EMC²
where information lives®

# EMC Business Continuity for Oracle Database 11g/10g

## Enabled by EMC CLARiiON CX4 and EMC Celerra Using FCP and NFS

Reference Architecture

EMC Global Solutions

EMC² PROVEN SOLUTIONS

vmware® TECHNOLOGY ALLIANCE PARTNER

Published December 2009



Part number: H6391.2

## Table of Contents

## Reference architecture overview

**Document purpose**

This document provides an overview of the architecture of a physically booted solution for midsize enterprises that uses the EMC® Celerra® unified storage platform and Oracle RAC 11*g* and 10*g* on Linux over FCP and NFS.

Information in this document can be used:

- As the basis for a solution build, white paper, best practices document, or training
- By other EMC organizations (for example, the technical services or sales organization) as the basis for producing documentation for a technical services or sales kit

**Solution purpose**

The purpose of this solution is to:

- Improve the performance, scalability, flexibility, and resiliency of an Oracle software stack that is physically booted on normal hardware by connecting multiple protocols to one storage platform as follows:
  - Fibre Channel Protocol (FCP) and Oracle ASM to access high-demand, low-latency storage elements
  - NFS to access all other storage elements
- Facilitate and reduce the risk of migrating existing Oracle Database 10*g* installations to 11*g* by providing documentation of best practices.
- Reduce cost by migrating an online production Oracle Database mounted over FCP to a target database mounted over NFS with no downtime and minimal performance impact. Improve the performance of an Oracle 11*g* or 10*g* production database by offloading all performance impacts of database operations, such as backup and recovery, using:
  - EMC Replication Manager®
  - EMC SnapView™

  These demonstrate significant performance and manageability benefits in comparison to normal Oracle Recovery Manager (RMAN) backup and recovery.
- Provide disaster recovery capability using:
  - EMC RecoverPoint with CLARiiON® splitters
  - EMC MirrorView™/Asynchronous over iSCSI

  These demonstrate significant performance and manageability benefits in comparison to normal Oracle Data Guard disaster recovery.
- Provide the capability to clone a running production database with minimal performance impact and no downtime using SnapView clones and Replication Manager.

**The business challenge**

Midsize enterprises face the same challenges as their larger counterparts when it comes to managing database environments. These challenges include:

- Rising costs
- Control over resource utilization and scaling
- Lack of sufficient IT resources to deploy, manage, and maintain complex environments at the departmental level
- The need to reduce power, cooling, and space requirements

Unlike large enterprises, midsize enterprises are constrained by smaller budgets and cannot afford a custom, one-off solution. This makes the process of creating a database solution for midsize enterprises even more challenging than for large enterprises.

**The solution**

This solution demonstrates how organizations can:

- Deploy a solution using a combination of the NFS and FCP protocols on the Celerra.
  - FCP is used for high-I/O and low-latency database objects (notably the datafiles, tempfiles, online redo logfiles, and controlfiles).
  - NFS is used for all other database objects (consisting basically of the flashback recovery area, disk-based backups, archive logs, and CRS files).
  - Manageability advantages are obtained by using a combination of FCP and NFS. Specifically, archived logs and backups can be accessed through a normal file system interface rather than ASM. Further, another clustered file system is not required for the CRS files. This simplifies the software installation and configuration on the database servers.
- Avoid investing in additional FC infrastructure by implementing a blended solution that uses both FCP and NFS to access storage elements.
- Work with different protocols in the blended solution to migrate an online production Oracle Database mounted over FCP to a target database mounted over NFS, with no downtime and minimal performance impact on the production database.
- Maximize the use of the database-server CPU, memory, and I/O channels by offloading performance impacts from the production server during:
  - Backup and recovery by using EMC Replication Manager or EMC SnapView
  - Disaster recovery operations by using EMC RecoverPoint or EMC MirrorView/A
- Reduce the complexity of backup operations and eliminate the need to implement scripted solutions by using EMC Replication Manager.
- Save time and maximize system uptime when migrating existing Oracle Database 10*g* systems to Oracle Database 11*g*.
- Implement a disaster recovery solution with MirrorView/A over iSCSI that reduces costs and complexity by using IP as the network protocol.
- Use EMC SnapView to free up the database server's CPU, memory, and I/O channels from the effects of operations relating to backup, restore, and recovery. SnapView clones also help in creating test/development systems without any impact on the production environment.

## Solution components

**Introduction**

This section describes the components of the solution, and explains some of the key terminology and concepts that are used in this document.

**Key terms defined**

**Solution**
A solution is a complete stack of hardware and software upon which a customer would choose to run their entire business or business function.

**Solution attributes**

A solution attribute addresses the entire solution stack, but does so in a way relating to a discrete area of testing. For example, performance testing is a solution attribute.

**Solution component**

A solution component addresses a subset of the solution stack that consists of a discrete set of hardware or software, and focuses on a single IT function. For example, backup and recovery, and disaster recovery are both solution components.

A solution component can be either “basic” or “advanced.”

**Basic solution component**

A basic solution component uses only the features and functionality provided by the Oracle stack. Examples are RMAN for backup and recovery and Data Guard for disaster recovery.

**Advanced solution component**

An advanced solution component uses the features and functionality of EMC hardware or software. Examples are EMC SnapView for backup and recovery and EMC MirrorView for disaster recovery.

**Solution attributes**

The following table describes the solution attributes that are included in this solution:

| Component | Description |
|---|---|
| Scale-up OLTP | Using an industry-standard OLTP benchmark against a single database instance, comprehensive performance testing is performed to validate the maximum achievable performance using the solution stack of hardware and software. |
| Resiliency | The purpose of resiliency testing is to validate the fault-tolerance and high-availability features of the hardware and software stack. Faults are inserted into the configuration at various layers in the solutions stack. Some of the layers where fault tolerance is tested include: Oracle RAC node, Oracle RAC node interconnect port, storage processors, and Data Movers. |

**Solution components**

The following table describes the solution components that are included in this solution:

| Component | Description |
|---|---|
| Basic Backup | This is backup and recovery using Oracle RMAN, the built-in backup and recovery tool provided by Oracle. |
| Advanced Backup | This is backup and recovery using EMC value-added software or hardware.<br>In this solution the following are used to provide Advanced Backup functionality:<br>• EMC Replication Manager<br>• EMC SnapView |
| Basic Protect | This is disaster recovery using Oracle Data Guard, Oracle's built-in remote replication tool. |
| Advanced Protect | This is disaster recovery using EMC value-added software and hardware:<br>In this solution the following are used to provide Advanced Protect functionality:<br>• EMC RecoverPoint with CLARiiON splitters<br>• EMC MirrorView/A over iSCSI |
| Test/Dev | A running production OLTP database is cloned with minimal, if any, performance impact on the production server, as well as no downtime. The resulting dataset is provisioned on another server for use for testing and development.<br>EMC Replication Manager is used to automate the Test/Dev process. |
| Migration | An online production Oracle database that is mounted over FCP/ASM is migrated to a target database mounted using NFS, with no downtime and minimal performance impact on the production database. |

## Technology solution

**Overview**

Two sites connected by a WAN are used in the solution, one site is used for production and the other site is used as a disaster recovery target. A Celerra is present at each site. Oracle RAC 11*g* or 10*g* for x86-64 is run on Red Hat Enterprise Linux or on Oracle Enterprise Linux.

FCP storage networks consisting of dedicated, redundant FCP switches are present at both sites. An EMC RecoverPoint cluster is also included at each site.

The solution also includes virtualized servers for use as Test/Dev, Basic Protect, and Advanced Protect targets. Virtualization of the test/dev and disaster recovery (DR) target servers is supported using VMware ESX Server.

**Blended FCP/NFS**

This is a blended FCP/NFS solution. Depending on the nature of the database object, either FCP or NFS is used to access it. The following table shows which protocol is used to access each database object.

| Database object | Type | Accessed using |
|---|---|---|
| Datafiles | High demand<br>Low latency | FCP |
| Online redo logfiles | | |
| Controlfiles | | |
| Tempfiles | | |
| Flashback recovery area | Low demand<br>High latency | NFS |
| Archive logs | | |
| Disk-based backups | | |
| CRS files | | |

**Production site**

The following components are present at the production site and are connected to the production FCP storage network and to the WAN:

- A Celerra (actually the CLARiiON back-end array)
- A physically booted four-node Oracle RAC 11*g* or 10*g* cluster
- A RecoverPoint cluster connected to the FCP storage network, and to the WAN

The Oracle RAC 11*g* or 10*g* servers are also connected to the client and RAC interconnect networks.

**Disaster recovery target site**

The disaster recovery target site consists of:

- A target Celerra (actually the CLARiiON back-end array) connected to the target FCP storage network
- A RecoverPoint cluster connected to the FCP storage network and to the WAN

**Connected to both sites**

The following are present at both sites:

- A VMware ESX server is connected to both the production and target FCP storage networks.
- A virtualized single-instance Oracle 11*g* or 10*g* server is used as:
  - The disaster recovery target for Basic Protect and Advanced Protect (DR site)
  - The target for Test/Dev (production site)

  The virtualized single-instance Oracle 11*g* or 10*g* target server accesses both the production and target FCP storage networks and is connected to the client WAN through virtualized connections on the virtualization server.
- A virtualized Replication Manager Server is responsible for handling replication tasks through the Replication Manager agent, which is installed on the production database servers.

The LUNs on the Celerra are discovered using Raw Device Mapping (RDM) on the target VMs.

**Storage layout**

The following table describes how each Oracle file type and database object is stored and accessed for this solution:

<table>
<tr><th>What</th><th>Protocol</th><th>Stored on</th><th>File-system type</th></tr>
<tr><td>Oracle datafiles</td><td rowspan="4">FCP</td><td rowspan="4">FC disk (LUNs)</td><td rowspan="4">RAID-protected ASM diskgroup</td></tr>
<tr><td>Oracle tempfiles</td></tr>
<tr><td>Oracle online redo logfiles</td></tr>
<tr><td>Oracle controlfiles</td></tr>
<tr><td>Voting disk</td><td rowspan="5">NFS</td><td rowspan="2">FC disk</td><td rowspan="5">RAID-protected NFS</td></tr>
<tr><td>OCR files</td></tr>
<tr><td>Archived logfiles</td><td rowspan="3">SATA II</td></tr>
<tr><td>Flashback recovery area</td></tr>
<tr><td>Backup target</td></tr>
</table>

High-performance database objects are accessed over an FCP network using redundant network switches.

| | |
|---|---|
| **ASM and ASMLib** | • Oracle ASM is used as the file system/volume manager. ASMLib is used to virtualize the LUNs on the database server.<br>• Oracle datafiles, tempfiles, and online redo logfiles are stored on separate LUNs that are mounted on the database server using ASM over FCP.<br>• Three ASM diskgroups are used — one diskgroup for datafiles and tempfiles, and two diskgroups for online redo logfiles.<br>• The online redo logfiles are mirrored across the two ASM diskgroups using Oracle software multiplexing. The controlfiles are mirrored across the online redo log ASM diskgroups.<br>• Each ASM diskgroup and its underlying LUNs are designed to satisfy the I/O demands of individual database objects. For example, RAID 5 is used for the datafiles and the tempfiles, but RAID 1 is used for the online redo logfiles. All of these diskgroups are stored on FC disks. |
| **Network architecture** | • TCP/IP and NFS provide network connectivity and file system semantics for NFS file systems on Oracle RAC 11*g* or 10*g*.<br>• Client virtual machines run on the VMware ESX server. They are connected to a client network.<br>• Client, RAC interconnect, and redundant TCP/IP storage networks consist of dedicated network switches and virtual local area networks (VLANs).<br>• The RAC interconnect and storage networks consist of trunked IP connections to balance and distribute network I/O. Jumbo frames are enabled on these networks. |

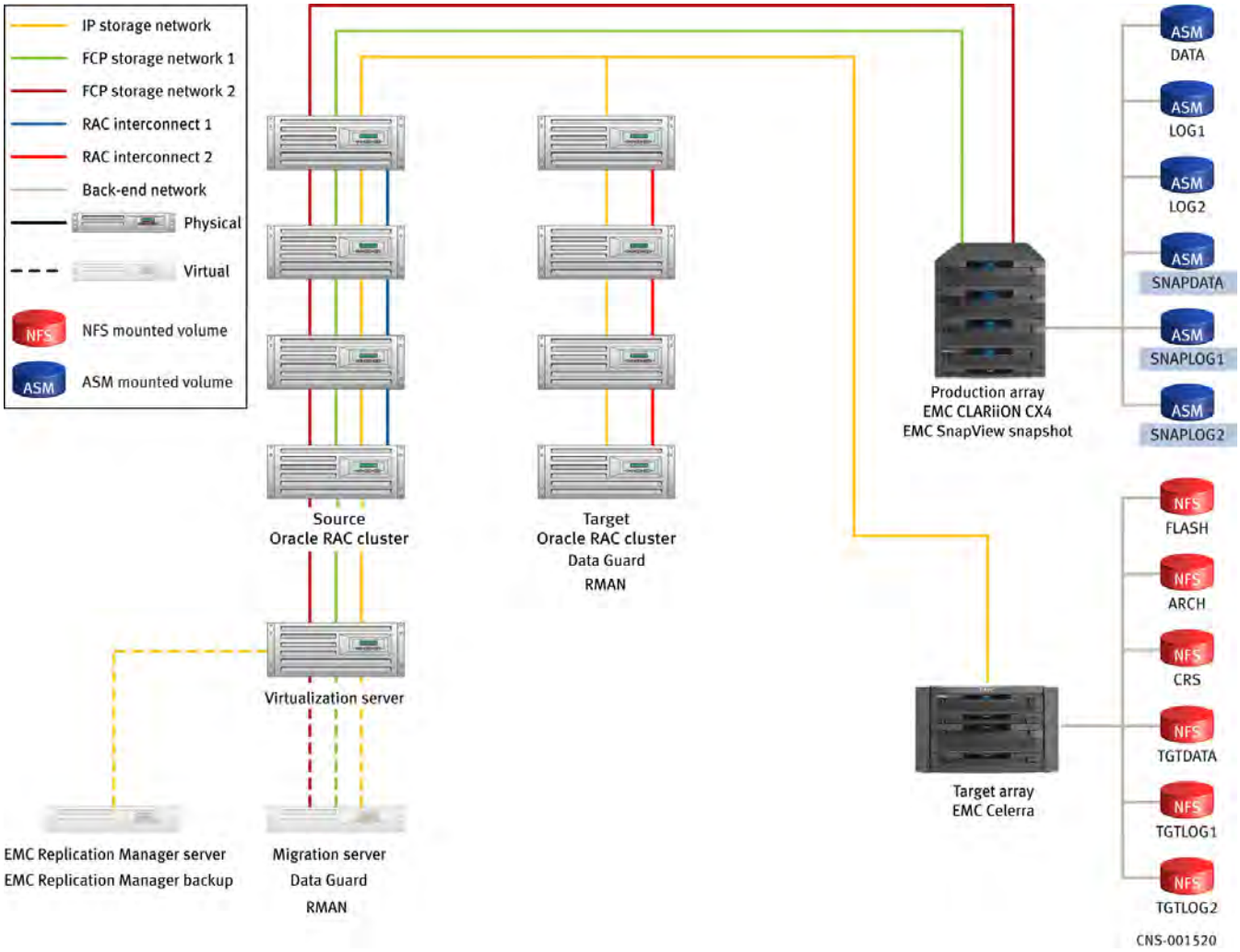
## Migration

**Introduction**

The ability to migrate an Oracle database across storage protocols is a frequent customer request. The EMC Oracle CSV group has tested and validated a solution component for migrating an online production Oracle database, which is mounted over FCP/ASM to a target database mounted using NFS. This is performed with minimal performance impact on the production database and no downtime.

**Migration diagram**

The following illustration is a high-level view of the migration component.



**Migrating an online Oracle database**

These steps were followed to perform the migration operation:

| Step | Action |
|---|---|
| 1 | Using EMC Replication Manager, a consistent backup of the running production database is performed on the CLARiiON utilizing CLARiiON SnapView snapshot. |
| 2 | This backup is mounted (but not opened) on the migration server in this case a VMware VM (a physically booted server would also work). The NFS target array is also mounted on the migration server. |
| 3 | Using Oracle Recovery Manager (RMAN), a backup of this database is taken onto the target location. This backup is performed as a database image, so that the datafiles are written directly to the target NFS mount. |
| 4 | The migration server is then switched to the new database, which has been copied by RMAN to the NFS mount. |
| 5 | The target database is set in Data Guard continuous recovery mode, and Data Guard log ship/log apply is used to catch the target database up to the production version. |
| 6 | Once the target database is caught up to production, Data Guard failover can be used to retarget to the target database. If appropriate networking configuration is performed, clients will see no downtime when this operation occurs. |

The end result, as stated above, is that the production FCP-mounted database can be migrated to NFS with minimal performance impact and no downtime.

## Key EMC components

**Introduction**

This section provides an overview of the EMC technologies that are used in this solution.

**EMC Celerra unified storage platform**

The Celerra is a remarkably versatile device. Celerra includes a world-class network-attached storage (NAS) array combined with the functionality and performance of the leading midrange storage area network (SAN) array. Celerra provides both NAS and SAN functionality and performance without compromise.

The key features provided by Celerra are described in the following table:

| Feature | Provided by |
|---|---|
| SAN storage | Fibre Channel Protocol (FCP) through the back-end EMC CLARiiON storage array |
| NAS | Network File System (NFS) and<br>Common Internet File System (CIFS) protocols |
| iSCSI storage | Celerra's Data Movers |

**EMC Replication Manager**

EMC Replication Manager manages EMC point-in-time replication technologies through a centralized management console. Replication Manager coordinates the entire data replication process – from discovery and configuration to the management of multiple application consistent disk-based replicas. Replication Manager can auto-discover your replication environment and enable streamlined management by scheduling, recording, and cataloging replica information including auto-expiration.

Replication Manager integrates with the Oracle database server and provides an easy interface to create and manage Oracle replicas.

**EMC SnapView**

EMC SnapView lets you create local point-in-time snapshots and complete data clones for testing, backup, and recovery operations. With SnapView, you can create multiple copies of production data on your CLARiiON quickly and easily. SnapView enables you to:

- Carry out non-disruptive backup operations with space-saving SnapView snapshots
- Restore data quickly and easily with SnapView clones
- Accelerate application development by quickly providing the most current data for testing

**EMC RecoverPoint**

EMC RecoverPoint provides integrated continuous remote replication (CRR) and continuous data protection (CDP) for on-demand protection and recovery at any point in time. RecoverPoint's advanced capabilities include policy-based management, application integration, and WAN acceleration.

RecoverPoint can be implemented to provide a single, unified solution to protect and replicate data across heterogeneous storage. This enables the customer to simplify management and reduce costs, recover data at a local or remote site at any point in time, and ensure continuous replication to a remote site without impacting performance.

---

**EMC CLARiiON splitter**

The CLARiiON (CX) splitter is a driver that can be directly installed on a CLARiiON. When an application writes to storage, the CX splitter splits the data and sends one copy of the data over FCP to the source-site RecoverPoint appliance (RPA), and a second copy to the storage array. These requests are acknowledged by the source RPAs and are then replicated to the remote RPAs over the IP WAN network.

The main advantages of CLARiiON array-based splitters are that they:

- Reduce the cost associated with the additional host-based splitter agents or specialized fabric
- Provide concurrent local and remote (CLR) protection. CLR replication can be performed on the same LUNs.

---

**EMC MirrorView/ Asynchronous over iSCSI**

EMC MirrorView software provides highly available data protection across a region or across the globe. MirrorView/Asynchronous (MirrorView/A) can be used to consolidate information to one centralized CLARiiON array for centralizing remote processes such as backups or disaster failover.

The advantages of using MirrorView/A over iSCSI to provide a disaster recovery solution are as follows:

**IP network**

The use of IP as the network protocol reduces costs and complexity. No conversion of FCP to IP is required.

**Simplified deployment**

The use of consistency technology allows you to deploy a disaster recovery solution with a more simplified Oracle Application Stack on the target side. Since Data Guard is not required on the target side, in many cases Oracle 11*g* or 10*g* Standard Edition can be used.

---

## Key third-party components

**Introduction**

This section provides an overview of the third-party technologies that are used in this solution.

**Oracle software stack**

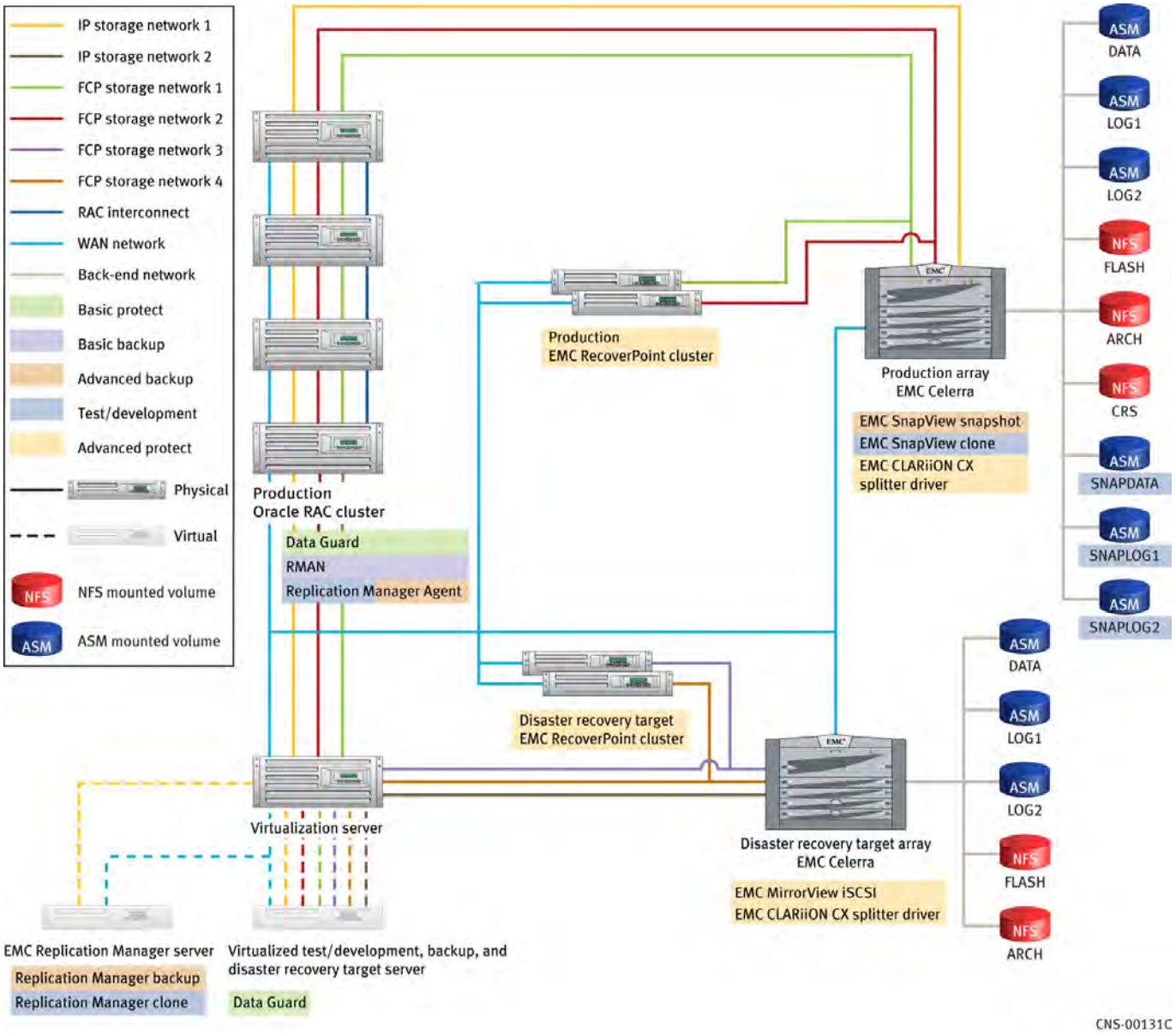
The Oracle software stack covered by this solution consists of:

- Oracle Database 10*g* or 11*g* using Real Application Clusters (RAC)
- Oracle Enterprise Linux
- Cluster Ready Services (CRS)
- Automatic Storage Management (ASM)

## Physical architecture

**Architecture diagram**

The following diagram depicts the overall physical architecture of the solution.



## Validated environment profile

| | |
|---|---|
| **Environment profile and test results** | For information on the validated environment profile and performance results, refer to the Proven Solution Guide for this solution. This information can be accessed on EMC Powerlink®, EMC.com, and EMC\|KB.WIKI.<br><br>If you do not have access to this content, contact your EMC representative. |

## Conclusion

**Summary**

This section provides a summary of the solution and of the business challenges that it addresses.

**Reduced total cost of ownership**

In any reasonable configuration, the database server's CPU is the most precious component of the entire architecture. Therefore, the over-arching principle of EMC's Oracle RAC 11*g* and 10*g* solutions for midsize enterprises is to free up the database server's CPU (as well as memory and I/O channels) from utility operations such as backup and recovery, disaster recovery staging, test/dev, and cloning. The highest and best use of the database server's CPUs is to parse and execute the SQL statements that are required by the application user.

**CPU usage**

This solution reduces the load on the database server CPU by using EMC Replication Manager and EMC SnapView to carry out a physical backup of an Oracle 11*g* or 10*g* production database while offloading all performance impacts of the backup operation from the production server.

**Blended FCP/NFS architecture**

The use of a blended FCP/NFS architecture provides very significant ease-of-use advantages. Provided that the customer has access to a Celerra unified storage platform, the amount of software that the customer is required to install and configure on the database server is significantly reduced.

The NFS protocol provides access to storage elements such as archive logs and disk-based backups using normal file-system semantics. This is a significantly simpler form of management than that provided through FCP and ASM.

The use of FCP allows for high-performance, low-latency I/O for databases that require high-performance storage.

**Advanced Backup and Recovery**

EMC Replication Manager is a comprehensive graphical application that provides Oracle storage replication using EMC storage technology. This eliminates the requirement for the customer to write scripts or to manually perform replication tasks. These tasks can now be fully automated and managed by Replication Manager.

One of the solution components using Replication Manager in the solution presented in this reference architecture included Advanced Backup and Recovery using CLARiiON SnapView snapshots.

**Advanced Protect**

**EMC RecoverPoint and CX splitters**

The advantages of Advanced Protect using RecoverPoint through the CX splitter driver are:

- **Cost savings:** RecoverPoint provides data compression of between two and fifteen times before shipping data over the wire (typical compression is approximately five times).
- **Superior data integrity:** RecoverPoint logs every change and maintains those changes as long as the journal volume is capable of tracking them (this depends on change rate and the size of the journal).
- **No disruption to replication of production data:** RecoverPoint allows the customer to rapidly provision a copy of the database to test disaster recovery or other tasks, without any disruption to the ongoing replication of the production data.
- **Simultaneous local and remote replication:** RecoverPoint is unique in that it allows for simultaneous replication targets from the same source.
- **Minimal performance impact:** The performance testing carried out by EMC demonstrates that RecoverPoint has minimal, if any, impact on performance.

**Advanced Protect using EMC MirrorView/A through iSCSI**

The advantages of Advanced Protect using MirrorView/A through iSCSI are:

- **Simplified deployment:** The use of consistency technology allows the customer to deploy a disaster recovery solution with a more simplified Oracle Application Stack on the target side. Since Data Guard is not required on the target side, in many cases Oracle 11*g* or 10*g* Standard Edition can be used.
- **IP network:** The use of IP as the network protocol reduces costs and complexity. No conversion of FCP to IP is required.

**Test/Dev**

The ability to deploy a writeable copy of the production database is required by many customers. The process of provisioning this copy must create minimal, if any, performance impact on the production database server. Also, absolutely no downtime can be tolerated. The Test/Dev solution component documented here provides this deployment ability.

**Robust performance and scaling**

The resiliency testing carried out by EMC ensures that the database configuration is reliable. High availability is used at every major layer of the solution, including the database server, NAS file server, and back-end SAN array. By testing the fault tolerance of all of these layers, the ability of the application to withstand hardware failures with no downtime is assured.

The performance testing carried out by EMC utilizes an industry-standard OLTP benchmark, but does so without exotic tunings that are not compliant with best practices. In addition, real-world configurations that the customer is likely to deploy are used. This enables the customer to be reasonably assured that the configuration that they choose to run their application will do so predictably and reliably.

**Next steps**

EMC can help accelerate assessment, design, implementation, and management while lowering the implementation risks and costs of an end-to-end solution for an Oracle Database 11*g* or 10*g* environment.

To learn more about this and other solutions contact an EMC representative or visit http://www.emc.com/solutions/application-environment/oracle/solutions-for-oracle-database.htm.

## Appendix: Hardware and software resources

**Hardware** The hardware used to validate the solution is listed below.

| Equipment | Quantity | Configuration |
|---|---|---|
| EMC Celerra unified storage platforms (includes an EMC CLARiiON CX4 back-end storage array) | 2 | • 2 Data Movers<br>• 4 GbE network connections per Data Mover<br>• 2 or 3 FC shelves<br>• 1 SATA shelf<br>• 30 or 45 73 GB FC disks (depending on configuration)<br>• 15 500 GB SATA disks<br>• 1 Control Station<br>• 2 storage processors<br>• DART version 5.6.44-4 |
| Gigabit Ethernet switches | 5 | 24 ports per switch |
| FCP switches | 2 | • 16 ports<br>• 4 Gb throughput |
| Database servers (Oracle RAC 11*g*/10*g* servers) | 4 | • 2 2.66 GHz Intel Pentium 4 quad-core processors<br>• 24 GB of RAM<br>• 2 146 GB 15k internal SCSI disks<br>• 2 onboard GbE Ethernet NICs<br>• 2 additional Intel PRO/1000 PT quad-port GbE Ethernet NICs<br>• 2 SANblade QLE2462-E-SP 4 Gb/s dual-port FC HBAs (4 ports in total) |
| EMC RecoverPoint appliances (RPA) | 4 | • 2 Dell 2950 servers per site<br>• QLA2432 HBA cards |
| Virtualization server (VMware ESX server) | 1 | • 4 2.86 GHz AMD Opteron quad-core processors<br>• 32 GB of RAM<br>• 2 146 GB 15k internal SCSI disks<br>• 2 onboard GbE Ethernet NICs<br>• 3 additional Intel PRO/1000 PT quad-port GbE Ethernet NICs<br>• 2 SANblade QLE2462-E-SP 4 Gb/s dual-port FC HBAs (4 ports in total) |

**Software** The software used to validate the solution is listed below.

| Software | Version |
|---|---|
| Oracle Enterprise Linux | 4.7 |
| VMware ESX Server/vSphere | 4.0 |
| Oracle VM | 2.1.1 |
| Microsoft Windows Server 2003 Standard Edition | 2003 |
| Oracle RAC Enterprise Edition | 11*g* or 10*g* |
| Oracle Database Standard Edition | 11*g* or 10*g* (11*g* Version 11.1.0.7.0) |
| Quest Benchmark Factory for Databases | 5.8.1 |
| EMC Celerra Manager Advanced Edition | 5.6 |
| EMC Navisphere® Agent | 6.26.0.2.24 |
| EMC PowerPath® | 5.0.0 (build 157) |
| EMC FLARE® | 04.28.000.5.504 |
| EMC DART | 5.6.44-4 |
| EMC Navisphere Management | 6.28 |
| EMC RecoverPoint | 3.0 SP1 |
| EMC Replication Manager | 5.2.0.0 |
| EMC MirrorView | 6.7 |
| EMC CLARiiON splitter driver | 04.28.006.6.003 |