

Technical Brief

ReFS: Is It a Game Changer?

Presented by: Rick Vanover, Director, Technical Product Marketing & Evangelism, Veeam



OVERVIEW

Backing up data is more important than ever, as data centers store larger volumes of information and organizations face various threats such as ransomware and other digital risks. Microsoft's Resilient File System or ReFS offers a more robust solution than the old NT File System. In fact, Microsoft has stated that ReFS is the preferred data volume for Windows Server 2016. ReFS is an ideal solution for backup storage. By utilizing the ReFS BlockClone API, Veeam has developed Fast Clone, a fast, efficient storage backup solution. This solution offers organizations peace of mind through a more advanced approach to synthetic full backups.

CONTEXT

Rick Vanover discussed Microsoft's Resilient File System (ReFS) and described how Veeam leverages this technology for its Fast Clone backup functionality.

KEY TAKEAWAYS

Resilient File System is a Microsoft storage technology that can transform the data center.

Resilient File System or ReFS is a valuable Microsoft storage technology for data centers. Some of the key differences between ReFS and the NT File System (NTFS) are:

 ReFS provides many of the same limits as NTFS, but supports a larger maximum volume size. ReFS and NTFS support the same maximum file name length, maximum path name length, and maximum file size. However, ReFS can handle a maximum volume size of 4.7 zettabytes, compared to NTFS which can only support 256 terabytes.

Feature	ReFS	NTFS
Maximum file name length	255 Unicode characters	255 Unicode characters
Maximum path name length	32K Unicode characters	32K Unicode characters
Maximum file size	18 EB (exabytes)	18 EB (exabytes)
Maximum volume size	4.7 ZB (zettabytes)	256 TB

Limits

- The most common functions are available on both ReFS and NTFS. These include BitLocker encryption, cluster shared volume (CSV) support, soft links, failover cluster support, access-control lists, USN journal, changes notifications, junction points, mount points, reparse points, volume snapshots, file IDs, oplocks, sparse files, and named streams.
- A few features are only available in ReFS, and not in NTFS. These include BlockClone, sparse VDL, and real-time tier optimization on Storage Spaces Direct.
- Similarly, a few features are available in NTFS and not in ReFS. These include file system compression, file system encryption, data deduplication, transactions, hard links, object IDs, short names, extended attributes, disk quotas, bootable media, NTFS storage tiers, and support on removable media.



ReFS offers data integrity, resilience, availability, speed, and efficiency.

ReFS was introduced in Windows Server 2012 and is a component of Windows Server 2016. It can be deployed via three different models:

- 1. Simple ReFS Volume. This provides fast cloning, spaceless full backups, and data integrity streams.
- 2. Shared Storage Spaces (S2S). This suffers from poor performance (parity) and poor capacity (mirror).
- 3. Storage Spaces Direct (S2D). This adds real-time storage tiering with SSD. While it resolves S2S performance issues, it requires the Datacenter Edition of Windows Server.

One of the advantages of ReFS is that it is built to scale. It can hold 16 exabytes of data, compared to NTFS, which only holds 256 terabytes. NTFS also has a 64TB chkdsk volume limit and max file fragments issues. ReFS does not suffer from these limitations.

Microsoft has stated that ReFS is the preferred data volume for Windows Server 2016. ReFS improvements in Windows Server 2016 include:

- Data integrity. Checksums protect all filesystem metadata, while optional checksums protect file data. Checksum verification occurs on every read of checksum-protected data during periodic background scrubbing. Healing occurs as soon as corruption is detected. ReFS uses alternate healthy versions to automatically repair corruption.
- Resilience and availability. ReFS is designed to stay online and keep data accessible. It performs repairs without taking volumes offline. Backups of critical metadata are automatically maintained on the volume. The online repair process consults backups if checksum-based repair fails.
- Speed and efficiency. Efficient VM checkpoints and backup are now possible since operations between parent and child VHDX are a ReFS metadata operation. This means reduced IO, increased speed, and lower time taken. It accelerates fixed and dynamic VHDX creation, lowering VM deployment times. ReFS provides near instantaneous VM storage provisioning.

With enterprise storage, petabytes of data are common. Any IoT application, for example, will generate huge amounts of unstructured data. ReFS can easily hold that information. *Rick Vanover, Veeam*

Data integrity streams in ReFS prevent and detect data corruption.

ReFS data integrity streams defend organizations in two ways:

- 1. **Prevent data corruption.** ReFS ensures consistency through allocate-on-write, so data blocks are never modified in place. When data integrity streams are enabled, this functionality is enabled by default for metadata and file data.
- 2. Detect silent corruption. The ReFS data integrity scanner detects and notifies organizations about silent data corruption or "bit rot." Notifications are sent via scrubber.

To set and check data integrity streams, PowerShell commandlets can be run against the entire volume or at the file level. The relevant commandlets are Get-FileIntegrity and Set-FileIntegrity.



The BlockClone API efficiently moves data that exists on disk.

The BlockClone API within ReFS provides intelligence within the file system. This can be used by applications, as well as core Windows Server services. For example, Hyper-V 2016 uses the BlockClone API to commit snapshots. The API can also be used outside of Windows.

ReFS is ideal for backup storage.

A best practice is the 3-2-1 rule of backup. This rule advocates keeping 3 different copies of data on 2 different media; 1 copy should be stored offsite and 1 should be stored offline. ReFS is a good "first stop" for backups, and then data can be sent to tape or the cloud.

Five reasons ReFS is a smart technology choice for backup storage:

- 1. Writes. ReFS is ideal for fat and slow hard drives because it behaves like a log-structured file system when data integrity streams are enabled. This eliminates the I/O blender effect.
- 2. Reads. As with writes, ReFS works well when reading from fat and slow hard drives. It was built to optimize reads via allocation and defrag policies. Enabling data integrity streams makes ReFS even better for slow storage.
- 3. Backup window. In a write I/O test, 4KB of random rights land on the disk in a very distributed manner when data integrity streams are enabled. This is a more efficient way of writing to a disk.



4. **Restores.** In a read I/O test, with data integrity streams turned on, 100% of random address requests land on the disk. This is perfect for running VMs directly from backup files.





 ReFS cluster format. The default allocation unit size for clusters in ReFS is 4K. This is not recommended, however, for use cases with large files such as backup storage. A better option is the 64K cluster size. Although this imposes a 10% capacity penalty, it also delivers better reliability and stability. A recent <u>TechNet</u> <u>blog post</u> offers cluster size recommendations for ReFS and NTFS.

Creating virtual disks on ReFS is nearly instantaneous.

Thanks to the BlockClone functionality in ReFS, VMs can be deployed almost instantly in Hyper-V. This encourages better Hyper-V administration practices. Many people thin provision VMs and select dynamic allocation in Hyper-V, because this is a fast way to create virtual disks. When many VMs are deployed on a volume and they all grow dynamically, however, the volume eventually fills up. That causes problems for every VM on the volume. The best practice is to configure VMs with fixed disk size and fixed geometry. Using ReFS to create virtual disks can help teams administer their Hyper-Vs more effectively.

ReFS can be deployed for free, if your organization already has Windows Server.

ReFS is available to organizations using Windows Server. Before deploying ReFS, however, organizations should be aware of the following:

- In Windows Server 2016, the BlockClone API is only available in ReFS 3.1.
- No option exists to natively migrate or perform an in-place upgrade from a previous version of ReFS or NTFS to the latest version.
- Windows 10 does not offer ReFS.
- Remember to keep Windows Updates current, since file system capabilities are updated through this mechanism.

Veeam leverages ReFS for its Fast Clone backup functionality.

Veeam protects approximately 13.0 million virtual machines for clients worldwide. Veeam supports an alwayson experience for any infrastructure, including on-premise, SaaS, cloud, and VCSP MSP. It also supports a wide range of services such as files, applications, virtual machines, and physical servers. Veeam Availability Suite provides non-stop business continuity, digital transformation agility, analytics, and visibility.

About Veeam





For its Fast Clone functionality, Veeam uses the ReFS BlockClone API and synthetic full backups. A traditional synthetic full backup is created from existing backup files in the backup repository. Synthetic full backups reduce stress on production storage, because data isn't read from that system. However, they generate a lot of IO on the backup repository as data is read and written to the synthetic full backup.



Synthetic Full Backup (Legacy)

With Veeam Fast Clone, the BlockClone API is used in conjunction with synthetic full backups. This API enables the backup to point to existing blocks. Both IOPS and capacity in the backup repository are reduced because there is no reading or writing. In addition, pointing is a fast process. Veeam's Fast Clone uses the BlockClone API in ReFS. The Veeam implementation shortens backup windows and uses storage more efficiently. Organizations can keep more full backups on disk.

Rick Vanover, Veeam



ADDITIONAL INFORMATION

- ReFS integration FAQs. Veeam has published a <u>YouTube video</u> which answers frequently asked questions about ReFS integration.
- Best practices for scaling back repositories with Veeam and Microsoft ReFS. This <u>white paper</u> can be downloaded from the Veeam website.
- ReFS and NTFS resources. A <u>comparison between NTFS and ReFS</u> is available online, as is a <u>ReFS overview</u>.

