Oracle Maximum
Availability Architecture

Oracle MAA Blueprints for Oracle Cloud Infrastructure (OCI) Deployments

Oracle Database High Availability in the Cloud

ORACLE WHITE PAPER | DECEMBER 2018

Disclaimer

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.

Table of Contents

Introduction 3

Oracle Cloud Infrastructure for MAA 4

- Availability Domain Isolation 4

High Availability Reference Architectures – Oracle Cloud Infrastructure - Database 5

Bronze: Development, Test, and Production Databases 8

- Bronze Summary 9

Silver: Departmental Databases 10

- Silver RAC Summary 11
- Silver Active Data Guard Fast-Start Failover Summary 12

Gold: Business and Mission Critical Databases 12

- Oracle Data Guard and Oracle Active Data Guard 13
- Gold Summary 15

Platinum: Extremely Critical Databases 15

- Oracle GoldenGate 16
- Edition-Based Redefinition 17
- Platinum Summary 17

Oracle Cloud Infrastructure – MAA Proof of Concepts & Configuration Practices 18

- Corruption Prevention, Detection, and Auto-Repair 18
- OCI Cloud Backup and Recovery with Object Storage 19
- Data Guard Fast-Start Failover 21
  - Configuring Fast-Start Failover 21

Fast-Start Failover Observer Placement Best Practices 24

Fast Start Failover with Far Sync 26

Fast-Start Failover Performance 26

(Active) Data Guard SYNC or ASYNC Evaluation 28

Data Guard Switchover to Reduce Downtime 29

Monitoring a Data Guard Configuration 30

Detecting a Transport or Apply Lag using Data Guard Broker 32

Data Guard Operations in OCI 33

Oracle GoldenGate 33

Conclusion 33

Introduction

Enterprises are under intense pressure to do more with less, reduce risk, and increase agility, flexibility, and security. The aggressive movement of information technology (IT) infrastructure and deployment of Database as a Service (DBaaS) on public and private clouds is a strategy that many enterprises are pursuing to accomplish these objectives.

Database cloud transformation drives cost savings by dramatically improving system utilization and reducing management overhead. DBaaS drives cost savings and increased agility through the standardization of IT infrastructure and processes. The Cloud enhances these benefits by enabling a more efficient utility model for computing.

All of the above initiatives, however, also incur business risk by amplifying the impact of downtime and data loss. The failure of a standalone environment used by a single developer or small work group is usually of limited impact. The failure of a critical application running in a traditional standalone environment is immediately felt by the business, but other applications can continue to run unaffected. In contrast, an outage of a consolidated environment supporting an organization's entire development staff, or multiple applications used by numerous departments, has a crippling effect on the business. Equally crippling would be an interruption in service at a cloud provider where such applications are running.

The Oracle Maximum Availability Architecture (Oracle MAA) and the MAA reference architectures provide the requisite level of standardization for all databases and Database-as-a-Service (DBaaS) where higher stability, lower downtime and better data protection are of interest. MAA reference architectures address the complete range of availability and data protection required by enterprises of all sizes and lines of business. All reference architectures are based upon a common platform that can be deployed on-premises or on cloud. This approach makes Oracle MAA simpler and less risky to move to the cloud.

This paper describes Oracle MAA reference architectures and the service level requirements that they address using the Oracle Cloud Infrastructure (OCI) Services. It furthermore, discusses some performance and availability results leveraging the existing OCI resources. The paper is most appropriate for a technical audience, that is, architects, directors of IT, and database administrators responsible for designing and implementing DBaaS and moving implementations to the cloud.

Oracle Cloud Infrastructure for MAA

OCI is Oracle's most advanced cloud infrastructure with all of the building blocks for database MAA to support our all Oracle Database customers including our most demanding Enterprise customers. The key building blocks include:

- Bare Metal Single Instance database systems with restart capabilities and redundant local storage
- Virtual Machine single instance database systems with restart capabilities and triple mirrored block storage
- Virtual Machine 2-node Real Application Clusters Systems
- Exadata systems (with various shapes, including quarter, half, and full racks) – best database platform
- Regions, Availability Domains (ADs), and Fault Domains (FDs) to provide outage isolation
- World class scalable networks
    - Secure, high bandwidth, and low latency within ADs and across ADs with Virtual Cloud (VCN) Peering with Public and fully Private subnets
    - Secure, high bandwidth across Regions with VCN Peering
- Scalable backup infrastructure with object storage

Please refer to the OCI documentation for more information and the latest developments within OCI infrastructure.

Cloud Infrastructure	Backup/Restore Options	RAC	ADG	Replication across ADs/Regions
OCI (BM)	Backup to OCI Object Storage (manual/automatic) Automatic backup copies across Availability Domains (ADs)		✓	Across ADs Across Regions via VCN peering
OCI (VM) (with SI or RAC)		✓*	✓	
Exa-OCI (X6/X7)		✓	✓	

* Oracle RAC VMs are placed in separate Fault Domains.

Figure 1 Summary: of OCI Database Infrastructure and Key MAA Components

Availability Domain Isolation

The physical infrastructure and software stack within Oracle Cloud Infrastructure is designed from the ground up to provide failure isolation within and across Availability Domains (AD). Oracle Cloud Infrastructure Service is hosted in Regions and possibly in multiple Availability Domains (ADs). A region is a localized geographic area, and an Availability Domain is one or more data centers located within a region. A region can be composed with three Availability Domains.

All the Availability Domains in a region are connected to each other by a low latency, high bandwidth network, which makes it possible to provide highly available connectivity to the Internet and to customer premises, and to build replicated systems in multiple Availability Domains for both high availability and disaster recovery. Availability Domains have Fault Domains. A Fault Domain is a grouping of hardware and infrastructure within an Availability

Domain. Each AD contains three Fault Domains (FDs). Fault Domains distribute your Oracle Real Application Clusters (Oracle RAC) database nodes and instances so that they are not on the same physical hardware within a single Availability Domain providing another beneficial level of fault isolation. A hardware failure or compute hardware maintenance that affects one Fault Domain does not affect instances in other Fault Domains.

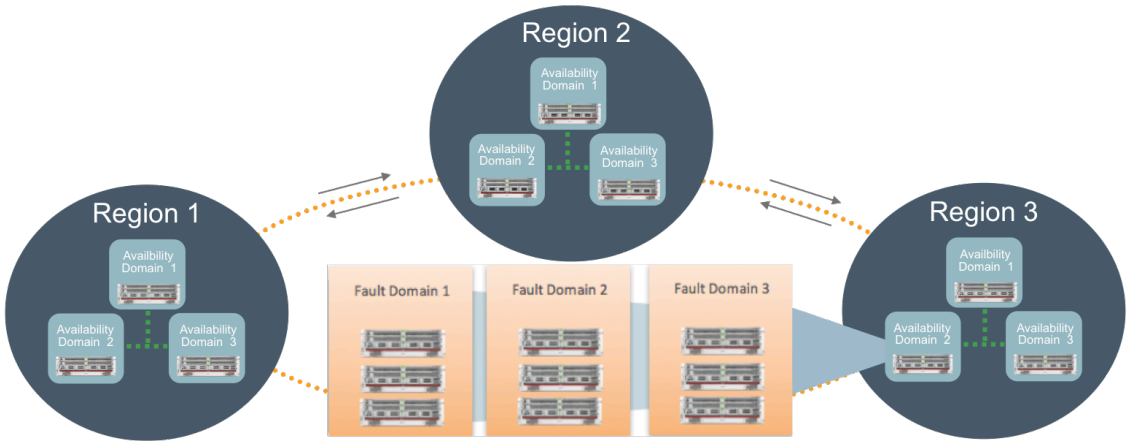


Figure 2 Oracle Cloud Infrastructure Regions, Availability and Fault Domains

Oracle recommends using VCN Peering, which is the process of connecting two VCNs in different regions when configuring Oracle Data Guard across regions. The peering allows the VCNs' resources to communicate using private IP addresses without routing the traffic over the internet or through your on-premises network. Without peering, a given VCN requires an internet gateway and additional public IP addresses. Refer to the OCI documentation for the latest information on regions that support Remote VCN Peering.

High Availability Reference Architectures – Oracle Cloud Infrastructure - Database

Oracle MAA best practices define four High Availability (HA) reference architectures that address the complete range of availability and data protection required by enterprises of all sizes and lines of business. The reference architectures are designated PLATINUM, GOLD, SILVER, and BRONZE as described in Figure 3.

Each architecture uses an optimal set of Oracle HA capabilities to reliably achieve a given service level at the lowest cost and complexity. They address all types of unplanned outages including data corruption, component failure, system, and site outages, as well as planned downtime due to maintenance, migrations, or other purposes. MAA also provides reference architectures for applications using Oracle Sharding and can provide unlimited scalability and unsurpassed availability within OCI if your application is designed and customized for sharded architectures. The MAA Oracle Sharding architecture is described in a different paper.

Figure 3 Oracle Database High Availability Reference Architectures

MAA reference architectures are based on a common infrastructure optimized for the Oracle Database that enables customers to dial in the level of HA appropriate for different service level requirements. This makes it simple to move a database from one HA-tier to the next, should business requirements change, or from one hardware platform to another, or from on-premises to the Oracle Cloud.

Bronze Reference Architecture is appropriate for databases where a simple restart of the database instance, node, or VM and restore from backup is 'HA and DR enough'. Bronze uses HA capabilities such as server and Oracle Clusterware monitoring and restart capabilities included with OCI and Oracle Standard and Enterprise Edition. Bronze is a single instance Oracle 11g database or single instance Oracle 12c or higher Multitenant database for additional consolidation, simplicity, and pluggable database agility. With Multitenant and pluggable databases, customers can relocate a PDB with PDB Relocate or refresh a PDB with PDB Hot Cloning. Oracle Multitenant is an option for database consolidation (multiple pluggable databases in a single container to reduce operational expenses by managing many databases as one, and to reduce capital costs by increasing consolidation density). Bronze relies upon Oracle-optimized backups to OCI object storage using Oracle Recovery Manager (RMAN) to provide data protection within the same region Backups are automatically replicated to another AD for additional isolation and protection.

Silver Reference Architecture is designed for databases that can't afford to wait for a cold restart or a restore from backup should there be an unrecoverable database server outage. Silver builds on the same functionality as the Bronze architecture and adds capabilities that provide a choice of two different patterns for enhancing availability.

1. The primary pattern is to use Oracle RAC. Oracle RAC is an active-active clustering technology for minimal or zero downtime in the event of database instance or server failure, and zero downtime for the most common software updates (operating system, periodic DB/GI software updates). Silver implements best practices to ensure application service failover. Just as with Bronze, RMAN provides database-optimized backups to protect data and restore availability should there be a complete database or cluster outage. A two-node Oracle RAC configuration is available in VM database systems; Oracle's premier Exadata Database Machine provides quarter, half, and full Exadata rack options. If Fault Domains exist in a given AD, then the Oracle RAC compute nodes are placed in different FDs for additional fault isolation for two-node RAC VMs. Exadata remains Oracle's best MAA database platform with Oracle RAC and its additional HA, data protection, HA quality-of-service, and management benefits not found in OCI RAC VM. For more information, refer to: http://www.oracle.com/technetwork/database/features/availability/exadata-maa-best-practices-155385.html

2. An alternative pattern uses Data Guard Fast-Start failover to maintain a local but separate synchronized copy of the production database for HA across ADs, or across Fault Domains if only one AD is available. Each AD is fault tolerant, independent, and isolated. Data Guard or a decoupled standby database copy provides HA for a broader set of database outages, including data corruptions, human error, cluster outages, network faults, and database upgrades. Data Guard synchronous replication with automatic database failover also provides an additional level of data protection. Data Guard is included with Oracle Database Enterprise Edition; there are no additional licensed products or cloud services beyond those used for Bronze databases. This alternative Silver Reference Architecture does NOT provide simple, minimal, or zero downtime for the most common software updates, which is why the Oracle RAC pattern is the preferred MAA Silver reference architecture solution.

Gold Reference Architecture is well suited for service level requirements that cannot tolerate downtime from database, cluster, data corruptions, and site failures and major database upgrade. The Gold reference architecture builds upon the Silver reference architecture Oracle RAC pattern by adding a standby database with Oracle Active Data Guard across availability domains, or across regions if regional protection is required for DR. The primary and standby database systems should be configured symmetrically to ensure that performance service levels are similar after Data Guard role transitions. Some customers may start with fewer licensed cores on the standby during recovery and burst after role transition, but the trade off is a delay in meeting performance SLAs. Data Guard Fast-Start failover must be configured to maintain lowest Recovery Time Objective (RTO). You can configure Data Guard Fast-Start failover with zero data loss across ADs or across regions by using SYNC or Far SYNC transport. Alternatively, data loss is minimal with ASYNC transport and Data Guard Max Performance protection mode.

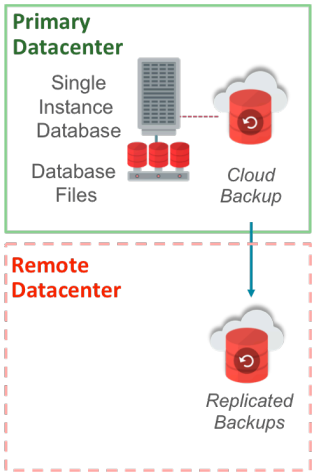
Oracle Active Data Guard across ADs or across regions exists today. You can configure Oracle RAC primary and standby using Oracle RAC VMs or Oracle Exadata Database Machines. Local and Remote VCN (Virtual Connection Network) Peering provides a secure, high bandwidth network across Availability Domains and regions. Refer to the OCI documentation for the latest information about ADs, Fault Domains, and Regions, and their associated VCN pairing support. If the Remote VCN peering option is not available, then the public Internet backbone can be used together with Oracle Net and TDE encryption; however, all regions should have this capability now or very soon.

Platinum Reference Architecture builds on the existing Gold architecture with Oracle GoldenGate for zero downtime upgrades and migrations, and Edition-Based Redefinition for zero downtime application upgrades. The Platinum reference architecture delivers substantial value for the most critical applications where downtime is not an option. Platinum requires the same on-premises features and products as Gold, plus Edition Based Redefinition and Oracle GoldenGate

The following sections describe all reference architectures specific to Oracle Cloud Infrastructure.

Bronze: Development, Test, and Production Databases

The Bronze architecture (shown in Figure 4) provides basic database service protection at the lowest absolute cost. A reduced level of HA and data protection is accepted in exchange for reduced cost and implementation complexity.



Bronze Summary	• *Single instance database with backups & auto-restart capabilities with Oracle Clusterware* • *Optional replication of backup* ◦ *OCI replicates backups across another data center or Availability Domain* • *Restore from backup to resume service following unrecoverable outages*
Features	• *Oracle Restart Capabilities (enable with Oracle Clusterware)* • *Multitenant Database with PDB features (12c+ DBs)* • *Online Maintenance (available manually)* • *Corruption Protection (enabled)* • *Flashback Technologies (enabled)* • *Recovery Manager and Cloud Storage* • *Recovery Appliance (future, available for Cloud at Customer)*

Figure 4: Bronze HA Reference Architecture

Bronze is based on a Single Instance Oracle Database with database server and database instance restart capabilities. For OCI, it can be a bare metal physical server or single instance VM. Each database installation has Oracle Clusterware installed. If the server or VM fails, a restart attempt is automatic. Once the database server is restarted, Oracle Clusterware restarts the Oracle instance, listener and associated service. When a machine becomes unusable or the database unrecoverable, the recovery time objective (RTO) is a function of how quickly a replacement system can be provisioned and a backup is restored. In a worst-case scenario of a complete AD outage there will be additional time required to perform these tasks to restore in a remote location.

Oracle Recovery Manager (RMAN) is used to perform regular backups of the Oracle Database from OCI object storage. The database backups are automatically replicated across ADs. The potential for data loss, also referred to as the recovery point objective (RPO), if there is an unrecoverable outage, is equal to the last available backup up to available contiguous set of archives needed for recovery. Daily backups and frequent archive backups to Cloud object storage can reduce RPO. Database backups to the Oracle Cloud can be leveraged should a disaster strike your existing availability domain.

Bronze uses the following capabilities included with the Oracle Database Enterprise Edition:

- Oracle Clusterware automatically restarts the database, the listener, and other Oracle components after a hardware or software failure, or whenever a database host computer restarts. Oracle Clusterware is pre-installed in all OCI database instances. MAA recommends using Oracle Clusterware managed services for all applications that connect to the database using these services.
- Oracle corruption protection checks for physical corruption and logical intra-block corruptions. In-memory corruptions are detected and prevented from being written to disk, and in many cases can be repaired automatically. For details see Preventing, Detecting, and Repairing Block Corruption for the Oracle Database. For OCI, default database configurations have `DB_BLOCK_CHECKSUM=FULL` enabled. MAA recommends enabling `DB_BLOCK_CHECKING` to MED or FULL if performance impact is minimal.
- Automatic Storage Management (ASM) is an Oracle-integrated file system and volume manager that includes local mirroring to protect against disk failure. MAA recommends that all Bare Metal deployments

and Exadata Database Machines use High Redundancy disk groups. For Exadata systems, it is the only option. For VM database systems, it is automatically deployed with External Redundancy with triple mirroring on block storage.

- Oracle Flashback Technologies provide fast error correction at a level of granularity that is appropriate to repair an individual transaction, a table, or the full database. For OCI database configurations, flashback database is enabled with a couple of hours of flashback retention by default. You can increase the `DB_FLASHBACK_RETENTION_TARGET` if you require longer window to flashback from logical corruption.
- Oracle Recovery Manager (RMAN) enables low-cost, reliable backup and recovery. In OCI, a backup to the OCI object storage can be enabled. The backup APIs have appropriate defaults to enable good backup/restore rates. Performance observations can be found in the MAA observations section below.
- Online maintenance includes online redefinition and reorganization for database maintenance, online file movement, and online patching. These operations can be performed manually.
- Continuous Availability / Application Continuity best practices should be established early on in the application design and configuration process, which will allow better service continuity after any minor outage or planned maintenance activity. Refer to Application Checklist for Continuous Service on the Autonomous Database MAA paper which covers the following
  - Using Clusterware-managed Oracle Services and Fast Application Notification
  - Draining and Rebalancing Sessions for Planned Maintenance
  - Transparent Application Failover (TAF)
  - Application Continuity (AC)
  - Transparent Application Continuity (TAC)

The above practices become more relevant with higher service level MAA reference architectures, especially any architecture that contain Oracle RAC or Data Guard Fast-Start Failover with zero data loss.

Bronze Summary

Table 1 summarizes RTO and RPO service level requirements for the Bronze reference architecture.

TABLE 1: BRONZE RECOVERY TIME (RTO) AND POTENTIAL DATA LOSS (RPO)

Event	Downtime- RTO	Potential Data Loss -RPO
Disk failure	Zero	Zero
Recoverable Database Instance Failure	Minutes	Zero
Recoverable Database Server Failure	Minutes to Hour	Zero
Data corruption, unrecoverable instance, server, database or site failure	Hours to day	Since last backup
Online File Move, Database reorganization or redefinition, Online Patching for eligible one-off patches	Zero	Zero
Hardware and software maintenance and patching	Minutes to Hour	Zero
Database upgrades (patch-sets and full releases)	Minutes to Hour	Zero
Platform migrations	Hours to a day	Zero
Application upgrades that modify back-end database objects	Hours to a day	Zero

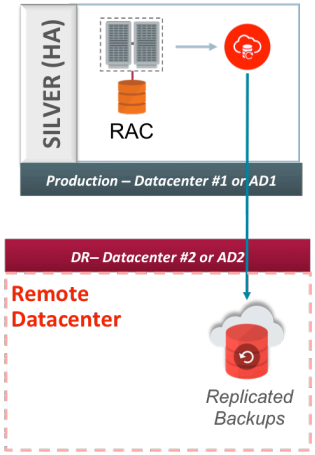
Bronze Requirements: Cloud deployments require a minimum of Oracle Enterprise DBaaS (PaaS) and Oracle OCI Object Storage. If Oracle Multitenant is used for database consolidation, then on-premises deployment also requires a license for Oracle Multitenant and cloud deployment requires subscription of at least Oracle High Performance option for Oracle Cloud Infrastructure - Database).

Silver: Departmental Databases

The Silver reference architecture (Figure 5) is designed for databases that can't afford to wait for a cold restart or a restore from backup should there be an unrecoverable database outage. Silver begins with the same functionality of Bronze and adds capabilities that provide a choice of two different patterns for additional HA.

The MAA recommended silver pattern uses Oracle RAC to enable automatic failover to a second active Oracle instance for HA, and provides a potential zero downtime for the most common set of software updates. Oracle RAC is available on OCI with Oracle RAC VMs and with Exadata.

The alternative pattern uses Data Guard database replication with automatic failover to a completely synchronized copy of the production database in a different availability domain for HA. Similar to Bronze, backups will be sent to the local OCI object storage and replicated to another AD object storage automatically.



Silver Summary	• *Active-Active clustering with Oracle RAC* • *All nodes active at all times* • *Real-time failover*

Features	• *Zero-Downtime rolling maintenance across Oracle RAC instances (in different Failure Domains)*
	  • *Hardware and OS maintenance*
	  • *Qualified Oracle Database patches*
	• *Available in the Oracle Cloud and ExaCC*
	• Benefit from the "Application Checklist for Continuous Service"

Figure 5: Silver HA Reference Architecture Option 1

Silver uses the following capabilities in addition to Bronze, included with Oracle Database Enterprise Edition:

- Oracle RAC enables an Oracle database to run across a cluster of servers, providing fault tolerance, performance, and scalability with no application changes necessary. When there's an instance or node failure, database downtime is essentially zero. Furthermore, Oracle Clusterware automatically restarts failed Oracle RAC instances and managed resources like Oracle listeners. For OCI, Oracle RAC is available as a 2-node Oracle RAC VM or with Exadata.
- With Oracle RAC, the customer has the ability to manage planned maintenance without user interruption and reduce service brownout for instance and node failures. Refer to Application Checklist for Continuous Service for more information.

Silver RAC Summary

Table 2 summarizes RTO and RPO service level requirements for the Silver reference architecture. Highlighted are the additional benefits when compared to the Bronze architecture.

TABLE 2: SILVER RAC RECOVERY TIME (RTO) AND POTENTIAL DATA LOSS (RPO)

Event	Downtime- RTO	Potential Data Loss -RPO
Disk failure	Zero	Zero
Recoverable or unrecoverable RAC instance failure	Seconds	Zero
Recoverable or unrecoverable RAC server failure	Seconds	Zero
Data corruptions, unrecoverable database, Availability Domain or Regional failure	Hours to day	Since last backup
Fault Domain failure (RAC nodes can be configured on separate fault domains within an AD)	Seconds	Zero
Online File Move, Database reorganization or redefinition, Online Patching for eligible one-off patches	Zero	Zero
Hardware and software maintenance and patching	Zero	Zero
Database upgrades (patch-sets and full releases)	Minutes to Hour	Zero
Platform migrations	Hours to a day	Zero
Application upgrades that modify back-end database objects	Hours to a day	Zero

Silver Requirements (Oracle RAC: Cloud deployments require a minimum of Oracle Enterprise DBaaS (PaaS) and Oracle Oracle Cloud Infrastructure Object Storage. Similar to Bronze, if Oracle Multitenant is used for database consolidation then on-premises deployment also requires a license for Oracle Multitenant and cloud deployment requires a minimum of Oracle High Performance DBaaS (Paas).

An alternative Silver MAA pattern uses Data Guard Fast-Start failover to maintain a local but separate synchronized copy of the production database for HA across availability domains, or across Fault Domains when only AD exists.

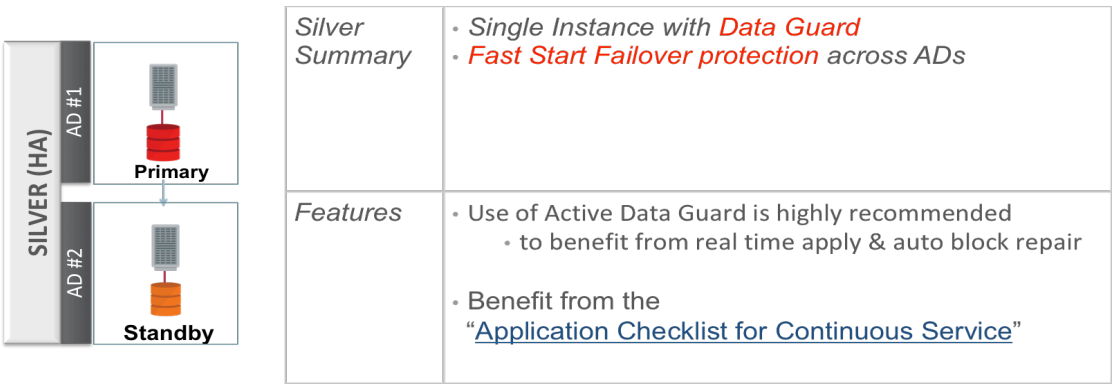


Figure 6: Silver HA Reference Architecture

Each Availability Domain is fault tolerant, independent, and isolated. Data Guard or a decoupled standby database copy provides HA for a broader set of database outages, including data corruptions, human error, cluster outages, network faults, and database upgrades. Data Guard synchronous replication with automatic database failover also provides an additional level of data protection

Silver Active Data Guard Fast-Start Failover Summary

Table 3 summarizes RTO and RPO service level requirements for the Silver reference architecture. Highlighted are additional benefits when compared to the Bronze table. MAA recommends setting up Data Guard Fast-Start failover with Max Availability protection mode. Using a SYNC transport across ADs has minimal performance impact due to high bandwidth network and very low latency between ADs. Refer to MAA Observations later in the paper. From Table 3, Data Guard may provide additional protection for more outages because it is a separate database; however, hardware and software maintenance, which can be quarterly, is not zero or near zero.

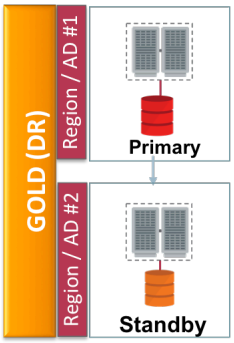
TABLE 3: ALTERNATIVE SILVER WITH ADG FSFO RECOVERY TIME (RTO) AND POTENTIAL DATA LOSS (RPO)

Event	Downtime- RTO	Potential Data Loss -RPO
Disk failure	Zero	Zero
Recoverable or unrecoverable RAC instance failure	Seconds to minute	Zero with SYNC
Recoverable or unrecoverable RAC server failure	Seconds to minute	Zero with SYNC
Data corruptions, unrecoverable database, Availability Domain or Regional failure (depends if standby is in another region)	Seconds to minute	Zero with SYNC
Online File Move, Database reorganization or redefinition, Online Patching for eligible one-off patches	Zero	Zero
Hardware and software maintenance and patching	Minutes to Hour	Zero
Database upgrades (patch-sets and full releases)	Seconds to minute	Zero
Platform migrations	Seconds to minute	Zero
Application upgrades that modify back-end database objects	Hours to a day	Zero

Alternative Silver Requirements (Data Guard Fast Start Failover): Cloud deployments require a minimum of Oracle Enterprise DBaaS (PaaS) and Oracle Oracle Cloud Infrastructure Object Storage. Similar to Bronze, if Oracle Multitenant is used for database consolidation then on-premises deployment also requires a license for Oracle Multitenant, and cloud deployment requires a minimum of Oracle High Performance DbaaS (Paas).

Gold: Business and Mission Critical Databases

The Gold reference architecture is ideal for service level requirements that cannot tolerate data center failures or possibly even regional site failures. It builds upon Silver with Oracle RAC but requires an Oracle Active Data Guard standby database that provides better data protection by enabling auto-block repair for physical data corruptions, and a remote standby to address complete database, cluster, or data center outage.



Gold Summary	• Active-Active clustering with Oracle RAC
	• All nodes active in each Availability Domain (AD)
	  • Real-time failover
	• Real-time data protection, HA & DR using Active Data Guard
	  • Best corruption protection
	  • Zero or near-zero data loss
	  • Offload read-only and backups

Features	• Minimal Downtime for Database Upgrades using DBMS_Rolling or transient logical standby • Automatic DB failover with potential zero data loss • Benefit from the "Application Checklist for Continuous Service"

Figure 7: Gold MAA Reference Architecture

Oracle Data Guard and Oracle Active Data Guard

Oracle Data Guard synchronizes one or more physical copies (standby databases) to eliminate single point of failure for a production database (the primary database). Oracle RAC (when available) enables multiple Oracle instances (running on separate compute nodes) to share access to the same Oracle Database. Data Guard maintains synchronization of completely separate Oracle databases, each having their own Oracle instance.

Data Guard provides the following capabilities:

- Data Guard synchronous replication and Maximum Availability protection modes are used to provide zero data loss protection required for an HA solution. Data Guard transmits changes made on a primary database to a standby database in real-time. Changes are transmitted directly from the log buffer of the primary to minimize propagation delay and overhead, and in order to completely isolate the standby database from corruptions that can occur in the I/O stack of a production database.
- The primary database and its standby copy can be deployed locally in the same region but in different Availability Domains or data centers. Each Availability Domain has its own power, cooling, network, servers, and storage.
- In addition to providing failover options in case there's a database, storage, or availability domain failure, Data Guard performs continuous Oracle validation to ensure that corruption is not propagated from the primary to the standby database. It detects physical and logical intra-block corruptions that can occur independently at either primary or standby databases. It is also unique in enabling run-time detection of silent lost-write corruptions (lost or stray writes that are acknowledged by the I/O subsystem as successful). For more details see My Oracle Support Note 1302539.1 – Best Practices for Corruption Detection, Prevention, and Automatic Repair.
- Data Guard Fast-Start Failover provides automatic database failover. A Data Guard standby is a running Oracle database, it does not need to be restarted to transition to the primary role. An automatic database failover can complete in less than 60 seconds, even on heavily loaded systems. Fast-Start Failover provides HA by eliminating the delay required for an administrator to be notified and respond to an outage.
  - Data Guard uses role-specific database services and the same Oracle client notification framework used by Oracle RAC to ensure that applications quickly drop their connections to a failed database and automatically reconnect to the new primary database. Role transitions can also be executed manually using either a command line interface or in the cloud console. To achieve the integrated transparent client

failover with the fastest role transition times, refer to Application Checklist for Continuous Service and Role Transition Best Practices: Data Guard and Active Data Guard.

- Data Guard performs complete, one-way physical replication of an Oracle database with the following characteristics: high performance, simple to manage, support for all data types, applications, and workloads such as DML, DDL, OLTP, batch processing, data warehouse, and consolidated databases. Data Guard is closely integrated with Oracle RAC, ASM, RMAN, and Oracle Flashback technologies.
- Primary and standby systems are exact physical replicas, enabling backups (in the future) to be offloaded from the primary to the standby database. A backup taken at the standby can be used to restore either the primary or standby database. This provides administrators with flexible recovery options without burdening production systems with the overhead of performing backups. Today OCI backups cannot currently executed on the standby.
- Standby databases can be used to upgrade to new Oracle Patch Sets (for example, patch release 12.1.0.2.180417 to 12.2.01.180417) or new Oracle releases (for example, release 12.2 to 18.1) in a rolling manner. This is done by first upgrading the standby and then switching production to run on the new version. Total downtime is limited to the time required to switch the standby database to the primary production role and transition users to the new primary after maintenance has been completed. The new optimized transient logical standby automated process incurs less than 15 seconds of downtime. Refer to Database Rolling Upgrade using Data Guard or Automated Database Upgrades using Oracle Active Data Guard and DBMS_ROLLING for 12.2 and later database versions.

For more background on why Oracle recommends database replication using Data Guard or Active Data Guard rather than storage-based remote mirroring solutions (for example, SRDF, Hitachi TrueCopy, and so on) refer to an in-depth discussion in Oracle Active Data Guard vs. Storage Remote Mirroring.

Oracle Active Data Guard is a superset of the capabilities that are provided by Oracle Data Guard. The Gold reference architecture uses the following advanced features of Oracle Active Data Guard:

- Choice of zero or near-zero data loss protection. If zero data loss upon a failover is not a requirement, you can choose to deploy Oracle Active Data Guard with asynchronous replication to the remote DR site with Remote VCN peering. If zero data loss is required, then an Oracle Active Data Guard Far Sync instance can be deployed. A far sync instance uses a light-weight forwarding mechanism to enable zero data loss failover even when primary and standby databases are hundreds or thousands of miles apart, without impacting primary database performance. Far sync instances are simple to deploy and transparent to operate. A far sync instance can also be used in combination with the Oracle Advanced Compression Option to enable off-host transport compression to conserve network bandwidth and reduce RPO.
- Offload of read-only workload to an Oracle Active Data Guard standby database open read-only while replication is active. An up-to-date active standby database is ideal for offloading ad-hoc queries and reporting workloads from the production database. This increases ROI in standby systems and improves performance for all workloads by using capacity that would otherwise be idle. It also provides continuous application validation that standby databases are ready to support production workload should an outage occur.
- Fast incremental backups from the standby database using an RMAN block change tracking file. Fast incremental backups complete up to 20x faster than traditional incremental backups. Today OCI backups cannot currently executed on the standby.
- Automatic repair of block-level corruption caused by intermittent random I/O errors that can occur independently at either primary or standby databases. Oracle Active Data Guard retrieves a good copy of the block from the opposite database to perform the repair. No application changes are required and impact of the corruption is transparent to the user.

Gold Summary

RTO and RPO service level requirements addressed by Gold are summarized in Table 4. With Gold, you can reap the benefits of Oracle RAC and Oracle Active Data Guard providing comprehensive data protection and low RTO and RPO for all outages except application upgrades.

TABLE 4: GOLD RECOVERY TIME (RTO) AND POTENTIAL DATA LOSS (RPO)

Event	Downtime- RTO	Potential Data Loss -RPO
Disk failure	Zero	Zero
Recoverable or unrecoverable RAC instance failure	Seconds	Zero
Recoverable or unrecoverable RAC server failure	Seconds	Zero
Data corruptions, unrecoverable database, Availability Domain or Regional failure (depends if standby is in another region)	Seconds to Minute	Zero with SYNC
Online File Move, Database reorganization or redefinition, Online Patching for eligible one-off patches	Zero	Zero
Hardware and software maintenance and patching	Zero	Zero
Database upgrades (patch-sets and full releases)	Seconds	Zero
Platform migrations	Seconds to Minute	Zero
Application upgrades that modify back-end database objects	Hours to day	Zero

Gold Requirements: On premises deployment as a DR site requires Oracle Enterprise Edition, Oracle Active Data Guard, Oracle Multitenant (optional for database consolidation) and Oracle Enterprise Manager life-cycle management, diagnostic, and tuning packs. Cloud deployment requires a minimum of Oracle Extreme Performance DBaaS (PaaS) or Exadata cloud services. Gold also uses Oracle Database Backup Cloud services.

Platinum: Extremely Critical Databases

The Platinum reference architecture (shown in Figure 7) builds upon the Gold architecture by deploying an extra level of redundancy and several advanced HA capabilities. Platinum is ideal for applications that have extremely low, if any, tolerance for downtime or data loss. In this architecture, all of the application best practices must be incorporated as described in Continuous Availability - Application Checklist for Continuous Service on the Autonomous Database as a prerequisite.

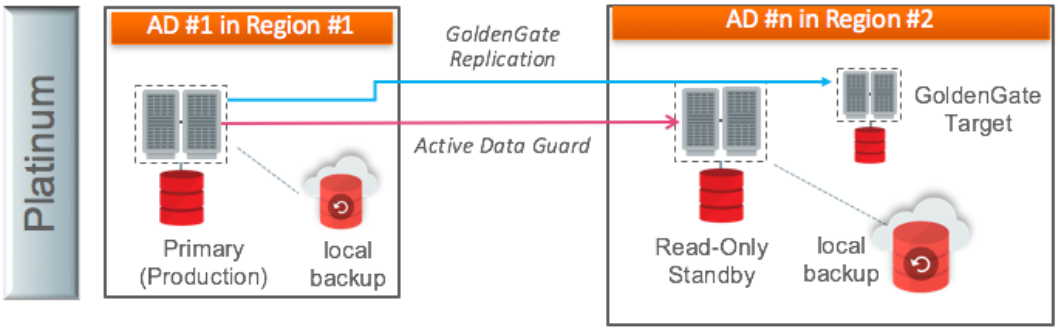


Figure 8: Platinum HA Reference Architecture

Platinum uses Oracle GoldenGate and Edition-Based Redefinition to enable zero downtime maintenance, migrations, and application upgrades.

Oracle GoldenGate

Oracle GoldenGate enables logical replication to maintain a synchronized copy (target database) of the production database (source database). The target database contains the same data, but is a different database from the source (for example, backups are not interchangeable). Oracle GoldenGate logical replication is a more sophisticated process that has a number of prerequisites that do not apply to Data Guard physical replication. In return for these prerequisites, Oracle GoldenGate provides unique capabilities to address advanced replication requirements. Refer to MAA Best Practices: Oracle Active Data Guard and Oracle GoldenGate for additional insights on the tradeoffs of each replication technology and requirements that may favor the use of one versus the other, or the use of both technologies in a complementary manner.

The Platinum reference architecture uses Oracle GoldenGate bi-directional replication to implement zero downtime maintenance and migrations. In such a scenario:

- Maintenance is first implemented at a target database.
- Source and target are synchronized across versions using Oracle GoldenGate replication.
- Once the new version or platform is synchronized and stable, bi-directional replication enables users to gradually migrate to the new platform with zero downtime. Users naturally terminate their connections to the database operating at the prior version when their work is complete and new connections are directed to the database running the new version. Bi-directional replication keeps old and new versions in sync throughout the migration process. This also provides for a fast fall back option should any unanticipated issues arise with the new version as workload grows.

Though the process is not trivial, Oracle GoldenGate bi-directional replication may also be used for application upgrades that modify back-end database objects. Developer-level knowledge of the database objects being modified or added by the new release is required in order to enable Oracle GoldenGate to replicate across versions. Implementing cross-version mapping using Oracle GoldenGate replication is required for each new release of the application.

Bi-directional replication can also be used to increase availability service levels when a continuous read-write connection to multiple copies of the same data is required. It is important to note that bi-directional replication is not application transparent. It requires conflict detection and resolution when changes are made to the same record at the same time in multiple databases. It also requires careful consideration of the impact of different failure states and replication lag.

It is also important to note that Oracle GoldenGate replication is an asynchronous process that is not able to provide the same zero data loss protection as Data Guard and Oracle Active Data Guard. This is one reason why Data Guard and Oracle Active Data Guard are used by the Silver, Gold, and Platinum reference architectures to provide HA during unplanned outages, because there is an assumption that an HA event should result in zero data loss. Oracle GoldenGate replication may be used in place of Data Guard or Oracle Active Data Guard where zero data loss protection is not a concern.

Many of our Platinum MAA customers use both Data Guard Fast-Start Failover and GoldenGate complementary to achieve zero downtime upgrade and migration solution and still maintain zero data loss for database failures. Refer to Transparent Role Transitions with Oracle Data Guard and Oracle GoldenGate for more details.

There is no concern for data loss during planned maintenance when Oracle GoldenGate is used, as long as the production copy of the database is protected by a Data Guard standby.

Edition-Based Redefinition

Edition-Based Redefinition enables online application upgrades that require changes to database objects that would otherwise require the database to be offline. EBR enables all changes to be implemented while the previous version of the application and the database remain online. When the upgrade process is complete, the pre-upgrade application and the post-upgrade application can be used at the same time against the same copy of the Oracle database. Existing sessions can continue to use the pre-upgrade version until their users decide to end them, and new sessions can use the post-upgrade version. When there are no longer any sessions using the pre-upgrade version of the application, the pre-upgrade version can be retired.

EBR enables online application upgrades in the following manner:

- Code changes are installed in the privacy of a new edition.
- Data changes are made safely by writing only to new columns or new tables not seen by the old edition. An editioning view exposes a different projection of a table into each edition to allow each to see just its own columns.
- A cross-edition trigger propagates data changes made by the old edition into the new edition's columns, and vice-versa.

Similar to Oracle GoldenGate zero downtime application upgrades, the use of EBR requires deep knowledge of the application and a non-trivial effort on the part of the developer to incorporate it. Unlike Oracle GoldenGate, there is a one-time investment to implement EBR. From that point forward, minimal effort is required to use EBR for subsequent new releases of the application. EBR has shown to be usable even for the most complex applications; for example, Oracle E-Business Suite 12.2 uses EBR for online patching. EBR is a feature included with Oracle Database at zero additional cost.

Platinum Summary

RTO and RPO service level requirements addressed by the Platinum reference architecture are summarized in Table 5. The assumption is that application continuity is required and can mask outages. Also, Oracle GoldenGate and Edition-Based Redefinition are leveraged for zero application downtime.

TABLE 5: PLATINUM RECOVERY TIME (RTO) AND POTENTIAL DATA LOSS (RPO)

Event	Downtime- RTO	Potential

		Data Loss -RPO
Disk failure	Zero	Zero
Recoverable or unrecoverable RAC instance failure	Zero or Seconds	Zero
Recoverable or unrecoverable RAC server failure	Zeor or Seconds	Zero
Data corruptions, unrecoverable database, Availability Domain or Regional failure (depends if standby is in another region)	Zero or Seconds	Zero with SYNC
Online File Move, Database reorganization or redefinition, Online Patching for eligible one-off patches	Zero	Zero
Hardware and software maintenance and patching	Zero	Zero
Database upgrades (patch-sets and full releases)	Zero or Seconds	Zero
Platform migrations	Zero	Zero
Application upgrades that modify back-end database objects	Zero	Zero

Platinum Requirements: On-premises deployment as a DR site requires Oracle Enterprise Edition, Oracle RAC, Oracle Active Data Guard, Oracle GoldenGate, Oracle Multitenant (optional for database consolidation), and Oracle Enterprise Manager life-cycle management, diagnostic, and tuning packs. Cloud deployments require a minimum of Oracle Extreme Performance DBaaS (PaaS) or Exadata cloud services and Oracle GoldenGate cloud service. Platinum also uses Oracle Database Backup Cloud services.

Oracle Cloud Infrastructure – MAA Proof of Concepts & Configuration Practices

This section highlights some of the very promising performance and HA benefits that were observed and validated.

Corruption Prevention, Detection, and Auto-Repair

The ability of the Oracle Database to prevent, detect and automatically repair data corruption is common across all MAA reference architectures. Each check is unique to the Oracle Database, using specific knowledge of Oracle data blocks and redo structures to increase HA and to protect data by preventing the spread of data corruptions that can occur in memory or due to faults in the I/O stack. These capabilities are summarized in Table 6. The “Type”-column in Table 6 indicates when validations for physical and logical corruption are performed.

- Manual checks are initiated by the administrator or at regular intervals by a scheduled job that performs periodic checks.
- Runtime checks are automatically executed on a continuous basis by background processes while the database is open.
- Background checks are run on a regularly scheduled interval, but only during periods when resources would otherwise be idle.

TABLE 6: CORRUPTION PRENTION, DETECTION, AND AUTO-REPAIR

Type	Reference	Capability	Physical Block Corruption	Logical Block Corruption

	Architecture (s)			
Manual	All	Dbverify, Analyze	Physical block checks	Logical checks for intra-block and inter-object consistency
Automatic with OCI backup APIs	All	RMAN	Physical block checks during backup and restore	Intra-block logical checks
Runtime	Silver – Pattern 2, Gold and Platinum	Data Guard, Active Data Guard	Physical block checking at standby Strong isolation between primary and standby eliminates single point of failure Automatic database failover	Detect lost-write corruption, auto shutdown and failover Intra-block logical checks at standby
Runtime	Gold and Platinum	Active Data Guard	Automatic repair of physical corruptions	
Runtime	All	Oracle block checksum and block checking Enabled by default for all OCI Data Guard deployed systems.	In-memory block and redo checksum	In-memory intra-block logical checks
Runtime	All	ASM	Automatic corruption detection and repair using local extent pairs when using ASM software redundancy (Exadata) instead of external redundancy (RAC VMs).	
Runtime	All – Exadata only[1]	Exadata[1]	HARD checks on write	HARD checks on write
Background	All – Exadata only[1]	Exadata[1]	Automatic Disk Scrub and Repair	

Note that HARD validation and the Automatic Disk Scrub and Repair are unique to Exadata storage. HARD validation ensures that Oracle Database does not write physically corrupt blocks to disk. Automatic Hard Disk Scrub and Repair inspects and repairs hard disks with damaged or worn out disk sectors (cluster of storage) or other physical or logical defects periodically when there are idle resources. Exadata sends a request to ASM to repair the bad sectors by reading the data from another mirror copy. By default the hard disk scrub runs every two weeks.

OCI Cloud Backup and Recovery with Object Storage

The following is a discussion about the best practices when backing-up or recovering to or from the Oracle Cloud Backup Service with Object Storage. Unless stated otherwise, the MAA recommended settings are the default settings used by the OCI cloud backup APIs starting with the November 2018 database infrastructure software.

- Ensure you have installed the Oracle Cloud Backup module from OTN and you configure your RMAN environment properly. With automatic backups, this is already taken care of by the backup agent.

```
RMAN>CONFIGURE CHANNEL DEVICE TYPE 'SBT_TAPE' PARMS
'SBT_LIBRARY=/home/oracle/OPC/lib/libopc.so,
ENV=(OPC_PFILE=/u01/products/db/12.1/dbs/opcodbs.ora)';
```

- To optimize data transfer while minimizing CPU overhead, MAA recommends the following default settings
    a. Set the RMAN `COMPRESSION` to LOW, except when the associated tablespace is already using OLTP or HCC compression. Set RMAN `COMPRESSION` to MED provides more value but incurs more CPU utilization.

```
RMAN> CONFIGURE COMPRESSION ALGORITHM 'LOW';

RMAN>BACKUP DEVICE TYPE SBT AS COMPRESSED BACKUPSET DATABASE PLUS ARCHIVELOG
FORMAT '%d_%U';
```

    b. Set RMAN `PARALLELISM` equivalent to 4 per database server. For 2 node Oracle RAC, set `PARALLELISM` to 8.

```
RMAN> CONFIGURE DEVICE TYPE 'SBT_TAPE' PARALLELISM 8 BACKUP TYPE TO
BACKUPSET;
```

- Use `MULTISECTION` default setting of 64 GBs,

  The purpose of multisection backups (available starting with Oracle Database 11g) is to enable RMAN channels to back up a single large file in parallel. RMAN divides the work among multiple channels, with each channel backing up one file section in a file. Backing up a file in separate sections can improve the performance of backups of large data files. For example, suppose that the tablespace contains a single datafile of 800 MB and assume that four SBT channels are configured, with the `PARALLELISM` setting for the SBT device set to 4. You can break up the datafile in this tablespace into file sections as shown below. For smaller datafiles, the `SECTIONSIZE` setting is ignored.

- Use the "weekly full and daily incremental" strategy

  The goal of an incremental backup is to back up only those data blocks that have changed since a previous backup. This has a lot of benefits, but prior to moving toward this standard approach, you should evaluate if your RTO requirements can still be met. The advantages of this strategy are:

    a. Reduce the amount of time needed for daily backups. Since backup times are shorter, you have an option to backup more frequently as well to reduce RPO.
    b. Reduce network usage and network bandwidth requirements when backing up over a network.
    c. Reduce backup overhead and read I/Os.

  The trade off is that the restore and recovery time is longer because you must restore the previous cumulative backup and subsequent incremental plus redo to recover the database.

- Backup archives every hour to reduce RPO or data loss

  Archive backups are automatic and executed every hour on any archives that have not been backed up previously. Since archives are automatically managed by the Fast Recovery Area, there are no user action to manage or purge the archives.

A more detailed Cloud OCI backup/restore write up will be published in a subsequent paper. Figure 8 serves as an example of MAA performance observations. OCI backup/restore APIs are being changed to use the MAA default recommendations. Most of the performance numbers were based on Exadata on OCI. It was stated that smaller OCI compute or Oracle RAC VM shapes may have network and object storage resource management controls, which implies that some backup/restore rates may be lower.

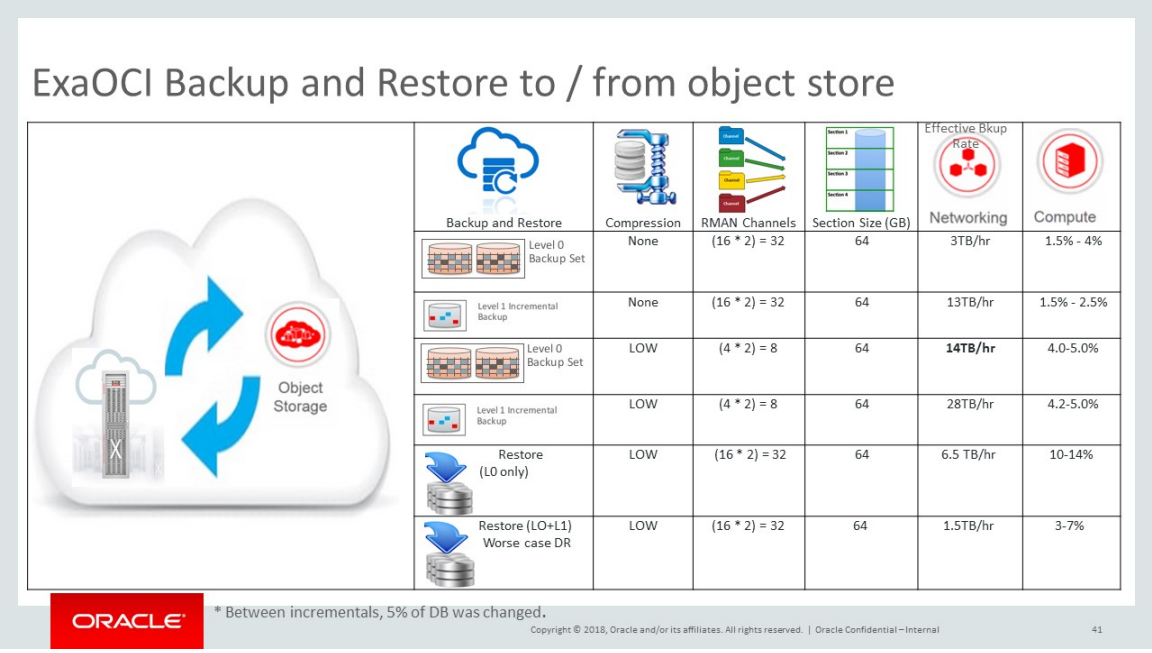


Figure 9: Backup/Restore Performance Observations

Data Guard Fast-Start Failover

Oracle MAA best practices require deploying Data Guard Fast-Start Failover for all Data Guard cloud deployments that require low recovery time objectives. Deploying Fast Start Failover in OCI can be manually configured using the following steps.

Configuring Fast-Start Failover

1. If they do not already exist, create Oracle Net aliases that connect to the primary and the standby databases. If this is an Oracle RAC configuration, the aliases should connect using the SCAN name.

```
chicago =
(DESCRIPTION =
 (ADDRESS_LIST =
   (ADDRESS=(PROTOCOL= TCP)
    (HOST=prmy-scan)(PORT=1521)))
(CONNECT_DATA =
  (SERVER = DEDICATED)
   (SERVICE_NAME = chicago)))
boston =
(DESCRIPTION =
 (ADDRESS_LIST =
   (ADDRESS=(PROTOCOL= TCP)
    (HOST=stby-scan)(PORT=1521)))
(CONNECT_DATA =
  (SERVER = DEDICATED)
   (SERVICE_NAME = boston)))
```

2. Create static SID entries for listeners for both the primary and standby. If Oracle Clusterware is installed, add static SID entries into the local node listener.ora, located in the grid infrastructure home on all hosts in the configuration. **Note that Static "_DGMGRL" entries are no longer needed as of Oracle Database 12.1.0.2 and later in Oracle Data Guard Broker configurations that are managed by Oracle Restart, Oracle RAC One Node, or Oracle RAC as the Broker uses the Oracle Clusterware or Oracle Restart to restart an instance** (See MOS note 1387859.1 for more information).

```
LISTENER = (DESCRIPTION =
     (ADDRESS_LIST=(ADDRESS=(PROTOCOL=tcp)(HOST=host_name)
     (PORT=port_num))))
SID_LIST_LISTENER=(SID_LIST=(SID_DESC=(SID_NAME=sid_name)
     (GLOBAL_DBNAME=db_unique_name_DGMGRL.db_domain)
     (ORACLE_HOME=oracle_home)
     (ENVS="TNS_ADMIN=oracle_home/network/admin"))
```

You should also be aware of the following additional considerations in an Oracle Clusterware environment.

- The static service must be set in the listener.ora file in the `GRID_HOME` of all the nodes.
- The `ORACLE_HOME` listener.ora parameter in the static service definition must be set to the `ORACLE_HOME` of the instance, not the `GRID_HOME`.
- The `ENVS` listener.ora parameter must be used in the static service definition to explicitly set the `TNS_ADMIN` environment variable to the appropriate network admin directory, which is usually the network admin directory of the Oracle Home for the Oracle database.
- In an Oracle RAC One Node or Policy Managed Oracle RAC environment, the `SID_NAME` of each possible instance must be specified in `SID_LIST`. The `SID_NAME` of each instance must match the `INSTANCE_NAME` database initialization parameter of that instance.

3. Restart or reload all of the listeners where the above modification was made (primary and standby nodes).

4. On a primary host connect with DGMGRL and create the configuration.

```
[oracle@exa503 /etc]$ dgmgrl sys/password
DGMGRL> create configuration 'dg_config' as primary database is 'chicago' connect
identifier is chicago;

Configuration "dg_config" created with primary database "chicago"

DGMGRL> add database 'boston' as connect identifier is boston;

Database "boston" added

DGMGRL> enable configuration;
Enabled.
```

5. Verify that the configuration was created successfully by using the `SHOW CONFIGURATION` command.

```
DGMGRL> show configuration;
Configuration - dg_config
Protection Mode: MaxPerformance
Databases:
chicago - Primary database
boston - Physical standby database

Fast-Start Failover: DISABLED

Configuration Status:
SUCCESS
```

6. Flashback database is required to reinstate a failed primary after a failover role transition. Enable flashback on both the primary and standby.

```
DGMGRL> CONNECT sys/<password>@chicago
DGMGRL> SQL "ALTER DATABASE FLASHBACK ON";
DGMGRL> CONNECT sys/<password>@boston
DGMGRL> EDIT DATABASE boston SET STATE=APPLY-OFF;
DGMGRL> SQL "ALTER DATABASE FLASHBACK ON";
DGMGRL> EDIT DATABASE boston SET STATE=APPLY-ON;
```

   Per Oracle's recommendation, set `DB_FLASHBACK_RETENTION_TARGET` to 120 minutes (default value is 1440) if only used for fast start reinstantiation following a Data Guard-based failover. Note that if Flashback Database serves the additional function of providing fast point in time recovery for protection against user error and corruption, an extended flashback retention period should be set for the amount of time deemed necessary to achieve these goals.

7. If the primary and standby are release 11.2.0.4, and you want to use Maximum Availability, set the `LogXptMode` database property for both the primary and target standby databases to `SYNC`. If the primary and standby are release 12.1.0.1 or 12.2.0.1 you can set the `LogXptMode` database property for both the primary and target standby databases to `FASTSYNC`.

```
DGMGRL> EDIT DATABASE 'chicago' SET PROPERTY LogXptMode=FASTSYNC;
DGMGRL> EDIT DATABASE 'boston' SET PROPERTY LogXptMode=FASTSYNC;
DGMGRL> EDIT CONFIGURATION SET PROTECTION MODE AS MaxAvailability;
```

   If you are more concerned about the performance of the primary database than zero data loss, or if the latency between primary and standby is unpredictable and not consistently <2 ms, enable fast-start failover and set the configuration protection mode to Maximum Performance. In this mode you must consider how much data loss is acceptable in terms of seconds and set the `FastStartFailoverLagLimit` configuration property accordingly. This property specifies the amount of data, in seconds, that the target standby database can lag behind the primary database in terms of redo applied. If the standby database's redo apply point is within that many seconds of the primary database's redo generation point, a fast-start failover is allowed. In the following example it is permissible to lose up to 60 seconds worth of redo during a failover in this mode.

```
DGMGRL> EDIT DATABASE 'boston' SET PROPERTY LogXptMode=ASYNC;
DGMGRL> EDIT DATABASE 'chicago' SET PROPERTY LogXptMode=ASYNC;
DGMGRL> EDIT CONFIGURATION SET PROTECTION MODE AS MaxPerformance;
DGMGRL> EDIT CONFIGURATION SET PROPERTY FastStartFailoverLagLimit=60;
```

8. Set the `FastStartFailoverThreshold` property to specify the number of seconds you want the observer and target standby database to wait (after detecting the primary database is unavailable) before initiating a failover.

```
DGMGRL> EDIT CONFIGURATION SET PROPERTY FastStartFailoverThreshold = <SECONDS>;
```

   A fast-start failover occurs when the Observer and the Standby database both lose contact with the production database for a period of time that exceeds the value set for `FastStartFailoverThreshold`, and when both parties agree that the state of the configuration is synchronized (Maximum Availability), or that the lag is not more than the configured `FastStartFailoverLagLimit` (Maximum Performance). An optimum value for `FastStartFailoverThreshold` weighs the trade-off between the fastest possible failover (thus minimizing downtime), and unnecessarily triggering failover due to fleeting network irregularities or other short-lived events that do not have material impact on availability. The default value set when Fast-Start Failover is enabled is 30 seconds. Recommended settings for `FastStartFailoverThreshold` are:

- Single instance production database, low latency, reliable network = 15 seconds
- Exadata Oracle RAC production database on OCI = minimum 30 seconds.
- Oracle RAC VM production database on OCI = minimum 60 seconds.

9. Enable fast-start failover with DGMGRL by issuing the `ENABLE FAST_START FAILOVER` command while connected to any database in the broker configuration, including on the observer computer.

```
DGMGRL> ENABLE FAST_START FAILOVER;
```

Fast-Start Failover Observer Placement Best Practices

In an ideal state fast-start failover is deployed with the primary, standby, and observer database, each within their own separate availability domain. However, configurations with only two availability domains, or even a single availability domain are supported. Described below are the three options and how to configure for observer placement:

- Option 1: AD1: Primary, AD2: Observer , AD3: Standby
- Option 2 with only 2 ADs: AD1: Primary and Observer in another fault domain, AD2: Standby on separate hardware and storage
    A. Role transition: primary and standby change roles and observer moves to the primary's AD. In 18.1, use preferred observer and move observer behavior for failover operation.
    B. Consideration for this option is that if AD1 is lost then no automatic failover will occur. The failover would have to be performed manually after disabling FSFO.
    C. If observer goes down in the primary (AD1), then attempt to restart the observer in AD1.
- Option 3: only 1 AD in the region
    a. Primary, standby, and observer are separate fault domains.
    b. Consideration that if you lose the entire AD you have no HA.

Observers should be installed and run on a computer system that is separate from the primary and standby systems. Observers are very lightweight and can be put on the smallest OCI VM shape. For Oracle Database releases 11.2 and 12.1, you must have a custom script to monitor the observer and attempt to restart on current target or in an alternate target.

Ideally, you should run the observer on a system that is on a separate AD from the primary and standby databases. You can also install the observer on the same network as the application to represent the same application connectivity. If a third, independent location is not available, then locate the observer in the primary AD on a separate fault domain and isolate the observer as much as possible from failures affecting the standby database.

To start an observer, you must be able to log in to DGMGRL with an account that has the SYSDG or SYSDBA privilege. An observer is an OCI client that connects to the primary and target standby databases using the same SYS credentials you used when you connected to the Oracle Data Guard configuration with DGMGRL. Note that the following example starts the observer using the `IN BACKGROUND` clause that is only available in Oracle Database release12.2. If you are starting the observer in version 12.1 or 11.2, omit that clause.

```
DGMGRL> sys/welcome1@boston
DGMGRL> START OBSERVER IN BACKGROUND
FILE IS /net/sales/dat/oracle/broker/fsfo.dat
```

```
LOGFILE IS /net/sales/dat/oracle/broker/observer.log
CONNECT IDENTIFIER IS sales_p
```

If your primary and standby database are release 12.2.0.1 you can register up to three observers to monitor a single Data Guard broker configuration. Each observer is identified by a name that you supply when you issue the `START OBSERVER` command. You can also start the observers as a background process.

```
DGMGRL> sys/welcome1@boston
DGMGRL> start observer number_one in background;
```

On the same host or a different host you can start additional observers for High Availability (release 12.2.0.1):

```
DGMGRL> sys/welcome1@boston
DGMGRL> start observer number_two in background;
```

Only the master observer can coordinate fast-start failover with the Data Guard broker. All other registered observers are considered to be backup observers.

If the observer was not placed in the background (release 12.2 only) then the observer is a continuously executing process that is created when the `START OBSERVER` command is issued. Thus, the command-line prompt on the observer computer does not return until you issue the `STOP OBSERVER` command from another DGMGRL session. To issue commands and interact with the broker configuration, you must connect through another DGMGRL client session.

The following conditions will trigger a fast-start failover:

- Database instance failure (or last instance failure in an Oracle RAC configuration)
- Shutdown abort (or shutdown abort of the last instance in an Oracle RAC configuration)
- Data files taken offline due to I/O errors
- When both the Data Guard observer and the standby database lose their network connection to the production database, and when the standby database confirms that it is in a synchronized state.
- A user configurable condition is one in which the user can specify a condition for which a FSFO is provoked. It is recommend that you leave these user specified condition at the default values. The conditions that can be set up for automatic failover are:
    a) Datafile offline (write error)
    b) Corrupted Dictionary
    c) Corrupted Controlfile
    d) Inaccessible Logfile
    e) Stuck Archiver
    f) ORA-240 (control file enqueue timeout)

Should one of these conditions be detected, the Data Guard observer fails over to the standby, and the primary is shut down, regardless of how the Data Guard broker attribute `FastStartFailoverPmyShutdown` is set. Note that for user specified conditions the fast start failover threshold is ignored and the failover proceeds immediately.

Following unplanned downtime on a primary database that requires a failover, full fault tolerance is compromised until the standby database is reestablished. Full database protection should be restored as soon as possible.

Reinstating databases is automated if you are using Data Guard fast-start failover, and failed instances are automatically restarted. After a fast-start failover completes, the observer automatically attempts to reinstate the original primary database as a standby database. Reinstatement restores high availability to the broker configuration so that, if the new primary database fails, another fast-start failover can occur. The reinstated database can act as

the fast-start failover target for the primary database, making a subsequent fast-start failover possible. The standby database is a viable target of a failover when it begins applying redo data received from the new primary database. If you want to prevent automatic reinstatement (for example, to perform diagnostic or repair work after failover has completed), set the `FastStartFailoverAutoReinstate` configuration property to `FALSE`, but only for debugging purposes.

The `FastStartFailoverAutoReinstate` configuration property controls whether the observer should automatically reinstate the original primary after a fast-start failover occurred, because a fast-start failover was initiated due to the primary database being isolated for longer than the number of seconds specified by the `FastStartFailoverThreshold` property. In some cases, an automatic reinstatement might not be wanted until further diagnostic or recovery work is done. Automatic reinstate is only for failovers that result due to primary isolation and instance crash, not for user defined failovers.

To reinstate the original primary database, the database must be started and mounted, but it cannot be opened.

Fast Start Failover with Far Sync

A far sync instance with fast start failover is configured, as shown in the previous section, using redo routes and the ADD FAR_SYNC DGMGRL command. Consider the following best practices for layout of a configuration with both FSFO and a far sync instance:

Deployment Configuration: 2 regions with two ADs within each region.

- Initial Primary Region will have the primary database in AD1 and HA far sync instances and HA observer in AD2.
- Initial Standby Region will have the standby database in AD1 and HA far sync instances and HA observer (used after role change) in AD2
    a) Far sync uses Oracle RAC for the lowest brownout and reconciliation to zero data loss (1.5 mins). If using Oracle RAC for the far sync instance is not feasible then use alternate destinations understanding that the time to get back to zero data loss after a far sync failure will be higher.
    b) For the HA observer MAA recommends to use at least 2 observer targets in the same primary region but different ADs

Deployment Configuration: 2 Regions and only 1 AD in each region

- Initial Primary regions will have the primary database and 2 servers to host far sync instances and observers. Place the far sync servers in different fault domains compared to primary database.
- Initial Standby Region will have the standby database and 2 servers to host far sync instances and observers (when there is a role change). Best to place the potential far sync servers in different fault domains compared to the standby database.

For more information on far sync instance sizing and far sync Data Guard architectures, refer to Oracle Active Data Guard Far Sync - Zero Data Loss at Any Distance

Fast-Start Failover Performance

With an application using Oracle client failover best practices, the MAA team observed failover outage times of just 8 seconds and an overall downtime of less than 15 seconds for single instance primary and standby in OCI.

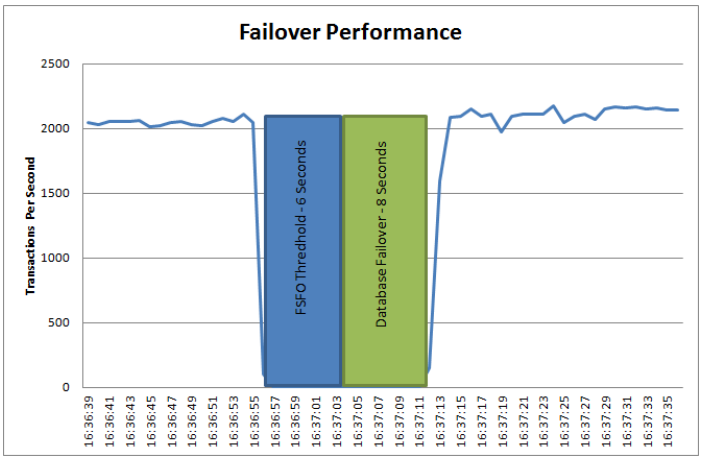


Figure 10: Oracle Database Failover Performance on Bare Metal

Oracle Active Data Guard is always recommended and is a requirement for the Gold and Platinum MAA reference architectures, for its additional benefits such as auto-block repair of physical data corruptions and ability to offload backups and reads to the real-time physical standby database. As represented in Figure 10, Read-Write transactions of the primary can incur a low downtime impact of 15 seconds while the Read-Only transactions on the read-only standby can incur an even lower downtime impact of 6 seconds, maintaining very good availability for all your business transactions.

Concluding, by deploying Data Guard FSFO minimal read-write and read-only downtime can be achieved after primary database failure.

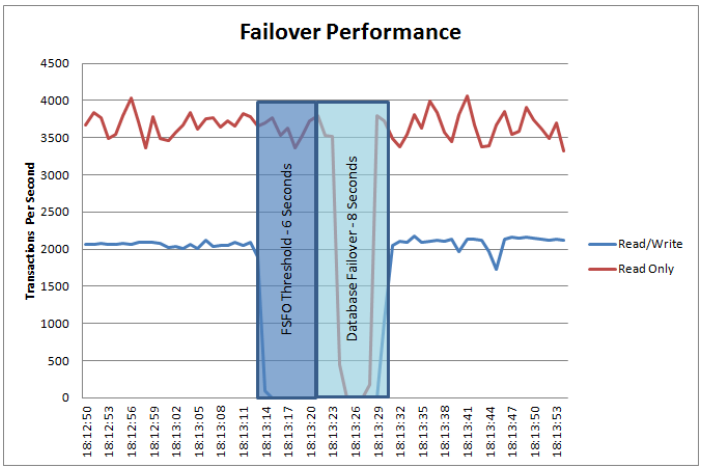


Figure 11: Oracle Database Failover Performance on OCI single instance with single instance standby

MAA recommends using a higher default FSFO threshold for Oracle RAC (60 seconds) and Exadata (30 seconds). The overall failover times can be still less than 1 minute for Exadata within OCI.

(Active) Data Guard SYNC or ASYNC Evaluation

In support of synchronous transport and zero data loss Data Guard configurations, Oracle OCI supports high bandwidth and low latency networks between availability domains suitable to run mission critical applications.

Oracle Data Guard SYNC transport with FSFO is typically configured to provide zero data loss HA. When configured with heavy OLTP Swingbench-generated workload that is 5 times higher than typical OLTP workload, test results have shown that Oracle OCI can easily support synchronous transport with very little overhead on the primary and with near zero data lag on the standby. Table 7 provides an overview of all test results.

	ASYNC	FAST SYNC	SYNC
Redo Rate (MB/sec)	14.93	14.48	14.39
Block Changes/sec (KB/sec)	96.92	94.3	93.86
Txn Rate	2082	2025	2018
% Difference from ASYNC	N/A	97%	97%

Table 7: Oracle OCI – Data Guard Comparisons

Figure 11 shows that enabling SYNC transport had very little performance impact on a very intensive OLTP application with 15 MB/sec redo rate. Typical OLTP application generates less than 3 MB/sec. OLTP transactions are impacted by the SYNC replication due to two-phase commit requirement. However, in Oracle OCI, no application impacts were observed.

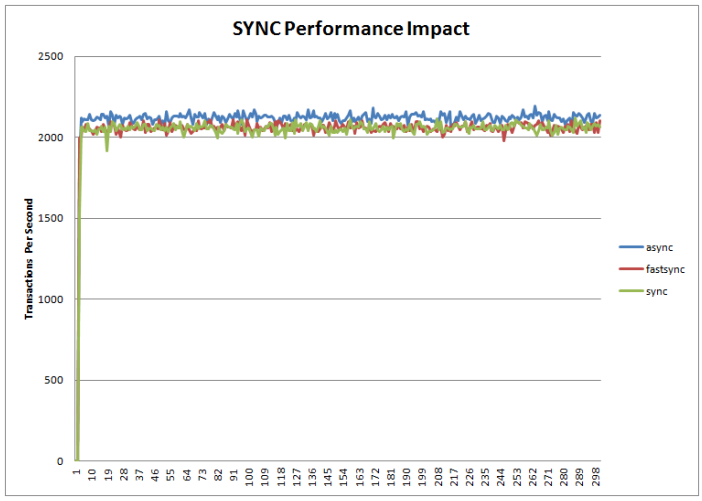


Figure 12:Oracle Database Synchronous Replication Performance on Bare Metal Infrastructure

Data Guard Switchover to Reduce Downtime

Last but not least, Data Guard can help to reduce downtime for planned maintenance activities such as applying patches, Release Updates, and even database upgrades. For details on planned downtime such as standby-first patching, Data Guard switchovers and transient logical standby, refer to MAA Best Practices documentation Reducing Downtime for Planned Maintenance .

The application downtime from the time the service switches to the new primary database, including the Data Guard switchover time, can be less than 2 minutes. With Oracle Database 12c, downtime was observed to be less than 13 seconds in Oracle OCI. For additional planned maintenance documents, please refer to Role Transition Best Practices: Data Guard and Active Data Guard, Database Rolling Upgrade using Data Guard and Automated Database Upgrades using Oracle Active Data Guard and DBMS_ROLLING.

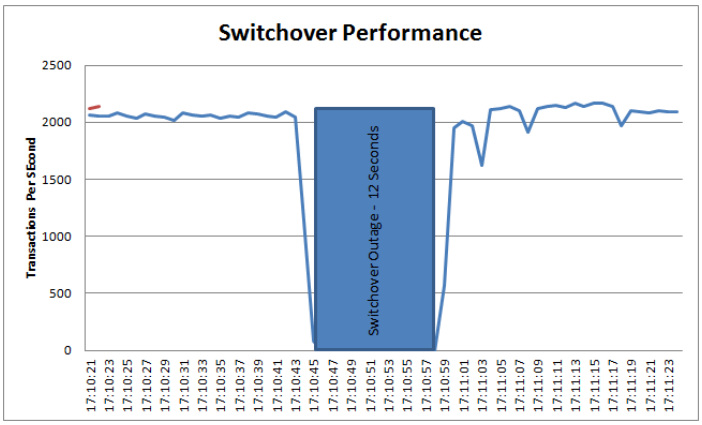


Figure 13: Oracle Database Switchover Performance on Bare Metal Infrastructure

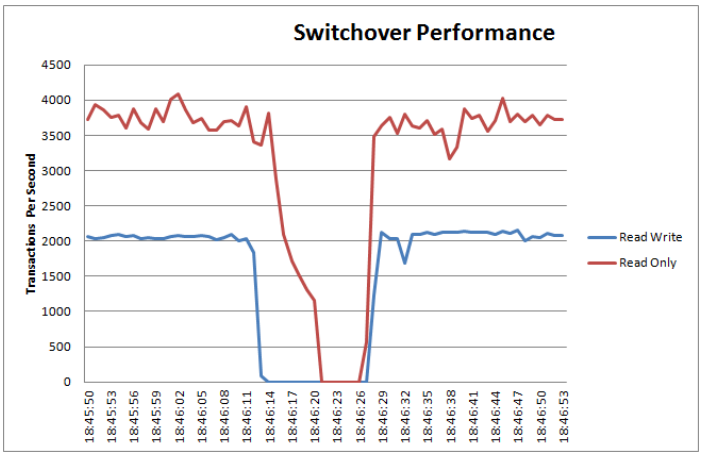


Figure 14: Oracle Database Switchover Performance on Bare Metal Infrastructure

Monitoring a Data Guard Configuration

The broker issues a health check once a minute and updates the configuration status. To force a health check to occur immediately issue the `SHOW CONFIGURATION VERBOSE` command.

On a primary database, the health check determines if the following conditions are met:

- Database is in the state specified by the user, as recorded in the broker configuration file
- Database is in the correct data protection mode
- Database is using a server parameter file
- Database is in the `ARCHIVELOG` mode
- Redo transport services do not have any errors
- Database settings match those specified by the broker configurable properties
- Redo transport settings match those specified by the redo transport-related properties of the standby databases
- Current data protection level is consistent with configured data protection mode
- Primary database is able to resolve all gaps for all standby databases

On a standby database, the health check determines whether the following conditions are met:

- Database is in the state specified by the user, as recorded in the broker configuration file
- Database is using a server parameter file
- Database settings match those specified by the broker configurable properties
- Database guard is turned on when the database is a logical standby database
- Primary and target standby databases are synchronized or within lag limits if fast-start failover is enabled

To identify any warnings on the overall configuration, show the status using the `SHOW CONFIGURATION` command:

```
DGMGRL> show configuration;

Configuration – dg

  Protection Mode: MaxPerformance
  Members:
  tin - Primary database
    can - Physical standby database

Fast-Start Failover: DISABLED

Configuration Status:
SUCCESS   (status updated 18 seconds ago)
```

If the configuration status is `SUCCESS`, everything in the broker configuration is working properly. However, if you see a status of `WARNING` or `ERROR`, then something is wrong in the configuration. Additional error messages will accompany the `WARNING` or `ERROR` status that should be used to identify current issues. The next step is to examine each database in the configuration to narrow down what the specific error is related to.

To identify the warnings on the primary database, show its status using the SHOW DATABASE command.

```
DGMGRL> show database tin

Database - tin

  Role:               PRIMARY
  Intended State:     TRANSPORT-ON
  Instance(s):
    tin1
    tin2

Database Status:
SUCCESS
```

If the database status is SUCCESS then the database is working properly. However, if you see a status of WARNING or ERROR, then something is wrong in the database. Additional error messages will accompany the WARNING or ERROR status that should be used to identify current issues. Repeat the same SHOW DATABASE command on the standby database and assess any error messages.

In addition to the above commands, Data Guard broker features a VALIDATE DATABASE command with Oracle Database 12c Release 1 and later:

```
DGMGRL> validate database tin
  Database Role:    Primary database
  Ready for Switchover:  Yes
DGMGRL> validate database can;

  Database Role:     Physical standby database
  Primary Database:  tin

  Ready for Switchover:  No
  Ready for Failover:    Yes (Primary Running)

  Capacity Information:
    Database  Instances        Threads
    tin       2                2
    can       1                2
    Warning: the target standby has fewer instances than the
    primary database, this may impact application performance

  Standby Apply-Related Information:
    Apply State:      Not Running
    Apply Lag:        Unknown
    Apply Delay:      0 minutes
```

The VALIDATE DATABASE command does not provide a SUCCESS or WARNING status and must be examined to determine if any action needs to be taken.

It is recommended that you run the VALIDATE DATABASE command after the broker configuration has been created, and prior to and after any role transition operation.

The VALIDATE DATABASE command performs the following checks:

- Whether there is missing redo data on a standby database
- Whether flashback is enabled
- The number of temporary tablespace files configured
- Whether an online data file move is in progress
- Whether online redo logs are cleared for a physical standby database

- Whether standby redo logs are cleared for a primary database
- The online log file configuration
- The standby log file configuration
- Apply-related property settings
- Transport-related property settings
- Whether there are any errors in the Automatic Diagnostic Repository (for example, control file corruptions, system data file problems, user data file problems)

Detecting a Transport or Apply Lag using Data Guard Broker

Given enough resources, network bandwidth in particular, a Data Guard standby can maintain pace with very high workloads. In cases where resources are constrained, the standby can fall behind, resulting in a transport or apply lag. A transport lag is the amount of data, measured in time that the standby has not received from the primary. An apply lag is the difference, in elapsed time, between when the last applied change became visible on the standby and when that same change was first visible on the primary.

When using the Data Guard broker, the transport or apply lag can be viewed by using the `SHOW DATABASE` command and referencing the standby database , as shown here.

```
DGMGRL> show database orclsb
Database - orclsb

  Role:               PHYSICAL STANDBY
  Intended State:     APPLY-ON
  Transport Lag:      0 seconds (computed 0 seconds ago)
  Apply Lag:          0 seconds (computed 1 second ago)
  Average Apply Rate: 792.00 KByte/s
  Real Time Query:    ON
  Instance(s):
    orclsb1 (apply instance)
    orclsb2

Database Status:
SUCCESS
```

The Data Guard broker `TransportDisconnectedThreshold` database property (default of 0 in release 11.2 and 30 seconds for releases 12.1 and 12.2) can be used to generate a warning status for a standby when the last communication from the primary database exceeds the value specified by the property. The property value is expressed in seconds. The follow is an example of the warning when a disconnection has occurred:

```
DGMGRL> show database orclsb;

Database - orclsb

  Role:               PHYSICAL STANDBY
  Intended State:     APPLY-ON
  Transport Lag:      0 seconds (computed 981 seconds ago)
  Apply Lag:          0 seconds (computed 981 seconds ago)
  Average Apply Rate: 12.00 KByte/s
  Real Time Query:    OFF
  Instance(s):
    orclsb1 (apply instance)
    orclsb2

  Database Warning(s):
    ORA-16857: member disconnected from redo source for longer than specified
threshold
```

The Data Guard broker also has configurable database properties that can be used to generate warnings when a transport or apply lag exceed a user defined value.

- The `ApplyLagThreshold` configurable database property generates a warning status for a logical or physical standby when the database's apply lag exceeds the value specified by the property. The property value is expressed in seconds. A value of 0 seconds results in no warnings being generated when an apply lag exists. As a best practice, Oracle recommends setting `ApplyLagThreshold` to at least 15 minutes (default of 0 in version 11.2 and 30 seconds for versions 12.1 and 12.2).

- The `TransportLagThreshold` configurable database property can be used to generate a warning status for a logical, physical, or snapshot standby when the database's transport lag exceeds the value specified by the property. The property value is expressed in seconds. A value of 0 seconds results in no warnings being generated when a transport lag exists. As a best practice, Oracle recommends setting `TransportLagThreshold` (default of 0 in release 11.2 and 30 seconds in release 12.1 and 12.2) to at least 15 minutes.

Data Guard Operations in OCI

OCI console and OCI APIs are available to enable Data Guard switchover, failover, and reinstate. As more functionality is added, the MAA team will continue to ensure MAA practices are in place.

Oracle GoldenGate

With Oracle GoldenGate, you can set up an active/active replica for zero downtime migrations, upgrades, or online changes. You can set up Oracle GoldenGate on Oracle OCI manually for deploying the Platinum MAA reference architecture. Refer to Oracle GoldenGate Performance Best Practices and Transparent Role Transitions with Oracle Data Guard and Oracle GoldenGate.

Conclusion

Enterprises need solutions that address the full continuum of requirements for data protection, availability, and disaster recovery. Oracle MAA best practices define four HA reference architectures: BRONZE, SILVER, GOLD, and PLATINUM that address the most typical HA SLAs.

Oracle OCI infrastructure provides an ultimate scalable and available network, compute, and storage environment to support all Oracle applications and databases that require any of the above MAA reference architectures.

Moreover, Oracle OCI's high bandwidth and low latency storage and network infrastructure, along with ability to deploy single instance databases, Oracle RAC Databases. Exadata Database Machine and Data Guard Fast-Start Failover configurations make OCI the best cloud infrastructure to deploy MAA. As more OCI features are added, the MAA team will continue to ensure that MAA architectures, configurations, and life cycle operations are incorporated.

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