Lenovo

TPC Express Benchmark™ AI
Full Disclosure Report

ThinkSystem SR650 V3

with 1x ThinkSystem SR650 V3
using

Anaconda Business

running on

Red Hat Enterprise Linux 8.9 (Ootpa)

TPCx-AI Version	2.0.0
Report Edition	First
Report Submitted	September 5, 2024

First Edition - September 2024

***Lenovo (Lenovo)**, 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.

Lenovo and the Lenovo Logo are trademarks of Lenovo 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 Lenovo and any other company.

TPC Express Benchmark™ AI, TPCx-AI, and AIUCpm@3, are registered certification marks of the Transaction Processing Performance Council.

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

Abstract

Lenovo conducted the TPC Express Benchmark™ AI (TPCx-AI) on the ThinkSystem SR650 V3. The software used included Anaconda Business. This report provides full disclosure of the results. All testing was conducted in conformance with the requirements of the TPCx-AI Standard Specification, Revision 2.0.0.

Configuration Overview

Test Sponsor	Node(s)	Operating System
Lenovo	1x ThinkSystem SR650 V3 (Server)	Red Hat Enterprise Linux 8.9 (Ootpa)

Metrics Overview

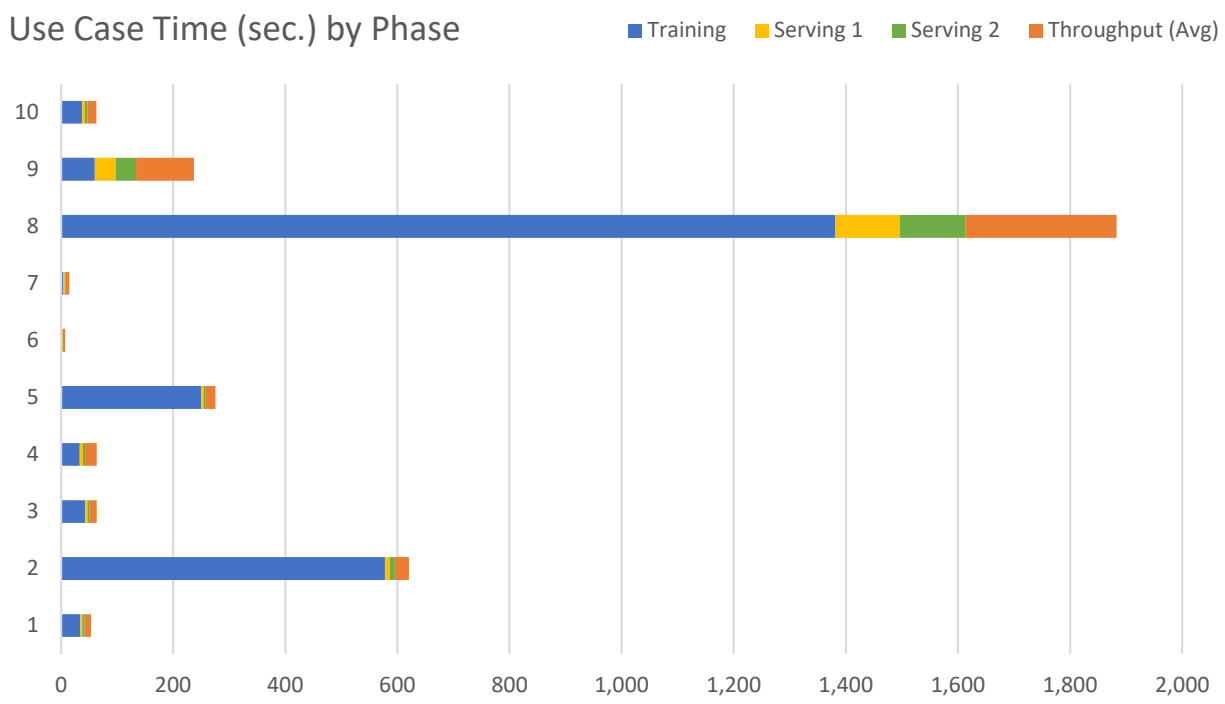
Total System Cost	Performance	Price/Performance	Availability Date
$121,475 USD	527.30 AIUCpm@3	230.38 USD $/AIUCpm@3	September 5, 2024

Executive Summary

The Executive Summary follows on the next several pages.

Lenovo	ThinkSystem SR650 V3		TPCx-AI 2.0.0 TPC Pricing 2.9.0 Report Date Sep. 05, 2024
TPCx-AI Performance **527.30 AIUCpm@3**	Total System Cost **$121,475 USD**	Price/Performance **$230.38 USD/AIUCpm@3**	Availability Date **September 5, 2024**

Framework	Operating System	Other Software	Scale Factor	Streams
Anaconda Business	Red Hat Enterprise Linux 8.9 (Ootpa)	N/A	3	140



Physical Storage / Scale Factor	Scale Factor / Physical Memory	Main Data Redundancy Model
640.00	0.01	RAID1

Servers: Total Processors/Cores/Threads	1 2 / 96 / 192
Server Type	1x ThinkSystem SR650 V3 (Server)
Processors	2x INTEL(R) XEON(R) PLATINUM 8558P
Memory	512 GiB
Storage Controller	1x M.2 SATA/x4 NVMe 2-Bay Adapter
Storage Device	2x 960 GB NVMe (VROC RAID1)
Network Controller	1x Intel E810-DA1 12/25 GbE SFP28 2-Port

ThinkSystem SR650 V3

TPCx-AI	2.0.0
TPC Pricing	2.9.0
Report Date	Sep. 05, 2024

Description	Part Number	Source	List Price	Qty	Extended Price	1-Yr. Maintenance
Server Hardware						
ThinkSystem SR650 V3 Configure-To-Order, includes:	7D76CTO1WW	1	$68,833	1	$68,833	
ThinkSystem V3 2U 24x2.5" Chassis	BLKK			1		
Intel Xeon Platinum 8558P 48C 350W 2.7GHz Processor	BYWA			2		
ThinkSystem 32GB TruDDR5 Performance+ 5600MHz (2Rx8) RDIMM	BZ4U			16		
ThinkSystem M.2 SATA/x4 NVMe 2-Bay Adapter	BM8X			1		
Intel VROC (VMD NVMe RAID) Premium for M.2	BS7N			1		
ThinkSystem M.2 7450 PRO 960GB Read Intensive NVMe PCIe 4.0 x4 NHS SSD	BKSR			2		
ThinkSystem Intel E810-DA2 10/25GbE SFP28 2-Port OCP Ethernet Adapter	BCD4			1		
ThinkSystem V3 2U x16/x16/E PCIe Gen4 Riser1 or 2	BLKM			1		
ThinkSystem 1800W 230V Platinum Hot-Swap Gen2 Power Supply v2	BMUF			1		
4.3m, 10A/100-250V, C13 to IEC 320-C14 Rack Power Cable	6263			1		
ThinkSystem 2U V3 Performance Fan Module	BLL6			6		
ThinkSystem Toolless Friction Rail v2	BK7W			1		
ThinkSystem 2U V3 EIA Latch Standard	BQQ2			1		
2m Passive DAC SFP+ Cable	A51P			1		
ThinkSystem SR650 V3 MB	BLL0			1		
RAID Configuration	2302			1		
ThinkSystem 2U V3 EIA right with FIO	BQQ6			1		
ThinkSystem 2U MS Air Duct Filler(For 2U Gap)	B8MP			2		
ThinkSystem 2U Main Air Duct	BP46			1		
MS 2FH Riser Filler	BC4X			1		
ThinkSystem 2U MS 3FH Riser 1&2 Cage w/Label1	BPKD			1		
ThinkSystem 2U MS 3FH Riser Filler	B8MM			1		
ThinkSystem SR650 V3 Firmware and Root of Trust Security Module v2	BYQL			1		
Red Hat Linux w/Lenovo Support	7S0FCTO1WW			1		
RHEL Server Physical or Virtual Node, 2 Skt Premium Subscription w/Lenovo Support 1Yr	S0N8			1		
Premier Essential - 3Yr 24x7 4Hr Resp + YDYD SR650 V3	5PS7B07067	1	$1,829	1		$1,829
				Subtotal	**$68,833**	**$1,829**
Server Software						
Anaconda Business Subscription (1 year)		2	$50,000	1	$50,000	
Anaconda Business Premium Support, including 24x7 support		2	$35,000	1		$35,000
				Subtotal	**$50,000**	**$35,000**
Infrastructure						
ThinkSystem Pref. Pro II USB Keyboard - US English	7ZB7A05522	1	$29	1	$29	
ThinkSystem Optical Wheel Mouse - USB	7M57A04698	1	$39	1	$39	
Lenovo D22e-20 21.45-inch FHD Monitor	66D2KCC6US	1	$80	1	$80	
				Subtotal	**$148**	**$0**
				Total	**$118,981**	**$36,829**
Dollar Volume Discount (See Note 1)	48.49%	1			$34,335	

Pricing: 1 = Lenovo; 2 = Anaconda

* Discount applies to all line items where Pricing=1; pricing is for these or similar quantities. Discounts for similarly sized configurations will be similar to what is quoted here, but may vary based on the specific components priced.

Audited by Doug Johnson, InfoSizing

Total System Cost (USD): $121,475

AIUCpm@3: 527.30

$/AIUCpm@3: $230.38

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.

Lenovo	ThinkSystem SR650 V3	TPCx-AI	2.0.0
		TPC Pricing	2.9.0
		Report Date	Sep. 05, 2024

Numerical Quantities

AIUCpm@3	**527.30**	T_{Load}	1.16
Scale Factor	3	T_{LD}	1.16
Streams	140	T_{PTT}	51.82
		T_{PST1}	6.10
Kit Version	1.0.3.1	T_{PST2}	6.16
Execution Status	OK	T_{PST}	6.16
Accuracy Status	Pass	T_{TT}	0.37

Test Times

Overall Run Start Time	2024-08-06 19:13:47.925
Overall Run End Time	2024-08-06 20:15:25.841
Overall Run Elapsed Time	3,697.916
Load Test Start Time	2024-08-06 19:16:35.263
Load Test End Time	2024-08-06 19:16:36.442
Load Test Elapsed Time	1.179
Power Training Start Time	2024-08-06 19:16:36.445
Power Training End Time	2024-08-06 19:57:00.335
Power Training Elapsed Time	2,423.890
Power Serving 1 Start Time	2024-08-06 19:57:00.338
Power Serving 1 End Time	2024-08-06 20:00:05.448
Power Serving 1 Elapsed Time	185.110
Power Serving 2 Start Time	2024-08-06 20:00:05.451
Power Serving 2 End Time	2024-08-06 20:03:11.984
Power Serving 2 Elapsed Time	186.533
Scoring Start Time	2024-08-06 20:04:29.415
Scoring End Time	2024-08-06 20:06:47.478
Scoring Elapsed Time	138.063
Throughput Start Time	2024-08-06 20:06:47.501
Throughput End Time	2024-08-06 20:15:25.836
Throughput Elapsed Time	518.335

Lenovo	ThinkSystem SR650 V3	TPCx-AI	2.0.0
		TPC Pricing	2.9.0
		Report Date	Sep. 05, 2024

Numerical Quantities (continued)

Use Case Times & Accuracy

Use Case	Training (sec)	Serving 1 (sec)	Serving 2 (sec)	Throughput (avg)	Accuracy
UC01	34.608	3.461	3.772	11.234	0.000
UC02	577.798	8.746	8.739	24.992	0.452
UC03	43.201	3.936	3.877	12.417	3.506
UC04	33.175	5.548	5.425	19.224	0.700
UC05	249.976	4.177	4.122	16.763	0.050
UC06	1.443	0.713	0.729	3.968	0.488
UC07	4.285	1.907	1.917	6.065	0.957
UC08	1,381.207	115.259	116.541	269.872	0.761
UC09	60.194	36.785	36.722	103.243	1.000
UC10	37.918	4.484	4.605	15.732	0.816

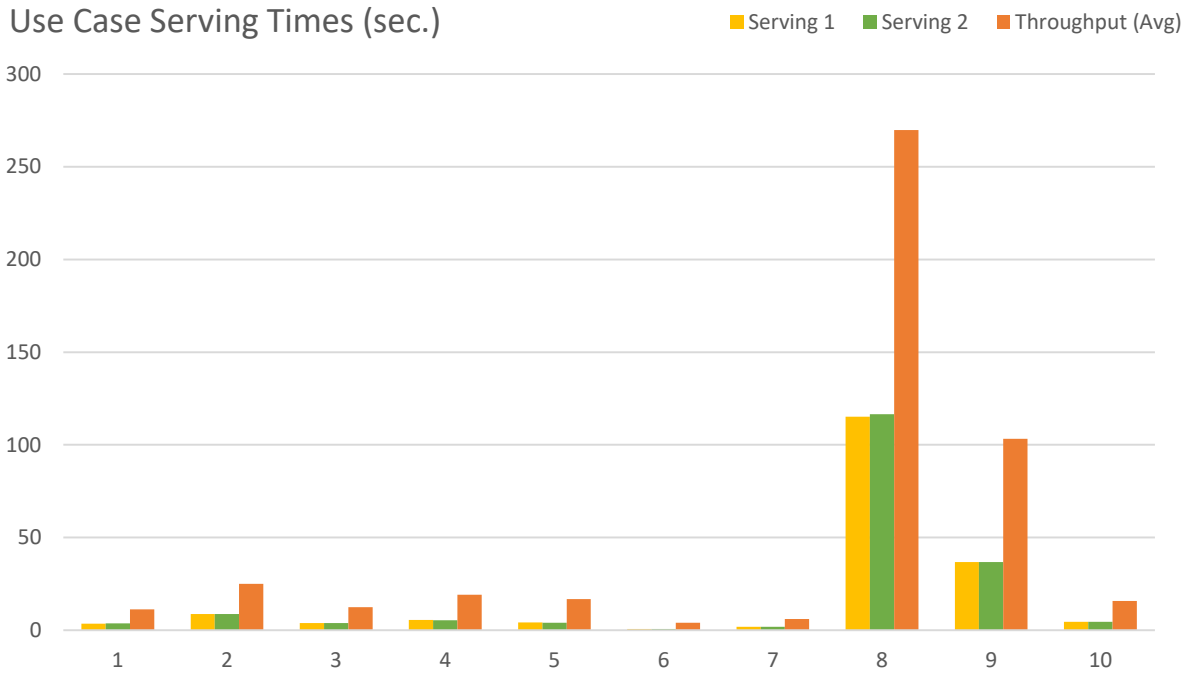


Table of Contents

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/ 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.

Clause 1 – General Items

1.1 Test Sponsor

This benchmark was sponsored by Lenovo.

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:	1		
Processors/Cores/Threads:	2/96/192	Storage Devices:	2
Total Memory:	512 GiB	Storage Capacity:	1,920 GB

	Server
Server	1x ThinkSystem SR650 V3
Procs/Cores/Threads:	2/96/192
Processor Model:	2x INTEL(R) XEON(R) PLATINUM 8558P
Memory:	512 GiB
Storage Controller:	1x M.2 SATA/x4 NVMe 2-Bay Adapter
Storage Devices:	2x 960 GB NVMe (VROC RAID1)
Network Controller:	1x Intel E810-DA1 12/25 GbE SFP28 2-Port

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 ThinkSystem SR650 V3	ventura-spr05	All	2x 960 GB NVMe (VROC RAID1)	OS, Data

Table 2-1 Software Components and Dataset Distribution

2.2 File System Implementation

A local file system provided by Red Hat Enterprise Linux 8.9 (Ootpa) / Anaconda Business 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

Anaconda Business consisted of the following components.

Component	Version
python	3.9
setuptools	59.8
pandas	1.5.3
scikit-learn	1.2.2
xgboost	1.7.4
numpy	1.23.5
nose	1.3.7
scipy	1.10.1
statsmodels	0.13.5
patsy	0.5.3
tqdm	4.65
keras	2.11.0
tensorflow	2.11.0
joblib	1.2.0
pyyaml	6.0
matplotlib	3.7.1
jinja2	3.1.2
pycryptodome	3.16
scikit-surprise	1.1.3
librosa	0.8.1
opencv	4.5.3
imbalanced-learn	0.10.1
tensorflow-addons	0.19.0

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 flies that were modified to facilitate system, platform, and framework differences.

TPCx-AI Kit Version	1.0.3.1
Modified File	Description of Changes
tools/python/dataRedundancyInformation.sh	Added platform-specific data collection.
tools/python/load.sh	cp's done in background.

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	8.746	8.739	30.227	18.122	22.097	22.253
	5	4.177	4.122	21.570	18.931	13.039	13.116
	9	36.785	36.722	92.554	96.387	126.993	106.330
Machine Learning	1	3.461	3.772	11.483	8.630	10.126	16.270
	3	3.936	3.877	14.431	11.014	4.449	13.249
	4	5.548	5.425	16.655	25.879	16.867	25.977
	6	0.713	0.729	5.127	3.573	3.300	2.874
	7	1.907	1.917	5.145	5.386	4.710	5.270
	8	115.259	116.541	239.339	263.758	291.261	282.889
	10	4.484	4.605	16.029	16.553	13.649	13.566

Type	UC ID	T5	T6	T7	T8	T9	T10
Deep Learning	2	24.234	24.538	20.380	23.856	19.494	22.745
	5	12.208	20.936	11.671	17.838	22.705	22.137
	9	133.773	87.793	95.652	91.215	96.697	116.982
Machine Learning	1	11.912	12.264	12.156	9.363	11.714	7.608
	3	11.356	14.853	14.954	9.993	16.215	18.835
	4	17.967	17.375	18.841	16.207	17.243	15.295
	6	4.457	3.539	4.392	3.545	4.149	2.427
	7	8.084	6.890	5.472	3.746	5.556	5.787
	8	243.192	281.137	293.710	262.929	277.874	262.187
	10	17.539	18.861	13.987	16.842	17.227	13.710

Type	UC ID	T11	T12	T13	T14	T15	T16
Deep Learning	2	19.247	34.662	12.867	21.308	24.382	20.287
	5	22.737	13.909	13.964	22.652	10.941	17.265
	9	104.921	104.240	129.743	94.393	90.677	100.787
Machine Learning	1	13.705	12.054	10.703	9.710	13.606	11.834
	3	8.401	7.685	14.204	11.016	16.291	12.813
	4	21.294	19.226	17.635	20.159	21.987	22.579
	6	5.555	6.987	0.898	4.810	3.315	4.102
	7	4.905	6.205	5.966	5.772	7.316	8.903
	8	281.146	281.675	280.623	287.527	273.902	274.258
	10	14.399	16.520	14.500	19.021	17.042	20.336

Type	UC ID	T17	T18	T19	T20	T21	T22
Deep Learning	2	16.028	22.901	20.239	28.158	20.953	19.724
	5	13.936	12.809	24.090	19.754	20.001	12.086
	9	87.782	104.314	89.947	115.689	127.118	114.058
Machine Learning	1	12.607	9.762	12.984	11.508	10.929	9.016
	3	10.719	10.142	16.853	9.960	13.574	11.862
	4	21.725	21.066	17.770	20.775	19.800	16.849
	6	4.994	5.052	5.201	4.126	3.800	3.356
	7	6.391	5.705	6.009	5.345	5.804	6.451
	8	254.669	265.208	288.152	245.280	249.572	275.223
	10	12.888	21.517	13.668	14.747	17.368	13.917

Type	UC ID	T23	T24	T25	T26	T27	T28
Deep Learning	2	25.621	19.635	20.859	32.794	18.233	20.299
	5	19.692	12.656	15.093	17.352	9.162	13.327
	9	96.079	113.491	93.603	90.957	124.702	109.712
Machine Learning	1	11.311	9.780	12.147	9.863	15.141	9.458
	3	11.205	11.863	13.342	11.192	11.357	14.480
	4	23.484	18.121	18.472	15.953	17.443	15.091
	6	5.869	3.994	3.744	4.893	6.764	2.933
	7	7.931	5.327	4.629	6.495	5.497	4.780
	8	268.950	296.250	285.772	268.411	284.404	288.267
	10	17.857	10.854	18.878	17.720	13.154	12.696

Type	UC ID	T29	T30	T31	T32	T33	T34
Deep Learning	2	19.725	29.862	10.393	20.943	16.432	34.678
	5	18.606	14.371	13.472	22.517	18.329	15.097
	9	107.672	115.332	129.428	103.643	108.205	106.038
Machine Learning	1	10.748	9.907	9.937	10.453	16.223	10.554
	3	13.319	14.450	12.638	10.391	10.575	10.990
	4	22.450	20.138	20.630	23.761	25.202	16.991
	6	5.193	3.632	4.953	3.404	3.196	3.539
	7	7.035	6.656	6.285	8.048	4.857	5.904
	8	240.917	246.876	279.012	259.957	274.981	267.264
	10	14.173	15.020	16.103	12.563	14.387	11.379

Type	UC ID	T35	T36	T37	T38	T39	T40
Deep Learning	2	37.857	22.376	28.882	33.140	17.753	20.404
	5	12.637	14.946	22.336	16.200	19.625	12.396
	9	92.295	88.950	91.528	105.139	114.615	107.036
Machine Learning	1	9.962	16.194	12.847	9.877	11.649	11.298
	3	10.794	10.140	12.089	9.924	10.909	11.106
	4	22.183	21.845	21.284	22.888	16.218	16.463
	6	4.717	3.048	3.632	2.375	3.151	3.202
	7	4.908	7.467	5.337	6.519	5.808	5.571
	8	289.625	284.921	250.915	276.685	288.747	279.678
	10	13.717	15.806	16.768	13.022	13.886	12.087

Type	UC ID	T41	T42	T43	T44	T45	T46
Deep Learning	2	27.706	23.745	29.864	24.902	31.180	26.025
	5	23.076	34.802	13.736	20.269	21.197	16.543
	9	104.784	93.631	100.442	107.245	106.121	94.499
Machine Learning	1	9.047	8.746	11.540	11.713	10.590	15.821
	3	10.469	13.560	10.960	10.022	11.517	12.969
	4	22.517	21.058	24.520	20.856	19.327	20.217
	6	2.360	4.581	3.767	2.556	4.492	5.981
	7	5.664	5.040	7.430	5.956	8.841	7.771
	8	268.367	245.857	276.704	269.850	276.147	272.613
	10	20.536	18.362	14.654	18.715	9.689	18.083

Type	UC ID	T47	T48	T49	T50	T51	T52
Deep Learning	2	26.286	25.615	17.086	18.311	25.140	19.406
	5	13.487	23.832	11.287	11.639	17.605	13.362
	9	102.367	102.235	95.574	101.428	122.924	87.953
Machine Learning	1	13.653	15.285	9.067	13.122	5.560	13.258
	3	10.312	13.849	12.014	10.633	15.346	9.838
	4	22.114	20.392	18.460	16.344	17.251	18.965
	6	3.316	4.623	3.478	2.974	5.169	4.514
	7	5.287	4.826	7.521	5.169	6.853	7.817
	8	272.063	260.628	260.648	259.072	271.275	245.658
	10	17.154	16.306	14.863	12.750	12.443	14.462

Type	UC ID	T53	T54	T55	T56	T57	T58
Deep Learning	2	33.921	39.147	26.513	16.333	18.729	38.915
	5	17.306	16.098	10.751	12.397	13.727	15.050
	9	108.020	96.445	93.041	89.234	105.675	93.765
Machine Learning	1	9.225	11.741	8.637	12.505	9.278	12.609
	3	11.760	14.268	10.954	15.342	10.974	13.748
	4	19.454	16.333	17.671	16.975	25.041	17.813
	6	4.756	3.573	4.151	3.789	3.371	3.404
	7	3.697	5.036	6.024	5.477	5.210	4.851
	8	274.428	266.514	285.909	259.637	280.746	294.747
	10	14.128	14.246	18.974	17.576	19.921	7.520

Type	UC ID	T59	T60	T61	T62	T63	T64
Deep Learning	2	21.859	27.904	19.788	24.585	25.500	17.362
	5	12.195	19.416	13.473	18.151	15.301	12.983
	9	93.463	95.743	111.208	88.007	114.097	96.967
Machine Learning	1	12.230	9.054	11.152	12.616	12.397	10.466
	3	9.111	12.672	16.880	13.931	14.400	10.435
	4	23.626	23.657	22.906	21.897	14.036	17.488
	6	3.330	3.687	5.239	3.316	3.149	3.130
	7	7.044	7.456	6.416	7.723	6.078	4.052
	8	250.933	242.383	271.777	286.753	282.616	276.369
	10	19.906	16.753	18.184	13.583	13.180	13.928

Type	UC ID	T65	T66	T67	T68	T69	T70
Deep Learning	2	20.583	36.375	30.295	31.548	26.830	21.187
	5	23.198	12.365	13.389	16.465	14.380	17.712
	9	100.468	110.760	130.825	94.020	83.648	97.072
Machine Learning	1	9.113	9.757	9.346	9.590	11.302	10.016
	3	9.184	11.223	12.248	13.790	10.833	17.623
	4	16.179	13.195	18.961	21.618	20.689	16.164
	6	3.212	4.820	3.302	4.796	3.231	3.714
	7	8.189	4.966	4.638	5.505	8.445	6.684
	8	245.333	265.919	258.183	249.790	291.558	279.496
	10	13.809	13.948	15.585	17.501	15.947	20.436

Type	UC ID	T71	T72	T73	T74	T75	T76
Deep Learning	2	22.940	25.502	34.389	25.162	26.931	25.968
	5	17.680	15.951	14.704	23.943	11.704	16.616
	9	98.334	96.707	101.307	94.395	106.421	121.866
Machine Learning	1	16.653	13.055	12.873	9.699	10.001	11.759
	3	6.286	13.590	13.336	12.421	11.950	13.018
	4	20.601	26.754	16.695	17.674	16.759	25.445
	6	4.099	3.942	3.105	2.898	3.286	3.778
	7	5.169	6.708	4.627	6.147	5.183	4.045
	8	286.801	237.398	271.102	291.376	280.408	254.528
	10	20.815	12.134	14.905	14.507	14.189	17.963

Type	UC ID	T77	T78	T79	T80	T81	T82
Deep Learning	2	27.805	25.799	20.032	21.709	27.838	32.512
	5	17.621	19.155	18.349	17.732	23.806	16.526
	9	91.874	129.696	114.666	92.200	106.047	104.910
Machine Learning	1	9.427	3.906	10.423	9.619	9.883	8.079
	3	17.959	5.907	10.647	8.990	9.143	12.262
	4	23.121	21.616	17.871	19.868	16.594	20.121
	6	3.505	4.034	2.561	4.529	4.919	3.044
	7	5.220	6.130	5.385	6.093	6.640	8.401
	8	250.775	277.428	289.055	283.898	270.819	276.685
	10	14.381	10.804	11.255	16.498	18.302	17.456

Type	UC ID	T83	T84	T85	T86	T87	T88
Deep Learning	2	31.876	16.333	19.123	19.027	34.451	19.178
	5	13.836	20.740	21.326	13.032	18.149	11.696
	9	108.726	102.351	112.985	124.443	105.676	108.986
Machine Learning	1	8.374	10.668	12.066	13.043	13.070	9.933
	3	14.862	12.761	16.923	12.886	9.975	9.924
	4	17.324	23.662	16.668	19.809	21.089	16.936
	6	4.624	4.340	3.119	4.166	5.470	3.550
	7	6.384	4.423	5.212	8.212	6.938	5.048
	8	276.834	277.012	268.723	241.129	267.783	289.946
	10	15.180	16.995	14.072	18.858	14.453	12.195

Type	UC ID	T89	T90	T91	T92	T93	T94
Deep Learning	2	26.474	33.599	19.062	21.492	19.522	19.833
	5	12.570	19.021	17.837	12.781	14.467	17.348
	9	121.594	109.287	108.719	98.764	98.119	115.463
Machine Learning	1	12.336	15.435	14.451	13.119	10.248	11.046
	3	12.986	12.326	9.585	10.463	12.998	14.144
	4	22.602	12.890	17.452	14.418	17.032	16.118
	6	3.785	3.632	3.516	3.900	3.446	5.279
	7	5.237	5.510	6.854	7.274	5.560	5.320
	8	259.609	276.817	272.747	290.539	286.716	274.043
	10	16.563	15.769	18.976	16.769	13.255	16.871

Type	UC ID	T95	T96	T97	T98	T99	T100
Deep Learning	2	27.105	17.949	33.820	32.616	24.142	16.985
	5	17.295	12.534	16.027	13.566	21.299	14.715
	9	105.064	93.729	94.296	94.045	110.031	116.167
Machine Learning	1	14.732	10.463	13.757	9.514	12.831	10.911
	3	13.730	11.398	12.039	11.199	15.014	12.750
	4	16.480	17.691	15.181	21.936	15.364	20.301
	6	4.359	4.487	3.263	4.105	3.189	4.216
	7	6.552	5.043	6.362	4.635	7.255	6.601
	8	239.197	269.074	246.378	293.187	246.844	274.668
	10	18.385	11.290	18.014	13.757	16.089	18.093

Type	UC ID	T101	T102	T103	T104	T105	T106
Deep Learning	2	27.717	20.936	24.687	25.479	22.497	27.787
	5	20.012	22.972	17.227	15.123	16.547	16.904
	9	100.362	97.831	116.333	86.784	113.933	106.077
Machine Learning	1	12.381	11.106	13.439	9.434	12.577	7.990
	3	14.392	14.932	8.084	9.907	9.995	16.398
	4	20.501	20.537	17.409	24.716	16.002	20.865
	6	4.537	2.747	3.599	4.045	3.889	3.671
	7	5.457	5.360	7.442	5.630	7.953	5.236
	8	249.590	282.084	273.350	248.388	243.940	275.162
	10	22.701	13.570	17.324	12.227	15.350	20.755

Type	UC ID	T107	T108	T109	T110	T111	T112
Deep Learning	2	25.745	30.338	21.066	36.824	26.629	37.417
	5	16.013	16.085	20.170	12.427	16.725	19.749
	9	99.951	95.269	126.016	91.935	90.281	115.622
Machine Learning	1	16.885	9.831	9.576	9.675	8.431	7.666
	3	16.748	10.440	10.912	12.538	16.081	13.557
	4	14.641	20.810	17.128	16.668	27.982	6.923
	6	3.709	3.613	3.238	4.066	4.818	3.533
	7	5.378	4.794	5.739	6.198	6.222	6.333
	8	268.436	260.155	265.798	295.420	278.075	278.909
	10	12.721	13.755	13.430	15.106	17.780	14.056

Type	UC ID	T113	T114	T115	T116	T117	T118
Deep Learning	2	29.334	27.432	25.595	20.140	31.900	18.874
	5	14.256	22.028	14.653	17.171	19.149	11.487
	9	93.472	98.659	91.535	113.930	98.729	87.526
Machine Learning	1	9.555	9.497	15.141	10.804	9.631	9.199
	3	18.898	15.296	11.702	10.234	14.834	13.164
	4	15.987	18.404	15.445	17.722	21.104	19.010
	6	6.068	3.573	3.225	4.844	4.626	3.678
	7	8.113	4.518	5.449	8.387	6.883	7.181
	8	279.911	253.800	286.842	276.101	253.289	253.620
	10	13.569	14.048	24.254	19.711	15.803	12.623

Type	UC ID	T119	T120	T121	T122	T123	T124
Deep Learning	2	20.249	20.258	27.725	27.267	29.145	31.622
	5	13.553	20.848	13.635	17.539	13.319	13.322
	9	97.959	96.229	107.283	91.256	91.495	98.936
Machine Learning	1	10.969	13.093	15.690	10.510	11.428	12.293
	3	13.418	15.939	8.804	13.888	15.699	13.178
	4	21.718	17.241	18.454	25.439	14.903	15.720
	6	3.573	4.872	3.796	4.164	3.288	3.475
	7	5.974	8.174	5.318	5.118	5.150	6.132
	8	282.895	237.374	271.296	273.539	258.716	273.683
	10	15.266	18.875	17.386	17.448	13.188	16.047

Type	UC ID	T125	T126	T127	T128	T129	T130
Deep Learning	2	29.073	37.126	20.711	25.629	22.514	37.429
	5	18.458	13.675	15.714	9.758	12.612	17.728
	9	102.181	96.111	118.237	94.793	112.913	92.255
Machine Learning	1	10.826	9.488	11.024	15.490	13.741	12.421
	3	14.003	10.331	13.523	15.411	10.815	7.508
	4	20.609	20.417	16.441	18.818	16.781	15.957
	6	5.254	3.986	4.035	3.371	3.278	3.873
	7	6.377	4.752	6.331	8.274	5.633	5.363
	8	248.121	246.691	244.026	269.240	289.199	290.746
	10	19.538	18.183	18.912	18.263	14.890	13.321

Type	UC ID	T131	T132	T133	T134	T135	T136
Deep Learning	2	24.872	25.006	33.566	21.675	21.447	20.373
	5	13.255	22.120	12.351	25.079	19.990	22.896
	9	87.325	102.384	108.475	97.497	93.783	108.270
Machine Learning	1	9.967	15.491	3.806	10.396	11.871	12.149
	3	15.360	12.709	11.675	15.386	9.909	13.883
	4	19.523	12.231	16.867	19.897	27.097	17.795
	6	3.918	4.087	3.659	4.406	4.697	4.770
	7	5.155	6.358	2.427	6.317	5.638	8.389
	8	279.291	282.870	294.872	261.843	265.282	268.761
	10	15.976	14.178	17.050	17.611	19.059	14.698

Type	UC ID	T137	T138	T139	T140
Deep Learning	2	24.297	24.043	19.244	34.832
	5	16.415	21.666	21.597	19.271
	9	92.985	120.875	93.970	109.520
Machine Learning	1	14.003	12.599	10.437	5.142
	3	12.779	15.484	10.602	13.977
	4	18.945	25.191	24.240	15.656
	6	3.945	5.318	3.858	4.669
	7	7.815	2.146	8.798	7.549
	8	276.780	268.170	251.357	273.848
	10	19.278	9.582	12.474	16.580

Table 3-2 Use Case Elapsed Times

3.4 SUT Validation Test Output

Validation Run Report

AIUCpm@1	258.07	T_{Load}	0.70
Scale Factor	1	T_{LD}	0.70
Streams	140	T_{PTT}	31.92
		T_{PST1}	4.22
Kit Version	1.0.3.1	T_{PST2}	4.26
Execution Status	OK	T_{PST}	4.26
Accuracy Status	OK	T_{TT}	0.31

Test Times

Overall Run Start Time	2024-08-06 18:15:06.370
Overall Run End Time	2024-08-06 19:13:22.856
Overall Run Elapsed Time	3,496.486
Load Test Start Time	2024-08-06 18:17:36.023
Load Test End Time	2024-08-06 18:17:36.734
Load Test Elapsed Time	0.711
Power Training Start Time	2024-08-06 18:17:36.736
Power Training End Time	2024-08-06 18:55:15.394
Power Training Elapsed Time	2,258.658
Power Serving 1 Start Time	2024-08-06 18:55:15.396
Power Serving 1 End Time	2024-08-06 18:58:02.728
Power Serving 1 Elapsed Time	167.332
Power Serving 2 Start Time	2024-08-06 18:58:02.731
Power Serving 2 End Time	2024-08-06 19:00:57.036
Power Serving 2 Elapsed Time	174.305
Scoring Start Time	2024-08-06 19:02:12.284
Scoring End Time	2024-08-06 19:06:07.200
Scoring Elapsed Time	234.916
Throughput Start Time	2024-08-06 19:06:07.223
Throughput End Time	2024-08-06 19:13:22.851
Throughput Elapsed Time	435.628

(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	0.310	<=	0.50	Pass
3	mean_squared_log_error	4.575	<=	5.40	Pass
4	f1_score	0.701	>=	0.65	Pass
5	mean_squared_log_error	0.013	<=	0.50	Pass
6	matthews_corrcoef	0.462	>=	0.19	Pass
7	median_absolute_error	0.891	<=	1.80	Pass
8	accuracy_score	0.715	>=	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)
1	2	960	1,920

Total Storage (GB)	1,920
Scale Factor	3
Data Storage Ratio	640.00

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	512	512

Scale Factor	3
Total Memory (GiB)	512
SF / Memory Ratio	0.01

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	527.30	AIUCpm@3
TPCx-AI Price/Performance Metric	230.38	$/AIUCpm@3
TPCx-AI Scale Factor	3	
TPCx-AI Stream Count	140	

Test Times

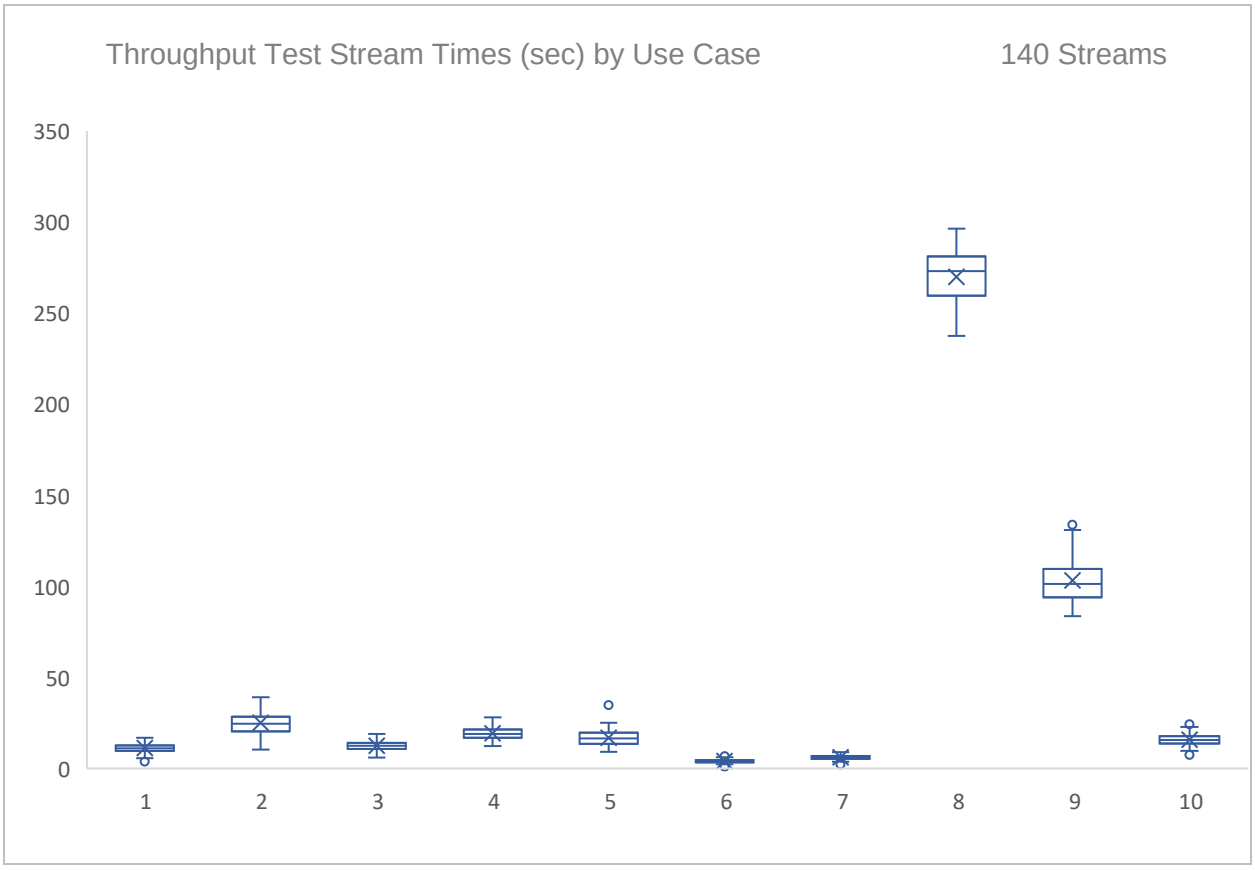
Overall Run Start Time	2024-08-06 19:13:47.925
Overall Run End Time	2024-08-06 20:15:25.841
Overall Run Elapsed Time	3,697.916
Load Test Start Time	2024-08-06 19:16:35.263
Load Test End Time	2024-08-06 19:16:36.442
Load Test Elapsed Time	1.179
Power Training Start Time	2024-08-06 19:16:36.445
Power Training End Time	2024-08-06 19:57:00.335
Power Training Elapsed Time	2,423.890
Power Serving 1 Start Time	2024-08-06 19:57:00.338
Power Serving 1 End Time	2024-08-06 20:00:05.448
Power Serving 1 Elapsed Time	185.110
Power Serving 2 Start Time	2024-08-06 20:00:05.451
Power Serving 2 End Time	2024-08-06 20:03:11.984
Power Serving 2 Elapsed Time	186.533
Scoring Start Time	2024-08-06 20:04:29.415
Scoring End Time	2024-08-06 20:06:47.478
Scoring Elapsed Time	138.063
Throughput Start Time	2024-08-06 20:06:47.501
Throughput End Time	2024-08-06 20:15:25.836
Throughput Elapsed Time	518.335

Accuracy Metrics

Use Case	Metric Name	Metric	Criteria	Threshold	Status
1	N/A	0.000	N/A	0.00	Pass
2	word_error_rate	0.452	<=	0.50	Pass
3	mean_squared_log_error	3.506	<=	5.40	Pass
4	f1_score	0.700	>=	0.65	Pass
5	mean_squared_log_error	0.050	<=	0.50	Pass
6	matthews_corrcoef	0.488	>=	0.19	Pass
7	median_absolute_error	0.957	<=	1.80	Pass
8	accuracy_score	0.761	>=	0.65	Pass
9	accuracy_score	1.000	>=	0.90	Pass
10	accuracy_score	0.816	>=	0.70	Pass

5.2 Throughput Test Stream Times

The following chart shows the minimum, 1st quartile, median, mean (X), 3rd 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.

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

InfoSizing
The Right Metric For Sizing IT

TPC
Certified Auditor

Danny Zhang
Performance Development Manager
7001 Development Drive
Morrisville, NC 27560

September 5, 2024

I verified the TPC Express Benchmark™ AI v2.0.0 performance of the following configuration:

Platform:	1x ThinkSystem SR650 V3
Operating System:	Red Hat Enterprise Linus 8.9 (Ootpa)
Additional Software:	Anaconda Business

The results were:

Performance Metric	**527.30 AIUCpm@3**	
Secondary Metrics	T_{LD}	1.16
	T_{PTT}	51.82
	T_{PST}	6.16
	T_{TT}	0.37

System Under Test	**1x ThinkSystem SR650 V3 with:**		
CPUs	2x Intel Xeon Platinum 8558P		
Memory	512 GiB		
Storage	***Qty***	***Size***	***Type***
	2	960 GB	NVMe (VROC RAID1)

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 v2.0.0.
- 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 3 GB.
- The generated dataset used for testing was protected by VROC RAID1.

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

- 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:

This result was produced using v1.0.3.1 of the TPC-provided kit and is being published under v2.0.0 of the benchmark. The TPC allows this because there were no technical changes made to the benchmark between these two versions and the only change to the TPC-provided kit was the version string.

Respectfully Yours,

Doug Johnson, Certified TPC Auditor

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

Third-Party Price Quotes

Anaconda

ANACONDA.

Anaconda Business: Support

To: **Ray Engler Jr., Lenovo and TPC Consortium | rengler@lenovo.com**
From: Nanette George, Staff Product Marketing Manager, Anaconda | ngeorge@anaconda.com
Re: **Anaconda Support Quote for Publishing TPC Performance Benchmark Testing**

Effective Date: August 2, 2024

Our reliable support team is one reason Anaconda is the platform of choice for one million organizations, including 93% of Fortune 500 companies.

This is a description of levels of service for a **one-year subscription** to Anaconda Business, which includes support. Terms will remain valid for **120 days** following the above effective date.

Premium Support

Included Services	
Anaconda Business, based on this configuration*	$50,000
Anaconda Business Premium Support, including 24×7 support Includes acknowledgment within 4 hours of issues being reported. Once resolution is identified, it will be provided to the customer within 4 hours.	$35,000
Named support contacts, including a dedicated technical account manager (TAM)	5

* This price is subject to change if the software configuration and/or number of users changes.

Software Packages Supported

This support quote applies to all packages and versions of open-source software packages included in the Anaconda repository: https://repo.anaconda.com/pkgs/main

Lenovo's TPCx-AI Benchmark Hardware Configuration

- 1 server, Lenovo ThinkSystem SR650 V3
- 2 processors, Intel Xeon Platinum 8558P 48C 350W 2.7GHz Processor
- 16x32GB DIMMs
- 2 SSDs, ThinkSystem M.2 7450 PRO 960GB Read Intensive NVMe PCIe 4.0 x4 NHS SSD, RAID'd w/VROC
- 1 ThinkSystem Intel E810-DA2 10/25GbE SFP28 2-Port OCP Ethernet Adapter
- RHEL

Anaconda Inc. • 901 E 6th St, Suite 102 • Austin, TX, 78702, USA
Friedrichstrasse 123 • 10117 Berlin, Germany

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 Lenovo	
Sponsor/ModelOptimization/…	Details of model optimization.
Sponsor/ModifiedKitFiles/…	2 modified file(s).
Sponsor/Tuning/…	All tuning files used.