

Exadata – aus "alt" mach "neu"!

Frank Schneede
Oracle Deutschland B. V. & Co.KG
Hannover

Schlüsselworte

Exadata, Upgrade, Database Machine, X3-2, X3-8, Datawarehouse, OLTP, Konsolidierung

Einleitung

Die Exadata Database Machine wurde als erstes einer Reihe von sogenannten Engineered Systems auf der Oracle Open World 2008 vorgestellt. Anfangs noch skeptisch als Oracles erster Schritt ins Hardware-Geschäft beäugt, erkannten die Kunden schnell die Vorzüge einer solchen vertikal integrierten Lösung. Regelmäßig wurde die Exadata Hardware in den Folgejahren aktualisiert und neue Funktionen wurden in der Exadata Software implementiert. So wurde auf der diesjährigen Oracle Open World bereits die vierte Generation der Exadata Database Machine vorgestellt.

Die ersten Exadata Kunden stehen mittlerweile vor der Herausforderung, ihre Exadata Maschinen erweitern oder durch neue Modelle ablösen zu müssen. Die Exadata Technologie ist so ausgelegt, dass eine Erweiterung älterer Modelle mit neuen Komponenten problemlos möglich ist. Neue Versionen der Exadata Software laufen ebenfalls auf älteren Modellen.

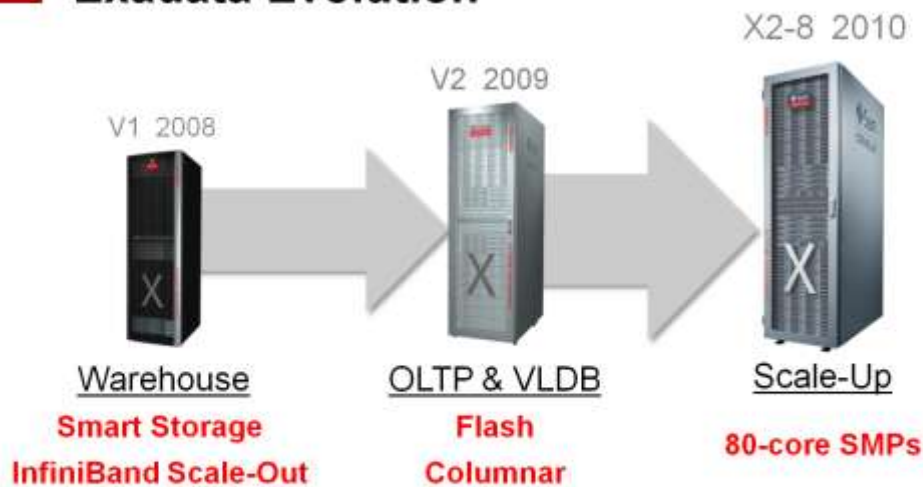
Dieser Vortrag stellt die neue Exadata X3 vor und zeigt die Möglichkeiten des Zusammenspiels alter und neuer Exadata Komponenten auf.

Exadata – ein Blick zurück

Die Exadata Database Machine V1 war das erste Modell der Exadata Produktfamilie. Mit einer komplett vorkonfigurierten und standardisierten Lösung bestehend aus Datenbank Servern und Storage Servern, die mit schnellen InfiniBand Netzwerk verbunden sind, richtete sich das Angebot primär an Datawarehouse Kunden. Datawarehouse Systeme zeichnen sich erfahrungsgemäß durch hohe Anforderungen an die I/O Bandbreite aus, die in herkömmlichen Architekturen nur mit sehr hohem Aufwand umgesetzt werden können. Insbesondere im Laufe des Lebenszyklus eines wachsenden Datawarehouse haben Systeme oft mit zunehmenden Performanceproblemen zu kämpfen. Eine Lösung dieses oft als „*Bandbreiten-Dilemma*“ bezeichneten Phänomens bietet die Exadata Database Machine.

Die Exadata ist als „*well-balanced system*“ so aufgebaut, dass sich eingesetzte CPU-, Memory-, I/O- und Netzwerkressourcen immer in einem optimalen Verhältnis befinden und kein Engpass die Systemperformance beeinträchtigt. Das gilt insbesondere auch bei Aufrüstungen bzw. Erweiterungen der Maschine. Die leistungsstarke Hardware der Exadata Database Machine wird zusätzlich unterstützt durch die Exadata Storage Server Software. Der Storage Server übernimmt einen Teil der Aufgaben, die in herkömmlichen Systemen durch Datenbankserver geleistet werden. Datenintensive Aufgaben wie das Scannen und Filtern großer Datenmengen werden im Storage Grid ausgeführt. Diese Operationen werden als „*Smart Scan*“ bezeichnet und stellen den grössten Innovationsschritt in der ersten Generation der Exadata Software dar.

Exadata Evolution



ORACLE

Abb. 1: Exadata Evolution

Die im Jahr 2009 vorgestellte zweite Generation der Exadata bestand erstmals vollständig aus eigenen Sun/Oracle Komponenten. Sie erweiterte das Einsatzspektrum der Maschine auf OLTP- und VLDB-Systeme. Das wurde möglich durch die Exadata Smart Flash Cache Technologie. Geeignete Daten werden transparent für die Anwendung auf schnellen Flash Speicherkarten abgelegt und können so sehr performant verarbeitet werden. Die Flash Technologie der Exadata hilft dadurch, Random I/O Operationen zu beschleunigen.

Die neu entwickelte Exadata Hybrid Columnar Compression ist mit der Exadata Database Machine V2 vorgestellt worden. Durch die Kombination einer spaltenorientierten Speicherung mit dem bislang ausschließlich verwendeten zeilenorientierten Speicherformat können sehr gute Komprimierungsfaktoren erreicht werden. Vier unterschiedliche Algorithmen tragen dazu bei, einen Kompromiss zwischen hoher Komprimierung und Rechenlast durch Komprimieren bzw. Dekomprimieren zu finden, der für die meisten Anwendungsfälle geeignet ist. Mit der Hybrid Columnar Compression – kurz HCC genannt – können bis zu 15 mal mehr Daten auf der Exadata Database Machine gespeichert werden, als dies ohne Komprimierung möglich ist. Dadurch können auch größte Datenbanken (VLDB) problemlos auf einer Exadata Database Machine betrieben werden.

Auf der Oracle Open World 2010 wurde die Exadata X2 Familie vorgestellt. Neben der Aktualisierung der Hardware auf neue Prozessoren und Verbesserungen der Exadata Storage Server Software wurde das neue Top-Modell der Exadata X2 Familie eingeführt. In der Exadata Database Machine X2-8 Full Rack werden zwei Datenbankserver eingesetzt, die mit jeweils 2 Intel 8-Core Prozessoren ausgestattet sind und über jeweils 1TB Hauptspeicher verfügen. Mit diesem Leistungsangebot können auf Exadata Database Machine X2-8 große Datenbank Konsolidierungen umgesetzt werden. Ebenso ist die Exadata X2-8 für den Betrieb in mehreren voll ausgebauten Racks prädestiniert.

Das Exadata Angebot bestehend aus der Exadata X2-2 Quarter, Half und Full Rack sowie der Exadata X2-8 wurde dann Anfang 2011 durch die Einführung der Exadata Storage Expansion Racks, die es ebenfalls als Quarter, Half oder Full Rack gibt, erweitert. Mit den Storage Expansion Racks, die lediglich Storage Server und InfiniBand Netzwerkkomponenten aber keine Datenbankserver beinhalten, können Kunden ihre Exadata nun praktisch unbegrenzt erweitern. Das ist vor allem dann sinnvoll, wenn zum Beispiel eine Information Lifecycle Management (ILM) Strategie umgesetzt werden soll, viele unstrukturierte Daten verarbeitet werden müssen oder zusätzlicher Platz für die Anfertigung von Backups bereitgestellt werden muss.

OOW 2012: Vorstellung der Exadata X3 Familie

Seit der Vorstellung der Exadata X2 Familie hat es immer wieder kleinere Aktualisierungen und Verbesserungen der Exadata Produktfamilie gegeben. So wurde Ende 2011 eine Aktualisierung der Exadata X2-8 auf Intel 10-Core Prozessoren vorgenommen, die in SandyBridge Mikroprozessorarchitektur aufgebaut sind. Gleichzeitig wurde die Größe der High Capacity Festplatten von 2TB auf 3TB erhöht.

Auf der diesjährigen OOW wurde nun der nächste Meilenstein in der Exadata Historie vorgestellt: die Exadata X3.

Exadata X3 Details

- Größer
 - Architektur wie Exadata X2
 - Preis wie Exadata X2
- Schneller

Alle X3 Database Maschinen

4X mehr Flash Memory	22 TB Flash Memory pro Rack
20X mehr Schreib-Performance	Exadata Smart Flash Write Cache
33% mehr I/O Bandbreite	100 GB/sec running SQL
10% bis 30% Stromersparnis	Bis 3 Kilowatt Ersparnis pro Rack

X3-2 Datenbank Server

33% schnellere Datenbank CPUs	8-Core Xeon® SandyBridge E5-2690
75% mehr Memory	1 TB bis 2 TB pro Rack
mit 10Gb Ethernet ins Data Center	40X 10Gb Ports pro Rack

ORACLE

Abb. 2: Exadata X3: Gegenüberstellung zur X2

Die bewährte Architektur bestehend aus Datenbankserver-Grid, Stageserver-Grid und InfiniBand Netzwerk bleibt auch in der neuen Exadata X3 erhalten, ebenso bleibt der Preis der Hardware unverändert.

Die Server, die in den neuen Exadata X3 Modellen verwendet werden, basieren nun alle auf der aktuellen SandyBridge Mikroprozessorarchitektur, die in der Exadata X2-8 bereits seit Ende 2011 verbaut wurden. In den Datenbankservern, die in der Exadata X3-2 verwendet werden, sind nun zwei

Intel Xeon ® E5-2690 Prozessoren verbaut, die über jeweils 8 Prozessorkerne verfügen, zudem wurde der **Hauptspeicher** von bisher 96GB auf **128GB** vergrößert. Die Storage Server nutzen zwei **Intel Xeon ® E5-2630L** Prozessoren mit 6 Prozessorkernen, der **Hauptspeicher** wurde von 24GB auf **64GB** vergrößert. Zusätzlich zur CPU- und Hauptspeichererweiterung stehen in der Exadata Database Machine X3-2 wesentlich mehr Ethernet Ports zum Anschluss an das kundeneigene Netzwerk zur Verfügung.

Die entscheidende Veränderung in der Exadata X3 befindet sich allerdings im Storage Server. Dieser verfügt nämlich nun mit **1,6TB Exadata Smart Flash Cache** über eine mehr als vier mal so große Kapazität wie das Vorgängermodell! Es werden in einem Exadata Storage Server X3-2 vier neue **Sun Flash Accelerator F40 PCIe** Karten verwendet, die mit eMLC (enterprise Multi-Level Cell) NAND Modulen bestückt sind. Die neuen F40 Flash Karten haben eine geringere Latenz und unterstützen auf diese Weise die hohen Performanceanforderungen I/O intensiver OLTP Anwendungen. Mit insgesamt 22,4TB verfügbarem Flash Cache kann in einem voll ausgebauten Rack praktisch eine ganze Datenbank im Flash Cache gehalten werden!

Write-Back Flash Cache

Der vergrößerte Exadata Smart Flash Cache kann mit der Exadata X3 in zwei unterschiedlichen Modi betrieben werden. Der bislang verwendete *WriteThrough* Modus ist defaultmäßig aktiviert und unterstützt in erster Linie Leseoperationen. Für Anwendungen mit hoher Schreiblast kann der Flash Cache in den Modus *WriteBack* versetzt werden. Dazu muss der Flash Cache entfernt und neu angelegt werden, was jedoch im laufenden Betrieb geschehen kann:

```
CellCLI> drop flashcache
Flash cache cel14_FLASHCACHE successfully dropped

CellCLI> alter cell shutdown services cellsrv
Stopping CELLSRV services...
The SHUTDOWN of CELLSRV services was successful.

CellCLI> alter cell flashCacheMode = WriteBack
Cell cel14 successfully altered

CellCLI> alter cell startup services cellsrv
Starting CELLSRV services...
The STARTUP of CELLSRV services was successful.

CellCLI> create flashcache all
Flash cache cel14_FLASHCACHE successfully created
```

Im Modus *WriteBack* ist der Cache persistent, das heißt auch bei einem Neustart des Storage Servers oder des Gesamtsystems geht der Inhalt des Cache nicht verloren. Die Verwaltung des Flash Cache erfolgt im Hauptspeicher des Storage Servers, der, wie oben bereits ausgeführt, aus diesem Grunde stark vergrößert worden ist. Die Bedienung der Datenbank kann praktisch immer aus dem Cache heraus erfolgen, ein Schreiben auf die Platte erfolgt unter Umständen erst mit großer Verzögerung.

Voraussetzung für die Nutzung des Write-Back Flash Cache ist **Exadata Storage Server 11.2.3.2.0** oder jünger und die Aktualisierung des ASM mit **Datenbank Patch für Exadata 11.2.0.3 BP 9** oder jünger. Wenn diese Bedingungen vorliegen, kann der Write-Back Flash Cache sogar auf den alten Exadata Modellen V2 oder X2 genutzt werden. Natürlich sind die zu erwartenden

Performancesteigerungen nicht so hoch, da die in diesen Modellen verbauten Flash Accelerator F20 Karten etwas geringere Leistungsdaten haben. Immerhin sollte die Anzahl der Schreiboperationen auf den alten Maschinen noch um Faktor 10 gegenüber Schreiboperationen auf Festplatte gesteigert werden. Auf der neuen Hardware sollte sogar ein Faktor 20 erreicht werden können!

Exadata Einstiegsmodell: Eighth Rack

Die Einstiegshürde in die Exadata Technologie erschien bislang hoch; selbst bei nur teilweiser Lizenzierung der Maschine war die Exadata sowohl in Bezug auf die angebotene Leistung als auch in Hinblick auf die Beschaffungskosten für viele Kunden nicht bedarfsgerecht dimensioniert. Aus diesem Grund wurde auf der OOW 2012 ein neues Exadata Modell vorgestellt: **Exadata Database Machine X3-2 Eight Rack**:



Niedrige Kosten

NEU Exadata X3-2 Eighth Rack

Schneller als 2010 Quarter-Rack, 2008 Half-Rack

- **Exadata Einstiegskonfiguration**
 - HW Listenpreis \$200K: 60% vom Quarter Rack
 - 16 Database Cores, 54 TB Disk, 2.4 TB PCI Flash
 - HA Konfiguration mit allen Exadata Funktionen
- **Exadata Extreme Performance für kleinere Workloads, Entwicklung, Test, Disaster Recovery**
- **Hardware entspricht Quarter-Rack**
 - Hälfte der CPU Cores, Disks und Flash Cards disabled
 - Hälfte der Datenbank und Exadata Software Lizenzen
- **Upgrade auf Quarter-Rack durch Skript**

ORACLE

Abb. 3: Exadata X3-2 Eight Rack

Im Einstiegsmodell stehen alle Hochverfügbarkeitsfeatures uneingeschränkt zur Verfügung, es werden wie in einem Quarter Rack zwei Datenbankserver und drei Storage Server betrieben, die über InfiniBand Netzwerk verbunden sind. Der Unterschied liegt darin, dass die Hälfte der Systemressourcen über Software-Mechanismen abgeschaltet sind, RAC Cluster und ASM sind jedoch nach wie vor so aufgesetzt, dass kein *Single Point of Failure* existiert! Es stehen demnach also

- **16** statt 32 **Prozessorkerne** im Datenbank Grid
- **18** statt 36 **Prozessorkerne** im Storage Grid
- **10,8TB** bzw. **54TB** statt 21,6TB bzw. 108TB **Festplattenkapazität** für High Performance bzw. High Capacity Festplatten (6 von 12 Festplatten pro Storage Server sind aktiviert)
- **2,4TB** statt 4,8TB **Flash Cache** Kapazität (6 von 12 Flash Karten sind aktiviert)

in der **Exadata Database Machine X2-3 Eighth Rack** zur Verfügung.

Mit dem Einstiegsmodell ist für viele potentielle Kunden die Einstiegshürde in die Exadata Technologie gesenkt worden. Neben dem günstigeren Anschaffungspreis wird durch Verringerung der

laufenden Prozessorkerne eine Einsparung bei den Lizenzkosten für die Oracle Datenbank und Optionen (RAC, Partitioning) erreicht. Bei Bedarf kann das Eighth Rack mittels eines Skriptes im laufenden Betrieb auf ein Quarter Rack aufgerüstet werden.

Exadata – maximale Interoperabilität

Um für die Anwender ein Höchstmaß an Investitionssicherheit zu bieten, ist die Exadata Database Machine seit der Exadata V2 so ausgelegt, dass unterschiedliche Exadata Generationen problemlos miteinander betrieben werden können. Mehrere deutsche Exadata Anwender haben so bereits ihr Einstiegsmodell erweitert oder aufgerüstet. Für eine Erweiterung ohne den Einsatz zusätzlicher Datenbankserver stehen unterschiedliche Optionen bereit:

- Exadata Storage Server X3-2: Der Storage Server muss in ein kundeneigenes Rack eingebaut werden. Für den Anschluss an die Exadata Database Machine und den Betrieb der Gesamtlösung ist der Kunde verantwortlich.
- Exadata Storage Expansion X3-2 Quarter/Half/Full Rack: Die Erweiterung wird komplett vorbereitet in einem eigenen Rack ausgeliefert und durch Oracle ACS in Betrieb genommen. Für Quarter bzw. Half Racks stehen analog zur Exadata Database Machine Ausbaustufen bereit.

Eine Erweiterung mittels Exadata Storage Expansion Rack ist sinnvoll, wenn keine zusätzliche Rechenleistung im Datenbank Grid erforderlich ist. Das kann unter anderem vorkommen, wenn die eingesetzte Anwendung auf einem RAC Cluster nicht gut skaliert oder lediglich eine online Archivierung historischer Daten gewünscht ist. Bei Kunden, die ihre Exadata in ein eigenes Rack (sogenanntes „*Re-Racking*“) umgebaut haben, ist das Halfrack die größtmögliche Ausbaustufe. Danach besteht nur die Möglichkeit, mit einem weiteren Half Rack oder eben durch den beschriebenen Einsatz von Expansion Racks zu erweitern.

Falls die Erweiterung des verfügbaren Plattenplatzes allein nicht ausreicht, sondern weitere Datenbankserver für die Bewältigung zusätzlicher Aufgaben eingesetzt werden müssen, hat der Kunde die Wahl zwischen mehreren Upgrade Kits:

- Exadata Database Machine X3-2 Eighth Rack to Quarter Rack Upgrade: Diese Upgrade Variante ist die einzige, die ohne die Installation neuer Hardware auskommt. Die Erweiterung wird durch Skripte vorgenommen und kann im laufenden Betrieb durchgeführt werden.
- Exadata Database Machine X3-2 Quarter Rack to Half Rack Upgrade: Die existierende Exadata Database Machine wird um 2 Datenbankserver, 4 Stageserver und einen InfiniBand Switch erweitert. Die Durchführung des Upgrades wird durch Oracle ACS durchgeführt. Eine Downtime ist für die Anbindung der Server notwendig, die Konfiguration kann dann im laufenden Betrieb geschehen.
- Exadata Database Machine X3-2 Half Rack to Full Rack Upgrade: Die existierende Exadata Database Machine wird um 4 Datenbankserver und 7 Stageserver erweitert. Die Durchführung des Upgrades wird durch Oracle ACS durchgeführt. Eine Downtime ist für die Anbindung der Server notwendig, die Konfiguration kann dann im laufenden Betrieb geschehen. Darüberhinaus können weitere Full Racks angeschlossen werden. Auf diese Weise ist es möglich, bis zu acht Racks ohne den Einsatz zusätzlicher Hardware zu verbinden.

Die verschiedenen Upgrade Kits und Exadata Generationen können in einer Maschine verwendet werden, wie Abbildung 4 anschaulich zeigt. Es ist lediglich zu beachten, dass der in den Upgrade Kits verbaute Festplattentyp dem in der Exadata Database Machine entspricht.

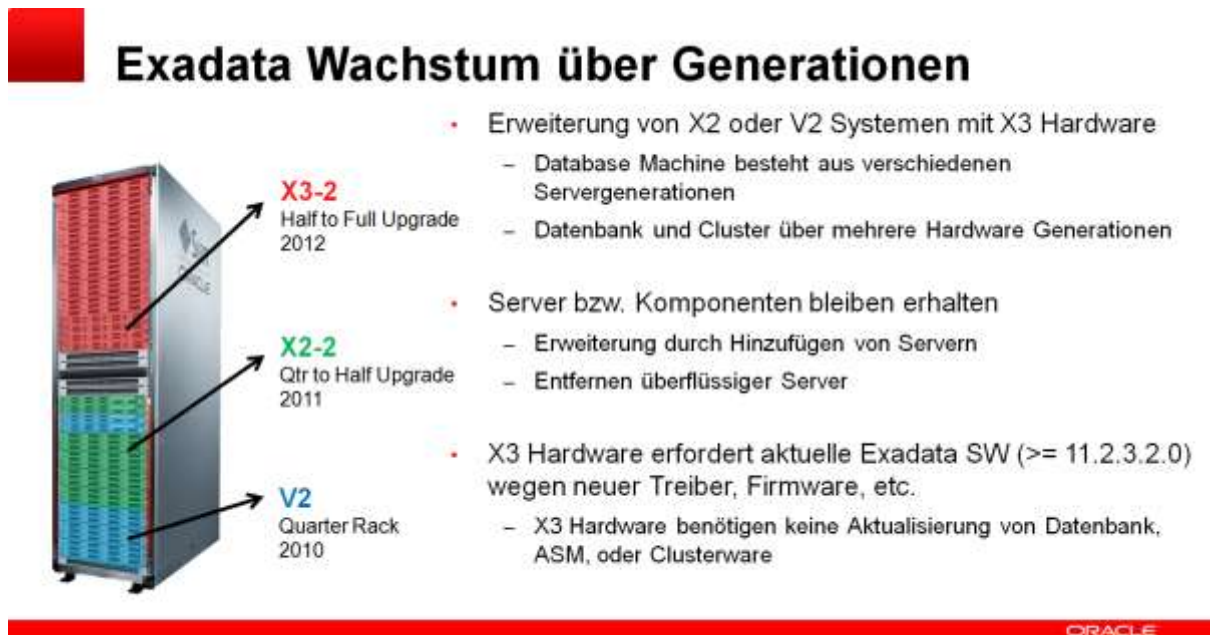


Abb. 4: Exadata X3-2 Upgrade Pfade

Eine Exadata Database Machine kann also theoretisch aus drei unterschiedlichen Maschinengenerationen bestehen. Bei dieser Systemkonfiguration gibt es allerdings einen wichtigen Punkt zu beachten: Die Maschine wird als eine Einheit in den Support genommen, der Support endet fünf Jahre (sogenanntes *EOS Datum*), nachdem das letzte Exemplar der ältesten Exadatageneration ausgeliefert worden ist! Für das in Abbildung 4 dargestellte Szenario bedeutet das, dass der Support für das gesamte System im Oktober 2015 endet. Zu diesem Zeitpunkt bestünde allerdings die Möglichkeit, die nachgerüsteten Komponenten, die nach wie vor im Support sind, in ein separates kundeneigenes Rack umzubauen und an eine andere Exadata Database Machine anzuschließen, die noch im Support ist. Zum jetzigen Zeitpunkt (November 2012) sind die letzten Exadata X2 noch in der Auslieferung, der Support für diese Exadata Modelle läuft also circa bis November 2017. Aus den oben angeführten Gründen heraus setzen die Kunden im Allgemeinen auf zwei unterschiedliche Upgrade Strategien:

1. Aufrüstung der bestehenden Exadata mit dem dazugehörigen Upgrade Kit, die dann alle gemeinsam das gleiche EOS Datum haben.
2. Kauf des neuen Exadata Modells und Nutzung der alten Produktionsumgebung als Disaster-Absicherung oder als Test- und Entwicklungsplattform

ASM Konfiguration mit Upgrade High Capacity Festplatten

Kunden, die eine Exadata Database Machine mit großen High Capacity Festplatten betreiben, müssen unter Umständen eine Besonderheit bei der Durchführung eines Upgrade beachten. High Capacity Festplatten wurden seit der Vorstellung der Exadata V2 im September 2009 bis Januar 2012 mit einer Kapazität von 2TB ausgeliefert. Seit Januar 2012 beträgt die Kapazität jedoch 3TB. Wenn der Kunde

also in 06/2012 ein Upgrade seiner Exadata Database Machine Quarter Rack X2-2 (Auslieferung 03/2011) durchführen möchte, so ist es notwendig, die Diskgruppen im ASM entsprechend zu konfigurieren. Auf diese Weise entsteht in den vier Storage Servern, die mit dem Upgrade zur bestehenden Exadata Database Machine hinzugefügt werden, eine weitere Diskgroup. Abbildung 5 zeigt das ASM Layout schematisch.

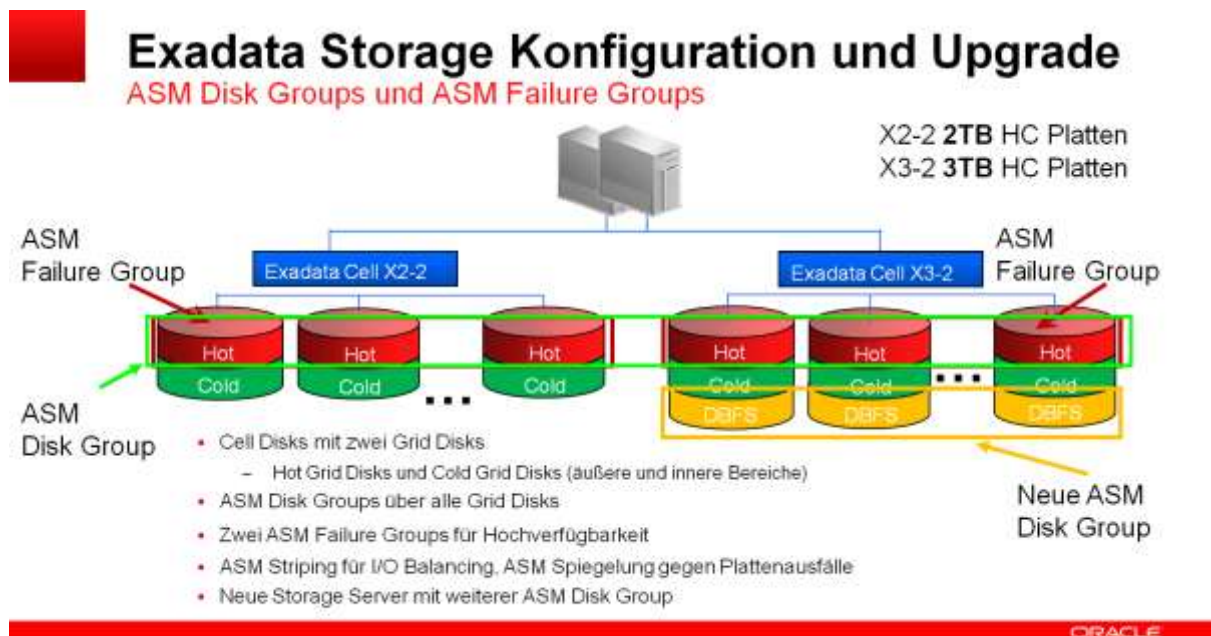


Abb. 5: Exadata ASM Layout bei Upgrade mit 3TB High Capacity Festplatten

Die neu entstandene Disk Group kann für verschiedene Zwecke genutzt werden. Eine Verwendung als Database Filesystem (DBFS) ist ebenso denkbar wie eine Nutzung als Disk Group für eine weitere Anwendung, die auf die erweiterte Exadata Database Machine gebracht werden kann.

Fazit

Mit der neuen Exadata Database Machine X3 wurde das Achtel Rack vorgestellt, mit dem Kunden den kostengünstigsten Einstieg in die Nutzung der Exadata Technologie vornehmen können. Durch die unterschiedlichen Möglichkeiten der Erweiterbarkeit der Exadata Database Machine auch über Exadata Generationen hinweg, ist bedarfsgerechtes Wachstum und gleichzeitig Investitionssicherheit sichergestellt. Viele Exadata Anwender der ersten Modellgenerationen haben bereits ihre Umgebungen auf die dargestellte Weise erweitert.

Mit der Exadata WriteBack Flash Technologie wird das Einsatzspektrum der Exadata um sehr schreibintensive OLTP Anwendungen erweitert. Die Verwendung des WriteBack Flash Cache ist auch mit älteren Exadata Modellen möglich.

Weiterführende Informationen:

- <http://www.oracle.com/us/products/database/exadata/overview/index.html>
- <http://www.oracle.com/us/products/database/exadata-db-machine-x3-2-1851253.pdf>
- <http://www.oracle.com/us/products/database/exadata-tech-overview-wp-1867437.pdf>
- <http://www.gartner.com/technology/reprints.do?id=1-1CEIIXJ&ct=121005&st=sb>
- <http://www.oracle.com/us/products/database/exadata-reference-booklet-400018.pdf>

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