Introduction to the National Analysis Facility

ATLAS-D MC Tutorial 2011 —
17th March 2011

Wolfgang Ehrenfeld
DESY

Content

National Analysis Facility

`http://naf.desy.de`

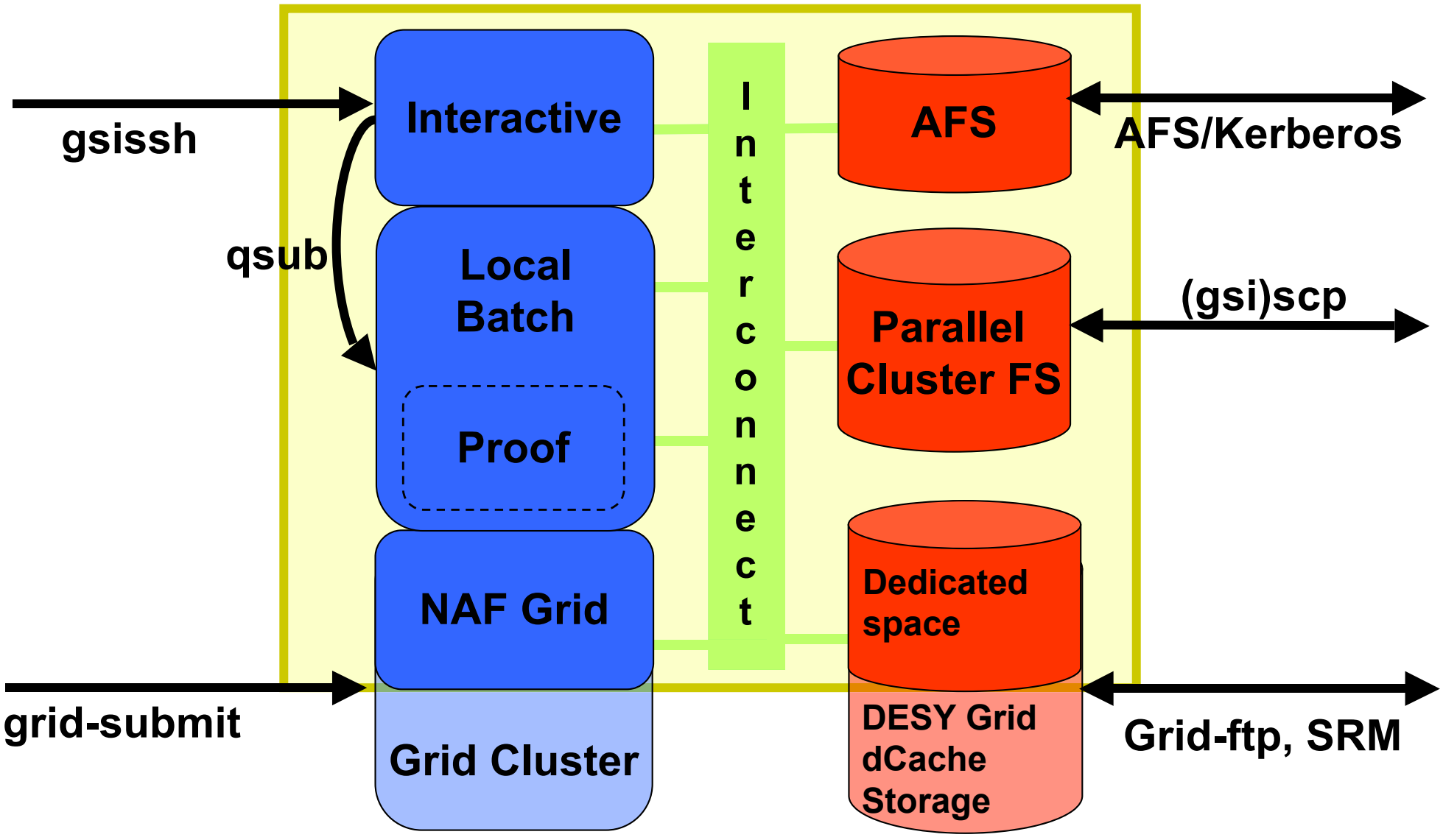
The National Analysis Facility (NAF) is part of the Helmholtz Alliance "Physics at the Terascale" (`http://terascale.de`).

The computing resources (storage and CPU) should be of the size of about 1.5 average Tier2, but with more emphasise on data storage.

The NAF provides:

- additional Grid resources
- interactive resources

to German particle physicists.


gsissh
Interactive
qsub
Local Batch
Proof
NAF Grid
Grid Cluster
Interconnect
AFS
Parallel Cluster FS
Dedicated space
DESY Grid dCache Storage
AFS/Kerberos
(gsi)scp
grid-submit
Grid-ftp, SRM

Interactive Login

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

- No new password to remember
- Getting an AFS token needs slightly more work

Steps to login in:

1. setup Grid User Interface (UI)
2. create a valid proxy in version 4 (option -rfc)
3. log into ATLAS NAF login server: atlas.naf.desy.de and you will be forwarded to one of the ATLAS work group server (load balancing)

Interactive Login

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

- No new password to remember
- Getting an AFS token needs slightly more work

Steps to login in:

1. setup Grid User Interface (UI)
2. create a valid proxy in version 4 (option -rfc)
3. log into ATLAS NAF login server: atlas.naf.desy.de and you will be forwarded to one of the ATLAS work group server (load balancing)

1. `source /afs/desy.de/project/glite/UI/etc/profile.d/grid-env.sh`
2. `grid-proxy-init -rfc`
3. `gsissh -X atlas.naf.desy.de`

Software

DESY provides operating system (SL5 64bit) and general software:

- Grid User Interface (UI)
- ROOT, pyROOT
- python 2.5

We provide ATLAS specific software:

- ATLAS kits and production releases
- Ganga, panda client, dq2
- scripts

UI, ROOT, python, Ganga, dq2, . . . can be set up via *ini* script.

Initial ATLAS software setup is done with:

```
ini atlas
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 → email to ATLAS NAF support

Your also have an AFS scratch directory:

- e.g. `/afs/naf.desy.de/user/e/efeld/scratch`
- comes without backup
- for small files, e.g. log files, histograms, . . .
- Quota: 1GB,
  can be increased → email to ATLAS NAF support

There is space for additional user, project and group space.

- with or without backup → 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 *cell name*
- this is the preferred direction!

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 via `naf_token` script (convert grid proxy into kerberos 5 ticket via heimdal)

Lustre

Lustre is a clustered file system for fast, multi client access.

- high bandwidth (200 MB/s write/ 600 MB/s read) to large Storage (O(10TB))
- copy data from Grid, process data, save results to AFS or Grid
- in the NAF Lustre space is scratch space (no backup and data life time of a fwe month)

In contrast to dCache it behaves like a normal file system:

`/scratch/hh/lustre/atlas` → 79 TB
`/scratch/zn/lustre/atlas` → 20 TB
`/scratch/zn/lustre/atlas2` → 36 TB

- we would like to automatically delete files not used for a few month!

Batch System - SGE

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

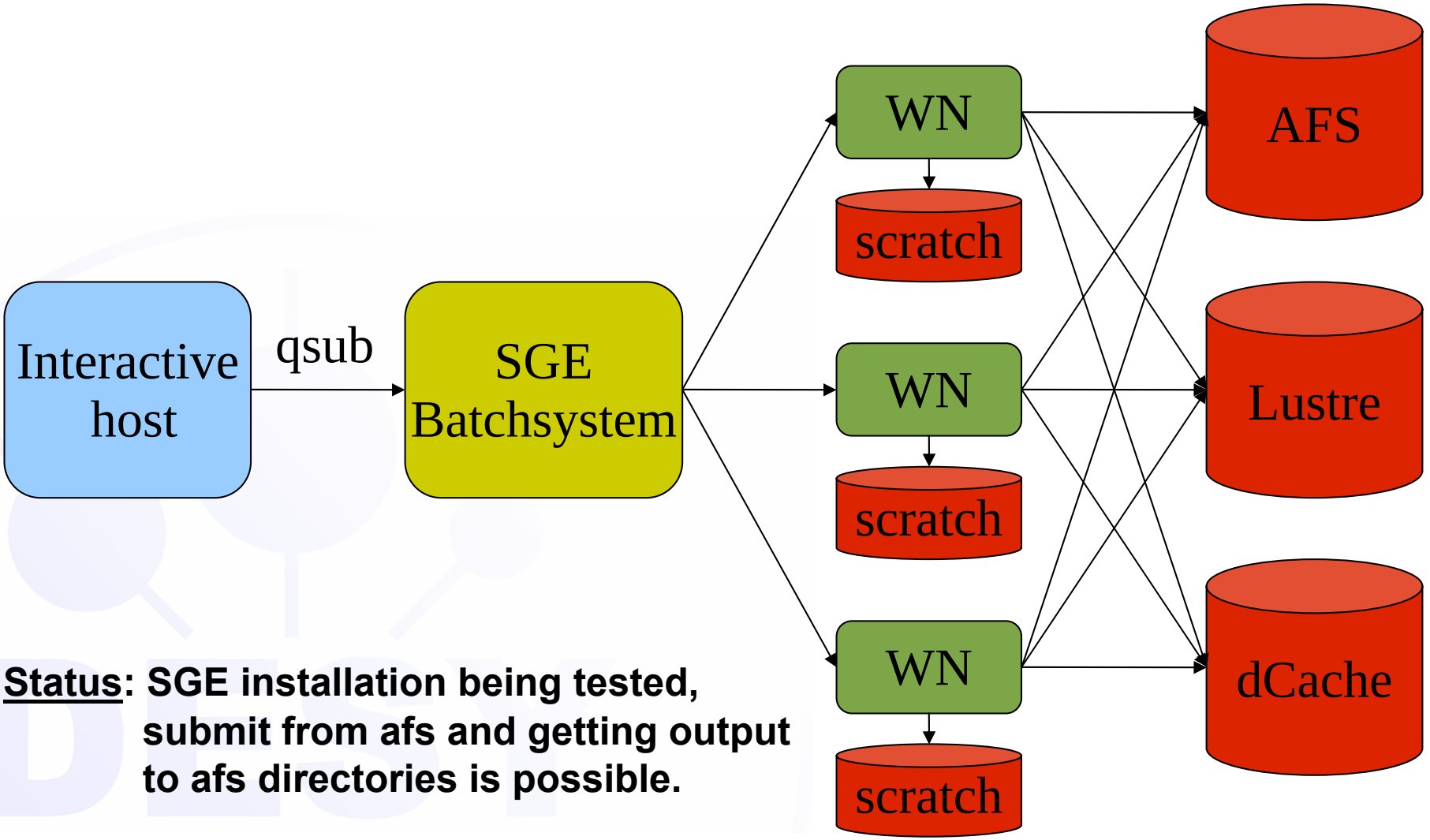
It is similar to LSF (CERN) and PBS for the available features and commands (`qsub`, `qstat`), but has a different concept for managing resources, e.g. CPU time and memory.

Specifications:

- $\sim$ 170 hosts or $\sim$ 1600 slots
- ATLAS share is $\sim$ 30%
- 8/12 cores per host
- 100 GB disk space per host
- 2-4 GB memory per core

You can see the job meta data of your finished jobs at:
`https://www-zeuthen.desy.de/dv-bin/batchssl/stat`

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_cpu=24:00:00
- virtual memory: -l h_vmem=1G
- 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.

Additional Tools/Services

The NAF admins and the ATLAS NAF team provides some tools, scripts and services to ease the use of the NAF:

naf_token:	script, which converts a Grid proxy into a NAF AFS tokens
autoproxy:	tools and service to automatically extend Grid proxy
dctools:	tools to mimic `/pnfs` file system from dCache without the actual mount (needs valid Grid proxy)
scripts:	creation of requirements file, get_tag, dq2-poolFCjobO.py

Support and Documentation

As always this is the hard part!

Support:

- non-experiment specific: `naf-helpdesk@desy.de`
- ATLAS specific:
  - `naf-atlas-support@desy.de`
  - HN: `gridkaCloudUserSupport`

User Communication:

- NAF User Committee: `http://naf.desy.de/nuc`
  Marcello Barisonzi, Wolfgang Ehrenfeld

Documentation (feel free to contribute):

- general NAF: `http://naf.desy.de`
- ATLAS@NAF: `http://naf.desy.de/atlas`

Summary

Log in!

In the rest of this tutorial you will:

- log into the NAF,
- setup athena,
- and run the HelloWorld example