

The NAF: National Analysis Facility at DESY.

The NAF concept and its place

In the German HEP field

In the Global Grid

The building blocks, with an emphasis on

Usage of VOMS for Grid-type resources

Interactive and batch cluster and integration with PROOF

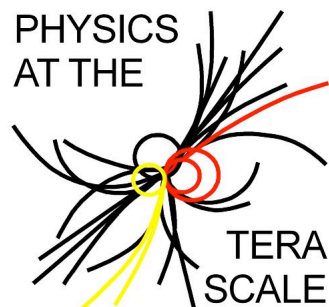
Usage of Grid-Proxies to access workgroup servers, AFS and dCache data

The usage and operation of Lustre for fast data access.

Running experience

[Andreas Haupt](#), [Yves Kemp](#) (DESY)

Prague, CHEP 2009, 24.3.2009



Helmholtz Alliance

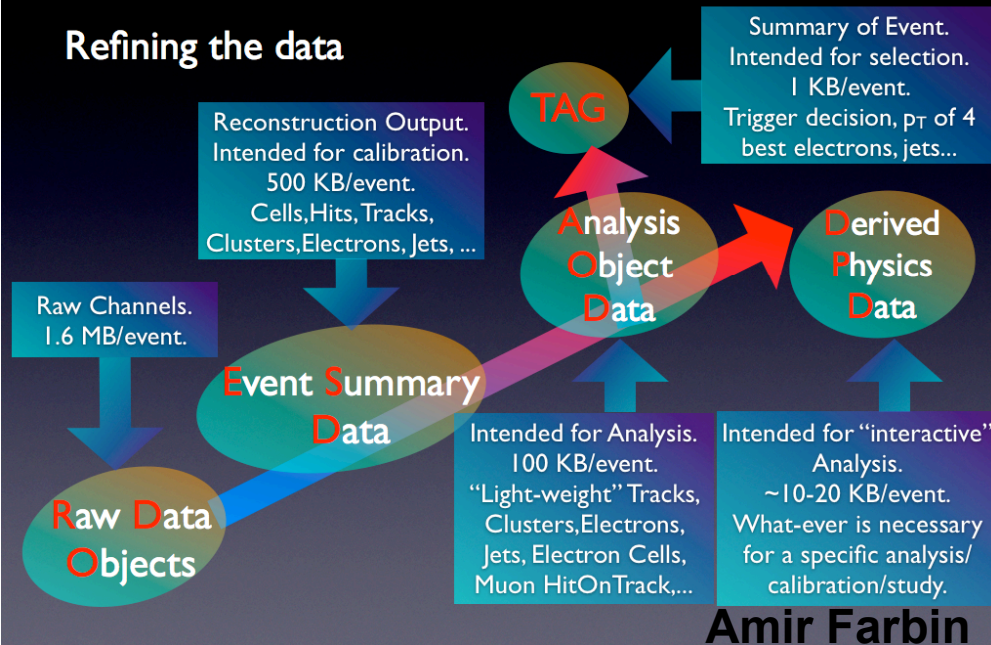


- > Helmholtz Alliance: Physics at the Terascale
- > Collaboration between ~20 german universities and research centers
 - People working on “Tera-eV physics”, e.g. LHC, ILC, and respective theorists
- > Many different research fields
 - Physics Analysis, Detector Technologies, Accelerator Physics, Grid Computing
- > <http://www.terascale.de> for more information

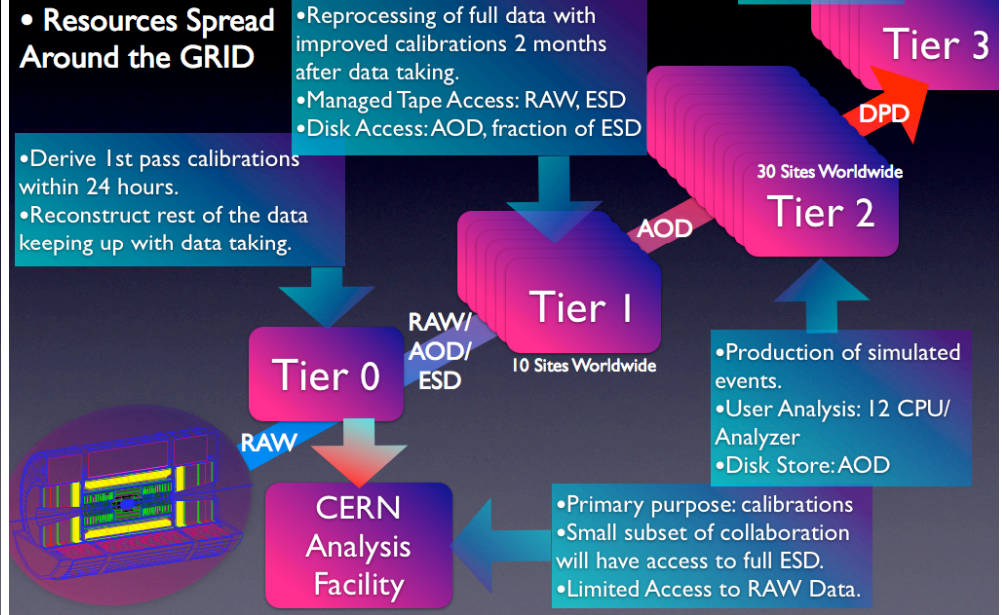
- > NAF is part of the Grid Computing research topic
 - Give users of the German institutes of the LHC and ILC experiments additional resources for analysis
 - Atlas, CMS, LHCb and ILC
 - Size ~1.5 average LHC Tier2, with emphasis on data storage
 - Intended as distributed facility, starting at DESY (with its two sites Hamburg and Zeuthen)

Place in the Computing Models (e.g. Atlas)

The Event Data Model



The Computing Model



- > NAF concept not foreseen by the LHC computing models
- > Different event descriptions and formats
 - At different stages of the Tier Model
- > End-user ready Analysis Format (e.g. AOD, DPD): At the end of the chain
 - Located at Tier2/Tier3 in the Grid Chain
- > Analysis Facility should ease access to these data: Place NAF at a Tier2!

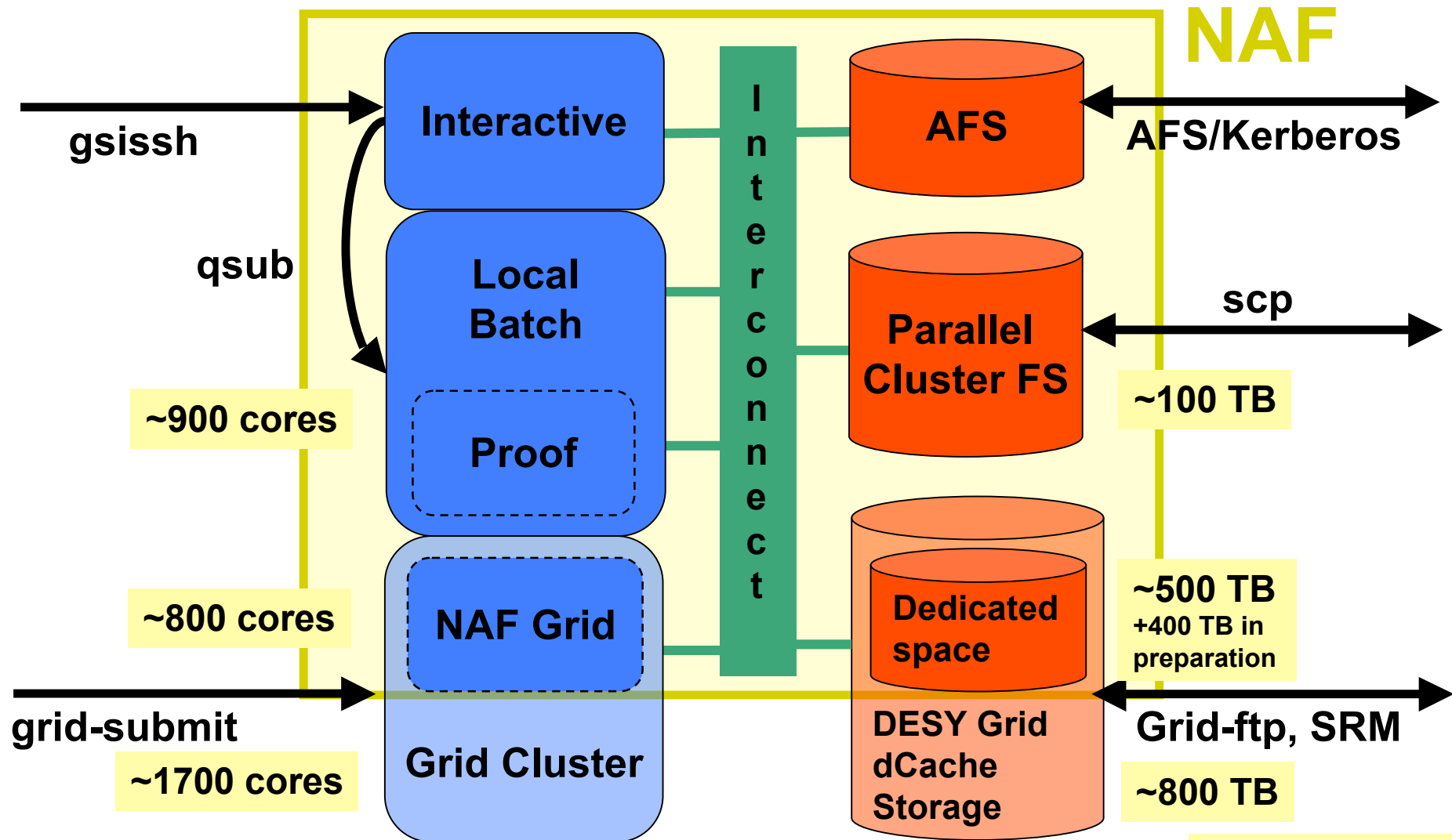
DESY Grid
Poster ID 328

Design key words

- Discussions and requirement papers from experiments
- Local Access to Grid Data
- Interactive \leftrightarrow Grid Integration
- Fast Turn-Around for Analysis
- Distributed Ansatz
- New Methods for Analysis
- Additional Grid CPU&storage



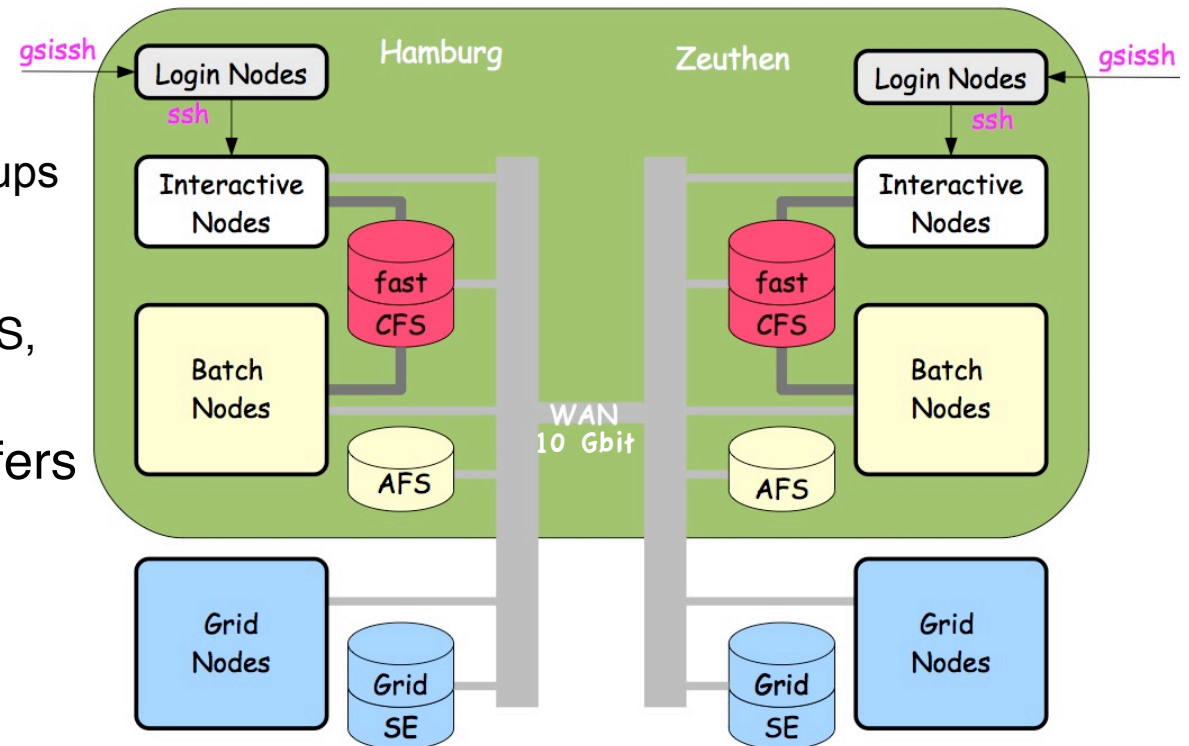
Building blocks



Status@1.4.09

And the implementation at two sites

- > Common effort between the DESY sites Hamburg & Zeuthen
 - Only legacy: Existing Grid CPU and storage infrastructure
 - Distributed over the DESY sites
 - Operated by the DESY IT/DV groups
- > Use DESY know-how
 - Build new instances: Registry, AFS, ...
- > Controlled WAN connections offers unique possibility to explore a distributed Analysis Facility
 - Offer best possible performance
 - Offer transparent access

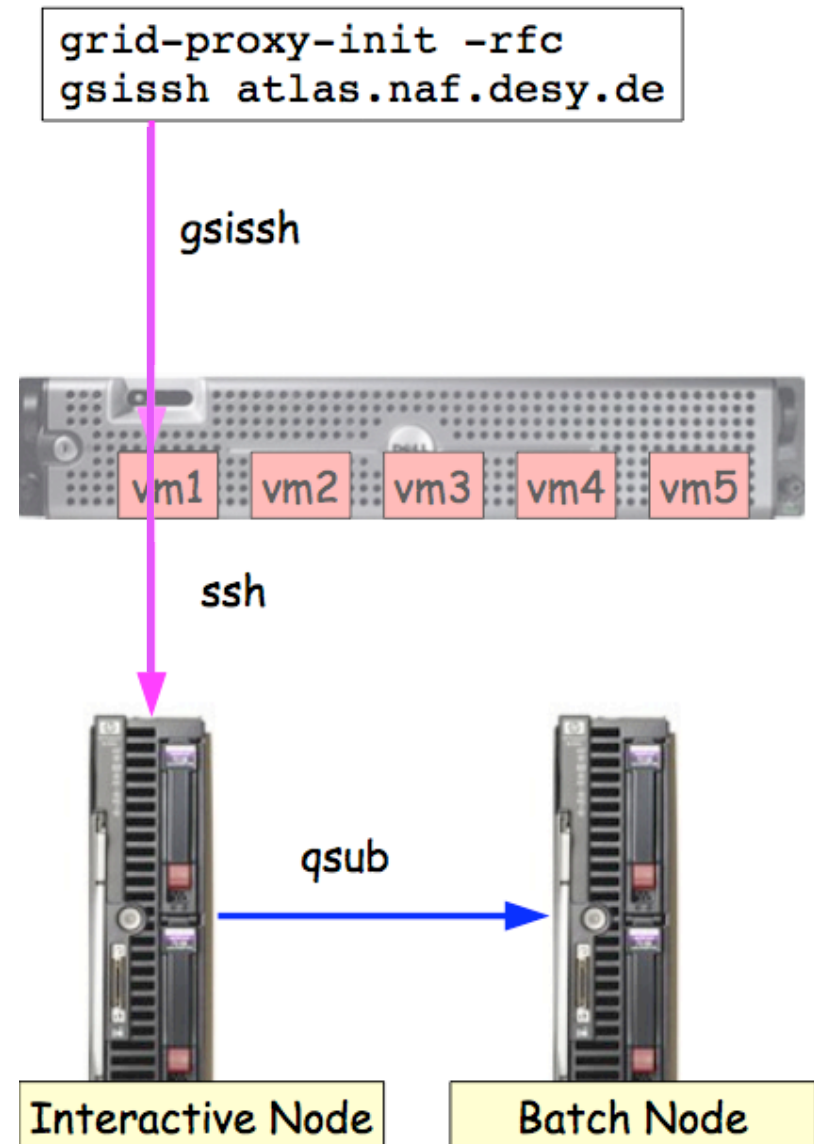


Usage of VOMS

- > DESY Grid infrastructure is generic and multipurpose.
- > Example CMS:
 - DESY Grid supports CMS as a WLCG Tier-2, has pledged resources
 - DESY Grid also supports German CMS members in NAF context in addition
- > Same Grid infrastructure for all, logical shares for different purposes
 - `cms:/cms` Normal Tier2 MoU User (20% of MoU share)
 - `cms:/cms/Role=production` Tier2 MoU production (80% of MoU share)
 - `cms:/cms/Role=lcgadmin` Software admin (basically high priority, no share)
 - `cms:/cms/dcems` German NAF Grid Users ([separate share](#))
- > Similar for Grid storage: Dedicated areas for German users
 - Within VO: All members can *read* all data from their VO

Access to Workgroup servers and AFS

- > *Idea:* Everyone doing LHC and ILC analysis has a Grid certificate
 - Can we use them to log into the interactive NAF?
- > Krb5 ticket & AFS Token generated from proxy certificate
 - Login node: failover&load-balanced
 - Each login node only serves one VO
- > Login node automatically redirects to Interactive Node
 - gsissh transparent to user
- > Possibility to get AFS token also outside of NAF
 - Remote access to AFS via proxy
- > Using Heimdahl Kerberos implementation!

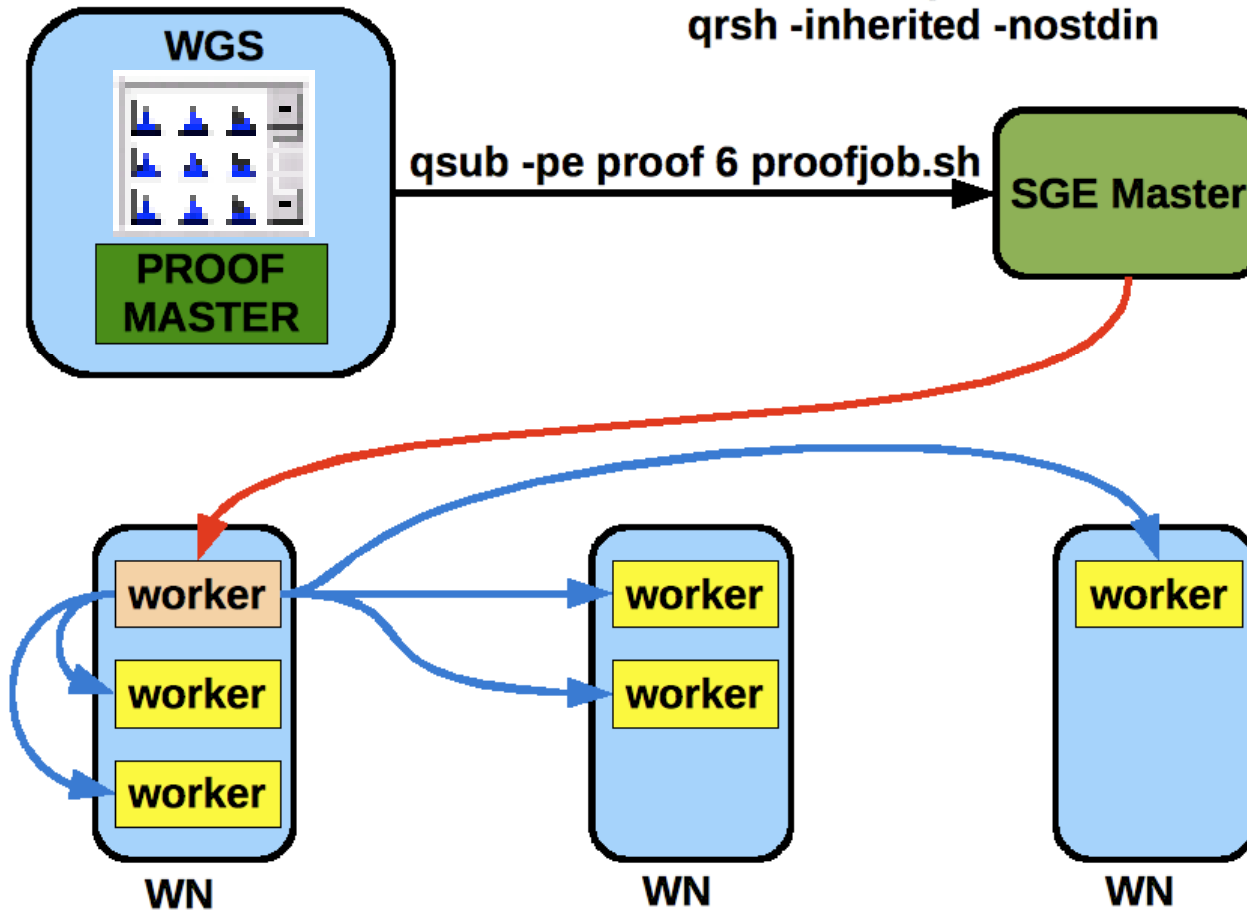


PROOF & SGE

 reserved slots by SGE

 starts master job

 starts slave jobs via
qsh -inherited -nostdin



- > Allows massive parallelization of analysis jobs
- > Keep interactive "ROOT prompt"
- > Used mainly by CMS (Uni HH)
- > Allows for multi-user and multi-group operations
- > Accounting & security possible

PROOF&SGE:
Poster ID 66

- > Tags - summary physics data for events
 - Efficient selection of interesting events
 - Direct navigation to these events
- > 2 formats
 - ROOT files: useful as indices to event
 - Relational Database: useful for querying: The format of choice!
- > 1 kB/event, includes pointers to AOD, ESD and RAW data files
- > TAG DB @ DESY: Currently being set up

Secure and consistent access to SE data and metadata

> Storage elements in Grid context:

- Should only use certificate/VOMS based authentication and authorization

> Protocols

- Gridftp (same as before)
- Gsidcap: Same as dcap, but with GSI authz
- GSI authz: Minimal overhead: $+O(500ms)$ per session

> E.g. ROOT supports gsidcap

> Meta-Data handling (e.g. file browsing)

- No file system mount necessary
- Secure GSI tools available



Further integrating X509 in NAF

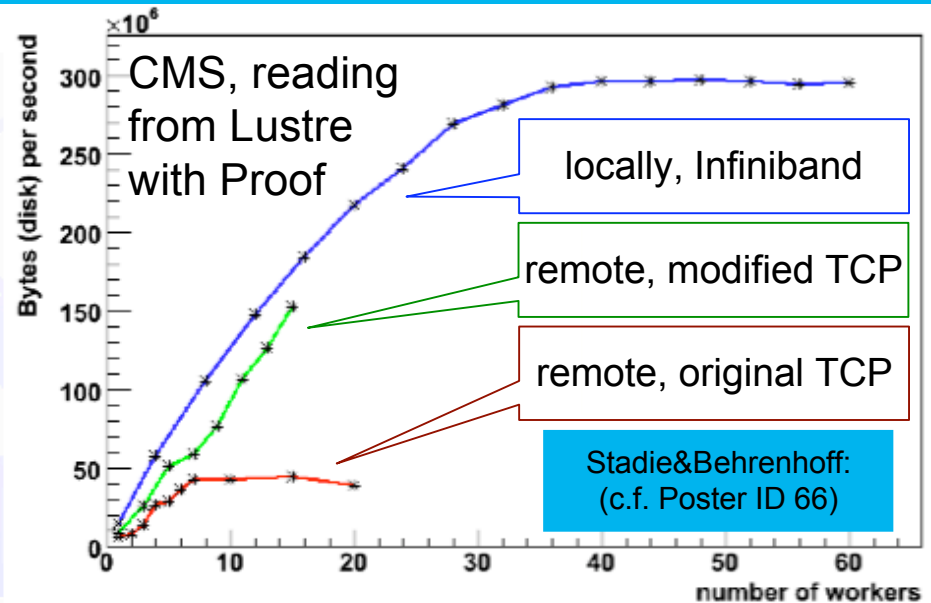
Certificate of Authenticity

- Gsish and X509 proxy used to login to NAF
 - Login proxy not available on NAF
- Users also need X509 proxy in the NAF
 - To access data e.g. via gsidcap
 - To submit Grid jobs
- Test implementation using MyProxy
 - User submits long MyProxy once
 - Automatically retrieved using k5 authz
 - Users: Set `$X509_USER_PROXY` is sufficient to use renewal



Storage Access: Speed

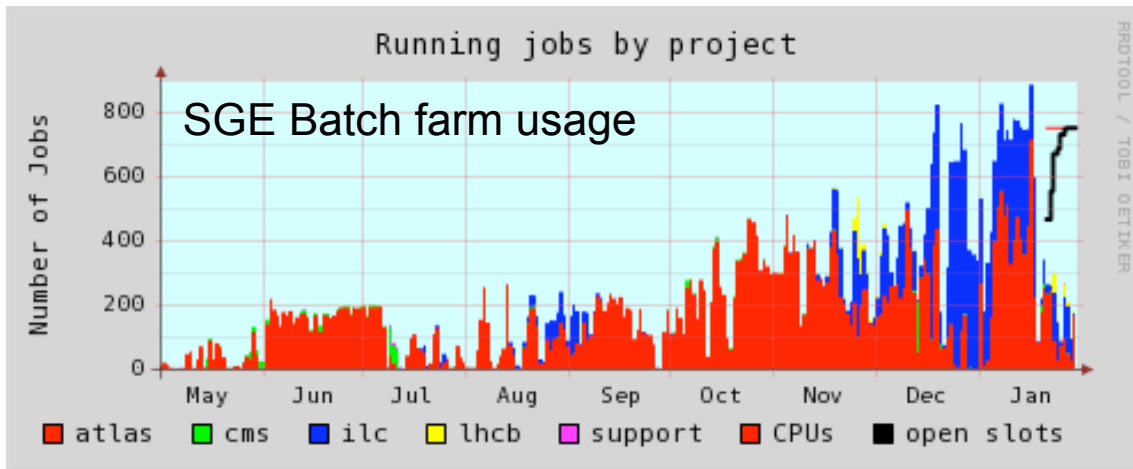
- > Grid Cluster and dCache optimized for overall performance
 - Analysis: Individual job should run as fast as possible
- > Lustre: Performance should be optimal for analysis
 - Used by experiments for different purposes
 - Infiniband connection “feels very fast”
 - WAN connection to Lustre also fast: TCP optimizations



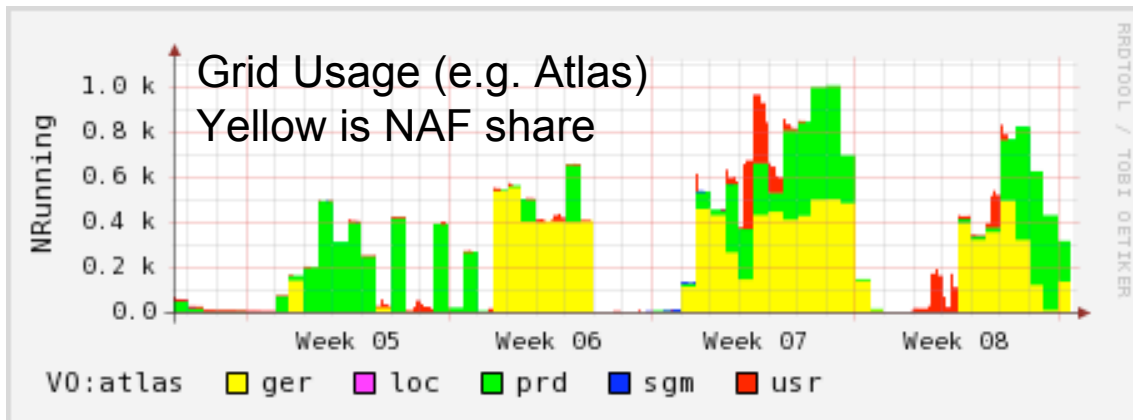
- > Need for Application benchmarks
 - MB/s is not necessarily the right measure!
 - Some benchmarks from CMS
 - Atlas recently developed HammerCloud: A suite of analysis applications that monitor performance
 - Will use this package to evaluate and optimize dCache and Lustre performance
 - Other real-life benchmarks welcome!

HammerCloud:
Talk ID 141

Running experience



- > Both Grid and interactive resources used!
- > Running smooth!
- > Support and contact to users important!



- NAF User Committee: Regular meetings and exchanges between users and operators
- Dedicated experiment mailing lists for experiment-specific questions
- First- and second-level support for fabric-related problems by DESY IT&DV

Summary and Outlook

- > The NAF is working: ~300 registered users
- > Hardware resources already substantial, enlargement in 2009
- > Generic approach:
 - All analysis workflows supported
 - All communities supported on one infrastructure
- > More information, documentation and links to support:
 - <http://naf.desy.de/>
- > **We all are waiting for our first great challenge:
The first LHC colliding-beam data!**
- > Questions? Comments? Welcome!
 - Yves.Kemp@desy.de / Andreas.Haupt@desy.de