11i on RAC: Sweating the Details

Practical overview of Technical Details of Implementing RAC for 11i E-Business Suite

Acronyms and Glossary Terms

- RAC: Real Application Clusters
- OATM: Oracle Application Tablespace Model
- CRS: Cluster Ready Services
- ASM: Automatic Storage Management
- 3rd Party Clusterware: HP Servicegaurd, Veritas, SunClusters
- TAF - Transparent Application Failover
- PCP - Parallel Concurrent Processing
- OMF - Oracle Managed Files
- OCR - Oracle Cluster Registry

RAC - Making a Business Case

- Real Application Clusters (RAC) is a single database running multiple instances across a cluster of servers
- RAC provides several business case scenarios
  - Availability
  - Scalability
  - Performance
  - Costs

RAC - Making a Business Case

- Business case should always take into account the cost-benefit of technical complexity
  - RAC ain't cheap
    - Commodity Hardware and Storage is cheap
    - Oracle expertise and License isn't cheap
  - RAC ain't easy
    - Easy to set init.ora parameter to "cluster=TRUE"
    - Is not easy to lay foundation for sharing storage, network, and server configuration

RAC on 11i - Making a BETTER Business Case

- 11i EBS implementations can realize these benefits of going RAC
  - Scalability when on smaller Linux boxes
  - Performance
    - when ASM is used
    - When OLAP can be removed
- Hardware Cost savings
  - Major cost savings realized in using Linux based servers
  - Significant storage costs using NAS over SAN
- Note: TAF not available on 11i EBS

RAC Technical Considerations

- Should perform a Technical Readiness Assessment
- Committed to the technology??
  - Server Platform
  - Storage Platform
  - ORACLE
- Complexity affects entire enterprise
- Resources available -- all cost $$$'s
  - Hardware -- do you have the servers, network, and storage infrastructure
  - Expertise
  - Training -- no class or certification can prepare you for real life operational scenarios

Technical Drill-down: Hardware

- Linux based (Intel / AMD / Itanium)
  - Most documentation/support available from Oracle
  - Very Good Hardware Offerings by all vendors
  - RH4/5 & OEL
- Proprietary
  - IBM's AIX on Power
  - HP-UX's Itanium
  - Sun Solaris
    - UltraSparc T2
    - T1 Coolthreads
- Blades and OracleVM -- another presentation
- RAC requires additional memory on all platforms

Technical Drill-down: Storage

- Platform
  - NAS and iSCSI (Highest Recommendation)
    - Best suited for sharing storage
  - SAN -- highest cost
    - FibreChannel throughput
    - Multipathing
- Configuration
  - CFS
  - Raw/Block
    - Devices shared between RAC'd database nodes
      - Physical
      - Logical
  - Always Stripe And Mirror Everything (SAME)

Technical Drill-down: Network

- Speed
  - 1GigE vs. 10GigE vs. Infiniband
  - Dedicated switches
- Hardware
  - Redundant NIC's and switches
  - Bonding / Teaming / APA
- Public IP's
- Private IP's (interconnect)
- VIP's

11i EBS Considerations

- ATG RUP Levels
- Architecture / Topologies
  - DB nodes
  - Apps nodes
- EBS Modules and Access
  - OLTP Functions
    - Module affinity per RAC node
    - Can single node handle requirement of 1 module
      - Heavy Inventory or OM?
  - OLAP Functions
- Customizations

RAC node requirements

- For Failover, quorum must be > 2

- RAC Quorum requirement
  - Keeping Quorum requires 3 nodes
    - 2 nodes: 1 node failure, remaing node != quorum
    - 3 nodes: 1 node failure, 2 nodes keep quorum
  - Personal Favorite: 5 nodes*
    - 3 OLTP nodes
    - 2 OLAP nodes

11i - Mixing Workloads?

- On-Line Transaction Processing (OLTP) Internal Access / Performance

  - External

    - iStore or iProcurement 24hour requirement

- On-Line Analytical Processing (OLAP)

  - RAC permits the use of Separate instance for reporting

  - Failover requirement (uptime SLA?)

Instance Strategy

- For RAC, do you have the Required Hardware and Real Estate
  - Require a mirror RAC instance for integration testing
    - Model Office
    - Integration Testing cycles increase
  - Test and Dev do not require RAC, but can/should be ASM
  - RMAN standard may require shared file system large enough for multiple online backups

Total commitment to testing

- Load Testing
  - LoadRunner
  - SwingBench -- customize to fit 11i
  - Orion Storage Tests (no HPUX)
- Failover Testing
  - Multiple scenarios
    - Server failure
    - Network failure
    - Instance failure

Software Config: O/S

- Kernel Parameters
- Clock sync between servers
- Linux: hangcheck timer
- 3rd party clusterware
  - I/O Fencing to prevent "split-brain" issues
  - VolMgr for Logical Raw Volumes
    - Enables Multipathing
  - NIC card failover
  - CRS Control disabled/deferred

Software Config: Clusterware

- Oracle Clusterware
  - Separate ORA_CRS_HOME
  - 3rd Party Clusterware
    - Integration
    - Certification
  - Host Equivalancy
  - Network
  - Access to shared devices for OCR and Voting
  - cluvfy

Software Config: Data

- Cluster file system (CFS)
- Shared Raw
  - Physical Raw
  - Logical Raw
- ASM
  - Best Practice (even without RAC)
    - Put on shared raw (physical or logical)
  - Abstraction to physical
    - Asmlib = linux
    - Asmcmd

ASM Config Notes

- Separate ASM Home
- Separate ASM Listener
    - On 11i, autoconfig creates LISTENER_<servername>
    - Using the LISTENER_<servername> created by ASM isn't sufficient
    - Rename default listener from LISTENER_<servername> to ASMLISTENER_<servername>
- Mirroring (High, medium, low)
- Use spfile

ASM operations

- RMAN Proficiency Required
- Asmlib (Linux) and asmcmd for command line reference
- Adding datafiles
  - Create datafiles using "size" & "autoextend off"
  - Watch growth via tool
    - OEM
- +FRA
  - Size this for Online Redo, and archivelogs
- Archivelogs
  - Use +FRA
  - Run RMAN backup frequently with remove option since you don't have easy visibility to growth

ASM caveats and alerts

- Only one ASM instance per server
    - Shared by multiple instances on same server
- Separate ORACLE_HOME's
    - Patching separate
    - If only using one instance, you can install this in same ORACLE_HOME
- Log writer will be write (by default) in synchronous mode
- Physical standby based on a file system must watch out for any of the following affecting the recovery
    - New datafiles added to primary instance
    - New Oracle directories created on primary instance
    - Utl_file related operations that are ASM based

RDBMS configuration

- 10gR2
- OATM
- Each RAC instance requires a separate UNDO tablespace
- Usually do not register with CRS so root cause analysis can be done before startup after failure
- Use spfile to share configuration changes between instances
- Use ifile for custom settings since autoconfig may overwrite pfile

ASM Migration

- Diskgroups
  - Strategy -- size disks for diskgroup according to growth strategy.
  - RBAL factor
  - Size Diskgroups appropriately
    - +DATA
    - +FRA (FlashbackRecoveryArea)
  - OMF
    - Onlinelogs
    - Datafiles
    - Controlfiles
    - tempfiles

RAC configuration

- Rconfig ??
- Node1: Adcfgclone
  - Dbtier
  - Recreates control file
  - Fails due to bug 5936694 (Metalink Note 421121.1)
  - Adpreclone dbtier
- RAC enable db1
  - Cluster = TRUE
  - Add logfile thread
  - Add untotbs2
  - Enable thread 2
- Node2: Adcfgclone

Apps Tier

- Autoconfig to recognize new db nodes
  - Uses fnd_nodes & fnd_database_instances
  - Creates new tns entries for DBn
- Edit context file
  - Two_tasks = DB_806_balance
  - Cp_two_task = DB1 or DB2 for PCP

Operations

- OEM
- Network monitoring
  - Interconnect: watch for max < 70%
- Clusterware
  - Srvctl & crsctl
  - Crs_stat -t
- ASM
- RDBMS
  - Monitor gc wait events

Backups / Recovery / Clone

- OCR and Voting disk backups
- RMAN
  - RMAN files on shared or mirrored drives between servers (source and destination)
- Adcfgclone
  - dbtier

Cutover Options

- **Big Bang**
  - Do CRS Install, 10gR2 upgrade, ASM Migration, RAC config across multiple nodes in single outage window

- **Incremental**
  - Upgrade to 10gR2
  - Implement ASM
  - Implement CRS & Configure RAC across multiple nodes

- **Pre-configure RAC and ASM as Standby**
  - Install CRS, ASM, and 10gR2 in RAC configuration on new hardware
  - Create STBY instance in clustered environment
  - Rollforward until cutover
  - Switchover and do RAC configuration

Helpful options

- Init.ora
  - _fast_start_instance_recovery_target
- RDBMS and 11i Database listener under CRS control
  - Reboot loops?
  - Root cause analysis before startup?

Additional Recommendations

- Define SLA's for performance and availability for each service or application
- Use Grid Control to manage CRS, ASM, Database and Application.
- All changes to the production environment must be previously tested on a separate environment
- Apply changes to one system element at a time, first on test then on production.
- Keep a detailed change log
- Implement services where possible to manage workload for OLAP and customizations
- Configure OSWatcher to have handy information about the OS layer in case of need, see on Metalink Note:301137.1, OS Watcher User Guide
- Configure RDA to have handy information to Oracle Support in case of need, see on Metalink Note:314422.1, Remote Diagnostic Agent Getting Started. RDA 4.5+ includes RAC data collection capability. It can be used in place of RAC Diagnostics tool RACDDT.
- Establish support mechanisms and escalation procedures.
- Make sure DBA's have well tested procedures about how to deal with problems and collect required diagnostics.
- Use Racdiag.sql to check database during normal behavior and be able to compare results, see on Metalink Note:135714.1

Alternatives to RAC

- OracleVM
- Larger Server w/ replication for OLAP reporting
  - NAS gives manual failover

Wrapping it Up

- Loose ends

- Closing Remarks

- Q & A

- Contact Me: kurt.forshee@gmail.com

- Thank You!!