EMC RECOVERPOINT CDP: CONTINUOUS DATA PROTECTION FOR OPERATIONAL RECOVERY

Applied Technology

Abstract
This white paper discusses the continuous data protection capabilities of EMC® RecoverPoint/CL, RecoverPoint/EX and RecoverPoint/SE. RecoverPoint transparently captures all changes generated by one or more applications and stores copies of these changes in a disk-based journal that is also used for point-in-time recovery. RecoverPoint’s instant recovery enables customers to quickly recover data lost due to physical or logical corruption and then restart their databases and applications.

May 2012
Copyright © 2009, 2012 EMC Corporation. All Rights Reserved.

EMC believes the information in this publication is accurate 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.

VMware and ESX are registered trademarks or trademarks of VMware, Inc. in the United States and/or other jurisdictions. All other trademarks used herein are the property of their respective owners.

Part Number h6181.3
# Table of Contents

**Executive summary** ................................................................. 4  
  Audience .................................................................................. 5  

**Data protection challenges** ....................................................... 5  
  Protect and recover .................................................................. 6  
  Data protection options .............................................................. 7  
  RecoverPoint CDP is complementary to backup technologies .......... 8  
  Benefits of RecoverPoint CDP .................................................... 9  

**EMC RecoverPoint operations** ................................................... 11  
  Managing RecoverPoint .............................................................. 11  
  System architecture .................................................................. 12  
   RecoverPoint appliance .......................................................... 13  
   Splitter .................................................................................... 13  
   Journal .................................................................................... 14  
   Application servers .................................................................. 14  
   Heterogeneous storage arrays .................................................. 15  
   SAN ......................................................................................... 15  
  Preserving data integrity during failure scenarios ......................... 16  
   Managing RecoverPoint appliance failure ................................ 16  
   Managing lost connectivity to source volumes ......................... 16  
   Managing lost connectivity to target volumes ........................... 17  
   Managing host or write splitter failures .................................... 17  

**RecoverPoint CDP use case scenarios** ....................................... 17  
  Case 1: Operational recovery ................................................... 17  
  Case 2: Repurposing ................................................................. 18  
  Case 3: Data migration .............................................................. 20  
  Case 4: Individual file or folder level recovery ............................ 21  

**Conclusion** ............................................................................. 22  

**References** ............................................................................. 22
Executive summary

Today's businesses are facing an ever-increasing amount of data that can overwhelm existing storage management solutions. The foundation of today's business operations is data. Data is created, stored, massaged, transformed, mined, modified, deleted, and utilized continuously across an organization. This data is present on many systems, on hard disks drives in laptops, in flash memory on USB key drives, and in large server farms sharing array-based storage on a SAN. What does all this data have in common? All of this data shares the same high cost of managing it, sharing it, and protecting it from damage and corruption.

Data protection is no longer as simple as copying yesterday's changed files to tape. Explosive technology growth has resulted in complex environments with dedicated storage area networks housing an assortment of intelligent storage devices. Businesses are looking for ways to ensure timely business restart from data corruption or system failure.

Critical data changes occur throughout the day, and to protect this data customers are frequently turning to new technology such as continuous data protection (CDP). Often touted for its ability to provide instant restoration and granular recovery, CDP also solves the backup window challenge by eliminating the backup window itself by enabling the creation of backup copies anytime, day or night, without affecting online operations.

The EMC® RecoverPoint family provides local and remote protection with CDP-based recovery. Local replication is sold as RecoverPoint CDP (continuous data protection), remote replication is sold as RecoverPoint CRR (continuous remote replication), and when both local and remote replication is used it is called concurrent local and remote (CLR) data protection. All of these are solutions that can dramatically improve the operational and disaster recovery capabilities of your environment.

RecoverPoint CDP intercepts every block written by an application that flows across the SAN and saves a copy of the block in a disk-based journal, which improves application recovery by enabling a DVR-like point-in-time recovery of data. Unlike host-based or array-based replication solutions, RecoverPoint is an appliance-based, out-of-band data protection solution designed to protect all SAN-based production data regardless of the host or storage array. RecoverPoint enables customers to simplify their point-in-time data recovery while centrally managing their recovery processes.

RecoverPoint is available in three versions, RecoverPoint/SE, RecoverPoint/EX and RecoverPoint/CL. RecoverPoint/SE is the entry offering that simplifies replication and continuous data protection for a single VNX™ series, CLARiiON® CX™, CX3 UltraScale™, CX4, or Celerra® unified storage array or between two VNX series, CLARiiON, or Celerra unified storage arrays. RecoverPoint/EX is the offering that simplifies continuous data protection and replication by using array-based write splitting for the Symmetrix VMAX 20K, Symmetrix 40K with Enginuity™ 5876, Symmetrix VMAX 10K with Enginuity 5875 or 5876, VPLEX Local and VPLEX Metro with GeoSyncrony 5.1, VNX™ series, Celerra® unified, and CLARiiON® CX3 UltraScale™
and CX4 series of arrays. RecoverPoint/CL is the full-featured offering that adds support for intelligent fabrics, heterogeneous servers, and heterogeneous storage platforms. RecoverPoint/SE can be upgraded to RecoverPoint. Except where noted, all references to RecoverPoint refer to RecoverPoint/CL, RecoverPoint/EX and RecoverPoint/SE. This white paper focuses on RecoverPoint’s use of CDP; additional EMC white papers are available covering the other capabilities of RecoverPoint.

This white paper is designed to give the reader an introduction to CDP concepts, explore how EMC RecoverPoint CDP operates, and give some use-case scenarios for RecoverPoint CDP.

**Audience**

This white paper is intended for systems integrators, systems administrators, and members of the professional services community. This paper serves as an overview to the RecoverPoint CDP software and discusses key software functions and features. The References section provides related information, including administrator guides, proven solutions, and white papers.

**Data protection challenges**

Protecting data is the key to a successful businesses operation. However, implementing real-time application recovery for critical data is not a simple proposition. The first step, even before analyzing data protection solutions, is to understand the current business processes and develop a clear set of objectives and plans that reflect what is required to safeguard against any disaster that could make the data at the primary site unavailable.

An evaluation of the data utilized by your business applications must be completed as part of designing a protection solution. The reason is that if the production volumes go offline due to a disaster, and your business processes must be restarted from a recovered image, how much delay and data loss can you tolerate before you are unable to restart production? According to the U.S. National Archives and Records Administration in Washington, D.C., 93 percent of companies that lost their corporate data for 10 days or more due to a disaster filed for bankruptcy within one year of the disaster. Of those companies, 50 percent filed for bankruptcy immediately. A PricewaterhouseCoopers survey calculated that a single incident of data loss costs businesses an average of $10,000 per hour.

There are two general principles that govern all recovery: recovery point objective (RPO) and return to operations (RTO). RPO defines how much data you are willing to lose when you recover data. For example, if you back up daily, the RPO would be 24 hours, which is the maximum amount of data loss that could occur between backup images. RTO defines the amount of time it takes to restart affected business applications from the recovered data. For example, once the data is recovered, it is necessary to restart the business applications based on the recovered data. This usually involves checking the recovered data for consistency, performing applicable
log processing, starting the application with the recovered data, and then re-creating any missing data due to your RPO.

When evaluating a data protection solution, it is important to look at all of the capabilities of the solution. One of the first challenges to examine is the overall performance of the data protection solution. For example, will the solution handle highly transactional applications? Does the solution support the different RPO and RTO requirements for the different applications that are used? The recovery time of various solutions will also vary. As an example, recovery from a set of backup tapes may be on the order of hours to days but recovery from a host-based or array-based snapshot may take less time, on the order of minutes or days. In most use cases, complete systems and application recovery using RecoverPoint will take only minutes.

A second area to evaluate is the costs inherent to the solution. These are not the fixed cost of the solution, but cover other items such as additional costs required to manage multiple backups or snapshot images. Also important to evaluate is the cost for data loss and application downtime. If there is a critical business application, such as real-time financial transactions, the business cannot afford to lose any data in the event of a disaster and it may be very important that the application comes back online in a matter of seconds without any noticeable impact to the end users. For other applications, a delay of a few minutes or hours may be tolerable. RecoverPoint dramatically reduces the amount of disk storage required for data protection and can support a zero RPO with its continuous data protection capability.

It is also important to evaluate the operational management impact of a data protection solution. Choosing the right recovery point is important to reduce the RTO. If you select an application-aware recovery point, you may reduce considerably the data consistency checking required by the application. If multiple point solutions are used, such as in a federated database environment, it is important to choose a product that can ensure consistent RPO and RTO across all the applications. Finally, while some solutions are ideal for data protection, they may offer little in the way of application-aware integration or may be challenged when it comes to supporting a test and development environment. RecoverPoint has application-aware integration with specific integration points for common Microsoft applications such as Exchange, SharePoint, and SQL Server as well as for Oracle Database environments. A simple stand-alone as well as a complex federated environment can easily be supported with RecoverPoint. Finally, RecoverPoint is also integrated with other EMC software products, including NetWorker® and Replication Manager, and it offers a rich application programming interface that can be used to integrate RecoverPoint into existing customer configurations.

**Protect and recover**

Companies are driven to develop operational recovery capabilities that can protect their email, business applications, images, and database environments. Using RecoverPoint, customers can reduce their RPO to zero, ensuring no loss of data. With RecoverPoint’s instant recovery capability, their RTO can be reduced to seconds,
minutes, or hours as compared to the hours, days, or weeks of alternative solutions. RecoverPoint enables the ability to have an application-consistent or crash-consistent recovery point with the granularity of a single write. Finally, RecoverPoint enables the federation of multiple critical applications across any supported storage infrastructure enabling a true application environment restart.

**Data protection options**

Mission-critical applications usually require recovery aligned to the available RPO and RTO. For example, within any customer environment, there may be multiple applications, each with different data protection objectives for RPO and RTO. Common solutions for some of these applications include:

- Daily operational backups for 24-hour operation protection with weekly full backups for longer-term archive
- Using periodic disk-based snapshots with remote replication to protect data in event of disaster in a local site when the business needs to fail over to a remote location and be up and running in a short timeframe
- Using synchronous or asynchronous replication to enable quick recovery in the event of physical disk loss, particularly in test and development environments

All of these solutions have challenges. The nightly backup may fail, or a data loss may occur 12 hours into the new backup period. If data loss occurs 12 hours into the new backup period, any new data created since the last backup is lost, since the system can be rolled back only to the last recovery point.

**Daily Backup:** Recovery point every 24 hours

**Snapshot:** Recovery point every 3 hours

**Sync and Async Replication:** Latest image replicated, but susceptible to logical corruption

**Figure 1. Data protection options**

Disk-based snapshots provide a smaller recovery point window, usually as short as three hours; however a gap still exists in recovery between snapshots. Synchronous and asynchronous replication ensures that there is no recovery window; however, both the production and mirrored data can be impacted by logical corruption.
Figure 2. Continuous data protection with application bookmarks

There is an approach, called continuous data protection, which is of value to any type of environment that has short recovery objectives, including database or messaging applications such as Microsoft SQL Server or Exchange, Oracle, or SAP.

Continuous data protection uses a journal-based architecture that captures time-indexed recovery points, taking small aperture snapshots as small as a single write. Using this journal, CDP can ensure data recovery back to any point in time. Users can bookmark recovery points to recover back to specific points in time, such as the close of a quarter or a pre-patch state. It is also possible to create application-aware I/O bookmarks.

Rather than trying to “force” a particular technology to protect specific environments, customers now have more choice in applying the most appropriate protection technology to their specific RPO and RTO objectives.

**RecoverPoint CDP is complementary to backup technologies**

Companies that have invested in tape or disk-based backup technology can protect those investments when they implement RecoverPoint. RecoverPoint complements existing backup products, such as NetWorker, by providing an online, near instantaneous recovery of recently changed information. RecoverPoint can recover an any-point-in-time image of the data volumes as long as the data resides in its journal. RecoverPoint extends the journal data by consolidating images in the journal as they age into daily, weekly or monthly recovery points. This enables customers to utilize RecoverPoint’s recovery capabilities to roll back data to points in time when data was last in a consistent recoverable state.

One use for this image would be to use the CDP replica as the source volume for a traditional backup operation. For example, using RecoverPoint the customer can roll back the data to any point in time and present it to a backup or media server in the SAN. This server would then back up the data and then have RecoverPoint resynchronize the CDP replica to the (production) source data once the backup is complete.

Additionally you can use RecoverPoint to provide a richer backup environment; for example, you may want to audit the data before the backup is performed. This is usually difficult with live data but can be easily implemented using RecoverPoint, where the latest image can be presented to the auditing server for checking and then presented to a different server where it can be backed up.
Benefits of RecoverPoint CDP

There are several features that RecoverPoint brings to a data protection environment. These include the following:

• **Improved reliability**

  With RecoverPoint, the management GUI can be used to quickly determine that RecoverPoint has protected the application data. Additionally, the GUI has wizards that utilize RecoverPoint’s instant recovery capability, enabling quick access to protected application data, such as for quickly testing (“fire drill”) the data necessary for application recovery.

• **Reduces operational costs**

  RecoverPoint saves changes to application data in the RecoverPoint journal. This journal can be on a different tier of storage or on a different array than the production volumes. The journal is compressed and RecoverPoint can also consolidate recovery points. All of these capabilities enable RecoverPoint to retain more recovery points than equivalent array, host, or tape-based technologies at the least cost.

• **Data centralization**

  Today’s customers are refusing to be locked into a single vendor for their storage and server solution. Customers typically have more than a single application in their data center. RecoverPoint provides protection and recovery for data regardless of the server or storage that contains the application. RecoverPoint supports all common open system platforms including many flavors of UNIX and Windows, and server virtualization solutions such as Microsoft Hyper-V and VMware. RecoverPoint supports EMC Symmetrix® and EMC CLARiiON storage arrays as well as Fibre Channel storage from vendors such as Hitachi, HP, HDS, IBM, and NetApp.

• **Avoid tape media loss**

  Backup tapes are a convenient and widely used method for backing up files. Yet when backup tapes contain sensitive information, losing control of the tapes—either through loss or theft—can put a company in violation of federal or state regulations. It is also true that a tape in transit is especially vulnerable to loss or theft. The loss of a tape can potentially compromise thousands to millions of customers or clients and subject the company to lawsuits as well. By using RecoverPoint, a company can avoid the risk of tape loss or theft.

• **Instant recovery**

  Customers see value to repurposing protected data, such as for backup acceleration, disaster testing, and cloning of production environments for development and test environments. Using RecoverPoint, application data can be recovered to an exact point in a matter of seconds. Once the data is recovered, it can then be manipulated without impact to production applications or the recovered copy. Any changes made to the data are only temporary, and will be
automatically discarded by RecoverPoint when access to the data is no longer required.

- **Zero RPO**

  RecoverPoint captures every change to production volumes and logs those changes for later recovery. The RecoverPoint CDP replica volumes are identical to the production volumes, which ensure a complete restoration with no data loss. Additionally, the RecoverPoint journal enables DVR-like any-point-in-time recovery, allowing you to restore data to a point in time before a corruption event.

- **Integrated consolidation enables RecoverPoint to augment existing backup technologies**

  The user can set a policy that tells RecoverPoint to maintain a zero RPO for a set number of days and then consolidate the data to a single recovery point per day and maintain that for a number of days, and then to weekly recovery points and after another set of weeks move to a monthly recovery point. RecoverPoint will retain a minimum of 12 hours of zero RPO recovery points, and given sufficient journal space, an unlimited number of zero RPO recovery points can be retained. A tape-based backup that is used on a daily basis risks data loss since it only gives a 24-hour RPO. Using RecoverPoint in conjunction with your tape-based backup eliminates worries about the RPOs, and helps eliminate worries about your current backup process.

- **Application-consistent recovery**

  Application recovery is becoming more complex and time-consuming. When considering data protection solutions, ensure that you select a solution that includes application consistency with the CDP recovery process. RecoverPoint includes utilities that integrate with Microsoft Exchange Server through VSS and Microsoft SQL Server through VDI. Additionally, RecoverPoint ships with application programming and scripting interfaces that enable business-specific processes to be integrated with RecoverPoint.

- **Non-invasive to the application or server that is being protected**

  RecoverPoint is designed to minimize any impact to a production's I/O throughput or CPU load. The RecoverPoint write splitter (which can reside on the host, fabric, or CLARiiON array) intercepts write operations to the target volume at “write-speed,” which ensures that there is no write performance degradation seen by the application or its storage arrays.

- **Out-of-band solution**

  RecoverPoint is an appliance-based, out-of-band product. It resides in the SAN, but I/Os generated by the application server do not flow though the RecoverPoint appliance. Therefore, an unlikely failure of the RecoverPoint appliance will not affect your application’s access to storage.

- **Scalable, reliable platform**
RecoverPoint is built on top of an industry-standard server platform that is always deployed in a clustered configuration. Each RecoverPoint appliance cooperates in managing the data protection and recovery processes for the site. The appliance is built on a high-availability architecture that ensures that the failure of a single appliance will not affect ongoing data protection operations. Additionally, RecoverPoint performance can be scaled by adding additional appliances to the cluster without disruption of ongoing operations.

- **Supports business policies and service level agreements**
  Companies assign different values to their different applications. RecoverPoint CDP supports policies that allow differing RPOs and RTOs on a per-application basis.

- **Integrated local and remote business continuity solution**
  RecoverPoint provides both local and remote protection using the same software and appliance architecture. A customer who starts with RecoverPoint CDP can extend their protection by adding RecoverPoint CRR for remote replication and CLR for concurrent local and remote data protection.

**EMC RecoverPoint operations**
RecoverPoint CDP is designed to replicate every change at a block level on one or more SAN volumes (source volumes) residing in one or more storage arrays. It allows the replicated volumes (target volumes) to reside in one or more heterogeneous storage arrays.

RecoverPoint maintains a transactionally consistent journal for each consistency group defined within a RecoverPoint system. This journal allows convenient rollback to any point in time, enabling instantaneous recovery for application environments. Additionally, RecoverPoint also provides application-aware integration that provides intelligent application-consistent recovery points for multiple third-party applications such as Microsoft Exchange and Microsoft SQL Server.

**Managing RecoverPoint**
RecoverPoint is managed using a Java-based graphical user interface (GUI), called the RecoverPoint Management Application. Additionally, RecoverPoint can be managed through a command line interface and can be controlled though EMC Replication Manager and EMC NetWorker. The following figure shows the RecoverPoint Management Application. This GUI is platform-independent and can be launched from any Windows-based Internet browser. You can view the health and status of RecoverPoint at a single glance, while quickly obtaining the protection status by clicking on a consistency group of interest.
Figure 3. RecoverPoint Management Application

To automate RecoverPoint operations, use the RecoverPoint CLI. This CLI can be invoked in an *interactive* mode, or in a programmatic mode. Interactive mode offers automatic help, command option prompting, and automatic command completion. Programmatic mode is used inside of scripts, where the script input may be generated and the script may further process output.

System architecture

The specific components of EMC RecoverPoint CDP are shown in Figure 4. This configuration shows the source and target volumes can reside on different disk subsystems. The RecoverPoint storage configuration is flexible and supports storage arrays such as EMC Symmetrix and EMC CLARiiON, as well as third-party arrays from multiple vendors such as Hitachi, IBM, NetApp, and others. The journal storage used for RecoverPoint can reside on the same storage array as the source or target volumes or it can reside on a separate array.
**Figure 4. EMC RecoverPoint architecture**

**RecoverPoint appliance**

The RecoverPoint appliance (RPA) runs the EMC RecoverPoint software inside a secure environment built from an industry-standard server platform. An RPA manages all aspects of data protection for a storage group, including capturing changes, maintaining the images in the journal volumes, and performing image recovery. Moreover, one appliance can manage multiple storage groups, each with differing policies.

There are at least two active RPAs that constitute a RecoverPoint CDP cluster and at least four active RPAs that constitute a RecoverPoint CRR cluster. The individual appliances in a cluster are called a node; so for local replication you have between two and eight nodes in the cluster and for remote replication you have four to 16 nodes in a cluster. All RPAs in a cluster have identical functionality. In normal operation, all RPAs are active all of the time. Consequently, if one of the RPAs in a cluster goes down, EMC RecoverPoint supports immediate *switchover* of the functions of that appliance to one or more of the remaining RPAs.

**Splitter**

The RecoverPoint write splitter driver is system software that can reside on the application host, on the array, or in the SAN fabric (through special hardware developed by Brocade and Cisco). The function of the write splitter driver is to “split” or “mirror” an application’s write traffic, so that the written data is sent to both the production storage volumes and to the RPA.

With the host-based write splitter, the write splitter driver (KDriver) is low-level driver software that is installed on all hosts that access the volumes that contain the data to be protected by RecoverPoint. The KDriver carries out this activity efficiently, with little perceptible impact on host performance, since all CPU-intensive processing necessary for data protection is performed by the RPA.
With the array-based write splitter, the write splitting is performed inside the Symmetrix VMAX 10K, Symmetrix VMAX 20K, Symmetrix VMAX 40K, VPLEX, VNX or CLARiiON series array. The RecoverPoint Symmetrix write splitter runs on Enginuity 5876 for Symmetrix VMAX 20K and Symmetrix VMAX 40K and on Enginuity 5875 or 5876 for Symmetrix VMAX 10K, the RecoverPoint VNX/CLARiiON write splitter is supported on all VNX series arrays and the RecoverPoint VPLEX write splitter runs on Gensyncrony 5.1 on VPLEX Local and VPLEX Metro. The VNX/CLARiiON write splitter is also supported on the CLARiiON CX4-120, CX4-240, CX4-480, and CX4-960 with FLARE® 28 or 29, and on the CLARiiON CX3-10, CX3-20, CX3-40, and CX3-80 with FLARE 26 or later patch code, as well as for the block volumes presented by the backing CLARiiON CX4 array on a Celerra unified NS-120, NS-480, and NS-960 or presented by the backing CLARiiON CX3 array on a Celerra unified NS20, NS40, and NS80.

Array write splitting creates an integrated offering for the Symmetrix VMAX, VPLEX, VNX, CLARiiON CX4, and CLARiiON CX3 series arrays in the heterogeneous enterprise that offers unprecedented levels of functionality, manageability, flexibility, and performance. This enables users to replicate their resident volume data in heterogeneous server environments, across any distance, with zero impact to the primary I/O path or applications. The array-based write splitter is operating system-independent and is recommended for all RecoverPoint configurations.

Alternatively, the splitter function can be carried out inside an intelligent fabric, such as provided by the Brocade FAP protocol on the Connectrix® AP-7600B or with a Connectrix MDS 9000 director with the SSM module or the MDS 18/4 Multi-Services blade and the Cisco SANTap protocol. When the write splitter function is performed using an intelligent fabric or an array-based splitter, a host-based KDriver is not required.

Journal

Journal volumes hold data that is waiting to be distributed to target volumes and also retain copies of the data previously distributed to the target volumes. Each consistency group has its own journal volumes that allow for differing retention periods across consistency groups. Each consistency group has two journal volumes, one assigned to the target volumes, and one assigned to the source volumes. Journal volumes reside on both in order to support production failover from the current active source volume to the target volume.

Storage efficiency is realized in the history journal by retaining only the changes between journal entries. Additionally, the journal volume is also compressed, resulting in even more storage savings. Source and target images are always consistent upon completing the distribution of each write change.

Application servers

RecoverPoint is an out-of-band, block-level solution. As such, it is agnostic to the server operating system, file system, and volume management environment. RecoverPoint is designed to support heterogeneous server platforms, including many
of the common Linux, UNIX, and Windows operating systems. Additionally, RecoverPoint also supports Microsoft Hyper-V and VMware® ESX® for protection of both the server and its virtual machines.

**Heterogeneous storage arrays**

Only EMC provides the breadth and depth of choices for local data protection that is shown in Figure 5. RecoverPoint provides data protection with guaranteed consistency across a data center's heterogeneous server, network, and storage environment, protecting and optimizing an organization's valuable infrastructure environment. This infrastructure can be used to protect critical application environments running on more expensive storage. Support for multiple arrays per site is not supported with RecoverPoint/SE. Support for non-EMC arrays is not supported with RecoverPoint/EX and RecoverPoint/SE.

Within a single SAN, RecoverPoint can replicate and protect volumes regardless of the source or destination arrays. As the next figure shows, RecoverPoint supports a variety of array configurations within the same site.

![Diagram of array configurations](image)

**Figure 5. Current RecoverPoint CDP deployment options for a single site**

**SAN**

RecoverPoint configurations are required to have each RPA connected to the host and storage through separate (A/B) SANs or fabrics. Each RPA has four Fibre Channel ports. For a minimum two-node RecoverPoint cluster, two Fibre Channel ports from both appliances will be connected to the A-fabric, and the other two FC ports from both appliances will connect to the B-fabric. Hosts must have multipath capability,
which can be supplied by EMC PowerPath® or other qualified multipathing products, and any storage array should have its LUNs presented though at least two ports on the array.

RecoverPoint supports any of the Layer-2 SAN environments present on the EMC Support Matrix. This includes SAN infrastructures based on Brocade, Cisco, and McDATA configurations. For SANs based on the Brocade and Cisco director-class switches, RecoverPoint also supports intelligent fabric splitting provided by the Brocade FAP protocol and the Cisco SANTap service. The FAP protocol is supported in the Connectrix AP-7600B switch. The SANTap service is supported on the Storage System Module or 18/4 Multi-Protocol blades.

**Preserving data integrity during failure scenarios**

RecoverPoint is designed to preserve data integrity in the event of a variety of failure scenarios. RecoverPoint leverages consistency groups that enable it to:

- Preserve the data integrity and consistency across all members of the consistency group
- Minimize the amount of time that data is unavailable to the end users

To maintain data consistency across replication pairs in a consistency group, RecoverPoint software protects data integrity in the event of the following failures:

- RecoverPoint appliance failure
- Lost connectivity to one or more source volumes
- Lost connectivity to one or more target volumes
- Host connectivity failure or host splitter failure

**Managing RecoverPoint appliance failure**

A minimum of two RecoverPoint appliances are required for RecoverPoint CDP. EMC does not support single-appliance RecoverPoint configurations. RecoverPoint appliances operate as a highly coupled, active/active, shared nothing cluster. The failure of a RecoverPoint appliance will be detected and its workload will be automatically distributed among the remaining RecoverPoint appliances without data loss.

When the appliance fault is corrected or a replacement appliance is brought online, the workload will be automatically returned to the original appliance. This capability allows for a nondisruptive upgrade of the RecoverPoint appliance. When an appliance is taken offline for maintenance the remaining appliance will take over its operations. When the appliance maintenance is completed, it is added back into the RecoverPoint cluster and will resume its original operations.

**Managing lost connectivity to source volumes**

A RecoverPoint consistency group will maintain data consistency across multiple volumes that reside on one or more local storage systems. If one of the volumes on
the local system becomes unresponsive the attached splitter will detect this and will notify one of the RPAs. RecoverPoint will “pause” the protection process for the affected consistency group, leaving the replicated volumes in a fully consistent and recoverable state. The user can either correct the fault or remove the affected replication pair from the consistency group so that other replication pairs can update successfully. Once the failure is resolved, the administrator can restart the replication process.

Managing lost connectivity to target volumes

A RecoverPoint consistency group will maintain data consistency across multiple volumes that can reside on one or more local storage systems. If one of the target volumes fails, then CDP protection for all of the volumes in the storage group will be paused until the volume is repaired. If connectivity to the SAN for a local host fails, and a host-based splitter is being used, then the host’s volumes cannot be monitored for changes. In both cases, the consistency of the replicated volumes will be unaffected, and CDP operations will continue for the remaining replication pairs in the consistency group. When connectivity is restored, the affected target volumes will be automatically resynchronized with the host volumes.

Managing host or write splitter failures

If connectivity to a local host fails, and a host-based write splitter is being used, then volumes accessible by the failed host will not be monitored for changes. However, the consistency of all of the CDP target volumes will be unaffected, and CDP operations will continue for the remaining volumes in the consistency group. If the host write splitter crashes, then the host goes into an unmanageable state, and the RecoverPoint appliance can be configured to pause the CDP process for all members of the consistency group or continue protection for the unaffected volumes.

RecoverPoint CDP use case scenarios

This section covers several use case scenarios for using EMC RecoverPoint CDP for local business continuance.

Case 1: Operational recovery

RecoverPoint enables a simplification of the backup, restoration, and disaster recovery for many applications, such as Microsoft Exchange. By utilizing RecoverPoint consistency groups, all of the volumes associated with Exchange, including the Exchange Storage Group Volumes and Exchange Log Volumes, are managed with the same policy and their recovery information is stored in the same journal. Additionally, RecoverPoint supports the Microsoft Volume Shadow Copy Services (VSS) APIs to ensure that there is a supportable recovery point for Exchange. The flexibility of RecoverPoint allows either an individual Exchange storage group to be protected or for the entire Exchange server to be protected.
Using RecoverPoint ensures that there is no disruption to the production database when making replica copies of the production volumes. Additionally, all volumes are fully consistent, with write order fidelity maintained at all times. This ensures that the replica volumes can be used for a variety of purposes, such as fast mailbox recovery, backup to tape with no impact to production, and compliance readiness testing.

The RecoverPoint journal, along with RecoverPoint’s virtual access, enables near instantaneous recovery of critical data at previous points in time without affecting data protection for the production volumes. This ensures that a company will be able to maintain applications in case of a local disaster that corrupts the production volumes.

**Case 2: Repurposing**

RecoverPoint capabilities can be used to quickly repurpose the production copy for uses such as backup or cloning of a point-in-time image of an application suite while the production systems remain online. This repurposed image can be used for a variety of purposes, such as backup to tape, data mining, development and testing of product updates, or compliance.

If multiple applications are involved, then RecoverPoint makes it simple to create a group set that periodically tags each consistency group’s image with the same bookmark. When it is time to clone the suite, each consistency group would be recovered back to the same exact point in time using the RecoverPoint recover browser and selecting the identical bookmark for each storage group. This is shown in Figure 7.
Figure 7. Using group sets for application suite repurposing

Once the images have been recovered back to the appropriate point in time, the application would be considered cloned and database consistency and integrity are ensured. At this point the volumes can be backed up, or mounted on the development and test systems for reads, writes, and changes or they could be transported to a different system using traditional SAN- or WAN-based tools.

Application data repurposing is a natural extension of RecoverPoint. Application production will not be impacted by data repurposing; instead an array snapshot is performed on a RecoverPoint point-in-time image. The clone that results is a concurrent, consistent copy of the production data. Finally, RecoverPoint’s ability to recover to any point in time means that the clone can be refreshed at any point with data from earlier or later points in time, without impact to the production application.

Once the volumes are rolled back to a specific point in time, the image can be used by array-specific applications such as EMC’s TimeFinder® or SnapView™ as shown in Figure 8.
Case 3: Data migration

RecoverPoint can be used to provide an application and host independent data migration between storage arrays, to redeploy storage volumes, or to consolidate from heterogeneous storage arrays. RecoverPoint consistency groups ensure that the data at the target volumes represents an identical copy of the production volumes. Additionally, RecoverPoint will perform both the initial copy as well as maintain the synchronization between the source and target volumes using SAN resources without impacting the production application.

By keeping the source and target volumes continuously synchronized, the migration of the application server to the new storage can be performed during prescheduled application maintenance windows. Finally, RecoverPoint CDP adds a second level of protection in case of a migration or storage failure by having the target volume and the CDP journal available as a real-time backup of the production data. Using RecoverPoint to migrate data between arrays will speed the migration process and maximize application availability when compared to other solutions.

Figure 8. Using TimeFinder with RecoverPoint CDP cloning

Figure 9. Deployment scenarios
Case 4: Individual file or folder level recovery

Data corruption of a file or folder at the production site may occur as a result of human error, a rolling disaster, or a machine failure. RecoverPoint protects data at the block level, recording every write into the history journal. Since it is a block-level product, it is unaware of the file system or volume contents so the administrator must use a recovery server to identify the appropriate recovery point and extract the files from the recovered image and move them back to the production server.

It may be difficult to determine the exact moment at which data corruption began. RecoverPoint's recovery wizards make it easy to test recovery at different points in time until you can identify the latest one that was created prior to the data corruption. Once identified, the recovered file or folder can be copied back to the source site using external means (such as FTP, copying it to an external media device, or any other means available).

One of the benefits of using RecoverPoint is that the production data continues to be protected during the recovery, and the recovery process does not affect normal application operations. To recover from data corruption the following steps would be followed:

1. Access an image that is near the time at which you believe the data corruption first occurred, as shown in Figure 10.

![Image](image_url)

**Figure 10. Selecting the image for recovery**

2. Using RecoverPoint virtual access recovery, the volume(s) will be instantly available and visible in the SAN. At the recovery server mount each volume that is in the consistency group.

3. On the recovery server the user would test and verify that the data is the correct version and that it is not corrupted on this image. If the data is still corrupted then recover and test an earlier version until you are able to identify a valid image that occurred shortly before the beginning of the corruption on the production volume.
4. After recovering a valid image, copy the recovered data to the production system. Mounting a share from the production server and copying the files and folders to the share can do this. Alternatively, a removable media device, such as a USB key, can be used to transport the files to the production server.

5. Once the files are copied over, you can shut down the recovery server and then restart the replication between the source and target volumes. Since virtual access was used, both the source and target volumes remained synchronized during the recovery process.

Conclusion

EMC RecoverPoint offers both continuous data protection and continuous remote replication products. RecoverPoint CDP provides continuous protection for local SAN-attached storage. Supporting a DVR-like any-point-in-time recovery capability, RecoverPoint CDP is designed to protect your critical business processes and improve their operational recovery with minimal impact to your production environment. With support for consistency groups, RecoverPoint CDP provides true CDP capability with no data loss and full write order consistency for protected volumes that can span multiple heterogeneous storage systems and servers.

References

More information on EMC RecoverPoint can be found at the RecoverPoint page on EMC.com and in the following documents on the EMC Powerlink® website:

- *Introduction to EMC RecoverPoint 3.5: New Features and Functions — Applied Technology* white paper
- *EMC RecoverPoint Administrator’s Guide* (Powerlink only)
- *EMC RecoverPoint Family Overview — A Detailed Review* white paper
- *Improving Microsoft Exchange Server Recovery with EMC RecoverPoint — Applied Technology* white paper
- *Using EMC RecoverPoint Concurrent Local and Remote for Operational and Disaster Recovery — Applied Technology* white paper