FUTURE OF DATACENTER STORAGE

Carol Wilder, carol.a.wilder@intel.com

Niels Reimers, niels.reimers@intel.com

LEGAL NOTICES/DISCLAIMER

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Learn more at intel.com, or from the OEM or retailer.

No computer system can be absolutely secure.

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors.

Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit http://www.intel.com/performance.

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. For more complete information about performance and benchmark results, visit http://www.intel.com/performance.

Cost reduction scenarios described are intended as examples of how a given Intel- based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction.

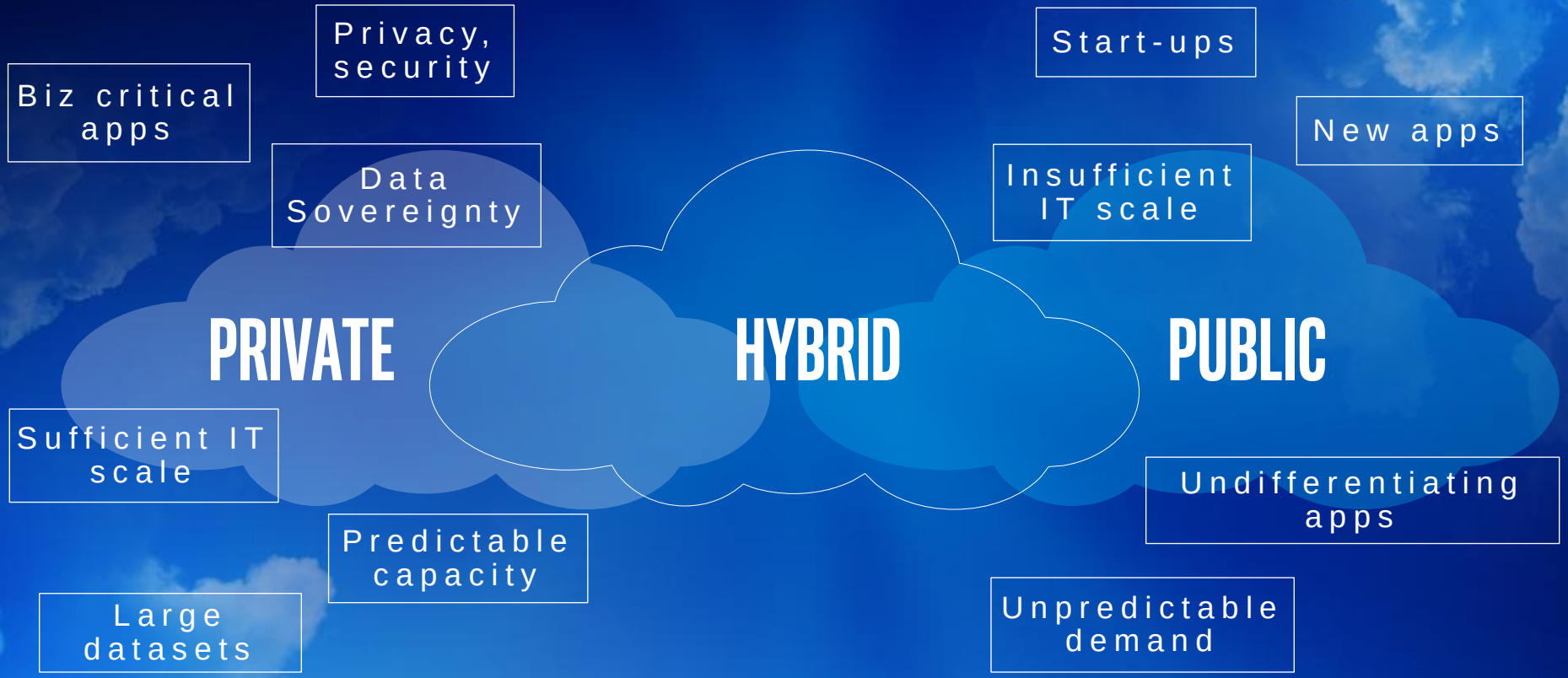
Results have been estimated or simulated using internal Intel analysis or architecture simulation or modeling, and provided to you for informational purposes. Any differences in your system hardware, software or configuration may affect your actual performance.

Intel does not control or audit third-party benchmark data or the web sites referenced in this document. You should visit the referenced web site and confirm whether referenced data are accurate.

Intel, the Intel logo and others are trademarks of Intel Corporation in the U.S. and/or other countries. *Other names and brands may be claimed as the property of others.

EXTRAORDINARY NEW DEMANDS FACE THE DATA CENTER

A VARIETY OF REQUIREMENTS ACROSS APPS



STORAGE AND DATA ACROSS INFRASTRUCTURES

TRADITIONAL STORAGE

CHALLENGE

Seamlessly access data anywhere, at anytime, on any device, at the required performance

CLOUD STORAGE

THE STORAGE MODERNIZATION JOURNEY

Modernize

Storage Intelligence

- Thin-provisioning
- De-duplication
- Compression
- Encryption
- Multi-tenancy

Hard Drive to Solid State Drive

- SSD Moore's Law
- **Storage Intelligence**
- Performance
- Emergence of new storage arrays:
  - All Flash
  - Hybrid

Automate

Architecture & Workload

Unifying Cloud & Enterprise

- Scale Up Architecture
- Scale Out Architecture
- Hyperconverged
- Software Defined Storage
  - Orchestrated management
  - Service Level Agreements

Orchestrate

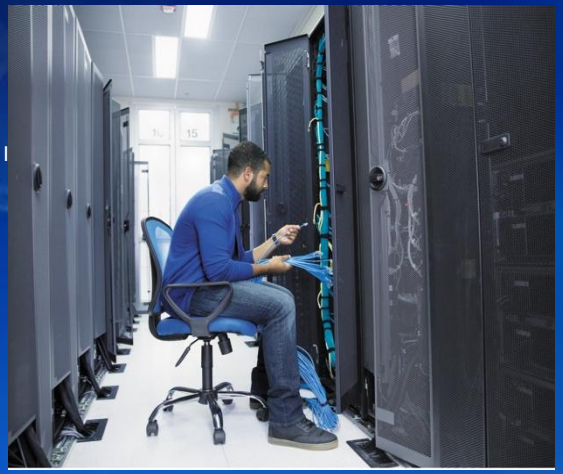
Seamless integration across architecture, media and vendors

- Network Function Virtualization
- Software Defined Infrastructure
- Shift of the IO bottleneck
- Business alignment

Storage Modernization enhances existing and new environments enabling seamless data services in the datacenter

INTEL ENABLING STORAGE MODERNIZATION

CREATING SEAMLESS DATA SERVICES IN THE DATACENTER



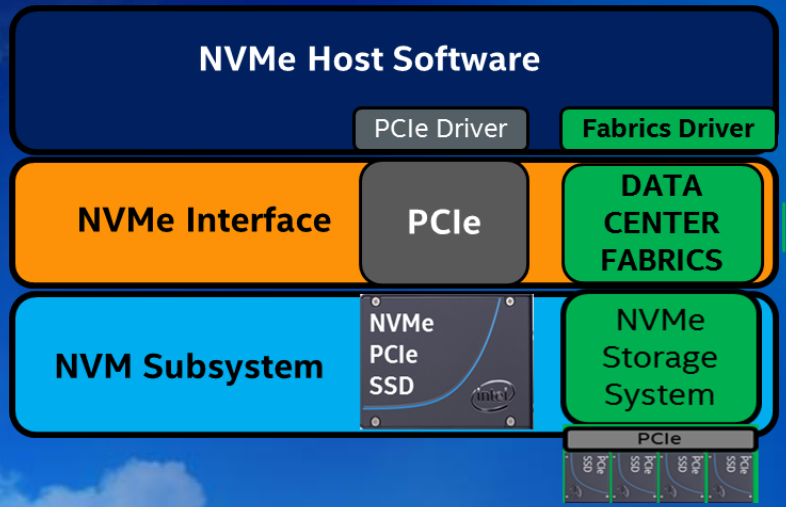
Server Based
Storage Architectures

Media Transition

Software Defined

WHAT IS NVM EXPRESS OVER FABRICS?

- Industry standard definition of NVMe over Data Center Fabrics
- Shares the same base Architecture and NVMe Host Software as PCIe
- Enables NVMe scale-out and low latency operations on Data Center Fabrics



Non-PCIe Transports and Data Center Fabrics

RDMA Fabrics
Ethernet (iWARP/RoCE)
InfiniBand™
Intel OmniPath™ V1

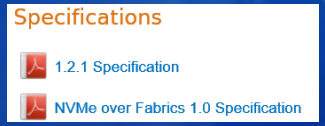
Fibre Channel (T11)

Future Fabrics
Intel OmniPath™ V2

NVMe over Fabrics is the most efficient Block Interface on Data Center Fabrics

NVME OVER FABRICS STANDARDIZATION AND ENABLING

- Industry-wide NVMexpress.org TWG community defined NVMe Fabrics
  - Dozens of companies participated in the definition over last ~18 months
- NVMe Over Fabrics Specifications were released on June 6th 2016
  - Downloadable from www.nvmexpress.org/specifications



- Linux Kernel Host and Target Drivers available
  - www.nvmexpress.org/drivers

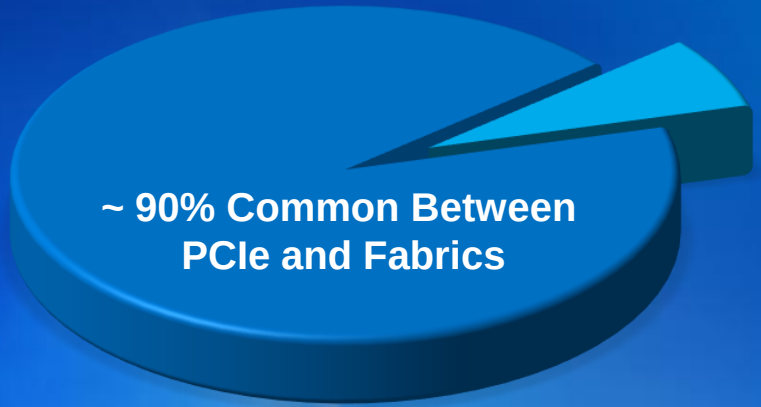
  - Integrated and tested with Linux 4.7-RC1 Kernel, plan is to upstream with 4.8 Kernel
- Intel SPDK User-Level NVMe over Fabrics Target
  - Downloadable from https://github.com/spdk

COMMONALITY BETWEEN NVME ON PCIE AND FABRICS

The vast majority of NVMe is leveraged as-is for Fabrics

- NVMe host interface, NVM Subsystem, Controllers, Namespaces, Commands, Registers/Properties, Power States, Asynchronous Events, Reservations, etc.
- Allows for use of common NVMe Host software with very thin fabric dependent layers

Primary differences reside in the discovery and queuing mechanisms



Differences	PCI Express® (PCIe)	Fabrics
Identifier	Bus/Device/ Function	NVMe Qualified Name (NQN)
Discovery	Bus Enumeration	Discovery and Connect commands
Queuing	Memory-based	Message-based
Data Transfers	PRPs or SGLs	SGLs only, added Key

NVME OVER FABRICS NVME COMMAND / COMPLETION DELIVERY

Fabric Capsules are messages with “encapsulated” **common** NVMe content

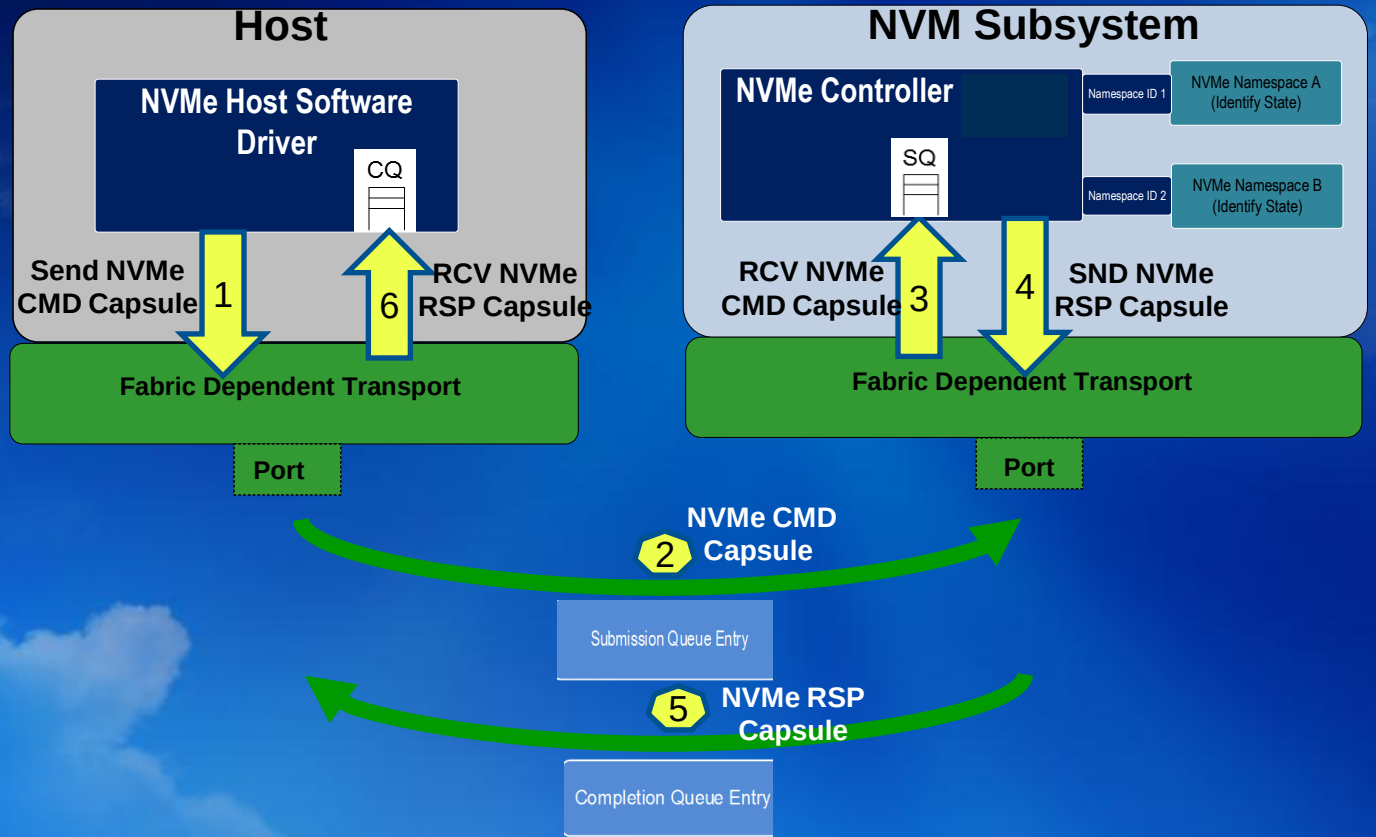
Host to NVMe Controller

NVMe Controller to Host

Data Section may be sent within the Capsule (as shown) or via a fabric type dependent data transfer mechanism, example RDMA_READ/RDMA_WRITE

Data Transfer to/from Host resident buffer

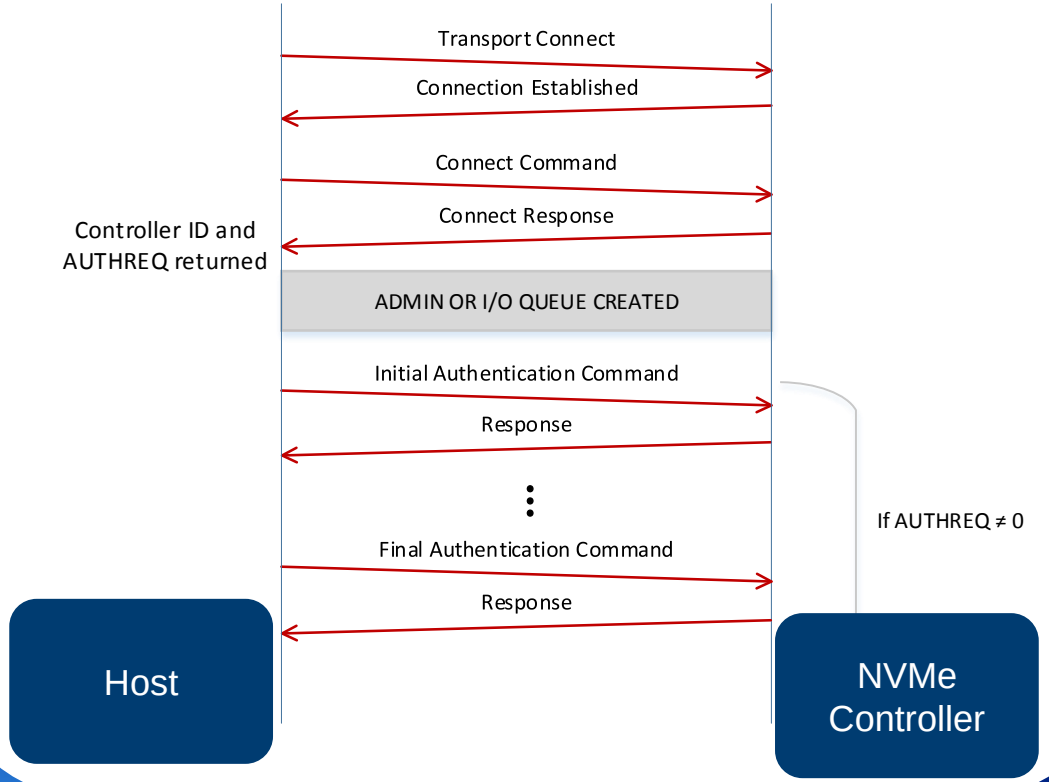
NVME OVER FABRICS CAPSULE EXCHANGE EXAMPLE



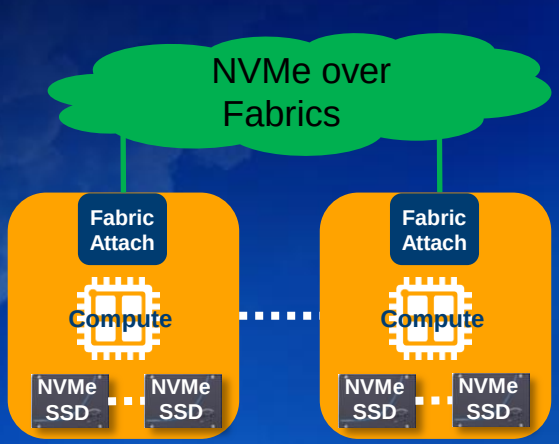
NVME QUEUE CREATION USING FABRIC CONNECT COMMAND

- Create a fabric-dependent transport connection
- Send a Command Capsule with Fabric Connect Operation (AdminQ Connect establishes an “association” to an NVMe Controller)
- <opt> Send Authentication Fabric Commands
- AdminQ or IOQ Ready for NVMe Commands

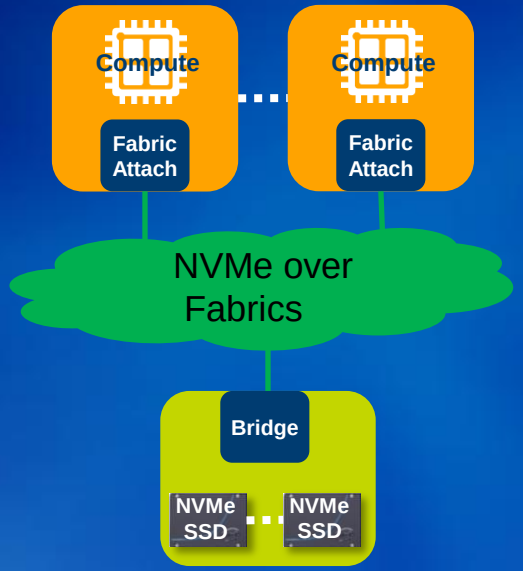
Figure 1: Queue Creation Flow



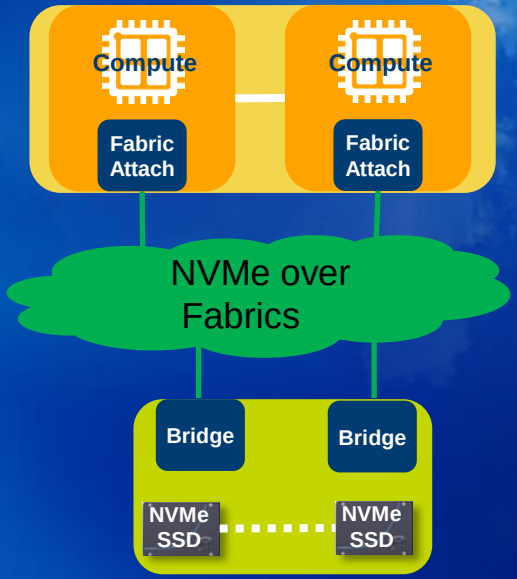
SOME OF THE USE CASES FOR NVME OVER FABRICS



Hyperconverged Cloud

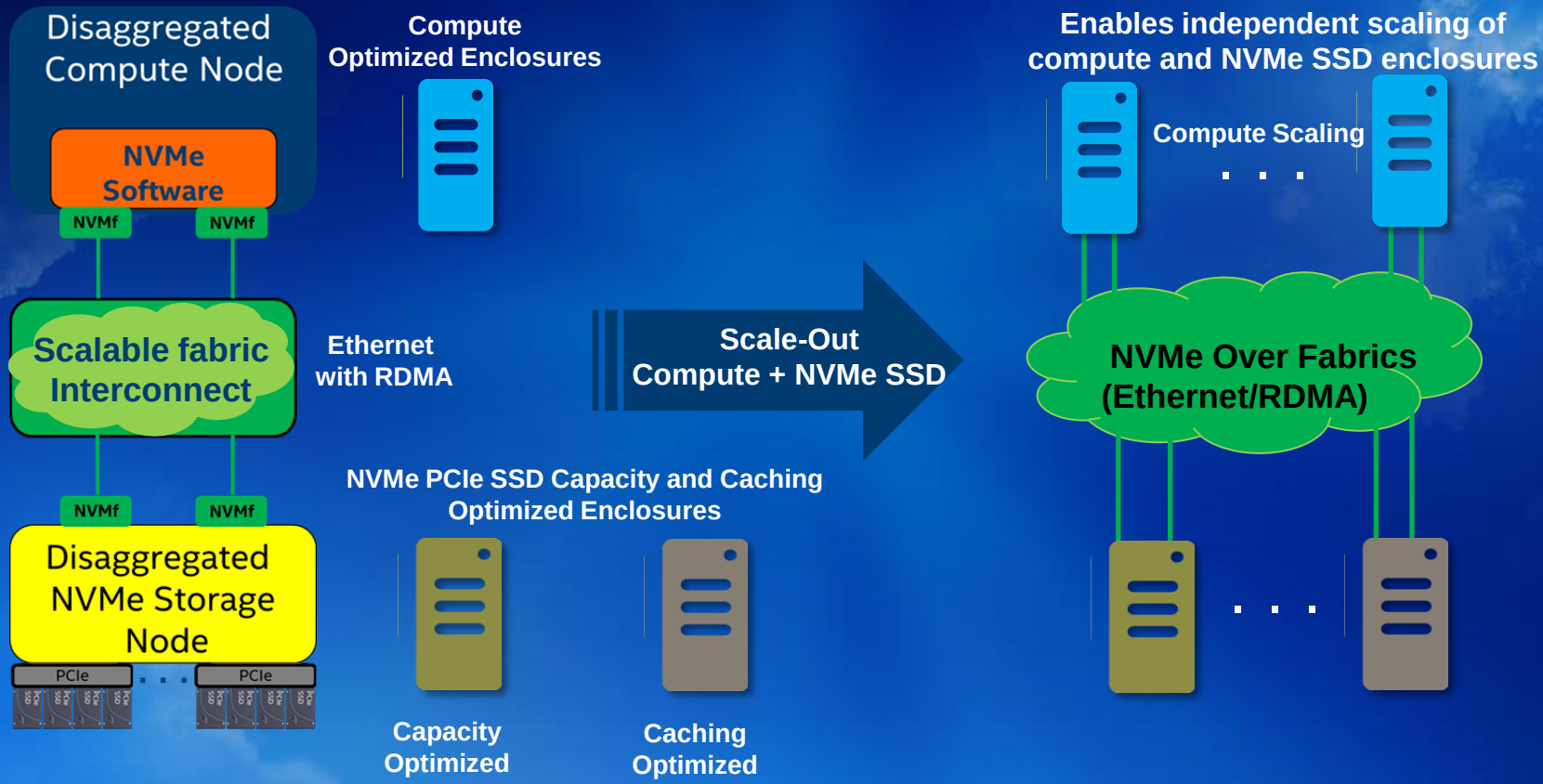


Disaggregated Cloud



External Storage

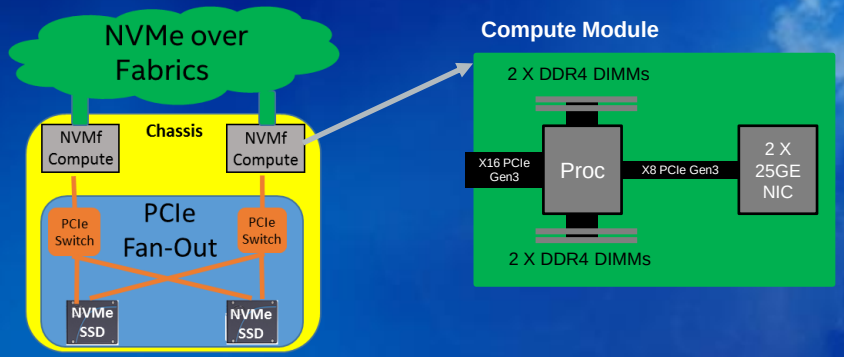
NVME OVER FABRICS CLOUD STORAGE TAXONOMY



DISAGGREGATED NVME OVER FABRICS ENCLOSURE OPTIONS

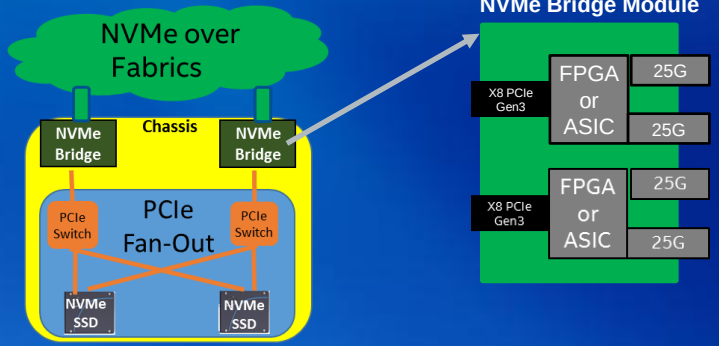
Intel Processor Based

- NVMf Target S/W (Front-End)
- SPDK or Linux Target S/W
- Typically have storage abstractions
  - Endurance, Security, High Avail. (RAID, Compute fail-over, …)



NVMe Bridge Based

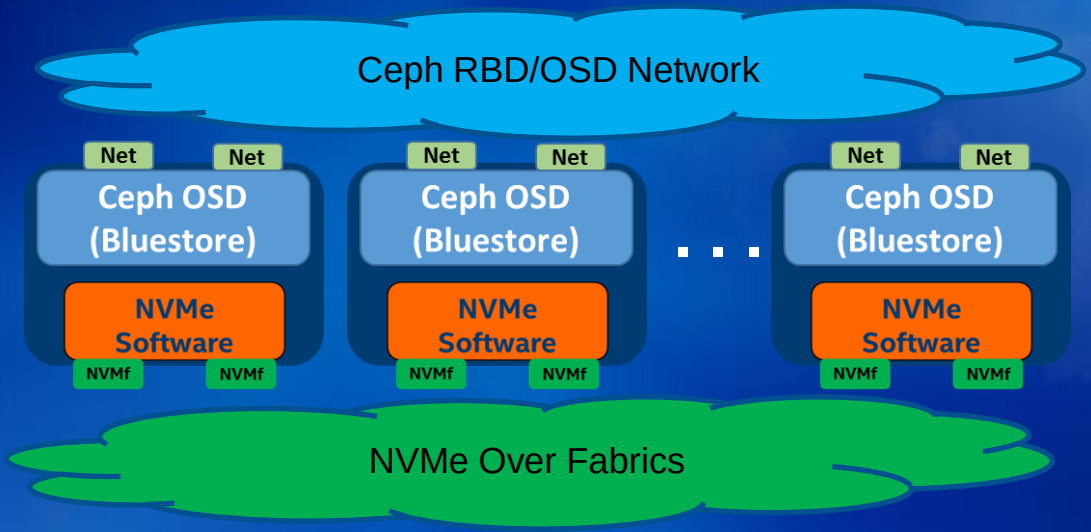
- Bridges NVMf to PCIe NVMe
- Discrete or integrated into RNIC
- High IOPS/Low Power
- JBOD like functionality with limited SSD sharing capabilities
- PCIe SSD sharing features will help
  - Multiple namespaces, SR_IOV, CRB, ROC, …



CLOUD STORAGE USAGE

Model (Disaggregated OSD block-store)

- Ceph OSD Software runs on disaggregated compute nodes (2 Socket Intel® Xeon™ E5)
- NVMe software enables access to disaggregated NVMe SSDs



- Capacity and Performance optimized NVMe SSD enclosures

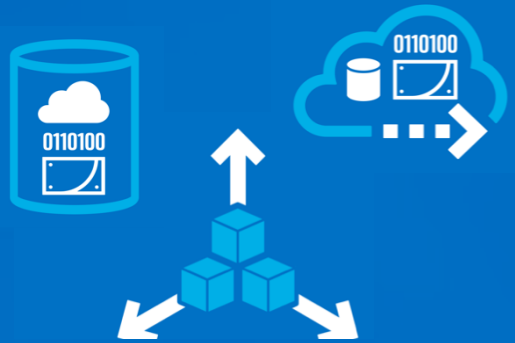
ENABLE SEAMLESS DATA SERVICES ACROSS CLOUDS

ENTERPRISE CLOUDS

DATA SERVICES

PUBLIC CLOUDS

Seamless Data Services: Agile, Automated, and Secure



Unified, Comprehensive Services Management & Orchestration – IaaS to SaaS

THE FUTURE OF STORAGE STARTS HERE

ACCELERATE THE PACE OF CHANGE

OPTIMIZE STORAGE SOLUTIONS WITH BUILDERS LABS

ENABLE SEAMLESS DATA SERVICES ACROSS CLOUDS AND TO THE EDGES

INTEL

EXPERIENCE NEXT GENERATION STORAGE

INVEST

in accelerating next generation storage solutions

EMBRACE

Intel® Solid State Drive & Intel® Optane technology for high performance

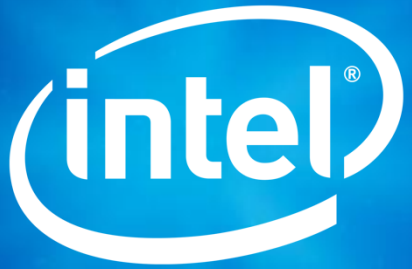
ALIGN

to accelerate software-defined storage

INTEL® STORAGE BUILDERS NETWORKING RECEPTION

BAYSHORE ROOM SECOND FLOOR, SEPTEMBER 21ST 6:00-8:00 PM

Join us for a mixer to enjoy an evening full of networking and to leverage business opportunities.


intel