

HPC Systems Anatomy & Storage

National Supercomputer Centre (NSC), Linköping University

SNIC-PRACE training

Online @NSC 20th Apr 2021, 10:00 - ca. 15:00

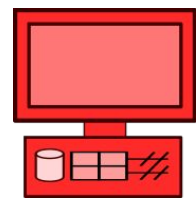
When & Why to use HPC?

HPC = High Performance Computing

- **High number** of simulation or data analysis jobs
- The jobs are **too large** for a desktop/laptop
- Used in most research fields today
 - Numerical weather prediction
 - Climate simulations
 - Flow simulations
 - Materials science
 - Many disciplines within Chemistry, Physics, Biology
 - ...

Desktop PC vs HPC

Tetralith: 1908 nodes, Sigma: 110 nodes



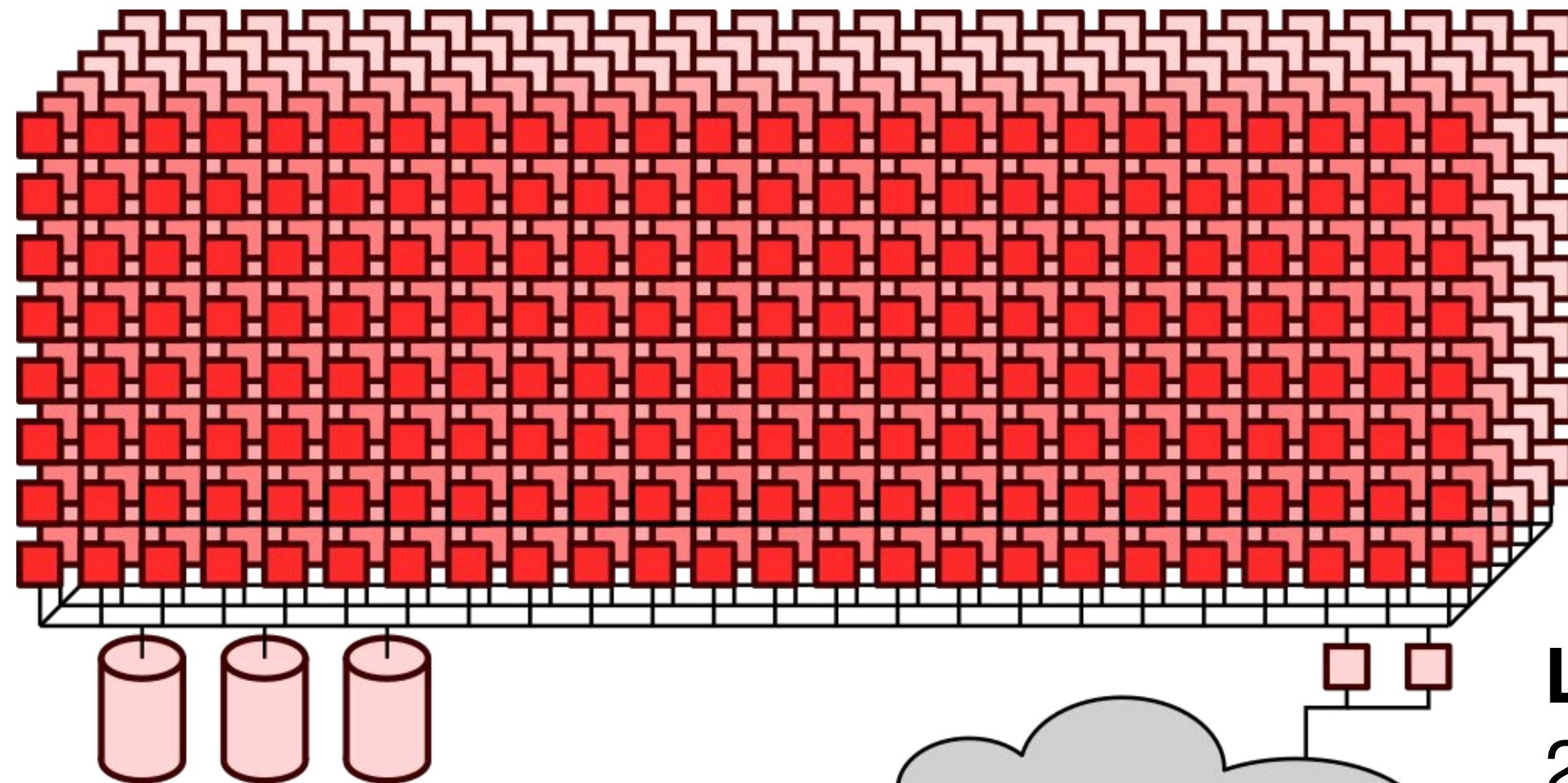
Desktop/laptop:

8 cores

16 GiB RAM

Windows, MacOS (Unix), Linux

1 user



Work node:

32 cores

96 (384) GiB RAM

Linux

Omni-Path network, interconnect

1 - few users at a time

Login nodes:

2 nodes

Linux

Many users

Internet

Desktop PC vs HPC

Single PC/laptop

Win, MacOS, -nix

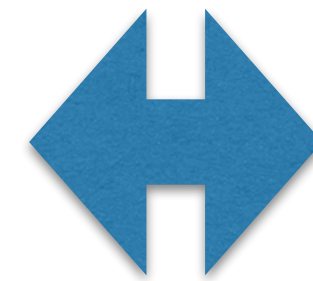
1 user - not shared

“ok” cores

“ok” RAM

Newest CPU?

1 gamer GPU?



Many nodes w/ fast interconnect

Linux

Many users on login nodes - shared

More cores

More RAM

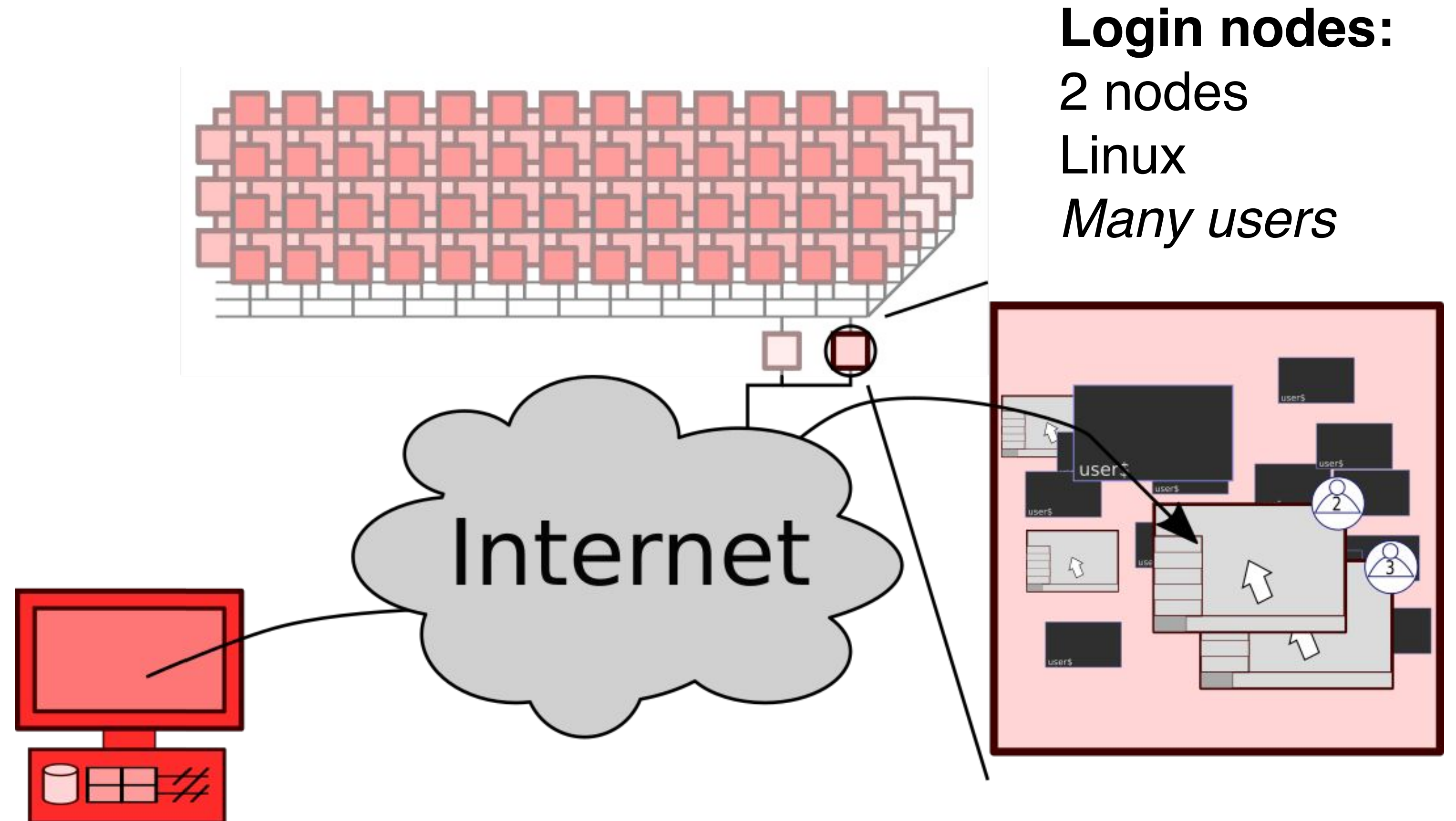
Cost efficient CPU

Many high-end GPUs?

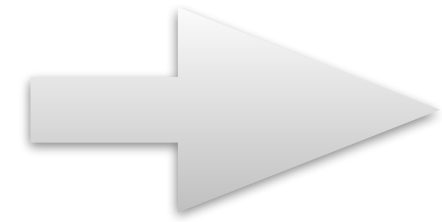
Note that HPC isn't always the best or fastest solution... ..it depends

@Login Node

- Typical access: using ssh
- For graphics, use ThinLinc
- *Many users* share login node
- Be mindful of login node usage
- Work node access via queue system (Slurm)



Some Basics



- **Linux**, see e.g. [guide](#) and [forum](#) + Simple **bash** scripting
 - ▶ Basic commands: `cd`, `pwd`, `ls`, `mkdir`, `mv`, `grep`, `less`, `cat`, `man`, ...
- **Common tools**
 - ▶ Text editors: `vi`, `gedit`, `emacs`, `nano`, ...
 - ▶ Plotting graphs: `gnuplot`, `grace`, ...
 - ▶ Analysis (basic/complex): `python`, `R`, `Matlab`, ...
- **Useful things**
 - ▶ Persistent terminal session: `screen`, `tmux`
 - ▶ Check compute usage [NSC]: `projinfo`
 - ▶ Check disk usage [NSC]: `snicquota`

Storage

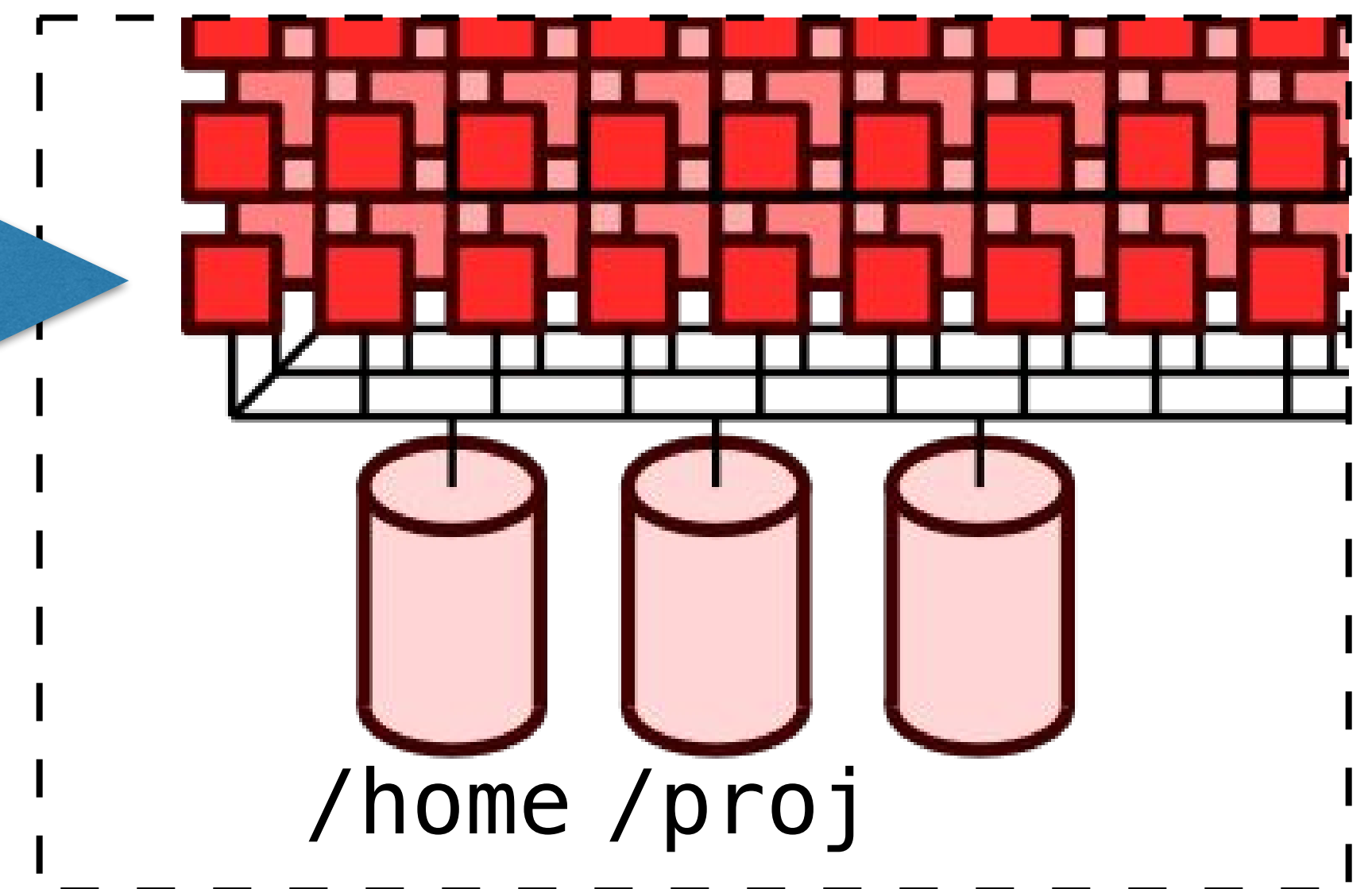
Three types of storage areas available:

Personal home folder, /home/username **small 20 GiB**

Project storage, owned by PI, /proj/ourstuff **large, \geq 500 GiB**

Work node local scratch disk (during run)

Heavy input/output /scratch/local
Large temp. files



Mounted everywhere

Backup & Snapshots

Three types of storage areas available:

Personal home folder, /home/username

Project storage, owned by PI, /proj/ourstuff

Work node local scratch disk (during run)

Recover deleted files?

Backup?

Snapshot?

yes!

yes!

no!

yes!

no!

no!

Backup > slow tape-drive > staff needs to check

Snapshot [7 days] > on disk > you can check directly

Backup your Data

- **Make your own backup** (also of work computer)!

- Avoid learning it the hard way...

- Use e.g. `rsync` `$ man rsync`

```
$ rsync -av username@tetralith.nsc.liu.se:datafolder .
```

- Your Univ. IT-dep. might help
- Data is **never 100% safe**, always some risk

Files & Folders

Permissions, who can read/write/execute a folder or file?

```
$ ls -l
-rw-rw-r-- 1 weiol pg_nsc          0 Apr 19 20:35 testingfile
```

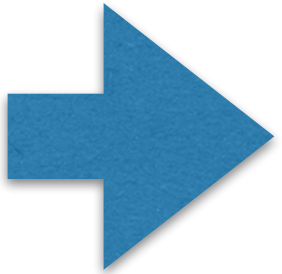
[-/d, file/dir], user [rwx], group [rwx], all [rwx]

Change permissions/group/owner using: chmod, chgrp, chown

Tip: Create a link to your project storage:

```
$ pwd
/home/username
$ ln -s /proj/ourproject/users/username ourproject
$ ls -l
total 1
lrwxrwxrwx 1 username group 26 Apr 18 01:06 ourproject -> /proj/ourproject/users/username
$ cd ourproject
```

Further Notes

- **Quota**, both file **size** & **numbers**!
 - Compress: tar and zip \$ `tar cfz Results.tar.gz Results`
- Odd sudden failures? Check your quota! \$ `snicquota`
- **Center storage:**
 - OK for analysis & post-processing
 - Not for long-term archiving (SNIC)  Univ. IT-dep?



Basic Security

- Unique password (non-trivial, but not overly complicated)
- Suspicion that your account is compromised -> contact NSC
- Don't hesitate to contact us!
- Sharing accounts is not allowed (accounts are personal)
Share files e.g. by managing project memberships and use /proj