

HIGH PERFORMANCE RESEARCH COMPUTING

HPRC Primer

Using the Slurm Scheduler on the ACES Cluster

February 06, 2024

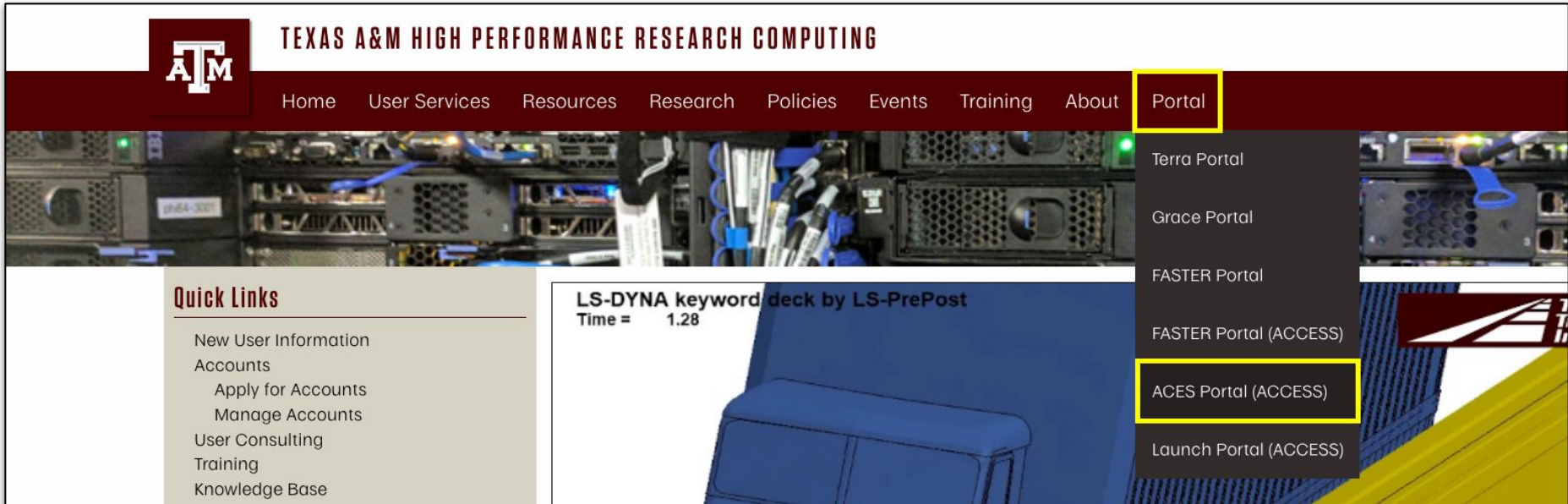


High Performance
Research Computing

DIVISION OF RESEARCH

Accessing the ACES Portal

- HPRC webpage: hprc.tamu.edu
- Aces portal shortcut: portal-aces.hprc.tamu.edu
- Requires ACCESS ID!



The screenshot shows the Texas A&M High Performance Research Computing website. The header features the ATM logo and the text "TEXAS A&M HIGH PERFORMANCE RESEARCH COMPUTING". The navigation menu includes links for Home, User Services, Resources, Research, Policies, Events, Training, About, and Portal. The Portal link is highlighted with a yellow box. A dropdown menu is open under Portal, listing Terra Portal, Grace Portal, FASTER Portal, FASTER Portal (ACCESS), ACES Portal (ACCESS) (highlighted with a yellow box), and Launch Portal (ACCESS). The main content area includes a "Quick Links" section with links for New User Information, Accounts, User Consulting, Training, and Knowledge Base. A "LS-DYNA keyword deck by LS-PrePost" section is also visible, showing a time of 1.28. The background of the website features images of server racks and a blue chair.

Accessing ACES via the HPRC Portal (ACCESS)

Navigate to portal-aces.hprc.tamu.edu to get to the ACCESS CILogon OpenID Connect page.

Log-in using your ACCESS credentials.

The screenshot shows the 'Consent to Attribute Release' section of the ACCESS portal. It lists the information requested: CILogon user identifier, name, email address, and username/affiliation. Below this is a 'Select an Identity Provider' dropdown menu with 'ACCESS CI (XSEDE)' selected. A yellow box highlights this dropdown. A 'Log On' button is visible below the dropdown. At the bottom, there is a footer with links for help and privacy policy.

The screenshot shows the 'Login to CILogon' page. It features input fields for 'ACCESS Username' and 'ACCESS Password', a 'Don't Remember Login' checkbox, and a 'Login' button. To the right is the CILogon logo and text: 'CILogon facilitates secure access to CyberInfrastructure (CI)'. Below the logo are links for XSEDE account users, registration, password recovery, and help.

A close-up of the 'Select an Identity Provider' dropdown menu, showing 'ACCESS CI (XSEDE)' as the selected option. A yellow box highlights the entire dropdown area.

Select the Identity Provider appropriate for your account.

Shell Access via the HPRC Portal

Once in the Portal,
select at the top:
“Clusters” → “aces Shell
Access”

- *shell* is also called *terminal* or *command line*

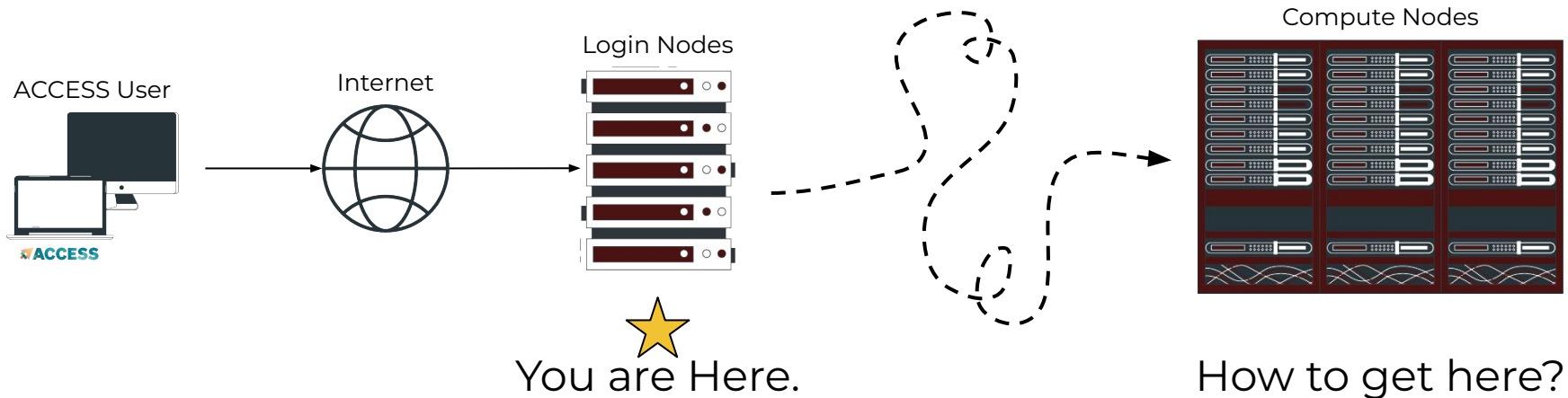


Hands-On Activity - 2 Minutes

- Connect to ACES now through the portal using portal-aces.hprc.tamu.edu
- Get a shell on the ACES cluster from the Clusters menu

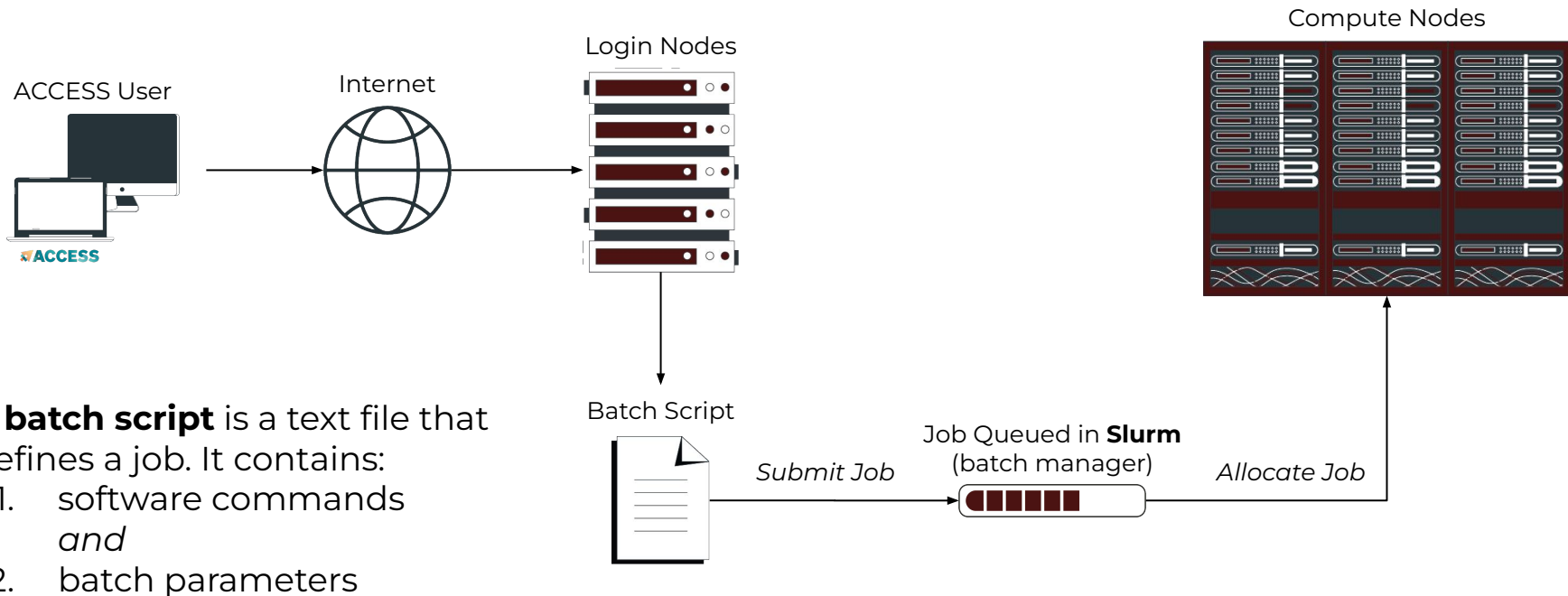
What is the hostname of the machine you connected to?

Batch Computing on the Clusters



- Types of nodes:
 - Login node - a shared machine for light editing
 - Compute node - an allocated machine for heavy computation

Batch Computing on ACES cluster



A **batch script** is a text file that defines a job. It contains:

1. software commands
and
2. batch parameters

Sample Job Script Structure

```
#!/bin/bash

##NECESSARY JOB SPECIFICATIONS
#SBATCH --job-name=hello_world
#SBATCH --time=00:15:00
#SBATCH --ntasks=2
#SBATCH --ntasks-per-node=2
#SBATCH --nodes=1
#SBATCH --mem=3G
#SBATCH --output=hello_world_log.%j

# load required module(s)
module purge
module load GCCcore/11.3.0
module load Python/3.10.4
python hello_world.py

# Job Environment variables
echo $SLURM_JOBID
echo $SLURM_SUBMIT_DIR
echo $TMPDIR
echo $SCRATCH
```

← This is a single-line comment and not run as part of the script.

These parameters describe your job to the job scheduler. The lines starting with #SBATCH are NOT comments! See the [Knowledge Base](#) for more info.

Whatever commands or scripts you want to run. Here, we set up the modules we need for our environment, run a python program, and print out some environment variables.

Important Batch Job Parameters

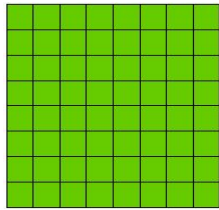
Slurm	Comment
<code>#SBATCH --time=HH:MM:SS</code>	Specifies the time limit for the job. Must specify seconds SS on ACES
<code>#SBATCH --ntasks=x</code>	Total number of tasks (cores) for the job.
<code>#SBATCH --ntasks-per-node=xx</code>	Specifies the maximum number of tasks (cores) to allocate per node
<code>#SBATCH --mem=xxxxM</code> or <code>#SBATCH --mem=xG</code>	Sets the maximum amount of memory (MB) per <i>node</i> . G for GB is supported on ACES
<code>#SBATCH --nodes=x</code>	Specifies the number of nodes to use

(These usually go in your job script file)

<https://hprc.tamu.edu/kb/Helpful-Pages/Batch-Translation/#job-specifications>

Mapping Jobs to Cores per Node on ACES

A.

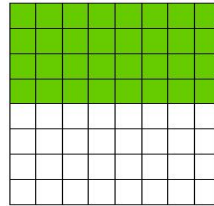


64 cores on
1 compute node

```
#SBATCH --ntasks=64  
#SBATCH --ntasks-per-node=64
```

Preferred Mapping
(if applicable)

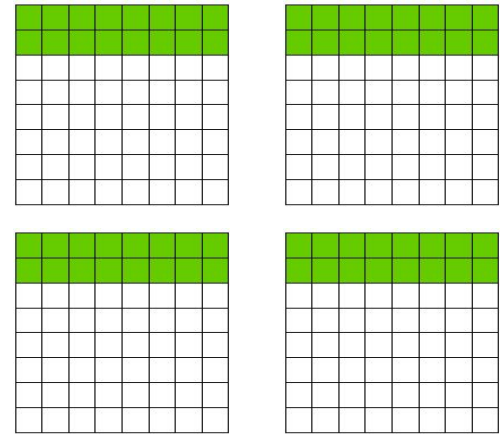
B.



64 cores on
2 compute nodes

```
#SBATCH --ntasks=64  
#SBATCH --ntasks-per-node=32
```

C.



64 cores on
4 compute nodes

```
#SBATCH --ntasks=64  
#SBATCH --ntasks-per-node=16
```

Pop Quiz

```
#SBATCH --job-name=JobExample2
#SBATCH --time=48:00:00
#SBATCH --ntasks=64
#SBATCH --ntasks-per-node=16
#SBATCH --nodes=4
#SBATCH --mem=48G
#SBATCH --output=stdout.%J
#SBATCH --error=stderr.%J
```

How many cores is this job requesting?

A. 1024

B. 64

C. 640

D. 4

Job Submission and Tracking

Slurm	Description
<code>sbatch jobfile</code>	Submit jobfile to batch system
<code>squeue [-u user_name] [-j job_id]</code>	List jobs
<code>scancel job_id</code>	Kill a job
<code>sacct -X -j job_id</code>	Show information for a job (can be when job is running or recently finished)
<code>sacct -X -S YYYY-MM-DD</code>	Show information for all of your jobs since YYYY-MM-DD

(These are command-line commands,
not parts of your job script)

<https://hprc.tamu.edu/kb/Helpful-Pages/Batch-Translation/#job-specifications>

Job Environment Variables

Each job has access to several self-referential variables:

- **\$SLURM_JOBID** = job id
- **\$SLURM_SUBMIT_DIR** = directory where job was submitted from
- **\$TMPDIR** = /work/job.\$SLURM_JOBID

You can also still use non-Slurm variables like:

- **\$SCRATCH** = /scratch/user/NetID

<https://hprc.tamu.edu/kb/Helpful-Pages/Batch-Translation/#environment-variables>

Submit a Job and Check Job Status

Submit job

```
sbatch example01.job
```

```
Submitted batch job 6853258
(from job_submit) your job is charged as below
  Project Account: 122792016265
  Account Balance: 1687.066160
  Requested SUs:   3
```

matching JOBID

Check status

```
squeue -u netID
```

JOBID	NAME	USER	PARTITION	NODES	CPUS	STATE	TIME	TIME_LEFT	START_TIME	REASON	NODELIST
6853258	jobname	someuser	xlong	2	96	RUNNING	3-07:36:50	16:23:10	2023-01-23T17:27:3	None	c [180,202]

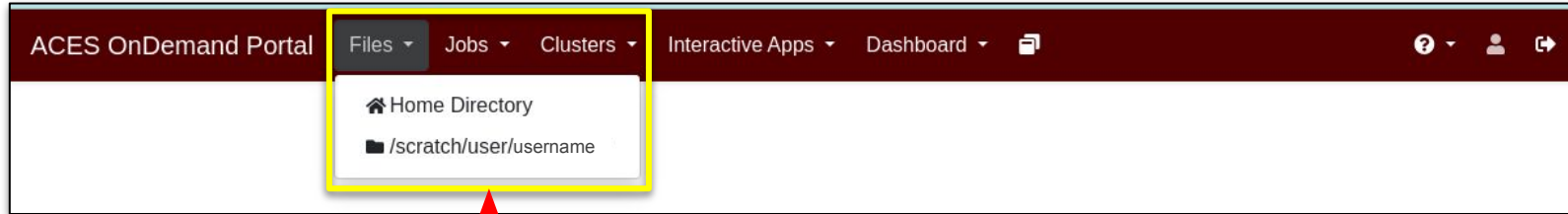
Hands-On Activity- 10 mins

There are example Files located at `/scratch/training/slurm_scheduler`

1. Copy these files to your home directory
2. Edit a batch file.
3. Submit a batch file using `sbatch`.
4. Check that the job is running in a Slurm queue with `squeue`.
5. Check the contents of the output file.

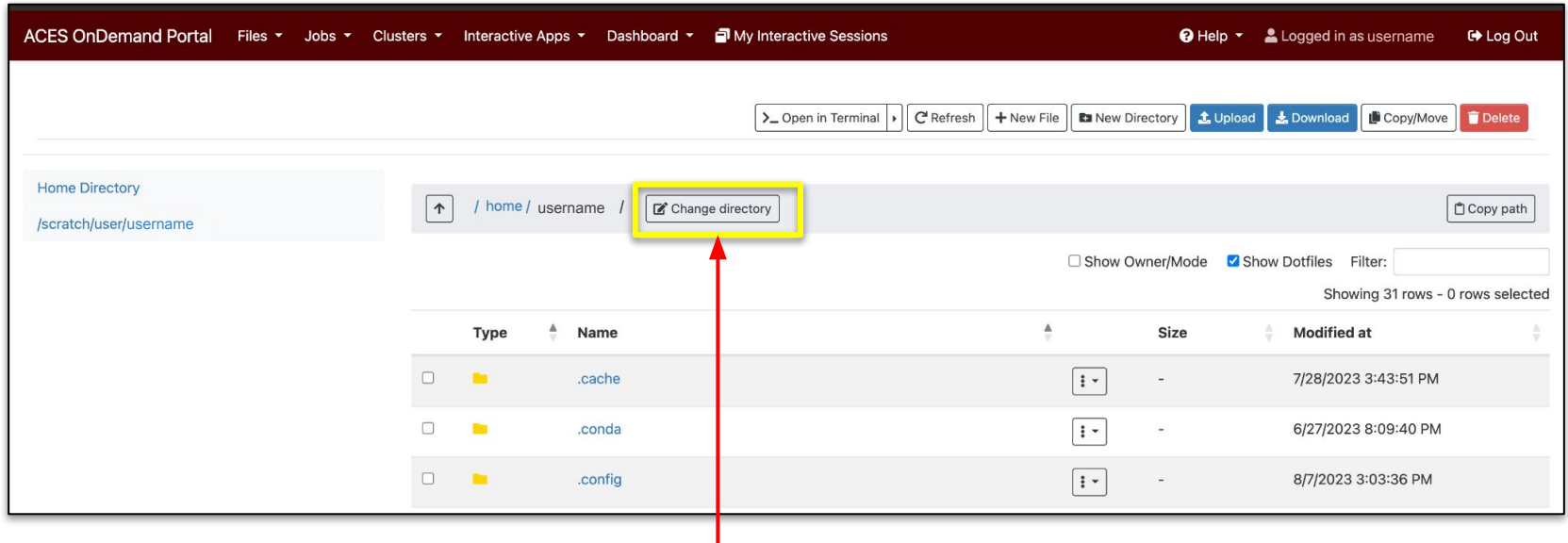
We are going to see above steps one-by-one in action in following slides.

Open the File Navigator



In the portal, choose one of the options under “Files”

Navigating to Training Directory



ACES OnDemand Portal Files Jobs Clusters Interactive Apps Dashboard My Interactive Sessions Help Logged in as username Log Out

> Open in Terminal Refresh + New File New Directory Upload Download Copy/Move Delete

Home Directory
/scratch/user/username

↑ / home / username / **Change directory** Copy path

Show Owner/Mode Show Dotfiles Filter:

Showing 31 rows - 0 rows selected

Type	Name	Size	Modified at
<input type="checkbox"/>	.cache	-	7/28/2023 3:43:51 PM
<input type="checkbox"/>	.conda	-	6/27/2023 8:09:40 PM
<input type="checkbox"/>	.config	-	8/7/2023 3:03:36 PM

Click on “change directory” and type
/scratch/training/slurm_scheduler

Copy the Example Files

ACES OnDemand Portal Files Jobs Clusters Interactive Apps Affinity Groups Dashboard My Interactive Sessions Help Logged in as username Log Out

>_ Open in Terminal Refresh + New File + New Directory Upload Download Copy/Move Delete

Home Directory
/scratch/user/username

↑ /scratch/training/slurm_scheduler/ Change directory Copy path

Show Owner/Mode Show Dotfiles Filter:

Showing 2 rows - 2 rows selected

Type	Name	Size	Modified at
✓ <input checked="" type="checkbox"/>	hello_world.py	73 Bytes	9/8/2023 3:54:45 PM
✓ <input checked="" type="checkbox"/>	hello_world.slurm	432 Bytes	9/8/2023 2:39:03 PM

Check box the two files and click Copy/Move

Return Home

The screenshot shows the ACES OnDemand Portal interface. At the top, there is a navigation bar with links for Files, Jobs, Clusters, Interactive Apps, Affinity Groups, Dashboard, and My Interactive Sessions. A user is logged in as 'username' and can log out. Below the navigation bar, there are several action buttons: Open in Terminal, Refresh, New File, New Directory, Upload, Download, Copy/Move, and Delete. The main area displays a file management dialog for copying or moving files from the directory `/scratch/training/slurm_scheduler` to the current directory. The dialog lists two files: `hello_world.py` and `hello_world.slurm`. Below the dialog, there is a 'Home Directory' button highlighted in yellow, with a red arrow pointing to it. The current directory path is `/scratch/user/username`. To the right of the dialog, there is a file list table showing the details of the selected files.

Type	Name	Size	Modified at
✓	hello_world.py	73 Bytes	9/8/2023 3:54:45 PM
✓	hello_world.slurm	432 Bytes	9/8/2023 2:39:03 PM

Select the directory to copy to

Paste the Example Files

ACES OnDemand Portal Files Jobs Clusters Interactive Apps Affinity Groups Dashboard My Interactive Sessions Help Logged in as username Log Out

Open in Terminal Refresh New File New Directory Upload Download Copy/Move Delete

Copy or move the files below from `/scratch/training/slurm_scheduler` to the current directory:

- hello_world.py
- hello_world.slurm

Copy Move

Home Directory
/scratch/user/username

/ home /username / Change directory Copy path

Show Owner/Mode Show Dotfiles Filter: Showing 2 of 28 rows - 0 rows selected

Type	Name	Size	Modified at
Directory	ACES_FundamentalsOfRProgramming	-	9/26/2023 10:56:01 AM
File	module.avail.aces	16 KB	9/18/2023 11:20:41 AM

The files on the right will change to show the directory you chose.
Hit “Copy” to actually copy the files to that directory.

View and Edit the Example Files

ACES OnDemand Portal Files Jobs Clusters Interactive Apps Dashboard

Open in Terminal Refresh New File New Directory Upload Download Copy/Move Delete

Home Directory
/scratch/user/username

/ home / username / Change directory Copy path

Show Owner/Mode Show Dotfiles Filter: Showing 2 of 21 rows - 0 rows selected

Type	Name	Size	Modified at
File	hello_world.py	73 Bytes	9/11/2023 10:49:24 AM
File	hello_world.slurm		9/11/2023 10:49:24 AM

View Edit Rename Download Delete

powered by OPEN OnDemand OnDemand version: 3.0.0

Open a Terminal (another option)

ACES OnDemand Portal Files Jobs Clusters Interactive Apps Affinity Groups Dashboard My Interactive Sessions Help Logged in as username Log Out

> Open in Terminal Refresh + New File New Directory Upload Download Copy/Move Delete

Home Directory /scratch/user/username

↑ / scratch / training / slurm_scheduler / Change directory Copy path

Show Owner/Mode Show Dotfiles Filter:

Showing 2 rows - 0 rows selected

Type	Name	Size	Modified at
<input type="checkbox"/>	hello_world.py	73 Bytes	9/8/2023 3:54:45 PM
<input type="checkbox"/>	hello_world.slurm	432 Bytes	9/8/2023 2:39:03 PM

Hands on Activity- Cont.

1. Investigate the two example files
 - Make a small edit to personalize.
(e.g., “Hello MyName”)
2. Submit the batch file using:
`sbatch hello_world.slurm`
3. Check the job status using:
`squeue -u $USER`
4. Once the job is completed, inspect the output file:
`hello_world_log.<job_id>`

Consumable Computing Resources

- Resources which we can specify in a job/slurm file:

- Processor cores
- Memory
- Wall time
- GPU

- Other resources:

- SUs
- Software license/token
 - Use `license status` to query
 - License Checker:
hprc.tamu.edu/kb/Software/useful-tools/License_Checker/

Find available license for "FEMZIP":

```
license_status -s FEMZIP
```

```
License status for FEMZIP:
```

```
-----  
|License Name          | # Issued| # In Use|# Available|  
-----  
|komp                  |      100|        0|        100|  
-----
```

More information on this command:

```
license_status -h
```

Key:

Your inputs in green

Terminal output in blue

How Does Slurm Assign Jobs?

- Job submissions are auto-assigned to batch *queues* (also called *partitions*) based on the resources requested
 - number of cores/nodes and walltime limit
 - specific resources requested
- Some jobs can be directly submitted to a queue:
 - If gpu nodes are needed, use the gpu partition/queue:
`#SBATCH --partition=gpu`

<https://hprc.tamu.edu/kb/User-Guides/Common/BatchProcessing/#batch-queues>

sinfo: Info for Node/Partition

To check the status of the nodes/partitions:

sinfo

PARTITION	AVAIL	TIMELIMIT	JOB_SIZE	NODES (A/I/O/T)	CPUS (A/I/O/T)
cpu*	up	7-00:00:00	1-64	17/48/13/78	1537/4703/1248/7488
gpu	up	2-00:00:00	1-15	14/1/2/17	1025/415/192/1632
atsp	up	2-00:00:00	1	0/2/2/4	0/192/192/384
bittware	up	2-00:00:00	1	0/0/2/2	0/0/192/192
d5005	up	2-00:00:00	1	0/2/0/2	0/192/0/192
memverge	up	2-00:00:00	1	0/5/3/8	0/480/288/768
nextsilicon	up	2-00:00:00	1	0/1/1/2	0/96/96/192
staff	up	2-00:00:00	1-110	31/59/20/110	2562/6078/1920/10560

For the NODES and CPUS columns:

A = Active (in use by running jobs)

I = Idle (available for jobs)

O = Offline (unavailable for jobs)

T = Total

pestat : Processor Status

- **pestat** allows you to check the status of the nodes on ACES
- -p allows you to show a specific partition
- Example:

```
pestat -p gpu -G
```

```
GPU GRES (Generic Resource) is printed after each jobid
Print only nodes in partition gpu
```

Hostname	Partition	Node State	Num CPU Use/Tot	CPUload (15min)	Memsize (MB)	Freemem (MB)	GRES/node	Joblist
ac036	gpu	alloc	72 96	2.95	500000	492233*	gpu:h100:2	<jobid>
ac037	gpu	idle	0 96	0.00	500000	479355	gpu:h100:2	
ac038	gpu	drain*	0 96	0.00	500000	489374	gpu:h100:2	
ac039	gpu	idle	0 96	0.00	500000	511650	gpu:h100:2	
ac040	gpu	idle	0 96	0.00	500000	463374	gpu:h100:2	
ac046	gpu	mix	0 96	0.00	500000	442789	gpu:h100:2	
ac047	gpu	alloc	96 96	1.97*	500000	489464	gpu:h100:2	<jobid>
ac064	gpu	idle	0 96	0.00	500000	509287	gpu:a30:6	

Job Memory Requests on ACES

- If you request more resources than is allowed, Slurm will reject the job
- To check the maximum requestable amount, use:

`maxconfig`

```
ACES partitions:  cpu  gpu  atsp  bittware  d5005  memverge  nextsilicon
ACES GPUs in gpu partition:  a30:6  h100:2
Showing max parameters (cores, mem, time) for partition cpu
```

```
#!/bin/bash
#SBATCH --job-name=my_job
#SBATCH --time=7-00:00:00
#SBATCH --nodes=1
#SBATCH --ntasks-per-node=1
#SBATCH --cpus-per-task=96
#SBATCH --mem=488G
#SBATCH --output=stdout.%x.%j
#SBATCH --error=stderr.%x.%j
```

Check your Service Unit (SU) Balance

- List the SU Balance of your Account(s) with:

```
myproject
```

```
=====
List of YourNetID's Project Accounts
-----
| Account | FY | Default | Allocation | Used & Pending SUs | Balance | PI |
-----
| 1228000223136 | 2023 | N | 10000.00 | 0.00 | 10000.00 | Doe, John |
-----
| 1428000243716 | 2023 | Y | 5000.00 | -71.06 | 4928.94 | Doe, Jane |
-----
| 1258000247058 | 2023 | N | 5000.00 | -0.91 | 4999.09 | Doe, Jane |
-----
```

- Run `myproject -d <Account#>` to change default project account
(replace <Account#> with your number!)
- Run `myproject -h` to see more options

https://hprc.tamu.edu/kb/User-Guides/AMS/Service_Unit/

<https://hprc.tamu.edu/kb/User-Guides/AMS/UI/>

Check your Service Unit (SU) Balance

The screenshot shows the ACES OnDemand Portal interface. The top navigation bar includes links for Files, Jobs, Clusters, Interactive Apps, and Dashboard. A dropdown menu for the Dashboard link is open, showing the ACES Dashboard option, which is highlighted by a red arrow. Below this, the ACES Dashboard page is displayed. It features a header with the ACES logo, the title 'ACES DASHBOARD', and buttons for 'Create Help Ticket' and 'Request Software'. The main content area is divided into two sections: 'CLUSTER STATISTICS' and 'SUMMARY'. The 'CLUSTER STATISTICS' section contains two donut charts for 'Node Utilization' and 'Core Utilization', with a legend indicating 'Allocated' (dark red), 'Mixed' (olive green), and 'Idle' (grey). The 'SUMMARY' section contains a table of account balances, with a red arrow pointing to the 'Balance' column. The table has the following data:

Account ↑↓	Default ↑↓	Allocation ↑↓	Used ↑↓	Balance ↑↓
155062417651	Set Default	20000	-0.53	19999.47

Below the table, there is a section for 'Disk Quotas'.

Planned Charging Scheme

Resource	Service Units (per hour)	ACCESS Credits (per hour)
Intel SPR / Icelake	1	0.1
NVIDIA H100	240	30
NVIDIA A30 and Intel PVC GPUs	120	15
Bittware Agilex FPGA	200	25
Intel D5005 FPGA	50	6.2
NEC Vector Engine	150	18.7
NextSilicon coprocessor	100	12.5
Graphcore IPU Classic	120	15
Graphcore IPU Bow	150	18.7
Intel Optane Memory	150	18.7

Continued Learning

[HPRC YouTube](#)

[HPRC Homepage](#)

[ACES Quick Start Guide](#)

[ACES Portal \(ACCESS users\)](#)

help@hprc.tamu.edu

Need Help?

First check the [FAQ](#)

- [ACES User Guide](#)
- Email your questions to help@hprc.tamu.edu

Help us help you -- we need more info

- Which Cluster
- Username
- Job id(s) if any
- Location of your jobfile, input/output files
- Application used if any
- Module(s) loaded if any
- Error messages
- Steps you have taken, so we can reproduce the problem



High Performance
Research Computing
DIVISION OF RESEARCH

Thank you.

Any questions?