



Supercomputer Center



**Department of
Higher Education**

Ohio Technology Consortium



CryoSPARC Deployment at OSC: Lessons Learned and Challenges

Heechang Na

Scientific Applications Operations Manager

MW-RCD 2026, Rosemont, IL



**Department of
Higher Education**

Ohio Supercomputer Center

OUTLINE

- Introduction: Cryo-EM
- CryoSPARC: challenges on HPC
- Deployment with Virtual Machine (VM)
- Deployment with Kubernetes (K8s)
- Summary

INTRODUCTION: CRYO-EM

CRYOGENIC ELECTRON MICROSCOPY: CRYO-EM

- Imaging technique for 3D structures of biological molecules like proteins and viruses in atomic resolution
- **You cannot just take it in high resolution!**
 - Electron Energy bound
 - Biological samples
 - Image reconstruction
- **2017 Nobel prize in Chemistry**
 - Dubochet, Frank, and Henderson

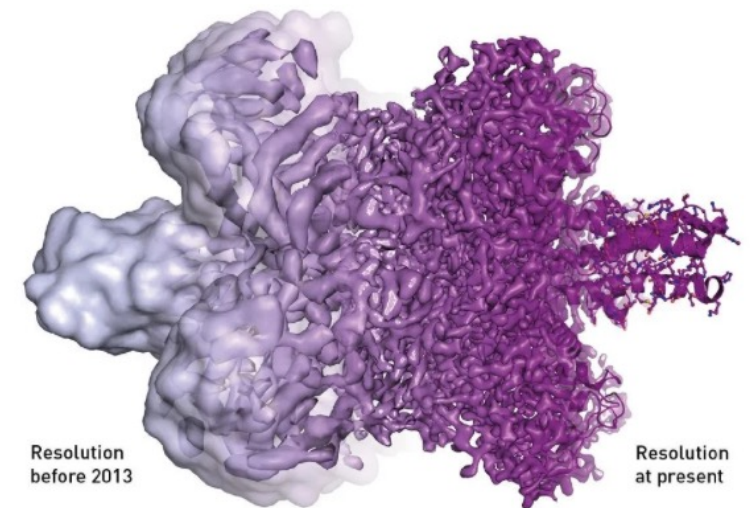


Figure 6. The electron microscope's resolution has radically improved in the last few years, from mostly showing shapeless blobs to now being able to visualise proteins at atomic resolution. © Martin Högbom/The Royal Swedish Academy of Sciences

CRYOSPARC

CRYOSPARC

- End-to-end computational pipeline tool for Cryo-EM data processing
 - Workflow from raw images to final 3D structures
 - Easy to use: intuitive GUI with web front end
 - Navigate the data and results in GUI
 - Drag and drop features
 - Submit jobs and monitor
 - Free option available for academia
 - Widely used with good community support

CRYOSPARCK

The screenshot displays the CryoSPARC web interface. At the top, the workspace is identified as 'P3 Soottikal_testing' under the 'Workspaces' tab. The main area shows a 'Data Processing' section with several job cards: 'J29 Curate Exposures', 'J28 Patch CTF', 'J26 Patch Motion Corr.', 'J25 Import Movies', and 'J23 3D Var'. Each card includes a thumbnail and a '27 days ago' timestamp. Below the jobs, a progress bar shows the status of various processing steps: Imports (2), Motion Correction (2), CTF Estimation (2), Exposure Curation (2), Particle Picking (5), Extraction (2), Particle Curation (4), 3D Reconstruction (1), 3D Refinement (2), Conformational Variability (2), Postprocessing (2), Utilities (1), Instance Testing Utilities (2), and Validation (4). A 'View 33 Jobs' button is located at the bottom of this section.

The right sidebar provides detailed information about the workspace:

- Details:** Title: Soottikal_testing; Tags: No tags applied; Created: Apr 2, 2026, 11:49:54 AM; Created By: Admin True; Created Version: v4.7.1; Last Accessed: Apr 29, 2026, 11:20:33 AM; Last Accessed By: Admin True; Directory: /fs/ess/PZS0710/cryosparc/CS-soottikal-testing
- Data Management:** Size: 61.09 GB; Intermediate Results Size: 5.27 GB; Size Last Updated: Apr 7, 2026, 4:20:41 PM
- Description:** (Collapsed)
- Sharing:** Not shared; Admin True is the Owner; A 'Share' button is available.

At the bottom of the sidebar, there are 'Actions' and 'View Project' buttons. The bottom status bar shows '1 Workspaces' and a user profile icon for 'AT'.

CRYOSPARCK INSTALLATION

- Master and worker configuration
 - Users can access the data and submit jobs from the master (server)
 - The jobs are running on the worker (compute nodes)
 - Two options: Standalone or HPC
- Challenges
 - Persistent access is desirable
 - Cannot share the instance with other groups
 - Raw data and job data are shared within an instance
 - Workflow with multiple jobs
 - A job with fixed resources is not optimal
 - Open OnDemand option

DEPLOYMENT WITH VM

CRYOSPARC WITH VM: OLD WAY

- **Deploy Master in VM for each group**
 - Required group accessible storage space: project space
 - Users submit Slurm jobs from the master
 - The Slurm jobs inherit the user identity of the master process running in the VM
 - Community account
 - Required license code from the research group
 - The instance is accessible from any web browser
 - OSC authentication before CryoSPARC authentication
- **Deploy Worker on each cluster**
 - As a software package for the research group, without a module
 - Other packages are provided separately: DeepEMhancer, Topaz

CRYOSPARC WITH VM: OLD WAY

- **Customized lanes for each group**
 - Lane: a Slurm template
 - Separated lanes for cpu/gpu/walltime at first
 - Flex lane available
- **Deployed 20 instances so far, and the requests are increasing**
 - Doesn't scale with VMs
 - Difficult to update
 - VM host is not flexible
- **K8s deployment would be an option!**

DEPLOYMENT WITH KUBERNETES

CRYOSPARC WITH K8S: NEW WAY

- **Deploying Master node on K8s**

- Containerized the Master node using a private Harbor registry
 - Instead of installing it separately for each group
- Provisioning via Puppet using prepared Helm charts
- Provided flexible lanes with Instance and Target defaults
- Ensured stateful data persistence for database and project files using PVCs
 - PVC: PersistentVolumeClaim
 - A shared project storage space is mounted as a PVC
 - Each group has its own PVC for its instance: Stateful data persistence

CRYOPARC WITH K8S: NEW WAY

- **Deploying Worker on cluster**
 - Single/unified installation serving all k8s Master instances
 - Consistent/standardized environments for all users
 - Additional packages are available for all users
 - DeepEMhancer, Topaz
 - Enables consistent documentation

CRYOSPARC WITH K8S: ADVANTAGES

- **Streamlined deployment pipeline via Puppet**
 - Identify the target project and community account
 - Configuration repository update for Puppet
 - Update config manifests with deployment details
 - CryoSPARC version, project/community account, port, license code
- **Flexible and scalable deployment**
 - No more manual installation required
 - Scalable version updates with minimal operational overhead
 - Flexible resource management
 - Efficient handling of instance failures

SUMMARY

CRYOSPARC DEPLOYMENT AT OSC

- **CryoSPARC is end-to-end pipeline tool for Cryo-EM data processing**
 - **Persistent master access is desirable**
 - **Traditional HPC model doesn't work well**
- **OSC deployed ~20 instances on VMs**
 - **Now, OSC is deploying on K8s cluster for scalable deployment**
 - **Next step: migration, user documentation, and fine-tuning**

THANK YOU

OSC.EDU



**Department of
Higher Education**

Ohio Supercomputer Center





Supercomputer Center



**Department of
Higher Education**

Ohio Technology Consortium

OSC.EDU