



HAMBURG • ZEUTHEN

The National Analysis Facility:

Structure

First experiences

Next Steps

Yves Kemp for the NAF collaboration

DESY IT Hamburg & DV Zeuthen

3.3.2008

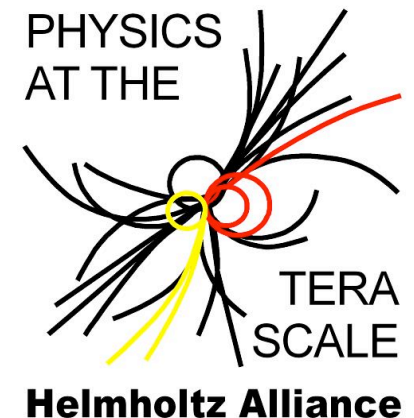
DPG Frühjahrstagung

Freiburg

NAF and the TeraScale Alliance



- **Motivations for a NAF**
 - **LHC Tier Model and Grid technology established for T0-T2**
 - Well proven for production and large-scale analysis
 - **Will the Tier Model work for small-scale analysis?**
 - **Need new methods (non-Grid) for large-scale analysis?**
 - **Reflect the physics cooperation within Germany in IT facilities**
- **The frame of the NAF:**
 - **Part of the Helmholtz Alliance**
 - **Only accessible by german research groups for LHC/ILC tasks**
 - **Planned for a size of about 1.5 av. Tier 2, but with more data**
 - **Starting as joint activity @ DESY, adding more sites later**



Considerations for NAF design

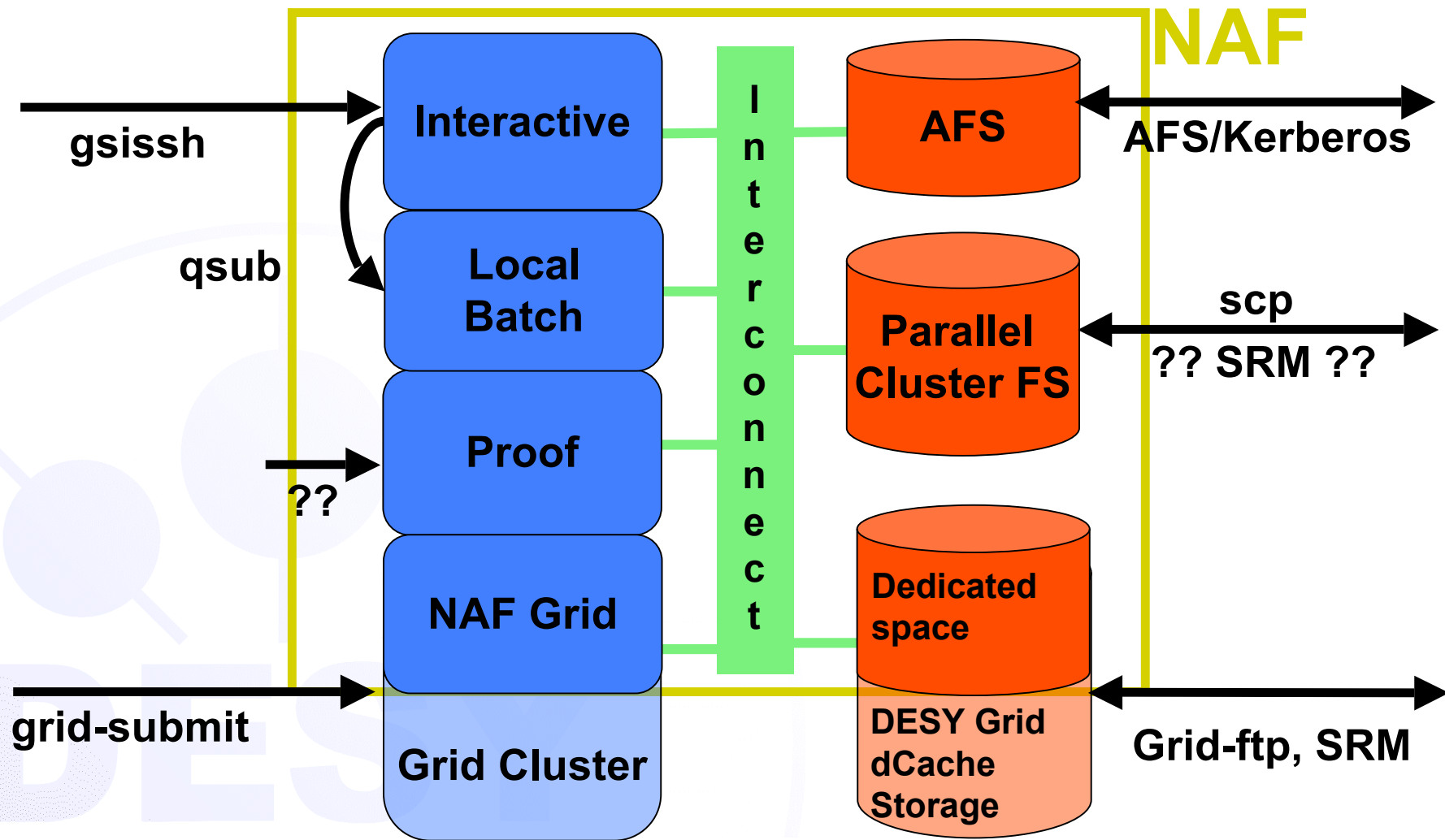
- **What is needed for efficiently doing analysis? What is missing?**
 - Not only technical, but also conceptual & organizational questions
- **Where is the boundary the Grid and my Local Facility?**
 - What can/must be done on the Grid?
 - What can/must be done locally (without Grid)?
- **How to handle the different worlds Grid and Local in a consistent, secure and high-performance way?**
- **The Data is the Central Part**
 - For Analysis, you want “one big heap” of data ...
 - ... that you can dig very efficiently ...
 - ... with different and changing tools and programs
- **Together with experiments, we try to answer these questions**
 - Not always easy, as no cookbook answers ready...

Starting with Atlas & CMS



- **Requirement papers. Some points:**
- **Interactive login**
 - Code development & testing, Experiment SW and tools
 - Uniform access
 - Central registry
- **Personal/group storage**
 - AFS home directories (and access to other AFS cells)
- **High-capacity /High-bandwidth storage**
 - Grid & local (with backup)
 - Grid-part: Enlargement of the T2 part
- **Batch-like resources:**
 - Local access: short queue, for testing purpose
 - Fast response wanted
 - Large part (only) available via Grid-mechanisms
- **Hosted Data:**
 - AODs (Full set in case for Atlas, maybe trade some for ESD?)
 - TAG database
 - User/Group data
- **Additional services**
 - PROOF farm, with connection to high-bandwidth storage
- **Flexible setup**
 - Allows reassignment of hosts between different types of services

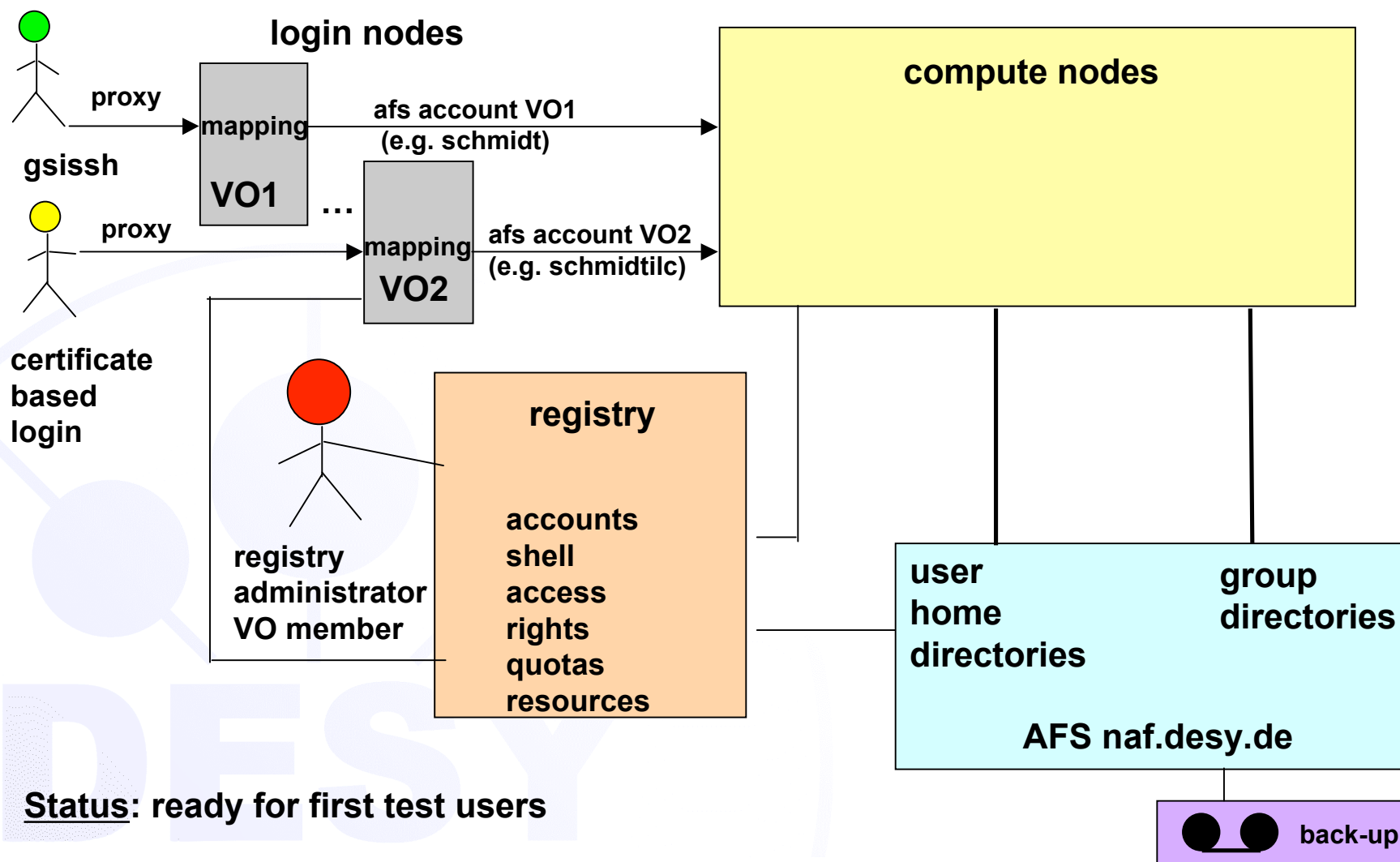
First sketch of the infrastructure



The first hardware

- **Hardware Computing:**
 - 6 x 16 DualCPU-Quadcore Blades (HP-Proliant BL460c)
 - 2GB RAM/core, 146 GB HD/Blade
 - Infiniband HCA
- **Hardware Storage:**
 - 7 x SUN thumpers (17.5TB/box at raid 6) for dCache pools & Lustre
 - 8 x DELL Poweredge 2950 with 8x146 GB SAS Disks for infrastructure and AFS
- **Other hardware: Racks, Infiniband Switch, Infrastructure servers, ...**
- **Hardware assignment flexible:**
 - **Following needs of experiments**

NAF login, interactive



IO and Storage

- **New AFS cell: naf.desy.de**
 - User & Working group directories
 - Special software area
 - Safe and distributed storage→ Working



- **Cluster File System**
 - High Bandwidth ($O(\text{GB/s})$) to large Storage ($O(10\text{TB})$)
 - Copy data from Grid, process data, save results to AFS or Grid
 - “Scratch-like” space, lifetime t.b.d., but longer than typical job→ Currently under construction



- **dCache**
 - Well-known product and access methods
 - Central entry point for data import and exchange
 - Special space for german users→ In production use



First peek @ Login&AFS



HAMBURG • ZEUTHEN

```
Terminal — ssh — 74x30
[lxplus217] /afs/cern.ch/user/k/kemp > grid-proxy-init -rfc
Your identity: /O=GermanGrid/OU=DESY/CN=Yves Kemp
Enter GRID pass phrase for this identity:
Creating proxy .....
..... Done
Your proxy is valid until: Fri Feb 29 20:08:59 2008
[lxplus217] /afs/cern.ch/user/k/kemp > gsissh tcsh6-vm1.naf.desy.de
Last login: Fri Feb 29 08:07:06 2008 from lxplus215.cern.ch
[kemp@tcsh6-vm1]~% klist
Ticket cache: FILE:/tmp/krb5cc_1006_v1Ez6A
Default principal: kemp@NAF.DESY.DE

Valid starting    Expires          Service principal
02/29/08 08:09:15  03/01/08 09:09:15  krbtgt/NAF.DESY.DE@NAF.DESY.DE
                renew until 03/30/08 09:09:15

Kerberos 4 ticket cache: /tmp/tkt1006
klist: You have no tickets cached
[kemp@tcsh6-vm1]~% grid-proxy-info
subject   : /O=GermanGrid/OU=DESY/CN=Yves Kemp/CN=955184393/CN=935890847
issuer    : /O=GermanGrid/OU=DESY/CN=Yves Kemp/CN=955184393
identity  : /O=GermanGrid/OU=DESY/CN=Yves Kemp
type      : RFC 3820 compliant impersonation proxy
strength  : 512 bits
path      : /tmp/x509up_p7560.fileE0B16B.1
timeleft  : 11:59:36
[kemp@tcsh6-vm1]~% █
```

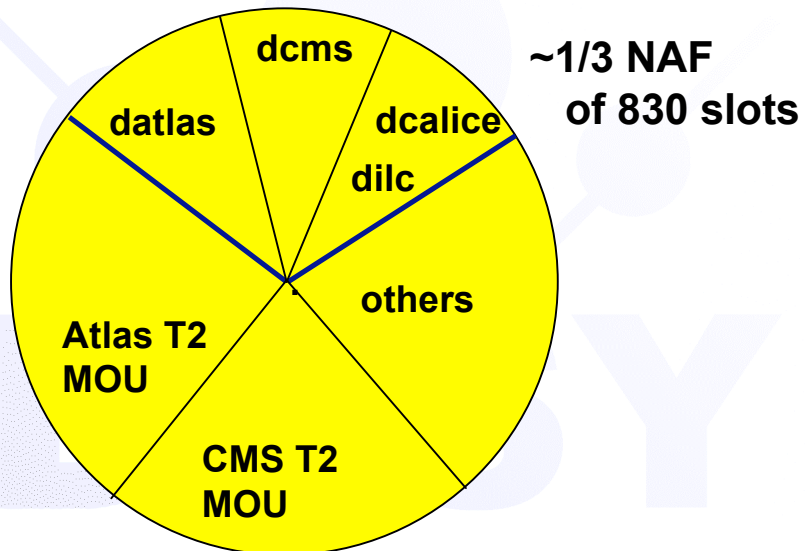
```
Terminal — ssh — 74x30
[kemp@tcsh6-vm1]~% ssh tcx031
Last login: Fri Feb 29 08:11:25 2008 from tcsh6-vm1.naf.desy.de
[kemp@tcx031]~% ls -l
total 6
drwxr-xr-x 3 kemp support 2048 Feb 28 10:37 private
drwxr-xr-x 2 kemp support 2048 Feb 28 10:37 public
drwxr-xr-x 2 kemp support 2048 Feb 28 10:37 scripts
[kemp@tcx031]~% pwd
/afs/nafe.desy.de/user/k/kemp
[kemp@tcx031]~% ls -l /afs/cern.ch/user/k/kemp/public
total 1252
-rw-r--r-- 1 899 1399 1282048 Feb 29 08:12 plots.root
[kemp@tcx031]~% echo hostname | qsub
Your job 42 ("STDIN") has been submitted
[kemp@tcx031]~% cat STDIN.o42
tcx036.naf.desy.de
[kemp@tcx031]~% free
              total        used        free      shared    buffers     cached
Mem:           16633032     675612     15957420          0       234320       371188
-/+ buffers/cache:        70104     16562928
Swap:           8385888          0       8385888
[kemp@tcx031]~% █
```



Batch @ NAF

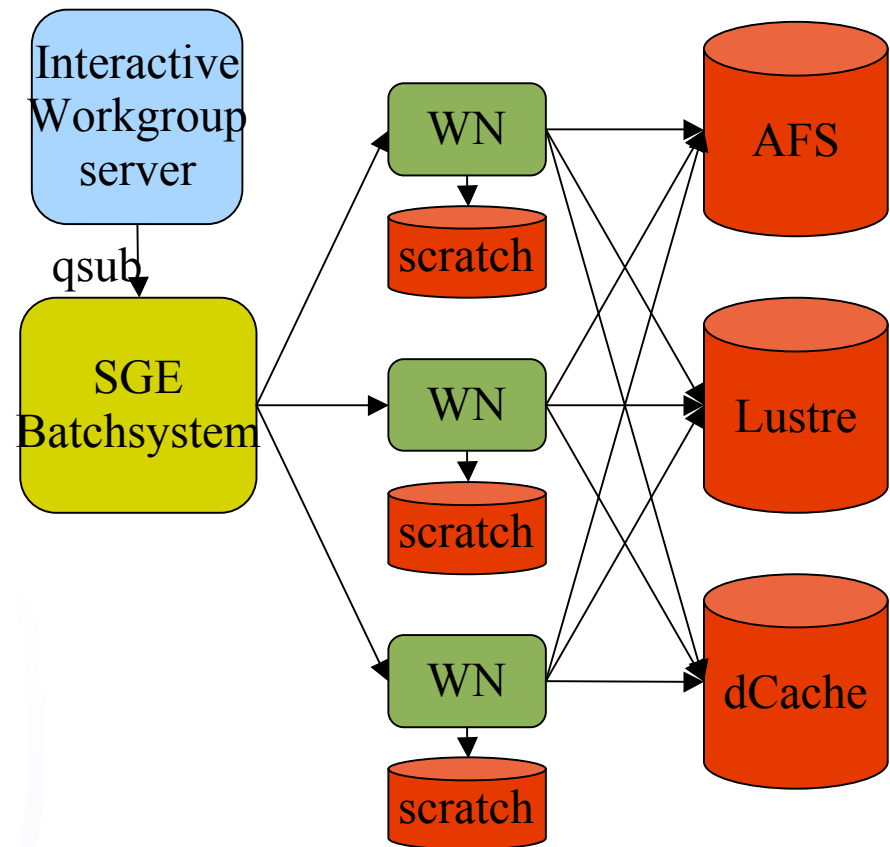
■ Grid Batch

- Integrated in Desy Grid Infrastructure (grid-ce3.desy.de)
- Dedicated Fairshare with higher priority
- Access using VOMS proxies
- In use, already first results!



■ Local Batch

- Ready for testing



- **Experiment specific software: Grid and Interactive world:**
 - DESY provides space and tools
 - Experiments install their software themselves
 - Because of current nature of Grid and Interactive parts: Two different areas
- **Common software:**
 - Grid world: Standard worker node installation
 - Interactive world:
 - Workgroup server installation: Compilers, debuggers...
 - No Browser, Mailclient,
 - ROOT, CERNLIB??? (Are the ones shipped with the experiment frameworks OK?)
- **Operation System:**
 - Currently all Grid WNs on SL4
 - InteractiveSL4 (64bit) (maybe some SL5). No SL3

- **Technical aspects:**
 - DESY can provide technical tools like mailing lists, request tracker, wiki, hypernews... if needed
 - GGUS might be integrated
- **Organisational aspects**
 - The experiments **MUST** provide first level support
 - Filter user questions
 - Transmit fabric issues to NAF admins
 - DESY will provide second level support
- **We NAF operators need fast feedback: NAF Users Board**
 - E.g. two experienced users from each experiment for technical advisory and fast feedback + 2 NAF operators

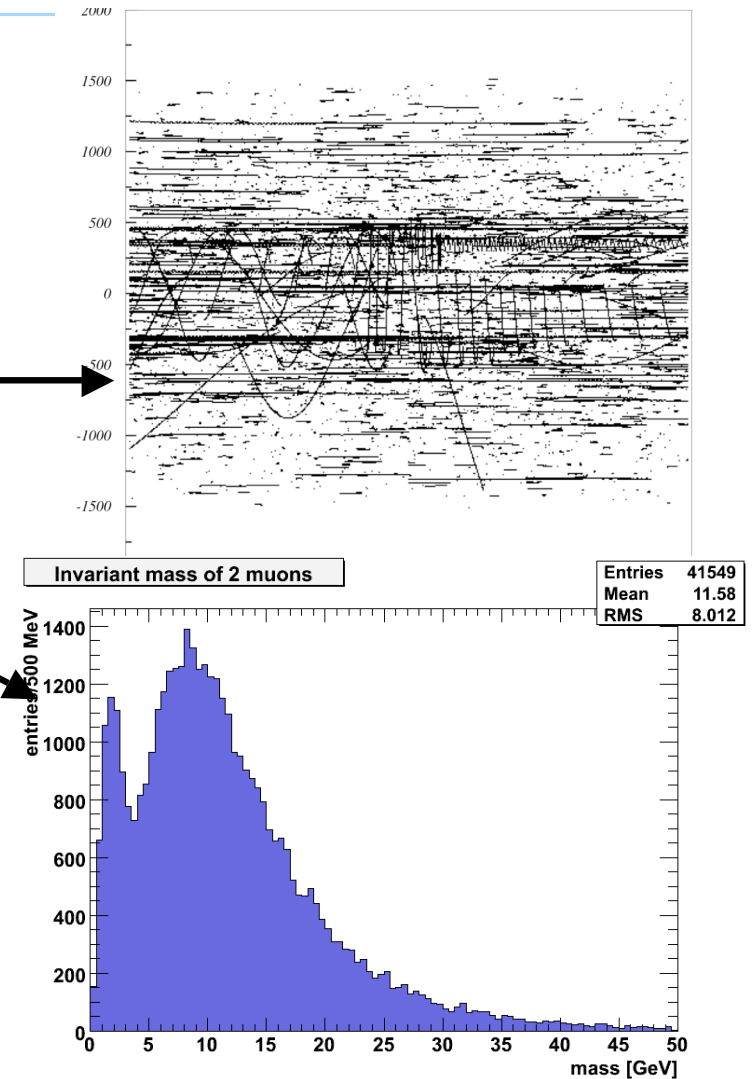
Timeline



- **Some parts already work: Grid-Batch, Grid-Storage**
- **For AFS and first interactive environment (login,WGS, local batch)**
 - Three steps planned:
- **Alpha:**
 - Started end of February
 - With chosen local people for fast interaction
- **Beta:**
 - ~two weeks later
 - Experiments/HGF Grid Project Board nominates experienced test users
- **Public operation:**
 - When beta phase is finished successfully
 - All German users
- **For other components: Similar process**
 - Lustre: Hopefully soon a test instance
 - **PROOF: Need more interaction with experiment experts, not in near future**

First user experience: Feedback

- Main users using the NAF Grid share:
 - Adrian Vogel (ILC): ~8k h, 10k Jobs
 - Machine-induced background studies, full Geant4 detector simulation
 - Manuel Giffels (CMS): ~70 h, 100 Jobs
 - Private background production (bb->2mu)
 - Walter Bender, Daisuke Tornier (CMS):
 - Exotica, private Alpgen production
- Experience:
 - VOMS group works fine, no problems



Summary & Outlook



- **Grid part of the NAF already working and used for physics!**
- **Interactive part of the NAF being tested by chosen local people**
- **Missing ingredients:**
 - **Lustre: Still some work to do, but working hard**
 - **PROOF: Schedule unclear, need interaction with experiments**
- **Some open questions**
 - **NAF Users Board**
 - **Experiments: First-Level-Support, Databases**