

Dell EMC Ready Solutions for HPC Lustre Storage

Forrest Ling

**HPC Enterprise Technologist at Dell EMC
Greater China**

2018.10.23

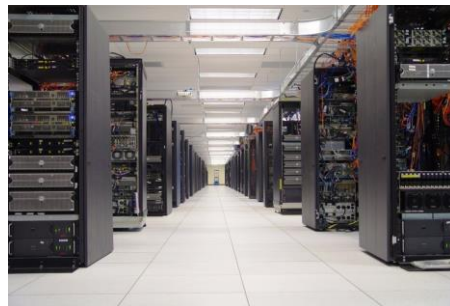


Dell EMC Supports HPC Open Source Software

Support Open Source Software projects



Prove Open Source Software in Dell EMC HPC Labs



Organize HPC Open Source Software Forums



Provide HPC Platforms running Open Source Software

Dell (China) Awarded “HPC Open Source Software Technical Practice Innovation”

THANKS to :

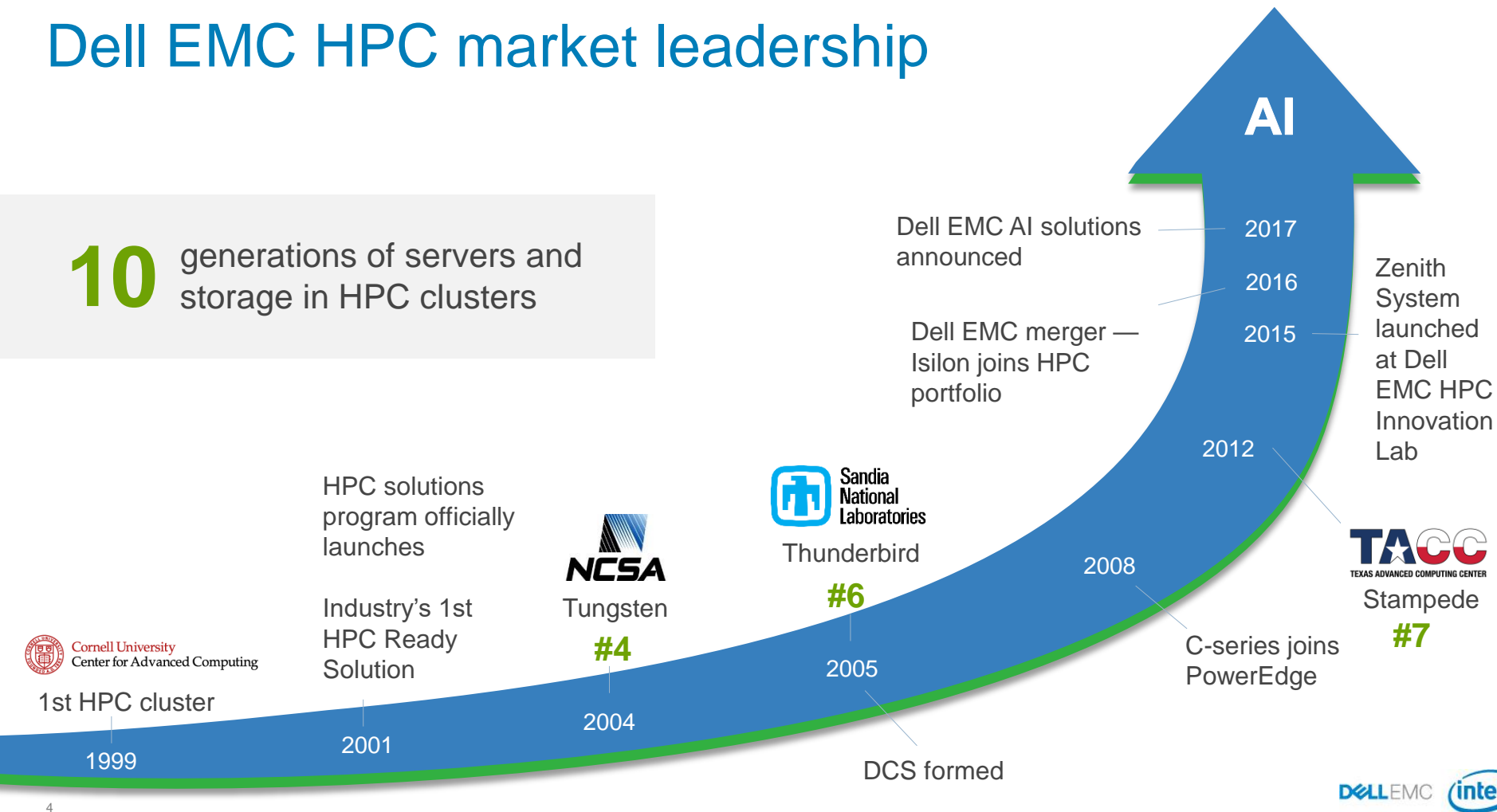
- CCF TCHPC*
- OpenHPC
- OpenSFS
- OpenPBSpro
- Singularity
- ...

CCF TCHPC – China Computer Federation, Technical Committee of HPC



Dell EMC HPC market leadership

10 generations of servers and storage in HPC clusters



PowerEdge HPC Server portfolio



C6420

Maximizes density, scalability, and energy efficiency per U for high performance hyperscale workloads

C6320p: Intel® Xeon Phi™ Processor (KNL)



C4140

2 socket, ultra-dense, 4 GPU rack server



R640/R440

Ideal combination for dense scale out data center computing and storage in a 1U/2S platform



R740, R740xd

Ideal for applications requiring best-in-class storage performance, high scalability, and density. Support up to 3 double-wide GPUs.



R840, R940, R940xa

Ideal for mission-critical applications and real-time data and analytics

R940xa: Support up to 4 double-wide GPUs or up to 8 FPGAs.



MX7000

Dense compute and optimal memory throughput for demanding HPC workloads

Purpose built for HPC

Infrastructure and I/O

Large memory

Modular

powerful performance | density | efficiency

Additional options: [DellEMC.com/servers](https://dell EMC.com/servers)

AMD processor-based PowerEdge HPC Server portfolio



R6415

Storage flexibility with up to 10 PCIe NVMe drives. Up to 32 cores in a single socket. 16 DIMMs bringing up to 2TB of total memory capacity. 128 lanes of PCIe 3.0 connectivity



R7415

Up to 24 flash drives, 32 total cores and 16 DDR4 DIMMs.



R7425

Up to 30% superior HPC performance. 4x more NVMe for extreme performance storage requirements. Three double wide GPUs, plus up to 8 NVMe drives. Up to 4TB memory capacity optimized.

IO/Storage nodes

Memory bandwidth-intensive compute

powerful performance | density | efficiency

Additional options: [DellEMC.com/servers](https://www.dell.com/emc/servers)

HPC Storage portfolio

GPFS

Available through DDN and Arcstream partnerships

Lustre

Building in-house support capability in partnership with Intel; high-end support in partnership with DDN and ClusterStor

BeeGFS, Ceph and others

PowerVault and PowerEdge platforms can power a variety of storage capabilities



HPC NFS Storage

High availability storage system with up to 480TB of raw storage capacity
Red Hat HA software stack



HPC Lustre Storage

Lustre storage starting from 960TB raw storage per object storage server pair and 14.9GB/s of write and 16.8GB/s of read throughput. Scale out performance and capacity with additional OSS pairs



Isilon

Scale-out NAS storage to store, protect and analyze unstructured data



Elastic Cloud Storage (ECS)

All the benefits of a public cloud while keeping cost under control

File system

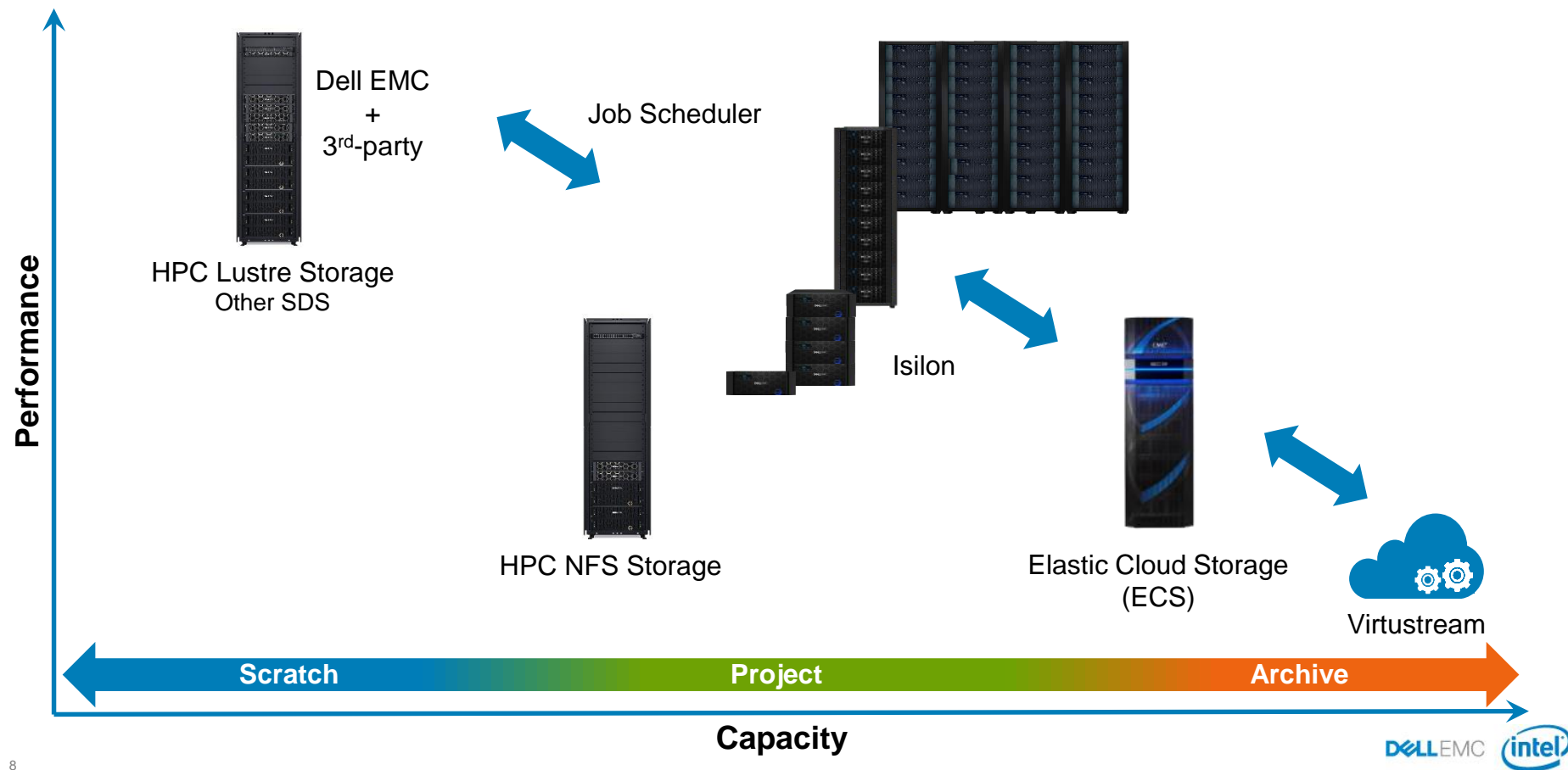
High performance and flash

Archival

powerful performance | efficiency | scalability

Additional options: http://i.dell.com/sites/doccontent/shared-content/data-sheets/en/Documents/Dell_Storage_Family_Portfolio.pdf

HPC Storage comparison



HPC fabric portfolio

Dell EMC Ethernet

- Utilized for base management and/or primary interconnect
- Options range from 1GbE to 100 GbE
- Low complexity



- Smaller lower cost HPC solutions
- In-house expertise for Ethernet management

Mellanox InfiniBand

- High speed, low latency fabric for optimized communication
- Established HPC fabric
- Virtual Protocol Interconnect (VPI) offers flexibility for converged fabric



- Proven high performance fabric
- Connect to existing InfiniBand fabric for HPC storage or compute

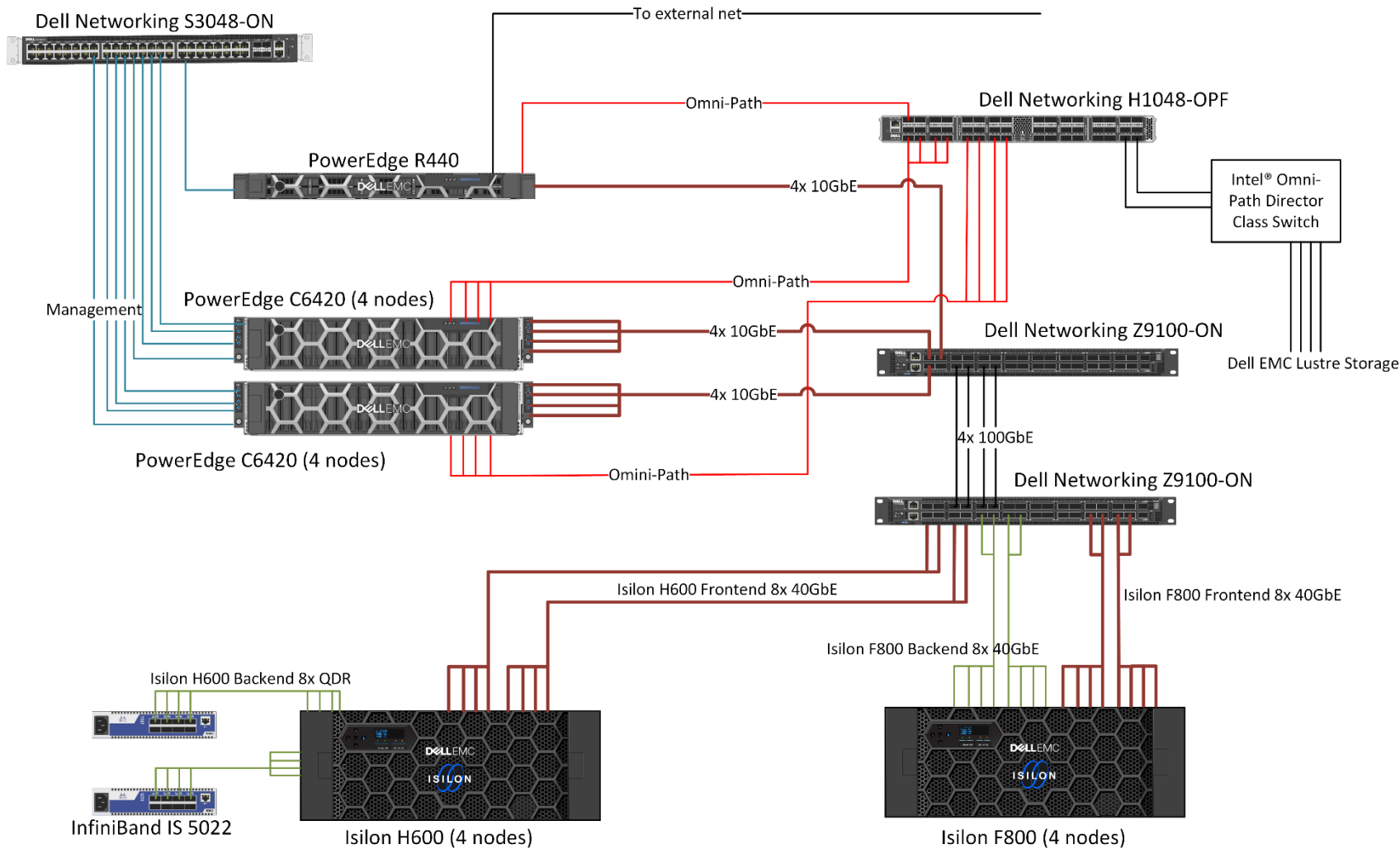
Dell EMC H-Series

- Designed specifically for HPC
- Higher switch chip port count for added density
- CPU-fabric integration

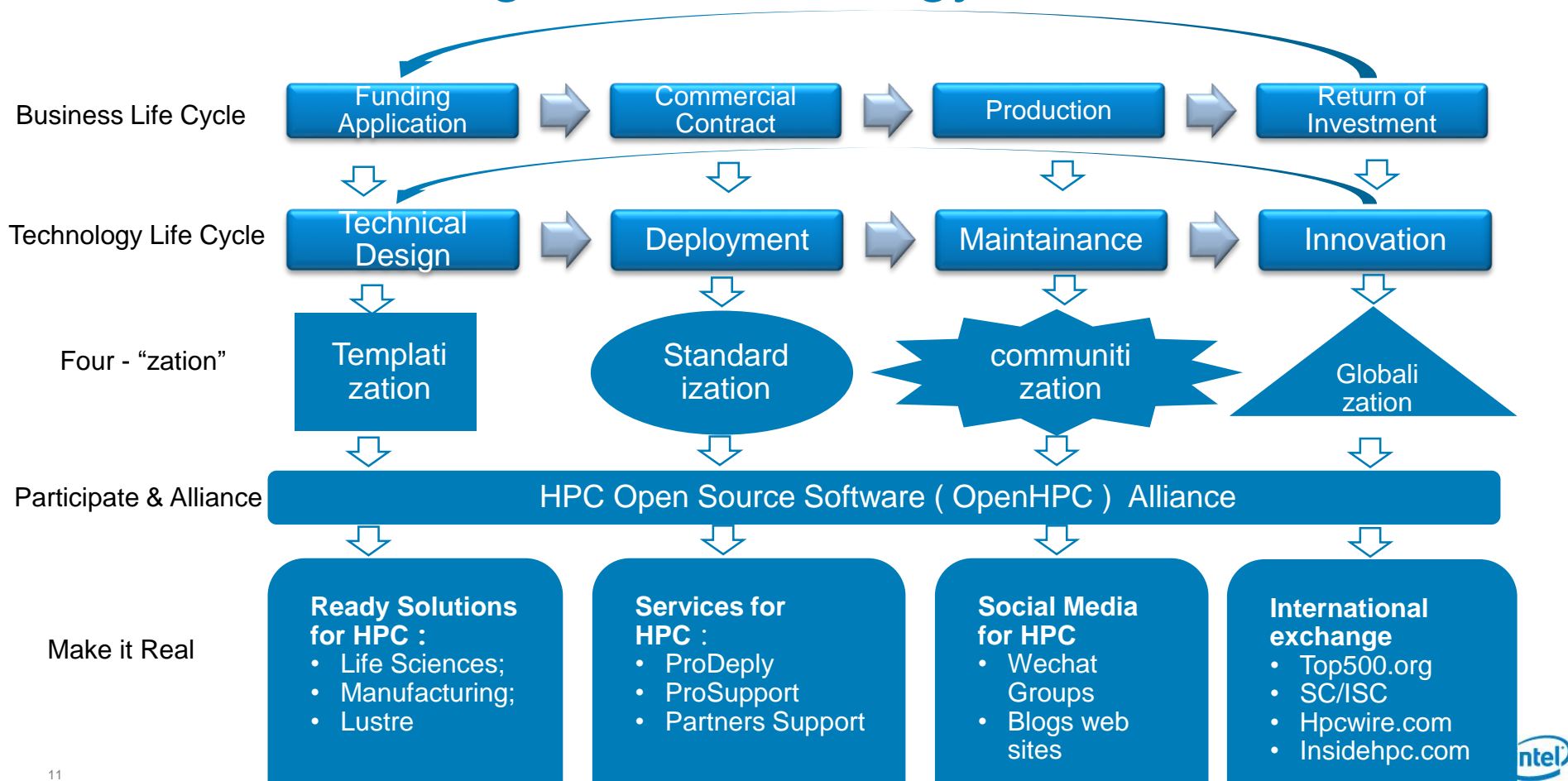


- Maximizes FLOPs/\$ keeping costs in balance
- Fewer switch hops in fat-tree

options included in all Dell EMC Ready Solutions for HPC or custom designed opportunities



Dell EMC “4 Stages” Methodology Grow HPC in China



Simplified design, configuration and ordering

Dell EMC has developed several modular Ready Solutions for HPC



Dell EMC
Ready Solution
for HPC



Dell EMC
Ready Solution
for HPC
Life Sciences



Dell EMC
Ready Solution
for HPC Digital
Manufacturing



Dell EMC
Ready Solution
for HPC
Research



Dell EMC
Ready Solution
for HPC NFS
Storage

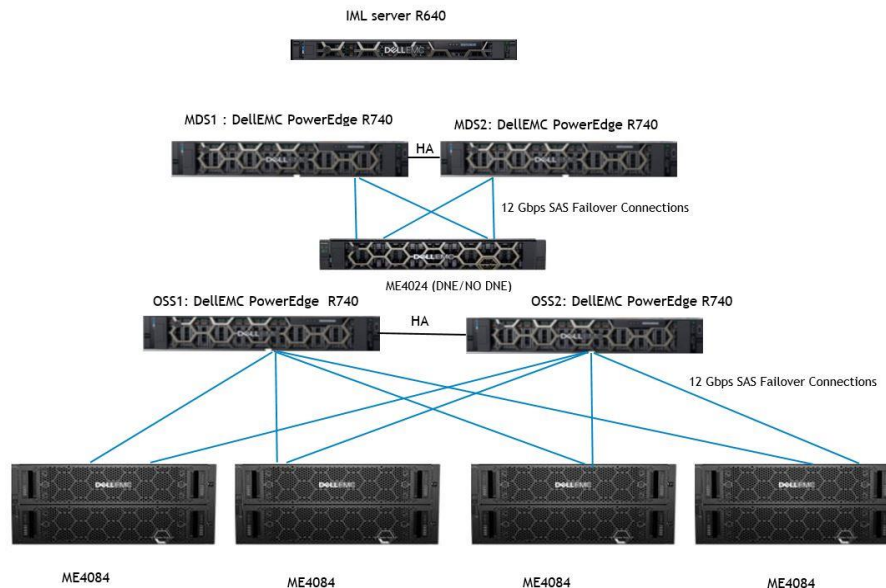


Dell EMC
Ready Solution
for HPC Lustre®
Storage

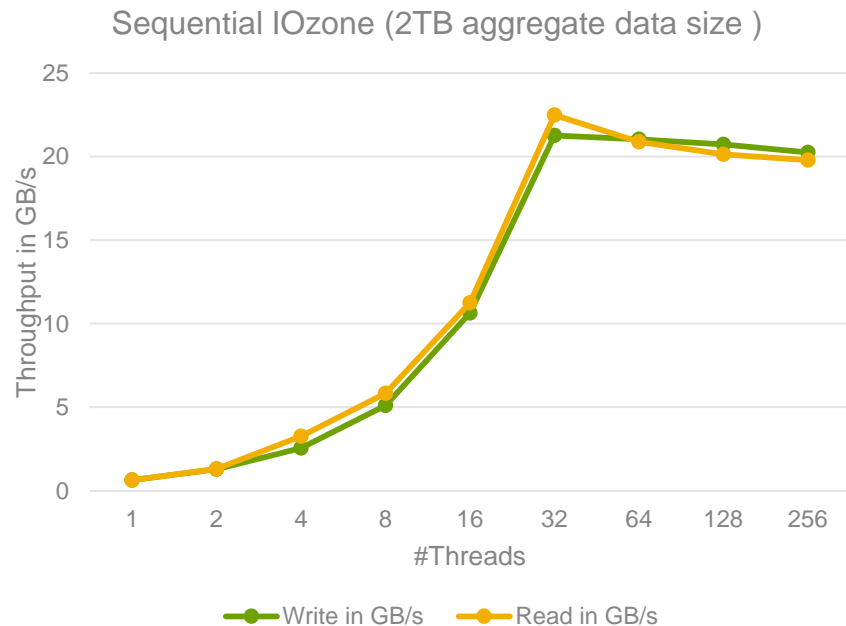


Upcoming Lustre solution – PowerVault ME storage

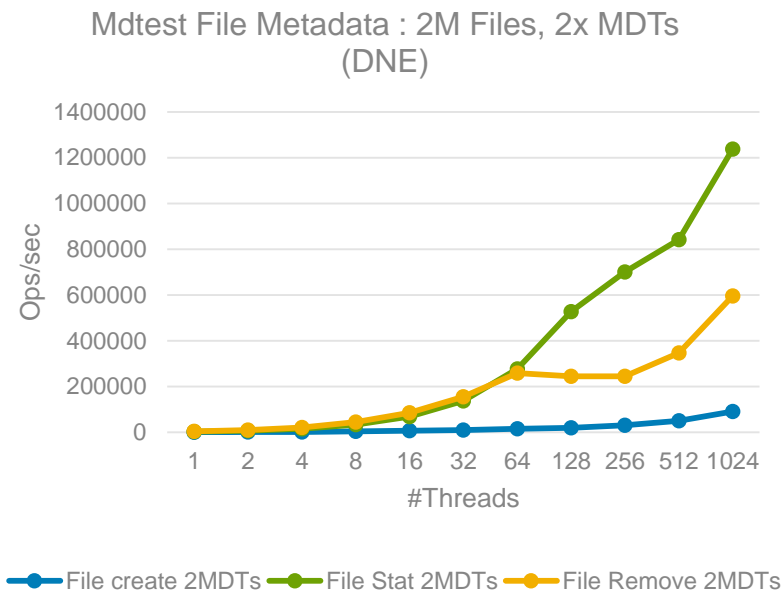
Server configuration	
IML, MDS and OSS server model	One DellEMC PowerEdge R640 and Four DellEMC PowerEdge R740.
Processor	IML Server: Dual Intel Xeon Gold 5118 @ 2.3GHz MDS and OSS Servers: Dual Intel Xeon™ Gold 6136 @ 3.00GHz
Memory	IML Server: 12 x 8GB 2666MT/s DDR4 RDIMMs. MDS and OSS Servers: 24 x 16GiB 2666MT/s DDR4 RDIMMs.
InfiniBand HCA (slot 8)	Mellanox ConnectX-5 EDR PCIe adapter.
External storage controller (slot 1,2,4 and 5)	Four Dell 12Gbps SAS HBAs.
Storage configuration	
Object Storage Enclosure	Four DellEMC PowerVault ME4084 enclosures fully populated for 336 drives. 2.69PB raw capacity solution if equipped with 8TB SAS drives.
Metadata Storage Enclosure	DellEMC PowerVault ME4024 enclosure fully populated with 24 x 960GB SAS SSD
RAID controllers	Duplex RAID controllers in the DellEMC ME4084 enclosure. Duplex RAID controllers in the DellEMC ME4024 enclosure.
Hard Disk Drives	84 - 8TB 7200 rpm NL SAS drives per ME4084 enclosure.(ME4 supports up to 12TB HDDs) 24 – 960GB SAS SSDs per ME4024 enclosure.



Initial Performance Data



Sequential IOzone 2TB Aggregate Data Size



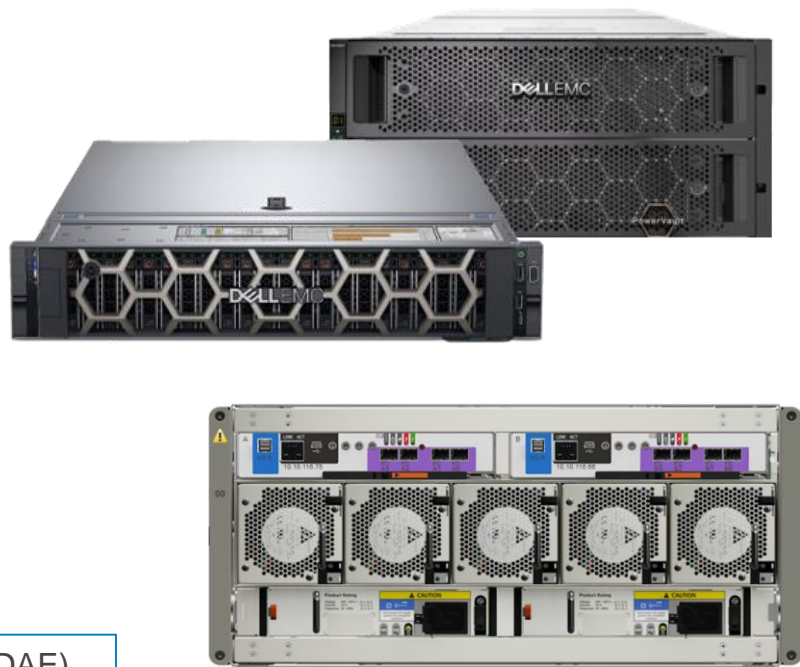
File MDTest 2M files, 2 MDTs in DNE

Upcoming Lustre solution – PowerVault ME storage

Category	Improvement over previous generation	Lustre with ME4	Lustre with MD3
Number of hard drives in base config	40%	336 drives	240 drives
Capacity	43%	Up to 4PB	Up to 2.8 PB
Peak Sequential Writes	42%	21.3 GB/s	15 GB/s
Peak Sequential Reads	25%	22.5 GB/s	18 GB/s
Peak File create metadata	107%	≈ 60K op/s	≈ 29K op/s
Peak File remove metadata	50%	≈ 240K op/s	≈ 160K op/s
Peak File stat metadata	23%	≈ 669K op/s	≈ 544K op/s

ME4012: 12-drive RBOD (DPE)
ME4024: 24-drive RBOD (DPE)
ME4084: 84-drive RBOD (DPE)

ME412: 12-drive Expansion (DAE)
ME424: 24-drive Expansion (DAE)
ME484: 84-drive Expansion (DAE)



HPC Solution Support and Deployment Services



Deployment

ProDeploy
for HPC



1

**HPC Add-on:
Individual
nodes**

2

**HPC Add-on:
M1000e**

3

**HPC Add-on:
Storage**



Support

Asset-level support

ProSupport

or

**ProSupport
Plus**



Solution support

**ProSupport
Add-on
for HPC**

Supported Hardware and Software Technology and Local Services Partners

Cluster Management



Operating System



Networking



Storage



Server



Local Partners





Thanks!