

A Introduction to The Center for Advanced Computing

Brock Palen brockp@umich.edu

TBD

Outline

Resources

Mechanics: Usage

- Resources
 - Hardware
 - Default Software
- 2 Mechanics: Usage
 - Compiling programs
 - The Batch System
- The Scheduler
 - Understanding the Scheduler
 - Scheduler Commands
- 4 Summary
 - Resources and Access
 - Job Management
 - Contact



Hardware

Compute Hardware

- 1 Altix node, 32 cores
- 586 Opteron nodes, over 1760 cores
- 400+ nodes on CAEN Grid
- Gigabit networking, Myrinet networking, Infiniband networking
- Upto 96GB of memory (64GB public) for SMP work

Nyx

Resources

0000000

- nyx is the Opteron cluster;
- Login: nyx-login.engin.umich.edu

Mechanics: Usage

- Currently has 7TB NFS file system
- Running RedHat Enterprise Linux 4

Bighouse: Available to Aero Space Dept



Hardware: bighouse

Bighouse

Resources

00000000

bighouse is our Itanium SMP machine;

Mechanics: Usage

- Login: bighouse.engin.umich.edu
- Shares nyx's 7TB NFS file system
- Running SUsE Linux Enterprise Server 10
- ProPack 5 from SGI
- Only available for benchmarking (Private)



The Scheduler

Hardware: Grid

Resources

00000000

nodes are great for paramater sweaps, hundereds of jobs etconly for engin accounts

CAEN Grid

- 400+ Nodes, Dual Core
- All nodes have 2GB Ram
- FAST Single CPU's
- Some Parallel Ability
- Short Jobs Only







Mechanics: Usage

The Scheduler

Summary

Software

Nyx Defaults

- OpenMPI
- PGI Compilers
- PBS commands

Bighouse Defaults

- Message Passing Toolkit (MPT)
- Intel Compilers
- PBS commands

Grid Defaults

- OpenMPI
- PGI Compilers
- Intel Compilers

1. Show Example



lanipulating Software All CAC systems use modules to control software. Users can and should write their own modules if needed.

Manipulating Software

Resources

0000000

Mechanics: Usage

All CAC systems use modules to control software. Users can and should write their own modules if needed.

The Scheduler

module commands

- module list Show loaded modules
- module load modulename Load *modulename* for use
- module avail modulename Show available versions of module modulename
- module rm modulename Remove currently loaded module





Module Fun

00000000

1. example suing fftw follows

Module Customization

- /privatemodules/default Allows users to change their default modules.
- /privatemodules/module/version Holds user created module

Mechanics: Usage

• man modulefile

Module Example

v List Loaded Modules

module list

- List Loaded Modules module list
- Show All Modules module avail
- Show All Versions of openmpi module avail openmpi
- Load FFTWmodule load fftw
- Show Variables defined by FFTW module show fftw

Module Example

- List Loaded Modules module list
- Show All Modules module avail
- Show All Versions of openmpi module avail openmpi
- Load FFTW

 module load fftw
- Show Variables defined by FFTW module show fftw

Module Example

Resources

0000000

- List Loaded Modules module list
- Show All Modules module avail
- Show All Versions of openmpi module avail openmpi
- Load FFTW
- Show Variables defined by FFTW

Module Example

- List Loaded Modules module list
- Show All Modules module avail
- Show All Versions of openmpi module avail openmpi
- Load FFTW module load fftw
- Show Variables defined by FFTW

u Show Variables defined by FFTW module show fftw echo \$FFTW_LINK

Module Example

- List Loaded Modules module list
- Show All Modules module avail
- Show All Versions of openmpi module avail openmpi
- Load FFTW module load fftw
- Show Variables defined by FFTW module show fftw echo \$FFTW_LINK

Tools

- All of the standard GNU/Linux tools are also available: make, autoconf, awk, sed, Perl, Python,
- We support emacs, vi{m}, and nano (a pico-like editor) on
- Only use notepad on Windows!

Mechanics: Usage

•00000000000

• If made on windows fix with dos2unix filename

0013

 All of the standard GNU/Linux tools are also available: make, autoconf, awk, sed, Perl, Python,

The Scheduler

- We support emacs, vi{m}, and nano (a pico-like editor) on the clusters. etc.
- Only use notepad on Windows!

Mechanics: Usage

•00000000000

• If made on windows fix with dos2unix filename

Tools

 All of the standard GNU/Linux tools are also available: make, autoconf, awk, sed, Perl, Python,

The Scheduler

- We support emacs, vi{m}, and nano (a pico-like editor) on the clusters. etc.
- Only use notepad on Windows!

Mechanics: Usage

•00000000000

• If made on windows fix with dos2unix filename

- 1. The following applies to the default modules
- 2. Grid: both compilers support OpenMP

The Scheduler

Compile Code

Nyx

Resources

- Use: mpicc, mpiCC, mpif90 for MPI code
- Use: pgcc, pgCC, pgf90 with -mp for OpenMP Code

Bighouse

- Use: icc, icpc, ifort with -lmpi for MPI code
- Use: icc, icpc, ifort with -openmp for OpenMP code

CAEN Grid

- Use: mpicc, mpiCC, mpif90 for MPI code
- Serial code: Intel or PGI commands are valid.



Compile Example

Resources

Copy Code: cp ~brockp/mpicodes.tar ~
tar -xvf mpicodes.tar
cd ~/mpicodes

Serial Code

- Fortran 90 pgf90 -fastsse -o f90hello helloworld.f90
- C
 pgcc -fastsse -o chello helloworld.c

Compile Example Ctd.

1. 'man make' Make lets you manage large bits of code. Works for all source types

—Compile Example Ctd.

Mechanics: Usage

• mpicc -fastsse -o c_ex01 c_ex01.c

Resources

1. 'man make' Make lets you manage large bits of code. Works for all source types

```
MPI Code
  make
 • mpirun -np 2 c_ex01
 • Thats it... Ok not really
  • make clean
```

Compile Example Ctd.

1. 'man make' Make lets you manage large bits of code. Works for all source types

MPI Code

Resources

- make
- \bullet mpirun -np 2 c_ex01
- Thats it... Ok not really
- make clean
- mpicc -fastsse -o c_ex01 c_ex01.c

reduction to the PBS Batch System

FBS

All access to the companie coding (reverting other than the page road) is visit backed System

We use a system called Forage, it, it defined from PBS

The hards system code for a system called Forage as it, it defined from PBS

The shade system code for a system coding if and where jobs can run

The seal response for the system decided if and where jobs can run

These as a region system code for grant and the system of t

 Resources
 Mechanics: Usage

 0000000
 00000000

The Scheduler

Introduction to the PBS Batch System

PBS

- All access to the compute nodes (everything other than the login node) is via the batch system
- We use a system called Torque, it is derived from PBS
- The batch system controls access to queues
- The scheduling system decides if and where jobs can run
- There is a single public queue: cac
- There are many private queues for people who own or rent nodes
- If you don't know use the route queue



reduction to the PBS Batch System

FBS

All access to the compute nodes (resysting other than the lagge node) is visit batch hystem.

We use a system called Torque, it is defined from PBS.

We use a system called Torque, it is defined from PBS.

The last this year notice access to question.

There is a negleg palicie guest or as.

There is a negleg palicie quest for paged who one or not not not.

All the participations are paged with one or not not not.

All you don't show use the resiste queste.

 Resources
 Mechanics: Usage

 0000000
 00000000

The Scheduler

Summar

Introduction to the PBS Batch System

PBS

- All access to the compute nodes (everything other than the login node) is via the batch system
- We use a system called Torque, it is derived from PBS
- The batch system controls access to queues
- The scheduling system decides if and where jobs can run
- There is a single public queue: cac
- There are many private queues for people who own or rent nodes
- If you don't know use the route queue

This steps to using the batch system are:

**O Cracles a batch file: this is a bard (E-15 lines) text file with more contained to the commands to the commands to the command to the comma

troduction to the PBS Batch System

Mechanics: Usage ○○○○●○○○○○ he Scheduler

Summary

Introduction to the PBS Batch System

PBS Files

Resources

The steps to using the batch system are:

- Create a batch file: this is a short (5-15 lines) text file with some batch commands and the commands to run your program
- Submit the file to the batch system
- Oheck on the status of your job
- Delete your job if you want to cancel it

Creating a PBS Batch File

```
#!/bin/sh
#PBS -N 1-cpu
#PBS -l nodes=1,walltime=1:00:00
#PBS -M brockp@umich.edu
```

Creating a PBS Batch File

```
#!/bin/sh
#PBS -N 1-cpu
#PBS -l nodes=1,walltime=1:00:00
```

Creating a PBS Batch File

```
#!/bin/sh
#PBS -N 1-cpu
#PBS -l nodes=1,walltime=1:00:00
#PBS -m abe
#PBS -M brockp@umich.edu
```

Creating a PBS Batch File

```
#!/bin/sh
#PBS -N 1-cpu
#PBS -l nodes=1,walltime=1:00:00
#PBS -m abe
#PBS -M brockp@umich.edu
#PBS -q route
```

Mechanics: Usage

The Scheduler

Creating a PBS Batch File

```
#!/bin/sh
#PBS -N 1-cpu
#PBS -l nodes=1,walltime=1:00:00
#PBS -m abe
#PBS -M brockp@umich.edu
#PBS -q route
#PBS -j oe
```

Creating a PBS Batch File

```
#!/bin/sh
#PBS -N 1-cpu
#PBS -l nodes=1,walltime=1:00:00
#PBS -m abe
#PBS -M brockp@umich.edu
#PBS -q route
#PBS -j oe
#PBS -V
```

Mechanics: Usage

The Scheduler

Creating a PBS Batch File

Resources

```
#!/bin/sh
#PBS -N 1-cpu
#PBS -l nodes=1,walltime=1:00:00
#PBS -m abe
#PBS -M brockp@umich.edu
#PBS -q route
#PBS -j oe
#PBS -V
cat $PBS_NODEFILE
```

```
Resources Mechanics: Usage
```

The Scheduler

Creating a PBS Batch File

```
#!/bin/sh
#PBS -N 1-cpu
#PBS -l nodes=1,walltime=1:00:00
#PBS -m abe
#PBS -M brockp@umich.edu
#PBS -q route
#PBS -j oe
#PBS -V
cat $PBS_NODEFILE
cd ~/input1dir/
mcnp5.mpi i=input o=output r=restart
```

```
#PBS -1 nodes=8:ppn=2,walltime=8:00:00
mpirum -mp 16 mcmp5.mpi i=imput2 o=output2 r=restart2
```

Creating a PBS Batch File

#PBS -M brockp@smich.edu #PBS -m ae

#1/bin/sh

#PBS -N mcmp-8x2

cd \${HOME}/input2/ echo "I ran on: ' cat \$PBS_NODEFILE

Resources

Creating a PBS Batch File

Mechanics: Usage

A more complicated example:

```
#!/bin/sh
#PBS -N mcnp-8x2
#PBS -l nodes=8:ppn=2,walltime=8:00:00
#PBS -q route
#PBS -M brockp@umich.edu
#PBS -m ae
#PBS -j oe
#PBS -V
cd ${HOME}/input2/
echo "I ran on: "
cat $PBS_NODEFILE
mpirun -np 16 mcnp5.mpi i=input2 o=output2 r=restart2
```

The Scheduler





Mechanics: Usage Resources

The Scheduler

Submitting, Checking, and Deleting Batch Jobs

• After you create your PBS script, you need to submit it:

```
$ qsub mcnp.q
542.nyx-login.engin.umich.edu
```

After you submit your script, you can check on the status of

• If you want to delete your job:



```
CAC Intro

Mechanics: Usage

The Batch System
Submitting, Checking, and Deleting Batch Jobs
```

```
Submitting, Checking, and Deleting Batch Jobs

• After you create your PES copts, you need to submit it:

• After you would your script, you can check on the state of your job.

• After you would your script, you can check on the state of your job.

• After you would you script, you can check on the state of your job.

• After you would you would you would not be you you the state of your job.

• After you would not you would not be you you the state of your job.

• After you would not you would not be you you the state of your job.

• After you would not you would not be you you would not you wou
```

Resources Mechanics: Usage

The Scheduler

Summai

Submitting, Checking, and Deleting Batch Jobs

After you create your PBS script, you need to submit it:

```
$ qsub mcnp.q
542.nyx-login.engin.umich.edu
```

 After you submit your script, you can check on the status of your job:

• If you want to delete your job:

```
$ adel 54
```



```
Submitting, Checking, and Deleting Batch Jobs

a After you create your PES cript, you need to submit it:

**The you can't your person, you can check on the status of your job.

**The you can't you go you can check on the status of your job.

**The your can't you can't you you can check on the status of your job.

**The your can't you can't you can check on the status of your job.

**The your can't you c
```

Resources 00000000 Mechanics: Usage

The Scheduler

Summar

Submitting, Checking, and Deleting Batch Jobs

• After you create your PBS script, you need to submit it:

```
$ qsub mcnp.q
542.nyx-login.engin.umich.edu
```

 After you submit your script, you can check on the status of your job:

• If you want to delete your job:

```
$ qdel 542
```



PBS Email

Resources

Mechanics: Usage ○○○○○○○○○ The Scheduler

Summar

PBS will send an email at the start and end of your job if you use the -m and -M options in your PBS script. The email after a job completes successfully looks like:

Date: Sun, 30 Apr 2006 12:50:17 -0400
From: adm <adm@nyx-login.engin.umich.edu>
To: "Palen, Brock E"

Frockp@umich.edu>
Subject: PBS JOB 542.nyx-login.engin.umich.edu

PBS Job Id: 542.nyx-login.engin.umich.edu
Job Name: mcnp-8x2
Execution terminated
Exit_status=0
resources_used.cput=13:17:26
resources_used.mem=1220672kb
resources_used.vmem=11146704kb
resources_used.walltime=00:49:57



PBS Example Job

cd ~/mpicodes
nano run
Edit #PBS -M
Ctl+o
Ctl+x
qsub run

Mechanics: Usage ○○○○○○○○○○

Mechanics: Usage

Interactive Jobs

Resources

Interactive Jobs

The CAC has CPU's for jobs 15 minutes or less

These CPU's can be used for testing PBS scripts and debugging code

Interactive jobs allow users to interact with the shell on a remote node

Interactive Jobs

Resources

Interactive Jobs

The CAC has CPU's for jobs 15 minutes or less These CPU's can be used for testing PBS scripts and debugging code

Interactive jobs allow users to interact with the shell on a remote node

Interactive Jobs

Resources

Interactive Jobs

The CAC has CPU's for jobs 15 minutes or less

These CPU's can be used for testing PBS scripts and debugging code

Interactive jobs allow users to interact with the shell on a remote node

Mechanics: Usage ○○○○○○○○○ The Scheduler

Summar

Interactive Jobs

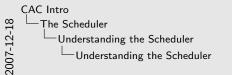
Resources

Interactive Jobs

The CAC has CPU's for jobs 15 minutes or less

These CPU's can be used for testing PBS scripts and debugging code

Interactive jobs allow users to interact with the shell on a remote node



and where. There are many factors that go into the scheduler's

m Maximum number of CPUs in use by one person: depends or

Mechanics: Usage Resources

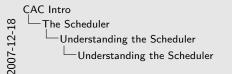
The Scheduler

Understanding the Scheduler

The scheduler determines what jobs can run, when they can run, and where. There are many factors that go into the scheduler's decision.

Limits and Priority

- Maximum number jobs eligible for scheduling: 4
- Maximum number of CPUs in use by one person: depends on queue
- Maximum number of jobs in the queue at one time: no limit
- Who you are: user and group level priorities
- How long you've waited: the longer you wait, the higher your
- Your recent usage (fairshare): People with less usage over the



and where. There are many factors that go into the scheduler's

- Maximum number of CPUs in use by one person: depends of

- a How long you've waited: the longer you wait, the higher your
 - a Your recent usage (fairshare): People with less usage over the past month will have a higher priority than those with a lot of

Resources

Mechanics: Usage

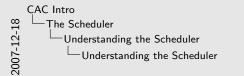
The Scheduler

Understanding the Scheduler

The scheduler determines what jobs can run, when they can run, and where. There are many factors that go into the scheduler's decision.

Limits and Priority

- Maximum number jobs eligible for scheduling: 4
- Maximum number of CPUs in use by one person: depends on queue
- Maximum number of jobs in the queue at one time: no limit
- Who you are: user and group level priorities
- How long you've waited: the longer you wait, the higher your priority
- Your recent usage (fairshare): People with less usage over the past month will have a higher priority than those with a lot of usage





Understanding the Scheduler

Mechanics: Usage

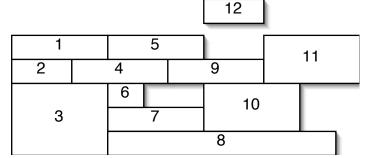
- Reservations
 - Advance reservations: holds nodes for users or groups
 - Job reservations: scheduler will reserve nodes for the next several jobs in each queue

The Scheduler

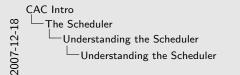
Backfill

Resources

 If the reservations leave holes in the schedule, they may be filled by short jobs that otherwise would have waited.









Understanding the Scheduler

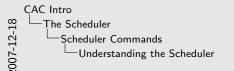
- Reservations
 - Advance reservations: holds nodes for users or groups

The Scheduler

- Job reservations: scheduler will reserve nodes for the next several jobs in each queue
- Backfill
 - If the reservations leave holes in the schedule, they may be filled by short jobs that otherwise would have waited.

1 5 11 2 4 9 11 3 7 8





derstanding the Scheduler

There are several commands that can give you insight into the school derisions

- abovq shows the state of the queue at that moment in time, showing the running jobs in order of soonest to finish to longest to finish; the idle jobs in order of priority; and the blocked jobs in the order they were submitted
- diagnose -p shows the factors that go into computing the priority for all of the idle jobs check tob robnumber — for idle jobs this will show why the
- job can't start

 u shougtart jobsumber this makes a (ooor) estimate of
- when the job will start

Resources

Mechanics: Usage

The Scheduler

Understanding the Scheduler

There are several commands that can give you insight into the scheduler's decisions.

- showq shows the state of the queue at that moment in time, showing the running jobs in order of soonest to finish to longest to finish; the idle jobs in order of priority; and the blocked jobs in the order they were submitted
- diagnose -p shows the factors that go into computing the priority for all of the idle jobs
- checkjob jobnumber for idle jobs this will show why the job can't start
- showstart *jobnumber* this makes a (poor) estimate of when the job will start



Summary

Resources

Summary

- Resources
 - Lots of CPUs
 - A reasonable amount of software
 - Watch or subscribe to http://cac.engin.umich.edu for updates
- Access
 - All access is via the SSH family of commands: ssh, sftp, scp
 - There are lots of clients for these commands for the different platforms
 - There is no graphical access, everything is via the command line

Summary

Resources

Summary Ctd.

- Job Submission
 - Every job needs a PBS script file
 - Two most important commands: qsub and qstat -au uniqname
- Job Scheduling
 - Scheduling depends on a lot of factors, it is best to submit jobs and let the scheduler optimize for their start.

Summary

Resources

Summary Ctd.

- News: http://cac.engin.umich.edu
 - RSS feed
 - New of changes, outages, other pertinent piece of information
- Contact: cac-support@umich.edu
 - Questions or concerns should be sent here (not to an individual) since this is read by six people. The odds of a quick reply are best this way.
 - We aren't parallel programmers, but we'll do what we can to help.

