*Cisco Systems, Inc.*

---

# TPC Express Benchmark™ HS (TPCx-HS)

# Full Disclosure Report

# for

# Cisco UCS Integrated Infrastructure for Big Data and Analytics

# (with 17 UCS C240M5 Servers)

# using

# Cloudera Enterprise Edition V5.13.0

# and

# Red Hat Enterprise Linux Server Release 7.4

---

**First Edition**

**January 30, 2018**

***Cisco Systems, Inc. (Cisco)****, 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 as a result of these and other factors. Therefore, the TPC Express Benchmark™ HS 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.*

*Cisco and the Cisco Logo are trademarks of Cisco Systems, Inc. and/or its affiliates in the U.S. and other countries. A listing of Cisco's trademarks can be found at www.cisco.com/go/trademarks. Third party trademarks mentioned are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company.*

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

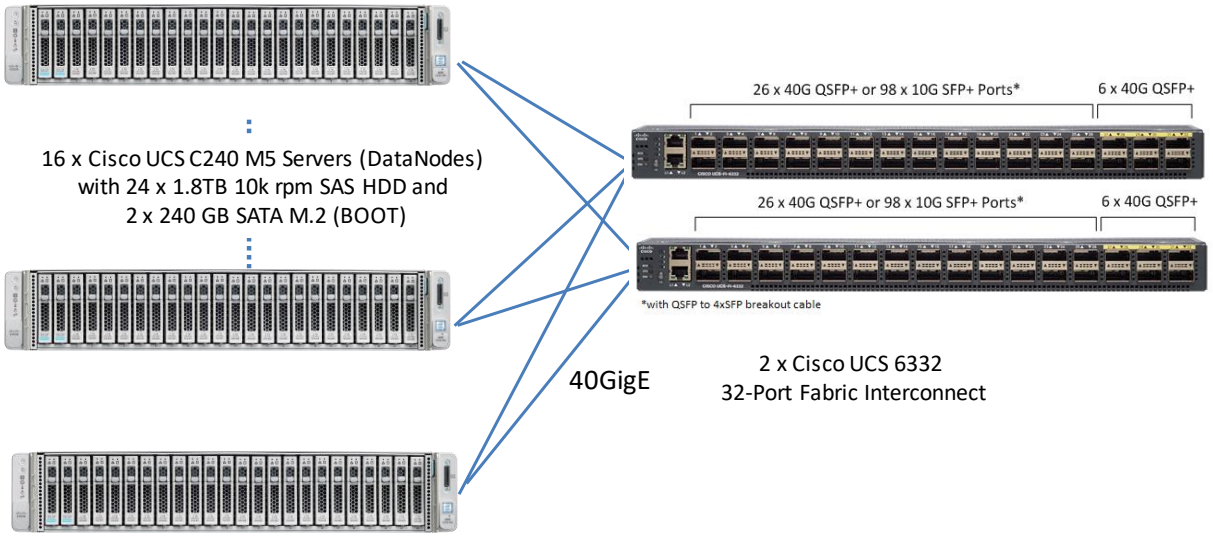
*The Cisco products, services or features identified in this document may not yet be available or may not be available in all areas and may be subject to change without notice. Consult your local Cisco business contact for information on the products or services available in your area. You can find additional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

| | | |
|---|---|---|
| | **Cisco UCS Integrated Infrastructure for Big Data and Analytics** | TPCx-HS Rev. v2.0.2<br>TPC-Pricing Rev. v2.2.0<br>Report Date:<br>**January 30, 2018** |

| Total System Cost | TPCx-HS Performance Metric | Price/Performance |
|---|---|---|
| **656,096 USD** | **11.13**<br>HSph@10TB | **58,948.43 USD**<br>$/HSph@10TB |

| Scale Factor | Apache Hadoop Compatible Software | Operating System | Other Software | Availability Date |
|---|---|---|---|---|
| 10 | Cloudera Enterprise Edition V5.13.0 | Red Hat Enterprise Linux Server Release 7.4 | n/a | January 30, 2018 |

## System Configuration




| Physical Storage/Scale Factor: 70.30 | Scale Factor/Physical Memory: 3.14 |
|---|---|

| | |
|---|---|
| Servers: | 17 x UCS C240M5 |
| Total Processors/Cores/Threads | 34/476 /952 |
| Server Configuration: | Per node: |
| Processors | 2 x Intel® Xeon® Gold Processor 6132, 2.60 GHz, 19.25 MB L3 |
| Memory | 192GB |
| Storage Controller | 1 x Cisco 12G SAS Modular Raid Controller |
| Storage Device | 24 x 1.8TB 10K RPM SAS HDD (data nodes)<br>2 x 1.8TB 10K RPM SAS HDD (name node)<br>2 x 240GB SATA (all nodes, boot disk) |
| Network | 1 x Cisco VIC 1387 Dual Port 40Gb QSFP CNA MLOM |
| Connectivity: | 2 x Cisco UCS 6332 32-Port Fabric Interconnect |
| Total Rack Units: | (17*C240M5)+(2*FI) = (17 * 2) + (2 * 1) = 36RU |

| Description | Part Number | Source | Unit Price | Qty | Extended Price | 3 Year Maint. Price |
|---|---|---|---|---|---|---|
| UCS C240 M5 24 SFF + 2 rear drives w/o CPU,mem,HD,PCIe,PS | UCSC-C240-M5SX | 1 | $ 4,399.00 | 17 | $ 74,783.00 | |
| 2.6 GHz 6132/140W 14C/19.25MB Cache/DDR4 2666MHz | UCS-CPU-6132 | 1 | $ 7,000.00 | 34 | $ 238,000.00 | |
| 16GB DDR4-2666-MHz RDIMM/PC4-21300/single rank/x4/1.2v | UCS-MR-X16G1RS-H | 1 | $ 900.00 | 204 | $ 183,600.00 | |
| Riser 1 incl 3 PCIe slots (x8, x16, x8); slot 3 req CPU2 | UCSC-PCI-1-C240M5 | 1 | $ 199.00 | 17 | $ 3,383.00 | |
| Cisco VIC 1387 Dual Port 40Gb QSFP CNA MLOM | UCSC-MLOM-C40Q-03 | 1 | $ 2,192.00 | 17 | $ 37,264.00 | |
| 240GB SATA M.2 | UCS-M2-240GB | 1 | $ 535.00 | 34 | $ 18,190.00 | |
| Cisco UCS 1600W AC Power Supply for Rack Server | UCSC-PSU1-1600W | 1 | $ 800.00 | 34 | $ 27,200.00 | |
| Power Cord, 125VAC 13A NEMA 5-15 Plug, North America | CAB-9K12A-NA | 1 | $ - | 34 | $ - | |
| Ball Bearing Rail Kit for C220 & C240 M4 & M5 rack servers | UCSC-RAILB-M4 | 1 | $ 220.00 | 17 | $ 3,740.00 | |
| IMC SW (Recommended) latest release for C-Series Servers. | CIMC-LATEST | 1 | $ - | 17 | $ - | |
| Heat sink for UCS C240 M5 rack servers 150W CPUs & below | UCSC-HS-C240M5 | 1 | $ - | 34 | $ - | |
| UCS C-Series M5 SFF drive blanking panel | UCSC-BBLKD-S2 | 1 | $ - | 56 | $ - | |
| C240 M5 PCIe Riser Blanking Panel | UCSC-PCIF-240M5 | 1 | $ - | 17 | $ - | |
| Mini Storage carrier for M.2 SATA/NVME (holds up to 2) | UCS-MSTOR-M2 | 1 | $ - | 17 | $ - | |
| Super Cap for UCSC-RAID-M5, UCSC-MRAID1GB-KIT | UCSC-SCAP-M5 | 1 | $ - | 17 | $ - | |
| Super Cap cable for UCSC-RAID-M5HD | CBL-SC-MR12GM5P | 1 | $ - | 17 | $ - | |
| Cisco 12G Modular RAID controller with 4GB cache | UCSC-RAID-M5HD | 1 | $ 2,900.00 | 17 | $ 49,300.00 | |
| Cisco ONE Data Center Compute Opt Out Option | C1UCS-OPT-OUT | 1 | $ - | 17 | $ - | |
| UCS SP 1.8 TB 12G SAS 10K RPM SFF HDD (4K) 4Pk | UCS-SP-H1P8TB-4X | 1 | $ - | 96 | $ - | |
| 1.8 TB 12G SAS 10K RPM SFF HDD (4K) | UCS-SP-H1P8TB | 1 | $ 1,500.00 | 384 | $ 576,000.00 | |
| 1.8 TB 12G SAS 10K RPM SFF HDD (4K) | UCS-HD18TB10K4KN | 1 | $ 1,827.00 | 2 | $ 3,654.00 | |
| SmartNet total care 3YR support (24x7x4) | CON-3OSP-C240M5A2 | 1 | $ 2,364.00 | 17 | | $40,188 |
| | | | | | | |
| UCS SP Select 2 x 6332 FI | UCS-SP-FI6332-2X | 1 | $ - | 1 | $ - | |
| (Not sold standalone) UCS 6332 1RU FI/12 QSFP+ | UCS-SP-FI6332 | 1 | $ 22,000.00 | 2 | $ 44,000.00 | |
| UCS 6332 Power Supply/100-240VAC | UCS-PSU-6332-AC | 1 | $ - | 4 | $ - | |
| 40GBASE-CR4 Passive Copper Cable, 3m | QSFP-H40G-CU3M | 1 | $ - | 16 | $ - | |
| QSFP40G BiDi Short-reach Transceiver | QSFP-40G-SR-BD | 1 | $ - | 8 | $ - | |
| UCS Manager v3.1 | N10-MGT014 | 1 | $ - | 2 | $ - | |
| UCS 6332 Fan Module | UCS-FAN-6332 | 1 | $ - | 8 | $ - | |
| UCS 6332 Chassis Accessory Kit | UCS-ACC-6332 | 1 | $ - | 2 | $ - | |
| 40GBASE-CR4 Passive Copper Cable, 3m | QSFP-H40G-CU3M= | 1 | $ 250.00 | 20 | $ 5,000.00 | |
| Power Cord, 125VAC 13A NEMA 5-15 Plug, North America | CAB-9K12A-NA | 1 | $ - | 4 | $ - | |
| 3rd Gen FI Per port License to connect C-direct only | UCS-LIC-6300-40GC= | 1 | $ 1,388.00 | 20 | $ 27,760.00 | |
| 3YR SMARTNET24X7X4OS (Not sold standalone) UCS6332 | CON-3OSP-SPFI6332 | 1 | $ 6,678.00 | 1 | | $6,678 |

| | Cisco UCS Integrated Infrastructure for Big Data and Analytics | TPCx-HS Rev. v2.0.2<br>TPC-Pricing Rev. v2.2.0<br>Report Date:<br>January 30, 2018 |
|---|---|---|

| Description | Part Number | Source | Unit Price | Qty | Extended Price | 3 Year Maint. Price |
|---|---|---|---|---|---|---|
| Cisco R42612 standard rack, w/side panels | RACK2-UCS2 | 1 | $ 6,341.00 | 1 | $ 6,341.00 | |
| Red Hat Enterprise Linux Server, 3Y 24x7 | CON-ISV1-EL2S2V3A | 1 | $ 3,897.00 | 17 | $66,249 | |
| Cloudera Enterprise Edition, 3Y 24x7 | UCS-BD-CEBN-GD= | 1 | $14,057.00 | 17 | $238,969 | |
| | | | | Total | $1,603,433.00 | $46,866 |
| Large Purchase Discount [1] | 61% for products and 35% for service | 1 | | | $ (978,094.13) | $ (16,403.10) |
| Acer V206HQL-LED monitor -20" (Inc 2 spares) | UM.1V6AA.A02 | 2 | $ 79.99 | 3 | $ 239.97 | |
| Logitech USB Corded Keyboard/Mouse Combo MK120 (Inc 2 spares) | 920-002565 | 2 | $ 17.99 | 3 | $ 53.97 | |

Pricing:1 = Cisco, 2 = CDW

[1] All discounts are based on US list prices and for similar quantities and configurations. The discounts are based on the overall specific components pricing from respective vendors 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 Doug Johnson of InfoSizing**

| | |
|---|---|
| **Three-Year Cost of Ownership** | **$656,096** |
| **HSph@10TB** | **11.13** |
| **$/HSph@10TB** | **$58,948.43** |

Prices used in TPC benchmarks reflect the actual prices a customer would pay for a one-time purchase of the stated components. 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 components. For complete details, see the pricing sections of the TPC benchmark specifications. If you find that the stated prices are not available according to these terms, please inform at pricing@tpc.org. Thank you.

## Numerical Quantities

Performance Run – Run 1

| | |
|---|---|
| Scale Factor | 10TB |
| | |
| Run Start Time | 2018-01-23 19:02:09.000 |
| Run End Time | 2018-01-23 19:55:59.000 |
| Run Elapsed Time | 3,232.000 |
| | |
| HSGen Start Time | 2018-01-23 19:02:11.000 |
| HSGen End Time | 2018-01-23 19:13:42.000 |
| HSGen Elapsed Time | 693.141 |
| | |
| HSSort Start Time | 2018-01-23 19:13:45.000 |
| HSSort End Time | 2018-01-23 19:50:23.000 |
| HSSort Elapsed Time | 2,199.504 |
| | |
| HSValidate Start Time | 2018-01-23 19:50:27.000 |
| HSValidate End Time | 2018-01-23 19:55:59.000 |
| HSValidate Elapsed Time | 333.302 |

Repeatability Run – Run 2

| | |
|---|---|
| Scale Factor | 10TB |
| | |
| Run Start Time | 2018-01-23 20:01:04.000 |
| Run End Time | 2018-01-23 20:50:33.000 |
| Run Elapsed Time | 2,971.000 |
| | |
| HSGen Start Time | 2018-01-23 20:01:06.000 |
| HSGen End Time | 2018-01-23 20:12:27.000 |
| HSGen Elapsed Time | 682.876 |
| | |
| HSSort Start Time | 2018-01-23 20:12:29.000 |
| HSSort End Time | 2018-01-23 20:45:35.000 |
| HSSort Elapsed Time | 1,986.172 |
| | |
| HSValidate Start Time | 2018-01-23 20:45:38.000 |
| HSValidate End Time | 2018-01-23 20:50:33.000 |
| HSValidate Elapsed Time | 296.144 |

## Run Reports

Run Report for Performance Run – Run 1

==========

TPCx-HS Performance Metric (HSph@SF) Report

| | | |
|---|---|---|
| Test Run 1 Details | Total Time = | 3232 |
| | Total Size = | 100000000000 |
| | Scale-Factor = | 10 |

TPCx-HS Performance Metric (HSph@SF): 11.1395

==========

Run Report for Repeatability Run – Run 2

==========

TPCx-HS Performance Metric (HSph@SF) Report

| | | |
|---|---|---|
| Test Run 2 Details | Total Time = | 2971 |
| | Total Size = | 100000000000 |
| | Scale-Factor = | 10 |

TPCx-HS Performance Metric (HSph@SF): 12.1182

==========

# Table of Contents

# 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 v2.0.2.

The test was conducted at a Scale Factor of 10 with 17 UCS C240M5 Servers running Cloudera Enterprise Edition V5.13.0 on Red Hat Enterprise Linux Server Release 7.4.

**Measured Configuration**

| Company Name | Cluster Node | Virtualization | Operating System |
|---|---|---|---|
| Cisco Systems, Inc. | UCS C240M5 | n/a | Red Hat Enterprise Linux Server Release 7.4 |

**TPC Express Benchmark© HS Metrics**

| Total System Cost | HSph@10TB | Price/Performance | Availability Date |
|---|---|---|---|
| 656,096 USD | 11.13 | 58,948.43 USD | January 30, 2018 |

# Preface

## TPC Express Benchmark™ HS Overview

*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. The 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 asses a broad range of system topologies and implementation of Hadoop clusters. The TPCx-HS can be used to asses 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-H 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 Cisco Systems, 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.*

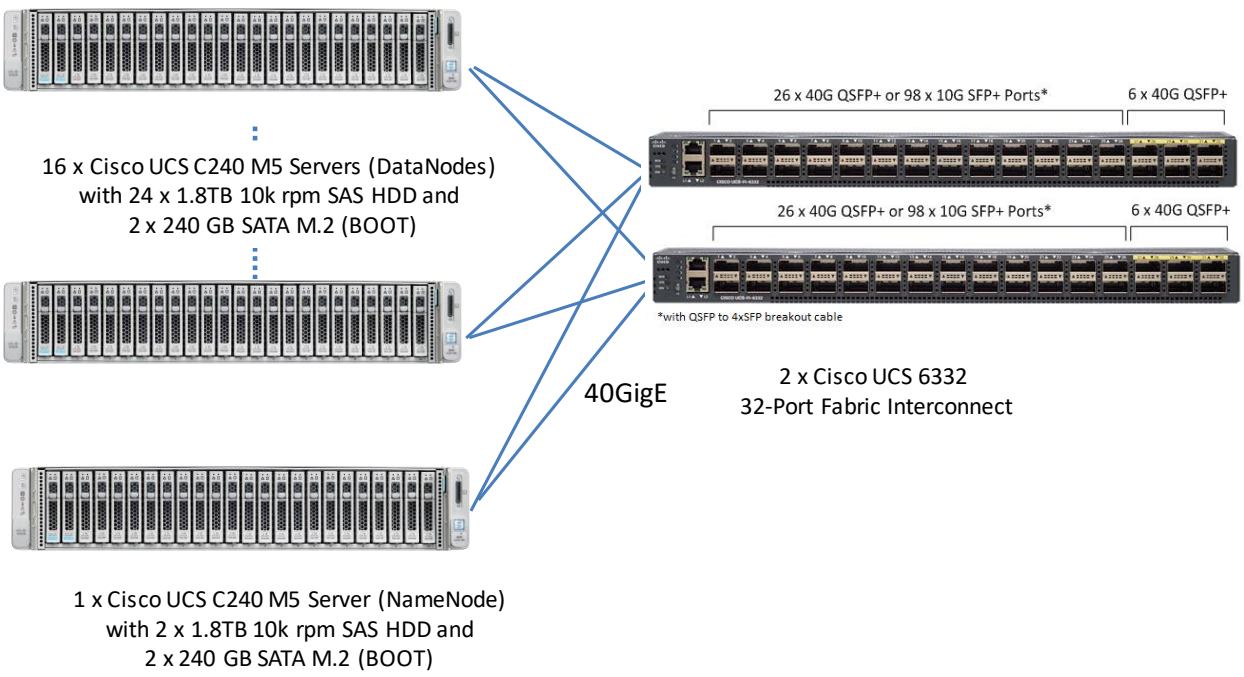
The supporting files contain the parameters and options used to configure the components involved in this benchmark.

## 1.3 Configuration Diagrams

*7.4.4 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.*

**Measured Configuration**




The measured configuration consisted of:

- Total Nodes: 17
- Total Processors/Cores/Threads: 34/476/952
- Total Memory: 3.19TB
- Total Number of Storage Drives/Devices: 420
- Total Storage Capacity: 702.96TB

Server nodes details:

- 17 x UCS C240M5 Servers, each with:
  - Processors/Cores/Threads: 2/28 /56
  - Processor Model: 2 x Intel® Xeon® Gold Processor 6132, 2.60 GHz, 19.25 MB L3
  - Memory: 192GB
  - Controller: 1 x Cisco 12G SAS Modular Raid Controller
  - Drives:
    - 24 x 1.8TB 10K RPM SAS HDD (data nodes)
    - 2 x 1.8TB 10K RPM SAS HDD (name node)
    - 2 x 240 GB SATA (all nodes, boot disk)
  - Network: 1 x Cisco VIC 1387 Dual Port 40Gb QSFP CNA MLOM

Network connectivity detail:

- 2 x Cisco UCS 6332 32-Port Fabric Interconnect

The distribution of software components over server nodes is detailed in section 1.5.

**Priced Configuration**

There are no differences between the priced and measured configurations.

## 1.4 Dataset Distribution

*The distribution of dataset across all media must be explicitly described.*

Table 1.4 describes the distribution of the dataset across all media in the system.

**Table 1.4: Dataset Distribution**

| Server Node | Controller | Disk Drive | Description of Content |
|---|---|---|---|
| 1 | Cisco 12G Modular RAID controller with 4GB cache | 1-2 (HDD) | Data, Temp |
| 1 | Embedded RAID PCH SATA | 0 (2 SSD, RAID-1) | Operating system, Root, Swap, Hadoop Master |
| 2-17 | Cisco 12G Modular RAID controller with 4GB cache | 1-24 (HDD) | Data, Temp |
| 2-17 | Embedded RAID PCH SATA | 0 (2 SSD, RAID-1) | Operating system, Root, Swap, Hadoop Master |

## 1.5 Software Components Distribution

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

Table 1.5 describes the distribution of the software components across the system.

**Table 1.5: Software Component Distribution**

| | Map/Reduce | | HDFS | | ZooKeeper |
|---|---|---|---|---|---|
| Node | Resource Manager | Node Manager | NameNode | DataNode | QuorumPeer |
| 1 | X | | X | | X |
| 2-3 | | X | | X | X |
| 4-17 | | X | | X | |

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

Cloudera Enterprise Edition V5.13.0 (fully HDFS compatible at the API level).

*Map/Reduce implementation and corresponding version must be disclosed.*

Cloudera Enterprise Edition V5.13.0 (compatible equivalent to Hadoop 2.6.0).

# Clause 2: Workload Related Items

## 2.1 Hardware & Software Tunable

*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 =              3232
                        Total Size =      100000000000
                        Scale-Factor =              10

TPCx-HS Performance Metric (HSph@SF):          11.1395

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


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

Test Run 2 Details      Total Time =              2971
                        Total Size =      100000000000
                        Scale-Factor =              10

TPCx-HS Performance Metric (HSph@SF):          12.1182

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

## 2.3 Benchmark Kit Identification

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

Kit Version 2.0.2

| File | MD5 |
|---|---|
| BigData_cluster_validate_suite.sh | 58c13ddb98a2d1228f2df10f4a087a71 |
| HSDataCheck.sh | a621d8a94db9c6015f3463557664ef21 |
| TPCx-HS-master.jar | 492cbc51a1a60c28b43d96c79d08683d |
| TPCx-HS-master.sh | b110d48eab67e21fd88822beb376b1d3 |

## 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.

**Table 3.1: Storage Device Capacity**

| Qty | Capacity (GB) | Total (GB) |
|---|---|---|
| 386 | 1,800 | 694,800 |
| 34 | 240 | 8,160 |
| **Total Storage (TB)** | | **702.96** |

Scale Factor = 10

**Data Storage Ratio =** (Storage / SF) **= 70.30**

## 3.2 Memory Ratio

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

Total Configured Memory = 3.19TB

**Scale Factor to Memory Ratio =** (SF / Memory) **= 3.14**

# Clause 4: Scale Factors and Metrics

## 4.1 HSGen Time

*The HSGen time must be disclosed for Run1 and Run2.*

| | Run1 | Run2 |
|---|---|---|
| HSGen | 693.141 | 682.876 |

## 4.2 HSSort Time

*The HSSort time must be disclosed for Run1 and Run2.*

| | Run1 | Run2 |
|---|---|---|
| HSSort | 2,199.504 | 1,986.172 |

## 4.3 HSValidate Time

*The HSValidate time must be disclosed for Run1 and Run2.*

| | Run1 | Run2 |
|---|---|---|
| HSValidate | 333.302 | 296.144 |

## 4.4 HSDataCheck Times

*Both HSDataCheck times must be disclosed for Run1 and Run2.*

| | Run1 | Run2 |
|---|---|---|
| HSDataCheck (pre-Sort) | 3.000 | 2.000 |
| HSDataCheck (post-Sort) | 4.000 | 3.000 |

## 4.5 Performance & Price-Performance

*The performance metric (HSph@SF) must be disclosed for Run1 and Run2. Price-performance metric ($/HSph@SF) must be disclosed for the performance run.*

| | Run1 | Run2 |
|---|---|---|
| HSph@10TB | 11.13 | 12.11 |

| | |
|---|---|
| $/HSph@10TB | 58,948.43 USD |

# Auditors' Information and Attestation Letter

*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 attestation letter is included in the next two pages.

InfoSizing
The Right Metric For Sizing IT

TPC Transaction Processing Performance Council
Certified Auditors

Raghunath Nambiar
Cisco Systems Inc.
3800 Zanker Road
San Jose, CA 95134

January 30, 2018

I verified the TPC Express Benchmark™ HS v2.0.2 performance of the following configuration:

| | |
|---|---|
| Platform: | Cisco UCS Integrated Infrastructure for Big Data and Analytics (with 17 UCS C240M5 Servers) |
| Operating System: | Red Hat Enterprise Linux Server 7.4 |
| Apache Hadoop Compatible Software: | Cloudera Enterprise Edition V5.13.0 |

The results were:

| | |
|---|---|
| **Performance Metric** | **11.13 HSph@10TB** |
| Run Elapsed Time | 3,232.00 Seconds |

| **Cluster** | **17 UCS C240M5 Servers, each node with:** | | |
|---|---|---|---|
| CPUs | 2 x Intel Xeon Gold Processor 6132 (2.60 GHz, 14-core, 19.25 MB L3) | | |
| Memory | 192 GB | | |
| Storage | ***Qty*** | ***Size*** | ***Type*** |
| | 24 | 1.8TB | 10K RPM SAS HDD (Data Nodes) |
| | 2 | 1.8TB | 10K RPM SAS HDD (Name Node) |
| | 2 | 240GB | SATA (All Nodes, boot disk) |

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 v2.0.2
- 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 10TB
- The generated dataset and the sorted dataset were replicated 3-ways

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:

None.

Respectfully Yours,

Doug Johnson, Certified TPC Auditor

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

# Supporting File Index

The following index outlines the information included in the supporting files archive.

| Clause | Description | Archive File Pathname |
|---|---|---|
| Clause 1 | Parameters and options used to configure the system | SupportingFilesArchive\Clause1 |
| Clause 2 | Configuration scripts & Run report | SupportingFilesArchive\Clause2 |
| Clause 3 | System configuration details | SupportingFilesArchive\Clause3 |

# Third Party Price Quotes

## CDW