

Prácticas BigData

HIVE

1. HiveServer 2 y conexiones remotas

- Vamos ahora a trabar con el servidor HiveServer 2
- Primero debemos configurar la siguiente propiedad en el fichero hive-site.xml. La ponemos a FALSE. Hay que buscarla, no crearla de nuevo.
- De esa forma nos conectamos como usuario hadoop y desactivamos ciertas acciones de seguridad. En un curso más avanzado podríamos ver otras opciones, para esta formación es suficiente.

```
<property>
  <name>hive.server2.enable.doAs</name>
  <value>>false</value>
  <description>
    Setting this property to true will have HiveServer2 execute
    Hive operations as the user making the calls to it.
  </description>
</property>
```

- Arrancamos el servidor desde el directorio bbdd que hemos creado antes

```
hiveserver2 &
```

- Nos conectamos con beeline. No hace falta poner usuario

```
beeline
beeline> !connect jdbc:hive2://localhost:10000
Connected to: Apache Hive (version 2.3.2)
Driver: Hive JDBC (version 2.3.2)
Transaction isolation: TRANSACTION_REPEATABLE_READ
0: jdbc:hive2://nodo1:10000>
```

-
- Vamos a probar ahora que podemos conectarnos en remoto desde otro nodo al tener HIVESERVER2
- Copiamos el directorio hive al nodo2, en la misma ruta, es decir en /opt/hadoop
- Copiamos también el .bashrc al /home/hadoop del segundo nodo. Esto nos permite tener la misma configuración

- Entramos en beeline y le ponemos la dirección del nodo1

```
beeline> !connect jdbc:hive2://nodo1:10000
```

```
Connecting to jdbc:hive2://nodo1:10000
```

```
Enter username for jdbc:hive2://nodo1:10000:
```

```
Enter password for jdbc:hive2://nodo1:10000:
```

```
Connected to: Apache Hive (version 2.3.2)
```

```
Driver: Hive JDBC (version 2.3.2)
```

```
Transaction isolation: TRANSACTION_REPEATABLE_READ
```

- Comprobamos que vemos lo mismo

```
show databases;
```

```
+-----+
```

```
| database_name |
```

```
+-----+
```

```
| default      |
```

```
| ejemplo      |
```

```
| prueba       |
```

```
+-----+
```

```
3 rows selected (0,438 seconds)
```

-
- Vamos a probar ahora que podemos conectarnos en remoto desde otro nodo al tener HIVESERVER2
- Copiamos el directorio hive al nodo2, en la misma ruta, es decir en /opt/hadoop
- Copiamos también el .bashrc al /home/hadoop del segundo nodo. Esto nos permite tener la misma configuración
- Entramos en beeline y le ponemos la dirección del nodo1

```
beeline> !connect jdbc:hive2://nodo1:10000
```

```
Connecting to jdbc:hive2://nodo1:10000
```

```
Enter username for jdbc:hive2://nodo1:10000:
```

```
Enter password for jdbc:hive2://nodo1:10000:
```

```
Connected to: Apache Hive (version 2.3.2)
```

```
Driver: Hive JDBC (version 2.3.2)
```

```
Transaction isolation: TRANSACTION_REPEATABLE_READ
```

- Comprobamos que vemos lo mismo

```
show databases;
```

```
+-----+
```

database_name
default
ejemplo
prueba
3 rows selected (0,438 seconds)