

How to back up VMware 3.0.1: File-level and raw virtual machine file backup alternatives

By James Davidge and James Geis, Forsythe

With virtualization technologies taking a real hold in the IT world, the physical server layer is being transitioned to a virtual server infrastructure. As a result, many areas of management and maintenance that were previously straightforward and well-understood are now overshadowed with questions about best practices and alternatives. The key area that my customers often ask about is best methods and practices around backing up and restoring virtual machine instances.

In this tip, I'll focus on backup alternatives for a standalone VMware 3.0.1 environment and a method of performing backups to allow for easy restoration of this type of virtual environment. It will detail the processes of both file level and raw virtual machine file backup alternatives, and will give examples of how to configure and set up the backups for each of these backup methods.

Two types of backups in a virtual environment

There are two different types of backups in a virtual environment:

File level backup of the virtual machine's partitions is identical to the current process of backing up a physical machine's local or SAN connected disk. It is usually done by installing a backup agent in the virtual machine and backing up files to a backup server via the network. This process is good for virtual machine file level restoration, but lacks the ability to do a bare metal type of restoration if there is a VMFS3 disk failure, a virtual machine is accidentally removed from a disk, or the entire machine needs to be restored.

Backup of the raw virtual machine VMDK and configuration files stored on VMFS3 partitions allows for a complete (bare metal) virtual machine restoration. This is by far the fastest way to restore an entire virtual machine, but would be a bit cumbersome to use for file level restoration.

Misconceptions about backup agents

Many new ESX administrators believe that reliable bare metal backups of virtual machines can be accomplished by installing a Linux-based backup agent on the VMware ESX server itself. In reality, this is not a best or recommended practice. Pristine backups cannot be obtained with this

process on an active (running) virtual machine because the virtual machine isn't aware of backups taking place at the ESX server level, and, inevitably, the virtual machine will perform writes of new data to the disk file while the backup is in progress. Thus, this method can potentially result in an incomplete backup, services not starting, or virtual machine database corruption. In the worst case, the ESX will report that the virtual disk file is damaged and/or unreadable when attempting to start a restored virtual machine.

Also, performing backups through the ESX server uses CPU and memory which will not be available to virtual machines. This can cause performance issues on all virtual machines running on the server during the backup window.

Best practices for virtual machine backups

The next section will describe the best practices for virtual machine file level backups and virtual machine bare metal backups for a stand-alone ESX server environment configuration. My next article will focus on enterprise-class ESX environments and some advanced backup and replication techniques to further enhance backups.

The stand-alone single server environment with only local disk

We'll start with a single ESX server with no SAN connections, using a local disk to store the running virtual machines (VMFS3 file system), as shown in Figure 1. While this is definitely not an optimal configuration, some companies use this setup to kick the tires of virtualization and are not ready to integrate the ESX server into a back-end SAN or NAS based infrastructure.

Figure 1. VMware environment

Figure 1 details

Backup Server

Local disk drives. (Large enough to hold raw virtual machine files)

Running OPENSSE services

Backup software installed for file level backups.

ESX Server

Two physical Gigabyte Ethernet cards

One dedicated for management–(Bare Metal backup copies will happen through this interface.)

One Dedicated for virtual machine communications–(File Level backups will happen through this interface.)

Backup software installed for file level backups.

Single Disk partitioned as:

100 Meg boot partition

20 Gig root partition

2 Gig Swap partition

2 Gig log partition

200 Gig VMFS3 partitions where the virtual machine files are stored

Two running virtual machines

VM1

VM2

Virtual machine file level backup

There isn't much choice with the configuration in *Figure 1*, but to install backup agents on each of the virtual machines and to use the same backup process that is used for the physical server layer in the environment, as in *Figure 2* below.

Figure 2. Virtual machine client-based backup flow

Virtual machine raw file backup

File level backups are great, but it could happen that a drive fails and you need to get the entire virtual machine restored to a new server. With file level backups, the process for restoring a virtual machine would be:

Create a new virtual machine and install a base OS or deploy it from a template.

Install the backup agents on the virtual machine.

Restore the server's system state, services, and file systems.

Reboot the server.

This is very time consuming, and the server could still experience issues because of domain membership or other special tweaks that may have been made and/or not applied correctly.

The process for restoring a virtual machine's raw files is as follows:

Restore the virtual machine core files to any ESX Server.

Register the virtual machine with the server

Power on the virtual machine.

This process is much less time consuming and will have the failed virtual machine up and running in a fraction of the time of the file restore method.

The following process can be used to back up the raw virtual machine files from the ESX server and to insure a pristine backup of your virtual machine files. The trick to backing up the raw virtual machine files while the virtual machines are running is to use the integrated snapshot capabilities of the VMware ESX server. The snapshot command performs the following to the target virtual machine(s):

Tells the virtual machine to dump its current cache and memory.

Creates temporary delta files and VMDK used by the virtual machine to write any new data while the snapshot is in place and backups are being performed.

Freezes and takes the lock off of the primary virtual machine's VMDK file(s). This allows for a clean backup of the virtual machine's disk files with no need to install backup agents on the ESX server..

The virtual machine continues to run as normal and service is not disrupted.

To perform a snapshot of a virtual machine, you need to ensure that the VMware tools have been installed and are running on the target virtual machine. VMware tools have program modules that interact with the virtual machine's operating system and issue commands to dump memory and freeze any disk writes during the snapshot process.

Before taking the snapshot. Since the delta and swap files won't be backed up, you first perform a directory list and pipe the output to a file in order to script the backup process. Issue the following command to create this file:

```
ls /vmfs/volumes/storage1/VM1 > /tmp/filestocopy
```

Create the snapshot of the virtual server (VM1). Execute the following command from the console of the ESX server:

```
/usr/bin/VMware-cmd
```

```
/vmfs/volumes/storage1/VM1/VM1.vmx
```

```
createsnapshot dailybackups "Backups snapshot" 1
```

Breakdown of the command line:

```
/usr/bin/VMware-cmd-application executable.
```

```
/vmfs/volumes/storage/VM1/VM1.vmx-
```

name of the virtual machines configuration file.

Createsnapshot–command to create the snapshot

Dailybackups–name of the snapshot

"Backups snapshot"–Entry that is logged in the ESX and Virtual Center environment.

1–tell the virtual machine to dump memory.

Figures 3a and 3b show directory listings of the raw virtual machine disk files before and after the snapshot command

is issued. Note that in the second listing there are some new delta files created to hold any new disk or memory writes.

Figure 3a. Pre-snapshot directory listing

Figure 3b. Post-snapshot directory listing

Back up the virtual machine files. Now that the virtual machine is snapped, you can back up the virtual machine files. As noted in the environment overview, we have set up OpenSSH services on the backup server to securely copy the raw virtual machines files from the ESX server to the backup server. This eliminates the need to put a backup agent on the ESX server, minimizes the time to transfer the files, and minimizes the length of time that the load is placed on the ESX server.

Back up all files in the virtual machine's directory, excluding delta or snapshot files. Remember the file created prior to the snapshot—use this file with the SFTP command to transfer all of the pre-snapshot virtual machine files.

To perform the backup:

Execute the SFTP process by entering the following on the ESX server: `sftp -b /tmp/filestocopy root@".FTPServerIP`

Type in the backup server's root password.

Files will now copy to the backup server.

Note that with this configuration of the network layer of the ESX environment, you will use the management console network interface for this copy, and not the Virtual Machine Ethernet interface. This eliminates any possibility of affecting virtual machine performance due to network card utilization. *Figure 4* is a graphical representation of the snapshot-based backup flow utilized by this process.

Figure 4. Snapshot-based backup flow

Commit changes from the delta files. Once the backup is complete, commit any changes (writes) that have been placed in the delta files back into the original VMDK files. To do this, run the following command, which commits the changes from the delta files back into the VMDK files, and then deletes the snapshot and delta files from the disk:

```
/usr/bin/VMware-cmd  
/vmfs/volumes/storage/VM1/VM1.vmx removesnapshot
```

Breakdown of the command line:

- `/usr/bin/VMware-cmd`—application executable.
- `/vmfs/volumes/storage/VM1/VM1.vmx`—name of the virtual machines configuration file.
- `removesnapshot`
- commit all delta files to virtual machine and delete snapshot delta files

That's it. Your raw files are now on the backup server, which can be backed up to tape or replicated to your DR site for safe keeping.

Perl script for the raw file backup

I've written a perl script that will perform this entire raw file backup process. Just download the file from the following link and save it in the `/tmp` directory of the ESX server.

Download the Perl script text document.

Disclaimer: This script is presented as-is. The writer, SearchServerVirtualization.com and TechTarget and/or Forsythe are not responsible for any potential damage the script may cause either in it's current state or if modified by the reader. The script has been tested in a lab environment but is not utilized for production use.

The command line to execute this script is:

```
perl snapvm.pl
```

Recommended third-party backup tool

One of the best third party tools to perform RAW virtual machine file backups is `esxRanger` from Vizioncore. A separate server is required to host this application, so this solution may involve additional hardware and software expenses. Information on `esxRanger`.

esxRanger Professional is the recognized industry-standard backup and restore solution for virtualized environments, and continues to lead the way in establishing disaster recovery standards for VMware Infrastructure 3 users. Administrators can schedule regular backups of either the full image, or just the differential, at appropriate intervals while the virtual machine is still running. Images can be stored either locally in the SAN or sent as compressed files over a WAN to remote locations to support disaster recovery strategies. *esxRanger Professional* also enables restore of individual files efficiently with an explorer tree through which users can easily identify and "grab" needed files.

More backup instructions to come

The next chapter of *How to backup VMware 3.0.1* will discuss backup methods for SAN-connected VMware ESX servers with Virtual Center as a management point. It will review technologies such as SAN-based snapshots, consolidated backup, and replication techniques for these types of environments. Stay tuned!

Published on SearchServerVirtualization.com, May 13, 2007.

James Davidge, solutions architect at Forsythe Solutions Group, possesses over 19 years of experience in enterprise strategy, analysis, design and implementation of Microsoft, Citrix and virtualization technologies. Davidge oversees assessment, design, and implementation planning phases of large scale

project to insure delivery of enterprise class solutions to Forsythe's customers.

James E. Geis is director of integrated solutions development for Forsythe. Geis developed Forsythe's unique information management framework—the roadmap Forsythe uses for information management, storage and server consulting engagements. He currently is focused on developing solutions to address the impact of virtualization on computing technology and the data center.



800-843-4488 | www.forsythe.com

© 2007 Forsythe Solutions Group, Inc. All Rights Reserved. Contents may not be reproduced, in part or in whole, without prior written permission from Forsythe.