

Virtually Effortless Backup for VMware Environments

Abstract: Backup and disaster recovery (DR) strategies protect data from corruption and failures — and ensure that if these types of events occur, companies aren't forced out of business. Data protection becomes even more important in a virtualized server environment, because each physical server serves as a foundation for multiple virtual servers. If a physical machine fails, multiple applications and processes are impacted.

However, standard agent-based backup processes aren't well suited to virtual server environments. VMware Consolidated Backup (VCB) improves virtual server backup, but there are a number of limitations. FalconStor Software has a solution that works with VCB while overcoming these limitations. Using this solution, you can quickly and efficiently back up virtual machines and retain transactional consistency, without management complexity.

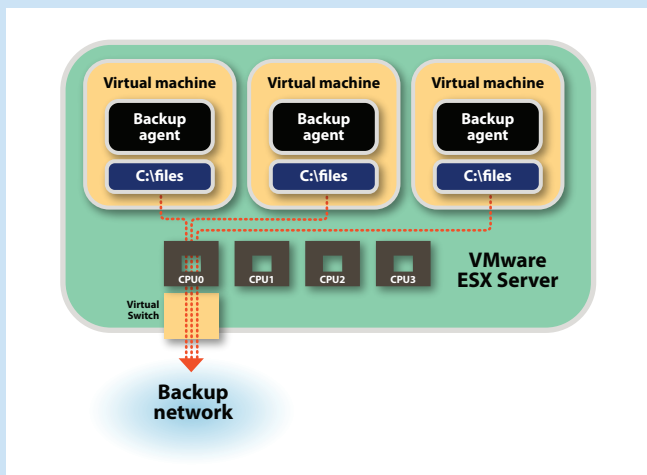
Why Standard Backup Processes Don't Work with Virtual Servers

The most common method of backing up virtual machines is to use backup software agents inside each virtual machine at the operating system level. These are the same decades-old agents used for physical server backups, so the process is well understood — but it doesn't fit the virtual machine paradigm for several reasons:

CPU and memory constraints. Running backups requires processor and memory resources as well as high I/O rates for both IP and storage networks. By design, these resources are much more highly utilized in a virtual server implementation than in a standard server farm. In fact, this more efficient utilization of resources is a key reason for virtualizing servers in the first place. For example, by consolidating multiple virtual machines onto fewer physical resources, you might increase a server's utilization from 15% to 95%, maximizing its usefulness. This offers a much better return on your investment in the server infrastructure, but it squeezes out the CPU and memory capacity needed for running agent-based backups. The upside of running physical servers inefficiently is that it leaves plenty of CPU and memory for running backups.

Downtime and complexity. Using standard backup procedures, the only way to create a clean, transactionally consistent image is to shut down the virtual machine before backing it up. This brings production to a halt while the virtual machine is shut down and storage-level snapshots are created. Coordinating the shut-down, snapshot, and backup processes requires either daily manual effort by your staff, or extremely complex scripting for each virtual machine individually. Productivity is reduced for users and IT.

Standard VM backup



Virtual machine backup is commonly done by installing standard backup agents within each virtual machine. Using this method, the backup agents can perform file-level backups from within the hosted operating system. This adds to software licensing costs and puts significant strain on system processing by funneling all backup traffic through a single CPU. With this method, full VMDK image restore is not possible.

Network bottleneck. Finally, VMware backups require that all network I/O processing be handled by a virtual switch, which operates via a single CPU. Therefore, even if your VMware ESX server includes eight CPUs, only one of them can assist with processing backup I/O. This results in a bottleneck that can impact your entire system. For this reason, a traditional network-based backup load is one of the worst loads for a VMware ESX server to handle.

Agent-Based File-Level and Image-Level Backup/Restore

Backup agents can perform file-level or image-level virtual machine backups. For file-level backup, the backup server can run as a virtual machine within the same VMware ESX server that is being backed up, or the backup server can be a separate physical server. Backup agents are installed within each individual guest application. These configurations place a heavy load on the VMware ESX server. In addition, they have an impact on the LAN, are difficult to manage, and may result in significant additional licensing costs for backup agents. A full virtual machine restore is not possible using this model.

Image-based backups place a Linux backup agent in the VMware Service Console. The backup agent "sees" folders that contain virtual machine files and executes folder-level backup. File-level restore is not available. While a full virtual machine restoration is possible, image-based backup has a negative effect on the VMware ESX server load, LAN, and manageability.

Image-level VM backup

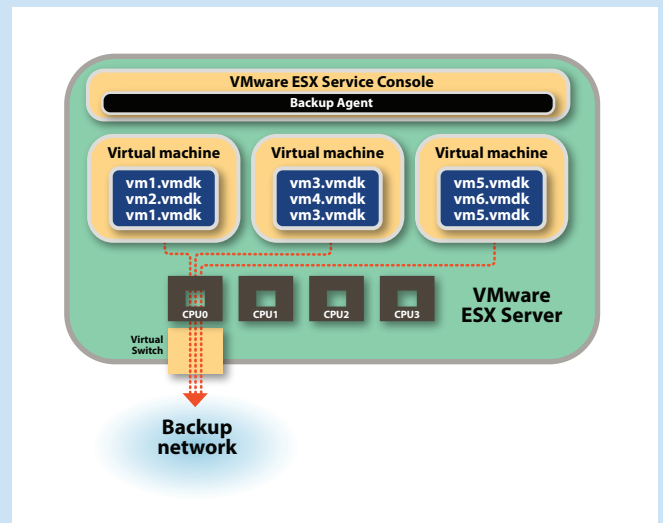


Image-level backup of virtual machines can be done by installing a Linux-based backup agent within the VMware ESX Service Console. This allows the agent to see the virtual machine files (such as VMDK files) and back them up in their entirety. Because the backup agent cannot see the files at the OS level, file-by-file backup and restore is not possible. This method causes significant strain on system processing by funneling all backup traffic through a single CPU.

Furthermore, crash-consistency of data is a challenge. Backing up files while a VMware ESX server is running — a “hot” backup — is commonly done to avoid interrupting operations. Although this provides a data image for restore, that image is crash-consistent. Should you need to restore, the system performs a consistency check as if a crash occurred. It can take a long time to resume operations while applications perform consistency checks. Databases, messaging applications, and file systems are likely to require extensive, cumbersome, and time-consuming repairs. You have an image to restore, but it will require costly downtime and make meeting your recovery time objectives (RTO) and recovery point objectives (RPO) unlikely.

VMware Consolidated Backup (VCB)

To address the limitations of traditional agent-based backups, VMware designed VMware Consolidated Backup (VCB), which has two key advantages: 1) it eliminates the need for backup agents on the virtual machines during the backup process and 2) backup is moved from the LAN to the SAN. The elimination of agents during backup makes it simpler and more flexible. However, agents are still required inside virtual machines for file restore. VCB speeds up the backup process by moving it to the SAN, thereby removing impact from the LAN. Both file- and image-level backups can be executed using VCB, but each requires a separate operation. You cannot achieve both file-level and image-level backup in a single operation.

With VCB, a Microsoft Windows server runs your backup software and functions as a VCB “proxy server.” VMware snapshots are used as the backup source. For file-level backup, VCB takes a snapshot, mounts it on the proxy server, and backs up to the tape library from there. Data is moved quickly over the SAN connection. For image-level backup, this snapshot is used to copy the virtual machine to the proxy server. The backup software then moves those files from the proxy server to tape. Depending on file size, the copy process can be very time-consuming.

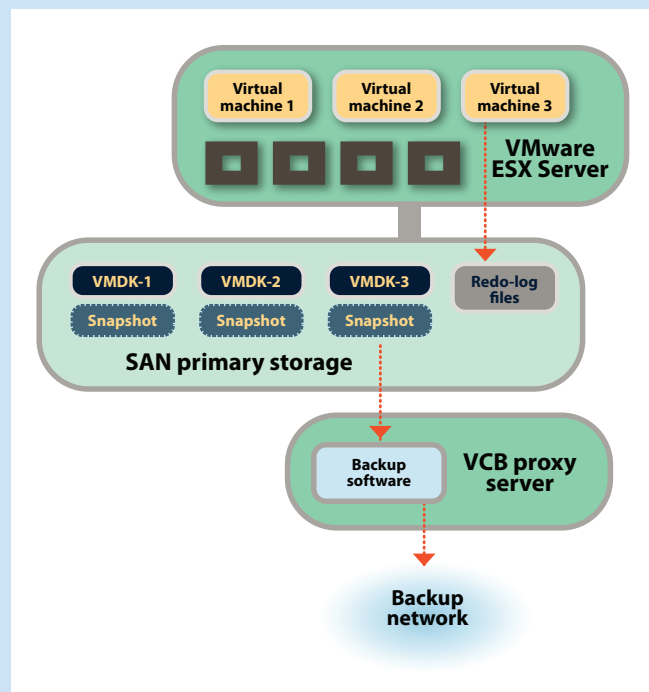
Limitations of VCB

The VCB snapshot process presents some challenges, including increased complexity, a delay in resuming full operations, and limited scalability.

Performance impact. As snapshots of virtual machines are taken, the virtual machines continue to run on the VMware ESX server. However, data is no longer written to the primary storage device. Instead, in order to make sure the snapshot image remains static, VMware ESX Server creates a separate “delta disk” or “redo log” to capture any new data written during the backup process. Once the backup is complete, the changes captured in the redo log must be written back to the primary storage device, a process that can significantly impact performance of the production environment.

Backup window. While recommitting the redo log, a second redo log is created temporarily to capture the new writes occurring while the previous snapshot flushes all of its data. This flushing process can take several hours or even days if the VM experiences heavy I/O during the flushing process. In fact, because production I/O is given

VCB backup



VCB backup eliminates the backup impact from virtual machine processing by moving the backup to a VCB Proxy Server. However, there is still significant impact on storage I/O. During the backup, all new changes are written to the redo-log. When the backup completes, this data must be written to the primary storage, which impacts performance. As a result, VCB backup does not scale well.

priority, committing of redo logs may be timed-out or canceled. There can also be a cascading effect: while the second layer of redo logs is being committed to disk, a third layer may be required. If the flushing has not been completed by the next day’s scheduled backup, it may be necessary to skip the backup job that day, leaving data highly vulnerable and well beyond recovery point service level agreements (SLAs). This type of bottleneck makes VCB a risky option, particularly for enterprise environments.

Limited scalability. Because the VMware Virtual Machine File System (VMFS) is a clustered/shared file system, virtual machines in a single VCB backup job may not all run on the same VMware ESX server, even though their virtual disks may be hosted on the same VMFS volume. This complex geography causes conflicts that limit the scalability of VCB — it can only back up four to five virtual machines at a time. In addition, it is not recommended that virtual machines be moved using VMotion or Distributed Resource Scheduler (DRS) while the snapshot is active, as data loss can occur. Therefore, if a backup job takes four hours, the virtual machine should not be moved to another VMware ESX host for at least that long. This can create unbalanced CPU and/or memory loads among VMware ESX servers that reside in the same DRS cluster while also decreasing system protection.

While VCB is an improvement over traditional backup methods on virtual machines, it is not a panacea. Remaining challenges include:

- > Due to the complexities of the snapshot process, it may take a long time for ongoing production writes to be recommitted after the backup.
- > Movement of virtual machines must be halted during the backup and recommit processes, limiting flexibility and resiliency.
- > The VMware ESX server is spared the heavy processing load of traditional backup methods, but disk resources are still significantly impacted with VCB.
- > Backup agents are no longer needed for backup, but they are needed for restore (in the virtual machine for file-level restore, or in the service console for image-level restore). This adds complexity and licensing costs back into the restore operation.
- > Image-level restores must be moved to a temporary location and then imported to the VMware ESX server, adding to the recovery time.
- > Image-level restores are only crash-consistent.
- > You must configure your SAN appropriately to zone the proxy server to the same LUNs as the VMware ESX servers.

Eliminate Problems with FalconStor Technology

FalconStor Software, the only provider of TOTALLY Open™ data protection solutions, has developed solutions that dramatically improve VMware backup. The FalconStor® HyperTrac™ Backup Accelerator for VMware Consolidated Backup (FalconStor HyperTrac for VMware) and FalconStor Application Snapshot Director (ASD) for VMware can be used with FalconStor Network Storage Server (NSS) to improve data protection, increase backup efficiency, save time, eliminate impact on VMware ESX servers, and ensure that data is fully recoverable.

Improving VMware Snapshots with FalconStor ASD

As we have seen, virtual machine backup presents two major challenges. The first is a crash-consistent image that can result in data corruption, data loss, or an extremely slow recovery process delineated by consistency checks, repairs, and application reinstallation. The second is that backups require a heavy utilization of resources — including snapshot use for the duration of the backup. This can significantly impact production operations.

FalconStor ASD addresses both of these issues. It ensures the transactional integrity of your data and eliminates the need for consistency checks and repairs. FalconStor ASD coordinates the snapshot process between the virtualized application and the VMware ESX server, and works in conjunction with VMware Site Recovery Manager (SRM). In addition, it reduces the impact on production operations because the VMware snapshot is used for only a few seconds, not for the duration of the backup process.

Application awareness. FalconStor ASD works in conjunction with storage virtualized by FalconStor NSS, letting you provision and grow storage resources easily across multiple arrays and connection

protocols, bringing your storage the same flexibility that VMware brings to your servers. FalconStor ASD is installed on the VMware ESX Service Console, and snapshot agents are installed within each virtual machine. Application-specific agents for various messaging and database applications (Microsoft Exchange and SQL Server and databases from Oracle, IBM, SAP, and others) place these applications into backup mode and ensure that current transactions are written to disk. In addition, a Microsoft Windows or Linux file system agent ensures that the file system is placed in backup mode and that transactions are flushed out of cache. This ensures that all disk activity has stabilized before the snapshot is executed. These application- and file system-specific agents ensure 100% data consistency and eliminate the need for checks and repairs during restore.

VMware Infrastructure awareness. FalconStor ASD communicates actively with the VMware vCenter Server to understand the relationship between each VMware ESX server and its associated virtual machines. Since the virtual machines can move among servers, FalconStor ASD must remain aware of each running virtual machine's location. Snapshots are initiated at the LUN (or LUN group) level, and snapshot notifications are sent only to the specific VMware ESX servers that are involved in any particular LUN or LUN group being snapped. All of this is automated, with no need for operational intervention.

Better backup and recovery. These consistent, low-impact snapshots can collectively serve as a data source for VCB to free up your production volume. A VMware ESX server dedicated to backup operations mounts the snapshot image, which can then be backed up directly from that dedicated server. This ensures that data is consistent and both the primary server and primary data volumes are completely released from backup processing.

Combining these transaction-consistent FalconStor snapshots with VCB offers an important benefit: consistently recoverable image-level backup. File-level backup can be performed by mounting an image hosted by a dedicated VMware ESX server. High-speed backup is now possible over iSCSI or Fibre Channel (FC) SAN connections instead of over the LAN. In addition, the VMware snapshot “delta disk” is only needed for a few seconds. Primary storage is not impacted, and no SAN configurations are required to properly zone primary storage. Your backup process is more reliable, faster, and has little impact on your primary storage. In the event of a failure, the speed of these processes enables you to implement recovery (e.g., quick, easy return to production) instead of restore, which involves the lengthy process of streaming the data back to a target.

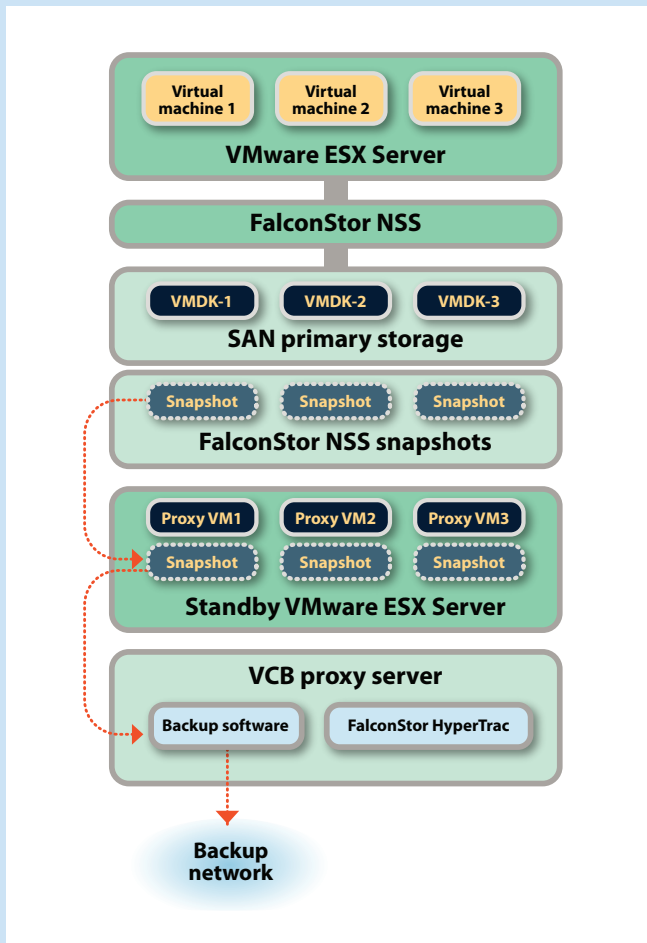
FalconStor HyperTrac Backup Accelerator for VMware Consolidated Backup

FalconStor HyperTrac for VMware is an extension of FalconStor NSS that enables you to offload VCB backup processes entirely from the production VMware ESX servers, freeing up valuable resources in your production environment and eliminating any impact on storage services. The backup window is eliminated, allowing backups to complete through VCB at any time — even during your busiest hours.

While VCB offers SAN-based backup support, it still inefficiently uses network and storage resources while VMware ESX servers are placed in hot backup mode. In addition, data changes are accumulated and

must be written back to the primary storage after backup is complete. The complex geography reduces scalability and can result in hours or days spent waiting for production writes to be recommitted in order for the process to complete. FalconStor HyperTrac for VMware eliminates the change tracking entirely, eliminating the delay as well as the production impact.

Backup with FalconStor HyperTrac for VMware



FalconStor HyperTrac for VMware enhances the VCB model by eliminating the impact on storage I/O. By using a snapshot image created by FalconStor NSS, it eliminates the performance impact of VMware snapshots and the need for a redo-log. As a result, backups can be run at any time with no impact to production systems. Because all processing is handled by dedicated systems, backup is a fast process.

The FalconStor solution essentially creates a “VCB island” that consists of a VCB proxy server and one VMware ESX server that hosts placeholder machines for each production virtual machine. These proxy components treat the FalconStor NSS snapshots like primary disks. Because these virtual machines are only serving as placeholders, they are always powered off. As a result, there is no production I/O hitting them and no need to create a VMware snapshot or redo logs. In addition, because the proxy machines are hosted by a single VMware ESX server, there are no conflicts that reduce scalability. While extremely fast and efficient backups are executed directly from FalconStor NSS over the SAN, production servers are free to perform their primary functions. As a result, FalconStor HyperTrac for VMware can extend VCB capability to enterprise environments and full VMware ESX server farms. Backup jobs are no longer limited to a handful at a time. This allows you to take full advantage of your backup infrastructure.

As a result, VMware ESX production servers are not impacted. Furthermore, there is no impact on the standby VMware ESX server. Consequently, the standby VMware ESX server can be shared by other virtual machines and other VCB backup jobs. Management is extremely simple, and no manual scripts are needed. In addition, because FalconStor HyperTrac for VMware is a transparent extension of FalconStor NSS, it can be implemented without any changes to your existing VCB or backup software applications. This method is both application- and virtual machine-aware, making it more consistent and immediately more easily restored than virtual machine-aware-only VCB snapshots.

Summary

Although VMware environments offer tremendous flexibility and increased resource utilization, protecting virtual machine data can be complicated, unreliable, and intrusive. The combined power of FalconStor technologies provides a solution that enables virtual storage and fast, simple, reliable backups without interrupting operations. Today’s organizations can count on FalconStor Software to deliver the storage intelligence that their storage environments need.

For more information, visit www.falconstor.com or contact your local FalconStor representative.

Corporate Headquarters
USA
+1 631 777 5188
sales@falconstor.com

European Headquarters
France
+33 1 39 23 95 50
infoeurope@falconstor.com

Asia-Pacific Headquarters
Taiwan
+866 4 2259 1868
infoasia@falconstor.com

