Alibaba Cloud Computing Ltd.

TPC Benchmark™ DS

Full Disclosure Report

for

Alibaba Cloud E-MapReduce Serverless Spark

(with 2112 E-MapReduce Serverless Spark CU)

using

E-MapReduce Serverless Spark esr-4.5.1

and

Alibaba Cloud Linux 3.2104 U11 (OpenAnolis Edition)

First Edition

September 1, 2025

First Edition – September 1, 2025

Table of Contents

Abstract

This document contains the methodology and results of the TPC Benchmark™ DS (TPC-DS) test conducted in conformance with the requirements of the TPC-DS Standard Specification, Revision 4.0.0.

The test was conducted at a Scale Factor of 100000GB with 2112 CUs running Alibaba Cloud E-MapReduce Serverless Spark esr-4.5.1 on Alibaba Cloud Linux 3.2104 U11 (OpenAnolis Edition).

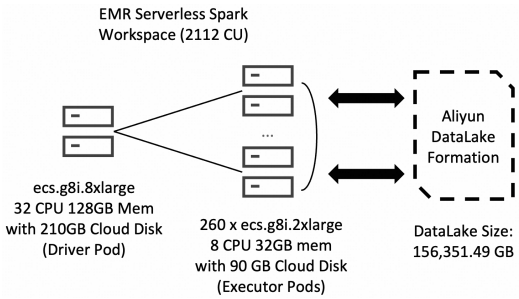
Measured Configuration

Company Name	Cluster Node	Database Software	Operation System
Alibaba Cloud Computing Ltd.	Alibaba Cloud E-MapReduce Serverless Spark CU	Alibaba Cloud E-MapReduce Serverless Spark esr-4.5.1	Alibaba Cloud Linux 3.2104 U11 (OpenAnolis Edition)

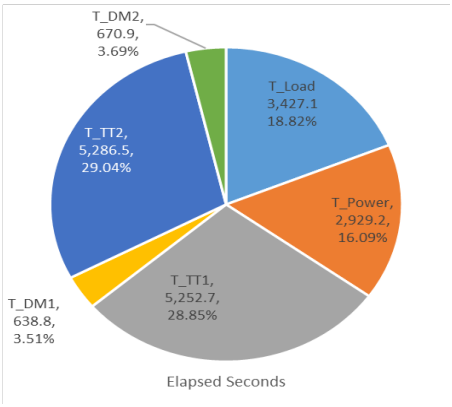
TPC Benchmark™ DS Metrics

Total System Cost (CNY)	TPC-DS Throughput (QphDS@100000)	Price/Performance (CNY / kQphDS@100000)	Availability Date
¥4,081,477	**65,684,477**	**¥62.14**	**2025-10-20**

Alibaba Cloud	**Alibaba Cloud E-MapReduce Serverless Spark**		TPC-DS: 4.0.0 TPC-Pricing: 2.9.0 Report Date: Sep. 1, 2025	
Total System Cost	TPC-DS Throughput	Price / Performance	System Availability Date	
¥4,081,477 CNY	**65,684,477** QphDS@100000	**¥62.14** CNY/kQphDS@100000	**2025-10-20**	
Dataset Size[1]	Database Manager	Operation System	Other Software	Cluster
100,000 GB	E-MapReduce Serverless Spark esr-4.5.1	Alibaba Cloud Linux 3.2104 U11 (OpenAnolis Edition)	N/A	Yes



Benchmarked Configuration



Elapsed Time

Load includes backup = No	RAID = No

System Configuration:	**Alibaba Cloud E-MapReduce Serverless Spark**
Servers:	1 x ecs.g8i.8xlarge + 260 x ecs.g8i.2xlarge
Total Processors/Cores/Threads:	2,112 vCPUs (threads)
Total Memory:	8,448 GB
Total Storage[2]:	179,961.49 GB
Storage Ratio[3]:	1.8
Server Configuration:	**Per Driver Pod (1)**
Processors:	Intel(R) Xeon(R) Platinum 8575C
Memory:	128 GB
Network:	Bandwidth: 12 Gbps, Packet forwarding rate: 3,000,000
Storage Device:	1 x 210 GB SSD Cloud Disk
Server Configuration:	**Per Executor Pod (260)**
Processors:	220 Intel(R) Xeon(R) Platinum 8575C 40 Intel(R) Xeon(R) Platinum 8475B
Memory:	32 GB
Network:	Bandwidth: 6 Gbps, Packet forwarding rate: 1,600,000
Storage Device:	1 x 90 GB SSD Cloud Disk

1. Dataset Size includes only raw data (i.e., no temp, index, redundant storage space, etc.).
2. Total Storage = 210 (Driver pod) + 90 * 260 (Executor pods) + 156,351.49 (Aliyun DataLake Formation) = 179,961.49 GB
3. Storage Ratio = Total Storage / SF = 179,961.49 GB / 100,000 GB = 1.8

Alibaba Cloud	**Alibaba Cloud E-MapReduce Serverless Spark**	TPC-DS: 4.0.0 TPC-Pricing: 2.9.0 Report Date: Sep. 1, 2025

Description	Part Number	Src	Unit Price (CNY)	Qty	Ext. Price (CNY)	3-Year Maint. (CNY)
Licensed Compute and Software Services						
EMR Serverless Spark						
2112 CU, one month pre-pay	(China North 6)	1	295,680	12	3,548,160	included
Aliyun DataLake Formation						
Metadata Management	(China North 6)	1	0	12	0	included
Metadata API Request	(China North 6)	1	43.62	12	523.44	Included
Data Storage	(China North 6)	1	23,452.73	12	281,432.76	Included
Throughput	(China North 6)	1	20,421.96	12	245,063.52	Included
	Licensed Compute and Software Services Sub-Total				**4,075,179.72**	**0.00**
Other Components						
Lenovo P14H Laptop (includes 2 spares)		2	2,099.00	3	6,297.00	
	Other Components Sub-Total				**6,297.00**	**0.00**

1 = Alibaba Cloud, 2 = Tmall.com	**1-Year Cost of Ownership:**	**4,081,477**
All Licensed Services prices are per month and based on 1-year pre-paid subscriptions.	**QphDS@100000:**	**65,684,477**
Audited by Doug Johnson, InfoSizing	**¥/kQphDS@100000:**	**62.14**

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.

Alibaba Cloud	Alibaba Cloud E-MapReduce	TPC-DS: 4.0.0 TPC-Pricing: 2.9.0 Report Date: Sep. 1, 2020

Metrics Details:

Name	Value	Description / Unit
SF	100,000	Scale Factor 100TB
S	8	Total Throughput Streams
Sq	4	Streams / Throughput Test
Q	396	Queries / Throughput Test
T_LD	0.0381	hours@100000
T_PT	3.2547	hours@100000
T_TT	2.9276	hours@100000
T_DM	0.3639	hours@100000

Secondary Details:

Name	Value	Unit
T_Load	3,427.1	seconds@100000
T_Power	2,929.2	seconds@100000
T_TT1	5,252.7	seconds@100000
T_TT2	5,286.5	seconds@100000
T_DM1	638.8	seconds@100000
T_DM2	670.9	seconds@100000

Test Timeline:

Test	Start	End	Seconds	(hh:mm:ss)
Load	2025-08-23 13:40:58.574	2025-08-23 14:38:06.000	3,427.055	00:57:07
Audit/Admin	2025-08-23 14:38:05.632	2025-08-23 15:47:02.908	4,137.276	01:08:57
Power	2025-08-23 15:47:02.997	2025-08-23 16:35:52.172	2,929.175	00:48:49
TT-1	2025-08-23 16:35:52.269	2025-08-23 18:03:24.949	5,252.680	01:27:33
DM-1	2025-08-23 18:03:24.968	2025-08-23 18:14:03.730	638.762	00:10:39
TT-2	2025-08-23 18:14:03.828	2025-08-23 19:42:10.149	5,286.321	01:28:06
DM-2	2025-08-23 19:42:10.168	2025-08-23 19:53:20.995	670.827	00:11:11

Stream	Start	End	Seconds	(hh:mm:ss)
Power - 0	2025-08-23 15:47:02.997	2025-08-23 16:35:52.000	2929.175	00:48:49
TT-1 - 1	2025-08-23 16:35:52.000	2025-08-23 17:58:50.289	4,978.020	01:22:58
TT-1 - 2	2025-08-23 16:35:52.270	2025-08-23 18:01:40.598	5,148.328	01:25:48
TT-1 - 3	2025-08-23 16:35:52.269	2025-08-23 18:03:24.949	5,252.680	01:27:33
TT-1 - 4	2025-08-23 16:35:52.270	2025-08-23 18:01:24.145	5,131.875	01:25:32
DM-1 - 1	2025-08-23 18:03:24.968	2025-08-23 18:08:35.099	310.131	00:05:10
DM-1 - 2	2025-08-23 18:08:35.112	2025-08-23 18:14:03.730	328.618	00:05:29
TT-2 - 5	2025-08-23 18:14:03.828	2025-08-23 19:42:08.241	5,284.413	01:28:04
TT-2 - 6	2025-08-23 18:14:03.828	2025-08-23 19:41:50.850	5,267.022	01:27:47
TT-2 - 7	2025-08-23 18:14:03.828	2025-08-23 19:42:10.149	5,286.321	01:28:06
TT-2 - 8	2025-08-23 18:14:03.828	2025-08-23 19:41:14.130	5,230.302	01:27:10
DM-2 - 3	2025-08-23 19:42:10.168	2025-08-23 19:47:39.328	329.160	00:05:29
DM-2 - 4	2025-08-23 19:47:39.342	2025-08-23 19:53:20.995	341.653	00:05:42

Timing Intervals for Each Query (In Seconds)

Query	Stream 0	Stream 1	Stream 2	Stream 3	Stream 4	Min	25th	Median	75th	Max	Stream 5	Stream 6	Stream 7	Stream 8	Min	25th	Median	75th	Max
1	17.7	31.1	18.5	23.8	24.8	18.50	22.48	24.30	26.38	31.10	13.9	30.0	20.9	18.4	13.90	17.28	19.65	23.18	30.00
2	24.2	63.3	53.3	65.7	9.4	9.40	42.33	58.30	63.90	65.70	52.6	88.9	94.1	8.0	8.00	41.45	70.75	90.20	94.10
3	11.4	15.0	14.8	8.8	11.2	8.80	10.60	13.00	14.85	15.00	15.8	8.8	10.1	9.6	8.80	9.40	9.85	11.53	15.80
4	61.0	74.8	80.2	83.8	96.4	74.80	78.85	82.00	86.95	96.40	80.3	81.6	109.7	83.5	80.30	81.28	82.55	90.05	109.70
5	34.1	83.3	50.0	40.4	51.2	40.40	47.60	50.60	59.23	83.30	74.0	73.9	39.3	103.1	39.30	65.25	73.95	81.28	103.10
6	12.1	17.8	15.4	16.6	11.4	11.40	14.40	16.00	16.90	17.80	21.4	24.5	20.8	17.5	17.50	19.98	21.10	22.18	24.50
7	26.0	25.9	19.8	30.0	38.1	19.80	24.38	27.95	32.03	38.10	29.4	21.7	22.5	26.9	21.70	22.30	24.70	27.53	29.40
8	13.8	16.0	12.4	13.6	13.7	12.40	13.30	13.65	14.28	16.00	14.4	13.7	15.6	19.4	13.70	14.23	15.00	16.55	19.40
9	35.3	78.7	67.6	89.0	49.7	49.70	63.13	73.15	81.28	89.00	72.7	58.9	127.7	105.9	58.90	69.25	89.30	111.35	127.70
10	18.2	26.0	36.6	31.4	26.6	26.00	26.45	29.00	32.70	36.60	31.7	34.7	31.2	21.2	21.20	28.70	31.45	32.45	34.70
11	35.0	52.0	65.3	53.2	48.7	48.70	51.18	52.60	56.23	65.30	58.1	58.2	60.6	62.7	58.10	58.18	59.40	61.13	62.70
12	8.4	10.2	7.3	7.3	14.0	7.30	7.30	8.75	11.15	14.00	9.0	9.4	9.0	8.6	8.60	8.90	9.00	9.10	9.40
13	21.1	28.2	34.0	30.4	37.7	28.20	29.85	32.20	34.93	37.70	32.8	36.1	24.0	40.0	24.00	30.60	34.45	37.08	40.00
14	84.8	32.8	129.5	240.5	190.0	32.80	105.33	159.75	202.63	240.50	188.1	243.2	90.9	141.1	90.90	128.55	164.60	201.88	243.20
15	14.8	14.6	27.0	20.0	20.2	14.60	18.65	20.10	21.90	27.00	19.5	19.3	21.2	19.7	19.30	19.45	19.60	20.08	21.20
16	7.9	144.6	7.7	150.0	229.2	7.70	110.38	147.30	169.80	229.20	95.4	117.6	8.6	106.1	8.60	73.70	100.75	108.98	117.60
17	22.9	45.5	37.4	26.4	38.3	26.40	34.65	37.85	40.10	45.50	38.8	37.9	37.6	36.9	36.90	37.43	37.75	38.13	38.80
18	21.8	39.1	38.4	35.2	24.7	24.70	32.58	36.80	38.58	39.10	26.2	35.7	32.9	45.3	26.20	31.23	34.30	38.10	45.30
19	13.6	16.1	16.0	17.3	26.9	16.00	16.08	16.70	19.70	26.90	31.7	15.9	19.2	16.3	15.90	16.20	17.75	22.33	31.70
20	8.9	9.7	8.2	11.2	8.3	8.20	8.28	9.00	10.08	11.20	8.6	7.8	9.4	10.1	7.80	8.40	9.00	9.58	10.10
21	6.9	7.6	6.9	5.3	11.0	5.30	6.50	7.25	8.45	11.00	8.7	8.9	16.8	6.4	6.40	8.13	8.80	10.88	16.80
22	13.1	8.3	11.8	9.9	11.3	8.30	9.50	10.60	11.43	11.80	12.1	8.3	10.9	9.0	8.30	8.83	9.95	11.20	12.10
23	108.5	192.3	206.3	150.4	215.8	150.40	181.83	199.30	208.68	215.80	313.0	228.3	267.4	388.8	228.30	257.63	290.20	331.95	388.80
24	82.7	181.1	216.0	244.0	156.6	156.60	174.98	198.55	223.00	244.00	185.4	238.9	158.9	167.4	158.90	165.28	176.40	198.78	238.90
25	17.9	32.8	28.3	29.3	36.5	28.30	29.05	31.05	33.73	36.50	20.9	33.2	31.4	27.4	20.90	25.78	29.40	31.85	33.20
26	15.6	29.8	18.3	24.0	27.9	18.30	22.58	25.95	28.38	29.80	24.0	25.1	66.4	22.4	22.40	23.60	24.55	35.43	66.40
27	17.6	23.0	22.6	26.8	39.5	22.60	22.90	24.90	29.98	39.50	25.2	31.1	19.3	18.4	18.40	19.08	22.25	26.68	31.10
28	60.8	127.1	122.6	182.8	168.6	122.60	125.98	147.85	172.15	182.80	125.0	96.6	124.7	134.6	96.60	117.68	124.85	127.40	134.60
29	56.2	68.4	77.9	55.0	84.5	55.00	65.05	73.15	79.55	84.50	71.4	73.7	68.9	84.8	68.90	70.78	72.55	76.48	84.80
30	12.3	16.1	19.6	24.0	18.5	16.10	17.90	19.05	20.70	24.00	18.4	25.1	17.7	9.5	9.50	15.65	18.05	20.08	25.10
31	20.2	24.8	46.7	30.8	28.3	24.80	27.43	29.55	34.78	46.70	25.3	34.5	29.3	32.5	25.30	28.30	30.90	33.00	34.50
32	7.9	9.0	13.6	9.5	13.7	9.00	9.38	11.55	13.63	13.70	11.5	9.3	12.5	14.0	9.30	10.95	12.00	12.88	14.00
33	19.9	18.8	20.7	20.4	13.2	13.20	17.40	19.60	20.48	20.70	26.1	19.9	19.3	18.2	18.20	19.03	19.60	21.45	26.10
34	28.9	35.9	26.9	38.9	28.1	26.90	27.80	32.00	36.65	38.90	28.1	31.6	31.3	39.0	28.10	30.50	31.45	33.45	39.00
35	28.6	46.3	52.7	25.7	46.6	25.70	41.15	46.45	48.13	52.70	46.4	48.4	37.6	55.8	37.60	44.20	47.40	50.25	55.80
36	21.9	26.0	29.9	33.3	33.0	26.00	28.93	31.45	33.08	33.30	22.8	25.9	30.2	31.0	22.80	25.13	28.05	30.40	31.00
37	22.1	36.0	66.0	43.9	55.7	36.00	41.93	49.80	58.28	66.00	51.3	28.4	75.2	62.7	28.40	45.58	57.00	65.83	75.20
38	55.9	75.3	97.4	105.1	96.0	75.30	90.83	96.70	99.33	105.10	97.9	92.5	75.7	86.2	75.70	83.58	89.35	93.85	97.90
39	39.5	42.5	45.9	50.0	48.1	42.50	45.05	47.00	48.58	50.00	30.2	44.8	39.2	42.9	30.20	36.95	41.05	43.38	44.80
40	17.9	59.2	49.0	33.5	20.5	20.50	30.25	41.25	51.55	59.20	49.6	23.5	19.3	36.7	19.30	22.45	30.10	39.93	49.60
41	4.4	6.9	3.2	3.5	3.3	3.20	3.28	3.40	4.35	6.90	4.4	3.4	3.0	3.5	3.00	3.30	3.45	3.73	4.40
42	10.1	8.3	11.8	7.9	8.2	7.90	8.13	8.25	9.18	11.80	8.6	9.8	9.1	9.0	8.60	8.90	9.05	9.28	9.80
43	10.7	21.1	10.6	26.2	15.7	10.60	14.43	18.40	22.38	26.20	21.1	16.8	11.8	21.3	11.80	15.55	18.95	21.15	21.30
44	20.5	34.0	96.3	85.5	26.1	26.10	32.03	59.75	88.20	96.30	24.0	51.4	49.8	39.7	24.00	35.78	44.75	50.20	51.40
45	17.2	19.9	22.9	17.4	21.2	17.40	19.28	20.55	21.63	22.90	19.8	13.7	15.5	19.6	13.70	15.05	17.55	19.65	19.80
46	18.6	23.8	31.2	36.9	26.3	23.80	25.68	28.75	32.63	36.90	39.2	32.0	29.1	29.7	29.10	29.55	30.85	33.80	39.20
47	28.0	9.8	56.4	57.6	40.3	9.80	32.68	48.35	56.70	57.60	77.5	40.6	55.7	48.2	40.60	46.30	51.95	61.15	77.50
48	20.7	22.8	18.6	25.7	31.3	18.60	21.75	24.25	27.10	31.30	31.3	38.1	50.2	34.5	31.30	33.70	36.30	41.13	50.20
49	23.3	32.1	27.3	19.6	20.5	19.60	20.28	23.90	28.50	32.10	31.7	28.1	25.3	38.1	25.30	27.40	29.90	33.30	38.10
50	82.6	151.3	253.6	210.1	200.5	151.30	188.20	205.30	220.98	253.60	125.0	187.6	173.2	141.6	125.00	137.45	157.40	176.80	187.60
51	33.4	43.7	52.7	33.9	65.2	33.90	41.25	48.20	55.83	65.20	36.7	54.4	61.3	55.6	36.70	49.98	55.00	57.03	61.30
52	6.6	14.9	8.0	9.1	6.3	6.30	7.58	8.55	10.55	14.90	8.2	5.9	11.2	13.3	5.90	7.63	9.70	11.73	13.30
53	14.3	24.6	17.5	21.2	13.5	13.50	16.50	19.35	22.05	24.60	12.1	24.1	13.8	30.4	12.10	13.38	18.95	25.68	30.40
54	17.2	20.5	25.2	21.7	27.6	20.50	21.40	23.45	25.80	27.60	29.3	19.5	21.5	17.0	17.00	18.88	20.50	23.45	29.30
55	5.5	10.3	7.4	9.5	21.6	7.40	8.98	9.90	13.13	21.60	6.0	20.2	16.1	5.7	5.70	5.93	11.05	17.13	20.20
56	16.8	17.8	23.5	23.1	25.8	17.80	21.78	23.30	24.08	25.80	21.0	22.1	19.5	22.2	19.50	20.63	21.55	22.13	22.20
57	26.5	49.1	42.2	43.1	51.9	42.20	42.88	46.10	49.80	51.90	49.2	64.8	8.2	29.6	8.20	24.25	39.40	53.10	64.80
58	15.5	18.5	16.6	19.2	13.8	13.80	15.90	17.55	18.68	19.20	21.8	12.6	16.8	14.6	12.60	14.10	15.70	18.05	21.80
59	32.2	58.9	39.0	73.1	7.3	7.30	31.08	48.95	62.45	73.10	10.4	83.0	107.7	43.9	10.40	35.53	63.45	89.18	107.70
60	17.8	27.4	20.0	22.0	18.8	18.80	19.70	21.00	23.35	27.40	29.1	32.1	34.7	22.3	22.30	27.40	30.60	32.75	34.70
61	15.0	18.4	26.7	23.6	19.2	18.40	19.00	21.40	24.38	26.70	26.8	18.7	22.9	23.0	18.70	21.85	22.95	23.95	26.80
62	21.1	42.2	20.0	19.0	16.9	16.90	18.48	19.50	25.55	42.20	19.0	27.4	24.5	26.8	19.00	23.13	25.65	26.95	27.40
63	13.0	16.6	10.5	13.8	21.4	10.50	12.98	15.20	17.80	21.40	16.9	26.2	17.5	18.2	16.90	17.35	17.85	20.20	26.20
64	113.7	265.1	265.5	208.3	249.8	208.30	239.43	257.45	265.20	265.50	279.7	204.4	321.1	224.1	204.40	219.18	251.90	290.05	321.10
65	34.4	62.5	62.8	54.0	62.3	54.00	60.23	62.40	62.58	62.80	66.0	54.6	72.8	60.7	54.60	59.18	63.35	67.70	72.80
66	19.7	24.0	24.0	21.9	27.0	21.90	23.48	24.00	24.75	27.00	17.7	25.6	23.0	27.4	17.70	21.68	24.30	26.05	27.40
67	70.6	169.6	131.8	213.2	127.1	127.10	130.63	150.70	180.50	213.20	180.2	205.5	191.2	143.4	143.40	171.00	185.70	194.78	205.50
68	15.4	15.3	14.6	11.8	17.3	11.80	13.90	14.95	15.80	17.30	22.6	17.4	14.0	16.2	14.00	15.65	16.80	18.70	22.60
69	16.3	25.6	20.7	34.0	23.3	20.70	22.65	24.45	27.70	34.00	14.2	25.1	25.1	6.7	6.70	12.33	19.65	25.10	25.10
70	21.1	28.2	20.9	34.0	31.8	20.90	26.38	30.00	32.35	34.00	38.6	23.1	27.4	27.3	23.10	26.25	27.35	30.20	38.60
71	26.2	23.7	23.4	23.5	35.1	23.40	23.48	23.60	26.55	35.10	37.2	38.4	30.5	30.8	30.50	30.73	34.00	37.50	38.40
72	56.5	124.5	89.5	101.8	112.8	89.50	98.73	107.30	115.73	124.50	122.3	103.3	121.9	116.9	103.30	113.50	119.40	122.00	122.30
73	16.0	19.5	22.8	13.7	18.8	13.70	17.53	19.15	20.33	22.80	21.8	16.8	13.2	26.8	13.20	15.90	19.30	23.05	26.80
74	53.1	52.0	46.5	52.7	42.4	42.40	45.48	49.25	52.18	52.70	44.1	100.9	54.2	95.1	44.10	51.68	74.65	96.55	100.90
75	70.9	133.2	224.6	162.7	138.0	133.20	136.80	150.35	178.18	224.60	139.9	108.2	171.1	121.6	108.20	118.25	130.75	147.70	171.10
76	46.0	87.0	75.2	115.6	86.3	75.20	83.53	86.65	94.15	115.60	116.4	89.3	88.3	103.9	88.30	89.05	96.60	107.03	116.40
77	18.7	25.5	19.6	17.4	17.9	17.40	17.78	18.75	21.08	25.50	22.5	26.3	27.1	17.9	17.90	21.35	24.40	26.50	27.10
78	108.0	288.1	254.3	167.9	307.4	167.90	232.70	271.20	292.93	307.40	253.9	209.2	228.1	228.9	209.20	223.38	228.50	235.15	253.90
79	15.9	38.6	28.3	17.1	21.3	17.10	20.25	24.80	30.88	38.60	33.0	37.0	24.9	25.0	24.90	24.98	29.00	34.00	37.00
80	45.2	89.6	78.4	103.2	71.9	71.90	76.78	84.00	93.00	103.20	104.8	80.7	75.7	80.7	75.70	79.45	80.70	86.73	104.80
81	15.2	20.1	12.8	24.7	23.2	12.80	18.28	21.65	23.58	24.70	14.6	21.3	18.3	14.6	14.60	14.60	16.45	19.05	21.30
82	33.2	124.6	74.2	81.4	56.6	56.60	69.80	77.80	92.20	124.60	91.5	60.9	95.0	80.9	60.90	75.90	86.20	92.38	95.00
83	14.4	15.1	16.9	17.8	15.3	15.10	15.25	16.10	17.13	17.80	10.5	14.4	14.8	18.1	10.50	13.43	14.60	15.63	18.10
84	12.3	20.1	14.8	13.1	13.8	13.10	13.63	14.30	16.13	20.10	17.5	28.2	32.0	22.3	17.50	21.10	25.25	29.15	32.00
85	24.5	27.3	34.7	26.5	21.4	21.40	25.23	26.90	29.15	34.70	21.2	33.0	35.5	26.8	21.20	25.40	29.90	33.63	35.50
86	13.7	13.2	20.3	19.3	11.8	11.80	12.85	16.25	19.55	20.30	16.9	17.1	22.1	10.1	10.10	15.20	17.00	18.35	22.10
87	51.4	113.4	98.4	85.0	73.7	73.70	82.18	91.70	102.15	113.40	82.2	100.6	103.4	89.5	82.20	87.68	95.05	101.30	103.40
88	17.0	36.6	38.2	13.9	31.3	13.90	26.95	33.95	37.00	38.20	38.9	36.6	34.6	34.2	34.20	34.50	35.60	37.18	38.90
89	14.2	28.8	22.5	16.3	27.3	16.30	20.95	24.90	27.68	28.80	19.5	25.5	18.5	16.8	16.80	18.08	19.00	21.00	25.50
90	18.0	17.7	23.6	21.7	28.9	17.70	20.70	22.65	24.93	28.90	18.5	17.6	21.3	22.4	17.60	18.28	19.90	21.58	22.40
91	16.9	24.9	17.7	17.4	22.0	17.40	17.63	19.85	22.73	24.90	16.3	20.2	18.5	18.5	16.30	17.95	18.50	18.93	20.20
92	11.8	6.2	8.2	8.8	8.2	6.20	7.70	8.20	8.35	8.80	13.3	8.3	8.0	8.5	8.00	8.23	8.40	9.70	13.30

93	111.3	174.1	260.2	222.0	229.5	174.10	210.03	225.75	237.18	260.20	262.3	181.3	201.0	287.8	181.30	196.08	231.65	268.68	287.80
94	39.2	61.7	71.1	70.8	67.9	61.70	66.35	69.35	70.88	71.10	71.0	69.3	69.5	48.5	48.50	64.10	69.40	69.88	71.00
95	85.5	110.8	152.1	178.5	102.5	102.50	108.73	131.45	158.70	178.50	171.5	182.3	162.1	175.4	162.10	169.15	173.45	177.13	182.30
96	25.7	26.6	18.4	28.3	47.1	18.40	24.55	27.45	33.00	47.10	17.4	33.6	33.4	35.6	17.40	29.40	33.50	34.10	35.60
97	27.0	44.5	49.9	48.8	86.2	44.50	47.73	49.35	58.98	86.20	65.3	39.8	58.3	37.2	37.20	39.15	49.05	60.05	65.30
98	17.6	23.0	21.2	16.0	21.8	16.00	19.90	21.50	22.10	23.00	20.4	17.9	27.3	14.1	14.10	16.95	19.15	22.13	27.30
99	19.0	44.2	35.2	34.3	30.7	30.70	33.40	34.75	37.45	44.20	28.1	19.5	52.0	24.5	19.50	23.25	26.30	34.08	52.00

Timing Intervals for Each Refresh Function (In Seconds)

Function	DM-1		DM-2						
ID	R-Run 1	R-Run 2	R-Run 3	R-Run 4	Min	25th	Median	75th	Max
DF_CS	121.5	134.4	138.4	125.4	121.50	124.43	129.90	135.40	138.40
DF_I	9.7	7.6	19.7	8.9	7.60	8.58	9.30	12.20	19.70
DF_SS	210.4	222.5	232.2	237.3	210.40	219.48	227.35	233.48	237.30
DF_WS	111.6	135.0	140.1	127.6	111.60	123.60	131.30	136.28	140.10
LF_CR	26.4	37.2	28.6	29.8	26.40	28.05	29.20	31.65	37.20
LF_CS	65.1	69.0	64.6	65.6	64.60	64.98	65.35	66.45	69.00
LF_I	24.0	17.1	22.9	24.8	17.10	21.45	23.45	24.20	24.80
LF_SR	28.4	25.5	27.9	29.1	25.50	27.30	28.15	28.58	29.10
LF_SS	71.5	51.5	69.2	75.4	51.50	64.78	70.35	72.48	75.40
LF_WR	25.9	26.2	24.9	27.9	24.90	25.65	26.05	26.63	27.90
LF_WS	58.7	65.2	56.0	57.5	56.00	57.13	58.10	60.33	65.20

Preface

TPC Benchmark™ DS Overview

The TPC Benchmark™ DS (TPC-DS) is a decision support benchmark that models several generally applicable aspects of a decision support system, including queries and data maintenance. The benchmark provides are presentative evaluation of performance as a general purpose decision support system.

This benchmark illustrates decision support systems that:

- Examine large volumes of data;
- Give answers to real-world business questions;
- Execute queries of various operational requirements and complexities (e.g., ad-hoc, reporting, iterative OLAP, data mining);
- Are characterized by high CPU and IO load;
- Are periodically synchronized with source OLTP databases through database maintenance functions.
- Run on “Big Data” solutions, such as RDBMS as well as Hadoop/Spark based systems.

A benchmark result measures query response time in single user mode, query throughput in multi user mode and data maintenance performance for a given hardware, operating system, and data processing system configuration under a controlled, complex, multi-user decision support workload.

The purpose of TPC benchmarks is to provide relevant, objective performance data to industry users. To achieve that purpose, TPC benchmark specifications require benchmark tests be implemented with systems, products, technologies and pricing that:

a) Are generally available to users;
b) Are relevant to the market segment that the individual TPC benchmark models or represents (e.g., TPC-DS models and represents complex, high data volume, decision support environments);
c) Would plausibly be implemented by a significant number of users in the market segment modeled or represented by the benchmark.

In keeping with these requirements, the TPC-DS database must be implemented using commercially available data processing software, and its queries must be executed via SQL interface. 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, which improve benchmark results but not real-world performance or pricing, are prohibited.

TPC benchmark results are expected to be accurate representations of system performance. Therefore, there are specific 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 benchmark requirements.

Further information is available at http://www.tpc.org/

General Items

0.1 Test Sponsor

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

This benchmark was sponsored by Alibaba Cloud Computing Ltd.

0.2 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:

- *Database Tuning Options*
- *Optimizer/Query execution options*
- *Query processing tool/language configuration parameters*
- *Recovery/commit options*
- *Consistency/locking options*
- *Operating system and configuration parameters*
- *Configuration parameters and options for any other software component incorporated into the pricing structure*
- *Compiler optimization options*

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

The Supporting File Archive contains the Operating System and DBMS parameters used in this benchmark.

0.3 Configuration Diagrams

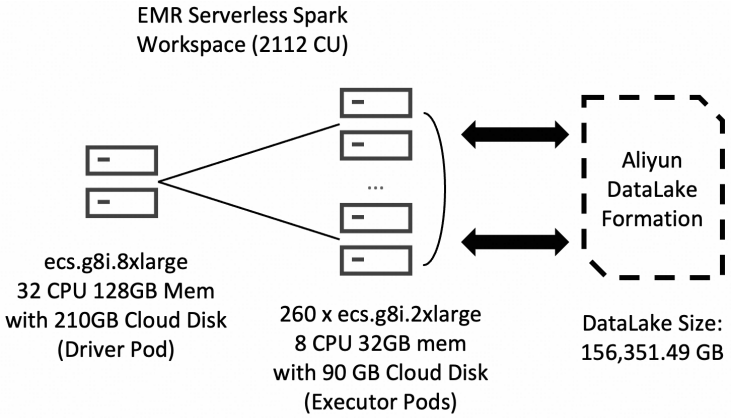
Diagrams of both measured and priced configurations must be provided, accompanied by a description of the differences. This includes, but is not limited to:

- *Number and type of processors*
- *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, including routers, workstations, terminals, etc., that were physically used in the test or are incorporated into the pricing structure.*
- *Type and the run-time execution location of software components (e.g., DBMS, query processing tools/languages, middle-ware components, software drivers, etc.).*

Measured Configuration

Following is an illustration of the measured configuration.

Figure 0.3: Diagram of the Measured Configuration



- The **Driver Pod** receives requests and dispatches them to the Executor Pods.
- The **Executor Pods** executes the queries.
- The **Aliyun DataLake Formation** stores and manages data.

The following table shows the configuration of the measured system.

Table 0.3: Configuration of the measured system.

Configuration	Driver Pod	Executor Pod
Instance Count	1	256
Instance Type	ecs.g8i.8xlarge	ecs.g8i.2xlarge
vCPU per instance	32	8
Memory per instance	128GB	32GB
Storage per instance	210GB SSD Cloud Disk	90GB SSD Cloud Disk
Processors	Intel(R) Xeon(R) Platinum 8575C	220 Intel(R) Xeon(R) Platinum 8575C 40 Intel(R) Xeon(R) Platinum 8475B
Network	Bandwidth: 12 Gbps, Packet forwarding rate: 3,000,000	Bandwidth: 6 Gbps, Packet forwarding rate: 1,600,000

Configuration	Aliyun DataLake Formation
Storage Capacity	156,352GB

Priced Configuration

There are no differences between the priced and measured configurations.

Clause 2: Logical Database Design Related Items

2.1 Database Definition Statements

Listings must be provided for the DDL scripts and must include all table definition statements and all other statements used to set up the test and qualification databases.

The Supporting File Archive contains the table definitions and all other statements used to set up the test and qualification databases.

2.2 Physical Organization

The physical organization of tables and indices within the test and qualification databases must be disclosed. If the column ordering of any table is different from that specified in Clause2.3 or 2.4, it must be noted.

The store_sales, store_returns, catalog_sales, catalog_returns, web_sales, web_returns and inventory are partitioned. The partition columns for these tables respectively are ss_sold_date_sk, sr_returned_date_sk, cs_sold_date_sk, cr_returned_date_sk, ws_sold_date_sk, wr_returned_date_sk and inv_date_sk.

2.3 Horizontal Partitioning

If any directives to DDLs are used to horizontally partition tables and rows in the test and qualification databases, these directives, DDLs, and other details necessary to replicate the partitioning behavior must be disclosed.

Horizontal partitioning is used on store_sales, store_returns, catalog_sales, catalog_returns, web_sales, web_returns and inventory tables and the partitioning columns are ss_sold_date_sk, sr_returned_date_sk, cs_sold_date_sk, cr_returned_date_sk, ws_sold_date_sk, wr_returned_date_sk and inv_date_sk. The partition granularity is by day.

2.4 Replication

Any replication of physical objects must be disclosed and must conform to the requirements of Clause 2.5.3.

No replication was used.

Clause 3: Scaling and Database Population

3.1 Initial Cardinality of Tables

The cardinality (e.g., the number of rows) of each table of the test database, as it existed at the completion of the database load (see Clause 7.1.2) must be disclosed.

Table 3.1 lists the cardinality of each table as they existed upon completion of the build.

Table 3.1 Initial Number of Rows

Table Name	Row Count
call_center	60
catalog_page	50,000
catalog_returns	14,411,635,868
catalog_sales	144,001,341,635
customer	100,000,000
customer_address	50,000,000
customer_demographics	1,920,800
date_dim	73,049
household_demographics	7,200
income_band	20
inventory	1,965,337,830
item	502,000
promotion	2,500
reason	75
ship_mode	20
store	1,902
store_returns	28,783,592,630
store_sales	288,017,344,252
time_dim	86,400
warehouse	30
web_page	5,004
web_returns	7,193,353,603
web_sales	71,986,550,431
web_site	96

3.2 Distribution of Tables and Logs Across Media

The Aliyun DataLake Formation provides an abstraction over the cloud object store with Paimon which stores both the data and the change log on the underlying object store (Aliyun Object Storage Service). This provides an equivalent guarantee as the underlying storage and data distribution is handled transparently by OSS.

3.3 Mapping of Database Partitions/Replications

The mapping of database partitions/replications must be explicitly described.

Neither database partitions nor replications are mapped to specific devices.

3.4 Implementation of RAID

Implementations may use some form of RAID. The RAID level used must be disclosed for each device. If RAID is used in an implementation, the logical intent of its use must be disclosed

Aliyun OSS Standard ZRS (Zone-redundant storage) stores multiple copies of your data across multiple zones in the same region. Your data is still accessible even if a zone becomes unavailable. Aliyun OSS Standard LRS (Locally redundant storage) stores multiple copies of your data on multiple devices of different facilities in the same zone. LRS provides data durability and availability even if hardware failures occur.

3.5 DBGEN Modifications

The version number (i.e., the major revision number, the minor revision number, and third tier number) of dsdgen must be disclosed. Any modifications to the dsdgen source code (see Appendix B:) must be disclosed. In the event that a program other than dsdgen was used to populate the database, it must be disclosed in its entirety.

Dsdgen version 4.0.0 was used. Two minor changes are made to the dsdgen tool. To reduce the dsdgen execution time, the dsdgen code is wrapped as a spark job. The wrapper does not change any of the TPC-provided code. Patches for dsdgen tool and the wrapper with source codes were included in the Supporting Files.

3.6 Database Load time

The database load time for the test database (see Clause 7.4.3.7) must be disclosed.

The database load time was 3,427.1 seconds.

3.7 Data Storage Ratio

The data storage ratio must be disclosed. It is computed by dividing the total data storage of the priced configuration (expressed in GB) by SF corresponding to the scale factor chosen for the test database as defined in Clause 3.1. The ratio must be reported to the nearest 1/100th, rounded up. For example, a system configured with 96 disks of 2.1 GB capacity for a 100GB test database has a data storage ratio of 2.02.

Total local storage Capacity (Disk) = 210 (Driver pod) + 90 * 260 (Executor pods) = 23,610GB

Aliyun DataLake Formation storage used = 156,351.49GB

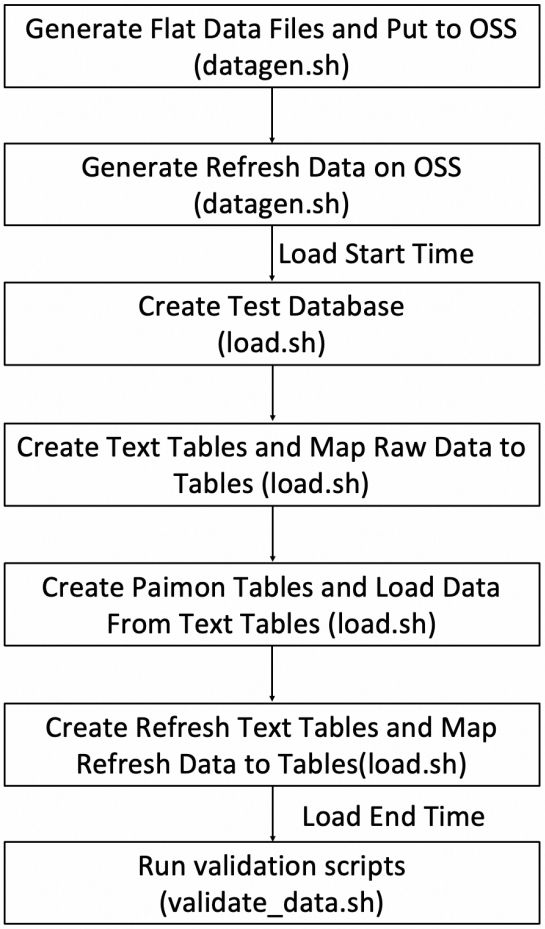
The data storage ration is (23,610 + 156,351.49) / 100,000 = 1.8

3.8 Database Load Mechanism Details and Illustration

The details of the database load must be disclosed, including a block diagram illustrating the overall process. Disclosure of the load procedure includes all steps, scripts, input and configuration files required to completely reproduce the test and qualification databases.

The tables were loaded as shown in Figure 3.8. All of the related source code and scripts are included in the Supporting Files.

Figure 3.8: Block Diagram of Database Load Process



The database load time is (load end time – load start time).

3.9 Qualification Database Configuration

Any differences between the configuration of the qualification database and the test database must be disclosed.

The qualification database is created using the same scripts as the test database with the following exceptions:

- The Scale factor is adjusted to 1GB
- The script create_qual_text_tables.sql is used instead of create_text_tables.sql to build the database in an alternate location within the Aliyun DataLake Formation.

All of the related source code and scripts are included in the Supporting Files.

Clause 4 and 5: Query and Data Maintenance Related Items

4.1 Query Language

The query language used to implement the queries must be identified.

SQL was the query language used to implement the queries.

4.2 Verifying Method of Random Number Generation

The method of verification for the random number generation must be described unless the supplied dsdgen and dsqgen were used.

A spark wrapper based on TPC-supplied dsdgen version 4.0.0 and dsqgen version 4.0.0 were used.

4.3 Generating Values for Substitution Parameters

The method used to generate values for substitution parameters must be disclosed. The version number (i.e., the major revision number, the minor revision number, and third tier number) of dsqgen must be disclosed.

TPC supplied dsqgen version 4.0.0 was used to generate the substitution parameters:

```
./dsqgen -directory ../query_templates -input ../query_templates/templates.lst -scale 100000 -streams 9 -output_dir ../../queries -dialect sparksql -rngseed $SEED
```

4.4 Query Text and Output Data from Qualification Database

The executable query text used for query validation must be disclosed along with the corresponding output data generated during the execution of the query text against the qualification database. If minor modifications have been applied to any functional query definitions or approved variants in order to obtain executable query text, these modifications must be disclosed and justified. The justification for a particular minor query modification can apply collectively to all queries for which it has been used. The output data for the power and Throughput Tests must be made available electronically upon request.

Supporting Files Archive contains the actual query text and query output. Following are the modifications to the query.

The following MQM are used:

- Use vendor specific string concatenation operator. (MQM c.3)
  - Q5
  - Q66
  - Q80
  - Q84
- Use vendor-specific syntax of date expressions. (MQM f.1)
  - Q5
  - Q12
  - Q16
  - Q20
  - Q21
  - Q32
  - Q37
  - Q40

  - Q72
  - Q77
  - Q80
  - Q82
  - Q94
  - Q95
  - Q98
- Use back quotes instead of double quotes to delimit column names. (MQM e.1)
  - Q16
  - Q32
  - Q50
  - Q62
  - Q92
  - Q94
  - Q95
  - Q99

Query results are inserted in a file (Clause 4.2.5) using an external table with column delimiter

- Q34 with an external table named q34_result_[s](stream[s])
- Q39 with an external table named q39_result_[s](stream[s])
- Q64 with an external table named q64_result_[s](stream[s])
- Q71 with an external table named q71_result_[s](stream[s])
- Q98 with an external table named q98_result_[s](stream[s])

The Supporting Files Archive contains the full set of executable query text template used in the test.

4.5 Query Substitution Parameters and Seeds Used

All the query substitution parameters used during the performance test must be disclosed in tabular format, along with the seeds used to generate these parameters.

The Supporting Files Archive contains the query substitution parameters and seed used in the test.

4.6 Refresh Setting

All query and refresh session initialization parameters, settings and commands must be disclosed.

The Supporting Files Archive contains the query and scripts, along with initialization parameters and settings.

4.7 Source Code of Refresh Functions

The details of how the data maintenance functions were implemented must be disclosed (including source code of any non-commercial program used).

The Supporting Files Archive contains the source code implementing the refresh functions.

4.8 Staging Area

Any object created in the staging area (see Clause 5.1.8 for definition and usage restrictions) used to implement the data maintenance functions must be disclosed. Also, any disk storage used for the staging area must be priced, and any mapping or virtualization of disk storage must be disclosed.

No staging area was used.

Clause 6: Data Persistence Properties Related Items

The results of the data accessibility tests must be disclosed along with a description of how the data accessibility requirements were met.

In this benchmark, the Data Accessibility requirements are met by providing documentation of the Data Accessibility features supported by the benchmark configuration.

The Aliyun DataLake Formation provides an abstraction over the cloud object store with Paimon which stores both the data and the change log on the underlying object store (Aliyun Object Storage Service). This provides an equivalent guarantee as the underlying storage and data distribution is handled transparently by OSS.

Public Documentation

This benchmark result was produced using the Aliyun OSS[1] as the underlying storage. The reliability features of the object stores are provided in the following documents:

- General Aliyun OSS Documentation
  Aliyun OSS – 12 9's of durability
- Information on how data availability is achieved in Aliyun OSS
  Aliyun OSS

TPC-DS Data accessibility requirements

The Data Accessibility Documentations must describe how data redundancy is accomplished within the SUT. Following are some examples of such description:

- Data Objects are stored on redundant devices (e.g. RAID 1, RAID 5)
- Data Objects are redundantly stored on multiple storage devices in the same facility.
- Data Objects are redundantly stored across multiple facilities.
- Data Objects are redundantly stored across data centers in multiple regions.

The following features must be supported by the SUT and described in the Data Accessibility Documentation:

- Synchronous writes: The redundant writes of multiple copies of Data Objects to multiple storage devices are executed synchronously.
- Automatic repair: Any loss of redundancy of a Data Object is automatically repaired without any operator intervention.

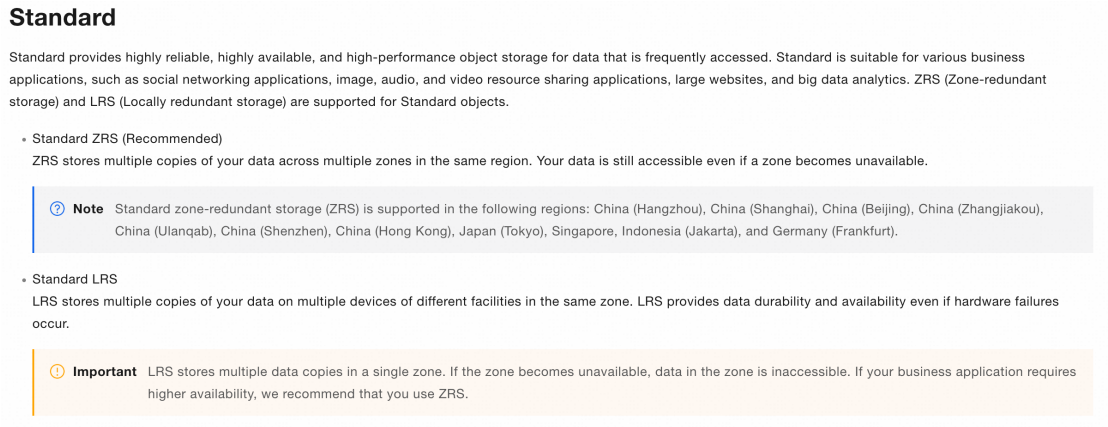
Aliyun OSS Standard ZRS (Zone-redundant storage) stores multiple copies of your data across multiple zones in the same region. Your data is still accessible even if a zone becomes unavailable. Aliyun OSS Standard LRS (Locally redundant storage) stores multiple copies of your data on multiple devices of different facilities in the same zone. LRS provides data durability and availability even if hardware failures occur.

Aliyun OSS Standard ZRS provides 99.9999999999% (12 9's) of data durability of objects over a given year. Aliyun OSS Standard LRS provides 99.999999999% (11 9's) of data durability of objects over a given year. This durability level corresponds to an average annual expected loss of 0.000000001% of objects. For example, if you store 10,000,000 objects with Aliyun OSS, you can on average expect to incur a loss of a single object once every 10,000 years.

Documentation: Aliyun OSS Storage classes

Screen-Capture

April 15, 2025



Standard

Standard provides highly reliable, highly available, and high-performance object storage for data that is frequently accessed. Standard is suitable for various business applications, such as social networking applications, image, audio, and video resource sharing applications, large websites, and big data analytics. ZRS (Zone-redundant storage) and LRS (Locally redundant storage) are supported for Standard objects.

- Standard ZRS (Recommended)
  ZRS stores multiple copies of your data across multiple zones in the same region. Your data is still accessible even if a zone becomes unavailable.

  Note Standard zone-redundant storage (ZRS) is supported in the following regions: China (Hangzhou), China (Shanghai), China (Beijing), China (Zhangjiakou), China (Ulanqab), China (Shenzhen), China (Hong Kong), Japan (Tokyo), Singapore, Indonesia (Jakarta), and Germany (Frankfurt).

- Standard LRS
  LRS stores multiple copies of your data on multiple devices of different facilities in the same zone. LRS provides data durability and availability even if hardware failures occur.

  Important LRS stores multiple data copies in a single zone. If the zone becomes unavailable, data in the zone is inaccessible. If your business application requires higher availability, we recommend that you use ZRS.

Storage class	Redundancy type	Data durability
Standard	ZRS (Recommended)	99.9999999999%
IA		
Archive		
Standard	LRS	99.999999999%
IA		
Archive		
Cold Archive		

References

[1] https://www.alibabacloud.com/en/product/object-storage-service?_p_lc=1

[2] https://www.alibabacloud.com/help/en/oss/user-guide/overview-53/

[3] https://www.alibabacloud.com/help/en/legal/latest/object-storage-service-service-level-agreement

Clause 7: Performance Metrics and Execution Rules Related Items

7.1 System Activity

Any system activity on the SUT that takes place between the conclusion of the load test and the beginning of the performance test must be fully disclosed including listings of scripts or command logs.

The only activity between the end of the load test and the beginning of the performance test was the generation of the executable query text.

7.2 Test Steps

The details of the steps followed to implement the performance test must be disclosed.

The Supporting Files Archive contains the scripts and logs.

7.3 Timing Intervals for Each Query and Refresh Function

The timing intervals defined in Clause 7 must be disclosed.

See the Executive Summary at the beginning of this report.

7.4 Throughput Test Result

For each Throughput Test, the minimum, the 25th percentile, the median, the 75th percentile, and the maximum times for each query shall be reported.

See the Executive Summary at the beginning of this report.

7.5 Time for Each Stream

The start time and finish time for each query stream must be reported.

See the Executive Summary at the beginning of this report.

7.6 Time for Each Refresh Function

The start time and finish time for each data maintenance function in the refresh run must be reported for the Throughput Tests

See the Executive Summary at the beginning of this report.

7.7 Performance Metrics

The computed performance metric, related numerical quantities and the price/performance metric must be reported.

QphDS@100000 = 65,684,477

See the Executive Summary at the beginning of this report for more detail.

Clause 8: SUT and Driver Implementation Related Items

8.1 Driver

A detailed textual description of how the driver performs its functions, how its various components interact and any product functionalities or environmental settings on which it relies must be provided. All related source code, scripts and configuration files must be disclosed. The information provided should be sufficient for an independent reconstruction of the driver.

EMR Serverless Spark uses spark-beeline as the client. It connects to the Spark Thrift Server running in driver pod by JDBC. The command is:

spark-beeline -u “jdbc:hive2://emr-spark-gateway-cn-beijing.data.aliyun.com:443/tpcds_oss_paimon_100000_ss_test” -f sqlfile

The Spark Thrift Server accepts SQL queries from the spark-beeline clients and processes the queries. The driver pod manages multiple executor pods. All queries are compiled on the driver and then submitted to the executors as a job. When the job finishes, the driver takes the result from the executors and sends it to spark-beeline.

The Supporting Files Archive contains all the command, scripts and logs.

8.2 Implementation Specific Layer (ISL)

If an implementation specific layer is used, then a detailed description of how it performs its functions, how its various components interact and any product functionalities or environmental setting on which it relies must be provided. All related source code, scripts and configuration files must be disclosed. The information provided should be sufficient for an independent reconstruction of the implementation specific layer.

No Implementation Specific Layer was used.

8.3 Profile-Directed Optimization

If profile-directed optimization as described in Clause 7.2.10 is used, such use must be disclosed. In particular, the procedure and any scripts used to perform the optimization must be disclosed.

Profile-directed optimization was not used.

Clause 9: Pricing Related Items

9.1 Hardware and Software Used

A detailed list of hardware and software used in the priced system must be reported. The rules for pricing are included in the current revision of the TPC Pricing Specification located on the TPC website (http://www.tpc.org)

A detailed list of all licensed services, hardware and software, is provided in the Executive Summary of this report.

9.2 Availability Date

The System Availability Date (see Clause 7.6.5) must be the single availability date reported on the first page of the executive summary. The full disclosure report must report Availability Dates individually for at least each of the categories for which a pricing subtotal must be. All Availability Dates required to be reported must be disclosed to a precision of 1 day, but the precise format is left to the test sponsor.

The total system will be available on 2025-10-20.

9.3 Country-Specific Pricing

Additional Clause 7 related items may be included in the full disclosure report for each country specific priced configuration.

The configuration is priced for the Chinese market.

Clause 11: Audit Related Items

Auditor’s Information and Attestation Letter

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.

This benchmark was audited by: Doug Johnson, of InfoSizing.

InfoSizing
The Right Metric For Sizing IT

TPC
Certified Auditor

Keyong Zhou
Cloud Valley, NO.1008 Dengcai Street, Xihu District,
Hangzhou, China

September 9, 2025

I verified the TPC Benchmark™ DS v4.0.0 performance of the following configuration:

Platform:	Alibaba Cloud E-MapReduce Serverless Spark (with 1 + 260 Pods)
Operating System:	Alibaba Cloud Linux 3.2104 U11 (OpenAnolis Edition)
Database Manager:	Alibaba Cloud E-MapReduce Serverless Spark esr-4.5.1

The results were:

Performance Metric	**65,684,477**	**QphDS@100000**
Secondary Metrics		
T_{Load}	3,427.1	seconds@100000
T_{Power}	2,929.2	seconds@100000
T_{TT1}	5,252.7	seconds@100000
T_{TT2}	5,286.5	seconds@100000
T_{DM1}	638.8	seconds@100000
T_{DM2}	670.9	seconds@100000

Servers	**1x ecs.g8i.8xlarge (Driver Pod); 260x ecs.g8i.2xlarge (Executor Pods) with:**		
CPUs	1x Intel® Xeon® Platinum 8575C (Driver Pod) 1x Intel Xeon Platinum 8575C (220 Executor Pods) 1x Intel Xeon Platinum 8475B (40 Executor Pods)		
Memory	128 GiB (Driver Pod); 32 GiB (Executor Pods)		
Storage	***Qty***	***Size***	***Type***
	1	210 GiB	SSD Cloud Disk (Driver Pod)
	1	90 GiB	SSD Cloud Disk (Executor Pods)

Storage	**Aliyun DataLake Formation**
Max Total Data Size	156,351.49 GiB

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

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

The following verification items were given special attention:

- The database records were defined with the proper layout and size.
- The database population was generated using DSDGen v4.0.0
- The database was properly scaled to 100,000 GB and populated accordingly.
- The primary and secondary metrics were correctly measured and reported.
- The query templates were produced using approved minor query modifications and/or query variants.
- The query substitution parameters were generated using DSQGen v4.0.0.
- The execution of the queries against the qualification database produced compliant output.
- The tests were driven and sequenced according to the requirements.
- Each throughput test comprised 4 query streams.
- The execution times for queries and data maintenance functions were correctly measured and reported.
- The data accessibility test was satisfied through documentation.
- The system pricing was verified for major components and maintenance.
- The major pages from the FDR were verified for accuracy.

Additional Audit Notes:

None.

Respectfully Yours,

Doug Johnson, Certified TPC Auditor

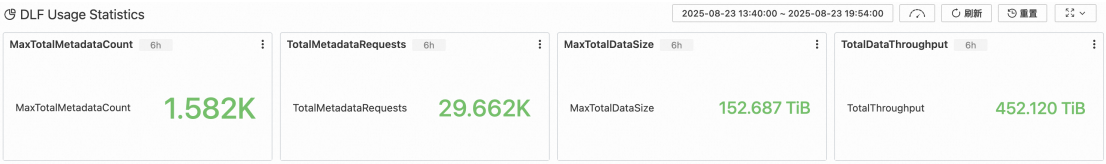
63 Lourdes Dr. | Leominster, MA 01453 | 978-343-6562 | www.sizing.com

Supporting Files Index

Clause	Description	Archive File Pathname
Clause 3	Database create and load scripts, SQL scripts for table creation and validation	SupportingFiles/Clause_3/
	The code for the Spark wrapper of dsdgen	SupportingFiles/Clause_3/datagen
	Patches for data generation tools	SupportingFiles/Clause_3/patches/tools/
Clause 4	The script to execute qualification test	SupportingFiles/Clause_4/
	Patches for query templates	SupportingFiles/Clause_4/patches/query_templates/
	SQL for qualification queries	SupportingFiles/Clause_4/queries/
	Output from executing qualification queries	SupportingFiles/Clause_4/output/
Clause 5	Data maintenance execution scripts and logs files	SupportingFiles/Clause_5/
	SQL scripts for DM functions for stream [s]	SupportingFiles/Clause_5/mtsqls_[s]/
	Data file with delete dates	SupportingFiles/Clause_5/delete/
		SupportingFiles/Clause_5/inventory_delete/
Clause 6	Data accessibility documentation	SupportingFiles/Clause_6/DA_Documentation.docx
Clause 7	Performance test scripts and logs	SupportingFiles/Clause_7/
	Query text for query [q] in stream [s]	SupportingFiles/Clause_7/stream_[s]_queries/query_[q].sql
	Output of query [q] in stream [s]	SupportingFiles/Clause_7/stream_[s]_results/query_[q].out
Clause 8	EMR Serverless Spark Configuration Inventory	SupportingFiles/Clause_8/

Appendix A: Purchase Page of Creating Alibaba Cloud E-MapReduce Serverless Spark with 1-Year Subscription

Appendix B: Aliyun DataLake Formation Resource Utilization



Appendix C: Third Party Price Quotes

Lenovo P14H Laptop: