

Prácticas BigData

HUE

1. Configuración de HUE

- Una vez terminado el proceso sin errores, podemos acceder al directorio

```
cd /opt/hadoop/hue
```

- Accedemos al directorio de configuración de HUE
- /opt/hadoop/hue/desktop/conf
- Dentro, debemos tener el fichero hue.ini
- Debemos configurar los valores para:
 - HDFS
 - YARN
 - HIVE
- No hace falta ninguno más porque no tenemos otros productos instalados y no funcionarían.
- Debemos poner correctamente la dirección. En nuestro caso, debería bastar con sustituir localhost por "nodo1" y cambiar el puerto de 8220 a 9000 en la parte de HDFS
- La parte de HDFS debe quedar similar a la siguiente

```
# Settings to configure your Hadoop cluster.

#####

[hadoop]

# Configuration for HDFS NameNode
# -----

[[hdfs_clusters]]
# HA support by using HttpFs

[[[default]]]
# Enter the filesystem uri
fs_defaultfs=hdfs://nodo1:9000

# NameNode logical name.
```

```
## logical_name=
```

```
# Use WebHdfs/HttpFs as the communication mechanism.
```

```
# Domain should be the NameNode or HttpFs host.
```

```
# Default port is 14000 for HttpFs.
```

```
webhdfs_url=http://nodo1:50070/webhdfs/v1
```

- La parte de YARN debe ser similar a la siguiente

```
[[yarn_clusters]]
```

```
[[[default]]]
```

```
# Enter the host on which you are running the ResourceManager
```

```
resourcemanager_host=nodo1
```

```
# The port where the ResourceManager IPC listens on
```

```
resourcemanager_port=8032
```

```
# Whether to submit jobs to this cluster
```

```
submit_to=True
```

```
# Resource Manager logical name (required for HA)
```

```
## logical_name=
```

```
# Change this if your YARN cluster is Kerberos-secured
```

```
## security_enabled=false
```

```
# URL of the ResourceManager API
```

```
## resourcemanager_api_url=http://nodo1:8088
```

```
# URL of the ProxyServer API
```

```
## proxy_api_url=http://nodo1:8088
```

```
# URL of the HistoryServer API
```

```
## history_server_api_url=http://nodo1:19888Se hace con el siguiente comando.
```

- Y por último, la parte de HIVE debe poner lo siguiente

```
[beeswax]
```

```
# Host where HiveServer2 is running.
# If Kerberos security is enabled, use fully-qualified domain name (FQDN).
hive_server_host=nodo1

# Port where HiveServer2 Thrift server runs on.
hive_server_port=10000

# Hive configuration directory, where hive-site.xml is located
hive_conf_dir=/opt/hadoop/hive/conf
```

- Por último, y muy importante, debemos activar WEBHDFS en nuestro cluster, lo que permite hacer llamadas vía HTTP al cluster.
- Modificamos el fichero hdfs-site y añadimos la siguiente propiedad

```
<property>

  <name>dfs.webhdfs.enabled</name>

  <value>true</value>

</property>
```

- Y en el fichero core-site.xml añadimos la siguiente propiedad

```
<property>

  <name>hadoop.proxyuser.hue.hosts</name>

  <value>*</value>

</property>

<property>

  <name>hadoop.proxyuser.hue.groups</name>

  <value>*</value>

</property>
```

- Paramos el cluster

- Copiamos los ficheros de configuración al resto de nodos
- Arrancamos el cluster

2. Arrancar y probar HUE

- Para arrancar el cluster debemos ejecutar el siguiente comando (sería interesante añadir este directorio al PATH de Linux para no tener que buscarlo). Con la opción -d lo lanzamos en background

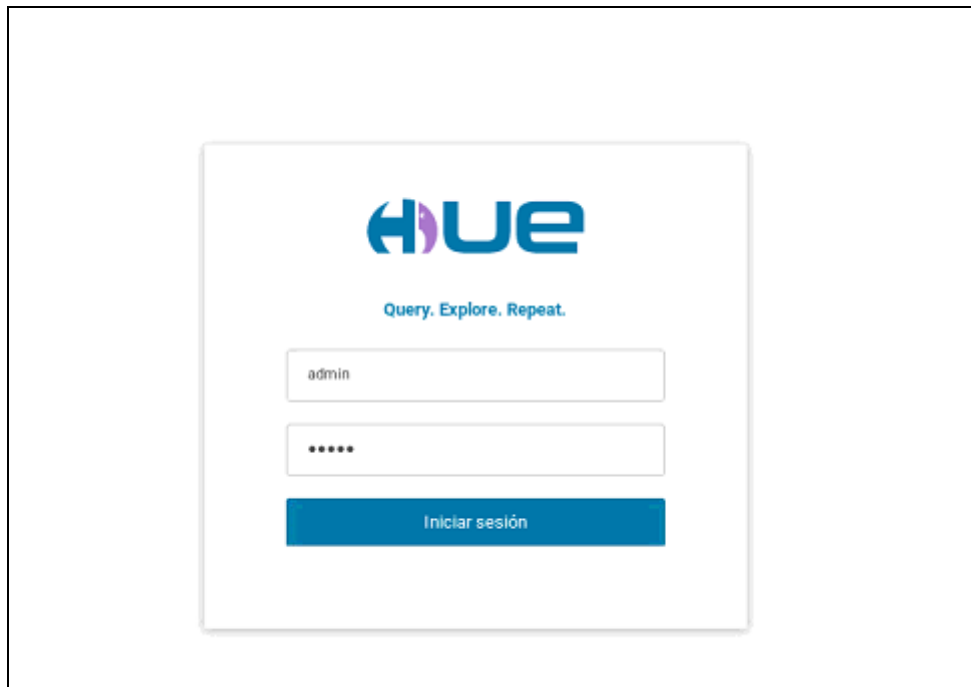
```
/opt/hadoop/hue/build/env/bin/supervisor -d
```

- Si todos va bien debemos tener el proceso funcionando

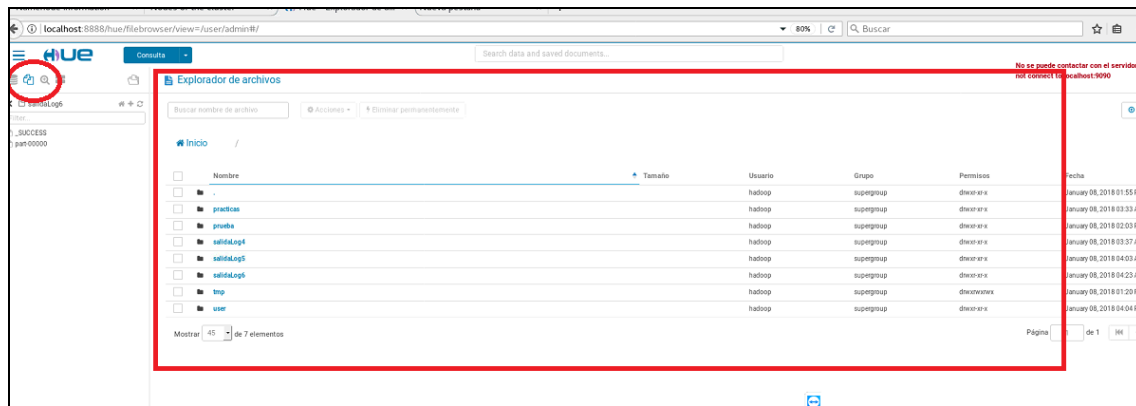
```
ps -ef | grep supervisor
```

```
hadoop      16257      1      1  11:57  ?                00:00:00
/opt/hadoop/hue/build/env/bin/python2.7
/opt/hadoop/hue/build/env/bin/supervisor -d
```

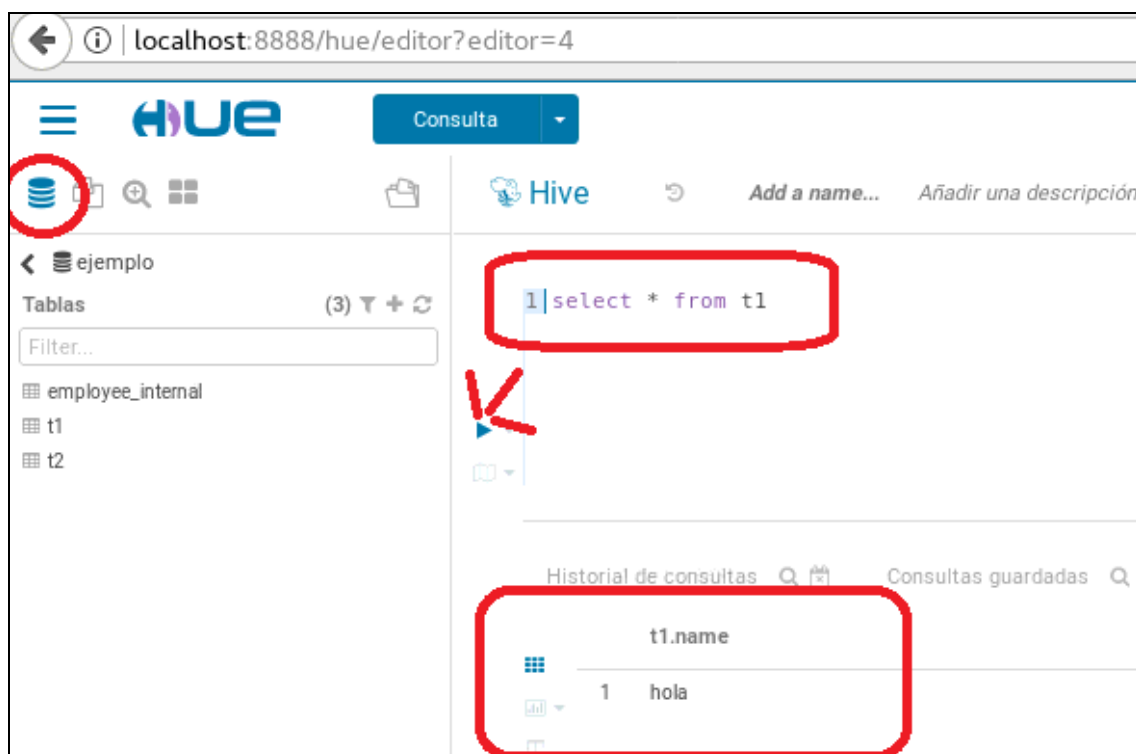
- Ahora abrimos el firefox y nos conectamos por el puerto 8888
- En la primera pantalla se nos pide establecer un usuario para la herramienta, en este caso lo llamo admin, puedes poner el nombre que quieras



- Una vez dentro, tenemos múltiples opciones, vamos a acceder al Browser de HDFS, donde podemos ver lo que tenemos.



- También podemos acceder a HIVE y ver lo que hemos hecho en las prácticas anteriores



- Desde la opción Consulta → Editor podemos acceder a numerosos editores de distintos productos.
- Como vemos, podemos hacer bastantes cosas con la herramienta