

Documentación del Laboratorio 0: AWS-EMR

Nombre: Liz Oriana Rodrigues Cruz
Fecha: 27 de Mayo

Prerrequisitos:

1. Creación de par de claves

<input type="checkbox"/>	key1	rsa	2022/04/24 20:42 GMT-5
<input checked="" type="checkbox"/>	Lab3-0	rsa	2022/05/27 12:07 GMT-5

2. Creación de S3 - Bucket

El URI de S3 es: <s3://lab3-0-bucket/salida/>

lab3-0-bucket [Info](#)

[Objetos](#) | [Propiedades](#) | [Permisos](#) | [Métricas](#) | [Administración](#) | [Puntos de acceso](#)

Objetos (1)
Los objetos son las entidades fundamentales que se almacenan en Amazon S3. Puede utilizar el [inventario de Amazon S3](#) para obtener una lista de todos los objetos de su bucket. Para que otras personas obtengan acceso a sus objetos, tendrá que concederles permisos de forma explícita. [Más información](#)

Copiar URI de S3 Copiar URL Descargar Abrir Eliminar **Acciones**

[Crear carpeta](#) [Cargar](#)

1

<input type="checkbox"/>	Nombre	Tipo	Última modificación	Tamaño	Clase de almacenamiento
<input type="checkbox"/>	salida/	Carpeta	-	-	-

Creación del Cluster

[Crear clúster](#) [Ver detalles](#) [Clonar](#) [Finalizar](#)

Filter: [Todos los clústeres](#) Clústeres: 3 (todos cargados)

	Nombre	ID	Estado
<input type="checkbox"/> ▶	Lab3-0-Cluster	j-33U2RY3LFVN5M	Esperando Preparado para el clúster

Comprobamos la conexión SSH

Se realiza en el directorio donde este el archivo Lab3-0.pem utilizando el comando:

```
ssh -i ./Lab3-0.pem hadoop@ec2-44-200-55-196.compute-1.amazonaws.com
```

```
MacBook-Pro-de-liz:Lab3-0 liz.rodriguez$ ssh -i ./Lab3-0.pem hadoop@ec2-44-200-55-196.compute-1.amazonaws.com
The authenticity of host 'ec2-44-200-55-196.compute-1.amazonaws.com (44.200.55.196)' can't be established.
ED25519 key fingerprint is SHA256:dMLV2YX4RuqEqT4vhT8dBvE2AkuUAaMKvTy5k9k1MY.
This key is not known by any other names
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added 'ec2-44-200-55-196.compute-1.amazonaws.com' (ED25519) to the list of known hosts.
Last login: Sat May 28 22:00:43 2022

 _ _ | _ _ | _ )
 _ | ( _ _ /   Amazon Linux 2 AMI
 _ _ | \ _ _ | _ _ |

https://aws.amazon.com/amazon-linux-2/
49 package(s) needed for security, out of 88 available
Run "sudo yum update" to apply all updates.
-bash: warning: setlocale: LC_CTYPE: cannot change locale (UTF-8): No such file or directory

EEEEEEEEEEEEEEEEEEEE MMMMMMM      MMMMMMM RRRRRRRRRRRRRR
E::::::::::::::::::::E M::::::::M      M::::::::M R::::::::::::R
EE::::::::EEEEEEEE::::E M::::::::M      M::::::::M R::::::::RRRRR::::R
E::::E      EEEEE M::::::::M      M::::::::M RR::::R R::::R
E::::E      M::::::::M M::::::::M M::::::::M R::::R R::::R
E::::EEEEEEEE M::::::::M M::::::::M M::::::::M R::::RRRRR::::R
E::::::::::::E M::::::::M M::::::::M M::::::::M R::::::::RR
E::::EEEEEEEE M::::::::M M::::::::M M::::::::M R::::RRRRR::::R
E::::E      M::::::::M M::::::::M M::::::::M R::::R R::::R
E::::E      EEEEE M::::::::M      M M::::::::M R::::R R::::R
EE::::::::EEEEEEEE::::E M::::::::M      M::::::::M R::::R R::::R
E::::::::::::E M::::::::M      M::::::::M RR::::R R::::R
EEEEEEEEEEEEEEEEEEEE MMMMMMM      MMMMMMM RRRRRRR      RRRRRR

[hadoop@ip-172-31-2-3 ~]$ 0
-bash: 0: command not found
[hadoop@ip-172-31-2-3 ~]$
```

Configuraciones los puertos

Configuración de bloqueo del acceso público

Bloquear acceso público

Act. [Cambiar](#)

Excepciones

Un clúster se puede lanzar con las reglas del grupo de seguridad que permiten

Editar

Rango de puertos ▲

22

8888-8888

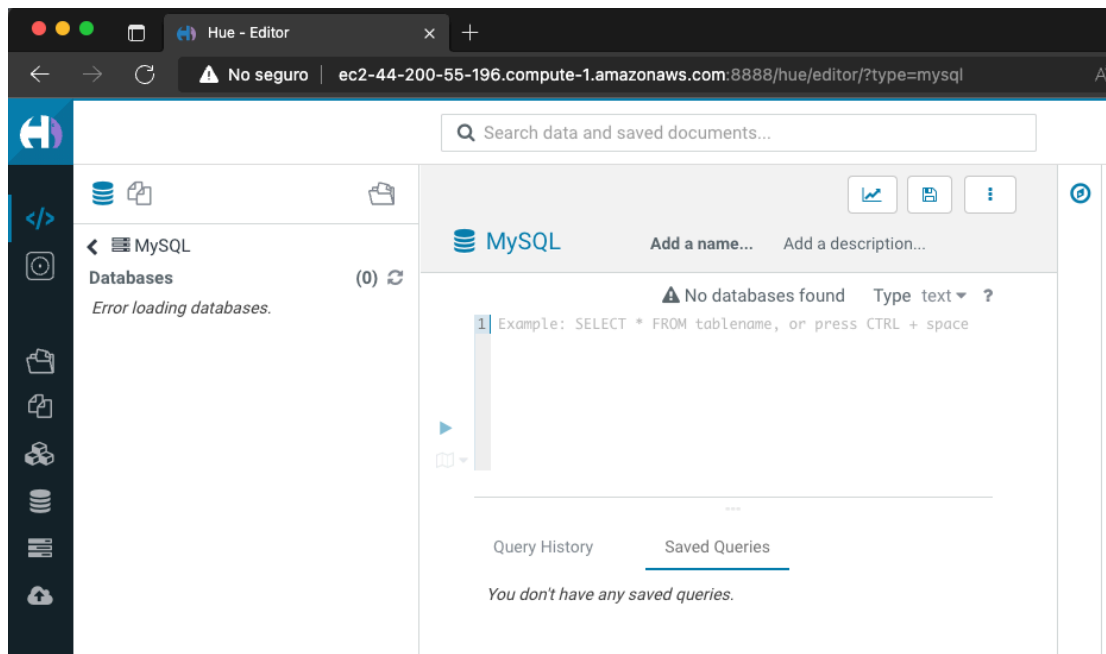
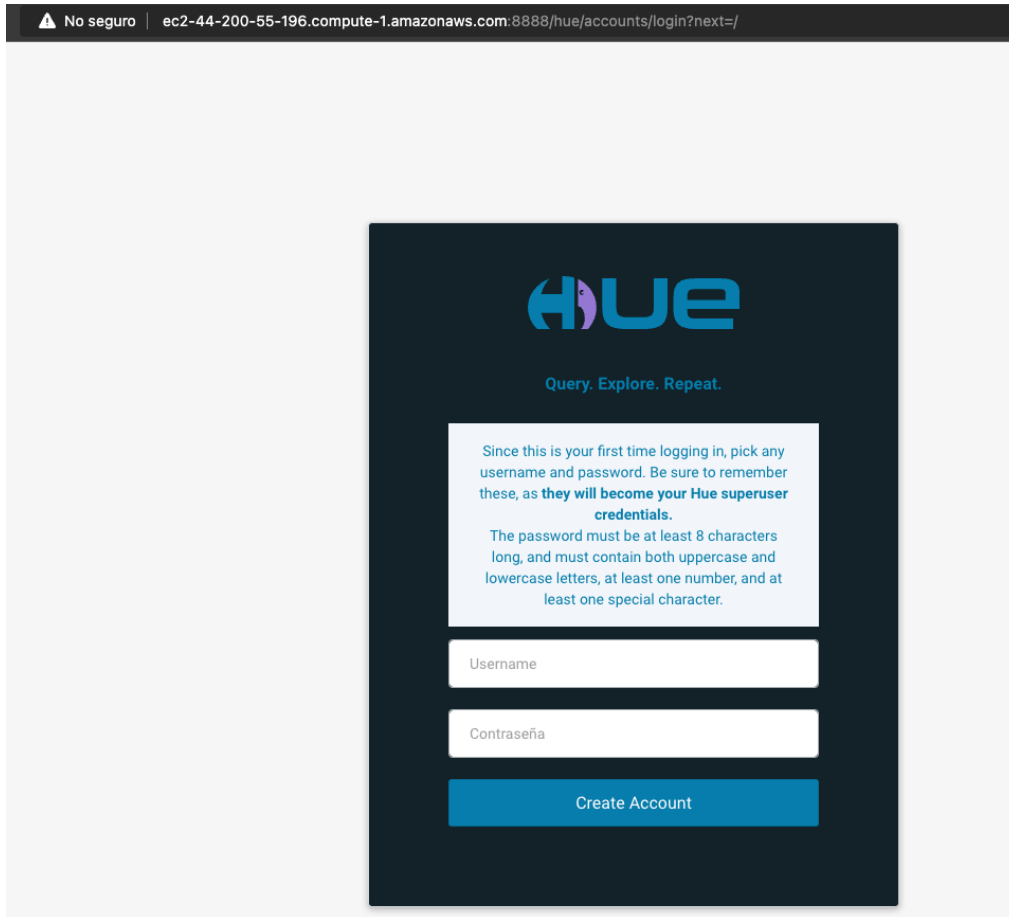
8890-8890

9443-9443

Acceso a HUE

Accedemos a la pagina por medio de la URL:

<http://ec2-44-200-55-196.compute-1.amazonaws.com:8888/hue/accounts/login?next=/>



Acceso a JupyterHub

Accedemos a la pagina por medio de la URL:

<https://ec2-44-200-55-196.compute-1.amazonaws.com:9443/hub/login?next=%2Fhub%2F>

The image shows two screenshots of a JupyterHub environment. The top screenshot is the login page, which has a dark header with the JupyterHub logo and a central orange 'Sign in' box. The box contains fields for 'Username:' and 'Password:', and a 'Sign in' button. The browser's address bar shows the URL: `ec2-44-200-55-196.compute-1.amazonaws.com`.

The bottom screenshot shows the JupyterLab interface after login. It features a top navigation bar with 'Logout' and 'Control Panel' buttons. Below this is a 'Files' tab, a 'Running' tab, and a 'Clusters' tab. A message says 'Select items to perform actions on them.' with 'Upload', 'New', and a refresh icon. Below this is a file browser showing a folder '/' and a message 'The notebook list is empty.' with columns for 'Name', 'Last Modified', and 'File size'.

The bottom part of the screenshot shows the JupyterLab code editor. It has a menu bar (File, Edit, View, Insert, Cell, Kernel, Widgets, Help) and a toolbar with icons for file operations and execution. The code area shows the following:

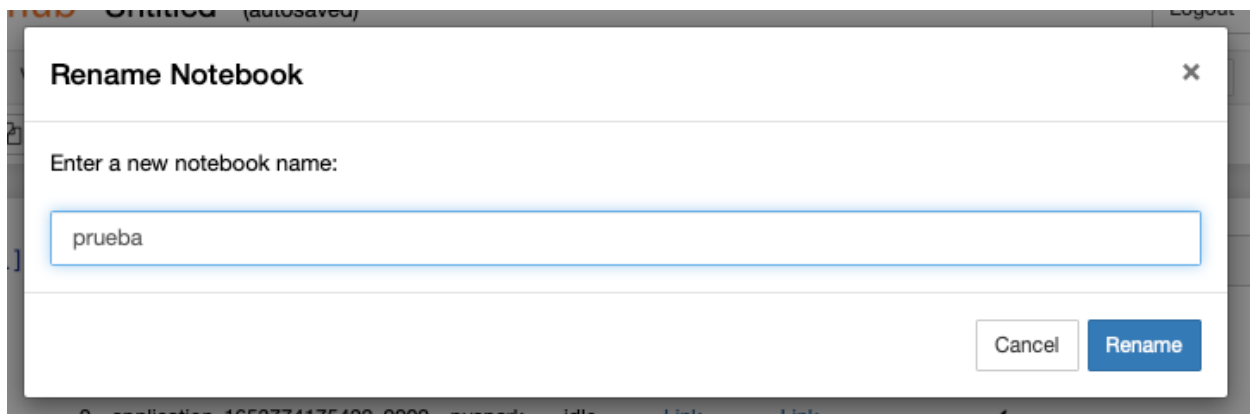
```
In [1]: spark
Starting Spark application


| ID | YARN Application ID            | Kind    | State | Spark UI             | Driver log           | Current session? |
|----|--------------------------------|---------|-------|----------------------|----------------------|------------------|
| 0  | application_1653774175422_0002 | pyspark | idle  | <a href="#">Link</a> | <a href="#">Link</a> | ✓                |


SparkSession available as 'spark'.
<pyspark.sql.session.SparkSession object at 0x7fc08d9e1750>

In [2]: sc
<SparkContext master=yarn appName=livy-session-0>

In [ ]:
```



Rename Notebook

Enter a new notebook name:

prueba

Cancel Rename

Acceso a Zeppelin

Accedemos a la pagina por medio de la URL:

<http://ec2-44-200-55-196.compute-1.amazonaws.com:8890/#/notebook/2H53UWDWC>

