

# TPC Express Benchmark™ HS Full Disclosure Report

## Dell PowerEdge R6515

(with 16x PowerEdge R6515 Servers; 1x PowerEdge R7515 Server)

Running

Cloudera Enterprise 7.1.4

on

SUSE Linux Enterprise Server 12 SP5

TPCx-HS Version  
Report Edition  
Report Submitted

2.0.3  
First  
March 15, 2021

**First Edition - March 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 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](http://www.dell.com). Actual performance and environmental costs of Dell products will vary depending on individual customer configurations and conditions.*

**Copyright© 2021 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 R6515	Cloudera Enterprise 7.1.4	SUSE Linux Enterprise Server 12 SP5

TPC Express Benchmark™ HS Metrics			
Total System Cost	HSph@3TB	Price/Performance	Availability Date
\$1,229,447	34.52	\$35,615.50	April 15, 2021

## Executive Summary

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

	<h1 style="text-align: center;">Dell PowerEdge R6515</h1>		TPCx-HS	2.0.3
			TPC Pricing	2.6.0
			Report Date	Mar. 15, 2021
Availability Date	TPCx-HS Performance	Price/Performance	Total System Cost	
<b>April 15, 2021</b>	<b>34.52 HSph@3TB</b>	<b>\$35,615.50 \$/HSph@3TB</b>	<b>\$1,229,447 USD</b>	
System Under Test Configuration Overview				
Scale Factor	Hadoop Software	Operating System	Other Software	
3	Cloudera Enterprise 7.1.4	SUSE Linux Enterprise Server 12 SP5	None	
<div style="display: flex; justify-content: space-between;"> <div style="width: 45%;">  <p><b>16 x Dell PowerEdge R6515 (Data Nodes)</b>                      1x AMD EPYC 75F3 32-Core Processor                      512 GB (8x 64GB RDIMM 3200 MT/s Dual Rank)                      2x 240 GB SSD SATA m.2 Hot-plug Drive                      5x 3.2 TB, NVMe, 2.5 SFF Drive                      1x 480 GB SATA SSD (not used in testing)                      1x Mellanox Dual Port ConnectX-5 100 GbE QSFP28 NIC                      1x Broadcom Gigabit Ethernet BCM5720 (External Connectivity)                      1x Broadcom 25GbE NIC Mezzanine</p> <p style="text-align: center;">...</p>   <p><b>1 x Dell PowerEdge R7515 (Master Node)</b>                      1x AMD EPYC 75F3 32-Core Processor                      512 GB (8x 64GB RDIMM 3200 MT/s Dual Rank)                      2x 240 GB SSD SATA m.2 Hot-plug Drive                      2x 480 GB MICRON MTFDDAK480TDC 2.5-inch Form Factor                      1x Mellanox Dual Port ConnectX-5 100 GbE QSFP28 NIC (Cluster Connectivity)                      Broadcom Gigabit Ethernet BCM5720 (External Connectivity)</p> </div> <div style="width: 50%;">  <p style="text-align: center;">Dell S3048F-ON Networking 48p 1GbE Switch</p>  <p style="text-align: center;">Mellanox MSN2700 Spectrum-based 32-port 100GbE Open Ethernet (1U)</p> </div> </div>				
Physical Storage/Scale Factor: 88.37		Scale Factor/Physical Memory: 0.35		
Total Number of Servers:		17 (16x PowerEdge R6515; 1x PowerEdge R7515)		
Total Processors/Cores/Threads:		17/544/1,088		
Server Configuration:	Per PowerEdge R6515	Per PowerEdge R7515		
Processors	1x AMD EPYC 75F3 32-Core	1x AMD EPYC 75F3 32-Core		
Memory	512 GiB	512 GiB		
Storage Controller	BOSS-S1	BOSS-S1, HBA330 Mini		
Storage Device	2x 240GB M.2 SATA SSD	2x 240GB M.2 SATA SSD		
	5x 3.2TB NVMe	2x 480GB SATA 6Gbps SSD		
Network	1x Mellanox Dual-port 100 Gb QSFP	1x Mellanox Dual-port 100 Gb QSFP		
	1x Broadcom 25 Gb NIC Mezzanine			
Connectivity:	1x MSN2700 32-port 100 GbE Switch; 1x S3048F-ON 48p 1 GbE Switch			
Total Rack Units:	16x(1U)+1x(2U)+1x(1U) +1x(1U) = 16U +2U + 1U + 1U = 20U			

	<h1 style="margin: 0;">Dell PowerEdge R6515</h1>		TPCx-HS	2.0.3				
			TPC Pricing	2.6.0				
			Report Date	Mar. 15, 2021				
Description		Part Number	Source	Unit Price	Qty	Extended Price	3 Yr. Maint. Price	
<b>HARDWARE COMPONENTS</b>								
<b>PowerEdge R6515 - Data Nodes</b>					1	\$69,360.66	16	\$1,109,770.56
2.5 Chassis		379-BDTF			1		16	
PowerEdge R6515 Server		210-ASVR			1		16	
No Trusted Platform Module		461-AADZ			1		16	
2.5" Chassis with up to 10 Hard Drives, including up to 8 SAS/SATA or 9 NVME Drives		321-BERT			1		16	
AMD EPYC 75F3 3.3GHz, 32C/64T, 256M Cache (280W) DDR4-3200		338-BZRK			1		16	
Heatsink for CPU equal to or greater than 180W		412-AAAC			1		16	
Performance Optimized		370-AAIP			1		16	
3200MT/s RDIMMs		370-AEVR			1		16	
Unconfigured RAID		780-BCDS			1		16	
HBA330 12Gbps SAS HBA Controller, Minicard		405-AAJU			1		16	
Performance BIOS Settings		384-BBBL			1		16	
UEFI BIOS Boot Mode with GPT Partition		800-BBDM			1		16	
No Additional Mid Fan		384-BBSO			1		16	
Dual, Hot Plug, Redundant Power Supply (1+1), 550W		450-AGUJ			1		16	
Riser Config 2, 2x16 LP PCIe slot		330-BBNN			1		16	
PowerEdge R6515 Motherboard, with 2 x 1Gb Onboard LOM (BCM5720) V2		384-BCNQ			1		16	
iDRAC9,Enterprise x5		385-BBOT			1		16	
Broadcom 57414 Dual Port 25GbE SFP28 LOM Mezz Card		540-BCKU			1		16	
LCD Bezel for x10 chassis		350-BBXX			1		16	
BOSS controller card + with 2 M.2 Sticks 240G (RAID 1),LP		403-BCHI			1		16	
No Quick Sync		350-BBKR			1		16	
iDRAC,Legacy Password		379-BCSG			1		16	
iDRAC Group Manager, Enabled		379-BCQV			1		16	
Enterprise Linux OS, Non Factory Installed, Requires Subscription Selection		605-BBFL			1		16	
SUSE Linux Enterprise Server, 1-2 SKT w Unlimited VMs, L3 Subscription, 3 Year		528-CHFF			1		16	
No Media Required		605-BBFN			1		16	
ReadyRails Sliding Rails Without Cable Management Arm		770-BCJI			1		16	
No Internal Optical Drive		429-AAIQ			1		16	
No Systems Documentation, No OpenManage DVD Kit		631-AACK			1		16	
PowerEdge R6515 Shipping		340-CMZJ			1		16	
PowerEdge R6515 x4 or x10 Shipping Material		343-BBMS			1		16	
PowerEdge R6515 CCC Marking		389-DUHM			1		16	
Basic Next Business Day 36 Months		709-BBFL			1		16	included
Prosupport Plus and 4Hr Mission Critical Initial, 36 Month(s)		865-BBNF			1		16	included
64GB RDIMM, 3200MT/s, Dual Rank		370-AEVP			1		128	
480GB SSD SATA Mix Use 6Gbps 512 2.5in Hot-plug AG Drive, 3 DWPD, 2628 TBW		400-AZUT			1		16	
Dell 3.2TB, NVME, Mixed Use Express Flash, 2.5 SFF Drive, U.2, PM1725b with Carrier		400-BEFE			1		80	
Power Cord - C13, 3M, 125V, 15A (North America, Guam, North Marianas, Philippines, Samoa, Vietnam)		450-AALV			1		32	
Mellanox ConnectX-5 EX Dual Port 40/100GbE QSFP28 Adapter, PCIe Low Profile		540-BCIX			1		16	
SUSE Linux Enterprise Server, 1-2 SKT w Unlimited VMs, L3 Subscription, 3 Year, Digitally Fulfilled		528-CHFF			1		16	included
GCP Operations Management		929-8509			1		16	
(continued on next page)								

	<h1 style="margin: 0;">Dell PowerEdge R6515</h1>				TPCx-HS	2.0.3
					TPC Pricing	2.6.0
					Report Date	Mar. 15, 2021
(continued from previous page)						
Description	Part Number	Source	Unit Price	Qty	Extended Price	3 Yr. Maint. Price
<b>PowerEdge R7515 - Master Node</b>			1 \$47,268.70	1	\$47,268.70	
2.5 Chassis	379-BDTF		1	1		
SAS/SATA/NVMe Capable Backplane	379-BDSW		1	1		
PowerEdge R7515 Server	210-ASVQ		1	1		
No Trusted Platform Module	461-AADZ		1	1		
Chassis with up to 24 x 2.5" Hard Drives Including Maximum of 12 NVME Drives	321-BERS		1	1		
AMD EPYC 75F3 3.3GHz, 32C/64T, 256M Cache (280W) DDR4-3200	338-BZRK		1	1		
Standard Heatsink	412-AASE		1	1		
Performance Optimized	370-AAIP		1	1		
3200MT/s RDIMMs	370-AEVR		1	1		
Unconfigured RAID	780-BCDS		1	1		
HBA330 12Gbps SAS HBA Controller, 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, with 2 x 1Gb Onboard LOM (BCM5720) V2	384-BCNR		1	1		
iDRAC9,Enterprise x5	385-BBOT		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 Selection	605-BBFL		1	1		
SUSE Linux Enterprise Server, 1-2 SKT w Unlimited VMs, L3 Subscription, 3 Year	528-CHFF		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, 1 DWPD, 876 TBW	400-AXTV		1	2		
Jumper Cord - C13/C14, 4M, 250V, 12A (North America, Guam, North Marianas, Philippines, Samoa)	492-BBDV		1	2		
Mellanox ConnectX-5 EX Dual Port 40/100GbE QSFP28 Adapter, PCIe Low Profile	540-BCIX		1	1		
SUSE Linux Enterprise Server, 1-2 SKT w Unlimited VMs, L3 Subscription, 3 Year, Digitally Fulfilled	528-CHFF		1	1		included
(continued on next page)						


	<h1>Dell PowerEdge R6515</h1>	TPCx-HS	2.0.3
		TPC Pricing	2.6.0
		Report Date	Mar. 15, 2021

(continued from previous page)

Description	Part Number	Source	Unit Price	Qty	Extended Price	3 Yr. Maint. Price
Mellanox® Passive Copper cable, ETH 100GbE, 100Gb/s, QSFP28, 2m, Black, 30AWG, CA-N	AA319963		\$96.67	2	\$193.34	
APC NetShelter SX Deep Enclosure - Rack - cabinet - black - 24U - 19-inch	A7522217		\$1,129.99	1	\$1,129.99	
APC Basic Rack PDU AP9567 - 0U - 120V NEMA 5-15 Input / (14) NEMA 5-15 Output	A7541364		\$189.99	1	\$189.99	
ASUS VE248Q 24 Inch LED monitor - Widescreen Full HD Monitor	A6732050		\$136.99	3	\$410.97	
Logitech MK200 Media Keyboard and Mouse Combo	A6859396		\$24.99	3	\$74.97	
C2G 6ft Cat6 Snagless Unshielded (UTP) Ethernet Network Patch Cable - Black - patch cable - 6 ft - black	A7052140		\$5.99	17	\$101.83	
Dell C13 to C14, PDU Style, 250 V Power Cord, North America - 6.5 feet	450-ACHI		\$11.93	34	\$405.62	
32-port Mellanox Spectrum SN2700 - Switch - L3 - managed - 32 x 100 Gigabit QSFP28	A8628811		\$20,008.99	1	\$20,008.99	
Mellanox M-1 Global Support Gold Plus. 2hr committed TAC response (24x7). 3 years	A9878902		\$3,234.06	1		\$3,234.06
<b>PowerSwitch S3048-ON</b>			\$15,066.51	1	\$15,066.51	
Dell Networking S3048-ON, 48x 1GbE, 4x SFP+ 10GbE ports, Stacking, IO to PSU air, 1x AC PSU, DNOS 9	210-AEDM			1		
ProSupport Plus: Mission Critical 4-Hour 7x24 On-Site Service with Emergency Dispatch, Initial Year	802-7419			1		
ProSupport Plus: 7x24 HW/SW Tech Support and Assistance, 3 Year	802-7434			1		included
ProSupport Plus: Mission Critical 4-Hour 7x24 On-Site Service with Emergency Dispatch, 2 Year Extended	802-7435			1		included
<b>HARDWARE COMPONENTS Subtotal</b>					\$1,197,855.53	
<b>SOFTWARE COMPONENTS</b>						
Cloudera CDP - 3 year 24x7 support			\$40,800.00	17	\$693,600.00	
<b>SOFTWARE COMPONENTS Subtotal</b>					\$693,600.00	
<b>TOTAL</b>						<b>\$1,891,455.53</b>
Large Purchase Discount (35%)*						-662,009.44

Pricing: 1 = Dell; * Discount applies to all line items where Key = 1. Discount based upon total system cost as purchased by a regular customer. <b>Audited by Doug Johnson, InfoSizing</b>	<b>Three-Year Cost of Ownership: \$1,229,447</b> <b>HSph@3TB: 34.52</b> <b>\$ / HSph@3TB: \$35,615.50</b>
---	---

*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](mailto:pricing@tpc.org). Thank you.*

	<h1>Dell PowerEdge R6515</h1>	TPCx-HS	2.0.3
		TPC Pricing	2.6.0
		Report Date	Mar. 15, 2021
<h2>Numerical Quantities</h2>			
<h3>Performance Run – Run 1</h3>			
Scale Factor			3TB
Run Start Time	2021-02-26 12:15:04.000		
Run End Time	2021-02-26 12:20:14.000		
Run Elapsed Time			313.000
HSGen Start Time	2021-02-26 12:15:05.000		
HSGen End Time	2021-02-26 12:16:06.000		
HSGen Elapsed Time			62.323
HSSort Start Time	2021-02-26 12:16:10.000		
HSSort End Time	2021-02-26 12:19:35.000		
HSSort Elapsed Time			205.211
HSValidate Start Time	2021-02-26 12:19:39.000		
HSValidate End Time	2021-02-26 12:20:14.000		
HSValidate Elapsed Time			36.254
<h3>Repeatability Run – Run 2</h3>			
Scale Factor			3TB
Run Start Time	2021-02-26 12:25:19.000		
Run End Time	2021-02-26 12:30:21.000		
Run Elapsed Time			305.000
HSGen Start Time	2021-02-26 12:25:20.000		
HSGen End Time	2021-02-26 12:26:21.000		
HSGen Elapsed Time			62.584
HSSort Start Time	2021-02-26 12:26:25.000		
HSSort End Time	2021-02-26 12:29:42.000		
HSSort Elapsed Time			197.317
HSValidate Start Time	2021-02-26 12:29:46.000		
HSValidate End Time	2021-02-26 12:30:21.000		
HSValidate Elapsed Time			36.218





# Dell PowerEdge R6515

TPCx-HS	2.0.3
TPC Pricing	2.6.0
Report Date	Mar. 15, 2021

## Run Reports

### Run Report for Performance Run – Run 1

#### TPCx-HS Performance Metric (HSph@SF) Report

Test Run 1 Details	Total Time =	313
	Total Size =	30000000000
	Scale-Factor =	3

TPCx-HS Performance Metric (HSph@SF): 34.5224

### Run Report for Repeatability Run – Run 2

#### TPCx-HS Performance Metric (HSph@SF) Report

Test Run 2 Details	Total Time =	305
	Total Size =	30000000000
	Scale-Factor =	3

TPCx-HS Performance Metric (HSph@SF): 35.4191

	<b>Dell PowerEdge R6515</b>	TPCx-HS 2.0.3 TPC Pricing 2.6.0 Report Date Mar. 15, 2021
---	-----------------------------	---

### Revision History

Date	Edition	Description
March 15, 2021	First	Initial Publication

# Table of Contents

Abstract..... 3

Executive Summary ..... 3

Table of Contents.....11

Clause 0 – Preamble .....12

    0.1 TPC Express Benchmark™ HS Overview.....12

Clause 1 – General Items .....13

    1.1 Test Sponsor .....13

    1.2 Parameter Settings .....13

    1.3 Configuration Diagrams .....13

        1.3.1 Measured Configuration.....14

        1.3.2 Priced Configuration .....14

    1.4 Dataset Distribution.....15

    1.5 Software Components Distribution.....15

Clause 2 – Workload Related Items.....16

    2.1 Hardware & Software Tunables .....16

    2.2 Run Report .....16

    2.3 Benchmark Kit Identification.....16

    2.4 Benchmark Kit Changes .....17

Clause 3 – SUT Related Items.....18

    3.1 Data Storage Ratio .....18

    3.2 Memory Ratio .....18

Clause 4 – Metrics Related Items .....19

    4.1 HSGen Time.....19

    4.2 HSSort Time .....19

    4.3 HSValidate Time .....19

    4.4 HSDataCheck Times .....19

    4.5 Performance & Price-Performance .....19

Auditor’s Information & Letter of Attestation .....20

Supporting Files Index .....23

Third-Party Price Quotes.....24

# 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](http://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](http://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

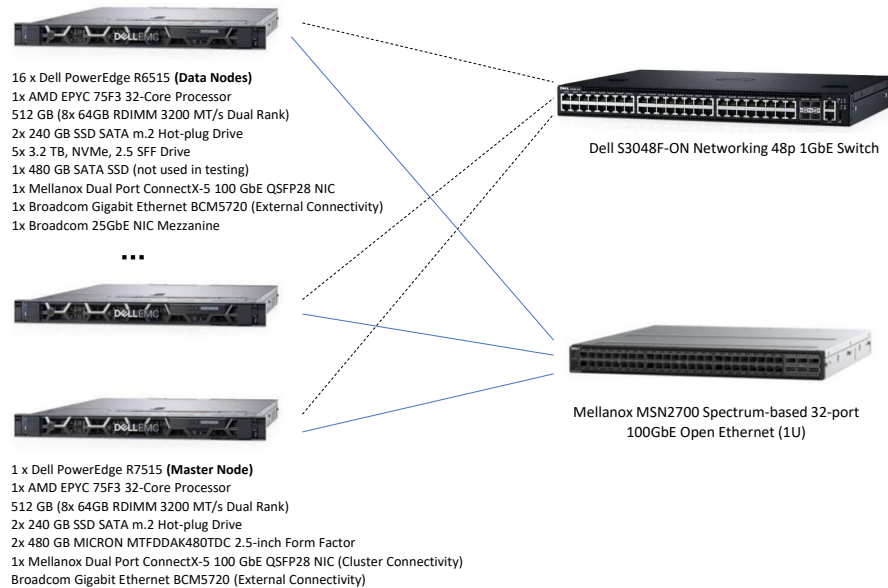


Figure 1-1 Measured Configuration

The measured configuration consisted of:

- Total Nodes: 17 (16x PowerEdge R6515; 1x PowerEdge R7515)
- Total Processors/Cores/Threads: 17/544/1,088
- Total Memory: 8.50TiB
- Total Number of Storage Drives/Devices: 116
- Total Storage Capacity: 265.12TB

Server node details:

16x PowerEdge R6515 Servers, each with:

- Processors/Cores/Threads: 1/32/64
- Processor Model: AMD EPYC 75F3 32-Core
- Memory: 512 GiB
- Controller: BOSS-S1
- Drives:
  - 2x 240GB M.2 SATA SSD
  - 5x 3.2TB NVMe
- Network:
  - 1x Mellanox Dual-port 100 Gb QSFP
  - 1x Broadcom 25 Gb NIC Mezzanine

1x PowerEdge R7515 Servers, each with:

- Processors/Cores/Threads: 1/32/64
- Processor Model: AMD EPYC 75F3 32-Core
- Memory: 512 GiB
- Controller: BOSS-S1, HBA330 Mini
- Drives:
  - 2x 240GB M.2 SATA SSD
  - 2x 480GB SATA 6 Gbps SSD
- Network: Mellanox Dual-port 100 Gb QSFP

Network connectivity detail:

- 1x MSN2700 32-port 100 GbE Switch; 1x S3048F-ON 48p 1GbE Switch

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
1	Boss Controller card	sda	Operating System, Root, swap, Hadoop Master
2-3	Boss Controller card	sda	Operating System, Root, swap, Hadoop Master
2-3	NVMe	nvme0n1, nvme1n1, nvme2n1, nvme3n1, nvme4n1	Data, Temp
4-17	Boss Controller card	sda	Operating System, Root, swap, Hadoop Master
4-17	NVMe	nvme0n1, nvme1n1, nvme2n1, nvme3n1, nvme4n1	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
	Resource Manager	Node Manager	NameNode	DataNode	QuorumPeer
1	X		X		X
2-3					X
4-17		X		X	

*Table 1-2 Software Component Distribution*

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

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

*Map/Reduce implementation and corresponding version must be disclosed.*

Cloudera Enterprise 7.1.4 (compatible equivalent to Hadoop 3.1.1.7).

## 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 =                313
                        Total Size =                30000000000
                        Scale-Factor =                3

TPCx-HS Performance Metric (HSph@SF):                34.5224

=====
```

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

Test Run 2 Details      Total Time =                305
                        Total Size =                30000000000
                        Scale-Factor =                3

TPCx-HS Performance Metric (HSph@SF):                35.4191

=====
```

### 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

The required data protection was provided by HDFS Erasure Coding rather than the default three-way data replication. A policy of RS-6-3-1024k was used. Therefore, each block group consisted of 6 data blocks and 3 parity blocks. Each block within a given block group was placed on a different node thus ensuring the required data protection.

To collect the necessary data for auditing, the HSDDataCheck.sh script was modified. In accordance with the TPCx-HS Standard Specification, this change received prior approval from the TPCx-HS subcommittee.

## 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)
34	240 GB	8.16
2	480 GB	0.96
80	3.2 TB	256.00
<b>Total Storage (TB)</b>		<b>265.12</b>

*Table 3-1 Storage Device Capacities*

Scale Factor = 3

**Data Storage Ratio = (Total Storage (TB) / SF) = 88.37**

### 3.2 Memory Ratio

*The Scale Factor to memory ratio must be disclosed.*

Total Configured Memory (TiB) = 8.50

**Scale Factor to Memory Ratio = (SF / Total Memory(TiB)) = 0.35**

## Clause 4 – Metrics Related Items

### 4.1 HSGen Time

The HSGen time must be disclosed for Run1 and Run2.

	Run 1	Run 2
HSGen	62.323	62.584

Table 4-1 HSGen Times

### 4.2 HSSort Time

The HSSort time must be disclosed for Run1 and Run2.

	Run 1	Run 2
HSSort	205.211	197.317

Table 4-2 HSSort Times

### 4.3 HSValidate Time

The HSValidate time must be disclosed for Run1 and Run2.

	Run 1	Run 2
HSValidate	36.254	36.218

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)	4.000	4.000
HSDataCheck (post-sort)	4.000	4.000

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	34.52	35.41

Table 4-5 Performance Metrics

Run 1 Price-Performance: 35,615.50 \$/ 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](http://www.tpc.org).

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



Mohan Rakkam  
 Senior Principal Engineer, PowerEdge Technical Marketing  
 Dell Inc.  
 1 Dell Way  
 Round Rock, TX 78682

March 2, 2021

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

Platform: Dell PowerEdge 6515 (16x PowerEdge 6515, 1x PowerEdge 7515)  
 16x Dell PowerEdge 6515 Servers (Data Nodes)  
 1x Dell PowerEdge 7515 Server (Master Node)

Operating System: SUSE Linux Enterprise Server 12 SP5  
 Apache Hadoop Cloudera Enterprise 7.1.4

Compatible Software:

The results were:

**Performance Metric 34.52 HSph@3TB**  
 Run Elapsed Time 313.00 Seconds

Cluster	16x PowerEdge 6515, 1x PowerEdge 7515 with:		
CPU	1x AMD® EPYC 75F3 32-Core Processor (all nodes)		
Memory	512 GiB (all nodes)		
Storage	<b>Qty</b>	<b>Size</b>	<b>Type</b>
	2	240 GB	M.2 SATA SSD (all nodes)
	5	3.2 TB	NVMe (data nodes)
	2	480 GB	SATA 6Gbps SSD (Master Node)

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 3 TB

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

- The generated dataset and the sorted dataset were erasure coded with a policy of RS-6-3-1024k
- 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](http://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

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