

## Configurar Apache - Zeppelin

Lo mas recomendable es visitar la dirección oficial “<https://zeppelin.apache.org/>” para detectar las particularidades de la descarga y la configuración y pinchar en la parte de “Quick start” donde vienen los pasos para poder configurarlo correctamente.

Necesitamos descargamos:

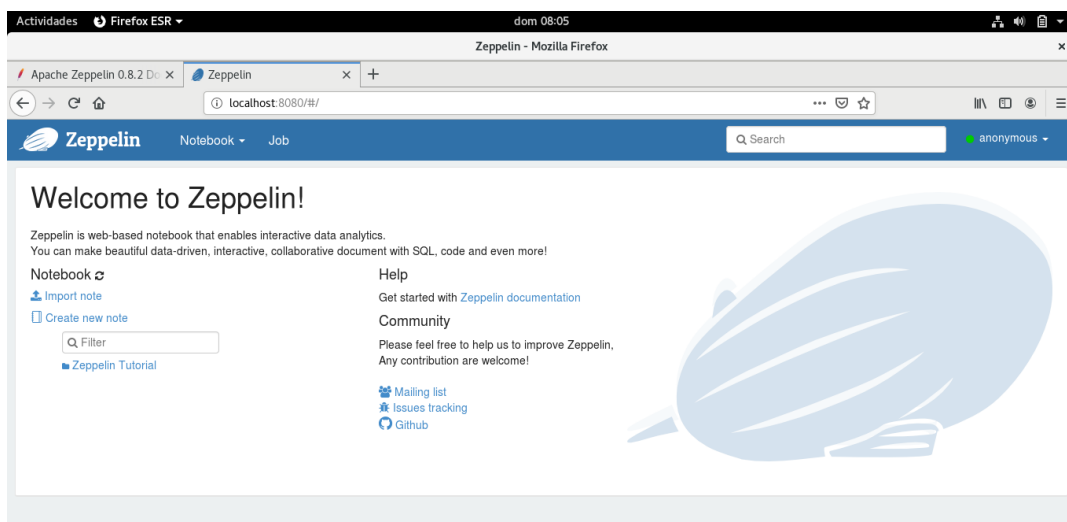
1. Spark (última versión estable)
2. Zeppelin (es importante bajarnos la versión donde venga -bin-all.tgz que contiene todo para trabajar)
3. JDK (lo más recomendable es bajar la versión 1.8 aunque funciona a partir de la 1.7)

Descargamos y descomprimos el archivo “zeppelin-0.8.2-bin-all.tgz” una vez descomprimido nos colocamos con un terminal en la ruta descomprimida.

Lanzamos el siguiente comando al estar en linux → bin/zeppelin-daemon.sh start

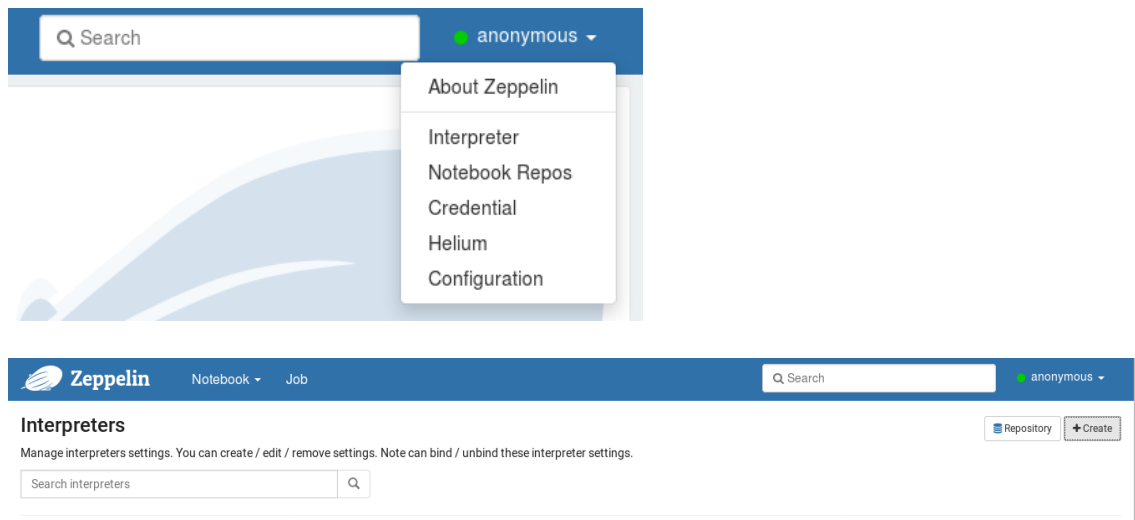
```
keepcoding@debian: ~/zeppelin-0.8.2-bin-all
Archivo  Editar  Ver  Buscar  Terminal  Ayuda
keepcoding@debian:~$ ls
Descargas      scala-2.12.1.deb
Documentos     snap
Escritorio     spark-2.0.2-bin-hadoop2.7
IdeaProjects   spark-2.0.2-bin-hadoop2.7.tgz
Imágenes       spark-2.4.4-bin-hadoop2.7
kafka_2.11-2.4.0  spark-2.4.4-bin-hadoop2.7.tgz
kafka_2.11-2.4.0.tgz  stale_outputs_checked
Música         Videos
Plantillas     zeppelin-0.8.2-bin-all
Público        zeppelin-0.8.2-bin-all.tgz
keepcoding@debian:~$ cd zeppelin-0.8.2-bin-all/
keepcoding@debian:~/zeppelin-0.8.2-bin-all$ bin/zeppelin-daemon.sh start
Zeppelin start [ OK ]
keepcoding@debian:~/zeppelin-0.8.2-bin-all$
```

Una vez realizado este comando, zeppelin se ha ejecutado en un servidor (en nuestro caso local) y por tanto accederemos a través de un navegador a la siguiente dirección para poder acceder a zeppelin → localhost:8080 (siendo el 8080 el puerto por el cual se ejecuta zeppelin, si tuviéramos algo ejecutándose en ese puerto tendríamos que ir a la carpeta donde descomprimimos zeppelin y en el directorio /conf/zeppelin-env.sh.template abríamos este archivo y escribimos “ZEPELIN\_PORT:NUEVO\_PUERTO” para que la aplicación se ejecutara en otro puerto distinto.



Ahora vamos a asociar zeppelin con spark.

Vamos al usuario anonymus en zeppelin y en el desplegable seleccionamos interpreter, una vez ahi dentro pinchamos en +create para empezar a configurar spark con zeppelin asociando asi un nuevo interprete.



En interpreter\_name ponemos el nombre del interprete que elijamos (en nuestro caso spark2) y en la opcion interpreter\_group → spark y podemos dar a save.

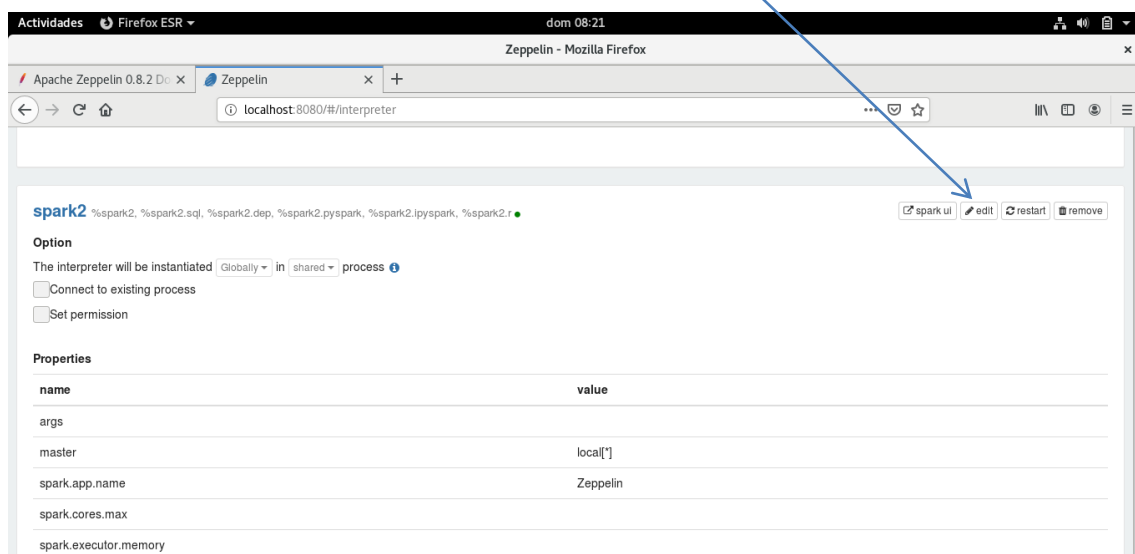
### Create new interpreter

Interpreter Name

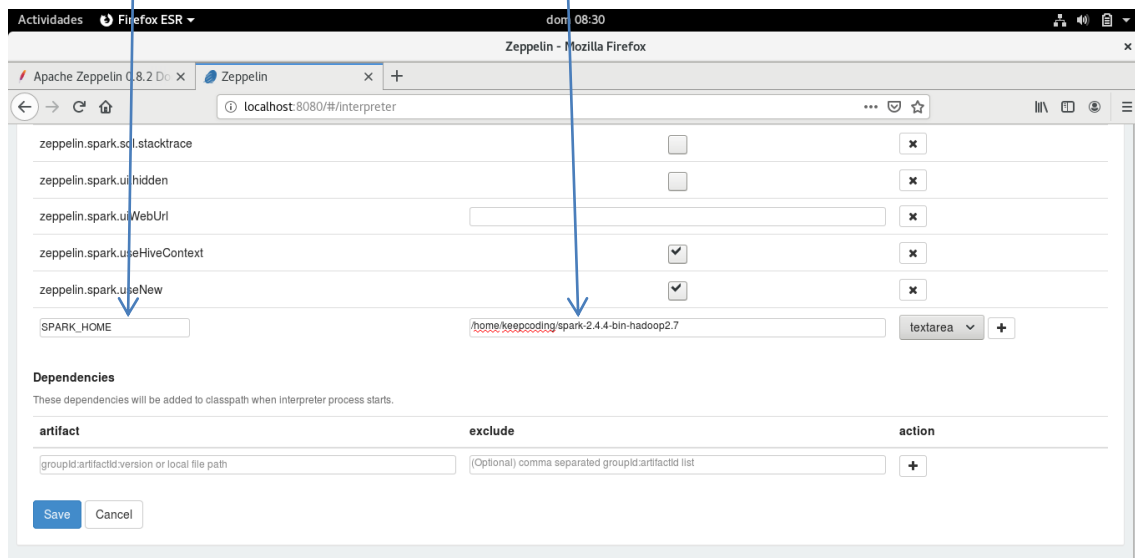
Interpreter group

En esta parte lo que hemos realizado es digamos realizar un “envoltorio” del nuevo interprete al que ahora le daremos el contenido que deseamos.

Nos vamos de nuevo a anonymus → interpreter → spark2 → edit y en la parte de propiedades, justo encima de donde pone “dependencies” escribimos los siguientes textos:



SPARK\_HOME y en el cuadro de la misma línea escribimos la ruta donde se encuentre el directorio de SPARK (en nuestro caso → /home/keepcoding/spark-2.4.4-bin-hadoop2.7) lo creamos y le damos a SAVE.



Ya podemos visualizar nuestra nueva propiedad con el intérprete que queríamos y reiniciamos el intérprete para que podamos utilizar apache con la versión que nos bajamos de Spark.

Nos aseguramos que el intérprete está bien asociado, pinchamos en Notebook → Create new note .

Create New Note

**Note Name**

**Default Interpreter**

spark2

Use '/' to create folders. Example: /NoteDirA/Note1

Create

Seleccionamos en el icono “interpreter binding” y verificamos que nuestro nuevo interprete no se encuentre en “gris” si así fuera pincharíamos en el y reiniciamos de nuevo el intérprete.

En un paragraph escribimos la sentencia : `sc.version` y nos aseguramos de que nos devuelve la versión de spark que queríamos asociar.

