Networking Flash Technology
“Showdown”

Rob Davis, J Metz, Motti Beck, Brandon Hoff, Peter Onufryk
Why Network Flash Based Storage?

- There are advantages to shared storage
  - Better utilization:
    - capacity, rack space, power
  - Scalability
  - Manageability
  - Fault isolation

- Shared storage requires a Network
Agenda

- **PCIe** Networked Flash Storage
  - Peter Onufryk, Microsemi(Microchip), NVM Solutions Fellow
- **InfiniBand** Networked Flash Storage
  - Motti Beck, Mellanox, Sr. Dir. Enterprise Market Development
- **Fibre Channel** Networked Flash Storage
  - Brandon Hoff, Principle Software Architect, Emulex Connectivity Division, Broadcom
- **Ethernet** Networked Flash Storage
  - J Metz, Cisco, R&D Engineer, Advanced Storage, Office of the CTO, UCS Systems Group
- Panel Discussion
Faster Storage Needs a Faster Network

Flash SSDs move the Bottleneck from the Disk to the Network
Eliminate the Bottleneck

Networked Storage

Storage Media

Protocol and Network

Access Time (micro-sec)

Faster Wire Speeds

More Efficient Protocol

Flash Memory Summit 2018
Santa Clara, CA
PCle® Fabric

Peter Onufryk
Microsemi Corporation
PCIe Fabric

- Storage Functions
  - Dynamic partitioning (drive-to-host mapping)
  - NVMe shared I/O (shared storage)

- Host-to-Host Communications
  - RDMA
  - Ethernet emulation

- Manageability
  - NVMe controller-to-host mapping
  - PCIe path selection
  - NVMe management

- Fabric Resilience
  - Supports link failover
  - Supports fabric manager failover
The PCIe Latency Advantage

Latency data from Z. Guz et al., "NVMe-over-Fabrics Performance Characterization and the Path to Low-Overhead Flash Disaggregation" in SYSTOR '17
The PCIe Advantage

Other Flash Storage Networks

PCIe Fabric
### PCIe Fabric Characteristics

<table>
<thead>
<tr>
<th>Property</th>
<th>Ideal Characteristic</th>
<th>PCIe Fabric</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost</td>
<td>Free</td>
<td>Low</td>
<td>• PCIe built into virtually all hosts and NVMe drives</td>
</tr>
<tr>
<td>Complexity</td>
<td>Low</td>
<td>Medium</td>
<td>• Builds on existing NVMe ecosystem with no changes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• PCIe fabrics are an emerging technology</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Requires PCIe SR-IOV drives for low-latency shared storage</td>
</tr>
<tr>
<td>Performance</td>
<td>High</td>
<td>High</td>
<td>• High bandwidth</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• The absolute lowest latency</td>
</tr>
<tr>
<td>Power consumption</td>
<td>None</td>
<td>Low</td>
<td>• No protocol translation</td>
</tr>
<tr>
<td>Standards-based</td>
<td>Yes</td>
<td>Yes</td>
<td>• Works with standard hosts and standard NVMe SSDs</td>
</tr>
<tr>
<td>Scalability</td>
<td>Infinite</td>
<td>Limited</td>
<td>• PCIe hierarchy domain limited to 256 bus numbers</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• PCIe has limited reach (cables)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• PCIe fabrics have limited scalability (less than 256 SSDs and 128 hosts)</td>
</tr>
</tbody>
</table>
InfiniBand Networked Flash Storage

Superior Performance, Efficiency and Scalability

Motti Beck – Sr. Director Enterprise Market Development, Mellanox Technologies
The Need for Intelligent and Faster Interconnect

Faster Data Speeds and In-Network Computing Enable Higher Performance and Scale

Must Wait for the Data
Creates Performance Bottlenecks

Analyze Data as it Moves!
Higher Performance and Scale
InfiniBand Technical Overview

- **What is InfiniBand?**
  - InfiniBand is an open standard, interconnect protocol developed by the InfiniBand® Trade Association: [http://www.infinibandta.org/home](http://www.infinibandta.org/home)
  - First InfiniBand specification was released in 2000

- **What does the specification includes?**
  - The specification is very comprehensive
  - From physical to applications

- **InfiniBand SW is open and has been developed**
  - [http://www.openfabrics.org/index.html](http://www.openfabrics.org/index.html)
RDMA enabled Networking Powers Modern Storage Platforms

Higher Performance, Higher Efficiency and Higher Scalability
NVMe over Fibre Channel

Brandon Hoff
Principal Software Architect
Emulex Connectivity Division, Broadcom Inc.
• Fibre Channel storage shows strong growth in capacity
  – Fibre Channel Storage capacity shipped is larger than all other types of external storage combined
• The adoption of All Flash Arrays and NVMe storage will drive the need for faster networks
• iSCSI is the dominate technology block over Ethernet
• Logical transition
  • SCSI FCP transitions to NVMe/FC
  • iSCSI transitions to NVMe/TCP
Dual protocol SANs enable low risk NVMe adoption

- Get NVMe performance benefits while migrating incrementally “as-needed”
- Migrate application volumes 1 by 1 with easy rollback options
- Interesting dual-protocol use cases
- Full fabric awareness, visibility and manageability with existing Brocade Fabric Vision technology
NVMe over Fibre Channel Performance on a A700s single node
Ethernet-Networked Flash Storage

J Metz, Ph.D
R&D Engineer, Advanced Storage
Cisco Systems
@drjmetz
Storage Perspective

• There is a “sweet spot” for storage
  • Depends on the workload and application type
  • No “one-size fits all”

• What is the problem to be solved?
  • Deterministic or non-deterministic?
  • Highly scalable or highly performant?
  • Level of manageability?

• Understanding “where” the solution fits is critical to understanding “how” to put it together
Ethernet Roadmap

- **How to go faster**
  - Different modulation techniques
  - Different data rate/lanes chosen
- **New Signaling methods**
  - Pulse Amplitude Modulation 4 vs. Non Return to Zero (NRZ)
- **New Form Factors**
  - Multi-lane interfaces

Source: ethernetalliance.org
Summary

• Ethernet
  • General Purpose network designed to solve many, many problems and do it well
  • Flexible for all but the most extreme conditions
  • Largest ecosystem of developers, vendors, and users
  • From the smallest system to the largest, there is no other networking technology more suited, or best understood