

**NAF Status**

**Action Items**

**Open Issues**

Yves Kemp  
NUC TelCo  
9.12.2009

# Status and items during last month

- > Old Lustre instance in Hamburg is finally decommissioned
  - Space being converted right now into dCache pool nodes
- > New Hamburg instance showed instabilities twice
  - ~22.11: Defective RAM exchanged. Not clear whether this was the (only) cause, still under investigation
  - ~1.12: SL4&Zeuthen clients (the ones with TCP/IP connection) troubled. One server NIC hung. Also under investigation, have some options (MSI, TSO).
- > 25.11. downtime
  - Longer than expected: Lustre nodes take storage from DDN, and DDN maintenance took longer than expected (because intervention started with delay).
  - Work was successful.
- > 23.11-24.11: Cooling maintenance in Zeuthen: Large parts of Grid and NAF shutdown



## Action item: AFS & CMT

- > CMT: Experimental patch from CERN deployed on one machine, undergoing testing by Wolfgang Ehrenfeld
  - Experimental means experimental.
  - If we provide it, only on SL5, and limited to (a subset of) Atlas-WGS
  - Potential stability issues on the clients, clients have to be rebuild for new kernels
  
- > AFS scratch creation:
  - Best implementation not yet clear, under discussion.
  - Good side effect: Consistent naming scheme (e.g. \$HOME/scratch)
  - Problems: Automatic creation (pre-)binds resources, existing \$HOME/scratch (or whatever) directories
  - Maybe start with small amount, admins can increase?
  - Maximal size in AFS also under discussion.

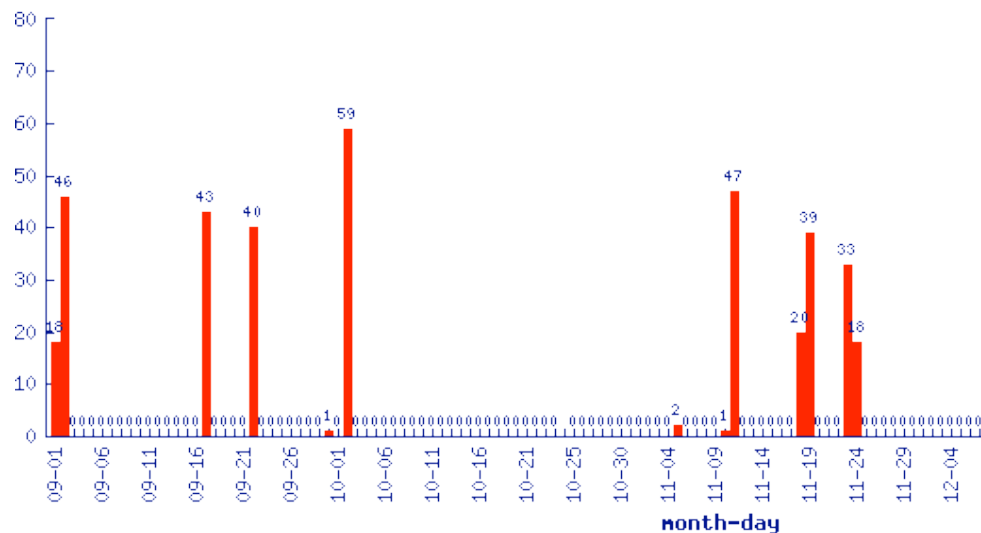
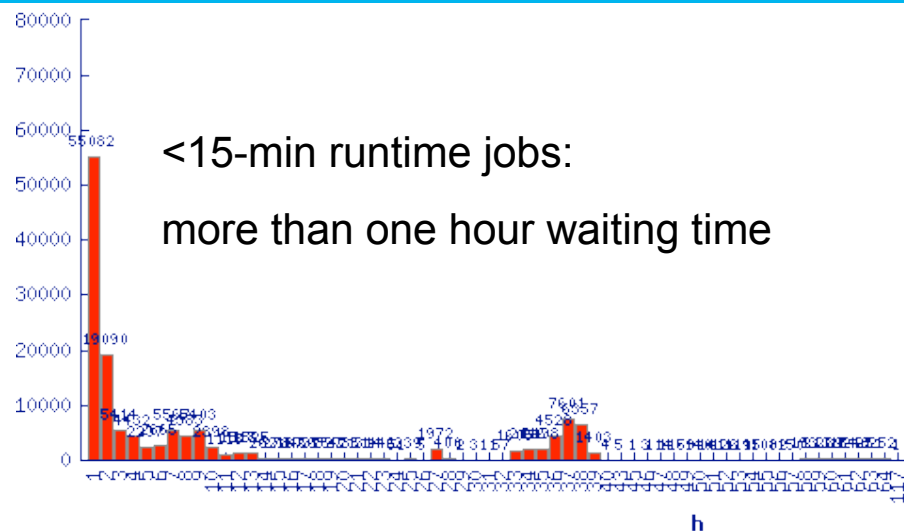


## Action item: Lustre, Quotas etc.

- > Background: Any Lustre quota needs an accounting “DB” initialized.
  - Should/Has to be done with Lustre unmounted
  - Has been done on Zeuthen Lustre instance during 25.11. Downtime
  - Has not (yet) been done on new HH Lustre
- > This alone can be used as “fast du” and is working.
  - Easy monitoring tool. If it works in HH, will be brought to users attention.
- > Now the next step: Quotas e.t.c.:
  - Use cases difficult. We need some internal discussion on how to best implement things.
- > Downtime beginning January? (Zeuthen needs 1d-downtime anyway in January)
- > Cleanup-Tool:
  - Like tmpwatch, based on atime, run e.g. once per week.
  - Caveat: Stability and performance not clear. Can only test “in vivo”
  - “Admins should not delete users data”. Experiments MUST make clear that we delete on their behalf!
  - Would knowledge of “big users” be enough?



# Action items: Monitoring



RAM: Used/Requested for one user

Caveat: Mean over whole day, if many jobs die, might be a bad measure.



# Action Item: Batch

## > Memory limit:

- For 16 GByte nodes:  $(16 \text{ GB Total} - 1 \text{ GB System}) / (10 \text{ concurrent jobs}) = 1.5 \text{ GB/Job}$
- 1.5 GByte job default will be implemented soon.
- 2.0 GByte: need at least 21 GB/system
- Will be done only when such systems majority in the NAF

## > Multicore docu: Has been reworked, some example switches presented.

- Generally: Docu has been reworked, thanks!

## > Multicore monitoring: Difficult (SGE buggy...), not yet there.



# Action Items: Others

## > Group profiles:

- Use case: “Need something to point to my AFS-Scratch” could be fixed consistent naming scheme

## > SL5 migration:

- Status of the experiments?
- Worker nodes?

## > Network between NAF and Heidelberg (and other institutes)

- Internal buffer size in ssh/scp small in older versions (including SL5.4 ones)
- On current SL4 machines: ~2-3 MB/s NAF->UniKarlsruhe
- On Fedora12: ~10 MB/s NAF->UniKarlsruhe (still not much, probably maximum possible)
- dCache should be used for transfer of massive data amounts



# Resource institute map

- > Reviewers like plots with the geographical distribution of people in the NAF
  - Was asked, producing the plot difficult (going manually through list and deciding about affiliation based on DN, guesses....
  - Do not need this very often though, must not be 100% accurate.
- > There is an affiliation/institute field in the registry and in the webmask.
- > Proposal:
  - Webmask will have a drop-down list with all institutes, according to Cert-DN, the best match is highlighted
  - Experiment admins should have a look whether this seems reasonable, eventually change.
  - Existing users: Fill using Cert-DN, some (<10%) remaining should be filled in manually

