

RAC oder Data Guard

Was ist denn jetzt eine HA Lösung?

Johannes Ahrends



- Oracle Spezialist seit 1992
 - 1992: Presales bei Oracle in Düsseldorf
 - 1999: Projektleiter bei Herrmann & Lenz Services GmbH
 - 2005: Technischer Direktor ADM Presales bei Quest Software GmbH
 - 2011: Geschäftsführer CarajanDB GmbH
- 2011 → Ernennung zum Oracle ACE
- Autor der Bücher:
 - Oracle9i für den DBA, Oracle10g für den DBA, Oracle 11g Release 2 für den DBA
- DOAG Themenverantwortlicher Datenbankadministration, Standard Edition
- Hobbies:
 - Drachen steigen lassen (Kiting) draußen wie drinnen (Indoorkiting)
 - Motorradfahren (nur draußen)



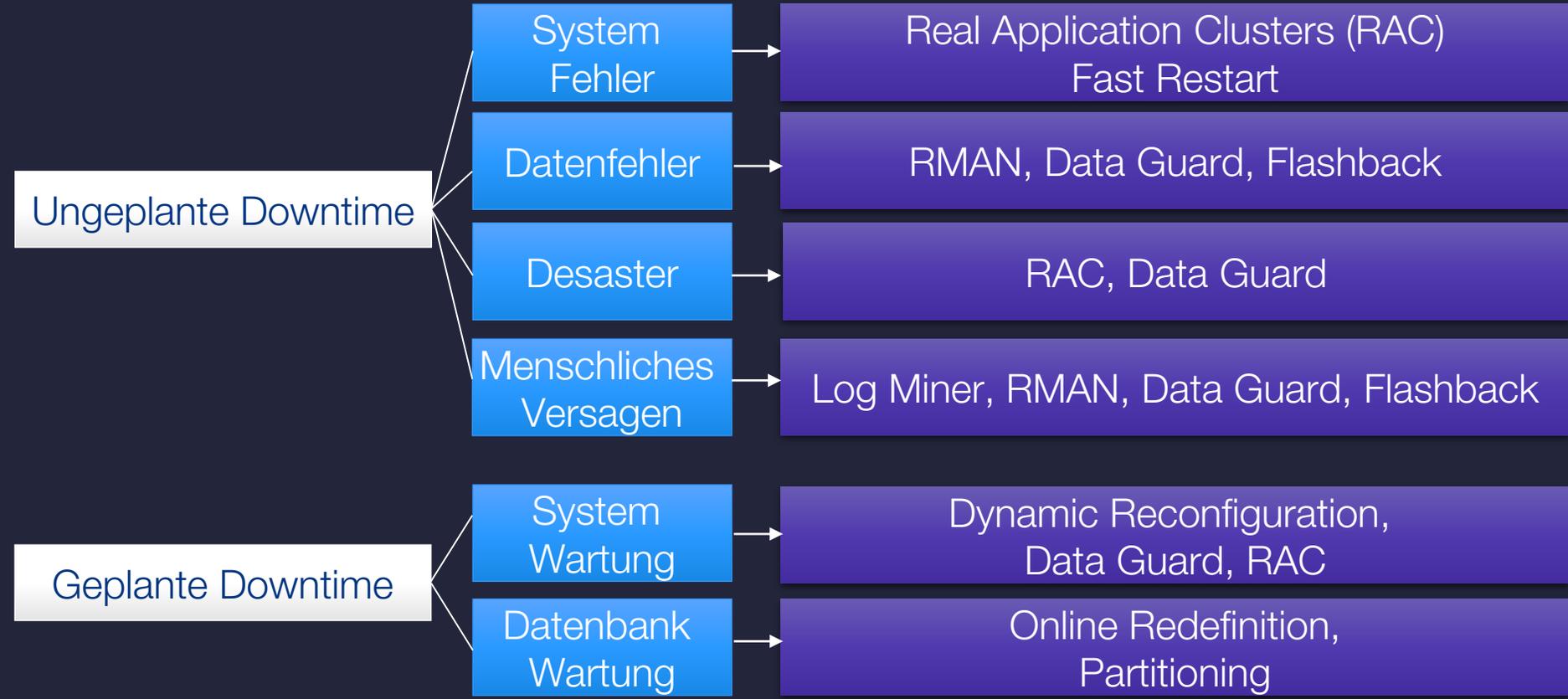


High Availability

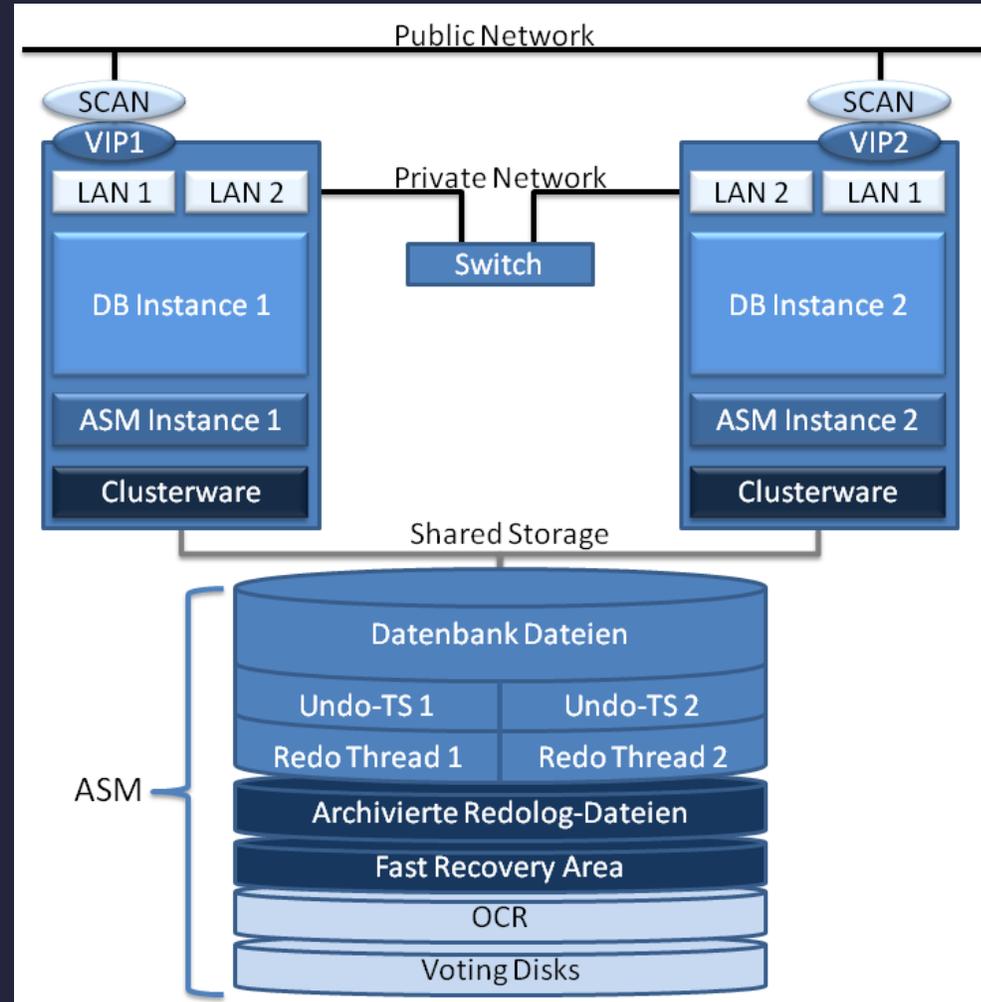
Was ist Hochverfügbarkeit?

Hochverfügbarkeit (englisch *high availability*, *HA*) bezeichnet die Fähigkeit eines Systems, trotz Ausfalls einer seiner Komponenten mit einer hohen Wahrscheinlichkeit (oft 99,99 % oder besser) den Betrieb zu gewährleisten. In Abgrenzung zur Fehlertoleranz kann es bei dem Betrieb im Fehlerfall zu einer Unterbrechung kommen. (Wikipedia)

HA Optionen aus Sicht von Oracle



- DIE Hochverfügbarkeitslösung für Oracle Datenbanken
 - Datenbankinstanz auf mind. zwei Guests
 - Voraussetzung: Shared Storage
- Vorteil:
 - Ausgereifte Technologie
- Nachteil:
 - Aufwändige Implementierung
 - Latenzgefahr bei Metrocluster
 - Lizenzkosten (bei Standard Edition Two enthalten)



- Voraussetzung für Cluster Lösungen
 - Real Application Cluster
 - RAC One Node
 - Oracle Cluster Filesystem
 - Oracle ASM
 - Oracle ASM Cluster Filesystem
 - 3rd Party Anwendungen

- Unabhängig von der Oracle Datenbank Software
- Eigene Software-Installation
- Clusterware Release \geq Oracle Datenbankrelease
- Rolling Upgrade möglich
 - Oracle 12.1.0.2 \rightarrow Oracle 18 funktioniert

- Public → Physikalische IP-Adresse an NIC gebunden, wird nicht benutzt!
- Private → Physikalische IP-Adresse an NIC gebunden ausschließlich für Interconnect
- Virtual → je Knoten eine Adresse, die beim Ausfall des Knotens geschwenkt wird
- Scan → in der Regel ein Name mit 3 Adressen für den gesamten Cluster. Für die Kommunikation mit der Anwendung

- Wichtigste Komponente im RAC
- Niemals den Default Service (= Name der Datenbank bzw. PDB) verwenden
- Servicename definiert Failover-Verhalten

```
srvctl add service -database <DBNAME> -service <SERVICENAME> -preferred  
<INSTANCE1> -available <INSTANCE2> -tafpolicy BASIC -role Primary -policy  
AUTOMATIC -failovertype SESSION -failovermethod BASIC -failoverretry 60 -  
failoverdelay 3 -notification TRUE
```

```
srvctl start service -database <DBNAME> -service <SERVICENAME>
```

- Keine RAC-spezifischen Einträge notwendig

```
<TNSALIAS> =  
  (DESCRIPTION =  
    (LOAD_BALANCE = ON)  
    (FAILOVER = ON)  
    (ADDRESS = (PROTOCOL = TCP) (HOST = <scanname>) (PORT = 1521))  
    (CONNECT_DATA =  
      (SERVER = DEDICATED)  
      (SERVICE_NAME = <SERVICENAME>)  
    )  
  )  
)
```

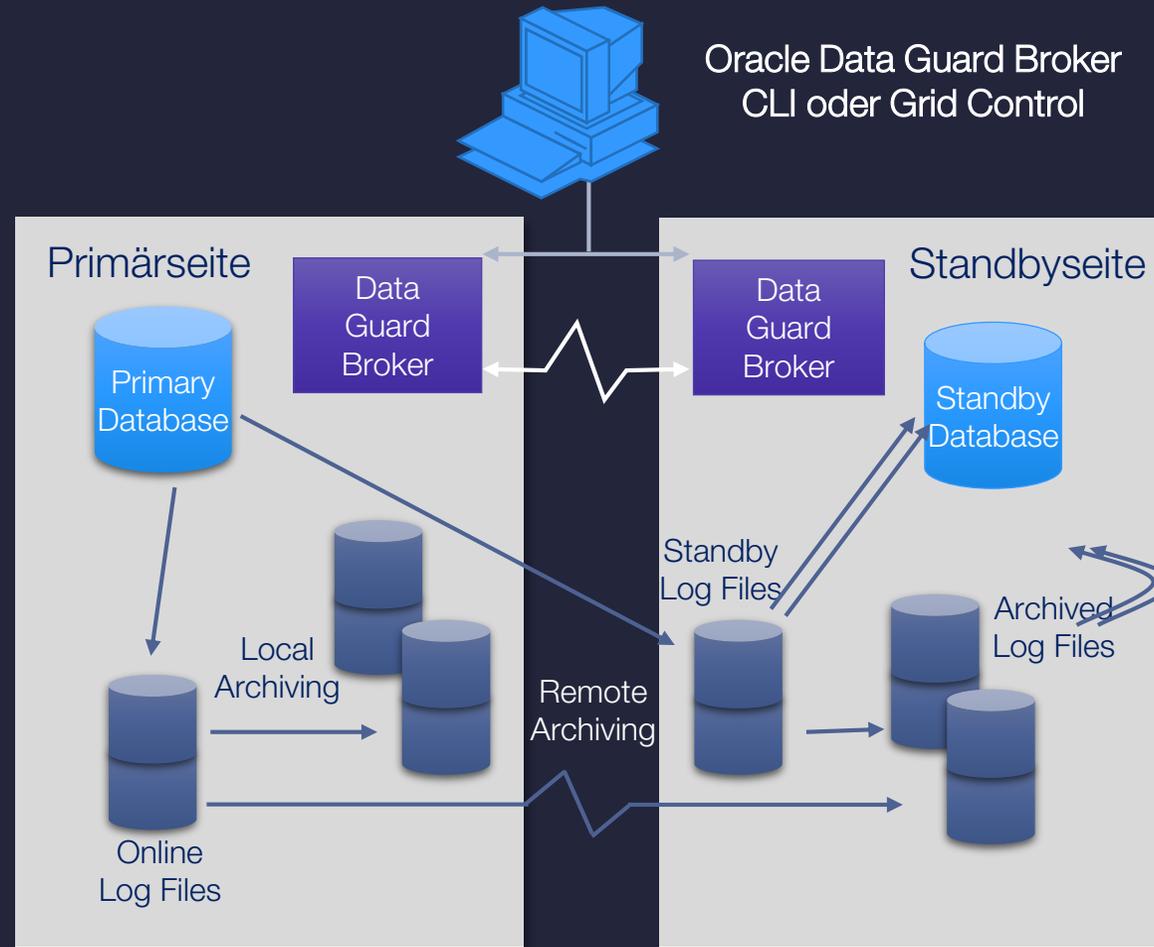
- Service Drain → Service wird geschwenkt und ein Knoten wird langsam geleert

```
srvctl relocate service -database <DBNAME> -service <SERVICENAME> -oldinst  
<INSTANCE1> -newinst <INSTANCE2>
```

Desaster Recovery

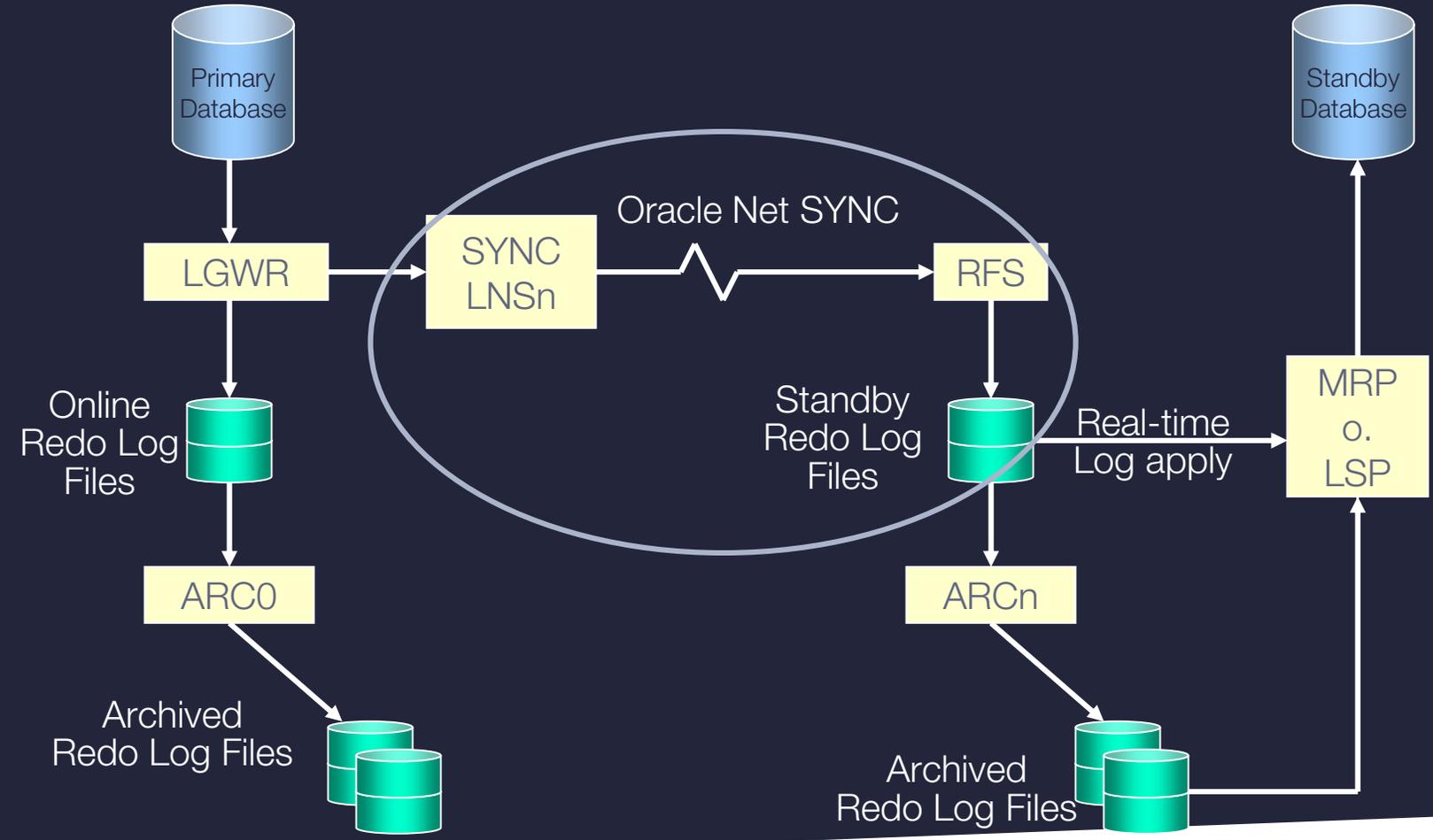
- DIE Disaster Recovery Lösung für Oracle Datenbanken
 - Redolog-Informationen wird in eine zweite Datenbank „recovered“
 - Voraussetzung: Oracle Enterprise Edition
- Vorteil:
 - Ausgereifte Technologie
 - Mehr als eine „Standby“ Datenbank möglich
 - Kein Shared Storage notwendig
- Nachteil:
 - Aufwändige Implementierung
 - Ggf. manuelle Umschaltung (ohne Observer)

Data Guard



- Maximum Protection
 - Es dürfen auf keinen Fall Transaktionen verloren gehen
 - „Transaction Completed“ erst, wenn in primary und eine Standby Redolog geschrieben wurde
 - RPO=0
- Maximum Performance
 - Die primäre Datenbank darf durch die Standby Datenbank nicht beeinflusst werden
 - „Transaction Completed“, wenn in primary Redolog geschrieben wurde
- Maximum Availability
 - Maximum Protection, solange die Standby zur Verfügung steht
 - Maximum Performance, wenn Standby mehr als <n> Sekunden nicht verfügbar ist

Maximum Availability Mode



- Maximum Availability Mode oft benutzt
 - Kein Datenverlust beim Ausfall eines Servers / Clusters
- Observer wird selten eingesetzt
 - Manuelles Umschalten im Fehlerfall notwendig
 - Umschaltdauer / Downtime = Alarm + Telefon + Online + Fehleranalyse + Failover
 - In der Regel nicht unter 2 Stunden!
 - In der Praxis oft irrelevant, weil sowieso alle Komponenten „weg“ sind

- Services müssen gestartet werden
 - Dedizierte Standby Services nur bei Active Data Guard

```
<TNSALIAS> =  
  (DESCRIPTION =  
    (CONNECT_TIMEOUT = 15) (RETRY_COUNT=20) (RETRY_DELAY=3)  
    (TRANSPORT_CONNECT_TIMEOUT = 3)  
    (ADDRESS_LIST =  
      (LOAD_BALANCE = ON)  
      (FAILOVER = ON)  
      (ADDRESS = (PROTOCOL = TCP) (HOST = <scaname1>) (PORT = 1521))  
      (ADDRESS = (PROTOCOL = TCP) (HOST = <scaname2>) (PORT = 1521))  
    )  
    (CONNECT_DATA =  
      (SERVER = DEDICATED)  
      (SERVICE_NAME = <SERVICENAME>)  
    )  
  )  
)
```

- Funktioniert im laufenden Betrieb sehr gut
 - Switchover mit Data Guard
 - Reconnect innerhalb ca. 30 Sekunden
- Was passiert bei Maintenance
 - SCAN-Adresse kann nicht mehr aufgelöst werden
 - Timeout beim Abfragen des Hostnamens
 - → `TRANSPORT_CONNECT_TIMEOUT` sorgt dafür, dass nach 3 Sekunden auf den anderen Host geschaltet wird.
Hinweis: Gilt für jeden Connect!

- Active Data Guard
- Global Services über beide Server / Cluster möglich
- gemeinsame Verwaltung

VMware

- Verschieden Lösungen:
 - vSphere vMotion
 - Verschieben von Guest auf alternativen Host ohne Downtime
 - Unterstützung von Maintenance (Server, Hypervisor)
 - vSphere DRS (Distributed Resource Scheduler)
 - Automatische Verwaltung von Guests (Load Balancing, Green-IT)
 - vSphere HA
 - Automatischer Restart von Guests nach Fehlern
 - Überwachung von virtuellen Maschinen (Heartbeat)
 - ...
- Schnell und einfach zu implementieren

- Standard im Bereich VMware
 - Guest wird einfach auf einen anderen Server verschoben
 - Voraussetzung: Shared Storage
 - Alternativ: Storage vMotion → Storage wird zusätzlich gespiegelt
- Vorteil
 - Einfach zu implementieren
 - Unabhängig von der der Anwendung / Datenbank
- Nachteil
 - Im Fehlerfall muss der Guest auf dem Zielsystem erst gestartet werden

Purpose

Explain to customers how Oracle supports our products when running on VMware

Scope & Application

For Customers running Oracle products on VMware virtualized environments.
No limitation on use or distribution.

Support Status for VMware Virtualized Environments

Oracle has not certified any of its products on VMware virtualized environments. Oracle Support will assist customers running Oracle products on VMware in the following manner: Oracle will only provide support for issues that either are known to occur on the native OS, or can be demonstrated not to be as a result of running on VMware.

If a problem is a known Oracle issue, Oracle support will recommend the appropriate solution on the native OS. If that solution does not work in the VMware virtualized environment, the customer will be referred to VMware for support. When the customer can demonstrate that the Oracle solution does not work when running on the native OS, Oracle will resume support, including logging a bug with Oracle Development for investigation if required.

If the problem is determined not to be a known Oracle issue, we will refer the customer to VMware for support. When the customer can demonstrate that the issue occurs when running on the native OS, Oracle will resume support, including logging a bug with Oracle Development for investigation if required.

NOTE: Oracle has not certified any of its products on VMware. For Oracle RAC, Oracle will only accept Service Requests as described in this note on Oracle RAC 11.2.0.2 and later releases.

Support Position for Oracle Products Running on VMWare Virtualized Environments (Doc ID 249212.1)

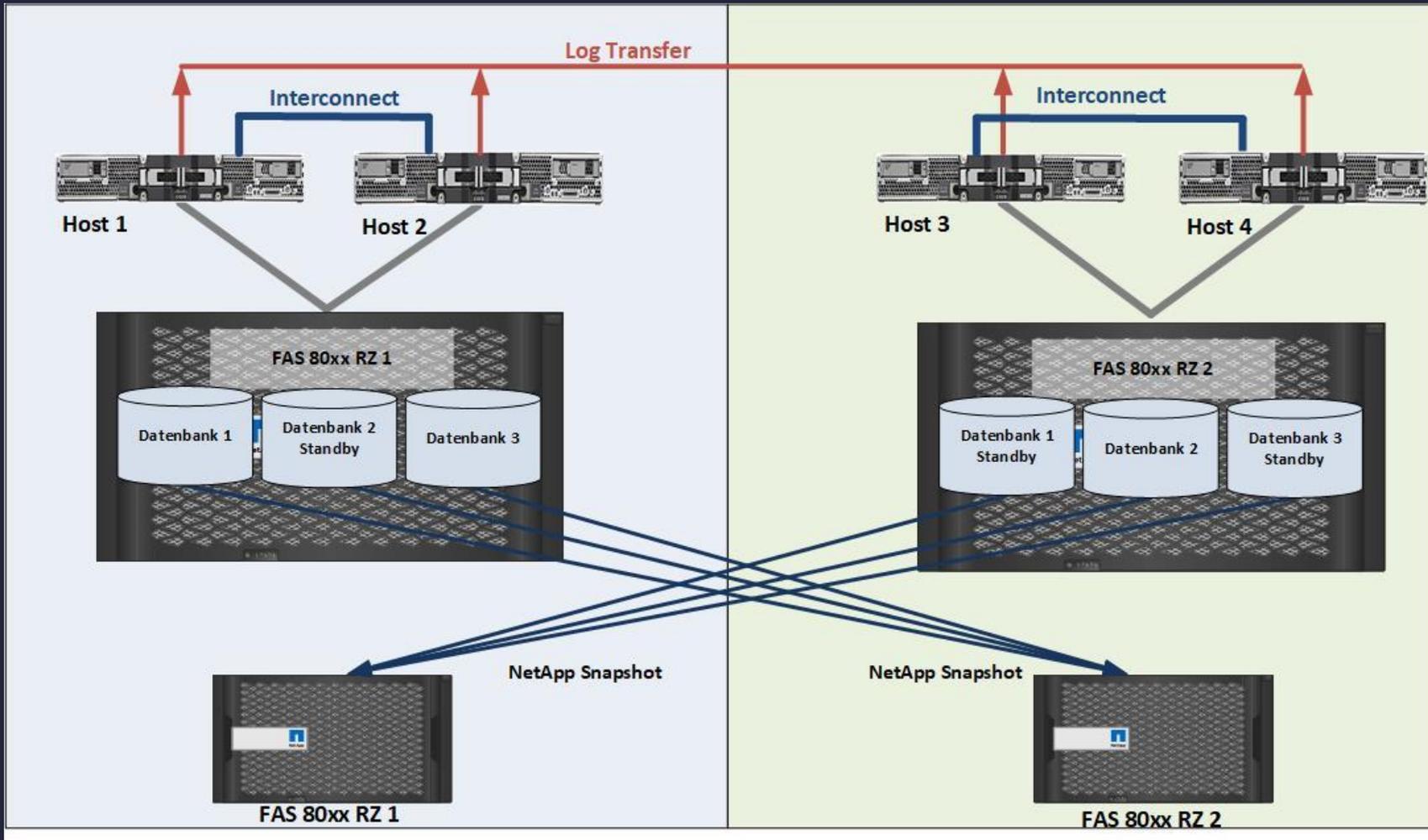
Projektbeispiele

- Ursprüngliche Architektur:
 - Oracle Real Application Clusters (bis 11.2.0.4) auf AIX LPARs
 - Je zwei LPARs pro Stage (4 Stages)
 - Drei Datenbanken pro Cluster (Schemakonsolidierung)
- 2016: Umstellung auf VMware
 - Eigener Oracle VMware Cluster
 - Ein Guest pro Datenbank
 - > 200 Guests
 - Keine Grid Infrastructure

- RAC
 - + Leistungsfähige Server
 - Aufwändige Terminfindung
 - Patch-Downtime bis zu 10 Stunden pro Cluster
- VMware
 - + Patch-Downtime < 15 Minuten
 - + Einfachere Terminfindung
 - + Keine Abhängigkeit von Hardware-Änderungen (vMotion, DRS, HA)
 - Anzahl Server verlängert Patchdauer

- Migration von 8 Oracle 11g Datenbank 6 Knoten Stretched RAC nach
 - Oracle 12c (12.2.0.1)
 - 2 Knoten RAC
 - Data Guard
 - Multitenant
 - 4 Stages (Maintenance, Test, Vorproduktion, Produktion)
 - 2 Projekte
- Start des Projektes Migration August 2015
- Start des Projektes Multitenant Januar 2016

Schematische Darstellung



- RAC
 - + Rolling Upgrade funktioniert
 - + Sehr robust
- Data Guard
 - + Sehr schnell aufgebaut
 - Probleme bei RZ-Ausfall
- Multitenant
 - + Hohe Selbstverantwortung in den Abteilungen (lokale DBAs)
 - Noch nicht komplett Data Guard fähig
 - Neue Konzepte bei übergreifenden Abfragen (vorher Schema, jetzt PDBs)

Zusammenfassung

Ist RAC Hochverfügbar?

- JA
 - System läuft trotz Ausfall einer Komponente weiter
 - Automatischer Reconnect der Anwendung möglich, ansonsten Unterbrechung
 - Patches lassen sich im laufenden Betrieb (Rolling) einspielen

Ist Data Guard Hochverfügbar?

- Nein
 - Das System läuft nicht weiter sondern muss explizit (Befehl) oder implizit (Observer) gestartet werden
 - Unterbrechung / Warten erforderlich (neuer Status der Datenbank)

Ist VMware Hochverfügbar?

- JA
 - VMware HA Funktionalität spiegelt einen Guest auf einen anderen Server
 - Beim Ausfall eines Servers wird der Guest automatisch zum aktiven Part
- NEIN
 - VMware vMotion verschiebt einen Guest auf einen anderen Server
 - Die Datenbank muss neu gestartet werden

Was wird benötigt?

- Frage 1:

„Wie lange darf die Anwendung maximal nicht zur Verfügung stehen?“

- Frage 2:

„Wie viel Datenverlust ist im Fehlerfall tolerierbar?“

- Frage 3:

„Wie lang darf ein Maintenancefenster sein bzw. wie viele sind erlaubt?“

Oracle Java Virtual Machine

- OJVM in der Datenbank
- Oracle Security Alert for CVE-2018-3110

CVE#	Component	Package and/or Privilege Required	Protocol	Remote Exploit without Auth.?	CVSS VERSION 3.0 RISK (see Risk Matrix Definitions)									Supported Versions Affected	Notes
					Base Score	Attack Vector	Attack Complex	Privs Req'd	User Interact	Scope	Confidentiality	Integrity	Availability		
CVE-2018-3110	Java VM	Create Session	Oracle Net	No	9.9	Network	Low	Low	None	Changed	High	High	High	11.2.0.4, 12.1.0.2, 12.2.0.1, 18	

MOS Note: 2217053.1

RAC Rolling Install Process for the "Oracle JVM Component Database PSU" (OJVM PSU) Patches

*"Beginning with the Jan2017 OJVM PSU patchset for 11.2.0.4 and for 12.1.0.2, this document defines a few specific situations where the OJVM PSU patchset can be postinstalled into each database **while the database remains in unrestricted "startup" mode**. This will allow a "Conditional Rolling Install" ability for the OJVM PSU patchsets for 11.2.0.4 and for 12.1.0.2"*

- Oracle 12.2. mit Release Update 180717 (Juli 2018)
- Bug 26430323 : ORA-07445[KKQGBPCHECKVALIDITYDP] WHEN SELECTING FROM DUAL
- “As you are running your 12.2.0.1 version on Windows platform, one off patches are not available for Windows. On Windows bug fixes are released in quarterly patch bundles. The bug will be fixed in Windows although cannot give you a timeframe when this will happen.”

- Überlegen Sie gut, was für Sie wichtig ist:
- Permanente Verfügbarkeit → RAC
- Geringer Datenverlust → Data Guard
- Einfache Verwaltung → VMware

- 11. Oktober DOAG Regio NRW: Oracle Datenbank Cloning mit Snapshots bzw. Copy On Write
- 12. Oktober DOAG Datenbank Webinar: Oracle RAC und NFS, ist das eine Alternative?

- 20.11. 14 Uhr Die Zukunft des DBS: über oder unter den Wolken?
- 20.11. 18 Uhr Open DB Mic Session
- 21.11. 14 Uhr Das Battle – Oracle vs. PostgreSQL
- 21.11. 16 Uhr Multitenant in der Standard Edition
- 22.11. 09 Uhr Oracle 18c XE
- 23.11. DOAG Schulungstag „Oracle XE, die kostenlose Oracle Datenbank“
- 21.11 10 Uhr Oracle Monitoring – Reicht der OEM (Fabian Greis)
- 21.11 15 Uhr Nur 3 Oracle Homes patchen bei 200 produktiven Datenbanken (Andre Lünsmann, Barmenia AG)

- Experten mit über 25 Jahren Oracle Erfahrung
- Firmensitz in Erftstadt bei Köln
- Spezialisten für
 - Oracle Datenbank Administration
 - Hochverfügbarkeit (RAC, Data Guard, Failsafe, etc.)
 - Einsatz der Oracle Standard Edition
 - Oracle Migrationen (HW, Unicode, Standard Edition)
 - Replikation
 - Performance Tuning
 - Datenbank Cloning (Delphix, Actifio, CloneDB)
- Fernwartung
- Schulung und Workshops (Oracle, Toad)



- E-Mail: johannes.ahrends@carajandb.com
- Homepage: www.carajandb.com
- Adresse:
 - CarajanDB GmbH
Siemensstraße 25
50374 Erftstadt
- Telefon:
 - +49 (22 35) 1 70 91 84
 - +49 (1 70) 4 05 69 36
- Twitter: [carajandb](https://twitter.com/carajandb)
- Facebook: [johannes.ahrends](https://www.facebook.com/johannes.ahrends)
- Blogs:
 - blog.carajandb.com
 - www.toadworld.com