Buzz Fibrechannel - To 16G and Beyond!

Mike Blair – mblair@cisco.com
David Lytle – dlytle@brocade.com

7 February 2013 (3:00pm – 4:00pm)

Session 13012
Abstract

• In this jointly presented session, the major players in storage networking will discuss:
  
  • FICON speed roadmap - per the standards.
  • Current customer trends in bandwidth utilization.
  • Do you need 8G, do you need 16G, do you need 32G?
  • How does zHPF play into FICON speeds?
  • What about FCoE - how does this play into FICON?

• At the end, there will be time for Q&A.
Agenda

• Trends and Drivers
  • Bandwidth Drivers
  • Fibre Channel Speed Evolution

• FICON Influences
  • Channel Speed Evolution
  • zHPF

• 16G and Beyond
  • Fibre Channel Roadmap
  • FCoE

• Let’s Talk about Light
  • Modal Dispersion
  • Light in Flight
  • Measuring Light Signals
Bandwidth Drivers
Fibre Channel Speed Evolution

TRENDS AND ROADMAPS
What is driving bandwidth demand?

- Applications increasing in scale and number
- Server virtualization
- Multi-core processors
- Large Memory increases
- Solid State Disks (SSD)
- Faster PCIe rates

The Internet Minute demands greater bandwidth and faster deployment from telecommunication manufacturers, operators and service providers. (Courtesy of Intel)
Prolific Applications

Server Virtualization

- Applications keep growing in number and breadth
- Multiple servers need to access data from shared storage
- Database applications drive bandwidth
- Server virtualization creates multiple virtual machines (VMs) for each application, so each physical server is producing more Input/Output (I/O)

© 2012-2013 Brocade & Cisco - For San Francisco's Spring SHARE 2013 Attendees
Faster Processors

- IBM has the Power7 that has 8 cores and supports 50 GBps of peak IO and directly interconnects 32 of these processors on a server.
- NehalemEX has 8 cores and 16 threads and uses Intel QuickPath Interconnect at 6.4 Gigatransfers per second delivering up to 25 GigaBytes/second (GBps).
- AMD has 8-core and 16 core processors that support 32 threads and HyperTransport 3.0 to support 4.8 gigaTransfers/second.
- Sun’s UltraSparrcT3 chip has 16 cores and supports up to 128 threads.
- A single, multi-processor server supports 10s or 100s of cores.
Increased Memory in Servers

- Memory has limited virtual servers in the past
- Server performance and number of VMs is dependent on memory capacity in servers
  - Gartner: Midrange servers averaged 32GB of memory in 2009 and were expected to triple to 96GB in 2012
  - Registered Dual-Inline Memory Modules (LRDIMM) already come in 32GB packaging
  - Dell’s 2U PowerEdge R710 supports 144GB of memory
  - Sun SPARC M9000-64 offers 4TB memory capacity
  - VMWARE supports 1TB/server and 255GB/VM
- Memory drives more applications that drive more storage I/O traffic
SSDs – Solid State Drives

- Performance of applications is limited by multiple factors with disk drive latency being one factor
- Order of magnitude improvements in performance
  - While traditional spinning disk drive seek times are in the millisecond range, SSD seek times are in the microsecond range
  - SSDs often referred to as Tier-0 storage while disk drives are Tier-1
  - Capacities in the hundreds of GBs per drive
  - Very energy efficient compared to spinning disks
  - Most SSDs provide over 50,000 I/Os per second per drive
- Texas Memory Systems RamSan-630 storage system supports 500,000 IOPS and 8 GBps (64 Gbps) of throughput

<table>
<thead>
<tr>
<th></th>
<th>Latency</th>
<th>Drive IOPs</th>
<th>Array IOPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDD</td>
<td>2-10 mS</td>
<td>100-300</td>
<td>400-40,000</td>
</tr>
<tr>
<td>SSD</td>
<td>50-250 uS*</td>
<td>40k-150k</td>
<td>50k-500k</td>
</tr>
</tbody>
</table>

* This is based on Flash memory and multiple parallel processing
PCIe Continues Ramping

- PCIe 2.0 increases in speed to support dual ported 16G FC HBAs
- PCIe 3.0 will support quad ported 16G FC HBAs
- But they use multiple lanes (wire links) to do it

<table>
<thead>
<tr>
<th>Number of Lanes</th>
<th>Speed per Lane (MBps)</th>
<th>Directional Bandwidth (Gbps)</th>
<th>Ports Supported</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCIe -1.0</td>
<td>4</td>
<td>250</td>
<td>8</td>
</tr>
<tr>
<td>PCIe -1.0</td>
<td>8</td>
<td>250</td>
<td>16</td>
</tr>
<tr>
<td>PCIe -2.0</td>
<td>4</td>
<td>500</td>
<td>16</td>
</tr>
<tr>
<td>PCIe -2.0</td>
<td>8</td>
<td>500</td>
<td>32</td>
</tr>
<tr>
<td>PCIe -3.0</td>
<td>4</td>
<td>1000</td>
<td>32</td>
</tr>
<tr>
<td>PCIe -3.0</td>
<td>8</td>
<td>1000</td>
<td>64</td>
</tr>
</tbody>
</table>
More Applications Drive more Bandwidth

- 16G FC was designed for servers over the next few years that will use these technologies

More Applications  More Bandwidth  More Data

More Virtual Machines  Multi-core Processors  More Memory  Faster PCIe  Faster SSDs

Complete your sessions evaluation online at SHARE.org/SanFranciscoEval

© 2012-2013 Brocade & Cisco - For San Francisco's Spring SHARE 2013 Attendees
The Evolution of Fibre Channel Speeds

- Five generations of Fibre Channel have been delivered to the market
- Speed doubling about every 3-years
- Fibre Channel dominates the storage market
Generations of Fibre Channel

The newest speed in Fibre Channel - Keep it Serial Stupid

<table>
<thead>
<tr>
<th>Generation</th>
<th>1&lt;sup&gt;st&lt;/sup&gt; Gen</th>
<th>2&lt;sup&gt;nd&lt;/sup&gt; Gen</th>
<th>3&lt;sup&gt;rd&lt;/sup&gt; Gen</th>
<th>4&lt;sup&gt;th&lt;/sup&gt; Gen</th>
<th>5&lt;sup&gt;th&lt;/sup&gt; Gen</th>
<th>6&lt;sup&gt;th&lt;/sup&gt; Gen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electrical / Optical Module</td>
<td>1GFC / GBIC / SFP</td>
<td>2GFC / SFP</td>
<td>4GFC / SFP</td>
<td>8GFC / SFP+</td>
<td>16GFC / SFP+</td>
<td>32GFC / SFP+</td>
</tr>
<tr>
<td>Electrical Speeds (Gbps)</td>
<td>1 lane at 1.0625</td>
<td>1 lane at 2.125</td>
<td>1 lane at 4.25</td>
<td>1 lane at 8.5</td>
<td>1 lane at 14.025</td>
<td>1 lane at 28.05</td>
</tr>
<tr>
<td>Encoding</td>
<td>8b/10b</td>
<td>8b/10b</td>
<td>8b/10b</td>
<td>8b/10b</td>
<td>64b/66b</td>
<td>64b/66b</td>
</tr>
</tbody>
</table>

Complete your sessions evaluation online at SHARE.org/SanFranciscoEval
FC Dominates the Backbone Storage Network

What is the predominant storage network backbone you use?

- By asking for the predominant storage network, we find where the heart of storage networking lies. With 82% selecting FC, the answer is clear.
- From our latest technology roadmap, non-FC storage network technologies are used in greater percentages than appear in this chart. FCoE is in use by 8% of respondents, and 10Gbps Ethernet (used for storage) is in use by 31%.

- Fibre Channel (FC): 82%
- 1Gbps Ethernet: 7%
- 10Gbps Ethernet: 4%
- iSCSI: 3%
- FCoE: 2%
- InfiniBand: 0%
- Other: 1%

Source: The InfoPro, 1H CY 2012
Full Sample: FC Switch Ports – Types and Usage

8GB Still Arriving; Bulk on 4GB

Of these total FC switch ports, break out the types and usage as a percentage:

- ISL
- Server or Device

- ISL 2GB: 23%
- ISL 8GB: 12%
- ISL 4GB: 65%

- Server or Device 1GB: 3%
- Server or Device 4GB: 60%
- Server or Device 8GB: 8%

Source: The InfoPro, Q2 CY 2011

(5/31/11): Full Sample. n=198.
8Gbps Fibre Channel Implementation

Source: The InfoPro, 1H CY 2012
Past and Forecast Adoption of 16G FC

Source: Dell Oro, Q1 2013
Channel Speed Evolution
zHPF

FICON INFLUENCES
Mainframe and Open Systems Time Lines

- 1964: System 360
- 1964-1970: System 360
- 1973: Dynamic Channel Reconnect
- 1974: OS/VS 2 MVS
- 1974: Intel 8080
- 1978: ILM Storage Mgmt (HSM/SMS)
- 1980: Physical Partitions
- 1980: 1st SAN through ESCON Directors
- 1988: Virtual Memory
- 1988: ESCON Storage Mgmt
- 1989: 1st SAN through ESCON Directors
- 1989: Dynamic Channel Reconnect
- 1989: ILM Storage Mgmt (HSM/SMS)
- 1990: 1st HA SAN Director
- 1992: Win 3.1
- 1992: 1st SAN Switch
- 1997: Server VMware
- 2000: FICON GA
- 2000: 1st HA SAN Director
- 2001: 248-port HA SAN Director
- 2006: 1st FC Director with 256 ports
- 2007: Windows Vista
- 2009: FICON Express 8
- 2009: Window 7
- 2010: 2009 System z196
- 2011: FICON Express 8S
- 2011: System z114
- 2012: System zEC12
- 2012: Windows 8
- 2012: 2012 Windows 8
- 2010: 2010 System z196
- 2009: Windows 7
- 2007: Windows Vista
- 2000: FICON GA
- 1998: 248-port HA ESCON Director
- 1998: 1st HA SAN Director
- 1988: 1st SAN through ESCON Directors
- 1988: ILM Storage Mgmt (HSM/SMS)
- 1980: Physical Partitions
- 1980: Dynamic Channel Reconnect
- 1974: OS/VS 2 MVS
- 1973: Virtual Memory
- 1964: System 360

Complete your sessions evaluation online at SHARE.org/SanFranciscoEval

© 2012-2013 Brocade & Cisco - For San Francisco's Spring SHARE 2013 Attendees
FICON performance – Start I/Os
Historical Actuals

I/Os per second
4k block size, channel 100% utilized

Complete your sessions evaluation online at SHARE.org/SanFranciscoEval
FICON performance on System z

I/O driver benchmark
I/Os per second
4k block size
Channel 100% utilized

ESCON 1200
FICON Express4 14000
and FICON Express2 31000
FICON Express8 and FICON Express2 20000
FICON Express8 52000
FICON Express8S 92000

NEW 77% increase

I/O driver benchmark
MegaBytes per second
Full-duplex
Large sequential
read/write mix

FICON Express4 350
4 Gbps z10
FICON Express4 520
8 Gbps z10
FICON Express8 620
8 Gbps z10
FICON Express8S 770
8 Gbps z10
FICON Express8S 620
8 Gbps z10
FICON Express8S 1600
8 Gbps z10

NEW 108% increase

© Copyright IBM Corporation 2012

Complete your sessions evaluation online at SHARE.org/SanFranciscoEval

© 2012-2013 Brocade & Cisco- For San Francisco's Spring SHARE 2013 Attendees
FCP performance on System z

I/Os per second
Read/writes/mix
4k block size, channel 100% utilized

FE4 4 Gbps
FE4 4 Gbps
FE8 8 Gbps
FE8 8 Gbps
FE8S 8 Gbps

10% increase

MegaBytes per second (full-duplex)
Large sequential
Read/write mix

FE4 4 Gbps
FE8 8 Gbps
FE8S 8 Gbps

95% increase

Complete your sessions evaluation online at SHARE.org/SanFranciscoEval
What’s Ahead?

Fibre Channel Roadmap
FCoE

16G AND BEYOND
The Benefits of 16GFC

- 16GFC is 100% faster than 8GFC and 40% faster than 10GE and leads to these benefits
  - Higher performance lets servers process more data
  - Fewer links to do the same job
  - Easier cable and device management
  - Less power consumption per bit

Innovation | Performance | Scalability | Environment Friendly
Characteristics of 16GFC

- Double the throughput over backplanes, 100 meters and 10 kilometers

Fibre Channel Physical Interfaces 5 (FC-PI-5) standardized 16GFC

<table>
<thead>
<tr>
<th>Speed Name</th>
<th>Throughput (MB/sec)</th>
<th>Line Rate (Gbps)</th>
<th>Encoding</th>
<th>Retimers in the module</th>
<th>Transmitter Training</th>
<th>OM1/2/3/4 Link Distance (meters)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1GFC</td>
<td>100</td>
<td>1.0625</td>
<td>8b/10b</td>
<td>No</td>
<td>No</td>
<td>300/500/860/*</td>
</tr>
<tr>
<td>2GFC</td>
<td>200</td>
<td>2.125</td>
<td>8b/10b</td>
<td>No</td>
<td>No</td>
<td>150/300/500/*</td>
</tr>
<tr>
<td>4GFC</td>
<td>400</td>
<td>4.25</td>
<td>8b/10b</td>
<td>No</td>
<td>No</td>
<td>50/150/380/400</td>
</tr>
<tr>
<td>8GFC</td>
<td>800</td>
<td>8.5</td>
<td>8b/10b</td>
<td>No</td>
<td>No</td>
<td>21/50/150/190</td>
</tr>
<tr>
<td>10GFC</td>
<td>1200</td>
<td>10.53</td>
<td>64b/66b</td>
<td>Yes</td>
<td>No</td>
<td>33/82/300/*</td>
</tr>
<tr>
<td>16GFC</td>
<td>1600</td>
<td>14.025</td>
<td>64b/66b</td>
<td>Yes</td>
<td>Yes</td>
<td>15/35/100/125</td>
</tr>
</tbody>
</table>

* FC-PI-5 didn’t standardize distances for OM4 fiber for 1/2/10GFC
Dual Codecs for Backward Compatibility

• To improve the efficiency of the protocols, 16GFC only uses 64b/66b coding that is 98% efficient:
  • 8b/10b codes used for 2/4/8G FC are 80% efficient
  • 16G FC signals cannot use the 8b/10b encoders

• To be backward compatible with 2/4/8G FC, 16G FC ASICs need to support both 8b/10b and 64b/66b coder/decoders (codec) on each link

• During speed negotiation, the transmitter and receiver switch back and forth between the speeds (and the corresponding codecs) until the fastest speed is reached for a given link
Speed Negotiation

- During speed negotiation, the speed dependent switch routes the initialization sequence to the appropriate encoder
  - 64b/66b for 16G FC
  - 8b/10b for 2/4/8G FC
  - The coupler sends the signals from one of the encoders to the SFP+

```
Fibre Channel ASIC

<table>
<thead>
<tr>
<th>Upper Level Processing + buffers</th>
<th>Speed Dependent Switch</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For 16G FC

<table>
<thead>
<tr>
<th>64b/66b Encoder</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coupler</td>
</tr>
<tr>
<td>8b/10b Encoder</td>
</tr>
</tbody>
</table>

For 2/4/8G FC

<table>
<thead>
<tr>
<th>SFP+</th>
</tr>
</thead>
</table>
```
Speed Negotiation Examples

- After Speed Negotiation, the chosen encoder remains static and the link works at the fastest supported speed.

For 16G FC

- 64b/66b Encoder
- 16GFC SFP+
- Link runs at fastest speed of 8GFC so uses 8b/10b encoding
- 8b/10b Encoder
- For 2/4/8GFC

For 2/4/8G FC

- 8b/10b Encoder
- 16GFC SFP+
- Link runs at fastest speed of 16GFC so uses 64b/66b encoding
- 64b/66b Encoder
- For 16G FC

Complete your sessions evaluation online at SHARE.org/SanFranciscoEval

© 2012-2013 Brocade & Cisco- For San Francisco’s Spring SHARE 2013 Attendees
Speeds Double Through 2020

- 32G FC Standard Stabilizing
- 64G FC Standard Starting
Converged Fabric

Why?

- Fewer CNAs (Converged Network adapters) instead of NICs, HBAs and HCAs
- Limited number of interfaces for Blade Servers / Rack Mounted Servers

![Diagram illustrating converged fabric traffic](image-url)
FCoE Enablers

- 10Gbps Ethernet
- Lossless Ethernet
  - Matches the lossless behavior guaranteed in FC by B2B credits
- Ethernet jumbo frames
  - Max FC frame payload = 2112 bytes

Normal ethernet frame, ethertype = FCoE

Same as a physical FC frame
Primary FCoE Use Case

Before Unified I/O

- LAN
- SAN A
- SAN B
- IP Switch
- IP Switch
- FC Switch
- FC Switch
- FC HBAs (FC Traffic)
- NICs (Ethernet Traffic)

After Unified I/O

- SAN
- Storage
- Corporate LAN
- Top-of-Rack Switches
- Converged CNA's (CEE/FCoE Traffic)
- FCoE Switch
- FCoE Switch

Disk array, FICON, and tape will require Fibre Channel for years to come.
FCoE Standards

FCoE is currently an approved standard – but is still evolving

FCoE is Fibre Channel
Standard Approved
June 4th 2009
FC-BB-5

Still Evolving
Adding / Refining:
Distance support
Some recovery cases
Support for Z require
FCIP
FC-BB-6
Current Platforms Supporting FCoE

- Blade servers
  - Cisco
  - IBM
  - HP
- Pizza Box PCs
- IBM Power System Announced, Feb 13, (FCoE and 16G FC)\(^1\)
  - The two-port PCIe2 16 Gb Fibre Channel Adapter for the POWER7+ 710/720/730/740/750/760.
  - The four-port PCIe2 Converged Network Adapter (CNA) for the POWER7+ 710/720/730/740/750/760.
- NetApp Filers
- EMC VMAX (for SRDF only currently)
Past and Forecast Switch Port Sales

Source: Dell Oro, Q1 2013

FC Switch Ports
FCoE Switch Ports

Source: Dell Oro, Q1 2013
10G on LAN On Motherboard (LOM) A Game Changer

10Gbe Deployment Costs

<table>
<thead>
<tr>
<th>SFP+ Top of Rack</th>
<th>10GBASE-T Top of Rack</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fiber</td>
<td>10GBASE-T LOM</td>
</tr>
<tr>
<td>$243 per port</td>
<td>$305 per port</td>
</tr>
<tr>
<td>Fiber</td>
<td>10GBASE-T NIC</td>
</tr>
<tr>
<td>$778</td>
<td>$305 per port</td>
</tr>
<tr>
<td>$256 per port</td>
<td>$5</td>
</tr>
<tr>
<td>$1277 per port TOR</td>
<td>$566 per port TOR</td>
</tr>
<tr>
<td>Direct Attach</td>
<td>10GBASE-T LOM</td>
</tr>
<tr>
<td>$243 per port</td>
<td>$256 per port</td>
</tr>
<tr>
<td>$135</td>
<td>$5</td>
</tr>
<tr>
<td>$634 per port TOR</td>
<td>$310 per port TOR</td>
</tr>
</tbody>
</table>

Complete your sessions evaluation online at SHARE.org/SanFranciscoEval
FC/FCoE Bandwidth Roadmaps

- **FCoE uses Ethernet as its physical transport and is used predominately for**
  - Enterprise data center
  - Converged SAN/LAN networks

- **FC is the predominate Enterprise SAN inter-connect**

- **Roadmap To 100G FCoE**
  - 10G FCoE
  - 40G FCoE

- **Roadmap To 128G FC**
  - 2Gb FC
  - 4Gb FC
  - 8Gb FC
  - 16Gb FC
  - 32Gb FC
  - 2/4/8/10/16 Gb FC and 10 Gb Ethernet/FCoE use the same typical optical/copper assemblies (i.e. OM2, OM3, OM4, twin Ax) with the same SFP+ module connection

**TOTAL Investment Protection!**

Complete your sessions evaluation online at SHARE.org/SanFranciscoEval
Storage Technology Hype Cycle Curve

In 2010, FCoE

Source: Gartner, 2010
Storage Technology Hype Cycle Curve

In 2010, Data Deduplication

Source: Gartner, 2010
Storage Technology Hype Cycle Curve

Source: Gartner, 2010

In 2010, SSD
Reflection / Discussion

• This is all interesting data-points but –
  • When do your applications need greater than 8G?
  • When will the servers have higher speed availability?
  • When will the Disks / Tapes / VTLs have higher speeds?
  • What and When are the Technology inflection points?
  • Which Technology(s) will have solutions sooner?
Modal Dispersion
Light in Flight
Measuring Light Signals

LET’S TALK ABOUT LIGHT
Modal dispersion is a distortion mechanism occurring in multimode fibers in which the signal is spread in time because the propagation velocity of the optical signal is not the same for all modes.

Modal dispersion limits the bandwidth and distance of multimode fibers.

OM3 and OM4
FC Storage Networking Terminology
Fiber Channel Links

- **Photos of Modal dispersion**
  - As you can see, a beam of light travels from side to side as it travels from one end of the cable to the other. This is how fibre optics can transmit data across long distances while not confined to being straight line of sight paths.

Light enters the cable

Light carries through the cable with a little dispersion

Without the cable light dispersion happens quickly
Light in Flight

There is now a camera that can take a trillion frames per second. Below is a photo of light in flight from a laser pointer. The distance of the light shown below is the total distance that light travels in atmosphere in a Femtosecond.

A femtosecond (10^{-15} seconds) is one quadrillionth, or one millionth of one billionth of a second. Put another way: a femtosecond compares to a second, as a second compares to 30 million years.

---

**Bullet of Light**

1,000,000 x Faster

<table>
<thead>
<tr>
<th>Unit</th>
<th>Size</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>attosecond</td>
<td>10^{-18} s</td>
<td>shortest time now measurable by scientists</td>
</tr>
<tr>
<td>femtosecond</td>
<td>10^{-15} s</td>
<td>pulse width on world’s fastest lasers</td>
</tr>
<tr>
<td>picosecond</td>
<td>10^{-12} s</td>
<td>switching time of the world’s fastest transistor</td>
</tr>
<tr>
<td>nanosecond</td>
<td>10^{-9} s</td>
<td>time for molecules to fluoresce</td>
</tr>
<tr>
<td>microsecond</td>
<td>10^{-6} s</td>
<td>length of time of a high-speed, strob light flash</td>
</tr>
<tr>
<td>millisecond</td>
<td>0.001 s</td>
<td>time for a housefly’s wing flap</td>
</tr>
</tbody>
</table>
Measuring Light Signals

- Technology is pushing our capabilities to measure the data in a light signal.
- 20 picoseconds is about our technological capability to be able to measure the rising and falling of light in pulse in order to determine the information that the light pulse is carrying – but a femtosecond of time can carry a lot of data.
References

• Fibre Channel Standard
  • www.t11.org

• Fibre Channel Industry Association
  • www.fcia.com

• Storage Networking Industry Association
  • www.snia.org

• Ethernet Alliance
  • www.ethernetalliance.org
THANK YOU!

SHARE, San Francisco, February 2013
Buzz Fibrechannel - To 16G and Beyond!

Session 13012

Thank You For Attending Today!
5 = “Aw shucks. Thanks!”
4 = “Mighty kind of you!”
3 = “Glad you enjoyed this!”
2 = “A Few Good Nuggets!”
1 = “You Got a nice nap!”

Complete your sessions evaluation online at SHARE.org/SanFranciscoEval

© 2012-2013 Brocade & Cisco- For San Francisco’s Spring SHARE 2013 Attendees