Nanoseconds, Stubborn SAS, and Other Takeaways from the Flash Memory Summit 2019

Every year at the Flash Memory Summit held in Santa Clara, CA, attendees get a firsthand look at the technologies that will impact the next generation of storage. This year many of the innovations centered on forthcoming interconnects that will better deliver on the performance that flash offers today. Here are DCIG's main takeaways from this year's event.

Takeaway #1 – Nanosecond Response Times Demonstrated

PCI Express (PCIe) fabrics can deliver nanosecond response times using resources (CPU, memory, storage) situated on different physical enclosures. In meeting with PCIe provider, [Dolphin Interconnect Solutions](#), it demonstrated how an application could access resources (CPU, flash storage & memory) on different devices across a PCIe fabric in nanoseconds. Separately, [GigaIO](#) [announced](#) 500 nanosecond end-to-end latency using its [PCIe FabreX](#) switches. While everyone else at the show was boasting about microsecond response times, Dolphin and GigaIO introduced nanoseconds into the conversation. Both these companies ship their solutions now.

Takeaway #2 – Impact of NVMe/TCP Standard Confirmed

Ever since we heard the industry planned to port NVMe-oF to TCP, DCIG [thought](#) this would accelerate the overall adoption of NVMe-oF. Toshiba confirmed our suspicions. In discussing

its [Kumoscale](#) product with DCIG, it shared that it has seen a 10x jump in sales since the industry ratified the NVMe/TCP standard. This stems from all the reasons DCIG stated in a previous [blog entry](#) such as TCP being well understood, Ethernet being widely deployed, its low cost, and its use of existing infrastructure in organizations.

Takeaway #3 — Fibre Channel Market Healthy, Driven by Enterprise All-flash Array

According to FCIA leaders, the Fibre Channel (FC) market is healthy. FC vendors are selling 8 million ports per year. The enterprise all-flash array market is driving FC infrastructure sales, and 32 Gb FC is shipping in volume. Indeed, DCIG's research revealed 37 all-flash arrays that support 32 Gb FC connectivity.

Front-end connectivity is often the bottleneck in all-flash array performance, so doubling the speed of those connections can double the performance of the array. Beyond 32 Gb FC, the FCIA has already ratified the 64 Gb standard and is working on the 128 Gb FC. Consequently, FC has a long future in enterprise data centers.

FC-NVMe brings the benefits of NVMe-oF to Fibre Channel networks. FC-NVMe reduces protocol overhead, enabling GEN 5 (16 Gb FC) infrastructure to accomplish the same amount of work while consuming about half the CPU of standard FC.

Takeaway #4 — PCIe Will Not be Denied

All resources (CPU, memory and flash storage) can connect with one another and communicate over PCIe. Further, using PCIe eliminates the need for introducing the overhead associated with storage protocols (FC, InfiniBand, iSCSI, SCSI). All these resources talk the PCIe protocol. With the PCIe 5.0 standard formally ratified in May 2019 and discussions about PCIe 6.0 occurring, the future seems bright for the growing

adoption of this protocol. Further, AMD and Intel having both thrown their support behind it.

Takeaway #5 — SAS Will Stubbornly Hang On

DCIG's research finds that over 75% of AFAs support 12Gb/second SAS now. This predominance makes the introduction of 24G a logical next step for these arrays. A proven, mature, and economical interconnect, few applications can yet drive the performance limits of 12Gb, much less the forthcoming 24G standard. Adding to the likelihood that 24G moves forward, the SCSI Trade Association (STA) reported that the recent 24G plug fest went well.

Editor's Note: This blog entry was updated on August 9, 2019, to correct grammatical mistakes and add some links.

Hackers Say Goodbye to Ransomware and Hello to Bitcoin Mining

Ransomware gets a lot of press — and for good reason — because when hackers break through your firewalls, encrypt your data, and make you pay up or else lose your data, it rightfully gets people's attention. But hackers probably have less desire than most to be in the public eye and sensationalized ransomware headlines bring them unwanted attention. That's why some hackers have said goodbye to the uncertainty of a payout associated with getting a ransom for your data and instead look to access your servers to do some bitcoin mining using your CPUs.

A week or so ago a friend of mine who runs an Amazon Web Services (AWS) consultancy and reseller business shared a story with me about one of his clients who hosts a large SaaS platform in AWS.

His client had mentioned to him in the middle of the week that the applications on one of his test servers was running slow. While my friend was intrigued, he did not at the time give it much thought. This client was not using his managed services offering which meant that he was not necessarily responsible for troubleshooting their performance issues.

Then the next day his client called him back and said that now all his servers hosting this application — test, dev, client acceptance, and production — were running slow. This piqued his interest, so he offered resources to help troubleshoot the issue. The client then allowed his staff to log into these servers to investigate the issue

Upon logging into these server, they discovered that all instances running at 100% also ran a Drupal web application. This did not seem right, especially considering that it was early on a Saturday morning when the applications should mostly be idle.

After doing a little more digging around on each server, they discovered a mysterious multi-threaded process running on each server that was consuming all their CPU resources. Further, the process also had opened up a networking port to a server located in Europe. Even more curious, the executable that launched the process had been deleted after the process started. It was as if someone was trying to cover their tracks.

At this point, suspecting the servers had all been hacked, they checked to see if there were any recent security alerts. Sure enough. On March 28, 2018, Drupal issued a security advisory that if you were not running Drupal 7.58 or Drupal

8.5.1, your servers were vulnerable to hackers who could remotely execute code on your server.

However, what got my friend's attention is that these hackers did not want his client's data. Rather, they wanted his client's processing power to do bitcoin mining which is exactly what these servers had been doing for a few days now on behalf of these hackers. To help their client, they killed the bitcoin mining process on each of these servers before calling his client to advise them to patch Drupal ASAP.

The story does not end there. In this case, his client did not patch Drupal quickly enough. Sometime after they killed the bitcoin mining processes, another hacker leveraged that same Drupal security flaw and performed the same hack. By the time his client came to work on Monday, there were bitcoin mining processes running on those servers that again consumed all their CPU cycles.

What they found especially interesting was how the executable file that the new hackers had installed worked. In reviewing their code, the first thing it did was to kill any pre-existing bitcoin mining processes started by other hackers. This freed all the CPU resources to handle bitcoin mining processes started by the new hackers. The hackers were literally fighting each other over access to the compromised system's resources.

Two takeaways from this story:

1. *Everyone is rightfully worried about ransomware but bitcoin mining may not hit corporate radar screens.* I

doubt that hackers want the FBI, CIA, Interpol, MI6, Mossad, or any other criminal justice agency hunting them down any more than you or I do. While hacking servers and "*stealing*" CPU cycles is still a crime, it probably is much further down on the priority list of most companies as well as these agencies.

A bitcoin mining hack may go unnoticed for long periods of time

and may not be reported by companies or prosecuted by these criminal justice agencies even when reported because it is easy to perceive this type of hack as a victimless crime. Yet every day the hacker's bitcoin mining processes go unnoticed and remain active, the more bitcoin the hackers earn. Further, one should assume hackers will only become more sophisticated going forward. Expect hackers to figure out how to install bitcoin mining processes that run **without** consuming all CPU cycles so these processes remain running and unnoticed for longer periods of time.

2. ***Hosting your data and processes in the cloud does not protect your data and your processes against these types of attacks.*** [AWS](#) has all the utilities available to monitor and detect these rogue processes. That said,

organizations still need someone to implement these tools and then monitor and manage them.

Companies may be relieved to hear that some hackers have stopped targeting their data and are instead targeting their processors to use them for bitcoin mining. However, there are no victimless crimes. Your pocket book will still get hit in cases like this as Amazon will bill you for using these resources.

In cases like this, if companies start to see their [AWS](#) bills going through the roof, it may not be the result of their businesses. It may be their servers have been hacked and they are paying to finance some hacker's bitcoin mining operation. To avoid this scenario, companies should ensure they have the right internal people and processes in place to keep their applications up-to-date, to protect infrastructure from attacks, and to monitor their infrastructures whether hosted on-premise or in the cloud.

Real-World Performance Testing Can Help Savvy Organizations Future Proof Their Emerging Flash Infrastructure

Almost all size organizations now view flash as a means to accelerate application performance in their infrastructure … and for good reason. Organizations that deploy flash typically see increases in performance by factor of up to 10x. But while

many all-flash storage arrays can deliver these increases in performance, savvy organizations must prepare to do more than simply increase workload performance. They need to identify solutions that help them better troubleshoot their emerging flash infrastructure as well as future proof their investment in flash by better modeling anticipated application workloads on all-flash arrays being evaluated before they are acquired.

One of the big advantages of all-flash arrays is that they make it much easier for organizations to improve the performance of almost any application regardless of its type. However the ease in which these all-flash arrays accelerate performance also may prompt organizations to lower their guard and fail to consider all of the potential pitfalls that accompany the deployment of such an array. One can just as easily over-provision an all-flash array as a disk-based array. Given the price per GB differences between the two, the cost penalty for over-provisioning all-flash arrays can be very significant.

Common pitfalls that DCIG hears about include:

- The all-flash array works fine at first but performance unexpectedly drops. This leaves everyone wondering, "What is the root cause of the problem?" The all-flash array? The storage network? The server? The application? Or some other component?
- An organization starts by putting a few or even one high-performance application on the all-flash array. It works so well that all of sudden everyone in the organization wants to put their applications on the array so performance on the all-flash array begins to suffer.

Performance analytics software can help in both of these cases as the recently released Load DynamiX 5.0 Storage Performance Analytics solution helps to illustrate. In the first scenario mentioned above, Load DynamiX provides a workload analyzer

that examines performance in existing networked storage environments (FC/iSCSI now, CIFS/NFS coming in 1H2016.) This analyzer pulls performance data from the production storage arrays as well as from the Ethernet or FC switches so organizations can visualize existing storage workloads.

The Load DynamiX software then more importantly equips organizations to analyze these workloads as it automates this task using a combination of real-time and historical views of the data. By comparing IOPs, throughput, latency, read/write and random/sequential workload mixes among many others, it can begin to paint a picture of what is actually going on in the environment and identify the root cause of the performance bottleneck. This type of automation and insight becomes especially important when performance bottlenecks occur intermittently and at seemingly random and unpredictable intervals.

Yet maybe what makes the Load DynamiX solution particularly impressive is that after it captures these various pools of performance data, organizations can use it to optionally recreate the same behavior in their labs. In this way, they can experiment and trial possible solutions to the problem in a lab environment without tampering with the production environment and potentially making the situation worse. This gives IT organizations the opportunity to identify a viable solution and verify it works in their lab so they have a higher degree of confidence it will work in their production environment before they start the process of actually implementing the proposed fix.

This ability to capture and model workloads also becomes a very handy feature to have at one's disposal when trialing new all-flash arrays as one organization recently discovered. It used Load DynamiX to first capture current performance data on its existing environment and then ran it against six (6) all-flash arrays under consideration.

As it turns out, all six (6) of them achieved the desired sub-2ms response times that they were hoping and expecting to get (*as opposed to the 10ms response times that they were seeing using their existing disk-based array*) when each of these all-flash arrays was tested using the company's existing Oracle-based application workloads as Chart 1 illustrates.
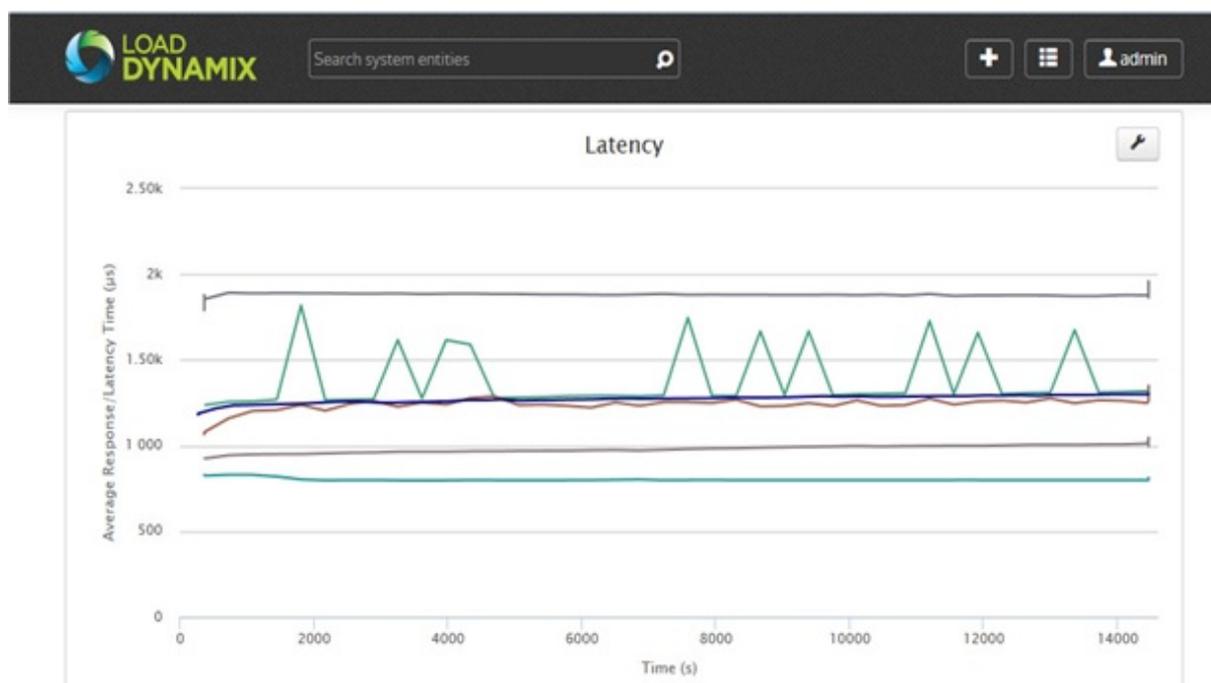

Chart 1

However the organization then did something very clever. It fully expected that over time the workloads on the all-flash array would increase for the reasons cited above — perhaps by as much as 10x in the years to come. To model those anticipated increases, it again used Load DynamiX to simulate a 10x increase in application workload performance. When measured against this 10x increase in workload, substantial performance differences emerged between the various all-flash arrays as Chart 2 illustrates.
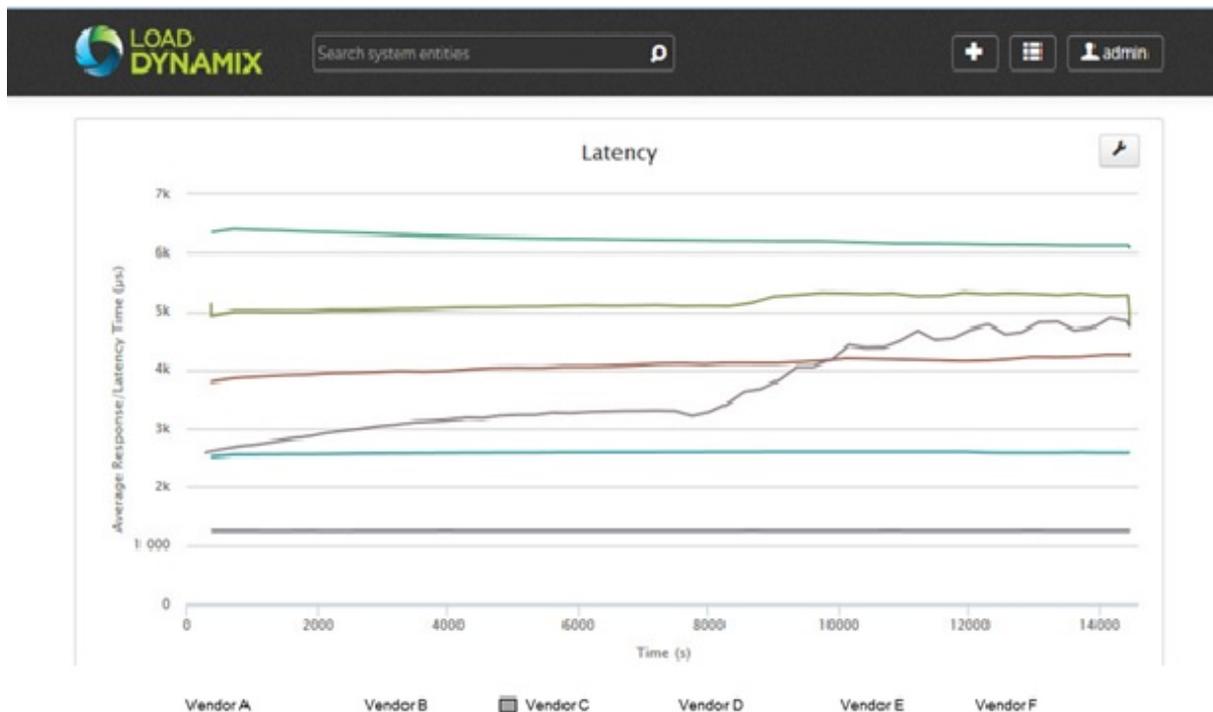
Chart 2

Under this 10x increase in workload, all of the all-flash arrays still outperformed the disk-based array. However only one of these arrays was able to deliver the sustained sub-2ms response times that this organization wanted its all-flash array solution to deliver over time. While a variety of factors came into play that account for these lower performance numbers, , it is noteworthy that all of these all-flash arrays except one had compression and deduplication turned on. As such, as applications workloads increased, it is conceivable and logical to conclude that these data reduction technologies begin to extract a heavier performance toll.

All-flash arrays have been a boon for organizations as they eliminate many of the complex, mind-numbing tasks that highly skilled individuals previously had to perform to coax the maximum amount of performance out of disk-based arrays. However that does not mean performance issues no longer exist once flash is deployed. Using performance analytics software like the Load Dynamix 5.0 Storage Performance Analytics solution, organizations can now better trouble-shoot both their legacy and new all-flash environment as well as make

better, more informed choices about all-flash arrays so they can better scale them to match their anticipated increases in workload demands.

The Performance of a $500K Hybrid Storage Array Goes Toe-to-Toe with Million Dollar All-Flash and High End Storage Arrays

On March 17, 2015, the Storage Performance Council (SPC) **updated** its "Top Ten" list of SPC-2 results that includes performance metrics going back almost three (3) years to May 2012. Noteworthy in these updated results is that the three storage arrays ranked at the top are, in order, a high end mainframe-centric, monolithic storage array (the **HP XP7**, OEMed from Hitachi), an all-flash storage array (from startup Kaminario, the **K2 box)** and a hybrid storage array (Oracle ZFS Storage **ZS4-4 Appliance**). Making these performance results particularly interesting is that the hybrid storage array, the Oracle ZFS Storage ZS4-4 Appliance, can essentially go toe-to-toe from a performance perspective with both the million dollar HP XP7 and Kaminario K2 arrays and do so at approximately half of their cost.

Right now there is a great deal of debate in the storage industry about which of these three types of arrays — all-flash, high end or hybrid — can provide the highest levels of performance. In recent years, all-flash and high end storage

arrays have generally gone neck-and-neck though all-flash arrays are generally now seen as taking the lead and pulling away.

However, when price becomes a factor (*and when isn't price a factor?*) such that enterprises have to look at price and performance, suddenly hybrid storage arrays surface as very attractive alternatives for many enterprises. Granted, hybrid storage arrays may not provide all of the performance of either all-flash or high end arrays, but they can certainly deliver superior performance at a much lower cost.

This is what makes the recently updated Top Ten results on the SPC website so interesting. While the breadth of arrays covered in the published SPC results by no means cover every storage array on the market, they do provide enterprises with some valuable insight into:

- How well hybrid storage arrays can potentially perform
- How comparable their storage capacity is to high-end and all-flash arrays
- How much more economical hybrid storage arrays are

In looking at these three arrays that currently sit atop the SPC-2 [Top Ten](#) list and how they were configured for this test, they were comparable in one of the ways that enterprises examine when making a buying decision. For instance, all three had comparable amounts of raw capacity.

Raw Capacity

High-End   HP XP7
                        230TB
All-Flash   Kaminario K2
    179TB
Hybrid Oracle ZFS Storage ZS4-4 Appliance
175TB

Despite using comparable amounts of raw capacity for testing purposes, they got to these raw capacity totals using decidedly different media. The high end, mainframe-centric HP XP7 used 768 300GB 15K SAS HDDs to get to its 230TB total while the all-flash Kaminario K2 used 224 solid state drives (SSDs) to get to its 179TB total. The Oracle ZS4-4 stood out from these other two storage arrays in two ways. First, it used 576 300GB *10K* SAS HDDs. Second, its storage media costs were a fraction of the other two. Comparing strictly list prices, its media costs were only about 16% of the cost of the HP XP7 and 27% of the cost of the Kaminario K2.

These arrays also differed in terms of how many and what types of storage networking ports they each used. Both the HP XP7 and the Kaminario K2 used a total of 64 and 56 8Gb FC ports respectively for connectivity between the servers and their storage arrays. The Oracle ZS4-4 only needed 16 ports for connectivity though it used Infiniband for server-storage connectivity as opposed to 8Gb FC. The HP XP7 and Oracle ZS4-4 also used cache (512GB and ~3TB respectively) while the Kaminario K2 used no cache at all. It instead used a total of 224 solid state drives (SSDs) packaged in 28 flash nodes (8-800GB SSDs in each flash node.)

This is not meant to disparage the configuration or architecture of any of these three different storage arrays as each one uses proven technologies in the design of their arrays. Yet what is notable is the end results when these three arrays in these configurations are subjected to the same SPC2 performance benchmarking tests.

While the HP XP7 and Kaminario K2 came out on top from an overall performance perspective, it is interesting to note how well the Oracle ZS4-4 performs and what its price/performance ratio is when compared to the high end HP XP7 and the all-flash Kaminario K2. It provides 75% to over 90% of the performance of these other arrays at a cost per MB that is up to 46% less.

SPC-2 Top 10 Performance Results					
Vendor	Date	GB/s	$/MB/s	TBs	Price
HP XP7	2014	43.01	$28.30	192.76	$1,217,462!
Kaminario K2	2013	33.47	$29.79	60.12	$997,348!
Oracle ZS4-4	2015	31.48	$17.09	30.96	$538,050!
Oracle ZS3-4	2013	17.24	$22.53	31.61	$388,472!
Oracle ZS3-2	2014	16.20	$12.08	24.20	$195,916!
Fujitsu Eternus DX8700	2013	16.04	$79.40	71.40	$1,275,163!
IBM DS8870	2012	15.42	$131.21	30.92	$2,023,742!
IBM Storwize V7000	2012	14.58	$129.14	74.49	$1,883,036!
NEC M700	2013	14.41	$25.10	53.55	$361,612!
Hitachi VSP	2012	13.15	$95.38	129.11	$1,254,093!
HP P9500 XP	2012	13.15	$88.34	129.11	$1,161,504!
Hitachi HUS VM	2014	11.27	$32.64	193.67	$368,065!

Source: "Top Ten" SPC-2 Results, https://www.storageperformance.org/results/benchmark_results_spc2_top-ten

It is easy for enterprises to become enamored with all-flash arrays or remain transfixed on high-end arrays because of their proven and perceived performance characteristics and benefits. But these recent SPC-2 performance benchmarks illustrate that hybrid storage arrays such as the Oracle ZFS Storage ZS4-4 Appliance can deliver levels of performance that are comparable to million-dollar all-flash and high-end arrays at half of their cost which are numbers that any enterprise can take to the bank.

Performance Testing as a

Technology Validation Method is as Complex as It Sounds

One of the main objectives of every DCIG Buyer's Guide is to help enterprises and/or their buyers create a short list of technology products that align with their specific business or technical needs. But once they have a short list of products that meet those needs, they still need some criteria to help them make the right choice from among those products. While there is no silver bullet that guarantees they make the "best" selection, performance testing is an option to which organizations often turn to validate a choice in technology though using this method is as complex as it sounds.

Every Buyer's Guide that DCIG produces has a number of objectives that DCIG's analysts endeavor to achieve to the best of their abilities to include:

- Identify products in a particular market segment
- Gather the most up-to-date information about each product
- Present product information accurately
- Score and rank product features in a way that generally represents end user priorities

However perhaps chief among these objectives, which is not listed above, is to help organizations arrive at a short list of 3-5 products. From that list, they may then choose one that is appropriate for them to acquire and implement in their environment. The issue then becomes, "How does an organization go from this short list of 3-5 products to selecting the most appropriate for their environment?"

Organization employ a number of tactics to do so that include:

- Conduct further research into the features of that product

- Functional testing
- Price comparisons
- Prioritizing supported product features according to their needs
- Performance testing

Of these five listed above (and there are others,) performance testing is often the logical and best means to arrive at the best solution. This is especially true in enterprise environments. Selecting a solution that is eventually determined to be undersized could prompt them to buy more products for their environment. This could in turn make the solution more complex and costly to implement and manage both short and long term. This undersized solution may even result in it not delivering the anticipated value and them being forced to bring in an entirely new solution to replace it.

Alternatively, an organization may choose to over-engineer the proposed solution to account for potential spikes in application performance that never materialize. This approach can again prompt organizations to invest too much in a solution as a means to insure themselves against acquiring an undersized solution.

Performance testing theoretically eliminates both of these possibilities as organizations may first test the solution to make sure it will work as intended in their environment. On the surface this approach sounds great but performance testing as a means to do a technology validation is as complex as it sounds for at least three reasons.

1. *It is often difficult if not impossible to simulate real world workloads in testing and development environments.* Real-world workloads are called that for a reason. It is difficult to simulate the types of peaks and valleys of performance that they generate in lab environments intended for testing and development. While technologies such as virtualization, snapshots and instant recoveries

certainly make it easier and more practical for organizations to run production applications in testing and development environments, they rarely can perfectly reproduce the exact conditions to which production workloads are subjected.

2. ***Setting up performance tests and interpreting performance results is a specialized skill.*** Assuming an organization can create an environment that closely mimics their production environment, they still need to make sure they set up the test correctly so that they get meaningful results. To set up the performance test, capture the output and then interpret the results is a very specialized skill which organizations do not necessarily possess.

3. ***It is time consuming.*** Organizations today are often extremely pressed for time and sometimes it is easier and potentially even less expensive to buy an overbuilt, over-engineered solution than to invest time in testing a variety of solutions to identify the perfect fit.

Performance testing sounds great in theory and can be an excellent means to identify the best product for your environment. Yet performance testing is as complex as it sounds even when the performance testing produces the desired results due to the time and effort required to execute upon it. In a forthcoming blog entry I will examine an appliance targeted at enterprises that helps them overcome these historical obstacles associated with performance testing of storage solutions being considered for acquisition.