Zerto

VERSION 1

July 2015

Protecting Oracle Databases with Zerto Virtual Replication

Table of Contents

1 INTRODUCTION

1.1 Overview

This document covers the best practices of protection Oracle databases in a virtualized environment using Zerto Virtual Replication 4.0 onwards. No warranty or support is provided for any example scripts within this document.

Zerto Virtual Replication, the industry’s first hypervisor-based replication solution for VMware environments, delivers robust business continuity and disaster recovery (BC/DR) which now includes support for Microsoft Hyper-V and can leverage Amazon Web Services as a target environment. It is different from any other BC/DR technology because it has the fastest and most efficient replication combined with fully automated failover, failback and non-disruptive DR testing for production workloads. It is the foundation for any cloud, public, private or hybrid, as it is the only solution which can effectively federate these IT strategies. With Zerto Virtual Replication, customers can easily migrate production workloads across different infrastructures to increase efficiencies. Simple to install, natively multi-tenant, and intuitive to use, Zerto Virtual Replication is ideally suited for private, public and hybrid BC/DR and long term retention for complete data protection.

1.2 Challenges of Protecting Oracle Databases

Oracle Databases are run in 98% of the Fortune 500 and form the backbone of critical applications and business operations. If the Oracle Database availability was impacted due to a disaster or logical failure then the productivity is interrupted, data potentially lost, corporate image tarnished and share price potentially devalued.

The protection of Oracle Database Servers is therefore critical to any organization size and their protection made ever more challenging by large amounts of data change being written on a constant basis inside large databases sometimes terabytes in size. Many different BC/DR technologies can be used to protect Oracle Database Servers, common examples being storage based replication, log shipping or database mirroring software. Each of these different technologies introduces the following risks to the recovery of Oracle Database Servers:

- Multiple solutions require separate skill sets relying on the knowledge of individuals to recover Oracle Database Servers also introducing complexity
- Log shipping and database mirroring software require DBAs to manage and maintain protection in addition being present to manage a disaster recovery plan
- There is no visibility or integration to the virtualized environment running the Oracle Database increasing operational complexity
- Manual recovery operations result in Recovery Time Objectives (RTOs) reliant on the responsiveness and knowledge of individuals
- Testing recovery is complex, time consuming and often requires production downtime making testing ad-hoc and in-frequent
- Constant maintenance can be required to maintain protection introducing risk of time windows of no protection and increases the cost of management overheads

When protecting Oracle Database Servers the application utilizing the databases are often overlooked or not included in the same protection technology or recovery plan. This leads to complications in recovery, lack of visibility into the protection status of the entire application stack, such as SAP, and can make disaster recovery testing incomplete as the entire stack is not tested together leading to increased risk in not being able to recover in a timely and efficient manner.

1.3 Benefits of Protecting Oracle Databases with Zerto

Zerto Virtual Replication resolves all the challenges of protecting Oracle Database servers running in virtualized environments by enabling:

- Simple centralized management of BC/DR from within the hypervisor across all protected VMs using a consistent control interface and one skillset with no DBA knowledge required
- No production changes needed to configure Oracle Database Server protection
- Recovery Point Objectives (RPOs) of seconds with no performance impact
- Application consistent point in time recovery using split-second hot backup mode quiescing
- Journal based protection maintaining write-order fidelity of all writes to all disks within the Oracle VM including database and redo log files
- VM-level integration, KPIs, reporting and protection status alerts
- Multi-VM consistent application recovery for multi-tier Oracle based applications like SAP
- Recovery Time Objectives (RTOs) of minutes, to previous points in time down to increments every few seconds up to 14 days in the past
- No-impact disaster recovery testing of Oracle Database servers and application servers to fully validate the recovery process in working hours in minutes
- Protect Physical or Virtual Raw Device Mappings (RDMs) to virtual disks or pre-provisioned RDMs in the target site
- Temporary Database initial replication, without subsequent change, saving 50% of replication traffic

Some of the largest businesses in the world rely on Zerto to protect their Oracle Database Servers as it simplifies BC/DR while providing the enterprise-class protection and delivers the SLAs required for protecting mission critical business systems like Oracle Database Servers.

1.4 How Zerto Virtual Replication Works

Zerto Virtual Replication Zerto Virtual Replication utilizes Virtual Replication Appliances (VRAs), deployed one per hypervisor host to replicate, protect and recover Virtual Machines (VMs) making the replication storage agnostic. A VRA is required on each hypervisor host to allow continued protection for VMs migrating between hosts and to remove the need for complicated sizing exercises as it forms a scale-out architecture. VRAs are deployed and managed from the Zerto Virtual Manager (ZVM) interface and require no downtime for installation or maintenance.

VMs are replicated by creating a Virtual Protection Group (VPG) in the ZVM interface and selecting the VMs to be protected. A VM can only exist in one protection group at once and VPGs are typically created on a per application basis. All of the VMs placed in a VPG are recovered as together to the same point in time enabling consistent recovery of Multi-VM applications. Write-order fidelity is maintained between all of the VMs and virtual disks within a VPG.

Zerto utilizes 4 different replication modes in order to protect the Oracle Database VMs in a VPG consistently, these are:

1. Continuous Data Protection
2. Bitmap-Sync
3. Initial-Sync
4. Delta-Sync

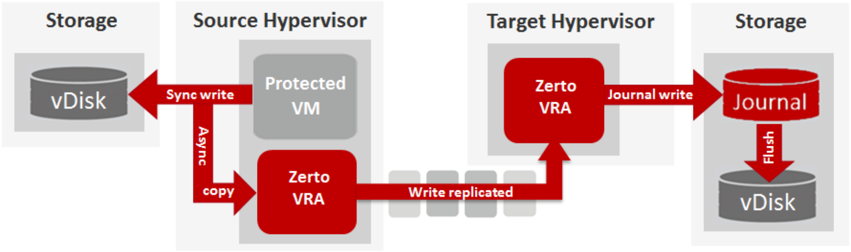
Zerto automatically switches between the replication modes depending on the protection situation. Following is an explanation of the replication mode and when each is used:

Continuous Data Protection

Protected VM writes are copied and sent asynchronously to local VRA, in the memory of the hypervisor, while the write continues to be processed on the protected site (*Image.1*). The local VRA then compresses the write using an LZ4 algorithm and replicates the write to the recovery site VRA delivering a typical RPO of sub 20 seconds on all write operations. The VRA does not interrupt the processing of the write in the protected site, or use any form of snapshot for the replication, thus the VRA is technically not able to slow down the performance of the protected VM. By not requiring confirmation of the write to the VRA it does not introduce any latency.

On the recovery site the write is written to a journal managed by the VRA. Each protected virtual machine has its own journal and every few seconds, a checkpoint is written to the journals. These checkpoints maintain write order fidelity and are crash-consistent with the ability to schedule application-consistent checkpoints on a frequency that does not impact production applications. The journal is stored in the target site and consists of a thin virtual disk on a per VM basis that automatically grows and shrinks depending on the data change rate within the time period specified to keep the journal. After the time period specified on the journal is reached the writes are flushed to the copy of the disk in the target site.

Image.1



Bitmap-Sync

Replication switches from continuous data protection to bitmap-sync to ensure no performance impact in the protected VM in certain scenarios. A bitmap-sync can occur if the write rate on the source disk becomes too high for the VRA to maintain its in-memory buffer, the IP link goes offline between the sites or if the link becomes saturated. Bitmap-sync utilizes a change block tracking mechanism stored in the memory of the hypervisor to maintain an index of the pointers to the most recently changed blocks. Nothing is written to the source storage and the index range of pointers increases over time rather than storing a backlog of changes or increasing the memory usage.

During a bitmap-sync the VM is still recoverable to the point in time before the bitmap-sync occurred and once finished the continuous data protection will re-commence automatically. During a bitmap-sync the RPO will steadily increase until the cause of the bitmap-sync is resolved. Upon resolution the local VRA with reference the latest bitmap data, read the most recently changed blocks on the source storage, compress, replicate and insert the changed blocks in the recovery site journal and automatically self-heal the continuous replication. At this point the RPO will typically return to sub 20 seconds. It is possible to configure a maximum RPO alert to notify any breach in the SLAs defined on a VPG due to a bitmap sync being triggered.

Initial-Sync

When a VM is first protected the source site VRA will read the VM disks, copy, compress and then replicate the data to create an exact copy of the disks in the target site. The read rate will automatically scale up and down depending on the latency to the source or target storage and all of the replication traffic is compressed using a built in LZ4 compression engine. This ensures that the initial-sync is completed as quickly as possible, subject to available bandwidth, without impacting the performance of the protected VM.

To reduce initial-sync time taken it is possible to pre-seed using a pre-imported backup copy of the disks which then utilizes a delta sync as explained below. Once an initial sync has completed the target disks are in a crash consistent state ready for recovery and for continuous data protection to commence.

Delta-Sync

A delta-sync reads both the source and target disks, compares the differences then compresses and replicates the differences to bring the target disk into a consistent state. A delta-sync is utilized when pre-seeding an initial sync replication or after configuring reverse replication.

The delta-sync completes as fast as the storage read rate allows in both source and target, subject to the latency to remove any performance impact, and is also dependent on the amount of data changed. After a delta-sync has completed the target disks are in a crash consistent state for recovery and for continuous data protection to commence.

2 PRODUCTION SUPPORT

2.1 VMware Support

As per the Oracle Databases on VMware Best Practices Guide, *Ref.1*, VMware states the following with regards to Oracle Database support:

"Oracle has a support statement for VMware products that is honored around the world. While there has been much public discussion about Oracle's perceived position on support for VMware virtualization, our experience is that Oracle Support upholds its commitment to customers, including those using VMware virtualization in conjunction with Oracle products.

The following are some of the key facts about Oracle Support:

- *Oracle RAC support is now included for Database 11.2.0.2 and later.*
- *Known issues – Oracle Support will accept customer support requests for Oracle products running on VMware virtual infrastructure if the reported problem is already known to Oracle. This is crucial—if you are running Database 9i, 10g, or another product with a long history, the odds are in your favor that Oracle has seen your problem before. If they have already seen it, they will accept it.*
- *New issues – Oracle Support reserves the right to ask customers to prove that "new issues" attributed to Oracle are not a result of an application being virtualized. This is reasonable, as this is essentially the same policy that other ISVs use to some degree. It is key to look at the history of Oracle Support with regard to new issues.*
- *Certification – VMware vSphere is a technology that resides under the certified Oracle stack (unlike other virtualization technologies that alter the OS and other elements of the stack). As a result, Oracle cannot certify VMware virtual infrastructure. However, VMware is no different in this regard from an x86 server—Oracle doesn't certify Dell, HP, IBM, or Sun x86 servers.*

VMware recommends that customers take a logical approach and test Oracle's support statement. Begin with pre-production systems, and as issues are encountered and SRs are filed, track Oracle's response. Our experience is that customers see no difference in the quality and timeliness of Oracle Support's response."

2.2 Hyper-V Support

Microsoft Hyper-V 2012 onwards fully supports running Oracle Database VMs, *Ref.2*. Zerto supports a minimum of Hyper-V Server 2012 R2 onwards and so Oracle Database VMs on this hypervisor version onwards can be protected by Zerto Virtual Replication.

2.3 AWS Support

AWS supports running Oracle Database VMs, *Ref.3*. As Zerto support replicating to AWS it is supported to protect Oracle Database VMs to AWS for recovery as EC2 instances.

2.4 Zerto Support

Zerto supports protecting VMs running any operating system, only automatic IP-reconfiguration utilizes a specific list of supported Operating Systems. Therefore Oracle databases running on either Linux or Windows as the Operating System can be protected using Zerto Virtual Replication.

As Zerto replicates from within the hypervisor, it resides under the certified Oracle stack and so Oracle cannot certify Zerto just as it cannot certify VMware. Zerto support does not cover the recovery of any application inside a VM or its Operating System (OS), only the recovery of the VM itself. For application or OS specific support the appropriate support vendor should be used.

3 ORACLE CONFIGURATION

3.1 Automatic Storage Management (ASM)

ASM provides a clustered file system with volume management for Oracle databases. Disk groups are limited by the performance of the slowest member and so it is recommended to place all protected virtual disks used in a disk group on storage with a similar I/O profile.

Oracle best practice dictates that two ASM disk groups should be created, one with a sequential I/O profile for log files, and one with a random I/O profile for datafiles. By selecting an Oracle database VM for protection Zerto will automatically protect all virtual disks assigned to the VM, therefore all disk groups will be included in the protection. As Zerto maintains write-order fidelity of all the disks within a VM, and between VMs in the same Virtual Protection Group, protecting a VM with multiple disk groups will recover all the disk groups to the exact same point in time. The journal will never present a checkpoint available for selection where the disk groups are at a different point in time.

In a VMware environment it is not recommend to use ASM failure groups, as per the Oracle Databases on VMware Best practices guide:

"Do not use Oracle ASM failure groups. Oracle failure groups consume additional CPU cycles and can operate unpredictably after suffering a disk failure. When using external redundancy, disk failures are transparent to the database and consume no additional database CPU cycles, because this is offloaded to the storage processors."

3.2 Oracle Clustered File System (OCFS)

The OCFS is a shared disk cluster technology for Linux used with Oracle Real Application Clusters and was used prior to ASM released in Oracle 10g. Zerto does not support protecting VMs utilizing shared disk cluster technologies for Oracle databases and therefore does not support OCFS. It is recommended to use ASM as the clustering technology.

3.3 Oracle Data Guard

Oracle Data Guard is an application based database protection solution for BC/DR that requires a target VM to be permanently running in order to provide protection of Oracle databases. It requires a separate license and it is not required when protecting Oracle databases with Zerto. Zerto performs the same continuous protection as that provided by Oracle Data Guard without the complexity of managing an additional solution that often requires manual intervention to maintain replication. The other key differentiators of Zerto vs Oracle Data Guard are:

- Zerto does not have any VMs running in the target site for the replica VM, reducing complexity
- Zerto only requires the Oracle database VM to be protected in a VPG and is much simpler to configure than Oracle Data Guard
- Zerto allows failover testing to an isolated network without breaking the replication or shutting down production, the only way to test failover with Data Guard is to failover

- Zerto failover testing allows applications that utilize the Oracle databases access in the failover network, applications cannot utilize Oracle databases in a failover test with Data Guard
- Failover orchestration is combined with all of the protected VMs rather than a different solution for Oracle
- Only one skillset is needed to manage replication and perform a recovery rather than multiple skillsets if using Oracle Data Guard
- Zerto enables advanced features around replication such as bandwidth throttling, real-time KPIs, SLAs, prioritization of replication traffic all of which are missing from Data Guard
- Recovery to previous points in time achieved by selecting a point in time in the Zerto GUI, no complex database operations are required

If desired an Oracle database can be protected by both Oracle Data Guard and Zerto simultaneously, as Oracle Data Guard runs inside the OS and Zerto runs in the hypervisor there is no conflict between the two solutions.

3.1 Oracle Real Application Clusters (RAC)

Oracle RAC utilizes shared virtual disks for CRS, voting, data and redo logs between 4 VM nodes. Zerto does not support protecting VMs utilizing shared disks, therefore it is not possible to protect Oracle RAC nodes using Zerto Virtual Replication.

4 ORACLE LICENSING

4.1 Licensing Overview

All advice given on the topic of licensing Oracle databases serve solely as recommendations and do not represent any legally binding certification of compliance for Oracle licensing. For further clarification it is recommended to contact your Oracle licensing manager or utilize a 3rd party Oracle licensing specialist.

The basic tenant of an Oracle Licensing Service Agreement (OLSA) is that Oracle databases need to be licensed where installed and/or running, not where it can run. This phrasing can cause ambiguity when configuring virtualization and disaster recovery for virtualized Oracle databases.

The storage on which an Oracle database VM runs, or the hypervisor hosts that have access has no bearing on the OLSA and so shared storage configurations are not a factor needs to be configured in the OLSA.

4.2 Production Site Licensing

Oracle is typically licensed per physical processor and it is required to license all of the processors on the hypervisor host on which the Oracle database VM was installed. The Oracle VM should be kept running on this host for this reason. It is recommended to license all of the processors in a second hypervisor host to allow High Availability failover, vMotion or Live Migration between hypervisor hosts without breaching Oracle licensing terms.

If HA and DRS rules are utilized to keep the Oracle databases VMs on the two hypervisor hosts that have Oracle licensing, then the hypervisor hosts can form part of a larger cluster. It is not a requirement to maintain a separate Oracle hypervisor cluster, or license all of the hosts in a cluster, only ensure the hosts on which the Oracle database VMs run are licensed.

It is technically possible to only license the processors in one hypervisor host, but the Oracle VMs would need to be shutdown when performing host maintenance and would not have any high availability capabilities to avoid breaching licensing terms and therefore this is not recommended.

4.3 Recovery Site Licensing

Zerto maintains a replica of the VM data and a journal of the data being changed within the time frame specified in the recovery site. No VM is registered or powered-on in the inventory until a Failover, Move or Test operation is initiated. From a licensing perspective a Zerto replica VM is a cold offline backup, but this does not mean the recovery site hypervisor host should not be licensed.

If a failover test is performed then the Oracle database VM is running in the recovery site host and by the basic tenant of the OLSA this then requires the processors to be licensed. As regular failover testing is recommended to ensure a successful recovery from a disaster it is therefore recommended to license one recovery site hypervisor host. Zerto replication should be configured to only replicate the Oracle VMs to the licensed hypervisor host. It is possible to change the target host in the ZVM to continue replication during hypervisor host maintenance without invalidating Oracle licensing, but a Failover, Move or Test operation should not be performed until the replication target has been changed back to the Oracle licensed hypervisor host.

Zerto failover test VMs cannot be vMotion between hosts in a recovery site and therefore it is not required to license all of the hosts in the recovery site hypervisor cluster, neither is it required to maintain a separate cluster for Oracle VMs in a recovery scenario.

If a Failover or Move operation is performed then HA and DRS rules should be applied to the failover VM to ensure the VM only runs on licensed hosts in the recovery site. This can be done manually or as part of a post-failover script on the VPG.

If an Oracle database VM is never failed over, migrated or run in a test failover operation then the recovery site hypervisor does not need to be licensed. However, due to all of the reasons above Zerto recommends to license one hypervisor host in the recovery site.

5 ZERTO BEST PRACTICES

5.1 Performance Impact of Replication

Zerto replicates only the block level changes on a per-VM basis and is asynchronous throughout. Only a copy of the writes are sent asynchronously to the local VRA in the memory of the hypervisor host. There is no delay on write to and from the local storage as a copy is being sent to the VRA and therefore no measurable performance impact on protecting an Oracle VM with Zerto Virtual Replication.

5.2 Oracle Temporary Database Optimization

Zerto utilizes a SWAP disk feature to optimize the replication traffic of Oracle temporary databases and it is enabled on a per-virtual disk basis. When enabled an initial copy of the virtual disk is replicated to the recovery site, but no subsequent changes to the disk are replicated. This ensures that any application that uses the virtual disk for transient data, I.E an Oracle temporary database, will see the disk and original files in a recovery scenario without requiring any manual reconfiguration. The SWAP disk feature typically saves 50%+ of replication traffic, journal space usage and VRA load.

It is therefore recommended to utilize the Zerto SWAP disk feature on all Oracle temporary databases that should be kept on a separate virtual disk to allow its use without impacting the replication of non-transient data. It is possible to turn the SWAP disk feature on and off by simply editing the VPG settings and so the impact of this feature on bandwidth usage and recovery operations can be tested easily. Performing a force-sync on a VPG will update the contents of any SWAP disk enabled virtual disk in the recovery site.

5.3 WAN Sizing

All Oracle database VMs should follow standard Zerto sizing recommendations by using the Zerto WAN sizing tool, available on the self-service portal, to capture the write rate on all the disks in the Oracle database VM to calculate the required bandwidth between the sites.

If backups are made on the Oracle database VM to a local disk it is recommended to have the backups stored on a separate virtual disk indicated as a SWAP disk during Zerto VPG configuration, just as recommended for temporary databases.

Failure to configure this can result in Oracle backups causing nightly bitmap-syncs as the VM is writing a copy of the entire database to a backup file which is being replicated. In this scenario if the data set is too large for the WAN link or for the VRA to keep up then the VM will enter a bitmap-sync mode.

5.4 Database Consistency in Windows

Zerto replicates continuously maintaining the write-order fidelity of all the block level changes between all of the VMs in a VPG and all of the disks within each VM which means that the replication is always crash consistent. To configure database consistency in Oracle VMs running Windows it is recommended to install and configure the Zerto PowerShell CMDlets from the self-service portal.

A PowerShell script can then be scheduled, *Ref.4*, to create database consistent points in time (checkpoints) with the following process:

1. Windows Task Scheduler starts PowerShell script
2. PowerShell script loads user configured variables
3. Connects to the database using SQLPLUS
4. Places the database into backup mode
5. Inserts a checkpoint into the journal for the Oracle VPG using the Zerto PowerShell cmdlet indicating the database consistent point in time
6. Resumes normal database operations by ending backup mode

Following is an example PowerShell script for creating database consistent points in time in Windows:

```
# Start of script - Configure the variables below
$OracleUser = "system"
$OraclePassword = "password"
$OracleServer = "OracleServer"
$ZVMServerIP = "192.168.0.30"
$ZVMPort = "9080"
$ZVMUser = "administrator"
$ZVMPassword = "password"
$VPGName = "Oracle"
$CheckpointTag = "HotBackup"
# Nothing to configure below here, loading Zerto PowerShell Commands
add-pssnapin "Zerto.PS.Commands"
# Connecting to Oracle and placing the database in hot backup mode
$OracleConnectionString = $OracleUser + "/" + $OraclePassword + "@" + $OracleServer
sqlplus $OracleConnectionString
ALTER DATABASE BEGIN BACKUP;
EXIT
# Inserting a checkpoint in the Zerto journal for the defined VPG
Set-Checkpoint -VirtualProtectionGroup $VPGName -Tag $CheckpointTag -ZVMIP $ZVMServerIP -ZVMPort $ZVMPort -Username $ZVMUser -Password
$ZVMPassword
# Connecting to Oracle and resuming normal database operations
sqlplus $OracleConnectionString
ALTER DATABASE END BACKUP;
EXIT
# End of script
```

The frequency on which database consistent points in time can be inserted is subject to the performance of the Oracle databases and services, not Zerto. In theory Zerto could support database consistent checkpoints being inserted every 10 seconds, as Zerto simply inserts a point in time marker and the database consistency is handled by the Oracle hot backup mode. It is recommend to start with daily database consistent points in time and build up to reach the tolerance of the databases and services.

Upon recovering a Windows Oracle database VM from a database consistent checkpoint it is required to release the database from backup mode to continue normal operation as it will not start automatically on boot. This can be done with the following example command, *Ref.5*:

```
sqlplus / as sysdba
startup mount;
alter database end backup;
alter database open;
```

5.5 Database Consistency in Linux

Zerto does not provide an agent for managing database consistency in Linux VMs. In order to insert database consistent points in time in Linux VMs it is recommended to use the Zerto PowerShell CMDlets from the self-service portal, installed on a Zerto Virtual Manager Windows server VM, with a PowerShell SSH server to allow remote execution from Linux VMs.

The process for executing a database consistent point in time using this technique is as follows:

1. Oracle Linux VM runs a bash script that enters the database into backup mode
2. Bash script connects via SSH to ZVM PowerShell server and executes a PowerShell script
3. PowerShell script on the ZVM inserts a checkpoint into the Zerto journal indicating the database consistent point in time
4. Bash script disconnects SSH session
5. Bash script resumes normal database operations by ending backup mode

Following is an example Linux script for performing this operation:

```
#!/bin/bash
# Placing database in backup mode
sqlplus "/as sysdba" <<EOF
alter database begin backup;
exit;
EOF
# Running remote PowerShell script on the PowerShell server via SSH
ssh administrator@192.168.0.115 '.\oraclevpg.ps1' exit
ENDSSH
# Resuming database operations
sqlplus "/as sysdba" <<EOF
alter database end backup;
exit;
EOF
# End of script
```

To enable remote SSH access to PowerShell CMDlets on the ZVM it is recommend to use PowerShell server, *Ref.6*:

http://www.powershellserver.com/download/

Following is an example PowerShell script to be run on the ZVM, initiated from the Linux VM via SSH:

```
#
# Start of script - Configure the variables below
#
$ZVMServerIP = "192.168.0.30"
$ZVMPort = "9080"
$ZVMUser = "administrator"
$ZVMPassword = "password"
$VPGName = "Oracle"
$CheckpointTag = "HotBackup"
#
# Inserting the checkpoint into the journal
#
Set-Checkpoint -VirtualProtectionGroup $VPGName -Tag $CheckpointTag -ZVMIP $ZVMServerIP -ZVMPort $ZVMPort -Username $ZVMUser -Password
$ZVMPassword
#
# End of script
#
```

Upon recovering a Linux Oracle database VM from a database consistent checkpoint it is required to release the database from backup mode to continue normal operation as it will not start automatically on boot. This can be done with the following example command:

```
sqlplus / as sysdba
startup mount;
alter database end backup;
alter database open;
```

5.6 VRA Configuration

It is recommended to install a VRA on each source and target hypervisor node that the protected Oracle VM can run on. Failure to deploy a VRA on each hypervisor will result in a break in replication if the Oracle VM is moved to a hypervisor without a VRA deployed.

If multiple Oracle VMs that are all highly transactional in their I/O profile require protection it is recommended to distribute the load evenly between hypervisor hosts and therefore VRAs to minimize the I/O load placed on a single appliance.

5.7 VPG Design

VPG design in protecting Oracle database VMs is dependent on a number of factors. When configuring VPGs it is important to understand the following rules:

1. A VM can only be protected by 1 VPG at once
2. A VPG can contain VMs across any number of clusters, hosts and storage
3. Recovery, move, backup and test operations are performed on per VPG basis
4. It is not possible to recover only 1 VM out of many in a VPG without using an offsite-clone as a workaround
5. All VMs and virtual disks in a VPG are recovered to the exact same timestamp
6. Recovering multiple Oracle database VMs to the same database consistent point in time in the same VPG requires all databases to pause simultaneously and only 1 command to insert a checkpoint can be run
7. Boot ordering is configurable on VMs within a VPG, not between VPGs

Given the above rules, the design of VPGs is dependent on a number of factors including the granularity of recovery required, the interdependencies between databases and the applications using the databases.

General best practice dictates to place all of the VMs that form a single application in the same VPG. If multiple applications use the same database and all require consistency then 1 large VPG should be created. If the application VMs don’t require consistency to the database VM then separate VPGs for the applications and the database VM should be created. If multiple database VMs update each other continuously and are interdependent then they should be placed in the same VPG for cross-database consistency.

Once a VPG design has been implemented it is possible to change the VPG of a protected VM by removing the VM from the VPG, selecting to keep the target disks then add the VM to the new VPG with the pre-seed option to minimize the disruption in protection.

Boot ordering within the VPG should be applied to ensure the Oracle Database VMs boot with a sufficient time period to bring database services online before any application servers that use the database services.

5.8 VPG SLAs

The VPG SLAs should be configured to allow a sufficient amount of data retention for recovery to previous points in time, give sufficient leeway on the RPO to allow spikes in data change rates without raising false alerts and finally alert on the frequency of testing on a reasonable schedule. For best practice the following VPG SLA recommendations should be observed:

Journal SLA

Start with a 1 hour journal and after 1 hour of production data has been replicated use the size of the journal to calculate the space required over the time period the business dictates as an SLA requirement for data retention. The journal should then be increased to the desired time frame for retention. Further journaling recommendations are covered in the journal section of this document.

RPO Alert

Start with a 5 minute RPO and leave the VPG protected for at least 1 working day in this configuration. Configure email alerts to be notified immediately of any RPO SLA breaches during this time period. After 24 hours the RPO should be altered up or down depending on the SLA requirement of the business which can be dictated by the RPO observed during this period. I.E if the RPO regularly hits 5 minutes but typically lasts for 10 minutes before returning to normal then the RPO alert and SLA should be changed to 15 minutes or the bandwidth between the sites should be increased to reduce the occurrence of the bitmap-syncs.

It is not recommended to change the RPO alert to less than 20 seconds as this leaves no leeway for any spikes in data change or network connectivity and can result in unnecessary RPO alerts being generated. Neither should the RPO alert be configured to the exact SLA defined by the business as this gives no leeway on the replication alert vs SLA defined. I.E if the business requires an RPO of 30 minutes then best practice would be to configure the RPO alert at 20 minutes to allow a comfortable leeway. If the RPO alert was constantly breached in this scenario then the situation can be remedied by increasing available bandwidth without breaching the business SLAs.

Test Frequency Alert

The default test frequency of 6 months is recommended when first configuring the protection of a VPG. It should be increased or decreased based on the frequency of testing required by the business defined SLAs.

5.9 Journaling

When configuring journaling for Oracle Database VMs it is recommended to appropriately size the amount of data change before setting the desired journal history length. Once the data change rate of 1 hour of production use has been calculated then a datastore, of the same I/O profile as the datastore containing the replica VM, data should be used to store the journal. The datastore should have enough space to accommodate the data change rate over the time period specified. To cope with spikes in data change there should be at least 30% of disk space free on the datastore and the maximum journal size should be configured on each journal to ensure no journal can fill the datastore specified.

Performance of the journal datastore is important as the protected VM data is inserted with the same I/O profile and the journal can be heavily used during the pre-commit and commit phase of a failover, before committing the failover.

Maintaining space availability on the journal datastore is also key to maintaining replication. Filling the datastore will break replication, whereas multiple journals hitting their size limit will reduce the time frame of the journal, it will not break the replication. This is why it is recommended to set maximum journal size limits for Oracle Database VMs given their often critical part of a disaster recovery plan.

5.10 Datastore Configuration

It is recommended to use a separate datastore for storing the journal and the replica VM data. The datastore configuration of the protected Oracle database VM should be matched in the recovery site to ensure the same performance in a recovery scenario. As Zerto is storage agnostic it is possible to replicate from any to any storage and therefore any to any LUN configuration, but this does not take into context the performance requirements of the protected VMs.

The number of VMs running on a datastore as recommended by the storage vendor should also be taken into to consideration if replicating multiple VMs to a consolidated number of datastores in a target site.

5.11 Networking Configuration

The networking configuration of the Oracle Database VMs running in production should be mirrored in the recovery site. If a different IP address is required in the recovery site the VPG can be configured on a per virtual NIC basis to automatically change the IP address as part of a failover, move and test operation. If an IP change is required then ensure the listener.ora and tnsnames.ora configuration files do not have IP addresses hard coded rather than using hostnames.

For failover testing a mirror of the production networking should be configured in an isolated network segment to allow failover testing with no impact on production.

5.12 Failover Testing

Failover testing is recommended to validate the recovery process, the recovery time objective, data consistency and operation of the Oracle Database VM in a recovery scenario. Failover testing of Oracle Database VMs can be done in working hours as there is no shutdown in production or break in the replication when performing a failover test operation. Ensure isolated networks are specified for failover testing to ensure the Oracle Database VMs cannot communicate with production applications and clients, otherwise the failover test can have a production impact as the failover test VM could receive database updates from production that will be deleted when the failover test operation is stopped.

The size of a failover test operation for Oracle Database VMs is simply the size of the changes written to the test VMs during the test. All changes from the failover test VMs are deleted when the test is stopped and it is not recommend to run a failover test for longer than the journal history is configured. If a longer term test is required then the offsite-clone feature should be utilized to create an entire VM-level clone of the VMs in the VPG onto any datastore from any specified point in time available in the journal.

If recovering to a database consistent point in time the failover test operation allows validation of the Oracle Database starting in backup mode to verify the consistency and then subsequent validation of removing the database from backup mode and mounting the database to ensure a successful recovery. It is recommended to perform test failovers to both crash and database consistent points in time to verify the recovery of both on a per Oracle database basis. The ability to recover from either should be recorded and included in a disaster recovery plan for future decision making in a failover scenario when deciding what point in time to recover to and the amount of data loss to accept.

If IP address changes are required as part of a failover or move operation then they should also be performed in the failover test operation to ensure a successful recovery. This allows the IP change process to be verified and the full disaster recovery scenario to be simulated.

5.13 Failover, Move & Failback

In order to ensure a successful failover, move and failback operation the commit policy feature should always be utilized. A default commit policy of manual is recommended to allow an indefinite time period to verify the successful recovery of Oracle Database VMs before deleting the journal of change. After a successful recovery has been confirmed the VPG should be committed to the point in time selected. If the recovery is un-successful then a different point in time should be selected from the journal for the next recovery.

Before performing a move operation for Oracle Database VMs it is recommended to take a database consistent checkpoint using the provided example scripts found in this document. This ensures that there is a guaranteed point in time in the journal if the database services do not shut down in a clean state.

In a failover operation the option exists to failover to the most recent point in time to minimize data loss or to failover to the last database consistent point in time and accept the data loss for the time interval between. No recommendation is given as to which should be selected as this is both a user and business decision. Using the commit policy feature gives an option of trying the most recent crash consistent point in time and if this does not work then the ability to failover to the last database consistent point in time.

It is recommended to ensure sufficient leeway is built into the RTO SLA defined to the business to allow for the VMs to be powered on, database services started and the recovery verified as successful before re-commencing business operations. I.E if the VM and VM service RTO takes 15 minutes, a recommended RTO SLA to the business would be 45 minutes to allow verification and subsequent recovery to a previous point in time in the event of the first point in time not recovering in a working state.

A sufficient time period should be allowed when configuring failback to allow for the delta-sync process to read both source and target disks then replicate the changes. The speed of a delta-sync operation in Oracle Database VMs is dependent on the following factors:

1. Amount of data changed since the failover operation
2. Speed of the IP link and available bandwidth for replicating the change
3. Speed of both the source and target storage
4. Size of the virtual disks and amount of free space
5. The latency to the source and target storage

The delta-sync process does not run as fast as the VRAs can run, it runs as fast as possible without impacting the performance of the Oracle Database VM. This is done by dynamically scaling up and down the read rate to the protected VM disks depending on the latency to the storage to ensure no performance impact of the delta-sync process. This ensures the performance of Oracle Database VMs in a recovery scenario is the highest priority with subsequent failback to production a background task that can be left to run until the VPGs are in a protected status ready for failback.

A failback operation is typically a planned migration back to production and so it is recommended to perform this out of working hours as it involves shutting down the VMs now running in the disaster recovery site. The same process followed for migration should be adhered to by taking a database consistent point in time before initiating the failback procedure. The failback should also be tested by a test failover process to validate the process and recovery time objective first.

6 VMWARE BEST PRACTICES

6.1 VMware Recommendation Summary

Recommendations
Create a computing environment optimized for vSphere
Create golden images of optimized operating systems using vSphere cloning technologies
Upgrade to ESX 4 (minimum)
Allow vSphere to choose the best virtual machine monitor based on the CPU and guest operating system combination
Set memory reservations equal to the size of the Oracle SGA
Use large memory pages
Use as few virtual CPUs (vCPUs) as possible
Enable hyperthreading for Intel Core i7 processors
Enable jumbo frames for IP-based storage using iSCSI and NFS
Create dedicated datastores to service database workloads
Use VMware vSphere VMFS for single instance Oracle database deployments
Align VMFS properly
Use Oracle automatic storage management
Use your storage vendor’s best practices documentation when laying out the Oracle database
Avoid silos when designing the storage architecture
Use Paravirtualized SCSI adapters for Oracle data files with demanding workloads
Use the VMXNET family of Paravirtualized network adapters
Separate infrastructure traffic from virtual machine traffic for security and isolation
Use NIC teaming for availability and load balancing
Take advantage of Network I/O Control to converge network and storage traffic onto 10GbE
Use vCenter and/or the esxtop/resxtop utility for performance monitoring in the virtual environment
To minimize time drift in virtual machines follow guidelines in relevant VMware Knowledge Base articles

7 REFERENCES

Ref.1 Oracle and VMware best practices
https://www.vmware.com/files/pdf/partners/oracle/Oracle_Databases_on_VMware_-_Best_Practices_Guide.pdf

Ref.2 Oracle Hyper-V Support
https://blogs.oracle.com/cloud/entry/oracle_and_microsoft_join_forces

Ref.3 Oracle AWS Support
http://www.oracle.com/technetwork/topics/cloud/faq-098970.html#support

Ref.4 Scheduling a PowerShell script
http://blogs.technet.com/b/heyscriptingguy/archive/2012/08/11/weekend-scripter-use-the-windows-task-scheduler-to-run-a-windows-powershell-script.aspx

Ref.5 Recovery from database consistent points in time
http://www.n2ws.com/blog/ec2_consistent_oracle_backup.html

Ref.6 Getting Started with PowerShell Server
http://www.powershellserver.com/getting-started/