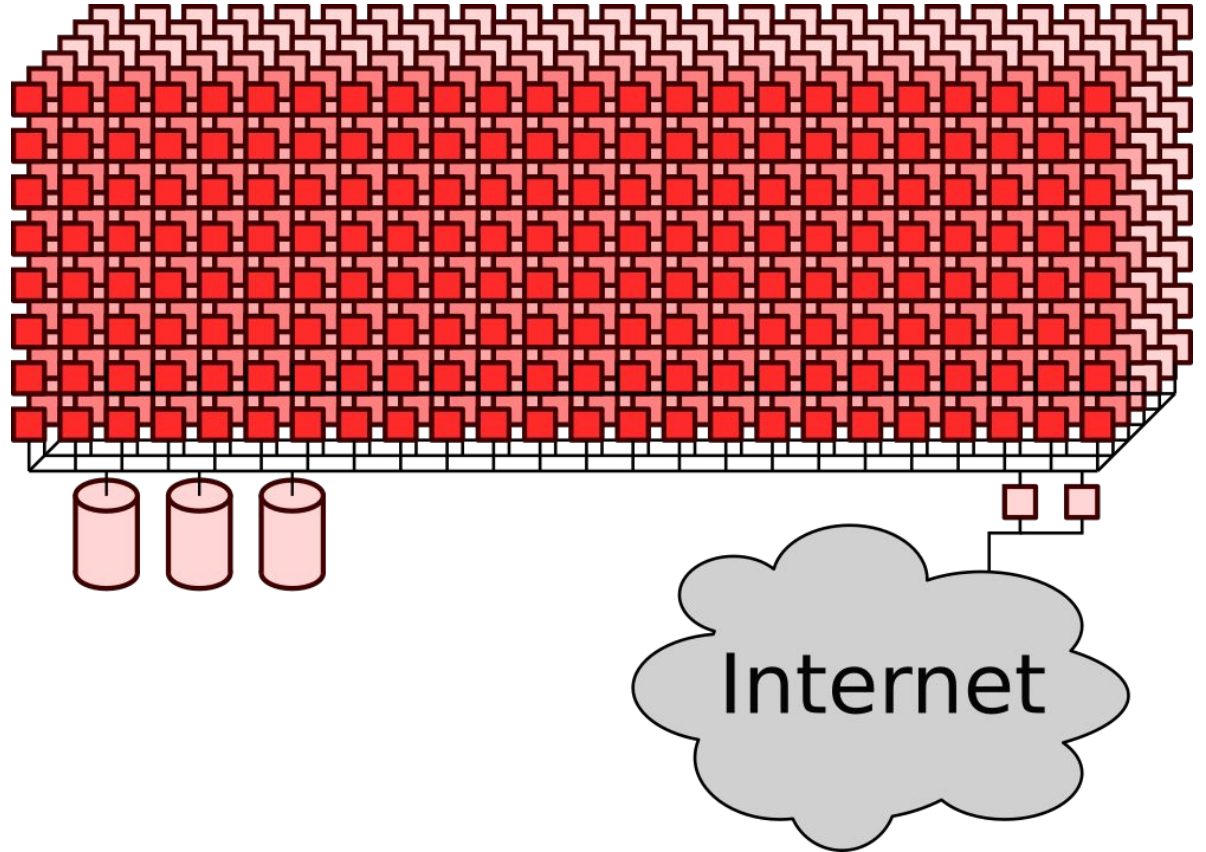


# 2. HPC: Getting started

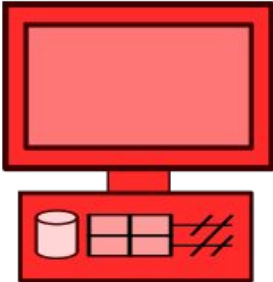
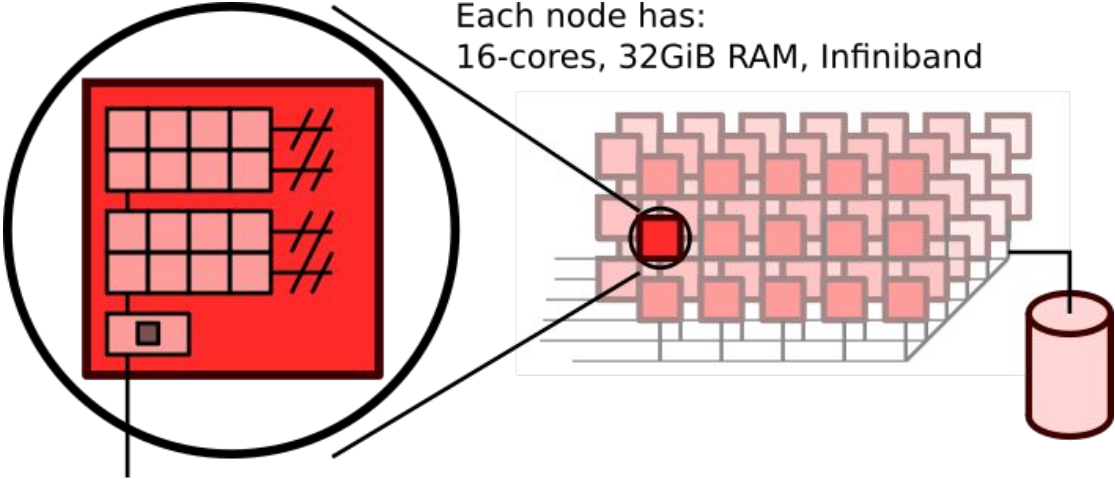
Peter Kjellström

National Supercomputer Centre (NSC)

# Desktop PC vs HPC: Scale



# Desktop PC vs HPC: Scale



Typical PC has  
4-cores, 8GiB RAM

# Desktop PC vs HPC

- Shared resource
  - Your desktop is your own, it's not shared
  - Typical national level HPC systems are shared by hundreds of researchers
- Workflow
  - On a desktop the interfaces are direct and immediate
  - On an HPC resource work typically happens in “batch mode”. That is, most work is prepared and queued just to be run at a later time when resources are available

# Linux - the operating system on most HPC machines

The operating system tries to balance sharing and separation while providing flexibility to new hardware/software/use-cases.

Where to find information:

- **man *command*** gives a listing of the manual for *command*
- Internet
  - Linux guides
  - Forums
- Most commands will accept “--help” and/or “-h”
- Ctrl-c interrupts / stops programs on the command line

Linux guide, e.g.: [http://www.dsl.org/cookbook/cookbook\\_toc.html](http://www.dsl.org/cookbook/cookbook_toc.html)

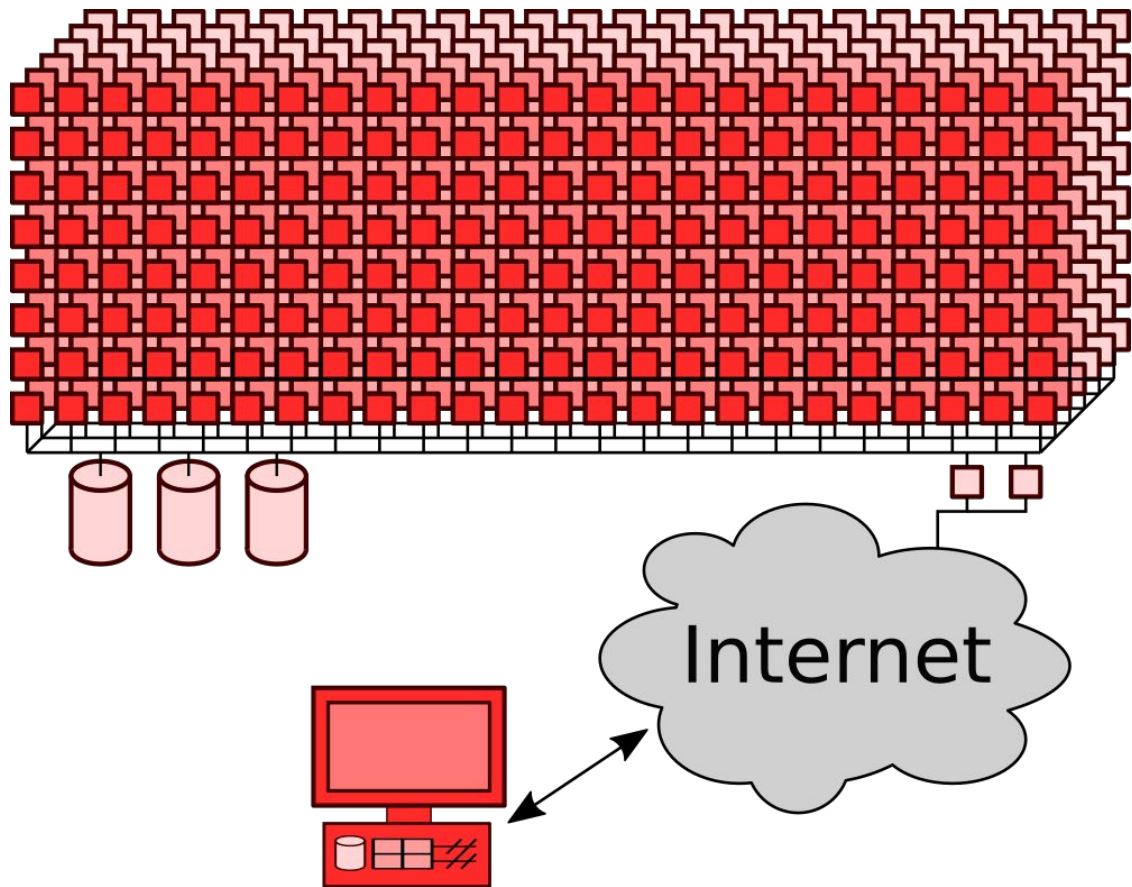
Linux forum e.g.: <https://stackoverflow.com/questions/tagged/linux>

# Access to HPC: login

Triolith has several (currently two) login nodes

Login nodes sit on the edge between the real compute resource and the network

Storage visible across the entire system



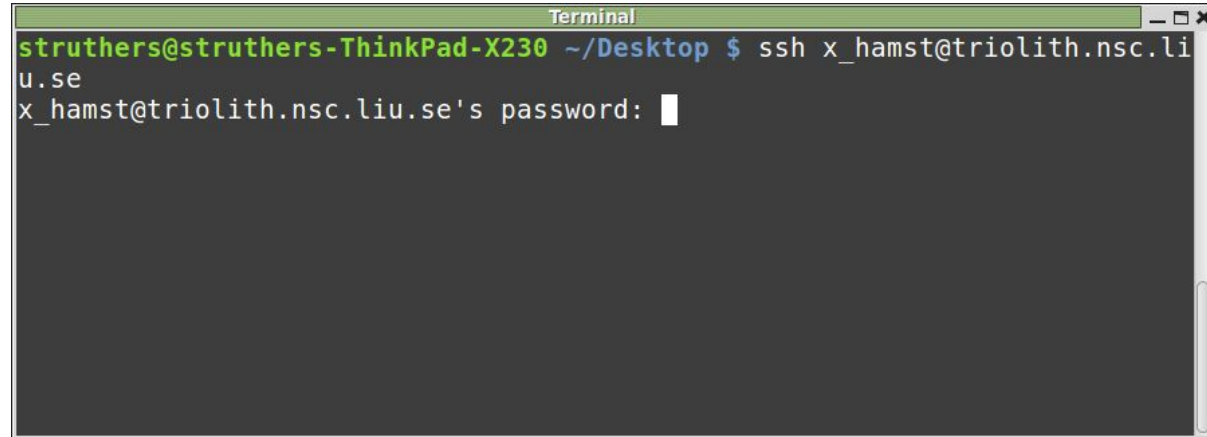
# Access to HPC: login the classical way (ssh)

Typical login via terminal from Linux / Mac:

```
ssh username@triolith.nsc.liu.se
```

Windows: can use PuTTY

Note: if you want to end up on a specific login node use `triolith1.nsc.liu.se` or `triolith2.nsc.liu.se`

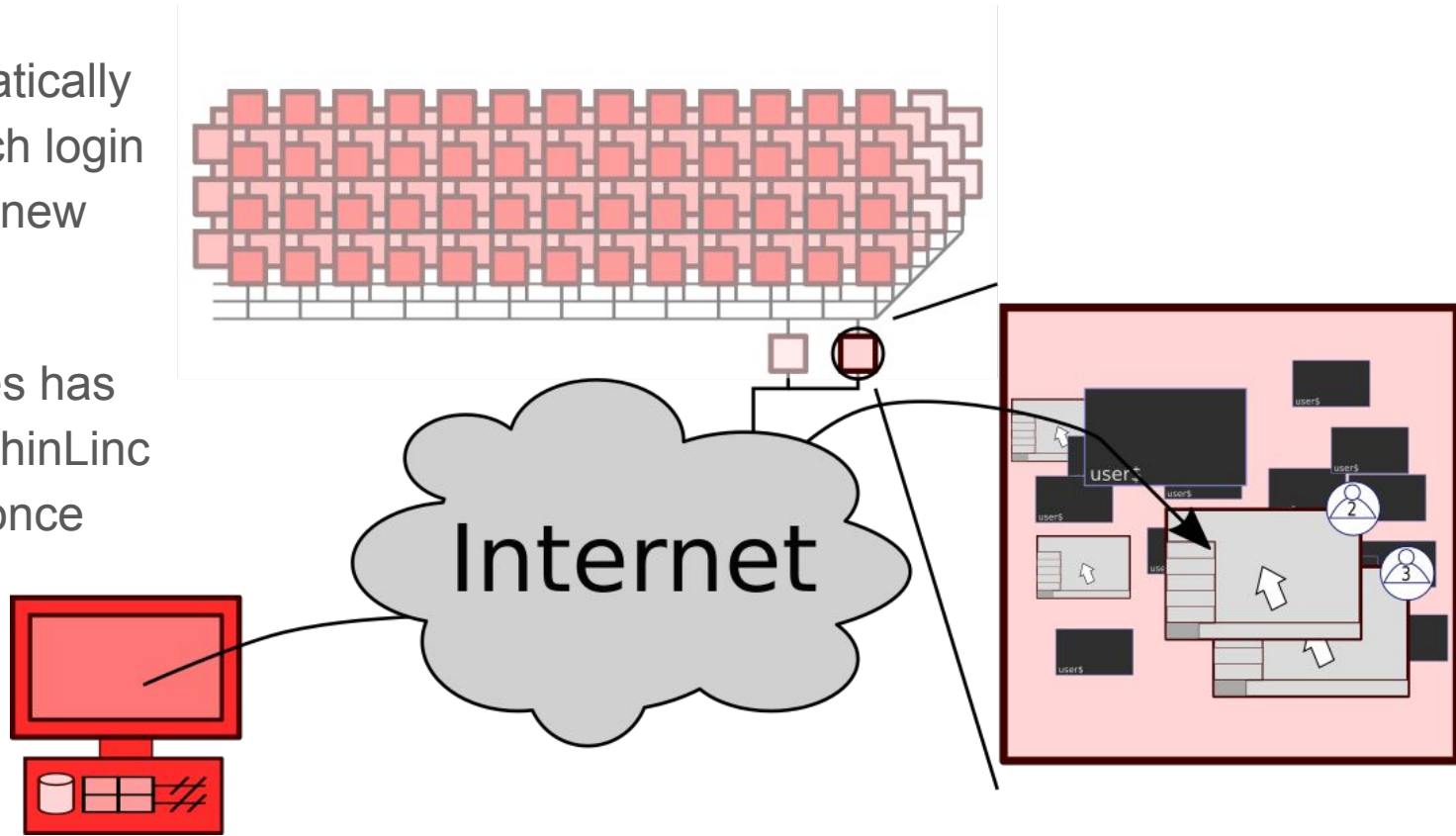
A terminal window titled "Terminal" with a green title bar. The window shows a command prompt on a Linux system. The prompt is "struthers@struthers-ThinkPad-X230 ~/Desktop \$". The user has entered the command "ssh x\_hamst@triolith.nsc.liu.se". The terminal output shows "u.se" on the next line, followed by "x\_hamst@triolith.nsc.liu.se's password:" and a cursor. The terminal background is dark gray, and the text is light gray/green.

```
Terminal  
struthers@struthers-ThinkPad-X230 ~/Desktop $ ssh x_hamst@triolith.nsc.li  
u.se  
x_hamst@triolith.nsc.liu.se's password: █
```

# Access to HPC: login using virtual desktop (ThinLinc)

ThinLinc automatically determines which login node is best for new sessions

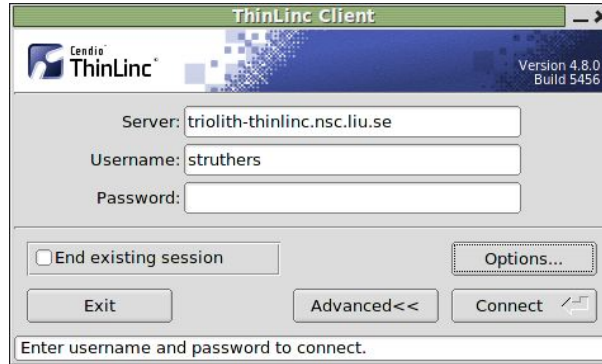
Each login nodes has many ssh and ThinLinc users active at once





# ThinLinc: The ThinLinc client

First step: you need a ThinLinc client installed on your computer.

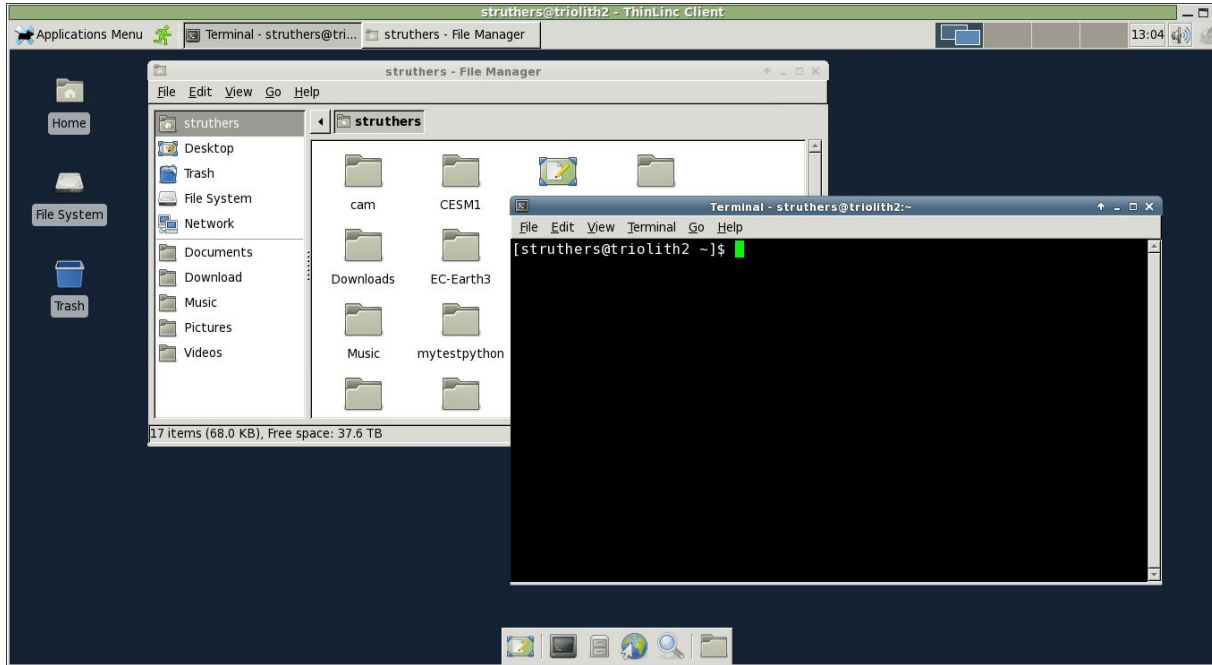


ThinLinc download: <https://www.cendio.com/thinlinc/download>

Cendio webpage: <https://www.cendio.com/thinlinc/what-is-thinlinc>

NSC documentation: <https://www.nsc.liu.se/support/graphics/>

# ThinLinc: The ThinLinc desktop



# ThinLinc: Advantages

- Provides a desktop interface to the system
- Allows for persistent sessions

(similar functionality can be had in a text terminal using screen or tmux)

- **NSC recommends** ThinLinc for graphical applications (better than X-forwarding)
- Hardware accelerated 3D graphics possible for some application (vglrun)

# Common tools

- Text editors
  - vi, gedit, nedit, emacs, nano, kwrite
- Plotting graphs
  - gnuplot, grace
- Analysis (basic/complex)
  - Python, R, Matlab, Mathematica (local LiU license)

# Security: Basics

You should use a **unique password**. The password you choose should be non-trivial but doesn't have to be overly complicated.

**If you suspect that your account is compromised contact NSC.** Don't be afraid to contact NSC support if you notice anything suspicious.

Sharing accounts is not allowed, that is, accounts are personal. If you want to share files manage project memberships and use /proj

# Security: key-pair authentication

Key-pair or public-key authentication is an alternative to traditional passwords.

Things to consider:

- Can be much more secure than regular password authentication.
- Can be less secure if used incorrectly (understand before use).
- Allows multiple logins without reentering password/pass phrase.

Initially every account starts with just password login possible.

# Files: Basic introduction

There are three storage areas available to Triolith and Gamma users:

1. Your personal home directory (e.g */home/x\_abcde*).
2. Project storage directories (e.g */proj/myproject*). You can have access to multiple different project directories. Owned by the project (PI)
3. The node local disk in each compute node (only during jobs).

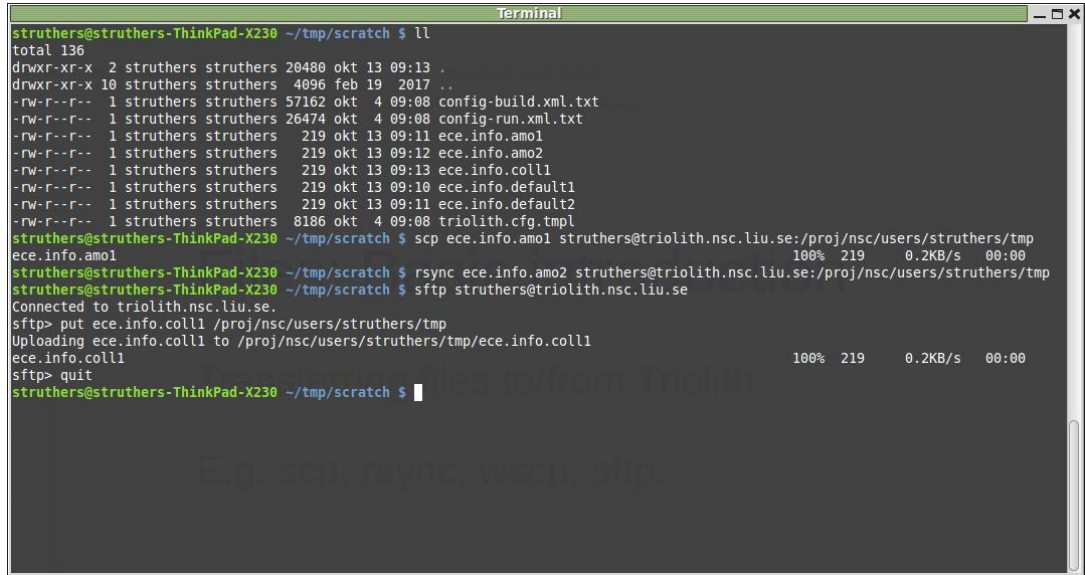
Notes:

- The command `snicquota` will show how much space is available to you in each location.
- Project storage is linked to specific project allocations and life time

# Files: Transferring data to/from Triolith

- Linux: scp, rsync, sftp
- Mac: Filezilla, scp
- Windows: PuTTY, Filezilla, winscp

Hint: use tar/compression when transferring large numbers of small files

A terminal window titled "Terminal" showing a series of commands and their outputs. The user is in a directory ~/tmp/scratch. They run 'll' to list files, showing a directory listing with permissions, owners, sizes, and dates. Then they run 'scp ece.info.amo1 struthers@triolith.nsc.liu.se:/proj/nsc/users/struthers/tmp/ece.info.amo1', which shows progress at 100%. Next, they run 'sftp struthers@triolith.nsc.liu.se', which connects to the remote host. They then run 'put ece.info.coll1 /proj/nsc/users/struthers/tmp', showing progress at 100%. Finally, they run 'quit' to end the sftp session.

```
struthers@struthers-ThinkPad-X230 ~/tmp/scratch $ ll
total 136
drwxr-xr-x  2 struthers struthers 20480 okt 13 09:13 .
drwxr-xr-x 10 struthers struthers  4096 feb 19  2017 ..
-rw-r--r--  1 struthers struthers 57162 okt  4 09:08 config-build.xml.txt
-rw-r--r--  1 struthers struthers 26474 okt  4 09:08 config-run.xml.txt
-rw-r--r--  1 struthers struthers   219 okt 13 09:11 ece.info.amo1
-rw-r--r--  1 struthers struthers   219 okt 13 09:12 ece.info.amo2
-rw-r--r--  1 struthers struthers   219 okt 13 09:13 ece.info.coll1
-rw-r--r--  1 struthers struthers   219 okt 13 09:10 ece.info.default1
-rw-r--r--  1 struthers struthers   219 okt 13 09:11 ece.info.default2
-rw-r--r--  1 struthers struthers   8186 okt  4 09:08 triolith.cfg.tpl
struthers@struthers-ThinkPad-X230 ~/tmp/scratch $ scp ece.info.amo1 struthers@triolith.nsc.liu.se:/proj/nsc/users/struthers/tmp/ece.info.amo1
100% 219   0.2KB/s  00:00
struthers@struthers-ThinkPad-X230 ~/tmp/scratch $ rsync ece.info.amo2 struthers@triolith.nsc.liu.se:/proj/nsc/users/struthers/tmp
struthers@struthers-ThinkPad-X230 ~/tmp/scratch $ sftp struthers@triolith.nsc.liu.se
Connected to triolith.nsc.liu.se.
sftp> put ece.info.coll1 /proj/nsc/users/struthers/tmp
Uploading ece.info.coll1 to /proj/nsc/users/struthers/tmp/ece.info.coll1
ece.info.coll1                               100% 219   0.2KB/s  00:00
sftp> quit
struthers@struthers-ThinkPad-X230 ~/tmp/scratch $
```



# Files: data safety and file recovery

Data is never completely safe, there is always a small chance that it can be lost.

- */home* storage is backed up has snapshots
- */proj* storage **nope** has snapshots

Recovering files lost due to mistakes is best done using snapshots:

- Read-only, point in time **snapshots** are taken at certain intervals. These can be used to recover deleted files and are kept for 24h (hourly snapshots) or 7 days (daily snapshots).

```
$ ls /home/.snapshots
```

Data safety: <https://www.nsc.liu.se/support/storage/snic-centrestorage/is-my-data-safe/>

Recovering files: <https://www.nsc.liu.se/support/storage/snic-centrestorage/recover-deleted-files/>

# Files: Node local storage

- Each compute node has a local hard disk (450GiB).
- The environment variable `$SNIC_TMP` points to a writable directory (*/scratch/local*) on the local disk that you can use.
- Anything stored on the local disk is **deleted** when your job ends.

Node local storage is available for storing temporary files that are only needed during a batch job. This storage is **not shared** between compute nodes.