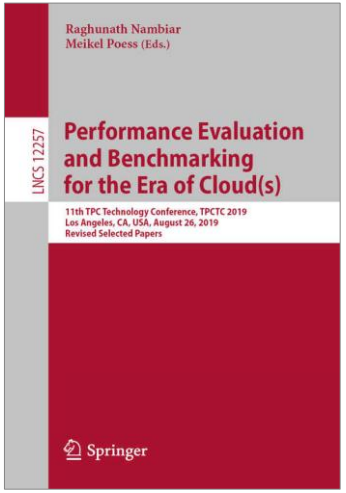
peakmarks

peakmarks® Workloads and Key Performance Metrics



peakmarks® showcased its software at the 2019 TPC Technology Conference in Los Angeles.

peakmarks® Version 10.2

May 2024

Copyright Information

Safe Harbor Statement

All performance data were determined with the peakmarks® software under certain conditions and do not necessarily correspond to the manufacturer's specifications.

Abbreviations and Metrics

[MBps]	megabyte per second
[GBps]	gigabyte per second
[dbps]	database blocks per second
[rbps]	redo blocks per second
[dbpt]	database blocks per transaction
[s]	seconds
[ms]	milliseconds
[µs]	microseconds
[IOPS]	I/O operations per second
[qps]	queries per second
[rps]	rows per second
[tps]	transactions per second
[kBpt]	kilobyte per transaction
[Mops]	million operations per second
Nodes	number of cluster nodes
Jobs	number of workload processes
BuCache	Database Buffer Cache
FlCache	Database or Exadata Flash Cache

In the following reports, the key performance metrics are marked red.

Naming Conventions

Database name	ORA19C / ORA21C / ORA23c	
Instance names	ORA19C / ORA21C / ORA23C	for a single instance
	ORA19C1 / ORA21C1 / ORA23C1	for RAC instance 1
	ORA19C2 / ORA21C2 / ORA23C2	for RAC instance 2
peakmarks PDB	PMK	
Connect string SYSTEM user	system/manager@SYSAWR	
Connect string peakmarks user	bench/bench@PMK	
peakmarks® base directory	../pmk	

Agenda

1 Platform Description

2 Workload Overview

3 Server Systems

4 Storage Systems

5 Log Writer Background Processes

6 Database Writer Background Processes

7 Data Load

8 Data Analytics

9 Transaction Processing

10 PL/SQL Application Programs

11 Summary of Scripts and Commands

peakmarks

Performance is not everything.
But without performance, everything is worth nothing.

Platform description

Processor

	Oracle Exadata X7-2 Flex Single Instance	Oracle Exadata X7-2 Flex 2-Node RAC Cluster
Processor type	Intel Xeon 8160, 2.1 – 3.7 GHz	Intel Xeon 8160, 2.1 – 3.7 GHz
Launch date	Q3/2017	Q3/2017
#cores per cpu	24	24
Multithreading	Enabled, 2-fold	Enabled, 2-fold
#threads per socket	48	48
Memory type	DDR4, 2.6 GHz	DDR4, 2.6 GHz

Server

	Oracle Exadata X7-2 Flex Single Instance	Oracle Exadata X7-2 Flex 2-Node RAC Cluster
Server system	1 x Exadata X7-2 Scale Out Database Server	2 x Exadata X7-2 Scale Out Database Server
Launch date	2017	2017
Technology	Bare metal	Bare metal
#sockets, total	2	4
#cores, total	48	96
#threads, total	96	192
DRAM capacity total [GByte], total	768	1,536
DRAM capacity per core [GByte]	8	8
Operating System	OEL	OEL
Connectivity, per server system	InfiniBand 2 x 40 Gbps	InfiniBand 2 x 40 Gbps

Storage

	Oracle Exadata X7-2 Flex Single Instance	Oracle Exadata X7-2 Flex 2-Node RAC Cluster
Storage system	3 x Exadata X8 Scale-Out Storage Server High Capacity	3 x Exadata X8 Scale-Out Storage Server High Capacity
DRAM capacity, total	192 GByte	192 GByte
PMEM capacity, total	-	-
Flash capacity, total raw	76.8 TByte	76.8 TByte
Disk capacity, total raw	504 TByte	504 TByte
Connectivity, per storage system	InfiniBand 2 x 40 Gbps	InfiniBand 2 x 40 Gbps
File system	ASM high redundancy ASM allocation unit 4 MByte	ASM high redundancy ASM allocation unit 4 MByte
Compression	No	No
Deduplication	No	No

Database

	Oracle Exadata X7-2 Flex Single Instance	Oracle Exadata X7-2 Flex 2-Node RAC Cluster
Database version	Oracle 19.20 Enterprise Edition	Oracle 19.20 Enterprise Edition
Database block size	8 kByte	8 kByte
Log Modus	NOARCHIVELOG	NOARCHIVELOG
DataGuard	No	No
REDO Log Files, per instance	4 x 4 GByte without multiplexing	4 x 4 GByte without multiplexing
SGA size, per instance	384 GByte	384 GByte
PGA size, per instance	150 GByte	150 GByte
CPU_COUNT, per instance	96	96

peakmarks® Software

	Oracle Exadata X7-2 Flex Single Instance	Oracle Exadata X7-2 Flex 2-Node RAC Cluster
Version	10.2	10.2
Build	231115	231115
Database size	8 TByte	2 x 4 TByte

Notes:

- To ensure full transparency, the peakmarks® Software generates individual Oracle AWR reports for each single performance test. In Oracle AWR reports, the idle wait event "enq: UL - contention" indicates process synchronization by the peakmarks® master process and does not cause wait states of workload execution processes.
- peakmarks® shows slightly better performance results because peakmarks® is the inner snapshot around tests while AWR is the outer snapshot for performance statistics.

peakmarks

Simple. Representative. Fast.

Workload Overview

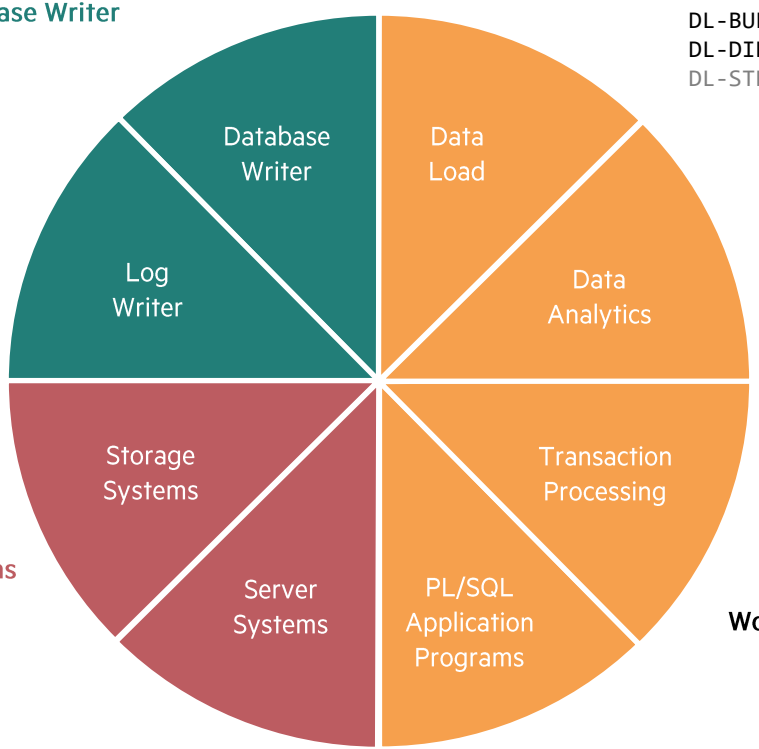
Note: Important workloads are shown in black, while gray workloads provide additional information. All workloads are generated with Oracle code within the Oracle environment.

Workload for Database Writer
- DBWR-THR

Workloads for Log Writer
- LGWR-LAT1
- LGWR-LAT25
- LGWR-LAT125
- LGWR-THR

Workloads for Storage Systems
- STO-READ
- STO-OFFLOAD
- STO-RANDOM
- STO-SCATTER

Workloads for Server Systems
- SRV-MIXED
- SRV-QUERY1
- SRV-QUERY25
- SRV-REPORT
- SRV-SCAN



Workloads for Data Load
- DL-BUFFER
- DL-DIRECT
- DL-STREAM

Workloads for Data Analytics
- DA-STORAGE
- DA-OFFLOAD
- DA-ROWSTORE
- DA-COLSTORE

Workloads for Transaction Processing
- TP-MIXED1
- TP-MIXED2
- TP-REPORT
- TP-LOOKUP
- TP-LIGHT
- TP-MEDIUM
- TP-HEAVY

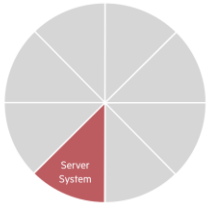
Workloads for PL/SQL-Application Programs
- PLS-MIXED
- PLS-ADD
- PLS-BUILTIN
- PLS-FIBO
- PLS-PRIME

peakmarks

Stop guessing. Start measuring.

peakmarks

Workloads to determine the Server Performance in Database Operations



Motivation

The server performance significantly impacts the performance of all database operations.

The goal is to

- Validate the performance capabilities (speed, throughput, scalability) of server components in database operation: processors, main memory, and internal memory channels
- Determine the impact of server virtualization, multithreading, NUMA effects, and encryption on server performance
- Optimize database license and maintenance costs for server system (per-core performance matters)

Notes

- Some cloud service providers do not publish their server components and configurations. Components and configurations of cloud services are subject to change without any prior notice.
- Customers need to know the per-core performance, significantly impacting application performance and the required Oracle licenses. In many cases, Oracle licensing costs far exceed infrastructure costs.
- Multithreaded servers often show misleading CPU utilization and do not deliver consistent per-core performance with increasing load.

Key Performance Metrics

- SQL query throughput in queries per second [qps]
- SQL query response time in milliseconds [ms]
- Logical reads in database blocks per second [dbps]
- SQL buffer cache scan rate in megabytes per second [MBps]

Server System Workloads

Description

Workload	Measurement Unit	Action
SRV-QUERY1	[qps] [ms]	Latency-oriented look-up query – select 1 row via index, e.g., select customer, account, product, order, invoice. This workload shows maximum query throughput and minimum response time for simple queries.
SRV-QUERY25	[qps] [ms]	Data volume-oriented look-up query – select Ø 25 rows via index, e.g., select last month's bank account bookings; select item list of order. This workload shows maximum query throughput and minimum response time for more complex queries.
SRV-REPORT	[dbps]	Online Report – select Ø 125 rows via index, e.g., select last month's cell phone call records. This workload shows maximum logical read throughput.
SRV-SCAN	[MBps]	Full table scan. Search for data without index support. This workload shows a maximum database buffer cache scan rate.

Notes

- All accessed data is completely stored in the database buffer cache. There are no I/O operations, and all SRV workloads are CPU-bound.
- These kinds of queries are generic to all applications in all industries.

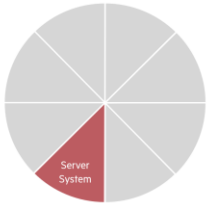
Description

Workload	Measurement Unit	Action
SRV-MIXED	[qps] [ms]	Mixed queries and full table scans on cached data. This complex workload comprises the equally weighted simple workloads SRV-QUERY1, SRV-QUERY25, SRV-REPORT, and SRV-SCAN. SPEC numbers are not always representative of Oracle database operations. SRV-MIXED is the most representative peakmarks workload used to determine server system performance in Oracle database operations.

Notes

- All accessed data is completely stored in the database buffer cache. There are no I/O operations, and all SRV workloads are CPU-bound.
- These kinds of queries are generic to all applications in all industries.
- Peakmarks uses the result of this workload to compare ARM, Intel Xeon, AMD EPYC, IBM POWER, and IBM Z processors.

Server System Performance



Workload SRV-QUERY1 – Simple look-up query, highest throughput, lowest response time

Single Instance

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	Queries total [qps]	Queries per cpu [qps]	Response time [ms]	Log reads total [dbps]	Log reads per cpu [dbps]	BuCache read [%]	Elapsed time [s]
2	7	SRV-QUERY1	1	1	2	1	0	98	111,223	111,223	0.009	445,066	445,066	99.94	181
	8	SRV-QUERY1	1	24	26	25	0	74	2,083,530	86,814	0.011	8,312,219	346,342	100.00	182
	9	SRV-QUERY1	1	48	51	50	1	49	3,267,839	68,080	0.015	12,991,193	270,650	100.00	181
	10	SRV-QUERY1	1	72	75	74	1	25	3,234,717	44,927	0.022	12,842,742	178,371	100.00	182
	11	SRV-QUERY1	1	96	98	97	1	2	3,346,258	34,857	0.029	13,272,540	138,256	100.00	182

2-node Cluster

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	Queries total [qps]	Queries per cpu [qps]	Response time [ms]	Log reads total [dbps]	Log reads per cpu [dbps]	BuCache read [%]	Elapsed time [s]
9	7	SRV-QUERY1	2	2	2	1	0	98	233,345	116,673	0.009	933,557	466,779	99.95	180
	8	SRV-QUERY1	2	48	25	25	0	75	4,178,134	87,044	0.011	16,669,620	347,284	100.00	182
	9	SRV-QUERY1	2	96	51	50	1	49	6,520,208	67,919	0.015	25,944,413	270,254	100.00	182
	10	SRV-QUERY1	2	144	75	73	1	25	6,501,930	45,152	0.022	25,846,679	179,491	100.00	182
	11	SRV-QUERY1	2	192	98	96	1	2	6,795,993	35,396	0.028	26,987,811	140,562	100.00	182

Workload SRV-QUERY25 – More complex query, highest throughput, lowest response time

Single Instance

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	Queries total [qps]	Queries per cpu [qps]	Response time [ms]	Log reads total [dbps]	Log reads per cpu [dbps]	BuCache read [%]	Elapsed time [s]
2	12	SRV-QUERY25	1	1	2	1	0	98	25,985	25,985	0.038	728,475	728,475	99.96	182
	13	SRV-QUERY25	1	24	26	25	0	74	443,080	18,462	0.054	12,392,509	516,355	100.00	182
	14	SRV-QUERY25	1	48	51	50	0	49	776,193	16,171	0.062	21,643,524	450,907	100.00	182
	15	SRV-QUERY25	1	72	75	74	1	25	889,468	12,354	0.080	24,759,956	343,888	100.00	182
	16	SRV-QUERY25	1	96	98	97	1	2	989,822	10,311	0.096	27,505,028	286,511	100.00	182

2-node Cluster

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	Queries total [qps]	Queries per cpu [qps]	Response time [ms]	Log reads total [dbps]	Log reads per cpu [dbps]	BuCache read [%]	Elapsed time [s]
9	12	SRV-QUERY25	2	2	2	1	0	98	53,726	26,863	0.037	1,506,933	753,467	99.96	180
	13	SRV-QUERY25	2	48	25	25	0	75	902,585	18,804	0.053	25,235,088	525,731	100.00	182
	14	SRV-QUERY25	2	96	51	50	0	49	1,577,747	16,435	0.061	43,988,376	458,212	100.00	182
	15	SRV-QUERY25	2	144	75	73	1	25	1,802,631	12,518	0.079	50,170,859	348,409	100.00	182
	16	SRV-QUERY25	2	192	97	96	1	3	2,000,281	10,418	0.096	55,576,848	289,463	100.00	182

Workload SRV-REPORT – Online Report, max throughput of Logical Reads

Single Instance

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	Queries total [qps]	Queries per cpu [qps]	Response time [ms]	Log reads total [dbps]	Log reads per cpu [dbps]	BuCache read [%]	Elapsed time [s]
2	2	SRV-REPORT	1	1	2	1	0	98	6,670	6,670	0.150	852,999	852,999	100.00	181
	3	SRV-REPORT	1	24	26	25	0	74	109,872	4,578	0.217	14,013,303	583,888	100.00	182
	4	SRV-REPORT	1	48	51	50	0	49	193,727	4,036	0.246	24,627,202	513,067	100.00	182
	5	SRV-REPORT	1	72	76	74	1	24	229,886	3,193	0.312	29,152,184	404,891	100.00	182
	6	SRV-REPORT	1	96	99	98	1	1	258,873	2,697	0.369	32,756,754	341,216	100.00	182

2-node Cluster

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	Queries total [qps]	Queries per cpu [qps]	Response time [ms]	Log reads total [dbps]	Log reads per cpu [dbps]	BuCache read [%]	Elapsed time [s]
9	2	SRV-REPORT	2	2	2	1	0	98	13,884	6,942	0.144	1,775,598	887,799	100.00	180
	3	SRV-REPORT	2	48	25	25	0	75	221,925	4,623	0.215	28,290,092	589,377	100.00	182
	4	SRV-REPORT	2	96	50	49	0	50	391,932	4,083	0.244	49,811,699	518,872	100.00	182
	5	SRV-REPORT	2	144	76	74	1	24	461,731	3,206	0.310	58,545,252	406,564	100.00	182
	6	SRV-REPORT	2	192	99	97	1	1	525,654	2,738	0.363	66,502,195	346,366	100.00	182

Workload SRV-SCAN – Scan-Rate in Oracle Buffer Cache

Single Instance

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	Scan rate total [MBps]	Scan rate per cpu [MBps]	Log reads total [dbps]	Log reads per cpu [dbps]	BuCache read [%]	Elapsed time [s]
2	17	SRV-SCAN	1	1	2	1	0	98	2,658	2,658	340,660	340,660	99.99	181
	18	SRV-SCAN	1	24	26	25	0	74	53,452	2,227	6,840,661	285,028	100.00	183
	19	SRV-SCAN	1	48	51	50	0	49	93,348	1,945	11,918,097	248,294	100.00	183
	20	SRV-SCAN	1	72	75	74	0	25	112,195	1,558	14,164,011	196,722	100.00	183
	21	SRV-SCAN	1	96	98	97	0	2	127,023	1,323	16,081,687	167,518	100.00	183

2-node Cluster

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	Scan rate total [MBps]	Scan rate per cpu [MBps]	Log reads total [dbps]	Log reads per cpu [dbps]	BuCache read [%]	Elapsed time [s]
9	17	SRV-SCAN	2	2	2	1	0	98	5,543	2,772	710,561	355,281	99.99	182
	18	SRV-SCAN	2	48	25	25	0	75	108,285	2,256	13,855,451	288,655	100.00	183
	19	SRV-SCAN	2	96	51	50	0	49	187,009	1,948	23,871,434	248,661	100.00	183
	20	SRV-SCAN	2	144	75	74	1	25	227,210	1,578	28,705,443	199,343	100.00	183
	21	SRV-SCAN	2	192	98	97	1	2	258,579	1,347	32,710,285	170,366	100.00	183

Workload SRV-MIXED – Mixed queries and full table scans

Single Instance

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	Queries total [qps]	Queries per cpu [qps]	Response time [ms]	Log reads total [dbps]	Log reads per cpu [dbps]	BuCache read [%]	Elapsed time [s]
3	2	SRV-MIXED	1	4	5	4	0	95	141,813	35,453	0.028	2,159,436	539,859	100.00	182
	3	SRV-MIXED	1	24	26	25	0	74	662,709	27,613	0.036	10,399,503	433,313	100.00	182
	4	SRV-MIXED	1	48	51	50	0	49	1,113,074	23,189	0.043	17,843,103	371,731	100.00	182
	5	SRV-MIXED	1	72	75	74	1	25	1,221,603	16,967	0.059	20,503,328	284,768	100.00	182
	6	SRV-MIXED	1	96	98	97	1	2	1,383,477	14,411	0.069	23,460,145	244,377	100.00	182

2-node Cluster

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	Queries total [qps]	Queries per cpu [qps]	Response time [ms]	Log reads total [dbps]	Log reads per cpu [dbps]	BuCache read [%]	Elapsed time [s]
10	2	SRV-MIXED	2	8	5	4	0	95	272,844	34,106	0.029	4,290,051	536,256	100.00	182
	3	SRV-MIXED	2	48	25	25	0	75	1,331,390	27,737	0.036	21,115,772	439,912	100.00	183
	4	SRV-MIXED	2	96	51	50	0	49	2,238,224	23,315	0.042	36,500,369	380,212	100.00	183
	5	SRV-MIXED	2	144	75	74	1	25	2,442,773	16,964	0.058	42,067,495	292,135	100.00	182
	6	SRV-MIXED	2	192	98	97	1	2	2,736,163	14,251	0.070	47,252,138	246,105	100.00	182

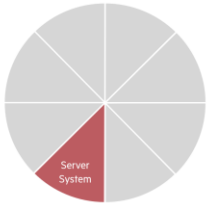
Note

If the number of jobs exceeds the number of cores, the throughput increases only slightly, but the response time increases significantly.

Server System Performance

Selected comparative figures – For training purposes only

In many cases, the per-core performance is relevant for licensing



Workload SRV-MIXED – Mixed queries and full table scans

Ampere Altra A1
32c, 32t, Launch 2021

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	Queries total [qps]	Queries per cpu [qps]	Response time [ms]	Log reads total [dbps]	Log reads per cpu [dbps]	BuCache read [%]	Elapsed time [s]
3	2	SRV-MIXED	1	4	13	12	0	87	70,019	17,505	0.057	1,549,780	387,445	100.00	182
	3	SRV-MIXED	1	8	25	25	0	75	138,227	17,278	0.058	3,129,406	391,176	100.00	182
	4	SRV-MIXED	1	16	50	49	1	50	275,590	17,224	0.058	5,987,945	374,247	100.00	182
	5	SRV-MIXED	1	24	75	74	1	25	417,896	17,412	0.057	8,961,438	373,393	100.00	182
	6	SRV-MIXED	1	32	99	98	1	1	550,974	17,218	0.058	11,643,341	363,854	100.00	182

Note
The Ampere Altra A1 processor does not use hyperthreading and provides predictable and scalable performance, even under high load.

Exadata X9M-2
2s, 64c, 128t, Launch 2021

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	Queries total [qps]	Queries per cpu [qps]	Response time [ms]	Log reads total [dbps]	BuCache read [%]	Elapsed time [s]
22	1	SRV-MIXED	1	4	3	3	0	97	156,927	39,232	0.025	2,646,526	100.00	301
	2	SRV-MIXED	1	32	25	25	0	75	1,042,878	32,590	0.031	15,816,338	100.00	301
	3	SRV-MIXED	1	64	50	50	0	50	1,876,485	29,320	0.034	27,521,600	100.00	301
	4	SRV-MIXED	1	96	75	74	1	25	2,088,995	21,760	0.046	30,505,022	100.00	301
	5	SRV-MIXED	1	128	98	96	1	2	2,271,821	17,749	0.056	33,549,232	100.00	302

Workload SRV-MIXED – Mixed queries and full table scans

Exadata X9M-8
8s, 192c, 384t
Launch 2021

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	Queries total [qps]	Queries per cpu [qps]	Response time [ms]	Log reads total [dbps]	Log reads per cpu [dbps]	BuCache read [%]	Elapsed time [s]
2	27	SRV-MIXED	1	4	2	1	0	98	126,555	31,639	0.032	1,641,272	410,318	100.00	301
	28	SRV-MIXED	1	96	25	24	0	75	1,981,945	20,645	0.048	24,422,632	254,402	100.00	303
	29	SRV-MIXED	1	192	47	46	1	53	2,586,784	13,473	0.073	37,831,820	197,041	100.00	305
	30	SRV-MIXED	1	288	67	65	1	33	2,662,710	9,246	0.106	42,466,861	147,454	100.00	307
	31	SRV-MIXED	1	384	83	82	1	17	2,522,164	6,568	0.150	44,448,615	115,752	100.00	305

Lenovo SR850 V2
4 s, 112c, 224t
Launch 2021

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	Queries total [qps]	Queries per cpu [qps]	Response time [ms]	Log reads total [dbps]	Log reads per cpu [dbps]	BuCache read [%]	Elapsed time [s]
10	22	SRV-MIXED	1	4	2	2	0	98	135,913	33,978	0.029	2,193,687	548,422	0.00	301
	23	SRV-MIXED	1	56	25	25	0	75	1,524,515	27,223	0.037	21,028,968	375,517	0.00	302
	24	SRV-MIXED	1	112	49	49	0	51	2,613,720	23,337	0.042	32,724,991	292,187	0.00	304
	25	SRV-MIXED	1	168	73	72	0	27	2,781,261	16,555	0.059	33,463,008	199,185	0.00	306
	26	SRV-MIXED	1	224	96	95	0	4	2,847,476	12,712	0.077	34,354,228	153,367	0.00	305

Workload SRV-MIXED – Mixed queries and full table scans

IBM POWER9
32c, 256t, Launch 2018

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	Queries total [qps]	Queries per cpu [qps]	Response time [ms]	Log reads total [dbps]	Log reads per cpu [dbps]	BuCache read [%]	Elapsed time [s]
3	37	SRV-MIXED	1	4	4	2	3	96	52,620	13,155	0.076	1,059,105	264,776	100.00	181
	38	SRV-MIXED	1	128	48	48	0	52	855,659	6,685	0.149	16,169,406	126,323	100.00	182
	39	SRV-MIXED	1	256	90	90	0	10	1,064,694	4,159	0.239	19,250,904	75,199	100.00	182

Notes

- The Oracle performance on POWER9 processors does not correlate with SPEC numbers when compared to Intel Xeon.
- POWER9 server achieves 33,271 queries and scans per second per core.

IBM LinuxOne z14
24c, 24t, Launch 2018

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	Queries total [qps]	Queries per cpu [qps]	Response time [ms]	Log reads total [dbps]	Log reads per cpu [dbps]	BuCache read [%]	Elapsed time [s]
4	15	SRV-MIXED	1	4	11	10	1	89	153,056	38,264	0.013	3,290,704	822,676	100.00	182
	16	SRV-MIXED	1	24	49	47	2	51	726,598	30,274	0.019	15,523,759	646,823	100.00	183
	17	SRV-MIXED	1	48	97	94	3	3	784,725	16,348	0.036	16,489,034	343,521	100.00	184

Notes

- The z14 can be configured with and without multithreading.
- LinuxOne server achieves 32,697 queries and scans per second per core.

License Costs per Query

	Ampere Altra A1	Intel Xeon 8358	AMD EPYC 9J14	IBM POWER9	IBM z14
Launch	2021	2021	2023	2018	2018
Clock rate in [GHz]	<= 3.3	2.6 – 3.4	2.6 – 3.7	3.55 – 3.9	4.5
License costs per Core in [USD]	11,875	23,750	23,750	47,500	47,500
Thread performance					
▪ Low CPU utilization	17,505	39,232	46,696	13,155	38,264
▪ High CPU utilization	17,218	17,749	15,585	4,159	16,348
▪ Performance decline	2%	55%	67%	68%	57%
Core performance & cost at full CPU utilization					
▪ Max queries per core	17,218	35,498	33,170	33,271	32,697
▪ License cost per query in [USD]	0.69	0.67	0.72	1.43	1.45

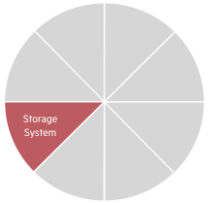
Note
Based on peakmarks® workload SRV-MIXED.

peakmarks

Stop guessing. Start measuring.

peakmarks

Workloads to determine the Storage Performance in Database Operations



Motivation

The storage performance significantly impacts the performance of all database operations.

The goal is to

- Validate the performance capabilities of the storage system in database operations
- Validate the impact of storage technologies like deduplication, compression, encryption, replication, off-loading, tiering, etc., on database performance and storage space utilization
- Optimize storage system license and maintenance costs

Note
Some cloud service providers do not publish their storage components and configurations. Components and configurations of cloud services are subject to change without notice.

Key Performance Metrics

- SQL sequential read throughput in megabytes per second [MBps]
- SQL random read throughput in I/O operations per second [IOPS]
- SQL I/O service time in milliseconds [ms]
- SQL random write throughput in database blocks per second [dbps]

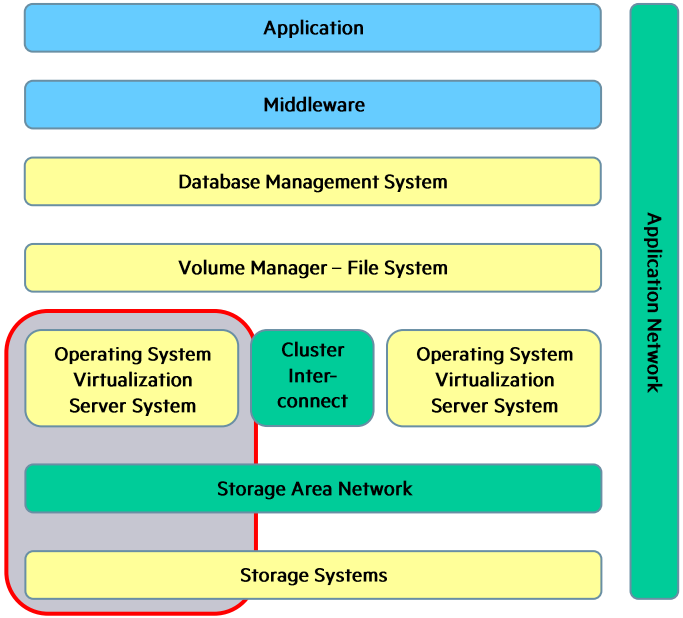
Storage System Workloads

Description

Workload	Measurement Unit	Action
STO-READ	[MBps]	SQL statements performing a sequential table scan using conventional storage devices. This workload delivers maximum sequential throughput using conventional storage.
STO-OFFLOAD	[MBps]	SQL statements performing a sequential table scan using smart-scan offload technology. This workload delivers maximum sequential throughput using smart-scan technology on Oracle Engineered Systems.
STO-RANDOM	[IOPS] [ms]	SQL statements reading/updating blocks via index access for different read/write ratios. The workload parameter specifies the write ratio in %. The following values are supported {0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100}. This workload delivers maximum random I/O operations per database server and/or storage system.
STO-SCATTER	[dbps]	SQL statement performing scattered block writes bypassing the buffer cache. The workload STO-SCATTER writes the blocks by foreground processes; in contrast, the DBWR-THR workload writes the blocks by background processes.

Notes
- Alle STO workloads are I/O-bound.
- These kinds of storage workloads are generic to all applications in all industries.

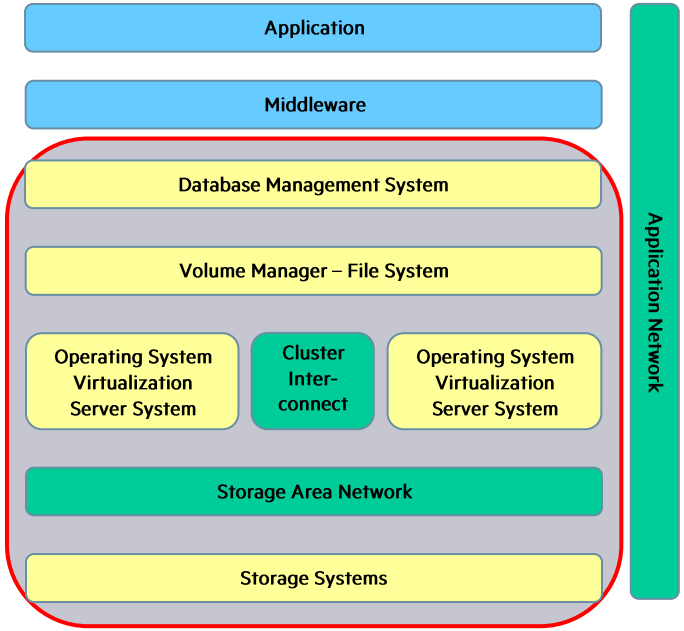
Server I/O Performance



Traditional I/O performance tools do not provide accurate I/O performance numbers for database services. The reason for that is the complexity of database I/O.

- Tools like vdbench, iometer, fio, Orion (Oracle I/O Numbers), etc. generate I/O system calls, but no further I/O processing
- These tools are unable to analyze the performance of computational storage arrays like Exadata and its offload functions
- I/O throughput and service time measured this way are not representative of database systems
- Oracle itself does not use Orion for performance testing of their Engineered Systems

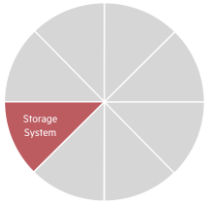
Database I/O Performance



The peakmarks® Software generates I/O load from Oracle, so-called SQL-generated I/O's.

- Most complex I/O operation
- Database buffer cache management
  - ✓ find a free slot
  - ✓ replace older blocks
  - ✓ local block access synchronization (buffer cache)
  - ✓ global block access synchronization (cache fusion)
  - ✓ database block consistency checks

Storage System Performance



Workload STO-READ – Sequential Read using a conventional Storage System

Single Instance

Run	Test	Workload	Wri [%]	Nodes	Jobs	DOP	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Phys reads total [IOPS]	IO time read [ms]	Phys reads total [MBps]	Phys reads FlCache [MBps]	FlCache read [%]	BuCache read [%]	Elapsed time [s]
4	1	STO-READ	0	1	1	4	2	1	1	98	0	3,636	0.912	3,619	3,619	100.00	0.00	173
	2	STO-READ	0	1	4	4	2	1	1	98	0	3,821	4.046	3,802	3,802	100.00	0.00	197
	3	STO-READ	0	1	8	4	3	2	1	97	0	6,222	4.938	6,204	6,204	100.00	0.00	181
	4	STO-READ	0	1	12	4	3	2	1	97	0	6,116	7.422	6,103	6,103	100.00	0.00	225
	5	STO-READ	0	1	16	4	3	2	1	97	0	6,282	9.992	6,269	6,269	100.00	0.00	219
	6	STO-READ	0	1	20	4	3	2	1	97	0	6,173	12.442	6,160	6,160	100.00	0.00	208

2-node Cluster

Run	Test	Workload	Wri [%]	Nodes	Jobs	DOP	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Phys reads total [IOPS]	IO time read [ms]	Phys reads total [MBps]	Phys reads FlCache [MBps]	FlCache read [%]	BuCache read [%]	Elapsed time [s]
11	1	STO-READ	0	2	2	4	2	1	1	98	0	6,628	1.015	6,531	6,531	100.00	0.00	175
	2	STO-READ	0	2	8	4	2	1	1	98	0	11,147	2.724	11,078	11,078	100.00	0.00	175
	3	STO-READ	0	2	16	4	3	2	1	97	0	11,682	5.297	11,631	11,631	100.00	0.00	183
	4	STO-READ	0	2	24	4	3	2	1	97	0	12,024	7.740	11,981	11,981	100.00	0.00	188
	5	STO-READ	0	2	32	4	3	2	1	97	0	12,118	10.107	12,067	12,067	100.00	0.00	193
	6	STO-READ	0	2	40	4	3	2	1	97	0	12,174	12.646	12,132	12,132	100.00	0.00	204

Workload STO-OFFLOAD – Sequential Read using offload technology

Single Instance

Run	Test	Workload	Wri [%]	Nodes	Jobs	DOP	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Phys reads total [IOPS]	IO time read [ms]	Phys reads total [MBps]	Phys reads FlCache [MBps]	FlCache read [%]	BuCache read [%]	Elapsed time [s]
4	33	STO-OFFLOAD	0	1	1	1	1	1	0	99	0	23,611	0.456	23,526	23,526	100.00	0.00	171
	34	STO-OFFLOAD	0	1	4	1	1	1	1	99	0	52,103	0.519	51,923	51,923	100.00	0.00	172
	35	STO-OFFLOAD	0	1	8	1	2	1	1	98	0	65,782	0.627	65,556	65,556	100.00	0.00	174
	36	STO-OFFLOAD	0	1	12	1	2	1	1	98	0	68,821	0.916	68,586	68,586	100.00	0.00	174
	37	STO-OFFLOAD	0	1	16	1	2	1	1	98	0	70,700	1.198	70,458	70,458	100.00	0.00	175
	38	STO-OFFLOAD	0	1	20	1	2	1	1	98	0	70,076	1.531	69,835	69,835	100.00	0.00	177

2-node Cluster

Run	Test	Workload	Wri [%]	Nodes	Jobs	DOP	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Phys reads total [IOPS]	IO time read [ms]	Phys reads total [MBps]	Phys reads FlCache [MBps]	FlCache read [%]	BuCache read [%]	Elapsed time [s]
11	33	STO-OFFLOAD	0	2	2	1	2	1	1	98	0	45,224	0.478	44,931	44,931	100.00	0.00	171
	34	STO-OFFLOAD	0	2	8	1	1	1	1	99	0	64,711	0.668	64,303	64,303	100.00	0.00	172
	35	STO-OFFLOAD	0	2	16	1	2	1	1	98	0	69,557	1.185	69,122	69,122	100.00	0.00	173
	36	STO-OFFLOAD	0	2	24	1	2	1	1	98	0	72,848	1.749	72,394	72,394	100.00	0.00	174
	37	STO-OFFLOAD	0	2	32	1	1	1	1	99	0	72,396	2.372	71,943	71,943	100.00	0.00	175

Note
The X8 data sheet specifies 75 GBps scan rate for QRHC.

Workload STO-RANDOM – Random I/O, 100% read

Single Instance

Run	Test	Workload	Wri [%]	Nodes	Jobs	DOP	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Phys reads total [IOPS]	IO time read [ms]	Phys reads total [MBps]	Phys reads FlCache [MBps]	FlCache read [%]	BuCache read [%]	Elapsed time [s]
4	65	STO-RANDOM	0	1	1	1	2	1	1	98	0	39,917	0.206	312	312	100.00	1.68	181
	67	STO-RANDOM	0	1	8	1	11	6	3	89	0	270,179	0.230	2,111	2,111	100.00	1.16	183
	69	STO-RANDOM	0	1	16	1	21	12	6	79	0	478,121	0.285	3,735	3,735	100.00	1.36	182
	71	STO-RANDOM	0	1	24	1	30	17	8	70	0	612,490	0.400	4,785	4,785	100.00	1.54	182
	73	STO-RANDOM	0	1	32	1	35	20	10	65	0	675,594	0.627	5,278	5,278	100.00	1.61	182
	75	STO-RANDOM	0	1	40	1	39	22	11	61	0	707,590	0.936	5,528	5,528	100.00	1.81	182
	79	STO-RANDOM	0	1	56	1	42	24	12	58	0	719,556	1.657	5,622	5,622	100.00	2.01	182

Note
The peak IOPS throughput means a high service time. Today (2024) , random I/O service time (8 kbyte block size) for database systems should be below 500 microseconds.

2-node Cluster

Run	Test	Workload	Wri [%]	Nodes	Jobs	DOP	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Phys reads total [IOPS]	IO time read [ms]	Phys reads total [MBps]	Phys reads FlCache [MBps]	FlCache read [%]	BuCache read [%]	Elapsed time [s]
11	65	STO-RANDOM	0	2	2	1	2	1	1	98	0	67,955	0.214	544	544	100.00	-0.33	183
	67	STO-RANDOM	0	2	16	1	13	7	4	87	0	476,058	0.259	3,743	3,743	100.00	1.65	182
	69	STO-RANDOM	0	2	32	1	21	12	6	79	0	745,340	0.334	5,855	5,855	100.00	1.88	182
	71	STO-RANDOM	0	2	48	1	24	11	8	76	0	1,168,859	0.581	9,179	9,179	100.00	1.71	183
	73	STO-RANDOM	0	2	64	1	26	12	9	74	0	1,287,733	0.857	10,122	10,122	100.00	2.11	182
	79	STO-RANDOM	0	2	112	1	32	15	11	68	0	1,399,140	2.008	11,001	11,001	100.00	2.85	182
	81	STO-RANDOM	0	2	128	1	32	15	11	68	0	1,410,819	2.489	11,094	11,094	100.00	3.00	182
	87	STO-RANDOM	0	2	176	1	35	17	12	65	0	1,425,543	3.955	11,212	11,212	100.00	3.10	182

Note
The X8 data sheet specifies 1,194,000 read IOPS for QRHC.

Workload STO-RANDOM – Random I/O, 80% read

Single Instance
ASM high redundancy

Run	Test	Workload	Wri [%]	Nodes	Jobs	DOP	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Phys reads total [IOPS]	IO time read [ms]	Phys reads total [MBps]	Phys reads FlCache [MBps]	BuCache read [%]	FlCache read [%]	Elapsed time [s]
4	97	STO-RANDOM	20	1	1	1	2	1	1	98	0	35,488	0.211	278	278	100.00	15.90	182
	99	STO-RANDOM	20	1	8	1	11	6	3	89	0	224,047	0.289	1,750	1,750	100.00	15.24	182
	101	STO-RANDOM	20	1	16	1	19	11	5	81	0	364,894	0.457	2,851	2,851	100.00	15.43	183
	103	STO-RANDOM	20	1	24	1	24	14	7	76	0	443,129	0.706	3,462	3,462	100.00	16.03	183
	105	STO-RANDOM	20	1	32	1	27	16	7	73	0	479,279	1.060	3,744	3,744	100.00	16.72	183
	107	STO-RANDOM	20	1	40	1	30	17	8	70	0	503,026	1.475	3,930	3,930	100.00	17.07	183
	111	STO-RANDOM	20	1	56	1	32	19	9	68	0	518,558	2.411	4,051	4,051	100.00	17.62	183

Note
The peak IOPS throughput means a high service time. Today (2024) , random I/O service time (8 kbyte block size) for database systems should be below 500 microseconds.

2-node Cluster
ASM high redundancy

Run	Test	Workload	Wri [%]	Nodes	Jobs	DOP	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Phys reads total [IOPS]	IO time read [ms]	Phys reads total [MBps]	Phys reads FlCache [MBps]	FlCache read [%]	BuCache read [%]	Elapsed time [s]
21	1	STO-RANDOM	20	2	2	1	2	1	1	98	0	66,545	0.227	524	524	100.00	17.99	182
	3	STO-RANDOM	20	2	16	1	10	5	3	90	0	423,127	0.336	3,319	3,319	100.00	15.93	183
	5	STO-RANDOM	20	2	32	1	17	9	5	83	0	662,583	0.534	5,208	5,208	100.00	16.78	183
	7	STO-RANDOM	20	2	48	1	21	11	6	79	0	781,877	0.817	6,153	6,153	100.00	17.56	183
	9	STO-RANDOM	20	2	64	1	24	13	7	76	0	845,871	1.105	6,651	6,651	100.00	18.18	184
	11	STO-RANDOM	20	2	80	1	26	14	7	74	0	883,718	1.471	6,951	6,951	100.00	18.69	184
	13	STO-RANDOM	20	2	96	1	27	15	8	73	0	904,607	1.862	7,116	7,116	100.00	18.91	183
	14	STO-RANDOM	20	2	102	1	27	15	8	73	0	909,024	1.998	7,151	7,151	100.00	18.93	183

Workload STO-SCATTER – Random Write foreground processes

Single Instance
ASM high redundancy

Run	Test	Workload	Wri [%]	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Phys writes total [dbps]	Phys writes total [IOPS]	Phys writes total [MBps]	Phys writes FlCache [MBps]	FlCache write [%]	Elapsed time [s]
4	129	STO-SCATTER	100	1	1	2	1	1	98	0	45,039	11,253	353	353	100.00	180
	130	STO-SCATTER	100	1	4	6	2	2	94	0	151,756	37,897	1,188	1,188	100.00	182
	131	STO-SCATTER	100	1	8	11	5	4	89	0	200,457	50,079	1,570	1,570	100.00	181
	132	STO-SCATTER	100	1	12	15	7	5	85	0	235,887	58,943	1,847	1,418	76.78	182
	133	STO-SCATTER	100	1	16	18	10	6	82	0	220,527	55,146	1,727	1,254	72.61	182

2-node Cluster
ASM high redundancy

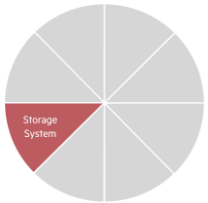
Run	Test	Workload	Wri [%]	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Phys writes total [dbps]	Phys writes total [IOPS]	Phys writes total [MBps]	Phys writes FlCache [MBps]	FlCache write [%]	Elapsed time [s]
11	129	STO-SCATTER	100	2	2	2	1	1	98	0	89,822	22,449	703	703	100.00	180
	130	STO-SCATTER	100	2	8	6	2	2	94	0	261,262	65,283	2,181	2,046	93.79	182
	131	STO-SCATTER	100	2	16	8	3	3	92	0	282,461	70,633	2,212	2,212	100.00	181
	132	STO-SCATTER	100	2	24	10	5	4	90	0	347,930	87,010	2,725	2,725	100.00	182
	133	STO-SCATTER	100	2	32	11	5	4	89	0	356,007	89,081	2,788	2,788	100.00	182
	134	STO-SCATTER	100	2	40	12	6	4	88	0	359,573	90,020	2,816	2,816	100.00	182

Notes

- The X8 data sheet specifies 1,088,000 write IOPS for QRHC
- The actual write rate to the storage system is 3x higher due to ASM's high-redundancy (3 x 359,573 = 1'078'719).

Storage System Performance

Selected comparative figures – For training purposes only



Workload STO-OFFLOAD – Sequential Read using offload technology

Exadata X5-2 QR HC
36c, 72t, Launch 2015

Run	Test	Workload	Wri [%]	Nodes	Jobs	DOP	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Phys reads total [IOPS]	IO time read [ms]	Phys reads total [MBps]	Phys reads FlCache [MBps]	FlCache read [%]	BuCache read [%]	Elapsed time [s]
6	56	STO-OFFLOAD	0	1	1	1	2	1	1	98	0	10,644	0.236	9,501	9,501	100.00	0.00	174
	57	STO-OFFLOAD	0	1	4	1	3	2	1	97	0	23,849	0.456	21,211	21,211	100.00	0.00	175
	58	STO-OFFLOAD	0	1	8	1	4	2	1	96	0	29,332	0.794	26,319	26,319	100.00	0.00	179
	59	STO-OFFLOAD	0	1	12	1	4	3	1	96	0	31,223	1.159	27,953	27,953	100.00	0.00	179
	60	STO-OFFLOAD	0	1	16	1	4	3	1	96	0	31,904	1.519	28,503	28,503	100.00	0.00	181

Exadata X9-2M QRHC
64c, 128t, Launch 2021

Run	Test	Workload	Wri [%]	Nodes	Jobs	DOP	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Phys reads total [dbps]	Phys reads total [IOPS]	Phys reads total [MBps]	FlCache read [%]	Elapsed time [s]
42	33	STO-OFFLOAD	0	1	1	1	1	1	0	99	0	6,321,983	49,577	49,390	100.00	285
	34	STO-OFFLOAD	0	1	4	1	1	1	0	99	0	16,678,679	130,774	130,302	100.00	285
	35	STO-OFFLOAD	0	1	8	1	2	1	0	98	0	19,073,872	149,537	149,015	100.00	286
	36	STO-OFFLOAD	0	1	12	1	2	1	1	98	0	19,626,132	153,903	153,329	100.00	287

Workload STO-RANDOM – Random I/O, 100% read

Exadata X5-2 QR HC
36c, 72t, Launch 2015

Run	Test	Workload	Wri [%]	Nodes	Jobs	DOP	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Phys reads total [IOPS]	IO time read [ms]	Phys reads total [MBps]	Phys reads FlCache [MBps]	BuCache read [%]	FlCache read [%]	Elapsed time [s]
35	1	STO-RANDOM	0	1	1	1	3	2	1	97	0	21,612	0.292	170	170	14.52	100.00	183
	2	STO-RANDOM	0	1	16	1	26	18	5	74	0	279,746	0.346	2,186	2,186	19.24	100.00	183
	3	STO-RANDOM	0	1	32	1	47	34	9	53	0	440,057	0.481	3,438	3,438	19.66	100.00	183
	4	STO-RANDOM	0	1	48	1	63	46	12	37	0	523,620	0.664	4,091	4,091	19.63	100.00	183
	5	STO-RANDOM	0	1	64	1	78	56	15	22	0	624,536	1.108	4,879	4,879	19.70	100.00	183
	6	STO-RANDOM	0	1	68	1	80	57	16	20	0	640,685	1.249	5,005	5,005	19.66	100.00	183

Exadata X9-2M QRHC
64c, 128t, Launch 2021

Run	Test	Workload	Wri [%]	Nodes	Jobs	DOP	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Phys reads total [dbps]	Phys reads total [IOPS]	IO time read [ms]	Phys reads total [MBps]	BuCache read [%]	FlCache read [%]	Elapsed time [s]
44	65	STO-RANDOM	0	1	1	1	1	1	0	99	0	64,222	62,520	0.137	502	58.33	32.73	301
	67	STO-RANDOM	0	1	8	1	7	5	1	93	0	573,934	569,553	0.056	4,484	14.06	27.84	301
	69	STO-RANDOM	0	1	16	1	13	9	2	87	0	1,028,768	1,024,525	0.060	8,037	9.01	27.27	302
	71	STO-RANDOM	0	1	24	1	19	13	3	81	0	1,384,287	1,380,049	0.069	10,815	7.34	27.76	302
	73	STO-RANDOM	0	1	32	1	25	18	5	75	0	1,755,477	1,751,230	0.074	13,715	6.35	28.05	302
	75	STO-RANDOM	0	1	40	1	31	23	6	69	0	2,066,438	2,062,184	0.081	16,144	5.82	27.73	302
	77	STO-RANDOM	0	1	48	1	38	28	7	62	0	2,278,435	2,274,177	0.093	17,800	5.60	28.23	302

Note
Exadata X9M-2 uses an additional PMEM cache; these results are achieved after the PMEM cache warmup.

Workload STO-RANDOM – Random I/O, 100% read

Run	Test	Workload	Wri [%]	Nodes	Jobs	DOP	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Phys reads total [dbps]	Phys reads total [IOPS]	IO time read [ms]	Phys reads total [MBps]	BuCache read [%]	FlCache read [%]	Elapsed time [s]
45	1	STO-RANDOM	0	4	4	1	3	2	1	97	0	150,316	146,673	0.046	1,174	4.64	95.49	302
	2	STO-RANDOM	0	4	24	1	10	7	2	90	0	843,164	839,192	0.058	6,587	1.66	100.00	303
	3	STO-RANDOM	0	4	48	1	17	13	3	83	0	1,447,285	1,447,100	0.063	11,307	2.25	100.00	302
	4	STO-RANDOM	0	4	72	1	25	19	4	75	0	2,363,664	2,363,645	0.068	18,466	2.15	100.00	302
	5	STO-RANDOM	0	4	96	1	30	23	4	70	0	2,936,253	2,936,250	0.078	22,939	2.23	100.00	303
	6	STO-RANDOM	0	4	120	1	35	29	4	65	0	3,998,866	3,998,865	0.099	31,241	2.09	100.00	302
	7	STO-RANDOM	0	4	144	1	42	34	5	58	0	4,130,953	4,130,950	0.116	32,273	2.25	100.00	302
	8	STO-RANDOM	0	4	168	1	46	39	5	54	0	4,156,891	4,156,889	0.129	32,476	2.53	100.00	302
	9	STO-RANDOM	0	4	172	1	47	39	5	53	0	4,112,775	4,112,773	0.128	32,131	2.62	100.00	302
	10	STO-RANDOM	0	4	172	1	47	39	5	53	0	4,147,268	4,147,267	0.129	32,401	2.56	100.00	302

Note
Exadata X8M-2 uses an additional PMEM cache; these results are achieved after the PMEM cache warmup.

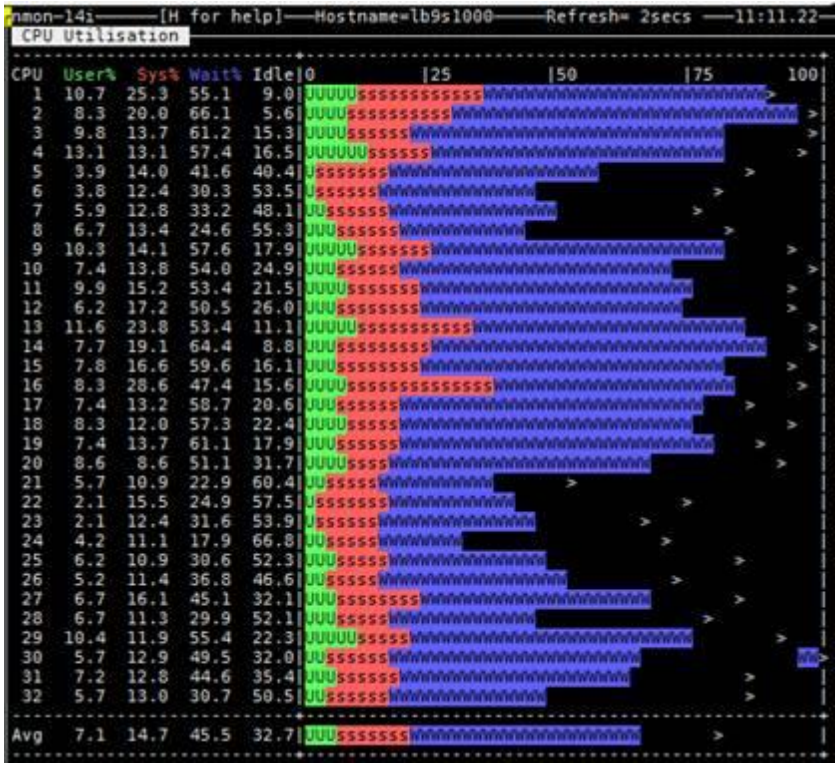
I/O Storage Path Optimization

Default configurations are not good enough

- I/O wait should be less than 10%

Example with PCI attached flash NVMe

- Database block size 8 Kbyte
- iowait > 45%



I/O Storage Path Optimization

Performance-by-Design approach

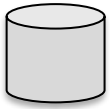
- Each layer optimized

Example with PCI attached flash NVMe

- 1.1 Mio IOPS with Oracle
- Database block size 8 KByte
- CPU completely utilized
- kernel mode 49%
- iowait < 1%

Storage System Performance

As per design, tiered storage cannot provide predictable and consistent performance



Database, 16 Tbyte

Database Buffer Cache, 2 x 768 Gbyte

XRMEM Cache, 3 x 1.25 = 3.75 Tbyte, avg access time per database block 20 µs

Flash Cache, 12 x 6.8 = 81.6 Tbyte, avg access time per database block 200 µs

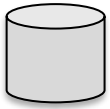
Usable HDD Storage, 36 x 22 / 3 = 264 Tbyte, avg access time per database block 15 ms

Example

16 Tbyte OLTP Database on Exadata X10M Quarter Rack High Capacity with 1.5 Tbyte RAM per database server; 50% of RAM for SGA.

Storage System Performance

As per design, tiered storage cannot provide predictable and consistent performance

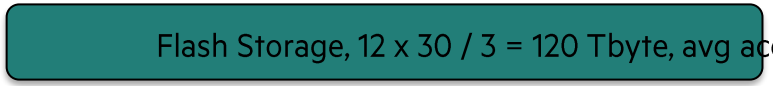


Database, 16 Tbyte

Database Buffer Cache, 2 x 768 Gbyte

XRMEM Cache, 3 x 1.25 = 3.75 Tbyte, avg access time per database block 20 µs

Flash Cache, 12 x 6.8 = 81.6 Tbyte, avg access time per database block 200 µs



Flash Storage, 12 x 30 / 3 = 120 Tbyte, avg access time per database block >200 µs

Example

16 Tbyte OLTP Database on Exadata X10M Quarter Rack Extreme Flash with 1.5 Tbyte RAM per database server; 50% of RAM for SGA.

peakmarks

Stop guessing. Start measuring.

Workloads to determine the Log Writer Performance

Motivation

Log Writer processes are primarily responsible for transaction management and database consistency in case of failures.

These processes are critical to overall Oracle performance, especially with a high transaction load.

The goal is to

- Optimize throughput and latency of REDO log writers
- Validate the impact of Oracle Data Guard on local transaction performance
- Validate the impact of several other factors on log writer performance, such as data deduplication, data compression, usage of ASM redundancy level, etc.

Log Writer Workloads

Key Performance Metrics

- SQL commit throughput in transactions per second [tps]
- SQL commit latency in milliseconds [ms]
- Log writer throughput in megabyte per second [MBps]

Log Writer Workloads for different transaction sizes

Description

Workload	Measurement Unit	Action
LGWR-LAT1	[tps] [ms]	Small-sized transaction; workload uses COMMIT WRITE WAIT IMMEDIATE. This workload shows the maximum commit rate and the minimum commit latency for small-sized transactions generating around 1 kByte redo data per transaction.
LGWR-LAT25	[tps] [ms]	Medium-sized transaction; workload uses COMMIT WRITE WAIT IMMEDIATE. This workload shows the maximum commit rate and the minimum commit latency for medium-sized transactions generating around 25 kByte redo data per transaction.
LGWR-LAT125	[tps] [ms]	Large-sized transaction size; workload uses COMMIT WRITE WAIT IMMEDIATE. This workload shows the maximum commit rate and the minimum commit latency for larger-sized transactions generating around 125 kByte redo data per transaction. This transaction size is typical for payment systems in banking applications.
LGWR-THR	[MBps]	Huge transaction; workload uses COMMIT WRITE NOWAIT BATCH. This workload shows the maximum redo data written by log writer processes.

Notes

- These kinds of transactions are generic to all applications in all industries.
- The actual REDO data volume per transaction can vary depending on the database size and the usage of RAC technology. Therefore, the exact value of this metric is explicitly shown in the peakmarks reports.

Log Writer Performance

Log Writer Performance

Workload LGWR-LAT1 – Small transactions, max commit throughput, low commit latency

Single Instance
ASM high redundancy

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	#prc LGWR	Commit thrput [tps]	Commit latency [ms]	LogFile sync [ms]	REDO data [rbps]	REDO data [IOPS]	REDO data [MBps]	REDO data [kBpt]	FlCache write [%]	Elapsed time [s]
5	1	LGWR-LAT1	1	1	1	1	0	99	0	9	4,074	0.245	0.175	12,761	4,121	6	1.383	99.91	180
	5	LGWR-LAT1	1	16	5	4	1	95	0	9	31,745	0.501	0.399	104,918	4,200	43	1.397	100.00	182
	9	LGWR-LAT1	1	32	8	6	1	92	0	9	46,056	0.695	0.549	153,440	3,114	63	1.400	100.00	181
	13	LGWR-LAT1	1	48	9	8	1	91	0	9	60,490	0.786	0.630	194,777	2,750	82	1.392	100.00	183
	17	LGWR-LAT1	1	64	13	10	2	87	0	9	73,004	0.872	0.698	229,257	2,492	99	1.390	100.00	183
	21	LGWR-LAT1	1	80	15	13	2	85	0	9	82,733	0.963	0.768	254,905	2,271	112	1.386	100.00	182
	25	LGWR-LAT1	1	96	18	15	2	82	0	9	92,293	1.034	0.834	280,277	2,096	125	1.383	100.00	182

2-node cluster
ASM high redundancy

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	#prc LGWR	Commit thrput [tps]	Commit latency [ms]	LogFile sync [ms]	REDO data [rbps]	REDO data [IOPS]	REDO data [MBps]	REDO data [kBpt]	FlCache write [%]	Elapsed time [s]
12	1	LGWR-LAT1	2	2	2	1	1	98	0	18	6,165	0.323	0.234	19,487	6,258	8	1.396	99.98	182
	3	LGWR-LAT1	2	32	6	4	1	94	0	18	55,534	0.572	0.461	197,131	10,794	77	1.417	100.00	183
	5	LGWR-LAT1	2	64	8	6	2	92	0	18	88,145	0.722	0.584	300,785	9,020	121	1.404	100.00	182
	7	LGWR-LAT1	2	96	11	9	2	89	0	18	110,529	0.863	0.695	364,364	7,960	151	1.401	100.00	183
	9	LGWR-LAT1	2	128	14	11	2	86	0	18	131,324	0.969	0.781	421,304	7,288	179	1.395	100.00	182
	11	LGWR-LAT1	2	160	16	13	2	84	0	18	153,826	1.031	0.833	478,342	5,434	209	1.388	87.62	183
	13	LGWR-LAT1	2	192	18	14	2	82	0	18	172,065	1.104	0.882	526,155	3,995	234	1.390	100.00	184
	15	LGWR-LAT1	2	224	20	16	3	80	0	18	183,894	1.187	0.937	558,979	3,749	250	1.393	90.58	186
	17	LGWR-LAT1	2	256	23	19	3	77	0	18	195,328	1.276	1.002	590,846	3,570	266	1.394	77.85	186
	19	LGWR-LAT1	2	288	25	21	3	75	0	18	203,230	1.356	1.073	610,915	3,350	276	1.392	100.00	189
	21	LGWR-LAT1	2	320	27	22	3	73	0	18	209,735	1.443	1.152	627,172	3,097	285	1.391	100.00	192

Log Writer Performance

Workload LGWR-LAT25 – Medium-sized transactions, max commit throughput, low commit latency

Single Instance
ASM high redundancy

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	#prc LGWR	Commit thrput [tps]	Commit latency [ms]	LogFile sync [ms]	REDO data [rbps]	REDO data [IOPS]	REDO data [MBps]	REDO data [kBpt]	FlCache write [%]	Elapsed time [s]
5	33	LGWR-LAT25	1	1	2	1	1	98	0	9	1,491	0.671	0.279	86,824	1,547	41	27.941	100.00	182
	35	LGWR-LAT25	1	8	6	4	1	95	0	9	7,923	1.005	0.594	434,862	2,540	203	26.220	100.00	182
	37	LGWR-LAT25	1	16	8	6	1	92	0	9	11,221	1.419	0.946	576,299	1,637	269	24.577	100.00	182
	39	LGWR-LAT25	1	24	9	7	1	91	0	9	12,320	1.943	1.398	621,435	1,222	291	24.155	85.26	181
	41	LGWR-LAT25	1	32	10	8	1	91	0	9	13,077	2.439	1.841	647,394	1,032	303	23.728	100.00	182
	43	LGWR-LAT25	1	40	10	8	1	90	0	9	13,792	2.896	2.236	662,045	899	310	23.028	100.00	182

2-node cluster
ASM high redundancy

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	#prc LGWR	Commit thrput [tps]	Commit latency [ms]	LogFile sync [ms]	REDO data [rbps]	REDO data [IOPS]	REDO data [MBps]	REDO data [kBpt]	FlCache write [%]	Elapsed time [s]
12	33	LGWR-LAT25	2	2	2	1	1	98	0	18	2,513	0.795	0.346	140,985	2,623	66	26.904	100.00	181
	34	LGWR-LAT25	2	16	6	4	1	94	0	18	14,430	1.103	0.627	783,458	4,669	366	25.962	100.00	182
	35	LGWR-LAT25	2	32	9	6	1	91	0	18	21,224	1.500	0.937	1,085,797	5,023	505	24.388	94.45	182
	36	LGWR-LAT25	2	48	11	8	2	89	0	18	25,362	1.885	1.223	1,287,709	5,078	599	24.165	91.30	182
	37	LGWR-LAT25	2	64	12	9	2	88	0	18	27,291	2.333	1.527	1,379,385	4,247	642	24.094	99.99	182
	38	LGWR-LAT25	2	80	13	10	2	87	0	18	29,219	2.722	1.803	1,419,974	3,741	662	23.191	92.39	182
	. . .																		
	46	LGWR-LAT25	2	208	20	16	3	80	0	18	34,348	5.951	3.197	1,628,515	2,367	763	22.745	96.69	185
	47	LGWR-LAT25	2	224	21	17	3	79	0	18	35,495	6.161	3.139	1,678,448	2,383	787	22.690	93.76	186
	48	LGWR-LAT25	2	240	21	17	3	79	0	18	36,241	6.499	3.236	1,716,774	2,393	805	22.736	96.83	184
	49	LGWR-LAT25	2	256	22	18	3	78	0	18	37,078	6.664	3.184	1,767,036	2,477	828	22.876	100.00	188

Log Writer Performance

Workload LGWR-LAT125 – Large transactions, max commit throughput, low commit latency

Single Instance
ASM high redundancy

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	#prc LGWR	Commit thrput [tps]	Commit latency [ms]	LogFile sync [ms]	REDO data [rbps]	REDO data [IOPS]	REDO data [MBps]	REDO data [kBpt]	FlCache write [%]	Elapsed time [s]
5	65	LGWR-LAT125	1	1	2	1	0	98	0	9	442	2.264	0.594	128,767	510	61	140.747	100.00	182
	67	LGWR-LAT125	1	8	7	6	1	93	0	9	2,380	3.352	1.432	628,442	1,137	295	127.115	91.17	182
	69	LGWR-LAT125	1	16	10	8	1	90	0	9	2,967	5.374	2.745	735,389	710	346	119.384	93.43	182
	71	LGWR-LAT125	1	24	12	10	1	88	0	9	3,621	6.597	3.426	858,429	552	404	114.318	100.00	182
	73	LGWR-LAT125	1	32	14	11	2	86	0	9	3,992	7.980	4.515	932,410	500	439	112.703	90.00	182
	75	LGWR-LAT125	1	40	14	11	2	86	0	9	4,189	9.502	5.100	977,609	386	461	112.679	90.47	182
	77	LGWR-LAT125	1	48	15	12	2	85	0	9	4,377	10.917	5.578	1,013,383	337	478	111.823	86.62	182
	79	LGWR-LAT125	1	56	16	13	2	84	0	9	4,593	12.146	6.302	1,032,064	328	487	108.542	100.00	182

2-node cluster
ASM high redundancy

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	#prc LGWR	Commit thrput [tps]	Commit latency [ms]	LogFile sync [ms]	REDO data [rbps]	REDO data [IOPS]	REDO data [MBps]	REDO data [kBpt]	FlCache write [%]	Elapsed time [s]
12	65	LGWR-LAT125	2	2	2	1	1	98	0	18	764	2.608	0.667	216,567	888	102	136.954	100.00	181
	66	LGWR-LAT125	2	16	9	7	1	91	0	18	4,910	3.242	1.072	1,319,552	4,452	619	129.016	100.00	182
	67	LGWR-LAT125	2	32	12	9	2	88	0	18	6,440	4.953	2.096	1,625,758	3,213	762	121.222	96.26	182
	68	LGWR-LAT125	2	48	15	12	2	85	0	18	7,629	6.254	2.632	1,824,650	2,691	857	114.963	100.00	182
	69	LGWR-LAT125	2	64	18	14	2	82	0	18	8,784	7.254	3.061	2,062,175	2,442	969	112.942	99.99	182
	70	LGWR-LAT125	2	80	20	16	2	80	0	18	9,458	8.405	3.706	2,211,406	2,082	1,040	112.590	92.25	182
	71	LGWR-LAT125	2	96	21	17	3	79	0	18	10,053	9.483	4.083	2,349,924	1,898	1,106	112.626	100.00	182
	72	LGWR-LAT125	2	112	22	18	3	78	0	18	10,368	10.727	4.423	2,406,002	1,815	1,132	111.847	92.74	182
	. . .																		
	77	LGWR-LAT125	2	192	24	19	3	76	0	18	11,583	16.378	5.499	2,594,696	1,520	1,222	108.050	100.00	183

Workload LGWR-THR – Maximum log writer throughput

Single Instance
ASM high redundancy

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	#prc LGWR	Commit thrput [tps]	Commit latency [ms]	LogFile sync [ms]	REDO data [rbps]	REDO data [IOPS]	REDO data [MBps]	REDO data [kBpt]	FlCache write [%]	Elapsed time [s]
5	97	LGWR-THR	1	1	2	2	0	98	0	9	321	3.110	0.577	191,519	115	91	288.742	71.45	180
	98	LGWR-THR	1	4	7	5	1	93	0	9	1,133	3.523	2.800	646,194	420	305	275.788	100.00	181
	99	LGWR-THR	1	8	10	8	1	90	0	9	1,804	4.416	4.313	988,157	487	466	264.721	100.00	182
	100	LGWR-THR	1	12	17	13	2	83	0	9	2,263	5.274	8.242	1,200,251	352	567	256.588	100.00	182
	101	LGWR-THR	1	16	17	13	2	83	0	9	2,571	6.197	11.501	1,328,855	461	628	249.973	100.00	182

2-node cluster
ASM high redundancy

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	#prc LGWR	Commit thrput [tps]	Commit latency [ms]	LogFile sync [ms]	REDO data [rbps]	REDO data [IOPS]	REDO data [MBps]	REDO data [kBpt]	FlCache write [%]	Elapsed time [s]
12	97	LGWR-THR	2	2	2	1	1	98	0	18	488	4.092	1.348	270,752	468	128	268.396	100.00	181
	98	LGWR-THR	2	16	9	7	1	91	0	18	3,026	5.261	1.693	1,614,870	2,279	760	257.316	100.00	181
	99	LGWR-THR	2	32	14	11	2	86	0	18	4,501	7.090	3.784	2,302,606	1,883	1,085	246.868	100.00	181
	100	LGWR-THR	2	48	16	13	2	84	0	18	4,986	9.615	6.496	2,437,707	1,578	1,150	236.091	100.00	181
	101	LGWR-THR	2	64	18	14	2	82	0	18	5,154	12.388	8.322	2,491,403	1,446	1,175	233.485	100.00	182
	102	LGWR-THR	2	80	18	14	3	82	0	18	5,281	15.100	10.556	2,497,715	1,366	1,178	228.475	100.00	181
	103	LGWR-THR	2	96	19	15	3	81	0	18	5,502	17.362	12.947	2,556,880	1,401	1,206	224.497	95.20	182
	104	LGWR-THR	2	112	20	16	3	80	0	18	5,558	20.076	14.167	2,576,229	1,313	1,215	223.921	100.00	181

Log Writer Performance

Selected comparative figures – For training purposes only

Log writer latency is important for transactional database systems

Log Writer Performance

Workload LGWR-LAT1 – Small transactions, max commit throughput, low commit latency

Exadata X5-2 QR HC
36c, 72t, Launch 2015
ASM normal redundancy

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	#prc LGWR	Commit thrput [tps]	Commit latency [ms]	LogFile sync [ms]	REDO data [rbps]	REDO data [IOPS]	REDO data [MBps]	REDO data [kBpt]	FlCache write [%]	Elapsed time [s]
15	171	LGWR-LAT1	1	1	2	1	1	98	0	9	1,936	0.516	0.346	6,081	1,975	2	1.256	98.61	181
	175	LGWR-LAT1	1	24	12	10	2	88	0	9	34,809	0.687	0.492	104,645	4,893	43	1.260	99.07	183
	179	LGWR-LAT1	1	48	21	17	2	79	0	9	56,187	0.849	0.619	163,642	4,252	69	1.255	99.10	183
	183	LGWR-LAT1	1	72	27	23	3	73	0	9	72,820	0.985	0.738	203,446	3,602	89	1.246	98.84	182
	187	LGWR-LAT1	1	96	28	24	3	72	0	9	86,251	1.110	0.846	234,961	1,966	105	1.248	98.40	181
	191	LGWR-LAT1	1	120	34	29	3	66	0	9	99,227	1.204	0.932	266,360	1,803	120	1.243	98.02	183
	193	LGWR-LAT1	1	132	36	31	4	64	0	9	102,513	1.282	1.004	274,091	1,692	124	1.243	98.50	183

Exadata X9-2M QRHC
64c, 128t, Launch 2021
ASM normal redundancy

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Commit throughput [tps]	Commit latency [ms]	REDO data [rbps]	REDO data [IOPS]	REDO data [MBps]	REDO data [kBpt]	LogFile sync [ms]	FlCache write [%]	Elapsed time [s]
60	1	LGWR-LAT1	1	1	1	1	0	99	0	11,202	0.089	48,535	11,213	21	1.920	0.029	19.13	302
	2	LGWR-LAT1	1	4	3	3	1	97	0	40,181	0.099	175,761	27,723	74	1.890	0.034	20.79	302
	3	LGWR-LAT1	1	8	6	5	1	94	0	79,227	0.101	347,887	34,581	146	1.890	0.035	22.12	302
	4	LGWR-LAT1	1	12	8	7	1	92	0	103,916	0.115	459,456	31,870	192	1.890	0.039	22.97	317
	5	LGWR-LAT1	1	16	10	9	1	90	0	138,660	0.115	612,493	31,390	256	1.890	0.044	23.05	302
	6	LGWR-LAT1	1	20	12	11	1	88	0	162,252	0.123	715,840	28,614	299	1.890	0.049	23.20	302
	7	LGWR-LAT1	1	24	14	12	1	86	0	173,856	0.137	769,589	26,790	321	1.890	0.058	23.70	302
	8	LGWR-LAT1	1	28	16	14	1	84	0	189,576	0.147	835,184	22,759	350	1.890	0.059	24.05	301
	9	LGWR-LAT1	1	32	16	14	1	84	0	193,878	0.165	850,051	19,667	358	1.890	0.071	24.68	301

Note

- Log writer latency benefits from Exadata X9 persistent memory (PMEM).

Log Writer Performance

Lenvov x86, 4 sockets Launch 2021
Netapp AFA, dedup on
ASM external redundancy

Workload LGWR-LAT1 – Small transactions, max commit throughput, low commit latency

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Commit throughput [tps]	Commit latency [ms]	REDO data [rbps]	REDO data [IOPS]	REDO data [MBps]	REDO data [kBpt]	LogFile sync [ms]	FlCache write [%]	Elapsed time [s]
61	1	LGWR-LAT1	1	1	1	0	0	99	0	4,106	0.244	17,864	4,115	8	2.000	0.172	0.00	302
	3	LGWR-LAT1	1	16	1	1	0	99	0	14,113	1.131	61,048	1,809	26	1.890	0.985	0.00	301
	5	LGWR-LAT1	1	32	1	1	0	99	0	19,172	1.669	81,302	1,224	35	1.870	1.448	0.00	302
	7	LGWR-LAT1	1	48	2	2	0	98	0	26,707	1.797	110,145	1,138	49	1.880	1.585	0.00	300
	9	LGWR-LAT1	1	64	3	2	0	97	0	34,018	1.881	138,081	1,087	62	1.870	1.650	0.00	303
	11	LGWR-LAT1	1	80	4	3	0	96	0	41,232	1.940	165,525	1,055	75	1.860	1.685	0.00	301
	13	LGWR-LAT1	1	96	5	5	0	95	0	47,983	2.000	191,405	1,023	87	1.860	1.733	0.00	302
	15	LGWR-LAT1	1	112	7	7	0	93	0	55,491	2.018	220,026	1,015	101	1.860	1.703	0.00	301
	17	LGWR-LAT1	1	128	7	7	0	93	0	61,814	2.069	244,620	991	112	1.860	1.745	0.00	302
	19	LGWR-LAT1	1	144	9	8	0	91	0	68,314	2.099	270,010	975	124	1.860	1.776	0.00	301
	21	LGWR-LAT1	1	160	10	10	0	90	0	74,449	2.148	295,137	952	136	1.870	1.786	0.00	301

Note

- Log writer latency suffers from storage systems data reduction technologies.

peakmarks

Stop guessing. Start measuring.

Workloads to determine the Database Writer Performance

Motivation

Database writer performance for buffer management is critical to overall database performance, especially for transaction systems with a high rate of updates and applications that primarily load data via the buffer cache.

The goal is to

- Optimize the number of database writer processes
- Validate the impact of several factors on database writer performance, e.g., ASM redundancy level

Database Writer Workloads

Key Performance Metrics

- Database writer throughput in database blocks per second [dbps]

Database Writer Workloads

Description

Workload	Measurement Unit	Action
DBWR-THR	[dbps] [MBps]	Massive block changes in the buffer cache; workload uses COMMIT WRITE NOWAIT BATCH. This workload shows the maximum number of changed blocks written back to the storage system by database writer processes. The number of database writer processes is a configurable Oracle instance parameter. The workload DBWR-THR writes the blocks by background processes; in contrast, the STO-SCATTER workload writes the blocks by foreground processes.

Database Writer Performance

Workload DBWR-THR – Maximum database writer throughput

Single Instance
ASM high redundancy

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	#prc DBWR	Phys writes total [dbps]	Phys writes total [IOPS]	Phys writes total [MBps]	REDO data [MBps]	Phys writes FlCache [MBps]	FlCache write [%]	Elapsed time [s]
6	312	DBWR-THR	1	1	3	2	0	97	0	12	9,582	2,290	84	9	76	90.72	181
	313	DBWR-THR	1	4	9	7	1	91	0	12	40,428	13,938	349	32	323	92.56	182
	314	DBWR-THR	1	8	17	14	2	83	0	12	86,579	30,474	735	57	703	95.56	183
	315	DBWR-THR	1	12	23	20	2	77	0	12	126,578	42,584	1,075	83	1,029	95.69	182
	316	DBWR-THR	1	16	31	26	3	69	0	12	155,222	67,526	1,322	105	1,251	94.62	183

2-node cluster
ASM high redundancy

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	#prc DBWR	Phys writes total [dbps]	Phys writes total [IOPS]	Phys writes total [MBps]	REDO data [MBps]	Phys writes FlCache [MBps]	FlCache write [%]	Elapsed time [s]
13	1	DBWR-THR	2	2	4	2	1	96	0	24	39,488	27,076	373	62	373	100.00	181
	2	DBWR-THR	2	8	9	6	2	91	0	24	197,333	34,425	1,778	227	1,778	100.00	182
	3	DBWR-THR	2	16	16	11	3	84	0	24	300,089	83,328	2,928	366	2,643	90.27	182
	4	DBWR-THR	2	24	23	16	5	77	0	24	328,586	182,857	3,047	462	2,735	89.77	183
	5	DBWR-THR	2	32	25	16	6	75	0	24	304,372	241,604	2,858	462	2,702	94.54	182

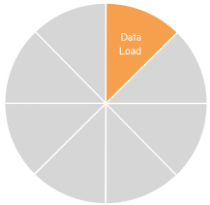
Note

- This configuration used the default configuration of 12 database writer processes per instance.
- The actual write rate to the storage system is 3x higher due to ASM's high redundancy.

peakmarks

Stop guessing. Start measuring.

Workloads to determine the Data Load Performance



Motivation

Data load performance capabilities are essential for all kinds of database applications:

- transaction processing systems
- data warehouse systems and data analytics systems

The amount of data is increasing, and the time window for loading the data decreases.

The goal is to

- Optimize data load throughput
- Validate the impact of several factors on data load performance

Data Load Workloads

Key Performance Metrics

- SQL data load rate in megabytes per second [MBps] or in rows per second [rps]

Data Load Workloads

Description

Workload	Measurement Unit	Action
DL-BUFFER	[MBps] [rps]	Insert of program generated data via default buffer cache into Oracle tables, includes maintenance of 1 unique and 2 non-unique indexes, uses COMMIT WRITE WAIT IMMEDIATE. The workload parameter specifies the number of rows per INSERT; the following values are supported {1, 2, 3, 4, 5, 10, 20, 25, 50, 100}. This data load workload is mainly used in Transaction Processing systems.
DL-DIRECT	[MBps] [TBph]	Insert data copying from other tables (CTAS technology) bypassing the buffer cache, includes maintenance of 1 unique index, uses COMMIT WRITE NOWAIT BATCH and NOLOGGING. This data load workload is mainly used in Data Warehouse systems.

DL-STREAM	[MBps]
	[rps]
		Insert data using a memory-optimized buffer pool, including maintenance of 1 unique index, using PL/SQL procedure to flush the buffer pool for memory-optimized tables. The workload parameter specifies the number of rows per INSERT; the following values are supported {1, 2, 3, 4, 5};
		This data load workload is mainly used in Data Analytics and Internet-of-things (IOT) systems.
		This workload is only available if the memory-optimized row store for fast ingest is configured. This feature was introduced in Oracle 19c. Locking problems occur occasionally in 19c with this workload . . .
		`SYS@ORA19C SQL> shutdown immediate`
		`ORA-62187: timeout occurred while waiting for lock to flush object 165013 data object 165179`

Notes

- These data load workloads are generic to all applications in all industries.
- The data load throughput depends mainly on the method (buffered, direct, streamed) and the transaction size (number of rows between commits), which is configurable.
- The technology used for workload DL-STREAM (memory-optimized tables) is available only on Oracle Exadata Systems; set _exadata_feature_on = true to use this feature on other platforms.

Data Load Performance

Workload DL-BUFFER – Data load via buffer cache

Single Instance
ASM high redundancy

Run	Test	Workload	TX size [rpt]	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Load rate total [rps]	Load rate total [MBps]	Phys writes total [IOPS]	Phys writes total [MBps]	REDO data [MBps]	FlCache write [%]	Elapsed time [s]
7	1	DL-BUFFER	5	1	1	2	1	0	98	0	13,243	4	2,971	24	13	99.99	182
	5	DL-BUFFER	5	1	16	7	5	1	93	0	113,484	32	8,516	232	106	100.00	182
	9	DL-BUFFER	5	1	32	10	8	2	90	0	162,879	47	14,351	347	147	100.00	181
	13	DL-BUFFER	5	1	48	13	11	2	87	0	192,029	55	12,846	443	177	100.00	182
	17	DL-BUFFER	5	1	64	17	13	2	83	0	213,652	61	14,987	528	198	100.00	182
	21	DL-BUFFER	5	1	80	18	14	2	82	0	233,276	67	17,176	603	214	100.00	182
	25	DL-BUFFER	5	1	96	19	16	3	81	0	241,635	69	19,886	674	223	100.00	182

2-node cluster
ASM high redundancy

Run	Test	Workload	TX size [rpt]	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Load rate total [rps]	Load rate total [MBps]	Phys writes total [IOPS]	Phys writes total [MBps]	REDO data [MBps]	FlCache write [%]	Elapsed time [s]
14	1	DL-BUFFER	5	2	2	2	1	1	98	0	25,151	7	5,796	46	24	100.00	182
	3	DL-BUFFER	5	2	32	7	5	1	93	0	204,245	58	14,866	398	193	100.00	183
	5	DL-BUFFER	5	2	64	11	8	2	89	0	303,210	87	18,297	617	275	100.00	182
	7	DL-BUFFER	5	2	96	13	10	2	87	0	377,273	108	19,309	828	347	100.00	182
	9	DL-BUFFER	5	2	128	17	13	2	83	0	420,796	120	23,709	988	392	100.00	182
	11	DL-BUFFER	5	2	160	20	16	3	80	0	452,938	130	28,397	1,123	416	100.00	182
	13	DL-BUFFER	5	2	192	23	18	3	77	0	476,194	136	32,492	1,248	439	100.00	184

Data Load Performance

Workload DL-DIRECT – Data load bypassing the buffer cache

Single Instance
ASM high redundancy

Run	Test	Workload	TX size [rpt]	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Load rate total [rps]	Load rate total [MBps]	Phys writes total [IOPS]	Phys writes total [MBps]	REDO data [MBps]	FlCache write [%]	Elapsed time [s]
7	33	DL-DIRECT	125,000	1	1	2	1	0	98	0	514,291	147	1,048	225	34	99.99	182
	34	DL-DIRECT	125,000	1	4	5	4	1	95	0	1,892,914	542	3,521	827	124	100.00	181
	35	DL-DIRECT	125,000	1	8	8	7	1	92	0	3,054,551	874	7,001	1,338	200	100.00	181
	36	DL-DIRECT	125,000	1	12	9	8	1	91	0	3,438,067	984	8,908	1,512	225	100.00	182
	37	DL-DIRECT	125,000	1	16	10	8	1	90	0	3,507,254	1,003	12,293	1,756	230	88.43	181
	38	DL-DIRECT	125,000	1	20	11	9	1	89	0	3,690,651	1,056	16,658	1,707	242	95.47	182
	39	DL-DIRECT	125,000	1	24	13	10	2	87	0	3,824,905	1,094	17,680	1,708	251	94.93	182

2-node cluster
ASM high redundancy

Run	Test	Workload	TX size [rpt]	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Load rate total [rps]	Load rate total [MBps]	Phys writes total [IOPS]	Phys writes total [MBps]	REDO data [MBps]	FlCache write [%]	Elapsed time [s]
10	376	DL-DIRECT	125,000	2	2	2	2	1	98	0	637,699	182	1,406	285	41	85.68	181
	377	DL-DIRECT	125,000	2	8	7	6	1	93	0	2,107,202	603	4,954	959	137	89.02	182
	378	DL-DIRECT	125,000	2	16	12	10	1	88	0	3,658,710	1,047	9,843	1,720	239	88.58	183
	379	DL-DIRECT	125,000	2	24	17	15	2	83	0	4,797,117	1,372	14,372	2,306	314	88.81	183
	380	DL-DIRECT	125,000	2	32	22	19	2	78	0	5,865,405	1,678	18,022	2,730	384	89.15	182
	381	DL-DIRECT	125,000	2	40	25	21	2	75	0	6,843,925	1,958	22,572	3,222	448	88.64	183
	382	DL-DIRECT	125,000	2	48	27	23	3	73	0	7,366,348	2,108	24,804	3,472	482	87.92	183
	383	DL-DIRECT	125,000	2	56	29	25	3	71	0	7,820,652	2,238	28,584	3,728	512	87.81	183
	384	DL-DIRECT	125,000	2	64	30	26	3	70	0	8,020,544	2,295	30,415	3,849	526	87.49	183

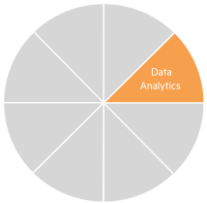
Note
The X8 data sheet specifies a peak load rate of 7.5 GBph = 2,133 MBps for QRHC.

peakmarks

Stop guessing. Start measuring.

peakmarks

Workloads to determine the Data Analytics Performance



Motivation

In general, data analytics operations cause full table scans. The performance of full table scan operation depends on the location of data in the storage hierarchy (storage, memory) and the technology used to boost scan performance (smart scan for the data location storage system, in-memory column store for the data location memory).

The goal is to

- Optimize data scan throughput
- Validate the impact of several factors on data analytics performance
  - Smart scan offload technology
  - In-memory column store technology
- Optimize Oracle license and maintenance costs

Key Performance Metrics

- SQL data scan rate in megabytes per second [MBps] or rows per second [rps]

Description

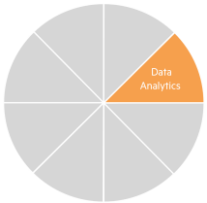
Workload	Measurement Unit	Action
DA-STORAGE	[MBps] [rps]	Simple aggregate after full table scan, using conventional storage.
DA-OFFLOAD	[MBps] [rps]	Simple aggregate after full table scan, using smart-scan offload technology.
DA-ROWSTORE	[MBps] [rps]	Simple aggregate after full table scan, using row-store.
DA-COLSTORE	[MBps] [rps]	Simple aggregate after full table scan, using column-store. This workload uses in-memory optimized arithmetic operations introduced in 18c if supported by the underlying processor. Starting in Oracle 19.8, the base level of the in-memory option supports up to 16 GByte in-memory column store without license costs.

Note
These kinds of analytic workloads are generic to all applications in all industries.

Data Analytics Performance

In many cases, the per-core performance is relevant for licensing

Intelligent storage solutions may cause additional license costs



Workload DA-STORAGE – Data scan using a conventional storage technology

Single Instance

Run	Test	Workload	Nodes	Jobs	DOP	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Scan rate total [rps]	Scan rate total [MBps]	FlCache read [%]	BuCache read [%]	Elapsed time [s]
8	1	DA-STORAGE	1	1	4	2	1	0	98	0	9,539,317	3,245	100.00	0.00	171
	3	DA-STORAGE	1	8	4	3	2	1	97	0	18,418,358	6,265	100.00	0.00	172
	5	DA-STORAGE	1	16	4	4	2	1	96	0	18,483,254	6,287	100.00	0.00	173
	7	DA-STORAGE	1	24	4	4	2	1	96	0	18,492,392	6,290	100.00	0.00	173
	9	DA-STORAGE	1	32	4	4	2	1	96	0	18,495,409	6,291	100.00	0.00	174

2-node cluster

Run	Test	Workload	Nodes	Jobs	DOP	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Scan rate total [rps]	Scan rate total [MBps]	FlCache read [%]	BuCache read [%]	Elapsed time [s]
15	1	DA-STORAGE	2	2	4	2	1	1	98	0	16,450,340	5,601	100.00	0.00	171
	3	DA-STORAGE	2	16	4	3	2	1	97	0	34,464,955	11,734	100.00	0.00	172
	5	DA-STORAGE	2	32	4	4	2	1	96	0	35,703,125	12,155	100.00	0.00	172
	7	DA-STORAGE	2	48	4	4	2	1	96	0	36,270,748	12,348	100.00	0.00	172
	9	DA-STORAGE	2	64	4	4	3	1	96	0	36,527,567	12,436	100.00	0.00	173
	11	DA-STORAGE	2	80	4	4	3	1	96	0	36,640,231	12,474	100.00	0.00	174
	13	DA-STORAGE	2	96	4	4	3	1	96	0	36,765,045	12,516	100.00	0.00	174

Workload DA-OFFLOAD – Data scan using a offload storage technology

Single Instance

Run	Test	Workload	Nodes	Jobs	DOP	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Scan rate total [rps]	Scan rate total [MBps]	FlCache read [%]	BuCache read [%]	Elapsed time [s]
8	33	DA-OFFLOAD	1	1	1	1	0	0	99	0	51,028,896	17,363	100.00	0.00	171
	35	DA-OFFLOAD	1	8	1	1	1	0	99	0	113,256,101	38,535	100.00	0.00	171
	37	DA-OFFLOAD	1	16	1	2	1	1	98	0	126,550,960	43,059	100.00	0.00	171
	39	DA-OFFLOAD	1	24	1	2	1	1	98	0	156,881,791	53,379	100.00	0.00	172
	41	DA-OFFLOAD	1	32	1	2	1	1	98	0	166,308,748	56,587	100.00	0.00	172
	43	DA-OFFLOAD	1	40	1	3	1	1	98	0	169,935,247	57,820	100.00	0.00	172

2-node cluster

Run	Test	Workload	Nodes	Jobs	DOP	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Scan rate total [rps]	Scan rate total [MBps]	FlCache read [%]	BuCache read [%]	Elapsed time [s]
15	33	DA-OFFLOAD	2	2	1	1	1	0	99	0	58,173,231	19,819	100.00	0.00	171
	35	DA-OFFLOAD	2	16	1	2	1	1	98	0	164,581,594	56,070	100.00	0.00	171
	37	DA-OFFLOAD	2	32	1	2	1	1	98	0	187,076,316	63,733	100.00	0.00	171
	39	DA-OFFLOAD	2	48	1	2	1	1	98	0	195,758,609	66,691	100.00	0.00	171
	41	DA-OFFLOAD	2	64	1	2	1	1	98	0	196,795,769	67,044	100.00	0.00	172
	43	DA-OFFLOAD	2	80	1	2	1	1	98	0	199,305,704	67,899	100.00	0.00	172
	45	DA-OFFLOAD	2	96	1	2	1	1	98	0	201,142,122	68,525	100.00	0.00	172

Workload DA-ROWSTORE – Data scan using buffer cache row store

Single Instance

Run	Test	Workload	Nodes	Jobs	DOP	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Scan rate total [rps]	Scan rate total [MBps]	FlCache read [%]	BuCache read [%]	Elapsed time [s]
8	65	DA-ROWSTORE	1	1	1	2	1	0	98	0	7,548,718	2,281	0.00	99.73	180
	66	DA-ROWSTORE	1	24	1	26	25	0	74	0	180,623,599	54,565	0.00	100.00	182
	67	DA-ROWSTORE	1	48	1	51	50	0	49	0	308,723,162	93,263	0.00	100.00	183
	68	DA-ROWSTORE	1	72	1	75	74	0	25	0	376,972,319	113,881	0.00	100.00	183
	69	DA-ROWSTORE	1	96	1	97	96	0	3	0	417,263,505	126,052	0.00	100.00	182

2-node cluster

Run	Test	Workload	Nodes	Jobs	DOP	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Scan rate total [rps]	Scan rate total [MBps]	FlCache read [%]	BuCache read [%]	Elapsed time [s]
15	65	DA-ROWSTORE	2	2	1	2	1	0	98	0	15,223,603	4,600	0.00	99.74	181
	66	DA-ROWSTORE	2	48	1	25	25	0	75	0	362,635,020	109,529	0.00	100.00	183
	67	DA-ROWSTORE	2	96	1	50	49	0	50	0	617,217,180	186,421	0.00	100.00	183
	68	DA-ROWSTORE	2	144	1	75	73	1	25	0	754,465,316	227,875	0.00	100.00	183
	69	DA-ROWSTORE	2	192	1	96	94	1	4	0	845,157,690	255,266	0.00	100.00	183

Note

The scan rate in MBps corresponds to the amount of data in the scanned table. This value depends on the data model. In our data model, the row length is approx. 300 bytes.

Workload DA-COLSTORE – Data scan using column store

Single Instance

Run	Test	Workload	Nodes	Jobs	DOP	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Scan rate total [rps]	Scan rate total [MBps]	FlCache read [%]	BuCache read [%]	Elapsed time [s]
8	70	DA-COLSTORE	1	1	1	2	1	0	98	0	4,820,827,589	1,460,407	0.00	70.46	181
	71	DA-COLSTORE	1	24	1	26	25	0	74	0	65,379,202,103	19,816,074	0.00	100.00	181
	72	DA-COLSTORE	1	48	1	52	50	1	48	0	58,633,796,758	17,787,953	0.00	100.00	181
	73	DA-COLSTORE	1	72	1	76	74	1	24	0	60,924,913,364	18,495,364	0.00	100.00	181
	74	DA-COLSTORE	1	96	1	97	95	1	3	0	62,339,614,528	18,936,557	0.00	100.00	181

2-node cluster

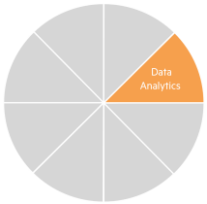
Run	Test	Workload	Nodes	Jobs	DOP	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Scan rate total [rps]	Scan rate total [MBps]	FlCache read [%]	BuCache read [%]	Elapsed time [s]
15	70	DA-COLSTORE	2	2	1	2	1	0	98	0	10,861,614,976	3,287,494	0.00	99.99	182
	71	DA-COLSTORE	2	48	1	26	25	0	74	0	141,771,840,591	42,916,756	0.00	100.00	182
	72	DA-COLSTORE	2	96	1	52	50	1	48	0	116,215,766,166	35,198,349	0.00	100.00	181
	73	DA-COLSTORE	2	144	1	75	72	1	25	0	125,054,889,772	37,885,745	0.00	100.00	181
	74	DA-COLSTORE	2	192	1	96	93	1	4	0	131,717,969,993	39,916,054	0.00	100.00	182

Note

The scan rate in MBps corresponds to the amount of data in the scanned table. This value depends on the data model. In our data model, the row length is approx. 300 bytes.

Data Analytics Performance

Selected comparative figures – For training purposes only



Ampere Altra A1
32c, 32t, Launch 2021

Workload DA-ROWSTORE – Data scan using buffer cache row store

Run	Test	Workload	Nodes	Jobs	DOP	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Scan rate total [rps]	Scan rate total [MBps]	FlCache read [%]	BuCache read [%]	Elapsed time [s]
4	1	DA-ROWSTORE	1	1	1	3	3	0	97	0	11,640,683	3,516	0.00	100.00	180
	2	DA-ROWSTORE	1	8	1	25	25	0	75	0	84,455,021	25,508	0.00	100.00	182
	3	DA-ROWSTORE	1	16	1	50	50	0	50	0	161,212,678	48,691	0.00	100.00	181
	4	DA-ROWSTORE	1	24	1	75	74	0	25	0	226,474,766	68,402	0.00	100.00	182
	5	DA-ROWSTORE	1	32	1	99	99	0	1	0	282,846,522	85,429	0.00	100.00	182

Notes
- The scan rate in MBps corresponds to the amount of data in the scanned table. This value depends on the data model. In our data model, the row length is approx. 300 bytes.
- The Ampera Altra A1 has no multithreading and delivers predictable and scalable performance up to high CPU utilization.

Ampere Altra A1
32c, 32t, Launch 2021

Workload DA-COLSTORE – Data scan using column store

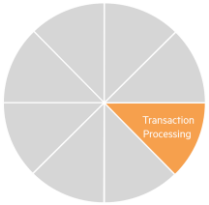
Run	Test	Workload	Nodes	Jobs	DOP	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Scan rate total [rps]	Scan rate total [MBps]	FlCache read [%]	BuCache read [%]	Elapsed time [s]
4	6	DA-COLSTORE	1	1	1	3	3	0	97	0	2,857,569,742	910,025	0.00	100.00	181
	7	DA-COLSTORE	1	8	1	25	25	0	75	0	20,092,145,123	6,398,904	0.00	100.00	182
	8	DA-COLSTORE	1	16	1	50	49	1	50	0	36,221,467,587	11,536,106	0.00	100.00	181
	9	DA-COLSTORE	1	24	1	75	74	1	25	0	52,822,677,226	16,823,559	0.00	100.00	182
	10	DA-COLSTORE	1	32	1	100	98	1	0	0	70,196,452,198	22,356,905	0.00	100.00	182

Notes
- The scan rate in MBps corresponds to the amount of data in the scanned table. This value depends on the data model. In our data model, the row length is approx. 300 bytes.
- The Ampera Altra A1 has no multithreading and delivers predictable and scalable performance up to high CPU utilization.

peakmarks

Stop guessing. Start measuring.

Workloads to determine the Transaction Processing Performance



Motivation

For capacity planning reasons, it is necessary to know the performance characteristics of a platform for transactions of varying complexity. Transaction Processing is the most complex database operation.

The goal is to

- Optimize the transaction throughput and transaction response time
- Validate the impact of several factors on transaction throughput and response time:
  - database size
  - buffer cache size
  - transaction size
  - I/O random read service time
  - log writer latency

Key Performance Metrics

- SQL transaction throughput in transactions per second [tps]
- SQL transaction response time in milliseconds [ms]

Description

Workload	Measurement Unit	Action
TP-REPORT	[tps]	Online report of transaction processing application.
	[ms]	SELECT Ø 25 rows via index.
TP-LOOKUP	[tps]	Fast lookup query.
	[ms]	SELECT single row via index, e.g., SELECT an account, product.
		If configured, this workload uses tables in the memory-optimized row store for fast look-up introduced in 18c. Otherwise, this workload uses conventional tables.

Note
These transaction processing workloads are generic to all applications in all industries.

Description

Workload	Measurement Unit	Action
TP-LIGHT	[tps]	Light transaction type.
	[ms]	SELECT/UPDATE single row via index, e.g., SELECT/UPDATE an account, product, or order with different SELECT/UPDATE ratios using SELECT FOR UPDATE locking. The workload parameter specifies the update ratio in %; the following values are supported {0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100}.
		This workload shows maximum transaction throughput and minimum transaction response time.
TP-MEDIUM	[tps]	Medium transaction type.
	[ms]	SELECT/UPDATE Ø 25 rows via index, e.g., SELECT/UPDATE last month's bank account bookings with different SELECT/UPDATE ratios using SELECT FOR UPDATE locking. The workload parameter specifies the update ratio in %; the following values are supported {0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100}.
TP-HEAVY	[tps]	Heavy transaction type.
	[ms]	SELECT/UPDATE Ø 125 rows via index, e.g., SELECT/UPDATE last month's cell phone call records with different SELECT/UPDATE ratios using SELECT FOR UPDATE locking. The workload parameter specifies the update ratio in %; the following values are supported {0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100}.

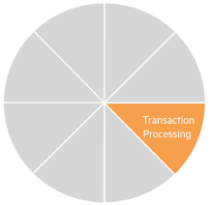
Note
These transaction processing workloads are generic to all applications in all industries.

Description

Workload	Measurement Unit	Action
TP-MIXED1	[tps] [ms]	A read-intensive mix of different transaction types. Logical reads: 83% read, 17% write; avg 256-byte REDO per transaction. This workload is a complex workload that is composed of the equally weighted simple workloads TP-REPORT and TP-LOOKUP, TP-MEDIUM (with 40% UPDATE) and DL-BUFFER (with 2 rpt).
TP-MIXED2	[tps] [ms]	A write-intensive mix of different transaction types. Logical reads: 65% read, 35% write; avg 1,725-byte REDO per transaction. This workload is a complex workload that is composed of the equally weighted simple workloads TP-LIGHT (with 40% UPDATE), TP-MEDIUM (with 30% UPDATE), TP-HEAVY (with 20% UPDATE), and DL-BUFFER (with 3 rpt).
TP-CUSTOM	[tps] [ms]	A configurable and customized mix of different transaction types. Default configuration compatible to peakmarks Version 9 TP-MIXED.

Notes
- TP-MIXED1 and TP-MIXED2 are the most representative peakmarks workloads for determining Oracle transaction processing performance capabilities on a specific platform.
- TP-MIXED1 achieves much higher transaction rates and CPU utilization than TP-MIXED2.
- This kind of transaction processing workload is generic to all industry applications.
- Peakmarks provides several performance reports for TP workloads: kpm_tp.sql (used in this presentation) shows overall transaction performance, kpm_tpplus.sql provides more detailed information, and kpm_tpio.sql provides I/O information.

Transaction Processing Performance



Workload TP-REPORT – online report, avg 25 rows per query

Single Instance

Run	Test	Workload	Upd [%]	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Transactions total [tps]	Response time [ms]	IO time read [ms]	REDO data [kBpt]	LogFile sync [ms]	BuCache read [%]	FlCache read [%]	Elapsed time [s]
9	2	TP-REPORT	N/A	1	1	1	1	0	99	0	748	1.337	0.207	1.802	1.279	18.68	100.00	181
	10	TP-REPORT	N/A	1	32	24	13	7	76	0	18,797	1.698	0.262	0.581	1.192	22.71	100.00	182
	14	TP-REPORT	N/A	1	48	34	19	10	66	0	24,563	1.944	0.296	1.022	1.179	23.53	100.00	183
	18	TP-REPORT	N/A	1	64	43	25	12	57	0	28,264	2.249	0.343	0.739	1.167	24.32	100.00	182
	22	TP-REPORT	N/A	1	80	49	29	13	51	0	30,795	2.584	0.400	0.573	1.157	25.01	100.00	182
	26	TP-REPORT	N/A	1	96	54	32	14	46	0	31,923	2.990	0.474	0.701	1.159	24.92	100.00	183

2-node cluster

Run	Test	Workload	Upd [%]	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Transactions total [tps]	Response time [ms]	IO time read [ms]	REDO data [kBpt]	LogFile sync [ms]	BuCache read [%]	FlCache read [%]	Elapsed time [s]
16	2	TP-REPORT	N/A	2	2	2	1	1	98	0	1,293	1.541	0.218	1.279	1.113	29.46	100.00	181
	6	TP-REPORT	N/A	2	32	15	8	5	85	0	17,836	1.784	0.250	0.934	1.030	34.60	100.00	182
	10	TP-REPORT	N/A	2	64	25	13	8	75	0	30,408	2.092	0.281	0.778	0.939	39.99	100.00	182
	14	TP-REPORT	N/A	2	96	32	17	10	68	0	38,278	2.494	0.314	0.556	0.920	41.09	100.00	182
	18	TP-REPORT	N/A	2	128	38	21	11	62	0	44,761	2.840	0.355	0.557	0.925	40.80	100.00	183
	22	TP-REPORT	N/A	2	160	43	24	13	57	0	51,247	3.101	0.394	0.498	0.886	43.45	100.00	183

Workload TP-LOOKUP – look-up query with minimum response time

Single Instance

Run	Test	Workload	Upd [%]	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Transactions total [tps]	Response time [ms]	IO time read [ms]	REDO data [kBpt]	LogFile sync [ms]	BuCache read [%]	FlCache read [%]	Elapsed time [s]
9	34	TP-LOOKUP	N/A	1	1	1	1	0	99	0	3,357	0.298	0.244	0.841	0.080	74.93	99.99	181
	38	TP-LOOKUP	N/A	1	16	17	17	0	83	0	1,257,350	0.013	0.571	0.396	0.000	100.00	97.75	182
	42	TP-LOOKUP	N/A	1	32	34	33	0	66	0	2,031,891	0.016	0.540	0.355	0.000	100.00	97.86	181
	46	TP-LOOKUP	N/A	1	48	50	49	1	50	0	2,487,948	0.019	0.391	0.307	0.000	100.00	99.05	181
	50	TP-LOOKUP	N/A	1	64	66	65	1	34	0	2,594,656	0.025	0.862	0.329	0.000	100.00	99.10	182
	54	TP-LOOKUP	N/A	1	80	82	81	1	18	0	2,727,060	0.029	0.823	0.324	0.000	100.00	97.90	181
	58	TP-LOOKUP	N/A	1	96	97	95	1	3	0	2,846,073	0.034	3.771	0.336	0.000	100.00	97.84	181

2-node cluster

Run	Test	Workload	Upd [%]	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Transactions total [tps]	Response time [ms]	IO time read [ms]	REDO data [kBpt]	LogFile sync [ms]	BuCache read [%]	FlCache read [%]	Elapsed time [s]
16	34	TP-LOOKUP	N/A	2	2	1	1	1	99	0	5,452	0.366	0.240	1.155	0.086	73.48	100.00	182
	38	TP-LOOKUP	N/A	2	32	17	17	0	83	0	2,480,910	0.013	0.516	0.380	0.000	100.00	98.67	182
	42	TP-LOOKUP	N/A	2	64	33	33	0	67	0	4,035,779	0.016	0.639	0.412	0.000	100.00	98.22	182
	46	TP-LOOKUP	N/A	2	96	50	49	1	50	0	4,936,726	0.019	0.593	0.345	0.000	100.00	99.10	181
	50	TP-LOOKUP	N/A	2	128	66	65	1	34	0	5,218,022	0.024	0.864	0.431	0.000	100.00	98.26	182
	54	TP-LOOKUP	N/A	2	160	83	81	1	17	0	5,407,780	0.029	0.865	0.359	0.000	100.00	99.10	182
	58	TP-LOOKUP	N/A	2	192	96	94	1	4	0	5,703,844	0.033	1.045	0.376	0.000	100.00	98.34	182

Workload TP-LIGHT – maximum transaction throughput and minimum response time (20% update)

Single Instance

Run	Test	Workload	Upd [%]	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Transactions total [tps]	Response time [ms]	IO time read [ms]	REDO data [kBpt]	LogFile sync [ms]	BuCache read [%]	FlCache read [%]	Elapsed time [s]
9	66	TP-LIGHT	20	1	1	1	1	0	99	0	1,903	0.525	0.241	0.511	0.313	67.61	100.00	181
	70	TP-LIGHT	20	1	16	13	8	3	88	0	42,182	0.378	0.235	0.185	0.298	77.82	100.00	182
	74	TP-LIGHT	20	1	32	25	17	5	75	0	98,024	0.325	0.257	0.184	0.266	83.47	100.00	182
	78	TP-LIGHT	20	1	48	34	24	7	66	0	138,530	0.345	0.278	0.226	0.261	83.81	100.00	182
	82	TP-LIGHT	20	1	64	42	29	9	58	0	174,940	0.364	0.294	0.285	0.258	83.88	100.00	182
	86	TP-LIGHT	20	1	80	51	35	10	49	0	206,960	0.384	0.308	0.281	0.254	83.93	100.00	182
	90	TP-LIGHT	20	1	96	58	40	12	42	0	236,484	0.404	0.323	0.258	0.247	83.91	100.00	182
	94	TP-LIGHT	20	1	112	64	45	13	36	0	258,264	0.431	0.350	0.255	0.245	83.92	100.00	182
	97	TP-LIGHT	20	1	124	68	47	14	32	0	270,624	0.456	0.374	0.306	0.244	83.94	100.00	182

2-node cluster

Run	Test	Workload	Upd [%]	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Transactions total [tps]	Response time [ms]	IO time read [ms]	REDO data [kBpt]	LogFile sync [ms]	BuCache read [%]	FlCache read [%]	Elapsed time [s]
16	66	TP-LIGHT	20	2	2	1	1	1	99	0	3,922	0.508	0.243	0.368	0.297	72.21	99.99	182
	70	TP-LIGHT	20	2	32	15	9	3	85	0	109,183	0.291	0.265	0.188	0.254	86.33	100.00	182
	74	TP-LIGHT	20	2	64	25	17	6	75	0	204,611	0.311	0.291	0.290	0.248	86.62	100.00	182
	78	TP-LIGHT	20	2	96	34	23	7	66	0	288,283	0.331	0.322	0.278	0.241	87.17	100.00	182
	82	TP-LIGHT	20	2	128	41	28	9	59	0	354,492	0.359	0.355	0.350	0.236	87.23	100.00	182
	86	TP-LIGHT	20	2	160	47	32	10	53	0	404,681	0.393	0.389	0.294	0.234	87.22	100.00	182
	90	TP-LIGHT	20	2	192	52	35	11	48	0	440,736	0.432	0.432	0.444	0.232	87.22	100.00	183
	94	TP-LIGHT	20	2	224	56	38	12	44	0	471,001	0.465	0.469	0.415	0.231	87.23	100.00	185
	97	TP-LIGHT	20	2	248	59	40	13	41	0	490,885	0.490	0.494	0.357	0.230	87.26	100.00	186

Workload TP-MEDIUM – maximum transaction throughput and minimum response time (20% update)

Single Instance

Run	Test	Workload	Upd [%]	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Transactions total [tps]	Response time [ms]	IO time read [ms]	REDO data [kBpt]	LogFile sync [ms]	BuCache read [%]	FlCache read [%]	Elapsed time [s]
9	98	TP-MEDIUM	20	1	1	2	1	1	98	0	701	1.427	0.217	0.415	4.637	46.68	100.00	182
	102	TP-MEDIUM	20	1	16	20	11	6	80	0	7,828	2.036	0.271	0.179	4.949	41.24	100.00	182
	106	TP-MEDIUM	20	1	32	32	18	9	68	0	12,867	2.478	0.335	0.260	4.891	42.44	100.00	182
	110	TP-MEDIUM	20	1	48	42	23	12	58	0	16,438	2.908	0.399	0.260	4.843	42.53	100.00	182
	114	TP-MEDIUM	20	1	64	49	28	14	51	0	19,174	3.324	0.476	0.314	4.774	43.82	100.00	182
	118	TP-MEDIUM	20	1	80	56	32	15	44	0	20,853	3.821	0.564	0.357	4.741	44.83	100.00	181
	122	TP-MEDIUM	20	1	96	60	36	16	40	0	21,953	4.357	0.668	0.325	4.717	45.19	100.00	182
	126	TP-MEDIUM	20	1	112	64	38	17	36	0	22,621	4.935	0.780	0.331	4.709	45.08	100.00	182
	129	TP-MEDIUM	20	1	124	66	40	17	34	0	22,914	5.391	0.872	0.770	4.704	45.06	100.00	182

2-node cluster

Run	Test	Workload	Upd [%]	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Transactions total [tps]	Response time [ms]	IO time read [ms]	REDO data [kBpt]	LogFile sync [ms]	BuCache read [%]	FlCache read [%]	Elapsed time [s]
16	98	TP-MEDIUM	20	2	2	2	1	1	98	0	1,328	1.504	0.225	0.398	4.393	59.51	100.00	182
	102	TP-MEDIUM	20	2	32	19	10	6	81	0	13,275	2.399	0.327	0.258	4.799	49.90	100.00	182
	106	TP-MEDIUM	20	2	64	28	15	8	72	0	20,211	3.148	0.465	0.366	4.682	53.37	100.00	182
	110	TP-MEDIUM	20	2	96	34	18	10	66	0	25,088	3.811	0.599	0.390	4.598	55.41	100.00	182
	114	TP-MEDIUM	20	2	128	38	21	11	62	0	28,319	4.499	0.721	0.377	4.566	55.36	100.00	182
	118	TP-MEDIUM	20	2	160	42	23	12	58	0	30,829	5.170	0.841	0.352	4.547	55.05	100.00	182
	122	TP-MEDIUM	20	2	192	44	25	13	56	0	32,644	5.853	0.989	0.416	4.534	55.26	100.00	182
	126	TP-MEDIUM	20	2	224	45	25	13	55	0	33,527	6.550	1.130	0.480	4.526	55.23	100.00	185
	129	TP-MEDIUM	20	2	248	45	26	13	55	0	33,979	7.085	1.243	0.573	4.519	55.37	100.00	186

Workload TP-HEAVY – maximum transaction throughput and minimum response time (20% update)

Single Instance

Run	Test	Workload	Upd [%]	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Transactions total [tps]	Response time [ms]	IO time read [ms]	REDO data [kBpt]	LogFile sync [ms]	BuCache read [%]	FlCache read [%]	Elapsed time [s]
9	130	TP-HEAVY	20	1	1	3	1	1	97	0	312	3.207	0.221	0.359	21.771	51.87	100.00	181
	134	TP-HEAVY	20	1	16	25	14	7	75	0	3,013	5.298	0.359	0.384	22.687	44.65	100.00	182
	138	TP-HEAVY	20	1	32	41	23	12	59	0	4,369	7.298	0.584	0.430	22.791	42.40	100.00	182
	142	TP-HEAVY	20	1	48	51	29	14	49	0	4,797	9.975	0.947	0.542	22.886	41.14	100.00	182
	146	TP-HEAVY	20	1	64	56	33	15	44	0	4,908	12.996	1.372	0.379	22.951	40.27	100.00	181

2-node cluster

Run	Test	Workload	Upd [%]	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Transactions total [tps]	Response time [ms]	IO time read [ms]	REDO data [kBpt]	LogFile sync [ms]	BuCache read [%]	FlCache read [%]	Elapsed time [s]
16	130	TP-HEAVY	20	2	2	3	1	1	97	0	702	2.848	0.228	0.560	20.356	66.94	100.00	182
	132	TP-HEAVY	20	2	16	14	7	4	86	0	3,675	4.337	0.350	0.331	21.683	57.34	100.00	182
	134	TP-HEAVY	20	2	32	21	11	7	79	0	5,600	5.693	0.547	0.866	21.702	55.56	100.00	181
	136	TP-HEAVY	20	2	48	26	14	8	74	0	6,504	7.350	0.780	0.579	21.763	54.44	100.00	181
	138	TP-HEAVY	20	2	64	29	15	9	71	0	6,863	9.300	1.068	0.730	21.962	52.10	100.00	181
	140	TP-HEAVY	20	2	80	30	16	9	70	0	6,984	11.410	1.390	0.497	22.063	50.89	100.00	181
	142	TP-HEAVY	20	2	96	31	16	9	69	0	7,012	13.635	1.692	0.885	22.085	50.61	100.00	182

Workload TP-MIXED1 – Read-intensive mix of different transaction types

Single Instance

Run	Test	Workload	Upd [%]	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Transactions total [tps]	Response time [ms]	IO time read [ms]	REDO data [kBpt]	LogFile sync [ms]	BuCache read [%]	FlCache read [%]	Elapsed time [s]
10	2	TP-MIXED1	N/A	1	4	4	2	1	96	0	7,864	0.507	0.218	0.239	1.854	65.26	100.00	182
	3	TP-MIXED1	N/A	1	8	7	4	2	93	0	37,713	0.211	0.225	0.246	0.746	77.72	100.00	183
	5	TP-MIXED1	N/A	1	16	16	11	3	84	0	271,708	0.058	0.246	0.350	0.162	92.51	100.00	183
	7	TP-MIXED1	N/A	1	24	27	18	6	73	0	414,529	0.058	0.281	0.466	0.128	93.52	100.00	182
	9	TP-MIXED1	N/A	1	32	33	22	7	67	0	548,974	0.058	0.298	0.582	0.113	93.71	100.00	182
	11	TP-MIXED1	N/A	1	40	38	26	8	62	0	610,997	0.065	0.314	0.692	0.113	93.28	100.00	182

2-node cluster

Run	Test	Workload	Upd [%]	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Transactions total [tps]	Response time [ms]	IO time read [ms]	REDO data [kBpt]	LogFile sync [ms]	BuCache read [%]	FlCache read [%]	Elapsed time [s]
17	2	TP-MIXED1	N/A	2	8	4	2	1	96	0	13,979	0.570	0.240	0.228	2.152	70.11	100.00	181
	5	TP-MIXED1	N/A	2	32	17	11	4	83	0	573,264	0.055	0.292	0.483	0.131	94.14	100.00	182
	9	TP-MIXED1	N/A	2	64	28	19	6	72	0	1,062,532	0.060	0.359	0.800	0.100	94.46	100.00	182
	13	TP-MIXED1	N/A	2	96	37	26	7	63	0	1,512,411	0.063	0.429	1.067	0.083	95.06	100.00	182
	17	TP-MIXED1	N/A	2	128	45	32	8	55	0	1,875,286	0.068	0.496	1.315	0.074	95.35	100.00	182
	21	TP-MIXED1	N/A	2	160	53	39	9	47	0	2,098,252	0.076	0.561	1.559	0.071	95.41	100.00	182
	25	TP-MIXED1	N/A	2	192	60	44	10	40	0	2,233,273	0.085	0.614	1.786	0.071	95.32	100.00	183
	29	TP-MIXED1	N/A	2	224	66	49	11	34	0	2,317,660	0.094	0.668	1.968	0.070	95.25	100.00	186

Workload TP-MIXED2 – Write-intensive mix of different transaction types

Single Instance

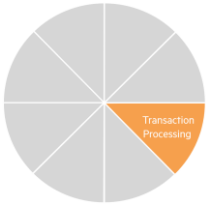
Run	Test	Workload	Upd [%]	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Transactions total [tps]	Response time [ms]	IO time read [ms]	REDO data [kBpt]	LogFile sync [ms]	BuCache read [%]	FlCache read [%]	Elapsed time [s]
11	2	TP-MIXED2	N/A	1	4	6	3	2	94	0	5,335	0.744	0.224	0.239	2.722	57.81	100.00	183
	5	TP-MIXED2	N/A	1	16	18	10	5	82	0	15,428	1.030	0.271	0.400	2.793	54.65	100.00	182
	9	TP-MIXED2	N/A	1	32	29	17	8	71	0	23,148	1.376	0.324	0.647	2.822	51.32	100.00	182
	13	TP-MIXED2	N/A	1	48	37	22	10	63	0	29,660	1.610	0.366	0.884	2.761	49.69	100.00	182
	17	TP-MIXED2	N/A	1	64	45	27	12	55	0	32,958	1.931	0.427	1.109	2.841	48.29	100.00	182
	21	TP-MIXED2	N/A	1	80	52	32	13	48	0	35,374	2.249	0.498	1.400	2.881	47.98	100.00	182
	24	TP-MIXED2	N/A	1	92	57	35	14	43	0	36,508	2.506	0.561	1.596	2.901	47.56	100.00	182

2-node cluster

Run	Test	Workload	Upd [%]	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Transactions total [tps]	Response time [ms]	IO time read [ms]	REDO data [kBpt]	LogFile sync [ms]	BuCache read [%]	FlCache read [%]	Elapsed time [s]
18	2	TP-MIXED2	N/A	2	8	6	3	2	94	0	10,444	0.762	0.250	0.267	2.827	65.73	100.00	183
	3	TP-MIXED2	N/A	2	16	10	5	3	90	0	17,037	0.933	0.274	0.387	2.917	63.63	100.00	182
	5	TP-MIXED2	N/A	2	32	16	9	4	84	0	26,667	1.192	0.327	0.594	2.932	62.20	100.00	182
	7	TP-MIXED2	N/A	2	48	20	11	6	80	0	32,956	1.450	0.382	0.810	2.895	59.73	100.00	182
	9	TP-MIXED2	N/A	2	64	23	13	7	77	0	37,798	1.685	0.437	1.021	2.880	58.63	100.00	182
	11	TP-MIXED2	N/A	2	80	26	14	7	74	0	41,402	1.922	0.496	1.223	2.942	58.27	100.00	182
	13	TP-MIXED2	N/A	2	96	28	16	8	72	0	45,473	2.098	0.550	1.414	2.908	58.41	100.00	182
	15	TP-MIXED2	N/A	2	112	30	17	8	70	0	49,139	2.266	0.605	1.605	2.865	58.46	100.00	182
	17	TP-MIXED2	N/A	2	128	32	18	9	68	0	51,392	2.478	0.660	1.770	2.854	57.81	100.00	182

Transaction Processing Performance

Selected comparative figures – For training purposes only



Workload TP-REPORT – online report with minimum response time

Exadata X5-2 QRHC, Launch 2015
Database size 1 Tbyte

Run	Test	Workload	Upd [%]	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Transactions total [tps]	Response time [ms]	IO time read [ms]	REDO data [kBpt]	LogFile sync [ms]	BuCache read [%]	FlCache read [%]	Elapsed time [s]
6	2	TP-REPORT	N/A	1	1	2	1	1	98	0	371	2.697	0.307	0.871	0.594	52.34	100.00	302
	4	TP-REPORT	N/A	1	8	9	7	1	92	0	4,550	1.749	0.307	0.421	0.825	74.87	100.00	303
	6	TP-REPORT	N/A	1	16	17	14	2	83	0	9,208	1.723	0.291	0.440	0.661	73.43	100.00	304
	8	TP-REPORT	N/A	1	24	25	20	3	75	0	12,931	1.845	0.300	0.458	0.732	72.26	100.00	304
	10	TP-REPORT	N/A	1	32	33	27	4	67	0	16,963	1.873	0.299	0.470	0.668	71.29	100.00	305
	12	TP-REPORT	N/A	1	40	41	33	5	59	0	20,150	1.972	0.309	0.474	0.595	71.07	100.00	304
	14	TP-REPORT	N/A	1	48	49	40	6	51	0	22,905	2.078	0.318	0.482	0.733	70.57	100.00	304
	15	TP-REPORT	N/A	1	52	52	42	7	48	0	23,366	2.210	0.330	0.492	0.726	69.83	100.00	304

Exadata X5-2 QRHC, Launch
2015 Database size 24 Tbyte

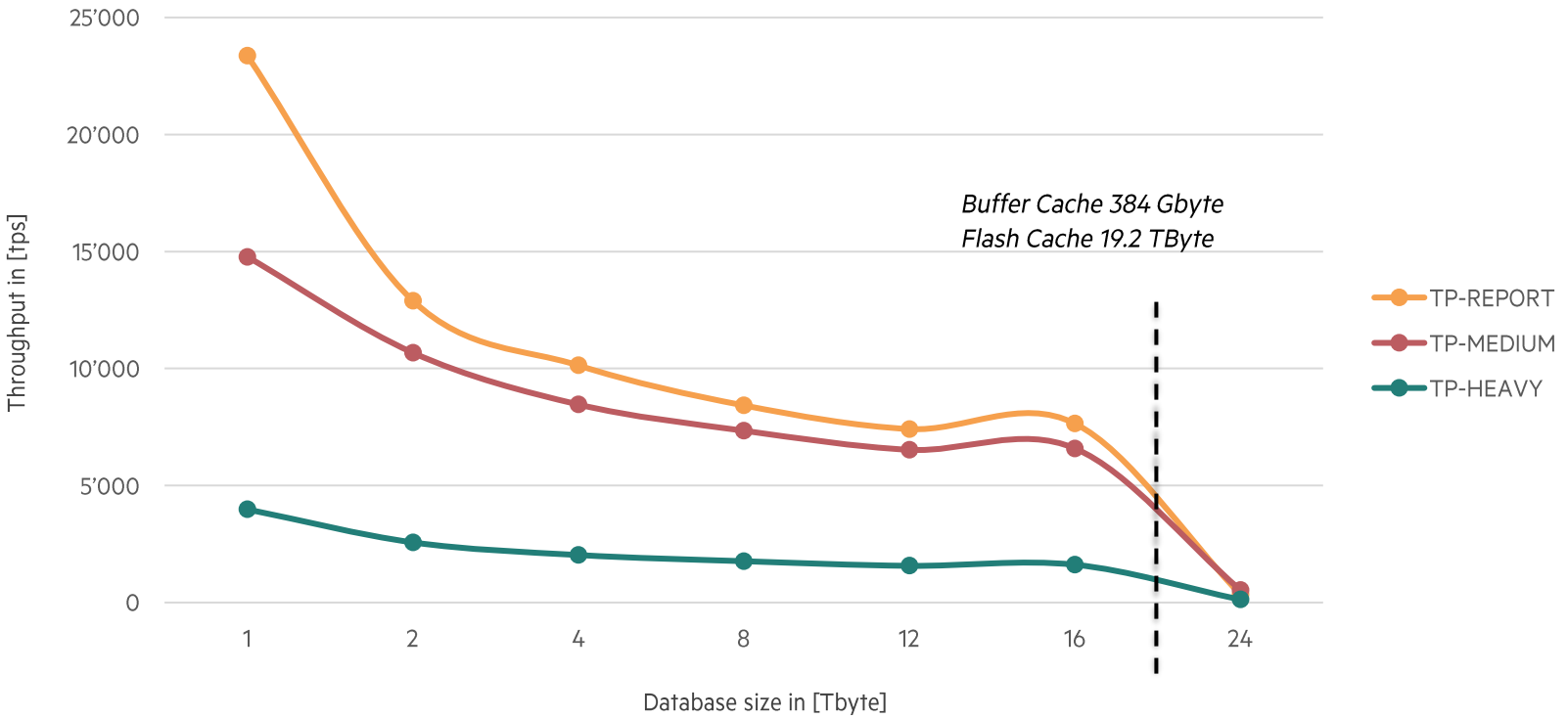
Run	Test	Workload	Upd [%]	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Transactions total [tps]	Response time [ms]	IO time read [ms]	REDO data [kBpt]	LogFile sync [ms]	BuCache read [%]	FlCache read [%]	Elapsed time [s]
18	2	TP-REPORT	N/A	1	1	1	1	0	99	0	42	23.693	1.185	1.527	0.685	23.09	87.85	300
	3	TP-REPORT	N/A	1	4	2	1	0	98	0	172	23.024	1.447	1.519	0.489	15.30	89.54	305
	4	TP-REPORT	N/A	1	8	2	1	0	98	0	172	41.141	3.043	1.322	0.922	19.10	93.75	466
	5	TP-REPORT	N/A	1	12	3	2	1	97	0	347	31.526	1.972	1.350	0.816	17.18	91.65	353
	6	TP-REPORT	N/A	1	16	2	2	1	98	0	248	41.182	2.788	1.330	0.680	17.29	90.80	653

Note

- Transaction throughput and transaction response times collapse if active data does not fit into the high-performance tier.

Transaction Processing Performance

Impact of ratio buffer cache size / database size – Impact of storage tiering



Workload TP-MIXED1 – Read-intensive mix of different transaction types

Exadata X10M QRHC, Launch 2023
Database size 16 Tbyte; 2 x 768 Gbyte SGA

Run	Test	Workload	Upd [%]	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Transactions total [tps]	Response time [ms]	IO time read [ms]	LogFile sync [ms]	REDO data [kBpt]	BuCache read [%]	FlCache read [%]	Elapsed time [s]
13	2	TP-MIXED1	N/A	2	8	2	1	0	98	0	233,655	0.034	0.195	0.202	0.152	94.24	100.00	182
	5	TP-MIXED1	N/A	2	72	10	7	2	90	0	2,290,866	0.031	0.271	0.614	0.071	95.81	100.00	183
	8	TP-MIXED1	N/A	2	144	16	12	3	84	0	4,145,383	0.035	0.362	0.915	0.057	96.30	100.00	182
	11	TP-MIXED1	N/A	2	216	20	15	3	80	0	5,896,048	0.036	0.472	1.242	0.046	96.83	100.00	182
	14	TP-MIXED1	N/A	2	288	23	17	4	77	0	7,531,551	0.038	0.575	1.559	0.038	97.19	100.00	182
	17	TP-MIXED1	N/A	2	360	26	20	4	74	0	8,673,697	0.041	0.689	1.852	0.036	97.40	100.00	182
	20	TP-MIXED1	N/A	2	432	29	22	4	71	0	9,683,596	0.044	0.808	2.134	0.033	97.57	100.00	182
	23	TP-MIXED1	N/A	2	504	31	24	4	69	0	10,478,403	0.048	0.935	2.404	0.031	97.72	100.00	182
	26	TP-MIXED1	N/A	2	576	33	27	4	67	0	11,163,923	0.051	1.041	2.620	0.029	97.83	100.00	182

Notes

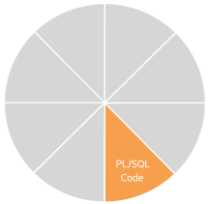
- Buffer Cache hit rate > 95% is not good enough.
- CPUs have some reserves, even in multithreaded mode.

```
Automatic Database Diagnostic Monitor (ADDM)

Findings and Recommendations
----------------------------

Finding 1: Unusual "User I/O" Wait Event - event "cell single block physical read: flash cache"
Finding 2: Unusual "User I/O" Wait Event - event "cell list of blocks physical read"
Finding 3: Commits and Rollbacks - event "log file sync"
Finding 4: Undersized SGA - The SGA was inadequately sized in some instances ...
```

peakmarks

Stop guessing. Start measuring.

Workloads to determine the PL/SQL Application Performance



Motivation

Mission-critical systems use PL/SQL to encapsulate essential functions for efficient execution as close to the data as possible. This aspect is critical in cloud environments with separate databases and application servers to reduce traffic between both environments.

The goal is to

- Validate processor performance capabilities to execute PL/SQL code

Note

- Separating application and data may cause performance problems, e.g., each roundtrip between Microsoft Azure (application) and Oracle Cloud Infrastructure (data) takes around 2 ms.

Key Performance Metrics

- PL/SQL operation throughput in a million operations per second [Mops]
- PL/SQL computing time to process algorithms in seconds [s]

Description

Workload	Measurement Unit	Action
PLS-ADD	[Mops]	Addition of numbers using different numeric data types. The workload parameter specifies the data type; the following values are supported {SI, SF, PI, NU, DA}. This workload shows the impact of the PL/SQL data type on the performance of simple arithmetic operations.
PLS-BUILTIN	[Mops]	Data type-specific operations, including SQL built-in functions, based on core banking and telco billing applications. The workload parameter specifies the data type; the following values are supported {SI, SF, PI, NU, VC}. This workload shows the impact of the PL/SQL data type on the performance of typical operations.

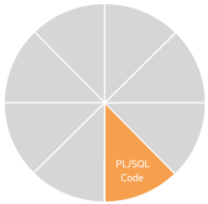
Note

- The following abbreviations for PL/SQL data types are used: SI = SIMPLE_INTEGER, SF = SIMPLE_FLOAT, PI = PLS_INTEGER, NU = NUMBER, DA = DATE, VC = VARCHAR2.

Description

Workload	Measurement Unit	Action
PLS-MIXED	[Mops]	Mixed data type-specific operations, including SQL built-in functions. This complex workload is composed of the equally weighted simple workloads PLS-ADD (NUMBER), PLS-ADD (PLS_INTEGER), PLS_BUILTIN (NUMBER), and PLS_BUILTIN (VARCHAR2). PLS-MIXED is the most representative peakmarks workload to determine an Oracle database server's PL/SQL performance capabilities.
PLS-FIBO	[s]	Calculation of Fibonacci number N. The workload parameter specifies N; the following values are supported {39, 40, 41, 42, 43, 44}. This workload shows the single-thread performance of a simple recursive algorithm implemented in PL/SQL.
PLS-PRIME	[s]	Calculation of first N prime numbers. The workload parameter specifies N; the following values are supported {1000, 10000}. This workload shows the single-thread performance of a simple algorithm implemented in PL/SQL.

PL/SQL Application Program Performance



Workload PLS-MIXED – complex workload with most used PL/SQL operations and data types

Single Instance

Run	Test	Workload	Parameter	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	Operations total [Mops]	Operations per cpu [Mops]	Elapsed time [s]
14	1	PLS-MIXED	N/A	1	4	5	4	0	95	366.51	91.63	182
	2	PLS-MIXED	N/A	1	24	25	25	0	75	1,997.48	83.23	182
	3	PLS-MIXED	N/A	1	48	50	50	0	50	3,226.36	67.22	182
	4	PLS-MIXED	N/A	1	72	75	74	0	25	3,118.72	43.32	182
	5	PLS-MIXED	N/A	1	96	98	97	0	2	3,298.58	34.36	182

2-node cluster

Run	Test	Workload	Parameter	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	Operations total [Mops]	Operations per cpu [Mops]	Elapsed time [s]
20	1	PLS-MIXED	N/A	2	8	5	4	0	95	737.47	92.18	182
	2	PLS-MIXED	N/A	2	48	25	25	0	75	4,004.28	83.42	182
	3	PLS-MIXED	N/A	2	96	50	49	0	50	6,447.20	67.16	182
	4	PLS-MIXED	N/A	2	144	74	73	0	26	6,937.48	48.18	183
	5	PLS-MIXED	N/A	2	192	95	94	0	5	6,572.97	34.23	183

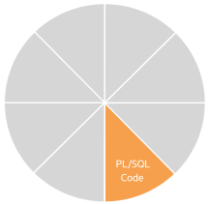
Single Instance

Workload PLS-FIBO and PLS-PRIME – single thread computing time of algorithms

Run	Test	Workload	Parameter	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	Operations total [Mops]	Operations per cpu [Mops]	Elapsed time [s]
13	31	PLS-FIBO	42	1	1	2	1	0	98	0.00	0.00	71
	32	PLS-PRIME	8000	1	1	2	1	0	98	0.00	0.00	80

PL/SQL Application Program Performance

Selected comparative figures – For training purposes only



PL/SQL Application Performance

IBM POWER9
32c, 256t, Launch 2018

Workload PLS-MIXED – complex workload with most used PL/SQL operations and data types

Run	Test	Workload	Parameter	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	Operations total [Mops]	Operations per cpu [Mops]	Elapsed time [s]
4	1	PLS-MIXED	0	1	32	12	12	0	88	1,610.51	50.33	182
	2	PLS-MIXED	0	1	64	24	24	0	76	2,774.12	43.35	182
	4	PLS-MIXED	0	1	128	47	47	0	53	3,229.68	25.23	182
	6	PLS-MIXED	0	1	192	70	70	0	30	3,325.56	17.32	182
	8	PLS-MIXED	0	1	256	89	89	0	11	3,746.85	14.64	182

IBM POWER9
32c, 256t, Launch 2018

Workload PLS-FIBO and PLS-PRIME – single thread computing time of algorithms

Run	Test	Workload	Parameter	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	Operations total [Mops]	Operations per cpu [Mops]	Elapsed time [s]
4	12	PLS-FIBO	42	1	1	3	0	3	97	0.00	0.00	171
	18	PLS-PRIME	8000	1	1	3	0	3	97	0.00	0.00	110

peakmarks

Simple. Representative. Fast.

Summary of Scripts and Commands for Monitoring peakmarks® Key Performance Metrics

Summary of Scripts and Commands

Scripts to monitor peakmarks key performance metrics

SQL> @kpm_all

SQL> @kpm_query
SQL> @kpm_scan
SQL> @kpm_ioread
SQL> @kpm_iowrite
SQL> @kpm_lgwr
SQL> @kpm_dbwr
SQL> @kpm_dl
SQL> @kpm_da
SQL> @kpm_tp
SQL> @kpm_tpplus
SQL> @kpm_tpio
SQL> @kpm_pls

peakmarks Mission

Identify Key Performance Metrics for Oracle Database Platforms.

On-Premises and in the Cloud.

For Quality Assurance, Evaluations, and Capacity Planning.