

Calibration of LOFAR data on the cloud

Jose Sabater Montes (jsm at roe dot ac dot uk)
Institute for Astronomy, University of Edinburgh

with S. Sanchez, J. Garrido, J. E. Ruiz, P. Best, L.
Verdes-Montenegro and the LOFAR collaboration

Challenges

- User data calibration
 - Example:
 - ◇ 10 hours full resolution → ~20 TB
 - ◇ 2 CPU years to run the calibration
 - Experimental pipeline
- LOFAR calibration software
 - Difficult to install (this is improving quickly)
 - Continuous development

Possible solution

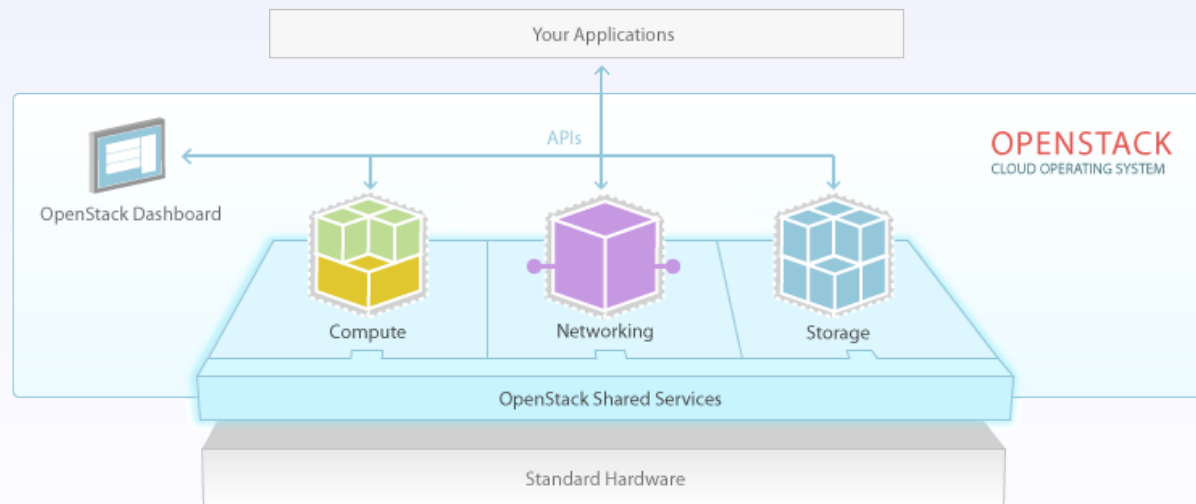
- Parallelizable:
 - Deal with a large amount of data in a reasonable time.
- Flexible:
 - Adapt the infrastructure (“hardware”) to different calibration strategies
 - Deal with quickly changing software

Cloud computing

- Infrastructure as a Service (IaaS)

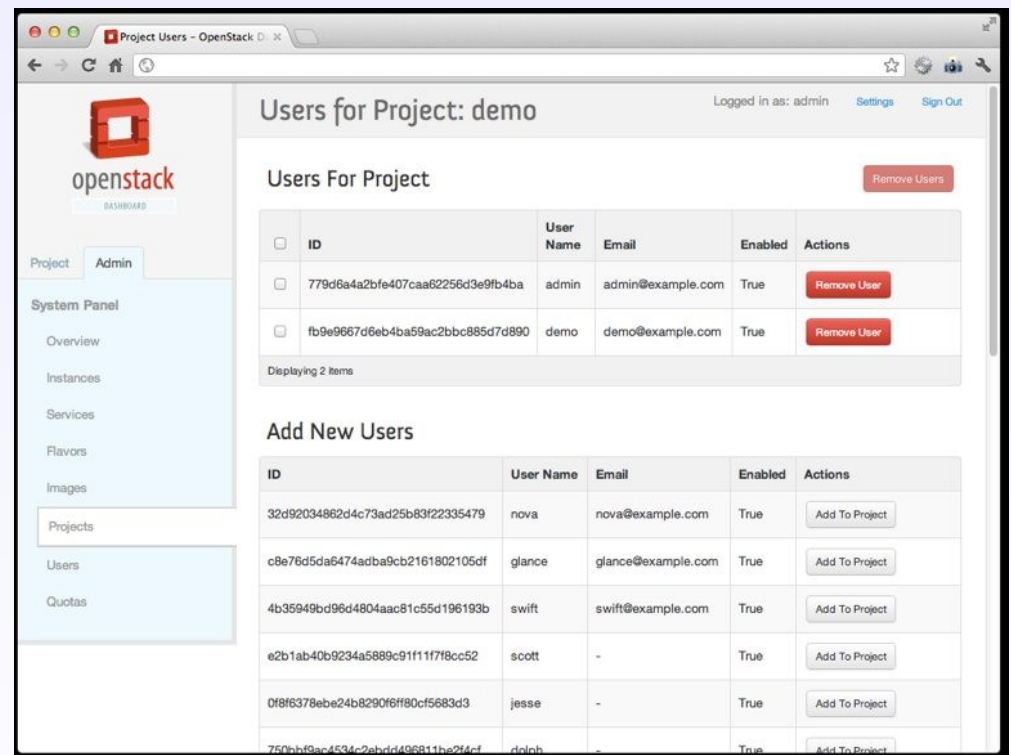
- Examples:

- Amazon Web Services (EC2, S3, etc), Google Compute Engine, RackSpace...
- Eucalyptus, OpenStack...



Tests on IBERCLOUD

- Cloud infrastructure provided by Ibergrid
- Implemented with OpenStack

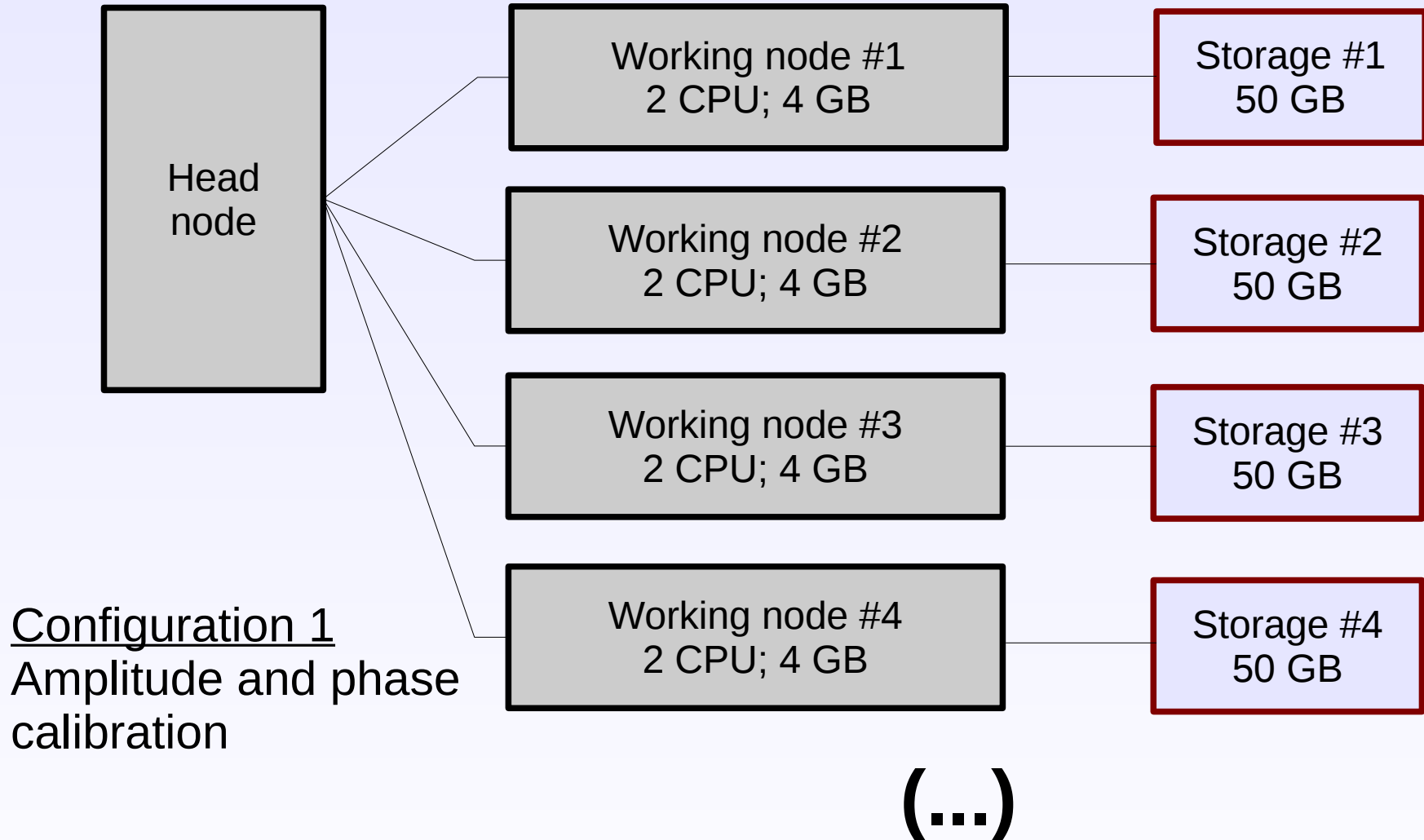


The screenshot shows the OpenStack Project Users interface. The page title is "Project Users - OpenStack D...". The user is logged in as "admin". The main content area is titled "Users for Project: demo" and contains a table of users for the project. The table has columns for ID, User Name, Email, Enabled, and Actions. There are two users listed: "admin" and "demo". Below the table, there is a section for "Add New Users" with a table of users that can be added to the project. The table has columns for ID, User Name, Email, Enabled, and Actions. There are six users listed: "nova", "glance", "swift", "scott", "jesse", and "dolph".

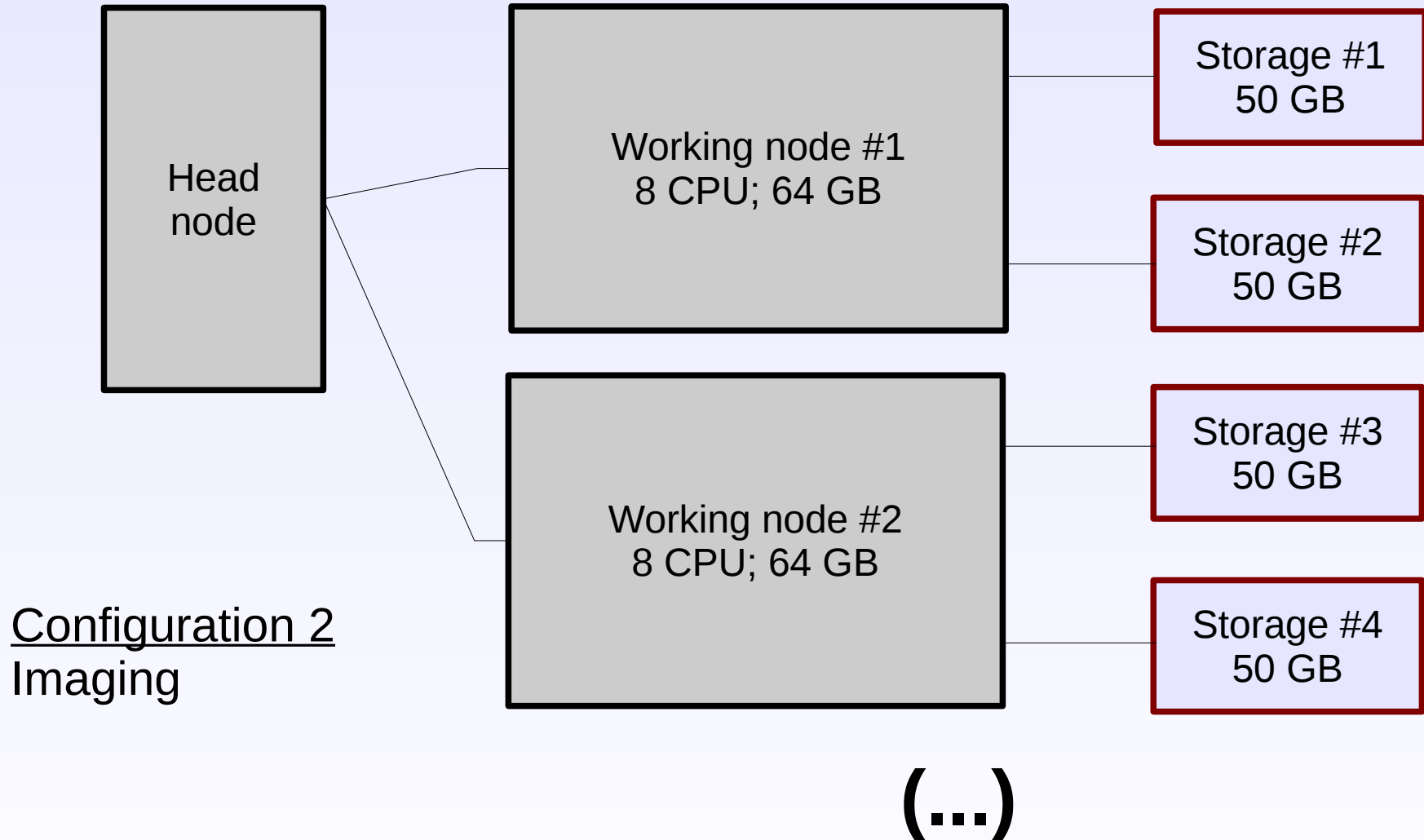
ID	User Name	Email	Enabled	Actions
779d8a4a2bf407caa62256d3e9fb4ba	admin	admin@example.com	True	Remove User
fb9e9667d6eb4ba59ac2bbc885d7d890	demo	demo@example.com	True	Remove User

ID	User Name	Email	Enabled	Actions
32d92034862d4c73ad25b83f22335479	nova	nova@example.com	True	Add To Project
c8e76d5da8474adba9cb2161802105df	glance	glance@example.com	True	Add To Project
4b35949bd96d4804aac81c55d196193b	swift	swift@example.com	True	Add To Project
e2b1ab40b9234a5889c91f117f8cc52	scott	-	True	Add To Project
0f8f6378ebe24b8290f6ff80cf5683d3	jesse	-	True	Add To Project
750bbf9ac4534e2abd4468911be2f4cf	dolph	-	True	Add To Project

LOFAR pilot in Ibercloud



LOFAR pilot in Ibercloud



Current status

- Current status:
 - LOFAR instance images created
 - Virtual on-demand cluster working
 - IPython parallel used to orchestrate the calibration on the nodes
- Next steps:
 - Additional testing
 - Use in production
 - Adapt to the new European Grid Infrastructure (EGI) Federated Cloud service.

Conclusions

- Cloud infrastructure to calibrate LOFAR data:
 - Elastic on-demand resource consumption
 - Parallelization - Ability to deal with big data
 - Flexibility - Quick development of innovative strategies