

Optimizing Disaster Recovery with VMware Site Recovery Manager and EMC Celerra Replicator V2

Applied Technology

Abstract

Disaster recovery requirements are critical to applications deployed on virtual infrastructures. This white paper outlines a solution that leverages the native replication capabilities of the EMC® Celerra® product family and VMware Site Recovery Manager. Celerra Replicator™ delivers both reliable and predictable recovery point objectives, while VMware Site Recovery Manager offers accelerated recovery time objectives.

August 2008

Copyright © 2008 EMC Corporation. All rights reserved.

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED “AS IS.” EMC CORPORATION MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on EMC.com

All other trademarks used herein are the property of their respective owners.

Part Number H5700

Optimizing Disaster Recovery with
VMware Site Recovery Manager and
EMC Celerra Replicator V2
Applied Technology

Table of Contents

Executive summary	4
Introduction	4
Audience	5
Terminology	5
Overview	7
Site Recovery Manager technical overview	8
Site Recovery Manager restrictions	11
Celerra Replicator V2 technical overview	11
Meeting service level agreements (SLAs)	11
Efficient use of IP network infrastructure	12
Multi-protocol support	12
Multi-site replication	12
Celerra Replicator V2 – how it works	12
Celerra Replicator V2 and Site Recovery Manager	14
Architectural overview	14
Configuration considerations	14
Celerra and SRM use cases	15
Disaster recovery test	15
Disaster recovery	15
Migration	15
Shared disaster recovery protection	16
Use cases not easily supported by SRM	16
Frequent data center switch-over	16
Multi-site disaster recovery	16
Graceful failover and reverse replication	16
Celerra and SRM failback	16
Conclusion	18
References	18

Executive summary

Today, organizations rely heavily on their information technology infrastructure. If IT services were lost, many organizations would lose customers and some would go out of business. Having a comprehensive, effective and efficient disaster recovery plan can help assure long-term success. New federal laws for safeguarding information increase the importance of IT infrastructure protection and bring greater relevance of service-level management paradigms.

Application downtime can cost organizations up to several hundred thousand dollars per hour. Imagine how devastating it can be to a company with a data center that is offline for an extended time period due to a disaster or unforeseen event. Effectively protecting and recovering critical applications quickly is essential for continued business operation.

Over the last few years, the rapid adoption and growth of virtualized infrastructures have been sweeping IT. The majority of the virtualized infrastructure market uses VMware because it delivers a reliable and mature solution set for server consolidation, more effective application provisioning models and huge improvements in efficiency and power consumption. Many applications that corporations seek to protect are migrating to a virtualized server model. The value of these applications and their data do not change when running above VMware, so the same disaster recovery requirements, methodologies and practices set forth by the business in their service level guidelines must adapt to support this new operational model.

The growing ubiquity of the virtualization layer presents a compelling opportunity to add functionality at this layer, and provide broad reaching disaster recovery orchestration. An additional benefit is improved recovery times from a single point of management for all the applications running on VMware.

Site Recovery Manager (SRM) integrates with all key EMC platforms and replication products, allowing customers to deliver a truly service-oriented and comprehensive disaster recovery methodology with a rapid, reliable and predictable recovery process, taking risk and worry out of disaster recovery. With the release of the VMware Site Recovery Manager, a plug-in for VMware ESX Infrastructure 3 and the EMC® Celerra® NS Series iSCSI EMC Celerra Replicator™ Adapter for VMware SRM, customers can implement disaster recovery for ESX Server virtual machines using Celerra NS Series storage systems running Celerra-based Replication V2 and Celerra SnapSure™ replication software.

Introduction

This white paper is intended to provide a technical overview of the new VMware Site Recovery Manager technology and the Celerra replication tools that it integrates with. Using a business use case structure, the paper articulates:

- How SRM and Celerra Replicator deliver customer value.
- How the technologies work in steady state operation and in the VMware disaster recovery context.
- The required components for a functional system.
- Considerations and best practices when integrating the two technologies.

It does not cover the details of product configuration, although this is covered in other materials available from EMC and VMware that are referenced at the end of this document.

Audience

The audience for this paper includes storage and VMware administrators, technical architects, and IT managers. A high level knowledge of VMware, Celerra, and disaster recovery concepts is assumed.

Terminology

Celerra Replicator V2: The licensed Celerra Replication feature that integrates with Site Recovery Manager.

Datastore: A storage location for virtual machine files, either VMFS or NFS, which serves as a virtual representation of an underlying pool of physical storage resources.

Datastore group: A datastore or group of datastores that are represented by an iSCSI LUN or a set of associated iSCSI LUNs on a Celerra. SRM automatically builds datastore groups based on the LUNs that are replicated.

Disaster recovery: Disaster recovery is the process of rebuilding data from a backup image, and then explicitly applying subsequent logs to roll the data state forward to a designated point of consistency. The mechanism to create recoverable copies of the data depends on the database and applications.

EMC Celerra NS Series iSCSI EMC Celerra Replicator Adapter for VMware Site Recovery Manager (SRM): A software package that allows the VMware Site Recovery Manager (SRM) to implement disaster recovery for ESX Server virtual machines using Celerra NS Series storage systems running Celerra-based replication v2 and Celerra SnapSure replication software.

EMC SnapSure: Celerra Network Server software feature that enables you to create and manage checkpoints, which are point-in-time, logical images of a production file system (PFS)

Experimental Support: Experimental support means that the feature is fully coded, functional, and expected to work, but has not been tested enough for full VMware support to be declared.

Failback: The process of returning to the protected site after implementing a failover following DR.

Failover: Initiated from SRM, the failover process will initiate a recovery plan and restart those VMs with their associated storage resources at the recovery site.

iSCSI: iSCSI ("Internet SCSI") is a protocol that allows clients (called initiators) to send SCSI commands to SCSI storage devices (targets) on remote servers. It is a Storage Area Network (SAN) protocol, allowing organizations to consolidate storage into data center storage arrays with block disk access over an IP network.

Protected Site: The source (production) VMware virtual data center.

Protection Group: Sets of VMs that are associated together are placed in a common protection group.

Raw Device Mapping (RDM): RDM includes a combination of a pointer, which is a .vmdk file that resides on a VMFS volume, and a physical raw device that the .vmdk file points to. RDM physical compatibility mode is used if importing a replicated device that was replicated or used by a physical server into a virtual environment, or in some cases where physical device level actions are required, for example, storage level snaps and replication.

Recovery Plan: An individual failover scenario. SRM may contain multiple recovery plans with different levels of granularity.

Recovery Point Objective (RPO): A business service level specifying the amount of data that is acceptable for an application to lose when recovering from a disaster.

Recovery Site: The target (secondary or failover) VMware virtual data center.

Recovery Time Objective (RTO): A business service level specifying the amount of down- time that is acceptable for an application prior to being re-started following the declaration of a disaster.

Resource Pools: Resource pools in VMware are used to hierarchically partition available CPU and memory resources within standalone hosts or within each DRS cluster.

Shadow Virtual Machine: A virtual machine that has been activated as part of a SRM failover process.

Silvering: Silvering is the process of telling the primary disk set to copy its data over to the backup mirror so that it is synchronized with the primary disk set.

Simulated Failover: Simulated failover refers to the process of testing a failover using the “Test Failover” command in Site Recovery Manager and leveraging snapshots at the remote site.

Site Recovery Manager: Site Recovery Manager (SRM) is a software plug-in that extends the VMware VirtualCenter console and interfaces with array-based replication tools (SRDF[®] for Symmetrix[®], MirrorView[™] for CLARiiON[®], RecoverPoint for open network replication and Celerra Replicator).

Snapshots: A snapshot is a logical copy of a file system or disk storage device, as they were at a particular point in time. Snapshots do not require the full space required for a mirror copy and typically employ a “copy on first write” methodology, where old data is copied out of the volume as it is overwritten to maintain the point in time copy of the storage object.

Time Out of Sync: The goal time difference at any point between the data image at the primary site and the restartable image at the secondary site (in minutes).

Virtual Infrastructure Client (VI Client): An interface that allows administrators and users to connect remotely to the VirtualCenter Management Server or individual ESX Server installations from any Windows platform.

Virtual Machine (VM): A virtualized x86 PC on which a guest operating system and associated application(s) run. A VM is also a set of discrete files that primarily include a .vmx configuration and one or more .vmdk virtual disk files.

VirtualCenter Management Server: VirtualCenter delivers centralized management, operational automation, resource optimization, and high availability to IT environments. Virtualization-based distributed services provided by VMotion, DRS, and HA equip the dynamic data center with unprecedented levels of serviceability, efficiency, and reliability. Automated resource optimization with DRS aligns available resources with predefined business priorities while streamlining labor and resource intensive operations. Migration of live virtual machines with VMotion makes the maintenance of IT environments nondisruptive. HA enables cost-effective application availability independent of hardware and operating systems. VirtualCenter delivers the high levels of simplicity, efficiency, security, and reliability required to manage virtualized IT environment of any size.

VMware ESX Server: VMware ESX Server is the foundation for delivering virtualization-based distributed service to IT environments. A core building block of VMware Infrastructure, ESX Server is a robust, production-proven virtualization layer that abstracts processor, memory, storage and networking resources into multiple virtual machines running side-by-side on the same server. Sharing hardware resources across a large number of virtual machines increases hardware utilization and dramatically decreases capital and operating cost. Virtual machines can be equipped with high availability, resource management, operational automation and security features that improve service levels even to the most resource-intensive mission critical applications. ESX Server delivers the highest levels of performance, scalability, and robustness required for enterprise IT environments.

VMware Virtual Machine File System (VMFS): VMware VMFS is a high-performance cluster file system for ESX Server virtual machines. Each virtual machine is encapsulated in a small set of files and VMFS is the default storage system for these files on physical SCSI disks and partitions. VMFS efficiently stores the entire virtual machine state in a central location to simplify virtual machine provisioning and administration. VMFS is a cluster file system that allows multiple ESX Servers to access the same virtual machine storage concurrently. A cluster file system is required for the virtualization-based distributed infrastructure services delivered by VMware VirtualCenter, VMware VMotion, VMware Distributed Resource Scheduler (DRS), and VMware High Availability (HA). It is used for holding VM images (vmdks) and is required for production support of SRM.

VMware Virtual Machine: Representation of a physical machine by software. A virtual machine has its own set of virtual hardware (for example, RAM, CPU, NIC, or hard disks) on which an operating system and applications are loaded. The operating system looks for a consistent and normalized set of hardware regardless of the actual physical hardware components. VMware virtual machines contain advanced hardware features, such as 64-bit computing and virtual symmetric multiprocessing.

Some optional components of VMware Infrastructure are:

VMware DRS: VMware DRS dynamically allocates and balances computing capacity across a collection of hardware resources aggregated into logical resource pools. VMware DRS continuously monitors utilization across resource pools and intelligently allocates available resources among the virtual machines based on pre-defined rules reflecting business needs and changing priorities. When a virtual machine experiences an increased load, VMware DRS automatically allocates additional resources by redistributing virtual machines among the physical servers. VMware DRS optimizes IT environments to align resources with business goals while ensuring flexibility and efficient utilization of hardware resources.

VMware HA: VMware HA provides an easy-to-use, cost-effective high availability for applications running in virtual machines. In the event of server failure, the affected virtual machines are automatically restarted on other production servers with spare capacity. HA minimizes downtime and IT service disruption while eliminating the need for dedicated standby hardware and installation of additional software. VMware HA provides a uniform high availability across the entire virtualized IT environment without the cost and complexity of failover solutions tied to either operating systems or specific applications.

VMware VMotion: VMware VMotion enables the live migration of running virtual machines from one physical server to another with zero downtime, continuous service availability, and complete transaction integrity. Live migration of virtual machines enables companies to perform hardware maintenance without scheduling downtime and disrupting business operations. VMotion also allows virtual machines to be continuously and automatically optimized within resource pools for maximum hardware utilization, flexibility, and availability. VMotion is a key enabling component of the dynamic, automated, and self-optimizing data center.

Overview

When considering disaster recovery (DR) service levels, customers tend to look at two critical business parameters: the recovery point objective and the recovery time objective. The recovery point objective (RPO) defines the maximum amount of data an application can tolerate losing in a disaster and the recovery time objective (RTO) defines the maximum amount of time it can take to restart applications at a recovery site. As a best practice, these parameters should be defined for each application requiring protection with remote replication solutions. Once they are defined, the IT organization must deliver solutions that can meet these objectives for each application. It is always desirable to define RPO and RTO values of zero (that is, zero data lost and zero time to restore), although this is too cost-prohibitive for all but the most mission-critical applications.

Many applications can be protected with a common mechanism by leveraging a storage array-based replication tool. This reduces the management overhead of managing replication through the application or at the server OS level. Celerra Replicator provides efficient snapshot-based, asynchronous, remote data replication of Celerra NAS file systems and Celerra iSCSI LUNs over IP networks. The EMC Celerra Replicator solution is highly efficient and scalable, supporting tens of terabytes and more than a thousand replicated objects. All this is done with a highly flexible, adaptive mechanism to ensure RPO service levels are met, even when many replicated objects are configured with different RPOs.

RPO and RTO definitions are typically applied at the application level, and need to be met regardless of whether the application is implemented in a physical server environment or in a virtualized server environment. Within a replication solution based on an external storage platform, failover of the application data is managed more effectively when compared to local or direct-attached storage, but it is still a significant task. Additionally, the actual failover of the server environment, the applications themselves, the

infrastructure services and the networks are highly manual processes that are defined in runbooks. These runbooks are problematic for a number of reasons:

- They can be hundreds of pages long.
- They can become outdated quickly due to changes in the protection needs of an enterprise.
- It is difficult to train others on how to use runbooks.
- They are prone to errors and often take more than a day to execute.

For these reasons, VMware SRM provides tremendous value. As a common server architecture layer, VMware can fulfill the role of disaster recovery orchestration, allowing failover to be codified into an automated set of steps for all of the servers, applications and networks under VMware control.

Site Recovery Manager technical overview

Site Recovery Manager (SRM) is a software package that extends the VMware VirtualCenter console and interfaces with array-based replication tools. SRM extends VirtualCenter by introducing a new management UI and allows administrators to build and conduct DR test and DR failover scenarios (previously managed with complex, manual runbooks) that can be executed at the push of a button. These scenarios may cover the failover of a single server or a whole site, allowing greater flexibility in configuration.

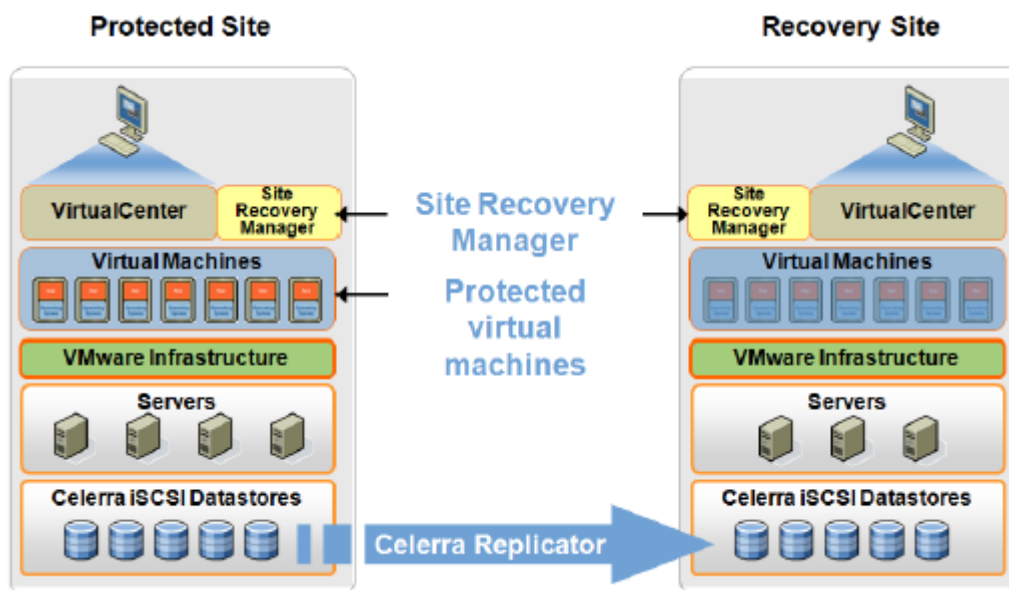


Figure 1 SRM components

SRM simplifies and automates the key elements of disaster recovery: setting up disaster recovery plans, testing those plans, and executing failover when a data center disaster occurs. The benefits of using SRM include:

- **Accelerating recovery with automated processes.** Traditional disaster recovery solutions leave many organizations unable to meet recovery time and recovery point objectives. The slow and often manual recovery processes common in traditional disaster recovery solutions are prone to errors and result in frequent failures. VMware Site Recovery Manager automates the recovery process to ensure that it is executed rapidly and accurately.
- **Ensuring reliable recovery with thorough automation and easier testing.** Testing disaster recovery plans and ensuring that they are executed correctly are critical to making recovery reliable. However, testing is difficult with traditional solutions due to the high cost, complexity and disruption associated

with tests. Another challenge is ensuring that staff are trained and prepared to successfully execute the complex process of recovery.

Site Recovery Manager helps you overcome these obstacles by enabling realistic, frequent tests of recovery plans and eliminating common causes of failures during recovery. It provides built-in capabilities for executing realistic, non-disruptive tests without the cost and complexity of traditional disaster recovery testing. Because the recovery process is automated, you can also ensure that the recovery plan will be carried out correctly in both testing and failover scenarios. Site Recovery Manager also leverages VMware Infrastructure to provide hardware-independent recovery to ensure successful recovery even when recovery hardware is not identical to production hardware.

- **Taking control of your disaster recovery plans.** Until now, keeping recovery plans and the runbooks that documented them accurate and up-to-date has been practically impossible due to the complexity of plans and the dynamic environment in today's data centers. Adding to that challenge, traditional solutions do not offer a central point of management for recovery plans and make it difficult to integrate the different tools and components of disaster recovery solutions.

VMware Site Recovery Manager simplifies and centralizes the creation and ongoing management of disaster recovery plans. Site Recovery Manager turns traditional oversized disaster recovery runbooks into automated plans that are easy to manage, store and document. Additionally, Site Recovery Manager is tightly integrated with VMware Infrastructure 3 (<http://www.vmware.com/products/vi/>), so you can create, manage and update recovery plans from the same place that you manage your virtual infrastructure.

Site Recovery Manager operates on a paired site relationship, for example, one site protects another, and requires that you have access to both the protected site and the recovery site.

The protected site (primary) refers to the production environment where the virtual components that are responsible for business continuity reside. The recovery site (secondary) is the location where the protected site will fail over to in the event of a disaster.

For each pair of sites that comprise the disaster recovery relationship (protected and recovery sites), the following requirements must be met:

VMware provided components:

- The minimum number of ESX Server hosts, running 3.02, 3.5 or 3i versions needed to provide sufficient resources for recovery.
- A VirtualCenter management server running version 2.5 or later.
- Site Recovery Manager with the following installed:
 - Site Recovery Manager 1.0.
 - Java Runtime Environment version 1.6 or later, available from the Powerlink® Software Downloads and Licensing section.
 - Celerra Replicator Adapter 1.0 available from the VMware website.
- Virtual Infrastructure Client with SRM Plug-in for management of SRM service on VirtualCenter

EMC provided components:

- An EMC Celerra NS Series storage array with the following features:
 - DART 5.6.36 or later.
 - Replicator v2.0.
 - SnapSure.
 - iSCSI protocol.
 - Enough storage to host the replications and snapshots of the iSCSI LUNs hosting the VMFS datastores.

VMware ESX itself does not provide a remote replication capability and is therefore dependent on storage array-based replication tools. In order to integrate with array-based replication, each of the supported

storage vendors provides a Storage Replication Adapter (SRA), which is used by SRM to communicate required actions to the storage component LUNs. The supported actions within SRM include:

- Discovery of and association to the arrays at the protected and recovery sites.
- Discovery of datastores hosted on replicated iSCSI LUNs and their associations with the protected VMs.
- Test Failover, used to test the DR process without impacting the protected or recovery site environments.
- Failover, used to physically failover the whole or a subset of the VMware environment from the protected site to the recovery site.

The storage array-based replication software is configured independent of SRM, through the storage array element manager interface. The key element of SRM configuration is the codifying of the DR process. At the protected site, the codified process results in one or more protection groups that identify the virtual components that the company is the most concerned with for maintaining its business continuity. At the recovery site, recovery plans, encapsulations of one or more protection groups at the protected site, are created to define the failover process for DR. For example, the recovery plan “Data Center Complete” could contain “Mission Critical” and/or “Business Critical” protection groups. For each recovery plan:

- Define protection groups of associated VMs that need to be recovered together, for example “Infrastructure” (Active Directory or DNS), “Mission Critical”, “Business Critical”, etc.
- Define actions prior to a test or failover at the recovery site such as closing down or suspending low-priority VMs to free recovery resources at the recovery site.
- Define the allocation of resources and any networking changes required by the virtual switches.
- Build call outs; actions that cause test or failover process to pause and present instructions to the administrator, or specify scripts in the recovery process.
- Identify finite values of time or specific numbers of heartbeats to wait for Virtual Machines to respond after their power-on process is complete.

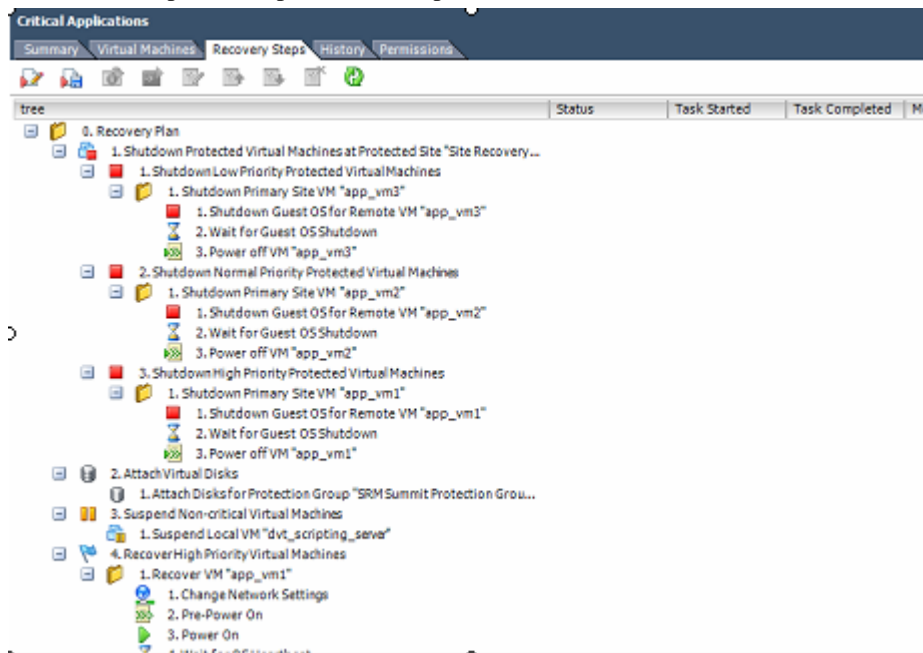


Figure 2 Sample SRM recovery plan

SRM does not automatically trigger the failover of a recovery plan, rather human intervention is required to evaluate and declare a disaster, then the recovery plan is initiated within SRM. Subsequent to storage

failover initiation, SRM executes the earlier steps and automatically manages the mapping of compute resources, network resources and the recovery of VMFS, associated datastores and ultimately virtual machines and their applications to a recovery site.

Site Recovery Manager restrictions

At the first release of SRM 1.0, the following VMware restrictions exist with SRM:

- Multi-Site DR is not supported by SRM (even if supported by the array based replication tools).
- NFS-based datastores are not supported by SRM.
- Raw Device Mapping (RDM) volumes are supported as experimental.
- SRM automated failback is not supported - in SRM V1.0, it is not possible to maintain failback recovery plans concurrently with failover recovery plans. As such, a full failover test of the VMware environment with SRM is not recommended, rather the test failover command should be used to validate the failover process. If a full failover test is required, failback can be affected manually (see [Celerra and SRM failback](#) on page 16), or by rebuilding SRM protection groups and recovery plans for the Recovery site back to the Protected site and conducting the failover process in the same manner as the full failover test.

Celerra Replicator V2 technical overview

EMC Celerra IP Replicator V2 is an extremely useful and easy-to-use feature for protecting traditional physical NAS and iSCSI environments as well as SRM enabled virtual environments. It uses a snapshot-based approach to provide an efficient data replication solution over Internet Protocol (IP) networks. Replicator V2 is a versatile feature with advanced replication capabilities that provides a solution for the disaster recovery, migration, data repurposing and data distribution use cases and applies to both Virtual and Physical environments. It supports all major IP storage protocols, multi-site replication and can be used to create copies of production data for use in backup, application testing and development.

With Celerra Replicator, users can set granular recovery point objectives (RPOs) for each of the objects being replicated, allowing business compliance service levels to be met especially when scaling a large NAS infrastructure.

Celerra Replicator can maintain up to 1,024 replication sessions per Data Mover. Administrators can protect extremely large ESX deployments or have finer granularity in segmenting their protection groups.

Meeting service level agreements (SLAs)

With Replicator V2, both the production data and replication data are accessible at all times. When defining a replication session, administrators can implement the required recovery point objective (RPO) by setting a single parameter called Time-Out-of-Sync. This parameter defines the goal time difference at any point, between the data image at the primary site and the re-startable image at the secondary location (in minutes).

Replicator V2 maintains the RPO service level through the use of an adaptive scheduling algorithm. It uses the bandwidth settings, the incoming data loads, and historical information on prior data transfers to determine the size and frequency of updates. In the unlikely event that a service level cannot be met, Replicator notifies the administrator through an event alert and adjusts its update rate to meet the previously defined RPO.

Replicator V2 is designed to be extremely robust. In the event of a power failure, the replication process will continue from where it left off once the system restarts. Should the network connection fail, Replicator V2 continues to track file system modifications through replication snaps and will catch up when connectivity is restored. The administrator is notified that service levels have been missed for the duration of the outage.

Efficient use of IP network infrastructure

Celerra Replicator V2 runs over standard IP LAN and WAN infrastructures. This simplifies configuration and management of replication sessions and allows customers to deploy remote replication with only IP networking skills.

Corporate WANs typically have limited bandwidth availability and administrators must ensure that the replication tools do not affect production network traffic by monopolizing the available network bandwidth. Replicator V2 is designed to use network bandwidth efficiently. It uses differential snapshots to send only changed block-level data over the wire. The system calculates a delta set between the two snaps, packages up the changes and sends groups of changes to the remote Celerra. By only sending changed data, Replicator V2 limits its use of valuable bandwidth. In addition, Celerra is responsible for managing the scheduling of many concurrent transfers in the most efficient manner, providing a simple hands-off management model. This ensures that RPOs are met, even when variable settings are configured, avoiding the need to set network quality of service at the object level.

Even with the efficient network utilization, some corporate networks specify hard limits on the amount of bandwidth that an application can consume. When defining interconnects between the primary and the secondary Celerra, Replicator V2 allows administrators to set bandwidth throttling parameters that specify a schedule of times, days and bandwidth limits. This capability ensures that Replicator V2 will not consume any more bandwidth than it has been allocated.

Multi-protocol support

With Replicator V2, administrators can replicate iSCSI LUNs. However, Celerra is a multi-protocol storage system that not only supports iSCSI LUNs, but CIFS and NFS file systems as well. Replicator V2 can create a copy of any network-attached file system, regardless of protocol. This capability ensures that Replicator V2 can be used to protect ESX Servers using direct-attached NFS storage and all other non-virtualized applications.

Note: At first release, SRM does not support NFS-based datastores and manual failover of NFS-based VMs is required.

Multi-site replication

Replicator V2 embraces an effective strategy for multi-site disaster recovery scenarios. Replicator not only maintains consistent copies between two Celerras as defined in SRM, but also offers additional features for one-to-N and cascading replication.

One-to-N replication allows a file system or iSCSI LUN to be replicated from a single source to up to four remote locations. This model is essentially multiple instances of 1-to-1 replication sessions but each session has its own RPO schedule and is managed separately.

Cascading or multi-hop replication allows a single file system or iSCSI LUN to be replicated from a source site (Site A) to a secondary site (Site B) and from there to a tertiary site (Site C). As with one-to-N replication, the administrator manages the tertiary hop (B to C) as an independent session.

Note: As of the first release, SRM does not support cascading and multi-hop replication.

Celerra Replicator V2 – how it works

Celerra Replicator creates an asynchronous, point-in-time copy of the production file system, VDM, or iSCSI LUN at the remote site. Local replication follows the same concept as remote replication, except that the production data and the replicas reside on the same Celerra.

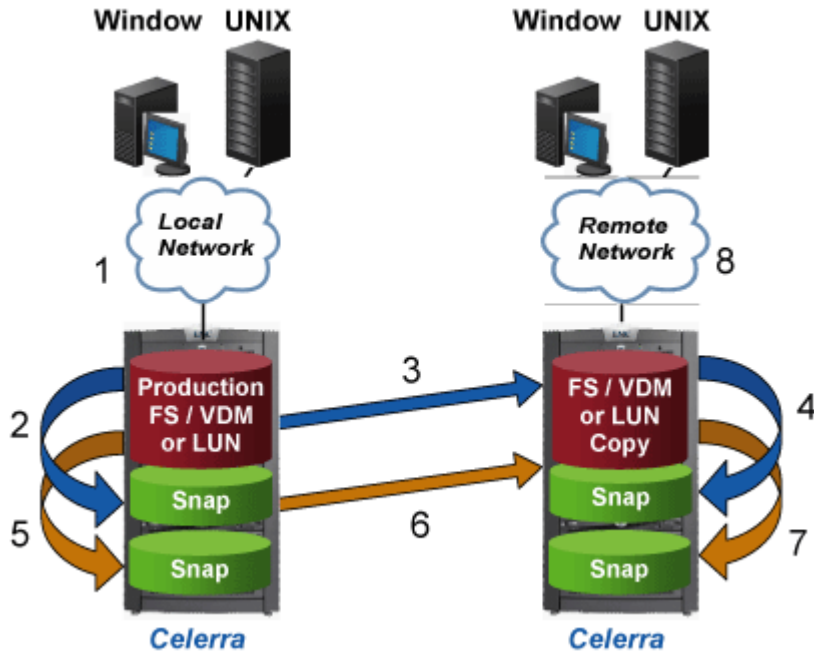


Figure 3 Celerra Replicator V2 replication process

To keep the two systems synchronized, changes made to the production data (primary) at a local site are transferred to a copy (secondary) at the remote site over an IP network. This function works automatically by setting parameters, or manually by executing commands.

The steps involved in remote replication are described below and shown in [Figure 3](#).

1. Throughout this process, network clients read and write to the production file systems, VDMs, or LUNS through a blade at the local site without interruption.
2. Replication takes an initial synchronization point with a SnapSure snap.
3. The user sets the replication parameters, which include the source/target LUN and the time-out-of-sync (RPO) value.
4. The data is either silvered over the IP network, or replicated to an intermediate Celerra that is known as a *swing-box*.
5. When the silvering process is completed, the destination object is snapped ensuring a consistent snap exists on the source and target Celerra. Replication is now established and the system can run in steady state mode. After the user sets replication parameters, the system tracks modifications to the production file system/iSCSI LUNs by internal snapshots used by Replicator V2.
6. When it decides a snap needs to be taken to ensure the data can be moved to the remote location in a timely manner to ensure the RPO service level is met, a second internal snap is taken.
7. The system calculates a delta set from the two snaps, packages up the changes and sends groups of changes to the remote Celerra, which get applied directly to the target object.
8. When the transmission is complete, the remote system takes an internal snap to ensure the two systems have a common snap for use in case of recovery.

-
- Throughout this process, network clients can read the copies at the remote site. For testing purposes, a writeable snap can be made with the remote copy.

Celerra Replicator V2 and Site Recovery Manager

This section explores an architectural overview of the solution, configuration considerations, Celerra and SRM use cases, use cases not supported by SRM, and Celerra and SRM failback.

Architectural overview

Celerra is composed of a number of key components. Celerra X-Blades are responsible for serving I/O to NAS files and iSCSI LUNs, and in addition provide the intelligence and transport mechanism for the Replicator technology. The Celerra Control Station provides a secure management interface for administration purposes. When configuring SRM, the Control Station of the Celerra hosting the protected site datastores is used to connect and authenticate with the Celerra. This connection is also used when the SRM requests an action (for example, failover or test failover) from the Celerra through the Storage Replication Adapter.

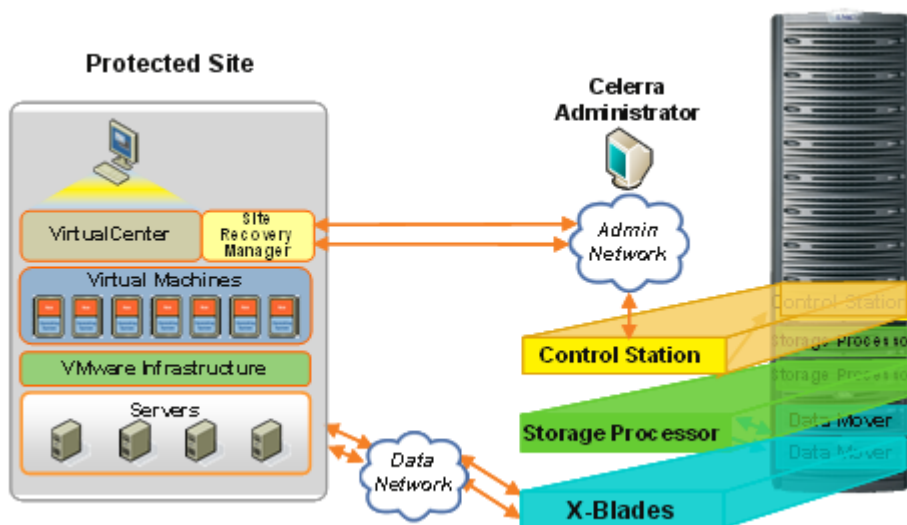


Figure 4 Celerra operation within a VMware environment

Configuration considerations

Configuration best practices for Celerra usage with VMware are covered in greater detail in other documents (specifically, *Using EMC Celerra IP Storage with VMware Infrastructure 3*) and is beyond the scope of this paper. However, the following are high-level considerations for using replication with SRM:

- I/O load and performance:** I/O loads should be analyzed carefully so that VMs with heavy I/O loads are presented from separate iSCSI LUNs.
- VMs failed over as single entities:** Since an SRM protection group is designed to protect a datastore, not VMs, VMs that need to be failed over or tested as a single entity should reside in a single datastore. Examples include: Active Directory and DNS infrastructure, Tier 1 applications, and Tier 2 applications.
- Capacity planning for the recovery site:** Administrators should consider that the available resources at the recovery site may not be equal to those available at the protected site.

Celerra and SRM use cases

This section covers a number of use cases, highlighting the value of the Celerra Replicator and SRM integration.

Disaster recovery test

SRM is integrated with the Celerra local snap capability, SnapSure. Disaster recovery tests are affected by using the test failover command (sometimes called simulated disaster recovery) against a recovery plan and/or a subset of protection groups. During this process, the production VMs at the protected site continue to run and the replication connection remains active for all the replicated iSCSI LUNs. When the test failover command is run, Site Recovery Manager issues requests for the Celerra at the recovery site to take a writeable snap using the local replication feature licensed at the recovery site. These snaps are discovered and mounted per the definitions of the recovery plan and any pre-power on scripts or callouts are executed. VMs are powered on and any post-power on scripts or callouts are executed.

The same recovery plan is used for the test as for the real failover so users can be confident that the test process is as close to a real failover as possible without actually failing over the environment. Companies realize a greater level of confidence in knowing that their users are trained on the DR process and can be assured that users can execute the process consistently and correctly each time. Users do have the ability to add a layer of test-specific customization to the workflow that is only executed during a test failover to handle scenarios where your test may have differences from your actual failover scenario. If VM power-on is successful, the Site Recovery Manager test process is complete. Should the user need to start applications and perform tests, then they are at liberty to do so. Prior to cleaning up the test environment, SRM uses a system call out to pause the simulated failover. At this point, users should verify that their test environment is consistent with the expected results; upon verification, the user acknowledges the call out and the test Failover process concludes by powering down and unregistering VMs, demoting and deleting the Celerra writeable snaps, and restarting any suspended VMs at the recovery site.

Disaster recovery

Failover is much like test failover, except rather than leveraging snaps at the recovery site while keeping the primary site running, the storage array is physically failed over to a remote location and the actual recovery site LUNs are brought online, and VMs are powered on. VMware will attempt to close down the protected site VMs if they are active when the failover command is issued, although if the protected site is destroyed, VMware will be unable to complete this task. Site Recovery Manager will not allow a VM to be active on both sites. Celerra Replicator has an adaptive mechanism that attempts to ensure recovery point objectives are met, even with varying VMware workloads, so users can be confident that the crash consistent datastores that are recovered by SRM meet their pre-defined service level specifications.

Bear in mind that whenever an SRM failover command is issued from SRM, considerations will need to be made to the process of failback, covered elsewhere in this paper.

Migration

SRM can be effectively leveraged to provide a simple migration capability for use with a major data center migration. Where VMware is extensively implemented, configuring storage level replication and building an SRM recovery plan for all the VMware assets that require migration from the old data center to the new data center allows all the values of SRM: risk mitigation due to the ability to test the migration scenario, as well as a single button, high speed failover (migrate) process through SRM. Celerra Replicator transparently supports both physical and virtual environments and so this scenario can even address the requirements of physical server migration and non-VMware NAS data, by issuing call-out scripts from the SRM recovery plan to effect any non-VMware migration actions, although these call outs will have to be manually tested. When the migration is planned, the customer can either test migration by using the test failover feature of SRM or execute the recovery plan to migrate to the new data center. All this occurs

while SRM manages and executes the call outs and scripts during the failover activity. After the failover, the old data center can be de-commissioned and the infrastructure can be re-deployed elsewhere. This migration process can also be implemented in a piecemeal process, a few applications at a time if required, whereby servers can be re-utilized more effectively between the new and old data center, merely by codifying a number of more granular SRM recovery plans.

Shared disaster recovery protection

Companies with a pervasive virtual infrastructure that do not have the resources to create a complete protection site can share the cost by collaborating with a second organization facing the same issue. In this scenario, each entity acts as both the protected site for its own data and the recovery site for the data from the other company. This solution results in a lower implementation cost and reduced operating cost for implementing the VMware automated DR solution.

Use cases not easily supported by SRM

Frequent data center switch-over

Some large sites are structured to allow a failover to a secondary site with full application, infrastructure and client switch-over. While this is possible with SRM, it is not a simple task to merely flip between the protected site and the recovery site with the SRM failover command. The reason for this is that SRM does not support failback. SRM can maintain only one relationship between protected and recovery sites and when the failover is initiated, the recovery site becomes the primary site. To leverage SRM to fail back, the relationships between this site and the new failback site needs to be rebuilt. Another option would be to maintain a manual process for failback. Read further for more details on the failback process.

Multi-site disaster recovery

While Celerra Replicator (and some of the other replication vendors) supports multi-site disaster recovery, the first release of SRM does not. The reason is because SRM supports only a single protected and recovery site relationship.

Graceful failover and reverse replication

Celerra Replicator supports the ability to coordinate the synchronization of any data at the protected site with the recovery site prior to failover. However, this option is not supported by the SRM SRA. Whenever a failover is issued, the connection is immediately severed, regardless of the state of the protected site. It is vital for an administrator to fully understand the implications of a failover decision. When the failover is a graceful switch-over, the protected site must be gracefully closed down, manually.

Celerra and SRM failback

The process of failing back a previously failed over Celerra/SRM environment, requires a manual process. This process is similar to the process automatically performed by SRM for the failover. The following steps describe the use case where the primary site is lost and no replica or data set exists at the failback site. After DR failover and the original failed site is rebuilt (data does not exist at the new failback site), there is a two-step process for Celerra side failback and VMware side failback, outlined as follows.

Celerra failback procedure (in the following steps, the primary Celerra is the Celerra at the operating failover site and the secondary Celerra is the Celerra at the new failback site)

1. Enable all requisite licenses on the secondary Celerra (iSCSI, SnapSure and Replicator V2).
2. Establish a trusted relationship between primary and secondary Celerras.

-
3. Create the Replicator interconnect between the primary and secondary Celerras.
 4. Create file systems, iSCSI targets and read-only iSCSI LUNs on the secondary Celerra.
 5. Perform LUN masking and start the iSCSI service on the secondary Celerra.
 6. Establish a replication session for each iSCSI LUN from the primary to secondary Celerra and allow the initial replications to complete.

NOTE: Celerra supports up to 16 concurrent replication sessions in “initial synchronization status”.

7. Reverse the direction of replication, flushing any changes from the primary to the secondary Celerra.

NOTE: For a failover where the data at the primary site is intact, the Celerra failback procedure does not require the re-creation and initial synchronization of the iSCSI LUNs from the recovery site to the protected site; it is merely the reversal and synchronization of the changes from the failover (recovery) site. The following VMware failback process is still required in this scenario.

VMware failback procedure (in the following steps, the protected site is the failback site containing the secondary Celerra in the process above and the recovery site is the original failover site containing the primary Celerra):

- At the recovery site, complete the following steps:
 1. Remove the recovery plan from SRM.
 2. Unregister the failed over VMs in the inventory view.
 3. Rescan the storage in the Site Recovery toolbar (that now has replication configured with the source in the recovery site and the target in the protected site due to the Celerra failback procedure performed above).
- At the (new) protected site, complete the following steps:
 1. Remove the appropriate protection groups from SRM.
 2. Unregister the previously failed over virtual machines.
 3. Rescan the storage for the ESX servers being failed back in SRM.
 4. Rename the failed back datastore to its original name (the replicated datastore will appear at the new protected site with an unrecognized name, e.g. “snap...xxxx” and needs to be changed).
 5. Add the protected virtual machines into the SRM inventory from the datastore browser in SRM.
 6. Re-create the protection groups as for prior to the failover.
 7. At the protected site, rebuild the recovery plans in SRM.

For more detail on these steps, see *EMC Celerra NS Series iSCSI EMC Celerra Replicator Adapter for VMware Site Recovery Manager Release Notes*. EMC is currently evaluating future mechanisms to provide automated failback.

Conclusion

The innovative technology of EMC Celerra Replicator V2 supports flexible levels of protection, without distance limitations and performance degradation. With its unique architecture, powerful data recovery features, and business-driven approach, Celerra Replicator V2 offers superior levels of local and remote data protection and business continuity to organizations running VMware ESX. Organizations implementing Celerra Replicator V2 with the VMware ESX Server are expected to see the following benefits:

- Full support for Celerra Replicator for iSCSI LUNs, integrated with VMware Site Recovery Manager
- Ability to leverage Replicator adaptive recovery point objective optimization with VMware Infrastructure optimizing recovery time objectives
- Complete support for VMware physical-to-virtual and virtual-to-virtual replication models (using RDMs that are currently experimental in ESX)
- Compatibility with other VMware technologies, including DRS, HA, Storage VMotion, and VMotion
- Supports replication between VMFS volumes as well as between RDM volumes in physical mode (experimental support from VMware)
- Simple and quick planned or unplanned failover for virtual machines and their data without distance limitation
- Out-of-band processing for replication that ensures that the performance of the ESX Server and its virtual machines are not impacted by Celerra Replicator
- Rapid and simple replication for virtual machines and their data to an alternate location and instantly accessible for disaster recovery or for recovery from logical corruption
- Ability to leverage local replication for operational or application recovery of a virtual machine while still maintaining remote replication to provide protection in case of a site-wide disaster

References

The following documents provide additional, relevant information. Access to these documents is based on your login credentials. If you do not have access to the following content, contact your EMC representative:

- *Customer Presentation: EMC Disaster Recovery with VMware Site Recovery Manager*
- *Customer Presentation: Information Infrastructure for VMware Data Protection*
- *EMC Extends Disaster Protection Capabilities for VMware Environments*
- *EMC Extends Innovation for VMware Environments*
- *Praetorian Financial Enhances Data Protection with EMC and VMware*
- *EMC and VMware: The Ultimate Disaster Recovery Solution*
- *EMC Celerra NS Series iSCSI EMC Celerra Replicator Adapter for VMware Site Recovery Manager Release Notes*
- *EMC Extends Disaster Protection Capabilities for VMware Environments*
- *EMC Extends Innovation for VMware Environments*
- *Customer Presentation: EMC Disaster Recovery with VMware Site Recovery Manager*
- *Video: EMC and VMware: The Ultimate Disaster Recovery Solution*
- *White Paper: VMware ESX Server Backup and Replication on EMC Celerra NS Series – Best Practices Planning*
- *Celerra Network Server Using Celerra Replicator (V2) – Technical Module*
- *Celerra Network Server Configuring iSCSI Targets on Celerra – Technical Module*

-
- *EMC Celerra Series EMC Celerra Replicator Adapter for VMware Site Recovery Manager Release Notes*
 - *Disaster Recovery with VMware Site Recovery Manager and EMC Celerra – Solution Overview*
 - *White Paper: Using EMC Celerra IP Storage with VMware Infrastructure 3 over iSCSI and NFS – Best Practices Planning*