

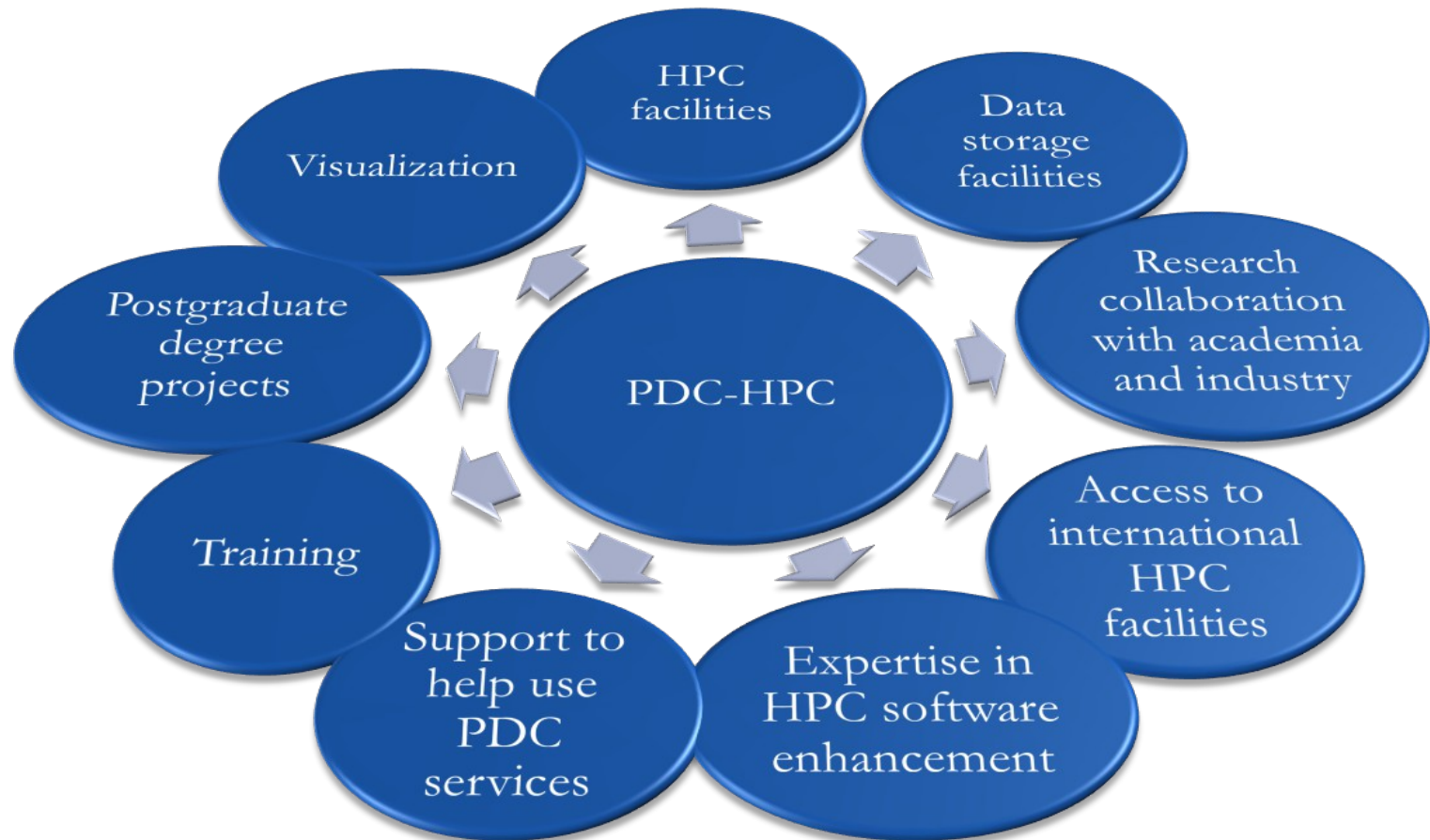
Introduction to the PDC environment



PDC Center for High Performance Computing
KTH, Sweden

Rossen Apostolov
(rossen@kth.se)
Application expert at PDC

PDC Offers...



PDC Key Assets: First-Line Support and System Staff

First-line support

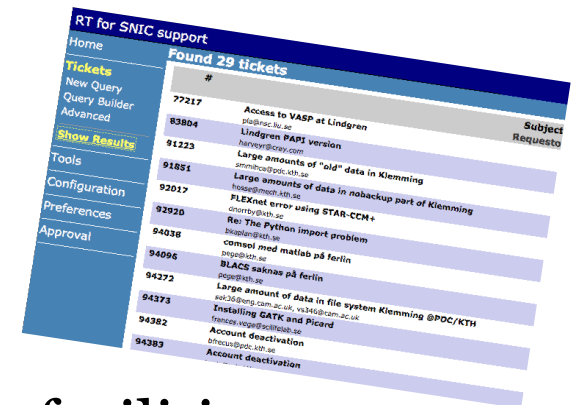


ROYAL INSTITUTE
OF TECHNOLOGY

Helps you have a smooth start to using PDC's resources and provides assistance if you need help while using our facilities

System staff: System managers/administrators

Ensure that PDC's HPC and storage facilities run smoothly and securely



PDC's Key Assets: HPC Application Experts

PDC-HPC application experts hold PhD degrees in different scientific fields and are experts in HPC.

Together with researchers, they optimize, scale and enhance scientific codes for the next generation supercomputers.



ROYAL INSTITUTE
OF TECHNOLOGY



Jonathan Vincent
Computational Physics



Rossen Apostolov
Computational Chemistry



Michael Djurfeldt
Computational Neuroscience



Henric Zazzi
Bioinformatics/Genetics



Jaime Axel Rosal Sandberg
Computational Chemistry



Jing Gong
Scientific Computing

Access to PDC resources



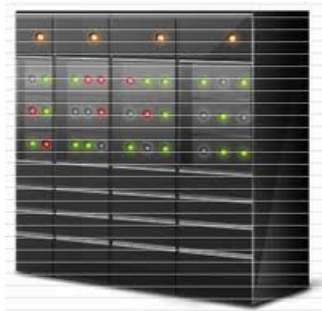
- User account
- Time allocation (CAC)
 - Which clusters you can access
 - How many jobs you can run per month
- Time allocation requirements
 - Can be personal or shared within a group
 - Every user must belong to at least one time allocation



ROYAL INSTITUTE
OF TECHNOLOGY

Infrastructure at PDC

What is a cluster



Cluster

Type of nodes

Login nodes

Compute nodes

Dedicated nodes

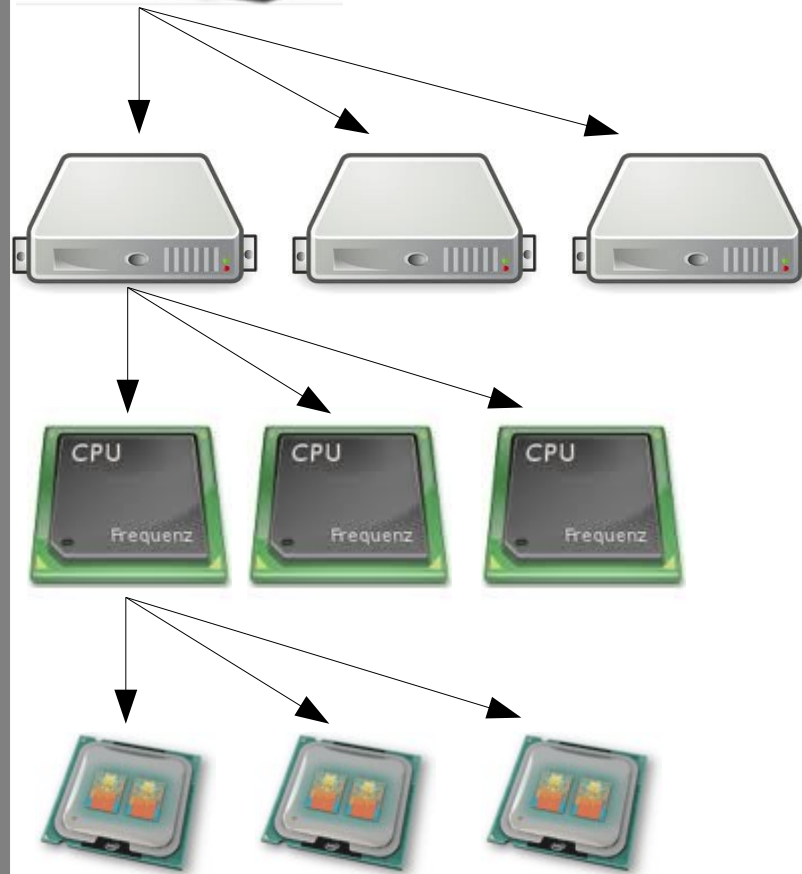
Interactive nodes

Shared nodes

Exclusive nodes

Transfer nodes

Service nodes



Nodes

Similar to computer without keyboard and monitor

Processors

Cores

Beskow

Hardware

1676 nodes

32 cores/node (53,632 total)

2.3 GHz Haswell

64 GB RAM



ROYAL INSTITUTE
OF TECHNOLOGY

- Large jobs have higher priority (>512 core/job)
- Queue limit is 24 hours
- Runs the SLURM queue system
- Partially reserved for PRACE, SCANIA etc.
 - Large allocations: 75%
 - Medium allocations: 20%
- Lifetime: end of 2018

Tegner

Hardware

65 node GPU cluster
nVIDIA Tesla K80s and
Quadro K420s
512GB-2TB RAM



- Intended for pre-/post-processing and GPU simulations and visualization
- Runs the Moab/Torque queue system
- In production since July 2015



**ROYAL INSTITUTE
OF TECHNOLOGY**

How to login

Kerberos



ROYAL INSTITUTE
OF TECHNOLOGY

- Is an authentication protocol originally developed at MIT
- PDC uses kerberos together with **SSH** for login
- **Ticket**
 - Proof of users identity
 - Users use password to obtain tickets
 - Tickets are cached on users computer for a specified duration
 - Tickets should be created on your local computer
 - As long as tickets are valid there is no need to enter password
- **Realm**
 - all resources available to access
 - example: NADA.KTH.SE
- **Principal**
 - Unique identity to which kerberos can assign tickets.
 - example: *username*@NADA.KTH.SE

Kerberos commands

- **kinit** - proves your identity
- **klist** - list your kerberos tickets
- **kdestroy** - destroy your kerberos ticket file
- **kpasswd** - change your kerberos password



ROYAL INSTITUTE
OF TECHNOLOGY

```
> kinit -f username@NADA.KTH.SE  
> klist -Tf
```

```
Credentials cache : FILE:/tmp/krb5cc_500
```

```
Principal: username@NADA.KTH.SE
```

```
Issued          Expires          Flags    Principal
```

```
Mar 25 09:45 Mar 25 19:45 FI krbtgt/NADA.KTH.SE@NADA.KTH.SE
```

```
Mar 25 09:45 Mar 25 19:45 FA afs/pdc.kth.se@NADA.KTH.SE
```

Login using kerberos tickets



ROYAL INSTITUTE
OF TECHNOLOGY

- Get a 7 days forwardable ticket on your local system
`kinit -f -l 7d username@NADA.KTH.SE`
- Forward your ticket via ssh
`ssh username@clustername.pdc.kth.se`
- Replace *clustername*...
 - Beskow login node: `beskow-login2.pdc.kth.se`
 - Tegner login node: `tegnor.pdc.kth.se`
- You will have reached the cluster
- Always create a kerberos ticket on your local system

Login from any computer



ROYAL INSTITUTE
OF TECHNOLOGY

- You can reach PDC from any computer or network
- The kerberos implementation heimdal can be installed on most operating systems
 - Linux
 - Windows
 - Mac
- Follow the instructions for your operating system

www.pdc.kth.se/resources/software/login-1



ROYAL INSTITUTE
OF TECHNOLOGY

File systems

AFS

Andrew File System



- AFS is a global file system accessible everywhere
/afs/pdc.kth.se/home/username 1st letter/username
- Your home directory is located in AFS
- Oldfiles folder contain yesterdays backup of your files
- You cannot run jobs from AFS on Beskow
- Follow the instructions for your operating system
www.pdc.kth.se/resources/software/file-transfer/file-transfer-with-afs

Lustre



- Massively parallel distributed file system
- Very high performance
- No backup
- No personal quota. Move your data when finished
- Always start and run your programs in lustre
- Home directory:

/cfs/klemming/nobackup/username 1st letter/username

Types of nodes



- Login nodes
 - Do not run computer intensive jobs here
- Interactive nodes
 - Node will be reserved just for you
 - Should be logged into directly and not via the login node
- Dedicated nodes
 - Reserve using the queue system
 - Node will be reserved just for you

Modules

- Used to load a specific software into your environment



ROYAL INSTITUTE
OF TECHNOLOGY

- **module avail [*software_name*]**
Lists available softwares
- **module add *software***
loads *software*
- **module show *software***
shows information about *software*
- **module list**
Lists currently loaded softwares
- **module swap *frommodule tomodule***
Swaps *frommodule* to *tomodule*



**ROYAL INSTITUTE
OF TECHNOLOGY**

How to run jobs

Prior to starting a job



- Get a forwardable kerberos ticket from **local computer**

```
kinit -f -l 7d username@NADA.KTH.SE
```

- Forward your ticket via ssh

```
ssh username@clustername.pdc.kth.se
```

- Copy your code to your AFS directory

/afs/pdc.kth.se/home/u/username

SLURM queue system



- Queue system to run jobs
 - Installed on both Beskow and Tegner
- Installed by default, no need to load module
- **salloc** – to submit a job to a dedicated node or book an interactive node
- More information at...
www.pdc.kth.se/resources/computers/beskow/how-to/run

SLURM queue system



ROYAL INSTITUTE
OF TECHNOLOGY

- To book a dedicated node

```
$ salloc -t <1:30:00> -N <nodes> [-A summer-2015]  
--reservation=summer-2015-08-11 --nodes=1 [script/command]
```

- To remove a submitted job

```
$ scancel jobid
```

- Show my running jobs

```
$ squeue -u <username>
```

- Submit scripts

```
$ sbatch <script>
```

Run interactively (Beskow)



ROYAL INSTITUTE
OF TECHNOLOGY

- MPI jobs

```
$ aprun -n <cores> [-N <nodes>] ./hello_mpi_c
```

- You have to book a node before using *aprun -n ...*, otherwise it will fail to execute!



ROYAL INSTITUTE
OF TECHNOLOGY

Compiling and Running on Tegner

GNU Compilers

- (optional) module add gcc/4.9.2
- Compile serial jobs



ROYAL INSTITUTE
OF TECHNOLOGY

```
gfortran -o hello hello.f  
gcc -o hello_mpi hello_mpi.c
```

- Compile MPI jobs

```
module add gcc/4.9.2 openmpi/1.8-gcc-4.9  
mpif90 -fopenmp -o hello_mpi hello_mpi.f  
mpicc -fopenmp -o hello_mpi hello_mpi.c
```

Intel Compilers

- Compile serial jobs

```
module add i-compilers  
ifort -o hello hello.f  
icc -o hello_mpi hello_mpi.c
```



ROYAL INSTITUTE
OF TECHNOLOGY

- Compile MPI jobs

```
module add i-compilers intelmpi  
mpiifort -openmp -o hello_mpi hello_mpi.f  
mpiicc -openmp -o hello_mpi hello_mpi.c  
mpiicpcp -openmp -o hello_mpi hello_mpi.cpp
```

Running on Tegner



ROYAL INSTITUTE
OF TECHNOLOGY

- For interactive usage book with e.g.

```
$ srun -t <1:30:00> -N <nodes> [-A summer-2015]  
--reservation=summer-2015-08-11 --nodes=1 ./hello_mpi
```

- Submit a script to dedicated nodes

```
sbatch ./job.sh
```

- Where job.sh is ...

Example batch script for Tegner



ROYAL INSTITUTE
OF TECHNOLOGY

```
#!/bin/bash -l
#SBATCH -J myjob
# Specify the account and the reservation (check the date!)
#SBATCH -A summer-2015
#SBATCH --reservation=summer-2015-08-19
# 10 minute wall-clock time will be given to this job
#SBATCH -t 1:30:00
# Number of nodes
#SBATCH --nodes=1
# set tasks per node to 24 to disable hyperthreading
#SBATCH --ntasks-per-node=24

# load intel compiler and mpi
module load i-compilers intelmpi

# Run program
mpirun -n 48 ./hello_mpi
```

NB: No space between # and SBATCH! Otherwise it's treated as a comment.

Compiling and Running CUDA programs on Tegner



ROYAL INSTITUTE
OF TECHNOLOGY

- To compile cuda programs the `nvcc` command should be used. The `-arch` flag to optimise for the hardware we have should also be used. e.g.

```
module add cuda
```

```
nvcc -arch=sm_37 -O2 hello.cu -o hello.x
```

- To run the program, we need to select nodes that have GPU cards. We can pass the selection either as an argument to `srun` (interactive use):

```
srun --nodes=1 -t 00:10:00 --ntasks-per-  
node=1 -gres=--gres=gpu:K420:1 ./hello_gpu
```

- or as an SBATCH option:

```
#SBATCH --gres=gpu:K80:2
```



ROYAL INSTITUTE
OF TECHNOLOGY

Compiling and Running on Beskow

Compilers

- Select programming environment with:

```
module swap PrgEnv-cray PrgEnv-gnu (or PrgEnv-intel)
```

- (optional) `module load cray-libsci`

- Compile serial or MPI jobs only with

```
ftn -o hello hello.f
```

```
cc -o hello_mpi hello_mpi.c
```

```
CC -o hello_mpi hello_mpi.cpp
```

- The wrappers *ftn*, *cc* and *CC* automatically link to the correct libraries and add needed flags



Running on Beskow



ROYAL INSTITUTE
OF TECHNOLOGY

- For interactive usage book with e.g.

```
$ salloc -t <1:30:00> -N <nodes> [-A summer-2015]  
--reservation=summer-2015-08-11 --nodes=1
```

- Then execute with e.g.

```
aprun -n 24 ./hello_serial
```

- Alternatively, submit a script to dedicated nodes

```
sbatch ./job.sh
```

- Where job.sh is...

Example batch script for Beskow



ROYAL INSTITUTE
OF TECHNOLOGY

```
#!/bin/bash -l  
#SBATCH -J myjob  
# Specify the account and the reservation (check the date!)  
#SBATCH -A summer-2015  
#SBATCH --reservation=summer-2015-08-19  
# 10 minute wall-clock time will be given to this job  
#SBATCH -t 1:30:00  
# Number of nodes  
#SBATCH --nodes=1  
# set tasks per node to 24 to disable hyperthreading  
#SBATCH --ntasks-per-node=24  
  
# load intel compiler and mpi  
module load i-compilers intelmpi  
  
# Run program  
aprun -n 48 ./hello_mpi
```

More information about running jobs



ROYAL INSTITUTE
OF TECHNOLOGY

<https://www.pdc.kth.se/resources/computers/beskow/how-to>

<https://www.pdc.kth.se/resources/computers/tegner/how-to>

PDC support



ROYAL INSTITUTE
OF TECHNOLOGY

- A lot of question can be answered via our web
www.pdc.kth.se/support
- The best way to contact us is via e-mail
www.pdc.kth.se/about/contact/support-requests
 - The support request will be tracked
 - Write descriptive subject line
 - For follow ups always include support number
 - **[SNIC support #NNNNN]**
 - Do not make new support cases by replying to old tickets