ORACLE Database 19c

Oracle Real Application Clusters (RAC) on Oracle Database 19c

ORACLE WHITE PAPER / FEBRUARY 2019

ORACLE

TABLE OF CONTENTS

EXECUTIVE OVERVIEW

Oracle Real Application Clusters (RAC) is an option to the Oracle Database that provides high availability (HA) and scalability to the Oracle Database without requiring any application changes. There is no other solution in the market that provides all the features as Oracle RAC.

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

Oracle RAC customers also benefit from reduced CAPEX incurred during costly lift and replace hardware cycles. Using Oracle RAC, new hardware can be provisioned along with old hardware, to create a pool of clusters running multiple instances of Oracle RAC. All these operations can be performed without incurring any downtime.

Oracle RAC has evolved over the years from a database only option to a comprehensive stack that manages resources outside the database such as storage, network and CPU. Oracle RAC ships with Oracle RAC “Family of Solutions” which is an integrated suite of products that help manage the entire stack from Oracle Database to the network, storage and operating system. This tight integration ensures that customers benefit from Oracle RAC scalability and seamless failover regardless of deployment in public, private or hybrid cloud environments.

Oracle RAC contributes immensely to the “autonomous” character of the Oracle Autonomous Database. New features such as Autonomous Health Framework and existing features such as Oracle Cache Fusion algorithms are continuously enhanced utilizing the latest in hardware innovations and industry trends such as cloud computing and machine learning. These features are discussed in detail in the technical section.

OVERVIEW OF ORACLE REAL APPLICATION CLUSTERS (RAC)

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 the same model or capacity. This saves on capital expenditures allowing 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 server failure. It is important to note that (almost) all applications such as Oracle Applications, PeopleSoft, Siebel, SAP run without any changes on Oracle RAC.

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, and client connections. 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.

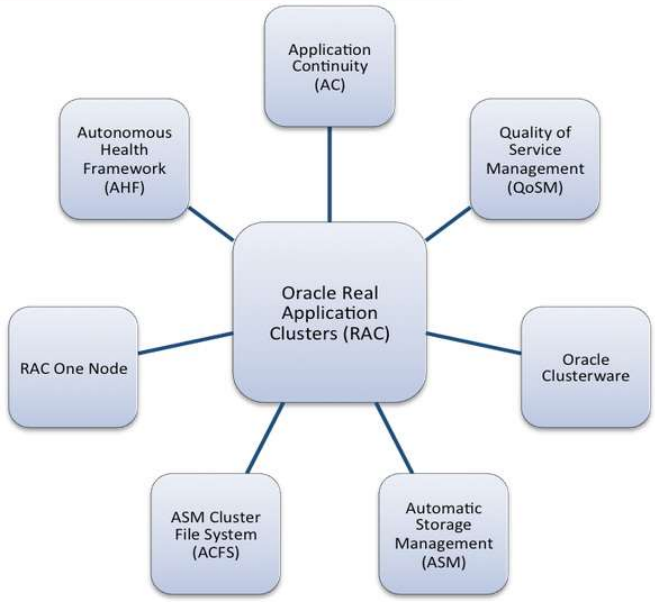

Figure 1. The Oracle RAC family of solutions

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 19C

Oracle Clusterware 19c enhances the new deployment options for easier management and deployments of large pool of clusters. The new architecture called Oracle Cluster Domain would enable individual clusters to dedicate

their resources to the database or application as management tasks such as deployment, storage management, performance monitoring is delegated to run on a pre-defined Cluster called the Domain Services Cluster.

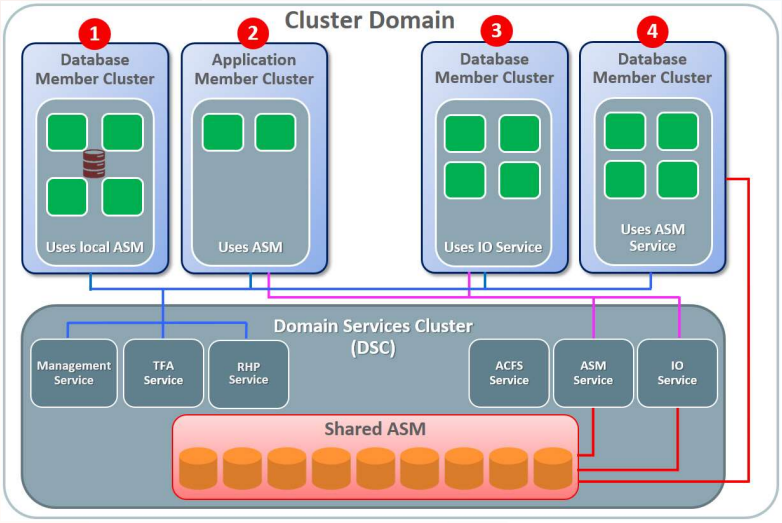

Figure 2. Oracle Cluster Domain

As shown in figure 2 above, a Cluster Domain consists of a single Domain Services Cluster (DSC) and one or more 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 (DSC). 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.

CHOOSING A DEPLOYMENT MODEL.

The choice of deployment no longer depends on the installation type as in previous releases. Oracle Clusterware now allows 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 such as 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

Oracle Automatic Storage Management (ASM) is the recommended volume manager that 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 19

Oracle ASM 19 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 such as (a) modifiable redundancy at individual database file level via File Groups, (b) Cloning/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 such as backup and synchronization across sites.

ACFS provides a Cluster file system which customers can use to store this data. They can additionally use ACFS features such as Tagging, Replication and Snapshots to ease their data management activities. New in 19c 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 in anticipation of future requirements.

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

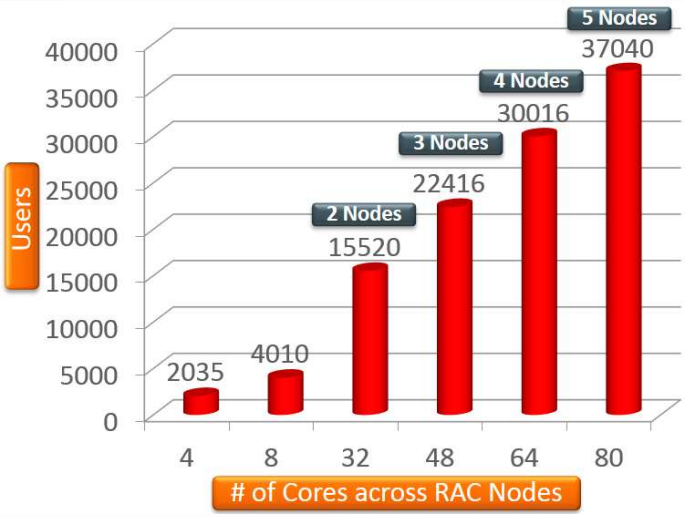

Figure 3. Scalability without application code changes

Oracle RAC provides all the availability and scalability features without any manual intervention while providing atomicity, consistency, isolation and durability (ACID) properties that 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 called Cache Fusion. DBA's can focus on adding business logic to their application while Oracle RAC automatically takes care of all their transactional needs. DBAs can 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 servers. This synchronization allows multiple users sessions to execute concurrent transactions on either instance of the Oracle RAC Database without incurring stale reads. There are 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 algorithms include

- **Database Reliability Framework (DRF):** Introduced in Oracle 19c RAC, DRF attempts to detect problems early before that problem can cause disruption in service. The idea is to detect problems and identify root cause accurately. Once a problem is identified, an action is implemented automatically. Action include resizing internal memory structures or changing the priority of Oracle RAC processes depending on the identified problem. The

algorithm considers multiple metrics to reach a conclusion. Consider a system which has high “redo waits” with no I/O contention based on the metrics collected over time. If there is enough CPU resource then a possible action plan for reducing the “redo waits” is to move the LGWR process to higher priority to ensure it gets enough CPU. DRF would take this action after carefully considering all the metrics. This results in problem resolution with minimal service disruption before the problem multiplies over time and affects database availability. There are around 50 critical events monitored today and this will continue to be enhanced in future releases.

- **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 with hardware vendors utilizing newer technologies such as SSDs and NVME. These devices have dramatically lower I/O latencies, so sometimes it may be beneficial to read blocks from disk rather than transmit 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.

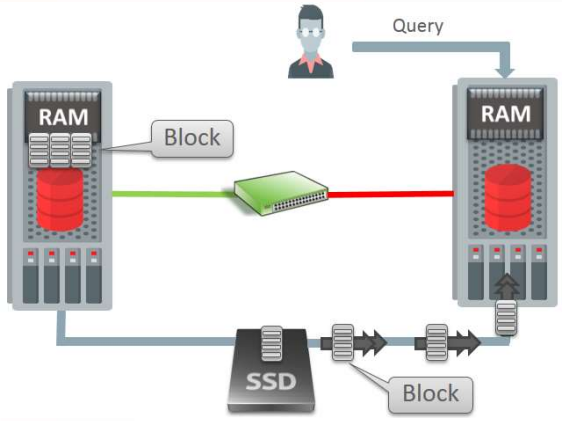

Figure 4. Scalability without application code changes

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 servers as Cache Fusion synchronization messages are only sent to those nodes where the PDBs / services are active.

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 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 command:

```
Alter sequence sequence_name scale;
```

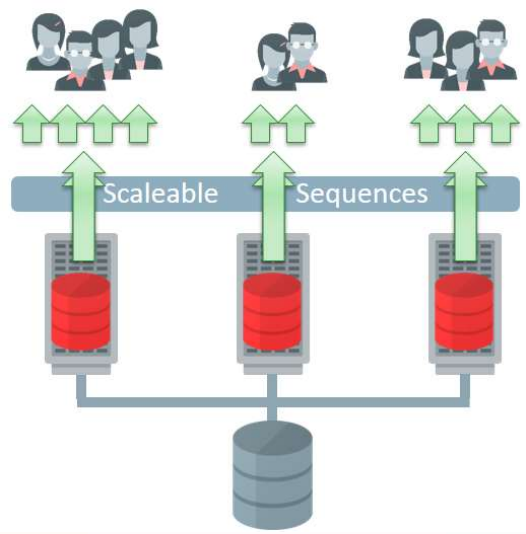

Figure 5. Scalable sequences in Oracle RAC

Service Oriented Buffer Cache

The Service Oriented Buffer Cache feature 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 19c, 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:

- **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.
- **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.

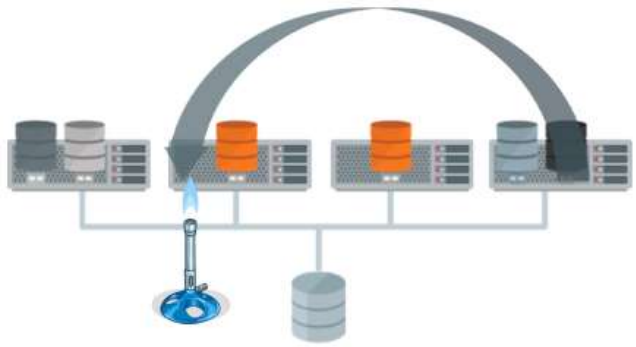

Figure 6. Pre-warm the cache before service failover

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 are 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 such as hardware failure and planned events such as application of firmware, operating system patches. It is essential that both planned and unplanned failures are (a) least disruptive to the underlying database and (b) be able to notify user sessions connected to the servers affected by the event, so that those sessions can gracefully reconnect to surviving database instance in the cluster. Oracle RAC 19c provides better availability by introducing new features and improving existing features introduced in previous releases.

Smart Reconfiguration

Servers 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 19c, 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

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 19c, 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.

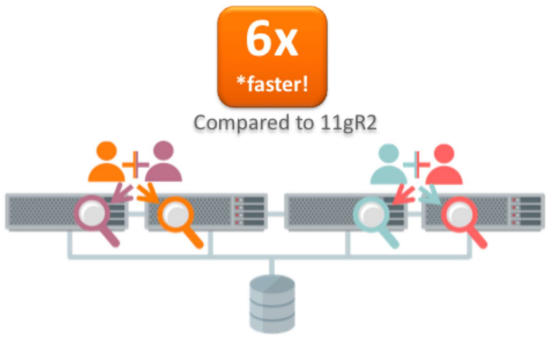

Figure 7: Recovery Buddy

Node Weighting

Oracle Node Weighting feature improves the algorithm used to choose the node(s) to be evicted during a node 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 19c, 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 but 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 server that has say more CPU or memory than other servers in the cluster. If all

the servers have same capacity in terms of CPU and memory, then a service can be defined as critical resulting in smarter node eviction.

Continuous Application Availability

Continuous Application Availability (CAA) in Oracle 19c 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 in anticipation of a node 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 such as provisioning, installation, configuration, performance tuning, log file management, patching were done individually one node at a time. This is inefficient and error prone. Features such as Oracle Fleet Provisioning and Patching (FPP) and Oracle Autonomous Health Framework (AHF) help manage these tasks efficiently over a pool of deployments.

Oracle Fleet Provisioning and Patching

Oracle Fleet Provisioning and Patching (FPP) previously known as Oracle Rapid Home Provisioning (RHP) provides an efficient and non-disruptive method to provision, patch and upgrade various layers of the Oracle software infrastructure on a fleet of servers. These layers include but are not limited to Oracle Grid Infrastructure, Oracle Database (RAC, RAC One Node and Single Instances), applications, and middleware.

FPP can also provision, upgrade and patch standalone clusters, domain services clusters and member clusters as shown in figure 8. FPP itself can be provisioned in a standalone deployment model or as part of the new cluster domain deployment model. FPP can help standardize customer software installations with 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 for deployed. For more information about FPP, visit oracle.com/goto/rhp

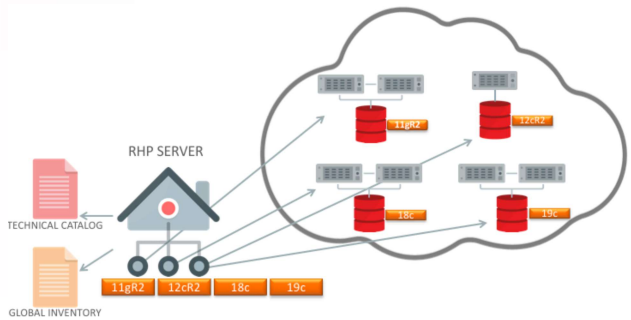

Figure 8. Rapid Home Provisioning

Oracle Autonomous Health Framework

Oracle Autonomous Health Framework (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 19c, 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

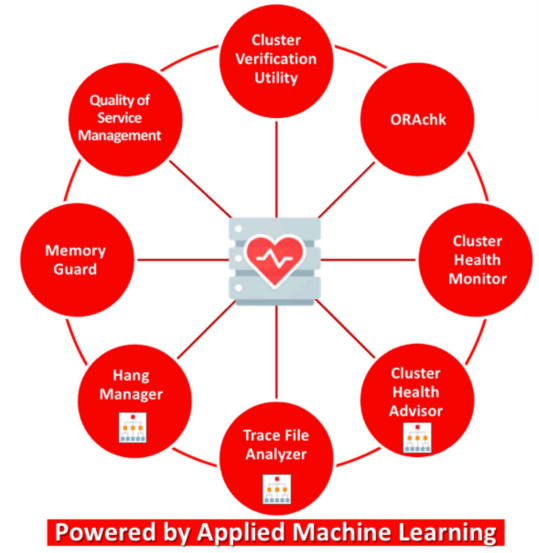

Figure 9. Components of the Oracle Autonomous Health Framework

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 Clusters 19c 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

Worldwide Headquarters
500 Oracle Parkway, Redwood Shores, CA 94065 USA

Worldwide Inquiries
TELE + 1.650.506.7000 + 1.800.ORACLE1
FAX + 1.650.506.7200
oracle.com

CONNECT WITH US

Call +1.800.ORACLE1 or visit oracle.com. Outside North America, find your local office at oracle.com/contact.

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

Integrated Cloud Applications & Platform Services

Oracle Real Application Clusters (RAC) on Oracle Database 19c
Updated February 2019

Author: Anil Nair