Introduction to VMware vSphere® Data Protection

TECHNICAL WHITE PAPER

vmware®

Table of Contents

Introduction

VMware vSphere® Data Protection™ is a backup and recovery solution for VMware® virtual machines. It is fully integrated with VMware vCenter Server™ and VMware vSphere Web Client, providing disk-based backup of virtual machines. It is available in two versions:

- vSphere Data Protection – Introduced with vSphere 5.1 and included with VMware vSphere Essentials Plus Kit and higher.
- vSphere Data Protection Advanced – A new offering that builds on the success of vSphere Data Protection, primarily with increased capacity (up to 8TB) and application-specific agents for Microsoft Exchange Server and SQL Server. These agents facilitate application-consistent backups and more granular backup and restore capabilities.

Benefits

- Fast, efficient backup and recovery for vSphere virtual machines
- Significantly reduced backup data-disk space requirements, with a patented, variable-length deduplication technology across all backup jobs
- Use of VMware vSphere Storage APIs – Data Protection, and Changed Block Tracking (CBT), to reduce load on the vSphere host infrastructure and minimize backup window requirements
- Full virtual machine restore—or "image-level" restore—and file-level restore (FLR), without the need for an agent to be installed in every virtual machine
- Simplified deployment and configuration using a virtual appliance form factor
- vSphere Web Client–utilized administration
- Appliance and data protection via a checkpoint and rollback mechanism
- Easy restoration of Windows and Linux files by an end user with the Web-based vSphere Data Protection Restore Client

Additional benefits with vSphere Data Protection Advanced

- Mission-critical Microsoft Exchange Server and SQL Server workload protection with agents designed specifically for these applications
- Dynamic capacity growth as backup requirements grow
- Upgrade from vSphere Data Protection to vSphere Data Protection Advanced

This paper presents an overview of the architecture, deployment, configuration and management of vSphere Data Protection and vSphere Data Protection Advanced.

Architectural Overview

vSphere Data Protection and vSphere Data Protection Advanced require vCenter Server 5.1 or higher. vCenter Server can be the traditional Microsoft Windows implementation or the Linux-based vCenter Server Appliance. vSphere Data Protection and vSphere Data Protection Advanced support backing up virtual machines on vSphere versions 4.1 and higher. Web browsers must be enabled with Adobe Flash Player to access the vSphere Web Client and vSphere Data Protection functionality. See vSphere documentation for a list of Web browsers currently supported with vSphere Web Client.

vSphere Data Protection is deployed as a prebuilt, Linux-based virtual appliance. It supports as many as 100 virtual machines per appliance. vSphere Data Protection Advanced supports as many as 400 virtual machines per appliance. A maximum of 10 appliances per vCenter Server instance is supported.

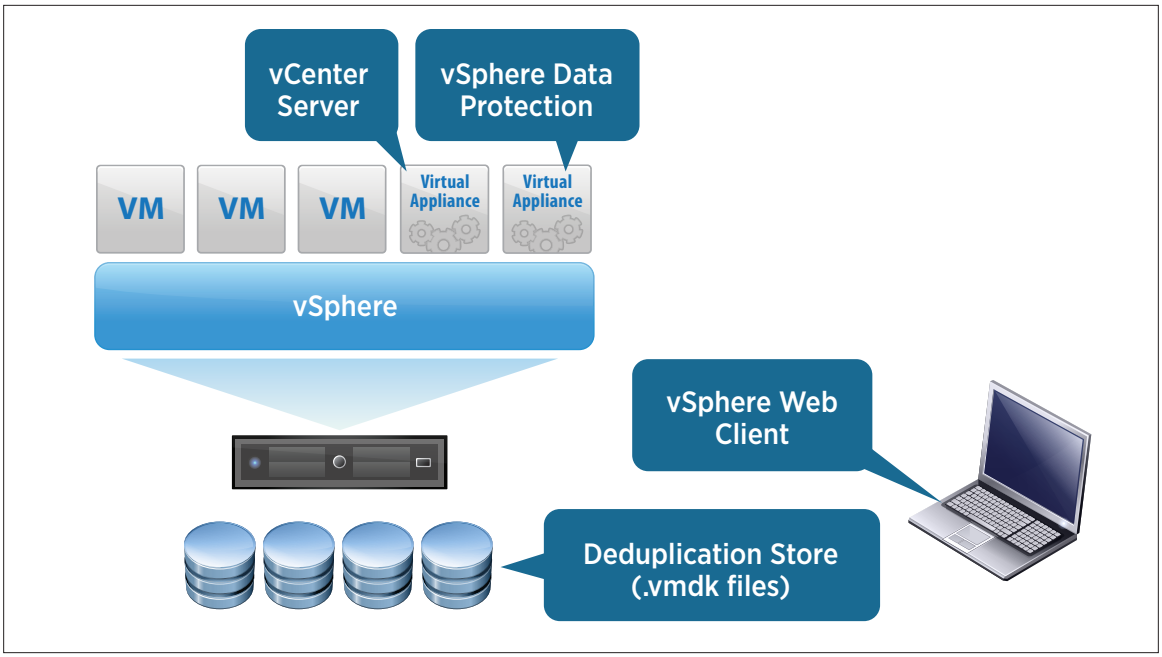

Figure 1. vSphere Data Protection Components

A vSphere Data Protection and vSphere Data Protection Advanced appliance is deployed by default with four processors. Figure 2 outlines the respective default storage and memory configurations. The actual amount of (thick-provisioned) storage consumed by an appliance is greater than the capacity shown in the "Deduplicated Backup Data Storage Capacity" column. This additional storage capacity is required for items such as the appliance guest operating system (OS), the vSphere Data Protection application and integrity checks. Thin-provisioned storage can also be utilized when deploying a vSphere Data Protection appliance.

	DEDUPLICATED BACKUP DATA STORAGE CAPACITY	MEMORY
vSphere Data Protection	.5TB	4GB
vSphere Data Protection	1TB	4GB
vSphere Data Protection	2TB	4GB
vSphere Data Protection Advanced	2TB	6GB
vSphere Data Protection Advanced	4TB	8GB
vSphere Data Protection Advanced	6TB	10GB
vSphere Data Protection Advanced	8TB	12GB

Figure 2. vSphere Data Protection and vSphere Data Protection Advanced Storage and Memory Configurations

When deploying vSphere Data Protection, plan adequately to help ensure proper sizing, because extra storage capacity cannot be added after the appliance has been deployed. In contrast, vSphere Data Protection Advanced enables dynamic provisioning of additional capacity (up to 8TB total). When determining storage capacity requirements, a number of items should be considered, such as number of protected virtual machines, amount of data being backed up, retention periods and typical data change rates.

Deployment and Configuration

vSphere Data Protection and vSphere Data Protection Advanced are deployed using vSphere Web Client or VMware vSphere Client™ from a prepackaged Open Virtualization Archive (.ova) file. The .ova files are labeled to easily identify the amount of backup data storage capacity deployed with the appliance.

After the appliance has been deployed and powered on, a Web browser is used to access the vSphere Data Protection–configured user interface (UI) and perform the initial configuration. The first time you connect to the vSphere Data Protection–configured UI, it will be running in "install mode." With the "install mode" wizard, items such as IP address, hostname, DNS, time zone and vCenter Server connection information are configured. Upon successful completion of the "install mode" wizard, the appliance will need to be rebooted. This reboot can take up to 30 minutes to complete as the appliance finishes initial configuration.

After the initial configuration, the vSphere Data Protection–configured utility runs in "maintenance mode." In this mode, the vSphere Data Protection–configured UI is utilized to perform functions such as starting and stopping services on the appliance, collecting logs, and rolling back the appliance to a previous valid configuration state, which will be discussed later in this document.

Administration

The vSphere Web Client is used to create and maintain backup jobs and to perform virtual machine and application restores, reporting and configuration.

Figure 3. vSphere Data Protection Advanced in vSphere Web Client

Backup

Creating and editing a backup job is accomplished using the **Backup** tab of the UI in vSphere Web Client. Individual virtual machines can be selected for backup. Containers such as datacenters, clusters and resource pools can also be selected for backup.

Backup jobs can be scheduled daily, weekly or monthly. Each job runs once on the day it is scheduled and begins when the backup window opens.

The retention policy can be defined in a few ways—for example, for x number of days or until a specific date. A custom retention policy can also be defined.

Figure 4. vSphere Data Protection Retention Policy Configuration

After a backup job has been created, it can be edited or deleted; it is also possible to clone the backup job. Cloning can be useful if, for example, the backup administrator wants to easily duplicate an existing custom retention policy for a new set of virtual machines. The administrator can clone the existing backup job and edit the selected virtual machines in the new—that is, clone—backup job.

vSphere Data Protection Advanced can protect and restore Microsoft Exchange Server and SQL Server. An application-specific agent is installed in the guest OS of a virtual machine. vSphere Data Protection Advanced utilizes this agent to back up and restore the Exchange or SQL Server application and databases. Using these agents with vSphere Data Protection Advanced also enables application-consistent backups and provides support for other options such as full, differential or incremental backups and multistreaming backups.

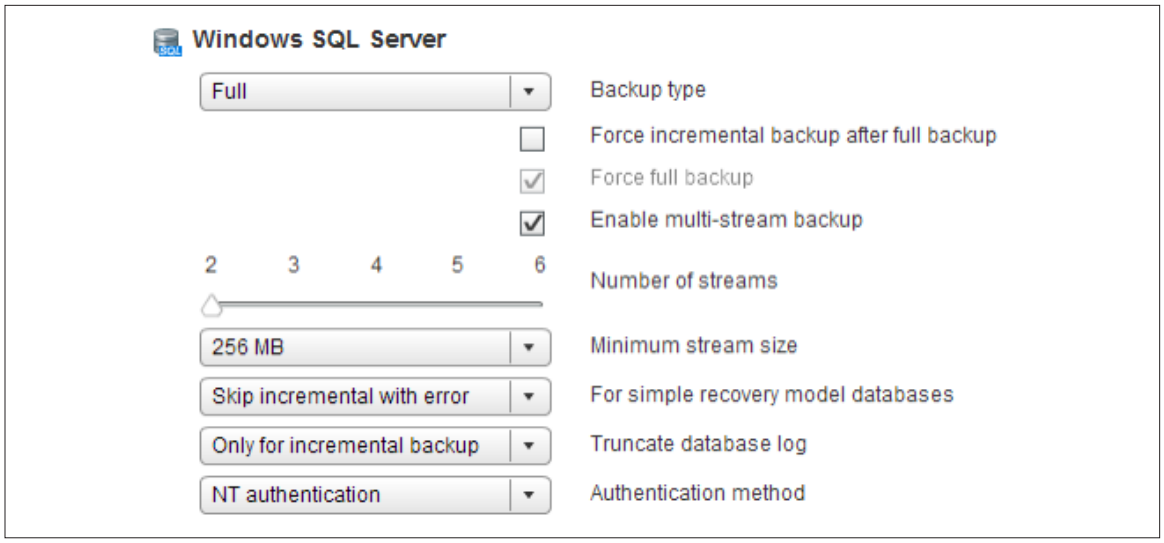

Figure 5. Application Backup Job Options in vSphere Data Protection Advanced

The initial backup of a virtual machine or application takes more time because all of the data for that virtual machine is being backed up. Subsequent backups take less time because vSphere Data Protection and vSphere Data Protection Advanced utilize CBT in vSphere.

Restore

Restoring an entire virtual machine is performed using the Restore tab of the UI of vSphere Data Protection and vSphere Data Protection Advanced in vSphere Web Client. The administrator can browse the list of protected virtual machines and select one or more restore points.

vSphere Data Protection and vSphere Data Protection Advanced offer fast and efficient recovery by leveraging CBT. When restoring an entire virtual machine, workloads of both a full image restore and a recovery leveraging CBT are evaluated. In some cases, the change rate since the last backup is very high, so the overhead of a CBT analysis operation would be more costly than that for a full image restore. vSphere Data Protection intelligently determines which method will result in the fastest virtual machine recovery time.

To prevent overwriting an existing virtual machine, a new virtual machine name and destination datastore can be specified. Choosing a restore location other than the original will result in a full image restore—CBT is not leveraged. There is also the option to power on the virtual machine and reconnect its NIC after the restore has been completed. This is useful for performing rehearsals to verify that existing backups can be restored successfully.

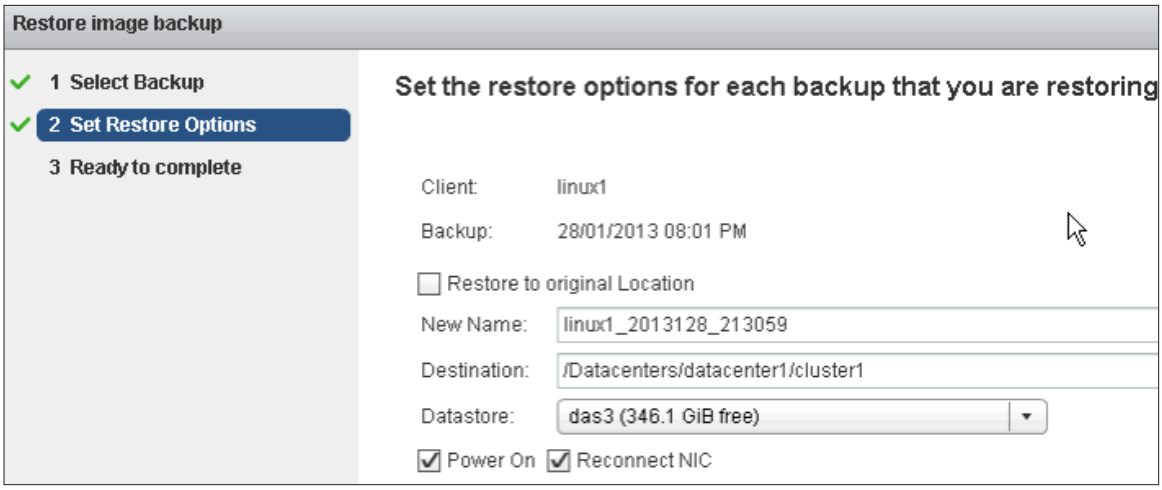

Figure 6. Specifying a New Name and Location for the Restored Virtual Machine

With vSphere Data Protection and vSphere Data Protection Advanced, it is also possible to restore individual files and folders/directories within a virtual machine. A file-level recovery is performed using a Web-based tool called vSphere Data Protection Restore Client. The process enables end users to conduct restores on their own, without the assistance of a backup administrator, by selecting a restore point and browsing the file system as it looked at the time that backup was done. They locate the item(s) to be recovered, select a destination for the restored items and start the recovery. The progress of the restore job can be monitored in vSphere Data Protection Restore Client.

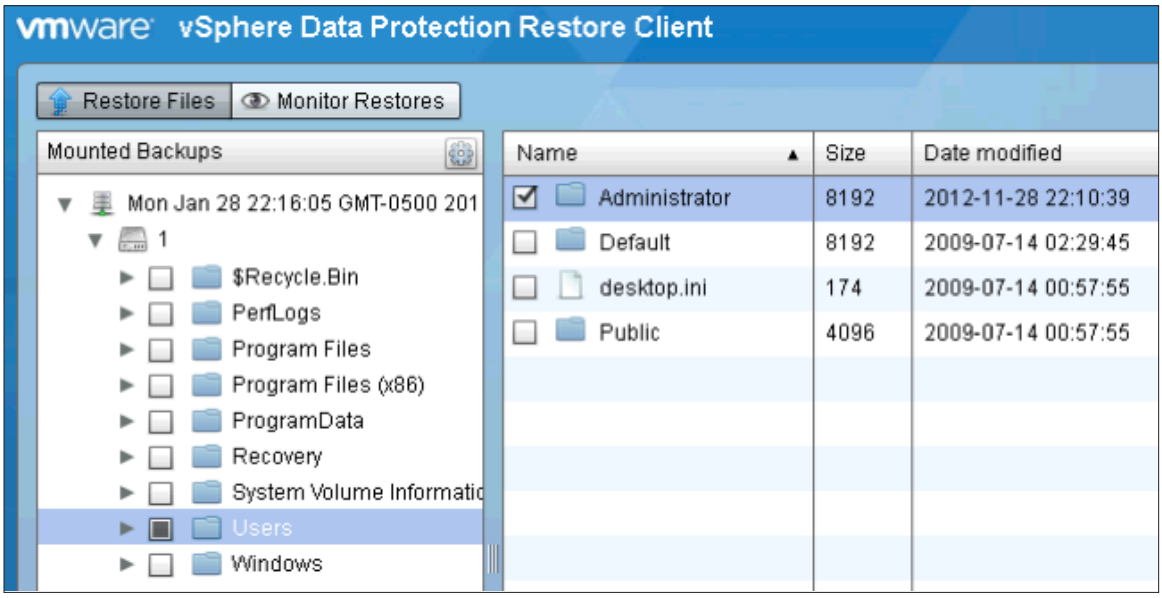

Figure 7. vSphere Data Protection Restore Client

Exclusive to vSphere Data Protection Advanced is the ability to restore Microsoft Exchange Server and SQL Server applications and their associated databases. Figure 8 shows an individual database selected for recovery.

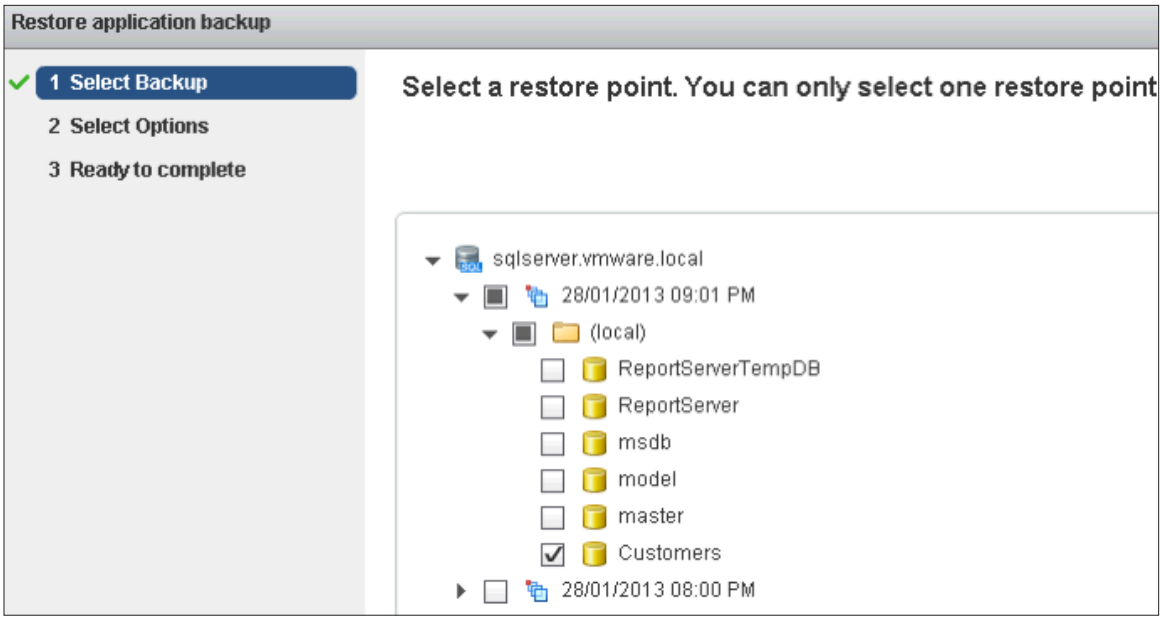

Figure 8. Individual Database Selected for Recovery in vSphere Data Protection Advanced

Another example is restoration to a recovery database (RDB) in Exchange Server 2010. The Exchange Server administrator can recover an individual mailbox or mailbox folder from the RDB.

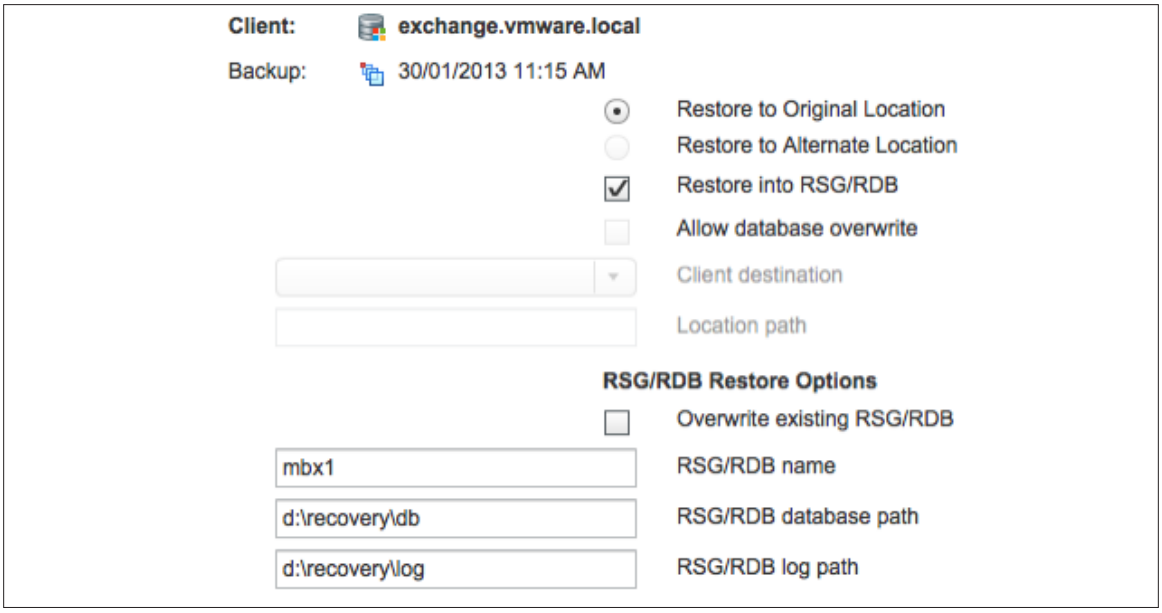

Figure 9. Exchange Server Restore into a Recovery Database in vSphere Data Protection Advanced

Reporting

The Reports tab in vSphere Data Protection and vSphere Data Protection Advanced displays a variety of information: appliance status, used capacity, backup job and virtual machine backup details, and so on. There are links to the event console and task console for additional information and for troubleshooting purposes. The list of virtual machines can be filtered using several criteria; for example, virtual machine name or date of the most recent backup. The Details section displays specifics about the virtual machine selected in the list of clients—items such as virtual machine name, guest OS, backup status and date of last successful backup.

In addition to the reporting capabilities of the UI, vSphere Data Protection and vSphere Data Protection Advanced can be configured to send email reports. These email reports can be scheduled at a specific time, once per day on any or all days of the week. Similar to the UI, these email messages contain details on the vSphere Data Protection appliance, backup jobs and the virtual machines that are backed up.

Avoiding Backup Data Corruption

vSphere Data Protection and vSphere Data Protection Advanced contain a checkpoint and rollback mechanism. A checkpoint is a system-wide backup of the vSphere Data Protection appliance that is performed to help protect the appliance from risks that might cause data corruption. An unexpected appliance power-off is an example of this. In this case, the appliance would roll back to the last validated checkpoint. Any backup jobs performed after that checkpoint would be lost, but data corruption—that is, loss of all backup information—likely would be avoided.

Checkpoint tag	Date	Valid
cp.20130128180932	01/28/2013 01:09:32 PM, EST -0500	Validated
cp.20130129040231	01/28/2013 11:02:31 PM, EST -0500	Validated

Figure 10. vSphere Data Protection Rollback Checkpoints

Summary

Data protection is a key component of any business continuity plan. VMware vSphere Data Protection and vSphere Data Protection Advanced provide efficient solutions for protecting a VMware virtual machine infrastructure including mission-critical applications such as Microsoft Exchange Server and SQL Server. Deployment is quick and easy. Administration is performed using a Web-based graphical UI integrated with the VMware vSphere Web Client. End users can restore files without requiring assistance from a backup administrator. A checkpoint and rollback protection system is included to help ensure that backup data is available for restoration when data loss occurs or disaster strikes.

About the Author

Jeff Hunter is a senior technical marketing manager with a focus on business continuity solutions at VMware. He has been employed at VMware for more than five years. Prior to Jeff's tenure at VMware, he spent several years assisting with the implementation and administration of VMware virtual infrastructures at two Fortune 500 companies.

VMware, Inc. 3401 Hillview Avenue Palo Alto CA 94304 USA Tel 877-486-9273 Fax 650-427-5001 www.vmware.com