18c ORACLE Database

Oracle Real Application Clusters (RAC) on Oracle Database 18c

ORACLE WHITE PAPER | FEBRUARY 2018

ORACLE

Executive Overview

Executive Overview

Oracle Real Application Clusters (RAC) is the premium solution utilized by customers to provide High Availability and Scalability to the Oracle Database. There is simply no other solution in the market that combines all the features that Oracle RAC provides without requiring any changes to the application.

Oracle RAC initially focused on providing best of class database services has steadily evolved over the years and now provides a comprehensive High Availability (HA) stack that can be used in Public, Private and Hybrid Clouds to provide high availability, scalability, flexibility and agility for any application. The Oracle RAC "Family of Solutions" is an integrated suite of products to manage the availability and scalability requirements of both applications and databases.

Oracle RAC and its Family of Solutions contribute immensely to the "autonomous" character of the Oracle Autonomous Database. These include new features like Autonomous Health Framework and existing features like Oracle Cache Fusion algorithms which are further enhanced utilizing the latest in hardware innovations and industry trends like cloud computing and Machine Learning. Many of these features are discussed in detail in the technical section.

Enterprises today continue to generate more and more data at an unprecedented pace. Couple that with the need for businesses to rapidly analyze this growing data and identify trends to react quickly to changing market conditions requires increasing amount of computing power. Oracle RAC is perhaps the only solution that provides linear horizontal scalability without application code change and along with the Oracle RAC Family of Solutions provide a unique solution for end to end data management. In the latest release, Oracle RAC provides a host of new features and improvements to existing features, many of which work autonomously below the covers without any need for manual intervention or calibration.

These features can be broadly classified as features that provide

- Better Scalability
- Better Availability
- Better Efficiency and Management functionality to manage pool of clusters

Oracle Real Application Clusters - Overview

Oracle Database with the Oracle Real Application Clusters (RAC) option allows multiple instances running on different servers to access the same physical database stored on shared storage. The database spans multiple hardware systems and yet appears as a single unified database to the application. This enables the utilization of commodity hardware to reduce total cost of ownership and to provide a scalable computing environment that supports various application workloads. If additional computing capacity is needed, customers can add additional nodes instead of replacing their existing servers. The only requirement is that servers in the cluster must run the same operating system and the same version of Oracle. They do not have to be of the same capacity. This saves on capital expenditures as this allows customers to buy servers with latest hardware configurations and use it alongside existing servers. This architecture also provides high availability as RAC instances running on different nodes provides protection from a node failure. It is important to note that (almost) all applications including Oracle Applications, PeopleSoft, Seibel, SAP etc run without any changes on Oracle RAC database.

Customer's requirement for database availability and scalability continue to increase as customers cannot afford any downtime in their environments anymore. These requirements are not isolated to just databases but include other critical components such as servers, network, client connections etc. Furthermore, there is a need for an intelligent resource manager that is able to redirect incoming workloads dynamically to nodes which are idle or in some cases more capable in terms of computing power and memory. The Oracle RAC Family of Solutions provides an integrated product bundle to ensure all these requirements are met. Oracle RAC Family of Solutions is comprised of the following components

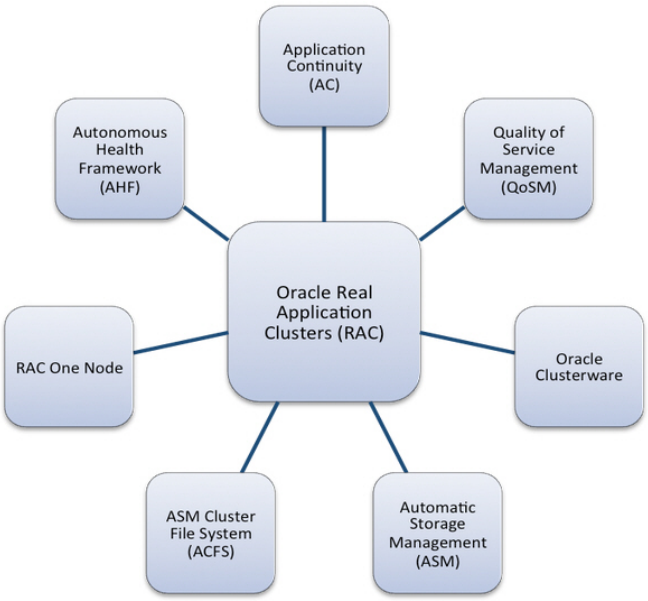

Figure 1: Oracle family of solutions

The functionality provided by the Oracle RAC Family of Solutions can be used by licensed Oracle RAC or Oracle RAC One Node customers without any additional charge.

Oracle Clusterware

Oracle Clusterware is the technology that transforms a server farm into a cluster. Oracle Clusterware is a complete, free-of-charge clustering solution that can be used with Oracle RAC, RAC One Node and even Single instance Oracle databases. Oracle Clusterware provides node membership, node fencing and optimal resource placement.

New in Oracle Clusterware 18c

Oracle Clusterware 18c enhances the new deployment options initially introduced in Oracle 12c Release 2 for easier management and deployments of large pool of clusters. The new architecture called Oracle Cluster Domain would free individual clusters to dedicate all its resources to the database or application as management tasks like deployment, storage management, performance monitoring is delegated to run on a pre-defined Cluster called the Domain Services Cluster.

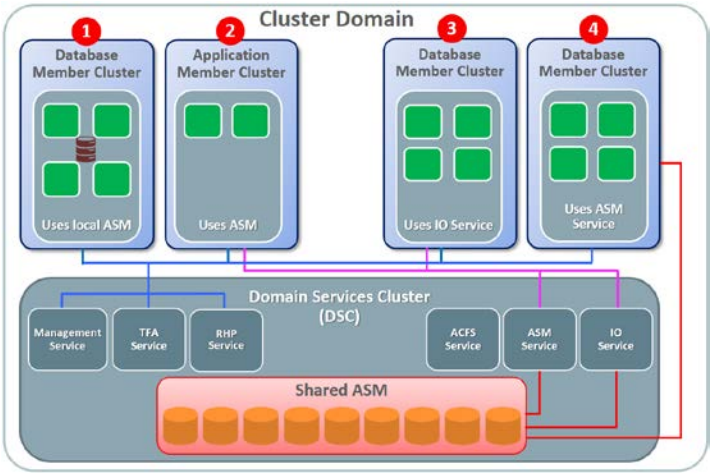

Figure 2: Oracle Cluster Domain

As shown in figure 2 above, a Cluster Domain consists of a single Domain Services Cluster (DSC) and one or many Member Clusters. DSC provides many services which can be utilized by Member Clusters. There are four types of member clusters

1. Database Member Cluster with high performance local storage that is not shared with other member clusters.
2. Application Member Cluster typically hosting applications.
3. Database Member Cluster accessing ASM storage using the ASM I/O services provided by the Domain Services Cluster. This is also referred to as indirect access.
4. Database Member cluster accessing ASM storage on the Domain Services Cluster directly via SAN storage. This is also referred to as direct access.

Regardless of the member cluster type, all of them benefit from the centralized Management Repository services, Trace File Analyzer services and other services provided by the DSC. Oracle 18c allows converting a standalone cluster to a member cluster.

Note that Oracle 18c can also be deployed as a Standalone cluster as in previous releases

Choosing a deployment model.

The choice of deployment no longer depends on the installation type as in previous releases. Oracle 18c allows both new installations to be deployed as a Cluster Domain model while allowing standalone Clusters to be converted to

Member Clusters. It is important to note that Oracle Clusterware licensing remains the same for both deployment models. Some aspects to consider when choosing a deployment model are:

- Cluster Domain architecture delegates the management aspects of Member Clusters to the DSC. This optimizes the Member Cluster management in terms of both provisioning and performance management. Resources like CPU and memory on the Member Cluster can now be dedicated to the database computing needs resulting in cost savings for the customer.
- Cluster Domain architecture provides a unified consolidated storage solution via the DSC. This model makes it easier to provision new databases using the Oracle ASM cloning feature. Storage consolidation using the Cluster Domain deployment model benefits vastly from the new Database oriented storage management features introduced in Oracle ASM.
- Centralized data collection facilities provided by Autonomous Health Framework (AHF) in the DSC allow the Member Clusters behavior to be analyzed using Machine learning capabilities used by AHF which continuously monitors the Member Clusters. This functionality can in many cases prevent a problem before it occurs. For example, AHF can detect anomalies between real time performance counters and expected values to notify system admin of impending performance issues while generating targeted diagnosis and corrective actions.

For more information about Oracle Clusterware, visit oracle.com/goto/clusterware

Oracle Automatic Storage Management (ASM)

Oracle Automatic Storage Management (ASM) is the recommended volume manager which can be used for both, Oracle RAC and Single Instance Oracle Databases. Oracle ASM simplifies storage management through the principle of “stripe and mirror everything” (SAME). Intelligent mirroring capabilities allow administrators to define 2-or 3-way mirrors to protect vital data. When a read operation identifies a corrupt block on a disk, Oracle ASM automatically relocates the valid block from the mirrored copy to an uncorrupted portion of the disk.

New in Oracle ASM 18c

Oracle ASM 18c enhances the database-oriented storage management introduced in previous releases via the new ASM Flex Disk Group. Customers can now convert to the Flex Disk Group and take advantage of the enhanced management capabilities of Flex Disk Group like (a) modifiable redundancy at individual database file level via File Groups (b) snapshot capabilities and (c) quota management at the database level for consolidated environments. The ability to create snapshots on pluggable databases (PDBs) without relying on the snapshot capabilities of the underlying storage enables DBAs to rapidly provision databases. ASM snapshots are executed at the database level without the need for downtime or any additional manual recovery steps. Additionally modifiable redundancy allows database administrators to start with a conservative mirroring strategy and change the redundancy in future depending on business needs.

Once the Disk Groups are converted to the new Flex Disk Group, storage administrators can utilize quota management to set storage space limits at the database level which helps in consolidation environments as it prevents single database from utilizing all the space in the Flex Disk Group. For more information about Oracle Automatic Storage Management, visit oracle.com/goto/asm

Oracle ASM Cluster File System (ACFS)

Oracle ACFS complements ASM file management capabilities by providing a POSIX-compatible file system to store general purpose and database files. Oracle database has for a long time provided column types to store blobs, XMLs and text files etc. However due to application or business requirements, customers needed a file system to store such data. Obviously storing this data outside the Oracle database requires customers to manually plan for data management activities like backup, synchronization across sites etc.

ACFS provides a Cluster file system which customers can use to store this data. They can additionally use ACFS features like Tagging, Replication and Snapshots to ease their data management activities. New in 18c is support for bidirectional snapshots and even better integration with Oracle Data Guard when using ACFS to store data files. Customers can additionally utilize ACFS tagging feature to add custom tags to their data and retrieve tags using a command line or using tagging API calls directly from their application. ACFS snapshots use copy-on-write (COW) on generic systems without relying on specialized storage resulting in these snapshots consuming minimal space. For more information about Oracle ACFS, visit oracle.com/goto/acfs

Better Scalability

Database sizes have grown tremendously over the years and data volumes continue to grow at a very rapid pace while businesses require database to store transactions instantaneously and at the same time need faster responses to analytical queries. Oracle RAC help with these requirements, In fact the scalability features in Oracle RAC make it possible to start with a smaller footprint and scale out as needed resulting in massive savings as customers do not need to provision a large environment upfront.

Figure 3 below depicts the horizontal scale out of Oracle RAC running a SAP sales and distribution (SD) module benchmark.

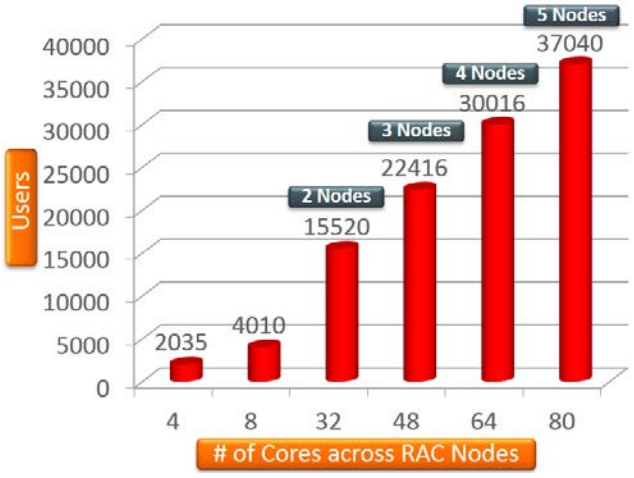

Figure 3: Scalability without Application Code changes

Oracle RAC provides all the availability and scalability features without any manual intervention while providing (Atomicity, Consistency, Isolation and Durability) ACID properties that all applications require. Applications connecting to any instance of the RAC Database do not have to worry about stale reads as that is automatically taken care of by Oracle RAC component Cache Fusion. DBA's can focus on adding business logic into their application while Oracle RAC automatically takes care of all their transactional needs. DBAs could also utilize other Oracle Database features such as Data Guard in conjunction with Oracle RAC for disaster recovery, Oracle In-Memory DB for advanced analytics and the Multitenant feature for consolidation with little or no configuration changes to the application. In fact, the benefits provided by these database features are further augmented when used along with Oracle RAC

Improved Algorithms

Oracle RAC scalability is a result of optimized algorithms that have kept up with the latest advancements in technology. Oracle Cache Fusion which is a component of Oracle RAC is the magic that works behind the scenes to

synchronize the cache of all the instances running on different nodes. This synchronization allows multiple users sessions to run concurrent transactions on either instance of the Oracle RAC Database without incurring stale reads. There are absolutely no manual steps required for this concurrency management. Oracle Cache Fusion will automatically spawn additional background processes as needed depending on the incoming workload and node capacity without any need for manual intervention. The improvements in Cache Fusion algorithms include

- Service Isolation: This feature improves performance by reducing Cache Fusion operations for PDBs / Services not offered in all instances. This is particularly useful when certain PDBs are open on a subset of nodes as Cache Fusion synchronization messages are only sent to those nodes where the PDBs / Services are up and running.
- Choose the right path: Until now, Cache Fusion solely utilized the private network to synchronize the cache, as rotating disk performance has been traditionally slower. However, storage access performance has improved recently as a result of hardware vendors utilizing newer technologies such as SSDs and NVME. These devices have dramatically low I/O latencies, so sometimes it may be beneficial to read blocks from disk rather than ship them over the private network. As shown in Figure 4, Cache Fusion monitors the network performance and storage I/O statistics on an ongoing basis and will utilize the more efficient path as needed. Note that this is done automatically and on an ongoing basis, without the need for DBA intervention.

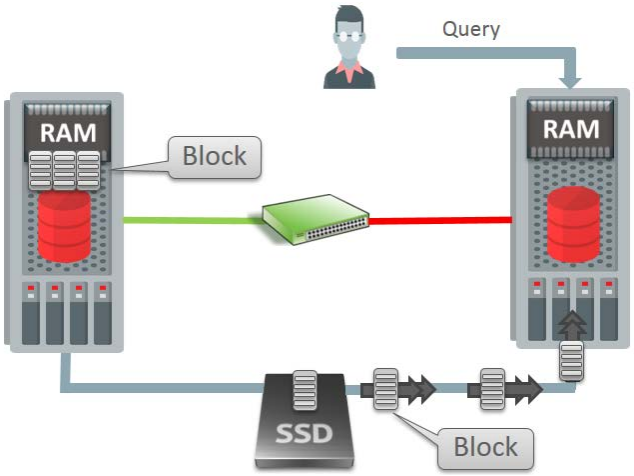

Figure 4: Scalability without application code changes

Scalable Sequences

Applications use Sequences to generate unique numbers. Scalable sequences optimize the sequence generation by using a unique combination of instance number and session number to reduce the impact of Index leaf block contention during massive loads. Scalable sequences feature will result in improved performance of workloads suffering from sequence generation contention. This is perhaps one of the few features that is not automatically enabled as it requires some intervention by DBA's to ensure this does not change their implemented business logic. However DBA's can easily convert existing sequences to the new Scalable sequence using the simple "*Alter sequence sequence_name scale*;" command.

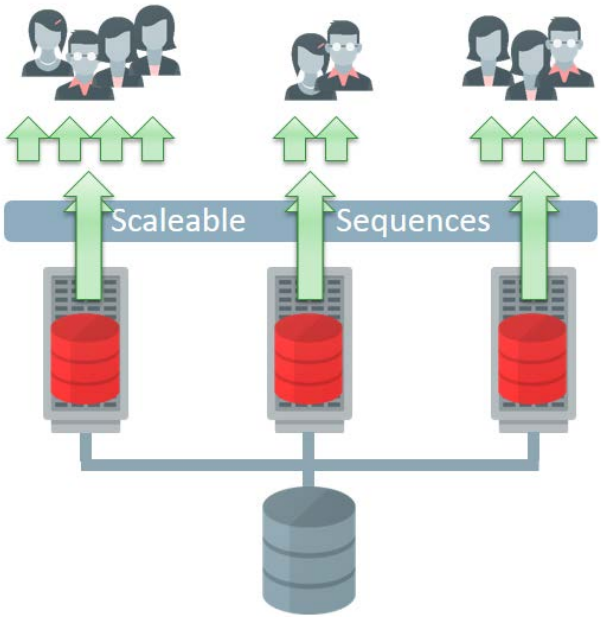

Figure 5: Scalable Sequences in Oracle RAC

Service Oriented Buffer Cache

Service Oriented Buffer Cache initially introduced in Oracle RAC 12c Release 2 essentially reduces and in many cases eliminates “physical reads” after a planned singleton service failover. Prior to this feature, singleton service failover response time is affected as buffers that were cached on the failed instance have to be re-read from the disks incurring the overhead of physical reads. In Oracle RAC 18c, Cache Fusion maintains an in-memory hash that tracks the blocks in the buffer cache and the service name used by the sessions to connect. This information is persisted automatically into the data dictionary so that it can be used after instance restart. Cache Fusion uses the in-memory hash in two ways:

1. Resource Mastering optimization: Resource mastering of a resource cached in the buffer is only considered on the node where the service that the session used to access the resource is running, This results in improved performance as it eliminates the need for sending additional messages on the private network for resource change operations.
2. Pre-Warm the Buffer Cache: During planned maintenance, when a singleton service is failed over, Cache Fusion will pre-warm the buffer of the instance to which the service is going to failover. This reduces the physical reads that the sessions would have otherwise incurred, resulting in consistent performance for those failed over sessions as shown in figure 6.

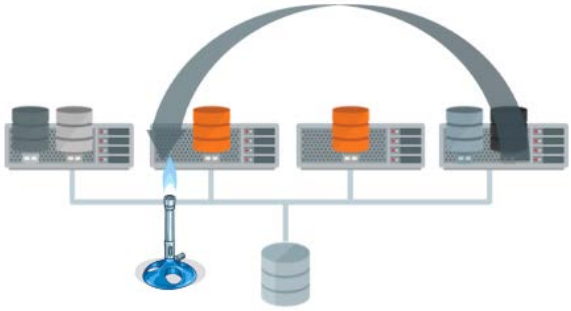

Figure 6: Pre-Warm cache before service failover

Better Availability

The cost of downtime to businesses regardless of application or database downtime has increased over the years. These costs increase dramatically in consolidated environments as multiple applications and databases will be affected by any downtime. Planning redundancy for different layers in the datacenter is critical to achieving availability. Contingencies have to be planned for unplanned events like hardware failure and planned events like application of firmware, operating system patches. It is essential that both planned and un-planned failures are (a) least disruptive to the underlying database and (b) be able to notify user sessions connected to the nodes affected by the event, so that those sessions can gracefully reconnect to surviving database instance in the cluster. Oracle RAC 18c provides better availability by introducing new features and improving existing features introduced in Oracle RAC 12c Release 2.

Smart Reconfiguration

Nodes leaving or joining a cluster result in a reconfiguration which is essentially a synchronization event to recover all the changes made by the failed instance. Oracle RAC has over the years reduced the time that the sessions wait on this event during reconfiguration. In Oracle RAC 18c, smart reconfiguration reduces the impact of service disruption from both planned and unplanned operations utilizing features such as Recovery Buddy and Service Isolation resulting in up to 4 times faster reconfiguration than previous releases.

Recovery Buddy

Figure 7: recovery Buddy

Recovery Buddy feature greatly reduces the time sessions have to wait during reconfiguration. In prior releases, Oracle RAC instances identified and recovered the changes made by the failed instance by reading the redo logs. The surviving instances have to wait until the changes from the failed instance are read and applied. However

reading these changes from storage is time consuming. Recovery Buddy feature optimizes this by allocating a Buddy Instance for every RAC Node. These Buddy instances track the block changes in the local SGA in a hash table. This allows recovery to proceed faster as the changes are read from the hash table instead of reading from storage. In Oracle RAC 18c, the buddy instance gets the necessary recovery locks to be applied in batch mode instead of requesting locks one at a time resulting in faster recovery.

Reader Nodes

Oracle Grid Infrastructure 12c introduced the concept of Oracle Flex cluster in which nodes can be classified as a Hub node or a Leaf node. Oracle Grid Infrastructure12c Release 2 added the feature to run read-only instances on Leaf nodes. Oracle 18c enhances this feature by automatically starting an instance in either read only or read write mode depending on the node type defined during installation. Any configured services are also started on the appropriate instance. Database sessions using the read-only or read-write service to connect is redirected to the appropriate instance without any manual intervention. Nodes can be converted from Hub Node to Leaf Node or vice versa. Once converted, a restart of the clusterware stack brings up the instance and services in the appropriate mode automatically. Reader node feature improves availability as database instances running on Hub nodes are not affected by the Leaf nodes joining or leaving the cluster.

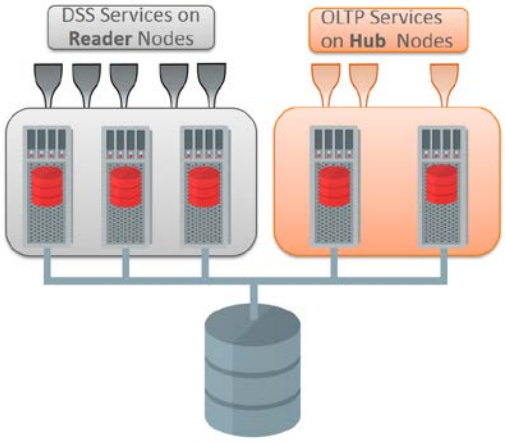

Figure 8 Reader Nodes with Read write Nodes

Since Leaf node restart does not disrupt database activity on Hub nodes, this feature can be used by customers to scale up and improve the performance of DSS queries by adding Leaf nodes running read only instances without impacting database availability. In addition, nodes with large number of CPU and memory can be configured to use Oracle Parallel Query (PQ) on Read Only Instances so they can use the additional memory for typical parallel query activities like sorting or hash-joins. These nodes can be moved between different RAC clusters depending on business needs without any disruption to the database. DBA's can opt to configure "Local Temporary Tablespace" so PQ slaves running massive sorts or hash joins can spill to disk if they run out of physical memory. Read only nodes can also utilize Oracle In-memory feature to transparently accelerate analytical queries. Figure 8 shows a five node RAC instance with two Hub nodes and three Leaf nodes.

Node Weighting

Oracle Node Weighting feature introduced in Oracle RAC 12c Release 2 changed the algorithm used to choose the node(s) to be evicted during an eviction when the split brain results in a cluster to be split into equal half's. In the past, the candidate node to be evicted was chosen based on the node number. In Oracle 18c, this algorithm is further improved as we consider several additional factors including the criticality of workloads that are currently

running, singleton services and additional secondary failures. This feature does not require any manual intervention, however DBA's can optionally configure a node or service to be marked critical based on business needs. This optional step can be used to configure criticality for a specific node that has say more CPU or memory than other nodes, while a service level definition can be used for a critical service that is deemed more important than other services.

Continuous Application Availability

Continuous Application Availability (CAA) in Oracle 18c seamlessly drains and migrates database sessions during planned or un-planned maintenance. The maintenance itself could be a planned operation for replacing failed hardware component or a node may just crash due to the component failure. CAA feature ensures that regardless of nature of node failure, database sessions connected to the failed nodes are migrated to the surviving nodes. This is done transparently without the need of special drivers. For more information about Continuous Application Availability, visit oracle.com/goto/ac

Efficient Management of a Pool of Clusters

Oracle RAC Family of Solutions provide an integrated set of tools to help enterprises with end-to-end life cycle management tasks that can be used to administer a pool of clusters efficiently. Until recently, tasks like provisioning, installation, configuration, performance tuning, log file management, patching were done individually one node at a time. This is inefficient and error prone. Features like Oracle Rapid Home provisioning (RHP) and Oracle Autonomous Health Framework (AHF) help manage these tasks efficiently over a pool of deployments.

Rapid Home Provisioning (RHP)

Oracle RHP provides an efficient and non-disruptive method to provision, patch and upgrade various layers of the Oracle software infrastructure; these layers include but are not limited to Oracle Grid Infrastructure, Oracle Database (RAC, RAC One Node and Single Instances), Applications and Middleware.

RHP can also provision, upgrade and patch Standalone clusters, Domain Services clusters and Member clusters as shown in figure 9. RHP itself can be provisioned in a standalone deployment model or as part of the new Cluster Domain deployment model. RHP can help standardize customer software install by the use of gold images. Essentially customers can create an environment or site specific gold image which can then be used as the standard image to be deployed. For more information about RHP, visit oracle.com/goto/rhp

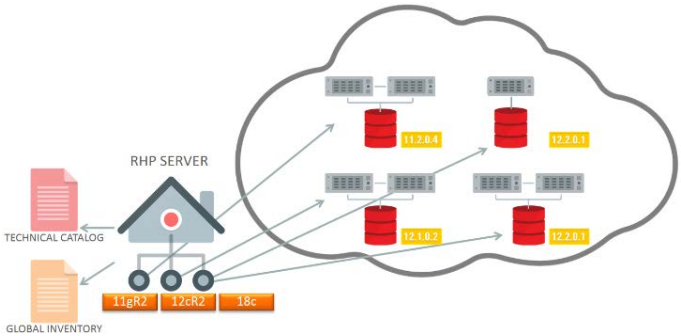

Figure 9 Rapid Home Provisioning

Oracle Autonomous Health Framework (AHF)

Oracle (AHF) presents the next generation of tools as components, which work together autonomously 24x7 to keep database systems healthy and running while minimizing human reaction time. Utilizing machine-learning technologies, Oracle AHF provides early warning or automatically solves operational runtime issues faced by Database and System administrators in the areas of availability and performance. In Oracle Database 18c, AHF extends this functionality and performance across nodes, databases and clusters with the following new features:

- Oracle Cluster Health Advisor Cross-Database and Cross-Cluster analysis to support targeting external issues to a specific database instance or node resulting in higher confidence diagnosis of impending problems and improved preventive actions.
- Introduction of Oracle Trace File Analyzer Service to assist DBA's and system administrators in proactively monitoring or rapidly diagnosing trace files across multiple member clusters from a centralized repository.
- ORAchk and EXAchk have been rewritten with a focus on performance and extensibility resulting in a 3x speed improvement and smaller resource footprint.
- Oracle Database QoS Management now supports automatic policy set provisioning when adding databases to existing clusters improving provisioning and management in fleet or cloud deployments.

For more information about AHF, visit oracle.com/goto/ahf

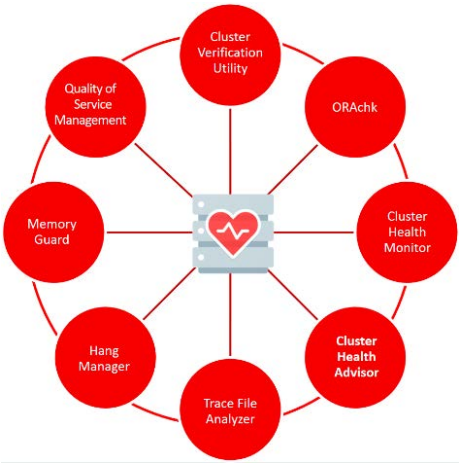

Figure 10 Components of Autonomous Health Framework

Conclusion

Scalability and availability requirements for businesses regardless of OLTP, DSS, Data warehouse or even converged Hybrid Transactional Analytical Processing (HTAP) systems are at all-time highs. Business cannot afford any downtime. They also need these systems to be scalable so they can analyze the vast volumes of data quickly to react to changing customer demands. Oracle Real Application Clusters along with the Oracle RAC Family of Solutions makes it easier for customers to take advantage of the Business continuity, availability, scalability, flexibility and agility requirements so they can rapidly adapt to changing business needs. Oracle Real Application Cluster 18c continues on this path by providing significant enhancements in all of the areas that matter most for businesses to succeed. Improved algorithms in all the components of the stack make Oracle RAC the database virtualization solution of choice for your IT infrastructure.

Oracle Corporation, World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065, USA

Worldwide Inquiries
Phone: +1.650.506.7000
Fax: +1.650.506.7200

CONNECT WITH US

blogs.oracle.com/oracle
facebook.com/oracle
twitter.com/oracle
oracle.com

Integrated Cloud Applications & Platform Services

Oracle Real Application Clusters - Technical Overview
February 2018
Author: Anil Nair
Contributing Authors: Markus Michalewicz