

High Performance and Distributed Computing for Big Data

Unit 3: AWS - Deploying a Jupyter Notebook in the Cloud

Jordi Mateo Fornés jordi.mateo@udl.cat

Universitat Rovira i Virgili and Universitat de Lleida

Today's session

Today's session

1. Practical session: Deploying in the Cloud
 - Apply learned knowledge to deploy an application in the Cloud.
2. Local vs Cloud deployment
 - Compare the difference between local and Cloud deployment.
3. AWS Services
 - Learn about AWS's different services and how to use them.

Today's session

1. Practical session: Deploying in the Cloud
 - Apply learned knowledge to deploy an application in the Cloud.
2. Local vs Cloud deployment
 - Compare the difference between local and Cloud deployment.
3. AWS Services
 - Learn about AWS's different services and how to use them.

Just for the record:

- I am not promoting AWS but using it due to its popularity and as an AWS educator.
- AWS is utilized for practicality and as an example.
- Adapt the knowledge to the cloud provider relevant to your professional context.

Requirements

All the students should have an AWS Educate account. Check your email for the invitation and follow the instructions to create your account.



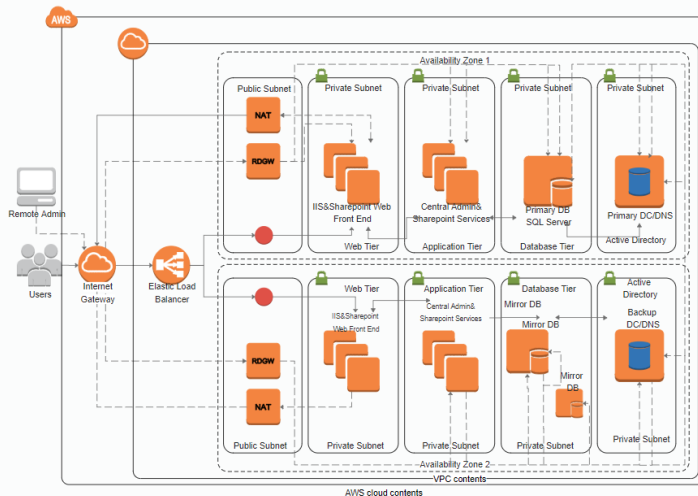
Before we start, all the students are invited to complete the AWS Cloud Foundations course. This course is available in the AWS Educate platform and it is a great introduction to the AWS services. The course is not mandatory, but it is highly recommended.

This introductory course is intended for students who seek an overall understanding of cloud computing concepts, independent of specific technical roles. It provides a detailed overview of cloud concepts, AWS core services, security, architecture, pricing, and support.

This course helps you to prepare for the AWS Certified Cloud Practitioner exam. The exam is not mandatory, but it is a great way to validate your knowledge and to have a certification that can be useful for your professional career.

Introduction to AWS

What is AWS?



Which are the main services?

AWS Services

Deployment & Management

Application Services



Amazon SQS



Amazon ElasticTranscoder



Amazon SES



Amazon AppStream



Amazon CloudSearch

Mobile Services



Amazon Cognito



Amazon Mobile Analytics



Amazon SNS

Enterprise Applications



Amazon WorkDocs



Amazon WorkSpaces



Amazon WorkMail

Application Services

Administration & Security



AWS DirectoryService



AWS IAM



AWS Trusted Advisor



AWS Config



AWS CloudTrail



Amazon CloudWatch

Deployment & Management



Amazon CloudFormation



AWS OpsWorks



AWS CodeDeploy

Analytics



Amazon Kinesis



AWS Data Pipeline



Amazon EMR

Foundation Services

Compute



Amazon EC2



AWS Lambda

Storage & Content Delivery



Amazon CloudFront



Amazon Glacier



AWS Storage Gateway



Amazon Content Delivery

Database



Amazon DynamoDB



Amazon RDS



Amazon Redshift



Amazon Elastic Cache

Networking



Amazon Route 53



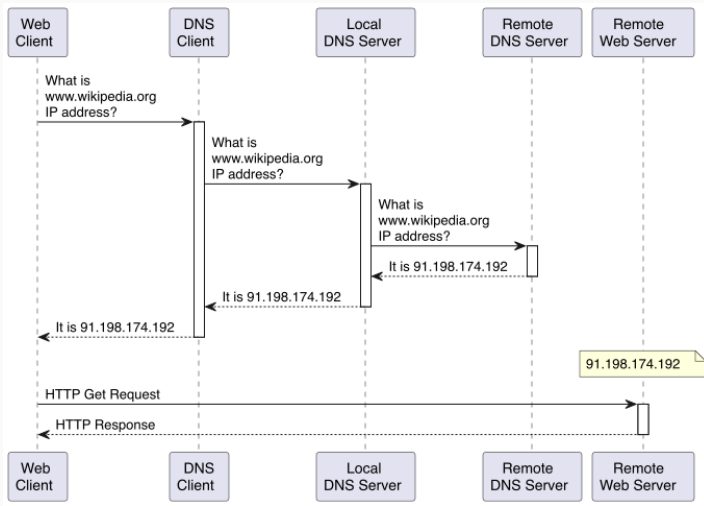
Amazon VPC



AWS Direct Connect

Using the Cloud

Client-Server Model in the Cloud



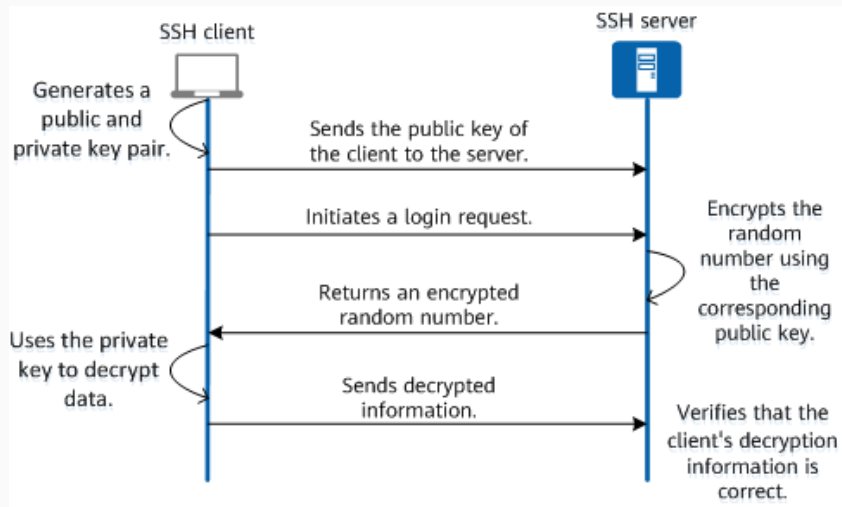
Client-Server Model

The client-server model is a distributed application structure dividing tasks between servers (providers of resources/services) and clients (service requesters).

In the cloud

In the cloud, the client is the user, and the server is the cloud provider. For example, in AWS, the client can be a user using the AWS Management Console, and the server is the AWS infrastructure.

SSH Protocol: Secure Cloud Communication



Creating a SSH Key Pair

Open a terminal or a powershell and type the following command:

```
ssh-keygen -t type
```

The `-t` option specifies the type of key to create:

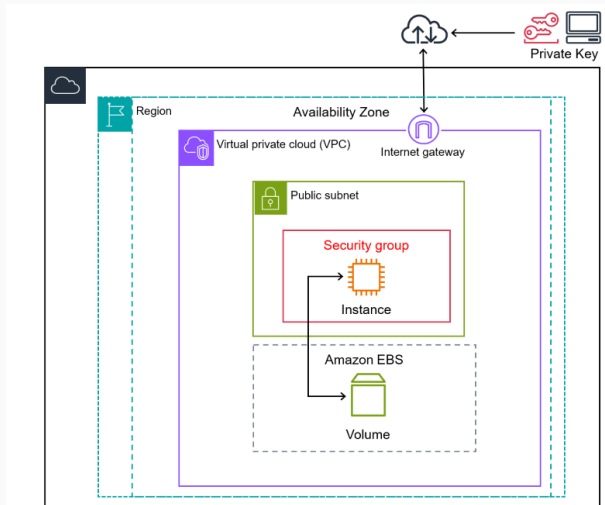
- `rsa`
- `dsa`
- `ecdsa`
- `ed25519` (recommended)

The command will prompt you to enter a file in which to save the key. **Press Enter** to save the key in the default location. The command will also prompt you to enter a passphrase. You can enter a passphrase or leave the passphrase empty. This command will create a public and a private key in the default location.

EC2 - Deploying a Jupyter Notebook

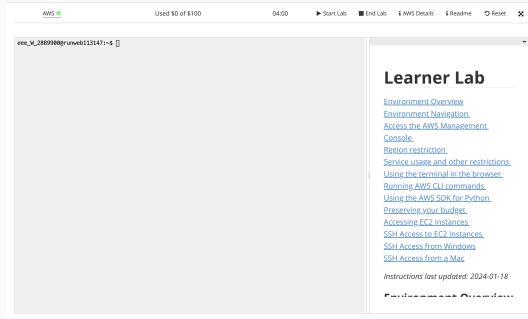
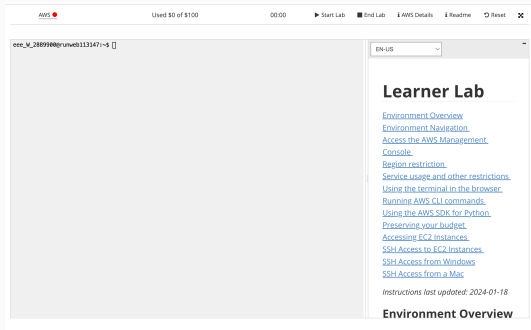
What is EC2?

EC2 is like a virtual machine in the cloud.



Staring the lab

1. Log in to the AWS Educate platform with the account you created. AWS Learner Lab
2. Go to Modules, and click Launch AWS Academy Learning LAB, and then click on start lab. Once the lab is started, you will see a green circle. *It may take a few minutes to start the lab.*



Be aware that the AWS Educate platform provides you with a temporary AWS account. This account is valid for a limited time and is not a real AWS account. Only some AWS services are available, and IAM is impossible to configure for security reasons.

Be aware that the AWS Educate platform provides you with a temporary AWS account. This account is valid for a limited time and is not a real AWS account. Only some AWS services are available, and IAM is impossible to configure for security reasons.

However, we can use AWS CLI credentials to communicate with different services. Also, we have the **LAB_ROLE** if we need to indicate a role.

Be aware that the AWS Educate platform provides you with a temporary AWS account. This account is valid for a limited time and is not a real AWS account. Only some AWS services are available, and IAM is impossible to configure for security reasons.

However, we can use AWS CLI credentials to communicate with different services. Also, we have the **LAB_ROLE** if we need to indicate a role.

To get the credentials, go to the AWS Educate platform and click on Account Details. You will see the AWS CLI credentials.

AWS Management Console

You will see the **AWS Management Console** if you click on the green circle. This is the main interface to manage the AWS services.

The screenshot displays the AWS Management Console interface. At the top, a dark navigation bar includes the AWS logo, a 'Services' menu, a search bar, and a keyboard shortcut '[Option+S]'. On the right side of the bar, it shows the current region as 'N. Virginia' and the user's profile 'voclabs/user3075385+Jordi_Mateo @ 6546-5438-9363'. Below the navigation bar, the main content area is titled 'Console Home'. It features a 'Reset to default layout' button and an 'Add widgets' button. The dashboard is organized into several sections: 'Recently visited' on the left lists 'DynamoDB', 'RDS', 'Amazon SageMaker', and 'IAM', with a 'View all services' link at the bottom. The 'Applications' section on the right shows '0' applications for the 'us-east-1 (Current Region)' and includes a 'Create application' button. Below this, a table header is visible with columns for 'Name', 'Description', 'Region', and 'Originating account'. The bottom row contains three sections: 'Welcome to AWS' with a 'Getting started with AWS' link, 'AWS Health' showing '0' open issues and '0' scheduled changes, and 'Cost and usage' displaying 'Current month costs' as '\$0.00' and 'Total costs per month' as '\$0.00015'. A footer bar at the very bottom contains links for 'CloudShell', 'Feedback', and copyright information for 2024 Amazon Web Services.

Console Home [Info](#)

[Reset to default layout](#) [+ Add widgets](#)

Recently visited [Info](#)

- DynamoDB
- RDS
- Amazon SageMaker
- IAM

[View all services](#)

Applications (0) [Info](#)

Region: US East (N. Virginia)

[Create application](#)

us-east-1 (Current Region)

< 1 >

Name	Description	Region	Originating account
No applications Get started by creating an application. Create application			

[Go to myApplications](#)

Welcome to AWS

[Getting started with AWS](#)
Learn the fundamentals and find valuable information to get the most out of AWS.

AWS Health [Info](#)

Open issues
0 Past 7 days

Scheduled changes
0 Upcoming and past 7 days

Cost and usage [Info](#)

Current month costs
\$0.00

Forecasted month end costs
There isn't enough historical data to

Total costs per month
Cost (USD)
\$0.00015
\$0.0001

CloudShell Feedback © 2024, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

Using custom key pair

AWS provides a key pair to connect to the EC2 instance. However, we can use our key pair.

1. Go to search and write Key Pairs.
2. Click on Key Pairs.
3. Click on Actions and then on Import Key Pair.
4. Fill the form with the following settings:
 - Name: aws-yourname (aws-instructor)
 - Browse and select the public key file we created before.
5. Import the key pair.

Creating an EC2 instance

1. Click on the Services and then on EC2.
2. Launch an instance.
3. Fill the form with the following settings:
 - Name: Jupyter Notebook
 - Image: Amazon Linux 2 AMI (HVM) - SSD Volume Type
 - Architecture: 64-bit (x86)
 - Type: t2.micro
 - Key pair: use the key pair created before. (aws-instructor)
 - Network: default VPC

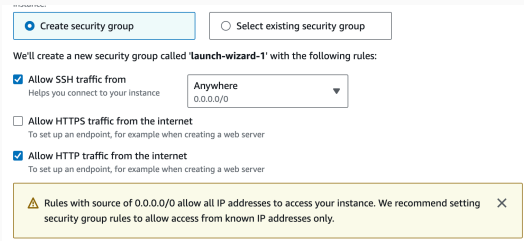
Security Group

This instance requires a security group that allows traffic on port 22 (SSH) and port 80 (HTTP).

- SSH is used to connect and manage the instance.
- HTTP is used to access the Jupyter Notebook from the browser.

Mark the checkbox to create a new security group and fill the form with the following settings:

- Allow SSH from anywhere
- Allow HTTP traffic from the internet




☒ Create security group ☐ Select existing security group

We'll create a new security group called 'launch-wizard-1' with the following rules:





- ☒ Allow SSH traffic from
Helps you connect to your instance
Anywhere
0.0.0.0/0
- ☐ Allow HTTPS traffic from the internet
To set up an endpoint, for example when creating a web server
- ☒ Allow HTTP traffic from the internet
To set up an endpoint, for example when creating a web server

⚠ Rules with source of 0.0.0.0/0 allow all IP addresses to access your instance. We recommend setting security group rules to allow access from known IP addresses only. ✕

Instance Details

 Services

[Option+S]



N. Virginia

voclabs/user3075385=Jordl_Mateo @ 6546-5438-9363

EC2 Dashboard

EC2 Global View

Events

Console-to-Code [Preview](#)

▼ Instances

Instances

Instance Types

Launch Templates

Spot Requests

Savings Plans

Reserved Instances

Dedicated Hosts

Capacity Reservations [New](#)

▼ Images

AMIs

AMI Catalog

▼ Elastic Block Store

Volumes

Snapshots


Lifecycle Manager

▼ Network & Security

Security Groups


Elastic IPs


Instances (1/1) [Info](#)

 [Connect](#) [Instance state](#) [Actions](#) [Launch instances](#)

[Any state](#)

[X](#) [Clear filters](#)

< 1 > 


<input checked="" type="checkbox"/>	Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone	Public IP
<input checked="" type="checkbox"/>	Jupyter Noteb...	i-0a94b7c50c1acc1cd	 Running	t2.micro	-	View alarms	us-east-1c	ec2-3-87-

Instance: i-0a94b7c50c1acc1cd (Jupyter Notebook)

[Details](#) [Status and alarms \[New\]\(#\)](#) [Monitoring](#) [Security](#) [Networking](#) [Storage](#) [Tags](#)

▼ Instance summary [Info](#)

Instance ID

 i-0a94b7c50c1acc1cd (Jupyter Notebook)

IPv6 address

-


Hostname type

IP name: ip-172-31-27-105.ec2.internal


Answer private resource DNS name

IPv4 (A)


Auto-assigned IP address

 3.87.76.117 [Public IP]


Public IPv4 address

 3.87.76.117 [open address](#)

Instance state

 Running


Private IP DNS name (IPv4 only)

 ip-172-31-27-105.ec2.internal


Instance type

t2.micro


VPC ID

 vpc-0d621aaeaa1843bc3 [open](#)

Private IPv4 addresses

 172.31.27.105

Public IPv4 DNS

 ec2-3-87-76-117.compute-1.amazonaws.com [open address](#)


Elastic IP addresses

-

AWS Compute Optimizer finding

[Opt-in to AWS Compute Optimizer for recommendations.](#)

[Learn more](#)

 CloudShell

[Feedback](#)

© 2024, Amazon Web Services, Inc. or its affiliates. [Privacy](#) [Terms](#) [Cookie preferences](#)

Connecting to the instance

1. Open a terminal (mac,linux) or a powershell (windows).
2. Use the following command to connect to the instance. Replace the public DNS with the public DNS of your instance.
 - **aws-instructor** is the full path to the private key file. *A simple solution is to open the shell in the same directory as the private key file.*
 - **ec2-user** is the default user for the Amazon Linux AMI. For other AMIs, the user can be different.

```
ec2-user@ip-172-31-27-105:~  
Last login: Thu Feb 15 08:34:01 on ttys003  
> ssh -i aws-instructor ec2-user@ec2-3-87-76-117.compute-1.amazonaws.com  
  
The authenticity of host 'ec2-3-87-76-117.compute-1.amazonaws.com (3.87.76.117)'  
can't be established.  
ED25519 key fingerprint is SHA256:vHjhi2ndPIdqmQ9qJuXpkb3tqw+mvUNIAMTBRIOPST4.  
This key is not known by any other names.  
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes  
Warning: Permanently added 'ec2-3-87-76-117.compute-1.amazonaws.com' (ED25519) t  
o the list of known hosts.  
  
_#_  
~\_####_ Amazon Linux 2  
~~ \_#####\  
~~ \###| AL2 End of Life is 2025-06-30.  
~~ \#/_____  
~~ V~' '~>  
~~~~ A newer version of Amazon Linux is available!  
~~~~_*_/_____  
~/ /___/ Amazon Linux 2023, GA and supported until 2028-03-15.  
_/m'/_____ https://aws.amazon.com/linux/amazon-linux-2023/  
  
4 package(s) needed for security, out of 8 available  
Run "sudo yum update" to apply all updates.  
[ec2-user@ip-172-31-27-105 ~]$
```

```
ssh -i aws-instructor ec2-user@ec2-3-87-76-117.compute-1.amazonaws.com
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
```

Installing Software in the EC2 instance

It is a good practice to update the instance before installing new Software.

```
sudo yum update -y
```

- By default, the instance we selected comes with Python 3.7. However, we will install a newer version (Python 3.9).
- We are going to make an installation script.

```
vi install-python39.sh  
# press i to insert content  
# copy content in the next slide  
# press esc  
# write :wq to save and exit  
bash install-python39.sh
```

install-python39.sh

```
#!/bin/bash
PYTHON_VERSION=3.9.0
PYTHON_TAG=$(echo $PYTHON_VERSION | cut -d. -f1-2)
sudo yum install -y gcc openssl11-devel bzip2-devel libffi-devel sqlite-devel
wget https://www.python.org/ftp/python/$PYTHON_VERSION/Python-$PYTHON_VERSION.tgz
tar -zxvf Python-$PYTHON_VERSION.tgz
cd Python-$PYTHON_VERSION
sudo ./configure --enable-optimizations
sudo make altinstall
if [ -f "/usr/local/bin/python${PYTHON_TAG}" ]; then
    sudo ln -sf "/usr/local/bin/python${PYTHON_TAG}" /usr/bin/python3
fi
"/usr/local/bin/python${PYTHON_TAG}" -m pip install --upgrade awscli --user
"/usr/local/bin/python${PYTHON_TAG}" -m pip install --upgrade pip
cd ..
sudo rm -rf Python-$PYTHON_VERSION
sudo rm -rf Python-$PYTHON_VERSION.tgz
```

Creating a virtual environment

```
python3.9 -m venv jupyter-env
```

To manually activate the virtual environment, use the following command:

```
source jupyter-env/bin/activate
```

To automatically activate the virtual environment, add the following line to the `.bashrc` file:

```
echo "source jupyter-env/bin/activate" >> ~/.bashrc
```

```
(jupyter-env) [ec2-user@ip-172-31-28-96 ~]$
```

Installing Jupyter Notebook

1. Install Jupyter Notebook in the instance.

```
pip install jupyter
```

2. Upgrade pip to avoid warnings.

```
pip install --upgrade pip
```

3. Create a directory for the notebooks.

```
mkdir notebooks  
cd notebooks
```

Running Jupyter Notebook (I)

- Running on localhost (default): Not accessible from the internet.

```
jupyter notebook
```

- Running on the public IP: Accessible from the internet.

```
jupyter notebook --ip=? --port=?
```

- ip:
 - 0.0.0.0 (default): Listen on all IP addresses.
 - Public IP: Obtain the public IP from the instance and use it.
- port:
 - 8888 (default): The default port for the Jupyter Notebook. **This port is not opened by the security group.**

Opening the port 8888

1. Search for Security Groups (EC2). Click on **launch-wizard-1**.
2. Edit the inbound rules and add a new rule with the following settings:
 - Type: Custom TCP
 - Port Range: 8888
 - Source: Anywhere (0.0.0.0/0)
3. Delete the old rule for HTTP (port 80).
4. Save the changes.

EC2 > Security Groups > sg-04dec016fbffc6eb6 - launch-wizard-1 > Edit inbound rules

Edit inbound rules [info](#)

Inbound rules control the incoming traffic that's allowed to reach the instance.

Inbound rules [info](#)

Security group rule ID	Type info	Protocol info	Port range info	Source info	Description - optional info
sg-00ad2188b6eb6136	SSH	TCP	22	Custom... 0.0.0.0/0	Delete
sg-087910e2078bc4b6e	HTTP	TCP	80	Custom... 0.0.0.0/0	Delete
-	Custom TCP	TCP	8888	Any... 0.0.0.0/0	Delete

[Add rule](#)

EC2 > Security Groups > sg-04dec016fbffc6eb6 - launch-wizard-1

sg-04dec016fbffc6eb6 - launch-wizard-1 [Actions](#)

Details

Security group name launch-wizard-1	Security group ID sg-04dec016fbffc6eb6	Description launch-wizard-1 created 2024-02-15T11:55:21.850Z	VPC ID vpc-0d621a9aa18d3bc1
Owner 654654389963	Inbound rules count 2 Permission entries	Outbound rules count 1 Permission entry	

[Inbound rules](#) [Outbound rules](#) [Tags](#)

Inbound rules (2) [Manage tags](#) [Edit inbound rules](#)

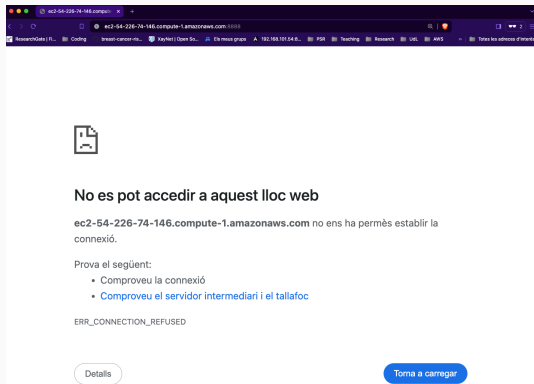
Search

	Name	Security group rule...	IP version	Type	Protocol	Port range	Source
<input type="checkbox"/>	-	sg-00ad2188b6eb6c01...	IPv4	SSH	TCP	22	0.0.0.0/0
<input type="checkbox"/>	-	sg-00eb4946b0b9ac5...	IPv4	Custom TCP	TCP	8888	0.0.0.0/0

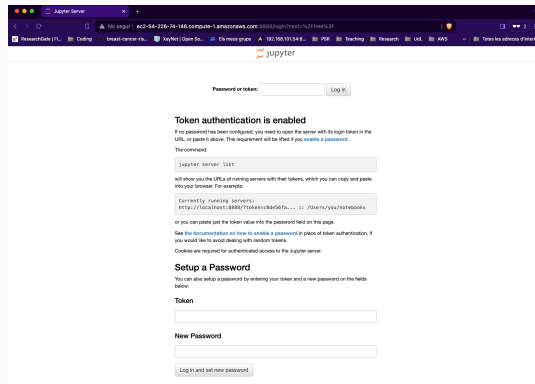
Running Jupyter Notebook (II)

```
jupyter notebook --ip=0.0.0.0 --port=8888  
# In the browser  
http://ec2-3-87-76-117.compute-1.amazonaws.com:8888
```

Before running the command



After running the command

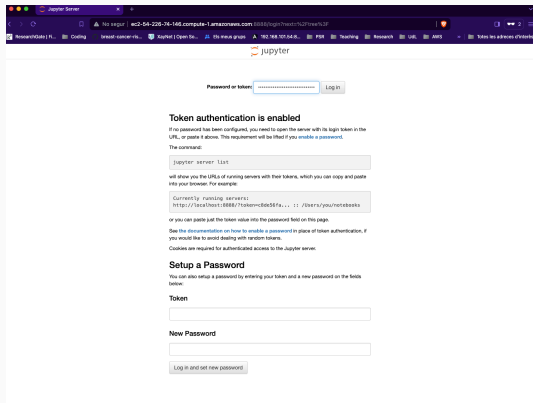


Accessing the Jupyter Notebook

1. Copy the token from the terminal. Then, log in to the Jupyter Notebook from the browser.

```
ec2-user@ip-172-31-27-105:~/notebooks
[I 2024-02-15 16:20:35.961 ServerApp] http://127.0.0.1:8888/tree?token=03c33d65561688656f4d870407bef3977c7a34376d8fe5d3
[I 2024-02-15 16:20:35.961 ServerApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation).
[W 2024-02-15 16:20:35.969 ServerApp] No web browser found: Error('could not locate runnable browser').
[C 2024-02-15 16:20:35.969 ServerApp]

To access the server, open this file in a browser:
file:///home/ec2-user/.local/share/jupyter/runtime/jpserver-3354-open.html
Or copy and paste one of these URLs:
http://172.31.27.105:8888/tree?token=03c33d65561688656f4d870407bef3977c7a34376d8fe5d3
http://127.0.0.1:8888/tree?token=03c33d65561688656f4d870407bef3977c7a34376d8fe5d3
[I 2024-02-15 16:20:35.991 ServerApp] Skipped non-installed server(s): bash-language-server, dockerfile-language-server-nodejs, javascript-typescript-languageserver, jedi-language-server, julia-language-server, pyright, python-language-server, python-lsp-server, r-languageserver, sql-language-server, texlab, typescript-language-server, unified-language-server, vscode-css-languageserver-bin, vscode-html-languageserver-bin, vscode-json-languageserver-bin, yamllanguage-server
[I 2024-02-15 16:20:42.444 ServerApp] 302 GET / (@91.126.177.34) 0.74ms
[I 2024-02-15 16:20:42.545 JupyterNotebookApp] 302 GET /tree? (@91.126.177.34) 0.89ms
```



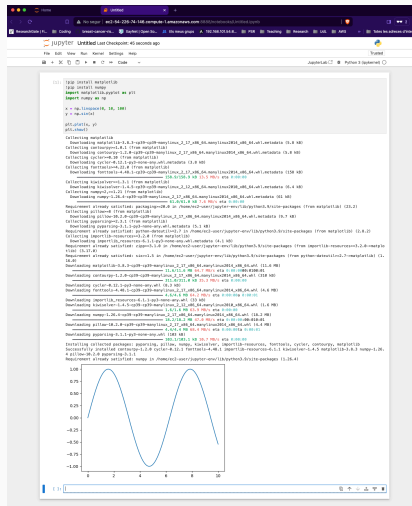
Running python code in the notebook

Click on New and then on Python 3. Write the following code and run it.

```
!pip install matplotlib
!pip install numpy
import matplotlib.pyplot as plt
import numpy as np

x = np.linspace(0, 10, 100)
y = np.sin(x)

plt.plot(x, y)
plt.show()
```



Common Issues:

- *Browser Protocol:* Ensure that the correct browser protocol is being used. In our case, we are using http, not https.
- *EC2 Instance:* Please note, that the public IP can change is a dynamic setting.

Conclusion

RECAP: Summary of the session

1. We have learned the primary servers in AWS, such as EC2.

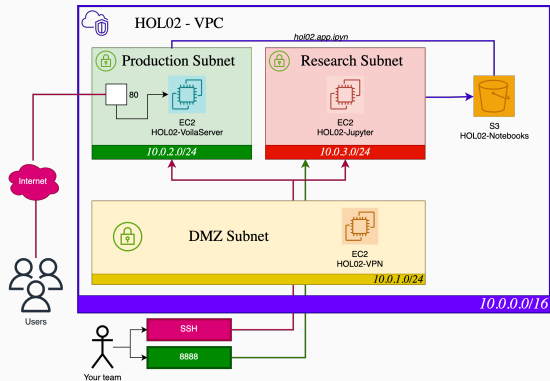
RECAP: Summary of the session

1. We have learned the primary servers in AWS, such as EC2.
2. We have learned how to deploy a Jupyter Notebook in EC2. We have learned about security groups, key pairs, and how to connect to the instance using SSH.

Tasks

Homework

HandsOnLab02: Deploying a hybrid infrastructure for researchers in AWS



For this hands-on lab, you will need to follow the instructions: HOL02