

### **HPC** Introduction

High-performance computing

2024-11-25

Martin Dahlö martin.dahlo@nbis.se

Enabler for Life Science











# **Objectives**

### What is PDC, what it provides

Projects at PDC

How to access PDC

Jobs and queuing systems

How to use the resources of PDC



Parallelldatorcentrum http://www.pdc.kth.se



Parallelldatorcentrum http://www.pdc.kth.se

- 1 278 nodes
  - 128 cores each, 163 584 in total
  - 128 GB RAM, also 256, 512, 1024, 2048



Parallelldatorcentrum http://www.pdc.kth.se

- 1 278 nodes
  - 128 cores each, 163 584 in total
  - 128 GB RAM, also 256, 512, 1024, 2048
- 18 PB fast parallel storage



# Parallelldatorcentrum http://www.pdc.kth.se

- 1 278 nodes
  - 128 cores each, 163 584 in total
  - 128 GB RAM, also 256, 512, 1024, 2048
- 18 PB fast parallel storage
- Bioinformatics software





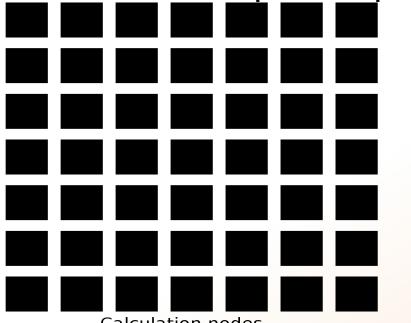
The basic structure of supercomputer







The basic structure of supercomputer



Calculation nodes

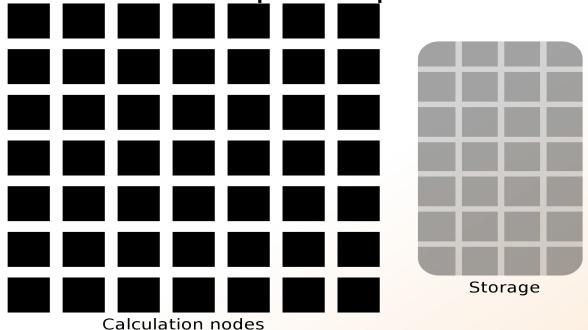


node = computer





The basic structure of supercomputer





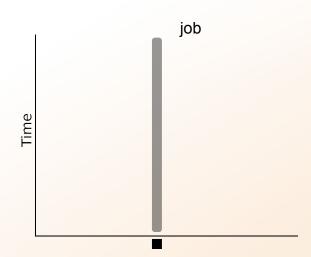
node = computer





The basic structure of a supercomputer

Parallel computing
Not one super fast







The basic structure of a supercomputer

Parallel computing
Not one super fast





# **Objectives**

What is PDC, what it provides

**Projects at PDC** 

How to access PDC

Jobs and queuing systems

How to use the resources of PDC



PDC provides its resources via

# projects



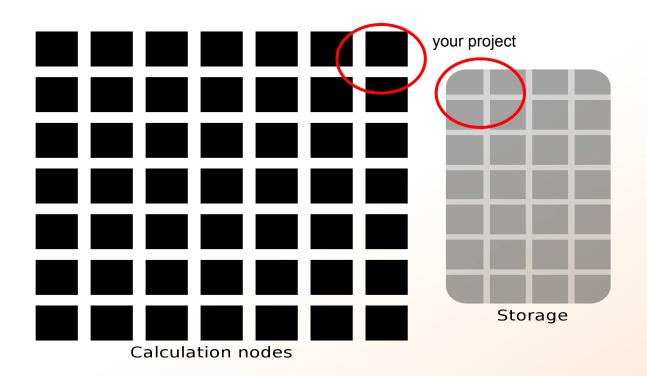
PDC provides its resources via

# projects

compute (core-hours/month)

storage (GB)







### Two separate projects:

### NAISS compute:

cluster **Dardel** 

2000 - 100 000+ core-hours/month

**512** GB storage

### **NAISS Storage**:

storage system **Klemming** 

1 - 100+ TB storage



Start Rounds Resources Support

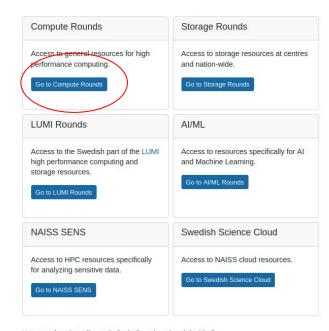
Login

Your are not logged in.

Start / Rounds

#### Rounds

Resources are made available through rounds, in which projects proposals are made. First, you need to select the type of round to use:



You can also view all rounds (including closed and decided).





Start Rounds

Resources Support Login

Your are not logged in.

Start / Rounds / Compute Rounds

#### **Compute Rounds**

HPC resources are made available through compute rounds. There are three different sizes of NAISS allocated rounds, as well as local rounds at centres. Select the size of round to use:

#### **NAISS Small Compute** NAISS Medium Compute Monthly evaluation of proposals during Up to 10 000 core-hours/month. Continuous evaluation of proposals the year. To apply, you must be a during the year. To apply, you must be scientist in Swedish academia, at least a scientist in Swedish academia, at at the level of assistant professor. least at the level of PhD student. Go to NAISS Medium Compute Go to NAISS Small Compute **NAISS Large Compute** Centre Local Compute Evaluation of proposals twice a year Local resources that are not allocated with a peer-review procedure. To apply, via NAISS. Conditions are described you must be a scientist in Swedish per round. academia, at least at the level of assistant professor.





Start / Rounds / NAISS Small Compute 2024

#### **NAISS Small Compute 2024**

Start Rounds

Resources Support

Login

Your are not logged in.

#### Open for Proposals

To apply, you must be a scientist in Swedish academia, at least at the level of PhD student.

#### Deadlines and Decisions

Proposals are processed weekly. Note that staff will be on vacation during the summer and proposals submitted in July will be processed at a reduced pace.

This round is open for proposals until 2025-01-01 00:00.

Create New Proposal for NAISS Small Compute 2024

#### Resources

	Resource	Centre	Upper Limit	Available	Unit	Note
•	Alvis	C3SE	1 000	80 000	GPU-h/month	The Alvis resource is dedicated for AI/ML research.
•	Tetralith	NSC	10	1 500	x 1000 core-h/month	
١	Dardel	PDC	10	1 720	x 1000 core-h/month	
•	Dardel- GPU	PDC	200	6 160	GPU-h/month	
Þ	Rackham	UPPMAX	10	1 500	x 1000 core-h/month	Restrictive policy for NEW projects on Rackham.

Click > above to show more information about the resource.



# **Objectives**

What is PDC what it provides

**Projects at PDC** 

**How to access PDC** 

Jobs and queuing systems

How to use the resources of PDC



### How to access PDC

SSH to Dardel

ssh -Y your username@dardel.pdc.kth.se

Requires setting up either SSH keys or Kerberos
<a href="https://www.pdc.kth.se/support/documents/basics/quickstart.html#how-to-log-in">https://www.pdc.kth.se/support/documents/basics/quickstart.html#how-to-log-in</a>



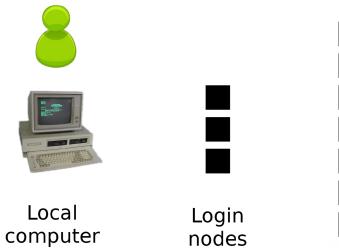
## How to access PDC

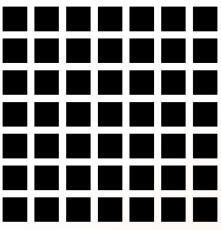
### SSH to Dardel

```
user@computer ~ $ ssh -Y username@dardel.pdc.kth.se
Last login: Mon Nov 11 10:19:30 2024 from icm-42-29.bmc.uu.se
    --== Welcome to Dardel! ==--
username@login1 ~ $
```



## SSH

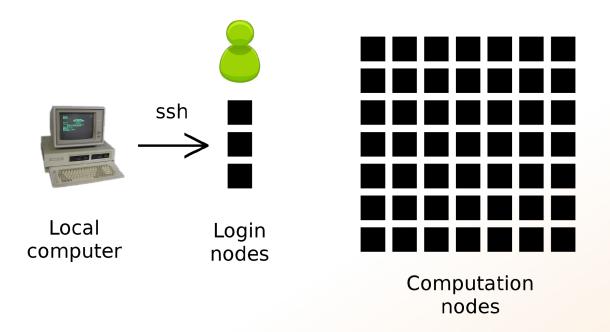




Computation nodes



## SSH





## How to use PDC

# Login nodes

use them to access PDC, never use them to run **jobs** 

# Calculation nodes

do your work here - testing and running, not accessible directly, SLURM (queueing system) gives you access



# **Objectives**

What is PDC, what it provides

**Projects at PDC** 

**How to access PDC** 

Jobs and queuing systems

How to use the resources of PDC





# Job (computing)

From Wikipedia, the free encyclopedia

For other uses, see <u>Job (Unix)</u> and <u>Job stream</u>.

In <u>computing</u>, a **job** is a unit of work or unit of execution (that performs said work). A component of a job (as a unit of work) is called a <u>task</u> or a <u>step</u> (if sequential, as in a <u>job stream</u>).

As a unit of execution, a job may be concretely identified with a single <u>process</u>, which may in turn have subprocesses (<u>child processes</u>; the process corresponding to the job being the <u>parent process</u>) which perform the tasks or steps that comprise the work of the job; or with a <u>process group</u>; or with an abstract reference to a process group, as in <u>Unix job control</u>.

SciLifeLab

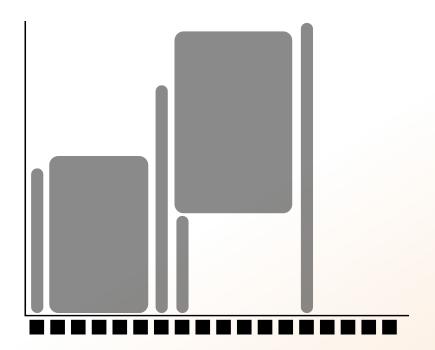
Read/open files

Do something with the data

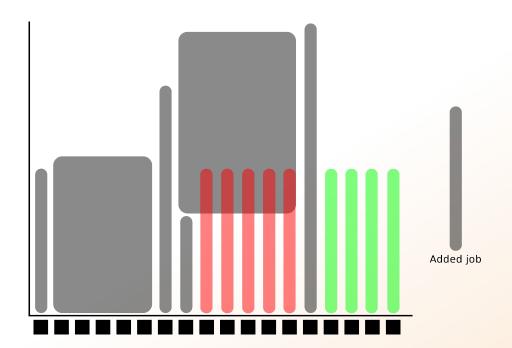
Print/save output



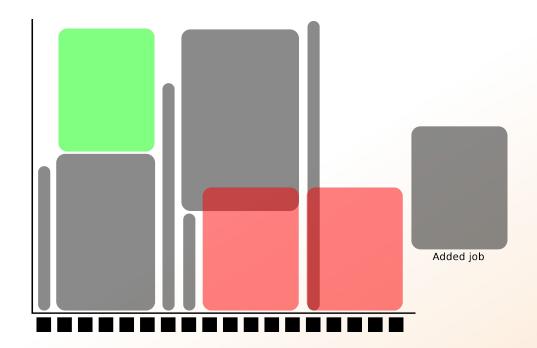
















workload manager job queue batch queue job scheduler

**SLURM** (Simple Linux Utility for Resource Management) free and open source



# **Objectives**

What is PDC, what it provides

**Projects at PDC** 

**How to access PDC** 

Jobs and queuing systems

How to use the resources of PDC





1) Ask for resource and run jobs manually
For testing, possibly small jobs, specific
programs needing user input while running

2) Write a script and submit it to SLURM
Submits an automated job to the job queue,
runs when it's your turn





### 1) Ask for resource and run jobs manually

book a node/core



ssh to the node



run programs



### 1) Ask for resource and run jobs manually

```
salloc -A naiss2099-99-999 -p shared -c 1 -t 00:05:00
```

- salloc commandmandatory job parameters:
- -A project ID (who "pays")
- **-p** partition
- -c number of cores
- **-t** time



salloc -A naiss2099-99-999 -p shared -c 1 -t 00:05:00

**-A** example course project naiss2099-99-999 you have to be a member



salloc -A naiss2099-99-999 -p shared -c 1 -t 00:05:00

- **-A** example course project naiss2099-99-999 you have to be a member
- -p shared = parts of 1 node 1 node = 128 cores)
  main = whole nodes 1 hour walltime = 128 core-hours



salloc -A naiss2099-99-999 -p shared -c 1 -t 00:05:00

- **-A** example course project naiss2099-99-999 you have to be a member
- -p shared = parts of 1 node 1 node = 128 cores)
  main = whole nodes 1 hour walltime = 128 core-hours
- -c number of cores (default value = 1)



salloc -A naiss2099-99-999 -p shared -c 1 -t 00:05:00

- -A example course project naiss2099-99-999 you have to be a member
- -p shared = parts of 1 node 1 node = 128 cores)
  main = whole nodes 1 hour walltime = 128 core-hours
- -c number of cores (default value = 1)
- -t format hh:mm:ss
  or dd-hh:mm:ss
  default value= 7-00:00:00
  jobs killed when time limit reaches always overestimate ~ 50%





#### Information about your jobs

squeue -u <user>

```
username@login1 ~ $ squeue -u username
  JOBID PARTITION NAME USER ST TIME NODES NODELIST(REASON)
  5236781 shared interact username R 20:41 1 nid002582
username@login1 ~ $
```



SSH to a calculation node (from a login node)



username@login1 ~ \$ salloc -A naiss2099-99-999 -t 01:00:00 -p shared -c 1



username@login1 ~ \$ salloc -A naiss2099-99-999 -t 01:00:00 -p shared -c 1

salloc: Pending job allocation 5236781

salloc: job 5236781 queued and waiting for resources

salloc: job 5236781 has been allocated resources

salloc: Granted job allocation 5236781

salloc: Nodes nid002582 are ready for job

username@login1 ~ \$



username@login1 ~ \$ salloc -A naiss2099-99-999 -t 01:00:00 -p shared -c 1

salloc: Pending job allocation 5236781

salloc: job 5236781 queued and waiting for resources

salloc: job 5236781 has been allocated resources

salloc: Granted job allocation 5236781

salloc: Nodes nid002582 are ready for job

username@login1 ~ \$ squeue -u username



username@login1 ~ \$ salloc -A naiss2099-99-999 -t 01:00:00 -p shared -c 1

salloc: Pending job allocation 5236781

salloc: job 5236781 queued and waiting for resources

salloc: job 5236781 has been allocated resources

salloc: Granted job allocation 5236781

salloc: Nodes nid002582 are ready for job

#### username@login1 ~ \$ squeue -u username

JOBID PARTITION NAME USER ST TIME NODES NODELIST (REASON)

5236781 shared interact username R 0:10 1 nid002582

username@login1 ~ \$



username@login1 ~ \$ salloc -A naiss2099-99-999 -t 01:00:00 -p shared -c 1

salloc: Pending job allocation 5236781

salloc: job 5236781 queued and waiting for resources

salloc: job 5236781 has been allocated resources

salloc: Granted job allocation 5236781

salloc: Nodes nid002582 are ready for job

username@login1 ~ \$ squeue -u username

JOBID PARTITION NAME USER ST TIME NODES NODELIST (REASON)

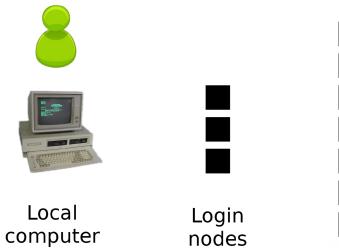
5236781 shared interact username R 0:10 1 nid002582

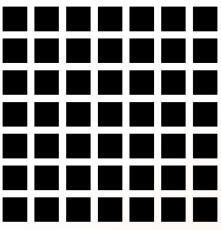
username@login1 ~ \$ ssh -Y nid002582





# SSH

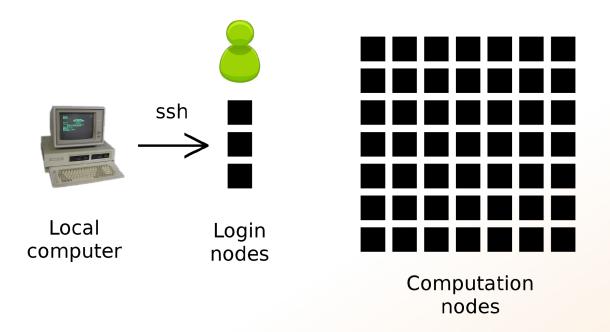




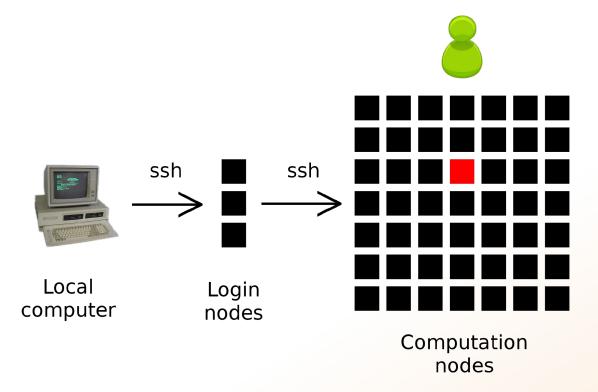
Computation nodes



# SSH











### 2) Write a script and submit it to SLURM

put all commands in a text file - script



tell SLURM to run the script (use the same job parameters)





### 2) Write a script and submit it to SLURM

put all commands in a text file - script

```
#! /bin/bash -l
#SBATCH -A XXXXXXX
#SBATCH -p shared
#SBATCH -J Template_script
#SBATCH -t 01:00:00

# load some modules
module load bioinfo-tools
# go to some directory
cd ~/testarea

# do something
echo Hello world!
```



#! /bin/bash -l



### 2) Write a script and submit it to SLURM

#### put all commands in a text file - script

```
#SBATCH -A XXXXXXX

#SBATCH -p shared

#SBATCH -J Template_script

#SBATCH -t 01:00:00

# load some modules

module load bioinfo-tools

# go to some directory

cd ~/testarea

# do something

echo Hello world!
```

job parameters

tasks to be done





### 2) Write a script and submit it to SLURM

#### put all commands in a text file - script

```
#! /bin/bash -1
#SBATCH -A naiss2099-99-999
#SBATCH -p shared
#SBATCH -J sample_001.fq_alignment
#SBATCH -t 7-00:00:00

# load some modules
module load bioinfo-tools bwa

# go to project directory
cd /cfs/klemming/projects/supr/naiss2099-99-999/
# align reads
bwa mem ref/human_reference.fa raw/sample_001.fq | samtools sort -o results/sample_001.aligned.sorted.bam
```



### 2) Write a script and submit it to SLURM

tell SLURM to run the script (use the same job parameters)

sbatch test.sh



### 2) Write a script and submit it to SLURM

tell SLURM to run the script (use the same job parameters)

sbatch test.sh

sbatch - command
test.sh - name of the script file



#### 2) Write a script and submit it to SLURM

tell SLURM to run the script (use the same job parameters)

**sbatch** -A naiss2099-99-999 -p main -c 8 -t 60:00:00 test.sh



```
username@login1 ~/test $ 11
-rw---- 1 username username 209 Nov 11 13:43 test.sh
username@login1 ~/test $
```



```
username@login1 ~/test $ 11
-rw---- 1 username username 209 Nov 11 13:43 test.sh
username@login1 ~/test $ cat test.sh
```



```
username@login1 ~/test $ 11
-rw----- 1 username username 209 Nov 11 13:43 test.sh
username@login1 ~/test $ cat test.sh
#! /bin/bash -1
#SBATCH -A naiss2099-99-999
#SBATCH -p shared
#SBATCH -J echo_test
#SBATCH -t 00:01:00

# go to home directory
cd ~/test

# print stuff
echo "Hello, this will be printed to the slurm-<jobID>.out"
username@login1 ~/test $
```



```
username@login1 ~/test $ 11
-rw-----    1 username username    209 Nov 11 13:43 test.sh
username@login1 ~/test $ cat test.sh
#! /bin/bash -1
#SBATCH -A naiss2099-99-999
#SBATCH -p shared
#SBATCH -J echo_test
#SBATCH -t 00:01:00

# go to home directory
cd ~/test

# print stuff
echo "Hello, this will be printed to the slurm-<jobID>.out"
username@login1 ~/test $ sbatch test.sh
```



```
username@login1 ~/test $ 11
-rw----- 1 username username 209 Nov 11 13:43 test.sh
username@login1 ~/test $ cat test.sh
#! /bin/bash -1
#SBATCH -A naiss2099-99-999
#SBATCH -p shared
#SBATCH -J echo_test
#SBATCH -t 00:01:00

# go to home directory
cd ~/test

# print stuff
echo "Hello, this will be printed to the slurm-<jobID>.out"
username@login1 ~/test $ sbatch test.sh
Submitted batch job 5831681
username@login1 ~/test $
```



```
username@login1 ~/test $ 11
-rw----- 1 username username 209 Nov 11 13:43 test.sh
username@login1 ~/test $ cat test.sh
#! /bin/bash -1
#SBATCH -A naiss2099-99-999
#SBATCH -p shared
#SBATCH -J echo_test
#SBATCH -t 00:01:00

# go to home directory
cd ~/test

# print stuff
echo "Hello, this will be printed to the slurm-<jobID>.out"
username@login1 ~/test $ sbatch test.sh
Submitted batch job 5831681
username@login1 ~/test $ 11
```



```
username@login1 ~/test $ 11
-rw----- 1 username username 209 Nov 11 13:43 test.sh
username@login1 ~/test $ cat test.sh
#! /bin/bash -l
#SBATCH -A naiss2099-99-999
#SBATCH -p shared
#SBATCH -J echo test
#SBATCH -t 00:01:00
# go to home directory
cd ~/test
# print stuff
echo "Hello, this will be printed to the slurm-<jobID>.out"
username@login1 ~/test $ sbatch test.sh
Submitted batch job 5831681
username@login1 ~/test $ 11
-rw----- 1 username username 252 Nov 11 13:43 slurm-5831681.out
-rw----- 1 username username 209 Nov 11 13:43 test.sh
username@login1 ~/test $
```



```
username@login1 ~/test $ 11
-rw----- 1 username username 209 Nov 11 13:43 test.sh
username@login1 ~/test $ cat test.sh
#! /bin/bash -l
#SBATCH -A naiss2099-99-999
#SBATCH -p shared
#SBATCH -J echo test
#SBATCH -t. 00:01:00
# go to home directory
cd ~/test
# print stuff
echo "Hello, this will be printed to the slurm-<jobID>.out"
username@login1 ~/test $ sbatch test.sh
Submitted batch job 5831681
username@login1 ~/test $ 11
-rw----- 1 username username 252 Nov 11 13:43 slurm-5831681.out
-rw----- 1 username username 209 Nov 11 13:43 test.sh
username@login1 ~/test $ cat slurm-5831681.out
```



```
username@login1 ~/test $ 11
-rw----- 1 username username 209 Nov 11 13:43 test.sh
username@login1 ~/test $ cat test.sh
#! /bin/bash -l
#SBATCH -A naiss2099-99-999
#SBATCH -p shared
#SBATCH -J echo test
#SBATCH -t. 00:01:00
# go to home directory
cd ~/test
# print stuff
echo "Hello, this will be printed to the slurm-<jobID>.out"
username@login1 ~/test $ sbatch test.sh
Submitted batch job 5831681
username@login1 ~/test $ 11
-rw----- 1 username username 252 Nov 11 13:43 slurm-5831681.out
-rw----- 1 username username 209 Nov 11 13:43 test.sh
username@login1 ~/test $ cat slurm-5831681.out
Hello, this will be printed to the slurm-<jobID>.out
username@login1 ~/test $
```



# **Software**

100+ programs installed

Managed by a 'module system' Installed, but hidden Manually loaded before use



# **Software**

#### 100+ programs installed

Managed by a 'module system' Installed, but hidden Manually loaded before use

```
module avail
module load <module name>
module unload <module name>
module list
module spider <word>
```

- Lists all available modules
  - Loads the module
  - Unloads the module
  - Lists loaded modules
  - Searches all modules after 'word'



# **Software**

Most bioinfo programs hidden under bioinfo-tools Load bioinfo-tools first, then program module

username@login1 ~ \$ module load cufflinks



# Most bioinfo programs hidden under bioinfo-tools Load bioinfo-tools first, then program module

#### username@login1 ~ \$ module load cufflinks

```
Lmod has detected the following error: These module(s) or
extension(s) exist but cannot be loaded as requested:
"cufflinks"
```



# Most bioinfo programs hidden under bioinfo-tools Load bioinfo-tools first, then program module

```
username@login1 ~ $ module load cufflinks
Lmod has detected the following error: These module(s) or
extension(s) exist but cannot be loaded as requested:
"cufflinks"
username@login1 ~ $ module load bioinfo-tools
```



# Most bioinfo programs hidden under bioinfo-tools Load bioinfo-tools first, then program module

```
username@login1 ~ $ module load cufflinks
Lmod has detected the following error: These module(s) or
extension(s) exist but cannot be loaded as requested:
"cufflinks"
username@login1 ~ $ module load bioinfo-tools
username@login1 ~ $ module load cufflinks
```



# Most bioinfo programs hidden under bioinfo-tools Load bioinfo-tools first, then program module

```
username@login1 ~ $ module load cufflinks
Lmod has detected the following error: These module(s) or
extension(s) exist but cannot be loaded as requested:
"cufflinks"
username@login1 ~ $ module load bioinfo-tools
username@login1 ~ $ module load cufflinks
Or
```

username@login1 ~ \$ module load bioinfo-tools cufflinks

#### username@login1 ~ \$ module avail

	/pdc/software/eb/modules/all				
angsd/0.940		googletest/1.8.1		ncurses/6.4	(L,D)
ant/1.10.14		gperf/3.0.4		nextflow/24.04.2	
apr/1.7.0		gperf/3.1	(D)	ninja/1.10.2	
argtable/2.13		groff/1.23.0		ninja/1.11.0	
<pre>autoconf-archive/2022.02.11</pre>		gs1/2.3		ninja/1.11.1	
autoconf/2.69		gs1/2.7.1	(D)	ninja/1.12.1	(D)
autoconf/2.71	(D)	guile/2.0.14		ntcard/1.2.1	
<pre>autodiff/0.5.10</pre>		gzip/1.10		onetbb/2021.5.0	
autodiff/0.6.5	(D)	gzip/1.12		opari2/2.0.6	
automake/1.16.1		gzip/1.13	(D)	openexr/3.2.0	
automake/1.16.4		haploeagle/2.4.1		openjpeg/2.5.0	
automake/1.16.5	(D)	harfbuzz/8.2.2		osmesa/21.3.7	
autotools/20220317		hdf/4.2.16-2		ospray/2.4.0	
bamm/2.5.0		help2man/1.49.2		otf2/2.3	
bamsifter/2.15.0		hifiasm/0.19.7		otf2/3.0.3	(D)
bamtools/2.5.2		highway/1.2.0		paml/4.9j	
barrnap/0.9		hisat2/2.2.1		pango/1.50.14	
bazel/6.3.1		htslib/1.15.1		parafly/0.1.0	
bbmap/38.61b		htslib/1.20	(L,D)	paralle1/20230422	
bbmap/39.01		hwloc/2.6.0		parasail/2.6.2	
bbmap/39.06	(D)	hwloc/2.9.0		patchelf/0.14.5	
bcftools/1.15.1		hwloc/2.11.1	(D)	pbbam/1.0.7	
bcftools/1.20	(D)	icu/69.1		pbcopper/1.8.0	
bc12fastq2/2.20.0		icu/74.1	(D)	pblat/2.5.1	
bedtools/2.31.0		imagemagick/7.1.0-32		pcre/8.45	
bifrost/1.0.6.4		imath/3.1.9		pcre2/10.40	



#### **PDC Commands**

#### projinfo

#### username@login1 ~ \$ projinfo

\$HOME folder

Path: /cfs/klemming/home/u/username

Storage: 12.34 GiB
Number of files: 83135

Project info for all projects for user: username

Information for compute project: uppmax.staff (PI: username)

Test allocation for UPPMAX

Active from 2023-10-01 00:00:00 to 2027-01-01 00:00:00

Members: username, username01, username02

dardel: 10000 corehours/month, used 12.52% (1252 corehours) during the past 30 days

Information for compute project: naiss2024-5-11 (PI: username)

SNIC systems access for application experts

Active from 2024-01-09 00:00:00 to 2025-02-01 00:00:00

Members: username, username01

dardel: 2000 corehours/month, used 0.00% (0 corehours) during the past 30 days



• The difference between user account and project



- The difference between user account and project
- Login nodes are not for running jobs



- The difference between user account and project
- Login nodes are not for running jobs
- SLURM gives you access to the compute nodes when you specify a project that you are member of



- The difference between user account and project
- Login nodes are not for running jobs
- SLURM gives you access to the compute nodes when you specify a project that you are member of
- Do not ask for more cores/nodes than your job can actually use



- The difference between user account and project
- Login nodes are not for running jobs
- SLURM gives you access to the compute nodes when you specify a project that you are member of
- Do not ask for more cores/nodes than your job can actually use
- A job script usually consists of:

Job settings (-A, -p, -c, -t)

Modules to be loaded

Bash code to perform actions

Run a program, or multiple programs



#### Laboratory time! (again)

https://nbisweden.github.io/workshop-ngsintro/2411/topics/hpc/intro/lab\_hpc\_intro.html