

Patching Exadata Database Machine – the DBA gets it all

**Konrad Häfeli
Trivadis AG
CH-3014 Bern**

Schlüsselworte:

Oracle Database Machine, Exadata, Patching, Lifecycle Management

Einleitung

Sie haben schon eine Exadata Database Machine im Haus oder sind am Evaluieren? Dann haben sie sich sicher auch überlegt wie sie das Patchen der Systeme machen könnten. Infiniband Switches, Storage Cells, Database Servers mit ihren verschiedenen Funktionalitäten benötigen von Firmware über OS, bis Cluster und Datenbanksoftware, alles was auch in irgendeiner Form über die Zeit gepatched oder upgraded werden muss. Die Kompatibilitätsextrixen und Checklisten sind vorhanden, kann das aber wirklich alles der DBA machen oder braucht er noch Unterstützung?

In diesem Vortrag wird in einem Erfahrungsbericht aus einem Kundenprojekt aufgezeigt, wie eine Exadata V2 auf den aktuellen Softwarestand einer X2-2 gebracht wurde und welches die Erkenntnisse für den Lifecycle einer solchen Plattform sind.

System/Patch-Komponenten

Die Database Machine ist weder eine Appliance noch eine Black-Box. Sie ist ein vorkonfiguriertes System, welches nach der Lieferung auf die jeweiligen Kundenbedürfnisse angepasst werden muss (Anzahl und Grösse der Datenbanken, erstellen der DB-Services und der Clusterkonfiguration). Ebenso müssen die Daten geladen werden, entweder aus einem bestehenden System (z.B. datapump Import) oder durch Ladeprozesse in einer Datawarehouse Umgebung.

Die Frage wer verantwortlich ist für das Patchen und Upgraden des Systems ist schnell beantwortet. Es ist der Betreiber. Also Sie! Oracle hat für die recht komplexen Exadata Umgebung einige Neuerungen im Bereich Patch-Management versprochen und eingeführt. Welche Skills es braucht um diese Administrationsarbeiten durchzuführen wird nirgends genau beschrieben, schauen wir doch mal was da alles gemacht werden muss.

Eine Exadata Database Machine besteht aus folgenden Komponenten:

- Database Server
 - RDBMS und Grid Infrastructure Software-Homes
 - ➔ **Bundle Patches:** Diese werden monatlich (ab Oktober 2011, alle zwei Monate) herausgegeben. Sie beinhalten den aktuellen PSU (Patch Set Update) welcher selbst wieder den aktuellen CPU (Critical Patch Update) beinhaltet. Bundle Patches sind cumulativ und werden mit dem OPatch utility installiert.
- Exadata Storage Server
 - Operating System
 - Firmware
 - ➔ **Storage Server Patches:** Werden als „single downloadable“ File zur Verfügung gestellt und beinhalten Updates für das OS, den Kernel, Infiniband-Treiber, ILOM, Firmware und auch neue Funktionalitäten im Bereich Exadata Storage. Der Download beinhaltet auch das „minimal OP pack“ für den Datenbank Server, damit diese kompatibel zum Storage Server bleibt. Diese Patches werden quartalsweise herausgegeben und werden mit einer Utility namens „patchmgr“ installiert. Dies kann „rolling“ oder „no rolling“ geschehen.
- Infiniband
 - Switches
 - ➔ **Infiniband Switch Patches:** Diese beinhalten nur Oracle Anpassungen. Anforderungen und Instruktionen sind in der zentralen MOS Note 888828.1 (Abb. 1) aufgeführt. Sie kommen ein- bis zwei-mal jährlich heraus. Die Installationsart hängt von der jeweiligen Version ab.
- KVM und Cisco
 - ➔ Keine Oracle Patches

Die zentrale Patch-Informationen MOS Note ist die Nummer: 888828.1 (Abb. 1).

Da sind nebst den Guidelines für das Patching auch „Critical Issues“, aktuelle Versionen und die Kompatibilität untereinander beschrieben.

Database Machine and Exadata Storage Server 11g Release 2 (11.2) Supported Versions (Doc ID 888828.1)

Modified 25-MAY-2011 Type REFERENCE Status PUBLISHED(EXTERNAL) Priority 3 [To Bottom](#)

In this Document

- [Purpose](#)
- [Scope](#)
- [Database Machine and Exadata Storage Server 11g Release 2 \(11.2\) Supported Versions](#)
- [Exadata Patching Overview and Guidelines](#)
- [Critical Issues](#)
- Current Versions**
- [Version Compatibility](#)
- [Database Server Operating System](#)
- [InfiniBand OFED Software](#)
- [InfiniBand HCA](#)
- [Enumerated Patch List](#)
- [Patch Release History for Exadata Database Machine Components](#)
- [Current Version Table Revision History](#)
- [References](#)

Abb. 1: My Oracle Support Note: 888828.1

Patch Methoden

Patchen in diesem komplexen Umfeld war immer ein wenig schwierig. Über die verschiedenen Version hat Oracle neuen Funktionalitäten eingeführt um das einfacher, automatischer zu machen. Die Abbildung 2 zeigt die Neuerungen und die Version der Einführung:

Feature	Minimum Version
Cell Rolling Apply	11.2.1.3.1
Bundle Patch Merge	11.2.0.2
RAC Rolling Installable	11.2.0.1 GI_BP1 11.2.0.1 DB_BP9
OPatch Auto Installable	11.2.0.1 GI_BP4 11.2.0.2 BP2
EM Installable	11.2.0.1 DB_BP7 11.2.0.2 BP1
Dataguard Standby-First Installable	11.2.0.1 DB_BP8 11.2.0.2 BP1
OPlan	11.2.0.2 BP2

Abb. 2: Neue Funktionalität über die Versionen

Der Rolling Bundle Patch auf dem Datenbankserver kann seit der Version 11.2.0.2 mit der „opatch auto“ Option gestartet werden. Diese ermöglicht ein Einzelkommando-Patchen pro Datenbank-Knoten und beinhaltet das aktualisieren des RDBMS- und Grid-Homes. Dabei werden die Datenbank Instanzen heruntergefahren und die laufenden Services auf den oder die restlichen Knoten verteilt. Bei

den Storage Zellen updates muss entschieden werden ob online oder offline gearbeitet werden soll. Folgende Überlegungen könne da gemacht werden:

Rolling:

- Pro:
 - Keine Ausfallzeiten
 - Im Fehlerfall nur eine Zelle betroffen
- Contra
 - Lange Patchapplikationszeit (es wird sequentiell gearbeitet)
 - Keine oder weniger Redundanz während dem applizieren
 - 1 ½ bis 3 Stunden Aufwand pro Zelle (bei einem Full-rack mit 14 Zellen dauert das...)

Non Rolling:

- Pro
 - 1 ½ bis 3 Stunden Aufwand für alles
 - Reduziertes Patch-Zeitfenster
- Contra
 - System-Ausfall während dem applizieren
 - Im Fehlerfall sind alle Zellen betroffen

Patch-Administrator/Manager

Die Auflistung der zu patchenden Komponenten zeigt, dass das über den bisher üblichen Horizont eines Datenbank Administrators geht. Dazu hat man in der Vergangenheit immer einen ausgebildeten System Administrator herangezogen. Durch die mögliche Rollen-Trennung auf RDBMS und Grid-Infrastruktur Ebenen (Cluster und ASM), ist eine Aufteilung der Aufgaben schon vorbereitet worden. Oracle will mit dem Einsatz von Exadata unter anderem auch Betriebskosteneinsparungen möglich machen. Da stell ich nun die Frage:

- Braucht es wirklich einen System-Administrator für eine Exadata Umgebung?

Was hätte er zu tun?

- OS Unterhalt, nicht nur das minimal OS pack auf den DB-Servern
- Usermanagement (personalisierte User erstellen, sudo Konfigurationen)
- Crontab setup/enabling
- Sendmail Konfigurationen
- Netzwerk Konfigurationen (Anpassungen)
- ...

Eine Kombination der Aufgaben wäre möglich, d.h. auf jeden Fall eine Aufgabenbereicherung für den jeweiligen Administrator. Egal ob es ein Sys-Admin ist der sich in die DBA Aufgaben einarbeitet oder ob es ein DBA ist der System Administration macht. Im Exadata Umfeld werden die Standardaufgaben mittels Checklisten (MOS Notes) von Oracle abgearbeitet, das ist sich der DBA sicher gewohnt. Grosse Patchaufgaben geht man optimalerweise im Team an. Zu zweit ist die akribische Abarbeitung der Aufgaben mit dem 4-Augen-Prinzip erfolgsversprechender.

Exadata Patching V2 auf X2-2

Das Kundensystem war ein MAA (Maximum Availability Architecture) Quarter-Rack Setup. D.h. die Systeme sind nicht nur in sich redundant sondern auch noch über eine zweite Lokation hinweg mit Dataguard gespiegelt. Diese Konfiguration ist natürlich das besten in Sachen Disaster-Recovery, bringt aber auch sehr viele Vorteile im Patchen mit sich. Die Hardware ist eine V2 und die Softwarebasis stammt noch aus der Oracle-Konfiguration des Kundensetups von Januar 2010. Da das System in den 15 Monaten nie gepatched wurde war da natürlich grosser Nachholbedarf. Um somehr, dass die Quarterracks nun auf je ein Half-Rack erweitert werden sollten, was natürlich nur möglich ist, wenn das alte System denselben Patchstand hat wie das neue, welches auf der Basis X2-2 ist. Ebenfalls ist in der Kompatibilitäts-Matrix ersichtlich, dass für eine 11.2.0.2 Datenbank im Minimum der Storage-Server auf Version 11.2.2.x sein muss (die Ähnlichkeit der Nummerierung soll nicht stören, sie haben nichts miteinander zu tun). Jedenfalls zeichnet sich da ein recht aufwändiges und komplexes Patchszenario ab.

Wie geht man nun optimalerweise vor? Klar, Informationsbeschaffung, von Leuten die das evtl. schon gemacht haben. Da es solche nicht allzu viele gibt, und der Autor selbst schon das Patching einer V2 im Exadata Competence Center von Trivadis gemacht hat, war die eigene Erfahrung die wichtigste. Die hat aber auch klar gezeigt, dass das genaue Studium und die Einhaltung der MOS Notes Schritte am erfolversprechendsten ist. Auch hier ist 888828.1 der Einstiegspunkt zu jeglicher relevanten Information.

Infiniband Switch

Als erstes sollen die Infiniband Switches gepatched werden. Das eruiere der aktuellen Versionen bringt Klarheit über das anzuwendende Vorgehen:

```
# nm2version
NM2-36p version: 1.0.1-1
Build time: Sep 14 2009 12:52:51
ComExpress info:
Manufacturing Date: 2009.02.19
Serial Number: "NCD2T0059"
Hardware Revision: 0x0006
Firmware Revision: 0x0102
```

Nebst dem Exadata patch 12373676 für die Infiniband switch software 1.3.3-2 muss noch ein zusätzlicher Infiniband NM2-36p payload patch 11891229 appliziert werden. Diese Anwendung die dann eher nicht DBA Arbeit, ist bei Befolgung der Anleitung durchaus machbar.

Storage Server

Die MOS Note 1323958.1 informiert uns, dass mit dem Patch 12577723 die Exadata Storage Server software 11.2.2.3.2 heruntergeladen werden kann. Ein Check der Critical Issues, Note 1270094.1, versichert uns, dass die bekannten Probleme mit der Version 11.2.2.2 gefixt sind, was natürlich nicht heisst, dass nicht noch unbekannt Probleme auftreten können. Wichtig ist das System vor und nach dem Anpassen mittels Health-Check Scripts (MOS Note: 1070954.1) auf Best Practices zu überprüfen, damit Konfigurations-Abweichungen einfach festgestellt werden können. Auch bei diesen Scripts gibt es immer wieder Neuerungen. Aktuell ist es „exachk“ welches das HealthCheck ablöst (dies ist nur noch aus Rückwärtskompatibilität zu den HP basierten V1 Systemen vorhanden). Aufruf und ein kleiner Teil des Outputs sieht so aus:

```
# oracle@bdm2db01:~/exachk/ [EXATEST1] ./exachk -a
```

```

.
=====
Node name - bdm2db02
=====
Collecting - CPU Information
Collecting - CRS active version
Collecting - CRS oifcfg
Collecting - CRS software version
Collecting - Cluster interconnect (clusterware)
Collecting - Compute node PCI bus slot speed for infiniband HCAs Collecting
- Exadata storage cells [DBMV2]
Collecting - Kernel parameters

```

Beim Vorhandensein einer Site-redundanten Umgebung können nun alle Primary-DBs auf eine Location gelegt werden, damit das Standby System für den Upgrade bereit ist. Die Abfrage der aktuellen Version mit dem Kommando „imageinfo“ bringt den Status hervor.

```

[root@bdm1cel01 ~]# imageinfo

Kernel version: 2.6.18-128.1.16.0.1.el5 #1 SMP Tue Jun 30 16:48:30 EDT 2009
x86_64
Cell version: OSS_11.2.1.2.1_LINUX.X64_100131
Cell rpm version: cell-11.2.1.2.1_LINUX.X64_100131-1

Active image version: 11.2.1.2.1
Active image activated: 2010-03-01 16:24:56 +0100
Active image status: success
Active system partition on device: /dev/md5
Active software partition on device: /dev/md7

In partition rollback to 11.2.1.2.0: Possible

Cell boot usb partition: /dev/sdm1
Cell boot usb version: 11.2.1.2.1

Inactive image version: undefined
Rollback to the inactive partitions: Impossible

```

Danach kann aus dem Directory wo die Patchsoftware liegt (Stage-Depot muss auf einem DB-Server sein) die Voraussetzungen für den Patch geprüft werden:

```
# ./patchmgr -cells cell_group -patch_check_prereq
```

Und mittels des dcli utility (distributed command line interface) die Umgebung für den Patch vorbereitet werden. D.h. Stoppen des Clusters auf allen DB-Knoten und Stoppen der Cell-Server auf allen Storage-Zellen:

```

dcli -g dbs_group -l root \
    "/u01/app/11.2.0/grid/bin/crsctl stop crs -f"

dcli -g dbs_group -l root "ps -ef | grep grid"

dcli -g cell_group -l root \

```

```
"cellcli -e alter cell shutdown services all"
```

Danach kommt das effective Patchen der Zellen mit dem patchmgr:

```
# ./patchmgr -cells cell_group -patch
.
14:33-25-May:2011      :Working: DO: Check cells have ssh equivalence for
root user. Up to 10 seconds per cell ...
.
.
16:14-25-May:2011 5 of 5 :Working: DO: Check the state of patch on cells.
Up to 5 minutes ...
16:14-25-May:2011 5 of 5 :SUCCESS: DONE: Check the state of patch on cells.
```

Das dauert dann ca 1 ¼ Stunden und kann dann mit dem cleanup Kommando beendet werden. Dieser Vorgang upgradet auch das Betriebssystem auf die aktuelle Version. Hier war das:

```
# cat /etc/redhat-release
Enterprise Linux Enterprise Linux Server release 5.5 (Carthage)
```

DB Server minimal Pack

Der Storage Cell Upgrade ist noch nicht fertig, es müssen noch die DB-Knoten gepatcht werden, damit sie die neuen Storage-Server verstehen. Der Patch muss über die Console des ILOM durchgeführt werden. Als erstes müssen die Cluster-Services und der oswatcher gestoppt werden. Danach kann aus dem Patchdirectory (unzip db_patch_11.2.2.3.2.110520.zip) das install Script aufgerufen werden:

```
./install.sh -force
```

Je nachdem wie gut „Murphy“ drauf ist, funktioniert das bestens und nach dem reboot kommt an der Console der Server wieder zum Vorschein. In unserem Fall hat es leider nicht so gut geklappt, denn ein Server hatte einen defekten Diskkontroller, was natürlich gerade bei der Remote-Site in 80km Entfernung passieren musste, sodass wir mit dem System nur „halb“ hochfahren konnten und danach den Hardware-Support aufbieten mussten.

DB Server Upgrade

Einen DB-Server mit RDBMS und Grid Installations upzugraden ist dann schon das Kerngeschäft eines DBAs. Dies hat dann auch gut funktioniert, war aber meilenweit von einem Single-Button Patch entfernt. Die MOS Note 1315926.1 beschreibt den Ablauf der folgende Punkte beinhaltet:

- Vorbereitung des existierenden Environment
 - Es muss minimalbedingungen erfüllen bevor auf 11.2.0.2 upgradet werden kann
- Installation und Upgrade Grid Infrastructure auf 11.2.0.2
 - Immer als RAC rolling upgrade
- Installation Database 11.2.0.2 Software
 - In ein neues ORACLE_HOME Directory (lässt das alte ohne Änderung)
- Upgrade der Database to 11.2.0.2
 - Ergibt DB-weite Downtime

OS Upgrade des DB-Servers

Auch für diesen Ablauf gibt es eine beschreibenden MOS Note: 1284070.1. Es gibt 6 Schritte einzuhalten:

- Step 1: Obtain RPM bundle, stage needed RPMs, apply workaround(s)
- Step 2: Install updated kernel
- Step 3: Shutdown Services
- Step 4: Update additional RPMs
- Step 5: Update additional packages
- Step 6: Clean up, restart processes, healthcheck

Danach ist das OS von Oracle Linux 5.3 auf 5.5 angehoben.

Fazit

Auch sogenannte „engineered systems“ benötigen software maintenance. Damit der Aufwand in Grenzen gehalten werden kann, soll regelmässig, ein guter Intervall ist quartalsweise, das komplette System gepatched werden. Dieser Ablauf soll im Lifecycle-Management als Prozess definiert und beschrieben werden, dazu gehören Checklisten, damit die vielen Schritte ohne Fehler abgearbeitet werden können. Regelmässige Health-Checks mit den vorhandenen resp. dem einen oder anderen selbst geschriebenen Skript gibt die Gewissheit, dass das System in Ordnung ist. Die Patch-Prozesse müssen auch regelmässig den neuen Funktionalitäten angepasst werden, da diese immer besser und einfacher werden. Eine grosse Hilfe ist auch die Konfiguration eines MAA Systems, damit lässt sich der Patch-Stress und die Dauer reduzieren, da komplette Redundanz vorhanden ist. Diese Szenario das hier beschrieben wurde ist mit Sicherheit eines der komplizierteren, nichts destotrotz bereiten Sie ihres seriös vor, testen es aus und berücksichtigen sie auch Murphy, denn es können nebst Software-Fehler auch Hardware-Probleme auftreten, die mit Sicherheit im dümmsten Moment kommen.

Die Frage war noch wer nun das Patchen machen soll? Ich meine der DBA kann das sehr gut machen, wenn er Interesse hat über den Tellerrand zu schauen und neue Aufgaben anzugehen.

Ich wünsche ihnen viel Erfolg und Spass beim Lifecycle-Management ihres Exadata Systems

Kontaktadresse:

Konrad Häfeli

Trvadis AG
Papiermühlestrasse 73
CH-3014 Bern

Telefon: +41(0)31-928 09 60
Fax: +41(0)31-928 09 64
E-Mail: konrad.haefeli@trivadis.com
Internet: www.trivadis.com