**Requirements/ Libraries for download:**

python=2.7

pyspark=2.4.3

pandas=0.24.2

urllib3=1.25.3

spark=0.2.1

scipy=1.2.2

scikit-learn=0.20.3

azure=4.0.0

azureml-core=1.0.43

\*make sure your <JAVA\_HOME> is properly set.

InsightEdge=14.5.0 (from website)

Tableau=2019.2 (from website)

\*The computer that runs the Tableau should have installed maven, and add the gigaspaces-jdbc-client.jar to the Tableau drivers folder.

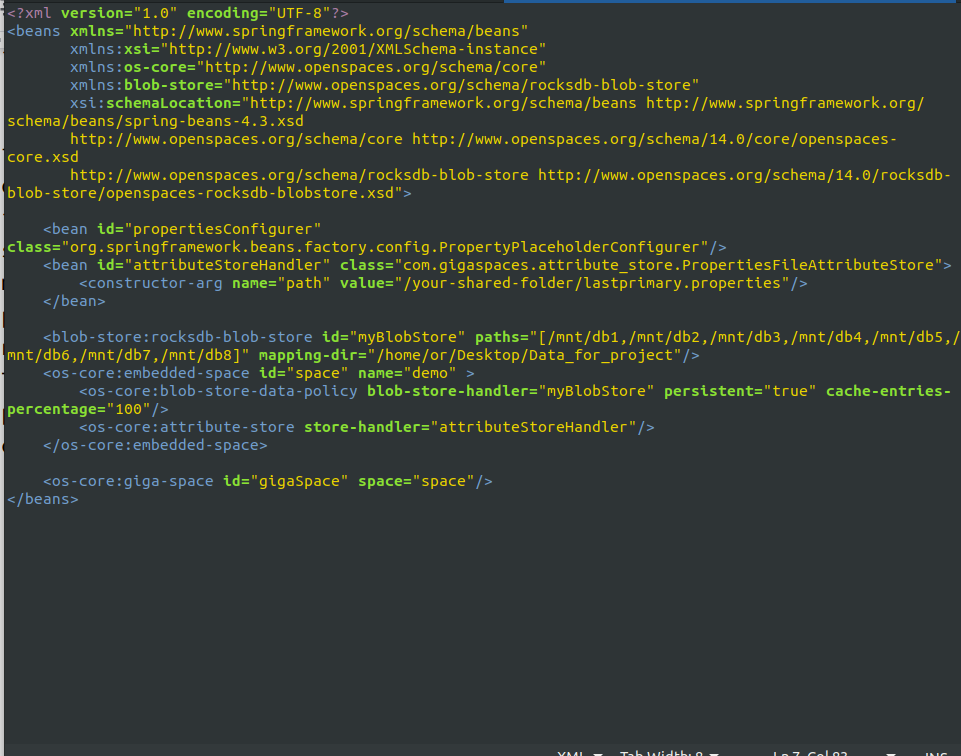
**What to change:**

First, in order to use the memoryXtend which allows data to be reloaded even after program restart, we first need to define to the program an external deployment processing unit to save the data to and also to load from:

Inside the gigaspaces-14.5 directory, go into the deploy folder, and then go into this path:

Rocksdb-pu -> META-INF -> spring

And go into the pu.xml file, and make sure that you rewrite it to look like this:



* **Note that mapping-dir is mapped to the folder that will act as the deployed processing-unit, then in the above line you are required to define the mnt paths.**

**How to run:**

Go into gigspaces-14.5 folder and then into the bin folder.

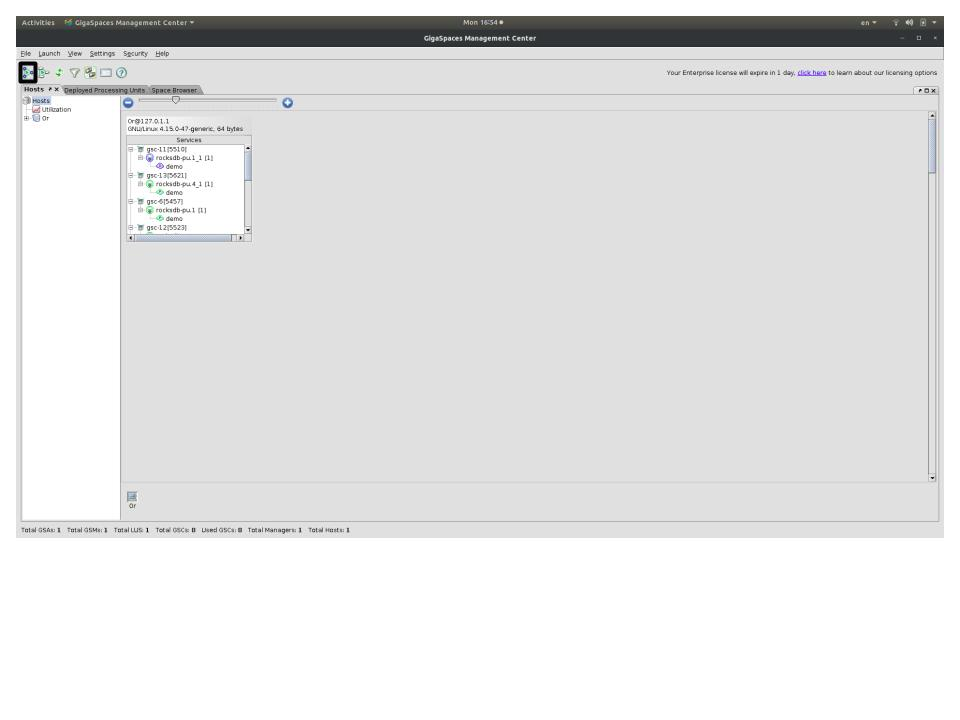
Now open two terminals and run each of the following commands each on different terminal:

./insightedge host run-agent --auto --gsc=8

./gs-ui.sh

**gs-ui**

In this ui, you will define the deployment processing unit, press on this button to do so:

now you need to define the deployment, make sure that XAP found the rocksdb that we defined, and that the processing unit is defined to partitioned, as well as 4 instances, and 1 back-up unit.

This tells the program to use in parallel 4 different GSMs, and have 1 back-up GSM for each one (resulting in 8 instances like the insightedge-command-line says).

**insightEdge command line:**

After a while after executing this command line, you will notice a pop-up internet window, at the address: <http://localhost:8099>

There you could notice that you can view the deployed processing units and make sure that they were loaded correctly.

Now switch to this address:

<http://localhost:9090>

Now you will get to zeppelin notebooks, in there you have files beginning in numbers : 1,2,3 - run them in the order of the numbers.

**Tableau Commands:**

1. Install InsightEdge maven artifacts (<GS\_HOME>\bin\insightedge maven install)

2. Run build-jdbc-client.cmd command (<GS\_HOME>\insightedge\tools\jdbc\build-jdbc-client.cmd)

3. Copy the created insightedge-jdbc-client.jar to C:\Program Files\Tableau\Drivers

4. Run <TABLEAU\_HOME>\bin\tableau.exe -DConnectPluginsPath=<GS\_HOME>\insightedge\tools\jdbc

5. In Tableau's "To a Server" menu, choose "Gigaspaces InsightEdge".

6. In Server field, fill the locator.

7. In Space field, write the space name.

8. All other fields are optional

9. Sign in