

Tiered Storage

Steve Lyness

VP, Solutions Architecture



CRAY

HPC Users Forum, Sept 11, 2019

CRAY
SLINGSHOT™

HPC STORAGE SYSTEM PARADIGM SHIFT

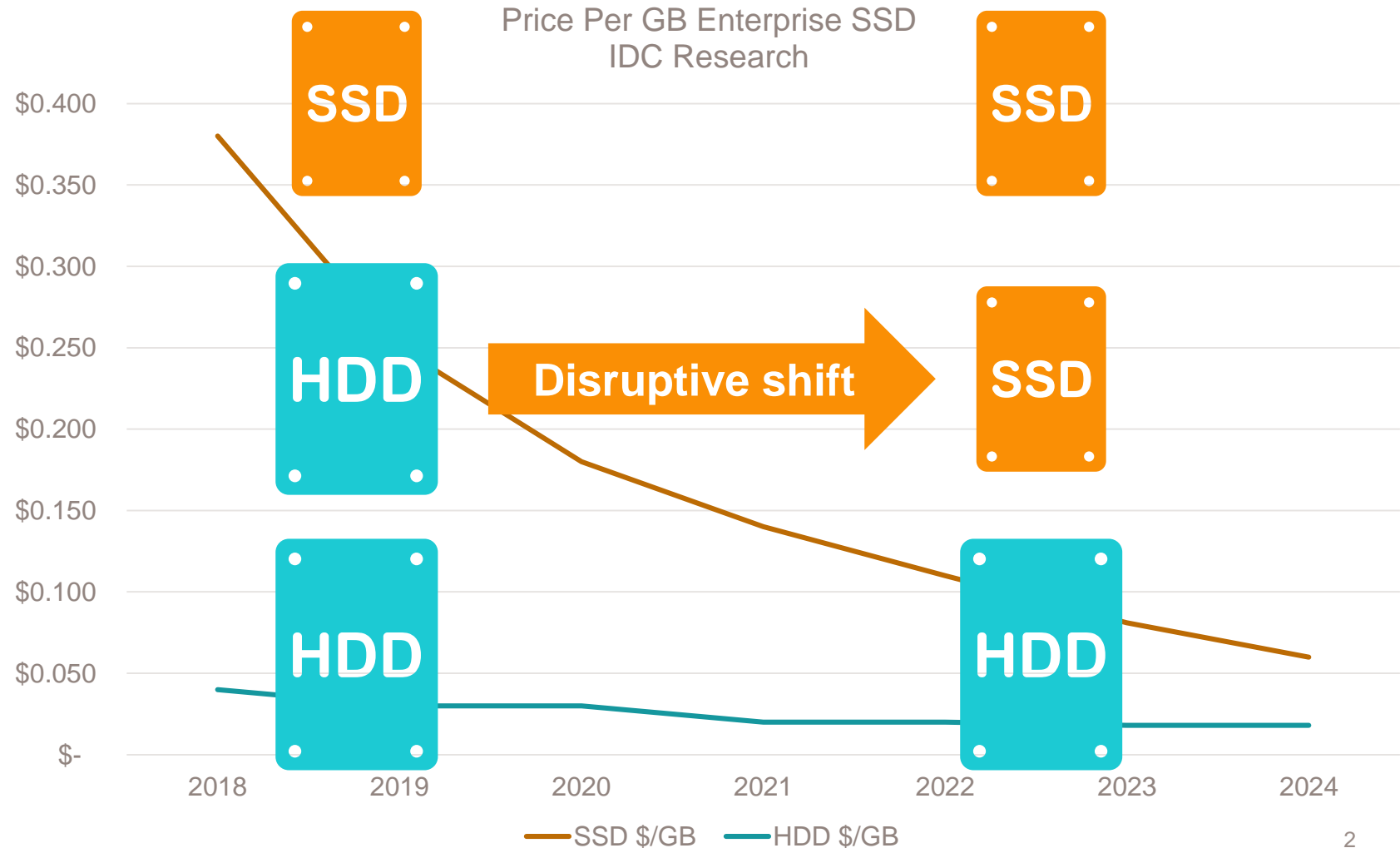
TODAY

EARLY 2020

Most cost-effective to deliver
IOPS
(IO operations/sec)

Most cost-effective to deliver
THROUGHPUT
(GB/sec)

Most cost-effective to deliver
CAPACITY
(Petabyte)

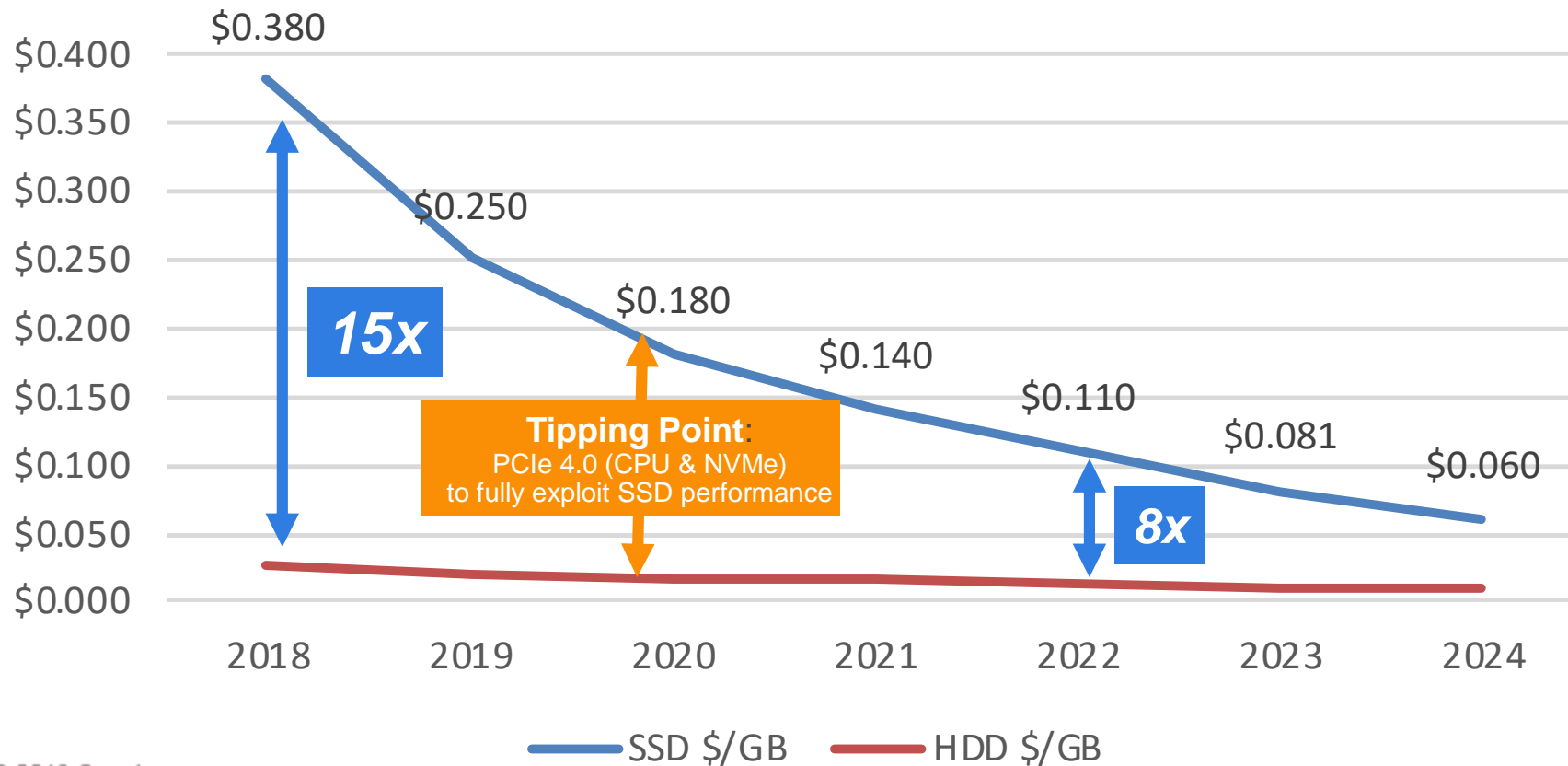


CHANGING ECONOMIES OF STORAGE MEDIA



> 60% of the cost of any storage system is in the media

\$\$ per GB - Enterprise drives
IDC

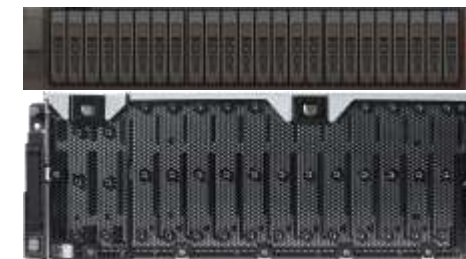


SSD Media =
PERFORMANCE
+ some initial capacity

HDD Media =
CAPACITY
+ some performance

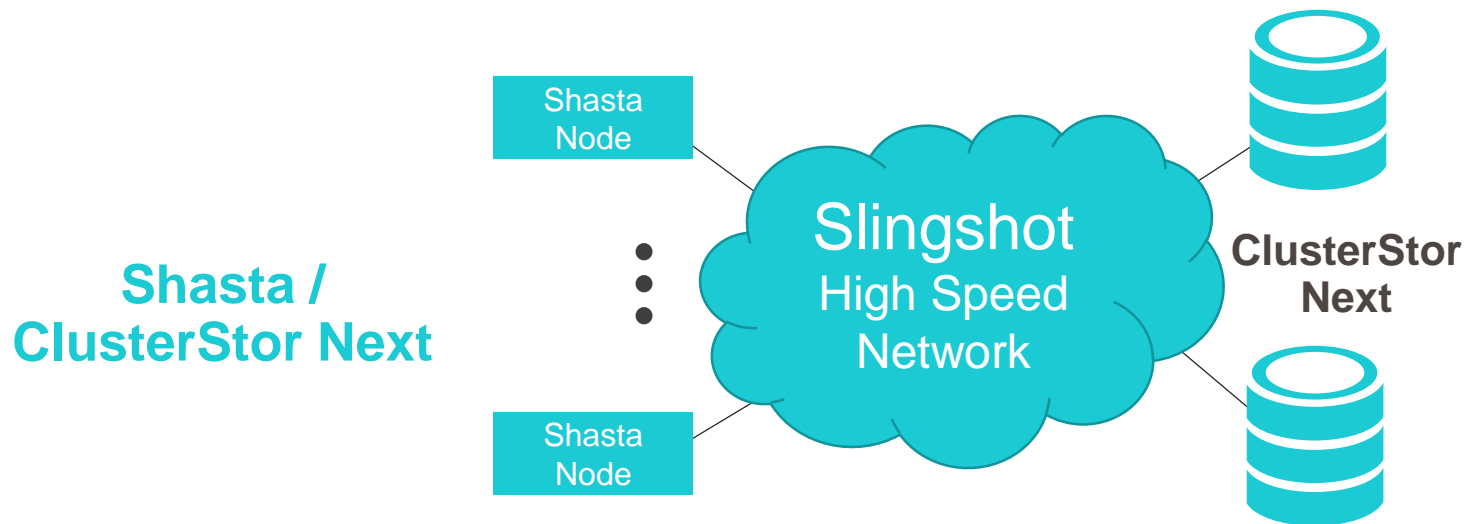
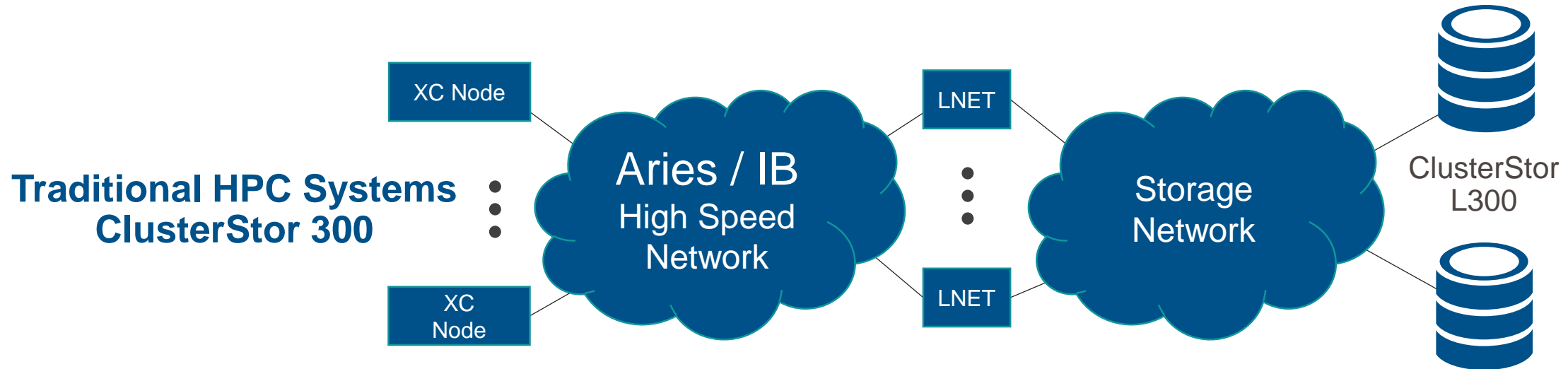
SOFTWARE =
Workflow-
accelerating
data placement
on the right media
at the right time

CLUSTERSTOR NEXT (CN) PORTFOLIO



	CN Metadata Array	CN Flash Array	CN Disk Array
Form Factor	2U24 PCIe 4.0 NVMe	2U24 PCIe 4.0 NVMe	4U106 SAS HDD
Servers	2	2	2
SSDs Per Enclosure	24	24	0
HDDs Per Enclosure	0	0	106
Optimal Use	Lustre Metadata	High BW/IOPS	Capacity
Max Per File System	8 (16 servers)	16 per rack	Configuration Dependent

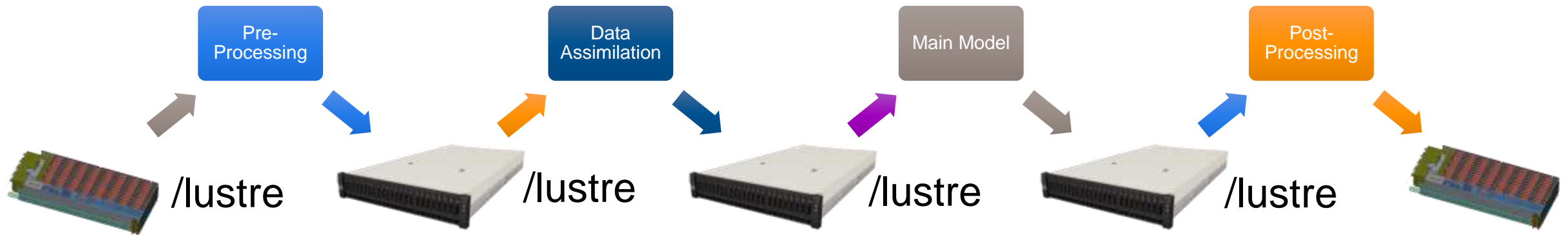
ClusterStor and Slingshot simplify Cray storage



Benefits:

- Lower cost
- Lower complexity
- Lower latency
- Improved small I/O performance

SCHEDULED TIERING WORKFLOW



- ClusterStor Data Services provides similar features
 - Integration with WLM for data promotion/demotion orchestration
 - Compute resources allocated after any requested data movement
 - High-speed flash directly on high-speed network
- DataWarp for the masses with >4-6x speed-up over disk

OPTIMIZING APPLICATION PERFORMANCE



Placement (FCS 2020)

What is this?

- Bridge capabilities from the L300 & L300F
- Use flash pool for temp data
- User can direct Lustre to move data up to the flash tier

Scheduled (2H 2020)

What is this?

- Bridge from DataWarp
- User inserts WLM directives into the job script
- Automated migration is performed outside of run time

Transparent (1H 2021)

What is this?

- A new class of service
- Read through, write back
- Transparent tiering
- No user impact

On Demand (roadmap)

What is this?

- Advancement from DataWarp
- Quality of service enhancement
- Dynamically provisioned via the WLM



PERLMUTTER



FIRST CLUSTERSTOR NEXT CUSTOMER

30 PB ALL FLASH

>4 TB/SEC

7M IOPS

3.2M file creates/sec.



>10TB/s with 1 EB*

*raw capacity



OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE
FOR U.S. DEPARTMENT OF ENERGY



CRAY



Argonne
NATIONAL
LABORATORY

Lab-wide Community and Campaign Storage

- 200+ PB
- 1+ TB/s
- 1st Half 2020 delivery
- Integrated with DAOS and Aurora in 2021



Call your local Cray Sales Rep for a
Deeper Dive on ClusterStor NEXT

