

# SURFsara e-infrastructures

Large scale computing & Big Data

Swammerdam Institute for Life Sciences Workshop, 17 Oct 2016

Natalie Danezi <[anatoli.danezi@surfsara.nl](mailto:anatoli.danezi@surfsara.nl)>

 software carpentry



**SURF** SARA

UNIVERSITEIT VAN AMSTERDAM

# SURF family

- 1971: SARA  
Stichting Academisch Rekencentrum Amsterdam, founded by UvA, VU, CWI
- 1984: National supercomputing center
- 2013: SURFsara



eScience Collaboration and Tools

National Research & Education Network

Commercial ICT Products & Services

Scientific Computing & Storage

An iceberg floating in the ocean. The tip of the iceberg is visible above the water surface, while the much larger, submerged part is visible below. The sky is blue with light clouds, and the water is a deep blue.

# High Performance Computing (HPC) in research

# Does my research fit in HPC?

- **Faster results**
- **Task repetition**
- **Higher accuracy**
- **Larger computational domains**
- **Larger volume of data**

<http://www.advancedgwt.com/groundwater-software/data-management-and-visualization/groundwater-desktop.html>

# The HPC Infrastructure

# More users, more resources

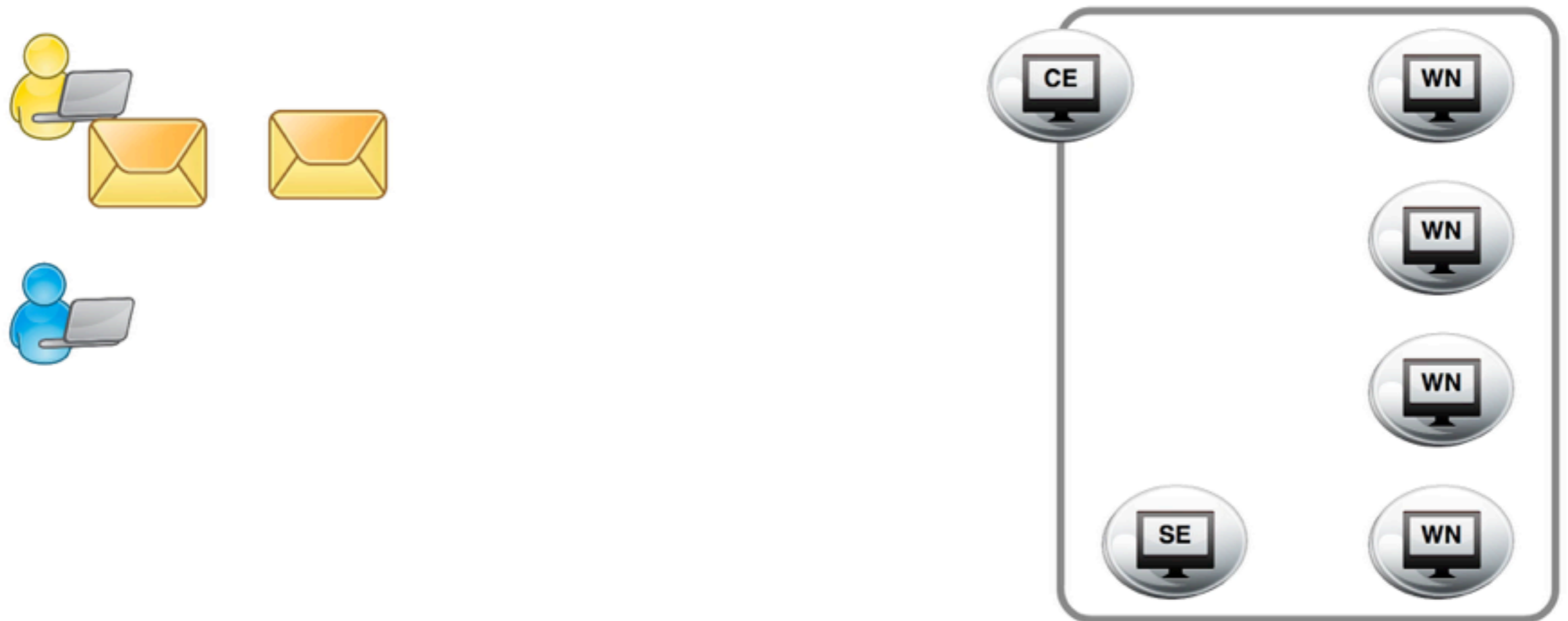
[http://web.grid.sara.nl/mooc/animations/single\\_user.html](http://web.grid.sara.nl/mooc/animations/single_user.html)



*Animation: EGI - SURFsara MOOC 2014*

# Lisa cluster: batch processing

<http://web.grid.sara.nl/mooc/animations/cluster.html>



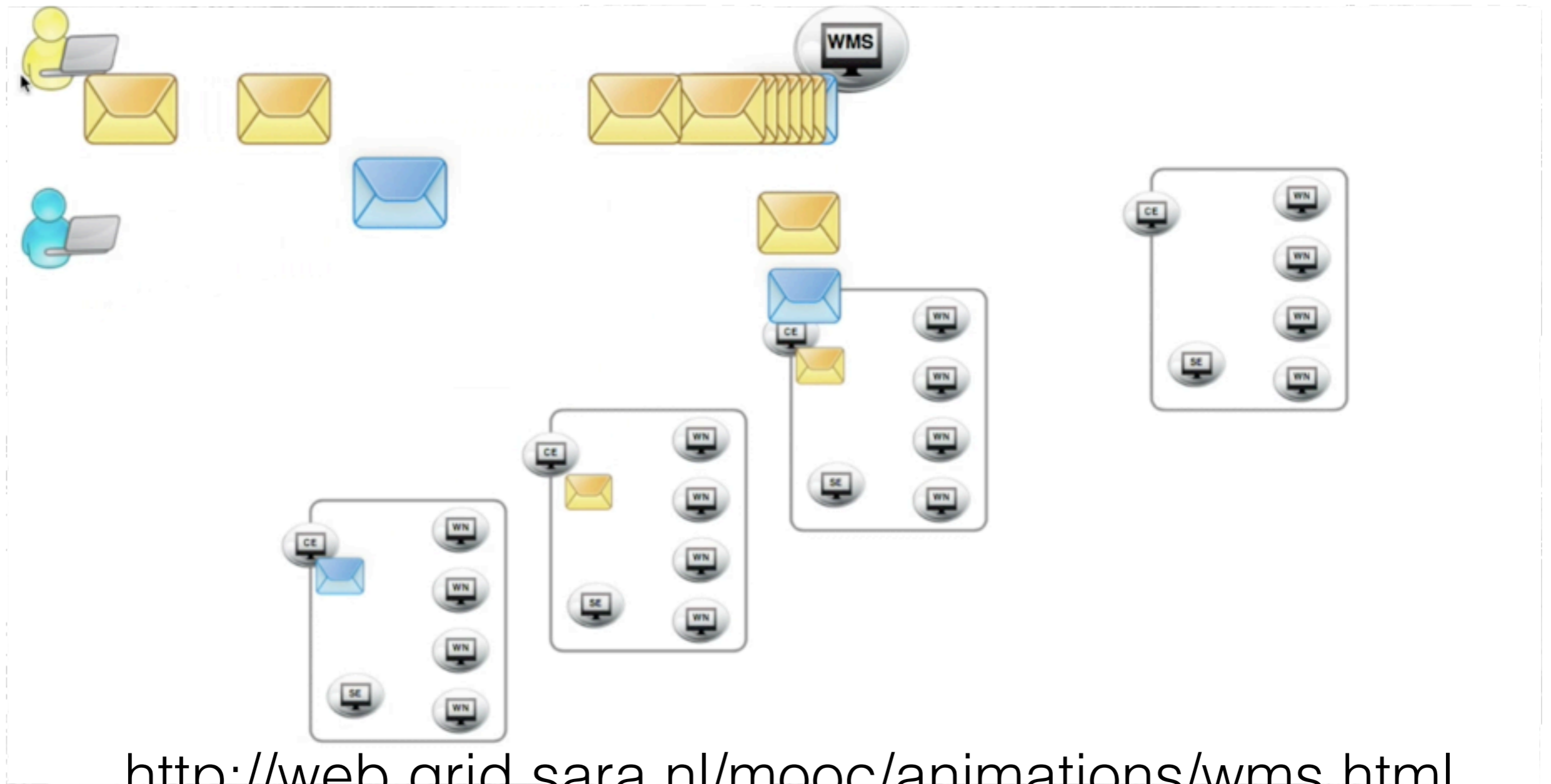
*Animation: EGI - SURFsara MOOC 2014*

# Lisa cluster: batch processing

- Calculation intensive applications
- Not data intensive applications
- Well supported software stack
- Relatively easy to start



# Life Science Grid: cluster of clusters



<http://web.grid.sara.nl/mooc/animations/wms.html>

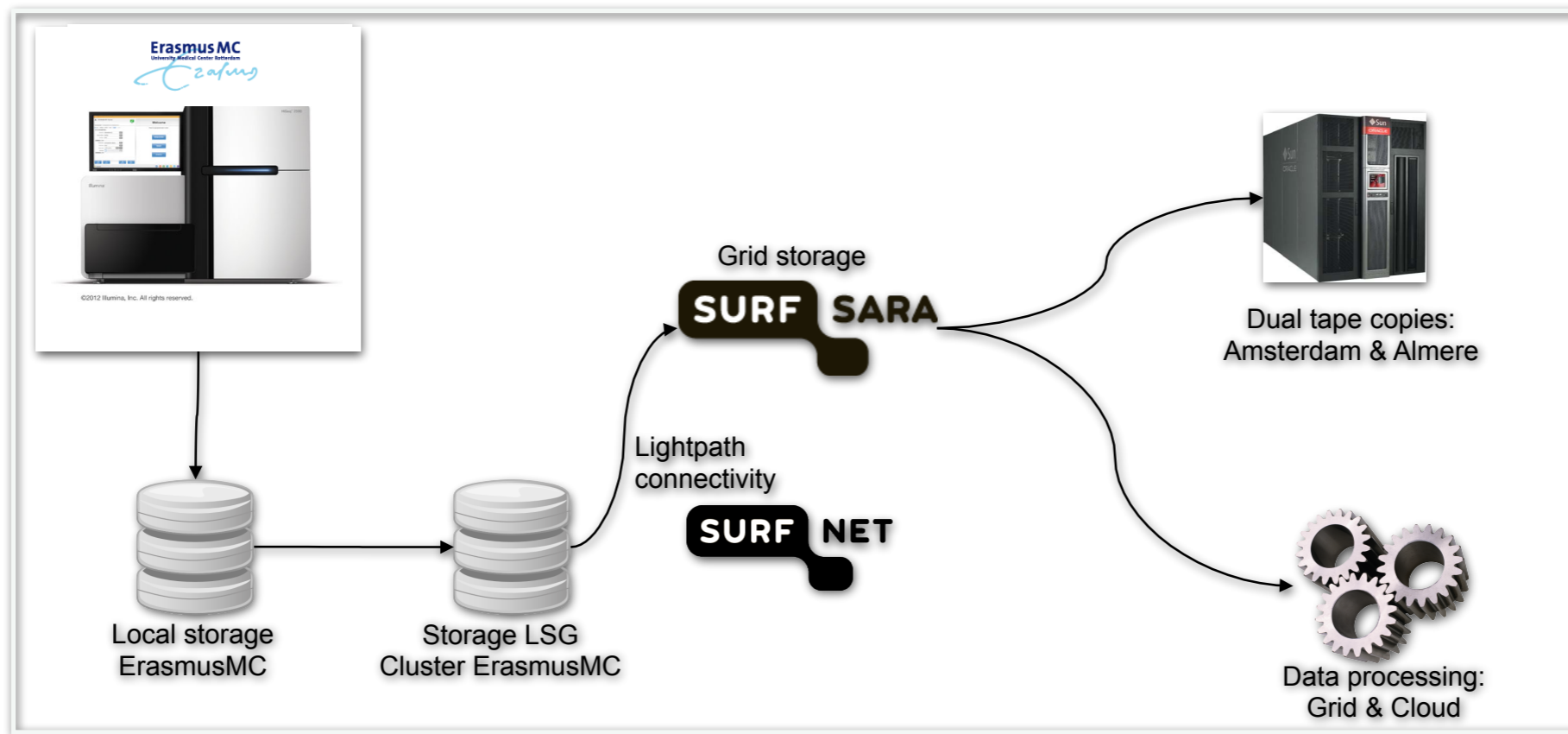
Animation: EGI - SURFsara MOOC 2014

# Life Science Grid: cluster of clusters

- Resources meant for life science researchers
- 11 local clusters (AMC, LUMC, WUR, TUD, RUG, ..)
- Capacity: +/- 12 000 cpu cores, peta bytes of storage
- Independent tasks: Parameter sweeps, Monte-Carlo, ..
- Linux experience required



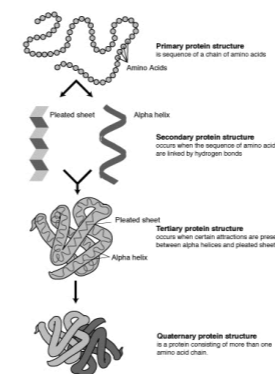
# Life Science Grid examples



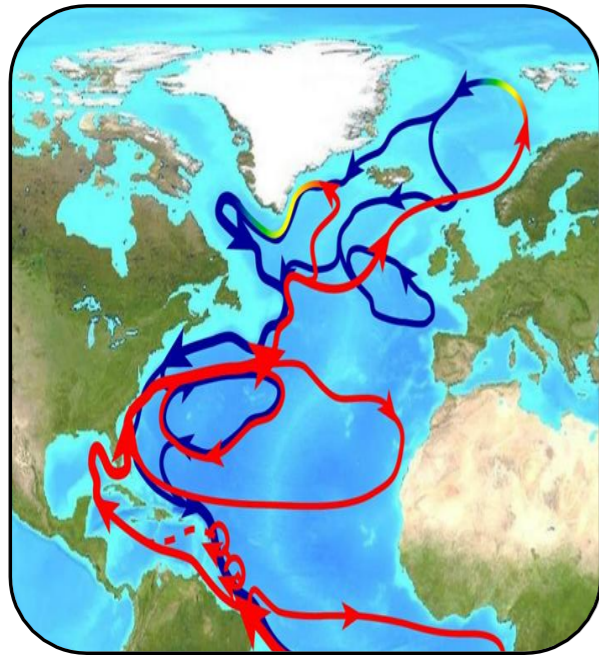
GoNL



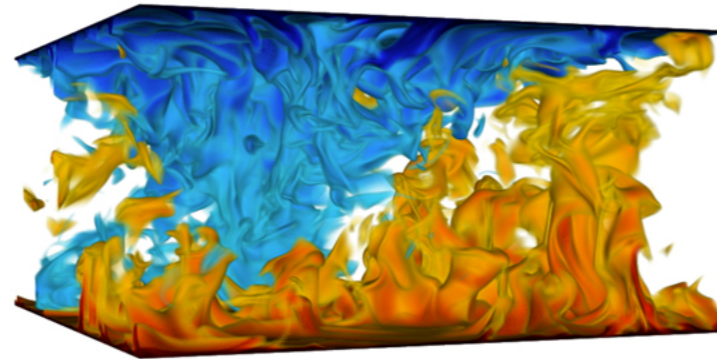
BBMRI.nl



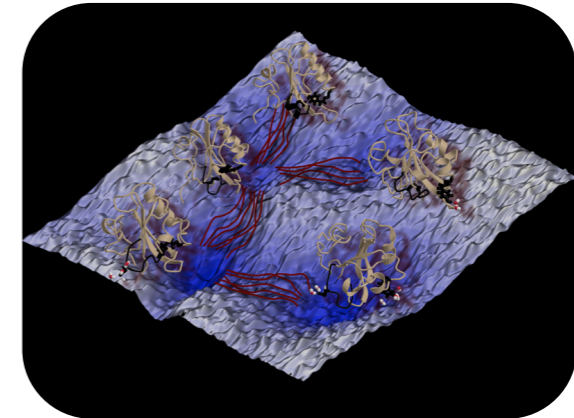
# Cartesius Supercomputer



eSALSA



Turbulence  
modelling



Protein  
structure

- Large memory, fast interconnect, fast and large I/O
- European collaborations: PRACE, HPC-Europa
- Climate models, cell simulations
- Programming experience required

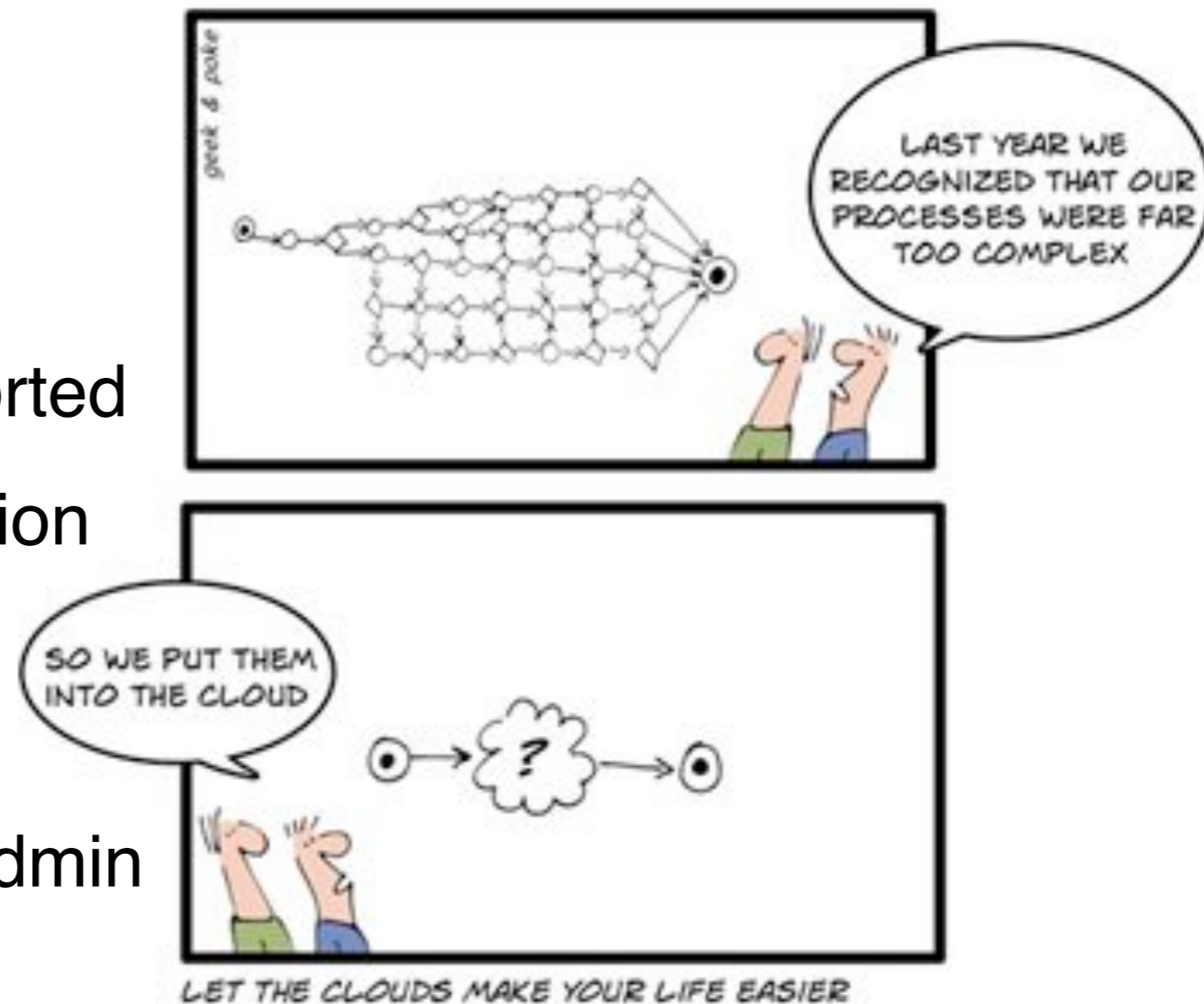
# Hadoop

## How to index the web

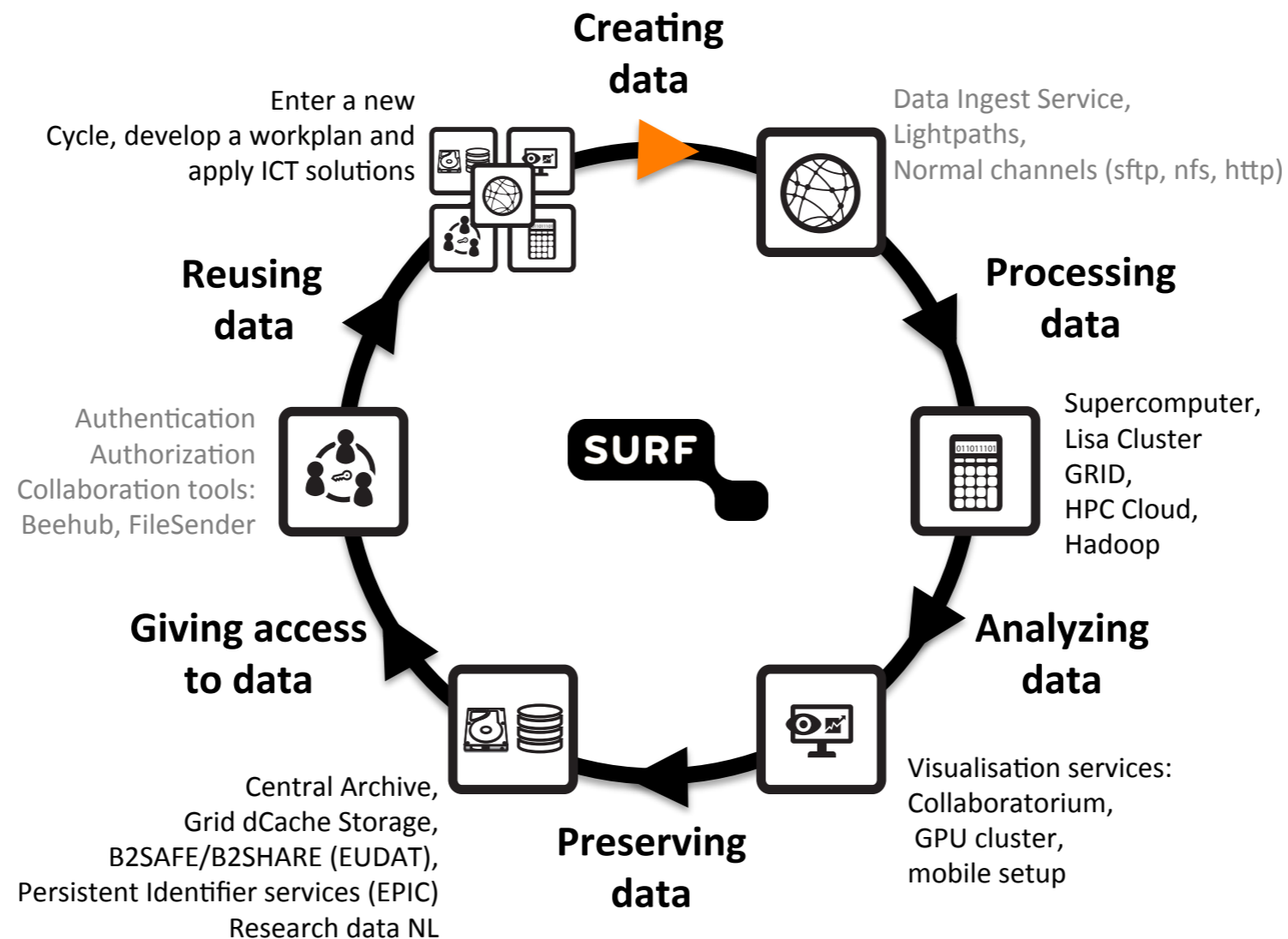
- 
- **Big Data**
  - **Exploration/mining of data**
  - **Map / Reduce**
  - **Programming experience required**

# HPC Cloud

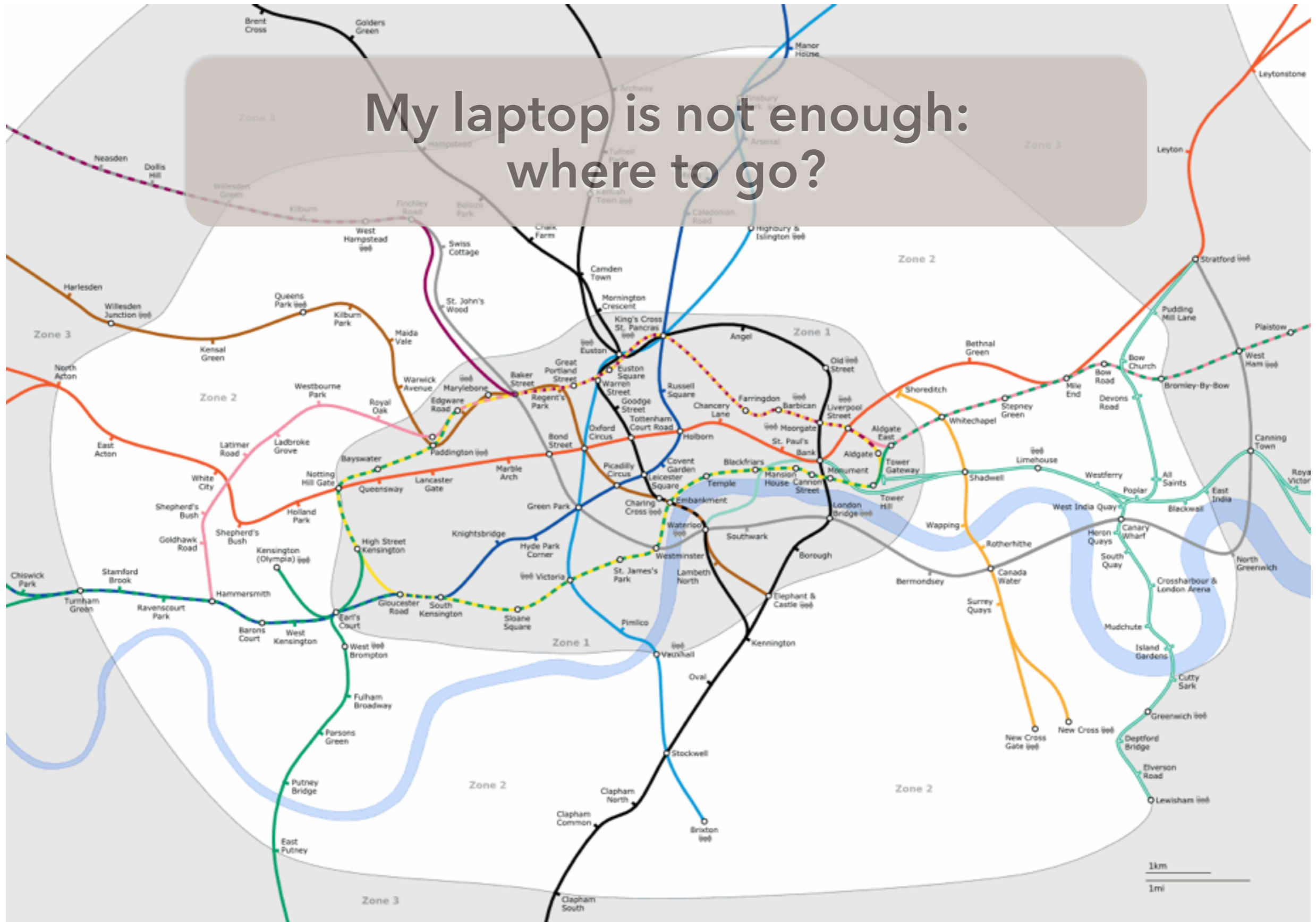
- Flexible & controllable
- Microsoft Windows supported
- Ideal for 3<sup>rd</sup> party application providers
- Graphical user interface
- User is also the system admin



# Data services



# My laptop is not enough: where to go?





# SURFsara services

- **Lisa** national compute cluster
- **Life Science Grid**: interconnected clusters across the Netherlands
- **Cartesius** national Supercomputer
- **Hanthi** Hadoop cluster
- **Oort** HPC Cloud cluster
- **Central Archive, Beehub, SURFdrive** for Data Services
- ..., Visualisation, Networking, Consultancy, Innovation

# Getting access



- For:
  - **Lisa National Cluster, Cartesius Supercomputer**
- Apply via:
  - **IRIS**

(NWO grant)

## SURF SARA Resource request form

Personal Resources Project

Applicant / Contact person

The principal researcher must have a permanent position at the organisation.

First name:

Last name:

Initials:

Gender:

Date of birth:

Title(s):

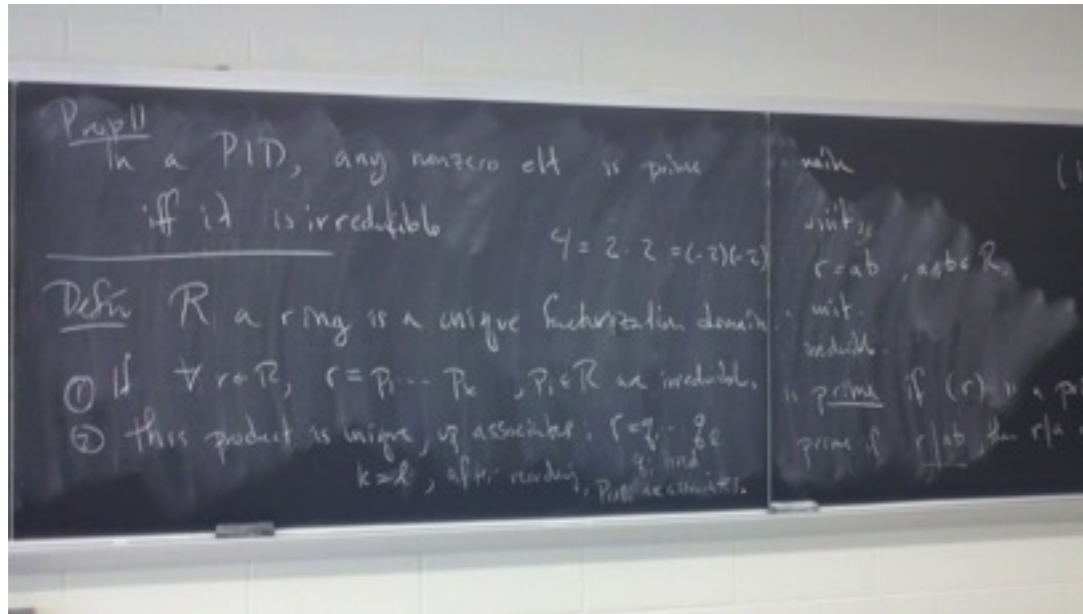
Function:

Institute:

- For:
  - **Grid, Hadoop, HPC Cloud, Data Services, Visualisation**
  - **Or, not sure what suits you best?**
- Apply via:
  - **<https://e-infra.surfsara.nl/>**

(SURF grant)

# Standard support



Bring your scientific problem

- We provide advice and support:
  - Getting access
  - Best practices
  - Design & optimisation
  - Integration to large scale



# Trainings, online tutorials

## SURFsara and UNESCO-IHE workshop: How water research benefits from HPC Cloud



UNESCO-IHE management delivered a hands-on course on their HPC Cloud and presented its benefits to the water sector.



# Questions?



<https://www.surf.nl/en/about-surf/subsidiaries/surfsara/>



[helpdesk@surfsara.nl](mailto:helpdesk@surfsara.nl)



020 800 1400