

Ressourcen Engpässe erkennen mit Exadata AWR

Frank Schneede

Oracle Deutschland B. V. & Co. KG

Hannover

Schlüsselworte

Exadata Database Machine X5, AWR, Statspack

Einleitung

Die Exadata Database Machine ist mittlerweile seit vielen Jahren als extrem leistungsfähige Plattform für die Oracle Datenbank bei zahlreichen Kunden im Einsatz. Anfänglich primär als System für eine einzelne besonders performance-kritische Anwendung genutzt, sieht man die Exadata zunehmend als Konsolidierungsplattform. Die Auslastung der Systeme steigt in Konsolidierungsszenarien erfahrungsgemäß an, daher kommt einer Identifizierung möglicher Bottlenecks oder Fehlerquellen im Betrieb mittlerweile eine besondere Bedeutung zu. Die für den Exadata Betrieb genutzten Werkzeuge wie Enterprise Manager Cloud Control oder auch AWR Reports liefern zwar Aussagen über besonders aufwändige SQL Statements auf Datenbankebene oder zeigen echte HW Ausfälle an, eine Analyse auf System- oder Storageebene hingegen war bislang nicht möglich.

Die seit kurzem verfügbaren Exadata Erweiterungen im AWR Report bzw. Active Reports bieten sehr gute Analysemöglichkeiten für eventuelle Ursachen von Performanceproblemen auf Exadata Systemebene.

Viele mögliche Fehlerquellen

Nach der produktiven Inbetriebnahme der Exadata Database Machine beim Kunden durchläuft die Installation üblicherweise einen Abnahmeprozess, so dass zu diesem Zeitpunkt davon ausgegangen werden kann, dass das Setup der Maschine den von Oracle festgelegten Best Practises entspricht. Der Ist-Zustand zum Zeitpunkt der Übergabe an den Kunden wird in einem Protokoll festgehalten, zu dem auch das Ergebnis eines exachk gehört, der alle Konfigurationsparameter auf System und Datenbankebene gegen Oracle Best Practises prüft. Die MOS-Note „[Oracle Exadata Database Machine exachk or HealthCheck \(Doc ID 1070954.1\)](#)“ wird stetig überarbeitet und enthält die aktuell gültige exachk Version.

Wenn es im Lebenszyklus der Exadata Installation zu undifferenzierten Performanceproblemen kommt, so kann es dafür viele unterschiedliche Ursachen geben. Eine mögliche Ursache ist häufig eine neue Datenbank, die wesentliche Teile der zur Verfügung stehenden I/O Bandbreite für sich beansprucht und dadurch andere Datenbanken beeinträchtigt.

Eine andere häufig in Projekten identifizierte Ursache ist eine plötzlich einsetzende hohe I/O Last, verursacht durch Wartungsoperationen wie Rebalancing und Backup oder auch einfach durch Ansteigen der Anwender- bzw. Anwendungslast. Da in der Exadata Daten nach dem SAME-Prinzip (Stripe And Mirror Everything) verteilt werden, kann sogar eine einzelne langsam arbeitende Festplatte die Systemperformance als Ganzes beeinträchtigen. Offload-Aktivitäten werden von den CPUs auf Storage-Ebene übernommen; auch hier kann sich die CPU-Leistung als Beeinträchtigung für die Gesamtsystemperformance herausstellen. Zu guter Letzt kann auch ein Konfigurationsunterschied (z. B. Verschieden SW Versionen auf Zellebene, Offline geschaltete Platten) zu undifferenzierten Performanceproblemen führen.

Viele herkömmliche Analysemethoden und –werkzeuge

Um undifferenzierten Performance problemen auf die Spur zu kommen, gab es auch bislang schon viele Ansätze, die zum Teil jedoch recht umständlich waren. Für die tiefgehende Analyse notwendige Daten stehen lediglich in Form von Metriken auf den Storage Servern zur Verfügung und finden keinen Niederschlag in Statistiken auf Datenbankebene. Aus diesem Grund wurden Werkzeuge entwickelt, die diese Daten zusammentragen und auswerten können. Eines dieser Werkzeuge ist das Analyseskript ExaWatcher.sh, das auf jedem Datenbank- und Storage-Server im Verzeichnis /opt/oracle.ExaWatcher/, das Systemdaten sammelt und das bisher verwendete OSWatcher Utility ab Exadata Version 11.2.3.3 ersetzt. Ein detaillierte Beschreibung des von ExaWatcher ist in der MOS-Note „[ExaWatcher utility on Exadata database servers and storage cells \(Doc ID 1617454.1\)](#)“ zu finden. Neben ExaWatcher Ausgaben können zusätzlich Trace-Files, Metriken oder Logdateien aus /var/log/* und vieles zur Analyse herangezogen werden.

Viele erfahrene DBAs nutzen auch gerne eigens entwickelte Skripte zur Eingrenzung von Problemen. Solche Skripte sind oftmals historisch gewachsen und können gerade die Spezifika einer Exadata Installation nur unzureichend abbilden. Ein bewährtes Mittel sind darüberhinaus AWR-Reports, die bislang jedoch nur Aussagen für einzelne Datenbanken beinhalteten und einen Rückschluß auf den Zustand des Gesamtsystems nur sehr mittelbar ermöglichten. War es in einer Konsolidierungsumgebung notwendig, auf AWR-Reports zurückzugreifen, konnte man nur anhand aller AWR-Reports aller konsolidierten datenbanken sinnvoll Analysen durchführen. Wenn man zum Beispiel mittels AWR-Reports eine Analyse eines Exadata Half Racks mit 4 Datenbanken durchführen möchte, müsste man für jede Instanz einen AWR Report zur Verfügung haben – das macht in diesem Beispiel 16 AWR Reports, die dann das gesamte System abbilden. Eine solche Analyse ist verständlicherweise sehr zeitaufwändig und umständlich!

Neue Analysemöglichkeiten durch Exadata AWR

Beginnend mit Oracle Database 12.1.0.2 und Exadata Storage Server 12.1.2.1.0 wurde die Möglichkeit geschaffen, dass die Inhalte von Metriken auf Zellebene in der Datenbank gesammelt werden. Es gab zwar bereits in vorherigen Datenbankversionen Views, die zum Beispiel die Konfiguration von Zellen beinhalteten, aber mit der aktuellen Datenbankversion hat es sehr umfangreiche Erweiterungen gegeben, wie folgendes Listing zeigt:

```
SQL> select table_name
       2  from dict
       3  where regexp_like(table_name, 'DBA.*(ASM|CEL)|^V\$\$CELL');
```

```
TABLE_NAME
-----
DBA_HIST_ASM_BAD_DISK
DBA_HIST_ASM_DISKGROUP
DBA_HIST_ASM_DISKGROUP_STAT
DBA_HIST_CELL_CONFIG
DBA_HIST_CELL_CONFIG_DETAIL
DBA_HIST_CELL_DB
DBA_HIST_CELL_DISKTYPE
DBA_HIST_CELL_DISK_NAME
DBA_HIST_CELL_DISK_SUMMARY
DBA_HIST_CELL_GLOBAL
DBA_HIST_CELL_GLOBAL_SUMMARY
DBA_HIST_CELL_IOREASON
DBA_HIST_CELL_IOREASON_NAME
DBA_HIST_CELL_METRIC_DESC
DBA_HIST_CELL_NAME
```

```

DBA_HIST_CELL_OPEN_ALERTS
V$CELL
V$CELL_CONFIG
V$CELL_CONFIG_INFO
V$CELL_DB
V$CELL_DB_HISTORY
V$CELL_DISK
V$CELL_DISK_HISTORY
V$CELL_GLOBAL
V$CELL_GLOBAL_HISTORY
V$CELL_IOREASON
V$CELL_IOREASON_NAME
V$CELL_METRIC_DESC
V$CELL_OFI_THREAD_HISTORY
V$CELL_OPEN_ALERTS
V$CELL_REQUEST_TOTALS
V$CELL_STATE
V$CELL_THREAD_HISTORY

```

33 Zeilen ausgewählt.

Während zum Beispiel v\$cell statische Informationen enthält, gibt es andere Views, wie zum Beispiel v\$cell_disk, deren Statistiken seit dem Start des cellsrv Prozesses fortgeschrieben werden. Views wie v\$cell_ofi_thread_history bieten eine Art von ASH-Report auf Exadata Zellenebene. Hierbei steht nicht nur der aktuelle Zustand zur Verfügung, sondern auch historische Informationen, mit Hilfe derer über die SQL-ID ein aussagekräftiges Bild für jeden denkbaren Zeitpunkt erzeugt werden kann.

Die View v\$cell_state enthält sehr viele Informationen, wobei die anzufragenden Statistiken von der eingesetzten Kombination von Oracle Datenbank und Exadata Software abhängen. Um die Vielfalt der gesammelten Daten zu illustrieren, sind einige wesentliche Statistiken exemplarisch in der folgenden Tabelle aufgelistet:

Statistik Type	Bedeutung
IOREASON	Schlüsselt jeden I/O auf der Zelle auf
RCVPORT	Empfangener Netzwerkverkehr
FLASHLOG	Details zur Nutzung des Flashlog Features
SENDPORT	Gesendeter Netzwerkverkehr
PREDIO	Informationen über Smart Scan
NPHYSDISKS	Anzahl der physischen Platten pro Zelle
PHASESTAT	Unterschiedliche Phasen des Smart Scan
LOCK	Mutex Waits pro Objekttyp
OFLGROUP	Offload Server Statistik

Natürlich können diese Informationen auch über eigene Skripte ausgewertet werden, das setzt allerdings sehr tiefes Know-How oder ein sehr gutes Buch voraus. Daher steht hier die Nutzung der Informationen durch AWR Exadata im Vordergrund.

AWR ist ein auf Datenbankebene weithin bekanntes und geschätztes Werkzeug, das allerdings die Lizenzierung des Diagnostic Pack erfordert. Daher war es nur naheliegend, AWR um Exadata Zelleninformationen zu erweitern, denn die Datenbank hat ja Zugriff auf alle Exadata Zellen, was eine Storage-Grid weite Datensammlung und Analyse ermöglicht. Zu beachten ist hierbei, dass die Zellenstatistiken grundsätzlich eine systemweite Sicht und somit auch die Auswirkungen durch andere Datenbanken darstellen und nicht nur die betrachtete Datenbank reflektieren!

AWR Reports können als html-File erzeugt werden oder active Report – eine reine textbasierte Reportvariante steht nicht zur Verfügung.

Im AWR Exadata Report einer einzelnen Instanz sind folglich auch andere Top Consumer zu sehen, das heißt I/O anderer Datenbanken auf dem gleichen System, die die eigene Performance möglicherweise beeinträchtigen. Um hohe eine hohe I/O Last analysieren zu können, wird festgestellt, welche Gründe für I/O vorliegen. Im Normalfall sollte man I/O von z. B. Smart Scan oder herkömmlichen Buffer Cache Reads erwarten – wenn nun schlecht balancierte Operationen vorkommen oder ein Backup durchgeführt wird, sieht man das im AWR Exadata Report. Da die Daten pro Exadata Zelle gesammelt werden, können auch hotspots identifiziert werden, die zu einer Ungleichverteilung von I/O und damit zu einer schlechteren Gesamtsystemperformance führen können.

AWR Exadata wird genutzt, um „Ausreisser“ zu identifizieren, also die Zelle oder Disk, die „anders“ ist, sei es in Hinblick auf deren Verhalten oder auf deren Konfiguration (SW Stände, Setup). Die Identifikation von „Ausreissern“ in Bezug auf Performanc erfolgt, indem unterschiedliche Metriken gegen Schwellenwerte oder gegen Mittelwerte verglichen wird.

Neue Passagen im AWR Exadata

Das Aufrufen eines AWR Reports entspricht dem bekannten Vorgehen mittels der Skripte awrrpt.sql für eine single instance und awrgrpt.sql für eine RAC Datenbank. Die Skripte liegen im Verzeichnis \$ORACLE_HOME/rdbms/admin. Soll ein Active Report erstellt werden, geschieht das überr das Skript perfhubrpt.sql, das ebenfalls im genannten Verzeichnis zu finden ist.

Im AWR Report sind die Exadata Informationen am Ende zu finden, die gewünschte Passgae ist über Links leicht erreichbar, das Menü ist in Abbildung 1 gezeigt.

Exadata Configuration and Statistics

- [Exadata Report Summary](#)
- [Exadata Server Configuration](#)
- [Exadata Server Health Report](#)
- [Exadata Statistics](#)

[Back to Top](#)

Abb. 1: Einstiegspunkt zu AWR Exadata

Die verfügbaren Informationen gliedern sich in

- Konfigurationsinformationen
- Status der Systemgesundheit
- Diverse Exadata Systemstatistiken

In den Konfigurationsdaten ist unter anderem zu sehen, wie die Softwarestände auf der Exadata sind, die Aufteilung der Festplatten und der Flashkarten. Abbildung 2 zeigt, dass auf allen Zellen die gleiche SW Versionen verwendet werden, wie es in der Regel auch sein sollte.

Exadata Storage Server Model

- Model Information of Servers
- Cpu Count refers to Logical CPUs, including Cores and HyperThreads

Model	CPU Count	Memory(GB)	# Cells	Cells
Oracle Corporation SUN SERVER X4-2L High Capacity	24	95	3	scao06celadm01; scao06celadm02; scao06celadm03

[Back to Exadata Server Configuration](#)
[Back to Top](#)

Exadata Storage Server Version

- Version Information of Packages on the Server

Package Type	Package Version	Cells
Kernel	2.6.39-400.128.9.el5uek.bug18492778	All
Cell	cell-12.1.2.1.0_LINUX.X64_140501-1	All
Offload	cellofi-11.2.3.3.0_LINUX.X64_131014.1	All
Offload	cellofi-12.1.2.1.0_LINUX.X64_140501	All

[Back to Exadata Server Configuration](#)
[Back to Top](#)

Abb. 2: Konfiguration der Exadata im AWR

In der Exadata Health Summary werden Alerts zusammengefasst, so dass hier unmittelbar Aktionen zum beheben des Problems abgeleitet werden können. Abbildung 3 zeigt eine Summary mit wenigen Problemen.

Exadata Server Health Report

- Exadata Alerts Summary
- Exadata Alerts Detail
- Exadata Non-Online Alerts

[Back to Top](#)

Exadata Alerts Summary

- Number of open alerts at the end snapshot

Cell Name	Total	Critical	Warning	Info
Total	6	0	3	3
fpnrcas01_adm	1	1	0	0
fpnrcas02_adm	1	1	0	0
fpnrcas03_adm	1	1	0	0
fpnrcas04_adm	1	1	0	0
fpnrcas05_adm	2	2	0	0

[Back to Exadata Server Health Report](#)
[Back to Top](#)

Exadata Alerts Detail

- Number of open alerts at the end snapshot
- Only the 10 most recent open alerts per cell are displayed
- Sorted by Cell Name, Begin Time desc

Cell Name	Alert Time	Severity	Status	Message
fpnrcas01_adm	04:17 15:00:23	critical	Y	"A power supply component is suspected of causing a fault. -exadata.br/>Component Name : /SYSPS2 -exadata.br/>Fault class : fault class:err power loss -exadata.br/>Fault message : http://www.sun.com/msg/SP130-8306-01"
fpnrcas02_adm	04:17 15:07:40	critical	Y	"A power supply component is suspected of causing a fault. -exadata.br/>Component Name : /SYSPS2 -exadata.br/>Fault class : fault class:err power loss -exadata.br/>Fault message : http://www.sun.com/msg/SP130-8306-01"
fpnrcas03_adm	04:17 15:19:26	critical	Y	"A power supply component is suspected of causing a fault. -exadata.br/>Component Name : /SYSPS2 -exadata.br/>Fault class : fault class:err power loss -exadata.br/>Fault message : http://www.sun.com/msg/SP130-8306-01"
fpnrcas04_adm	04:17 15:20:04	critical	Y	"A power supply component is suspected of causing a fault. -exadata.br/>Component Name : /SYSPS2 -exadata.br/>Fault class : fault class:err power loss -exadata.br/>Fault message : http://www.sun.com/msg/SP130-8306-01"
fpnrcas05_adm	08:28 12:33:57	critical	Y	"A genio component is suspected of causing a fault. -exadata.br/>Component Name : /SYSMP -exadata.br/>Fault class : fault severity:err data open -exadata.br/>Fault message : http://www.sun.com/msg/SP130-8306-01"
	13:23 14:18:20	critical	Y	"A genio component is suspected of causing a fault. -exadata.br/>Component Name : /SYSMP -exadata.br/>Fault class : fault class:err boot power-off requested -exadata.br/>Fault message : http://www.sun.com/msg/SP130-8306-01"

Abb. 3: Exadata Health Summary

Die interessanteren Informationen finden sich im Bereich der Exadata Statistiken. Hier werden unterschiedliche Aspekte analysiert, mit umfangreichen Erläuterungen versehen und sogar farblich gekennzeichnet. Abbildung 4 zeigt die verschiedenen Rubriken in diesem Abschnitt.

Exadata Statistics

- [Exadata Resource Statistics](#)
 - [Exadata Outlier Summary](#)
 - [Exadata OS Statistics Outliers](#)
 - [Exadata OS IO Statistics](#)
 - [Exadata OS IO Latency](#)
 - [Exadata OS CPU Statistics](#)
 - [Exadata Cell Server Statistics Outliers](#)
 - [Cell Server IOPS Statistics](#)
 - [Cell Server IO MB/s Statistics](#)
 - [Cell Server IO Latency Statistics](#)
 - [Exadata Outlier Details](#)
 - [Exadata OS Statistics Outlier Details](#)
 - [Exadata Cell Server Outlier Details](#)
 - [Exadata OS Statistics Top](#)
 - [Exadata OS IO Statistics](#)
 - [Exadata OS IO Latency Statistics](#)
 - [Exadata OS CPU Statistics](#)
 - [Exadata Cell Server Statistics Top](#)
 - [Cell Server IOPS Statistics](#)
 - [Cell Server IO MB/s Statistics](#)
 - [Cell Server IO Latency](#)
- [Exadata Smart Statistics](#)
 - [Smart IO](#)
 - [Flash Log](#)
 - [Flash Cache](#)
- [Exadata IO Reasons](#)
 - [Top IO Reasons by Requests](#)
 - [Top IO Reasons by MB](#)
- [Exadata Top Database Consumers](#)
 - [Top Databases by Requests](#)
 - [Top Databases by Throughput](#)
 - [Top Databases by Requests per Cell](#)
 - [Top Databases by Throughput per Cell](#)

Abb. 4: Exadata Health Summary

Mit Hilfe der Analyse von Ausreißern lassen sich viele Problemursachen aufspüren und Ansatzpunkte für deren Behebung finden. Abbildung 5 zeigt Ausreißer im Bereich Betriebssystem I/O der Festplatten.

Exadata OS IO Statistics - Outlier Disks

- These statistics are collected by the OS on the cells and are not restricted to this database or instance
- Outliers are disks whose average performance is outside the normal range, where normal range is ± 2 standard deviation
- Outlier disks must have a minimum of 10 IOPs. Idle disks are not considered for outlier analysis.
- Outliers for hard disks are displayed when Hard Disk IOPs exceeds 60.12 (1% of maximum capacity of 6,012)
- Outliers for flash disks are displayed when Flash Disk IOPs exceeds 4285.52 (1% of maximum capacity of 428,552)
- A ∇ and a dark yellow background indicates an outlier value below the low range
- A ∇ and a light red background indicates an outlier value above the high range
- % Total - Avg [IOPs | IO MB/s] of the disk as a percentage of total [IOPs | IO MB/s] for the disk type

Disk Type	Cell Name	Disk Name	# Disks	IOPs			IO MB/s			% Disk Utilization					
				% Total	Mean	Std Dev	Normal Range	% Total	Mean	Std Dev	Normal Range	Mean	Std Dev	Normal Range	
F373G	All	All	48		3,146.68	207.77	2,731.14 - 3,562.23		142.88	12.46	117.96 - 167.80		48.34	4.66	39.03 - 57.66
H3.6T	All	All	36		∇ 282.47	21.49	239.20 - 325.15		3.68	1.51	0.65 - 6.70		68.90	5.92	77.05 - 100.75
Outlier	sca06celadm03	CD_05_sca06celadm03			3.21	∇ 326.62			3.32	4.40			91.33		
Outlier	sca06celadm01	CD_11_sca06celadm01			2.47	∇ 250.68			2.51	3.32			∇ 75.24		

[Back to Exadata OS Statistics Outliers](#)
[Back to Exadata Resource Statistics](#)

Abb. 5: Exadata OS IO Statistics – outlier Disks

Da die Exadata Statistiken im AWR Exadata für das Gesamtsystem ermittelt werden, bietet der letzte Abschnitt im AWR Report aufschlußreiche Informationen zum Ressourcenverbrauch je Datenbank. Abbildung 6 zeigt die Top Konsumenten nach IO Requests und nach IO Bandbreite.

Top Databases by IO Requests

- The top databases by IO Requests are displayed
- At most 10 databases are displayed
- %Total - % of Total IO requests
- Total - total IO requests or IO throughput (Flash + Disk)
- Ordered by IO requests desc

DB Name	DBID	%Total	IO Requests			IO Throughput (MB)		
			Total Requests	Flash	Disk	Total MB	Flash	Disk
FINANCE	844757284	51	105,708,303	104,651,220	1,057,083	825,846	817,587	8,258
HR	4150816880	35	71,198,064	70,493,343	704,721	556,234	550,729	5,505
SALES	112018265	10	20,522,949	20,319,750	203,199	14,821,785	14,618,586	203,199
REPORT	3050038566	4	8,795,550	8,708,466	87,084	8,795,550	8,708,466	87,084
ASM	1	0	75,483	0	75,483	0	0	0

[Back to Exadata Top Database Consumers](#)

[Back to Exadata Statistics](#)

Top Databases by IO Throughput

- The top database by IO Throughput are displayed
- At most 10 database are displayed
- %Total - % of Total IO throughput
- Total - total IO throughput or IO requests (Flash + Disk)
- Ordered by IO Throughput desc

DB Name	DBID	%Total	IO Throughput (MB)			IO Requests		
			Total MB	Flash	Disk	Total Requests	Flash	Disk
SALES	112018265	59	14,821,785	14,618,586	203,199	20,522,949	20,319,750	203,199
REPORT	3050038566	35	8,795,550	8,708,466	87,084	8,795,550	8,708,466	87,084
FINANCE	844757284	3	825,846	817,587	8,258	105,708,303	104,651,220	1,057,083
HR	4150816880	2	556,234	550,729	5,505	71,198,064	70,493,343	704,721
ASM	1	0	0	0	0	75,483	0	75,483

[Back to Exadata Top Database Consumers](#)

[Back to Exadata Statistics](#)

Abb. 6: Exadata Top Consumer

Fazit

Durch das Sammeln von Exadata System- und Performancedaten in Datadictionary Views und deren Auswertung mittels AWR Reports oder als Active Reports sind die Analysemöglichkeiten von Performanceproblemen auf der Exadata in den letzten Releases sehr stark erweitert worden. Die in den AWR Reports enthaltenen Hinweise sind dazu geeignet, dass auch weniger versierte Anwender im AWR Report ein gutes Hilfsmittel zur Exadata Administration vorfinden.

Weiterführende Informationen

[Oracle Exadata Database Machine exachk or HealthCheck \(Doc ID 1070954.1\)](#)

[ExaWatcher utility on Exadata database servers and storage cells \(Doc ID 1617454.1\)](#)

[Oracle Exadata Database Machine System Overview Release 1 \(12.1\), What's New in Oracle Exadata Database Machine 12c Release 1 \(12.1.2.1.0\): Oracle Exadata Storage Server Software Performance Statistics in AWR Reports](#)

[Oracle Exadata Database Machine System Overview Release 1 \(12.1\), What's New in Oracle Exadata Database Machine 12c Release 1 \(12.1.2.2.0\): Oracle Exadata Storage Statistics in AWR Reports](#)

[Oracle Exadata Storage Server Software User's Guide 12c Release 1 \(12.1\), Appendix D:
Automatic Workload Repository Views for Oracle Exadata Storage Server Software](#)

Kontaktadresse:

Frank Schneede
Oracle Deutschland B. V. & Co. KG
Thurnithistr. 2
D-30519 Hannover

Telefon: +49 (0) 511 95787 - 250
Fax: +49 (0) 511 95787 - 100
E-Mail frank.schneede@oracle.com
Internet: www.oracle.com