Evaluation Report: Maximizing the Potential of All-Flash Storage with end-to-end HPE StoreFabric Gen 5 16Gb Fibre Channel

Executive Summary

At Demartek, we've been declaring for years that flash storage—specifically all-flash arrays—and high-speed storage networks are made for each other. Flash storage systems are just too fast, and capable of too much bandwidth, to deploy on slow networks. Optimal performance cannot be realized when the storage area network (SAN) cannot bring the real power of flash to the server. Return on investment (ROI) and application performance suffer when this happens. Gen 5 16GFC HBAs and switches reached the market in 2011, bringing enough performance to the SAN to harness the real potential that flash storage has to offer. Storage vendors were slower in rolling out Gen 5 16GFC targets, but these are now common enough that storage customers should make them a requirement for any all-flash system purchase.

Application owners understand the value of an infrastructure able to support gigabytes (GB) of data per second with sub-millisecond I/O latency. Real-time processing, bandwidth intensive data streams, and highly transactional applications demand a SAN that won't limit I/O from high-performance storage devices. HPE delivers end-to-end Gen 5 16GFC through OEM relationships with Fibre Channel industry leaders QLogic, Broadcom (Emulex), and Brocade. HPE StoreFabric Q and E-series HBAs, in tandem with StoreFabric B-series switches, puts 16 Gb/s of Fibre Channel per port at the host, switch, and storage targets to achieve optimal performance from all-flash storage systems.

Understanding that application owners and data center managers need compelling reasons to commit to upgrading older SAN technology, Demartek evaluated a common transactional database application on an all-flash storage system, running across both 8GFC and Gen 5 16GFC SANs. We feel that the data and conclusions presented in the paper should convince businesses deploying all-flash that a Gen 5 16GFC or better SAN is the only way to maximize the return on investment in high-performance storage systems. Older SAN technology simply robs all-flash products of valuable bandwidth and response time potential.

Key Findings

Transaction-intensive database workloads are sensitive to storage I/O latency and, in the case of particularly busy applications, storage bandwidth. This evaluation's Online Transaction Processing (OLTP) workload, deployed on an HPE 3PAR StoreServ 8450 All-Flash Storage System, saturated a single port of 8GFC bandwidth. This I/O limit drove up latency due to the queuing of I/O requests and placed a hard limit on the number of database transactions the server could process each second. By replacing 8GFC components with Gen 5 16GFC, we doubled the transaction rate by improving bandwidth and IOPs more than 70% while cutting latency by a similar percentage.

Demartek	Transactions per second	Bandwidth	IOPS	Average I/O Latency
8GFC	2625	770 MB/s	94000	0.5 ms
Gen 5 16GFC	5370	1330 MB/s	161600	0.16 ms
Percent Improved	104%	70%	72%	68%

Figure 1 – Performance metrics of OLTP workload and storage performance across a single port of Fibre Channel, measured from the application server.

This evaluation involved only the upgrade of 8GFC infrastructure with HPE StoreFabric Gen 5 16GFC HBAs and switch to an already configured application workload environment. The purpose was to highlight where legacy technology was throttling the throughput and increasing the response time of I/O from the HPE 3PAR StoreServ 8450 All-Flash Storage System. No application or hardware tuning was done to optimize application performance over Gen 5 16GFC. With additional application loading or tuning, performance would likely have been driven even higher.

HPE Gen 5 16GFC and HPE Smart SAN for 3PAR Software and StoreFabric SANs

Gen 5 16GFC has been available in the HPE StoreFabric Storage Networking product line for some time. The technology doubles the bandwidth and IOPS of older 8GFC devices and can cut latency in half. These are huge improvements for high-end storage systems and demanding applications. By OEMing proven Gen 5 16GFC technology from Brocade, QLogic and Broadcom (Emulex), HPE guarantees interoperability, performance, and scalability in the FC SAN. Backwards compatibility (to 8GFC & 4GFC) keeps older, legacy devices operational and supportable until it is feasible to upgrade or decommission them.

HPE has also taken advantage of the enhanced FC feature set that is part of Gen 5 16GFC infrastructure, to incorporate aspects of SAN management into what HPE has branded "Smart SAN." HPE Smart SAN offers simplified SAN deployment and management with a number of advanced features.

- Automated target driven zoning simplifies the zoning process
- Embedded in the firmware of Gen 5 16GFC HBAs, switches and 3 PAR Storage Systems—no additional software needed
- Control through the 3PAR CLI
- Easy troubleshooting from a single point in the fabric
- A diagnostic ready fabric

Transactional Database Workload Description and Test Platform

We could have deployed any of number of different workloads for this evaluation. However, database applications and their performance demands are well understood by most storage administrators, data center managers and application owners. Therefore, we selected a transactional database as the I/O generator for this test case. Relational databases and OLTP workloads represent a large portion of applications deployed in modern data centers and these applications can produce a considerable number of mid-sized I/O requests (8-16KB per I/O being common). Particularly busy systems will also demand a lot of storage bandwidth. Low transactional latency—the amount of time the application spends performing a transaction—is very important for these workloads. The time a storage system takes to respond to I/O will have a significant impact on the overall application latency. Where performance is critical, due to real-time processing needs or customer experience, application owners are increasingly choosing flash as the storage media.

To demonstrate the benefit of Gen 5 16GFC for a real application, as opposed to a modeled, synthetic and predictable I/O engine, real database and application software were deployed for this evaluation. Microsoft SQL Server, running on an HPE DL380 server with storage provided by an HPE 3PAR 8450 StoreServ All-Flash Storage System provided the test platform. The database was configured with real tables, indexes, stored procedures, etc… and populated with data to mimic a hypothetical financial institution's OLTP system. Microsoft Benchcraft executed a transaction profile modelled on this type of business' day-to-day processing activity. Roughly 90% of all transactions were read I/O with about 10% writes.

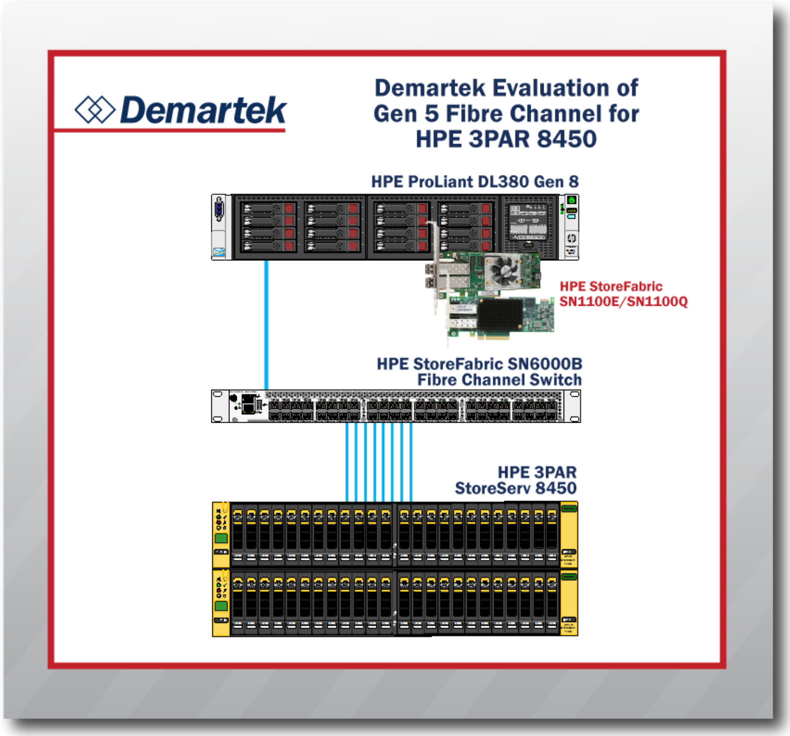

Figure 2 – The HPE Gen 5 16GFC testbed

A legacy 8GFC SAN including the server HBA and an HPE StorageWorks 8/40 8GFC switch were deployed as the baseline environment. The HPE 3PAR 8450 only comes with Gen 5 16GFC targets, so this was the sole exception to an end-to-end 8GFC SAN[1]. The workload was configured to reach single-lane 8GFC line rate bandwidth, without driving average latencies higher than a millisecond. When this was achieved, and metrics collected, the 8GFC SAN was replaced with a Gen 5 16GFC SAN (server HBA and SN6000B switch) and the workload re-run with the exact same parameters.

The decision to demonstrate this workload across a single lane/port of Fibre Channel was made retrospectively. When using two ports of Fibre Channel, upgrading to Gen 5 16GFC supplied the DL 380 with enough I/O to drive the CPUs to full utilization, creating a server bottleneck to application performance and making it impossible to calculate the true performance delta between the two generations of SAN technology.

[1] With storage targets already at Gen 5, I/O leaving the storage system already has the benefit of lower latency from technology improvements inherent in Gen 5 over 8GFC. The baseline will have better performance than a true end-to-end 8GFC SAN for this reason. The impact of 8GFC will not take effect until I/O packets reach the switch.

Performance Metrics

The three key metrics highlighted for SAN and storage system analysis are I/Os per second (IOPS), bandwidth, and latency:

- **IOPS** – I/Os per second – a measure of the total I/O operations (reads and writes) issued by the application servers.
- **Bandwidth** – a measure of the data transfer rate, or I/O throughput, measured in bytes per second or Megabytes per second (MBPS).
- **Latency** – a measure of the time taken to complete an I/O request, also known as response time. This is frequently measured in milliseconds (one thousandth of a second). Latency is introduced into the SAN at many points, including the server and HBA, SAN switching, and at the storage target(s) and media.

It is important to consider all three metrics when evaluating the performance of a SAN. These three contribute to how well the storage will support an application. IOPS drive bandwidth. The number of IOPS times the I/O request size determines the amount bandwidth delivered. The database application used for this evaluation performs predominantly 8-16 kilobyte I/Os.

Latency can be critical even though it doesn't necessarily have a direct effect on IOPS and bandwidth[2]. It does have a very significant effect on application performance and user experience. Unlike IOPS and bandwidth, where more is better, with latency, the goal is to keep it as low as possible. The impacts of latency vary with the workload deployed. Some applications have a greater tolerance, while other applications are negatively impacted by even small increases in latency. Batched processing and streaming I/O may be marginally affected by higher latency, whereas extremely fast response time is required to support applications attempting to achieve near real-time performance. So as to include all round-trip latency contributions, metrics were collected at the server for this evaluation.

All-flash storage systems have no problem achieving very high IOPS and bandwidths. These two metrics are really only of interest insomuch as bandwidth is used to identify the point where the 8GFC HBA (and switch port) reaches saturation. Going into this evaluation, we were aware that our workload would stress neither bandwidth nor IOPS of

[2] Latency may have an indirect effect on IOPS and bandwidth. Lower latency means I/Os are serviced by the SAN more quickly, which enables the server to process I/O in a shorter time interval. If this processing then demands additional I/O, IOPS and bandwidth will increase (up to SAN limits) to satisfy it.

the storage system. Latency has been emerging as the most critical of the three metrics for performance sensitive applications, particularly in the case of transaction-intensive workloads. Application service level agreements often set maximum permissible transaction latencies, which are directly dependent on I/O response times.

For the purpose of this evaluation, we also included an application based metric, database transactions per second, as a measure of the actual work accomplished by a real-world application.

Results and Analysis of Upgrading from 8GFC to Gen 5 16GFC

With latency being the most critical metric for many OLTP application SLAs, it is the first performance characteristic examined here. We supplied an initial, best case, 8GFC response time by setting workload intensity below line rate (Figure 3). This limits the negative effects of I/O queuing in order to show that even a non-saturated I/O channel will see a response time boost by upgrading to Gen 5 16GFC. We then increased application I/O until 8GFC line rate was achieved.

The baseline 8GFC SAN latency, using about 550 MB/s of bandwidth, was .32 milliseconds (320 microseconds if you prefer). When we saturated the 8GFC link, that latency went up to .65 milliseconds (650 microseconds). The first takeaway, of course, is that the HPE 3PAR StoreServ 8450 is a very fast storage system. Even with an old 8GFC SAN, this system has no trouble delivering sub-millisecond I/O response times. When 8GFC is replaced with Gen 5 16GFC, it gets even better. Steady-state response times were .2 milliseconds, a 38% improvement over the 550 MB/s workload I/O and 69% better than when the workload hit 8GFC line rate. (For the remainder of this analysis, we will only be considering the scenario where line rate was achieved.)

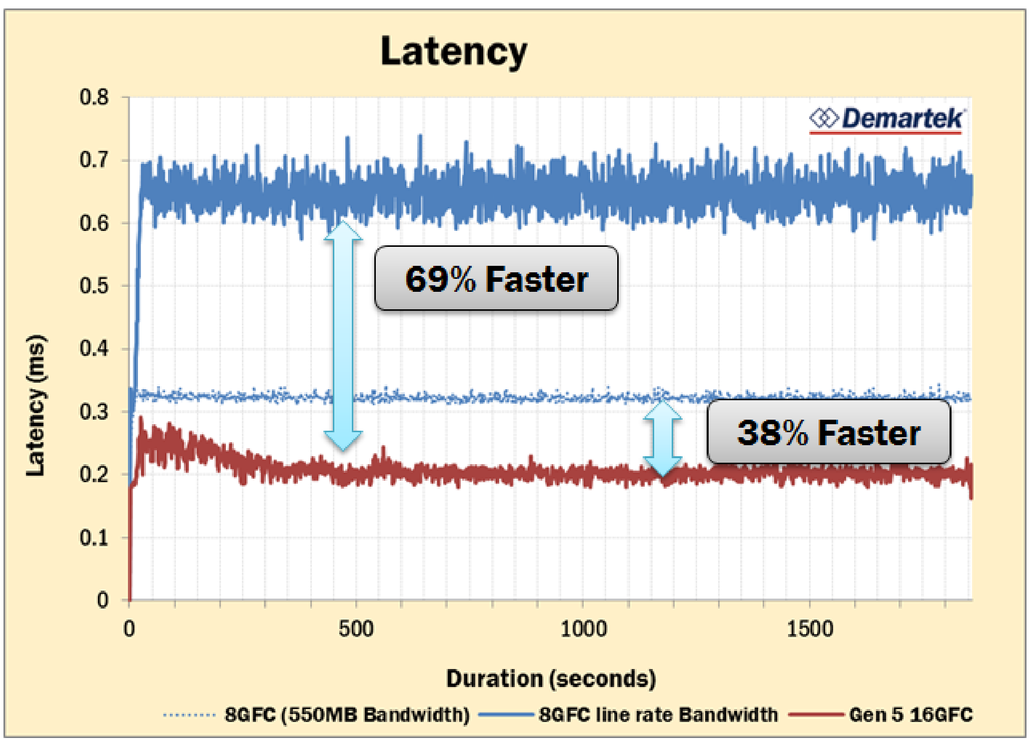

Figure 3 – I/O response times of a workload running in an 8GFC SAN, and Gen 5 16GFC SAN

It is difficult to hit line rate precisely. Real-world applications don't just stop when an I/O limit is reached, instead they queue up requests and the user experience begins to suffer. This happens in data centers every day. However, we don't want I/O requests in queue; we want them serviced. Obviously the flash is more than fast enough to do that; the requests just needed to be delivered quicker. This, in essence, is most of our argument about why 8GFC is a liability for all-flash storage systems.

High bandwidth and IOPS values are easily attained on flash storage, and the HPE 3PAR All-Flash StoreServ 8450 is no exception (Figure 4). 8GFC does exactly what we expect—it limits the application to 8 Gb/s of I/O. Swapping out for Gen 5 16GFC provided an immediate 73% improvement in the amount of data being transmitted. Again, it is clear that the server and the storage system were capable of supporting more I/O than a single channel of 8GFC could deliver.

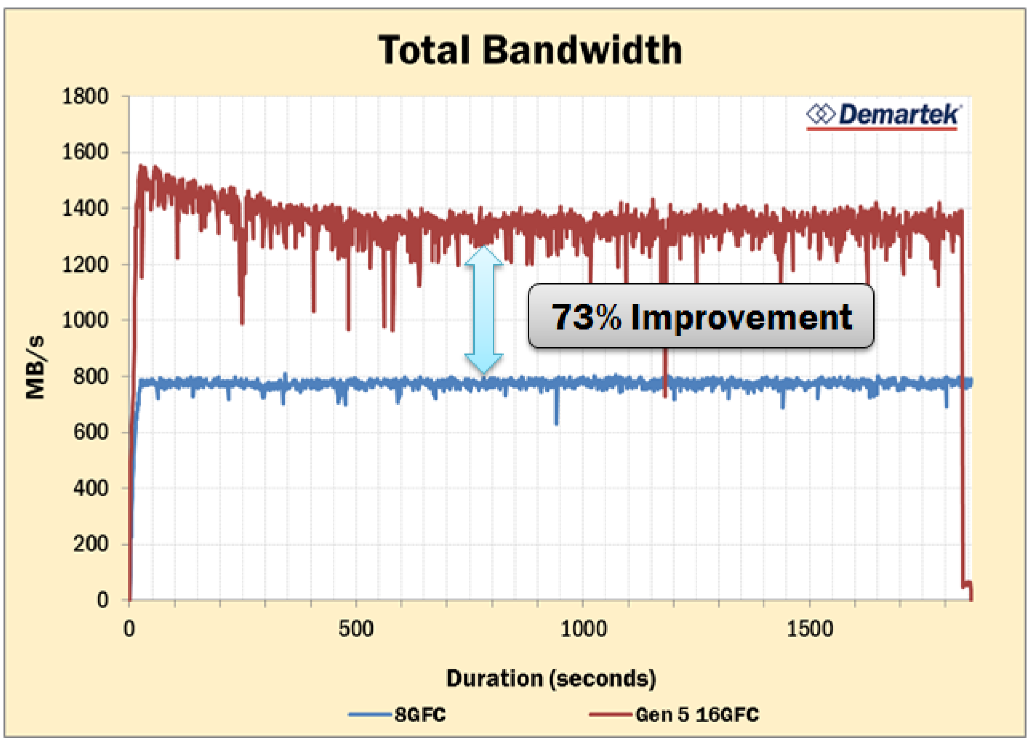

Figure 4 – Improvement in number of SQL Server Transactions per second

These storage metrics are useful for assessing the behavior of the compute system as a whole, but what does this mean with regard to real work performed by that system? Database transactions are the finite units of work executed by the application. We measured a 2x increase—twice the work—in transactional performance by upgrading to Gen 5 16GFC (Figure 5). The transactions executed by the application include dependencies on the completion of prior transactions, which makes this workload particularly sensitive to I/O latency. With the SAN able to deliver I/O faster, subsequent transaction requests are submitted and serviced by the database in a shorter time, driving the number of transactions per second even higher than the improvement in storage bandwidth alone would have provided.

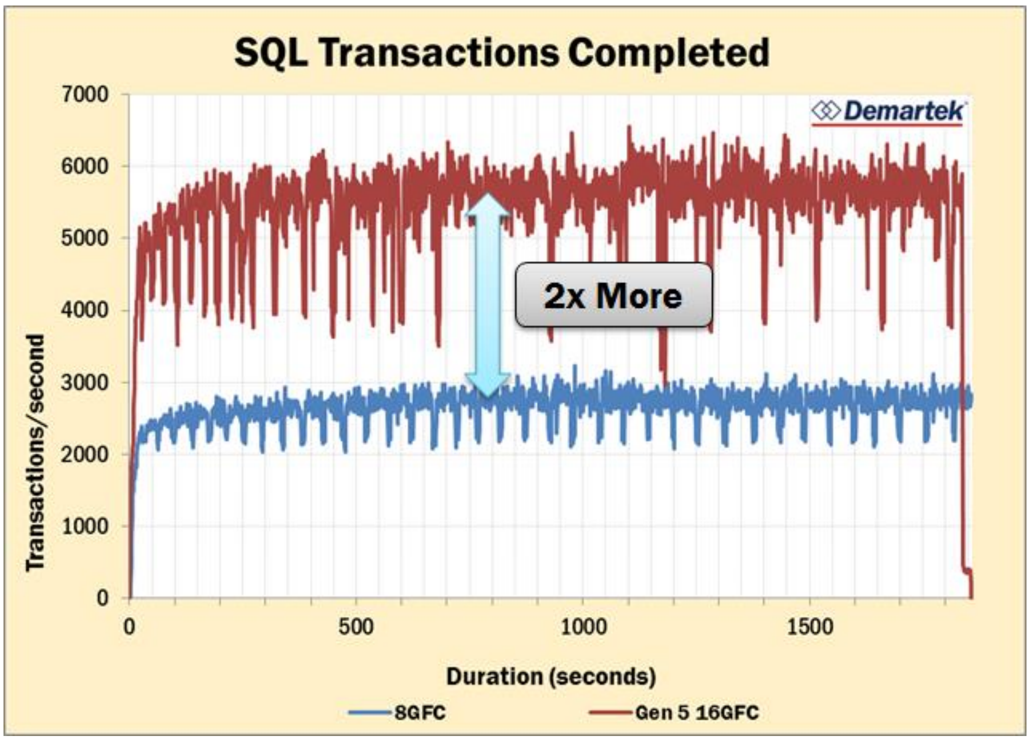

Figure 6 – Improvement in number of SQL Server Transactions per second

This evaluation was performed with a single port of the host-side Fibre Channel. The argument could be made that doubling the number of 8GFC ports will provide the same bandwidth benefits as a single Gen 5 16GFC port. While this is true, it would still not address the latency improvements of the newer technology. A bandwidth-heavy, sequential I/O workload like a data warehouse might scale in performance very effectively by just adding ports. However, a highly transactional application, such as demonstrated in this evaluation, is affected by latency and simply increasing bandwidth will never provide the same degree of scalability. Finally, increasing the number of ports drives up the port count for added cost and complexity across the SAN.

Summary and Conclusion

Customers with intense performance requirements are already invested in flash storage. It makes sense to capitalize on this investment with a SAN technology that enables a real return on that expense. Gen 5 16GFC supports the high bandwidth and IOPS of today's very busy workloads and, perhaps more importantly, permits high-end storage to deliver the very low I/O latency that time sensitive applications demand. This evaluation has attempted to make the case for upgrading the SAN by pointing out where older technology is throttling performance. We have also demonstrated that even when bandwidth limits are not reached, Gen 5 16GFC technology delivers a significant improvement in I/O response time. In real-time processing environments, time *is* money and time saved can be turned into a very real business advantage.

We didn't even begin to consider the additional return on server and blade investments beyond their ability to process more transactions. Naturally, as storage systems delivers I/O more efficiently, processor cores spend less time in wait states and more time doing work. This drives up server utilization, improving application response and increasing customer satisfaction. Add in port reductions on hosts and switches, plus management enhancements through HPE Smart SAN, for an even greater payoff. If extrapolated across a data center, an HPE StoreFabric Gen 5 16GFC SAN has the potential to considerably increase productivity while reducing deployment times and practically eliminating impacts from outages.

Customers are advised to upgrade 8GFC (and older) SANs as soon as possible. 8GFC, though still widely deployed, is a legacy technology. Modern storage systems support Gen 5 16GFC, and any enterprise deploying these devices cannot fully tap their potential without it. With only two generations of backward compatibility in the Fibre Channel specification, older SANs only have a limited supportable lifespan. With computers and storage continuing to improve at an amazing pace, it's time to relegate legacy technology to the dustbin or be left behind.

Appendix A – A Brief Summary of Gen 5 16 Gb/s Fibre Channel

Gen 5 16GFC provides an increased ceiling in terms of the possible performance for each connection, doubling the throughput of the previous generation (8GFC). Database administrators often tune their applications to take advantage of every possible increase in compute system speed, and Gen 5 16GFC connections provide an important performance improvement opportunity. Other benefits include a reduced number of links needed to achieve the same bandwidth as older technologies, reduced power consumption needed to achieve the same bandwidth, and fewer cables to manage. In addition, the higher speed allows fabrics to be connected with fewer inter-switch links (ISLs), which is especially helpful in large fabrics.

Gen 5 16GFC includes re-timers in the optical modules and transmitter training, features that improve link performance characteristics, electronic dispersion compensation, and backplane links.

Because of FC's backward compatibility, host servers can deploy 16 Gb/s FC HBAs and 16 Gb/s FC switches independently, as each works with the two previous generations (8 Gb/s and 4 Gb/s).

Table 1 – Fibre Channel Speed Characteristics[3]

Speed	Throughput (MBps)	Line Rate (Gb/s)	Encoding	Retimers in the module	Transmitter training
Gen 1 (1GFC)	100	1.0625	8b/10b	No	No
Gen 2 (2GFC)	200	2.125	8b/10b	No	No
Gen 3 (4GFC)	400	4.25	8b/10b	No	No
Gen 4 (8GFC)	800	8.5	8b/10b	No	No
Gen 5 (16GFC)	1600	14.025	64b/66b	Yes	Yes

[3] There was never a formal industry adoption of Gen 1 - Gen 4 as a naming convention. While prior generations of FC technology may be referred to by a generational nomenclature, no FC products have ever been marketed with Gen 1 – Gen 4 in the product name or description.

Table 2 – Fiber Optic Cable Link Distance

Speed	Multi-Mode				Single-Mode
	OM1	OM2	OM3	OM4	OS1
Gen 1 (1GFC)	300	500	860	*	10,000
Gen 2 (2GFC)	150	300	500	*	10,000
Gen 3 (4GFC)	50	150	380	400	10,000
Gen 4 (8GFC)	21	50	150	190	10,000
Gen 5 (16GFC)	15	35	100	125	10,000

* The link distance for OM4 fiber optic cables has not been defined for these speeds.

The original version of this document is available at: http://www.demartek.com/Demartek_Brocade_Emulex_QLogic_HPE_16GFC_All-flash_Benefits_Evaluation_2016-06.html on the Demartek website.