

Further indications for Hadoop installation

Dario Colazzo

Main steps

- There are many post on the Web indication step to follow in order to install Hadoop 2.7.3
- For instance:

<http://hadooprecipes.blogspot.fr/2015/09/setting-up-hadoop-271-on-mac-os-x.html>

- Pay attentions in particular to the following aspects

Crucial recommendations

- Be sure that the last Java version is installed, or at least that Java 7 is available
- With the pseudo-distributed installation you have several parallel/concurrent threads for HDFS and YARN daemons.
- These interact via ssh and without the need of typing a password
 - Use **dsa** public-private key generation as indicated in the tutorial
 - Be sure that remote login has been habilitated on your laptop. On Mac OS this can be done System Preference -> Sharing.
 - If 'ssh localhost' still requires a password, it may be that your machine is configured so that **rsa** keys are needed.
 - To this end follows the same key generation instructions of the tutorial by simply replacing 'dsa' with 'rsa'

Crucial recommendations

- When manipulating XML files for configuration be sure to use a 'neutral' editor, pico for instance. It is important that no hidden character is handed
- Do not forget to add 'export JAVA_HOME= your Java home path' inside hadoop-env.sh configuration file.
- 'your Java home path' can be retrieved in MacOs with `/usr/libexec/java_home`

Crucial recommendations

- When launching YARN it may be the case that the name of your machine is not associated to the localhost IP
 - In that case you need to modify your /etc/hosts' file.
 - In this case let me know and I will help you.
- Once HDFS and YARN are on, you can also type 'hadoop fs -ls /' to start exploring HDFS file-directly hierarchy. The same holds for other HDFS commands other than 'ls'