cryoSPARC Architecture & Management

A brief overview
Overview

• cryoSPARC System Architecture
• cryoSPARC Hardware Requirements
• Technical Troubleshooting in cryoSPARC
• Data Management
cryoSPARC System Architecture
Master-Worker

- cryoSPARC Master
  - cryoSPARC Web Application Server
  - cryoSPARC Command Core Server
  - cryoSPARC MongoDB Database

- cryoSPARC Worker
  - cryoSPARC GPU Compute Libraries
Typical cryoSPARC System Setups

- **Single Workstation** *(simplest)*
  - both the cryoSPARC master and cryoSPARC worker processes are run on the same machine.

- **Master- Worker** *(most flexible)*
  - the cryoSPARC master is installed on a lightweight machine, and the worker processes are installed on one or more GPU servers

- **Master with a Cluster integration** *(most scalable)*
  - the master node submits jobs to a cluster scheduler system (e.g., Slurm Workload Manager).
Single Workstation

Single cryoSPARC Workstation

- CPU
- Memory
- SSD Storage

Bulk Storage

NVIDIA GPU (0)
NVIDIA GPU (1)
NVIDIA GPU (2)
NVIDIA GPU (3)
Master-Worker cryoSPARC Setup

Master Instance
- CPU
- Memory
- SSD Storage

Worker Instance 1
- CPU
- Memory
- SSD Storage
- NVIDIA GPU (0)
- NVIDIA GPU (1)
- NVIDIA GPU (2)
- NVIDIA GPU (3)

Worker Instance 2
- CPU
- Memory
- SSD Storage
- NVIDIA GPU (0)
- NVIDIA GPU (1)
- NVIDIA GPU (2)
- NVIDIA GPU (3)

Bulk Storage

Master-Worker

Single Particle Cryo-EM Image Analysis Workshop

June 2020
Cluster Integration

Single Particle Cryo-EM Image Analysis Workshop

June 2020
Master- Worker Requirements

1) All nodes have access to a shared file system.
   - This file system is where the project directories are located, allowing all nodes to read and write intermediate results as jobs start and complete.

2) The master node has password-less SSH access to each of the worker nodes.
   - SSH is used to execute jobs on the worker nodes from the master node.

3) All worker nodes have TCP access to 10 consecutive ports on the master node (default ports are 39000-39010).
   - These ports are used for metadata communication via HTTP Remote Procedure Call (RPC) based API requests.
Hardware Considerations

- CPU
- SSD
- Network
- GPU
Technical Troubleshooting

Debugging failed jobs in cryoSPARC

- Inspect the Job’s Streamlog
- Inspect the logs from the command line
- Use the discussion forum
- Reach out to us!
Data Management in cryoSPARC
Archiving and Exporting Results

- Sharing within an instance
- Exporting and Archiving a Project
- Exporting a Job
- Exporting Particles (as an output result group)
- Exporting Particles (to use in another program)
Thank you!

Questions?

Unlock the potential of cryo-EM with cryoSPARC™

CryoSPARC is a state-of-the-art scientific software platform for cryo-electron microscopy (cryo-EM) used in research and drug discovery pipelines.

Get cryoSPARC™
Version 2.15
Released: May 13, 2020