Re-Introduction to Ethernet Networked Storage

August 4, 2016
SNIA Presenters

Rob Davis
VP of Technology
Mellanox

Fred Zhang
Product Marketing Manager
Intel
About SNIA

SNIA-At-A-Glance

160 unique member companies

3,500 active contributing members

50,000 IT end users & storage pros worldwide

Learn more: snia.org/technical
The material contained in this presentation is copyrighted by the SNIA unless otherwise noted.

Member companies and individual members may use this material in presentations and literature under the following conditions:

- Any slide or slides used must be reproduced in their entirety without modification.
- The SNIA must be acknowledged as the source of any material used in the body of any document containing material from these presentations.

This presentation is a project of the SNIA.

Neither the author nor the presenter is an attorney and nothing in this presentation is intended to be, or should be construed as legal advice or an opinion of counsel. If you need legal advice or a legal opinion please contact your attorney.

The information presented herein represents the author's personal opinion and current understanding of the relevant issues involved. The author, the presenter, and the SNIA do not assume any responsibility or liability for damages arising out of any reliance on or use of this information.

NO WARRANTIES, EXPRESS OR IMPLIED. USE AT YOUR OWN RISK.
The Evolution of Storage Devices

IBM Almaden Research Facility
First Disk Drive - RAMAC 350 - 1956

Almaden Lobby
First Disk Drive - RAMAC 350 - 1956

Almaden Lobby
First Disk Drive - RAMAC 350 - 1956

Almaden Lobby
Evolution of Disk Drive

5 MB
Latency ~25MS

60 years

10 TB
Latency ~4MS

~$81,000/MB

~$60/TB
First Solid-State Drive (SSD) - Bulk Core - 1976

- 8 memory boards, 256KB of RAM, total 2MB
- Access time 750μsec
- Cost $9700 in 1977, equivalent to $40,000 today
Evolution of SSDs

- RAM: 750 usec
- NAND: ~25 usec read, ~250 write
- 2MB to 4TB in 40 years
- ~$20,000/MB today
- ~$300/TB
At First SSDs Fill the Performance Gap
Next Step - SANs

- Sharing storage across multiple servers
  - Better utilization: capacity, rack space, power
  - Scalability, management, fault isolation
Next Step - SANs

- Sharing storage across multiple servers
  - Better utilization: capacity, rack space, power
  - Scalability, management, fault isolation

- Fibre Channel standard started in 1988, with ANSI standard 1.0 approval in 1994

- First Fibre-channel connected storage product SPARCstorageArray, 1Gb/s
Fibre Channel SANs
SANs and LANs
iSCSI SANs

IETF ratified the iSCSI standard 2003
iSCSI SANs
iSCSI SANs
FCoE SANs

FCoE part of T11 FC-BB-5 was forwarded to ANSI for publication in 2009
FCoE SANs
FCoE SANs
FCoE SANs
Why Ethernet Storage Revival Now?

Data Explosion
Why Ethernet Storage Revival Now?
Capacity Needs Driving Lower Cost Networked Storage

FCoE
iSCSI
FC

Block
Capacity Needs Driving Lower Cost Networked Storage

- SMB (CIFS)
- NFS
- FCoE
- iSCSI
- FC
- Object
- Ceph
Capacity Needs Driving Lower Cost Networked Storage

- SMB (CIFS)
- NFS
- iSCSI
- FC
- FCoE
- Ceph
- Object

Block

File
Capacity Needs Driving Lower Cost Networked Storage

- SMB (CIFS)
- NFS
- iSCSI
- FC
- FC-E
- iSCSI

File

Block

Ceph

Object
Capacity Needs Driving Lower Cost Networked Storage
Capacity Needs Driving Lower Cost Networked Storage

Ethernet

SMB (CIFS) File

NFS

iSCSI

FC

FCoE

Ceph Object
Three major storage technology transitions
- SSDs replacing hard disk drives
- Scale-out, software-defined, or “Server-SAN” storage
- Hyperconverged
Cloud Driving Scale Out Storage into Data Center
Cloud Driving Scale Out Storage into Data Center

- Clouds abandoned traditional SAN
  - There is no Fibre Channel in the cloud
  - Different networking and storage models
- Enterprises are migrating to a cloud model
- New storage applications use scale-out
Scale Out Storage Growth

![Graph showing annual revenue projections for Traditional Enterprise Storage, Hyperscale Server SAN & Enterprise Server SAN from 2012 to 2027.](http://wikibon.org/wiki/v/The_Rise_of_Server_SAN#The_Business_&_Technology_Drivers_behind_Server_SAN)

- **Today**
- **Scale-Out**
- **Scale Up**
- **Ethernet Channel**
Hyperconverged

- Adaptation of scale-out storage
- Collapse separate compute & storage into one layer
  - Integrated compute-storage nodes
- IO requirements
  - Low latency, high bandwidth, Ethernet
Hyperconverged

- Adaptation of scale-out storage
- Collapse separate compute & storage into one layer
  - Integrated compute-storage nodes
- IO requirements
  - Low latency, high bandwidth, Ethernet
Hyperconverged Market

Spend on Converged Infrastructure Categories ($B)

Overall CAGR: 51%

Source: IDC #244924
Source: EMC Market Intelligence. TAM excludes Internal Enterprise Storage, DIY Hardware and Services in excess of EMC share
Why Ethernet Storage Revival Now?

Source: Wikibon 2014, from numerous sources
Why Ethernet Storage Revival Now?
SSD are Getting Faster – Much Faster

Storage Media Technology

Access Time in Micro Seconds

- HDD
- SSD
- PM

10,000x improvement
And so is Ethernet – Much Faster
And Ethernet Storage is Also Innovating

- Pure Bandwidth – up to 100Gb/s
  - Flash over Block, File and Object
And Ethernet Storage is Also Innovating

- Pure Bandwidth – up to 100Gb/s
  - Flash over Block, File and Object

- RDMA
  - RoCE, iWARP, InfiniBand
  - iSER
  - SMB Direct, NFSoRDMA
  - Ceph over RDMA
And Ethernet Storage is Also Innovating

- Pure Bandwidth – up to 100Gb/s
  - Flash over Block, File and Object

- RDMA
  - RoCE, iWARP, InfiniBand
  - iSER
  - SMB Direct, NFSoRDMA
  - Ceph over RDMA

- Persistent Memory (PM)
  - NVMe over Fabrics (NVMf)
  - PMf (NVDIMM, 3D-Xpoint, etc.)
Faster Wire Speeds is only $\frac{1}{2}$ the solution
Faster Wire Speeds is only $\frac{1}{2}$ the solution
Faster Wire Speeds is only $\frac{1}{2}$ the solution
Faster Protocol: NVMe

- NVMe: Optimized for flash and next-gen NV-memory
  - Traditional SCSI interfaces designed for spinning disk
  - NVMe bypasses unneeded layers
Faster Protocol: NVMe

- NVMe: Optimized for flash and next-gen NV-memory
  - Traditional SCSI interfaces designed for spinning disk
  - NVMe bypasses unneeded layers

- NVMe Flash Outperforms SAS/SATA Flash
  - 2x-2.5x more bandwidth, 40-50% lower latency, Up to 3x more IOPS

![Random Read/Write Performance Chart](chart.png)

- 750 Series (PCIe) vs. 730 Series (SATA)

- NVMe driver
- SAS driver

- 2.8 microseconds
- 6.0 microseconds
The idea is to extend the efficiency of the local NVMe interface over a fabric
- Ethernet or IB
- NVMe commands and data structures are transferred end to end

Capitalizes on RDMA for performance

https://www.brighttalk.com/webcast/663/185909
Faster Protocol: RDMA

- Efficient Data Movement (RDMA)
  - Higher Bandwidth
  - Lower Latency
  - More CPU Power For Applications

Kernel Bypass and Protocol Offload

Adapter based transport
Faster Protocol: RDMA

https://www.youtube.com/watch?v=u8ZYhUjSUoI
Many factors are changing the direction of storage networking in Ethernet’s favor

- Accelerating capacity requirements driving lower cost solutions
- SSD performance driving higher performance solutions
- Scale-Out storage and Hyperconverged driving single network solutions

Ethernet is VERY well positioned to take advantage of the new direction

- Already dominates low cost File and Object solutions
- Performance roadmap lines up well with SSD performance
- Innovative new protocols that tightly couple to new storage directions
After This Webcast

- Please rate this Webcast and provide us with feedback
- This Webcast and a PDF of the slides will be posted to the SNIA Ethernet Storage Forum (ESF) website and available on-demand
  - [www.snia.org/forums/esf/knowledge/webcasts](http://www.snia.org/forums/esf/knowledge/webcasts)
- A full Q&A from this webcast, including answers to questions we couldn't get to today, will be posted to the SNIA-ESF blog
  - [sniaesfblog.org](http://sniaesfblog.org)
- Follow us on Twitter @SNIAESF
Thanks!