peakmarks® Benchmark Software

August 2023

contact@peakmarks.com

Vorwort

Die peakmarks® Benchmark Software wurde anlässlich der 11th TPC Technology Conference 2019 in Los Angeles einem breiteren Benchmark Publikum vorgestellt[1]. Seither wurde die peakmarks® Benchmark Software an neue Technologien angepasst und steht nun in der Version 10 für alle Oracle Datenbank Plattformen ab Oracle 19c mit und ohne CDB-Architektur zur Verfügung. Diese Publikation ist eine aktualisierte Version der ursprünglichen TPC-Veröffentlichung.

Einleitung

Die peakmarks® Benchmark Software ermöglicht die schnelle und zuverlässige Ermittlung von allgemein verständlichen Performance Kennzahlen für Oracle Database Services, sowohl On-premises als auch in der Cloud.

Performance Kennzahlen von Oracle Database Services sind für verschiedene Aufgaben notwendig.

Evaluation. Selbst für die bekannten Cloud Services von Amazon, Google und Microsoft findet man keine vergleichbaren und nachvollziehbaren Performance Kennzahlen. Überlegungen zum Preis-/Leistungsvergleich der verschiedenen Angebote sind ohne Performance Kennzahlen nicht möglich. Performance Kennzahlen ermöglichen faktenbasierte Entscheidungen.

Kapazitätsplanung. Bei der Migration von Datenbanken auf neue Plattformen, sei es On-premises oder in der Cloud, helfen Performance Kennzahlen bei einer fundierten Kapazitätsplanung.

Lizenzkostenoptimierung. Lizenzkosten übersteigen häufig bei weitem die Infrastrukturkosten. Durch Optimierung der Infrastruktur können Lizenzkosten häufig deutlich reduziert werden. Die am besten geeigneten Infrastruktur-Komponenten werden mit Hilfe von Performance Kennzahlen schnell und eindeutig identifiziert.

Qualitätssicherung. Cloud Services können ihre Infrastruktur jederzeit und ohne Vorankündigung ändern. Durch eine regelmässige Überprüfung der Performance können die Leistungsversprechen der Anbieter periodisch überprüft werden - schnell und zuverlässig.

[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).

Anforderungen an Benchmark Tools

Huppler[2] beschreibt die fünf wichtigsten Eigenschaften für einen guten Benchmark: relevant, wiederholbar, fair, nachvollziehbar und wirtschaftlich. peakmarks® erfüllt all diese Anforderungen.

Aber auch andere Eigenschaften sind für die Kundenakzeptanz von Benchmark Tools entscheidend.

Einfachheit. Es muss einfach sein, die Benchmark Software zu installieren, den Benchmark durchzuführen und die Ergebnisse zu interpretieren. peakmarks® ist mit den Werkzeugen der Datenbank und ohne Betriebssystemskripte implementiert. Daher läuft peakmarks® unverändert überall dort, wo die Oracle Datenbanksoftware verfügbar ist. Jeder Datenbank-Administrator kann die Benchmark Software ohne zusätzliches Know-how einfach betreiben.

Geschwindigkeit. Die Installation, das Laden der Daten, die Verarbeitung der verschiedenen Workloads und die Auswertung der Performance Kennzahlen sollte schnell erfolgen. peakmarks® kann in wenigen Stunden installiert werden, einschliesslich aller Anpassungen der Datenbank. Die Ladezeit der Datenbank hängt von der Datenbankgrösse und der Performance der Infrastruktur ab. Der skalierbare Ladevorgang passt sich automatisch an die Leistungsfähigkeit der Datenbankplattform an. Auf leistungsfähigen Systemen werden Ladezeiten von 4 TByte pro Stunde gemessen. Komplette Benchmark-Läufe mit allen Workloads dauern zwischen 12 und 24 Stunden; die Ergebnisse sind sofort verfügbar. Ein umfassendes Benchmark Projekt kann innerhalb einer Woche abgeschlossen werden.

Verständliche Performance Metriken. Viele Benchmarks liefern nur eine einzige Leistungskennzahl. Dies vereinfacht den Vergleich verschiedener Systeme. Eine einzige komplex erstellte Metrik ist jedoch schwer zu interpretieren[3]. peakmarks® bietet eine Reihe von repräsentativen und leicht verständlichen Kennzahlen für verschiedene Aspekte. Konkrete Fragen zur Performance lassen sich so leichter beantworten. Leistungsengpässe und Fehlfunktionen werden schneller erkannt.

Unterschiedliche Lastsituationen. Oft ist nicht der Maximalwert einer Leistungskennzahl von Interesse, sondern der optimale Leistungsbereich, wo eine dauerhafte und vorhersehbare Leistungsabgabe erfolgt. peakmarks® analysiert die Performance eines Database Services in verschiedenen Lastsituationen. Workloads beginnen mit niedriger Last. Die Last wird kontinuierlich bis zur Systemsättigung erhöht. Auf diese Weise kann der optimale Leistungsbereich eines Database Services ermittelt werden.

Stetige Weiterentwicklung. Im Gegensatz zu vielen Open Source Benchmark Tools wird die peakmarks® Benchmark Software kontinuierlich weiterentwickelt. Die Anforderungen durch neue Hardware-Technologien wie *Flash Storage* und *Persistant Memory* fliessen genauso in die Weiterentwicklung ein, wie neue Datenbank Technologien (*Smart Scan*, *In-Memory Column Store*, *Memory-Optimized Tables*, etc.). Die Liste der Workloads wird ergänzt, sofern dies für die Vermessung der Infrastruktur und der Wahl der Lösungsarchitektur von Relevanz sind.

Für tiefergehende Betrachtungen zum systematischen Benchmarking von Cloud Services sei auf die Literatur von Bermbach, Wittern und Tai hingewiesen[4].

---

[2] 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).

[3] 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).

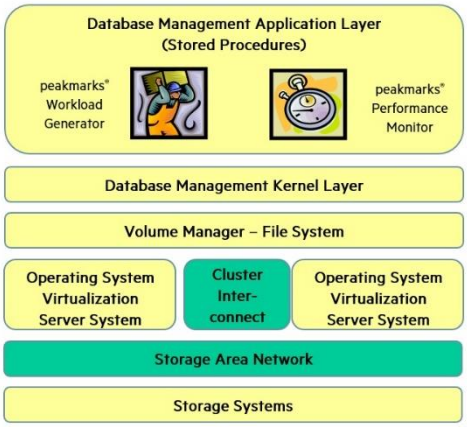
[4] Bermbach, Wittern, Tai.: *Cloud Service Benchmarking – Measuring Quality of Cloud Services from a Client Perspective*. Springer International Publishing 2017.

Die Architektur der peakmarks® Benchmark Software

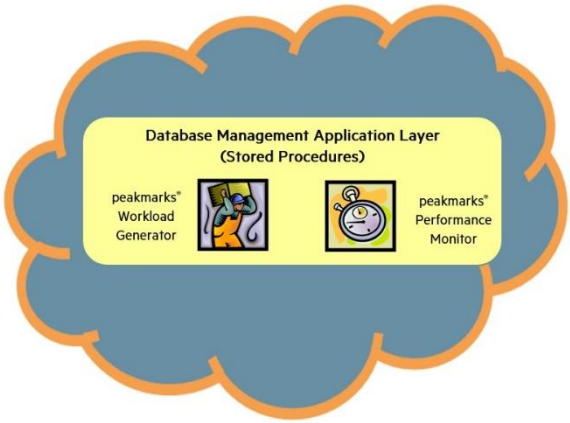
Die peakmarks® Benchmark Software ist in Form von Stored Procedures in der Datenbank gespeichert. Für den Zugang zum Database Server werden Tools wie Putty benötigt. Für den Transfer der Benchmark Software und den Austausch der Ergebnisse werden Tools auf der Basis von SCP eingesetzt.

Die peakmarks® Benchmark Software besteht aus zwei Komponenten. Ein Workload Generator erzeugt die Datenbanklast, sogenannte Workloads. Alle Workloads werden innerhalb der Datenbank erzeugt und von Datenbank Jobs ausgeführt.

Ein Performance Monitor sammelt vor und nach jedem einzelnen Performance Test alle relevanten Statistiken und zeigt unmittelbar das Ergebnis in tabellarischer Form an. Die Performance wird an der Schnittstelle zur Applikation gemessen, d.h. die gemessene Performance steht der Applikation direkt zur Verfügung.



peakmarks® Benchmark Software
On-Premise



peakmarks® Benchmark Software
in der Cloud

Über verschiedene peakmarks® Konfigurationsparameter kann die synthetisch Benchmark Datenbank an Kundenbedürfnisse angepasst werden.

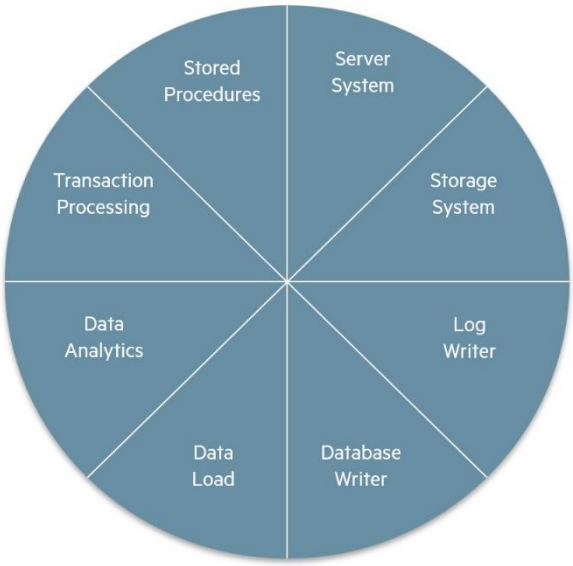
- Anzahl CPUs.
- Grösse Datenbank Buffer Cache.
- Datenbankgrösse zwischen 64 GByte und 64 TByte pro Datenbank Instanz.
- Optionaler Einsatz von Database Flash Cache für konventionelle Datenbank Server und Cell Flash Cache für Engineered Systems.

Die Laufzeit von Performance Tests kann zwischen 60 Sekunden und 12 Stunden festgelegt werden.

Optional kann Oracle Data Guard konfiguriert werden. Die Oracle Real Application Cluster Technologie mit mehreren Datenbank Servern wird ebenfalls unterstützt.

Vollständiger 360 Grad Performance Übersicht

Mit über 30 Workloads, sogenannte Micro-Benchmarks, in 8 Workload Gruppen ermöglicht die peakmarks® Benchmark Software einen repräsentativen und vollständigen Überblick zur Leistungsfähigkeit eines Oracle Database Services in allen Lastsituationen.



Die beiden wichtigsten Komponenten einer Plattform - Server System und Storage System - haben einen entscheidenden Einfluss auf die Performance aller Datenbank Operationen und bestimmen Oracle Lizenz- und Wartungskosten. Eine Gruppe von Workloads analysiert daher das Leistungsverhalten eines Server Systems im Oracle Datenbankbetrieb. Eine weitere Workload Gruppe ermittelt das Leistungsverhalten eines Storage Systems im Oracle Datenbankbetrieb.

Für einen reibungslosen Betrieb der Datenbank sind zwei Service-Prozesse von grosser Bedeutung. Log-Writer Prozesse sind für das Transaktions-Management verantwortlich, Database-Writer Prozesse für das Buffer-Management. Die peakmarks® Benchmark Software untersucht mit eigens entwickelten Workloads das Leistungsverhalten dieser beiden wichtigen Datenbank Service-Prozesse.

Für die Kapazitätsplanung sind Performance Kennzahlen repräsentativer Datenbank Operationen wie Data Load, Data Analytics und Transaction Processing notwendig. Dabei können unterschiedliche Technologien zum Einsatz kommen.

Die Data Load Workloads ermitteln den Durchsatz zum Laden von Daten bei Einsatz verschiedener Verfahren wie *Buffered Load*, *Direct Load* und *Stream Load* für IoT-Anwendungen.

Die Data Analytics Workloads untersuchen die Leistungsfähigkeit verschiedener Technologien zum beschleunigten Durchsuchen von nicht-indizierten Daten wie etwa der *Smart Scan* Technologie auf Oracle Engineered Systems oder der *In-Memory* Technologie, die auf allen Plattformen verfügbar ist.

Die Transaction Processing Workloads ermitteln den Transaktionsdurchsatz und das Antwortzeitverhalten für Transaktionen unterschiedlicher Komplexität.

Einige Applikationen kapseln zentral wichtige Business-Funktionen und Transaktionen in Stored Procedures. Die peakmarks® Benchmark Software bietet Workloads zur Analyse des Leistungsverhalten von PL/SQL Applikationsprogrammen auf unterschiedlichen Prozessoren an.

Alle wesentlichen Performance Kennzahlen werden in peakmarks® eigenen Reports dargestellt und in einem Performance Zertifikat zusammengefasst. Für eine detaillierte Analyse aller Performancetests werden automatisch Oracle AWR-Reports erzeugt.

Konfiguration von peakmarks® Performance Tests

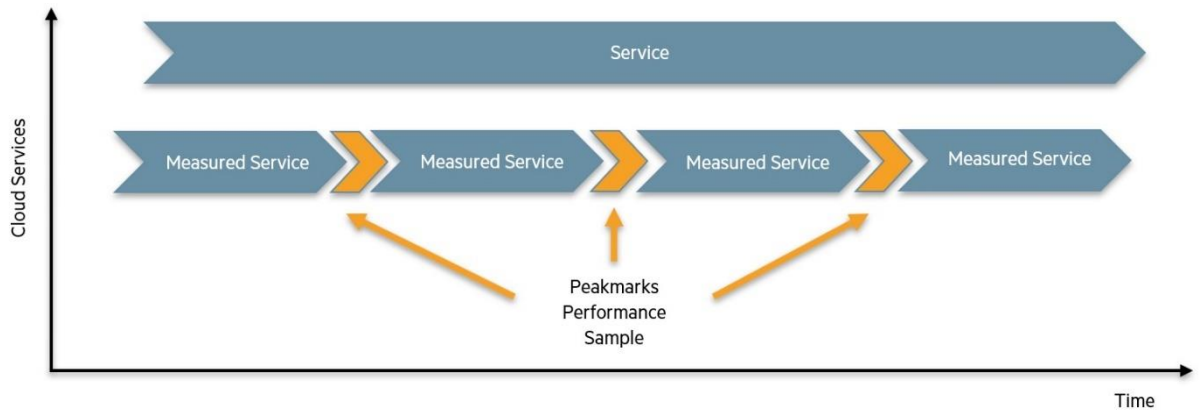
Für die Konfiguration eines einzelnen Performance Tests benötigt die peakmarks® Benchmark Software folgende Informationen:

- Auszuführender Workload.
- Für einige Workloads gibt es optionale Workload Parameter.
- Für einige Workloads kann die interne Oracle Parallelität (DOP) gewählt werden.
- Anzahl Clusterknoten, auf denen der Workload ausgeführt werden soll.
- Anzahl Prozesse, die den Workload ausführen. In einem Cluster werden die Prozesse *round-robin* über alle Clusterknoten gleichmässig verteilt.
- Angestrebte Laufzeit für den Performance Test.

Um einen Database Service vollständig zu vermessen, werden Workloads für unterschiedliche Lastsituationen konfiguriert. Dafür ist eine Sequenz von Performancetests notwendig, die auf verschiede Art und Weise konfiguriert werden kann.

Smart Benchmark Configuration. Die *Smart Benchmark Configuration* ist vollständig automatisiert. Sie ist das komfortabelste und schnellste Verfahren, um einen vollständigen Überblick zur Leistungsfähigkeit eines Oracle Database Services zu erhalten. Für jeden Workload wird eine Sequenz von Performance Tests (basierend auf dem Oracle Konfigurationsparameter CPU_COUNT, kann manuell übersteuert werden) automatisch generiert. Sollte bei zunehmender Last eine Sättigung des Systems erreicht werden, wird die Testsequenz nicht fortgesetzt.

Sample Benchmark Configuration. Die *Sample Benchmark Configuration* läuft ebenfalls vollständig automatisiert ab, führt aber nur eine ausgewählte Menge von vorkonfigurierten Performancetests durch. Diese Art von Benchmark Konfiguration wird gerne bei Cloud Services angewandt, um periodisch innerhalb eines Wartungsfensters von 60 Minuten zu überprüfen, ob der Database Service seine Performance Eigenschaften noch vollumfänglich beibehalten hat.



Manual Benchmark Configuration. Die *Manual Benchmark Configuration* bietet höchste Flexibilität bei der Konfiguration von Benchmark Tests. Jeder Parameter eines Performance Tests kann individuell gewählt werden. Dieses Verfahren wird bevorzugt von Engineers eingesetzt, um spezielle Lastsituationen im Detail zu analysieren.

Aufbau der peakmarks® Benchmark Reports

Masseinheiten und Abkürzungen

In den Benchmark Reports werden folgende Masseinheiten und Abkürzungen verwendet; die Spalten der jeweiligen Key Performance Metriken sind rot markiert.

Abkürzung	Bedeutung	Abkürzung	Bedeutung
[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
		[kBpt]	kiloByte per transaction
[IOPS]	I/O operations per second	[Mops]	million operations per seconds
BuCache	Oracle buffer cache	FlCache	Flash Cache (Database, Exadata)

Beispiele

Folgende Beispiele zeigen zwei tabellarische peakmarks® Benchmark Reports[5].

Das erste Beispiel zeigt das Leistungsverhalten des Systems bei Ausführung des Workloads SRV-QUERY25. In dem 2-Knoten Cluster verfügt jeder Server über 48 Cores und 96 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]
1	1	SRV-QUERY25	2	2	2	1	0	98	53,726	26,863	0.037	1,506,933	753,467	99.96	180
	2	SRV-QUERY25	2	48	25	25	0	75	902,585	18,804	0.053	25,235,088	525,731	100.00	182
	3	SRV-QUERY25	2	96	51	50	0	49	1,577,747	16,435	0.061	43,988,376	458,212	100.00	182
	4	SRV-QUERY25	2	144	75	73	1	25	1,802,631	12,518	0.079	50,170,859	348,409	100.00	182
	5	SRV-QUERY25	2	192	97	96	1	3	2,000,281	10,418	0.096	55,576,848	289,463	100.00	182

[5] Alle folgenden Benchmark Reports wurden auf dem aktuellen peakmarks® Referenzsystem ermittelt (Stand August 2023): ein Exadata Quarter Rack System mit zwei Exadata X7 Database Servern und drei Exadata X8 High Capacity Storage Servern in einer ASM «high redundancy» Konfiguration. Die technischen Details dieser Konfiguration können den entsprechenden Datenblättern entnommen werden.

Das zweite Beispiel zeigt das Leistungsverhalten des Systems bei Ausführung des Workloads STO-RANDOM (100% read).

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]
2	1	STO-RANDOM	0	2	2	1	2	1	1	98	0	67,955	0.214	544	544	100.00	-0.33	183
	2	STO-RANDOM	0	2	16	1	13	7	4	87	0	476,058	0.259	3,743	3,743	100.00	1.65	182
	3	STO-RANDOM	0	2	32	1	21	12	6	79	0	745,340	0.334	5,855	5,855	100.00	1.88	182
	4	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
	5	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
	6	STO-RANDOM	0	2	80	1	28	13	10	72	0	1,341,404	1.209	10,549	10,549	100.00	2.53	182
	7	STO-RANDOM	0	2	96	1	30	14	11	70	0	1,377,906	1.588	10,831	10,831	100.00	2.57	182
	8	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
	9	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
	10	STO-RANDOM	0	2	144	1	33	15	12	67	0	1,418,645	2.964	11,154	11,154	100.00	3.04	182
	11	STO-RANDOM	0	2	160	1	34	16	12	66	0	1,422,483	3.455	11,187	11,187	100.00	3.10	182
	12	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
	13	STO-RANDOM	0	2	192	1	35	17	12	65	0	1,428,829	4.482	11,236	11,236	100.00	3.23	182

Beschreibung der Performance Tests

Die erste Spalte eines jeden Benchmark Reports identifiziert den Benchmark Lauf (Spalte *Run*). Innerhalb eines Benchmark Laufs sind alle Performance Tests durchnummeriert (Spalte *Test*). Die dritte Spalte beschreibt den Workload. Optional gibt es je nach Workload zusätzliche Workload Parameter. Im zweiten Beispiel ist dies der prozentuale Anteil an Schreiboperationen (Spalte *Wri [%]*).

Die Spalte *Nodes* gibt an, wie viele Cluster Knoten in einem Test involviert sind. Die Spalte *Jobs* dokumentiert die Anzahl Prozesse auf allen beteiligten Cluster Knoten, die die Last erzeugen. Bei Tests in Clustern werden die Prozesse *round-robin* auf alle Cluster Knoten verteilt.

Bei einigen Workloads gibt es eine zusätzliche Spalte *DOP* (*degree of parallelism*), welche die Oracle interne Parallelisierung beschreibt.

Wahl der Messpunkte bei CPU-bound Performance Tests

Bei *CPU-bound* Workloads wie SRV-QUERY25 wählt die peakmarks® Benchmark Software automatisch fünf Messpunkte, um die Leistungsfähigkeit des Servers in unterschiedlichen Lastsituationen festzustellen: Single-Thread, 25% CPU-Auslastung, 50% CPU-Auslastung, 75% CPU-Auslastung, 100% CPU-Auslastung.

Wahl der Messpunkte bei anderen Workloads

Bei allen anderen Workloads wählt die peakmarks® Benchmark Software bis zu zweiunddreissig Messpunkte. Beginnend mit einem Prozess pro Server wird die Belastung automatisch gesteigert, bis eine Sättigung des Systems eintritt. Die Laststeigerung wird automatisch in Abhängigkeit vom Wert des Oracle Konfigurationsparameter CPU_COUNT ermittelt und kann manuell übersteuert werden.

Im zweiten Beispiel mit dem STO-RANDOM Workload wird die Messreihe nach dem dreizehnten Test abgebrochen, weil keine Leistungssteigerung mehr eintritt. Das System ist gesättigt.

CPU-Auslastung

Die folgenden vier Spalten beschreiben die CPU-Auslastung aller involvierten Server. Bei Workloads mit hohem I/O Anteil gibt es eine zusätzliche Spalte, welche den prozentualen CPU I/O Wait Anteil beschreibt (Spalte *CPU iow [%]*).

Performance Kennzahlen

Nun folgen Spalten mit den Performance Kennzahlen. Je nach Workload werden unterschiedliche Performance Kennzahlen aufgeführt. Im ersten Beispiel die Anzahl Abfragen (Spalte *Queries total [qps]*), die durchschnittliche Antwortzeit der Abfragen (Spalte *Response Time [ms]*) und die Anzahl logischer Datenbank Blockzugriffe im Datenbank Buffer Cache (Spalte *Log reads total [dbps]*).

Im zweiten Beispiel sind die beiden zentralen Performance Kennzahlen die Anzahl I/O Operationen (Spalte *Phys reads total [IOPS]*) und die I/O Servicezeit (Spalte *IO Time read [ms]*).

Bei den Performance Kennzahlen sind häufig zusätzliche Spalten zu finden, die die Leistung pro verwendete CPU anzeigen[6]. Diese Kennzahlen sind für Preis-/Leistungsvergleiche wichtig, um die Leistungsfähigkeit pro Core zu ermitteln. Oracle Lizenz- und Wartungskosten orientieren sich in der Regel an der Anzahl eingesetzter Sockets (Standard Edition) und Cores (Enterprise Edition).

Weitere Kennzahlen

Es folgen Spalten mit weiteren Kennzahlen. Diese Kennzahlen liefern häufig zusätzliche Informationen, um das Systemverhalten besser zu verstehen oder zur Kontrolle, ob der Workload auch optimal abgelaufen ist.

In unserem ersten Beispiel sind das die Oracle Buffer Cache Hitrate (Spalte *BuCache read [%]*) und die Verarbeitungszeit des Workloads (Spalte *Elapsed time [s]*). Im zweiten Beispiel findet man zusätzlich die Flash Cache Hitrate (Spalte *FlCache read [%]*).

Workloads für Server Systeme

peakmarks® bietet verschiedene Workloads für Server Systeme an. Die ausgewählten Server Workloads treten branchenübergreifend in allen Datenbank-Applikationen auf. Sie zeigen die *Real-World* Performance eines Server Systems im Oracle Datenbankbetrieb.

Alle Server Workloads greifen über SQL mit unterschiedlichen Zugriffsmustern auf Tabellen zu. Die betroffenen Tabellen liegen komplett im Buffer Cache. Es gibt nahezu keine I/O-Operationen, so dass die Performance dieser Workloads vollständig durch die CPU-Leistung limitiert ist.

Server Performance Tests geben Antworten auf Fragen nach der Skalierbarkeit von Servern, der Effizienz von Technologien wie Multithreading, Virtualisierung und Verschlüsselung von Daten. Für On-premises Plattformen ermöglichen sie einen einfachen Preis-/Leistungsvergleich von Prozessoren und Servern unter Berücksichtigung der Oracle Lizenzierung. Auch bei der Evaluation von Cloud Plattformen, bei

[6] Der Begriff CPU orientiert sich an der Oracle Definition im Sinne des Oracle Konfigurationsparameters CPU_COUNT. Bei Prozessoren ohne Multithreading entspricht eine CPU einem Core, bei Systemen mit Multithreading entspricht eine CPU einem Thread.

denen die Hardware-Komponenten nicht immer bekannt sind, liefern sie vergleichbare Kennzahlen zum Preis-/Leistungsverhältnis eines Cloud Anbieters.

Performance Kennzahlen für Server Systeme im Oracle Datenbankbetrieb

Zur Leistungsvermessung von Server Systemen im Oracle Datenbankbetrieb haben sich folgende Performance Kennzahlen bewährt:

- Durchsatz an SQL-Abfragen (*query throughput*) in *queries per second* [qps].
- Durchschnittliche Antwortzeit der SQL-Abfragen (*query response time*) in *milliseconds* [ms].
- Anzahl logischer Datenbank Blockzugriffe im Buffer Cache (*logical reads throughput*) in *database blocks per second* [dbps].
- SQL Buffer Cache Scan Rate in *MegaByte per second* [MBps].

Workloads zur Ermittlung der Server Performance im Oracle Datenbankbetrieb

Folgende Tabelle zeigt die verschiedenen Workloads, die zur Ermittlung der Server Performance im Oracle Datenbankbetrieb verwendet werden.

Workload	Aktion	Performance Kennzahl	Mass-einheit
SRV-QUERY1	Query Typ A: Selektiere 1 Datensatz via Index. Beispiel: Selektiere Konto, Produkt, Auftrag, Rechnung, etc. Dieser Workload zeigt den maximalen Query Durchsatz eines Servers und die minimale Query Antwortzeit an.	Query Durchsatz Query Antwortzeit	[qps] [ms]
SRV-QUERY25	Query Typ B: Selektiere durchschnittlich 25 Datensätze via Index. Beispiel: Selektiere Kontobuchungen der letzten Woche; Liste Artikel einer Bestellung etc.	Query Durchsatz Query Antwortzeit	[qps] [ms]
SRV-REPORT	Online Report: Selektiere durchschnittlich 125 Datensätze via Index. Dieser Workload entspricht einem einfachen Online Report und zeigt den maximalen Wert für *logical reads* eines Servers an. Beispiel: Online Report Mobilfunk Call Records vom letzten Monat, Online Report E-Banking mit Konto Mutationen vom letzten Monat, etc.	Blockzugriffe im Buffer Cache (logical reads)	[dbps]
SRV-SCAN	Datensuche ohne Index. Dies entspricht einem *full table scan*. Es wird nur ein Datensatz ermittelt, aber es muss die gesamte Datenmenge der Tabelle im Buffer Cache durchsucht werden.	Scan Rate im Buffer Cache	[MBps]

Workload	Aktion	Performance Kennzahl	Mass-einheit
SRV-MIXED	Komplexer Workload mit einem Mix von gleichgewichteten einfachen Workloads, die gleichzeitig ablaufen: ▪ SRV-QUERY1 ▪ SRV-QUERY25 ▪ SRV-REPORT ▪ SRV-SCAN SRV-MIXED zeigt die Server-Systemleistung im Oracle-Datenbankbetrieb bei einem repräsentativen Querschnitt von peakmarks® Server Workloads an.	Query Durchsatz Query Antwortzeit	[qps] [ms]

Beispiel: Workload SRV-QUERY1

Der Workload SRV-QUERY1 (1 Datensatz pro Abfrage) zeigt die maximal mögliche Anzahl Abfragen und die kürzeste Antwortzeit von Abfragen für ein Server System an.

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	1	SRV-QUERY1	2	2	2	1	0	98	233,345	116,673	0.009	933,557	466,779	99.95	180
	2	SRV-QUERY1	2	48	25	25	0	75	4,178,134	87,044	0.011	16,669,620	347,284	100.00	182
	3	SRV-QUERY1	2	96	51	50	1	49	6,520,208	67,919	0.015	25,944,413	270,254	100.00	182
	4	SRV-QUERY1	2	144	75	73	1	25	6,501,930	45,152	0.022	25,846,679	179,491	100.00	182
	5	SRV-QUERY1	2	192	98	96	1	2	6,795,993	35,396	0.028	26,987,811	140,562	100.00	182

Beispiel: Workload SRV-REPORT

Der Workload SRV-REPORT (Online-Report mit durchschnittlich 125 Datensätzen pro Abfrage) zeigt die maximale Anzahl der *logical reads* für einen Datenbank Server an.

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	11	SRV-REPORT	2	2	2	1	0	98	13,884	6,942	0.144	1,775,598	887,799	100.00	180
	12	SRV-REPORT	2	48	25	25	0	75	221,925	4,623	0.215	28,290,092	589,377	100.00	182
	13	SRV-REPORT	2	96	50	49	0	50	391,932	4,083	0.244	49,811,699	518,872	100.00	182
	14	SRV-REPORT	2	144	76	74	1	24	461,731	3,206	0.310	58,545,252	406,564	100.00	182
	15	SRV-REPORT	2	192	99	97	1	1	525,654	2,738	0.363	66,502,195	346,366	100.00	182

Beispiel: Workload SRV-SCAN

Der Workload SRV-SCAN (Abfrage mit einem *full table scan*, wobei 1 Treffer ermittelt wird) zeigt die Scan Rate im Buffer Cache an.

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]
3	21	SRV-SCAN	2	2	2	1	0	98	5,543	2,772	710,561	355,281	99.99	182
	22	SRV-SCAN	2	48	25	25	0	75	108,285	2,256	13,855,451	288,655	100.00	183
	23	SRV-SCAN	2	96	51	50	0	49	187,009	1,948	23,871,434	248,661	100.00	183
	24	SRV-SCAN	2	144	75	74	1	25	227,210	1,578	28,705,443	199,343	100.00	183
	25	SRV-SCAN	2	192	98	97	1	2	258,579	1,347	32,710,285	170,366	100.00	183

Beispiel: Workload SRV-MIXED

Der Workload SRV-MIXED wird gerne anstatt der anderen Workloads für Server Systeme verwendet, um schnell eine repräsentative Aussage zur Leistungsfähigkeit eines Server Systems im Oracle Datenbankbetrieb zu erhalten.

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	31	SRV-MIXED	2	8	5	4	0	95	272,844	34,106	0.029	4,290,051	536,256	100.00	182
	32	SRV-MIXED	2	48	25	25	0	75	1,331,390	27,737	0.036	21,115,772	439,912	100.00	183
	33	SRV-MIXED	2	96	51	50	0	49	2,238,224	23,315	0.042	36,500,369	380,212	100.00	183
	34	SRV-MIXED	2	144	75	74	1	25	2,442,773	16,964	0.058	42,067,495	292,135	100.00	182
	35	SRV-MIXED	2	192	98	97	1	2	2,736,163	14,251	0.070	47,252,138	246,105	100.00	182

Anmerkungen

Hauptspeicherzugriffe[7] ermöglichen ultimative Datenbank Performance. Beim Workload SRV-QUERY1 werden extrem kurze Antwortzeiten zwischen 9 und 28 µs ermittelt.

Irreführende CPU-Auslastung. Wenn Intel Xeon Prozessoren mit Multithreading betrieben werden, ist die Angabe zur CPU-Auslastung nicht immer aussagekräftig. Beim Workload SQL-QUERY1 werden im Test 3 alle Cores durch Prozesse belegt. Die CPU-Auslastung zeigt 51% an und es werden 6,520,208 Abfragen pro Sekunde verarbeitet. Beim Test 5 werden alle Threads durch Prozesse belegt, die CPU-Auslastung steigt auf 98% an und es werden 6,795,993 Abfragen pro Sekunde verarbeitet. Der Server ist nun vollständig ausgelastet. Zwischen den beiden Tests hat sich die CPU-Auslastung fast verdoppelt, die Durchsatz ist aber nur um weniger als 4% gestiegen. D.h. bereits im Test 3 betrug die tatsächliche CPU-Auslastung 96% und nicht 51%!

Viele Betriebssysteme werten bei der Berechnung der CPU-Auslastung Threads als Ressourcen, so wie Cores. Das sind sie aber nicht. Threads sind lediglich ein Konzept zur Verbesserung der parallelen Verarbeitung von Instruktionen.

Das Multithreading von Intel Xeon Prozessoren kann bei Oracle Workloads zu einer Durchsatz Verbesserung von bis zu 25% führen. Für Prozessorarchitekturen wie AMD EPYC, IBM POWER oder IBM z gelten andere Werte.

Skalierbarkeit bei zunehmender Belastung. Bei jedem der Server Workloads beobachten wir eine abnehmende Performance pro CPU bei zunehmender Belastung. So sinkt die Scan Rate pro CPU beim Workload SRV-SCAN von 2,772 MBps (Test 21, Single Thread Performance) auf 1,347 MBps (Test 25, Systemsättigung). Die Leistung pro CPU sinkt bei zunehmender Last auf ca. 50% des Spitzenwertes.

Wir beobachten dieses Verhalten vor allem bei Intel Xeon Prozessoren mit einer grossen Spreizung der Taktrate. In unserem Fall (Intel Xeon 8160) beträgt die maximale Taktrate 3.7 GHz. Bei zunehmender Belastung wird die Taktrate aus thermischen Gründen auf 2.1 GHz reduziert. Dies bedeutet eine Reduktion der Leistungsabgabe von 43%.

In der Regel weisen Intel Xeon Prozessoren mit hoher Anzahl Cores auch eine hohe Spreizung der Taktrate auf. Und umgekehrt gilt, dass Intel Xeon Prozessoren mit geringerer Anzahl Cores auch eine geringere Spreizung der Taktrate haben und damit eine stabilere Leistungsabgabe garantieren können.

[7] Zur Orientierung: die schnellsten Prozessoren (Stand August 2023) benötigen weniger als 0.1 Nanosekunden für einen DRAM Zugriff. Die Datenbank benötigt ca. 1 Mikrosekunde für einen Blockzugriff im Oracle Buffer Cache und zwischen 20 und 500 Mikrosekunden (je nach Storage Fabric und Protokoll) für einen Blockzugriff auf einem Flash Storage.

Server Evaluation für den Oracle Datenbankbetrieb. Um eine berechenbare (*predictable*) und dauerhafte (*persistent*) CPU-Performance in allen Lastbereichen zu erreichen, empfiehlt sich der Einsatz von Intel Xeon Prozessoren mit geringerer Anzahl Cores. Solche Prozessoren liefern auch die höchste Pro-Core Performance und verbessern daher das Preis-/Leistungsverhältnis eines Servers unter Berücksichtigung der Oracle Lizenz- und Wartungskosten.

Um Skalierbarkeitsanforderungen zu erreichen sind dann eventuell mehr Sockets notwendig. Dies führt nebenbei zu der Fähigkeit deutlich höhere Hauptspeicherkapazitäten zu verwenden, was wiederum zu einer weiteren Performanceverbesserung führen kann und dem Einsatz von *In-Memory* Technologien zugutekommt.

Workloads für Storage Systeme

Herkömmliche I/O-Benchmark Tools wie fio, vdbench, iometer, Orion zeigen oft Leistungswerte an, die im realen Datenbankbetrieb nicht erreicht werden. Der Grund dafür liegt in der Komplexität von Datenbank I/O-Operationen, die von diesen Tools nicht berücksichtigt wird.

Wird ein Datenblock gelesen, muss die Buffer Cache Verwaltung der Datenbank viele Aufgaben erfüllen:

- einen freien Platz für den Block finden.
- wenn kein freier Platz vorhanden ist, ältere Blöcke ersetzen.
- alle Datenbankprozesse synchronisieren, die gleichzeitig versuchen, freie Plätze im Buffer-Cache einzunehmen.
- bei Verwendung einer Shared-Disk Cluster-Architektur (Oracle Real Application Cluster) muss die Synchronisation clusterweit erfolgen, was im besten Fall den Austausch von Informationen (Messages) zwischen Cluster Knoten erfordert, im schlechtesten Fall zusätzlich den Austausch von Datenbankblöcken.
- und schliesslich werden Blöcke während der I/O-Übertragung auf ihre Integrität und Konsistenz überprüft. Oracle Konfigurationsparameter legen den Umfang der Integritäts- und Konsistenztests fest. Diese Tests können zu einer Mehrbelastung von Prozessoren und I/O System führen.

Die peakmarks® Benchmark Software erzeugt I/O-Last mit sogenannten SQL-generierten I/O-Operationen, um repräsentative Performance Kennzahlen für das Storage System im Oracle Datenbankbetrieb zu ermitteln. Bei sequenziellen Operationen kann zusätzlich die Intra-SQL Parallelität getestet werden, um die optimale SQL-Parallelität zu finden (Oracle DOP).

Storage Performance Tests geben Antworten auf Fragen der Effizienz von Protokollen wie NVMe *over fabrics* (IP, FC), RoCE (*RDMA over Converged Ethernet*) und Technologien wie Flash Storage, Persistant Memory, Storage Tiering, Datenverschlüsselung, Offload Funktionen intelligenter Storage Systeme und der Datenreduktion durch De-Duplikation und Komprimierung. Sie liefern fundierte Kennzahlen für Preis-/Leistungsvergleiche bei Storage Systemen für Oracle Datenbanken und bei der Evaluation geeigneter Storage Replikationsverfahren.

Performance Kennzahlen für Storage Systeme im Oracle Datenbankbetrieb

Zur Leistungsvermessung von Storage Systemen im Datenbankbetrieb werden folgende Performance Kennzahlen ermittelt:

- Durchsatz beim sequenziellen Lesen von Storage Systemen (*sequential read*) in *megabyte per second* [MBps]. Dabei werden typischerweise grosse I/O Einheiten von 1 MByte verwendet.
- Wahlfreier Zugriff von Storage Systemen (*random I/O*) in *I/O operations per second* [IOPS]. Dabei werden typischerweise kleinere I/O Einheiten wie Oracle Datenbankblöcke verwendet. Die Grösse von Oracle Datenbank Blöcken ist konfigurierbar: 2 KByte, 4 KByte, 8 KByte, 16 KByte, 32 KByte). Der Default-Wert beträgt 8 KByte.
- I/O Servicezeit beim wahlfreien Zugriff in *milliseconds* [ms].

Workloads zur Ermittlung der Storage Performance im Datenbankbetrieb

Folgende Tabelle zeigt die zur Verfügung stehenden Workloads der peakmarks® Benchmark Software zur Vermessung der Leistungsfähigkeit von Storage Systemen im Oracle Datenbankbetrieb.

Workload	Aktion	Performance Kennzahl	Mass-einheit
STO-READ	Durch SQL-Statements erzeugte sequenzielle Lese-operation.	Durchsatz *sequential* I/O	[MBps]
STO-OFFLOAD	Durch SQL-Statements erzeugte sequenzielle Lese-operation bei Verwendung intelligenter Storage Server mit offload Funktionalität (*smart scan*).	Durchsatz *sequential* I/O	[MBps]
STO-RANDOM	Durch SQL-Statements erzeugte wahlfreie Lese- und Schreiboperationen. Bei diesem Workload werden Lese-Operationen durch Vordergrundprozesse und Schreib-Operationen durch Hintergrundprozesse durchgeführt. Das Verhältnis von Datenbankblöcken, die gelesen bzw. geschrieben werden, ist konfigurierbar[8].	Durchsatz *random* I/O I/O Servicezeit	[IOPS] [ms]
STO-SCATTER	Durch SQL-Statements erzeugte wahlfreie Schreiboperationen. Bei diesem Workload werden Schreib-Operationen durch Vordergrundprozesse durchgeführt.	Durchsatz *random* I/O	[dbps]

[8] Oracle führt Schreiboperationen häufig als multiblock I/O Operation durch; d.h. mit einer I/O Operationen wird eine Liste von Blöcken geschrieben.

Beispiel: Workload STO-READ

Der Workload STO-READ zeigt den I/O Durchsatz zwischen Datenbank Server und Storage System bei sequenziellen Leseoperationen an. In diesem Beispiel wird eine Oracle DOP von 4 verwendet. Die Datenmenge wird aus dem Flash Cache geladen (Spalte *FlCache read*).

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	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
	7	STO-READ	0	2	48	4	3	2	1	97	0	12,108	15.116	12,061	12,061	100.00	0.00	194

Beispiel: Workload STO-OFFLOAD

Der Workload STO-OFFLOAD ist der gleiche wie STO-READ, nutzt aber die Intelligenz der Storage Systeme und verwendet die *Smart Scan* Technologie. Beim Einsatz der Smart-Scan Technologie kann auf die Angabe eines DOP verzichtet werden, da die Storage Server den Zugriff selbstständig optimieren.

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	11	STO-OFFLOAD	0	2	2	1	2	1	1	98	0	45,224	0.478	44,931	44,931	100.00	0.00	171
	12	STO-OFFLOAD	0	2	8	1	1	1	1	99	0	64,711	0.668	64,303	64,303	100.00	0.00	172
	13	STO-OFFLOAD	0	2	16	1	2	1	1	98	0	69,557	1.185	69,122	69,122	100.00	0.00	173
	14	STO-OFFLOAD	0	2	24	1	2	1	1	98	0	72,848	1.749	72,394	72,394	100.00	0.00	174
	15	STO-OFFLOAD	0	2	32	1	1	1	1	99	0	72,396	2.372	71,943	71,943	100.00	0.00	175

Beispiel: Workload STO-RANDOM

Der Workload STO-RANDOM zeigt den I/O Durchsatz zwischen Datenbank Server und Storage System bei wahlfreiem Zugriff (*random*) auf einzelne Datenbankblöcke an. In diesem Beispiel beträgt die Grösse der Oracle Datenbankblöcke 8 KByte.

Im ersten Benchmark Report (Test 21 bis 27) werden nur lesende Operationen vermessen. Beim zweiten Benchmark Report (Test 31 – 37) beträgt das Verhältnis von Lese- zu Schreiboperationen 80:20 (Spalte *Wri [%]*).

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	21	STO-RANDOM	0	2	2	1	2	1	1	98	0	67,955	0.214	544	544	100.00	-0.33	183
	22	STO-RANDOM	0	2	32	1	21	12	6	79	0	745,340	0.334	5,855	5,855	100.00	1.88	182
	23	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
	24	STO-RANDOM	0	2	96	1	30	14	11	70	0	1,377,906	1.588	10,831	10,831	100.00	2.57	182
	25	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
	26	STO-RANDOM	0	2	160	1	34	16	12	66	0	1,422,483	3.455	11,187	11,187	100.00	3.10	182
	27	STO-RANDOM	0	2	192	1	35	17	12	65	0	1,428,829	4.482	11,236	11,236	100.00	3.23	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]
4	31	STO-RANDOM	20	2	2	1	2	1	1	98	0	66,545	0.227	524	524	100.00	17.99	182
	32	STO-RANDOM	20	2	16	1	10	5	3	90	0	423,127	0.336	3,319	3,319	100.00	15.93	183
	33	STO-RANDOM	20	2	32	1	17	9	5	83	0	662,583	0.534	5,208	5,208	100.00	16.78	183
	34	STO-RANDOM	20	2	48	1	21	11	6	79	0	781,877	0.817	6,153	6,153	100.00	17.56	183
	35	STO-RANDOM	20	2	64	1	24	13	7	76	0	845,871	1.105	6,651	6,651	100.00	18.18	184
	36	STO-RANDOM	20	2	80	1	26	14	7	74	0	883,718	1.471	6,951	6,951	100.00	18.69	184
	37	STO-RANDOM	20	2	96	1	27	15	8	73	0	904,607	1.862	7,116	7,116	100.00	18.91	183

Anmerkungen

I/O-Bandbreite. Beim Workload STO-READ wird die übertragene Datenmenge in der Regel durch die I/O Bandbreite zwischen Server System und Storage System begrenzt.

I/O-Servicezeit. Beim Workload STO-RANDOM wird die I/O Servicezeit durch die Effizienz des I/O Stacks beeinflusst. Bei Einsatz von All-Flash-Arrays werden Servicezeiten von unter 500 Mikrosekunden erwartet. Durch neuartige Protokolle wie RoCE (*RDMA over converged Ethernet*) sind mittlerweile I/O Servicezeiten von unter 20 Mikrosekunden möglich, falls Daten im Cache des Storage Systems gefunden werden.

Log Writer Workloads

Log Writer Prozesse sind in erster Linie für die Transaktionsprotokollierung und die Wiederherstellung der Datenbank nach Systemausfällen verantwortlich. Diese Prozesse sind entscheidend für die Gesamtleistung von Oracle bei der Verarbeitung von Transaktionen. Die Latenzzeit der Transaktionsprotokollierung kann einen erheblichen Einfluss auf die Antwortzeit von Benutzertransaktionen haben.

Optional werden die Log Writer Prozesse auch für die Datenbankreplikation verwendet, um Standby-Datenbanken zu synchronisieren. Diese Technologie ist sehr beliebt bei Desaster Recovery-Lösungen. Die Replikation kann im synchronen oder asynchronen Modus erfolgen. Der Datentransfer zwischen Primär- und Standby-Datenbank kann optional verschlüsselt und/oder komprimiert werden. Bei der synchronen Replikation müssen lokale Transaktionen warten, bis auch die Standby-Datenbanken das Transaktionsprotokoll erhalten haben. Dies kann die lokale Transaktionsverarbeitung erheblich verzögern[9].

Performance Kennzahlen für Log Writer Prozesse

Die Performance Kennzahlen für Log Writer Prozesse sind:

- Anzahl SQL Commit Operationen in *transactions per second* [tps].
- Durchschnittliche Latenzzeit bei SQL Commit Operationen in *milliseconds* [ms].
- REDO-Durchsatz in *megabyte per second* [MBps].

[9] Beim Einsatz von Oracle Data Guard als BCP-Lösung wird bei *mission-critical* Systemen eine synchrone Replikation verwendet. Dies bedeutet für die Log Writer eine zusätzliche Latenz von mindestens 1 ms.

Workloads für Log Writer Prozesse

Um das Leistungsverhalten der Log Writer Prozesse für verschiedene Transaktionsgrössen zu analysieren, bietet die peakmarks® Benchmark Software unterschiedliche Workloads an.

Workload	Aktion	Performance Kennzahl	Mass-einheit
LGWR-LAT1 LGWT-LAT25 LGWR-LAT125	Transaktionen mit ca. 1, 25 oder 125 KByte REDO-Datenmenge pro Transaktion. Alle Transaktionen schliessen mit einem *COMMIT WAIT IMMEDIATE* ab.	SQL Commit Operationen SQL Commit Latenzzeit	[tps] [ms]
LGWR-THR	Transaktionen, die eine grosse REDO-Datenmenge erzeugten. Alle Transaktionen schliessen mit einem *COMMIT NOWAIT BATCH* ab.	Log Writer Durchsatz	[MBps]

Beispiel: Workload LGWR-LAT1

Der Workload LGWR-LAT1 zeigt die maximale Anzahl Commit-Operationen an, die das System verarbeiten kann. Die durchschnittliche Transaktionsgrösse, gemessen an der REDO-Datenmenge, liegt bei knapp 1.4 kByte (Spalte *REDO data [kBpt]*).

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	2	2	2	1	1	98	0	18	6,165	0.323	0.234	19,487	6,258	8	1.396	99.98	182
	2	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
	3	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
	4	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
	5	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
	6	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
	7	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
	8	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
	9	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
	10	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
	11	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

Beispiel: Workload LGWR-THR

Der Workload LGWR-THR zeigt die maximale REDO-Datenmenge an, die von den Log Writer Prozessen verarbeitet werden kann. Die Anzahl Log Writer Prozesse ist nicht konfigurierbar und wird von Oracle automatisch ermittelt.

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	12	LGWR-THR	2	2	2	1	1	98	0	18	488	4.092	1.348	270,752	468	128	268.396	100.00	181
	13	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
	14	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
	15	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
	16	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
	17	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
	18	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
	19	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

Anmerkungen

Technologien zur Datenreduzierung auf Speichersystemen. Einige Cloud Service Provider verwenden Storage Systeme mit Verfahren zur Datenreduktion (Deduplizierung, Komprimierung). Diese Technologien reduzieren die benötigte Storage Kapazität, können sich aber nachteilig auf die Latenzzeit von Log Writer Prozessen und den Transaktionsdurchsatz von OLTP-Systemen auswirken.

Database Writer Workload

Die Database Writer sind für die Verwaltung des Buffer Caches verantwortlich. Diese Prozesse sind entscheidend für die Gesamtleistung von Oracle, wenn viele Blöcke geändert werden, wie z.B. beim konventionellen Laden von Daten über den Buffer Cache oder bei intensiver Transaktionsverarbeitung.

Die Anzahl der Database Writer Prozesse wird von Oracle automatisch ermittelt. Sollte dieser Wert nicht ausreichend sein, kann er manuell angepasst werden.

Performance Kennzahlen für Database Writer Prozesse

Die Performance Kennzahl für Database Writer Prozesse ist

- Durchsatz an zurückgeschriebenen Datenbankblöcken in *database blocks per second* [dbps].

Workloads zur Ermittlung der Database Writer Performance

Zur Ermittlung der Database Writer Performance steht der Workload DBWR-THR zur Verfügung.

Workload	Aktion	Performance Kennzahl	Mass-einheit
DBWR-THR	Massive Blockveränderungen im Buffer Cache	Database Writer Durchsatz	[MBps]

Beispiel: Workload DBWR-THR

Der Workload DBWR-THR zeigt die maximale Anzahl Blöcke an, die von den Database Writer Prozessen zurückgeschrieben werden können.

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	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

Anmerkungen

ASM Redundanz. Bei der Konfiguration von ASM können verschiedene Redundanz Level definiert werden (*external*, *normal redundancy*, *high redundancy*). Dies hat erhebliche Auswirkungen auf die Performance schreibender Prozesse.

In unserem Beispiel wird der auf Exadata Systemen übliche Level «high redundancy» verwendet. D.h. dass die insgesamt dreifache Datenmenge geschrieben werden muss, z.B. im Test 4 insgesamt 3 x 328,586 Datenbankblöcke!

Data Load Workloads

Systemarchitekten und Kapazitätsplaner benötigen Performance Kennzahlen von Database Services bezüglich ihrer Fähigkeit, Daten zu laden. Dies ist besonders wichtig für Data Warehouse und Data Analytics Systeme, bei denen die Datenmengen immer grösser und die Zeitfenster für das Laden immer kleiner werden.

Oracle stellt verschiedene Technologien zum Laden von Daten zur Verfügung:

- konventionelles Laden über den Buffer-Cache; diese Technologie wird bevorzugt in *Transaction Processing Systems* verwendet.
- direktes Laden unter Umgehung des Buffer-Cache; dieses Verfahren wird vorwiegend im *Data Warehouse* und *Data Analytics* Umfeld eingesetzt, um grosse Datenmengen möglichst schnell zu laden.
- ab Oracle 19c steht ein neuartiges Ladeverfahren für IOT-Anwendungen (*Internet of Things*) zur Verfügung. Massendaten können schnell geladen werden, wobei aber die Transaktionssicherheit nicht immer garantiert wird. Wenn Aggregate wie Durchschnittswerte, etc. nicht auf jeden Datensatz angewiesen sind, ist dieses Ladeverfahren akzeptabel.

Kennzahlen für Data Load Performance

Für Data Load Workloads gibt es nur eine Performance Kennzahl:

- Durchsatz an geladenen Daten (*data load rate*) in *megabyte per second* [MBps] oder *terabyte per hour* ([TBph].

Workloads zur Ermittlung der Data Load Performance

Die peakmarks® Benchmark Software bietet Workloads für alle drei Ladetechniken an.

Workload	Aktion	Performance Kennzahl	Mass-einheit
DL-BUFFER	Die Daten werden programmgeneriert und konventionell via Buffer-Cache geladen. Es werden 3 zusätzliche Indizes beim Laden aktualisiert. Typischerweise werden nur wenige Datensätze pro Transaktion geladen. Der Ladeprozess verwendet beim Transaktionsabschluss *COMMIT WRITE WAIT IMMEDIATE*.	Data Load Durchsatz	[MBps]

Workload	Aktion	Performance Kennzahl	Mass-einheit
DL-DIRECT	Die Daten werden aus einer Datenquelle direkt unter Umgehung des Buffer Caches geladen. Typischerweise werden grosse Datenmengen pro Transaktion geladen. Es wird nur ein zusätzlicher Index beim Laden aktualisiert. Der Ladeprozess verwendet die *NOLOGGING* Option und beim Transaktionsabschluss *COMMIT WRITE NOWAIT BATCH*.	Data Load Durchsatz	[MBps]
DL-STREAM	Die Daten werden programmgeneriert und verwenden den *memory-optimized row store for write* für IOT-Anwendungen. Es wird nur ein zusätzlicher Index beim Laden aktualisiert. Der Ladeprozess verwendet kein Commit, sondern lediglich eine Flush-Operation des Memory-Optimized Pool.	Data Load Durchsatz	[MBps]

Beispiel: Workload DL-BUFFER

Der Workload DL-BUFFER zeigt die maximale Lade Rate an, die beim *buffered data load* erreicht wird. Dieses Verfahren zum Laden wird bevorzugt in transaktionsorientierten Systemen verwendet.

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	2	2	2	1	1	98	0	25,151	7	5,796	46	24	100.00	182
	2	DL-BUFFER	5	2	32	7	5	1	93	0	204,245	58	14,866	398	193	100.00	183
	3	DL-BUFFER	5	2	64	11	8	2	89	0	303,210	87	18,297	617	275	100.00	182
	4	DL-BUFFER	5	2	96	13	10	2	87	0	377,273	108	19,309	828	347	100.00	182
	5	DL-BUFFER	5	2	128	17	13	2	83	0	420,796	120	23,709	988	392	100.00	182
	6	DL-BUFFER	5	2	160	20	16	3	80	0	452,938	130	28,397	1,123	416	100.00	182
	7	DL-BUFFER	5	2	192	23	18	3	77	0	476,194	136	32,492	1,248	439	100.00	184

Beispiel: Workload DL-DIRECT

Der Workload DL-DIRECT zeigt die maximale Lade Rate an, die beim *direct data load* unter Umgehung des Buffer Caches erreicht wird. Dieses Verfahren zum Laden wird bevorzugt in Data Warehouse Systemen verwendet.

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	14	DL-DIRECT	125,000	2	2	2	1	0	98	0	940,670	269	1,917	411	62	100.00	181
	15	DL-DIRECT	125,000	2	16	8	7	1	92	0	5,540,540	1,585	12,476	2,437	362	100.00	182
	16	DL-DIRECT	125,000	2	32	11	9	1	89	0	6,851,879	1,960	18,829	3,021	448	100.00	182
	17	DL-DIRECT	125,000	2	48	13	11	2	87	0	6,923,907	1,981	28,546	3,085	454	100.00	182
	18	DL-DIRECT	125,000	2	64	13	10	2	87	0	6,742,719	1,929	36,474	3,025	443	100.00	182

Data Analytics Workloads

Systemarchitekten und Kapazitätsplaner benötigen Performance Kennzahlen von Database Services bezüglich ihrer Fähigkeit, grosse Datenmengen zu durchsuchen. Diese Anwendungen basieren typischerweise auf *full table scan* Operationen und werden durch keine Indexstrukturen unterstützt.

Die Performance von *full table scans* hängt von der Position der Daten in der Speicherhierarchie (Storage, Hauptspeicher) und der Technologie zur Steigerung der Scanleistung ab (*Smart Scan*, *In-Memory Column Store*).

Die *Smart Scan* Technologie steht nur auf Oracle Engineered Systems zur Verfügung und ist zusätzlich lizenzpflichtig. Die *In-Memory* Technologie steht auf allen Plattformen zur Verfügung und ist ab Oracle 21c bis zu einer gewissen Kapazität lizenzfrei.

Data Analytics Workloads geben Antworten auf Fragen der Effizienz beider Technologien und ermöglichen Preis-/Leistungsvergleiche.

Kennzahlen für Data Analytics Performance

Für Data Analytics Workloads gibt es nur eine Performance Kennzahl:

- Durchsatz an durchsuchten Daten (*data scan rate*) in *megabyte per second* [MBps] oder *rows per second* [rps]

Workloads zur Ermittlung der Data Load Performance

peakmarks® bietet Workloads zum Testen verschiedener Datenstandorte (Storage, Hauptspeicher) und zum Testen verschiedener Boost-Technologien (*Smart Scan*, *In-Memory Column Store*) an. Auch hier kann die Intra-SQL Parallelität verändert werden, um die optimale Datenbank Parallelität (Oracle DOP) beim Data Scan zu finden.

Workload	Aktion	Performance Kennzahl	Mass-einheit
DA-STORAGE	*Full table scan* mit Aggregat geringer Komplexität. Die Daten werden vom Storage System gelesen.	Data Scan Durchsatz	[MBps]
DA-OFFLOAD	*Full table scan* mit Aggregat geringer Komplexität. Die Daten werden vom intelligenten Storage System mit Offload Funktionalität (*smart scan*) gelesen.	Data Scan Durchsatz	[MBps]
DA-ROWSTORE	*Full table scan* mit Aggregat geringer Komplexität. Die Daten werden vom *Row Store* im Buffer Cache gelesen.	Data Scan Durchsatz	[MBps]
DA-COLSTORE	*Full table scan* mit Aggregat geringer Komplexität. Die Daten werden vom *Column Store* im Buffer Cache verarbeitet.	Data Scan Durchsatz	[rps]

Beispiel: Workload DA-STORAGE

Der Workload DA-STORAGE zeigt die maximale Scan Rate an, wenn alle Daten vom Storage gelesen werden müssen. Die Verarbeitung des Data Scans (Filteroperationen) findet auf dem Database Server statt.

Die Metrik *Scanned user data [MBps]* entspricht der tatsächlich durchsuchten Datenmenge. Die Metrik *Scanned user data [rps]* hängt vom Datenmodell ab und gilt für eine Satzlänge von 300 Byte (Satzlänge der synthetischen Tabellen der peakmarks® Benchmark Software).

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	2	2	4	2	1	1	98	0	16,450,340	5,601	100.00	0.00	171
	2	DA-STORAGE	2	16	4	3	2	1	97	0	34,464,955	11,734	100.00	0.00	172
	3	DA-STORAGE	2	32	4	4	2	1	96	0	35,703,125	12,155	100.00	0.00	172
	4	DA-STORAGE	2	48	4	4	2	1	96	0	36,270,748	12,348	100.00	0.00	172
	5	DA-STORAGE	2	64	4	4	3	1	96	0	36,527,567	12,436	100.00	0.00	173
	6	DA-STORAGE	2	80	4	4	3	1	96	0	36,640,231	12,474	100.00	0.00	174
	7	DA-STORAGE	2	96	4	4	3	1	96	0	36,765,045	12,516	100.00	0.00	174
	8	DA-STORAGE	2	112	4	4	3	1	96	0	36,773,204	12,519	100.00	0.00	174

Beispiel: Workload DA-OFFLOAD

Der Workload DA-OFFLOAD zeigt die maximale Scan Rate an, wenn alle Daten vom intelligenten Storage System verarbeitet werden (Filteroperationen). Lediglich die Treffermenge wird zum Database Server übertragen.

Diese Offload-Technologie ist nur auf Oracle Engineered Systems verfügbar und benötigt zusätzliche Lizenzen für deren Storage Server.

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	11	DA-OFFLOAD	2	2	1	1	1	0	99	0	58,173,231	19,819	100.00	0.00	171
	12	DA-OFFLOAD	2	16	1	2	1	1	98	0	164,581,594	56,070	100.00	0.00	171
	13	DA-OFFLOAD	2	32	1	2	1	1	98	0	187,076,316	63,733	100.00	0.00	171
	14	DA-OFFLOAD	2	48	1	2	1	1	98	0	195,758,609	66,691	100.00	0.00	171
	15	DA-OFFLOAD	2	64	1	2	1	1	98	0	196,795,769	67,044	100.00	0.00	172
	16	DA-OFFLOAD	2	80	1	2	1	1	98	0	199,305,704	67,899	100.00	0.00	172
	17	DA-OFFLOAD	2	96	1	2	1	1	98	0	201,142,122	68,525	100.00	0.00	172
	18	DA-OFFLOAD	2	104	1	3	1	1	97	0	201,183,310	68,539	100.00	0.00	172
	19	DA-OFFLOAD	2	112	1	2	1	1	98	0	200,329,561	68,248	100.00	0.00	172

Beispiel: Workload DA-ROWSTORE

Der Workload DA-ROWSTORE zeigt die maximale Scan Rate an, wenn alle Daten im Buffer Cache des Database Server verarbeitet werden.

Diese Form der Verarbeitung steht auch für die Oracle Standard Edition zur Verfügung. Eine grosse Hauptspeicherkapazität wirkt sich vorteilhaft aus, um möglichst viele Daten im Buffer Cache speichern zu können.

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	21	DA-ROWSTORE	2	2	1	2	1	0	98	0	15,223,603	4,600	0.00	99.74	181
	22	DA-ROWSTORE	2	48	1	25	25	0	75	0	362,635,020	109,529	0.00	100.00	183
	23	DA-ROWSTORE	2	96	1	50	49	0	50	0	617,217,180	186,421	0.00	100.00	183
	24	DA-ROWSTORE	2	144	1	75	73	1	25	0	754,465,316	227,875	0.00	100.00	183
	25	DA-ROWSTORE	2	192	1	96	94	1	4	0	845,157,690	255,266	0.00	100.00	183

Beispiel: Workload DA-COLSTORE

Der Workload DA-COLSTORE zeigt die maximale Scan Rate an, wenn alle Daten im *Column Store* des Database Server verarbeitet werden. Der Einsatz dieser *In-Memory* Technologie setzt entsprechende Lizenzen auf dem Datenbank Server voraus.

Es müssen nicht ganze Tabellen im *Column Store* abgelegt werden, sondern nur einzelne ausgewählte Spalten, die häufig als Suchkriterium verwendet werden. Durch Komprimierungsverfahren kann die im *Column Store* zu speichernde Datenmenge zusätzlich erhöht werden. Auch hier wirkt sich eine grosse Hauptspeicherkapazität vorteilhaft aus, um möglichst viele Daten im *Column Store* zu speichern.

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	31	DA-COLSTORE	2	2	1	2	1	0	98	0	10,861,614,976	3,287,494	0.00	99.99	182
	32	DA-COLSTORE	2	48	1	26	25	0	74	0	141,771,840,591	42,916,756	0.00	100.00	182
	33	DA-COLSTORE	2	96	1	52	50	1	48	0	116,215,766,166	35,198,349	0.00	100.00	181
	34	DA-COLSTORE	2	144	1	75	72	1	25	0	125,054,889,772	37,885,745	0.00	100.00	181
	35	DA-COLSTORE	2	192	1	96	93	1	4	0	131,717,969,993	39,916,054	0.00	100.00	182

Die Metrik *Scanned user data [rps]* entspricht der tatsächlich durchsuchten Datenmenge. Die Metrik *Scanned user data [MBps]* hängt vom Datenmodell ab und gilt für eine Satzlänge von 300 Byte (Satzlänge der synthetischen Tabellen der peakmarks® Benchmark Software).

Anmerkungen

Effizienz der verschiedenen Verfahren. Es ist unschwer zu erkennen, dass die *In-Memory* Technologie mit dem *Column Store* die höchste Performance aufweist und auch die *Smart Scan* Technologie weit hinter sich lässt. Die *In-Memory* Technologie hat sich daher in den letzten Jahren zur bevorzugten Technologie für Data Analytics Anwendungen entwickelt.

Preis-/Leistungsvergleich der verschiedenen Verfahren. Sowohl für die *Smart Scan* Technologie als auch die *In-Memory* Technologie sind zusätzliche Lizenzen notwendig. Es fällt auf, dass die *In-Memory* Technologie ein deutlich besseres Preis-/Leistungsverhältnis als die *Smart Scan* Technologie hat, wenn man die Scan Raten im Verhältnis zu den Lizenzkosten betrachtet.

Die *Smart Scan* Technologie arbeitet allerdings weitgehend transparent für Applikationen, es sind nur selten Änderungen am Code notwendig. Die *Smart Scan* Technologie kann mit der *In-Memory* Technologie auf den Storage Servern kombiniert werden und kann immense Datenkapazitäten verwalten. Die *Smart Scan* Technologie kann allerdings nur auf Oracle Engineered Systems genutzt werden.

Die *In-Memory* Technologie steht hingegen auf allen Plattformen zur Verfügung, benötigt keine spezielle Hardware und liefert ultimative Performance. Allerdings ist die Datenmenge durch die Hauptspeicherkapazität (DRAM) der Datenbank Server begrenzt. Auch müssen Eingriffe in die Applikation vorgenommen werden, um zu spezifizieren welche Spalten einer Tabelle in den *Column Store* aufgenommen werden sollen. Komprimierungsverfahren helfen dabei, die Kapazität des *Column Store* optimal zu nutzen.

Transaction Processing Workloads

Systemarchitekten und Kapazitätsplaner benötigen Performance Kennzahlen von Database Services bezüglich ihrer Fähigkeit typische Transaktionen ausführen zu können.

Transaction Processing ist eine äusserst komplexe Form der Datenverarbeitung. Alle Plattform Komponenten und Datenbank Serviceprozesse sind massgeblich involviert und müssen genau austariert werden. Transaction Processing ist die Königsdisziplin von Datenbank-Plattformen.

Die Transaction Processing Performance hängt von sehr vielen Faktoren ab. Die wichtigsten sind:

- Verhältnis von Datenbankgrösse und Buffer Cache Grösse, dass sich in der Buffer Cache Hitrate ausdrückt.
- I/O Performance beim Lesen von Daten.
- I/O Performance beim Schreiben des Transaktionsprotokolls durch REDO Log Writer Prozesse.
- I/O Performance beim Schreiben von Daten durch Database Writer Prozesse.

Die I/O Performance beim Lesen von Daten kann wiederum beeinflusst werden durch Storage Tiering und vielen anderen Faktoren. Die I/O Performance der Log Writer Prozesse kann durch die Data Guard Performance oder andere Technologien zur Spiegelung von Daten zwischen verschiedenen Data Center beeinflusst werden.

Kennzahlen für Transaction Processing Performance

Für Transaction Processing Workloads sind folgende Performance Kennzahlen von Interesse:

- Durchsatz von SQL-Transaktionen (*transaction throughput*) in *transactions per second* [tps].
- Durchschnittliche Antwortzeit der SQL-Transaktionen (*transaction response time*) in *milliseconds* [ms].

Übersicht der Transaction Processing Performance Workloads

Die peakmarks® Benchmark Software stellt verschiedene Arten von Workloads zur Verfügung:

- Lesende Workloads TP-LOOKUP und TP-REPORT.
- Mutierende Workloads mit unterschiedlicher Komplexität TP-LIGHT, TP-MEDIUM und TP-HEAVY. Bei diesen Workloads kann das Verhältnis zwischen lesenden und ändernden Operationen konfiguriert werden.
- Mix von Workloads mit SELECT-, UPDATE- und INSERT-Statements, wobei mehrere einfache TP-Workloads gleichzeitig und gleichgewichtet ablaufen und um einen DL-BUFFER Workload ergänzt werden.

  TP-MIXED1 ist ein lese-intensiver Transaktionsmix, TP-MIXED2 ein schreib-intensiver Workload.

  Diese beiden Workloads werden bevorzugt verwendet, um schnell eine repräsentative Aussage zur Transaction Processing Performance einer Oracle Datenbank Plattform zu ermitteln.

Workload	Aktion	Performance Kennzahl	Mass-einheit
TP-LOOKUP	Schnelle Look-up Query, welche die Existenz eines einzelnen Datensatzes prüft. Beispiel: Prüfe Kreditkarte, Auto Zulassung, Telefonnummer, etc. Falls auf der Plattform verfügbar, nutzt dieser Workload *memory-optimized tables* für einen schnellen Zugriff via hash key. Andernfalls werden konventionelle Tabellen mit einem *unique b-tree index* verwendet.	Durchsatz von Queries Antwortzeit von Queries	[tps] [ms]
TP-REPORT	Online Report, der durchschnittlich 25 Datensätze pro Query selektiert.	Durchsatz von Queries Antwortzeit von Queries	[tps] [ms]
TP-LIGHT	Selektiere/Mutiere einzelnen Datensatz via Index. Beispiel: Selektiere/Mutiere Konto, Produkt, Auftrag, Rechnung, etc. Transaktionen werden immer mit einem *COMMIT WRITE WAIT IMMEDIATE* abgeschlossen.	Durchsatz von Transaktionen Antwortzeit von Transaktionen	[tps] [ms]
TP-MEDIUM	Selektiere/Mutiere Ø 25 Datensätze via Index. Beispiel: Selektiere/Mutiere Kontobuchungen letzte Woche; Artikelliste einer Bestellung etc. Transaktionen werden immer mit einem *COMMIT WRITE WAIT IMMEDIATE* abgeschlossen.	Durchsatz von Transaktionen Antwortzeit von Transaktionen	[tps] [ms]
TP-HEAVY	Selektiere/Mutiere Ø 125 Datensätze via Index. Transaktionen werden immer mit einem *COMMIT WRITE WAIT IMMEDIATE* abgeschlossen. Beispiel: Selektiere/Mutiere Mobilfunk Call Records vom letzten Monat, etc.	Durchsatz von Transaktionen Antwortzeit von Transaktionen	[tps] [ms]
TP-MIXED1	Gleichgewichteter lese-intensiver Workload Mix: ▪ TP-REPORT ▪ TP-LOOKUP ▪ TP-MEDIUM (40% UPDATE) ▪ DL-BUFFER (2 rpt)	Durchsatz von Transaktionen Antwortzeit von Transaktionen	[tps] [ms]
TP-MIXED2	Gleichgewichteter schreib-intensiver Workload Mix: ▪ TP-LIGHT (40% UPDATE) ▪ TP-MEDIUM (30% UPDATE) ▪ TP-HEAVY (20% UPDATE) ▪ DL-BUFFER (3 rpt).	Durchsatz von Transaktionen Antwortzeit von Transaktionen	[tps] [ms]

Die Benchmark Reports für Transaction Processing Workloads umfassen sowohl die primären Leistungskennzahlen wie Durchsatz und Antwortzeit von Transaktionen als auch wichtige Einflussfaktoren: Buffer Cache Hitrate (Spalte *BuCache read [%]*), Flash Cache Hitrate - falls vorhanden (Spalte *FlCache read [%]*),

die Servicezeit beim Random Read (Spalte *IO time read [ms]*) und die Latenzzeit vom Log Writer beim Transaktionsabschluss (Spalte *Log File syn [ms]*).

Beispiel: Workload TP-MIXED1

Lese-intensiver Workload Mix TP-MIXED1.

Run	Test	Workload	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]
9	1	TP-MIXED1	2	8	4	2	1	96	0	13,979	0.570	0.240	0.228	2.152	70.11	100.00	181
	2	TP-MIXED1	2	24	14	9	3	86	0	310,522	0.077	0.271	0.412	0.200	91.53	100.00	183
	3	TP-MIXED1	2	40	19	13	4	81	0	685,432	0.058	0.298	0.542	0.127	93.91	100.00	182
	4	TP-MIXED1	2	56	26	17	5	74	0	943,014	0.059	0.341	0.721	0.105	94.38	100.00	182
	5	TP-MIXED1	2	72	30	21	6	70	0	1,179,671	0.061	0.378	0.877	0.094	94.61	100.00	182
	6	TP-MIXED1	2	88	35	24	7	65	0	1,403,307	0.062	0.412	1.001	0.087	94.91	100.00	182
	7	TP-MIXED1	2	104	39	27	7	61	0	1,601,678	0.065	0.446	1.113	0.080	95.14	100.00	182
	8	TP-MIXED1	2	120	43	31	8	57	0	1,791,183	0.067	0.485	1.264	0.076	95.33	100.00	182
	9	TP-MIXED1	2	136	47	34	9	53	0	1,919,583	0.070	0.518	1.387	0.073	95.36	100.00	182
	10	TP-MIXED1	2	152	51	37	9	49	0	2,013,175	0.075	0.546	1.502	0.072	95.32	100.00	182
	11	TP-MIXED1	2	168	55	40	10	45	0	2,108,923	0.079	0.573	1.616	0.072	95.32	100.00	182
	12	TP-MIXED1	2	184	58	43	10	42	0	2,171,213	0.084	0.605	1.739	0.072	95.28	100.00	182
	13	TP-MIXED1	2	200	61	45	10	39	0	2,262,549	0.087	0.624	1.829	0.071	95.27	100.00	183
	14	TP-MIXED1	2	208	63	47	11	37	0	2,281,498	0.090	0.641	1.881	0.071	95.28	100.00	184
	15	TP-MIXED1	2	216	64	48	11	36	0	2,335,079	0.091	0.653	1.940	0.070	95.30	100.00	185

Beispiel: Workload TP-MIXED2

Schreib-intensiver Workload Mix TP-MIXED2.

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]
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
	4	TP-MIXED2	N/A	2	24	13	7	4	87	0	23,605	1.011	0.293	0.473	2.749	62.96	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
	6	TP-MIXED2	N/A	2	40	18	10	5	82	0	30,199	1.318	0.353	0.699	2.893	60.82	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
	8	TP-MIXED2	N/A	2	56	21	12	6	79	0	35,255	1.578	0.411	0.922	2.912	59.27	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
	10	TP-MIXED2	N/A	2	72	24	13	7	76	0	39,469	1.814	0.468	1.145	2.923	58.60	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
	12	TP-MIXED2	N/A	2	88	27	15	8	73	0	43,347	2.017	0.521	1.304	2.924	57.94	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
	14	TP-MIXED2	N/A	2	104	29	16	8	71	0	48,129	2.148	0.572	1.479	2.849	58.52	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
	16	TP-MIXED2	N/A	2	120	31	18	9	69	0	49,413	2.413	0.636	1.693	2.894	57.91	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

Ein zusätzlicher Report liefert ergänzende Informationen zum I/O Verhalten.

Run	Test	Workload	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	CPU iow [%]	Transactions total [tps]	Logical reads [dbpt]	Logical writes [dbpt]	Physical reads [dbpt]	Physical writes [dbpt]	Phys reads total [IOPS]	Phys reads total [MBps]	Phys writes total [IOPS]	Phys writes total [MBps]	Elapsed time [s]
18	2	TP-MIXED2	2	8	6	3	2	94	0	10,444	21.204	11.376	7.267	2.561	75,680	616	30,931	242	183
	3	TP-MIXED2	2	16	10	5	3	90	0	17,037	23.482	11.774	8.541	2.918	145,085	1,181	53,268	443	182
	4	TP-MIXED2	2	24	13	7	4	87	0	23,605	22.884	11.230	8.476	2.874	199,417	1,627	69,387	600	182
	5	TP-MIXED2	2	32	16	9	4	84	0	26,667	25.522	12.018	9.648	3.060	256,338	2,080	80,782	722	182
	6	TP-MIXED2	2	40	18	10	5	82	0	30,199	25.502	11.891	9.993	3.049	300,517	2,450	89,758	813	182
	7	TP-MIXED2	2	48	20	11	6	80	0	32,956	26.087	11.935	10.505	3.026	344,581	2,804	96,018	882	182
	8	TP-MIXED2	2	56	21	12	6	79	0	35,255	26.570	12.018	10.823	3.001	379,739	3,090	101,000	937	182
	9	TP-MIXED2	2	64	23	13	7	77	0	37,798	26.592	11.902	11.001	2.925	413,558	3,361	104,637	981	182
	10	TP-MIXED2	2	72	24	13	7	76	0	39,469	27.422	12.118	11.352	2.926	445,688	3,617	108,548	1,026	182
	11	TP-MIXED2	2	80	26	14	7	74	0	41,402	27.513	12.188	11.482	2.872	472,848	3,831	111,366	1,060	182
	12	TP-MIXED2	2	88	27	15	8	73	0	43,347	27.535	12.134	11.582	2.783	499,383	4,047	112,232	1,078	182
	13	TP-MIXED2	2	96	28	16	8	72	0	45,473	27.787	12.121	11.557	2.735	522,949	4,238	115,367	1,113	182
	14	TP-MIXED2	2	104	29	16	8	71	0	48,129	27.249	11.882	11.304	2.625	541,540	4,378	116,818	1,134	182
	15	TP-MIXED2	2	112	30	17	8	70	0	49,139	27.756	11.967	11.531	2.630	564,340	4,559	119,305	1,159	182
	16	TP-MIXED2	2	120	31	18	9	69	0	49,413	28.158	12.108	11.852	2.658	583,253	4,710	120,994	1,178	182
	17	TP-MIXED2	2	128	32	18	9	68	0	51,392	27.674	11.934	11.676	2.606	597,772	4,831	123,293	1,202	182

Performance von PL/SQL-Anwendungsprogrammen

PL/SQL ist die bevorzugte Programmiersprache für komplexe Transaktionslogik und Algorithmen. PL/SQL Programme werden kompiliert im Datenbankserver gespeichert. Einige grosse Anwendungen, wie z.B. *Core Banking Systems*, sind weitgehend in PL/SQL implementiert.

Kennzahlen für PL/SQL Performance

Für PL/SQL Workloads sind folgende Performance Kennzahlen von Interesse:

- Anzahl der ausgeführten PL/SQL-Operationen pro Zeiteinheit in *million operations per second* [Mops].
- Ausführungszeit von PL/SQL Algorithmen in *seconds* [s].

Workloads zur Ermittlung der PL/SQL Performance

peakmarks® bietet Workloads an, um die Effizienz von PL/SQL Programmen auf einem bestimmten Prozessor zu testen. Diese Workloads sind vollständig durch die CPU-Leistung limitiert.

Workload	Aktion	Performance Kennzahl	Mass-einheit
PLS-ADD	Addition von Zahlen	Durchsatz PL/SQL Operationen	[Mops]
PLS-BUILTIN	Datentyp-spezifische Operationen einschliesslich SQL-Funktionen, basierend auf Core Banking Systemen und Telco Billing Systemen	Durchsatz PL/SQL Operationen	[Mops]
PLS-PRIME	Berechnung der ersten n Primzahlen	Verarbeitungszeit eines Algorithmus	[s]
PLS-FIBO	Berechnung der ersten n Fibonacci Zahlen (rekursiver Algorithmus)	Verarbeitungszeit eines Algorithmus	[s]
PLS-MIXED	Komplexer Workload mit einem Mix von gleichgewichteten einfachen Workloads, die gleichzeitig ablaufen: ▪ PLS_ADD mit Datentyp NUMBER ▪ PLS_ADD mit Datentyp PLS_INTEGER ▪ PLS_BUILTIN mit Datentyp NUMBER ▪ PLS_BUILTIN mit Datentyp VARCHAR2	Durchsatz PL/SQL Operationen	[Mops]

Beispiel: Workload PLS-MIXED

Der Workload PLS-MIXED zeigt den Durchsatz von unterschiedlichen PL/SQL Operationen mit verschiedenen Datentypen an.

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

Beispiel: Workload PLS-FIBO

Der Workload PLS-FIBO zeigt die Verarbeitungszeit für einen Algorithmus an, der Fibonacci-Zahlen (Spalte *N*) berechnet. Der verwendete Datentyp kann konfiguriert werden (Spalte *Data type*).

Run	Test	Workload	Parameter	Nodes	Jobs	CPU busy [%]	CPU user [%]	CPU sys [%]	CPU idle [%]	Operations total [Mops]	Operations per cpu [Mops]	Elapsed time [s]
19	31	PLS-FIBO	42	2	1	1	1	0	99	0.00	0.00	69
	32	PLS-PRIME	8000	2	1	1	1	0	99	0.00	0.00	79

Fazit

peakmarks® Benchmark Software bietet Systemingenieuren und Systemarchitekten ein robustes und umfassendes Benchmark Framework zur Bestimmung aussagekräftiger und verständlicher Performance Kennzahlen von Oracle Database Services – **On-Premise** und in der **Cloud**.

Die leicht verständlichen Performance Kennzahlen von peakmarks® machen Technologien, Konfigurationen, Komponenten und komplette Systeme hersteller- und anbieterübergreifend vergleichbar.

Die peakmarks® Benchmark Software deckt alle erdenklichen Workloads für einen Performancetest der Infrastruktur ab. Die Software wird ständig weiterentwickelt und an neue Datenbankversionen angepasst.

Vorteile für den IT-Betrieb. Viele IT-Organisationen können die Kosten für Database Services senken, indem sie sich für *best-in-class* Infrastrukturkomponenten oder *best-in-class* Cloud Services entscheiden und die Lizenz- und Wartungskosten minimieren.

Engpässe, Fehlkonfigurationen und Fehlfunktionen werden vor der Inbetriebnahme erkannt. Regelmässige Performance Analysen von Cloud Services stellen die Leistungsqualität sicher. peakmarks® Performance Kennzahlen sind eine solide Grundlage für die Kapazitätsplanung bei der Migration auf andere Plattformen oder Cloud Services.

Das Ergebnis sind Oracle Database Services mit einer berechenbaren (*predictable*) und dauerhaften (*persistent*) Performance in allen Lastsituationen.

Vorteile für IT-Hardware Anbieter und Service Provider. peakmarks® Benchmarks sind schnell und effizient und unterstützen den Prozess bei der Entwicklung und Vermarktung lizenzoptimierter Oracle Datenbank Lösungen mit bestem Preis-/Leistungs-Verhältnis.

Kunden und Interessenten profitieren von den verständlichen, benutzerfreundlichen und vergleichbaren Performance Kennzahlen, die zeitaufwändige und teure Proofs-of-Concepts reduzieren können.

peakmarks® Performance Kennzahlen tragen dazu bei, die Positionierung eines Anbieters im Vergleich zum Wettbewerb zu stärken.

Anhang: Performance Zertifizierung für Reference System

Oracle Database Service

Platform	Database Server, each	Storage Server High Capacity, each	Database	peakmarks® configuration
▪ Exadata Flex ▪ 2 database servers Exadata X7 ▪ 3 storage server Exadata X8 high capacity	▪ Intel Xeon 8160 ▪ 2 sockets, 2.1 – 3.7 GHz ▪ 48 cores ▪ 96 threads ▪ 768 GByte DDR4-2666, PCI Gen 3	▪ 12 x 14 = 168 TByte HDD (raw) ▪ 4 x 6.4 = 25.6 TByte flash cache ▪ ASM high redundancy	▪ Oracle 19.20 EE ▪ db block size 8 kByte ▪ RAC ▪ data guard off ▪ archiving off	▪ Version 10.1 ▪ Build 230801 ▪ database size 2 x 2 TByte

peakmarks® Key Performance Metrics for platform components in Oracle database operations

Category	Key Performance Metric	Response time	Throughput	peakmarks® Workload
Server System	Mix of queries and scans	0.029 – 0.070 ms	2,736,163 qps	SRV-MIXED
	Buffer cache logical read throughput	-	66,502,195 dbps	SRV-REPORT
	Buffer cache scan rate	-	258,579 MBps	SRV-SCAN

Category	Key Performance Metric	Throughput	peakmarks® Workload
Storage System	SQL sequential I/O	12,132 MBps	STO-READ
	SQL sequential I/O - using smart scan	72,394 MBps	STO-OFFLOAD
	SQL random read, 100% read, < 500 µs	745,340 IOPS	STO-RANDOM
	SQL random read, 80% read, < 500 µs	423,127 IOPS	STO-RANDOM
	SQL random write	360,793 dbps	STO-SCATTER

peakmarks® Key Performance Metrics for critical database background processes

Category	Key Performance Metric	Latency	Throughput	peakmarks® Workload
Log Writer (LGWR)	Small-sized transactions - commit throughput and latency	0.323 – 1.443 ms	209,735 tps	LGWR-LAT1
	Throughput of REDO data		1,443 MBps	LGWR-THR
Database Writer (DBWR)	Buffer cache writes		328,586 dbps	DBWR-THR

peakmarks® Key Performance Metrics for representative database operations

Category	Key Performance Metric	Throughput	peakmarks® Workload
Data Load	Transactional data load - using the buffer cache (5 rpt)	136 MBps 476,194 rps	DL-BUFFER
	Data Warehouse data load - bypassing the buffer cache	1,981 MBps 6,923,907 rps	DL-DIRECT

Category	Key Performance Metric	Throughput	peakmarks® Workload
Data Analytics	Data scan - using storage system	12,519 MBps 36,773,204 rps	DA-STORAGE
	Data scan - using smart scan	68,539 MBps 201,183,310 rps	DA-OFFLOAD
	Data scan - using row store	255,266 MBps 845,157,690 rps	DA-ROWSTORE
	Data scan - using column store	39,916,054 MBps 131,717,969,993 rps	DA-COLSTORE

Category	Key Performance Metric	Response time	Throughput	peakmarks® Workload
Transaction Processing	Read-intensive transaction mix	0.058 – 0.091 ms	2,335,079 tps	TP-MIXED1
	Write-intensive transaction mix	0.762 – 2.478 ms	51,392 tps	TP-MIXED2

peakmarks® Key Performance Metrics for PL/SQL application code

Category	Key Performance Metric	Throughput	Execution Time	peakmarks® Workload
PL/SQL	PL/SQL operations throughput	6,937 Mops	-	PLS-MIXED
	PL/SQL algorithm execution time Fibonacci NUMBER for n = 42		69 s	PLS-FIBO