubbard discusses his thoughts on iXsystems' responsibility toward the open source community for their contributions and how it draws the line between proprietary and open source code.

**Ben:** *Do you have a responsibility to return your code to the open source community? Where do you see the line between proprietary and open source?*

**Jordan:** That is a great question and I think a lot of companies struggle with it. Eventually this always comes down to personal taste and where you combine the moral and the technical arguments.

It is a very tricky one to make, and I try not to moralize about this because there's been a lot of invective on this topic. Returning the code makes the most sense, from primarily a technical perspective, and we can leave the morals out of it.

I have a strong tropism towards supporting that community and the warm fuzzy part of it. But even if you totally set that aside, I have almost never seen a case in which it did not make more sense to open source something, and put the commercial value-add into the support of crafting custom solutions for people, and make the source code not the battleground.

It always makes more sense, in 99.9 percent of the cases, to have as many eyeballs on that code as possible, to have this feeling of inclusion that you get from saying, *"We are working shoulder to shoulder with you folks on a variety of technical challenges."*

There are always more technical challenges than you can ever possibly accomplish, for any size company. I don't care if you're a Fortune 500 with 20,000 employees who are all full-time engineers; there will always be more to do than you can possibly do.

Being able to actually work with those folks without having these firewalls in place, and being able to commit your code into open repositories, is important. I've seen a lot of companies get that wrong where they say, *"We will open source it but only in intervals. We will throw it over the wall periodically."* And that should be good enough, right?

The fact is that is not good enough because what people external to the company really want is to work with you. They do not really actually want the code. That is the biggest irony of open source. It is really not about the code. It is about the collaboration.

I do not want to sound too touchy feely, but it really is about the people and the relationships and being able to talk freely, discuss what is speculative, what science experiments are currently ongoing, what is really going to be in the next version of the product.

That is why working on something like [FreeNAS](#) is so rewarding, because we do all of that out in the open. All of the source code is freely available. But more importantly, the mailing lists on which we discuss these things, the IRC channels in which new features are discussed, which nuances of the product are discussed, the forums, bug reporting systems – all of that is completely open.

Our own people use it. It is not double entry bookkeeping where there is the stuff that's discussed inside and furtively and without any oversight or collaborative opportunities, and the stuff that goes on externally. They literally work in the open. That is what multi-community and mindshare and people who are actually enthusiastic about things means.

This is truly an open source project. You guys are doing it all right here. We can send you requests for patches, we can discuss stuff that is clearly half-baked and sitting in a branch somewhere. That is fine because it is on a branch. It is not going to be in the main product for a while, but we can check that out, we can see what you are doing, we can offer suggestions, we can get involved. That is what open source really means: developers working shoulder to shoulder.

***In [Part I](#) of this interview series with iXsystems' CTO Jordan Hubbard, we take a look at some ways in which iX's value propositions set it apart from its competitors.***

***In [Part II](#) of this interview series, we discuss***

*iXsystems' ability to consult with their clients and how that practice helps them create more customizable storage appliance and server configurations.*

*In [Part III](#) of this interview series, we discuss how iXsystems is introducing and managing flash drives in its storage systems, and why Jordan believes that a hybrid storage approach is currently the best solution.*

*In [Part IV](#) of this interview series, Hubbard shares how companies in general and iXsystems specifically benefits short and long term for its developers doing work at home and in the FreeBSD kernel community.*

*In [Part VI](#) of this interview series, we discuss Jordan's ideas on if the open source community is a meritocracy, and what type of person has the chance to rise above the rest in the field.*

---

**When Microseconds Matter:  
Delivering Highly Available  
Inline Deduplication and  
Consistent Low Latencies at  
Scale; Interview with Thomas  
Isakovich, Nimbus Data  
Systems, Inc. Chief Executive**

# Officer and Founder, Part 2

In this second blog entry from our interview with [Nimbus Data](#) CEO and Founder Thomas Isakovich, we discuss *microsecond latencies and how the recently announced Gemini X-series scale-out all-flash platform performs against the competition.*

**DCIG:** *Could you address what kind of latencies we should expect to see with the Gemini X-series?*

**Thomas:** With the [Gemini X](#) latency is going to be around 100 microseconds, whereas the Gemini F can get as low as 50 microseconds. That latency is consistent regardless of the number of Flash Nodes, which is pretty impressive.

**DCIG:** *Is that difference even detectable for the end user?*

**Thomas:** We had a potential client who was looking at purchasing 100 units. Our product delivered, I think, something like 95 microseconds in their tests. The next vendor delivered 150, and then the next one delivered 190. Even though we cost a little bit more, they bought all of it with us, because to them, that app hits so hard that it's a 2x delta between 95 microseconds and 190 microseconds. One is twice as fast. So certainly not everybody, but a lot of these web-scale guys that run these Oracle databases – they get that much transaction load, they really do care.

Also, with us if you lose a controller there's no change in performance. The design is such that all the IO – on the regular configuration basis, anyway – is processed by one controller, even though it's active/active. It's what we call asymmetric active/active. So, if you lose a controller, then all the IO remains at the same rate; there's no change in the performance. If you lose a drive with our product, it's like a 5 or 10 percent hit.

With some other competitive offerings, if they lose a

controller, they'll lose half their performance as well. They truly relied on balancing every IO and LUN ownership between two controllers. So you lose one controller, and you lose half the horsepower.

**DCIG:** *When you talk about the Gemini X's ability to lose a controller without a significant performance impact, are you talking about losing one of the two controllers in a Flash Node or losing one of the Flash Directors?*

**Thomas:** Either. In either scenario there's no performance impact.

**DCIG:** *The metadata management happens at the Flash Node level?*

**Thomas:** Yes.

**DCIG:** *And each Flash Node is independent in terms of deduplication?*

**Thomas:** Correct. It's possible that we could add kind of global cross-node deduplication. I'll be honest though, at 4K I don't think that's really going to do much, because the metadata tables are already going to have so many 4K variations. I can't imagine there being a slight variation from one or the other very often.

If we ended up doing it globally, then we kind of lose scale-out in a sense because you put the burden for that calculation on the Flash Director. I don't think we're going to do it. Given how granular the deduplication is, I think we're fine. If we're doing, say, 64K granularity, then you can maybe make an argument that global deduplication would help. But at 4K, I think we're good.

In the [Part 1](#) of this interview series, Thomas Isakovich guided us through the development of the Nimbus Data [Gemini X-series](#), and where he sees it fitting into the current market.

In the final part of this interview series, we discuss the

*appeal of Gemini X to enterprise and cloud service providers.*

---

# **Effectively Leverage CPU, DRAM and Flash Cache in Enterprise Storage Array Performance, Part II**

Providing high levels of capacity is only relevant if a storage array can also deliver high levels of performance. The number of CPU cores, the amount of DRAM and the size of the flash cache are the key hardware components that most heavily influence the performance of a hybrid storage array. In this second blog entry in my series examining the Oracle ZS3 Series storage arrays, I examine how its performance compares to that of other leading enterprise storage arrays using published performance benchmarks.

The Oracle [ZS3 Series](#) ZS3-2 and ZS3-4 storage controllers scale to support substantially more of these three key performance engines than any of their competitors offer as illustrated in the first table below. However, superior hardware by itself does not guarantee superior performance—a sophisticated operating system and caching algorithms are necessary to extract maximum performance from the hardware.



	ORACLE ZS3-2	ORACLE ZS3-4	EMC VNX5400	EMC VNX8000	EMC ISILON S200	NETAPP FAS3250	NETAPP FAS6250	IBM V7000U (4 Node)
Max Raw Capacity (TBs)	768	3,500	750	3,000	3,110	2,880	5,760	1,920
CPU Cores	32	80	8	32	8	16	16	32
DRAM (GB)	512	2,048	32	256	96	40	144	64
Flash Cache (TB)	16	22	1	4.2	13.8	2	8	Not Published

Oracle ZS3 Series storage leverages a multi-threaded, SMP (Symmetric Multi-Processing) operating system and Hybrid Storage Pools intelligent data caching architecture and algorithm to ensure that up to 90% of “hot” IO is processed in DRAM – up to 2TB per system. Frequently accessed data is cached in flash – up to 22TB per system – and less frequently accessed data is read from disk when needed. The efficacy of the ZFS Appliance’s hardware/software combination is that it delivers performance that far exceeds its traditional competitors as demonstrated in industry benchmarks.

As shown in the next table, both Oracle ZS3 Series storage systems beat the NetApp FAS3250 filer (the only other comparable two-node system) in performance. In contrast, the EMC VNX 8000 is a seven node system and the Isilon 200 is a 56-node system cost significantly more and, even in those two cases, the ZS3-4 delivers lower latency than they do.

The EMC VNX5400 is probably the most ill-equipped of these arrays to meet enterprise performance demands short and long term. It combines the DART operating system from CLARiiON and the FLARE operating system from Celerra in one physical storage array. In it, each one remains a separate, distinct operating system that is converged under a virtual hypervisor. This architectural approach adds latency to storage processing and complexity to storage management.

	ORACLE ZS3-2	ORACLE ZS3-4	EMC VNX5400	EMC VNX8000	EMC ISILON S200	NETAPP FAS3250	NETAPP FAS6250	IBM V7000U (2 Node)
SPC-2 MBPS	+	17,224	+	+	+	+	+	3,133
SPECsfs IOPS (NFS)	210535	450702	+	580796	456223	100922	+	+
SPECsfs Latency	1.12ms	700µs	+	780µs	3.27ms	1.76ms	+	+

+ No third party performance benchmark information publicly available

Taken together these SPC-2 and SPECsfs results **show** that the ZS3 Series storage excels at a range of workloads from high-throughput streaming performance applications, such as data warehousing and business intelligence to latency sensitive applications such as databases.

However, this level of performance that the Oracle ZS3 Series can deliver is relevant only if enterprises need it. In Oracle's case, Oracle Database users have always sought higher I/O and throughput to drive their applications.

In Oracle's customer base the ZS3 Series storage will prove meaningful as it removes existing throughput and I/O bottlenecks and accelerates the performance of Oracle Database and applications. In addition, through Oracle's hardware and software co-engineering development, there are a number of unique integration points between the Oracle ZS3 Series storage and Oracle Database (covered below) that further drive performance, efficiency, and lower TCO. The Oracle ZS3 Series storage also helps resolve other data center performance issues, especially highly virtualized environments.

***In [Part I](#) in this series, I examine how the Oracle ZS3 Series provides the levels of storage capacity and support that enterprise organizations expect.***

***In [Part III](#) in this series, I examine how the Oracle ZS3 Series differentiates itself in virtualized environments and what specialized features it offers to overcome the emerging storage network bottleneck.***

To read the entire DCIG Special Report that examines the competitive advantage that the Oracle ZS3 Storage Series offers for enterprise hybrid storage arrays, please follow this [link](#).

---

# Gemini X All Flash Scale-Out Storage Ready to Replace HDD as Enterprise Tier One; Interview with Thomas Isakovich, Nimbus Data Systems, Inc. Chief Executive Officer and Founder, Part 1

Recognized as an innovator in storage system technology, Thomas Isakovich sat down with [DCIG](#) to discuss the development, capabilities, and innovation in [Nimbus Data's](#) latest release: the Gemini X. In this first blog entry, he guides us through the development of the X-series, and where he sees it fitting into the current market.

**DCIG:** *Can you tell us what is so different about the X-series?*

**Thomas:** In terms of availability, this is probably the most advanced product – well, it *is* the most advanced product we've ever made, because it builds on everything that we've been improving. It takes that [Gemini](#) technology and then amps it up with true scale-out capability that is managed by our all-new Flash Director device. We've been working on it for the past two years. It's been a real challenge and also a pleasure

developing it.

And, really, for us it completes the story. We believe we have the most competitive all-flash system currently in the market with the Gemini F. The only caveat being how do we scale to huge, huge sets of capacity? We now provide that with the Gemini X, and we've done it in a way that keeps the software and the hardware building blocks about 90 percent shared between the two platforms.

Customers can start with the F series and go to the X series later. There's a lot of commonality between the two. From a manageability perspective, that familiarity will be a big plus. I think from an all-flash array portfolio perspective, we've got customers covered from three terabytes to a petabyte now – from a \$50,000 entry point to multimillion dollar solutions – all on the same Nimbus Data technology.

The timing of this product from our perspective is pretty perfect because our sales force is increasingly encountering customers that want to do wholesale refreshes of their entire tier-1 infrastructure. Not just flash for individual applications like databases and VDI, but really observing all-flash as a potential contender for the entirety of the tier-1 infrastructure. So having the ability to scale is well-timed and we're excited to be putting it out there now.

***DCIG: Can you talk more about the deduplication and compression of Gemini X?***

***Thomas:*** The deduplication and compression is really a sneak preview to an important feature of our forthcoming HALO 2014 operating system that we'll announce later this spring.

One of the challenges in scaling an array that uses inline deduplication is managing the vast metadata hash table that is the result of that, and keeping it in a manner where it's very rapidly accessible. And, as you know, a lot of solutions consume inordinate amounts of RAM to hold all this. But it's

actually the RAM constraints of the controllers that may be limiting the ability for these inline deduplicating storage arrays to scale. So, many of those guys have been resorting to scale-out because, really, who's going to build an Intel server that can hold 20 terabytes of RAM? And even if it could, how do you protect it?

So we've come up with an algorithm here that effectively uses about 1/50th the RAM and can deliver the same 4K block size in-line deduplication. This is one of the reasons we can build such high scale systems in such a small size. The Gemini X takes advantage of that technology, and so will the Gemini F, as part of running the HALO 2014 OS.

**DCIG: What environments are you seeing that are pushing high IOPS?**

**Thomas:** It's definitely geared more toward folks that just need huge amounts of capacity in a single domain. A lot of our customers, like our biggest one, which has 100 Gemini systems – they have no interest in actually presenting that as a single logical name space. They really do want 100 different name spaces, because of the way they're doing their scale-out. But they're a very sophisticated cloud provider, they can do very specific fancy things. For general purpose enterprise that doesn't have that level of sophistication – they're used to having their 500 terabyte hard drives or whatever – they need something that can present as one big box, and that's where this guy plays.

For example, a Fibre Channel port on a good day can do 100,000 IOPS. You're not going to get a million IOPS into a single server unless you're prepared to stick ten Fibre Channel cards in that server – which is going to be a challenge – and then run everything perfectly parallelized and all this other stuff. So our thought process in supporting four million IOPS is that we're going to need to support dozens or maybe hundreds of physical machines. And at 100,000 IOPS a port,

that actually works out to about four million IOPS because you can have up to forty host ports on the Gemini X.

It's not so much that there's any one application that can come close to that, but you need to maintain a reasonable sort of IOPS-per-terabyte/IOPS-per-port kind of ratio. That's what our rationale is on achieving four million, because if you look at the Gemini F, by itself it's doing north of a million in a read-write balance scenario. So on an IOPS-per-terabyte basis, the Gemini F is actually better, because when you do a cluster scale-out like this, you're going to have at least a little bit of latency from the cluster grid. We've kept latency very low because of the Flash Director design.

*In part two of this interview series, we discuss micro-latency and how the Gemini X performs against the competition.*

---

## **A Primer on Today's Storage Array Types**

Anyone who managed IT infrastructures in the late 1990's or early 2000's probably still remembers how external storage arrays were largely a novelty reserved for high end enterprises with big data centers and deep pockets. Fast forward to today and a plethora of storage arrays exist in a variety of shapes and sizes at increasingly low price points. As such it can be difficult to distinguish between them. To help organizations sort them out, my blog entry today provides a primer on the types of storage arrays currently available on the market.

The large number of different storage arrays on the market today would almost seem to suggest that there are too many on

the market and that a culling of the herd is inevitable. While there may be some truth to that statement, storage providers have been forced to evolve, transform and develop new storage arrays to meet the distinctive needs of today's organizations. This has resulted in the emergence of multiple storage arrays that have the following classifications.

- **Enterprise midrange arrays.** These are the original arrays that spawned many if not all of the array types that follow. The primary attributes of these arrays are high availability, high levels of reliability and stability, moderate to high amounts of storage capacity and mature and proven code. Features that typify these arrays include dual, redundant controllers, optimized for block level traffic (FC & iSCSI), and hard disk drives (HDDs). These are generally used as general purpose arrays to host a wide variety of applications with varying capacity and performance requirements. (*The most recent DCIG Buyer's Guide on midrange arrays may be accessed via this [link](#).*)
- **Flash memory storage arrays.** These are the new speed demons of storage arrays. Populated entirely with flash memory, many of these arrays can achieve performance of 500,000 to 1 million IOPS with latency at under a millisecond.

*The two potential "gotchas" here are their high costs and relative immaturity of their code. To offset these drawbacks, many providers include compression and deduplication on their arrays to increase their effective capacity. Some also use open source versions of ZFS as a means to mature their code and overcome this potential client objection. Making these distinctively different from the other array types in this list of array types is their ability to manage flash's idiosyncrasies (garbage collection, wear leveling, etc.) as well as architecting their controllers to facilitate the faster throughputs that flash provides so they do not become*

a bottleneck. (The most recent DCIG Buyer's Guide on flash memory storage arrays may be accessed via this [link](#).)

- **Hybrid storage arrays.** These arrays combine the best of what both flash memory and midrange arrays have to offer. Hybrid storage arrays offer both flash memory and HDDs though what distinguishes them from a midrange array is their ability to place data on the most appropriate tier of storage at the best time. To accomplish this feat they use sophisticated caching algorithms. A number also use compression and deduplication to improve storage efficiencies and lower the effective price per GB of the array. (The most recent DCIG Buyer's Guide on hybrid storage arrays may be accessed via this [link](#).)
- **Private cloud storage arrays.** Private cloud storage arrays (sometimes referred to as scale-out storage arrays) are defined by their ability to dynamically add (or remove) more capacity, performance or both to an existing array configuration by simply adding (or removing) nodes to the array.

The appeals of these arrays are three-fold. 1.) They give organizations the flexibility to start small with only as much capacity and performance as they need and then scale out as needed. 2.) They simplify management since administrators only need to manage one logical array instead of multiple smaller physical arrays. 3.) Organizations can mitigate and often eliminate the need to migrate data to new arrays as the array automatically and seamlessly redistributes the data across the physical nodes in the logical array.

While these arrays possess many of the same attributes as public storage clouds in terms of their data mobility and scalability, they differentiate themselves by being intended for use behind corporate firewalls. (The most recent DCIG



*Buyer's Guide on private cloud storage arrays may be accessed via this [link](#).)*

- **Public cloud storage gateway arrays.** The defining characteristic of these storage arrays is their ability to connect to public storage clouds on their back end. Data is then stored on their local disk cache before it is moved out to the cloud on some schedule based upon either default or user-defined policies.

*The big attraction of these arrays to organizations is that it eliminates their need to continually scale and manage their internal storage arrays. By simply connecting these arrays to a public storage cloud, they essentially get the capacity they want (potentially unlimited but for a price) and they eliminate the painful and often time-consuming need to migrate data every few years. (A DCIG Buyer's Guide on this topic is scheduled to be released sometime next year.)*

- **Unified storage arrays.** Sometimes called converged storage arrays, the defining characteristic of these storage arrays is their ability to deliver both block (FC, iSCSI, FCoE) and file (NFS, CIFS) protocols from a single array. In almost every other respect they are similar to midrange arrays in terms of the capabilities they offer.

*The main difference between products in this space is that some use a single OS to deliver both block and file services while others use two operating systems running on separate controllers (this alternate architecture gave rise to the term "converged.") The "unified" name has stuck in large part because both block and file services are managed through a single (i.e. "unified") interface though the "converged" and "unified" terms are now used almost interchangeably.. (The most recent DCIG Buyer's Guide on midrange unified storage arrays may be accessed via this [link](#).)*

Organizations should take note that even though multiple storage array types exist, many storage arrays exist that satisfy multiple classifications. While no one array model yet ships that fits neatly into all of them, DCIG expects that by the end of 2014 there will be a number of storage array models that will. This becomes important to those organizations that want the flexibility to configure a storage array in a way that best meets their specific business and/or technical requirements while eliminating the need for them to buy another storage array to do so.