Cisco Systems, Inc.

TPC Express Benchmark™ HS (TPCx-HS)

Full Disclosure Report

for

Cisco UCS Integrated Infrastructure for Big Data

(with 32 Cisco UCS C240M4 Servers)

using

Cloudera Distribution for Apache Hadoop (CDH) 5.3.2

and

Red Hat Enterprise Linux Server 6.5

First Edition

October 23, 2015

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.

Copyright © 2015 Cisco Systems, Inc.

	Cisco UCS Integrated Infrastructure for Big Data	TPCx-HS Rev. 1.3.0 TPC-Pricing Rev. 2.1.0
		Report Date: October 23, 2015

Total System Cost	TPCx-HS Performance Metric	Price/Performance
861,868 USD	**23.42** HSph@30TB	**36,800.52 USD** $/HSph@30TB

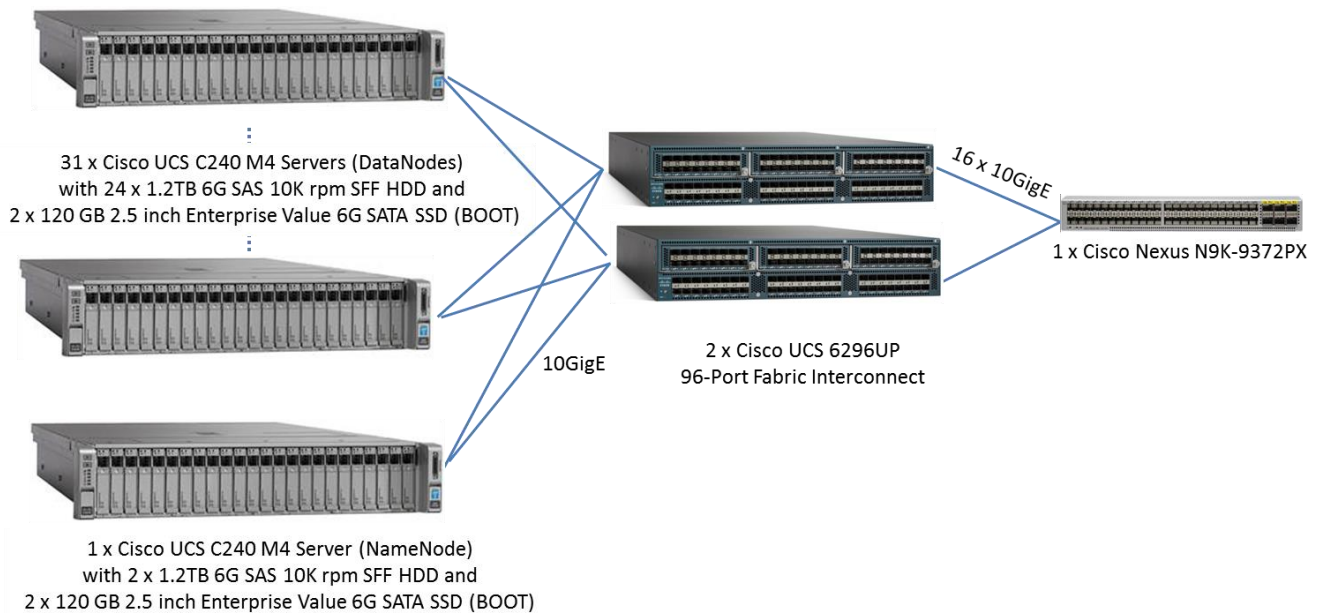| Scale Factor	Apache Hadoop Compatible Software	Operating System	Other Software	Availability Date
30TB	Cloudera Distribution for Apache Hadoop (CDH) 5.3.2	Red Hat Enterprise Linux Server 6.5	None	October 23, 2015

System Configuration

31 x Cisco UCS C240 M4 Servers (DataNodes) with 24 x 1.2TB 6G SAS 10K rpm SFF HDD and 2 x 120 GB 2.5 inch Enterprise Value 6G SATA SSD (BOOT)

16 x 10GigE

1 x Cisco Nexus N9K-9372PX

10GigE

2 x Cisco UCS 6296UP 96-Port Fabric Interconnect

1 x Cisco UCS C240 M4 Server (NameNode) with 2 x 1.2TB 6G SAS 10K rpm SFF HDD and 2 x 120 GB 2.5 inch Enterprise Value 6G SATA SSD (BOOT)

Physical Storage/Scale Factor: 30.1	Scale Factor/Physical Memory: 3.75

Servers: 32 x Cisco UCS C240M4 Server
Total Processors/Cores/Threads   64/768/1536

Server Configuration: Per node:
Processors   2 x Intel® Xeon® Processor E5-2680 v3, 2.50 GHz, 30MB L3
Memory   256GB
Storage Controller   1 x Cisco 12G SAS Modular Raid Controller, 2GB FBWC Cache
Storage Device   24 x 1.2TB 6G SAS 10Krpm SFF HDD (for 31 DataNodes)
2 x 1.2TB 6G SAS 10Krpm SFF HDD (for 1 NameNode)
2 x 120GB 2.5in. Enterprise Value 6G SATA SSD (boot disks, all nodes)
Network   1 x Cisco UCS VIC1227 VIC MLOM - Dual Port 10Gb SFP+

Connectivity: 2 x Cisco UCS 6296UP 96-Port Fabric Interconnect
1 x Cisco Nexus N9K-9372PX

Cisco UCS Integrated Infrastructure for Big Data

Description	Part Number	Source	Unit Price	Qty	Extended Price	3 Year Maint.
Hardware Components						
Cisco UCS CPA v3 Performance Optimized Bundle	UCS-SL-CPA3-P	1	$0.00	1	$0.00	
(Not sold Standalone)UCS C240 M4SX w/E52680v3,256GB mem	UCS-SA-C240M4SX-P	1	$48,005.00	16	$768,080.00	
2.50 GHz E5-2680 v3/120W 12C/30MB Cache/DDR4 2133MHz	UCS-CPU-E52680D	1	$0.00	32	$0.00	
16GB DDR4-2133-MHz RDIMM/PC4-17000/dual rank/x4/1.2v	UCS-MR-1X162RU-A	1	$0.00	256	$0.00	
120 GB 2.5 inch Enterprise Value 6G SATA SSD (BOOT)	UCS-SD120G0KSB-EV	1	$0.00	32	$0.00	
1.2 TB 6G SAS 10K rpm SFF HDD	UCS-HD12T10KS2-E	1	$0.00	384	$0.00	
Cisco 12Gbps SAS 2GB FBWC Cache module (Raid 0/1/5/6)	UCSC-MRAID12G-2GB	1	$0.00	16	$0.00	
Cisco 12G SAS Modular Raid Controller	UCSC-MRAID12G	1	$0.00	16	$0.00	
Cisco UCS VIC1227 VIC MLOM - Dual Port 10Gb SFP+	UCSC-MLOM-CSC-02	1	$0.00	16	$0.00	
Right PCI Riser Bd (Riser 1) 2onbd SATA bootdrvs+ 2PCI slts	UCSC-PCI-1C-240M4	1	$0.00	16	$0.00	
1200W / 800W V2 AC Power Supply for 2U C-Series Servers	UCSC-PSU2V2-1200W	1	$0.00	32	$0.00	
Power Cord 125VAC 13A NEMA 5-15 Plug North America	CAB-9K12A-NA	1	$0.00	32	$0.00	
Right PCI Riser Bd (Riser 1) 2onbd SATA bootdrvs+ 2PCI slts	UCSC-PCI-1C-240M4	1	$0.00	16	$0.00	
Heat sink for UCS C240 M4 rack servers	UCSC-HS-C240M4	1	$0.00	32	$0.00	
Supercap cable 250mm	UCSC-SCCBL240	1	$0.00	16	$0.00	
Ball Bearing Rail Kit for C220 M4 and C240 M4 rack servers	UCSC-RAILB-M4	1	$0.00	16	$0.00	
(Not sold standalone) UCS 6296UP 2RU FI w/18p LIC,16xCables	UCS-SA-BD-FI96	1	$22,517.00	2	$45,034.00	
UCS 6296UP Power Supply/100-240VAC	UCS-PSU-6296UP-AC	1	$0.00	4	$0.00	
10GBASE-CU SFP+ Cable 3 Meter	SFP-H10GB-CU3M	1	$0.00	32	$0.00	
UCS Manager v2.2	N10-MGT012	1	$0.00	2	$0.00	
UCS 6296UP Chassis Accessory Kit	UCS-ACC-6296UP	1	$0.00	2	$0.00	
UCS 6200 Series Expansion Module Blank	UCS-BLKE-6200	1	$0.00	6	$0.00	
UCS 6296UP Fan Module	UCS-FAN-6296UP	1	$0.00	8	$0.00	
Power Cord 125VAC 13A NEMA 5-15 Plug North America	CAB-9K12A-NA	1	$0.00	4	$0.00	
(Not sold Standalone)UCS C240 M4SX w/E52680v3,256GB mem	UCS-SA-C240M4SX-P	1	$48,005.00	16	$768,080.00	
2.50 GHz E5-2680 v3/120W 12C/30MB Cache/DDR4 2133MHz	UCS-CPU-E52680D	1	$0.00	32	$0.00	
16GB DDR4-2133-MHz RDIMM/PC4-17000/dual rank/x4/1.2v	UCS-MR-1X162RU-A	1	$0.00	256	$0.00	
120 GB 2.5 inch Enterprise Value 6G SATA SSD (BOOT)	UCS-SD120G0KSB-EV	1	$0.00	32	$0.00	
1.2 TB 6G SAS 10K rpm SFF HDD	UCS-HD12T10KS2-E	1	$0.00	384	$0.00	
Cisco 12Gbps SAS 2GB FBWC Cache module (Raid 0/1/5/6)	UCSC-MRAID12G-2GB	1	$0.00	16	$0.00	
Cisco 12G SAS Modular Raid Controller	UCSC-MRAID12G	1	$0.00	16	$0.00	
Cisco UCS VIC1227 VIC MLOM - Dual Port 10Gb SFP+	UCSC-MLOM-CSC-02	1	$0.00	16	$0.00	
Right PCI Riser Bd (Riser 1) 2onbd SATA bootdrvs+ 2PCI slts	UCSC-PCI-1C-240M4	1	$0.00	16	$0.00	
1200W / 800W V2 AC Power Supply for 2U C-Series Servers	UCSC-PSU2V2-1200W	1	$0.00	32	$0.00	
Power Cord 125VAC 13A NEMA 5-15 Plug North America	CAB-9K12A-NA	1	$0.00	32	$0.00	
Right PCI Riser Bd (Riser 1) 2onbd SATA bootdrvs+ 2PCI slts	UCSC-PCI-1C-240M4	1	$0.00	16	$0.00	
Heat sink for UCS C240 M4 rack servers	UCSC-HS-C240M4	1	$0.00	32	$0.00	
Supercap cable 250mm	UCSC-SCCBL240	1	$0.00	16	$0.00	
Ball Bearing Rail Kit for C220 M4 and C240 M4 rack servers	UCSC-RAILB-M4	1	$0.00	16	$0.00	
Cisco Smart Net 24X7X4 3Y UCS C240 M4S BD SP Server	CON-OSP-C240V4SP	1	$1,284.99	32		$41,119.68
2rd Gen FI License to connect C-direct only	UCS-L-6200-10G-C	1	925.00	60	$55,500.00	
10GBASE-CU SFP+ Cable 3 Meter	SFP-H10GB-CU3M	1	$100.00	64	$6,400.00	
Cisco R42610 standard rack w/side panels	RACK-UCS2	1	$3,429.00	2	$6,858.00	
Cisco Smart Net 24X7X4 3Y UCS 6296UP 2RU Fabrc Int/2 PSU/4 Fans	CON-SNTP-FI6296UP	1	$5,781.00	2		$11,562.00
Cisco Nexus 9372PX Switch	N9K-C9372PX	1	$22,500.00	1	$22,500.00	
Cisco SNTP Cisco Nexus 9372PX Switch, 36 24x7	CON-SNTP-9372PX		$5,256.00	1		$5,256.00
Hardware subtotal					**$1,672,452.00**	**$57,937.68**

Cisco UCS Integrated Infrastructure for Big Data

Description	Part Number	Source	Unit Price	Qty	Extended Price	3 Year Maint.
Software Components						
Red Hat Enterprise Linux Server, 3Y 24x7	CON-ISV1-RH2SUG3A	1	$2,397.00	32	$76,704.00	Inc.
Cloudera Enterprise Edition	UCS-BD-CBI85-3YR	1	$11,354.00	32	$363,328.00	Inc.
Software subtotal					**$440,032.00**	
Total					**$2,112,484.00**	**$57,937.68**
Acer V206HQLAbd - LED monitor - 20" (Inc. 2 spares)	UM.1V6AA.A02	2	$94.99	3	$284.97	
Logitech USB Corded Keyboard/Mouse Combo MK120 (Inc. 2 spares)	920-002565	2	$17.99	3	$53.97	
Large Purchase Discount [1] (61% for products and 35% for service)		1			($1,288,615.24)	($20,278.19)

Pricing:1 = Cisco, 2 = CDW.com

[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, Inc.

Three-Year Cost of Ownership	**$861,868**
HSph@30TB	**23.42**
$/HSph@30TB	**$36,800.52**

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.

Measurement Results for Performance Run

Scale Factor   30TB

Run Start Time   2015-10-14 02:56:52
Run End Time   2015-10-14 04:13:39
Run Elapsed Time   4,610.000

HSGen Start Time   2015-10-14 02:56:52
HSGen End Time   2015-10-14 03:15:40
HSGen Elapsed Time   1,129.107

HSSort Start Time   2015-10-14 03:15:42
HSSort End Time   2015-10-14 04:04:58
HSSort Elapsed Time   2,956.530

HSValidate Start Time   2015-10-14 04:05:01
HSValidate End Time   2015-10-14 04:13:39
HSValidate Elapsed Time   518.731

Measurement Results for Repeatability Run

Scale Factor   30TB

Run Start Time   2015-10-14 01:23:21
Run End Time   2015-10-14 02:40:05
Run Elapsed Time   4,607.000

HSGen Start Time   2015-10-14 01:23:21
HSGen End Time   2015-10-14 01:42:17
HSGen Elapsed Time   1,136.883

HSSort Start Time   2015-10-14 01:42:19
HSSort End Time   2015-10-14 02:31:26
HSSort Elapsed Time   2,947.100

HSValidate Start Time   2015-10-14 02:31:29
HSValidate End Time   2015-10-14 02:40:05
HSValidate Elapsed Time   516.998

Run Report for Performance Run

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

Test Run 2 details: Total Time = 4610
                    Total Size = 300000000000
                    Scale-Factor = 30.0000

TPCx-HS Performance Metric (HSph@SF): 23.4283


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

Run Report for Repeatability Run

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

Test Run 1 details: Total Time = 4607
                    Total Size = 300000000000
                    Scale-Factor = 30.0000

TPCx-HS Performance Metric (HSph@SF): 23.4429


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

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 1.3.0.

The test was conducted at a Scale Factor of 30TB with 32 Cisco UCS C240M4 Servers running Cloudera Distribution for Apache Hadoop (CDH) 5.3.2 on Red Hat Enterprise Linux Server 6.5.

Measured Configuration

Company Name	Cluster Node	Virtualization	Operating System
Cisco Systems, Inc.	Cisco UCS C240M4 Server	n/a	Red Hat Enterprise Linux Server 6.5

TPC Express Benchmark© HS Metrics

Total System Cost	HSph@30TB	Price/Performance	Availability Date
861,868 USD	23.42	36,800.52 USD	October 23, 2015

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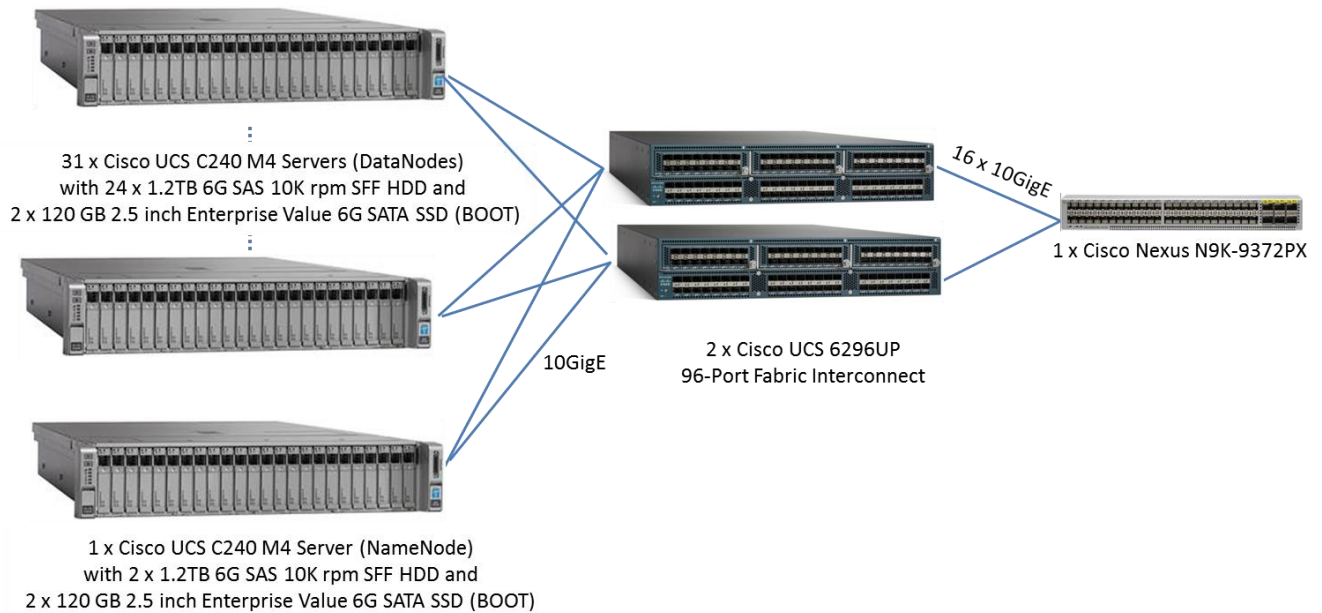
The supporting files contain the parameters and options used to configure the components involved in this benchmark.

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

31 x Cisco UCS C240 M4 Servers (DataNodes)
with 24 x 1.2TB 6G SAS 10K rpm SFF HDD and
2 x 120 GB 2.5 inch Enterprise Value 6G SATA SSD (BOOT)

16 x 10GigE

1 x Cisco Nexus N9K-9372PX

10GigE

2 x Cisco UCS 6296UP
96-Port Fabric Interconnect

1 x Cisco UCS C240 M4 Server (NameNode)
with 2 x 1.2TB 6G SAS 10K rpm SFF HDD and
2 x 120 GB 2.5 inch Enterprise Value 6G SATA SSD (BOOT)

The measured configuration consisted of:
- Total Nodes: 32
- Total Processors/Cores/Threads: 64/768/1536
- Total Memory: 8TB
- Total Number of Storage Drives/Devices: 810
- Total Storage Capacity: 902.88TB

Server nodes details:
- 32 x Cisco UCS C240M4 Servers, each with:
    - Processors/Cores/Threads: 2/24/48
    - Processor Model: 2 x Intel® Xeon® Processor E5-2680 v3, 2.50 GHz, 30MB L3
    - Memory: 256GB
    - Controller: 1 x Cisco 12G SAS Modular Raid Controller, 2GB FBWC Cache
    - Drives:
        - 24 x 1.2TB 6G SAS 10Krpm SFF HDD (for 31 DataNodes)
        - 2 x 1.2TB 6G SAS 10Krpm SFF HDD (for 1 NameNode)
        - 2 x 120GB 2.5in. Enterprise Value 6G SATA SSD (boot disks, all nodes)
    - Network: 2 x Cisco UCS VIC1227 VIC MLOM - Dual Port 10Gb SFP+

Network connectivity detail:
- 2 x Cisco UCS 6296UP 96-Port Fabric Interconnect
- 1 x Cisco Nexus N9K-9372PX

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	MegaRAID 3108	1-2 (HDD)	Data, Temp
2-32	MegaRAID 3108	1-24 (HDD)	Data, Temp
1-32	Intel Chipset Embedded SATA RAID	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: Dataset Distribution

	Map/Reduce		HDFS		ZooKeeper
Node	JobTracker	TaskTracker	NameNode	DataNode	QuorumPeer
1	x		x		x
2		x	x	x	x
3		x		x	x
4-32		x		x	

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

Cloudera Distribution for Apache Hadoop (CDH) 5.3.2 (fully HDFS compatible at the API level).

Map/Reduce implementation and corresponding version must be disclosed.

Cloudera Distribution for Apache Hadoop (CDH) 5.3.2 (compatible equivalent to Hadoop 2.7.0).

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.

- **Run1 Performance Summary**

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

Test Run 1 details: Total Time = 4607
                    Total Size = 300000000000
                    Scale-Factor = 30.0000

TPCx-HS Performance Metric (HSph@SF): 23.4429


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

- **Run2 Performance Summary**

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

Test Run 2 details: Total Time = 4610
                    Total Size = 300000000000
                    Scale-Factor = 30.0000

TPCx-HS Performance Metric (HSph@SF): 23.4283


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

2.3 Benchmark Kit Identification

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

The version number of the TPCx-HS kit used is 1.3.0. The md5sum for the TPCx-HS kit files used during the benchmark are:

TPCx-HS-master.jar 4ceaefc51c698c0733b57244b7760808

BigData_cluster_validate_suite.sh 58c13ddb98a2d1228f2df10f4a087a71

TPCx-HS-master.sh 70ba6b440de47b4e4a902bf4983ee4c1

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)
746	1,200	895,200
64	120	7,680
Total Storage (TB)		902.88

Scale Factor = 30TB

Data Storage Ratio = (Storage / SF) = **30.1**

3.2 Memory Ratio

The Scale Factor to memory ratio must be disclosed.

Total Configured Memory = 8TB

Scale Factor to Memory Ratio = (SF / Memory) = **3.75**

Clause 4: Scale Factors and Metrics

4.1 HSGen Time

The HSGen time must be disclosed for Run1 and Run2.

	Run1	Run2
HSGen	1,136.883	1,129.107

4.2 HSSort Time

The HSSort time must be disclosed for Run1 and Run2.

	Run1	Run2
HSSort	2,947.100	2,956.530

4.3 HSValidate Time

The HSValidate time must be disclosed for Run1 and Run2.

	Run1	Run2
HSValidate	516.998	518.731

4.4 HSDataCheck Times

Both HSDataCheck times must be disclosed for Run1 and Run2.

	Run1	Run2
HSDataCheck (pre-Sort)	2.000	2.000
HSDataCheck (post-Sort)	3.000	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@30TB	23.44	23.42

$/HSph@30TB	36,800.52 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 for InfoSizing, Inc.

     www.sizing.com
     20 Kreg Lane
     Manitou Springs, CO 80829
     719-473-7555.

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

October 20, 2015

I verified the TPC Express Benchmark^{TM} HS v1.3.0 performance of the following configuration:

Platform:	Cisco UCS Integrated Infrastructure for Big Data (Cisco UCS CPA v3) (with 32 Cisco UCS C240M4 Servers)
Operating System:	Red Hat Enterprise Linux Server 6.5
Apache Hadoop Compatible Software:	Cloudera Distribution for Apache Hadoop (CDH) 5.3.2

The results were:

Performance Metric   23.42 HSph@30TB
Run Elapsed Time   4,610.00 Seconds

Cluster   32 Cisco UCS C240M4 Servers

CPUs	2 x Intel Xeon Processor E5-2680 v3 (2.50 GHz, 12-core, 30 MB L3)
Memory	256 GB
Storage	

Qty	Size	Type
24	1.2 TB	10K rpm SFF HDD (31 DataNodes)
2	1.2 TB	10K rpm SFF HDD (1 NameNode)
2	120GB	SATA SSD (boot disks, all 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 v1.3.0
- 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 30TB
- 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,

Doug Johnson, Auditor

François Raab, President

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