Next Generation Storage Networking for Next Generation Data Centers

Dennis Martin
President, Demartek

Tuesday, September 16, 2014

Agenda

- ☐ **About Demartek**

- ☐ **Increased Bandwidth Needs for Storage**

- ☐ **Storage Interface Technology & Futures**
  - ☐ **Ethernet, Fibre Channel, SAS, Thunderbolt, USB**

- ☐ **Cabling – Fiber Optic and Copper**

- ☐ **Performance Results**

- ☐ **Demartek Free Resources**

2

SNIA SDC Session Evaluations

- Be sure to complete the session evaluations that will be emailed to you.

- The SNIA staff reads these and I also get a copy of the results for this session.

- SNIA is offering *a $100 American Express gift card* to a randomly selected individual who completes the evaluations.

Demartek Video

Click to view this one minute video
(available in 720p and 1080p)

Demartek YouTube Channel:
http://www.youtube.com/user/Demartek/videos

About Demartek

- Industry Analysis and ISO 17025 accredited test lab
- Lab includes servers, networking & storage
  - Ethernet: 1, 10 & 40 Gbps: NFS, SMB (CIFS), iSCSI, FCoE and SR-IOV
  - Fibre Channel: 4, 8 & 16 Gbps
  - Servers: 8+ cores, large RAM
  - Virtualization: VMware, Hyper-V, Xen, KVM
- We prefer to run real-world applications to test servers and storage solutions
- 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**
- **Bootstorm test with 90 VMs in one physical server**
  http://www.demartek.com/Demartek_Analysis_of_VDI_Storage_Performance_during_Bootstorm.html

Physical Server

Virtual Machines
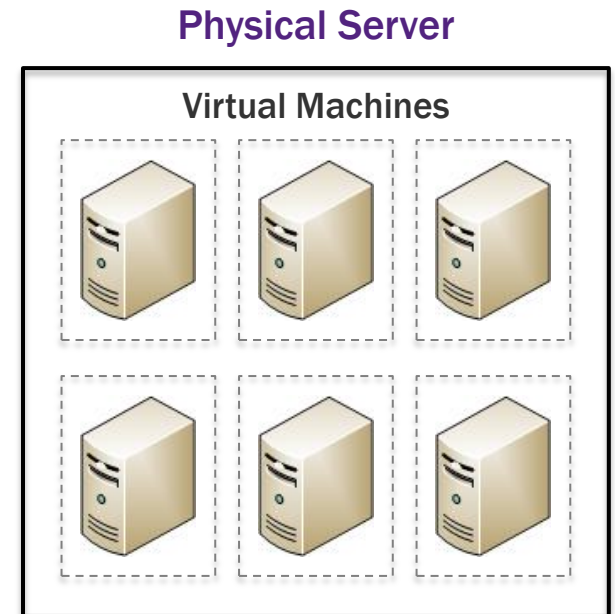
The Need for More Bandwidth
► New Hardware

- ❑ **New Generations of Servers**
  - ❑ **PCI Express 3.0 since 2012**
    - ❑ Up to 40 PCIe lanes per processor
  - ❑ **New servers support 10GbE on the motherboard**
  - ❑ **More cores per processor**
  - ❑ **Larger memory support (up to 1.5TB/processor)**
- ❑ **SSD**
  - ❑ **Are you deploying enterprise SSDs today?**

7

SD C 14

PCI-Express

- 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.

◇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

* This table available at http://www.demartek.com/Demartek_Interface_Comparison.html

- **PCIe 4.0 – In November 2011, the PCI-SIG announced the approval of 16 GT/s as the bit rate for PCIe 4.0.**
  - PCIe 4.0 specification Rev 0.5 targeted for 2H 2014*
  - PCIe 4.0 specification Rev 0.9 targeted for 1H 2016*
    - * Source: PCI-SIG

8

SDC 14

Ethernet
## ► 1GigE and 10GigE

Ethernet
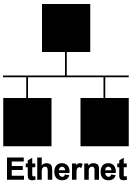
- ☐ **1GigE**
  - ☐ Not unusual to have 4, 6 or 8 NIC ports in a server
    - ☐ Consider the number of cables and PCIe slots used
  - ☐ Can be quad-port, dual-port or single-port

- ☐ **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 3.0 x4 or PCIe 2.0 x8
    - ☐ Single-port 10GigE NIC – PCIe 2.0 x4 or PCIe 1.0 x8
  - ☐ Adoption: blade servers yes, rack servers not so much

SDC 14

Ethernet
► 40GigE and 100GigE
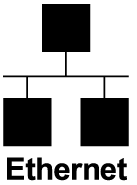
Ethernet

- **IEEE 802.3ba (40GigE & 100GigE) ratified June 2010**
- **The fastest Ethernet cables and connectors today are <u>10 Gbps</u> 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 connectors will soon be available**
  - 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 possibly available in 2014 or 2015
- **40 Gbps NICs require PCIe 3.0 x8 or x16 slot in the server**

SD C 14

Ethernet
► 25GigE
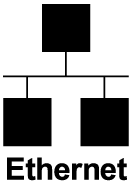
Ethernet

- ❑ **25Gb PHYs are beginning to appear**

- ❑ **Why not 25GbE over single-lane connection?**

- ❑ **25G Ethernet Consortium Announcement – July 1, 2014**
  - ❑ Arista Networks, Broadcom, Google, Mellanox and Microsoft
  - ❑ 25GbE and 50GbE specifications
  - ❑ www.25GEthernet.org

- ❑ **IEEE has announced a 25GbE study group – July 2014**
  - ❑ First use case: server interconnects

11

Fibre Channel
► 16 Gigabit (16GFC)

- **16GFC is backward compatible with 4GFC & 8GFC**
- **Uses 14 Gbps single-lane connectors**
  - Doubles speed of 8GFC due to newer 64b/66b encoding
- **First 16GFC switches and HBAs shipped in 2011**
  - Some of these HBAs can function as 10 Gb NICs
- **FC speeds and server slot requirements (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
► 32 Gigabit and 128 Gigabit ("Generation 6")

- **In February 2014, "Gen 6" Fibre Channel was announced**
- **32 Gbps single-lane connection ("32GFC")**
  - OM4 fiber-optic expected cable distance: 100m
- **128 Gbps parallel connection (4 x 32, "128GFCp")**
  - Initially used for switch-to-switch connections
- **Forward Error Correction (FEC)**
- **Energy Efficiency**
  - Power at transceiver is reduced when not in use ("dimmer switch")
- **Backward Compatible with 16GFC and 8GFC**
- **Products expected to be available in 2016**

Converged Networks

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

SAS – Serial Attached SCSI

- ❑ **12Gb/s SAS also known as SAS3**
- ❑ **12Gb/s began shipping in 2H 2013**
  - ❑ SAS HBAs and RAID controllers
  - ❑ Drives – SSDs and some HDDs
  - ❑ Some external storage arrays
- ❑ **Volume production ramp-up expected in 2014**
- ❑ **For best results use servers that support PCIe 3.0**
  - ❑ PCIe 3.0 x8 for typical 12Gb/s SAS adapter
- ❑ **12Gb/s SAS uses mini-SAS HD connectors**

SAS
► Mini-SAS HD connectors
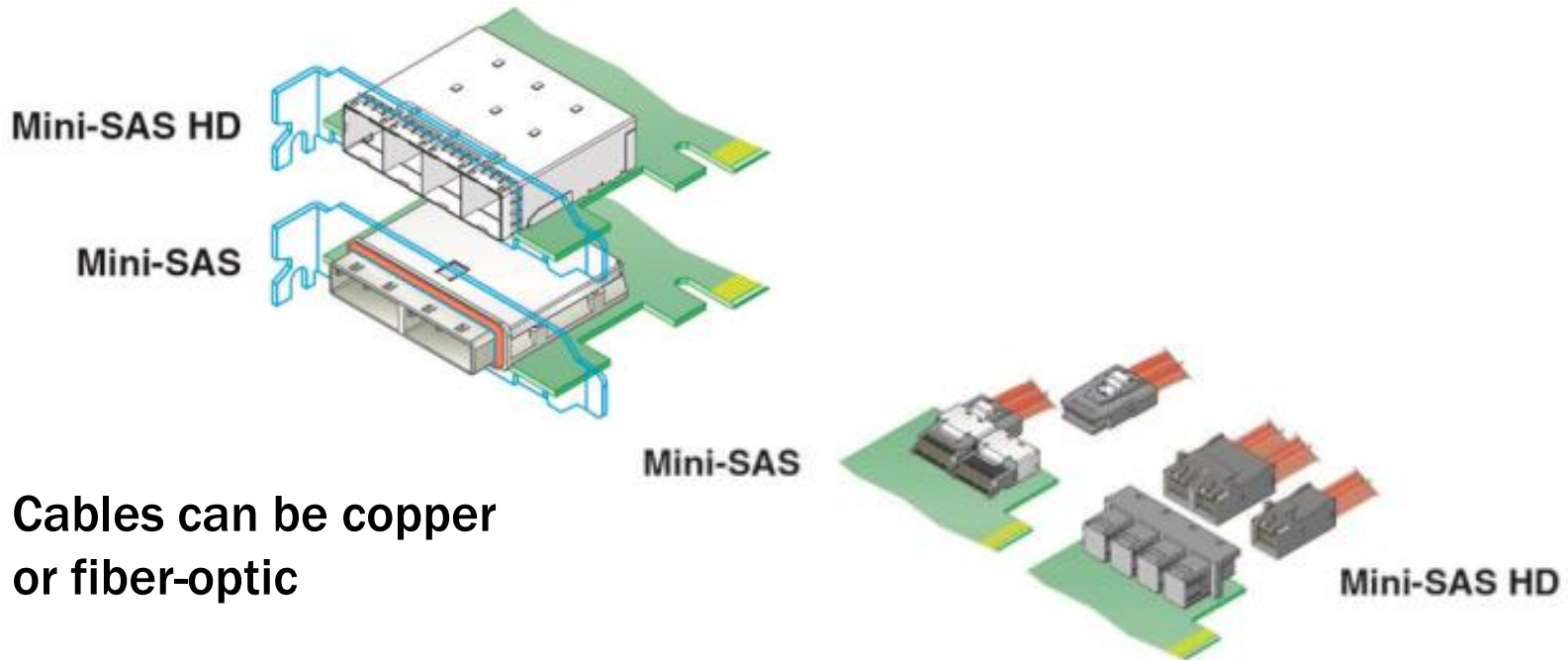
❑ Cables can be copper
   or fiber-optic

❑ See larger versions of these diagrams and information for other storage
   interfaces on the Demartek Storage Interface Comparison page:
   http://www.demartek.com/Demartek_Interface_Comparison.html

Thunderbolt™ 2

- **Doubles previous speed to 20 Gbps**
- **Target audience is media creators and editors who use premium laptops, desktops, workstations and peripherals that connect to them.**
  - Includes storage devices, especially SSDs
- **Currently limited to six (6) devices on one connection**
  - Devices can be daisy-chained
- **Available on motherboards now**
  - Add-in cards now available
- **Thunderbolt will support NVMe**
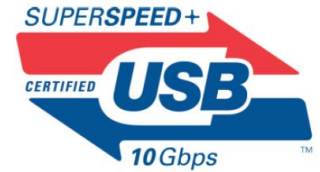- **Expect more activity during 2014 and 2015**

17

Thunderbolt™ 2

- **Thunderbolt 2 hubs are now available**
- **Thunderbolt 2 can be used to carry Ethernet at 10 Gbps**
  - Share files between PC and Mac
  - Thunderbolt 2 to 10GbE bridge devices connect to standard 10GbE switches

Sept. 10, 2014 at IDF 2014

USB 3.1

- **USB 3.1 specification completed July 2013**
  - Doubles speed to 10 Gbps (USB 3.0 is rated for 5 Gbps)
  - Works with existing USB 3.0 and 2.0 products
- **USB 3.1 Power Delivery**
  - Can deliver up to <u>100 watts</u>, bi-directionally
  - Can deliver audio/video, data and power concurrently
- **Media Agnostic USB protocol (USB over WiFi)**
  - Allows wireless devices and docking stations to communicate using the USB protocol
- **New USB Type-C bi-directional connector**
  - Similar in size to existing USB 2.0 micro-B
- **Products expected by end of year 2014**

19
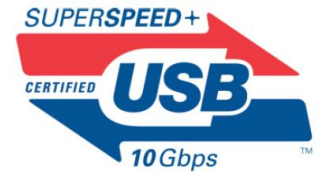
USB 3.1 Type-C Connector

- **Type-C Connector Specification completed in August 2014**
- **Entirely New Design with Smaller Size**
  - Tailored for emerging product designs
  - Robust enough for laptops and tablets; slim enough for mobile phones
- **Reversible plug orientation & cable direction**
- **All Type-C cables are electronically marked**
  - Can pass cable information to the device
- **Typical length: 1m**
  - Passive cables only today, but active cables could be built
  - No optical cables yet, but nothing prevents it

USB Type-C Cable & Connector

USB Type-C Cable and Connector Renderings

Provided by the USB 3.0 Promoter Group

TOP-MOUNT RECEPTACLE

MID-MOUNT RECEPTACLE

TYPE-C PLUG & CABLE

TOP-MOUNT RECEPTACLE
HYBRID

TOP-MOUNT RECEPTACLE
DUAL-ROW SMT

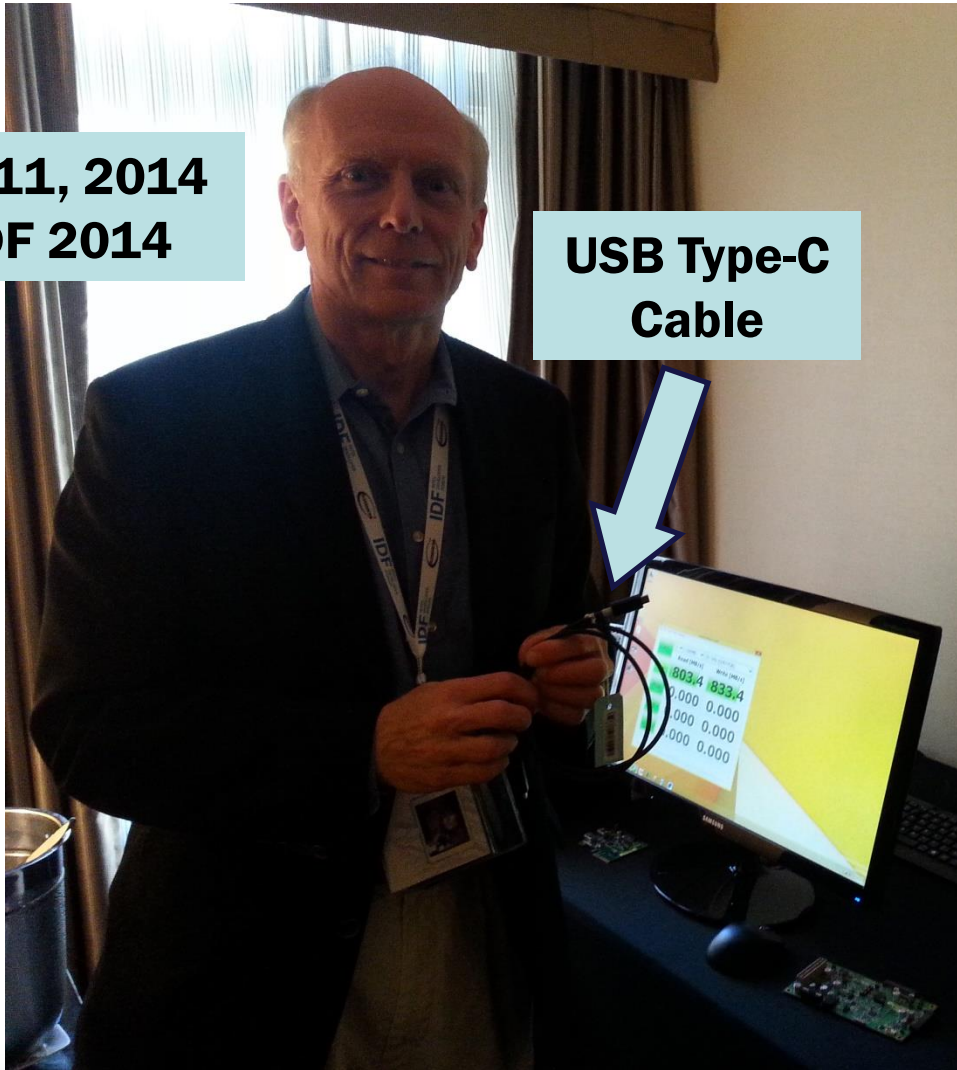MID-MOUNT RECEPTACLE
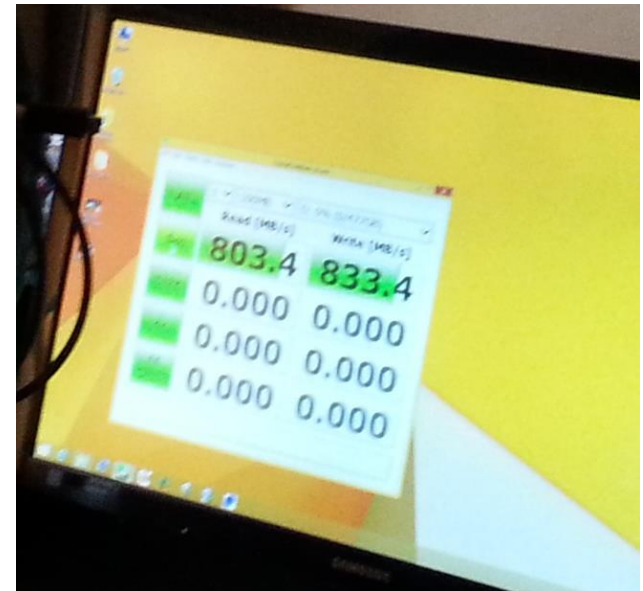HYBRID

MID-MOUNT RECEPTACLE
DUAL-ROW SMT

USB 3.1

Sept. 11, 2014 at IDF 2014

USB Type-C Cable

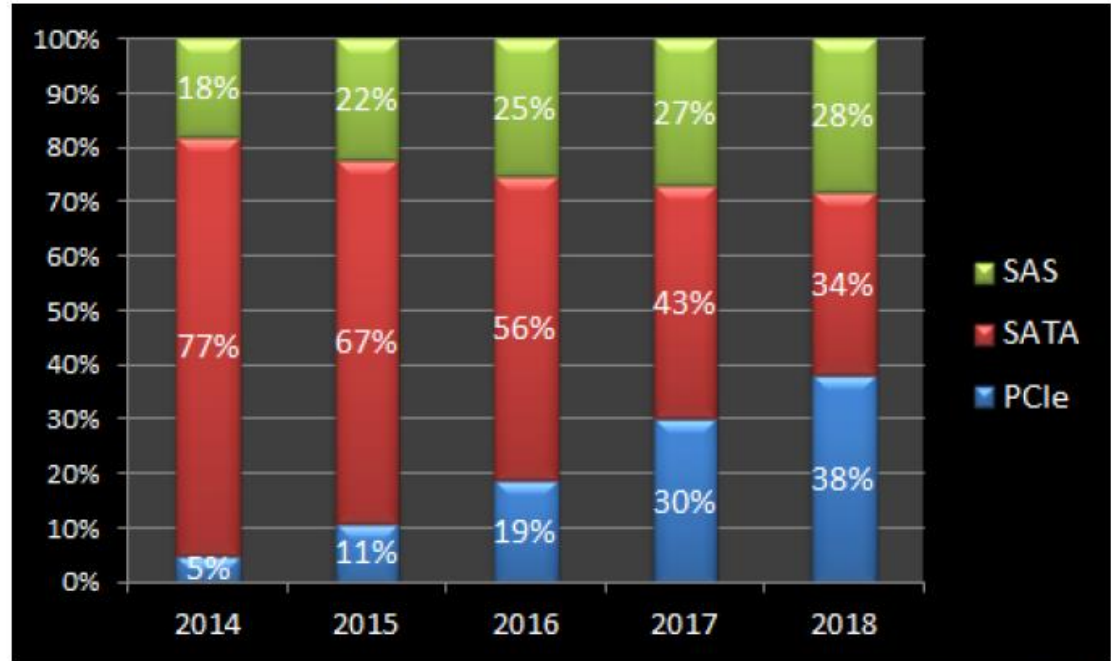**Single SSD running over USB 3.1 800+ MB/sec**

NVM Express (NVMe)

- Scalable host controller interface designed for enterprise and client systems that use PCI Express SSDs

- Designed with Flash memory and technologies coming after Flash memory in mind (non-volatile memory)

- Much faster (lower latency) software stack than existing storage stacks such as SAS and SATA

- In-box drivers for Windows and Linux now, others planned

- Expect SSD product announcements in 2014 and 2015
    - First shipping product available late March 2014

- Additional comments and explanation: http://www.demartek.com/Demartek_Comments_IDF2013_and_NVMe.html (2014 edition to be posted soon)

23

NVM Express (NVMe) Futures

- **PCI Express (PCIe) projected to be the leading enterprise SSD interface by 2018**

Enterprise SSD by Interface

	2014	2015	2016	2017	2018
SAS	18%	22%	25%	27%	28%
SATA	77%	67%	56%	43%	34%
PCIe	5%	11%	19%	30%	38%

Source: IDC

- **Expect NVMe to ship broadly in client SSD market in 2015.**

- **NVMe over Fabrics development underway. Goal is to run NVMe over network of choice within ~10 µs latency of local.**
  - NVMe works well with RDMA

24

Cabling Recommendations
► **Fiber Optic Cables (data center)**

☐ Fiber optic cabling service life: 15 – 20 years

☐ Recommendation: OM4 cables for current & future

  ☐ OM4 will support 40/100 GigE and 32GFC

◈ **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

* This table available at http://www.demartek.com/Demartek_Interface_Comparison.html

Cabling Recommendations
► Copper Cables (data center)

- ## 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 Recommendations
► 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

    ❑ Active copper can go longer distances than passive copper
    ❑ Active copper is thinner allows for better airflow than passive copper
    ❑ Active copper uses more power than passive copper
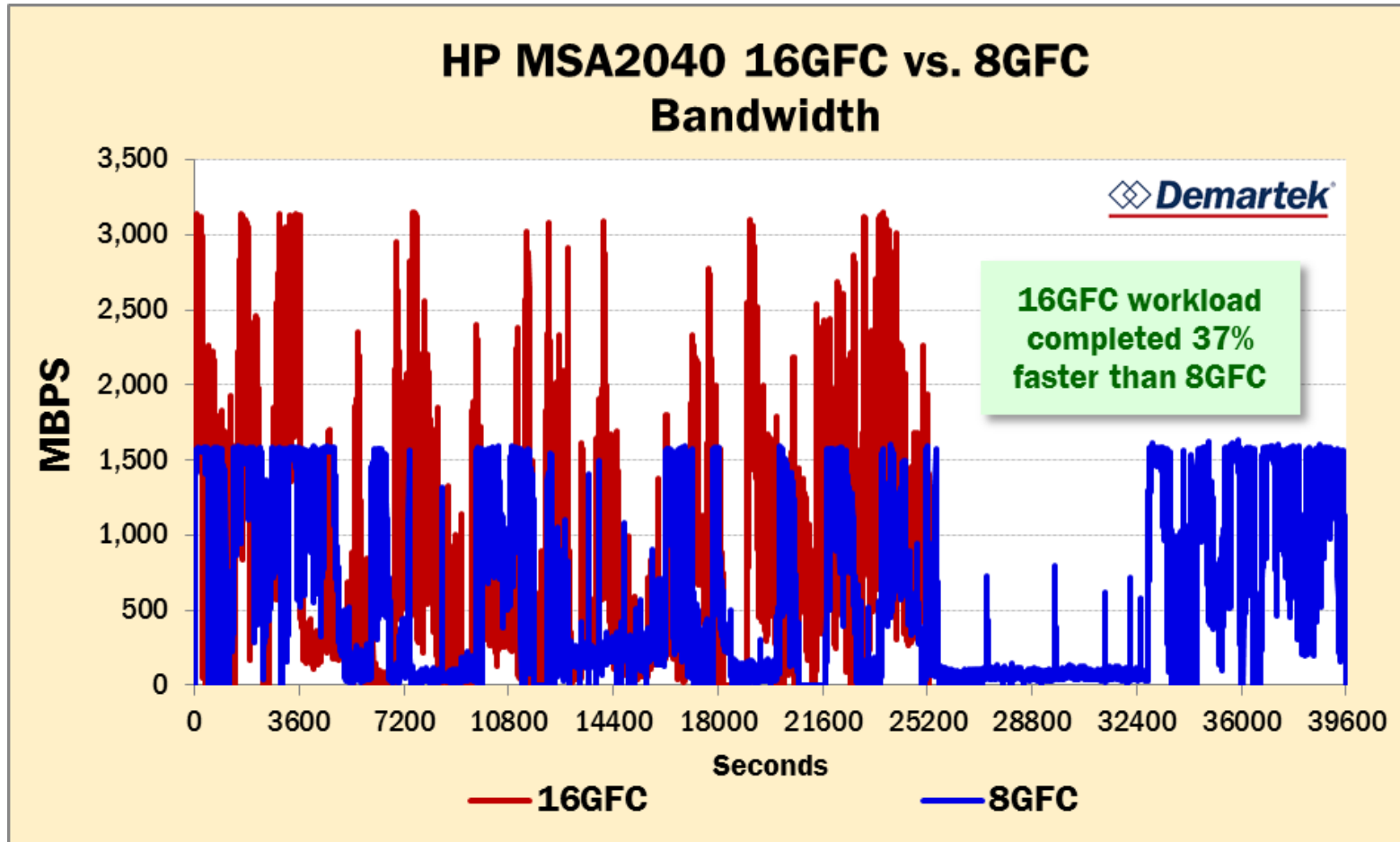
Connectors

Single-lane – SFP, SFP+
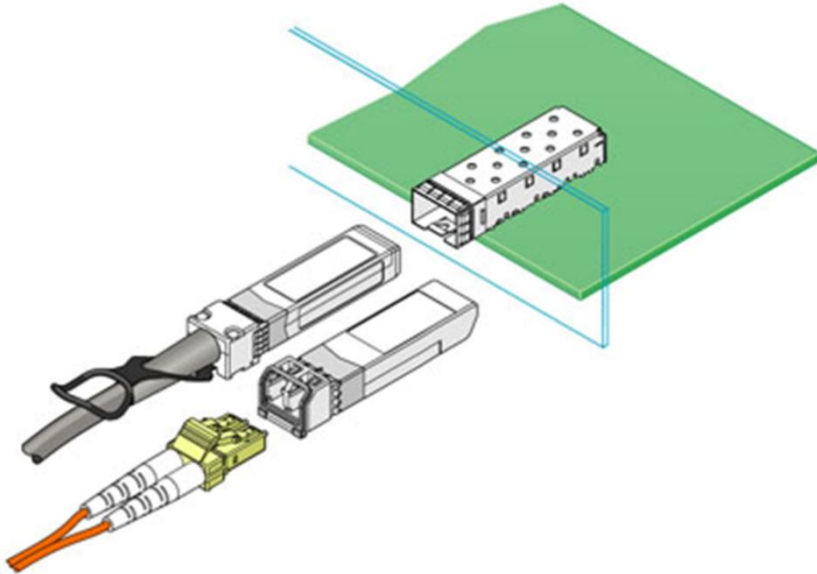
Four-lane – QSFP, QSFP+
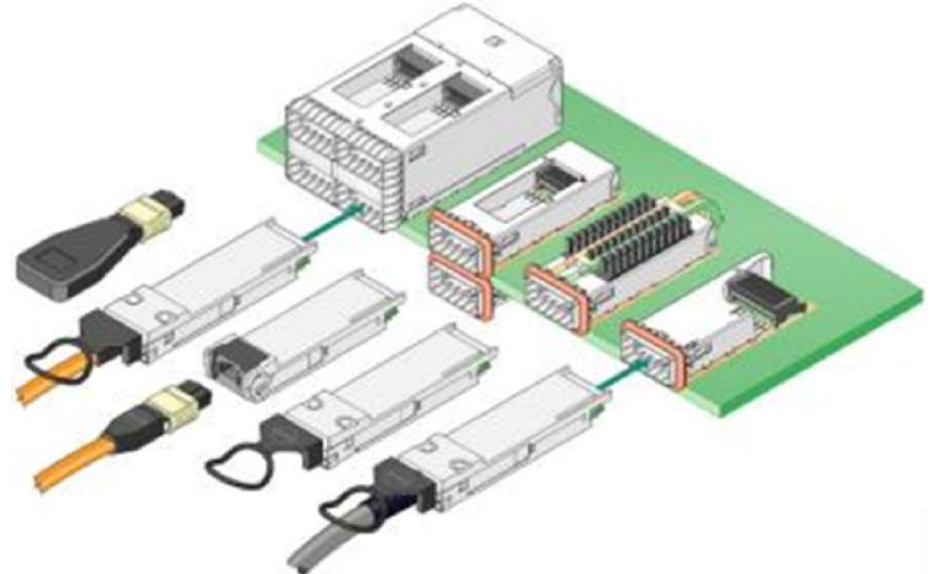
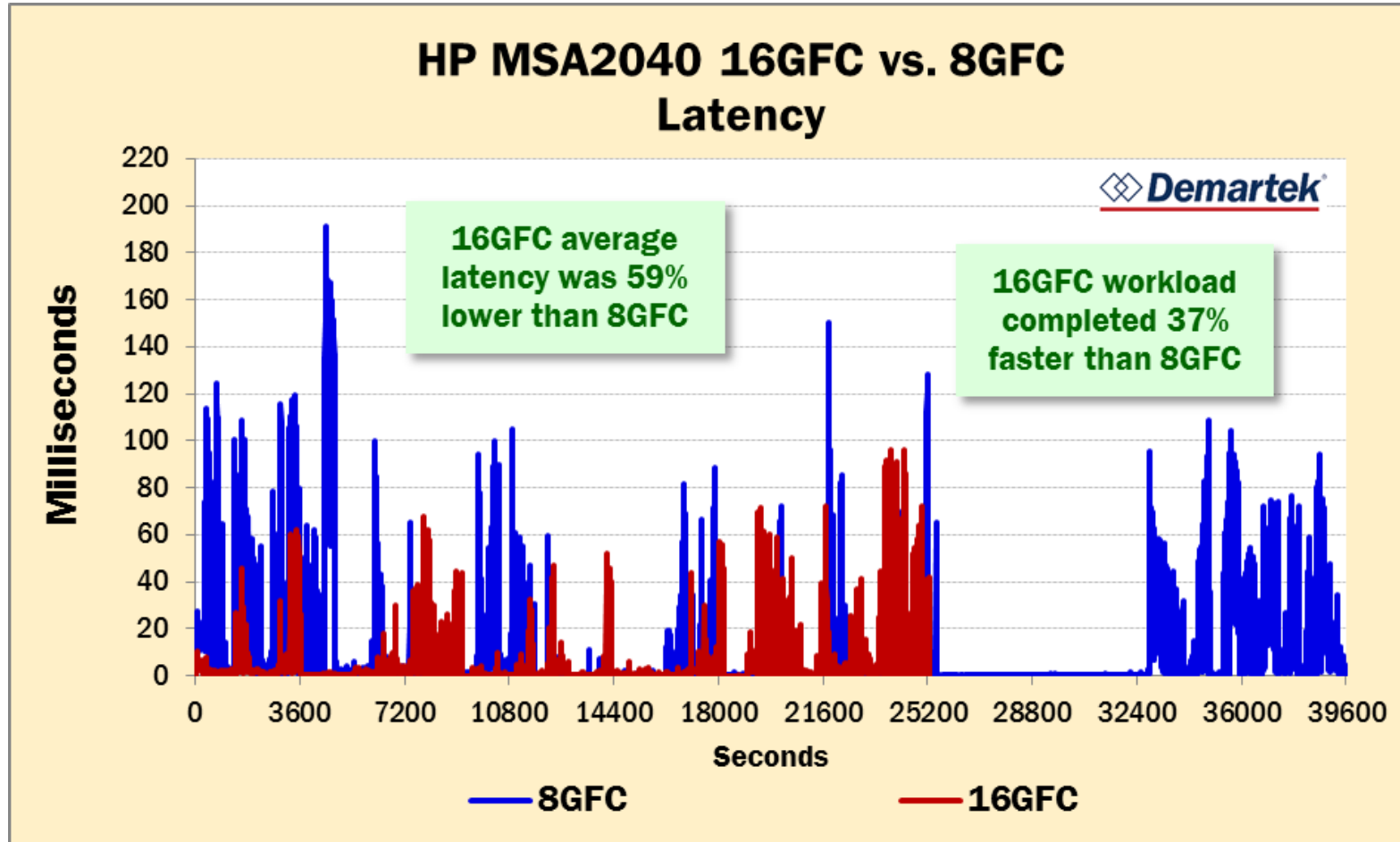|	SFP	SFP+	QSFP+
Ethernet	1GbE	10GbE	40GbE
Fibre Channel	1GFC, 2GFC, 4GFC	8GFC, 16GFC	–
Infiniband	–	–	QDR, FDR

Performance Example: 16GFC vs. 8GFC
► Bandwidth– SQL Server data warehousing workload

Performance Example: 16GFC vs. 8GFC
► Latency – SQL Server data warehousing workload

HP MSA2040 16GFC vs. 8GFC Latency

16GFC average latency was 59% lower than 8GFC
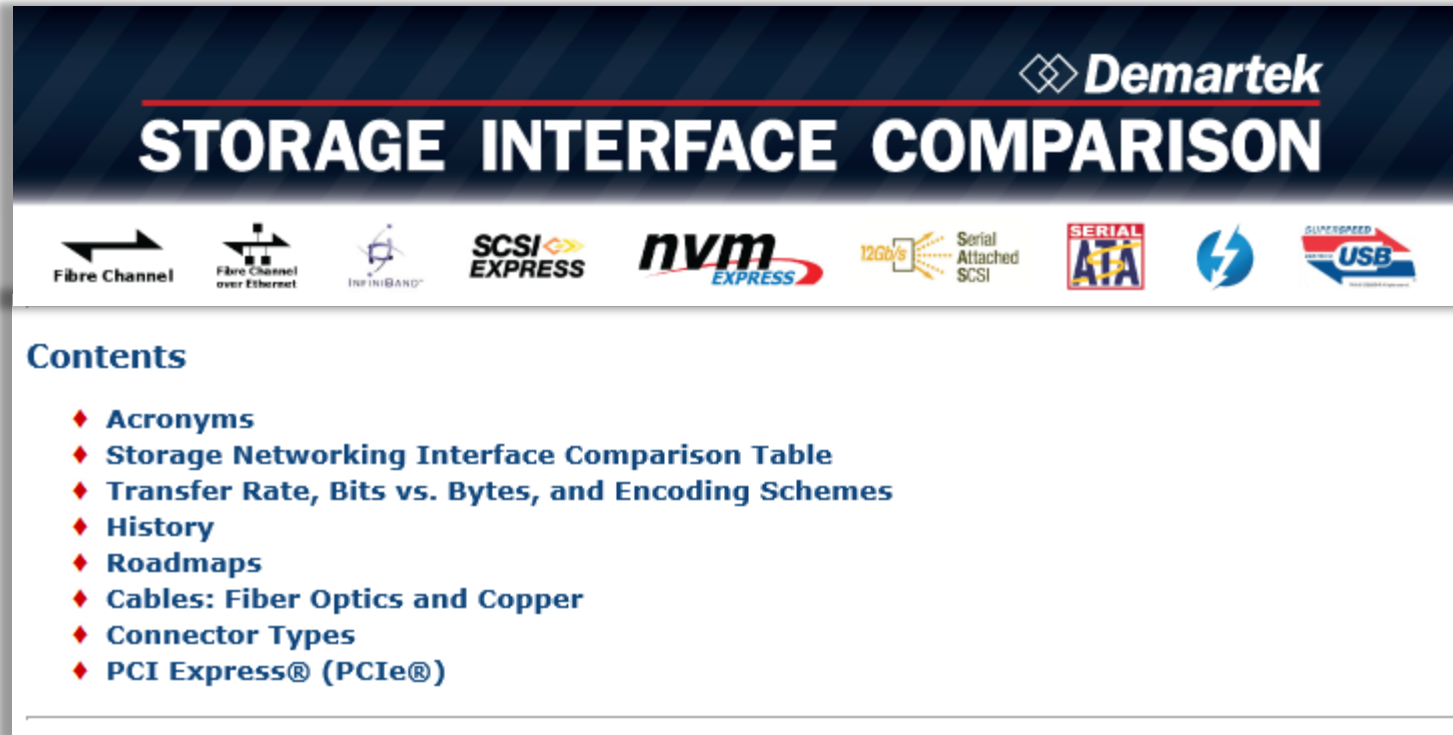
16GFC workload completed 37% faster than 8GFC

Demartek Free Resources

- Demartek comments on Flash Memory Summit 2014
  www.demartek.com/Demartek_Flash_Memory_Summit_2014_Commentary.html

- Demartek comments on CES 2014 (USB & Thunderbolt)
  www.demartek.com/Demartek_CES_2014.html

- Demartek comments on IDF2013 & NVMe (2014 version available soon)
  www.demartek.com/Demartek_Comments_IDF2013_and_NVMe.html

- Demartek SSD Deployment Guide
  www.demartek.com/Demartek_SSD_Deployment_Guide.html

- Demartek Video Library -
  http://www.demartek.com/Demartek_Video_Library.html

- Demartek FC Zone – www.demartek.com/FC

- Demartek iSCSI Zone – www.demartek.com/iSCSI

- Demartek SSD Zone – www.demartek.com/SSD

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

31

Storage Interface Comparison

- Downloadable interactive PDF version now available

- Search engine: "storage interface comparison"

- www.demartek.com/Demartek_Interface_Comparison.html

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

SNIA SDC Session Evaluations

- Be sure to complete the session evaluations that will be emailed to you.

- The SNIA staff reads these and I also get a copy of the results for this session.

- SNIA is offering *a $100 American Express gift card* to a randomly selected individual who completes the evaluations.

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.**

REAL-WORLD HANDS-ON RESEARCH & ANALYSIS

*also on the back of Dennis' business card

Powered by: AURASMA