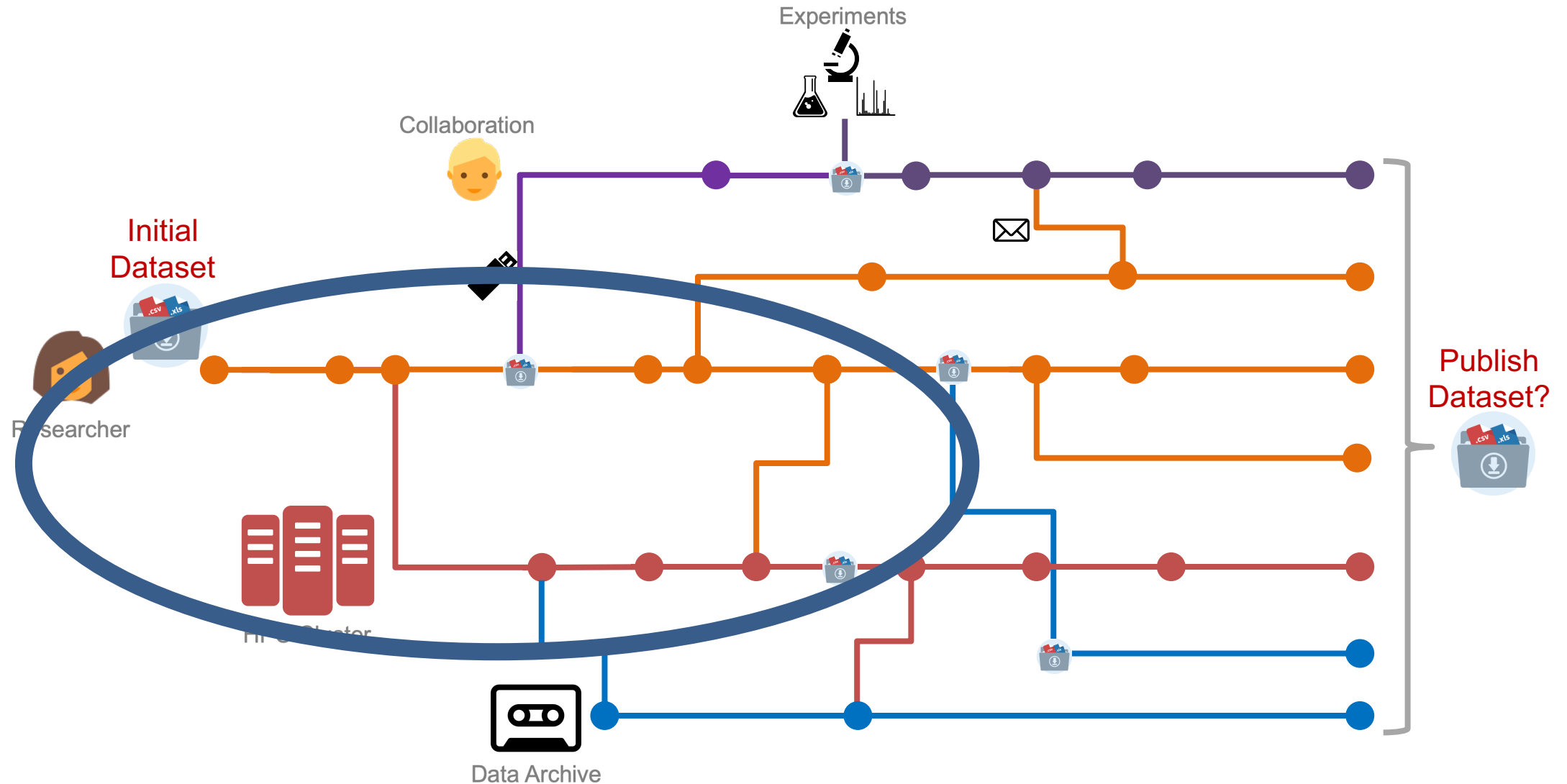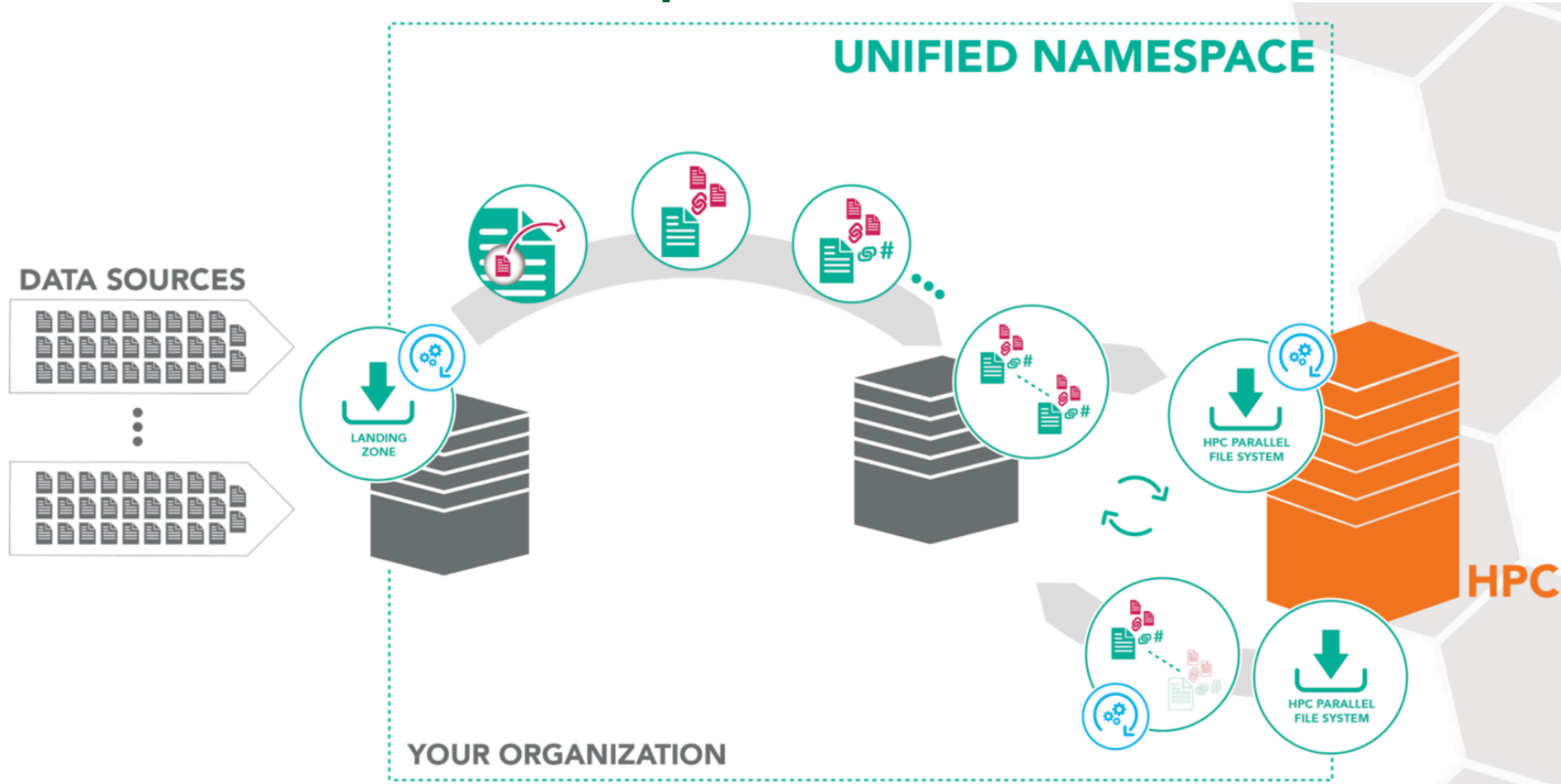# Compute cluster Lisa

Data, what's the problem?
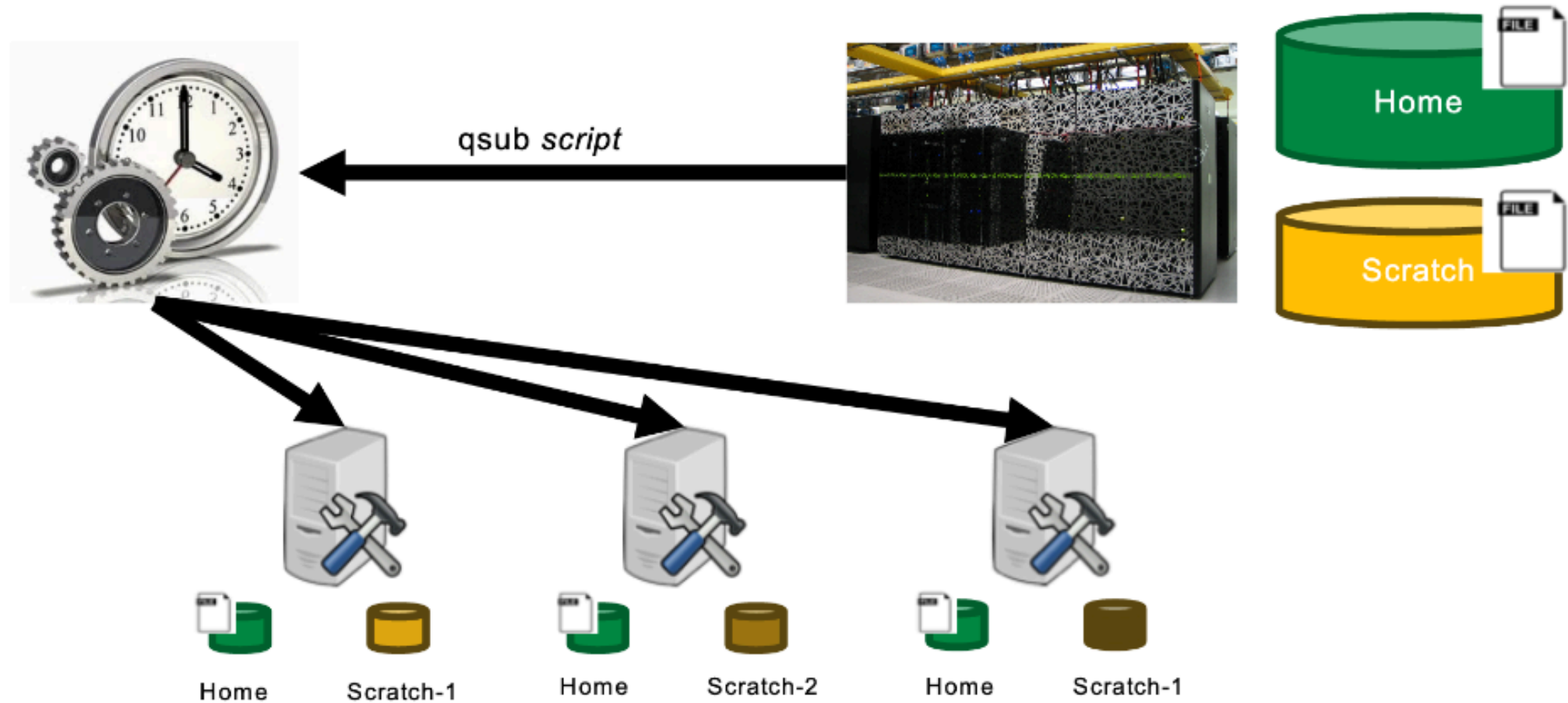
iRODS computation workflows

What, why, when Supercomputer?

- What is a super computer?

  - A fast computer

  - A large computer (memory/storage)

  - An expensive computer

- Why/when do you need a supercomputer?

  - If your task would take months/years on a normal PC

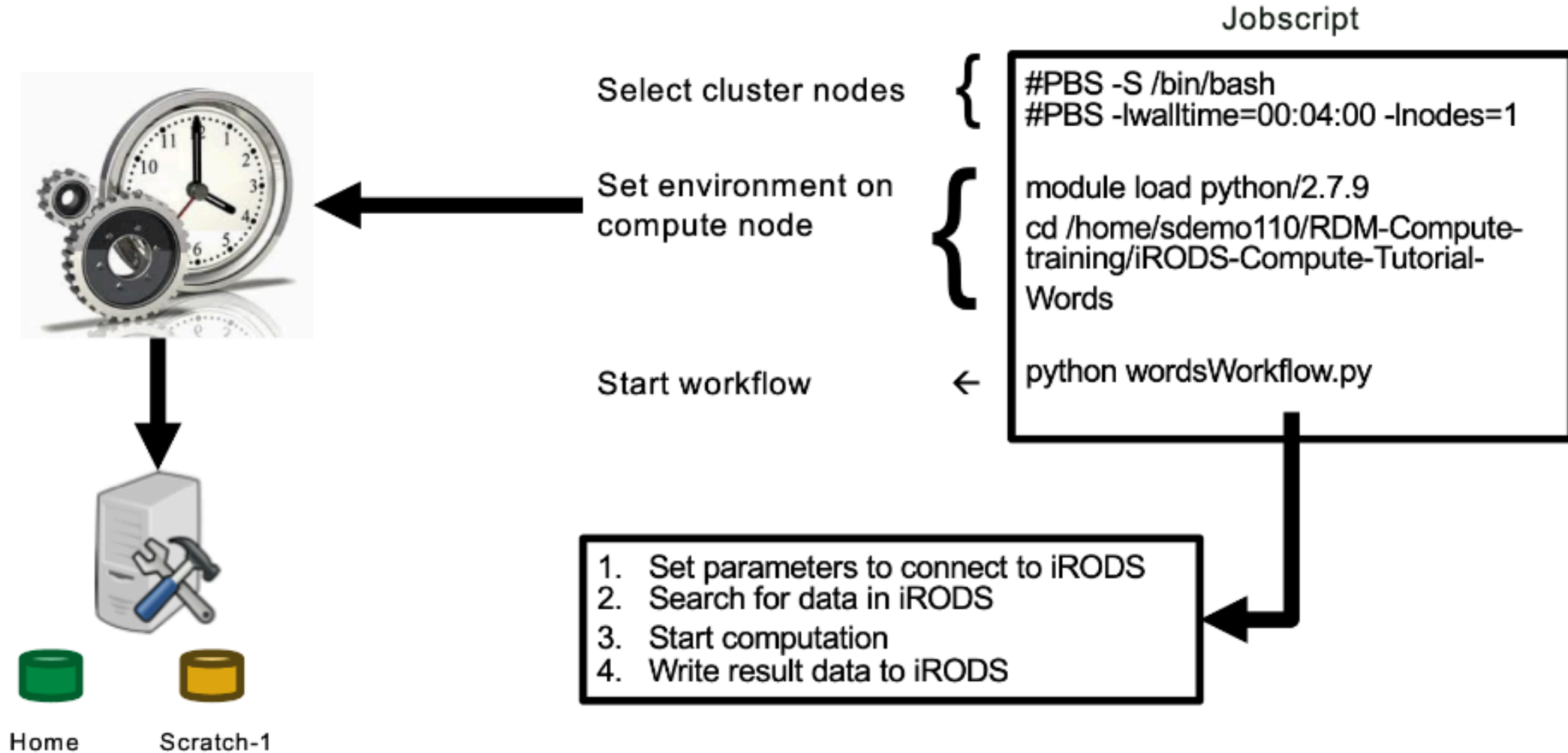 - If your task requires more space (memory/storage) than available on your PC

Compute clusters SURFsara

- SURFsara has Cartesius and Lisa computer clusters
- What is the difference?

  - Cartesius – Larger blocks (capability – fewer large scale jobs)
  - Lisa -- smaller blocks (capacity computing – more smaller scale jobs)

  - Cartesius – expensive expensive
  - Lisa – Cheaper, but still expensive

Lisa

qsub *script*

Workflow 1 - Wordcount

Jobscript

Select cluster nodes
```
#PBS -S /bin/bash
#PBS -lwalltime=00:04:00 -lnodes=1
```

Set environment on compute node
```
module load python/2.7.9
cd /home/sdemo110/RDM-Compute-
training/iRODS-Compute-Tutorial-
Words
```

Start workflow
```
python wordsWorkflow.py
```

1. Set parameters to connect to iRODS
2. Search for data in iRODS
3. Start computation
4. Write result data to iRODS

Home   Scratch-1

Optional Workflow 2- ACES

Jobscript

Select cluster nodes {
```
#PBS -S /bin/bash
#PBS -lwalltime=00:04:00 -lnodes=1

module load python/2.7.9
```

Set environment on compute node {
```
cd /home/sdemo110/RDM-Compute-
training/iRODS-Compute-Tutorial-Words
```

Set iRODS parameters
Download dataset from FigShare
10.6084/m9.figshare.3119248

Start workflow ← `python acesWorkflow.py`

1. Search for data in iRODS
2. Start computation
3. Write result data to iRODS

Home   Scratch-1

SURF SARA