Storage I/O Performance on VMware vSphere® 5.1 over 16 Gigabit Fibre Channel

Performance Study

TECHNICAL WHITE PAPER

vmware®

Table of Contents

Executive Summary

Fibre Channel (FC) interconnect speed has increased to 16 gigabits per second (Gbps). To take full advantage of this enhancement, VMware vSphere® 5.1 has added support for the 16Gb FC host bus adapter (HBA). This white paper illustrates the improvements in storage I/O throughput that vSphere gains when using a 16Gb FC HBA instead of an 8Gb FC HBA. In addition, this paper looks at the CPU efficiency of the vSphere host when handling the increased storage I/O.

Introduction

With the availability of 16Gb FC HBAs, larger block I/Os have doubled their throughput compared to the 8Gb per second bandwidth. The throughput has improved together with better CPU efficiency per I/O. For random I/Os with small block sizes, I/Os per second (IOPs) have increased because their throughput is not bounded by the FC link speed before reaching the 16Gbps wire speed limit.

Having doubled the throughput rates from the newly available 16Gb FC HBA, the vSphere 5.1 platform can support more storage devices and meet bandwidth requirements using the same number of FC links. The doubled bandwidth for large block transfers can benefit applications like VMware vSphere® Storage vMotion®. The improvement in random read IOPs at a 4KB block size, for example, can benefit application components like VMware® View™ clients. Applications are no longer limited by the existing 8Gb FC bandwidth to meet their peak performance requirements.

Setup

ESXi Host

- HP Intel Proliant DL370 G6 Server
- Two Quad Core Xeon W5580 Processors, 3.20GHz
- Emulex LPe16002 16Gb FC HBA initiator
- Emulex LPe12000 8Gb FC HBA initiator

EMC VNX7500 Storage Array

- 8Gb FC target ports connected to 16Gb FC SAN switch (Brocade DS-6510) for the LPe16002 initiator
- 8Gb FC target ports connected to 8Gb FC SAN switch (Brocade DS300) for the LPe12000 initiator
- A total of 32 solid-state drive (SSD) LUNs; the 256MB SSD cached LUNs had mirrored write cache enabled

Virtual Machine

- Guest OS: Windows 2008 R2, 64-bit
- Virtual disk: 256MB Mapped Raw LUN (RDM physical)
- Number of virtual disks: four
- SCSI controller: one PVSCSI virtual controller
- Number of virtual CPUs: one
- Virtual hardware version: 7
- Guest memory: 4096MB

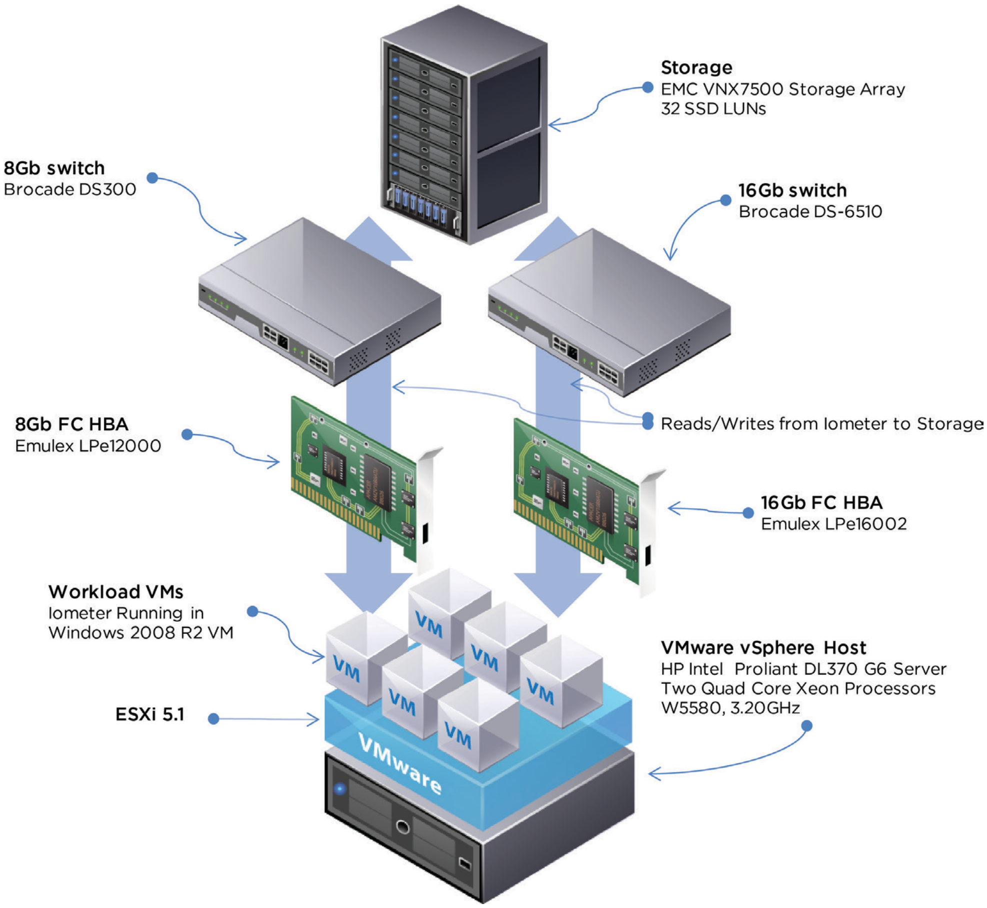

Figure 1. Test bed setup

Workload

A single Iometer worker in a 1-vCPU virtual machine was used to compare the performance between 16Gb FC and 8Gb FC HBAs. The Iometer worker drove 4 target LUNs with 32 outstanding I/Os. Sequential read and write I/Os were used to compare the throughput and CPU efficiency for different block sizes.

The same virtual machine configuration and Iometer workload were used for a multi-virtual machine comparison. In general, sequential I/Os have a higher throughput than random I/Os. However, the random I/Os in small I/O block sizes can also benefit from the higher bandwidth. In the multi-virtual machine configuration, we used random I/Os to show the benefit of using 16Gb FC with an incremental workload from one to eight virtual machines.

To demonstrate ESXi storage performance and the benefits from using a 16Gb FC HBA, the workload made use of mirrored write cache at the storage array. This configuration provided the most benefit from the cache at the storage array for the throughput test. Only the read I/Os were used in the multi-virtual machine configuration so that the workload did not stress the single storage array to sustain the best throughput performance needed by the test.

Each performance comparison was carried out over a single port of the LPe16002 (16Gb FC HBA) and then over a single port of the LPe12000 (8Gb FC HBA) to the backend storage array.

Results

Sequential Write Throughput and CPU Efficiency

The following is the comparison of the 16Gb FC HBA to the 8Gb FC HBA for sequential write I/Os using a single Iometer worker thread.

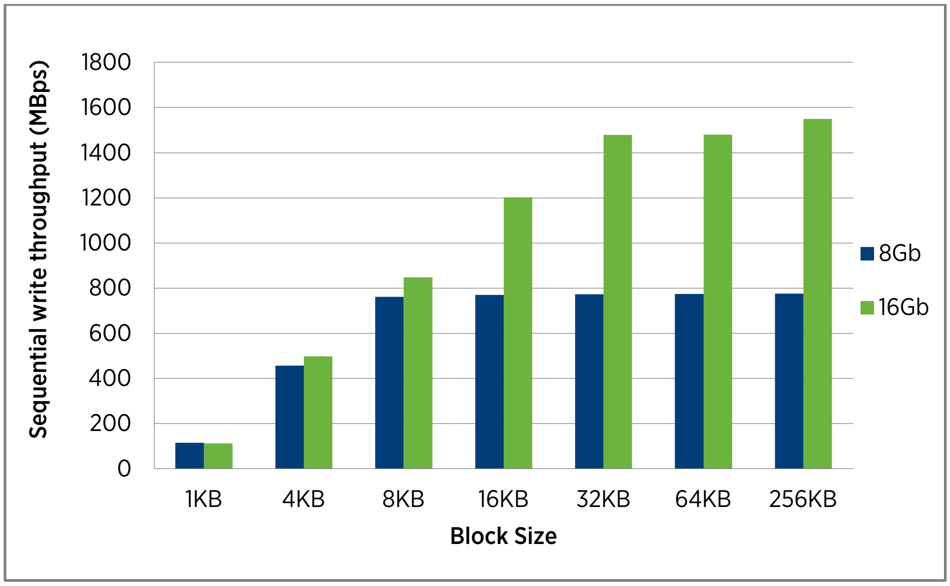

Figure 2. Sequential write throughput in megabytes per second

Figure 2 shows that larger block sizes (32KB and greater) come close to fully utilizing the 16Gb FC HBA. With a 32KB block size and above, the throughput for the 16Gb FC HBA is near its wire speed and is 90% to 100% faster than that of the 8Gb FC HBA.

Figure 3 shows CPU efficiency (also known as CPU cost), which is the measure of the amount of CPU resources used by ESXi to perform a given amount of I/O. The CPU efficiency for 16Gb FC is better than 8Gb FC by 10% to 15%.

The 16Gb FC HBA has better write throughput in smaller block sizes as well. It has better throughput by 56% for a 16KB block size, and better throughput by 10% for the 4KB and 8KB block sizes, all with approximately 5% better CPU efficiency.

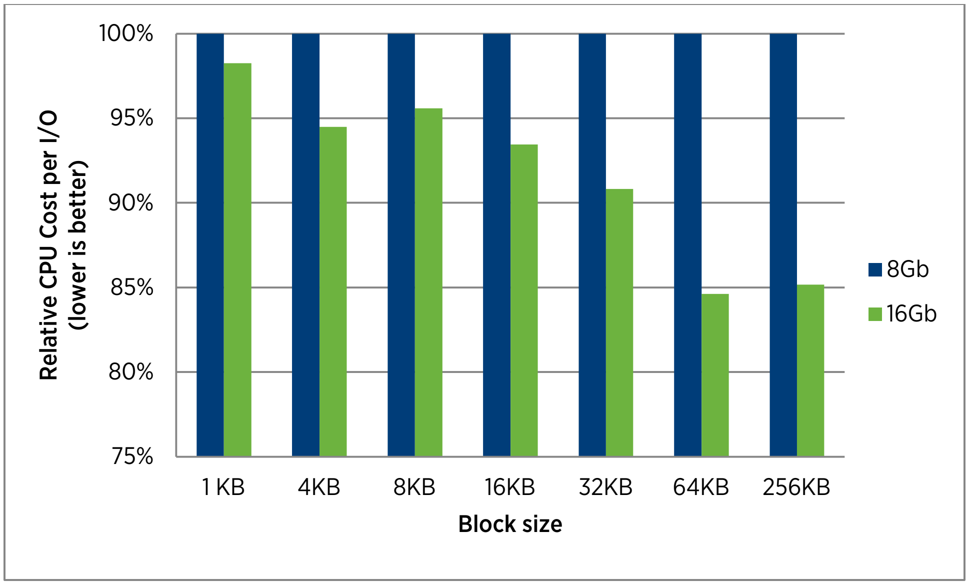

Figure 3. Sequential writes relative CPU cost per I/O (lower is better and means a higher CPU efficiency)

Sequential Read Throughput and CPU Efficiency

The following is the comparison of the 16Gb FC HBA to the 8Gb FC HBA for the sequential read I/Os using single Iometer worker thread.

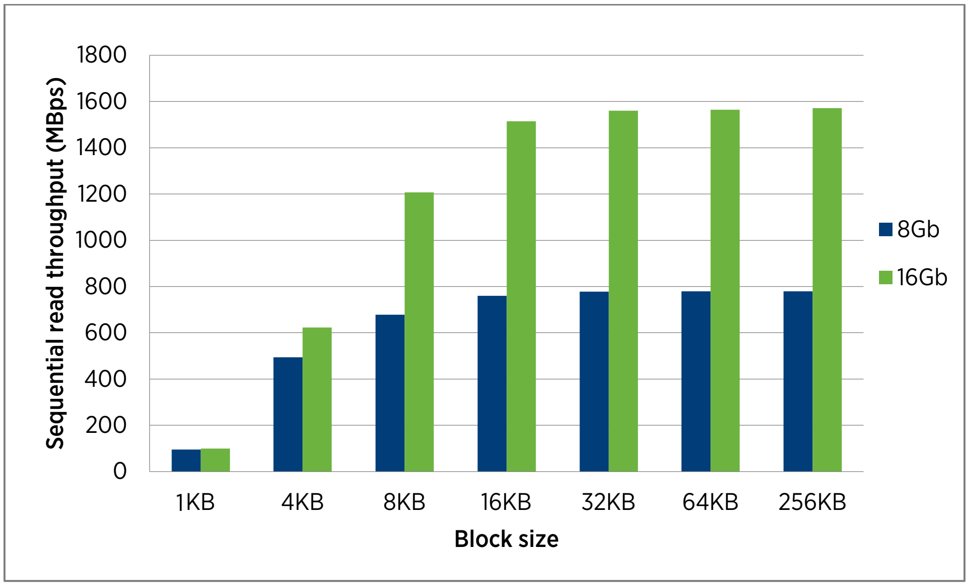

Figure 4. Sequential reads throughput in megabytes per second

Unlike the write I/Os that incur a small performance penalty as the result of having to write to both copies of the mirrored cache at the storage array, the read I/Os have better performance improvements starting at smaller I/O block sizes. Figure 4 shows that the read throughput has improved by 25% in the 4KB block size, and by 75% in the 8KB block size.

Figure 5 shows that all block sizes have a better CPU efficiency by about 7% to 15% compared to the 8Gb FC HBA.

For block sizes 16KB and above, both the 16Gb FC HBA and the 8Gb FC HBA reach their respective wire speeds.

At 1570 MBps, the 16Gb FC HBA has doubled its throughput compared to the 8Gb FC HBA at 780 MBps (Figure 4), and with a better CPU efficiency by about 8% to 10% (Figure 5).

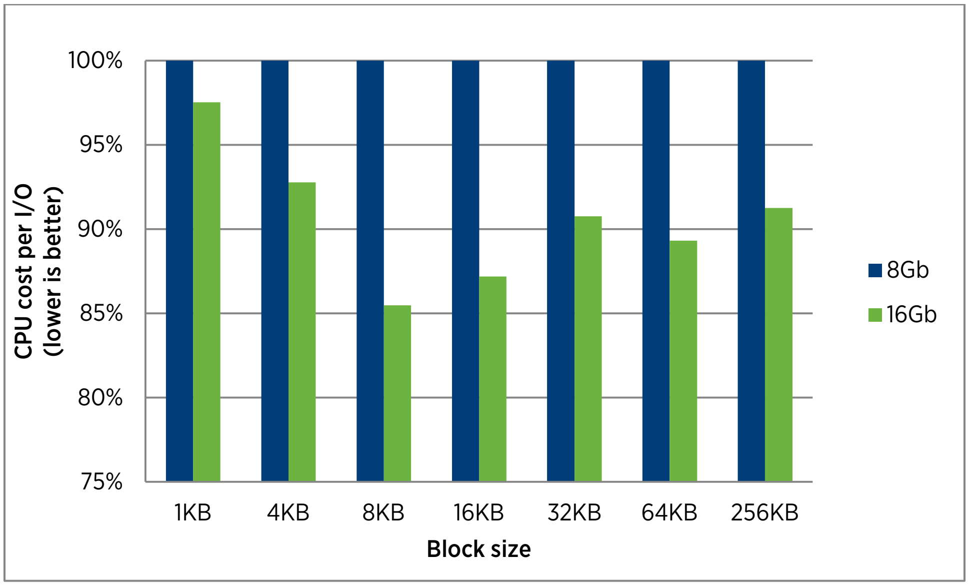

Figure 5. Sequential reads relative CPU cost per I/O (lower is better and means a higher CPU efficiency)

Random Read Throughput and I/Os Per Second with Multiple Virtual Machines

The following is the random I/O performance for the 16Gb FC HBA from the multi-VM workload using one Iometer worker per virtual machine

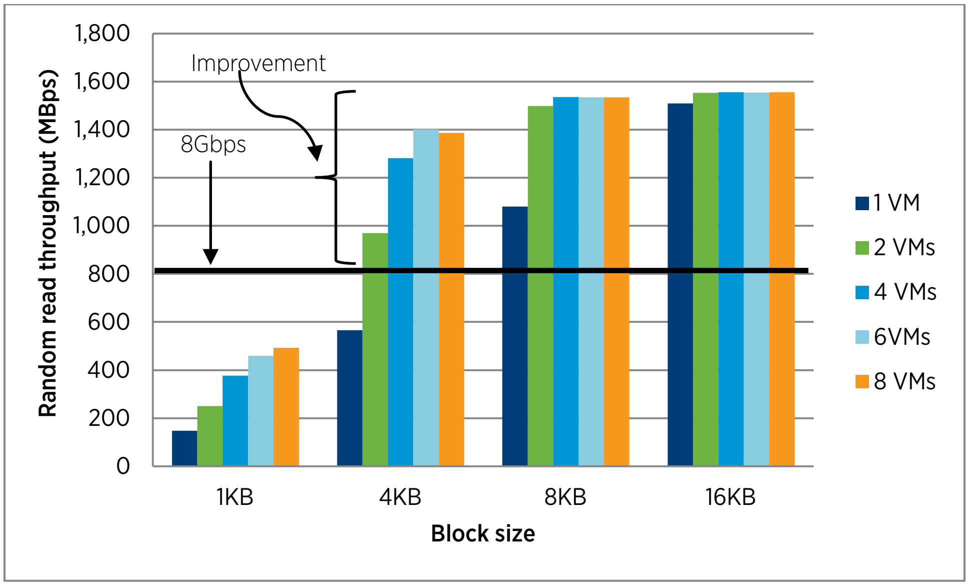

Figure 6. Random read throughput (MBps) for different block sizes and multiple virtual machines (VMs)

The random I/Os in small block sizes also benefit from using a 16Gb FC HBA. The extra bandwidth allows for much higher IOPs beyond the 8Gbps wire speed. For the 4KB block size, the total 350K IOPs has an accumulative throughput of 1400 MBps (Figure 6 and Figure 7). The 8KB I/Os have an accumulative throughput of 1535 MBps, which is near the 16Gbps wire speed (Figure 6).

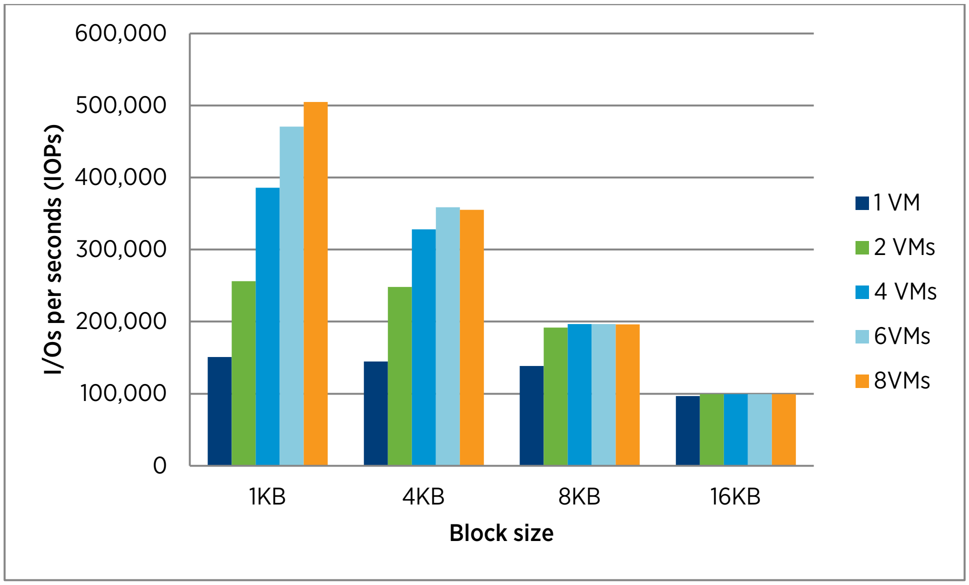

Figure 7. Random read IOPs for different block sizes and multiple virtual machines

Conclusion

By having the ability to use 16Gb FC HBAs, vSphere 5.1 gains higher I/O throughput and better CPU efficiency. Applications virtualized on vSphere 5.1 that generate storage I/O of various block sizes can take advantage of 16Gbps wire speed for better sequential and random I/O performance.

VMware, Inc. 3401 Hillview Avenue Palo Alto CA 94304 USA Tel 877-486-9273 Fax 650-427-5001 www.vmware.com