Mendocino SOFTWARE

The RecoveryONE Solution

Architecture Guide

Product Version 1.0
January 3, 2006

Part Number: ARGD-01B-010-C00

TABLE OF CONTENTS

About this Guide

The RecoveryONE™ Recovery Management Solution Architecture Guide contains the following information:

- The Introduction provides an overview of the RecoveryONE Recovery Management Solution
- Figure 1-1 provides a visual overview of the RecoveryONE architecture components
- Table 1 provides terms and definitions of the RecoveryONE architecture (and functionality) components
- The Data Handling Components Details section describes the relationship between the components involved in data collection, data storage, and data movement
- The Data Recovery Components Details section describes the relationship between the components involved in data recovery activities

Typographic Conventions

The following table shows the typographic conventions used in this guide.

Convention	Usage	Examples
`monospace`	Computer output, file, and directory names; equations, software elements, such as command names, options, and parameters	`start_prot -rs host1 sales`
`monospace` `(bold)`	User input	Enter **`restore`** to continue.
`monospace``(italic)`	Variable, user-specified name	`start_prot -rs` *`<recovery_server_name>`* *`<protection_set_name>`*
`bold`	New terms; graphical user interface menu choices, fields, and button names	On the Protected Server, all logs in the journal are **redo logs**. In the Custom Setup panel, click **OK**.
italic	Emphasis	The internal rollback window is *longer* than the business cycle.
#	Superuser prompt	`# pkgadd –d /path`
[ ]	Brackets indicate an optional argument	`ls [-a]`
|	A vertical bar separates mutually exclusive arguments	`mount [suid | nosuid]`

Introduction

The RecoveryONE™ Recovery Management Solution provides a quick, reliable, and intuitive system for continuous data protection (CDP) and data recovery management. Recovery time from corruption or disaster, to the optimal recovery point, is reduced to minutes from hours or days. The RecoveryONE solution accomplishes this through its unique ability to recover protected data down to the point in time preceding a logical corruption or hardware failure.

This document is for enterprise system administrators, storage administrators, and applications administrators who are interested in learning more about the RecoveryONE architecture. This guide assumes a medium to high-level familiarity with backup and recovery methodologies, terminology, and implementation.

Overview of the RecoveryONE Architecture Components

The RecoveryONE solution consists of the following components:

- A *Recovery Server* that manages the configuration, historical views, and policies for the components
- A *Protected Server* that hosts the applications and data to be protected
- A *Management Administrator* that supplies the GUI interface or CLI functionality to manage the entire recovery system
- A *Recovery Storage* system that is the time-addressable and event-addressable block storage
- A *Protected Storage* system that contains the data for the protected applications
- A *Private Storage* system that contains meta-data with configuration information for the Data Tap
- A *Historical Views* system that accesses the Historical Event Markers/Views
- An *Alternate Server* (optional) that is a server attached to the same SAN as the Recovery Server on which the historical volumes (Snapshots) are accessed

Figure 1-1 shows the RecoveryONE architecture components and how they communicate with one another across a Local Area Network (LAN) and a Storage Area Network (SAN).

Figure 1-1 RecoveryONE Architecture Components & Communication

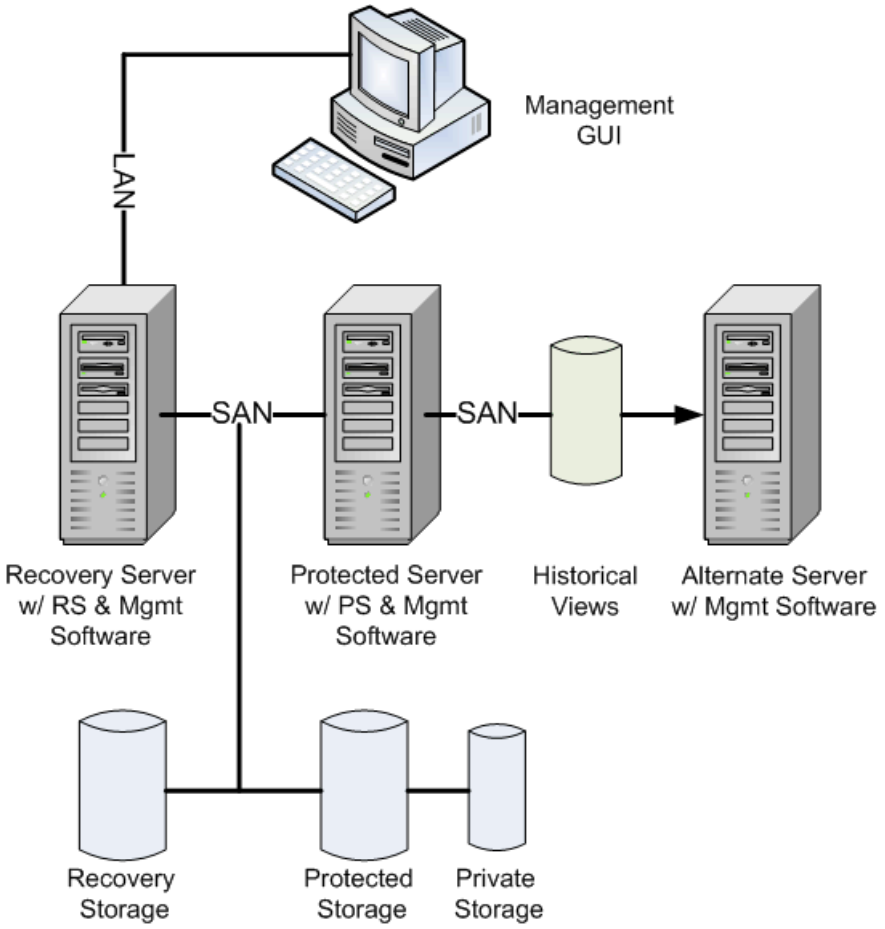


Table 1-1 provides the RecoveryONE terms and definitions.

Table 1-1 RecoveryONE Terms & Definitions

Component Terms	Definitions	Notes
Recovery Server	Server that provides continuous data protection services to one or more Protected Objects; it includes two components: the Mendocino Recovery Server Software and a defined reference platform.	
Protected Server	Server that hosts one or more applications that are being protected by the RecoveryONE solution.	

Management Administrator or Management GUI/CLI	RecoveryONE Management Graphical User Interface or Command Line Interface.	
Recovery Storage	Time-addressable and event-addressable block storage.	
Protected Storage	Storage used by the applications that are protected. The Protected Storage contains the Protected Objects.	
Private Storage	Storage used by the Protected Server software (Data Tap) component to store meta-data which contains configuration information.	
Historical Views	Historical point-in-time images of one or more Protected Objects that reside in Protection Sets, which can be accessed from the Protected Server or Alternate Server. Historical Views can be used to verify recovery points or off-host processing.	For more information, refer to the "Data Recovery Components" section in this guide.
Alternate Server	Any server that is attached by Fibre Channel to the same SAN that the Recovery Server is connected to and which accesses Snapshots (Historical Views).	Alternate Servers are optional.
Functionality Terms	**Definitions**	**Notes**
Protection Sets	Groups of interdependent protected objects on a Recovery Server for which the RecoveryONE solution provides continuous data protection.	Creating a Protection Set associates the Protected Objects with a Recovery Server and a Protected Server.
Protected Objects	Objects, such as a volumes or raw partitions or devices, which are protected through the RecoveryONE Recovery Management solution.	
Event Markers	A tag that marks a point in time between two writes in the data stream; this is a point in time or a point in process that is significant to you for the purposes of recovery.	Protection must be enabled for an existing Protection Set before you can place an Event Marker. For more information, refer to Chapter 5 in the *RecoveryONE Recovery Management User Guide*.
Snapshots	Same as Historical Views (see above).	An Event Marker must exist in the recovery timeline before a Snapshot can be taken. For more information, refer to Chapter 6 in the *RecoveryONE Recovery Management User Guide*.

Optimized Recovery Window	A window of time in which point in time images are available.	For more information, refer to the "Data Recovery Components" section in this guide.
APIT Window	A window of time in which point-in-time images are available. This window is the portion of the optimized recovery window which allows you to create historical views from any-point-in-time available (whether Event Markers pre-exist or not).	For more information, refer to the "Data Recovery Components" section in this guide.
Point In Time	A single, specific time on the recovery time line.	
Significant Point In Time	A point in the Protected Server's data stream, marked by an "Event Marker", that enables optimized recovery to a specific point in time. An Optimized Recovery Point corresponds to an event in the application process such as a database quiescence that is associated with completed transactions.	
Region Recovery	Region Recovery is the process of harvesting writes that did not make it over to a Recovery Server before a failure or loss of connection occurred.	For more information, refer to the "Protected Server" section in this guide.

Chapter 1: Data Handling Components

The following section describes the RecoveryONE data handling components in more detail. The components involved in data handling follow:

- Recovery Server
- Protected Server
- Alternate Server
- Management GUI
- Application Groups
- Protection Sets
- Recovery Storage, Protected Storage, and Private Storage

The RecoveryONE solution consists of a Recovery Server that is in communication with a Protected Server (and optionally an Alternate Server), which are controlled through the Management GUI or Command Line Interface. Your RecoveryONE recovery management system also contains the Recovery Storage, which stores data for the Recovery Server, Protected Storage, which stores the protected data (also known as Protection Sets), and Private Storage, which is used by the Data Tap.

In an active, fully-functioning recovery system, proprietary software monitors changes in the data and the changes are sent to protected storage. Also, a copy of all writes goes to the Recovery Server. If you employ an Alternate Server in your recovery environment, the Alternate Server is used to access historical views (also known as Snapshots).

Recovery Server

The Recovery Server is either a Linux or Windows machine and is responsible for managing the system configuration, historical views, and policies for your recovery environment. The Recovery Server communicates with the Protected Server by way of a SAN and stores a copy of all writes made to the Protected Server.

Protected Server

The Protected Server is either a Solaris or Windows machine and is the production server with one or more applications being protected. The Protected

Server copies all writes to the Recovery Storage and asynchronously sends meta-data to the Recovery Server

Alternate Server

The Alternate Server is either a Solaris or Windows machine and is responsible for accessing the historical views (Snapshots). It is attached to the same SAN as the Recovery Server. The Alternate Server is active in both Data Collection and Data Recovery activities.

Management GUI/Command Line Interface

The Management GUI/Command Line Interface is software running on a Windows machine that provides an easy-to-use interface with which to manage your recovery system. You use the GUI/Command Line to create and work with Protection Sets, insert and work with Event Markers, create and work with Snapshots, perform recovery operations, and manage application integration.

The Management GUI/Command Line communicates with the Recovery Server by way of a LAN.

Protection Sets

Protection Sets are made up of Protected Objects (volumes or raw partitions or devices) on the Protected Server. Application Groups are simply sets of Protected Objects (or Protection Sets) managed by the Protected Server. You create, modify, and delete Protection Sets by using the GUI or Command Line.

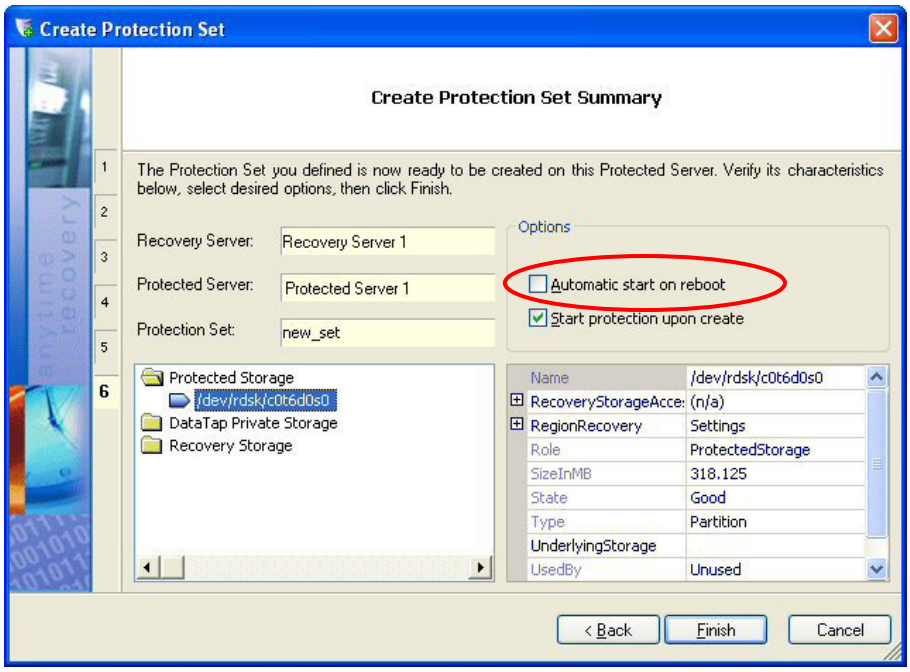
For more information on creating and working with Protection Sets, refer to the *RecoveryONE Recovery Management User Guide*.

Protection Sets & High Availability Software

You can control Protection Sets with HA software, which lets a third-party HA software suite control the system on which a Protection Set's device interfaces are made available to the application. However, because the behavior of HA-controlled Protection Sets is different than those not under HA control, this section describes both scenarios.

By default, Protection Sets are created such that they are not under HA control. If you want to use HA control software, be sure the “Automatic start on reboot” option is *disabled* in the Create Protection Set wizard (see Figure 1-2).

Figure 1-2 Enable HA Control Setting in Create Protection Set Wizard



Importing & Exporting HA-Controlled Protection Sets

The RecoveryONE solution lets you import and export Protection Sets to support High Availability software control of your Protection Sets.

- Importing a Protection Set instantiates the appropriate Data Tap devices and sets up the system configuration so that the protected device interfaces are available for use on the importing server. When a Protection Set is imported, any required Region Recovery is automatically performed. **Note:** A Protection Set can only be imported on a single server at a time; importing a single Protection Set on multiple hosts can corrupt data and can destroy the Protection Set's configuration.
- Exporting the Protection Set withdraws access to the devices from applications on the server and makes the Protection Set available for import by other servers.

HA-controlled Protection Sets differ from those not under HA control in one important aspect: HA-controlled Protection Sets are *never automatically imported* by the RecoveryONE subsystem. The HA software that is in control of the Protection Set "knows best" about what actions to take when a system reset occurs. In fact, if the HA software is going to fail over a Protection Set on

another server, it is in fact dangerous to have the RecoveryONE solution attempt to auto-import the Protection set.

Note: HA-controlled Protection Sets can be exported and imported using the `rxexport` and `rximport` commands; however, care should be taken to understand the interactions that such operations can cause with the HA software in control of a Protection Set. For more information about the rximport and rexexport commands, refer to the "Importing & Exporting Protection" section in Chapter 3 of the *RecoveryONE Recovery Management Solution Command Line User Guide*.

Importing & Exporting Non HA -Controlled Protection Sets

As noted earlier, by default, a Protection Set is created as *not* under HA control. For non HA-controlled Protection Sets, the process of importing the Protection Set on a server is handled by the RecoveryONE software. When first created, the Protection Set is (by definition) imported on the Protected Server where it was created. After a system restart (or a RecoveryONE subsystem restart), the RecoveryONE startup process imports all non-HA-controlled Protection Sets that were imported on the host at the time of the reset. These Protection Sets are said to have been *automatically imported*, or *auto-imported* by the RecoveryONE solution.

You can import and export non-HA- controlled Protection Sets explicitly. You do this from the command line using the `rximport` and `rxexport` utilities.

- The `rxexport <PSet-UUID>` command exports the user-specified Protection Set from the server where the command was issued and makes it available for import by other hosts. Once the Protection Set has been exported, it will not be automatically imported after a system reset or a RecoveryONE subsystem restart.
- The `rximport <PSet-UUID>` command causes the user-specified Protection Set to be imported on the server where the command was issued; this command also performs any recovery that is required and makes the protected devices available for use by applications on that server. Once successfully imported on the server, a system or RecoveryONE subsystem reset causes the Protection Set to be automatically imported after a system or RecoveryONE restart.

You can then enter the `rxshow` command to list the status of the Protection Sets known to the RecoveryONE subsystem on the server where the command is executed.

Recovery Storage, Protected Storage, & Private Storage

Recovery Storage is the storage used to store a replica of the protected storage and each write along with meta-data information used by the Recovery Server. The Recovery Storage is used by the Recovery Server when accessing and delivering historical views. Protected Storage contains the protected data; it is used by the applications that are being protected. The Private Storage stores configuration and other information about the Protection Sets and keeps a record of I/O operations to the Protected Storage that might not have been received by the Recovery Server.

Recovery Storage

The Recovery Storage is used by the Recovery Server to locate and access the Historical Views (Snapshots).

Recovery Storage & Raw Volumes versus Volume Manager Volumes

The Recovery Storage must be on raw volumes or non-volume manager volumes because Recovery Storage requires one contiguous LUN. Often the volume manager tools (for example, Veritas Volume Manager) contain multiple nested LUNS below a visible LUN, and this makes these volumes unsuitable for use in Recovery Storage.

Recovery Storage Discovery

The `mgclient` and the `mgserver` components automatically discover the LUNs on the FC SAN and the volumes as seen by the operating system. The discovery layer merges the two levels of discovery and presents it to the RecoveryONE system.

To uniquely identify the Recovery Storage LUNs across hosts, you must install the SNIA HBA API vendor libraries on the Recovery Server and the Protected Server (along with the HBA drivers). These libraries are distributed by the HBA vendors either in driver packages or in separate application kit packages. For example, if you load SANsurfer or HBAnyware you normally have these libraries. Note: SANsurfer and HBAanyware are not requirements of the RecoveryONE solution.

For information on how to download and install the proper libraries and Emulex and QLogic drivers, refer to the *RecoveryONE Recovery Management Hardware Guide*.

For more information on the steps for identifying and viewing Recovery storage, refer to Chapter 3 in the *RecoveryONE Recovery Management*

*Solution User Guide (*or the *RecoveryONE Recovery Management Solution Command Line User Guide).*

Protected Storage

The Protected Storage contains the protected data; it is used by the applications that are being protected. You assign which disks to protect in the GUI or Command Line.

Region Recovery

The Recovery Server provides the Protected Server with a group of locations for placing the data onto the Recovery Storage. Unlike traditional solutions, the RecoveryONE Server can be unavailable for some time before requiring a pause in protection or "blackout".

If the Recovery Server becomes unavailable or experiences a loss of connection to the Protected Server for a long period of time and the pre-allocated locations are filled up by Protected Server writes, the Protected Server continuously tracks locations of block changes. Once the Recovery Server is available again, a Region Recovery occurs which copies the changed blocks to the Recovery Storage. If over time the pre-allocated locations become full after a RecoveryONE Server failure, historical views are not available (a "blackout" exists) for that timeframe until Region Recovery has completed.

The Protected Server leverages a small amount of space in the Private Storage to track the changes when the pre-allocated locations are filled up in a RecoveryONE failure scenario. All I/Os in a given region are tracked using 1 bit as a measurement. The Region Recovery size (or density) is the number of bits mapped to each region. The Region Recovery setting balances size with performance. The value for Region Recovery is 32 MB (where every bit represents 32 MB of disk space). The smaller the region size, the finer the Statemap tracking is of that region, potentially increasing the amount of bits to "activate" during normal I/O, as a new region designated by a bit that is "deactivated" is accessed. .The advantage of having a smaller region size is that, potentially, less data is copied from the Protected Storage to the Recovery Storage during the region recovery process.

When you do a production rollback, you are undoing some of the last I/Os that occurred. If Region Recovery of the unsynchronized data is taking place when you attempt to do a production rollback, you can choose to let the Region Recovery process continue or discard the writes that occurred after the failure. If you let the Region Recovery process complete, you have an option to undo

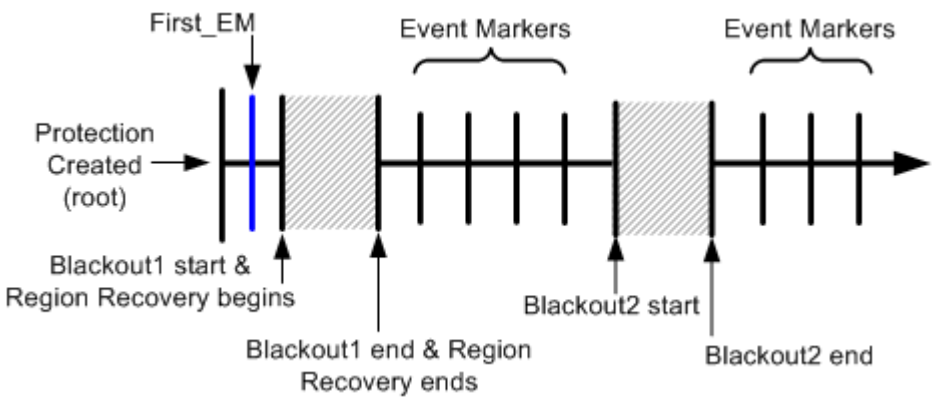
the production rollback process, enabling you to restore the image to the current point-in-time image. If you choose to not wait for the Region Recovery process to complete, you will have an option to undo the production rollback process, but will need to restore the image to the point-in-time before the blackout occurred. The GUI and CLI provide an estimate of how much data must be copied to complete Region Recovery.

Since production Rollback is the process whereby you make changes to your actual production data, this procedure requires careful consideration. The amount of data recovered depends on how far back in time you are rolling back to and how many writes occurred during the timeframe. For example, if you are rolling a Protection Set back to15 minutes prior and 20 MB of data has changed in that time period, in the Production Recovery process, you are only recovering that 20 MB of changed data.

Blackout Periods

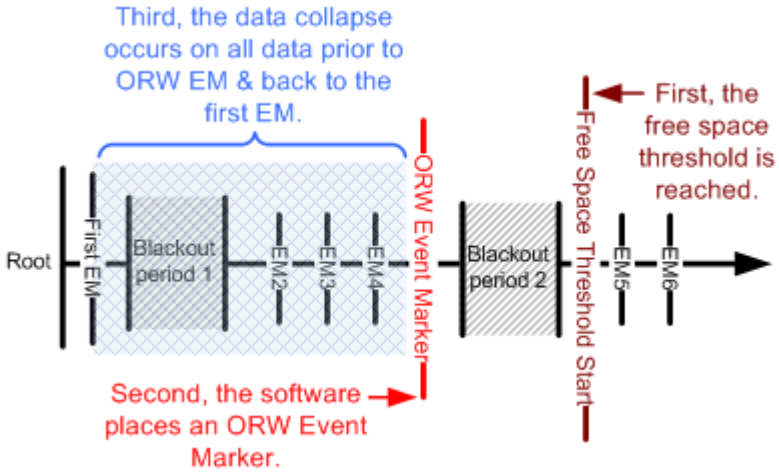
During Production Recovery, the Region Recovery function inserts two additional markers into the recovery timeline starting at the point where the oldest Event Marker exists. The new markers represent the blackout start and blackout end points of a blackout period. Blackout periods are portions of time that are not available for recovery. Blackout periods occur in two circumstances: when there was unavailable space for a write or when there was loss of connection between the Recovery Server and the Protected Server (such as when a machine goes down). Figure 1-3 shows a timeline that includes two blackout periods and illustrates the data collapse process.

Figure 1-3 Recovery Timeline with Blackout Periods



In Figure 1-3 above, the protection is created (at root) and an Event Marker, named First_EM, is placed. Region Recovery is performed at the beginning of Blackout_ 1. The Region Recovery process stops at the end of Blackout1. Figure 1-4 shows what happens to the timeline when the storage threshold is then reached.

Figure 1-4 Blackouts & Data Collapse Process



In Figure 1-4, the system chooses a point in time prior to the free space storage threshold being met and places an ORW_Event_Marker in the location. The data collapse process then begins from First_EM up to the ORW_Event_Marker. During this portion of the timeline, all of the Event Markers between First_EM and the ORW EM are deleted and you can not recover to any point in time during that section of the timeline.

Each time the free space storage threshold is met, another ORW Event Marker is recreated in a new spot along the recovery timeline and the data collapse

process repeats. For more information on the ORW and data collapse process, refer to the “Optimized Recovery Window” section in this guide.

Note: If your timeline includes Medium or High Priority Event Markers, the data is collapsed at those locations last and only after the ORW has expired or a Snapshot has been discarded. For more information, refer to “The ORW & Priority Event Markers” section in this guide.

Private Storage

The Private Storage stores configuration (and other information) about the Protection Sets and keeps a record of I/O operations to the Protected Storage that might not have been received by the Recovery Server.

Chapter 2: Data Recovery Components

The following section describes the RecoveryONE data recovery elements in more detail. The components involved in data recovery follow:

- Alternate Server
- Event Markers
- Snapshots
- Optimized Recovery Window
- Significant Points in Time

Alternate Server

The Alternate Server is a server that is responsible for accessing the historical views (Snapshots). It is attached to the same SAN as the Recovery Server. The Alternate Server software is a management component which auto-mounts the Historical Views for easier access.

Event Markers

Event Markers are tags that mark points in time or points in process that are significant to you for the purposes of recovery. For example, if your company experiences a catastrophic system failure, you can place a historical Event Marker that corresponds to the point in time directly *before* the failure and recover all of your company's protected data to that clean point in time.

You place Event Markers by using the Management GUI or Command Line. You can place an immediate Event Maker, which corresponds to the present moment in time or an historical Event Marker, which corresponds to some significant point in the past.

Event Markers have three levels of priority: Low, Medium, and High.

- Low Priority Event Markers are the default. Event Markers are assigned Low priority status when you use the default settings in the Create Event Marker dialog (or the "Preserve Event Marker" feature is disabled). The data represented by Low Priority Event Markers is collapsed first during the data collapse process.
- Medium Priority Event Markers are preserved within an active ORW. The data represented by Medium Priority Event Markers is only collapsed (as required by the system) after the ORW expires. Event

Markers are assigned Medium priority status when you select the "Preserve Event Marker" feature in the Create Event Marker dialog.

- High Priority Event Markers are automatically assigned this status after a Snapshot is created. Event Markers stay in High Priority status until you delete the associated Snapshot. Once a Snapshot is deleted, the Event Marker returns to the priority with which it was created (either Low or Medium). For example, if you place a medium priority Event Marker at a certain point in time and then take a Snapshot of the data at that point in time, the Event Marker is automatically designated as High Priority. When you delete the Snapshot later, the Event Marker returns to its former priority of Medium.

  Note: When you delete a Snapshot, all the writes that occurred to it are deleted, but this does not automatically delete the underlying Event Marker. Whether the Event Marker is removed or not depends on what the priority of the Event Marker was when it was created, whether your ORW has expired at the time the Snapshot is deleted, and the amount of available storage. For example, you decide to delete a Snapshot at which time the Event Marker returns to the status of Medium Priority. At that point, the ORW has not expired. The Event Marker representing that point in time remains in place until the ORW expires and the natural data collapse process ensues.

For information on how to set Event Markers, refer to Chapter 5 in the *RecoveryONE Recovery Management User Guide*.

Snapshots

Snapshots are point-in-time images of one or more Protected Objects that reside in Protection Sets. Snapshots help you validate whether a particular point in time is an acceptable place to recover to and can also be used for off-host processing. You place Snapshots by using the Management GUI or Command Line. The point in time for a Snapshot must be within the Optimal Recovery Window (For more information on the ORW, refer to the following section).

For more information on using Snapshots to perform off-host processing, refer to the "Creating a Block-Ordered Snapshot from a Virtual Snapshot" section in the *RecoveryONE Recovery Management User Guide*.

Before you can create a Snapshot, an Event Maker must exist for the date and time of the Snapshot that you want to create. For more information on setting Event Markers and taking Snapshots, refer to Chapters 5 and 6, respectively, in the *RecoveryONE Recovery Management User Guide*.

Optimized Recovery Window

To successfully manage the retrieval of historical views, your recovery environment must balance the size of the Optimized Recovery Window (ORW) with the amount of available storage.

Understanding the Optimized Recovery Window

The ORW is the user-specified window of time (in days) when the *any-points-in-time* can be presented. This is the window of time that the data restores (from Snapshots) are available to you before being discarded or overwritten. Determining the size of your ORW (and its effectiveness at presenting historical views) depends largely on balancing the amount of available storage (or free space) and the historical period of time in which the data can be recovered. Since you are striking a balance between these factors, the ORW often represents the *goal* for coverage, but it is not always the end result. The ORW default is five days.

To set an ORW, you create and enable protection for a Protection Set in the GUI or CLI, set an Event Marker that tracks the beginning of the protection phase, and then complete a Production Recovery (Region Recovery).

The `FREE_SPACE_THRESHOLD_START` and `FREE_SPACE_THRESHOLD_END` variables control at which percentage of the available storage the ORW begins and ends. As your protection system advances forward in time, the Recovery Storage begins to fill up with protected and historical data. The older data is typically of less interest after a certain amount of time has passed and the protected data has undergone extensive changes. When the Recovery Storage reaches the specified threshold, the free space commands facilitate data collapse, which is the process by which older (or unneeded) protected data is overwritten by newer protected data. You can preserve data within the ORW by setting the priority of Event Markers. For more information, refer to the "The ORW & Priority Event Markers" section.

The ORW & Priority Event Markers

The RecoveryONE solution provides a means of preserving data at specified points in time even if data collapse starts within the ORW. This is done by setting the priority of an Event Marker. If you set the priority of an Event Marker to Medium, the data is preserved at that point in time as long as the ORW is active. A High Priority Event Marker is automatically designated as such whenever you take a Snapshot of a point in time. High Priority Event Markers are not deleted until the Snapshot at that location is deleted. During the data collapse process, if any High Priority Event Marker exists, the

RecoveryONE solution will not overwrite the data at that point in time until you manually remove the marker. However, the data *around* the High Priority Event Marker is deleted by the RecoveryONE solution as necessary.

Figure 1-5 shows the High Priority Event Marker process. In the figure, you have an ORW specified, but the maximum storage threshold is reached before the ORW has expired, and the data collapse process begins. You have placed an Event Marker and taken a Snapshot of the data at 2:00 PM, which designates the Event Marker as High Priority. The data at all other points around 2:00 PM will be overwritten (as the system requires) until the minimum threshold for storage is met.

Figure 1-5 High Priority Event Marker Process

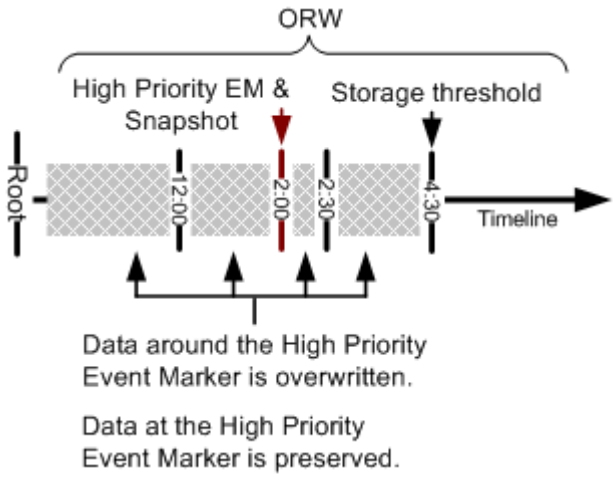
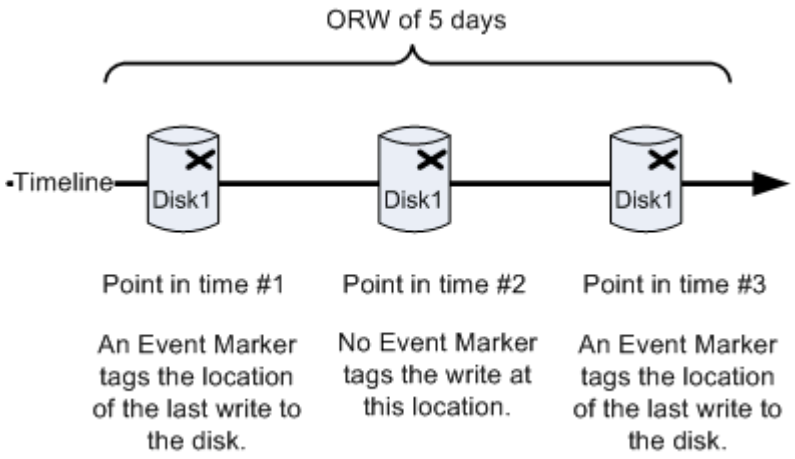


Figure 1-6 shows what happens when an ORW expires and a production recovery is required. In the figure, the disk has received various writes at different points in time along the recovery timeline. At the first point in time, the disk received two separate writes to the same location. A medium priority Event Marker exists at the point. At the second point in time, a new write occurred to the same location on the disk; however, no Event Marker exists for that point in time. At the third point in time, a write has again occurred in the same location on the disk and a medium priority Event Marker was placed there.

If you then want to perform a production recovery of your data, you can recover to points one and three because Event Markers exist for those points in time and the ORW has not expired. However, you can not roll back to point number two because no Event Marker exists for that location.

Figure 1-6 Timeline of Writes & Event Markers



The RecoveryONE solution provides three commands to govern the windows of time and thresholds used in presenting historical views: `ORW_WINDOW`, `FREE_SPACE_THRESHOLD_START`, and `FREE_SPACE_THRESHOLD_END`.

For more information on how to specify these commands and thresholds, refer to the "Managing the Optimized Recovery Window" section in Chapter 3 of the *RecoveryONE Recovery Management User Guide*.