

Increasing Recoverability of Critical Data with EMC Data Protection Advisor and Replication Analysis

A Detailed Review

Abstract

EMC® Data Protection Advisor (DPA) provides a comprehensive set of features designed to analyze replication activities and ensure that your data is protected and recoverable. Analyzing the host, application, and array configuration for gaps, DPA can capture these issues so that they can be addressed before a failure. This white paper outlines how Data Protection Advisor helps avoid the costs associated with lost data and reduces the risk of such loss, saves time and effort, and improves ROI for replication.

May 2010

Copyright © 2009, 2010 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 h6659.1

Table of Contents

Executive summary	4
Introduction	4
Audience	4
Overview	4
The challenge	4
The solution	5
Benefits of an automated solution	9
Going beyond replication	10
Conclusion	10
Supported environments	10
References	11

Executive summary

Businesses invest heavily in a range of data protection technologies to enable recovery under various failure and disaster scenarios. The goal of these technologies is to recover applications within specific time and data loss objectives. However, IT management and application owners have little visibility as to the ability of their data protection investments to recover their applications when needed. Therefore, they are often uncertain about their ability to meet the recovery objectives of the business.

Moreover, identifying and resolving recovery risks through testing and consulting is an error-prone and resource-intensive process. Overloaded staff and a skills shortage are putting the strategic goal of ensuring business continuity at risk. Furthermore, continuous access to critical information is now targeted by regulatory compliance audits, increasing pressure to provide positive proof of the ability to recover data.

Enterprises have to resort to manual and expensive methods such as custom scripting, infrequent disaster recovery testing, and external assessments to obtain an indication of their recoverability. Automating the process of analyzing application recoverability across data protection technologies dramatically reduces recovery failure risk, saves time and effort, and provides proof for auditors.

Introduction

This white paper outlines how EMC® Data Protection Advisor helps avoid the costs associated with lost data and reduces the risk of such loss, saves time and effort, and improves ROI for replication.

Audience

This white paper is aimed at mid- to high-level management responsible for storage technology costs, IT operations, DR planning/testing groups, and CIO-level management.

Overview

As the cost of downtime increases and with a rise in compliance and corporate governance initiatives, companies have had to evaluate the risk to business-critical applications. There are a number of strategies that can be employed to protect important data, and each has its strengths and weaknesses. Many businesses turn to replication to provide very high levels of availability. For the most business-critical data, companies rely on array-based replication. The cost of these solutions depends on a variety of factors including the amount of data, level of availability required, distance between recovery sites, and the choice of storage platform and replication technology, among others. Still, the cost of these solutions can be justified based on their proven track record of protecting data from loss, and businesses from the impact of that loss.

After deploying replication, organizations must ensure that the replication policies put in place continue to be effective over time. This is a continuous challenge as most environments are constantly changing. Small changes to the production environment such as increased capacity, or new servers and applications, can impact the recoverability of applications if replication policies or configurations are not adapted to the changed environment. There are a number of approaches that organizations use to ensure that their disaster recovery processes still work as designed. These include running regular DR tests, paying vendors for audits, and using stringent change management practices. All of these approaches take significant amounts of time and money, and are limited in their effectiveness.

The challenge

The challenge for businesses is to ensure that the replication solution remains accurate and effective in the face of the constant change that happens in every data center. Almost any change that modifies the existing application environment can have an impact on whether the replication/DR solution will work to meet the businesses requirements. Some things that can impact recoverability include:

- Application growth – New volumes, volumes on different arrays, volumes on the host

-
- Adding new applications – No available storage, no policy set, not done through the proper channels
 - Application upgrades – Upgrades can break existing procedures
 - Server and storage upgrades – Different mappings, different policies
 - LAN/SAN/WAN changes – New networks, new device mapping
 - Data center moves – Major/multiple configuration changes
 - Replication technology changes – Did everything get moved correctly?
 - Lack of DR testing - No validation that data is recoverable
 - Manual processes with changes captured manually – Were they entered correctly? Were all changes captured? How recently were they updated?
 - Application spread across multiple arrays

Some environments are simply very large and running heterogeneous hosts and applications, making it easy to miss a change and the potential impact to recoverability objectives.

To maintain the replication solution's effectiveness businesses periodically monitor the environment using manual checklists, spreadsheets, and manually created reports. The intent of these audits is to detect gaps in the solution before a failure occurs. Many companies run periodic DR tests to validate the ability to recover an application or complete data center. These events are manually intense and costly to run, driven higher by the scale of the test, and size of the environment. Large environments may dedicate full-time resources to ensure the recoverability of the company's data and applications. When dedicated resources are not available, consultants can be contracted to analyze the environment periodically. These consulting engagements can last a week to several months depending on need and can incur a large cost for each test.

All of these efforts are periodic and become invalid as soon as the next change happens — making the need for testing as constant as the ongoing changes. Several things are missing from these periodic manual efforts such as an automated and simple way to collect the information, get alerts on the status of the environment, and report on the status and configuration.

The solution

EMC Data Protection Advisor solves these problems and more. It automates the collection of data from applications, hosts, and arrays, constantly monitoring for exposures, and alerting on potential missed SLAs and gaps in the protection objectives. Data Protection Advisor (DPA) supports TimeFinder[®], SRDF[®], and RecoverPoint replication technologies on EMC Symmetrix[®] arrays and MirrorView[™], SAN Copy[™], SnapView[™], and RecoverPoint replication technologies on CLARiiON[®] arrays. Constant monitoring of the replication environment can significantly reduce the manual effort required and improve data availability. Finally, it provides a graphical map of the replication environment, enabling you to quickly see where the issues are, and how replication is set up for each host and application. The graphical map is an intuitive way to see the current state of the environment and the relationship between hosts, applications, and storage.

DPA provides:

- Replication analysis, which is a scalable, enterprise solution for monitoring replication environments to ensure that mission-critical applications are protected and recoverable
- Discovery and mapping
 - DPA discovers selected clients' applications, databases, and filesystems and maps them to physical storage devices. It maps all the copies (recovery points) of the primary data including snapshots, clones, and remote synchronous and asynchronous replicas.
 - After the discovery process, DPA correlates client, application, and storage mapping and other metadata to identify incomplete and inconsistent replicas that would result in a failed recovery.

- Display and report
 - DPA provides an intuitive graphical map of the relationship between the host and storage.
 - DPA presents the recoverability gaps and exposures in easy-to-understand reports and views that can be used by storage, database, and backup administrators to resolve issues.

There are a variety of replication error conditions that DPA can monitor for. Some examples include:

- Missed application volume, the replica is incomplete, and the application may not be recoverable
- Inconsistent replication technologies, and clone and snap images for the same host/application
- Different split times, impacting recoverability
- Different states, such as split versus synchronized

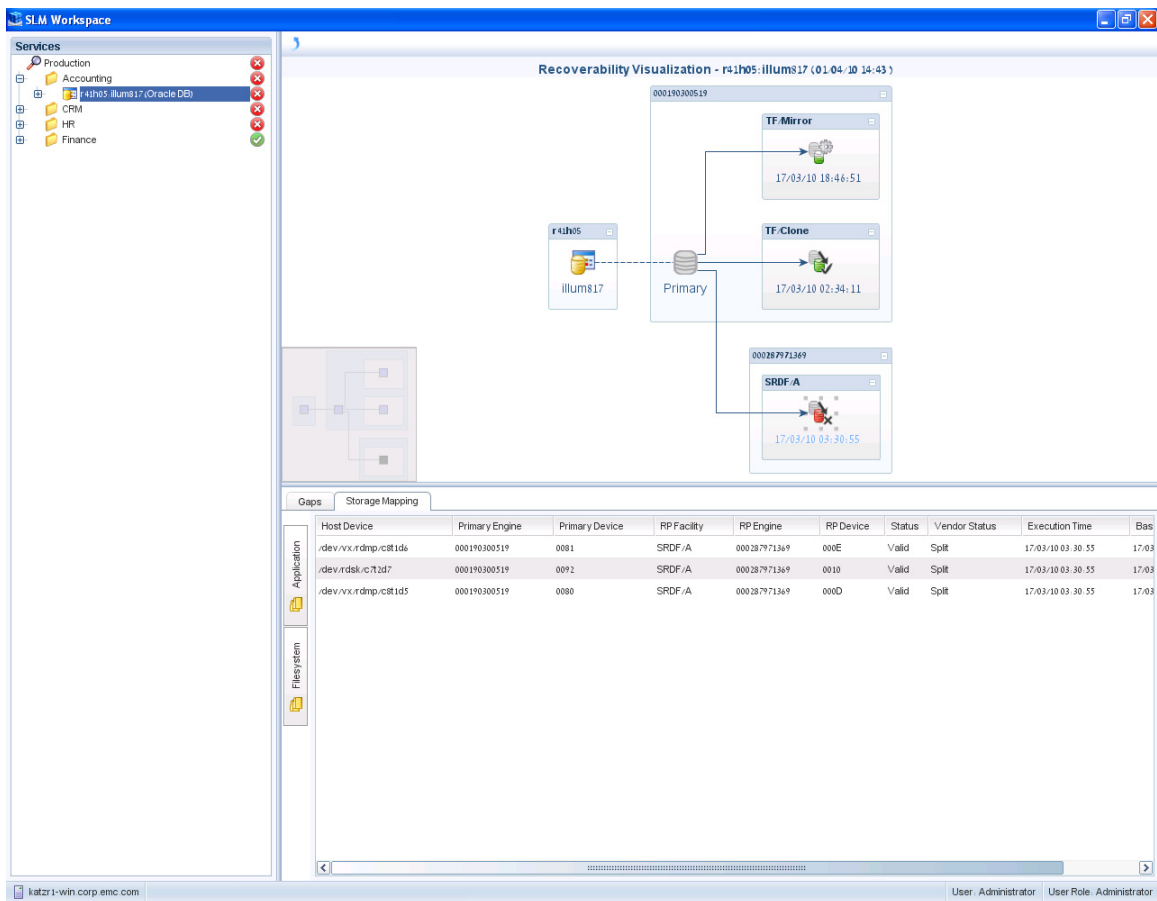


Figure 1. The DPA user interface

Figure 1 shows a topology view of a host residing on the primary array, with two copies in the same array and another in a second array. You can see very quickly that an issue has been identified with the remote copy. Below the map are a series of tabs that provide more detail on the hosts, arrays, and volumes in the map.

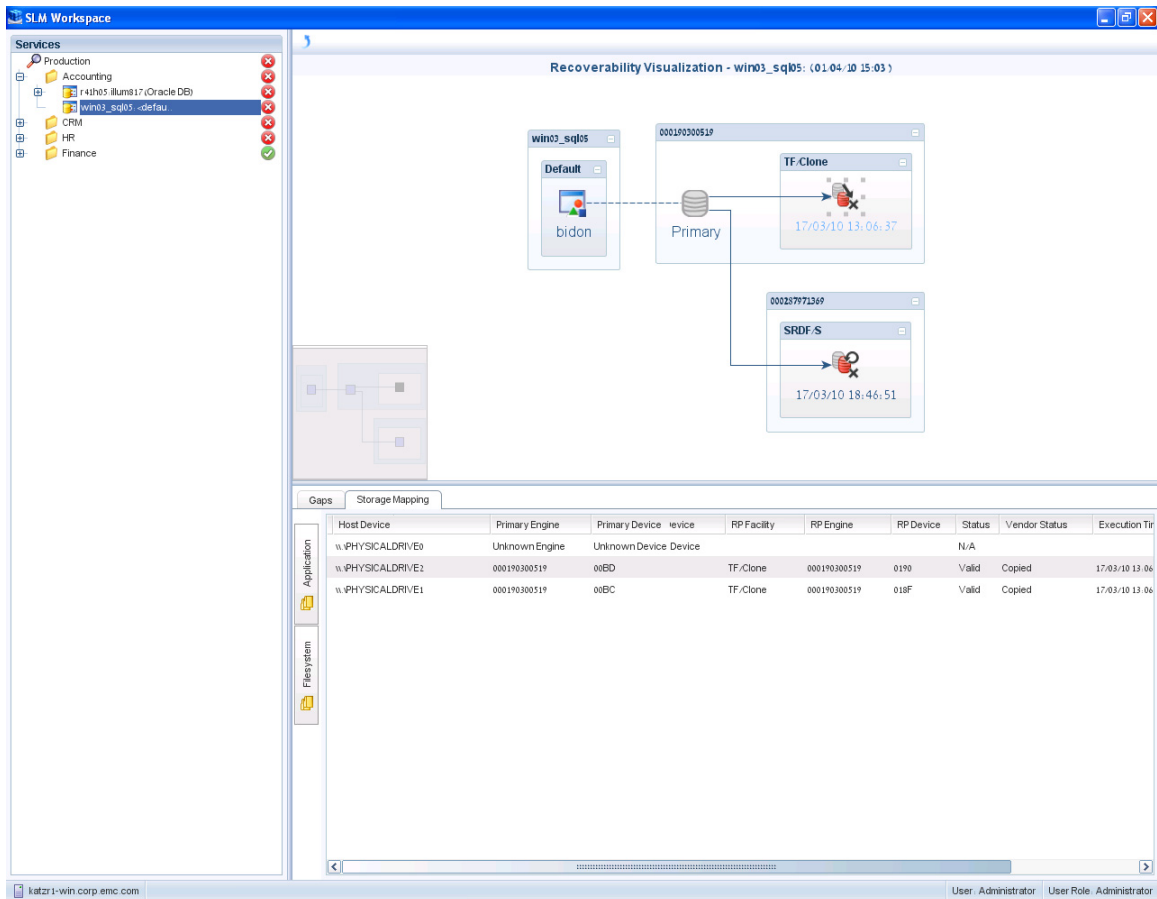


Figure 2. A partial replication where three volumes exist but only two are replicated

In Figure 2 above, the Recoverability Visualization is showing a host running SQL Server residing on a Symmetrix array. Both the local TimeFinder/Clone and remote SRDF/S copies are invalid replicas. By selecting the TimeFinder/Clone replica, DPA displays more details about the replica in the Storage Mapping tab to better understand the source of why it's invalid. The intelligent mapping within DPA shows that the host has three volumes that make up the SQL Server database; however, only two volumes are included in TimeFinder/Clone — identifying this as an invalid replica.

DPA has the ability to create Data Protection Policies. By creating and assigning these policies to business services/hosts/applications, DPA will check that those policies are enforced – meaning valid replicas exist, and those replicas are created in a timely manner that does not violate the RPO. Examples of such policies could be:

- Platinum policy – Continuous remote copy (RDF), remote copy (RBCV) refreshed every hour, local copy refreshed every hour, local snap refreshed every 10 minutes (total of six snaps)
- Gold policy – Continuous remote copy, local snap refreshed every 60 minutes (total of three snaps)
- Bronze policy – Local copy refreshed every hour

To see how this can assist with your environment, let's look at a few examples of automated checks:

Application-related checks

- Automatically detect a new application, such as Oracle, was added to a server. New databases or volumes can be overlooked; this check ensures all necessary data is replicated.
- Check that an application was not in backup mode during the replication.

-
- With Oracle, by detecting Oracle was up and not in backup mode, you are alerted that a replica is invalid.
 - Figure 1 shows an example of an application not in backup mode, while a BCV volume is split from the production volume.
 - Alert on missing archive logs. When a database was replicated but archive logs were not, the recovery can be limited.

Storage consistency checks

- Alert on non-consistent split.
 - Consistency is critical to recoverability of an application.
 - Some conditions that create inconsistency are different split times, different split states, or different replication technologies (clone versus snap).
- Alert on not using consistency groups.
 - With EMC TimeFinder[®] consistency groups can be used to split multiple volumes from two or more hosts simultaneously. By including this in the SLA for a host or application, DPA will catch if this option is not used.
 - SRDF[®] consistency groups should be used when an application resides on multiple arrays, or when the user wishes to split an application residing on multiple volumes. By setting this as the SLA for a host or application, DPA will catch if this option is not used.
 - DPA can also alert on non-consistent bookmarks created using EMC RecoverPoint. When the application resides on multiple consistency groups and one wishes to use EMC RecoverPoint to create a consistent bookmark, the RecoverPoint replication set should be used.

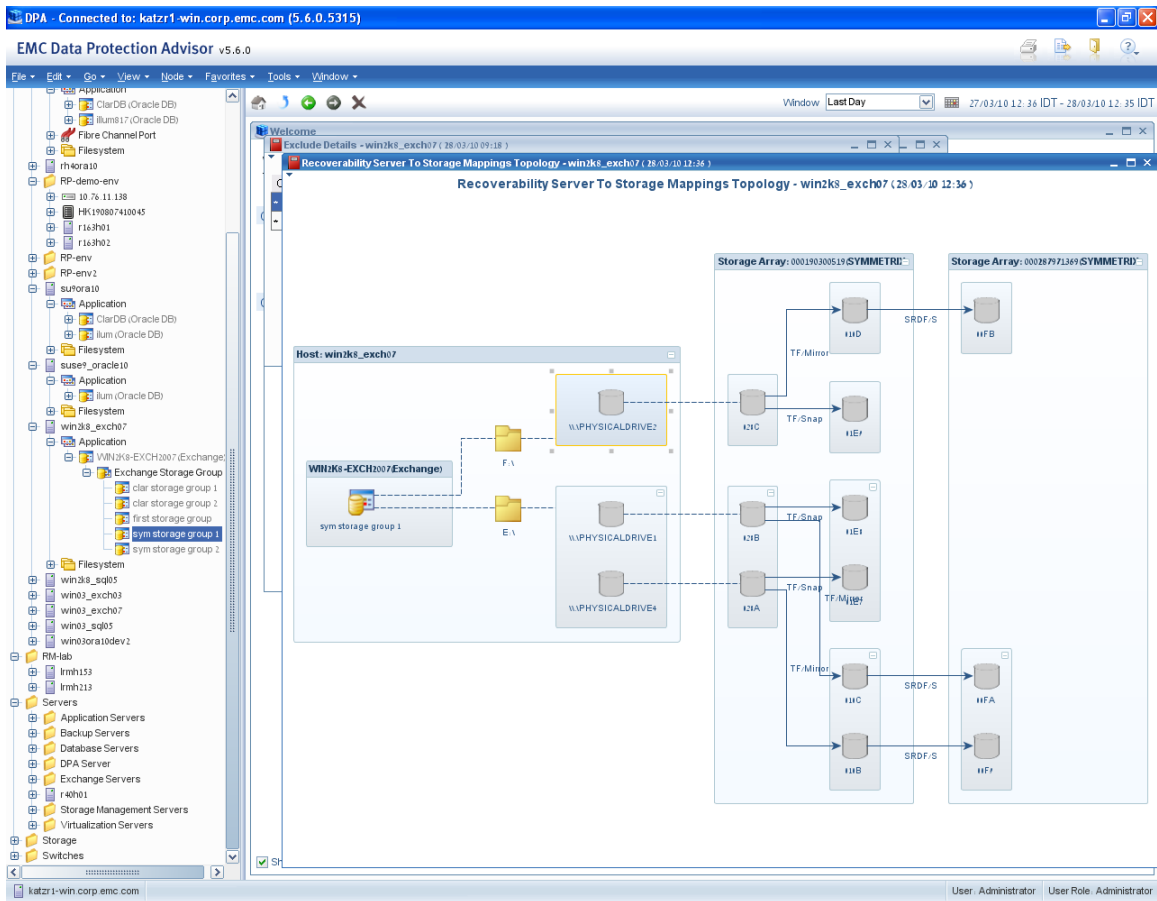


Figure 3. A server being mapped to local volumes and to a remote array; included are device names, volume groups, and array names

Benefits of an automated solution

Some areas of impact of an automated solution include:

- Manual data collection is eliminated
 - DPA automates the collection of data from both hosts and arrays, with collection running much more frequently than is possible manually.
 - A substantial portion of an FTE can be consumed with day-to-day activity and each DR test consuming a week or more of prep time for each test.
- Reduced time to complete audits
 - Audit impact comes from the automated data collection, and from graphical mapping showing the status of hosts and devices, as well as from the comprehensive reporting.
 - Customers have seen a 95 percent reduction in the time required to complete an audit using DPA, from 2 weeks down to 2 hours.
 - For quarterly audits, 8 weeks a year are saved.
- Consulting services costs
 - For periodic DR tests, consulting services are frequently employed. The duration varies, but costs can be \$15,000 per week. This is estimated to be \$60,000 per year for quarterly tests.
- Service providers reporting on achievement of SLA by customer

-
- Using DPA they can send those reports to their customers, with no extra work.
 - Penalties can be avoided by using DPA to track and ensure that RPO SLAs are met.
 - Costs will vary but one service provider was recently fined \$700,000 for failure to protect a single server.
 - Change management, including spreadsheet tracking of changes in the environment
 - A daily monitor of replication operations captures changes in configuration.
 - Configuration spreadsheets can be automatically generated from DPA, whenever they are needed.
 - DPA captures changes to the environment “as implemented,” not the proposed change.
 - Reducing the number of DR tests
 - By catching issues before a test, the success rate of each test will increase, and this can drive fewer tests annually.
 - This is estimated at one fewer test annually, based on quarterly tests.
 - Better storage utilization
 - Identifying unused copies that are sitting idle or are in use but not needed based on the SLA. For example, does a host or application really need three copies especially if the SLA only calls for one copy?
 - Identifying redundant copies. By providing a visual topology of all application copies to the application owner, the application owner can immediately detect redundant copies.
 - This storage space can be reclaimed to avoid additional storage purchases.
 - Avoiding data loss (the amount of cost saving depends on the organization)
 - The overall goal of replication, checks, and DR tests is specifically to avoid the loss of data and the loss of business that results.
 - Assuming \$10,000 per hour of downtime, a failure that takes two to three days to recover from tape due to a replication failure can cost \$480,000 to \$720,000 in lost business.
 - By catching changes, errors, and gaps in protection, data loss can be avoided before a failure.

Going beyond replication

Because data protection involves multiple software and hardware components, DPA monitors these components through data collection and analysis. Supporting the backup environment as well as the replication environment it is possible to set a comprehensive Protection Policy for systems and data. DPA can monitor, analyze and alert — whether the issue was a backup failure, multiple failures, or a replication issue — ensuring your data is protected properly.

Conclusion

DPA can save substantially on maintaining the recoverability of your data, through time savings, application availability, minimizing risk, optimization of resources, and consulting fees. The replication analysis and visualization capabilities in DPA provide the needed automation to simplify your daily operations, improve your data protection, and ensure that your most critical information is recoverable within the specified objectives.

Even cars as cheap as \$10,000 have continuous monitoring systems that let you know if your brakes or air bags have a problem. Why wouldn't you invest in a system to monitor the multimillion-dollar DR investment that your business depends on?

Supported environments

- EMC Symmetrix
 - TimeFinder: Mirror, Clone, and Snap
 - SRDF: Synchronous, Asynchronous, Concurrent, Cascading, and Star

-
- RecoverPoint CDP, CRR, and CLR
 - EMC CLARiiON
 - MirrorView/A and MirrorView/S
 - SnapView
 - SAN Copy
 - RecoverPoint CDP, CRR, and CLR
 - Operating systems
 - Sun Solaris 8, 9, 10
 - HP-UX 11iv1, 11iv2, 11iv3 (PA-RISC, IA64)
 - Linux (x86) - Red Hat 3.0, 4.0, 5.0 (AS/ES); SUSE 9, 10
 - AIX 5.2, 5.3, 6.1
 - Microsoft Windows (x86, x64) - Windows 2003, 2008
 - Applications
 - Oracle 8.1.x, 9.x, 10g, 11g
 - SQL Server 2000, 2005, 2008
 - Exchange 2000, 2003, 2007
 - Virtual hosts — DPA has the ability to map and alert on filesystems that are created on VMware, both when Raw Device Mapping (RDM) or VMFS is used.
 - In addition to the application awareness and topology mapping for the EMC RecoverPoint family, Data Protection Advisor also supports monitoring, alerting, and reporting capabilities for RecoverPoint appliances.

References

The following can provide additional information and can be found on Powerlink[®], EMC's password-protected customer- and partner-only extranet.

- *EMC Data Protection Advisor Reference Guide*
- *EMC Data Protection Advisor Administration Guide*
- *EMC Data Protection Advisor Installation Guide*
- *EMC Data Protection Advisor User Guide*
- *EMC Data Protection Advisor Compatibility Matrix*

For access to Evaluation licenses go to the Data Protection Advisor page on EMC.com:

<http://www.emc.com/products/detail/software/data-protection-advisor.htm>