# Understanding Oracle RAC on Sun Cluster 3.x

Kristien Hens

Technology System Engineer

Sun Microsystems Belgium

Mature Foundation of Success with Sun Cluster 3 and Oracle

- More than 36,000 Sun Cluster nodes installed and growing
  - Consistent sales growth qtr/qtr

- Sun Cluster is the most popular Solaris platform for Oracle RAC deployments

- Thriving Sun-Oracle relationship
  - Joint technology Development
  - Joint Support Centers
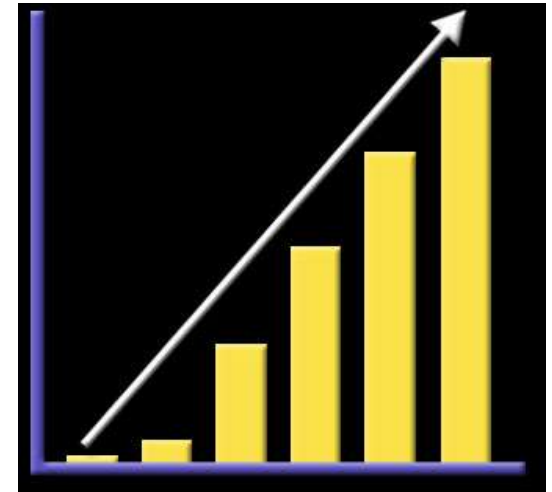
Sun Cluster 3 and Oracle

- Sun Cluster and Oracle co-exist in 2 ways:
  - HA Oracle
    - Oracle is 'cluster unaware'
    - Listener and Instance started and stopped through agent
    - Instance runs on one node, fails over to another
  - Oracle Real Application Clusters (9i or 10g)
    - Oracle is 'cluster aware'
    - Instances running on different nodes

HA Oracle

- 2 Resource Types (agents)
  - SUNW.oracle_server
  - SUNW.oracle_listener

- Instance and Listener run on one node, can failover to other node
  - When node dies
  - When Oracle server fault probe decides this

HA Oracle

- oracle_listener resource
  - Never failed over
  - Restarted when it fails over to other node

- One of our most elaborate agents
  - Fault monitor can react to many failures
  - Or you can decide not to probe anything

Oracle Real Application Clusters

- Oracle instances on all nodes

  – Instances synchronise information with each other through private interconnect

 – Instances never fail over to other nodes

Oracle Real Application Clusters

- Oracle RAC communicates with Sun Cluster to get the following information:

  - Cluster Membership

    - Node has left or joined cluster – necessary resources need to be set up or deleted in RAC

  - Cluster Interconnect

    - Interconnect used for RAC's cache fusion information

Oracle 9i RAC

- Multiple instances on different nodes
- Need for interaction with Cluster framework
  - Information about Cluster interconnect
  - Information about Cluster membership
- Oracle UDLM communicates with Sun Cluster ucmm
- Sun Cluster ucmm executes reconfiguration steps
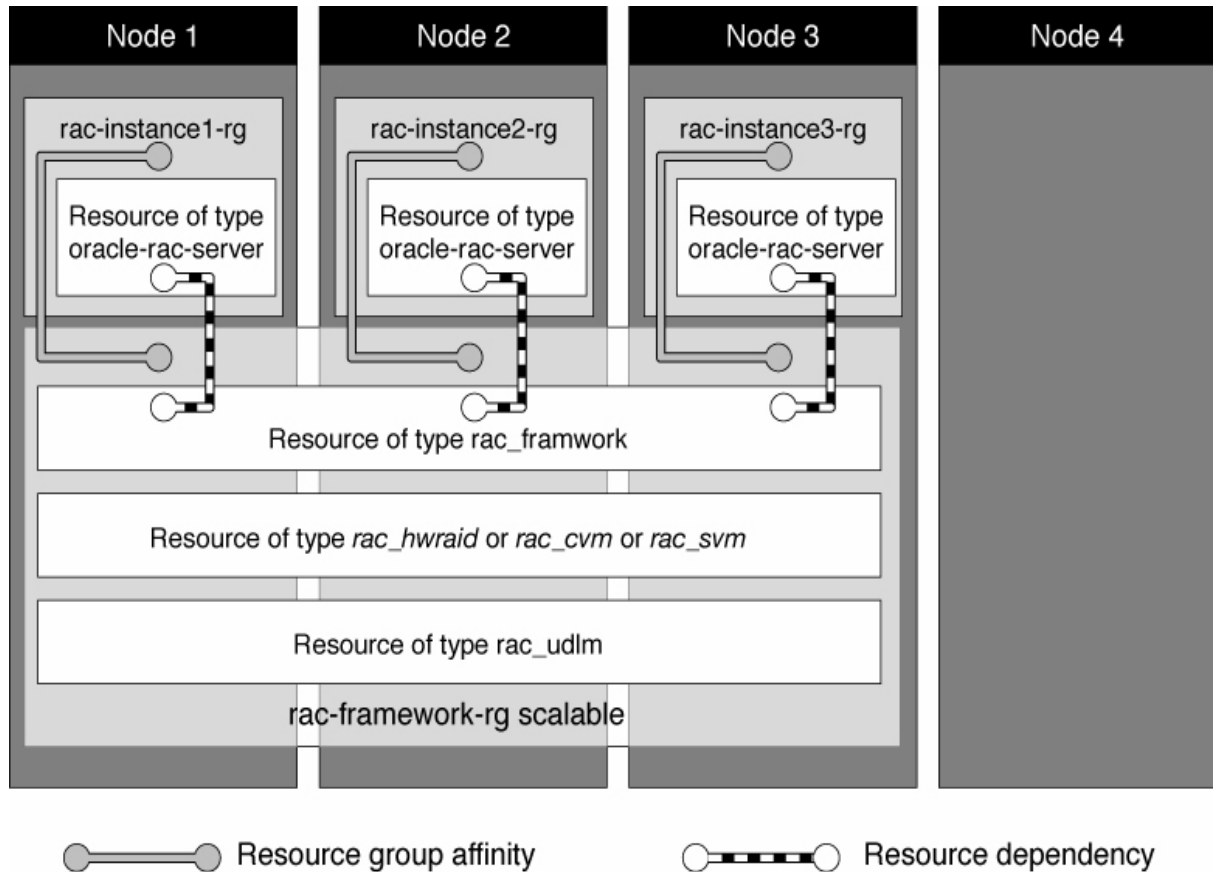
Oracle RAC on Sun Cluster
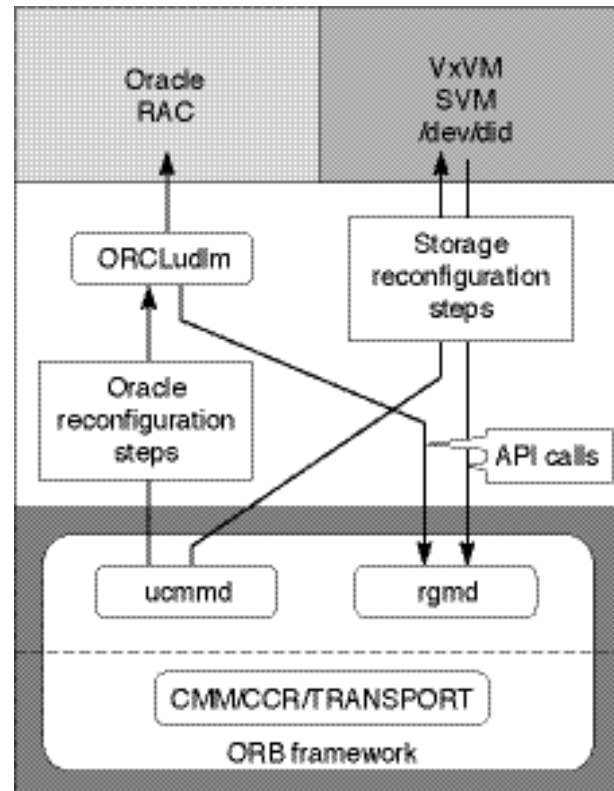
Improved Manageability for Oracle 9i RAC

- As of Sun Cluster 3.1 update 1: RAC manageability Framework

  - Resource Types to check if necessary daemons are running and if necessary reconfiguration steps have finished

  - Does NOT start/stop ucmm or dlm, just checks them

Improved HA for Oracle 9i RAC

- As of Sun Cluster 3.1 update 3: SUNW.oracle_rac_server resource type

  - Start and stops Oracle instances

  - No fault probe necessary

  - Leverages new Resource Dependencies model in update 3

Oracle RAC Cluster Packages

Improved Flexibility for Oracle RAC Cluster Packages

- Main goals:

  - Supply supported way to change properties

  - Status of RAC framework daemons visible in Sun Plex Manager, scstat...

  - Easy to install RAC on a subset of nodes

Oracle 10g RAC

- Ships with Cluster of its own (CRS)

  - Still Cluster layer necessary for RAC

  - Userland cluster

- Needs the Cluster Ready Services (CRS) daemons running

  - Will get private interconnect from CRS

Oracle 10g RAC

- Best in combinationwith Sun Cluster

  - CRS will rely on SC for cluster membership info, split brain detection

  - SC more reliable

Oracle RAC datafile options

- Currently supported:
  - VxVM with Cluster option (CVM)
  - Raw /dev/did devices
    - When there is Path and Disk redundancy
      - Not supported for all storage options
    - Functionally equivalent to CVM
    - Better reconfiguration times

Oracle RAC datafile options

- Future options:
  - Shared QFS
    - High-performance filesystem
    - Shared access to datafiles via SAN
    - Agent supplied with QFS 4.2
    - Advantages:
      - Near-raw performance
      - Look and feel of a filesystem

Oracle RAC datafile options

- ## Future options: Multi-Node Disk-Sets

- SVM with cluster functionality (shared disksets)

- Functionality similar to Veritas CVM

    - One master node, other nodes are slaves

    - Only master node updates metadbs

    - Freely included in Solaris as of Solaris 9 update 3
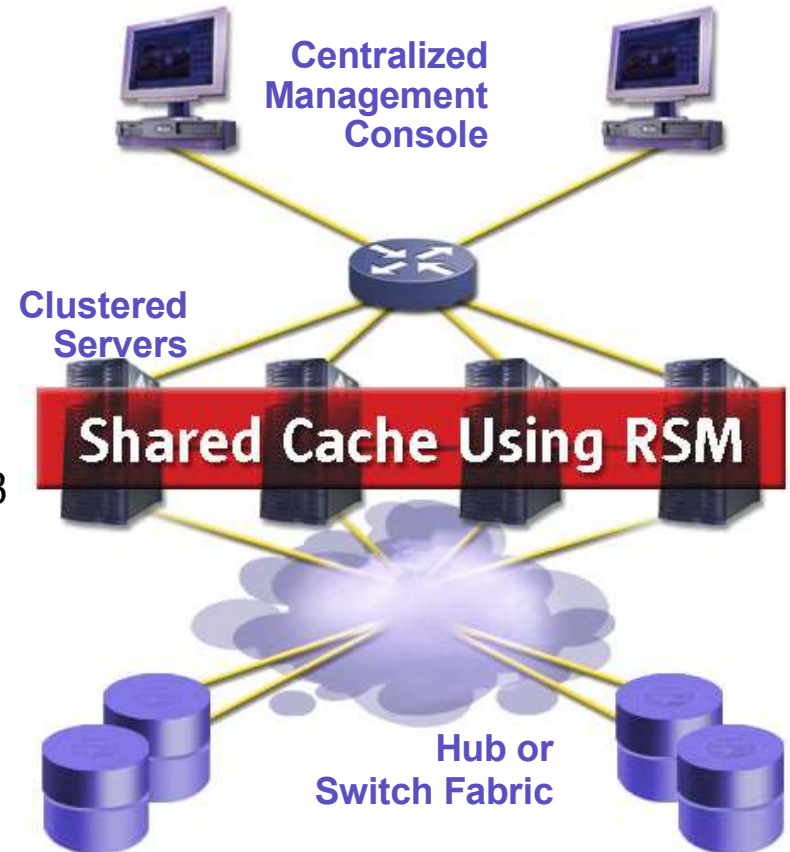
Oracle RAC Client Connections

- Oracle 9i RAC:

  - Use TAF on client side

    - Client will reconnect to other node when instance has died

  - Use shared addresses on all nodes

    - To avoid long failover times upon node crashes

    - Client will not reconnect, but since the SA will failover, client will end up on the other node

Benefits of Sun Cluster and Oracle RAC

Reasons for Sun-Oracle mutual market success

- Low Cost without sacrificing performance

  - No charge for Remote Shared Memory technology – high performance at low cost

  - Supported on 1-106 processor systems

  - ½ price promo unveiled in October 2003 savings of more than $10K with RSM

- Proven, time-tested solution

  - Over 10yrs of parallel DB experience

- Large customer base

  - British Telecom, Canal+, Travelocity, France Telecom, BACS, UHHS, PICC

Centralized Management Console

Clustered Servers

Shared Cache Using RSM

Hub or Switch Fabric

Sun Cluster 3 and Oracle10g Positioning

- Several advantages of deploying 10g RAC on Sun Cluster

  - Mature HA framework needs mature clustering framework

  - Possibility to deploy other HA applications on the same cluster

  - Flexible API

  - No need for identical hardware on both nodes

  - SPOFs exist in 10g without SC configurations

Understanding Oracle RAC on Sun Cluster 3.x

Sun microsystems