DMFUG 2018

DMF7 Architectural Overview and Design Introduction

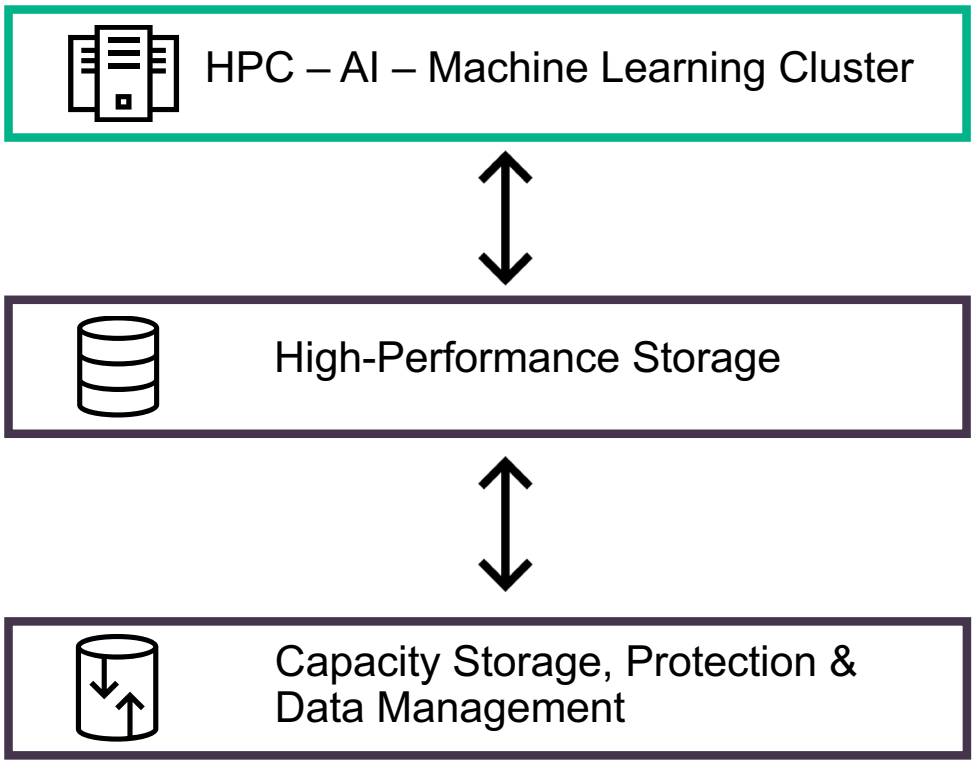
Zsolt Ferenczy

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.

HPC & AI | **Data Management** Technology Shifts & Strategies



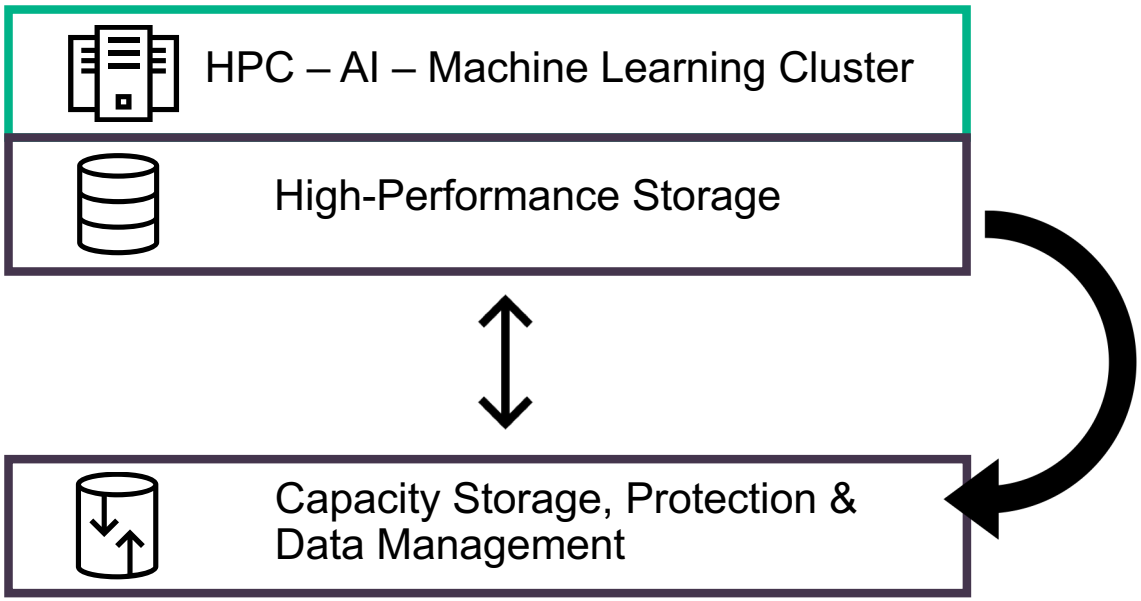
Key Takeaways

- Disaggregate and scale High-Performance Storage Tier *Independently* from Capacity Tier
- Collocate Performance Tier with Compute and Fabric
- Implement tiered Data Management for Capacity Scaling and Data Protection

Tiered Data Movement and Management are a Key Requirement – and HPE Data Management Framework (DMF) Meets That Need

Hewlett Packard Enterprise

HPC & AI | **Data Management** Technology Shifts & Strategies

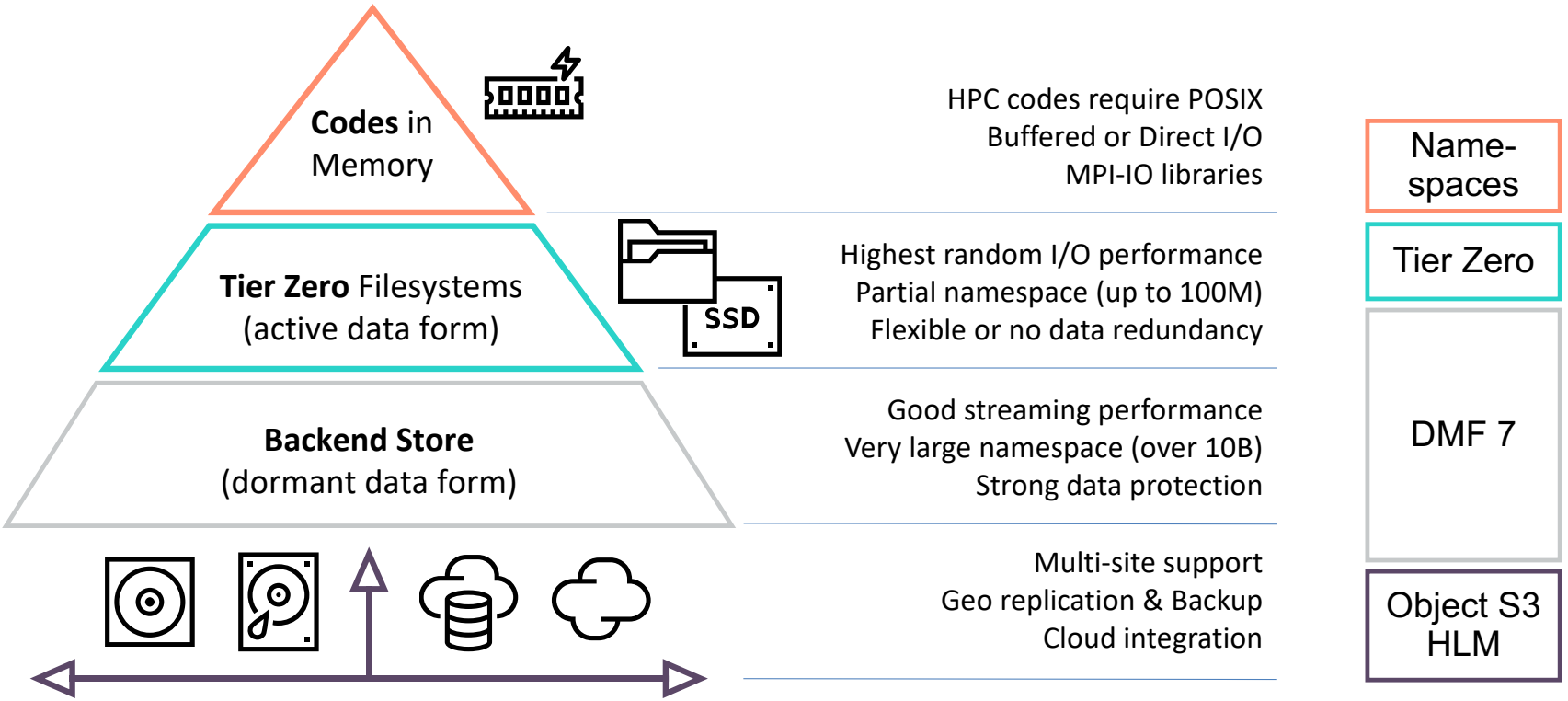


Key Takeaways

- **Transition From**
  - Enormous Static File Systems
  - Traditional Backups
  - Snapshots
- **Transition To**
  - Dynamic Namespaces
  - Searchable Metadata
  - Integrated Data Management & Provenance
  - Policy-based Protection

Hewlett Packard Enterprise

HPC Workflows | **Data Tiering Model** Active & Dormant Forms



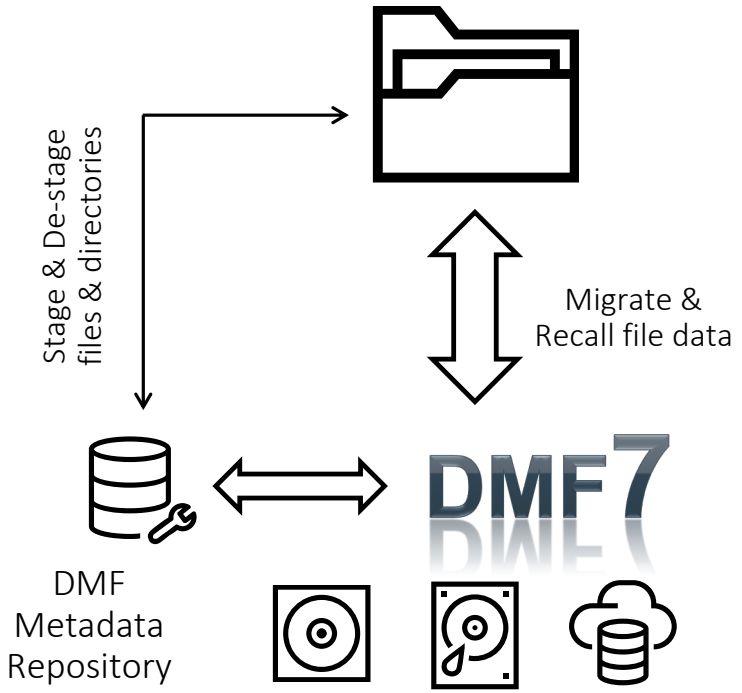
Data Management | **DMF 7** Data Staging & Dormant Form

Key Takeaways

Data Staging and Dormant Form

DMF 7 objects can exist in dormant form without filesystem representation. Only data that needs to be accessed has to be staged into a filesystem.

- **Just-in-Time Data Staging:** Individual objects or collections (data sets) can be staged into managed filesystem as files and directories either on demand by user or in coordination with job scheduler.
- **Increased Filesystem Performance:** De-staging inactive files and directories reduces namespace pollution and allows DMF managed filesystems contain unstructured file sprawl.
- **Enhanced Privacy and Protection:** Files not in active use can be completely de-staged and removed from managed filesystem preventing unauthorized access.



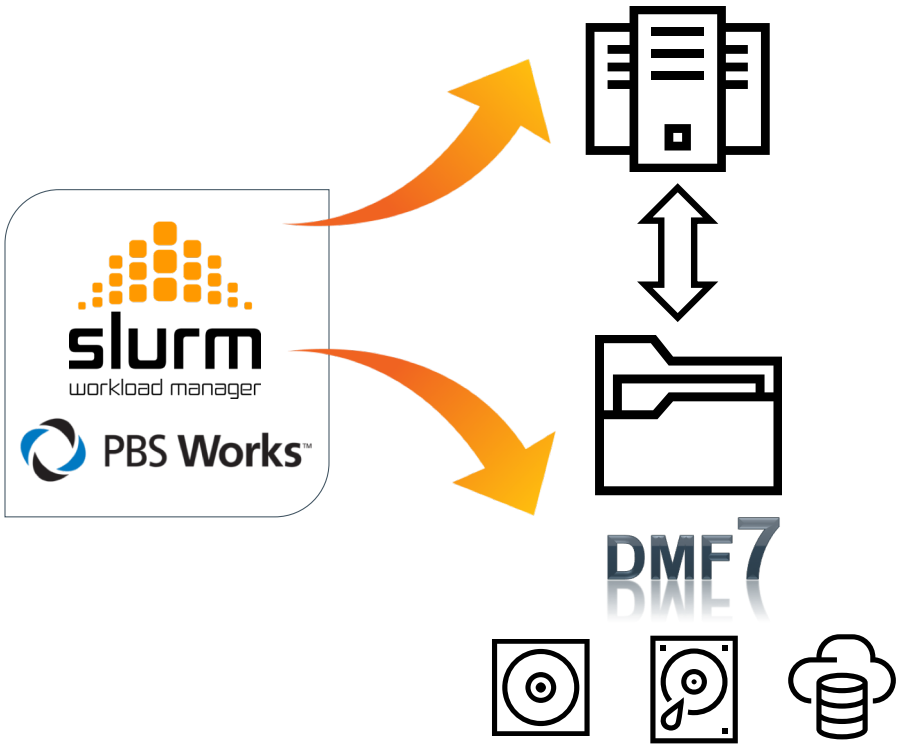
Data Management | **DMF 7** Job Scheduler Integration

Key Takeaways

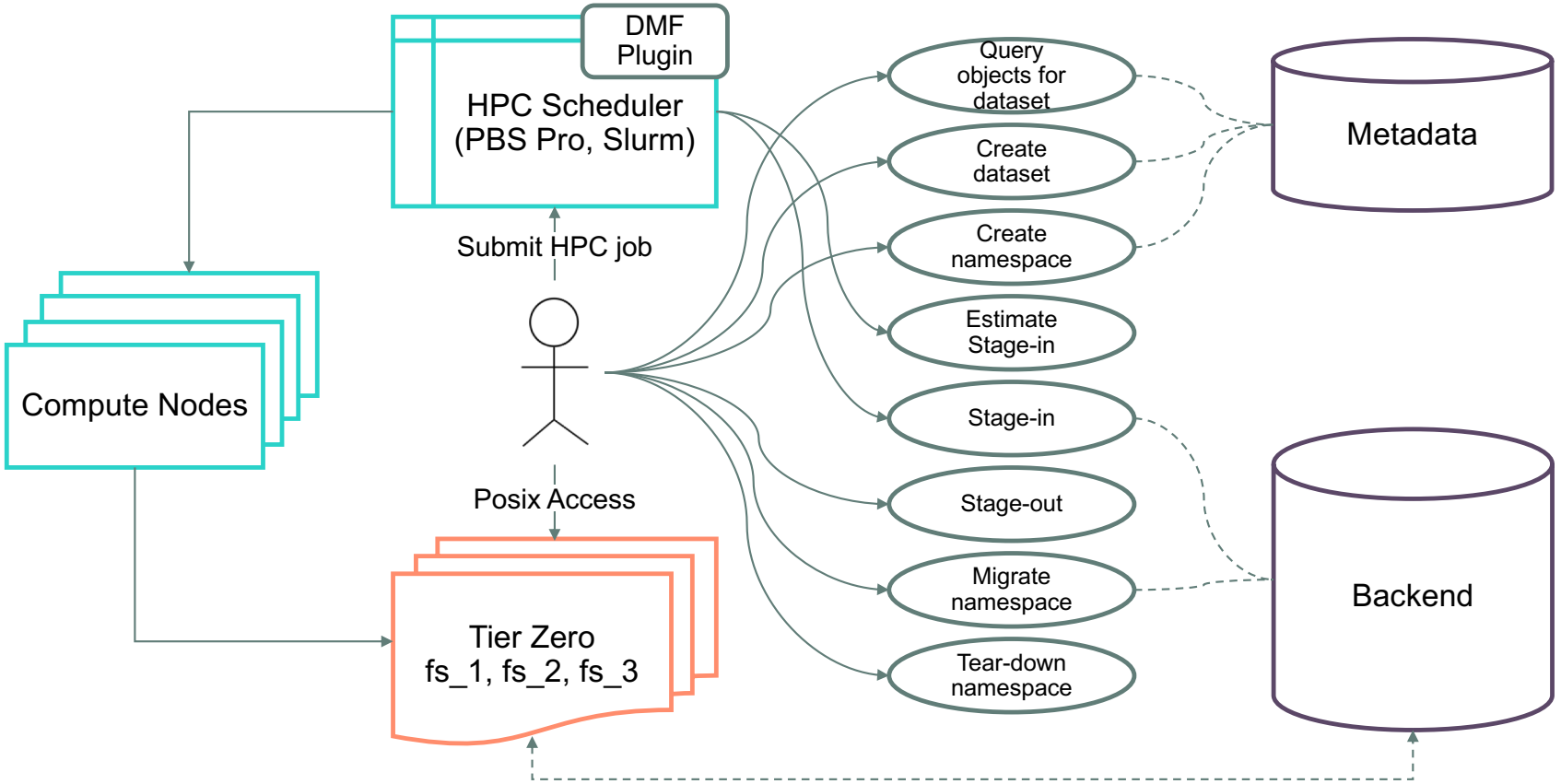
Job Scheduler Integration

DMF 7 is API-enabled to allow integration with job schedulers for operations such as data pre-staging to a high-performance flash tier in advance of job execution.

- **Data Pre-Staging or Recall Based on Metadata:** Job scheduler definitions can include information on required data sets that should be on the fastest tier of storage in advance of job initiation.
- **Data Set Definition:** Job administrators can define labeled "data sets" that are a collection of files/directories associated with a specific job type. This process can simplify job management and enable more easily reproducible results in the future.
- **Data Migration or De-Staging After Job Completion:** Data can be migrated or de-staged from high-performance storage based on automated policies – or job administrators can direct the system to migrate or de-stage data after job completion.

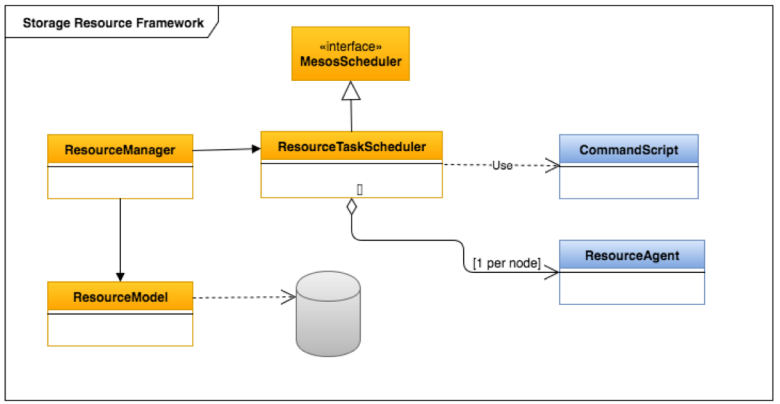


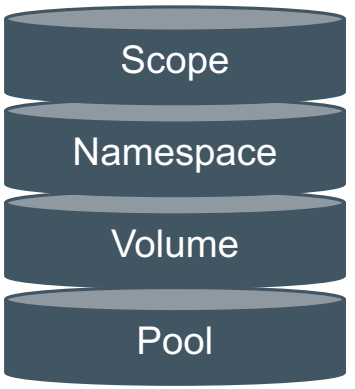
HPC Use Case | **Dynamic Namespaces** HPC User



Data Management | **DMF 7** Storage Resource Framework

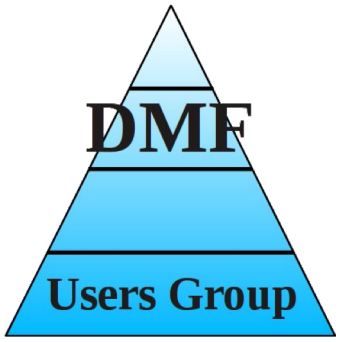
- SRF (Storage Resource Framework) is the DMF7 component responsible for managing all the storage resources known to DMF.
- DMF7 supports the following storage types:
  - Static HPE XFS filesystem
  - Static Lustre filesystem
  - Dynamic HPE XFS filesystem on a static volume
  - Fully Dynamic HPE XFS filesystem
- SRF performs the following actions:
  - Discover nodes, pools, parallel filesystem providers, static volumes, and static namespaces
  - Register nodes, pools, parallel filesystem providers, static volumes, and static namespaces with DMF
  - Define groups of compute nodes (scopes)
  - Create Dynamic Namespaces
  - Attach a Dynamic Namespace to a scope of compute nodes
  - Unregister, Detach, and Destroy





Hewlett Packard Enterprise

Hewlett Packard
Enterprise



Thank You