

NEW FEATURES IN RAC AND CLUSTERWARE 10G RELEASE 2

Dan Norris, [dnorris\(at\)itconvergence.com](mailto:dnorris@itconvergence.com), IT Convergence

INTRODUCTION AND OVERVIEW

Oracle 10g Release 2 introduced many new features with RAC and Clusterware (formerly known as Cluster Ready Services). Many of the new features are significant because they are entirely original and are not simply extensions to prior functionality. Here you will find technical information about several of these new features including some options for putting the new features to use.

BRIEF REVIEW OF ORACLE 10G RELEASE 1 FEATURES

In order to provide a frame of reference for those only familiar with 9i features, this section reviews a list of significant 10g Release 1 features.

One of the most significant features for 10g RAC environments was the introduction of Oracle Clusterware for all platforms. The Oracle Clusterware is mandatory for all RAC environments starting with 10g and it can be integrated with some other cluster software if desired.

Also significant for 10g databases is the introduction of automatic storage management (ASM). ASM is capable of directly managing database storage and eliminates the need to be concerned about placement of individual files.

ORACLE CLUSTERWARE NEW FEATURES

The clusterware introduced with 10g Release 2 is a major leap forward. It now includes the ability to monitor and operate on custom processes. Many Oracle experts anticipate that Clusterware modules or extensions for managing other Oracle and possibly non-Oracle applications like Oracle HTTP Server will appear in future releases.

In keeping with the company's past practices, Oracle Cluster Ready Services (CRS) from 10g Release 1 was renamed to Oracle Clusterware in 10g Release 2.

APPLICATION MONITORING AND FAILOVER

A very simple application programming interface (API) is now available to allow you to build programs to monitor, start, stop, and restart your custom applications. The API consists of three methods in a single script: start, stop and check.

The start method contains all the commands or steps necessary to start the external application so that it remains running. The stop method performs all the actions necessary to stop the process(es). Finally, the check method is responsible for determining if the application is available or not. The check method could simply look for the process in the process table, or it could attempt to retrieve a URL or just about anything else you wish.

This functionality can be very effective for monitoring and maintaining any application that you can monitor today using scripts or programmatic interfaces.

FILE REDUNDANCY FOR OCR AND VOTING DISKS

Another important new feature for Oracle Clusterware is the install-time choices for multiple voting disk and multiple Oracle Cluster Registry (OCR) locations. You can now choose up to three locations for voting disks and up to two locations for OCRs. This configuration is requested during the Oracle Clusterware installation dialogs.

With multiple locations for OCR, there is less chance of losing the OCR or suffering from filesystem corruption. There must be an odd number of voting disk locations in order to avoid a tie, so you must specify one or three voting disk locations.

RAC NEW FEATURES

Besides the important features in Clusterware, Oracle also enhanced many RAC database features and introduced several new ones.

CLONING SUPPORT

With the great improvements to Oracle Grid Control, it is now easier than ever to manage all servers together as though they were a single entity. The ability to clone a cluster node makes operations such as adding a node to the cluster very easy and much less prone to error than in previous releases.

Cloning is a practice commonly used within Oracle Applications environments. Cloning defines a set of steps in which a copy of the ORACLE_HOME and possibly database are made and all necessary changes are made in order to allow the new copy to function as the original. This usually involves changes to configuration files that specify the hostname or IP address, and other references to the original source's instance name or database name.

With 10g Release 2, Oracle allows cloning for RAC environments so that you can quickly and easily add nodes to the cluster.

SILENT INSTALL SUPPORT

Oracle has added features in this release allowing silent installations. The deployment framework also allows nodes to be added quickly and easily. While this feature may seem like a minor improvement for two- or three-node clusters, it significantly improves manageability for clusters with more than 20 nodes.

SEPARATE ORACLE_HOME FOR ASM

Oracle recommends that the ASM installation (if you use it) be performed into its own ORACLE_HOME directory. This allows more flexibility, manageability, and higher availability. When ASM is installed into a separate ORACLE_HOME (referred to as an ASM_HOME), it is not affected by operations that patch or otherwise require downtime for products in the database ORACLE_HOME. Additionally, if multiple instances can share ASM disk groups, having a dedicated ASM_HOME allows for storage consolidation for multiple instances.

The installer and DBCA tools have been enhanced to allow for installation and configuration of ASM independent of a database installation or configuration. The DBCA has a separate option for ASM operations allowing the DBA to manage storage in ASM without Oracle Enterprise Manager (OEM) or command-line tools. Of course, OEM and command-line tools are still available and have been enhanced to address all the new features in ASM as well.

Another significant ASM feature is the migration utility included with Oracle Enterprise Manager Grid Control 10g Release 2. This utility will migrate file system or raw device storage to ASM on the same host.

ORACLE ENTERPRISE MANAGER IMPROVEMENTS

Oracle Enterprise Manager can now be used to add and delete instances from a RAC database. It provides a much easier way to ensure that all the necessary steps are completed in the proper order. OEM is available in both Grid Control and Database Control.

OEM also allows a conversion to RAC from a single-instance database. Conversions to RAC are generally more complicated than just adding an instance, so this OEM facility will help ensure that all necessary steps are followed. The database installation includes a command-line utility to perform the conversion and that utility can be used directly instead of using it via OEM.

COMMIT PROPAGATION

In 10g Release 1, Oracle made significant improvements that allowed for very fast propagation of commits. Commit propagation is the operation that informs all instances in the cluster that data has been committed and invalidates the buffers in the each instance's cache if necessary. Prior to 10g Release 2, the speed of propagation was governed by the max_commit_propagation_delay parameter. Setting this parameter to 0 forces commits to be broadcast to all instances in the cluster immediately. This activity could cause performance degradation if used in excess. Leaving this parameter at its default value of 700 centiseconds allowed Oracle to use a high-performing Lamport scheme for commit propagation.

The max_commit_propagation_delay parameter has been officially deprecated in 10g Release 2, and the default behavior is to broadcast on commit. Due to performance enhancements in the algorithm, Oracle has made this change while maintaining the high performance standards required by customers.

NETWORKING NEW FEATURES

Some of the most exciting new features in the RAC environment have been added to the networking capabilities.

IMPROVED FCF INTEGRATION

Fast Application Notification (FAN) publishes information about high availability events. These events can be used by connection pool to provide Fast Connection Failover (FCF). Oracle has integrated this functionality directly in the ODP.NET and OCI clients in order to shorten the delay between a failure and a reconnection. The FAN notifications are provided by an Oracle Notification Service (ONS). ONS is a facility that these clients can be configured to access in order to receive the event notifications.

LOAD BALANCING ADVISORIES

Starting with 10g Release 2, Oracle Database has the ability to send information about its load back to connection pools via FAN notifications. These advisories can be used by some Oracle clients (like JDBC and OCI-based clients) to help determine where to connect and where to send any new SQL statements. The TNS listeners in the cluster receive the same advisories and use them to help direct incoming connection requests to the least-busy instance in the cluster.

Oracle has integrated the advisories with JDBC and OCI clients, but the advisory feature includes an API that can be used by any application to receive the advisory data and process in any way necessary.

RUNTIME CONNECTION LOAD BALANCING

Additionally, with JDBC and ODP.NET clients using connection pools, runtime load balancing can be employed to ensure that every database request is routed to the least-utilized instance in the cluster. On a technical level, the connection pool method for "getConnection" has been modified to obtain a connection to the least-loaded instance from the connection pool. This new functionality is tightly integrated with services, automatic workload repository and the load balancing advisory facilities in 10g.

CLUSTER VERIFICATION UTILITY

The cluster verification utility (CVU) is a tool to help troubleshoot a cluster, especially during installation. There are several different modes available for CVU, one for each stage of installation and one for each RAC environment component. This utility will check many aspects of the cluster environment, including connectivity between nodes, storage access from each node, system configuration like kernel settings, and user permissions.

A stage in the CVU is related to installation stage you're at and represents a group of component checks. The stages offered by CVU are pre-CRS install, post-CRS install, pre-database install, post-database install, pre-database configuration, and post-database configuration. The database installation component checks can be conducted individually for these components: system requirements, shared storage, cluster filesystem space, connectivity between nodes, OS user and permissions, node applications (VIP, ONS, GSD, and listener), and peer-to-peer comparison. Additional checks for the clusterware components include: cluster manager, Oracle Cluster Registry, and Oracle Clusterware. Each of the checks described here provide sufficient information on the test results and can provide additional information by using the "-verbose" flag.

For complete documentation on the CVU and its options, see Appendix A of the Oracle Clusterware and Oracle Real Application Clusters Administration and Deployment Guide 10g Release 2 (10.2), Part Number B14197-02.

REFERENCES

- Oracle Clusterware and Real Application Clusters Administration and Deployment Guide 10g Release 2 (10.2), Part Number B14197-02
- Oracle Database 10g Release 2 Automatic Storage Management New Features Overview, An Oracle Technical Whitepaper, April 2005, Ara Shakian and Nitin Vengurlekar
- Oracle Real Application Clusters 10g Release 2 presentation by Barb Lundhild and Phil Newlan
- Oracle Database 10g Release 2 New Features In Clustering, Barb Lundhild

FROM THE LAWYERS

The information contained herein should be deemed reliable but not guaranteed. The author has made every attempt to provide current and accurate information. If you have any comments or suggestions, please contact the author at [dnorris\(at\)itconvergence.com](mailto:dnorris(at)itconvergence.com).