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

(w/ 18x HPE ProLiant DL380 Gen10, 3x HPE ProLiant DL360 Gen10)

using

Cloudera for Apache Hadoop (CDH) 5.16.1

and

Red Hat Enterprise Linux Server 7.6

First Edition

May 1, 2019

***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.

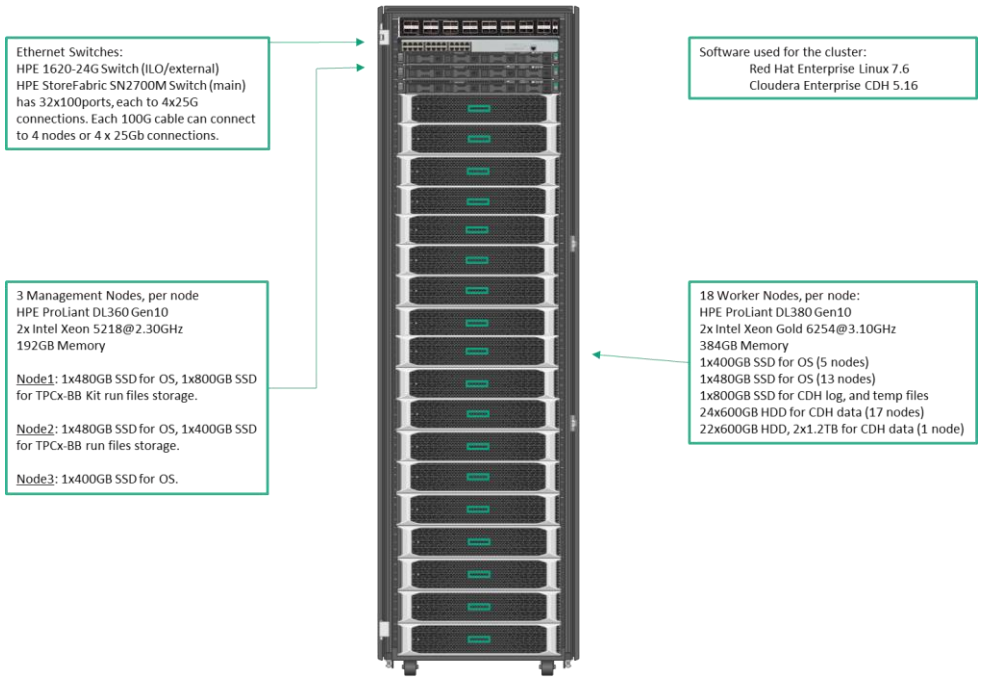
The HPE products, services or features identified in this document may not yet be available or may not be available in all areas and may be subject to change without notice. Consult your local HPE business contact for information on the products or services available in your area. You can find additional 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.

Hewlett Packard Enterprise	**Hewlett Packard Enterprise ProLiant DL for Big Data**	TPCx-BB Rev. v1.2.0 TPC-Pricing Rev. v2.4.0 Report Date: May 1, 2019

Total System Cost	TPCx-BB Performance Metric	Price/Performance
913,107 USD	**1,789.75** BBQpm@10000	**510.19 USD** $/BBQpm@10000

Framework	Operating System	Other Software	Availability Date	Scale Factor	Streams
Cloudera for Apache Hadoop (CDH) 5.16.1	Red Hat Enterprise Linux Server 7.6	None	May 1, 2019	10000	10

System Configuration



Physical Storage/Scale Factor: 28.56	Scale Factor/Physical Memory: 1.34

Servers: Total Processors/Cores/Threads	18x HPE ProLiant DL380 Gen10, 3x HPE ProLiant DL360 Gen10 42/744/1,488	
Server Configuration: Processors Memory Storage Controller Storage Device Network	**Per HPE ProLiant DL380 Gen10:** 2x Intel Xeon Gold 6254 @ 3.10GHz 384GB HPE Smart Array P408i-a SR 1x HPE 800GB (all nodes) 1x HPE 400GB, 24x HPE 600GB (5 nodes) 1x HPE 480GB, 24x HPE 600GB (12 nodes) 1x HPE 480GB, 24x HPE 600GB, 2x HPE 1.2TB (1 node) HPE Eth 10/25Gb 2P 640SFP28	**Per HPE ProLiant DL360 Gen10:** 2x Intel Xeon Gold 5218 @ 2.30GHz 192GB HPE Smart Array E208i-a SR 1x HPE 480GB, 1x HPE 800GB Node1 1x HPE 400GB, 1x HPE 480GB Node2 1x HPE 400GB Node3 HPE Eth 10/25Gb 2P 640SFP28
Connectivity:	HPE 1620-24G Switch (ILO connection) HPE StoreFabric SN2700M 100/25G (main connection)	

Description	Price Key	Part Number	Unit Price	Qty	Extended Price	3 Yr Maint Price
Server Hardware						
HPE DL360 Gen10 8SFF CTO Server	1	867959-B21	$2,099	3	$6,297	
HPE DL360 Gen10 Xeon-G 5218 Kit	1	P02592-B21	$2,019	3	$6,057	
HPE DL360 Gen10 Xeon-G 5218 FIO	1	P02592-L21	$2,019	3	$6,057	
HPE 16GB 1Rx4 PC4-2666V-R Smart Kit	1	835955-B21	$630	36	$22,680	
HPE 480GB 6G SATA RI-2 SFF SC SSD	1	804593-B21	$609	2	$1,218	
HPE 400GB 6G SATA WI-2 SFF SC SSD	1	804665-B21	$719	2	$1,438	
HPE 800GB 6G SATA WI-2 SFF SC SSD	1	804671-B21	$1,389	1	$1,389	
HPE 800W CS Platinum Plus AC Power Supply	1	865414-B21	$379	6	$2,274	
HPE 96W Smart Storage Battery 145mm Cbl	1	875241-B21	$99	3	$297	
HPE Eth 10/25Gb 2P 640SFP28	1	817753-B21	$779	3	$2,337	
HPE 3Y FC 24x7 DL360 Gen10 SVC	1a	H8QF0E	$1,811	3		$5,433
HPE iLO Adv incl 3yr TS U E-LTU	1a	E6U64ABE	$469	3		$1,407
HP W2082a 20-inch LED Monitor (1 + 2 spare)	2	#L8K84AA	$85	3	$255	
HPE USB US Keyboard/Mouse Kit	1	631341-B21	$29	3	$87	
			Subtotal		**$50,386**	**$6,840**
HPE DL380 Gen10 24SFF CTO Server	1	868704-B21	$2,759	18	$49,662	
HPE DL380 Gen10 High Perf Fan Kit	1	867810-B21	$239	18	$4,302	
HPE DL380 Gen10 2SFF Bay Kit	1	826687-B21	$249	18	$4,482	
HPE DL380 Gen10 Xeon-G 6254 Kit	1	P02517-B21	$6,065	18	$109,170	
HPE DL380 Gen10 Xeon-G 6254 Kit	1	P02517-L21	$6,065	18	$109,170	
HPE 32GB 2Rx4 PC4-2666T-R Kit	1	815100-B21	$1,170	216	$252,720	
HPE 1.2TB SAS 10K SFF SC HDD	1	718162-B21	$909	2	$1,818	
HPE 600GB SAS 6G 10K SFF SC DS HDD	1	872477-B21	$525	430	$225,750	
HPE 480GB 6G SATA RI-2 SFF SC SSD	1	804593-B21	$609	13	$7,917	
HPE 400GB 6G SATA WI-2 SFF SC SSD	1	804665-B21	$719	5	$3,595	
HPE 800GB 6G SATA WI-2 SFF SC SSD	1	804671-B21	$1,389	18	$25,002	
HPE Smart Array P408i-a SR	1	804331-B21	$599	18	$10,782	
HPE 12Gb DL380 Gen10 SAS Expander Card	1	870549-B21	$699	18	$12,582	
HPE 1600W FS Plat Ht Plg LH Pwr Sply Kit	1	830272-B21	$479	36	$17,244	
HPE Eth 10/25Gb 2P 640SFP28	1	817753-B21	$779	18	$14,022	
HPE 3Y FC 24x7 DL380 Gen10 SVC	1a	H8QP7E	$2,448	18		$44,064
HPE iLO Adv incl 3yr TS U E-LTU	1a	E6U64ABE	$469	18		$8,442
			Subtotal		**$848,218**	**$52,506**

(continued on next page)

(continued from previous page)

Description	Price Key	Part Number	Unit Price	Qty	Extended Price	3 Yr Maint Price
Network						
HPE 1620-24G Switch (1 + 2 spare)	1	JG913A	$315	3	$299	
HPE StoreFabric SN2700M 100GbE 32QSFP28 Switch	1	Q2F21A	$23,995	1	$23,995	
HPE 3Y PC 24x7 SN2700M Stg Switch SVC	1a	HT5Q6E	$7,414	1		$7,414
CAT6 UTP Ethernet Network Cable 7ft (42 + 5 spare)	3	2302	$2	47	$94	
HPE 100Gb QSFP28 to 4x25Gb SFP28 3m DAC (6 + 2 spare)	1	845416-B21	$699	8	$5,592	
			Subtotal		**$29,980**	**$7,414**
Rack						
HPE 42U 600x1075mm Adv G2 Kit Pllt Rack	1	P9K07A	$1,179	1	$1,179	
HPE 24A High Voltage Core Only Corded PDU	1	252663-D74	$259	2	$518	
			Subtotal		**$1,697**	**$0**
Server Software						
Cloudera Ent Data Eng Ed	1	P9V18AAE	$12,000	21	$252,000	
RHEL Svr 2 Sckt/2 Gst 3yr 24x7 E-LTU	1a	G3J30AAE	$3,702	21	$77,742	
			Subtotal		**$329,742**	**$0**
		Total Extended Price			**$1,260,023**	**$66,760**
		Total Discounts			**$413,676**	**$0**
Sales contact: HPE WW Headquarters, 3000 Hanover St., Palo Alto, CA 94304-1185 (650) 857-1501 or HPE: 855-472-5233		**Grand Total**			**$846,347**	**$66,760**

Pricing:1 = HPE; 2 = store.hp.com; 3 = monoprice.com

(1) All discounts are based on US list prices and for similar quantities and configurations. A 35% discount was based on the overall specific components pricing from vendor 1 (excluding 1a items) 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	**$913,107**
BBQpm@10000	**1,789.75**
$/BBQpm@10000	**$ 510.19**

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	10
SUT Validation Test	FAIL

Performance Run (Run 1)

Overall Run Start Time	2019-04-12 23:55:56.418
Overall Run End Time	2019-04-14 05:56:10.456
Overall Run Elapsed Time	108,014.038
Load Test Start Time	2019-04-12 23:55:56.418
Load Test End Time	2019-04-13 00:30:44.659
Load Test Elapsed Time	2,088.241
Power Test Start Time	2019-04-13 00:30:44.660
Power Test End Time	2019-04-13 05:02:30.251
Power Test Elapsed Time	16,305.591
Throughput Test Start Time	2019-04-13 05:02:30.251
Throughput Test End Time	2019-04-14 05:56:10.456
Throughput Test Elapsed Time	89,620.205
Performance Metric (BBQpm@ 10000)	1,789.75

Repeatability Run (Run 2)

Overall Run Start Time	2019-04-14 10:39:56.837
Overall Run End Time	2019-04-15 16:35:48.657
Overall Run Elapsed Time	107,751.820
Load Test Start Time	2019-04-14 10:39:56.837
Load Test End Time	2019-04-14 11:13:59.022
Load Test Elapsed Time	2,042.185
Power Test Start Time	2019-04-14 11:13:59.023
Power Test End Time	2019-04-14 15:45:58.576
Power Test Elapsed Time	16,319.553
Throughput Test Start Time	2019-04-14 15:45:58.577
Throughput Test End Time	2019-04-15 16:35:48.657
Throughput Test Elapsed Time	89,390.080
Performance Metric (BBQpm@ 10000)	1,793.02

Performance Run Report (Run 1)

****
TPCx-BB
Result
v1.2
****
INFO: T_LOAD = 2088.241
INFO: T_LD = 0.1 * T_LOAD: 208.8241
INFO: T_PT = 10822.5410086123
INFO: T_T_PUT = 89620.205
INFO: T_TT = 8962.0205
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 = 1789.75057336083

Repeatability Run Report (Run 2)

****
TPCx-BB
Result
v1.2
****
INFO: T_LOAD = 2042.185
INFO: T_LD = 0.1 * T_LOAD: 204.2185
INFO: T_PT = 10820.131815946
INFO: T_T_PUT = 89390.08
INFO: T_TT = 8939.008
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 = 1793.02254740899

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 21 nodes ({Node Model}) running Cloudera for Apache Hadoop (CDH) 5.16.1 on Red Hat Enterprise Linux Server 7.6.

Measured Configuration

Company Name	Cluster Node	Virtualization	Operating System
Hewlett Packard Enterprise Company	18x HPE ProLiant DL380 Gen10 3x HPE ProLiant DL360 Gen10	n/a	Red Hat Enterprise Linux Server 7.6

TPC Express Benchmark© Big Bench Metrics

Total System Cost	BBQpm@10000	Price/Performance	Availability Date
913,107 USD	1,789.75	510.19 USD	May 1, 2019

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.

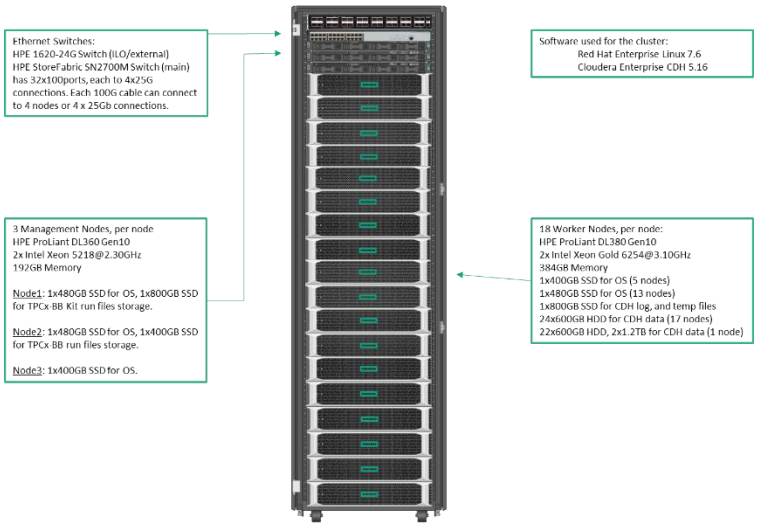
The Supporting Files Archive contains the parameters and options used to configure the components involved in this benchmark.

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: 21
- Total Processors/Cores/Threads: 42/744/1,488
- Total Memory: 7,488
- Total Number of Storage Drives/Devices: 473
- Total Storage Capacity: 285,600

Network connectivity detail:

- HPE 1620-24G Switch (ILO connection), HPE StoreFabric SN2700M 100/25G (main connection)

Server nodes details:

18x HPE ProLiant DL380 Gen10, each with:

- Processors/Cores/Threads: 2/36/72
- Processor Model: 2x Intel Xeon Gold 6254 @ 3.10GHz
- Memory: 384GB
- Controller: 1x HPE Smart Array P408i-a SR
- Drives:
  - 1x HPE 800GB (all nodes)
  - 1x HPE 400GB, 24x HPE 600GB (5 nodes)
  - 1x HPE 480GB, 24x HPE 600GB (12 nodes)
  - 1x HPE 480GB, 24x HPE 600GB, 2x HPE 1.2TB (1 node)
- Network: HPE Eth 10/25Gb 2P 640SFP28

3x HPE ProLiant DL360 Gen10, each with:

- Processors/Cores/Threads: 2/32/64
- Processor Model: 2 x Intel Xeon Gold 5218 @ 2.30GHz
- Memory: 192GB
- Controller: 1x HPE Smart Array E208i-a SR
- Drives:
  - 1x HPE 480GB, 1x HPE 800GB Node1
  - 1x HPE 400GB, 1x HPE 480GB Node2
  - 1x HPE 400GB Node3
- Network: HPE Eth 10/25Gb 2P 640SFP28

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/Hive Gateway/YARN Node Manager/Spark Gateway	18	N	pan-[04-21]	• HPE DL380 Gen10 • HW/SW Config (Intel Xeon Gold 6254, 2, 3.1GHz, 72) • Memory: 384GB • Storage: 24 x 600GB SAS HDD, 1 x 480GB or 400GB SSD, 1x800GB SSD • Network: HPE Eth 10/25Gb 2P 640SFP28 • OS: RHEL 7.5 • Cloudera CDH 5.16.1	OS: HPE 480GB or 400GB SSD, Intermediate/Shuffle/ Temp Data/ Distributed FS: 1 x 800GB SSD, 24 x HPE 600GB 6G SAS 10k HDD (except 2 x 1.2TB HDD replacement)
Cloudera Manager Node #1	HDFS Balancer/HDFS Namenode/Hive Gateway/Cloudera Management Services Alert Publisher/Clouder a Management Services Event Server/Hive Metastore Server/Hue Server/Cloudera Management Services/YARN JobHistory Server/YARN ResourceManager/ ZooKeeper Server/Spark Gateway/Spark History	1	N	pan-01	• HPE DL360 Gen10 Server • HW/SW Config (Intel(R) Xeon(R) Gold 5218 CPU @ 2.30GHz, 64) • Memory: 192GB • Storage: see storage • Network: HPE Eth 10/25Gb 2P 640SFP28 • OS: RHEL 7.5 • Cloudera CDH 5.16.1	1x480GB SSD for OS, 1x800GB SSD for TPCx-BB Kit run files storage.
Cloudera Manager Node #2	Hive Gateway/HiveServ er2/ZooKeeper Server	1	N	pan-02	• HPE DL360 Gen10 Server • HW/SW Config Intel(R) Xeon(R) Gold 5218 CPU @ 2.30GHz, 64) • Memory: 192GB • Storage: see storage • Network: HPE 560 SFP+10G NIC • OS: RHEL 7.5 • Cloudera CDH 5.16.1	1x480GB SSD for OS, 1x400GB SSD for TPCx-BB run files storage.
Cloudera Manager Node #3	HDFS SecondaryNameN ode/Hive Gateway/Cloudera Management Service Activity Monitor/ZooKeep er Server	1	N	pan-03	• HPE DL360 Gen10 Server • HW/SW Config (Intel(R) Xeon(R) Gold 5218 CPU @ 2.30GHz, 64) • Memory: 192GB • Storage: 1 x 800GB SSD • Network: HPE 560 SFP+10G NIC • OS: RHEL 7.5 • Cloudera CDH 5.16.1	1x400GB SSD for OS

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) 5.16.1 (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	1.6.0
MapReduce	2.6.0
ZooKeeper	3.4.5

2.4 Frameworks

Frameworks and Engine used in the benchmark should be disclosed.

Framework	Version
CDH	5.16.1
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 Report Summary (Performance Run)**

```
**********
TPCx-BB
Result
v1.2
**********
INFO: T_LOAD = 2088.241
INFO: T_LD = 0.1 * T_LOAD: 208.8241
INFO: T_PT = 10822.5410086123
INFO: T_T_PUT = 89620.205
INFO: T_TT = 8962.0205
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 = 1789.75057336083
```

- **Run2 Report Summary (Repeatability Run)**

```
**********
TPCx-BB
Result
v1.2
**********
INFO: T_LOAD = 2042.185
INFO: T_LD = 0.1 * T_LOAD: 204.2185
INFO: T_PT = 10820.131815946
INFO: T_T_PUT = 89390.08
INFO: T_TT = 8939.008
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 = 1793.02254740899
```

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.

Type	Query	Power	Stream 1	Stream 2	Stream 3	Stream 4	Stream 5
Structured	1	206.720	862.790	812.071	721.817	846.014	810.822
	6	521.266	2,883.775	2,531.468	2,568.889	2,444.161	2,715.734
	7	385.043	2,323.674	2,162.147	2,161.003	2,123.497	2,081.445
	9	381.192	1,895.952	2,142.065	1,963.814	1,904.470	1,276.402
	11	122.596	495.830	494.089	537.281	501.254	543.258
	13	193.140	937.745	965.936	924.334	1,121.531	979.368
	14	71.991	269.162	282.523	252.621	272.597	258.727
	15	132.208	397.652	429.082	383.406	308.614	319.392
	16	586.299	3,593.235	3,390.686	3,232.057	3,430.097	3,368.665
	17	306.093	931.895	884.735	945.943	937.456	825.689
	20	339.702	1,208.308	1,105.688	1,042.276	1,089.583	1,264.200
	21	585.710	2,624.056	2,836.457	2,962.783	2,794.596	2,791.406
	22	95.974	254.285	139.123	230.449	202.283	215.549
	23	155.428	316.628	298.610	333.406	323.148	357.437
	24	186.794	869.772	962.232	1,036.051	936.519	870.975
	25	468.315	2,526.593	2,179.094	2,876.168	2,359.924	3,128.860
	26	285.156	983.769	811.184	799.689	838.682	861.276
	29	220.276	1,006.357	1,029.928	951.549	1,032.212	995.507
Semi-structured	2	1,958.425	12,715.709	13,068.889	12,508.710	13,050.640	13,244.333
	3	983.923	6,388.101	6,716.164	6,158.359	6,740.480	6,847.151
	4	1,571.924	11,100.933	11,159.677	11,221.016	10,781.888	10,185.306
	5	585.832	3,167.528	3,163.484	3,223.844	3,132.780	3,187.345
	8	616.201	4,175.215	3,508.960	4,394.706	3,429.560	3,629.182
	12	329.626	1,839.650	1,691.846	1,751.204	1,887.481	1,911.926
	30	1,858.999	16,065.975	16,513.532	16,384.620	16,779.379	16,268.330
Unstructured	10	345.169	1,101.180	1,298.034	993.186	1,281.059	1,810.824
	18	1,858.293	5,457.653	5,789.087	5,321.518	5,357.025	5,822.524
	19	578.958	1,087.209	1,429.966	1,422.420	1,356.100	1,142.552
	27	64.750	162.200	187.768	198.927	180.726	203.768
	28	309.560	770.572	684.121	688.551	749.026	749.655

Type	Query	Stream 6	Stream 7	Stream 8	Stream 9	Stream 10
Structured	1	894.293	704.786	203.170	762.771	962.503
	6	2,869.740	2,504.121	2,695.783	2,634.598	2,702.072
	7	2,100.787	2,360.668	2,147.310	2,217.463	2,176.461
	9	1,976.845	1,825.312	2,031.485	1,871.161	2,301.613
	11	490.108	572.753	901.342	578.580	540.008
	13	963.142	869.575	1,019.075	917.865	980.080
	14	224.411	256.588	280.339	224.675	278.603
	15	323.108	276.165	398.021	336.087	284.782
	16	3,508.825	3,281.635	3,274.599	3,930.961	3,521.318
	17	810.267	569.301	678.143	829.994	897.499
	20	1,124.190	1,156.867	332.795	1,056.282	1,291.387
	21	2,614.662	2,732.223	2,849.163	2,553.934	2,686.270
	22	230.317	186.241	198.666	210.881	197.985
	23	401.098	278.967	157.033	288.818	361.304
	24	1,057.704	1,046.631	1,066.616	969.223	971.628
	25	2,670.513	2,558.724	2,801.084	2,242.975	2,576.802
	26	787.270	761.925	736.262	796.779	825.557
	29	1,255.067	1,004.123	949.148	960.244	1,022.137
Semi-structured	2	12,701.148	12,770.103	12,519.285	12,665.377	11,842.439
	3	6,775.622	6,339.044	6,759.836	6,829.001	6,227.394
	4	11,192.240	12,054.847	10,876.109	10,516.227	10,638.500
	5	3,046.760	3,207.853	3,359.224	3,093.460	3,213.493
	8	3,702.582	4,001.986	3,918.609	3,657.412	3,893.087
	12	1,587.598	1,686.074	1,668.525	1,889.716	1,767.526
	30	16,084.625	16,849.570	18,680.520	17,178.645	17,111.719
Unstructured	10	1,303.342	1,082.889	1,239.845	1,394.122	1,200.022
	18	5,665.332	5,806.906	5,753.948	5,588.455	5,805.538
	19	1,149.156	1,192.324	1,326.910	1,380.939	1,335.286
	27	191.707	178.769	148.637	182.938	217.746
	28	848.879	604.260	648.708	793.265	674.904

3.5 Validation Test Output

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

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

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/conf/engineSettings.conf
- bigBench-configs/hive/conf/engineSettings.sql
- bigBench-configs/hive/population/hiveCreateLoad.sql
- bigBench-configs/hive/queries/q01/engineLocalSettings.sql
- bigBench-configs/hive/queries/q02/engineLocalSettings.sql
- bigBench-configs/hive/queries/q03/engineLocalSettings.sql
- bigBench-configs/hive/queries/q04/engineLocalSettings.sql
- bigBench-configs/hive/queries/q05/engineLocalSettings.sql
- bigBench-configs/hive/queries/q06/engineLocalSettings.sql
- bigBench-configs/hive/queries/q07/engineLocalSettings.sql
- bigBench-configs/hive/queries/q08/engineLocalSettings.sql
- bigBench-configs/hive/queries/q09/engineLocalSettings.sql
- bigBench-configs/hive/queries/q10/engineLocalSettings.sql
- bigBench-configs/hive/queries/q11/engineLocalSettings.sql
- bigBench-configs/hive/queries/q12/engineLocalSettings.sql
- bigBench-configs/hive/queries/q13/engineLocalSettings.sql
- bigBench-configs/hive/queries/q14/engineLocalSettings.sql
- bigBench-configs/hive/queries/q15/engineLocalSettings.sql
- bigBench-configs/hive/queries/q16/engineLocalSettings.sql
- bigBench-configs/hive/queries/q17/engineLocalSettings.sql
- bigBench-configs/hive/queries/q18/engineLocalSettings.sql
- bigBench-configs/hive/queries/q19/engineLocalSettings.sql
- bigBench-configs/hive/queries/q20/engineLocalSettings.sql
- bigBench-configs/hive/queries/q21/engineLocalSettings.sql
- bigBench-configs/hive/queries/q22/engineLocalSettings.sql
- bigBench-configs/hive/queries/q23/engineLocalSettings.sql
- bigBench-configs/hive/queries/q24/engineLocalSettings.sql
- bigBench-configs/hive/queries/q25/engineLocalSettings.sql
- bigBench-configs/hive/queries/q26/engineLocalSettings.sql
- bigBench-configs/hive/queries/q27/engineLocalSettings.sql
- bigBench-configs/hive/queries/q28/engineLocalSettings.sql
- bigBench-configs/hive/queries/q29/engineLocalSettings.sql
- bigBench-configs/hive/queries/q30/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 included 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)
7	1	400	2,800
15	1	480	7,200
17	24	600	244,800
1	22	600	13,200
19	1	800	15,200
1	2	1,200	2,400

Total Storage (GB)	285,600
Scale Factor	10000
Data Storage Ratio	28.56

4.5 Scale Factor to Memory Ratio

The Scale Factor to memory ratio must be disclosed.

Nodes	Memory (GB)	Total (GB)
3	192	576
18	384	6,912

Scale Factor	10000
Total Memory (GB)	7,488
SF / Memory Ratio	1.34

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 1,789.75

5.2 Repeatability Run Metric

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

Repeatability Run
BBQpm@10000 1,793.02

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 510.19

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
10

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 06:00:14.038	108,014.038
Run 2	01 05:55:51.820	107,751.820

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	2,088.241	2,042.185
Power Test	16,305.591	16,319.553
Throughput Test	89,620.205	89,390.080

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

April 29, 2019

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

Platform:	Hewlett Packard Enterprise ProLiant DL for Big Data (w/ 18x HPE ProLiant DL380 Gen 10, 3x HPE ProLiant DL360 Gen10)
Operating System:	Red Hat Enterprise Linux Server 7.6
Apache Hadoop Compatible Software:	Cloudera for Apache Hadoop (CDH) 5.16.1

The results were:

Performance Metric	**1,789.75 BBQpm@10000GB**
Run Elapsed Time	01 06:00:14.038 (108,014.038 Seconds)

Cluster	**18x HPE ProLiant DL380 Gen 10 (Data nodes), 3x HPE ProLiant DL360 Gen10 (Management nodes)**		
CPUs	2 x Intel Xeon Gold 6254 (3.10 GHz, 18-core, 24.75 MB L3) (Data nodes) 2 x Intel Xeon Gold 5218 (2.30 GHz, 16-core, 22 MB L3) (Mgmt. nodes)		
Memory	384GB (Data nodes), 192GB (Mgmt. nodes)		
Storage	***Qty***	***Size***	***Type***
	7	400GB	6G SATA SSD (OS; 2 Mgmt. nodes, 5 Data nodes)
	15	480GB	6G SATA SSD (OS/misc, 2 Mgmt. nodes, 13 Data nodes)
	430	600GB	6G SAS 10K HDD (Data, Data nodes)
	19	800GB	6G SATA SSD (Temp, Data nodes, 1 Mgmt. node)
	2	1.2TB	SAS 10K HDD (Data, Data nodes)

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

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

- 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.

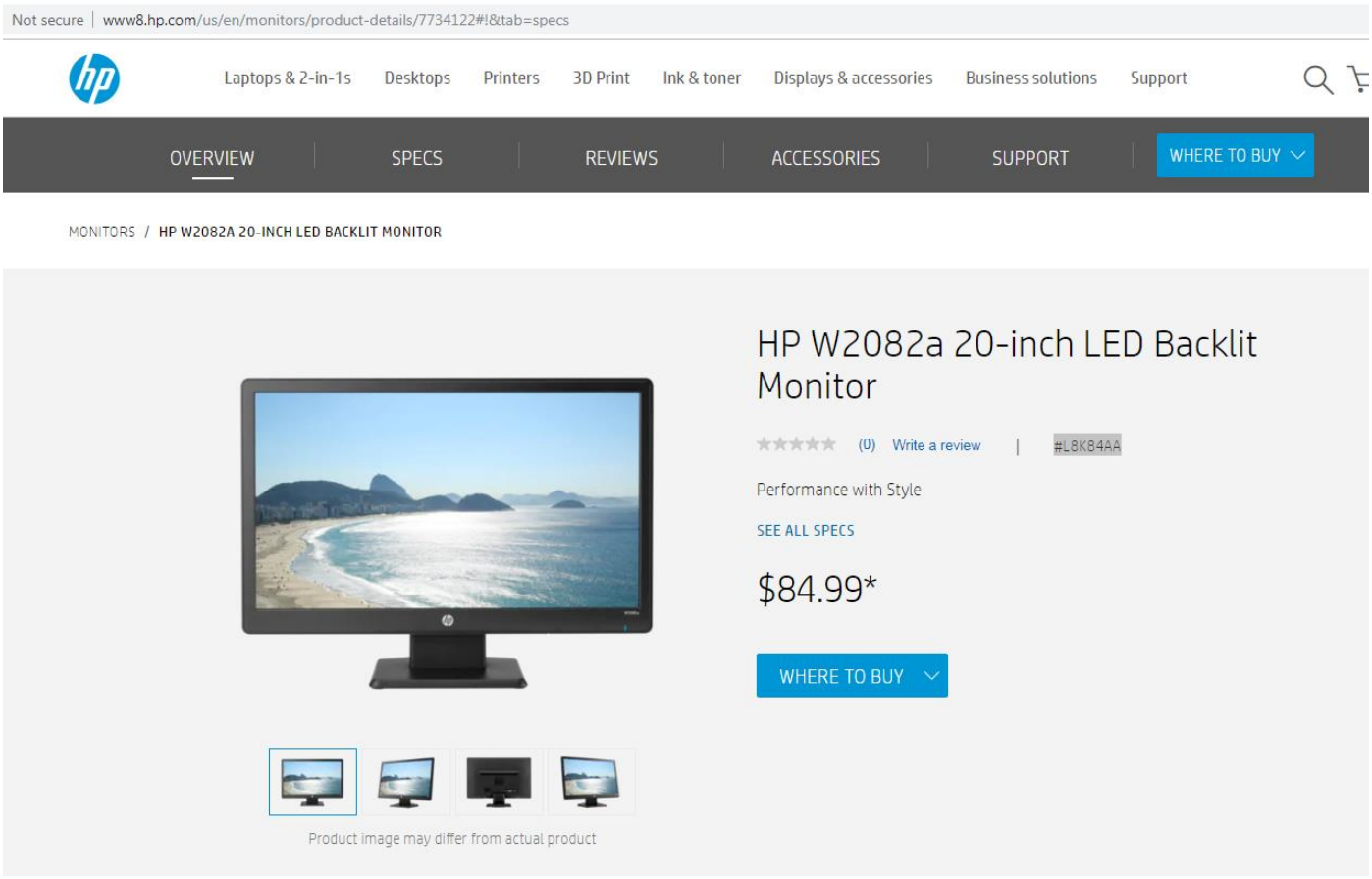
Respectfully Yours,

Doug Johnson, TPC Auditor

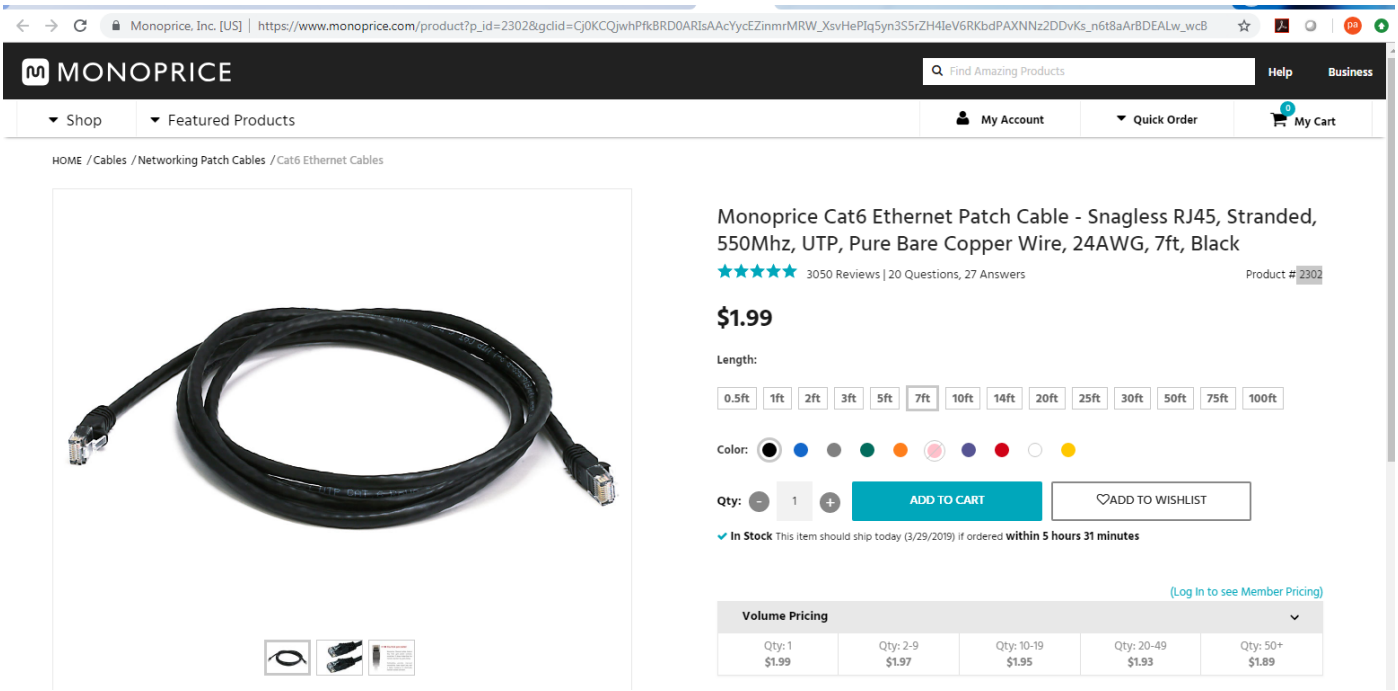
63 Lourdes Dr. | Leominster, MA 01453 | 978-343-6562 | www.sizing.com

Third Party Price Quotes

store.hp.com



monoprice.com



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-CLX-5-2019\
Validation Run Files	Supporting-Files-10TB-CLX-5-2019\logs-20190412-200803-hive-sf10000-Validation
Performance Run Files	Supporting-Files-10TB-CLX-5-2019\logs-20190415-163919-hive-sf10000-run2
Repeatability Run Files	Supporting-Files-10TB-CLX-5-2019\logs-20190414-055938-hive-sf10000-run1
Clause 3 - Workload Related Items	
Benchmark Generic Parameters	Supporting-Files-10TB-CLX-5-2019\logs-20190415-163919-hive-sf10000-run2\bigBench-configs\conf\userSettings.conf
Query Parameters used in the benchmark execution Settings	Supporting-Files-10TB-CLX-5-2019\logs-20190415-163919-hive-sf10000-run2\bigBench-configs\hive\conf\queryParameters.sql
Benchmark Global Framework Parameters Settings	Supporting-Files-10TB-CLX-5-2019\logs-20190415-163919-hive-sf10000-run2\bigBench-configs\hive\conf\engineSettings.sql
Benchmark Global Framework Parameters Settings	Supporting-Files-10TB-CLX-5-2019\logs-20190415-163919-hive-sf10000-run2\bigBench-configs\hive\conf\engineSettings.conf
Load Test script	Supporting-Files-10TB-CLX-5-2019\logs-20190415-163919-hive-sf10000-run2\bigBench-configs\hive\population\hiveCreateLoad.sql
Queries specific optimization parameters settings	Supporting-Files-10TB-CLX-5-2019\logs-20190415-163919-hive-sf10000-run2\bigBench-configs\hive\queries\q[01-30]\engineLocalSettings.conf
Queries specific optimization parameters settings	Supporting-Files-10TB-CLX-5-2019\logs-20190415-163919-hive-sf10000-run2\bigBench-configs\hive\queries\q[01-30]\engineLocalSettings.sql
Clause 4 - SUT Related Items	
Data Redundancy report	Supporting-Files-10TB-CLX-5-2019\data-redundancy-report.txt
Benchmark execution script	Supporting-Files-10TB-CLX-5-2019\run-all.sh
Hardware and Software Report from a representative node	Supporting-Files-10TB-CLX-5-2019\logs-20190415-163919-hive-sf10000-run2\logs-20190331-135757-hive-sf10000\run-logs\envInfo-pan-08
All Framework configuration files are included in the Supporting Files Archive	Supporting-Files-10TB-CLX-5-2019\logs-20190415-163919-hive-sf10000-run2\logs-20190331-135757-hive-sf10000\run-logs\envInfo-pan-08\envInfo-pan-08\hadoop
	Supporting-Files-10TB-CLX-5-2019\logs-20190415-163919-hive-sf10000-run2\logs-20190331-135757-hive-sf10000\run-logs\envInfo-pan-08\envInfo-pan-08\hive
	Supporting-Files-10TB-CLX-5-2019\logs-20190415-163919-hive-sf10000-run2\logs-20190331-135757-hive-sf10000\run-logs\envInfo-pan-08\envInfo-pan-08\spark
Clause 5 - Metric and Scale Factor Related Items	
Benchmark Performance Report	Supporting-Files-10TB-CLX-5-2019\logs-20190415-163919-hive-sf10000-run2\logs-20190331-135757-hive-sf10000\run-logs\BigBenchResult.log
Validation Test Report	Supporting-Files-10TB-CLX-5-2019\logs-20190412-200803-hive-sf10000-Validation\run-logs\BigBenchResult.log