Dell Inc.

TPC Express Benchmark™ Big Bench (TPCx-BB)

Full Disclosure Report

for

Dell PowerEdge 14G R640/R740xd

(with 19x R640/R740xd)

using

Cloudera CDH 5.13.1

and

Red Hat Enterprise Linux Server 7.3

First Edition

March 20, 2018

***Dell Inc. (Dell)**, 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™ BB 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.

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

TPC Benchmark™, TPCx-BB and BBQpm, are registered certification marks of the Transaction Processing Performance Council.

The Dell 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 Dell business contact for information on the products or services available in your area. You can find additional information via Dell's web site at www.dell.com. Actual performance and environmental costs of Dell products will vary depending on individual customer configurations and conditions.

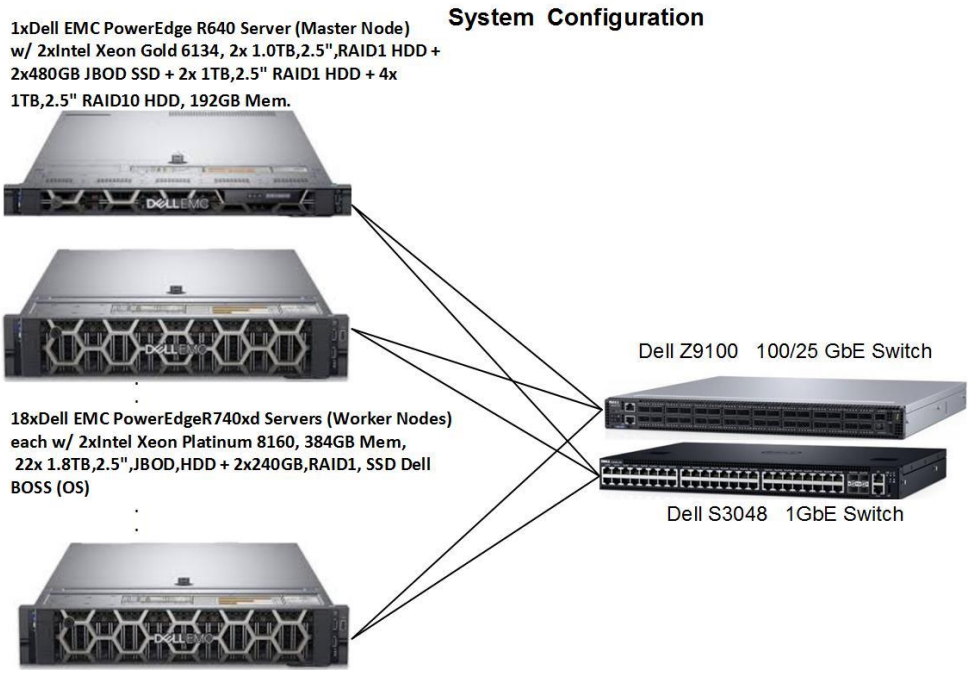
Copyright © 2018 Dell Inc.

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

DELL EMC	**Dell PowerEdge 14G R640/R740xd**	TPCx-BB Rev. v1.2.0 TPC-Pricing Rev. v2.2.0 Report Date: March 20, 2018
Total System Cost	TPCx-BB Performance Metric	Price/Performance
908,125 USD	**1,660.75** BBQpm@10000	**546.82 USD** $/BBQpm@10000

Framework	Operating System	Other Software	Availability Date	Scale Factor	Streams
Cloudera CDH 5.13.1	Red Hat Enterprise Linux Server 7.3	None	March 20, 2018	10000	8

System Configuration



Physical Storage/Scale Factor: 73.04	Scale Factor/Physical Memory: 1.41

Servers: Total Processors/Cores/Threads	19x R640/R740xd 38/880/1,760	
Server Configuration: Processors Memory Storage Controller Storage Device Network	**R640 (Master Node)** 2x Intel Xeon Gold 6134 @3.20GHz 192GB PERC H730P 12Gbps RAID Controller 2x 480GB SSD 8x 1TB HDD Mellanox ConnectX-4 Lx DP 25Gbe	**18x R740xd (Worker Nodes):** 2x Intel Xeon Platinum 8160 @2.10GHz 384GB HBA 330+ 12Gbps SAS Controller 2x 240GB SSD 22x 1.8TB HDD Mellanox ConnectX-4 Lx DP 25Gbe
Connectivity:	Dell Z9100 100/25 GbE Switch Dell S3048 1GbE Switch	

Description	Part Number	Key		Unit Price	Qty	ExtendedPrice	3 yr. Maint.Price
HARDWARE COMPONENTS							
PowerEdge R740XD Server	210-AKZR	1	$25,520.00	$63,779.47	18	$1,148,030.46	
Chassis with Up to 24 x 2.5" Hard Drives for 2CPU	321-BCPY	1	$0.00	$0.00	18		
No Trusted Platform Module	461-AADZ	1	1	$0.00	18		
PowerEdge R740XD Shipping	340-BLBE	1	1	$0.00	18		
Intel® Xeon® Platinum 8160 2.1G,24C/48T,10.4GT/s, 33M Cache,Turbo,HT (150W) DDR4-2666	338-BLMR	1	1	$0.00	18		
Intel® Xeon® Platinum 8160 2.1G,24C/48T,10.4GT/s, 33M Cache,Turbo,HT (150W) DDR4-2666	374-BBOM	1	1	$0.00	18		
2 Standard Heatsinks for greater than 125W CPUs (no MB or GPU)	412-AAIR	1	1	$0.00	18		
6 Performance Fans forR740/740XD	384-BBPZ	1	1	$0.00	18		
2666MT/s RDIMMs	370-ADNU	1	1	$0.00	18		
Performance Optimized	370-AAIP	1	1	$0.00	18		
32GB RDIMM, 2666MT/s, Dual Rank	370-ADNF	1	1	$0.00	216		
C7, Unconfigured RAID for HDDs or SSDs (Mixed Drive Types Allowed)	780-BCDS	1	1	$0.00	18		
HBA330+ Controller, 12Gbps Mini card	405-AANV	1	1	$0.00	18		
1.8TB 10K RPM SAS 12Gbps 512e 2.5in Hot-plug Hard Drive	400-ARXC	1	1	$0.00	396		
iDRAC9,Enterprise	385-BBKT	1	1	$0.00	18		
ReadyRails™ Sliding Rails Without Cable Management Arm	770-BBBQ	1	1	$0.00	18		
PowerEdge 2U Standard Bezel	325-BCHU	1	1	$0.00	18		
Power Saving Dell Active Power Controller	750-AABF	1	1	$0.00	18		
NEMA 5-15P to C13 Wall Plug, 125 Volt, 15 AMP, 10 Feet (3m), Power Cord, North America	450-AALV	1	1	$0.00	36		
Dual, Hot-plug, Redundant Power Supply (1+1), 1100W	450-ADWM	1	1	$0.00	18		
No Systems Documentation, No OpenManage DVD Kit	631-AACK	1	1	$0.00	18		
Red Hat Enterprise Linux Non Factory Install,x64,Req Lic&Sub Selection	421-4727	1	1	$0.00	18		
No Media Required	421-5736	1	1	$0.00	18		
Red Hat Enterprise Linux,2SKT,1 Physical OR 2Guest,3Yr PREMIUM SUB,No Media	634-BJBO	1	1	$0.00	18		
UEFI BIOS Boot Mode with GPT Partition	800-BBDM	1	1	$0.00	18		
BOSS controller card + with 2 M.2 Sticks 240G (RAID 1),FH	403-BBPT	1		$0.00	18		
PowerEdge R740 Shipping Material	343-BBFU	1		$0.00	18		
iDRAC Group Manager, Enabled	379-BCQV	1		$0.00	18		
iDRAC,Legacy Password	379-BCSG	1		$0.00	18		
Riser Config 2, 3 x8, 1 x16 slots	330-BBHB	1		$0.00	18		
Mellanox ConnectX-4 Lx Dual Port 25GbE DA/SFP rNDC	406-BBLG	1		$0.00	18		
Quick Sync 2 (At-the-box mgmt)	350-BBJU	1		$0.00	18		
INFO QS, DATA ANALYTICS BUNDLE	379-BBWM	1		$0.00	18		
3 Years ProSupport and Mission Critical 24x7, 4Hr Onsite Service	813-6068, 813-9295, 813-9296, 951-2015	1		$2,677.44	18		$48,193.92

Description	Part Number	Key	Unit Price	Qty	ExtendedPrice	3 yr. Maint.Price
PowerEdge R640 Server	210-AKWU	1	$31,531.69	1	$31,531.69	
2.5" Chassis with up to 10 Hard Drives and 3PCIe slots	321-BCQL	1	$0.00	1		
No Trusted Platform Module	461-AADZ	1	$0.00	1		
PowerEdge R640 Shipping	340-BKNE	1	$0.00	1		
Intel® Xeon® Gold 6134 3.2G,8C/16T,10.4GT/s, 24.75M Cache,Turbo,HT (130W) DDR4-2666	338-BLMF	1	$0.00	1		
Intel® Xeon® Gold 6134 3.2G,8C/16T,10.4GT/s, 24.75M Cache,Turbo,HT (130W) DDR4-2666	374-BBNZ	1	$0.00	1		
Standard Heatsink for 2 CPU	370-ABWE	1	$0.00	1		
8 Standard Fans for R640	384-BBQJ	1	$0.00	1		
2666MT/s RDIMMs	370-ADNU	1	$0.00	1		
Performance Optimized	370-AAIP	1	$0.00	1		
16GB RDIMM, 2666MT/s, Dual Rank	370-ADND	1	$0.00	12		
C7, Unconfigured RAID for HDDs or SSDs (Mixed Drive Types Allowed)	780-BCDS	1	$0.00	1		
PERC H730P RAID Controller, 2GB NV Cache, Minicard	405-AANT	1	$0.00	1		
480GB SSD SAS Mix Use 12Gbps 512n 2.5in Hot-plug Drive, PX05SV,3 DWPD,2628 TBW	400-ASEO	1	$0.00	2		
1TB 7.2K RPM NLSAS 12Gbps 512n 2.5in Hot-plug Hard Drive	400-ASHE	1	$0.00	8		
iDRAC9 Enterprise with OME Server Configuration Management	385-BBKT	1	$0.00	1		
No Internal Optical Drive for 10 or 24 HDD Chassis	429-AAIQ	1	$0.00	1		
ReadyRails™ Sliding Rails Without Cable Management Arm	770-BBBC	1	$0.00	1		
No Bezel for x10 chassis	350-BBBW	1	$0.00	1		
Power Saving Dell Active Power Controller	750-AABF	1	$0.00	1		
NEMA 5-15P to C13 Wall Plug, 125 Volt, 15 AMP, 10 Feet (3m), Power Cord, North America	450-AALV	1	$0.00	2		
Dual, Hot-plug, Redundant Power Supply (1+1), 750W	450-ADWS	1	$0.00	1		
No Systems Documentation, No OpenManage DVD Kit	631-AACK	1	$0.00	1		
Red Hat Enterprise Linux Non Factory Install,x64,Req Lic&Sub Selection	421-4727	1	$0.00	1		
No Media Required	421-5736	1	$0.00	1		
Red Hat Enterprise Linux,2SKT,1 Physical OR 2Guest,3Yr PREMIUM SUB,No Media	634-BJBO	1	$0.00	1		
UEFI BIOS Boot Mode with GPT Partition	800-BBDM	1	$0.00	1		
Energy Star	387-BBMK	1	$0.00	1		
PowerEdge R640 x4 and x10 Drive Shipping Material	340-BLUC	1	$0.00	1		
iDRAC Group Manager, Enabled	379-BCQV	1	$0.00	1		
iDRAC,Legacy Password	379-BCSG	1	$0.00	1		
Riser Config 4, 2x16 LP	330-BBGY	1	$0.00	1		
Mellanox ConnectX-4 Lx Dual Port 25GbE DA/SFP rNDC	406-BBLG	1	$0.00	1		
Quick Sync 2 (At-the-box mgmt)	350-BBKC	1	$0.00	1		
INFO QS, DATA ANALYTICS BUNDLE	379-BBWM	1	$0.00	1		
Dell Netshelter SX 42U Rack - 600mm Wide x 1070mm Deep	A7545497	1	$1,299.99	1	$1,299.99	
Logitech MK120 Keyboard and Mouse	A6999510	1	$13.99	1	$13.99	
Dell 24 Monitor	210-AIWG	1	$169.99	1	$169.99	
3 Years ProSupport and Mission Critical 24x7, 4Hr Onsite Service	813-9255, 813-9259, 813-9265, 989-3439	1	$2,505.31	1		$2,505.31

DELL EMC	**Dell PowerEdge 14G R640/R740xd**	TPCx-BB Rev. v1.2.0 TPC-Pricing Rev. v2.2.0 Report Date: March 20, 2018

Description	Part Number	Key		Unit Price	Qty	ExtendedPrice	3 yr. Maint.Price
Dell Networking Z9100-ON Pod Switch		1		$51,508.00	1	$51,508.00	
Dell Networking Z9100-ON, 32x QSFP28 and 2x SFP+ fixed ports, PSU to IO airflow, 2x AC PSUs, OS9	210-AETC	1		$0.00	1		
OS9 installed on Z9100-ON, with entitlement to OS10 Enterprise	634-BPDD	1		$0.00	1		
Dell Networking,100GbE Q28 to 4xSFP28 Passive Copper Breakout Cable,2 Meter	470-ABOS	1		$0.00	5		
Power Cord, 250V, 12A, 2 Meters, C13/C14, Qty 2	450-AASX, 450-AASX	1		$0.00	1		
Dell Networking Z9100-ON User Guide	631-AAPO	1	$21,156.00	$0.00	1		
Software, Rights to use L3 on OS9, Z9100-ON	634-BEBM	1	$0.00	$0.00	1		
No Installation	900-9997	1	$0.00	$0.00	1		
INFO QS, DATA ANALYTICS BUNDLE	379-BBWM	1	$0.00	$0.00	1		
ProSupport Plus: 3 Year Mission Critical 4HR Onsite Service	803-9323, 803-9474, 803-9483, 803-9503, 951-2015, 975-3461, 997-6306	1	$0.00	$9,820.00	1		$9,820.00
Dell Networking S3048-ON iDRAC Switch		1		$11,345.68	1	$11,345.68	
Dell Networking S3048-ON, 48x 1GbE, 4x SFP+ 10GbE ports, Stacking, IO to PSU air, 1x AC PSU, DNOS 9	210-AEDM	1		$0.00			
OS9 installed on S3048-ON, with entitlement to OS10 Enterprise	528-BBSY	1		$0.00			
Software, Rights to use L3 on OS9, S3048-ON	634-BDXE						
Force10, Power Cord, 250V, 12A, 2 Meters, C13/C14	450-AASX	1		$0.00			
Dell Networking S3048-ON User Guide	634-BCXR	1		$0.00			
ProDeploy Dell Networking S Series 3XXX	805-2398, 805-2399	1		$0.00			
Declined Remote Consulting Service	973-2426	1		$0.00			
INFO QS, DATA ANALYTICS BUNDLE	379-BBWM	1		$0.00			
3 year ProSupport 4hr response, 24x7	802-7389, 802-7394, 802-7400, 802-7403, 802-7404, 989-3439, 997-6306	1		$1,495.33	1		$1,495.33
HARDWARE COMPONENTS					**Subtotal**	$1,243,899.80	$62,014.56
SOFTWARE COMPONENTS							
Cloudera Enterprise Basic Edition, Node License, 24x7 3YR	A9081370	1	1	4,800	19	$91,200.00	
SOFTWARE COMPONENTS					**Subtotal**	$91,200.00	$0.00
Total						**$1,335,099.80**	**$62,014.56**
Large Purchase Discount (35%)*						-467,284.93	-21,705.10

Pricing:1 = Dell

(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	**$908,125**
BBQpm@10000	**1,660.75**
$/BBQpm@10000	**$ 546.82**

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

Scale Factor	10000
Streams	8
SUT Validation Test	PASS

Performance Run (Run 2)

Overall Run Start Time	2018-02-21 21:56:35.983
Overall Run End Time	2018-02-22 23:17:50.147
Overall Run Elapsed Time	91,274.164
Load Test Start Time	2018-02-21 21:56:35.983
Load Test End Time	2018-02-21 22:33:06.640
Load Test Elapsed Time	2,190.657
Power Test Start Time	2018-02-21 22:33:06.641
Power Test End Time	2018-02-22 03:48:56.072
Power Test Elapsed Time	18,949.431
Throughput Test Start Time	2018-02-22 03:48:56.073
Throughput Test End Time	2018-02-22 23:17:50.147
Throughput Test Elapsed Time	70,134.074
Performance Metric (BBQpm@ 10000)	1,660.75

Repeatability Run (Run 1)

Overall Run Start Time	2018-02-20 19:25:02.650
Overall Run End Time	2018-02-21 20:47:30.424
Overall Run Elapsed Time	91,347.774
Load Test Start Time	2018-02-20 19:25:02.651
Load Test End Time	2018-02-20 20:01:32.356
Load Test Elapsed Time	2,189.705
Power Test Start Time	2018-02-20 20:01:32.358
Power Test End Time	2018-02-21 01:16:22.830
Power Test Elapsed Time	18,890.472
Throughput Test Start Time	2018-02-21 01:16:22.831
Throughput Test End Time	2018-02-21 20:47:30.424
Throughput Test Elapsed Time	70,267.593
Performance Metric (BBQpm@ 10000)	1,660.79

Performance Run Report (Run 2)

```
************
TPCx-BB
Result
v1.2
************
INFO: T_LOAD = 2190.657
INFO: T_LD = 0.1 * T_LOAD: 219.0657
INFO: T_PT = 12863.5069618657
INFO: T_T_PUT = 70134.074
INFO: T_TT = 8766.75925
INFO: === Checking validity of the final result ===
INFO: OK: All required BigBench phases were performed.
INFO: OK: All 30 queries were running in the power test.
INFO: OK: All 30 queries were running in the first throughput test.
INFO: OK: Pretend mode was inactive. All commands were executed.
INFO: === Final result ===
INFO: VALID BBQpm@10000 = 1660.75446611642
```

Repeatability Run Report (Run 1)

```
************
TPCx-BB
Result
v1.2
************
INFO: T_LOAD = 2189.705
INFO: T_LD = 0.1 * T_LOAD: 218.9705
INFO: T_PT = 12838.6663454205
INFO: T_T_PUT = 70267.593
INFO: T_TT = 8783.449125
INFO: === Checking validity of the final result ===
INFO: OK: All required BigBench phases were performed.
INFO: OK: All 30 queries were running in the power test.
INFO: OK: All 30 queries were running in the first throughput test.
INFO: OK: Pretend mode was inactive. All commands were executed.
INFO: === Final result ===
INFO: VALID BBQpm@10000 = 1660.79427640449
```

Summary details of the run reports are shown above. For the complete run reports, see the Support Files Archive.

Table of Contents

ABSTRACT 10
PREFACE 11
CLAUSE 1: GENERAL ITEMS 12
1.1 TEST SPONSOR 12
1.2 PARAMETER SETTINGS 12
1.3 CONFIGURATION DIAGRAMS 12
CLAUSE 2: SOFTWARE COMPONENTS AND DATASET DISTRIBUTION 14
2.1 ROLES AND DATASET DISTRIBUTION 14
2.2 DISTRIBUTED FILE SYSTEM IMPLEMENTATION 14
2.3 ENGINE IMPLEMENTATION 15
2.4 FRAMEWORKS 15
2.5 APPLIED PATCHES 15
CLAUSE 3: WORKLOAD RELATED ITEMS 16
3.1 HARDWARE & SOFTWARE TUNABLE 16
3.2 KIT VERSION 16
3.3 RUN REPORT 16
3.4 QUERY ELAPSED TIMES 17
3.5 VALIDATION TEST OUTPUT 18
3.6 GLOBAL FRAMEWORK PARAMETERS 18
3.7 KIT MODIFICATIONS 19
CLAUSE 4: SUT RELATED ITEMS 20
4.1 SPECIALIZED HARDWARE/SOFTWARE 20
4.2 FRAMEWORK CONFIGURATION FILES 20
4.3 SUT ENVIRONMENT INFORMATION 20
4.4 DATA STORAGE TO SCALE FACTOR RATIO 20
4.5 SCALE FACTOR TO MEMORY RATIO 20
CLAUSE 5: METRICS AND SCALE FACTORS 21
5.1 PERFORMANCE RUN METRIC 21
5.2 REPEATABILITY RUN METRIC 21
5.3 PRICE-PERFORMANCE METRIC 21
5.4 SCALE FACTOR 21
5.5 STREAM COUNT 21
5.6 ELAPSED RUN TIMES 22
5.7 ELAPSED TEST TIMES 22
AUDITORS' INFORMATION AND ATTESTATION LETTER 23
THIRD PARTY PRICE QUOTES 26
SUPPORTING FILE INDEX 27

Abstract

This document contains the methodology and results of the TPC Express Benchmark™ Big Bench (TPCx-BB) test conducted in conformance with the requirements of the TPCx-BB Standard Specification, Revision v1.2.0.

The test was conducted at a Scale Factor of 10000 with 19 nodes (R640/R740xd) running Cloudera CDH 5.13.1 on Red Hat Enterprise Linux Server 7.3.

Measured Configuration

Company Name	Cluster Node	Virtualization	Operating System
Dell Inc.	19x R640/R740xd	n/a	Red Hat Enterprise Linux Server 7.3

TPC Express Benchmark© Big Bench Metrics

Total System Cost	BBQpm@10000	Price/Performance	Availability Date
908,125 USD	1,660.75	546.82 USD	March 20, 2018

Preface

TPC Express Benchmark™ Big Bench Overview

Big data analytics is a growing field of research and business. The significant decrease in the overall cost of hardware, the emergence of Open Source based analytics frameworks, along with the greater depth of data mining capabilities allows new types of data sources to be correlated with traditional data sources. For example, online retailers used to record only successful transactions on their website, whereas modern systems are capable of recording every interaction. The former allowed for simple shopping basket analysis techniques, while the current level of detail in monitoring makes detailed user modeling possible. The growing demands on data management systems and the new forms of analysis have led to the development of a new type of ***Big Data Analytics Systems (BDAS)****.*

Similar to the advent of ***Database Management Systems****, there is a vastly growing ecosystem of diverse approaches to enabling Big Data Analytics Systems. This leads to a dilemma for customers of* ***BDAS****, as there are no realistic and proven measures to compare different* ***BDAS*** *solutions. To address this, TPC has developed TPCx-BB (BigBench), which is an express benchmark for comparing* ***BDAS*** *solutions. The TPCx-BB Benchmark was developed to cover essential functional and business aspects of big data use cases. The benchmark allows for an objective measurement of* ***BDAS*** *System under Test, and provides the industry with verifiable performance, price/performance, and availability metrics.*

The TPCx-BB kit is available from the TPC website (see www.tpc.org for more information). Users must sign-up and agree to the TPCx-BB End User Licensing Agreement (EULA) to download the kit. All related work (such as collaterals, papers, derivatives) must acknowledge the TPC and include the TPCx-BB copyright. The TPCx-BB kit includes: TPCx-BB Specification document (this document), TPCx-BB Users Guide documentation, shell scripts to set up the benchmark environment, Java code to execute the benchmark workload, Data Generator, ***Query*** *files, and Benchmark Driver.*

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-BB models and represents a Big Data Analytics System such as Hadoop ecosystem or 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 Dell 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 components used by the SUT.*
- *Configuration parameters and options for Operating System and file system components used by the SUT.*
- *Configuration parameters and options for any other software components (e.g compiler optimization options) used by the SUT.*

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.

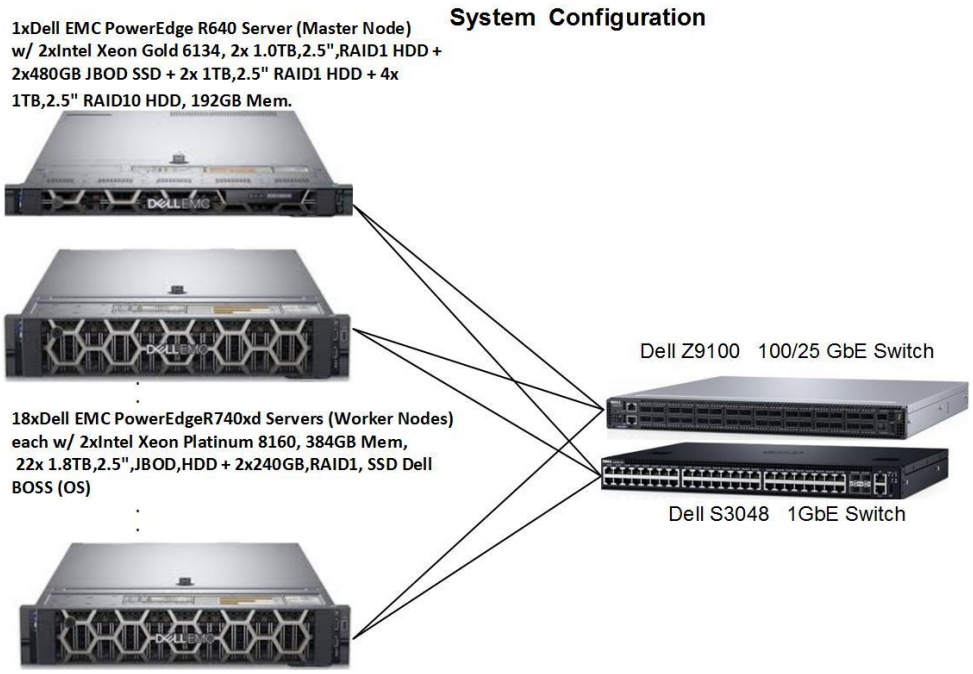
The Supporting Files Archive contains the parameters and options used to configure the components involved in this benchmark.

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: 19
- Total Processors/Cores/Threads: 38/880/1,760
- Total Memory: 7,104
- Total Number of Storage Drives/Devices: 442
- Total Storage Capacity: 730,400
- Network: Dell Z9100 100/25 GbE Switch, Dell S3048 1 GbE Switch

R640 (Master Node):

- Processors/Cores/Threads: 2/16/32
- Processor Model: 2x Intel Xeon Gold 6134 @3.20GHz
- Memory: 192GB
- Controller: PERC H730P 12Gbps RAID
- Drives:
  - 2x 480GB SSD
  - 8x 1TB HDD
- Network: Mellanox ConnectX-4 Lx DP 25GbE

18x R740xd (Worker Nodes), each with:

- Processors/Cores/Threads: 2/48/96
- Processor Model: 2x Intel Xeon Platinum 8160 @2.10GHz
- Memory: 384GB
- Controller: HBA 330+ 12Gbps SAS
- Drives:
  - 2x 240GB SSD
  - 22x 1.8TB HDD
- Network: Mellanox ConnectX-4 Lx DP 25GbE

The distribution of software components over server nodes is detailed in section 2.1.

Priced Configuration

There are no differences between the priced and measured configurations.

Clause 2: Software Components and Dataset Distribution

2.1 Roles and Dataset Distribution

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

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

Table 1.4: Software Components and Dataset Distribution

Server	Role(s)	Count	Virtual	Host Names	HW/SW Configuration	Storage Setup
Worker	HDFS Datanode HDFS Gateway HDFS NFS Gateway Hive Gateway Spark Gateway YARN/MR2 Gateway YARN Node Manager	18	N	r2xd[1-18] alias dn[1-18]	Dell PowerEdge R740xd Processor: 2x Intel Xeon Platinum 8160 Memory : 384 GB Storage: 22x 1.8TB (Data), 240GB BOSS (OS) Network: Mellanox 25GbE 2p Connectx-4lx OS: RHEL 7.3 Cloudera CDH 5.13	OS: 240GB BOSS, SAS, HDD Intermediate/ Shuffle /Temp Data Distributed FS: 22x 1.8 TB, JBOD, SAS
Master	HDFS Namenode HDFS Secondary Namenode HDFS Balancer HDFS NFS Gateway Hive Server Hive Meta Store Hive Gateway Hive Web MetHCat Cloudera Management Services Cloudera Manager Spark Gateway Spark History Server YARN/MR2 Gateway YARN JobHistory YARN Resource Manager Zookeeper	1	N	namenode1 alias nn1	Dell PowerEdge R640 Processor: 2x Intel Xeon Gold 6134 Memory : 192 GB Storage: 8x 1 TB (OS) Network: Mellanox 25GbE 2p Connectx-4lx OS: RHEL 7.3 Cloudera CDH 5.13	OS: 2x1 TB RAID1, SAS, HDD Intermediate/Temp Data: 2x480GB, JBOD,SAS,SSD Metadata: 2x1 TB RAID1, SAS, HDD PostgreSQL DB: 4x 1.0TB RAID10,SAS, HDD

2.2 Distributed File System Implementation

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

Cloudera CDH 5.13.1 (supporting Apache HDFS version 2).

2.3 Engine Implementation

The Engine implementation and corresponding version must be disclosed.

Component	Version
Hive	1.1.0
HDFS	2.6.0
YARN	2.6.0
Spark	1.6.0
MapReduce	2.6.0

2.4 Frameworks

Frameworks and Engine used in the benchmark should be disclosed.

Framework	Version
CDH	5.13.1
Hive	1.1.0
HDFS	2.6.0
YARN	2.6.0
Spark	1.6.0
MapReduce	2.6.0

2.5 Applied Patches

Any additional vendor supported patches applied to the SUT should be disclosed.

No additional patches were applied.

Clause 3: Workload Related Items

3.1 Hardware & Software Tunable

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

The Supporting Files Archive contains all configuration scripts.

3.2 Kit Version

Version number of the TPCx-BB kit must be included in the Report.

TPCx-BB Kit Version
v1.2

3.3 Run Report

The run report generated by TPCx-BB benchmark kit must be included in the Report.

The Supporting File Archive contains the full run report. Following are summary extracts from both runs.

- **Run1 Report Summary (Repeatability Run)**

  ****
  TPCx-BB
  Result
  v1.2
  ****
  INFO: T_LOAD = 2189.705
  INFO: T_LD = 0.1 * T_LOAD: 218.9705
  INFO: T_PT = 12838.6663454205
  INFO: T_T_PUT = 70267.593
  INFO: T_TT = 8783.449125
  INFO: === Checking validity of the final result ===
  INFO: OK: All required BigBench phases were performed.
  INFO: OK: All 30 queries were running in the power test.
  INFO: OK: All 30 queries were running in the first throughput test.
  INFO: OK: Pretend mode was inactive. All commands were executed.
  INFO: === Final result ===
  INFO: VALID BBQpm@10000 = 1660.79427640449

- **Run2 Report Summary (Performance Run)**

  ****
  TPCx-BB
  Result
  v1.2
  ****
  INFO: T_LOAD = 2190.657
  INFO: T_LD = 0.1 * T_LOAD: 219.0657
  INFO: T_PT = 12863.5069618657
  INFO: T_T_PUT = 70134.074
  INFO: T_TT = 8766.75925
  INFO: === Checking validity of the final result ===
  INFO: OK: All required BigBench phases were performed.
  INFO: OK: All 30 queries were running in the power test.
  INFO: OK: All 30 queries were running in the first throughput test.
  INFO: OK: Pretend mode was inactive. All commands were executed.
  INFO: === Final result ===
  INFO: VALID BBQpm@10000 = 1660.75446611642

3.4 Query Elapsed Times

Elapsed times of all power and throughput Queries needs to be reported from the Performance Run, grouped respectively as Structured, semi-structured and unstructured buckets.

Type	Query	Power	Stream 1	Stream 2	Stream 3	Stream 4	Stream 5	Stream 6	Stream 7	Stream 8
Structured	1	257.733	622.551	754.937	793.591	622.811	648.038	668.589	655.390	714.789
	6	566.061	1,922.269	1,794.328	1,680.277	1,412.759	1,704.682	1,879.070	1,890.389	1,897.916
	7	442.013	1,770.321	1,905.369	1,913.252	1,933.241	1,954.297	2,202.916	2,096.488	2,011.489
	9	435.917	1,797.316	1,965.272	1,906.117	1,789.720	443.933	1,598.892	1,613.709	1,501.877
	11	149.549	806.893	434.246	671.772	754.012	751.858	511.733	643.569	409.860
	13	191.970	984.599	843.846	770.377	1,107.896	944.387	1,062.246	420.390	983.617
	14	73.906	351.981	297.257	284.705	312.774	231.284	316.248	240.600	142.602
	15	144.493	408.409	390.898	459.378	441.321	180.879	380.587	374.316	456.146
	16	724.411	1,648.190	2,916.307	2,193.622	2,679.133	3,070.748	2,795.225	2,672.033	2,818.230
	17	389.924	837.936	959.877	809.592	834.826	751.797	875.746	794.875	793.919
	20	364.272	1,210.605	804.306	972.003	1,138.581	1,114.276	1,095.936	867.016	732.269
	21	688.362	2,488.014	2,513.538	2,997.144	3,251.408	2,264.866	2,449.230	2,797.824	2,500.101
	22	110.488	265.490	145.340	296.678	380.797	265.136	238.187	209.551	180.838
	23	178.101	660.899	424.282	374.256	365.719	375.237	422.326	251.193	229.431
	24	208.421	689.486	655.201	677.784	671.786	565.232	747.622	574.668	771.011
	25	456.199	1,584.653	2,368.969	1,842.368	1,704.727	1,429.424	1,941.115	1,493.167	1,929.750
	26	329.434	862.519	938.747	732.828	980.759	800.234	918.490	832.949	858.144
	29	250.548	710.306	841.886	605.246	943.851	1,075.095	888.490	1,015.887	801.644
Semi-structured	2	2,265.657	11,913.824	8,816.968	10,556.080	10,912.188	11,887.679	10,559.136	9,955.818	11,150.568
	3	1,201.008	4,493.970	5,389.304	3,759.932	5,520.639	5,619.587	5,411.430	4,417.545	4,415.417
	4	1,983.506	9,733.721	11,399.715	11,412.048	7,677.972	9,240.736	9,837.806	10,126.004	9,275.287
	5	671.093	2,213.113	2,532.097	2,625.120	2,377.139	2,506.385	2,420.109	2,248.190	2,772.059
	8	644.864	2,819.642	2,555.121	3,105.990	1,978.049	3,113.723	2,990.932	3,113.217	3,073.819
	12	833.449	2,191.017	1,625.266	1,927.481	1,908.976	1,854.796	2,161.179	1,942.538	2,135.450
	30	2,130.249	9,216.895	9,025.662	8,648.930	9,611.611	10,071.269	6,796.909	8,402.595	9,470.454
Unstructured	10	360.385	702.137	681.707	764.236	657.967	686.624	706.512	604.765	658.883
	18	1,450.739	3,987.235	4,083.152	3,780.977	4,167.162	3,950.097	4,177.956	4,182.840	4,128.277
	19	760.036	1,498.150	1,474.358	1,380.183	1,662.235	1,330.154	1,532.468	1,652.711	1,720.254
	27	102.114	171.406	166.861	165.520	184.307	158.209	162.544	149.126	200.269
	28	584.487	1,114.781	1,115.704	1,000.141	1,015.601	1,143.380	1,329.149	1,112.027	948.021

3.5 Validation Test Output

Output report from successful SUT Validation test must be included in the Report.

Query Number	Query Execution	Output Validation
1	PASS	PASS
2	PASS	PASS
3	PASS	PASS
4	PASS	PASS
5	PASS	PASS
6	PASS	PASS
7	PASS	PASS
8	PASS	PASS
9	PASS	PASS
10	PASS	PASS
11	PASS	PASS
12	PASS	PASS
13	PASS	PASS
14	PASS	PASS
15	PASS	PASS
16	PASS	PASS
17	PASS	PASS
18	PASS	PASS
19	PASS	PASS
20	PASS	PASS
21	PASS	PASS
22	PASS	PASS
23	PASS	PASS
24	PASS	PASS
25	PASS	PASS
26	PASS	PASS
27	PASS	PASS
28	PASS	PASS
29	PASS	PASS
30	PASS	PASS

3.6 Global Framework Parameters

Global Framework parameter settings files must be included in the Report.

The Supporting File Archive contains the global framework parameter settings files.

3.7 Kit Modifications

Test Sponsor kit modifications files must be included in the Report.

The following files were modified by the Test Sponsor to facilitate system, platform and Framework differences.

- bigBench-configs/conf/userSettings.conf
- bigBench-configs/hive/conf/engineSettings.conf
- bigBench-configs/hive/queries/q28/run.sh
- bigBench-configs/hive/queries/q[01-30]/engineLocalSettings.sql

Clause 4: SUT Related Items

4.1 Specialized Hardware/Software

Specialized Hardware/Software used in the SUT must be included.

No specialized hardware or software was used.

4.2 Framework Configuration Files

All Framework configuration files from SUT, for the performance run.

All Framework configuration files are included in the Supporting Files Archive.

4.3 SUT Environment Information

SUT environment info in form of envinfo.log from a representative worker node form every role in the server.

All envinfo.log files are include in the Supporting Files Archive.

4.4 Data Storage to Scale Factor Ratio

The data storage ratio must be disclosed.

Nodes	Disks	Size (GB)	Total (GB)
1	2	480	960
1	8	1,000	8,000
18	2	240	8,640
18	22	1,800	712,800

Total Storage (GB)	730,400
Scale Factor	10000
Data Storage Ratio	73.04

4.5 Scale Factor to Memory Ratio

The Scale Factor to memory ratio must be disclosed.

Nodes	Memory (GB)	Total (GB)
1	192	192
18	384	6,912

Scale Factor	10000
Total Memory (GB)	7,104
SF / Memory Ratio	1.41

Clause 5: Metrics and Scale Factors

5.1 Performance Run Metric

The Reported Performance Metric (BBQpm@SF for the Performance Run) must be disclosed in the Report.

Performance Run
BBQpm@10000 1,660.75

5.2 Repeatability Run Metric

The Performance Metric (BBQpm@SF) for the Repeatability Run must be disclosed in the Report..

Repeatability Run
BBQpm@10000 1,660.79

5.3 Price-Performance Metric

The Reported Performance Metric (BBQpm@SF for the Performance Run) must be disclosed in the Report.

Price / Performance
$BBQpm@10000 546.82

5.4 Scale Factor

The Scale Factor used for the Result must be disclosed in the Report.

Scale Factor
10000

5.5 Stream Count

The number of streams in the throughput run used for the Result must be disclosed in the Report.

Streams
8

5.6 Elapsed Run Times

The total elapsed time for the execution of the Performance Run and Repeatability Run must be disclosed in the Report.

Run	Elapsed Time	Seconds
Run 1	01 01:22:27.774	91,347.774
Run 2	01 01:21:14.164	91,274.164

5.7 Elapsed Test Times

The total time for each of the three tests must be disclosed for the Performance Run and the Repeatability Run.

Test	Performance Run	Repeatability Run
Load Test	2,190.657	2,189.705
Power Test	18,949.431	18,890.472
Throughput Test	70,134.074	70,267.593

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 01453
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
Certified Auditor

Mr. Nicholas Wakou
Dell Inc.
701 E. Parmer Ln. Bldg. 2
Austin, TX 78753

March 20, 2018

I verified the TPC Express Benchmark™ BB v1.2.0 performance of the following configuration:

Platform:	Dell PowerEdge 14G R640/R740xd (with 19x R640/R740xd)
Operating System:	Red Hat Enterprise Linux Server 7.3
Apache Hadoop Compatible Software:	Cloudera for Apache Hadoop (CDH) 5.13.1

The results were:

Performance Metric	**1,660.75 BBQpm@10000GB**
Run Elapsed Time	01 01:21:14.164 (91,724.164 Seconds)

Cluster	**1x R640 (Master Node), 18x R740xd (Worker Nodes)**		
CPUs	2 x Intel Xeon Gold 6134 (3.20 GHz, 8-core, 24.75 MB L3) (Master Node) 2 x Intel Xeon Platinum 8160 (2.10 GHz, 24-core, 33 MB L3) (Worker Nodes)		
Memory	192GB (Master Node); 384GB (Worker Nodes)		
Storage	***Qty***	***Size***	***Type***
	2	480 GB	SSD (Master Node)
	8	1 TB	HDD (Master Node)
	2	240 GB	SSD (Worker Nodes)
	22	1.8 TB	HDD (Worker 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 v1.2.0
- No modifications were made to any of the Java code
- Any and all modifications to shell scripts were reviewed for compliance
- The tested Scale Factor (10000GB) was confirmed to be valid for publication

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

- All validation queries executed successfully and produced compliant results
- No errors were reported during the run
- 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:

From the TPCx-BB Kit's README:

Q28 Depending on the Hadoop distribution version can fail automated Engine Validation due to empty space characters when the output is written to HDFS. Manually open the result file and validate the reference values and written values.

Query 28 failed automated Engine Validation. A manual validation was performed as part of this audit to confirm the only differences were due to white space.

Respectfully Yours,

Doug Johnson, TPC Auditor

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

Third Party Price Quotes

None.

Supporting File Index

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

Description	Archive File Pathname
Clause 1 - General Items	
The Supporting Files Archive contains the parameters and options used to configure the components involved in this benchmark	Supporting-Files-10TB-14G-xBB-03-2018\
Validation Run Files	Supporting-Files-10TB-14G-xBB-03-2018\
Performance Run Files	Supporting-Files-10TB-14G-xBB-03-2018\Performance-run-logs
Repeatability Run Files	Supporting-Files-10TB-14G-xBB-03-2018\Repeatability-run-logs
Clause 3 - Workload Related Items	
Benchmark Generic Parameters	Supporting-Files-10TB-14G-xBB-03-2018\ Performance-run-logs\bigBench-configs\conf\userSettings.conf
Query Parameters used in the benchmark execution Settings	Supporting-Files-10TB-14G-xBB-03-2018\ Performance-run-logs\bigBench-configs\hive\conf\queryParameters.sql
Benchmark Global Framework Parameters Settings	Supporting-Files-10TB-14G-xBB-03-2018\ Performance-run-logs\bigBench-configs\hive\conf\engineSettings.sql
Benchmark Global Framework Parameters Settings	Supporting-Files-10TB-14G-xBB-03-2018\ Performance-run-logs\bigBench-configs\hive\conf\engineSettings.conf
Load Test script	Supporting-Files-10TB-14G-xBB-03-2018\ Performance-run-logs\bigBench-configs \hive\population\ hiveCreateLoad.sql
Queries specific optimization parameters settings	Supporting-Files-10TB-14G-xBB-03-2018\ Performance-run-logs\bigBench-configs\hive\queries\q[01-30]\engineLocalSettings.conf
Queries specific optimization parameters settings	Supporting-Files-10TB-14G-xBB-03-2018\ Performance-run-logs\bigBench-configs\hive\queries\q[01-30]\engineLocalSettings.sql
Clause 4 - SUT Related Items	
Data Redundancy report	Supporting-Files-10TB-14G-xBB-03-2018\hdfs-data-redundancy-report.txt
Benchmark execution script	Supporting-Files-10TB-14G-xBB-03-2018\TPCxBB_Benchmarkrun.sh
Hardware and Software Report from a representative node	Supporting-Files-10TB-14G-xBB-03-2018\ Performance-run-logs\run- logs\envInfo-r2xd1.ignition.dell.com\envInfo.log
All Framework configuration files are included in the Supporting Files Archive	Supporting-Files-10TB-14G-xBB-03-2018\Performance-run-logs\run-logs\envInfo-r2xd1.ignition.dell.com\hadoop
	Supporting-Files-10TB-14G-xBB-03-2018\Performance-run-logs\run-logs\envInfo-r2xd1.ignition.dell.com\hive
	Supporting-Files-10TB-14G-xBB-03-2018\Performance-run-logs\run-logs\envInfo-r2xd1.ignition.dell.com\spark
Clause 5 - Metric and Scale Factor Related Items	
Benchmark Performance Report	Supporting-Files-10TB-14G-xBB-03-2018\Performance-run-logs\run-logs\BigBenchResult.log
Validation Test Report	Supporting-Files-10TB-14G-xBB-03-2018\Validation-run-logs\run-logs\BigBenchResult.log