This Big Data Benchmark for BigBench Specification ("Software") is maintained by Intel and Cloudera

About BigBench

The [https://github.com/intel-hadoop/Big-Data-Benchmark-for-Big-Bench|Big Data Benchmark for Big Bench]] specification based benchmark.

Run BigBench

You can run the testsuite [https://github.com/intel-hadoop/Big-Data-Benchmark-for-Big-Bench/wiki/HowToRunBigBench|BigBench] documentation web page.

You need to edit the conf/userSettings.conf file and do the changes according to your installation. E.g. Change the variable's content from

```
$ -export BIG_BENCH_HADOOP_CONF="/etc/hadoop/conf.cloudera.hdfs"
$ -export BIG_BENCH_HADOOP_LIBS_NATIVE="/opt/cloudera/parcels/CDH/lib/hadoop/lib/native"
```

To the below lines

```
$ +export BIG_BENCH_HADOOP_CONF="/usr/local/hadoop/etc/hadoop"
$ +export BIG_BENCH_HADOOP_LIBS_NATIVE="/usr/local/hadoop/lib/native"
```

A complete benchmark run with all stages, can be done by running (e.g., 4 map tasks, scale factor 100, 2 streams):

```
$ Big-Bench$ ./bin/bigBench runBenchmark -m 4 -f 100 -s 2
```

Issues

On Ubuntu, errors: "line 46: pssh: command not found".

Tried installing pssh package using the command

```
$ sudo apt-get install pssh
```

and created the below softlinks under /usr/bin/

```
parallel-ssh is pssh
parallel-scp is pscp
parallel-rsync is prsync
parallel-nuke is pnuke
parallel-slurp is pslurp
```

It looks like couple of error’s or issues need to be fixed
#1 Benchmark run terminated
Reason: An error occurred while running a command in phase CLEAN_ALL

#2 A warning

WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable