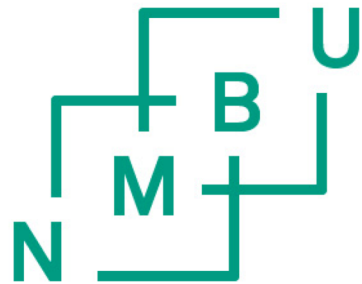


High performance computing in Norway

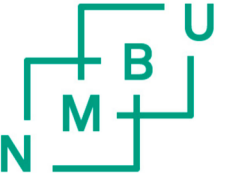
Åsmund Røhr Kjendseth

Faculty of Chemistry, Biotechnology, and Food Science



**Norwegian University
of Life Sciences**

Overarching theme



“Develop understanding of protein structure, function and catalysis”

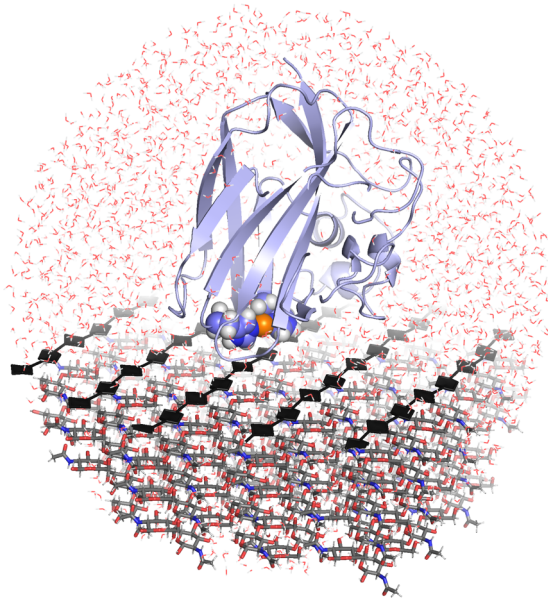
Enzymes – molecular machines



Ver. 1

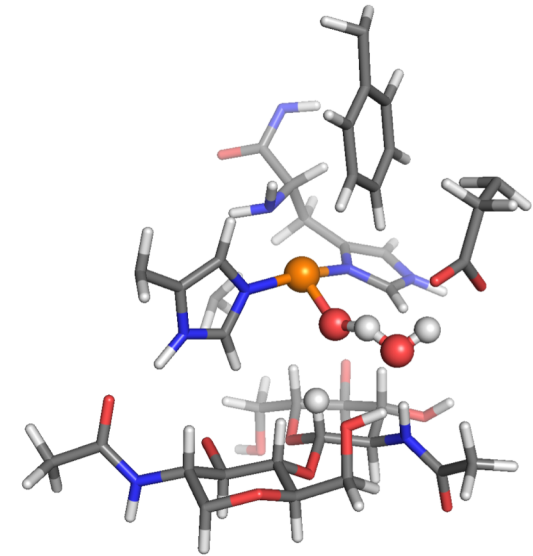


Ver. 2



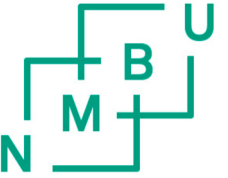
190 000 atoms!

Ver. 3



Increasing complexity when moving towards atomic scale

Supercomputers in biology



- ✓ Complicated research models may require **extensive computational resources**.
- ✓ Extensive computations may result in **huge amounts of raw data**.

National computational resources in Norway



UNINETT Sigma2 AS is a non-commercial company that manages the national infrastructure for computational science in Norway.

- ✓ High performance computing (HPC)
- ✓ Data storage

Collaboration between RCN, UiO, UiB, UiT and NTNU.
Linked to European computational infrastructure.

High performance computing



UiT:
Fram
(Stallo)

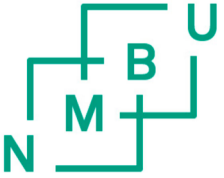
NTNU:
Vilje

UiB:
Hexagon

UiO:
Abel



Data storage

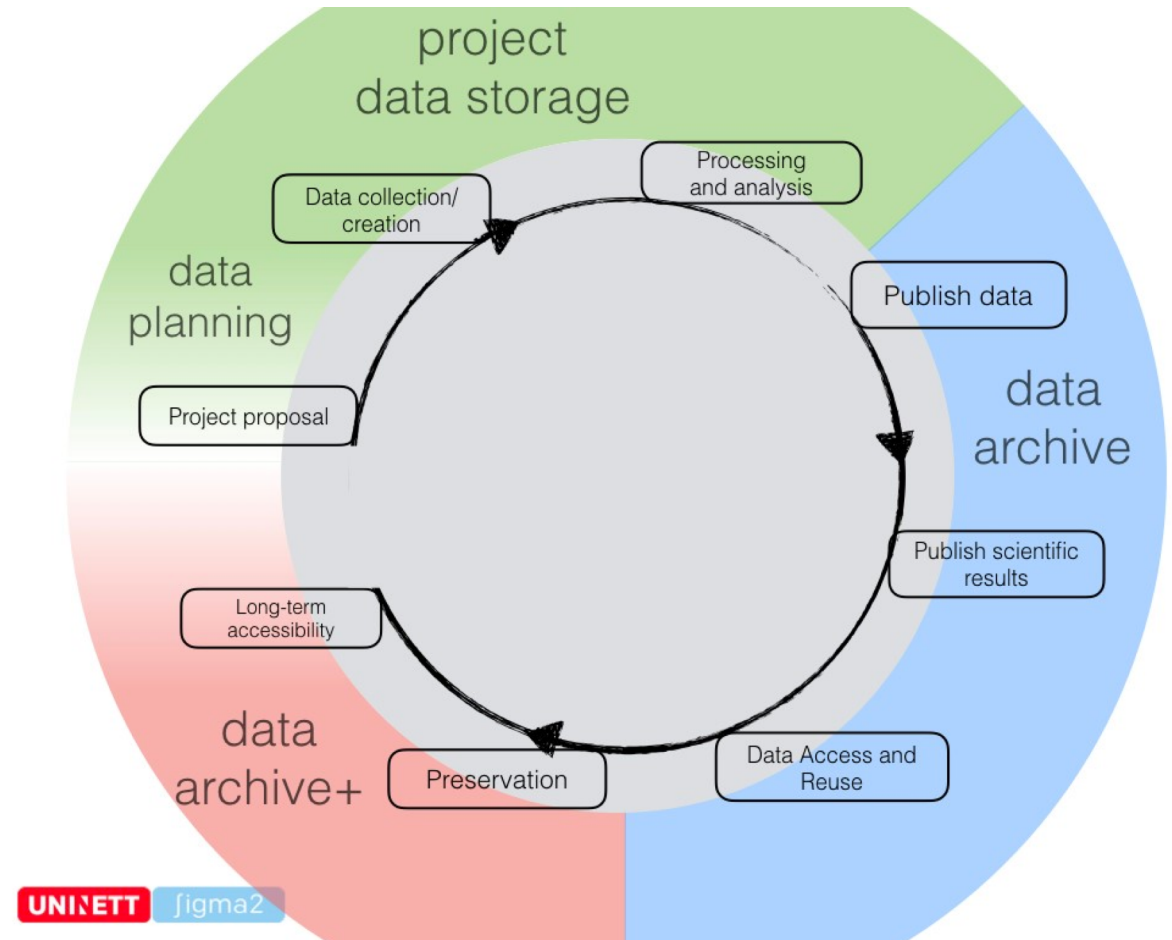


National e-Infrastructure for Research Data (NIRD)

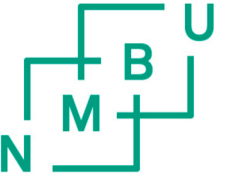
- ✓ Accommodates FAIR principles

Operating from 2017

~ 12 PB (12 000 terabyte)



The Abel computer cluster



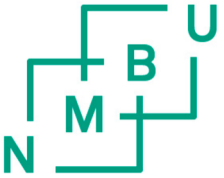
- ✓ More than 10 000 cores (about 2500 desktop computers)
- ✓ More than 650 nodes (16 cores pr. node)
- ✓ About 64 GB RAM pr. node

Software parallelization (efficiency) varies ...

Check out:

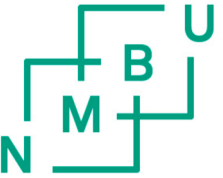
<http://www.uio.no/english/services/it/research/hpc/abel/help/software/>

Biology/bioinformatics software (Abel):



-
- | | | | | |
|-----------------|---------------------------------------|------------------------|----------------|------------------------------------|
| ■ 454apps | ■ denoiser | ■ LAMARC | ■ Orthograph | |
| ■ alleleCount | ■ dosageconverter | ■ MAFFT | ■ PAML | |
| ■ ABySS | ■ FastTree | ■ MACS2 | ■ Pandaseq | ■ STAR |
| ■ AmpliconNoise | ■ flash | ■ MAGMA | ■ ParsInsert | ■ Stringtie |
| ■ BEAGLE | ■ fqgrep | ■ MaSuRCA | ■ PAUP | ■ Subread |
| ■ BEAST | ■ Freesurfer | ■ MaxQuant | ■ PennSNV | ■ swarm |
| ■ BLAST | ■ FSL | ■ mcmcphase | ■ PhyloBayes | ■ TransDecoder |
| ■ BLAST+ | ■ GARLI | ■ metaxa2 | ■ PhyML | ■ TREEFINDER |
| ■ bowtie2 | ■ GATK | ■ MGLTools | ■ Picard-tools | ■ Trinotate |
| ■ busco | ■ Geneid | ■ Microbiome Utilities | ■ PLINK | ■ UCLUST |
| ■ cdbfasta | ■ hisat2 | ■ Migrate | ■ PLINKSEQ | ■ UNPHASED |
| ■ CD-HIT | ■ HMMER | ■ MIRA | ■ pplacer | ■ USEARCH |
| ■ CEGMA | ■ HUMAnN2 | ■ Molden | ■ ProtTest | ■ Velvet |
| ■ Circleator | ■ IMPUTE2 | ■ mothur | ■ QIIME | ■ vsearch |
| ■ Clearcut | ■ Infernal | ■ MrBayes | ■ RAxML | ■ <u>Wise2</u> (formerly GeneWise) |
| ■ ClonalOrigin | ■ Interproscan | ■ MUSCLE | ■ rtax | ■ VCFLIB |
| ■ ClustalW | ■ Irreproducible Discovery Rate (IDR) | ■ NCL | ■ SHAPEIT | |
| | | ■ ngsplot | ■ SPAdes | |
| | | ■ Novoalign | ■ structure | |
| | | | ■ Stacks | |

Other



Chemistry

- ADF
- AMBER
- AutoDock
- AutoDock Vina
- CP2K
- Gaussian
- LAMMPS
- MaterialsStudio
- NAMD
- QuantumEspresso
- VASP

Computational linguistics

- DyNet
- VISL CG-3

Geo Sciences

- ESyS-Particle
- FLEXPART
- OpenIFS
- WRF & WRF CHEM

Statistics

- R
- Stata

When do you need access to HPC

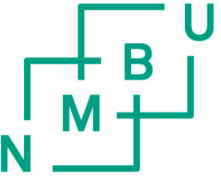


- ✓ When it feels like your laptop is melting
- ✓ When one program consumes ALL resources on your computer

CHECK THE ABEL SOFTWARE LIST!

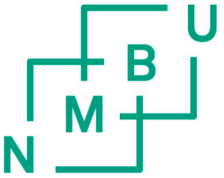


How to get access?



- ✓ Visit: <https://www.sigma2.no/content/apply-e-infrastructure-resources>
- ✓ Write a short but consistent proposal.
- ✓ Estimate how long the project period should be.
- ✓ Do some test-calculations so you can estimate the number of cpu-hours to apply for (e.g. 12 hours on 10 cores = 120 cpu hours)
- ✓ Small projects (~50 000 cpu hours) may be granted on short notice for testing purposes.

Training/courses



- ✓ Training courses are given regularly at UiO. Check out the Sigma2 site!
- ✓ HPC staff are in general supportive!

Norwegian Bioinformatics Platform

Search Site

Home About Help Desk Applications Training Storage News Publications Login

Hands-on Bioinformatics Workshop Organized in Trondheim by ELIXIR on January 25th

We are pleased to announce the next hands-on bioinformatics workshop organized in Trondheim by ELIXIR.

The topic is on how you can analyse microRNA sequencing data through the Norwegian e-Infrastructure for Life Sciences (NeLS) (nsls.bioinfo.no) and a ready-to-use microRNA analysis pipeline in Galaxy (galaxy-ntnu.bioinfo.no). Together, we will go through a complete microRNA sequencing analysis scenario.

Time: January 25th, 10:00-14:00

Place: St.Olavs Hospital, Neuro Centre, 3rd floor, room NS31 (Datasal)
MazeMap: <http://bit.ly/2Dids11>

Format: Hands-on. All participants need to bring their laptop.

Lecturers: Morten Beck Rye, Jostein Johansen, Kjetil Klepper

Register via google forms: [click here](#)

The workshop is limited to 20 people.

On behalf of ELIXIR - NTNU,
Jostein Johansen

Video Tutorials Available

[Click Here](#) to watch online video tutorials on NeLS-portal, NeLS-Galaxy and StoreBioInfo.

The bioinformatics platform provides research-based services and training to functional genomics users in Norway.

Join our Mailing List

Register for the bioinfo mailing list to receive updates

Analysis Workshop on Friday, 30-June-2017

Training events in 2018

Python workshop, Tromsø, 4-6 September, 2018

PDC Summer School, KTH, Sweden, 13-24 August, 2018

CSC Summer School in HPC 2018, Espoo, Finland 26 June - 4 July 2018

Cross-national training workshop on HPC and applications, Iceland 13 - 15 June 2018

CodeRefinery Workshop, Oslo, 12 - 14 June 2018

Software Carpentry Workshop, UiO 7 - 8 June 2018

NORBIS course: HPC in Bioinformatics, UiO, 16 - 27 April 2018

HPC for research training course, Oslo, 21 - 23. March, 2018

CSC Spring School on Computational Chemistry, Finland, 13 - 16 March 2018

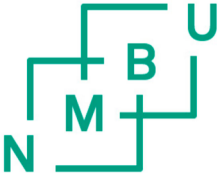
CodeRefinery Workshop, Trondheim, 27 February - 1 March 2018

Introduction to parallel programming, Trondheim, 06 - 15. February 2018

Research Bazaar 2018, Oslo, 7 - 9 February 2018

HPC course 2018.1, Bergen 25 - 26. January 2018

HPC at NMBU

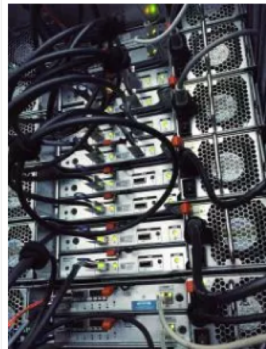


CIGENE computational unit

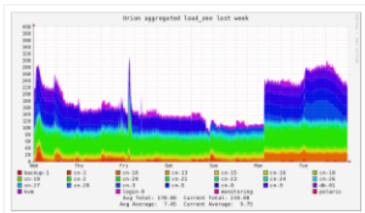


Bioinformatics is an essential and integrated part of CIGENE operations. The rapidly reduced cost of sequencing has increased demand of computational storage and analysis, and CIGENE has during the last years invested in a local computer cluster to meet this demand.

We currently maintain and administer a



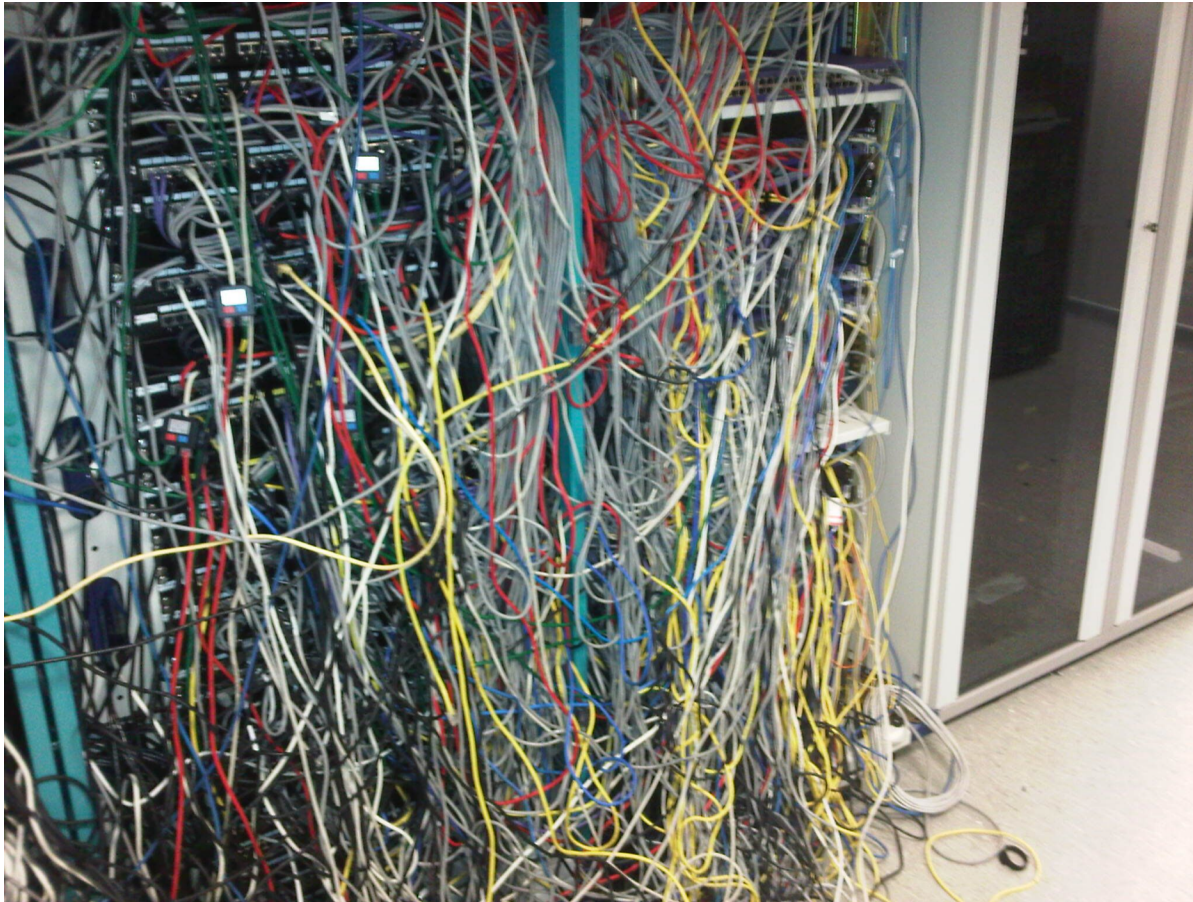
SLURM-based linux cluster called *Orion* with 580 CPUs, 4 TB RAM and 430 TB storage space. Access to Orion is provided to local users at the Norwegian University of Life Sciences (NMBU). CIGENE also hosts a *Galaxy* server to facilitate analysis of large-scale data in a friendly, reproducible environment; this is a collaboration with [NMBU's node in ELIXIR Norway](#).



Who use HPC at NMBU?

- ✓ Share experiences?
- ✓ Expand HPC toolboxes?
- ✓ Arrange workshops?
- ✓ Preach HPC to potential users and expand the community?

My experience with Management of Research Data



during the project

At project termination:

