NAF Status

- Status of SONAS successor
- PNFS problems on WNs
- AFS status - idea

Yves Kemp
10.8.2016 Hamburg

HELMHOLTZ ASSOCIATION

DESY

SONAS successor DUST ... update

- What we know and have
  - Hardware from IBM (and some DELL), software GPFS with native raid from IBM
  - ~1.8 PByte raw capacity (~1.3 PByte usable), already delivered
  - Setup needs some time: Currently fighting with dead-on-arrival hardware...fixed
  - Currently Will perform burn-in and initial load tests before pilot production
  - Currently working on quota management migration to AMFORA (the AFS quota mgmt tool) ... Almost done
- Prospected migration plan
  - Start with CMS ... Would also come in handy (hopefully) for the CMS-DAS school in September -> **Looks good, will be ready by ~15.8.**
  - Migrate the other groups N month later (N>1-2 && N<9)
  - IT does migration for a COMPLETE VO: We copy all data (rsync under the hood). Downtime only for final diff and re-mount. Max. 2 days downtime (e.g. in batch)
  - **Currently testing our migration tools. No estimate yet**

... One more thing on SONAS & DUST2

- Policy for data of expired accounts
- We will unlink data once an account is expired.
  - Unlink a fileset: No longer visible, but still existing on disks
- After 180 days: Removal of the fileset

- Better: Users should organize data inheritance ***BEFORE*** account expires

- ... Somewhat in line with what was decided in CUC
- Is the ICHEP rush over? We have ~30 TB free in Sonas

Local dCache space

- CMS:
  - Also user space on dCache is full (20 TB free of 1.3 PB usable)
  - After ICHEP rush: Can things be deleted?
- ATLAS:
  - Is fine

PNFS troubles on BIRD WNs

```
Active Internet connections (w/o servers)
Proto Recv-Q Send-Q Local Address
...
tcp        0      0 131.169.84.217:816      141.34.228.53:2049      ESTABLISHED
tcp        0      0 131.169.84.217:868      131.169.64.132:2049     ESTABLISHED
tcp        0      0 131.169.84.217:885      131.169.64.156:2049     ESTABLISHED
...
```

RPC-Port

SONAS

„NFS-Port“

Two dCache pool nodes

Source-IP:Source-Port

Dest-IP:Dest-Port

- A connection to every NFS server that the NFS client speaks to is opened. On a different source-RPC Port
  - One dCache pool node can host many pools – each pool is treated as one server
- The port range is 665-1023 = 358 possible values
  - ~340 dCache poolnodes, with many more pools
- What can happen – and *does* happen
  - Kernel runs out of free RPC ports on the client
- People notice this as „cannot access dCache/Sonas/NetApp on wnXYZ, but on wnABC or on wgsNN“
  - This is **NOT** a SONAS or dCache or NetApp **Server** problem!

DESY

PNFS troubles on BIRD WNs _ 2

- Hard to debug ... Intermittent error on WNs
  - The ports are freeed again after some timeout
- Very lucky to have an NFS server developer inhouse! (Tigran)
  - Got down to the root cause of the problem
- Long-Term-Actions: Tigran proposed a modification in the kernel
  - Unlikely that it will make it into RHEL 6 (=SL 6)
  - RHEL 7 might be possible
- Short-Term-Action: Monitoring, Mitigation
  - Monitoring of usage in place
  - Cool-Down script in place: Take a node offline if RPC-port-usage > 80%
  - **Not yet done (consequences unclear): Lower the threshold from 665 to e.g. 256**
  - **This has been done – no more reports of NFS troubles since!**

Status of AFS / OpenAFS

- Have you followed the CUC in the last three month?
- In a nutshell:
- OpenAFS project has trouble delivering working kernel modules for Ubuntu
  - There is some danger for RHEL derivatives, but currently we seem to be fine.
- Nevertheless, we must think about what comes after OpenAFS
- A questionaire was sent out to DESY groups. NAF users not explicitly asked.
  - ... But we can/should discuss this also here

One *idea* for $HOME batch in future:

- Each user@experiment
  - Will have local-cluster-store $HOME
  - Might have local-cluster-store $SCRATCH
  - AFS is there, and accessible. $HOME != /afs
  - Local-cluster-store only mounted within NAF+Batch. External access via sshfs or the like
- User mccartney works in experiment beatles and wings
- /nfs/somestore/beatles/home/mccartney
  - $HOME when entering through naf-beatles WGS and submitting jobs here
  - O(10GByte) Quota, with backup / snapshots
  - Provided by IT
- /nfs/somestore/wings/home/mccartney
  - $HOME when entering through naf-wings WGS and submitting jobs here
  - O(10GByte) Quota, with backup / snapshots
- /nfs/somestore/beatles/scratch/mccartney
  - O(1TByte) Quota, no backup/snapshots
  - Group has to pay for it