NVMF-302-1: Benefits and Use Cases for NVMe-oF

Steve McQuerry, Senior TME
PureStorage
NVMe Transports

- Memory
  - Data & Commands/Responses use shared memory
  - Example: PCI Express

- Message
  - Commands/Responses use Capsules
  - Data may use Capsules or Messages
  - Examples: Fibre Channel, TCP

- Message & Memory
  - Commands/Responses use Capsules
  - Data may use Capsules or Shared Memory
  - Examples: RDMA (InfiniBand, RoCE, iWARP)

Capsule = Encapsulated NVMe Command/Completion within a transport message
Data = Transport data exchange mechanism (if any)
NVMe on Arrays

Chassis PCIe

NVMe Backend PCIe/NVMe-oF

Shelf NVMe/RoCE

NVMe Frontend NVMe-oF
Backend NVMe transport examples

- **PCIe**
  - Good for connectivity within a chassis, outside of the chassis it has challenges (hardware, scale, operations)

- **NVMe/RoCE**
  - Uses same transport model as PCIe (direct memory) and mitigates challenges
  - Pure Storage has been shipping NVMe/RoCE on the backend as DirectFlash Shelf to scale capacity since 2018
Frontend NVMe transport examples

• NVMe/RoCE
  – Leverages Ethernet. Use cases include front end and back end deployments. Customer applications tend to be disaggregated POD designs

• NVMe/FC
  – Mature Fabric easy path to adoption for customers who want to leverage existing FC operations model

• NVMe/TCP
  – Generic Ethernet connectivity. Use cases include designs that leverage data network and storage network
Frontend NVMe Ethernet topologies

Single Hop

Spine Leaf
Disaggregated Designs

Rack Design

POD Design
Customer/Market Observations

- The OS ecosystem for NVMe-oF is still developing
- Many Fibre Channel customers are looking to adopt NVMe/FC on existing infrastructure
- Some customers are looking to move to an all Ethernet data center fabric
- NVMe/RoCE is works well in disaggregated rack and pod deployment scenarios
- NVMe/TCP is gaining mindshare and poised to be very popular in less deterministic Ethernet environments
Q & A