

Exadata Database Machine

Horizontal Scaling with Oracle Database

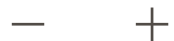


Christian Craft

July 31, 2021



Text Size 100%:



When computing systems and associated databases grow, they need to “scale” to a larger size. There are two dimensions to this growth:

Handling Progressively Larger **DATABASES**

Handling Progressively Larger **WORKLOADS**

Scaling for larger amounts of data requires more storage, but also requires the associated storage network capabilities to move data between servers and storage. Scaling for larger workloads requires more compute power. Sometimes the workload is automated processing, but we generally think of scaling the workload as increasing numbers of online users. It’s not uncommon for a computing system to experience larger databases and workloads at the same time, and thus require scaling in both dimensions.

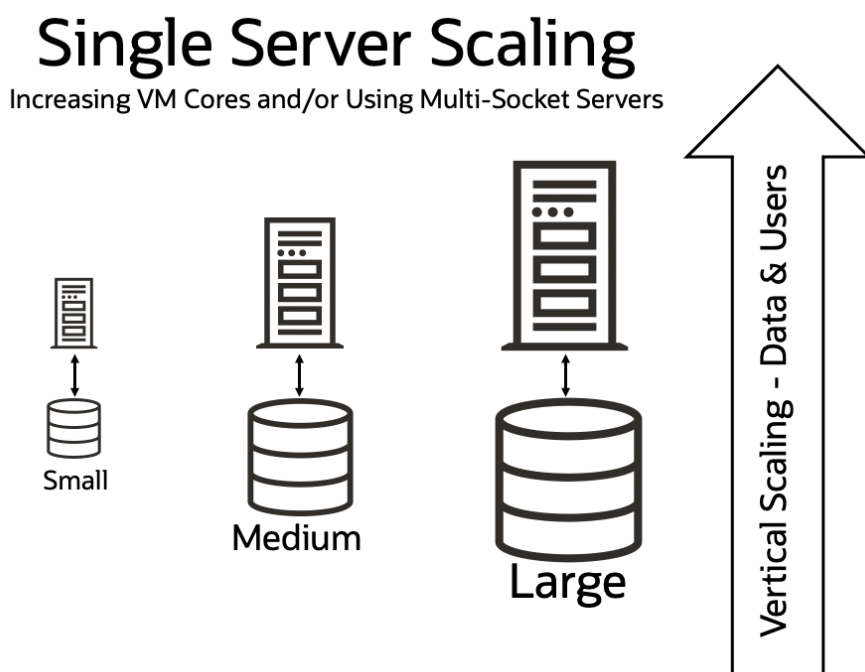
Scaling solutions come in two general forms: **Vertical** and **Horizontal**. Vertical scaling (also called “scale up”) occurs when the existing applications and databases are moved to a larger system. Horizontal scaling (also called “scale out”) occurs when additional systems are added to an existing configuration, and the data and workload is spread across them. Vertical scaling can be very disruptive, costly, and ultimately limited in size, whereas horizontal scaling can avoid these issues, provided the underlying software infrastructure is designed for horizontal scaling.

This article focuses on horizontal scaling with Oracle Databases, and the unique ways in which Oracle Database software supports horizontal scaling. But first, a bit more detail on vertical scaling, which is

available to any application and database.

Vertical Scalability

Scaling vertically means running a database on a single server, then moving to progressively larger and larger servers to handle larger amounts of data and workload. Oracle Database can run on anything from the smallest Virtual Machine (or Cloud “Instance”) to the largest Virtual Machine possible as well as massive bare-metal servers, with 4-sockets, 8-sockets, and even 16-socket servers. Scaling vertically involves MOVING the database to a larger machine, whether that’s a larger Virtual Machine or a larger Physical Machine.



Vertical scaling is sometimes referred to as scaling by forklift because it would literally require a forklift to move a larger server into a data center.

Horizontal Scalability

Horizontal scalability can be a more convenient (less disruptive) method of scaling compared to changing the size of a Virtual Machine (sometimes called an instance type in Cloud environments), increasing number of vCPU assigned to a Virtual Machine, or using a forklift to bring a larger physical server into a data center. Horizontal scalability means combining multiple smaller machines to construct a larger configuration. There are generally 3 methods for accomplishing horizontal scalability as follows:

- Compute Clusters

- Data Replication

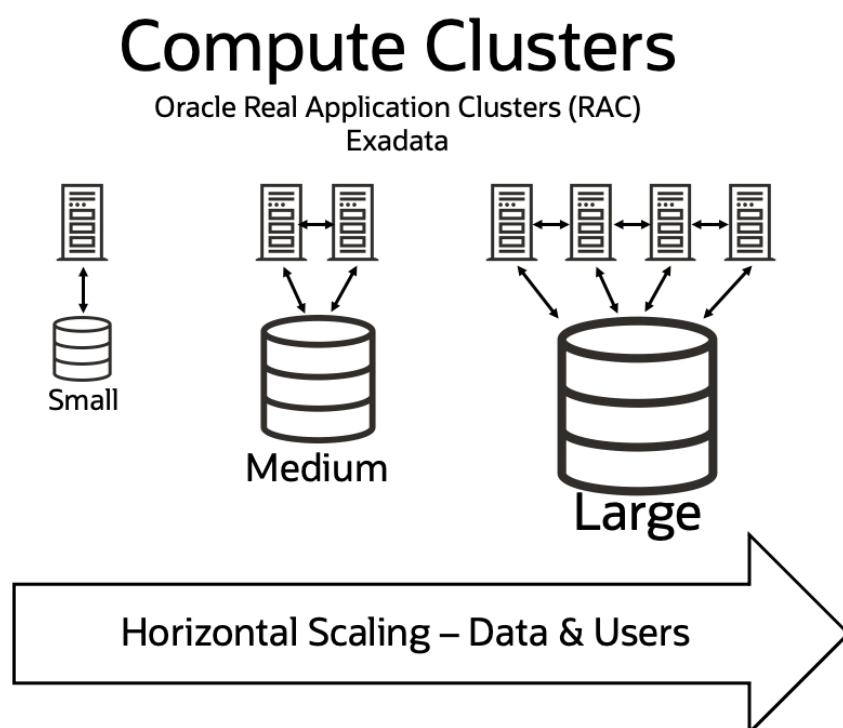
- Database Sharding

Each of these 3 architectures offer advantages, and there isn’t necessarily one “correct” approach for all cases. We will also see that these technologies can be combined (at least with Oracle Database), so

it's not necessarily a choice of one over the others.

Horizontal Scalability - Compute Clusters

The original concept of scaling horizontally involved using clusters of smaller computers combined with shared storage accessed simultaneously by processes running on all compute nodes of the cluster. When you need to grow, just add another compute node to the cluster and expand the shared storage pool. Oracle Real Application Clusters (RAC) was developed to support this scaling architecture and has reached the ultimate implementation in the Oracle Exadata platform (Exadata). Compute clusters require a high bandwidth, low latency interconnect network, and there are key innovations in Exadata that deliver the highest bandwidth (100 Gbps) and lowest latency (~10 µsec) in the industry. For more about how we achieve this, see this blog on [Persistent Memory and RDMA inside of Exadata](#).



Oracle Exadata Cloud Service ([Exadata in the Oracle Cloud](#)) in the X8M generation scales horizontally to 32 Database Servers with 3,200 vCPU and 64 Storage Servers with 3,162 TB of storage. To be clear, we're talking about a SINGLE database with a massive amount of data and massive workload or, more likely, a large number of databases consolidated together on the same Exadata Cloud Service system. Exadata Cloud Service is designed for high-end enterprise workloads, while Oracle Database Cloud Service (Oracle Database on commodity servers in the Oracle Cloud) supports 2-way RAC clusters for smaller deployments (up to 192 vCPU and 40TB of storage).

Availability & Scalability with RAC

[Oracle Real Application Clusters](#) (RAC) provides horizontal scalability while also increasing availability. Scaling horizontally by adding compute nodes to a RAC cluster has the added advantage of allowing a RAC database to continue operating during an outage of 1 or more compute nodes within the cluster. The surviving compute nodes are unaffected by the outage, and the users connected to the failed nodes reconnect to a surviving node after a short outage.

For continuous availability, [Oracle Application Continuity](#) in combination with RAC enables applications to continue operating without loss of connections. Application Continuity allows transactions to continue without rollback and without special retry logic in the application. RAC with Application Continuity has several advantages over another common approach, which is use of Live Migration of Virtual Machines. Application Continuity handles failures as well as proactive maintenance. Nodes of a RAC cluster can be shut down in a rolling fashion for proactive maintenance of hardware, O/S software, drivers, and even Oracle Database and Oracle Grid Infrastructure software. Virtual Machine Live Migration is only useful for proactively moving a running Virtual Machine from one physical node to another prior to proactive maintenance on the physical server or VM hypervisor. Live Migration does not help during failures and does not facilitate maintenance that require O/S or database instance restart.

Oracle Real Application Clusters (RAC) is the clear leader in scaling workloads across active/active compute clusters without application changes, including the most complex workloads such as ERP (Enterprise Resource Planning), OLTP (Online Transaction Processing), and Data Warehouse (Analytic) workloads. One of the most dramatic examples of scaling workloads on RAC is the experience of [PayPal scaling workloads with Oracle RAC](#).

Application Compatibility with RAC

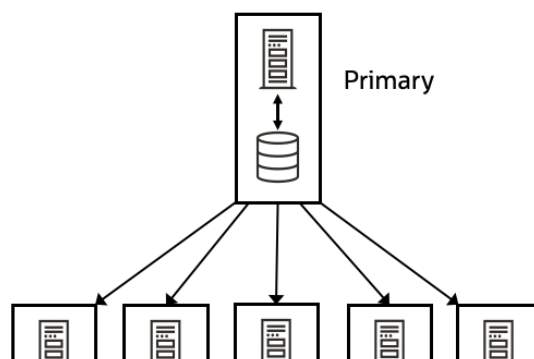
The vast majority of applications are compatible with RAC and can scale on RAC configurations. Exadata has been proven as the ultimate platform for scaling RAC databases due to innovations such as Exadata Smart Storage and high bandwidth (100 Gbps) low latency cluster interconnect using RDMA over Converged Ethernet (RoCE). Over 2,000 ISV applications have gone through certification of their applications on Exadata. RAC on Exadata is widely considered to be in the same category of transparency as “vertical” scalability, even though it achieves scalability by scaling horizontally.

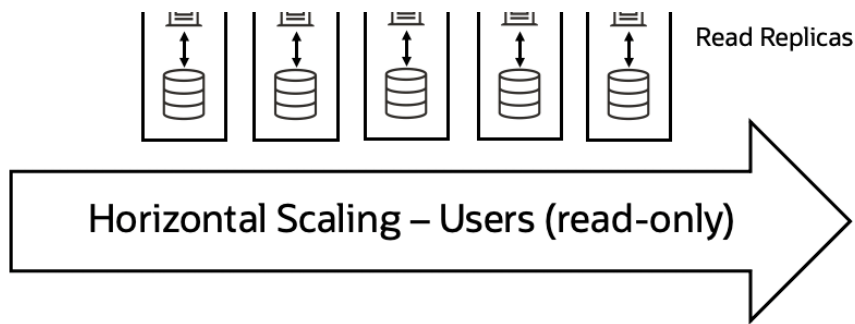
Horizontal Scalability – Data Replication (Read Replicas)

Data replication software maintains copies of a database on one or more systems across a network. The most common forms of replication are for read-only replicas, in support of queries or reports. Updates are applied to the central copy of the database and replicated to the read-only copies (read replicas). The use of read replicas helps with scalability of workload (i.e., users) but does not help with scalability of data. That is, more vCPU can be provisioned to allow reading of the data, but there can only be 1 read/write copy of the database. All copies of the database are the same size, so read replicas certainly are not the solution to scale in terms of data volume.

Read Replicas

Oracle Active Data Guard





[Oracle Active Data Guard](#) is data replication software for Oracle Database that allows up to 30 read replicas from a single primary database, as well as cascading replicas where one replica can feed other replicas. Of course, read replicas are read-only, so it's not necessarily "transparent" from an application perspective. Unlike some other databases, Oracle Database with Active Data Guard does allow DML redirection (changes are transparently redirected to the updatable copy), but it's not designed for large volumes of changes. DML redirection is really intended for read-mostly applications. Enabling low volume changes in a read-mostly application is sometimes sufficiently transparent to the application.

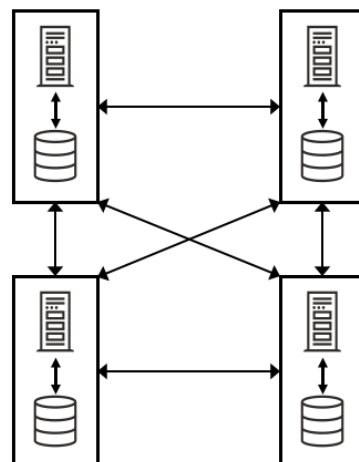
Oracle Active Data Guard provides both synchronous and asynchronous propagation of changes from the primary copy to the replicas, including multiple options to strike a balance between performance and availability. Performance of Active Data Guard is largely dependent on the distance between sites and the networking infrastructure since changes are propagated via redo logs rather than other solutions that propagate all changes at the storage layer (with far greater volume of change propagation).

Horizontal Scalability – Multi-Master Replication

Multi-Master replication goes beyond read replicas, with each replica database open in read/write mode. Multi-Master Replication relies on application-defined detection and resolution logic for any data synchronization conflicts that arise, and this logic can get quite complex if the data structures and workload are complex. Conflict detection and resolution logic should be treated as an integral part of the application and should be designed as part of the application (not an entirely separate component). As the name implies, multi-master replication involves 2 or more databases of the same size, so multi-master replication doesn't provide the ability to scale the volume of data in a database.

Multi-Master Replication

Oracle Golden Gate

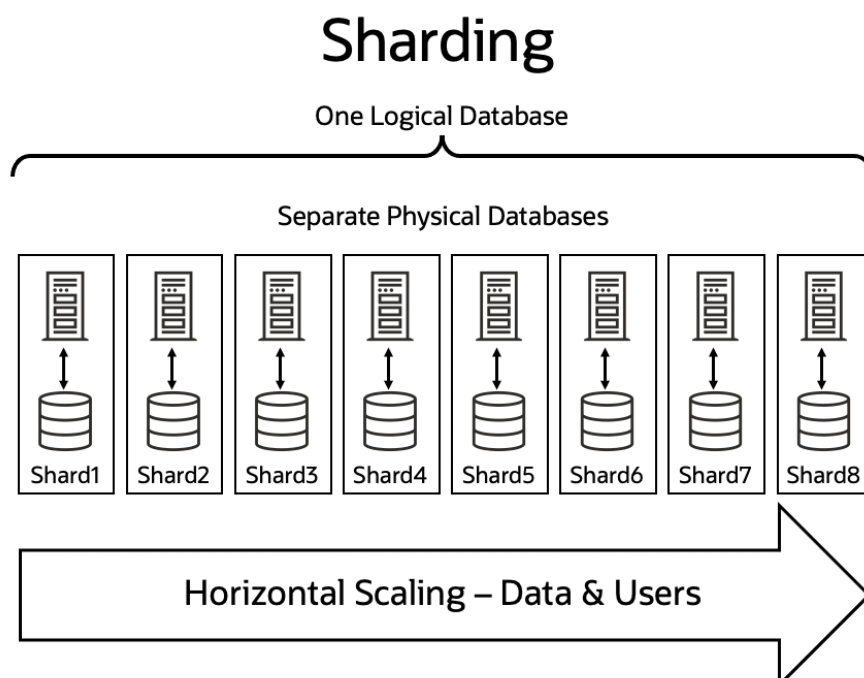


Horizontal Scaling – Users (read/write)

Multi-Master Replication allows multiple databases (often at multiple sites) to operate independently. Replication conflict detection and resolution logic handles updates to the database that occur separately at the independent sites. [Oracle Golden Gate](#) is the leading technology for multi-master replication in the industry, including cross-platform, cross-database, and cross-version heterogeneous data replication. Oracle Golden Gate is also used in other deployment models such as populating a Data Warehouse or Reporting Databases from one or more transactional databases, so Golden Gate is not exclusively a multi-master replication solution.

Horizontal Scalability – Database Sharding

Sharding involves breaking down a single logical database and spreading the data across multiple physical databases, or you can conceptually think of sharding in the opposite direction, combining multiple separate physical databases into one large logical database. While sharding was originally done within application code (I worked on several of these myself), database vendors like Oracle have built sharding features directly into the database to simplify building applications on sharded databases. Sharding provides scalability for both DATA as well as WORKLOAD but does require careful consideration during application development and implementation phases. While changes to application code and data structures are often minimal, sharding typically cannot be used with 3rd party packaged applications unless those applications specifically support sharding.



Oracle introduced [Sharding](#) in Database 12c, and the features of Oracle Sharding have become very sophisticated over the subsequent versions. Oracle Sharding provides a set of services that greatly simplify implementation compared to the old methods of sharding driven by application code. Oracle Sharding also provides capabilities that would be impractical to implement in application code such as intra-shard and inter-shard parallelism.

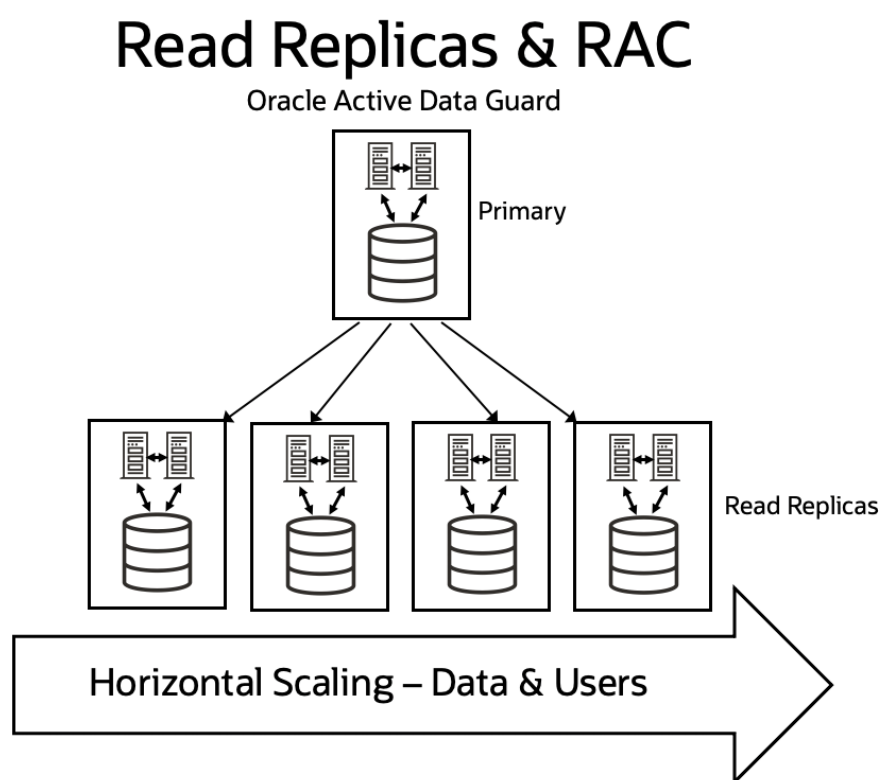
Unlike other solutions on the market, Oracle Sharding is designed to provide the scalability advantages of a multi-master replication solution, but with the added benefit of being able to scale the number of shards.

or sharding, combined with the full capabilities of the Oracle converged database. Sharding implementations are available on NoSQL databases, but those databases do not have the rich data modeling and workload support that Oracle has.

Combining RAC with Sharding & Replication

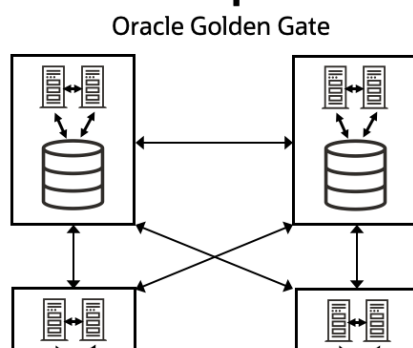
Oracle Real Application Clusters can be combined with Oracle Sharding to reduce the number of shards (by making each shard larger) and to increase availability of each shard. One of the most impressive examples of this is a customer who replaced 90,000 shards on a Cassandra database with just 400 shards on Oracle.

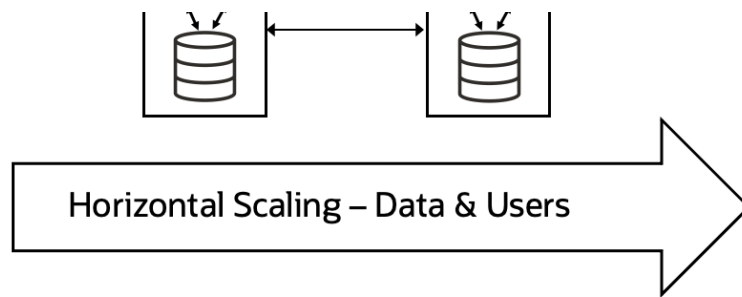
While up to 30 read replicas can be associated with 1 primary database using Oracle Active Data Guard for horizontal scalability, Real Application Clusters (RAC) can be used to increase the scale of the primary and replicas, as well as to increase the availability of each.



Multi-Master Replication enables the deployment of multiple distributed copies of the same data, with each site operating independently. Synchronization of the sites uses an “eventual consistency” model based on replication conflict detection and resolution logic. Each site should be configured for high availability using Oracle Real Application Clusters.

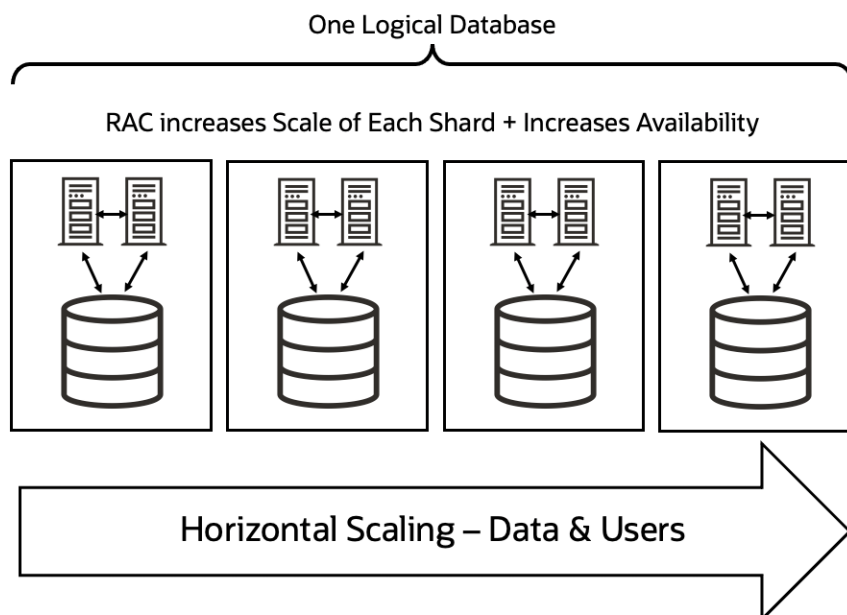
Multi-Master Replication & RAC





Oracle Sharding can be combined with Real Application Clusters to increase the scale of each shard, as well as to increase the availability of each shard. This is especially important as the size and criticality of the shards increases. The ability to have larger shards means managing fewer of them, such as managing tens of shards instead of hundreds. While downtime of a single shard doesn't impact other shards, the data within that shard is unavailable and users who depend on that shard are impacted.

Oracle Sharding & RAC



RAC for High Availability + Active Data Guard for Disaster Recovery

A common configuration is Oracle Real Application Clusters (RAC) for High Availability with Active Data Guard for Disaster Recovery (DR). Active Data Guard allows the DR copy to be open for reporting purposes, enabling businesses to see day-to-day value from their investment in a DR solution rather than it being strictly reserved for disasters. The Oracle Cloud automates this type of deployment, while more advanced implementations can be deployed via direct configuration of Data Guard in Oracle Cloud. Disaster Recovery copies are placed in another Availability Domain (AD) in Oracle Cloud regions that include multiple AD's. Advanced implementations can also place DR copies in another Oracle Cloud Region if needed.

Conclusion

Oracle Real Application Clusters is Oracle's long-standing and original horizontal scaling solution and also provides advantages in terms of database availability. Oracle can also scale horizontally using multiple read-replicas with Active Data Guard, including support for up to 30 read replicas for a single primary database as well as cascading read-replicas. Oracle Sharding provides the ability to scale horizontally by dividing a single logical database across multiple, independent physical databases. Oracle GoldenGate provides real-time data integration across heterogeneous databases, including support for multi-master replication with automatic conflict detection and resolution to provide eventual global consistency. In short, Oracle is the clear industry leader in terms of scalability, including both vertical and horizontal scaling.



Christian Craft

Senior Director, Product Management

Chris Craft is Senior Director of Product Management, focused on advanced Oracle database technologies including Oracle Converged Database, Exadata, Exadata Cloud Services, Autonomous Database, and Oracle's Zero Data Loss Recovery Appliance. Mr. Craft has served in multiple roles within Oracle including On-Site Customer Support, Service Delivery Management, Sales Consulting, and Product Management. His experience ranges from application development, data modeling, and performance benchmarking to managing customer relations, and leading teams of technical professionals. Mr. Craft has worked with over 500 Oracle customers across virtually all industries and including many of the largest customers in the world.

[< Previous Post](#)

[Next Post >](#)