Sugon

Dawning Information Industry Co., Ltd.

TPC Express Benchmark™ Big Bench (TPCx-BB)

Full Disclosure Report

for

Sugon Cluster

(with 33x I620-G30)

using

Cloudera for Apache Hadoop (CDH) 5.11.1

and

Red Hat Enterprise Linux Server 7.3

First Edition

January 22, 2018

Dawning Information Industry Co., Ltd. (Sugon)*, 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.

Sugon and the Sugon Logo are trademarks of Dawning Information Industry Co., Ltd. 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 Sugon and any other company.

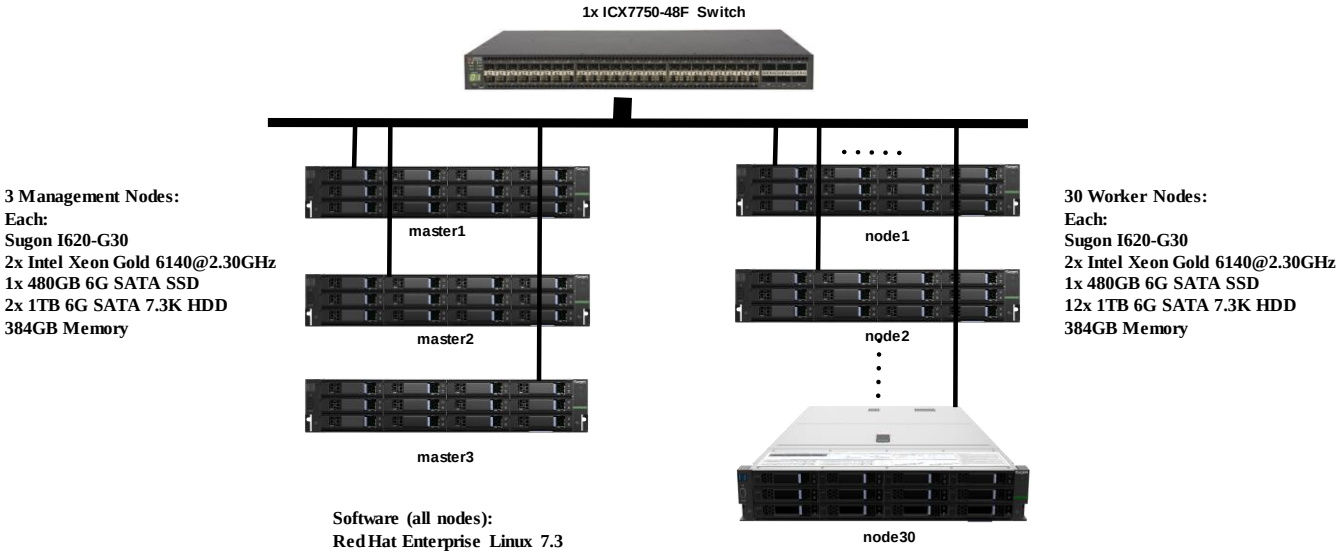
TPC Benchmark™, TPCx-BB and BBQpm, are registered certification marks of the Transaction Processing Performance Council.

The Sugon 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 Sugon business contact for information on the products or services available in your area. You can find additional information via Sugon's web site at www.sugon.com. Actual performance and environmental costs of Sugon products will vary depending on individual customer configurations and conditions.

Sugon	**Sugon Cluster**	TPCx-BB Rev. v1.2.0 TPC-Pricing Rev. v2.2.0 Report Date: January 22, 2018
Total System Cost	TPCx-BB Performance Metric	Price/Performance
1,041,770 USD	**3,383.95** BBQpm@30000	**307.86 USD** $/BBQpm@30000

Framework	Operating System	Other Software	Availability Date	Scale Factor	Streams
Cloudera for Apache Hadoop (CDH) 5.11.1	Red Hat Enterprise Linux Server 7.3	None	January 22, 2018	30000	4

System Configuration



Physical Storage/Scale Factor: 12.73	Scale Factor/Physical Memory: 2.37

Servers:	33x I620-G30
Total Processors/Cores/Threads	66/1,188/2,376
Server Configuration:	**Per I620-G30:**
Processors	2x Intel Xeon Gold 6140 @ 2.30 GHz
Memory	384 GB
Storage Controller	LSI 9361 RAID Card, 2GB SAS 12G (worker nodes only)
Storage Device	1x 480GB 6G SATA SSD 2x 1TB 6G SATA 7.3K RPM HDD (management nodes only) 12x 1TB 6G SATA 7.3K RPM HDD (worker nodes only)
Network	Intel 10-Gigabit SFI/SFP+ (dual port)
Connectivity:	Brocade ICX7750-48F Switch

Description	Part Number	Source	Unit Price ($)	Qty	Extended Price ($)	3 Year Maint. Price ($)
Server Hardware - Management Nodes						
I620-G30(with P20-D312 chassis)	61006780	1	5,737.28	3	17,211.84	
Intel Xeon Gold 6140	62000614	1	2,993.33	6	17,959.98	
32GB DDR4 2Rx8 2667MHz Memories	50000383	1	427.95	36	15,406.20	
Network Card, 2x10GE Port(82599)	24000829	1	228.97	3	686.91	
480 GB, 2.5inch 6Gb/s SSD	54000554	1	347.82	3	1,043.46	Included
1 TB, 7.2k 3.5inch 6Gb/s SATA	54000544	1	91.79	6	550.74	
4*3.5" Hard Disk Back plane	24001494	1	30.26	3	90.78	
550W Platinum AC power supply unit	33000186	1	63.21	6	379.26	
8038 8K Fan Module	3001379	1	5.90	12	70.80	
			SubTotal		53,399.97	Included
Server Hardware - Worker Nodes						
I620-G30(with P20-D312 chassis)	61006781	1	6,746.53	30	202,395.90	
Intel Xeon Gold 6140	62000614	1	2,993.33	60	179,599.80	
32GB DDR4 2Rx8 2667MHz Memories	50000383	1	427.95	360	154,062.00	
Network Card, 2x10GE Port(82599)	24000829	1	228.97	30	6,869.10	
480 GB, 2.5inch 6Gb/s SSD	54000554	1	347.82	30	10,434.60	
1 TB, 7.2k 3.5inch 6Gb/s SATA	54000544	1	91.79	360	33,044.40	Included
4*3.5" Hard Disk Back plane	24001494	1	30.26	90	2,723.40	
2*2.5" Rear Hard Disk Back plane Module	3000609	1	2.44	30	73.20	
LSI 9361 RAID Card, 2GB SAS 12G	24001041	1	295.77	30	8,873.10	
550W Platinum AC power supply unit	33000186	1	63.21	60	3,792.60	
8038 8K Fan Module	3001379	1	5.90	120	708.00	
			SubTotal		602,576.10	Included

{continues next page}

Sugon	Sugon Cluster	TPCx-BB Rev. v1.2.0 TPC-Pricing Rev. v2.2.0 Report Date: January 22, 2018

Description	Part Number	Source	Unit Price ($)	Qty	Extended Price ($)	3 Year Maint. Price ($)
Network						
BROCADE ICX7750-48F	96002001	1	26,962.50	1	26,962.50	Included
Avago AFBR-709SMZ 10G 850nm SFP+,Multi-mode Module	3000807	1	24.10	66	1,590.60	
			SubTotal		28,553.10	Included
Monitor/Keyboard/Mouse (includes 2 spares)						
Lenovo 60F5MAR6US Thinkvision T2224Z	0E6-009A-001M8	2	209.00	3	627.00	
Lenovo Professional Wireless Keyboard and Mouse Combo	9SIA6ZP5M99142	2	59.99	3	179.97	
			SubTotal		806.97	
Server Software						
Cloudera Ent Basic Ed 1y r 24x7		1	2,304.00	99	228,096.00	Included
RHEL Svr 2 Sckt/2 Gst 3y r 24x7 E-LTU		1	3,889.00	33	128,337.00	
			SubTotal		356,433.00	Included
			Total Extended Price		1,041,769.14	0.00

Pricing:1 = Sugon; 2 = newegg

(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	**$1,041,770**
BBQpm@30000	**3,383.95**
$/BBQpm@30000	**$ 307.86**

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	30000
Streams	4
SUT Validation Test	PASS

Performance Run (Run 2)

Overall Run Start Time	2017-12-11 17:34:06.821
Overall Run End Time	2017-12-12 21:18:04.834
Overall Run Elapsed Time	99,838.013
Load Test Start Time	2017-12-11 17:34:06.822
Load Test End Time	2017-12-11 19:19:25.921
Load Test Elapsed Time	6,319.099
Power Test Start Time	2017-12-11 19:19:25.924
Power Test End Time	2017-12-12 02:05:48.207
Power Test Elapsed Time	24,382.283
Throughput Test Start Time	2017-12-12 02:05:48.208
Throughput Test End Time	2017-12-12 21:18:04.834
Throughput Test Elapsed Time	69,136.626
Performance Metric (BBQpm@ 30000)	3,383.95

Repeatability Run (Run 1)

Overall Run Start Time	2017-12-10 13:04:02.994
Overall Run End Time	2017-12-11 16:35:03.876
Overall Run Elapsed Time	99,060.882
Load Test Start Time	2017-12-10 13:04:02.995
Load Test End Time	2017-12-10 14:50:56.042
Load Test Elapsed Time	6,413.047
Power Test Start Time	2017-12-10 14:50:56.045
Power Test End Time	2017-12-10 21:28:35.474
Power Test Elapsed Time	23,859.429
Throughput Test Start Time	2017-12-10 21:28:35.474
Throughput Test End Time	2017-12-11 16:35:03.876
Throughput Test Elapsed Time	68,788.402
Performance Metric (BBQpm@ 30000)	3,435.53

Performance Run Report (Run 2)

****
TPCx-BB
Result
v1.2
****
INFO: T_LOAD = 6319.099
INFO: T_LD = 0.1 * T_LOAD: 631.9099
INFO: T_PT = 13589.2217300351
INFO: T_T_PUT = 69136.626
INFO: T_TT = 17284.1565
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@30000 = 3383.95742649002

Repeatability Run Report (Run 1)

****
TPCx-BB
Result
v1.2
****
INFO: T_LOAD = 6413.047
INFO: T_LD = 0.1 * T_LOAD: 641.3047
INFO: T_PT = 13217.848388304
INFO: T_T_PUT = 68788.402
INFO: T_TT = 17197.1005
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@30000 = 3435.53763619344

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 30000 with 33 nodes (I620-G30) running Cloudera for Apache Hadoop (CDH) 5.11.1 on Red Hat Enterprise Linux Server 7.3.

Measured Configuration

Company Name	Cluster Node	Virtualization	Operating System
Dawning Information Industry Co., Ltd.	33x I620-G30	n/a	Red Hat Enterprise Linux Server 7.3

TPC Express Benchmark© Big Bench Metrics

Total System Cost	BBQpm@30000	Price/Performance	Availability Date
1,041,770 USD	3,383.95	307.86 USD	January 22, 2018

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 Dawning Information Industry Co., Ltd.

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.

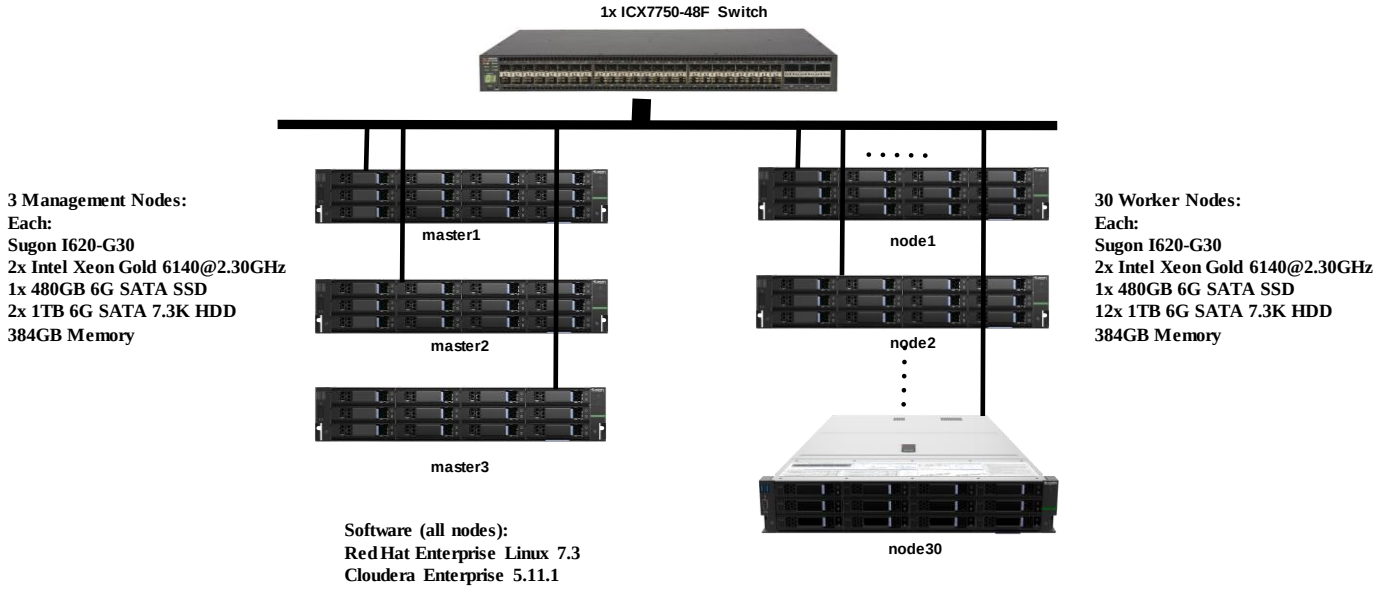
The Supporting Files Archive contains the parameters and options used to configure the components involved in this benchmark.

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: 33
- Total Processors/Cores/Threads: 66/1,188/2,376
- Total Memory: 12,672
- Total Number of Storage Drives/Devices: 399
- Total Storage Capacity: 381,840
- Network: Brocade ICX7750-48F Switch

33x I620-G30, each with:

- Processors/Cores/Threads: 2/36/72
- Processor Model: 2x Intel Xeon Gold 6140 @ 2.30 GHz
- Memory: 384
- Controller: LSI 9361 RAID Card, 2GB SAS 12G (worker nodes only)
- Drives:
    - 1x 480GB 6G SATA SSD
    - 2x 1TB 6G SATA 7.3K RPM HDD (management nodes only)
    - 12x 1TB 6G SATA 7.3K RPM HDD (worker nodes only)
- Network: Intel 10-Gigabit SFI/SFP+ (dual port)

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
Cloudera Manager Node #1	HDFS Gateway/Hive Gateway/Cloudera Management Services/Spark Gateway/YARN Gateway/ZooKeeper Server	1	N	master1	Sugon I620-G30 HW/SW Config (Intel Xeon Gold 6140, 2, 2.3GHz, 72) Memory: 384 GB Storage: 1x 480GB SSD, 2x 1TB HDD Network: 1x Intel 82599ES SFP+10G OS: RHEL 7.3 Cloudera CDH: 5.11.1	OS:480GB 6G SATA SSD Data Drive: 2x 1TB 6G 7200RPM SATA HDD
Cloudera Manager Node #2	HDFS Balancer/HDFS NameNode/Hive Gateway/Hive Metastore Server/HiveServer2/Spark Gateway/Spark History Server/YARN JobHistory Server/YARN ResourceManager/ZooKeeper Server	1	N	master2	Sugon I620-G30 HW/SW Config (Intel Xeon Gold 6140, 2, 2.3GHz, 72) Memory: 384 GB Storage: 1x 480GB SSD, 2x 1TB HDD Network: 1x Intel 82599ES SFP+10G OS: RHEL 7.3 Cloudera CDH: 5.11.1	OS:480GB 6G SATA SSD Data Drive: 2x 1TB 6G 7200RPM SATA HDD
Cloudera Manager Node #3	HDFS SecondaryNameNode/Hive Gateway/Spark Gateway/ZooKeeper Server	1	N	master3	Sugon I620-G30 HW/SW Config (Intel Xeon Gold 6140, 2, 2.3GHz, 72) Memory: 384 GB Storage: 1x 480GB SSD, 2x 1TB HDD Network: 1x Intel 82599ES SFP+10G OS: RHEL 7.3 Cloudera CDH: 5.11.1	OS:480GB 6G SATA SSD Data Drive: 2x 1TB 6G 7200RPM SATA HDD
Worker	HDFS DataNode/Hive Gateway/Spark Gateway/YARN NodeManager	30	N	node[1-30]	Sugon I620-G30 HW/SW Config (Intel Xeon Gold 6140, 2, 2.3GHz, 72) Memory: 384 GB Storage: 1x 480GB SSD, 12x 1TB HDD Network: 1x Intel 82599ES SFP+10G OS: RHEL 7.3 Cloudera CDH: 5.11.1	OS:480GB 6G SATA SSD Distributed FS/Intermediate/Shuffle/Temp Data:12x1TB 6G 7200RPM SATA HDD

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) 5.11.1 (fully HDFS compatible at the API level).

2.3 Engine Implementation

The Engine implementation and corresponding version must be disclosed.

Component	Version
HDFS	2.6.0
Hive	1.1.0
Spark	1.6.0
YARN	2.6.0
ZooKeeper	3.4.5
MapReduce	2.6.0

2.4 Frameworks

Frameworks and Engine used in the benchmark should be disclosed.

Framework	Version
CDH	5.11.1
HDFS	2.6.0
Hive	1.1.0
Spark	1.6.0
YARN	2.6.0
ZooKeeper	3.4.5
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 (Repeatability Run)**

  ****
  TPCx-BB
  Result
  v1.2
  ****
  INFO: T_LOAD = 6413.047
  INFO: T_LD = 0.1 * T_LOAD: 641.3047
  INFO: T_PT = 13217.848388304
  INFO: T_T_PUT = 68788.402
  INFO: T_TT = 17197.1005
  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@30000 = 3435.53763619344

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

  ****
  TPCx-BB
  Result
  v1.2
  ****
  INFO: T_LOAD = 6319.099
  INFO: T_LD = 0.1 * T_LOAD: 631.9099
  INFO: T_PT = 13589.2217300351
  INFO: T_T_PUT = 69136.626
  INFO: T_TT = 17284.1565
  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@30000 = 3383.95742649002

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
Structured	1	133.158	255.487	180.071	1,104.237	333.995
	6	504.012	656.059	844.632	712.873	1,219.731
	7	370.298	463.633	1,720.400	885.443	799.310
	9	347.772	1,131.905	995.954	1,897.420	3,944.746
	11	123.150	435.689	362.893	535.046	564.942
	13	221.368	1,137.759	568.006	1,576.759	646.508
	14	278.161	1,097.541	623.588	1,263.022	2,342.686
	15	107.316	239.354	433.422	248.101	361.611
	16	577.547	1,377.532	1,883.888	1,619.155	2,887.277
	17	179.367	1,050.261	2,214.842	725.179	436.065
	20	322.937	619.130	901.955	1,898.498	1,454.915
	21	916.687	3,077.213	1,498.286	2,448.447	3,489.843
	22	86.536	428.922	440.837	854.301	170.340
	23	337.928	1,025.759	626.049	821.483	500.375
	24	169.824	191.470	370.304	592.443	552.750
	25	551.980	2,621.670	1,820.608	3,702.129	4,949.731
	26	324.266	2,572.952	2,216.127	875.265	1,055.565
	29	204.510	858.270	995.889	1,028.848	1,187.764
Semi-structured	2	3,408.481	9,712.466	5,838.373	8,365.362	3,946.605
	3	1,985.344	2,787.241	3,952.044	3,270.789	4,780.645
	4	3,211.655	6,277.117	6,908.631	9,737.176	6,840.865
	5	658.920	1,231.037	3,270.363	2,356.776	1,530.171
	8	1,154.754	3,425.060	3,076.195	1,805.951	2,067.724
	12	1,068.570	4,296.789	2,739.306	3,483.993	4,077.008
	30	2,943.335	8,946.995	6,652.814	10,162.359	9,921.674
Unstructured	10	389.970	980.304	700.419	728.114	951.551
	18	2,181.756	5,498.452	7,692.190	3,698.366	5,275.237
	19	1,304.706	2,141.179	1,591.260	2,141.251	2,249.248
	27	95.562	99.703	176.792	167.615	222.577
	28	222.373	514.255	2,437.042	430.203	262.348

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/bigBench.properties
- bigBench-configs/conf/userSettings.conf
- bigBench-configs/hive/conf/engineSettings.conf
- bigBench-configs/hive/conf/engineSettings.sql
- bigBench-configs/hive/queries/q{01 – 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 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)
33	1	480	15,840
3	2	1,000	6,000
30	12	1,000	360,000

Total Storage (GB)	381,840
Scale Factor	30000
Data Storage Ratio	12.73

4.5 Scale Factor to Memory Ratio

The Scale Factor to memory ratio must be disclosed.

Nodes	Memory (GB)	Total (GB)
33	384	12,672

Scale Factor	30000
Total Memory (GB)	12,672
SF / Memory Ratio	2.37

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@30000 3,383.95

5.2 Repeatability Run Metric

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

Repeatability Run
BBQpm@30000 3,435.53

5.3 Price-Performance Metric

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

Price / Performance
$BBQpm@30000 307.86

5.4 Scale Factor

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

Scale Factor
30000

5.5 Stream Count

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

Streams
4

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 03:31:00.882	99,060.882
Run 2	01 03:43:58.013	99,838.013

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	6,319.099	6,413.047
Power Test	24,382.283	23,859.429
Throughput Test	69,136.626	68,788.402

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. Yvan Tian
Sugon Building, No. 36 Z-Park
No. 8 Dongbeiwang West Rd.
Haidian District, Beijing 100193
CHN

Jan 22, 2018

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

Platform:	Sugon Cluster (w/ 33x I620-G30)
Operating System:	Red Hat Enterprise Linux Server 7.3
Apache Hadoop Compatible Software:	Cloudera for Apache Hadoop (CDH) 5.11.1

The results were:

Performance Metric	**3,383.95 BBQpm@30000GB**
Run Elapsed Time	01 03:43:58.013 (99,838.013 Seconds)

Cluster	**33x I620-G30 Servers**		
CPUs	2 x Intel Xeon Gold 6140 (2.30 GHz, 18-core, 24.75 MB L3)		
Memory	384 GB		
Storage	***Qty***	***Size***	***Type***
	1	480 GB	6G SATA SSD (OS, all nodes)
	2	1 TB	6G SATA 7200 rpm HDD (Management nodes)
	12	1 TB	6G SATA 7200 rpm HDD (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
- The tested Scale Factor (30000GB) was confirmed to be valid for publication
- All validation queries executed successfully and produced compliant results

63 Lourdes Dr. | Leominster, MA 01453 | 978-343-6562 | 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:

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.

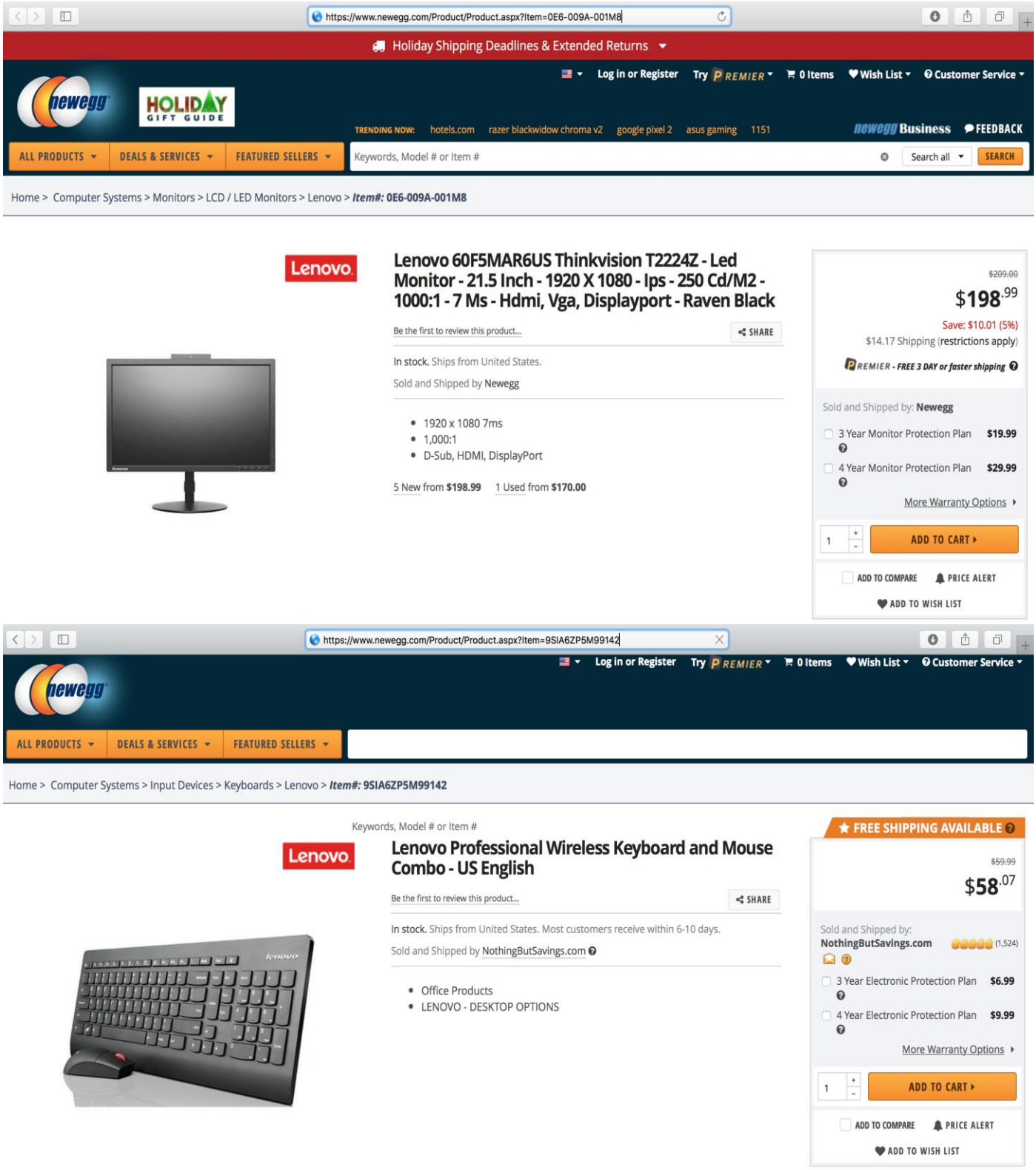
Respectfully Yours,

Doug Johnson, TPC Auditor

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

Third Party Price Quotes

Newegg



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	Support-Files-for-Sugon-33nodes-30TB
Validation Run Files	Support-Files-for-Sugon-33nodes-30TB\Validation-Run-logs-20171210-122003-hive-sf30000
Performance Run Files	Support-Files-for-Sugon-33nodes-30TB\Performance-Run-logs-20171212-213156-hive-sf30000
Repeatability Run Files	Support-Files-for-Sugon-33nodes-30TB\Repeatability-Run-logs-20171211-164903-hive-sf30000
Clause 3 - Workload Related Items	
Benchmark Generic Parameters	Support-Files-for-Sugon-33nodes-30TB\Performance-Run-logs-20171212-213156-hive-sf30000\bigBench-configs\conf\userSettings.conf
Query Parameters used in the benchmark execution Settings	Support-Files-for-Sugon-33nodes-30TB\Performance-Run-logs-20171212-213156-hive-sf30000\bigBench-configs\hive\conf\queryParameters.sql
Benchmark Global Framework Parameters Settings	Support-Files-for-Sugon-33nodes-30TB\Performance-Run-logs-20171212-213156-hive-sf30000\bigBench-configs\hive\conf\engineSettings.sql
Benchmark Global Framework Parameters Settings	Support-Files-for-Sugon-33nodes-30TB\Performance-Run-logs-20171212-213156-hive-sf30000\bigBench-configs\hive\conf\engineSettings.conf
Load Test script	Support-Files-for-Sugon-33nodes-30TB\Performance-Run-logs-20171212-213156-hive-sf30000\bigBench-configs\hive\population\hiveCreateLoad.sql
Queries specific optimization parameters settings	Support-Files-for-Sugon-33nodes-30TB\Performance-Run-logs-20171212-213156-hive-sf30000\bigBench-configs\hive\queries\q[01-30]\engineLocalSettings.conf
Queries specific optimization parameters settings	Support-Files-for-Sugon-33nodes-30TB\Performance-Run-logs-20171212-213156-hive-sf30000\bigBench-configs\hive\queries\q[01-30]\engineLocalSettings.sql
Clause 4 - SUT Related Items	
Data Redundancy report	Support-Files-for-Sugon-33nodes-30TB\hdfs-data-redundancy-report.txt
Benchmark execution script	Support-Files-for-Sugon-33nodes-30TB\publication-runs.sh
Hardware and Software Report from a representative node	Support-Files-for-Sugon-33nodes-30TB\Performance-Run-logs-20171212-213156-hive-sf30000\run-logs\envInfo-node1\envInfo.log
All Framework configuration files are included in the Supporting Files Archive	Support-Files-for-Sugon-33nodes-30TB\Performance-Run-logs-20171212-213156-hive-sf30000\run-logs\envInfo-node1\hadoop
	Support-Files-for-Sugon-33nodes-30TB\Performance-Run-logs-20171212-213156-hive-sf30000\run-logs\envInfo-node1\hive
	Support-Files-for-Sugon-33nodes-30TB\Performance-Run-logs-20171212-213156-hive-sf30000\run-logs\envInfo-node1\spark
Clause 5 - Metric and Scale Factor Related Items	
Benchmark Performance Report	Support-Files-for-Sugon-33nodes-30TB\Performance-Run-logs-20171212-213156-hive-sf30000\run-logs\BigBenchResult.log
Validation Test Report	Support-Files-for-Sugon-33nodes-30TB\Validation-Run-logs-20171210-122003-hive-sf30000\run-logs\BigBenchResult.log