

Grid Engine at DESY

Grid Engine at DESY



Pirmin Fix
Thomas Finner, Owen Synge
Grid Engine at DESY
HEPiX
Prague, 25.04.2012

Agenda

> Environment

- DESY
- BIRD
- NAF

> Upgrade

- Batch system comparison
- Installation
- Testing

> Future ideas

- Ideas
- Outlook
- Summary



DESY

BIRD

NAF



DESY Environment

➤ Several Batch Systems in use:

➤ Sun Grid Engine

- BIRD
- NAF

➤ Univa Grid Engine

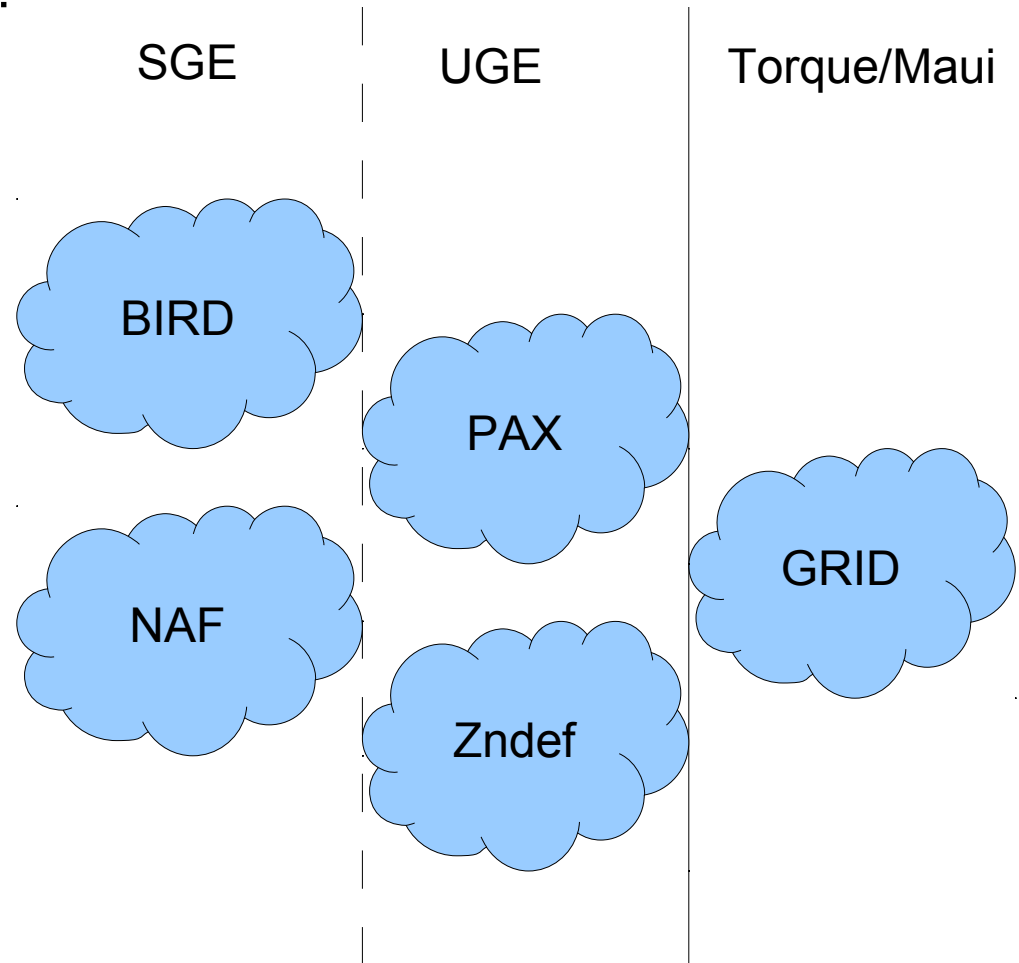
- PAX
- Zeuthen default

➤ Torque/Maui

- GRID

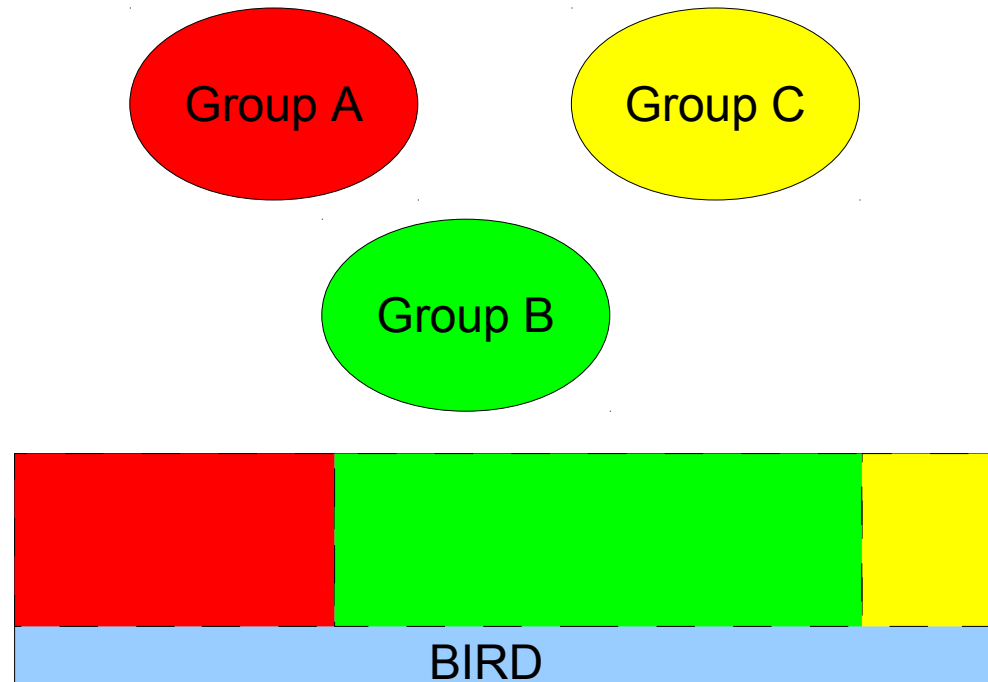
➤ Others

- Independent from IT

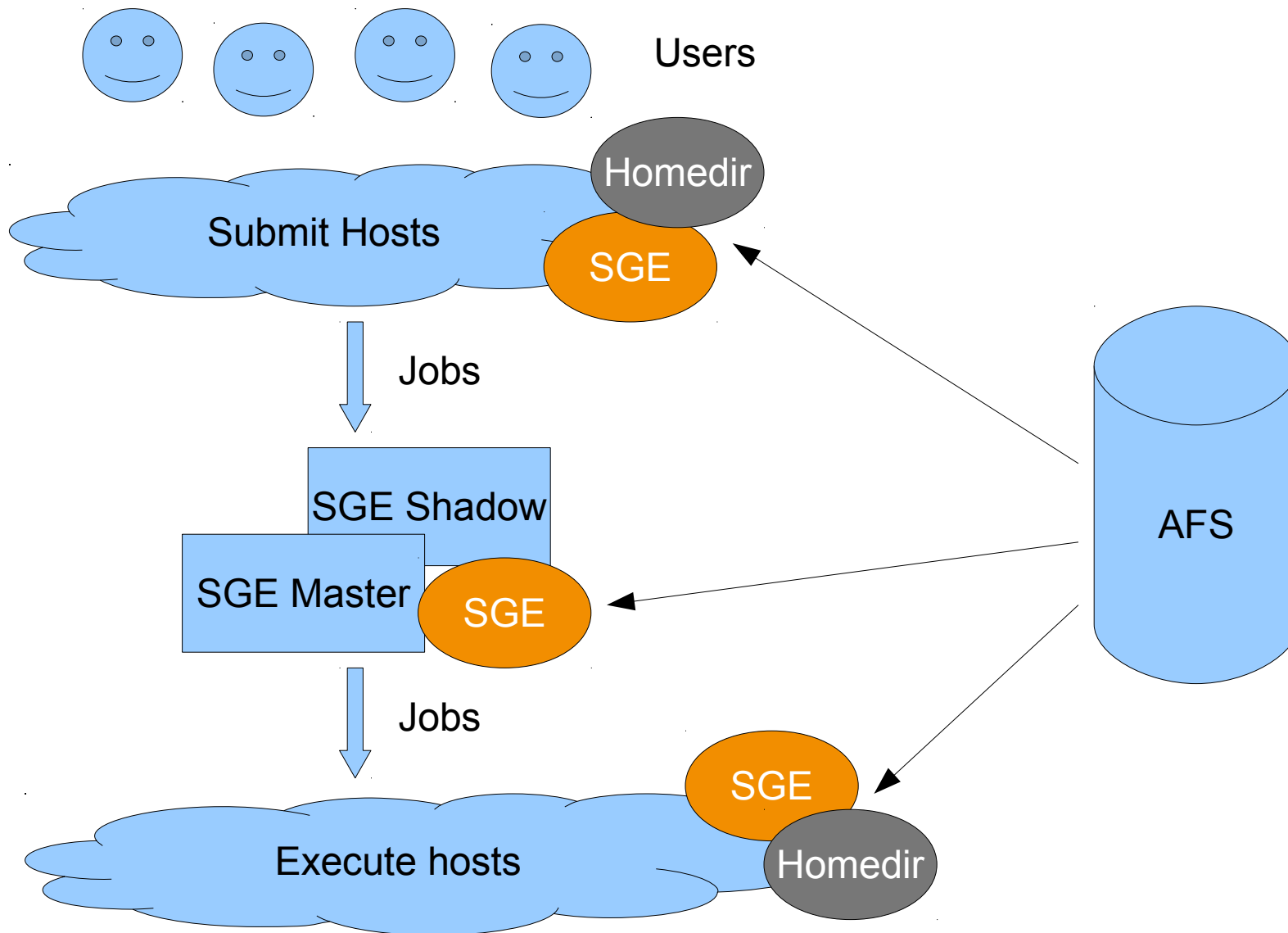


BIRD – Batch Infrastructure Resource at DESY

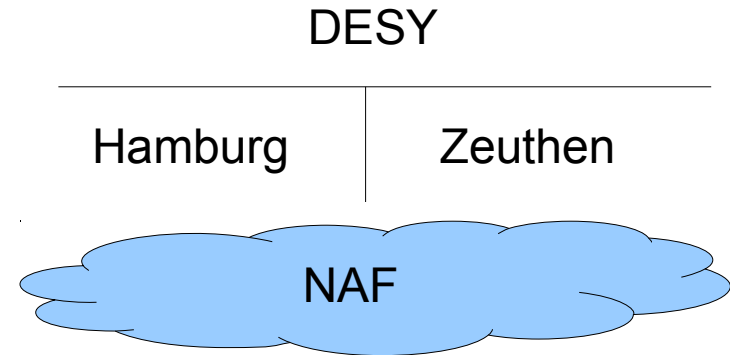
- Uses Sun Grid Engine
 - Submit hosts are publicly available
 - You need to be registered
- Open for registered DESY user
- Fair share
 - Buy your own resources and get guaranteed access
 - If somebody else doesn't need theirs, you can use that as well
- GPU Support



BIRD and SGE



- DESY Hamburg and DESY Zeuthen Co-Locations
- Several external projects/VOs



BIRD

Comparisons

Upgrade

Tests



Reasons for Upgrade

- From SGE 6.2u5 to SoGE 8.0.0d
- Benefit from development
- Bug fixes (and security)
- Free software
- Keeping the chance for further improvements
 - Not loosing compatibility to newer versions



- Oracle Grid Engine
- Univa Grid Engine (Supported)
- Univa Grid Engine Core
- Open Grid Scheduler (Grid Engine 2011.11)
- Son of Grid Engine



- Do cost money
- Provides support
 - Can help with special cases and need for development
 - We just have 'default' jobs
- More tools around SGE for Enhancement and Management
 - Our environment works as it is



> Univa Grid Engine Core

- Version from github (github.com/gridengine/gridengine)
- Didn't look active for about a year in public repository

> Open Grid Scheduler aka Grid Engine 2011

- Seems to include patches from SoGE
- Seems to be more conservative
- Didn't further investigate as SoGE works
- Might consider it for future use as Grid Engine derivatives are hopefully still interchangeable

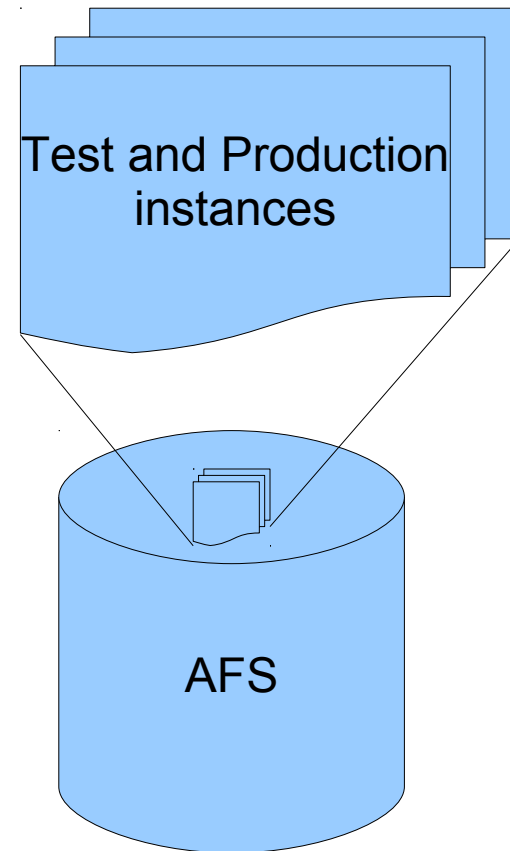
> Son of Grid Engine

- Recent changes in repository
- Seems to be active
- Based on comments from the mailing list seems to be more liberal to new features
- (provides rpm packages)



Installation

- Installed at same location as before
- AFS on every node
- Using export/import script for “old” configuration
- Own compilation
 - RPMs not possible due to our multi-version and central repository
 - Per SGE version a directory
 - Link to most recent version



Tests to perform before upgrade

- Test instance parallel to life instance
- Dependencies
 - hwloc not included in SoGE
 - Different repositories have different versions
 - Included in our distribution fixed to the newest version available
- Does the old config load?
- Is configuration used?
- Submitting
- Logging
- Kerberos Credentials
- AFS Tokens
 - To access home directory



More robust SoGE

GPU Support

Graphical Access

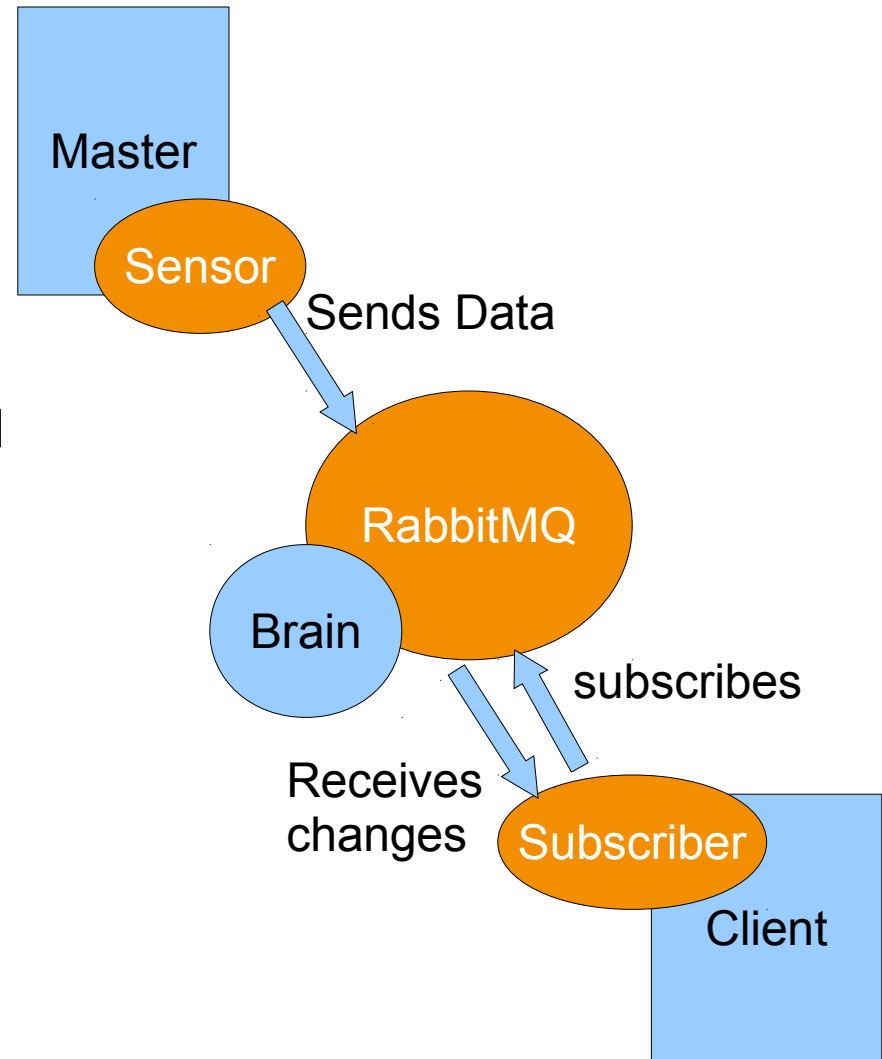
Substitution of Torque with SoGE in Grid Cluster

Cloud



More robust SoGE

- Move SGE to local disk
 - Become independent of AFS during bootup
- Some parts of SGE rely on shared filesystem
 - Some mechanism to update files in another way
 - Event based message passing system
- Work in progress



- Tools rely on specific versions
- Represent GPU as Consumable / Complex Resource/Configuration in SoGE
- Right now only one Job / GPU-Host
- S(o)GE has no control over GPU
 - Not like control over CPU & Mem
- Virtualization?



Graphical Access

- Interactive access to batch WN
- With GUI (for displaying graphs etc)
- Desired in NAF and BIRD
- Set display variable correctly
- S(o)GE uses own SSH client
- Performance
- Platform support
 - Some distributions don't allow remote applications to access X
 - User education

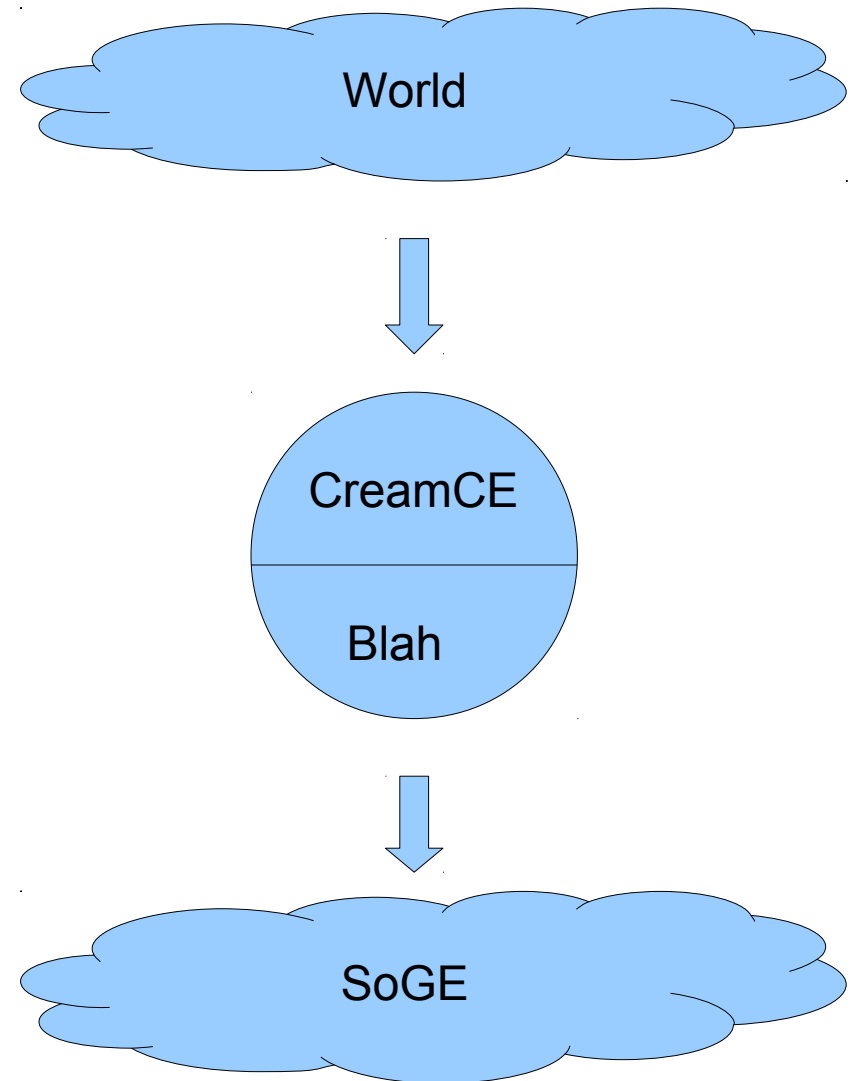


- Scaling issues
- Configuration of the Test Environment
 - User Management
 - Sandboxes (Movement of user files)
 - Accounting



SoGE GRID Integration progress

- EMI 1 CreamCE with EMI 1 Blah
 - Does not work
- EMI 1 CreamCE with EMI 2 Blah
 - Works
 - Looks promising
- EMI 2 CreamCE with EMI 2 Blah
 - Released 7th of may
 - We will evaluate



- Just a bunch of ideas
- Dynamic Queues for generating VMs
- Dynamic OS-Versions
- Individual Images
 - Virtualization Group
- Separate Hardware from Operating System
 - Eases hardware maintenance and responsibility distribution
- Elastic Infrastructure
- Integration in Scheduler



Summary

- SoGE is regularly updated
- Does work with our settings
- SoGE seems to work in the future with grid middleware
- Planning a robust server setup



- Will run SoGE in BIRD and maybe NAF
- Evaluating SoGE as Scheduler for GRID
- ...homogenization and central management of the batch infrastructure
- Future: Cloud integration



Thank You!

Questions?

