

# 10: Advertentie Database Systems 2003

# Oracle 9i RAC

## Architectuur en toepassing

*Database systemen verwerken soms terabytes aan transactiedata per jaar. In de meeste gevallen gaat het hier om kritieke data voor zowel bedrijven, instellingen als individuen in de samenleving. Al deze transactiedata zal op het juiste moment, op het juiste tijdstip beschikbaar moeten zijn om als informatie te kunnen dienen.*

*Om Oracle database systemen continu beschikbaar te maken kunnen diverse strategieën gekozen worden, zoals Oracle Multimaster Replicatie of Oracle 9i Real Application Clusters.*

Oracle 9i Real Application Cluster (9i RAC) is een oplossing voor 'cluster database systemen', die op basis van hardware clusters functioneert. Een 'cluster' is een groep van onafhankelijke server systemen (nodes) die door onderlinge samenwerking als één krachtige server wordt gepresenteerd.

Het doel van Oracle 9i RAC is het bieden van significante verbeteringen in termen van beschikbaarheid (high availability), schaalbaarheid (scalability) en beheersbaarheid (manageability):

- *Beschikbaarheid* - database systemen moeten continue beschikbaar zijn
- *Schaalbaarheid* - database systemen moeten in de gelegenheid zijn om on-line te kunnen groeien, bijvoorbeeld door het toevoegen van meer opslag capaciteit of het toevoegen van een nieuwe cluster deelnemer in de configuratie.
- *Beheersbaarheid* - diverse database instances moeten beheerd kunnen worden alsof er één database instance aanwezig is.

**Oracle 9i RAC wil significante verbeteringen bieden in beschikbaarheid, schaalbaarheid en beheersbaarheid**

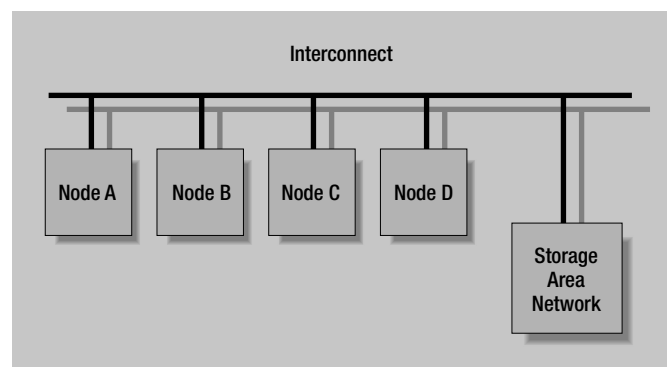
### Architectuur

De architectuur van een cluster database omgeving kan ingedeeld worden in twee componenten:

- Real Application Cluster System
- Real Application Cluster Software

De architectuur van de 'Real Application Cluster System' component bestaat uit minimaal twee of meer nodes die onderling verbonden zijn met een 'interconnect'. De nodes en interconnect zijn vervolgens gekoppeld aan een 'storage device' of 'shared disk subsystem', zoals een 'storage area network' (SAN).

Alle nodes in een cluster kunnen een eigen specifieke configuratie in aantal CPU's en geheugen hebben voor het verwerken van specifieke processen of applicaties. De interconnect dient als communicatiekanaal tussen de verschillende componenten in een cluster. Hiervoor wordt gebruik gemaakt van een 'inter-processes communication' (IPC) interconnect: een communicatiefaciliteit met een hoge bandbreedte en een lage latency.



Figuur 1. Architectuur van de Real Application Cluster System component

De hardware die als interconnect gebruikt kan worden kan zowel een Gigabit Ethernet verbinding als een fiber distributed data interface (FDDI) (glasvezel) zijn. Om geen single-point of failure in de cluster configuratie te hebben dient een back-up interconnect aanwezig te zijn in de cluster configuratie.

De architectuur van de 'Real Application Cluster Software' component bestaat uit de volgende onderdelen:

- Operating System - Dependent Clusterware
- Real Application Clusters Shared Disk componenten
- Real Application Clusters specifieke Daemon en Instance processen
- Global Cache en Global Enqueue Service

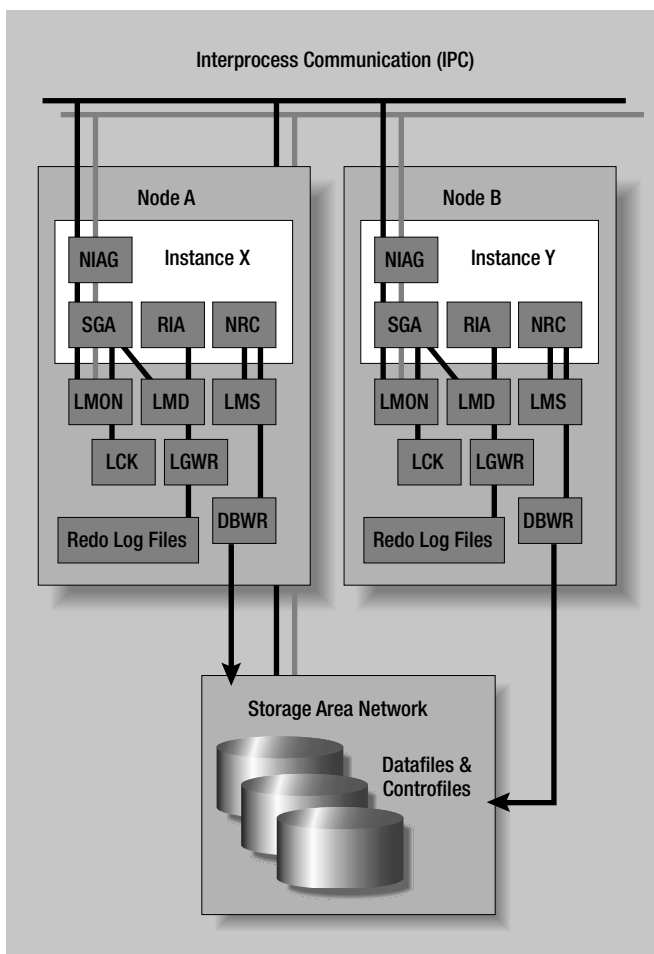
Real Application Clusters maken gebruik van operating system dependent (OSD) clusterware om het besturingssysteem te benaderen en voor cluster gerelateerde service afhandelingen, zoals de communicatie tussen de nodes voor het opstarten of afsluiten van een instance. De OSD beschikt over de volgende sub-componenten:

- Cluster Manager (CM)
- Node Monitor
  - De node monitor controleert de status van alle resources in de cluster, zoals de nodes, shared disks en interconnect hardware en software. De node monitor in de door Oracle geleverde Cluster Manager controleert daarnaast ook de Oracle Instances.

- Interconnect
  - De IPC software zorgt voor de communicatie tussen de nodes. Daarnaast wordt de IPC in een RAC gebruikt voor het transport van data blokken tussen de instances.

RAC-databases beschikken over dezelfde componenten als een single-instance database: één of meerdere control files, online redo log files, eventueel archive log files en datafiles. Doordat iedere node gebruik maakt van deze componenten, is het noodzakelijk dat deze bestanden op een shared disk worden geplaatst. RAC's kunnen hiervoor gebruik maken van cluster file system (CFS) of van RAW Devices.

***9i RAC configuraties zijn uitermate geschikt voor datawarehousing en online transaction processing (OLTP)***



Figuur 2. RAC - Specifieke Instance Processen

In een RAC omgeving bevindt zich een specifiek cluster proces op iedere node: global services daemon (GSD). De GSD communiceert met de cluster manager en handelt aanvragen van clients af om administratieve taken uit te voeren, zoals het opstarten of afsluiten van een instance. De database processen in een RAC omgeving komen overeen met de database processen in een single-instance omgeving: proces monitor (PMON), database writer (DBWRn), log writer (LGWR).

Aditionele RAC specifieke processen zijn:

- Global Cache Service Proces (LMSn)
  - Verzorgt de routing van berichten naar remote instances en beheert de toegang tot globale data blokken.
- Global Enqueue Service Monitor (LMON)
  - Monitoren van globale enqueues en resources in de cluster en het uitvoeren van global enqueue recovery operaties.
- Global Enqueue Service Daemon (LMD)
  - Beheren van globale enqueues en de toegankelijkheid van resources. Binnen een instance beheert het LMD proces inkomende remote resource aanvragen.
- Lock Proces (LCK)
- Diagnosability Daemon (DIAG)
  - Verzamelen van diagnostische gegevens van niet goed functionerende processen binnen de instances. Operaties die door dit proces worden uitgevoerd, worden opgenomen in de alert logfile

Groep	Platform	OS	9i RAC versie	Configuratie	Versie
Unix	HP-UX	11.0	9.0.1 64-bit	MC/Service Guard OPS Edition	11.09, 11.12, 11.13, 11.14
			9.2 64-bit	MC/Service Guard OPS Edition	11.13, 11.14
		11i	9.0.1 64-bit	MC/Service Guard OPS Edition	11.09, 11.12, 11.13, 11.14
			9.2 64-bit	MC/Service Guard OPS Edition	11.13, 11.14
	Sun Solaris (Sparc)	Solaris 8	9.0.1 32-bit/ 9.0.1 64-bit	Sun Cluster	3.0
				Veritas DBE/AC	3.5
			Fujitsu-Siemens PrimeCluster	V1.2A, V3.0, V4.0	
		Solaris 9	9.0.1 32-bit/ 9.0.1 64-bit	Sun Cluster	3.0
				Veritas DBE/AC	3.5
			Fujitsu-Siemens PrimeCluster	V4.0, V4.1	
	HP True64	5.1, 5.1a, 5.1b	9.0.1 64-bit 9.2 64-bit	True64 TruCluster	
	IBM AIX pSeries (RS/6000)	AIX 4.3.3	9.0.1 64-bit/ 9.2 64-bit	HACMP/ES CRM	4.4.x
AIX 5.1 (5L)		9.2 64-bit	HACMP/ES CRM	4.4.x, 4.5	
Linux	Red Hat	7.1	9.0.1	Oracle OSD Clusterware	9.0.1
		2.1 Advanced Server	9.0.1 / 9.2	Oracle OSD Clusterware	9.0.1 / 9.2
	SuSE	7.1 / 7.2	9.0.1	Oracle OSD Clusterware	9.0.1
	SLES 7	9.0.1/9.2	Oracle OSD Clusterware	9.0.1 / 9.2	
Windows	NT	4.0	9.0.1/9.2	Oracle OSD Clusterware	9.0.1 / 9.2
	2000 Server/ Advanced Server/ Datacenter Server	2000	9.0.1/9.2	Oracle OSD Clusterware	9.0.1 / 9.2

Figuur 3. Overzicht certificeringen Oracle 9i RAC

De 'Global Cache Service (GCS)' en 'Global Enqueue Service (GES)' zijn geïntegreerde componenten in een RAC die de toegang tot de shared database en shared resources in de database coördineren. Deze services onderhouden de consistentie en data integriteit. De GCS en GES op iedere instance maken gebruik van de IPC voor de communicatie tussen de instances en de overige componenten in de cluster. Indien deze services niet de benodigde informatie kunnen verkrijgen van een specifieke instance, zal de instance worden afgesloten om integriteit van de RAC database te bewaren.

### Configuratie certificering Oracle Real Application Cluster

Zoals uit de architectuur van RAC is te zien, functioneren hardware en software op cluster-omgevingen nauw samen met elkaar. Hierdoor worden bepaalde eisen en voorwaarden

gesteld aan de hardware om een RAC omgeving te kunnen inrichten. In het onderstaande schema is een overzicht opgenomen met de gecertificeerde hardware configuratie in combinatie met de Oracle 9i RAC software voor de meest gebruikte platformen: Unix, Linux en Windows.

### Toepassingsgebieden van Real Application Clusters

Het implementeren van RAC oplossingen kan op basis van diverse criteria en eisen een goede reden zijn om te gebruiken. Overeenkomende criteria en eisen in alle RAC implementatie trajecten zijn: hoge beschikbaarheid, schaalbaarheid en beheersbaarheid van de Oracle 9i Database. Een bijkomende voordeel van het gebruik van 9i RAC is dat bedrijfskritieke applicaties door de Cluster configuratie ook een hoge beschikbaarheid kent.

## RAC-databases beschikken over dezelfde componenten als een single-instance database

Oracle 9i RAC configuraties zijn uitermate geschikt voor Datawarehousing en online transaction processing (OLTP) omgevingen. In deze omgevingen wegen hoge beschikbaarheid, schaalbaarheid en beheersbaarheid eisen het zwaarst. Deze omgevingen kunnen verschillend van omvang zijn, zoals een databank voor regionale gegevens in een (deel)gemeente, verzekeringsmaatschappijen, banken en overige organisaties die over verschillende landen of zelfs over heel de wereld verspreid zijn.

### Een praktisch voorbeeld

Het internationale bedrijf X - dat we kennen van het artikel: 'Multimaster Replicatie' uit Optimize nr. 6 (december 2002) - wil een configuratie met 9i RAC gebruiken om de producten en diensten van haar ondernemingen in de vier landen (Nederland, Australië, Brazilië en Japan) up-to-date te houden. Hiertoe zal een datacenter opgericht worden in Brazilië, waar de RAC geplaatst zal worden. De applicatie die de regiokantoren en de hoofdkantoren in de diverse landen gebruiken is geheel in de 3-tier omgeving ontwikkeld en operationeel. Deze applicatie zal opgenomen worden in de cluster configuratie in de Datacenter, waardoor naast de database ook de applicatie een hoge beschikbaarheid kent. Bedrijf X maakte gebruik van HP-UX 64-bit systemen in de verschillende landen. Deze worden gecentraliseerd in Brazilië en opgenomen in een Cluster met MC/Service Guard OPS Edition met Oracle 9.2 Enterprise Edition. In figuur 4 is de opzet van 9i RAC voor bedrijf X opgenomen.

De manager van Bedrijf X ziet met de bovenstaande configuratie de volgende voor- en nadelen ten opzichte van de huidige Multimaster Replicatie configuratie. Dit kan hem helpen om een overweging te maken om wel of geen gebruik te maken van 9i RAC voor zijn onderneming.

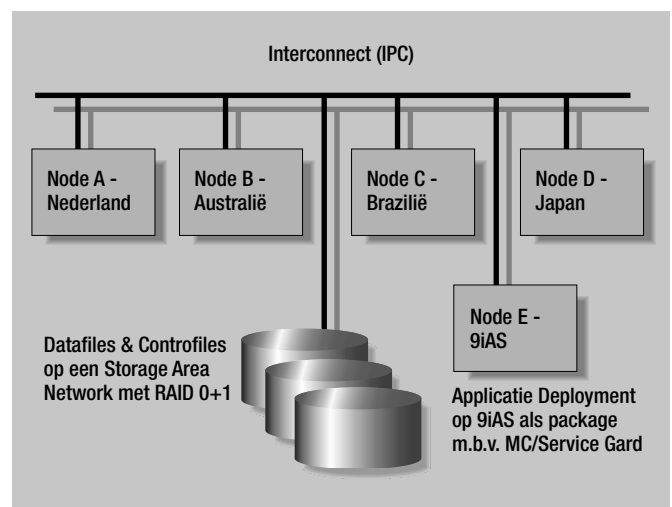
### Voordelen

- Door centralisatie van de servers naar Brazilië kunnen de volgende voordelen worden opgenoemd:
  - Er zijn geen additionele servers nodig om 9i RAC te kunnen gebruiken
  - Het fysieke onderhoud aan de servers gebeurt op één plaats: datacenter

- Het beheer kan worden gerealiseerd door minder personeel en alle servers zijn fysiek te benaderen door de beheerder(s).
- De communicatie tussen de nodes vindt met hoge bandbreedte plaats via de IPC
- Zowel de applicatie als de database voldoen nog steeds aan de eisen die bedrijf X had gesteld:
  - Hoge beschikbaarheid van zowel database als applicatie (7x 24 mogelijk)
  - Schaalbaarheid voor toekomstige groei
  - Beheersbaarheid (zowel lokaal als remote)
- Alle regiokantoren verspreid over de diverse landen hebben direct toegang tot de applicatie en database, waardoor conflicten met betrekking tot mutatie in data niet meer voor kunnen komen.
- Alle regiokantoren hebben direct de laatste gegevens over producten, diensten en klanten.
- Door gebruik te maken van de Oracle 9i Release 2 database heeft bedrijf X ook de voordelen van de 'new-features' op de database.

### Nadelen

- De communicatieverbindingen tussen de diverse landen en regiokantoren moeten eventueel aangepast of uitgebreid worden in verband met het gebruik van bandbreedte en redundantie.
- De servers staan op één centrale plaatst, waardoor fysieke uitwijk niet direct mogelijk is. Dit kan worden opgevangen door bijvoorbeeld een stand-by database in Australië op te zetten en te laten bijwerken met archive logfiles vanuit Brazilië.
- MC/Service Guard kennis is noodzakelijk om de cluster te kunnen beheren. Dit kan betekenen dat bedrijf X hiervoor iemand met deze kennis moet aannemen indien dit niet aanwezig is of met een derde partij een 'Service Level Agreement' (SLA) moet afsluiten om het beheer uit te laten voeren.



Figuur 4. Voorbeeld van een 9i RAC omgeving voor bedrijf X

## Voorzieningen

Door de bovenstaande configuratie te implementeren heeft bedrijf X voorzieningen getroffen voor:

- Fail-over / High Availability
  - Indien een database instance op een node niet goed functioneert, zal door de GCS en GES services de instance worden afgesloten en door middel van de Cluster Manager de beheerder op de hoogte worden gesteld. De database is ondanks het niet goed functioneren van een instance toch beschikbaar voor gebruikers.
  - Indien een hardware storing optreedt op één van de nodes, blijven zowel de applicatie als de database beschikbaar voor gebruikers.
  - Een storing op de node waar de applicatie op draait, kan met behulp van MC/Service Guard om een 'package switch' forceren, waardoor de applicatie op een andere node komt te draaien. Hierdoor merken gebruikers nauwelijks dat er problemen zijn en is zowel de applicatie als de database beschikbaar voor gebruikers. Na herstel van de storing kan de applicatie package weer handmatig terug worden geplaatst.
  - Doordat de interconnect redundant is uitgevoerd in de cluster omgeving is onderlinge communicatie gewaarborgd.

***Als de interconnect redundant is uitgevoerd in de cluster omgeving is onderlinge communicatie gewaarborgd***

- Load Balancing
  - Door de aanwezigheid van een aantal servers in de cluster, kunnen geheugen en CPU efficiënter gebruikt worden door load balancing toe te passen. Hierdoor wordt de draagkracht verdeeld over meerdere servers en kunnen meer gebruikers tegelijk de applicatie en database gebruiken.
  - De instances op de diverse nodes worden automatisch in een load-balance configuratie opgenomen in de 'Transport Network Service (TNS)' listener configuratie. Hierdoor wordt de RAC load balance feature automatisch aangepast indien een wijziging in de cluster plaatsvindt, zoals het toevoegen van een nieuwe node in de cluster.
- Schaalbaarheid
  - Bedrijf X kan met de huidige hardware overschakelen naar een RAC omgeving. In de toekomst kan hij bij uitbreiding van zijn services in andere landen additioneel nodes toevoegen.

- Naast de nodes in de cluster is ook de opslagvoorzieningen schaalbaar. De SAN kan in de toekomst uitgebreid worden door een nieuwe SAN toe te voegen in de cluster. Hierdoor kan de database blijven groeien zonder de bestaande configuratie te reorganiseren.

```
# LISTENER.ORA.ALPHA1 Network Configuration File:
/u01/app/oracle/product/9.0.1/
network/admin/listener.ora.alpha1
# Generated by Oracle configuration tools.
LISTENER =
  (DESCRIPTION_LIST =
    (DESCRIPTION =
      (ADDRESS_LIST =
        (ADDRESS = (PROTOCOL = TCP)(HOST = alpha1)(PORT = 1521))
      )
    )
  )

SID_LIST_LISTENER =
(SID_LIST =
  (SID_DESC =
    (SID_NAME = PLSExtProc)
    (ORACLE_HOME = /u01/app/oracle/product/9.0.1)
    (PROGRAM = extproc)
  )
  (SID_DESC =
    (SID_NAME = db9i1)
    (ORACLE_HOME = /u01/app/oracle/product/9.0.1)
  )
)

# LISTENER.ORA.ALPHA2 Network Configuration File:
/u01/app/oracle/product/9.0.1/
network/admin/listener.ora.alpha2
# Generated by Oracle configuration tools.
LISTENER =
  (DESCRIPTION_LIST =
    (DESCRIPTION =
      (ADDRESS_LIST =
        (ADDRESS = (PROTOCOL = TCP)(HOST = alpha2)(PORT = 1521))
      )
    )
  )

SID_LIST_LISTENER =
(SID_LIST =
  (SID_DESC =
    (SID_NAME = PLSExtProc)
    (ORACLE_HOME = /u01/app/oracle/product/9.0.1)
    (PROGRAM = extproc)
  )
  (SID_DESC =
    (SID_NAME = db9i2)
    (ORACLE_HOME = /u01/app/oracle/product/9.0.1)
  )
)
```

Figuur 5. Voorbeeld van een listener configuratie in 9i RAC

## Tools voor Real Application Clusters

De database componenten in RAC's kunnen worden geïmplementeerd met behulp van Oracle Universal Installer (OUI), Oracle Net Configuration Assistant en Database Configuration Assistant (DBCA). Met Oracle Universal Installer (OUI) wordt de Oracle 9i Software geïnstalleerd en ge-update. Oracle Net Configuration Assistant helpt mee met de configuratie van de listeners op de diverse nodes, waarvan een voorbeeld is weergegeven in figuur 5.

De Oracle Database Configuration Assistant zorgt voor de begeleiding om de database in de cluster aan te maken. Wanneer de database is aangemaakt kan op de volgende methode gecontroleerd worden of de database instances op alle nodes functioneert.

```
oracle db9i1> sqlplus /nolog
SQL*Plus: Release 9.0.1.0.0 - Production on Wed Feb 13 14:16:10
2002 (c) Copyright 2001 Oracle Corporation. All rights reserved.
SQL> connect / as sysdba
Connected.
SQL> select * from v$active_instances;

INST_NUMBER INST_NAME
-----
1 alpha1.domain.com:db9i1
2 alpha2.domain.com:db9i2
```

Figuur 6. Controle van 9i RAC database instances op diverse nodes

Het primaire gereedschap om een RAC omgeving te beheren is de 'Server Control (SVRCTL) Utility'. SRVCTL beheert verschillende configuratie gegevens die door andere tools gebruikt worden, zoals Oracle Enterprise Manager en Oracle Intelligent Agents.

Er zijn twee groepen van administratieve taken die met behulp van SRVCTL uitgevoerd kunnen worden:

- Cluster Database Taken
  - Starten en stoppen van de Cluster database
  - Starten en stoppen van de Cluster database instances
  - Het opvragen van de status van de cluster database of cluster database instance
- Cluster Database Configuratie Taken
  - Toevoegen en verwijderen van cluster database configuratie gegevens
  - Toevoegen aan en verwijderen van een database instance uit de cluster database configuratie
  - Verplaatsten van database instance in een cluster database configuratie
  - Het activeren en uitzetten van de omgeving van een instance in een cluster database configuratie

***Alle nodes in een cluster kunnen een eigen configuratie hebben wat betreft aantal CPU's en geheugen***

- Het activeren en uitzetten van de gehele cluster database in een cluster database omgeving.

## Conclusie

Het gebruik van Oracle 9i RAC is een goede oplossing voor datawarehouse en OLTP systemen, waarbij hoge eisen worden gesteld aan hoge beschikbaarheid, schaalbaarheid en beheersbaarheid van database systemen. Oracle 9i RAC omgevingen zijn gecentraliseerd en hebben daardoor zowel voor- als nadelen.

Om een Oracle 9i RAC omgeving te kunnen inrichten zijn minimaal de volgende componenten noodzakelijk:

- Cluster hardware, bestaande uit een shared opslagsysteem (SAN), interconnect (IPC), en minimaal twee server nodes.
- Gecertificeerde besturingssysteem en clusterware voor het beheren van de cluster
- Oracle 9i Enterprise Edition Software CD

In het volgende deel over 9i RAC zal het beheer van een 9i RAC omgeving onder de loep genomen worden. Onderwerpen die hierin aan bod zullen komen zijn onder meer het beheer van 9i RAC met Oracle Enterprise Manager en back-up & recovery met RMAN in een 9i RAC omgeving.

**Ing. Shaktiewant Adhien** is Oracle Applications DBA bij Qualogy Applications B.V. (e-mail: [sadhien@qualogy.com](mailto:sadhien@qualogy.com))