Hewlett Packard Enterprise

DATA STORAGE & MANAGEMENT AT SCALE FOR MAXIMUM COMPUTE ROI

Storage Technology Showcase
March 2020

CONFIDENTIALITY NOTICE

- **The information contained in this presentation** is proprietary to Hewlett Packard Enterprise (HPE) Company and is offered in confidence, subject to the terms and conditions of a Confidential Disclosure Agreement

- **HPE makes no warranties regarding the accuracy of this information**. This document contains forward looking statements regarding future operations, product development, product capabilities and availability dates. This information is subject to substantial uncertainties and is subject to change at any time without prior notification. Statements contained in this document concerning these matters only reflect Hewlett-Packard Enterprise's predictions and / or expectations as of the date of this document and actual results and future plans of Hewlett-Packard Enterprise may differ significantly as a result of, among other things, changes in product strategy resulting from technological, internal corporate, market and other changes. This is not a commitment to deliver any material, code or functionality and should not be relied upon in making purchasing decisions.

GOALS

Full-Speed Compute Execution

Managed Storage Costs

Data Protection & Management

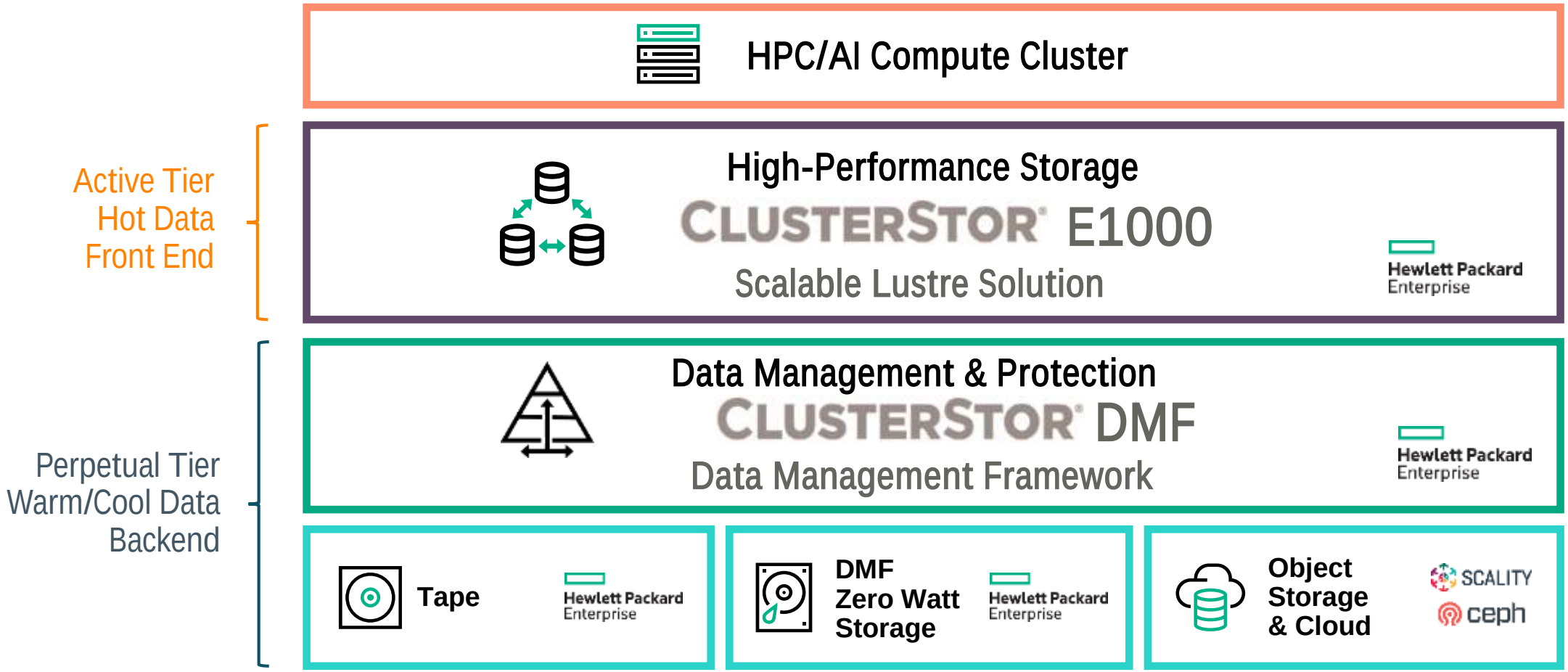
Reduced Administrator Workload

SOME NUMBERS...

10 20 31

HPC/AI COMPUTE & STORAGE

HPC Storage Supports an Optimized Compute Experience for Active Workloads



CLUSTERSTOR E1000

Scalable Lustre Solution

CLUSTERSTOR® **E1000 "AT A GLANCE"**

SSD

NVMe Gen 4 storage controllers

- Extracting up to 3.3 GB/sec per SSD
- Extracting more than 140 MB/sec per HDD
- Support for Ethernet, InfiniBand and Slingshot fabrics

ClusterStor Data Services

- Intelligent data movement between Lustre flash & disk storage tiers
- Maximizes performance & cost-efficiency

lustre

Lustre 2.12 LTS

- Choice of LDISKFS/GridRAID or ZFS
- Hardened by Cray

HDD

Dense media enclosures

- >1.2 PB usable of HDD per 4RU
- >230 TB usable of SSD per 2RU

View for ClusterStor 2.0

- Per job performance analysis
- Integrating system management

NEW END-TO-END PCIE 4.0 STORAGE CONTROLLER

Two embedded HA application modules with single socket AMD “Rome” **PCIe 4.0** CPU each

Up to 6 x 100/200 Gbps **PCIe 4.0** NICs (Slingshot, GbE, IB)

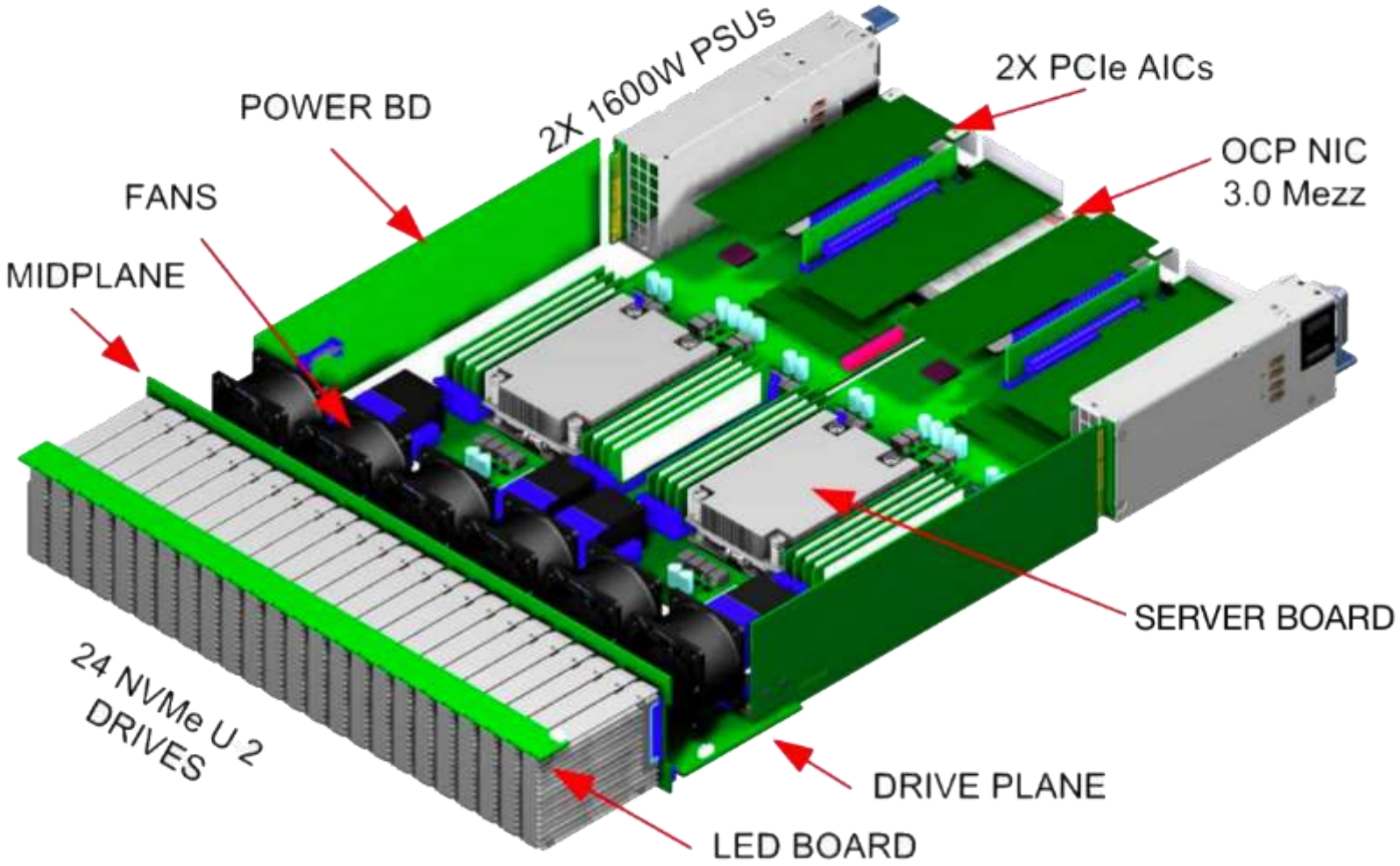
60 GB/sec Write*

- Up to 1.7M IOPs**

80 GB/sec Read*

- Up to 3M IOPs**

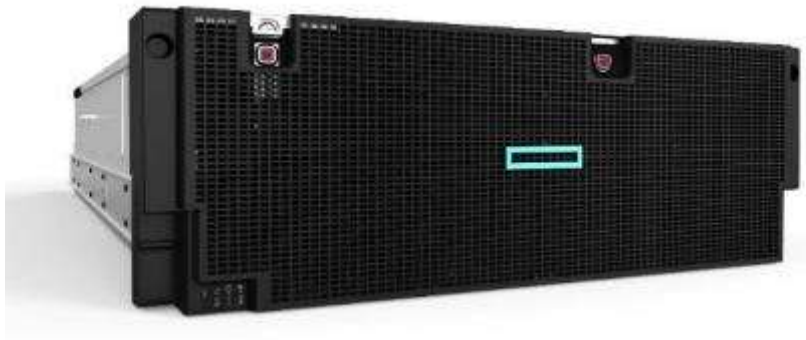
Up to 424 x 7.2K RPM HDD SAS-attached to the 2U enclosure



HPE D8000 ENCLOSURE

Capabilities and Attributes:

- 106 Drive Large Form Factor 4U JBOD (Spade Carrier)
- 12Gbps SAS3.0 E2E
- 12Gbps Daisy Chaining
- 12Gbps SAS/SSD/SATA Hard Drives
- Single and Dual Domain support (IOMs)
- Hot Swappable Fans, Hard Drives, IO Modules, PSUs
- 1,2, 4M cable length support
- Non-Disruptive JBOD FW flashing (online upgrade)
- Individual HDD Power control



NEW ZERO BOTTLENECK PCIE 4.0 STORAGE CONTROLLERS

Two embedded
HA application modules with
single socket AMD "Rome"
PCIe 4.0 CPU each

Up to 6 x 100/200 Gbps
PCIe 4.0 NICs
(Cray Slingshot, GbE, IB)

Up to 24 x dual ported
NVMe PCIe 4.0 SSD
in the 2U enclosure

Up to 424 x 7.2K RPM
HDD SAS-attached
to the 2U enclosure

Flexible Building Block for ClusterStor E1000

System Mgmt Unit
(5 SSD)

One
per file system

MetaData Unit
(24 SSD)

Up to 10
per file system

All Flash Unit
(24 SSD)

Disk Controller
(2 SSD)

SAS-attached 4U106
HPE D8000 HDD enclosures

Supported Storage Drives

NVMe Gen 4

3 DWPD:

- 1.6/3.2/6.4/12.8 TB

1 DWPD:

- 1.92/3.84/7.68/15.36 TB

7.2K RPM

4/6/10/12/14/16 TB

CLUSTERSTOR E1000 **FLEXIBILITY**



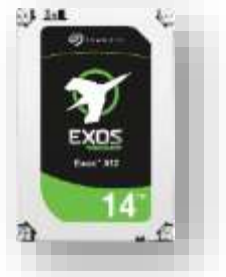
	Extreme Performance	HDD Performance	HDD Capacity
SSD Performance (read)	80 GB/s		
SSD Performance (write)	60 GB/s		
SSD Usable Capacity (3.2 TB)	55.3 TB		
HDD Performance		30 GB/s	30 GB/s
HDD Usable Capacity (14TB)		2.14 PB	4.27 PB
Network ports	6 x 200 Gbps	2 x 200 Gbps	2 x 200 Gbps
Height Rack Units	2	10	18

Base rack specs:
- 1,440 GB/sec (read)
- 1,080 GB/sec (write)

Expansion rack specs:
- 1,600 GB/sec (read)
- 1,200 GB/sec (write)

Base rack specs:
- 90 GB/sec read/write
- 7.5 PB usable capacity

Expansion rack specs:
- 120 GB/sec read/write
- 10 PB usable capacity

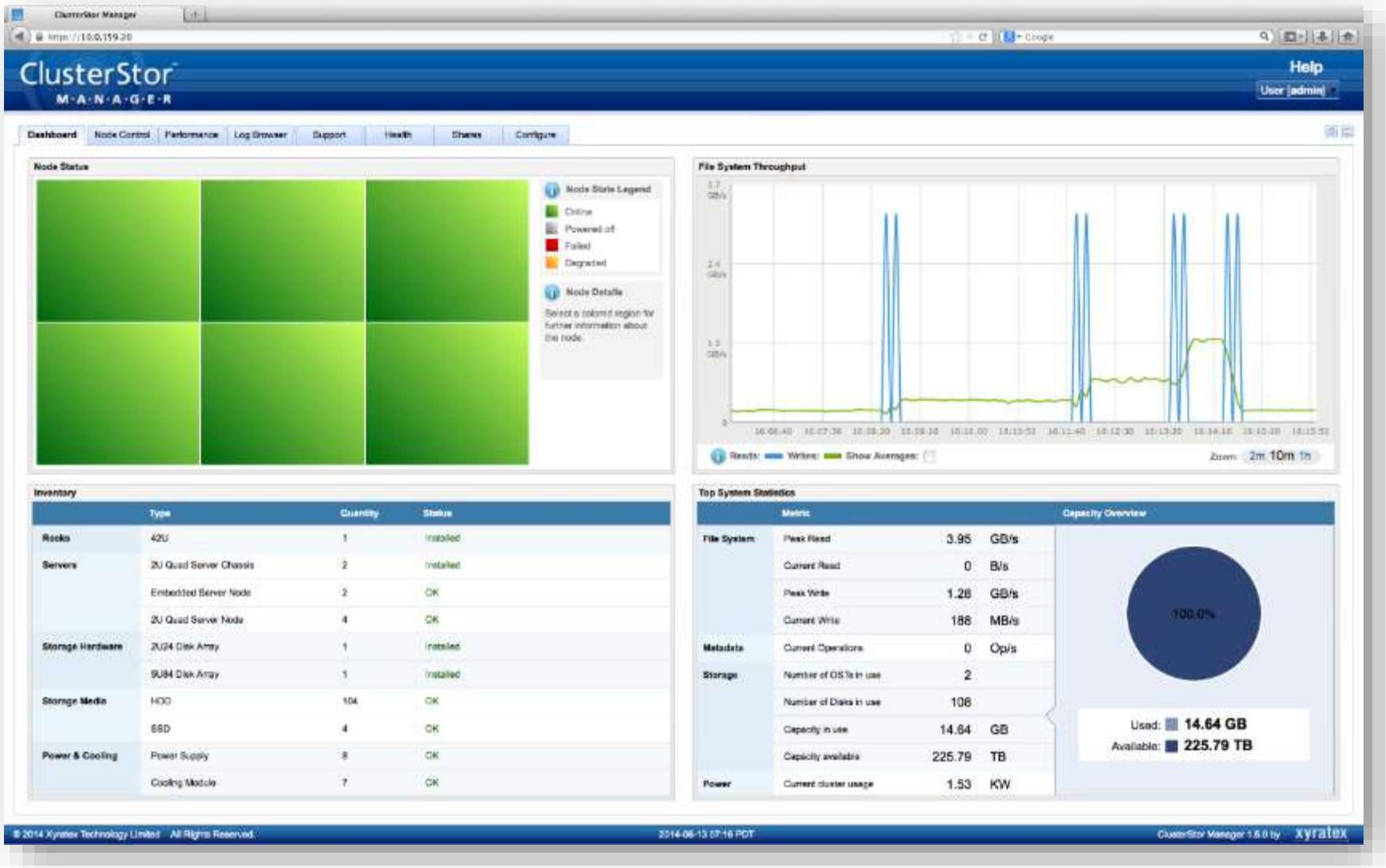


CLUSTERSTOR DATA-AT-REST ENCRYPTION

- E1000 will add Data Encryption-at-Rest (DEAR) support in the 2nd half of 2020
- E1000 DEAR support will leverage our field proven encryption support from ClusterStor L300
- E1000 DEAR support will be achieved with:
  - Self-encrypting drives (SEDs) HDDs and SSDs
  - Federal Information Processing Standard (FIPS) HDDs and SSDs
  - Fornetix Key Orchestration™ Appliances
  - AES256-bit encryption keys
- DEAR disk options are targeted to be:
  - 4/10/16TB SED or FIPS HDDs
  - 1.92/3.84/7.68/15.36TB SED SSDs
  - TBD FIPS SSD capacities – contact product management for further details.
- DEAR can be implemented on all nodes (MDUs, SSUs) within E1000 or only a subset of nodes
- DEAR support will not incur a performance penalty. The onboard hardware encryption of the drive ensures the encryption does not impact filesystem performance.

CLUSTERSTOR MANAGER

- Deep integration with HW infrastructure with rich telemetry data collection
- Fault identification, operator notification and guided navigation for repair
- Focus on ease of operator use



CLUSTERSTOR VIEW
PERFORMANCE MANAGEMENT

JOBS
Searchable fields for the most important data to the job step level

QUALITY
Know which jobs might be causing issues with visual cues

CLUSTERSTOR E1000: SUMMARY

Powerful

Highest file system performance in the industry for

- All Flash
- All Disk
- Hybrid system

Intelligent

Intelligent software that aligns the data flow with the workflow.

Intelligent hardware design that extracts the maximum performance from each storage drive.

Flexible

All Flash, All Disk and Hybrid configurations

Modeling & Simulation and ML/DL and High performance data analytics

Connects to compute from any vendor

Scalable

Start anywhere, scale to wherever you need.

Selected by 100% of the U.S. DOE exascale sites.

Easy

Easy to buy with HPE and Cray HPC/AI compute.

Easy to implement as full integrated, factory-tested system.

CLUSTERSTOR DMF

Data Management Framework

HPE DATA MANAGEMENT FRAMEWORK (DMF)

DATA MANAGEMENT AND PROTECTION AT SCALE FOR HPC

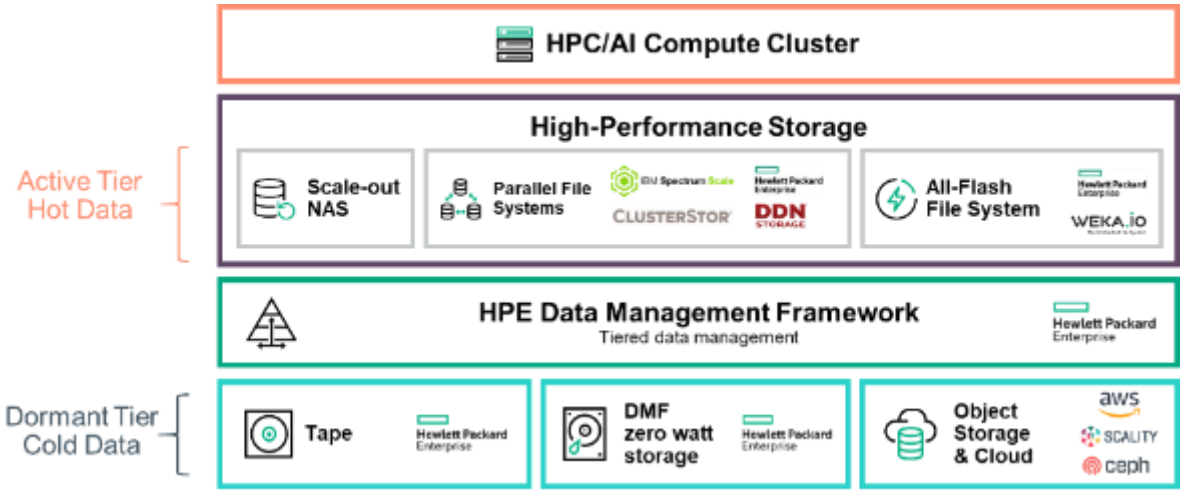
Solving HPC Storage Pain Points

- Backup and Protection of Exascale Datasets: Simple, effective and proven solution for backup, retention and disaster recovery for multi-petabyte data sets across on-prem and cloud
- Catalog and Organization of Massive File Sets: Supports multi-billion file systems with scalable data movement and search. File version management, extended attributes and retention.
- Meets HPC Users' "Need for Speed": Links to HPC job schedulers to automatically place HPC data assets on the fastest class of storage for the most efficient overall system throughput and ROI

Customer Deployments

NASA	• Over 20 Years in Production • Over 200 Petabytes of HPE Managed Data
ENI Oil & Gas	• Deployed with Cray ClusterStor • Tiering Data to HPE CEPH Storage
National Basketball Association	• Managing Player/Team Video Repository • Core Data Protection and Delivery System

A Complete Solution for HPC Data Management

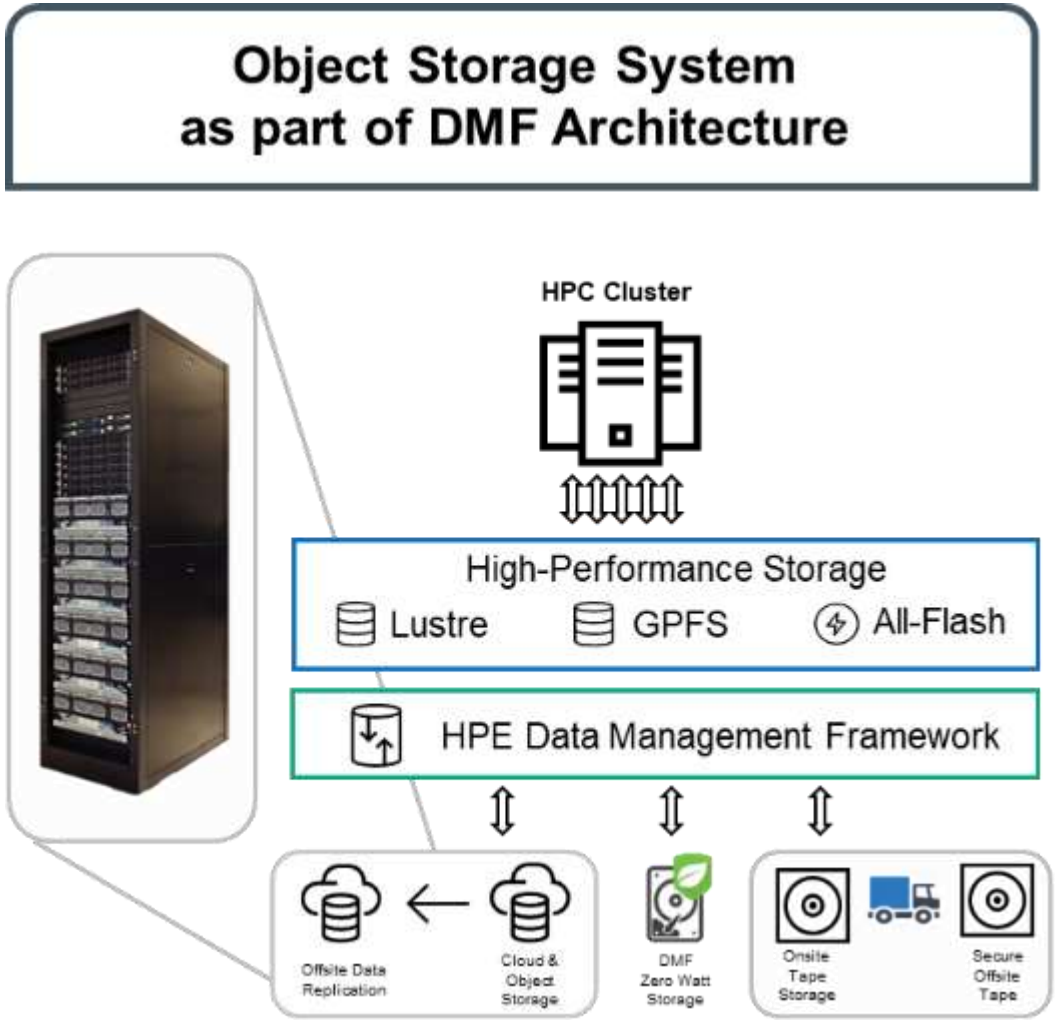


Simplify New Technology Integration & Ease Administrator Workloads

- Integrate cloud-based storage as part of the HybridHPC mix
- Seamlessly manage migration, validation and consolidation of massive data sets
- Perform the migration over a period of weeks or months with no impact to users or jobs
- Stage managed data to burst buffers or all-flash filesystems

DMF OBJECT STORAGE SUPPORT

- **Standards-based Integration**:
  - Use of S3 interface enables compatibility with Scality, CEPH, Amazon S3, DDN WOS, HGST Active Archive, NetApp StorageGrid, ECS, and open source alternatives
- **Scalability & Throughput**:
  - Scalable DMF connections to object storage environment
  - DMF Parallel Data Mover architecture with high availability, balancing and failover
- **Flexibility**:
  - Ability to blend object storage with alternative storage options including Zero Watt Storage (performance) or tape (off-site disaster recovery)



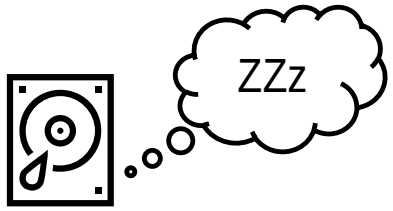
DMF TAPE STORAGE INTEGRATION

- DMF is certified with libraries from HPE, as well as Spectra Logic, IBM and Oracle (StorageTek)
  - Streams to tape drive at native rates, even for small files
  - Block ID positioning for fast seek
- Support for latest LTO-8 and Enterprise-class drive technology
- Advanced feature support for accelerated retrieval and automated library management
  - Recommended Access Order (RAO) and SpectraLogic's TAOS
  - Data Integrity Verification (DIV) and Logical Block Protection (LBP) available with Oracle T10k and IBM LTO drives

LTO-8 12TB
360MB/s

TS1160 20TB
400MB/s

ZERO WATT STORAGE ADVANCED DISK MANAGEMENT



Integrated Disk Power Control

- Disk Power Control is configurable by DMF administrator down to individual drive
- DMF automatically calculates power off delay based on drive SMART data and drive lifespan
- The maximum number of drives that can be simultaneously accessed are configurable by DMF administrator per virtual ”library”

Proactive Disk Scrubbing

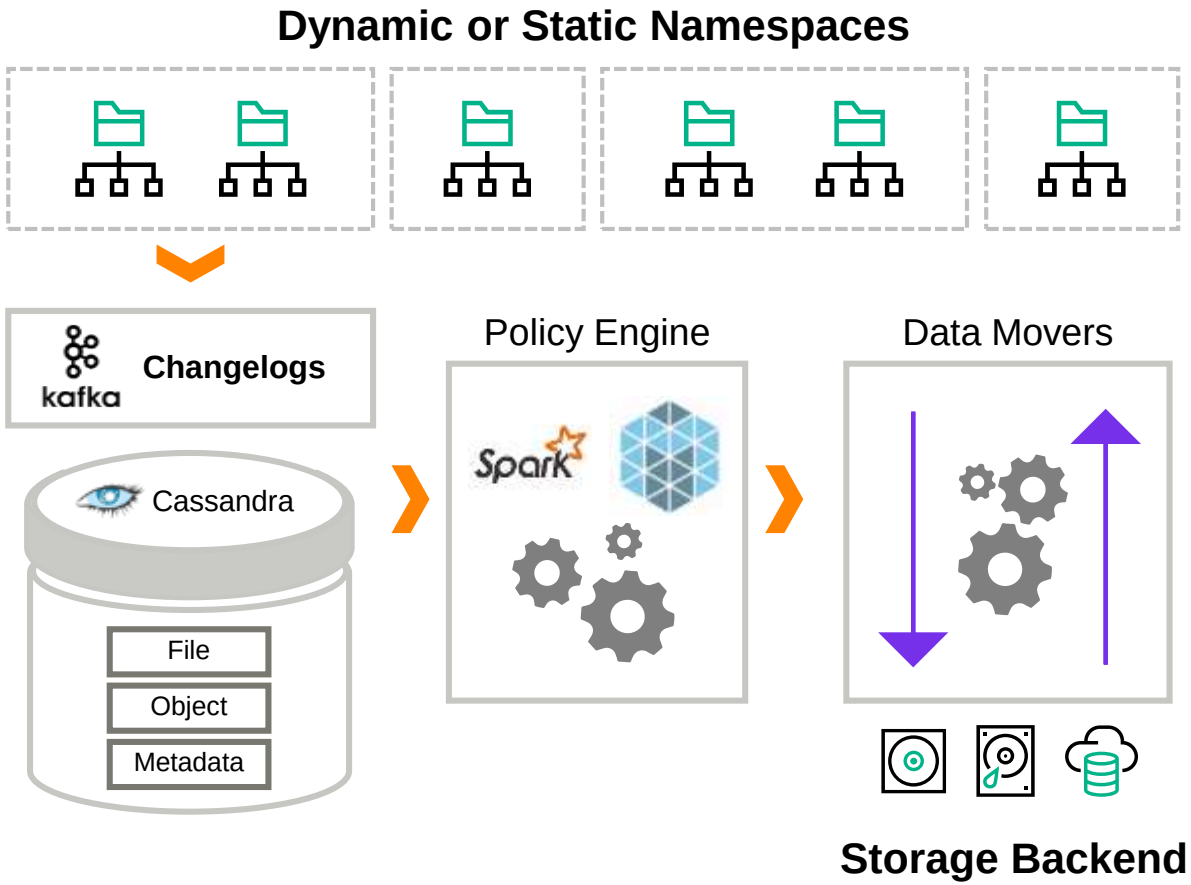
- All drives are proactively scrubbed in background on regular schedule
- Sector read errors that can be corrected by parity data will be automatically fixed
- Drives with uncorrectable errors are flagged for administrative action
- DMF records SMART data for each drive on mount and unmount
- Scrub progress and statistics are available to DMF administrator

106 HDDs
12Gb/s
4U

DMF: DATA MANAGEMENT AND PROTECTION IN A SINGLE SOLUTION

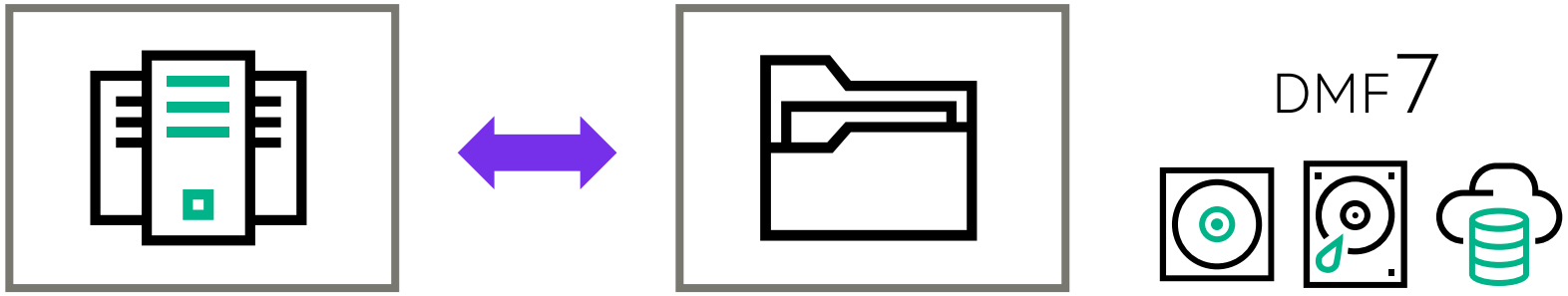
Managing Data Workflows

- Aware of file metadata in managed namespaces
- Allows queries & policies driven by file metadata
- Supports backend data versioning
- Uses “datasets” as easy handle to potentially very large sets of files and directory structures
- Can de-stage files to backend and stage data from backend into namespaces
- Dynamically creates and tears down namespaces
- Modular architecture, scalability and HA
- Built on modern open source frameworks designed for Big Data



INTEGRATED BACKUPS AND FILE VERSIONING

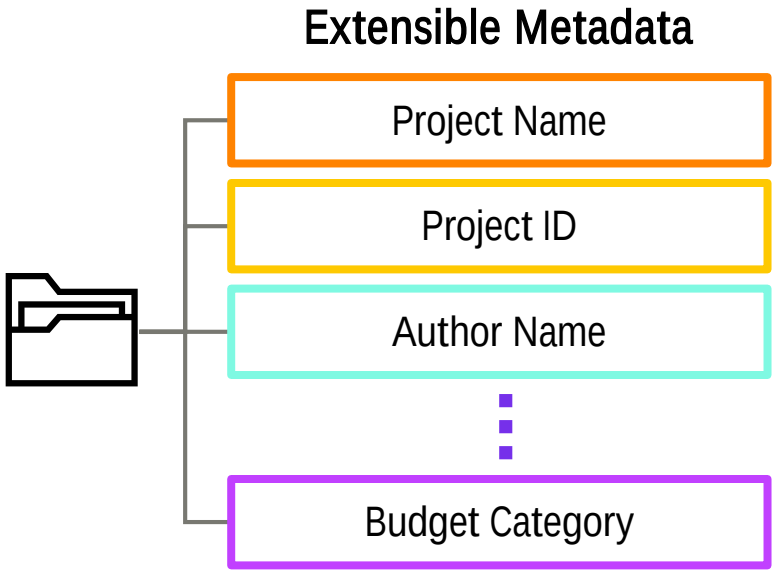
- Large volumes of data can be backed up via automated migrations without impacting users
  - Policies trigger migrations of all modified files & directories at specific intervals or after periods of inactivity
  - Filesystem namespace is periodically captured into backup datasets
- Users can view past versions of backup data sets and individual files and restore the desired version
  - Complete history of the evolution and contents of file systems maintained by DMF v7
- Replication of results for specific job runs, or for validating the correct operation of modified system codes, is enabled via Point in Time restoration of file systems from backup datasets



EXTENSIBLE METADATA SUPPORT

Data management flexibility and precision with extensible metadata

- DMF v7 is based on scalable metadata repository
- Repository functions as long-term data store for information about file system structure, attributes, contents and evolution over time
- Metadata repository supports POSIX extended attributes on files and directories, e.g. project name, project ID, etc.
- Queries can be run against metadata including extended attributes for precise and flexible selection of files, e.g. data set creation
- Additionally, policies can be run against the results of metadata queries for data movement, archiving, etc.

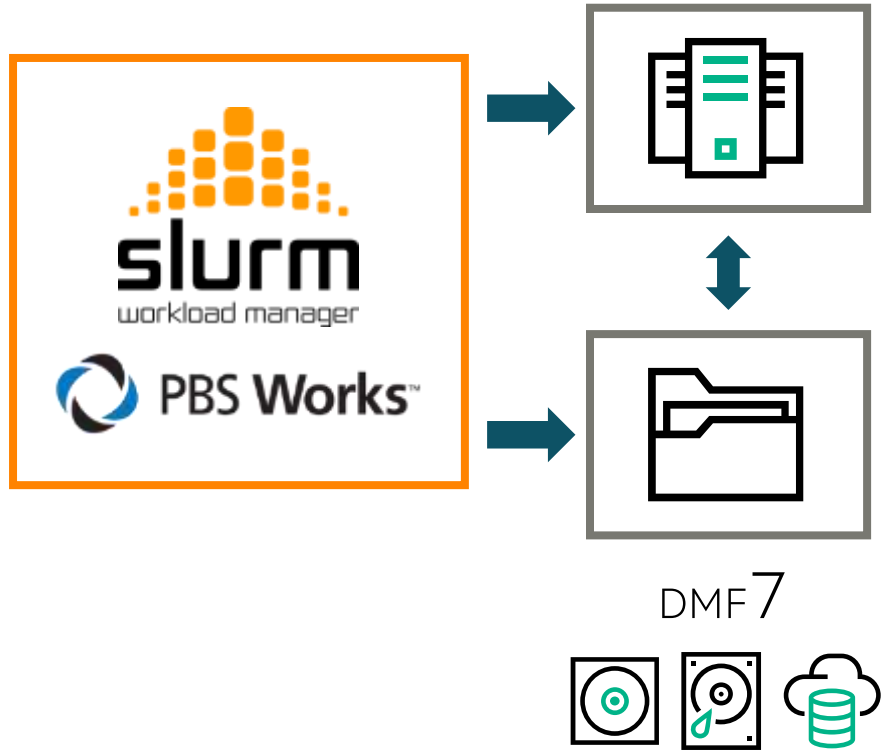


JOB SCHEDULER INTEGRATION

DMF v7 jobs can be scheduled via standard HPC job schedulers:

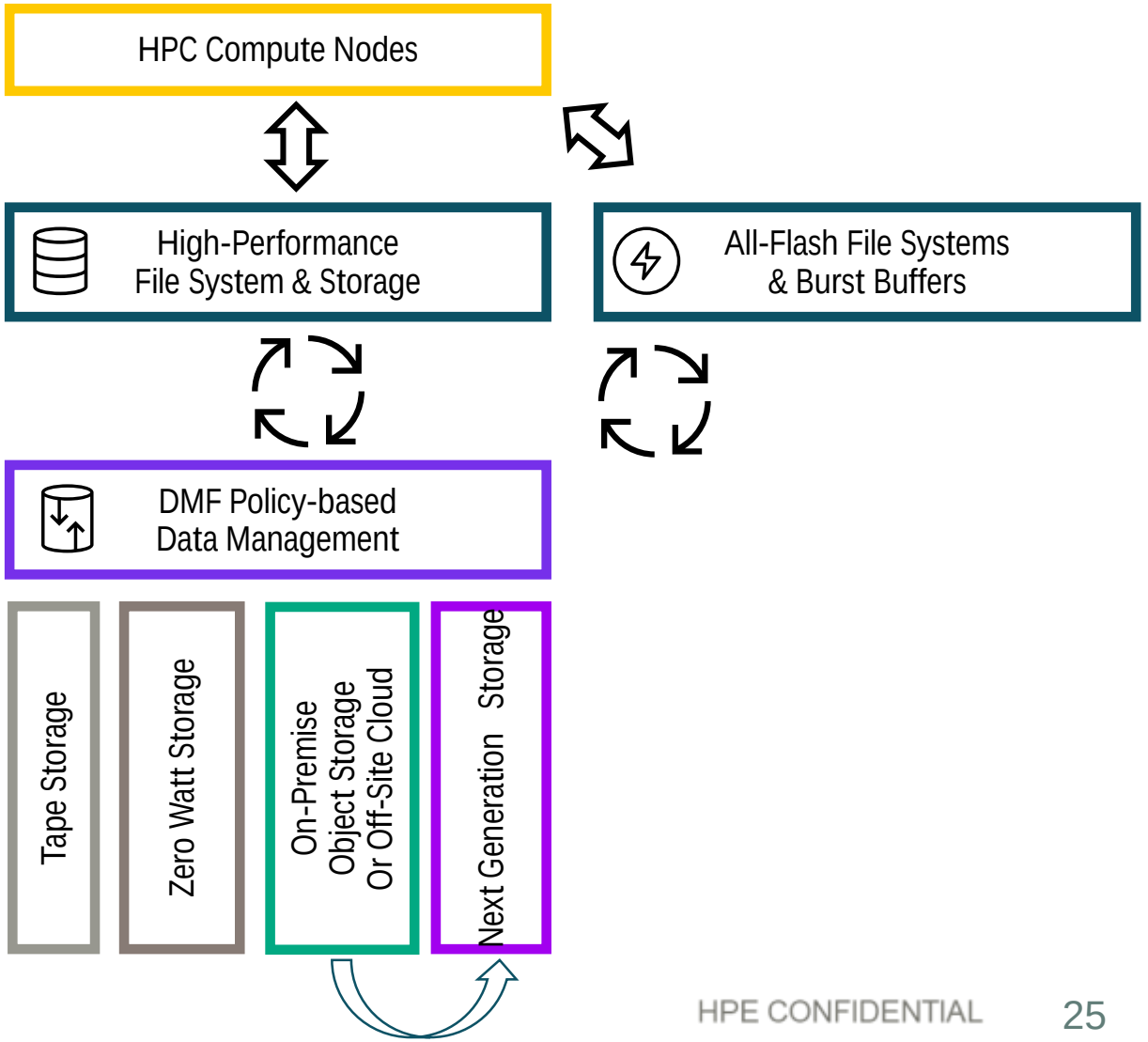
- DMF v7 is API-enabled for integration with job schedulers, e.g. Slurm, PBS Works
- Data set "labels" can be defined by an administrator to "name" a data set
  - Simplifies job management and reproducibility of results in the future.
- Job scheduler definitions can include information to place data sets on the fastest tier of storage in advance of job initiation
- After a job is done, its data set can be de-staged and migrated to a designated tier by policy

Metadata extensions, data set labels, and job scheduler integration simplify and streamline data workflows



CONTINUOUS TECHNOLOGY INTEGRATION

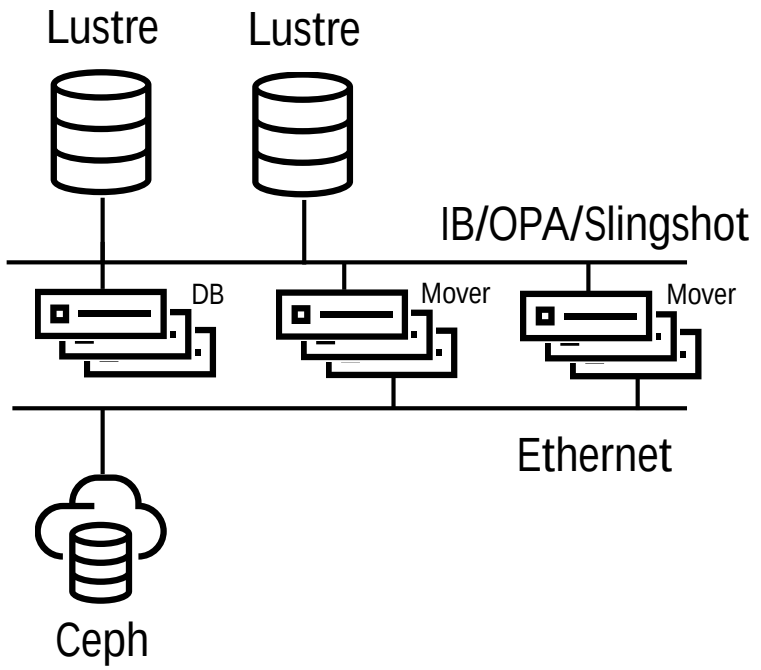
- Manage introduction of new storage technologies over time without disruption
  - Seamlessly manage migration, validation and consolidation of massive data sets
  - Perform the migration over a period of weeks or months with no impact to user data access
  - Stage managed data to burst buffers or all-flash filesystems



DMF 7 CUSTOMER EXAMPLE

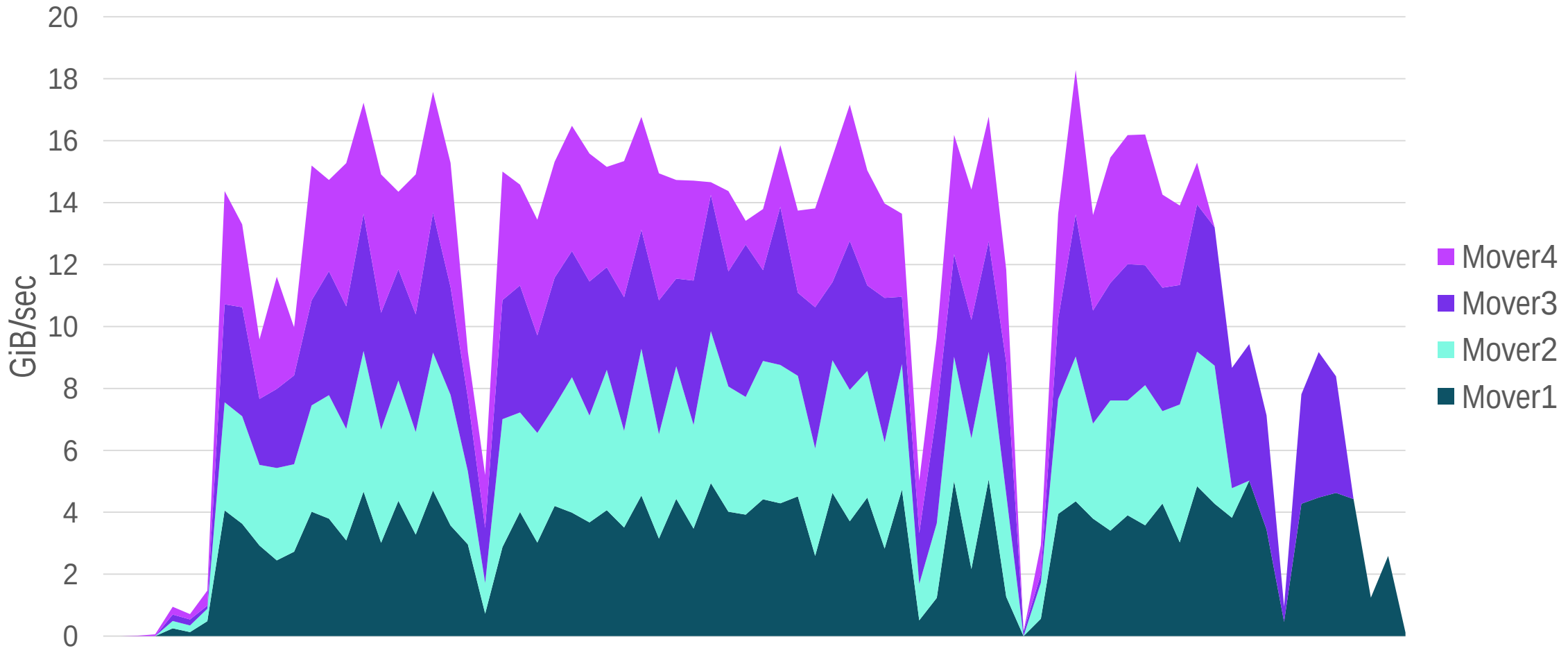
Multiple file system environment

- Multiple, highly active managed Lustre filesystems
- Tiering to object storage backend to free up space on front-end
  - >10GB/s sustained
- Moving datasets across managed filesystems
  - Including future technologies
- Copying files & directories to backend for disaster recovery
- Running reports
  - Understanding data consumption by user & group
  - Tracking data growth trends



DMF7 S3 MOVER PERFORMANCE

4 DL360 Mover Nodes | Lustre to Ceph



SUMMARY

Full-Speed Compute Execution

Managed Storage Costs

Data Protection & Management

Reduced Administrator Workload

THANK YOU

Mark.Seamans@hpe.com