SNIA ESF | ETHERNET STORAGE

EVERYTHING YOU WANTED TO KNOW ABOUT STORAGE, BUT WERE TOO PROUD TO ASK

Where Does My Data Go?

August 1, 2017
10:00 am PT

SNIA Legal Notice

- 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.

Today's Presenters

Scott Shadley
Micron

Rob Peglar
Symbolic IO

J Metz
Cisco

Chad Hintz
Cisco

SNIA-At-A-Glance

SNIA-at-a-Glance

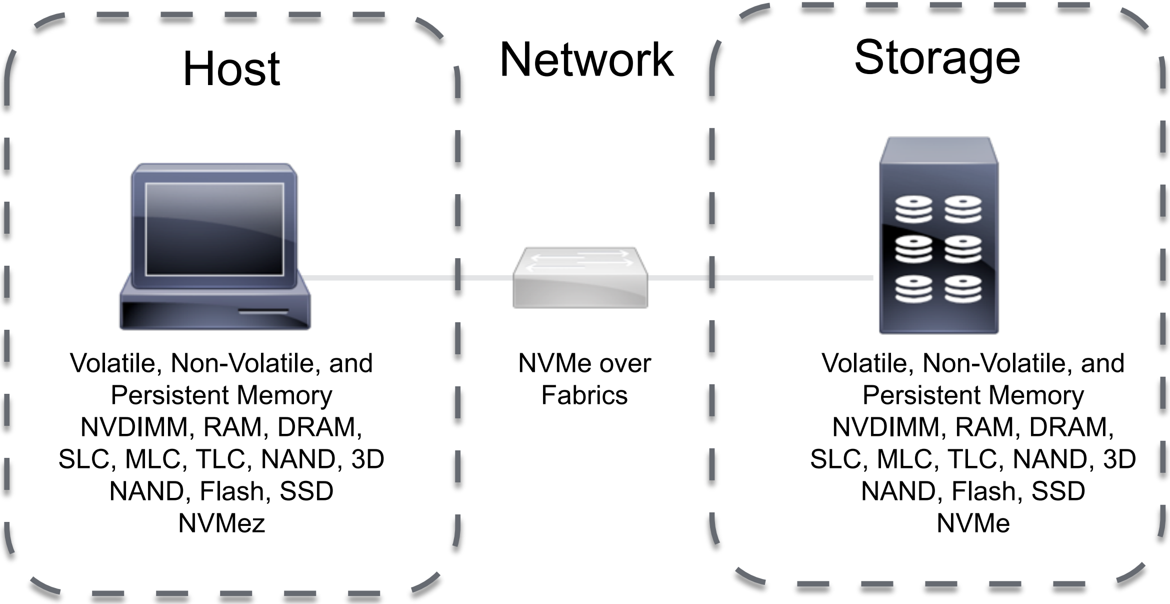
160
unique member companies

2,500
active contributing members

50,000
IT end users & storage pros worldwide

Agenda

- Volatile, Non-Volatile and Persistent Memory
- NVDIMM, RAM, DRAM, SLC, MLC, TLC, NAND, 3D NAND, Flash, SSDs, NVMe
- NVMe and NVMe over Fabrics



Volatile, Non-Volatile, and Persistent Memory

Scott Shadley, Micron

Definition

- Volatile Storage
  - SNIA Defines volatility as:
    - [Computer System] A property of data yielding the possibility that it will be obliterated if certain environmental conditions are not met.
    - For example, data held in DRAM is volatile, since if electrical power to DRAM is cut, the data in it is lost.
  - Another Definition would be:
    - Data Volatile and lost when the ability to 'refresh' the data is removed. Either by removing power, or 'refresh engine' issues

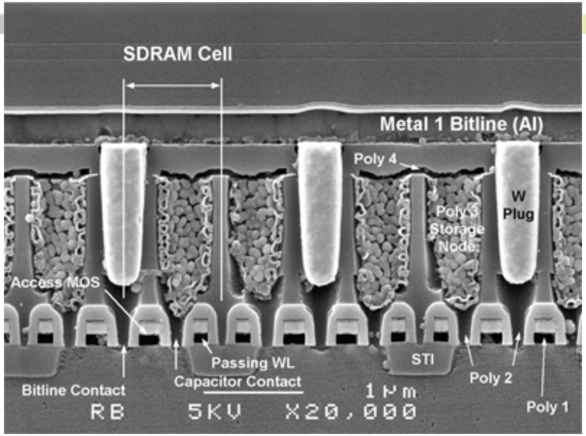
Volatile Storage

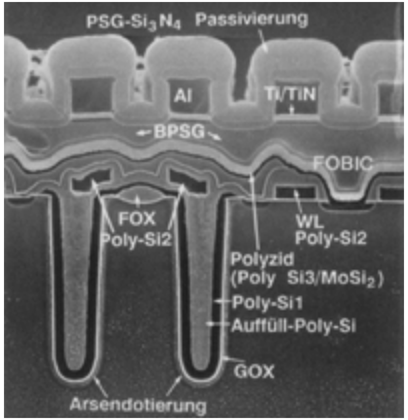
- Computer Memory is Volatile due to the capacitor used to store host data –

  Charged is a 1

  Discharged is a 0

- Capacitors discharge when no power is attached, or sustained
- EXTREMELY FAST READ/WRITE





Definition

SNIA ESF | ETHERNET STORAGE

- **Non-Volatile Storage**
  - SNIA defines non-volatility as:
    - [Storage System] The property of an electronic device that data is preserved even when electrical power is removed
  - Another Definition would be:
    - Data is Programmed and Erased from non-volatile data and can be powered or unpowered and not lose data once programed

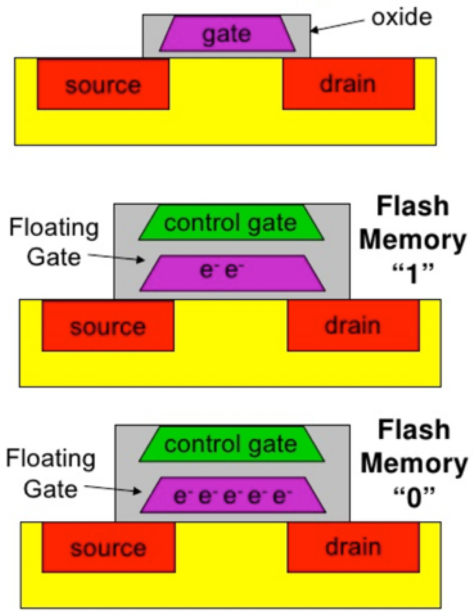
Non-Volatile

SNIA ESF | ETHERNET STORAGE

- Non-Volatile is Programmed, not "charged" like Volatile memory
  - A state change of a the transistor element is modified to denote a 0 or 1

    An 'Erased' location is a 1

    A 'Programmed' location is a 0
- Since the cell is isolated it is persistent
- LONG PROGRAM, FAST READ



Definition

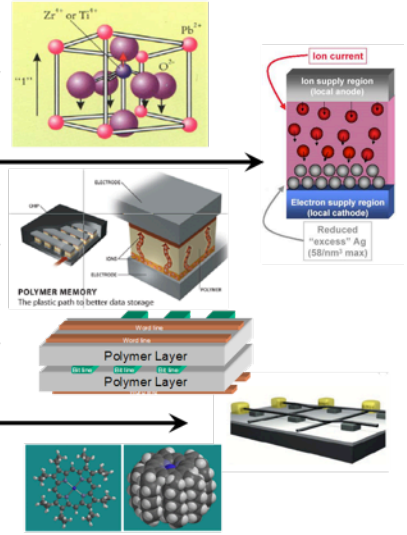
- Persistent Memory
  - SNIA Defines Persistence as:
    - [Computer System] A synonym for non-volatility, usually used to distinguish between data and metadata held in DRAM, which is lost when electrical power is lost, and data held on non-volatile storage
  - Another Definition would be:
    - A material used to store data that requires an alternate state of the material. Some versions today exists as MRAM, 3DXPoint, RRAM, STTRAM
- SMALLER READ/WRITE Differences, Not as fast as Volatile Memory

Persistent Memory

SNIA ESF | ETHERNET STORAGE

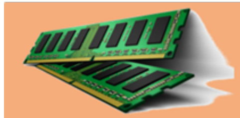
- These technologies require a ‘Change’ to the material to manage 0 or 1
- Material Scaling is challenged
- True Persistence is not clear on some techs

Material	Structure
Complex Metal Oxide	PCMO Material Derived from High Temp Superconductor
Programmable Metallization	Silver Dissolved in Chalcogenide
Resistive Polymer	“Special Formulation” Polymer
Ferro Electric Polymer	Polymer chains with a dipole moment
Carbon Nanotube	Cross Point Array of Nanotubes
Molecular	Molecules in CrossPoint Array



Intro to Product Comparisons

SNIA ESF | ETHERNET STORAGE



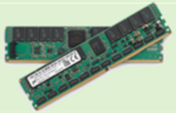
RDIMM/LRDIMM
(DRAM)
35ns

Fastest to Read and Write
Lowest Density

No Power, No DATA

New Technologies in Progress

NVDIMM-N
(DRAM/NAND)
35ns



NVDIMM-P
(DRAM/NVM)
200ns

NVME SSD
(3D XPoint)
15us

Has Fast Read, Slower Write
Densities Vary, but still lower
Data stays or is auto recovered

Mainstay of fast Storage Today
Interface dictates Speed
Largest Density options

NVME SSD
(NAND)
85us

SATA/SAS SSD
(NAND)
115us

Data stays, finite life

Historical Data Storage Solution

Hard Disk Drive
(Spinning Media)
10ms

Volatile ← Persistent — Non-Volatile → Magnetic

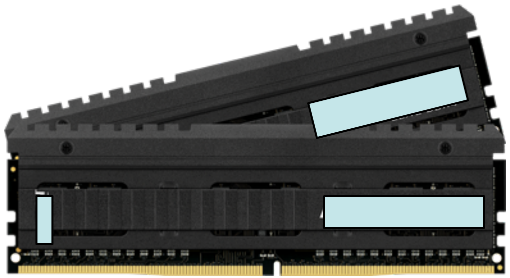
NVDIMM, RAM, DRAM, SLC, MLC, TLC, NAND, 3D NAND, Flash, SSDs, NVMe

Alphabet Soup?

Rob Peglar, Symbolic IO

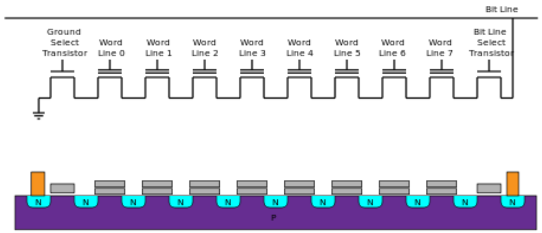
Where do We Start?

- How about a Taxonomy?
  - You just heard about different forms of memory
  - How are those forms deployed in practice?
  - Let's start with RAM
    - Random Access Memory – aka byte-addressable memory
    - Accessed over a memory channel – tied to a processor (CPU, GPU, FPGA, etc.)
    - Several forms of RAM, the most familiar of which is DRAM
      - DRAM – Dynamic RAM – a form of volatile memory using transistors and capacitors
      - DRAM die are typically deployed in Dual Inline Memory Modules – DIMMs
      - DRAM is also typically deployed in smaller form factors inside mobile/embedded devices

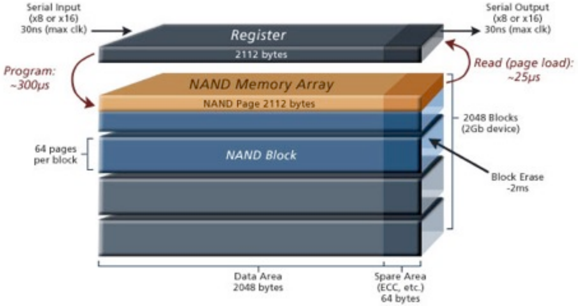


Flash Dance

- Can we get flashy wit it?
  - Yes we can
  - Flash memory is:
    - Non-volatile – as opposed to DRAM which is volatile
    - Two types of flash memory – NOR and NAND (not-or, not-and electronic circuits)
    - NAND flash forms the fundamental technology for several devices
    - NAND flash only holds blocks of data – not byte-addressable
    - NAND blocks organized into ‘pages’ – the erase/write (aka program) boundary
    - NAND cells have finite PE (program/erase) cycles – very important for reliability
    - NOR flash often found in small form-factor embedded devices (e.g. in autos)
      - Not nearly as dense as NAND, relatively more expensive than NAND



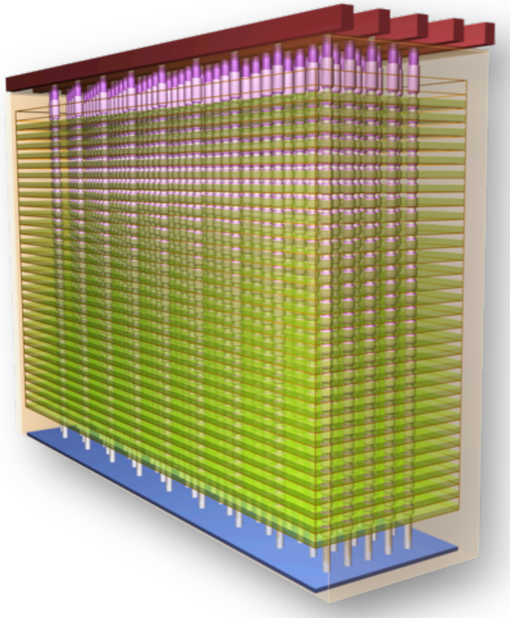
Two Bits, Four Bits, Six Bits, a Dollar



- What's all the fuss about bits per cell?
  - Let's talk NAND flash
    - SLC – single level cell – the first form of NAND
      - Two voltage levels inside the cell – representing 0 or 1
      - Nice and simple – and fast with high PE cycles – but not dense
    - MLC – multiple level cell – aka two-bits per cell (in practice)
      - Four voltage levels inside the cell – representing two bits – 00, 01, 10, 11
      - Slower than SLC and fewer PE cycles but more dense – it's a two-fer
    - TLC – triple level cell – three bits per cell
      - Eight voltage levels representing three bits – 000, 001, 010, 011, 100, 101, 110, 111
      - Slower than MLC and fewer PE cycles but more dense – ratio 3/2
    - QLC is under development – four bits per cell – you get the drift

Let's Get Vertical

- What the heck is 3D NAND?
  - Think of a skyscraper – multiple floors
    - 3D NAND has layers – all interconnected
    - These layers are extremely compact (vertically)
    - Each layer has a certain capacity
    - 3D NAND is typically either MLC or TLC design
    - 32, 48 and 64-layer on the market now
      - 96 and even 128-layer have been proposed
    - Excellent for use in very dense deployments
    - Lots of new research & development in 3D designs
    - 3D will be the basis for most new future products using NAND



It's All About the Storage

- What's an SSD?
  - Solid State Disk – all electronic, no moving parts
  - As opposed to the Hard Disk Drive (HDD) – electromechanical
    - SSDs are typically use NAND for the persistence layer
      - Some SSDs also have DRAM acting as a data cache for writes (and sometimes reads)
      - One SSD is made with another type of PM (3DX); MRAM, RRAM, STT-RAM future
    - SSDs are very diverse in terms of capacity, speed, endurance, & price
      - Have now surpassed HDDs in density (capacity per physical size)
    - SSDs come in various form factors
      - 2.5" form factor with U.2 connector – SATA, SAS, or NVMe
      - M.2 form factor – SATA or NVMe
      - Add-in card form factor – PCI-Express – NVMe

Last but Not Least

- What's this NVDIMM thing?
  - Non Volatile DIMM – in the memory channel
  - Three types are defined (by JEDEC)
    - -F is all NAND flash
    - -N is half DRAM, half NAND (depicted)
    - -P is little DRAM, big NAND
    - -F is block-addressable; -N and -P are byte-addressable
  - NVDIMMs are an example of persistent memory
    - Requires an external energy source to insure data in DRAM is persisted to NAND
    - Typically deployed in the form of supercapacitors
    - They kick in when AC power is removed from the system – special circuitry

NVDIMM, RAM, DRAM, SLC, MLC, TLC, NAND, 3D NAND, Flash, SSDs, NVMe

Wasn’t that Fun?

NVMe and NVMe over Fabrics

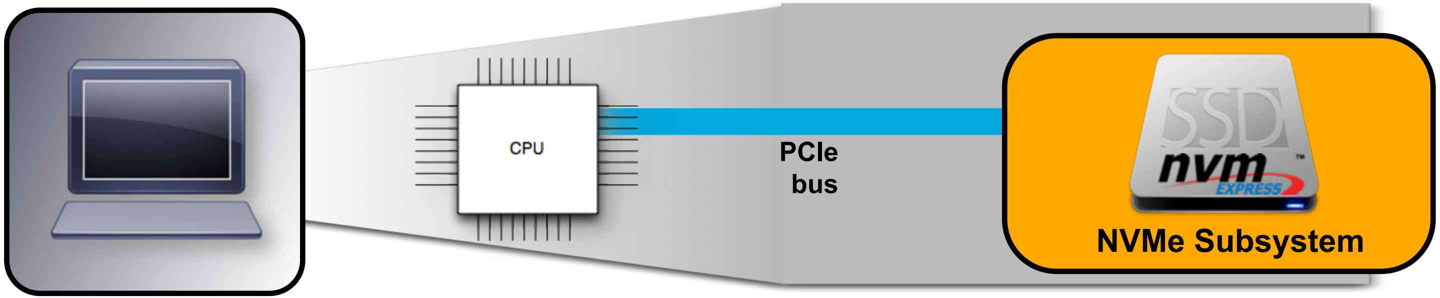
J Metz, Cisco

Definition

- **Non-Volatile Memory Express (NVMe)**
  1. [Storage System] A host controller interface with a register interface and command set designed for PCI Express®-based SSDs.
  2. [Organization] Technical body that defines and standardizes NVMe and NVMe over Fabrics (NVMe-oF) protocols

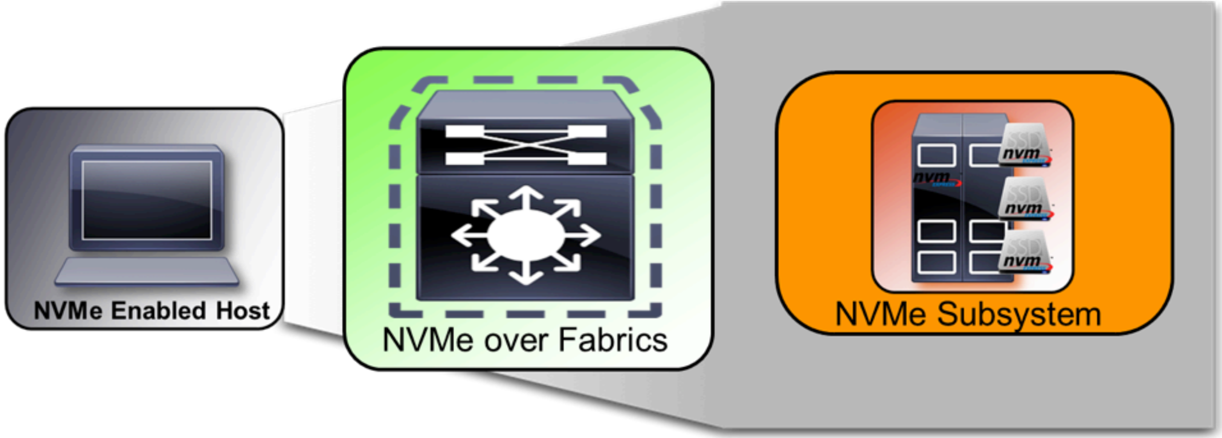
What is NVMe?

- A "Host Controller Interface"
  - Allows a computer (the host) to connect to storage devices
  - Standardizes the register set, feature set, and command set
- Designed for Non-Volatile Media
  - Architected from the "Ground Up"
  - Developed by an open industry consortium group
- PCIe-Based
  - Low-latency, High IOPS, Direct-Attached NVM storage



Definition

- **Non-Volatile Memory Express over Fabrics (NVMe-oF)**
  - An extension of the NVM Express controller interface to support message-based NVMe operations over a network fabric.



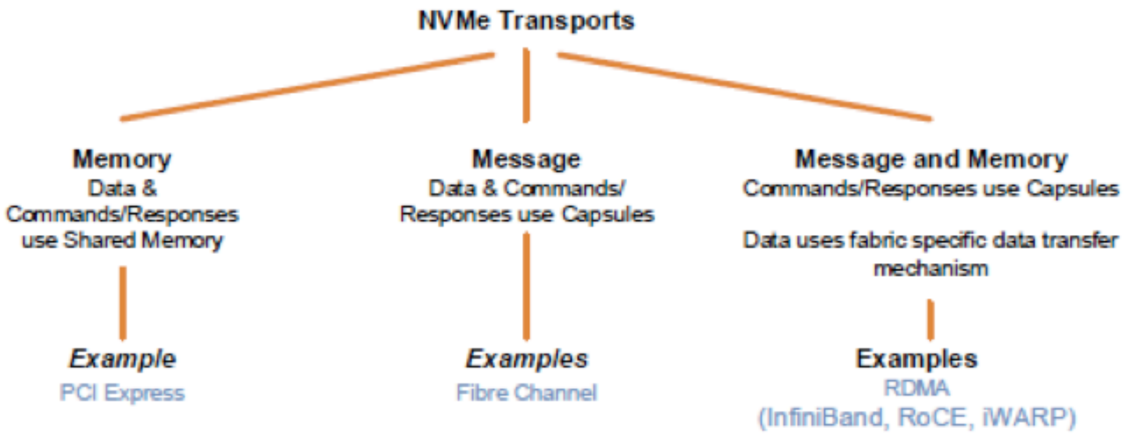
NVMe/NVMe-oF

- Non-Volatile Memory Express (NVMe)
  - Began as an industry standard solution for efficient PCIe attached non-volatile memory storage (e.g., NVMe PCIe SSDs)
  - Low latency and high IOPS direct-attached NVM storage
- NVMe over Fabrics (NVMe-oF)
  - Built on common NVMe architecture with additional definitions to support message-based NVMe operations
  - Standardization of NVMe over a range Fabric types
    - Initial fabrics; RDMA(RoCE, iWARP, InfiniBand™) and Fibre Channel

Different Models

- NVMe is a Memory-mapped, PCIe Model
- Fabrics is a message-based transport; no shared memory

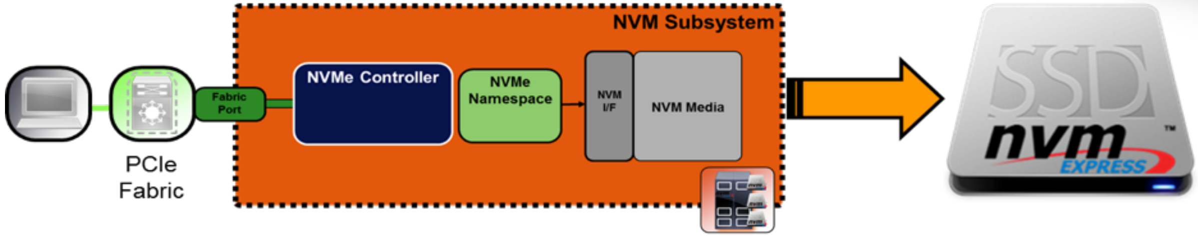
Figure 1: Taxonomy of Transports



NVMe Subsystem Implementations

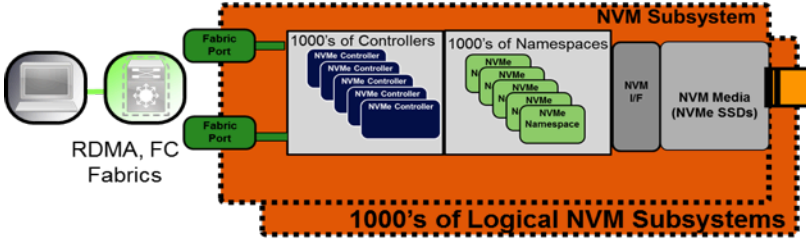
NVMe PCIe SSD Implementation
(single Subsystem/Controller)

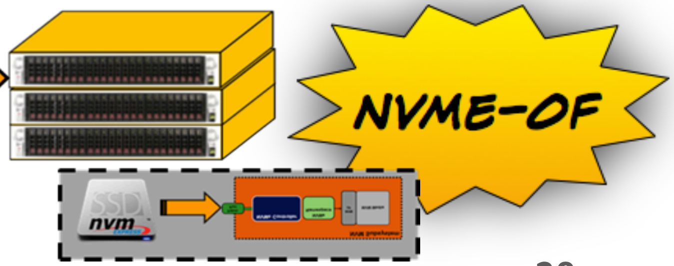
NVMe PCIe SSD



NVMe all NVM Storage Appliance Implementation
(1000’s of Subsystems/Controllers)

All NVM Appliance with PCIe NVMe SSDs





Summary

SNIA ESF | ETHERNET STORAGE

- Volatile, Non-Volatile and Persistent Memory fit different use cases. It’s important to factor in the speeds and volatility of each to make the right decision for each application.
- Memory comes in many different form factors: NVDIMM, DRAM, SLC, MLC, TLC, NAND, 3D NAND, Flash, SSDs and NVMe. Chose wisely to select the one that fits your use case.
- Lastly, we showed use cases for NVMe, a host controller interface for PCIe-based SSDs and NVMe-oF, an extension of NVMe to support operations over a fabric.

Other Storage Terms Got Your Pride? This is a Series!

SNIA ESF | ETHERNET STORAGE

Don’t miss our next one – September 28, 2017

Everything You Wanted To Know About Storage But Were Too Proud To Ask – Cyan

(Storage Management)

- SNMP
- SMI-S
- SNIA Swordfish
- Software-defined storage
- **Register at: https://www.brighttalk.com/webcast/663/269513**

Other Storage Terms Got Your Pride? This is a Series! Check Out:

- **Teal** – Buffers, Queues and Caches https://www.brighttalk.com/webcast/663/241275
- **Rosé** - All things iSCSI https://www.brighttalk.com/webcast/663/244049
- **Chartreuse** – The Basics: Initiator, Target, Storage Controller, RAID, Volume Manager and more https://www.brighttalk.com/webcast/663/215131
- **Mauve** – Architecture: Channel vs. Bus, Control Plane vs. Data Plane, Fabric vs. Network https://www.brighttalk.com/webcast/663/225777
- **Sepia** – Getting from Here to There https://www.brighttalk.com/webcast/663/249431
- **Vermillion – What if Programming and Networking Had a Storage Baby?** https://www.brighttalk.com/webcast/663/260321

Speaking of Series…Check out Storage Performance Benchmarking

- Storage Performance Benchmarking:
  1. Introduction and Fundamentals
  2. Solution under Test
  3. Block Components
  4. File Components

Watch them all on-demand at:
http://www.snia.org/forums/esf/knowledge/webcasts-topics

After This Webcast

SNIA ESF | ETHERNET STORAGE

- 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 at 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
- Follow us on Twitter @SNIAESF
- Need help with all these terms? Download the SNIA Dictionary http://www.snia.org/education/dictionary