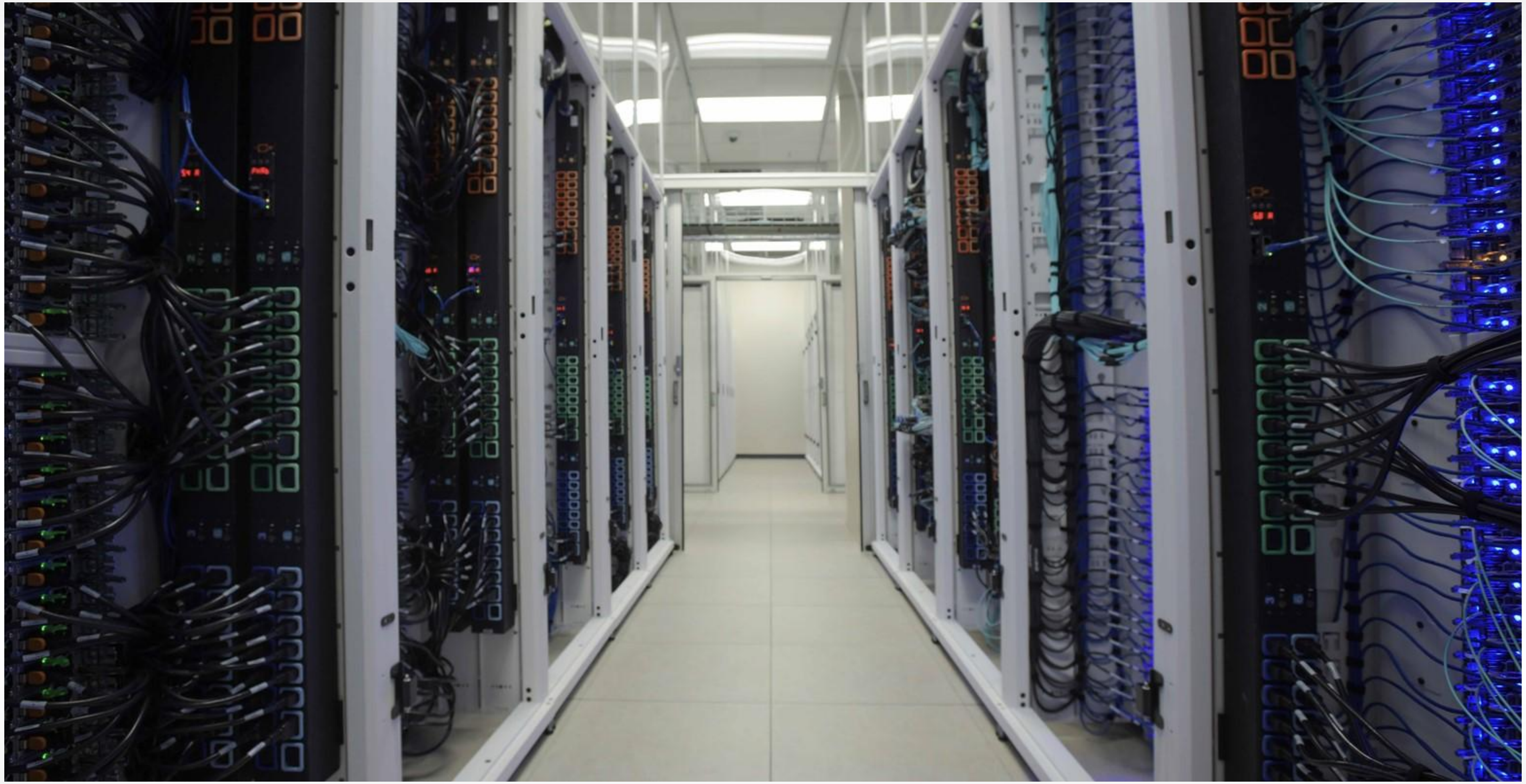
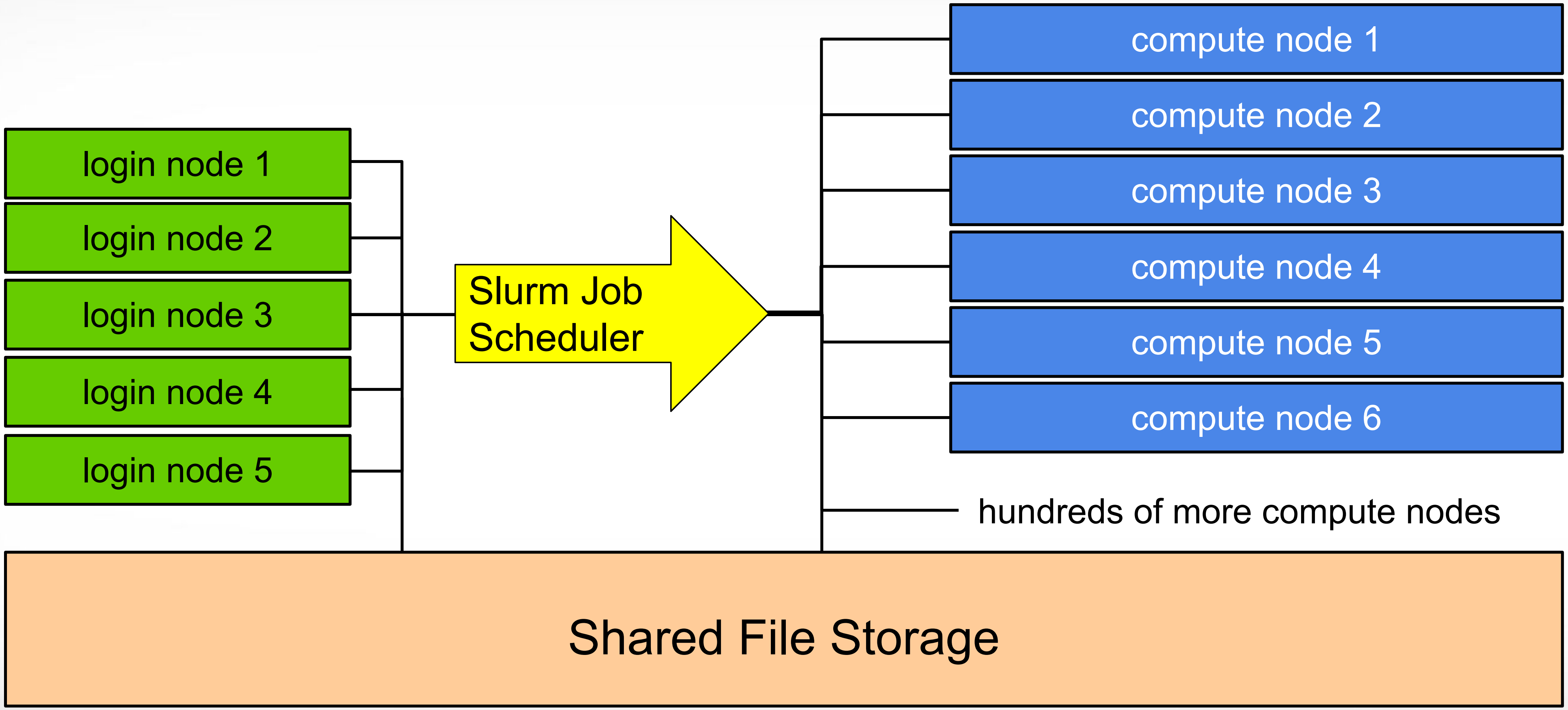
Introduction to Grace



HIGH PERFORMANCE RESEARCH COMPUTING
TEXAS A&M UNIVERSITY

Spring 2022

HPC Diagram



Grace Compute nodes on HPRC clusters are not connected to the internet but you can connect using 'module load WebProxy'

Grace Clusters Specs

	Grace
Compute Nodes	925
Total Cores	44,654
64GB memory nodes	-
384 GB memory nodes	800 (48 cores/node @ 3.00 GHz)
1TB memory nodes	-
2TB memory nodes	-
3TB memory nodes	8 (80 cores/node)
GPU nodes	117
login nodes	5
Production Date	Spring 2021

https://hprc.tamu.edu/resources

Accessing Grace

- SSH command is required for accessing Grace:
  - On campus:
    - `ssh -X netid@grace.hprc.tamu.edu`
    - `-X` enables viewing of image files and GUI applications
  - Off campus:
    - Set up and start VPN (Virtual Private Network):
      - http://u.tamu.edu/VPnetwork
    - Then: `ssh -X netid@grace.hprc.tamu.edu`
- SSH programs for Windows:
    - MobaXTerm (preferred, includes SSH and X11)
- Access through https://portal.hprc.tamu.edu
  - Menu “Clusters” => “Grace Shell Access”
- Grace has 5 login nodes. Check the bash prompt to see which you log into.
  -
    ```
    [netid@grace2 ~]$
    ```

https://hprc.tamu.edu/wiki/HPRC:Access

Grace Usage

- Login sessions that are idle for **60** minutes will be closed automatically
- Processes run longer than **60** minutes on login nodes will be stopped.
  - use data transfer nodes for large copy jobs
- **Do not use more than 8 cores on the login nodes!**
- Contact us for assistance installing software or install it yourself in your $SCRATCH directory.
  - **Do not use the sudo command.**
- HPRC applications must be renewed each fiscal year which ends August 31 each year.
  - https://hprc.tamu.edu/apply https://hprc.tamu.edu/policies
  - SUs expire at the end of each fiscal year
    - if you do not renew, you can still access your files on the clusters but you cannot submit any jobs
- Do not share your HPRC account with anyone

File Transfers with Grace

- Simple File Transfers:
  - scp: command line   (Linux, MacOS)
  - rsync: command line (Linux, MacOS); ***can resume transfer***
  - MobaXterm: GUI (Windows)
  - WinSCP: GUI (Windows)
  - Portal: https://portal.hprc.tamu.edu
- Use data transfer nodes for bulk data transfers (many large files):
  - scp and rsync data transfer processes will not timeout at 60 minutes
    - **`ssh grace-dtn1.hprc.tamu.edu`**
      `ssh grace-dtn2.hprc.tamu.edu`
  - Globus Connect
    - https://hprc.tamu.edu/wiki/SW:GlobusConnect
  - GridFTP

https://hprc.tamu.edu/wiki/HPRC:FileTransfers

File Systems and User Directories

Directory	Environment Variable	Space Limit	File Limit	Intended Use
/home/$USER	$HOME	10 GB	10,000	Small to modest amounts of processing.
/scratch/user/$USER	$SCRATCH	1 TB	250,000	Temporary storage of files for on-going computations. Not intended to be a long-term storage area.

- **`$HOME`** and **`$SCRATCH`** directories are not shared between HPRC clusters.
- View usage and quota limits using the command: **`showquota`**
- Quota and file limit increases will only be considered for scratch directory
- Request a group directory for sharing files.
- Do not share your home, or scratch directory.

Software

- See the Software wiki page for instructions and examples
  - https://hprc.tamu.edu/wiki/SW
  - https://hprc.tamu.edu/software/grace
  - https://hprc.tamu.edu/wiki/Bioinformatics
- License-restricted software
  - Contact license owner for approval
- Contact us for software installation help/request
  - User can install software in their $SCRATCH dir
  - **Do not use "*sudo*" command when installing software**

Software Modules

- Installed software applications are available as modules which are available to all users
  - (except for restricted modules)
- It's a good habit to purge unused modules before loading new modules.
- It is highly recommended to load a specific software version instead of the defaults
- **Avoid loading modules in your `~/.bashrc`**

```
module avail              # list all available modules (sometimes it is very slow)
                          # space bar down, page up/down, q to quit
                          # / for case sensitive search (similar to a UNIX man page)

module spider trinity     # case insensitive search for modules with ‘boost’ in name

module keyword assembly   # search module descriptions for keyword "assembly"
                          # some assembly software modules may be missed if
                          # keyword is not found in description (case insensitive)
```

Computing Environment

- Almost all software, applications, libraries, etc. are available as a module.
  - Module names have the format:

software_name / version - [Python-version]

`Trinity/2.12.0-Python-3.8.2`

- Some software is installed as an Anaconda environment

```
module load Anaconda3/2020.07
conda env list
source activate qiime2-2021.2
```

- Grace uses a modulefile hierarchy so library dependencies must be loaded prior to loading the software module

Module Loading Exercise

1.
```
module list                        # list all loaded modules
```

2.
```
module spider trinity              # search for available module names matching trinity
                                   # not case sensitive unless an exact match is found
```

3.
```
module spider Trinity/2.12.0-Python-3.8.2       # search a specific module version
```

```
          # see message about what additional modules need to be loaded first
You will need to load all module(s) on any one of the lines below before the "Trinity/2.12.0-Python-3.8.2" module is available to load.
      GCC/9.3.0  OpenMPI/4.0.3
```

4.
```
module load GCC/9.3.0  OpenMPI/4.0.3  Trinity/2.12.0-Python-3.8.2
```

5.
```
module avail bwa                   # show software with toolchain compatible with currently
                                   # loaded modules
```

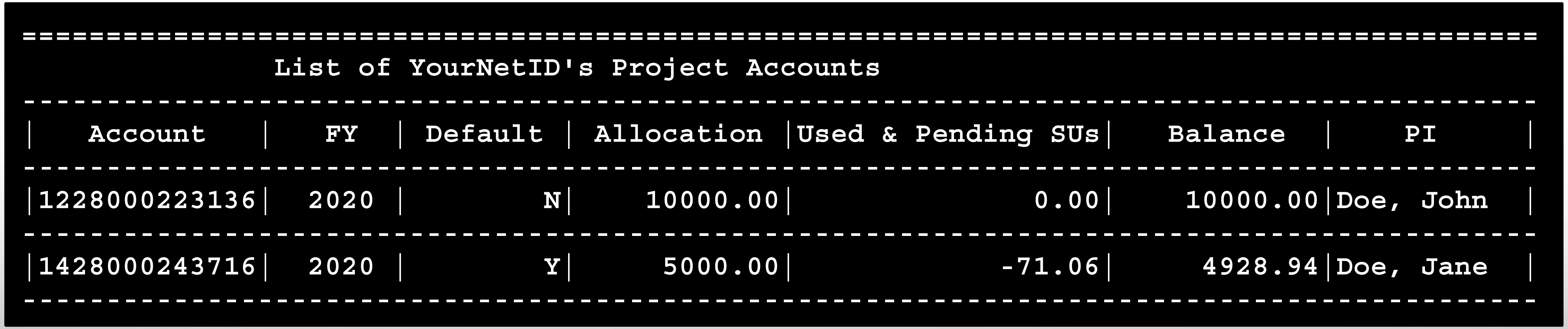
6.
```
module load BWA/0.7.17             # load a toolchain-compatible module
```

7.
```
module purge                       # unload all modules
```

Computing Resources

- Resources specified in a job file:
  - Processor cores, memory, wall time
- Service Unit (SU) - defined as one CPU core usage for one hour
  - Use "**`myproject`**" to see your balance
  - SUs expire at the end of each fiscal year; Aug 31 and must be renewed https://hprc.tamu.edu/apply

```
myproject
```



```
==========================================================================================
               List of YourNetID's Project Accounts
------------------------------------------------------------------------------------------
|   Account   |   FY  | Default | Allocation |Used & Pending SUs|   Balance  |    PI     |
------------------------------------------------------------------------------------------
|1228000223136|  2020 |        N|    10000.00|              0.00|    10000.00|Doe, John  |
------------------------------------------------------------------------------------------
|1428000243716|  2020 |        Y|     5000.00|            -71.06|     4928.94|Doe, Jane  |
------------------------------------------------------------------------------------------
```

https://hprc.tamu.edu/wiki/HPRC:AMS:Service_Unit

Grace: Examples of SUs charged based on Job Cores, Time and Memory Requested

A Service Unit (SU) on Grace is equivalent to one core or **7.5** GB memory usage for one hour.

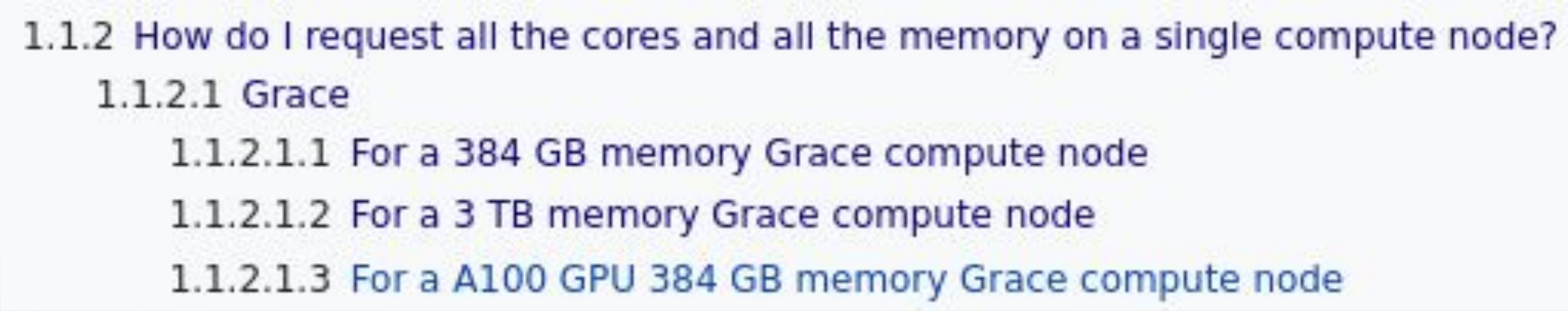
	Number of Cores	Total Memory (GB)	Hours	SUs charged
1.	1	7.5	1	1
2.	1	7.6	1	2
3.	1	360	1	48
4.	48	360	1	48

– SUs are charged at job submit time and the job's unused SUs are reimbursed if job finishes early

https://hprc.tamu.edu/wiki/HPRC:AMS:Service_Unit

See the HPRC Wiki on how to request all cores and all memory on a compute node

- Use all cores if the software you are using supports multi-core unless the software recommends fewer cores
- Use all cores if you are using all the memory even if the software only supports running on a single core
- When your job is complete, use the 'seff jobid' command to see job efficiency and use fewer cores and memory next time if efficiency was low.

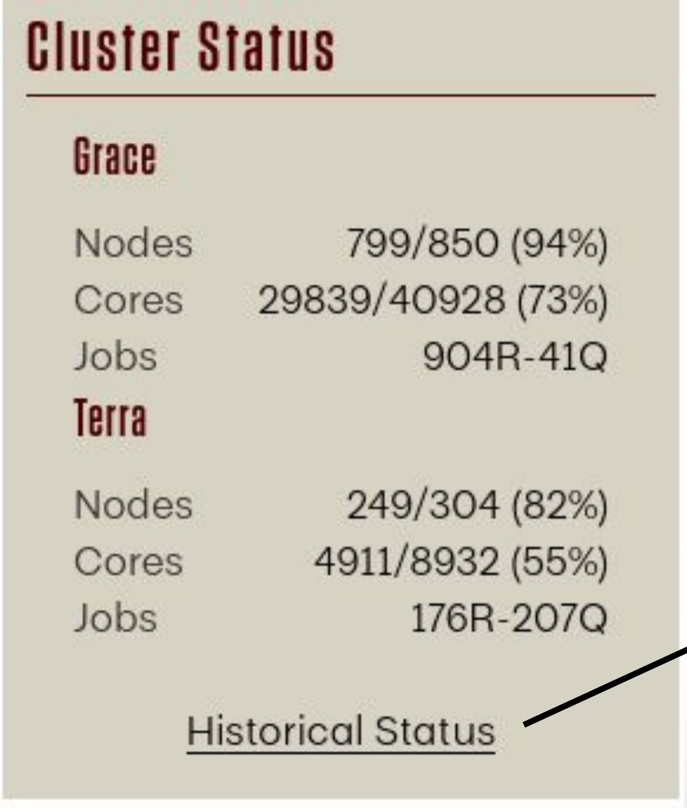


1.1.2 How do I request all the cores and all the memory on a single compute node?
- 1.1.2.1 Grace
  - 1.1.2.1.1 For a 384 GB memory Grace compute node
  - 1.1.2.1.2 For a 3 TB memory Grace compute node
  - 1.1.2.1.3 For a A100 GPU 384 GB memory Grace compute node

https://hprc.tamu.edu/wiki/Bioinformatics:FAQ

Historical HPRC Cluster Usage

hprc.tamu.edu



Cluster Status

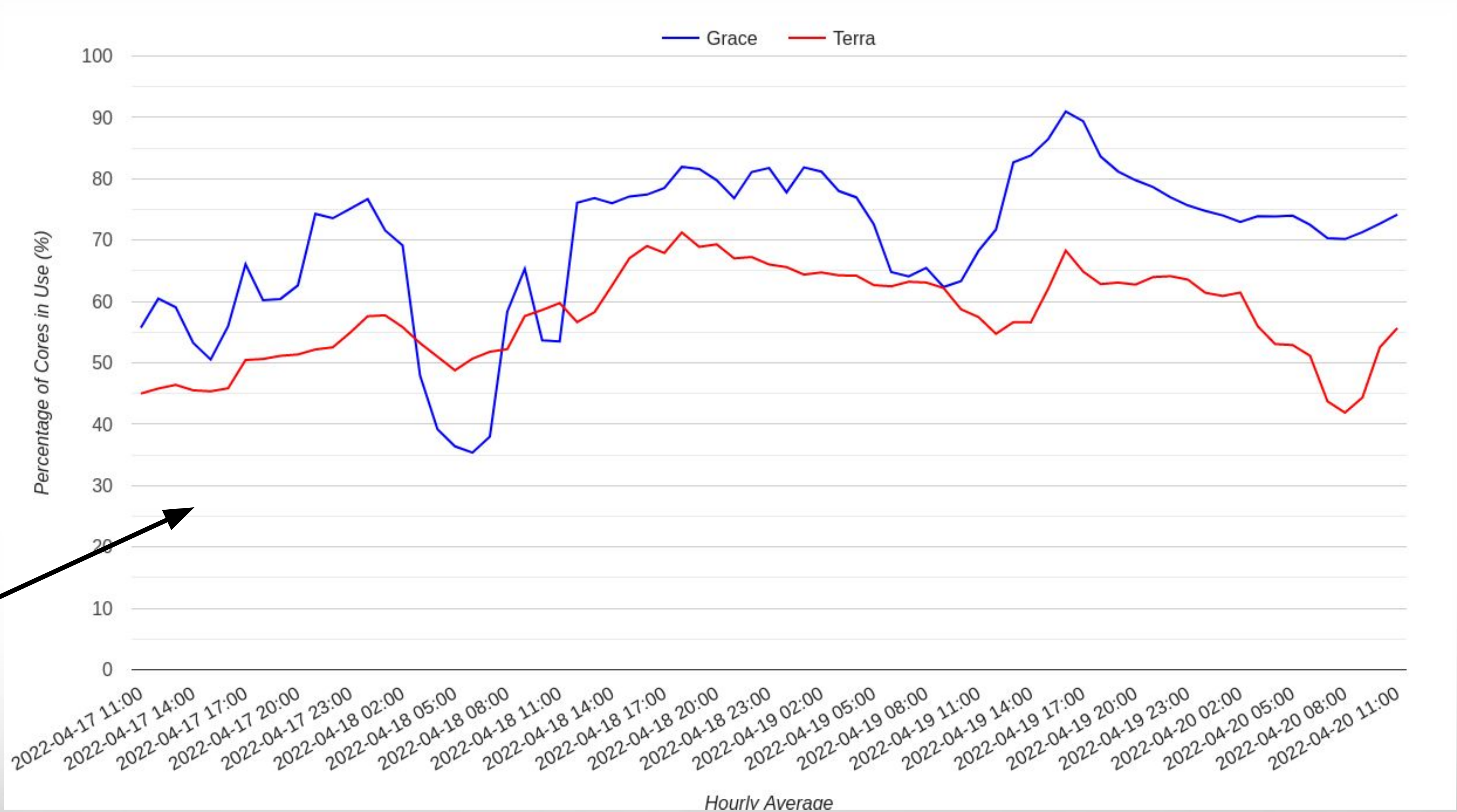
Grace

Nodes	799/850 (94%)
Cores	29839/40928 (73%)
Jobs	904R-41Q

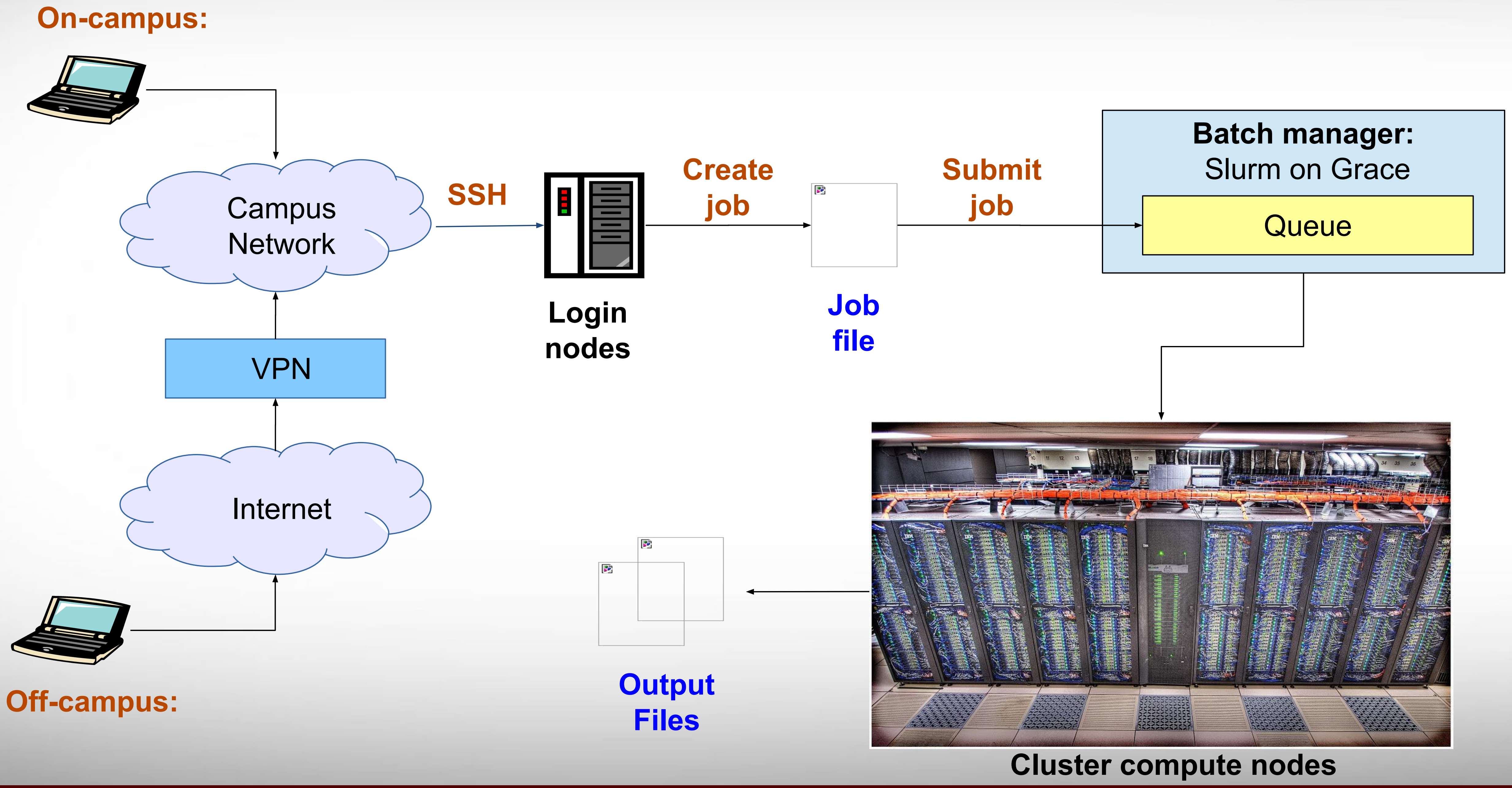
Terra

Nodes	249/304 (82%)
Cores	4911/8932 (55%)
Jobs	176R-207Q

Historical Status



Batch Computing on HPRC Clusters



On-campus:

Campus Network

SSH

Login nodes

Create job

Job file

Submit job

Batch manager:
Slurm on Grace

Queue

VPN

Internet

Off-campus:

Output Files

Cluster compute nodes

Slurm Job Scheduling

https://hprc.tamu.edu/files/training/2022/Spring/Slurm_Job_Scheduling_2022_spring.pdf

Batch Job Scripts

Sample Slurm Job Script Structure

```
# NECESSARY JOB SPECIFICATIONS
#!/bin/bash
#SBATCH --export=NONE                # do not export current env to the job
#SBATCH --job-name=my_job            # job name
#SBATCH --time=7-00:00:00            # max job run time dd-hh:mm:ss
#SBATCH --ntasks-per-node=1          # tasks (commands) per compute node
#SBATCH --cpus-per-task=48           # CPUs (threads) per command
#SBATCH --mem=360G                   # total memory per node
#SBATCH --output=stdout.%j           # save stdout to file
#SBATCH --error=stderr.%j            # save stderr to file


# OPTIONAL JOB SPECIFICATIONS
#BSUB -P 123456
#BSUB -u email_address
#BSUB -B -N

# load required module(s)
module load  GCC/9.3.0  OpenMPI/4.0.3  Trinity/2.12.0-Python-3.8.2

Trinity
```

These parameters describe your job to the job scheduler

This is single line comment and not run as part of the script

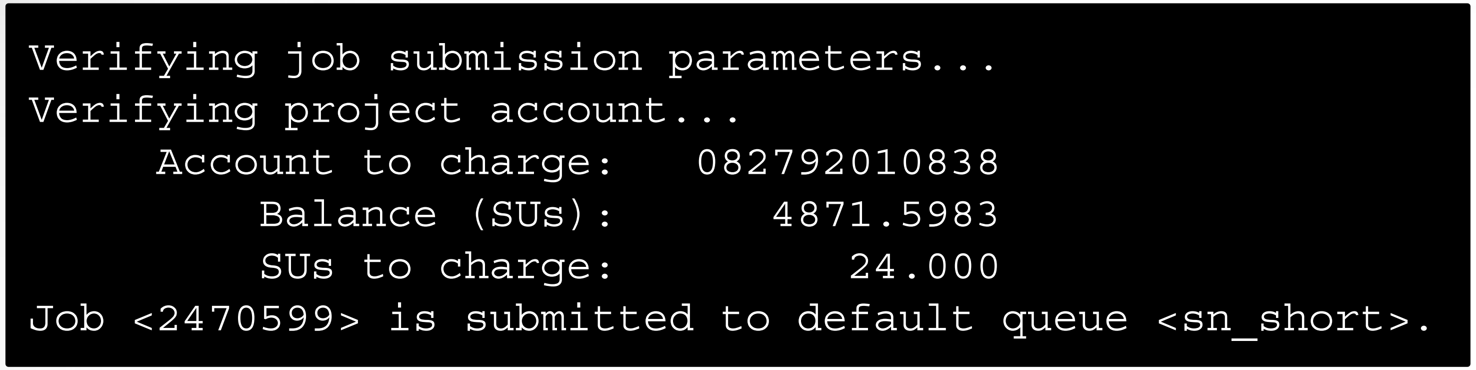
Load the required module(s) first

This is a command (task) that is executed by the job

Submitting Your Job and Check Job Status

Submit job

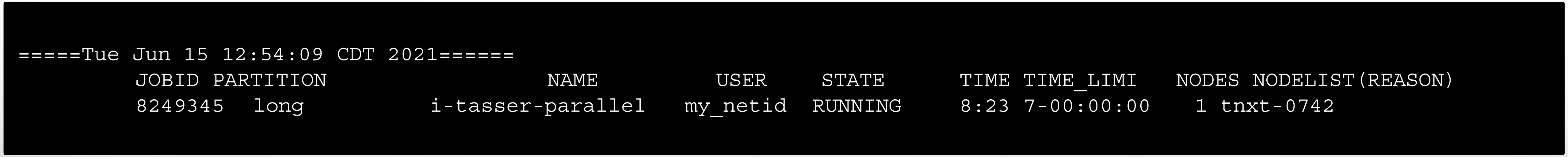
`sbatch run_trinity_grace.sh`



```
Verifying job submission parameters...
Verifying project account...
     Account to charge:   082792010838
         Balance (SUs):      4871.5983
         SUs to charge:         24.000
Job <2470599> is submitted to default queue <sn_short>.
```

Check status

`squeue -u $USER`

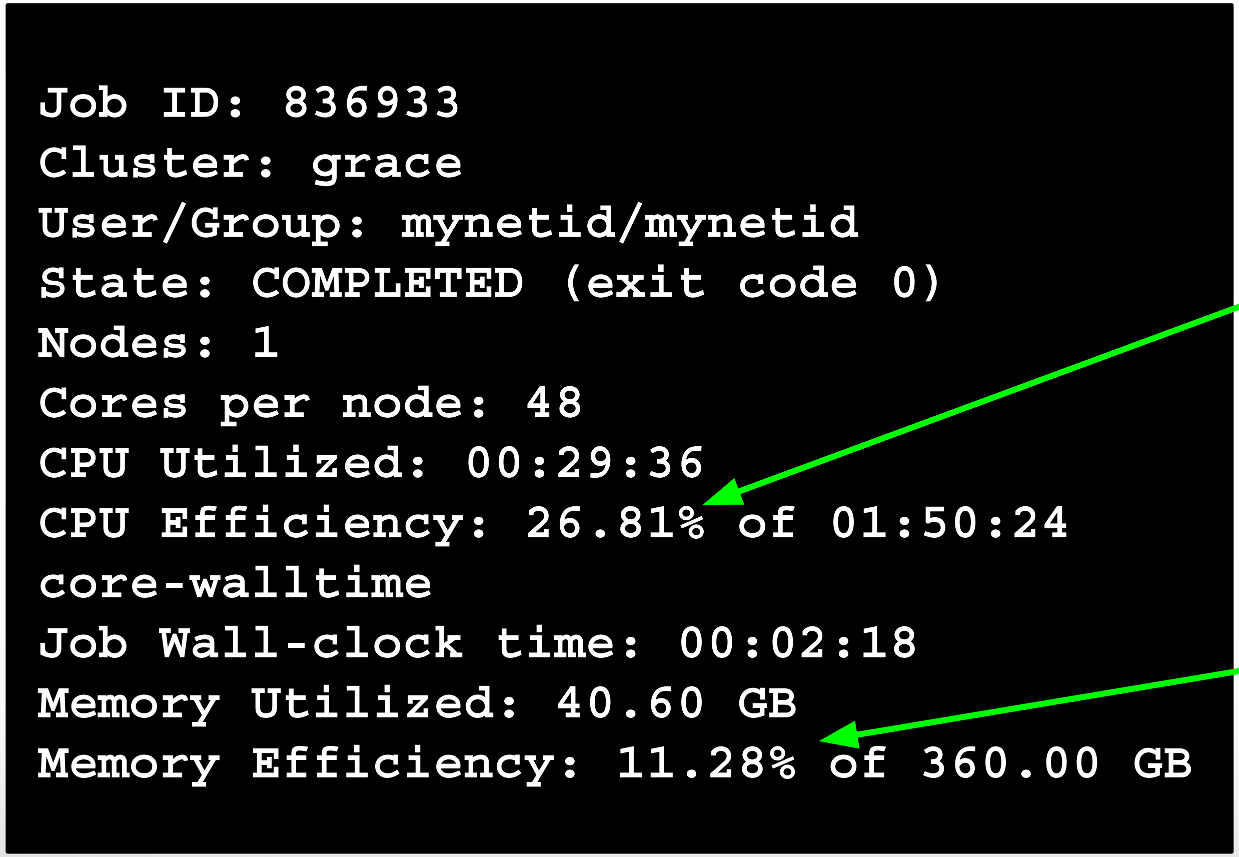


```
=====Tue Jun 15 12:54:09 CDT 2021======
         JOBID PARTITION                 NAME          USER    STATE      TIME TIME_LIMI   NODES NODELIST(REASON)
         8249345  long          i-tasser-parallel   my_netid  RUNNING     8:23 7-00:00:00   1 tnxt-0742
```

https://hprc.tamu.edu/files/training/2021/Spring/Slurm_Job_Scheduling_2021_spring.pdf

See Completed Job Efficiency Stats

- **`seff JOBID`**
  - will show CPU and Memory efficiency based on selected resources



```
Job ID: 836933
Cluster: grace
User/Group: mynetid/mynetid
State: COMPLETED (exit code 0)
Nodes: 1
Cores per node: 48
CPU Utilized: 00:29:36
CPU Efficiency: 26.81% of 01:50:24
core-walltime
Job Wall-clock time: 00:02:18
Memory Utilized: 40.60 GB
Memory Efficiency: 11.28% of 360.00 GB
```

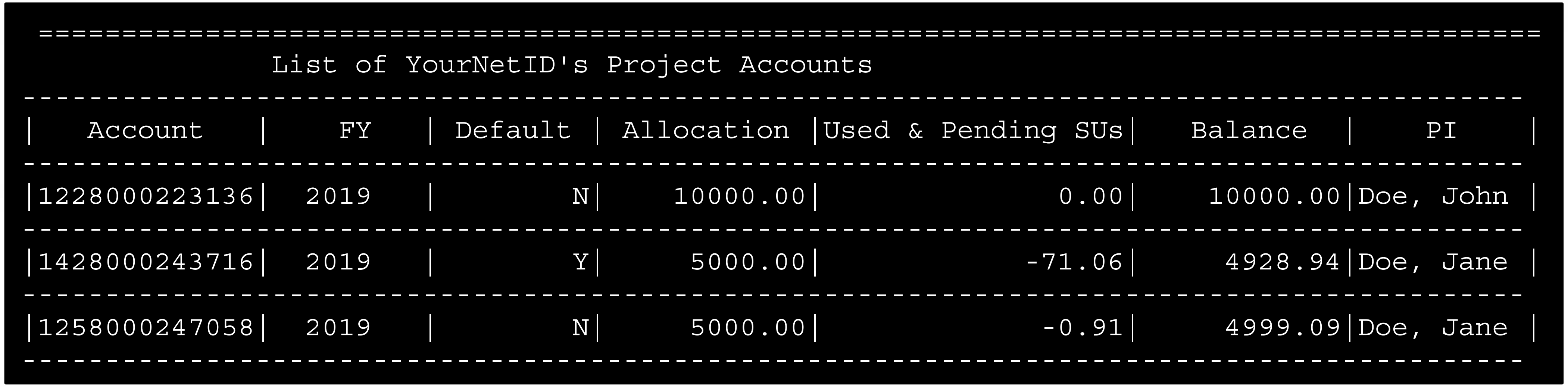
CPU load was at 100% for 27% of the run time or CPU load at 27% for 100% of the run time

max memory utilized was 11% of requested memory

Check your Service Unit (SU) Balance

- List the SU Balance of your Account(s)

```
myproject
```

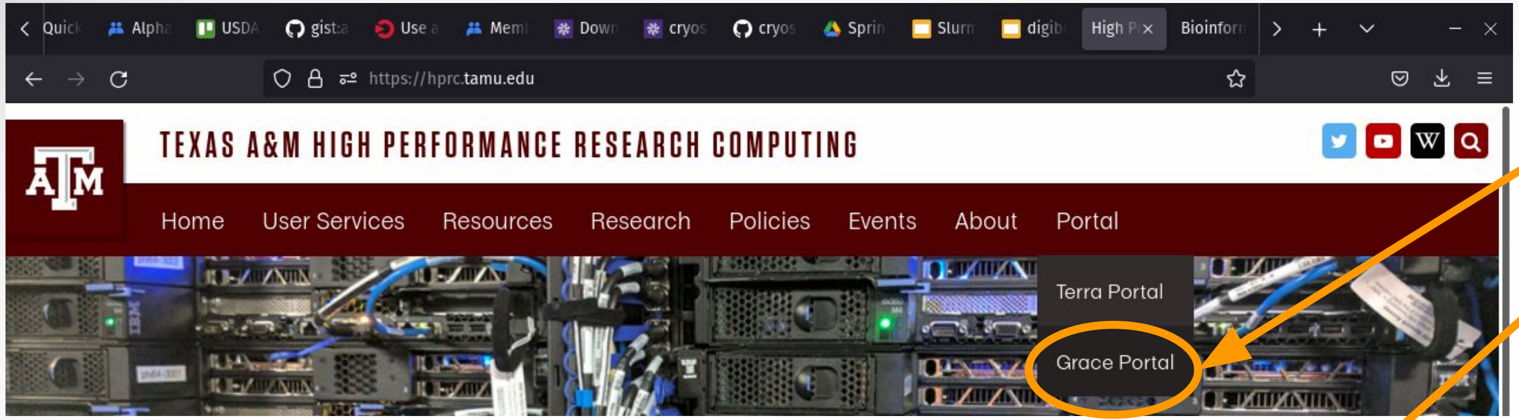


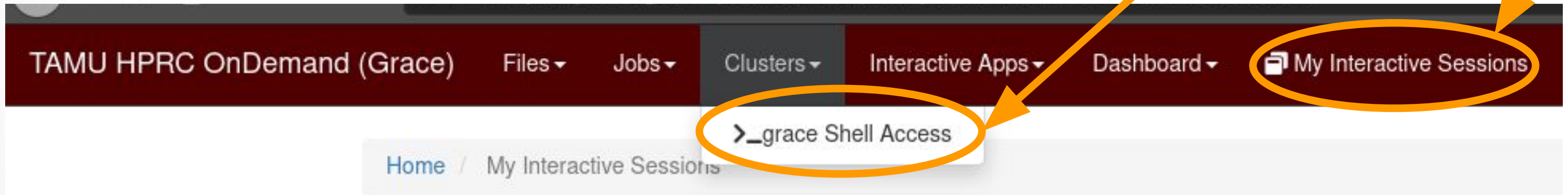
```
=========================================================================================
              List of YourNetID's Project Accounts
----------------------------------------------------------------------------------------
|   Account    |    FY    | Default | Allocation |Used & Pending SUs|   Balance   |    PI     |
----------------------------------------------------------------------------------------
|1228000223136|  2019    |        N|    10000.00|              0.00|     10000.00|Doe, John  |
----------------------------------------------------------------------------------------
|1428000243716|  2019    |        Y|     5000.00|            -71.06|      4928.94|Doe, Jane  |
----------------------------------------------------------------------------------------
|1258000247058|  2019    |        N|     5000.00|             -0.91|      4999.09|Doe, Jane  |
----------------------------------------------------------------------------------------
```

- To specify a project ID to charge in the job file
  - `#BSUB -P Account#`
- Run "**`myproject -d Account#`**" to change default project account
- Run "**`myproject -h`**" to see more options

https://hprc.tamu.edu/wiki/HPRC:AMS:Service_Unit
https://hprc.tamu.edu/wiki/HPRC:AMS:UI

HPRC Portal





The HPRC portal allows users to do the following

- Browse files on the Grace filesystem
- Access the Grace UNIX command line
  - no SUs charged for using command line
- Launch jobs
  - SUs charged
- Compose job scripts
- Launch interactive GUI apps (SUs charged)
  - Grace
    - IGV
    - MATLAB
    - VNC
    - Jupyter Notebook
    - JuypterLab
    - RStudio
- Monitor and stop running jobs and interactive sessions

HPRC Cluster Usage Notes

- Can test software on the login nodes but do not use more than 8 cores
  - processes get automatically killed on login node when they reach 60 minutes
  - login session gets automatically disconnected when idle for 60 minutes
- Transfer large data between HPRC clusters using the login nodes
  - transfer large data between HPRC clusters and sites off campus using the data transfer nodes
- Carefully read the stdout and stderr files for completed and failed jobs
  - most causes of failed jobs can be found in the log files
- Search the wiki and look at the FAQ pages for a solution before sending a helpdesk request
  - https://hprc.tamu.edu/wiki/HPRC:CommonProblems
  - https://hprc.tamu.edu/wiki/Bioinformatics:FAQ

Need Help?

- First check the FAQ https://hprc.tamu.edu/wiki/HPRC:CommonProblems
  - Grace User Guide https://hprc.tamu.edu/wiki/Grace
- Email your questions to **help@hprc.tamu.edu**.
  - Managed by a ticketing system
- Help us, help you -- we need more info
  - Which Cluster
  - UserID/NetID (*UIN is not needed!*)
  - Job id(s) if any
  - Location of your jobfile, input/output files
  - Software module(s) used if any
  - Error messages
  - Steps you have taken, so we can reproduce the problem
- Or visit us @ 114A Henderson Hall
  - Making an appointment is recommended.

HIGH PERFORMANCE RESEARCH COMPUTING

TEXAS A&M UNIVERSITY

Thank you.

Any question?