

## NAF status report

Andreas Haupt, Yves Kemp  
NUC, 14.4.2010

# Main problem: Lustre

- > Hamburg Lustre instances were affected
- > Started 27.3. evening (Saturday), and continued during Sunday, affecting mostly CMS.
- > On 29.3. noon (Monday), started also to hit ATLAS and ILC.
  - 29.3. 15:30: Emergency patch (announced via naf-announce)
  - Took longer than expected, new problems (ATLAS and CMS).
- > 30.3: Still not fixed. Contacted experts from SUN and DDN, closing batch system.
- > 31.4 16:00: Previous rescue attempts unsuccessful, recreate file system and restore metadata
- > 1.4. 14:00: Message: Everything works again!
- > First bug: Stability after upgrade to 1.8.2 (Lustre bug ID 22177 for experts), hitting MetaData server
- > Second bug: Misconfigured ext3/ext4-extends in patch to first bug, only hit newer FS like the HH ones.
- > **Thanks to NAF Lustre experts and also external ones (mainly DDN)!**



# Lustre, continued

- > No progress on ATLAS deletion tool.
  - “RobinHood” project used for similar purposes but also extremely slow: ~40h for a FS with similar #files, on dedicated machine with 4 threads... (NB: RobinHood doc warns “not to overload FS with scans”)
  - But investigating. Wolfgang suggested to contact the GPB: Any help by others in Germany (and elsewhere...) is helpful. Maybe someone else already solved that problem, and we are just not aware of it!
- > Lustre and SL4:
  - Lustre 1.6 (i.e. Lustre for SL4) EOL: June 2010
  - SL 4.9. comes after this date. Not clear, whether we will be able to build a Lustre client
- > Lustre in general:
  - Statement about Lustre future@Oracle postponed... think what you want about this.
- > Stability: Still issues with memory deadlocks on clients. Responsible for largest fraction of node failures.



## Lustre, continued 2

- > Experiments requirements (last NUC): First thoughts
- > ILC: Probably rather easy: Group space belongs to unix group ilc-adm
- > CMS: Probably similar: Group space belongs to unix group cms-adm. VO-Box like system and its implications still under discussion. Eventually, a second FS is needed, but not clear.
- > ATLAS: Hard to fulfill (see also deletion tool)
  - Rather certainly only possible with two separate FS, and this is something we do not want at the moment.
- > Secondary groups will come, so this is progressing.



# AFS scratch space

- > AFS scratch is there: \$HOME/scratch
  - for most users, some CMS users already had volumes with similar functionality, in accordance with Hartmut, they do not get this space.
- > 2 GB/user assigned as a start
- > Registry platform adapter automatically assigns AFS scratch to newly created accounts
- > 332 existing spaces: 20 ATLAS and 1 CMS non empty (as of 12.4.)
- > But extremely overcommitted space: Need to buy some new HW for AFS server.



## Other items:

- > SGE priorities:
  - Basically clear, already explained in last NUC, needs to be implemented
- > SL 5.5beta will come soon (next week?), test system will be provided. Which communication channel to be used for details?
- > Feedback to Twitter? We only see one follower. Only few posts, biggest issue (Lustre) communicated via naf-announce.

