BROCADE

Cloud Optimized Performance: I/O-Intensive Workloads Using Flash-Based Storage

Brocade continues to innovate by delivering the industry's first 16 Gbps switches for low latency and high transaction workloads.

Table of Contents

Introduction

As customers look to Fibre Channel storage area networks (SANs) for building private storage cloud services for their enterprise data centers, some of the key attributes are:

• Consolidated and highly virtualized pools of compute, storage and network resources

• Secure and efficient use of inter-fabric connectivity

• Lower capital and operational costs, higher asset utilization

• On-demand and efficient provisioning of application resources through automated management

Brocade has developed solutions to address these key attributes leveraging its seventh-generation "Condor3" ASIC features in the 16 Gbps platform, Fabric Operating System (FOS) v7.0, Brocade Network Advisor, Brocade fabric adapter technology, and is working with transceiver vendors to address these key requirements. In order to

enable customers to achieve these goals, Brocade is delivering key technologies (see Figure 1) that would allow customers to:

• Scale up/out based on business growth

• Secure data and optimize inter-data center bandwidth utilization

• Reduce CapEx and OpEx cost with built in diagnostics and SAN management tools

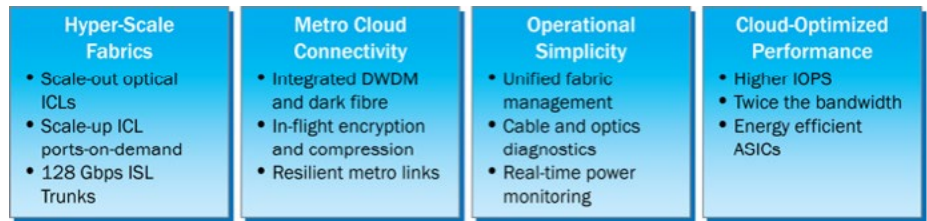• Optimize both bandwidth and Input/Output Operations Per Second (IOPS while being energy efficient

Hyper-Scale Fabrics	Metro Cloud Connectivity	Operational Simplicity	Cloud-Optimized Performance
• Scale-out optical ICLs • Scale-up ICL ports-on-demand • 128 Gbps ISL Trunks	• Integrated DWDM and dark fibre • In-flight encryption and compression • Resilient metro links	• Unified fabric management • Cable and optics diagnostics • Real-time power monitoring	• Higher IOPS • Twice the bandwidth • Energy efficient ASICs

Figure 1: Enabling Private Storage Clouds.

Cloud-Optimized Performance

This is one of four technical briefs addressing Brocade solutions for helping customers transition to private storage clouds. To cater to new emerging workloads of Private Cloud, Brocade has developed 16 Gbps switches to support both high IOPS and high bandwidth.

Evolution of Data Center Technologies

In the mid 1990s, enterprise storage was rarely the bottleneck in overall system performance. The advent of Fibre Channel interfaces made it particularly rare to find a server that could saturate even a single storage port. Customers would fan out storage controllers to dozens—or even hundreds—of servers to leverage their investment, and rarely would this approach cause storage to bog down.

But in the latter part of this decade, several technologies combined to change this dynamic. Game-changers included multi-core processors, faster PCI Express I/O bus speeds, increased speed and capacity of Random Access Memory (RAM), and server virtualization. Now, a single enterprise server can—and often will—exceed the capabilities of an enterprise storage array that uses traditional magnetic disks.

IT professionals are handling this by adopting faster storage media. According to the independent research organization Gartner, it is necessary for enterprises to move to Solid State Disk (SSD) if they want to attain the benefits of modern server architectures, as shown in Figure 2.
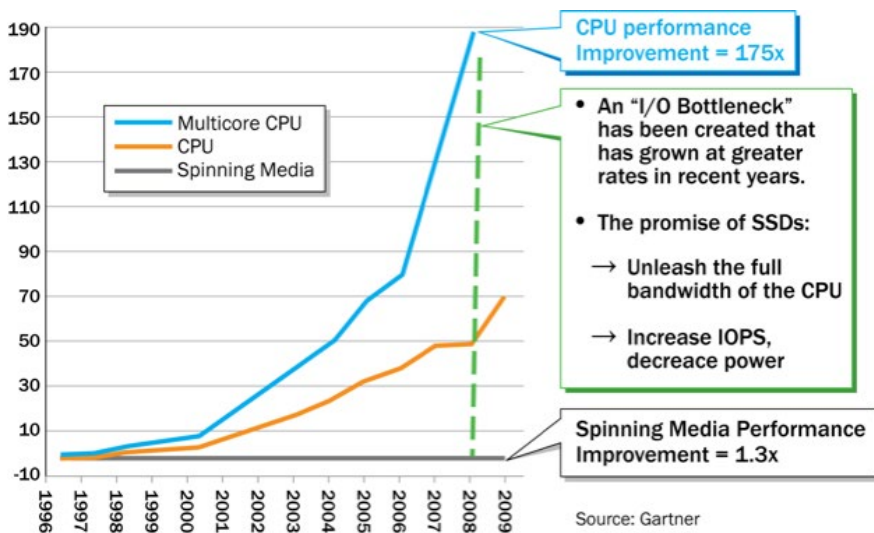
The chart shows three lines tracking performance from 1996 to 2009, with a y-axis ranging from -10 to 190.

Legend:
- Multicore CPU
- CPU
- Spinning Media

Annotations:
- CPU performance Improvement = 175x
- An "I/O Bottleneck" has been created that has grown at greater rates in recent years.
- The promise of SSDs:
  - → Unleash the full bandwidth of the CPU
  - → Increase IOPS, decreace power
- Spinning Media Performance Improvement = 1.3x

Source: Gartner

Figure 2: Gartner Analysis of SSD Potential.

The theory that SSD improves application performance works in the real world, but only if there are no other bottlenecks between the server and its SSD storage. The good news is that modern adapters, such as the Brocade 1860 Fabric Adapter, are capable of sustaining very high IOPS rates and throughput. While modern enterprise RAID controllers are fast, performance gains by adding Solid-State Drives is limited since the bottleneck is the controller itself. By deploying technologies referenced in Figure 2 along with Brocade's new lineup of 16 Gbps Fibre Channel products will enable customers to deploy infrastructure with no performance bottlenecks.

One benefit of moving to 16 Gbps Fibre Channel is obvious: If you increase the bandwidth between a server and an SSD array, you will have a fatter pipeline for data movement. This will reduce backup and restore times, speed up migrations, and benefit large I/O applications, such as video streaming or moving virtual machine images.

However, while SSDs can increase throughput, IOPS is the driving performance factor in many applications. This is especially true for customers who adopt SSDs for web farms to service a large user base with lots of small IOPS. Besides the pure bandwidth benefit for deploying 16 Gbps Fibre Channel, faster network is also beneficial for IOPS-intensive applications. The reason for this is that the Brocade 16 Gbps Fibre Channel ASICs have features specifically designed to reduce network delay and congestion. This includes more advanced load balancing and buffer management and fast cut-through routing.

Why Latency Matters?

To understand why delay matters so much in storage performance, it is necessary to understand how upper-layer storage protocols interact between endpoints. SCSI read operations are simple: the SCSI initiator initiates a request for data and the SCSI target responds with all of the requested data without further acknowledgements or round-trips across the connection. However, the SCSI write operation is more involved—it requires two round-trips between the SCSI initiator and target. The following example applies to all major storage protocols: SCSI, iSCSI, ATA, AoE, Fibre Channel, and so on.
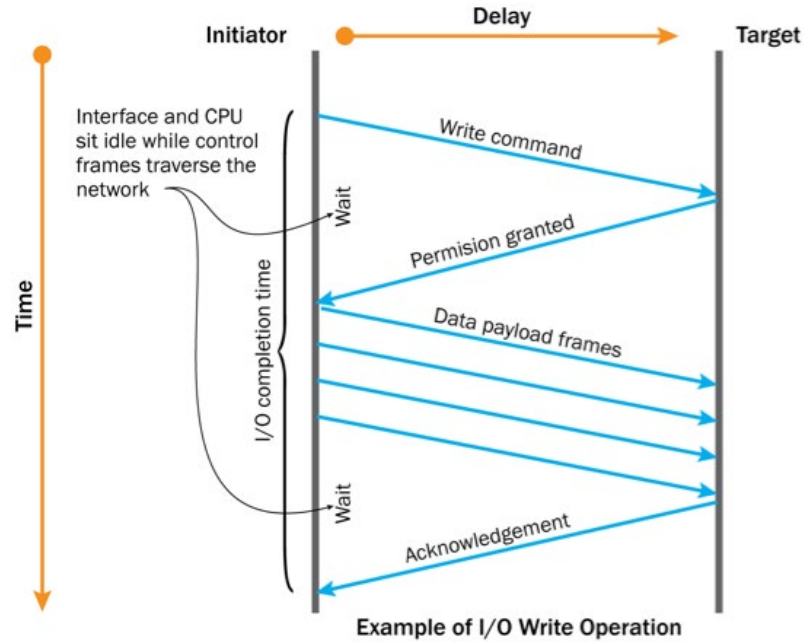
Figure 3: Anatomy of a Storage I/O at 8 Gbps Fibre Channel transfer rate.

Figure 3 shows the structure of a write between a server and an array. The initiator (server) is on the left, and it wants to write data to the target (storage) on the right. When it sends information to the target, delay is incurred, as represented on the x axis. Delay causes time to pass, as shown on the y axis. Intuitively, as delay increases, so does the time to get information between endpoints. What may be less intuitive is how this impacts application performance.

Looking at Figure 3, you can see that an I/O write consists of several commands surrounding a data transfer. First, the initiator asks for permission to send the write I/O. Then the target gives permission. The initiator must receive this before the I/O can be sent. After the I/O has been received, the storage will acknowledge completion. The time between the first command and the last is the I/O time, and it is this metric that determines how many IOPS the application can sustain.

The critical point is that the server cannot begin to transfer data without permission from the storage, and it cannot move on to its next task until the acknowledgement is received. This causes the server CPU to sit idle, waiting for storage activity to complete. These waiting periods are represented in the figure as triangles. If the delay between the server and storage increases, the triangles "open up" wider, which means fewer IOPS. Conversely, if delay is reduced, the triangles become narrower, which means higher IOPS, as shown in Figure 4.
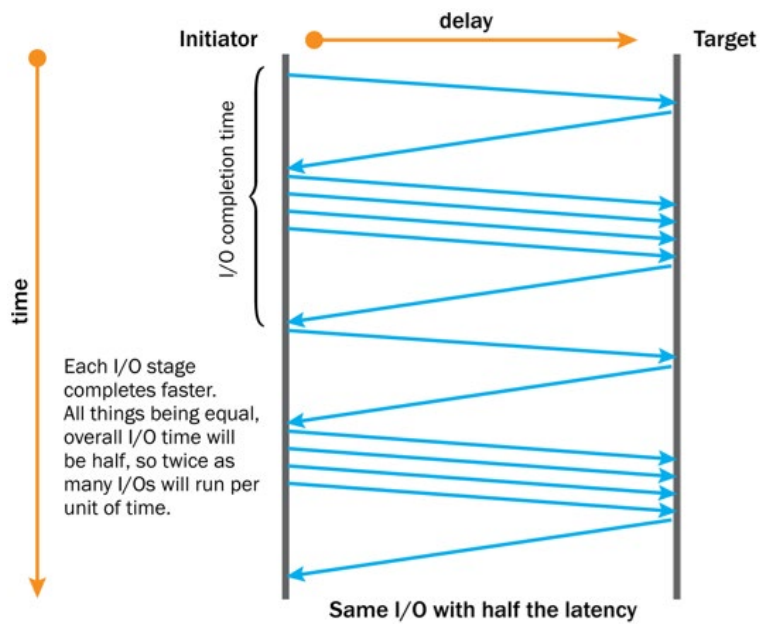
Figure 4: I/O Diagram After Upgrade to 16 Gbps Fibre Channel transfer rate.

Why 16 Gbps Fibre Channel Matters

A Brocade 16 Gbps switch ASIC can switch up to 420 million frames per second or 8.75 million frames per second per port (twice that of 8 Gbps switch ASICs), and this helps with high I/O transactions. A Brocade 16 Gbps Fabric Adapter can also drive over 500K IOPS per port. 16 Gbps Fibre Channel has lower delay because it takes half as long to get a frame onto the wire versus 8 Gbps Fibre Channel, or a quarter as long as 4 Gbps Fibre Channel. Looking back at Figure 4, you can see how a faster traversal of some frames would result in a faster turnaround time for commands, and thus an increase in IOPS.

Example: eCommerce Applications

Theory is important, but it is also helpful if new technologies produce benefits in the real world application. For example, if a customer running a server farm—where each server is processing transactions like eCommerce orders—where the goal is to minimize the number of servers and storage arrays required to support large customer order volumes migrating to a new lower physical footprint and low latency infrastructure may be the right solution. If each server supports a higher IOPS rate that means faster execution on each individual operation, and more operations handled by any given server.
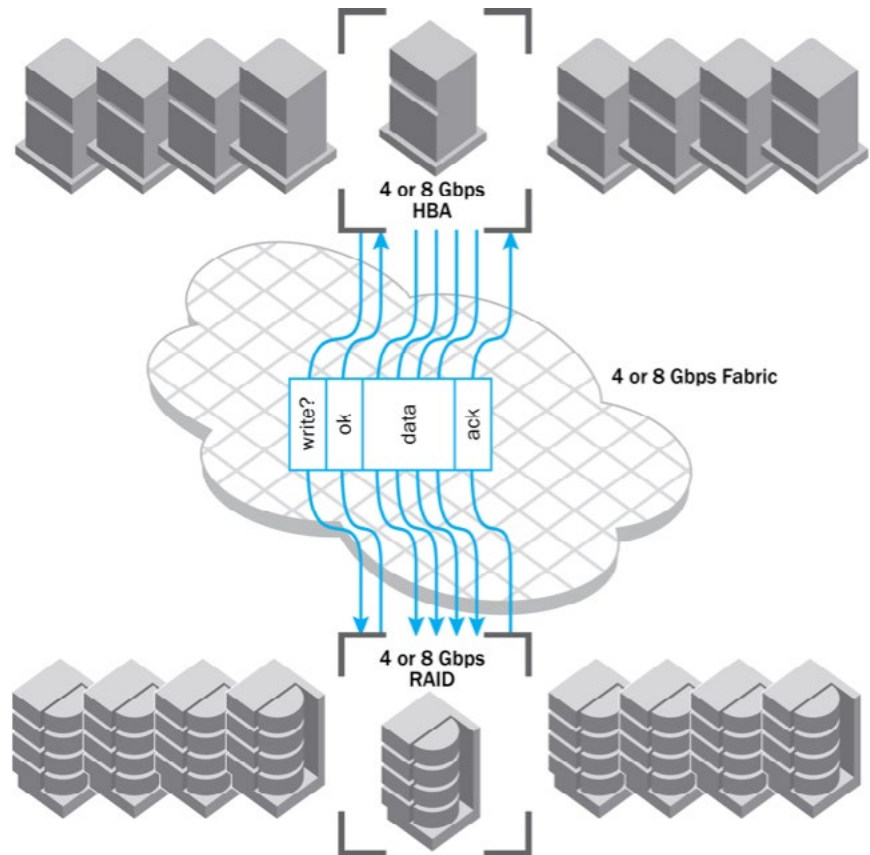
Figure 5: Traditional FC Server Farm I/O Illustration – single 4KB block write flow.

In a traditional web farm (figure 5) IOPS performance is achieved using RAID 5 and increasing cache size on the controller. Historically, there would be bottlenecks inside servers or disk arrays. For a typical single web transaction process a write to the database for an ecommerce order could take 28 microsecond I/O time (initial request, write data transfer and acknowledgement) of which the actual fabric latency would be (5 frames (3 control frames and 2 data frames for a 4KB/ block write) * 700 nanoseconds) or 3.5 microseconds in a switch or a Brocade DCX backbone with local switching.

However, with more and more consumers migrating to web based ecommerce applications, companies are moving to take advantage of faster network speeds, multi-core CPUs and SSD arrays to free up applications from the traditional bottlenecks and reduce the total datacenter footprint by consolidating multiple storage arrays to achieve IOPS requirements. The following figure 6 shows how the new architecture would look for high IOPS transactions.
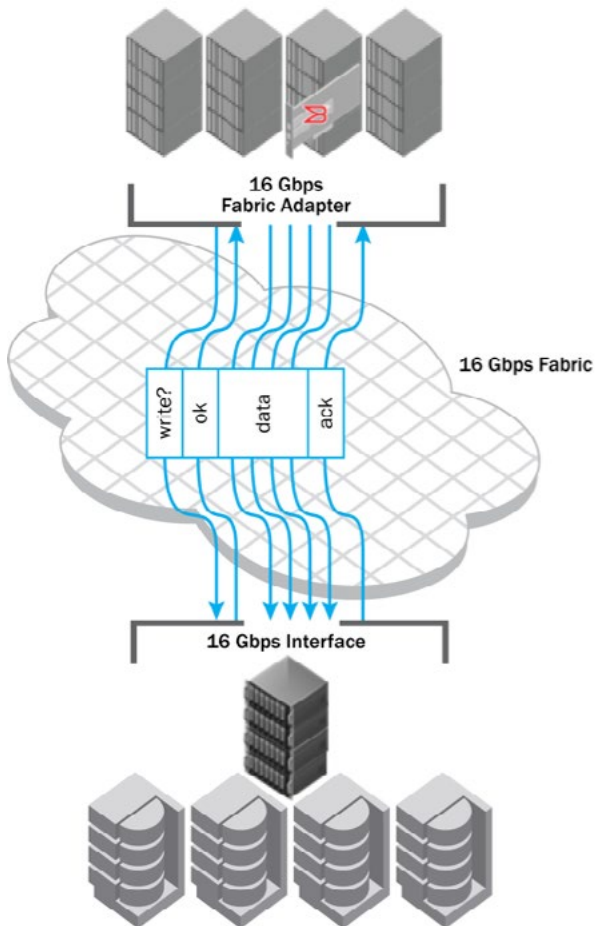

Figure 6: 16 Gbps Server Farm I/O Illustration – single 4kB write flow.

In this example, by going to the multi-core servers with Brocade 1860 Fabric Adapter, 16 Gbps switch, and SSD arrays twice as many frames can traverse the fabric within the same timeframe compared to a previous generation of Fibre Channel fabrics. This reduced the amount of time that servers spend in a wait state, and proportionally increases the volume of transactions that can be handled in the fabric. In the new architecture, the SSD array would handle the SCSI requests speeding up the response times for higher IOPS.

Of course, there are always delay variables outside of the SAN. After a server gets an I/O acknowledgement, it takes some time to get that information to the application, and upgrading to 16 Gbps will not impact this time. There is lag inside an application and even inside a SSD. Also, the benefit is dependent on I/O size—the larger the I/O, the more benefit from a higher 16 Gbps bandwidth, but the less benefit from its lower delay. The number of simultaneous outstanding I/Os is also a major factor.

The result is that real-world applications will not necessarily obtain a linear benefit from upgrading 8 Gbps to

16 Gbps. However, in most environments, it is reasonable to assume that a substantial benefit will accrue.
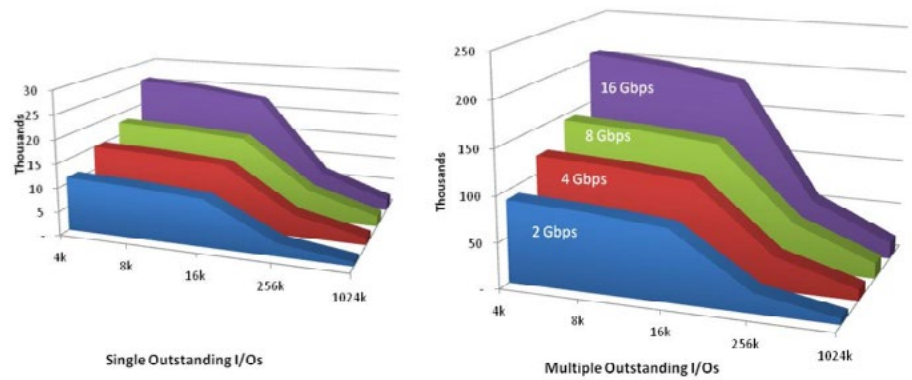
Figure 7: IOPS as a Function of Fabric Speed and Outstanding I/Os.

Figure 7 shows some fairly typical results from upgrading older storage to SSD in combination with Fibre Channel fabric upgrades as shown in Figure 6. Metrics are presented for a server that allows only a single outstanding I/O at a time and a multi-I/O server that supports eight outstanding I/Os at time. Running 4 kilobyte I/Os at 16 Gbps could increase the performance to over two hundred thousand IOPS per port with higher levels of outstanding I/Os achieving even higher IOPS levels.

Summary

Clearly, as SSDs proliferate and servers increase in speed, it will become more important to decrease network delay. With multi-core processors and faster SSD arrays, 16 Gbps Fibre Channel infrastructure applications no longer need to deal with bottlenecks in the server, fabric or the storage.

To find out more about the Brocade DCX 8510 Product family, the Brocade 6510, the Brocade 1860 Fabric Adapter, Fabric Operating System (FOS), and Brocade Network Advisor, contact your sales representative or visit: http://www.brocade.com/privatecloud.

Corporate Headquarters
San Jose, CA USA
T: +1-408-333-8000
info@brocade.com

European Headquarters
Geneva, Switzerland
T: +41-22-799-56-40
emea-info@brocade.com

Asia Pacific Headquarters
Singapore
T: +65-6538-4700
apac-info@brocade.com

BROCADE