



Hewlett Packard
Enterprise

Hewlett Packard Enterprise Company

TPC Express Benchmark™ Big Bench (TPCx-BB)

Full Disclosure Report

for

Hewlett Packard Enterprise ProLiant DL for Big Data

(with 9x HPE ProLiant DL380 Gen9, 3x HPE ProLiant DL360 Gen9)

using

Cloudera for Apache Hadoop (CDH) 5.6

and

Red Hat Enterprise Linux Server 6.7

First Edition

March 23, 2016

Hewlett Packard Enterprise Company (HPE), the Sponsor of this benchmark test, believes that the information in this document is accurate as of the publication date. The information in this document is subject to change without notice. The Sponsor assumes no responsibility for any errors that may appear in this document.

The pricing information in this document is believed to accurately reflect the current prices as of the publication date. However, the Sponsor provides no warranty of the pricing information in this document.

Benchmark results are highly dependent upon workload, specific application requirements, and system design and implementation. Relative system performance will vary as a result of these and other factors. Therefore, the TPC_{Express} Benchmark™ BB should not be used as a substitute for a specific customer application benchmark when critical capacity planning and/or product evaluation decisions are contemplated.

All performance data contained in this report was obtained in a rigorously controlled environment. Results obtained in other operating environments may vary significantly. No warranty of system performance or price/performance is expressed or implied in this report.

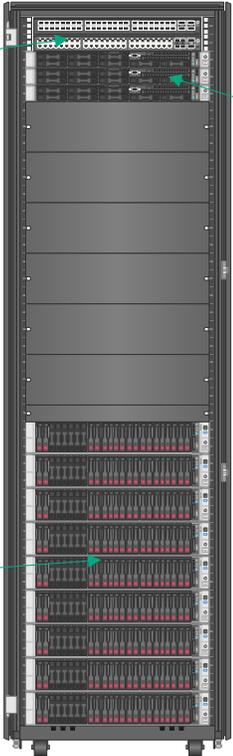
HPE and the HPE Logo are trademarks of Hewlett Packard Enterprise Company and/or its affiliates in the U.S. and other countries. Third party trademarks mentioned are the property of their respective owners. The use of the word partner does not imply a partnership relationship between HPE and any other company.

TPC Benchmark™, TPCx-BB and BBQpm, are registered certification marks of the Transaction Processing Performance Council.

The HPE products, services or features identified in this document may not yet be available or may not be available in all areas and may be subject to change without notice. Consult your local HPE business contact for information on the products or services available in your area. You can find additional information via HPE's web site at www.hpe.com. Actual performance and environmental costs of HPE products will vary depending on individual customer configurations and conditions.

Copyright © 2016 Hewlett Packard Enterprise Company

All rights reserved. Permission is hereby granted to reproduce this document in whole or in part provided the copyright notice printed above is set forth in full text or on the title page of each item reproduced.

 Hewlett Packard Enterprise		Hewlett Packard Enterprise ProLiant DL for Big Data			TPCx-BB Rev. 1.0.1 TPC-Pricing Rev. 1.7.0	
Total System Cost		TPCx-BB Performance Metric			Report Date: March 23, 2016	
322,428 USD		265.93 BBQpm@3000GB			1,212.46 USD \$/BBQpm@3000GB	
Framework	Operating System	Other Software	Availability Date	Scale Factor	Streams	
Cloudera for Apache Hadoop (CDH) 5.6	Red Hat Enterprise Linux Server 6.7	None	March 23, 2016	3000GB	2	
System Configuration						
<p>Ethernet Switch: HPE 1620-24G Switch (ILO connection) HP5900AF-48XGT-4QSFP+ (main connection)</p>					<p>3 Management Nodes Each: HPE ProLiant DL360 Gen9 2x Intel E5-2640 v3 2.60GHz 1x HPE 480GB SSD 1x HPE 800GB SSD (1 node) 128GB Memory</p>	
<p>9 Worker Nodes Each: HPE ProLiant DL380 Gen9 2x Intel E5-2697 v3 2.60GHz 1x HPE 480GB SSD 16x HPE 1TB HDD 128GB Memory</p>					<p>Software: Red Hat Enterprise Linux 6.7 Cloudera Enterprise 5.6</p>	
Physical Storage/Scale Factor: 50.19			Scale Factor/Physical Memory: 1.95			
Servers:		9x HPE ProLiant DL 380 Gen 9, 3x HPE ProLiant DL360 Gen9				
Total Processors/Cores/Threads		24/300/600				
Server Configuration:	Per HPE ProLiant DL 380 Gen 9:		Per HPE ProLiant DL360 Gen9:			
Processors	2x Intel Xeon E5-2697 v3 @ 2.60GHz		2x Intel Xeon E5-2640 v3 @ 2.60GHz			
Memory	128GB		128GB			
Storage Controller	HPE Smart Array P840/4G		Smart HBA H240ar			
Storage Device	1x HPE 480GB SSD 16x HPE 1 TB HDD		1x HPE 480GB SSD 1x HPE 800GB SSD (1 node)			
Network	HPE Ethernet 10Gb 2P 560FLR-SFP+ Adptr		HPE Ethernet 10Gb 2P 560FLR-SFP+ Adptr			
Connectivity:	HPE 1620-24G Switch, HPE 5900AF-48XG-4QSFP+ Switch					



**Hewlett Packard
Enterprise**

**Hewlett Packard Enterprise
ProLiant DL for Big Data**

TPCx-BB Rev. 1.0.1
TPC-Pricing Rev. 1.7.0

Report Date:
March 23, 2016

Description	Price Key	Part Number	Unit Price	Qty	Extended Price	3 Yr Maint Price
Server Hardware						
HPE DL360 Gen9 8SFF CTO Server	1	755258-B21	\$1,797	3	\$5,391	
HPE DL360 Gen9 E5-2640v3 Kit	1	755386-B21	\$1,329	3	\$3,987	
HPE DL360 Gen9 E5-2640v3 FIO Kit	1	755386-L21	\$1,329	3	\$3,987	
HPE 16GB 2Rx4 PC4-2133P-R Kit	1	726719-B21	\$379	24	\$9,096	
HP H240ar 12Gb 2-ports Int FIO Smart Host Bus Adapter	1	749976-B21	\$249	3	\$747	
HPE 500W FS Plat Ht Plg Pwr Supply Kit	1	720478-B21	\$309	6	\$1,854	
HPE 480GB 6G SATA MU-2 SFF SC SSD	1	832414-B21	\$1,189	3	\$3,567	
HPE 800GB 6G SATA MU-2 SFF SC SSD	1	804671-B21	\$3,038	1	\$3,038	
HPE Ethernet 10Gb 2P 560FLR-SFP+ Adptr	1	665243-B21	\$679	3	\$2,037	
HPE 3Y FC 24x7 DL360 Gen9 SVC	1	U7AL9E	\$1,404	3		\$4,212
HPE iLO Adv incl 3yr TS U E-LTU	1	E6U64ABE	\$469	3		\$1,407
HP W1972a 18.5-In LED Monitor (1 + 2 spare)	1	B7M13A8#	\$79	3	\$238	
HP PS/2 Keyboard And Mouse Bundle (1 + 2 spare)	1	B1T13AA#	\$27	3	\$82	
			Subtotal		\$34,023	\$5,619
HPE DL380 Gen9 24SFF CTO Server	1	767032-B21	\$2,107	9	\$18,963	
HPE DL380 Gen9 High Perf Fan Kit	1	719079-B21	\$239	9	\$2,151	
HPE DL380 Gen9 2SFF Bay Kit	1	724864-B21	\$149	9	\$1,341	
HPE DL380 Gen9 Secondary Riser	1	719073-B21	\$99	9	\$891	
HPE DL380 Gen9 E5-2697v3 Kit	1	719055-B21	\$2,599	9	\$23,391	
HPE DL380 Gen9 E5-2697v3 Kit	1	719055-L21	\$2,599	9	\$23,391	
HPE 16GB 2Rx4 PC4-2133P-R Kit	1	726719-B21	\$379	72	\$27,288	
HPE 1TB 6G SATA 7.2k 2.5in SC MDL HDD	1	655710-B21	\$499	144	\$71,856	
HPE 480GB 6G SATA MU-2 SFF SC SSD	1	832414-B21	\$1,189	9	\$10,701	
HPE Smart Array P840/4G Controller	1	726897-B21	\$1,249	9	\$11,241	
HPE 12Gb DL380 Gen9 SAS Expander Card	1	727250-B21	\$699	9	\$6,291	
HPE 800W FS Ti Ht Plg Pwr Supply Kit	1	720482-B21	\$409	18	\$7,362	
HPE Ethernet 10Gb 2P 560FLR-SFP+ Adptr	1	665243-B21	\$679	9	\$6,111	
HPE 3Y FC 24x7 DL380 Gen9 SVC	1	U7AE5E	\$1,821	9		\$16,389
HPE iLO Adv incl 3yr TS U E-LTU	1	E6U64ABE	\$469	9		\$4,221
			Subtotal		\$210,978	\$20,610
Network						
HPE 1620-24G Switch	1	JG913A	\$299	1	\$299	
HPE 5900AF-48XG-4QSFP+ Switch	1	JC772A	\$20,990	1	\$20,990	
HPE 1m Multi-mode OM3 LC/LC FC Cable	1	AJ834A	\$70	12	\$840	
HPE A58x0AF 650W AC Power Supply	1	JC680A	\$749	4	\$2,996	
HPE 58x0AF Bck(pwr)-Frt(ports) Fan Tray	1	JC682A	\$179	4	\$716	
			Subtotal		\$25,841	\$0
Rack						
HPE Universal Rack 11642 1075mm Shock Rack	1	H6J66A	\$1,699	1	\$1,699	
HPE 24A High Voltage Core Only Corded PDU	1	252663-D74	\$259	2	\$518	
			Subtotal		\$2,217	\$0



**Hewlett Packard
Enterprise**

**Hewlett Packard Enterprise
ProLiant DL for Big Data**

TPCx-BB Rev. 1.0.1
TPC-Pricing Rev. 1.7.0

Report Date:
March 23, 2016

Description	Price Key	Part Number	Unit Price	Qty	Extended Price	3 Yr Maint Price
Server Software						
Cloudera Ent Basic Ed 1yr 24x7	1	G7M27A	\$2,304	36	\$82,944	
RHEL Svr 2 Sckt/2 Gst 3yr 24x7 E-LTU	1	G3J30AAE	\$3,889	12	\$46,668	
Subtotal					\$129,612	\$0
Total Extended Price					\$402,671	\$26,229
Total Discounts					\$105,207	\$1,266
Sales contact: HPE WW Headquarters, 3000 Hanover St., Palo Alto, CA 94304-1185 (650) 857-1501 or H-P direct: 800-203-6748			Grand Total		\$297,464	\$24,963
Pricing: 1 = HPE ⁽¹⁾ All discounts are based on US list prices and for similar quantities and configurations. The discounts are based on the overall specific components pricing from respective vendors in this single quotation. Discounts for similarly sized configurations will be similar to those quoted here, but may vary based on the components in the configuration. Audited by Doug Johnson of InfoSizing, Inc.				Three-Year Cost of Ownership \$322,428 BBQpm@3000GB 265.93 \$/BBQpm@3000GB \$1,212.46		
Prices used in TPC benchmarks reflect the actual prices a customer would pay for a one-time purchase of the stated components. Individually negotiated discounts are not permitted. Special prices based on assumptions about past or future purchases are not permitted. All discounts reflect standard pricing policies for the listed components. For complete details, see the pricing sections of the TPC benchmark specifications. If you find that the stated prices are not available according to these terms, please inform at pricing@tpc.org . Thank you.						



Numerical Quantities

Scale Factor	3000GB
Streams	2
SUT Validation Status	PASS

Performance Run

Overall Run Start Time	2016-03-12 13:05:16.670
Overall Run End Time	2016-03-13 10:37:51.892
Overall Run Elapsed Time	73,955.222

Load Test Start Time	2016-03-12 13:05:16.671
Load Test End Time	2016-03-12 13:41:59.708
Load Test Elapsed Time	2,203.037

Power Test Start Time	2016-03-12 13:41:59.710
Power Test End Time	2016-03-12 21:20:59.935
Power Test Elapsed Time	27,540.225

Throughput Test Start Time	2016-03-12 21:20:59.936
Throughput Test End Time	2016-03-13 10:37:51.892
Throughput Test Elapsed Time	44,211.956

Performance Metric (BBQpm@SF)	265.93
-------------------------------	--------

Repeatability Run

Overall Run Start Time	2016-03-11 16:44:52.574
Overall Run End Time	2016-03-12 12:39:44.000
Overall Run Elapsed Time	71,691.426

Load Test Start Time	2016-03-11 16:44:52.575
Load Test End Time	2016-03-11 17:21:38.198
Load Test Elapsed Time	2,205.623

Power Test Start Time	2016-03-11 17:21:38.200
Power Test End Time	2016-03-12 00:55:08.677
Power Test Elapsed Time	27,210.477

Throughput Test Start Time	2016-03-12 00:55:08.678
Throughput Test End Time	2016-03-12 12:39:43.999
Throughput Test Elapsed Time	42,275.321

Performance Metric (BBQpm@SF)	272.99
-------------------------------	--------



Run Report – Run 1

```
*****
TPCx-BB
Result
v1.0.1
*****
INFO: T_LOAD = 2205.623
INFO: T_LD = 0.1 * T_LOAD: 220.56230000000002
INFO: T_PT = 18099.997606264762
INFO: T_T_PUT = 42275.321
INFO: T_TT = 21137.6605
INFO: === Checking validity of the final result ===
INFO: OK: All required BigBench phases were performed.
INFO: OK: All 30 queries were running in the power test.
INFO: OK: All 30 queries were running in the first throughput test.
INFO: OK: Pretend mode was inactive. All commands were executed.
INFO: === Final result ===
INFO: VALID BBQpm@3000 = 272.99597675236356
```

Run Report – Run 2

```
*****
TPCx-BB
Result
v1.0.1
*****
INFO: T_LOAD = 2203.037
INFO: T_LD = 0.1 * T_LOAD: 220.3037
INFO: T_PT = 18249.03919370655
INFO: T_T_PUT = 44211.956
INFO: T_TT = 22105.978
INFO: === Checking validity of the final result ===
INFO: OK: All required BigBench phases were performed.
INFO: OK: All 30 queries were running in the power test.
INFO: OK: All 30 queries were running in the first throughput test.
INFO: OK: Pretend mode was inactive. All commands were executed.
INFO: === Final result ===
INFO: VALID BBQpm@3000 = 265.9385341586281
```

Summary details of the run reports are shown above. For the complete run reports, see the Support Files Archive.

Table of Contents

ABSTRACT	9
PREFACE	10
CLAUSE 1: GENERAL ITEMS.....	11
1.1 TEST SPONSOR.....	11
1.2 PARAMETER SETTINGS	11
1.3 CONFIGURATION DIAGRAMS.....	11
CLAUSE 2: SOFTWARE COMPONENTS AND DATASET DISTRIBUTION	13
2.1 ROLES AND DATASET DISTRIBUTION.....	13
2.2 DISTRIBUTED FILE SYSTEM IMPLEMENTATION	13
2.3 ENGINE IMPLEMENTATION	14
2.4 FRAMEWORKS	14
2.5 APPLIED PATCHES	14
CLAUSE 3: WORKLOAD RELATED ITEMS	15
3.1 HARDWARE & SOFTWARE TUNABLE	15
3.2 KIT VERSION	15
3.3 RUN REPORT.....	15
3.4 QUERY ELAPSED TIMES	16
3.5 VALIDATION TEST OUTPUT.....	16
3.6 GLOBAL FRAMEWORK PARAMETERS.....	17
3.7 KIT MODIFICATIONS.....	17
CLAUSE 4: SUT RELATED ITEMS.....	18
4.1 SPECIALIZED HARDWARE/SOFTWARE	18
4.2 FRAMEWORK CONFIGURATION FILES	18
4.3 SUT ENVIRONMENT INFORMATION	18
4.4 DATA STORAGE TO SCALE FACTOR RATIO.....	18
4.5 SCALE FACTOR TO MEMORY RATIO	18
CLAUSE 5: METRICS AND SCALE FACTORS.....	19
5.1 PERFORMANCE RUN METRIC	19
5.2 REPEATABILITY RUN METRIC.....	19
5.3 PRICE-PERFORMANCE METRIC	19
5.4 SCALE FACTOR.....	19
5.5 STREAM COUNT.....	19
5.6 ELAPSED RUN TIMES.....	19
5.7 ELAPSED TEST TIMES.....	19
AUDITORS' INFORMATION AND ATTESTATION LETTER.....	21
SUPPORTING FILE INDEX.....	24

Abstract

This document contains the methodology and results of the TPC Express Benchmark™ Big Bench (TPCx-BB) test conducted in conformance with the requirements of the TPCx-BB Standard Specification, Revision 1.0.1.

The test was conducted at a Scale Factor of 3000GB with 12 9x HPE ProLiant DL380 Gen9, 3x HPE ProLiant DL360 Gen9 running Cloudera for Apache Hadoop (CDH) 5.6 on Red Hat Enterprise Linux Server 6.7.

Measured Configuration

Company Name	Cluster Node	Virtualization	Operating System
Hewlett Packard Enterprise Company	9x HPE ProLiant DL380 Gen9, 3x HPE ProLiant DL360 Gen9	n/a	Red Hat Enterprise Linux Server 6.7

TPC Express Benchmark© Big Bench Metrics

Total System Cost	BBQpm@3000GB	Price/Performance	Availability Date
322,428 USD	265.93	1,212.46 USD	March 23, 2016

Preface

TPC Express Benchmark™ Big Bench Overview

*Big data analytics is a growing field of research and business. The significant decrease in the overall cost of hardware, the emergence of Open Source based analytics frameworks, along with the greater depth of data mining capabilities allows new types of data sources to be correlated with traditional data sources. For example, online retailers used to record only successful transactions on their website, whereas modern systems are capable of recording every interaction. The former allowed for simple shopping basket analysis techniques, while the current level of detail in monitoring makes detailed user modeling possible. The growing demands on data management systems and the new forms of analysis have led to the development of a new type of **Big Data Analytics Systems (BDAS)**.*

*Similar to the advent of **Database Management Systems**, there is a vastly growing ecosystem of diverse approaches to enabling Big Data Analytics Systems. This leads to a dilemma for customers of **BDAS**, as there are no realistic and proven measures to compare different **BDAS** solutions. To address this, TPC has developed TPCx-BB (BigBench), which is an express benchmark for comparing **BDAS** solutions. The TPCx-BB Benchmark was developed to cover essential functional and business aspects of big data use cases. The benchmark allows for an objective measurement of **BDAS** System under Test, and provides the industry with verifiable performance, price/performance, and availability metrics.*

*The TPCx-BB kit is available from the TPC website (see www.tpc.org for more information). Users must sign-up and agree to the TPCx-BB End User Licensing Agreement (EULA) to download the kit. All related work (such as collaterals, papers, derivatives) must acknowledge the TPC and include the TPCx-BB copyright. The TPCx-BB kit includes: TPCx-BB Specification document (this document), TPCx-BB Users Guide documentation, shell scripts to set up the benchmark environment, Java code to execute the benchmark workload, Data Generator, **Query** files, and Benchmark Driver.*

The purpose of TPC benchmarks is to provide relevant, objective performance data to industry users. To achieve that purpose, TPC benchmark specifications require that benchmark tests be implemented with systems, products, technologies and pricing that:

- *Are generally available to users;*
- *Are relevant to the market segment that the individual TPC benchmark models or represents (e.g., TPCx-BB models and represents a Big Data Analytics System such as Hadoop ecosystem or Hadoop File-system API compatible systems);*
- *Would plausibly be implemented by a significant number of users in the market segment the benchmark models or represents.*

The use of new systems, products, technologies (hardware or software) and pricing is encouraged so long as they meet the requirements above. Specifically prohibited are benchmark systems, products, technologies or pricing (hereafter referred to as "implementations") whose primary purpose is performance optimization of TPC benchmark results without any corresponding applicability to real-world applications and environments. In other words, all "benchmark special" implementations that improve benchmark results but not real-world performance or pricing, are prohibited.

The rules for pricing are included in the TPC Pricing Specification and rules for energy measurement are included in the TPC Energy Specification.

Further information is available at www.tpc.org

Clause 1: General Items

1.1 Test Sponsor

A statement identifying the benchmark sponsor(s) and other participating companies must be provided.

This benchmark was sponsored by Hewlett Packard Enterprise Company

1.2 Parameter Settings

Settings must be provided for all customer-tunable parameters and options which have been changed from the defaults found in actual products, including by not limited to:

- *Configuration parameters and options for server, storage, network and other hardware components used by the SUT.*
- *Configuration parameters and options for Operating System and file system components used by the SUT.*
- *Configuration parameters and options for any other software components (e.g compiler optimization options) used by the SUT.*

Comment 1: In the event that some parameters and options are set multiple times, it must be easily discernible by an interested reader when the parameter or option was modified and what new value it received each time.

Comment 2: This requirement can be satisfied by providing a full list of all parameters and options, as long as all those that have been modified from their default values have been clearly identified and these parameters and options are only set once.

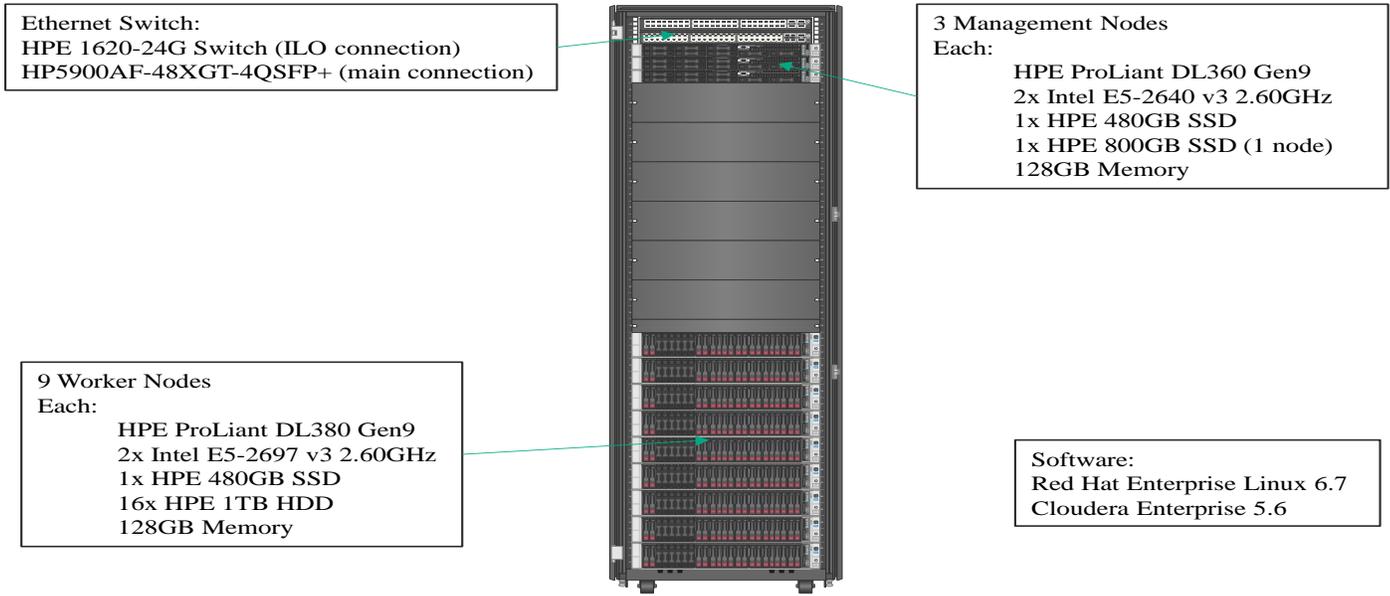
The Supporting Files Archive contains the parameters and options used to configure the components involved in this benchmark.

1.3 Configuration Diagrams

7.4.4 Diagrams of both measured and priced configurations must be provided, accompanied by a description of the differences. This includes, but is not limited to:

- *Total number of nodes used;*
- *Total number and type of processors used/total number of cores used/total number of threads used (including sizes of L2 and L3 caches);*
- *Size of allocated memory, and any specific mapping/partitioning of memory unique to the test;*
- *Number and type of disk units (and controllers, if applicable);*
- *Number of channels or bus connections to disk units, including their protocol type;*
- *Number of LAN (e.g., Ethernet) connections and speed for switches and other hardware components physically used in the test or are incorporated into the pricing structure;*
- *Type and the run-time execution location of software components.*

Measured Configuration



The measured configuration consisted of:

- Total Nodes: 12
- Total Processors/Cores/Threads: 24/300/600
- Total Memory: 1,536GB
- Total Number of Storage Drives/Devices: 157
- Total Storage Capacity: 150,560GB

Network connectivity detail:

- 2 x Cisco UCS 6296UP 96-Port Fabric Interconnect

Server nodes details:

9x HPE ProLiant DL380 Gen9, each with:

- Processors/Cores/Threads: 2/28/56
- Processor Model: 2x Intel Xeon E5-2697 v3 @ 2.60GHz
- Memory: 128GB
- Controller: 1 x HPE Smart Array P840/4G
- Drives:
 - 1 x HPE 480GB SSD
 - 16 x HPE 1TB HDD
- Network: HPE Ethernet 10Gb 2P 560FLR-SFP+ Adptr

3x HPE ProLiant DL360 Gen9, each with:

- Processors/Cores/Threads: 2/16/32
- Processor Model: 2 x Intel Xeon E5-2640 v3 2.60GHz
- Memory: 128GB
- Controller: Smart HBA H240ar
- Drives:
 - 1 x HPE 480GB SSD
 - 1 x HPE 800GB SSD (1 node)
- Network: HPE Ethernet 10Gb 2P 560FLR-SFP+ Adptr

The distribution of software components over server nodes is detailed in section 2.1.

Priced Configuration

There are no differences between the priced and measured configurations.

Clause 2: Software Components and Dataset Distribution

2.1 Roles and Dataset Distribution

The distribution of dataset across all media must be explicitly described.

The distribution of various software components across the system must be explicitly described.

Table 1.4 describes the distribution of the dataset across all media in the system.

Table 1.4: Software Components and Dataset Distribution

Server	Role(s)	Count	Virtual	Host Name(s)	HW/SW Configuration	Storage Setup
Worker	HDFS DataNode/Hive Gateway/YARN Node Manager	9	N	hsw[04-12]	<ul style="list-style-type: none"> HPE DL380 Gen9. HW/SW Config (Intel E5-2697v3, 2, 2.6GHz, 28) Memory: 128GB Storage: 16 x 1TB SATA HDD, 1 x 480GB SSD Network: HPE 560 SFP+10G NIC OS: RHEL 6.7 Cloudera CDH 5.6 	OS: HPE 480GB 6G SATA SSD, Intermediate/Shuffle/Temp Data/Distributed FS: x HPE 1TB 6G SATA 7.2k HDD
Cloudera Manager Node #1	Cloudera Manager/HDFS Balancer/HDFS Namenode/Hive Gateway/Hive Metastore Server/Hue Server/Cloudera Management Services/Oozie Server/YARN JobHistory Server/YARN ResourceManager/ZooKeeper Server	1	N	hsw01	<ul style="list-style-type: none"> HPE DL360 Gen9 Server HW/SW Config (Intel E5-2640v3, 2, 2.6GHz, 16) Memory: 128GB Storage: 1 x 480GB SSD Network: HPE 560 SFP+10G NIC OS: RHEL 6.7 Cloudera CDH 5.6 	OS: HPE 480GB 6G SATA SSD
Cloudera Manager Node #2	Hive Gateway/HiveServer2/ZooKeeper Server	1	N	hsw02	<ul style="list-style-type: none"> HPE DL360 Gen9 Server HW/SW Config (Intel E5-2640v3, 2, 2.6GHz, 16) Memory: 128GB Storage: 1 x 480GB SSD Network: HPE 560 SFP+10G NIC OS: RHEL 6.7 Cloudera CDH 5.6 	OS: HPE 480GB 6G SATA SSD
Cloudera Manager Node #3	HDFS SecondaryNameNode/Hive Gateway/Cloudera Management Service Activity Monitor/ZooKeeper Server	1	N	hsw03	<ul style="list-style-type: none"> HPE DL360 Gen9 Server HW/SW Config (Intel E5-2640v3, 2, 2.6GHz, 16) Memory: 128GB Storage: 1 x 480GB SSD Network: HPE 560 SFP+10G NIC OS: RHEL 6.7 Cloudera CDH 5.6 	OS: HPE 480GB 6G SATA SSD

2.2 Distributed File System Implementation

Distributed file system implementation and corresponding Hadoop File System API version must be disclosed.

Cloudera for Apache Hadoop (CDH) 5.6 (fully HDFS compatible at the API level).

2.3 Engine Implementation

The Engine implementation and corresponding version must be disclosed.

Component	Version
Hive	1.1.0
HDFS	2.6.0
YARN	2.6.0
Spark	1.5.0
MapReduce	2.6.0

2.4 Frameworks

Frameworks and Engine used in the benchmark should be disclosed.

Framework	Version
CDH	5.6.0
Hive	1.1.0
HDFS	2.6.0
YARN	2.6.0
Spark	1.5.0
MapReduce	2.6.0

2.5 Applied Patches

Any additional vendor supported patches applied to the SUT should be disclosed.

No additional patches were applied.

Clause 3: Workload Related Items

3.1 Hardware & Software Tunable

Script or text used to set for all hardware and software tunable parameters must be reported.

The Supporting Files Archive contains all configuration scripts.

3.2 Kit Version

Version number of the TPCx-BB kit must be included in the Report.

TPCx-BB Kit Version
1.0.1

3.3 Run Report

The run report generated by TPCx-BB benchmark kit must be included in the Report.

The Supporting File Archive contains the full run report. Following are summary extracts from both runs.

- **Run1 Performance Summary**

```
*****
TPCx-BB
Result
v1.0.1
*****
INFO: T_LOAD = 2205.623
INFO: T_LD = 0.1 * T_LOAD: 220.56230000000002
INFO: T_PT = 18099.997606264762
INFO: T_T_PUT = 42275.321
INFO: T_TT = 21137.6605
INFO: === Checking validity of the final result ===
INFO: OK: All required BigBench phases were performed.
INFO: OK: All 30 queries were running in the power test.
INFO: OK: All 30 queries were running in the first throughput test.
INFO: OK: Pretend mode was inactive. All commands were executed.
INFO: === Final result ===
INFO: VALID BBQpm@3000 = 272.99597675236356
```

- **Run2 Performance Summary**

```
*****
TPCx-BB
Result
v1.0.1
*****
INFO: T_LOAD = 2203.037
INFO: T_LD = 0.1 * T_LOAD: 220.3037
INFO: T_PT = 18249.03919370655
INFO: T_T_PUT = 44211.956
INFO: T_TT = 22105.978
INFO: === Checking validity of the final result ===
INFO: OK: All required BigBench phases were performed.
INFO: OK: All 30 queries were running in the power test.
INFO: OK: All 30 queries were running in the first throughput test.
INFO: OK: Pretend mode was inactive. All commands were executed.
INFO: === Final result ===
INFO: VALID BBQpm@3000 = 265.9385341586281
```

3.4 Query Elapsed Times

Elapsed times of all power and throughput Queries needs to be reported from the Performance Run, grouped respectively as Structured, semi-structured and unstructured buckets.

Query Type	Query Number	Power	Throughput	
		Stream 1	Stream 1	Stream 2
Structured	1	263.338	372.045	256.282
	6	725.020	1,692.274	801.858
	7	441.061	477.812	984.890
	9	2,801.596	7,379.498	5,157.169
	11	211.320	510.809	391.876
	13	359.765	377.606	932.580
	14	90.522	108.926	104.932
	15	127.301	260.189	177.258
	16	979.058	970.918	1,131.366
	17	323.898	485.679	885.398
	20	380.010	467.266	541.543
	21	566.188	651.425	662.291
	22	727.787	1,033.369	990.221
	23	591.306	1,292.198	771.677
Semi-structured	2	2,039.890	3,299.515	3,022.936
	3	1,019.385	1,229.323	1,039.372
	4	1,945.953	3,087.133	7,191.610
	5	2,009.083	3,990.962	2,453.640
	8	745.367	846.018	944.944
	12	602.007	1,146.786	715.773
	30	3,171.266	3,136.045	4,234.199
Unstructured	10	1,455.198	1,947.524	1,467.504
	18	3,024.994	3,203.121	4,375.076
	19	489.275	791.818	495.390
	27	108.838	1,531.010	109.166
	28	491.644	542.589	504.155

3.5 Validation Test Output

Output report from successful SUT Validation test must be included in the Report.

Query Number	Execution Successful	Validation Successful
1	TRUE	TRUE
2	TRUE	TRUE
3	TRUE	TRUE
4	TRUE	TRUE
5	TRUE	TRUE
6	TRUE	TRUE
7	TRUE	TRUE
8	TRUE	TRUE
9	TRUE	TRUE
10	TRUE	TRUE
11	TRUE	TRUE
12	TRUE	TRUE
13	TRUE	TRUE
14	TRUE	TRUE
15	TRUE	TRUE
16	TRUE	TRUE
17	TRUE	TRUE
18	TRUE	TRUE
19	TRUE	TRUE
20	TRUE	TRUE
21	TRUE	TRUE
22	TRUE	TRUE
23	TRUE	TRUE
24	TRUE	TRUE
25	TRUE	TRUE
26	TRUE	TRUE
27	TRUE	TRUE
28	TRUE	TRUE
29	TRUE	TRUE
30	TRUE	TRUE

3.6 Global Framework Parameters

Global Framework parameter settings files must be included in the Report.

The Supporting File Archive contains the global framework parameter settings files.

3.7 Kit Modifications

Test Sponsor kit modifications files must be included in the Report..

The following files were modified by the Test Sponsor to facilitate system, platform and Framework differences.

- bigBench-configs/conf/userSettings.conf
- bigBench-configs/hive/queries/q28/engineLocalSettings.sql

Clause 4: SUT Related Items

4.1 Specialized Hardware/Software

Specialized Hardware/Software used in the SUT must be included.

No specialized hardware or software was used.

4.2 Framework Configuration Files

All Framework configuration files from SUT, for the performance run.

All Framework configuration files are included in the Supporting Files Archive.

4.3 SUT Environment Information

SUT environment info in form of envinfo.log from a representative worker node from every role in the server.

All envinfo.log files are include in the Supporting Files Archive.

4.4 Data Storage to Scale Factor Ratio

The data storage ratio must be disclosed.

Nodes	Disks	Size (GB)	Total (GB)
3	1	480	1,440
1	1	800	800
9	1	480	4,320
9	16	1,000	144,000

Total Storage (GB)	150,560
Scale Factor	3,000
Data Storage Ratio	50.19

4.5 Scale Factor to Memory Ratio

The Scale Factor to memory ratio must be disclosed.

Nodes	Memory (GB)	Total (GB)
12	128	1,536

Scale Factor	3,000
Total Memory (GB)	1,536
SF / Memory Ratio	1.95

Clause 5: Metrics and Scale Factors

5.1 Performance Run Metric

The Reported Performance Metric (BBQpm@SF for the Performance Run) must be disclosed in the Report.

Performance Run	
BBQpm@3000	265.93

5.2 Repeatability Run Metric

The Performance Metric (BBQpm@SF) for the Repeatability Run must be disclosed in the Report..

Repeatability Run	
BBQpm@3000	272.99

5.3 Price-Performance Metric

The Reported Performance Metric (BBQpm@SF for the Performance Run) must be disclosed in the Report.

Price-Performance	
\$/BBQpm@3000	\$1,212.46

5.4 Scale Factor

The Scale Factor used for the Result must be disclosed in the Report.

Scale Factor
3000

5.5 Stream Count

The number of streams in the throughput run used for the Result must be disclosed in the Report.

Streams
2

5.6 Elapsed Run Times

The total elapsed time for the execution of the Performance Run and Repeatability Run must be disclosed in the Report.

Run	Elapsed Time	Seconds
Run 1	19:54:51.426	71,691.426
Run 2	20:32:35.222	73,955.222

5.7 Elapsed Test Times

The total time for each of the three tests must be disclosed for the Performance Run and the Repeatability Run.

Test	Performance Run	Repeatability Run
Load Test	2,203.037	2,205.623
Power Test	27,540.225	27,210.477
Throughput Test	44,211.956	42,275.321

Auditors' Information and Attestation Letter

The auditor's agency name, address, phone number, and Attestation letter must be included in the full disclosure report. A statement should be included specifying who to contact in order to obtain further information regarding the audit process.

This benchmark was audited by Doug Johnson for InfoSizing, Inc.

www.sizing.com
20 Kreg Lane
Manitou Springs, CO 80829
719-473-7555.

This benchmark's Full Disclosure Report (FDR) can be downloaded from www.tpc.org.

A copy of the auditor's attestation letter is included in the next two pages.



The Right Metric For Sizing IT

TPC Transaction Processing Performance Council

Certified Auditors



Paul Cao
Hewlett Packard Enterprise
11445 Compaq Center Dr West
Houston, TX 77070



March 23, 2016

I verified the TPC Express Benchmark™ BB v1.0.1 performance of the following configuration:

Platform: Hewlett Packard Enterprise ProLiant DL for Big Data
(with 9x HPE ProLiant DL380 Gen9 and 3x HPE ProLiant DL360 Gen9 Servers)
Operating System: Red Hat Enterprise Linux Server 6.7
Apache Hadoop Cloudera for Apache Hadoop (CDH) 5.6
Compatible Software:

The results were:

Performance Metric 265.93 BBQpm@3000GB

Run Elapsed Time 73,955.222 Seconds

Cluster 9x HPE DL380 Gen9, 3x HPE DL360 Gen9 Servers

CPU	DL380s	2 x Intel Xeon Processor E5-2697 v3 (2.60 GHz, 14-core, 35 MB L3)		
	DL360s	2 x Intel Xeon Processor E5-2640 v3 (2.60 GHz, 8-core, 20 MB L3)		
Memory		128 GB (all nodes)		
Storage		Qty	Size	Type
		1	480GB	SATA SSD (all nodes)
		16	1TB	7.2K rpm SATA HDD (DL380 nodes)
		1	800GB	SATA SSD (1 DL360 node)

In my opinion, these performance results were produced in compliance with the TPC requirements for the benchmark.

The following verification items were given special attention:

- All TPC-provided components were verified to be v1.0.1
- No modifications were made to any of the Java code
- Any and all modifications to shell scripts were reviewed for compliance
- The tested Scale Factor (3000GB) was confirmed to be valid for publication
- All validation queries executed successfully and produced compliant results

20 KREG LANE • MANITOU SPRINGS, CO 80829 • 719-473-7555 • WWW.SIZING.COM

- No errors were reported during the run
- The elapsed times for all phases and runs were correctly measured and reported
- The Storage and Memory Ratios were correctly calculated and reported
- The system pricing was verified for major components and maintenance
- The major pages from the FDR were verified for accuracy

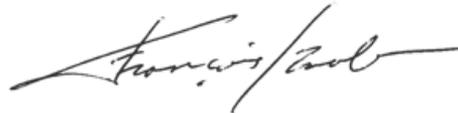
Additional Audit Notes:

None.

Respectfully Yours,



Doug Johnson, Auditor



François Raab, President

Supporting File Index

The following index outlines the information included in the supporting files archive.

Description	Archive File Pathname
Clause 1 - General Items	
The Supporting Files Archive contains the parameters and options used to configure the components involved in this benchmark	Supporting-Files-3TB-HSW-3-2016\
Validation Run Files	Supporting-Files-3TB-HSW-3-2016\Validation-run-logs-20160311-162432-hive-sf3000
Performance Run Files	Supporting-Files-3TB-HSW-3-2016\logs-20160313-104031-hive-sf3000
Repeatability Run Files	Supporting-Files-3TB-HSW-3-2016\Repeatability-run-logs-20160312-124215-hive-sf3000
Clause 3 - Workload Related Items	
Benchmark Generic Parameters	Supporting-Files-3TB-HSW-3-2016\logs-20160313-104031-hive-sf3000\bigBench-configs\conf\userSettings.conf
Query Parameters used in the benchmark execution Settings	Supporting-Files-3TB-HSW-3-2016\logs-20160313-104031-hive-sf3000\bigBench-configs\hive\conf\queryParameters.sql
Benchmark Global Framework Parameters Settings	Supporting-Files-3TB-HSW-3-2016\logs-20160313-104031-hive-sf3000\bigBench-configs\hive\conf\engineSettings.sql
Benchmark Global Framework Parameters Settings	Supporting-Files-3TB-HSW-3-2016\logs-20160313-104031-hive-sf3000\bigBench-configs\hive\conf\engineSettings.conf
Load Test script	Supporting-Files-3TB-HSW-3-2016\logs-20160313-104031-hive-sf3000\bigBench-configs\hive\population\hiveCreateLoad.sql
Queries specific optimization parameters settings	Supporting-Files-3TB-HSW-3-2016\logs-20160313-104031-hive-sf3000\bigBench-configs\hive\queries\q[01-30]\engineLocalSettings.conf
Queries specific optimization parameters settings	Supporting-Files-3TB-HSW-3-2016\logs-20160313-104031-hive-sf3000\bigBench-configs\hive\queries\q[01-30]\engineLocalSettings.sql
Clause 4 - SUT Related Items	
Data Redundancy report	Supporting-Files-3TB-HSW-3-2016\hdfs-data-redundancy-report.txt
Hardware and Software Report from a representative node	Supporting-Files-3TB-HSW-3-2016\logs-20160313-104031-hive-sf3000\run-logs\envInfo-hsw04\envInfo.log
All Framework configuration files are included in the Supporting Files Archive	Supporting-Files-3TB-HSW-3-2016\logs-20160313-104031-hive-sf3000\bigBench-configs\hadoop
	Supporting-Files-3TB-HSW-3-2016\logs-20160313-104031-hive-sf3000\bigBench-configs\hive
	Supporting-Files-3TB-HSW-3-2016\logs-20160313-104031-hive-sf3000\bigBench-configs\spark
Clause 5 - Metric and Scale Factor Related Items	
Benchmark Performance Report	Supporting-Files-3TB-HSW-3-2016\logs-20160313-104031-hive-sf3000\run-logs\BigBenchResult.log
Validation Test Report	Supportin-Files-3TB-HSW-3-2016\Validation-run-logs-20160311-162432-hive-sf3000\run-logs\BigBenchResult.log