

HIGH PERFORMANCE RESEARCH COMPUTING

ACES: Introduction to CryoSPARC for Cryo-EM Data Processing in Collaboration with the Laboratory for Biomolecular Structure and Dynamics



High Performance
Research Computing
DIVISION OF RESEARCH

Fall 2023



High Performance Research Computing | hprc.tamu.edu | NSF Award #2112356

CryoSPARC on ACES

- CryoSPARC Academic License ID
- Resources and Limitations
- CryoSPARC on the ACES Portal
- Group data directories
- CryoSPARC Tutorial

CryoSPARC Academic License ID

- CryoSPARC provides a personal Academic License ID free of charge for TAMU staff and students
- A CryoSPARC Academic License ID is required to launch the CryoSPARC ACES portal app
- Use your academic email address for the Academic License
 - <https://cryosparc.com/download>

Resources for Running CryoSPARC

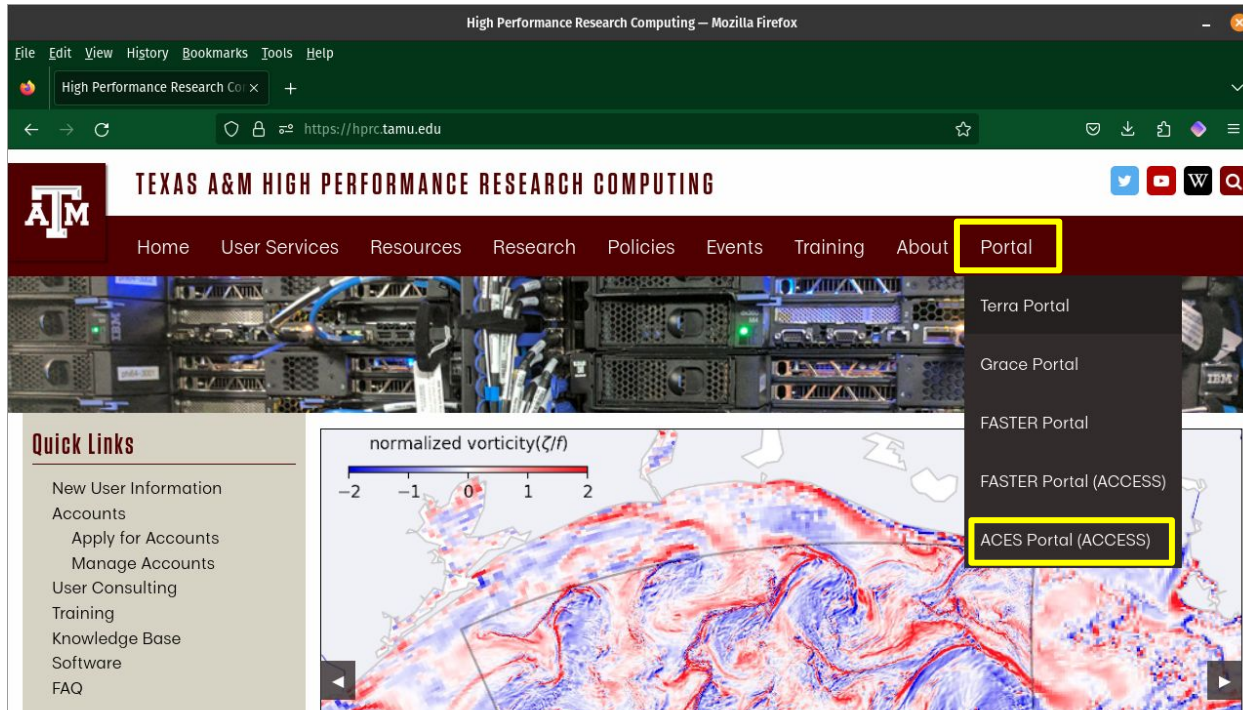
- CryoSPARC is available as an Interactive App on the ACES Portal
 - <https://portal-aces.hprc.tamu.edu>
- ACES hardware overview (can change since ACES is a composable resource cluster)
 - 30 x H100 GPUs (2 x H100 GPUs per node)
 - 6 x A30 GPUs (6 x A30 GPUs per node)
 - 70 x CPU only nodes
 - 10 x other specialized nodes not for CryoSPARC
- Specify enough time to allow your processing to complete
- If you launch a job for 24 hours and you finish your work in 12 hours, delete your job to make GPUs available for other jobs



Resource Limitations

- Can only launch one CryoSPARC portal session at a time
 - can run multiple CryoSPARC jobs within a portal session
- CPUs
 - Some CryoSPARC tasks do not require a GPU such as importing images
- GPUs
 - The GPU queue can get busy at times and it may take over an hour for your job to launch
 - Cancel your pending job if you will be away from your computer for a long time and it hasn't started yet
- Submitted jobs can have delayed start times due to unavailability of GPUs or scheduled system maintenance

Accessing the ACES Portal



The screenshot shows the Texas A&M High Performance Research Computing (HPRC) website in a Mozilla Firefox browser. The browser's address bar displays <https://hprc.tamu.edu>. The website header includes the TAMU logo and the text "TEXAS A&M HIGH PERFORMANCE RESEARCH COMPUTING". A navigation menu contains links for Home, User Services, Resources, Research, Policies, Events, Training, About, and Portal. The "Portal" link is highlighted with a yellow box, and a dropdown menu is open, listing "Terra Portal", "Grace Portal", "FASTER Portal", "FASTER Portal (ACCESS)", and "ACES Portal (ACCESS)". The "ACES Portal (ACCESS)" option is also highlighted with a yellow box. Below the navigation menu, there is a banner image of server racks and a "Quick Links" sidebar with items like "New User Information", "Accounts", "Apply for Accounts", "Manage Accounts", "User Consulting", "Training", "Knowledge Base", "Software", and "FAQ". A central visualization shows a map of Texas with a color scale for "normalized vorticity(ζ/f)" ranging from -2 to 2.

HPRC webpage: hprc.tamu.edu

Accessing ACES via the ACCESS Portal

Log-in using your ACCESS credentials.

The screenshot shows the ACCESS portal interface. At the top left is the ACCESS logo, and at the top right is the text "Powered By CILogon" with the CILogon logo. Below the header is a "Consent to Attribute Release" section with a dropdown arrow. The consent text reads: "TAMU FASTER ACCESS OOD requests access to the following information. If you do not approve this request, do not proceed." followed by a list of requested information: "Your CILogon user identifier", "Your name", "Your email address", and "Your username and affiliation from your identity provider". Below the consent section is a "Select an Identity Provider" section. It features a dropdown menu with "ACCESS CI (XSEDE)" selected. There is a checkbox for "Remember this selection" and a "Log On" button. A note below the button states: "By selecting 'Log On', you agree to the privacy policy." At the bottom of the page, there is a footer with links for "FAQs" and "help@cilogon.org".

The screenshot shows the ACCESS portal login interface. At the top left is the ACCESS logo, and at the top right is the CILogon logo. The main heading is "Login to CILogon". Below this are two input fields: "ACCESS Username" and "ACCESS Password". There is a checkbox for "Don't Remember Login" and a "Login" button. To the right of the login fields is a list of links: "Register for an ACCESS Account", "Forgot your password?", and "Need Help?". A note above the links states: "If you had an XSEDE account, please enter your XSEDE username and password for ACCESS login". At the bottom of the page, there is a footer with a link for "Click Here for Assistance".

This is a close-up of the "Select an Identity Provider" dropdown menu. The dropdown is open, showing the selected option "ACCESS CI (XSEDE)" with a question mark icon to its right.

Select the Identity Provider appropriate for your account.

ACES CryoSPARC Portal App

portal-aces.hprc.tamu.edu

TEXAS A&M HIGH PERFORMANCE RESEARCH COMPUTING

Home User Services Resources Research Policies Events Training About Portal

Terra Portal
Grace Portal
FASTER Portal
FASTER Portal (ACCESS)
ACES Portal (ACCESS)

Quick Links

- New User Information
- Accounts
- Apply for Accounts
- Manage Accounts

normalized vorticity(ζ/f)

TAMU HPRC OnDemand (ACES) Files Jobs Clusters Interactive Apps

Interactive Apps

GUI

- VNC
- Nextsilicon VNC
- Imaging
- CryoSPARC
- ImageJ
- cisTEM

CryoSPARC

This app will launch CryoSPARC on one ACES cluster compute node.

Version

4.2.1

CryoSPARC license id (required)

- You must get an individual license id from cryosparc.com/download
- After you click the blue Launch button below, it may take up to 5 minutes before CryoSPARC is available
- Use the following to login to your CryoSPARC session:
 - username: **admin@cryo.edu**
 - password: **admin**
- You can only launch one CryoSPARC portal app job but you can run multiple computational jobs within CryoSPARC

Number of hours (max 168)

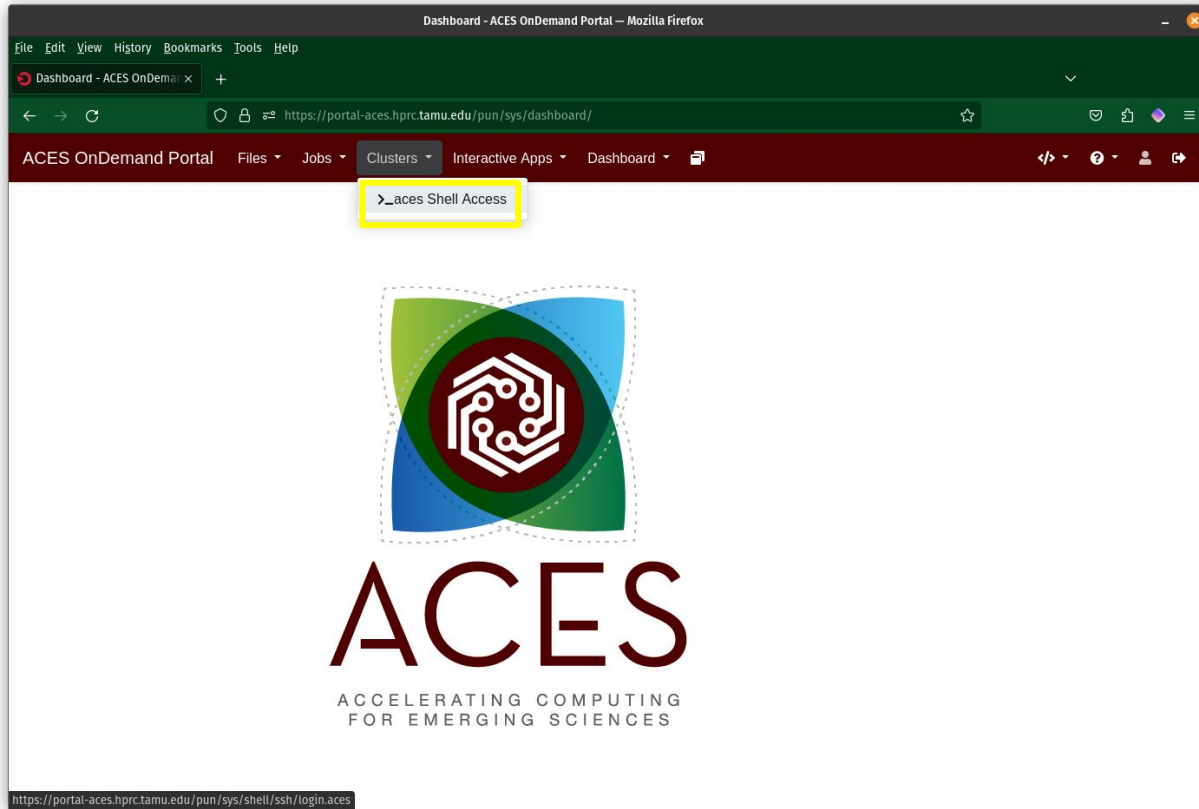
3

Node type

CPU only

Launch CryoSPARC the first time using 'CPU only' in order to initialize the database.

Shell Access via the ACES Portal



Display GPU Availability on ACES

See GPU configuration and current availability using the ACES portal shell access

- <https://portal-aces.hprc.tamu.edu>
- select "Clusters" then ">_aces Shell Access"
- at the command line prompt, enter:

gpuavail

CONFIGURATION	
NODE TYPE	NODE COUNT

gpu:h100:2	13
gpu:pvc:4	6
gpu:h100:4	1
gpu:a30:6	1

There are currently 15 compute nodes on ACES that have 2 x H100 GPUs attached to each

AVAILABILITY					
AVAILABILITY					
NODE NAME	GPU TYPE	GPU COUNT	GPU AVAIL	CPU AVAIL	GB MEM AVAIL

ac027	h100	4	1	72	248
ac048	h100	2	2	96	488
ac049	h100	2	1	88	408
ac050	h100	2	2	96	488
ac064	a30	6	6	96	488

ACES Maintenance

- ACES will be offline during routinely scheduled maintenances which may occur a few times each year
- Submitted jobs will remain PENDING and not run if the job script wall time overlaps with the maintenance start time
- When maintenance is complete, scheduled jobs will automatically begin running based on priority
- Use the **maintenance** command to see if a maintenance is scheduled and to display the time until the start of the maintenance

maintenance

```
The scheduled 35 hour ACES maintenance will start in:
```

```
3 days 21 hours 32 minutes
```

```
Scheduled jobs will not start if they overlap with this maintenance window.
```

ACES CryoSPARC Running Job

CryoSPARC (26167)

1 node

16 cores

Running

Host: >_ac036

Delete

Created at: 2023-11-02 12:54:31 CDT

Time Remaining: 2 hours and 53 minutes

Session ID: 40e97f89-f167-44f2-a9af-add6debd6302

Click Launch

Launch CryoSPARC

Increase Image Quality to 9

Image Quality

0 (low) to 9 (high)

View Only (Share-able Link)

ACES CryoSPARC Portal App

portal-aces.hprc.tamu.edu

TEXAS A&M HIGH PERFORMANCE RESEARCH COMPUTING

Home User Services Resources Research Policies Events Training About Portal

Terra Portal
Grace Portal
FASTER Portal
FASTER Portal (ACCESS)
ACES Portal (ACCESS)

Quick Links

- New User Information
- Accounts
- Apply for Accounts
- Manage Accounts

normalized vorticity(ζ/f)

TAMU HPRC OnDemand (ACES) Files Jobs Clusters Interactive Apps

Interactive Apps

GUI

- VNC
- Nextsilicon VNC
- Imaging
- CryoSPARC
- ImageJ
- cisTEM

CryoSPARC

This app will launch CryoSPARC on one ACES cluster compute node.

Version

4.2.1

CryoSPARC license id (required)

- You must get an individual license id from cryosparc.com/download
- After you click the blue Launch button below, it may take up to 5 minutes before CryoSPARC is available
- Use the following to login to your CryoSPARC session:
 - username: **admin@cryo.edu**
 - password: **admin**
- You can only launch one CryoSPARC portal app job but you can run multiple computational jobs within CryoSPARC

Number of hours (max 96)

3

Node type

H100 GPU

Number of GPUs

1

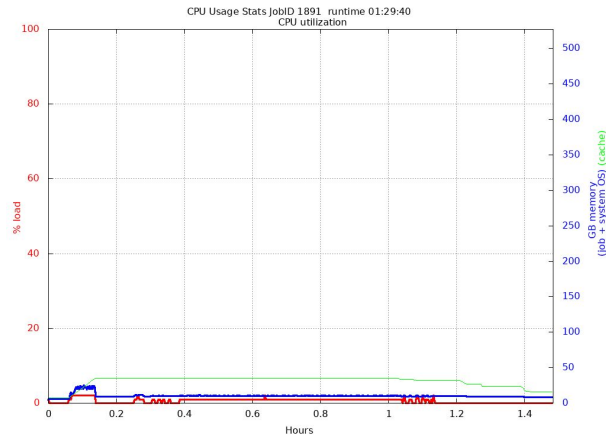
Delete your CryoSPARC "CPU only" job and launch a new CryoSPARC job using 1 x H100 GPU for the tutorial

CryoSPARC Visualization of Resource Utilization

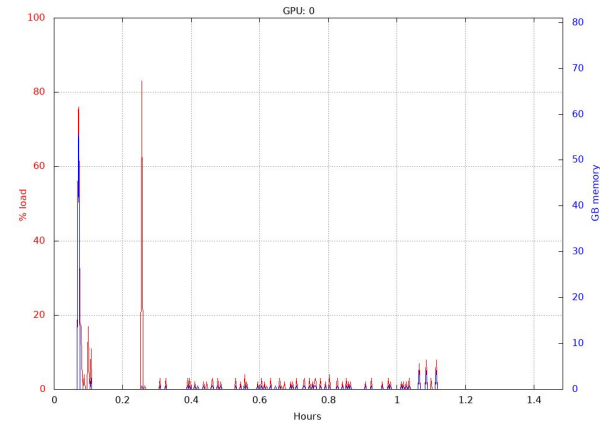
Review CPU and GPU usage for a Job

The `jobstats` utility automatically runs for each CryoSPARC portal job and monitors CPU and GPU resource usage and creates a graph for each.

- CPU stats monitors all cores regardless of how many were configured for the job
 - CryoSPARC jobs use a fraction of total cores and memory if selecting one of many GPUs on a compute node to allow others to use the additional GPUs on the compute node
- GPU stats will create a graph for each GPU that was configured for the job



stats_cpu.1891.png

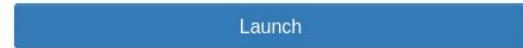


stats_gpu.1891.png

Review CPU and GPU usage for a CryoSPARC Job

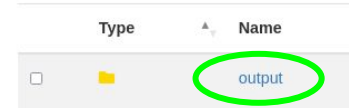
1. At the bottom of the CryoSPARC portal app, access the CryoSPARC data root directory
2. Click 'output'
3. Click 'Modified at' to sort newest session on top
4. Click the Name of the session you want to view based on the date
5. Click the .png image files to view the CPU and GPU usage stats

1.

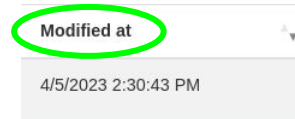


* The CryoSPARC session data for this session can be accessed under the **data root directory**.

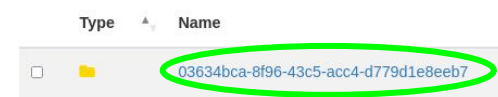
2.



3.



4.



5.



Group Data Directories

- Group data directories can be used to share input files or to have a group work space
- Send a request to the HPRC helpdesk to create a group
 - help@hprc.tamu.edu
 - provide a group name and usernames of members
 - group directories have their own disk quotas separate from individual users
- You will need to mount your group directory when launching the CryoSPARC portal app

Optional group directory to mount

- You must already be a member of the group
- Example values:
 - /junjiez
 - /scratch/group/davislab

CryoSPARC Tutorial