

#sosc2022

#sosc22

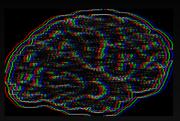
#sosc2022

#### 4th International School on Open Science Cloud

## Using S3 storage Fundamentals

**INFN Perugia - Diego Ciangottini** 

## Few notes before starting



Tools overview

- You will use during all the hands-on sessions a remote python environment
- The interface is a de facto standard nowadays: JupyterLab
  - It is a common solution to offer the user an interactive experience while working on remote resource
- Let's take a look briefly to the main features together now
- https://sosc.131.154.96.42.myip.cloud.infn.it/hub



# JupyterLab tour...

And btw, what you'll be using under the hood is this docker image: <u>ghcr.io/sosc-school/sosc22-jupyterlab:latest</u> created via:

https://github.com/SOSC-School/SOSC22-livesessions/blob/main/jupyterlab/Dockerfile

You can try if you feel nerd enough, to follow the today tutorial locally via docker.

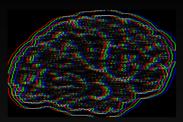
**DO NOT USE LOCALLY RUNNING CONTAINER FOR TUE AND THU WORKFLOW HANDS ON** 

They are using a cluster of machine setup, that you won't be able to access from your machine.

# Let's give a stab to S3 storage now



## **Objects come in buckets**



#### How can I explore data

- Buckets are the building blocks of a S3 object store
- They are "standalone" collection of objects that can have different set of permissions and quota
- Every object inside a bucket is identified by a key (also called "name")
- So in order to access a file in an object storage you'll need:
  - The server URL
  - The bucket name
  - The object key

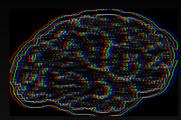
# S3 storage web interface



#### Main features

- <u>https://console.131.154.96.42.myip.cloud.infn.it/</u>
- Now you need to download the hands-on notebook from the school GitHub repository
  - <u>https://github.com/SOSC-School/SOSC22-livesessions</u>
- From there you can get your credentials and login
- Let's play a bit with this interface now

## **Uploading my "object"**



#### a.k.a. store files on my storage

- Choose your bucket
  - In this case your username
- Locate/create a local data to upload
- Go ahead with either webUI or any S3 client

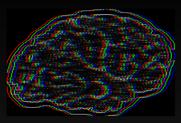
Rewind 🔊	Refresh 🖒	Upload 🟦
	Crea	ate new path ://

#### Upload you first file

[6]: data = open('src.C', 'rb')
s3.Bucket(username).put\_object(Key='test.jpg', Body=data)

[6]: s3.Object(bucket\_name='dciangot', key='test.jpg')

#### How can read my data?



#### Most common patterns

- You are now good to go
- You can share your file via a temporary pre-signed URL
- Also downloading it via the CLI or on of the many existing S3 python clients

C Share File
This is a temporary URL with integrated access credentials for sharing objects valid for up to 7 days.
The temporary URL expires after the configured time limit.
Active for
7Days0Hours0Minutes
Link will be available until: 12/04/2022 21:44:16 https://minio.131.154.96.42.myip.cloud.infn.it/dciangot/test.jpg?X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Am

Download a file ¶

[12]: with open('test-download.jpg', 'wb') as f: s3.download\_fileobj(username, 'workflow-with-template-ref/workflow-with-template-ref-t-512854155/main.log', f) Time to take a look...

