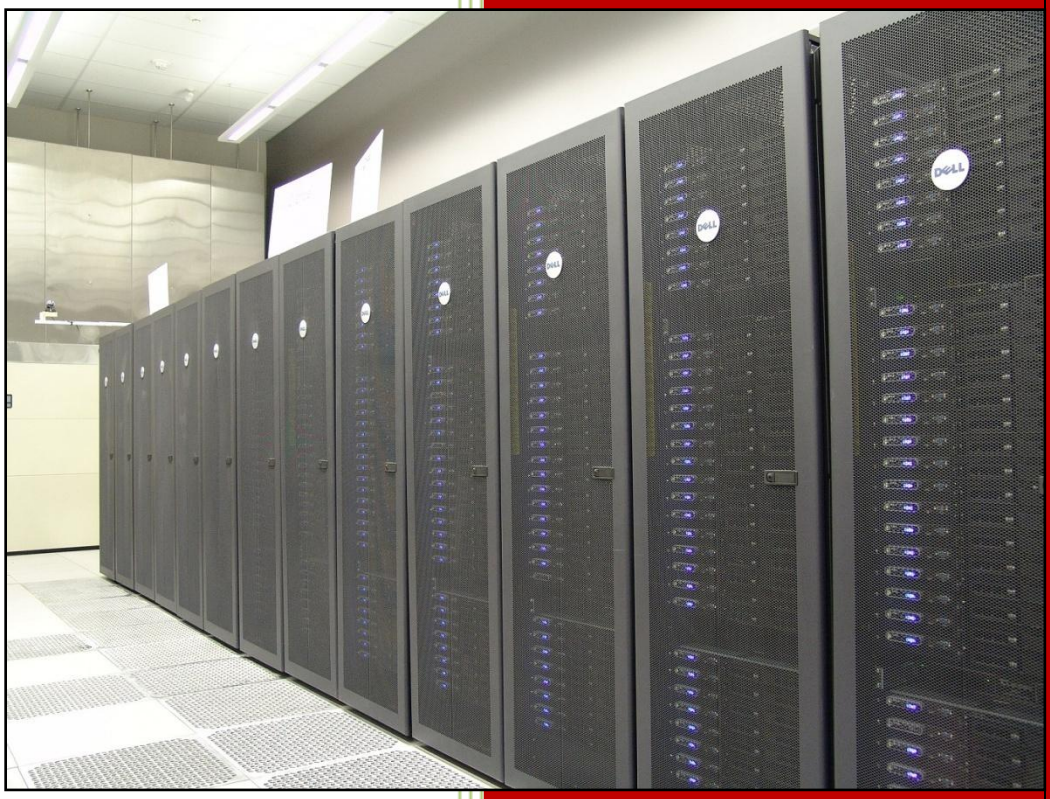
HPCC - Hrothgar

Getting Started User Guide – SURAgrid



Texas Tech University

Table of Contents

User Guide

Submitting R jobs to SURAgrid

1. Requesting Personal Certificates

To join SURAgrid and run jobs on its resource sites, you need a personal certificate as your grid identity. If you don't have a valid personal certificate, please follow the instruction below to make a request.

You are **STRONGLY** encouraged to use ***Firefox for both Windows or Macintosh*** as your web browser. In order to get your personal OSG certificate, you need to complete the following steps:

1. Point your web browser to the URL https://oim.grid.iu.edu/oim/certificaterequestuser.
2. Enter your contact information in the Contact Information field.
3. Enter your profile information in the Profile Information field.
4. Enter a password to be used for issuing your certificate and encrypting your private key. (**IMPORTANT: If you forget this password, you will not be able to issue your certificate and import it your browser after it is approved.**)
5. Select SURAGrid from the pick list in the Sponsor field.
6. Check the "I AGREE" box.
7. Click on the Submit button.

After you have submitted your request for an OSG certificate, your sponsor will receive an email from the OSG Certificate Registration Authority (run by the OSG Grid Operations Center or GOC) asking them to validate your request. After your request is approved, you will receive an email which contains a link to your certificate and private key. You need to download the file that contains your user certificate and key from the link to your local computer.(**IMPORTANT NOTE:** You must use the **SAME** browser on the **SAME** computer that you used to request the certificate when you import the certificate and private key.)

2. Importing Certificates/Key pair to your Web Browser

When you get your personal certificate, please install it in your web browser, it will function as an electronic credential that can be used by OSG secure websites. When a user access the websites, a valid certificate is required.

Firefox for Windows

It is recommended that you export your OSG certificate and private key as a PKCS#12 file. To export these items, follow the steps below:

1. Click on the **Tools** option at the top of the browser.
2. Select **Options** from the list.
3. Click on the **Advanced** tab.
4. Click on the **Encryption** tab.
5. Click on the **View Certificates** button.
6. Click on the **Your Certificates** tab.
7. Click on the Import button.
8. Select the certificate from the directory where you saved it (the download location).
9. Click on the Open button.

Then you should see a message of "Successfully imported your security certificate and private key".

Other Web Browsers

To find the details for importing your user certificate to your web browser, please see the instructions through the following links.

1. Importing User Certificate on Firefox
   https://confluence.grid.iu.edu/display/CENTRAL/Importing+User+Certificate+on+Firefox
2. Importing User Certificate on IE
   https://confluence.grid.iu.edu/display/CENTRAL/Importing+User+Certificate+on+IE
3. Importing User Certificate on Chrome
   https://confluence.grid.iu.edu/display/CENTRAL/Importing+User+Certificate+on+Chrome
4. Importing User Certificate on Safari
   https://confluence.grid.iu.edu/display/CENTRAL/Importing+User+Certificate+on+Safari
5. Importing User Certificate for Command Line Use
   https://confluence.grid.iu.edu/display/CENTRAL/Importing+User+Certificate+for+Command+Line+Use

3. Exporting Your Certificates/Key pair for use by Globus

In order to use your OSG certificate and private key on grid resources or submit machines, please copy your **file_name.p12** file to the ***$HOME/.globus*** directory (if you don't have a

.globus directory, you need to create it) on that machine, change its name to **usercred.p12** and set its permissions as follows.

```
# mkdir .globus
# chmod 700 .globus
# mv $HOME/.globus/file_name.p12 $HOME/.globus/usercred.p12
# chmod 400 $HOME/.globus/usercred.p12
```

4. Registering DN with SURAgrid VOMS

After your have received your personal grid certificate, you need to register your certificate's Distinguished Name (DN) with the SURAgrid Virtual Organization Membership Service (VOMS).

The following steps need to be followed to carry out this registration:

1. Using the same web browser into which you imported your grid certificate, point it to the SURAgrid VOMS server https://voms.hpcc.ttu.edu:8443/voms/suragrid.
2. Fill in the following fields:
    - First Name: Given Name
    - Last Name: Family Name
    - Your Institution: Institution
    - Phone: Phone number
    - Address: Address
    - Email Address: Email address
    - Read the policy and check the box of "I confirm I have read and agree with the terms expressed in the VO Acceptable Usage Policy document displayed above".
    - Click on the Submit button.

You will receive an automated email from the VOMS server within minutes.

- Click the link in the email to confirm your membership request.

Then, the VO administrator will receive an email for approval. After your request is approved, you will receive an email indicating that your membership request for VO SURAgrid has been approved.

If you have fussy email filters, you could prep them to accept the automated messages. The From field on both is:

```
* From: SURAgrid-VO-Admin@ttu.edu
```

The first email has the subject:

```
* Subject: [VOMS Admin] Your membership request for VO suragrid
```

The second email has the subject:

```
* Subject: [VOMS Admin] Your vo membership request for VO suragrid has
been approved
```

5. Preparing for Job Submissions

Log on to Hugin or Munin with your eRraider account,

```
 # ssh hugin.hpcc.ttu.edu
or
 # ssh munin.hpcc.ttu.edu
```

For how to logging on to a linux server using SSH on your Windows machine, please refer to the second section in the user guide *http://www.hpcc.ttu.edu/downloads/docs/HPCC-Hrothgar-Connecting%20to%20the%20Server%20-Windows.pdf*.

In your /home directory, run commands "mkdir .globus" and "chmod 700 .globus",

```
 # mkdir .globus
 # chmod 700 .globus
```

If you have your certificate in .p12 format, save it in the .globus directory as "usercred.p12", and run command "chmod 400 .globus/usercred.p12",

```
 # chmod 400 .globus/usercred.p12
```

If you don't have your certificate in .p12 format, but you have it in .pem format, please save the cert and key pair (usercert.pem and userkey.pem) in your .globus directory, and run command " openssl pkcs12 -export -inkey .globus/userkey.pem -in .globus/usercert.pem -out usercred.p12 -name "usercred" " to convert your .pem format certificate to .p12 format.


```
 # openssl pkcs12 -export -inkey .globus/userkey.pem -in
.globus/usercert.pem -out usercred.p12 -name "usercred"
```

6. Submitting simple test jobs to SURAgrid

The following two steps are the necessary steps for submitting jobs to a SURAgrid resource site.

Setup the grid environment,

```
 # . /usr/local/grid/setup.sh
or
 # source /usr/local/grid/setup.sh
```

Create a temporary proxy to be used for your job submissions, the proxy contains your grid identity,

```
# voms-proxy-init -voms suragrid

Enter GRID pass phrase for this identity:
Your identity: /DC=com/DC=DigiCert-Grid/O=Open Science
Grid/OU=People/CN=Amy Wang 8
```

```
Creating temporary proxy
................................................ Done
Contacting  voms.hpcc.ttu.edu:15003 [/DC=com/DC=DigiCert-Grid/O=Open
Science Grid/OU=Services/CN=voms/voms.hpcc.ttu.edu] "suragrid" Done
Creating proxy ........................................... Done
Your proxy is valid until Mon Apr 29 22:50:33 2013
```

IMPORTANT: If you have created a pass phrase for your certificate, when you generate a proxy, the system will ask you to enter the pass phrase.

Submit a simple test job to the grid resources,

```
 # globus-job-run antaeus.hpcc.ttu.edu/jobmanager-sge /bin/date
Mon Apr 29 11:00:58 CDT 2013

 # globus-job-run calclab-ce.math.tamu.edu/jobmanager-pbs /bin/hostname
bloc609-01

 # globus-job-run hurr.tamu.edu/jobmanager-pbs /bin/hostname
c0231

 # globus-job-run ce.grid.unesp.br/jobmanager-fork /usr/bin/id
uid=20079(suragrid) gid=9999(osg) groups=9999(osg)

 # globus-job-run gk04.swt2.uta.edu/jobmanager-fork /usr/bin/id
uid=601(suragrid) gid=502(grid) groups=502(grid)

 # globus-job-run fermigridosg1.fnal.gov/jobmanager-condor /bin/hostname
fnpc5006.fnal.gov

 # globus-job-run gk01.atlas-swt2.org/jobmanager-condor /bin/hostname
gk01.atlas-swt2.org
```


7. Submitting R jobs to SURAgrid

Before you start to submit jobs, please make sure you have followed the first two steps shown in the last section.

7.1. Submitting through Globus

First, you need to create a folder in the ***/state/partition1*** directory for the file transfers, and you need to add the full access permission to allow your job to write output files into the folder. Also, you need to save your R input file in the folder.

To create the folder, run

mkdir /state/partition1/mytest

To add the full access permission, run
chmod 777 /state/partition1/mytest

To copy the input file, under the directory that contains the file, run
cp rtest.R /state/partition1/mytest/

Then you need to create a job submission script that looks like the below. This job submission script is named "rtest.sh", in this example, "rtest.R" is the input file, you need to replace it with the name of your input file. The input file will be copied to the target resource site when the job submitted. "$OSG_SITE_NAME.rtest.output" is the named output file which will be generated by the job, you can name it with a different name.

```
#!/bin/bash
##rtest.sh
echo "Job started on $OSG_SITE_NAME node `hostname` in dir `pwd` at
`date`"

##Set up grid environment
source $GLOBUS_LOCATION/etc/globus-user-env.sh
if [ -f $OSG_APP/suragrid/etc/profile ]; then
. $OSG_APP/suragrid/etc/profile
else
echo "WARNING: cannot find OSG_APP/suragrid/etc/profile" 1>&2
fi

## Copy the input file from your local directory to the target site
globus-url-copy gsiftp://munin.hpcc.ttu.edu/state/partition1/mytest/rtest.R
file:///`pwd`/

##Start R with the input file "rtest.R", save the result to the output file
R --vanilla < rtest.R > `pwd`/rtest.output

## Copy the output file from the target site to your local directory
globus-url-copy file:///`pwd`/rtest.output
gsiftp://munin.hpcc.ttu.edu/state/partition1/mytest/$OSG_SITE_NAME.rtest.
output

## Clean up the saved and generated files on the target site
/bin/rm rtest.R
/bin/rm rtest.sh
/bin/rm rtest.output

echo "Job finished at `date`"
```

When you submit your job through Globus, both the input file and the job submission script need to be copied to the target resource site. The step to copy the input file can be included in the job submission script and to be done by the job, but, you HAVE TO manually copy the job submission script to the target resource site, and you need to add the execute access

permission to make it an executable, the copy through gsiftp doesn't carry the execute permission. To do so, you need to run

globus-url-copy file://`pwd`/rtest.sh gsiftp://ce.grid.unesp.br/~/rtest.sh

and

globus-job-run ce.grid.unesp.br/jobmanager-fork /bin/chmod +x rtest.sh

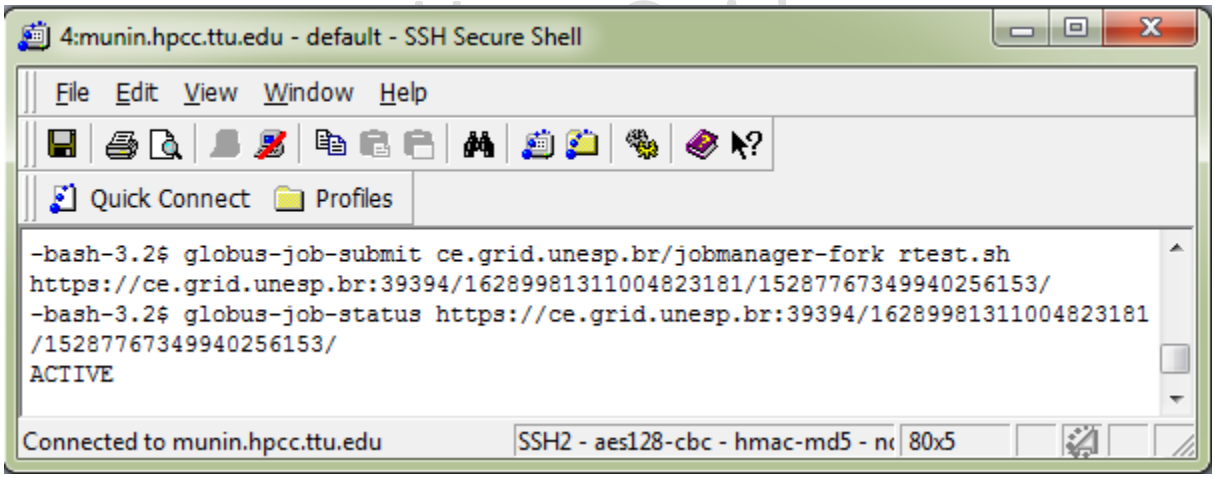
In these two commands above, "rtest.sh" is the job submission script for the job, you need to change it with the name of your job submission script. In this example, the hostname of the target resource site to run your job is "ce.grid.unesp.br", you need to change it if you want to submit your job to another site.

Next, run

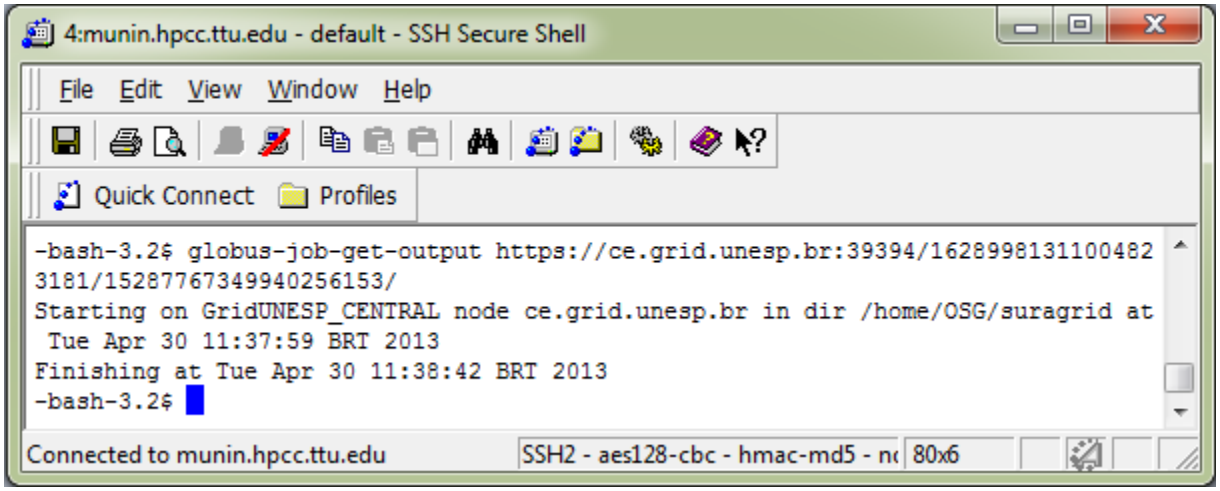
globus-job-submit ce.grid.unesp.br/jobmanager-fork rtest.sh

to submit your job to the target resource site. The command will return a URL that you'll need to use for checking the job status and output.

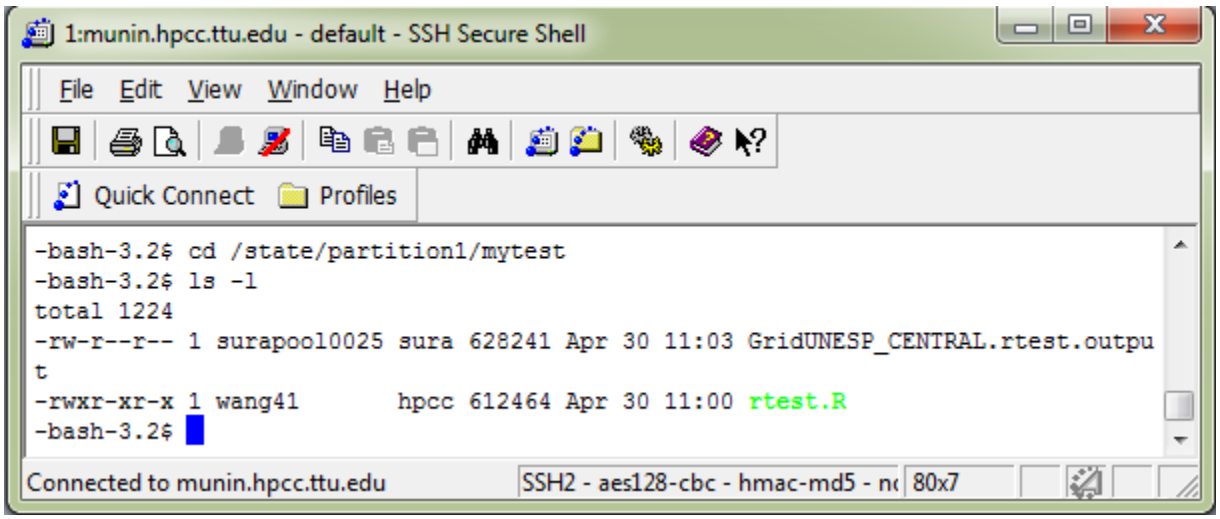
To check the job status, run the command ***globus-job-status*** following with the URL.



If the job status is shown as "ACTIVE", it means the job is still running, when the status changes to "DONE", the job is finished. Then you may run ***globus-job-get-output*** following with the URL to check the output, but please remember this output is NOT the result of running your R program, it's only the output for your job submission. The program result is saved in the folder ***/state/partition1/mytest*** as "$OSG_SITE_NAME.rtest.output".



In this example, the variable *$OSG_SITE_NAME* returned the site name "GridUNESP_CENTRAL", the output file is saved as GridUNESP_CENTRAL.rtest.output in folder ***/state/partition1/mytest***.



7.2. Submitting through Condor-g

When you submit a job through Condor-g, unlike submitting through Globus, you don't have to copy your job submission scripts and input files to the target resource sites, however, you have to submit the job from the directory that contains the job submission scripts and the input files.

For the GridFTP file transfer, you still need to create a folder in the ***/state/partition1*** directory and add the full access permission to allow your job to write output files in the folder.

To create the folder, run

mkdir /state/partition1/mytest

To add the full access permission, run

chmod 777 /state/partition1/mytest

Then you need to create a job submission script that looks like the below. This job submission script is named "rtest.sh", in this example, "rtest.R" is the input file, you need to replace it with the name of your input file. The input file will be copied to the target resource site when the job submitted. "$OSG_SITE_NAME.rtest.output" is the named output file which will be generated by the job, you can name it with a different name

rtest.sh:

```
#!/bin/bash
##rtest.sh
echo "Job started on $OSG_SITE_NAME node `hostname` in dir `pwd` at
`date`"

##Set up grid environment
source $GLOBUS_LOCATION/etc/globus-user-env.sh
if [ -f $OSG_APP/suragrid/etc/profile ]; then
. $OSG_APP/suragrid/etc/profile
else
echo "WARNING: cannot find OSG_APP/suragrid/etc/profile" 1>&2
fi

##Start R with the input file "rtest.R", save the result to the output file
R --vanilla < rtest.R > `pwd`/rtest.output

## Copy the output file from the target site to your local directory
globus-url-copy file://`pwd`/rtest.output
gsiftp://munin.hpcc.ttu.edu/state/partition1/mytest/$OSG_SITE_NAME.rtest.
output
```

```
## Clean up the saved and generated files on the target site
/bin/rm rtest.R
/bin/rm rtest.sh
/bin/rm rtest.output

echo "Job finished at `date`"
```

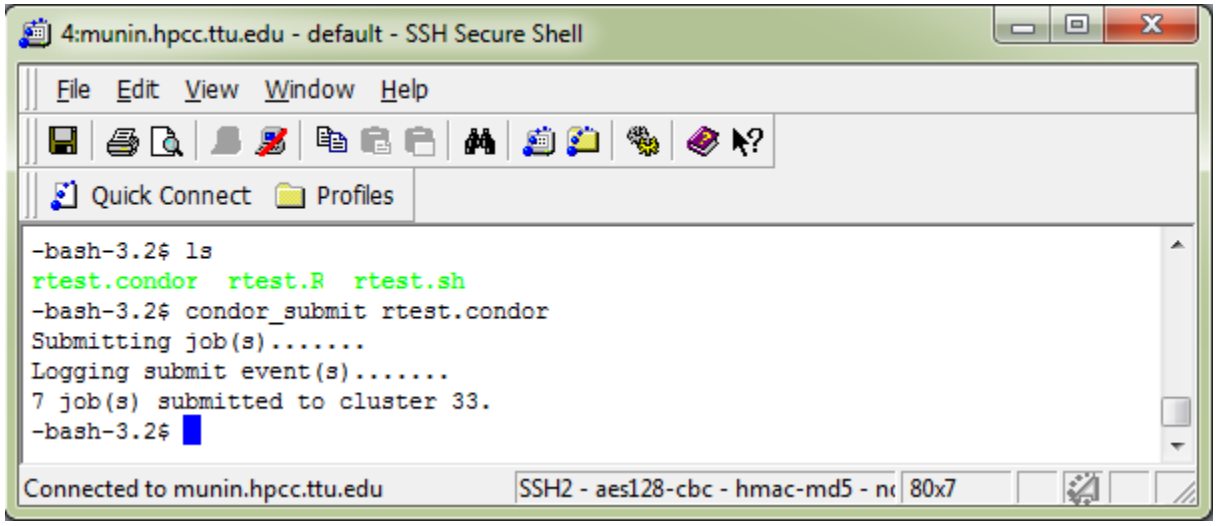
You also need to create another job submission script for submitting the job through Condor-g. The example script is shown below. This script is named "rtest.condor", the input file for this job is "rtest.R", you need to replace it with the name of your input file, and if the input file is not saved in your job submission directory, you need to specify the path to the input file. In this example script, "rtest.sh" is specified as the value of the variable "Executable", which means, this script "rtest.condor" will call "rtest.sh" and run it on the target resource sites when job is submitted.

rtest.condor:

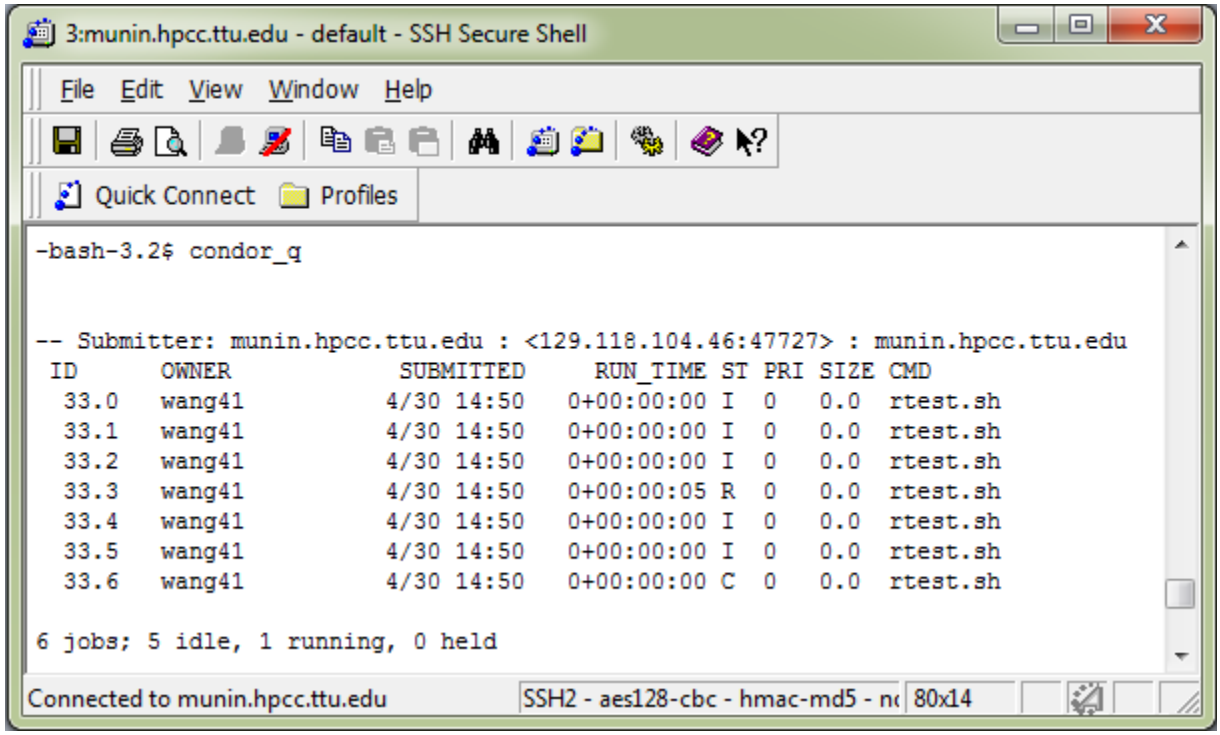
```
Universe                = grid
Notification            = never
Executable              = rtest.sh
Transfer_Executable     = True
Transfer_Input_Files    = rtest.R
Should_Transfer_Files   = YES
When_To_Transfer_Output = ON_EXIT
Globus_RSL              = (maxWalltime=60)

grid_resource           = gt2 antaeus.hpcc.ttu.edu/jobmanager-sge
Output                  = rtest-$(Cluster)-$(Process)-Antaeus.out
Error                   = rtest-$(Cluster)-$(Process)-Antaeus.err
Log                     = rtest-$(Cluster)-Antaeus.log
queue 1

grid_resource           = gt2 calclab-ce.math.tamu.edu/jobmanager-fork
Output                  = rtest-$(Cluster)-$(Process)-Calclab.out
Error                   = rtest-$(Cluster)-$(Process)-Calclab.err
Log                     = rtest-$(Cluster)-Calclab.log
queue 1
```

```
grid_resource               = gt2 hurr.tamu.edu/jobmanager-fork
Output                      = rtest-$(Cluster)-$(Process)-Brazos.out
Error                       = rtest-$(Cluster)-$(Process)-Brazos.err
Log                         = rtest-$(Cluster)-Brazos.log
queue 1

grid_resource               = gt2 ce.grid.unesp.br/jobmanager-fork
output                      = rtest-$(Cluster)-$(Process)-unesp.out
error                       = rtest-$(Cluster)-$(Process)-unesp.err
log                         = rtest-$(Cluster)-unesp.log
queue 1

grid_resource               = gt2 fermigridosg1.fnal.gov/jobmanager-fork
output                      = rtest-$(Cluster)-$(Process)-FNAL.out
error                       = rtest-$(Cluster)-$(Process)-FNAL.err
log                         = rtest-$(Cluster)-FNAL.log
queue 1

grid_resource               = gt2 gk04.swt2.uta.edu/jobmanager-fork
output                      = rtest-$(Cluster)-$(Process)-uta.out
error                       = rtest-$(Cluster)-$(Process)-uta.err
log                         = rtest-$(Cluster)-uta.log
queue 1

grid_resource               = gt2 gk01.atlas-swt2.org/jobmanager-fork
output                      = rtest-$(Cluster)-$(Process)-atlas.out
error                       = rtest-$(Cluster)-$(Process)-atlas.err
log                         = rtest-$(Cluster)-atlas.log
queue 1
```

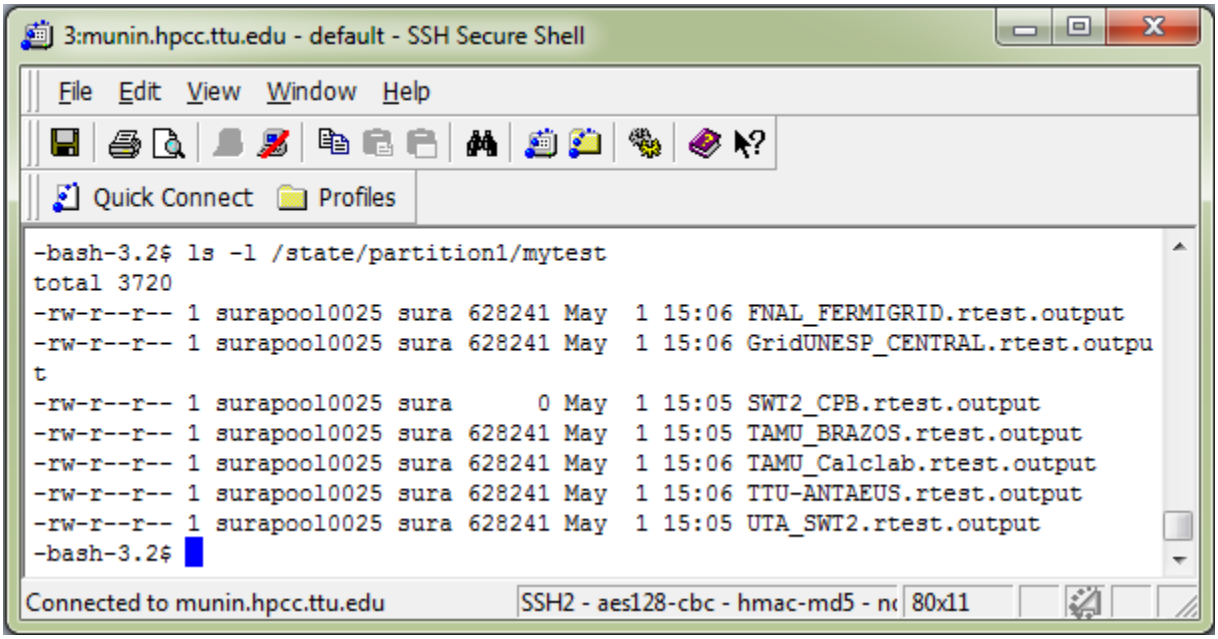
To submit the job to SURAgrid resource sites, run the command ***condor_submit rtest.condor***, replace "rtest.condor" with the name of your job submission script.



To check the job status, run the command ***condor_q***. The status "R" indicates the job is running, and the status "C" indicates the job finished and it's cleaning up.



As the snapshot shown below, the output files are saved in the folder ***/state/partition1/mytest***.



User Guide

Last updated: 05/01/2013

For Additional Assistance Contact: hpccsupport@ttu.edu

For Comments/Suggestions on user guide hpcc@ttu.edu