CLEMSON®

COMPUTING AND
INFORMATION TECHNOLOGY

Introduction to Palmetto

Galen Collier
Computational Scientist

Cyberinfrastructure Technology Integration (CITI)
Clemson University

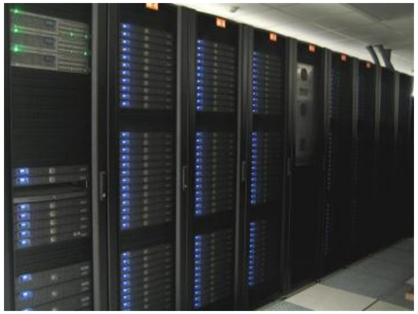
Jan. 17, 2012

The Palmetto Cluster

- Both Shared and Distributed configs
- Operates at over 96 TF/s
- #128 on November 2011 “Top 500” list
- #2 among public academic institutions
- 1,616 compute nodes (14,168 cores)
- Operating System: Scientific Linux 6
- Myrinet 10G network interconnect
- Data Storage:

  115 TB “scratch” filesystem

  > 72 TB of purchased backed-up storage



Node Specifications (1 of 2)

	Name	Count	Model	Processor	L2 Cache	Cores	Memory	Local Disk
compute node phase 1	node0001-0257	257	Dell PE 1950	Intel Xeon E5345 @2.33GHz x 2	4 MB	8	12 GB	80 GB (SATA)
compute node phase 2	node0258-0515	258	Dell PE 1950	Intel Xeon E5410 @2.33GHz x 2	6 MB	8	12 GB	80 GB (SATA)
compute node phase 3	node0516-0771	256	Sun X2200 M2 x64	AMD Opteron 2356 @ 2.3GHz x 2	4 MB	8	16 GB	250 GB (SATA)
compute node phase 4	node0772-1023,1108-1111	256	IBM dx340	Intel Xeon E5410 @2.33GHz x 2	6MB	8	16 GB	160 GB (SATA)
compute node phase 4.1	node1024-1107	84	IBM dx340	Intel Xeon E5410 @2.33GHz x 2	6MB	8	16 GB	160 GB (SATA)
compute node (former CCMS nodes)	node1112-1541	430	Sun X6250	Intel Xeon L5420 @2.5GHz x 2	6MB	8	32 GB	160 GB (SATA)
compute node phase 6	nodes 1553-1622	70	HP DL 165 G7	AMD Opteron 6172 @2.1GHz x 2	12MB	24	48 GB	250 GB (SATA)

Node Specifications (2 of 2)

	Name	Count	Model	Processor	L2 Cache	Cores	Memory	Local Disk
regular large shared memory systems	nodelm01-nodelm04	4	HP DL 580 G7	Intel Xeon 7542 @ 2.66 GHz x 4	18MB	24	512 GB	146 GB (SAS)
math sciences large memory	nodemath	1	HP DL 980 G7	Intel Xeon 7560 @ 2.66 GHz x 8		64	2 TB	

/scratch

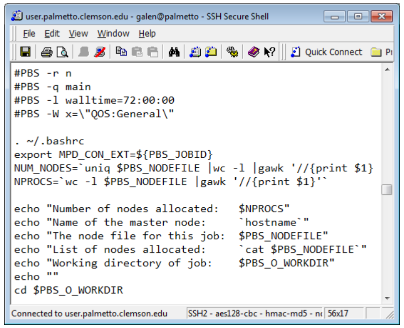
- Temporary “work” directory for all jobs
  (Move all files in, run jobs, move all files out)
- Open to all users, create a directory for yourself
- Specialized high-performance hardware
- Storage server software = OrangeFS
- Designed to handle very high I/O activity
- 115 TB of space, open to all users

Note: Try to keep the number of files per directory to less than 1,000

CLEMSON
COMPUTING AND
INFORMATION TECHNOLOGY

Accessing Palmetto

Command-line interface, any Secure Shell (ssh) client

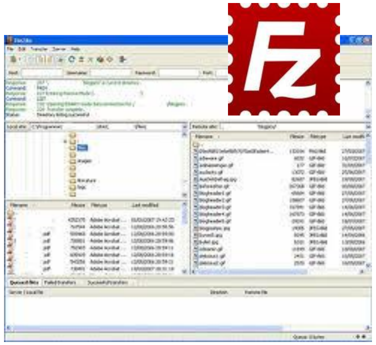

user.palmetto.clemson.edu - galen@palmetto - SSH Secure Shell
File Edit View Window Help
Quick Connect
Pı

```
#PBS -r n
#PBS -q main
#PBS -l walltime=72:00:00
#PBS -W x=\"QOS:General\"

. ~/.bashrc
export MPD_CON_EXT=${PBS_JOBID}
NUM_NODES=`uniq $PBS_NODEFILE |wc -l |gawk '//{print $1}
NPROCS=`wc -l $PBS_NODEFILE |gawk '//{print $1}'`

echo "Number of nodes allocated:   $NPROCS"
echo "Name of the master node:     `hostname`"
echo "The node file for this job:  $PBS_NODEFILE"
echo "List of nodes allocated:     `cat $PBS_NODEFILE`"
echo "Working directory of job:    $PBS_O_WORKDIR"
echo ""
cd $PBS_O_WORKDIR
```

Connected to user.palmetto.clemson.edu
SSH2 - aes128-cbc - hmac-md5 - nc
56x17



Transfer files to/from using FileZilla (or scp)

MOTD (/etc/motd)

```
--------------------------------------------------------------------------
        Welcome to the PALMETTO CLUSTER at CLEMSON UNIVERSITY

  * Please email ithelp@clemson.edu with questions or to report problems.
  * HPC webpage is http://citi.clemson.edu/hpc

  * The FIRST TUESDAY OF EACH MONTH, from 9:00am to 12:00noon,
    is reserved for system wide cluster maintenance.

  DO NOT RUN JOBS ON THE USER LOGIN NODE.  THEY WILL BE
  TERMINATED WITHOUT NOTICE.  NO EXCEPTIONS.

  Useful commands:
    checkquota                - get your current disk quota
    module avail              - list software packages
    qstat -xf jobid           - check status of jobid
    qstat -Qf queuename       - check status of queuename
    pbstop -m 24              - check status of jobs across the cluster
    cat /etc/hardware-table   - list hardware types: ram,cores,chip,etc.

  Palmetto User Guide:   http://desktop2petascale.org/resources/159

  Workaround for PBS and -k option:
  If you use "-k oe" or "-k e" or "-k o", you must set permissions
  on your /home dir to 711:    chmod 711 /home/userid
  otherwise your output will be blank at the job's end.

---------------------------- /etc/motd ----- Last Updated: 10-Jan-2012 ---
```

Passwordless SSH

```
ssh-keygen -t rsa
```

```
Generating public/private rsa key pair.
Enter file in which to save the key (/home/userid/.ssh/id_rsa):   [Enter]
Enter passphrase (empty for no passphrase):   [Enter]
Enter same passphrase again:   [Enter]
Your identification has been saved in /home/userid/.ssh/id_rsa.
Your public key has been saved in /home/userid/.ssh/id_rsa.pub.
The key fingerprint is: 64:72:2a:7b:20:fa:a7:0c:91:26:a6:43:85:0b:1c:21
```

```
cd .ssh
```

```
cp id_rsa.pub authorized_keys
```

(you can also test your configuration: `qsub -I`)

Copy Example Files

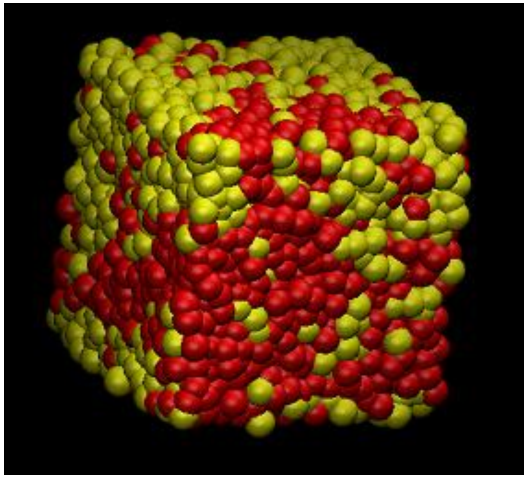
Copy the examples to your home (or scratch) directory:

```
cd ~
cp -r /scratch/galen/intro.palmetto ~
```

```
[galen@user001 ~]$ cd /scratch/galen/intro.palmetto
[galen@user001 intro.palmetto]$ ls
bashrc.example               gethostname.mpi.c
job.gethostname.bash         job.template.bash
examples                     hello_world_mpi.c
job.gethostname.mpi.bash     lammps-30Sep11
gethostname.c                hello_world_openmp.c
job.hello_world.omp.bash     pbsdsh
```

Try an Example Job

- LAMMPS is a molecular dynamics simulation program
- Cooling of a binary mixture
- 5,000 atoms (LJ interactions)
- 50,000 0.005 ps steps

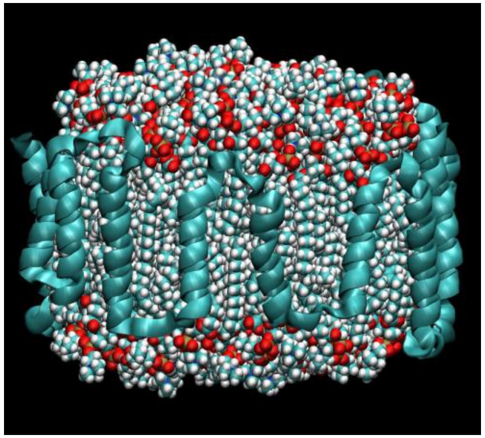


```
cd ~/intro.palmetto/examples/lammps.example
```

Edit job.lammps.bash (optional), then run the job:

```
qsub job.lammps.bash
```

Or try this one…



- NAMD is a molecular dynamics simulation program
- apoa1 system (over 92,000 atoms)
- 12 Å cutoff, PME every 4 steps

```
cd ~/intro.palmetto/examples/namd.example
```

Edit job.namd.bash (optional), then run the job:

```
qsub job.namd.bash
```

Setting Your Environment

- To get a list of available software modules:

```
module avail
```

- To use specific software every time you log-in, place the **`module add`** command in your ~/.bashrc file:

```
module add intel/12.0 mpich2/1.4
```

- List your loaded modules:

```
module list
```

- To clear-out your added modules:

  `module purge` or **`module clear`**

The PBS Queue System

PBS Pro 11.1 is the resource management service used on the Palmetto Cluster

- Enables you to make more efficient use of your time through scripting computational tasks
- PBS takes care of running these tasks and returning the results
- If the cluster is full, PBS holds your tasks and runs them when the resources are available
- PBS ensures fair sharing of cluster resources (policy enforcement)
- PBS ensures optimal/efficient use of available resources

PBS Commands

Command(s)	Description
qsub *job-script*	Submit a batch job
qsub –I *resources*	Submit an interactive job
qstat –u *userid*	Check status of all of your jobs
qhold *jobID*	Put a job on hold (before it starts)
qrls *jobID*	Release a job from hold status
qdel *jobID*	Delete a job, running or not
qselect *criteria*	Select jobs by specified criteria

qstat –xf *[job ID]*

```
Job Id: 817100.pbs01
    Job_Name = ucaccc
    resources_used.cpupercent = 791
    resources_used.cput = 10:50:29
    resources_used.mem = 270832kb
    resources_used.ncpus = 32
    resources_used.vmem = 1170952kb
    resources_used.walltime = 01:21:34
    exec_host = node1026/0*8+node1027/0*8+node1028/0*8+node1029/0*8
    Error_Path = user001.palmetto.clemson.edu:/scratch/galen/amber/ucaccc/ucaccc.stderr
    Output_Path = user001.palmetto.clemson.edu:/scratch/galen/amber/ucaccc/ucaccc.stdout
    Resource_List.mem = 44gb
    Resource_List.mpiprocs = 32
    Resource_List.ncpus = 32
    Resource_List.nodect = 4
    Resource_List.place = free:shared
    Resource_List.select = 4:ncpus=8:mpiprocs=8:mem=11gb
    Resource_List.walltime = 72:00:00
    jobdir = /scratch/galen
    Variable_List = PBS_O_SYSTEM=Linux,PBS_O_SHELL=/bin/bash,
        PBS_O_HOME=/home/galen,PBS_O_HOST=user001.palmetto.clemson.edu,
        PBS_O_LOGNAME=galen,PBS_O_WORKDIR=/scratch/galen/amber/ucaccc,
        PBS_O_LANG=en_US.UTF-8,
        PBS_O_PATH=/scratch/galen/protg/mmtsb.v.Jul-31-2009/perl:/scratch/gale
        n/protg/mmtsb.v.Jul-31-2009/bin:/usr/lib64/qt-3.3/bin:/opt/pbs/default/
        bin:/opt/gold/bin:/opt/condor/bin:/opt/condor/sbin:/usr/local/bin:/bin:
        /usr/bin:/usr/local/sbin:/usr/sbin:/sbin:/opt/mx/bin:/home/galen/bin,
        PBS_O_MAIL=/var/spool/mail/galen,PBS_O_QUEUE=workq
```

Some qsub Options

qsub -N name	Job name
qsub -q workq	Queue to assign job to
qsub -V	Export environment variables
qsub -v var=value	Expand upon environment variables
qsub -l (see below)	Resource list (hardware required)

chip_manufacturer=amd or chip_manufacturer=intel
chip_model=opteron or chip_model=xeon
chip_type=e5345 chip_type=e5410 (l5420, x7542, 2356, 6172, etc.)
node_manufacturer=dell (hp, ibm, sun, dell)

For all available qsub options, see PBS User Guide, p. 75

Example Batch Job Script File

```
#!/bin/bash
#PBS -N jobname
#PBS -l select=4:ncpus=8:mpiprocs=8
#PBS -l mem=11gb
#PBS -l chip_manufacturer=amd
#PBS -l walltime=2:00:00
#PBS -o stdout.txt
#PBS -e stderr.txt
#PBS -q workq
#PBS -M userid@clemson.edu

my.program.exe [arguments for my program]
```

Notification Parameters

```
#PBS -M [e-mail address]
```

e-mail address can be a list of email addresses separated by commas

`#PBS -m bean` or `-m be` or `-m e`

Mail options: send an email when the job **b**egins, **e**nds, is **a**borted, or **n**o notification

Handling Output Files

- The ‘standard output’ and ‘error output’ are sent to files named ***jobname.o987349*** and ***jobname.e987349***
(in your $PBS_O_WORKDIR directory).

- The following parameters can modify this behavior:

`#PBS -e [path/filename]`	error output file
`#PBS -o [path/filename]`	standard o/p
`#PBS -j eo`	merge the error and standard output
`#PBS -k eo`	keep error and standard output separate

Using an Interactive Job

- Useful for debugging applications, short tests, or for computational steering

```
qsub -I

qsub -I -l walltime=2:00:00

qsub -I -l select=1:ncpus=8:mem=11gb

qsub -I -l walltime=2:00:00,select=2:npcus=8

qsub -I -l select=2:ncpus=8:mpiprocs=8:mem=15gb,walltime=2:00:00

qsub -I -q bigmem -l select=1:ncpus=1:mem=64gb,walltime=4:00:00
```

checkqueuecfg

```
Default (workq) routing queue configs


MaxRun/MaxQue           0-2 hrs(quick)  2-24 hrs(short) 24-72 hrs(long)

==========================================================================

       1 nodes(single)  450/unlimited   450/unlimited   450/unlimited

    2-16 nodes(tiny)    30/unlimited    30/unlimited    30/unlimited

   17-64 nodes(small)   15/unlimited    15/unlimited    15/unlimited

  65-255 nodes(medium)  4/unlimited     4/unlimited     4/unlimited

256-1622 nodes(large)   1/unlimited     1/unlimited     1/unlimited
```

Killing Jobs

One, or a few jobs:

```
qdel [jobID] [jobID] [jobID] ...
```

Kill all of your jobs:

```
qselect -u $USER | xargs qdel
```

Kill all of your queued jobs:

```
qselect -u $USER -s Q | xargs qdel
```

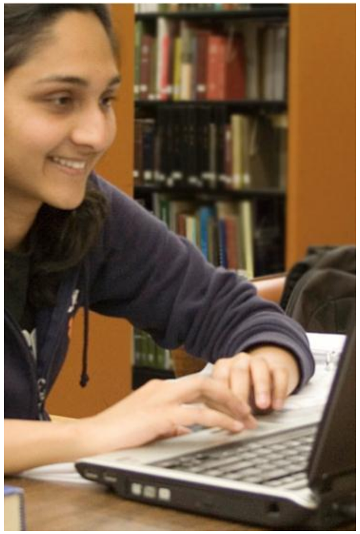
Kill all of your running jobs:

```
qselect -u $USER -s R | xargs qdel
```

CLEMSON
COMPUTING AND
INFORMATION TECHNOLOGY

Palmetto User Support

- “Help” requests or technical questions can be submitted to the Palmetto Admin Staff and anyone who can help will respond.



- Computational scientists are available for consultation and training.
- We can help with compiling code and installing programs.
- We can also help with developing proposals that make use of Palmetto resources.

Final Notes

Palmetto Cluster User’s Guide

PBS Professional 11.1 User Guide

- Do not run jobs on the head/login node (user001)
- Work in /scratch, then move all files when finished.
- When you log-in, read the MOTD
- Need help?

  E-mail ithelp@clemson.edu
  with the word “Palmetto” in the subject line…
  job ID or copy of your PBS job script helps a lot