TPC Express Benchmark™ HS
Full Disclosure Report

Supermicro Cluster
(with 16x AS-1114S-WN10RT Servers; 1x AS-1114S-WTRT Servers)

Running

CDP Private Cloud Base Edition 7.1.6
on
SUSE Linux Enterprise Server 12 SP5

TPCx-HS Version 2.0.3
Report Edition First
Report Submitted September 16, 2021
Super Micro Computer, Inc. (Supermicro), 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.

Supermicro and the Supermicro Logo are trademarks of Super Micro Computer, 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 Supermicro and any other company.

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

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

Copyright© 2021 Super Micro Computer, Inc.

All rights reserved. Permission is hereby granted to reproduce this document in whole or in part provided the copyright notice printed above is set forth in full text or on the title page of each item reproduced.
Abstract

This document contains the methodology and results of the TPC Express Benchmark™ HS (TPCx-HS) test conducted in conformance with the requirements of the TPCx-HS Standard Specification, Revision 2.0.3.

The benchmark results are summarized below.

<table>
<thead>
<tr>
<th>Measured Configuration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Company Name</td>
</tr>
<tr>
<td>Supermicro</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TPC Express Benchmark™ HS Metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total System Cost</td>
</tr>
<tr>
<td>$885,866</td>
</tr>
</tbody>
</table>

Executive Summary

The Executive Summary follows on the next several pages.
Supermicro Cluster

EXECUTIVE SUMMARY

TPCx-HS 2.0.3
Full Disclosure Report
Supermicro Cluster
Supermicro
Report Date
September 16, 2021

TPC Pricing
2.7.0

Report Date
Sep. 16, 2021

Availability Date
Currently Available

TPCx-HS Performance
27.54
HSp@1TB

Price/Performance
$32,166.53
$/HSp@1TB

Total System Cost
$885,866 USD

Physical Storage/Scale Factor: 447.92
Scale Factor/Physical Memory: 0.24

Server Configuration:

<table>
<thead>
<tr>
<th>Processors</th>
<th>Memory</th>
<th>Storage Device</th>
<th>Network</th>
<th>Connectivity</th>
<th>Total Rack Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Per AS-1114S-WN10RT</td>
<td>1x AMD EPYC 75F3</td>
<td>1x 1 TB NVMe (12 nodes)</td>
<td>1x Mellanox Dual-port ConnectX-5 Ex 100 GbE (all nodes)</td>
<td>1x SSE-C3632R 32-port 100 GbE</td>
<td>16x(1U)+1x(1U)+1x(1U)+1x(1U) = 16U+1U+1U+1U = 19U</td>
</tr>
</tbody>
</table>

| Total Processors/Cores/Threads: | 17 (16x AS-1114S-WN10RT; 1x AS-1114S-WTRT) | 17/544/1,088 |

System Under Test Configuration Overview

Scale Factor
1

Hadoop Software
CDP Private Cloud Base Edition 7.1.6

Operating System
SUSE Linux Enterprise Server 12 SP5

Other Software
None
## HARDWARE

### Data Nodes
- **H12SSW-NTR, CSE-116TS-R106WBP5**, RoHS
  - **Part Number**: AS-1114S-WN10RT
  - **Source**: MEM-DR432L-HL01-ER32
  - **Qty**: 1
  - **Unit Price**: $1,477.00
  - **Extended Price**: $23,632.00
- **Kioxia CM6 3.84TB NVMe PCIe 4x4 2.5” 15mm SIE 1DWPD**
  - **Part Number**: HDS-TUN-KCM6XRU1LT84
  - **Source**: AOC-MCK536A-CDAT
  - **Qty**: 1
  - **Unit Price**: $859.00
  - **Extended Price**: $3,396.00

### Network and Cables
- **42U Enclosure system**
  - **Part Number**: SRK-42SE-11
  - **Source**: AP7911B
  - **Qty**: 1
  - **Unit Price**: $1,516.30
  - **Extended Price**: $4,548.40

### Infrastructure
- **Logitech MK200 Media Keyboard and Mouse Combo**
  - **Part Number**: 920-002714
  - **Source**: N/A
  - **Qty**: 3
  - **Unit Price**: $41.50
  - **Extended Price**: $124.50

## TPC Pricing

**Report Date**: Sep. 16, 2021

**TPC Pricing**: 2.7.0

**TPCx-HS**: 2.0.3

**Supermicro Cluster**

**Report Date**: September 16, 2021

---

**HARDWARE**

(continued next page)
Supermicro Cluster

TPC Pricing 2.7.0

Report Date Sep. 16, 2021

(continued from previous page)

**SOFTWARE**

<table>
<thead>
<tr>
<th>Description</th>
<th>Part Number</th>
<th>Source</th>
<th>Qty</th>
<th>Unit Price</th>
<th>Extended Price</th>
<th>3 Yr. Maint. Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUSE Linux Enterprise Server, x86 &amp; x86-64, 1-2 Sockets or 1-2 Virtual Machines, Priority Subscription, 3 Year 874-006883</td>
<td>SFT-NV-SU2P3YBAC</td>
<td>1</td>
<td>17</td>
<td>$2,916.00</td>
<td>$49,572.00</td>
<td></td>
</tr>
<tr>
<td>Cloudera Data Platform Private Cloud Base Edition - Annual Subscription per Node for up to 16 Cores/128 GB RAM for compute and up to 48 TB for storage. BusinessLevel Support.</td>
<td>SMC-CDP-PVBASE-BUS</td>
<td>1</td>
<td>51</td>
<td>$9,600.00</td>
<td>$489,600.00</td>
<td></td>
</tr>
<tr>
<td>COMPUTE: price per CCU per year for compute in excess of 16 cores/128GB RAM per Node, where 1 CCU = 1 core + 8 GB RAM</td>
<td>SMC-CDP-COMPUTE</td>
<td>1</td>
<td>816</td>
<td>75</td>
<td>$61,200.00</td>
<td></td>
</tr>
</tbody>
</table>

**SOFTWARE Subtotals**

$600,372.00 $0.00

**Pricing:** 1 = Supermicro; 2 = APC; 3 = Amazon

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

**Three-Year Cost of Ownership:** $885,866

<table>
<thead>
<tr>
<th>Description</th>
<th>Key</th>
<th>Quantity</th>
<th>Unit Price</th>
<th>Extended Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUSE Linux Enterprise Server</td>
<td>1</td>
<td></td>
<td>$2,916.00</td>
<td>$49,572.00</td>
</tr>
<tr>
<td>Cloudera Data Platform Private Cloud Base Edition</td>
<td>1</td>
<td></td>
<td>$9,600.00</td>
<td>$489,600.00</td>
</tr>
<tr>
<td>COMPUTE: price per CCU per year for compute in excess of 16 cores/128GB RAM per Node</td>
<td>1</td>
<td></td>
<td>$61,200.00</td>
<td></td>
</tr>
</tbody>
</table>

**Audited by Doug Johnson, InfoSizing**

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.
### Numerical Quantities

#### Performance Run – Run 1

<table>
<thead>
<tr>
<th>Scale Factor</th>
<th>1TB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Run Start Time</td>
<td>2021-08-29 09:52:32.000</td>
</tr>
<tr>
<td>Run End Time</td>
<td>2021-08-29 09:54:40.000</td>
</tr>
<tr>
<td>Run Elapsed Time</td>
<td>131.000</td>
</tr>
<tr>
<td>HSGen Start Time</td>
<td>2021-08-29 09:52:33.000</td>
</tr>
<tr>
<td>HSGen End Time</td>
<td>2021-08-29 09:52:57.000</td>
</tr>
<tr>
<td>HSGen Elapsed Time</td>
<td>25.104</td>
</tr>
<tr>
<td>HSSort Start Time</td>
<td>2021-08-29 09:53:02.000</td>
</tr>
<tr>
<td>HSSort End Time</td>
<td>2021-08-29 09:54:20.000</td>
</tr>
<tr>
<td>HSSort Elapsed Time</td>
<td>79.062</td>
</tr>
<tr>
<td>HSValidate Start Time</td>
<td>2021-08-29 09:54:26.000</td>
</tr>
<tr>
<td>HSValidate End Time</td>
<td>2021-08-29 09:54:40.000</td>
</tr>
<tr>
<td>HSValidate Elapsed Time</td>
<td>15.890</td>
</tr>
</tbody>
</table>

#### Repeatability Run – Run 2

<table>
<thead>
<tr>
<th>Scale Factor</th>
<th>1TB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Run Start Time</td>
<td>2021-08-29 09:55:45.000</td>
</tr>
<tr>
<td>Run End Time</td>
<td>2021-08-29 09:57:52.000</td>
</tr>
<tr>
<td>Run Elapsed Time</td>
<td>130.000</td>
</tr>
<tr>
<td>HSGen Start Time</td>
<td>2021-08-29 09:55:46.000</td>
</tr>
<tr>
<td>HSGen End Time</td>
<td>2021-08-29 09:56:10.000</td>
</tr>
<tr>
<td>HSGen Elapsed Time</td>
<td>24.802</td>
</tr>
<tr>
<td>HSSort Start Time</td>
<td>2021-08-29 09:56:15.000</td>
</tr>
<tr>
<td>HSSort End Time</td>
<td>2021-08-29 09:57:32.000</td>
</tr>
<tr>
<td>HSSort Elapsed Time</td>
<td>78.652</td>
</tr>
<tr>
<td>HSValidate Start Time</td>
<td>2021-08-29 09:57:37.000</td>
</tr>
<tr>
<td>HSValidate End Time</td>
<td>2021-08-29 09:57:52.000</td>
</tr>
<tr>
<td>HSValidate Elapsed Time</td>
<td>16.475</td>
</tr>
</tbody>
</table>
Run Report for Performance Run – Run 1

TPC-HS Performance Metric (HSph@SF) Report

Test Run 1 Details
Total Time = 131
Total Size = 10000000000
Scale-Factor = 1

TPC-HS Performance Metric (HSph@SF): 27.5482

Run Report for Repeatability Run – Run 2

TPC-HS Performance Metric (HSph@SF) Report

Test Run 2 Details
Total Time = 130
Total Size = 10000000000
Scale-Factor = 1

TPC-HS Performance Metric (HSph@SF): 27.7008
Supermicro Cluster

Revision History

<table>
<thead>
<tr>
<th>Date</th>
<th>Edition</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>September 16, 2021</td>
<td>First</td>
<td>Initial Publication</td>
</tr>
</tbody>
</table>
# Table of Contents

Abstract............................................................................................................................................... 3
Executive Summary .......................................................................................................................... 3
Table of Contents .............................................................................................................................. 10
Clause 0 – Preamble .......................................................................................................................... 11
  0.1 TPC Express Benchmark™ HS Overview ............................................................................... 11
Clause 1 – General Items .................................................................................................................. 12
  1.1 Test Sponsor ............................................................................................................................ 12
  1.2 Parameter Settings .................................................................................................................. 12
  1.3 Configuration Diagrams ......................................................................................................... 12
    1.3.1 Measured Configuration ................................................................................................. 13
    1.3.2 Priced Configuration ...................................................................................................... 13
  1.4 Dataset Distribution ............................................................................................................... 14
  1.5 Software Components Distribution ....................................................................................... 14
Clause 2 – Workload Related Items .............................................................................................. 15
  2.1 Hardware & Software Tunables ............................................................................................. 15
  2.2 Run Report ............................................................................................................................. 15
  2.3 Benchmark Kit Identification ............................................................................................... 15
  2.4 Benchmark Kit Changes ........................................................................................................ 16
Clause 3 – SUT Related Items ........................................................................................................ 17
  3.1 Data Storage Ratio ................................................................................................................ 17
  3.2 Memory Ratio ........................................................................................................................ 17
Clause 4 – Metrics Related Items ................................................................................................... 18
  4.1 HSGen Time .......................................................................................................................... 18
  4.2 HSSort Time .......................................................................................................................... 18
  4.3 HSVValidate Time ................................................................................................................ 18
  4.4 HSDataCheck Times .............................................................................................................. 18
  4.5 Performance & Price-Performance ...................................................................................... 18
Auditor's Information & Letter of Attestation .................................................................................... 19
Supporting Files Index .................................................................................................................... 22
Third-Party Price Quotes ................................................................................................................. 23
  APC............................................................................................................................................... 23
  Amazon......................................................................................................................................... 24
Clause 0 – Preamble

0.1 TPC Express Benchmark™ HS Overview

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

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

The TPCx-HS kit is available from the TPC (See www.tpc.org/tpcx-hs 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 Super Micro Computer, Inc..

1.2 Parameter Settings

Settings must be provided for all customer-tunable parameters and options which have been changed from the defaults found in actual products, including by not limited to:

- Configuration parameters and options for server, storage, network and other hardware component incorporated into the pricing structure;
- Configuration parameters and options for operating system and file system component incorporated into the pricing structure;
- Configuration parameters and options for any other software component incorporated into the pricing structure;
- Compiler optimization options.

Comment 1: In the event that some parameters and options are set multiple times, it must be easily discernible by an interested reader when the parameter or option was modified and what new value it received each time.

Comment 2: This requirement can be satisfied by providing a full list of all parameters and options, as long as all those that have been modified from their default values have been clearly identified and these parameters and options are only set once.

The supporting files contain the parameters and options used to configure the components involved in this benchmark.

1.3 Configuration Diagrams

Diagrams of both measured and priced configurations must be provided, accompanied by a description of the differences. This includes, but is not limited to:

- Total number of nodes used;
- Total number and type of processors used/total number of cores used/total number of threads used (including sizes of L2 and L3 caches);
- Size of allocated memory, and any specific mapping/partitioning of memory unique to the test;
- Number and type of disk units (and controllers, if applicable);
- Number of channels or bus connections to disk units, including their protocol type;
- Number of LAN (e.g., Ethernet) connections and speed for switches and other hardware components physically used in the test or are incorporated into the pricing structure;
- Type and the run-time execution location of software components.
1.3.1 Measured Configuration

The measured configuration consisted of:

- **Total Nodes:** 17 (16x AS-1114S-WN10RT; 1x AS-1114S-WTRT)
- **Total Processors/Cores/Threads:** 17/544/1,088
- **Total Memory:** 4.25TiB
- **Total Number of Storage Drives/Devices:** 130
- **Total Storage Capacity:** 447.92TB

Server node details:

<table>
<thead>
<tr>
<th>16x AS-1114S-WN10RT Servers, each with:</th>
<th>1x AS-1114S-WTRT Servers, each with:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Processors/Cores/Threads: 1/32/64</td>
<td>• Processors/Cores/Threads: 1/32/64</td>
</tr>
<tr>
<td>• Processor Model: AMD EPYC 75F3</td>
<td>• Processor Model: AMD EPYC 75F3</td>
</tr>
<tr>
<td>• Memory: 256 GiB</td>
<td>• Memory: 256 GiB</td>
</tr>
<tr>
<td>• Drives:</td>
<td>• Drives:</td>
</tr>
<tr>
<td>o 1x 1 TB NVMe (12 nodes)</td>
<td>o 2x 1 TB NVMe</td>
</tr>
<tr>
<td>o 1x 960 GB NVMe (4 nodes)</td>
<td>Network:</td>
</tr>
<tr>
<td>o 7x 3.84 TB NVMe (all nodes)</td>
<td>o 1x Mellanox Dual-port ConnectX-5</td>
</tr>
<tr>
<td>• Network:</td>
<td>Ex 100 GbE (12 nodes)</td>
</tr>
<tr>
<td>o 1x Mellanox Dual-port ConnectX-5 Ex 100 GbE (12 nodes)</td>
<td></td>
</tr>
<tr>
<td>o 1x Mellanox Dual-port ConnectX-5 100 GbE (4 nodes)</td>
<td></td>
</tr>
<tr>
<td>o 1x Broadcom Dual-port 10 GbE (all nodes)</td>
<td></td>
</tr>
</tbody>
</table>

Network connectivity detail:

- 1x SSE-C3632R 32-port 100 GbE
- 1x E1031 48-port 1/10 GbE

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.

<table>
<thead>
<tr>
<th>Server Node</th>
<th>Controller</th>
<th>Disk Drive</th>
<th>Description of Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NVMe</td>
<td>nvme7n1</td>
<td>Operating System, Root, Swap, Hadoop Master</td>
</tr>
<tr>
<td>2-3</td>
<td>NVMe</td>
<td>nvme7n1</td>
<td>Operating System, Root, Swap, Hadoop Master</td>
</tr>
<tr>
<td>2-3</td>
<td>NVMe</td>
<td>nvme0n1, nvme1n1, nvme2n1, nvme3n1, nvme4n1, nvme5n1, nvme6n1</td>
<td>Data, Temp</td>
</tr>
<tr>
<td>4-17</td>
<td>NVMe</td>
<td>nvme7n1</td>
<td>Operating System, Root, Swap, Hadoop Master</td>
</tr>
<tr>
<td>4-17</td>
<td>NVMe</td>
<td>nvme0n1, nvme1n1, nvme2n1, nvme3n1, nvme4n1, nvme5n1, nvme6n1</td>
<td>Data, Temp</td>
</tr>
</tbody>
</table>

*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.

<table>
<thead>
<tr>
<th>Node</th>
<th>Map/Reduce</th>
<th>HDFS</th>
<th>ZooKeeper</th>
<th>Spark</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Manager</td>
<td>NameNode</td>
<td>DataNode</td>
<td>QuorumPeer</td>
</tr>
<tr>
<td>1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>2-3</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>4-17</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

*Table 1-2 Software Component Distribution*

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

CDP Private Cloud Base Edition 7.1.6 (fully HDFS compatible at the API level).

Map/Reduce implementation and corresponding version must be disclosed.

CDP Private Cloud Base Edition 7.1.6 (compatible equivalent to Hadoop 3.1.1.7.1.6.0-297).
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 = 131
Total Size = 10000000000
Scale-Factor = 1

TPCx-HS Performance Metric (HSph@SF): 27.5482
```

**Run Report for Run 2 – Repeatability Run**

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

Test Run 2 Details
Total Time = 130
Total Size = 10000000000
Scale-Factor = 1

TPCx-HS Performance Metric (HSph@SF): 27.7008
```

2.3 Benchmark Kit Identification

*Version number of TPCx-HS kit and checksum for HSGen, HSSort and HSValidate Programs must be reported.*

<table>
<thead>
<tr>
<th>File Name</th>
<th>MD5</th>
</tr>
</thead>
<tbody>
<tr>
<td>BigData_cluster_validate_suite.sh</td>
<td>57f7cd68251a9aba0f6648630f5fda</td>
</tr>
<tr>
<td>HSDataCheck.sh</td>
<td>bcf0b946a49d1249c9da174b5d9805f1</td>
</tr>
<tr>
<td>TPCx-HS-master_Spark.jar</td>
<td>19f3ce092066e056b884a85ee92fb7fc</td>
</tr>
<tr>
<td>TPCx-HS-master.sh</td>
<td>c619a0819571ecd00cd75d2b76ba8c64</td>
</tr>
</tbody>
</table>
2.4 Benchmark Kit Changes

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

To collect the necessary data for auditing, the HSDataCheck.sh script was modified. In accordance with the TPCx-HS Standard Specification, this change received prior approval from the TPCx-HS subcommittee.
Clause 3 – SUT Related Items

3.1 Data Storage Ratio

*The data storage ratio must be disclosed.*

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

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Capacity</th>
<th>Total (TB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>1 TB</td>
<td>12.00</td>
</tr>
<tr>
<td>4</td>
<td>960 GB</td>
<td>3.84</td>
</tr>
<tr>
<td>112</td>
<td>3.84 TB</td>
<td>430.08</td>
</tr>
<tr>
<td>2</td>
<td>1 TB</td>
<td>2.00</td>
</tr>
</tbody>
</table>

**Total Storage (TB)**

447.92

*Table 3-1 Storage Device Capacities*

Scale Factor = 1

**Data Storage Ratio** = (Total Storage (TB) / SF) = 447.92

3.2 Memory Ratio

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

Total Configured Memory (TiB) = 4.25

**Scale Factor to Memory Ratio** = (SF / Total Memory(TiB)) = 0.24
Clause 4 – Metrics Related Items

4.1 HSGen Time
The HSGen time must be disclosed for Run1 and Run2.

<table>
<thead>
<tr>
<th></th>
<th>Run 1</th>
<th>Run 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSGen</td>
<td>25.104</td>
<td>24.802</td>
</tr>
</tbody>
</table>

Table 4-1 HSGen Times

4.2 HSSort Time
The HSSort time must be disclosed for Run1 and Run2.

<table>
<thead>
<tr>
<th></th>
<th>Run 1</th>
<th>Run 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSSort</td>
<td>79.062</td>
<td>78.652</td>
</tr>
</tbody>
</table>

Table 4-2 HSSort Times

4.3 HSValidate Time
The HSValidate time must be disclosed for Run1 and Run2.

<table>
<thead>
<tr>
<th></th>
<th>Run 1</th>
<th>Run 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSValidate</td>
<td>15.890</td>
<td>16.475</td>
</tr>
</tbody>
</table>

Table 4-3 HSValidate Times

4.4 HSDataCheck Times
Both HSDataCheck times must be disclosed for Run1 and Run2.

<table>
<thead>
<tr>
<th></th>
<th>Run 1</th>
<th>Run 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSDataCheck (pre-sort)</td>
<td>5.000</td>
<td>5.000</td>
</tr>
<tr>
<td>HSDataCheck (post-sort)</td>
<td>6.000</td>
<td>5.000</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th></th>
<th>Run 1</th>
<th>Run 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSph@1TB</td>
<td>27.54</td>
<td>27.70</td>
</tr>
</tbody>
</table>

Table 4-5 Performance Metrics

Run 1 Price-Performance: 32,166.53 $/ 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 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.
Srini Bala  
Super Micro Computer, Inc.  
980 Rock Avenue,  
San Jose, CA 95131  
USA  

September 12, 2021  

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

Platform:  
Supermicro Cluster with:  
16x AS-1114S-WN10RT Servers (Data Nodes)  
1x AS-1114S-WTRT Server (Master Node)  

Operating System:  
SUSE Linux Enterprise Server 12 SP5  

Apache Hadoop:  
CDP Private Cloud Base Edition 7.1.6  

Compatible Software:  

The results were:  

**Performance Metric** 27.54 HSp@1TB  
Run Elapsed Time 131.00 Seconds  

<table>
<thead>
<tr>
<th>Cluster</th>
<th>16x AS-1114S-WN10RT, 1x AS-1114S-WTRT with:</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPUs</td>
<td>1x AMD® EPYC 75F3 32-Core Processor (all nodes)</td>
</tr>
<tr>
<td>Memory</td>
<td>256 GiB (all nodes)</td>
</tr>
<tr>
<td>Storage</td>
<td>Qty</td>
</tr>
<tr>
<td>2</td>
<td>1 TB</td>
</tr>
<tr>
<td>1</td>
<td>1 TB</td>
</tr>
<tr>
<td>1</td>
<td>960 GB</td>
</tr>
<tr>
<td>7</td>
<td>3.84 TB</td>
</tr>
</tbody>
</table>

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

The following verification items were given special attention:  

- All TPC-provided components were verified to be v2.0.3  
- No modifications were made to any of the Java code  
- Any and all modifications to shell scripts were reviewed for compliance  
- All checksums were validated for compliance
- The generated dataset was properly scaled to 1 TB
- The generated dataset and the sorted dataset were erasure coded with a policy of RS-6-3-1024k
- The elapsed times for all phases and runs were correctly measured and reported
- The Storage and Memory Ratios were correctly calculated and reported
- The system pricing was verified for major components and maintenance
- The major pages from the FDR were verified for accuracy

Additional Audit Notes:

None.

Respectfully Yours,

[Signature]

Doug Johnson, Certified TPC Auditor
## Supporting Files Index

<table>
<thead>
<tr>
<th>Clause</th>
<th>Description</th>
<th>Archive File Pathname</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clause 1</td>
<td>Parameters and options used to configure the system</td>
<td>SupportingFiles/Clause1</td>
</tr>
<tr>
<td>Clause 2</td>
<td>Configuration scripts and Run Report</td>
<td>SupportingFiles/Clause2</td>
</tr>
<tr>
<td>Clause 3</td>
<td>System configuration details</td>
<td>SupportingFiles/Clause3</td>
</tr>
</tbody>
</table>
Third-Party Price Quotes

APC

Rack PDU, Switched, 2U, 30A, 208V, (16)C13

AP7911B

$1,025.00

In Stock

Add To Cart

Add to Favorites

Compare

Get Support

View User Manual

Update Software/Firmware

View Product FAQs

Register Your Product

Our Partners

Show all

TPCx-HS 2.0.3
Full Disclosure Report
Supermicro
Supermicro Cluster
Report Date
September 16, 2021
Amazon

LONCEVON-12 inch IPS 1920x1080P Portable Small HDMI Monitor with VGA-AV-USB-BNC Port; External Small HDMI Screen LCD Display Monitor for Laptop Raspberry Pi/TV; Build in Dual Speakers, Remote.

Visit the LONCEVON Store

Price: $99.99 & FREE Returns

Get $50 off instantly: Pay $49.99 $49.99 upon approval for the Amazon Rewards Visa Card. No annual fee

Logitech Media Combo MK200 Full-Size Keyboard and High-Definition Optical Mouse (920-002714)

Price: $41.50 & FREE Returns

Get $10 off instantly: Pay $31.50 upon approval for the Amazon Store Card.

- Instant media and internet access
- Eight shortcut keys
- Full-size keyboard
- Comfortable, quiet typing
- Instant access to applications