Dell Inc.

TPC Express Benchmark™ Big Bench (TPCx-BB)

Full Disclosure Report

for

Dell PowerEdge R7515

(with 1x PowerEdge R7515; 10x PowerEdge R7515)

using

Cloudera CDP 7.1.4

and

SUSE Linux Enterprise Server 12 SP5

First Edition

March 15, 2021

***Dell Inc. (Dell)**, 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.

Dell and the Dell Logo are trademarks of Dell Inc. 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 Dell and any other company.

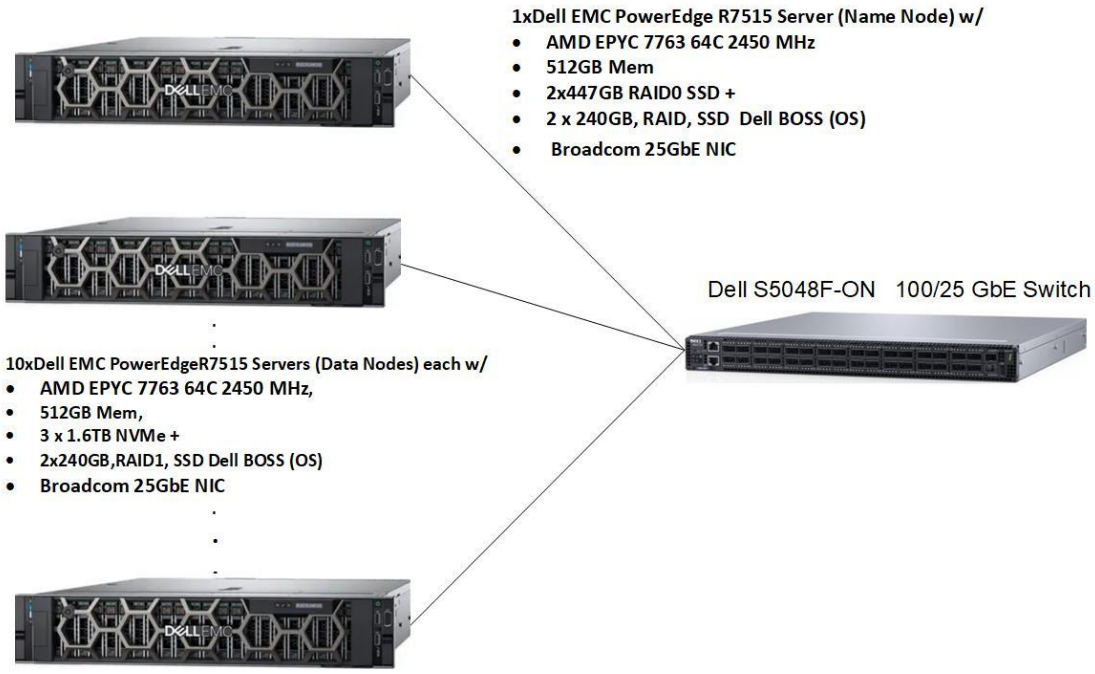
TPC Benchmark™, TPCx-BB and BBQpm, are registered certification marks of the Transaction Processing Performance Council.

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

DELL EMC	**Dell PowerEdge R7515**	TPCx-BB Rev. 1.5.0 TPC-Pricing Rev. 2.7.0 Report Date: March 15, 2021
Total System Cost	TPCx-BB Performance Metric	Price/Performance
753,292 USD	**1,544.13** BBQpm@3000	**487.85 USD** $/BBQpm@3000

Framework	Operating System	Other Software	Availability Date	Scale Factor	Streams
Cloudera CDP 7.1.4	SUSE Linux Enterprise Server 12 SP5	None	March 15, 2021	3000	7

System Configuration



Physical Storage/Scale Factor: 18.08	Scale Factor/Physical Memory: 0.53	Main Data Redundancy Model: 3-way replication

Servers: Total Processors/Cores/Threads	11x PowerEdge R7515 / PowerEdge R7515 11/704/1,408	
Server Configuration: Processors Memory Storage Controller Storage Device Network Controller	**1x PowerEdge R7515 (Name Node):** 1x AMD EPYC 7763 64-Core Processor 512 GiB HBA330 12Gbps SAS HBA 2x 240 GB BOSS M.2 SSD 2x 480 GB SATA SSD Broadcom 25 GbE NIC	**10x PowerEdge R7515 (Data Node):** 1x AMD EPYC 7763 64-Core Processor 512 GiB n/a 2x 240 GB BOSS M.2 SSD 3x 1.6 TB NVMe+ Broadcom 25 GbE NIC
Connectivity:	Dell S5048F-ON 100/25 GbE Switch	

Description	Part Number	Source	Unit Price	Qty	Extended Price	3 Yr. Maint. Price
HARDWARE COMPONENTS						
PowerEdge R7515 -Worker Nodes		1	$61,361.38	10	$613,613.80	
2.5 Chassis	379-BDTF	1		10		
NVMe Backplane	379-BDSX	1		10		
PowerEdge R7515 Server	210-ASVQ	1		10		
No Trusted Platform Module	461-AADZ	1		10		
2.5" Chassis with up to 24 NVMe Drives	321-BERW	1		10		
AMD 7763 2.45GHz,64C/128T,256M,280W,3200	338-BZRR	1		10		
Standard Heatsink	412-AASE	1		10		
Performance Optimized	370-AAIP	1		10		
3200MT/s RDIMMs	370-AEVR	1		10		
C30, No RAID for NVME chassis	780-BCDO	1		10		
No Controller	405-AACD	1		10		
CFI,STANDARD OPTION NOT SELECTED	692-BBBB	1		10		
Performance BIOS Settings	384-BBBL	1		10		
UEFI BIOS Boot Mode with GPT Partition	800-BBDM	1		10		
No Additional Mid Fan	384-BBSO	1		10		
Dual, Hot-plug, Redundant Power Supply (1+1), 1100W	450-ADWM	1		10		
Riser Config 2, 2 x 16 FH + 2 x 16 LP PCIe slot	330-BBNL	1		10		
PowerEdge R7515 Motherboard,MLK	384-BCUN	1		10		
iDRAC9,Enterprise 15G	385-BBOT	1		10		
PowerEdge R7515 Motherboard, with 2x1Gb Onboard LOM (BCM5720) V2	384-BCNQ	1		10		
PowerEdge 2U Standard Bezel	350-BBWP	1		10		
BOSS controller card + with 2 M.2 Sticks 240G (RAID 1),FH	403-BCHP	1		10		
No Quick Sync	350-BBKU	1		10		
iDRAC,Legacy Password	379-BCSG	1		10		
iDRAC Group Manager, Enabled	379-BCQV	1		10		
Enterprise Linux OS, Non Factory Installed, Requires Subscription	605-BBFL	1		10		
No Media Required	605-BBFN	1		10		
ReadyRails Sliding Rails	770-BBBQ	1		10		
No Internal Optical Drive	429-AAIQ	1		10		
No Systems Documentation, No OpenManage DVD Kit	631-AACK	1		10		
PowerEdge R7515 Shipping	340-CMZG	1		10		
PowerEdge R7515 Ship Material	340-CODN	1		10		
PowerEdge R7515 CCC Marking, No CE Marking	343-BBPQ	1		10		
GCP Operations Management	929-8509	1		10		
Basic Next Business Day 36 Months	709-BBFM	1		10		Included
Prosupport Plus and 4Hr Mission Critical Initial, 36 Month(s)	865-BBNF	1		10		Included
On-Site Installation Declined	900-9997	1		10		
64GB RDIMM, 3200MT/s, Dual Rank	370-AEVP	1		80		
1.6TB Enterprise NVMe Mixed Use AG Drive U.2 Gen4 with carrier	400-BKGF	1		30		
Jumper Cord - C13/C14, 4M, 250V, 12A (North America, Guam, North	492-BBDV	1		20		
Broadcom 57414 Dual Port 10/25GbE SFP28 Adapter, Mezz Card	540-BCKU	1		10		
SUSE Linux Enterprise Server, 1-2 SKT w Unlimited VMs, L3 Subscription, 3	528-CHFF	1		10		

(continued on next page)

(continued from previous page)

Description	Part Number	Source	Unit Price	Qty	Extended Price	3 Yr. Maint. Price
PowerEdge R7515 - Master Node		1	$53,965.14	1	$53,965.14	
2.5 Chassis	379-BDTF	1		1		
SAS/SATA Backplane	379-BDSS	1		1		
PowerEdge R7515 Server	210-ASVQ	1		1		
No Trusted Platform Module	461-AADZ	1		1		
2.5" Chassis with up to 24 Hot Plug Hard Drives	321-BERX	1		1		
AMD 7763 2.45GHz,64C/128T,256M,280W,3200	338-BZRR	1		1		
Standard Heatsink	412-AASE	1		1		
Performance Optimized	370-AAIP	1		1		
3200MT/s RDIMMs	370-AEVR	1		1		
No RAID	780-BCDI	1		1		
HBA330 12Gbps SAS HBA Controller (NON-RAID), Minicard	405-AAJU	1		1		
Performance BIOS Settings	384-BBBL	1		1		
UEFI BIOS Boot Mode with GPT Partition	800-BBDM	1		1		
High Performance Fan	750-AAWT	1		1		
Dual, Hot-plug, Redundant Power Supply (1+1), 1100W	450-ADWM	1		1		
Riser Config 2, 2 x 16 FH + 2 x 16 LP PCIe slot	330-BBNL	1		1		
PowerEdge R7515 Motherboard,MLK	384-BCUN	1		1		
iDRAC9,Enterprise 15G	385-BBOT	1		1		
PowerEdge R7515 Motherboard, with 2x1Gb Onboard LOM (BCM5720) V2	384-BCNQ	1		1		
PowerEdge 2U Standard Bezel	350-BBWP	1		1		
BOSS controller card + with 2 M.2 Sticks 240G (RAID 1),FH	403-BCHP	1		1		
No Quick Sync	350-BBKU	1		1		
iDRAC,Legacy Password	379-BCSG	1		1		
iDRAC Group Manager, Enabled	379-BCQV	1		1		
Enterprise Linux OS, Non Factory Installed, Requires Subscription	605-BBFL	1		1		
No Media Required	605-BBFN	1		1		
ReadyRails Sliding Rails	770-BBBQ	1		1		
No Internal Optical Drive	429-AAIQ	1		1		
No Systems Documentation, No OpenManage DVD Kit	631-AACK	1		1		
PowerEdge R7515 Shipping	340-CMZG	1		1		
PowerEdge R7515 Ship Material	340-CODN	1		1		
PowerEdge R7515 CCC Marking, No CE Marking	343-BBPQ	1		1		
GCP Operations Management	929-8509	1		1		
Basic Next Business Day 36 Months	709-BBFM	1		1		Included
Prosupport Plus and 4Hr Mission Critical Initial, 36 Month(s)	865-BBNF	1		1		Included
On-Site Installation Declined	900-9997	1		1		
64GB RDIMM, 3200MT/s, Dual Rank	370-AEVP	1		8		
480GB SSD SATA Read Intensive 6Gbps 512 2.5in Hot-plug AG Drive,	400-AXTV	1		2		
Jumper Cord - C13/C14, 4M, 250V, 12A (North America, Guam, North	492-BBDV	1		2		
Broadcom 57414 Dual Port 10/25GbE SFP28 Adapter, PCIe Low Profile	540-BBVK	1		1		
SUSE Linux Enterprise Server, 1-2 SKT w Unlimited VMs, L3 Subscription, 3	528-CHFF	1		1		

(continued on next page)

(continued from previous page)

Description	Part Number	Source	Unit Price	Qty	Extended Price	3 Yr. Maint. Price
PowerSwitch S5248-ON		1	$41,049.37	1	$41,049.37	
Dell EMC S5248F-ON Switch, 48x25GbE SFP28, 4x100GbE QSFP28,	210-APEX	1		1		
Dell EMC S52XX-ON Series User Guide	343-BBLP	1		1		
OS10 Enterprise, S5248F-ON	634-BRUN	1		1		
Dell Hardware Limited Warranty 1 Year	818-4856	1		1		Included
ProSupport Plus:Mission Critical 4-Hour 7x24 On-Site Service with Emerger	818-4898	1		1		Included
ProSupport Plus Mission Critical:7x24 HW/SW Technical Support and Assis	818-4901	1		1		Included
ProSupport Plus:Mission Critical 4-Hour 7x24 On-Site Service with Emerger	818-4902	1		1		Included
Dell Limited Hardware Warranty Extended Year(s)	975-3461	1		1		Included
Info 3rd Party Software Warranty provided by Vendor	997-6306	1		1		
On-Site Installation Declined	900-9997	1		1		
Dell Networking, Cable, SFP+ to SFP+, 10GbE, Active Optical (Optics include	470-ABME	1		11		
Power Cord, 125V, 15A, 10 Feet, NEMA 5-15/C13	450-AAFH	1		1		
Power Cord, 125V, 15A, 10 Feet, NEMA 5-15/C13	450-AAFH	1		1		
APC NetShelter SX Deep Enclosure - Rack - cabinet - black - 24U - 18-inch	A7522217	1	$1,129.99	1	$1,129.99	
APC Basic Rack PDU AP9567 - 0U - 120V NEMA 5-15 Input / (14) NEMA 5-15 Output	A7541362	1	$189.99	1	$189.99	
24 Inch LED monitor VE248Q - Widescreen Full HD Monitor	A6732050	1	$136.99	1	$136.99	
Logitech MK200 Media Keyboard and Mouse Combo	A6859396	1	$24.99	1	$24.99	
HARDWARE COMPONENTS				**Subtotal**	$710,110.27	
SOFTWARE COMPONENTS						
Cloudera CDP - 3 year 24x7 support		1	$40,800.00	11	$448,800.00	
SOFTWARE COMPONENTS				**Subtotal**	$448,800.00	
TOTAL						**$1,158,910.27**
Large Purchase Discount (35%)*						-$405,618.59

Pricing:1 = {Source 1}; 2 = {Source 2}

(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	**$753,292**
BBQpm@3000	**1,544.13**
$/BBQpm@3000	**$ 487.85**

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	3000
Streams	7
SUT Validation Test	PASS

Performance Run (Run 1)

Overall Run Start Time	2021-03-09 06:59:44.024
Overall Run End Time	2021-03-09 15:36:16.763
Overall Run Elapsed Time	30,992.739
Load Test Start Time	2021-03-09 06:59:44.024
Load Test End Time	2021-03-09 07:07:37.124
Load Test Elapsed Time	473.100
Power Test Start Time	2021-03-09 07:07:37.126
Power Test End Time	2021-03-09 08:41:09.654
Power Test Elapsed Time	5,612.528
Throughput Test Start Time	2021-03-09 08:41:09.655
Throughput Test End Time	2021-03-09 15:36:16.763
Throughput Test Elapsed Time	24,907.108
Performance Metric (BBQpm@ 3000)	1,544.13

Repeatability Run (Run 2)

Overall Run Start Time	2021-03-09 15:58:49.796
Overall Run End Time	2021-03-10 00:33:45.551
Overall Run Elapsed Time	30,895.755
Load Test Start Time	2021-03-09 15:58:49.796
Load Test End Time	2021-03-09 16:06:42.485
Load Test Elapsed Time	472.689
Power Test Start Time	2021-03-09 16:06:42.487
Power Test End Time	2021-03-09 17:40:33.047
Power Test Elapsed Time	5,630.560
Throughput Test Start Time	2021-03-09 17:40:33.047
Throughput Test End Time	2021-03-10 00:33:45.551
Throughput Test Elapsed Time	24,792.504
Performance Metric (BBQpm@ 3000)	1,547.29

Performance Run Report (Run 1)

```
************
TPCx-BB
Result
v1.5
************
INFO: T_LOAD = 473.1
INFO: T_LD = 0.1 * T_LOAD: 47.31
INFO: T_PT = 3344.72753844527
INFO: T_T_PUT = 24907.108
INFO: T_TT = 3558.15828571428
INFO: === Checking validity of the final result ===
INFO: OK: All required BigBench phases were performed.
INFO: OK: All 30 queries were running in the power test.
INFO: OK: All 30 queries were running in the first throughput test.
INFO: OK: Pretend mode was inactive. All commands were executed.
INFO: === Final result ===
INFO: VALID BBQpm@3000 = 1544.13535691466
```

Repeatability Run Report (Run 2)

```
************
TPCx-BB
Result
v1.5
************
INFO: T_LOAD = 472.689
INFO: T_LD = 0.1 * T_LOAD: 47.2689
INFO: T_PT = 3346.39272029584
INFO: T_T_PUT = 24792.504
INFO: T_TT = 3541.78628571428
INFO: === Checking validity of the final result ===
INFO: OK: All required BigBench phases were performed.
INFO: OK: All 30 queries were running in the power test.
INFO: OK: All 30 queries were running in the first throughput test.
INFO: OK: Pretend mode was inactive. All commands were executed.
INFO: === Final result ===
INFO: VALID BBQpm@3000 = 1547.29014415038
```

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 1.5.0.

The test was conducted at a Scale Factor of 3000 with 11 nodes (1x PowerEdge R7515, 10x PowerEdge R7515) running Cloudera CDP 7.1.4 on SUSE Linux Enterprise Server 12 SP5.

Measured Configuration

Company Name	Cluster Node	Virtualization	Operating System
Dell Inc.	1x PowerEdge R7515 10x PowerEdge R7515	n/a	SUSE Linux Enterprise Server 12 SP5

TPC Express Benchmark© Big Bench Metrics

Total System Cost	BBQpm@3000	Price/Performance	Availability Date
753,292 USD	1,544.13	487.85 USD	March 15, 2021

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.

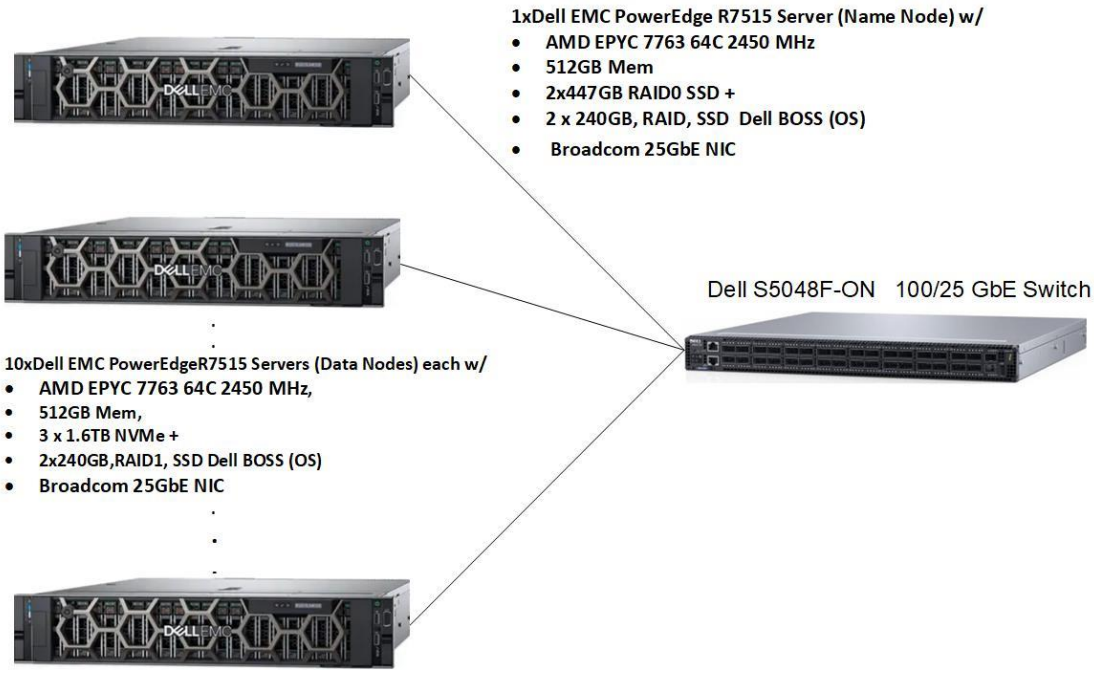
The Supporting Files Archive contains the parameters and options used to configure the components involved in this benchmark.

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



The measured configuration consisted of:

Total Nodes:	11
Total Processors/Cores/Threads:	11/704/1,408
Total Memory:	5,632
Total Number of Storage Devices:	54
Total Storage Capacity:	54,240
Network:	Dell S5048F-ON 100/25 GbE Switch

	1x PowerEdge R7515 (Name Node):	**10x PowerEdge R7515 (Data Node):**
Processors/Cores/Threads:	1/64/128	1/64/128
Processor Model:	1x AMD EPYC 7763 64-Core Processor	1x AMD EPYC 7763 64-Core Processor
Memory:	512 GiB	512 GiB
Storage Controller:	HBA330 12Gbps SAS HBA	n/a
Storage Devices:	2x 240 GB BOSS M.2 SSD 2x 480 GB SATA SSD	2x 240 GB BOSS M.2 SSD 3x 1.6 TB NVMe+
Network Controller:	Broadcom 25 GbE NIC	Broadcom 25 GbE NIC

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
Workers	HDFS DataNode Hive Gateway Spark Gateway Tez Getway YARN NodeManager	10	N	tpcxbb[2-11] Alias w[2-11]	• Dell PowerEdge R7515 • Processor: 1xAMD EPYC 7763 64-Core Processor • Memory : 512 GB • Storage: 3x 1.6TB NVMe(Data) + 2x240GB BOSS (OS) • Network: Broadcom Adv. Dual 25 GbE • OS: SUSE Linux Enterprise Server 12 SP5 • Cloudera Private Cloud Base (CDP) 7.1.4	OS: 240GB BOSS,SAS, SSD Intermediate/Shuffle /Temp Data Distributed FS: 3x1.6TB NVMe
Master	HDFS Balancer HDFS Namenode HDFS Secondary Namenode Hive Gateway Hive Metastore Server Hive on Tez Hive Server2 CMS Alert Publisher CMS Event Server CMS Host Monitor CMS Reports Monitor CMS Service Monitor YARN Queue Manager Store YARN Queue Manager Webapp Spark Gateway Spark History Server Tez Getway YARN JobHistory Server YARN ResourceManager Zookeeper Server	1	N	tpcxbb1 alias m1	• Dell PowerEdge R7515 • Processor: 1xAMD EPYC 7763 64-Core Processor • Memory : 512 GB • Storage: 2x447GB SSD + 2x240GB BOSS (OS) • Network: Broadcom Adv. Dual 25 GbE • OS: SUSE Linux Enterprise Server 12 SP5 • Cloudera Private Cloud Base (CDP) 7.1.4	OS: 240GB BOSS, SAS, SSD Metadata/Temp Data: 2x480GB, RAID0, SSD

2.2 Distributed File System Implementation

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

Cloudera CDP 7.1.4 (fully HDFS compatible at the API level).

2.3 Engine Implementation

The Engine implementation and corresponding version must be disclosed.

Component	Version
HDFS	3.1.1
YARN	3.1.1
MapReduce2	3.1.1
Spark2	2.4.0
Tez	0.9.1
Hive	3.1.0

2.4 Frameworks

Frameworks and Engine used in the benchmark should be disclosed.

Framework	Version
PvC Base	7.1.4
HDFS	3.1.1
YARN	3.1.1
MapReduce2	3.1.1
Spark2	2.4.0
Hive	3.1.3

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
v1.5

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
v1.5
**********
INFO: T_LOAD = 473.1
INFO: T_LD = 0.1 * T_LOAD: 47.31
INFO: T_PT = 3344.72753844527
INFO: T_T_PUT = 24907.108
INFO: T_TT = 3558.15828571428
INFO: === Checking validity of the final result ===
INFO: OK: All required BigBench phases were performed.
INFO: OK: All 30 queries were running in the power test.
INFO: OK: All 30 queries were running in the first throughput test.
INFO: OK: Pretend mode was inactive. All commands were executed.
INFO: === Final result ===
INFO: VALID BBQpm@3000 = 1544.13535691466
```

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

```
**********
TPCx-BB
Result
v1.5
**********
INFO: T_LOAD = 472.689
INFO: T_LD = 0.1 * T_LOAD: 47.2689
INFO: T_PT = 3346.39272029584
INFO: T_T_PUT = 24792.504
INFO: T_TT = 3541.78628571428
INFO: === Checking validity of the final result ===
INFO: OK: All required BigBench phases were performed.
INFO: OK: All 30 queries were running in the power test.
INFO: OK: All 30 queries were running in the first throughput test.
INFO: OK: Pretend mode was inactive. All commands were executed.
INFO: === Final result ===
INFO: VALID BBQpm@3000 = 1547.29014415038
```

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	Stream 6	Stream 7
Structured	1	102.790	710.654	531.274	745.765	100.145	511.804	1,629.931	332.454
	6	116.965	144.161	1,234.233	116.533	1,569.362	150.592	318.835	333.028
	7	41.229	234.819	2,868.781	466.703	335.749	1,296.175	1,867.152	29.444
	9	33.718	28.066	497.018	36.262	637.152	1,483.541	45.255	249.498
	11	32.505	70.159	128.508	29.374	1,999.001	559.303	727.220	700.910
	13	78.386	267.899	289.940	1,464.918	636.792	1,044.880	424.086	141.868
	14	29.756	361.193	1,174.492	25.970	38.429	496.657	156.216	175.256
	15	29.405	37.800	219.357	1,019.715	703.110	79.959	71.654	63.021
	16	66.171	117.049	133.210	1,598.838	1,004.321	120.900	89.225	212.856
	17	37.482	237.442	70.014	613.380	267.875	412.283	395.652	23.202
	20	432.522	861.718	664.273	428.393	997.455	607.849	486.324	927.780
	21	87.614	754.966	850.474	730.316	733.280	230.438	1,608.643	747.390
	22	29.716	757.284	71.594	30.795	1,479.225	1,178.499	70.023	203.208
	23	237.488	1,668.312	1,152.562	2,694.605	304.174	528.241	1,599.600	637.933
	24	35.709	180.611	1,019.683	116.641	1,714.445	1,435.541	79.401	1,662.401
	25	446.416	1,566.808	602.107	2,948.053	1,503.628	622.001	1,447.147	799.627
	26	578.314	826.228	802.571	932.756	618.695	677.628	1,845.553	1,093.717
	29	71.605	127.936	1,377.926	68.607	1,792.483	148.284	750.008	123.937
Semi-structured	2	690.726	845.651	809.968	1,103.877	698.074	699.825	851.174	700.179
	3	241.583	265.357	763.494	4,529.286	685.884	1,486.902	273.814	1,071.911
	4	379.164	420.537	388.014	387.575	456.127	416.149	2,771.687	688.819
	5	523.187	716.385	528.742	1,266.858	497.968	668.500	833.161	824.673
	8	231.066	3,923.941	1,441.176	234.264	977.044	1,295.804	678.166	2,645.005
	12	105.084	199.105	719.479	97.483	337.575	157.126	265.850	175.755
	30	398.364	531.684	640.336	915.449	430.649	714.093	439.271	2,335.792
Unstructured	10	152.183	727.221	158.519	719.958	352.907	180.003	1,217.677	183.012
	18	242.042	1,347.625	2,308.296	569.945	540.882	476.892	1,487.518	1,195.754
	19	61.489	2,725.246	362.180	60.558	76.059	70.295	230.791	97.112
	27	33.109	139.941	1,638.866	626.771	170.890	447.515	108.633	1,527.221
	28	66.710	517.674	80.027	327.448	318.181	75.970	280.168	82.375

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/hive/conf/engineSettings.conf
- bigBench-configs/hive/queries/q01/engineLocalSettings.sql
- bigBench-configs/hive/queries/q03/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/q13/engineLocalSettings.sql
- bigBench-configs/hive/queries/q14/engineLocalSettings.sql
- bigBench-configs/hive/queries/q15/engineLocalSettings.sql
- bigBench-configs/hive/queries/q17/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 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)
11	2	240	5,280
1	2	480	960
10	3	1,600	48,000

Total Storage (GB)	54,240
Scale Factor	3000
Data Storage Ratio	18.08

4.5 Scale Factor to Memory Ratio

The Scale Factor to memory ratio must be disclosed.

Nodes	Memory (GiB)	Total (GiB)
1	512	5,632

Scale Factor	3000
Total Memory (GB)	5,632
SF / Memory Ratio	0.53

Clause 5: Metrics and Scale Factors

5.1 Performance Run Metric

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

Performance Run
BBQpm@3000 1,544.13

5.2 Repeatability Run Metric

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

Repeatability Run
BBQpm@3000 1,547.29

5.3 Price-Performance Metric

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

Price / Performance
$BBQpm@3000 487.85

5.4 Scale Factor

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

Scale Factor
3000

5.5 Stream Count

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

Streams
7

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	00 08:36:32.739	30,992.739
Run 2	00 08:34:55.755	30,895.755

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	473.100	472.689
Power Test	5,612.528	5,630.560
Throughput Test	24,907.108	24,792.504

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
Certified Auditor

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

March 13, 2021

I verified the TPC Express Benchmark™ BB v1.5.0 performance of the following configuration:

Platform:	Dell PowerEdge R7515 (with 1x R7515, 10x R7515)
Operating System:	SUSE Linux Enterprise Server 12 SP5
Apache Hadoop Compatible Software:	Cloudera CDP 7.1.4

The results were:

Performance Metric	**1,544.13 BBQpm@3000GB**
Run Elapsed Time	00 08:36:32.739 (30,992.739 Seconds)

Cluster	**1x R7515 (Master Node), 10x R7515 (Worker Nodes)**		
CPUs	1x AMD EPYC 7763 64-Core Processor (2.45 GHz) (All Nodes)		
Memory	512GB (All Nodes)		
Storage	***Qty***	***Size***	***Type***
	2	240 GB	BOSS M.2 SSD (All Nodes)
	2	480 GB	SATA SSD (Master Node)
	3	1.6 TB	NVMe+ (Worker 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 v1.5.0.
- 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 (3000GB) 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:

None.

Respectfully Yours,

Doug Johnson, TPC Auditor

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

Third Party Price Quotes

All components are available directly through the Test Sponsor (Dell Inc.).

Supporting File Index

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

Description	Archive File Pathname
Clause 1 - General Items	
The Supporting Files Archive contains the parameters and options used to configure the components involved in this benchmark	Supporting-Files-3TB-R7515-03-2021\
Validation Run Files	Supporting-Files-3TB-R7515-03-2021\Validation-run logs
Performance Run Files	Supporting-Files-3TB-R7515-03-2021\Performance-run logs
Repeatability Run Files	Supporting-Files-3TB-R7515-03-2021\Repeatability-run logs
Clause 3 - Workload Related Items	
Benchmark Generic Parameters	Supporting-Files-3TB-R7515-03-2021\ Performance-run logs\bigBench-configs\conf\userSettings.conf
Query Parameters used in the benchmark execution Settings	Supporting-Files-3TB-R7515-03-2021\ Performance-run logs\bigBench-configs\hive\conf\queryParameters.sql
Benchmark Global Framework Parameters Settings	Supporting-Files-3TB-R7515-03-2021\ Performance-run logs\bigBench-configs\hive\conf\engineSettings.sql
Benchmark Global Framework Parameters Settings	Supporting-Files-3TB-R7515-03-2021\ Performance-run logs\bigBench-configs\hive\conf\engineSettings.conf
Load Test script	Supporting-Files-3TB-R7515-03-2021\ Performance-run logs\bigBench-configs \hive\population\ hiveCreateLoad.sql
Queries specific optimization parameters settings	Supporting-Files-3TB-R7515-03-2021\ Performance-run logs\bigBench-configs\hive\queries\q[01-30]\engineLocalSettings.conf
Queries specific optimization parameters settings	Supporting-Files-3TB-R7515-03-2021\ Performance-run-logs\bigBench-configs\hive\queries\q[01-30]\engineLocalSettings.sql
Clause 4 - SUT Related Items	
Data Redundancy report	Supporting-Files-3TB-R7515-03-2021\Performance-run logs\run-logs-data-redundancy-report.log
Benchmark execution script	Supporting-Files-3TB-R7515-03-2021\TPCxBB_Benchmarkrun.sh
Benchmark run script	Supporting-Files-3TB-R7515-03-2021\ TPCxBB_FullBenchmark_sequence_run
Hardware and Software Report from a representative node	Supporting-Files-3TB-R7515-03-2021\ Performance-run logs\run-logs\envInfo-tpcxbb02\envInfo.log
All Framework configuration files are included in the Supporting Files Archive	Supporting-Files-3TB-R7515-03-2021\Performance-run-logs\run-logs\envInfo-tpcxbb02\hadoop
	Supporting-Files-3TB-R7515-03-2021\Performance-run logs\run-logs\envInfo-tpcxbb02\hive
Clause 5 - Metric and Scale Factor Related Items	
Benchmark Performance Report	Supporting-Files-3TB-R7515-03-2021\Performance-run logs\run-logs\BigBenchResult.log
Validation Test Report	Supporting-Files-3TB-R7515-03-2021\Validation-run logs\run-logs\BigBenchResult.log