TPC EXPRESS BENCHMARK™ HS

(TPCx-HS)

Standard Specification

Version 1.1.

April 23, 2014

Transaction Processing Performance Council (TPC)

www.tpc.org

info@tpc.org

© 2014 Transaction Processing Performance Council

All Rights Reserved
Legal Notice

The TPC reserves all right, title, and interest to this document and associated source code as provided under U.S. and international laws, including without limitation all patent and trademark rights therein.

Permission to copy without fee all or part of this document is granted provided that the TPC copyright notice, the title of the publication, and its date appear, and notice is given that copying is by permission of the Transaction Processing Performance Council. To copy otherwise requires specific permission.

No Warranty

TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, THE INFORMATION CONTAINED HEREIN IS PROVIDED “AS IS” AND WITH ALL FAULTS, AND THE AUTHORS AND DEVELOPERS OF THE WORK HEREBY DISCLAIM ALL OTHER WARRANTIES AND CONDITIONS, EITHER EXPRESS, IMPLIED OR STATUTORY, INCLUDING, BUT NOT LIMITED TO, ANY (IF ANY) IMPLIED WARRANTIES, DUTIES OR CONDITIONS OF MERCHANTABILITY, OF FITNESS FOR A PARTICULAR PURPOSE, OF ACCURACY OR COMPLETENESS OF RESPONSES, OF RESULTS, OF WORKMANLIKE EFFORT, OF LACK OF VIRUSES, AND OF LACK OF NEGLIGENCE. ALSO, THERE IS NO WARRANTY OR CONDITION OF TITLE, QUIET ENJOYMENT, QUIET POSSESSION, CORRESPONDENCE TO DESCRIPTION OR NON-INFRINGEMENT WITH REGARD TO THE WORK.

IN NO EVENT WILL ANY AUTHOR OR DEVELOPER OF THE WORK BE LIABLE TO ANY OTHER PARTY FOR ANY DAMAGES, INCLUDING BUT NOT LIMITED TO THE COST OF PROCURING SUBSTITUTE GOODS OR SERVICES, LOST PROFITS, LOSS OF USE, LOSS OF DATA, OR ANY INCIDENTAL, CONSEQUENTIAL, DIRECT, INDIRECT, OR SPECIAL DAMAGES WHETHER UNDER CONTRACT, TORT, WARRANTY, OR OTHERWISE, ARISING IN ANY WAY OUT OF THIS OR ANY OTHER AGREEMENT RELATING TO THE WORK, WHETHER OR NOT SUCH AUTHOR OR DEVELOPER HAD ADVANCE NOTICE OF THE POSSIBILITY OF SUCH DAMAGES.

Trademarks

TPC Benchmark, TPC Expres, TPCx-HS, HSGen, HSDataCheck, HSSort, HSVldate and HSp are trademarks of the Transaction Processing Performance Council.
Acknowledgments

Developing a TPC benchmark for a new environment requires a huge effort to conceptualize research, specify, review, prototype, and verify the benchmark. The TPC acknowledges the work and contributions of the TPCx-HS subcommittee member companies in developing the TPCx-HS specification.

The TPCx-HS subcommittee would like to acknowledge the contributions made by the many members during the development of the benchmark specification. It has taken the dedicated efforts of people across many companies, often in addition to their regular duties. The list of contributors to this version includes Andrew Bond, Andrew Masland, Avik Dey, Brian Caufield, Chaitanya Baru, Da Qi Ren, Dileep Kumar, Jamie Reding, John Fowler, John Poelman, Karthik Kulkarni, Matthew Emmerton, Meikel Poess, Mike Brey, Mike Crocker, Paul Cao, Raghunath Nambiar, Reza Taheri, Simon Harris, Tariq Magdon-Ismail and Wayne Smith.

Document Revision History

<table>
<thead>
<tr>
<th>Date</th>
<th>Version</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>To be tracked after GC approval of the draft 1.0</td>
</tr>
</tbody>
</table>
TPC Membership
(as of April 2014)

Full Members

<table>
<thead>
<tr>
<th>Full Members</th>
<th>Full Members</th>
<th>Full Members</th>
<th>Full Members</th>
<th>Full Members</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cisco</td>
<td>cloudera</td>
<td>DELL</td>
<td>FUJITSU</td>
<td>HP</td>
</tr>
<tr>
<td>HITACHI</td>
<td>Huawei</td>
<td>IBM</td>
<td>Intel</td>
<td>MapR</td>
</tr>
<tr>
<td>Microsoft</td>
<td>NEC</td>
<td>Oracle</td>
<td>Red Hat</td>
<td>Sybase</td>
</tr>
<tr>
<td>Teradata</td>
<td>Unisys</td>
<td>VMware</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Associate Members

<table>
<thead>
<tr>
<th>TTA</th>
<th>IDEAS</th>
<th>SDSC</th>
<th>ITOM International Co.</th>
<th>University of California</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Table of Contents

**CLAUSE 1 : INTRODUCTION**

- 1.2 Preamble 7
- 1.2 TPCx-HS Kit and Licensing 7
- 1.3 General Implementation Guidelines 7
- 1.4 General Measurement Guidelines 8

**CLAUSE 2: WORKLOAD AND EXECUTION**

- 2.1 Benchmark Kit 9
- 2.2 Benchmark Workload 10
- 2.3 Benchmark Execution 10
- 2.4 Configuration and Tuning 12

**CLAUSE 3: SYSTEM UNDER TEST AND DRIVER**

- 3.1 System Under Test 13

**CLAUSE 4 SCALE FACTORS AND METRICS**

- 4.1 Scale Factor 15
- 4.2 Metrics 15
- 4.3 Performance Metric (HSPH@SF) 15
- 4.4 Price Performance Metric ($/HSPH@SF) 16
- 4.5 System Availability Date 16
- 4.6 Metric Comparison 16
- 4.7 Required Reporting Components 16

**CLAUSE 5: PRICING**

- 5.1 Priced System 17
- 5.2 User Interface Devices and Communications 17
- 5.3 Additional Operational Components 17
- 5.4 Software 18
- 5.5 Allowable Substitutions 18

**CLAUSE 6: ENERGY**

- 6.1 General Measurement Guidelines 19
- 6.3 TPCx-HS Energy Metric 21
- 6.4 Executive Summary and Full Disclosure Report 21

**CLAUSE 7: FULL DISCLOSURE**

- 7.1 Reporting Requirements 22
- 7.2 Format Guidelines 22
- 7.3 Full Disclosure Report 22
- 7.4 General Items 22
- 7.5 Clause 2 - Workload Related Items 26
- 7.6 Clause 3 - SUT Related Items 26
- 7.7 Clause 5 - Performance Metrics and Execution Rules Related Items 26
- 7.8 Clause 8 - Audit Related Items 26
- 7.9 Executive Summary 27
  - 7.9.1 Page Layout 27
- 7.9 Implementation Overview 27
- 7.10 Pricing Spreadsheet 29
7.11 Numerical Quantities Summary
7.12 TPCx-HS Run Report
7.13 Availability of the Full Disclosure Report
7.14 Revisions to the Full Disclosure Report

CLAUSE 8: AUDIT

8.1 General Rules
   8.1.2 TPCx-HS results can be used as the basis for new TPCx-HS results if and only if:

8.2 Audit Check List
   8.2.1 Clause 2 Related Items
   8.2.2 Clause 3 Related Items
   8.2.3 Clause 4 Related Items
   8.2.4 Clause 5 Related Items
   8.2.5 Clause 6 Related Items
CLAUSE 1: INTRODUCTION

1.1 Preamble

Big Data technologies like Hadoop have become an important part of the enterprise IT ecosystem. 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 runtime, Hadoop Filesystem API compatible systems and MapReduce layers. This workload can be used to assess a broad range of system topologies and implementation of Hadoop clusters. The TPCx-HS can be used to assess a broad range of system topologies and implementation methodologies in a technically rigorous and directly comparable, in a vendor-neutral manner.

1.2 TPCx-HS Kit and Licensing

The TPCx-HS kit is available from the TPC (See www.tpc.org/tpcx-hs for more information). User 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 (this document), TPCx-HS Users Guide documentation, Shell scripts to set up the benchmark environment, Java code to execute the benchmark load. See Clause 2.

1.3 General Implementation Guidelines

The purpose of TPC benchmarks are 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 Filesystem API compatible systems);
- Would plausibly be implemented by a significant number of users in the market segment the benchmark models or represents.

The system under test (SUT)’s fault-recovery attributes must transparently and automatically allow recovery from partial failures. Partial failure is defined as a failure that is limited to a given component (e.g., machines or hardware nodes) or phase of processing (e.g., map or reduce tasks). In case of a partial failure, the implementation will only allow restarting of the impacted portion of the work in progress. The work that has already successfully completed in earlier phases of processing in other non-failed components of the SUT can not be restarted. For example, if a node fails, only the tasks that were scheduled and ran on that node can be restarted.

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 following characteristics shall be used as a guide to judge whether a particular implementation is a benchmark special. It is not required that each point below be met, but that the cumulative weight of the evidence be considered to identify an unacceptable implementation. Absolute certainty or certainty beyond a reasonable doubt is not required to make a judgment on this complex issue. The question that must be answered is: "Based on the available evidence, does the clear preponderance (the greater share or weight) of evidence indicate that this implementation is a benchmark special?"

The following characteristics shall be used to judge whether a particular implementation is a benchmark special:

- Is the implementation generally available, externally documented, and supported?
- Does the implementation have significant restrictions on its use or applicability that limits its use beyond the TPCx-HS benchmark?
- Is the implementation or part of the implementation poorly integrated into the larger product?
- Does the implementation take special advantage of the limited nature of the TPCx-HS benchmark in a manner that would not be generally applicable to the environment the benchmark represents?
- Is the use of the implementation discouraged by the vendor? (This includes failing to promote the implementation in a manner similar to other products and technologies.)
- Does the implementation require uncommon sophistication on the part of the end-user, programmer, or system administrator?
- Is the implementation (including beta) being purchased or used for applications in the market area the benchmark represents? How many sites implemented it? How many end-users benefit from it? If the implementation is not currently being purchased or used, is there any evidence to indicate that it will be purchased or used by a significant number of end-user sites?

The rules for pricing are included in the TPC Pricing Specification located at www.tpc.org.

The rules for energy measurement are included in the TPC Energy Specification located at www.tpc.org.

1.4 General Measurement Guidelines

TPC benchmark results are expected to be accurate representations of system performance. Therefore, there are certain guidelines that are expected to be followed when measuring those results. The approach or methodology to be used in the measurements are either explicitly described in the specification or left to the discretion of the test sponsor. When not described in the specification, the methodologies and approaches used must meet the following requirements:

- The approach is an accepted engineering practice or standard;
- The approach does not enhance the result;
- Equipment used in measuring the results is calibrated according to established quality standards;
- Fidelity and candor is maintained in reporting any anomalies in the results, even if not specified in the TPC benchmark requirements.
CLAUSE 2: WORKLOAD AND EXECUTION

This clause defines workload and execution.

2.1 Benchmark Kit

2.1.1 Kit Contents

- The TPCx-HS kit contains the following:
  - TPCx-HS Specification document
  - TPCx-HS Users Guide documentation
  - Shell scripts to set up the benchmark environment
  - Java code to execute the benchmark load

2.1.2 Kit Usage

To submit a compliant TPCx-HS benchmark result, the test sponsor is required to use the TPCx-HS kit as provided.

The kit must be used as outlined in the TPCx-HS Users Guide.

The output of the TPCx-HS kit is called the **run report** which includes the following

- Version number of TPCx-HS kit
- Checksum for HSGen, HSDataCheck, HSSort and HSValidate Programs
- Output from HSGen, HSDataCheck (presort and post sort), HSSort and HSValidate Programs
- Cardinality of the test dataset presort and post sort
- Verification of data replication

2.1.3 Kit modification

2.1.3.1 Minor Shell Script Modifications

Minor modifications to the provided shell scripts in the TPCx-HS kit to facilitate operating system differences are allowed without TPC approval.

The following changes are considered minor modifications:

1. Shell script changes necessary for the kit scripts to execute on a particular operating system as long as the changes do not alter the execution logic of the script

2.1.3.2 Major Shell Script Modifications

All changes to the scripts included in the kit other than those allowed in section 2.1.3.1 would be considered major modifications. Major modifications must be approved by the TPC prior to being used in a benchmark submission.

It will be the judgment of the TPC members reviewing the submission or the TPCx-HS certified auditor (if being used) as to whether scripting changes are considered minor or major. If the test sponsor has any doubts they are encouraged to have the changes approved by the TPC prior to being used in a submission.

2.1.3.3 Java Code Modifications
No modifications are allowed to the java code provided in the TPCx-HS kit.

### 2.1.4 Future Kit Releases

The TPC will release future TPCx-HS benchmark kits at its discretion to fix bugs or add features. When a new kit version is released the TPC will release a timetable regarding the last date a benchmark submission can be made using the previous kit version. After this date only submissions using the new kit version will be considered, and submissions using the previous kit version will immediately be found non-compliant.

All new kit releases for this benchmark will be performance-neutral so the comparability to prior published results can be maintained. Performance-neutral in this context means that the new kit provides results that are no more than two percent faster or slower than the previous kit when run in the exact same configuration.

If the test sponsor would like new scripts or existing script changes to be included in a future release of the TPCx-HS benchmark kit, then the test sponsor can donate the scripts or script code changes to the TPC and work with the TPC to get them included in the next release.

If a test sponsor would like to see changes made to the java code of the kit then the changes should be provided to the TPC for potential inclusion in the next release of the TPCx-HS benchmark kit.

### 2.2 Benchmark Workload

The TPC Benchmark™HS (TPCx-HS) benchmark workload consists of the following modules:

- **HSGen** is a program to generate the data at a particular Scale Factor. (See Clause 4). HSGen is based on TeraGen
- **HSDataCheck** is a program to check the compliance of the dataset and replication.
- **HSSort** is a program to sort the data into a total order. HSSort is based on TeraSort [1]
- **HSValidate** is a program that validates the output is sorted. HSValidate is based on TeraValidate [1]


### 2.3 Benchmark Execution

A valid run will consist of five separate phases run sequentially. These phases may not overlap in their execution times. E.g. the start of Phase 2 may not begin until Phase 1 is complete and the start of Phase 3 may not begin until Phase 2 is complete etc. All phases are initiated by the <tpchs-master> script, which can be executed from any of the nodes in the SUT. The phases are listed below:

2. Dataset (See Clause 4) verification via HSDataCheck. The program is to verify the cardinality, size and replication factor of the generated data. If the HSDataCheck program reports failure then the run is considered invalid.
3. Running the sort using HSSort on the input data. This phase samples the input data and sorts the data.
4. Dataset (See Clause 4) verification via HSDataCheck. The program is to verify the cardinality, size and replication factor of the sorted data. If the HSDataCheck program reports failure then the run is considered invalid.

5. Validating the sorted output data via HSValidate. HSValidate validates the sorted data. This phase is not part of the primary metric but reported in the Full Disclosure Report. If the HSValidate program reports that the HSSort did not generate the correct sort order, then the run is considered invalid.

The completion time for each phase must be reported.

The benchmark test consists of two runs, Run1 and Run2, which must follow the Run Phases above. The performance run is defined as the run with the lower TPCx-HS Performance Metric. The repeatability run is defined as the run with the higher TPCx-HS Performance Metric. The reported performance metric is the TPCx-HS Performance Metric for the performance run.

No configuration or tuning changes are allowed between the two runs.

The benchmark execution phases are listed below and illustrated in Figure 1:

**Table 1: Figure 2-1 TPCx-HS Execution Phases**
2.3.1 No part of the SUT may be rebooted or restarted during or between the runs and any of the phases. If there is a non-recoverable error reported by any of the applications, operating system, or hardware in any of the five phases, the run is considered invalid. If a recoverable error is detected in any of the phases, and is automatically dealt with or corrected by the applications, operating system, or hardware then the run is considered valid. However, manual intervention by the test sponsor is not allowed. If the recoverable error requires manual intervention to deal with or correct then the run is considered invalid. A minimum of three way data replication must be maintained throughout the run.

2.4 Configuration and Tuning

The SUT cannot be reconfigured, changed, or re-tuned by the test sponsor during or between any of the five phases or between Run 1 and Run 2. Any manual tunings to the SUT must be performed before the beginning of Phase 1 of Run 1, and must be fully disclosed. Any automated changes or tuning performed by the OS or commercially available product between any of the phases is allowed. Any changes to default tunings or parameters of the applications, operating systems, or hardware of the SUT must be disclosed.
CLAUSE 3: SYSTEM UNDER TEST AND DRIVER

This clause defines the System Under Test (SUT) and the benchmark driver.

3.1 System Under Test

The tested and reported configuration is composed of those hardware and software components that are employed in the performance test and whose cost and performance are described by the benchmark metrics. Specifically, the SUT consists of (See figure 3.1):

- Devices, i.e. compute devices and/or data storage devices, including hardware and software components,
- Any hardware and software devices of all networks required to connect and support the SUT systems,
- Each compute device includes a benchmark specific software layer, the benchmark implementation, and other commercially available software products, and
- One benchmark specific driver, the benchmark driver. The driver may reside on one of the compute devices or on a separate system. In case the driver resides on a separate compute device, this device is part of the SUT.

3.1.1 Except for the benchmark implementation and the benchmark driver, all SUT components must be commercially available software or hardware products.

3.1.2 The source code of any non-commercially available components used to implement the SUT (such as scripts to configure the system, set tunables) must be disclosed.

3.1.3 The driver presents the workload to the SUT. The driver is a logical entity that can be implemented using one or more programs, processes, or systems.
This clause defines Scale Factor and Metrics.

4.1 Scale factor

The TPCx-HS follows a stepped size model. **Scale factor** (SF) used for the test dataset must be chosen from the set of fixed Scale Factors defined as follows:

1TB, 3TB, 10TB, 30TB, 100TB, 300TB, 1000TB, 3000TB, 10000TB.

The corresponding number of records are 10B, 30B, 100B, 300B, 1000B, 3000B, 10000B, 30000B, 100000B, where each record is 100 bytes generated by HSGen.

**Comment:** The TPC will continuously evaluate adding larger Scale Factors and retiring smaller Scale Factors based on industry trends.

4.2 Metrics

4.2.1 TPC-xHS defines three primary metrics:

1. HSph@SF, the Performance Metric, reflecting the TPCx-HS throughput; where SF is the Scale Factor (seed in Clause 4.1);
2. $/HSph@SF, the Price-Performance metric;
3. System availability date (see Clause 5);
4. When TPC-Energy option is chosen for reporting, the TPCx-HS energy metric reports the power per performance and is expressed as Watts/HSph@SF (see TPCx-Energy specification for additional requirements).

4.3 Performance Metric (HSph@SF)

The performance metric of the benchmark is HSph@SF, the effective sort throughput of the benchmarked configuration:

\[
HSph \@ SF = \frac{SF}{T / 3600}
\]

Where:

- SF is the Scale Factor (See Clause 4.1),
- T is the total elapsed time for the run in seconds

**Comment:** The performance run is defined as the run with the lower TPCx-HS Performance Metric (See Clause 2.3).
4.4 Price Performance Metric ($/HSph@SF)

The price-performance metric for the benchmark is defined as:

\[
\frac{P}{HSph \@ SF} = \frac{P}{HSph \@ SF}
\]

Where:

- \(P\) is the total cost of ownership of the SUT.

4.4.1 If a benchmark configuration is priced in a currency other than US dollars, the units of the price-performance metrics may be adjusted to employ the appropriate currency.

4.5 System Availability Date

The System Availability Date is defined in the TPC Pricing Specification.

4.6 Metric Comparison

A TPCx-HS Result is only comparable with other TPCx-HS Results of the same Scale Factor (See Clause 4).

4.6.1 Results at the different Scale Factors are not comparable, due to the substantially different computational challenges found at different data volumes. Similarly, the system price/performance may not scale down linearly with a decrease in dataset size due to configuration changes required by changes in dataset size.

4.6.2 If results measured against different Scale Factors appear in a printed or electronic communication, then each reference to a result or metric must clearly indicate the Scale Factors against which it was obtained. In particular, all textual references to TPCx-HS metrics (performance or price/performance) appearing must be expressed in the form that includes the Scale Factor as an integral part of the metric’s name; i.e. including the “@SF” suffix. This applies to metrics quoted in text or tables as well as those used to annotate charts or graphs. If metrics are presented in graphical form, then the Scale Factor on which metric is based must be immediately discernible either by appropriate axis labeling or data point labeling.

In addition, the results must be accompanied by a disclaimer stating: "The TPC believes that comparisons of TPCx-HS results measured against different Scale Factors are misleading and discourages such comparisons”.

4.7 Required Reporting Components

To be compliant with the TPCx-HS standard and TPC Policies, the URL to the benchmark result and Availability Date of the complete configuration must be included for all public references (See Clause 8.3 of TPC Policies).
CLAUSE 5: PRICING

This section defines the components, functional requirements of what is priced, and what substitutions are allowed. Rules for pricing the Priced System and associated software and maintenance are included in the TPC Pricing Specification located at www.tpc.org.

5.1 Priced System

The system to be priced shall include the hardware and software components present in the System Under Test (SUT), a communication interface that can support user interface devices, additional operational components configured on the test system, and maintenance on all of the above.

5.1.1 Calculation of the priced system consists of:
- Price of the SUT as tested and defined in Clause 3;
- Price of a communication interface capable of supporting the required number of user interface devices defined in Clause 5.2;
- Price of additional products (software or hardware) required for customary operation, administration and maintenance of the SUT for a period of 3 years;
- Price of all products required to create, execute, administer, and maintain the executables or necessary to create and populate the test environment.

5.1.2 Specifically excluded from the priced system calculation are:
- End-user communication devices and related cables, connectors, and switches;
- Equipment and tools used exclusively in the production of the full disclosure report;

5.2 User Interface Devices and Communications

The priced system must include the hardware and software components of a communication interface capable of supporting a number of user interface devices (e.g., terminals, workstations, PCs, etc.).

5.2.1 Only the interface is to be priced. Not to be included in the priced system are the user interface devices themselves and the cables, connectors and concentrators used to connect the user interface devices to the SUT. For example, in a configuration that includes an Ethernet interface to communicate with PCs, the Ethernet card and supporting software must be priced, but not the Ethernet cables and the PCs.

5.2.2 The communication interface used must be an industry standard interface, such as Ethernet, Token Ring, or RS232.

5.3 Additional Operational Components

5.3.1 Additional products that might be included on a customer installed configuration, such as operator consoles and magnetic tape drives, are also to be included in the priced system if explicitly required for the operation, administration, or maintenance, of the priced system.
5.3.2 Copies of the software, on appropriate media, and a software load device, if required for initial load or maintenance updates, must be included.

5.3.3 The price of all cables used to connect components of the system (except as noted Clause 5.3) must be included.

5.4 Software

All software licenses must be priced.

5.5 Allowable Substitutions

Substitution is defined as a deliberate act to replace components of the Priced Configuration by the test sponsor as a result of failing the availability requirements of the TPC Pricing Specification or when the part number for a component changes.

Comment: Corrections or "fixes" to components of the Priced Configuration are often required during the life of products. These changes are not considered Substitutions so long as the part number of the priced component does not change. Suppliers of hardware and software may update the components of the Priced Configuration, but these updates must not negatively impact the reported performance metric or numerical quantities more than two percent.

5.5.1 The following are not considered substitutions:

- software patches to resolve a security vulnerability
- silicon revision to correct errors
- new supplier of functionally equivalent components (i.e. memory chips, disk drives, ...)

5.5.2 Some hardware components of the Priced Configuration may be substituted after the test sponsor has demonstrated to the auditor's satisfaction that the substituting components do not negatively impact the reported performance metric or numerical quantities. All substitutions must be reported in the FDR and noted in the auditor's attestation letter. The following hardware components may be substituted:

- Durable Medium (i.e. disk drives) and Cables

Comment: Durable Medium is defined as a data storage medium that is inherently non-volatile such as a magnetic disk or tape.

Comment: If any hardware component is substituted then the result must be audited by a TPC certified Auditor (See Clause 8).
CLAUSE 6: ENERGY

This section contains the rules and methodologies for measuring and reporting energy metric in TPCx-HS benchmarks. This includes the energy consumption of system components associated with typical business information technology environments, which are characterized by:

- Energy consumption of servers
- Energy consumption of applications
- Energy consumption of other items that consume power and are required by the benchmark specification as components of the System Under Test (SUT).

Rules for the Energy configuration and associated software and maintenance are included in the TPC-Energy Specification located at [www.tpc.org](http://www.tpc.org).

6.1 General Measurement Guidelines

The TPCx-HS energy results are expected to be accurate representations of system performance and energy consumption. Therefore there are certain requirements which must be followed. The approach and methodology are explicitly detailed in this specification and the TPC Benchmark Standards, as defined in TPC-Energy.

When TPC-Energy option is chosen for reporting, power consumption is measured for Run 1 and Run 2. The energy metric reported is calculated for the **reported run**.

Figure 6.1: Power Measurement Interval

TPCx-HS Energy metric reports the power per performance and is expressed as Watts/HSph@SF.

6.2.1 Energy Calculation

The computation of the total energy consumption for the *performance run* must be disclosed. If the energy of the entire Priced Configuration is not derived from direct measurements, the methods for deriving the energy for components which were not measured must be disclosed. The average power consumption for the *performance run* must be disclosed. The TPC-Energy Primary Metric must be disclosed, including the calculation that is used to derive it.
If the optional TPC-Energy Secondary Metrics are reported, the components which are included in each subsystem must be identified. This can be achieved with separate lists to be included in the FDR or with a specific designation in the price spreadsheet (See Clause 7.10). Every component that consumes energy must be included in exactly one subsystem. For each defined subsystem, the calculations defined for the TPC-Energy Secondary Metrics must be reported, using the reported performance metric of the entire SUT and the energy consumption for each subsystem under report.

\[ P = \sum_{1 \leq i \leq m} p_i \]
\[ E = \int_0^T P(t) dt \]

Where \( p_i \) is power measurement of each subsystem \( i \) during the run.

For example \( p_i = p_{\text{storage}} \) is the power consumption of the storage subsystem for the run.

\( T \) is the time elapsed time for the performance run.

\( P(t) \) power consumption measured at time \( t \)

**Figure 6.2: Power Measurement for Subsystems**
6.2.2 For idle Power reporting please refer the TPC-Energy Specification.

6.3 TPCx-HS Energy Metric

The energy measurement from the reported run is used to compute the TPCx-HS power metric at the chosen Scale Factor. It must be computed as:

\[ \frac{E}{(T \times HSph@SF)} \]

Where:

- E is the energy consumption for the reported run
- T is the elapsed time in seconds for the reported run
- \( HSph@SF \) is the reported performance metric

6.3.1 The units of Energy Metric are reported to one digit after the decimal point, rounded to the nearest 0.1.

6.4 Reporting TPC-Energy Optional Metrics

When the TPC-Energy optional reporting is selected by the test sponsor, the additional requirements and format of TPC-Energy related items in the executive summary and FDR are included in the TPC Energy Specification, located at www.tpc.org.
Rules for reporting Pricing information are included in the TPC Pricing Specification located at www.tpc.org.

7.1 Reporting Requirements

7.1.1 A Full Disclosure Report (FDR) in pdf format and Executive Summary are required.

7.1.2 The intent of this disclosure is to simplify comparison between results and for a customer to be able to replicate the results of this benchmark given appropriate documentation and products.

7.2 Format Guidelines

7.2.1 While established practice or practical limitations may cause a particular benchmark disclosure to differ from the examples provided in various small ways, every effort should be made to conform to the format guidelines. The intent is to make it as easy as possible for a reviewer to read, compare and evaluate material in different benchmark disclosures.

7.2.2 All sections of the report, including appendices, must be printed using font sizes of a minimum of 8 points.

7.2.3 The Executive Summary must be included near the beginning of the full disclosure report.

7.3 Full Disclosure Report

The FDR should be sufficient to allow an interested reader to evaluate and, if necessary, recreate an implementation of TPCx-HS. If any sections in the FDR refer to another section of the report, the names of the referenced scripts/programs must be clearly labeled in each section. Unless explicitly stated otherwise “disclosed” refers to disclosed in the FDR.

Comment: Since the building test environment may consist of a set of scripts and corresponding input files, it is important to disclose and clearly identify, by name, scripts and input files in the FDR.

The order and titles of sections in the test sponsor’s full disclosure report must correspond with the order and titles of sections from the TPCx-HS standard specification (i.e., this document).

7.4 General Items

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

7.4.2 Settings must be provided for all customer-tunable parameters and options that have been changed from the defaults found in actual products, including but 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.

7.4.3 Explicit response to individual disclosure requirements specified in the body of earlier sections of this document must be provided.

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.

The following sample diagram illustrates a measured benchmark configuration using Ethernet, an external driver, and four processors each with two cores and four threads per node in the SUT. Note that this diagram does not depict or imply any optimal configuration for the TPCx-HS benchmark measurement.

Depending on the implementation of the SUT the Name Node, Job Tracker, Task Tracker, Data Nodes etc or the functional equivalents must be specified in the diagram.

Figure 7.1: Sample Configuration Diagram (the front system box describes one node)
Total Nodes: 8 X a243DX Server
Total Processors/Cores/Threads: 16/32/64
Total Memory: 1024GB
Total Number of Storage Drives/Devices: 96
Total Storage Capacity: 1TB

4 Nodes of
Processor/Cores/Threads: 2/4/8 X 3 GHz, 4 MB L2
Memory: 128 GB
Controller: 1 x SAS RAID Controller
Drives: 12 X 1TB 7.2K RPM LFF SATA
Network: 1GB Ethernet

4 Nodes of
Processor/Cores/Threads: 2/4/8 X 2.6 GHz, 4 MB L2
Memory: 64 GB
Controller: 1 x SAS RAID Controller
Drives: 6 X 1TB 7.2K RPM LFF SATA
Network: 1GB Ethernet

**Comment:** Detailed diagrams for system configurations and architectures can vary widely, and it is impossible to provide exact guidelines suitable for all implementations. The intent here is to describe the system components and connections in sufficient detail to allow independent reconstruction of the measurement environment. This example diagram shows homogeneous nodes. This does not preclude tests sponsors from using heterogeneous nodes as long as the system diagram reflects the correct system configuration.

7.4.5 The distribution of dataset across all media must be explicitly described using a format similar to that shown in the following example for both the tested and priced systems.

**Table 7-1: Sample Layout Description**
<table>
<thead>
<tr>
<th>Server</th>
<th>Controller</th>
<th>Disk Drive</th>
<th>Description of Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>40A</td>
<td>0</td>
<td>Operating system, root, swap, Hadoop Master</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1-12</td>
<td>Distributed file system Metadata</td>
</tr>
<tr>
<td>2</td>
<td>40A</td>
<td>0</td>
<td>Operating system, root, swap</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1-12</td>
<td>33% of Dataset</td>
</tr>
<tr>
<td>3</td>
<td>40A</td>
<td>0</td>
<td>Operating system, root, swap</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1-12</td>
<td>33% of Dataset</td>
</tr>
<tr>
<td>4</td>
<td>40A</td>
<td>0</td>
<td>Operating system, root, swap</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1-12</td>
<td>33% of Dataset</td>
</tr>
</tbody>
</table>

7.4.6 The distribution of various software components across the system must be explicitly described using a format similar to that shown in the following example for both the tested and priced systems.

**Table 7-2: Distribution of Software Components**

<table>
<thead>
<tr>
<th>Server</th>
<th>Software Component</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Name Node</td>
</tr>
<tr>
<td></td>
<td>Resource Manager</td>
</tr>
<tr>
<td>2</td>
<td>Data Node</td>
</tr>
<tr>
<td>3</td>
<td>Data Node</td>
</tr>
<tr>
<td>4</td>
<td>Data Node</td>
</tr>
</tbody>
</table>
7.5 Clause 2 – Workload Related Items

7.5.1 Script or text used to set for all hardware and software tunable parameters must be reported.

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

7.5.1 The run report generated by TPCx-HS benchmark kit must be reported.

7.6 Clause 3 – SUT Related Items

7.6.1 All hardware and software options must be reported.

7.6.2 The data storage ratio must be disclosed. It is computed by dividing the total physical data storage present in the priced configuration (expressed in TB) by the chosen Scale Factor as defined in Clause 4.1. Let \( r \) be the ratio. The reported value for \( r \) must be rounded to the nearest 0.01. That is, reported value = round\((r, 2)\). For example, a system configured with 96 disks of 1TB capacity for a 1TB Scale Factor has a data storage ratio of 96.

Comment: For the reporting of configured disk capacity, terabyte (TB) is defined to be \( 10^{12} \) bytes.

The Scale Factor to memory ratio must be disclosed. It is computed by dividing the Scale Factor by the total physical memory present in the priced configuration (see clause 3). Let \( r \) be this ratio. The reported ratio must be rounded to the nearest 0.1. That is, reported value = round\((r, 1)\). For example, a system configured with 1TB of physical memory for a 10TB Scale Factor has a memory ratio of 10.

7.7 Clause 5 - Performance Metrics and Execution Rules Related Items

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

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

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

7.7.4 The performance metric (HSph@SF) must be disclosed for Run1 and Run2. Price-performance metric ($/HSph@SF) must be disclosed for the performance run. See Clause 2.3 and Clause 4.

7.8 Clause 8 - Audit Related Items

7.8.1 If the benchmark is audited by an Independent Auditor, the auditor’s agency name, address, phone number, and attestation letter with a brief audit summary report indicating compliance must be included in the full disclosure report. A statement should be included specifying whom to contact in order to obtain further information regarding the audit process.
7.9 Executive Summary

The executive summary is meant to be a high level overview of a TPCx-HS implementation. It should provide the salient characteristics of a benchmark execution (metrics, configuration, pricing, etc.) without the exhaustive detail found in the FDR. When the TPC-Energy optional reporting is selected by the test sponsor, the additional requirements and format of TPC-Energy related items in the executive summary are included in the TPC Energy Specification, located at www.tpc.org.

The executive summary has three components:

- Implementation Overview
- Pricing Spreadsheet
- Numerical Quantities

7.10 Page Layout

Each component of the executive summary should appear on a page by itself. Each page should use a standard header and format, including:

- 1/2 inch margins, top and bottom;
- 3/4 inch left margin, 1/2 inch right margin;
- 2 pt. frame around the body of the page. All interior lines should be 1 pt.;

7.11 Implementation Overview

The implementation overview page contains five sets of data, each laid out across the page as a sequence of boxes using 1 pt. rule, with a title above the required quantity. Both titles and quantities should use a 9-12 pt. Times font unless otherwise noted.

7.11.1 The first section contains information about the sponsor and system identification

Table 7-3: Sponsor and System Identification

<table>
<thead>
<tr>
<th>Title</th>
<th>Font</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sponsor Name or Logo</td>
<td>16-20 pt. Bold (for Name)</td>
</tr>
<tr>
<td>System Identification</td>
<td>16-20 pt. Bold</td>
</tr>
<tr>
<td>Version Numbers for TPCx-HS, TPC-Pricing and TPC-Energy (if reported)</td>
<td>16-20 pt. Bold</td>
</tr>
<tr>
<td>Report Date</td>
<td>16-20 pt. Bold</td>
</tr>
</tbody>
</table>
Comment 1: It is permissible to use or include company logos when identifying the sponsor.

Comment 2: The report date must be disclosed with a precision of 1 day. The precise format is left to the test sponsor.

7.11.2 The second section contains the Total System Cost; and, TPCx-HS Performance Metric and Price/Performance for the performance run.

<table>
<thead>
<tr>
<th>Table 7-4 Test Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
</tr>
<tr>
<td>Total System Cost</td>
</tr>
<tr>
<td>TPCx-HS Performance Metric</td>
</tr>
<tr>
<td>Price/Performance</td>
</tr>
</tbody>
</table>

Depending on the currency used for publication this sign has to be exchanged with the ISO currency symbol

7.11.3 The third section contains detailed the system configuration

<table>
<thead>
<tr>
<th>Table 7-5: System Configuration Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
</tr>
<tr>
<td>Operating System</td>
</tr>
<tr>
<td>Other Software</td>
</tr>
</tbody>
</table>
System Availability Date
The Availability Date of the system, defined in Clause 0 of the TPC Pricing Specification.

Comment: The Software Version must uniquely identify the orderable software product referenced in the Priced Configuration (e.g., RALF/2000 4.2.1)

7.11.4 The fourth section contains the storage and memory ratios. See Clause 7.6.2

<table>
<thead>
<tr>
<th>Title</th>
<th>Precision</th>
<th>Font</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical Storage /Scale Factor</td>
<td>0.01</td>
<td>9-12 pt. Times</td>
</tr>
<tr>
<td>Scale Factor/Physical Memory</td>
<td>0.01</td>
<td>9-12 pt. Times</td>
</tr>
</tbody>
</table>

7.11.5 The fifth section contains the components, including

- Total number of nodes used/total number of processors used with their types and speeds in GHz/ total number of cores used/total number of threads used, See TPC Policies;
- Main and cache memory sizes;
- Network and I/O connectivity;
- Disk quantity and geometry.

7.10 Pricing Spreadsheet

The major categories in the Price Spreadsheet, as appropriate, are:

- Network(s)
- Server(s)/Node(s)
- Storage
- Software

Discounts (may optionally be included with above major category subtotal calculations).
7.11 Numerical Quantities Summary

The Numerical Quantities Summary page contains two sets of data, presented in tabular form, detailing the execution timings for the reported execution of the performance test. Each set of data should be headed by its given title and clearly separated from the other tables.

7.11.1 The first section contains measurement results from the benchmark execution.

<table>
<thead>
<tr>
<th>Item Title</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale Factor</td>
<td>1</td>
</tr>
<tr>
<td>Run Start Time</td>
<td>yyyy-mm-dd hh:mm:ss</td>
</tr>
<tr>
<td>Run End Time</td>
<td>yyyy-mm-dd hh:mm:ss</td>
</tr>
<tr>
<td>Run Elapsed Time</td>
<td>ss.sss</td>
</tr>
<tr>
<td>Start of HSGen</td>
<td>yyyy-mm-dd hh:mm:ss</td>
</tr>
<tr>
<td>End of HSGen</td>
<td>yyyy-mm-dd hh:mm:ss</td>
</tr>
<tr>
<td>HSGen Time</td>
<td>ss.sss</td>
</tr>
<tr>
<td>Start of HSSort</td>
<td>yyyy-mm-dd hh:mm:ss</td>
</tr>
<tr>
<td>End of HSSort</td>
<td>yyyy-mm-dd hh:mm:ss</td>
</tr>
<tr>
<td>HSSort Time</td>
<td>ss.sss</td>
</tr>
<tr>
<td>Start of HSValidate</td>
<td>yyyy-mm-dd hh:mm:ss</td>
</tr>
<tr>
<td>End of HSValidate</td>
<td>yyyy-mm-dd hh:mm:ss</td>
</tr>
<tr>
<td>HSValidate Time</td>
<td>ss.sss</td>
</tr>
</tbody>
</table>

7.11.2 Second section contains the measurement result for **repeatability run**. See Table 7-7 for contents and precision.
7.12 TPCx-HS Run Report

The run report from TPCx-HS must be included in page 4 of the Executive Summary.

7.13 Availability of the Full Disclosure Report

- The full disclosure report must be readily available to the public. The report and must be made available when results are made public. In order to use the phrase “TPC Benchmark HS”, the full disclosure report submitted electronically to the TPC using the procedure described in the TPC Policies and Guidelines document.
- The official full disclosure report must be available in English but may be translated to additional languages.

7.14 Revisions to the Full Disclosure Report

Revisions to the full disclosure documentation shall be handled as follows:

- Substitutions will be open to challenge for a 60 day period. No other portion of the FDR and supporting files archive are challengeable.
- During the normal product life cycle, problems will be uncovered that require changes, sometimes referred to as ECOs, FCOs, patches, updates, etc. When the cumulative result of applied changes causes the HSph rating of the system to decrease by more than two percent from the initially reported HSph, then the test sponsor is required to re-validate the benchmark results. The complete revision history is maintained following the timing interval section showing the revision date and description.
- Full disclosure report and supporting files archive revisions may be required for other reasons according to TPC policies (see TPC Policies).
CLAUSE 8: AUDIT

Rules for auditing Pricing information are included in the TPC Pricing Specification located at www.tpc.org.

When the TPC-Energy optional reporting is selected by the test sponsor, the rules for auditing of TPC-Energy related items are included in the TPC Energy Specification located at www.tpc.org.

8.1 General Rules

An independent audit or peer audit of the benchmark result is required before publication.

The vendor may choose an independent audit or peer audit.

8.1.1 The term independent is defined as “the outcome of the benchmark carries no financial benefit to the auditing agency other than fees earned directly related to the audit.” The auditing agency cannot have supplied any performance consulting under contract for the benchmark. The Independent Auditor must be certified by the TPC to audit TPCx-HS. The Independent Auditor must meet the following:

- The auditing agency cannot be financially related to the sponsor. For example, the auditing agency is financially related if it is a dependent division of the sponsor, the majority of its stock is owned by the sponsor, etc.
- The auditing agency cannot be financially related to any one of the suppliers of the measured/priced configuration.
- The auditor's attestation letter is to be made readily available to the public as part of the full disclosure report. A detailed report from the auditor is not required.

8.1.2 The term peer review is defined as the process of reviewing benchmark results for compliance prior to publication by a committee named the peer review committee. The peer review committee consists of 3 members from the TPCx-HS committee. Each member serves a period of 3 months. The membership will be rotated through the TPCx-HS membership. The submission is confidential to the peer review committee until the result is published. The peer review committee must complete the review in 10 business days. If no issues are raised in 10 days the result is considered valid.

8.1.2 TPCx-HS results can be used as the basis for new TPCx-HS results if and only if:

- The auditor or peer review committee ensure that the hardware and software products are the same as those used in the prior result;
- The auditor or peer review committee reviews the FDR of the new results and ensures that they match what is contained in the original sponsor's FDR;
- The auditor or peer review committee can attest to the validity of the pricing used in the new FDR.

Comment 1: The intent of this clause is to allow a reseller of equipment from a given supplier to publish under the re-seller's name a TPCx-HS result already published by the supplier.
8.2 Audit Check List

8.2.1 Clause 2 Related Items

8.2.1.1 Verify that the TPC provide kit is used

8.2.1.2 Verify that all phases are complete with no error in Run1 and Run2

8.2.1.3 Verify that all scripts and source code to implement the benchmark is included.

8.2.2 Clause 3 Related Items

8.2.2.1 Verify that all components of the SUT is publically available

8.2.2.2 Verify that all components of the SUT is included in the pricing

8.2.3 Clause 4 Related Items

8.2.3.1 Verify that the system is scaled as per the specification

8.2.3.2 Verify that the metrics are reported as per the precision requirements in clause 7

8.2.4 Clause 5 Related Items

8.2.4.1 Verify that the benchmark is in compliance with the TPC Pricing specification

8.2.5 Clause 6 Related Items

8.2.5.1 Verify that the benchmark is in compliance with the TPC Energy specification (if reported)

8.2.6 Clause 7 Related Items

8.2.6.1 Verify that full disclosure report and executive summary report are accurate and comply with the reporting requirements. This includes:

- Metric calculation;
- System availability;
- The diagrams of both measured and priced configuration.
- System pricing;
- The numerical quantity summary