

# Hive installation

Dario Colazzo

# Main steps

- There are many post on the Web indication step to follow in order to install Hive
- For instance this recent and rather complete one:

[http://www.bogotobogo.com/Hadoop/BigData\\_hadoop\\_Hive\\_Install\\_On\\_Ubuntu\\_16\\_04.php](http://www.bogotobogo.com/Hadoop/BigData_hadoop_Hive_Install_On_Ubuntu_16_04.php)

- It is for Linux but works for MacOS too (tested on my MacBook)
- Attention: incomplete indication for configuring hive-site.xml

# hive-site.xml

- The crucial steps concern local and HDFS paths
- Here is a snapshot of my configuration that should be clear enough (sometimes you see 'dario' but there you should put you user-name instead)

```
.....
<property>
  <name>hive.exec.stagingdir</name>
  <value>.hive-staging</value>
  <description>Directory name that will be created inside table locations in order to support HDFS encryption. This is replaces ${hive$
</property>
<property>
  <name>hive.exec.scratchdir</name>
  <value>/tmp/hive/${hive.session.id}_resources</value>
  <description>HDFS root scratch dir for Hive jobs which gets created with write all (733) permission. For each connecting user, an HD$
</property>
<property>
  <name>hive.exec.local.scratchdir</name>
  <value>/tmp/dario</value>
  <description>Local scratch space for Hive jobs</description>
</property>
<property>
  <name>hive.downloaded.resources.dir</name>
  <value>/tmp/dario_resources</value>
  <description>Temporary local directory for added resources in the remote file system.</description>
</property>
<property>
.....
```

- Of course I had to create directories /tmp/dario and /tmp/dario\_resources

# Metastore

- In case first attempts to launch Hive you may need to erase the 'metastore\_db' directory
- In my installation it is here  
`/usr/local/apache-hive-2.1.1-bin/bin`
- Go inside the 'metastore\_db' directory, erase everything ( by means of 'sudo rm -rf \*')
- Than go back to the /bin directory and erase the metastore ('rmdir metastore\_db')
- Attention: do this only in the installation/first launching phase, otherwise you will loose access to your databases!

# First test

- Download the u.data file for MovieLens data from the Moodle.
- Execute the following commands, and then you can start execution queries on the u\_data table.

```
hive> CREATE TABLE u_data (  
    userid INT,  
    movieid INT,  
    rating INT,  
    unixtime STRING)  
ROW FORMAT DELIMITED  
FIELDS TERMINATED BY '\t'  
STORED AS TEXTFILE;
```

```
hive> OAD DATA LOCAL INPATH '...your path to u.data.../u.data' OVERWRITE INTO TABLE u_data;
```