



# Using a PCIe-Based Switch Module to Enhance Enterprise Storage Architecture

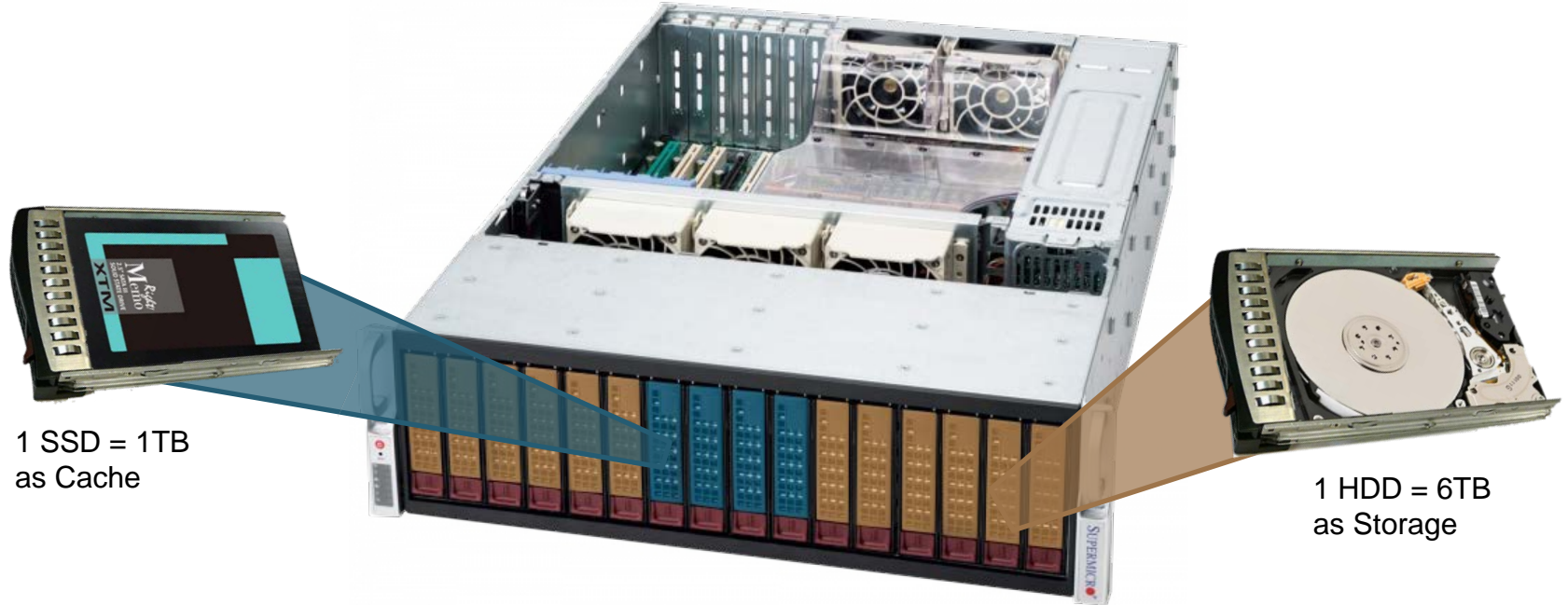
Chi-Lin Tom

Memoright

[chilintom@memoright-usa.com](mailto:chilintom@memoright-usa.com)

# Legacy Storage Solution Architecture

2U Server – 16 Slots  
12 HDD & 4 SSD



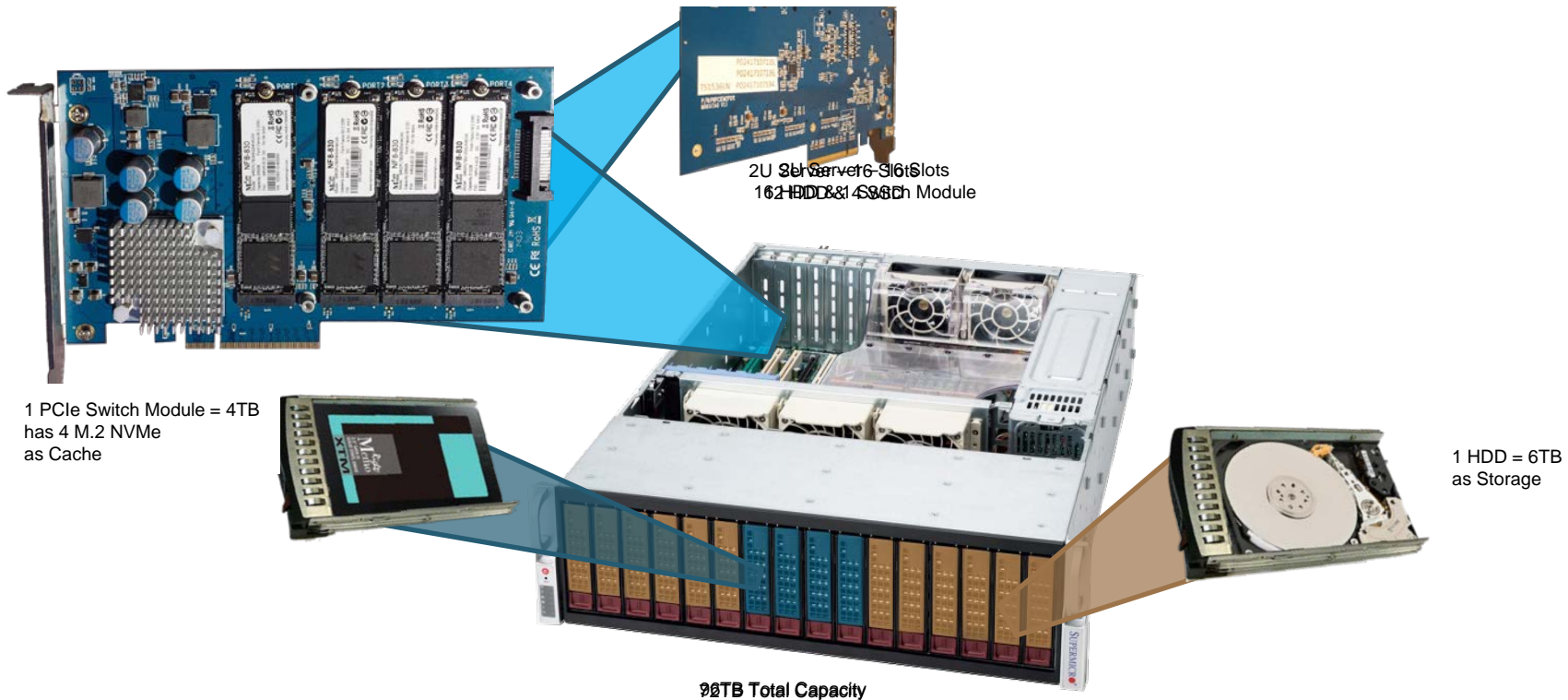
1 SSD = 1TB  
as Cache

1 HDD = 6TB  
as Storage

# Drawbacks of Legacy Storage Solution

- HBA (Host Bus Adaptor) quickly becomes a bottleneck and contention point.
- Some HDD slots (typically 4) are reserved for SSD as cache and the storage space is reduced.

# Legacy Storage Solution Architecture



# Advantages of Contemporary Storage Solution

- Increase storage space
- Reduce latency
- Improve performance
- Switching is full duplex in parallel with SAS HBA traffic
- Affinity of each M.2 to one or many NUMA CPU cores
- 4 M.2 adds flexibility of resource aggregation
- Provide PCIe pass through to VMs
- No driver needed



# Why M.2 PCIe NVMe?

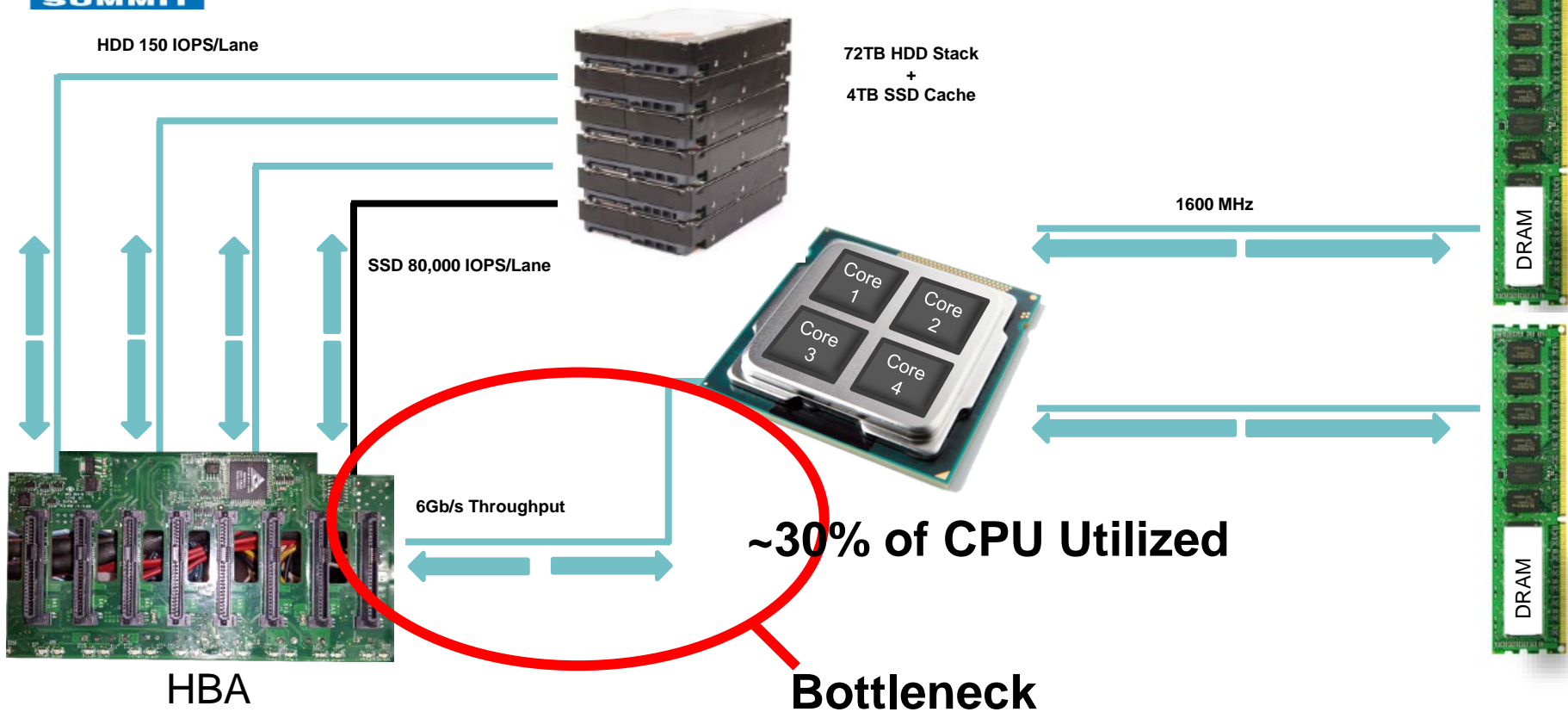
- Higher Bandwidth
  - SATA III SSD: 6Gb
  - Gen 2 x 2: 8Gb
  - Gen 3 x 2: ~16Gb
  - Gen 3 x 4: ~32Gb
- Low latency
- Native PCIe instead of AHCI Mode



# AHCI vs. NVMe

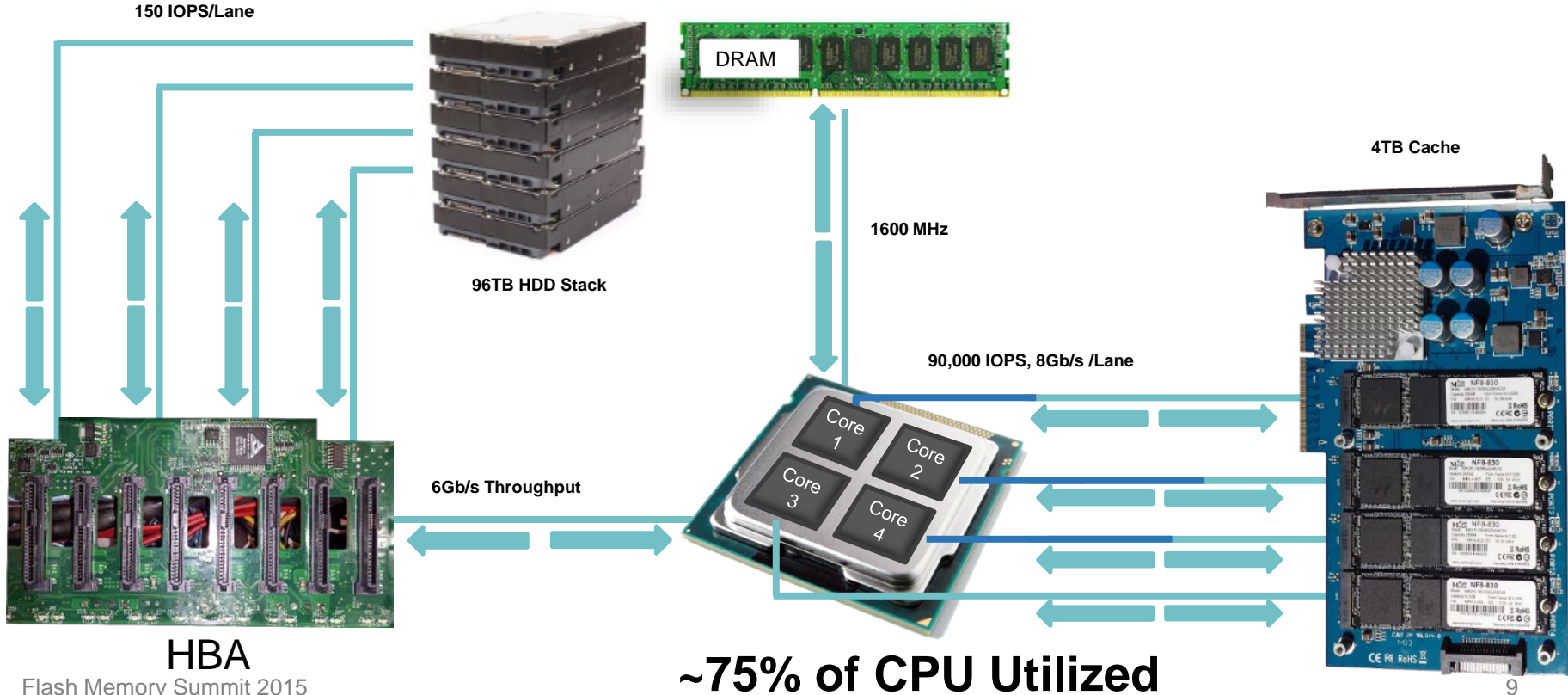
	AHCI	NVMe
Maximum Queue Depth	1 command queue; 32 commands per queue	65536 queues; 65536 commands per queue
Uncacheable register accesses (2000 cycles each)	6 per non-queued command; 9 per queued command	2 per command
MSI-X and interrupt steering	single interrupt; no steering	2048 MSI-X interrupts
Parallelism and multiple threads	requires synchronization lock to issue a command	no locking
Efficiency for 4 KB commands	command parameters require two serialized host DRAM fetches	gets command parameters in one 64 Bytes fetch

# Legacy Solution CPU Utilization





# Contemporary Solution CPU Utilization



# Storage CPU Usage

- Virtualization software often is starved for IOPs and PCIe pass-through can improve the CPU utilization from the average 30% to the desired 75%.

# Applications

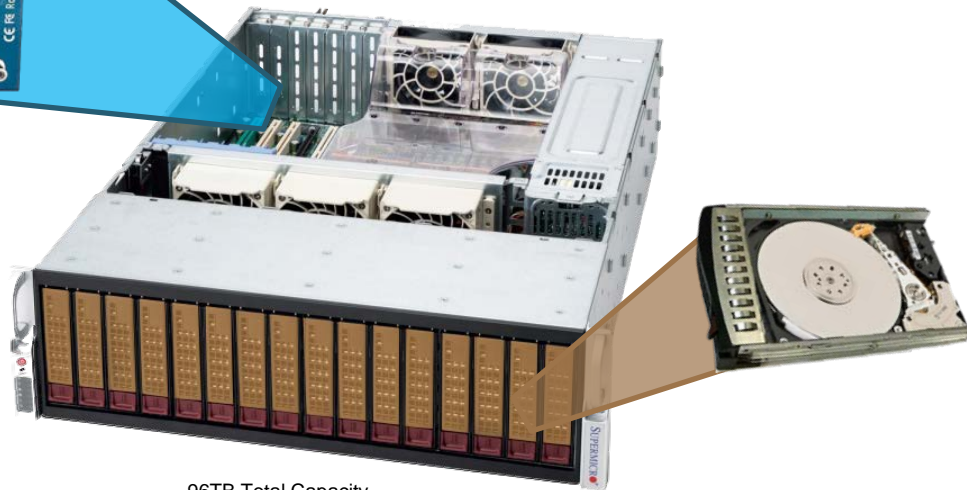
- Storage Appliance
- EDA
- NAS Gateway
- Others

# Application: Storage Appliance



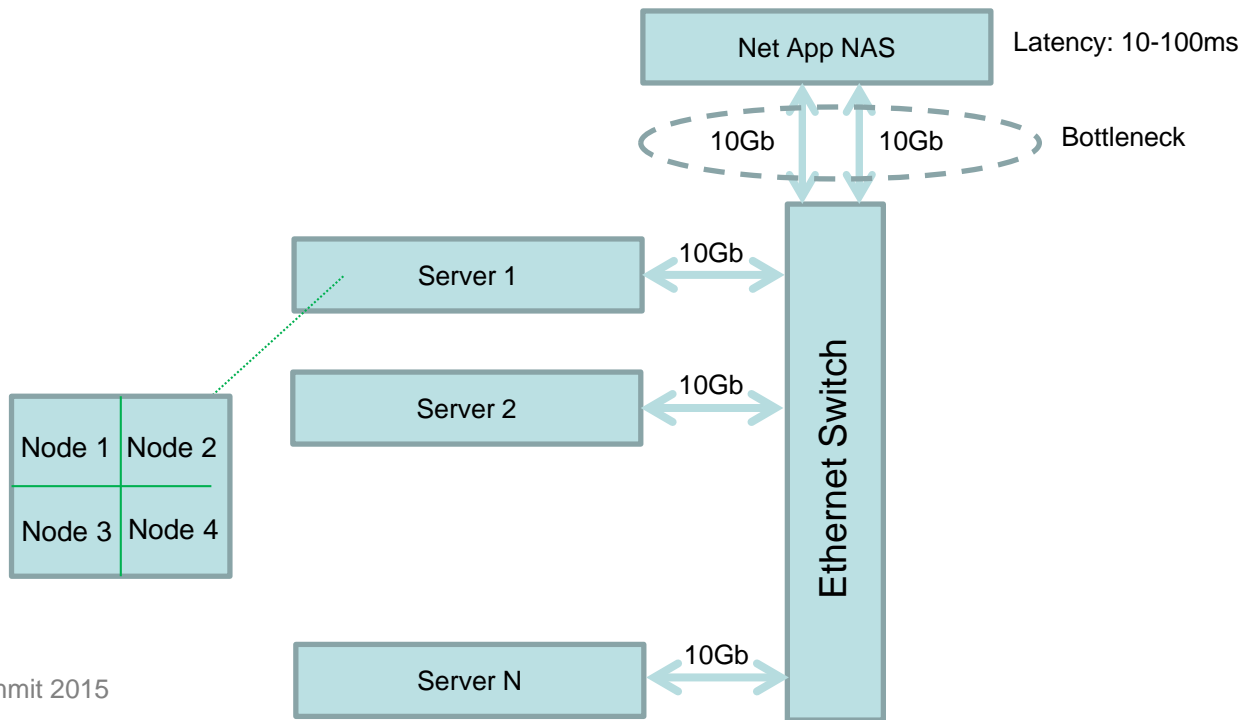
1 PCIe Switch Module = 4TB  
has 4 M.2 NVMe  
as Cache

2U Server – 16 Slots  
16 HDD & 1 Switch Module

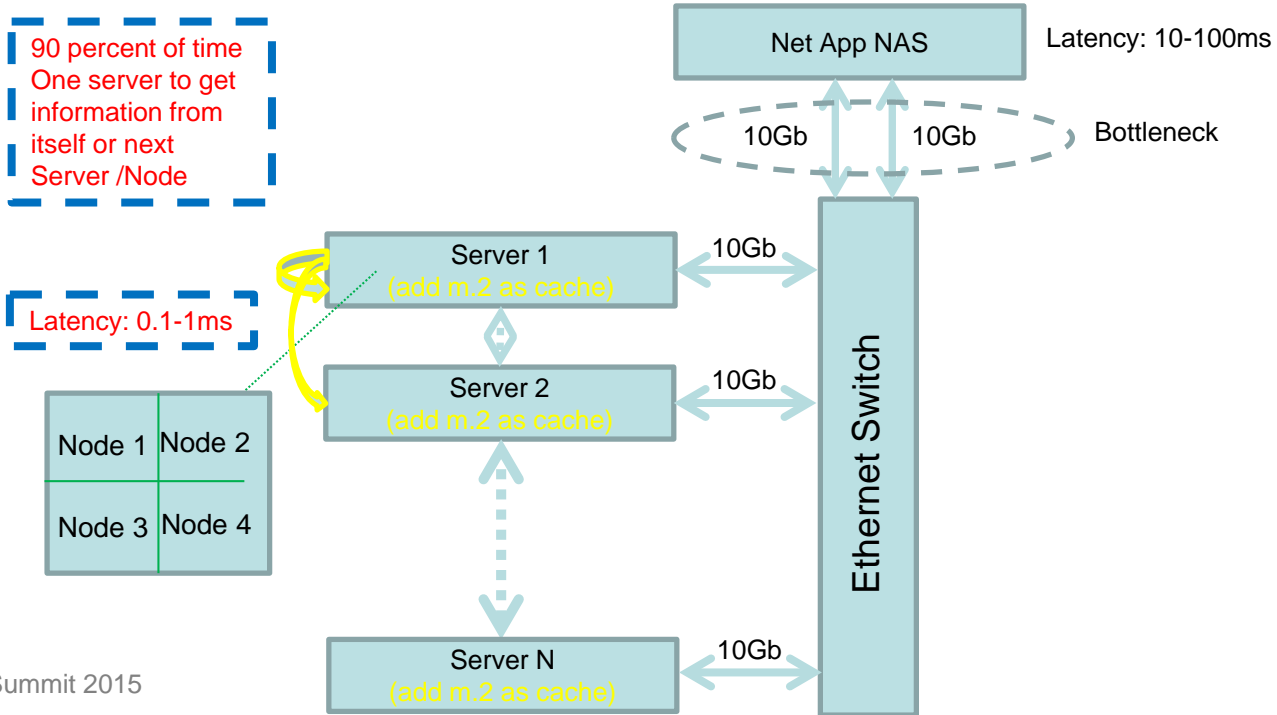


1 HDD = 6TB  
as Storage

# Applications: Electronic Design Automation (current)



# Applications: Electronic Design Automation (new)



# Application: NAS Gateway

<b>Access Protocols</b>	NFS	CIFS		
<b>Management</b>	Single Point of Administration	User Defined Policies		
<b>Availability</b>	N-Way Clustering	QoS		
<b>Data Services</b>	Compression	Deduplication	Replication	Encryption
<b>Caching</b>	In-Switch Metadata	DRAM Cache	NVMe Switched Fabric	
<b>SAN Integration</b>	iSCSI	FCoE	AoE	





# Thank You!