Evolution of interactive Analysis Facilities: From NAF to NAF 2

Andreas Haupt, Yves Kemp,
Friederike Nowak
DESY
CHEP 2013, October 14

Outline

National Analysis Facility

- General Ideas
- Setup
- Lessons Learned

National Analysis Facility 2

- Requirements
- Implementation
- Status
- Future Plans

NAF: General Ideas

Founding

- Give users of German institutes additional resources for analyses
- “Helmholtz Alliance: Physics at the Terascale” → Collaboration of ~20 institutes
- Initially ATLAS, CMS, LHCb, and ILC groups

Ideas

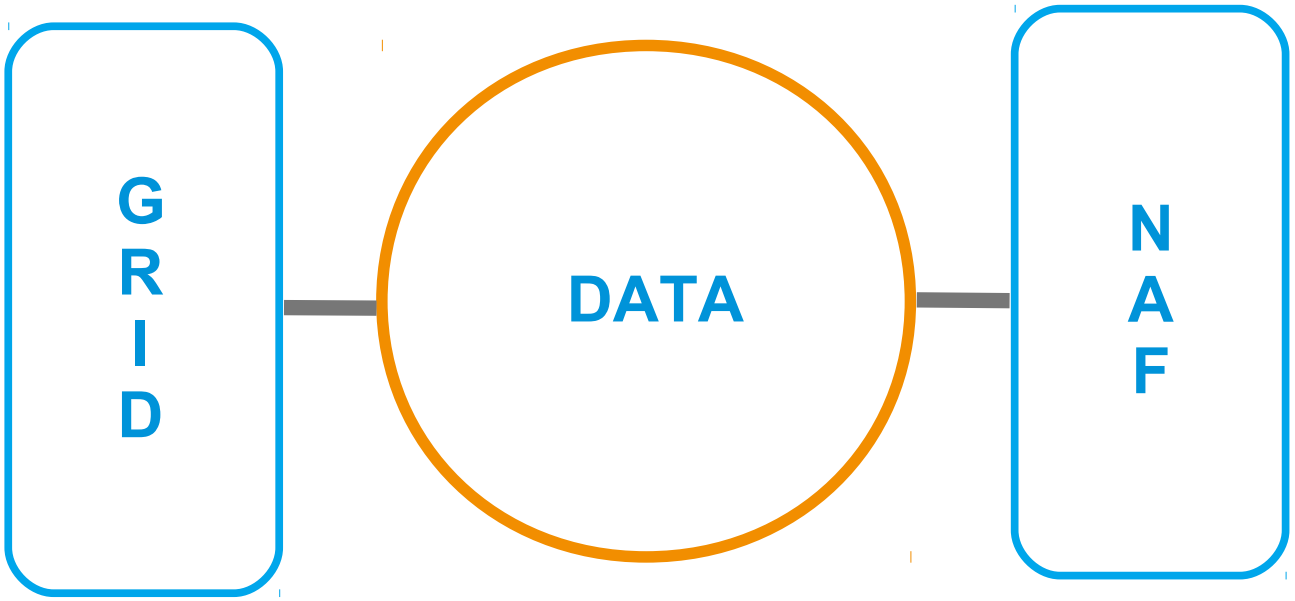
- Computing size : ~1 average Tier 2
- Focus of the Grid lies in optimized resource usage → well filled queues
- Analysts need their results fast, i.e. “interactively”
- Turn-around times <~ 1h
- Coupling (not integration!) to existing Tier 2 storage

Poster by A. Gellrich
“Job Scheduling in Grid Farms”

DESY

A Data Centric View

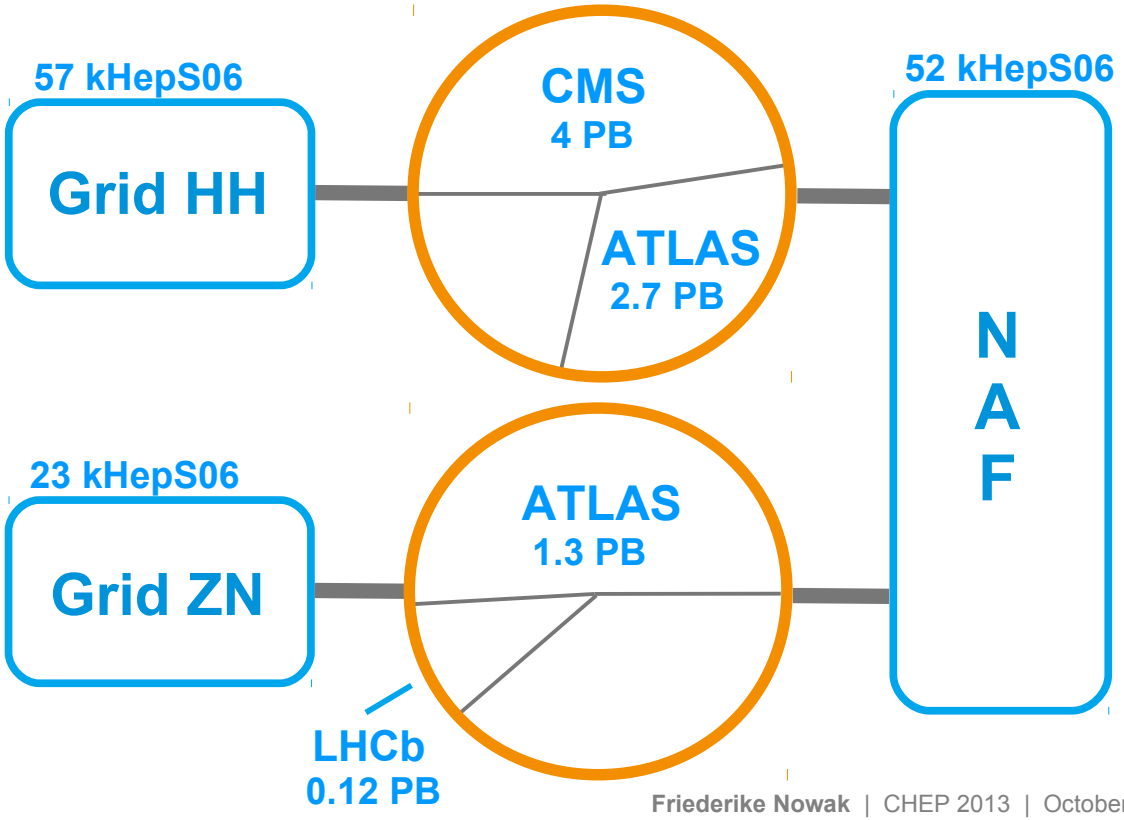
- Jobs go where the data is
- Grid well constructed for long running jobs
- NAF complementary: designed for short (analysis) jobs
- → NAF has to be placed where the data is!



NAF @ DESY

- Two DESY sites: Zeuthen and Hamburg
- Two multi-VO Tier 2
- > 8 PB (LHC) Storage installed

- CMS data completely in HH
- ATLAS data both in HH and ZN
- For ATLAS, NAF installation on both sites (divided by physics groups)

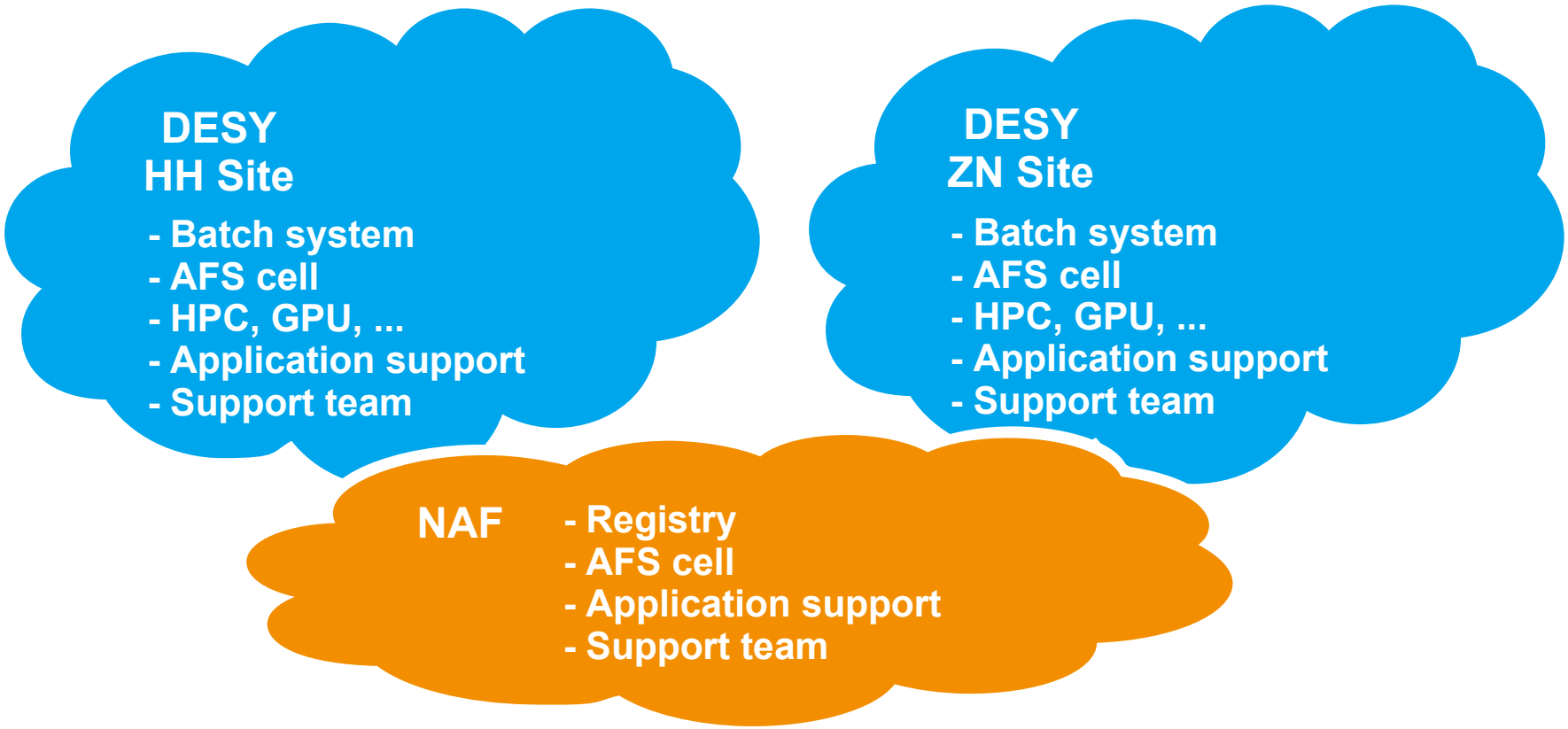


Thoughts About the Workflow

- Use NAF as one step within workflow
- Grid → NAF → local cluster → laptop
- Workgroup servers: compilers, debuggers, …
- No browser, mail client, ...

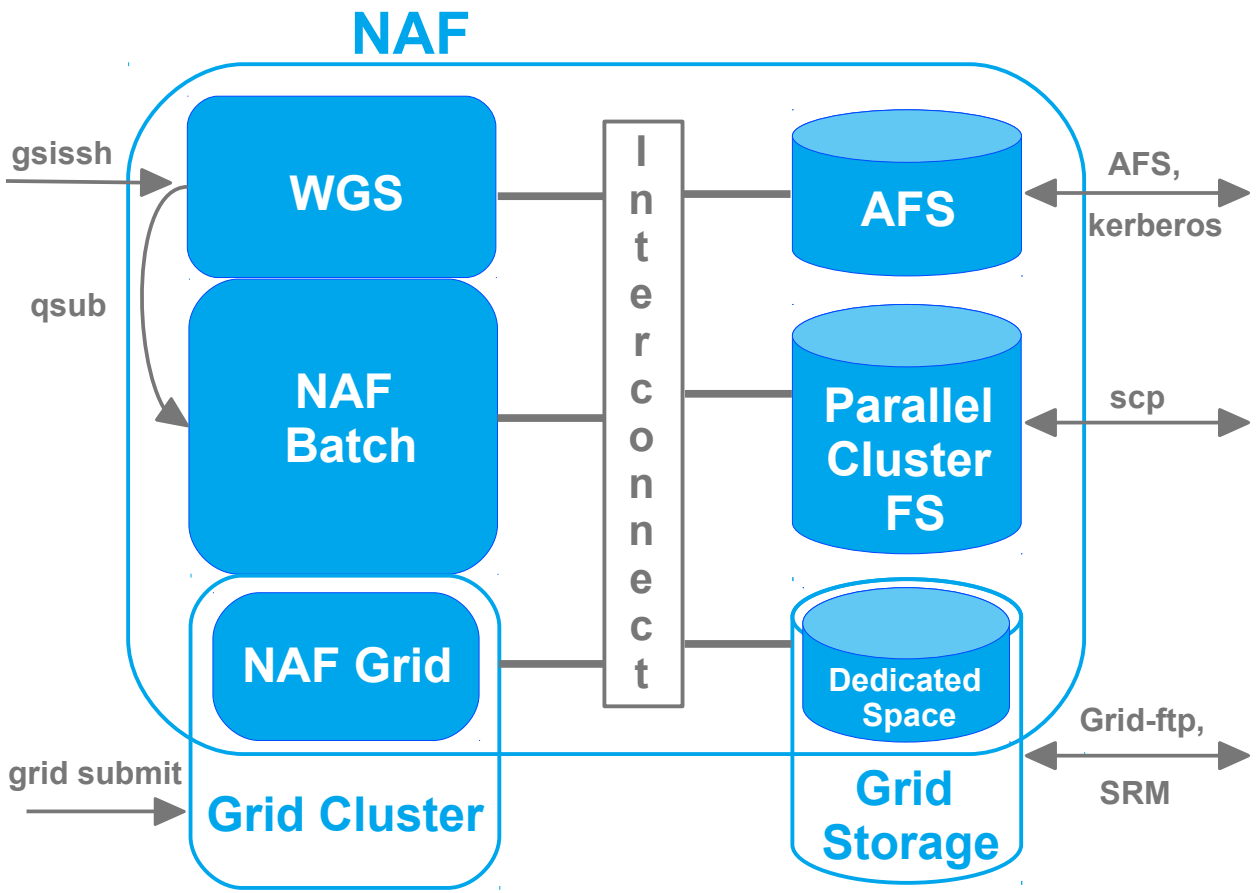
- Exchange tool : AFS
- Cross-mount in HH and ZN for ATLAS analyses
- Place NAF within own infrastructure (no site legacies)

NAF (Provider View)



- Own infrastructure for NAF (island solution)
- Own registry, application support, support team, config management, ...

NAF (User View)



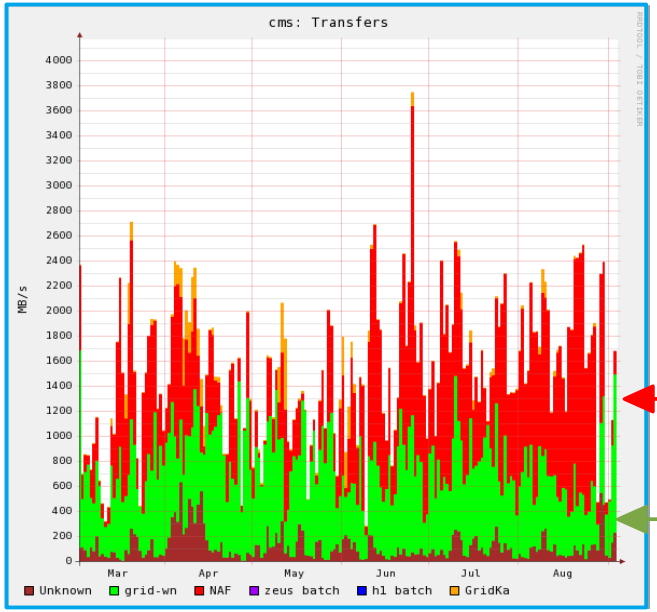
- Login via certificate
- SGE batch farm
- Own afs cell
- Parallel Cluster FS as scratch
- Dedicated space within grid storage
- Extra grid resouces with high priority for german users
- Placement in HH and ZN
- Well defined network

DESY

NAF Usage

- In usage since 2007
- ~500 accounts overall
- ~60 active users currently
- Main users: ATLAS, CMS
- 1/3 DESY Usage, rest by external Institutes
- Substancial amount of storage access through NAF

CMS dCache access since May

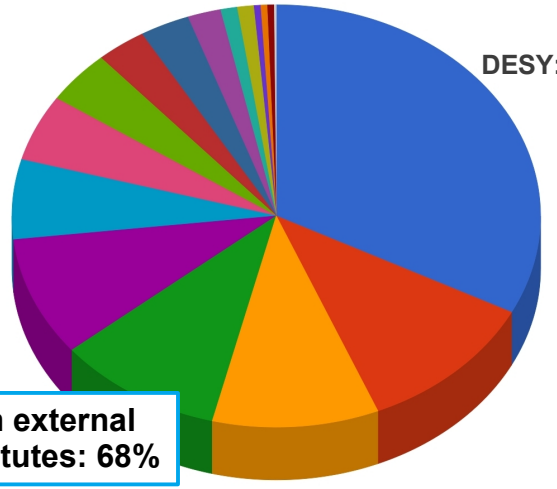


NAF

GRID

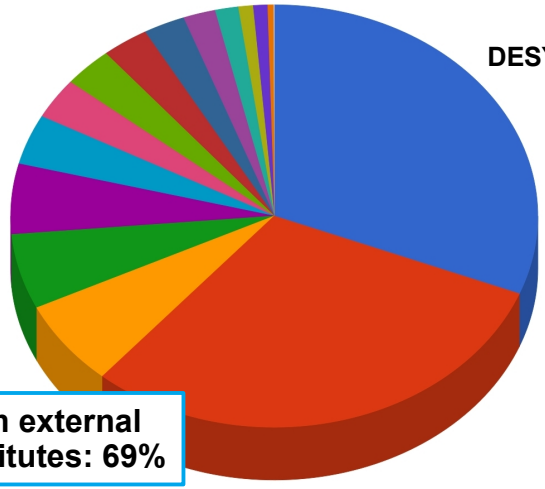
Lustre Usage by Institutes

1.12.2009-1.04.2013



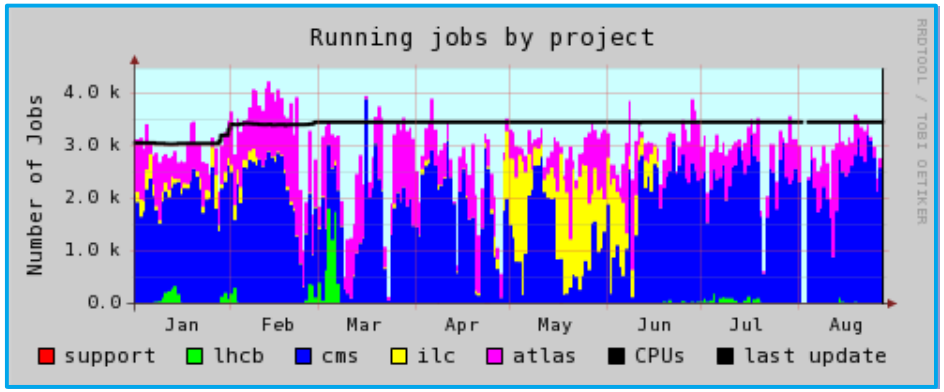
CPU Usage by Institutes

1.12.2009-1.04.2013

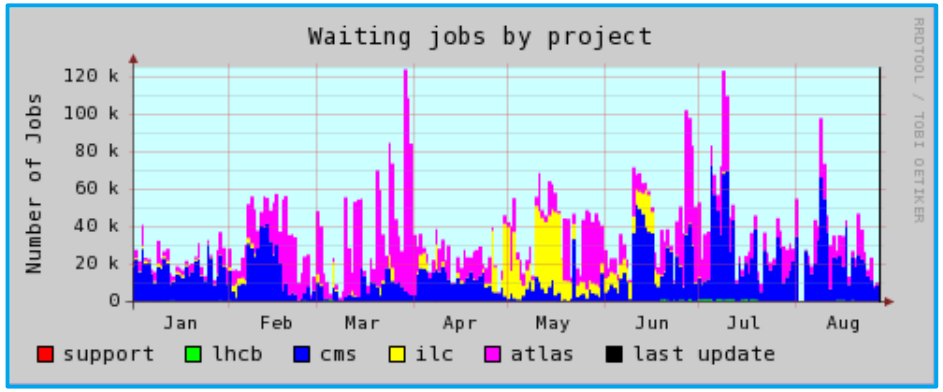


NAF Usage II

Running Jobs



Waiting Jobs



- > 3k job slots well used
- Up to 120 k waiting jobs at peak times!

- Data analysis with experiment software
- Data analysis with PROOF
- Code development and grid submission
- Private MC simulation
- MC generator tuning

Governance and Support

NAF User Committee (NUC)

- Represents physicists and coordinates resource usage
- Each experiment and DESY IT sends two members
- Ensures close collaboration between experiments and DESY IT

Support

- Delegation Model: DESY IT provides tools, experiments administrate resources (quota, experiment sw, ...)
- Fabric issues are handled by IT admins
- VO specific support provided by groups

Lessons Learned

- Own infrastructure too manpower intensive → need to consolidate
- Grid → NAF → local cluster → laptop: most user stay on NAF

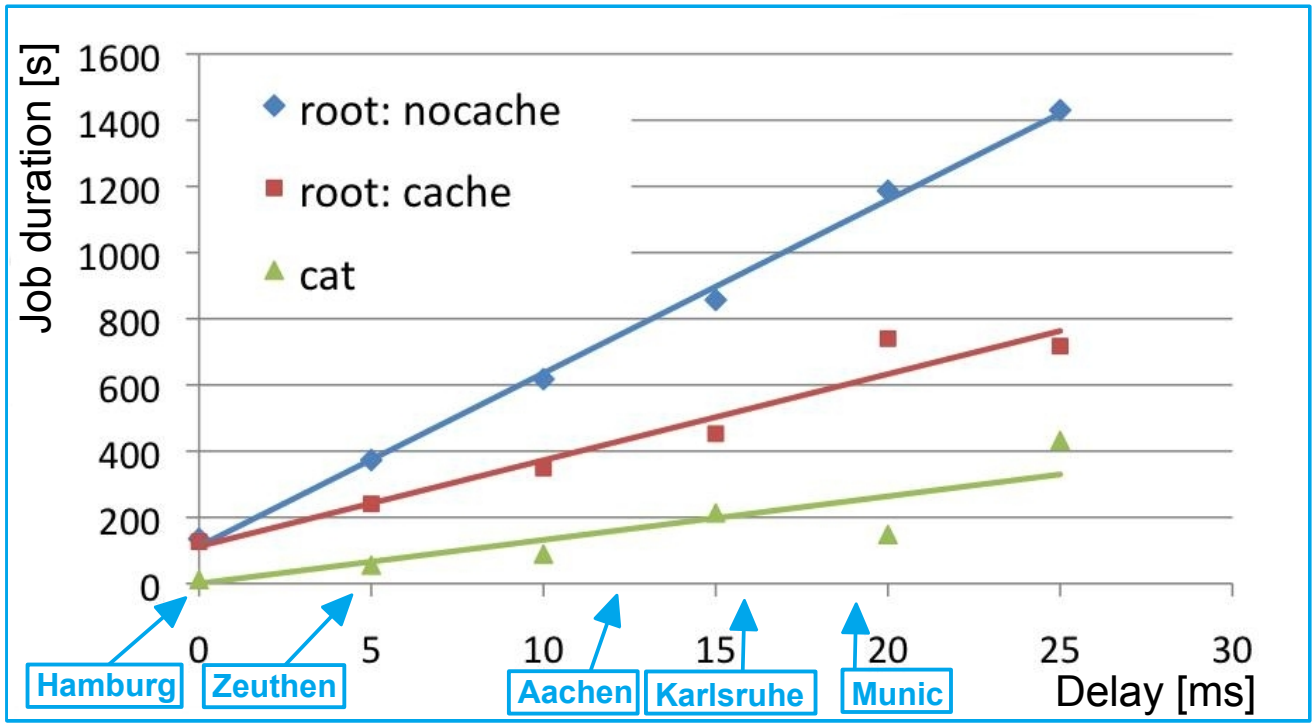
  → graphical tools/interactive access needed
- Since 2007: mobile devices

  → graphical work also in high latency networks
- Low general acceptance of grid certificates/X.509 (Non-HEP products)

Lessons Learned II

- Latency problems in analyses using both sites

Latency effects for ROOT file access via WAN with dCache NFS 4.1, lab test



- Also bandwith problem for other groups if such analysis is running

Requirements for NAF 2

Infrastructure

- Embedding in DESY Infrastructure increases efficiency
- Access to other DESY resources (HPC, GPU, …) possible if desired

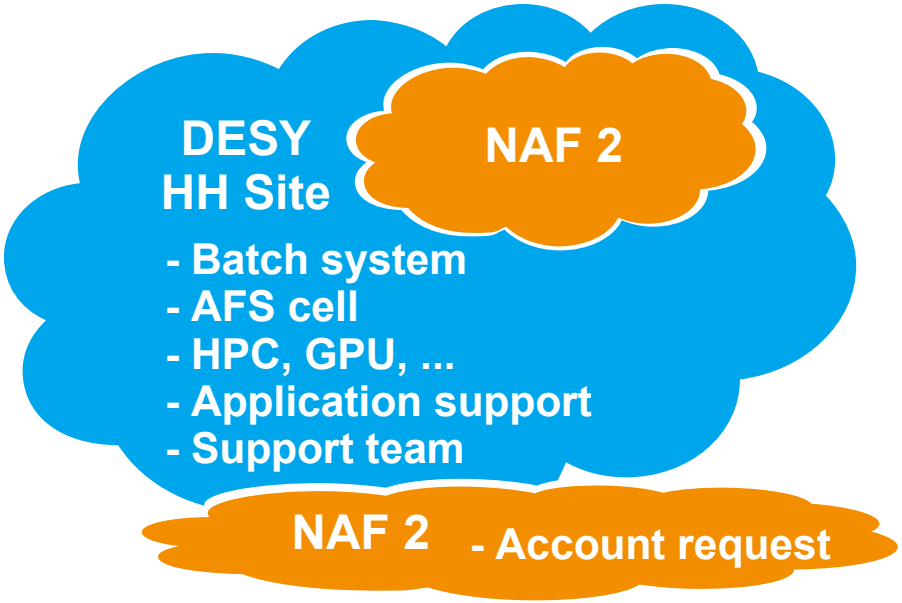
User Handling

- User registration should work with certificate for convenience
- Access to NAF 2 should work with std. user/password pair

Other

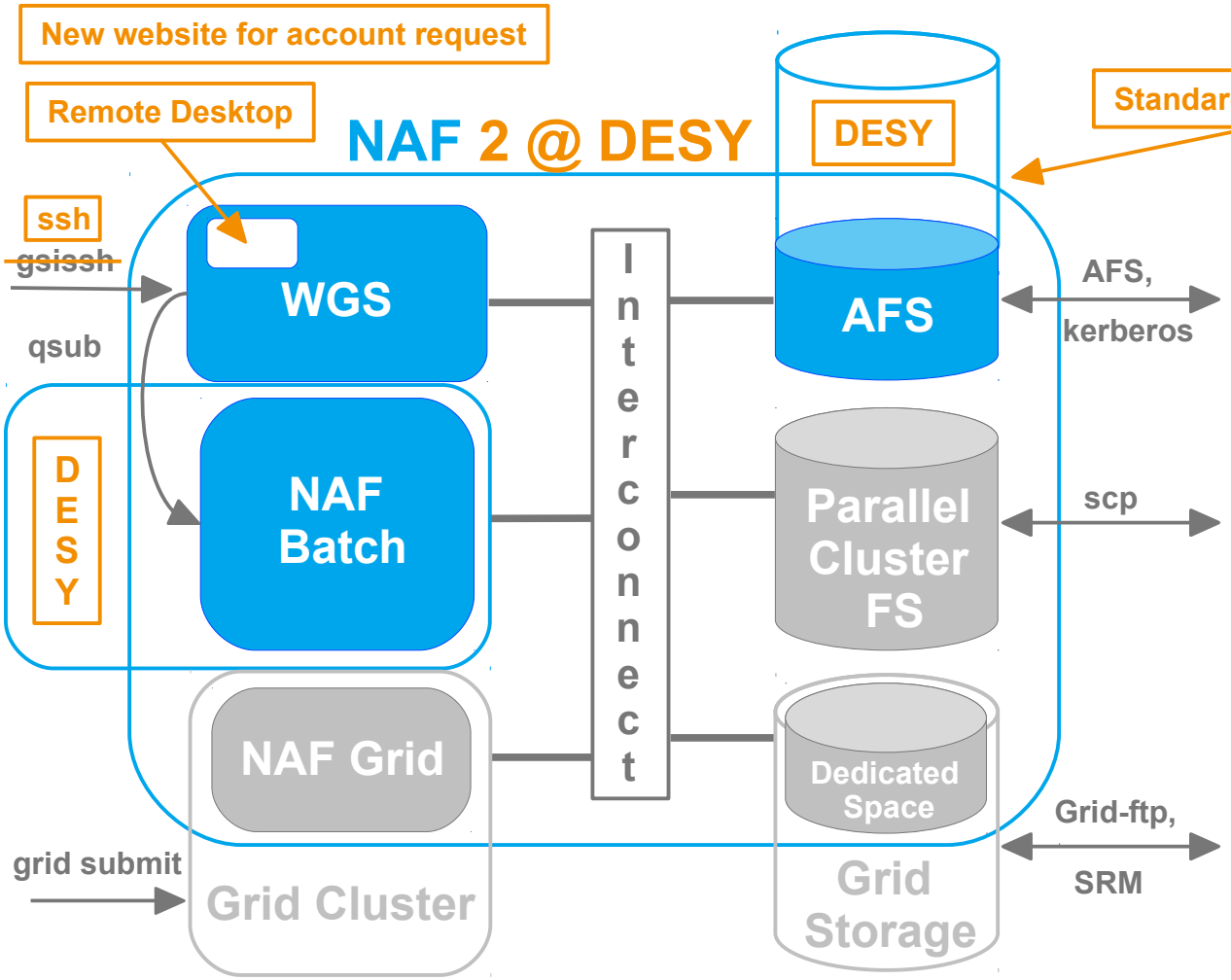
- People working remotely need to work graphically → provide tool to hide latency
- For each VO, provide resources on one site

NAF 2 Changes (Provider View)



- Integration of NAF into DESY infrastructure
- Additional Resources (HPC, GPU) easily accessible if needed
- Only account request left as disjunct entry point

NAF 2 Changes (User View)



- Infrastructure within DESY
- Login via ssh
- Remote desktop
- DESY batch farm
- DESY afs cell
- Concentration on one site
- New website for account request

Status

WGS

- ATLAS, CMS, ILC, Belle, Hera Fitter, ZEUS
- SL5 or SL6 (depending on VO)
- IT Managed

- ILC, Belle, HeraFitter, ZEUS fully on NAF 2
- ATLAS, CMS in migration process

Talk by C. Jung
"Otimization of data life cycles"

Batch

- Mixture of SL5 (~1600 cores) and SL6 (~1000 cores)
- NAF 1 nodes will eventually join
- Share will be ~4000 cores

Other

- Website for user registration in place
- 600 TB scratch space installed

DESY

Future Plans

- Shut down NAF 1 in 2013 / early 2014
- Grid-ftp access to scratch space if desired
- Add gsissh access if desired
- Provide extensive monitoring of resources
- Prepare for new communities
- And also for new Requirements (HPC, GPU, ...)

Summary

- To support national analysts, in 2007, the National Analysis Facility has been established
- It has proven to be a very important resource complementary to the grid
- After more than 5 years, concepts and setup were reviewed, to launch the successor NAF 2
- Important changes were
  - Embedding in DESY infrastructure
  - Inclusion of new technologies (NX, ...)
  - Inclusion of new experiments
- Currently in the phase of transition from NAF 1 to NAF 2
- NAF 2 already successfully in use by several experiments