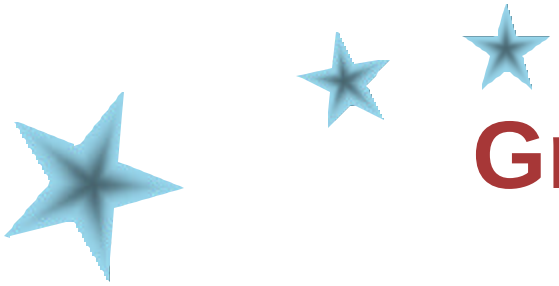


Grid architecture at PHENIX

Job monitoring and related stuff in multi cluster environment

Andrey Shevel@mail.chem.sunysb.edu

Tech talk
20th June 2003

Plan

- General PHENIX grid scheme
- Available Grid components
- Conceptions and scenario for multi cluster environment
- Job submission and job monitoring
- Live demonstration

Andrey Shevel@mail.chem.sunysb.edu

General scheme: jobs are planned to go where data are and to less loaded clusters

Partial Data Replica

SUNY RAM

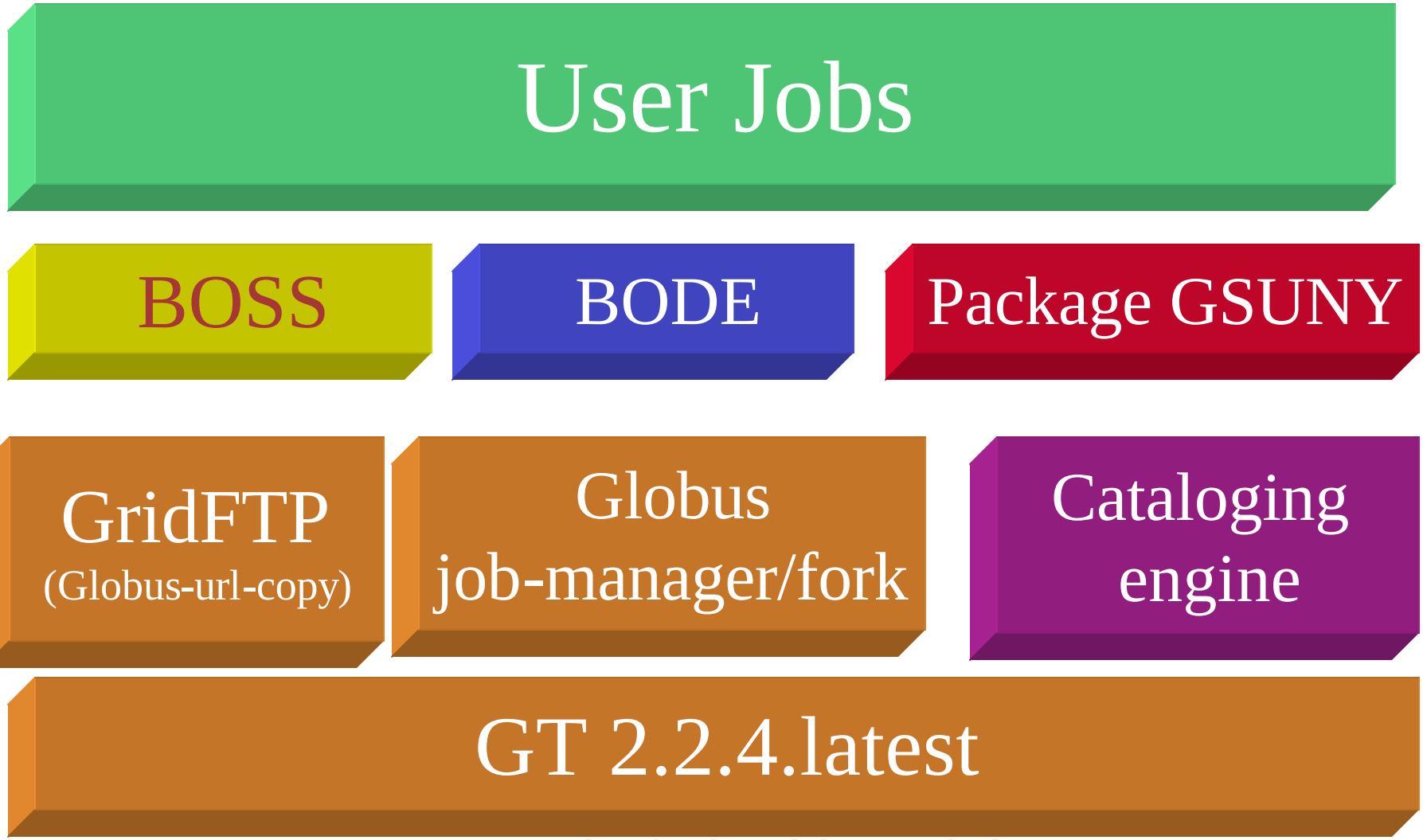
File Catalog

Main Data Repository

RCF

Andrey Shevel@mail.chem.sunysb.edu

Base subsystems for PHENIX Grid



Andrey Shevel@mail.chem.sunysb.edu

Conceptions

Major Data Sets
(physics or simulated data)

Minor Data Sets
(Parameters, scripts, etc.)

Master Job (script)
submitted by user

Satellite Job (script)
Submitted by Master Job

Input/Output Sandbox(es)

Andrey Shevel@mail.chem.sunysb.edu

The job submission scenario at remote Grid cluster

- To determine (to know) qualified computing cluster: available disk space, installed software, etc.
- To copy/replicate the major data sets to remote cluster.
- To copy the *minor data sets* (scripts, parameters, etc.) to remote cluster.
- To start the master job (script) which will submit many jobs with default batch system.
- To watch the jobs with monitoring system – BOSS/BODE.
- To copy the result data from remote cluster to target destination (desktop or RCF).

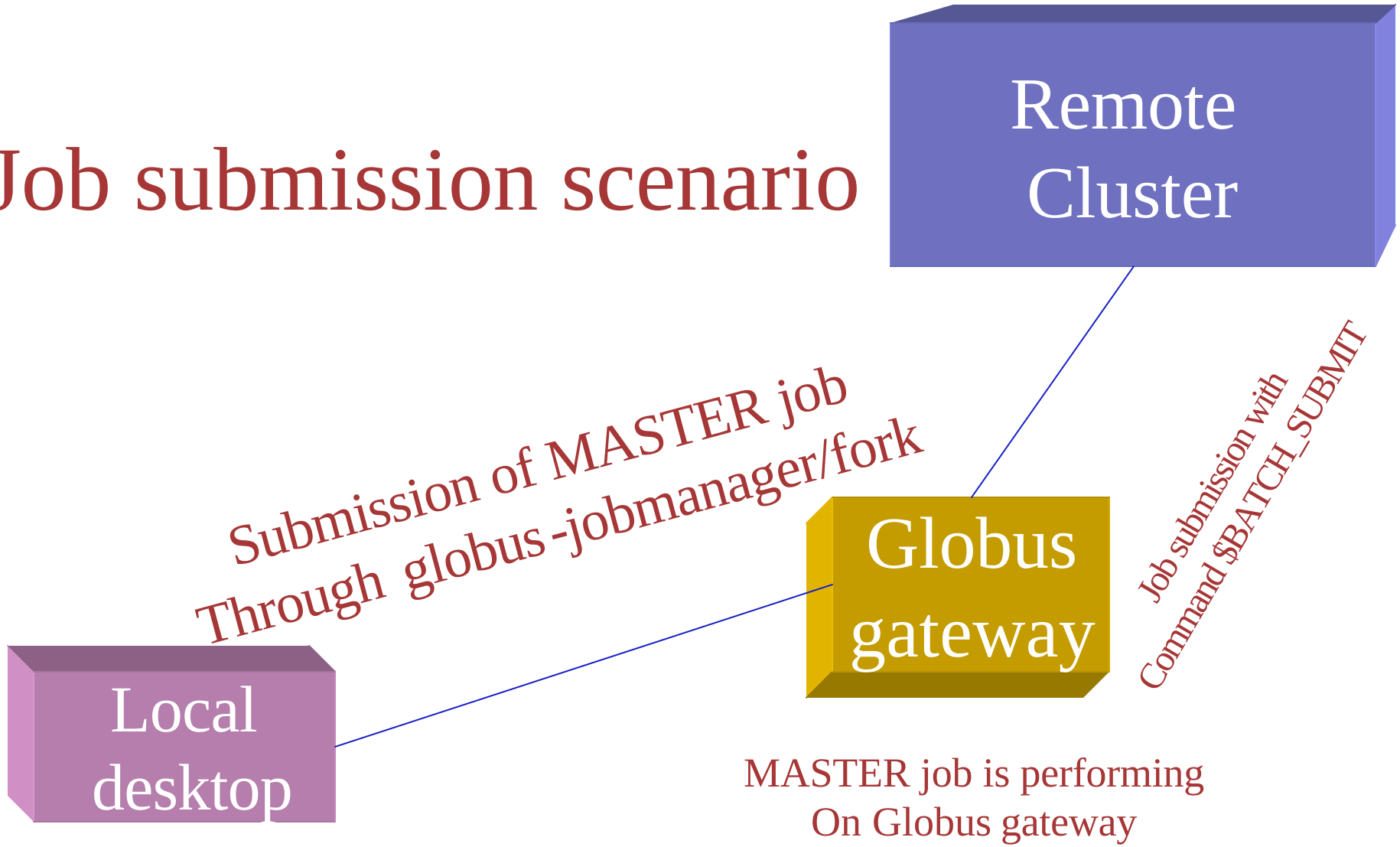
Andrey Shevel@mail.chem.sunysb.edu

Master job-script

- The master script is submitted from your desktop and performed on the Globus gateway (may be in *group account*) with using monitoring tool (it is assumed BOSS).
- It is supposed that the master script will find the following information in the environment variables:
  - CLUSTER_NAME – name of the cluster;
  - BATCH_SYSTEM – name of the batch system;
  - **BATCH_SUBMIT – command for job submission through BATCH_SYSTEM**.

Andrey Shevel@mail.chem.sunysb.edu

Job submission scenario



Andrey Shevel@mail.chem.sunysb.edu

Transfer the major data sets

- There are a number of methods to transfer major data sets:
  - The utility *bbftp* (whithout use of GSI) can be used to transfer the data between clusters;
  - The utility *gcopy* (with use of GSI) can be used to copy the data from one cluster to another one.
  - Any third party data transfer facilities (e.g. HRM/SRM).

Andrey Shevel@mail.chem.sunysb.edu

Copy the minor data sets

- There are at least two alternative methods to copy the minor data sets (scripts, parameters, constants, etc.):
  - To copy the data to /afs/rhic.bnl.gov/phenix/users/*user_account/*…
  - To copy the data with the utility *CopyMinorData* (part of package gsuny).

Andrey Shevel@mail.chem.sunysb.edu

Package gsuny
List of scripts

- General commands
  (**ftp://ram3.chem.sunysb.edu/pub/suny-gt-2/gsuny.tar.gz)**
  - **GPARAM – configuration description for set of remote clusters;**
  - **gsub** – to submit the job on less loaded cluster;
  - **gsub-data** – to submit the job where data are;
  - **gstat** – to get status of the job;
  - **gget** – to get the standard output;
  - **ghisj** – to show job history (which job was submitted, when and where);
  - **gping** – to test availability of the Globus gateways.

Andrey Shevel@mail.chem.sunysb.edu

Package gsuny
List of scripts (continued)

- **GlobusUserAccountCheck** – to check the Globus configuration for local user account.
- **gdemo** – to see the load of remote clusters.
- **gcopy** – to copy the data from one cluster (local hosts) to another one.
- **CopyMinorData** – to copy minor data sets from cluster (local host) to cluster.

Andrey Shevel@mail.chem.sunysb.edu

Job monitoring

- After the initial development of the description of required monitoring tool (https://www.phenix.bnl.gov/phenix/WWW/p/draft/shevel/TechMeeting4Aug2003/jobsub.pdf) it was found the packages:
  - **B**atch **O**bject **S**ubmission **S**ystem (**BOSS**) by Claudio Grandi http://www.bo.infn.it/cms/computing/BOSS/
  - Web interface BOSS DATABASE EXPLORER (BODE) by Alexei Filine http://filine.home.cern.ch/filine/

Andrey Shevel@mail.chem.sunysb.edu

Basic BOSS components

`boss` **executable:**

the BOSS interface to the user

MySQL database:

where BOSS stores job information

`jobExecutor` **executable:**

the BOSS wrapper around the user job

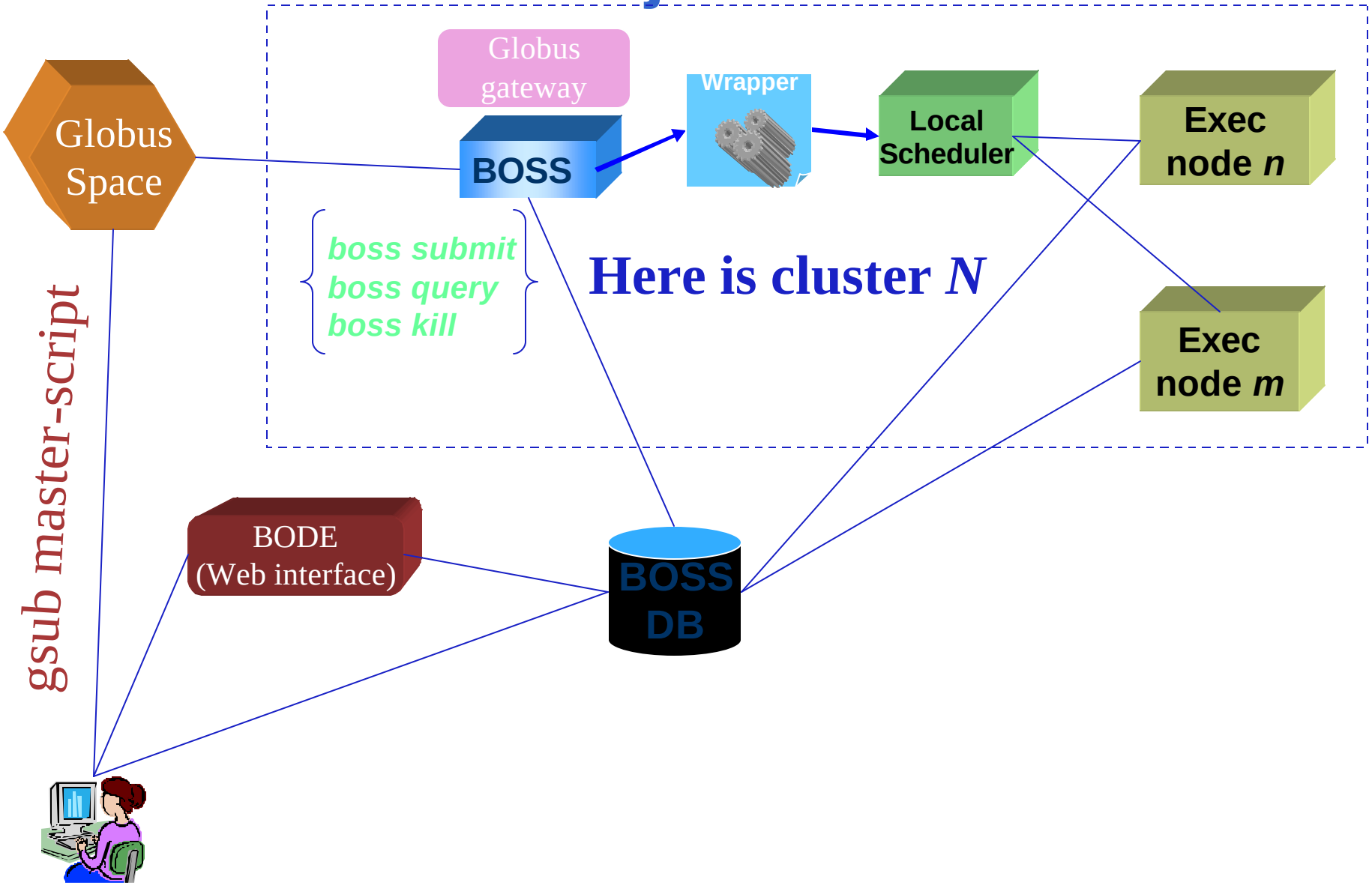
`dbUpdator` **executable:**

the process that writes to the database while the job is running

Interface to Local scheduler

Andrey Shevel@mail.chem.sunysb.edu

Basic job flow



Andrey Shevel@mail.chem.sunysb.edu

```
[shevel@ram3 shevel]$ CopyMinorData local:andrey.shevel unm:.
++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
    YOU are copying THE minor DATA sets
                --FROM--                                    --TO--
Gateway   =      'localhost'                                  'loslobos.alliance.unm.edu'
Directory =      '/home/shevel/andrey.shevel'         '/users/shevel/.'

Transfer of the file '/tmp/andrey.shevel.tgz5558' was succeeded
```

```
[shevel@ram3 shevel]$ cat TbossSuny
. /etc/profile
. ~/.bashrc
echo "                          This is master JOB"
printenv
boss submit -jobtype ram3master -executable ~/andrey.shevel/TestRemoteJobs.pl -stdout   \
~/andrey.shevel/master.out -stderr ~/andrey.shevel/master.err
```

gsub TbossSuny # submit to less loaded cluster

Andrey Shevel@mail.chem.sunysb.edu

Status of the PHENIX Grid

- Live info is available on the page
  http://ram3.chem.sunysb.edu/~shevel/phenix-grid.html
- The group account '*phenix*' is available now at
  - SUNYSB (rserver1.i2net.sunysb.edu)
  - UNM (loslobos.alliance.unm.edu)
  - IN2P3 (in process now**)**

Andrey Shevel@mail.chem.sunysb.edu

Organization	Grid gateway	Contact person	Status
BNL PHENIX (RCF)	**phenixgrid01.rcf.bnl.gov** **GT 2.2.4; LSF**	Dantong Yu	**tested**
SUNYSB (RAM)	**rserver1.i2net.sunysb.edu** **GT 2.2.3; PBS**	Andrey Shevel	**tested**
New Mexico	**loslobos.alliance.unm.edu** **GT 2.2.4; PBS**	Tim Thomas	**No PHENIX software.**
IN2P3 (France)	**ccgridli03.in2p3.fr** **GT 2.2.3; BQS**	Albert Romana	**tested**
Vanderbilt	**Grid gateway is not yet available for testing**	Indrani Ojha	**Not tested**

Andrey Shevel@mail.chem.sunysb.edu

Live Demo for BOSS Job monitoring

http://ram3.chem.sunysb.edu/~magda/BODE

User: guest
Pass: Guest101

Computing Meeting
13 October 2003

Andrey Shevel@mail.chem.sunysb.edu