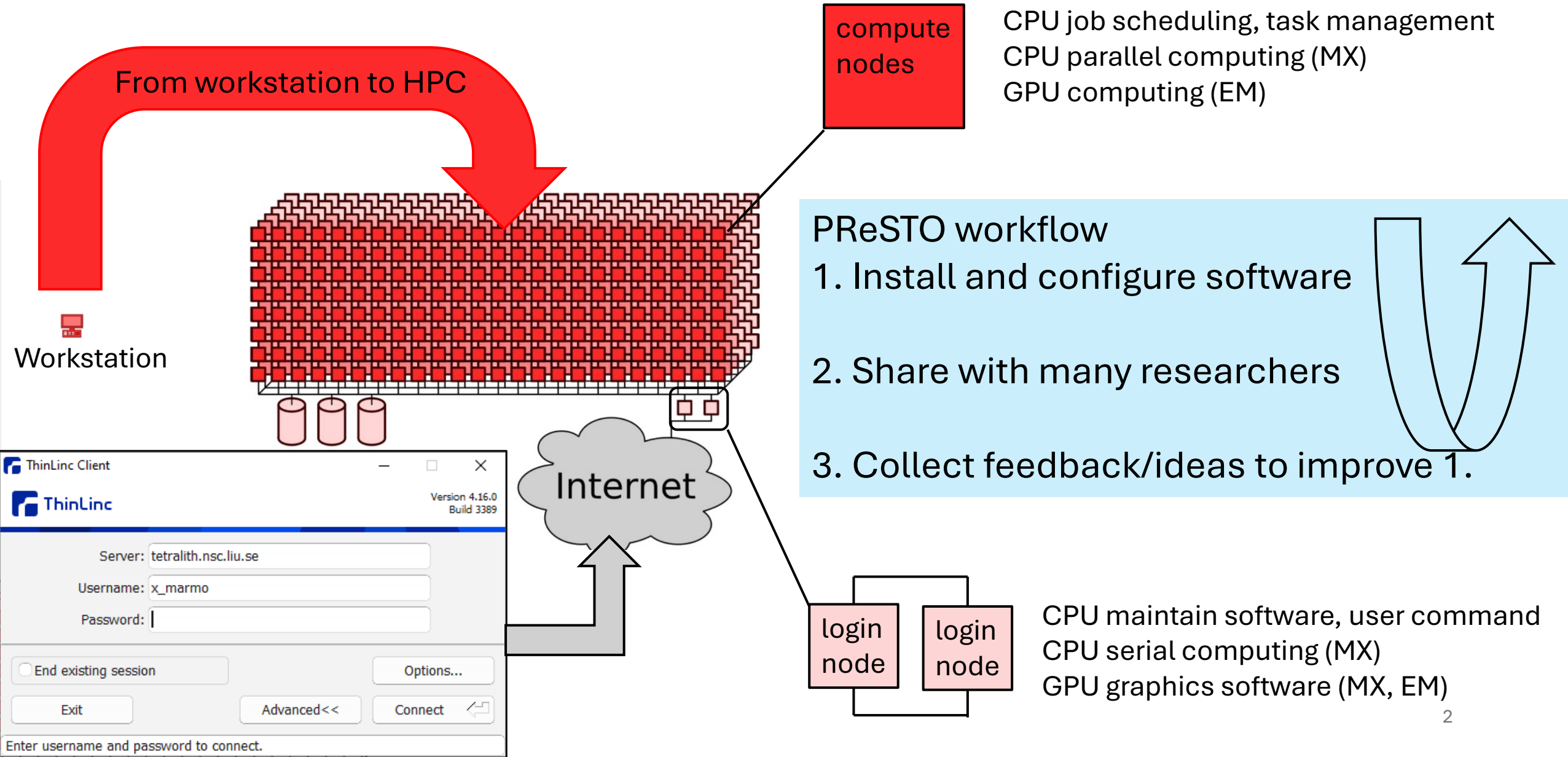


*PR*e*STO*

Rapid *PR*ot*E*in *ST*ructure c*O*mputation in an integrated *e*-*S*cience environment

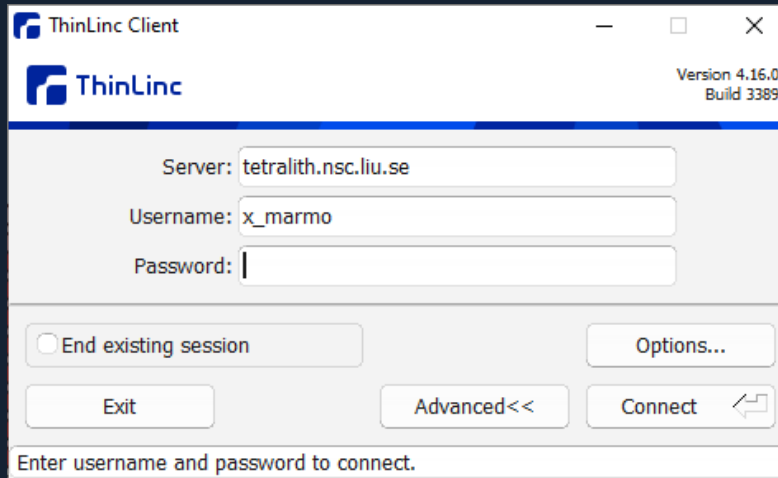
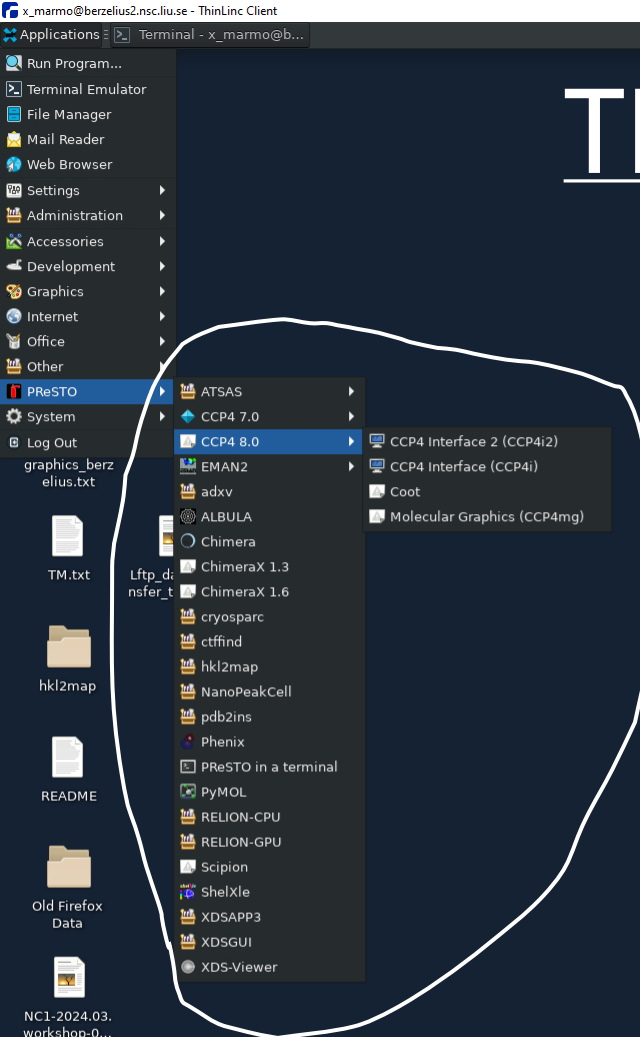
Enable structural biology software in Swedish HPC environments

PReSTO more about access than speed



Thinlinc Remote Desktop

- Available from any PC, Linux, MAC computer



PReSTO desktop menu
-software with GUIs

PReSTO available at four national clusters

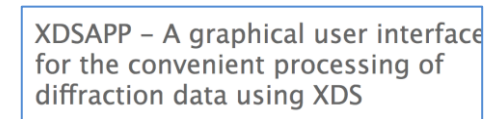
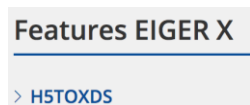
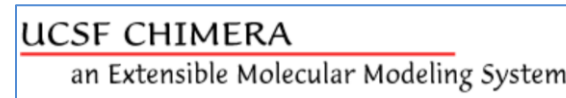
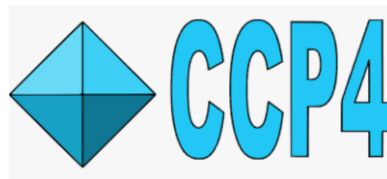
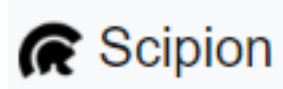
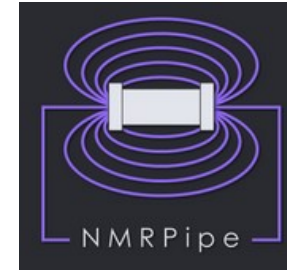
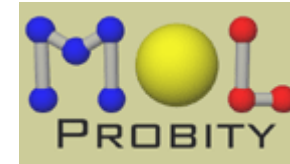
Software etc.	LUNARC (LU)	NSC (LiU)	NSC (LiU)	MAX IV (LU)
Resource	Cosmos	Tetralith	Berzelius	offline-fe1
Access	SUPR	SUPR	SUPR	DUO
EM - Relion	No	Yes	Yes	No
EM - CryoSPARC	Yes	Yes	Yes	No
EM - CryoDRGN	No	Yes	Yes	No
MX - CCP4	Yes	Yes	Yes ¹	Yes
XFEL - CrystFEL	Yes	Yes	Yes ¹	Yes
SAXS - ATSAS	Yes	Yes	Yes ¹	Yes
NMR - CCPN	No	Yes	Yes ¹	No

¹Installed for ISB course. Berzelius "GPU Usage Efficiency Policy" limit production use

PReSTO is a software distribution

1. Installation recipes for each application
2. Patches (modifications) for HPC environments
3. Application versions are grouped into **releases**
4. Desktop “PReSTO menu”
5. Contains around 60 applications for structural biology
 - a. MX (Macromolecular X-ray crystallography)
 - b. CryoEM (Cryo Electron Microscopy)
 - c. NMR (Nuclear Magnetic Resonance)
 - d. SAXS (Small Angle X-ray Scattering)
 - e. SSX (Serial Synchrotron X-ray crystallography)
 - f. XFEL (Xray Free Electron Laser)

Applications in PReSTO software distribution



Releases

- A particular set of applications and versions
- Allows for software testing prior to release
 - Are application versions compatible with each other?
- Multiple releases can be installed side-by-side
 - Users can choose to load older releases

Use of old software versions

PreSTO menu
load default modules

```
mochma@cosmos2:~  
[mochma@cosmos2 ~]$ module avail Phenix  
----- /sw/pkg/presto/modules/all -----  
Phenix/1.19.2-4158-Rosetta-3.10-PreSTO-11.4  
Phenix/1.20.1-4487-Rosetta-3.10-PreSTO-9.0  
Phenix/1.20.1-4487-Rosetta-3.10-PreSTO-9.2  
Phenix/1.20.1-4487-Rosetta-3.10-PreSTO-9.3  
Phenix/1.20.1-4487-Rosetta-3.10-PreSTO-9.4  
Phenix/1.21.1-5286-Rosetta-3.10-PreSTO-11.0  
Phenix/1.21.1-5286-Rosetta-3.10-PreSTO-11.1  
Phenix/1.21.1-5286-Rosetta-3.10-PreSTO-11.2  
Phenix/1.21.1-5286-Rosetta-3.10-PreSTO-11.3  
Phenix/1.21.1-5286-Rosetta-3.10-PreSTO-11.4 (D)  
Phenix/1.21.1-5207-Rosetta-3.10-PreSTO-10.0  
  
Where:  
D: Default Module
```

Phenix 1.19.2 for joint X-ray/Neutron refinement

```
Terminal - x_marmo@tetralith2:~  
File Edit View Terminal Tabs Help  
[x_marmo@tetralith2 ~]$ module avail PyMOL  
----- /software/presto/modules/all -----  
PyMOL/2.1.0-2-PreSTO  
PyMOL/2.3.0-1-PreSTO  
PyMOL/2.3.0-2-PreSTO  
PyMOL/2.5.0-PreSTO-toolchain-foss-2021a  
PyMOL/2.5.0-PreSTO-8.0  
PyMOL/2.5.0-PreSTO-9.0  
PyMOL/2.5.0-PreSTO-9.4  
PyMOL/2.5.0-PreSTO-9.6  
PyMOL/2.5.0-PreSTO-10.0  
PyMOL/2.5.0-PreSTO-11.0  
PyMOL/2.5.0-PreSTO-11.1  
PyMOL/2.5.0-PreSTO-11.2  
PyMOL/2.5.0-PreSTO-11.3  
PyMOL/2.5.0-1-PreSTO  
PyMOL/2.5.0-2-PreSTO  
PyMOL/2.5.0-3-PreSTO  
PyMOL/2.5.0-4-PreSTO  
PyMOL/2.5.0-5-PreSTO  
PyMOL/2.5.0-7-PreSTO  
----- /software/sse2/tetralith_el9/modules -----  
PyMOL/recommendation (D) PyMOL/2.5.0-hpcl-gcc-2022a-eb  
  
Where:  
D: Default Module  
[x_marmo@tetralith2 ~]$
```

Use of ATASAS-PyMol sas.py plugin

CryoSPARC Extensive validation test job

Lane (cluster submission script) name	berzelius mig	tetralith gpu	berzelius safe	cosmos gpua100i
Lane capacity GPU	1/7 NVIDIA A100 10GB VRAM...	1 NVIDIA T4	1 NVIDIA A100 40GB VRAM	1 NVIDIA A100 40GB VRAM
Lane CPU	2 cores	32 cores	16 cores	16 cores
Extensive Validation	01 hr. 17 min. 37 sec.	02 hr. 53 min. 43 sec.	48 min. 40 sec.	49 min. 01 sec.
Import Movies	22 sec.	56 sec.	23 sec.	24 sec.
Patch Motion Correction	06 min. 05 sec.	06 min. 34 sec.	03 min. 15 sec.	03 min. 41 sec.
Patch CTF Estimation	02 min. 06 sec.	02 min. 48 sec.	02 min. 15 sec.	02 min. 26 sec.
Manually Curate Exposures	09 sec.	17 sec.	09 sec.	09 sec.
Blob Picker	16 sec.	52 sec.	18 sec.	17 sec.
Inspect Particle Picks	08 sec.	13 sec.	06 sec.	05 sec.
Extract From Micrographs (CPU)	17 sec.	21 sec.	17 sec.	18 sec.
2D Classification (NEW)	03 min. 36 sec.	02 min. 37 sec.	01 min. 45 sec.	01 min. 45 sec.
Select 2D Classes	06 sec.	10 sec.	06 sec.	06 sec.
Template Picker	20 sec.	53 sec.	23 sec.	25 sec.
Inspect Particle Picks	07 sec.	12 sec.	07 sec.	06 sec.
Extract From Micrographs (CPU)	33 sec.	39 sec.	33 sec.	33 sec.
2D Classification (NEW)	06 min. 07 sec.	05 min. 04 sec.	04 min. 11 sec.	03 min. 58 sec.
Select 2D Classes	08 sec.	14 sec.	08 sec.	07 sec.
Particle Sets Tool				01 sec.
Ab-Initio Reconstruction	12 min. 38 sec.	11 min. 30 sec.	11 min. 30 sec.	09 min. 06 sec.
Homogeneous Refinement	12 min. 24 sec.	08 min. 55 sec.	06 min. 52 sec.	05 min. 09 sec.
Non-uniform Refinement	21 min. 57 sec.	18 min. 15 sec.	07 min. 06 sec.	07 min. 13 sec.
3D Classification	02 min. 16 sec.	02 min. 41 sec.	02 min. 06 sec.	02 min. 19 sec.
3D Variability	04 min. 53 sec.	05 min. 14 sec.	04 min. 22 sec.	04 min. 14 sec.

Slow Extensive validation jobs runs best as Cosmos due to Berzelius "GPU Usage Efficiency Policy" limitations
berzelius_safe line use 1 out of 8 GPUs, 16 out of 128 cores, 2 TB out of 16 TB SSD capacity per node

Real dataset comparison (Daniel Larsson)

	berzelius_mig	tetalith_gpu	berzelius_safe	tetalith_gpu
GPU and VRAM	1/7 NVIDIA A100 10GB VRAM...	1 NVIDIA T4 16 GB VRAM	1 NVIDIA A100 40GB VRAM	1 NVIDIA T4 16 GB VRAM
SSD	NVMe 15/56~0.27 TB SSD	NVMe SSD 2 TB	NVMe SSD 15/8 ~2 TB	NO SSD
CPU	2 cores	32 cores	16 cores	32 cores
RAM	32 GB	96 GB	1 TB	96 GB
GPU/h cost of run	0.25	1.0	1.0	1.0
CryoSPARC version	4.4.1	4.5.3	4.4.1	4.5.3
2D classification (6Å, 100 classes)	108 min (20 iterations)	15 hours ¹ (20 iterations)	40 min (20 iterations)	5 hours (20 iterations)
3D classification (6Å, 10 classes)	6.1 hours (185 iterations)	3.5 hours ² (183 iterations)	5.7 hours (185 iterations)	3+ days (120 iterations)
Homogeneous refinement ⁵	Failed	9.2 hours ³ (9 iterations 1.75Å)	9.0 hours (10 iterations 1.75Å)	17+ hours (5 iterations)
Non-uniform refinement ⁵	Failed	17.8 hours ⁴ (9 iterations, 1.80Å)	5.5 hours (7 iterations, 1.75Å)	not attempted
¹ 13 hours to sync to SSD, 2 threads, node n1282				
² 43 min to sync to SSD, 2 threads				
³ 33 min to sync to SSD				
⁴ 34 min to sync to SSD				
⁵ with higher-order CTF and per-particle corrections				

PReSTO 2025-2028?

PReSTO towards SciLifeLab NBIS – decision now in Dec 2024

Erik Lindahl new NAISS director

Expansions towards PDC-Dardel

PRe**STO** → **NBS**

The Swedish Research community!



Swedish National Large-scale experimental research Infrastructures

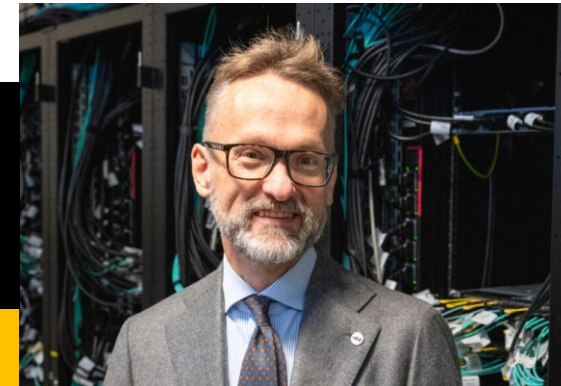
...approx 25% of them 😊 at Tällberg 2024

MAX IV

Cryo-EM National facility



Prof. Erik Lindahl new NAISS director
Relion and Gromacs developer
EuroHPC research advisory group
PRACE scientific steering committee



SwedNMR

National Compute resources

NSC

NAISS

National Supercomputer Centre

National Academic Infrastructure for Supercomputing in Sweden



PReSTO on Dardel...

Tim Schulte: “We may be able to establish a Relion-based processing platform at Dardel. Relion can be compiled on AMD GPUs. To facilitate particle picking, we could also include the cryolo picker which can be run in CPU-only mode at Dardel.”

+ Dardel

Stefan Fleischmann now split his time between SciLifeLab CryoEM and PDC hosting Dardel

Tim Schulte filed an application entitled ” Establishing Relion/cryolo installation on Dardel to support Swedish cryo-EM community”

- Dardel

Dardel GPU nodes are AMD and not NVIDIA.

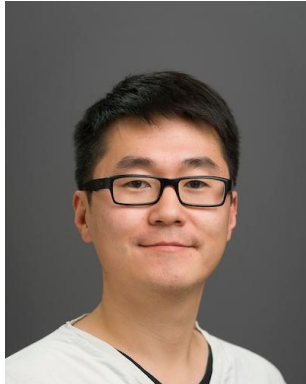
PReSTO enablers 2013-2024!



Cryo-EM National facility



Ana Gonzales



Daniel Larsson



André Graca



Tim Schulte



Karl Hörnell



Christian Luckey



Maria Sunnerhagen



Zdenek Matej



Jose delaRosa Trevin



Michael Hall



Piotr Draczkowski



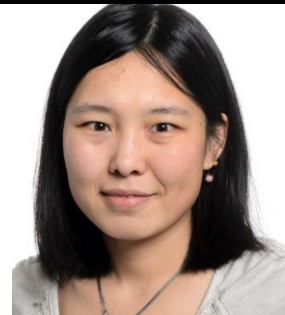
Sebastian Thorarensen



Torben Rasmussen



Marcos Acebes



Jie Nan



Stefan Fleischmann



Tapu Shaikh



More enablers at <https://www.nsc.liu.se/support/presto/enablers/>



THE END

Questions?

