Lustre at Durham

Lustre User Group meeting May 2024

Durham HPC Days

Alastair Basden

DiRAC / Durham University

HPC at Durham

- University service: Hamilton
- EPSRC service: BEDE
- DiRAC node: COSMA

- All use Lustre

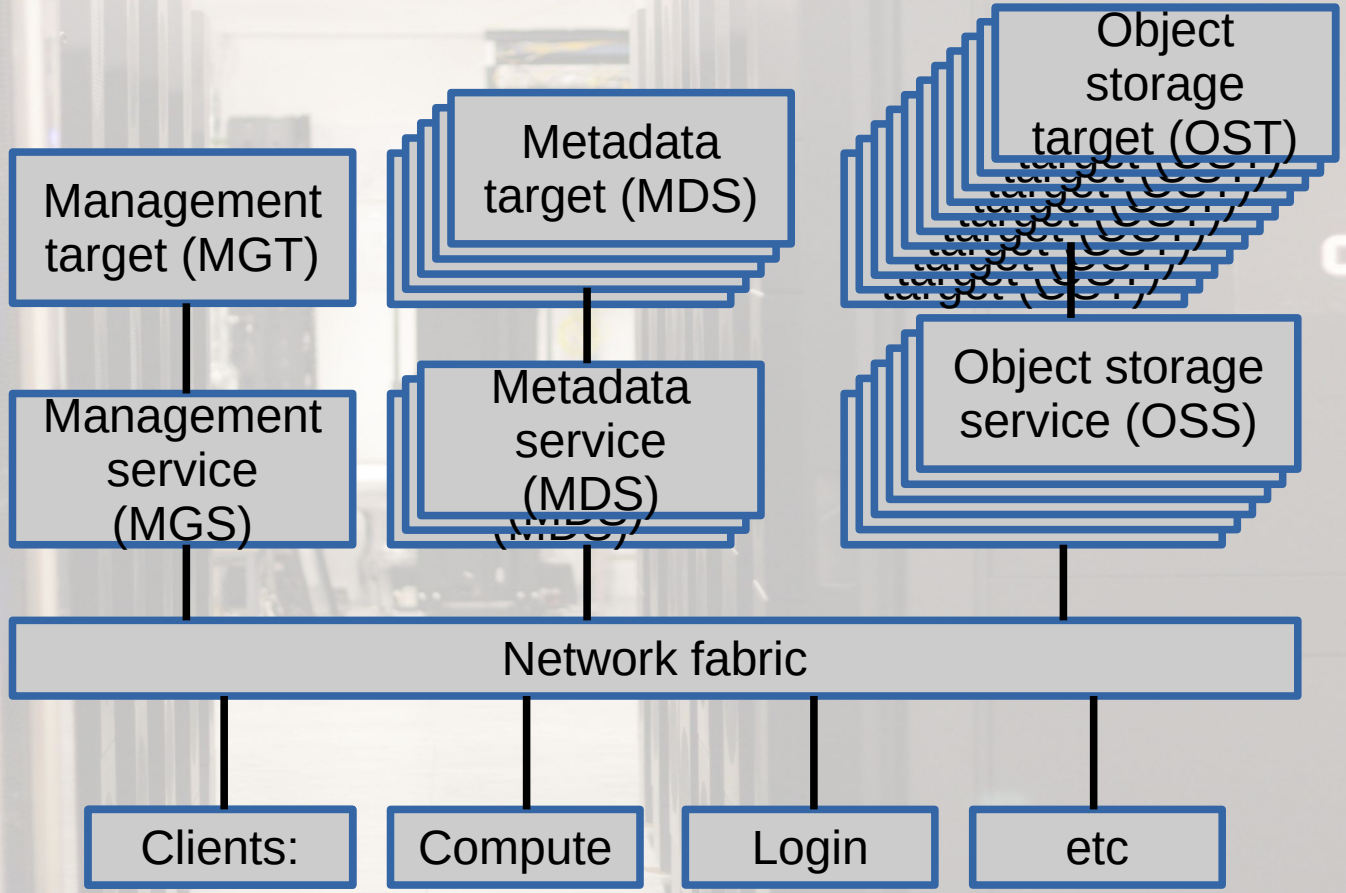
Alastair Basden
DiRAC / Durham University

Introduction to COSMA

- The DiRAC Memory Intensive node at Durham
  - STFC funded
  - 3 other DiRAC sites (Leicester, Cambridge, Edinburgh)
    - All using Lustre - DDN, HPE, Cray ClusterStore and Community

Alastair Basden
DiRAC / Durham University

Simplified Lustre overview



Alastair Basden
DiRAC / Durham University

Lustre on COSMA

- 8 Lustre systems:
  - Serving 3 HPC services
    - COSMA5 (Internal ICC/Physics system)
    - COSMA7 (DiRAC, 2018)
      - 4 Lustre file systems (1 almost-retired)
    - COSMA8 (DiRAC, 2021, 2023)
      - 2 Lustre file systems
  - And 1 other almost-retired system
    - For which the compute no longer exists
- All based on community Lustre 2.12.6

Alastair Basden
DiRAC / Durham University

COSMA5 (~2012)

- /cosma5 serves the ~300 node COSMA5 cluster
  - 1.5PB, Available read-only elsewhere
  - Installed in 2019 to replace the ageing original GPFS storage
    - Compacting 2.5 racks into 12U
- 2 servers (Dell R640)
  - 1 combined MGS, MDS, OSS (1 MDT)
  - 1 OSS

- 2 JBODs (Dell ME484)
  - 1 with 84x 12TB drives, 1 with 84x 16TB drives
  - 14 OSTs
- ZFS throughout
- No High Availability
  - If the MGS/MDS server fails, downtime required
  - If the OSS server fails, the OSTs can be migrated to the other server

Alastair Basden
DiRAC / Durham University

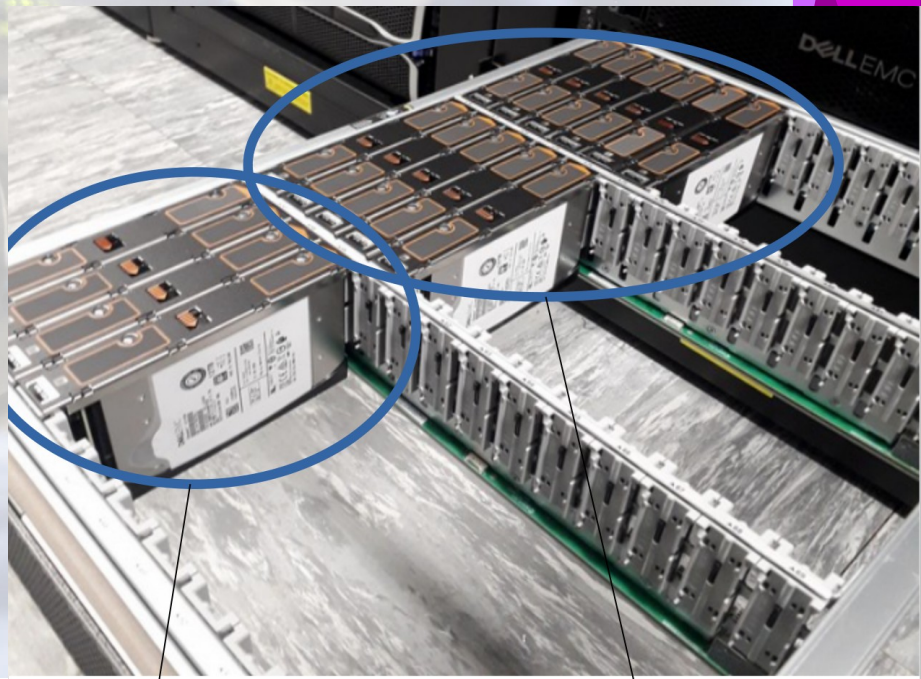
COSMA6 (~2012)

- /cosma6 served the now-retired COSMA6 cluster
  - ~500 nodes
  - 2.6PB
- 6 servers, Dell R640
  - 2 MDS, 4 OSS (paired)
- 4 ME484 JBODs (each 84 x 12TB drives)
  - 28 OSTs, ZFS
- 1 ME4024 raid controller for MDTs
  - 2 MDTs, ldiskfs

Alastair Basden
DiRAC / Durham University

ME484 JBOD aside

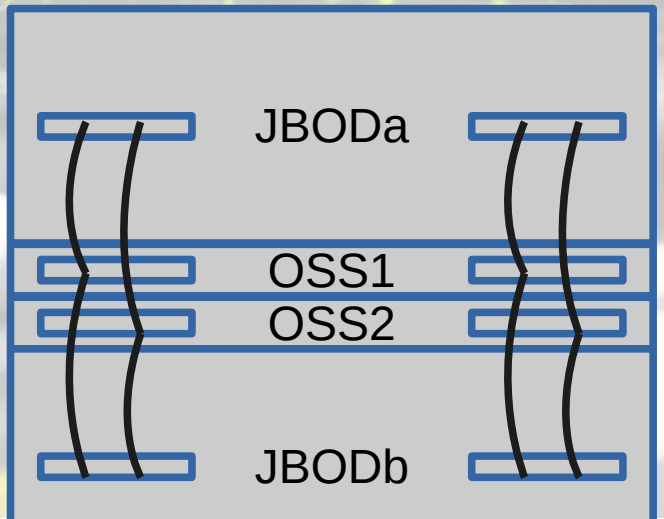
- SAS channel balance non-optimal
  - 4 SAS channels
    - 2 with 14 drives, 2 with 28 drives
  - This leads to a natural OST size of 12 drives



Alastair Basden
DiRAC / Durham University

OSS to JBOD cabling

- We cable the OSSs in High Availability pairs:
  - 2 JBODs are each connected to 2 OSSs
    - 4 SAS cables from each server (2 into to each JBOD)
    - Redundancy (multipath) for each SAS card and JBOD interface module



Alastair Basden
DiRAC / Durham University

COSMA7 (~2018)

- 4 storage systems, two entering retirement
  - 3 share the same MDS and 20 OSSs
    - /cosma7old (3PB)
      - 54 OSTs
    - /madfs (700TB)
      - 12 OSTs
    - ldiskfs throughout, 8TB drives
    - /snap7 (460TB)
      - 160 OSTs (3.2TB NVMe)
  - R640 servers, MD3460 arrays (controllers and JBODs)
  - /cosma7 is a new replacement for /cosma7old
- Serving the 452 node COSMA7 cluster
- No metadata High Availability (single server)

Alastair Basden
DiRAC / Durham University

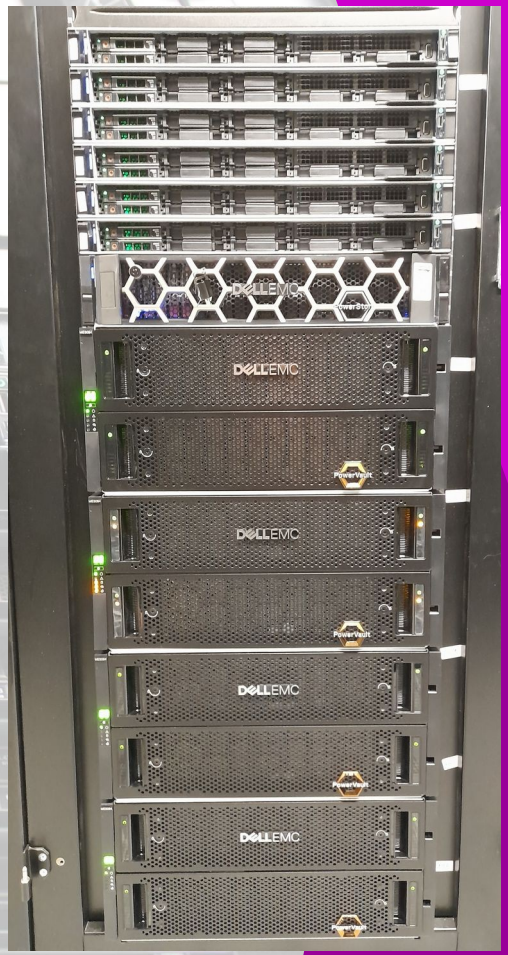
/snap7

- Fast checkpointing storage
  - Based on 160 3.2TB NVMe drives in 20 OSSs
    - 160 OSTs
  - Fastest storage in Europe at installation - 200GB/s
  - No redundancy
- Bespoke design for a particular use case
  - Checkpointing of cosmology simulations

Alastair Basden
DiRAC / Durham University

/cosma7

- New replacement storage for /cosma7
  - Installed in 2023
  - More cost-effective to purchase new storage than to extend the warranty on old storage for 2 years
  - 2 MDSs, 4 OSSs
    - Dell R6526 servers, Dell PowerStore and 4x ME5084 controllers
    - ldiskfs throughout
    - 3.4PB, 8 OSTs, 2 MDTs
  - Migration upon request



Alastair Basden
DiRAC / Durham University

COSMA8

- 2 file systems: /cosma8 and /snap8
  - /cosma8: 16PB, 4 MDSs, 18 OSSs
    - MDS using Dell R7525 backed by NVMe, ldiskfs
      - DRBD for data replication for High Availability
    - OSSs (R6525) paired with 18x ME484 JBODs with 16TB drives, ZFS
  - Has grown organically over time:
    - 3PB, 5PB, 12PB, 14PB, 16PB
      - Extensions are added live
    - Now at estimated 150GB/s
    - HDR200

Alastair Basden
DiRAC / Durham University

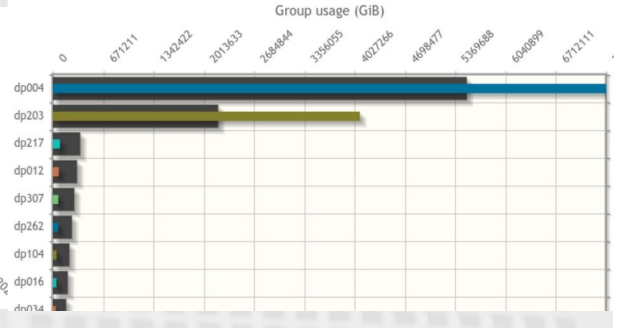
/snap8

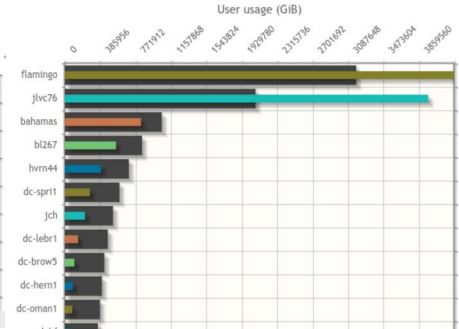
- /snap8: 1.2PB, 1 MDS, 24 OSSs
  - Dell R650 servers
  - 8x 6.4TB NVMe drives per OSS and MDS
  - 2x HDR100 per server
    - Dual IB card routing adds complexity
  - Apache Pass NVDIMMs to increase cache
  - Close to 400GB/s bandwidth
  - No redundancy on OSTs

Alastair Basden
DiRAC / Durham University

Monitoring

- Status monitoring using cronjobs
  - e.g. reporting of hardware failures
- Lustre exporter to grafana in development
  - To view user usage stats etc
- User usage dashboard





User management

- Nothing complicated:
  - New users get allocated a directory on the least-full MDT
  - Default quota allocation
    - This can be increased upon request
      - Different projects have different rules
  - Mount points are reasonably obvious for users
    - e.g. /cosma8
    - Cross-mounted between clusters where possible
      - Read-only

File system compression

- On ZFS OSTs:
  - lz4 compression
    - Achieves ~1.1-1.2x across the file system

Alastair Basden
DiRAC / Durham University

High availability (HA)

- Manual HA on some systems
  - /cosma6, /cosma7, /cosma8
  - Not on /cosma5, /cosma7old, /madfs, /snap7, /snap8
    - 2 for performance, rest due to design
- Pacemaker setup
  - Not used - causes more problems than it solves

Backup and archival

- No automatic backup of data
  - MDTs are backed up daily
- Archival to tape available on request

Problems

- Dual IB routing non-deterministic on startup
- Occasional ZFS MDT kernel panics
  - /cosma5 only

Alastair Basden
DiRAC / Durham University

Wish list

- ZFS rebuild times are large
  - >4 days
  - dRAID will probably solve this
  - We need to investigate Lustre 2.15
- Erasure coding would be nice
  - Particularly for the checkpointing (fast) file systems

Conclusions

- Lustre serves our use cases well
  - Cost effective and highly performant
- Lots of operational experience

Alastair Basden
DiRAC / Durham University