



# NVMe over Fabrics

September 20<sup>th</sup>, 2017

*Brandon Hoff, Broadcom*

*Brandon.Hoff@Broadcom.com*



# Agenda

- ❖ NVMe over Fabrics
  - ❖ Update
  - ❖ Market
  - ❖ Roadmap
  - ❖ NVMe-TCP
- ❖ The benefits of NVMe over Fabrics
- ❖ Transport Independent - Storage Architectures
- ❖ Use Cases

# The Value of Shared Storage and the 'need for speed'

- The cost of data-at-rest is no longer the right metric for storage TCO
  - The value of data is based on how fast it can be accessed and processed
- NVMe over Fabrics increases the velocity of data
  - Faster storage access enables cost reduction through consolidation
  - Faster storage access delivers more value from data
- SSDs are going to become much faster
  - 3D Xpoint Memory, 3D NAND, etc.
  - PMEM, Storage Class Memory, etc.
  - ... and innovation will continue



# NVMe for Fabrics Progress (NVMe-oF)

## NVMe-oF Developments (NVM Express)

**6/6/16** – Fabrics Working group published v1.0

**6/6/16** – NVMe Fabrics Linux Driver working group published Linux target and driver code for inclusion in Linux kernel.org

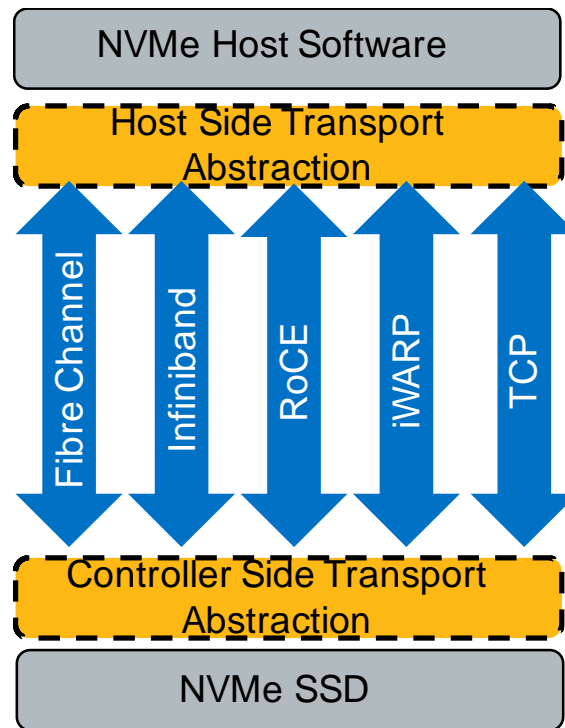
**8/8/16** – NVMe over Fabrics Panel at FMS 2016

**9/7/17** – NVMe over Fabrics Roadmap released and NVMe-TCP announced at FMS 2017

**2019** – NVMe-oF.next

## Additional Enterprise features being developed

- Boot from SAN
- Multi-pathing for failover & load balancing



# NVMe over Fabrics delivers for the External Block Storage Market

All Flash Arrays is a **\$6.8B Market** in 2017, growing at a **32% CAGR**.

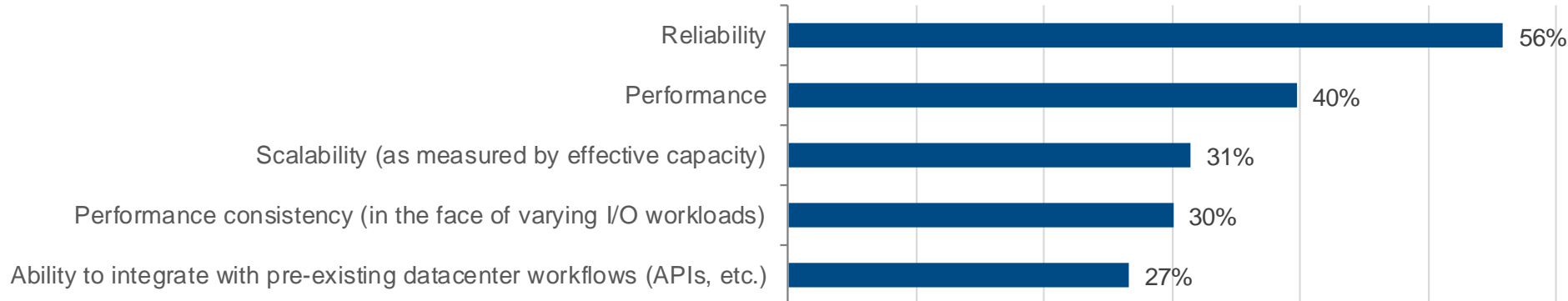
Only 13% of storage capacity shipped is DAS (inside the server), **87% of the total storage capacity shipped is external storage.**

NVMe-oF 1.0 was released in June 2016 and provides support for **RDMA and Fibre Channel**, plus **NVMe-TCP with 1.1**

As NVMe becomes adopted, NVMe-oF will enable applications **access to 1000's of NVMe drives with FC, RoCE, iWARP, IB, or TCP** as transport options.

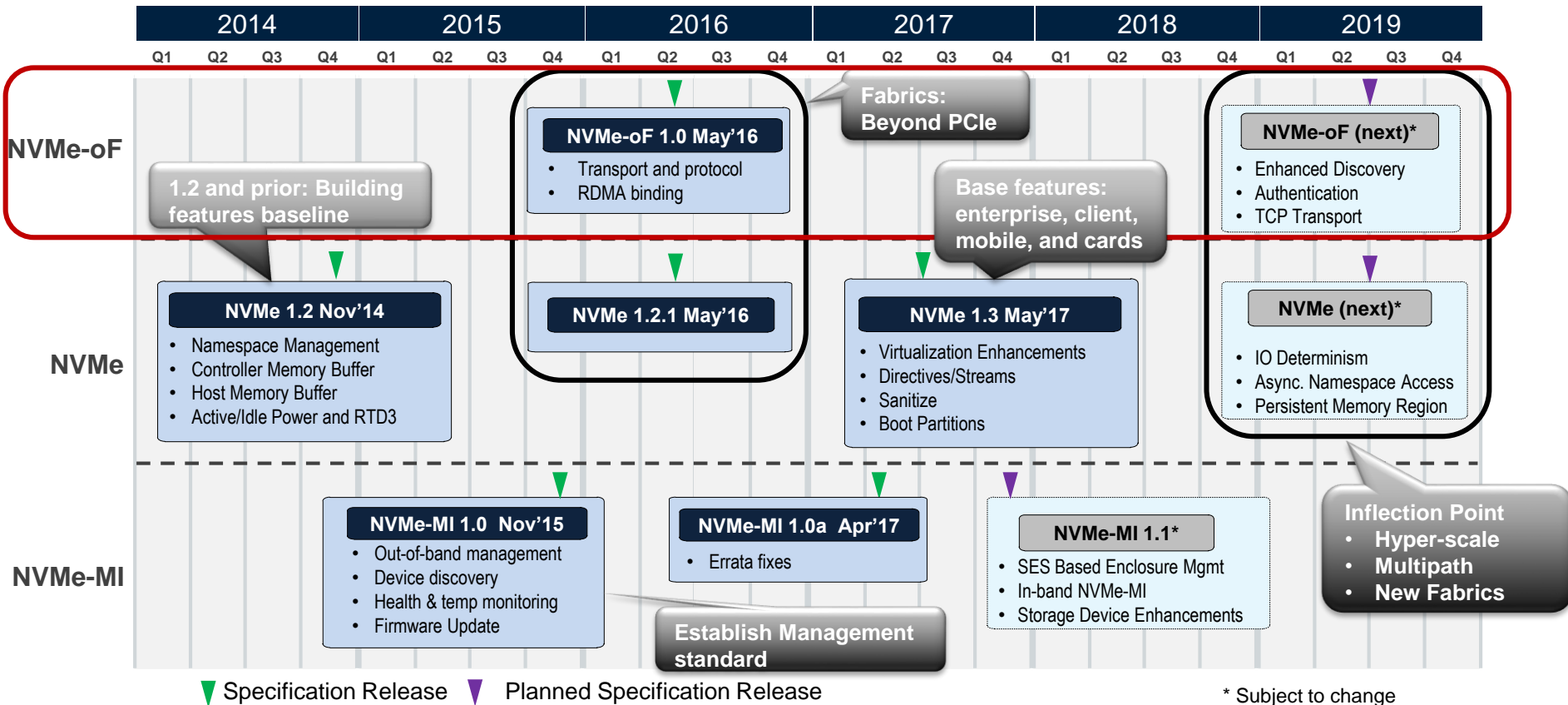
# What Drives AFA Purchases? (Top 5)

From the list below, please, select up to three most important criteria when purchasing/considering AFA



Source: IDC, All-Flash Array Adoption

# NVMe Roadmap

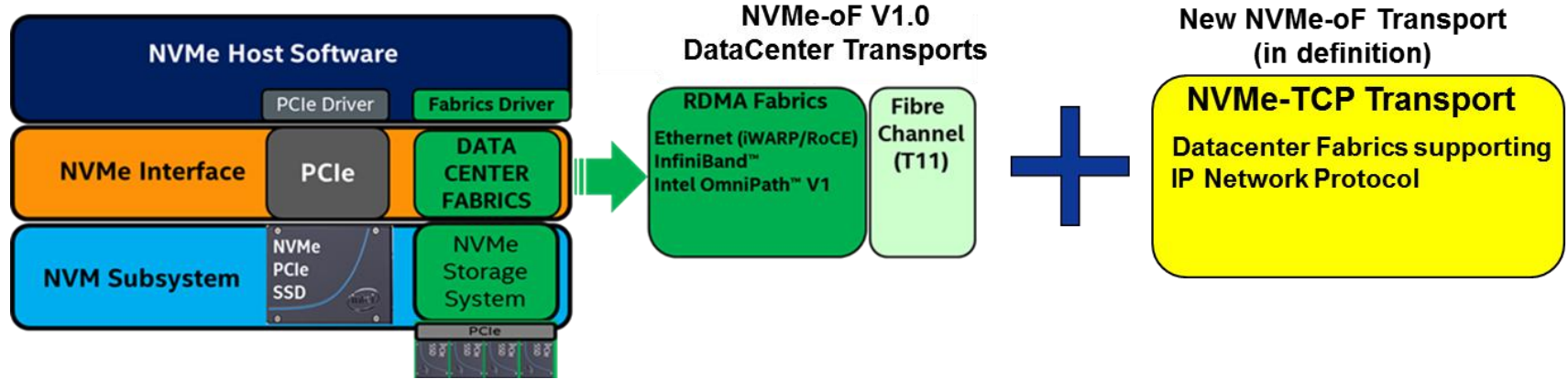


# NVMe-oF coming to a network near you

- NVMe-oF V1.0 enabled efficient end-2-end NVMe over RDMA and Fibre Channel networks
  - RDMA because of it's high efficiency and similar architecture characteristics
  - FC because of it's reliable credit based flow control and delivery mechanism
- What about existing IP network infrastructures?

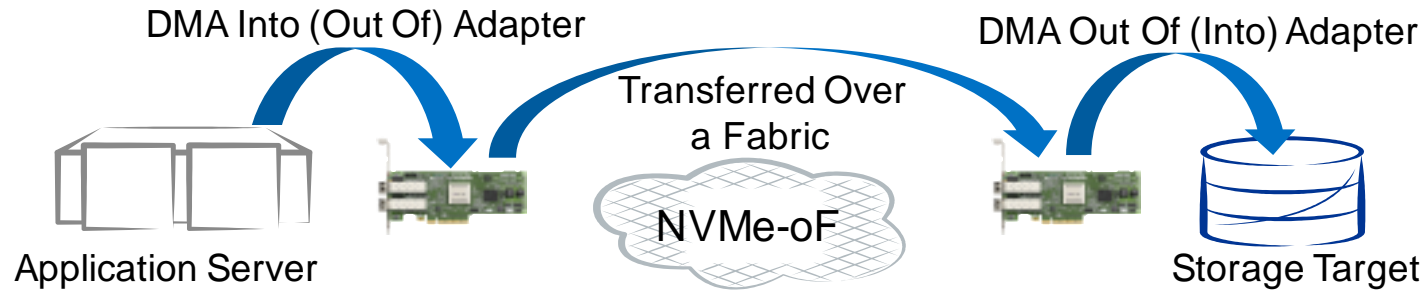


# NVMe-TCP Transport



- Enables the use of NVMe-oF over existing Datacenter IP networks
- Supports all of the NVMe-oF and NVMe Architecture features
- Layered over standard IETF TCP transport to allow software-only and/or hardware (accelerated/offloaded) implementations

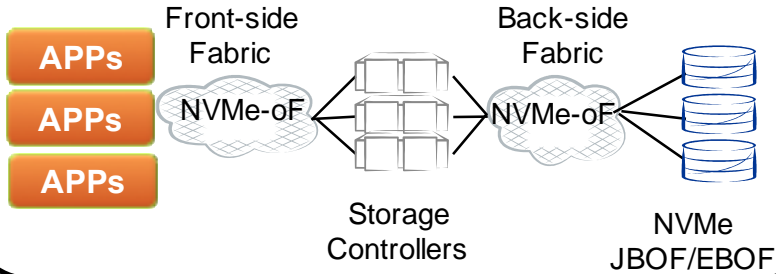
# Simplicity of NVMe over Fabrics



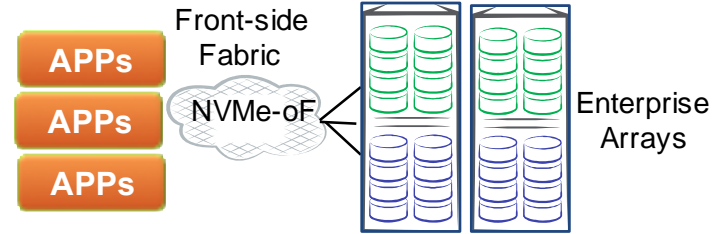
- NVMe-oF delivers a new level of performance for today's business-critical applications
- NVMe-oF is, by design, is transport agnostic:
  - Application developers can write to a single block storage stack and access NVMe over Fibre Channel, TCP, or RDMA networks
- Data is DMA'd in and out of the adapters to maximize performance
  - Zero copy is available **today** for Fibre Channel and RDMA protocols for improved performance and the new TLDK (new) for TCP can provide zero copy for TCP

# NVMe over Fabrics – Storage Architectures

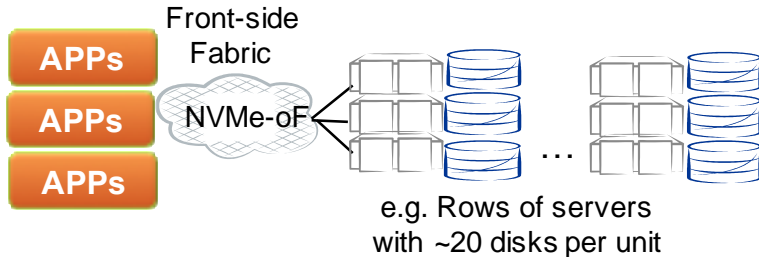
## End to NVMe and NVMe-oF Solutions



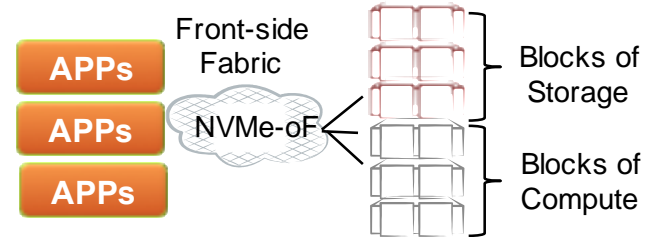
## Traditional SAN



## Server SAN/Disaggregated Storage



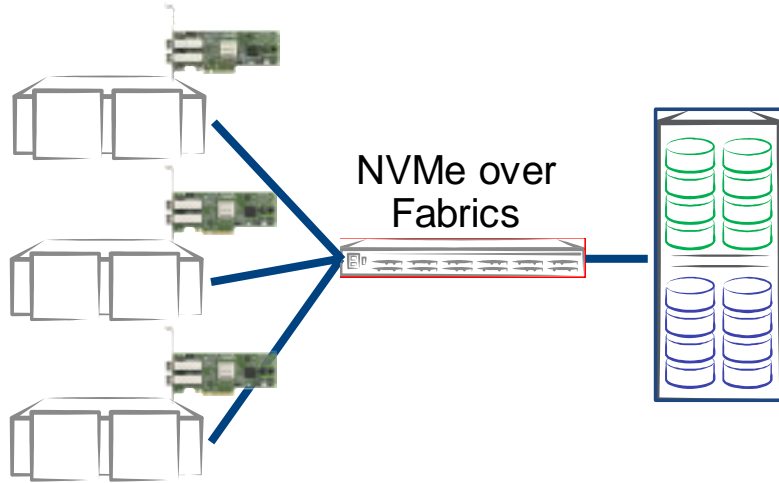
## Rack Scale/Scaleout/HyperScale



# Evolving the Fabric for a New Generation of Enterprise Applications

## Use Case 1: Faster Flash

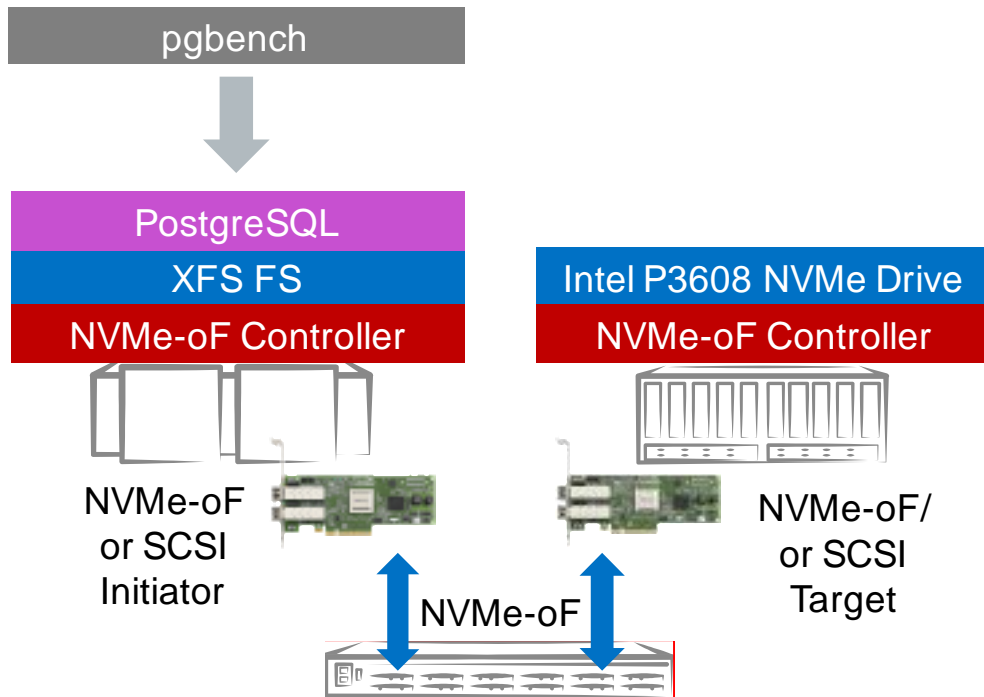
### Accelerated Enterprise Data Fabric



### Faster Flash

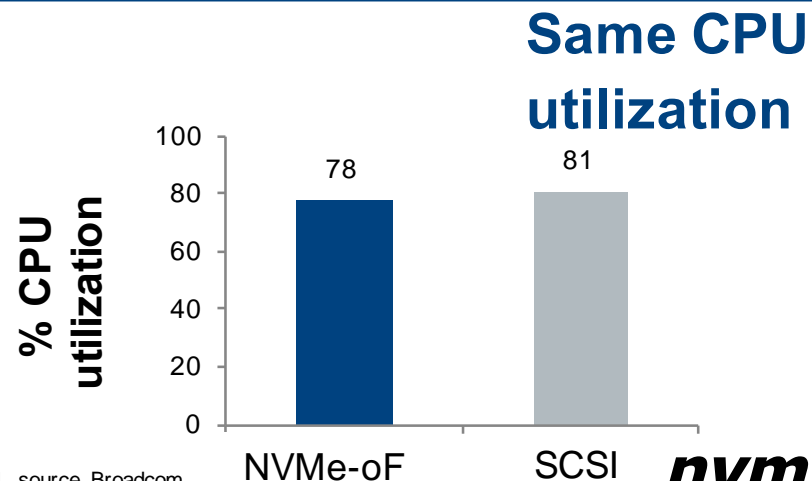
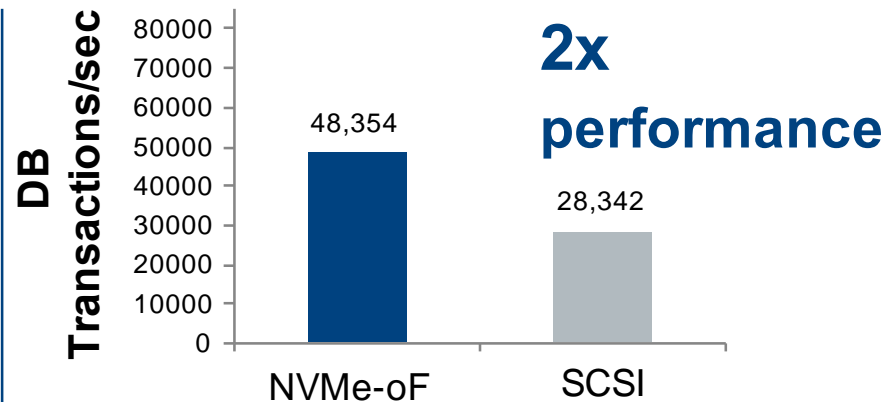
- Lower latency for NVMe-oF versus legacy SCSI
- 25GE/32GFC 100GE/128GFC Bandwidth to support 32G (PCIe) and faster NVMe drives
- Lower CPU Utilization
- Scales to 100's and 1000's of drives making more capacity available to applications

# Benefits of NVMe over Fabrics for SQL Workloads



Test Setup

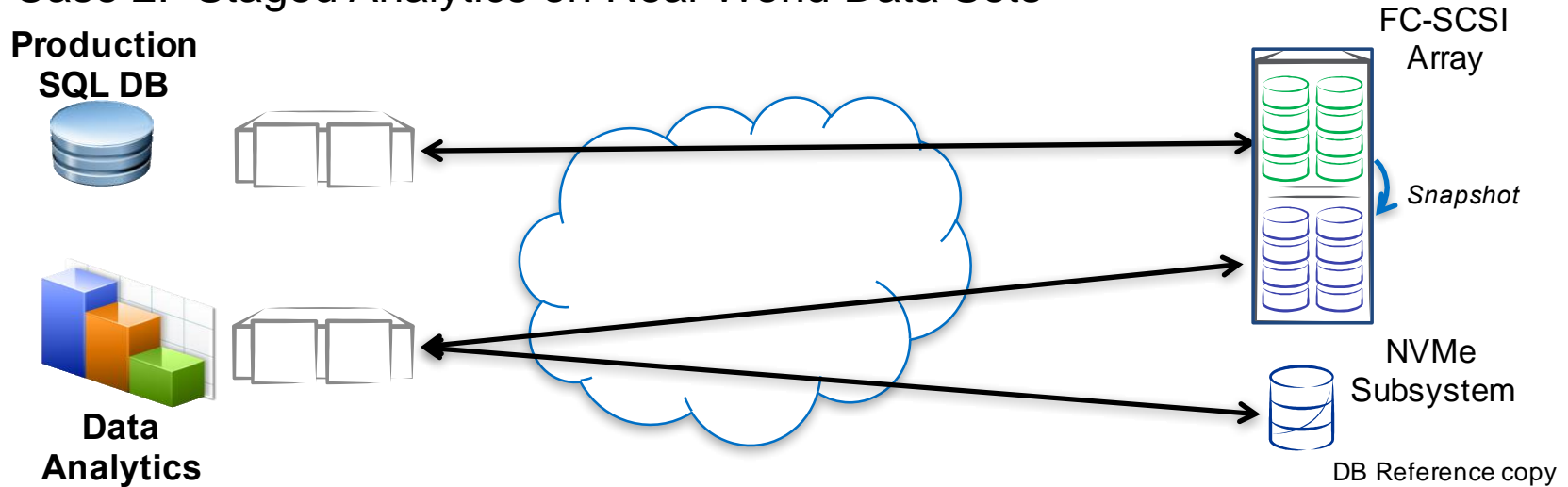
- Standard Servers
- SLES 12 SP3



Pgbench Read only Transaction workload, source Broadcom

# NVMe over Fabrics Concurrent IO Use Case

## Use Case 2: Staged Analytics on Real-World Data Sets



- 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 Analytics server to copy snapshot to Ultra-low latency NVMe-oF copy
- Run Data Analytics application using low latency NVMe-oF reference copy

# Integrating New Storage Technologies into the Datacenter

- Storage Class Memory (SCM) as a Cache
  - Adding Flash into a DIM to support low latency storage for NVMe over Fabrics
- Data Management for In-Memory Databases
  - Adding data management for business-critical data
  - Replication, deduplication, Tiering data off, etc.

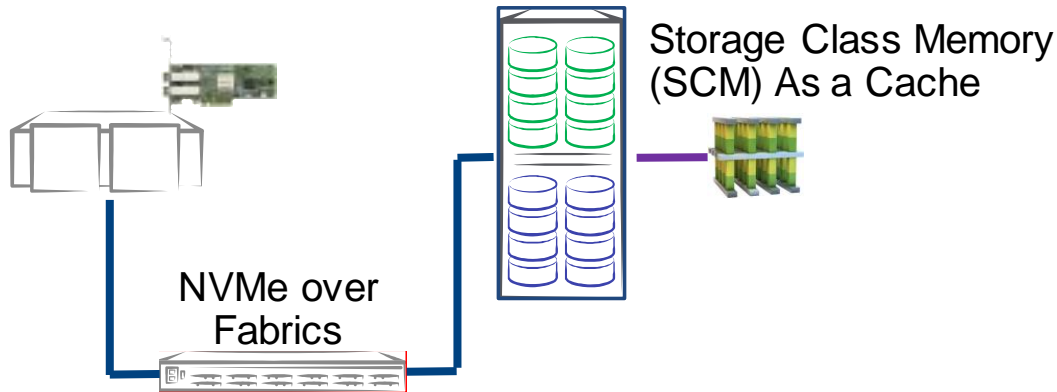
**“Storage-class memory technologies could alter price/performance and density of current storage and compute platforms while bringing both benefits and uncertainties to I&O leaders responsible for IT infrastructure planning,”**

**Gartner**<sup>®</sup>

# Creating the Fabric for a New Generation of Enterprise Applications

## Use Case 3: **Storage Class Memory as a storage array cache**

### New Media as Cache



### SCM as Cache

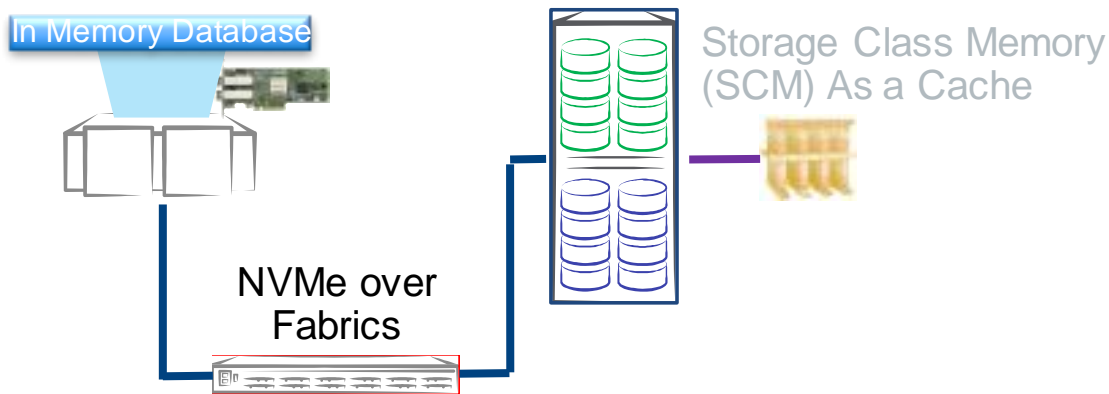
- Improved performance: 4x more IOPs
  - Leverage 32G NVMe
  - Benefit from low latency media
- NVMe over Fabrics fits well
- Highest performance, highest bandwidth, lowest latency, storage target available
  - Removes PCIe latency in storage for caching/fast storage



# Creating the Fabric for a New Generation of Enterprise Applications

## Use Case 4: Data Management for in-memory databases

### Enterprise Data Management



### Enterprise Data Management everywhere

- New Storage Tier: In-Memory DBs Databases
- NVMe over Fabrics enables
  - Snapshots
  - Data Tiering (e.g. key for HANA)
  - Availability
  - Workload Migration
- Eliminates another silo in the datacenter
- Eliminates stranded storage

# Resources from NVM Express and FMS

NVMe over Fabrics 2016 Webinar:

[https://www.brighttalk.com/webcast/12367/225125?utm\\_campaign=webcasts-search-results-feed&utm\\_content=NVMe%20over%20Fabrics&utm\\_source=brighttalk-portal&utm\\_medium=web](https://www.brighttalk.com/webcast/12367/225125?utm_campaign=webcasts-search-results-feed&utm_content=NVMe%20over%20Fabrics&utm_source=brighttalk-portal&utm_medium=web)

NVMe over Fabrics 2017 FMS Presentation(s):

<http://www.nvmexpress.org/nvme-content-from-fms-2017/>

NVMe over Fabrics Standard:

<http://www.nvmexpress.org/resources/specifications/>

NetApp's Presentation on SCM and PMEM from FMS:

<https://www.youtube.com/watch?v=1fQIJusiVdE>

