Intel Enterprise Edition of Lustre* HSM

Scaling capacity and performance without compromise using SGI® DMF™

Capacity, Performance & Reliability

Robert Mollard

Senior Storage Specialist, APAC

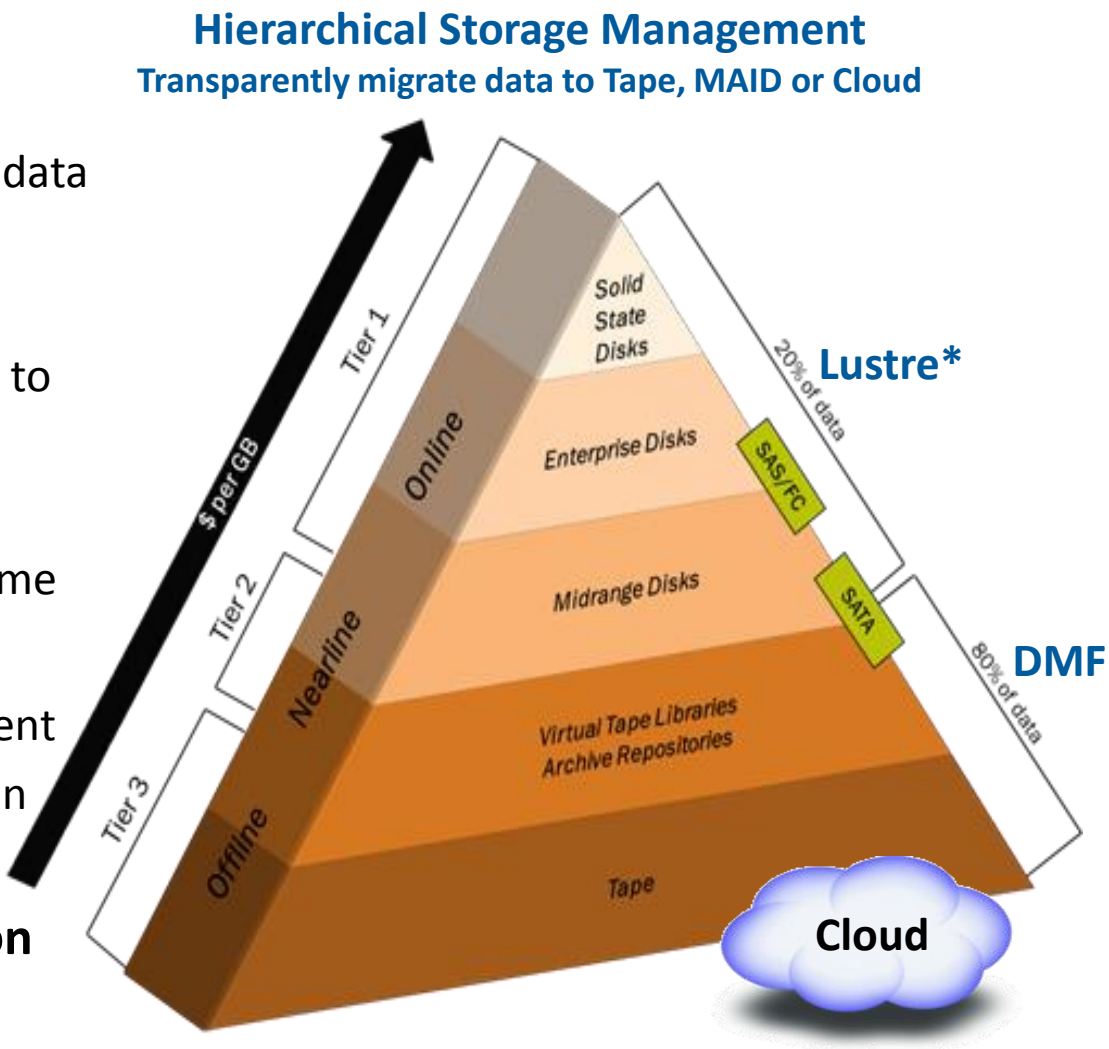
sgi

Agenda

- Hierarchical Storage Management
- Lustre Scalability with DMF (HSM)
- Tiered Data Management
- DMF – Start small and grow
- DMF Direct Archiving
- JBFS Fast Mount Cache
- Summary

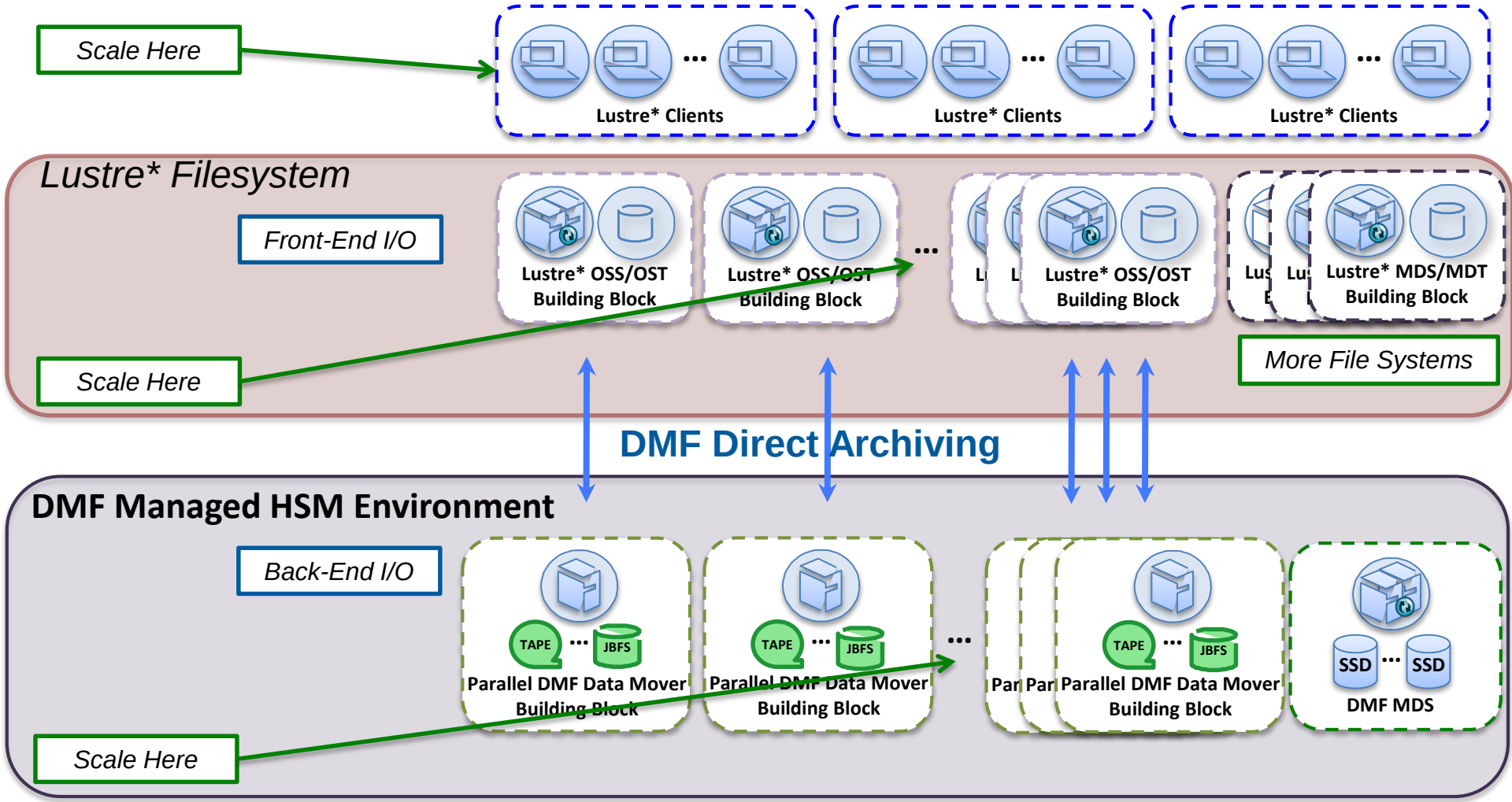
HSM | Data Migration Facility (DMF)

- **Data life cycle management**
  - DMF manages the placement of data within multiple tiers of storage
- **Automated data migration**
  - From expensive, production disk to 2nd or 3rd tier storage
- **Transparent to user**
  - All data appears on line all the time
- **Key Benefits**
  - DMF reduces tier 1 disk investment
  - DMF reduces power consumption
  - DMF protects data long term
- **SGI® DMF™ 25 years in production**

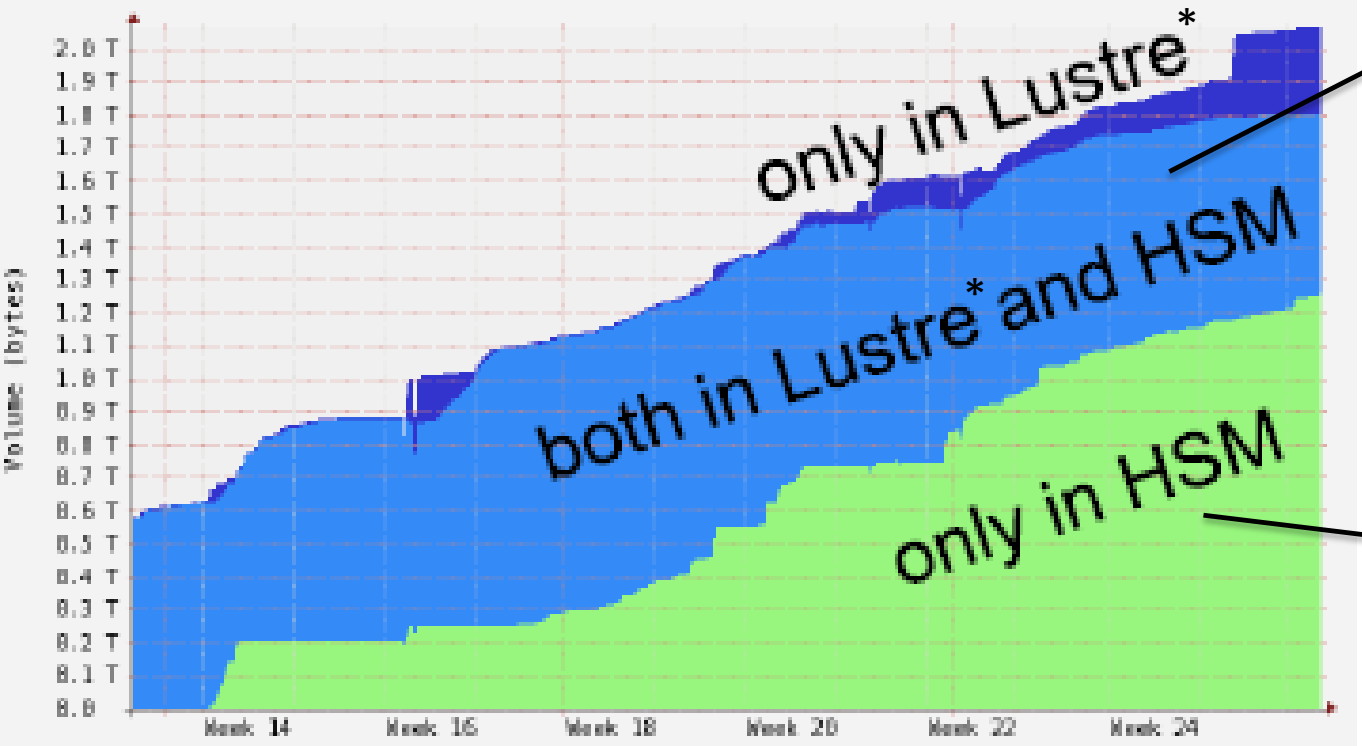


Scalability without compromise

Capacity, Performance & Reliability



Seamless Tiered Data Management



The most recent and active data is "live" in Lustre* and mirrored within DMF.
ALL DATA APPEARS ONLINE to users.

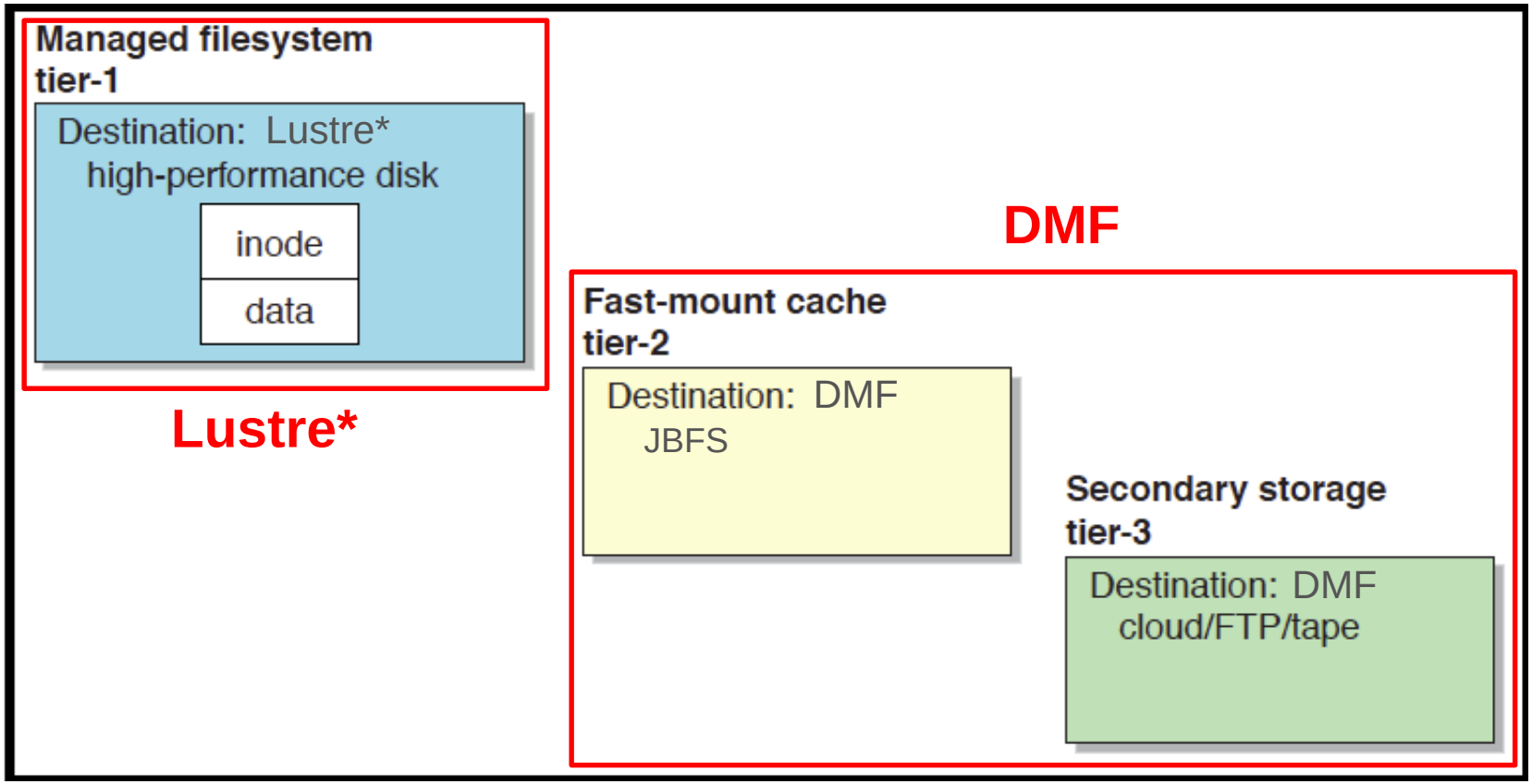
"Overflow" data is stored and protected within DMF on various cost-correct media types

Tiered Data Management

HSM perspective: **regular** file

User perspective: **online** file

Before migrating

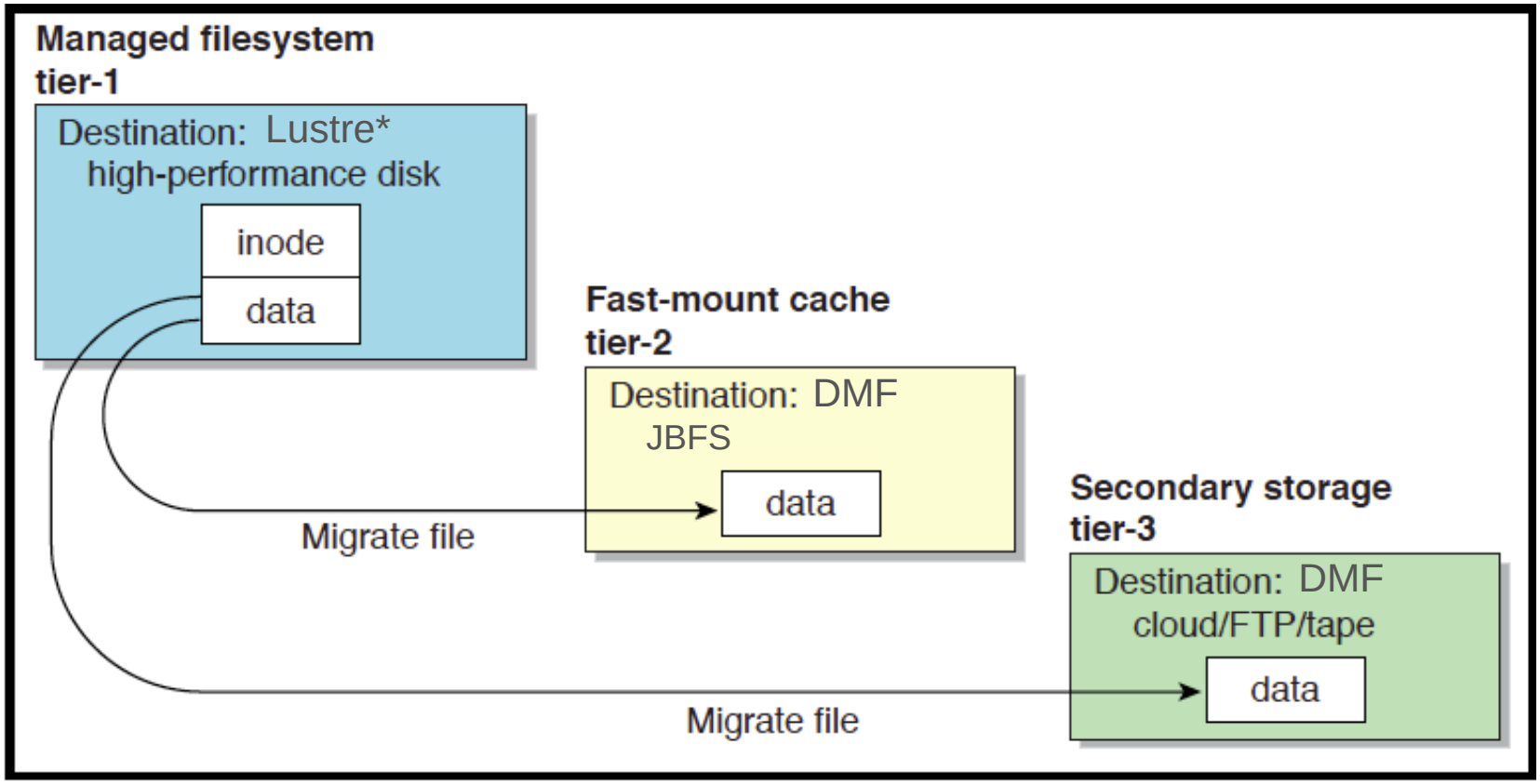


Tiered Data Management

HSM perspective: **dual-state** file

User perspective: **online** file

After migrating

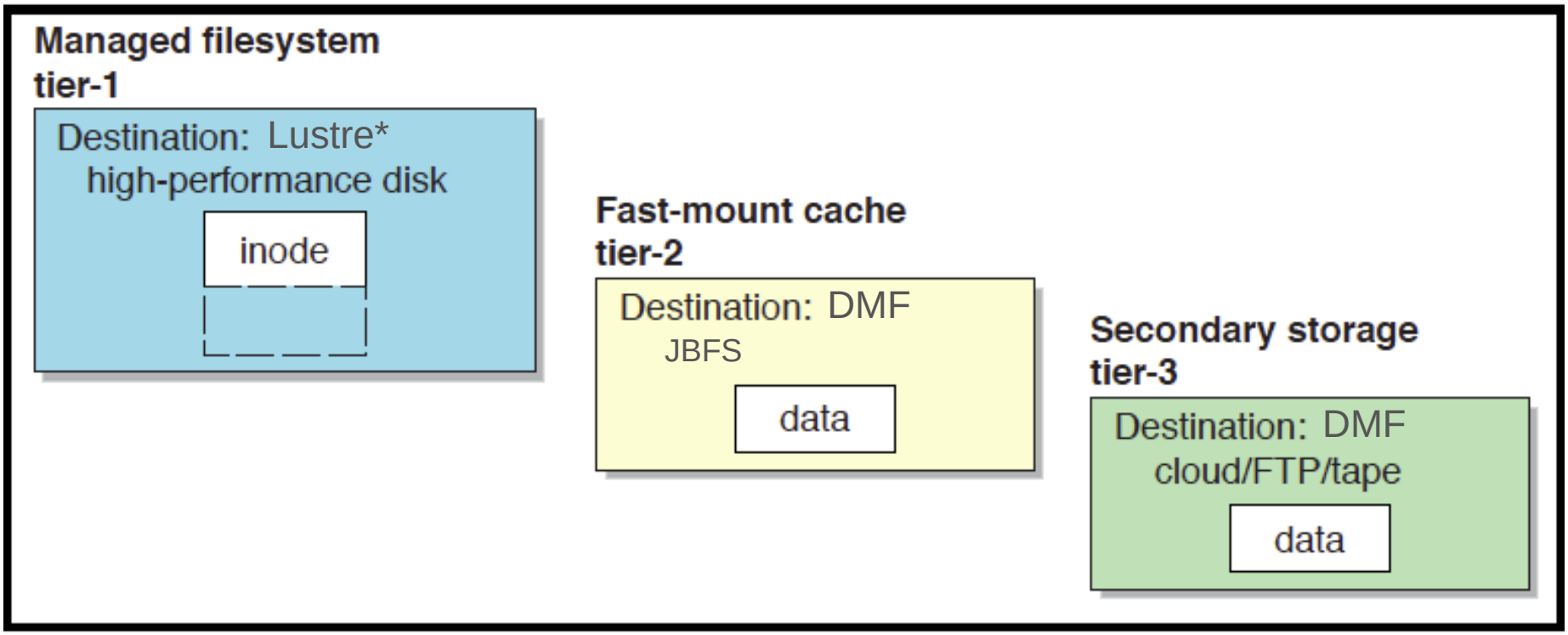


Tiered Data Management

HSM perspective: **offline** file

User perspective: **online** file

After freeing space

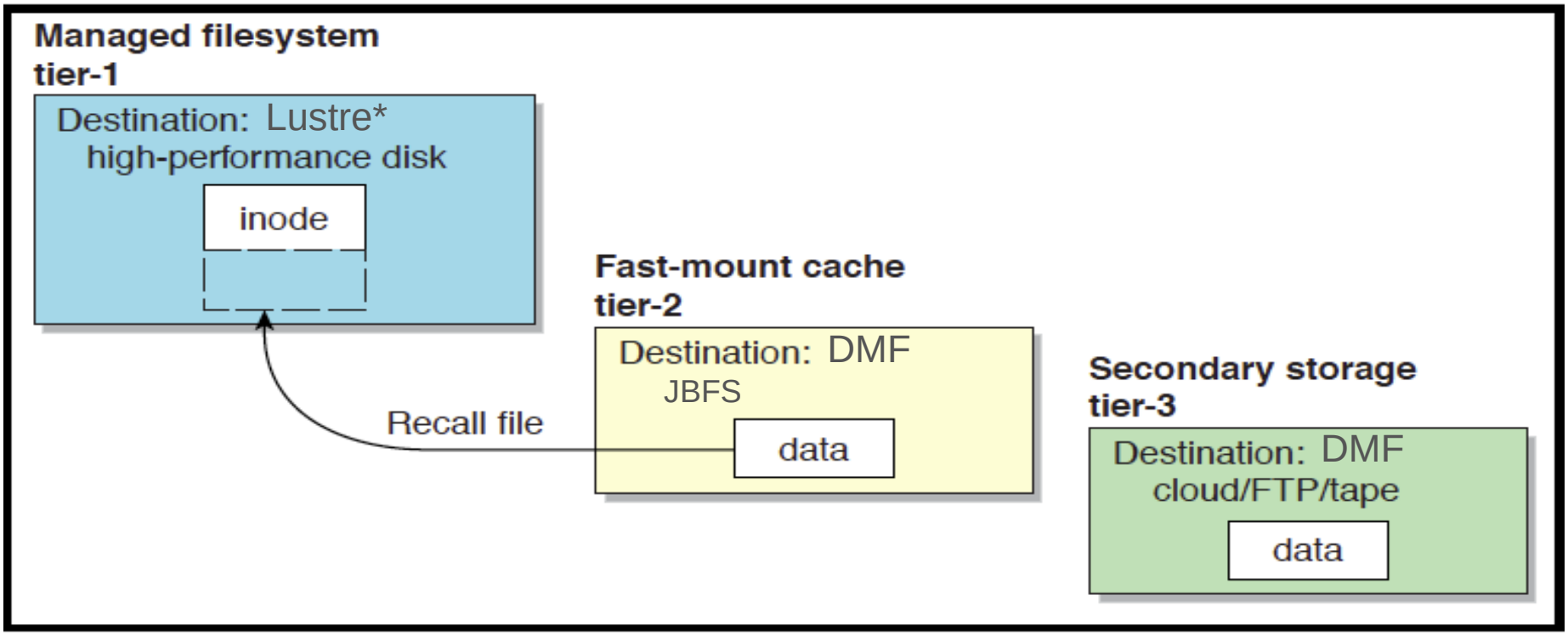


Tiered Data Management

HSM perspective: **unmigrating** file

User perspective: **online** file

Recalling file data from cache

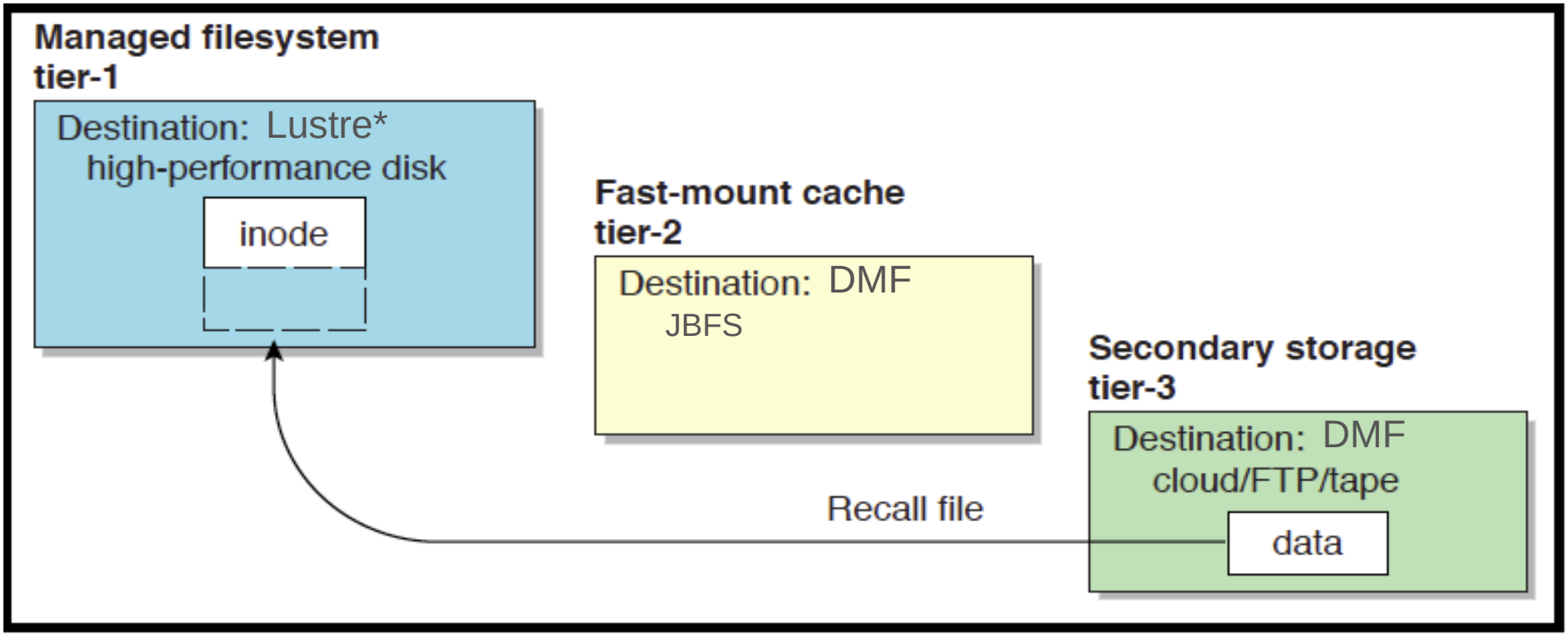


Tiered Data Management

HSM perspective: **unmigrating** file

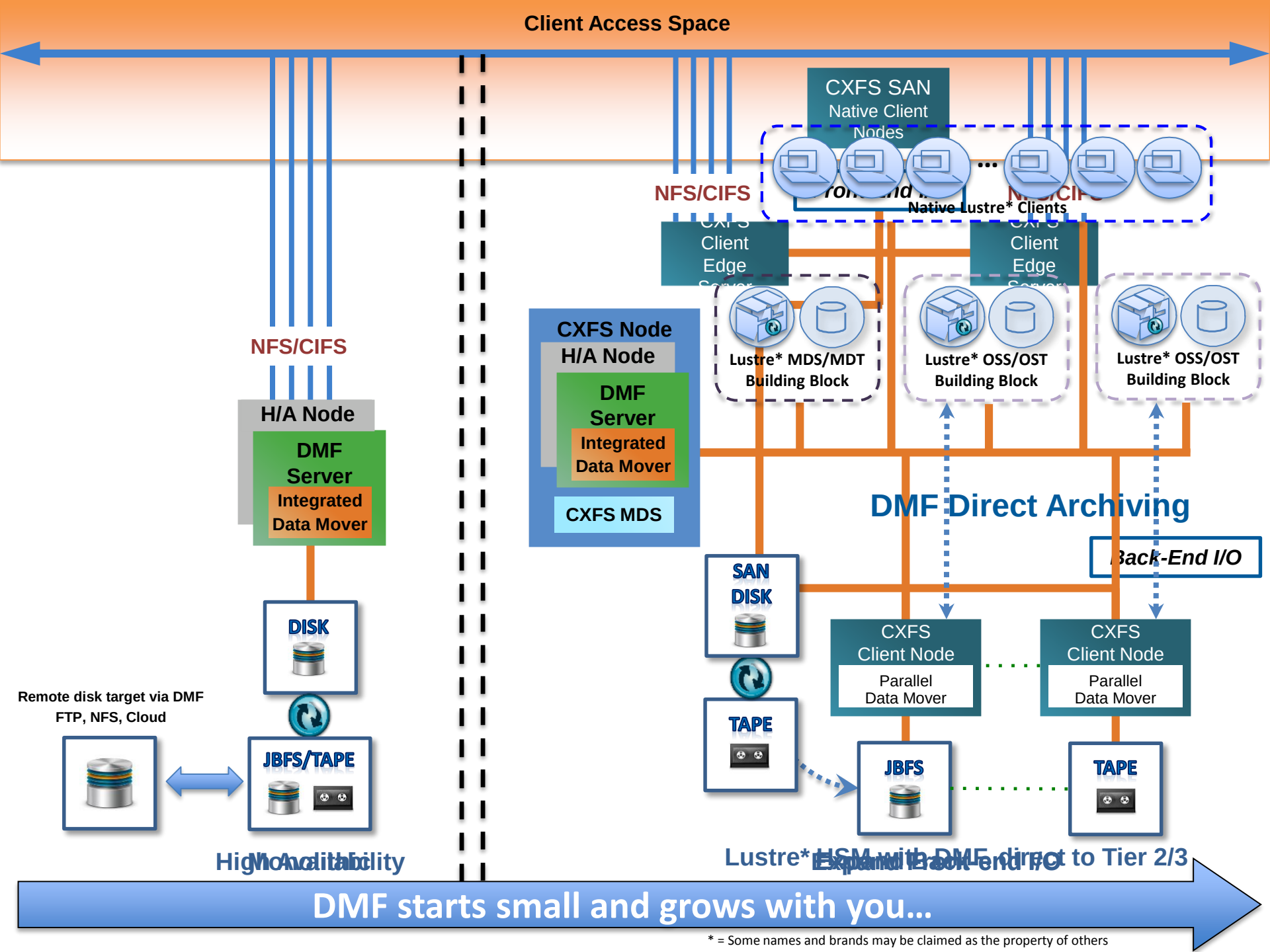
User perspective: **online** file

Recalling file data from cache

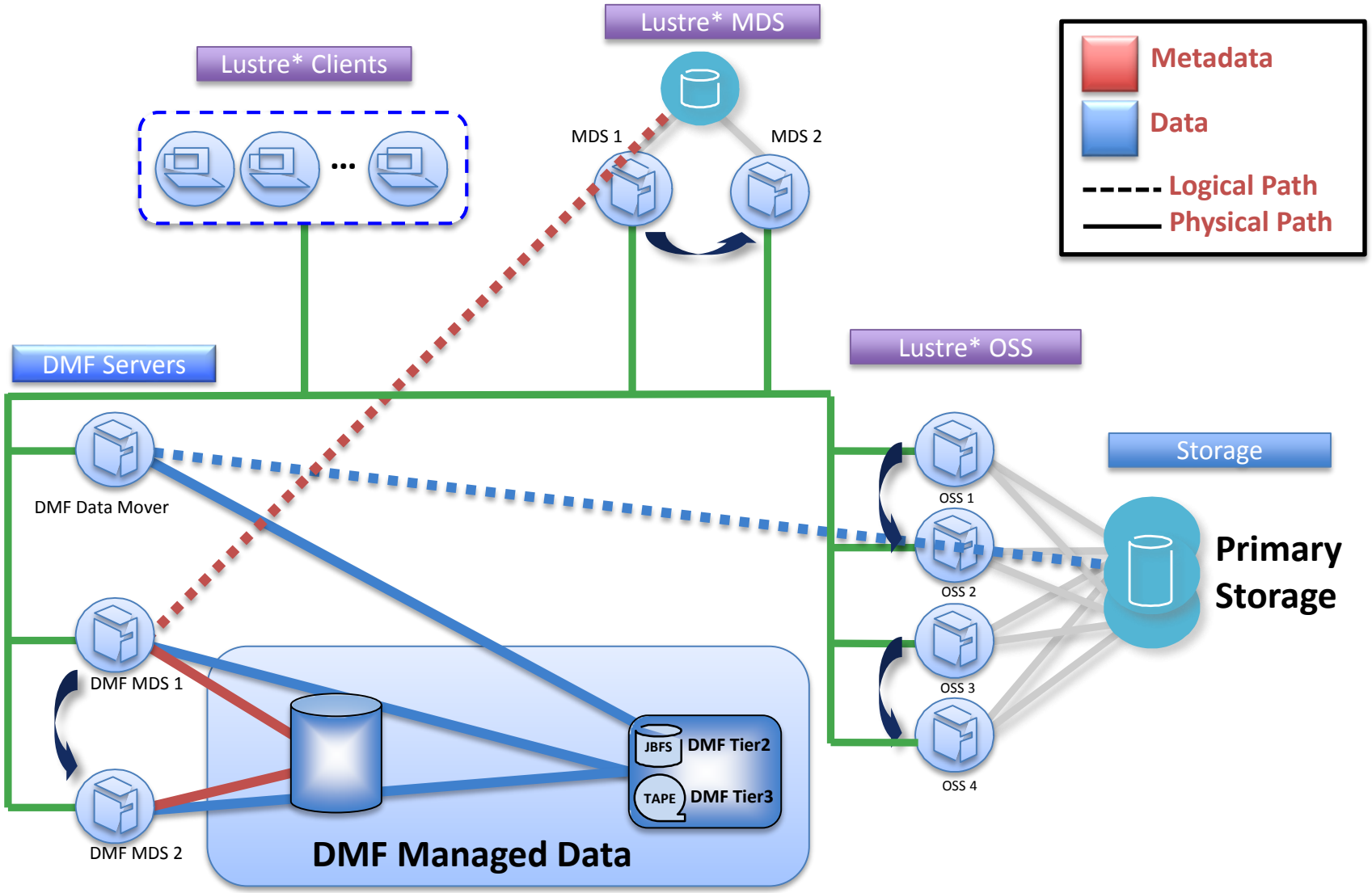


DMF Evolution

Start small and grow

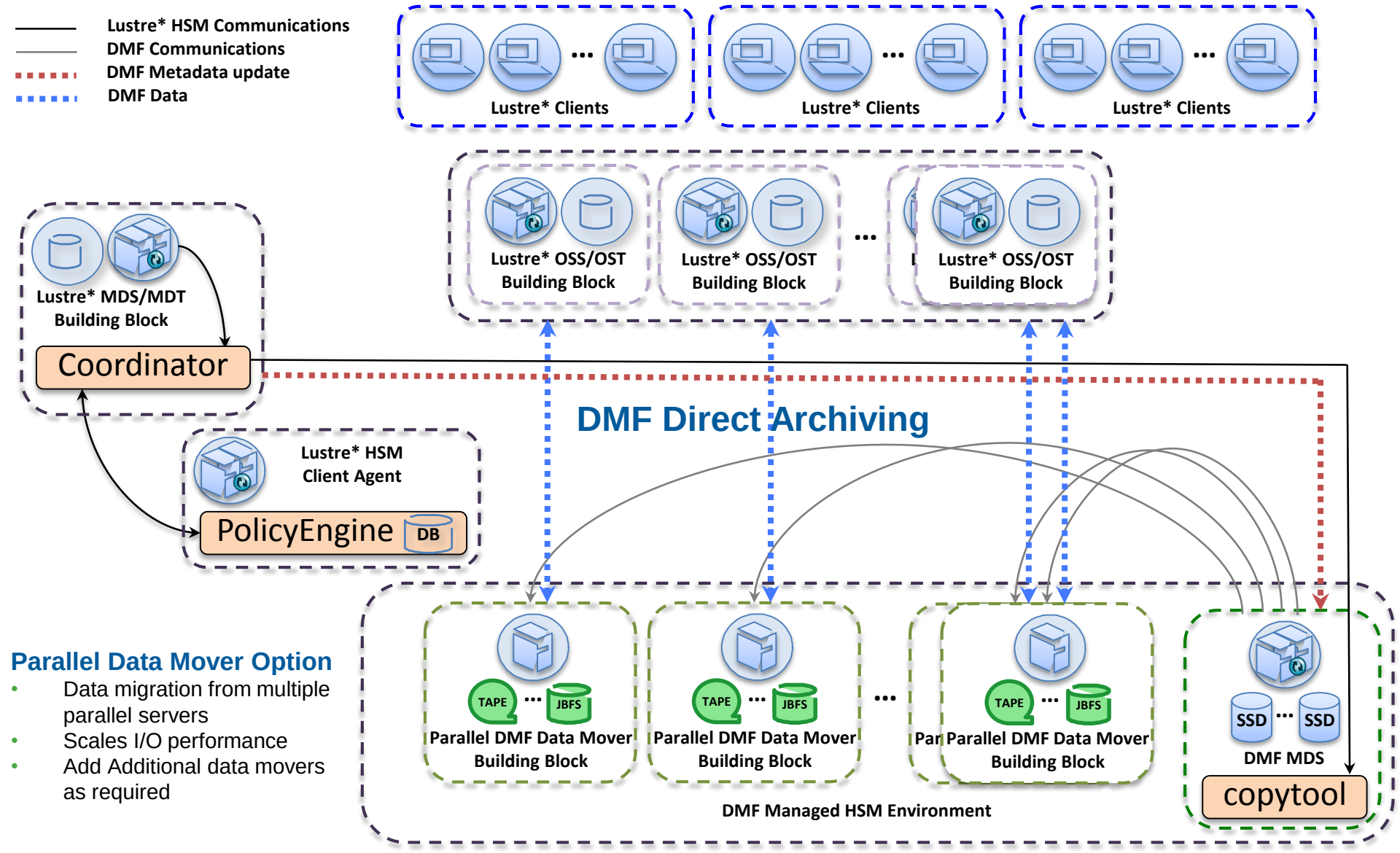

Client Access Space
NFS/CIFS
H/A Node
DMF Server
Integrated Data Mover
DISK
Remote disk target via DMF
FTP, NFS, Cloud
JBFS/TAPE
High Availability
CXFS Node
H/A Node
DMF Server
Integrated Data Mover
CXFS MDS
NFS/CIFS
CXFS SAN
Native Client Nodes
Native Lustre* Clients
CXFS Client Edge Server
CXFS Client Edge Server
Lustre* MDS/MDT Building Block
Lustre* OSS/OST Building Block
Lustre* OSS/OST Building Block
DMF Direct Archiving
Back-End I/O
SAN DISK
TAPE
CXFS Client Node
Parallel Data Mover
CXFS Client Node
Parallel Data Mover
JBFS
TAPE
Lustre* HSM with DMF direct to Tier 2/3
DMF starts small and grows with you…
* = Some names and brands may be claimed as the property of others

DMF Direct Archiving | Data Flow



Lustre* HSM | Communication & Data Flow

* = Some names and brands may be claimed as the property of others

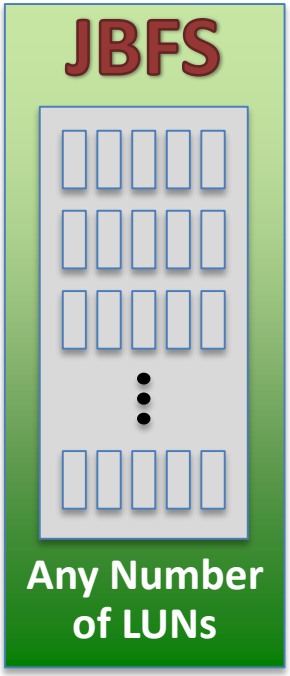


Parallel Data Mover Option

- Data migration from multiple parallel servers
- Scales I/O performance
- Add Additional data movers as required

JBFS | The OpenVault VTL for DMF

- JBFS is an acronym for JBOD File System
- JBFS provides mounting services
  - Serialised access to disk media
  - Independent from Linux disk mounts and file systems
- Disks treated like tapes mounted in tape drives
- The primary advantages
  - Mount performance
  - Low-cost scalable data throughput performance
  - Power management via ZeroWatt™



JBFS | Why A New File System?

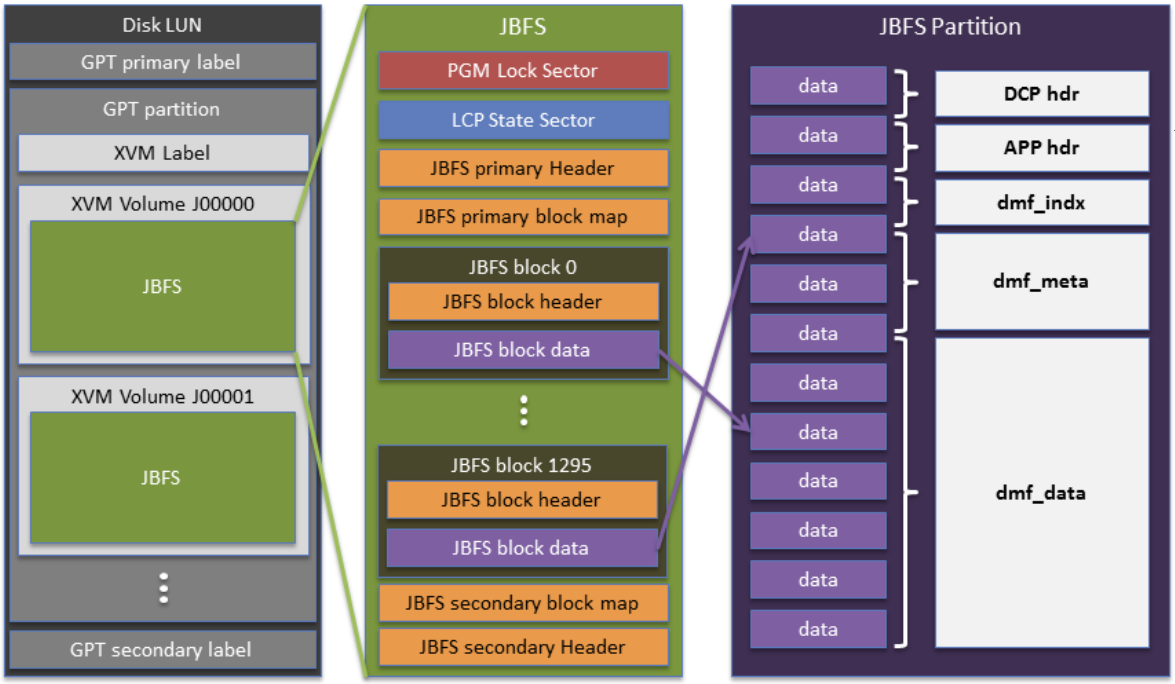
XFS (or any typical file system)	DMF Preferred	JBFS Provides
x A large number of objects	Small number of objects	✓ Small number of objects
x Object sizes change	Object sizes fixed	✓ Object sizes fixed
x Flexible object organisation	Fixed object organisation	✓ Fixed object organisation
x Primarily random access	Primarily sequential access	✓ Primarily sequential access
x Bursty access	Sustained access	✓ Sustained access
x Mount/dismount infrequently	Mount/dismount frequently	✓ Mount/dismount frequently

JBFS | Additional Benefits

- Recoverability
- Data Assurance
- High Performance
- Flexibility
- Power Management with Zero-Watt™
- JBFS API (Same as SGI Copan)

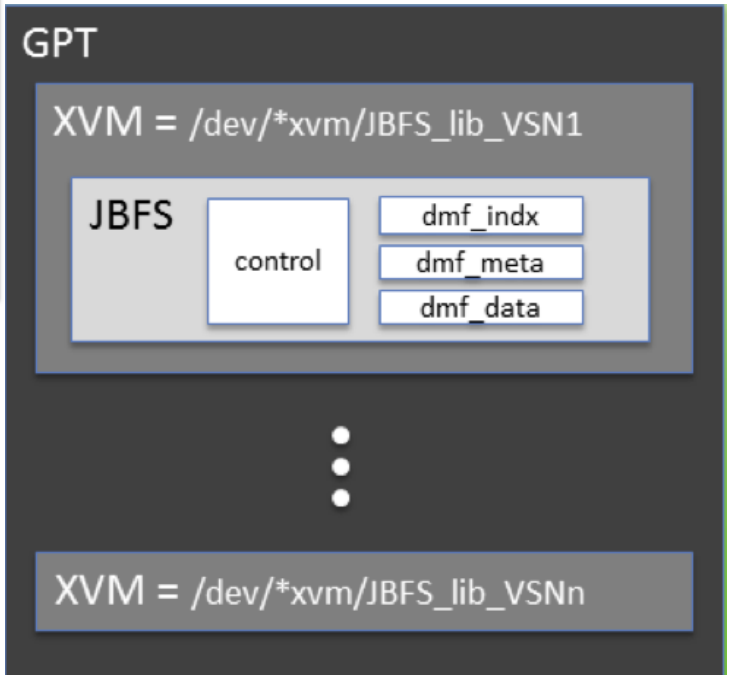
✧OpenVault includes a new DCP and a new LCP to manage JBFS volumes

JBFS | Disk Structure



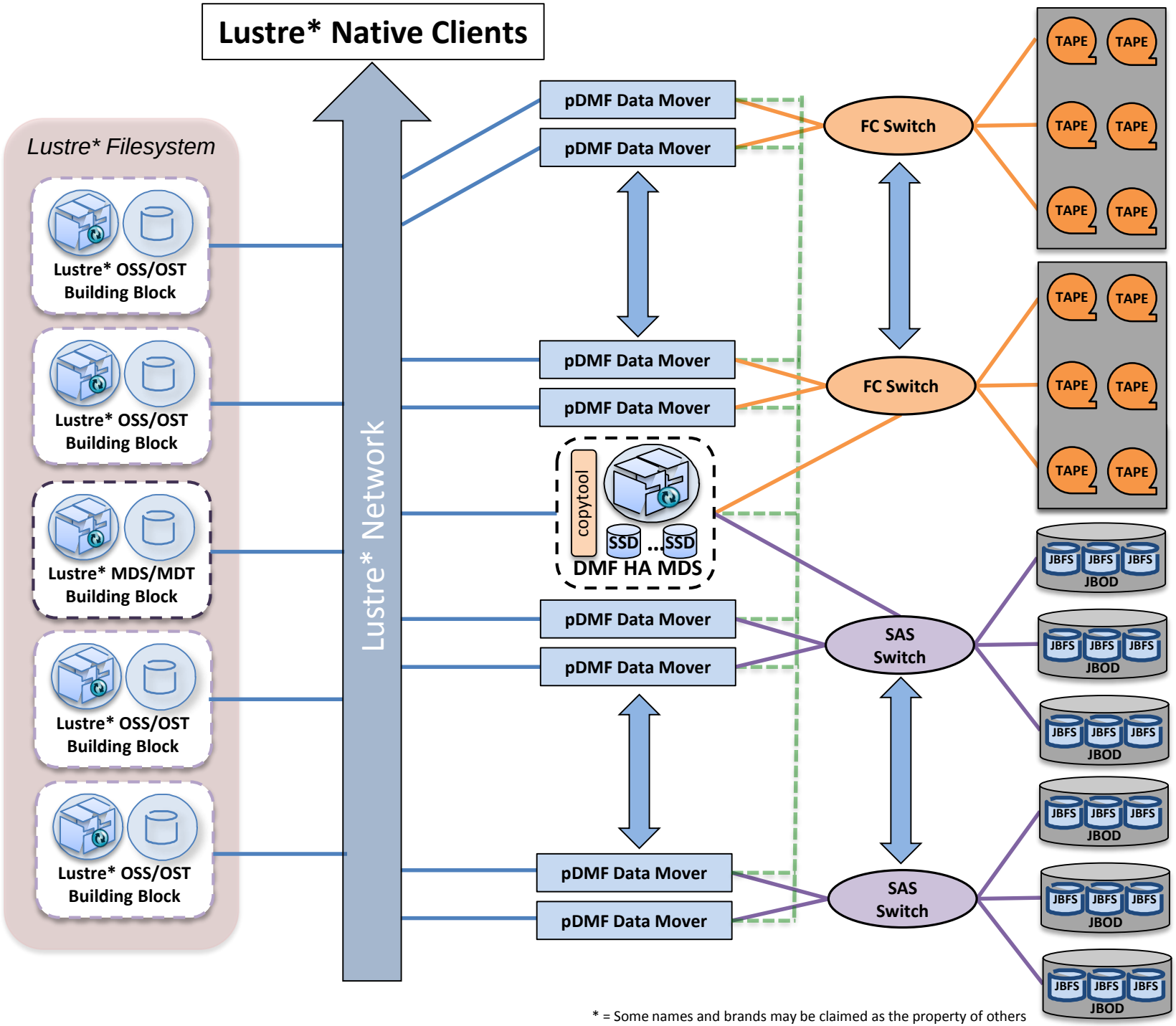
XVM volume name must be "JBFS_{lib}_{PCL}"

- JBFS – Fixed
- {lib} – OpenVault Library Name
- {PCL} – unique 6-character value [0-9A-Z]



Preparing a disk device for use with JBFS consists of three basic steps:

1. Apply the GPT labels
2. Apply the XVM labels
3. Apply the JBFS format


Lustre* Native Clients
Lustre* Filesystem
Lustre* OSS/OST Building Block
Lustre* OSS/OST Building Block
Lustre* MDS/MDT Building Block
Lustre* OSS/OST Building Block
Lustre* OSS/OST Building Block
Lustre* Network
pDMF Data Mover
pDMF Data Mover
pDMF Data Mover
pDMF Data Mover
copytool
SSD
SSD
DMF HA MDS
pDMF Data Mover
pDMF Data Mover
pDMF Data Mover
pDMF Data Mover
FC Switch
FC Switch
SAS Switch
SAS Switch
TAPE
JBFS
JBOD
* = Some names and brands may be claimed as the property of others

Summary and Key Points

- New Lustre* and DMF features allow cost effective scalability without compromising performance
- SGI DMF provides a high performance parallel HSM for Lustre* with direct archiving to tier 2/3 storage targets
- SGI DMF – JBFS delivers a tier 2 fast mount cache with built in power management$ capabilities
- The Result:
  - Cost effective capacity, reduced TCO (low cost/power storage tiers)
  - Proven long-term data protection (DMF – 25 years in production)
  - Improved operational procedures (simplified access to data)
  - Scalable performance within archive tiers (parallel DMF)

$ = on supported hardware

Questions & Responses

Robert Mollard
Senior Storage Specialist, Asia Pacific
rmollard@sgi.com