Benchmarking Database Cloud Services

May 2024

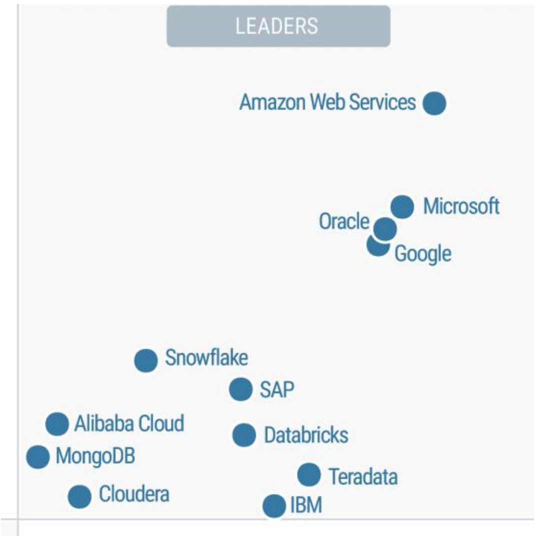
contact@peakmarks.com

Foreword

«Benchmarking Database Cloud Services» was published for the first time at the 11th TPC Technology Conference 2019 in Los Angeles[1]. The peakmarks® software presented there has since been further developed and adapted to new technologies. It is now available in version 10 for all Oracle database platforms from Oracle version 19c with the CDB architecture.

This publication is an updated version of the original TPC publication.

Introduction



Source: *Magic Quadrant for Cloud Database Management Systems*; Gartner, December 2022

Even for the well-known cloud market leader[2], there are no comparable and comprehensible performance indicators for the operation of Oracle databases.

Questions about the price-performance ratio, the scope of Oracle licenses required when using the BYOL model[3], capacity planning when transitioning from on-premises installations to the cloud, and the stability and scalability of a cloud service under maximum load, among other topics, cannot be answered without further information. Cloud service providers change their infrastructure components and configurations without prior notice. How do you ensure you still receive the same service as a customer?

Proof-of-concepts may not always be the most suitable approach due to their complexity, time-consuming nature, and high cost.

Peakmarks has developed benchmark software to quickly and reliably determine easy-to-understand key performance indicators for an Oracle platform - whether on-premises or in the cloud.

[1] Drozd: *Benchmarking Database Cloud Services*. In: Nambiar, R., Poess, M. (eds.) Performance Evaluation and Benchmarking for the Era of Cloud(s), Lecture Notes in Computer Science (LNCS), vol. 12257, pp. 139-153. Springer, Switzerland 2020.

[2] Gartner: *Magic Quadrant for Cloud Database Management*. Gartner Report, December 2022.

[3] BYOL (bring your own license) is a licensing model that allows customers to transfer existing on-premises licenses to the cloud.

The peakmarks® Software

The peakmarks® software allows for the quick and reliable determination of easily understandable performance indicators for Oracle Database Services, whether on-premises or in the cloud. These indicators are essential for various tasks.

Evaluation. Performance indicators enable fact-based decisions when selecting cloud services, components, configurations, and options.

Capacity planning. Performance indicators help with sound capacity planning when migrating databases to new platforms, whether on-premises or in the cloud.

License cost optimization. In many cases, license costs far exceed infrastructure costs. Optimizing the infrastructure can significantly reduce license costs. Performance indicators help quickly identify the most suitable infrastructure components.

Quality assurance. Cloud services can change their infrastructure at any time and without prior notice. By regularly checking performance, providers' performance promises can be verified periodically - quickly, reliably, and comprehensibly (Oracle AWR reports).

However, it is also important to check the performance characteristics of on-premises installations after making changes to the infrastructure, such as replacing components or modifying the configuration.

Requirements for Benchmark Tools

Huppler[4] describes a good benchmark's five most important characteristics: relevant, repeatable, fair, comprehensible, and economical. peakmarks® fulfills these requirements.

However, other features are also decisive for the customer's acceptance of benchmark tools.

Simplicity. Installing the software, running the benchmark, and interpreting the results must be easy.

peakmarks® is implemented with the tools of the database. Therefore, peakmarks® runs unchanged wherever the Oracle database software is available. Without additional software, without additional hardware, without additional scripting. Every database administrator can run the benchmark software.

Speed. Installation, data loading, processing of the various workloads, and evaluation of the performance indicators should be fast.

peakmarks® can be installed in a few hours, including all database adjustments. The synthetic benchmark database's loading time depends on the database's size and the infrastructure's performance. The scalable loading process can be adapted to the database platform's performance and reach up to 5 TByte per hour on powerful systems. Complete benchmark runs with all workloads take 12 and 24 hours; the results are available immediately. A comprehensive performance assessment can be completed within a week.

[4] Huppler: *The Art of Building a Good Benchmark*. In: Nambiar, R., Poess, M. (eds.) Performance Evaluation and Benchmarking, LNCS, vol. 5895, pp. 18-30. Springer, Heidelberg 2009.

Understandable performance metrics. Many benchmarks only provide a single performance metric. This simplifies the comparison of different systems. However, a single complex metric is difficult to interpret[5].

peakmarks® offers a range of representative and easy-to-understand key figures for various aspects. This makes it easier to answer specific questions about performance and detects performance bottlenecks and malfunctions more quickly.

Different load situations. It is often not the maximum value of a key performance indicator that is of interest but the optimum performance range where a permanent and predictable performance occurs.

peakmarks® analyzes a database service's performance in different load situations. Workloads start with a low load and continuously increase until the system is saturated. This method can determine a database service's optimum performance range.

Continuous further development. In contrast to many open-source benchmark tools, the peakmarks® software is continuously developed.

The requirements of new hardware technologies such as flash storage and persistent memory are incorporated into further development, as are new database technologies (smart scan, in-memory column store, memory-optimized tables, etc.). The list of workloads is supplemented where relevant for measuring the infrastructure and selecting the solution architecture.

For more in-depth considerations on the systematic benchmarking of cloud services, please refer to the literature by Bermbach, Wittern, and Tai[6].

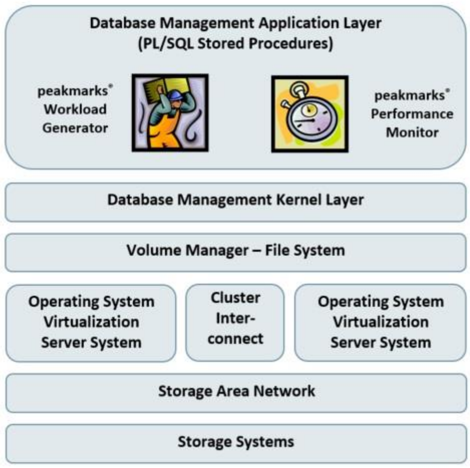
[5] Crolotte: *Issues in Benchmark Metric Selection*. In: Nambiar, R., Poess, M. (eds.) Performance Evaluation and Benchmarking 2009, LNCS, vol. 5895, pp. 146-152. Springer, Heidelberg 2009.

[6] Bermbach, Wittern, Tai: *Cloud Service Benchmarking - Measuring Quality of Cloud Services from a Client Perspective*. Springer International Publishing 2017.

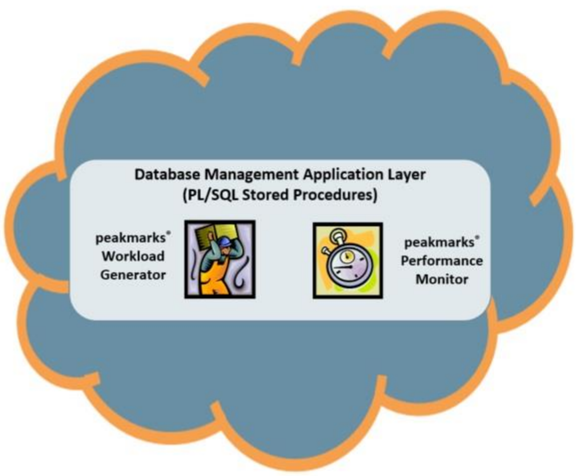
The Architecture of the peakmarks® Software

The peakmarks® software is stored in the database using stored procedures. Tools such as Putty are required to access the database server. Tools based on SCP are used to transfer the benchmark software and exchange the results. Otherwise, no additional hardware or software installations or scripting are required.

The peakmarks® software consists of two components. A workload generator creates the database load, the so-called workload. All workloads are generated within the database and executed by database jobs.



peakmarks® Software On-Premises



peakmarks® Software in the Cloud

A performance monitor collects all relevant statistics before and after each performance test and immediately displays the result in an easy-to-read tabular form. Oracle AWR reports are generated for a detailed analysis. Performance is measured at the application's interface, i.e., it is directly available to the application.

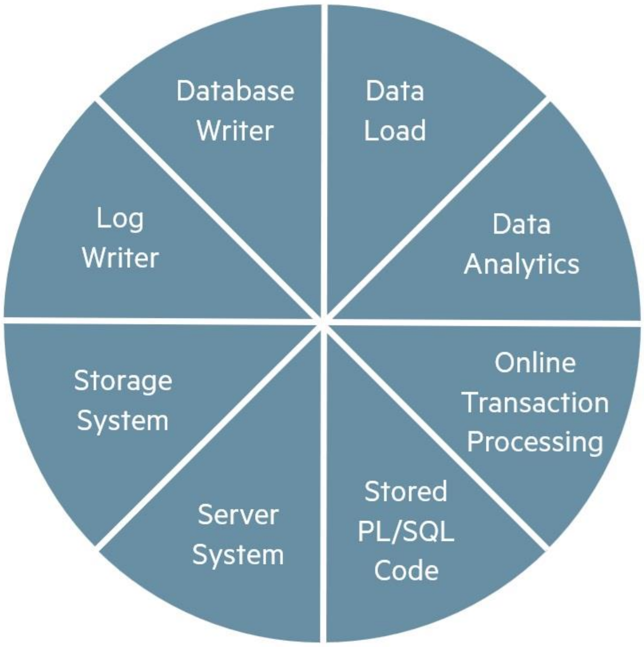
The synthetic benchmark database can be adapted to customer requirements using various peakmarks® configuration parameters:

- Number of CPUs.
- Size of the database between 64 GByte and 64 TByte per database instance.
- Size of the database buffer cache.
- Optional use of Database Flash Cache for conventional database servers or Cell Flash Cache for Oracle Engineered Systems.
- The duration of performance tests can be set between 1 minute and 12 hours.

Oracle features such as Database Encryption, Data Guard, or Real Application Clusters with several database servers are supported.

Complete 360-degree performance overview

With over 30 workloads, so-called micro-benchmarks, in 8 workload groups, the peakmarks® software provides a representative and complete performance overview of an Oracle Database Service in all load situations.



The two most important components of a platform - the server system and storage system - have a decisive influence on the performance of all database operations and determine Oracle license and maintenance costs. One group of workloads, therefore, analyzes the performance of a server system in Oracle database operation, and another determines the performance behavior of a storage system in Oracle database operation.

Two service processes significantly impact the smooth operation of the database. The log writer processes are responsible for transaction management, and the database writer processes are responsible for buffer management. The peakmarks® software examines the performance behavior of these two important database service processes with specially developed workloads.

Performance indicators for representative database operations, such as data load, data analytics, and transaction processing, are required for capacity planning. Different technologies can be used here. The data load workloads determine the throughput for loading data using methods such as buffered load, direct load, and stream load for IoT applications. The Data Analytics workloads examine the performance of various technologies for accelerated searching of non-indexed data, such as smart scan technology on Oracle Engineered Systems or in-memory technology, which is available on all platforms. The online transaction processing workloads determine the transaction throughput and response time behavior for transactions of different complexity. If configured, performance-enhancing technologies such as memory-optimized tables are also used here to process lookup queries quickly. Some applications encapsulate centrally important business functions and transactions in stored procedures. The peakmarks® software offers workloads for analyzing the performance behavior of PL/SQL application programs on different processors.

All key performance indicators are presented for system engineers in a clear and easy-to-read peakmarks® Performance Report and summarized in a peakmarks® Executive Summary for decision-makers[7]. Oracle AWR reports are automatically generated to provide a detailed analysis of all performance results.

[7] Examples of such reports can be found at www.peakmarks.com.

Configuration of peakmarks® Performance Tests

The peakmarks® software requires the following information to configure a single performance test:

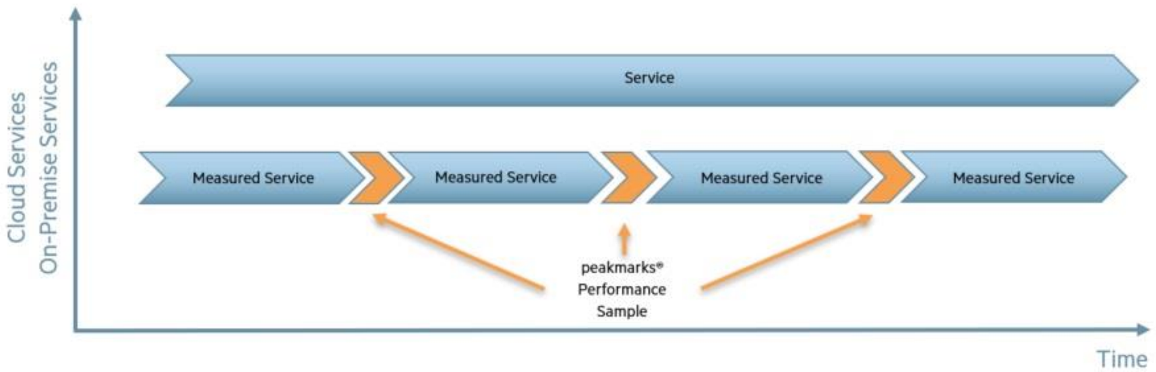
- Workload to be carried out.
- There are optional workload parameters for some workloads.
- Internal Oracle parallelism (DOP) can be selected for some workloads.
- Number of cluster nodes on which the workload is to be executed.
- Number of processes that execute the workload. In a cluster, the processes are distributed evenly across all cluster nodes in a round-robin fashion.
- Target runtime for the performance test.

To measure a database service completely, workloads are configured for different load situations. This requires a sequence of performance tests that can be configured in various ways.

Smart Benchmark Configuration. The Smart Benchmark Configuration is fully automated. It is the most convenient and fastest method for obtaining a complete overview of an Oracle Database Service's performance. A sequence of performance tests (based on the Oracle configuration parameter CPU_COUNT) is automatically generated for each workload. The test sequence is aborted if the system reaches saturation as the load increases.

Manual Benchmark Configuration. The Manual Benchmark Configuration offers maximum flexibility in the configuration of benchmark tests. Each parameter of a performance test can be selected individually. This method is preferred by engineers to analyze special load situations in detail.

Sample Benchmark Configuration. The Sample Benchmark Configuration also runs completely automatically but only performs a selected number of pre-configured performance tests. This benchmark configuration is often used with cloud services to periodically check within a 60-minute maintenance window whether the database service has fully retained its performance characteristics.



Structure of the peakmarks® Reports

Units of measurement and abbreviations

The following measurement units and abbreviations are used in the peakmarks® reports; the columns of the respective key performance metrics are marked in red.

Abbreviation	Meaning	Abbreviation	Meaning
[MBps]	megabyte per second	[s]	seconds
[GBps]	gigabyte per second	[ms]	milliseconds
[TBph]	terabyte per hour	[µs]	microseconds
[rps]	rows per second	[dbps]	database blocks per second
[qps]	queries per second	[dbpt]	database blocks per transaction
[tps]	transactions per second	[rbps]	redo blocks per second
[sps]	sql executes per second	[kBpt]	kiloByte per transaction
[IOPS]	I/O operations per second	[Mops]	million operations per seconds
BuCache	Oracle buffer cache	FlCache	Flash Cache (Database, Exadata)

Examples

The first example shows the system performance when running the server workload SRV-QUERY25. In the 2-node cluster, each server has 192 cores and 384 threads.

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	12	SRV-QUERY25	2	2	1	0	0	99	58,091	29,046	0.034	1,701,794	850,897	99.89	181
	13	SRV-QUERY25	2	192	24	24	0	76	4,110,983	21,411	0.046	115,336,687	600,712	100.00	182
	14	SRV-QUERY25	2	384	48	47	0	52	7,492,380	19,511	0.051	210,081,611	547,088	100.00	182
	15	SRV-QUERY25	2	576	70	69	0	30	8,400,864	14,585	0.068	235,458,379	408,782	100.00	182
	16	SRV-QUERY25	2	768	91	90	0	9	9,175,440	11,947	0.083	256,996,831	334,631	100.00	182

The second example shows the system performance when executing the STO-RANDOM storage workload (80% read, 20% write).

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	97	STO-RANDOM	20	2	2	1	1	1	0	99	0	77,338	0.191	605	605	100.00	53.12	182
	98	STO-RANDOM	20	2	16	1	5	4	1	95	0	537,260	0.222	4,227	4,227	100.00	24.52	183
	99	STO-RANDOM	20	2	32	1	7	6	1	93	0	912,569	0.262	7,182	7,182	100.00	21.56	183
	100	STO-RANDOM	20	2	48	1	9	7	1	91	0	1,220,600	0.292	9,606	9,606	100.00	20.62	183
	101	STO-RANDOM	20	2	64	1	10	8	2	90	0	1,359,028	0.309	10,695	10,695	100.00	20.50	183

Description of the performance reports

The first column of each benchmark report identifies the benchmark run (Run column). Within a benchmark run, all performance tests are numbered consecutively (Test column). The third column describes the workload. Optionally, there are additional workload parameters depending on the workload. In the second example, this is the percentage of write operations (column Wri [%]).

The Nodes column indicates the number of cluster nodes involved in a test. The Jobs column documents the number of processes on all cluster nodes that generate the load. For cluster tests, the processes are distributed round-robin to all cluster nodes.

For some workloads, there is an additional column, DOP (degree of parallelism), which describes the Oracle internal parallelization.

Choice of measurement points for CPU-bound performance tests

For CPU-bound workloads such as SRV-QUERY25, the peakmarks® software automatically selects five measurement points to determine the server's performance in different load situations: Single-thread, 25% CPU utilization, 50% CPU utilization, 75% CPU utilization, and 100% CPU utilization.

Choice of measuring points for other workloads

The peakmarks® software selects up to thirty-two measurement points for all other workloads. Starting with one process, the load automatically increases until the system reaches saturation. The load increase is determined automatically depending on the Oracle configuration parameter CPU_COUNT value and can be manually overridden. In the second example, the measurements are aborted after the fifth test with the STO-RANDOM workload because performance does not increase further. The system is saturated.

CPU utilization

The following four columns describe the CPU utilization of all servers involved. For workloads with a high I/O share, an additional column describes the percentage CPU I/O wait share (column CPU iow [%]).

Key performance indicators

Columns with the key performance indicators follow this. Depending on the workload, different performance key figures are listed. In the first example, the number of queries (Queries total [qps] column), the average response time of the queries (Response time [ms] column), and the number of logical database block accesses in the database buffer cache (Log reads total [dbps] column).

In the second example, the two key performance indicators are the number of I/O operations (Phys reads total [IOPS] column) and the I/O service time (IO Time read [ms] column).

The performance figures often include additional columns that show the performance per CPU used[8]. These key figures are important for price/performance comparisons to determine performance per thread or core. Oracle license and maintenance costs are generally based on the number of sockets (Standard Edition) and cores (Enterprise Edition) used.

Further key figures

This is followed by columns with further key figures. These key figures often provide additional information to understand the system behavior better or check whether the workload has run optimally.

In our first example, these are the Oracle buffer cache hit rate (BuCache read [%] column) and the workload's processing time (Elapsed time [s] column). The second example shows the Flash Cache Hitrate (column FlCache read [%]).

Workloads for Server Systems

peakmarks® offers various workloads for server systems. The selected server workloads occur in all database applications across all industries. They show the real-world performance of a server system in Oracle database operation.

All server workloads access tables via SQL with different access patterns. The tables concerned are completely in the buffer cache. There are virtually no I/O operations, meaning that the server performance completely limits the performance of these workloads.

Server performance tests answer questions about server scalability, the influence of NUMA effects, and the efficiency of technologies such as multithreading, virtualization, and data encryption. For on-premises platforms, they enable a simple price/performance comparison of processors and servers, taking Oracle licensing into account. Even when evaluating cloud platforms, where the hardware components are not always known, they provide comparable key figures on a cloud provider's price/performance ratio.

Performance key figures for server systems in Oracle database operation

The following performance indicators have proven useful for measuring the performance of server systems in Oracle database operation:

- Throughput of SQL queries (query throughput) in queries per second [qps].
- Average response time of SQL queries (query response time) in milliseconds [ms].
- Number of logical database block accesses in the buffer cache (logical reads throughput) in database blocks per second [dbps].
- SQL Buffer Cache Scan Rate in MegaByte per second [MBps].

[8] The term CPU is based on the definition of the Oracle configuration parameter CPU_COUNT. For processors without multithreading, one CPU corresponds to one core; for systems with multithreading, one CPU corresponds to one thread. In both cases, the performance specification corresponds to the maximum performance that a single application process can obtain. Only in the case of Oracle internal parallelization can an application process use the performance of several threads.

Workloads for determining server performance in Oracle database operation

The following table shows the workloads determining server performance in Oracle database operation.

Workload	Action	Performance metrics	Unit of measurement
SRV-QUERY1	Query type A: Select 1 data record via a unique index. Examples: Select account, product, order, invoice, etc. This workload shows the maximum query throughput of a server and the minimum query response time.	Query throughput Query response time	[qps] [ms]
SRV-QUERY25	Query type B: Select an average of 25 data records via a non-unique index. Examples: Select account postings from the last week; list items of an order, etc.	Query throughput Query response time	[qps] [ms]
SRV-REPORT	Online Report: Select an average of 125 data records via a non-unique index. Examples: Online report mobile phone call records from last month, online report e-banking with account mutations from last month, etc. This workload corresponds to a simple online report and shows the maximum value for logical reads of a server.	Block accesses in the buffer cache	[dbps]
SRV-SCAN	Data search without index. This corresponds to a full table scan. The table's entire data volume in the buffer cache must be searched sequentially.	Scan rate in the buffer cache	[MBps]
SRV-MIXED	Complex workload with a mix of equally weighted simple workloads running simultaneously: ▪ SRV-QUERY1 ▪ SRV-QUERY25 ▪ SRV-REPORT ▪ SRV-SCAN SRV-MIXED displays the server system performance in Oracle database operation for a representative cross-section of peakmarks® server workloads.	Query throughput Query response time	[qps] [ms]

Example: Workload SRV-QUERY1

The workload SRV-QUERY1 (1 data record per query) shows the maximum possible number of queries and the shortest response time of queries for a server system.

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	7	SRV-QUERY1	2	2	1	0	0	99	274,511	137,256	0.007	1,169,565	584,783	99.85	181
	8	SRV-QUERY1	2	192	25	24	0	75	15,764,764	82,108	0.012	63,131,433	328,810	100.00	182
	9	SRV-QUERY1	2	384	48	47	0	52	23,143,866	60,270	0.016	92,647,479	241,269	100.00	182
	10	SRV-QUERY1	2	576	70	69	0	30	26,333,715	45,718	0.022	105,406,443	182,997	100.00	182
	11	SRV-QUERY1	2	768	91	90	0	9	30,068,291	39,151	0.025	120,342,777	156,696	100.00	182

Example: Workload SRV-REPORT

The workload SRV-REPORT (online report with an average of 125 data records per query) shows a database server's maximum number of logical reads.

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	2	SRV-REPORT	2	2	1	0	0	99	16,569	8,285	0.120	2,194,460	1,097,230	100.00	181
	3	SRV-REPORT	2	192	24	24	0	76	980,029	5,104	0.195	125,547,207	653,892	100.00	182
	4	SRV-REPORT	2	384	48	47	0	52	1,739,893	4,531	0.219	222,763,561	580,113	100.00	182
	5	SRV-REPORT	2	576	71	70	0	29	2,018,212	3,504	0.284	258,193,416	448,252	100.00	182
	6	SRV-REPORT	2	768	91	90	0	9	2,121,253	2,762	0.360	270,964,318	352,818	100.00	182

Example: Workload SRV-SCAN

The SRV-SCAN workload (query with a full table scan) shows the scan rate in the buffer cache.

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]
4	17	SRV-SCAN	2	2	1	0	0	99	9,169	4,585	1,174,876	587,438	100.00	181
	18	SRV-SCAN	2	192	25	24	0	75	558,287	2,908	71,315,571	371,435	100.00	183
	19	SRV-SCAN	2	384	48	48	0	52	934,129	2,433	119,452,244	311,074	100.00	183
	20	SRV-SCAN	2	576	71	70	0	29	1,073,604	1,864	137,396,812	238,536	100.00	183
	21	SRV-SCAN	2	768	91	90	0	9	1,199,605	1,562	153,531,372	199,911	100.00	183

Example: Workload SRV-MIXED

The workload SRV-MIXED is often used to quickly obtain a representative statement on the performance of a server system in Oracle database operation, e.g., to compare common processor architectures such as ARM, AMD EPYC, IBM POWER, IBM Z or Intel Xeon. Performance figures according to SPEC[9] do not always correlate with the performance of a processor in Oracle database.

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]
5	2	SRV-MIXED	2	8	1	1	0	99	351,483	43,935	0.023	5,311,292	663,912	100.00	183
	3	SRV-MIXED	2	192	25	24	0	75	6,572,595	34,232	0.029	99,839,421	519,997	100.00	183
	4	SRV-MIXED	2	384	48	48	0	52	10,834,520	28,215	0.035	176,951,554	460,811	100.00	183
	5	SRV-MIXED	2	576	70	69	0	30	11,725,318	20,356	0.049	200,746,793	348,519	100.00	183
	6	SRV-MIXED	2	768	91	90	0	9	13,074,900	17,025	0.058	219,067,369	285,244	100.00	182

[9] See www.spec.org.

Notes

Main memory accesses[10] **enable ultimate database performance.** The SRV-QUERY1 workload determines extremely short response times between 7 and 25 µs.

Misleading CPU utilization. If processors are operated with multithreading, the information on CPU utilization is not always correct. In the SRV-QUERY1 workload, all cores are occupied by processes in Run 4, Test 9. The CPU utilization shows 50%, and 23,143,866 queries are processed per second. In test 11, all threads are occupied by processes, the CPU utilization increases to 91%, and 30,068,291 queries per second are processed. The server is now almost fully utilized. Between the two tests, CPU utilization increased by 82%, but throughput only by 30%. This means that in test 9, the actual CPU utilization was already 77%, not 50%!

When calculating CPU utilization, many operating systems treat threads as resources, just like cores with their hardware for executing instructions. But they are not. Threads are merely a concept for improving the parallel processing of instructions.

Scalability with increasing load. For each server workload, we observe a decreasing performance per CPU with increasing load. For example, the throughput per CPU for the SRV-MIXED workload drops from 43.935 qps (Run 5, Test 2, Single Thread Performance) to 17.025 qps (Run 5, Test 6, System Saturation). The performance per CPU drops to approx. 39% of the peak value as the load increases!

We observe this behavior in processors with a large clock rate spread. In this example, the maximum clock rate is 3.7 GHz. As the load increases, the clock rate is reduced to 2.6 GHz for thermal reasons. This alone means a 30% reduction in performance.

As a rule, processors with many cores also have a high clock rate spread. Conversely, processors with fewer cores have a lower clock rate spread and can guarantee a more predictable performance.

Server evaluation for Oracle database operation. Various aspects must be considered when evaluating servers. Processors with fewer cores are recommended to achieve predictable and persistent CPU performance in all load ranges. Such processors deliver the highest per-core performance and, therefore, improve a server's price/performance ratio, taking Oracle license and maintenance costs into account.

More sockets may then be necessary to meet scalability requirements. This also allows for significantly higher main memory capacities, leading to further performance improvement and benefits using in-memory technologies. However, NUMA effects are expected with such architectures in Oracle database operation.

Influence of the clock rate on performance. The clock rate of processors is irrelevant when comparing the performance of different architectures, as different tasks are performed during a cycle depending on the architecture.

[10] Just to recap, as of May 2024, the fastest processors take less than 0.1 nanoseconds for DRAM access. In the Oracle buffer cache, the database takes approximately 0.5 microseconds for block access, and for flash storage, it takes between 20 and 500 microseconds for block access, depending on the storage fabric and protocol.

Workloads for Storage Systems

Conventional I/O benchmark tools such as fio, vdbench, iometer, and Orion often show performance values not achieved in real database operations. This is due to the complexity of database I/O operations, which these tools do not take into account.

When a data block is read, the buffer cache management of the database must fulfill many tasks:

- find a free space for the block.
- if there is no free space, replace older blocks.
- synchronize all database processes simultaneously, trying to occupy free space in the buffer cache.
- when using a shared disk cluster architecture (Oracle Real Application Cluster), synchronization must take place cluster-wide, which in the best case requires the exchange of information (messages) between cluster nodes; in the worst case, additionally, the exchange of database blocks. Even the exchange of messages can lead to a limitation of scalability.
- finally, blocks are checked for integrity and consistency during I/O transfer. Oracle configuration parameters define the scope of the integrity and consistency tests, which can add additional load to processors and the I/O system.

The peakmarks® software generates I/O load with so-called SQL-generated I/O operations to determine representative performance figures for the storage system in Oracle database operations. The intra-SQL parallelism can also be tested for sequential operations to find the optimum SQL parallelism (Oracle DOP).

Storage performance tests provide answers to questions about the efficiency of protocols such as NVMe over fabrics (IP, FC), RoCE (RDMA over Converged Ethernet), and technologies such as flash storage, persistent memory, storage tiering, data encryption, offload functions of intelligent storage systems and data reduction through de-duplication and compression. They provide well-founded key figures for price/performance comparisons for storage systems for Oracle databases and for evaluating suitable storage replication processes.

Key performance indicators for storage systems in Oracle database operation

The following key performance indicators are determined to measure the performance of storage systems in database operation:

- Throughput for sequential reading of storage systems (sequential read) in megabytes per second [MBps]. Large I/O units of 1 MByte are typically used.
- Random access of storage systems in I/O operations per second [IOPS]. Smaller I/O units, such as Oracle database blocks, are typically used. The size of Oracle database blocks is configurable: 2 KByte, 4 KByte, 8 KByte, 16 KByte, 32 KByte. The default value is 8 KByte.
- I/O service time for random access in milliseconds [ms].

Workloads to determine the storage performance in database operation

The following table shows the peakmarks® workloads for measuring the performance of storage systems in Oracle database operation.

Workload	Action	Performance metrics	Unit of measurement
STO-READ	Sequential read operation generated by SQL statement.	Throughput sequential I/O	[MBps]
STO-OFFLOAD	Sequential read operation generated by SQL statement when using intelligent storage servers with offload functionality (smart scan).	Throughput sequential I/O	[MBps]
STO-RANDOM	Randomly scattered read and write operations of individual database blocks generated by SQL statement In this workload, read operations are carried out by foreground processes and write operations by background processes. The ratio of database blocks that are read or written is configurable[11].	Throughput random I/O I/O service time	[IOPS] [ms]
STO-SCATTER	Write operations generated by SQL statements. With this workload, write operations are carried out by foreground processes.	Throughput random I/O	[dbps]

Example: Workload STO-READ

The STO-READ workload shows the I/O throughput between the database server and storage system during sequential read operations. In this example, no Oracle DOP is used. The data volume is loaded from the flash cache (FlCache read column).

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	1	STO-READ	0	2	2	1	0	0	0	100	0	5,719	0.280	5,696	5,696	100.00	0.00	176
	2	STO-READ	0	2	16	1	1	1	0	99	0	19,389	0.563	19,332	19,332	100.00	0.00	242
	3	STO-READ	0	2	32	1	1	1	1	99	0	38,249	0.876	38,170	38,170	100.00	0.00	215
	4	STO-READ	0	2	48	1	2	2	0	98	0	29,089	1.153	28,967	28,967	100.00	0.00	298
	5	STO-READ	0	2	64	1	2	1	0	98	0	33,122	1.894	33,016	33,016	100.00	0.00	280
	6	STO-READ	0	2	80	1	2	1	1	98	0	35,254	2.019	35,139	35,139	100.00	0.00	294

Example: Workload STO-OFFLOAD

The STO-OFFLOAD workload is similar to STO-READ, but it uses the storage systems' intelligence and Smart Scan technology. When using Smart Scan technology, there is no need to specify a DOP, as the storage systems optimize access independently.

[11] Oracle often performs write operations as multiblock I/O operations, i.e. a list of blocks is written with one I/O operation.

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	33	STO-OFFLOAD	0	2	2	1	0	0	0	100	0	96,393	0.232	96,171	96,171	100.00	0.00	172
	34	STO-OFFLOAD	0	2	16	1	1	0	0	99	0	167,855	0.393	167,481	167,481	100.00	0.00	175
	35	STO-OFFLOAD	0	2	32	1	1	0	0	99	0	168,018	0.980	167,645	167,645	100.00	0.00	179
	36	STO-OFFLOAD	0	2	48	1	1	0	0	99	0	171,495	1.537	171,115	171,115	100.00	0.00	180
	37	STO-OFFLOAD	0	2	64	1	1	0	0	99	0	167,633	2.157	167,258	167,258	100.00	0.00	183

Example: Workload STO-RANDOM

The STO-RANDOM workload shows the I/O throughput between the database server and storage system with random access to individual database blocks. In this example, the size of the Oracle database blocks is 8 kByte.

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	65	STO-RANDOM	0	2	2	1	1	1	0	99	0	82,641	0.271	650	650	100.00	47.51	182
	67	STO-RANDOM	0	2	32	1	5	3	1	95	0	1,225,625	0.248	9,640	9,640	100.00	6.72	183
	69	STO-RANDOM	0	2	64	1	9	6	2	91	0	2,394,143	0.239	18,829	18,829	100.00	4.54	183
	71	STO-RANDOM	0	2	96	1	12	8	3	88	0	3,141,738	0.265	24,706	24,706	100.00	4.06	183
	73	STO-RANDOM	0	2	128	1	14	9	3	86	0	3,356,962	0.329	26,399	26,399	100.00	4.04	183
	75	STO-RANDOM	0	2	160	1	15	10	4	85	0	3,561,487	0.410	28,010	28,010	100.00	4.05	182
	77	STO-RANDOM	0	2	192	1	16	10	4	84	0	3,718,834	0.525	29,245	29,245	100.00	4.02	182
	79	STO-RANDOM	0	2	224	1	17	11	4	83	0	3,816,575	0.659	30,014	30,014	100.00	4.05	182
	81	STO-RANDOM	0	2	256	1	18	11	5	82	0	3,900,201	0.862	30,673	30,673	100.00	4.03	182

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	97	STO-RANDOM	20	2	2	1	1	1	0	99	0	77,338	0.191	605	605	100.00	53.12	182
	98	STO-RANDOM	20	2	16	1	5	4	1	95	0	537,260	0.222	4,227	4,227	100.00	24.52	183
	99	STO-RANDOM	20	2	32	1	7	6	1	93	0	912,569	0.262	7,182	7,182	100.00	21.56	183
	100	STO-RANDOM	20	2	48	1	9	7	1	91	0	1,220,600	0.292	9,606	9,606	100.00	20.62	183
	101	STO-RANDOM	20	2	64	1	10	8	2	90	0	1,359,028	0.309	10,695	10,695	100.00	20.50	183
	102	STO-RANDOM	20	2	80	1	10	7	2	90	0	1,293,471	0.346	10,179	10,179	100.00	21.09	183

Only read operations are measured in the first benchmark report (Run 6, Test 65 to 81). In the second benchmark report (Run 6, Test 97 - 102), the ratio of read-to-write operations is 80:20 (column Wri [%]).

Notes

I/O bandwidth. With the STO-READ workload, the amount of data transferred is usually limited by the I/O bandwidth between the server system and the storage system.

I/O service time. With the STO-RANDOM workload, the I/O service time **is** influenced by the efficiency of the I/O stack. Service times of less than 500 microseconds are expected when using all-flash arrays. Thanks to new protocols such as RoCE (RDMA over converged Ethernet), I/O service times of less than 20 microseconds are now possible if data is found in the respective cache of the storage system.

Tiered storage. The systems measured here have tiered storage with three cache layers. The I/O throughput and I/O service time depend heavily on which cache a block is in. Repeating the tests and "warming up" the cache layers could significantly improve throughput and I/O service time. Tiered storage systems do not deliver predictable performance per se.

Log Writer Workloads

Log Writer processes are responsible for transaction logging and database recovery after system failures. These processes are critical to Oracle's overall performance when processing transactions. Transaction logging latency can significantly impact user transaction response time.

Optionally, the Log Writer processes can also be used for replication to synchronize standby databases. This technology is very popular for disaster recovery solutions. Replication can take place in synchronous or asynchronous mode. The data transfer between the primary and standby databases can optionally be encrypted and/or compressed. With synchronous replication, local transactions must wait until the standby databases have received the transaction log. This can significantly delay local transaction processing[12].

Key performance indicators for log writer processes

The key performance indicators for Log Writer processes are:

- Number of SQL commit operations in transactions per second [tps].
- Average latency time for SQL commit operations in milliseconds [ms].
- REDO throughput in megabytes per second [MBps].

Workloads for log writer processes

To analyze the performance of the Log Writer processes for different transaction sizes, the peakmarks® software offers different workloads.

Workload	Action	Performance metrics	Unit of measurement
LGWR-LAT1 LGWT-LAT25 LGWR-LAT125	Transactions with approx. 1, 25, or 125 kByte REDO data volume per transaction. All transactions conclude with a COMMIT WAIT IMMEDIATE.	SQL Commit operations SQL Commit latency	[tps] [ms]
LGWR-THR	Transactions that generated a large amount of REDO data. All transactions conclude with a COMMIT NOWAIT BATCH.	Log Writer throughput	[MBps]

[12] When using Oracle Data Guard as a BCP solution, synchronous replication is used for *mission-critical* systems. This means an additional latency of at least 1 ms for the log writers.

Example: Workload LGWR-LAT1

The LGWR-LAT1 workload shows the maximum number of commit operations the system can process. The average transaction size, measured in terms of the REDO data volume, is just under 1.4 kByte (REDO data [kBpt] column).

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]
7	1	LGWR-LAT1	2	2	1	0	0	99	0	26	8,385	0.238	0.179	26,351	8,485	11	1.394	100.00	181
	4	LGWR-LAT1	2	72	3	2	0	97	0	26	121,052	0.592	0.491	422,741	10,712	168	1.421	100.00	182
	7	LGWR-LAT1	2	144	4	4	1	96	0	26	189,004	0.758	0.633	611,871	8,730	258	1.399	100.00	182
	10	LGWR-LAT1	2	216	5	4	1	95	0	26	223,780	0.959	0.790	699,801	5,950	305	1.398	100.00	182
	13	LGWR-LAT1	2	288	6	5	1	94	0	26	243,390	1.173	0.954	750,839	3,862	334	1.403	100.00	183
	16	LGWR-LAT1	2	360	8	6	1	92	0	26	276,053	1.298	1.080	841,050	3,346	378	1.404	93.44	182
	19	LGWR-LAT1	2	432	8	7	1	92	0	26	309,943	1.385	1.164	932,873	2,959	424	1.400	100.00	182
	22	LGWR-LAT1	2	504	8	7	1	92	0	26	323,753	1.550	1.231	969,780	2,895	442	1.399	100.00	183

Example: Workload LGWR-THR

The workload LGWR-THR shows the maximum amount of REDO data that can be processed by the Log Writer processes. The number of Log Writer processes is not configurable and is determined automatically by Oracle.

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]
7	97	LGWR-THR	2	2	1	1	0	99	0	26	598	3.344	0.685	349,037	636	165	282.312	100.00	182
	98	LGWR-THR	2	24	6	5	1	94	0	26	5,387	4.377	1.073	2,971,762	3,881	1,396	265.381	100.00	184
	99	LGWR-THR	2	48	7	6	1	93	0	26	7,972	6.002	2.501	4,181,495	2,660	1,969	252.936	100.00	182
	100	LGWR-THR	2	72	7	6	1	93	0	26	9,523	7.549	3.635	4,810,385	2,013	2,268	243.858	100.00	182
	101	LGWR-THR	2	96	8	6	1	92	0	26	10,108	9.455	4.870	4,999,847	1,761	2,358	238.906	100.00	181

Notes

Technologies for data reduction on storage systems. Some cloud service providers use storage systems with data reduction processes (deduplication, compression). These technologies reduce the required storage capacity but can negatively impact the latency of log writer processes and the transaction throughput of OLTP systems.

Database Writer Workload

The database writers are responsible for managing the buffer cache. These processes are crucial for Oracle's overall performance when many blocks are changed, e.g., during conventional data loading via the buffer cache or intensive transaction processing.

Oracle automatically determines the number of Database Writer processes. If this value is not sufficient, it can be adjusted manually.

Performance key figures for database writer processes

The key performance indicator for Database Writer processes is

- Throughput of database blocks written back in blocks per second [dbps] or megabyte per second [MBps]

Workloads for determining the database writer performance

The DBWR-THR workload is available to determine Database Writer performance.

Workload	Action	Performance metrics	Unit of measurement
DBWR-THR	Massive block changes in the buffer cache	Database Writer Throughput	[MBps]

Example: Workload DBWR-THR

The DBWR-THR workload shows the maximum number of blocks the Database Writer processes can write back.

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	#prc DBWR	#check points	Phys writes total [dbps]	Phys writes total [IOPS]	Phys writes total [MBps]	REDO data [MBps]	Phys writes FlCache [MBps]	FlCache write [%]	Elapsed time [s]
9	1	DBWR-THR	2	2	1	1	0	99	0	96	1,159	5,407	1,490	48	6	48	100.00	285
	2	DBWR-THR	2	8	4	3	1	96	0	96	74,972	296,745	40,583	2,558	226	2,558	100.00	182
	3	DBWR-THR	2	16	6	5	1	94	0	96	77,168	585,938	80,048	5,095	419	5,027	98.67	181
	4	DBWR-THR	2	24	7	5	1	93	0	96	40,958	703,502	85,162	6,168	636	6,168	100.00	183
	5	DBWR-THR	2	32	8	5	1	92	0	96	20,627	781,537	132,586	6,900	756	6,900	100.00	182
	6	DBWR-THR	2	40	9	6	2	91	0	96	13,616	844,609	181,547	7,753	845	7,484	96.53	183

Notes

ASM redundancy. Various redundancy levels can be defined when configuring ASM (external, normal redundancy, high redundancy). This has a significant impact on the performance of write processes. In our example, the ASM level high redundancy is used. The database measured the data volume of 844,609 database blocks per second specified in the peakmarks® report. However, ASM has to write three times the amount of data to various disk groups.

Data Load Workloads

System architects and capacity planners need performance metrics from database services regarding their ability to load data. This is particularly important for data warehouse and data analytics systems, where data volumes are increasing and loading time windows are decreasing.

Oracle provides various technologies for loading data:

- Conventional loading via the buffer cache; this technology is preferably used in transaction processing systems.
- Direct loading by bypassing the buffer cache; this method is mainly used in the data warehouse and data analytics environment to load large amounts of data as quickly as possible.
- As of Oracle 19c, a new loading process is available for IOT (Internet of Things) applications. Mass data can be loaded quickly, but transaction security is not always guaranteed. This loading technology is acceptable if aggregates, such as average values, etc., are not dependent on each data record.

Key figures for data load performance

There is only one performance indicator for data load workloads:

- Throughput of loaded data (data load rate) in megabyte per second [MBps] or terabyte per hour ([TBph].

Workloads for determining the data load performance

The peakmarks® software offers workloads for all three data load technologies.

Workload	Action	Performance metrics	Unit of measurement
DL-BUFFER	The data is program-generated and loaded conventionally via a buffer cache. 3 additional indices are updated when loading. Typically, only a few data records are loaded per transaction. The loading process uses COMMIT WRITE WAIT IMMEDIATE to complete the transaction.	Data load throughput	[MBps]
DL-DIRECT	The data is loaded directly from a data source, bypassing the buffer cache. Typically, large amounts of data are loaded per transaction. Only one additional index is updated during loading. The loading process uses the NOLOGGING option and COMMIT WRITE NOWAIT BATCH on transaction completion.	Data load throughput	[MBps]

Workload	Action	Performance metrics	Unit of measurement
DL-STREAM	The data is program-generated and uses the memory-optimized row store for data loading in IOT applications. Only one additional index is updated during loading. The loading process does not use a commit but only a flush operation of the memory-optimized pool.	Data load throughput	[MBps]

Example: Workload DL-BUFFER

The DL-BUFFER workload shows the maximum load rate achieved with the buffered data load. This loading method is preferably used in transaction-oriented systems.

The metric Load rate total [MBps] corresponds to the amount of data loaded. The metric Load rate total [rps] depends on the data model and applies here for a row length of 300 bytes (record length of the peakmarks® tables).

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	1	DL-BUFFER	5	2	2	1	0	0	99	0	32,882	9	7,595	57	31	100.00	183
	5	DL-BUFFER	5	2	32	2	2	0	98	0	255,386	73	17,982	476	236	100.00	182
	9	DL-BUFFER	5	2	64	3	2	0	97	0	382,675	109	18,572	773	363	100.00	182
	13	DL-BUFFER	5	2	96	4	4	1	96	0	497,906	142	23,741	1,005	451	100.00	183
	17	DL-BUFFER	5	2	128	6	5	1	94	0	590,132	169	29,693	1,238	534	100.00	183
	21	DL-BUFFER	5	2	160	7	5	1	93	0	651,361	186	36,340	1,460	598	100.00	182
	25	DL-BUFFER	5	2	192	7	6	1	93	0	710,089	203	44,101	1,686	660	100.00	183
	29	DL-BUFFER	5	2	224	8	7	1	92	0	747,770	214	51,627	1,875	698	100.00	183

Example: Workload DL-DIRECT

The DL-DIRECT workload shows the maximum load rate achieved with direct data loading bypassing the buffer cache. This loading method is preferably used in data warehouse systems.

Here, too, the metric Load rate total [MBps] corresponds to the amount of data loaded. The metric Load rate total [rps] depends on the data model and applies here for 300 bytes (row length of the peakmarks® tables).

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	33	DL-DIRECT	125,000	2	2	1	0	0	99	0	1,288,740	369	2,857	568	86	100.00	182
	34	DL-DIRECT	125,000	2	8	2	1	0	98	0	5,164,404	1,478	9,618	2,273	344	100.00	182
	35	DL-DIRECT	125,000	2	16	3	2	0	97	0	9,566,459	2,737	18,218	4,258	635	100.00	182
	36	DL-DIRECT	125,000	2	24	4	3	1	96	0	13,406,989	3,836	27,418	6,052	887	100.00	182
	37	DL-DIRECT	125,000	2	32	5	4	1	95	0	14,293,862	4,090	32,624	7,155	954	90.56	182
	38	DL-DIRECT	125,000	2	40	6	5	1	94	0	16,476,603	4,714	37,637	8,462	1,081	88.02	182
	39	DL-DIRECT	125,000	2	48	6	5	1	94	0	17,363,471	4,968	40,809	8,689	1,137	90.66	182
	40	DL-DIRECT	125,000	2	56	6	5	1	94	0	17,263,001	4,939	42,547	9,892	1,131	79.49	182

Data Analytics Workloads

System architects and capacity planners require performance metrics from database services regarding their ability to search large amounts of data. These applications are typically based on full table scan operations and are not supported by index structures.

Full table scans' performance depends on the data's position in the memory hierarchy (storage, main memory) and the technology used to increase scanning performance (Smart Scan, In-Memory Column Store).

Smart Scan technology is only available on Oracle Engineered Systems and requires an additional license. The In-Memory technology is available on all platforms and is license-free from Oracle 21c up to a certain capacity.

Data analytics workloads answer questions about both technologies' efficiency and enable price/performance comparisons.

Key figures for data analytics performance

There is only one key performance indicator for data analytics workloads:

- Throughput of scanned data (data scan rate) in megabyte per second [MBps] or rows per second [rps]

Workloads for determining the data analytics performance

peakmarks® offers workloads for testing different data locations (storage, main memory) and boost technologies (Smart Scan, In-Memory Column Store). Intra-SQL parallelism can also be changed here to find the optimum database parallelism (Oracle DOP) for the data scan.

Workload	Action	Performance metrics	Unit of measurement
DA-STORAGE	Full table scan with low complexity aggregate. The data is read from the storage system.	Data scan throughput	[MBps]
DA-OFFLOAD	Full table scan with low complexity aggregate. The intelligent storage system reads the data with offload functionality (smart scan).	Data scan throughput	[MBps]
DA-ROWSTORE	Full table scan with low complexity aggregate. The data is read from the row store of the buffer cache.	Data scan throughput	[MBps]
DA-COLSTORE	Full table scan with low complexity aggregate. The data is processed by the column store of the buffer cache.	Data scan throughput	[rps]

Example: Workload DA-STORAGE

The DA-STORAGE workload displays the maximum scan rate when all data must be read from the storage. The processing of the data scan (filter operations) takes place on the database server.

The metric Scan rate total [MBps] corresponds to the amount of data scanned. The metric Scan rate total [rps] depends on the data model and applies here for 300 bytes (row length of the peakmarks® tables).

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]
11	1	DA-STORAGE	2	2	1	1	0	0	99	0	2,866,393	1,494	100.00	0.00	176
	3	DA-STORAGE	2	32	1	3	2	1	97	0	86,138,194	29,786	100.00	0.00	176
	5	DA-STORAGE	2	64	1	6	5	1	94	0	101,988,920	35,174	100.00	0.00	177
	7	DA-STORAGE	2	96	1	9	7	1	91	0	110,251,520	37,984	100.00	0.00	178
	9	DA-STORAGE	2	128	1	10	9	1	90	0	115,716,037	39,839	100.00	0.00	179
	11	DA-STORAGE	2	160	1	11	10	1	89	0	123,544,507	42,498	100.00	0.00	180

Example: Workload DA-OFFLOAD

The DA-OFFLOAD workload shows the maximum scan rate when the intelligent storage system processes all data. Filter operations and aggregates can also be processed on the storage system. Only the number of hit rows is transferred to the database server.

The metric Scan rate total [MBps] corresponds to the amount of data scanned. The metric Scan rate total [rps] corresponds to the equivalent amount of data with a record length of 300 bytes (record length of the peakmarks® tables).

This offload technology is only available on Oracle Engineered Systems and requires additional licenses for their storage servers.

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]
11	33	DA-OFFLOAD	2	2	1	0	0	0	100	0	205,377,286	69,813	100.00	0.00	171
	34	DA-OFFLOAD	2	16	1	1	0	0	99	0	364,233,795	123,813	100.00	0.00	171
	35	DA-OFFLOAD	2	32	1	1	0	0	99	0	362,256,280	123,140	100.00	0.00	171
	36	DA-OFFLOAD	2	48	1	1	0	0	99	0	358,280,094	121,789	100.00	0.00	172
	37	DA-OFFLOAD	2	64	1	1	0	0	99	0	362,118,861	123,094	100.00	0.00	172

Example: Workload DA-ROWSTORE

The DA-ROWSTORE workload displays the maximum scan rate when all data is processed in the database server's buffer cache. This form of processing is also available for the Oracle Standard Edition. A large main memory capacity is advantageous for storing as much data as possible in the buffer cache.

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]
11	97	DA-ROWSTORE	2	2	1	1	0	0	99	0	29,030,670	9,255	0.00	100.00	182
	98	DA-ROWSTORE	2	192	1	25	24	0	75	0	1,867,213,752	562,610	0.00	100.00	183
	99	DA-ROWSTORE	2	384	1	48	47	0	52	0	3,090,006,400	930,714	0.00	100.00	183
	100	DA-ROWSTORE	2	576	1	70	69	0	30	0	3,607,255,308	1,086,438	0.00	100.00	183
	101	DA-ROWSTORE	2	768	1	87	86	0	13	0	3,953,619,845	1,190,720	0.00	100.00	183

Example: Workload DA-COLSTORE

The DA-COLSTORE workload displays the maximum scan rate when all data is processed in the database server's column store. This in-memory technology requires corresponding licenses on the database server.

It is unnecessary to store entire tables in the column store, but only individual selected columns frequently used as search criteria. Compression methods can also increase the amount of data stored in the column store. A large main memory capacity is also advantageous to store as much data as possible in the 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]
11	65	DA-COLSTORE	2	2	1	1	0	0	99	0	21,075,613,283	6,709,882	0.00	100.00	181
	68	DA-COLSTORE	2	48	1	7	6	0	93	0	471,912,961,217	150,245,219	0.00	100.00	182
	71	DA-COLSTORE	2	96	1	13	12	0	87	0	921,355,549,887	293,336,361	0.00	100.00	182
	74	DA-COLSTORE	2	144	1	19	19	0	81	0	1,361,359,924,528	433,422,239	0.00	100.00	182
	77	DA-COLSTORE	2	192	1	25	25	0	75	0	1,782,368,941,499	567,460,248	0.00	100.00	182
	80	DA-COLSTORE	2	240	1	31	30	0	69	0	2,170,081,743,882	690,897,496	0.00	100.00	182
	83	DA-COLSTORE	2	288	1	36	36	0	64	0	2,409,617,664,730	767,161,259	0.00	100.00	182
	86	DA-COLSTORE	2	336	1	42	42	0	58	0	2,603,380,187,024	828,852,194	0.00	100.00	182
	89	DA-COLSTORE	2	384	1	48	47	0	52	0	2,761,724,145,692	879,266,412	0.00	100.00	182

The metric Scan rate total [rps] corresponds to the amount of data scanned. The metric Scan rate total [MBps] depends on the data model and applies to 300 bytes (row length of the peakmarks® tables).

Notes

Efficiency of the various processes. It is evident that in-memory technology with the column store offers the highest performance, surpassing smart scan technology. Consequently, in-memory technology has become the preferred choice for data analytics applications in recent years..

Price/performance comparison of the various processes. Additional licenses are required for both Smart Scan technology and in-memory technology. If you consider the scan rates concerning the license costs, the in-memory technology has a significantly better price/performance ratio than the Smart Scan technology.

However, Smart Scan technology works largely transparently for applications, and changes to the code are rarely necessary. Smart Scan technology can be combined with in-memory technology on the storage servers and can manage immense data capacities. However, Smart Scan technology can only be used on Oracle Engineered Systems.

In-memory technology, on the other hand, is available on all platforms, requires no special hardware, and delivers ultimate performance. However, the amount of data is limited by the main memory capacity (DRAM) of the database server. As a rule, interventions must be made in the application in order to specify which columns of a table are to be included in the column store. Compression procedures help to make optimum use of the capacity of the column store.

Online Transaction Processing Workloads

System architects and capacity planners need performance metrics from database services regarding their ability to execute typical transactions.

Online Transaction Processing is an extremely complex form of data processing. All platform components and database service processes are significantly involved and must be precisely balanced. Online Transaction Processing is the supreme discipline of database platforms.

Transaction processing performance depends on many factors. The most important are

- Ratio of database and buffer cache size, expressed in the buffer cache hit rate.
- I/O performance when reading data.
- I/O performance when writing the transaction log by REDO Log Writer processes.
- I/O performance when writing data through Database Writer processes.

Storage tiering and many other factors can influence the I/O performance when reading data. The I/O performance of the log writer processes can be influenced by the data guard performance or other technologies for mirroring data between different data centers.

Key figures for transaction processing performance

The following performance indicators are of interest for transaction processing workloads:

- Throughput of SQL transactions (transaction throughput) in transactions per second [tps].
- Average response time of SQL transactions (transaction response time) in milliseconds [ms].

Overview of transaction processing performance workloads

The peakmarks® software provides different types of workloads:

- Reading workloads TP-LOOKUP and TP-REPORT.
- Mutating workloads of varying complexity TP-LIGHT, TP-MEDIUM and TP-HEAVY. The ratio between read and change operations can be configured for these workloads.
- Mix of workloads with SELECT, UPDATE, and INSERT statements, whereby several simple TP workloads run simultaneously and equally weighted and are supplemented by a DL-BUFFER workload.

  TP-MIXED1 is a read-intensive transaction mix, and TP-MIXED2 is a write-intensive workload.

  These two workloads are preferably used to quickly determine a representative statement on the transaction processing performance of an Oracle database platform.

Workload	Action	Performance metrics	Unit of measurement
TP-LOOKUP	Fast look-up query that checks the existence of a single data record. Examples: Check credit card, car registration, phone number, etc. If available on the platform, this workload uses memory-optimized tables for fast access via hash key. Otherwise, conventional tables with a unique b-tree index are used.	Query throughput Query response time	[tps] [ms]
TP-REPORT	Online report that selects an average of 25 data records per query.	Query throughput Query response time	[tps] [ms]
TP-LIGHT	Select/mutate individual data sets via index. Examples: Select/mutate account, product, order, invoice, etc. Transactions are always concluded with a COMMIT WRITE WAIT IMMEDIATE.	Transaction throughput Transaction response time	[tps] [ms]
TP-MEDIUM	Select/mutate Ø 25 data records via index. Examples: Select/mutate account postings from last week, article list of an order, etc. Transactions are always concluded with a COMMIT WRITE WAIT IMMEDIATE.	Transaction throughput Transaction response time	[tps] [ms]
TP-HEAVY	Select/mutate Ø 125 data records via index. Transactions are always concluded with a COMMIT WRITE WAIT IMMEDIATE. Examples: Select/mutate mobile phone call records from the last month, etc.	Transaction throughput Transaction response time	[tps] [ms]
TP-MIXED1	Balanced read-intensive workload mix: ▪ TP-REPORT ▪ TP-LOOKUP ▪ TP-MEDIUM (40% UPDATE) ▪ DL-BUFFER (2 rpt) Logical block access: 83% read, 17% write. Avg REDO per transaction 256-byte.	Transaction throughput Transaction response time	[tps] [ms]
TP-MIXED2	Balanced write-intensive workload mix: ▪ TP-LIGHT (40% UPDATE ▪ TP-MEDIUM (30% UPDATE) ▪ TP-HEAVY (20% UPDATE) ▪ DL-BUFFER (3 rpt) Logical block access: 65% read, 35% write. Avg REDO per transaction 1,725-byte.	Transaction throughput Transaction response time	[tps] [ms]

As a special feature, the peakmarks® software allows for defining a customer-specific transaction mix.

Workload	Action	Performance metrics	Unit of measurement
TP-CUSTOM	Configurable workload mix based on simple workloads.	Transaction throughput Transaction response time	[tps] [ms]

Example: Read-intensive workload TP-MIXED1

The overview report with the key performance indicators includes both the primary performance indicators such as throughput and response time of transactions as well as important influencing factors: Buffer cache hit rate (column BuCache read [%]), Flash cache hit rate (column FlCache read [%]), the service time for random read (column IO time read [ms]) and the latency time from the log writer on transaction completion (column Log file syn [ms]).

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

Example: Write-intensive workload TP-MIXED1

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]
14	2	TP-MIXED2	N/A	2	8	2	1	1	98	0	7,493	1.062	0.202	0.221	2.976	61.03	100.05	182
	3	TP-MIXED2	N/A	2	24	5	3	1	95	0	25,606	0.931	0.240	0.356	3.095	59.16	100.00	182
	4	TP-MIXED2	N/A	2	48	8	5	2	92	0	40,160	1.189	0.281	0.584	3.120	54.22	100.00	182
	5	TP-MIXED2	N/A	2	72	10	6	2	90	0	50,730	1.413	0.335	0.775	3.167	52.25	100.00	182
	6	TP-MIXED2	N/A	2	96	11	6	3	89	0	57,857	1.649	0.396	0.939	3.236	50.37	100.00	182
	7	TP-MIXED2	N/A	2	120	12	7	3	88	0	63,575	1.875	0.458	1.111	3.248	49.40	100.00	182
	8	TP-MIXED2	N/A	2	144	12	7	3	88	0	68,316	2.098	0.520	1.255	3.252	49.05	100.00	182
	9	TP-MIXED2	N/A	2	168	13	7	3	87	0	71,298	2.346	0.587	1.423	3.287	48.55	100.00	182
	10	TP-MIXED2	N/A	2	192	13	8	3	87	0	74,309	2.571	0.660	1.579	3.293	48.42	100.00	182
	11	TP-MIXED2	N/A	2	216	14	8	3	86	0	76,117	2.822	0.732	1.739	3.304	47.99	100.00	182

Notes

peakmarks® provides additional reports for analyzing the performance of the platform during online transaction processing with the following information: logical I/O per transaction (read, write), physical I/O per transaction (read, write), REDO data volume per transaction, number of checkpoints, etc.

The two examples show only moderate CPU utilization. The bottleneck can be quickly identified using AWR reports, which also contain recommendations from the Automatic Database Diagnostic Monitor (ADDM). In this case, the buffer cache is too small for a 16 TByte database.

Performance of PL/SQL application programs

PL/SQL is the preferred programming language for complex transaction logic and algorithms. PL/SQL programs are stored and compiled in the database server. Some large applications, such as core banking systems, are largely implemented in PL/SQL.

Key figures for PL/SQL performance

The following performance indicators are of interest for PL/SQL workloads:

- Number of PL/SQL operations executed per time unit in million operations per second [Mops].
- Execution time of PL/SQL algorithms in seconds [s].

Workloads for determining PL/SQL performance

peakmarks® offers workloads to test the efficiency of PL/SQL programs on a specific processor. These workloads are completely limited by the CPU performance.

Workload	Action	Performance metrics	Unit of measurement
PLS-ADD	Addition of numbers of different data types	Throughput PL/SQL operations	[Mops]
PLS-BUILTIN	Data type-specific operations, including SQL functions based on core banking systems and telco billing systems	Throughput PL/SQL operations	[Mops]
PLS-PRIME	Calculation of the first n prime numbers	Algorithm processing time	[s]
PLS-FIBO	Calculation of the first n Fibonacci numbers (recursive algorithm)	Algorithm processing time	[s]
PLS-MIXED	Complex workload with a mix of equally weighted simple workloads running simultaneously: ▪ PLS_ADD with data type NUMBER ▪ PLS_ADD with data type PLS_INTEGER ▪ PLS_BUILTIN with data type NUMBER ▪ PLS_BUILTIN with data type VARCHAR2	Throughput PL/SQL operations	[Mops]

Example: Workload PLS-MIXED

The PLS-MIXED workload shows the throughput of different PL/SQL operations with different 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]
16	1	PLS-MIXED	N/A	2	8	1	1	0	99	642.55	80.32	182
	2	PLS-MIXED	N/A	2	192	25	24	0	75	15,374.16	80.07	182
	3	PLS-MIXED	N/A	2	384	48	47	0	52	26,792.12	69.77	182
	4	PLS-MIXED	N/A	2	576	71	70	0	29	33,060.21	57.40	182
	5	PLS-MIXED	N/A	2	768	91	90	0	9	39,005.97	50.79	182

Example: Workload PLS-FIBO

The PLS-FIBO workload shows the processing time for an algorithm to calculate the Fibonacci number series unti N (column Parameter). This workload uses a recursive algorithm. The PLS-PRIME workload determines all prime numbers between 1 and N (column Parameter).

Run	Test	Workload	Parameter	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	Elapsed time [s]
15	31	PLS-FIBO	42	2	1	1	0	0	99	42
15	32	PLS-PRIME	8000	2	1	1	0	0	100	47

Conclusion

peakmarks® Software provides system engineers and system architects with a robust and comprehensive framework for determining meaningful and understandable performance metrics of Oracle Database Services - **on-premise** and in the **cloud**.

The peakmarks® performance indicators are easy to understand, allowing comparisons of technologies, configurations, components, and complete systems across manufacturers and providers.

The peakmarks® software covers all important workloads for assessing infrastructure performance. It is constantly being further developed and adapted to new database versions.

Benefits for IT operations. Many IT organizations can reduce the cost of database services by opting for best-in-class infrastructure components or cloud services and minimizing licensing and maintenance costs.

Bottlenecks, misconfigurations, and malfunctions are detected before commissioning. Regular performance analyses of cloud services ensure the quality of performance. peakmarks® performance indicators are a solid basis for capacity planning when migrating to other platforms or cloud services.

The result is Oracle Database Services with predictable and persistent performance in all load situations.

Advantages for IT hardware vendors and service providers. peakmarks® Performance Assessments are fast and efficient and support developing and marketing license-optimized Oracle database solutions with the best price/performance ratio.

Customers and interested parties benefit from the comprehensible, user-friendly, and comparable performance indicators, which can reduce time-consuming and expensive proofs of concept. peakmarks® performance indicators help strengthen the positioning of a provider in comparison to the competition.

Summary of Performance Results

In a peakmarks® Executive Summary, the most important results are summarized comprehensibly for decision-makers.

peakmarks® key performance metrics for platform components in Oracle database operation

Category	Key performance metrics	Response time	Throughput	peakmarks® Workload
Server System	Mix of queries and scans	0.023 - 0.058 ms	13,074,900 qps	SRV-MIXED
	Buffer cache logical read throughput	-	270,964,318 dbps	SRV-REPORT
	Buffer cache scan rate	-	1,199,605 MBps	SRV-SCAN

Category	Key performance metrics	Throughput	peakmarks® Workload
Storage System	SQL sequential I/O	35,139 MBps	STO-READ
	SQL sequential I/O - using smart scan	171,115 MBps	STO-OFFLOAD
	SQL random read, 8 kByte, 100% read, < 500 µs	3,668,554 IOPS	STO-RANDOM
	SQL random read, 8 kByte, 80% read, < 500 µs	1,359,028 IOPS	STO-RANDOM
	SQL random write	363,251 dbps	STO-SCATTER

peakmarks® key performance metrics for critical database background processes

Category	Key performance metrics	Latency time	Throughput	peakmarks® Workload
Log Writer	Commit throughput and latency	0.238 - 1.581 ms	331,844 tps	LGWR-LAT1
	Throughput of REDO data		2,358 MBps	LGWR-THR

Category	Key performance metrics	Throughput	peakmarks® Workload
Database Writer	Database blocks written	844,609 dbps	DBWR-THR

peakmarks® key performance metrics for representative database operations

Category	Key performance metrics	Throughput		peakmarks® Workload
Data Load	Data load throughput - using the buffer cache Transactional data load, 5 rpt	221 774,029	MBps rps	DL-BUFFER
	Data load throughput - bypassing the buffer cache Data Warehouse data load, bulk load 125,000 rpt	4,968 17,363,471	MBps rps	DL-DIRECT
	Data load throughput - streamed Internet-of-things data load, 100 rpt	- -	MBps rps	DL-STREAM

Category	Key performance metrics	Throughput		peakmarks® Workload
Data Analytics	Data scan throughput - using storage system	42,498 123,544,507	MBps rps	DA-STORAGE
	Data scan throughput - using smart scan	123,094 362,118,861	MBps rps	DA-OFFLOAD
	Data scan throughput - using row store	1,190,720 3,953,619,845	MBps rps	DA-ROWSTORE
	Data scan throughput - using column store	879,266,412 2,761,724,145,692	MBps rps	DA-COLSTORE

Category	Key performance metrics	Response time		Throughput		peakmarks® Workload
Online Transaction Processing	Transaction throughput	0.030 - 0.051	ms	11,163,923	tps	TP-MIXED1
	Transaction throughput	0.931 - 3.333	ms	78,822	tps	TP-MIXED2

peakmarks® Key Performance Metrics for PL/SQL application code

Category	Key performance metrics	Response time		Throughput		peakmarks® Workload
PL/SQL	PL/SQL operations throughput			39,005	Mops	PLS-MIXED
	PL/SQL algorithm execution time	42	s	-		PLS-FIBO (42)
	PL/SQL algorithm execution time	47	s	-		PLS-PRIME(8k)