Martinos Center Compute Clusters

- Intro
  - What are the compute clusters
  - How to gain access
  - Housekeeping
- Usage
  - Log In
  - Submitting Jobs
  - Queues
  - Request CPUs/vmem
  - Email Status
  - I/O
  - Interactive
  - Dependencies
    - Daisy Chain
    - Wrapper Script
    - In Progress
- Job Status
  - Running Jobs
    - - Show Job Status
    - - See Standard Output
  - Completed Jobs
  - Failed Jobs
  - Delete Jobs

Martinos Center Compute Clusters

Intro - What are the compute clusters?

- launchpad
  - At Needham Data Center
  - 127 nodes
    - ~115 “normal” nodes
      - Two 64 bit Intel Xeon quad cores
      - 56 GB RAM
    - ~12 GPU nodes
      - Available exclusively for GPU jobs
- tensor
  - In CNY “cold room”
  - 107 nodes
    - 89 “normal” nodes
      - Dual cores with 4GB RAM
    - 18 “big” nodes
      - Dual cores with 8GB RAM each

I will only talk about launchpad today. We recommend you primarily use launchpad. Use tensor if launchpad is full, if launchpad resources are overkill, or if the data for your jobs are local and are slowed by the I/O problem.

Martinos Center Compute Clusters

Intro - How to gain access

Email: clusteradm [at] nmr.mgh.harvard.edu
Let us know who/what/why/how you need access.

Martinos Center Compute Clusters

Intro - Housekeeping

Questions?:
- Any specific questions (extend the walltime of your jobs etc.) can be sent to us: clusteradm [at] nmr.mgh.harvard.edu
- General questions can be sent to the group for advice and help: batch-users [at] nmr.mgh.harvard.edu

Limits:
- We prefer each user to only use 100 job slots during normal usage.
- Evenings/weekends you may use up to 200 slots of CPU/vmem
- While there is a queue, we request you only use 50 CPU/vmem slots

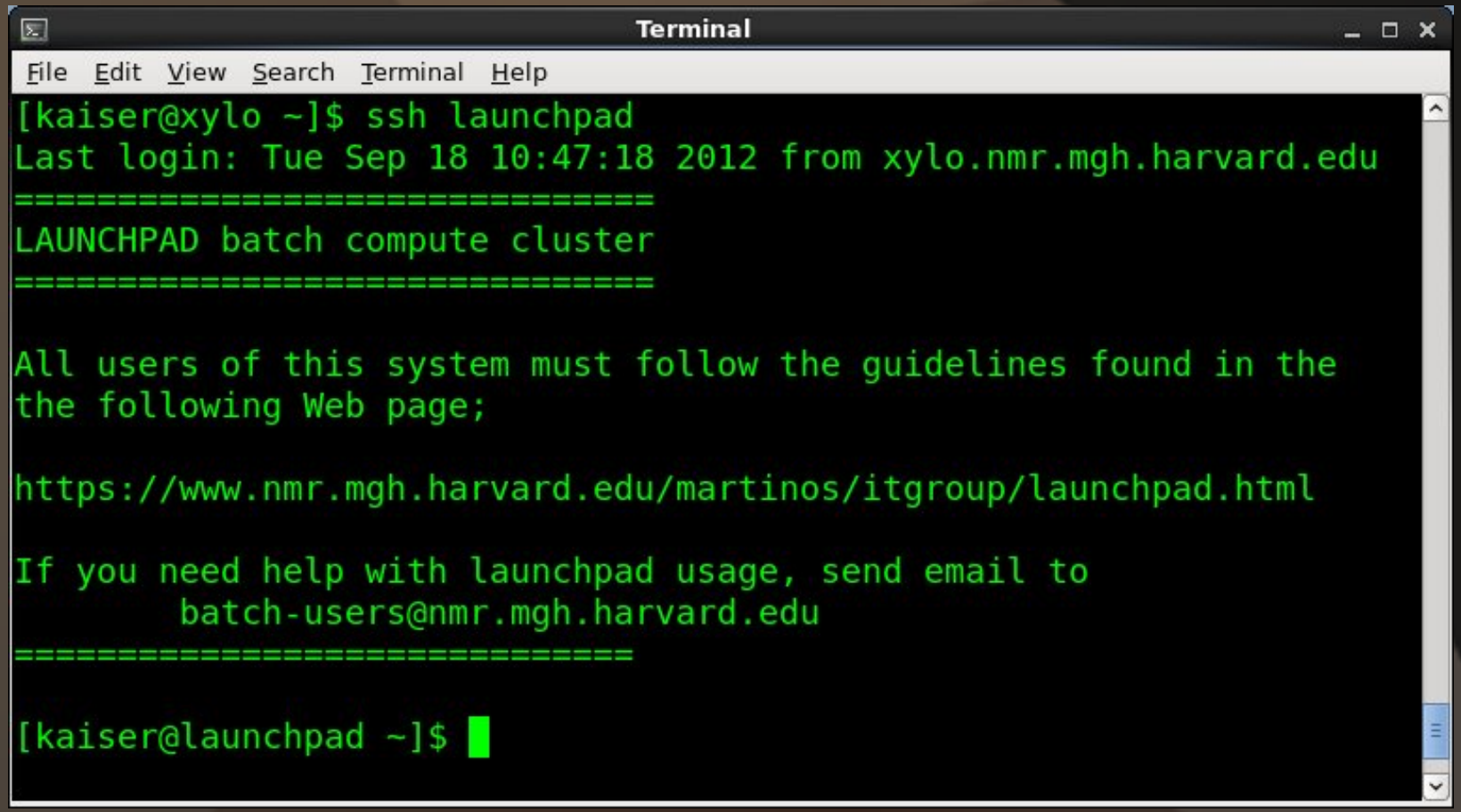
Do not run another directly on launchpad. Submit your jobs. Any programs found running on the master node will be deleted, no exceptions.

Matlab:
- There are only 60 Matlab licenses for the entire center. For this reason, we recommend any Matlab code submitted for execution should be “compiled” ahead of time. Please see the URL to the article on how to do it.
When the program is compiled, it doesn't use a matlab license and they are no longer under a matlab restriction.

Courtesy of coutu:
http://nmr.mgh.harvard.edu/martinos/itgroup/deploytool.html

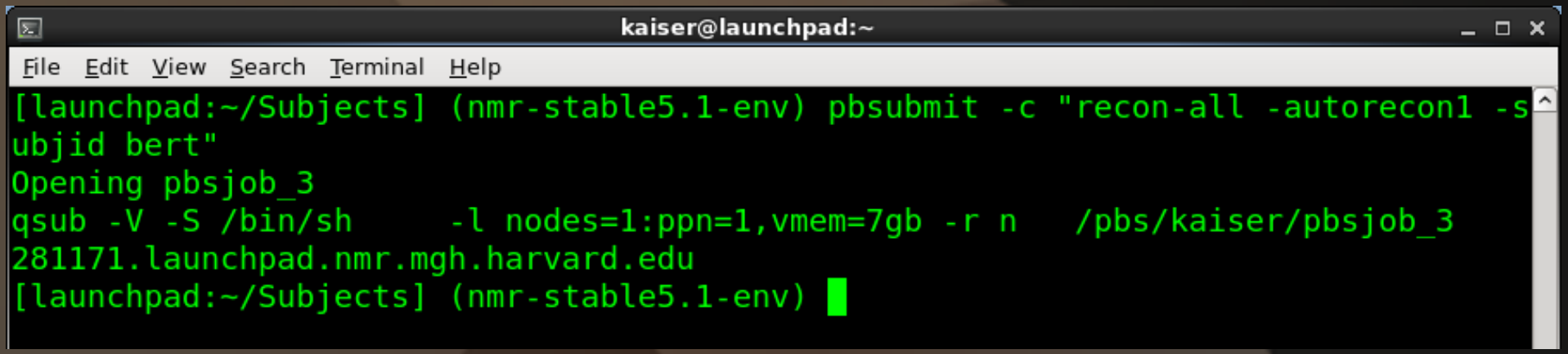
Martinos Center Compute Clusters

Usage - Log In



Martinos Center Compute Clusters

Usage – Submitting Jobs



pbsubmit is a wrapper script that:

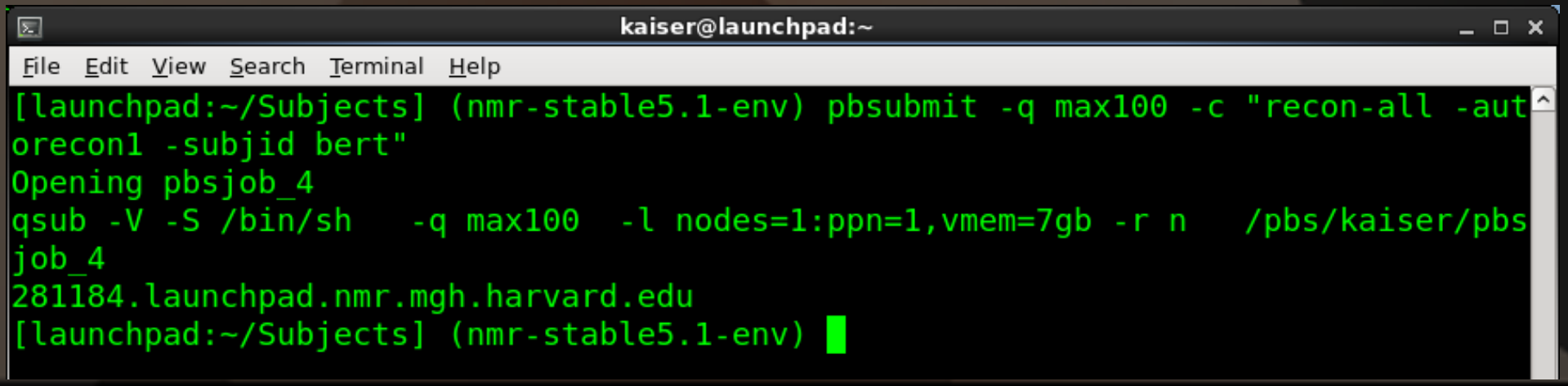
- formats the command that is executed (/pbs/kaiser/pbsjob_3)
- automatically selects the default settings (unless overridden)
  - Select the number of nodes (nodes=1)
  - Select the number of CPUs (ppn=1)
  - Select the amount of virtual memory (vmem=7gb)
- submits the job using the qsub command

pbsjob_3 is the Job Number
281171.launchpad.nmr.mgh.harvard.edu is the Job ID

Martinos Center Compute Clusters

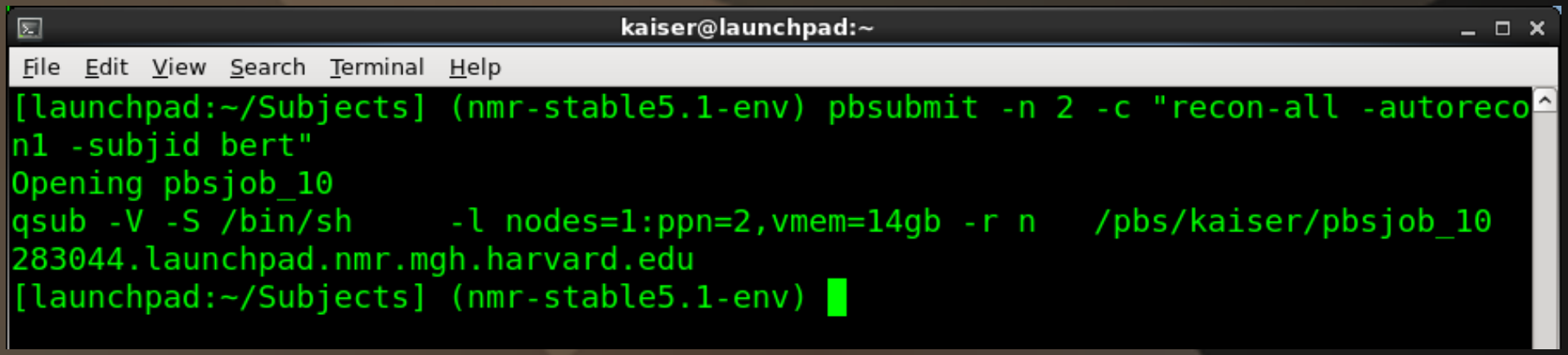
Usage - Queues

Queue	Priority	Max CPU/User	Description
default	10100	150	Walltime of 96 hours
p20	10200	Unlimited	
p30	10300	Unlimited	
GPU	90	Unlimited	GPU nodes
extended	8000	50	Walltime of 196 hours
matlab	10100	20	Limit of 60 matlab licenses for the Center
max10	10100	10	
max20	10100	20	
max50	10100	50	
max75	10100	75	
max100	10100	100	
max200	8000	200	



Martinos Center Compute Clusters

Usage - Request CPUs/vmem



Only request more CPUs or Virtual Memory if you need them.

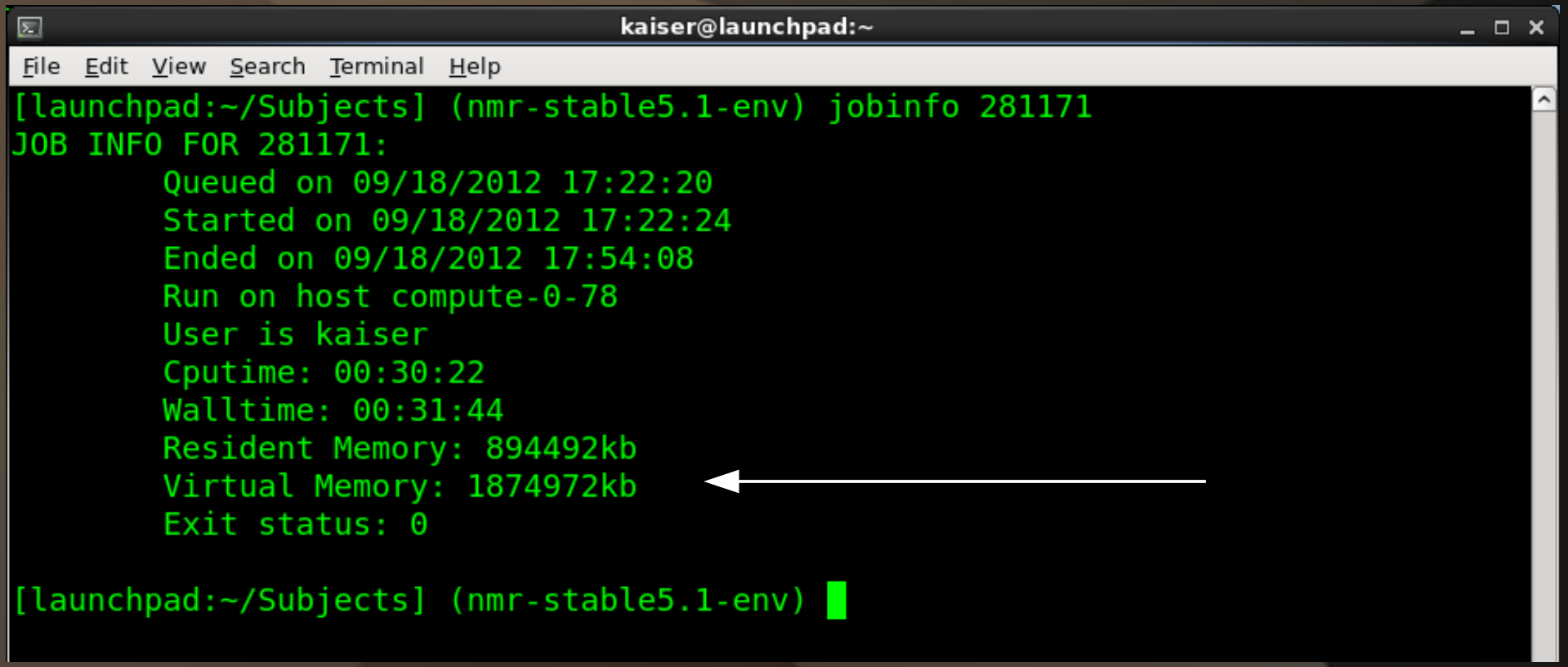
CPUs
- You should only request extra CPUs if the program you are running is multi-threaded.
- If you aren't sure if the program is multi-threaded, it probably isn't.

Virtual Memory
- Only request as much as you need.
- If you aren't sure how much you'll need, run a single test case. Start with the default of 7GB of vmem. If it fails due to a lack of memory, request 14GB. Then 21GB etc...

So, how much virtual memory did the job use?

Martinos Center Compute Clusters

Usage – Request CPUs/vmem



Only used 1.9GB of virtual memory. Safely under the default request of 7GB. No need to ask for more.

Limits – keep in mind that we prefer each user to only use 100 job slots during the day.
A job that requests 1 CPU and 14GB of vmem counts as two slots worth of resources.
Submit the jobs to the max50 queue ('-q max50') to self-regulate.

Martinos Center Compute Clusters

Usage - Email Status

```
kaiser@launchpad:~
File  Edit  View  Search  Terminal  Help
[launchpad:~/Subjects] (nmr-stable5.1-env) pbsubmit -m kaiser -c "recon-all -aut
orecon1 -subjid bert"
Opening pbsjob_5
qsub -V -S /bin/sh  -m abe -M kaiser   -l nodes=1:ppn=1,vmem=7gb -r n   /pbs/kai
ser/pbsjob_5
282509.launchpad.nmr.mgh.harvard.edu
[launchpad:~/Subjects] (nmr-stable5.1-env)
```

Start Execution:

```
pine
File  Edit  View  Search  Terminal  Help
  ALPINE 2.02(1266)   MESSAGE TEXT    INBOX     Message 2,

Date: Wed, 19 Sep 2012 13:09:30 -0400
From: adm <adm@nmr.mgh.harvard.edu>
To: kaiser@nmr.mgh.harvard.edu
Subject: PBS JOB 282509.launchpad.nmr.mgh.harvard.edu

PBS Job Id: 282509.launchpad.nmr.mgh.harvard.edu
Job Name:   pbsjob_5
Exec host:  compute-0-64/4
Begun execution
```

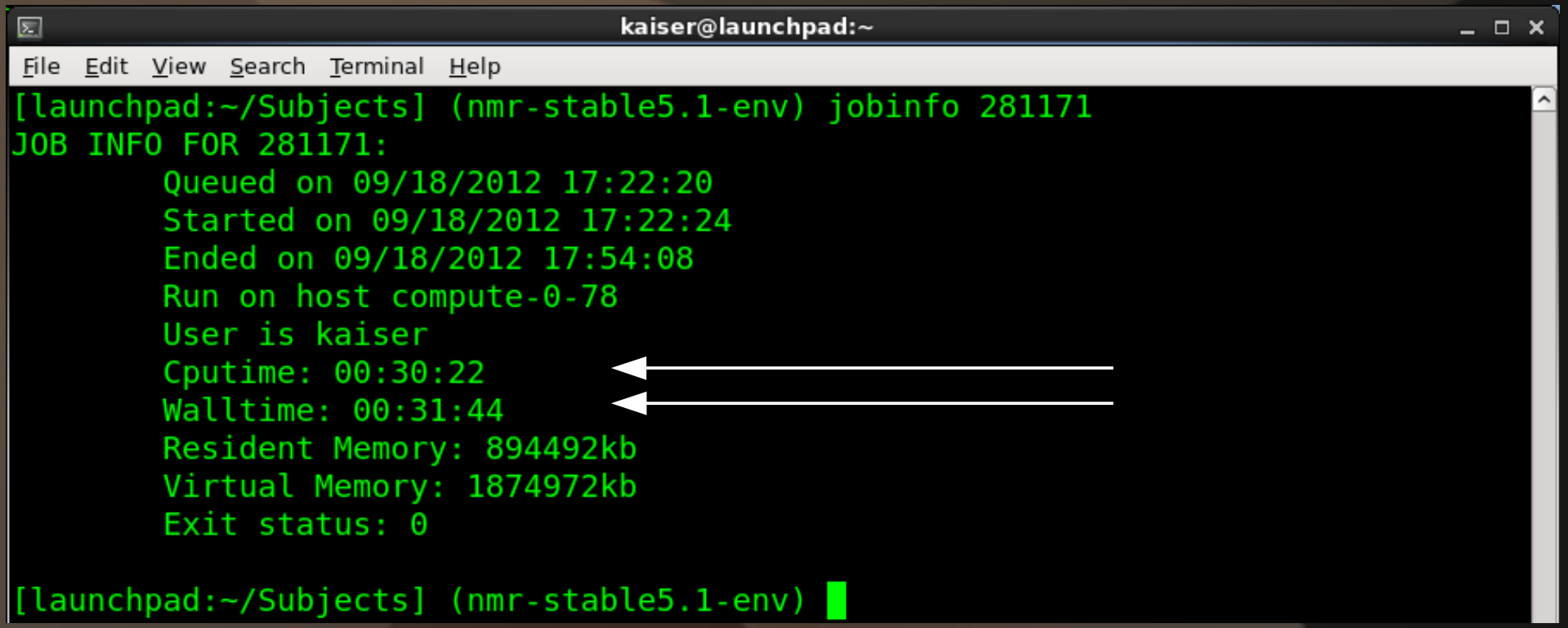
Finish Execution:

```
pine
File  Edit  View  Search  Terminal  Help
  ALPINE 2.02(1266)   MESSAGE TEXT    INBOX     Messag

Date: Wed, 19 Sep 2012 13:18:06 -0400
From: adm <adm@nmr.mgh.harvard.edu>
To: kaiser@nmr.mgh.harvard.edu
Subject: PBS JOB 282509.launchpad.nmr.mgh.harvard.edu

PBS Job Id: 282509.launchpad.nmr.mgh.harvard.edu
Job Name:   pbsjob_5
Exec host:  compute-0-64/4
Execution terminated
Exit_status=0
resources_used.cput=00:07:33
resources_used.mem=984016kb
resources_used.vmem=1964488kb
resources_used.walltime=00:08:36
```

Sends email to user (**replace 'kaiser' with your username**) on job start and finish
- To receive email only if job completes with an error, append '-e' to command line
- To receive email only upon job completion (error or no error), append '-f' to command line

Martinos Center Compute Clusters

Usage - I/O



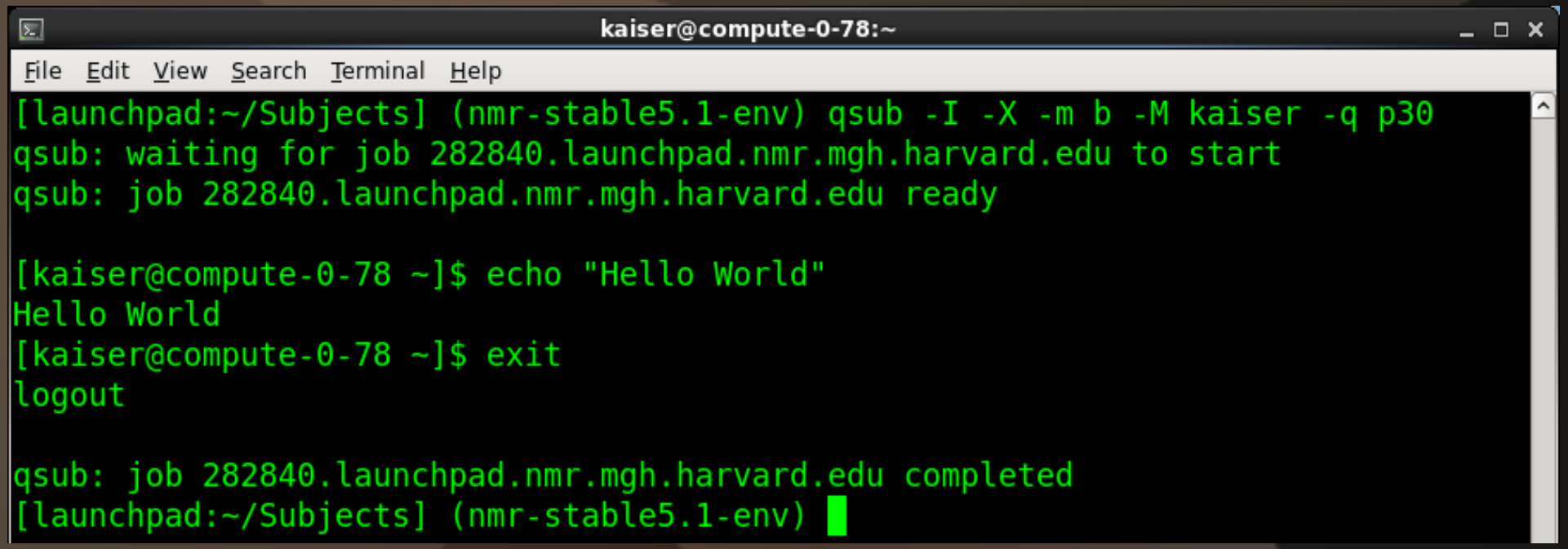
Compare CPUtime and Walltime. If Walltime is larger than CPUtime, time was wasted in I/O.

This job was run using data from my local machine. Over one minute was wasted transferring data back and forth between launchpad (in Needham) to my computer at Martinos.
Tips:

- Copy/move local data to /cluster/ directories before running jobs.
- Have scripts/programs write temp data to /cluster/scratch/
- Instead of launchpad, use tensor which lives in CNY.
- Space out submission of jobs so they don't all have large I/O needs at the same time.

Martinos Center Compute Clusters

Usage – Interactive Jobs



Use the qsub command to start an interactive job using the high priority p30 queue. You will receive an email when the job begins execution. **Replace 'kaiser' with your username!** Actively wait until the job is slated for execution. Don't immediately leave for lunch.

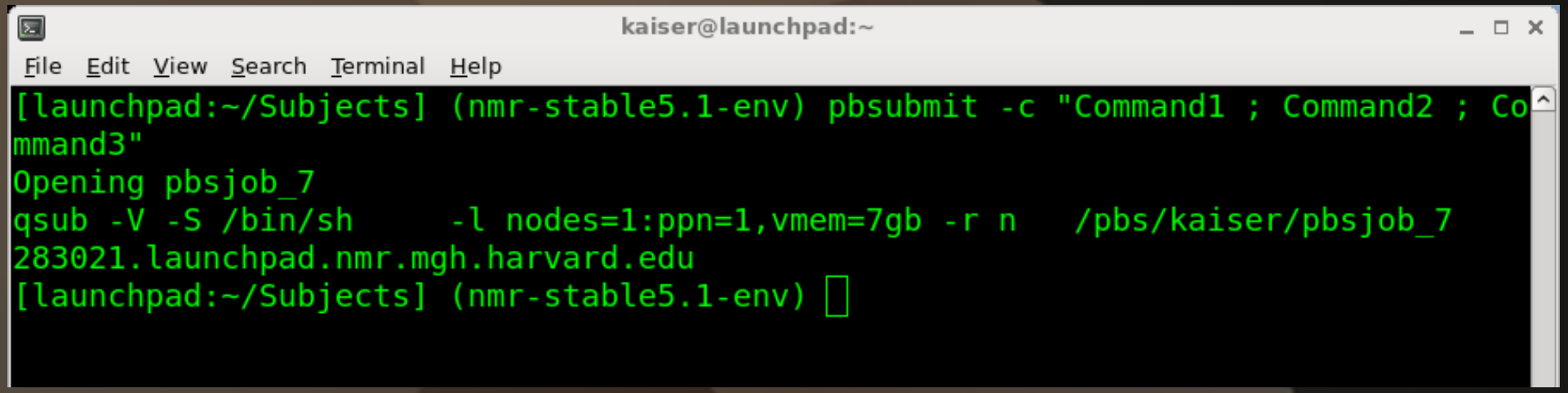
1. As soon as a slot becomes available, the job is assigned a Job ID and you are ssh'ed to the node where your job will execute.
2. Run your commands...
3. When completed, exit out of the node. Your job will not be completed until you exit.

Please attend to an interactive session. As soon as the job begins and you are ssh'ed into the node, you take up a job slot. Exit out of the node as soon as your commands are done. You will continue to take up a job slot until you exit out of the node.

Martinos Center Compute Clusters

Usage – Dependencies – Daisy Chain

If you have a series of commands that you want to execute in a row (one after another). The easiest way to do it is to daisy chain the commands together on the command line:



The commands are separated on the command line by a colon (;).

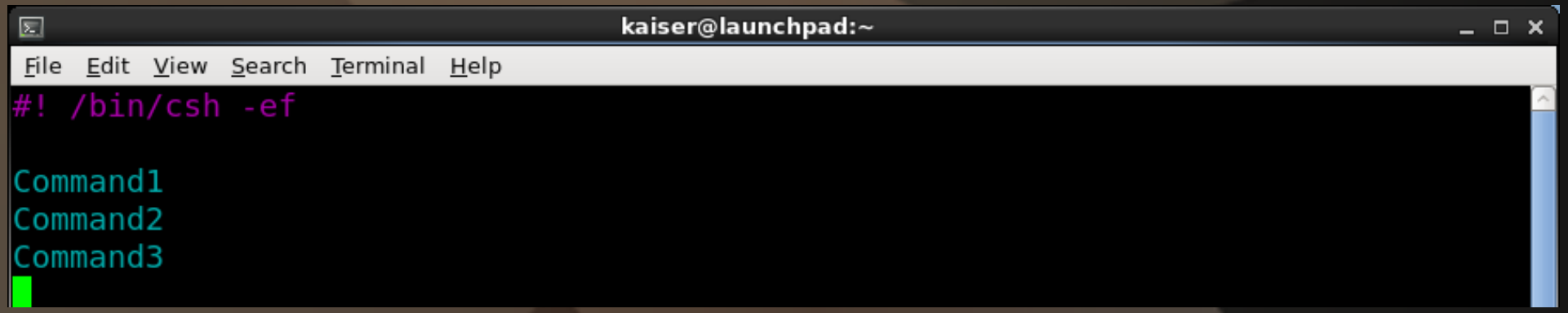
Each command will run even if one before it failed.

Replace Command1, Command2, Command3 with the specific commands you want to run

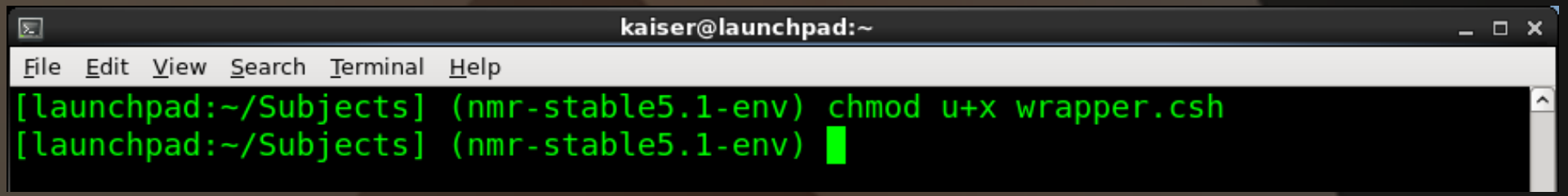
Martinos Center Compute Clusters

Usage – Dependencies – Wrapper Script

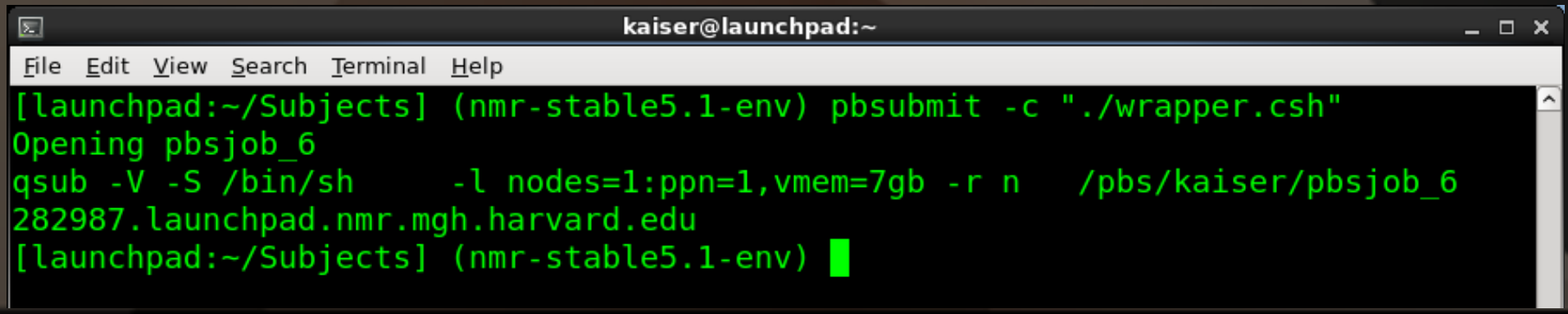
A more elegant way to do it is to write a wrapper script. Use a text editor to create a file called wrapper.csh with these contents:



The -e flag above, instructs the script to exit if any of the individual commands exit with an error.
Make the script executable:



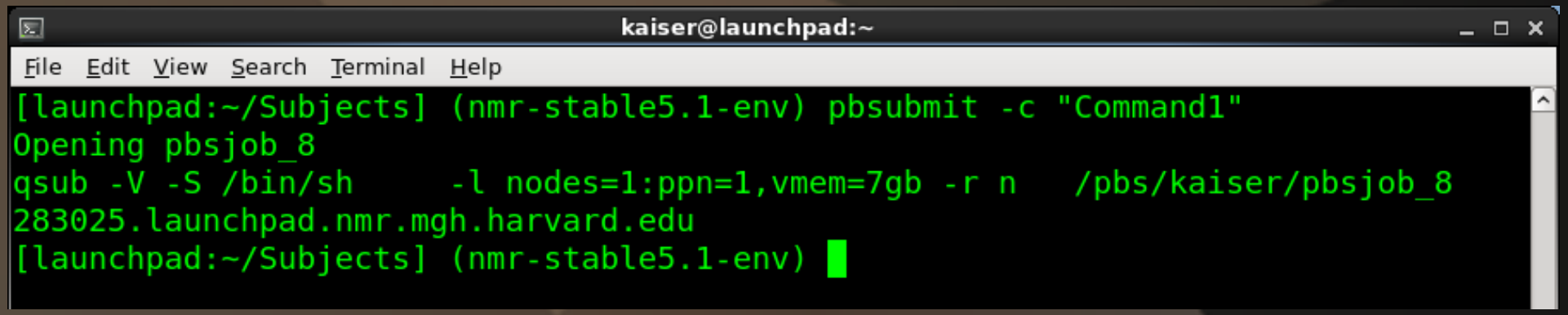
Submit the script for execution:



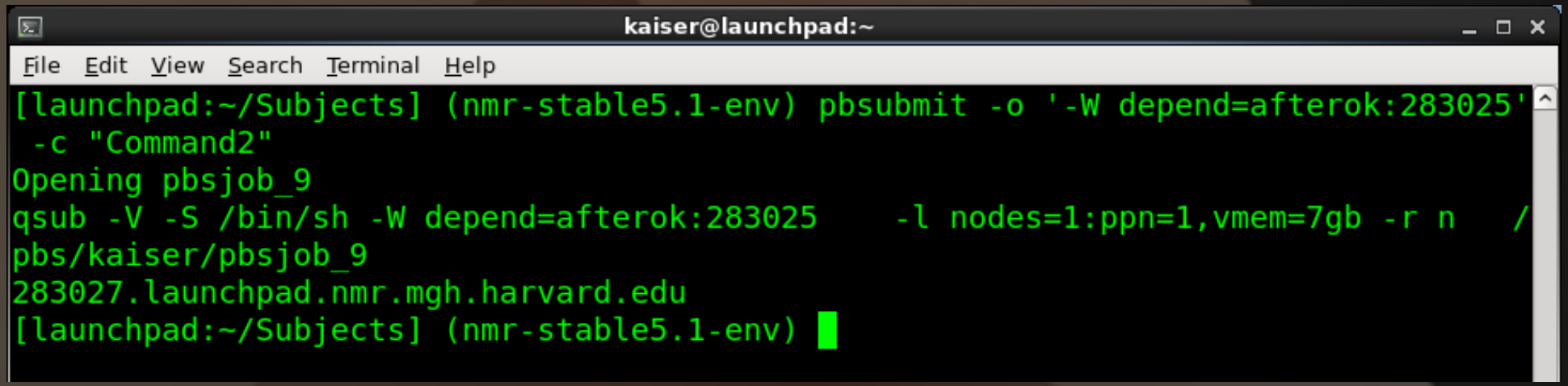
Martinos Center Compute Clusters

Usage – Dependencies – In Progress

If you already have a job running....



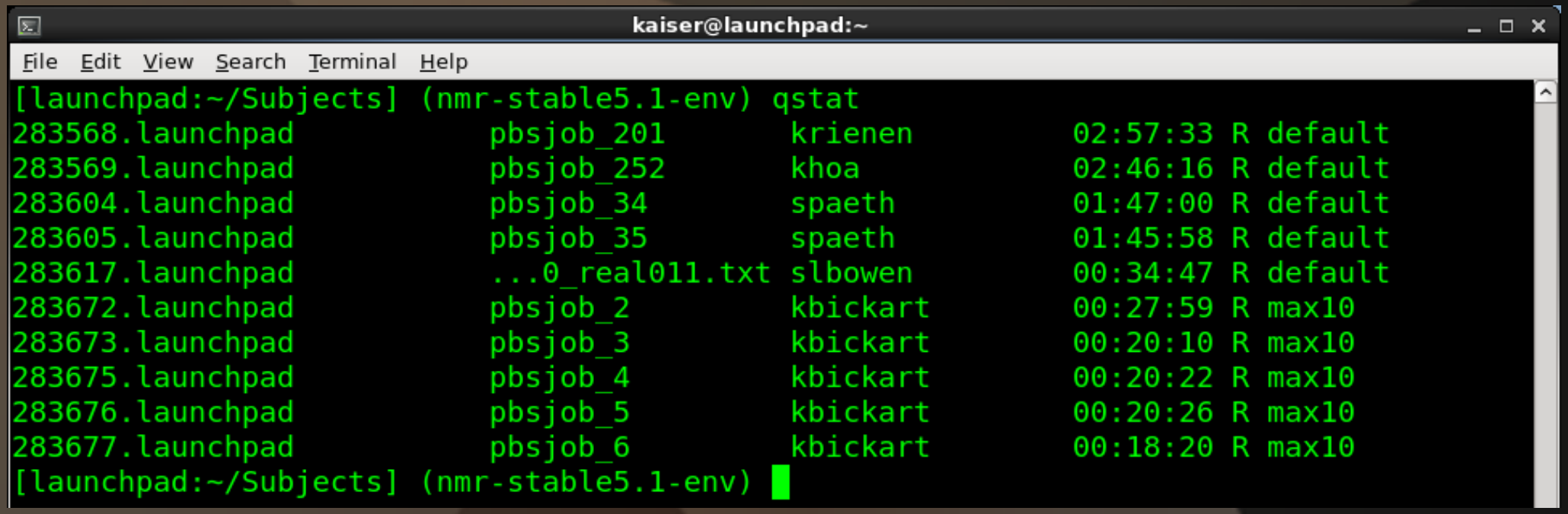
And you want to start another job that will run immediately after the first job completes without error:



This second job will be held until the first one completes without error. If the first job exits with an error, the second job will not run.

Martinos Center Compute Clusters

Job Status - Running Jobs – Show Job Status



Job ID - Job Name - User - CPUtime - State - Queue

[R]unning
[Q]ueued

Additional options:
To see just your jobs:
qstat -u <username>
qstat | grep <username>

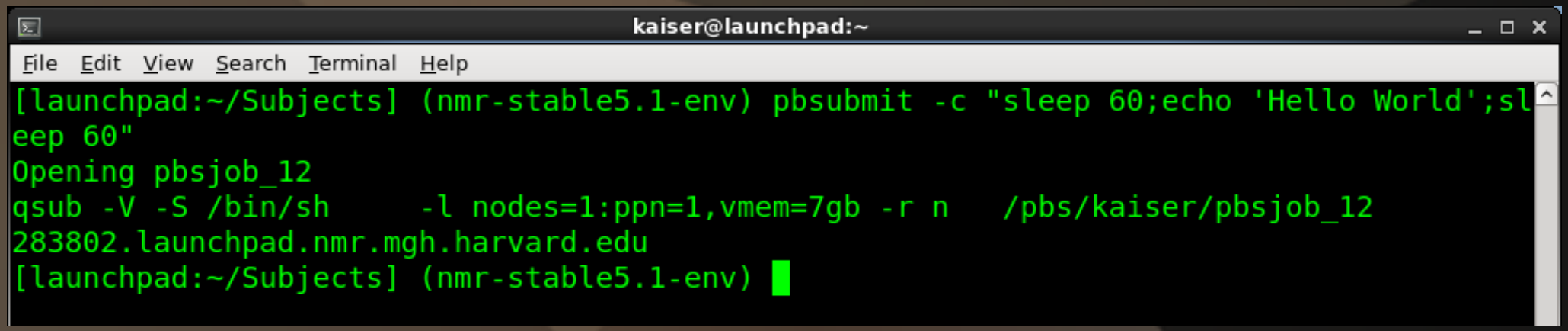
To get all your running and queued jobs:
qstat | grep <username> | grep -w R
qstat | grep <username> | grep -w W

Another Job Status command is 'showq':
To see just your jobs:
showq -u <username>
showq | grep <username>

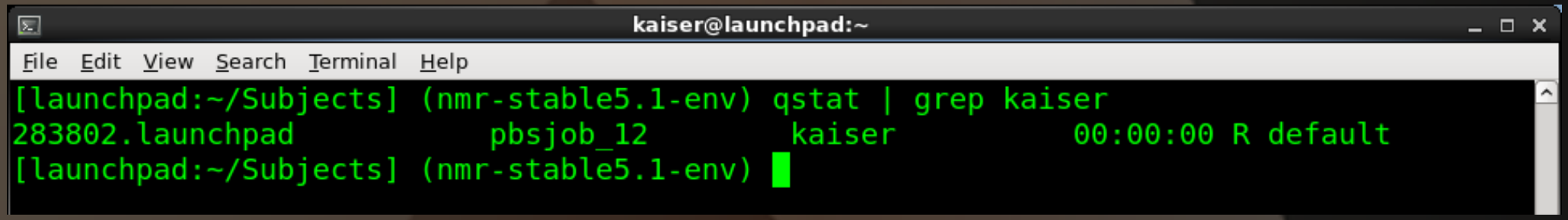
To get all your running and queued jobs:
showq -r -u <username>
showq -i -u <username>

Martinos Center Compute Clusters

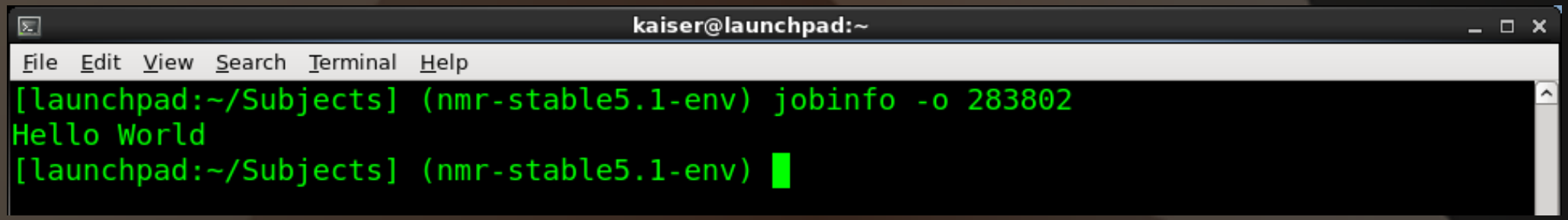
Job Status - Running Jobs – See Standard Output



Job is running:



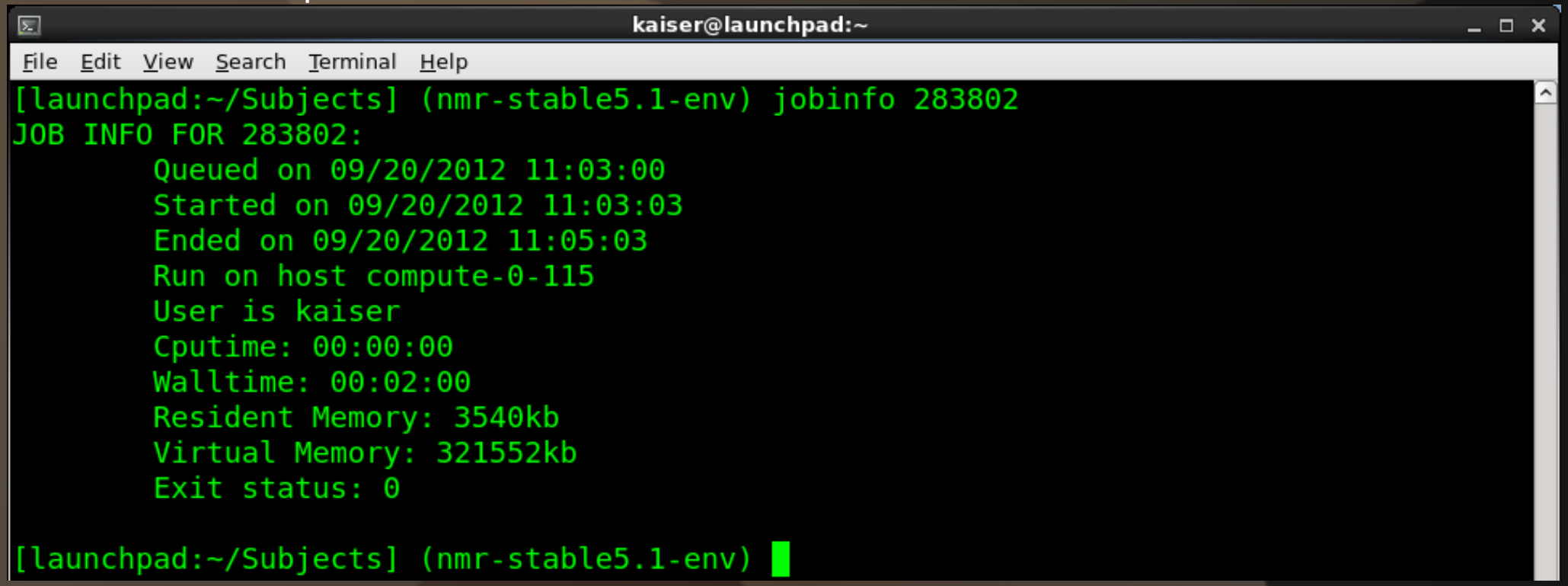
Check on the standard output of the job:



To see the standard error of an actively running job; 'jobinfo -e <Job ID>'

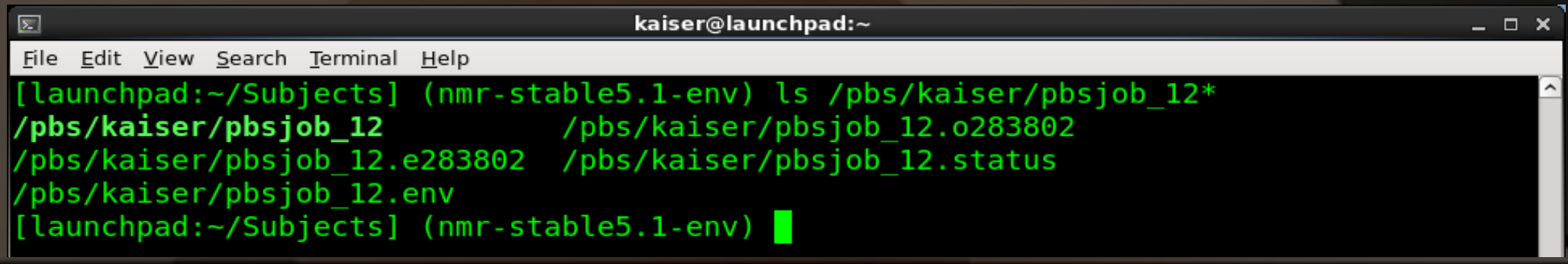
Martinos Center Compute Clusters

Job Status - Completed Jobs



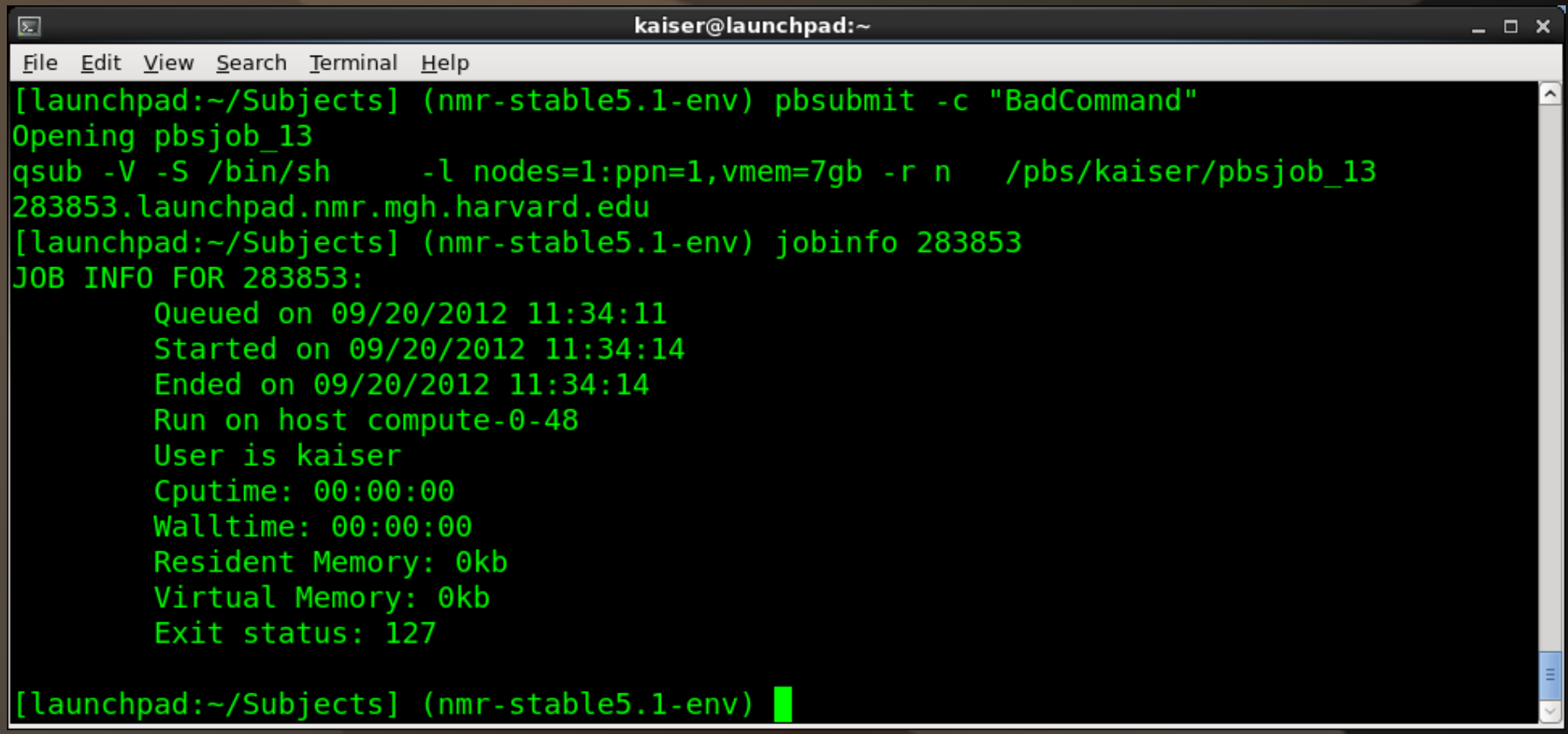
Check the Exit Status of the job. Zero means it successfully completed.

The job script, standard output, standard error and the exit status are all saved as separate text files in your pbs directory:



Martinos Center Compute Clusters

Job Status - Failed Jobs



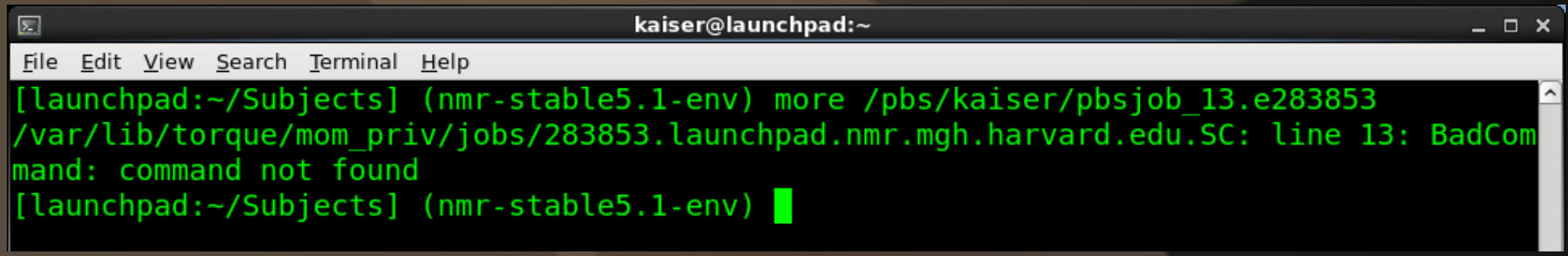
Ack! My job finished with an Exit Status of 127.

How do I troubleshoot???

Martinos Center Compute Clusters

Job Status - Failed Jobs

Check the standard error and standard output files for any hints:



Other Possible Hints:

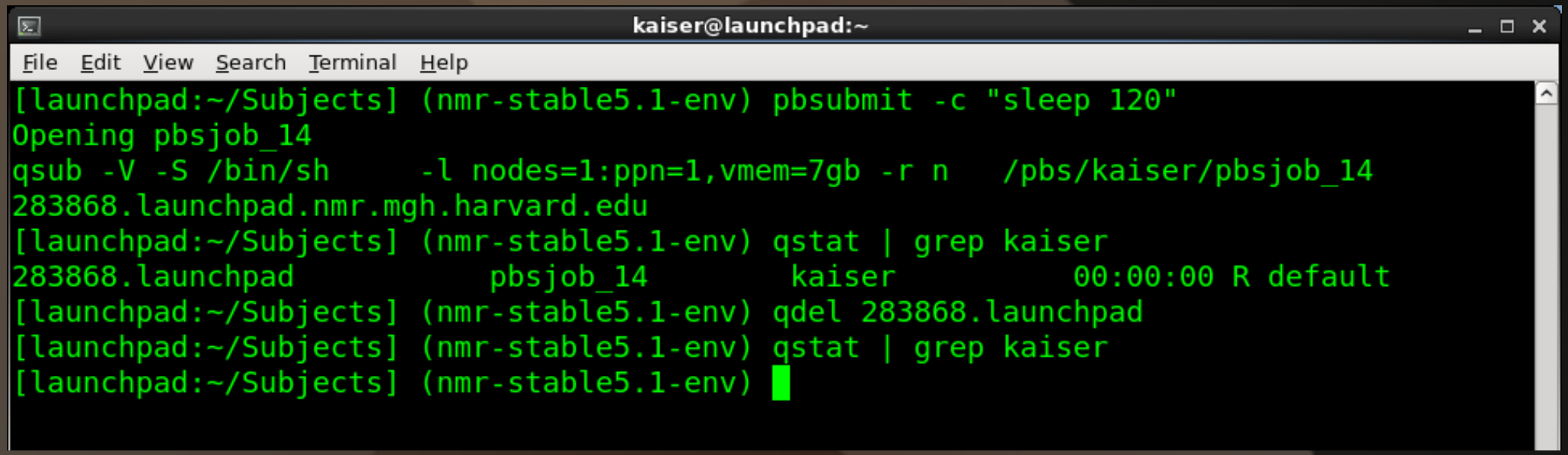
Resource Related
- Check vmem is under the requested amount (default: 7GB)
- Check walltime is under the requested amount (default: 96 hours)

Command Related
- Check standard error and standard output files!!
- If the program is home-made, was it compiled for the launchpad architecture?
- Test-run the command locally. If it breaks, the problem is probably not with the cluster.

Martinos Center Compute Clusters

Job Status - Delete Jobs

You submit a job, realize there is a mistake and want to delete it:



Martinos Center Compute Clusters

Adios