Profiling Pig query in Hadoop Cluster using Yourkit Profiler

Note: License is required to view gui but there is a 15 day trial version. Also, its free for pig contributors.

Source(Refer for more info): https://cwiki.apache.org/confluence/display/PIG/HowToProfile

1. On each machine create a yourkit directory (Make sure that the path is same in each machine)
   mkdir Yourkit

2. cd Yourkit

3. Download Youtkit for 64 bit kernel
   wget http://www.yourkit.com/download/yjp-11.0.2-linux.tar.bz2

4. Unzip YourKit
   tar xvjf yjp-11.0.2-linux.tar.bz2

5. Run pig query
   java -Dmapred.max.split.size=10000000
   -Dmapred.task.timeout=60000000
   -Dmapred.task.profile.maps=0-0
   -Dmapred.task.profile.reduces=0-0
    -Dmapred.task.profile=true
   -Dmapred.task.profile.params=-agentpath:/Yourkit/yjp-11.0.2/bin/linux-x86-xx/libyjpagent.so=dir=/tmp/snapshots,
   filters=/dev/null,
   tracing,
   disablej2ee
   -cp pig-0.9.0-core.jar:Hadoop/hadoop-0.xx.x/conf org.apache.pig.Main example.pig

6. After running the pig query, snap shot files are generated in /tmp/snapshots directory

7. Open the snapshot file viewer. It will ask for the license
   ./ YourKit/yjp-11.0.2/bin/yjp.sh

8. Open the snapshot files using GUI