Submitting Spark Applications

We are going to submit those applications using “client” deploy mode (**client mode**), since in our laptops we will not have a Apache Spark cluster running.

(Note, you need to have previously installed Apache Spark in your machine. See instructions).

In cluster mode, the Spark driver runs inside an application master process which is managed by cluster-manager on the cluster, and the client can go away after initiating the application. **In client mode**, the driver runs in the client process, and the application master is only used for requesting resources from spark cluster.

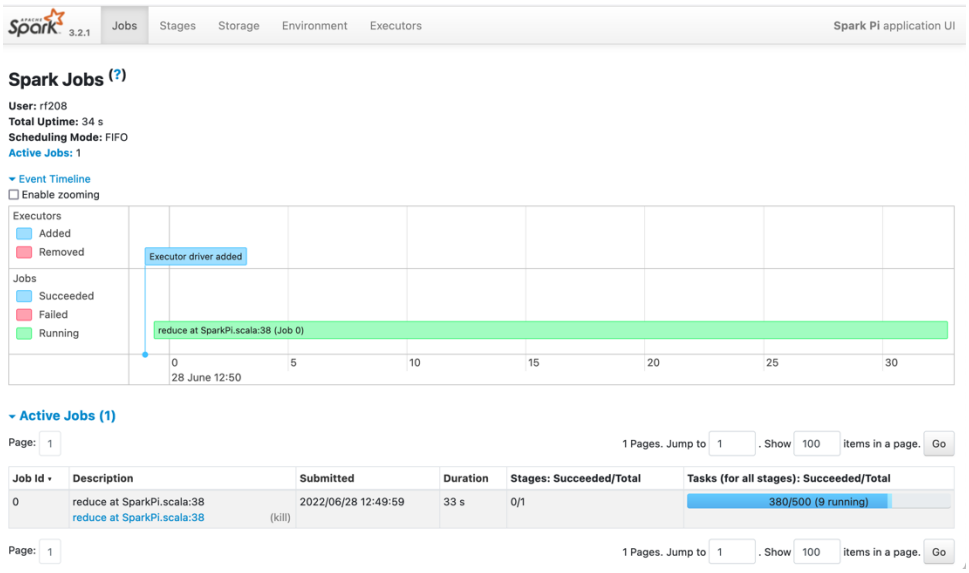
Submitting a the SparkPi Application:

1. Submit the ***SparkPi*** application (which is an example that comes with spark) to your Spark Cluster that you have started by using ***spark_submit_SparkPi*** script

   ./spark_submit_SparkPi.sh > result.txt

2. Check the result in result.txt:
   a. Pi is roughly 3.1403511403511404
3. To monitor the spark application run again the step 1 and go to the web browser: http://localhost:4040

   a. `The connection (localhost:4040) will ONLY alive while the application is running`



Submitting the WordCount Spark Application:

We can also submit the wordcount.py application stored in *Spark_Applications* folder

In cluster mode, the Spark driver runs inside an application master process which is managed by cluster-manager on the cluster, and the client can go away after initiating the application. **In client mode**, the driver runs in the client process, and the application master is only used for requesting resources from spark cluster.

>> ***./ spark_submit_wordcount.sh***

Important: Currently, the **standalone mode** does not support **cluster mode** for **Python applications**