Dell Inc.

TPC Express Benchmark™ Big Bench (TPCx-BB)

Full Disclosure Report

for

PowerEdge R730/R730xd

(with 13x R730/R730xd)

using

Cloudera CDH 5.9.0

and

Red Hat Enterprise Linux Server 7.2

First Edition

May 12, 2017

![DELL EMC]	**PowerEdge R730/R730xd**	TPCx-BB Rev. v1.2.0 TPC-Pricing Rev. v2.1.1
		Report Date: May 12, 2017

Total System Cost	TPCx-BB Performance Metric	Price/Performance
439,187 USD	**495.28** BBQpm@10000	**886.75 USD** $/BBQpm@10000

Framework	Operating System	Other Software	Availability Date	Scale Factor	Streams
Cloudera CDH 5.9.0	Red Hat Enterprise Linux Server 7.2	OpenJDK 64-bit Server 1.7.0_67	May 12, 2017	10000	2

System Configuration

System Configuration

1xDell R730 Server (Namenode)
w/ 2xIntel E5-2650 v4, 8x 1TB,3.5",6G,SATA HDD

12xDell R730xd Servers (Datanodes)
each w/ 2xIntel E5-2690 v4,
24x 1.2TB,2.5",6G,SAS + 2x300GB,2.5",6G, SAS (OS)

Networking:

2xDell Force10 S4810 10 GbE

Physical Storage/Scale Factor: 36.08	Scale Factor/Physical Memory: 3.00

Servers:	13x R730/R730xd
Total Processors/Cores/Threads	26/360/720

Server Configuration:	**R730 (Namenode):**	**12x R730xd (Datanodes):**
Processors	2x Intel E5-2650 v4 @ 2.20GHz	2x Intel E5-2690 v4 @ 2.60GHz
Memory	256 GB	256 GB
Storage Controller	Dell PERC H730 Mini	Dell PERC H730 Mini
Storage Device	8x 1TB, 3.5", 7200 rpm, 6G SATA	24x 1.2TB, 2.5", 10K rpm, 6G SAS 2x 300GB, 2.5", 6G SAS
Network	Intel x520 DP 10 Gb Intel x520 DP 10GbE + I350 DP 1GbE	Intel x520 DP 10GbE + I350 DP 1GbE
Connectivity:	2x Dell Force10 S4810	

DELL EMC

PowerEdge R730/R730xd

TPCx-BB Rev. v1.2.0
TPC-Pricing Rev. v2.1.1

Report Date:
May 12, 2017

Description	Part Number	Key		Unit Price	Qty	Extended Price	3 yr. Maint. Price
HARDWARE COMPONENTS							
PowerEdge R730xd Server	210-ADBC	1	$25,520.00	$37,964.00	12	$455,568.00	
PowerEdge R730/R730xd Motherboard	329-BCZK	1	$0.00	$0.00	12		
R730/xd PCIe Riser 2, Center	330-BBCO	1	1	$0.00	12		
R730/xd PCIe Riser 1, Right	330-BBCR	1	1	$0.00	12		
Intel X520 DP 10Gb DA/SFP+, + I350 DP 1Gb Ethernet, Network Daughter Card	540-BBBB	1	1	$0.00	12		
Chassis with up to 24, 2.5 Hard Drives and 2, 2.5" Flex Bay Hard Drives	350-BBFE	1	1	$0.00	12		
Performance BIOS Settings	384-BBBL	1	1	$0.00	12		
UEFI BIOS	800-BBDM	1	1	$0.00	12		
No RAID for H330/H730/H730P including Flex Bay Drives (1-24 HDDs or SSDs)	780-BBLS	1	1	$0.00	12		
PERC H730 Integrated RAID Controller, 1GB Cache	405-AAEG	1	1	$0.00	12		
Intel Xeon E5-2690 v4 2.6GHz,35M Cache,9.60GT/s QPI,Turbo,HT,14C/28T (135W) Max Mem 2400MHz	338-BJCT	1	1	$0.00	12		
Upgrade to Two Intel Xeon E5-2690 v4 2.6GHz,35M Cache,9.60GT/s QPI,Turbo,HT,14C/28T (135W)	374-BJES	1	1	$0.00	12		
16GB RDIMM, 2400 MT/s, Dual Rank, x8 Data Width	370-ACNX	1	1	$0.00	192		
2400MT/s RDIMMs	370-ACPH	1	1	$0.00	12		
Performance Optimized	370-AAIP	1	1	$0.00	12		
300GB 10K RPM SAS 6Gbps 2.5in Flex Bay Hard Drive,13G	400-AJPR	1	1	$0.00	24		
1.2 TB 10K RPM SAS 6Gbps 2.5in Hot-plug Hard Drive,13G	400-AEFO	1	1	$0.00	288		
Electronic System Documentation and OpenManage DVD Kit, PowerEdge R730/xd	631-AAJG	1	1	$0.00	12		

PowerEdge R730/R730xd

Description	Part Number	Key		Unit Price	Qty	Extended Price	3 yr. Maint. Price
ReadyRails Sliding Rails With Cable Management Arm	770-BBBR	1	1	$0.00	12		
Dual, Hot-plug, Redundant Power Supply (1+1), 750W	450-ADWS	1	1	$0.00	12		
C13 to C14, PDU Style, 12 AMP, 2 Feet (.6m) Power Cord, North America	492-BBDH	1	1	$0.00	12		
DIMM Blanks for System with 2 Processors	370-ABWE	1	1	$0.00	12		
Standard Heatsink for PowerEdge R730/R730xd	374-BBHM	1	1	$0.00	12		
Standard Heatsink for PowerEdge R730/R730xd	374-BBHM	1	1	$0.00	12		
INFO QS, 13G HADOOP BUNDLE	379-BBWM	1	1	$0.00	12		
4hr response, 24x7 Maintenance Package, R730xd				$2,529.00	12		$30,348.00
- Thank you choosing Dell ProSupport. For tech support, visit http://support.dell.com/ProSupport	911-6619	1		$0.00	12		
- Dell Limited Hardware Warranty Plus Service, Initial Year	978-4029	1		$0.00	12		
- ProSupport Mission Critical Package: Enhanced Services,3 Year	978-4042	1		$0.00	12		
- ProSupport Mission Critical: 7X24 HW / SW Tech Support and Assistance,3 Year	978-4043	1		$0.00	12		
- Mission Critical Package: 4-Hours 7X24 On-Site Service with Emergency Dispatch, 3 Year	978-4044	1		$0.00	12		
- On-Site Installation Declined	900-9997	1		$0.00	12		
- US Order	332-1286	1		$0.00	12		

PowerEdge R730/R730xd

Description	Part Number	Key	Unit Price	Qty	Extended Price	3 yr. Maint. Price
PowerEdge R730 Server	329-BCZK	1	$25,027.92	1	$25,027.92	
PowerEdge R730/R730xd Motherboard	591-BBCH	1	$0.00	1		
R730/xd PCIe Riser 2, Center	330-BBCO	1	$0.00	1		
R730 PCIe Riser 3, Left	330-BBCQ	1	$0.00	1		
R730/xd PCIe Riser 1, Right	330-BBCR	1	$0.00	1		
Intel X520 DP 10Gb DA/SFP+ Server Adapter, Low Profile	540-BBHY	1	$0.00	1		
Intel X520 DP 10Gb DA/SFP+, + I350 DP 1Gb Ethernet, Network Daughter Card	540-BBBB	1	$0.00	1		
Chassis with up to 8, 3.5" Hard Drives	350-BBEO	1	$0.00	1		
Performance BIOS Settings	384-BBBL	1	$0.00	1		
UEFI BIOS	800-BBDM	1	$0.00	1		
No RAID for H330/H730/H730P (1-16 HDDs or SSDs)	780-BBJS	1	$0.00	1		
PERC H730 Integrated RAID Controller, 1GB Cache	405-AAEG	1	$0.00	1		
Intel Xeon E5-2650 v4 2.2GHz,30M Cache,9.60GT/s QPI,Turbo,HT,12C/24T (105W) Max Mem 2400MHz	338-BJDV	1	$0.00	1		
Upgrade to Two Intel Xeon E5-2650 v4 2.2GHz,30M Cache,9.60GT/s QPI,Turbo,HT,12C/24T (105W)	374-BJDW	1	$0.00	1		
32GB RDIMM, 2400 MT/s, Dual Rank, x4 Data Width	370-ABUG	1	$0.00	8		
2400MT/s RDIMMs	370-ACPH	1	$0.00	1		
Performance Optimized	370-AAIP	1	$0.00	1		
1TB 7.2K RPM SATA 6Gbps 3.5in Hot-plug Hard Drive,13G	400-AEEZ	1	$0.00	8		
Electronic System Documentation and OpenManage DVD Kit, PowerEdge R730/xd	631-AAJG	1	$0.00	1		
DVD+/-RW, SATA, Internal	429-AAPS	1	$0.00	1		

Description	Part Number	Key	Unit Price	Qty	Extended Price	3 yr. Maint. Price
ReadyRails Sliding Rails With Cable Management Arm	770-BBBR	1	$0.00	1		
Dual, Hot-plug, Redundant Power Supply (1+1), 750W	450-ADWS	1	$0.00	1		
C13 to C14, PDU Style, 12 AMP, 2 Feet (.6m) Power Cord, North America	492-BBDH	1	$0.00	2		
DIMM Blanks for System with 2 Processors	370-ABWE	1	$0.00	1		
Standard Heatsink for PowerEdge R730/R730xd	374-BBHM	1	$0.00	1		
Standard Heatsink for PowerEdge R730/R730xd	374-BBHM	1	$0.00	1		
INFO QS, 13G HADOOP BUNDLE	379-BBWM	1	$0.00	1		
4hr response, 24x7 Maintanance Package, R730		1	$2,049.00	1		$2,049.00
- Thank you choosing Dell ProSupport. For tech support, visit http://support.dell.com/ProSupport	911-6619	1	$0.00	1		
- Dell Limited Hardware Warranty Plus Service, Initial Year	978-3603	1	$0.00	1		
- ProSupport Mission Critical Package: Enhanced Services,3 Year	978-3617	1	$0.00	1		
- ProSupport Mission Critical: 7X24 HW / SW Tech Support and Assistance,3 Year	978-3618	1	$0.00	1		
- Mission Critical Package: 4-Hours 7X24 On-Site Service with Emergency Dispatch, 3 Year	978-3619	1	$0.00	1		
- On-Site Installation Declined	900-9997	1	$0.00	1		
- US Order	332-1286	1	$0.00	1		

Description	Part Number	Key		Unit Price	Qty	Extended Price	3 yr. Maint. Price
Dell Force10 S4810 Switch		1	$21,156.00	$18,436.00	2	$36,872.00	
Force10, S4810P, 1RU, 48 x 10GbE SFP+, 4 x 40GbE QSFP+, 1 x AC PSU, 2 x FM, PSU to IO Panels (Reverse)	225-2477	1	$0.00	$0.00	2		
Dell Networking, Jumper Cord, 250V, 12A, 2 Meters, C13/C14, US	450-AASX	1	$0.00	$0.00	2		
Force10 Customer not deploying this switch in iSCSI or FCOE environment.	332-0139	1	$0.00	$0.00	2		
INFO QS, 13G HADOOP BUNDLE	379-BBWM	1	$0.00	$0.00	2		
Dell Networking, Cable, SFP+ to SFP+, 10GbE, Copper Twinax Direct Attach Cable, 3 Meter	470-AAGP	1	$0.00	$0.00	30		
4hr response, 24x7 Maintanance Package, S4810		1	$3,153.99	$3,153.99	2		$6,307.98
- Force10, User Documentation for S4810, DAO/BCC	331-6279	1	$0.00	$0.00	2		
- SW Support,Force10 Software ,3 Years	935-0103	1	$0.00	$0.00	2		
- ProSupport: 7x24 HW / SW Tech Support and Assistance, 3 Years	935-0143	1	$0.00	$0.00	2		
- Thank you choosing Dell ProSupport. For tech support, visit http://www.dell.com/support or call 1-800- 945-3355	989-3439	1	$0.00	$0.00	2		
- Dell Hardware Limited Warranty Initial Year	996-2670	1	$0.00	$0.00	2		
- Dell Hardware Limited Warranty Extended Year(s)	996-2760	1	$0.00	$0.00	2		
- On-Site Installation Declined	900-9997	1	$0.00	$0.00	2		
- Declined Remote Consulting Service	973-2426	1	$0.00	$0.00	2		
Dell Netshelter SX 42U Rack - 600mm Wide x 1070mm Deep	A7545497	1		$1,299.99	1	$1,299.99	
Logitech MK120 Keyboard and Mouse	A3974709	1		$16.99	1	$16.99	
HARDWARE COMPONENTS				**Subtotal**		$518,784.90	$38,704.98

Description	Part Number	Key		Unit Price	Qty	Extended Price	3 yr. Maint. Price
SOFTWARE COMPONENTS							
Red Hat Enterprise Linux,1-2SKT,3yr Premium Subscription,1 Virtual Guest	421-5721	1	1	$3,059.00	13	$39,767.00	
- Red Hat Enterprise Linux Non Factory Install, x64,Reqs Subscription Selection	605-BBIX	1	1	$0.00		$0.00	
- Red Hat Enterprise Linux 7.3 Media Only X86_64, No Subscription	340-AVFG	1	1	$0.00		$0.00	
Cloudera Enterprise Basic Edition, Node License, 24x7 3YR	A9081372	1	1	6,000	13	$78,000.00	
SOFTWARE COMPONENTS					Subtotal	$117,767.00	$0.00
Total						**$636,551.90**	**$38,704.98**
Large Purchase Discount (35%)*						-222,793.17	-13,546.74
ViewSonic VA2055Sa LED Monitor 20" (includes spares)	3701841	2		$89.99	3	269.97	

Pricing:1 = Dell; 2 = CDW (www.cdw.com)

[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

Three-Year Cost of Ownership	**$439,187**
BBQ pm@10000	**495.28**
$/BBQ pm@10000	**$ 886.75**

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	10000
Streams	2
SUT Validation Status	TRUE

Performance Run

Overall Run Start Time	2017-05-05 22:45:23.903
Overall Run End Time	2017-05-07 15:14:41.732
Overall Run Elapsed Time	145,757.829
Load Test Start Time	2017-05-05 22:45:23.904
Load Test End Time	2017-05-05 23:51:19.682
Load Test Elapsed Time	3,955.778
Power Test Start Time	2017-05-05 23:51:19.684
Power Test End Time	2017-05-06 14:36:07.138
Power Test Elapsed Time	53,087.454
Throughput Test Start Time	2017-05-06 14:36:07.140
Throughput Test End Time	2017-05-07 15:14:41.731
Throughput Test Elapsed Time	88,714.591
Performance Metric (BBQpm@ 10000)	495.282507443747

Repeatability Run

Overall Run Start Time	2017-05-03 23:08:35.781
Overall Run End Time	2017-05-05 15:26:35.283
Overall Run Elapsed Time	145,079.502
Load Test Start Time	2017-05-03 23:08:35.782
Load Test End Time	2017-05-04 00:14:50.426
Load Test Elapsed Time	3,974.644
Power Test Start Time	2017-05-04 00:14:50.428
Power Test End Time	2017-05-04 14:58:00.867
Power Test Elapsed Time	52,990.439
Throughput Test Start Time	2017-05-04 14:58:00.869
Throughput Test End Time	2017-05-05 15:26:35.282
Throughput Test Elapsed Time	88,114.413
Performance Metric (BBQpm@ 10000)	496.657229239438

Run Report – Run 1

```
************
TPCx-BB
Result
v1.2
************
INFO: T_LOAD = 3974.644
INFO: T_LD = 0.1 * T_LOAD: 397.4644
INFO: T_PT = 29163.26971375
INFO: T_T_PUT = 88114.413
INFO: T_TT = 44057.2065
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@10000 = 496.657229239438
```

Run Report – Run 2

```
************
TPCx-BB
Result
v1.2
************
INFO: T_LOAD = 3955.778
INFO: T_LD = 0.1 * T_LOAD: 395.5778
INFO: T_PT = 29131.838981562
INFO: T_T_PUT = 88714.591
INFO: T_TT = 44357.2955
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@10000 = 495.282507443747
```

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

Table of Contents

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 v1.2.0.

The test was conducted at a Scale Factor of 10000 with 13 nodes (R730/R730xd) running Cloudera CDH 5.9.0 Red Hat Enterprise Linux Server 7.2.

Measured Configuration

Company Name	Cluster Node	Virtualization	Operating System
Dell Inc.	13x R730/R730xd	n/a	Red Hat Enterprise Linux Server 7.2

TPC Express Benchmark© Big Bench Metrics

Total System Cost	BBQpm@10000	Price/Performance	Availability Date
439,187 USD	495.28	886.75 USD	May 12, 2017

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 Dell Inc.

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.
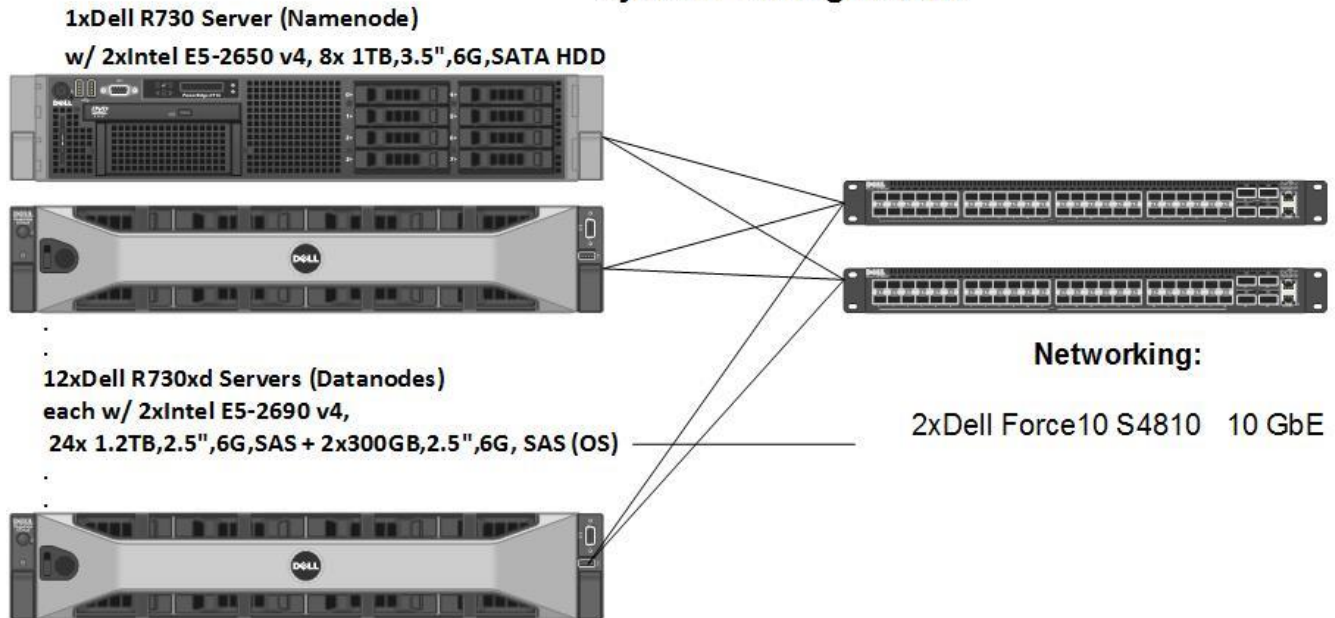
1.3 Configuration Diagrams

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

System Configuration

1xDell R730 Server (Namenode)

w/ 2xIntel E5-2650 v4, 8x 1TB,3.5",6G,SATA HDD

12xDell R730xd Servers (Datanodes)
each w/ 2xIntel E5-2690 v4,
24x 1.2TB,2.5",6G,SAS + 2x300GB,2.5",6G, SAS (OS)

Networking:

2xDell Force10 S4810 10 GbE

The measured configuration consisted of:
- Total Nodes: 13
- Total Processors/Cores/Threads: 26/360/720
- Total Memory: 3,328 GB
- Total Number of Storage Drives/Devices: 320
- Total Storage Capacity: 360,800 GB
- Network: 2x Dell Force10 S4810

R730 (Namenode):
- Processors/Cores/Threads: 2/24/48
- Processor: 2x Intel E5-2650 v4 @ 2.20GHz
- Memory: 256 GB
- Controller: Dell PERC H730 Mini
- Drives:
  - 8x 1TB, 3.5", 7200 rpm, 6G SATA
- Network:
  - Intel x520 DP 10 Gb
  - Intel x520 DP 10GbE + I350 DP 1GbE

12x R730xd (Datanodes) each with:
- Processors/Cores/Threads: 2/28/56
- Processor: 2x Intel E5-2690 v4 @ 2.60GHz
- Memory: 256 GB
- Controller: Dell PERC H730 Mini
- Drives:
  - 24x 1.2TB, 2.5", 10K rpm, 6G SAS
  - 2x 300GB, 2.5", 6G SAS
- Network:
  - Intel x520 DP 10GbE + I350 DP 1GbE

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 Names	HW/SW Configuration	Storage Setup
Worker	HDFS Datanode HDFS Gateway HDFS NFS Gateway Hive Gateway Spark Gateway YARN/MR2 Gateway YARN Node Manager	12	No	R3s1xd (1-12)	Dell PowerEdge R730xd Processor: 2x Intel E5-2690 v4 Memory: 256 GB Storage: 24x 1.2TB (Data), 2x 300GB (OS) Network: Intel x520 DP 10GbE + I350 DP 1GbE OS: RHEL 7.2 Cloudera CDH 5.9	OS: 300GB RAID1, SAS, HDD Intermediate / Shuffle / Temp Data Distributed FS: 24x 1.2 TB, JBOD, SAS
Master	HDFS Namenode HDFS Secondary Namenode HDFS Balancer HDFS NFS Gateway Hive Server Hive Meta Store Hive Gateway Hive Web MetHCat Clouder Management Services Cloudera Manager Spark Gateway Spark History Server YARN/MR2 Gateway YARN JobHistory YARN Resource Manager Zookeeper	1	No	r3s1	Dell PowerEdge R730 Processor: 2x Intel E5-2650 v4 Memory: 256 GB Storage: 8x 1 TB (OS) Network: Intel x520 DP 10GbE + I350 DP 1GbE, Intel x520 DP 10GbE OS: RHEL 7.2 Cloudera CDH 5.9	OS: 8x1 TB RAID10, SATA, HDD

2.2 Distributed File System Implementation

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

Cloudera CDH 5.9.0 (supporting Apache HDFS version 2).

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	1.6.0
MapReduce	2.6.0

2.4 Frameworks

Frameworks and Engine used in the benchmark should be disclosed.

Framework	Version
CDH	5.9.0
Hive	1.1.0
HDFS	2.6.0
YARN	2.6.0
Spark	1.6.0
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
v1.2

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
  v1.2
  ************
  INFO: T_LOAD = 3974.644
  INFO: T_LD = 0.1 * T_LOAD: 397.4644
  INFO: T_PT = 29163.26971375
  INFO: T_T_PUT = 88114.413
  INFO: T_TT = 44057.2065
  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@10000 = 496.657229239438
  ```

- **Run2 Performance Summary**

  ```
  ************
  TPCx-BB
  Result
  v1.2
  ************
  INFO: T_LOAD = 3955.778
  INFO: T_LD = 0.1 * T_LOAD: 395.5778
  INFO: T_PT = 29131.838981562
  INFO: T_T_PUT = 88714.591
  INFO: T_TT = 44357.2955
  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@ 10000 = 495.282507443747

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 Stream 1	Throughput Stream 1	Throughput Stream 2
Structured	1	348.434	926.967	409.467
	6	1052.014	2434.697	3869.293
	7	889.358	3980.956	1065.815
	9	1102.577	1666.111	2716.029
	11	221.547	268.329	223.642
	13	613.964	3101.933	1280.741
	14	109.544	110.012	128.274
	15	148.338	417.735	146.764
	16	2726.34	3248.206	6831.777
	17	563.813	2720.419	564.505
	20	682.654	977.311	748.701
	21	1697.216	2013.052	7170.01
	22	574.157	1056.408	1002.655
	23	295.54	538.378	315.159
	24	367.421	1969.282	382.561
	25	1456.822	1807.295	2634.341
	26	2106.322	5857.922	2622.13
	29	1597.414	1621.164	2342.914
Semi-structured	2	6035.585	7330.176	7545.4
	3	2452.554	5675.187	3052.584
	4	5312.899	9989.862	7420.411
	5	5722.3	7141.053	10827.788
	8	1260.173	2011.179	1580.659
	12	1281.36	2037.867	1763.45
	30	9069.147	13122.844	13305.469
Unstructured	10	362.856	403.44	434.819
	18	1680.195	3179.17	3055.653
	19	775.197	2175.866	1571.31
	27	125.256	121.113	152.172
	28	2456.388	810.595	2263.546

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/bigBench.properties
- bigBench-configs/conf/userSettings.conf
- bigBench-configs/hive/queries/q{1-30}/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 form 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)
1	8	1000	8000
12	24	1200	345600
12	2	300	7200

Total Storage (GB)	360,800 GB
Scale Factor	10000
Data Storage Ratio	36.08

4.5 Scale Factor to Memory Ratio

The Scale Factor to memory ratio must be disclosed.

Nodes	Memory (GB)	Total (GB)
13	256	3328

Scale Factor	10000
Total Memory (GB)	3,328 GB
SF / Memory Ratio	3.00

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@10000 495.28

5.2 Repeatability Run Metric

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

Repeatability Run
BBQpm@10000 496.65

5.3 Price-Performance Metric

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

Price / Performance
$BBQpm@10000 886.75

5.4 Scale Factor

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

Scale Factor
10000

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	01:16:17:59.502	145,079.502
Run 2	01:16:29:17.829	145,757.829

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	3,955.778	3,974.644
Power Test	53,087.454	52,990.439
Throughput Test	88,714.591	88,114.413

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, InfoSizing.

> www.sizing.com
> 63 Lourdes Drive
> Leominster, MA 01453
> 978-343-6562.

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.

InfoSizing
The Right Metric For Sizing IT

TPC Transaction Processing Performance Council
Certified Auditors

Mr. Nicholas Wakou
Dell Inc.
701 E. Parmer Ln. Bldg. 2
Austin, TX 78753

May 12, 2017

I verified the TPC Express Benchmark™ BB v1.2.0 performance of the following configuration:

Platform: Dell PowerEdge R730/R730xd (with 13x R730/R730xd)
Operating System: Red Hat Enterprise Linux Server 7.2
Apache Hadoop   Cloudera for Apache Hadoop (CDH) 5.9.0
Compatible Software:

The results were:

Performance Metric   **495.28 BBQpm@10000GB**
Run Elapsed Time   01:16:29:17.829 (145,757.829 Seconds)

Cluster	**1x R730 (Namenode), 12x R730xd (Datanodes)**		
CPUs	2 x Intel E5-2650 v4 (2.20 GHz, 12-core, 30 MB L3) (Namenode)		
	2 x Intel E5-2690 v4 (2.60 GHz, 14-core, 35 MB L3) (Datanodes)		
Memory	256GB (all nodes)		
Storage	Qty	Size	Type
	8	1 TB	7.2K rpm 6G SATA HDD (OS Namenode)
	2	300 GB	10K rpm 6G SAS HDD (OS Datanodes)
	24	1.2 TB	10K rpm 6G SAS HDD (Data Datanodes)

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 v1.2.0
- 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 (10000GB) was confirmed to be valid for publication
- All validation queries executed successfully and produced compliant results

- 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:

From the TPCx-BB Kit's README:

Q28 Depending on the Hadoop distribution version can fail automated Engine Validation due to empty space characters when the output is written to HDFS. Manually open the result file and validate the reference values and written values.

Query 28 failed automated Engine Validation. A manual validation was performed as part of this audit to confirm the only differences were due to white space.

Respectfully Yours,

Doug Johnson, TPC Auditor

63 Lourdes Dr. | Leominster, MA 01453 | 978-343-6562 | www.sizing.com

Third Party Price Quotes

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-10TB-BDW-5-2017\
Validation Run Files	Supporting-Files-10TB-BDW-5-2017\Validation-run-20170503-224412-hive-sf10000
Performance Run Files	Supporting-Files-10TB-BDW-5-2017\Performance-run-logs-20170507-151707-hive-sf10000
Repeatability Run Files	Supporting-Files-10TB-BDW-5-2017\Repeatability-run-logs-20170505-152900-hive-sf10000
Clause 3 - Workload Related Items	
Benchmark Generic Parameters	Supporting-Files-10TB-BDW-5-2017\ Performance-run-logs-20170507-151707-hive-sf10000\bigBench-configs\conf\userSettings.conf
Query Parameters used in the benchmark execution Settings	Supporting-Files-10TB-BDW-5-2017\ Performance-run-logs-20170507-151707-hive-sf10000\bigBench-configs\hive\conf\queryParameters.sql
Benchmark Global Framework Parameters Settings	Supporting-Files-10TB-BDW-5-2017\ Performance-run-logs-20170507-151707-hive-sf10000\bigBench-configs\hive\conf\engineSettings.sql
Benchmark Global Framework Parameters Settings	Supporting-Files-10TB-BDW-5-2017\ Performance-run-logs-20170507-151707-hive-sf10000\bigBench-configs\hive\conf\engineSettings.conf
Load Test script	Supporting-Files-10TB-BDW-5-2017\ Performance-run-logs-20170507-151707-hive-sf10000\bigBench-configs \hive\population\ hiveCreateLoad.sql
Queries specific optimization parameters settings	Supporting-Files-10TB-BDW-5-2017\ Performance-run-logs-20170507-151707-hive-sf10000\bigBench-configs\hive\queries\q[01-30]\engineLocalSettings.conf
Queries specific optimization parameters settings	Supporting-Files-10TB-BDW-5-2017\ Performance-run-logs-20170507-151707-hive-sf10000\bigBench-configs\hive\queries\q[01-30]\engineLocalSettings.sql
Clause 4 - SUT Related Items	
Data Redundancy report	Supporting-Files-10TB-BDW-5-2017\hdfs-data-redundancy-report.txt
Benchmark execution script	Supporting-Files-10TB-BDW-5-2017\TPCxBB_Benchmarkrun.sh
Hardware and Software Report from a representative node	Supporting-Files-10TB-BDW-5-2017\ Performance-run-logs-20170507-151707-hive-sf10000\run-logs\ envInfo-datanode1-r3s1.ignition.dell.com\envInfo.log
	Supporting-Files-10TB-BDW-5-2017\Performance-run-logs-20170507-151707-hive-sf10000\bigBench-configs\hadoop

All Framework configuration files are included in the Supporting Files Archive	Supporting-Files-10TB-BDW-5-2017\Performance-run-logs-20170507-151707-hive-sf10000\bigBench-configs\hive
	Supporting-Files-10TB-BDW-5-2017\Performance-run-logs-20170507-151707-hive-sf10000\bigBench-configs\spark

Clause 5 - Metric and Scale Factor Related Items

Benchmark Performance Report	Supporting-Files-10TB-BDW-5-2017\Performance-run-logs-20170507-151707-hive-sf10000\run-logs\BigBenchResult.log
Validation Test Report	Supporting-Files-10TB-BDW-5-2017\Performance-run-logs-20170507-151707-hive-sf10000\run-logs\BigBenchResult.log