HP Enterprise

The Experience On Running Industry First Big Data Benchmark TPCx-BB

Paul Cao

May 11, 2016
Big Data Benchmark
TPCx-BB Agenda

✓ What system you used.
✓ How easy/difficult it was to set the benchmark up.
✓ How much time you spent tuning.
✓ Did you run into any issues.
✓ Assessment of the benchmark.
✓ How well does it measure the performance of your system.
Hardware Configuration

Two TPC Official Publications
## Configuration 1

**Hewlett Packard Enterprise**

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Total System Cost</th>
<th>TPCx-BB Performance Metric</th>
<th>Price/Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HP ProLiant DL Gen9</strong></td>
<td><strong>371,977 USD</strong></td>
<td><strong>337.26 BBqpm@3000</strong></td>
<td><strong>1,102.94 USD</strong>/$BBqpm@3000</td>
</tr>
</tbody>
</table>

### Framework
- Cloudera for Apache Hadoop (CDH) 5.6
- Red Hat Enterprise Linux Server 6.7

### Operating System
- None

### Other Software
- None

### Availability Date
- March 31, 2016

### Scale Factor
- 3000

### Streams
- 2

### System Configuration
- **9 Worker Nodes**
  - HPE ProLiant DL380 Gen9
    - 2x Intel E5-2697A v4 @ 2.60GHz
    - 1x HPE 480GB SSD
    - 1x HPE 1TB HDD
    - 1x HPE 1.2TB SSD
    - 192GB Memory

- **3 Management Nodes**
  - HPE ProLiant DL360 Gen9
    - 2x Intel E5-2640 v4 @ 2.40GHz
    - 1x HPE 800GB SSD
    - 256GB Memory

### Physical Storage/Scale Factor
- 53.84

### Scale Factor/Physical Memory
- 1.20

### Server Configuration
- **24/348/696**

### Per HPE ProLiant DL380 Gen9:
- **192GB**
- HPE Smart Array P440ar
- 1x HPE 480GB SSD
- 1x HPE 1TB HDD
- 1x HPE 1.2TB SSD
- HPE Ethernet 10Gb 2P 560FLR-SFP+ Adptr

### Per HPE ProLiant DL360 Gen9:
- **256GB**
- Smart HBA H240ar
- 1x HPE 800GB SSD
- HPE Ethernet 10Gb 2P 560FLR-SFP+ Adptr

### Connectivity
- HPE 1620-24G Switch, HPE 5900AF-48XG-4QSFP+ Switch
Configuration 2

Hewlett Packard Enterprise

Hewlett Packard Enterprise ProLiant DL for Big Data

<table>
<thead>
<tr>
<th>Total System Cost</th>
<th>TPCx-BB Performance Metric</th>
<th>Price/Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>325,156 USD</td>
<td>TPCx-BB Rev. 1.0.1</td>
<td>Report Date:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>March 23, 2016</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TPC-Pricing Rev. 1.7.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>265.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BBOpenM@3000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1,222.72 USD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$/BBOpenM@3000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Framework</th>
<th>Operating System</th>
<th>Other Software</th>
<th>Availability Date</th>
<th>Scale Factor</th>
<th>Streams</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloudera for</td>
<td>Red Hat</td>
<td>None</td>
<td>March 23, 2016</td>
<td>3000</td>
<td>2</td>
</tr>
<tr>
<td>Apache Hadoop (CDH) 5.6</td>
<td>Enterprise Linux Server 6.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

System Configuration

9 Worker Nodes
Each:
- HPE ProLiant DL380 Gen9
  - 2x Intel E5-2697 v3 @2.60GHz
  - 1x HPE 480GB SSD
  - 16x HPE 1TB HHD
  - 128GB Memory

3 Management Nodes
Each:
- HPE ProLiant DL360 Gen9
  - 2x Intel E5-2640 v3 @2.60GHz
  - 1x HPE 480GB SSD
  - 1x HPE 300GB SSD (1 node)
  - 128GB Memory

Software:
- Red Hat Enterprise Linux 6.7
- Cloudera Enterprise 5.6

Physical Storage/Scale Factor: 50:19
Scale Factor/Physical Memory: 1.95

Servers:
- Total Processors/Threads: 24/3069600
- HPE ProLiant DL 380 Gen 9
  - 3x HPE ProLiant DL380 Gen9
  - 9x HPE ProLiant DL 380 Gen9

Per HPE ProLiant DL 350 Gen 9:
- 2x Intel Xeon E5-2640 v3 @2.60GHz
- 128GB
- HPE Smart Array P940ar
- 1x HPE 480GB SSD
- 1x HPE 1TB HHD
- HPE Ethernet 10Gb 2P 560FLR-SFP+ Adapter

Per HPE ProLiant DL360 Gen9:
- 2x Intel Xeon E5-2640 v3 @2.60GHz
- 128GB
- HPE Smart Array P940ar
- 1x HPE 480GB SSD
- 1x HPE 300GB SSD (1 node)
- HPE Ethernet 10Gb 2P 560FLR-SFP+ Adapter

Connectivity:
- HPE 1620-24G Switch, HPE 5900AF-48XGT-4QSFP+ Switch
✓ The Challenge

✓ Multiple nodes (12 nodes x 2)
✓ Hardware setup
✓ Software setup
✓ Runtime
✓ Tuning
✓ Debugging (hardware, software)
✓ Audit
✓ FDR
How To

1. Get your cluster ready with HW/SW

2. Setup Passwordless ssh for pssh to work.

3. Edit "userSettings.conf" and insert your intended variables.
   - export BIG_BENCH_DEFAULT_DATABASE="bigbench" */ Name of Hive Metastore */
   - export BIG_BENCH_DEFAULT_ENGINE="hive" */ Default framework Engine */
   - export BIG_BENCH_DEFAULT_MAP_TASKS="80" */ Number of map tasks to generate data, read extended readme on selecting one */
   - export BIG_BENCH_DEFAULT_SCALE_FACTOR="10" */ Scale Factor you would like to test 1000=1TB, 3000=3TB, 10000=10TB, 300000=30TB, 1000000=1PB */
   - export BIG_BENCH_DEFAULT_NUMBER_OF_PARALLEL_STREAMS="2" */ Number of concurrent streams to use during throughput phase, default is 2 */
   - export BIG_BENCH_DEFAULT_BENCHMARK_PHASE="run_query" */ Adjust this to whatever distribution of Hadoop you are using */
   - export BIG_BENCH_HADOOP_CONF="/etc/hadoop/conf.cloudera.hdfs" */ Adjust this to whatever distribution of Hadoop you are using */
   - export BIG_BENCH_HADOOP_LIBS_NATIVE="/opt/cloudera/parcels/CDH/lib/hadoop/lib/native" */ Adjust this to whatever distribution of Hadoop you are using */
   - export BIG_BENCH_DATAGEN_DFS_REPLICATION="1" */ Use this setting to select number of replicas for your generated data on your HDFS file system, "1" means no replica, "3" is HDFS default of 3 copies. WARNING:* This setting has no bearing on default HDFS replication for all other files, which is set to 3" */
   - export BIG_BENCH_STOP_AFTER_FAILURE="1" */ the default behaviour is to stop when a query error occurs, set this to 0 to keep on running when an error occurs*/Challenges */
4. Run $/bin/TPCxBB_Verification.sh */ Engine Validation Phase on SF1“ */

5. Run $/TPCxBB_Benchmarkrun.sh */To execute the Benchmark, run twice to obtain "Performance" and "Repeatability" numbers. */

6. Copy 3 set of logs from $../logs 1. From Engine Validation Phase logs. 2. Performance run logs. 3. Repeatability run logs. E.g logfile from one run:logs-20151029-135147-hive-sf1xxx.zip

7. Engage the auditor and submit the report for publication.
Assessment of the Benchmark

- Stability
- Scaling
- Challenges
- Audit
System Response to the Workload

- Better than expected
HPE ProLiant DL Clusters
Intel Haswell vs Broadwell
on Red Hat RHEL 6

$1,212

27% performance gain

$1,103

9% better in price/perf

BBQpm@3000

HPE ProLiant DL Cluster
+ Intel Haswell
+ Red Hat RHEL 6

HPE ProLiant DL Cluster
+ Intel Broadwell
+ Red Hat RHEL 6

$/BBQpm

tpc.org/3501, tpc.org/3502
Q & A
TPC Big Bench CPU Utilization % 2 vs. 4 Streams

Cpu Utilization

Total 67%

Cpu Utilization

Total 83%
TPC Big Bench Memory Utilization

Memory Utilization

Memory Utilization

Hewlett Packard Enterprise
TPC Big Bench Disk Bandwidth

Disk Bandwidth

Disk Bandwidth

Axis Title

Axis Title

Sum of rkB/s

Sum of wkB/s

Sum of rkB/s

Sum of wkB/s
TPC Big Bench Context Switches

Context Switches

Axis Title

Average of wa
Average of sy
Average of us
Average of cs

Context Switches

Axis Title

Average of wa
Average of sy
Average of us
Average of cs