HIGH PERFORMANCE RESEARCH COMPUTING

HPRC Primer

Introduction to FASTER and ACES: Composable Clusters

January 30, 2024

High Performance
Research Computing
DIVISION OF RESEARCH

FASTER System Description

[FASTER](#) (Fostering Accelerated Sciences Transformation Education and Research) is a 180-node Intel cluster from Dell with an InfiniBand HDR-100 interconnect and Liquid PCIe Gen4 for composing the GPUs. NVIDIA A100, A10, A30, A40 and T4 GPUs are available. The 180 compute nodes are based on the Intel Ice Lake processor.

Node Type			
	Login	**Compute**	**Large Memory**
Quantity	3 (2 for TAMU, 1 for ACCESS)	180	1
CPU	2 32-core (64 cores) Intel Xeon 8352Y	2 32-core (64 cores) Intel Xeon 8352Y	2 32-core (64 cores) Intel Xeon 8352Y
Memory	256 GB	256 GB	1 TB
Additional Cluster Specifications			
Disk Space	3.84 TB NVMe (local, /tmp) 5PB DDN Lustre (global)		
Composable Accelerators	**GPU**: 200 T4 16GB; 40 A100 40GB; 10 A10 24GB; 4 A30 24GB; 8 A40 48GB		
Interconnect	Mellanox HDR100 InfiniBand (MPI and storage) Liqid PCIe Gen4 (resource composability)		

https://hprc.tamu.edu/kb/Quick-Start/FASTER/

ACES System Description

[ACES](#) (Accelerating Computing for Emerging Sciences) is a Dell cluster with a rich accelerator testbed:

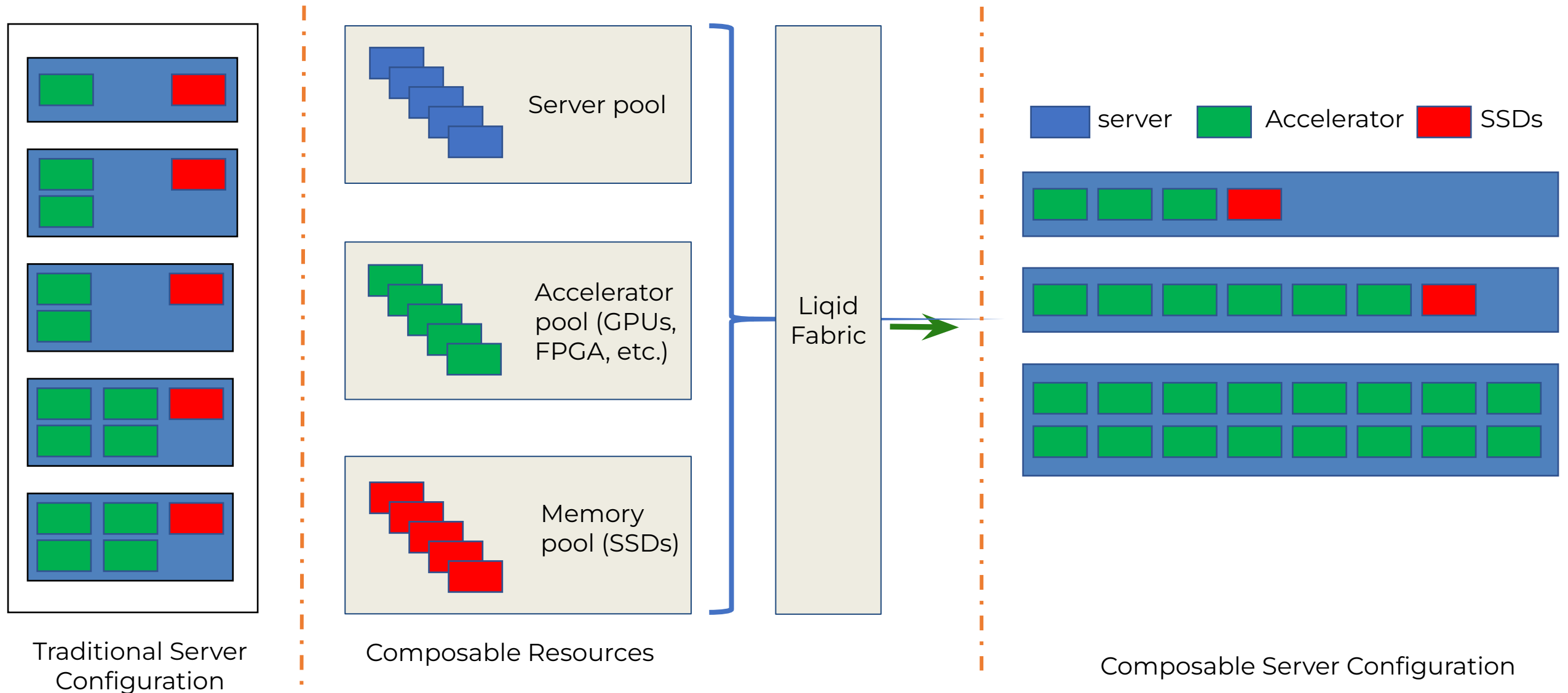|	
Compute nodes	• 110 Intel Sapphire Rapids Nodes • 17 Intel Ice Lake Nodes • 1 AMD Rome Graphcore Node with 16 Mk2 Colossus GC200 IPUs • 1 Intel Ice Lake Graphcore Node with 16 Bow-2000 IPUs • 1 Intel Cascade Lake Node with 8 NEC Vector Engine Type 20B-P cards
Accelerators	• 30 NVIDIA H100 GPUs (In 15 SPR nodes) • 2 NextSilicon coprocessors • 4 NVIDIA A30 GPUs • 48 Intel Optane Memory • 2 Bittware Agilex FPGAs • 120 Intel GPU Max 1100 • 2 Intel D5005 FPGAs (Formerly Ponte Vecchio)
Interconnect	NVIDIA Mellanox NDR200 InfiniBand (MPI and storage) Liqid PCIe Gen4 Fabrics (composability), Gen 5 Fabrics (Coming Soon)

The Intel Optane SSDs and all accelerators (except the Graphcore IPUs and NEC Vector Engines) are accessed using Liqid's composable framework via PCIe (Peripheral Component Interconnect express) Gen4 and Gen5 fabrics.

Visit https://hprc.tamu.edu/kb/User-Guides/ACES/ for more details.

Design: Composability at the Hardware Level

Traditional Server Configuration

Composable Resources

server   Accelerator   SSDs

Composable Server Configuration

Server pool

Accelerator pool (GPUs, FPGA, etc.)

Memory pool (SSDs)

Liqid Fabric

Pop Quiz

Which accelerators are available on FASTER?

A. T4   B. A100   C. A10

D. A40   E. A and B   F. All the above

Accessing FASTER/ACES via the HPRC Portal

- HPRC webpage: hprc.tamu.edu
  - TAMU: portal-faster.hprc.tamu.edu
  - ACCESS: portal-faster-access.hprc.tamu.edu

Using the Portal - Shell Access

Convenient shell access anywhere with a web browser

Starts an in-browser Linux terminal on FASTER

Accessing FASTER via SSH

- SSH (Secure SHell) allows users to establish a connection between their local machine and the TAMU HPRC clusters.
- SSH Programs

Operating System	Windows	MacOS	Linux
Programs	MobaXTerm* PuTTY SSH Windows Subsystem for Linux (WSL) Windows Command Prompt	Terminal*	Terminal*
	* Recommended		

- SSH Command `ssh [NetID]@faster.hprc.tamu.edu`

- TAMU (on-campus: Duo 2FA; off-campus: VPN and Duo 2FA)
- ACCESS (ssh not available) https://hprc.tamu.edu/kb/User-Guides/FASTER/ACCESS-CI/
- ACES cannot be accessed via SSH

Hands-On Activity - 2 Minutes

1. Please try to access FASTER or ACES through the portal.

2. What message do you see when you log on to the shell?

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 large files for on-going computations. Not intended to be a long-term storage area.

`$SCRATCH` is shared between FASTER and Grace clusters.

View file usage and quota limits using the command: `showquota`

Do NOT share your home or scratch directories. Request a group directory for sharing files.

https://hprc.tamu.edu/kb/User-Guides/FASTER/Filesystems_and_Files/

Using the Portal -Dashboard

View your account quota and projects

OOD Dashboard: FASTER

Easily view Cluster utilization, Storage Quotas & Allocation Balances

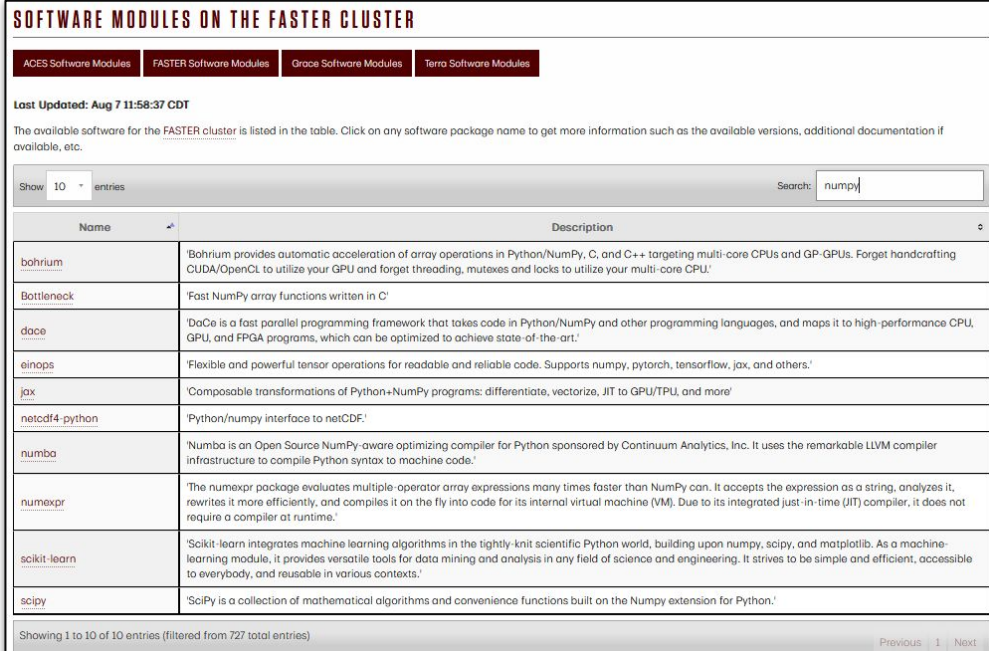Quota and file limit increases will only be considered for scratch directories

Preferred way to request *Quota Increases*

Hands-On Activity - 2 Minutes

1. Please try to access the dashboard now through the portal.

2. Where is the button to request a quota increase?

Software

- See the Software Knowledge Base page
  https://hprc.tamu.edu/kb/Software/
  for instructions and examples

- Search for software modules on
  https://hprc.tamu.edu/software/faster/

- License-restricted software
  - Contact help@hprc.tamu.edu

- Contact HPRC for software installation help/request
  - User can install software in their home/scratch directory
  - Do NOT run the `sudo` command when installing software

Software: Application Modules

- Installed applications are made available with the module system
- FASTER uses a software hierarchy inside the module system
- In this hierarchy, the user loads a compiler "toolchain" which then makes available software built with that compiler toolchain

```
module list                        # shows which software is ready

module load GCC/11.2.0             # load GCC compiler version 11.2.0

module list                        # show which software is now ready

module load OpenMPI/4.1.1          # load OpenMPI version 4.1.1

module list                        # show which software is now ready
```

Software: Modules and Toolchains

- Toolchains are what we call groups of compilers & libraries
- There is a variety of toolchains on the clusters:

```
intel/2020b
iomkl/2021a
foss/2022a
GCCcore/12.1.0
```

(more than just these versions)
(varies by cluster, as well)

- Module management:

```
module spider        # Search for modules and their dependencies
module purge         # removes all loaded modules from environment
```

Hands-On Activity - 2 Minutes

1. Search for and load the following module:

   `GCC/12.3.0`

2. Next remove all your current modules.

Batch Computing on HPRC Clusters

On-Campus User

Campus Network

Submit Job

SSH

Job Queued in SLURM (batch manager)

Off-Campus User

Internet

VPN

Login Nodes

Create Job

Job Script

Compute Nodes

ACCESS User

Internet

A batch job script is a text file that contains Unix and software commands and batch manager job parameters.

Output Files
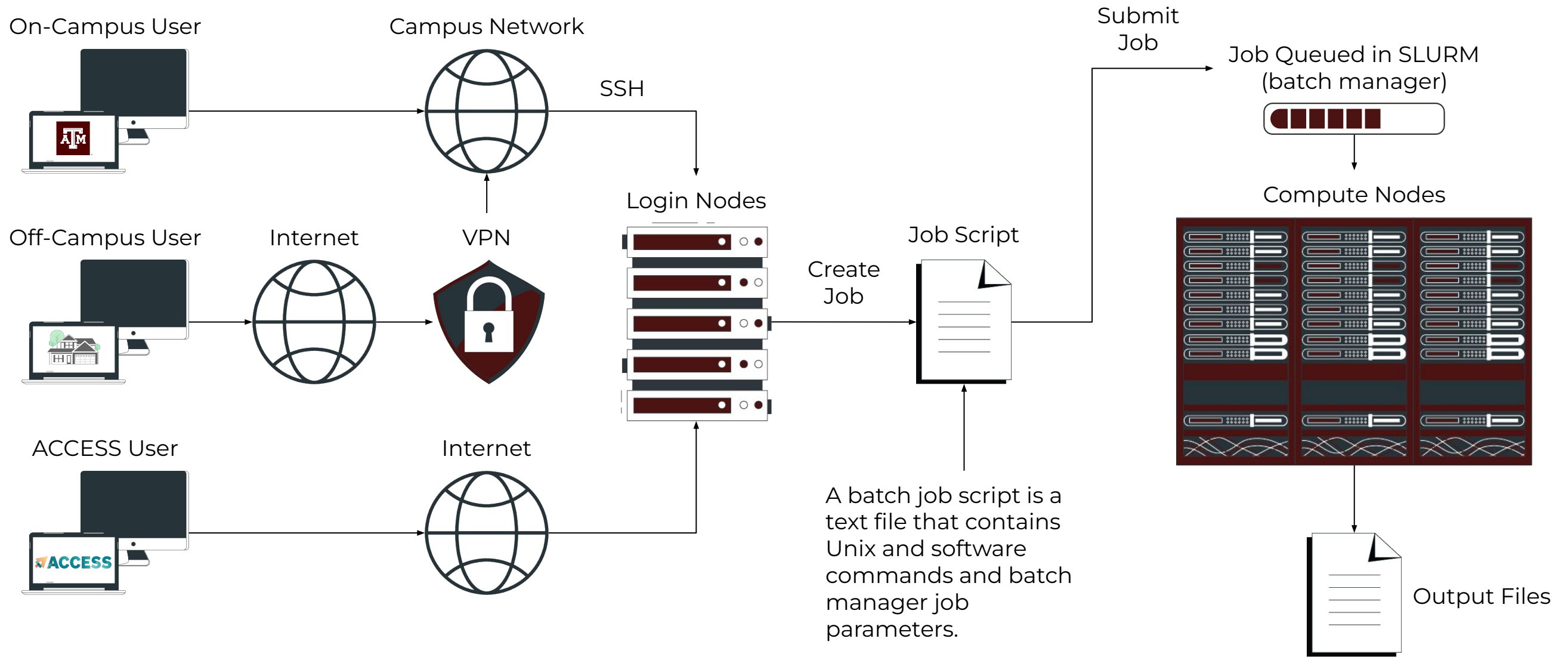
Batch Queues

- Job submissions are auto-assigned to batch queues based on the resources requested (e.g. cpu vs. gpu).
- Use 'sinfo' to check on their status:

```
[ashwin_k@faster1 ~]$ sinfo
PARTITION       AVAIL   TIMELIMIT     JOB_SIZE    NODES(A/I/O/T)      CPUS(A/I/O/T)
cpu*            up      7-00:00:00    1-32        83/21/12/116        5123/1533/768/7424
gpu             up      7-00:00:00    1-32        6/36/10/52          98/2590/640/3328
atsp            down    2-00:00:00    1           0/8/2/10            0/512/128/640
staff           up      2-00:00:00    1           89/65/24/178        5221/4635/1536/11392
```

FASTER

```
[u.ak118300@aces-login3 ~]$ sinfo
PARTITION       AVAIL   TIMELIMIT     JOB_SIZE    NODES(A/I/O/T)      CPUS(A/I/O/T)
cpu*            up      7-00:00:00    1-64        5/36/15/56          418/3518/1440/5376
gpu             up      2-00:00:00    1-15        14/0/2/16           29/1315/192/1536
pvc             up      2-00:00:00    1-15        2/7/11/20           96/768/1056/1920
atsp            up      2-00:00:00    1           0/3/1/4             0/288/96/384
bittware        up      2-00:00:00    1           0/2/0/2             0/192/0/192
d5005           up      2-00:00:00    1           0/2/0/2             0/192/0/192
gpu-hybrid      up      2-00:00:00    1-infinite  15/7/13/35          121/1991/1248/3360
```

ACES

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

Additional information:
https://hprc.tamu.edu/kb/User-Guides/FASTER/Batch/#batch-queues

Sample Job Script Structure

```
#!/bin/bash
##NECESSARY JOB SPECIFICATIONS
#SBATCH --export=NONE
#SBATCH --get-user-env=L
#SBATCH --job-name=JobExample1
#SBATCH --time=01:30:00
#SBATCH --ntasks=1
#SBATCH --mem=2G
#SBATCH --output=stdout.%j


##OPTIONAL JOB SPECIFICATIONS
#SBATCH --account=123456
#SBATCH --mail-type=ALL
#SBATCH --mail-user=email_address


# load required module(s)
module purge
module load GCCcore/11.3.0
module load Python/3.10.4

# Run your program
python my_program.py
```

These parameters describe your job to the job scheduler. The lines starting with #SBATCH are NOT comments!

Account number to be charged (if you're not using a default account).

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

Hands-On Activity

1. Navigate to `/scratch/training/Intro-to-Faster`

2. Copy "`hello_world.job`" to your home directory

Submit Your Job and Check Job Status

Submit job

`sbatch example01.job`

```
Submitted batch job 161997
(from job_submit) your job is charged as below
                Project Account: 122792016265
                Account Balance: 1687.066160
                Requested SUs:    3
```

Check status

`squeue -u username` (or `squeue --me` as a shortcut for your own jobs)

JOBID	NAME	USER	PARTITION	NODES	CPUS	STATE	TIME	TIME_LEFT	START_TIME	REASON	NODELIST
64039	somejob	someuser	medium	4	112	PENDING	0:00	20:00	2017-01-30T21:00:4	Resources	
64038	somejob	someuser	medium	4	112	RUNNING	2:49	17:11	2017-01-30T20:40:4	None	tnxt-[0401-0404]

Hands-On Activity - 5 Minutes

1. Create a submission file for a serial job; you can use the template shown previously or create your own.

2. Submit this job file using sbatch.

3. Check which queue the job is in.

4. When your job completes, check the contents of the ExOutput.txt file.

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|  2023  |       N|   10000.00|              0.00|   10000.00|Doe, John |
-------------------------------------------------------------------------
|142800024371 6|  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
- 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/

(reminder)

OOD Dashboard: FASTER

Easily view Cluster utilization, Storage Quotas & Allocation Balances

Quota and file limit increases will only be considered for scratch directories

Preferred way to request *Quota Increases*

Hands-On Activity - 2 Minutes

1. Use myproject to check the SU balance of your accounts.

2. Use the dashboard to check the same information.

Need Help?

First check the FAQ: https://hprc.tamu.edu/kb/FAQ/Accounts/
- FASTER User Guide: https://hprc.tamu.edu/kb/User-Guides/FASTER/
- ACES User Guide: https://hprc.tamu.edu/kb/User-Guides/ACES/
- Email your questions to help@hprc.tamu.edu

Help us help you -- when you contact us, tell us:
- Which Cluster you're using
- Your 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

Thank you.

Any questions?