

Oracle Exadata Database Machine – next Generation

Frank Schneede, ORACLE Deutschland B. V. & Co. KG

Auf der diesjährigen OpenWorld stellte Oracle die dritte Generation der Exadata Database Machine vor. Neben der Aktualisierung der eingesetzten Hardware für die bekannten Modelle wurde das neue Modell Exadata Database Machine X2-8 präsentiert, das sich primär an Betreiber großer OLTP-Datenbanken richtet. Der Artikel gibt einen Überblick über das gesamte Angebot, das den Kunden für nahezu jedes Einsatzszenario der Oracle-Datenbank interessante Möglichkeiten bietet.

Exponentiell wachsende Datenmengen stellen aktuelle Datenbank-Systeme und die darauf laufenden Anwendungen zunehmend vor große Herausforderungen, wie unabhängige Untersuchungen („Information, Unplugged: 2009 OAUG ResearchLine Survey on Enterprise Application Information Lifecycle Management“) zeigen. Zu diesem Datenwachstum tragen einerseits gesetzliche Aufbewahrungsvorschriften bei, andererseits steigt auch das Bedürfnis der Anwender in Hinblick auf Granularität und Aktualität der gewünschten Informationen.

Mit herkömmlichen Architektur-Ansätzen ist der großen Datenmengen kaum Herr zu werden, da systemimmanente „Bottlenecks“ (SAN, Controller, Anzahl Platten etc.) den I/O-Durchsatz begrenzen. Die Exadata-Technologie setzt mit ihrer Grid-Architektur genau hier an: Datenintensives Suchen und Lesen findet bereits im Storage-Grid statt, durch die Filterung und Weitergabe nur der relevanten Informationen („Query Offloading“) wird das gesamte I/O-System entlastet. Die Datenbank profitiert ebenfalls von diesem Vorgehen, denn dort erfolgt nur noch die rechenintensive Verarbeitung der durch Query Offloading reduzierten (Teil-)Ergebnismenge. Freigewordene CPU-Ressourcen im Database-Grid können für andere Zwecke genutzt werden.

Die geschilderte Reduzierung des I/O-Volumens und die parallele Abarbeitung der I/O-Requests im Storage führen zu enorm hohen I/O-Durchsatzraten, die mit dem Einsatz herkömmlicher NAS- oder SAN-Storage-Systeme nicht beziehungsweise nur mit sehr großem Aufwand erreichbar

sind. Der Grundgedanke der Exadata-Technologie hat weiterhin Gültigkeit, aktualisierte Hardware sorgt für noch mehr I/O-Durchsatz.

X2 statt V2

Vor zwei Jahren sorgte die erste Version der Exadata Database Machine für einiges Aufsehen. Der Ansatz, standardisierte, kostengünstige Hardware-Komponenten für Datenbank, Netzwerk und Storage mit einem innovativen Software-Konzept zu verknüpfen und als Gesamtlösung für den Betrieb von Oracle-Datenbanken anzubieten, war vollkommen neu und stieß bei den Kunden auf großes Interesse.

Die nachfolgende Exadata Database Machine V2 war bislang in verschiedenen Ausbaustufen verfügbar, die sich lediglich in der Anzahl der verbauten Komponenten unterschieden. So gab es Quarter-, Half- und Full-Rack-Varianten der Database Machine, die sich aus 2, 4 oder 8 Datenbankservern mit 3, 7 oder 14 Storage-Zellen zusammensetzten. In den Servern waren jeweils Quad-Core-Prozessoren eingebaut. Diese 2-Socket-Systeme heißen zukünftig X2-2, wobei das „X“ für die Intel-Chipfamilie, „2“ für die Exadata-Generation V2 und „-2“ für 2-Socket-Database-Server steht.

Neben den aktualisierten 2-Socket-Systemen steht nun ergänzend eine Exadata Database Machine X2-8 zur Verfügung, die – der Namenskonvention folgend – aus 8-Socket-Database-Servern besteht.

Sowohl die X2-2-Modelle als auch das X2-8-Modell können wahlweise mit „High Performance“ (600-GB-15.000-RPM-SAS-Festplatten) oder „High Capa-

city“ (2-TB-7.200-RPM-SAS-Festplatten) betrieben werden.

Exadata Storage Server X2-2

In allen Modellen der Exadata Database Machine kommt als Storage Server ein Sun Fire X4270 M2 Server zum Einsatz. Zwei neue 6-Core Intel Xeon Prozessoren L5640 (2,26 GHz) sorgen für zusätzliche CPU-Kapazität, es stehen damit 50 Prozent mehr Cores als bisher auf den Storage-Zellen zur Verfügung.

Die Exadata Storage Server können wahlweise mit 600-GB- oder 2-TB-SAS-Platten ausgestattet sein. Je Storage Server ergibt sich somit eine Rohdatenkapazität von 7,2 (High Performance SAS) oder 24 TB (High Capacity SAS), die durch den Einsatz von Exadata Hybrid Columnar Compression noch deutlich ansteigt.

Die mit der Exadata Version 2 eingeführten Flash-Karten sind unverändert geblieben. Vier davon sind in einem Storage Server und bieten insgesamt 384 GB Speicherplatz, der transparent für geeignete Objekte zum Einsatz kommt. Mit der Flash-Technologie ist es auch möglich, Transaktionslasten mit OLTP-Charakteristik auf der Database Machine abzubilden. Die eingebauten Dual Port InfiniBand-Karten sorgen für einen bidirektionalen Durchsatz von 40 Gb/sec in jeder Richtung zwischen Datenbank-Servern und Storage Zellen.

Extreme Performance

In der Exadata Database Machine X2-2 kommt neben dem beschriebenen Storage-Server ein Sun Fire X4170 M2 Database Server zum Einsatz. Dieser

unterscheidet sich von dem bisher eingesetzten Modell in folgenden Komponenten:

- 2 x 6-Core Intel Xeon X5670 Prozessoren (2,93 GHz)
- 96 GB Hauptspeicher
- 4 x 300 GB 10000 RPM SAS-Festplatten
- Dual-Port 10 GbE PCIe 2.0 Netzwerkkarten
- Dual-Port QDR InfiniBand HCA

Durch die neuen Prozessoren im Database- und Storage-Grid steigt der I/O-Durchsatz auf 1,8 GB/sec (je Storage Server). Damit kann eine Exadata Database Machine X2-2 Full Rack nun 25 GB/sec Daten verarbeiten. Der vergrößerte Hauptspeicher trägt den gestiegenen Anforderungen an die Datenbank SGA Rechnung; die um das Doppelte vergrößerten internen Festplatten bieten mehr Freiraum für die Oracle-Home-Verzeichnisse, Logdateien und vieles mehr.

Neue 10-GbE-Netzwerkkarten ermöglichen eine leichte Integration der Exadata Database Machine in schnelle Unternehmensnetzwerke.

Das Modell X2-2 gibt es nach wie vor als Quarter-, Half- und Full-Rack. Mit diesen drei Grundmodellen ist eine Konfiguration für nahezu jede Datenbankgröße verfügbar. Die lineare Skalierbarkeit der Exadata-Technologie stellt sicher, dass Kunden ihre Systeme bedarfsgerecht erweitern können. Investitionssicherheit wird für die Kunden großgeschrieben: Ein Anwender, der bereits heute auf eine Exadata Database Machine V2 Quarter Rack setzt, kann diese problemlos um ein Upgrade-Kit auf ein zur Hälfte aus neuen Komponenten bestehendes Half-Rack aufrüsten.

Als Betriebssystem wird weiterhin auf Oracle Enterprise Linux gesetzt, das in der Version 5.5 zum Einsatz kommt.

Konsolidierung, große OLTP-Systeme und Warehouses

Anwender, denen die Leistungsdaten der beschriebenen Exadata-Modelle nicht ausreichen, werden an der Exadata Database Machine X2-8 Gefallen

finden (siehe Abbildung 1). Zwei Sun Fire X4800 Database Server bieten ein Höchstmaß an Rechenleistung. Jeder der Server besteht aus folgenden Komponenten:

- 8 x 8-Core Intel Xeon X7560 Prozessoren (2,93 GHz)
- 1TB Hauptspeicher
- 8 x 300 GB 10000 RPM SAS Festplatten
- 2 x Network Express Modules (NEM)
- 4 x Dual-Port 10 GbE PCIe 2.0 Fabric Express Module (FEM)
- 4 x Dual-Port QDR InfiniBand PCIe 2.0 Express Module

Dieses High-End-Modell der Exadata-Produktfamilie richtet sich an Kunden, die sehr hohe Ansprüche an Rechenleistung und die SGA-Größe haben. Das können zum Beispiel große OLTP-Systeme oder Konsolidierungsprojekte sein, für die ein einzelnes Rack nicht ausreicht.

Es stehen in einem Full-Rack insgesamt 2 TB Hauptspeicher und 128 Cores zur Verfügung. Zu beachten ist allerdings die Verfügbarkeit des Systems, wenn nur ein Full-Rack eingesetzt wird. Bei einem Knotenausfall oder einer Wartung stünde dann nur die Hälfte der Systemgesamtleistung zur Verfügung.

Die Exadata Database Machine X2-8 wird mit zwei unterschiedlichen Betriebssystemen angeboten, Oracle Enterprise Linux 5.5 mit dem Oracle Linux Unbreakable Enterprise Kernel oder Solaris Express 11. Der Unbreakable Enterprise Kernel ist eine Weiterentwicklung des Kernels 2.6.32, in dem Maßnahmen zur Steigerung der Performance umgesetzt worden sind. Hierzu gehören unter anderem Verbesserungen im Non-Uniform Memory Access Scheduling (NUMA) und in der Speicherverwaltung, im Kernel-Locking-Verhalten und in der Netzwerk-Skalierbarkeit. Anwender, die strategisch auf das Betriebssystem Solaris setzen, können alternativ Solaris Express 11 als Betriebssystem auf dem Datenbank-Knoten wählen, sobald diese Version gegen Ende dieses Jahres verfügbar ist. Die beiden Betriebssystemvarianten sind besonders für den Einsatz in großen Maschinen wie der

Exadata Database Machine X2-8 optimiert worden. Als Datenbank-Software ist auf dem High-End-Modell X2-8 die Version 11.2.0.2 erforderlich, die seit Mitte September verfügbar ist.

Neue Software-Funktionen

Die Exadata Storage Server Software Version 11.2.2.1, die mit den neuen Modellen der Database Machine ausgeliefert wird, enthält ebenfalls einige Verbesserungen.

Wesentliche Erweiterungen hat es im Bereich der Systemsicherheit gegeben. In dem zur Exadata-Software gehörenden neuen Oracle Enterprise Linux 5.5 wurden unnötige Linux Packages auf Datenbank- und Storage-Servern eliminiert und Services (nfs, ftp, telnet etc.) auf den Storage-Servern deaktiviert. Zusätzlich wurde eine Firewall-Funktionalität eingeführt, die mit vordefinierten Regeln auf den Exadata-Storage-Zellen implementiert ist. Über IP-Tabellen besteht die Möglichkeit, die Firewall auf den Datenbank-Servern den Kundenanforderungen anzupassen.



Abbildung 1: Exadata Database Machine X2-8

Die Database Machine verfügt nun über eine „Calling Home“-Funktion, mit der automatisch Service Requests (ASR) für bestimmte Hardware-Probleme möglich sind (siehe Abbildung 2). Diese Funktion ermöglicht eine schnellere Fehlerbehebung, möglicherweise sogar ohne dass der Anwender den Fehler schon bemerkt hat.

Die Hardware-Fehlermeldungen sind stark verbessert worden, sodass der Administrator den Status des Systems erkennen und eventuell einleitende Maßnahmen besser umsetzen kann. Zusätzlich sind die verbauten InfiniBand-Switches durch den Integrated Lights Out Manager (ILOM) in die Systemüberwachung integriert worden, aus der dann SNMP-Benachrichtigungen entstammen.

Fehlerhafte Platten oder solche, auf denen sich ein Fehler ankündigt, werden proaktiv aus der ASM-Disk-Gruppe entfernt, ebenso fehlerhafte Flash-Karten aus dem Exadata Smart Flash Cache.

Bestandteil der neuesten Exadata-Software ist darüber hinaus die Aktualisierung aller Firmware-Komponenten für Flash, Festplattencontroller und ILOM. Neben den Verbesserungen der Exadata Storage Server Software bietet natürlich die Oracle-Datenbank in der Version 11.2.0.2 einige neue Funktionen, die für Exadata besonders interessant sind:

- Asynchrones I/O beim Rebalancing

Tabelle 1

		X2-8 Full Rack	X2-2 Full Rack	X2-2 Half Rack	X2-2 Quarter Rack
Raw Disk ¹	High Perf Disk	100 TB	100 TB	50 TB	21 TB
	High Cap Disk	336 TB	336 TB	168 TB	72 TB
Raw Flash ¹		5.3 TB	5.3 TB	2.6 TB	1.1 TB
User Data ² (assuming no compression)	High Perf Disk	28 TB	28 TB	14 TB	6 TB
	High Cap Disk	100 TB	100 TB	50 TB	21 TB
Raw Disk Data Bandwidth ^{3,5}	High Perf Disk	25 GB/s	25 GB/s	12.5 GB/s	5.4 GB/s
	High Cap Disk	14 GB/s	14 GB/s	7 GB/s	3 GB/s
Raw Flash Data Bandwidth ^{3,5}		50 GB/s	50 GB/s	25 GB/s	11 GB/s
Disk IOPS ^{4,5}	High Perf Disk	50.000	50.000	25.000	10.800
	High Cap Disk	25.000	25.000	12.500	5.400
Flash IOPS ^{4,5}		1.000.000	1.000.000	500.000	225.000
Data Load Rate ⁵		5 TB/hr	5 TB/hr	2.5 TB/hr	1 TB/hr

1. Rohdatenkapazität; 1GB=109 bytes, 1TB=1012 bytes
2. Benutzerdatenkapazität ohne Komprimierung unter Berücksichtigung Spiegelung (ASM normale Redundanz) und unter Abzug von Speicherplatz für Temp, Log, Undo und Indizes. Tatsächliche Kapazität hängt von der Anwendung ab. 1TB=240bytes
3. Spitzendurchsatz ohne Komprimierung
4. IOPS auf der Basis von 8K I/O Requests
5. Tatsächliche Performance hängt von der Anwendung ab

- Erweiterungen beim Offloading (OR-Prädikate, IN-Listen)
- Paralleles Insert-Select mit Order-By-Klausel sortiert die Daten beim Laden so, dass Storage-Indizes bei Abfragen besseren Nutzen bringen
- Im Resource Management werden feste Grenzwerte für CPU bzw. I/O je Consumer Group gesetzt, ebenso gelten nun Grenzwerte beim Parallel Statement Queuing je Consumer Group

Fazit

Die Tabelle 1 zeigt zusammenfassend die technischen Kenngrößen der verschiedenen Modelle auf. Mit den neuen Modellen der Exadata Database Machine stehen für jedes Einsatzszenario, sei es im Data-Warehouse-, OLTP- oder Konsolidierungsbetrieb, hochperformante Alternativen zur Verfügung. Das neue Modell X2-8 erweitert das Portfolio und bietet Kunden mit besonders hohen Anforderungen im Multi-Rack-Betrieb eine Lösung. Die in den Maschinen eingesetzten neuen Software-Versionen des Exadata Storage Servers 11.2.2.1 und der Oracle-Datenbank 11.2.0.2 helfen dabei, die Möglichkeiten der Exadata-Technologie noch besser auszunutzen.

Weiterführende Informationen

- <http://www.oracle.com/technetwork/database/exadata/index.html?ssSourceSiteId=ocomen>
- <http://www.oracle.com/us/technologies/linux/unbreakable-enterprise-kernel-ds-173416.pdf>
- <http://www.oracle.com/asr/index.html>

Kontakt:

Frank Schneede
frank.schneede@oracle.com

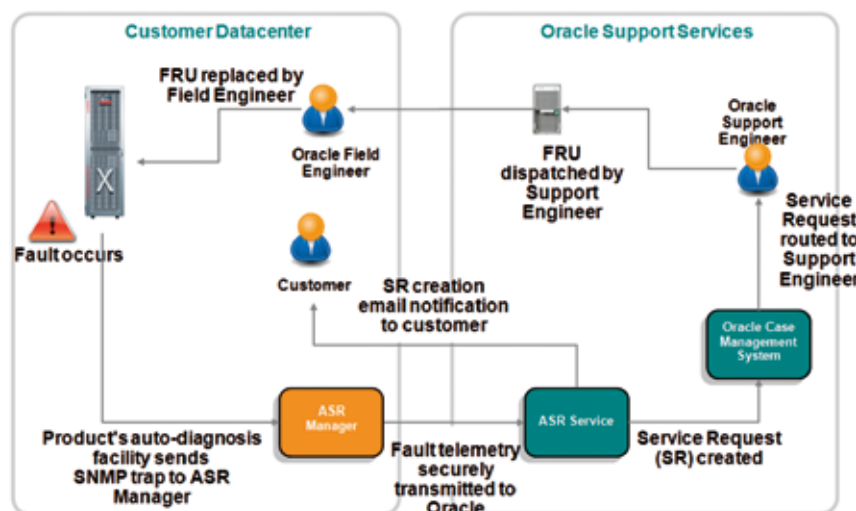


Abbildung 2: Automatischer Service Request (ASR)