make history.

THE UNIVERSITY of ADELAIDE

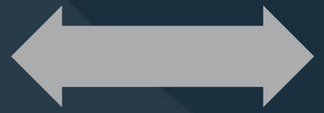
Setting up Cryo-EM tools and workflows on HPC Phoenix – lessons learnt

Dr Fabien Voisin

Who are we?

Adelaide Microscopy
Fiona Whelan
Ashley Slattery
Bryant Roberts

Division Research and Infrastructure
Fabien Voisin
Bowen Chen



Information Technology and digital services
Jason Tan
Tim Bowen
Peter Koleff

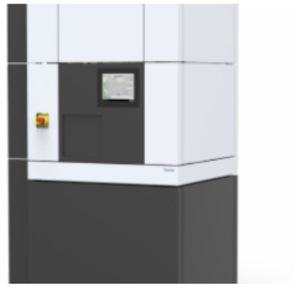
THE UNIVERSITY of ADELAIDE

Cryo-TEM requirements

- Large data acquisitions (TBs of raw data)
- Live data analysis to check quickly the quality of the data before further analysis
- Installation of software for data post-processing (e.g Cryosparc, Relion…), GPUs are required!
- Sending the data/results securely to collaborators outside the University

THE UNIVERSITY of ADELAIDE

UoA Cryo-TEM Microscope



FEI Glacios 200kV Cryo-Transmission Electron Microscope

Please contact Dr Fiona Whelan for further information.
Located at Frome Road.

Collapse All

Information on the FEI Glacios 200kV Cryo-Transmission Electron Microscope

The Thermo Scientific Glacios Cryo-Transmission Electron Microscope (Cryo-TEM) delivers a complete cryo-electron microscopy (cryo-EM) solution to a broad range of scientists. It features 200 kV XFEG optics, the industry-leading Autoloader (cryogenic sample manipulation robot), and the same innovative automation for ease of use as on the Thermo Scientific Krios G4 Cryo-TEM. The Glacios Cryo-TEM bundles all this into a smaller footprint that simplifies installation.

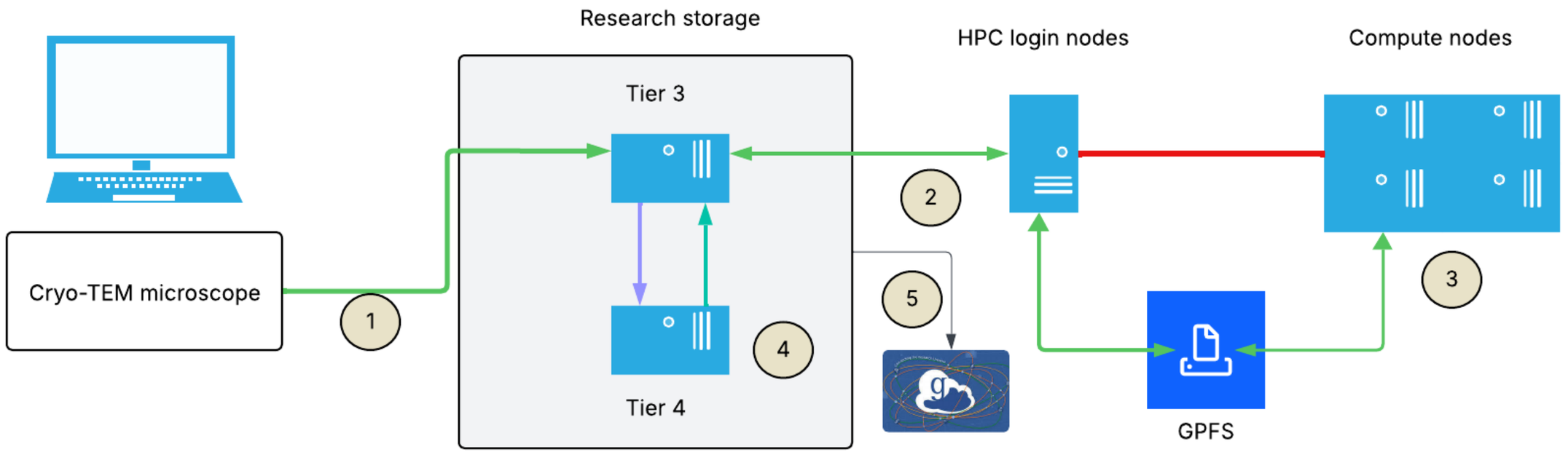
HPC Phoenix

Cluster	Node Type	Nodes	Node Specification	Total Specification	Performance	GoLive
Phoenix HPC1	CPU	105	2x Intel(R) Xeon(R) Platinum 8360Y CPU, 36 cores @ 2.4GHz, 255000 MiB memory(around 256 GB per node)	7560 CPUs, 26 TB of memory	N/A	2022-2023
	CPU High Memory	8	2x Intel(R) Xeon(R) Platinum 8360Y CPU, 36 cores @ 2.4GHz, 1990000 MiB memory (around 2TB per node)	576 CPUs, 16 TB of memory	N/A	2022-2023
	GPU	50	2x Intel(R) Xeon(R) Platinum 8360Y CPU, 36 cores @ 2.4GHz, 515000 MiB memory, 4x Nvidia A100-SXM4-40GB	3600 CPU, 200 GPUs, 25 TB of memory	N/A	2022-2023
SYSTEM TOTAL		163		**11736 cores, 200 GPU accelerators, 242 TB memory**		

Leverage the capabilities of our existing HPC to accelerate analysis and outputs

THE UNIVERSITY of ADELAIDE

High level architecture



1. Data acquisition is copied from CryoTEM workstation to Research storage
2. Data is continuously pulled from Research storage onto HPC GPFS.
3. Data processing is done onto HPC compute nodes and results transferred back to Research storage.
4. Inactive data sent to lower storage tier.
5. Data sent to external collaborators through GLOBUS

THE UNIVERSITY of ADELAIDE

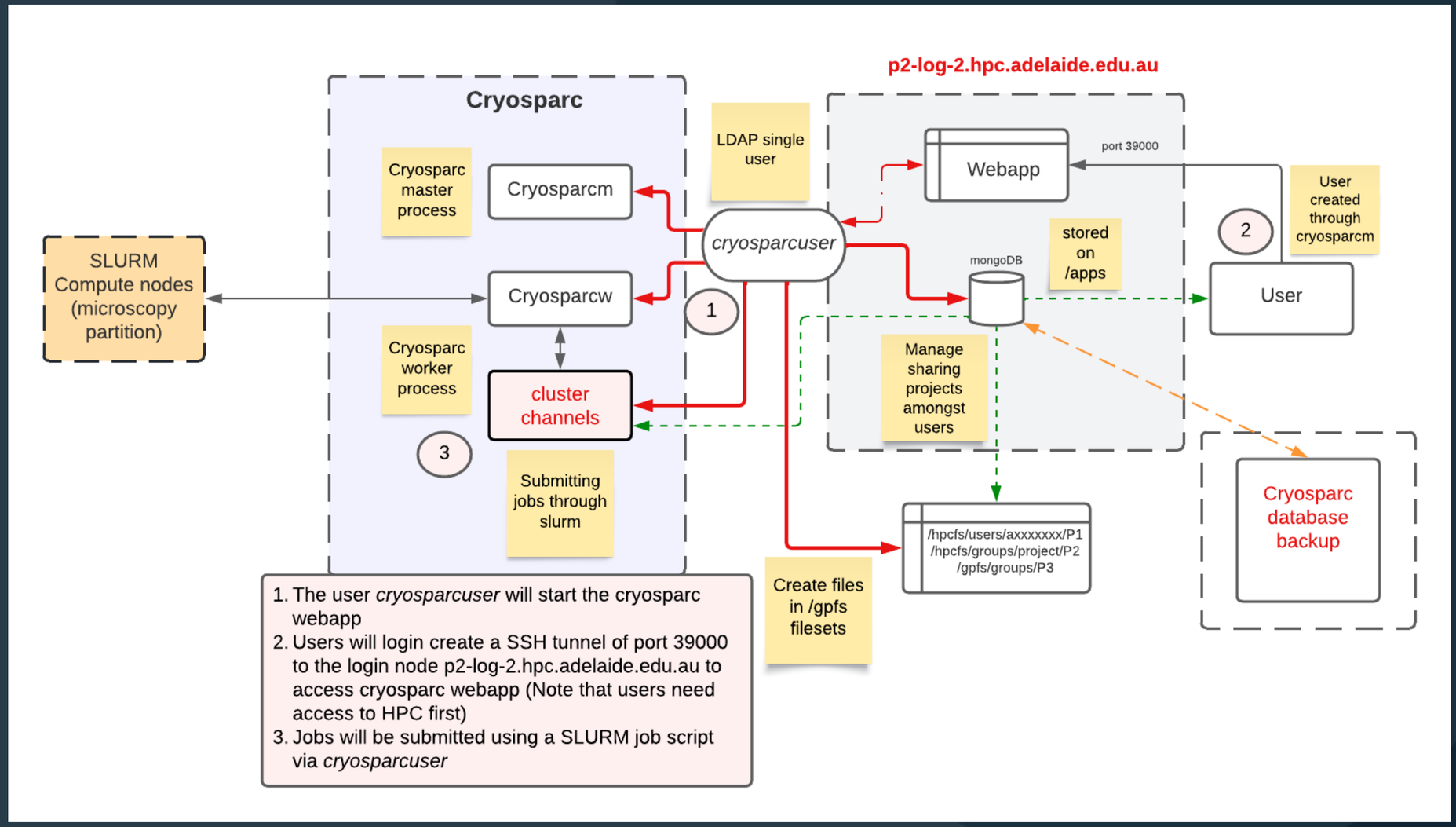
Cryo-TEM data processing apps in HPC

Cryosparc
• Data analysis solution for single-particle analysis. • Machine learning approach to particle picking with **Topaz** • https://guide.cryosparc.com/

RELION
• Software package that employs an empirical Bayesian approach for electron cryo-microscopy (cryo-EM) structure determination • https://relion.readthedocs.io/en/release-5.0/

THE UNIVERSITY of ADELAIDE

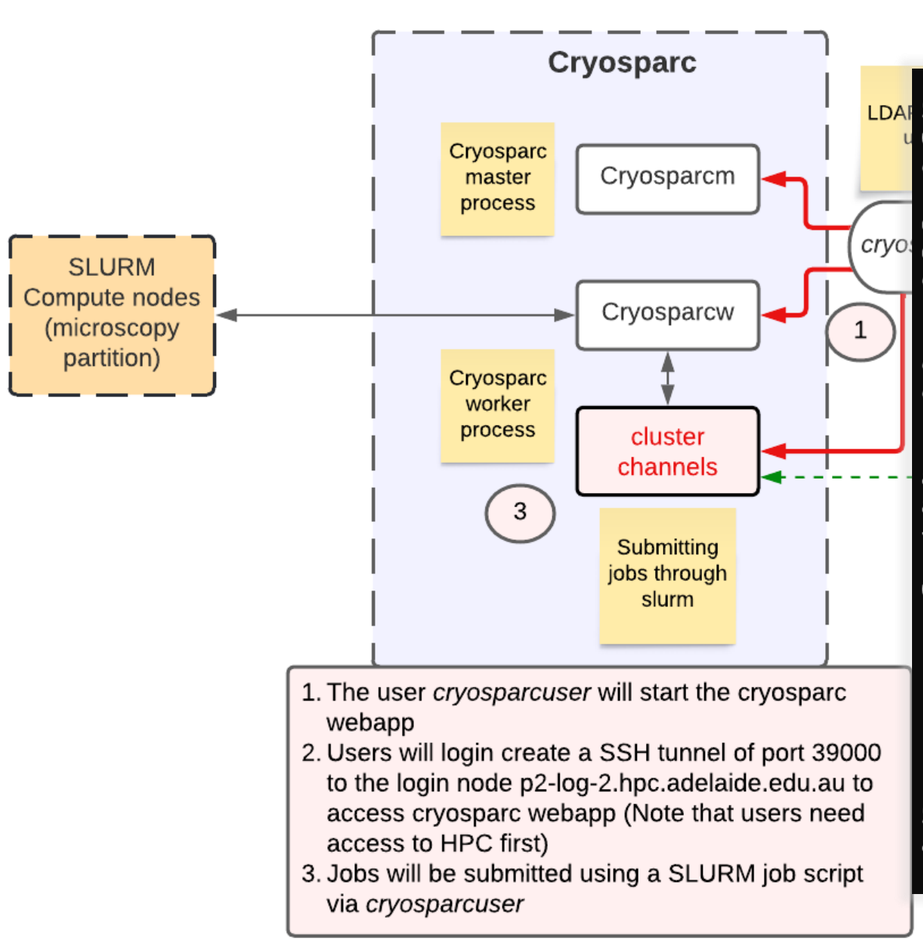
Cryosparc application high level setup



THE UNIVERSITY of ADELAIDE

Cryosparc high level setup

1 Starting up the cryosparc webapp



1. The user *cryosparcuser* will start the cryosparc webapp
2. Users will login create a SSH tunnel of port 39000 to the login node p2-log-2.hpc.adelaide.edu.au to access cryosparc webapp (Note that users need access to HPC first)
3. Jobs will be submitted using a SLURM job script via *cryosparcuser*

p2-log-2.hpc.adelaide.edu.au

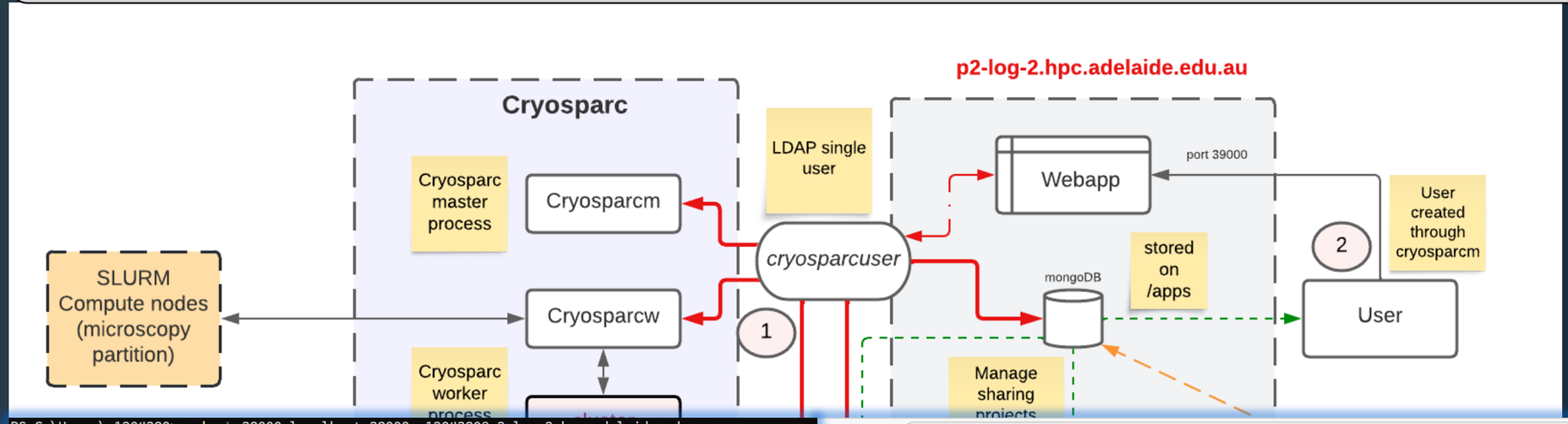
```
(base) [cryosparcuser@p2-log-2 ~]$ cryosparcm start
Starting CryoSPARC System master process...
CryoSPARC is not already running.
configuring database...
    configuration complete
database: started
database OK
command_core: started
    command_core connection succeeded
    command_core startup successful
command_vis: started
command_rtp: started
    command_rtp connection succeeded
    command_rtp startup successful
app: started
app_api: started
---------------------------------------------------

CryoSPARC master started.
 From this machine, access CryoSPARC and CryoSPARC Live at
    http://localhost:39000

 From other machines on the network, access CryoSPARC and CryoSPARC Live at
    http://p2-log-2:39000


Startup can take several minutes. Point your browser to the address
and refresh until you see the CryoSPARC web interface.
(base) [cryosparcuser@p2-log-2 ~]$
```

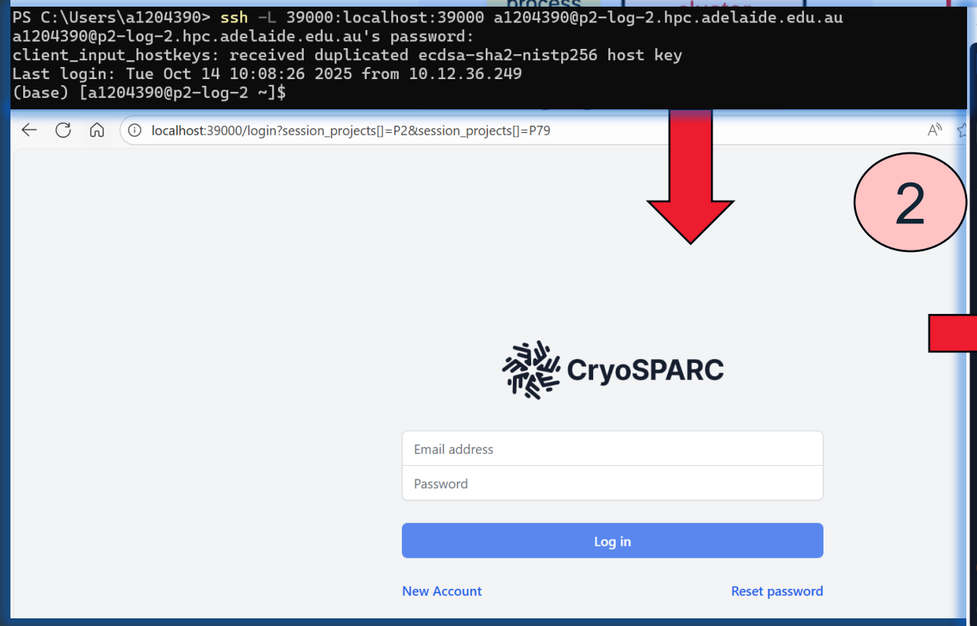
1

THE UNIVERSITY of ADELAIDE

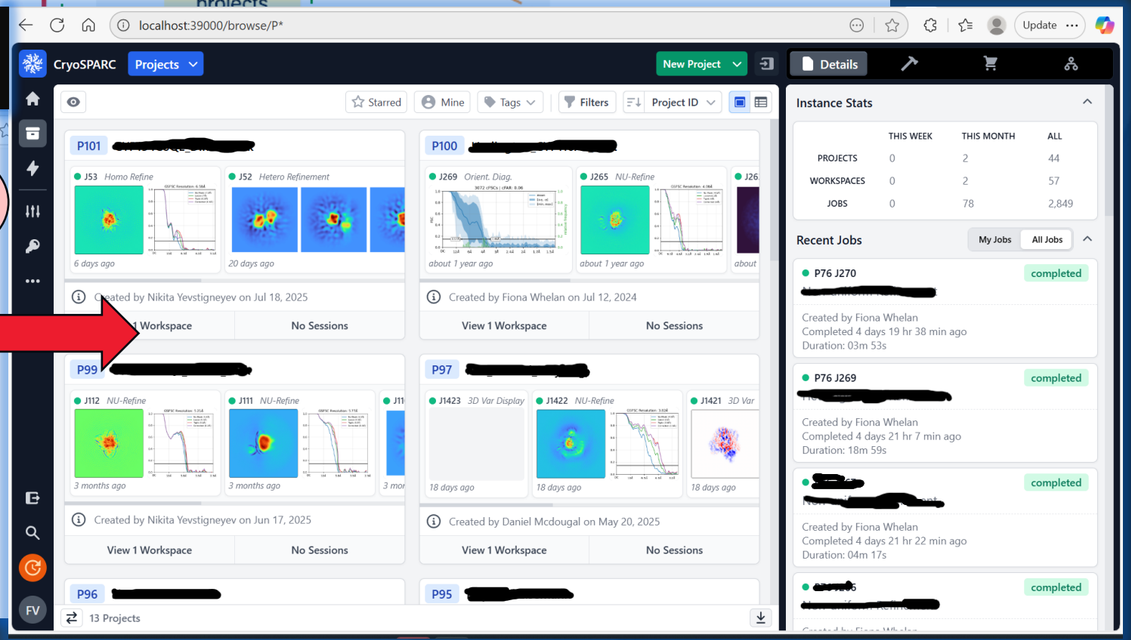
Cryosparc high level setup

2 Accessing the cryosparc webapp through SSH tunneling





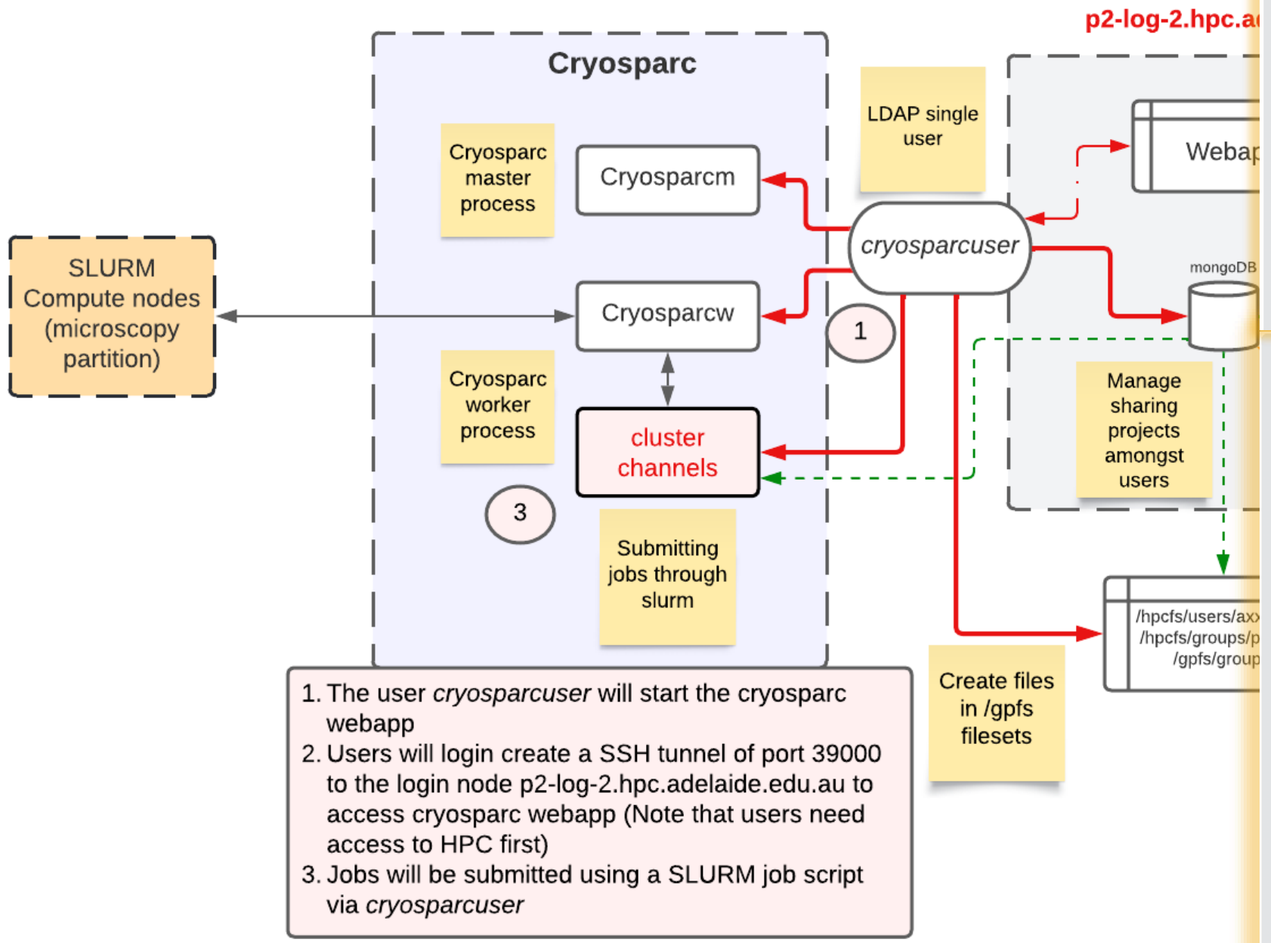
2



THE UNIVERSITY of ADELAIDE

Cryosparc high level setup

3 Configure a cluster lane for Cryosparc



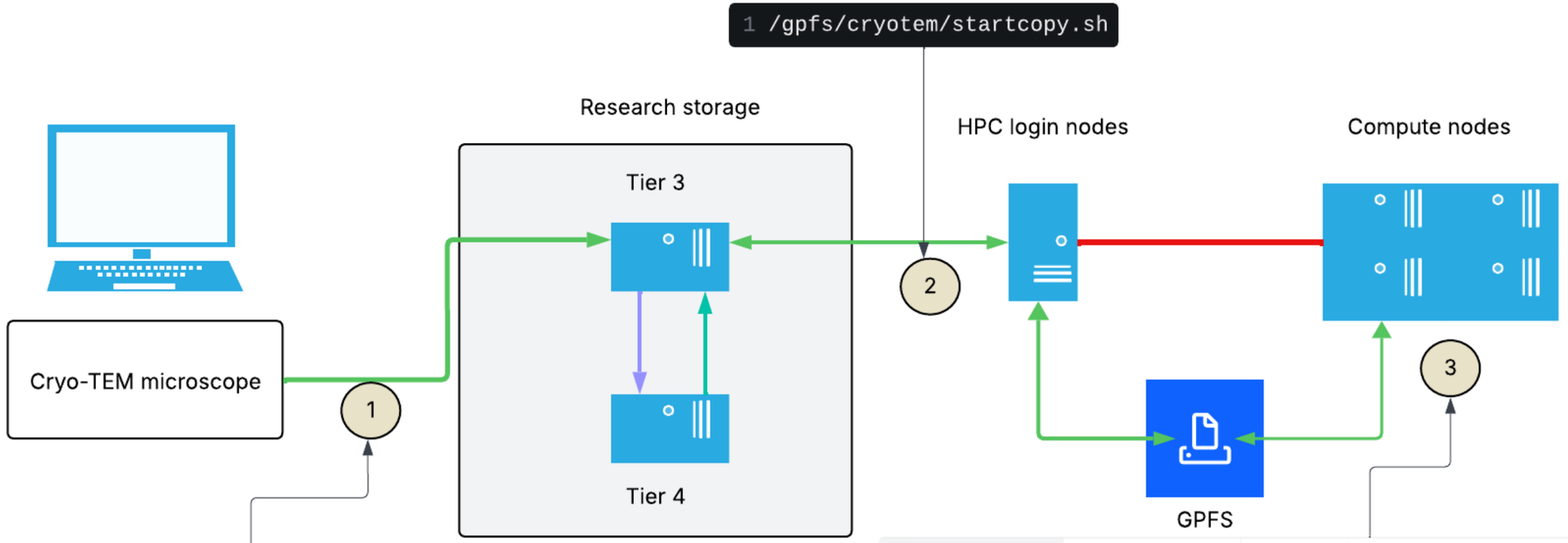
3

```
(base) [a1204390@p2-log-1 hpc1]$ cat cluster_info.json
{
    "name": "hpc1",
    "title": "hpc1",
    "worker_bin_path":
"/apps/skl/software/CryoSPARC/3.2/cryosparc_worker/bin/cryosparcw",
    "send_cmd_tpl": "{{ command }}",
    "qsub_cmd_tpl": "sbatch {{ script_path_abs }}",
    "qstat_cmd_tpl": "squeue -j {{ cluster_job_id }}",
    "qdel_cmd_tpl": "scancel  {{ cluster_job_id }}",
    "qinfo_cmd_tpl": "sinfo --format='%.8N %.6D %.10P %.6T %.14C %.5c %.6z %.7m %.7G
%.9d %20E'",
    "cache_path": null,
    "cache_quota_mb": null,
    "cache_reserve_mb": 10000
}
```

```
(base) [a1204390@p2-log-1 hpc1]$ cat cluster_script.sh
#!/bin/bash
#SBATCH --job-name=cryosparc_{{ project_uid }}_{{ job_uid }}
#SBATCH --partition=microscopy
#SBATCH --output={{ job_log_path_abs }}
#SBATCH --error={{ job_log_path_abs }}
#SBATCH --nodes=1
#SBATCH --mem={{ (ram_gb*6400)|int }}MB
#SBATCH --ntasks-per-node=1
#SBATCH --cpus-per-task={{ num_cpu }}
#SBATCH --gres=gpu:{{ num_gpu }},tmpfs:500G
#SBATCH --gres-flags=enforce-binding
#SBATCH --time=2-00:00:00
export CRYOSPARC_SSD_PATH=${TMPDIR}
module load CUDA/12.2.2
module load  CUDAcompat/12.2-535.161.08
srun {{ run_cmd }}
```

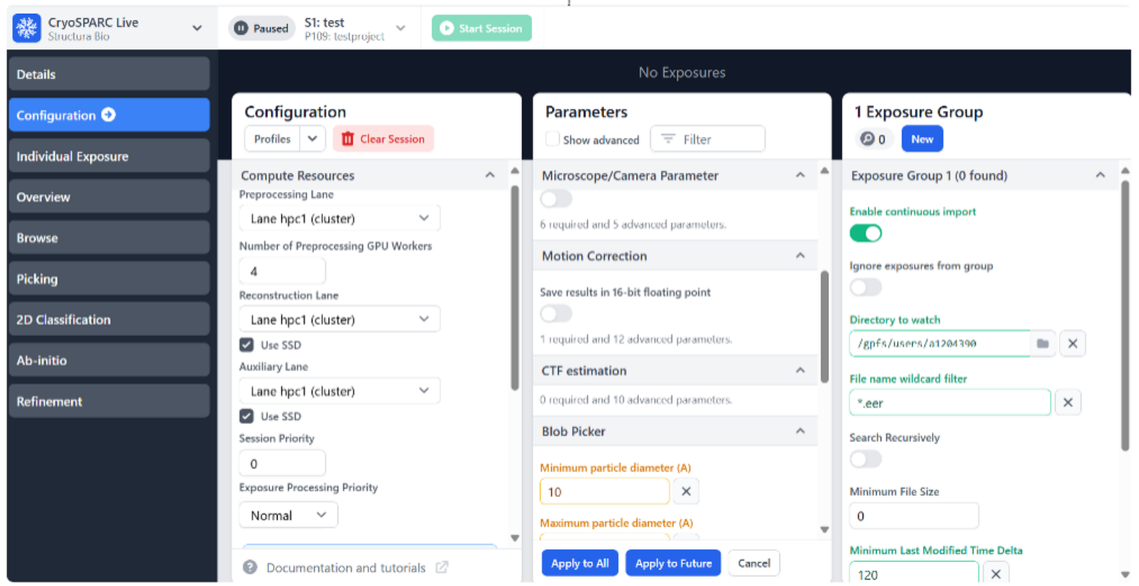
```
(base) [cryosparcuser@p2-log-1 ~]$ cd /apps/skl/software/CryoSPARC/3.2/cryosparc_master/hpc1/
(base) [cryosparcuser@p2-log-1 hpc1]$ cryosparcm cluster connect
```

Cryosparclive…setup



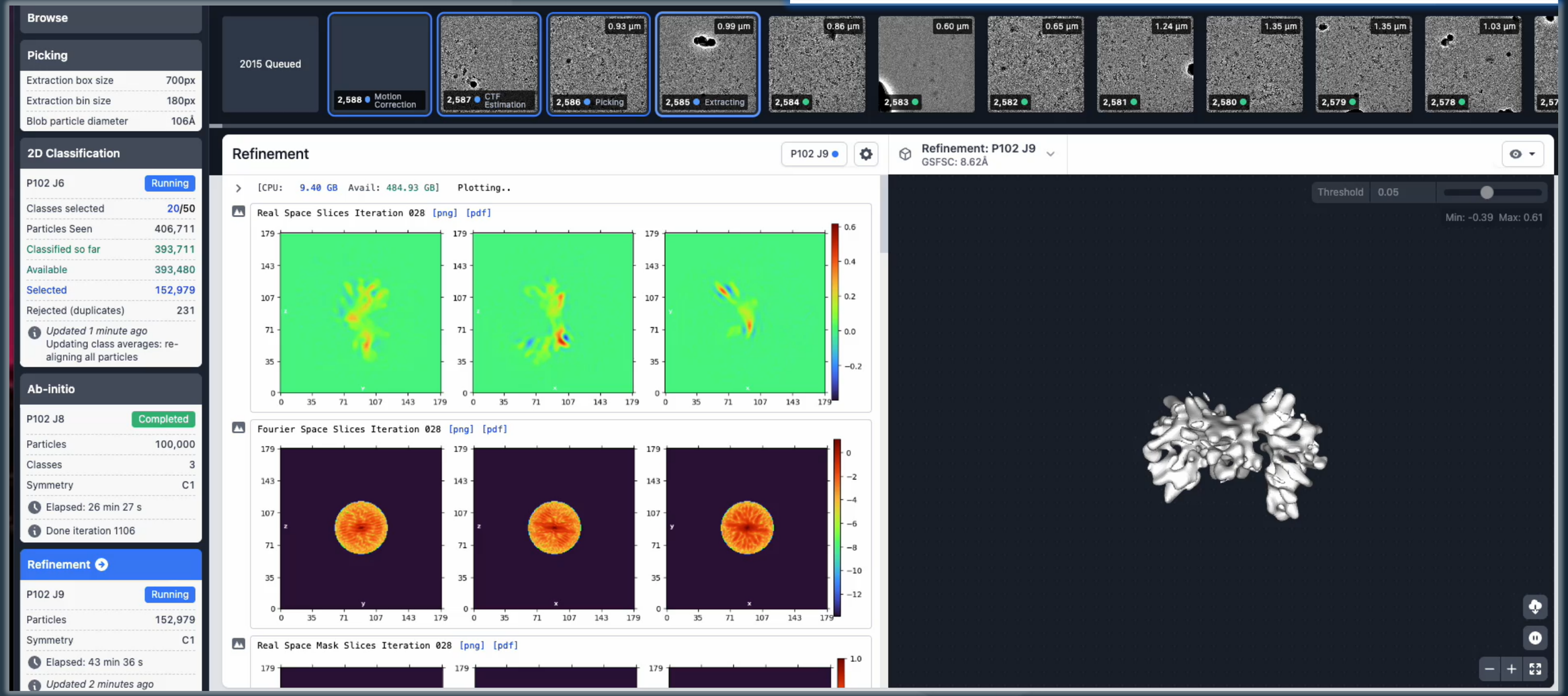

```
$Date = Get-Date -format yyyyMMdd

robocopy $SourceFolder $DestinationFolder /E /R:10 /W:5 /MT /maxage:$date /MOT:2
```

HPC copy scripts available in Appendix



Cryosparclive…live

```
(base) [cryosparcuser@p2-log-2 cryosparcuser]$ squeue --me
             JOBID PARTITION     NAME     USER ST       TIME  NODES NODELIST(REASON)
           6043054 microscop cryospar cryospar  R    2:35:43      1 p2-gpu-39
           6043062 microscop cryospar cryospar  R    2:35:43      1 p2-gpu-39
           6043099 microscop cryospar cryospar  R    2:35:43      1 p2-gpu-40
           6043149 microscop cryospar cryospar  R    2:35:43      1 p2-gpu-40
           6050330 microscop cryospar cryospar  R       1:28      1 p3-gpu-42
           6050252 microscop cryospar cryospar  R      30:24      1 p3-gpu-41
           6043582 microscop cryospar cryospar  R    2:29:44      1 p3-gpu-41
```



Cryosparc on HPC

Advantage	Issues	Potential resolution?
✓ Instantaneous access to scalable compute resources (thanks to strategic reservation of compute nodes) + concurrent analysis.	⚠ Data transfer is still a ***manual*** process. Multiple steps required to ensure Cryosparclive can successfully fetch data	Looking at Active File Management (AFM) to cache the data from archival storage to GPFS.
✓ Access to high end GPUs with High performance scratch storage and 100+Gbps interconnect	⚠ Projects separated by users/group but *cryosparcuser* needs Posix access to all raw data that must be imported –> **users can potentially access other users input files.**	• Containerized Cryosparc? • LDAP integrated Cryosparc?
✓ Ability to split A100 GPUs into smaller multi-instance GPUs for increased throughput	✗ Cryosparc webapp currently needs IT or HPC personnel to restart in case something goes wrong -> **Potential delays**	• Systemd restart script • Setuid script allowing clean processes kill and service restart by Cryo-TEM admin
✓ Cryosparc webapp can be started in other login nodes in case one login node is down	✗ Overall, more control from microscopy staff and users over data tiering and access	

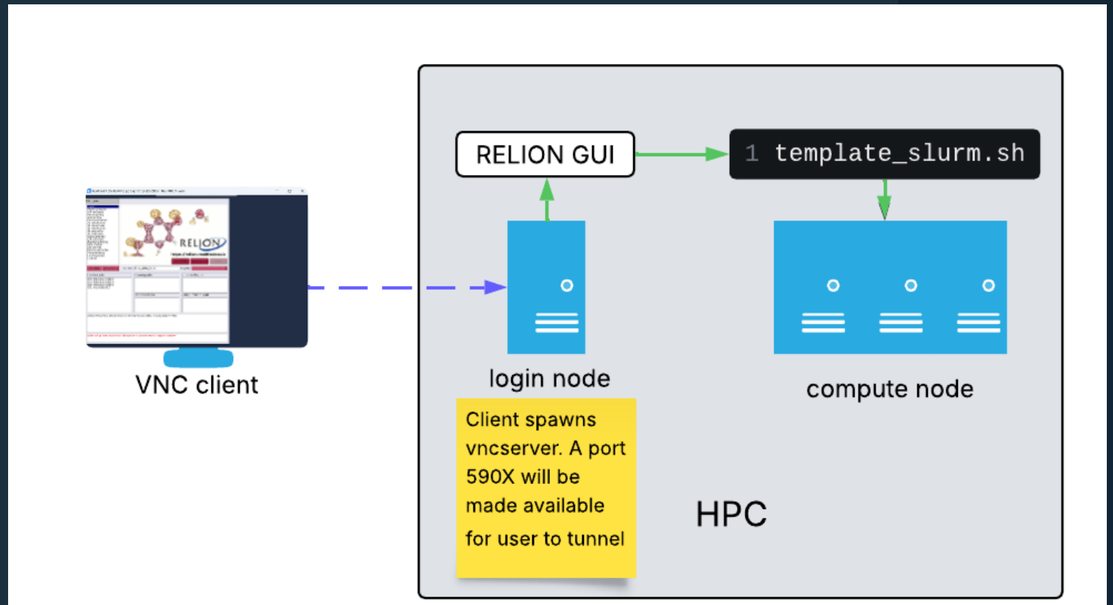
THE UNIVERSITY of ADELAIDE

Relion installation and setup

Installation and setup

- Requires VNC server (or anything that can provide a graphic interface) if user interface is a requirement.
- Used **EasyBuild** installation tools for dependencies but user personal conda environment for RELION installation.
- **Use template job scripts available to users for RELION GUI to submit jobs on compute nodes. Must add for multi-nodes and GPUS:**

```
export RELION_QSUB_EXTRA_COUNT=2
export RELION_QSUB_EXTRA1="Number of nodes:"
export RELION_QSUB_EXTRA1_DEFAULT="1"
export RELION_QSUB_EXTRA2="Number of GPUs:"
export RELION_QSUB_EXTRA2_DEFAULT="0"
```




```
(base) [a1204390@p2-log-1 cryotem]$ cat template_slurm.sh
#!/bin/bash
#SBATCH -N XXXnodesXXX
#SBATCH --ntasks-per-node=XXXmpinodesXXX
#SBATCH --partition=XXXqueueXXX
#SBATCH --error=XXXerrfileXXX
#SBATCH --gres=gpu:XXXextra2XXX

module load RELION
module load Ghostscript/9.56.1-GCCcore-11.3.0
export OMP_NUM_THREADS=XXXthreadsXXX
mpiexec -np XXXmpinodesXXX --bind-to-core --map-by
ppr:XXXmpinodesXXX:node:pe=XXXthreadsXXX -x OMP_NUM_THREADS  XXXcommandXXX
```

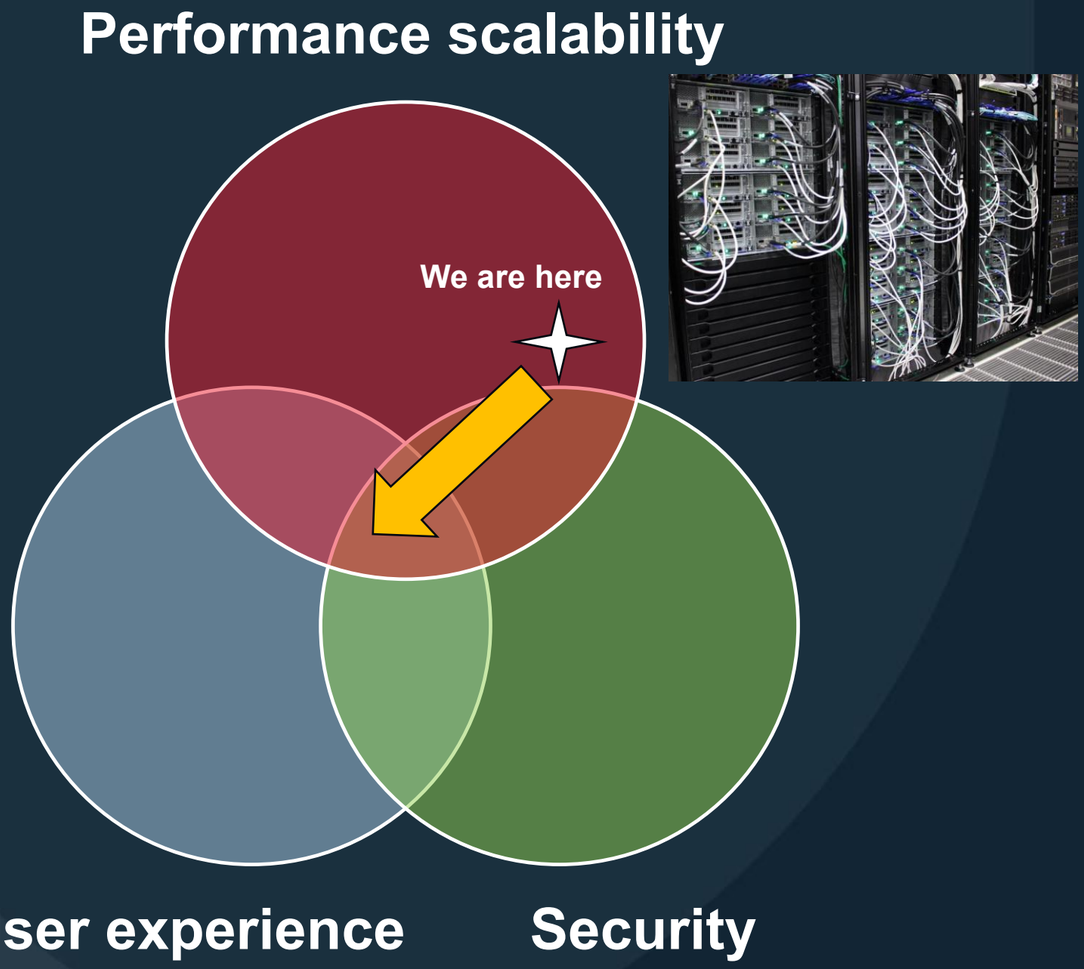
RELION GUI on HPC

Advantage	Issues	Potential resolution?
✓ Template job script good to avoid inefficient resources setup	⚠ VNC on login nodes can be tedious and with limited features (as opposed to proper VDIs)	Open on Demand?
✓ Leverage HPC application specialist for tedious software installation.	⚠ **Relion workflows may have dependencies to other modules and thus need constant support to ensure new jobs does not fail due to missing libraries/python packages.**	- Singularity image and template script to call singularity images to deal with complexity of installation

THE UNIVERSITY of ADELAIDE

The bigger picture….

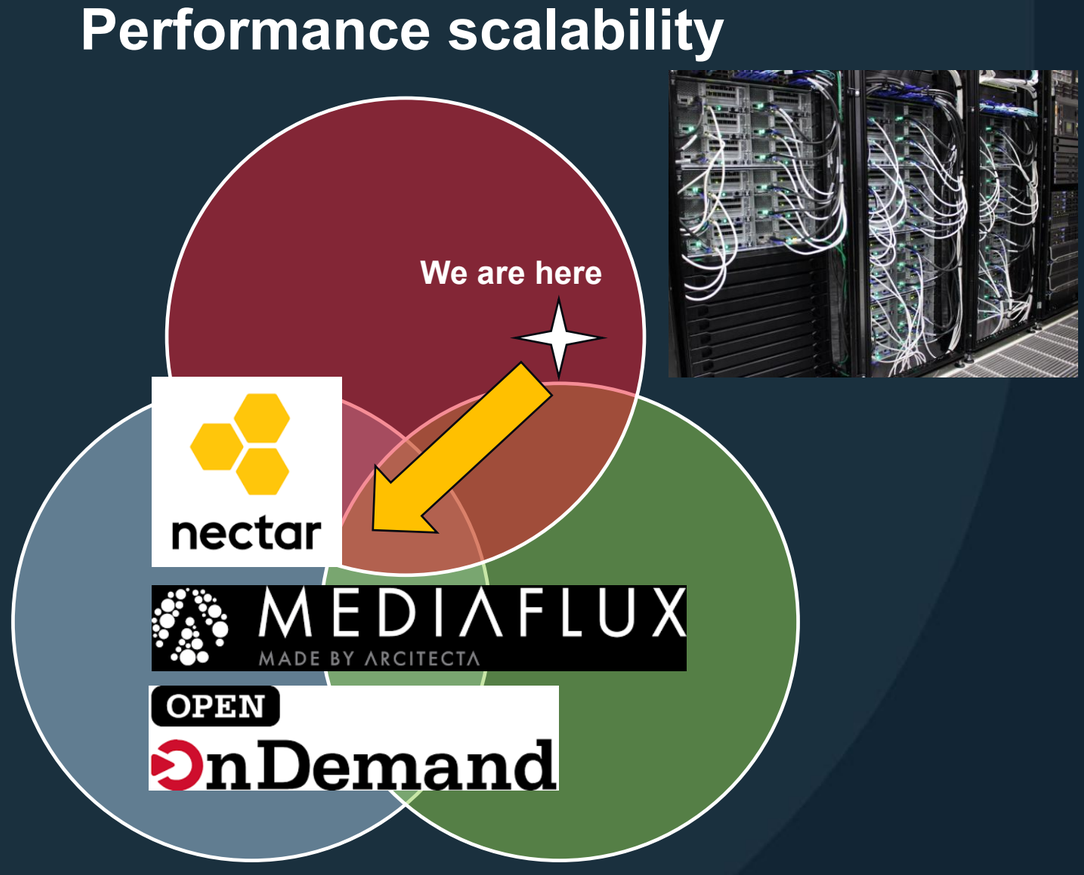
- Current setup shows great capability to scale.
- Steep learning curve for new users to fully take advantage of such infrastructure
  - Need scripts/solutions for better user experience in both compute and data management AND/OR
  - Sacrifice performance for better user experience AND/OR
  - Integrate with other e-infrastructure (e,g Nectar VMs)?



THE UNIVERSITY of ADELAIDE

The bigger picture….

- Current setup shows great capability to scale.
- Steep learning curve for new users to fully take advantage of such infrastructures
  - Need scripts/solutions for better user experience in both compute and data management AND/OR
  - Sacrifice performance for better user experience AND/OR
  - Integrate with other e-infrastructure (e,g Nectar VMs)?



THE UNIVERSITY of ADELAIDE

make history.

THE UNIVERSITY of ADELAIDE

Thank you

Appendix : HPC startcopy script

```
(base) [a1204390@p2-log-2 cryotem]$ cat startcopy.sh
#!/bin/bash

#Check if an existing session is running"
IS_RUNNING=$(screen -ls | grep robocopy)
ANSWER=""
if [[ ! ${IS_RUNNING} == "" ]]
then
    echo -e "It seems that robocopy instance is already running\nPlease stop robocopy
    exit 1
else
    while [[ ! ${ANSWER} =~ [nNyY] ]]
    do
        echo -e "Do you want to remove the content?[yY/nN]"
        read ANSWER
        if [[ ${ANSWER} =~ [yY] ]]
        then
            echo -e "Removing files now in /gpfs/cryotem/cryosparc_live"
            rm -rf /gpfs/cryotem/cryosparc_live/*
        elif [[ ${ANSWER} =~ [nN] ]]
        then
            echo "keeping the files but please make sure to eventually remove them\n"
        fi
    done
    echo "Starting robocopy in the background"
    screen -S robocopy -d -m  /gpfs/cryotem/livecopy.sh
fi
```

```
(base) [a1204390@p2-log-2 cryotem]$ cat livecopy.sh
#!/bin/bash
#This script will read the folder to copy, the config must be full path
TARGET="/gpfs/cryotem/cryosparc_live/"
while true
do
    sleep 10
FOLDERARRAY=($(grep \#FOLDER copypath.config | awk '{print $NF}'))
          for FOLDER in ${FOLDERARRAY[@]}
          do
              echo Start tranferring content ${FOLDER} into ${TARGET} >>/gpfs/cryotem/
              cp -ru ${FOLDER} ${TARGET} &
          done
wait
done
```

```
(base) [a1204390@p2-log-2 cryotem]$ cat copypath.config.example
#FOLDER /uofaresstor/cryotem/user_data/<examplepath>/
```

startcopy.sh :

1/ first verify whether the older livedata can be discarded.

2/ Create a *screen* session that copies the data from archival storage to HPC storage

3/ copypath.config helps for finer selection of folders to be copied across.

THE UNIVERSITY of ADELAIDE