# Job monitoring tool for PHENIX

Andrey Shevel
*State University of New York at Stony Brook, NY 11794, USA*

This document is intended for user or manager who needs to keep track of the jobs running in multi cluster computing environment. The same technology could be applied to keep track of any important commands which are performed on different computers (home desktop, laptop, computing clusters, etc.). Actually the configuration described here is prepared for the collaboration PHENIX. However almost all components (with slight change) could be used anywhere.

## 1. Foreword

In this document we describe the procedure how to install and use PHENIX job monitoring system for job monitoring needs in multi cluster environment. This document was prepared on the base of previous materials: [3] and [4].

## 2. Introduction to user multi-cluster environment

In multi-site job submission, a lager number of job monitoring and control functions will be made available. For example a user will be able to:

- query which of his jobs are currently running and where they are running (on which site or cluster);
- query whether or not there is output from his job (directory file, cluster);
- keep valuable information about the job on a database.

It is envisaged that rather detailed information about each job in the multi-cluster environment [including its log file] will be stored in a database. A web-based interface provides primary interaction to this database as well as control panels to query various jobs.

## 3. Job submission model

In order to do job monitoring we need to follow several conventions about the job submission. We distinguish several types of computers where in general user can be logged:

- user desktop - the computer where user submits his jobs;
- cluster gateway - the computer through which jobs are submitted to the associated cluster.

To permit the job monitoring in PHENIX we have to create PHENIX job monitoring environment:

- to have the database server to keep track of monitored jobs;
- to install required software and configuration files inside the account[1].

The start up package for different clusters are available at the directory `ftp://ram3.chem.sunysb.edu/pub/Monitoring/`. Which file has to be used for which cluster is shown in the table I on the page 2.

---

[1]Here we assume that the software is deployed in user account. Also every user has the separate database on the database server to keep the information about his jobs

**Revision: Thu Jun 17 17:06:03 EDT 2004**

Table I The gateways and names of the clusters

| Organization | Gateway | Cluster name | Monitoring package |
|---|---|---|---|
| BNL | phenixgrid01.rcf.bnl.gov | RCF | `FTPDIRECTORY/RCF_PHENIX_BOSS.tgz`[a] |
| SUNYSB | rserver1.i2net.sunysb.edu | RAM | `FTPDIRECTORY/BOSS_SUNY_PHENIX.tgz` |
| IN2P3 | ccali.in2p3.fr | CCF | `FTPDIRECTORY/CCF_BOSS_PHENIX.tgz` |
| UNM | loslobos.alliance.unm.edu | UNM | `FTPDIRECTORY/BOSS_UNM_PHENIX.tgz` |
| VU | vmplogin.scc.center.vanderbilt.edu | Vampire | `FTPDIRECTORY/BOSS_VAMPIRE_PHENIX.tgz` |
| PNPI | pcfarmf.pnpi.spb.ru | PNPI | `FTPDIRECTORY/PNPI_BOSS_PHENIX.tgz` |

[a] FTPDIRECTORY=`ftp://ram3.chem.sunysb.edu/pub/Monitoring`

## 4. Installation

1. Please log onto the required computing facility where you plan to use Monitoring Tool (laptop, desktop, cluster). It is assumed that your computing facility is running under Linux.
2. Please do the command

   `wget ftp://ram3.chem.sunysb.edu/pub/Monitoring/BOSS_VAMPIRE_PHENIX.tgz`

   to download the distributing binary version for Monitoring Tool.
3. `tar zxvf BOSS_VAMPIRE_PHENIX.tgz`
4. `cd BOSS_VAMPIRE_PHENIX` here you could see something like below

```
[shevelay@vmps04 ~/BOSS_VAMPIRE_PHENIX]$ ls -lt
total 24
-rwx------    1 shevelay rhic         8843 Jun  9 17:12 VAMPIRE_Batch_Conf.bash
-rwx------    1 shevelay rhic         1209 Jun  9 17:04 vampire-BossAddPBS
-rw-------    1 shevelay rhic          344 May  4 18:07 CATALOG_INFO
-rwx------    1 shevelay rhic          278 May  4 17:41 vampire-BossDelPBS
```

5. `./VAMPIRE_Batch_Conf.bash`

   You will see something like below

```
[shevelay@vmps04 ~/BOSS_VAMPIRE_PHENIX]$ ./VAMPIRE_Batch_Conf.bash

./VAMPIRE_Batch_Conf.bash:
+-------------------------------------------------------------------------------------+
|       The script is dedicated to check everything and to setup the user account     |
|       for the use of the system BOSS. |

|        This version of the script is configured for 'Vanderbilt University'|

| |
                Hostname          = 'vmps04.vampire'
                User account name = 'shevelay'
                Local Date/Time   = 'Wed Jun  9 17:47:14 CDT 2004'
| |
+-------------------------------------------------------------------------------------+


We have to set following environment variables:
```

```
Cluster nikname            - GRIDVM_CLUSTER_NAME = 'VAMPIRE'
Cluster batch system       - GRIDVM_BATCH_SYSTEM = 'PBS'
Command for job submission - GRIDVM_BATCH_SUBMIT = 'boss submit -scheduler vampire-pbs '

----- You have no directory ''.
      ----- You can start the installation of the BOSS now
Please answer: [no]/yes   yes
```

Please answer 'yes'.

6. After your answer 'yes' the gzipped tarball will be copied from anonymous ftp server ram3.chem.sunysb.edu and performed default configuration. To see what is exactly will be done one might look at the script. Also in this step two files will be corrected: `~/.bashrc` and `~/.cshrc`. The following lines will be added for file `~/.bashrc`

```
#BOSS45730092003    ! Do not deletethis line !
#-----------------------------------------------------------------------+
#            This block ot lines is dedicated to BOSS           |
#    It is prepared by the script './VAMPIRE_Batch_Conf.bash'.
#    This file is dedicated to be used as "~/.bashrc".  |
#    Date = Wed Jun  9 18:19:06 CDT 2004
#     Initial version Tue Sep 30 17:46:03 EDT 2003 by A. Shevel@bnl.gov  |
#-----------------------------------------------------------------------+
#


source /home/shevelay/PHENIX_JOB_MONITORING/bossenv.sh         # Please check this file
export GRIDVM_CLUSTER_NAME="VAMPIRE"
export GRIDVM_BATCH_SYSTEM="PBS"
export GRIDVM_BATCH_SUBMIT="boss submit -scheduler vampire-pbs "
#End addition BOSS45730092003
```

For the file `~/.cshrc` it will look like

```
#BOSS45730092003    ! Do not deletethis line !
#-----------------------------------------------------------------------+
#            This block ot lines is dedicated to BOSS and Globus env     |
#    This file is dedicated to be used as "~/.cshrc"  |
#    It is prepared by the script './VAMPIRE_Batch_Conf.bash'.
#    Date = Wed Jun  9 18:28:35 CDT 2004
#     Initial version Tue Sep 30 17:46:03 EDT 2003 by A. Shevel@bnl.gov  |
#-----------------------------------------------------------------------+
#

   source /home/shevelay/PHENIX_JOB_MONITORING/bossenv.csh         # Please check this fil
   setenv GRIDVM_CLUSTER_NAME VAMPIRE
   setenv GRIDVM_BATCH_SYSTEM PBS
   setenv GRIDVM_BATCH_SUBMIT "boss submit -scheduler vampire-pbs "
#End addition BOSS45730092003
```

At this point it is better to do 'logout' and after that to login back to the same cluster. You have to see now several environment variables:

- `GRIDVM_CLUSTER_NAME`

**Revision: Thu Jun 17 17:06:03 EDT 2004**

- `GRIDVM_BATCH_SYSTEM`
- `GRIDVM_BATCH_SUBMIT`

It is easy to see it by the command line

```
printenv | grep GRIDVM
```

If you do not see the mentioned above environment variables please make sure that appropriate profile is taken into account (performed).

The set of values for all environment variables on all clusters are available at the table II on the page 6. The current set of available clusters is shown in the table I on the page 2.

7. Test the monitoring system
   If all previous steps were performed successfully (no error messages) you could test the monitoring system.

   (a) please do
   ```
   boss
   ```
   you have to see
   ```
   Usage:
   boss command [command options]
   command:
   version [v]          : print boss version string
   registerScheduler    : register a new scheduler
   registerJob          : register a new type of job
   showSchedulers       : list registered schedulers
   showJobTypes         : list registered jobtypes
   deleteScheduler      : delete a scheduler
   deleteJobType        : delete a jobtype
   declare              : declare a job without submitting
   submit  [s]          : submit a job
   query   [q]          : get info about one or more jobs
   SQL                  : SQL query
   kill    [k]          : kill a job
   recoverJob           : update job info on the DB from journal
                          or from job STDOUT/ERR
   delete  [d]          : delete job entries from the DB
   purge   [p]          : purge old job ebtries from the DB
   ```

   (b) In this step you have to declare so called **BOSS scheduler**s. It could be done with scripts (we continue the line with our above example with cluster VAMPIRE):
   - `~/BOSS_VAMPIRE_PHENIXvampire-BossAddFork`
   - `~/BOSS_VAMPIRE_PHENIX/vampire-BossAddPBS`

   (c) please download and perform the simple test
   ```
   wget ftp://ram3.chem.sunysb.edu/pub/Monitoring/TestBoss.tgz
   ```
   and do following commands
   ```
   tar zxvf TestBoss.tgz
   cd TestBoss
   tcsh StartJob
   ```
   You have to see the messages like below
   ```
   Job ID 151
   Write log to /Users/shevel/BOSS_COMMAND_STORE/T13281.log
   Scheduler ID is 612630.ram0
   ```
   It does mean that PHENIX monitoring system assigned test job number 151. The same job has local scheduler (PBS) identification '612630.ram0'.
   The whole process can be performed on different cluster

**Revision: Thu Jun 17 17:06:03 EDT 2004**

```
rcas2071->wget ftp://ram3.chem.sunysb.edu/pub/Monitoring/TestBoss.tgz
--15:15:18--  ftp://ram3.chem.sunysb.edu/pub/Monitoring/TestBoss.tgz
           => 'TestBoss.tgz'
Connecting to 192.168.1.4:3128... connected.
Proxy request sent, awaiting response... 200 OK
Length: 1,694 [application/x-tar]

100%[===============================================================================>

15:15:18 (1.62 MB/s) - 'TestBoss.tgz' saved [1694/1694]

rcas2071->tar zxvf TestBoss.tgz
TestBoss/
TestBoss/ST
TestBoss/StartJob
TestBoss/README
TestBoss/BossArchive_141.tgz
rcas2071->cd TestBoss
rcas2071->tcsh StartJob
Job ID 152
Write log to /phenix/u/shevel/BOSS_COMMAND_STORE/T20774.log
Scheduler ID is 720468
```

(d) Testing the monitoring one command
You could adapt and test the script `$BOSSDIR/G_Command` for you needs to keep trace of the critical commands. Below you see only example

```
G_Command cp /etc/ssh/ssh_config ~/.ssh/.
```

for above command you will see the output like below

```
[shevel@ram3 BossPhenix]$ G_Command cp /etc/ssh/ssh_config ~/.ssh/

/home/shevel/bin/G_Command:

   The script is intended to perform any command through BOSS.
   In other words to store all information about the command to
   the database.

Job ID 153
Write log to /home/shevel/BOSS_COMMAND_STORE/Command.log.20446
Scheduler ID is jobExecutor::153::20458
```

If everything was performed fine you have working client part for PHENIX job monitoring.

## 5. Final results

Finally a manager has to add you account name to the Web panel BODE [2]. After that you can see your database records.

## 6. Troubleshuting

### 6.1. Different account names on different clusters

It may happen that on different clusters you have different account name. In this situation you have to determine which is you preferable account name and use it everywhere in PHENIX job monitoring system.

In particular you have to see the generated configuration file `$BOSSDIR/BossConfig.clad` and check you account names inside. All of them have to be the same as your preferable account name.

**Revision: Thu Jun 17 17:06:03 EDT 2004**

Table II The environment variable values on the clusters

| Gateway | GRIDVM_CLUSTER_NAME | GRIDVM_BATCH_SYSTEM | GRIDVM_BATCH_SUBMIT |
|---|---|---|---|
| phenixgrid01.rcf.bnl.gov | PHENIX | LSF | boss submit -scheduler rcf-lsf |
| rserver1.i2net.sunysb.edu | SUNY_NCG | PBS | boss submit -scheduler local-suny-pbs |
| ccali.in2p3.fr | CCF | BQS | boss submit -scheduler ccf-bqs |
| loslobos.alliance.unm.edu | UNM | PBS | |
| vmplogin.scc.center.vanderbilt.edu | VAMPIRE | PBS | boss submit -scheduler vampire-pbs |
| pcfarmf.pnpi.spb.ru | PNPI | SGE | boss submit -scheduler pnpi-sge |

If you need more deep understnding for underlying subsystems please see docs on appropriate subsystem: `$BOSSDIR/doc/BOSS-UserGuide-3.3.3.pdf` and [2].

## 7. Using the BOSS for job submission

As we saw before the simple interactive command (or script) can be submitted with `G_Command`. If you need to submit batch job with BOSS you need to use line:

```
$GRIDVM_BATCH_SUBMIT -executable user_script -stdout user_output -stdin user_input -stderr u
```

In general not all the parameters are mandatory, for example, you may skip -stdin if do not need for that.

After job submission the information about such the event will be put into the database on the server ram3.chem.sunysb.edu (in this version of the monitoring system).

Access to the stored information might be done with CLI interface (see command **boss**). You can use alternative way to see what is in the BOSS database by using the Web panel `http://ram3.chem.sunysb.edu/BODE`.

## References


[1] BOSS (Batch Object Submission System) provides an easy to use book keeping system for jobs running on a Linux computing farm. Different job types can be registered to the BOSS System, allowing the storage on a database of information specific to the task which is being performed by the job itself. This information is used both for job monitoring and for book-keeping.
`http://www.bo.infn.it/cms/computing/BOSS/`

[2] Web interface for BOSS - source codes (CERN CMS case) `http://filine.home.cern.ch/filine/bode/bode-2.1.1.tgz`
Web interface for BOSS - documentaion (CERN CMS case)
`http://filine.home.cern.ch/filine/bode/doc/Contents.html`
the server at SUNYSB is available at `http://ram3.chem.sunysb.edu/BODE`

[3] The proposal for job moniutoring at PHENIX
`ftp://ram3.chem.sunysb.edu/pub/Monitoring/jobsub.pdf`

[4] Joined STAR/PHENIX meeting
`ftp://ram3.chem.sunysb.edu/pub/Monitoring/monitor-star-3.ppt`


**Revision: Thu Jun 17 17:06:03 EDT 2004**