Cluster-user meeting

Tom de Geus
Bart Vossen
Jeroen van Beeck
Leo Wouters

L.H.G.Wouters@tue.nl
MaTeCluster@tue.nl

www.mate.tue.nl/~cluster

- Cluster manual
- This presentation
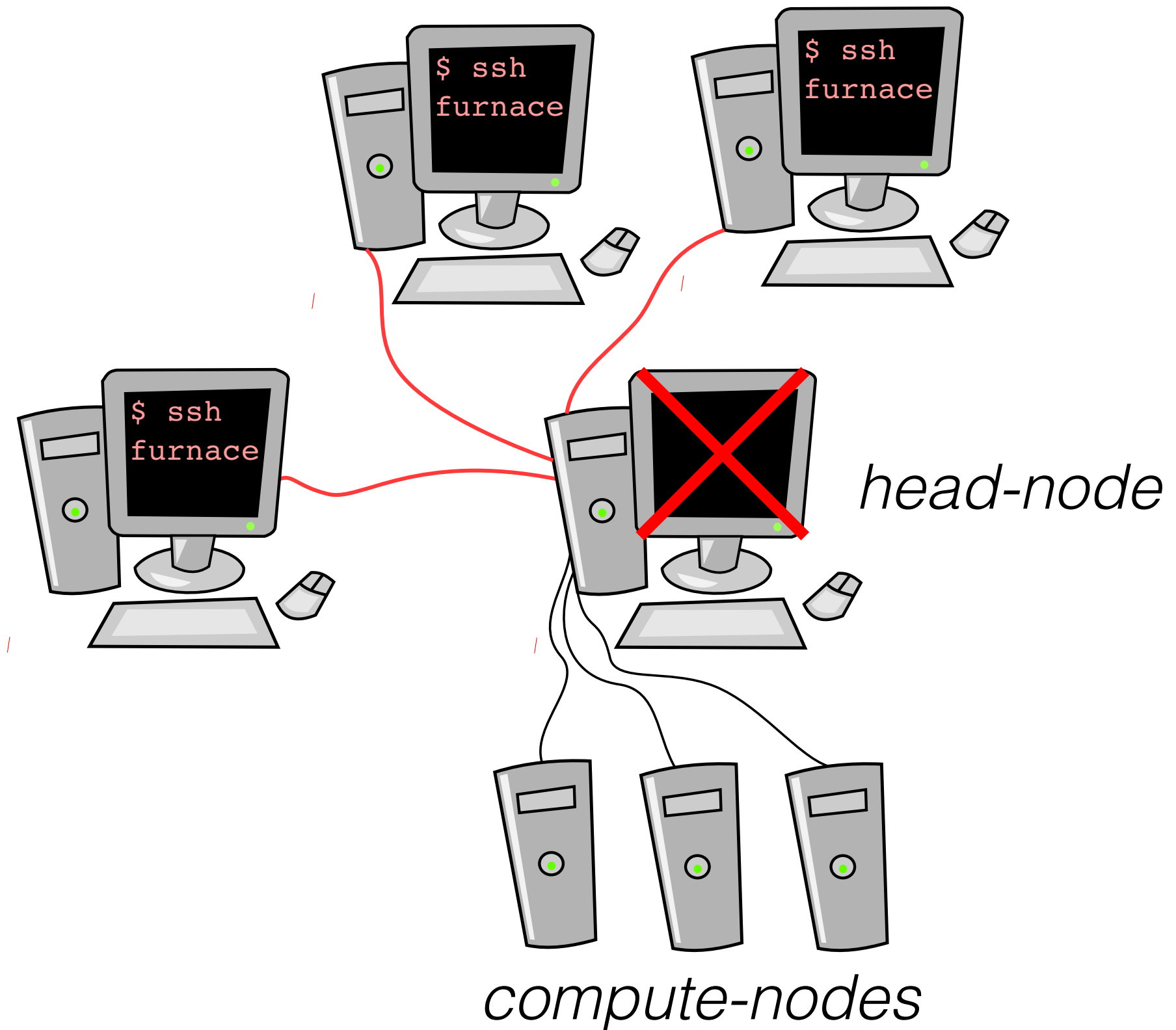- …

Feel free to interrupt at any time

Outline

increase knowledge
and
stimulate discussion/collaboration

1. Computing cluster, the basics…
2. Monitoring jobs
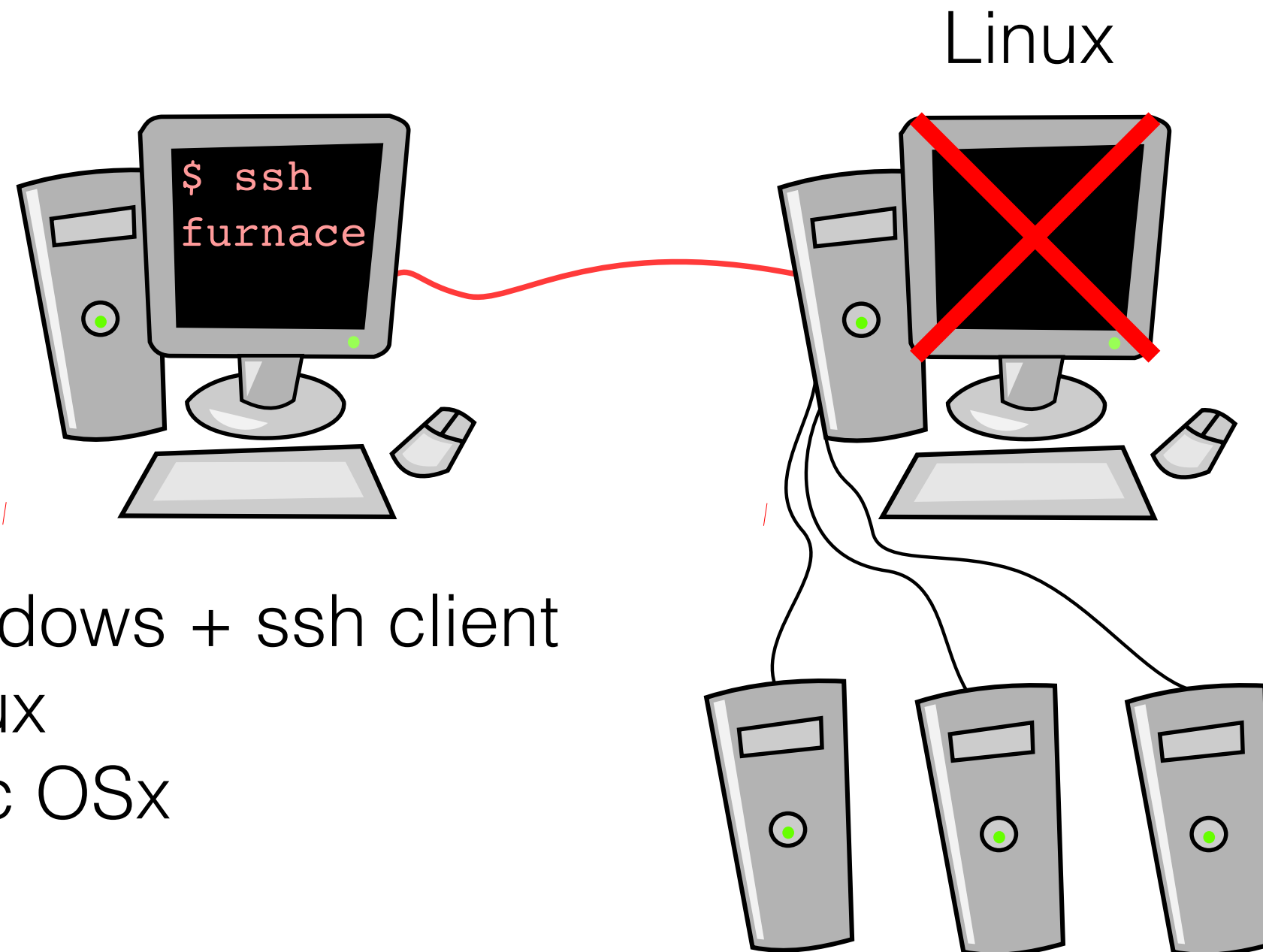3. Etiquette
4. Getting help
5. Future

The cluster

Interfacing with the cluster

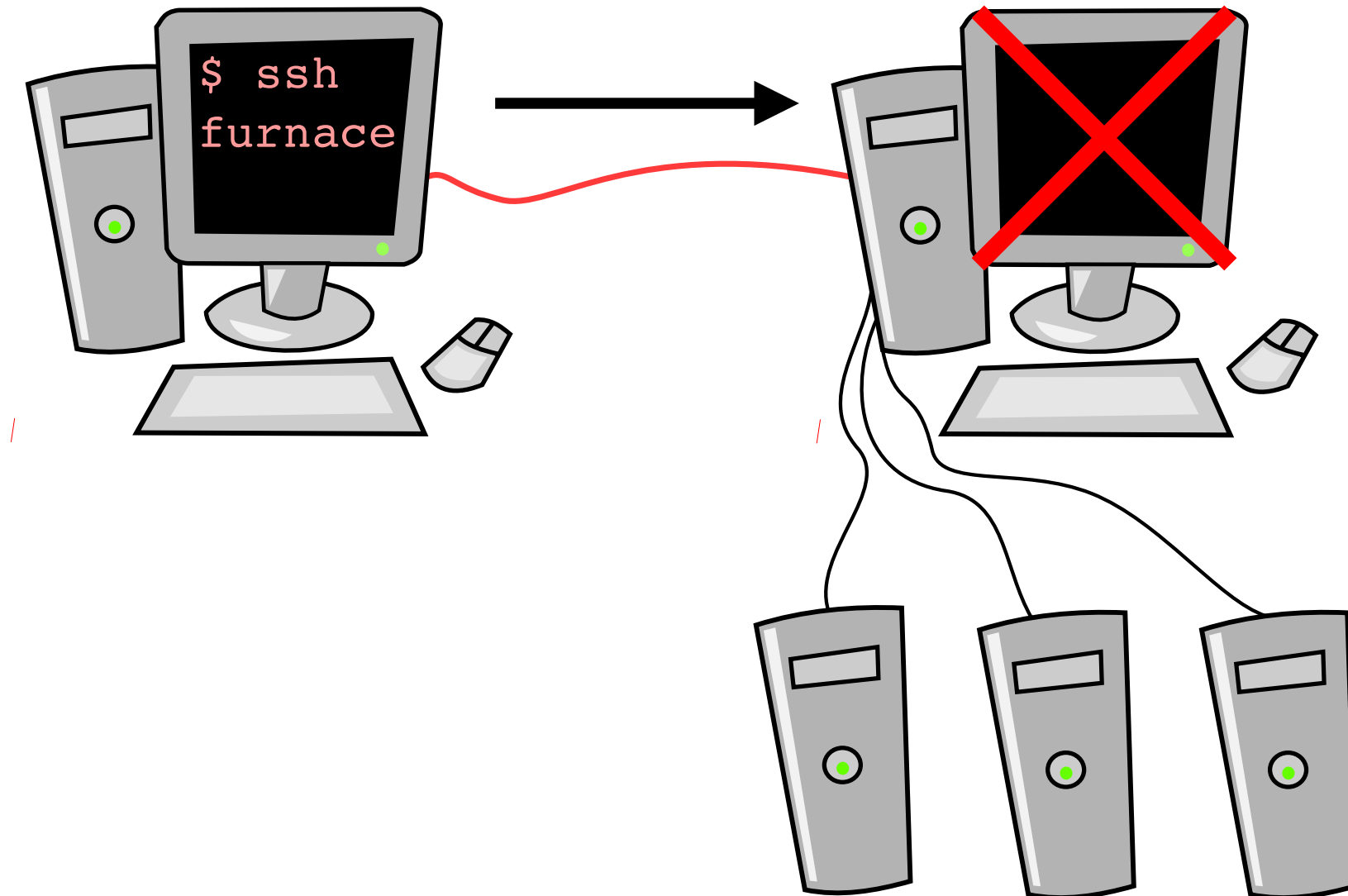Linux

```
$ ssh
furnace
```

- Windows + ssh client
- Linux
- Mac OSx

Connecting

```
$ ssh tdegeus@furnace.wfw.wtb.tue.nl
```

Connecting: details
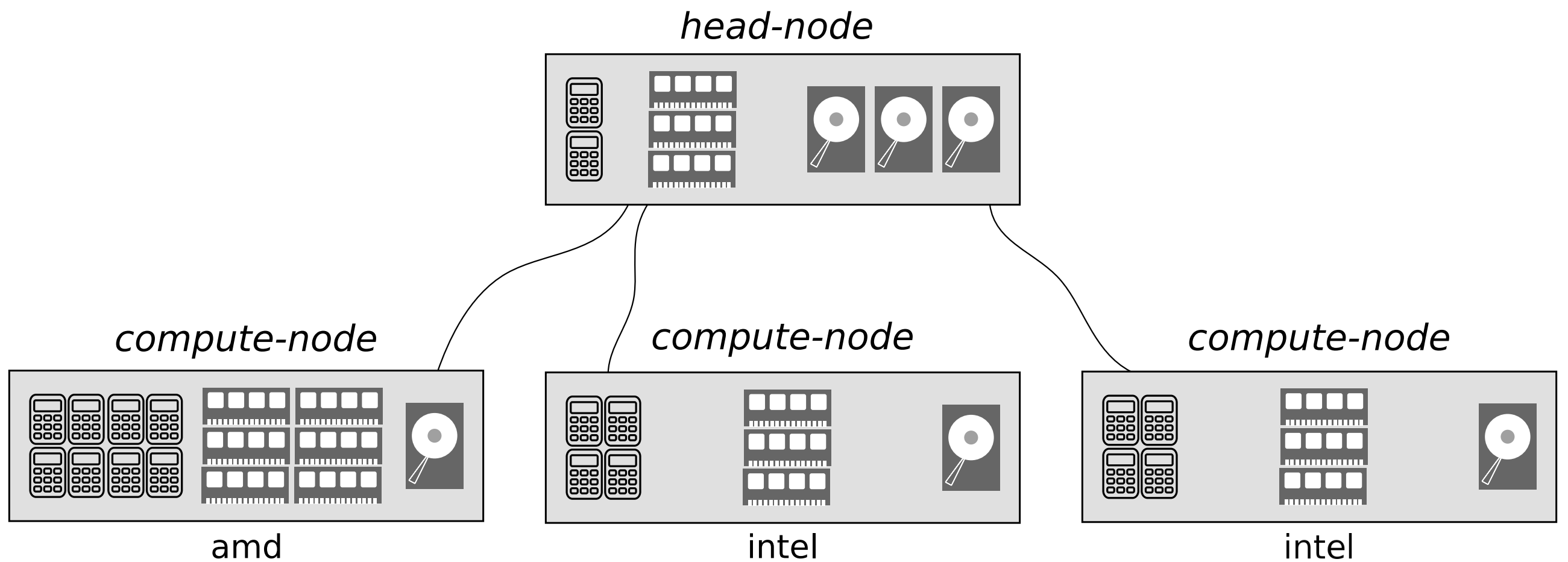
head-node

compute-node

amd

compute-node

intel

compute-node

intel

Connecting: details

```
$ ssh tdegeus@furnace.wfw.wtb.tue.nl

$ pwd
/home/tdegeus
```

head-node

compute-node

compute-node

compute-node

amd

intel

intel

Connecting: details

```
$ ssh tdegeus@furnace.wfw.wtb.tue.nl

$ pwd
/home/tdegeus
```

head-node

compute-node

compute-node

compute-node

amd
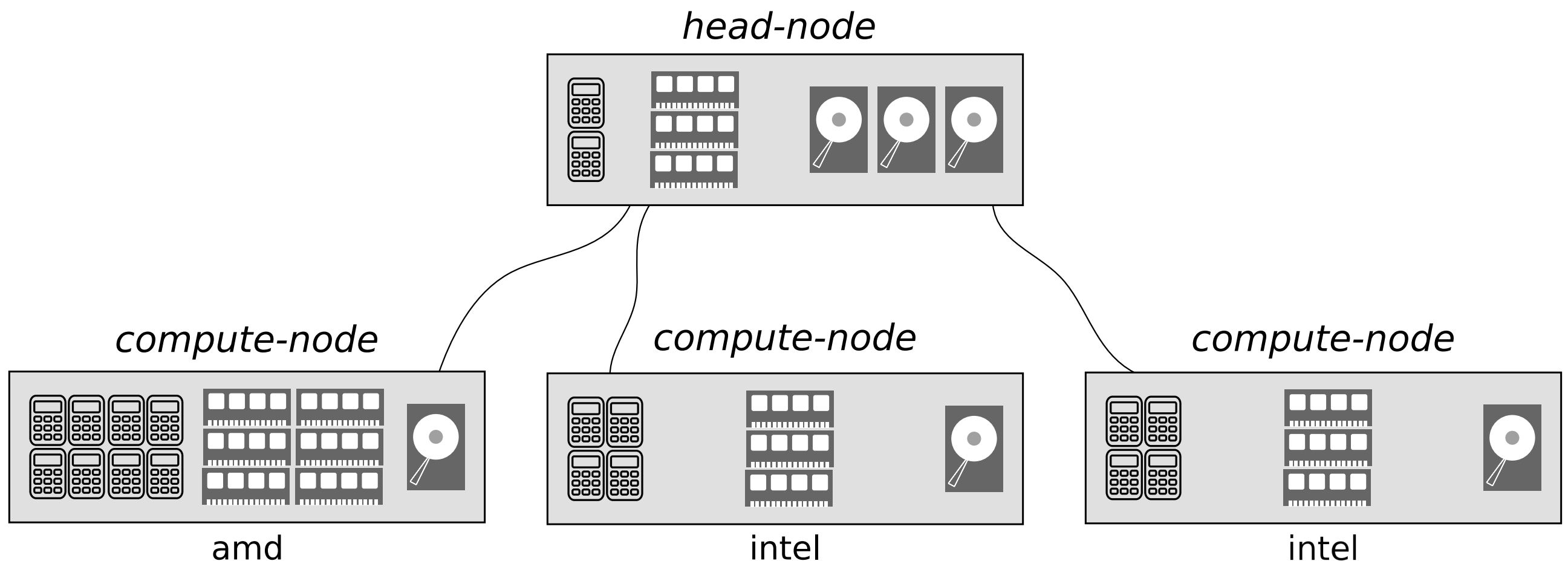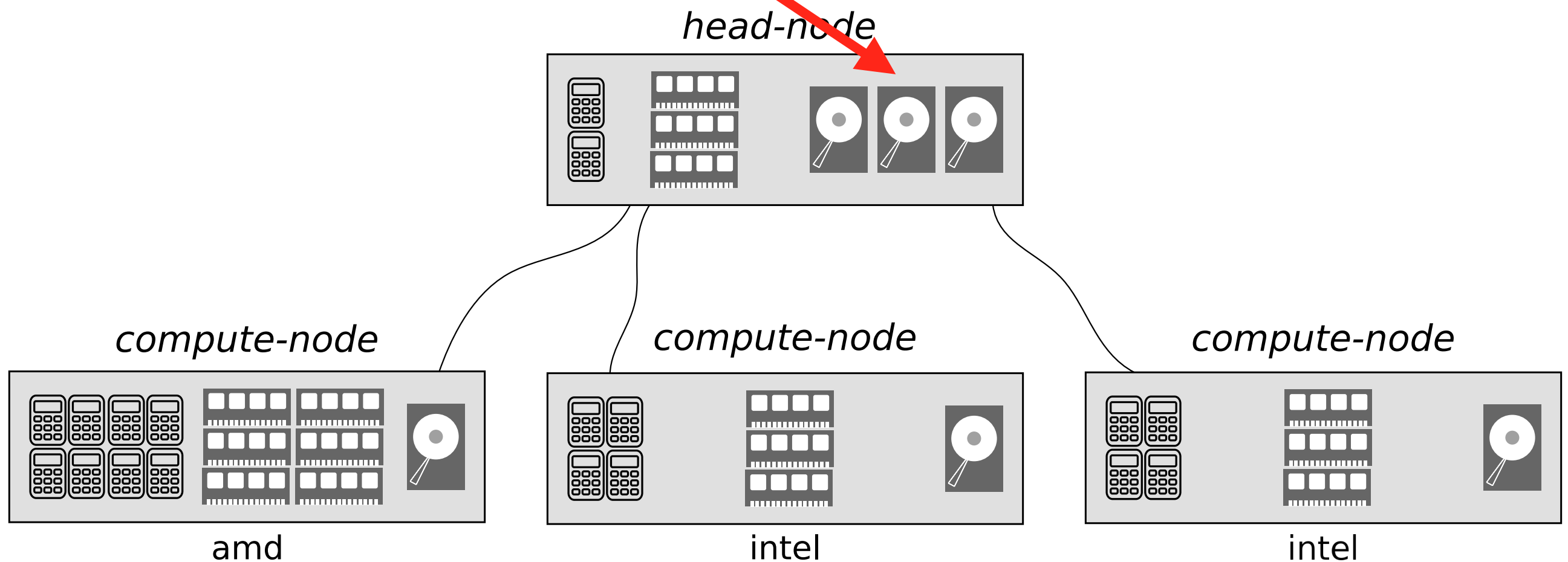
intel

intel

Two types of nodes
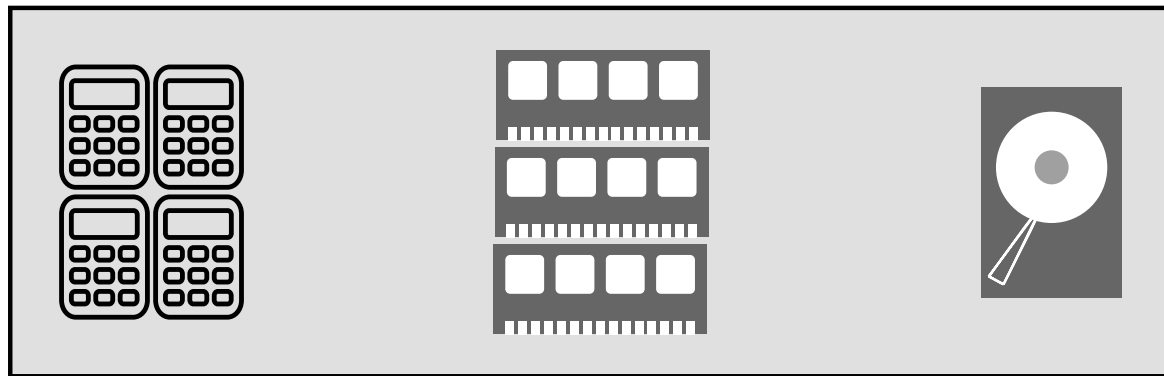
intel

- Fast
- 8 cpus per node
- Limited amount of memory
  - 48 GB
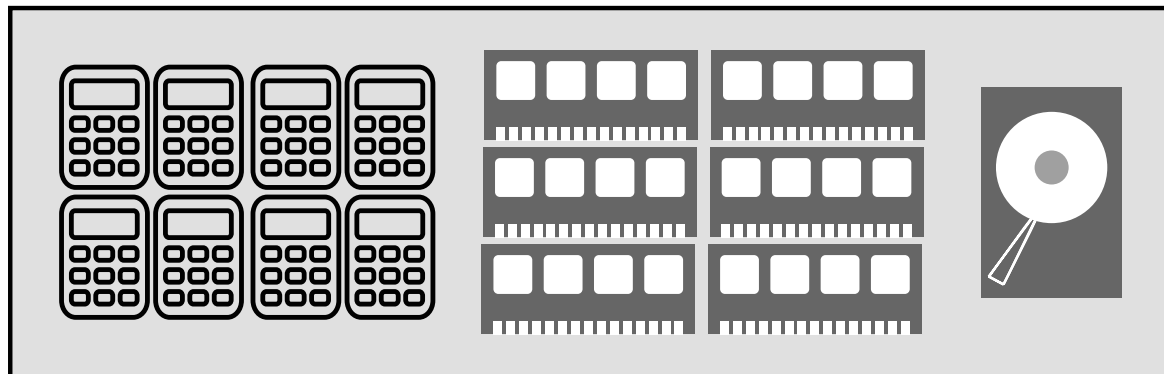
amd

- Slow
- 24 cpus per node
- Large amount of memory
  - 132 GB

The scheduler assigns jobs, once running:
everything is shared

Limited resources
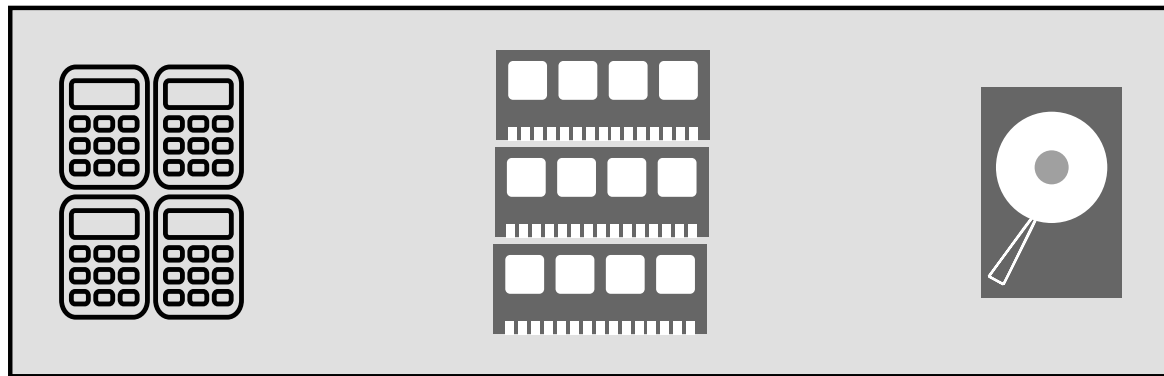
everything is shared

intel

CPUs:
- use > claim: every job slows down
- claim > use: blocking part of the cluster

Out of memory / out of disk space:
node is killed, all jobs fail!

Claiming resources

- Claim (a number) of cpus
  - `#PBS -l nodes=1:ppn=1:intel`   *most jobs*
  - `#PBS -l nodes=1:ppn=2:amd`   *heavy memory*

- Heavy memory jobs, claim memory
  - `#PBS -l pmem=20gb`
  - `#PBS -l pvmem=20gb`

Claim what you use and use what you claim

Submitting a job

```
$ qsub myjob.pbs
```

head-node

scheduler

compute-node

compute-node

compute-node

Submitting a job

```
$ qsub myjob.pbs
```

```bash
#!/bin/bash
#PBS -S /bin/bash
#PBS -j oe
#PBS -o pbs.out
#PBS -l nodes=1:ppn=1:intel

# print the working directory
pwd

# print a personal string
echo "Hello world"
```

head-node

scheduler

compute-node   *compute-node*   *compute-node*
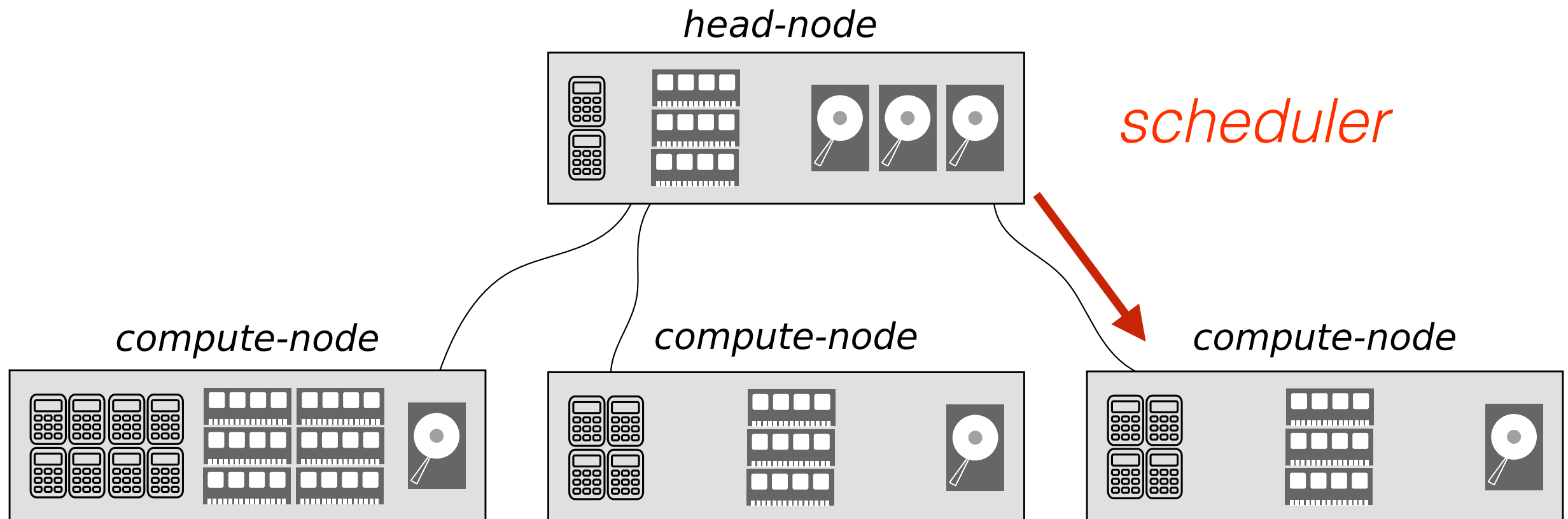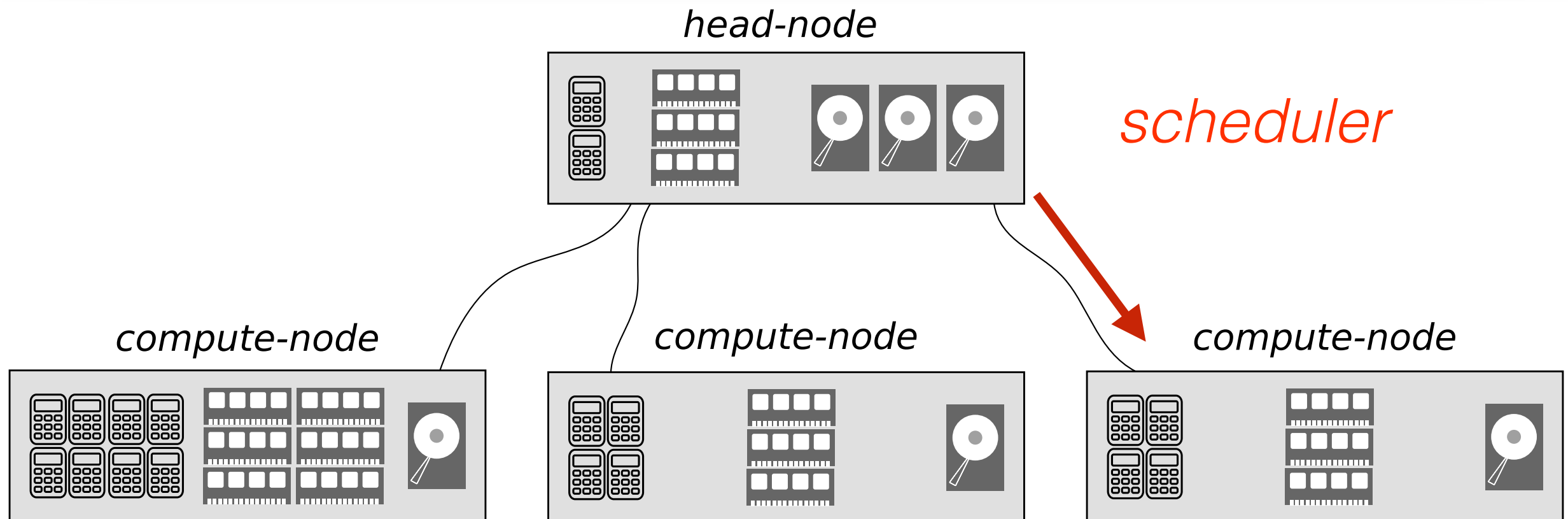
Submitting a job

```
$ qsub myjob.pbs
```

```
#!/bin/bash
#PBS -S /bin/bash          script interpreter
#PBS -j oe
#PBS -o pbs.out            scheduler directives
#PBS -l nodes=1:ppn=1:intel

# print the working directory
pwd                        program

# print a personal string
echo "Hello world"
```

head-node
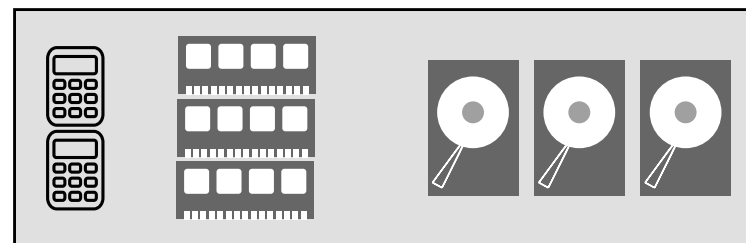
scheduler
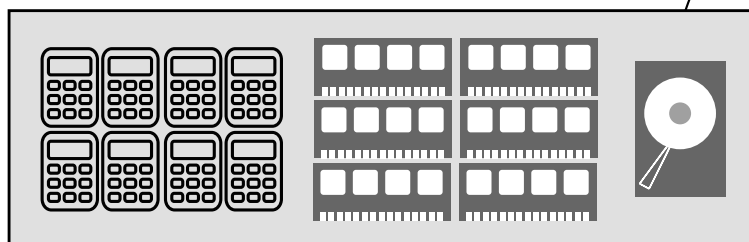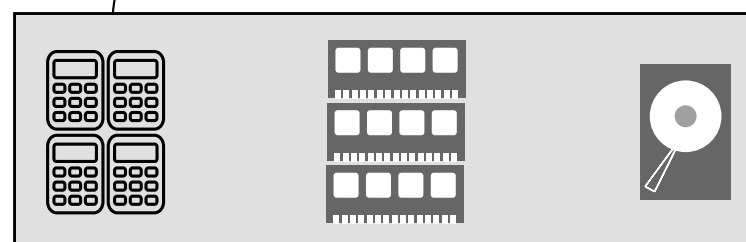
compute-node   *compute-node*   *compute-node*
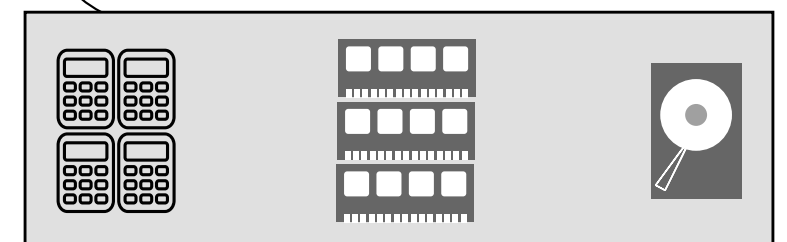
Monitoring

```
$ myqstat

jid     , owner     , job name      , host , cpus    , mem   , S , time , score
------  , --------- , ------------- , ---- , ------  , ----  , - , ---- , -----
202169  , tdegeus   , myjob.pbs     ,   11 , 1:1:i   , 1mb   , R ,  10s ,  1.00
...
```

> 30/100 GB
risky

<<1 waste
=1   good
>1   harming
     others

Etiquette: monitor resources

Monitor:

- memory usage
- CPU usage
- disk usage

Etiquette: monitor resources

```
$ myqstat -N --long


Node, State         , Type , Ctot, Cused, Cfree, Load , Mtot , Mused, Mem%, HDtot, HDfree, HD% , Btot, Bin  , Bout
----, -------------, -----, ----, -----, -----, -----, -----, -----, ----, -----, ------, ----, ----, -----, ----
   0, job-exclusive,   amd,   24,    24,     0,  1.24, 132gb,  18gb, 0.09, 177gb, 110gb, 0.38,  8kb,   8kb, 329b
   1, job-exclusive,   amd,   24,    24,     0,  1.54, 132gb,  19gb, 0.10, 177gb, 120gb, 0.32,  8kb,   8kb, 326b
   2, job-exclusive,   amd,   24,    24,     0, 24.00, 132gb,  11gb, 0.06, 419gb, 378gb, 0.10,  8kb,   8kb, 380b
   3, job-exclusive,   amd,   24,    24,     0,  1.51, 132gb,  20gb, 0.10, 177gb, 149gb, 0.16,  8kb,   8kb, 320b
   5, job-exclusive,   amd,   24,    24,     0,  1.33, 132gb,  23gb, 0.11, 177gb, 145gb, 0.18,  8kb,   8kb, 320b
   6, job-exclusive,   amd,   24,    24,     0, 24.00, 132gb,  15gb, 0.08, 177gb, 151gb, 0.15,  8kb,   8kb, 353b
   7, job-exclusive,   amd,   24,    24,     0, 24.00, 132gb,  16gb, 0.08, 177gb, 157gb, 0.12,  8kb,   8kb, 344b
   8, job-exclusive,   amd,   24,    24,     0, 24.00, 132gb,  12gb, 0.06, 177gb, 161gb, 0.09,  8kb,   8kb, 339b
   9, job-exclusive,   amd,   24,    24,     0, 24.00, 132gb,  10gb, 0.05, 177gb, 161gb, 0.09,  8kb,   8kb, 332b
  10,         free,   amd,   24,     6,    18,  1.18, 132gb,  19gb, 0.10, 419gb, 388gb, 0.07,  8kb,   8kb, 319b
  12,         free, intel,    8,     5,     3,  5.00,  49gb,  13gb, 0.12, 419gb, 368gb, 0.12,  8kb,   8kb, 296b
  11,         free, intel,    8,     2,     6,  2.00,  49gb,   1gb, 0.01, 419gb, 366gb, 0.13,  8kb,   8kb, 289b
  14, job-exclusive, intel,    8,     8,     0,  8.00,  49gb,   4gb, 0.03, 904gb, 850gb, 0.06,  8kb,   8kb, 280b
  13, job-exclusive, intel,    8,     8,     0,  8.00,  49gb,   4gb, 0.03, 177gb, 155gb, 0.12,  8kb,   8kb, 239b
  16, job-exclusive, intel,    8,     8,     0,  8.00,  49gb,   4gb, 0.04, 177gb, 147gb, 0.17,  8kb,   8kb, 242b
  15, job-exclusive, intel,    8,     8,     0,  8.00,  49gb,   5gb, 0.04, 177gb, 143gb, 0.19,  8kb,   8kb, 271b
  18, job-exclusive, intel,    8,     8,     0,  1.33,  49gb,   2gb, 0.02, 177gb,  89gb, 0.50, 21mb, 587kb, 20mb
  17, job-exclusive, intel,    8,     8,     0,  6.34,  49gb,  11gb, 0.10, 177gb, 137gb, 0.23,  8kb,   8kb, 270b
  20, job-exclusive, intel,    8,     8,     0,  1.53,  49gb,  19gb, 0.16, 177gb, 130gb, 0.27,  8kb,   8kb, 239b
  21, job-exclusive, intel,    8,     8,     0,  1.76,  41gb,  18gb, 0.17, 177gb, 142gb, 0.20,  8kb,   8kb, 253b
  22, job-exclusive, intel,    8,     8,     0,  8.00,  49gb,   4gb, 0.04, 177gb, 141gb, 0.20,  8kb,   8kb, 239b
  19,         free, intel,    8,     2,     6,  2.00,  49gb,   2gb, 0.02, 177gb, 109gb, 0.39,  8kb,   8kb, 274b
  23,         down, intel,    8,      ,      ,      ,      ,      ,     ,      ,       ,     ,     ,      ,
  24,         free, intel,    8,     2,     6,  2.00,  49gb,   2gb, 0.02, 177gb, 128gb, 0.28,  8kb,   8kb, 248b
  25, job-exclusive, intel,    8,     8,     0,  1.61,  49gb,  19gb, 0.17, 419gb, 319gb, 0.24,  8kb,   8kb, 237b
  26, job-exclusive, intel,    8,     8,     0,  3.19,  49gb,   6gb, 0.05, 177gb, 155gb, 0.12,  8kb,   8kb, 256b
  27, job-exclusive, intel,    8,     8,     0,  8.00,  49gb,   5gb, 0.04, 419gb, 346gb, 0.17,  8kb,   8kb, 243b
  28, job-exclusive, intel,    8,     8,     0,  8.03,  49gb,   4gb, 0.04, 177gb, 121gb, 0.32,  8kb,   8kb, 241b
  29, job-exclusive, intel,    8,     8,     0,  8.00,  49gb,   4gb, 0.03, 177gb, 134gb, 0.24,  8kb,   8kb, 244b
  30, job-exclusive, intel,    8,     8,     0,  8.00,  41gb,   4gb, 0.04, 177gb, 133gb, 0.25,  8kb,   8kb, 244b
  31, job-exclusive, intel,    8,     8,     0,  8.00,  49gb,   5gb, 0.04, 419gb, 377gb, 0.10,  8kb,   8kb, 248b
  32,         free, intel,    8,     1,     7,  1.00,  49gb,   2gb, 0.01, 177gb, 135gb, 0.24,  8kb,   8kb, 253b
-------------------------------
number of CPUs total    : 416 ( 240 amd / 176 intel )
number of CPUs offline  :   8 (   0 amd /   8 intel )
number of CPUs online   : 408 ( 240 amd / 168 intel )
number of CPUs working  : 362 ( 222 amd / 140 intel )
number of CPUs free     :  46 (  18 amd /  28 intel )
```
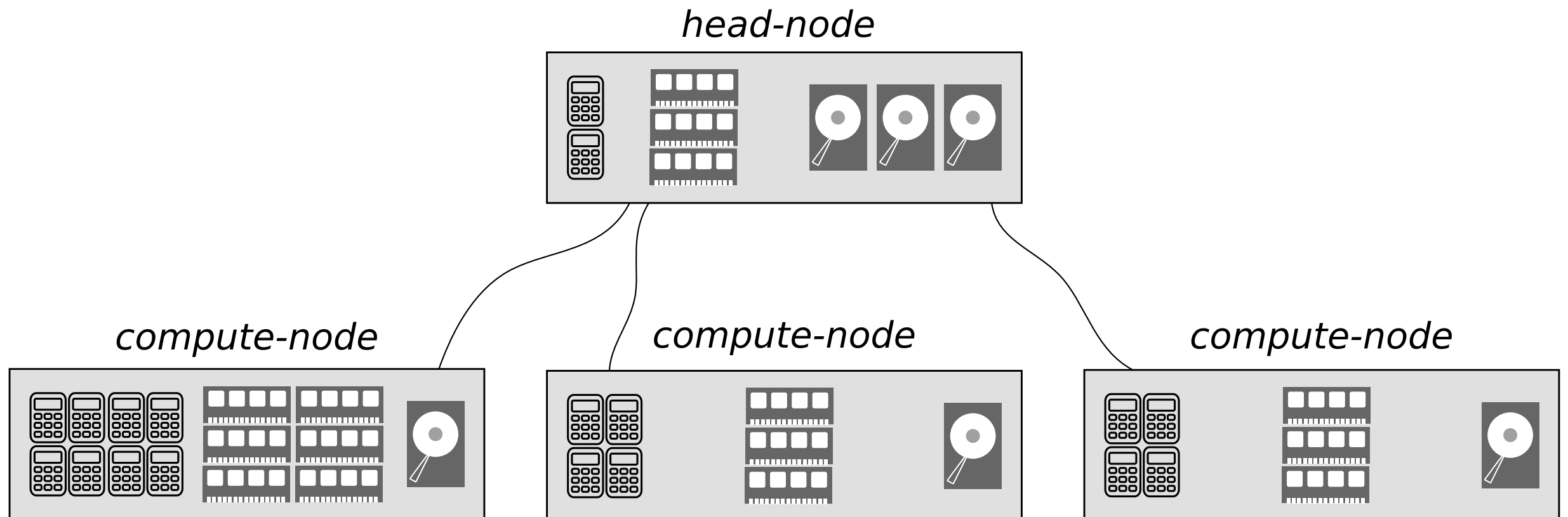
Etiquette: monitor resources

Monitor:

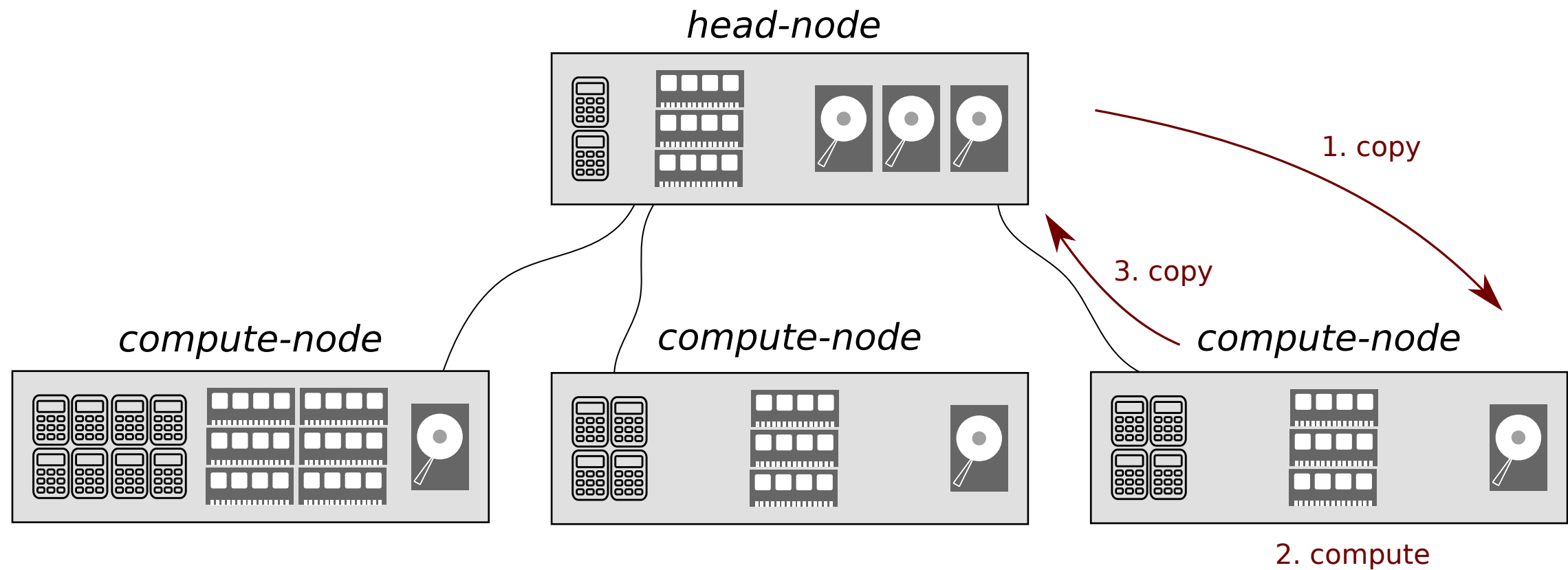- memory usage
- CPU usage
- disk usage

```
$ ssh compute-0-11

$ top
$ df -h
```

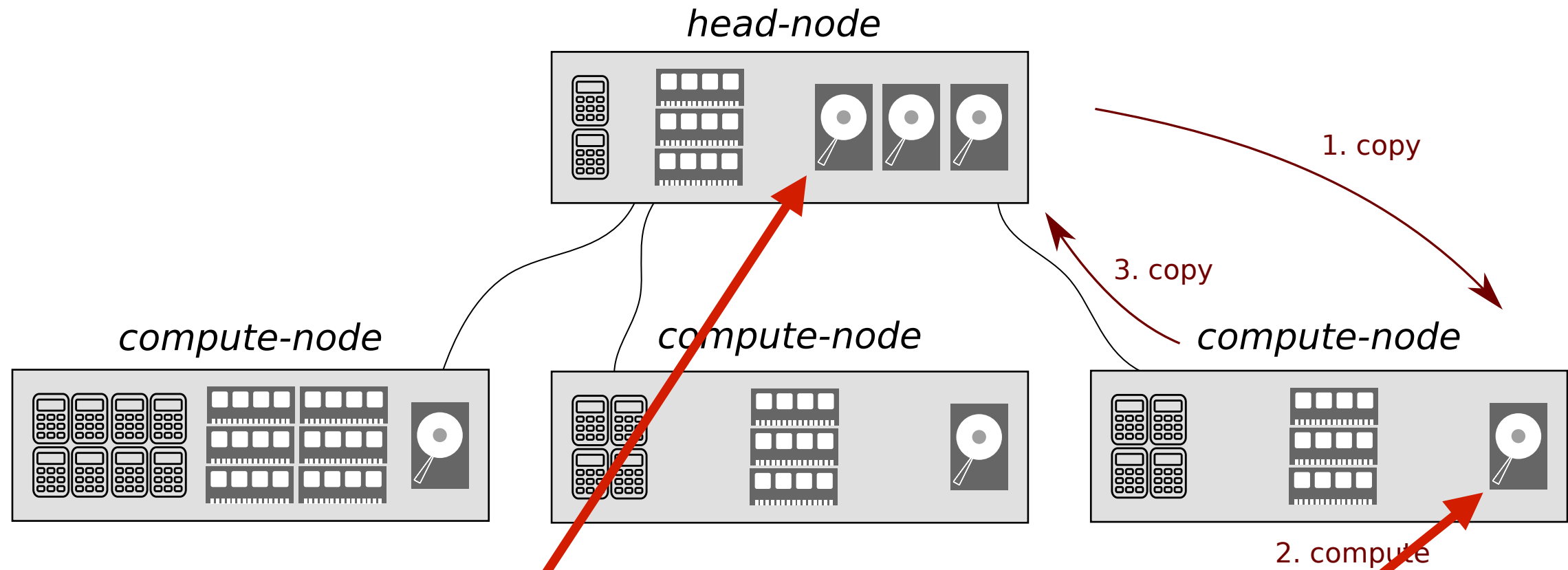www.furnace.wfw.wtb.tue.nl/ganglia

Etiquette: minimize network usage

Etiquette: minimize network usage

head-node

compute-node   *compute-node*   *compute-node*

1. copy

3. copy

2. compute

```
$ ssh compute-0-11

$ pwd
/home/tdegeus

$ cd /state/partition1/tdegeus
$ du -hs *
$ ls -lh
```

Stop a job

```
$ myqstat

jid      , owner     , job name      , host , cpus    , mem   , S , time , score
------   , --------- , ------------- , ---- , ------  , ----  , - , ---- , -----
202169 , tdegeus   , myjob.pbs     ,    11 ,  1:1:i , 1mb   , R ,  10s ,  1.00
...

$ qdel 202169
```

Monitor files

```
$ rocks run host "ls /state/partition1/`whoami`"
compute-0-30:
compute-0-19:
compute-0-31:
compute-0-15:
compute-0-27:
compute-0-17:
compute-0-32:
compute-0-26:
compute-0-16:
compute-0-13:
compute-0-21:
compute-0-18:
compute-0-20:
compute-0-28:
compute-0-29:
compute-0-3:
compute-0-5:
compute-0-24:
compute-0-25:
compute-0-6:
compute-0-10:
compute-0-1:
compute-0-11:
compute-0-0:
compute-0-14:
compute-0-22:
compute-0-8:
compute-0-9:
compute-0-7:
compute-0-2:
compute-0-12:
compute-0-23: down
```

head-node

1. copy

3. copy

compute-node   *compute-node*   *compute-node*

2. compute
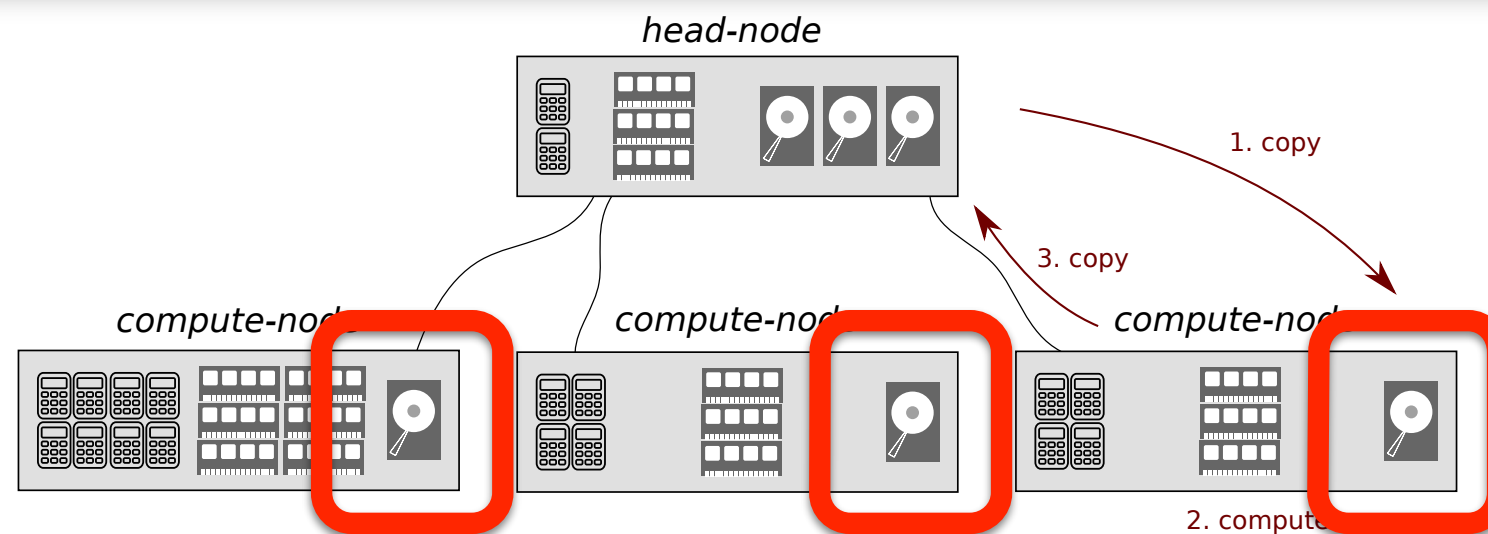
Etiquette

- **Monitor your jobs and resources**:
  - CPU usage ("score")
  - Memory usage
  - Disk usage
  - Network usage
- When in doubt: **ask for help**
- If you plan on running a large simulation: **discuss with Leo Wouters first**
- **Claim what you use and use what you claim**
- Don't claim too much at once

Help others, share your experience

Collaboration

Full Name
E-mail
username @furnace
Software

Bart Vossen
B.G.Vossen@tue.nl
bart
- Linux/Bash
- Matlab
- MSC.Marc
  + subroutines

Jeroen van Beeck
J.v.Beeck@tue.nl
jeroen
- Linux/Bash
- Python
- Matlab
- MSC.Marc
  + subroutines

Tom de Geus
T.W.J.d.Geus@tue.nl
tdegeus
- Linux/Bash
- Python
- C/C++
- MSC.Marc

Future

- …
- …
- …

Getting help

www.mate.tue.nl/~cluster   `$ myqstat —help`

- Cluster manual
- This presentation
- …

MaTeCluster@tue.nl

- Bart Vossen (B.G.Vossen@tue.nl)
- Jeroen van Beeck (J.v.Beeck@tue.nl)
- Tom de Geus (T.W.J.d.Geus@tue.nl)

L.H.G.Wouters@tue.nl

- Leo Wouters
- Obtain account