WHITE PAPER
Bigtera VirtualStor™ Scaler
Intel® Xeon® Processor D-1500 Product Family
Intel® SSD Data Center Series
Topic: Software Defined Storage

Deploy Bigtera VirtualStor™ Scaler based on Intel® Xeon® Processor D-1500 Product Family Servers & Intel® SSDs

Bigtera VirtualStor™ Scaler leverages the Intel® Xeon® processor D-1500 product family and Intel® Data Center SSDs, to build the high-performance and cost-effective storage solution.

"Enterprise IT needs to deploy a scalable storage cluster optimized for performance and cost."

Execution Summary

Enterprise IT need a scalable, high-performance and cost-effective storage solution that can handle the rapid data growth and satisfy different types of workload. To address the challenges of enterprise IT to deploy a scalable storage cluster optimized for performance and cost, Bigtera* and Intel performed extensive testing to characterize optimized configurations for deploying Bigtera VirtualStor™ Scaler on Intel® Xeon® processor D-1500 product family servers & Intel® SSD.

Overview

Per IDC reports, the digital universe is doubling its size every two years, and by 2020, there will be more than 44 zettabytes, or 44 trillion gigabytes of data in the world[1]. This is having a substantial impact on enterprise IT requirements: enterprise IT will need storage solutions that are scalable, high performance, and cost effective. Unfortunately, these enterprises have outgrown their traditional infrastructures due to following challenges:

- **Business complexity.** As business processes get more sophisticated, different applications demand different characteristics for storage, such as protocols, performance, availability, etc. Thus, users must purchase different kind of storage systems across their workflows, creating "storage islands," where these systems have required high implementation costs, but have low utilization.

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.

- **Rapid data growth.** Taking media and broadcasting industry as example: The demand for high definition and 4K video continues to rise, placing pressure on storage capacity and scalability. A 90-minute standard definition TV program can consume 540 GB of space, while a high-quality 90-minute reality show's raw materials might consume terabytes (TBs) of storage. Because the traditional storage systems are unable to seamlessly scale-out, storage islands continue to grow, further increasing the complexity of managing infrastructure, data, and content.
- **The silo effect.** In many enterprises, digital data should move with the workflow. With multiple storage islands, data must be manually copied between heterogeneous environments, increasing workflow complexity, duplicating data, and wasting time and resources.

To address these challenges, Bigtera* and Intel are working together to provide the software-defined storage solutions that are based on the Intel® Xeon® processor D-1500 product family and Intel® Data Center SSDs, and Bigtera VirtualStor™ Scaler. Bigtera* software-defined storage solution, paired with Intel technologies, can promise the scale-out capacity with economics and high performance.

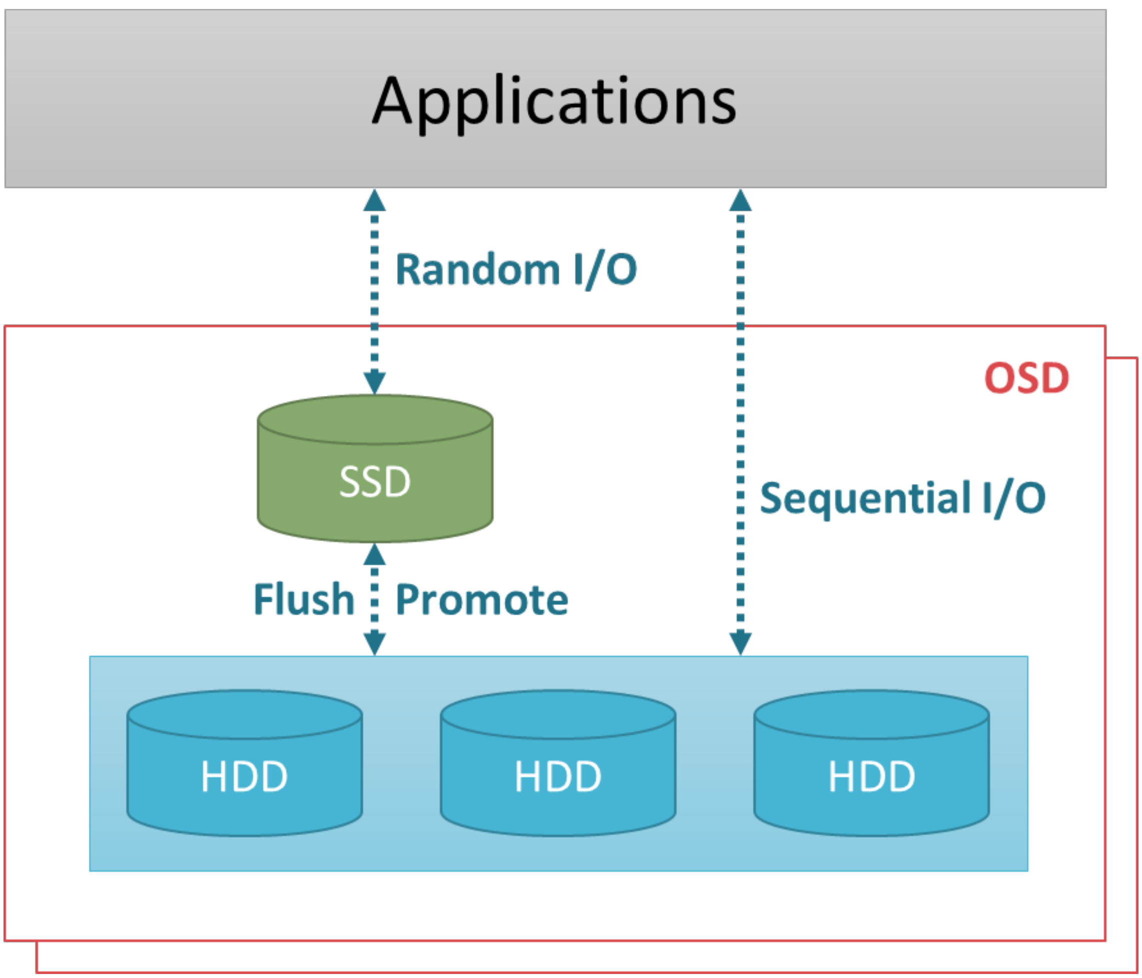

Figure 1. Caching data flow for VirtualStor™ Scaler.

Bigtera VirtualStor™ Scaler product overview

VirtualStor™ Scaler, deployed with a bare metal Intel® Xeon® processor D-1500 product family server, provides customers with a high-performance scale-out storage solution that allows them to pay as they grow. Running on Intel® technologies, VirtualStor™ Scaler utilizes a highly scalable and available architecture, and provides the flexibility to specify the storage type (Network Attached Storage (NAS), Storage Area Network (SAN), or object storage), performance (input and output operations per second (IOPS) and throughput), and efficiency, all while delivering resilient and secure capacity.

Caching within Object Storage Device (OSD) (Figure 1)

As the world leader in software-defined storage technology, Bigtera* revolutionized the storage industry by a new approach to leverage the high-performance solid state drive (SSD) to optimize the performance for random input and output operations (I/O).

Unlike other software-defined storage providers put many SSD OSD together into a cache pool, Bigtera VirtualStor™ Scaler provides the OSD caching ability that can deliver the random I/O caching by taking advantages of high-performance SSD. In addition, the sequential I/O will skip the SSD and directly read/write the hard drives. Usually the SSD is a magnitude faster than the spinning hard drives for random I/O, but with less advantages in sequential read/writes. Therefore, VirtualStor™ Scaler can intelligently serve random I/O and sequential I/O differently, to optimize the overall performance.

Intel® Xeon® Processor D-1500 Product Family

Bigtera VirtualStor™ Scaler delivers a scalable, high-performance software-defined storage solution, while serving multiple kinds of storage types (NAS, SAN, or object storage) with efficiency requires sufficient processing capability on a cost-effective hardware platform. The Intel® Xeon® processor D-1500 product family is designed to overcome this challenge. It is Intel's 3rd-generation 64-bit system-on-a-chip (SoC) and the first Intel® Xeon® processor D-1500 product family SoC based on Intel 14nm silicon technology. It delivers balanced computing performance with low power consumption. It also can support both extreme storage density and exceptional power efficiency to hyper-scale cloud data-centers.

Balance the performance and power consumption

The Intel® Xeon® processor D-1500 product family boasts hardware and software scalability from two to eight cores in a TDP of ~19 to 65 W. Utilizing similar development tools and processes as Intel® Core™, Intel® Atom™, and other Intel® Xeon® processors, the Intel® Xeon® processor D-1500 product family delivers broad application compatibility and software consistency from the data center to the edge. The reliable Intel x86 64-bit software support helps save time, cost, and validation.

It's an ideal platform for supporting a complete product line that meets a range of IT requirements, including Intel Platform Storage Extensions, which allow smarter and more cost-effective storage solutions through integrated technologies that accelerate data movement, protect data, and simplify data management.

Intel® Solid State Drive Data Center Family

Intel® SSD Data Center family drives can offer high performance with low latency and high endurance for enterprise and service provider needs. NVM-based Intel® SSDs also can offer amazing performance for storage needs.

Intel® Advanced Encryption Standard New Instructions (Intel® AES-NI)

Intel® AES-NI implements strong encryption and decryption while greatly reducing the associated

STORAGE NODE SPECIFICATION OVERVIEW	
CPU	1x Intel® Xeon® Processor D-1500 Product Family @ 2.00GHz / 6 Cores
Memory	1x 32 GB DDR4
Network	1x 10GbE port from Intel® Xeon® Processor
Storage Controller	1x LSI* 9207-8i
OS Disk	1x Intel® SSD DC S3700 200GB
Journal and Cache Disk	1x Intel® SSD DC S3500 480GB
Data Disk	10x Seagate* ST3000NC000 3TB 7200RPM SATA HDD

Table 1. Storage Node Specifications.

processing time that strong security[2] requires. Therefore, using Intel® AES-NI to protect data can ensure security without compromising the performance.

Intel® Reference Architectures for Bigtera VirtualStor™ Scaler

This reference architecture uses the Intel® Xeon® processor D-1500 product family reference design as Storage Nodes in the Bigtera VirtualStor™ Scaler cluster. Table 1 provides an overview of the storage node specifications.

This reference architecture is based on the standard Intel X86 design. It is customizable to support different combinations to satisfy the market requirements.

Bigtera VirtualStor™ Scaler Configurations

There are no two identical companies in the world, and no single configuration can fit all companies and their applications. For example, totally different approaches would be taken to design the storage of the lowest cost per gigabytes, or the one of the highest performance. Even for a performance-oriented storage, there will be different configurations for high-throughput workloads and high-IOPS workloads. In this reference architecture, two different Scaler configurations are used to demonstrate how to optimize the performance under different kinds of workloads.

High Throughput configuration is designed for high-throughput workloads, and these workloads are usually characterized by larger data blocks and sequential I/O. Many software-defined storage solutions perform very well in this use case. To optimized for highest MB per second, Bigtera* suggests using two SSDs for OSD journals. Also, deploy one OSD per one SATA hard disk, so totally there are 10 OSDs in a storage node.

High IOPS configuration is designed for high IOPS workloads, and these workloads are usually characterized by small-block random I/O (e.g. 4KB random I/O). For the storage to server virtualization environments, its I/O patterns are usually more IOPS-intensive. To optimize for high IOPS, Bigtera* highly recommends leveraging the speed and high performance of Intel® SSD. In this configuration, Bigtera* uses two SSDs for OSD journals and the cache of large capacity SATA hard drives. In addition, five hard drives are composed to one redundant array of independent disk (RAID) 0 disk, as one OSD. Thus, there are two OSDs in a storage node.

Test Configuration

This section presents the configurations of Bigtera VirtualStor™ Scaler test cluster. Intel deployed Bigtera VirtualStor™ Scaler in a 4-node cluster as shown on Figure 2.

Each storage node has 10 HDDs and two SSDs, and one 10GbE link to the network switch. Table 1 summarizes the details of the storage node specifications.

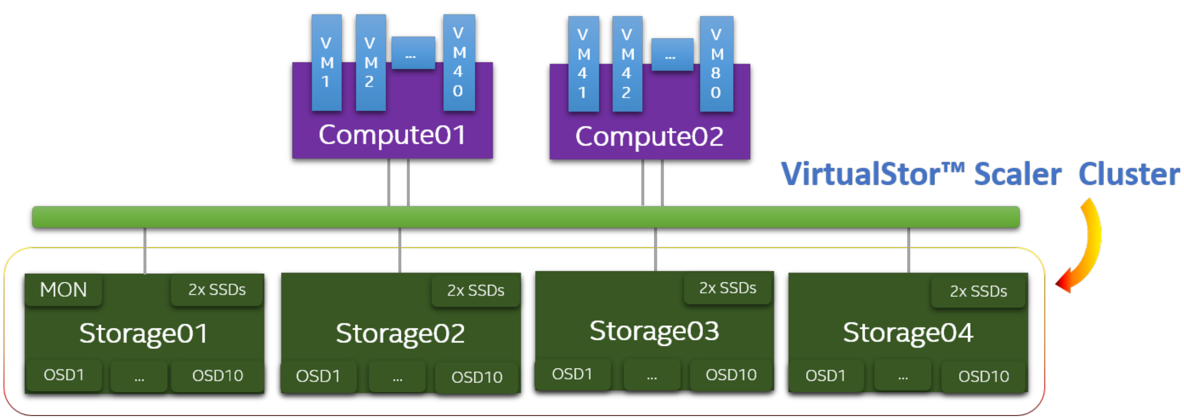

Figure 2. Test Configuration for VirtualStor™ Scaler Cluster.

2 No computer system can provide absolute security under all conditions. Built-in security features available on select Intel® processors may require additional software, hardware, services and/or an Internet connection. Results may vary depending upon configuration. Consult your system manufacturer for more details. For more information. See http://security-center.intel.com/.

COMPUTE NODE SPECIFICATION OVERVIEW	
CPU	2 x Intel® Xeon® Processor E5-2680 v2 @ 2.80GHz / 20 Cores
Memory	128 GB (8GB * 16 DDR3 1333 MHZ)
NIC	2 x 10Gb from Intel® Ethernet Controller X540-AT2
Disks	1 x Seagate* 500G SATA HDD for OS

Table 2. Compute Node Specifications.

Two 2U Intel® Xeon® processor E5-2680 v2 servers are used for the test clients (Compute Notes). 40 virtual machine (VM) instances are created on each physical compute node to emulate the concurrent access cases. Each VM will have one 10GB RADOS (Reliable Autonomic Distributed Object Storage) block device (RBD) volume, and run a benchmarking tool, fio (http://git.kernel.dk/?p=fio.git), to measure the aggregated performance of Bigtera VirtualStor™ Scaler nodes.
In this test, we will gradually increase the number of VM instances from 20 VMs to 80 VM. Then, we should be able to observe the changes of performance when the tested data set gets larger than the size of SSD cache.

Table 2 summarizes the details of the compute node specifications.

Performance of High Throughput Configurations

Performance of High Throughput Configurations is evaluated for sequential I/O workloads of large block-size. All the tests were run with two replicas. Journal data were stored in different partitions of the two SSDs, and OSD data were stored on 10 HDDs. Throughput and latency across the tests are measured by fio under CeTune framework (https://github.com/01org/CeTune). Associated CPU, memory and network utilization are also observed.

Throughput close to hardware limit

For sequential reads (Figure 3) of large block size (128KB), the four node VirtualStor™ Scaler cluster keeps the consistent performance gains from 20 to 80 concurrent VMs, and peaks at 4109 MB/s. On the other hand, for sequential writes (Figure 4) of large block size (128KB), the four node VirtualStor™ Scaler cluster keeps the consistent performance gains from 20 to 80 concurrent VMs, and peaks at 1407 MB/s. After analysis, sequential read write performance is close to the raw hard drive performance, which means the throughput is optimized for the hardware.

Performance of High IOPS Configurations

Performance of High IOPS Configurations is evaluated for random I/O workloads of small block-size. All the tests were run with two replicas. Journal data and cache data were stored in different partitions of the two SSDs. OSD data were stored on 10 HDDs, which five HDDs were composed to one RAID0 disk. IOPS and latency across the tests are measured by fio under CeTune framework (https://github.com/01org/CeTune). Associated CPU, memory and network utilization are also observed.

Random read 213K IOPS

For random reads (Figure 5) of small block size (4KB), the four node VirtualStor™ Scaler cluster peaks at over 213K IOPS with 20 concurrent VMs. The IOPS went downward from 213K to 172K IOPS, as the concurrent VMs were increased. When VMs increase, the dataset is also increased. Memory cache got more and more saturated, and not all hot data could be in memory cache. Thus, lower performance was observed when more VMs are running.

For random writes (Figure 6) of small block size (4KB), the four node VirtualStor™ Scaler cluster maintains the consistent performance around 31K to 32K IOPS, and peaks at over 32.7K IOPS with 30 concurrent VMs. After further observation, CPU utilization is almost full during the test, which means it already reached the limit of computing power.

Latency and IO depth

To further validate if the VirtualStor™ Scaler cluster can provide small enough latency for random IO, another test was designed to measure the latency changes when the IO depth increased exponentially. In the test, 20 VM instances were created on the compute nodes, and increase each VM's IO depth from 1 to 64. During the tests of random reads and random writes, IOPS and latency were measured.

For random reads (Figure 7) of small block size (4KB), the four node VirtualStor™ Scaler cluster gradually increased its latency from 0.478 milliseconds, and peaks at 6 milliseconds when iodepth = 64. For random writes (Figure 8) of small block size (4KB), the four node VirtualStor™ Scaler cluster gradually increased its latency from 1.825 milliseconds, and peaks at 41 milliseconds when iodepth = 64.

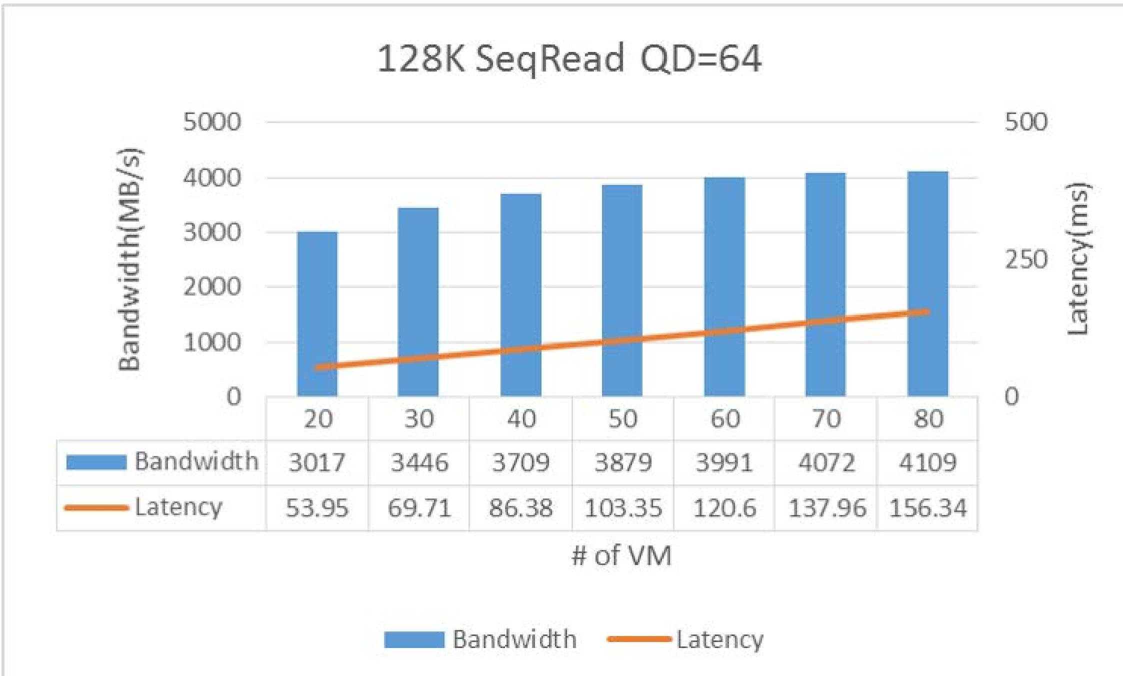

Figure 3. Sequential Read Bandwidth on a 4-node VirtualStor™ Scaler Cluster.

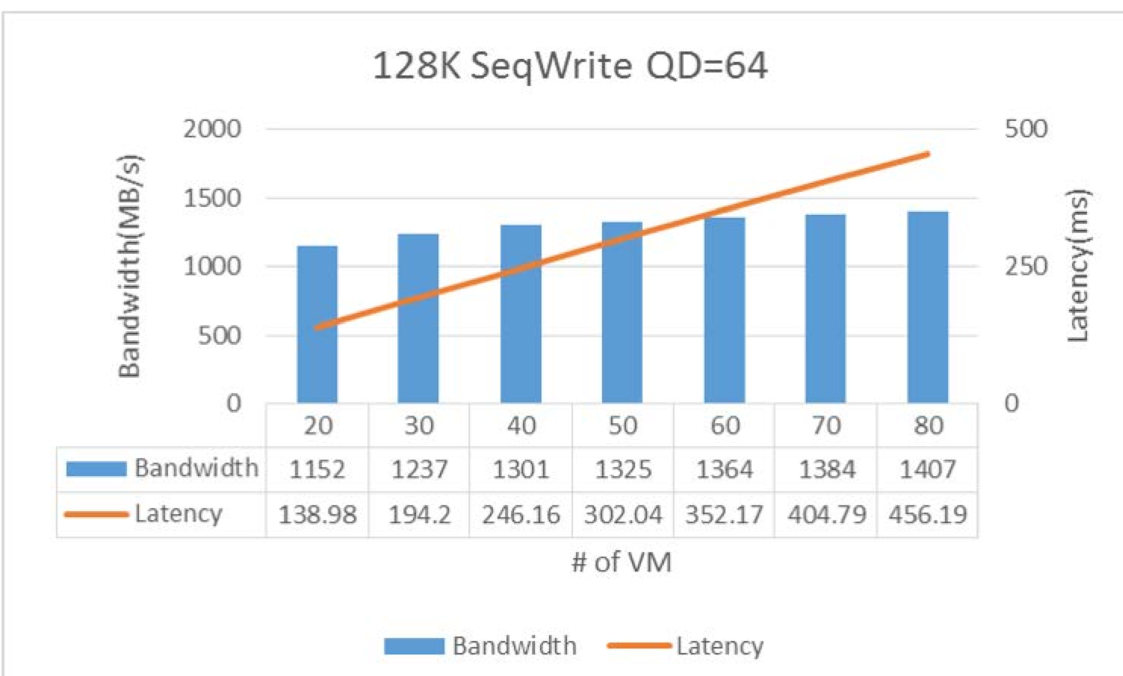

Figure 4. Sequential Write Bandwidth on a 4-node VirtualStor™ Scaler Cluster.

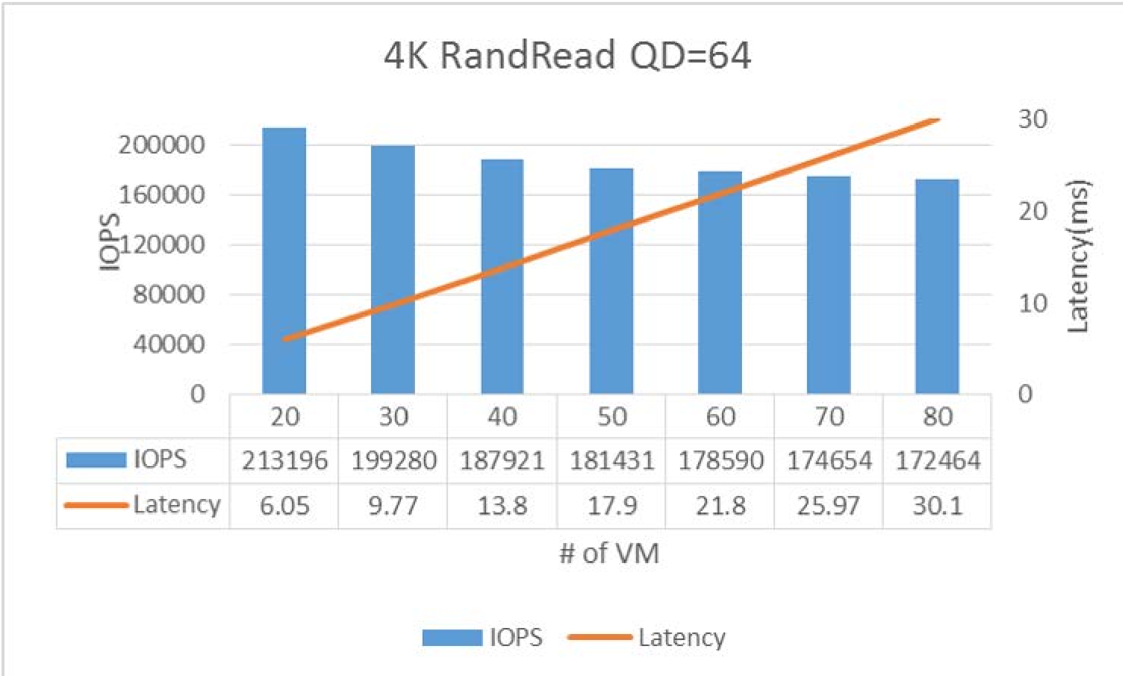

Figure 5. Random Read IOPS on a 4-node VirtualStor™ Scaler Cluster.

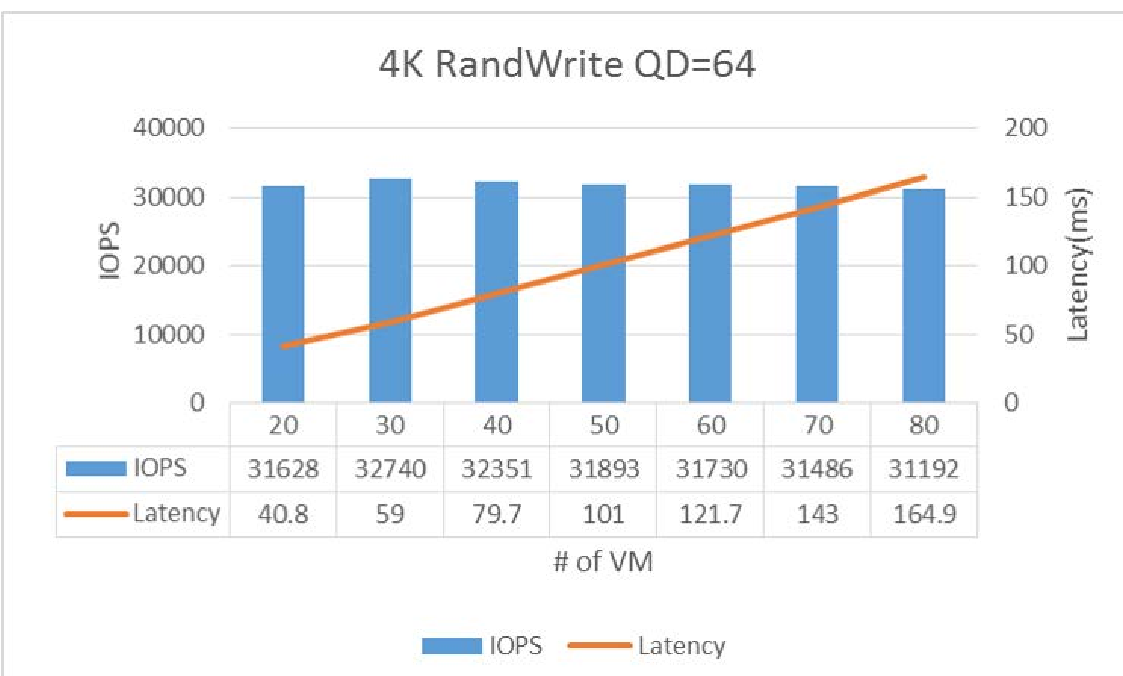

Figure 6. Random Write IOPS on a 4-node VirtualStor™ Scaler Cluster.

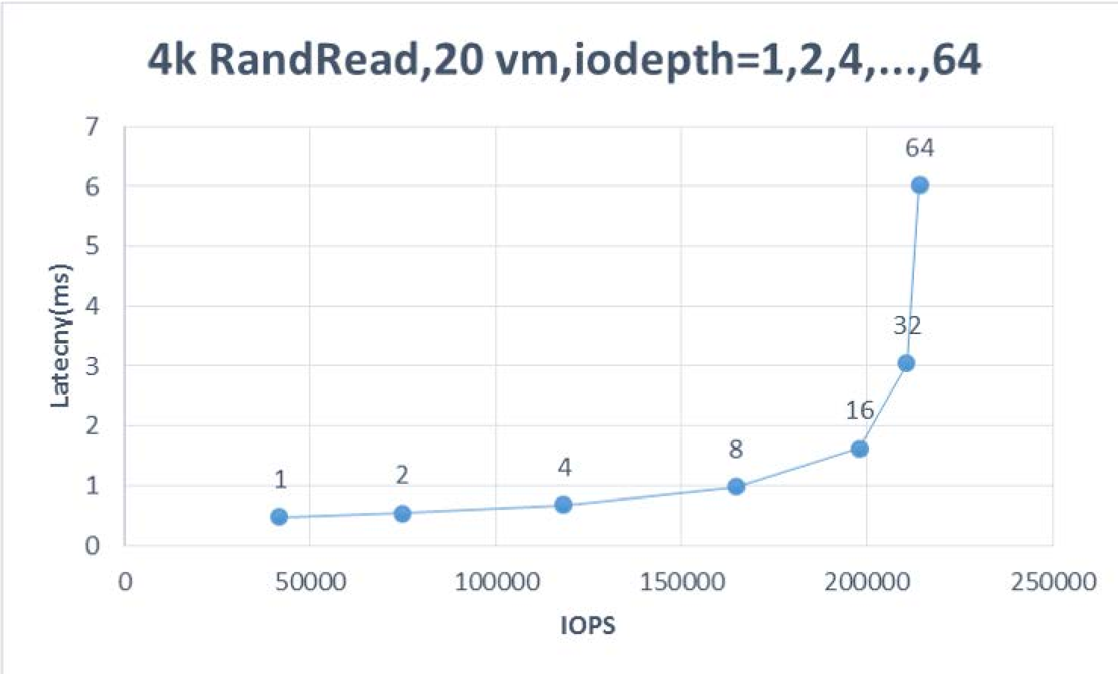

Figure 7. Random Read IOPS and Latency on a 4-node VirtualStor™ Scaler Cluster.

Figure 8. Random Write IOPS and Latency on a 4-node VirtualStor™ Scaler Cluster.

Conclusion

It's time for modern organizations to eliminate the pain points associated with traditional storage.

Legacy storage infrastructures not only fail to meet business needs, but also unnecessarily increase IT costs. The reference architectures and two storage configurations outlined in this document offer optimized throughput and IOPS, under a cost-effective hardware platform. Also, like the spirit of software-defined infrastructure, each architecture's hardware and software can be flexibly customized to fit various kinds of demands. For example, if the customer needs more storage capacity, additional storage nodes can be installed or replaced with the hard disks of larger capacity. If the customer needs greater performance, on the hardware side, the traditional spinning disks can be replaced with SSD, and additional storage nodes also can help. On the software side, Bigtera VirtualStor™ Scaler can offer reference configurations to fit high-throughput workloads or high-IOPS workloads under the economic hardware platform.

Bigtera VirtualStor™ Scaler can offer high-performance, and its features like Unified Storage, Multitenancy, and Quality of Service can satisfy the needs of business complexity without pain. Scale-out expansion, seamless migration, and cost-effective storage virtualization, can help enterprises consolidate their storage islands, and address the challenges of rapid data growth. Data migration and tunneling between NAS storage pool and object storage pool is also the recipe to solve the silo effects in companies.

Get More Information

To find about Bigtera* software defined storage (SDS) solutions, visit **http://www.bigtera.com**

To learn more about Intel storage products and technologies, visit **http://www.intel.com/storage**