Concurrent NVMe and SCSI over Fibre Channel

Curt Beckmann
Brocade
Presentation Topics

- Background on NVMe
- How Concurrent SCSI and NVMe Lowers Risk
- Use Cases for NVMe over Fibre Channel
- The status of the FC-NVMe Protocol
- Recap and Q & A
Presentation Topics

- Background on NVMe
- How Concurrent SCSI and NVMe Lowers Risk
- Use Cases for NVMe over Fibre Channel
- The status of the FC-NVMe Protocol
- Recap and Q & A
The NVMe Ecosystem

- NVMe: NonVolatile Memory Express
  - An alternative to SCSI
  - Overseen by NVMExpress.org
  - Initially for PCIe (hence “Express”)
- NVMe 1.x is now 4+ years old
  - widely supported for PCIe SSDs
- NVMe over Fabrics 1.0 came in June 2016
NVMe server-side stack efficiency

Most focus is on lower latency, but parallelism is also important

- To reap all the benefit of low latency SSDs, we need a fast server-side stack with lower CPU utilization
- NVMe was designed from scratch with no baggage
  - Fewer commands, no tape support, better queuing

**Diagram Description**

- **Linux NVMe Stack**
  - User App
  - VFS
  - OS scheduling & CTX switch
  - Block Driver
  - Device Driver

- **Test system measured IOPS**
  - 3.5
  - ~10K cycles
  - ~3 usecs
  - 1.02M

- **Cores consumed for 1M IOPs**
  - 10
  - ~27K cycles
  - ~10 usecs

**Measurement Details**

- Measurements taken on Intel Core i5-2500K 3.3GHz 6MB L3 cache quad-core desktop processor
- Using RHEL 6.0 2.6.32-51.71-kernel
- Testing and measurement by Intel
NVMe includes Enhanced Queuing

- Supports more and deeper queues
  - Up to 64K queues
  - 64K commands per queue
- Queuing is integrated into the protocol rather than layered on top (as with SCSI)
  - Reduces server-side context switching, eliminating vast wasted clock cycles
Three NVMe phases

1. Basic (PCIe-based)
   NVMe in servers
   (last few years)

2. Basic NVMe
   in storage backend
   (production!)

3. NVMe over Fabrics
   (demoing now)

NVMe SSD modules can plug directly into servers
Or they can plug into PCIe expansion slots
Enterprise NVMe Requires a Network

Today’s focus is on Fibre Channel as the NVMe fabric

- Fibre Channel connects 80+% of today’s All Flash Arrays

- For reference, short comments on other fabrics
  - RoCEv2, iWARP and IB are RDMA-based
    - but not compatible with each other
  - NVMe over FCoE leverages FC-NVMe
    - Requires a DCB (“lossless”) network for Ethernet hop(s)
  - NVMe over TCP (“iNVMe”) is *not* RDMA-based
    - Recently proposed by Intel, Facebook, etc.
    - iSCSI-like, runs on existing commodity
Presentation Topics

- Background on NVMe
- How Concurrent SCSI and NVMe Lowers Risk
- Use Cases for NVMe over Fibre Channel
- The status of the FC-NVMe Protocol
- Recap and Q & A
Reducing risk accelerates adoption

- Enterprise storage teams are risk averse
  - Enterprise storage must meet SLA
  - Enterprise storage assets are “high value”
- Downtime or data loss can mean big OpEx
  - Teams plan ahead to minimize risk
- Perception of risk will cause reluctance
  - Flip side: reducing risk makes investment attractive
Fibre Channel is the premium fabric

- Fibre Channel was born in the Ethernet heyday
  - ATM, Token Ring faded as FC was adopted
  - FC focused on DC-scale, lossless, storage centric

- A robust “test-and-support” ecosystem culture
  - Enterprise array vendors test / sell / support SAN
  - Value is about more than just product features
Most Enterprises Already Use Flash

- Most Storage Environments are “brownfield”
  - Changing, duplicating infrastructure is costly, painful
- Storage teams already rely on Fibre Channel for
  - Performance and visibility
  - Tried-and-true fabric services
    - Name services, Discovery, Zoning, Flow Control
- Fibre Channel carries SCSI and NVMe at once
  - This reduces many kinds of risk
Protect Investments and Data: No Rip-and-Replace

Using the same infrastructure **dramatically** lowers NVMe adoption risk

- FC-NVMe & FC-SCSI coexist...
  - In the same server
  - On the same HBA
  - On the same driver
  - On the same SAN

- FC-NVMe & FC-SCSI use...
  - The same fabric services
  - The same management tools
  - The same diagnostic tools

- FC-NVMe & FC-SCSI offer...
  - The same full-matrix testing
  - Trusted support models
  - Trusted vendor relationships
Migrate applications with confidence

- Migrate volume by volume
- Migrate each volume based on its specific SLA needs
- Rollback models are simple. Individual volumes can be migrated in short windows
  - Easy to plan and implement

---

NCMe Traffic
SCSI Traffic
NVMe can be adopted now rather than much later

- Full scale app migration to new infrastructure is a major lift
  - Plan big windows for each app

- Big windows means that months will be needed to move all apps
  - Delays other projects that need big windows, progress slow & painful

- Net: if new infrastructure is required, storage teams will defer adoption
  - Or: if today’s infrastructure can be leveraged, teams can adopt quickly
Presentation Topics

- Background on NVMe
- How Concurrent SCSI and NVMe Lowers Risk
- Use Cases for NVMe over Fibre Channel
- The status of the FC-NVMe Protocol
- Recap and Q & A
Traditional latency focused storage apps

- NVMe reduces latency of enterprise storage features. As the latency savings of the NVMe decrease, the latency-sensitive apps, like financial apps, will be motivated to move to NVMe over FC.
Traditional IOPS focused storage apps

- NVMe’s enhanced queuing capabilities enable more parallelism, even for existing apps. As the number of server cores and threads grows, and the number of VMs explodes, there is increasing need to exploit the potential parallelism in solid state targets. NVMe is designed to enable that.
New latency focused “memory” apps

- The ultra-low latency of SSDs is corresponding with new apps, such as data mining or machine learning. These apps have voracious appetites for low latency persistent memory, beyond what fits in a server, and yet they often do not require the “enterprise features” of traditional enterprise storage. We might call these “memory-oriented” applications. They can leverage the low latency of Fibre Channel to access massive low latency NVMe arrays.
FC-NVMe / FC-SCSI dual protocol usage

- Database app maintains high value database on high SLA legacy array
- Data mining app requires super low latency reference image of DB
- Regularly Snapshot DB in legacy array
- Use Data mining server to copy snapshot to Ultra-low latency NSID
- Run Data mining application using low latency NSID reference copy
Presentation Topics

- Background on NVMe
- How Concurrent SCSI and NVMe Lowers Risk
- Use Cases for NVMe over Fibre Channel
- The status of the FC-NVMe Protocol
- Recap and Q & A
FC-NVMe Protocol Status

- T11 approved FC-NVMe spec in August
  - T11 formally sends it to INCITS Sept/Oct
- UNH Plugfest in June
  - 12 vendors participated
  - Next UNH plugfest will be in October
- Other plugfests happen frequently
Recent Demo Rack, UNH Interop Fabric

Target Emulator 32/16G FC-NVMe

Storage System 32/16 G FC

Storage System FC-NVMe 16/8 G FC

FC Storage

Server1/HBA 32G FC

Server2/HBA 16G FC

Server3/HBA 16G FC

Slow Drain Server/HBA

Native FC

NPV or AG

SANBLAZE (NVMe)

HP-3PAR (FCP)

EMULEX

QLOGIC

BROCADE G620

CISCO MDS

TELEDYNE LECROY

VIAVI

Amphenol

FCIA Demo Rack

2017 Storage Developer Conference. © Brocade. All Rights Reserved.
“Conservative” storage vendors acting…

- Many vendors have announced NVMe “in the backend”
- NetApp demoed NVMe over Fibre Channel in August
- Others have shared plans to support NVMe over FC
- Others are dropping hints
Presentation Topics

- Background on NVMe
- The status of the FC-NVMe Protocol
- How Concurrent SCSI and NVMe Lowers Risk
- Use Cases for NVMe over Fibre Channel
- Recap and Q & A
Recap

- NVMe over Fibre Channel offers concurrent NVMe / SCSI
- Concurrent FC-NVMe / FC-SCSI reduces risk & cost
  - Familiar fabric services, tools, vendor support
  - Lower risk makes NVMe easier to adopt
- Concurrent FC-NVMe / FC-SCSI offers new use cases
- The FC-NVMe Protocol Specification is stable
- Vendor support is ramping
Q & A

Who has a question?

Didn’t take notes? No worries…

You can download the eBook here:
dummies.com/go/nvmeoverfibrechannelfordummiesbrocadespecialedition