

Grid computing

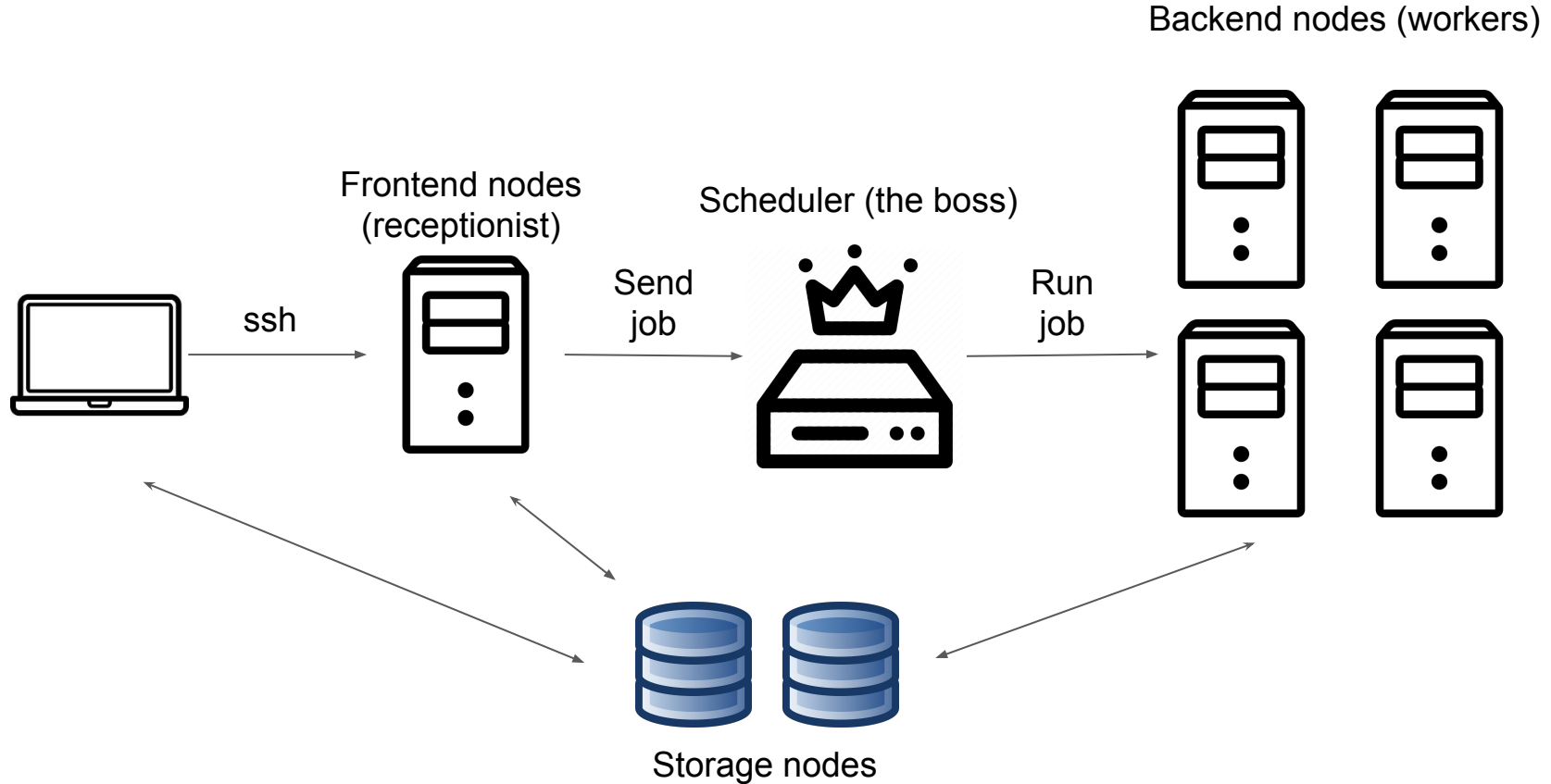
Ing. Stanislav Smatana

Grid

- Collection of computers, often geographically distributed, pooling resources to increase computational power
- We will talk primarily about czech national grid infrastructure Metacentrum
- Allows you to ask for resources (memory, processors, HDD space) you could not easily achieve on your computer



Structure of a grid supercomputer



Connect to frontend node

ssh username@host

HOST	HOME DIRECTORY
zuphux.cerit-sc.cz	/storage/brno3-cerit/home/
skirit.ics.muni.cz	/storage/brno2/home/
alfrid.meta.zcu.cz	/storage/plzen1/home/
tarkil.grid.cesnet.cz	/storage/praha1/home/
nympha.zcu.cz	/storage/plzen1/home/
minos.zcu.cz	/storage/plzen1/home/
perian.ncbr.muni.cz	/storage/brno2/home/

Send a job request

`qsub -l walltime=01:00:00 -l select=1:ncpus=6:mem=4gb:scratch_local=10gb -l`

Maximal job running time (hh:mm:ss)

Number of computers you want (usually 1)

How many processors do you need ?

How much memory do you need ?

Do you need special fast storage ? If so, how much ?

I want interactive job !

The diagram illustrates the components of the `qsub` command. The command is shown in a single line: `qsub -l walltime=01:00:00 -l select=1:ncpus=6:mem=4gb:scratch_local=10gb -l`. Below the command, several red text annotations are connected to specific parts of the command by arrows. An upward-pointing arrow from 'Maximal job running time (hh:mm:ss)' points to `walltime=01:00:00`. A downward-pointing arrow from 'Number of computers you want (usually 1)' points to `select=1`. An upward-pointing arrow from 'How many processors do you need ?' points to `ncpus=6`. A downward-pointing arrow from 'How much memory do you need ?' points to `mem=4gb`. An upward-pointing arrow from 'Do you need special fast storage ? If so, how much ?' points to `scratch_local=10gb`. A downward-pointing arrow from 'I want interactive job !' points to the final `-l` flag.

Execute

- After your job has been granted resources, you can enter commands into the terminal like on you do on your Linux !
- There is plenty of software pre-installed on the Metacentrum
- To prevent clashes, software is accessed through **modules**

```
module add fastQC-0.11.5
```

↑
Name of the module to add.

Getting data to the grid

```
scp path_to_local_file username@storage_server:~/path/to/folder
```

storage-brno3-cerit.metacentrum.cz	<i>/storage/brno3-cerit/</i>
storage-plzen1.metacentrum.cz	<i>/storage/plzen1/</i>
storage-brno12-cerit.metacentrum.cz	<i>/storage/brno12-cerit/</i>
storage-praha1.metacentrum.cz	<i>/storage/praha1/</i>

Tips and Tricks

- It is also possible to run a script of choice instead of interactive sessions
- Do not compute on front nodes !
- Always use absolute paths (readlink -f file)
- Further resources:
 - https://wiki.metacentrum.cz/wiki/Beginners_guide
 - https://wiki.metacentrum.cz/wiki/Working_with_data
 - https://wiki.metacentrum.cz/wiki/FAQ/Grid_computing