



Benchmarking Database Cloud Services

Mai 2024

contact@peakmarks.com

Vorwort

«Benchmarking Database Cloud Services» wurde erstmals anlässlich der 11th TPC Technology Conference 2019 in Los Angeles publiziert¹. Die darin vorgestellte peakmarks® Software wurde seitdem weiterentwickelt und an neue Technologien angepasst. Sie steht nun in der Version 10 für alle Oracle Datenbank Plattformen ab Oracle Version 19c mit der CDB-Architektur zur Verfügung.

Diese Publikation ist eine aktualisierte Version der ursprünglichen TPC-Veröffentlichung.

Einleitung



Selbst für die bekannten Cloud Marktführer² gibt es keine vergleichbaren und nachvollziehbaren Leistungskennzahlen für den Oracle Datenbankbetrieb.

Fragen zum Preis-Leistungsverhältnis, zum Umfang der Oracle Lizenzen bei Verwendung des BYOL-Modells³, zur Kapazitätsplanung beim Wechsel von On-Premises Installationen in die Cloud, zur Stabilität und Skalierbarkeit bei maximaler Belastung eines Cloud Services und vieles mehr können nicht ohne Weiteres beantwortet werden. Cloud-Service-Anbieter ändern ihre Infrastrukturkomponenten und -konfigurationen ohne Vorankündigung. Wie überprüft man als Kunde, dass man immer noch die gleiche Leistung erhält?

Proof-of-Concepts erscheinen nicht immer als das richtige Werkzeug, denn sie sind aufwendig, langwierig und teuer.

Peakmarks hat eine Benchmark Software entwickelt, um schnell und zuverlässig einfach verständliche Kennzahlen zur Leistungsfähigkeit einer Oracle Plattform – sei es On-Premises oder in der Cloud – zu ermitteln.

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*) ist ein Lizenzmodell, bei dem Kunden bestehende On-Premises Lizenzen in der Cloud nutzen können.

Die peakmarks® Software

Die peakmarks® 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. Performance Kennzahlen ermöglichen faktenbasierte Entscheidungen bei der Auswahl von Cloud Services, Komponenten, Konfigurationen und Optionen.

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 in vielen Fällen bei weitem die Infrastrukturkosten. Durch Optimierung der Infrastruktur können Lizenzkosten deutlich reduziert werden. Die am besten geeigneten Infrastruktur-Komponenten werden mit Hilfe von Performance Kennzahlen schnell 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, zuverlässig und nachvollziehbar (Oracle AWR-Reports).

Aber auch bei On-Premises Installationen können die Performance Eigenschaften nach Veränderungen der Infrastruktur (Austausch von Komponenten, Änderung der Konfiguration, etc.) verifiziert werden.

Anforderungen an Benchmark Tools

Huppler⁴ beschreibt die fünf wichtigsten Eigenschaften für einen guten Benchmark: relevant, wiederholbar, fair, nachvollziehbar und wirtschaftlich. peakmarks® erfüllt diese Anforderungen.

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

Einfachheit. Es muss einfach sein, die Software zu installieren, den Benchmark durchzuführen und die Ergebnisse zu interpretieren.

peakmarks® ist mit den Werkzeugen der Datenbank implementiert. Daher läuft peakmarks® unverändert überall dort, wo die Oracle Datenbanksoftware verfügbar ist. Ohne zusätzliche Software, ohne zusätzliche Hardware, ohne zusätzliches Scripting. Jeder Datenbank-Administrator kann die Benchmark Software 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 synthetischen Benchmark-Datenbank hängt von der Datenbankgrösse und der Performance der Infrastruktur ab. Der skalierbare Ladevorgang kann an die Leistungsfähigkeit der Datenbankplattform angepasst werden und erreicht auf leistungsfähigen Systemen bis zu 5 TByte pro Stunde. Komplette Benchmark-Läufe mit allen Workloads dauern zwischen 12 und 24 Stunden; die Ergebnisse sind sofort verfügbar. Ein umfassendes Performance Assessment kann innerhalb einer Woche abgeschlossen werden.

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.

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 interpretieren⁵.

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® Software kontinuierlich weiterentwickelt.

Die Anforderungen durch neue Hardware-Technologien wie Flash Storage und Persistent 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 hingewiesen⁶.

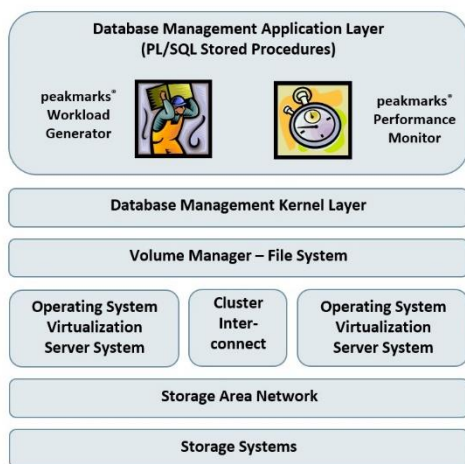
⁵ Crolotte: *Issues in Benchmark Metric Selection*. In: Nambiar, R., Poess, M. (eds.) *Performance Evaluation and Benchmarking 2009*, LNCS, vol. 5895, pp. 146–152. Springer, Heidelberg 2009.

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

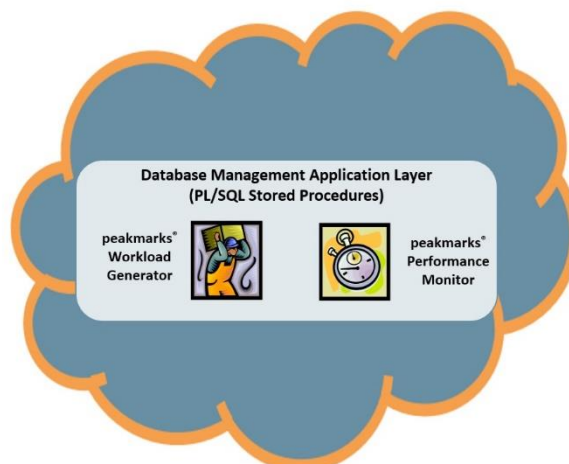
Die Architektur der peakmarks® Software

Die peakmarks® 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. Ansonsten sind keinerlei zusätzliche Hardware- oder Software-Installationen und auch kein Scripting notwendig.

Die peakmarks® 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.



peakmarks® Software On-Premises



peakmarks® Software in der Cloud

Ein Performance Monitor sammelt vor und nach jedem einzelnen Performance Test alle relevanten Statistiken und zeigt unmittelbar das Ergebnis in leicht lesbarer tabellarischer Form an. Für detaillierte Analysen werden Oracle AWR-Reports erzeugt. Die Performance wird an der Schnittstelle zur Applikation gemessen, d.h. die gemessene Performance steht der Applikation direkt zur Verfügung.

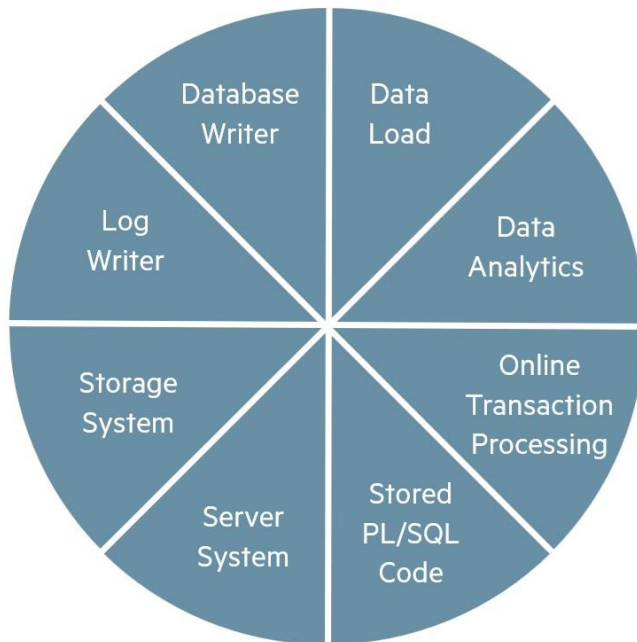
Über verschiedene peakmarks® Konfigurationsparameter kann die synthetische Benchmark Datenbank an Kundenbedürfnisse angepasst werden:

- Anzahl CPUs.
- Grösse der Datenbank zwischen 64 GByte und 64 TByte pro Datenbank Instanz.
- Grösse des Datenbank Buffer Caches.
- Optionaler Einsatz von Database Flash Cache für konventionelle Datenbank Server oder Cell Flash Cache für Oracle Engineered Systems.
- Die Laufzeit von Performance Tests kann zwischen 1 Minute und 12 Stunden festgelegt werden.

Oracle Funktionen wie Database Encryption, Data Guard oder Real Application Cluster mit mehreren Datenbank Servern werden unterstützt.

Vollständige 360-Grad Performance Übersicht

Mit über 30 Workloads, sogenannte Micro-Benchmarks, in 8 Workload Gruppen ermöglicht die peakmarks® 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® Software untersucht mit eigens entwickelten Workloads das Leistungsverhalten dieser beiden wichtigen Datenbank Service-Prozesse.

sucht 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 Online Transaction Processing Workloads ermitteln den Transaktionsdurchsatz und das Antwortzeitverhalten für Transaktionen unterschiedlicher Komplexität. Sofern konfiguriert, werden auch hier leistungssteigernde Technologien wie z.B. memory-optimized tables für eine schnelle Verarbeitung von Look-Up Abfragen eingesetzt. Einige Applikationen kapseln zentral wichtige Business-Funktionen und Transaktionen in Stored Procedures. Die peakmarks® Software bietet Workloads zur Analyse des Leistungsverhalten von PL/SQL Applikationsprogrammen auf unterschiedlichen Prozessoren an.

Alle wesentlichen Performance Kennzahlen werden für System Engineers in einem übersichtlichen und leicht zu lesenden peakmarks® Performance Report aufbereitet und in einem peakmarks® Executive Summary für Entscheidungsträger zusammengefasst⁷. Für eine detaillierte Analyse aller Performance Resultate werden automatisch Oracle AWR-Reports erzeugt.

⁷ Beispiele für solche Reports sind auf www.peakmarks.com zu finden.

Konfiguration von peakmarks® Performance Tests

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

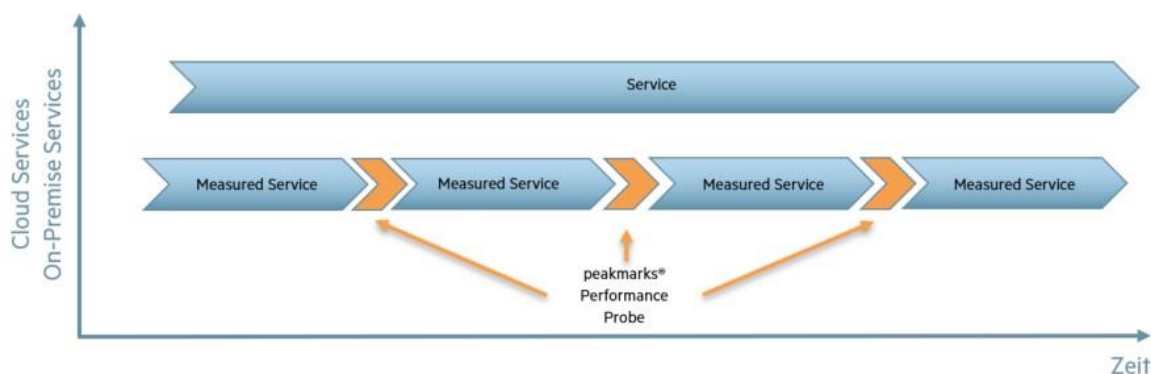
- Auszuführender Workload.
- Für einige Workloads gibt es optionale Workload Parameter.
- Für einige Workloads kann die interne Oracle Parallelität (DOP) gewählt werden.
- Anzahl Clusterknoten, auf denen der Workload ausgeführt werden soll.
- Anzahl Prozesse, die den Workload ausführen. In einem Cluster 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 verschiedene Art und Weise konfiguriert werden kann.

Smart Benchmark Configuration. Die Smart Benchmark Configuration ist vollumfänglich 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) automatisch generiert. Sollte bei zunehmender Last eine Sättigung des Systems erreicht werden, wird die Testsequenz abgebrochen.

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.

Sample Benchmark Configuration. Die Sample Benchmark Configuration läuft ebenfalls vollständig automatisiert ab, führt aber nur ausgewählte und vorkonfigurierte Performancetests durch. Diese Art der 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.



Aufbau der peakmarks® Reports

Masseinheiten und Abkürzungen

In den peakmarks® 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
[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)

Beispiele

Das erste Beispiel zeigt das Leistungsverhalten eines Systems bei Ausführung des Server Workloads SRV-QUERY25. In dem 2-Knoten Cluster verfügt jeder Server über 192 Cores und 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

Das zweite Beispiel zeigt das Leistungsverhalten des Systems bei Ausführung des Storage Workloads STO-RANDOM (80% read, 20% write, Datenbank Blockgrösse 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	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

Beschreibung der Performance Reports

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 Workload Prozesse auf allen beteiligten Cluster Knoten. 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® 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® 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 fünften 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 anzeigen⁸. Diese Kennzahlen sind für Preis-/Leistungsvergleiche wichtig, um die Leistungsfähigkeit pro Thread bzw. 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 FICache 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 Server-Leistung limitiert ist.

Server Performance Tests geben Antworten auf Fragen nach der Skalierbarkeit von Servern, dem Einfluss von NUMA-Effekten, 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 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].

⁸ Der Begriff CPU orientiert sich an der Definition des Oracle Konfigurationsparameters CPU_COUNT. Bei Prozessoren ohne Multithreading entspricht eine CPU einem Core, bei Systemen mit Multithreading entspricht eine CPU einem Thread. In beiden Fällen entspricht die Leistungsangabe der maximalen Performance, die ein einzelner Anwendungsprozess beziehen kann. Nur im Falle der Oracle internen Parallelisierung kann ein Applikationsprozess die Performance mehrerer Threads nutzen.

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 Metrik	Mass-einheit
SRV-QUERY1	Query Typ A: Selektiere 1 Datensatz via unique index. Beispiele: 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 non-unique index. Beispiele: 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 non-unique index. Beispiele: Online Report Mobilfunk Call Records vom letzten Monat, Online Report E-Banking mit Konto Mutationen vom letzten Monat, etc. Dieser Workload entspricht einem Online Report und zeigt den maximalen Wert für logical reads eines Servers an.	Blockzugriffe im Buffer Cache	[dbps]
SRV-SCAN	Datensuche ohne Index. Dies entspricht einem full table scan. Es wird die gesamte Datenmenge der Tabelle im Buffer Cache sequenziell durchsucht.	Scan Rate Buffer Cache	[MBps]
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]
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

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]
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

Beispiel: Workload SRV-SCAN

Der Workload SRV-SCAN (Abfrage mit einem full table scan) 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]
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

Beispiel: Workload SRV-MIXED

Der Workload SRV-MIXED wird gerne verwendet, um schnell eine repräsentative Aussage zur Leistungsfähigkeit eines Server Systems im Oracle Datenbankbetrieb zu erhalten, z.B. zum Vergleich gängiger Prozessorarchitekturen wie ARM, AMD EPYC, IBM POWER, IBM Z oder Intel Xeon. Performance Kennzahlen nach SPEC⁹ korrelieren nicht immer mit der Performance eines Prozessors im Oracle Datenbankbetrieb.

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 Siehe www.spec.org.

Anmerkungen

Hauptspeicherzugriffe¹⁰ ermöglichen ultimative Datenbank Performance. Beim Workload SRV-QUERY1 werden extrem kurze Antwortzeiten zwischen 7 und 25 µs ermittelt.

Irreführende CPU-Auslastung. Werden Prozessoren mit Multithreading betrieben, ist die Angabe zur CPU-Auslastung nicht immer zutreffend. Beim Workload SRV-QUERY1 werden im Run 4, Test 9 alle Cores durch Prozesse belegt. Die CPU-Auslastung zeigt 50% an und es werden 23,143,866 Abfragen pro Sekunde verarbeitet. Beim Test 11 werden alle Threads durch Prozesse belegt, die CPU-Auslastung steigt auf 91% an und es werden 30,068,291 Abfragen pro Sekunde verarbeitet. Der Server ist nun nahezu vollständig ausgelastet. Zwischen den beiden Tests ist die CPU-Auslastung um 82% gestiegen, der Durchsatz aber nur um 30%. D.h. im Test 9 betrug die CPU-Auslastung bereits 77%, nicht 50%!

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

Skalierbarkeit bei zunehmender Belastung. Bei allen Server Workloads beobachten wir eine abnehmende Performance pro CPU bei zunehmender Auslastung. Der Durchsatz pro CPU sinkt beim Workload SRV-MIXED von 43,935 qps (Run 5, Test 2, Single Thread Performance) auf 17,025 qps (Run 5, Test 6, Systemsättigung). Die Leistung pro CPU sinkt bei zunehmender Last auf ca. 39% des Spitzenwertes!

Wir beobachten dieses Verhalten bei Prozessoren mit einer grossen Spreizung der Taktrate. In diesem Fall beträgt die maximale Taktrate 3.7 GHz. Bei zunehmender Belastung wird die Taktrate aus thermischen Gründen auf 2.6 GHz reduziert. Dies allein bedeutet bereits eine Reduktion der Leistung von 30%.

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

Server Evaluation für den Oracle Datenbankbetrieb. Bei der Server Evaluation sind einige Aspekte zu beachten. Um eine berechenbare (predictable) und dauerhafte (persistent) CPU-Performance in allen Lastbereichen zu erreichen, sind Prozessoren mit geringerer Anzahl Cores vorteilhaft. Solche Prozessoren liefern 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 eventuell mehr Sockets notwendig. Dies ermöglicht deutlich höhere Hauptspeicherkapazitäten, was wiederum zu einer Performanceverbesserung führen kann und dem Einsatz von In-Memory Technologien zugutekommt. Allerdings sind bei solchen Architekturen im Oracle Datenbankbetrieb NUMA-Effekte zu erwarten.

Einfluss der Taktrate auf die Leistung. Die Taktrate von Prozessoren ist für den Leistungsvergleich verschiedener Architekturen irrelevant, da je nach Architektur während einem Zyklus unterschiedliche Aufgaben erledigt werden.

¹⁰ Zur Orientierung: die schnellsten Prozessoren (Stand Mai 2024) benötigen weniger als 0.1 Nanosekunden für einen DRAM Zugriff. Die Datenbank benötigt ca. 0.5 Mikrosekunden 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.

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. Bereits der Austausch von Messages kann zu einer Begrenzung der Skalierbarkeit führen.
- 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® 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, Persistent 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 peakmarks® Workloads zur Vermessung der Leistungsfähigkeit von Storage Systemen im Oracle Datenbankbetrieb.

Workload	Aktion	Performance Metrik	Mass-einheit
STO-READ	Durch SQL-Statement erzeugte sequenzielle Leseoperation.	Durchsatz sequential I/O	[MBps]
STO-OFFLOAD	Durch SQL-Statement erzeugte sequenzielle Leseoperation bei Verwendung intelligenter Storage Server mit off-load Funktionalität (smart scan).	Durchsatz sequential I/O	[MBps]
STO-RANDOM	Durch SQL-Statement erzeugte zufällig gestreute Lese- und Schreiboperationen einzelner Datenbankblöcke. 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 ¹¹ .	Durchsatz random I/O I/O Servicezeit	[IOPS] [ms]
STO-SCATTER	Durch SQL-Statements erzeugte zufällig gestreute Schreiboperationen. Bei diesem Workload werden Schreib-Operationen durch Vordergrundprozesse durchgeführt.	Durchsatz random I/O	[dbps]

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 keine Oracle DOP 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]
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

Beispiel: Workload STO-OFFLOAD

Der Workload STO-OFFLOAD ist ähnlich 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 Systeme den Zugriff selbstständig optimieren.

¹¹ Oracle führt Schreiboperationen häufig als multiblock I/O Operation durch; d.h. mit einer I/O Operationen wird eine Liste von Blöcken geschrieben.

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

Beispiel: Workload STO-RANDOM

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

Im ersten Benchmark Report (Run 6, Test 65 bis 81) werden nur lesende Operationen vermessen. Beim zweiten Benchmark Report (Run 6, Test 97 – 102) beträgt das Verhältnis von Lese- zu Schreiboperationen 80:20 (Spalte Wri [%]).

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 jeweiligen Cache des Storage Systems gefunden werden.

Tiered Storage. Die hier vermessenen Systeme verfügen über einen Tiered Storage mit drei Cache Layern. Der I/O Durchsatz und die I/O Servicezeit hängen stark davon ab, in welchem Cache ein Block gefunden wird. Durch wiederholte Ausführung der Tests und ein Aufwärmen der Cache Layer konnten sowohl Durchsatz als auch I/O Servicezeit deutlich verbessert werden. Tiered Storage Systeme liefern per se keine verlässliche Performance.

Log Writer Workloads

Log Writer Prozesse sind 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 für die Datenbankreplikation verwendet, um Standby-Datenbanken zu synchronisieren. Diese Technologie ist sehr beliebt bei Disaster 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 die Standby-Datenbanken das Transaktionsprotokoll erhalten haben. Dies kann die lokale Transaktionsverarbeitung verzögern¹².

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].

Workloads für Log Writer Prozesse

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

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

¹² 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.

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]
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

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]
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	2,600	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

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] oder megabyte per second [MBps]

Workloads zur Ermittlung der Database Writer Performance

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

Workload	Aktion	Performance Metrik	Masseneinheit
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	#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

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 ASM Level high redundancy verwendet. Die im peakmarks® Report angegebene Datenmenge von 844,609 database blocks per second wurde in der Datenbank gemessen. ASM muss aber tatsächlich die dreifache Datenmenge auf verschiedene Disk Groups schreiben.

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® Software bietet Workloads für alle drei Ladetechniken an.

Workload	Aktion	Performance Metrik	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]
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]

Workload	Aktion	Performance Metrik	Mass-einheit
DL-STREAM	Die Daten werden programmgeneriert und verwenden den memory-optimized row store zum Laden von Daten in 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.

Die Metrik Load rate total [MBps] entspricht der tatsächlich geladenen Datenmenge. Die Metrik Load rate total [rps] hängt vom Datenmodell ab und gilt hier für eine Satzlänge von 300 Byte (Satzlänge der peakmarks® Tabellen).

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

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.

Auch hier entspricht die Metrik Load rate total [MBps] der tatsächlich geladenen Datenmenge. Die Metrik Load rate total [rps] hängt vom Datenmodell ab und gilt hier für eine Satzlänge von 300 Byte (Satzlänge der peakmarks® Tabellen)

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

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 Analytics 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 Metrik	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 im Row Store des Buffer Cache gelesen.	Data Scan Durchsatz	[MBps]
DA-COLSTORE	Full table scan mit Aggregat geringer Komplexität. Die Daten werden im Column Store des 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 Scan rate total [MBps] entspricht der tatsächlich durchsuchten Datenmenge. Die Metrik Scan rate total [rps] hängt vom Datenmodell ab und gilt hier für eine Satzlänge von 300 Byte (Satzlänge der peakmarks® Tabellen).

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

Beispiel: Workload DA-OFFLOAD

Der Workload DA-OFFLOAD zeigt die maximale Scan Rate an, wenn alle Daten vom intelligenten Storage System verarbeitet werden. Dabei werden können Filteroperationen und auch Aggregate auf dem Storage System verarbeitet werden. Lediglich die Treffermenge wird zum Datenbank-Server übertragen.

Die Metrik Scan rate total [MBps] entspricht der tatsächlich durchsuchten Datenmenge. Die Metrik Scan rate total [rps] entspricht der äquivalenten Datenmenge, bei einer Satzlänge von 300 Byte (Satzlänge der peakmarks® Tabellen).

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]
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

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]
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

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]
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

Die Metrik Scan rate total [rps] entspricht der tatsächlich durchsuchten Datenmenge. Die Metrik Scan rate total [MBps] hängt vom Datenmodell ab und gilt für eine Satzlänge von 300 Byte (Satzlänge der peakmarks® Tabellen).

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. Hier müssen in der Regel Eingriffe in die Applikation vorgenommen werden, um zu spezifizieren welche Spalten einer Tabelle in den Column Store aufgenommen werden sollen. Optionale Komprimierungsverfahren helfen dabei, die Kapazität des Column Store optimal zu nutzen.

Online Transaction Processing Workloads

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

Online Transaction Processing ist eine äusserst komplexe Form der Datenverarbeitung. Alle Plattform Komponenten und Datenbank Serviceprozesse sind massgeblich involviert und müssen genau austariert werden. Online 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 Replikation von Daten zwischen verschiedenen Data Center beeinflusst werden.

Kennzahlen für Transaction Processing Performance

Für Online 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® 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 Metrik	Mass-einheit
TP-LOOKUP	Schnelle Look-up Query, welche die Existenz eines einzelnen Datensatzes prüft. Beispiele: 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.	Query Durchsatz Query Antwortzeit	[tps] [ms]
TP-REPORT	Online Report, der durchschnittlich 25 Datensätze pro Query selektiert.	Query Durchsatz Query Antwortzeit	[tps] [ms]
TP-LIGHT	Selektiere/Mutiere einzelnen Datensatz via Index. Beispiele: 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. Beispiele: 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. Beispiele: 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) Logical block access: 83% read, 17% write. Avg REDO per transaction 256-byte.	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) Logical block access: 65% read, 35% write. Avg REDO per transaction 1,725-byte.	Durchsatz von Transaktionen Antwortzeit von Transaktionen	[tps] [ms]

Als Besonderheit bietet die peakmarks® Software die Möglichkeit, einen kundenspezifischen Transaktionsmix festzulegen.

Workload	Aktion	Performance Metrik	Mass-einheit
TP-CUSTOM	Konfigurierbarer Workload Mix auf der Basis von einfachen Workloads.	Durchsatz von Transaktionen	[tps]
		Antwortzeit von Transaktionen	[ms]

Beispiel: Lese-intensiver Workload TP-MIXED1

Der Übersichtsreport mit den zentralen Performance Kennzahlen umfasst sowohl die primären Leistungskennzahlen wie Durchsatz und Antwortzeit von Transaktionen als auch wichtige Einflussfaktoren: Buffer Cache Hitrate (Spalte BuCache read [%]), Flash Cache Hitrate (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]).

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

Beispiel: Schreib-intensiver 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

Anmerkungen

peakmarks® stellt zusätzliche Reports zur Analyse des Leistungsverhaltens der Plattform bei Online Transaction Processing mit folgenden Informationen zur Verfügung: logical I/O pro Transaktion (read, write), physical I/O pro Transaktion (read, write), REDO Datenmenge pro Transaktion, Anzahl Checkpoints, etc.

Die beiden Beispiele zeigen nur eine moderate CPU-Auslastung. Mit Hilfe von AWR-Reports, die auch Empfehlungen des Automatic Database Diagnostic Monitor (ADDM) enthalten, kann der Engpass schnell festgestellt werden. In diesem Fall ein zu kleiner Buffer Cache für eine 16 TByte Datenbank.

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 Anwendungen, wie z.B. Core Banking Systems, nutzen intensiv PL/SQL, um Algorithmen und Transaktionslogik möglichst nah an den Daten verarbeiten zu können.

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 Metrik	Mass-einheit
PLS-ADD	Addition von Zahlen verschiedener Datentypen	Durchsatz von PL/SQL Operationen	[Mops]
PLS-BUILTIN	Datentyp-spezifische Operationen einschliesslich SQL-Funktionen, basierend auf Core Banking Systemen und Telco Billing Systemen	Durchsatz von 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 von 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]
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

Beispiel: Workload PLS-FIBO

Der Workload PLS-FIBO zeigt die Verarbeitungszeit für die Berechnung der Fibonacci Zahlenfolge bis N (Spalte Parameter) an. Dabei wird ein rekursiver Algorithmus verwendet. Der Workload PLS-PRIME ermittelt alle Primzahlen beginnend bei 1 bis N (Spalte 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

Fazit

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

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

Die peakmarks® Software deckt alle wichtigen Workloads für ein Performance Assessment 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 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 verlässlichen (predictable) und dauerhaften (persistent) Performance in allen Lastsituationen.

Vorteile für IT-Hardware Anbieter und Service Provider. peakmarks® Performance Assessments 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.

Zusammenfassung der Performance Ergebnisse

In einem peakmarks® Executive Summary werden die wichtigsten Ergebnisse in verständlicher Form für Entscheidungsträger zusammengefasst.

peakmarks® Key Performance Metriken für Plattform Komponenten im Oracle Datenbankbetrieb

Kategorie	Key Performance Metrik	Antwortzeit	Durchsatz	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

Kategorie	Key Performance Metrik	Durchsatz	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 Metriken für kritische Datenbank Hintergrund Prozesse

Kategorie	Key Performance Metrik	Latenzzeit	Durchsatz	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

Kategorie	Key Performance Metrik	Durchsatz	peakmarks® Workload
Database Writer	Database blocks written	844,609 dbps	DBWR-THR

peakmarks® Key Performance Metriken für repräsentative Datenbank-Operationen

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

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

Kategorie	Key Performance Metrik	Antwortzeit	Durchsatz	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 für PL/SQL application code

Kategorie	Key Performance Metrik	Antwortzeit	Durchsatz	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)