

EMC Avamar Backup and Recovery for VMware Environments

Applied Technology

Abstract

This white paper describes components of the VMware vSphere and VMware View solutions and discusses options for protecting these environments using EMC[®] Avamar[®] with global source-based data deduplication.

August 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 h6396.2

Table of Contents

Executive summary	4
Introduction.....	4
Audience	4
VMware Infrastructure	4
VMware vSphere virtual data center operating system	4
VMware View desktop technology	5
VMware Infrastructure components	6
Virtualization deployment options	6
EMC Avamar – Never back up the same data twice.....	7
Avamar for VMware	8
Solutions	9
Avamar for virtual infrastructures	9
Guest-based backup	9
Image backup based on vStorage APIs for Data Protection.....	11
vCenter integration.....	12
Avamar for VMware View	13
Data Protection Strategy No. 1 for a VMware View environment	13
Data Protection Strategy No. 2 for a VMware View environment	13
Recovering VMware View components	15
Avamar deployment options	16
Centralized virtualization	16
Distributed virtualization	16
Conclusion	17
References	18

Executive summary

VMware products offer the industry's first cloud operating system virtualization suite that allows enterprises and small businesses alike to transform, manage, and optimize their IT infrastructure through virtualization. VMware Infrastructure delivers comprehensive virtualization, management, resource optimization, application availability, and operational automation capabilities in an integrated offering.

As VMware Infrastructure is the industry's most widely deployed virtualization solution, it is important to ensure that virtual machines deployed in both the data center and remote offices must be protected against failure. Extending data protection to virtual machines is thus an important function. In the virtualized environment provided by VMware Infrastructure, there are many ways to improve the convenience and reliability of data protection, each with its particular advantages and challenges.

EMC® Avamar® is an enterprise-class backup and recovery / disaster recovery solution that is optimized for VMware virtual infrastructures that uses unique deduplication technology to back up VMware Infrastructure components efficiently and removes the traditional burden of backup on the shared resources.

Introduction

This white paper discusses details of the Avamar solution and various ways of providing data protection for the various VMware portfolio products.

Audience

The information in this white paper is primarily intended for business application administrators and backup systems administrators who are responsible for architecting, deploying, and protecting a VMware View environment. They must have a working knowledge of the components that comprise a VMware View solution, including Microsoft Active Directory (AD), VMware View Manager (formerly Virtual Desktop Manager), and VMware Infrastructure (vCenter and ESX).

VMware Infrastructure

VMware vSphere virtual data center operating system

VMware's flagship product, VMware Infrastructure, coupled with VMware's comprehensive roadmap of groundbreaking new products provide a virtual data center operating system (OS) for IT environments of all sizes. The virtual data center OS addresses customers' needs for flexibility, speed, resiliency, and efficiency by transforming the data center into an "internal cloud" – an elastic, shared, self-managing and self-healing utility that can federate with external clouds of computing capacity, freeing IT from the constraints of static hardware-mapped applications. The virtual data center OS guarantees appropriate levels of availability, security, and scalability to all applications independent of hardware and location. Just like the single server OS was an indispensable part of the traditional IT stack, the virtual data center OS is an indispensable platform for business computing of the future.

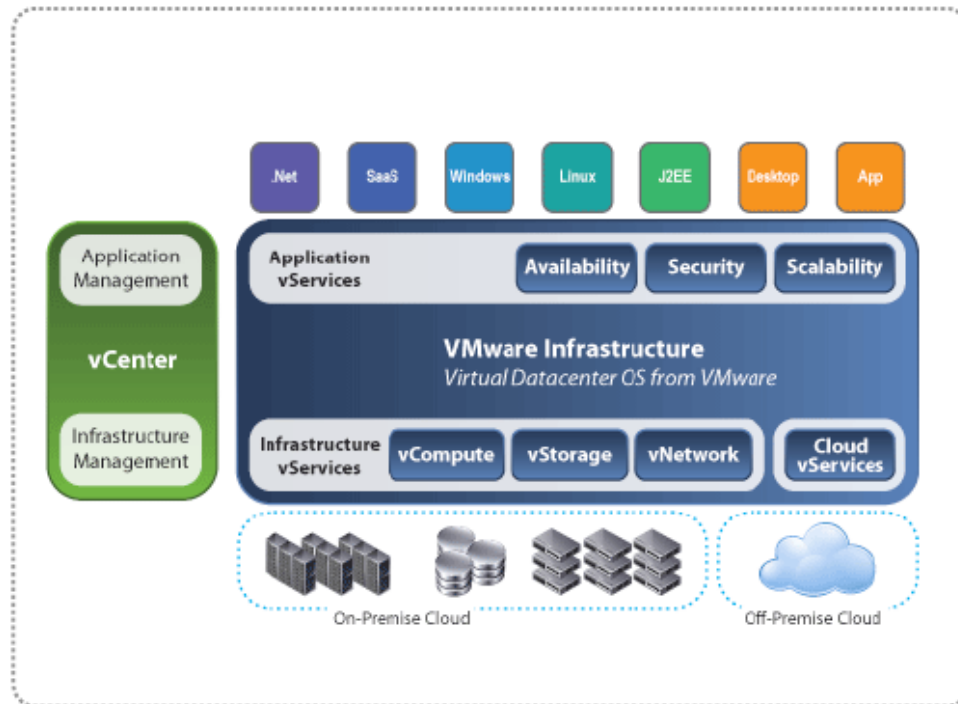


Figure 1. VMware vSphere virtual data center OS

VMware View desktop technology

Business is more dependent on technology than ever before, yet more frustrated by its inflexibility. With an increasing mobile and globally dispersed workforce using multiple devices on multiple platforms, they struggle to connect to their data and applications across a tangle of web, desktop, and server-based solutions. IT struggles to retrofit and manage the tightly bound single-purpose system of an OS, applications, and hardware. And when a user breaks or misplaces a computer, then productivity stops, security is breached, and intellectual property can be lost. Reconnecting is no easy task: Days are lost bringing users back online and weeks go by as IT tries to recoup lost information.

The desktop of the future will not be a single physical device but a collection of different devices and environments. Applications and data may be located across a combination of locations, for example, a virtual desktop running on a server, a home notebook computer, and a webmail account. End users want the same view regardless of what device they use to connect to their desktop or where their applications and data are located – the user wants a universal client. IT organizations on the other hand want to simplify management and take control of desktops and applications cost-effectively. A universal client is the next evolution in desktop computing, including a virtual desktop infrastructure.

Decouple applications, data, and the operating system from the hardware and deliver them to the user rather than to a device. Give users a personal view of their applications and data, whether they're on a thin client or laptop, in the office, or on the road. Intelligently deliver applications and data to any device and allow users to focus on their jobs rather than on the tools. Balance the requirements of your business with the needs of your users and create a seamless experience where applications and data follow the user and not the device.

The monolithic model of tightly coupled hardware, an operating system, and applications cannot keep up with today's global economy while complying with business, regulatory, and security objectives. You need flexible solutions to drive your infrastructure based on the needs of your business. See how to improve your desktop management and:

- Deliver personalized applications and desktops to a user, not a device
- Enable flexibility with control

- Centrally manage and secure user desktop environments

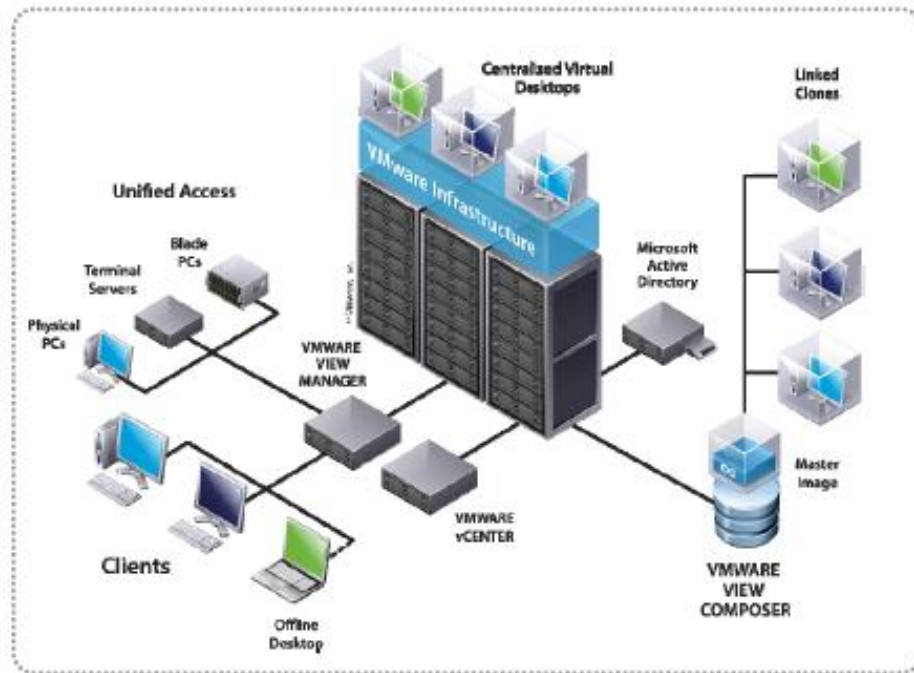


Figure 2. Example of a VMware View environment

VMware Infrastructure components

VMware Infrastructure includes the following major components:

- **VMware ESX** — A production-proven virtualization layer run on physical servers that abstracts processor, memory, storage, and networking resources to be provisioned to multiple virtual machines.
- **VMware Virtual Machine File System (VMware VMFS)** — A high-performance cluster file system for virtual machines.
- **vCenter Management Server** — The central point for configuring, provisioning, and managing a virtualized IT infrastructure.
- **Virtual Infrastructure Client (VI Client)** — An interface that allows administrators and users to connect remotely to the VirtualCenter Management Server or individual ESX installations from any Windows PC.
- **VMware VMotion** — Enables the live migration of running virtual machines from one physical server to another with zero downtime, continuous service availability, and complete transaction integrity.
- **VMware vStorage APIs for Data Protection** — Takes the backup load off the ESX host, eliminates the backup window, removes backup traffic from the LAN, and eliminates the need to run backup agents inside virtual machines to perform image-level and file-level backups of virtual machine data.

Virtualization deployment options

Depending on business needs and network reliability, physical servers running VMware Infrastructure can be deployed centrally in a data center or locally in a remote office.

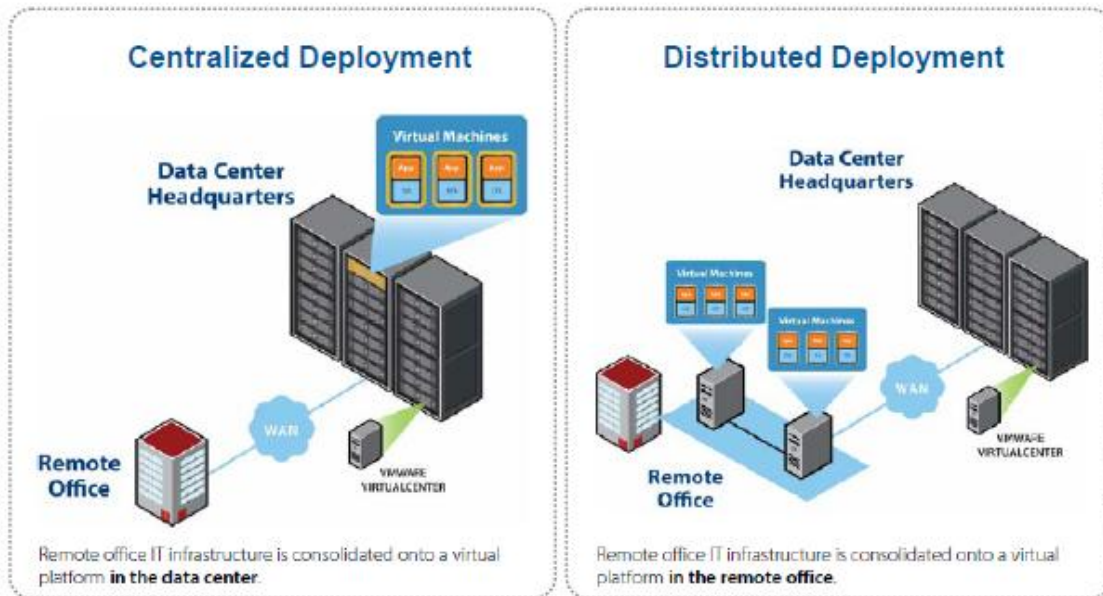


Figure 3. Deployment approaches for a virtual infrastructure

- Centralized deployment** – VMware Infrastructure consolidates remote office IT infrastructure onto a virtual platform in the data center, and remote offices access server and desktop workloads over a secure network connection. This option is ideal for organizations with reliable, high-bandwidth, low-latency network links, or organizations that have implemented a wide-area data services solution for application acceleration across the wide area network (WAN). This deployment option maximizes server consolidation ratios and cost savings.
- Distributed deployment** – VMware Infrastructure consolidates remote office IT infrastructure onto a virtual platform in the remote office, and a central IT staff manages server and desktop workloads remotely from the data center. This option is ideal for organizations with unreliable network links, or organizations that require physical servers to be located close to the end user. This deployment option maintains application performance regardless of WAN speed and availability.

EMC Avamar – Never back up the same data twice

Revolutionize your backup by moving less data to solve your toughest VMware, NAS, remote office, and desktop/laptop backup challenges

Traditional backup solutions require a rotational schedule of full and incremental backups, which move a significant amount of redundant data week over week. Because of the unnecessary data movement, enterprises are often faced with backup windows that roll into production hours, network constraints, and too much storage under management. In VMware Infrastructure environments, server consolidation can mean overlapping backup windows and heavy impact on hardware resources.

Over the last decade, disk storage has been used to augment traditional backup approaches, but disk solutions that are designed to replace tape libraries and media solve only a fraction of the data protection challenges faced by enterprises.

EMC Avamar backup and recovery software with integrated source, global data deduplication solves the challenges associated with traditional backup, enabling fast, efficient protection for remote offices, VMware environments, and data center LAN / NAS servers.

Unlike with traditional backup solutions, EMC Avamar identifies redundant data segments at the source — before they are transferred across the network. By moving only new, unique subfile data segments, Avamar

delivers fast daily full backups while reducing the required daily network bandwidth by up to 500x. This capability allows companies to utilize existing network bandwidth for backup and disaster recovery of remote offices and data centers, despite slow or congested networks and infrastructure. Data can be encrypted both in flight and at rest for added security, and centralized management makes protecting hundreds of remote offices easy and efficient.

By storing just a single instance of each subfile data segment globally, EMC Avamar also reduces total back-end storage by up to 50x, enabling cost-effective disk-based recovery over extended periods of time. Although EMC Avamar backs up data to disk, it can also work with existing tape and traditional backup software such as EMC NetWorker®. In addition, EMC Avamar's grid architecture provides online scalability, and its patented redundant array of independent nodes (RAIN) technology provides high availability.

Avamar for VMware

Avamar software quickly and efficiently protects VMware Infrastructure environments by reducing the size of backup data within and across virtual machines —using agents in the virtual machines or on the VMware vStorage APIs for Data Protection proxy server. For virtual machine backups, Avamar eliminates traditional backup bottlenecks caused by the large amount of redundant data that must pass through the same set of shared resources — the physical server's CPU, Ethernet adapter, memory, and disk storage. Avamar reduces the traditional backup load — up to 200 percent weekly — to as little as 2 percent over the same day period, dramatically reducing backup times and resource utilization.

Key Avamar benefits include:

- Up to 10x faster daily full backups
- Up to 500:1 reduction in required daily network bandwidth
- Up to 50:1 reduction in required global backup storage media
- Encryption of backup data in flight and at rest
- Fault tolerance across Avamar nodes and elimination of single points of failure using the patented RAIN technology
- Scalable grid architecture
- Daily server integrity and data recoverability checks
- Simple one-step recovery
- Flexible deployment options, including EMC Avamar Data Store and EMC Avamar Virtual Edition for VMware (a virtual appliance)
- Improved physical server consolidation ratios

Figure 4 represents a comparison of full backups using traditional methods versus full backups using Avamar deduplication technology. The left side of each graph represents the impact on the shared resource (CPU / network / disk) using traditional backup solutions, and the right side of each graph represents the impact of the Avamar solution on the shared resource. This reduced impact on the shared resources of an ESX server when running Avamar at the guest or vStorage API level allows users to easily meet backup windows and lessen the network bandwidth requirements for a backup and recovery infrastructure.

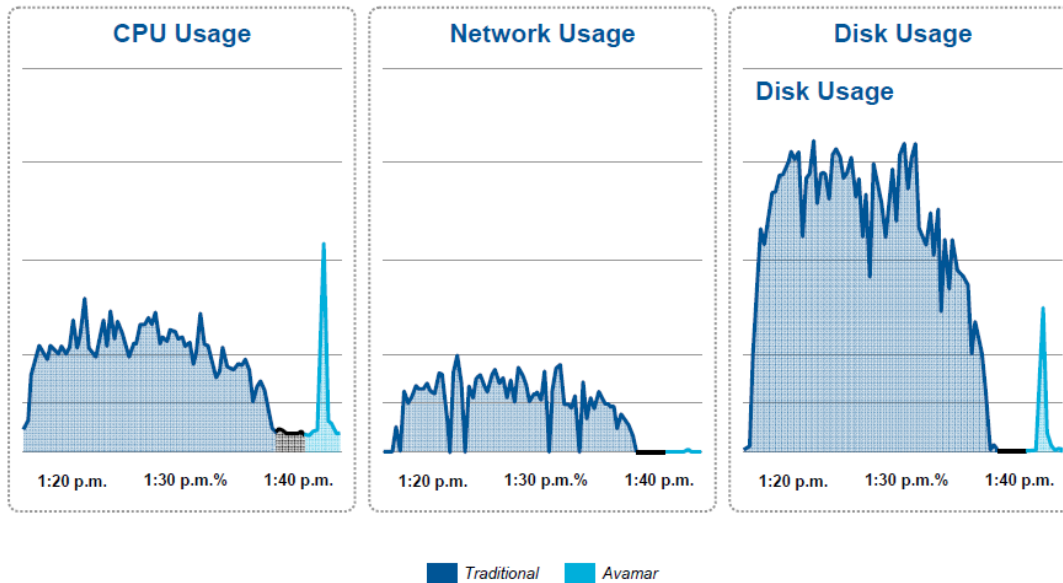


Figure 4. Backup for a VMware guest – Avamar vs. traditional backup

Solutions

Avamar for virtual infrastructures

Different data center environments favor different solutions. EMC Avamar in a VMware Infrastructure environment offers you the flexibility of implementing your data protection solution in a variety of ways:

- **Guest/file-level backup:** Avamar Agent is installed inside the virtual machine.
- **Image-level backup via vStorage APIs for Data Protection:** Avamar Agent is installed on a proxy server.

Guest-based backup

Guest-level backup involves installing the lightweight Avamar Agent inside each virtual machine.

Backup configuration for this method is no different from that for a physical server. Usually no scripting is needed for this type of backup. Configuration beyond basic client setup might be needed to support a specific application, such as Microsoft SQL Server or Exchange, or Oracle. The main advantages of this procedure are as follows:

- Highest level of data deduplication
- Support for backup of applications inside the virtual machines
- Support for partial or file-level restores
- Identical backup methods for physical and virtual machines
- No requirement for advanced scripting or VMware software knowledge
- Unchanged day-to-day procedures for backing up

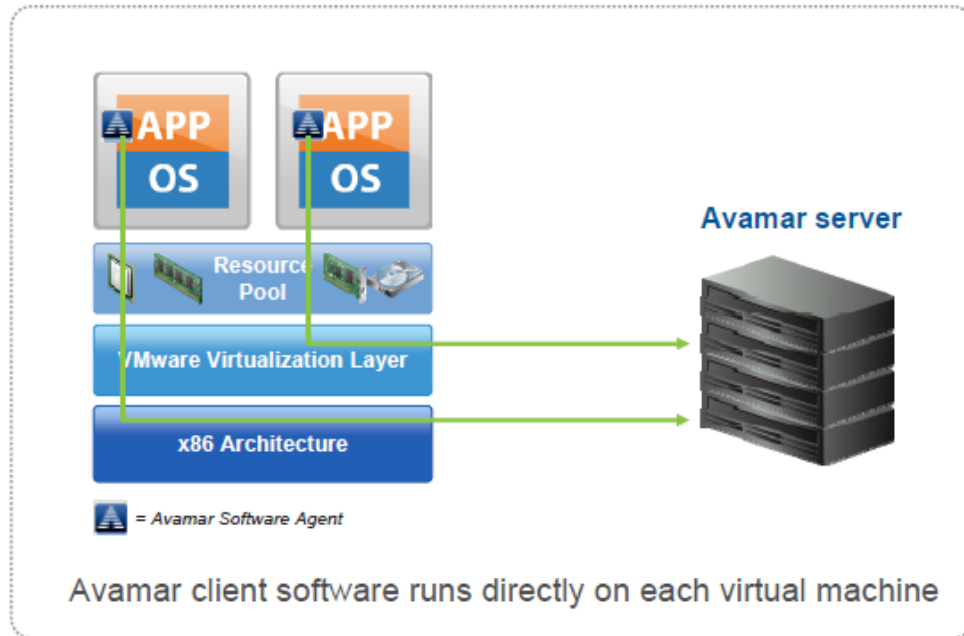


Figure 5. Guest-based backup

Restoring virtual machines in this configuration is easily accomplished. Figure 6 shows the method for restoring a full virtual machine image.

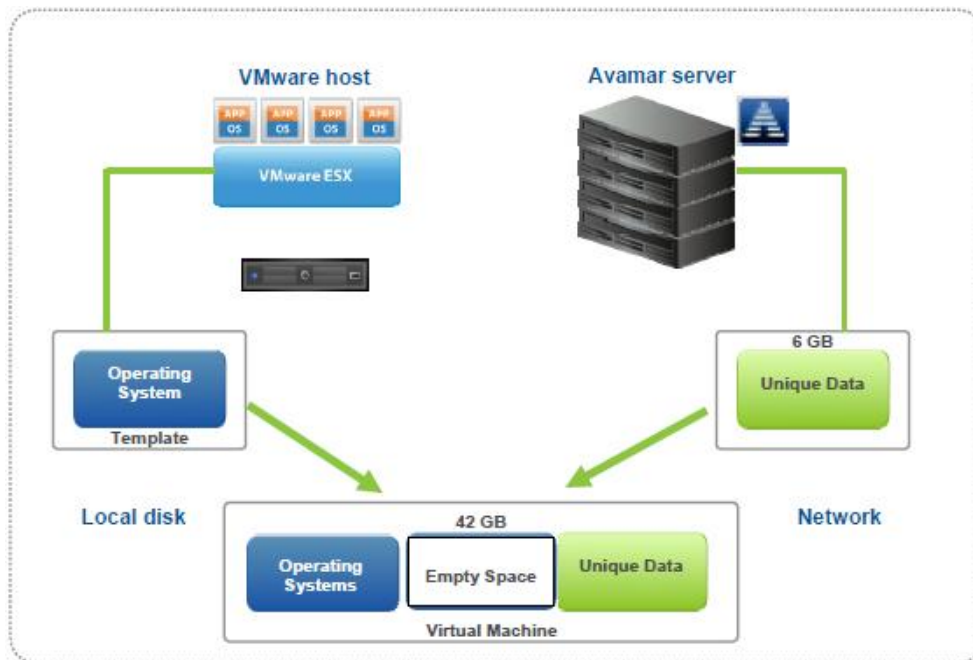


Figure 6. Restoring an entire virtual machine

To perform a full system restore (sometimes called a bare-metal restore), the system must create a virtual machine with the operating system and Avamar Agent installed. This step is most easily achieved either by using a template or a ready-made operating system image. To restore an entire file system on the server, take the following steps:

1. Deploy a new virtual machine from a template or image.

2. Power on the virtual machine and register it with the Avamar server.
3. Perform a redirected restore to the new virtual machine.

Image backup based on vStorage APIs for Data Protection

VMware's vStorage APIs for Data Protection enable LAN-free backup and offload the backup workload to a backup proxy server. The vStorage APIs' proxy server can mount a virtual machine's .vmdk files, and provide either a .vmdk backup or a file-level backup (Windows and Linux only) to provide recoverability of the entire image or set of files. Using the Avamar Agent to back up the mounted virtual machine disks, Avamar provides data deduplication at both the file level and the .vmdk level.

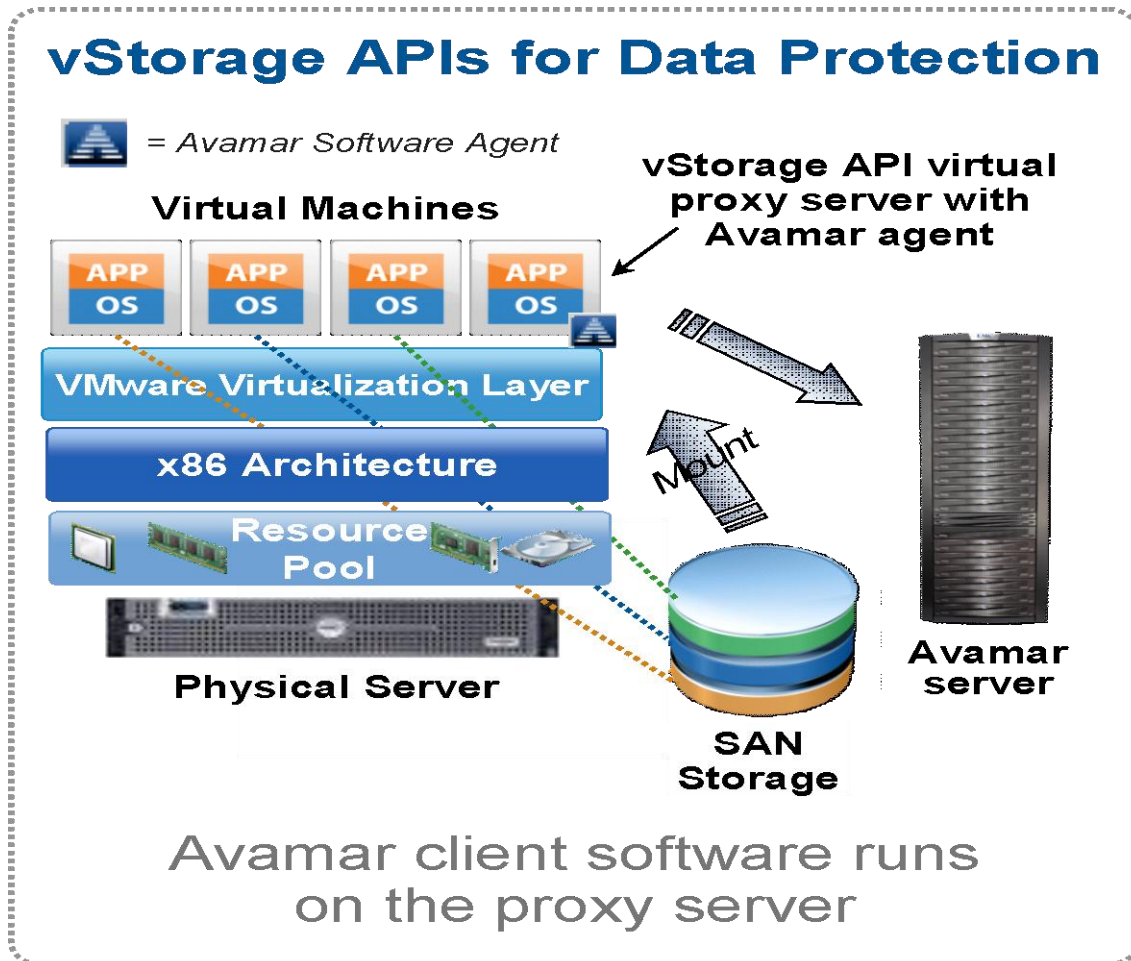


Figure 7. vStorage API backup

VMware's vStorage APIs for Data Protection consist of a set of utilities and APIs that work in conjunction with Avamar Agent software and the Avamar Interoperability Module (AVIM). The Avamar Agent and the AVIM run on the proxy to provide the backup services. Actual backup of the virtual machine happens on the backup proxy server. One backup server can provide backup services to many virtual machines on multiple ESX hosts. The backups run efficiently (hotadd) when all the virtual machines are stored on storage area networks (SANs) accessible by the proxy.

Avamar integration with the vStorage APIs for Data Protection and the AVIM leverages the vStorage APIs to create snapshots and to mount and unmount the snapshots — point-in-time copies of the running virtual machines. When the Avamar solution initiates a backup according to the schedule and policy you specify, the Avamar Agent on the proxy server initiates the backup activity.

The advantages of using Avamar and the vStorage APIs for Data Protection include:

- Provides full image backups of running virtual machines
- Utilizes efficient transport (hotadd), avoiding copying entire virtual disk images over the network
- Provides file-level restores from image-level backups for Windows and Linux
- Deduplicates within and across .vmdk files
- Uses changed block tracking for faster backups
- Minimizes network traffic by deduplicating and compressing data
- Eliminates the need to manage backup agents in each virtual machine for most scenarios

vCenter integration

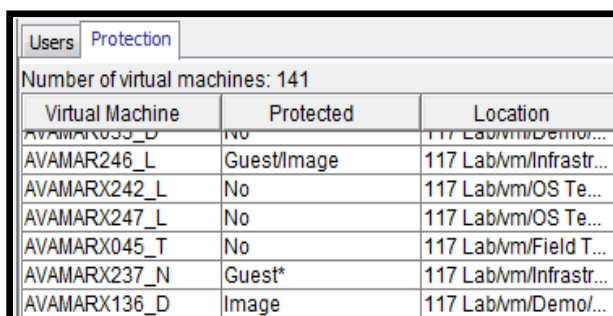
Avamar 5.0 provides unique integration capabilities with VMware’s vCenter management utility. The integration to the vCenter APIs allows the Avamar Management Console to routinely query a single instance or group of vCenter instances, and provide key data protection information that will simplify management of backup and recovery activities.

Key Avamar 5.0 features with vCenter include:

- Integration provides discovery of VMs and their associated groups in the Avamar UI
- Ability to add individual VMs or groups and define “image” and/or “guest” backup policies
- Ability to define multiple VMware Image Proxies
- Ability to initiate VM image or guest backup/restore operations
- Ability to monitor backup/restore operations in the Activity Monitor
- Ability to view VM protection status (guest, image, or none)

Key benefits of the Avamar 5.0 vCenter integration include:

- Simple views of whether VMs have been backed up or not, and the simple ability to remediate
- Shows *how* a VM was backed up (guest, VM, or not at all) and when
- Automatically adding a backup policy to virtual machines as they are added



Virtual Machine	Protected	Location
AVAMAR035_D	No	117 Lab/vm/Demo/...
AVAMAR246_L	Guest/Image	117 Lab/vm/Infrastr...
AVAMARX242_L	No	117 Lab/vm/OS Te...
AVAMARX247_L	No	117 Lab/vm/OS Te...
AVAMARX045_T	No	117 Lab/vm/Field T...
AVAMARX237_N	Guest*	117 Lab/vm/Infrastr...
AVAMARX136_D	Image	117 Lab/vm/Demo/...

Figure 8. Determine a data protection approach by virtual machine

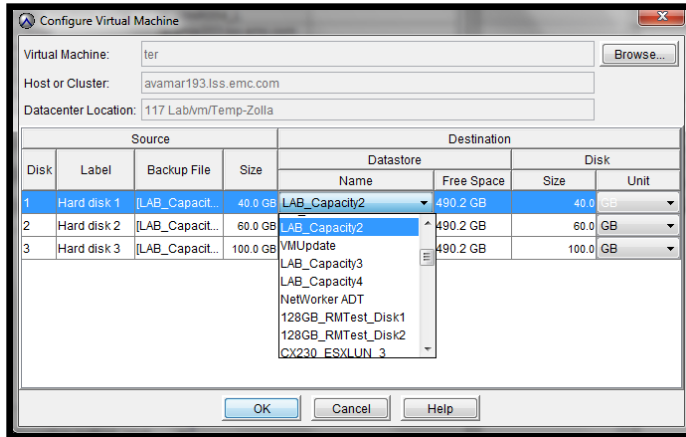


Figure 9. Ability to recover .vmdk files to different storage

Avamar for VMware View

Data Protection Strategy No. 1 for a VMware View environment

There are two strategies for protecting an entire VMware View environment. The first approach for protecting a VMware View infrastructure can be accomplished by using a combination of hardware and software to make a simultaneous copy of the LUNs where the virtual machine desktops are stored, as well as the View Manager application and configuration information. This solution allows for a total recovery of the entire environment.

Requirements for this approach involve making a duplicate copy of the VMware View environment for presentation to the backup solution. Recovering a VMware View environment in this scenario requires the following steps:

1. Restoring the data contained on the LUNs of all components to new storage of equal size
2. Rebuilding Active Directory, View Manager, and ESX
3. Connecting them to the recovered storage
4. Restarting the applications
5. Verifying configuration

Data Protection Strategy No. 2 for a VMware View environment

The second approach for protecting an entire VMware View infrastructure is to protect the key components independently using the Avamar client software agents. This is considered the best option for enterprises that want to avoid the infrastructure cost of duplicating the entire VMware View solution for presentation to the backup environment. Figure 10 represents the steps toward protecting the key VMware View components.

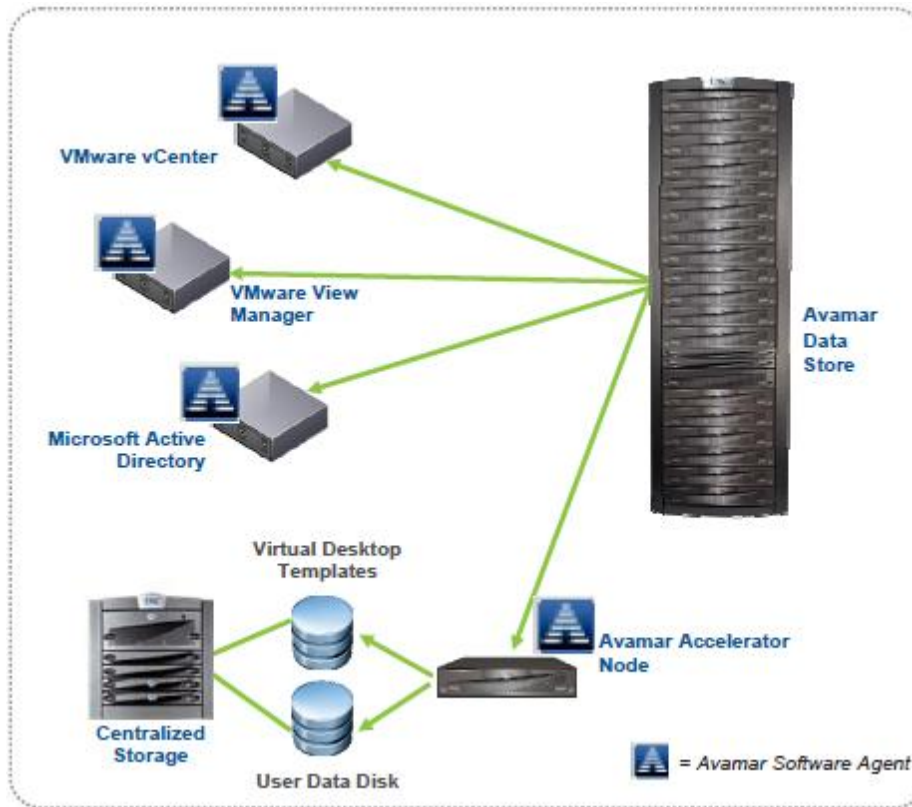


Figure 10. Individual component protection

Microsoft Active Directory (AD)

- Install the Windows Avamar client on the AD server.
- To restore AD to the same hardware, Avamar uses the ntbakup utility to export AD data for backup.
- To restore AD to different hardware, a combination of Avamar and EMC HomeBase™ is required.

VMware View Manager server

- Install the Windows Avamar client on the VDM View Manager server.
- Using a pre-script, Avamar calls the VDMexport.exe application to export LDAP configuration information located on the AD server to a flat file.
- Avamar backs up the LDAP flat file and the View Manager application located in the following location: \VMWare\VDMS

VMware vCenter

- Protecting vCenter requires a Windows Avamar client running on the server to protect the application and a database plug-in to protect the VC database.

Note: If the database is stored on a different server, a second Avamar client with a database plug-in on the database server is required.

Virtual desktop templates

As a best practice, user home directories and virtual desktop templates should be stored on a centralized shared storage device.

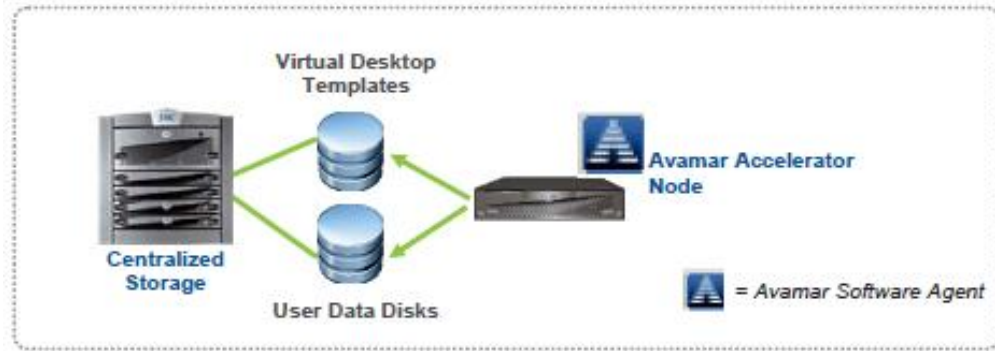


Figure 11. Desktop templates' and user home directories' data protection

Avamar data protection uses an accelerator node to protect SAN/NAS storage devices. Accelerator nodes can access data on the storage solution using NFS, CIFS, or NDMP.

Recovering VMware View components

This solution allows for the recovery of the individual components that make up the VMware View solution. Individual components can then be manually brought back into the VMware View solution based on the level of recovery required. Components are as follows:

Microsoft Active Directory (AD)

- Recover AD using the procedures documented in the *EMC Avamar System Administration Guide*.
- It is important to note that a HomeBase/Avamar solution is required to recover AD to different hardware. If recovering to identical hardware, use ntbakup as described in the *EMC Avamar System Administration Guide*.

VMware vCenter

- Install or recover the vCenter application.
- Install or recover the vCenter database application (SQL Server 2005 by default).
- Recover the vCenter database using the Avamar database plug-in.

VMware View Manager server

- Install or recover the View Manager application.
- Recover the View Manager AD schema flat file and import it into AD using the VDMimport.exe utility.

Note: This step is required only if the AD server was recovered without the View Manager configuration data in its schema.

ESX VMware View servers

- Rebuild the ESX infrastructure that will be used for virtualization of the desktops.
- Install the View Manger ESX Server Agent on the ESX servers.

User data disks

- Recover the user data disks to the centralized storage solution.

Virtual desktop templates

- Recover the virtual desktop templates to an NFS share or directly to the ESX infrastructure.
- Import the templates using vCenter.

Avamar deployment options

Centralized virtualization

Avamar offers flexibility in solution deployments, depending on the specific use case and recovery requirements. There are two convenient deployment options—EMC Avamar Data Store and EMC Avamar Virtual Edition.



Figure 12. Centralized deployment options

For enterprise virtual environments where applications are deployed and managed centrally, Avamar Agents are used at the guest or vStorage API level, and data is stored on an Avamar Data Store, a highly available, prepackaged backup and recovery solution that integrates Avamar software with EMC-certified hardware for streamlined deployment. Once data has been moved to Avamar Data Store during the process, it can then be replicated to a second facility for disaster recovery purposes.

Distributed virtualization

For environments that have standardized on a VMware virtual infrastructure at remote sites, EMC offers the EMC Avamar Virtual Edition for VMware: the industry's first deduplication virtual appliance for backup, recovery, and disaster recovery. Avamar Virtual Edition enables a complete Avamar server to be deployed as a virtual appliance on an existing ESX server, leveraging existing disk storage (SAN, iSCSI, DAS). Backup and recovery are encapsulated and virtualized, and VMotion is supported for deployment flexibility, reducing demands on IT staff. Avamar Virtual Edition also provides cost-effective disaster recovery through secure, efficient replication.

For non-virtualized remote office environments, an entry-level Avamar Data Store configuration is ideal when fast, local backup and recovery are priorities but the sites have not yet been virtualized.

In both cases, replication can be used to get remote data back to Avamar Data Store in a centralized data center where data-center-class data protection can be applied.

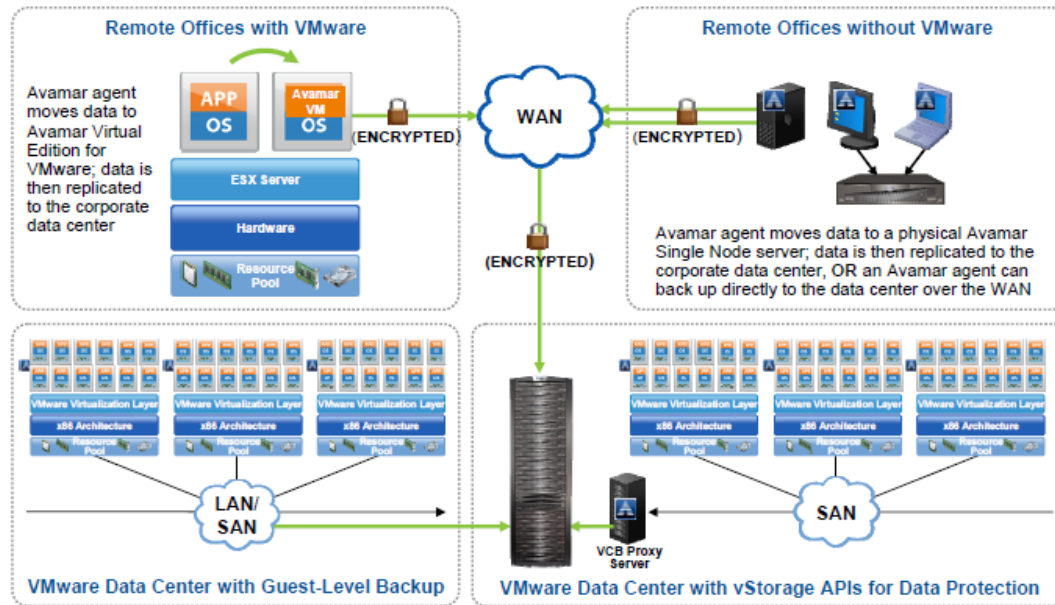


Figure 13. Avamar deployment options

Conclusion

The purpose of this white paper was to define the individual components of the VMware Infrastructure and VMware View solutions and discuss various approaches to protecting those components using Avamar backup and recovery technology.

In regards to protecting VMware virtual infrastructures, the predominant approach today has been to deploy the Avamar client software at the *guest level*, and realize the full benefits of deduplication at the source. This approach removes the costly burden that traditional backup approaches place on the shared resources, enables up to 10x faster daily full backups, and dramatically decreases the network bandwidth required to conduct daily full backups (up to 500:1). By removing the backup bottleneck, Avamar also optimizes server consolidation ratios and decreases overall infrastructure costs.

The second approach is to deploy Avamar in conjunction with the vStorage APIs for Data Protection, and utilize the Avamar deduplication capabilities at the proxy server level. This approach provides all of the benefits of deduplication at the source (lower infrastructure costs, faster backups, less network bandwidth required), but also removes 100 percent of the backup and recovery burden from the production servers onto the proxy server. Avamar's deduplication capabilities are exceptional in these environments when backing up entire .vmdk images, as the Avamar server will quickly remove redundant data relative to the operating system, patches, application, and so on, and will provide dramatically reduced infrastructure costs.

For VMware View environments, there were two approaches that were discussed. The first approach, which involves creating a duplicate storage-based copy of all of the underlying LUNs, is the least invasive to the applications during the backup process, as there is virtually no impact to the production application components; however, it does require a mirrored storage infrastructure that can be split and presented to the Avamar backup solution. The second approach involves using Avamar client software agents to back up each individual component of the VMware View infrastructure, and does not require a duplicate storage pool. In this scenario, the client agents will need to quiesce the database elements momentarily during the backup process, but will provide fast and reliable deduplicated backups of the data. In both scenarios, the unparalleled deduplication capabilities of the Avamar solution will drive down overall infrastructure costs dramatically, and will enable fast and reliable restore of any or all key components of the VMware View environment.

In summary, EMC Avamar with integrated global source data deduplication can dramatically decrease infrastructure costs related to media and network requirements, improve backup performance, and create efficiencies that will increase application consolidation ratios in many cases. Data availability will be increased as a result of the many data protection features built into the Avamar architecture, and restore capabilities will be dramatically improved as compared to traditional tape-based methodologies.

For more information on best practices for protecting a VMware Infrastructure and VMware View environment as well as the return on investment of an Avamar backup and recovery solution, please contact your local EMC or EMC Partner sales executive, or visit us directly at EMC.com.

References

For additional information on technologies discussed in this paper, the following resources are recommended:

- VMware VMware vSphere and VMware product pages on VMware.com
<http://www.vmware.com/products/vsphere/>
<http://www.vmware.com/products/view/>
- Avamar product page on EMC.com
<http://www.emc.com/products/family/avamar-family.htm>
- Powerlink[®], EMC's customer- and employee-only extranet (registration required)
<http://powerlink.emc.com>