Introduction to the National Analysis Facility

ATLAS-DESY NAF and FDR Tutorial — 12th May 2008

Wolfgang Ehrenfeld
University of Hamburg/DESY

Content

National Analysis Facility
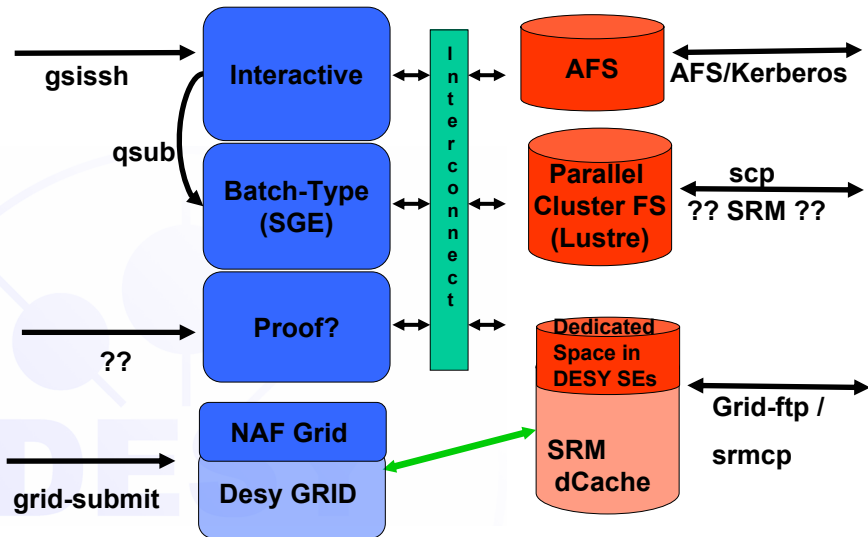
```
http://naf.desy.de
```

The National Analysis Facility (NAF) is part of the Strategic
Helmholtz Alliance (`http://terascale.de`) for German
particle physics.

Planned for a size of about 1.5 average Tier2, but with more
data.

The NAF provides:
- additional Grid resources
- interactive resources

NAF: Schematic basic layout

Accessing NAF Grid Resources

At the moment the additional Grid resources (worker nodes) are located at the DESY Hamburg site (`DESY-HH`). There is no dedicated NAF queue, but user jobs with the ATLAS-D role (`/atlas/de`) get a higher priority.

Proxy creation:
```
voms-proxy-init -voms atlas:/atlas/de
```

Proxy info:
```
voms-proxy-info -all

[snip]
attribute :  /atlas/de/Role=NULL/Capability=NULL
[snip]
```

Interactive Login

Authentification at the NAF will be done via a Grid proxy!

Advantages: No new password to remember.
Disadvantage: How to extend (get) AFS token?

Steps to login in:

1. setup recent Grid User Interface (UI) (Glite >= 3.1)
2. create a valid proxy in version 4 (option -rfc)
3. log into ATLAS NAF login server: atlas.naf.desy.de
4. log into ATLAS work group server: atlas-wgs[XX]

Interactive Login

Authentification at the NAF will be done via a Grid proxy!

Advantages: No new password to remember.
Disadvantage: How to extend (get) AFS token?

Steps to login in:

1. setup recent Grid User Interface (UI) (Glite >= 3.1)
2. create a valid proxy in version 4 (option -rfc)
3. log into ATLAS NAF login server: atlas.naf.desy.de
4. log into ATLAS work group server: atlas-wgs[XX]

1. ```
   source
   /afs/desy.de/project/glite/UI/etc/profile.d/grid-env.sh
   ```
2. `voms-proxy-init -rfc`
3. `gsi-ssh atlas.naf.desy.de`
4. `ssh atlas-wgs`

Software

DESY provides operating system (SL4/64bit) and general software:

- Grid User Interface (UI)
- ROOT

We provide ATLAS specific software:

- ATLAS kits and production releases
- EventView group areas
- Ganga

UI, ROOT and Ganga can be set up via the *ini script*.

Initial ATLAS software setup is done with
```
/afs/naf.desy.de/group/atlas/software/scripts/atlas_setup.py
   --create
```

Storage Overview

- AFS
- dCache
- lustre
- local disc

AFS

Your home directory is in AFS:

- e.g. `/afs/naf.desy.de/user/e/efeld`
- comes with backup (expensive storage)
- Quota: 500MB,
  can be increased $\rightarrow$ email to ATLAS NAF support

In addition there are resources for additional user, project and group space.

- with or without backup $\rightarrow$ expensive or cheap storage
- send request to ATLAS NAF support

AFS File Access

From NAF to outside:

- `scp/sftp`,rsync
- direct access to other AFS cells: use klog -cell NAME

From outside to NAF:

- `gsiscp/gsisftp` to login server (`atlas.naf.desy.de`)
- direct access to users `public` directory
- direct access to full user directory in the future
  (convert grid proxy into kerberos 5 ticket via heimdal)

dCache

dCache is the storage system used by DESY for the T2 and the NAF (Similar to Castor at CERN).

There are three protocols to access files in the dCache:

- native dCache (needs /pnfs mount)
- dcap (not very secure)
- gsidcap (needs valid grid proxy with VOMS extension)

Files can be accessed using:
  dccp (provided by Grid UI), root, athena, ...

- /pnfs mount only on ATLAS NAF work group server
- /pnfs mount read only in the future
- gsidcap the only allowed protocol in the future

Cluster File System - lustre

Ideas from IT:

- High Bandwidth (O(GB/s)) to large Storage (O(10TB))
- Copy data from Grid, process data, save results to AFS or Grid
- "Scratch-like" space, lifetime t.b.d., but longer than typical job

In contrast to dCache it looks like a normal file system:
```
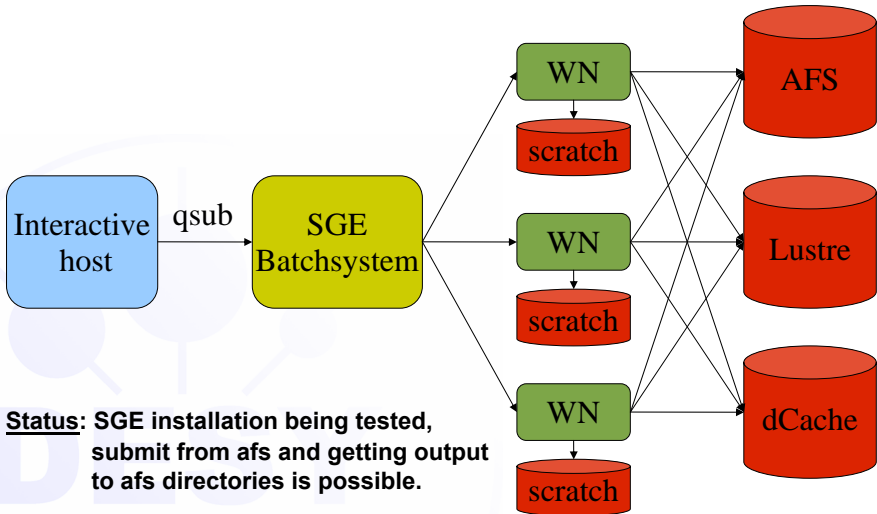/scratch/lustre-1.6/atlas
```

Batch System - SGE

The local batch system is running Sun Grid Engine (SGE).

It is similar to LSF (CERN) and PBS for the available features and commands (qsub, qstat), but has a different concept for managing resources, e.g. CPU time.

Specifications:

- $\sim$ 20 hosts (number growing)
- 8 cores per host
- 100 GB disk space per host
- 2 GB memory per core

Batch System

Status: **SGE installation being tested,**
submit from afs and getting output
to afs directories is possible.

Batch System - Resource

Time, memory, disk space and more are treated as consumable resources. If they are used up, the job is killed. If you request too much resources, the maximum number of jobs per host can not be filled up.

Hence you need to request the right amount of resources.

Examples:

- wall clock time: -l h_rt=24:00
- virtual memory: -l h_vmem=1GB
- stack for threads (needed for athena): -l h_stack=10M
- scratch disk space on worker mode: -l h_fsize=10G

Either specify resources on the command line or use the special comment (#$) in your job script.

Support and Documentation

As always this is the hard part!

Support:
- non-experiment specific: `naf-helpdesk@desy.de`
- ATLAS specific:
  - HN: `gridkaCloudUserSupport`
  - in the next days: `naf-atlas-support@desy.de`
  - in the future: `naf-atlas-helpdesk@desy.de`

Documentation (some is already written):
- general: `http://naf.desy.de`
- ATLAS@NAF: `http://naf.desy.de/atlas`

Summary

Check it out!

Summary

Check it out!

ATLAS NAF Wiki: Getting Started