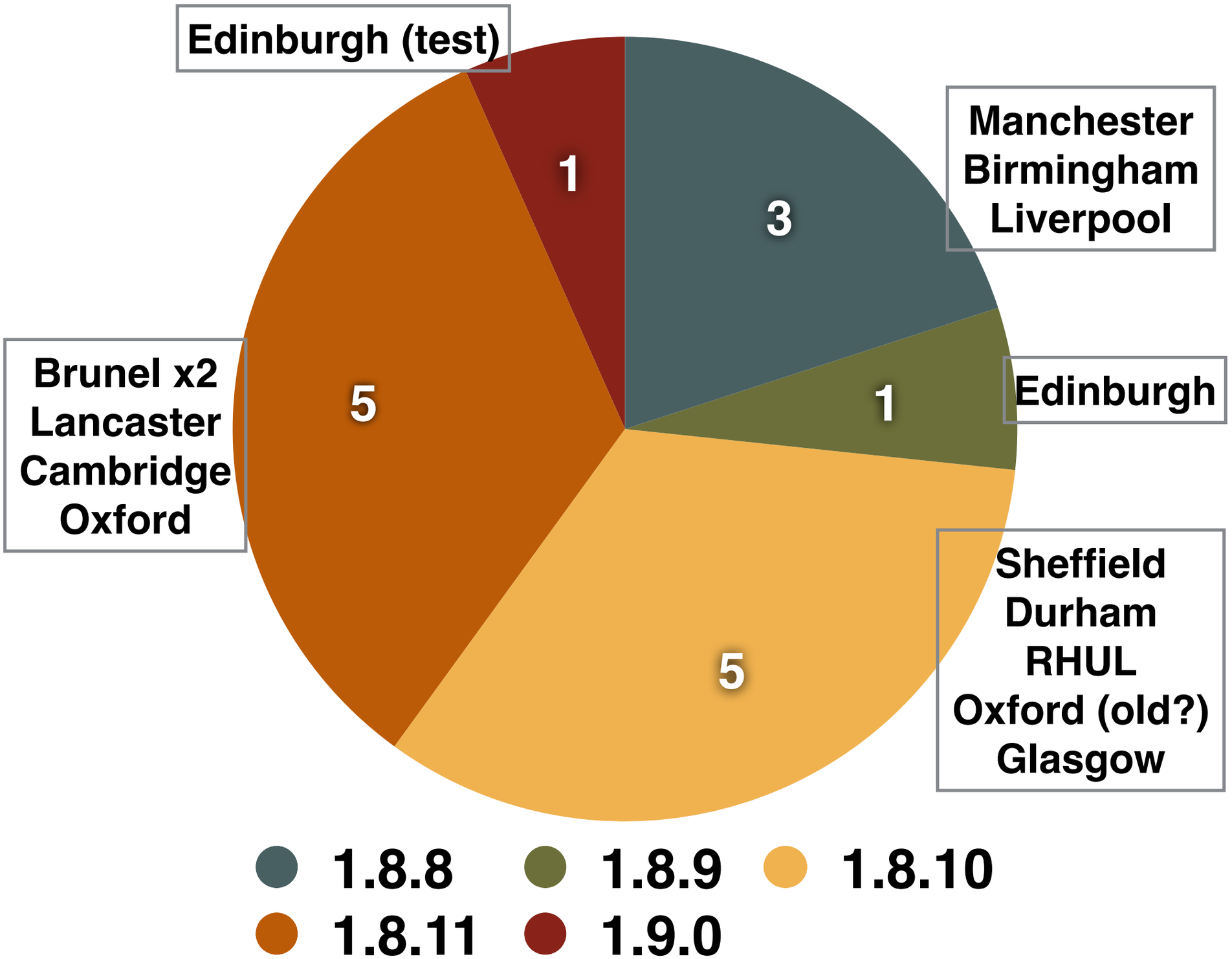
DPM
UK Status 2016

Sam Skipsey
on behalf of GridPP Storage Group

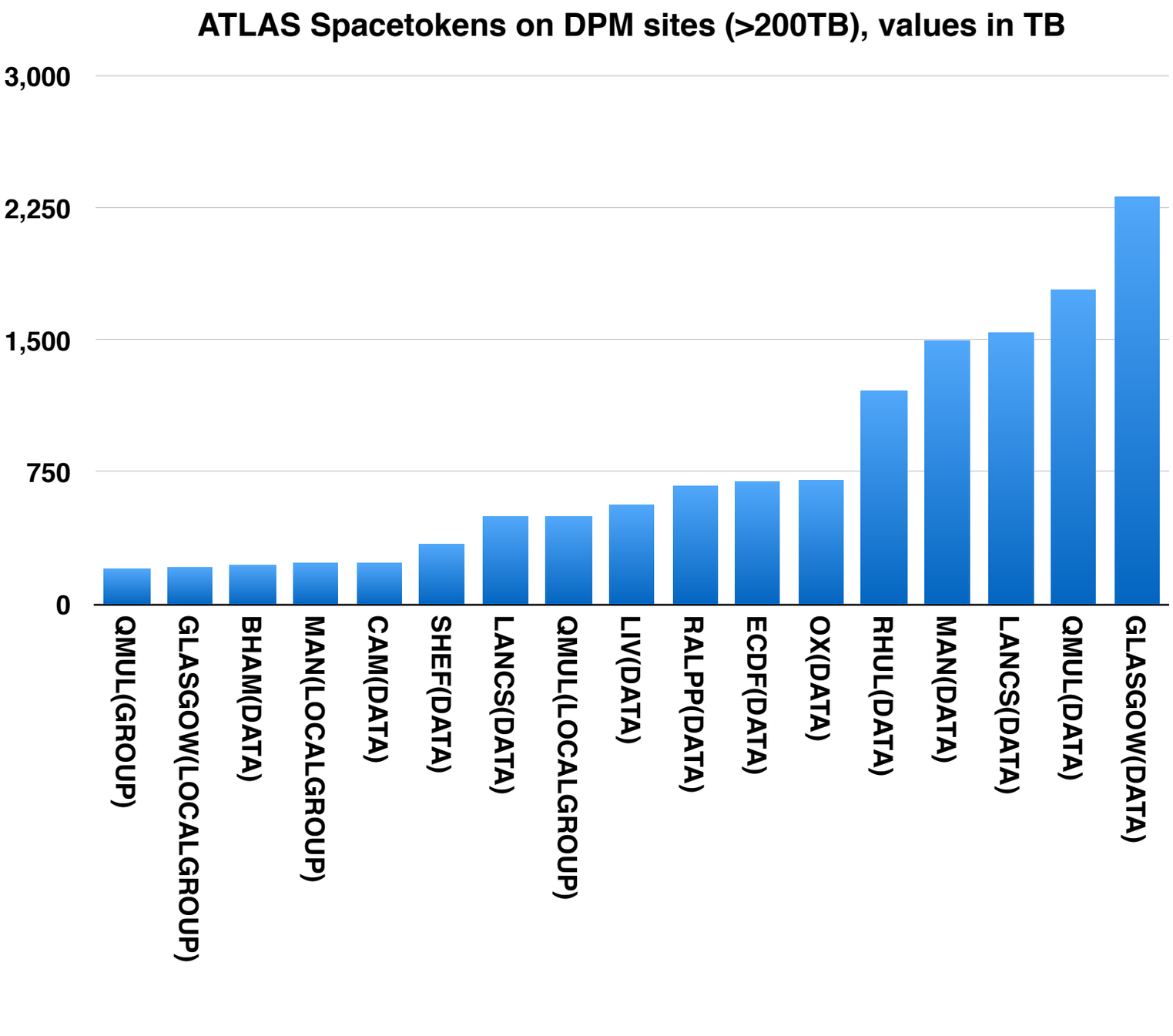
Itinerary

- Status
- Concerns
- Wishlists
- Future Look T2s

Status - DPM releases



Status - Storage Size



ATLAS Storage

Tier-2 site view (2016 pledge)

Site (click for detail)	Pledge (TB)	Actual (TB)	Not-free (TB)	Free (TB)	Usage
UKI-LT2-Brunel	0	33	28	5	
UKI-LT2-RHUL	963	1418	1218	201	
UKI-NORTHGRID-LANCS-HEP	1229	2100	1604	496	
UKI-NORTHGRID-LIV-HEP	486	735	567	168	
UKI-NORTHGRID-MAN-HEP	1080	1750	1546	204	
UKI-NORTHGRID-SHEF-HEP	387	393	323	70	
UKI-SCOTGRID-DURHAM	0	44	38	6	
UKI-SCOTGRID-ECDF	540	891	700	190	
UKI-SCOTGRID-GLASGOW	1305	2583	2178	405	
UKI-SOUTHGRID-BHAM-HEP	189	335	232	104	
UKI-SOUTHGRID-CAM-HEP	207	285	214	71	
UKI-SOUTHGRID-OX-HEP	729	823	711	112	
Tier-2 total	8987	14835	11876	2959	

Concerns and comments

- DPM Namespace is not reconstructible from PFNs
  - That is, the filesystem path (SFN) is abstract, and does not resemble the SURL path.
    - /dpm/gla.scotgrid.ac.uk/vo/dir1/myfile.stuff
    - disk001:/store1/vo/2016-11-26/myfile.stuff
- This is to allow multiple replicas (1 file in multiple locations) and decouple DPNS metadata from filesystem metadata for simplicity… but do we actually use that abstraction?

Concerns and comments

- "DPM fairly straightforward to manage, but perhaps puppet not the best solution post YAIM"

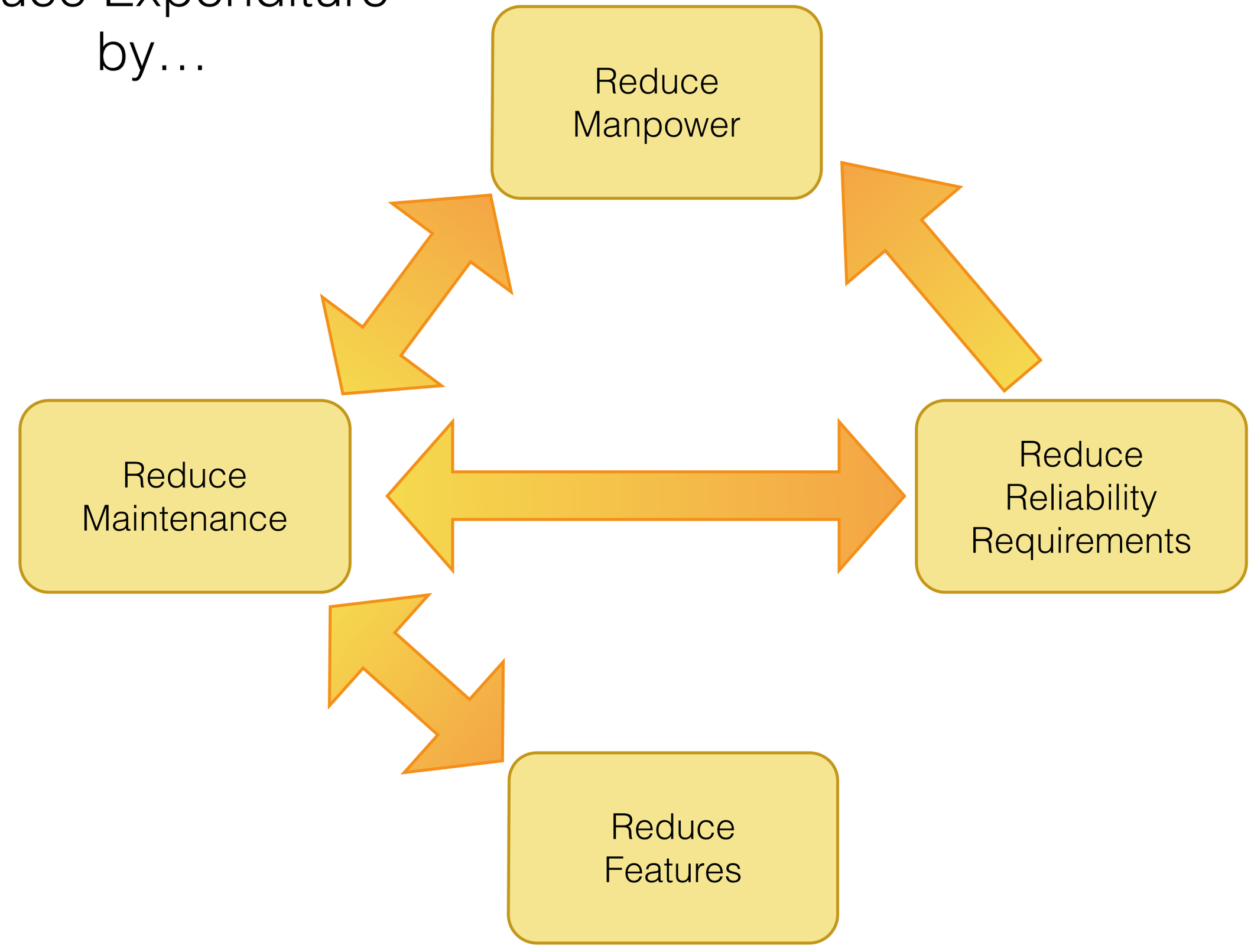
Various Wishlist Items

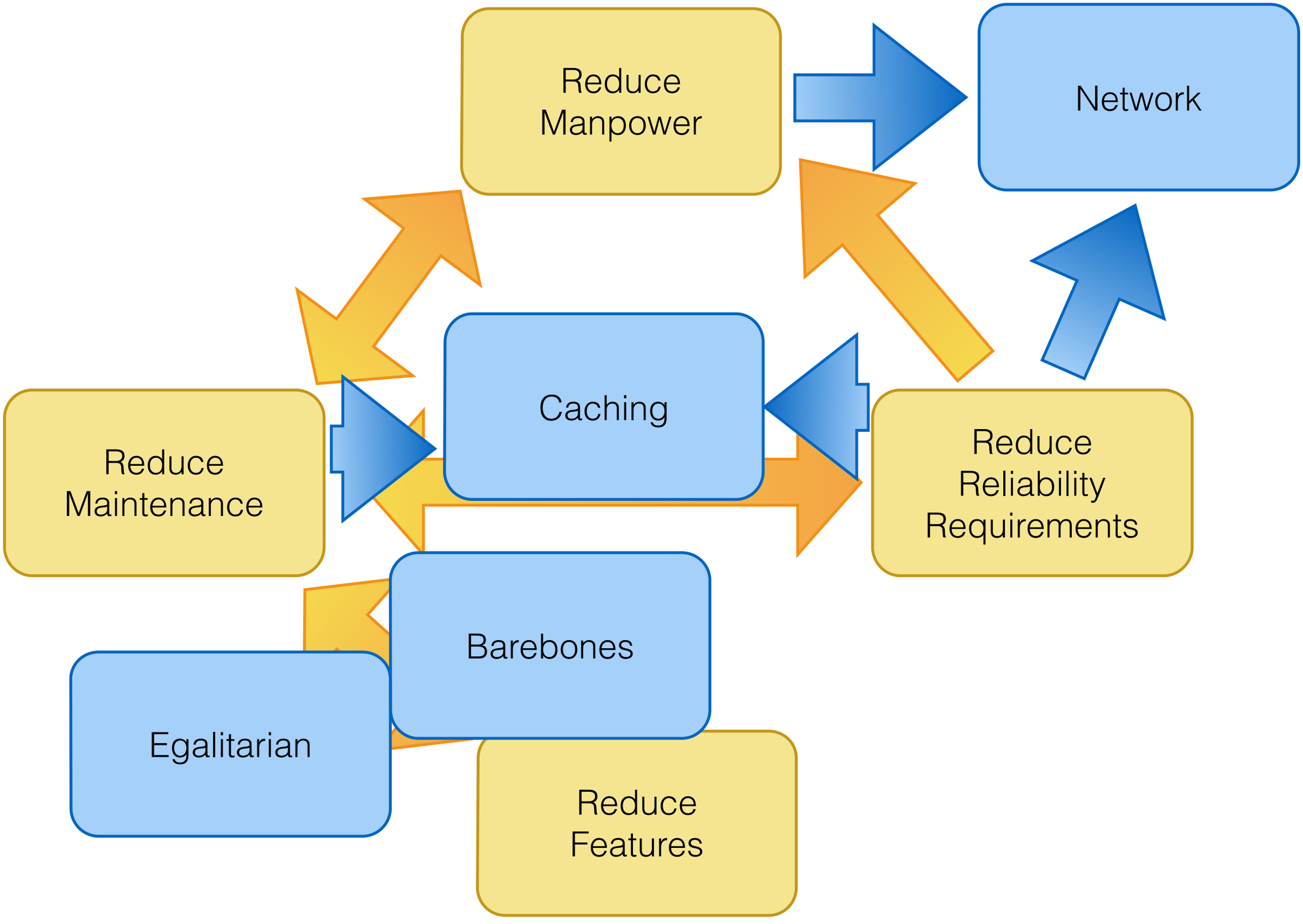
- Dual / HA head node
- Centos 7 [we have this]
- dpm-drain issues continue
  - "dpm-drain is slow"
  - "dmlite-shell drain is slow"

Future T2 Storage Evolution

- Complex topic
- multiple stakeholders
- political dimensions

Reduce Expenditure by…




Reduce
Manpower
Network
Caching
Reduce
Maintenance
Reduce
Reliability
Requirements
Barebones
Egalitarian
Reduce
Features

Caching /
Read only

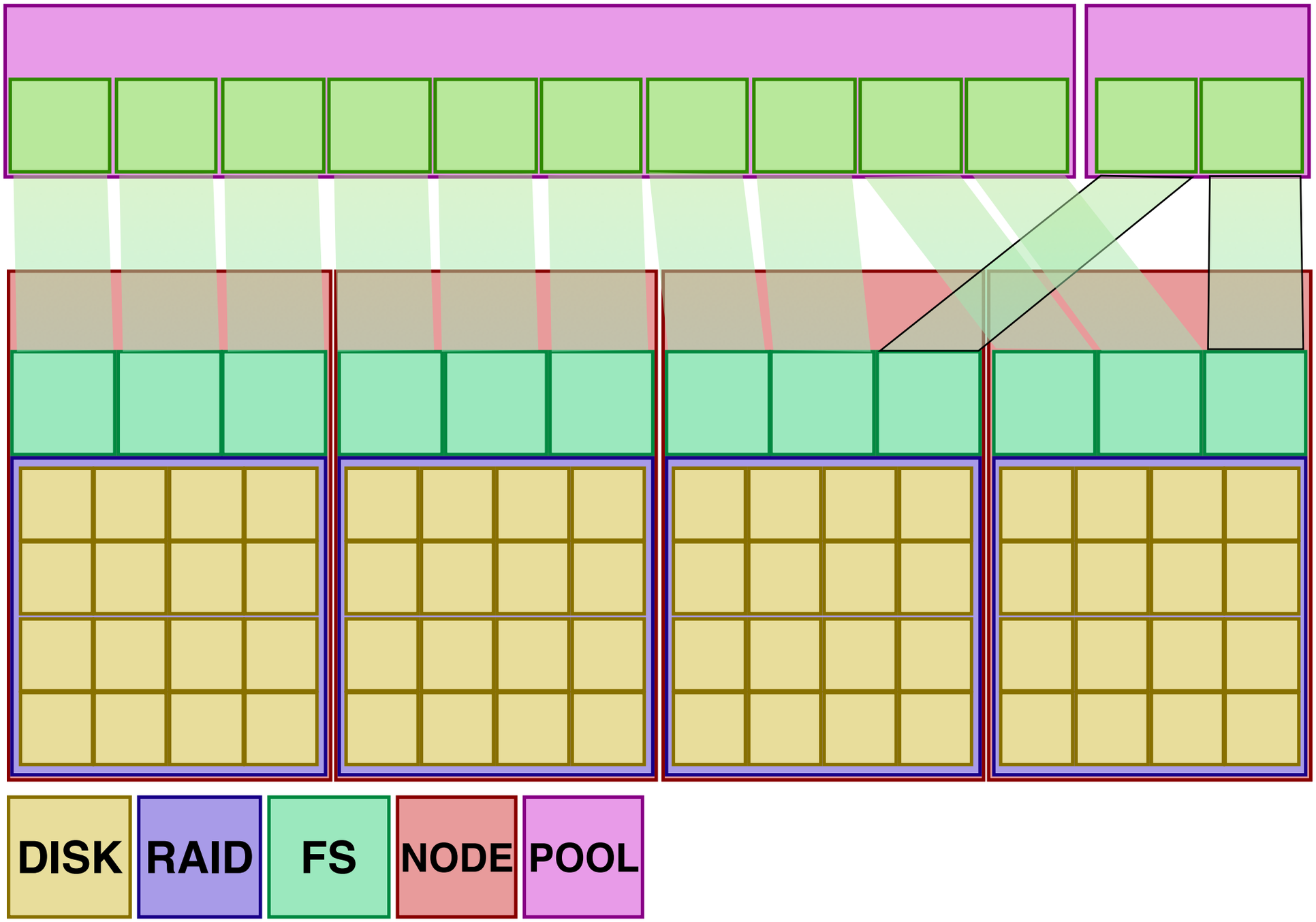
T2C Caching

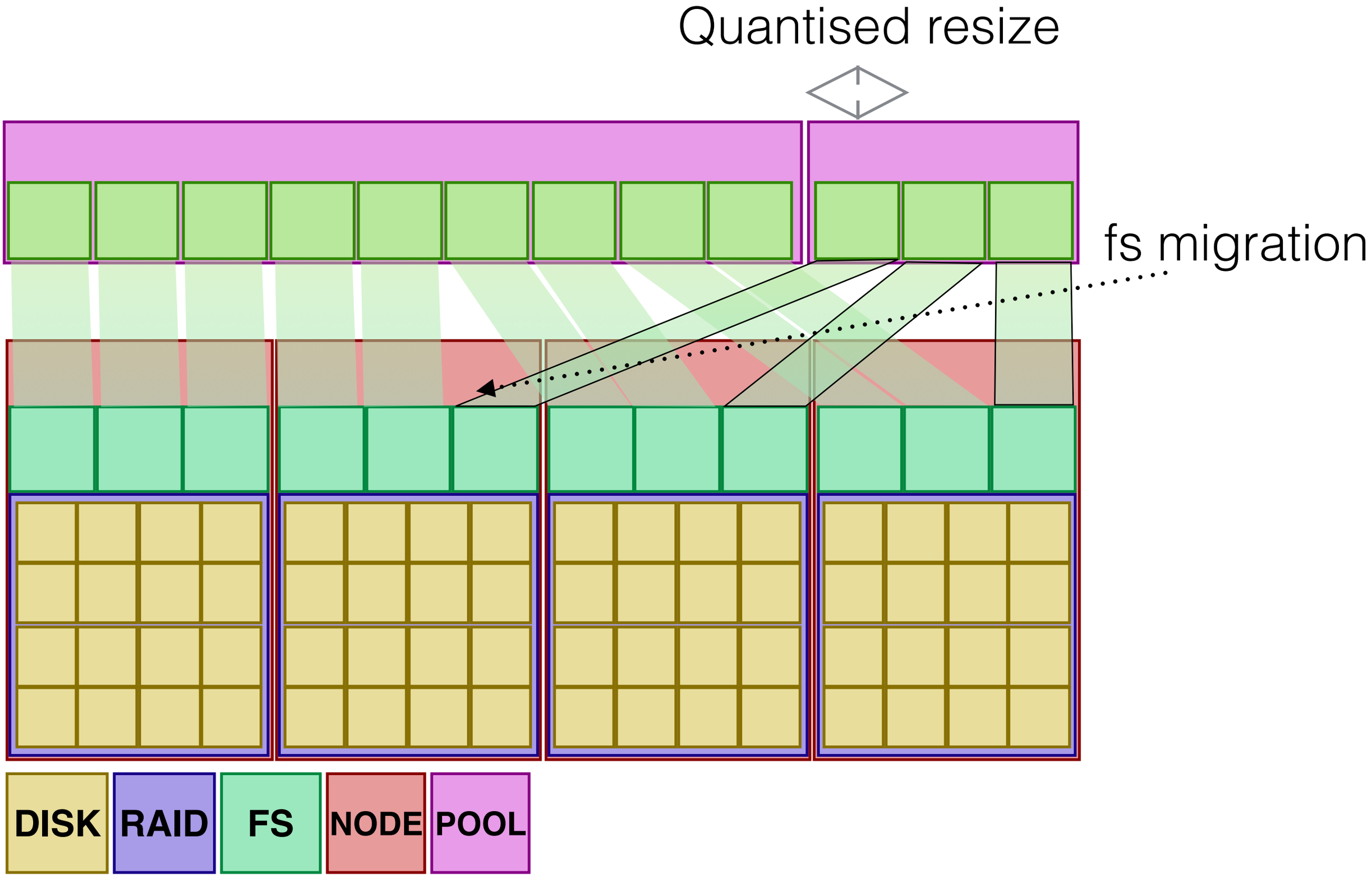
- WLCG Data Caching Group loosely "coordinating".
- Two solutions:
  - Xrootd - mature, not widely used?, compatible with xrootd federation
  - "DPM" - conceptual, in development, protocol agnostic, but DPM specific.
- Easy transition (for sites with DPM or Xrootd services)

DPM Pool simplification with ZFS.

- Current T2 storage is specialist - big RAID-6 arrays.
- Work by Marcus on ZFS, software 'RAID':
  - faster, more flexible, more extensible than HW RAID
  - small efficiency gains via compression (~4%)
- Small step (server local changes).
- (Also heads off RAID6 scaling limit @ 8TB disks)

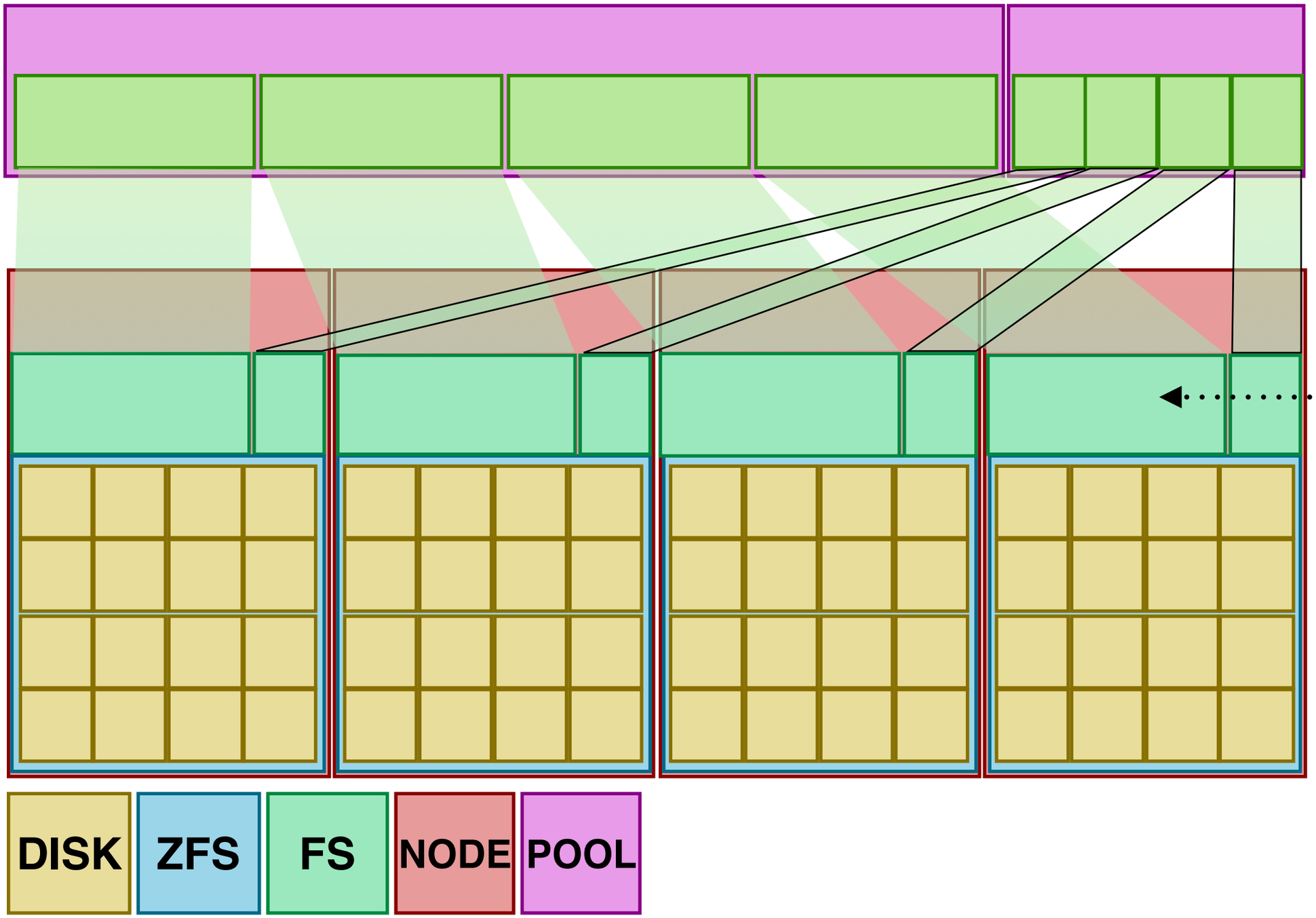
"Traditional" DPM with POSIX



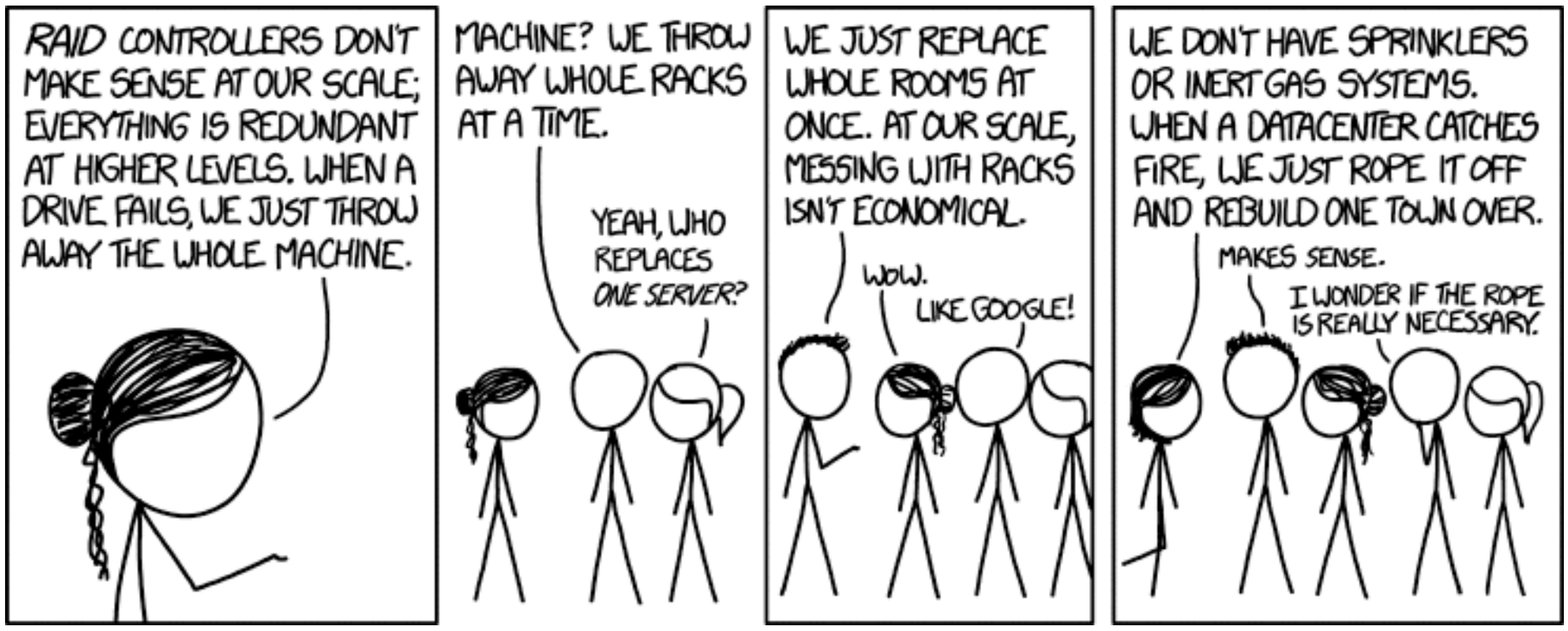

Quantised resize
fs migration
DISK
RAID
FS
NODE
POOL

"Novel" DPM with ZFS Datasets

DPM does not directly control pool space



Scaling up?



http://xkcd.com/1737/