Solutions for AD Backup and Restore:

Picking the right type of solution for Active Directory Backup and Restore

WHITE PAPER

Semperis
Directory Services Protection Platform™

Preface

Backing up and restoring Active Directory Domain Controllers has caused serious headaches for many administrators. When virtualization was introduced, whether Hyper-V, vSphere or XenServer, virtualizing Domain Controllers causes even more problems than non-virtualized Domain Controllers in this area. Now, in the cloud era, it feels the heap of challenges around Domain Controllers only increases further.

Solving these problems before they harm the business is key, because in most environments when the Active Directory goes down, the entire network grinds to a halt. Authenticating into vCenter, your Single Sign-On-enabled cloud applications or even starting up the complete Hyper-V cluster can be extremely challenging without Active Directory.

But even if you've planned everything meticulously and deployed it flawlessly, there's still the need to operate the environment. A big piece in operations is continuity planning and the two big pieces in Active Directory continuity planning are replication and proper restores. For proper restores you'll need proper backups.

This whitepaper discusses proper backups for Active Directory Domain Controllers. It discusses the three main types of solutions in the market today to perform backups, and how to pick the solution you need to make sure the backups of your Active Directory Domain Controllers are the foundation of your restore ambitions.

About the author

Sander Berkouwer is a MCSA, MCSE, MCITP and has been a Microsoft Most Valuable Professional (MVP) on Directory Services and Enterprise Mobility for the last 9 years.

He is also a Veeam Vanguard. Sander calls himself an Active Directory aficionado and has done everything with Active Directory and Azure Active Directory... including decommissioning. Sander blogs on The Things that are better left Unspoken and ServerCore.Net.

Disclaimer

All content provided in this whitepaper is for informational purposes only and is provided "AS IS" with no warranties and confers no rights. The author makes no representations as to the accuracy or completeness of any information in this document or found by following any link in this document. The author will not be liable for any errors or omissions in this information nor for the availability of this information. The author will not be liable for any losses, injuries, or damages from the display or use of this information. This policy is subject to change at any time.

An overview of current solutions for Active Directory Backup and Restore

In the current market, we're able to distinguish two types of solutions for Active Directory backup and restore:

1. Host-based backup and restore solutions
2. Agent-based backup and restore solutions

The distinction is clearly on the way the backup is created and may be restored. Obviously, the first type of solution is only applicable to virtualized Domain Controllers.

Another distinction can also be made between solutions that backup to on-premises repositories and ones that backup to cloud repositories and allow for restoration into the cloud provider's Infrastructure-as-a-Service environment.

While some solutions offer all kinds of backup and restore methods and targets, throughout this whitepaper, we'll focus on two solutions: Azure Site Recovery Services and Semperis Directory Services Protector for Active Directory. This way, we'll showcase the strengths of two strong visions on Active Directory Backup and Restore and how they might benefit organizations to accomplish these tasks.

Agent-based vs. host-based

A decade ago, when virtualization wasn't widespread, agent-based backups were the way to go for all systems. There would be one central backup repository. Agents would be needed to be installed and configured on each system you wanted to backup. If the system was a special kind of system, like a SQL Server or Active Directory Domain Controller, a specific add-on agent license was often needed. Most backup solutions offered a centralized console for monitoring backups, creating boot media and initiating restores.

These agent-based solutions are still around and they've been developed further. However, in the meantime, a new type of backup solution has emerged on the wings of server virtualization: host-based backups. This type of backup can only be used for virtual machines. In this type of solution, the virtualization host is responsible for the backup, instead of an agent in the virtual machine. It most commonly works together with storage providers and Microsoft's virtual shadow copy service to create snapshots of virtual machines.

Now, whenever I mention snapshots to Active Directory folks, people start to cringe. For a long time, Active Directory restores suffered from inadequate backup solutions, that were mere snapshot solutions. I'll explain how these solutions mangle Active Directory Domain Controllers in the next chapter.

On-premises vs. the cloud

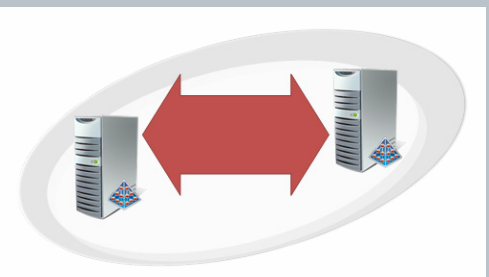
Five years ago, we couldn't imagine the cloud taking off like it did. For most organizations the question is not if they'll consume their e-mail functionality from the cloud, but when. While e-mail is a good example of highly-coveted cloud functionality, we're seeing more and more organizations consume the functionality of the backup repository from the cloud. It makes sense from an agility and cost-perspective, but I guess we'll have to dig a bit deeper into cloud-based backup solutions to find out if they actually make sense for Active Directory Domain Controllers. That's the last chapter of this whitepaper.

Challenges related to Domain Controller backups and restores

Over the years, we've seen a lot of issues with image-based backups of Domain Controllers. When you're running an environment with a single Domain Controller you won't encounter issues with this kind of backups, but when you're running multiple Domain Controllers (a best practice), you might encounter USN Rollbacks. Also, when you don't take into account your tombstone lifetime, you might end up with Lingering objects.

USN Rollbacks

Domain Controllers replicate changes. Whenever a change occurs on a Domain Controller, the Unique Serial Number (USN) of that Domain Controller increases. Each Domain Controller records the USNs it sees of its replication partners. This is recorded in the High Watermark Table. Replication partners are denoted using Invocation IDs in this table. The combination of USN and Domain Controller is captured as the up-to-dateness vector.
When you restore a Domain Controller to an earlier state, you would restore the USN to an earlier state. This is called an USN rollback.
Since its replication partners have seen a future USN for the Domain Controller, no changes will be replicated out until the restored Domain Controller reaches the USN recorded in the High Watermark Table. The effect is that user accounts and computer accounts that are created on the restored domain controller do not exist on replication partners. Or, the password updates that originated there do not exist on replication partners.



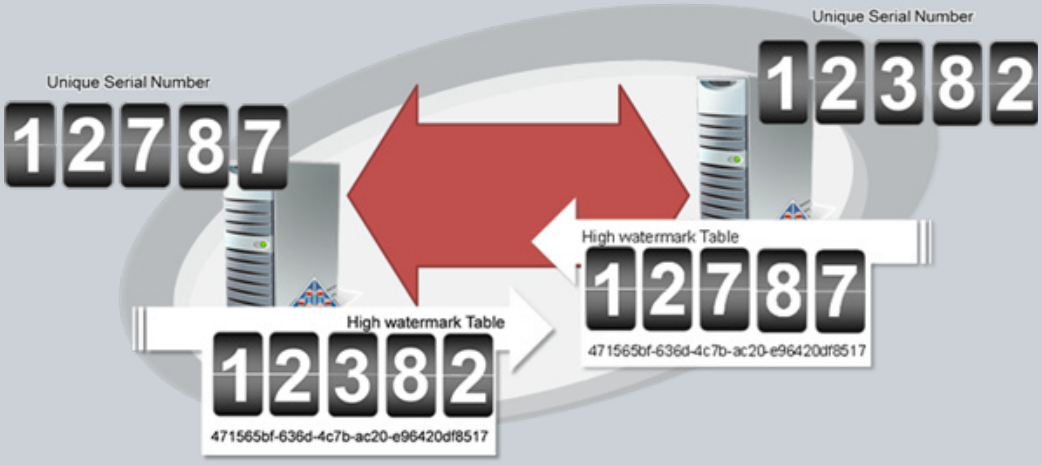
An example

Fabrikam has a Hyper-V host running Windows Server 2008 R2. Two virtual guests running Windows Server are hosted on the Hyper-V host. These two servers are Domain Controllers for the domain fabrikam.local. These two virtual servers are named DC1.fabrikam.local and DC2.fabrikam.local and are located within the same Active Directory domain and site. Replication occurs without problems.

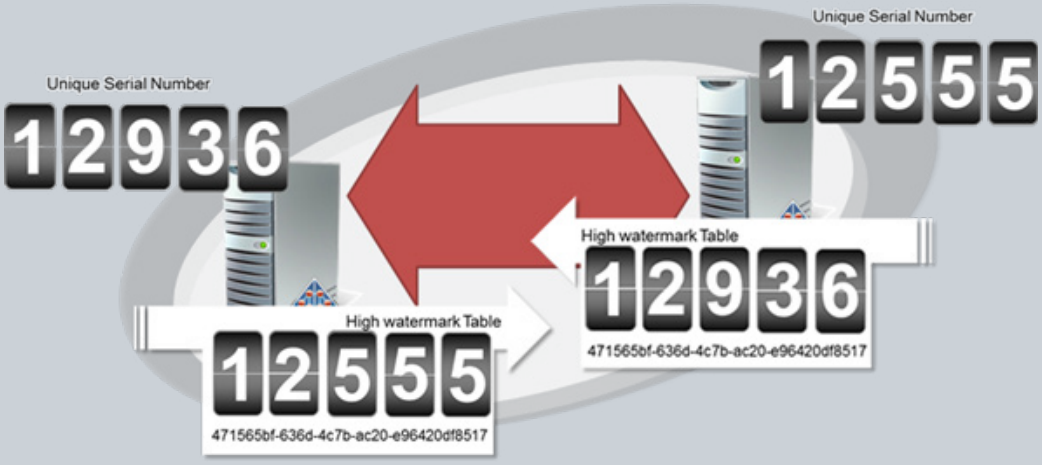
Looking at the High Watermark tables and Up-to-dateness vectors on each Domain Controllers, the following information becomes apparent:

Domain Controller name	USN	InvocationID
DC1.fabrikam.local	12787	de235686-7bc1-4412-941a-4f6e7e248be1
DC2.fabrikam.local	12382	471565bf-636d-4c7b-ac20-e96420df8517

This means DC1 knows all changes from its replication partner DC2 with InvocationID 471565bf-636d-4c7b-ac20-e96420df8517 up to USN 12382. DC2 knows all changes from DC1 with Invocation-ID de235686-7bc1-4412-941a-4f6e7e248be1. This looks like this:

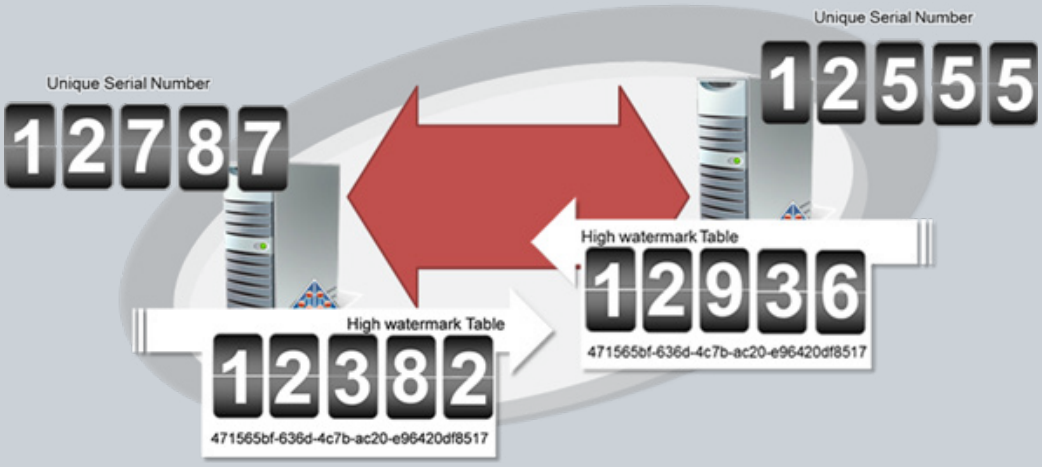


Suppose we make a backup of DC1 at this point and after the backup we make some changes inside Active Directory on DC1. We create some users, reset some passwords en create a couple of computer accounts. After replication the situation looks like this:



As you can clearly see the changes get replicated without problems and the Unique Serial Numbers (USNs) on both Domain Controllers get updated.

Now the time has come to restore DC1 from the backup. Restoring the Domain Controller to this previous state will roll back the Unique Serial Number of this Domain Controller to the value it had at the time of the backup. Graphically, this looks like this:



Now we have a problem.

DC2 knows DC1 has replicated changes all through USN 12936. Even worse, it knows the last changes have originated from DC1, so it's not replicating back these changes. DC1 and DC2 both own their version of the truth. As you make changes in Active Directory on DC1, these changes do not replicate to DC2, until the USN of DC1 reaches a higher value than the USN recorded in the High Watermark Table on DC2.

Users might have different passwords, determined by the Domain Controller they authenticate against. Some user and computer accounts might not even exist, depending on the authenticating Domain Controller, resulting in weird problems and possible security issues.

To prevent real harm, when an Active Directory Domain Controller detects another Domain Controller has been reverted to a previous version, as compared to the up-to-dateness vector, it will stop replicating to that Domain Controller.

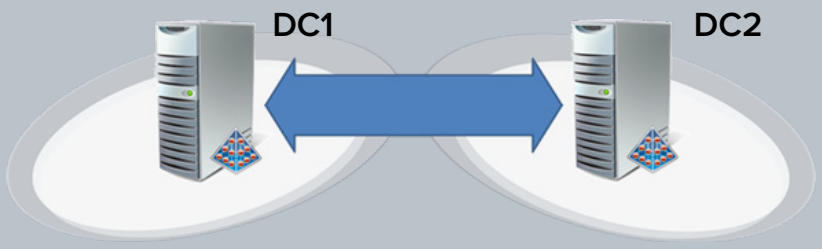
Lingering objects

When you delete an object in Active Directory it doesn't get deleted, it gets tombstoned. In this process all but its most critical attributes (objectGUID, objectSid, nTSecurityDescriptor uSNChanged and sIDHistory) are stripped and the changes are replicated between Domain Controllers. Only after the tombstone lifetime, the object gets deleted. This deletion takes place every 12 hours by the Garbage Collection process per Domain Controller.

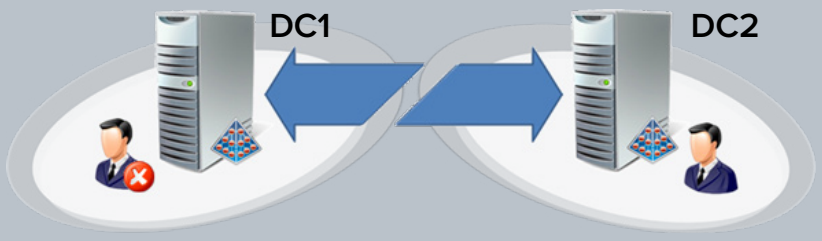
In normal situations, the tombstone process allows Domain Controllers to have sufficient time to replicate the tombstones. However, when you restore a Domain Controller to a point in time beyond the tombstone lifetime, the process may fail and objects that you expect to have been deleted may still exist on some Domain Controllers. These objects are called lingering objects.

An example

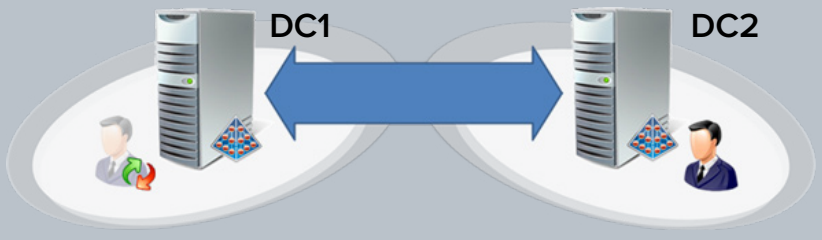
Contoso has two locations. On each location a Hyper-V host exists, running Windows Server 2008 R2. On each of these hosts, a virtual Domain Controller keeps the domain corp.constoso.com up and running. The two Domain Controllers are called dc1.corp.contoso.com and dc2.corp.contoso.com. They are placed within different Active Directory sites of the same Active Directory domain. Replication occurs without problems over the WAN link between the two locations. Admins work in the location where DC1 resides. The location with DC2 has a dial-in server.



The WAN connection between the two locations isn't very reliable (that's an understatement). One day when the WAN connection was failing an admin deleted a user object on DC1. This user object belongs to a salesperson who was found guilty of fraud. Because the WAN connection was down, this change was not replicated. At night both Domain Controllers were successfully backed up. In this backup, the following depicts the state of the user account:



The WAN connection eventually came back online, but unfortunately that night was the last night backups of DC2 were successful. Seven months later one of the Hyper-V hosts fails. It is the host responsible for DC2. A new Hyper-V host gets placed and all virtual machines are restored to their last backup. Now, the situation looks like this:



In the past seven months the objects tombstone event was replicated between DC1 and DC2.
After the tombstone lifetime had passed, the Garbage Collection process on both DC1 and DC2 had deleted all references to the user object.

Now, unfortunately, DC2 was restored to a moment in time where the user account still existed and DC1 did not receive the tombstone event. Suppose the USN rollback would have been taken care off, the object tombstone event will never reach DC2, since the object does not exist anymore on DC1.

The roaming user belonging to the deleted user account would be able to access resources through the dial-up server after this time. This, of course, poses a security risk.

Event ID 2089

While the above mentioned USN Rollbacks and Lingering Objects are truly frightening circumstances you don't want to find yourself or your organization in, you probably found this whitepaper, because you found a warning in the event viewer.

By default, Active Directory Domain Services logs Event ID 2089 when the Active Directory partitions have not successfully been backed up in a while. This period of time, by default, is set to half the tombstone lifetime.

About the tombstone lifetime

As explained as part of Lingering Objects, the tombstone lifetime is the specified period of time between replication of a deletion and the actual trigger for the garbage collection process to delete the object from the database.

Active Directory environments that were first setup using Windows 2000 Server or Windows Server 2003 as the Operating System for the Domain Controllers, as well as Active Directory environments that were migrated from Windows NT 4, have a default tombstone lifetime of 60 days. Active Directory environments that were first setup using Windows Server 2003 R2, or up, as the Operating System for the Domain Controllers have a default tombstone lifetime of 180 days.

The tombstone lifetime can be configured by an admin.

Picking the right solution

Getting the basics right

Backups are great. However, for Active Directory admins using Microsoft native tools, they should serve as a last resort. Microsoft has built in many great features into Windows Server that allow admins to avoid having to restore objects, to rely less on backups and to allow them to granularly restore objects and containers in Active Directory, if needed.

The below two features are built-in features of Windows Server. They're neither perfect nor fool-proof. While a strong Active Directory Backup and Restore solution would instill even more confidence into Active Directory admins to make changes, I strongly encourage Active Directory admins to turn these two features on for increased resiliency:

Protected from accidental deletion

In Active Directory since Windows Server 2008, objects and containers can be protected from accidental deletion. While this functionality looks like a simple option in the graphical Active Directory Users and Computers (**dsa.msc**) and Active Directory Administrative Center (**dsac.exe**) user interfaces, under the hood a set of 'Deny' access control entries (ACEs) is applied to the

object itself and its parent object for the 'Everyone' security principal.

Its basic functionality is to throw an Access denied error when you try to delete the object, either through the built-in graphical user interfaces, Active Directory PowerShell module or any other 3rd party or self-built tool. Before you can actually perform the deletion of the object, the checkmark has to be removed from the object, removing the 'Deny' ACE for 'Everyone'.

Any Organizational Unit (OU) created with the graphical tools and through the Active Directory PowerShell module will be automatically protected from accidental deletion. OUs created using **ldifde.exe** or **csvde.exe** will not. It's strongly recommended to protect all OUs throughout the Active Directory environment from accidental deletion. The Active Directory Best Practices Analyzer will keep pointing this out, too.

By default, only OUs are protected from accidental deletion. However, when you configure clustering, the cluster computer object is protected, by default, too. It's the only object, besides OUs, that we encounter in the field, that is protected from accidental deletion. I strongly encourage you to protect other critical objects throughout the Active Directory environment from accidental deletion, too.

Please note that the 'Protected from Accidental Deletion' feature only adds resiliency when admins delete objects. It does not offer rollback of modifications of attributes for these objects.

Active Directory Recycle Bin

The Active Directory Recycle Bin feature, that has been available for Active Directory since Windows Server 2008 R2. However, it wasn't until Windows Server 2012 that Microsoft shipped a graphical tool to restore objects and containers, when they introduced the Active Directory Administrative Center (**dsac.exe**). As Windows Server 2012 introduced a slew of Active Directory features, many organizations adopted the Active Directory Recycle Bin in that point of time.

While the Active Directory Recycle Bin is of no use when you need to perform a forest recovery, it might just save you a lot of time multiple times per week to recover deleted user objects, including their group memberships.

Under the hood, a new 'isdeleted' state was added to objects. Only after this period of time, which, by default, is equal to the tombstone lifetime, objects enter the 'tomstone', which is denoted as the 'isrecycled' state. During the 'isdeleted' state, objects keep their group memberships and other attributes, but are filtered out from view in the graphical Active Directory Users and Computers (**dsa.msc**) and Active Directory Administrative Center (**dsac.exe**) user interfaces, when using the Active Directory PowerShell module and 3rd party tools that support the feature. (When these tools use the native Active Directory calls, they support the feature.)

Please note that the Active Directory Recycle Bin, just like the 'Protected from Accidental Deletion' feature, only adds resiliency when admins delete objects. It does not offer rollback of modifications of attributes for these objects.

Benefits of host-based backups

Virtualization adds many benefits to running datacenters. Optimization of resources is often mentioned as a cost saving benefit. Backups, however, can also be added to the list of benefits. Virtualization Host-based backups can be a more cost-effective way to backup virtual servers than the traditional methods:

Fewer agents equals fewer licenses and fewer hassle

In a virtualized datacenter with a decent virtualization ratio a serious number of virtualized servers run on each virtualization host. Using the traditional backup method means installing and managing an agent to each virtual server. This method requires a license and program on each server. With host-based backups a license and program is only required per virtualization host or per environment.

Bandwidth optimization means less networking hardware

Host-based backups represent a bandwidth optimization when run through separate backup networks. In traditional non-virtualized environments to run backups over a dedicated backup network requires Network Interface Cards (NICs) for each server and networking infrastructure to tie everything together. Host-based backups require a single NIC per virtualization host. This equals less networking equipment.

Resource optimization means a smaller backup window

In traditional non-virtualized environments, backups are initiated sequentially from the media server(s) or written to a centralized backup location simultaneously. Host-based backups deliver the best of both worlds. Backups of virtual servers are initiated sequentially by each backup host and written simultaneously to a central backup location. Especially when deploying bandwidth optimization techniques, like change block tracking and deduplication, this represents a low backup overhead load per virtualization host, while benefiting from an optimized backup window.

Virtualized backups mean virtualized restores

A host-based backup is an ideal path to perform a fast disaster recovery restore. Since a virtualization host known everything about the configuration of its virtual guests, it can backup both the contents of the disks and the accompanying configuration, including Processor, RAM and NIC settings. Restoring this backup package can then be done on any virtualization host, running the backup program. In a non-virtualized datacenter, you'd have to purchase expensive and typical hardware running roughly the same specs as the original box. In an environment with backup agents you would either have to purchase expensive Disaster Recovery licenses or configure a virtual server with the same specifications, install the Operating System and the backup agent on it and then perform a restore.

Disadvantages of host-based backups

Of course, there are also drawbacks to the host-based model:

Host-based backups mean full host-based restores

Unless the host-based backup solution also incorporates (on-the-fly) agents, to restore Active Directory objects is to restore an entire Active Directory Domain Controller and then export the required information to live Domain Controllers, in case of granular restore needs. Tools like **csvde.exe** and **ldifde.exe** are often used in these scenarios to export and then import the required attributes for objects.

Licensing less-optimized environments may be more expensive

Backup licenses for host-based backups are not free. Typically, a license designed for virtualization hosts is six times costlier than a backup agent for non-virtualization hosts. In environments where the ratio of virtual machines per virtualization host is below six, it may be more expensive, from a license perspective, to run host-based backups.

Not all virtualization platforms are the same

While most Active Directory admins think of virtual Domain Controllers as running on their virtualization platforms. However, some virtual Domain Controllers may be running in public cloud environments. While Azure Infrastructure-as-a-Service is based on Hyper-V as the virtualization platform, Microsoft does not offer a way for its customers to benefit from host-based backups of the virtual machines running on it.

Benefits of agent-based backups

In environments with non-virtualized Domain Controllers, host-based backups cannot be used. In these scenarios, agents will need to be installed (just in time) onto Domain Controllers and typically they create snapshots of Active Directory and (parts of) the System State.
This describes the difference between host-based backups and agent-based backups, although the distinction is usually not this clear. Most host-based backups also offer agents to circumvent the drawback of having to restore full Domain Controllers by leveraging (just in time) restore agents to Domain Controllers.
However, the benefits of agent-based backups can be made crystal clear:

There's no distinction between virtual and physical Domain Controllers

From an agent's point of view, it doesn't matter if the Domain Controller is a physical box or a virtual machine running on your virtualization platform or as a cloud-based virtual machine; it will treat all Domain Controllers the same.

Granularity of backing up

When agents back up Domain Controllers, organizations can choose which parts of the Domain Controller to backup. This, as an example, allows organizations to misuse Domain Controllers as hosts offering other functionality as well, like a file-based source code repository, but not backup the repository as often as the Domain Controller functionality.

Granularity of restores

Restoring is where agent-based backup solutions outshine host-based backups. Unfortunately, this is also where some vendors leave their customers in the dark. Restoring Domain Controllers is somewhat of an art form to get right. Cleanups, additional steps and additional checks need to be performed around restores to make sure the outcome is Active Directory is in a healthier state.

About Semperis ADFR

Most vendors offer guides for Active Directory restores, but one vendor goes beyond. Semperis' offers an Active Directory Forest Recovery solution (ADFR) as part of its Directory Services Protection Platform offers local caching and off-site storage of Domain Controller backups and does away with all the complexity in Active Directory restores in three scenarios:

1. **Logical corruption of the Active Directory database on a Domain Controller**
2. **Attacks directed at Domain Controllers, like ransomware**
3. **Active Directory Forest Restore**

Semperis' ADFR is responsible for the Active Directory Forest Restore part on the platform's functionality. I feel Semperis ADFR Server is the prime example of how backups and restores should be done from a pure Active Directory disaster recovery point of view. However, an organization's IT environment typically consists of more than Active Directory... That's why another agent, like the Azure Recovery Services agent, might also be installed (just in time) to create backups, too.

In terms of agents, one agent, typically, doesn't exclude another agent, although they don't always get along nicely when they perform their backups jobs simultaneously.

The advanced features of Semperis' protection for the other two parts of functionality of the platform have a slightly different approach with many benefits of their agent-based approach: Their agent keeps a local copy of all the changes in Active Directory – not just deletions – and this local copy can be used to restore attributes, objects, and even entire organizational unit structures in mere seconds. Leveraging this feature, that taps into Active Directory through its native APIs, an Active Directory admin using Semperis' platform is able to beat the recovery times of other Active Directory Backup and Restore solutions hands-down.

Security is built-in

Most backup agents I come across offer security features, like encryption of data in transit and data at rest. With agents, encryption of backups can occur in a very early stage and most agent encrypt the backup data the nanosecond they access it.

Disadvantages of agent-based backups

Of course, there are also drawbacks to the host-based model:

The quality of the agent determines the quality of backups and restores

Because every backup and restore depends on the agent, the agent needs to be of high quality. When agents for a backup solution are present on a Domain Controller all the time, and contain a vulnerability, this might prove to be the right attack angle on your Active Directory... just like we've seen happen with vulnerable printer drivers and vulnerable anti-malware solutions.

Additionally, the amount of manual actions an Active Directory admin has to perform after restoring a Domain Controller in certain scenarios can also be used as a quality denominator.

Backups are sequential, mostly

Because agents communicate to a central repository, and mostly get prompted to create a backup by a central orchestration solution, you can expect most backups to be performed sequentially; one Domain Controller after another.

This will dramatically increase the backup window. Where a host-based backup solution uses the optimized networking stack of the storage system and you'd be tempted to add a second network interface adapter in Domain Controllers, too, this is not a solution, unless the Active Directory admin and networking admin work together. Otherwise, you might end up with Domain Controllers registering the wrong IP addresses for name resolution, etc.

Most of the time, though, networking is not even the issue, as backup agents come with optimizations like deduplication. However, these features require CPU cycles, in return.

Security based on Active Directory

The communications between agents and centralized repositories are encrypted, but for some agents, the certificates used for encryption might be issued by a Certification Authority (CA), based on Active Directory Certificate Services (AD CS), that might be configured as an enterprise CA within Active Directory Domain Services. This Catch-22 situation might make it really hard to restore.

Semperis Protection Platform

Semperis' Forest Recovery solution distributes backup jobs to each domain controller, allowing independent generation of backups. In addition, backups are compressed (compression ratio can be controlled), optionally encrypted an automatically copied to regional distribution points, without relying on orchestration from the central server.

The solution generates it's own internal certificates that allow it to secure communication between management server and Domain Controller agent without relying on an external CA or signing authority, and the restore process is fully automated, without the need for manual intervention in most cases.

Benefits of cloud-based backups

The cloud offers agile computing, networking and storage. Cloud-based backups have become popular fast. There's a couple of reasons behind this trend:

Pay-as-you-go

Typically, backup repositories are scaled in terms of storage using a factor to compensate for future growth. Throughout the lifecycle of the repository, this means a part of the storage of the repository is structurally left unused. It has been paid for, though.

For cloud-based backups, organizations typically pay for storage that is actually used by the backups. In some cases, storage deduplication and tiering are taken into account to lower the monthly bill even more.

Off-site, by definition

Many organizations struggle to find a suitable solution to get their backups off-site. However, by definition, a cloud-based backup repository is off-site. This helps organizations to mitigate the risks associated with their physical locations.

Gentlemen, test your restores

Testing a restore is trivial with a cloud-based backup solution. Since restoring means spinning up a virtual machine in the cloud provider's datacenter, it can be cost-effective.

About Azure Backup

Microsoft has seen the potential for this feature and makes it really easy to perform a test restore of your protected resources in the Azure portal.

However, as we found out with these test restores, Microsoft does not support FRS-replicated SYSVOL shares on protected Domain Controllers. It has never been part of their test beds and subsequently didn't make it in the service. If you want to make the best of Azure Recovery Services and Azure Backup with Domain Controllers, be sure to make DFSR the replication method for your SYSVOL shares.

Disadvantages of cloud-based backups

Where to restore?

The way a cloud-based backup makes the most sense is when you're able to restore the backed up system to the cloud provider. Of course, you could set up at least one system in your (new) datacenter – or actually any location with an abundance of bandwidth –, install the backup agent or boot up from the bare-metal recovery media and restore the system while transferring the required data over the Internet.

Some organizations don't feel confident to run Active Directory Domain Controllers in public cloud environment, yet. But that's where other organizations benefit when they restore. We're seeing reports of organizations lifting and shifting to Azure Infrastructure-as-a-Service through its recovery services.

Concluding

There are many backup solutions available that will make backups of your Active Directory Domain Controllers. These solutions can be put in the following categories:

1. Host-based backup solutions
2. Agent-based backup solutions
3. Cloud backup solutions

Your organization's policies, security standards and restore needs dictate the best solution, based on its needs.

While an Active Directory-aware backup solution enables consistent backups of your Active Directory, true Active Directory backup and restore solutions offer more than just consistent restores of Active Directory; granular, easy and meaningful restores, saving Active Directory admins time and hassle.