



RAC auf Windows

Installation und Betrieb:
Tipps und Tricks

Herrmann & Lenz Services GmbH

Herrmann & Lenz Solutions GmbH

- Erfolgreich seit 1996 am Markt
- Firmensitz: Burscheid (bei Leverkusen)
- Beratung, Schulung und Betrieb/Fernwartung rund um das Thema Oracle Datenbanken
- Schwerpunktthemen: Hochverfügbarkeit, Tuning, Migrationen und Troubleshooting
- Herrmann & Lenz Solutions GmbH
 - Produkt: Monitoring Module
 - Stand auf Ebene 2



Überlegungen

- Strategische Plattform: Windows
- Know-How im eigenen Unternehmen
- Anforderungen:
 - Hochverfügbarkeit
 - Skalierbarkeit
- Ergebnis: RAC auf Windows





Installation

BS - Installation

- Auf jeden Fall: 64bit, wenn irgend möglich!
- Grid Infrastructure ab 11.2.0.1 nicht verfügbar für 32bit Windows
- Zertifizierungsmatrix beachten!
- Installation über RDP:
 - Win2003: `mstsc /v:servername /f /console`
 - Win2003 R2: `mstsc /v:servername /f /admin`
 - Ab Win2008: `mstsc /v:servername /f`



BS – Firewall & Virens Scanner

- Windows-Firewall und Virens Scanner während der Installation und während Patch- oder Add-Node-Vorgängen deaktivieren!
- Firewall **nie** für das private Netzwerk aktivieren.
- Ggfs. für die Firewall entsprechende Ausnahmen konfigurieren
- Während Virus-Scans auf OCFS-Platten: Clusterware stoppen



BS - Umgebungsvariablen

- Im RAC-Umfeld: `ORACLE_HOME` nicht als Umgebungsvariable setzen!
- Soll ein bestimmtes `ORACLE_HOME` verwendet werden, im Aufruf die `PATH`-Variable anpassen
- `TEMP` bzw. `TMP`-Variablen sowohl für den Installationsbenutzer als auch für das System auf feste Verzeichnisse setzen (keine Variablen)



Plattenplatz

- Software-Installation
 - Out of Place-Upgrade ab 11.2 für GI vorgeschrieben, für DB empfohlen
 - lieber doppelten Platz einplanen, um auch für den nächsten Patchset gerüstet zu sein



Plattenplatz

- Shared Storage
 - Bei Verwendung von ASM: **eine** erweiterte Partition mit logischem Laufwerk pro ASM-Disk
 - Erstellung mit `diskpart.exe`, später Labeln mit `asmtool(g).exe`
 - Write Cache auf Knoten-Ebene muss deaktiviert werden! Auf SAN-Ebene ok, wenn Caching für alle Knoten sichtbar



Benutzer (1)

- Berechtigungen
 - Lokale Admin-Rechte (Domänen-Admin-Rechte nicht erforderlich); Berechtigung durch **direkte** Mitgliedschaft in der Administratoren-Gruppe, **keine** Vererbung!
 - Domänenbenutzer (empfohlen seit 11.2) oder lokaler Benutzer möglich; wenn lokal: identische Benutzernamen und Passwörter



Benutzer (2)

- Erforderliche Benutzerrechte:



- Manipulation der Remote Registry
- Verwaltung des Auditing- und Security-Log
- Anmeldung als Dienst / Batch Job
- Memory-Quoten für einen Prozess anpassen
- Als Teil des Betriebssystems handeln



Windows 2008

- Windows 2008: UACL deaktivieren (wenn möglich)
- Zumindest für Administratoren die Ausführung von Programmen als Administrator ohne Nachfrage gestatten



User Account Control: Admin Approval Mode for the Built-in Administrator account	Disabled
User Account Control: Allow UIAccess applications to prompt for elevation without using the...	Disabled
User Account Control: Behavior of the elevation prompt for administrators in Admin Approv...	Elevate without prompting
User Account Control: Behavior of the elevation prompt for standard users	Prompt for credentials
User Account Control: Detect application installations and prompt for elevation	Enabled
User Account Control: Only elevate executables that are signed and validated	Disabled
User Account Control: Only elevate UIAccess applications that are installed in secure locations	Enabled
User Account Control: Run all administrators in Admin Approval Mode	Disabled



Namenskonventionen (1)

- Gewohnheit: Bei der Windows-Administration gibt es keine Case Sensitivity
- Bei RAC schon – egal ob GI 11gR2 oder Clusterware 10gR2 / 11gR1!
- Hostnamen müssen identisch geschrieben werden (also nicht NODE1 für den ersten und node2 für den 2. Knoten). Ansonsten evtl. später Probleme mit dem ONS



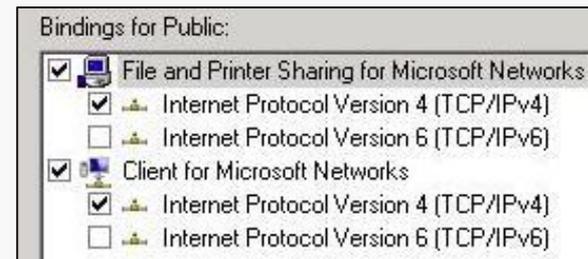
Namenskonventionen (2)

- Best Practice für Netzwerk-Schnittstellennamen: Umbenennung in kürzere / sprechendere Namen als "LAN Connection 1" etc.
- Bug: Nicht die Namen PUBLIC und PRIVATE (alles Großbuchstaben) verwenden! Alle anderen Schreibweisen (z.B. Public) sind ok.
- Auch hier gilt: Namensgebung identisch auf allen Knoten



Netzwerk

- Public Network muss an 1. Stelle in der Bindungsreihenfolge stehen
- Keine Unterstützung von IPv6 (auch nicht in 11.2)!
 - Deaktivieren oder niedriger als IPv4 priorisieren
- Ab Windows 2003R2: SNP-Features "Receive Side Scaling" und "TCP/IP Offload" deaktivieren, falls vorhanden



Netzwerk – Multicast und Redundant Interconnect

- Multicast (ab 11.2.0.2)
 - Soll laut GI-Installationshandbuch für die IP-Ranges 224.0.0.0 /24 und 230.0.1.0 /24 ermöglicht werden
 - Wird aber tatsächlich nicht verwendet (im Gegensatz zu Linux / UNIX)
 - Redundant Interconnect Usage existiert **nicht** für die Windows-Plattform!



Netzwerk - DNS

- In den Eigenschaften der öffentlichen (public) Netzwerkschnittstelle sollte die dynamische Registrierung deaktiviert werden (Registerkarte DNS)
- Öffentliche Namen und VIPs manuell im DNS registrieren



DNS			
HLDC			
Forward-Lookupzonen			
hl.intranet			
w2k8-vm2-vip	Host		192.168.149.18
w2k8-vm2	Host		192.168.149.16
w2k8-vm1-vip	Host		192.168.149.17
w2k8-vm1	Host		192.168.149.15



Registry-Parameter

- **LocalAccountTokenFilterPolicy**
 - Zur Verwendung des `net use`-Befehls auf die administrativen Freigaben der entfernten Knoten
 - **REG_DWORD** in `HKLM\Software\Microsoft\Windows\CurrentVersion\Policies\System` (**Wert=1**)
- **DisableDHCPMediaSense**
 - Entkopplung einer IP-Adresse vom Netzwerkadapter bei Verbindungsabbruch zum Switch
 - **REG_DWORD** in `HKLM\System\CurrentControlSet\Services\Tcpip\parameters` (**Wert=1**)



Registry-Parameter (2)

- MaxNegPhaseCorrection (Wert=0)
 - Verhindert, dass der Windows-Zeitserver die Zeit in die Vergangenheit anpasst
 - REG_DWORD unter HKLM\SYSTEM\CurrentControlSet\Services\W32Time\Config
 - Anschließend Aktivierung an der Kommandozeile:
`W32tm /config /update`
- Alternativ kann ab 11.2 auch der Cluster-Zeitserver ctssd verwendet werden (dann aber Windows-Zeitserver deaktivieren!)





Betrieb

ORACLE-Tools

- Windows 2008: "Run as Administrator"-Kontext für diverse ORACLE-Tools erforderlich (dbca, netca...)
 - Kontext ist bei Nutzung der Startmenü-Icons bereits implementiert
 - Kommandozeile muss jedoch eine Administrator-Kommandozeile sein



Windows-Dienste

Name	Description	Status	Startup Type
Oracle Object Service		Started	Manual
Oracle Process Manager			Automatic
OracleASM1020TNSListenerLISTENER_Prac01		Started	Manual
OracleASMService+ASM1		Started	Manual
OracleClusterVolumeService		Started	Automatic
OracleCRService		Started	Manual
OracleCSService		Started	Manual
OracleEVMSERVICE		Started	Manual

10gR2 / 11gR1

Name	Descrip...	Status	Startup T...
Oracle Object Service		Started	Automatic
Oracle Process Manager (instance1)		Started	Automatic
Oracle VMRACDB1 VSS Writer Service			Manual
OracleASMService+ASM1		Started	Manual
OracleDBConsoleVMRACDB		Started	Automatic
OracleJobSchedulerVMRACDB1			Disabled
OracleMTSRecoveryService		Started	Automatic
OracleOHService		Started	Automatic
OracleOraCrs11g_home1TNSListener		Started	Manual
OracleOraCrs11g_home1TNSListenerLISTENER_SCAN1			Manual
OracleOraCrs11g_home1TNSListenerLISTENER_SCAN2		Started	Manual

11gR2



Parameterdateien

- Listener, ASM
 - Bis einschließlich 11.1: zugehörig zur Datenbank
 - Ab 11.2: Zugehörig zur Grid Infrastructure
- Parameterdateien daher ab 11.2 in
GI_HOME\network\admin
- TNS_ADMIN sollte auf dieses Verzeichnis zeigen (auch für DB_HOME-Verzeichnisse)



Parameterdateien – sqlnet.ora

- Ab 11.2:
 - NAMES.DIRECTORY_PATH muss EZCONNECT enthalten
 - SQLNET.AUTHENTICATION_SERVICES = (NTS)
 - Ansonsten keine Anmeldung der Instanzen beim SCAN-Listener



Parameterdateien – tnsnames.ora

- Ab 11.2:
 - Clients \geq 11.2.0.1:
 - keine Host-Adressliste mehr erforderlich; SCAN wird als Host eingetragen
 - Clients $<$ 11.2.0.1:
 - können nicht mit mehreren IP-Adressen hinter einem einzigen Hostnamen umgehen
 - Verwendung einer Host-Adressliste mit den SCAN-IP-Adressen
 - "alte" RAC-Syntax mit Host-VIPs ebenfalls möglich



CREATE DATABASE mit dem dbca

- Achtung: mit dem dbca erzeugte Skripte zur Erstellung einer Datenbank funktionieren nicht korrekt (direkte Erstellung einer Datenbank mit dbca schon)!
- In den Skripten fehlt die Erstellung der erforderlichen Windows-Dienste auf den entfernten Knoten
- Manuelle Eingriffe sind erforderlich (alle Versionen einschl. 11.2.0.2)





Anwendungsbeispiel einer benutzerdefinierten Cluster-Ressource

Failover und Gratuitous ARP

- Szenario: RAC-Server und Clients in unterschiedlichen Netzwerk-Segmenten, verbunden mit Cisco-Switchen
- Ausfall eines Knotens; Failover der VIP auf den verbleibenden Knoten
 - Sofort sichtbar innerhalb des RAC-Segments
 - Erst nach mehreren Minuten sichtbar innerhalb des (unterschiedlichen) Client-Segments
- Problem: ARP-Cache auf dem Switch wird nicht aktualisiert bzw. Information über den Failover wird erst nach Ablauf eines definierten Intervalls (hier: 8 Minuten) vom Switch in andere Segmente weitergegeben.



Gratuitous ARP

- Laut Microsoft ist dies beabsichtigtes Verhalten, um falsche ARP-Einträge für doppelte IP-Adressen zu vermeiden
- MS-Failover-Cluster umgeht das Problem "durch internen Call, die API ist nicht veröffentlicht" (MS Support)
- Reduzierung des Intervalls auf dem Switch oder Verschiebung des RAC in das Client-Segment sind keine Optionen
- Workaround: eigene Cluster-Ressourcen!



Benutzerdefinierte Cluster-Ressourcen

- Erstellung einer eigenen Cluster-Ressource für jede Knoten- und SCAN-VIP
- Action-Skripte gewährleisten die Leerung des lokalen ARP-Caches sowie einen Ping seitens der geschwenkten VIP auf das Default Gateway, sobald eine VIP den Status ONLINE erreicht



Benutzerdefinierte Cluster-Ressourcen

- Resultat: geschwenkte VIPs sind für Clients aus anderen Netzwerksegmenten sofort erreichbar
- Action-Skript muss Return-Codes für Start-, Stopp- und Status-Kommandos zurückgeben.
- Beispiel Action-Skript, auf eigenen Bedarf anzupassen:

```
GI_HOME\crs\demo\action_scr.bat
```



Benutzerdefinierte Cluster-Ressourcen

- Beispiel Hilfs-Ressource zum Refresh einer SCAN-VIP:

```
crsctl add resource ora.scan1_vip.app
-type cluster_resource
-attr
"ACTION_SCRIPT='C:\script\refresh_scan1_vip.cmd'
, PLACEMENT='restricted',
SERVER_POOLS='Generic', CHECK_INTERVAL='60',
START_DEPENDENCIES='hard(ora.scan1.vip)
pullup(ora.scan1.vip)',
STOP_DEPENDENCIES='hard(ora.scan1.vip)',
RESTART_ATTEMPTS='2', ACL='owner:nt
authority\system:rwx,pgrp::r-x,other:---'"
```



Cluster-Ressourcen: Application VIP

- Basierend auf Action-Scripts können auch Ressourcen zur Administration eigener Anwendungen und zugehörige Application-VIPs konfiguriert werden:

```
appvipcfg create -network=1  
-ip=192.168.149.21 -vipname=ora.myapp.vip  
-user="nt authority\system"
```



Fazit

- Traditionell gibt es viele Vorurteile gegen Oracle (insbesondere RAC) auf Windows
- Diese können nach unseren Erfahrungen so nicht bestätigt werden
- Wünschenswert: etwas mehr "Zuwendung" seitens Oracle / Oracle Support
- Empfehlung wie immer: Verwendung der Plattform, über die das meiste Wissen im Haus vorhanden ist!



Hilfreiche Dokumentation

- MOS:
 - Doc ID # 811271.1: RAC & Clusterware Best Practices (Windows)
 - Doc ID # 810394.1: RAC & Clusterware Best Practices (Generic)
 - Doc ID # 887522.1: 11gR2 SCAN explained
 - Doc ID # 948456.1: Pre 11g DB-Issues in 11gR2 GI Environment
 - ...
- Oracle Dokumentation
 - Release Notes für Windows
 - GI / Real Application Clusters Installation Guide für Windows



Fragen & Kontakt

- susanne.jahr@hl-services.de
- <http://www.hl-services.de>
- Hier in der Ausstellung Ebene 2 (gelb) direkt gegenüber der Rolltreppe

