

Dell EMC and Toshiba Power Nutanix to Deliver High-Performance and High-Scalability in Enterprise Workloads



Executive Summary

Hyper-converged Infrastructure (HCI) is becoming a significant services platform in the datacenter, and with good reason. Abstracting processing, storage and network resources to create a fully software-defined datacenter is the logical progression of the decade-long virtualization trend. Companies like Nutanix are leading the way, enabling IT departments to create and expand application environments through the simple addition of common off-the-shelf servers. Eliminating the expense and complexity of deploying and managing a datacenter full of disparate devices with a mixture of management interfaces plus multiple networking and storage protocols makes it easy to deploy and support an on-premise cloud.

As one of the top providers of enterprise servers, the Dell EMC PowerEdge product line is a recognized standard when it comes to datacenter deployments. Powered by Toshiba PX05S 3840GB 12Gb/s SAS SSDs, the PowerEdge platform is an ideal building block for a Nutanix private cloud. Dell EMC's partnership with Nutanix extends to the XC series appliances—one and two rack unit PowerEdge servers certified specifically to support the Nutanix HCI solution.

Demartek deployed a four-node cluster of Dell EMC PowerEdge XC730xd servers to test the performance and scalability of Microsoft SQL Server workloads in a

Nutanix Hyper-V environment. Nutanix managed in-chassis Toshiba PX05S SSDs provided the all flash software-defined storage layer serving I/O demands of the testbed. With three server nodes initially in an HCI cluster, baseline metrics were established and the number of virtual machines steadily increased. The fourth server was then added to the cluster and additional virtual machines installed to evaluate ease of use in growing and supporting a Nutanix cluster, and the performance impact, if any of doing so.

Key Findings

- > The amount of database work performed (as measured by the number of database transactions) by the Dell EMC Nutanix cluster scaled linearly and predictably as virtual machines were deployed within the cluster.
- > The scaling of work performed by the cluster itself was not impacted as the cluster size was expanded.
- > Toshiba PX05S SSD maintained storage I/O response times in the low milliseconds as I/O requirements were steadily increased.

Dell EMC XC servers with Toshiba 12Gb/s SAS SSDs Power High-Performance with Nutanix Hyper-converges solution

Dell EMC PowerEdge Servers

Dell EMC is an industry leader in the enterprise server market. The PowerEdge family of the servers supports the latest memory technology and Intel Xeon processors in one, two, or four rack unit form factors. With the XC series of hyper-converged appliances, one and two rack unit PowerEdge servers are certified by Dell EMC and Nutanix for complete compatibility and supportability with Nutanix, including processors, memory, drives, and interface cards. Coming pre-installed with a choice of hypervisors and the Nutanix software, a fully functional Nutanix HCI environment can be configured out-of-the-box in under an hour.

This evaluation deployed four Dell EMC XC730xd servers. Each server was provisioned with Intel Xeon E5-2650 v4 CPUs, 256 GB RAM, and 12 Toshiba PX05S SSDs. Microsoft Windows Server with Hyper-V was chosen as the hypervisor.

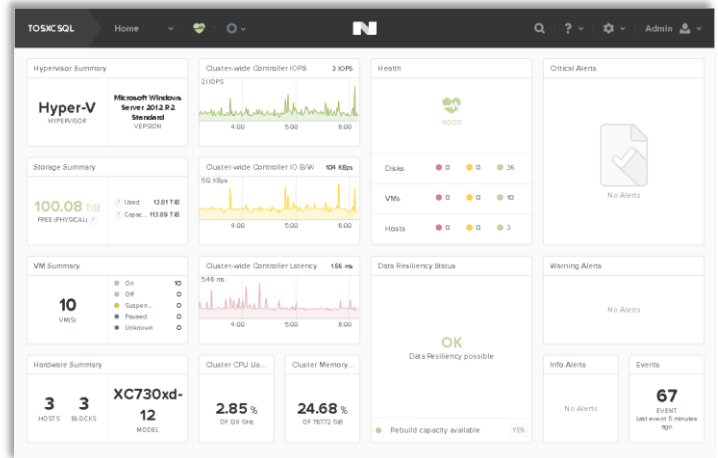
Toshiba PX05S 12Gb/s SAS SSDs

Toshiba recently released the PX05S SSD, a 12Gb/s SAS SSD available in 200GB to 3840GB formatted capacities. Models within this drive family are available for write intensive or mixed-use read/write applications with variable levels of endurance to allow customers to select the drive type tailored to their specific needs. Twelve drives in a 3840GB formatted capacity were provisioned in each of the four Dell EMC servers to present more than 150TB of storage to Nutanix.

Nutanix HCI

Nutanix produces one of the top HCI solutions on the market today. Installed on off-the-rack server hardware, Nutanix ties processing, memory, and storage into a single virtualized service with an easy to use management interface. By using commodity servers as building blocks, clusters can be provisioned to the size needed at the time of implementation and grown server-by-server, keeping costs and complexity down. Software-defined storage and a high-speed TCP-IP network connecting all cluster nodes maximize storage performance without the complexity of a traditional

SAN and multiple networking protocols. Nutanix is extremely flexible, supporting Windows Hyper-V and VMware ESX hypervisors, or be deployed with its own Acropolis hypervisor. These design options minimize the amount of specialized skills needed to set up and manage a Nutanix HCI solution and permit customers to fit Nutanix into their environment in a manner that best meets their specific requirements.

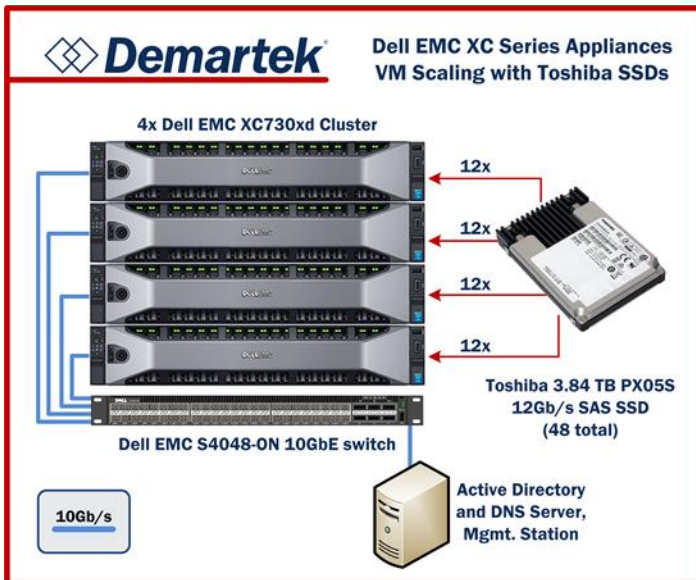


The Nutanix Prism management interface

As a server hosting a Nutanix controller virtual machine is added to a Nutanix cluster, all storage on that server can be included in a Nutanix storage pool. This storage pool is shared across the cluster as network attached storage (using the SMB protocol in the case of a Windows Hyper-V cluster) and virtual hard drives provisioned from it for virtual machine operating systems and data storage. When data is written, multiple copies of that data are spread across the cluster nodes to ensure continued availability in the event that any node goes offline. The Nutanix strategy is to also migrate a copy of any data requested by read-I/O to the node hosting the virtual machine reading that data. This eliminates cross-node reads, reducing network traffic and improving storage I/O response times. Write I/O performance will be affected by the need to place new or modified data in multiple locations making high-speed storage and networking essential for a cluster supporting applications with strict service level requirements.

Dell EMC XC servers with Toshiba 12Gb/s SAS SSDs Power High-Performance with Nutanix Hyper-converges solution

The Evaluation Configuration



The evaluation environment consisted of four Dell EMC PowerEdge R730xd servers. A single PowerEdge R230 management station also functioned as the cluster's Active Directory server but was not considered part of the system under test. Including a Dell EMC S4048-ON 10GbE switch, this entire setup consumed ten rack units of space.

Evaluation Parameters and Goals

The intention of this evaluation was to consider Dell EMC XC730xd servers and Toshiba PX05S drives as the hardware layer for a Nutanix HCI cluster in Microsoft Hyper-V. As such, this is a system, or solution, evaluation, tuned to perform the most work within test parameters and not highlight any single aspect of the platform over any other. The engine chosen to drive computing and I/O in the cluster was a Microsoft SQL Server OLTP database and application with both components hosted on each virtual machine. The I/O profile of this workload mirrored industry standard read-heavy OLTP benchmarking workloads with a read-to-write transaction ratio of roughly ten to one.

Testing proceeded in two phases. The first phase was

scale-up testing. We define scale-up as creating an initial configuration and steadily increasing the processing and I/O demands placed on it. Three cluster nodes were deployed as this initial configuration and three Microsoft Windows Server 2016 virtual machines with SQL Server 2016 were installed upon them. Each virtual machine was given 16 virtual processors and 64GB of virtual memory. A single workload was run on a single virtual machine and then on one virtual machine per cluster node. This was followed by two virtual machines per node executing the same workload, and finally three virtual machines per node. Scaling up the workload was stopped at three virtual machines per node as this consumed 75% of each cluster node's physical memory.

For the second phase, a scale-out scenario was created by adding a fourth node to the cluster (by scaling out, we mean expanding the initial cluster through increasing the number of servers that make it up). Three more SQL Server virtual machines, again with the same OLTP workload, were installed on this cluster node. All twelve virtual machines then ran the same OLTP application simultaneously.

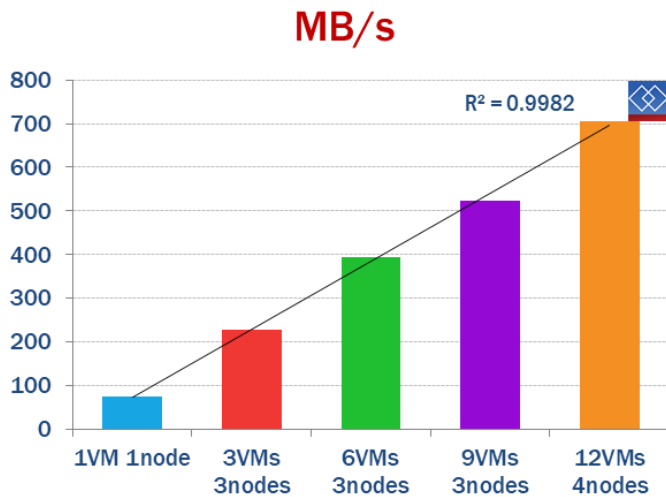
The objective of these test cases was to evaluate the performance of the basic cluster as the density of virtual machines and number of I/O requests increased (scaled up), with a particular interest in storage I/O bandwidth and response time plus database transactions. With a pattern established from the scale-up test cases, the addition of another cluster node would determine the impact or benefit of growing the cluster as application demand grows beyond that which can be supported by the initial configuration. If storage and application performance scales in a linear fashion, we can have some confidence in the ability of Nutanix, as backed by Dell EMC and Toshiba, to support workload growth over time in a real-world datacenter.

Metrics were collected within each virtual machine through Microsoft Windows Performance monitor and all metrics aggregated post collection.

Dell EMC XC servers with Toshiba 12Gb/s SAS SSDs Power High-Performance with Nutanix Hyper-converges solution

Results and Analysis

Storage bandwidth scaled in a linear fashion, delivering a predictable performance pattern across the three node cluster, and continuing to do so as the fourth node was added.

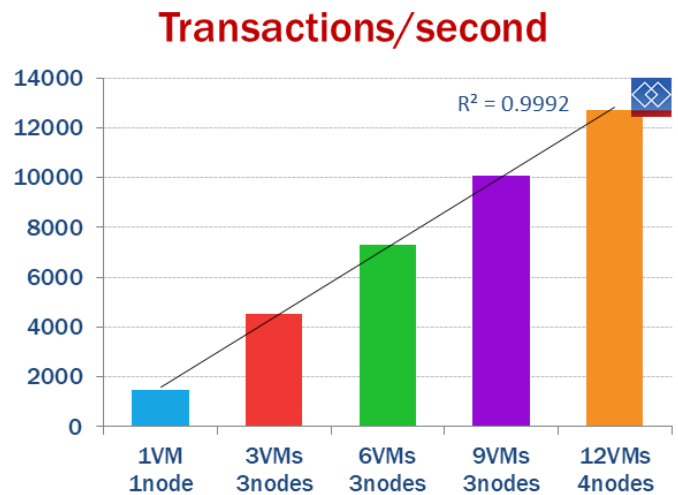


Calculating a simple linear regression of the I/O bandwidth data, an R^2 value¹ of very nearly 1.0 highlights how little variance there is between the actual data and the best fit trend line. This analysis allows us to reasonably predict performance anywhere along the trend line and project behavior beyond that line if the Nutanix cluster were scaled out to additional nodes.

The database applications have nearly 64GB of virtual memory to use as a cache. The job of those twelve SSDs is to get data off of media and into memory as quickly as possible. As fast as SSDs are, reads from memory are faster still and deliver an even better application user experience. As a general rule, databases prefer to build a large memory cache and operate as much as possible on data in the memory read cache rather than depend on storage, which until the advent of fast non-volatile memory such as flash drives, has been relatively slow. The throughput represented in this graph represents data that is either not in the large memory read cache

or represents data written by the application to the storage pool.

The actual work done by the cluster can be represented by the number of database transactions over time. As the workload scales up and out, a similar performance pattern can be seen as identified in the storage bandwidth data.



In fact, analysis of the scaling behavior of the number of database transactions processed each second as virtual machines and cluster nodes are added is even more predictable. Each three virtual machine increase is accompanied by a very precise increase of twenty seven to twenty eight hundred transactions per second². The R^2 value calculated in this metric's analysis also shows an extremely tight correlation to the trend line. We have no reason to believe that this trend would not continue if additional cluster nodes were provisioned.

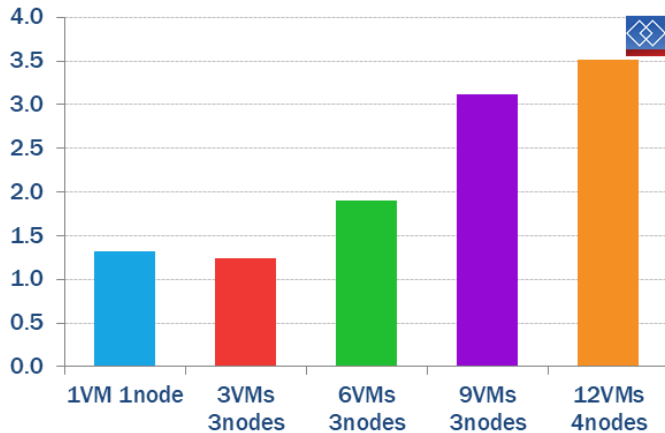
Application user experience will be impacted by response times. Our data collection methodology did not provide a metric for database transaction response time, but we can extrapolate from the I/O latencies recorded on the virtual drives provisioned from the Nutanix storage pool.

¹ In statistical modeling, R^2 , or r^2 , is known as the coefficient of determination and is used to evaluate how closely observed data correlates with a model.

² Given that different workload profiles generate different types and sizes of transactions, the total number of transactions is of lesser interest that the delta created by workload scaling.

Dell EMC XC servers with Toshiba 12Gb/s SAS SSDs Power High-Performance with Nutanix Hyper-converges solution

Avg Response Time (ms)



A single virtual machine saw an average storage I/O response time of roughly 1.25 milliseconds for the duration of the workload. As the number of virtual machines was scaled to three (one per cluster node) response time was essentially unchanged, indicating that Nutanix data locality is indeed minimizing cross-node reads and that the number of writes is taxing neither the network, drive controllers, or the SMB protocol at this level of I/O intensity.

Response times increased slightly to just under 2 milliseconds with two virtual machines per node and then to slightly over 3 milliseconds with three virtual machines running the workload on each node. Correlating this data with the storage bandwidth chart, we can be confident that the number of 8KB database I/O requests needed to generate 400 MB/s to 500MB/s is not going to stress the controllers or media of twelve SSDs per server. Remembering that Nutanix uses file shares for storage, we would suspect that we are seeing SMB protocol and network overhead in the slightly elevated latency as more data traffic begins to move across the wire. We note that with the addition of the fourth cluster node, the latency increase delta was much reduced, likely due to the addition of resources to balance out the increase of three more virtual machines—which is the whole point of scale-out architecture (i.e: expanding resources to support increasing demand). Had further nodes been added, we speculate that the latency curve may have flattened

around or below 4ms.

The question to ask now is what these response times mean to the application workload? This is ultimately a measure of whether business defined services levels are met or not. Legacy storage systems would be physically unable to achieve a 4ms response time for this workload, let alone anything under 2 ms (10 to 20 milliseconds would be more likely for magnetic hard disk drives used in this capacity, and possibly higher given the NAS storage protocol involved). By that measure, the SSDs appear to be giving a reasonably good account for themselves. Ordinarily, we would expect a chassis full of enterprise SSDs to deliver sub-millisecond response times, but given the nature of HCI and its requirement to distribute data across a cluster over Ethernet, we're satisfied with these results. Low millisecond I/O response times will translate into very fast application response times and we speculate that our hypothetical users will be satisfied.

Dell EMC XC servers with Toshiba 12Gb/s SAS SSDs Power High-Performance with Nutanix Hyper-converges solution

Summary and Conclusion

HCI takes the traditional datacenter storage tier, often expensive SAN and NAS arrays, and discards it in favor of consolidating local server storage into virtual pools to share across scalable clusters of servers. As demand for capacity or performance increases, servers are added to the cluster and drives to the pool. The simplicity of scaling, ease of use, and speed of deployment make HCI appealing to many businesses.

Nutanix is a key player in this space and their specific brand of HCI includes features to help them stand out from the crowd. Multi-hypervisor support, data locality, the Prism management interface, and Dell certified hardware in the form of PowerEdge XC-series servers with Toshiba PX05S 12Gb/s SAS SSDs are just a few of the reasons why Dell customers are giving Nutanix serious consideration as an HCI platform.

To get the most out of an HCI deployment, high performance servers and storage are a must. Dell has been an industry leader for a long time and the current generation of the Dell servers continues that trend. PowerEdge servers form an excellent backbone for Nutanix and the PowerEdge XC-series guarantees compatibility of server and storage, making deployment

of a virtual datacenter quick and easy. With powerful Toshiba SSDs in a current generation Dell XC-series server, enterprise applications like Microsoft SQL Server perform and scale in such a linear fashion that data engineers and architects are able to accurately predict the size and configuration of a suitable Nutanix cluster.

Based on our experience with other similar clustered virtual machine scaling environments, the Toshiba SSDs are a key contributor to the overall linear scaling that we observed. From the trends shown in results from this test, we expect that this type of configuration could support a substantially larger cluster configuration.

As HCI evolves, we may well see more and more traditional datacenters transition into private or hybrid clouds. Whether this becomes a dominant services tier, or another tool the IT toolbox can be debated. What's clear is that enterprises desiring to deploy a Nutanix cloud can count on Dell EMC to supply the hardware to power it. Contact your Dell EMC sales representative for more information about Dell EMC's commitment to the Nutanix platform and more information about the Dell EMC/Toshiba/Nutanix partnership.

The most current version of this report is available at www.demartek.com/Dell-XC-Toshiba on the Demartek website.

Dell and PowerEdge are trademarks of Dell, Inc.

Toshiba is a registered trademark of Toshiba, Inc.

Nutanix is a registered trademark of Nutanix, Inc

Demartek is a registered trademark of Demartek, LLC.

All other trademarks are the property of their respective owners.