

Information about our Clusters

francis.tru.ca

marie.tru.ca

To connect to any of these systems port 2200 must be used.

francis.tru.ca

This system has ten (10) nodes (servers) with eight processors each (each node is a dual quadcore). A single Gaussian03 (g03) job could use up to eight (8) processors (although this number of processors might not be the best choice – see the Performance Comparison section on page 2).

Linux OS: A flavour of RedHat 64 bit

Master node: Quadcore Xeon 2.83 GHz, 2 GB RAM, 2 x 500 GB HB

Server type A (1 server): Dual quadcore Xeon 2.83 GHz, 16 GB RAM, 2 TB HB (queue 0)

Server type B (3 servers): Dual quadcore Xeon 2.83 GHz, 8 GB RAM, 1 TB HB (queues 1 to 3)

Server type C (6 servers): Dual quadcore Xeon 2.83 GHz, 4 GB RAM, 500 GB HB (queues 4 to 9)

There is a queue for every node. The full name for queue 0 is: all.q@compute-0-0.local

The full name for queue *i* is: all.q@compute-0-*i*.local

Web interface for monitoring francis: <https://francis.tru.ca>

Any updates about the cluster system will be posted here. In addition, you can see what the cluster is currently doing, and there are help docs here too.

To go to node *i* type “ssh compute-0-*i*” and to return to the previous place type “exit”.

You can access your home directory is the usual way from any node.

To access your scratch space in any node, log into it and type “cd /scr/<username>”

“/scr” is a simplified (and equivalent) notation for “/state/partition1”

To know the available scratch space in a given node log into it and type “df -h”.

The output of this command will also tell you about your available home space.

The scratch space in node 0 is 1.8 TB, in nodes 1 to 3 is 904 GB and in nodes 4 to 9, as well as in the master node, is 446 GB.

marie.tru.ca

This system is a 4-processor **PQS QuantumCube**. It has two nodes with two processors each. A single Gaussian03 (g03) job could use up to two (2) processors.

Linux OS: Mandrake 9.2

4 - 2.8 GHz Xeon processors on dual motherboards,

RAM: 4 GB per node (8 GB total); HD: 2 x 120 GB;

The change of passwords in marie must be done twice, first on the master node (n1) and then on the slave node (n2).

To go from n1 to n2 type “rlogin n2” and to return to n1 type “exit”.

Performance Comparison

DFT test job for a system with 11 heavy atoms (MPWB1K/6-311G(d,p), 258 basis functions)

4 consecutive jobs: a partial optimization followed by a frequency calculation with the 6-31G basis set, and the full optimization of a TS and frequency calculation.

Number of processors*	Job CPU time
1	8 hours 4' (4h 9' + 50' + 1h 59' + 1h 6')
2	4 hours 15' (2h 14' + 26' + 1h 2' + 33')
4	2 hours 34' (1h 11' + 30' + 36' + 17')
4^b	2 hours 51' (1h 24' + 17' + 45' + 25')
6	1 hours 50' (54' + 11' + 32' + 13')
8	1 hours 38' (51' + 9' + 28' + 10')

* No other jobs are running on the same node when the test job was running

^b The other four processors of the node are being used by another job

These jobs took 14 hours (6h 52' + 1h 18' + 3h 13' + 2h 36') in marie with 1 processor.

It makes sense to use up to **4 processors** for this type of calculations. 6 processors could also be used (although the time gained is not that significant) but it does not make sense to use 8 processors.

Useful commands when running jobs to run in any of the clusters above:

qstat -f Will show you all the queues in the system

e.g.,

```
[nmora@francis ~]$ qstat -f
queueName          qtype used/tot. load_avg arch      states
-----
all.q@compute-0-0.local  BIP  0/8    0.00  lx26-amd64
-----
all.q@compute-0-1.local  BIP  0/8    0.00  lx26-amd64
-----
all.q@compute-0-2.local  BIP  0/8    0.00  lx26-amd64
-----
all.q@compute-0-3.local  BIP  0/8    0.00  lx26-amd64
-----
all.q@compute-0-4.local  BIP  0/8    0.00  lx26-amd64
-----
all.q@compute-0-5.local  BIP  0/8    0.00  lx26-amd64
-----
all.q@compute-0-6.local  BIP  0/8    0.00  lx26-amd64
-----
all.q@compute-0-7.local  BIP  2/8    2.03  lx26-amd64
  62  0.60500 ethene.com nmora    r  03/27/2008 15:56:02  2
-----
all.q@compute-0-8.local  BIP  1/8    1.00  lx26-amd64
  53  0.50500 etheneO-ts nmora    r  03/27/2008 09:00:25  1
-----
all.q@compute-0-9.local  BIP  0/8    0.00  lx26-amd64
```

```
[nmora@n1 nmora]$ qstat -f
queueName          qtype used/tot. load_avg arch      states
-----
n1.q               BIP  0/2    0.00  glinux
-----
n1p0.q             BIP  0/1    0.00  glinux
-----
n1p1.q             BIP  1/1    0.00  glinux
  6666 -100 test-aceto nmora    r  03/27/2008 23:07:23 MASTER
-----
n2.q               BIP  0/2    0.01  glinux
-----
n2p0.q             BIP  0/1    0.01  glinux
-----
n2p1.q             BIP  0/1    0.01  glinux
-----
```

In the first example (francis.tru.ca) there are two jobs running, one on queue 7 (all.q@compute-0-8.local) that uses 2 of the 8 available processors and has jobid=62, and another one on queue 8 using one processor, with jobid=53. In the second example (marie.tru.ca) there is only one job running using a single processor, with jobid=6666.

qdel <jobid> Removes the job with jobid= <jobid> from the queue where it is running or from the waiting-to-run stage.

e.g., From the first example above:

```
[nmora@francis ~]$ qdel 53 62
```

This will stop the execution (or delete from the queue) of the jobs with ID 53 and 62.

To run Gaussian03 jobs:

Use the rung03 script.

To obtain a full description of the script rung03 in any of the clusters just type “rung03”. The information that follows will be displayed.

****** HELP TEXT FOR RUNG03 ******

This is the script used in francis.tru.ca to run Gaussian'03 jobs.

Correct Syntax: `rung03 [options] <filename>`

Basic Usage: To run a Gaussian job type "rung03" and the filename of the Gaussian input file excluding the ".com" extension. For example if the input file name is "o2.com" then type "rung03 o2".

Note that the g03 input file must have extension ".com"

Output Files: This script names Gaussian output files by taking the Gaussian input file base name then adding the SGE job id and the ".log" extension. For example "o2.com" becomes "o2_555.log". Output files are created in the same directory as the input file.

Resources: The number of processors and the amount of memory to use should be specified in the Gaussian input file, this script will read those values from the input file. If nothing is specified then one processor and the default amount of memory will be used.

Up to eight (8) processors can be used for a single job even though this might not be the most effective way of using the resources available.

The number of processors to be used must be specified in the input file using the link 0 keyword "%nproc=#".

Memory should be specified by using the link 0 keyword "%mem=#".

Run Files: The run script, SGE standard and error outputs, Gaussian scratch files, read-write files etc. are all stored in /scr/USERNAME/RUNDIR where USERNAME is the name of the user submitting the job and RUNDIR has the same name as the output file name excluding the ".log" extension. If the Gaussian job finishes normally this directory is automatically deleted.

The folder "/home" (of about 446 GB) can be accessed from any node, but the scratch folders cannot. There is a scratch folder (/scr) per node.

If the check point (chk) files are stored in a scratch folder,

(using "%chk=/scr/USERNAME/<filename>" make sure that

jobs that use the same chk file run in the same node where the file is.

Please clean your scratch folders periodically.

When a Gaussian job terminates with an error or it is manually cancelled, its scratch folder is not automatically deleted and this must be done by hand to avoid using all the space available in disk.

Hence periodically delete the files you do not need to keep and compress (gzip) the ones you need to keep. This is also applicable to the files you keep in your home directory: /home/USERNAME.

To inspect and clean your scratch folder in node *i*, log into by typing “ssh compute-0-*i*” (to return to the previous place type “exit”), then type “cd /scr/<username>”.

A numbering code is used before the name of the scratch folder of a given Gaussian job. The prefix "00" indicates a job that is running or has been manually cancelled, and the prefix "11" indicates a Gaussian job with an error termination.

Queues: When a Gaussian job is submitted it will be run in an available queue, if no queue is available it will wait until one is free. There are ten queues named: all.q@compute-0-*i*.local, with *i* going from 0 to 9, with eight processors in each. Each queue can have up to eight jobs running at a time.

Options: The following options have been implemented:

-q <name> Specifies that the job must be run in a specific queue where <name> is the name of the queue to run in. If the queue is occupied the job will wait until it is free.

e.g., `rung03 -q all.q@compute-0-0.local <inputfile>`

This line will submit the execution of the g03 input file <inputfile.com> to the queue *all.q@compute-0-0.local* (node 0) which identifies the node with the largest amount of memory and hard disk.

For jobs to run on nodes 0 and 1, the queue name must be explicitly indicated.

To know the queue names type “qstat -f”.

-M Requests that an email be sent to <username>@tru.ca when the job is done.

-m <address> Requests that an e-mail be sent to the <address> indicated when the job is done.

e.g., `rung03 -M <inputfile>`

e.g., `rung03 -m frank@rhubarb.com <inputfile>`

When the job finishes an e-mail like the one that follows is received:

```
Job 97 (test.1707011.g03) Complete
User      = nmora
Queue     = all.q@compute-0-4.local
Host      = compute-0-4.local
Start Time = 03/29/2008 17:07:03
End Time  = 03/29/2008 18:18:41
User Time = 01:03:24
System Time = 00:07:37
Wallclock Time = 01:11:38
CPU       = 01:24:28
Max vmem  = 1.138G
Exit Status = 1
```

****** HELP TEXT FOR RUNG03 ******

This is the script used in marie.tru.ca to run Gaussian'03 jobs.

Correct Syntax: `rung03 [options] <filename>`

Basic Usage: To run a Gaussian job type "rung03" and the filename of the Gaussian input file excluding the ".com" extension. For example if the input file name is "o2.com" then type "rung03 o2".

Note that the g03 input file must have extension ".com".

Output Files: This script names Gaussian output files by taking the Gaussian input file base name then adding the SGE job id and the ".log" extension. For example "o2.com" becomes "o2_555.log". Output files are created in the same directory as the input file.

Resources: The number of processors and the amount of memory to use should be specified in the Gaussian input file, this script will read those values from the input file. If nothing is specified then one processor and the default amount of memory will be used. Only one or two processors can be used for each job in this system, this can be specified in the input file using the link 0 keyword "%nproc=#". Memory should be specified by using the link 0 keyword "%mem=#mb" to allocate "#" megabytes of memory up to 1700 MB on node 1 and up to 1900 MB on node 2.

Run Files: The run script, SGE standard and error outputs, Gaussian scratch files, read-write files etc. are all stored in /scr/USERNAME/RUNDIR where USERNAME is the name of the user submitting the job and RUNDIR has the same name as the output file name excluding the ".log" extension. If the Gaussian job finishes normally this directory is automatically deleted.

There is a scratch folder (/scr) per node with a capacity of around 200 GB each. Please clean them periodically. When a Gaussian job terminates with an error or it is manually cancelled, its scratch folder is not automatically deleted and this must be done by hand to avoid using all the space available in disk. Hence periodically delete the files you do not need to keep and compress (gzip) the ones you need to keep. This is also applicable to the files you keep in your home directory: /home/USERNAME.

The folder "/home" (of about 9 GB) is shared between the two nodes, but the scratch folders are not. If the checkpoint (chk) files are stored in either scratch folder, make sure that jobs that use the same chk file run in the same node where the file is.

A numbering code is used before the name of the scratch folder of a given

Gaussian job. The prefix "00" indicates a job that is running or has been manually cancelled, and the prefix "11" indicates a Gaussian job with an error termination.

To inspect and clean the scratch folder in the slave node (n2) type "rlogin n2" and to return to the master node (n1) type "exit".

Queues: When a Gaussian job is submitted it will be run in an available queue, if no queue is available it will wait until one is free. There are four queues named n1p0.q, n1p1.q, n2p0.q, n2p1.q, each representing one of four processors, two on each of the two nodes. Each queue should only have one process running in it at a time.

Options: The following options have been implemented:

-q <name> Specifies that the job must be run in a specific queue where <name> is the name of the queue to run in. If the queue is occupied the job will wait until it is free.

e.g., `rung03 -q n1p1.q <filename>`

This line will submit the execution of the g03 input file *filename.com* to the queue n1p1.q
To know the queue names just type: `qstat -f`

-p Jobs are normally assigned a low priority of -100, a job can be given a higher priority of 0 by adding the "-p" option to the `rung03` command. Higher priority jobs will be run before lower priority jobs. Each user may only have one higher priority job running or in the queue at any time.