Introduction to Python on ASU's HPC systems

Gil Speyer
Lead Scientific Software Engineer
RTO Scientific Software Engineering

Rebecca Belshe
Software Application Analyst
RTO Scientific Software Engineering

Jason Yalim
Postdoctoral Scholar
RTO Scientific Software Engineering

ASU Research Computing
Arizona State University

August 31st, 2020

ASU Knowledge Enterprise
Arizona State University

Core Research Facilities

Getting Started

- To use ASU's Agave system, create an account:
  - Go to http://links.asu.edu/getHPC
  - For "Affiliated Faculty":
    - If faculty, list yourself
    - If a student or student worker, list your professor or supervisor
    - If other, list the speaker of this presentation or ask rtshelp@asu.edu for help

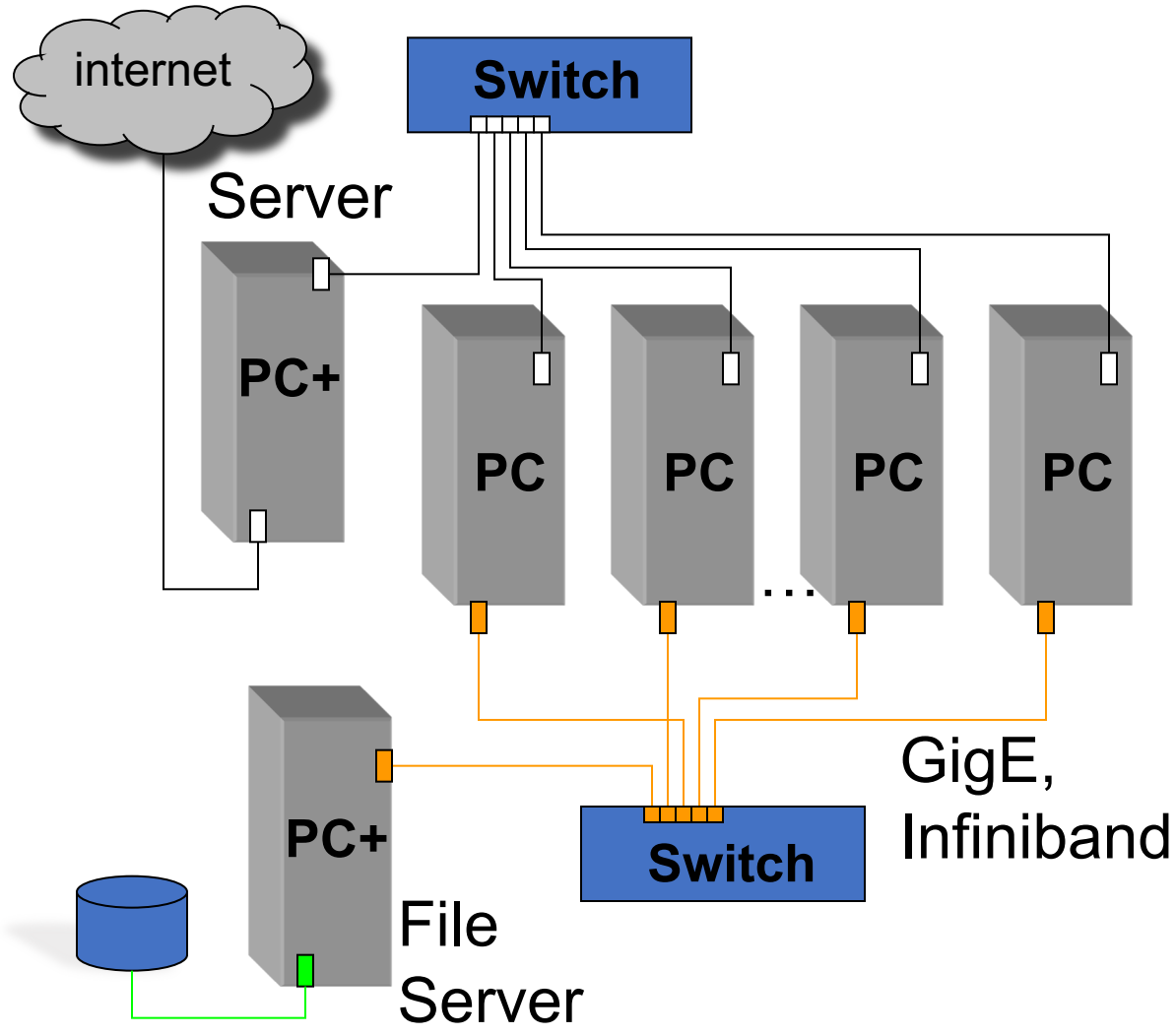
- For access outside ASU campuses, download the VPN:
  - Go to sslvpn.asu.edu and install the Cisco AnyConnect Client
  - In the connect window, enter "sslvpn.asu.edu"
  - Sign in with ASURITE and password
  - Faculty and staff may also need Duo Two-Factor Authentication

What is Agave?

- Agave is a High-Performance Computing (HPC) cluster
- This cluster architecture uses thousands of compute servers, also called nodes, to help users optimize their research
- This optimization can happen in several ways:
  - Free the researcher's local machine by running the program on a separate server
  - Allow for parallelization or for multiple operations, or "jobs", to run in parallel
  - Access to graphics processing units (GPUs)
  - Access to high memory computation and storage
  - Overall greater and more refined research through HPC

Image Credit: NetApp

How Does a Cluster Work?

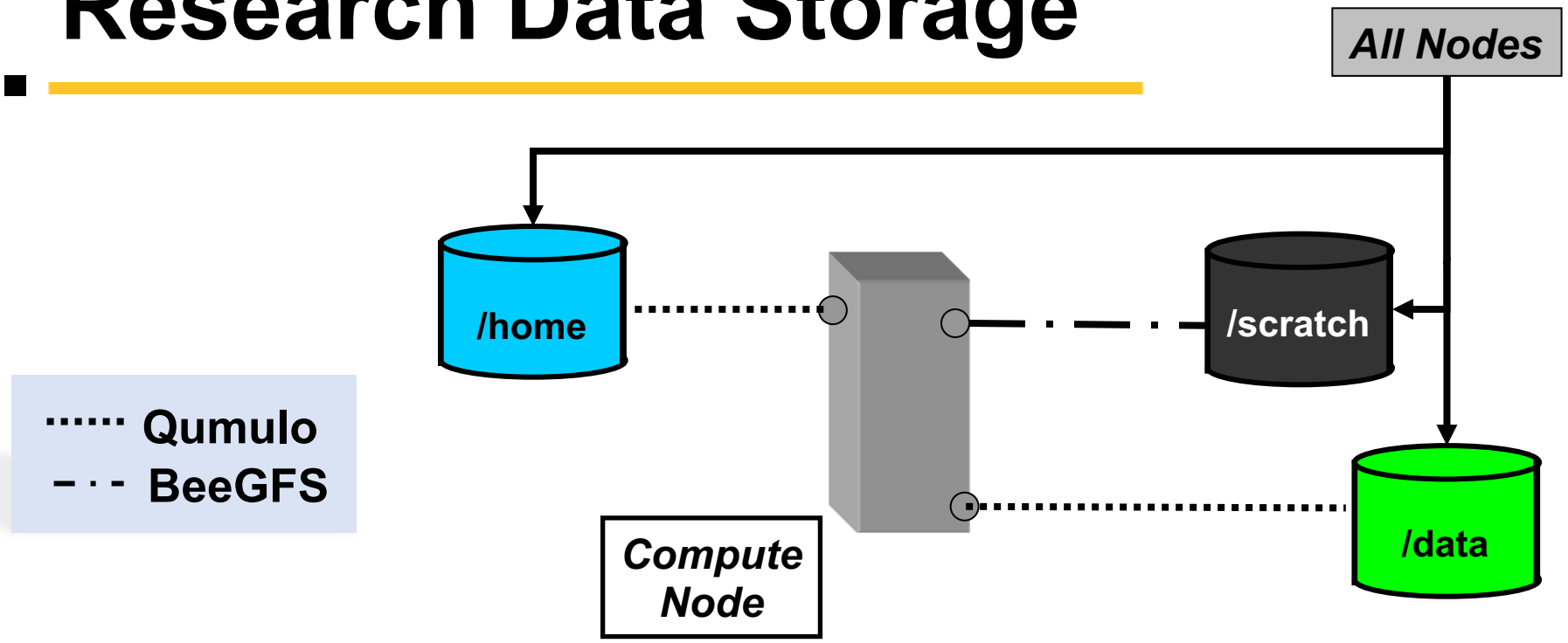


(Adv. HPC System)

Agave System Information

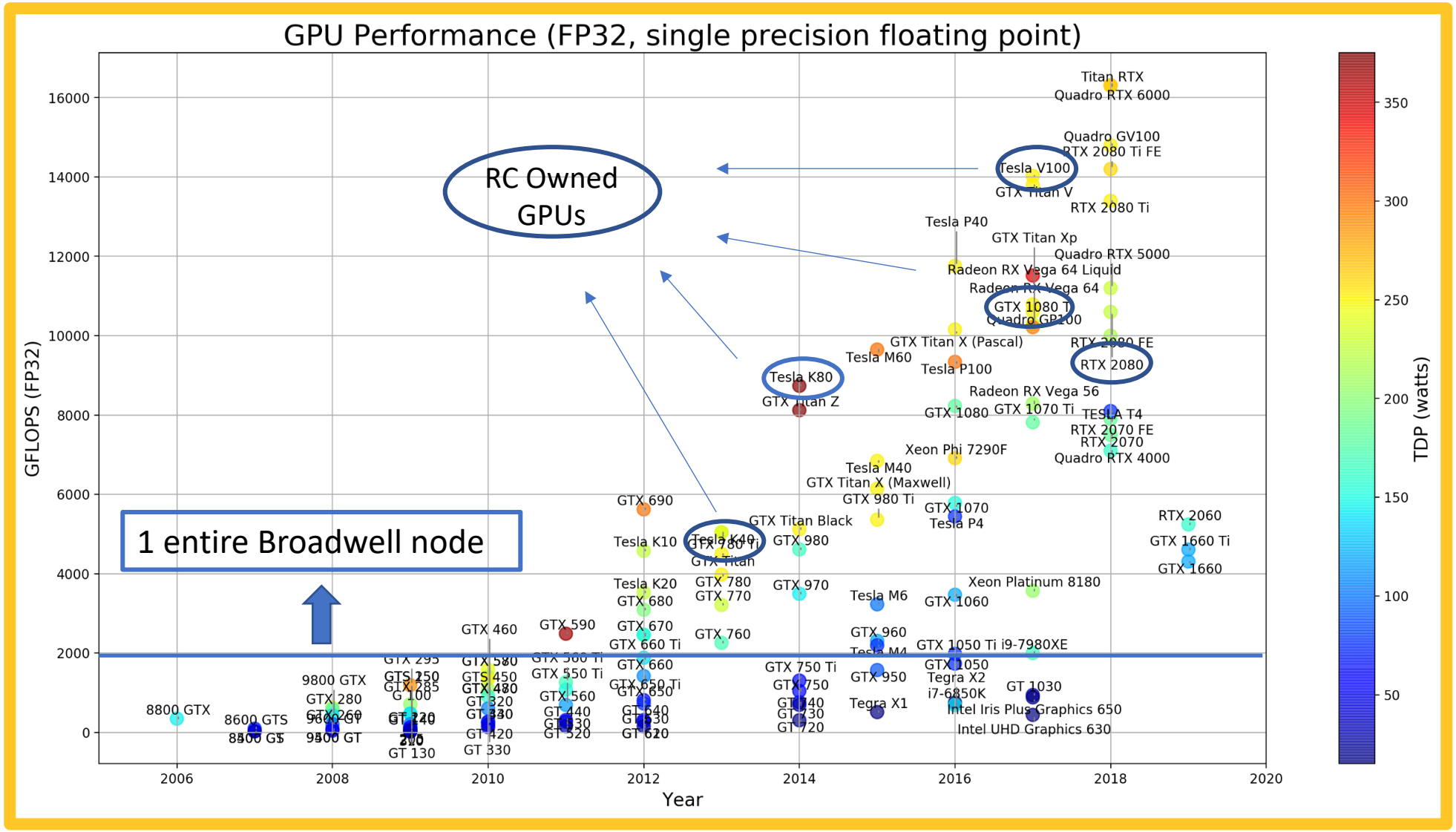
- Agave ~7.5K networked Broadwell cores (Intel CPU)
  - >300 trillion (theoretical) floating point operations/s (TFLOPs)
  - >34TB aggregate RAM
  - >6PB aggregate disk
  - Omnipath Interconnect – 100Gb/s
  - Located in ISTB1
- Additional ~3K cores in various partitions
- ~5K Xeon Phi cores
- GPU partitions (~300 GPUs)
- High memory nodes (1TB, 1.5TB, 2TB)
- Remember, Agave is a true **cluster** architecture
  - Each Broadwell node has 28 processors and 128GB of RAM (~4.5GB/CPU).
  - Programs needing more resources **must** use parallel programming
  - Normal, single processor applications **do not go faster** on Agave

Research Data Storage

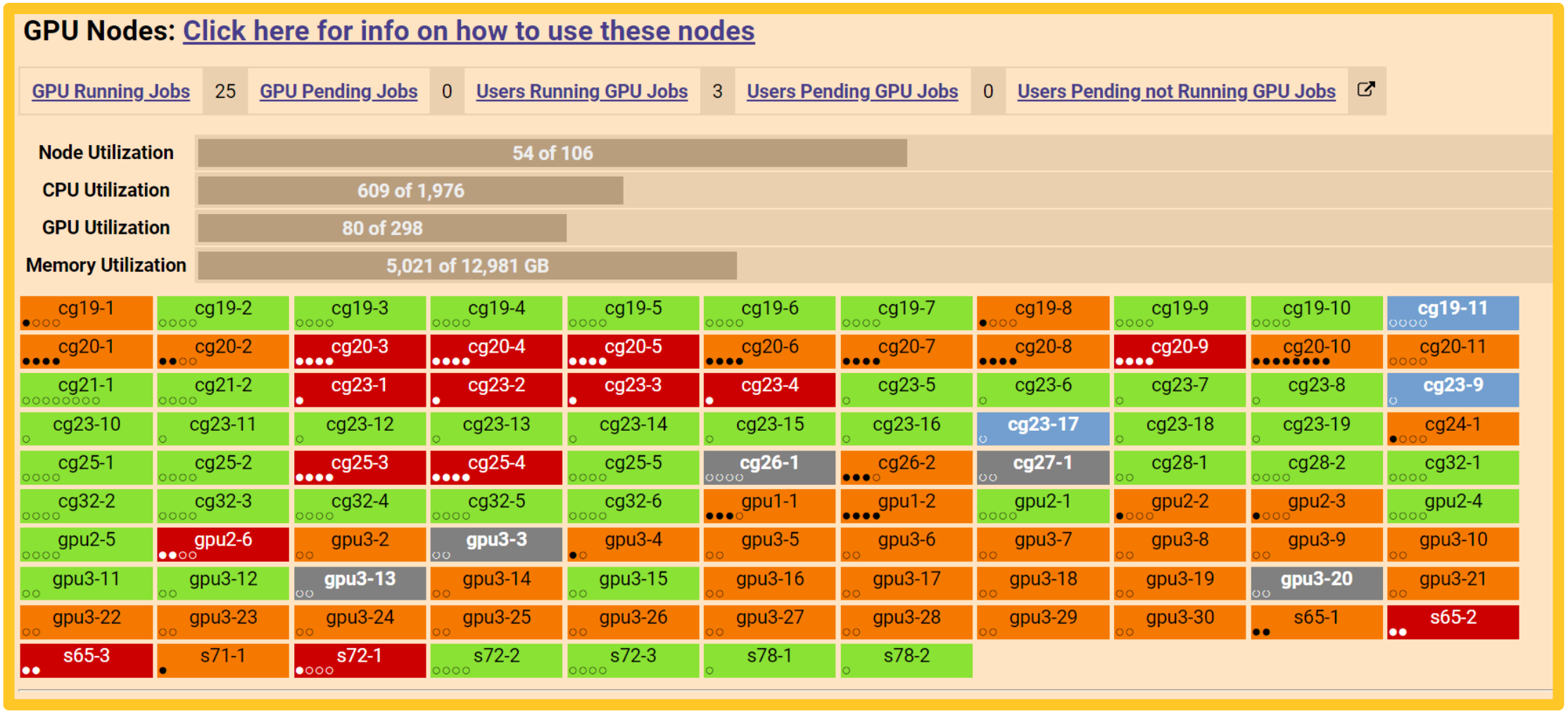


Mount point	User Quota	Lifetime
/home	100 GB	Project
/scratch	None	30 days
/data	Purchased	$50/TB/year

Graphical Processing Units (GPU) on Agave and Performance



Agave's GPU Nodes



Legend:

Green – node is free

Orange – node is mixed between taken and free

Red – node is taken

Blue – node is being drained

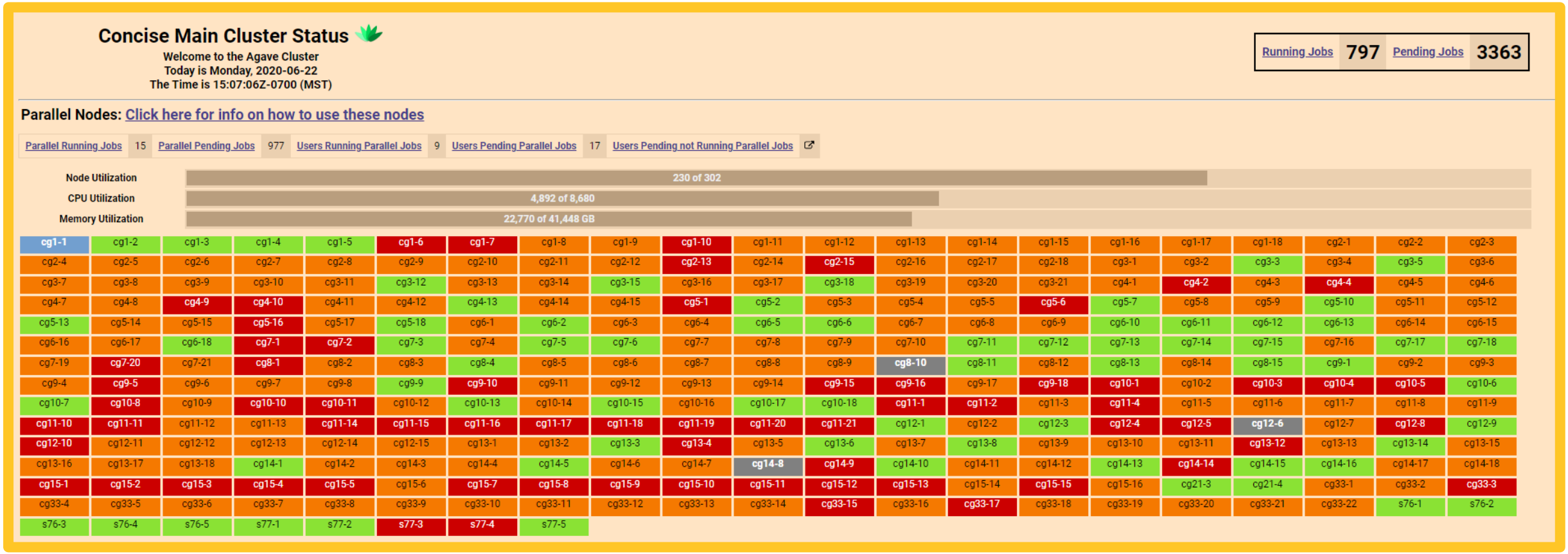
Grey – node is offline

● – GPU is taken

○ – GPU is free

Agave Status Page

- To see the status of Agave and to check which nodes are free, check out http://links.asu.edu/status



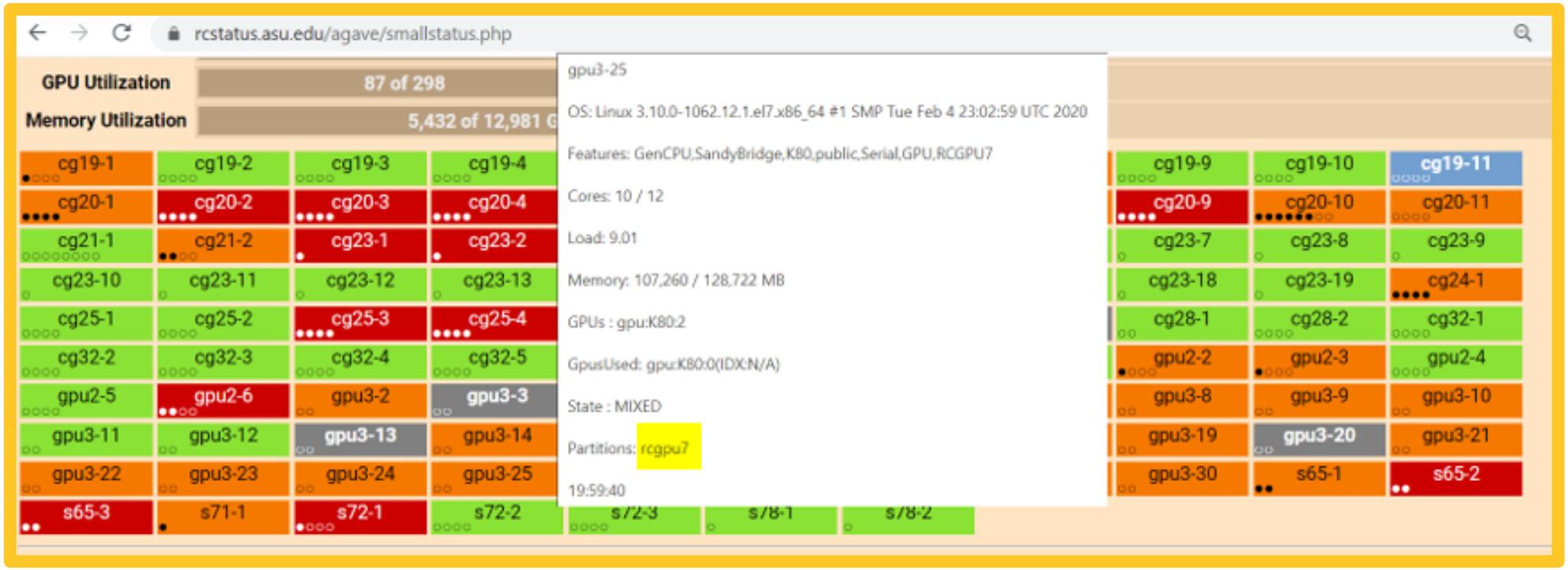
Legend:

Green – node is free **Orange** – node is mixed between taken and free **Red** – node is taken

Blue – node is being drained **Grey** – node is offline

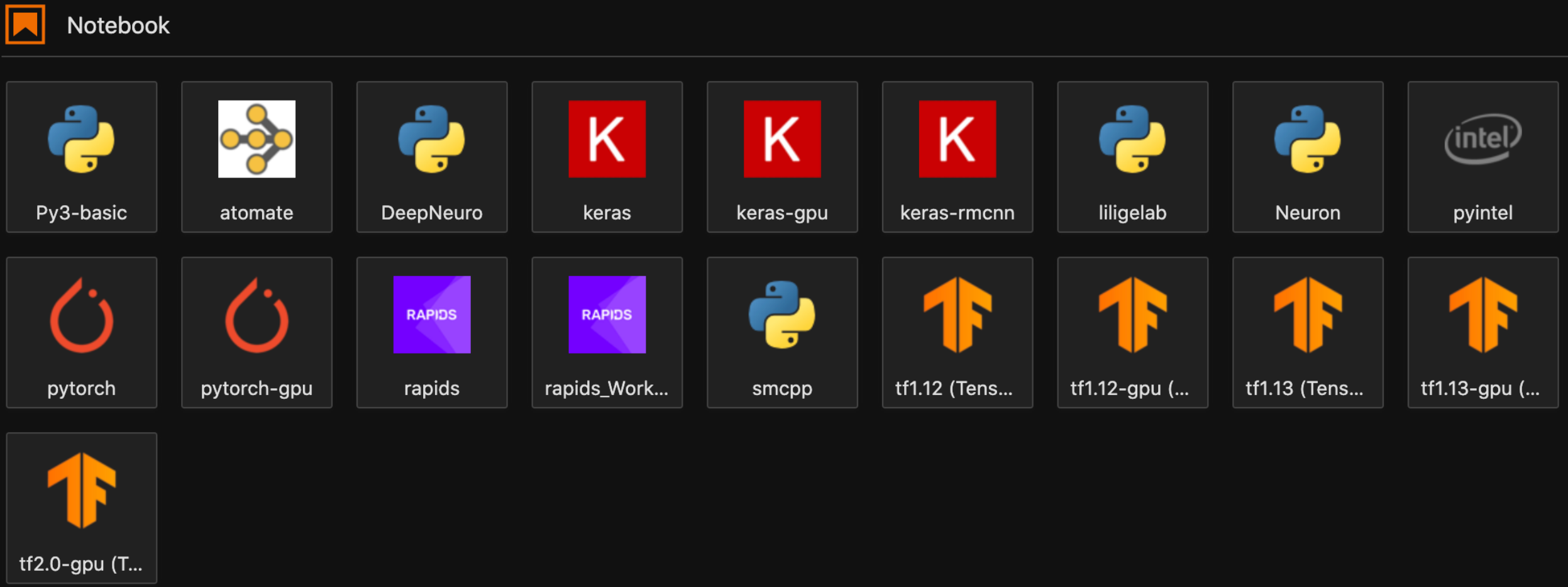
More on Agave's GPU Nodes

- Currently, Research Computing's GPU partition contains over 100 GPU nodes with nearly 300 GPUs collectively
- This GPU partition is broken into two subpartitions:
  - 1) private GPUs
  - 2) public GPUs

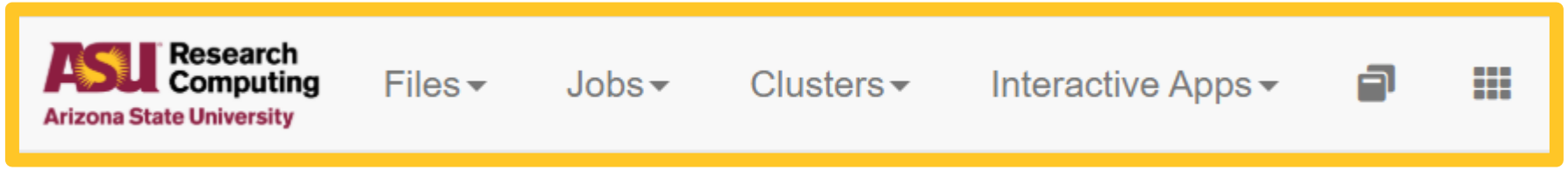


Popular Python Software

Notebook

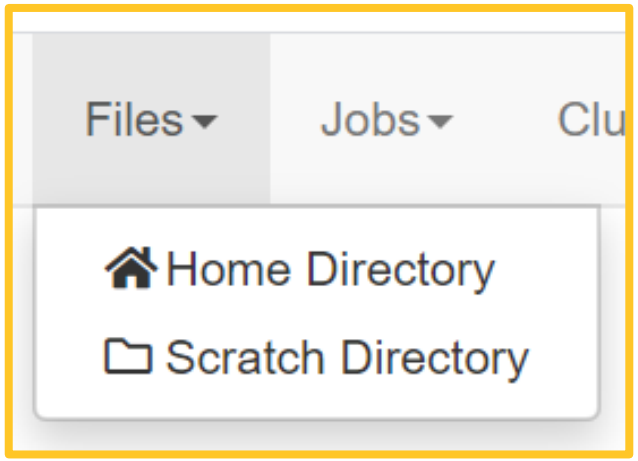


Using Agave through OPEN OnDemand

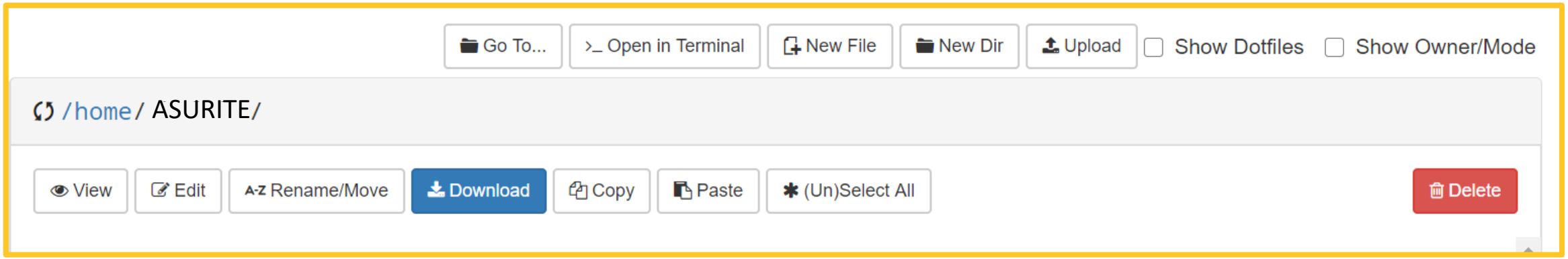


- Supercomputing through the web
- Once connected to VPN: go to https://login.rc.asu.edu
- Single sign on with ASURITE and password
- From here users can:
  - Manage file systems Files
  - Create and monitor jobs Jobs
  - Access the shell Clusters
  - Access interactive apps Interactive Apps
  - View and manage interactive sessions
  - View all apps
  - Request help Help

Browser-Based File Management

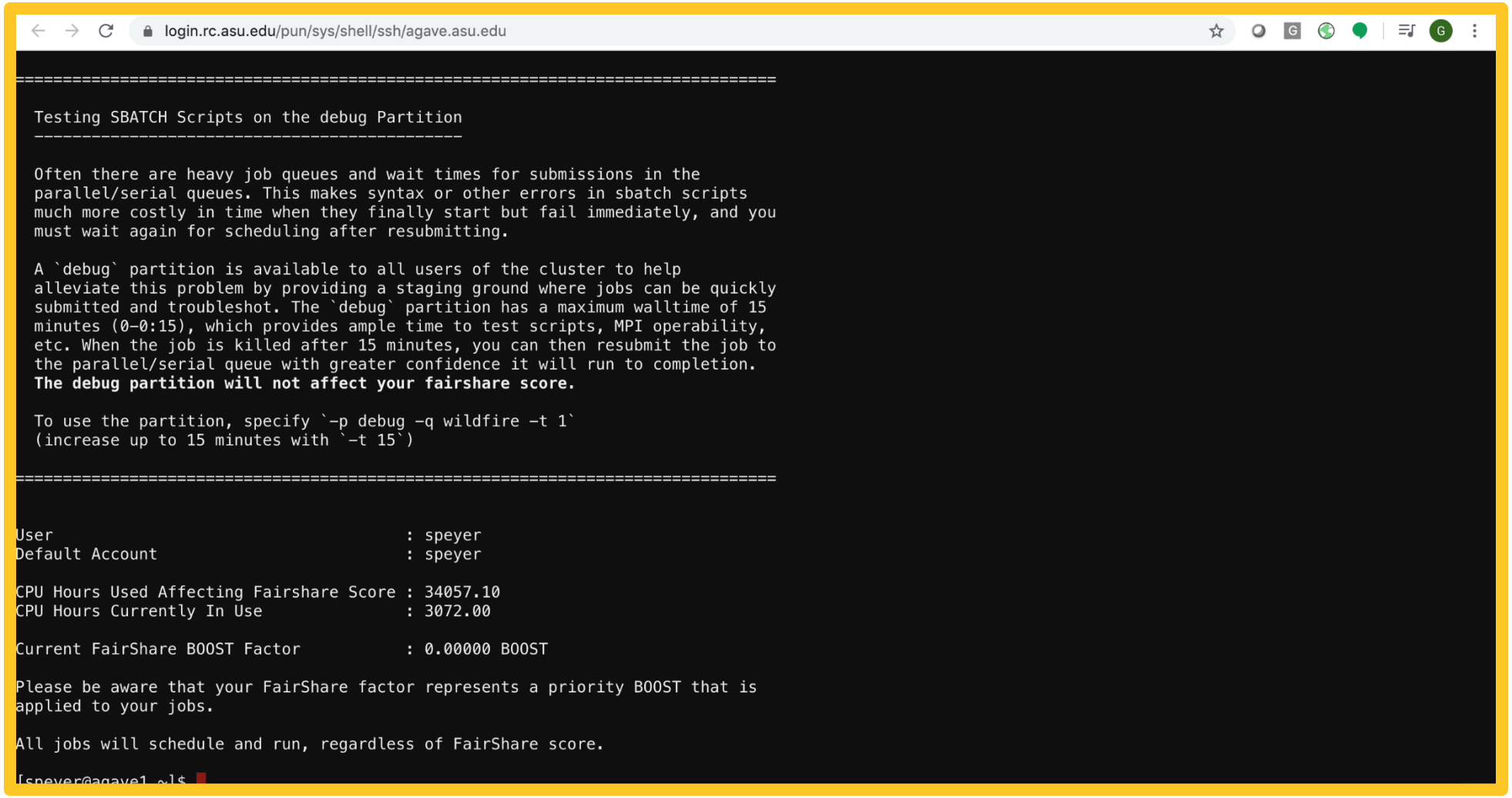


- Manage files between local machine, home, and scratch with click and drop interface
- See more under Available File Systems

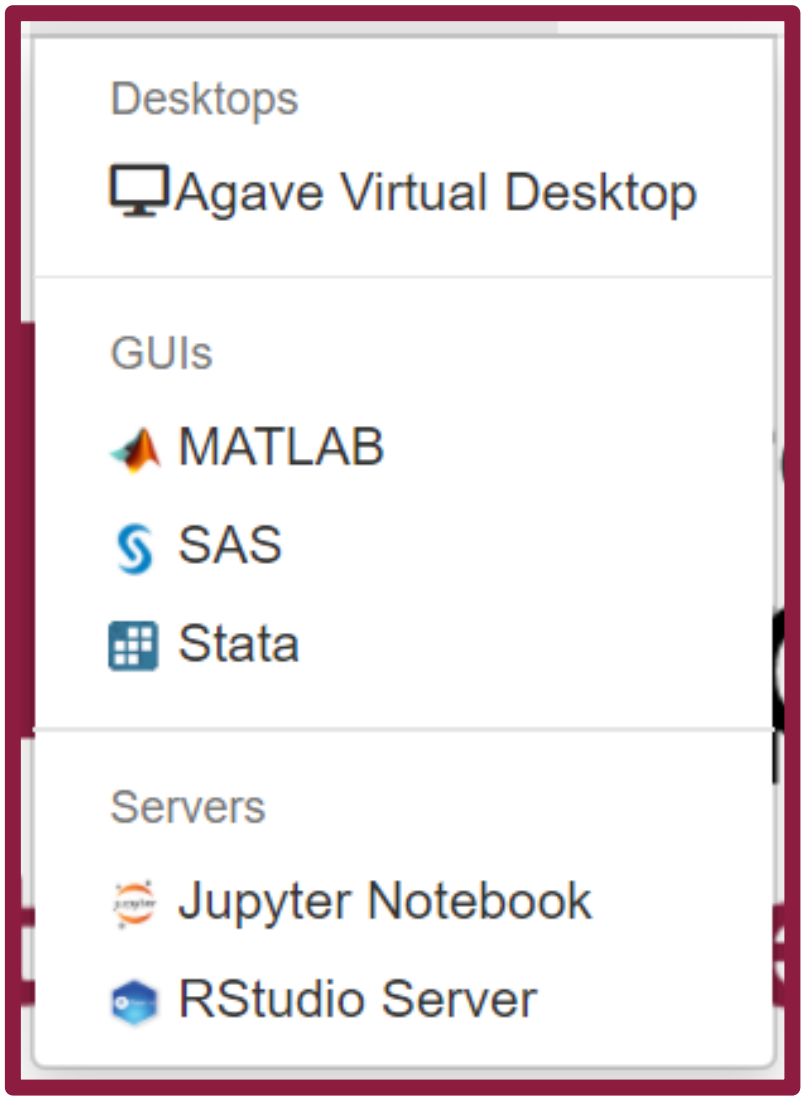


Browser-Based Job Submission

Browser-Based Shell



Browser-Based Virtual Desktops



Select Jupyter tricks

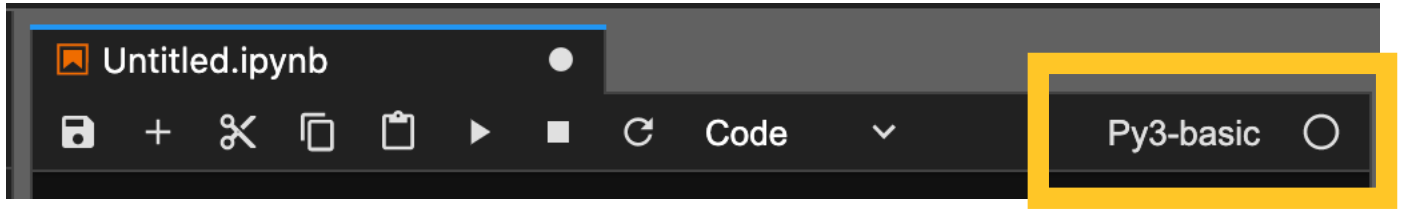
Change from Jupyter lab to notebook midsession:

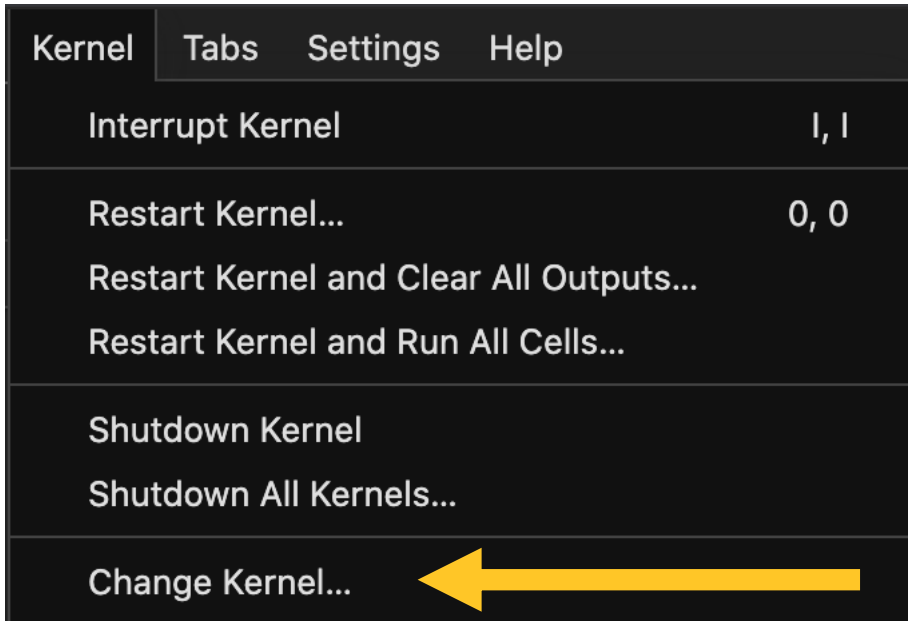
https://login.rc.asu.edu/node/cg1-2.agave.rc.asu.edu/38457/lab?

->

https://login.rc.asu.edu/node/cg1-2.agave.rc.asu.edu/38457/tree

Change the kernel of a notebook:





Scheduling, Priority, and Fairshare

- Jobs in the queue are scheduled based on priority through the SLURM scheduler
- Priority is computed from a linear combination of factors: job age, job size, job quality of service, and FairShare with admin defined weights
- FairShare ranges from 0 (lowest priority) to 1 (highest priority)
- Calculated from usage

$$\text{FairShare} = 2^{-usage/(10{,}000\ core\ hours)}$$

- $usage$ decays exponentially, half-life of 1 week

Other Resources and Services

Additional advantages of Research Computing:

- Consultation in optimizing research
- Consultation and opportunities to purchase hardware through Research Computing--
  - Research Computing maintains the hardware, handles OS, and provides software installations for researchers
- Offers XSEDE Campus Champions allocations
- Offers in-depth workshops for HPC applications, XSEDE, and Carpentries

Contribute! (http://links.asu.edu/services)

Dell Pre-Approved Compute Nodes

Compute Node Types	Processor	Core Count	RAM	GPU	Price
Standard Intel	Intel Cascade Lake	52 cores	192GB	N/A	$9,555.87
High Memory	Intel Cascade Lake	52 cores	1,536GB	N/A	$20,801.46
GPU Option 1	Intel Cascade Lake	52 cores	192GB	(4) 16GB V100	$37,242.80
GPU Option 2	Intel Cascade Lake	52 cores	192GB	(4) 32GB V100	$43,331.73
Standard AMD	AMD EPYC	64 cores	256GB	N/A	$11,450.04

Fall Workshops (links.asu.edu/learn)

Introduction to Linux Command Line Interface:

- **Aug 17**, 3-4 PM, Zoom.
- **October 19**, 3-4 PM, Zoom.

Matlab for HPC and on GPU:

- **Aug 24**, 2-3 PM, Zoom.

Python HPC Applications Series:

- **Aug 31**, 2-3 PM, *Python Intro for HPC* (**Python 1**), Zoom.
- **Sep 8**, 2-3 PM, *Python Data Handling for HPC* (**Python 2**), Zoom.
- **Oct 5**, 3-4 PM, *Python Machine Learning for HPC* (**Python 3**), Zoom.
- **Nov 11**, 2-3 PM, *Python Deep Learning for HPC* (**Python 4**), Zoom.

Google Drive, Globus, and HPC:

- **Aug 31**, 3-4 PM, Zoom.

R for HPC and GPU Applications:

- **Sep 14**, 2-3 PM, Zoom.

General GPU Applications:

- **Oct 5**, 2-3 PM, Zoom.

Introduction to HPC and the Research Computing Cluster:

- **Aug 17**, 2-3 PM, Zoom.
- **Sep 21**, 2-3 PM, Zoom.
- **Nov 30**, 2-3 PM, Zoom.
- **Dec 14**, 2-3 PM, Zoom.

Quick Links

For assistance: rtshelp@asu.edu
Office Hours: 1-3:30 Tu (and W during Academic semesters)
Zoom: https://asu.zoom.us/j/9598636843 Password: 456456
Slack Channel: #rc-support
Documentation Page: http://links.asu.edu/docs
Discourse Site (for common errors and fixes):
http://links.asu.edu/discourse
Agave Status Page:
http://links.asu.edu/status
Mailing List:
http://links.asu.edu/MailingList

Thank you

Please visit

researchcomputing.asu.edu

for more information

ASU Research Computing
Arizona State University

August 31st 2020

ASU Knowledge Enterprise
Arizona State University

Core Research Facilities