

TPC Express Benchmark™ HS Full Disclosure Report

Transwarp Data Hub

(with 16x PRO-983 Servers; 6x PRO-995 Servers)

Running

TDH 7.0.1

on

CentOS Linux 8

First Edition - January 2023

Transwarp Technology (Shanghai) Co., Ltd. (Transwarp), 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.

Transwarp and the Transwarp Logo are trademarks of Transwarp Technology (Shanghai) Co., Ltd. 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 Transwarp and any other company

TPC Express Benchmark™ HS, TPCx-HS, and HSph, are registered certification marks of the Transaction Processing Performance Council.

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

Copyright© 2023 Transwarp Technology (Shanghai) Co., Ltd.

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
Transwarp	PRO-983	TDH 7.0.1	CentOS Linux 8


TPC Express Benchmark™ HS Metrics			
Total System Cost	HSph@3TB	Price/Performance	Availability Date
¥3,265,525	39.42	¥82,839.30	January 4, 2023

Executive Summary

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

	<h1 style="text-align: center;">Transwarp Data Hub</h1>		TPCx-HS 2.0.3
			TPC Pricing 2.8.0
			Report Date Jan. 04, 2023
Availability Date January 4, 2023	TPCx-HS Performance 39.42 HSph@3TB	Price/Performance ¥82,839.30 ¥ / HSph@3TB	Total System Cost ¥3,265,525 RMB
System Under Test Configuration Overview			
Scale Factor 3	Hadoop Software TDH 7.0.1	Operating System CentOS Linux 8	Other Software None
<div style="display: flex; justify-content: space-around; align-items: flex-start;"> <div style="text-align: center;">  <p>4 x Transwarp Big Data Appliance TxData-4 Server 2 x Intel® Xeon® Gold 5218R Processor 256GB (8 x Samsung 32G DDR4 RECC 2933MT/s) 1 x Intel S4510 M.2-480G SSD 4 x Intel P4510 U.2-2T NVMe SSD 1 x Quad-port 10Gb Ethernet Card, 2x10Gb SFP+, 2x 10Gb RJ45 1 x Mellanox 100G EDR IB Dual-Port QSFP28 Network Adapter</p> </div> <div style="text-align: center;">  <p>1x Mellanox MSB7800-ES2F</p> </div> <div style="text-align: center;">  <p>1x H3C S5110-52P-SI</p> </div> <div style="text-align: center;">  <p>6 x Transwarp Big Data Appliance TxData-2G server 2 x Intel® Xeon® Gold 6330 Processor 256GB (8 x Samsung 32G DDR4 RECC 3200MT/s) 2 x Intel P4510 U.2 1T SSD 6 x Intel P5510 U.2-3.84T NVMe SSD 1 x Intel I350 2-port 1Gb Network Adapter 1 x Mellanox 100G EDR IB Dual-Port QSFP28 Network Adapter</p> </div> </div>			
Physical Storage/Scale Factor: 95.31		Scale Factor/Physical Memory: 0.55	
Total Number of Servers:		22 (16x PRO-983; 6x PRO-995)	
Total Processors/Cores/Threads:		44/976/1,952	
Server Configuration:	Per PRO-983	Per PRO-995	
Processors	2x Intel® Xeon® Gold 5218R CPU @ 2.10GHz	2x Intel® Xeon® Gold 6330 CPU @ 2.00GHz	
Memory	256 GiB	256 GiB	
Storage Controller	On Board	On Board	
Storage Device	1x 480 GB M.2 SSD 4x 2 TB NVMe SSD	2x 1 TB SSD 6x 3.84 TB NVMe SSD	
Network	1x Mellanox 100G EDR IB Dual-Port QSFP28 1x Quad-port 10GB (2x SFP+, 2x RJ45)	1x Mellanox 100G EDR IB Dual-Port QSFP28 1x Intel I350 2-port 1 Gb	
Connectivity:	1x Mellanox MSB7800-ES2F, 1x H3C S5110-520-Si		
Total Rack Units:	(4x3U) + (6x3U) + (1x1U) + (1x1U) = 12U + 18U + 1U + 1U = 32U		

		<h1 style="text-align: center;">Transwarp Data Hub</h1>			TPCx-HS	2.0.3	
					TPC Pricing	2.8.0	
					Report Date	Jan. 04, 2023	
Description		Part Number	Source	Qty	Unit Price	Extended Price	3 Yr. Maint. Price
Hardware							
Transwarp Big Data Appliance TxData-4 Server							
Transwarp Big Data Appliance TxData-4 Server, 2U, 4 Nodes		PRO-983		1	4 ¥ 47,450.00	¥ 189,800.00	
Intel® Xeon® Gold 5218R Processor 27.5M Cache, 2.10 GHz, 20C/40T		PRO-872		1	32 ¥ 13,390.00	¥ 428,480.00	
Samsung 32G DDR4 RECC 2933MT/s		PRO-754		1	128 ¥ 1,570.00	¥ 200,960.00	
Intel S4510 M.2-480G SSD		PRO-632		1	16 ¥ 1,350.00	¥ 21,600.00	
Intel P4510 U.2-2T NVMe SSD		PRO-667		1	64 ¥ 4,200.00	¥ 268,800.00	
Mellanox 100G EDR IB Dual-Port QSFP28 Network Adapter		PRO-556		1	16 ¥ 8,400.00	¥ 134,400.00	
Quad-port 10Gb Ethernet Card, 2x10Gb SFP+, 2x 10Gb RJ45 Network Adapter		PRO-554		1	16 ¥ 4,950.00	¥ 79,200.00	
Transwarp 4-hour 7x24 On-site Service, 3 years		PRO-100		1	1 ¥ 66,162.00		¥ 66,162.00
						¥ -	
Transwarp Big Data Appliance TxData-2G server							
Transwarp Big Data Appliance TxData-2G server, 2U, single node		PRO-995		1	6 ¥ 24,450.00	¥ 146,700.00	
Intel® Xeon® Gold 6330 Processor, 42M Cache, 2.00 GHz, 28C, 56T		PRO-881		1	12 ¥ 22,560.00	¥ 270,720.00	
Samsung 32G DDR4 RECC 3200MT/s		PRO-764		1	48 ¥ 1,670.00	¥ 80,160.00	
Intel P4510 U.2 1T SSD		PRO-633		1	12 ¥ 2,100.00	¥ 25,200.00	
Intel P5510 U.2 3.84T NVMe SSD		PRO-668		1	36 ¥ 6,200.00	¥ 223,200.00	
Intel I350 2-port 1Gb Network Adapter		PRO-518		1	6 ¥ 700.00	¥ 4,200.00	
Mellanox 100G EDR IB Dual-Port QSFP28 Network Adapter		PRO-556		1	6 ¥ 8,400.00	¥ 50,400.00	
Transwarp 4-hour 7x24 On-site Service, 3 years		PRO-100		1	1 ¥ 40,029.00		¥ 40,029.00
						¥ -	
Network and Cables							
Mellanox MSB7800-ES2F Switch-IB 2 Based EDR InfiniBand 1U Switch 36 QSFP28 Ports		MSB7800-ES2F		1	1 ¥ 159,840.00	¥ 159,840.00	
Support and Warranty - 3 Year for MSB7800-ES2F		SUP-MSB7800-ES2F		1	1 ¥ 15,984.00		¥ 15,984.00
Mellanox® Passive Copper cable, IB EDR, up to 100Gb/s, QSFP28, 5m, Black, 26AWG				1	22 ¥ 2,795.00	¥ 61,490.00	
H3C S5110-52P-SI 52-port Gigabit Ethernet Switch,with 1 year Support and Warranty		S5110-52P-SI		1	3 ¥ 2,800.00	¥ 8,400.00	
						¥ -	
Infrastructure							
42U Enclosure system				1	1 ¥ 4,000.00	¥ 4,000.00	
24" LED Monitor				1	3 ¥ 1,500.00	¥ 4,500.00	
Keyboard and Mouse				1	3 ¥ 100.00	¥ 300.00	
						¥ -	
Software							
Transwarp Data Hub 7.0.1 Subscription Edition - 3 Years				1	22 ¥ 30,000.00		¥ 660,000.00
Support Service - 3 Years for CentOS 8				1	22 ¥ 5,500.00		¥ 121,000.00
Subtotals						¥ 2,362,350.00	¥ 903,175.00
Pricing: 1 = Transwarp * 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				Three-Year Cost of Ownership: ¥3,265,525 HSph@3TB: 39.42 ¥ / HSph@3TB: ¥82,839.30			
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.							

 <p>星 环 科 技</p>	<h1>Transwarp Data Hub</h1>	TPCx-HS 2.0.3 TPC Pricing 2.8.0 Report Date Jan. 04, 2023
<h2>Numerical Quantities</h2>		
<h3>Performance Run – Run 1</h3>		
Scale Factor	3TB	
Run Start Time	2022-09-02 21:24:22.000	
Run End Time	2022-09-02 21:28:54.000	
Run Elapsed Time	274.000	
HSGen Start Time	2022-09-02 21:24:23.000	
HSGen End Time	2022-09-02 21:25:05.000	
HSGen Elapsed Time	43.177	
HSSort Start Time	2022-09-02 21:25:09.000	
HSSort End Time	2022-09-02 21:28:26.000	
HSSort Elapsed Time	198.740	
HSValidate Start Time	2022-09-02 21:28:32.000	
HSValidate End Time	2022-09-02 21:28:54.000	
HSValidate Elapsed Time	24.453	
<h3>Repeatability Run – Run 2</h3>		
Scale Factor	3TB	
Run Start Time	2022-09-02 21:29:59.000	
Run End Time	2022-09-02 21:34:29.000	
Run Elapsed Time	273.000	
HSGen Start Time	2022-09-02 21:30:00.000	
HSGen End Time	2022-09-02 21:30:35.000	
HSGen Elapsed Time	35.971	
HSSort Start Time	2022-09-02 21:30:39.000	
HSSort End Time	2022-09-02 21:33:48.000	
HSSort Elapsed Time	191.441	
HSValidate Start Time	2022-09-02 21:33:54.000	
HSValidate End Time	2022-09-02 21:34:29.000	
HSValidate Elapsed Time	37.534	



Transwarp Data Hub

TPCx-HS	2.0.3
TPC Pricing	2.8.0
Report Date	Jan. 04, 2023

Run Reports

Run Report for Performance Run – Run 1

TPCx-HS Performance Metric (HSph@SF) Report

Test Run 1 Details	Total Time =	274
	Total Size =	30000000000
	Scale-Factor =	3

TPCx-HS Performance Metric (HSph@SF): 39.4218

Run Report for Repeatability Run – Run 2

TPCx-HS Performance Metric (HSph@SF) Report

Test Run 2 Details	Total Time =	273
	Total Size =	30000000000
	Scale-Factor =	3

TPCx-HS Performance Metric (HSph@SF): 39.5778


	<h1>Transwarp Data Hub</h1>	TPCx-HS 2.0.3 TPC Pricing 2.8.0 Report Date Jan. 04, 2023
Revision History		
Date	Edition	Description
January 4, 2023	First	Initial Publication

Table of Contents

Abstract..... 3

Executive Summary 3

Table of Contents..... 9

Clause 0 – Preamble10

 0.1 TPC Express Benchmark™ HS Overview.....10

Clause 1 – General Items11

 1.1 Test Sponsor11

 1.2 Parameter Settings11

 1.3 Configuration Diagrams11

 1.3.1 Measured Configuration.....12

 1.3.2 Priced Configuration12

 1.4 Dataset Distribution.....13

 1.5 Software Components Distribution.....13

Clause 2 – Workload Related Items.....14

 2.1 Hardware & Software Tunables14

 2.2 Run Report14

 2.3 Benchmark Kit Identification.....14

 2.4 Benchmark Kit Changes14

Clause 3 – SUT Related Items.....15

 3.1 Data Storage Ratio15

 3.2 Memory Ratio15

Clause 4 – Metrics Related Items16

 4.1 HSGen Time.....16

 4.2 HSSort Time16

 4.3 HSValidate Time16

 4.4 HSDataCheck Times16

 4.5 Performance & Price-Performance16

Auditor’s Information & Letter of Attestation17

Supporting Files Index20

Third-Party Price Quotes.....21

Clause 0 – Preamble

0.1 TPC Express Benchmark™ HS Overview

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

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

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

The purpose of TPC benchmarks is to provide relevant, objective performance data to industry users. To achieve that purpose, TPC benchmark specifications require that benchmark tests be implemented with systems, products, technologies and pricing that:

- Are generally available to users;
- Are relevant to the market segment that the individual TPC benchmark models or represents (e.g., TPCx-HS models and represents Hadoop run-time and Hadoop File-system API compatible systems);
- Would plausibly be implemented by a significant number of users in the market segment the benchmark models or represents.

The use of new systems, products, technologies (hardware or software) and pricing is encouraged so long as they meet the requirements above. Specifically prohibited are benchmark systems, products, technologies or pricing (hereafter referred to as "implementations") whose primary purpose is performance optimization of TPC benchmark results without any corresponding applicability to real-world applications and environments. In other words, all "benchmark special" implementations that improve benchmark results but not real-world performance or pricing, are prohibited.

The rules for pricing are included in the TPC Pricing Specification and rules for energy measurement are included in the TPC Energy Specification. Further information is available at www.tpc.org.

Clause 1 – General Items

1.1 Test Sponsor

A statement identifying the benchmark sponsor(s) and other participating companies must be provided

This benchmark was sponsored by Transwarp Technology (Shanghai) Co., Ltd..

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

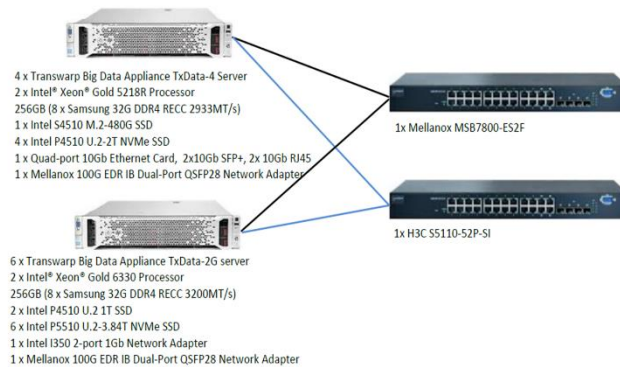


Figure 1-1 Measured Configuration

The measured configuration consisted of:

- Total Nodes: 22 (16x PRO-983; 6x PRO-995)
- Total Processors/Cores/Threads: 44/976/1,952
- Total Memory: 5.50TiB
- Total Number of Storage Drives/Devices: 128
- Total Storage Capacity: 285.92TB

Server node details:

- 16x PRO-983 Servers, each with:
 - Processors/Cores/Threads: 2/40/80
 - Processor Model: Intel® Xeon® Gold 5218R CPU @ 2.10GHz
 - Memory: 256 GiB
 - Controller: On Board
 - Drives:
 - 1x 480 GB M.2 SSD
 - 4x 2 TB NVMe SSD
 - Network
 - 1x Mellanox 100G EDR IB Dual-Port QSFP28
 - 1x Quad-port 10GB
- 6x PRO-995 Servers, each with:
 - Processors/Cores/Threads: 2/56/112
 - Processor Model: Intel® Xeon® Gold 6330 CPU @ 2.00GHz
 - Memory: 256 GiB
 - Controller: On Board
 - Drives:
 - 2x 1 TB SSD
 - 6x 3.84 TB NVMe SSD
 - Network:
 - 1x Mellanox 100G EDR IB Dual-Port QSFP28
 - 1x Intel I350 2-port 1 Gb

Network connectivity detail:

- 1x Mellanox MSB7800-ES2F
- 1x H3C S5110 S5110-520-Si

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
1-16	SATA	sda	Operating System, Root, Swap
17-22	NVME	raid1 (nvme6n1, nvme7n1)	Operating System, Root, Swap
12-14	NVME	nvme0n1, nvme1n1	Hadoop Master
1-16	NVME	nvme0n1, nvme1n1, nvme2n1, nvme3n1	Data, Temp
17-22	NVME	nvme0n1, nvme1n1, nvme2n1, nvme3n1, nvme4n1, nvme5n1	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	Spark		TDFS		ZooKeeper
	Master	Worker	NameNode	DataNode	QuorumPeer
1	X				
6-10					X
12-14			X		
1-22		X		X	

Table 1-2 Software Component Distribution

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

TDH 7.0.1 (fully HDFS compatible at the API level).

Map/Reduce implementation and corresponding version must be disclosed.

TDH 7.0.1 (compatible equivalent to Hadoop 2.7.7).

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 =                274
                        Total Size =                30000000000
                        Scale-Factor =                3

TPCx-HS Performance Metric (HSph@SF):                39.4218

=====
    
```

```

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

Test Run 2 Details      Total Time =                273
                        Total Size =                30000000000
                        Scale-Factor =                3

TPCx-HS Performance Metric (HSph@SF):                39.5778

=====
    
```

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                faeff3091759aac98080be4e39f7896a
TPCx-HS-master_Spark.jar        19f3ce092066e056b884a85ee92fb7fc
TPCx-HS-master.sh                c619a0819571ecd00cd75d2b76ba8c64
    
```

2.4 Benchmark Kit Changes

No modifications were made to the TPC-provided kit.

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)
16	480 GB	7.68
64	2 TB	128.00
12	1 TB	12.00
36	3.84 TB	138.24
Total Storage (TB)		285.92

Table 3-1 Storage Device Capacities

Scale Factor = 3

Data Storage Ratio = (Total Storage (TB) / SF) = 95.31

3.2 Memory Ratio

The Scale Factor to memory ratio must be disclosed.

Total Configured Memory (TiB) = 5.50

Scale Factor to Memory Ratio = (SF / Total Memory(TiB)) = 0.55

Clause 4 – Metrics Related Items

4.1 HSGen Time

The HSGen time must be disclosed for Run1 and Run2.

	Run 1	Run 2
HSGen	43.177	35.971

Table 4-1 HSGen Times

4.2 HSSort Time

The HSSort time must be disclosed for Run1 and Run2.

	Run 1	Run 2
HSSort	198.740	191.441

Table 4-2 HSSort Times

4.3 HSValidate Time

The HSValidate time must be disclosed for Run1 and Run2.

	Run 1	Run 2
HSValidate	24.453	37.534

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)	6.000	6.000

Table 4-4 HSDataCheck Times

4.5 Performance & Price-Performance

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

	Run 1	Run 2
HSph@3TB	39.42	39.57

Table 4-5 Performance Metrics

Run 1 Price-Performance: 82,839.30 ¥/HSph@3TB

Auditor's Information & Letter of Attestation

The auditor's agency name, address, phone number, and Attestation letter must be included in the full disclosure report. A statement should be included specifying who to contact in order to obtain further information regarding the audit process.

This benchmark was audited by Doug Johnson, InfoSizing.

www.sizing.com
63 Lourdes Drive
Leominster, MA 10453
978-343-6562

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

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



Jun Zheng
 Transwarp Technology (Shanghai) Co., Ltd.
 Floor 11 & 12, Block B, No. 88Hongcao Road
 Xuhui District, Shanghai
 China

January 3, 2023

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

Platform: Transwarp Data Hub with:
 16x PRO-983 Servers
 6x PRO-995 Servers

Operating System: CentOS Linux 8
 Apache Hadoop TDH 7.0.1

Compatible Software:

The results were:

Performance Metric 39.42 HSph@3TB
 Run Elapsed Time 274.00 Seconds

Cluster	16x PRO-983, 6x PRO-995 with:		
CPUs	2x Intel® Xeon® Gold 5218R CPU @ 2.10GHz (PRO-983 nodes) 2x Intel® Xeon® Gold 6330 CPU @ 2.00GHz (PRO-995 nodes)		
Memory	256 GiB (all nodes)		
Storage	Qty	Size	Type
	1	480 GB	M.2 SSD (PRO-983 nodes)
	4	2 TB	NVMe SSD (PRO-983 nodes)
	2	1 TB	SSD (PRO-995 nodes)
	6	3.84 TB	NVMe SSD (PRO-995 nodes)

In my opinion, these performance results were produced in compliance with the TPC requirements for the benchmark.

The following verification items were given special attention:

- All TPC-provided components were verified to be v2.0.3
- No modifications were made to any of the Java code
- Any and all modifications to shell scripts were reviewed for compliance

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

- All checksums were validated for compliance
- The generated dataset was properly scaled to 3 TB
- The generated dataset and the sorted dataset were replicated 3-ways
- The elapsed times for all phases and runs were correctly measured and reported
- The Storage and Memory Ratios were correctly calculated and reported
- The system pricing was verified for major components and maintenance
- The major pages from the FDR were verified for accuracy

Additional Audit Notes:

None.

Respectfully Yours,

A handwritten signature in black ink that reads "Doug Johnson". The signature is written in a cursive style with a long horizontal flourish extending to the right.

Doug Johnson, Certified TPC Auditor

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

Supporting Files Index

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

Third-Party Price Quotes

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