APRIL 2-4, 2013

COMPUTERWORLD SNIA

SNW

Intelligent Architecture for the Data-Driven Business

Analyst Perspective:
Next Generation Storage Networking for Next Generation Data Centers

Dennis Martin
President, Demartek

The original version of this presentation is available here:
http://www.demartek.com/Demartek_Presenting_SNWUSA_2013-04.html

Agenda

- About Demartek
- What drives the need for increased bandwidth?
- Ethernet – 10 Gigabit and futures
- Fibre Channel – 16 Gigabit and futures
- Converged networks (DCB, FCoE, etc.)
- Network (I/O) virtualization (NPIV, SR-IOV, etc.)
- Cabling considerations and recommendations
- Demartek free resources

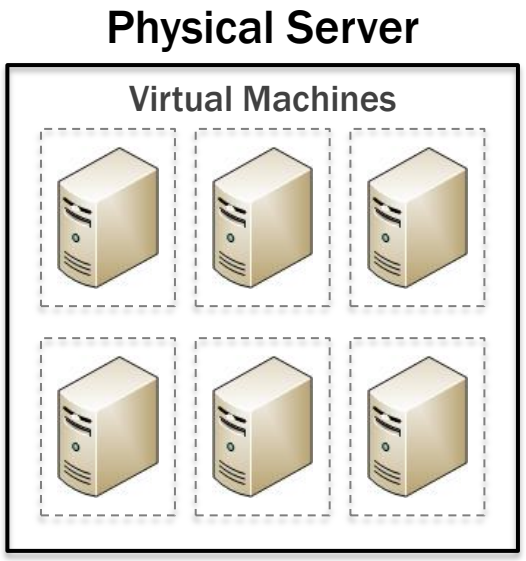
About Demartek

- Industry analysis with on-site test lab
- Lab includes servers, networking and storage infrastructure
  - Fibre Channel – 4, 8 & 16 Gbps
  - Ethernet – 1 & 10 Gbps: NFS, SMB (CIFS), iSCSI & FCoE
  - Servers – 8+ cores, large RAM
  - Virtualization – VMware, Hyper-V, Xen
- We prefer to run real-world applications to test servers and storage solutions
  - Currently testing SSD, 10GbE, 16GFC and other technologies
- Website: www.demartek.com

The Need for More Bandwidth

► Server and Application Growth

- Server virtualization
  - How many VMs per physical server do you deploy?
  - Compare the number of VMs today vs. one and two years ago
- Application growth
  - Applications processing more data today



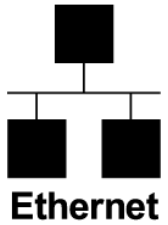
The Need for More Bandwidth

► New Hardware

- New generation of servers (1H 2012)
  - PCI-Express 3.0
    - Up to 40 PCIe lanes per processor
  - New servers support 10GigE on the motherboard
  - More cores per processor, more memory
- SSD
  - Are you deploying enterprise SSDs today?

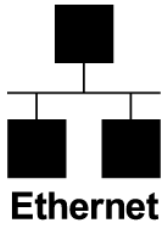
Ethernet

► 10 Gigabit (10GigE or 10GbE)



- 10 Gigabit specification was ratified in 2002
  - 10GBASE-T specification was ratified in 2006
- Adoption of 10GigE increasing over the last 2-3 years
- Blade server chassis have two or more 10GigE ports
- Connector types:
  - SFP+ – Fiber-optic cables and direct attach copper
  - RJ45 – Twisted-pair “traditional Ethernet” (10GBASE-T)
    - 10GBASE-T a few years behind SFP+ but expected to gain acceptance relatively quickly
  - Some 10GigE switches support both SFP+ and RJ45

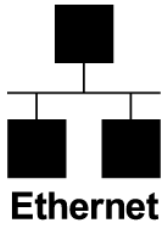
Ethernet



► 1GigE vs. 10GigE

- 1GigE
  - Not unusual to see 4, 6 or 8 NIC ports in a server
    - Requires 4, 6, or 8 Ethernet cables
  - Used for management and network/storage traffic
  - 1GigE NICs can be quad-port, dual-port or single-port
  - Can consume two, four or more I/O slots in a server
- 10GigE
  - A dual-port 10GigE NIC provides bandwidth and failover
  - Good choice for 1U servers that have few I/O slots
  - Slot requirements
    - Dual-port 10GigE NIC – PCIe 2.0 x8
    - Single-port 10GigE NIC – PCIe 2.0 x4 or PCIe 1.0 x8

Ethernet



► 40GigE and 100GigE

- IEEE 802.3ba (40GigE and 100GigE) ratified in June 2010
- The fastest Ethernet cables and connectors today are 10 Gbps per lane or channel
- Higher speeds today are achieved by bundling
  - 40GigE today = 4 x 10 Gbps together
  - 100GigE today = 10 x 10 Gbps together
- 25 Gbps connector samples became available in 2012
  - These connectors support up to 28 Gbps (“25/28G”)
  - 100GigE (future) = 4 x 25 Gbps together
  - 250GigE (future) = 10 x 25 Gbps together
  - End-user products with 25 Gbps expected in 2013 or 2014

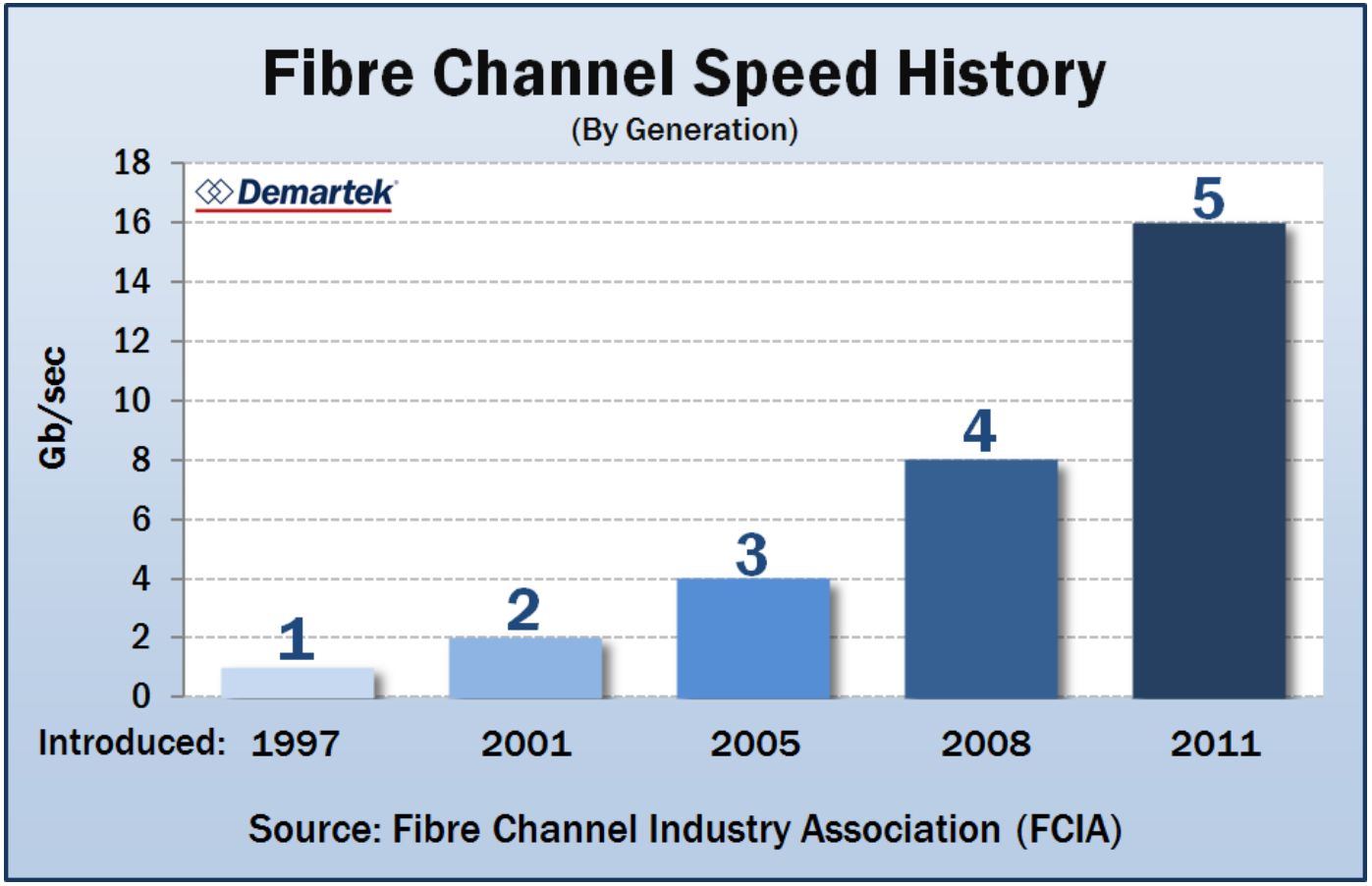
Fibre Channel

► 16 Gigabit (16GFC)

- 16 GFC is backward compatible with 4 GFC & 8 GFC
- Uses 14 Gbps single-lane connectors
  - Doubles speed of 8 GFC due to newer 64b/66b encoding
- The first 16 GFC switches and HBAs shipped in 2011
  - Some of these HBAs can also function as 10 Gb NICs
- 16 GFC storage targets becoming available
- Fibre Channel speeds and server slots (dual-port)
  - 4 Gb: PCI-X 2.0, PCIe 1.0
  - 8 Gb: PCIe 2.0 x4 or PCIe 1.0 x8
  - 16 Gb: PCIe 3.0 x4 or PCIe 2.0 x8

Fibre Channel

► SAN Interface



Fibre Channel

Doubles in speed every 3-4 years

FCoE (10 Gb) introduced in 2009

Fibre Channel

► Disk Drive Interface

- 4 Gb/s was the highest speed for FC interface disk drives

- HDD/SSD vendors have moved to 6 Gb/s SAS for enterprise drives
  - First 12 Gb/s SAS drives announced in Spring 2012

Fibre Channel

► 32 Gigabit and 64 Gigabit

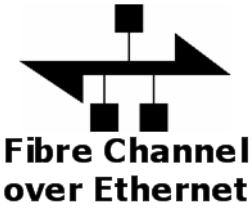
- Formal statement of direction:

  "The INCITS Technical Committee T11 is currently working on the 32 GFC Fibre Channel specifications. The 32 GFC specifications are going to letter ballot in April and should be complete this year. TheT11 committee is also investigating a multi-lane 128 GFC interface that is based on the 32GFC work. Work has not yet begun in T11 for developing the 64 GFC specifications, but 64 GFC is on the FCIA Speed roadmap."

  Steve Wilson, Director of Technology and Standards, Brocade and INCITS Technical Committee T11 Chairman

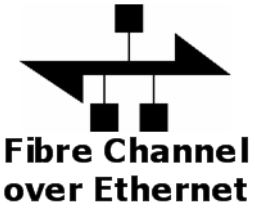
- 32 GFC will use 28 Gbps connectors (25/28G), and will double the speed of 16 GFC

Converged Networks



- Combined LAN and SAN networks
  - Lossless features of Fibre Channel with ubiquity of Ethernet
  - Within a rack (short-term)
  - Entire infrastructure (long-term)
- DCB – Data Center Bridging
  - Enhanced Ethernet to support FC storage traffic and more
- FCoE – Fibre Channel over Ethernet
  - First major application for DCB
- CNA – Converged Network Adapter
  - Supports 10 Gb Ethernet and 10 Gb FCoE at the same time on the same cable

Converged Networks



► Organizational Issues

- In typical large shops today, networking and storage are separate departments
  - Networking – Dynamic (more changes)
  - Storage – Stable (fewer changes)
- Other areas of convergence
  - Consider voicemail & email
- Those that learn networking and storage will be in the best position
  - *It doesn't hurt to be multi-lingual*

Network (I/O) Virtualization

- Virtualizing the ***I/O path*** between a server and an external device
- De-couple the logical from the physical
  - Hardware can be split into smaller logical units
  - Hardware can be represented as multiple units
  - Hardware can be combined into larger units
- Could apply to anything that performs I/O or works with an I/O adapter in a server, such as:
  - Ethernet Network Interface Cards (NICs) and switches
  - Disk Controllers (including RAID controllers)
  - Fibre Channel Host Bus Adapters (HBAs) and switches
  - SSDs mounted on internal cards

Network Virtualization

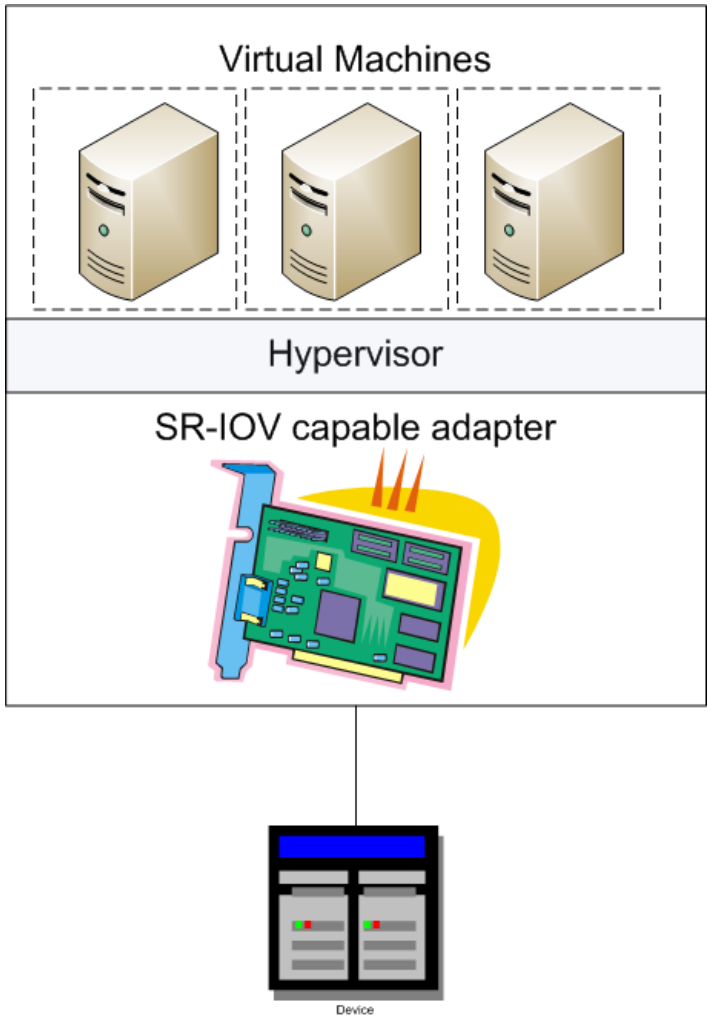
► Existing Forms

- NIC Teaming
  - A virtual NIC composed of two or more physical NICs
- Virtual LAN
  - Multiple, smaller logical LANs within a physical LAN infrastructure
- Fibre Channel NPIV
  - Multiple logical N_Port IDs sharing one physical N_Port
- Virtual SAN Fabrics
  - Multiple, smaller logical SANs within a physical SAN infrastructure

Network Virtualization

► Single-root I/O Virtualization (SR-IOV)

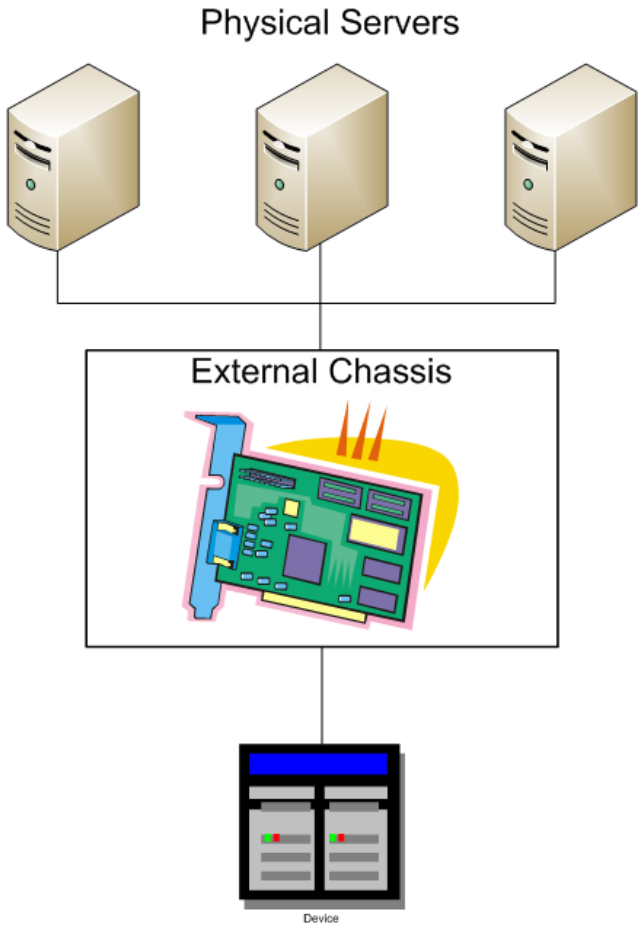
- Multiple VMs sharing one I/O adapter
- Bandwidth of the I/O adapter is shared among the VMs
- Virtual adapters created and managed by SR-IOV adapter (not hypervisor)
- Improved performance for VMs and their apps (near-native) by offloading I/O management and mapping functions to the adapter



Network Virtualization

► Multi-root I/O Virtualization (MR-IOV)

- Multiple servers & VMs sharing one I/O adapter
- Bandwidth of the I/O adapter is shared among the servers
- The I/O adapter is placed into a separate chassis
- Bus extender cards are placed into the servers



Network Virtualization

► SR-IOV Virtual Functions (VF)

- Virtual functions are the way that the adapter makes multiple versions of itself visible to the Hypervisor
- The Hypervisor assigns a VF to a guest
  - Many VFs can be created per physical port
  - Hypervisor has no visibility into the VF
- Guest sees a new adapter that it can use for anything that adapter can do

Network Virtualization

► SR-IOV Availability Today

- Available today for Ethernet only
  - Recommended for high-traffic VMs
- SR-IOV Support in Operating System & Hypervisor:
  - Citrix XenServer
  - RHEL 6 KVM
  - VMware vSphere 5.1
    - vMotion, Storage vMotion not supported with SR-IOV
  - Windows Server 2012 (Hyper-V and guest)
    - Live Migration is supported with SR-IOV

Network Virtualization

► SR-IOV Dependencies

- Hardware
  - Many, but not all PCIe 2.0 and 3.0 servers meet these criteria
  - Processor (CPU) support (specific support for SR-IOV)
  - Motherboard support (chipset, etc.)
  - BIOS support
  - SR-IOV capable NIC
- Software
  - Hypervisor IOV-enabled virtual switch
  - VF driver for the guest O.S.
  - Windows registry: IovEnableOverride in parent partition

Network Virtualization

► NPIV and Virtual Fibre Channel

- Available with Windows Server 2012 Hyper-V
  - Similar in concept to SR-IOV but for Fibre Channel
  - Supported by most FC HBAs (requires NPIV support)
    - NPIV is enabled by default in some FC HBAs, not in others
  - Requires support by the guest O.S. (Windows 2008R2/2012)
- Procedure
  - Assign a pair of virtual WWPNs to a guest O.S.
  - Add the WWPNs to the zoning and storage LUN masking
- Benefits
  - VMs have their own virtual FC HBAs, just like physical servers
  - VMs can be moved and take their FC storage with them

Encoding Schemes

- 8b/10b
  - For every 8 bits, adds 2 bits for command and control
  - 20% overhead = (10-8)/10
- 64b/66b
  - Used by 10 GigE and 16 GFC
  - For every 64 bits, adds 2 bits for command and control
  - 3% overhead = (66-64)/66
- 128b/130b
  - Used by PCIe 3.0
  - For every 128 bits, adds 2 bits for command and control
  - 1.5% overhead = (130-128)/130

PCI-Express

- Bus used in modern computers for I/O adapters
- Measured in gigatransfers/second (GT/s)
  - Bandwidth specified by indicating number of lanes such as “x1”, “x2”, etc., and generally spoken as “by 1”, “by 2”, etc.
- **PCIe 4.0** – In November 2011, the PCI-SIG announced the approval of 16 gigatransfers per second as the bit rate for the next generation of PCIe architecture, known as PCIe 4.0. Final specifications are expected in 2014-2015.

Demartek	GT/s	Encoding	x1	x2	x4	x8	x16
PCIe 1.x	2.5	8b/10b	250 MB/s	500 MB/s	1 GB/s	2 GB/s	4 GB/s
PCIe 2.x	5	8b/10b	500 MB/s	1 GB/s	2 GB/s	4 GB/s	8 GB/s
PCIe 3.x	8	128b/130b	1 GB/s	2 GB/s	4 GB/s	8 GB/s	16 GB/s

Fibre Channel Adapter Specifications

- Fibre Channel can run in full-duplex mode, but storage protocols generally operate in half-duplex mode
  - Throughput numbers below are half-duplex (one-way)
- Host Adapter Requirements below are for dual-port cards

Demartek	Throughput (MBps)	Encoding	Line Rate (Gbaud)	Host Adapter Requirements
1GFC	100	8b/10b	1.0625	PCI-X
2GFC	200	8b/10b	2.125	PCI-X
4GFC	400	8b/10b	4.25	PCI-X 2.0 or PCIe 1.0 (x4)
8GFC	800	8b/10b	8.5	PCIe 1.0 (x8) or PCIe 2.0 (x4)
16GFC	1600	64b/66b	14.025	PCIe 2.0 (x8) or PCIe 3.0 (x4)

Cabling Recommendations

► Fiber Optic Cables

- Fiber optic cabling service life – 15 to 20 years
- ***Recommendation*** – OM4 cables for current and future
  - OM4 will support 40/100 GigE and higher speeds of FC

Demartek	OM1	OM2	OM3	OM4
Jacket color	Orange	Orange	Aqua	Aqua
1 Gb/s	300m	500m	860m	–
2 Gb/s	150m	300m	500m	–
4 Gb/s	70m	150m	380m	400m
8 Gb/s	21m	50m	150m	190m
10 Gb/s	33m	82m	Up to 300m	Up to 400m
16 Gb/s	15m	35m	100m	125m

Cabling Recommendations

► Copper Cables

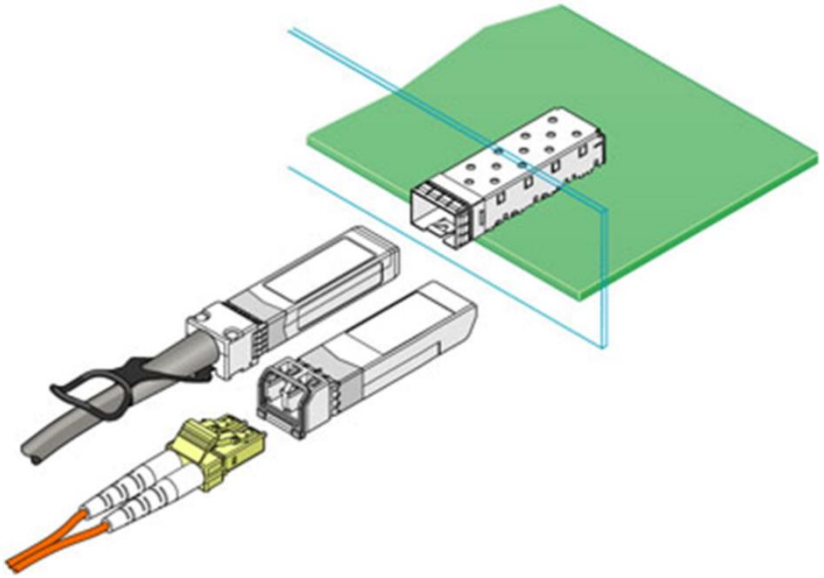
- 10 GigE – SFP+ Copper
  - SFP+ copper cables are known as Direct Attach Copper (DAC)
  - SFP+ “transceiver” is directly attached to the cable
  - Common lengths of 10 GigE DAC are 3 and 5 meters
- 10 GigE – RJ45 / 10GBASE-T
  - Cables must be certified to at least 500MHz to ensure 10GBASE-T compliance
  - ***Recommendation*** – Cat6a & Cat7 up to 100 meters
  - Cat6 can be used up to 55 meters, but should be tested first
  - Cat5e is not recommended for 10 GigE

Cabling Future Technology Outlook

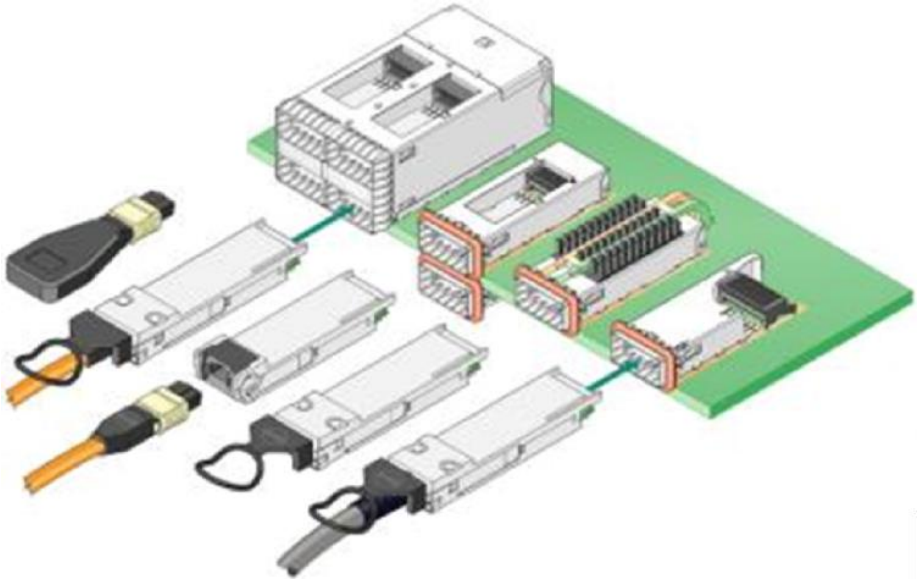
- As interface speeds increase, expect increased usage of fiber-optic cables and connectors for most interfaces
  - At higher Gigabit speeds, passive copper cables and interconnects experience “amplitude loss” and become too “noisy” except for short distances (within a rack or to adjacent racks)
  - Expect to see “active copper” for some higher-speed connection types

Connectors

Single-lane – SFP, SFP+



Four-lane – QSFP, QSFP+



Demartek	SFP	SFP+	QSFP+
Ethernet	1GbE	10GbE	40GbE
Fibre Channel	1GFC, 2GFC, 4GFC	8GFC, 16GFC	–
Infiniband	–	–	QDR, FDR

Demartek Free Resources

- Demartek FC Zone
  - www.demartek.com/FC
- Demartek FCoE Zone
  - www.demartek.com/FCoE
- Demartek iSCSI Zone
  - www.demartek.com/iSCSI
- Demartek SSD Zone
  - www.demartek.com/SSD
- Demartek SSD Deployment Guides
  - iSCSI and SSD available now, 16GFC coming…

Performance reports, Deployment Guides and commentary available for free download.

- www.demartek.com/Demartek_Interface_Comparison.html
- Or search for "storage interface comparison" in your favorite search engine

Free Monthly Newsletter

Demartek publishes a free monthly newsletter, ***Demartek Lab Notes***, highlighting recent reports, articles and commentary.

Look for the newsletter sign-up at:
www.demartek.com/Newsletter

Thank You!

Dennis Martin, President
dennis@demartek.com
www.linkedin.com/in/dennismartin

Demartek

(303) 940-7575
www.demartek.com
http://twitter.com/Demartek
www.youtube.com/Demartek
Skype: Demartek

To learn more about Demartek:

- Download the Aurasma App (Android/iPhone)
- Search and follow "Demartek"
- View image below with viewfinder.

*also on the back of Dennis' business card

Powered by: AURASMA