Inspur Cloud Company

TPC Express Benchmark™ HS
Full Disclosure Report

InspurCloud Data Cloud

(with 22x Enginetech EG420-G40 Servers)

Running

InspurCloud Data Cloud Platform 5.1.0

On

CentOS Linux 7.6

TPCx-HS Version	2.0.3
Report Edition	First
Report Submitted	January 8, 2024

First Edition - January 2024

Inspur Cloud Company (Inspur Cloud)*, 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 issubject 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.

Inspur Cloud and the Inspur Cloud Logo are trademarks of Inspur Cloud Company and/or its affiliates in the U.S. and other countries. Third party trademarks mentioned are the property of theirrespective owners.The use of the word partner does not imply a partnership relationship between Inspur Cloud and any other company

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

The Inspur Cloud products, services or features identified in this document may not yet be available or may not be available in all areas and maybe subject to change without notice. Consult your local Inspur Cloud business contact for information on the products or services available in your area. You can find additional information via Inspur Cloud's web site at cloud.inspur.com. Actual performance and environmental costs of Inspur Cloud products will vary depending on individual customer configurations and conditions.

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
Inspur Cloud	InspurCloud Data Cloud	InspurCloud Data Cloud Platform 5.1.0	CentOS Linux 7.6

TPC Express Benchmark™ HS Metrics			
Total System Cost	HSph@1TB	Price/Performance	Availability Date
$663,462	50.00	$13,269.24	January 8,2024

Executive Summary

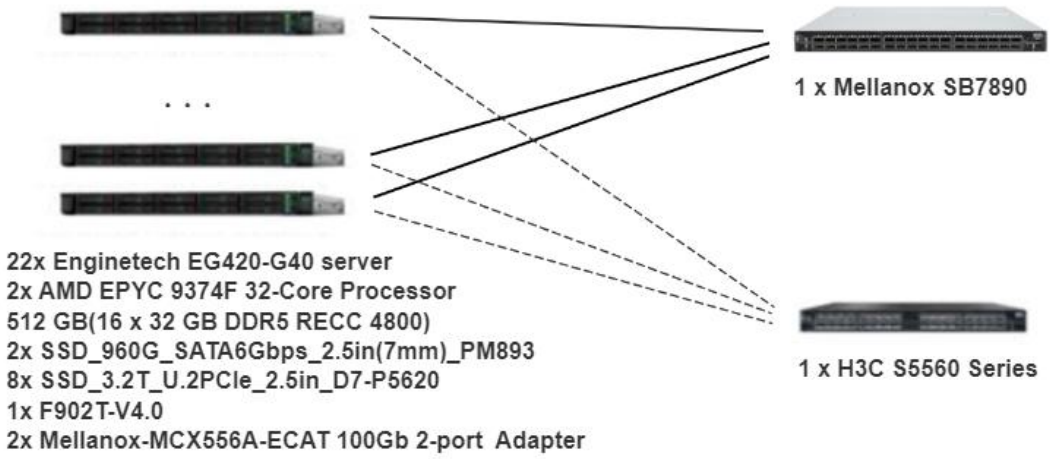
The Executive Summary follows on the next several pages.

inspur 浪潮云	**InspurCloud Data Cloud**	TPCx-HS 2.0.3 TPC Pricing 2.8.0 Report Date Jan. 8, 2024

Availability Date	TPCx-HS Performance	Price/Performance	Total System Cost
January 8,2024	**50.00** **HSph@1TB**	**$13,269.24** **$ / HSph@1TB**	**$663,462 USD**

System Under Test Configuration Overview

Scale Factor	Hadoop Software	Operating System	Other Software
1	InspurCloud Data Cloud Platform 5.1.0	CentOS Linux 7.6	N/A



Physical Storage/Scale Factor: 604.45	Scale Factor/Physical Memory: 0.09
Total Number of Servers:	22 (22x Enginetech EG420-G40 Servers)
Total Processors/Cores/Threads:	44/1408/2,816

Server Configuration:	22x Enginetech EG420-G40 Servers
Processors	2x AMD EPYC 9374F 32-Core Processor
Memory	512 GiB
Storage Device	2x 960 GB SATA SSD 8x 3.2 TB NVMe SSD
Network	1x F902T-V4.0 2x Mellanox-MCX556A-ECAT
Connectivity:	1x Mellanox SB7890 1x H3C S5560 Series
Total Rack Units:	22 (2U) + 1 (1U) + 1 (2U) = 47U

InspurCloud Data Cloud

TPCx-HS	2.0.3
TPC Pricing	2.8.0
Report Date	Jan. 8, 2024

Description	Price Key	Part Number	Unit Price	Qty	Extended Price	3 YrMaint Price
Server Hardware						
Enginetech EG420-G40 Server	1	P54199-B21	$5,566	22	$122,452	
AMD EPYC 9374F 3.85GHz 32-core 320W Processor	1	P54199-B21	$2,514	44	$110,616	
2U Passive CPU Heat Sink for AMD Socket SP5 Processors	1	SNK-P0083P	$42	44	$1,848	
Middle Cooling Fan for 2U Hyper-S Systems 80x80x38mm 13.5K RPM	1	FAN-0209L4-1	$28	88	$2,464	
32GB_DDR5_RECC_4800B_2R*8(M321R4GA3BB6-CQKMS)	1	M321R4GA3BB6	$140	352	$49,280	
SSD_960G_SATA6Gbps_2.5in(7mm)_PM893(MZ7L3960HCJR-00B7C)	1	MZ7L3960HCJR	$140	44	$6,160	
SSD_3.2T_U.2PCIe_2.5in_D7-P5620(SSDPF2KE032T1N1)	1	SSDPF2KE032T1N1	$420	176	$73,920	
1600W redundant single output power supply with inp	1	PWS-1K63A-1R	$210	44	$9,240	
F902T-Gigabit dual port network card	1	F902T	$84	22	$1,848	
Mellanox-MCX556A-ECAT 100Gb 2-port Adapter	1	MCX556A-ECAT	$698	44	$30,712	
Software						
InspurCloud Data Cloud Platform 5.1.0 Subscription Edition - 3 Years	1		$10475	22		$230,450
Inspur Cloud 7x24 On-site Service, 3 years	1					
Other Hardware Components						
Mellanox SB7890	1	SB7890	$23,040	1	$23,040	
H3C S5560 Series Switch	1	S5560	$1,120	1	$1,120	
Keyboards and mouses	1		$32	1	$32	
Morintor	1		$280	1	$280	
		Subtotals	**$ 433,012.00**			**$230,450.00**

Pricing: 1 = Inspur Cloud;

* All discounts are based on US list prices and for similar quantities and configurations. A 35% discount was based on the overall specific components pricing from vendor 1 in this single quotation. Discounts for similarly sized configurations will be similar to those quoted here, but may vary based on the components in the configuration.

Audited by

Three-Year Cost of Ownership: $663,462

HSph@1TB: 50.00

$ / HSph@1TB: $13,269.24

Sales contact: No.1036 Inspur Road, High-Tech Zone, Jinan City, Shandong Province.Tel:400-607-6657

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 pastor 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.

	InspurCloud Data Cloud	TPCx-HS	2.0.3
		TPC Pricing	2.8.0
		Report Date	Jan. 8, 2024

Numerical Quantities

Performance Run - Run 1

Scale Factor	1TB
Run Start Time	2024-01-07 00:08:02.000
Run End Time	2024-01-07 00:09:11.000
Run Elapsed Time	72.000
HSGen Start Time	2024-01-07 00:08:02.000
HSGen End Time	2024-01-07 00:08:14.000
HSGen Elapsed Time	14.090
HSSort Start Time	2024-01-07 00:08:17.000
HSSort End Time	2024-01-07 00:08:59.000
HSSort Elapsed Time	43.452
HSValidate Start Time	2024-01-07 00:09:02.000
HSValidate End Time	2024-01-07 00:09:11.000
HSValidate Elapsed Time	10.050

Repeatability Run - Run 2

Scale Factor	1TB
Run Start Time	2024-01-07 00:10:16.000
Run End Time	2024-01-07 00:11:25.000
Run Elapsed Time	72.000
HSGen Start Time	2024-01-07 00:10:16.000
HSGen End Time	2024-01-07 00:10:28.000
HSGen Elapsed Time	14.110
HSSort Start Time	2024-01-07 00:10:31.000
HSSort End Time	2024-01-07 00:11:13.000
HSSort Elapsed Time	43.450
HSValidate Start Time	2024-01-07 00:11:16.000
HSValidate End Time	2024-01-07 00:11:25.000
HSValidate Elapsed Time	9.788

	InspurCloud Data Cloud	TPCx-HS	2.0.3
		TPC Pricing	2.8.0
		Report Date	Jan. 8, 2024

Run Reports

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

Test Run 1 Details	Total Time =	72
	Total Size =	10000000000
	Scale-Factor =	1.0000
	Framework =	Spark

TPCx-HS Performance Metric (HSph@SF): 50.0000

==========

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

Test Run 2 Details	Total Time =	72
	Total Size =	10000000000
	Scale-Factor =	1.0000
	Framework =	Spark

TPCx-HS Performance Metric (HSph@SF): 50.0000

==========

	InspurCloud Data Cloud	TPCx-HS	2.0.3
		TPC Pricing	2.8.0
		Report Date	Jan. 8, 2024

Revision History

Date	Edition	Description
Jan 8, 2024	First	Initial Publication

Table of Contents

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 (Seewww.tpc.org/tpcx-hsfor 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 Inspur Cloud Information Technology 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 Priced Configuration

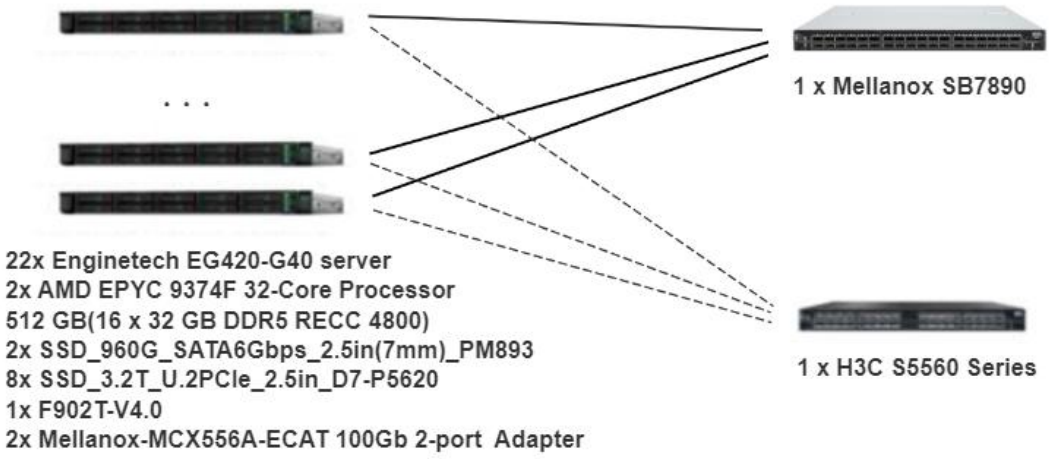


Figure 1-1 Priced Configuration

The priced configuration consists of:

- Total Nodes: 22 (22x Enginetech EG420-G40 Servers)
- Total Processors/Cores/Threads: 44/1408/2,816
- Total Memory: 11TiB
- Total Number of Storage Drives/Devices: 220
- Total Storage Capacity: 604.45TB

Server node details:

- 22x Enginetech EG420-G40 Servers, each with:
  - Processors/Cores/Threads: 2/64/128
  - Processor Model: AMD EPYC 9374F 32-Core Processor
  - Memory: 512 GiB
  - Drives: 2x 0.94 TB SSD 8x 3.2 TB NVMe
  - Network: 1x F902T-V4.0
    2x Mellanox-MCX556A-ECAT

Network connectivity detail:

- 1x Mellanox SB7890 (cluster connectivity)
- 1x H3C S5560 Series (admin)

The distribution of software components over server nodes is detailed in section1.5.

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

Server Node	Controller	Disk Drive	Description of Content
1-22	SATA	sda	Operating System, Root
1,6	NVMe	nvme0n1	Hadoop Master
1-22	NVMe	nvme0n1, nvme1n1, nvme2n1, nvme3n1, nvme4n1, nvme5n1, nvme6n1, nvme7n1	Data, Temp

Table 1-1Dataset Distribution

1.5 Software Components Distribution

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

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

	Spark		HDFS		ZooKeeper
Node	master	worker	NameNode	DataNode	QuorumPeer
1			X		X
2-3					X
4	X				
6			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.

InspurCloud Data Cloud Platform 5.1.0 (fully HDFS compatible at the API level).

Map/Reduce implementation and corresponding version must be disclosed.

InspurCloud Data Cloud Platform 5.1.0 (compatible equivalent to Hadoop 3.1.1.3.1.0.0-78).

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 =                72
                        Total Size =       10000000000
                        Scale-Factor =          1.0000
                        Framework =              Spark

TPCx-HS Performance Metric (HSph@SF):          50.0000

=====================================================


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

Test Run 2 Details      Total Time =                72
                        Total Size =       10000000000
                        Scale-Factor =          1.0000
                        Framework =              Spark

TPCx-HS Performance Metric (HSph@SF):          50.0000

=====================================================
```

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
HSDataCheck.sh	faeff3091759aac98080be4e39f7896a
TPCx-HS-master_Spark.jar	19f3ce092066e056b884a85ee92fb7fc
TPCx-HS-master.sh	b776e15d2d187186ea7911d9ce87e3a7

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

Quantity	Capacity	Total (TB)
44	960 GB	41.25
176	3.2 TB	563.2
Total Storage (TB)		**604.45**

Table 3-1 Storage Device Capacities

Scale Factor = 1

Data Storage Ratio = (Total Storage (TB) / SF) = **604.45**

3.2 Memory Ratio

The Scale Factor to memory ratio must be disclosed.

Total Configured Memory (TiB) = 11

Scale Factor to Memory Ratio = (SF / Total Memory(TiB)) = **0.09**

Clause 4 - Metrics Related Items

4.1 HSGen Time

The HSGen time must be disclosed for Run1 and Run2.

	Run 1	Run 2
HSGen	14.090	14.110

Table 4- 1 HSGen Times

4.2 HSSort Time

The HSSort time must be disclosed for Run1 and Run2.

	Run 1	Run 2
HSSort	43.450	43.450

Table 4-2 HSSort Times

4.3 HSValidate Time

The HSValidate time must be disclosed for Run1 and Run2.

	Run 1	Run 2
HSValidate	10.050	9.788

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)	3.000	3.000
HSDataCheck (post-sort)	3.000	3.000

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@1TB	50.00	50.00

Table 4-5 Performance Metrics

Run 1 Price-Performance: 13,269.24 $/ HSph@1TB

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

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

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 Inspur Cloud.