Low Latency Evaluation of Fibre Channel, iSCSI and SAS Host Interfaces

Evaluation report prepared under contract with LSI Corporation

Introduction

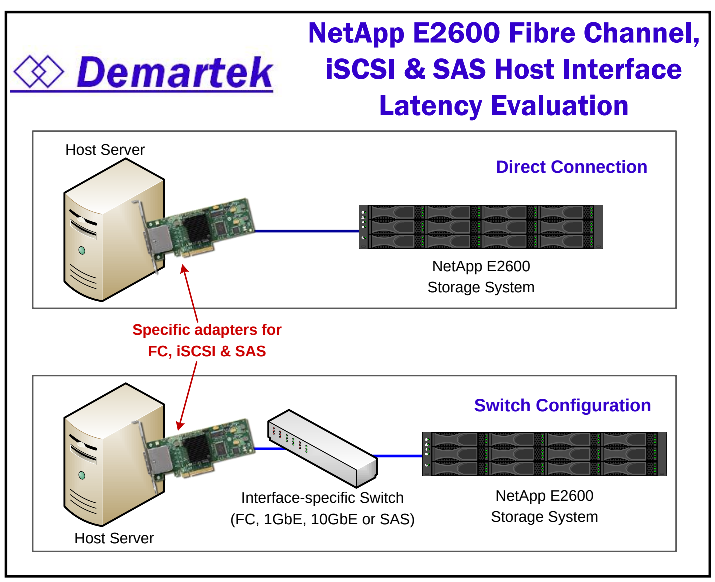
IT professionals see Solid State Disk (SSD) products as a means to obtain higher performance and better response times. Having higher IOPS and shorter latencies than HDD, SSDs are being deployed in external storage arrays in conjunction with, or replacing HDDs. While high IOPS is frequently the metric for assessing drive and array performance, in many cases latency is as important. For example, in an OLTP application, short response times are critical. A transaction may require many successive queries to a database, where each query depends on answers returned from the previous query. User response times are entirely dependent on how quickly storage can return answers to the queries.

While SSDs can enable short response times, the connection to the storage array may limit these responses. An increasing number of current generation storage systems are providing multiple types of host interfaces for the same storage system, giving customers a wide variety of choices for connectivity to these storage systems. Block host interfaces, including Fibre Channel, iSCSI and SAS are all becoming increasingly available for today's storage systems.

LSI commissioned Demartek to compare four block interface and speed combinations on the same storage system for basic performance. At first glance, one might assume that the interfaces with the highest transfer rates would provide the best overall performance, but we found that this was not the case.

Evaluation Environment

A NetApp E2600 storage system was used in this evaluation because it supports all four host interface types and can ensure an apple-to-apples comparison of the interface types. The same host server was used for the entire test, using appropriate adapters for each interface type. We included two basic types of configurations for each interface type. The first configuration was a direct connection between the host server and the storage system. To represent the basis of a SAN, the



second configuration included an interface-specific switch between the host server and the storage system. As database applications frequently require short response times, we used Microsoft SQLIO as a workload generator for random and sequential read and write operations. In order to measure minimum latencies, a queue depth of one was used for all workloads.

Evaluation Summary

We found that the best overall performance for this storage system was with the 6Gb SAS host connection, followed very closely by the 8Gb Fibre Channel host connection. This is due to the low latency characteristics for 6Gb SAS connections, as measured from the application host server. Although 6Gb SAS has a lower line rate than 8Gb Fibre Channel or 10Gb iSCSI connections, its overall protocol overhead is lower than the other interfaces tested, and makes an excellent choice as a host interface for storage systems.

A Case for SSDs

With the variety of host interfaces available for storage systems today, it is reasonable to ask which interfaces satisfy today's performance needs and which interfaces will satisfy future performance needs. During the 1990's, we saw increases in the hard disk drive (HDD) rotations per minute (RPM) about every four years, until around the year 2000, when enterprise disk drives achieved 15,000 RPM, which appeared to meet the needs of that time period. However, the maximum practical disk drive RPM speed has remained at that level for a decade. Today's best enterprise disk drives spin at 15,000 RPM and have an average response time (latency) of approximately two milliseconds (2ms). The slower disk drives have higher latencies. This means that no read or write operation from or to the surface of a disk drive can be completed in less than 2ms, on average. Although we can get more I/Os per second by adding hard disk drives to a configuration, we appear to have hit the limit of electro-mechanical spinning disk drive latency performance. For these types of devices, the host interfaces are usually able to keep up with the performance of the disk drives on the back end.

In this decade we are seeing increasing acceptance of solid state drives (SSDs), which measure their latencies in microseconds, which are significantly smaller (faster) than hard disk drives. With the advent of SSDs, the question of which host interface to use becomes more interesting, as these newer SSD devices have much faster response times.

Array Response Times

In general, higher IOPS results in shorter latency. However, latency cannot be shorter than the storage device's response time. It is important to note that striping the data across many HDDs will increase IOPS because all the drives are working simultaneously, but the disk's rotation rate still governs the array's response time. A good explanation of latencies in RAID5 configurations is available at http://en.wikipedia.org/wiki/Standard_RAID_levels#RAID_5_latency. It describes how array striping will increase latency to double the average latency of a HDD.

Testing Overview

We measured the IOPS and response time (latency) of the same server host and NetApp E2600 storage system. These interfaces were evaluated:

- 8Gb Fibre Channel (FC)
- 10Gb iSCSI
- 1Gb iSCSI
- 6Gb SAS

Our expectation was that the FC connection would have the highest IOPS performance. The next highest could be either 10Gb iSCSI or 6Gb SAS. With 10Gb throughput capability, it would seem that iSCSI would have an edge over SAS, especially for large data transfers. In general, iSCSI suffers from the overhead of packing data inside internet protocol but hardware offload engines are supposed to mitigate any delays. SAS has become the de facto interface for connectivity to enterprise storage drives, so its performance for external connectivity was expected to be high, possibly comparable to 10Gb iSCSI.

To obtain the highest IOPS and lowest latency from the E2600, data transactions were constrained to only use the controller's cache DDR memory. This removed SSDs from potentially limiting the measurement results. In many database applications, query results are needed for subsequent queries, so a queue depth of one was used to ensure only one outstanding transaction. Since the E2600 can support a small number of servers with direct connections, we measured both a direct connection configuration and through a switch.

Latency and IOPS Commentary

Measured latencies are shown in Figure 1 and IOPS using random transactions in Figure 2. Random transactions tend to be more prevalent than sequential transactions when SSDs are used. For latency, it shows that with the exception of 1GE iSCSI, latencies are below 1 ms. Both 1Gb and 10Gb iSCSI are significantly lower in IOPS compared to SAS and FC.

The resolution of SQLIO doesn't resolve latencies below 1ms. To view performance in this area, we derived the latency using Little's Law. Since the queue depth is one, latency is the reciprocal of the IOPS. Figure 8 shows the calculated latency using the measured IOPS. Since 1Gb iSCSI is significantly slower than the other interfaces, it was dropped from the calculation

The 6Gb SAS interface generally performed the best, but in some cases, the 8Gb FC interface performed the best of the four interfaces, especially at 64KB and higher block sizes. As expected, the 1Gb iSCSI had the lowest performance in all categories. These results are for this particular storage system, and other storage systems that support all four of these interfaces may provide different results. However, this test shows that 6Gb SAS can be an excellent choice for a host interface when performance and latency are important.

Although iSCSI running at 10GbE has a theoretical advantage in wire speed over the 8Gb FC and 6Gb SAS interfaces, the overhead of the iSCSI, TCP and IP layers reduces the overall performance for 10Gb iSCSI.

Performance Commentary

Of the 24 combinations of random and sequential, read and write block sizes, the directly connected 6G SAS interface had the highest results in 16 of the combinations, and was approximately equal to the next best performing interface in 2 of the combinations. In all cases, the 6G SAS interface outperformed the 10G and 1G iSCSI interfaces for each test combination. Also in all cases, the 8G FC interface outperformed the 10G and 1G iSCSI interfaces.

The difference in performance between the 6G SAS and 8G FC was found primarily in the block sizes. For 8KB through 32KB block sizes, the directly connected 6G SAS interface performed better than the 8G FC interface for random and sequential, read and write operations. The 8G FC interface generally performed better than 6G SAS at 64KB and higher block sizes. For sequential writes, the directly connected 6G SAS interface performed equal to or better than the 8G FC interface for all block sizes.

As a host interface, 6G SAS performed equal to or better than the next best performing interface (8G FC) in approximately 2/3 of the tests. This clearly shows that 6G SAS can provide superior performance for many workload scenarios.

Latency Summary

Microsoft SQLIO records the latencies by the percentage of results in each 1ms interval up through 23ms then one last category for 24 or more milliseconds. For these measurements, the better results have a higher percentage of data points in the 0-1ms range (the more to the left the better). The results up to 5ms are shown in this graph. The full graph is shown in Appendix 2.

The directly connected 6G SAS interface had the most number of latency results in the 0-1ms category, followed very closely by switched SAS and 8G FC. More than 99% of the 6G SAS and 8G FC data points are within the 0-2ms categories, followed by more than 97% for 10G iSCSI and only 34% for 1G iSCSI. Looking just at the latencies in the 0-1ms and 0-2ms ranges, 6G SAS and 8G FC have nearly identical results overall. The chart of all the latency measurements is found in Figure 6 (Appendix 2) of this report.

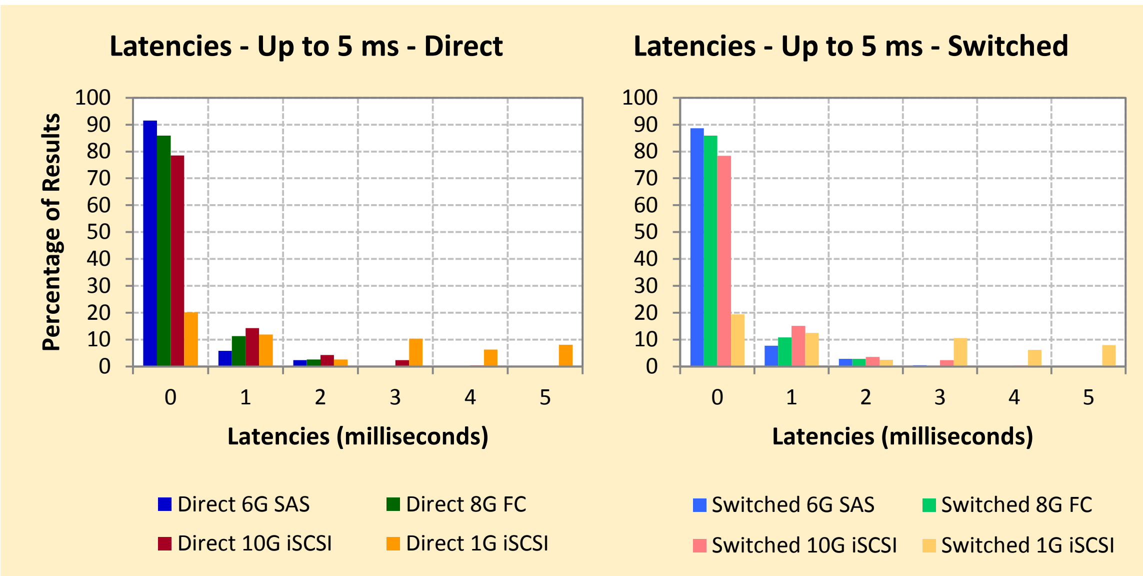


Figure 1 – Latencies Up to 5 ms

Host Interface		Percentage of Latencies (ms)							
		0	1	2	3	4	5	0-1	0-2
6G SAS	Direct	91.50%	5.88%	2.42%	0.04%	0.00%	0.00%	97.38%	99.88%
	Switched	88.63%	7.79%	2.88%	0.46%	0.08%	0.04%	96.42%	99.29%
8G FC	Direct	85.96%	11.25%	2.58%	0.13%	0.00%	0.00%	97.21%	99.79%
	Switched	85.88%	10.88%	2.83%	0.21%	0.00%	0.00%	96.75%	99.58%
10G iSCSI	Direct	78.54%	14.25%	4.25%	2.42%	0.29%	0.04%	92.79%	97.04%
	Switched	78.46%	15.08%	3.54%	2.42%	0.29%	0.04%	93.54%	97.08%
1G iSCSI	Direct	20.08%	11.88%	2.67%	10.38%	6.29%	8.08%	31.96%	34.63%
	Switched	19.42%	12.54%	2.50%	10.54%	6.17%	7.96%	31.96%	34.46%

Table 1 – Percentage of Latencies by Host Interface

Appendix 1 – IOPS with Random I/O and Sequential I/O

The charts below show the IOPS results for random and sequential I/O, separated into the direct connections and the switched connections.

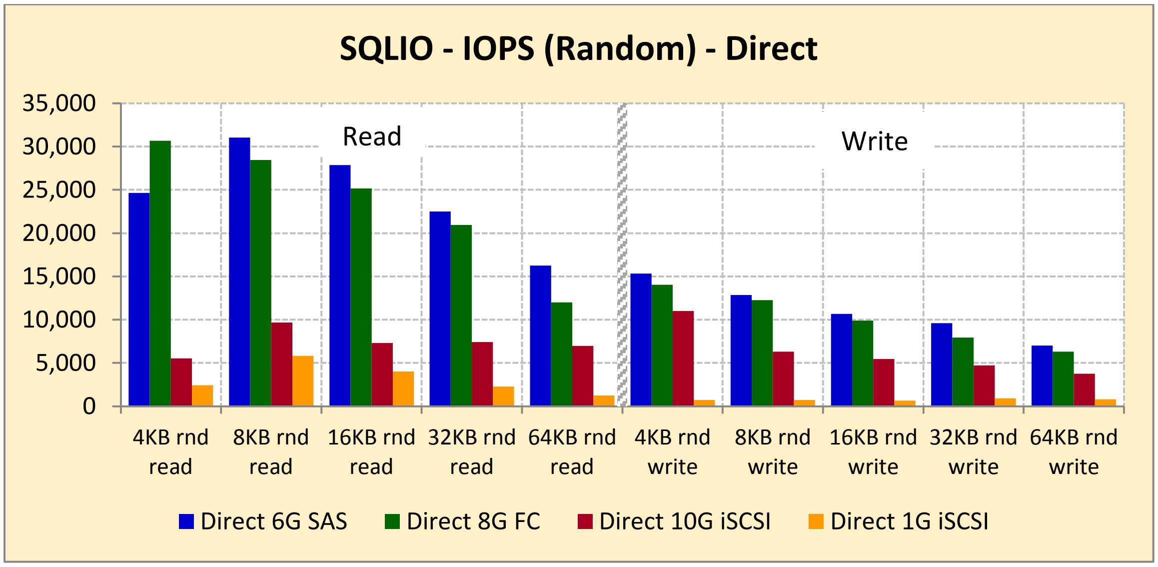


Figure 2 – SQLIO Random IOPS – Direct Connect

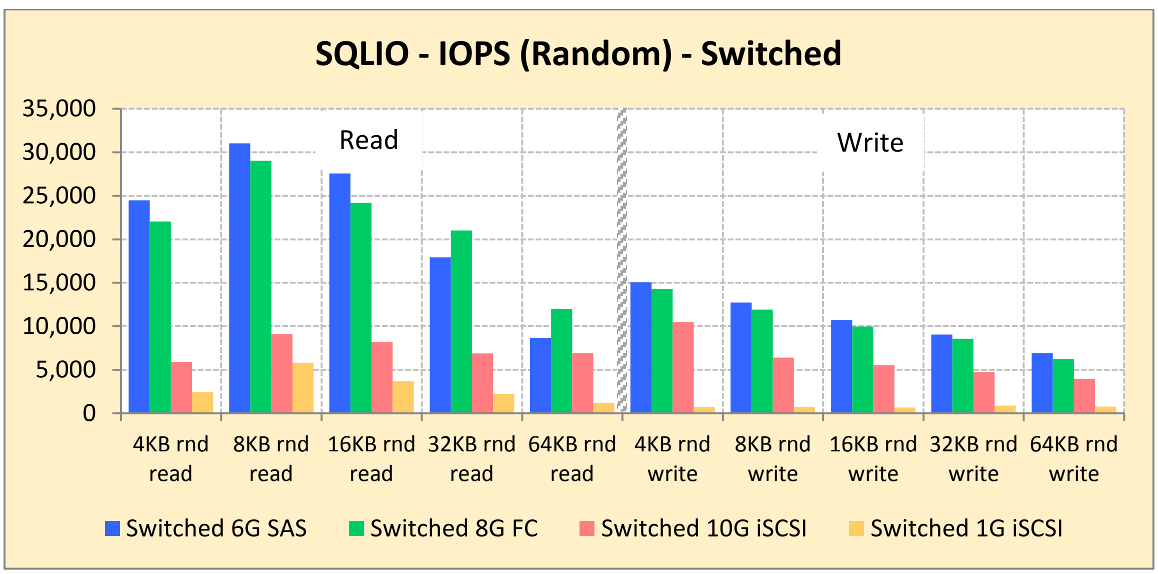


Figure 3 – SQLIO Random IOPS – Switch Connected

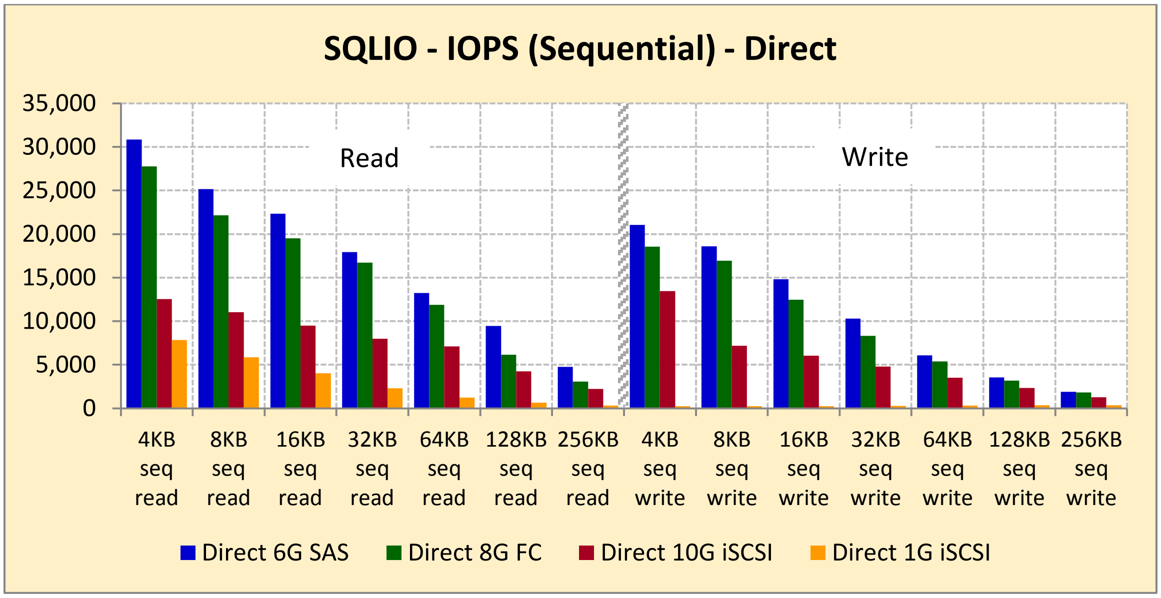


Figure 4 - SQLIO Sequential IOPS - Direct Connect

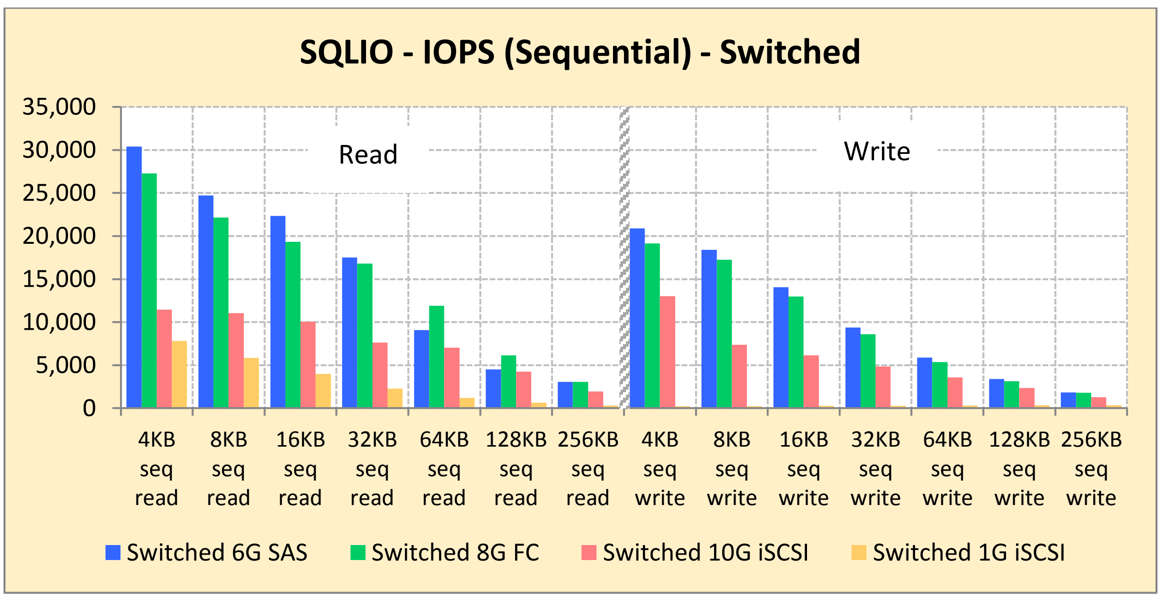


Figure 5 - SQLIO Sequential IOPS - Switch Connected

Appendix 2 – Latencies

The chart below shows all the latency results. The lowest latencies were for 6G SAS, followed closely by 8G Fibre Channel. The 1G iSCSI latencies were the highest (worst), as expected.

The best results are indicated by the highest data points in the left-most columns. The latency distribution graph shown in Figure 1 includes the first six columns (up to 5ms) of this graph.

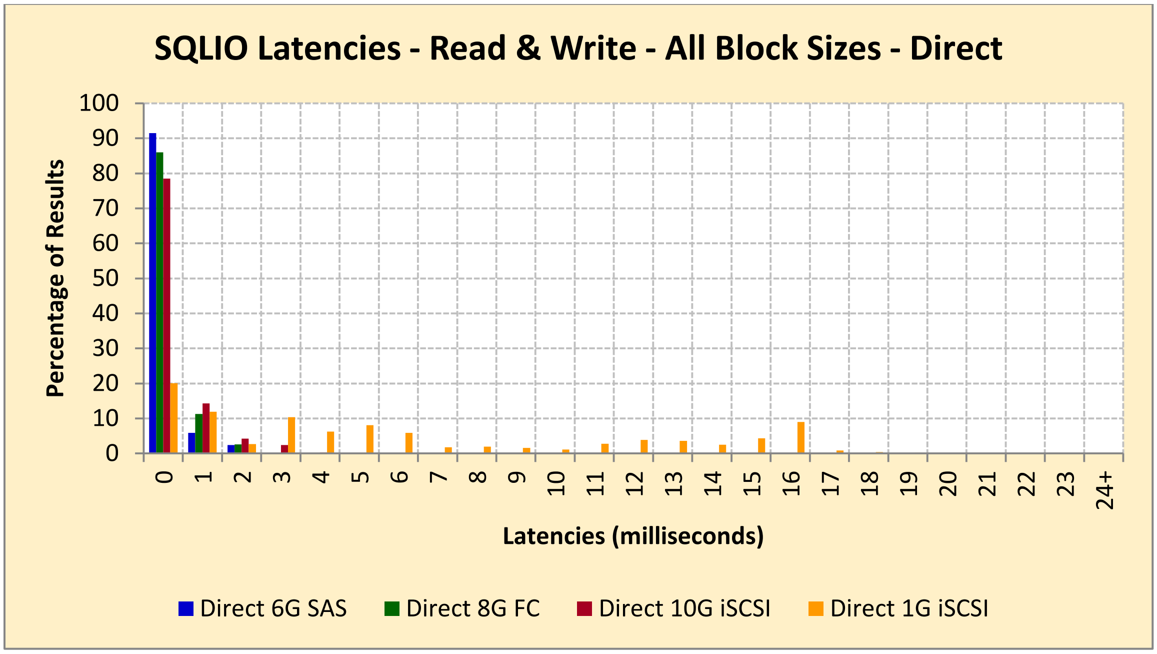


Figure 6 – SQLIO Latencies – Read & Write – All Block Sizes – Direct Connect

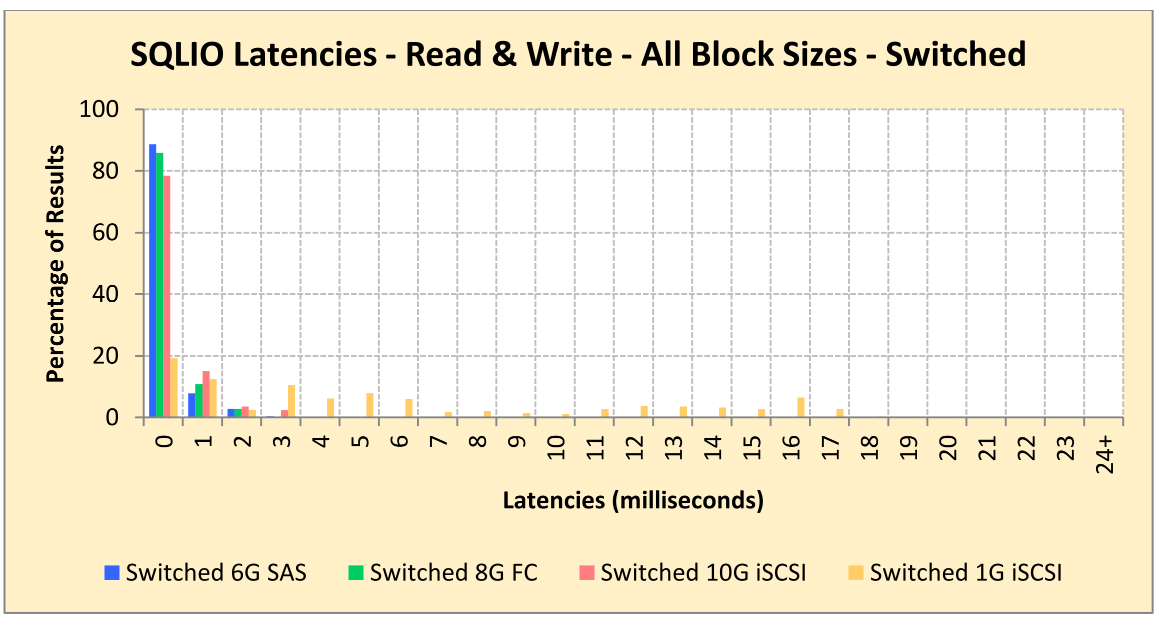


Figure 7 – SQLIO Latencies – Read & Write – All Block Sizes – Switch Connected

To get the latency in milliseconds of an individual I/O we take the inverse of the IOPs using the following equation:

$$Computed\ Latency = \left(\frac{1}{IOPS}\right) * 1000$$

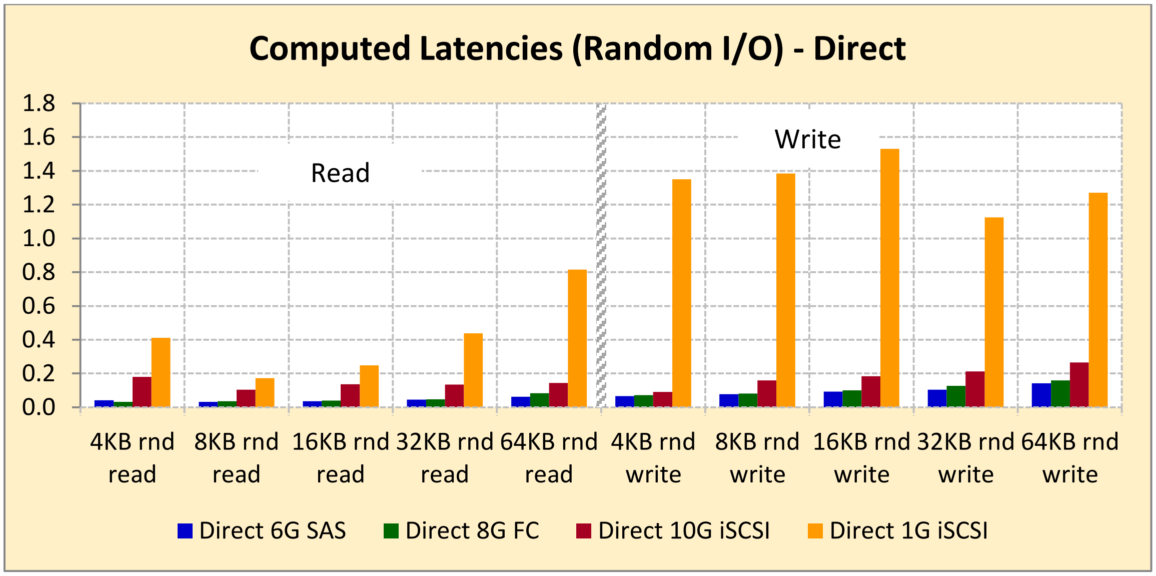


Figure 8 – Computed Latencies – Random I/O – Direct

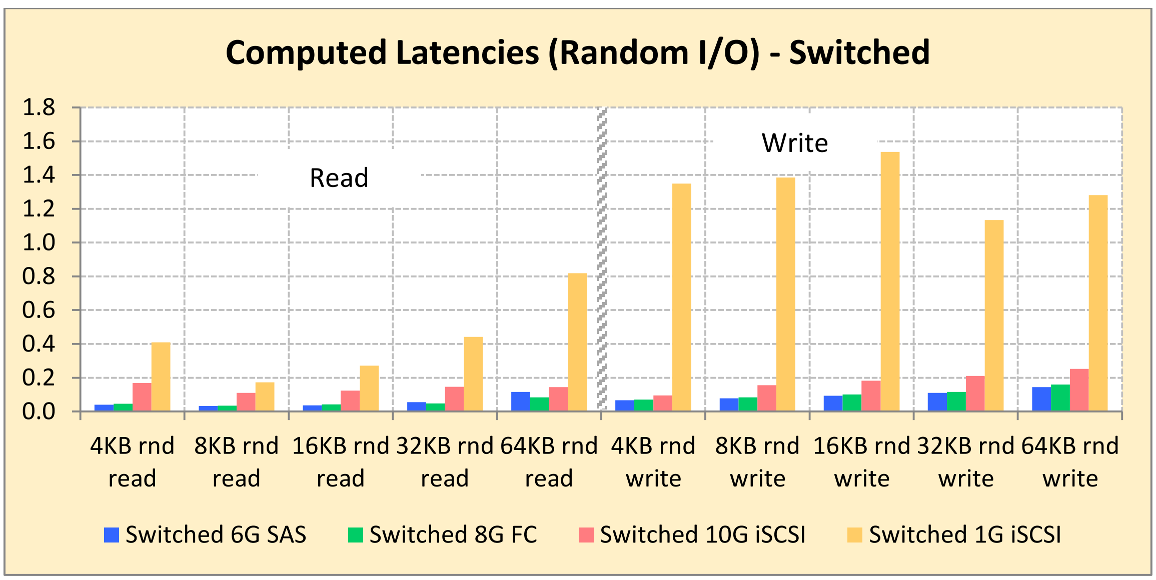


Figure 9 – Computed Latencies – Random I/O – Switch Connected

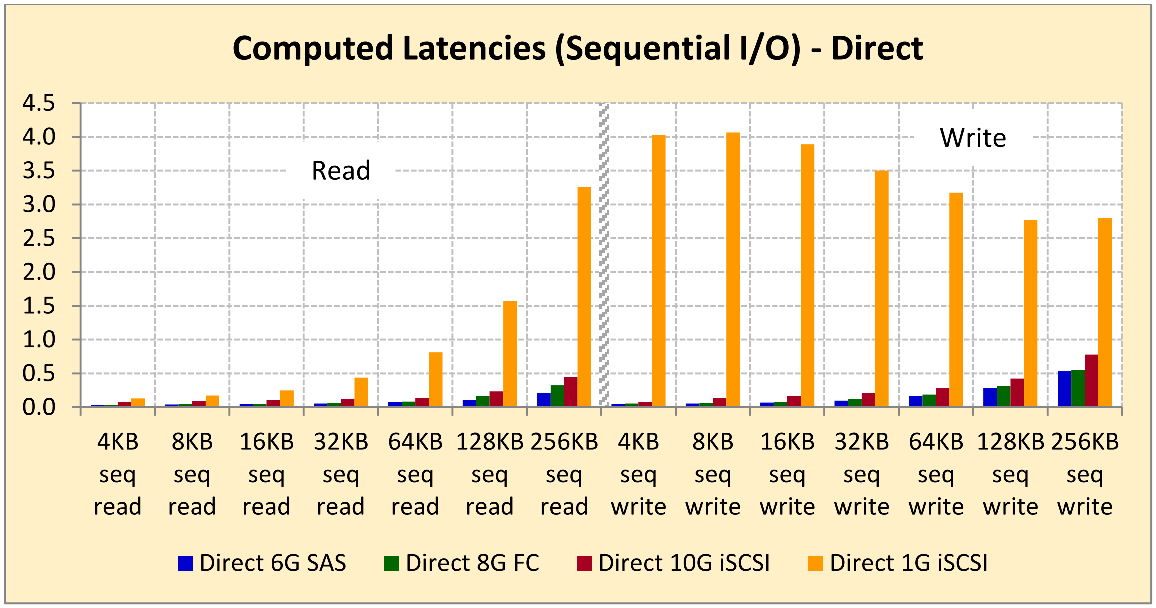


Figure 10 – Computed Latencies – Sequential I/O – Direct Connect

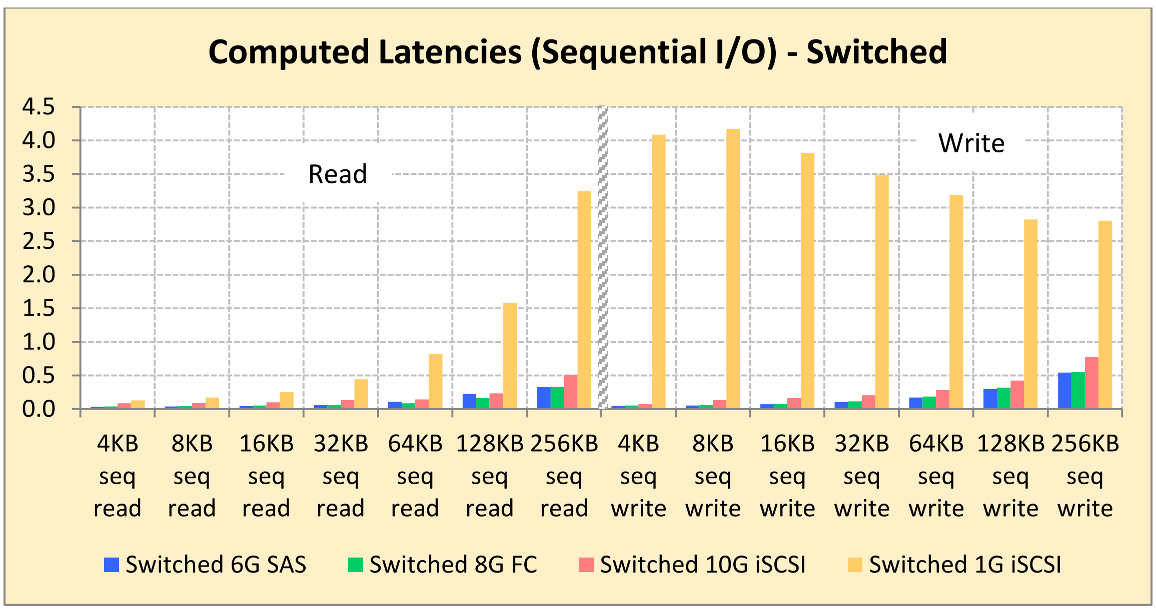


Figure 11 – Computed Latencies – Sequential I/O – Switch Connected

Appendix 3 – Evaluation Environment

These tests were conducted in the Demartek lab in Arvada, Colorado using Demartek server, storage and networking infrastructure.

<u>Server</u>

- IBM x3650, Dual Intel Xeon E5345, 2.33GHz, 8 total cores
- 32GB RAM (DDR2, ECC)
- Internal SAS disk array for boot volume
- Windows Server 2008 R2 SP1
- Microsoft SQLIO I/O workload generator

<u>Adapters</u>

- 6G SAS: LSI 9200-8e SAS HBA
- 8G FC: Emulex LPe12002 FC HBA
- 10G iSCSI: Intel X520 Server Adapter
- 1G iSCSI: Motherboard NIC

<u>Switches</u>

- 6G SAS: LSI SAS6160, 16-port
- 8G FC: Brocade 300, 24-port
- 10GbE: Cisco Nexus 5020, 40-port
- 1GbE: Dell PowerConnect 2748, 48-port

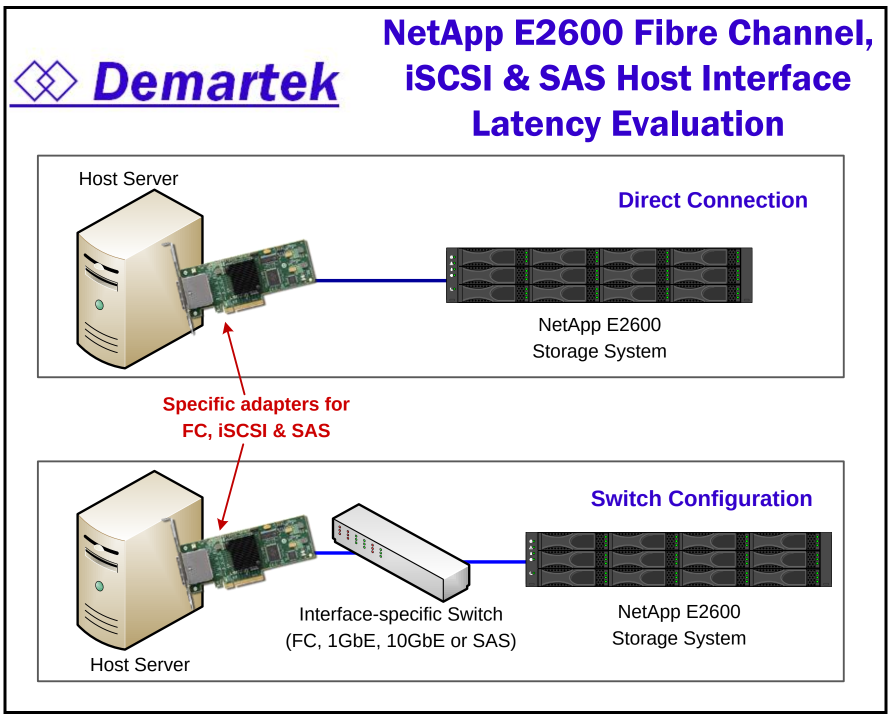


Figure 12 – Evaluation Configurations