Intel® Solutions
Reference Architecture

Cost Effective Scalable Storage for the Enterprise

Intel® and VMware* Deliver Software-Defined Storage Solutions

By John Hubbard

Intel® Xeon® Processor E5-2600 v3 Product Family

Audience and Purpose

This solutions blueprint provides reference architecture information for those striving to address, and scale with, tomorrow's Enterprise application and storage requirements utilizing existing infrastructure along with cost-effective, off-the-shelf Intel components. This document outlines Intel's reference architecture joined with VMware* Virtual SAN* (VMware's Software Defined Storage) optimized for mid-range tier-two workloads and test/development environments. The workload is made up of multiple active virtual machines (VMs) reading and writing storage in a completely random fashion. Intel's reference architecture, based on Intel® Processors, Network Adapters and Solid State Drives, creates a hyper-converged, multi-host, resilient, distributed storage environment. This paper describes the topology, hardware, software, best practices and performance metrology intended to simplify the deployment and management of cost-effective scalable storage.

Please note that application requirements can vary significantly and it is up to the reader to design a storage environment to best suit their needs.

This document assumes the reader has a basic knowledge of VMware ESXi and vCenter* Server.

January 2015

Table of Contents

Executive Summary

With the industry shift towards the Software Defined Data Center (SDDC) there is a growing demand for resilient solutions that scale quickly with minimal effort. It has been begrudgingly accepted that scaling Data Center infrastructure is neither simple, nor timely and requires extensive collaboration and investment. Collaboration between the necessary engineers (application, network, storage, virtualization, server, electrical, HVAC), managers (Data Center and project), and other essential personnel is a daunting task on its own. Considering the collaboration challenge, the trend toward reductions in IT spending, and the potential that solutions to one problem may create others, scaling can be a very daunting endeavor.

As digital content growth continues to accelerate, companies are striving to cope by scaling-out storage infrastructure. VMware's Virtual SAN combines Data Center servers with storage and networking in the same nodes. This hyper-convergence enables companies to scale bandwidth, performance, compute and storage, with a lower total cost of ownership (TCO) and with minimal effort. This solutions blueprint outlines scale-out reference architectures built on Intel technology powering VMware Virtual SAN for various usage models such as VDI and Test/Dev.

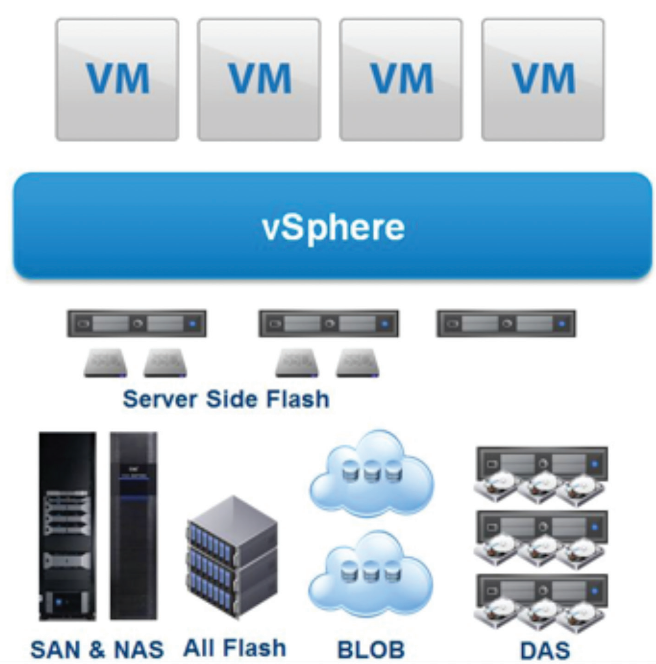
Introduction

IDC reports that the digital universe is doubling in size every two years, and by the year 2020, there will be more than 44 zettabytes of digital information worldwide.[1] For enterprises to remain successful amidst the onslaught of massive data, IT departments need storage solutions that are flexible, scalable, and cost effective. Unfortunately, traditional storage architectures often fall short of delivering on these capabilities, causing bottlenecks, underutilization, and application workload inefficiencies. Traditional storage methods also require a significant investment, create complex management tasks for IT departments, and lack the scalability required to keep up with data demands. Moreover, traditional storage systems are not easily re-provisioned on the fly, and they don't have the flexibility needed to adapt to fast-changing user and business demands. Server virtualization has transformed the way that IT is managed and delivered as it simultaneously places new demands on storage in terms of input/output (I/O) performance, latency, and scalability.

To address these storage challenges, Intel and VMware are working together to provide foundational technologies that deliver intelligent storage solutions based on the Intel® Solid-State Drive (SSD) Data Center (DC) Family, 10 gigabit Intel® Ethernet Converged Network Adapters, Intel® Xeon® processor E5 family, and VMware Virtual SAN. The software-defined storage of Virtual SAN, paired with Intel technologies, can transform the data center from isolated systems with limited scalability and inefficient utilization to a dynamic, fully automated, and flexible infrastructure that encompasses servers, storage, and network components.

VMware* Virtual SAN Product Overview

Virtual SAN, embedded inside the VMware vSphere* kernel, utilizes a highly available architecture that can withstand failures at the disk, server, and network level with no data loss, using built-in redundancy mechanisms that transparently store multiple copies of the data across disks and host servers. Virtual SAN uses a policy-based approach to storage management that allows IT administrators to specify storage attributes—such as capacity, performance, and availability—in the form of simple policies that are associated with individual virtual machines or virtual disks. Storage can then be instantly provisioned and automatically configured according to the assigned policies. Virtual SAN also dynamically self-tunes and load-balances to meet the policies of each individual virtual machine, thereby adapting



to ongoing workload condition changes and helping to make sure that service-level agreements (SLAs) are met throughout the virtual machine lifecycle.

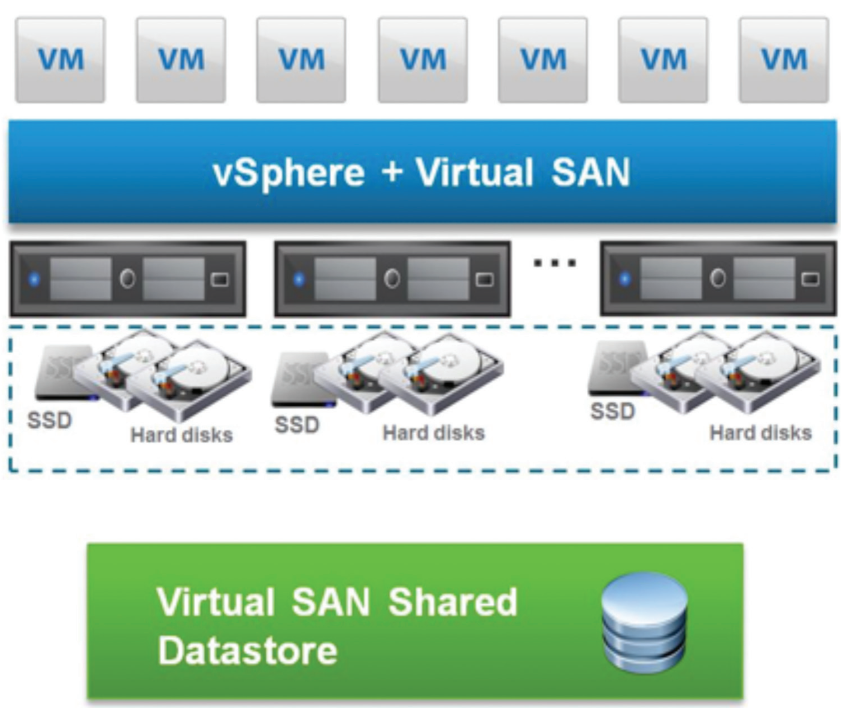
Virtual SAN makes it easy to provision and manage storage for virtual machines. IT professionals can quickly provision storage directly from the VMware vSphere Web Client with just a few clicks, taking advantage of a self-tuning system that automatically optimizes itself to deliver the appropriate SLAs based on the requirements for each virtual machine—without any downtime.

Unlike traditional storage arrays, Virtual SAN does not require large upfront investments to get started. A Virtual SAN Datastore can be created with as few as three servers and can be used on any server based on standard Intel® Xeon® processors.

Virtual SAN also delivers a lower total cost of ownership (TCO) by optimizing available performance and capacity on standard servers. Plus, Virtual SAN shares hardware resources used for capital expenditures, power, and cooling, which reduces energy costs. This means that existing high-dollar SAN devices can be freed up to perform the mission-critical work for which they were originally intended. Meanwhile, Virtual SAN enables organizations to increase their storage performance and capacity incrementally and predictably.

A High-Performance Storage Foundation

When paired with VMware Virtual SAN, Intel technologies like the Intel® Xeon® processor E5 family, 10 gigabit Intel Ethernet Converged Network Adapters, and the Intel SSD® Data Center Family, create the foundation for innovative software-defined storage solutions.



Intel® Xeon® Processor E5 v2 and v3 Families

VMware Virtual SAN delivers an efficient, scalable, and cost effective software-based storage solution, but running multiple virtual machine workloads—whether compute or storage workloads—on virtualized servers requires sufficient processing capability and increased I/O bandwidth, as well as additional memory. The Intel® Xeon® processor E5-2600 v2 product family is designed to overcome this challenge with up to eight cores. The Intel® Xeon® processor E5-2600 v3 product family improves upon v2 with up to 18 cores and DDR4 Memory support with up to 68 GB/s of bandwidth. Both families provide built-in Intel® Virtualization Technology (Intel® VT) to further improve on the versatility of software-based storage solutions.[2]

Another built-in technology, Intel® Integrated I/O, helps remove server, storage, and network bottlenecks to decrease latency and increase data throughput while enabling energy-efficient performance for the most demanding workloads.

Running VMware Virtual SAN on servers built with Intel® Xeon® processors can reduce an enterprise's total cost of ownership (TCO) while increasing flexibility and scalability with intelligent, automated storage provisioning and management.

Intel® Solid-State Drive Data Center Family

Intel SSDs have emerged as the ideal storage solution for data centers. SSDs have no moving platters or actuator arms that can fail, use semiconductor-based, non-volatile memory, which makes them superior to hard-disk drives (HDDs), and produce less heat and noise than HDDs.

The Intel® SSD Data Center Family can retain data even with unexpected loss of power and balances fast read/write speeds with optimized CPU utilization. The Intel® SSD Data Center Family typically has about one-third of the average latency and 1,300 times the performance of HDDs.[3] In particular, the Intel® SSD DC S3700 Series and Intel® SSD DC P3700 Series offer full end-to-end data protection, consistent performance with low latencies, AES 256-bit encryption for enhanced data protection, high write endurance, and high capacities for growing storage needs.

VMware designed Virtual SAN to harness SSD capabilities for high performance read caching and write buffering. By clustering server direct-attached storage with Intel SSDs, Virtual SAN creates a distributed, shared datastore at the hypervisor layer that is designed and optimized for virtual machines. As the needs of virtualized applications change, the hypervisor is uniquely positioned to make I/O optimizations and intelligent data-placement decisions to optimize application performance.

Intel® Ethernet Converged Network Adapters

Based on Intel's unified networking vision for supporting all LAN data and storage traffic on a common Ethernet infrastructure, the 10 gigabit Intel® Ethernet Converged Network Adapter X520 product family can dramatically lower latency and increase the data center's performance versus traditional storage architectures. Intel® X520 adapters provide twice the server bandwidth, which can reduce the necessary energy to power each rack by up to 45 percent. Intel® X520 adapters provide the scalable, high-throughput features necessary to meet the demands of combined compute and storage workloads on Virtual SAN systems while reducing infrastructure costs by 15 percent.[4]

It is considered best practice to ensure that sufficient PCIe lanes are available to the network cards and their ASICs to prevent the artificial limiting of network performance. Per PCI-SIG, a single PCIe Gen3 lane supports 1 GBps or 8 Gbps, thereby doubling the bandwidth of PCIe Gen2. Thus a single 10 gigabit port is artificially limited to 8 Gbps with a single PCIe Gen3 lane. Having a minimum of four PCIe Gen3 or eight PCIe Gen2 lanes available to dual 10 gigabit network cards is recommended.[5]

Intel Reference Architectures for VMware Virtual SAN

All reference architectures support the same base build which includes the network adapter, storage controller and OS Disk. Their differences are in the processor family, platform, memory, storage configuration and VMware ESXi Build. Please note that ESXi 5.5 Update 2 was selected for the new environments whereas ESXi 5.5 Update 1 was selected for existing environments. This decision was based on the expectation that new environments are likely to utilize the newest available version of ESXi whereas established environments may not.

Existing Environments with Intel® Xeon® E5 v2 Family

CPU: Intel® Xeon® E5-2690 v2
Memory: 128GB DDR3 RAM
Motherboard: Intel® 2600GZ Server Board
Network: Intel® Ethernet Server Adapter X520-DA2
OS Disk: Intel DC S3700 SSD
Storage Controller: LSI 9207-8i HBA
VSAN Storage: Hybrid

Emerging Environments with Intel® Xeon® E5 v3 Family

CPU: Intel® Xeon® E5-2699 v3
Memory: 128GB DDR4 RAM
Motherboard: Intel® S2600WT Server Board
Network: Intel® Ethernet Server Adapter X520-DA2
OS Disk: Intel DC S3700 SSD
Storage Controller: LSI 9207-8i HBA
VSAN Storage: Hybrid

High Performance “All Flash” Environments with Intel® Xeon® E5 v3 Family (Experimental configuration)

CPU: Intel® Xeon® E5-2699 v3
Memory: 128GB DDR4 RAM
Motherboard: Intel® S2600WT Server Board
Network: Intel® Ethernet Server Adapter X520-DA2
OS Disk: Intel DC S3700 SSD
Storage Controller: LSI 9207-8i HBA
VSAN Storage: “All Flash”

Virtual SAN Storage Configurations

Two fundamentally different storage configurations were used to demonstrate the varying performance benefits between the reference architectures.

Hybrid storage is the combination of solid state drives and rotational hard disk drives. Virtual SAN was designed around hybrid storage topologies, utilizing the speed and efficiency of solid state drives with benefits of large capacity rotational media.

All Flash storage configurations are comprised solely of solid state drives and offer very high performance and reliability. Substituting SATA solid state drives for rotation media not only dramatically changes the performance characteristics; it reduces management complexity and enables more applications to coexist. Although VMware does not officially support All Flash Virtual SAN configurations, these configurations do make excellent proof of concepts and development environments.

Test Configuration

Test Topology and Methodology

Virtual machines are spread out across all four hosts. Each VM runs dynamo and connects across a non-routed private network to a remote system acting as a harness or controller. The harness system is running IOMeter and directing the VMs to generate disk IO to their virtual disks.

A total of five different tests are run on each Virtual SAN cluster. The first test consists of one powered-on VM on each of four hosts generating disk IO, for a total of four powered-on VMs. The second test doubles the number of powered-on VMs to two per host, for a total eight powered-on VMs. This doubling trend continues until each of the four hosts has 16 powered-on VMs, for a total of 64 IO generating, powered-on VMs in the fifth test.

Each version of the test was run three times and the results were averaged. Before each new test version, the Write Buffer and Read Cache were cleared from the Virtual SAN Datastore on each host. Clearing the Write Buffer and Read Cache between test versions is necessary to prevent previous tests’ information from altering or ‘coloring’ the next set of tests. The Write Buffer and Read Cache were not cleared between the three instances of the same test, but were cleared between the different tests. For example, when the third instance of the four-VM test was completed the Write Buffer and Read Cache were cleared and the first instance of eight-VM test began.

- Storage Policy
  - Number of Disk Stripes: 4 Disks
  - Number of Failures to Tolerate: 1
- Windows 2008 R2
- 2 vCPU, 4GB vRAM
- VMDKs*
  - OS Disk: 40GB Thin Provision
  - Data Disk: 7.5GB Thick

Virtual Machines

A "golden VM" was built and then cloned and distributed evenly across the hosts and Virtual SAN Datastore.

The Storage Policy for each VM is also the same; striping each of the VM's virtual disks or VMDKs across all Datastore disks on the host.

The 2nd policy "Number of Failures to Tolerate: 1", enables the VM to survive one failure such as a host going offline. This 2nd policy duplicates every IO generated by the VM to an additional host. This means that every write IO is to two hosts and every read IO can be serviced by either host. This policy in particular comes at a penalty; sacrificing performance by nearly doubling the write time for the sake of resiliency.

Intel® IOMeter 1.1

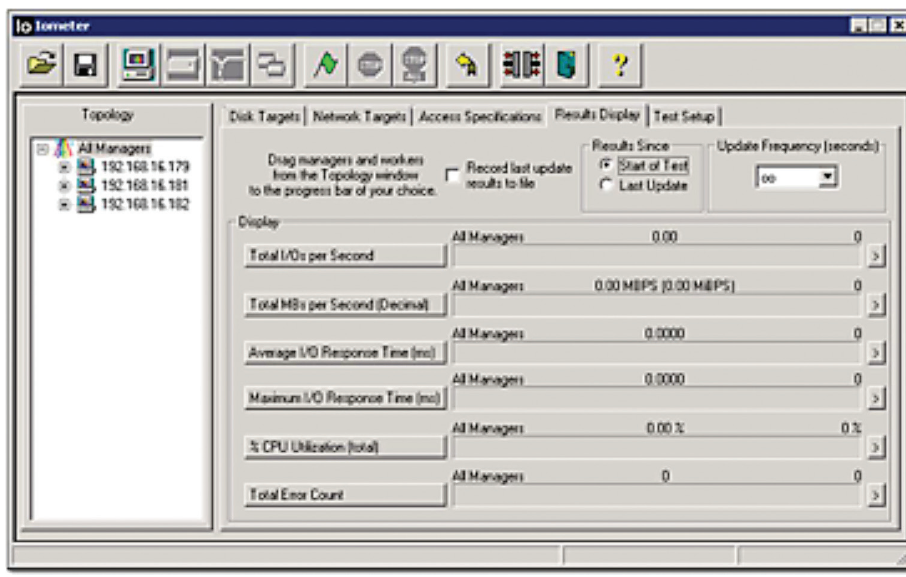
All IO was generated by Intel using Intel® IOMeter version 1.1. The workload is 100% Random, 70% Read (30% Write) with a 4K transfer size. The span size is set to 7.5 GB or the size of the data disk attached to the VM. The workload ramps for 30 minutes and then runs for 90 minutes.

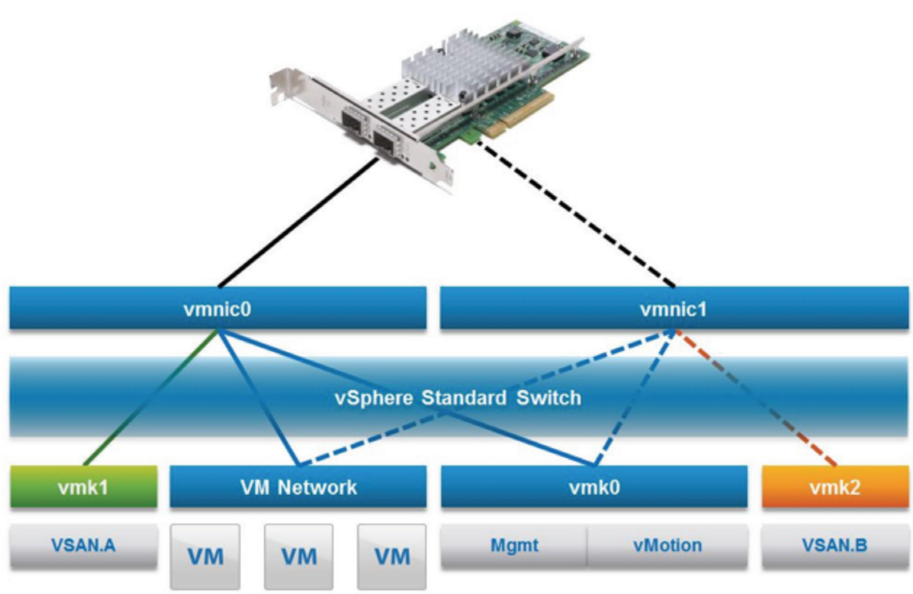
Virtual SAN Host Networking

One vSwitch is used to connect both 10 Gb vmnics as uplinks. Four port groups, three of which are VM Kernel Adapters (or VMKs) are used to carry the varying traffic. Different VLANs are also used to segment traffic to the respective port groups and VMKs. Utilizing two 10 Gb links to carry all traffic greatly simplifies the network requirements by reducing the number of network adapters, network cables, and switch ports required.

The VM Network port group and vmk0 are configured to load balance across either vmnic using the "Route based on originating virtual port" method. All management and vMotion traffic is carried across vmk0.

Virtual SAN traffic is split between two VMKs; vmk1 and vmk2, each on a different VLAN. The default failover order has been modified to include one vmnic as Active and the other as Unused. For example, vmk1 has identified vmnic0 as Active and vmnic1 as Unused, vmk2 has identified vnmnic0 as Unused and vmnic1 as Active. This configuration also lends itself to iSCSI port binding and MPIO if iSCSI targets reside on the same Virtual SAN VLANs.





All Flash Virtual SAN††

While the idea of a Virtual SAN Datastore sounds very powerful, Virtual SAN requires that a host have a minimum of one SSD for cache and one HDD for storage which precludes All Flash or All SSD Virtual SAN configurations. However, because it is possible to designate how disk drives are reported in ESXi it is possible to modify certain rules, thereby allowing you to identify SSDs as HDDs. The following commands mark all drives matching a specific model, this in turn creates the "Disable SSD" Rule.

```
# esxcli storage nmp satp rule add -s VMW_SATP_LOCAL -d naa.5001517972e66581 -o disable_ssd
# esxcli storage core device list -d naa.5001517972e66581 | grep SSD
   Model: INTEL SSDSC2BA40
   Is SSD: true
# esxcli storage core claiming unclaim -t device -d naa.5001517972e66581
# esxcli storage core claimrule load
# esxcli storage core claimrule run
# esxcli storage core device list -d naa.5001517972e66581 | grep SSD
   Model: INTEL SSDSC2BA40
   Is SSD: false
```

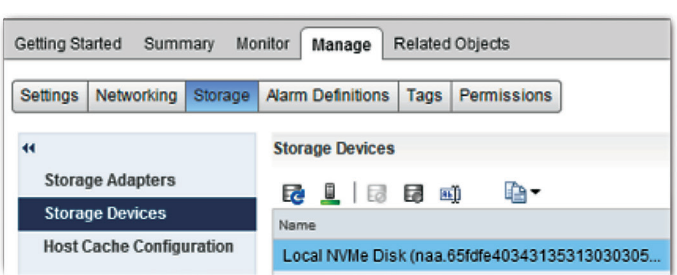
Determining Disk Drive IDs

The IDs of the drives can be retrieved in multiple locations either with the vSphere Clients or the ESXi command line. With the vSphere Web Client, one method is to navigate to the "Manage" tab of an ESXi host, and select "Storage Devices" under the "Storage" section.

Using the ESXi command line, multiple commands are available. Some of those commands are listed below.

```
ls /dev/disks
esxcfg-scsidevs -l
esxcli storage core device list
```

†† Experimental Configuration, not supported by VMware



ESXi Versions and Optimizations

The Intel® Xeon® E5 v2 Family is using ESXi 5.5 Update 1 (Build 1623387). This version was publicly released on March 11th, 2014. Update 1 was chosen to run on the Intel® Xeon® E5 v2 Family as it represented the most likely choice for existing deployments.

The Intel® Xeon® E5 v3 Family is using ESXi 5.5 Update 2 (Build 2009162). This version was publicly released on September 9th, 2014. Update 2 was chosen to run on the Intel® Xeon® E5 v3 Family as it represents the most recent version available and the most likely choice for new hardware deployments. VMware had made some adjustments to the Virtual SAN code in this release. These changes affected congestion limits and recovery in the event of a host failure and were designed with smaller SSDs in mind. Intel worked with VMware's VSAN performance team and isolated two commands to improve performance. Those commands are listed below and do not require a reboot to take effect.

```
esxcfg-advcfg -s 8 /LSOM/lsomLogCongestionLowLimitGB
esxcfg-advcfg -s 12 /LSOM/lsomLogCongestionHighLimitGB
```

NVMe Support

Currently NVMe devices require an NVMe driver to be supplied by the drive manufacturer. Intel has developed an NVMe driver compatible with VMware ESXi 5.5. This driver is experimental and not supported by VMware at this time. The Intel development driver is available to Intel partners and customers upon request.

Intel is actively working to certify Intel® PCIe solid-state drives and the Intel® NVMe driver with VMware for ESXi 5.5 and future releases. Once certified, the Intel® SSD DC P3700 Series and others will be listed on the VMware Compatibility Guide and NVMe driver publicly available.

Performance

Performance of each Virtual SAN cluster was carefully measured and reported. Before presenting the performance values of the outlined reference architectures, it is important to understand how performance is measured and how it is to be interpreted.

Performance Metrology

Below is a reference of how the information is represented and its significance.

IO: An Input/Output Operation; either a read or write.

IOPS: Input/Output Operations per second; reflects the raw number of transactions that complete in one second.

Latency: Latency reflects the round trip time it takes for a single transaction to complete and is measured in milliseconds (ms).

Round Trip Time: A round trip time is characterized by the time it takes to generate an IO, propagate that IO, process that IO, and acknowledge that the IO transaction has completed.

QoS: Quality of Service; represented as the percentage of total IOs measured in a given time frame that complete within a specific time. For example, 50% of the IOs generated over the course of one second completed in one millisecond or less.

Service Level: A service level is a guarantee of raw performance within a given QoS. For example, the environment can generate 10,000 IOPS which complete in 10ms or less 99.99% of the time.

Interpreting the Graphs

IOPS: The higher the value the better.

Quality of Service: Values are differentiated by color representing the number of concurrent VMs and are plotted in as (x,y) coordinates. The x-axis (horizontal) is measured in milliseconds and represents the round trip time of an IO. The y-axis (vertical) is measured as the percentage that an IO completes within a given time as compared to the total number of IOs.

Higher performance guarantees or a percentage of IOs completing in less time, will plot closest the top left corner of the graph. Conversely, lower performance guarantees or a percentage of IOs requiring more time to complete, will plot closer to the lower right corner of the graph.

Quality of Service <2ms: The data set has been reduced for clarity to focus on transactions requiring 2ms or less to complete.

Quality of Service <30ms: The dataset has been reduced for clarity to focus on transactions requiring 30ms or less to complete.

Comparing Performance and Quality of Service across the three reference architectures.

1. Existing: Intel® Xeon® E5 v2, Intel® SSD DC S3700 + HDD
2. Emerging: Intel® Xeon® E5 v3, Intel® SSD DC S3700 + HDD
3. High Performance "All Flash" with NVMe: Intel® Xeon® E5 v3, Intel® SSD DC P3700† + Intel® SSD DC S3500††

† NVMe support coming soon

†† Experimental Configuration, not supported by VMware

Existing Environments

Hybrid Storage

Intel® Xeon® E5 v2 Family

SSD Cache

4x 800GB Intel® SSD DC S3700 Series
3.2 TB Total SSD Capacity

Datastore

16x 1.2 TB SAS 10K RPM HDD Seagate ST1200MM0017
16 TB Total Capacity (8 TB Usable)

Performance[6]

Throughput

The four node Virtual SAN cluster maintains linear performance gains and peaks at over 11,500 IOPS with 64 concurrent VMs.

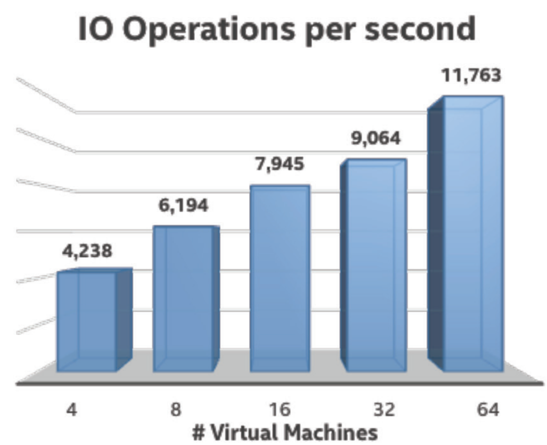
Quality of Service

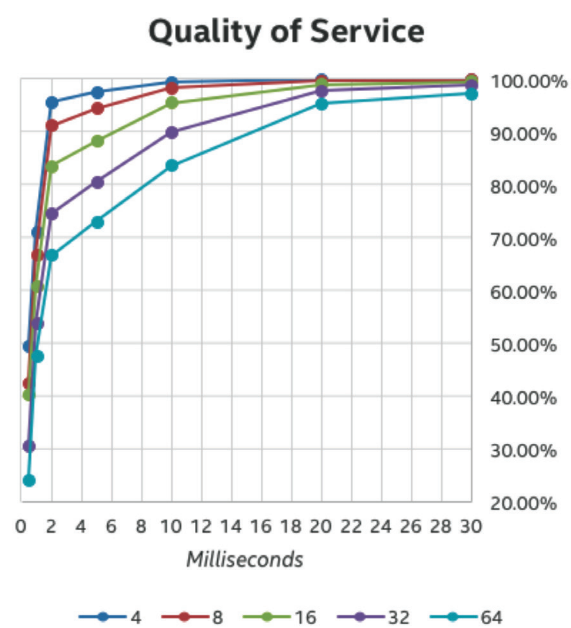
At 1ms or less the cluster can guarantee service levels at 71% with four concurrent VMs. As the number of concurrent VMs increase service levels decrease between roughly three and seven percent.

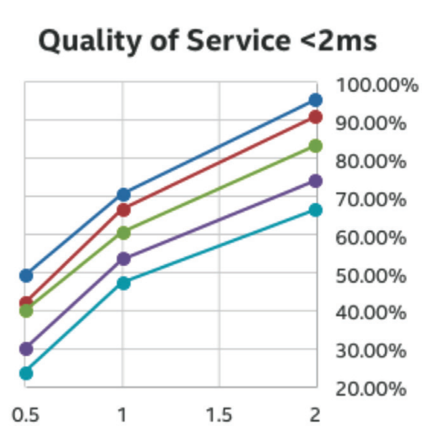
At 2ms or less the cluster can guarantee service levels at 95.5% with four concurrent VMs. As the number of concurrent VMs increase, QoS decreases between four and eight percent dropping to 66.6% with 64 concurrent VMs.

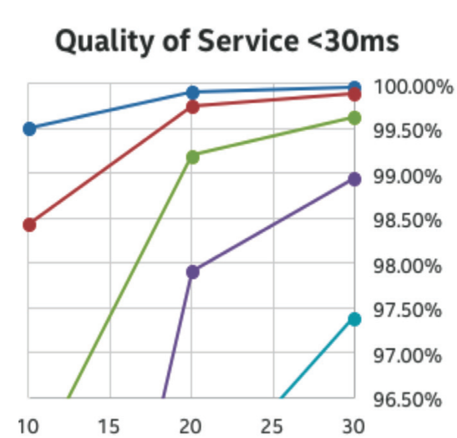
At 10ms or less the cluster can guarantee service levels at 99.5% with four concurrent VMs. With 64 concurrent VMs QoS decreases by 16% reaching a service level of 83%.

At 30ms or less service levels reach 99.9% and are much tighter varying by 2.5% between four and 64 concurrent VMs.









Emerging Environments

Hybrid Storage

Intel® Xeon® E5 v3 Family

SSD Cache

4x 800GB Intel® SSD DC S3700 Series
3.2 TB Total SSD Capacity

Datastore

16x 1.2 TB SAS 10K RPM HDD Seagate ST1200MM0017
16 TB Total Capacity (8 TB Usable)

Performance[6]

Throughput

The four node Virtual SAN cluster maintains linear performance gains and peaks at over 17,000 IOPS with 64 concurrent VMs. An improvement of 67% or 5,500 IOPS over the hybrid storage with the Intel® Xeon® E5 v2 Family.

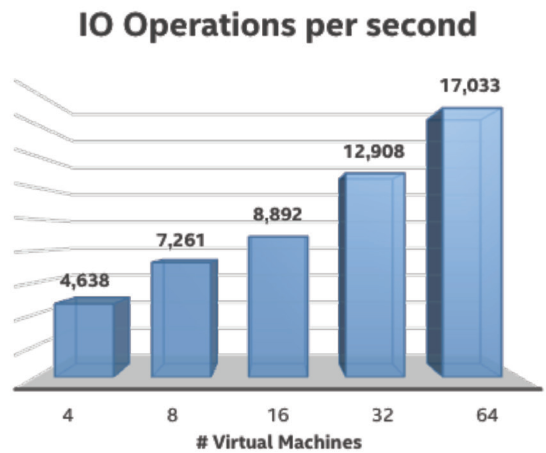
Quality of Service

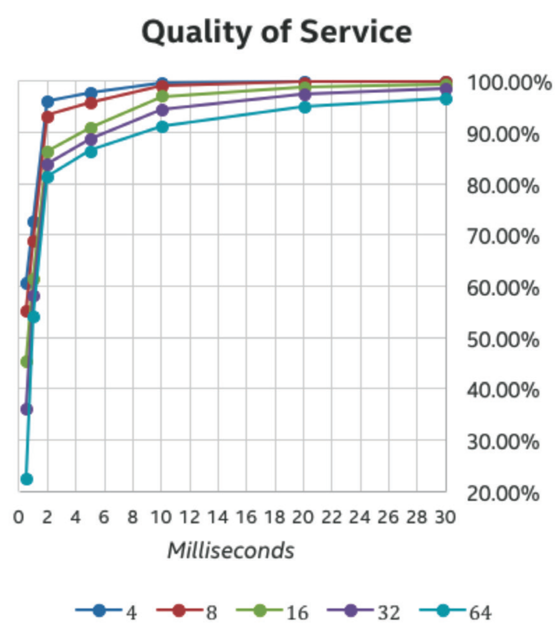
At 1ms or less the cluster can guarantee service levels at 72.5% with four concurrent VMs. As the number of concurrent VMs increase exponentially service levels decrease between roughly three and five percent.

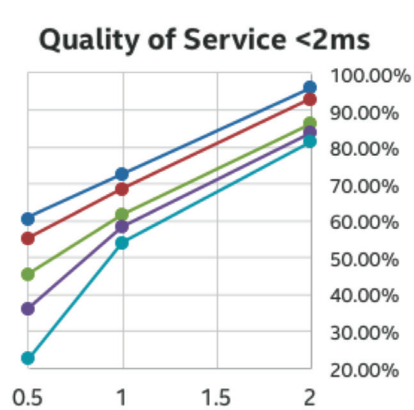
At 2ms or less the cluster can guarantee service levels at 96.2% with four concurrent VMs. As the number of concurrent VMs increase exponentially, QoS decreases between two and seven percent dropping to 81.4% with 64 concurrent VMs.

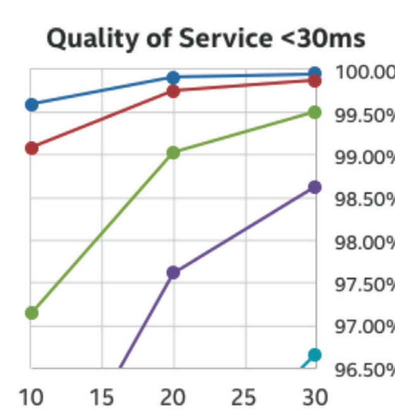
At 10ms or less the cluster can guarantee service levels at 99.5% with four concurrent VMs. With 64 concurrent VMs QoS decreases by 11% reaching a service level of 86%.

At 30ms or less service levels reach 99.9% and are much tighter varying by 3.3 % between four and 64 concurrent VMs.









High Performance "Experimental"

All Flash

Intel® Xeon® E5 v3 Family

SSD Cache

4x 800GB Intel® SSD DC P3700 Series with NVMe
3.2 TB Total SSD Capacity

Datastore

16x 800GB Intel® SSD DC S3500 Series
12.8 TB Total Capacity (6.4 TB Usable)

Performance[6]

Throughput

The four node Virtual SAN cluster maintains nearly linear performance gains until 32 concurrent VMs and peaks at over 39,500 IOPS with 64 VMs. An improvement of 232% or 22,500 IOPS over hybrid storage with the same Intel® Xeon® E5 v3 Family.

Quality of Service

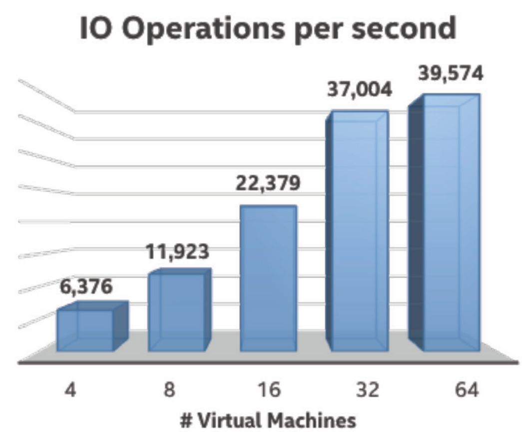
At 1ms or less the cluster can guarantee service levels at 76.5% with four concurrent VMs. As the number of concurrent VMs increase exponentially service levels decrease between roughly one and nine percent.

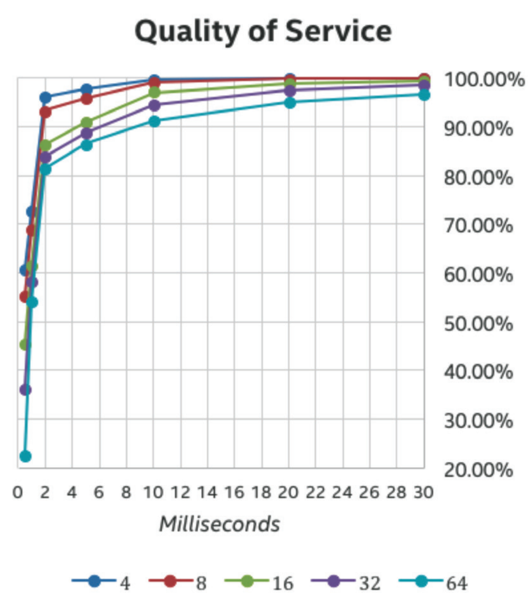
At 2ms or less the cluster can guarantee service levels at 99.3% with four concurrent VMs. With the exception of 64 VMs, as the number of concurrent VMs increase exponentially, QoS decreases between only one and three percent dropping to 86.2% with 64 concurrent VMs.

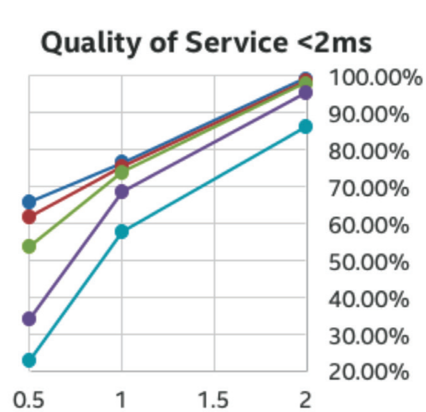
At 5ms or less the cluster can guarantee service levels between 99.8% and 99.5% with four and 32 concurrent VMs respectively. 64 concurrent VMs see a service level of 95.3%.

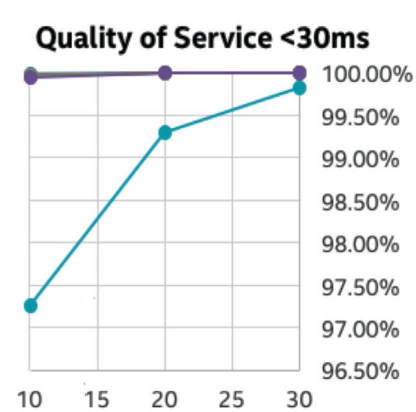
At 10ms or less the cluster can guarantee service levels between 99.98% and 99.95% with four and 32 concurrent VMs respectively. 64 concurrent VMs see a service level of 97.25%.

At 30ms or less service levels reach 99.997% and are much tighter varying by less than 0.2% between four and 64 concurrent VMs.









Three Reference Architectures

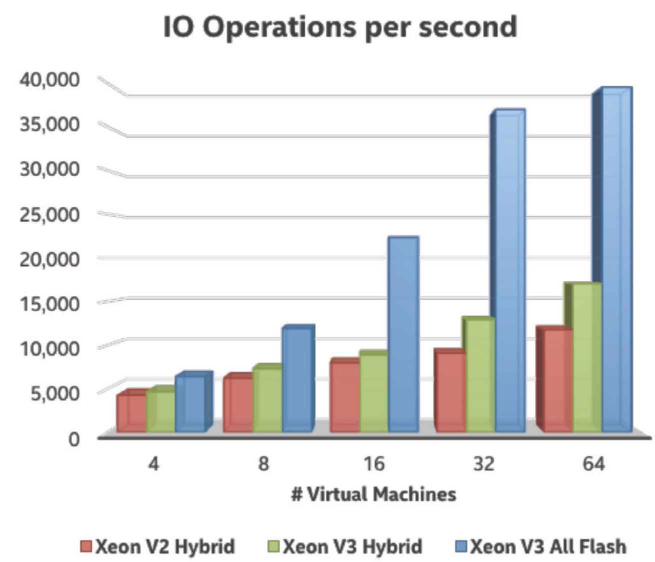
Performance[6]

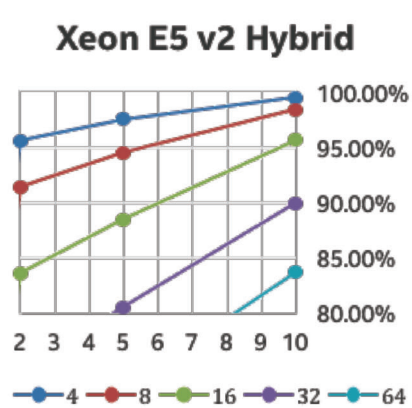
Throughput

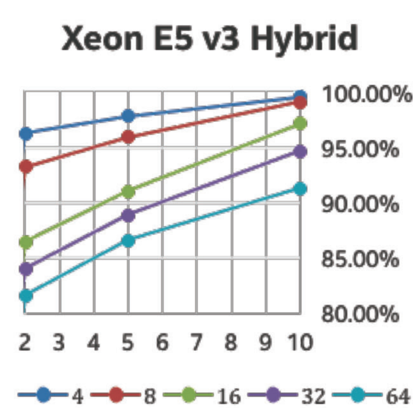
Each reference architecture's number of IOPS increase as the number of concurrent VMs increase exponentially. Comparing each Hybrid Storage reference architecture, the most value is realized with 32 and 64 concurrent VMs. With eight VMs or greater the throughput of the All Flash reference architecture performance increases dramatically and generates over 3x the IOPS with 64 VMs when compared to Hybrid Storage with the Intel ® Xeon® E5 v2 Family.

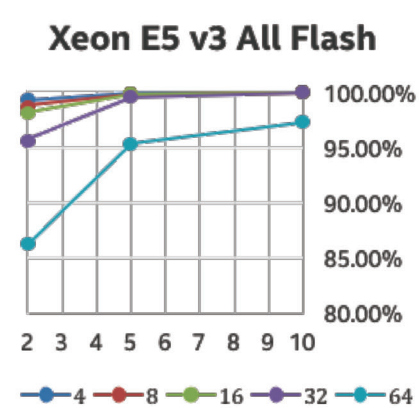
Quality of Service 2ms – 10ms

Applications with service level requirements requiring QoS less than two to 10 milliseconds will see most value between two and 10 milliseconds.









The Sweet Spot

The best overall balance between IOPS throughput and QoS is realized with the All Flash reference architecture. With 16 concurrent VMs, the four node Virtual SAN cluster produces over 22,300 IOPS with each transaction completing in two ms or less with a service level of 98%. At five ms or less that number increases to nearly three nines at 99.81%. With 32 concurrent VMs the IOPS nearly double to over 37,000 IOPS with each transaction completing in five ms or less with a service level of 99.6%. At 10ms or less service levels barely miss four nines with 32 concurrent VMs or less at 99.94%.

Synthetic Workloads and Applications

Not every application was created equal and as such these reference architectures should only be used as a baseline. The IOmeter workload used in this paper ran a single worker with a queue depth of one on each virtual machine. For all intents and purposes this represents a single threaded application that must wait for each IO to complete before queuing the next IO. Many applications are multithreaded and utilize queue depths greater than one. As such, those applications can generate more storage IO and receive greater throughput. The IOPS results in this paper could be increased, perhaps dramatically, if high queue depth workloads were used.

.

Conclusion

The three reference architectures outlined in this document offer tremendous performance and quality of service at varying service levels. Each architecture's hardware and software can also be modified to suit many application needs. For example, on the hardware side, if an application requires additional capacity, additional disks can be installed or even replaced with larger capacity ones. If that application also needs greater performance rotational media can be replaced with solid state drives.

On the software side, the administrator can leverage different VM Storage Policies and additional VSAN disk groups. A VM Storage Policy can be modified to isolate an inefficient application to fewer spinning disks, or the same policy could be used to increase the amount of read cache available to that VM. Additional VSAN disk groups, each requiring at least one SSD and one rotational disk, can be created to increase performance even greater than what was outlined in this paper.

For VDI environments that are read heavy and access the same content repeatedly, crafting a storage policy that allocates a higher percentage of read cache for the VMs would increase storage IO responsiveness and provide a more beneficial user experience.

Test and development (Test/Dev) environments may benefit greatly from Virtual SAN's relatively quick deployment time of hours, instead of the days or week of collaboration required when involving various Data Center staff.

Similarly, the DMZ environment can also benefit; not just due to deployment simplicity, but even more so due to the isolation of storage from the production environment. VMs that support services in the DMZ can have their storage segregated from the rest of the existing storage infrastructures such as Fibre Channel and SAN.

Looking forward, PCIe SSDs with NVMe can reduce latency and increase performance dramatically and with a better cost per IOPS. With the High Performance reference architecture, the Intel® SSD DC P3700 (PCIe) and Intel® SSD DC S3500 (SATA) more than double the performance of the same platform using Hybrid Storage and with 99% QoS service levels at 2ms.

Intel and VMware Deliver Better Data Storage Together

The partnership of Intel technology and VMware Virtual SAN provide a competitive edge through an acceleration of service delivery using an agile software-based storage solution. With Virtual SAN running on the Intel® Xeon® processor E5 v2 family, the Intel SSD Data Center Family, and 10 Gigabit Intel Ethernet Converged Network Adapters, enterprises can efficiently manage vast, ever-expanding storage requirements. While enterprise gains the loyalty of their customers through the fulfillment of mission-critical SLAs, greater data center flexibility, and scalability, Intel and VMware maintain the trust of enterprise by helping reduce initial investment costs and total cost of ownership.

1. IDC. The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things, April 2014. http://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm.
2. Intel® Virtualization Technology requires a computer system with an enabled Intel® processor, BIOS, and virtual machine monitor (VMM). Functionality, performance, or other benefits will vary depending on hardware and software configurations. Software applications may not be compatible with all operating systems. Consult your PC manufacturer. For more information, visit http://www.intel.com/go/virtualization.com.
3. Intel® Solid State Drive technology with VMware Virtual SAN delivers 2x the IOPS at 1/3rd the latency of hard disk drives. Source: "Unleashing the power of VMware's Virtual SAN on the latest industry standard high performance NVMe/PCIe SSDs", VMworld, August 2014.
4. Source: Max. I/O R/W bandwidth comparing Intel® Xeon® E5-2680 vs. Intel® Xeon® X5670 and 10 x 1GbE ports to 2 x 10GbE ports.
5. PCI Express* 3.0 Frequently Asked Questions, PCI-SIG: https://www.pcisig.com/specifications/pciexpress/resources/PCIe_3_0_External_FAQ_Nereus.pdf.
6. Data based on Intel® SSD DC S3700 and DC S3500 Series data sheets. Results have been estimated based on internal Intel analysis and are provided for informational purposes only. Any difference in system hardware or software design or configuration may affect actual performance. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as IOMeter*, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. Source: Internal Testing Configuration: See http://www.intel.com/ssd for detailed products specifications and testing configurations.

Disclaimers

Printed in USA 141216tlm

Order number 331407-002