Using EMC VNX Unified Storage Replication Consistency Technology in Oracle 11gR2 and ASM Deployments

Validation Test Report

EMC
Corporate Headquarters
Hopkinton, MA 01748-9103
1-508-435-1000
www.EMC.com

Published February 2011

Using EMC VNX Unified Storage Replication Consistency Technology in Oracle 11gR2 and ASM Deployments Validation Test Report

Part number h8169

Contents

Introduction

Customers need solutions that support different connectivity requirements and provide functionality that increases efficiency, such as unified Fully Automated Storage Tiering (FAST), unified replication, and application protection (security, rollback recovery, and compliance**).** EMC offers centralized storage, combining NAS and SAN under a single point of management. The VNX™ series is a member of the VNX family and is EMC's next-generation midtier unified storage platform, combining all the benefits of Celerra® and CLARiiON® into a single product line.

The following table shows the different VNX series products and what current models they replace.

Table 1. CLARiiON/Celerra transition to VNX series

Class	Model name	Attributes	Current product
7000	VNX7500	• Maximum performance • Maximum scale • Maximum connectivity • FAST enabled	CLARiiON CX4-960 and Celerra NS-960
5000	VNX5700 VNX5500 VNX5300 VNX5100™	• Balanced performance • High capacity and scaling • FAST enabled • Nondisruptive upgrades to a higher platform	• CLARiiON CX4-480 and Celerra NS-480 • CLARiiON CX4-240 • CLARiiON CX4-120 and Celerra NS-120 • CLARiiON AX4 (Fibre Channel)

The following figure shows the new generation of EMC® unified storage.

Figure 1. EMC VNX unified storage system product family

EMC VNX series is designed for midtier-to-enterprise storage environments that require advanced features, flexibility, and configurability. Users who are dealing with factors such as exponential data growth, virtualization, and an IT industry shift from a storage focus to a business focus demand storage solutions that are simple, efficient, and powerful, which are all characteristics provided by the EMC VNX series. How are the three critical requirements covered? The following points provide a brief explanation.

Simple

- **Centralized management** with EMC Unisphere™ for file, block, and object; and with software services
- **Optimized for virtual applications** with VMware and Hyper-V integration

Efficient

- **Efficient data services** with Block Compression, File Deduplication and Compression, Virtual Provisioning™, and application-centric protection
- **Automated tiering** with FAST for Virtual Pools (FAST VP), which automates and optimizes data on Flash, SAS, and near-line (NL)-SAS drives.

Powerful

- **Flash optimized**, allowing customers to use Flash drives as extendable cache and in storage pools

VNX unified storage is intuitive, greatly reducing the time needed to manage and tune storage**.** Unisphere is very easy to learn, and FAST VP automates the performance analysis and management of volumes. EMC's automation of performance and efficiency optimization allows customers to meet service-level agreements with little interaction and stress. Customers can truly consolidate many applications onto the same EMC storage, define service-level agreements for each, and let the VNX series' "set-it-and-forget-it" self-optimization do the rest. Customers achieve a lower total cost of ownership with EMC's "Flash 1st" strategy. FAST Cache keeps the hottest data at the Flash level, and FAST VP brings efficiency and keeps unused data on the cheaper/slower tiers. On a whole, the VNX series gives customers capacity savings of up to 50 percent by leveraging features such as File Deduplication, Block Compression, and thin provisioning.

As in previous generations of VNX (then known as CLARiiON), the SnapView™, MirrorView™/Synchronous (MV/S), and MirrorView/Asynchronous (MV/A) storage-based replication options are available. To ensure compatibility between VNX's latest replication technologies and Oracle's latest offerings (Oracle Database 11gR2), extensive testing was done in EMC's Engineering labs. Replication testing was done on an EMC VNX5100 model, but the validation extends to all models in the VNX family that support SnapView and MirrorView. The operating system on the server was Oracle Enterprise Linux 5 (OEL) update 4; however, this same methodology is valid for other versions of Linux that are in alignment with the *EMC Support Matrix* (ESM).

During replication, data content consistency across multiple LUNs comprising a database must be maintained. As such, testing focused on using the consistency feature available with SnapView, MirrorView/S, and MirrorView/A in conjunction with Oracle 11gR2's flashback database and restore point features to create a valid point-in-time backup of an Oracle database that spans multiple LUNs, without having to put the database into hot backup mode during replication.

EMC VNX5100 test environment (Linux)

Host machines

Two Dell 2950 servers were used, one for production and one for the backup host, with the following components:

Operating system	OEL5 U4
	Kernel 2.6.18-92.el5xen
Memory	33 GB
CPU	8 x dual-core Intel Xeon CPU, X5355 @ 2.66 GHz

Naviacli	navicli-6.28.10.3.11-1
Naviagent	naviagent-6.28.10.3.11-1
EMC PowerPath®	EMCpower.LINUX-5.3.1.00.000-111

Storage systems

One VNX5100 storage array was used for the primary system with the following components:

Processors	2
Memory size	4 GB per SP
Number of disks	75 x 268 GB @ 15k rpm SAS drives
Base software	VNX OE 05.31.000.4.012

One VNX5100 storage array (MirrorView only) was used for the secondary system with the following components:

Processors	2
Memory size	4 GB per SP
Number of disks	75 x 268 GB @ 15k rpm SAS drives
Base software	VNX OE 05.31.000.4.012

Oracle software

Oracle Database version	Oracle 11*g* Release 2 (11gR201)
ASM	oracleasm-2.6.18-92.el5-2.0.4-1.el5

Storage system configuration

Six 4+1 RAID 5 LUNs were created on the EMC VNX unified storage array and configured to be managed by Automatic Storage Management (ASM), Oracle's file system and volume manager. These six production LUNs were distributed between four disk groups (using external redundancy) as follows:

- **DATA_DGRP:** Consists of two LUNs for all database files and control files
- **FLASH_DGRP:** Consists of one LUN for the Fast Recovery Area (Flash Recovery Area is referred to as Fast Recovery Area from 11gR2 onward)
- **ARCHIVE_DGRP**: Consists of one LUN for the archive logs
- **REDO_DGRP**: Consists of two LUNs for the online redo logs

Snapshots, clones, and remote mirrored pairs were created on the appropriate primary and secondary storage arrays. For MirrorView replication, two consistency groups were created out of the six remote mirrored pairs. One consistency group consists of just the mirrored pair for the archive logs, and the other consistency group consists of the remaining five mirrored pairs.

Test scenarios

An online backup test that simulates bank transactions of about 400,000 rows being updated over a period of time was used for this purpose. For each test run, updates were ongoing when the replicated copy was split from its source. To further stress-test the system, ASM disk group rebalancing was in progress with ongoing updates when the replicated copy was split from its source. The two scenarios tested for each replication technology were:

- Storage consistency, flashback on, database not in hot backup mode, no disk group rebalance
- Storage consistency, flashback on, database not in hot backup mode, with disk group rebalance

To ensure that there was no timing issue, each test scenario was repeated a minimum of five times.

Test details

To enable flashback as well as archiving to a location outside the Fast Recovery Area, the following was added to the init*.ora parameter files of the production and backup instance:

DB_RECOVERY_FILE_DEST_SIZE = *100G*
DB_RECOVERY_FILE_DEST = '+*FLASH_DGRP'*
LOG_ARCHIVE_DEST_1 = 'LOCATION=+*ARCHIVE_DGRP'*

Archive logging and flashback were then enabled using the following SQL*Plus commands:

ALTER DATABASE ARCHIVELOG;
ALTER DATABASE FLASHBACK ON;

It is necessary to archive to a separate location outside the Fast Recovery Area so that the disk group containing the archive logs could be split after the backup was taken. This archived log can then be used to roll a backup forward in time. The following sections detail steps executed on the production host to capture a point-in-time copy of the database and on the secondary host to recover and roll the database forward.

Production host

1. While the production database was actively being updated but before taking a snapshot or fracturing the replicated LUNs, create a restore point using the following SQL*Plus command:

   CREATE RESTORE POINT *bkup1* GUARANTEE FLASHBACK DATABASE;

   A "restore point" is a name (example, *bkup1*) associated with a committed System Change Number (SCN) of the database corresponding to the time the restore point was created. This restore point can later be used to flash the database back to the time of creation of the restore point without the need to determine the SCN.

   For tests involving ASM disk group rebalancing, start the rebalance process right after creating the restore point. The following SQL*Plus commands initiate the rebalance process:

   ALTER DISKGROUP *DATA_DGRP* REBALANCE POWER 2 NOWAIT;
   ALTER DISKGROUP *REDO_DGRP* REBALANCE POWER 2 NOWAIT;
   ALTER DISKGROUP *FLASH_DGRP* REBALANCE POWER 2 NOWAIT;
   ALTER DISKGROUP *ARCHIVE_DGRP* REBALANCE POWER 2 NOWAIT;

2. Once a restore point has been captured, start a snapshot session or fracture of the replicated LUNs. The LUN containing the archive logs is not part of this step. Clones and mirrors were verified to be in a consistent or synchronized state before the fracture. The following host-based Navisphere® Secure CLI (naviseccli) command creates a replica of the production LUNs for the different replication options.

   #### SnapView snapshot to start a consistent snapshot session

   naviseccli -User *admin_user* -Password *admin_passwd* -scope 0 -h *SP_name* snapview -startsession *session_name* -snapshotname *lun1,lun2,lun3,lunN* -persistence -consistent

SnapView clone to start a consistent clone fracture

naviseccli -User *admin_user* -Password *admin_passwd* -scope 0 -h *SP_name* snapview -consistentfractureclones –CloneGroupNameCloneId *name1 cloneid1 name2 cloneid2 ... nameN cloneidN*

MirrorView/Synchronous to start a consistent group fracture

naviseccli -User *admin_user* -Password *admin_passwd* -scope 0 -h *SP_name* mirror -sync –fracturegroup –name *group_name*

MirrorView/Asynchronous to start a consistent group fracture

naviseccli -User *admin_user* -Password *admin_passwd* -scope 0 -h *SP_name* mirror -async –fracturegroup –name *group_name*

At no point was the production database put in hot backup mode before it was snapped or fractured.

3. After the production database has been snapped or fractured, perform a log switch and capture the highest SCN to be used for recovery:

   ALTER SYSTEM ARCHIVE LOG CURRENT;

   SELECT 'NextChange', next_change# from v$log_history where recid=(select max(recid) from v$log_history);

4. After the log switch completes, the LUN containing the archive log is then snapped or fractured depending on the replication option selected. The same Navisphere Secure CLI commands were used but with just the archive log LUN.

Backup host

1. In order for a backup host to access the replicated database, the snapshots, fractured clones, or secondary images must be assigned to a storage group that is connected to the backup host. Use the following method to enable a backup host to access the replicated database:

SnapView snapshots

- Add the snapshots to a storage group that is connected to the backup host.
- Issue applicable operating system commands to scan for new devices, making them visible to the backup host.
- Activate the snapshot to enable the backup host to access the newly visible devices:

  admsnap activate –s *session_name*

SnapView clones

- Add the clones to a storage group that is connected to the backup host.
- Issue applicable operating system commands to scan for new devices, making them visible and accessible to the backup host.

MirrorView (MV/S and MV/A) secondary images

Using mirror promote:

- With this VNX OE release please use the NaviCLI commands to create mirrors, secondary images, and consistency groups. In later releases of VNX OE, the Unisphere GUI will be enabled to support these actions.
- Perform a local consistency group promote of the secondary images that are on the secondary array.

MirrorView/Synchronous:

naviseccli -User *admin_user* -Password *admin_passwd* -scope 0 -h *SP_name* mirror -sync –promotegroup –name *group_name* –type local

MirrorView/Asynchronous:

naviseccli -User *admin_user* -Password *admin_passwd* -scope 0 -h *SP_name* mirror -async –promotegroup –name *group_name* –type local

- Add the promoted LUNs to a storage group that is connected to the backup host.
- Issue applicable operating system commands to scan for new devices, making them visible and accessible to the backup host.

Using SnapView snapshots:

- Start a consistent snap session.
- Add the snapshots to a storage group that is connected to the backup host.
- Issue applicable operating system commands to scan for new devices, making them visible to the backup host.
- Activate the snapshot to enable the backup host to access the newly visible devices.

Using SnapView clones:

- Start a consistent clone fracture.
- Add the clones to a storage group that is connected to the backup host.
- Issue applicable operating system commands to scan for new devices, making them visible and accessible to the backup host.

2. Mount the disk groups using the following SQL*Plus commands so that they are available to the Oracle database instance running on the backup host:

```
ALTER DISKGROUP  DATA_DGRP      MOUNT;
ALTER DISKGROUP  REDO_DGRP      MOUNT;
ALTER DISKGROUP  FLASH_DGRP     MOUNT;
ALTER DISKGROUP  ARCHIVE_DGRP   MOUNT;
```

3. Restore and recover the database using the following SQL*Plus commands:

```
SQL> STARTUP MOUNT;
SQL> FLASHBACK DATABASE TO RESTORE POINT  bkup1;
SQL> RECOVER AUTOMATIC DATABASE UNTIL CHANGE  scn USING BACKUP CONTROLFILE;
SQL> SHUTDOWN
SQL> STARTUP MOUNT
SQL> ALTER DATABASE OPEN RESETLOGS;
SQL> EXIT
```

Test matrix and results

Feature	Test	Results			Times run
		Start snap session	Clone fracture	MV/S and MV/A fracture	
CONSISTENT	Flashback **ON** Database **NOT** in **HOT BACKUP** mode	Successfully restored and rolled forward	Successfully restored and rolled forward	Successfully restored and rolled forward	5
	Flashback **ON** Database **NOT** in **HOT BACKUP** mode Disk group rebalance in progress	Successfully restored and rolled forward	Successfully restored and rolled forward	Successfully restored and rolled forward	5

Summary

Test results show that the consistency features of SnapView and MirrorView for the EMC VNX storage system, when used in conjunction with Oracle 11gR2's flashback database and restore point features, enable creation of a usable and valid Oracle database backup. This combination of EMC and Oracle technologies to create a consistent backup eliminates the need to put the database in hot backup mode. Because there is no need to put the database in hot backup mode prior to replication, the production database is not impacted and the Oracle database replication process is simplified.