NFS access to GPFS at DLS, our experience

Frederik Ferner <frederik.ferner@diamond.ac.uk>

Diamond Light Source Ltd

December 5, 2019

diamond

Outline

- Diamond
- Filesystems and their use
- Latest Implementation
  - CES
  - CES NFS problems
    - Cosmetic
    - Real Issues
    - Dealbreaker
- Solution for DLS

Diamond

Diamond Light Source (DLS) is the UK's national synchrotron facility. It is located at the Harwell Science and Innovation Campus in Oxfordshire, UK.

- Synchrotron
- many beamlines/detectors
- users/scientists need quick feedback on data quality for experiment planing

High Performance Filesystems at DLS

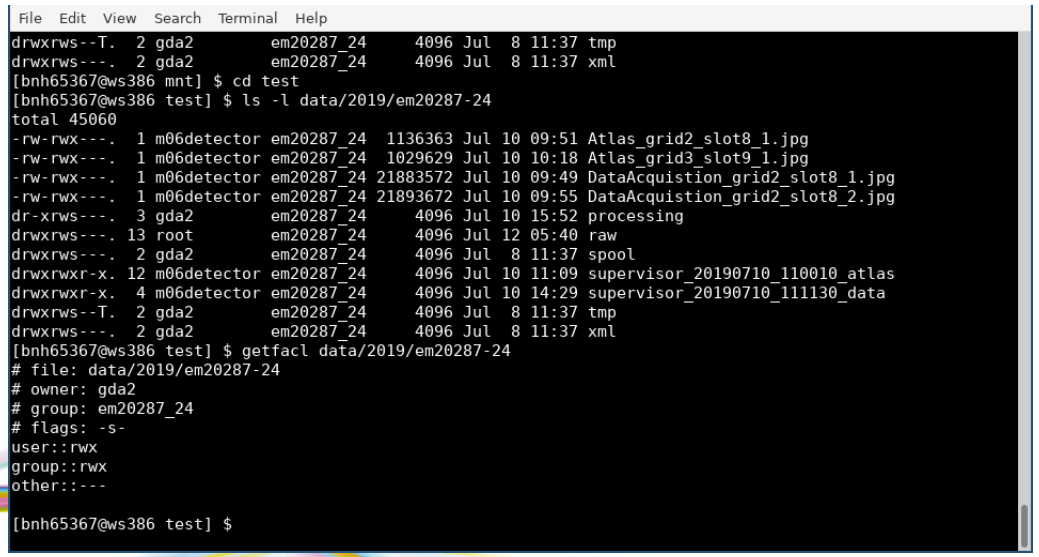
3 GPFS file systems

- Shared use with Data Acquisition, Processing and Visualisation in parallel
- Scientists want to see data on Desktops
- NFS (and CIFS) access required

GPFS03 — Initial CES setup

- GPFS 5.0.2.X
- 4 nodes
- CES
- SMB CES service proved not possible due to POSIX ACLs

ACLs (POSIX)



ACLs over NFS

- fully enforced (good)
- not displayed on NFS (CES)

Impatient Scientist Scenario

- Detector creates files. File count over NFS lagging behind by initially seconds, increases with number of files in directory, eventually writer completes 10000 files but NFS still listing only 7000-8000 files hours later.
- expected NFS behaviour???
- eventually hotfix....

I/O errors

Some time after restarting everything, user suddenly receive I/O errors over NFS.

- ganesha process hitting limit of open FDs (No failover trigger?)
- FD leak?

https://www.ibm.com/support/pages/ibm-spectrum-scale-nfs-operations-may-fail-io-error

random CES IP failovers

With the changes for the FD issue in place, NFS ganesha now so busy that it causes IP failovers? Clients hanging...

file corruption over NFS

Users reported that their data collection were broken, the data they see is not what they expect!

- files on native GPFS client OK
- files on NFS different => not OK
- only some files affected
- consistent across all NFS clients
- peristent for hours (i.e. identified a file at midnight, still same issue next morning)

Still unresolved to date.

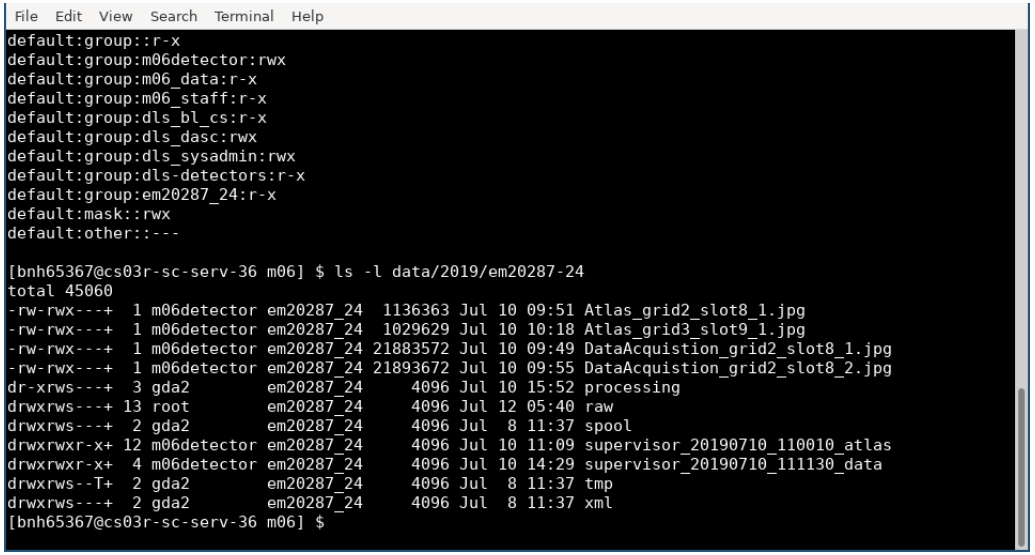
Documentation?

Best practice documents how to set CES up for this? => 404

Solution for DLS

Giving up on CES (for now??????), replacing it with kernel NFSd solution

Problem solved!



Questions

Questions?