

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

Spring 2024



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 Interactive 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

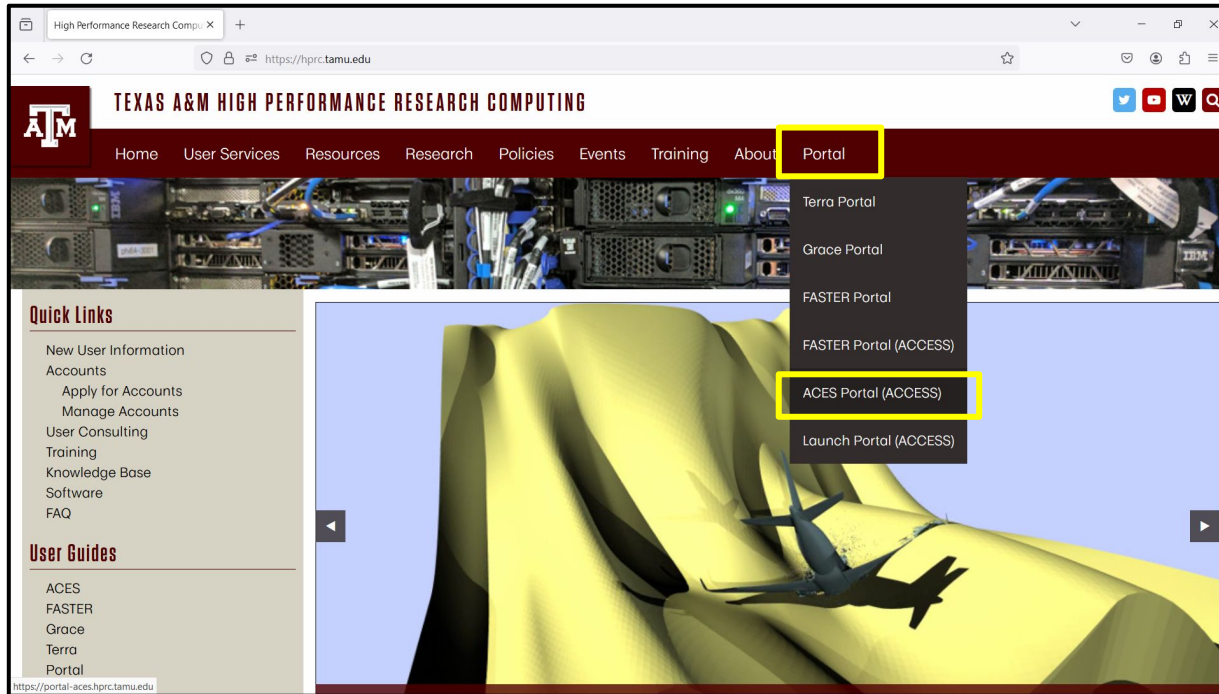
- CryoSPARC is available as an Interactive App on the ACES Portal
 - <https://portal-aces.hprc.tamu.edu>
- ACES hardware overview (can change since ACES is a composable resource cluster)
 - 30 x H100 GPUs (2 x H100 GPUs per node)
 - 6 x A30 GPUs (6 x A30 GPUs per node)
 - 70 x CPU only nodes
 - 10 x 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 resources 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, such as importing images, do not require a GPU
- 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 HPRC ACES Portal



HPRC webpage: hprc.tamu.edu

Accessing ACES via the ACCESS Portal

Log-in using your ACCESS credentials.

ACCESS Powered By **CILogon**

Consent to Attribute Release

TAMU FASTER ACCESS OOD requests access to the following information. If you do not approve this request, do not proceed.

- Your CILogon user identifier
- Your name
- Your email address
- Your username and affiliation from your identity provider

Select an Identity Provider

ACCESS CI (XSEDE)

Remember this selection

Log On

By selecting "Log On", you agree to the [privacy policy](#).

For questions about this site, please see [FAQs](#) or send email to help@clilogon.org.
Know your responsibilities using the CILogon Service. See [acknowledgments](#) or [support](#) for this site.

ACCESS

Login to CILogon

ACCESS Username

ACCESS Password

Don't Remember Login

Login

CILogon

CILogon facilitates secure access to CyberInfrastructure (CI).

- If you had an XSEDE account, please enter your XSEDE username and password for ACCESS login
- Register for an ACCESS Account
- Forgot your password?
- Need Help?

[Click Here for Assistance](#)

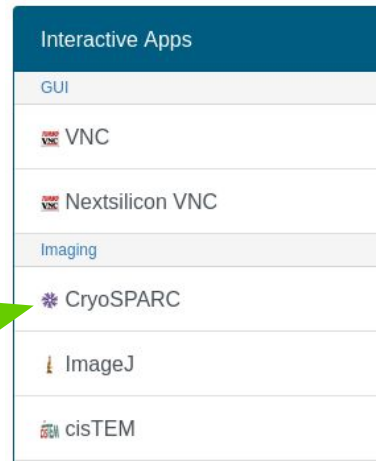
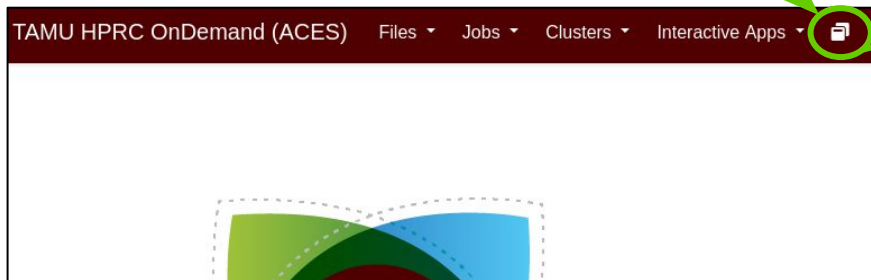
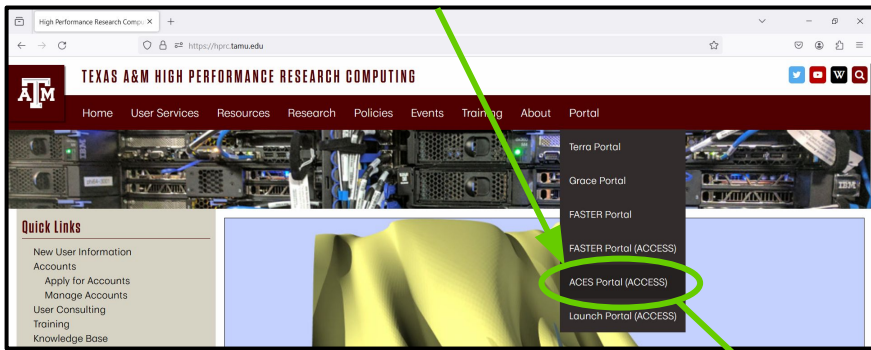
Select an Identity Provider

ACCESS CI (XSEDE)

Select the Identity Provider appropriate for your account.

ACES CryoSPARC Portal App

portal-aces.hprc.tamu.edu



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)

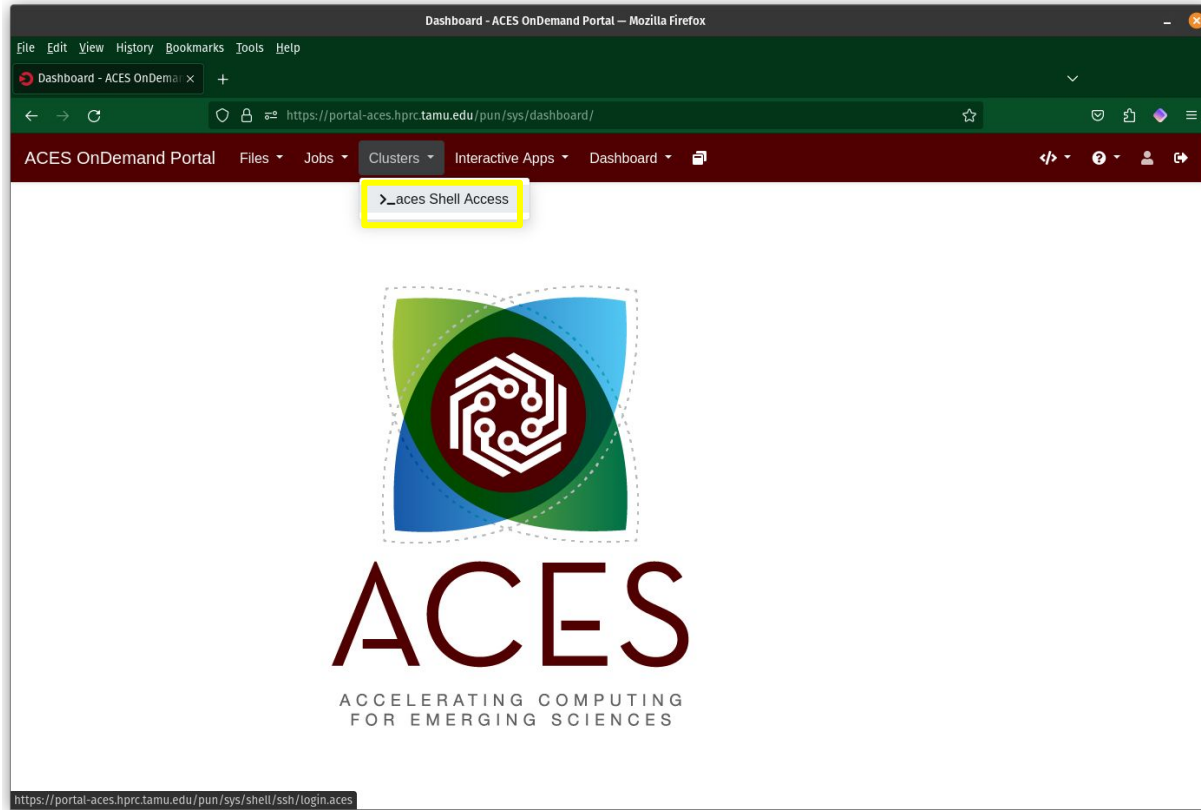
3

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**
- the CONFIGURATION will change frequently since ACES is a composable resource cluster

Based on this output, there are 9 compute nodes on ACES that have exactly 2 x H100 GPUs attached to each

<https://hprc.tamu.edu/kb/Software/useful-tools/gpuavail>

gpuavail

CONFIGURATION

NODE	NODE
TYPE	COUNT
gpu:h100:2	9
gpu:pvc:4	9
gpu:h100:4	4
gpu:a30:2	2

AVAILABILITY

NODE	GPU	GPU	GPU	CPU	GB MEM
NAME	TYPE	COUNT	AVAIL	AVAIL	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 (138403) 1 node | 48 cores | Running

Host: [>_ac095](#) Delete

Created at: 2024-03-20 12:46:52 CDT

Time Remaining: 2 hours and 56 minutes

Session ID: [f55b9a74-fb5e-4cea-b988-0b268204f35b](#)

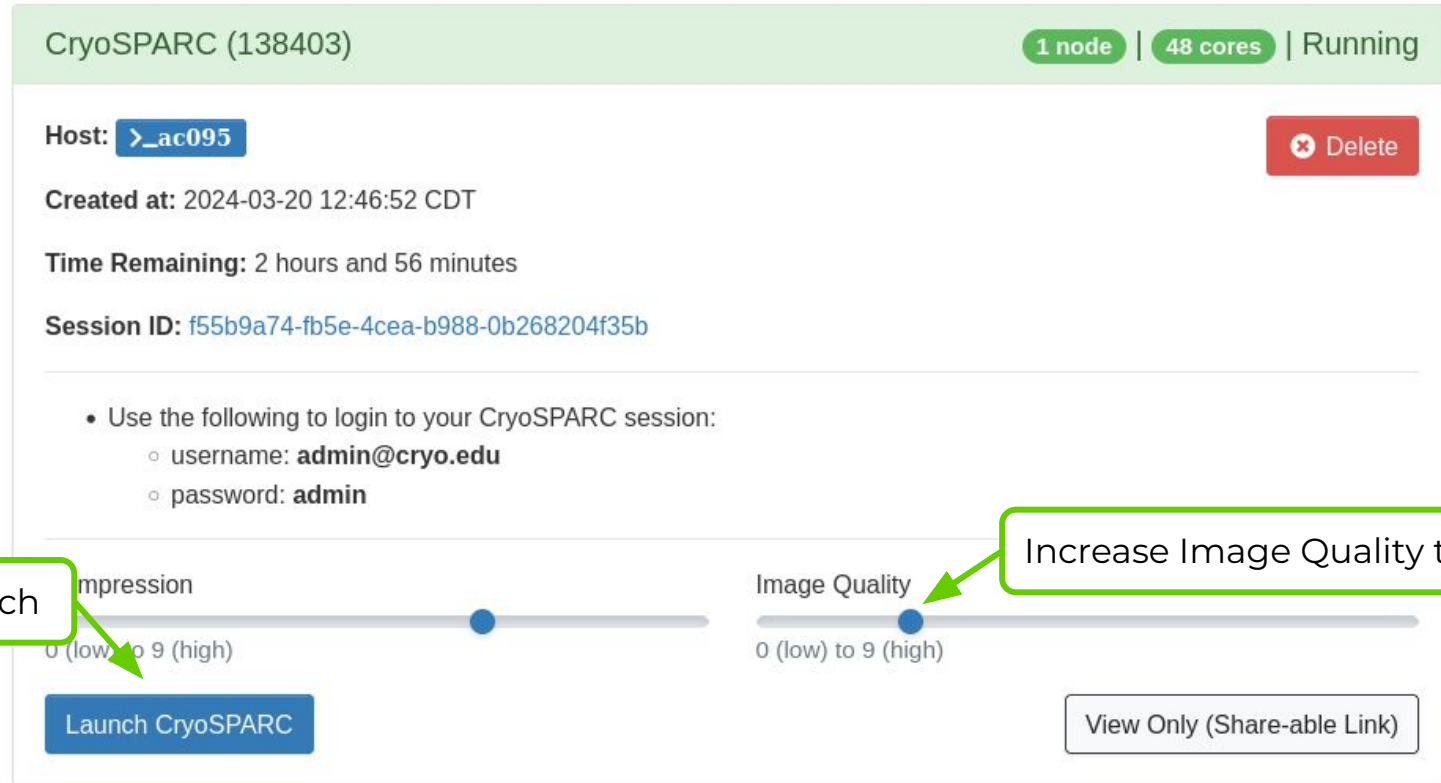
- Use the following to login to your CryoSPARC session:
 - username: **admin@cryo.edu**
 - password: **admin**

Impression 0 (low) to 9 (high)

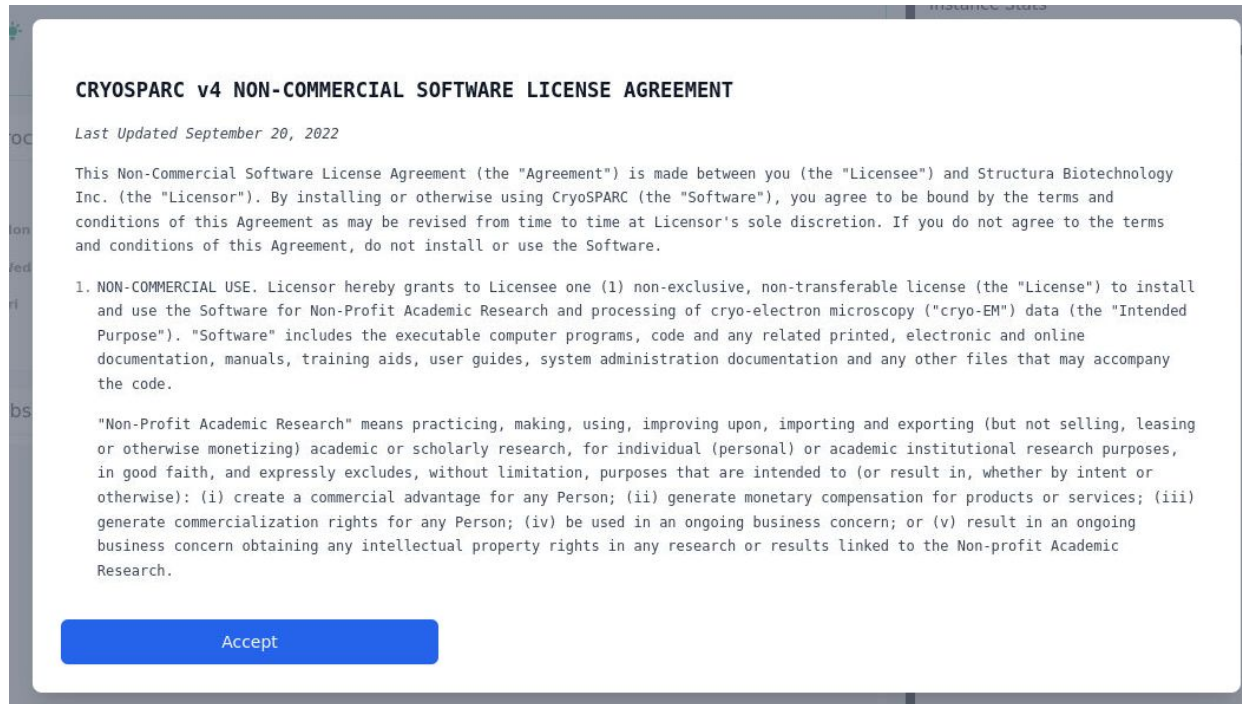
Image Quality 0 (low) to 9 (high)

Click Launch Increase Image Quality to 9

Launch CryoSPARC View Only (Share-able Link)

The image shows a web interface for a CryoSPARC job. At the top, it displays 'CryoSPARC (138403)' with status indicators for '1 node', '48 cores', and 'Running'. Below this, the host is identified as '>_ac095' with a 'Delete' button. The job was created on 2024-03-20 at 12:46:52 CDT and has 2 hours and 56 minutes remaining. The session ID is 'f55b9a74-fb5e-4cea-b988-0b268204f35b'. A list of instructions for logging in is provided, including the username 'admin@cryo.edu' and password 'admin'. At the bottom, there are two sliders: 'Impression' and 'Image Quality', both ranging from 0 (low) to 9 (high). A blue 'Launch CryoSPARC' button is positioned below the 'Impression' slider, and a 'View Only (Share-able Link)' button is to the right. Two green callout boxes with arrows point to the 'Launch CryoSPARC' button and the 'Image Quality' slider, with the text 'Click Launch' and 'Increase Image Quality to 9' respectively.

Read and Accept the License



CRYOSPARC v4 NON-COMMERCIAL SOFTWARE LICENSE AGREEMENT

Last Updated September 20, 2022

This Non-Commercial Software License Agreement (the "Agreement") is made between you (the "Licensee") and Structura Biotechnology Inc. (the "Licensor"). By installing or otherwise using CryoSPARC (the "Software"), you agree to be bound by the terms and conditions of this Agreement as may be revised from time to time at Licensor's sole discretion. If you do not agree to the terms and conditions of this Agreement, do not install or use the Software.

1. NON-COMMERCIAL USE. Licensor hereby grants to Licensee one (1) non-exclusive, non-transferable license (the "License") to install and use the Software for Non-Profit Academic Research and processing of cryo-electron microscopy ("cryo-EM") data (the "Intended Purpose"). "Software" includes the executable computer programs, code and any related printed, electronic and online documentation, manuals, training aids, user guides, system administration documentation and any other files that may accompany the code.

"Non-Profit Academic Research" means practicing, making, using, improving upon, importing and exporting (but not selling, leasing or otherwise monetizing) academic or scholarly research, for individual (personal) or academic institutional research purposes, in good faith, and expressly excludes, without limitation, purposes that are intended to (or result in, whether by intent or otherwise): (i) create a commercial advantage for any Person; (ii) generate monetary compensation for products or services; (iii) generate commercialization rights for any Person; (iv) be used in an ongoing business concern; or (v) result in an ongoing business concern obtaining any intellectual property rights in any research or results linked to the Non-profit Academic Research.

Accept

then exit and start a new CryoSPARC job this time with a GPU

ACES CryoSPARC Portal App

portal-aces.hprc.tamu.edu

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

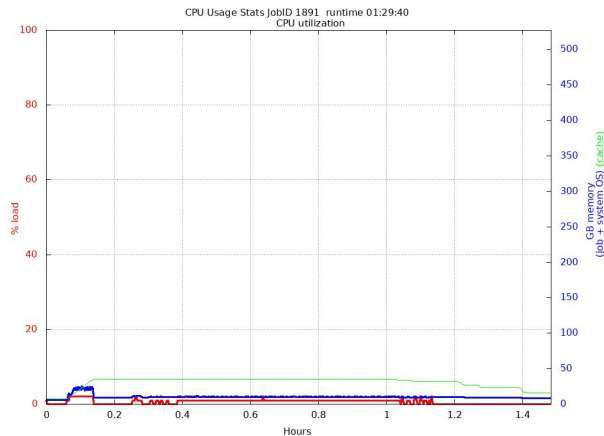
Delete your CryoSPARC "CPU only" job and launch a new CryoSPARC job using 1 x H100 GPU for the interactive tutorial

CryoSPARC Visualization of Resource Utilization

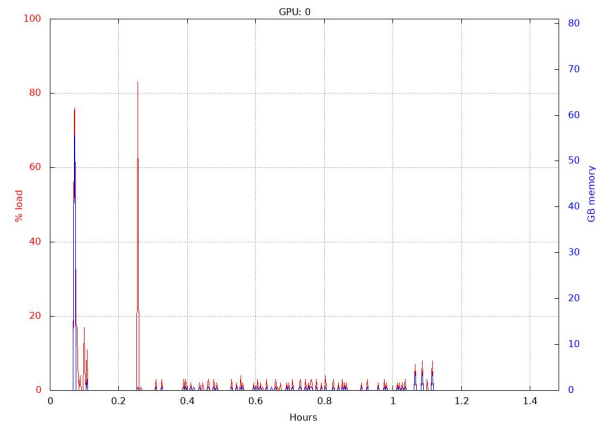
Review CPU and GPU usage for a Job

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 users to request the other GPUs on the compute node
- GPU stats will create a graph for each GPU that was configured for the job



stats_cpu.1891.png

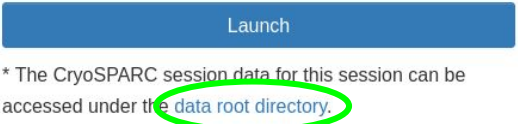
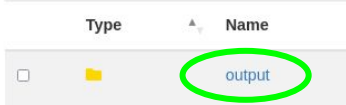
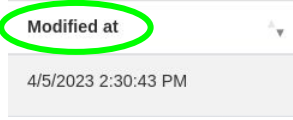
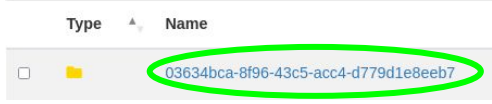



stats_gpu.1891.png

<https://hprc.tamu.edu/kb/Software/useful-tools/jobstats>

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
 - a. CPU usage stats are for the entire GPU node so it will include any other jobs running on that node
 - b. GPU stats are exclusive to your job

1. 
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

presented by:

Gaya Yadav

Cryo-EM Technical Director at the
Laboratory for Biomolecular
Structure and Dynamics