1. T570 Ubuntu 18.04安装

export JAVA\_HOME=/usr/lib/jvm/jdk1.8.0\_181

export JRE\_HOME=${JAVA\_HOME}/jre

export CLASSPATH=.:${JAVA\_HOME}/lib:${JRE\_HOME}/lib

export SCALA\_HOME=/usr/share/scala-2.11

export MAVEN\_HOME=/usr/local/idea/plugins/maven/lib/maven3

export SPARK\_HOME=/usr/local/spark-2.2.1

export FLINK\_HOME=/usr/local/flink-1.5.5

export PATH=${PATH}:${JAVA\_HOME}/bin:${JRE\_HOME}/bin:${MAVEN\_HOME}/bin:${SCALA\_HOME}/bin

1) local flink:

$> cd $FLINK\_HOME

$> ./bin/start-cluster.sh //start flink on localhost, [http://localhost:8081](http://localhost:8081/)

$>./bin/stop-cluster.sh //stop flink on localhost

$> mvn package exec:java -Dexec.mainClass=gtl.data.sdb.beam.BeamSDBApp -Dexec.args="--inputFile=/home/vincent/gtl/data/dat/counties/county\_small.tsv --outputFile=/home/vincent/gtl/data/dat/temp/county\_small\_temp --runner=FlinkRunner" -Pflink-runner

=or==================================================================

$> mvn package -Pflink-runner

$> mvn exec:java -Dexec.mainClass=gtl.data.sdb.beam.BeamSDBApp -Dexec.args="--inputFile=/home/vincent/gtl/data/dat/counties/county\_small.tsv --outputFile=/home/vincent/gtl/data/dat/temp/county\_small\_temp\_flink --runner=FlinkRunner --flinkMaster=localhost:8081 --filesToStage=target/geobeam-bundled-1.0-SNAPSHOT.jar" -Pflink-runner

=reference==================================================================

https://beam.apache.org/get-started/quickstart-java/#run-wordcount

https://beam.apache.org/documentation/runners/flink/

2) local spark:

$> cd $SPARK\_HOME

$> ./bin/start-all.sh //start spark on localhost, [http://localhost:8080](http://localhost:8081/)

$>./bin/stop-all.sh //stop spark on localhost

$> mvn package exec:java -Dexec.mainClass=gtl.data.sdb.beam.BeamSDBApp -Dexec.args="--inputFile=/home/vincent/gtl/data/dat/counties/county\_small.tsv --outputFile=/home/vincent/gtl/data/dat/temp/county\_small\_spark --runner=SparkRunner" -Pspark-runner

=reference==================================================================

https://beam.apache.org/get-started/quickstart-java/#run-wordcount

<https://beam.apache.org/documentation/runners/spark/>

[http://spark.apache.org/docs/latest/#launching-on-a-cluster](http://spark.apache.org/docs/latest/" \l "launching-on-a-cluster)

<http://spark.apache.org/docs/latest/spark-standalone.html>

<http://spark.apache.org/docs/latest/running-on-yarn.html>

1. install git

sudo apt update

sudo apt install git

<https://www.cnblogs.com/chenwolong/p/GIT.html>

1. 运行flink版本的WordCount
2. 运行spark版本的WordCount
3. 运行beam版本的WordCount

1)vincent@H:~/wordcount/beam$ mvn archetype:generate \

-DarchetypeGroupId=org.apache.beam \

-DarchetypeArtifactId=beam-sdks-java-maven-archetypes-examples \

-DarchetypeVersion=2.8.0 \

-DgroupId=org.example \

-DartifactId=word-count-beam \

-Dversion="0.1" \

-Dpackage=org.apache.beam.examples \

-DinteractiveMode=false

2)vincent@H:~/wordcount/beam$ mvn compile exec:java -Dexec.mainClass=org.apache.beam.examples.WordCount -Dexec.args="--inputFile=pom.xml --output=counts" -Pdirect-runner

3)vincent@H:~/wordcount/beam$ mvn compile exec:java -Dexec.mainClass=org.apache.beam.examples.WordCount -Dexec.args="--runner=FlinkRunner --inputFile=pom.xml --output=counts" -Pflink-runner

4)vincent@H:~/wordcount/beam$ mvn package exec:java -Dexec.mainClass=org.apache.beam.examples.WordCount -Dexec.args="--runner=FlinkRunner --flinkMaster=<flink master> --filesToStage=target/word-count-beam-bundled-0.1.jar --inputFile=/path/to/quickstart/pom.xml --output=/tmp/counts" -Pflink-runner

5)vincent@H:~/wordcount/beam$mvn compile exec:java -Dexec.mainClass=org.apache.beam.examples.WordCount -Dexec.args="--runner=SparkRunner --inputFile=pom.xml --output=counts" -Pspark-runner

1. 运行geospark的range查询并记录时间
2. 运行geobeam的range查询并记录时间