

Exadata Upgrade auf 12.2 – Upgrade mit Herausforderungen

Timo Giese
Fiducia & GAD IT AG
Karlsruhe

Schlüsselworte

Exadata, RAC, OVM, Virtualisierung, GridInfrastructure, 12.2, Patching

Einleitung

Im Zuge der Einführung unserer neuen virtualisierten Exadata X6-2 Systeme haben wir uns zum Ziel gesetzt, diese vor Produktivsetzung noch einem Upgrade zu unterziehen. Bei der Planung für das Upgrade ergab sich die Möglichkeit gleich auf das Exadata Release 12.2 zu wechseln, anstatt ein Upgrade auf das aktuelle 12.1er Release durchzuführen. Nach intensiver Diskussion der zu verwendenden Version haben wir uns entschieden das Upgrade auf das neue 12.2er Release zu wagen, trotz der Gefahr das Probleme beim Upgrade und im Betrieb auftreten können. Im Nachfolgendem werden die Schritte zur Vorbereitung, Durchführung und einzelnen Herausforderungen während des Upgrade Prozesses dargestellt.

Ist-Situation

Die zu patchende Exadata Infrastruktur ist wie folgt aufgebaut:

2 Eighth-Rack Exadatas bilden einen Dataguardverbund

jede Exadata ist virtualisiert und mit einer virtuellen Maschine pro Computernote bestückt

Die aktuelle Patchversion ist 12.1.2.3.2.160721

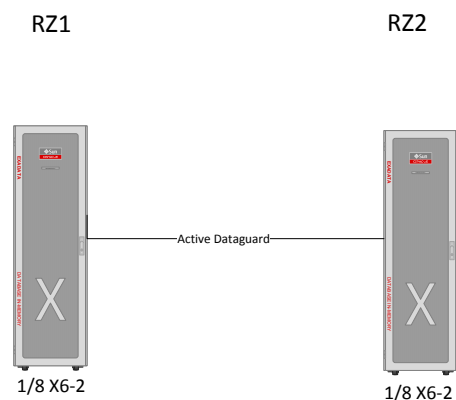


Abb. 1: schematischer Aufbau der bestehenden Umgebung

Auswahl Patchrelease und Patchtermin

Im Januar 2017 haben wir uns entschieden die Planung für das Patchen der Exadatas anzugehen. Da dies unseren ersten Exadatas sind, haben wir beschlossen den Patch mit Unterstützung des Oracle ACS durchzuführen. Nach Absprache mit dem ACS-Support hat sich ein Patchtermin für Mitte März herauskristallisiert.

Als nächstes stand das zu verwendende Patchrelease zur Diskussion. Dabei waren wir in der Lage zwischen dem Exadata Release 12.1 und dem neuen Major Release 12.2 zu wählen. Das Release 12.2 ist das Erste und laut Oracle auf allen zertifizierten Exadata-Konfigurationen getestet. Die Oracle ACS Einschätzung zu diesem Release zeigt ein anderes Bild. Als Indikator wurde uns der Hinweis gegeben, dass es noch Probleme beim Upgrade und im Betrieb geben kann. Nach reiflicher Überlegung ob es sinnvoll ist auf das neueste Release zu gehen oder auf dem bewährten noch zu bleiben haben wir uns entschieden das Upgrade auf 12.2 (12.2.1.1.0.170126.2) zu wagen. Ein maßgeblicher Grund war, dass aktuell auf diesen Exadatas keine produktive Datenbank war und wir uns durch das neue Release das Major Upgrade in der Zukunft sparen, was eventuell dann bei produktiven Datenbanken doch zu Problemen führen kann.

Patchablauf

Der Patchverlauf ist in folgende Schritte unterteilt:

Zuerst muss die GridInfrastructure 12.1, die aktuell den Patchstand Juli 2016 hat, auf mindestens Oktober 2016 gebracht werden. Dies ist eine Grundvoraussetzung damit das Exadata 12.2 Release überhaupt eingespielt werden kann. Wir werden in diesem Zuge die Grid-Infrastructure und das Datenbank auf den Patchstand Januar 2017 anheben.

Im nächsten Schritt wird der Cellserver gepatcht. Danach folgt die VM (domU) und im Anschluss die dom0. Abschließend zum Patch der Exadata-Maschine werden die Infiniband-Switche mit einem Firmware-Upgrade versorgt. Danach wird die GridInfrastructure auf 12.2 angehoben.

Zum Schluss erstellen wir noch ein neues Datenbank-Home für die Version 12.2, damit wir für zukünftige Anforderungen schon die Basis zur Verfügung haben.

Pretasks

Zur Vorbereitung für den Patchday gibt es noch einige Schritte durchzuführen. Beginnend mit dem Erstellen des „exachks“. Dieser muss im Fall einer virtualisierten Exadata immer für die dom0 und die domU (VM) auf dem 1. Knoten erstellt werden. Kritische Probleme im resultierenden Exachk-Report, die eventuell einen erfolgreichen Patch verhindern, müssen vorher bereinigt werden.

Im Anschluss an den Exachk werden weitere Prechecks mit dem „patchmgr“ durchgeführt. Hierfür ist es nötig, dass vorab die benötigten Patche von der Oracle-Website heruntergeladen werden und auf den Knoten der Exadata zur Verfügung stehen.

Benötigt werden folgende Komponenten (für alle Upgrade-Steps):

```
Storage server software (12.2.1.1.0.170126.2) and InfiniBand switch
software(2.2.4-3) Patch 21052028: EXADATA 12.2.1.1.0 (MOS NOTE 2207148.1) -
> p21052028_122110_Linux-x86-64.zip
```

```
Database server bare metal / domU ULN
exadata_dbserver_12.2.1.1.0_x86_64_base OL6 channel ISO image
(12.2.1.1.0.170126.2) -> p25463013_122110_Linux-x86-64.zip
```

Database server dom0 ULN exadata_dbserver_dom0_12.2.1.1.0_x86_64_base OVM3 channel ISO image (12.2.1.1.0.170126.2) -> p25463019_122110_Linux-x86-64.zip

Exadata database server update - patchmgr Patch 21634633:
DBSERVER.PATCH.ZIP ORCHESTRATOR PLUS DBNU - ARU PLACEHOLDER ->
p21634633_122110_Linux-x86-64.zip

OPatch latest update for 11.2, 12.1 and 12.2 _ Patch 6880880: OPatch patch of version 12.2.0.1.8 for Oracle software releases 12.1.0.x (installer) and 12.2.0.x (OCT 2016) -> p6880880_121010_Linux-x86-64.zip

For DBhome: 12.1.0.2.170117 - Jan 2017 (Patch 24968615) plus OJVM PSU - Database 12.1.0.2.170117 Patch 24917972

Grid Software clone version 12.2.0.1.0 -> p25528839_122110_Linux-x86-64.zip

V839960-01.zip Oracle Database 12c Release 2 (12.2.0.1.0) for Linux x86-64

Bundle-Patch Januar 2017 p24968615_121020_Linux-x86-64.zip

Oneoff Patch DB Version 12.1 p24917972_121020_Linux-x86-64.zip

Alle Patche bgzl. Cellserver domU(VM) und dom0 werden direkt auf dem Computenode in der dom0 unter /EXAVMIMAGES/patches bereitgestellt und ausgepackt.

OPatch, BundlePatch und Oneoff Patches für die Datenbank und GridInfrastructure werden in der domU unter /u01/patches abgelegt.

Sind alle Sourcen auf der Exadata verfügbar werden die precheck-runs ausgeführt.

dom0:

```
cd /EXAVMIMAGES/patches/YUM/dbserver_patch_5.170214/
```

```
./patchmgr -dbnodes ~/all_dom0_except_firstnode -precheck -  
nomodify_at_prereq -log_dir auto -target_version 12.2.1.1.0.170126.2 -  
iso_repo /EXAVMIMAGES/patches/YUM/p25463019_122110_Linux-x86-64.zip
```

...

```
2017-07-04 14:00:56 +0200 :ERROR : dbnodeupdate.sh precheck failed  
on one or more nodes
```

```
SUMMARY OF WARNINGS AND ERRORS FOR dba01:
```

```
dba01: # The following file lists the commands that would have been  
executed for removing rpms when specifying -M flag. #
```

```
dba01: ERROR: There are DOMU images not yet upgraded to a UEK4 kernel  
image. You must upgrade all DOMU images to a UEK4 kernel image prior to  
upgrading DOM0.
```

dba01: Warning: This Exadata update includes an Oracle Linux 5 to Oracle Linux 6 update.

dba01: Warning: Although the system was analyzed for custom rpms, any other custom installed software (such as tar-balls) cannot be detected.

dba01: The following known issues will be checked for but require manual follow-up:

2017-07-04 14:00:58 +0200 :ERROR : DONE: dbnodeupdate.sh precheck on dba01

2017-07-04 14:01:09 +0200 :INFO : SUMMARY FOR ALL NODES:

2017-07-04 14:01:09 +0200 :ERROR : dba01 has state: FAILED

domU (von dom0 aus):

```
cd /EXAVMIMAGES/patches/YUM/dbserver_patch_5.170214/
```

```
./patchmgr -dbnodes ~/all_node_vm -precheck -nomodify_at_prereq -log_dir  
auto -target_version 12.2.1.1.0.170126.2 -iso_repo  
/EXAVMIMAGES/patches/YUM/p25463013_122110_Linux-x86-64.zip
```

...

2017-07-04 14:25:08 +0200 :ERROR : dbnodeupdate.sh precheck failed
on one or more nodes

SUMMARY OF WARNINGS AND ERRORS FOR exa01vm01:

exavm01: # The following file lists the commands that would have been
executed for removing rpms when specifying -M flag.

exavm01: Warning: Active network mounts found on this DB node.

exavm01: ERROR: Oracle Grid Infrastructure home must be running release
12.1.0.2.161018 (Oct 2016) Proactive Bundle Patch or later before updating
to Exadata 12.2.1.1.0 or later.

exavm01: ERROR: Preventive abort of update due to dependency issues or a
missing package that will fail the update

exa01vm01: The following known issues will be checked for but require
manual follow-up:

2017-07-04 14:25:11 +0200 :ERROR : DONE: dbnodeupdate.sh precheck on exavm01

2017-07-04 14:25:21 +0200 :INFO : SUMMARY FOR ALL NODES:

2017-07-04 14:25:21 +0200 :ERROR : exavm01 has state: FAILED

...

Cellserver:

```
cd /EXAVMIMAGES/patches/CELL/patch_*
./patchmgr -cells all_cells -cleanup
./patchmgr -cells all_cells -patch_check_prereq
```

Infiniband:

```
cd /EXAVMIMAGES/patches/CELL/patch_*
./patchmgr -ibswitches ibswitches.lst -upgrade -ibswitch_precheck
```

Nachdem wir alle Checks durchgeführt haben und keine Auffälligkeiten die patchverhindernd sind fanden, standen wir dem Patchday optimistisch gegenüber.

Patchday

Der Tag beginnt mit dem Einspielen des Januar 2017 Bundle Patches für die GridInfrastructure und die Datenbank ohne Auffälligkeiten. Damit sind die Vorbedingungen für das Upgrade des Exadata Machine Stacks gelegt.

Nächster Schritt ist der Patch der Cellserver. Beim erneuten Precheck kommt unerwartet das Ergebnis „ERROR: Patch prerequisite checks failed“. Mit dem zusätzlichen Hinweis die Logfiles des „patchmgr“ nach dem genauen Fehler zu durchforsten.

Ergebnis der Analyse war, dass der precheck ein Problem mit der Abfrage eines der konfigurierten NTP-Severn eines Cellserver hatte. Nach Analyse des Patchmgr-Script-Aufrufes ergab sich, dass auch bei Verfügbarkeit mehrerer NTP-Server das Skript Fehler meldet, wenn einer dieser nicht erreichbar ist.

script: dostep.sh.tmpl

```
# Check NTP servers.
...
    # Exclude local clock
    if [ "$ntp_server" != '127.127.1.0' ]; then
        ((ntp_servers_cnt++))
        # It takes about 20 seconds with -t 5 if ntp server is not
        responding
        echo "Check NTP server: ntpdate -t 5 -q -d $ntp_server"
        ntpdate -t 5 -q -d $ntp_server
        if [ $? -ne 0 ]; then
            echo "[ERROR] NTP server $ntp_server is not responding"
            ret_code=$((ret_code | 1 ))
        fi
    fi
done
fi
if [ $ntp_servers_cnt -eq 0 ]; then
```

```
    echo "[WARNING] No NTP server(s) found in /etc/ntp.conf"
fi
```

...

Mittels „ipconf“ auf dem entsprechenden Cellserver haben wir den problematischen NTP-Server entfernt und ein erneuter precheck führte zum Erfolg.

Danach den Patchmgr mit „nohup ./patchmgr -cells ~/all_cells -patch &“ laufen lassen und warten bis alle Cellserver gepatcht sind.

Das nächste Problem lässt nicht lange auf sich warten. Während der Patchmgr die Cellserver patcht haben wir ein Auge auf die Logfiles und sehen, dass der Patchprozess auf dem 3. Cellserver einfach nicht weiterläuft.

Nach Anzeige der abhängigen Prozesse von pstree ist zu erkennen, dass der Patchmgr beim Aufruf des „hostname“ commands auf dem 3. Cellserver hängt. Da keine weiteren Abhängigkeiten bestehen, beschließen wir den Prozess zu killen

```
pstree -a 335709
```

```
patchmgr ./patchmgr -cells /root/all_cells -patch
```

```
    +-ssh -o BatchMode=yes -o ConnectionAttempts=1 -o ConnectTimeout=60
    -q -l root exacel03 hostname
```

```
ps -axwww |grep 352637
```

```
352637 pts/2    S          0:00 ssh -o BatchMode=yes -o ConnectionAttempts=1 -o
ConnectTimeout=60 -q -l root exacel03 hostname
```

```
kill 352637
```

Danach läuft das Cellpatching erfolgreich bis zum Ende.

Im Anschluss daran fügen wir den herauskonfigurierten NTP-Server wieder mit „ipconf“ hinzu.

Als nächstes werden die domU's (VMs) gepatcht „./patchmgr -dbnodes ~/all_node_vm -backup -iso_repo /EXAVMIMAGES/patches/YUM/p25463013_122110_Linux-x86-64.zip -target_version 12.2.1.1.0.170126.2 -allow_active_network_mounts“ .

Bei diesem Schritt gab es keine Überraschungen.

Im Anschluss steht das Patchen der dom0 auf dem Programm.

```
„nohup ./patchmgr -dbnodes ~/all_dom0_except_firstnode -upgrade -log_dir auto -target_version
12.2.1.1.0.170126.2 -iso_repo /EXAVMIMAGES/patches/YUM/p25463019_122110_Linux-x86-
64.zip -nobackup &“
```

Der Patch läuft erfolgreich durch, jedoch stellen wir fest, dass kein Infiniband in den VMs verfügbar ist. Nach weiteren langen Analysen zeigt sich, dass das „rdma“-Kernel Modul in der dom0 nicht geladen wird. Das manuelle Laden via modprobe führt zu einem kurzfristigen Erfolg, jedoch überlebt dies keinen Reboot. Dafür haben wir als Workaround die Datei „/etc/exadata/security/pset.xml“ angepasst, in der das Flag für das Starten des „rdma-Services“ auf „off“ stand.

```
...  
<system_service name="rdma">  
<value>on</value>  
</system_service>  
...
```

Ein erneuter Reboot führt auch mit diesem Workaround zum Erfolg. Als letzten Schritt des Exadata Machine Patchings stehen noch die Infiniband-Switches an. Diese werden mit „patchmgr -ibswitches ibswitches.lst -upgrade &“ gepatcht. Dabei passiert es, dass ein Switch nach einem Reboot nicht mehr startet. Ein Wiederbeleben via ILOM-Console führt nicht zum Erfolg und deshalb wird der Switch im Nachgang durch einen Neuen ersetzt. Dieser kommt natürlich nicht mit der korrekten Firmwareversion und muss noch auf die gleiche Version wie auf den anderen bereits gepatchten upgegraded werden. Dieses Mal funktioniert das Ganze.

Nun steht das Upgrade der GridInfrastructure auf Version 12.2 an. Die Software wird mittels den Gold-Images in die domU eingehängt, danach wird aus dem 12.2er Softwarehome „gridSetup.sh -skipRemoteCopy“ gestartet und als Option „Upgrade“ ausgewählt. Im Verlauf des Upgrades ergeben sich verschiedene Probleme wie z.B. die falsche „osdba“-Gruppe, falsches ORACLE_BASE, etc. Am Ende gelingt auch dieses Upgrade, jedoch haben wir im Nachgang die GridInfrastructure neu aufgesetzt, um die Fehler zu korrigieren.

Zum Schluss folgt noch die Installation der Oracle Datenbanksoftware 12.2 in ein neues Home, welches ohne Probleme vonstattengeht.

Fazit

Wir haben hart gekämpft und die Exadata bezwungen. Die Erfahrungen die wir gemacht haben, helfen uns dabei die Plattform besser einzuschätzen. Der Upgradeprozess ist nicht so stabil wie wir es erwartet haben, obwohl Oracle diese laut eigener Aussage für alle zertifizierten Konfigurationen testet.

Kontaktadresse:

Timo Giese
Fiducia & GAD IT AG
Fiduciastr. 20
D-76227 Karlsruhe

Telefon:	+49 (0) 721-4004 1017
E-Mail	timo.giese@fiduciagad.de
Internet:	www.fiduciagad.de