

Vom Cluster zum Grid  
–  
Upgrade von Oracle RAC

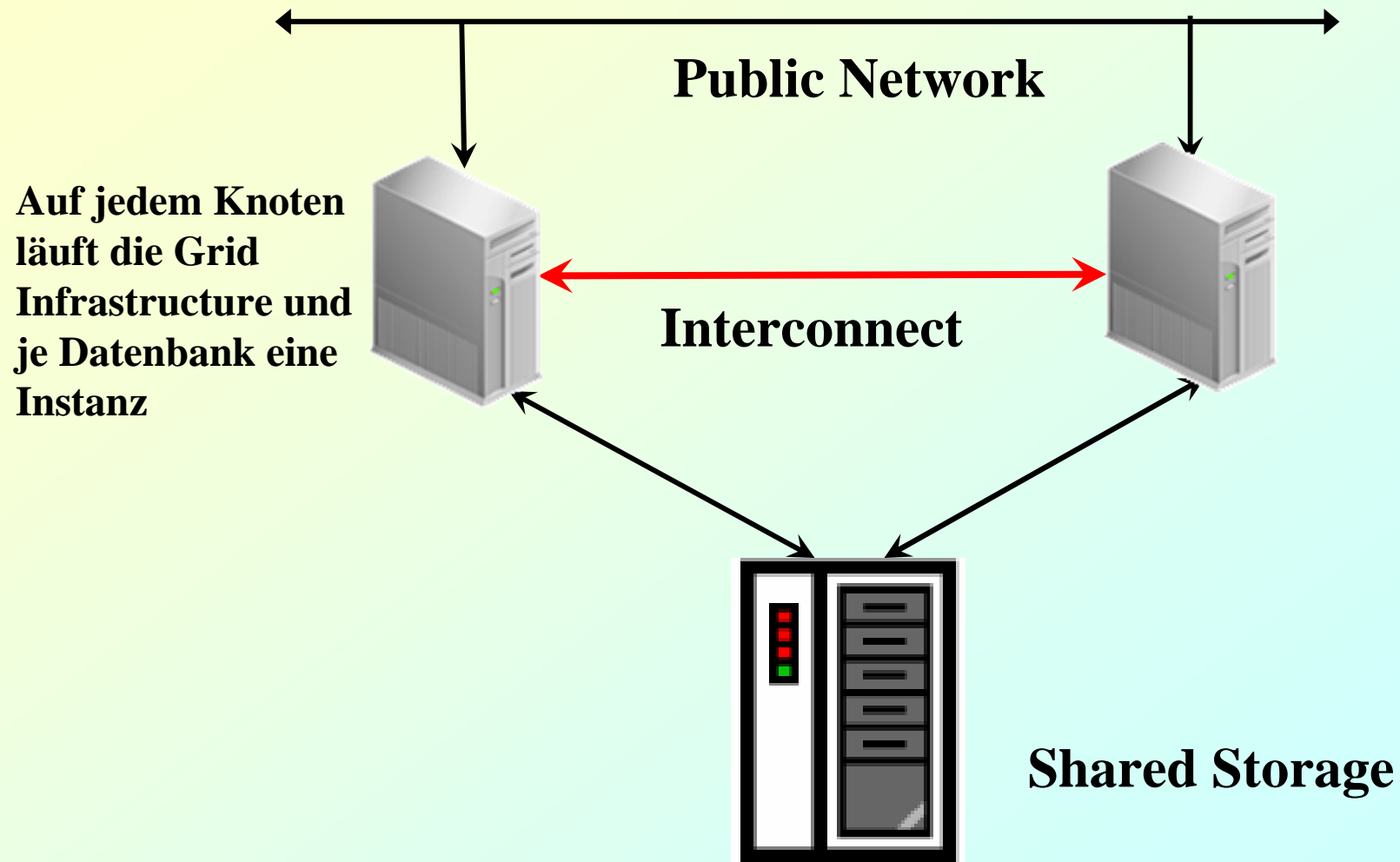
Dr. Frank Haney

DOAG-Konferenz 2010

## Inhalt

- Besonderheiten der Systemarchitektur
- Herausforderungen an ein effizientes Szenarium
- Vom 9i Cluster Manager zu 10g CRS - Reminiszenzen
- Upgrade der Clusterware nach 11 Release 1
- Besonderheiten der Grid Infrastructure 11g Release 2
- Ablauf
- Konfiguration
- Probleme
- Verifikation des Erfolgs
- Bemerkungen zu Datenbankinstallation und Upgrade

## RAC – Prinzip



# RAC-Upgrade – Herausforderungen

## Allgemeine Herausforderungen:

- Ø CRS muß immer den höchste Release-Stand haben
- Ø Unter Umständen müssen Datenbanken verschiedener Releases parallel in einem Cluster betrieben werden
- Ø Minimierung der Downtime für Cluster, Datenbank und Applikationen

## Plattformspezifische Probleme:

- Ø Schaffung identischer lokaler Pfade für die Installationen (keine lokalen Platten)
- Ø Alle Knoten teilen sich eine OS-Installation, d.h. schreiben in das gleiche root-Dateisystem
- Ø Jede Installation (CRS, Datenbank, Agent etc.) gehört einem eigenen OS-User
- Ø Jede Installation hat eigenes `/etc/oraInst.loc` = globales Inventory

## RAC-Upgrade – Szenarien

### **Neue Hardware steht zur Verfügung**

Migration mit Export/Import

#### **Vorteile:**

- ∅ Keine Downtime für Cluster und Datenbanken
- ∅ Die alte Umgebung wird nicht beeinträchtigt.
- ∅ Die neue Clusterware bzw. Grid Infrastructure kann sauber separat installiert und getestet werden, bevor die Datenbanken nach und nach auf das neue Release gebracht werden.

#### **Nachteile:**

- ∅ Die Notwendigkeit separater Hardware und die damit verbundenen Kosten
- ∅ Es wird mehr Platz im Storage benötigt.
- ∅ Die Downtime für die aktuell zu migrierende Datenbank ist hoch, weil Export und Import auch mit Datapump relativ langsam

### **Neue Hardware steht *nicht* zur Verfügung**

Upgrade des Clusters

Upgrade der Datenbanken mit Export/Import, manuell oder mit dem DBUA

## Vom 9i-Cluster Manager zur 10g-Clusterware

### **Neue Hardware stand *nicht* zur Verfügung**

#### **Installation der Clusterware (CRS) – zwei Möglichkeiten**

- Ø Mit Upgrade der Konfigurationsdatei des 9i-Clusters zur 10g OCR
- Ø Völlig autonome Installation der 10g Clusterware (Die Installation hat keine “Kenntnis” davon, daß auf den Knoten bereits der Cluster Manager 9i installiert und konfiguriert ist. )

#### **Nachteile beider Varianten:**

- Ø Gleichzeitige Downtime für alle Datenbanken zur Installation der 10g Clusterware (CRS)
- Ø Mögliche Konflikte zwischen 9i und 10g Cluster im laufenden Betrieb

## Von 9 nach 10 – Das Problem

Bei der Ausführung von `root .sh` auf dem ersten Knoten wird ein Upgrade der Konfigurationsdatei des 9i-Clusters zur 10g-OCR versucht. Das schlägt fehl:

```
PROT-8: Failed to import data from specified file to the cluster registry  
Oracle Cluster Registry configuration upgraded successfully
```

In `ocrconfig.log` steht:

```
[ OCRCONF][2546082016]ocrconfig starts...  
[ OCRCONF][2546082016]Upgrading OCR data  
[ OCRRAW][2546082016]propriogid:1: INVALID FORMAT  
[ OCRRAW][2546082016]ibctx:1:ERROR: INVALID FORMAT  
[ OCRRAW][2546082016]proprinit:problem reading the bootblock or superbloc 22
```

`srvctl config` zeigt danach die 9i-Datenbanken, aber keine Instanzen sind registriert, man kann keine hinzufügen und den Eintrag auch nicht aus der OCR löschen!

Service Request lief von 4/2007–9/2008, letztlich ohne Resultat!

**Gewähltes Szenarium:** Völlig *autonome* Installation der Clusterware 10g und schrittweises, *manuelles* Upgrade der Datenbanken

## Von 9 nach 10 – Zeitbedarf

<b>Migrationsschritt</b>	<b>Downtime</b>
Installation Clusterware 10g	2h
Patchen Clusterware nach 10.2.0.4	Keine
Installation und Patchen Datenbanksoftware 10g	Keine
Vorbereitung einer 9i-Datenbank für das Upgrade	Keine
Upgrade <i>einer</i> Datenbank nach 10g	4h
Anpassung Oracle Applications (eBusiness-Suite)	2h

### Bemerkungen:

1. Die Installation der Clusterware erfordert eine Downtime für **alle** Datenbanken im Cluster.
2. Das eigentliche DB-Upgrade bedeutet Downtime **nur** für die jeweilige Datenbank.



## Von 10.2 nach 11.1 – I

Ein Rolling Upgrade der Clusterware von 10g nach 11g ist möglich, ein Rolling Upgrade der Datenbank noch nicht.

### 1. Vorbereitung

- Blackout in Grid Control setzen (Cluster, Datenbankinstanzen)
- Bestandsverzeichnis in `/etc/OraInst.loc` richtig setzen.
- Datenbankinstanz und Nodeapps auf dem ersten Knoten stoppen

```
srvctl stop instance -d db_name -i instance_name  
srvctl stop nodeapps -n node_name
```

- EM-Agent auf dem ersten Knoten stoppen

```
emctl stop agent
```

- In das Quellverzeichnis der Clusterware 11g Software wechseln und als root folgendes Skript ausführen:

```
upgrade/preupdate.sh -crshome $CRS_HOME  
-crsuser crs_user
```

## Von 10.2 nach 11.1 – II

### 2. Durchführung

- OUI aus dem Quellverzeichnis der Clusterware 11g Release 1 als Eigentümer der Clusterware aufrufen: `./runInstaller`
  - Prüfen, ob in das richtige Oracle Home installiert wird
  - Knoten auswählen, der auf 11g gebracht werden soll
  - Am Ende der Installation als `root` folgendes Skript ausführen:
  - `$ORA_CRS_HOME/install/rootupgrade`
  - Das Skript startet die CRS, die nodeapps und alle HA-Applikationen auf dem Knoten
3. Schritte 1 und 2 nacheinander auf den anderen Knoten durchführen
  4. Blackout in Grid Control beenden.
  5. Eventuelles DB-Upgrade

**Fazit:** Einfach und problemlos (11.1 ist eher 10.2.0.5)

Keine Downtime für Cluster und Applikationen

**Zeitbedarf:** < 1h (ohne DB-Upgrade)

## Grid Infrastructure 11g Release 2

Schritt von 11.1 nach 11.2 ist deutlich größer als der nach 11.1!

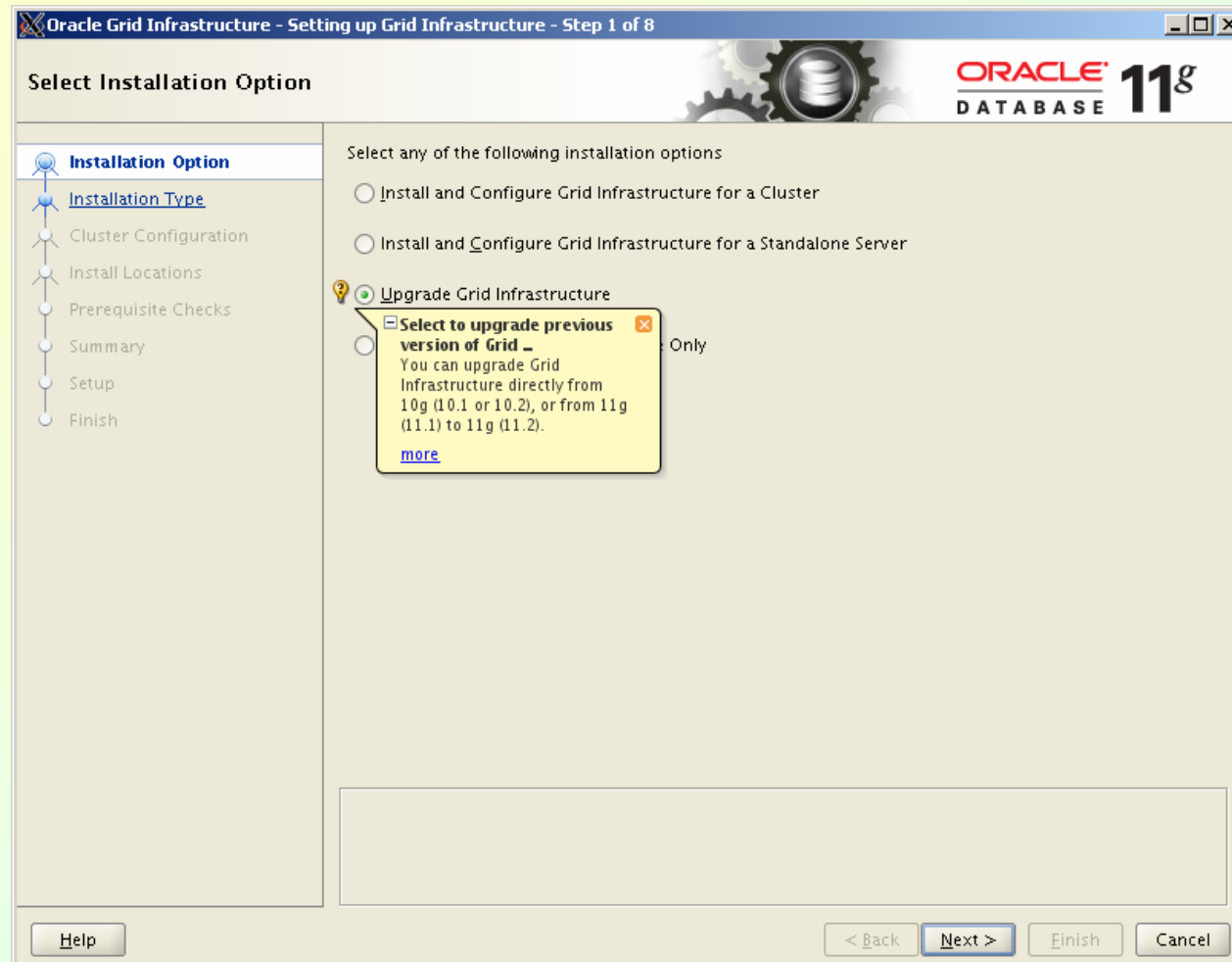
Was ist neu aus der Sicht des Upgrades?

- Ø Völlig neue Architektur (Prozeßstruktur etc.)
- Ø ASM jetzt gebündelt mit Grid Infrastructure
- Ø SCAN Listener erforderlich
- Ø Auch beim Upgrade Installation in ein neues Home
- Ø Eigener Cluster Time Synchronization Service (CTSS)
- Ø OCR und Voting können in ASM gespeichert werden
- Ø Eigenes ASM Cluster File System (ACFS) für Nichtdatenbankdateien
- Ø Backups der Voting Disk jetzt automatisch
- Ø Erweiterte Funktionalität von CVU (OUI-Integration)
- Ø Erweiterte Syntax von `srvctl` und `crsctl`

## Vorbereitung

- Ø SUSE-Release wenigstens 10 Update 2 wegen der nötigen Packages
- Ø Richtiges `/etc/oraInv.loc` setzen
- Ø Altes CRS 10.2.0.x läuft und ist voll funktional
- Ø Das neue GRID\_HOME darf **kein** Unterverzeichnis des HOME des Users oder irgendeines ORACLE\_BASE sein. (Berechtigungen!)
- Ø Neues GRID\_HOME muß **separat** (parallel) zum alten ORA\_CRS\_HOME sein.
- Ø Wenigstens eine IP als SCAN-VIP wird benötigt. (In produktiven Umgebungen 3) Diese SCAN-VIP muß im DNS aufgelöst werden. Eintrag in `/etc/hosts` reicht nicht! (Bei 3 nur im DNS.)
- Ø Die alten Skripte `/etc/init.*` werden von `rootupgrade.sh` überschrieben!!! Vorher sichern!
- Ø Das neue init-Skript `/etc/init.ohasd` enthält knotenspezifische Einträge! Eindeutigkeit durch symbolische Links herstellen. Gleiches gilt für die Konfigurationsdatei `/etc/ohasd`.
- Ø Die Verzeichnisse `/opt/oracle` und `/opt/ORCLfmap` müssen spezifisch für jeden Knoten sein. Eindeutigkeit durch symbolische Links herstellen.

# Start



# SCAN Listener

Oracle Grid Infrastructure - Setting up Grid Infrastructure - Step 4 of 10

Grid Infrastructure SCAN Information

ORACLE 11g  
DATABASE

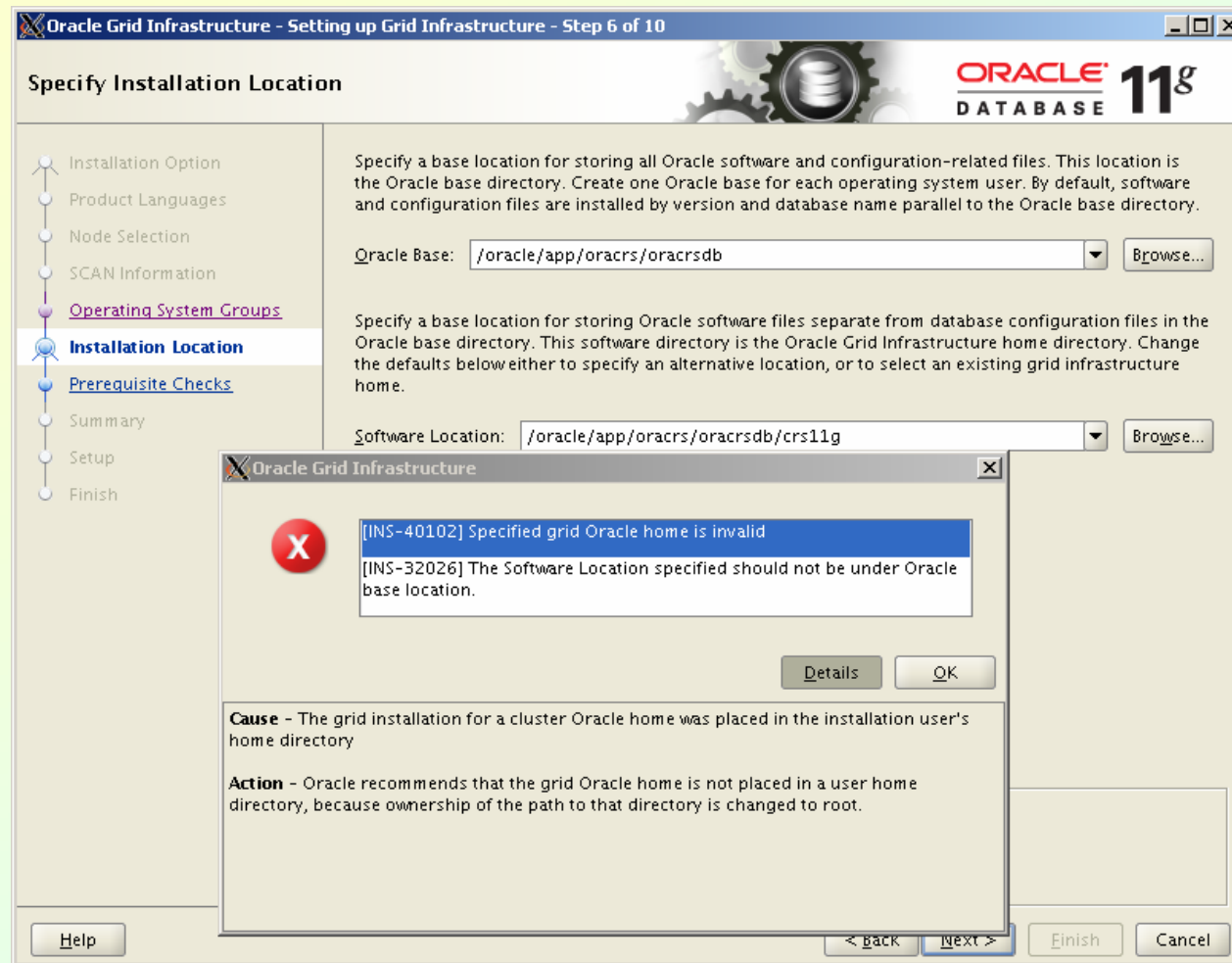
- Installation Option
- Product Languages
- Node Selection
- SCAN Information**
- Operating System Groups
- Installation Location
- Prerequisite Checks
- Summary
- Setup
- Finish

Cluster Name: lobocrs

SCAN Name:   
Single ClientSCAN names allow clients to connect to the cluster with a single hostname in their connect string

SCAN Port:

# GRID\_HOME



# Voraussetzungen prüfen

Oracle Grid Infrastructure - Setting up Grid Infrastructure - Step 7 of 10

Perform Prerequisite Checks

ORACLE DATABASE 11g

Some of the minimum requirements for installation are not completed. Review and fix the issues listed in the following table, and recheck the system.

Check Again Fix & Check Again Show Failed All Nodes Ignore All

Checks	Status	Fixable
Network Time Protocol (NTP)	Failed	

**Details**

Network Time Protocol (NTP) - This task verifies cluster time synchronization on clusters that use Network Time Protocol (NTP).

Operation Failed on Nodes: [lolek](#)

Verification result of failed node: lolek  
Expected Value : n/a  
Actual Value : n/a  
[Back to Top](#)

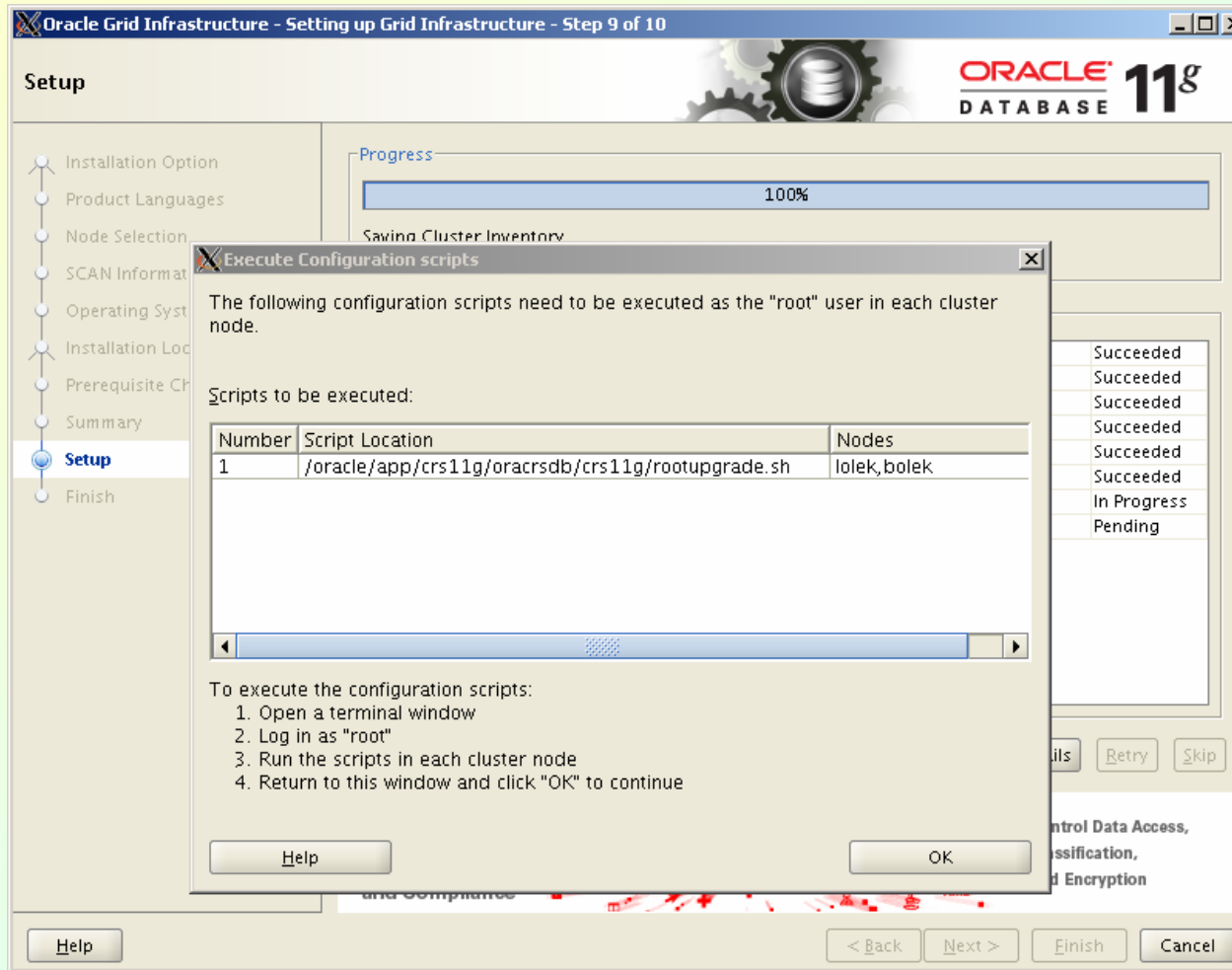
This task... protocol (NTP). [\(more\)](#)

Operation Failed on Nodes: [lolek](#)

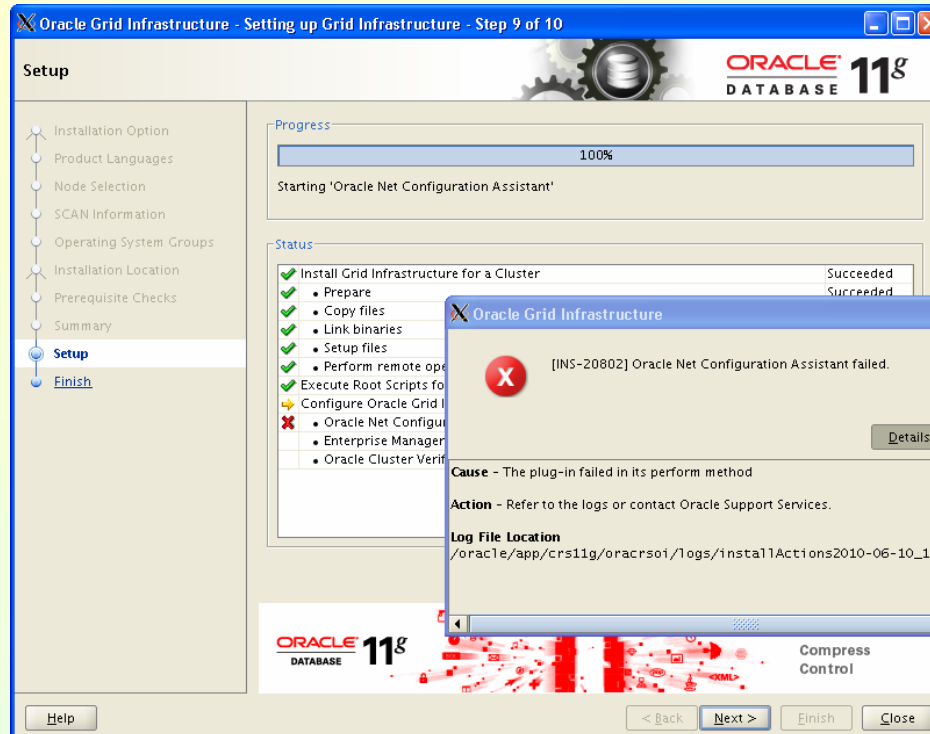
Help < Back Next > Finish Cancel



# rootupgrade.sh



# Konfiguration I



**INFO:** Oracle Net Services Configuration:

**INFO:** Listener "LB10" will be migrated to Oracle home: /oracle/app/crs11g/oracrsdb/crs11g

**INFO:** Failed migrating Oracle Net Services configuration. The exit code is 1.

**Ursache:** Listener heißt nicht LISTENER

**Lösung:** Listener mit 10g-NETCA entfernen und mit 11.2-NETCA neu anlegen.

## Konfiguration II

INFO: Checking if CTSS Resource is running on all nodes...  
INFO: CTSS resource check passed  
INFO: Querying CTSS for time offset on all nodes...  
INFO: Query of CTSS for time offset passed  
INFO: Check CTSS state started...  
INFO: **CTSS is in Observer state.** Switching over to clock synchronization checks using NTP  
INFO: Starting Clock synchronization checks using Network Time Protocol(NTP)...  
INFO: NTP Configuration file check started...  
INFO: NTP Configuration file check passed  
INFO: Checking daemon liveness...  
INFO: **Liveness check failed for "ntpd"**  
INFO: Check failed on nodes:  
INFO: bolel,lolek  
INFO: PRVF-5415 : Check to see if NTP daemon is running failed  
INFO: Clock synchronization check using Network Time Protocol(NTP) failed  
INFO: PRVF-9652 : Cluster Time Synchronization Services check failed  
INFO: Post-check for cluster services setup was unsuccessful on all the nodes.



### **Lösung:**

- 1) wenn externe Zeitsynchronisation wie NTP verwendet wird, dann muß diese richtig konfiguriert sein und laufen. oder
- 2) Wenn der CTSSD für die Zeitsynchronisation sorgen soll, darf kein anderer Dienst dafür konfiguriert sein! Also `/etc/ntp.conf` oder `/etc/xntp.conf` entfernen.

## Konfiguration III

### Problem Interface in segmentierten Netzen:

```
>oifcfg getif
eth3 192.168.144.0 global cluster_interconnect
>srvctl config nodeapps -a
VIP exists.:bolek
VIP exists.: /172.18.144.88/172.18.144.88/255.255.255.192
VIP exists.:lolek
VIP exists.: /172.18.144.87/172.18.144.87/255.255.255.192
```

### Lösung:

Siehe Metalink 276434.1: Modifying the VIP or VIP Hostname of a 10g or 11g Oracle Clusterware Node

```
>oifcfg getif -global
eth3 192.168.144.0 global cluster_interconnect
eth1 172.18.144.64 global public
>srvctl config nodeapps -a
VIP exists.:bolek
VIP exists.: /172.18.144.88/172.18.144.88/255.255.255.192/eth1
VIP exists.:lolek
VIP exists.: /172.18.144.87/172.18.144.87/255.255.255.192/eth1
```

## Resultat Cluster

Ressourcen werden vom init-Prozeß gestartet:

```
crsctl stat res -t -init
```

```
-----  
NAME                TARGET  STATE        SERVER        STATE_DETAILS  
-----  
Cluster Resources  
-----  
ora.asm  
  1                OFFLINE OFFLINE  
ora.crsd  
  1                ONLINE   ONLINE      lolek  
ora.cssd  
  1                ONLINE   ONLINE      lolek  
ora.cssdmonitor  
  1                ONLINE   ONLINE      lolek  
ora.ctssd  
  1                ONLINE   ONLINE      lolek        ACTIVE:0  
ora.diskmon  
  1                ONLINE   ONLINE      lolek  
ora.evmd  
  1                ONLINE   ONLINE      lolek  
ora.gipcd  
  1                ONLINE   ONLINE      lolek  
ora.gpnpd  
  1                ONLINE   ONLINE      lolek  
ora.mdnsd  
  1                ONLINE   ONLINE      lolek
```

## Resultat lokale Ressourcen

Ressourcen werden vom CRS gestartet:

```
oracrs@lolek:/dbinst/stage/Oracle/RDBMS/11.2.0.1_x86-64/grid>crsctl stat res -t
```

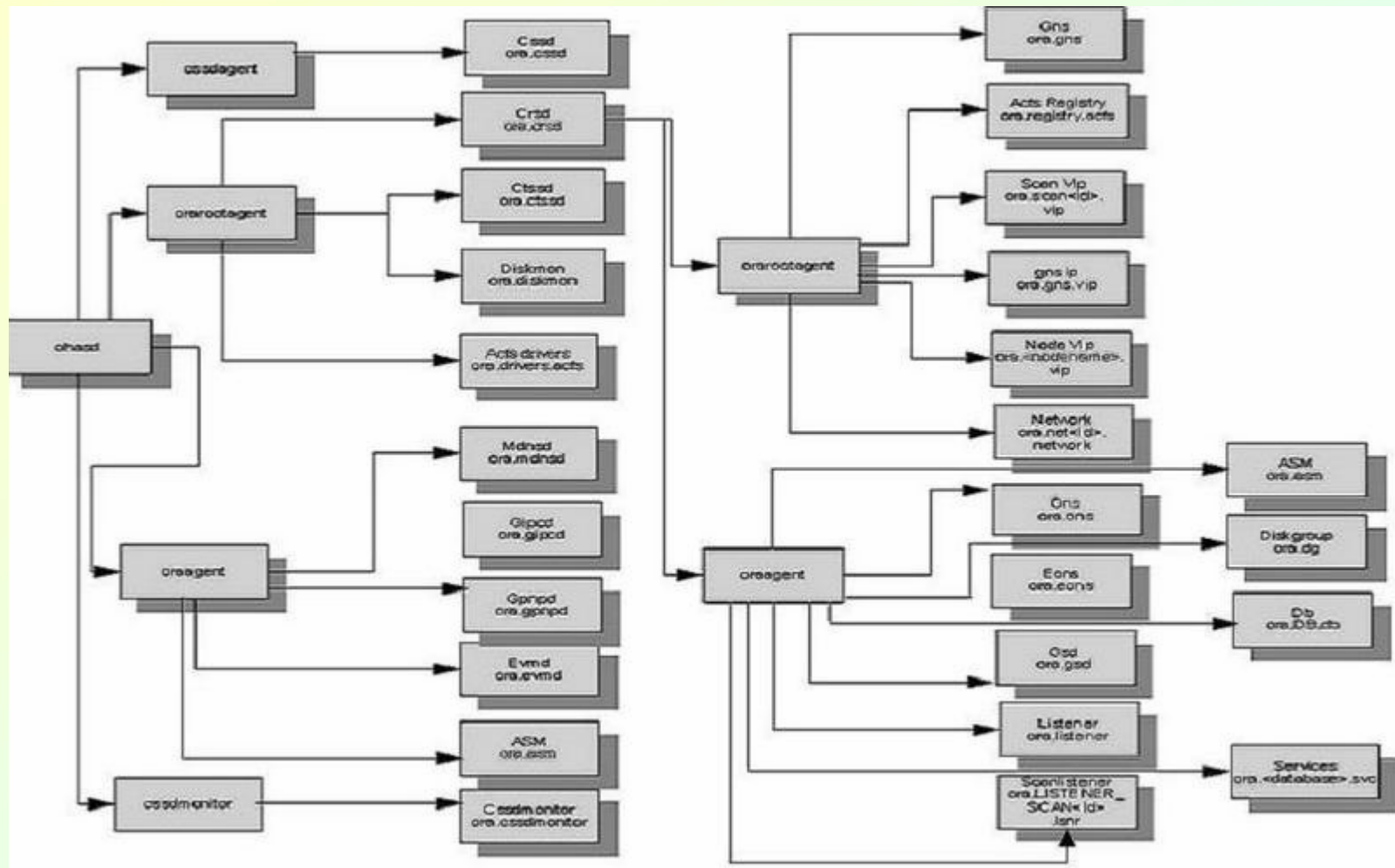
```
-----  
NAME                TARGET  STATE        SERVER        STATE_DETAILS  
-----  
Local Resources  
-----  
ora.asm  
                ONLINE  ONLINE       bolelek       Started  
                ONLINE  ONLINE       lolek        Started  
ora.eons  
                ONLINE  ONLINE       bolelek  
                ONLINE  ONLINE       lolek  
ora.gsd  
                OFFLINE OFFLINE       bolelek  
                OFFLINE OFFLINE       lolek  
ora.net1.network  
                ONLINE  ONLINE       bolelek  
                ONLINE  ONLINE       lolek  
ora.ons  
                ONLINE  ONLINE       bolelek  
                ONLINE  ONLINE       lolek
```

## Resultat clusterweite Ressourcen

-----  
Cluster Resources  
-----

```
ora.LISTENER_SCAN1.lsnr
  1          ONLINE  ONLINE          lolek
ora.bolek.vip
  1          ONLINE  ONLINE          bolek
ora.lb10.db
  1          ONLINE  ONLINE          lolek          Open
  2          ONLINE  ONLINE          bolek          Open
ora.lb20.db
  1          ONLINE  ONLINE          lolek          Open
  2          ONLINE  ONLINE          bolek          Open
ora.lolek.vip
  1          ONLINE  ONLINE          lolek
ora.oc4j
  1          ONLINE  ONLINE          lolek
ora.scan1.vip
  1          ONLINE  ONLINE          lolek
```

# Prozeßstruktur und Startreihenfolge





## Prozeßstruktur und Startreihenfolge

```
ohasd->orarootagent->ora.cssdmonitor : Überwacht CSSD und node health.
      ->ora.ctss : Cluster Time Synchronization Services Daemon
      ->ora.crsd->oraagent
                                ->ora.LISTENER.lsnr
                                ->ora.LISTENER_SCAN.lsnr
                                ->ora.ons
                                ->ora.eons
                                ->ora.asm
                                ->ora.DB.db
      ->orarootagent->ora.nodename.vip
                                ->ora.net1.network
                                ->ora.gns.vip
                                ->ora.gnsd
                                ->ora.SCAN.vip

->cssdagent ->ora.cssd: Cluster Synchronization Service

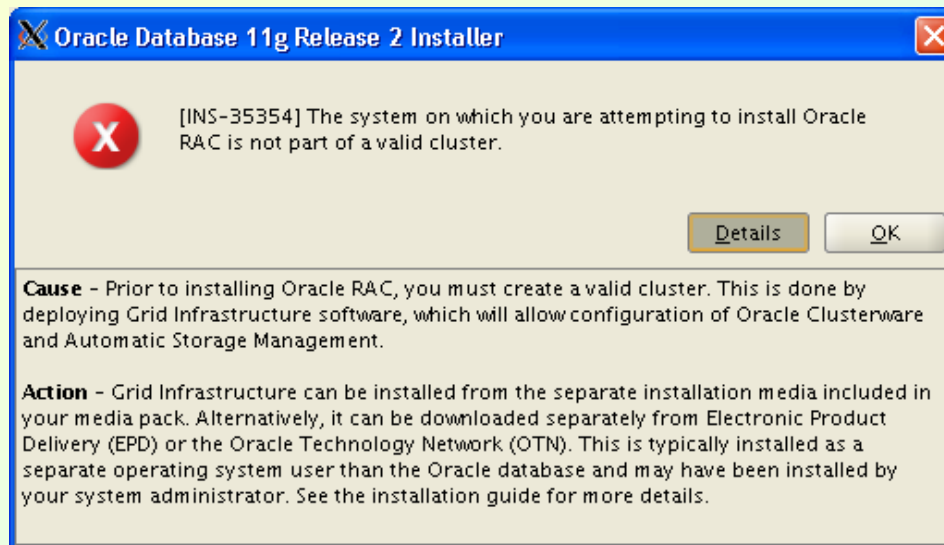
->oraagent ->ora.mdnsd: Für DNS lookup
           ->ora.evmd
           ->ora.asmd
           ->ora.gpnpd: Grid Plug and Play
```

Quelle: Metalink Note 1053147.1

## Besonderheiten der Datenbankinstallation

**Grid Infrastructure 11.2.0.1 läuft und ist voll funktional.**

**Problem:** Cluster wird nicht erkannt. (Verschiedene Inventories!)



**Lösung:**

Zunächst den OUI mit folgender Syntax aufrufen

```
./runInstaller -silent -ignoreSysPrereqs -attachHome  
ORACLE_HOME="/oracle/app/crs11g/oracrsdb/crs11g"  
ORACLE_HOME_NAME="Ora11g_gridinfrahome1" LOCAL_NODE='lolek'  
CLUSTER_NODES=lolek,bolek CRS=true
```

## Besonderheiten des Datenbankupgrade

**Skript utlu112i.sql muß zwingend in der alten Umgebung gelaufen sein.**  
(Nimmt Änderungen am Data Dictionary vor)

**Upgrade ohne RAC (nur als Single Instance) in der neuen Umgebung**  
**STARTUP UPGRADE** geht nicht bei aktivem RAC!

**Also:**

```
ALTER SYSTEM SET CLUSTER_DATABASE=false SCOPE=spfile;
```

```
host srvctl stop database -d LB10
```

**STARTUP UPGRADE**

```
@?/rdbms/admin/catupgrd.sql
```

```
ALTER SYSTEM SET CLUSTER_DATABASE=true SCOPE=spfile;
```

```
SHUTDOWN IMMEDIATE
```

```
host srvctl remove database -d LB10
```

```
host srvctl add database -d LB10 -o $ORACLE_HOME
```

```
host srvctl add instance -d LB10 -i LB10-01 -n lolek
```

```
host srvctl add instance -d LB10 -i LB10-02 -n bolek
```

```
host srvctl start database -d LB10
```

## Fazit

**Das Upgrade der Clusterware zur Grid Infrastructure ist eine komplexe Maßnahme, und es gilt hier besonders:**

**Testen, Testen, Testen**

### **Literatur:**

**Oracle® Grid Infrastructure** Installation Guide

**Oracle® Clusterware** Administration and Deployment Guide

**Oracle® Real Application Clusters** Administration and Deployment Guide

**Oracle® Database** Upgrade Guide

Dr. Frank Haney

info@haney.it

Tel.: 03641-210224

**ORACLE**

**CERTIFIED  
PROFESSIONAL**