

# TPC Express Benchmark™ AI Full Disclosure Report

## PowerEdge R6625

with 1x PowerEdge R6625; 3x PowerEdge R6625  
using

Cloudera SEL Data Platform Private Cloud  
Base Edition

running on

Red Hat Enterprise Linux 8.6

TPCx-AI Version  
Report Edition  
Report Submitted

1.0.2  
First  
November 10, 2022

**First Edition - November 2022**

*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, TPC Express Benchmark™ AI 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™ AI, TPCx-AI, and AIUCpm@100, 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© 2022 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

Dell conducted the TPC Express Benchmark™ AI (TPCx-AI) on the PowerEdge R6625. The software used included Cloudera SEL Data Platform Private Cloud Base Edition. This report provides full disclosure of the results. All testing was conducted in conformance with the requirements of the TPCx-AI Standard Specification, Revision 1.0.2.

## Configuration Overview

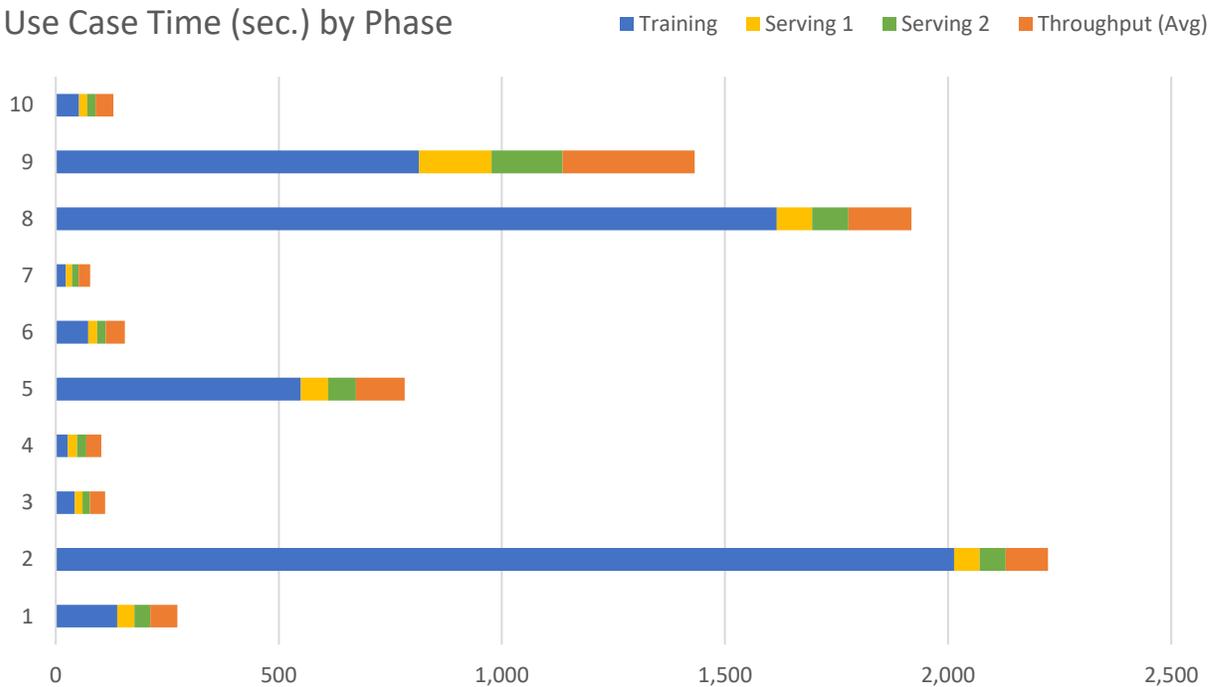
Test Sponsor	Node(s)	Operating System
Dell	1x PowerEdge R6625 (Primary Node) 3x PowerEdge R6625 (Worker Node)	Red Hat Enterprise Linux 8.6

## Metrics Overview

Total System Cost	Performance	Price/Performance	Availability Date
\$309,091 USD	868.49 AIUCpm@100	355.90 USD \$/AIUCpm@100	February 22, 2023

# Executive Summary

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

		<h1>PowerEdge R6625</h1>		TPCx-AI 1.0.2 TPC Pricing 2.8.0 Report Date Nov. 10, 2022																																																							
TPCx-AI Performance  <b>868.49</b> <b>AIUCpm@100</b>	Total System Cost  <b>\$309,091 USD</b>	Price/Performance  <b>\$355.90</b> <b>USD/AIUCpm@100</b>	Availability Date  <b>February 22, 2023</b>																																																								
Framework Cloudera SEL Data Platform Private Cloud Base Edition	Operating System Red Hat Enterprise Linux 8.6	Other Software N/A	Scale Factor 100	Streams 9																																																							
<h3>Use Case Time (sec.) by Phase</h3>  <table border="1"> <caption>Approximate Use Case Time (sec.) by Phase</caption> <thead> <tr> <th>Phase</th> <th>Training (sec)</th> <th>Serving 1 (sec)</th> <th>Serving 2 (sec)</th> <th>Throughput (Avg) (sec)</th> </tr> </thead> <tbody> <tr><td>1</td><td>100</td><td>50</td><td>50</td><td>50</td></tr> <tr><td>2</td><td>2000</td><td>50</td><td>50</td><td>50</td></tr> <tr><td>3</td><td>50</td><td>50</td><td>50</td><td>50</td></tr> <tr><td>4</td><td>50</td><td>50</td><td>50</td><td>50</td></tr> <tr><td>5</td><td>500</td><td>50</td><td>50</td><td>50</td></tr> <tr><td>6</td><td>50</td><td>50</td><td>50</td><td>50</td></tr> <tr><td>7</td><td>50</td><td>50</td><td>50</td><td>50</td></tr> <tr><td>8</td><td>1600</td><td>50</td><td>50</td><td>50</td></tr> <tr><td>9</td><td>800</td><td>200</td><td>200</td><td>300</td></tr> <tr><td>10</td><td>50</td><td>50</td><td>50</td><td>50</td></tr> </tbody> </table>					Phase	Training (sec)	Serving 1 (sec)	Serving 2 (sec)	Throughput (Avg) (sec)	1	100	50	50	50	2	2000	50	50	50	3	50	50	50	50	4	50	50	50	50	5	500	50	50	50	6	50	50	50	50	7	50	50	50	50	8	1600	50	50	50	9	800	200	200	300	10	50	50	50	50
Phase	Training (sec)	Serving 1 (sec)	Serving 2 (sec)	Throughput (Avg) (sec)																																																							
1	100	50	50	50																																																							
2	2000	50	50	50																																																							
3	50	50	50	50																																																							
4	50	50	50	50																																																							
5	500	50	50	50																																																							
6	50	50	50	50																																																							
7	50	50	50	50																																																							
8	1600	50	50	50																																																							
9	800	200	200	300																																																							
10	50	50	50	50																																																							
Physical Storage / Scale Factor <b>787.20</b>	Scale Factor / Physical Memory <b>0.04</b>	Main Data Redundancy Model Replication 3, RAID 1																																																									
Servers: Total Processors/Cores/Threads	4 8 / 256 / 512																																																										
Server Type	1x PowerEdge R6625 (Primary Node)	3x PowerEdge R6625 (Worker Node)																																																									
Processors	2x AMD EPYC 9354 32-Core Processor GHz	2x AMD EPYC 9354 32-Core Processor GHz																																																									
Memory	384 GiB	768 GiB																																																									
Storage Controller	1x PERC H965i	1x PERC H965i																																																									
Storage Device	2x 240 GB M.2 SSD; 2x 3.84 TB NVMe	2x 240 GB M.2 SSD; 6x 3.84 TB NVMe																																																									
Network Controller	1x Mellanox ConnectX-4, 25GbE, 2-port	1x Mellanox ConnectX-4, 25GbE, 2-port																																																									
Connectivity	1x Mellanox SN2400 100/25 GbE (Switch)																																																										

	<h1>PowerEdge R6625</h1>	TPCx-AI	1.0.2
		TPC Pricing	2.8.0
		Report Date	Nov. 10, 2022

Description	Part Number	Source	List Price	Qty	Extended Price	1-Yr. Maintenance
<b>Hardware</b>						
PowerEdge R6625 Server - Primary Node	210-ATCF	1	\$67,602.00	1	\$67,602.00	
2.5 Chassis	379-BDTF	1	0	1		
NVMe Backplane	379-BDSX	1	0	1		
Trusted Platform Module 2.0 V3	461-AAIG	1	0	1		
C03-03 : 8x U.2 G4 RAID - Low Z (FPERC 12)	321-BIIN	1	0	1		
MOD,PRC,9354,2.7,GOA,32C,XXX,QB	338-CGMZ	1	0	1		
MOD,PRC,9354,2.7,GOA,32C,XXX,QB	338-CGMZ	1	0	1		
Performance Optimized	370-AAIP	1	0	1		
4800MT/s RDIMMs	370-AHCL	1	0	1		
16GB RDIMM, 4800MT/s Single Rank	370-AGZO	1	0	24		
C31, No RAID with NVMe and front PERC	379-BEGI	1	0	1		
PERC H965i with floating brackets for lowz	405-ABDN	1	0	1		
Front PERC Mechanical Parts,rear load	750-ACFQ	1	0	1		
No Hard Drive	400-ABHL	1	0	1		
3.84TB Enterprise NVMe Read Intensive AG Drive U.2 Gen4 with carrier	400-BKGL	1	0	2		
Performance BIOS Settings	384-BBBL	1	0	1		
UEFI BIOS Boot Mode with GPTPartition	800-BBDM	1	0	1		
High Performance Fan for CPUgreater than or equal to 180W(2 CPU)	750-ADJI	1	0	1		
Dual, Hot-plug, Fully Redundant Power Supply (1+1), 1400W, Mixed Mode, NAF	450-AIQX	1	0	2		
Power Cord - C13, 3M, 125V, 15A (North America, Guam, North Marianas, Philippines, Samoa, Vietnam)	450-AALV	1	0	1		
Riser Config 2, 1 x 16 LP PCIe slot (CPU1), 2 x 16 LP PCIe slot (CPU2)	330-BBNR	1	0	1		
PowerEdge R6625 Motherboard	384-BCWP	1	0	1		
Broadcom 5720 Dual Port 1GbE Optional LOM	540-BDKD	1	0	1		
Mellanox Technologies MT27710 Family [ConnectX-4 Lx], 25 GbE	540-BBVV	1	0	1		
iDRAC9,Enterprise 15G	385-BBOT	1	0	1		
Dell EMC Luggage Tag (x8 or x10 chassis)	350-BBXP	1	0	1		
Standard Bezel	325-BCHH	1	0	1		
No Quick Sync	350-BBXM	1	0	1		
iDRAC,Factory Generated Password	379-BCSF	1	0	1		
<b>Red Hat Enterprise Linux 8.6 (Ootpa), kernel 4.18.0-372.9.1.el8.x86_64</b>	605-BBFL	1	0	1		
ReadyRails Sliding Rails Without Cable Management Arm or Strain Relief Bar	770-BECD	1	0	1		
Cable Management Arm	770-BDMT	1	0	1		
BOSS-S2 controller card + with 2 M.2 240GB (RAID 1)	403-BCMG	1	0	1		
No Systems Documentation, NoOpenManage DVD Kit	631-AACK	1	0	1		
PowerEdge R6625 Shipping Material 4	340-COXQ	1	0	1		
PowerEdge R6625 CCC Marking, No CE Marking	389-DTIQ	1	0	1		
US Order	332-1286	1	0	1		
Dell Hardware Limited Warranty Plus On-Site Service	828-3901	1	\$200.00	1		\$200.00
ProSupport Mission Critical:4-Hour 7x24 On-Site Service with Emergency Dispatch 3 Years	828-3855	1	\$940.00	1		\$940.00
ProSupport Mission Critical:7x24 HW / SW Technical Support and Assistance 3 Years	828-3847	1	\$1,816.00	1		\$1,816.00

(continued on the next page)

	<h1>PowerEdge R6625</h1>	TPCx-AI 1.0.2
		TPC Pricing 2.8.0
		Report Date Nov. 10, 2022

(continued from the previous page)

Description	Part Number	Source	List Price	Qty	Extended Price	1-Yr. Maintenance
PowerEdge R6625 Server - Worker Nodes	210-ATCF	1	\$103,571.00	3	\$310,713.00	
2.5 Chassis	379-BDTF	1	0	3		
NVMe Backplane	379-BDSX	1	0	3		
Trusted Platform Module 2.0 V3	461-AAIG	1	0	3		
C03-03 : 8x U.2 G4 RAID - Low Z (FPERC 12)	321-BIIN	1	0	3		
MOD,PRC,9354,2.7,GOA,32C,XXX,QB	338-CGMZ	1	0	3		
MOD,PRC,9354,2.7,GOA,32C,XXX,QB	338-CGMZ	1	0	3		
Performance Optimized	370-AAIP	1	0	3		
4800MT/s RDIMMs	370-AHCL	1	0	3		
32GB RDIMM, 4800MT/s Dual Rank	370-AGZP	1	0	72		
C31, No RAID with NVMe and front PERC	379-BEGI	1	0	3		
PERC H965i with floating brackets for lowz	405-ABDN	1	0	3		
Front PERC Mechanical Parts,rear load	750-ACFQ	1	0	3		
No Hard Drive	400-ABHL	1	0	3		
3.84TB Enterprise NVMe Read Intensive AG Drive U.2 Gen4 with carrier	400-BKGL	1	0	18		
Performance BIOS Settings	384-BBBL	1	0	3		
UEFI BIOS Boot Mode with GPTPartition	800-BBDM	1	0	3		
High Performance Fan for CPUgreater than or equal to 180W(2 CPU)	750-ADJI	1	0	3		
Dual, Hot-plug, Fully Redundant Power Supply (1+1), 1400W, Mixed Mode, NAF	450-AIQX	1	0	3		
Power Cord - C13, 3M, 125V, 15A (North America, Guam, North Marianas, Philippines, Samoa, Vietnam)	450-AALV	1	0	6		
Riser Config 2, 1 x 16 LP PCIe slot (CPU1), 2 x 16 LP PCIe slot (CPU2)	330-BBNR	1	0	3		
PowerEdge R6625 Motherboard	384-BCWP	1	0	3		
Broadcom 5720 Dual Port 1GbE Optional LOM	540-BDKD	1	0	3		
Mellanox Technologies MT27710 Family [ConnectX-4 Lx], 25 GbE	540-BBVV	1	0	3		
iDRAC9,Enterprise 15G	385-BBOT	1	0	3		
Dell EMC Luggage Tag (x8 or x10 chassis)	350-BBXP	1	0	3		
Standard Bezel	325-BCHH	1	0	3		
No Quick Sync	350-BBXN	1	0	3		
iDRAC,Factory Generated Password	379-BCSF	1	0	3		
<b>Red Hat Enterprise Linux 8.6 (Ootpa), kernel 4.18.0-372.9.1.el8.x86_64</b>	605-BBFL	1	0	3		
ReadyRails Sliding Rails Without Cable Management Arm or Strain Relief Bar	770-BECD	1	0	3		
Cable Management Arm	770-BDMT	1	0	3		
BOSS-S2 controller card + with 2 M.2 240GB (RAID 1)	403-BCMG	1	0	3		
No Systems Documentation, NoOpenManage DVD Kit	631-AACK	1	0	3		
PowerEdge R6625 Shipping Material 4	340-COXQ	1	0	3		
PowerEdge R6625 CCC Marking, No CE Marking	389-DTIQ	1	0	3		
US Order	332-1286	1	0	3		
Dell Hardware Limited Warranty Plus On-Site Service	828-3901	1	\$200.00	3		\$600.00
ProSupport Mission Critical:4-Hour 7x24 On-Site Service with Emergency Dispatch 3 Years	828-3855	1	\$940.00	3		\$2,820.00
ProSupport Mission Critical:7x24 HW / SW Technical Support and Assistance 3 Years	828-3847	1	\$1,816.00	3		\$5,448.00
Keyboard and Optical Mouse, USB, Black, English	570-AAKV, 580-ADJC	1	\$12.00	1	\$12.00	
Dell 24 Monitor	210-AIWG	1	\$169.99	1	\$169.99	
Mellanox MSN2410-CB2F 25GbE/100GbE Switch	920-9N112-00F7-0X2	2	\$14,242.00	1	\$14,242.00	
NVIDIA ENT Business Critical Support Services for SN2000 - 12 Months	780-C20N0Z +P2CMI12	2	\$1,140.00	1		\$1,140.00
				<b>Subtotal</b>	\$392,738.99	\$12,964.00
<b>Software</b>						
Cloudera SEL Data Platform Private Cloud Base Edition Annual Subscription per node Business Level Support AB352445		1	\$15,384.62	4		\$61,538.48
				<b>Subtotal</b>	\$0.00	\$61,538.48
Large Purchase Discount (35%)*					-\$132,473.95	-\$25,676.87
				<b>Total</b>	\$260,265.04	\$48,825.61

Pricing: 1 = Dell; 2 = Nvidia

\* Discount applies to all line items where Key = 1. Discount based upon total system cost as purchased by a regular customer.

**Audited by Doug Johnson, InfoSizing**

**Total System Cost (USD): \$309,091**

**AIUCpm@100: 868.49**

**\$/AIUCpm@100: \$355.90**

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.

	<h2>PowerEdge R6625</h2>		TPCx-AI	1.0.2
			TPC Pricing	2.8.0
			Report Date	Nov. 10, 2022
<u>Numerical Quantities</u>				
<b>AIUCpm@100</b>	<b>868.49</b>	$T_{Load}$	314.20	
Scale Factor	100	$T_{LD}$	314.20	
Streams	9	$T_{PTT}$	170.45	
Kit Version	1.0.2	$T_{PST1}$	35.37	
Execution Status	Pass	$T_{PST2}$	35.42	
Accuracy Status	Pass	$T_{PST}$	35.42	
		$T_{TT}$	12.01	
Test Times				
Overall Run Start Time	2022-10-28 23:26:16.411			
Overall Run End Time	2022-10-29 01:51:40.373			
Overall Run Elapsed Time	8,723.962			
Load Test Start Time	2022-10-28 23:33:33.068			
Load Test End Time	2022-10-28 23:38:49.333			
Load Test Elapsed Time	316.265			
Power Training Start Time	2022-10-28 23:38:49.334			
Power Training End Time	2022-10-29 01:08:08.741			
Power Training Elapsed Time	5,359.407			
Power Serving 1 Start Time	2022-10-29 01:08:08.743			
Power Serving 1 End Time	2022-10-29 01:16:28.456			
Power Serving 1 Elapsed Time	499.713			
Power Serving 2 Start Time	2022-10-29 01:16:28.457			
Power Serving 2 End Time	2022-10-29 01:24:45.222			
Power Serving 2 Elapsed Time	496.765			
Scoring Start Time	2022-10-29 01:28:44.798			
Scoring End Time	2022-10-29 01:33:28.795			
Scoring Elapsed Time	283.997			
Throughput Start Time	2022-10-29 01:33:28.799			
Throughput End Time	2022-10-29 01:51:40.373			
Throughput Elapsed Time	1,091.574			



# PowerEdge R6625

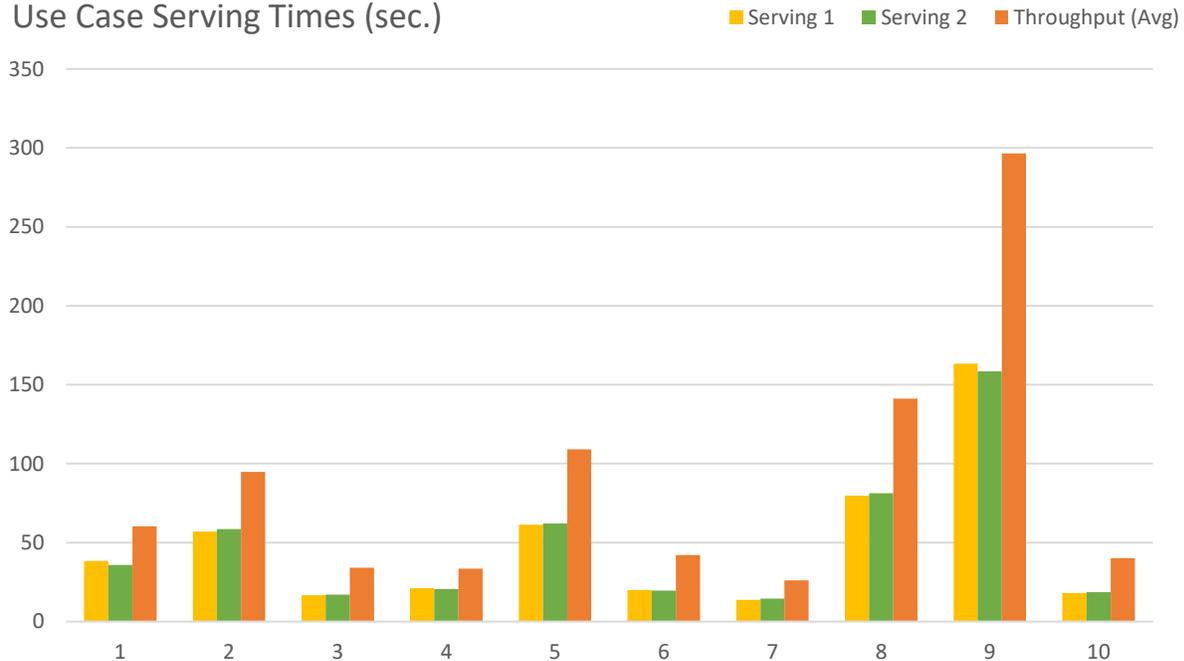
TPCx-AI 1.0.2  
 TPC Pricing 2.8.0  
 Report Date Nov. 10, 2022

Numerical Quantities (continued)

Use Case Times & Accuracy

Use Case	Training (sec)	Serving 1 (sec)	Serving 2 (sec)	Throughput (avg)	Accuracy
UC01	137.870	38.369	35.873	60.275	0.000
UC02	2,013.347	57.043	58.527	94.730	0.394
UC03	43.034	16.747	16.970	34.070	3.909
UC04	27.216	21.109	20.599	33.473	0.694
UC05	549.776	61.344	62.142	109.071	0.038
UC06	72.815	20.078	19.546	42.207	0.204
UC07	23.066	13.662	14.485	26.081	1.307
UC08	1,615.781	79.722	81.280	141.162	0.755
UC09	813.859	163.349	158.454	296.504	0.990
UC10	52.425	18.085	18.708	40.255	0.816

Use Case Serving Times (sec.)



# Table of Contents

- Abstract..... 3
- Executive Summary ..... 3
- Table of Contents..... 9
- Clause 0 – Preamble .....11
  - 0.1 TPC Express Benchmark™ AI Overview .....11
- 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 Differences Between the Measured and the Priced Configurations.....14
- Clause 2 – SW Components & Data Distribution .....15
  - 2.1 Roles and Dataset Distribution.....15
  - 2.2 File System Implementation.....15
  - 2.3 Execution Engine, Frameworks, Driver & Libraries .....16
  - 2.4 Applied Patches.....16
- Clause 3 – Workload Related Items.....17
  - 3.1 Hardware & Software Tuning.....17
  - 3.2 Kit Version & Modifications .....17
  - 3.3 Use Case Elapsed Times.....17
  - 3.4 SUT Validation Test Output .....18
  - 3.5 Configuration Parameters .....19
- Clause 4 – SUT Related Items.....20
  - 4.1 Specialized Hardware/Software .....20
  - 4.2 Configuration Files .....20
  - 4.3 SUT Environment Information.....20
  - 4.4 Data Storage to Scale Factor Ratio.....20
  - 4.5 Scale Factor to Memory Ratio.....20
  - 4.6 Output of Tests .....20
  - 4.7 Additional Sponsor Files .....20
  - 4.8 Model Optimizations .....20
- Clause 5 – Metrics and Scale Factor .....21
  - 5.1 Reported Performance Metrics .....21

5.2 Throughput Test Stream Times.....22

Auditor's Information .....23

Third-Party Price Quotes.....26

    Nvidia .....26

Supporting Files Index .....27

# Clause 0 – Preamble

## 0.1 TPC Express Benchmark™ AI Overview

Artificial intelligence (AI) has become a key transformational technology of our times. Advances in neural networks and other machine learning techniques have made it possible to use AI on a variety of use cases. From the public sector to aerospace, defense and academia, new and improved ways to use AI techniques are changing the way we harness data and analytics. This along with advances in compute, interconnect and memory technologies have made possible to solve complicated challenges that will ultimately benefit customers in production datacenter and cloud environments.

Abundant volumes of rich data from text, images, audio and video are the essential starting point for creating a benchmark that would represent the myriad of use cases and customers. TPC Express Benchmark™ AI (TPCx-AI) is created in keeping with the TPC tradition of emulating real world AI scenarios and data science use cases. Unlike most other AI benchmarks, the TPCx-AI uses a diverse dataset and is able to scale across a wide range of scale factors. TPCx-AI may later expand with additional use cases and add additional flexibility for a greater variety of implementations.

The benchmark defines and provides a means to evaluate the System Under Test (SUT) performance as a general-purpose data science system that:

- Generates and processes large volumes of data.
- Trains preprocessed data to produce realistic machine learning models.
- Conducts accurate insights for real-world customer scenarios based on the generated models.
- Can scale to large scale distributed configurations.
- Allows for flexibility in configuration changes to meet the demands of the dynamic AI landscape.

The benchmark models real-life examples of companies and public-sector organizations that use a range of analytics techniques, both AI and more traditional machine learning approaches, as well as the potential application of these techniques in situations like those in which they have already been successfully deployed. In addition, the benchmark measures end to end time to provide insights for individual use cases, as well as throughput metrics to simulate multiuser environments for a given hardware, operating system, and data processing system configuration under a controlled, complex, multi-user AI or machine learning data science workload.

The purpose of TPC benchmarks is to provide relevant, objective performance data to industry users. To achieve that purpose, TPC benchmark specifications require benchmark runs 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-AI models and represents complex, high data volume, decision support environments).
- Would plausibly be implemented.

The TPCx-AI kit is available from the TPC website (see [www.tpc.org/tpcx-ai/](http://www.tpc.org/tpcx-ai/) for more information). Users must sign up and agree to the TPCx-AI 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-AI copyright. The TPCx-AI kit includes: TPCx-AI Specification document (this document), TPCx-AI Users Guide (README.md) documentation, scripts to set up the benchmark environment, code to execute the benchmark workload, Data Generator, use case related files, and Benchmark Driver.

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.

Further information is available at [www.tpc.org](http://www.tpc.org).

## Clause 1 – General Items

### 1.1 Test Sponsor

This benchmark was sponsored by Dell Inc..

### 1.2 Parameter Settings

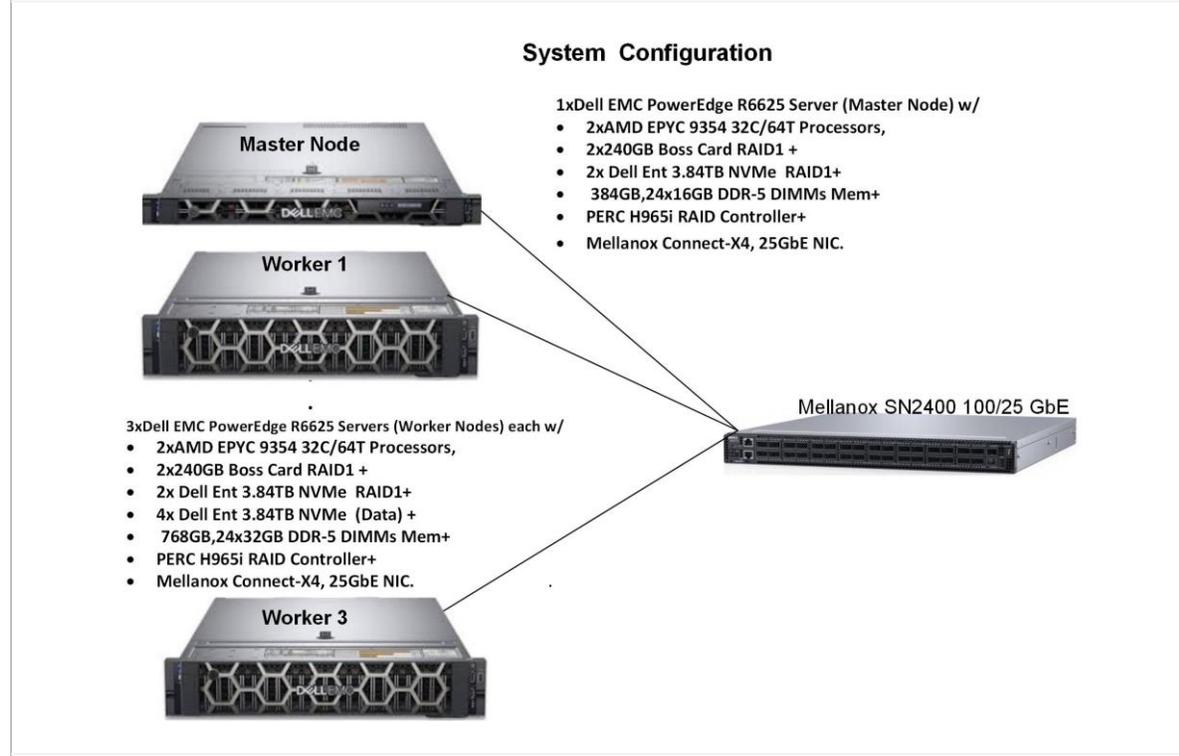
The [Supporting Files Archive](#) contains the parameters and options used to configure the components involved in this benchmark.

### 1.3 Configuration Diagrams

The measured configuration diagram is shown below. In addition, any differences between the measured and the priced configurations are described.

### 1.3.1 Measured Configuration

Nodes:	4		
Processors/Cores/Threads:	8/256/512	Storage Devices:	28
Total Memory:	2,688 GiB	Storage Capacity:	78,720 GB



	<u>Primary Node</u>	<u>Worker Node</u>
Server	1x PowerEdge R6625:	3x PowerEdge R6625:
Procs/Cores/Threads:	2/32/64	2/32/64
Processor Model:	2x AMD EPYC 9354 32-Core Processor	2x AMD EPYC 9354 32-Core Processor
Memory:	384 GiB	768 GiB
Storage Controller:	1x PERC H965i	1x PERC H965i
Storage Devices:	2x 240 GB M.2 SSD 2x 3.84 TB NVMe	2x 240 GB M.2 SSD 6x 3.84 TB NVMe
Network Controller:	1x Mellanox ConnectX-4, 25GbE, 2-port	1x Mellanox ConnectX-4, 25GbE, 2-port
Network:	1x Mellanox SN2400 100/25 GbE (Switch);	

The distribution of software components over server nodes is detailed in [Clause 2](#).

### 1.3.2 Differences Between the Measured and the Priced Configurations

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

## Clause 2 – SW Components & Data Distribution

### 2.1 Roles and Dataset Distribution

Table 2-1 describes the distribution of the dataset across all media in the SUT.

Server	Host Name	SW Services	Storage	Contents
1x PowerEdge R6625	genoa- namenode	Core Configuration Gateway Core Configuration Storage Operations HDFS Balancer HDFS NameNode HDFS SecondaryNameNode Hive Gateway Hive Metastore Server Hive on Tez Gateway Hive on Tez HiveServer2 Cloudera Management Service Alert Publisher Cloudera Management Service Event Server Cloudera Management Service Host Monitor Cloudera Management Service Service Monitor YARN Queue Manager Store YARN Queue Manager Webapp Spark Gateway Spark History Server Tez Gateway YARN JobHistory Server YARN ResourceManager ZooKeeper Server	2x 240 GB M.2 SSD 2x 3.84 TB NVMe	OS Kit Metadata
3x PowerEdge R6625	genoa- datanode[1-3]	Core Configuration Gateway HDFS DataNode Hive Gateway Hive on Tez Gateway Spark Gateway Tez Gateway YARN NodeManager	2x 240 GB M.2 SSD 2x 3.84 TB NVMe 4x 3.84 TB NVMe	OS Kit Metadata Distributed FS

Table 2-1 Software Components and Dataset Distribution

### 2.2 File System Implementation

A distributed file system provided by Red Hat Enterprise Linux 8.6 / Cloudera SEL Data Platform Private Cloud Base Edition was used for data generation and the Load Test. The data set was not relocated after generation and before the Load Test.

## 2.3 Execution Engine, Frameworks, Driver & Libraries

Cloudera SEL Data Platform Private Cloud Base Edition consisted of the following components.

Component	Version
HDFS	3.1.1
YARN	3.1.1
MapReduce2	3.1.1
Spark	2.4.7

*Table 2-2 Software Components*

For a detailed listing of installed libraries, please see the envInfo logs in the [Supporting Files](#).

## 2.4 Applied Patches

No additional vendor-supported patches were applied to the SUT.

## Clause 3 – Workload Related Items

### 3.1 Hardware & Software Tuning

The [Supporting Files](#) archive contains all hardware and software configuration scripts.

### 3.2 Kit Version & Modifications

Table 3-1 shows the version of the TPCx-AI used to produce this result along with any kit files that were modified to facilitate system, platform, and framework differences.

TPCx-AI Kit Version	1.0.2
<u>Modified File</u>	<u>Description of Changes</u>
None – See Auditor’s Note	N/A

Table 3-1 Kit Version & Modifications

### 3.3 Use Case Elapsed Times

Below are the elapsed times for each use case. Use cases are grouped based on whether they use Deep Learning or Machine Learning techniques.

Type	UC ID	P1	P2	T1	T2	T3	T4
Deep Learning	2	57.043	58.527	79.801	88.754	125.529	83.976
	5	61.344	62.142	92.827	329.121	74.355	78.609
	9	163.349	158.454	192.889	227.040	239.524	621.365
Machine Learning	1	38.369	35.873	79.794	41.665	56.769	63.172
	3	16.747	16.970	18.970	60.999	17.571	73.574
	4	21.109	20.599	26.773	26.217	22.965	25.023
	6	20.078	19.546	58.069	26.912	48.706	20.400
	7	13.662	14.485	30.832	14.600	18.889	14.797
	8	79.722	81.280	123.652	96.315	212.879	81.044
	10	18.085	18.708	85.286	24.877	22.791	18.552

Type	UC ID	T5	T6	T7	T8	T9
Deep Learning	2	106.953	84.235	106.727	75.736	100.862
	5	81.649	73.008	96.349	65.355	90.370
	9	233.982	185.405	198.775	575.209	194.346
Machine Learning	1	44.788	38.501	40.115	39.376	138.296
	3	19.980	20.810	57.003	18.525	19.196
	4	60.078	39.838	36.812	40.552	23.001
	6	53.785	67.748	31.659	27.521	45.063
	7	28.477	40.555	28.323	41.036	17.222
	8	117.065	214.673	168.365	132.591	123.878
	10	38.291	33.033	52.657	19.293	67.512

Table 3-2 Use Case Elapsed Times

### 3.4 SUT Validation Test Output

<u>Validation Run Report</u>			
AIUCpm@1	18.25	T <sub>Load</sub>	262.90
Scale Factor	1	T <sub>LD</sub>	262.90
Streams	9	T <sub>PTT</sub>	64.41
Kit Version	1.0.2	T <sub>PST1</sub>	22.04
Execution Status	Pass	T <sub>PST2</sub>	22.21
Accuracy Status	Pass	T <sub>PST</sub>	22.21
		T <sub>TT</sub>	3.10
Test Times			
Overall Run Start Time	2022-10-28 22:24:30.197		
Overall Run End Time	2022-10-28 23:17:37.431		
Overall Run Elapsed Time	3,187.234		
Load Test Start Time	2022-10-28 22:27:27.070		
Load Test End Time	2022-10-28 22:31:52.028		
Load Test Elapsed Time	264.958		
Power Training Start Time	2022-10-28 22:31:52.029		
Power Training End Time	2022-10-28 22:55:58.920		
Power Training Elapsed Time	1,446.891		
Power Serving 1 Start Time	2022-10-28 22:55:58.922		
Power Serving 1 End Time	2022-10-28 23:00:06.861		
Power Serving 1 Elapsed Time	247.939		
Power Serving 2 Start Time	2022-10-28 23:00:06.863		
Power Serving 2 End Time	2022-10-28 23:04:13.826		
Power Serving 2 Elapsed Time	246.963		
Scoring Start Time	2022-10-28 23:08:10.885		
Scoring End Time	2022-10-28 23:12:47.552		
Scoring Elapsed Time	276.667		
Throughput Start Time	2022-10-28 23:12:47.556		
Throughput End Time	2022-10-28 23:17:37.431		
Throughput Elapsed Time	289.875		
(continued on next page)			

Validation Run Report (continued)

Accuracy Metrics					
Use Case	Metric Name	Metric	Criteria	Threshold	Status
1	N/A	0.000	N/A	0.00	Pass
2	word_error_rate	0.471	<=	0.50	Pass
3	mean_squared_log_error	5.141	<=	5.40	Pass
4	f1_score	0.697	>=	0.65	Pass
5	mean_squared_log_error	0.110	<=	0.50	Pass
6	matthews_corrcoef	0.224	>=	0.19	Pass
7	median_absolute_error	1.715	<=	1.80	Pass
8	accuracy_score	0.701	>=	0.65	Pass
9	accuracy_score	1.000	>=	0.90	Pass
10	accuracy_score	0.817	>=	0.70	Pass

### 3.5 Configuration Parameters

The [Supporting Files](#) archive contains all Global Benchmark Parameter and Use Case Specific Parameter settings.

## Clause 4 – SUT Related Items

### 4.1 Specialized Hardware/Software

No Specialized Hardware/Software was used in the SUT.

### 4.2 Configuration Files

The [Supporting Files](#) archive contains all configuration files.

### 4.3 SUT Environment Information

All envInfo.log files are included in the [Supporting Files](#) archive.

### 4.4 Data Storage to Scale Factor Ratio

The details of the Data Storage Ratio are provided below.

Node Count	Disks	Size (GB)	Total (GB)
4	2	240	1,920
1	2	3,840	7,680
3	6	3,840	69,120

Total Storage (GB) 78,720

Scale Factor 100

Data Storage Ratio 787.20

### 4.5 Scale Factor to Memory Ratio

The details of the Memory to Scale Factor Ratio are provided below.

Nodes	Memory (GiB)	Total (GiB)
1	384	384
3	768	2,304

Scale Factor 100

Total Memory (GiB) 2,688

SF / Memory Ratio 0.04

### 4.6 Output of Tests

The [Supporting Files](#) archive contains the output files of all tests.

### 4.7 Additional Sponsor Files

The [Supporting Files](#) archive contains any additional files that were used.

### 4.8 Model Optimizations

The [Supporting Files](#) archive contains any model optimization files that were used.

# Clause 5 – Metrics and Scale Factor

## 5.1 Reported Performance Metrics

### Metric Overview

TPCx-AI Performance Metric	868.49	AIUCpm@100
TPCx-AI Price/Performance Metric	355.90	\$/AIUCpm@100
TPCx-AI Scale Factor	100	
TPCx-AI Stream Count	9	

### Test Times

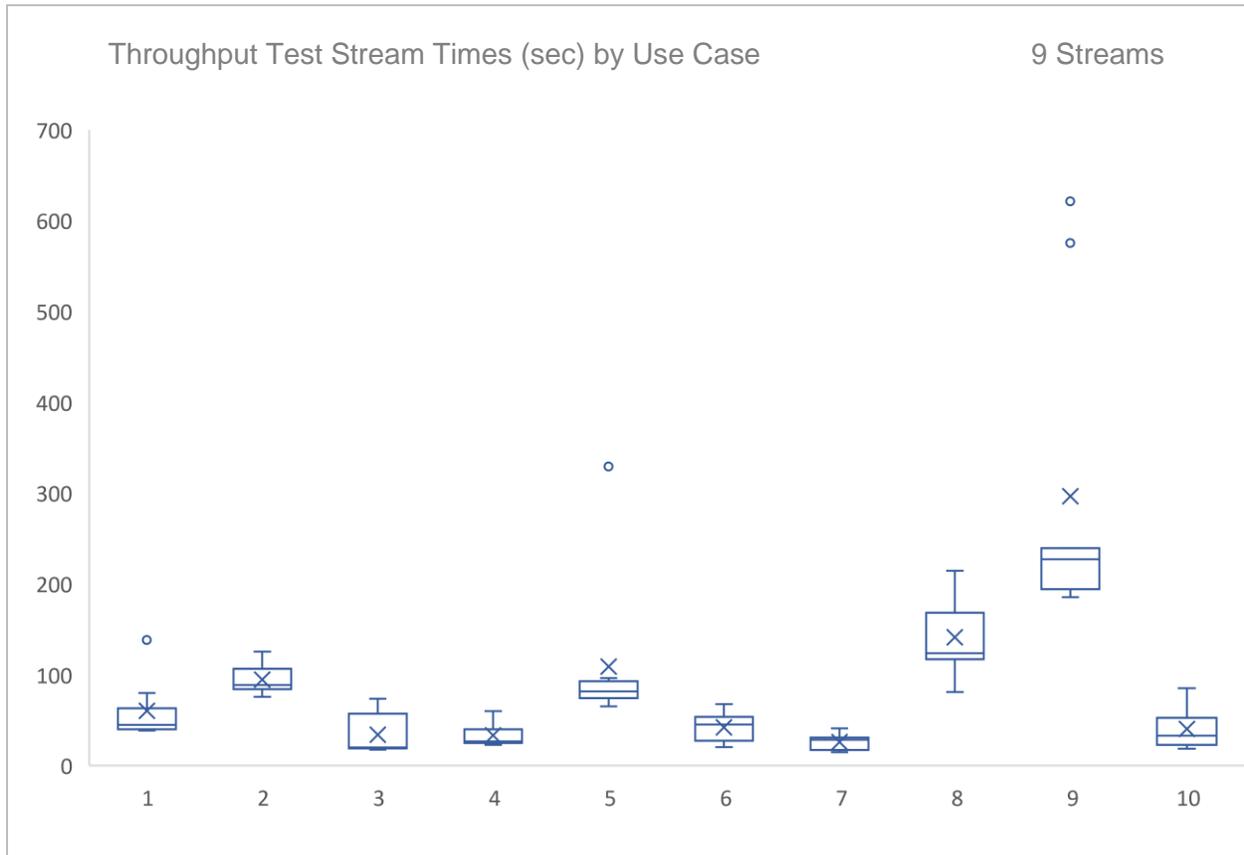
Overall Run Start Time	2022-10-28 23:26:16.411
Overall Run End Time	2022-10-29 01:51:40.373
Overall Run Elapsed Time	8,723.962
Load Test Start Time	2022-10-28 23:33:33.068
Load Test End Time	2022-10-28 23:38:49.333
Load Test Elapsed Time	316.265
Power Training Start Time	2022-10-28 23:38:49.334
Power Training End Time	2022-10-29 01:08:08.741
Power Training Elapsed Time	5,359.407
Power Serving 1 Start Time	2022-10-29 01:08:08.743
Power Serving 1 End Time	2022-10-29 01:16:28.456
Power Serving 1 Elapsed Time	499.713
Power Serving 2 Start Time	2022-10-29 01:16:28.457
Power Serving 2 End Time	2022-10-29 01:24:45.222
Power Serving 2 Elapsed Time	496.765
Scoring Start Time	2022-10-29 01:28:44.798
Scoring End Time	2022-10-29 01:33:28.795
Scoring Elapsed Time	283.997
Throughput Start Time	2022-10-29 01:33:28.799
Throughput End Time	2022-10-29 01:51:40.373
Throughput Elapsed Time	1,091.574

Accuracy Metrics

Use Case	Metric Name	Metric	Criteria	Threshold	Status
1	N/A	0.000	N/A	0.00	Pass
2	word_error_rate	0.394	<=	0.50	Pass
3	mean_squared_log_error	3.909	<=	5.40	Pass
4	f1_score	0.694	>=	0.65	Pass
5	mean_squared_log_error	0.038	<=	0.50	Pass
6	matthews_corrcoef	0.204	>=	0.19	Pass
7	median_absolute_error	1.307	<=	1.80	Pass
8	accuracy_score	0.755	>=	0.65	Pass
9	accuracy_score	0.990	>=	0.90	Pass
10	accuracy_score	0.816	>=	0.70	Pass

## 5.2 Throughput Test Stream Times

The following chart shows the minimum, 1<sup>st</sup> quartile, median, mean (X), 3<sup>rd</sup> quartile, and maximum stream times by use case for the Throughput Test. Outliers are marked with “o”.



## Auditor's Information

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 can be downloaded from [www.tpc.org](http://www.tpc.org).

A copy of the auditor's attestation letter is included in the next two pages.



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

November 8, 2022

I verified the TPC Express Benchmark™ AI v1.0.2 performance of the following configuration:

Platform: 1x Dell PowerEdge R6625; 3x Dell PowerEdge 6625  
 Operating System: Red Hat Enterprise Linux 8.6  
 Additional Software: Cloudera SQL Data Platform Private Cloud Base Edition

The results were:

**Performance Metric 868.49 AIUCpm@100**

Secondary Metrics	T <sub>LD</sub>	314.20
	T <sub>PTT</sub>	170.45
	T <sub>PST</sub>	35.42
	T <sub>TT</sub>	12.01

**System Under Test 1x Dell PowerEdge R6625; 3x Dell PowerEdge 6625 with:**

CPU	2x AMD EPYC 9354 32-Core Processor (all nodes)		
Memory	384 GiB (Primary Node); 768 GiB (Worker Nodes)		
Storage	<b>Qty</b>	<b>Size</b>	<b>Type</b>
	2	240 GB	M.2 SATA (all nodes)
	2	3.84 TB	NVMe (primary node)
	6	3.84 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 v1.0.2.
- All checksums were validated for compliance.
- Any modifications to shell scripts were reviewed for compliance.
- No modifications were made to any of the Java code.
- The generated dataset was properly scaled to 100 GB.

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

- The generated dataset used for testing was protected by Replication 3.
- 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:

Two files were erroneously reported as having incorrect checksum. This is due to a minor issue in the TPC-provided kit. The TPCx-AI Subcommittee is aware of this and will correct it in a future release of the kit.

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

# Third-Party Price Quotes

## Nvidia



Date: 11/3/2022  
 Quote # Q-863709  
 Opportunity # O-313351  
 Expiration Date 2/1/2023

Advanced Micro Devices (AMD)

**End Customer**  
 Advanced Micro Devices (AMD)  
 United States

**Pricing Request Type**

**NPN Solution Provider**  
 Direct

**NVIDIA Salesperson**  
 Martin McNarney  
 mcnarney@nvidia.com

Qty	Part Number	Reference Part Number	Description	Term (Year)	Unit Price	Discount (%)	Sale Price	Total
1	920-9N112-00F7-0X2	MSN2410-CB2F	Mellanox Spectrum based 25GbE/100GbE 1U Open Ethernet switch with Oryx, 48 SFP28 ports and 8 QSFP28 ports, 2 Power Supplies (AC), x86 CPU, short depth, P2C airflow, Rail Kit		\$14,242.00	0.00	\$14,242.00	\$14,242.00
1	780-C20N0Z+P2CM112		NVIDIA ENT Business Critical Support Services for SN2000 - 12 Months	1	\$1,140.00	0.00	\$1,140.00	\$1,140.00

**Net Total** \$15,382.00

### NOTES

- All pricing is in USD and subject to change. Pricing does not include currency conversion fees, taxes or VAT, and other considerations that may affect the final price that you pay.
- Purchase Order Receipt Date AND Customer Request Date (CRD) must be earlier than Quote expiration date.
- Distributor Purchase Order to NVIDIA:
  - Quote number listed in this document must be included and can only be used for one unique order and may not be used in any other orders.
  - All products, quantities, pricing, reseller and end customer information must align with those on the referenced, valid quote.
  - Reseller, Reseller Contact, End Customer information
  - Must include full company name with no acronyms or abbreviations, address, first and last name of the product end customer contact, and a valid email address with a domain that matches the company name.
- Please reference your NVIDIA Price List for all ordering rules.
- This document serves as a pricing information to the Distributor.
- This quotation is subject to the terms and conditions specified in the applicable signed agreement between NVIDIA and Partner. In the absence of such signed agreement, NVIDIA's Standard Terms & Conditions will apply.
- Accordingly, the products are offered under the applicable terms and conditions, and this quotation is expressly conditional on acceptance of such terms and conditions. No additional or conflicting terms and conditions will apply without NVIDIA's prior, express written consent, and any such additional or conflicting terms and conditions on partner's purchase order, acknowledgement or other business form are hereby rejected by NVIDIA.

2788 San Tomas Expressway | Santa Clara, CA 95051 | T 408.486.2000 | F 408.486.2200 | www.nvidia.com

# Supporting Files Index

The Supporting Files archive for this disclosure contains the following structure.

Supporting Files Directory	Description
CheckIntegrity/...	Output of CHECK_INTEGRITY test (if the phase is not done as part of the Validation and Performance Test).
PerformanceTest/...	Performance Test output files.
ValidationTest/...	Validation Test output files.
Additional files used by Dell	
Sponsor/ModelOptimization/...	Details of model optimization.
Sponsor/ModifiedKitFiles/...	0 modified file(s). See Auditor's Note.
Sponsor/Tuning/...	All tuning files used.