Lustre and Robinhood

Joseph C. Nemeth

Technical Lead
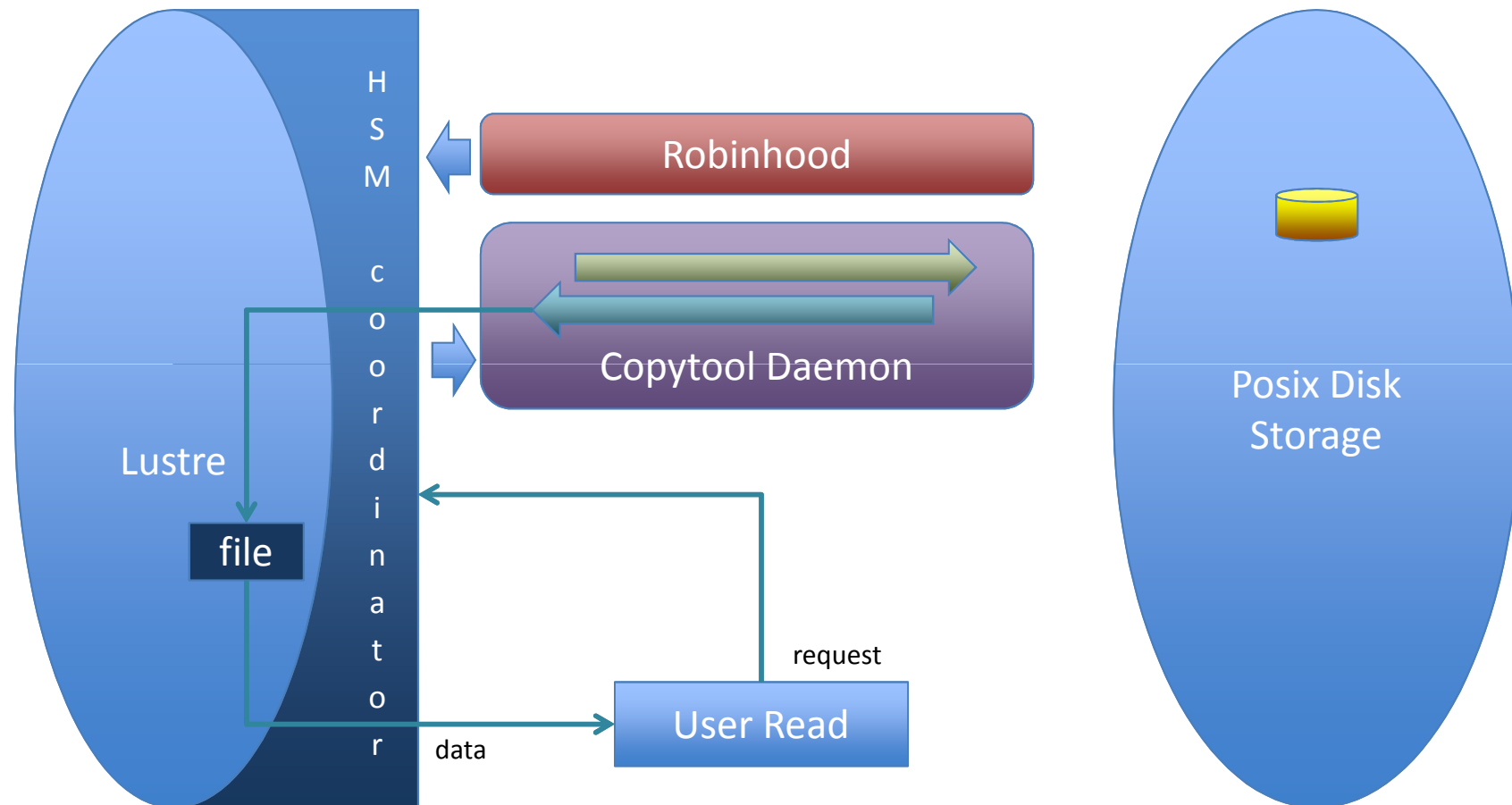
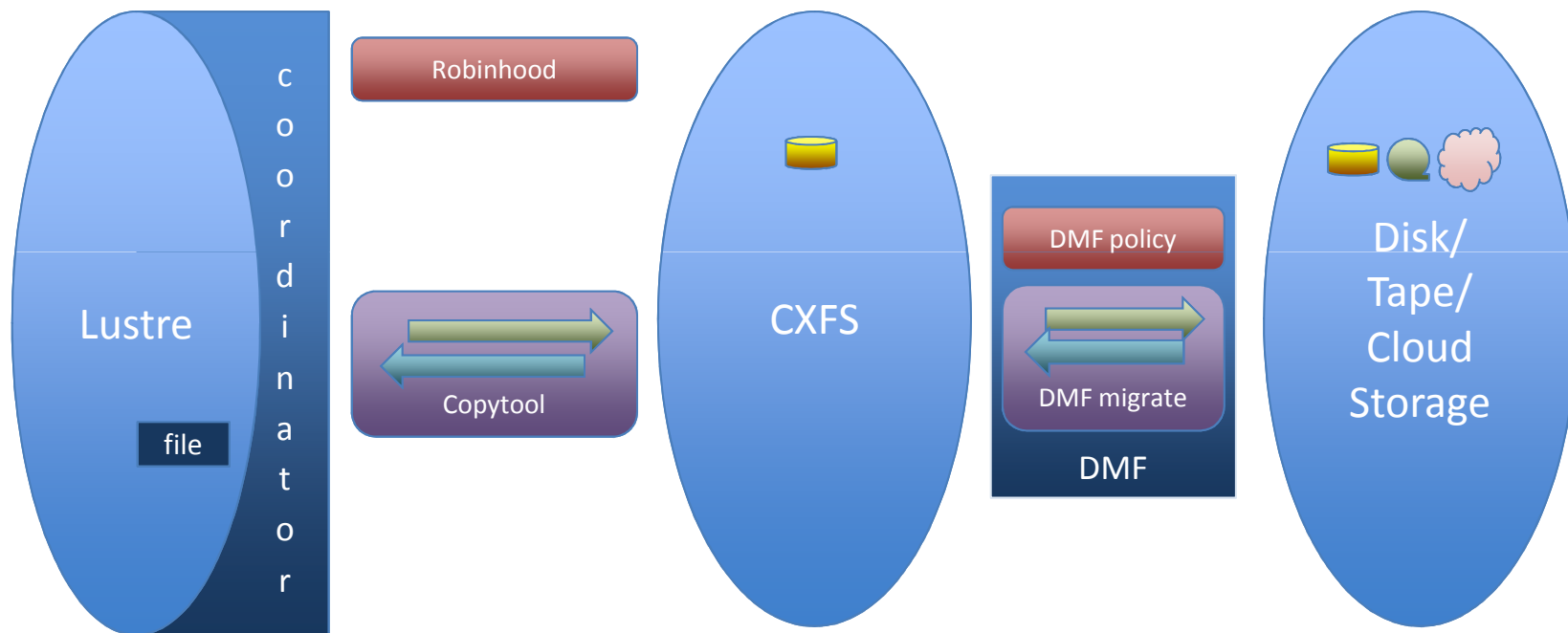# The DMF Model

CXFS

file

DMF policy engine

DMF migrate engine

DMF

DMAPI

request
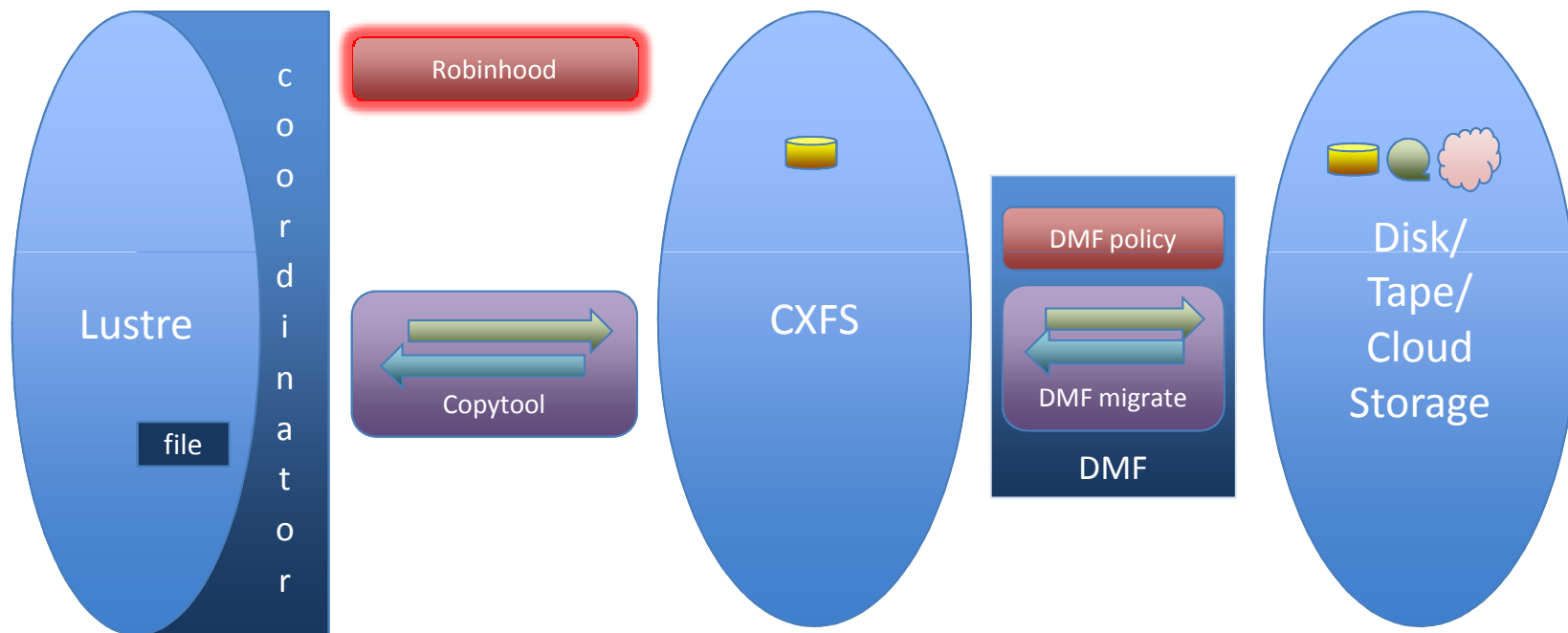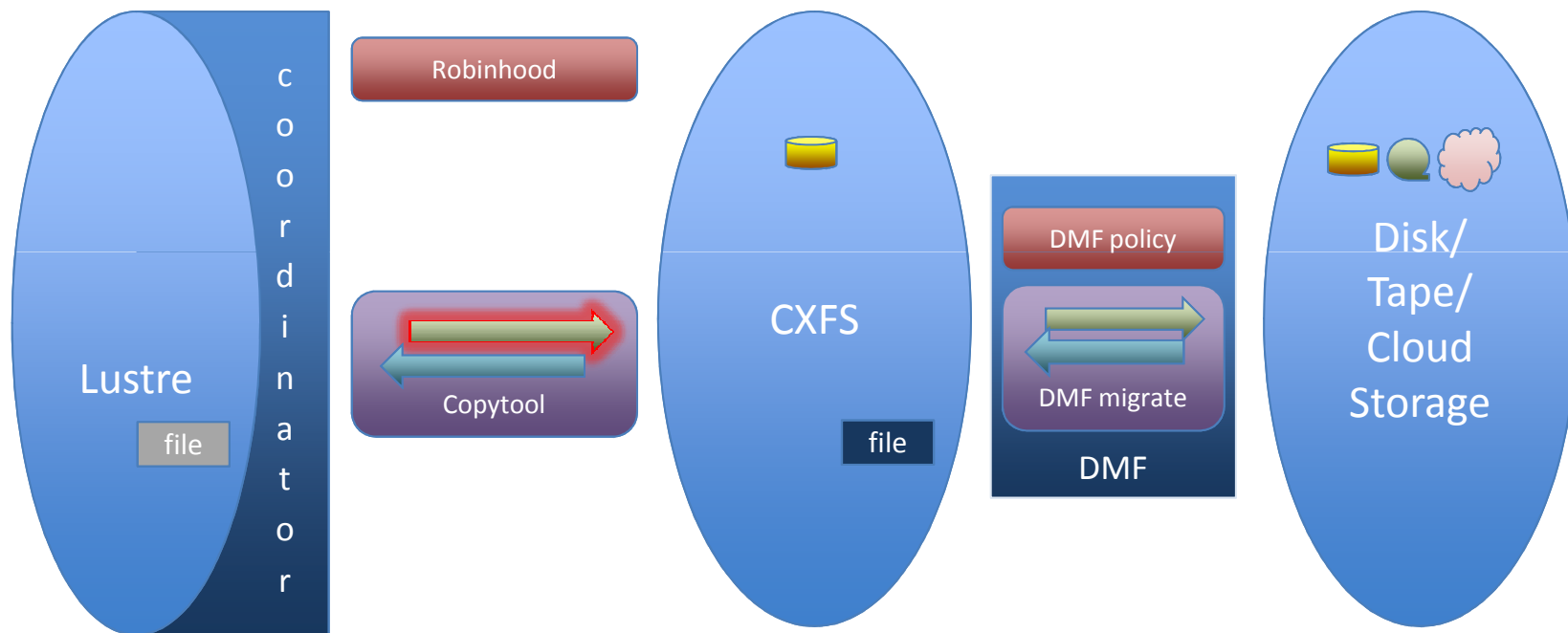
User Read

data

Disk/
Tape/
Cloud
Storage

sgi

The Lustre Model

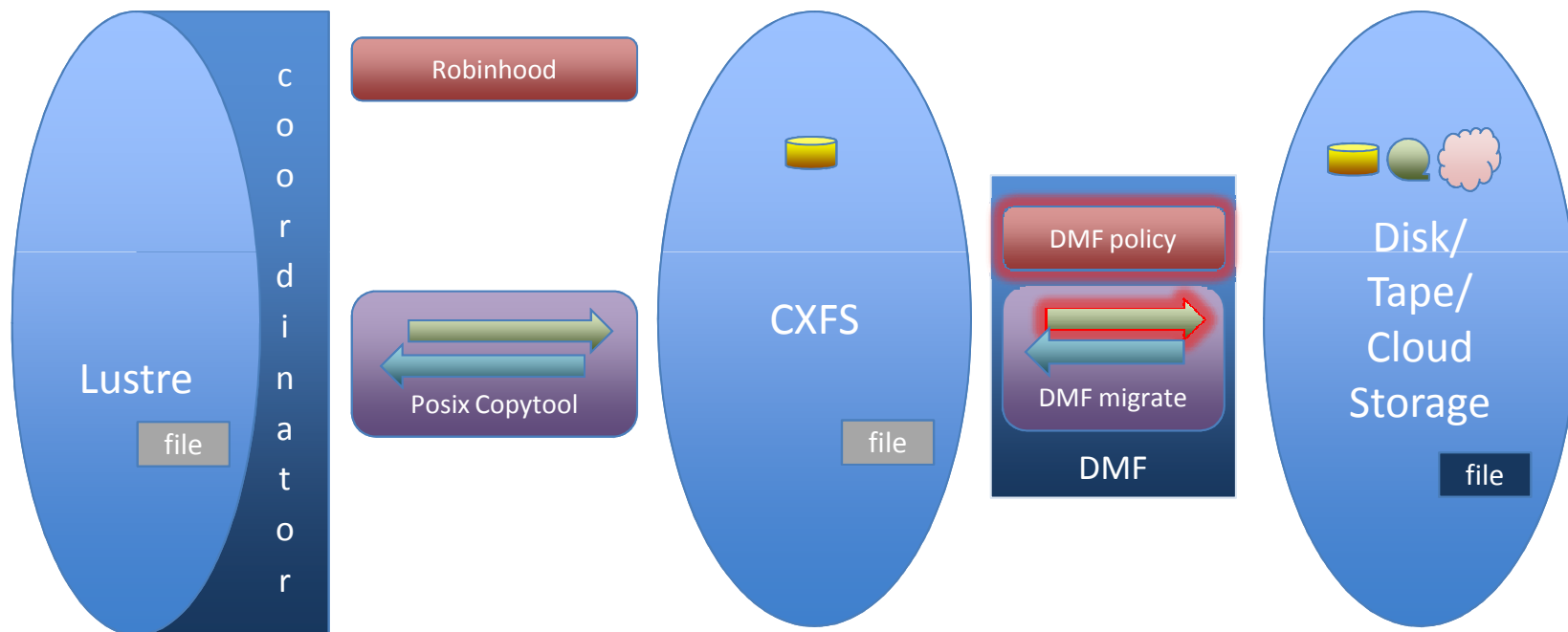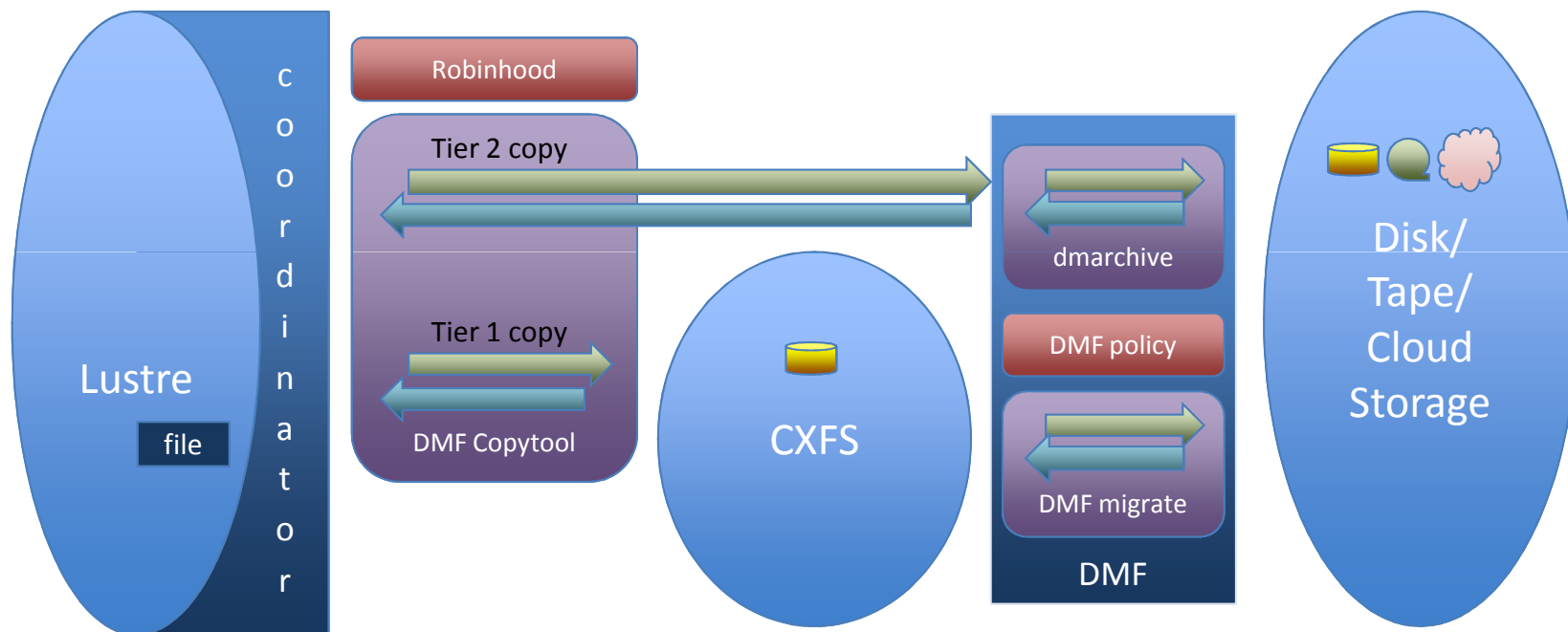# Simple Lustre/DMF Model

Simple Lustre/DMF Model

Simple Lustre/DMF Model

sgi

Simple Lustre/DMF Model

Better Lustre/DMF Model
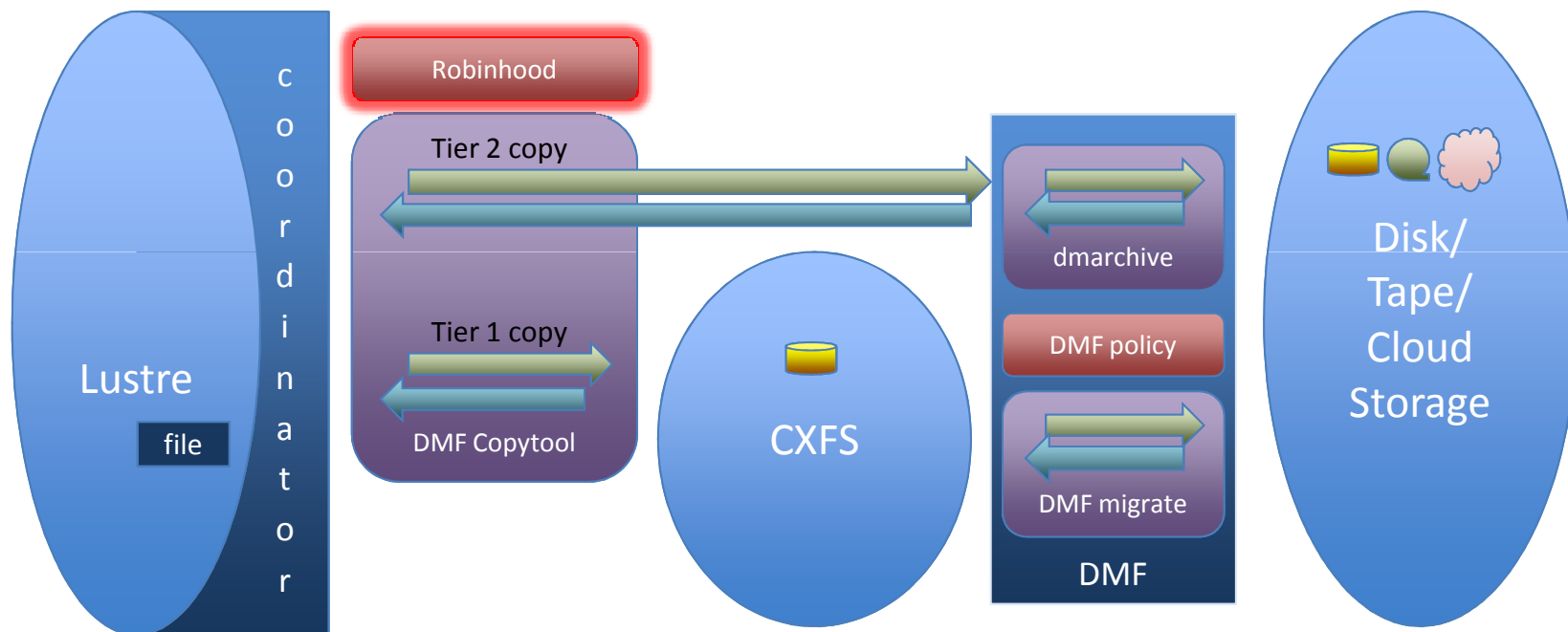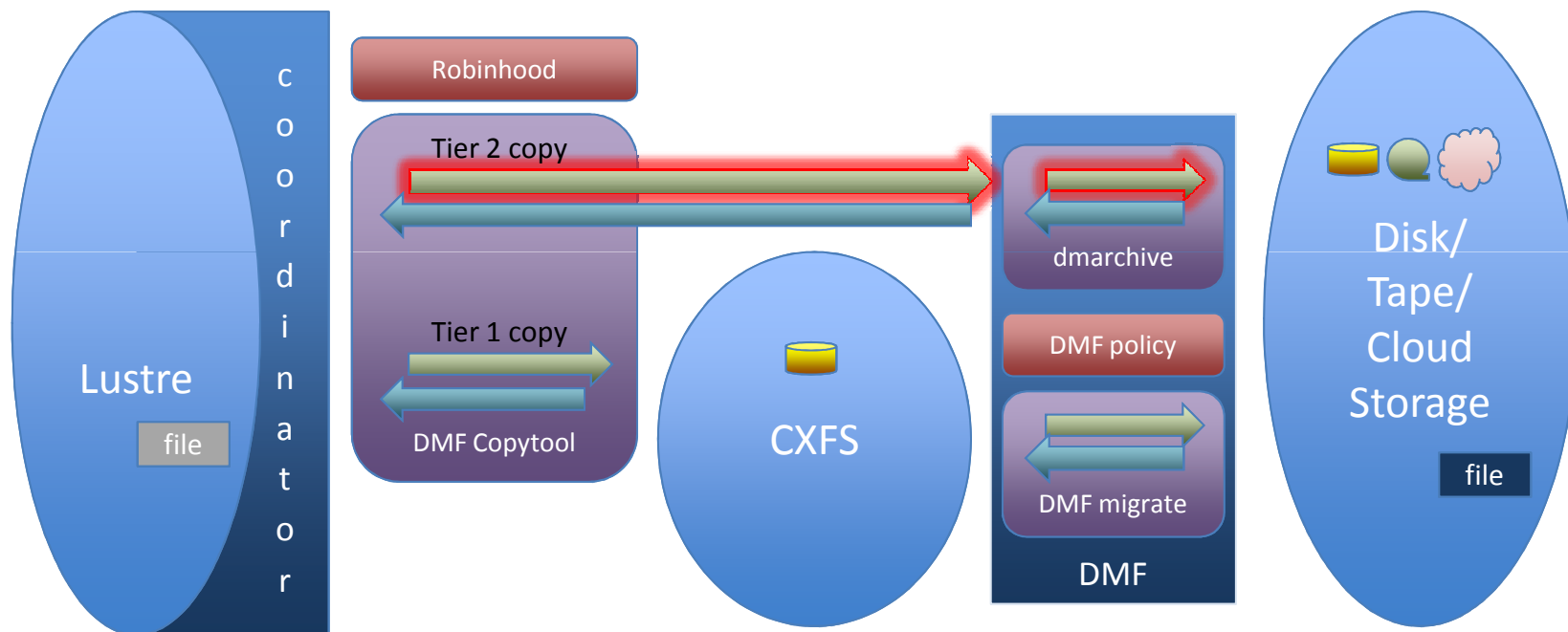
Better Lustre/DMF Model

Better Lustre/DMF Model

Better Lustre/DMF Model

Better Lustre/DMF Model

Lustre

file

coordinator

Robinhood

Tier 2 copy

Tier 1 copy

DMF Copytool

CXFS

file

dmarchive

DMF policy

DMF migrate

DMF

Disk/
Tape/
Cloud
Storage

Better Lustre/DMF Model

Lustre

file

coordinator

Robinhood

Tier 2 copy

Tier 1 copy

DMF Copytool

CXFS

file

dmarchive

DMF policy

DMF migrate

DMF

Disk/ Tape/ Cloud Storage

file

Basic DMF Configuration

OSS nodes

MDS nodes

Lustre client nodes

Lustre

Robinhood   SSD

copytool   Tier 1   XFS

Tier 2

DMF srv   dmarchive
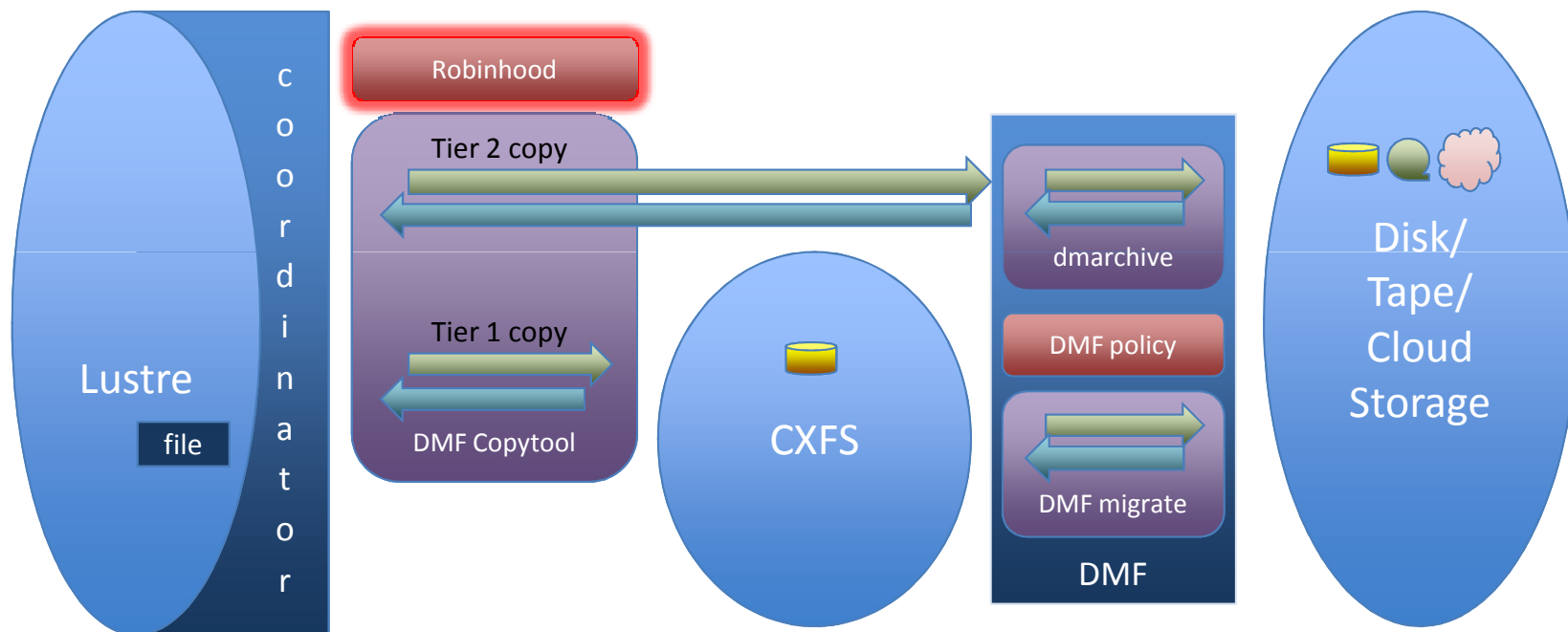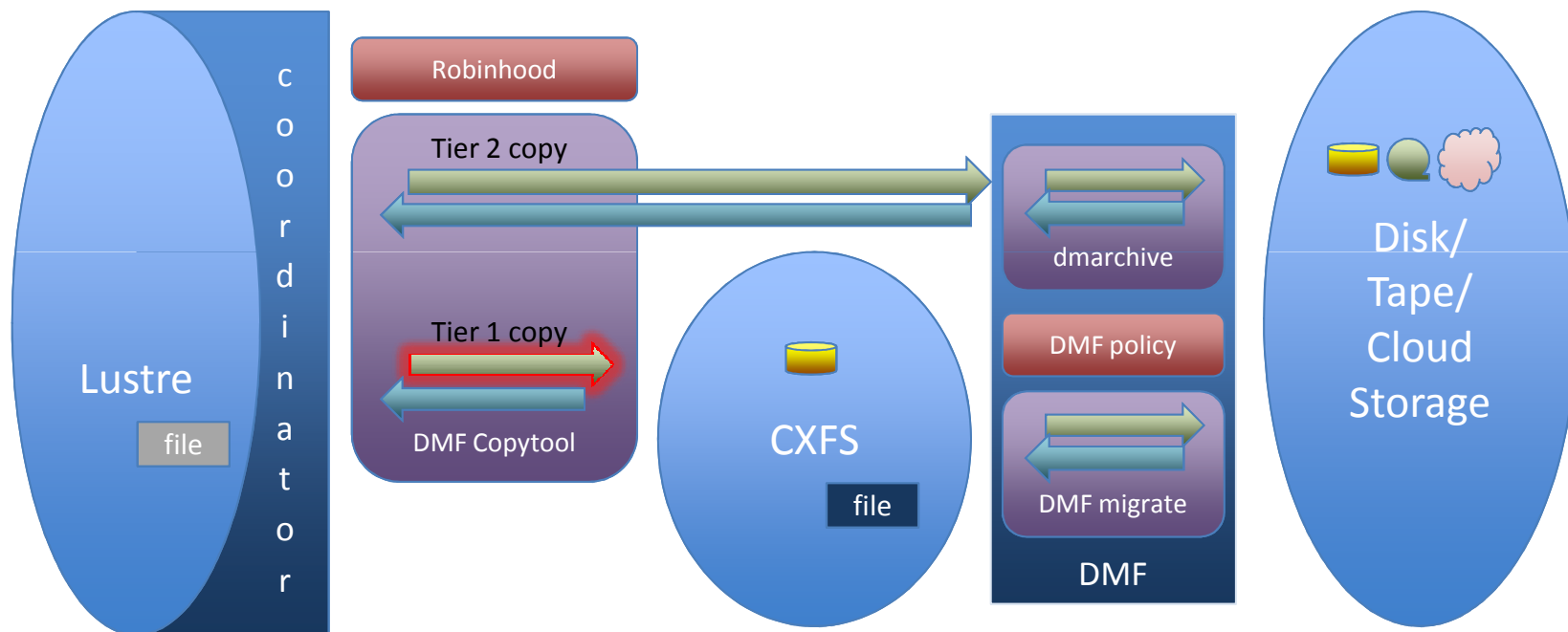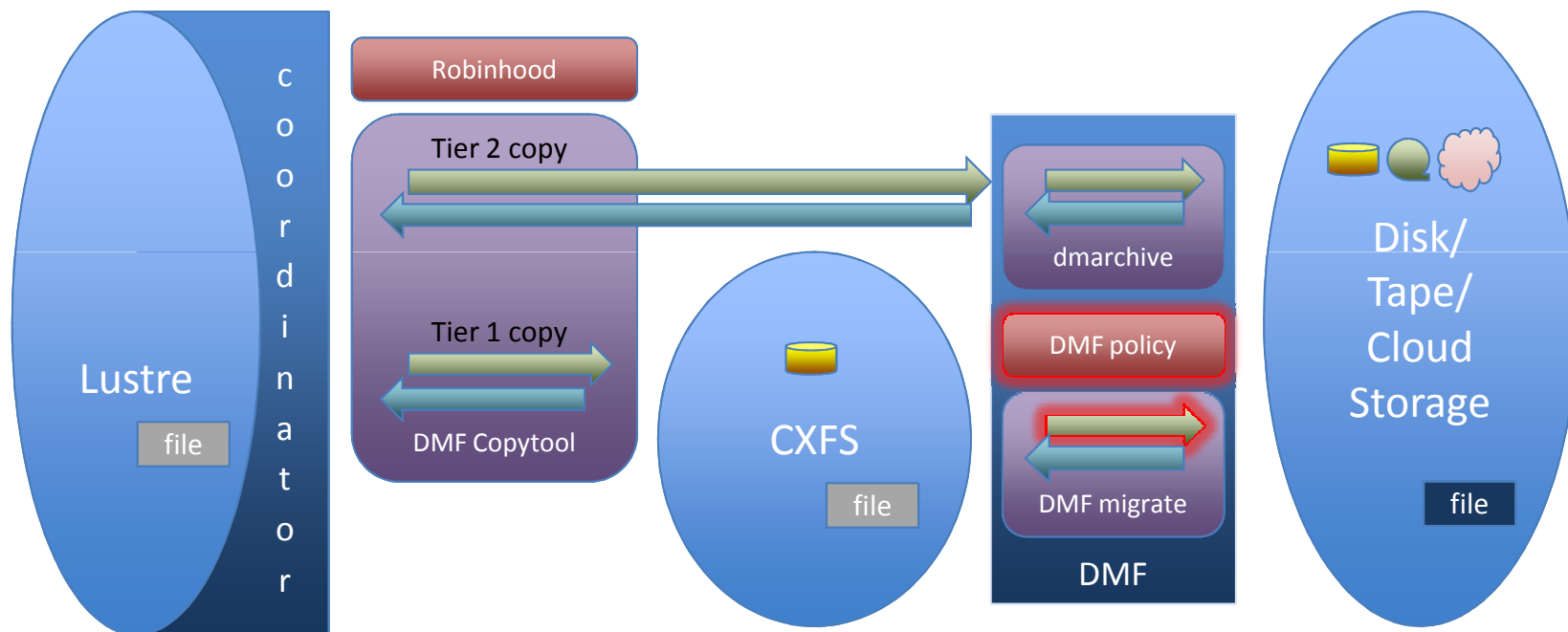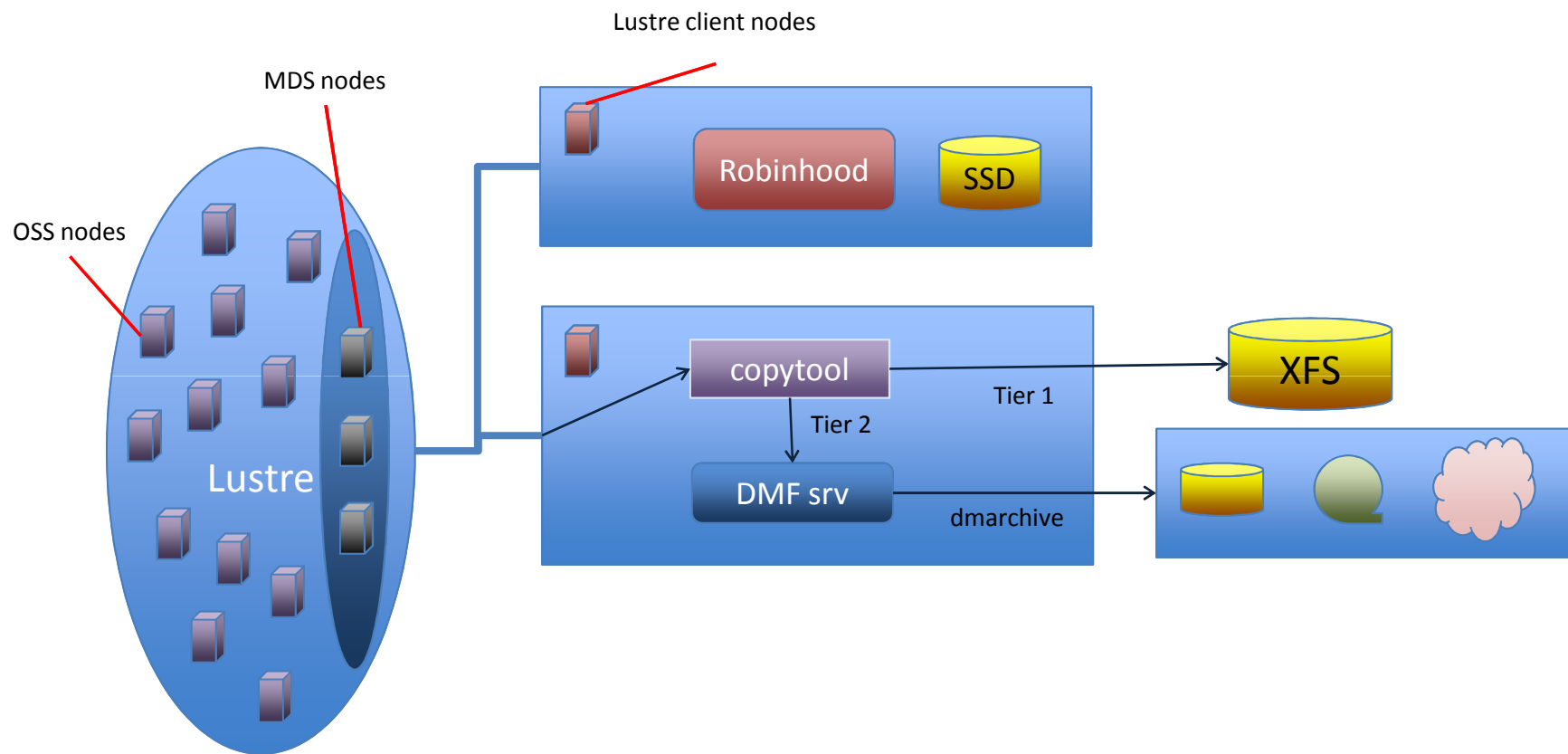
Parallel DMF Configuration
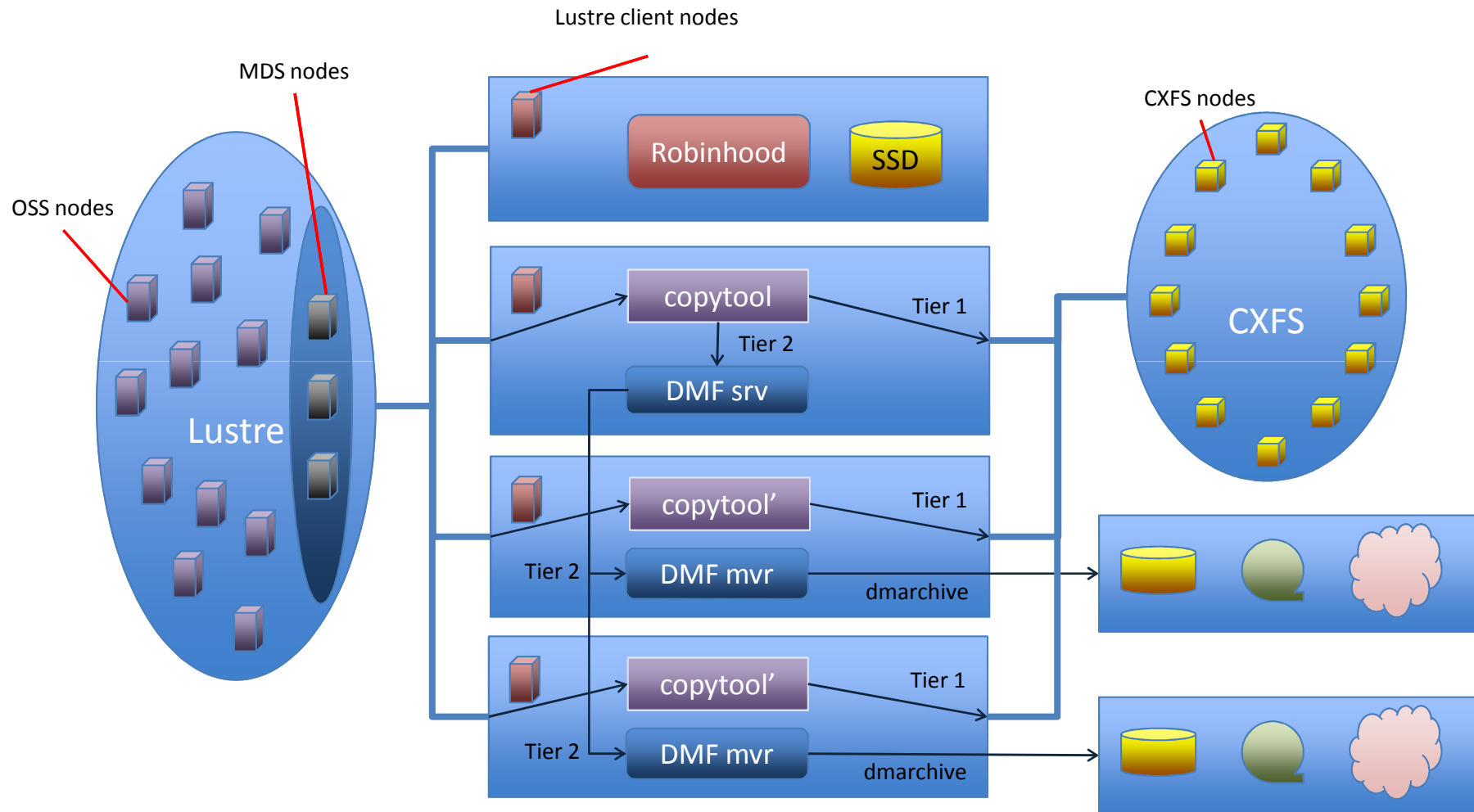
Tuning Lustre

- max_requests
  - maximum concurrent copy operations

- active_request_timeout
  - time to wait before giving up on a non-responsive copy operation

- policy
  - NBR (Non-Blocking Restore)
  - NRA (No Retry Action)

Tuning the Copytool

- --archive
  - advanced use of archive numbers
- --bandwidth
  - total bandwidth limitation
- --chunk-size
  - match to Lustre stripe size
- --abort-on-error
  - useful when troubleshooting
- --update-interval
  - frequency with which to update the coordinator

Disaster Recovery

- What if the Lustre FS is destroyed?

- The archive is flat, with coded file names

- The archive shadow directory

- Copytool option --import

- Files are created in OFL state

Robinhood

- First-time Lustre FS scan

- The Lustre changelog

- Differences between DMF and Robinhood
  - DMF is a free-space manager with archiving
  - Robinhood is an archiver with free-space mgmt

Robinhood Actions

- Archive files ("migrate")
  - copy files from Lustre to DMF, putting files in the DUL state

- Release files ("purge")
  - release Lustre data, putting files in the OFL state

- Remove files
  - delete stale archive copies

Robinhood Archive Policies

- Periodic scan for archive candidates
  - period is configurable
- Files selected by per-file rules
  - rules same for archive and purge
- All files selected are copied
  - can be limited by file count
  - can be limited by total transfer size

sgi

Robinhood Purge Policies

- Periodic scan for "purge triggers"

- Four kinds of trigger

  – Lustre-wide

  – per-OST

  – per-user

  – per-group

- Trigger specifies max and min usage

- Files selected for purge by per-file rules

- File release count can be limited

Robinhood File Selection

- Properties
  - tree
  - name
  - fullpath
  - type
  - owner or group
  - size
  - last_access (age)
  - last_mod (age)
  - OST pool
  - xattr

- Operators
  - ==
  - !=
  - <
  - <=
  - >=
  - >

- Conjunctions
  - or
  - and
  - not
  - grouping

Robinhood Delete Policies

- Periodic scan for deletion

- "Grace period" for stale archive files

- File delete count can be limited