

TPC Express Benchmark™ HS Full Disclosure Report

PowerEdge R7515

(with 9x PowerEdge R7515 Servers; 1x PowerEdge R6525 Servers)

Running

Cloudera Enterprise Edition 6.3

on

Red Hat Enterprise Linux Server 7.6

TPCx-HS Version
Report Edition
Report Submitted

2.0.3
First
September 17, 2019

First Edition - September 2019

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 because of these and other factors. Therefore, the TPC Express Benchmark™ V 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 Express Benchmark™ HS, TPCx-HS, and HSph, 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.

Copyright© 2019 Dell Inc.

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

Abstract

This document contains the methodology and results of the TPC Express Benchmark™ HS (TPCx-HS) test conducted in conformance with the requirements of the TPCx-HS Standard Specification, Revision 2.0.3.






The benchmark results are summarized below.


Measured Configuration			
Company Name	Cluster Node	Hadoop Software	Operating System
Dell	PowerEdge R7515	Cloudera Enterprise Edition 6.3	Red Hat Enterprise Linux Server 7.6

TPC Express Benchmark™ HS Metrics			
Total System Cost	HSph@3TB	Price/Performance	Availability Date
\$447,883	11.62	\$38,544.15	October 15, 2019

Executive Summary

The [Executive Summary](#) follows on the next several pages.

	<h1>PowerEdge R7515</h1>		TPCx-HS 2.0.3
			TPC Pricing 2.4.0
			Report Date Sep. 17, 2019
Availability Date October 15, 2019	TPCx-HS Performance 11.62 HSpH@3TB	Price/Performance \$38,544.15 \$ / HSpH@3TB	Total System Cost \$447,883 USD
System Under Test Configuration Overview			
Scale Factor 3	Hadoop Software Cloudera Enterprise Edition 6.3	Operating System Red Hat Enterprise Linux Server 7.6	Other Software N/A
<div style="display: flex; justify-content: space-around; align-items: flex-start;"> <div style="text-align: center;">  <p>...</p> <p>9x Dell PowerEdge R7515 Servers (Data Nodes)</p> <ul style="list-style-type: none"> • 1x AMD EPYC 7542 32-Core Processor • 512 GB (8x 64GB RDIMM 3200MT/s Dual Rank) • 2x240GB SSD SATA m.2 Hot-plug Drive • 1x Mellanox Dual Port 100GbE QSFP28 NIC • 3x Dell 1.6TB, NVMe, 2.5 SFF Drive <p>...</p>  <p>...</p>  <p>1x Dell PowerEdge R6525 Server (Master Node)</p> <ul style="list-style-type: none"> • 2x AMD EPYC 7502 32-Core Processor • 512 GB (8x 64GB RDIMM 3200MT/s Dual Rank) • 2x240GB SSD SATA 2.5in Hot-plug Drive • 1x Broadcom Adv. Dual 25Gb Ethernet NIC (cluster connectivity) • 1x Embedded Broadcom Gigabit NIC (external connectivity) </div> <div style="text-align: center;">  <p>Mellanox SN2700</p> </div> </div>			
Physical Storage/Scale Factor: 16.00		Scale Factor/Physical Memory: 0.60	
Total Number of Servers:		10 (9x PowerEdge R7515; 1x PowerEdge R6525)	
Total Processors/Cores/Threads:		11/352/704	
Server Configuration:	Per PowerEdge R7515	Per PowerEdge R6525	
Processors	1x AMD EPYC 7542 32-Core Processor	2x AMD EPYC 7502 32-Core Processor	
Memory	512 GiB	512 GiB	
Storage Controller	Perc H740P	Perc H740P	
Storage Device	2x 240GB SATA SSD m.2 3x Dell 1.6TB NVMe	2x 240GB SATA SSD m.2	
Network	Mellanox Dual Port 100GbE	Broadcom Adv. Dual 25Gb	
Connectivity:	Mellanox SN2700 (32-port, 100Gbe)		
Total Rack Units:	(9xR7515)+(1xR6525)+(1xMSN2700) = 18 + 1 + 1 = 20RU		

	<h1>PowerEdge R7515</h1>	TPCx-HS	2.0.3
		TPC Pricing	2.4.0
		Report Date	Sep. 17, 2019

Description	Part Number	Source	Unit Price	Qty	Extended Price	3 Yr. Maint. Price
HARDWARE COMPONENTS						
PowerEdge R7515				1	\$41,778.17	\$376,003.53
PowerEdge R7515 Server	210-ASVQ	1	\$0.00	1	\$0.00	
PowerEdge R6515/R7515 Motherboard, with 2 x No Trusted Platform Module	384-BCEL	1	\$0.00	1	\$0.00	
PowerEdge R7515 Shipping	461-AADZ	1	\$0.00	1	\$0.00	
PowerEdge R7515 Shipping Material,DAO w/CC Cert Marking	340-CODN,	1	\$0.00	1	\$0.00	
Standard Heatsink	343-BBNT	1	\$0.00	1	\$0.00	
3200MT/s RDIMMs	412-AASE	1	\$0.00	1	\$0.00	
Performance Optimized	370-AEVR	1	\$0.00	1	\$0.00	
No Media Required	370-AAIP	1	\$0.00	1	\$0.00	
iDRAC9, Express X5	605-BBFN	1	\$0.00	1	\$0.00	
iDRAC Group Manager, Disabled	385-BBOU	1	\$0.00	1	\$0.00	
iDRAC,Factory Generated Password	379-BCQY	1	\$0.00	1	\$0.00	
Single, Hot-plug Power Supply (1+0), 495W	379-BCSF	1	\$0.00	1	\$0.00	
NEMA 5-15P to C13 Wall Plug, 125 Volt, 15 AMP,	450-ADWP	1	\$0.00	1	\$0.00	
PowerEdge 2U Standard Bezel	450-AALV	1	\$0.00	1	\$0.00	
No Quick Sync	350-BBWP	1	\$0.00	1	\$0.00	
Performance BIOS Setting	350-BBKU	1	\$0.00	1	\$0.00	
ReadyRails Sliding Rails With Cable Management Arm	384-BBBL	1	\$0.00	1	\$0.00	
No Systems Documentation, No OpenManage DVD Kit	770-BBBR	1	\$0.00	1	\$0.00	
US Order	631-AAACK	1	\$0.00	1	\$0.00	
Basic Deployment Dell Server R Series 1U/2U	332-1286	1	\$0.00	1	\$0.00	
Chassis with up to 24x2.5" Drives	804-6747	1	\$0.00	1	\$0.00	
No Internal Optical Drive for x10 or greater HDD Chassis	379-BDTF	1	\$0.00	1	\$0.00	
Riser Config 2, 2 x 16 FH + 2 x 16 LP PCIe slot	429-AAIQ	1	\$0.00	1	\$0.00	
NVMe Backplane	330-BBNL	1	\$0.00	1	\$0.00	
No Additional Mid Fan	379-BDSX	1	\$0.00	1	\$0.00	
2.5" Chassis with up to 24 NVMe Drives	384-BBSO	1	\$0.00	1	\$0.00	
AMD EPYC 7542 2.90GHz, 2C/64T, 128M Cache (225W) DDR4-3200	321-BERW	1	\$0.00	1	\$0.00	
64GB RDIMM, 3200MT/s, Dual Rank	338-BTTY	1	\$0.00	1	\$0.00	
C30, No RAID for NVME chassis	370-AEVP	1	\$0.00	8	\$0.00	
No Hard Drive	780-BCDO	1	\$0.00	1	\$0.00	
Dell 1.6TB, NVMe, Mixed Use Express Flash, 2.5 SFF Drive, U.2, PM1725b with Carrier	400-ABHL	1	\$0.00	1	\$0.00	
Mellanox ConnectX-5 EX Dual Port 40/100GbE	400-BEFC	1	\$0.00	3	\$0.00	
BOSS controller card + with 2 M.2 Sticks 240G	540-BCIU	1	\$0.00	1	\$0.00	
No Controller	403-BCHM	1	\$0.00	1	\$0.00	
UEFI BIOS Boot Mode with GPT Partition	405-AAACD	1	\$0.00	1	\$0.00	
Keyboard and Optical Mouse, USB, Black, English	800-BBDM	1	\$0.00	1	\$0.00	
3 Years ProSupport Plus Mission Critical 4Hr Onsite Service	570-AAKV,	1	\$12.00	3	\$36.00	
Dell 24 Monitor - E2417H	827-1296,	1	\$5,299.00	9	\$47,691.00	
Mellanox MSN2700-CS2F, 32 Port QSFP 100Gbe	E2417H	1	\$169.99	3	\$509.97	
1m (3ft) Dell DAC-Q28-100G-1M Compatible 100G QSFP28 Passive Direct Attach Copper Twinax Cable	MSN2700-		\$35,913.00	1	\$35,913.00	
3m (10ft) Dell DAC-Q28-45FP28-25G-3M Compatible 100G QSFP28 to 4x25G SFP28 Passive Direct Attach Copper Breakout Cable	Q28-PC01		\$29.00	11	\$319.00	
C13 to C14, PDU Style, 12 AMP, 2 Feet (.6m) Power Cord, North America	Q-		\$93.00	3	\$279.00	
APC NetShelter SX 24U 600mm x 1070mm Deep Enclosure	492-BBDI	1	\$20.00	22	\$440.00	
Rack PDU, Basic, Zero U, 15A, 120V, 5-15 input, (14) 5-15 output	A7545498	1	\$1,129.99	1	\$1,129.99	
	A7541364	1	\$174.99	1	\$174.99	
Subtotal					\$414,805.48	\$47,691.00

(continued on next page)

	<h1>PowerEdge R7515</h1>	TPCx-HS	2.0.3
		TPC Pricing	2.4.0
		Report Date	Sep. 17, 2019

(continued from previous page)

PowerEdge R6525 Server

PowerEdge R6525 Server	210-ATCF	1	\$749.00	1	\$749.00
BOSS controller card + with 2 M.2 sticks 240G (RAID 1), LP	403-BCHM	1	\$879.00	1	\$879.00
PowerEdge R6525 Motherboard, with 2 x 1Gb Onboard LOM (BCM5720)	384-BCGI	1	\$0.00	1	\$0.00
Very High Performance Fan for 10 x 2.5" or 10 x 2.5" + 2 x 2.5" chassis	384-BCGH	1	\$399.00	1	\$399.00
Chassis with up to 10 x 2.5" Hot Plug Hard Drives (2CPU & XGM1)	321-BEUP	1	\$370.00	1	\$370.00
No Trusted Platform Module	461-AADZ	1	\$0.00	1	\$0.00
PowerEdge R6525 Shipping Label, FCC2, PSU, 800W, w/o BIS	343-BBOL	1	\$0.00	1	\$0.00
PowerEdge R6525 x4 or x10 Drive Shipping Material	340-COGL	1	\$99.00	1	\$99.00
AMD EPYC 7502 2.00GHz/2.55GHz, 32C/64T, 64M Cache (180W) DDR4-2666	338-BSWD	1	\$5,599.00	2	\$11,198.00
Heatsink for 1st CPU, greater than or equal to 180W	412-AASH	1	\$0.00	1	\$0.00
Heatsink for 2nd CPU, greater than or equal to 180W	412-AASJ	1	\$0.00	1	\$0.00
3200MT/s RDIMMs	370-AEVR	1	\$0.00	1	\$0.00
64GB RDIMM, 3200MT/s Dual Rank	370-AEVP	1	\$2,439.00	8	\$19,512.00
Performance Optimized	370-AAIP	1	\$0.00	1	\$0.00
Unconfigured RAID	780-BCDS	1	\$0.00	1	\$0.00
PERC H745 for Chassis up to 10 drives	405-AAUZ	1	\$1,199.00	1	\$1,199.00
No Media Required	421-5736	1	\$0.00	1	\$0.00
iDRAC9,Enterprise	385-BBKT	1	\$489.00	1	\$489.00
iDRAC Group Manager, Disabled	379-BCQY	1	\$0.00	1	\$0.00
iDRAC, Factory Generated Password	379-BCSF	1	\$0.00	1	\$0.00
Riser Config 2, 1 x 16 LP PCIe slot (CPU1), 2 x 16 LP PCIe slot (CPU2)	330-BBNR	1	\$349.00	1	\$349.00
No Internal Optical Drive	429-ABBF	1	\$0.00	1	\$0.00
Dual, Hot-plug, Redundant Power Supply (1+1), 1400W	450-AIQZ	1	\$1,699.00	1	\$1,699.00
No Bezel	350-BBBW	1	\$0.00	1	\$0.00
Dell EMC Luggage Tag (x8 or x10 chassis)	350-BBXP	1	\$0.00	1	\$0.00
No Quick Sync	350-BBXM	1	\$49.00	1	\$49.00
Performance BIOS Settings	384-BBBL	1	\$0.00	1	\$0.00
UEFI BIOS Boot Mode with GPT Partition	800-BBDM	1	\$0.00	1	\$0.00
ReadyRails Sliding Rails With Cable Management Arm	770-BCKT	1	\$189.00	1	\$189.00
No Systems Documentation, No OpenManage DVD Kit	631-AACK	1	\$0.00	2	\$0.00
US Order	332-1286	1	\$0.00	1	\$0.00
ProSupport and 4Hr Mission Critical, 36 Month(s)	[865-BBNB]	1	\$4,661.00	1	\$4,661.00
No Installation	[900-9997]	1	\$0.00	1	\$0.00
NEMA 5-15P to C13 Wall Plug, 125 Volt, 15 AMP, 10 Feet (3m), Power Cord, North America	450-AALV	1	\$0.00	2	\$0.00

Sub Total					\$37,180.00	\$4,661.00
------------------	--	--	--	--	-------------	------------

HARDWARE COMPONENTS					Subtotal	\$451,985.48	\$52,352.00
----------------------------	--	--	--	--	-----------------	--------------	-------------


SOFTWARE COMPONENTS							
Cloudera Enterprise Data Engineering (CDH 6.3), 3yr Node License	CEDEN-	1	\$14,769.24	10	\$147,692.40		
Red Hat Linux Registration Document, No Subscription	340-AVFG	1	\$0.00	10	\$0.00		
Enterprise Linux OS, Non Factory Installed, Requires Subscription Selection	605-BBFL	1	\$0.00	10	\$0.00		
Red Hat Enterprise Linux, 1-2SKT, Physical Node, 3YR Premium Subscription, with up to 1 Virtual Guest	605-BBH0	1	\$3,702.00	10	\$37,020.00		

SOFTWARE COMPONENTS					Subtotal	\$184,712.40	\$0.00
----------------------------	--	--	--	--	-----------------	--------------	--------

Total					\$636,697.88	\$52,352.00
Large Purchase Discount (35%)*					-222,844.26	-18,323.20

Pricing: 1 = Dell	Three-Year Cost of Ownership:	\$447,883
* Discount applies to all line items where Key = 1. Discount based upon total system cost as purchased by a regular customer.	HSph@3TB:	11.62
Audited by Doug Johnson, InfoSizing	\$/ HSph@3TB:	\$38,544.15

Prices used in TPC benchmarks reflect the actual prices a customer would pay for a one-time purchase of the stated Line Items. 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 Line Items. For complete details, see the pricing section of the TPC Benchmark Standard. If you find that the stated prices are not available according to these terms, please inform the TPC at pricing@tpc.org. Thank you.

	PowerEdge R7515	TPCx-HS 2.0.3
		TPC Pricing 2.4.0
		Report Date Sep. 17, 2019

Numerical Quantities		
Performance Run – Run 1		
Scale Factor		3TB
Run Start Time	2019-09-09 21:02:53.000	
Run End Time	2019-09-09 21:18:18.000	
Run Elapsed Time		929.000
HSGen Start Time	2019-09-09 21:02:54.000	
HSGen End Time	2019-09-09 21:06:11.000	
HSGen Elapsed Time		198.808
HSSort Start Time	2019-09-09 21:06:17.000	
HSSort End Time	2019-09-09 21:16:55.000	
HSSort Elapsed Time		639.108
HSValidate Start Time	2019-09-09 21:17:01.000	
HSValidate End Time	2019-09-09 21:18:18.000	
HSValidate Elapsed Time		78.278
Repeatability Run – Run 2		
Scale Factor		3TB
Run Start Time	2019-09-09 21:20:25.000	
Run End Time	2019-09-09 21:35:49.000	
Run Elapsed Time		927.000
HSGen Start Time	2019-09-09 21:20:26.000	
HSGen End Time	2019-09-09 21:23:47.000	
HSGen Elapsed Time		202.124
HSSort Start Time	2019-09-09 21:23:53.000	
HSSort End Time	2019-09-09 21:34:17.000	
HSSort Elapsed Time		625.454
HSValidate Start Time	2019-09-09 21:34:23.000	
HSValidate End Time	2019-09-09 21:35:49.000	
HSValidate Elapsed Time		86.889



PowerEdge R7515

TPCx-HS	2.0.3
TPC Pricing	2.4.0
Report Date	Sep. 17, 2019

Run Reports

Run Report for Performance Run – Run 1

TPCx-HS Performance Metric (HSph@SF) Report

Test Run 1 Details	Total Time =	929
	Total Size =	30000000000
	Scale-Factor =	3

TPCx-HS Performance Metric (HSph@SF): 11.6279

Run Report for Repeatability Run – Run 2

TPCx-HS Performance Metric (HSph@SF) Report

Test Run 2 Details	Total Time =	927
	Total Size =	30000000000
	Scale-Factor =	3

TPCx-HS Performance Metric (HSph@SF): 11.6504

	<h1>PowerEdge R7515</h1>	TPCx-HS 2.0.3 TPC Pricing 2.4.0 Report Date Sep. 17, 2019
---	--------------------------	---

Revision History

Date	Edition	Description
September 17, 2019	First	Initial Publication

Table of Contents

Abstract..... 3

Executive Summary 3

Table of Contents.....10

Clause 0 – Preamble11

 0.1 TPC Express Benchmark™ HS Overview.....11

Clause 1 – General Items12

 1.1 Test Sponsor12

 1.2 Parameter Settings12

 1.3 Configuration Diagrams12

 1.3.1 Measured Configuration.....13

 1.3.2 Priced Configuration13

 1.4 Dataset Distribution.....14

 1.5 Software Components Distribution.....14

Clause 2 – Workload Related Items.....15

 2.1 Hardware & Software Tunables15

 2.2 Run Report15

 2.3 Benchmark Kit Identification.....15

 2.4 Benchmark Kit Changes15

Clause 3 – SUT Related Items.....16

 3.1 Data Storage Ratio16

 3.2 Memory Ratio16

Clause 4 – Metrics Related Items17

 4.1 HSGen Time.....17

 4.2 HSSort Time17

 4.3 HSValidate Time17

 4.4 HSDataCheck Times17

 4.5 Performance & Price-Performance17

Auditor’s Information & Letter of Attestation18

Supporting Files Index21

Third-Party Price Quotes.....22

Clause 0 – Preamble

0.1 TPC Express Benchmark™ HS Overview

The TPC Express Benchmark™ HS (TPCx-HS) was developed to provide an objective measure of hardware, operating system and commercial Apache Hadoop File System API compatible software distributions, and to provide the industry with verifiable performance, price-performance and availability metrics. The benchmark models a continuous system availability of 24 hours a day, 7 days a week.

Even though the modeled application is simple, the results are highly relevant to hardware and software dealing with Big Data systems in general. TPCx-HS stresses both hardware and software including Hadoop run-time, Hadoop File-system API compatible systems and MapReduce layers. This workload can be used to assess a broad range of system topologies and implementation of Hadoop clusters. TPCx-HS can be used to assess a broad range of system topologies and implementation methodologies in a technically rigorous and directly comparable and vendor-neutral manner.

The TPCx-HS kit is available from the TPC (See www.tpc.org/tpcx-hs for more information). Users must sign-up and agree to the TPCx-HS User Licensing Agreement (ULA) to download the kit. Re-distribution of the kit is prohibited. All related work (such as collaterals, papers, derivatives) must acknowledge the TPC and include TPCx-HS copyright. The TPCx-HS Kit includes: TPCx-HS Specification document, TPCx-HS Users Guide documentation, shell scripts to set up the benchmark environment and Java code to execute the benchmark load.

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-HS models and represents Hadoop run-time and 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 component incorporated into the pricing structure;*
- *Configuration parameters and options for operating system and file system component incorporated into the pricing structure;*
- *Configuration parameters and options for any other software component incorporated into the pricing structure;*
- *Compiler optimization options.*

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 contain the parameters and options used to configure the components involved in this benchmark.

1.3 Configuration Diagrams

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.*

1.3.1 Measured Configuration

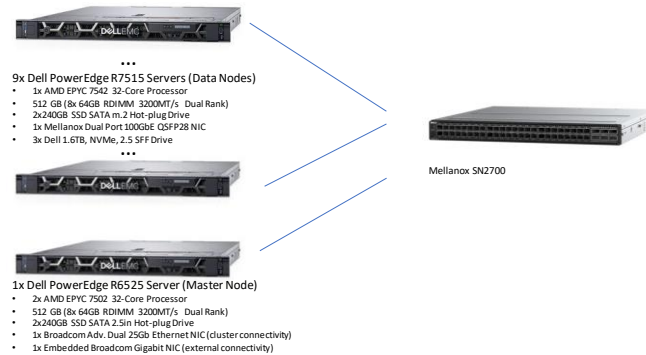


Figure 1-1 Measured Configuration

The measured configuration consisted of:

- Total Nodes: 10 (9x PowerEdge R7515; 1x PowerEdge R6525)
- Total Processors/Cores/Threads: 11/352/704
- Total Memory: 5.00TiB
- Total Number of Storage Drives/Devices: 47
- Total Storage Capacity: 48.00TB

Server node details:

- 9x PowerEdge R7515 Servers, each with:
 - Processors/Cores/Threads: 1/32/64
 - Processor Model: AMD EPYC 7542 32-Core Processor
 - Memory: 512 GiB
 - Controller: Perc H740P
 - Drives:
 - 2x 240GB SATA SSD m.2
 - 3x Dell 1.6TB NVMe
 - Network: Mellanox Dual Port 100GbE
- 1x PowerEdge R6525 Servers, each with:
 - Processors/Cores/Threads: 2/64/128
 - Processor Model: AMD EPYC 7502 32-Core Processor
 - Memory: 512 GiB
 - Controller: Perc H740P
 - Drives:
 - 2x 240GB SATA SSD m.2
 - Network: Mellanox Dual Port 100GbE

Network connectivity detail:

- Mellanox SN2700 (32-port, 100Gbe)

The distribution of software components over server nodes is detailed in section 1.5.

1.3.2 Priced Configuration

There are no differences between the priced configuration and the measured configuration.

1.4 Dataset Distribution

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

Table 1-1 describes the distribution of the dataset across all media in the system.

Server Node	Controller	Disk Drive	Description of Content
10	Boss controller card	sda	Operating System, Root, Swap, Hadoop Master
1-2	Boss controller card	sda	Operating System, Root, Swap, Hadoop Master
1-2	NVMe	nvme0n1,nvme1n1	Data, Temp
3-9	Boss controller card	sda	Operating System, Root, Swap, Hadoop Master
3-9	NVMe	nvme0n1,nvme1n1	Data, Temp

Table 1-1 Dataset Distribution

1.5 Software Components Distribution

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

Table 1-2 Describes the distribution of the software components across the system.

Node	Map/Reduce		HDFS		ZooKeeper	Spark
	Resource Manager	Node Manager	NameNode	DataNode	QuorumPeer	HistoryServer
10	X		X		X	X
1-2		X		X	X	
3-9		X		X		

Table 1-2 Software Component Distribution

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

Cloudera Enterprise Edition 6.3 (fully HDFS compatible at the API level).

Map/Reduce implementation and corresponding version must be disclosed.

Cloudera Enterprise Edition 6.3 (compatible equivalent to Hadoop 3.0.0).

Clause 2 – Workload Related Items

2.1 Hardware & Software Tunables

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

The Supporting File Archive contains all configuration scripts.

2.2 Run Report

The run report generated by TPCx-HS benchmark kit must be reported.

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

```
Run Report for Run 1 – Performance Run
=====
TPCx-HS Performance Metric (HSph@SF) Report

Test Run 1 Details      Total Time =                929
                        Total Size =                30000000000
                        Scale-Factor =                3

TPCx-HS Performance Metric (HSph@SF):                11.6279
=====
```

```
Run Report for Run 2 – Repeatability Run
=====
TPCx-HS Performance Metric (HSph@SF) Report

Test Run 2 Details      Total Time =                927
                        Total Size =                30000000000
                        Scale-Factor =                3

TPCx-HS Performance Metric (HSph@SF):                11.6504
=====
```

2.3 Benchmark Kit Identification

Version number of TPCx-HS kit and checksum for HSGen, HSSort and HSValidate Programs must be reported.

```
Kit Version                2.0.3

File                        MD5
-----
BigData_cluster_validate_suite.sh  57f7cd68251a9aba0feb6648630ff5da
HSDDataCheck.sh                bcf0b946a49d1249c9da174b5d9805f1
TPCx-HS-master_MR2.jar         492cbc51a1a60c28b43d96c79d08683d
TPCx-HS-master.sh              c619a0819571ecd00cd75d2b76ba8c64
```

2.4 Benchmark Kit Changes

HSDDataCheck.sh was modified at the auditor’s direction to collect fsck data.

Clause 3 – SUT Related Items

3.1 Data Storage Ratio

The data storage ratio must be disclosed.

Table 3-1 describes the details of the storage devices configured on the system and their capacity.

Quantity	Capacity	Total (TB)
20	240 GB	4.80
27	1.6 TB	43.20
Total Storage (TB)		48.00

Table 3-1 Storage Device Capacities

Scale Factor = 3

Data Storage Ratio = (Total Storage (TB) / SF) = **16.00**

3.2 Memory Ratio

The Scale Factor to memory ratio must be disclosed.

Total Configured Memory (TiB) = 5.00

Scale Factor to Memory Ratio = (SF / Total Memory(TiB)) = **0.60**

Clause 4 – Metrics Related Items

4.1 HSGen Time

The HSGen time must be disclosed for Run1 and Run2.

	Run 1	Run 2
HSGen	198.808	202.124

Table 4-1 HSGen Times

4.2 HSSort Time

The HSSort time must be disclosed for Run1 and Run2.

	Run 1	Run 2
HSSort	639.108	625.454

Table 4-2 HSSort Times

4.3 HSValidate Time

The HSValidate time must be disclosed for Run1 and Run2.

	Run 1	Run 2
HSValidate	78.278	86.889

Table 4-3 HSValidate Times

4.4 HSDataCheck Times

Both HSDataCheck times must be disclosed for Run1 and Run2.

	Run 1	Run 2
HSDataCheck (pre-sort)	6.000	6.000
HSDataCheck (post-sort)	6.000	6.000

Table 4-4 HSDataCheck Times

4.5 Performance & Price-Performance

The performance metric (HSph@SF) must be disclosed for Run 1 and Run 2. Price-performance metric (\$/HSph@SF) must be disclosed for the performance run.

	Run 1	Run 2
HSph@3TB	11.62	11.65

Table 4-5 Performance Metrics

Run 1 Price-Performance: 38,544.15 \$/ HSph@3TB

Auditor's Information & Letter of Attestation

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 10453
978-343-6562

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

A copy of the auditor's Letter of Attestation follows.



Cindy Stap
 Senior Manager
 Solution Performance Analysis
 1 Dell Way PS2-39
 Round Rock, TX 78682

September 16, 2019

I verified the TPC Express Benchmark™ HS v2.0.3 performance of the following configuration:

Platform: Dell PowerEdge R7515 (9x PowerEdge R7515, 1x PowerEdge R6525)
 Operating System: Red Hat Enterprise Linux Server 7.6
 Apache Hadoop Cloudera Enterprise Edition 6.3 (using MapReduce)
 Compatible Software:

The results were:

Performance Metric 11.62 HSph@3TB
 Run Elapsed Time 929.00 Seconds

**Cluster 9x Dell PowerEdge R7515 (data nodes),
 1x PowerEdge R6525 (master node) each node with:**

CPU	1x AMD EPYC 7542 (2.90 GHz, 32-core, 16 MB L3) (Data nodes) 2x AMD EPYC 7502 (2.50 GHz, 32-core, 16 MB L3) (Master node)		
Memory	512 GiB (all nodes)		
Storage	Qty	Size	Type
	2	240GB	SATA SSD m.2 (All nodes)
	3	1.6TB	NVMe (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 v2.0.3
- No modifications were made to any of the Java code
- Any and all modifications to shell scripts were reviewed for compliance
- All checksums were validated for compliance
- The generated dataset was properly scaled to 3TB

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

- The generated dataset and the sorted dataset were replicated 3-ways
- 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,

A handwritten signature in cursive script that reads "Doug Johnson". The signature is written in black ink and has a long, sweeping horizontal line extending to the right.

Doug Johnson, Certified TPC Auditor

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

Supporting Files Index

Clause	Description	Archive File Pathname
Clause 1	Parameters and options used to configure the system	SupportingFiles/Clause1
Clause 2	Configuration scripts and Run Report	SupportingFiles/Clause2
Clause 3	System configuration details	SupportingFiles/Clause3

Third-Party Price Quotes

None.