



Server-side Data Migration

Migrating data from an existing volume to a StorSimple volume

September 2, 2015

Copyright and Notices

This document is provided "as-is". Information and views expressed in this document, including URL and other Internet Web site references, may change without notice. Some examples depicted herein are provided for illustration only and are fictitious. No real association or connection is intended or should be inferred.

This document does not provide you with any legal rights to any intellectual property in any Microsoft product. You may copy, use and modify this document for your internal, reference purposes.

© 2015 Microsoft Corporation. All rights reserved.

Microsoft, Windows Azure, StorSimple, Hyper-V, Internet Explorer, Silverlight, SQL Server, Windows, Windows PowerShell, and Windows Server are trademarks of the Microsoft group of companies. All other trademarks are property of their respective owners.

The DiskBoss data management product is the property of Flexense Ltd. (see <http://www.flexense.com/> and <http://www.diskboss.com>).

Contents

- Copyright and Notices.....2
- Contents.....3
- Introduction.....5
- Intended audience5
- Migration methods5
- Migrating data from a NAS.....5
- Migrating data from a SAN.....6
- File copy utilities.....6
- Data migration process with Robocopy6
- Tools for data verification7
- Best practices for using Robocopy with StorSimple.....8
 - Pros and cons of using Robocopy9
- Data migration process with DiskBoss9
 - DiskBoss facts9
 - Software download and compatibility10
 - Installing DiskBoss10
 - Simple file sync.....11
 - Building a complex execution command to enable tiered file sync14
 - Scheduled sync job creation18
 - DiskBoss data verification22
- Data migration process with disk mirroring22
 - Considerations prior to creating the disk mirror.....22
 - Creating a mirrored volume22
 - Removing the mirror25

Pros and cons of using disk mirroring	27
References	28

Introduction

This document outlines different methods available for copying and synchronizing the data from existing file server volumes to volumes created on the StorSimple appliance. You should plan the migration process and identify all cut-over times before using any of the outlined methods. Your planning process should involve choosing an appropriate data migration method after reviewing the pros and cons of each method. A detailed planning guide for using File and Storage Services to migrate data is available [from Microsoft TechNet](https://technet.microsoft.com/en-us/library/jj863566.aspx) (<https://technet.microsoft.com/en-us/library/jj863566.aspx>).

Intended audience

This document is intended for information technology (IT) professionals and knowledge workers who are responsible for operating and deploying file servers in a managed environment.

Migration methods

This document considers the following commonly used data migration methods. Other data migration methods and tools are outside the scope of this document.

- **Disk mirroring** – Synchronizing data between two volumes by using RAID on the host server
- **File copy utilities** – Copying data by using file copy utilities such as Robocopy or third party applications such as [DiskBoss](#).

Subsequent sections outline the detailed implementation steps, as well as the pros and cons of each method.

Note: When creating volumes on the StorSimple appliance and mounting these volumes, follow the best practices defined in the StorSimple documentation.

Note: Use of Robocopy or DiskBoss are optional and not mandatory.

Migrating data from a NAS

A network-attached storage (NAS) device uses a Common Internet File System (CIFS) or a Network File System (NFS) share to serve data directly to a user. Vendors for NAS devices include NetApp, Isilon, and EMC Data Movers. Migrating data from a NAS device involves copying data from the existing device to the volumes mounted on the Windows server, and then routing users to the new data location. This process requires extensive planning and the use of a file copy utility to migrate data, followed by a cutover from the existing NAS device to the Windows server, where the data will be served from StorSimple volumes.

Migrating data from a SAN

Data from a storage area network (SAN) is served by a host on which the SAN volume is mounted. The host server, commonly a Windows server, uses iSCSI or Fibre Channel protocol to mount a volume from a SAN device. You can use the disk mirroring option available on the server that is hosting the SAN volume to mirror the data. You can also use file copy utilities to migrate data from the SAN volumes mounted on the Windows server to a different Windows server that is mounting a volume from the StorSimple appliance.

File copy utilities

File copy and data migration utilities can be used to copy data from existing volumes to the new volumes created on the StorSimple appliance and mounted on a Windows server. Windows servers have a built-in file copy utility called Robocopy installed on the system. The following 2 sections use Robocopy and DiskBoss as examples, to outline the steps required for migrating data from existing volumes. If you use any other utilities, please refer to the documentation for the utilities for best practices. Each section of this paper provides Robocopy command recommendations and best practices.

Data migration process with Robocopy

A typical data migration process involves the following steps:

1. Planning
2. Initial data copy
3. Incremental sync of data
4. Final data copy and cutover
5. Mapping shares/applications to new storage location

PLANNING: You should carefully plan a data migration. Be sure to consider the cutover dates and the total amount of data being copied. Back up the configuration information for accessing existing storage (such as share names, share permissions, and so on). Plan a maintenance window for the final cutover phase so that the existing data is not modified during this phase and all file locks are released.

INITIAL DATA COPY: This operation involves copying the bulk of the data from the existing storage location to the new volume mounted on the StorSimple appliance. During this operation, we recommend copying older data first and then copying the recently created data. This ensures that older data tiers to the lower tiers on the StorSimple appliance while retaining the most recently modified data on the SSD tier of the appliance.

In Robocopy, use the following command:

```
Robocopy<source> <destination> /COPYALL /S /MIR /ZB /MINAGE:30 /R:2 /W:5 /MT:32
```

- **/COPYALL** – copy all the attributes of the files. This includes; data, security/NTFS access control lists (ACLs), attributes, timestamps, owner information, and auditing information.
- **/S** – include subfolders.

- `/MIR` – mirror a directory tree (this is equivalent to `/E` plus `/PURGE`).
- `/MINAGE:30` – exclude files that are newer than 30 days.
- `/R:2` – retry twice on error (default is 1000000).
- `/W:5` – wait for 5 seconds before retry (default is 30).
- `/MT:32` – use 32 threads (default is 8 and the value can be between 1 and 128).
- `/ZB` – use restartable mode; if access is denied use backup mode.

We recommend that you include the retry and wait options because the files may be open during the copy operation, which will result in an error. The copy operation will continue by skipping the open file or the file on which it encounters an error. We also recommend that you log the operation for better performance (log options are not included in the above example). Run Robocopy in administrator mode to ensure it has enough privileges to copy data and all attributes.

INCREMENTAL SYNC OF DATA: This operation copies the changes made to the files after the initial copy operation is performed. This will also synchronize any files that were open and not available for copying during the initial copy operation. You should run an incremental sync on a regular basis to reduce the amount of data that needs to be copied during the final sync operation. During the incremental sync, any files that were deleted on the original data volume can be deleted on the destination so that the data remains synchronized.

In Robocopy, use the following command:

```
Robocopy<source> <destination> /COPYALL /S /MIR /ZB /R:2 /W:5 /MT:32
```

Note: `/MINAGE:30` has been removed so that files which were created recently are included in the sync operation.

FINAL DATA COPY AND CUTOVER: This process involves the final sync of data from existing storage to the new volume created on the StorSimple appliance. You should perform this phase during a scheduled downtime to avoid modification to the data during the final copy process. During this phase, we recommend that you disable user access and close any open handles to the files to ensure that data is copied. After the final copy is complete, you can give users access to the destination by changing the path to existing shares or by changing the drive letter of the new volume.

In Robocopy, use the following command:

```
Robocopy<source> <destination> /COPYALL /S /MIR /ZB /R:2 /W:5 /MT:32
```

Tools for data verification

After you have used Robocopy to perform the file-level data migration, you must verify your data to ensure its integrity. You should compare data from the original share to data on the new share. You should perform this activity before you retire the original share.

You can download and use the [PowerShell File Checksum Integrity Verifier \(PsFCIV\)](#) for data verification. A PsFCIV tracks the integrity status of your files by calculating cryptographic hashes over a file (or files) and writing them into an FCIV-compatible XML database. You can determine whether files were changed since the last check ran.

Important

To avoid pulling the data from the cloud, you should plan to run the verification tool in batches based on the local capacity of your appliance during the data copy operation. Local capacity of the appliance will vary based on the StorSimple appliance model you have purchased. For more information on the local capacity of various models, refer to the [StorSimple appliance datasheets](#).

There are various other third-party tools available that you can use for data verification and comparison. (You may have to purchase or pay the license fee to use a third-party tool.)

Best practices for using Robocopy with StorSimple

When you use Robocopy, pay attention to the following best practices:

- When running Robocopy with StorSimple, use the multi-thread feature (**/MT**) with a maximum of 32 multiple threads. Robocopy supports up to 128 MT; however, the 32 thread recommendation is to ensure that when migrating the data to StorSimple, you leave enough CPU cycles on the source appliance/server for your primary dataset I/Os.
- Use:
 - **/MON:n** – Monitor the source; run again if there are more than *n* changes.
 - **/MOT:m** – Monitor the source; run again in *m* minutes time, if changed.
 - **Log** – use the different log options as needed. More information on logging switches is available [here](#)
 - **/J** : – copy using unbuffered I/O (recommended for large files).
- Depending on the scenario, it is best to have multiple Robocopy jobs running from different hosts and writing on StorSimple. For example, instead of running one Robocopy job with 32 threads, run 4 different Robocopy jobs with 8 threads from separate hosts. This helps improve the copy performance because you will use multiple iSCSI sessions for the copy operation. A single iSCSI session has a limit of 32 queue depth.
- DO NOT USE the **/MOVE** or **/MOV** command to preserve the original copy. After you have verified the contents and your users have had the opportunity to verify their data, take the original copy offline for a couple of weeks, and if there are no complains of missing data, then you can delete the original copy.

Pros and cons of using Robocopy

Pros	Cons
Enables data copy from NAS devices where data mirroring is not possible.	Needs extensive planning for the entire migration process.
Allows consolidating of data from multiple locations.	Data is copied over the network if a different server is used for running the file copy application.
Allows copying of old data first facilitating the prospering tiering of old data to the cloud.	Open files cannot be copied using Robocopy.
Cloud snapshots taken are usable and will contain all the files copied till the cloud snapshot was taken.	Shares may need to be re-mapped.

Data migration process with DiskBoss

Important

DiskBoss is a third party utility that we use as an example in this documentation. It is optional, as you can use any other utility as well. As with any third party utility, we recommend that you read the latest version of the relevant user manual, as its functionality and user interface may have changed since the user manual was initially created. Should you decide to use DiskBoss as your utility, you can review the [DiskBoss' user manual](#).

The intent of this section is to show how to go through the process of migrating data with DiskBoss.

There are a few factors that are key to success when you perform a live file share migration to a share in a StorSimple volume:

- Minimize the down or offline time and operational burden and provide an estimated cut-off window.
- Ensure data, security, and structural integrity.
- Migrate data by different last accessed time stamps to ensure proper tiering.
- Ensure that you have the ability to audit and do checksums or data verification after a file has been copied.
- After the migration, make sure that you can cleanly uninstall the application.

DiskBoss facts

Diskboss has the following features and limitations:

- Option to preserve directories timestamps for file synchronization operations.
- Option to disable I/O buffering when transferring large files.
- Ability to export/import commands in command groups.

- Four different file copy modes for file copy operations.
- Ability to control the file verification mode for file copy operations.
- Performance optimization options for file copy operations by fully taking advantage of the SMB3.x capabilities
- Fault tolerant file copy mode.
- Ability to exclude directories for file copy operations.
- DiskBoss allocates an aligned memory buffer for each file synchronization thread according to the configured number of threads and the buffer size.

For example, when you configure a 4 MB and 32 threads, DiskBoss will allocate 128 MB of system memory for I/O buffers, which is acceptable for servers with enough memory. However, be careful not to use DiskBoss on a production server with all the memory used by other applications.

- The server hosting DiskBoss should have at least 256 MB of free memory for DiskBoss Ultimate and 512 MB for DiskBoss Server. For more on system requirements, please refer to [DiskBoss help](#).
- By default, file sync commands are created with the preview mode enabled, which may cause some issues for customers trying to synchronize millions of files in the preview mode. Consider disabling preview mode for all file sync commands for large directories. With scheduled periodic jobs, you might accidentally close the desktop application or log out from the server and then the DiskBoss desktop application will be closed by the operating system. To use scheduled periodic jobs, you need to keep the DiskBoss desktop application running or use the DiskBoss Server version, which runs in the background as a service. The drawback of running as a service is that many companies do not allow enough permissions for service accounts, or for service accounts to have backup operator rights.

Software download and compatibility

As this is a user-generated and controlled operation, you should use the appropriate version of [DiskBoss](#). We used DiskBoss Ultimate x64 in our tests and examples because it provides more flexibility than the service version and it does not require a service account with necessary rights. Gaining access to such a service account can be difficult because of security and compliance policies.

DiskBoss Download Links

[DiskBoss Ultimate 32-Bit](#)

[DiskBoss Ultimate 64-Bit](#)

[DiskBoss Server 32-Bit](#)

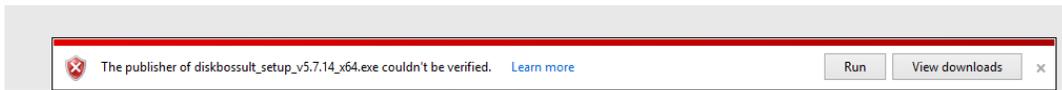
[DiskBoss Server 64-Bit](#)

Installing DiskBoss

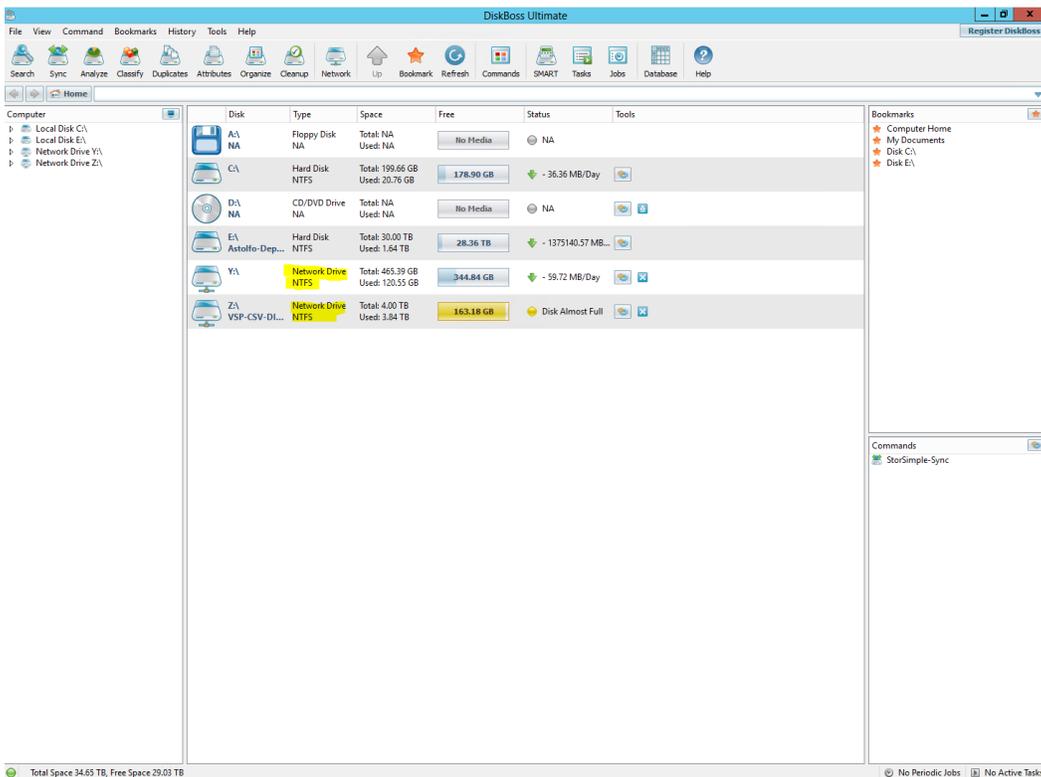
As most operating systems are 64-bit you will get better performance if you use the 64-bit version of DiskBoss. You can verify this by the x64 extension after the DiskBoss version number. In this case, we selected DiskBoss Ultimate v5.7.14 **x64**.

To install DiskBoss

1. Double-click the installer. The following publisher warning will appear.



2. If your organization and/or company's policies allow it, run the program and install the application.
3. Shut down DiskBoss and use the appropriate credentials to mount your network shares.
4. Start DiskBoss Ultimate, and your shares should be visible. In the example, Y and Z are properly identified as network drives.

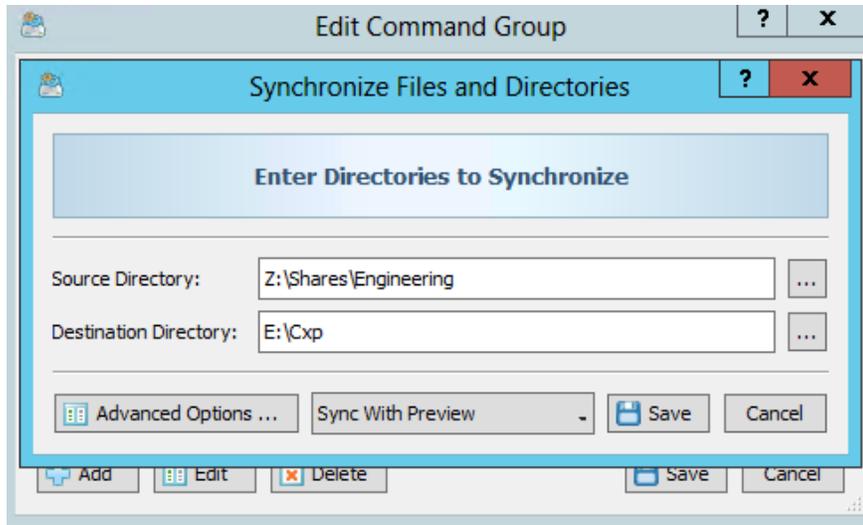


Simple file sync

[Use the following procedure to create a simple sync command.](#) enter the source and destination folders you want to migrate, and add a copy of files that have not been accessed for more than 45 days.

To create the simple sync command

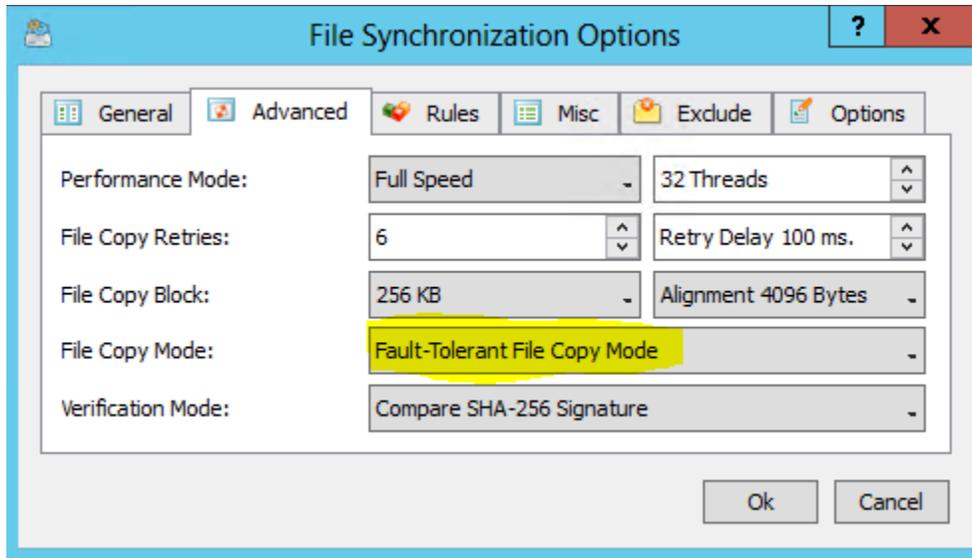
1. Click **Add** to add a new command.
2. Select the source and destination.



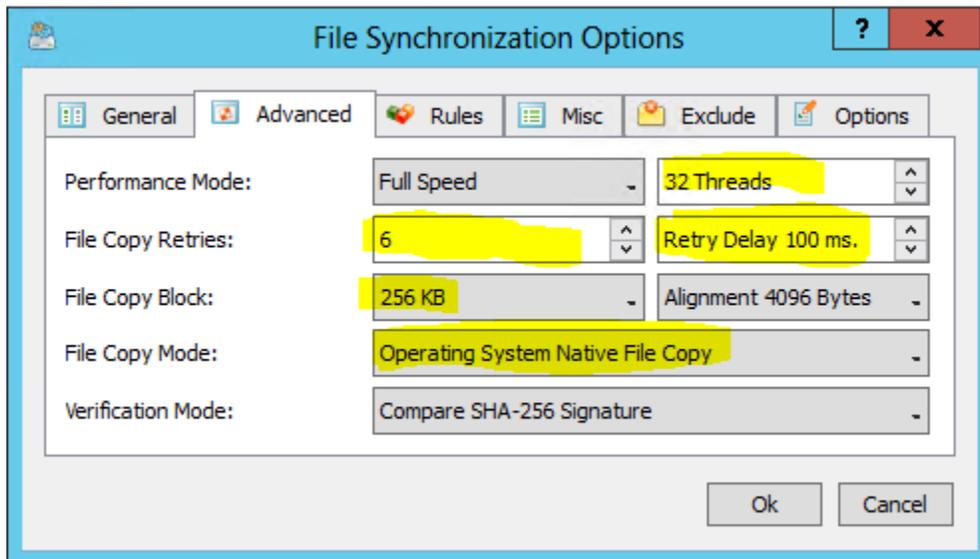
3. Click the **Advanced Options** button.
4. On the **Advanced** tab, under file copy mode, test both the **Fault-Tolerant File Copy Mode** and the **Operating System Native File Copy Mode** with the same options.

DiskBoss has the following copy modes:

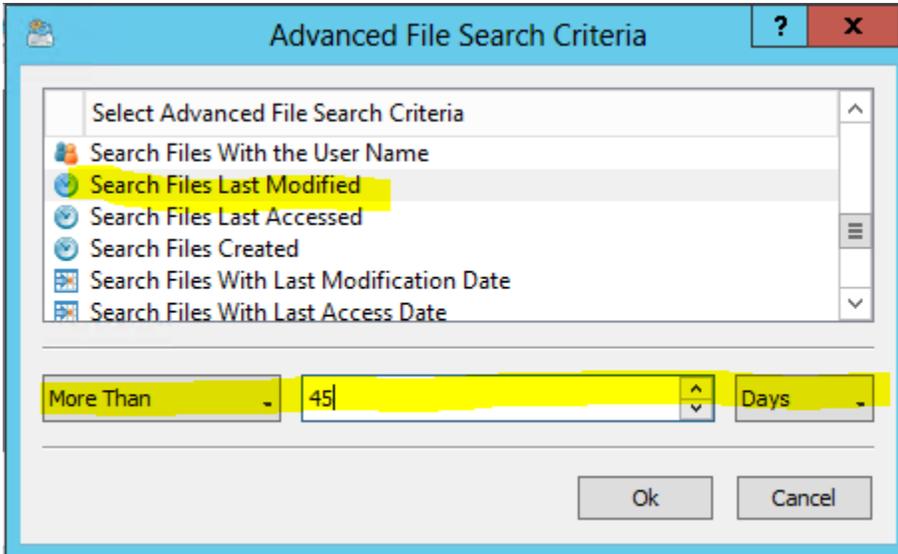
- **Active File Copy Mode** – In this file copy mode, DiskBoss automatically switches between different file copy modes depending on the size and the type of each file. Small files are transferred using a regular, buffered file copy mode, while large files are transferred using a non-buffered file copy mode which, when the hardware platform supports it, enables the write-through I/O mode.
- **Regular, Buffered File Copy Mode** – In this file copy mode, DiskBoss performs regular, buffered file copy operations according to the specified file I/O block size and memory alignment.
- **Fault-Tolerant File Copy Mode** – In this file copy mode, DiskBoss creates a temporary file for each file that should be copied, and only after a successful file copy operation, replaces the original file with the new one. If a network failure occurs during a file copy operation, the original files will remain in place without any changes. See the following illustration.



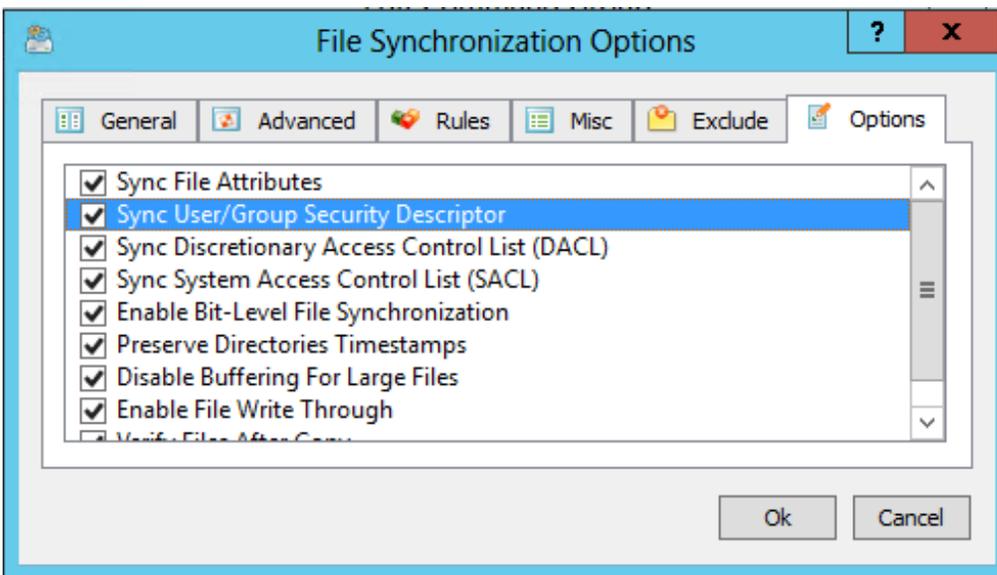
- **Operating System Native File Copy Mode** – This file copy mode is optimized for modern operating systems such as Windows Server 2012 and later. This file copy mode supports SMB direct, direct RDMA file transfers, multi-channel, and write-through file I/O operations. To take full advantage of all performance optimizations when synchronizing files via the network, Windows Server 2012 or newer should be installed on both sides.



5. Click **Ok**.
6. In the **Advanced Files Search Criteria** dialog box, select **More than**, **45**, and **Days** from the drop-down boxes. Click **Ok**.



7. On the **File Synchronization Options** tab, select the file sync options. Enable the **Preserve Directories Timestamps**. (Be sure to enable this option before you run the initial sync operation.) Click **Ok**.



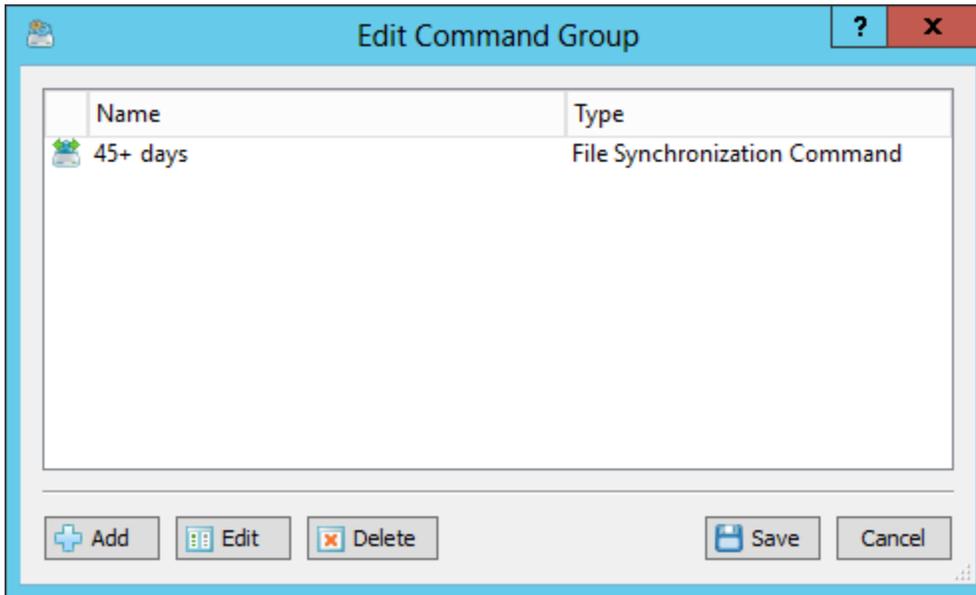
8. Save your command and execute it by right clicking in the commands pane.

Building a complex execution command to enable tiered file sync

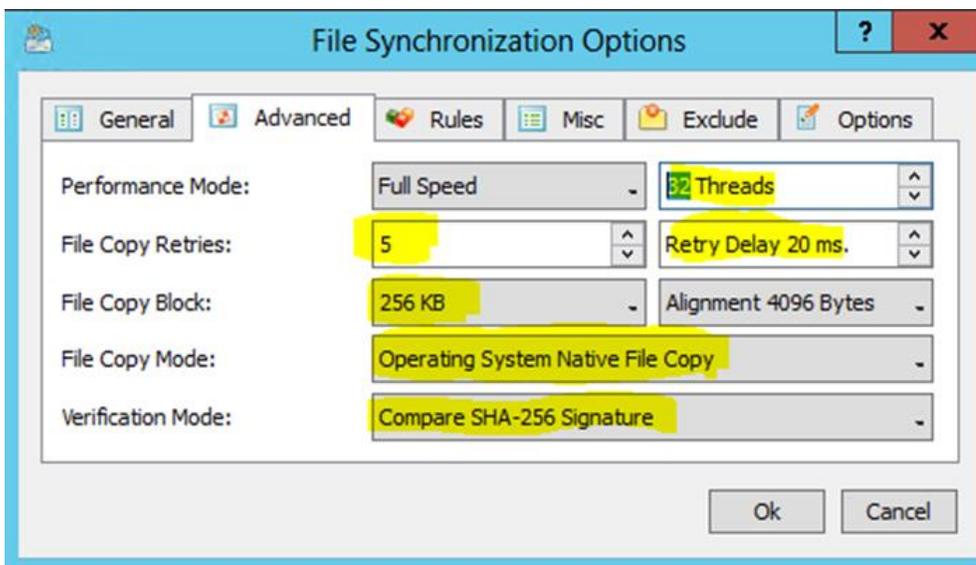
By right-clicking on the **Tiered File Sync** command group you can edit the simple command that you created, and add additional tasks. When you right-click, the **Edit Command Group** dialog box appears.

To build the complex command

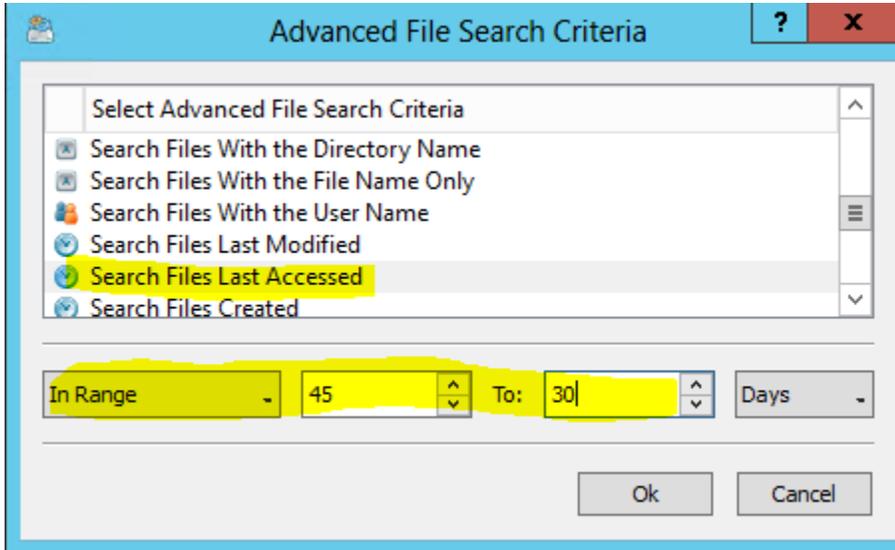
1. Select the command and click **Edit**.



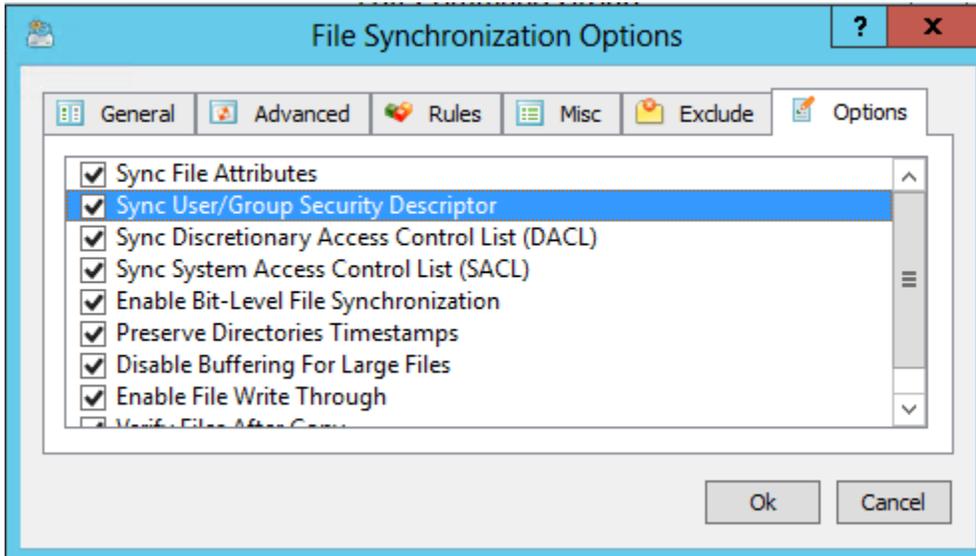
2. Add a 30 day to 45 day range by repeating the simple file sync steps, but now selecting a range.
3. The following illustration shows the file synchronization options before you add the range.



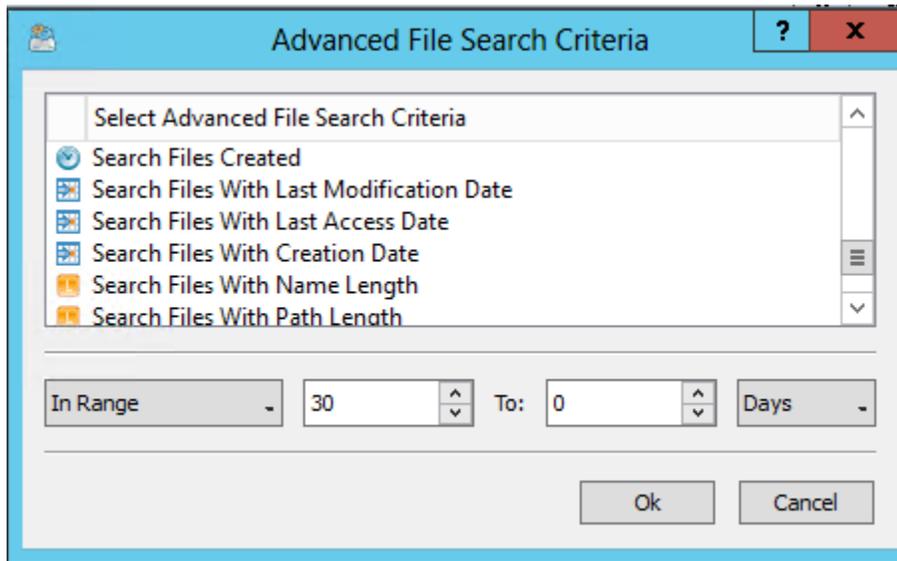
To add the 15 day to 30 day range, select **In Range**, **45**, and **30** from the drop-down boxes, as shown in the following illustration. Click **OK**.



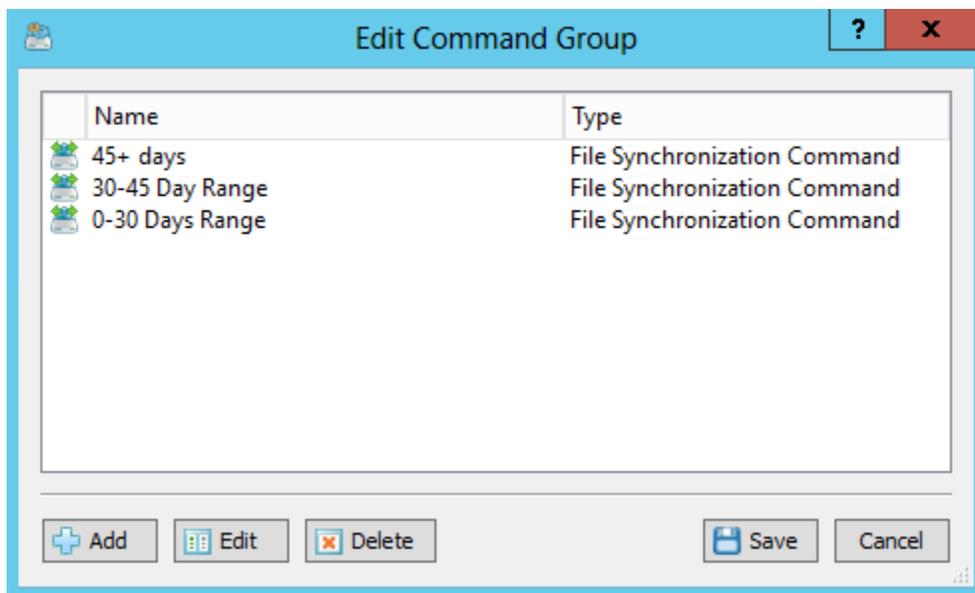
4. On the **Options** tab, select the file sync options. Click **Ok**.



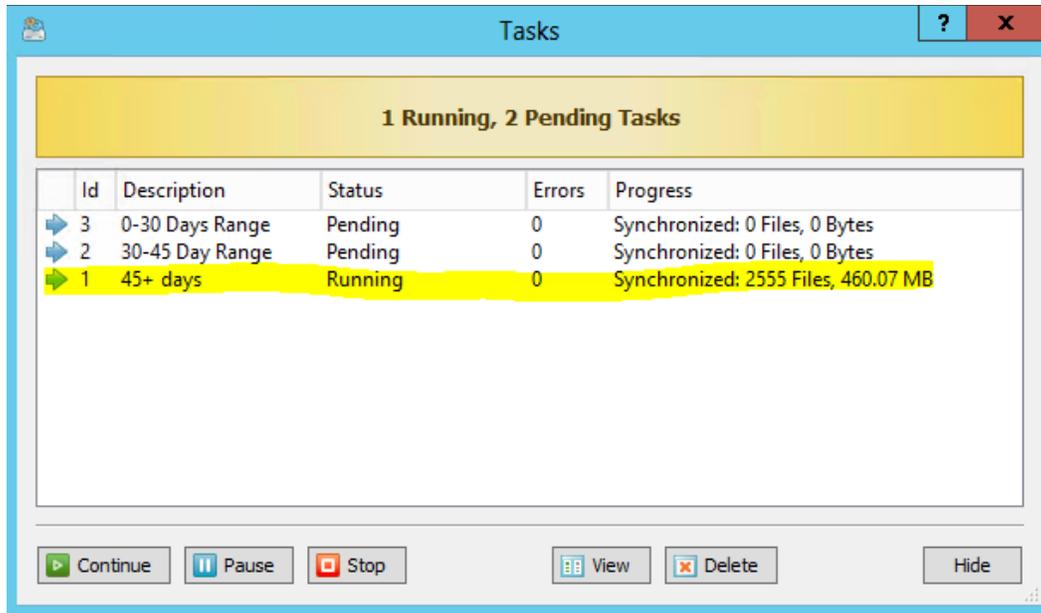
5. Add the 0 to 30 days range by repeating steps 1–2 and selecting a range of 30–0.



6. Click **Ok** and save the command.



7. Execute the tiered command group by right-clicking.

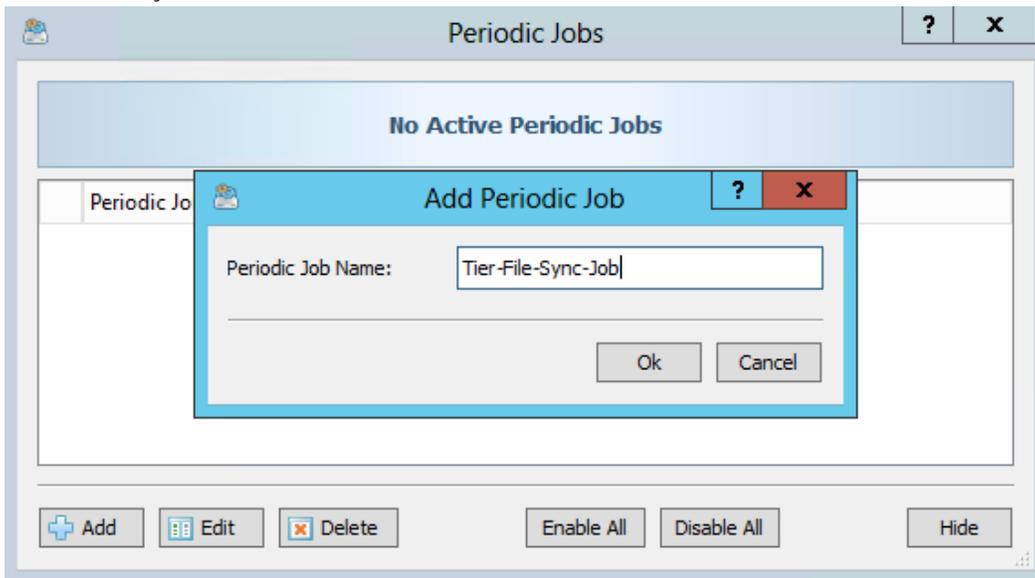


Scheduled sync job creation

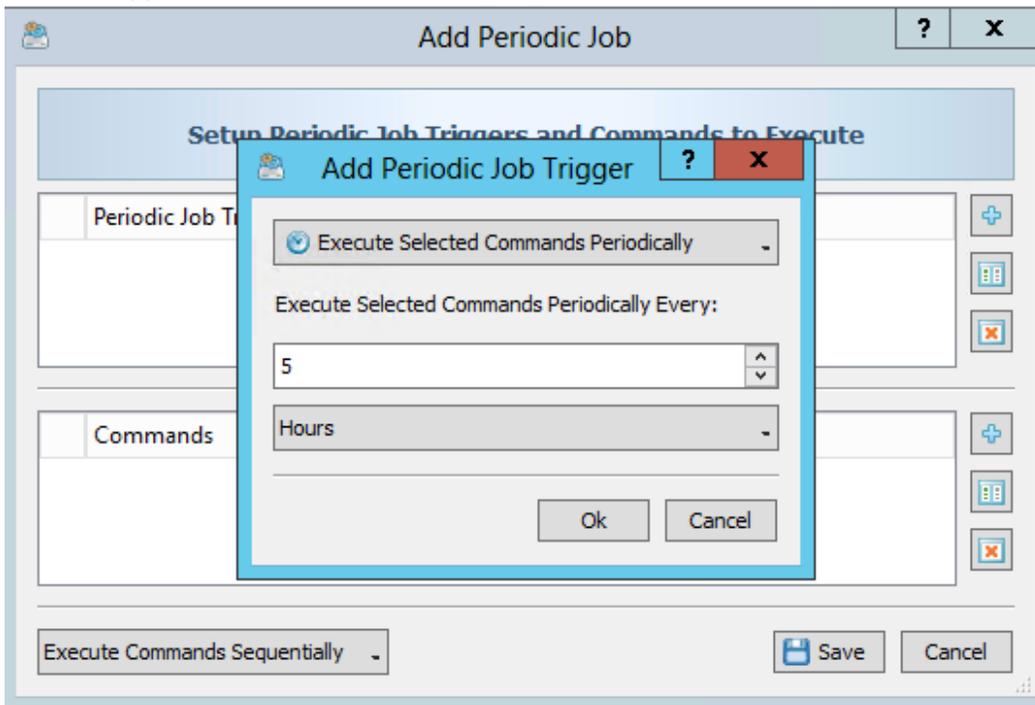
This test validates the ability to run the tiered command job at set intervals until you are ready for the final sync. Depending on your network throughput, latency, number of files and directories and their size; and activity, you should carefully estimate the intervals at which you want to run the command group. A good reference point would be to set the interval to $1/3^{\text{rd}}$ of the initial sync, as you do not want to over run your network and affect user experience.

To complete this task

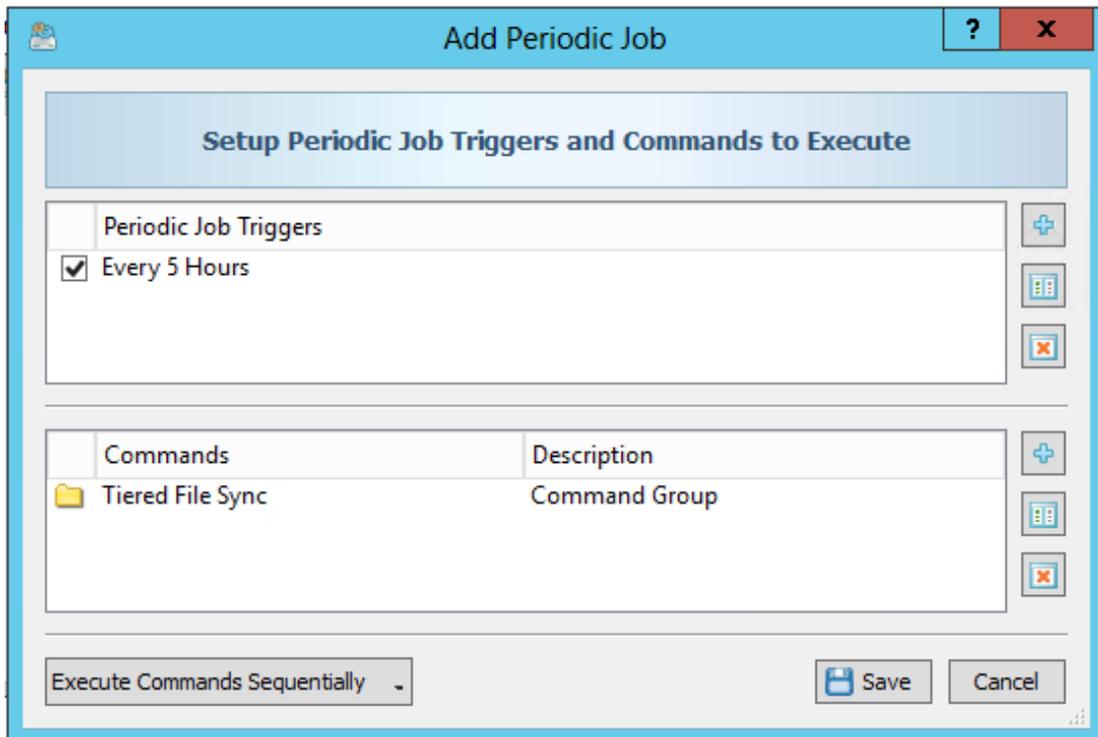
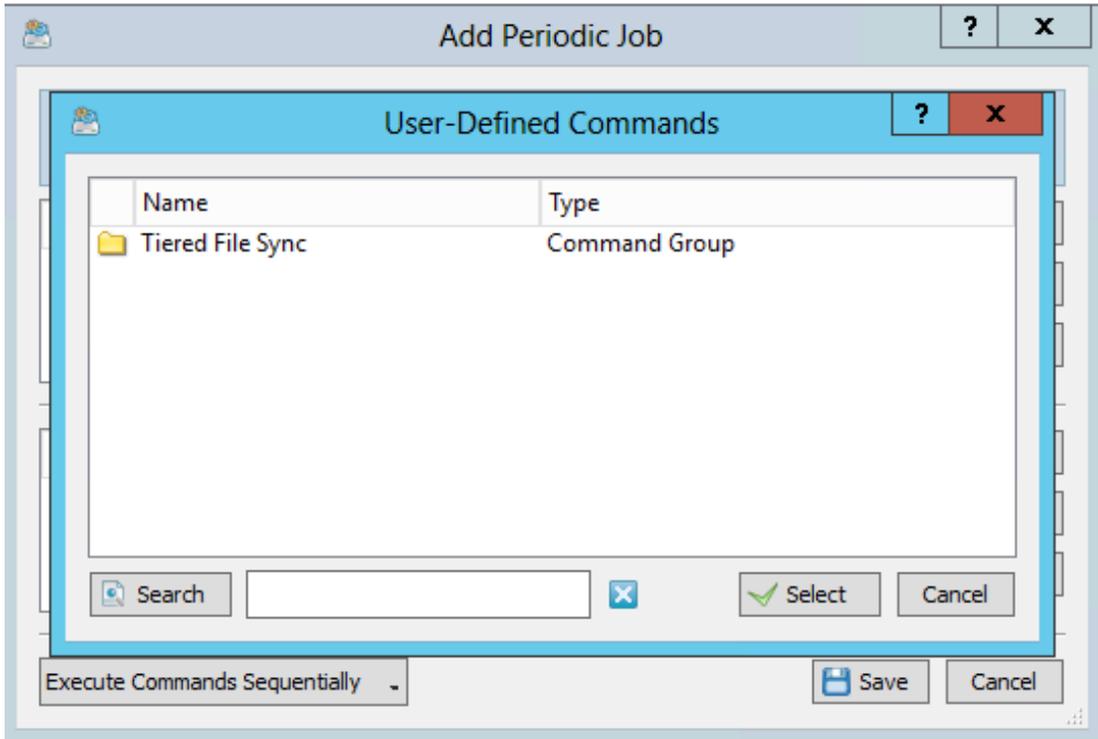
1. Create a new job



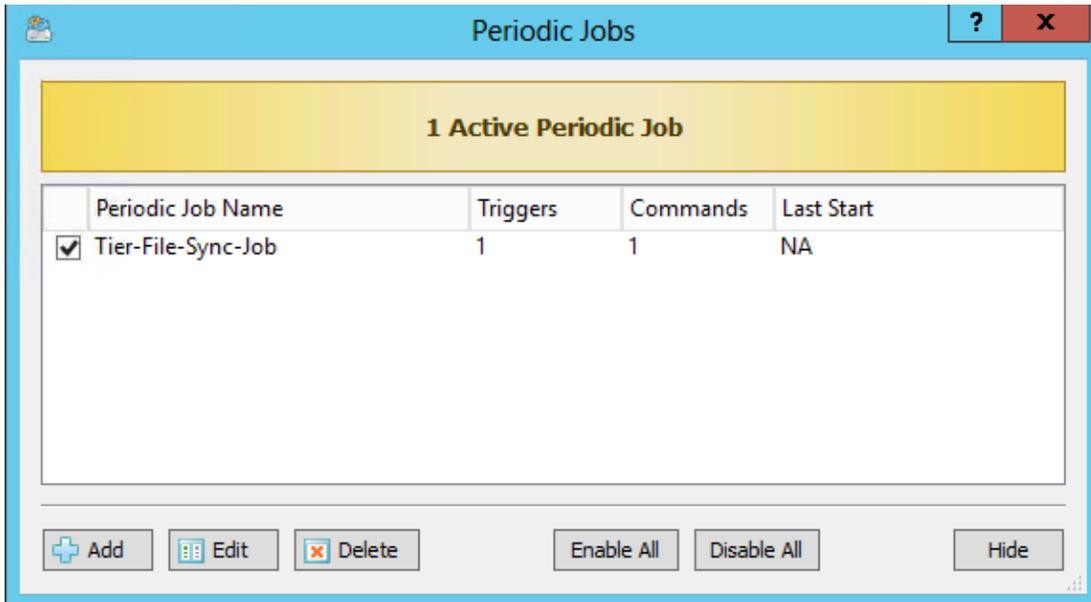
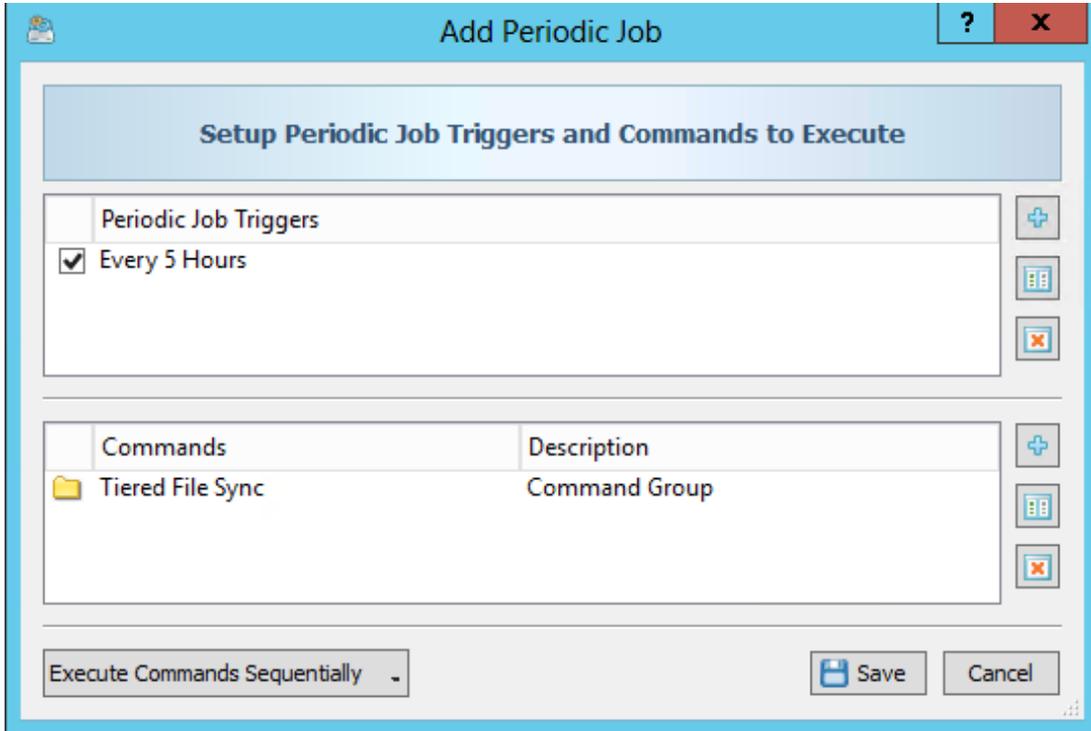
2. add the trigger,



3. then add the job command.



4. Save and enable the job.



5. Save the job and then run it.

DiskBoss data verification

When the option to verify files after copy is selected, DiskBoss calculates a hash signature of the source file while the file is copied to the destination. After the file transfer is completed and the destination file is closed, DiskBoss reads the destination file again, calculates a hash signature of the destination file, and then checks if both signatures are identical.

On the **Advanced** options tab, you can select one of the following hash signatures or CRC checksums:

- **SHA-256 Signature** – the default; the most reliable signature, but requires more CPU resources.
- **SHA-1 Signature** – a good enough signature; requires less CPU resources.
- **MD5 Signature** – a simple hash signature; requires less CPU resources.
- **CRC64 Checksum** – a 64-Bit cyclic redundancy check.
- **CRC32 Checksum** – a 32-Bit cyclic redundancy check.

For multi-threaded file sync operations, verification is performed in parallel using a number of CPUs according to the number of configured file synchronization threads.

Data migration process with disk mirroring

Windows Server provide disk mirroring (RAID1) for volumes by using tools built into the operating system. The disk mirror feature allows the synchronization of existing data in volumes/disks mounted on the Windows server to be mirrored, and any subsequent writes to be performed to both disks in the mirror. This process allows administrators to synchronize all the data and then break the mirror to use the new volume independently.

Considerations prior to creating the disk mirror

A disk mirror can be created only if the disk/volume, to be mirrored, is present on the Windows server. The disk must be converted from a basic disk to a dynamic disk prior to setup of the mirror. Mirroring of disks cannot be set up for any share mounted as a drive letter. The existing data disk/volume should be formatted as an NTFS volume.

Note: You might not be able to use this method if the disk geometries are not the same.

To find out if your disk geometries align please use [diskpart](#), [MSINFO32](#), or [NTFSInfo](#).

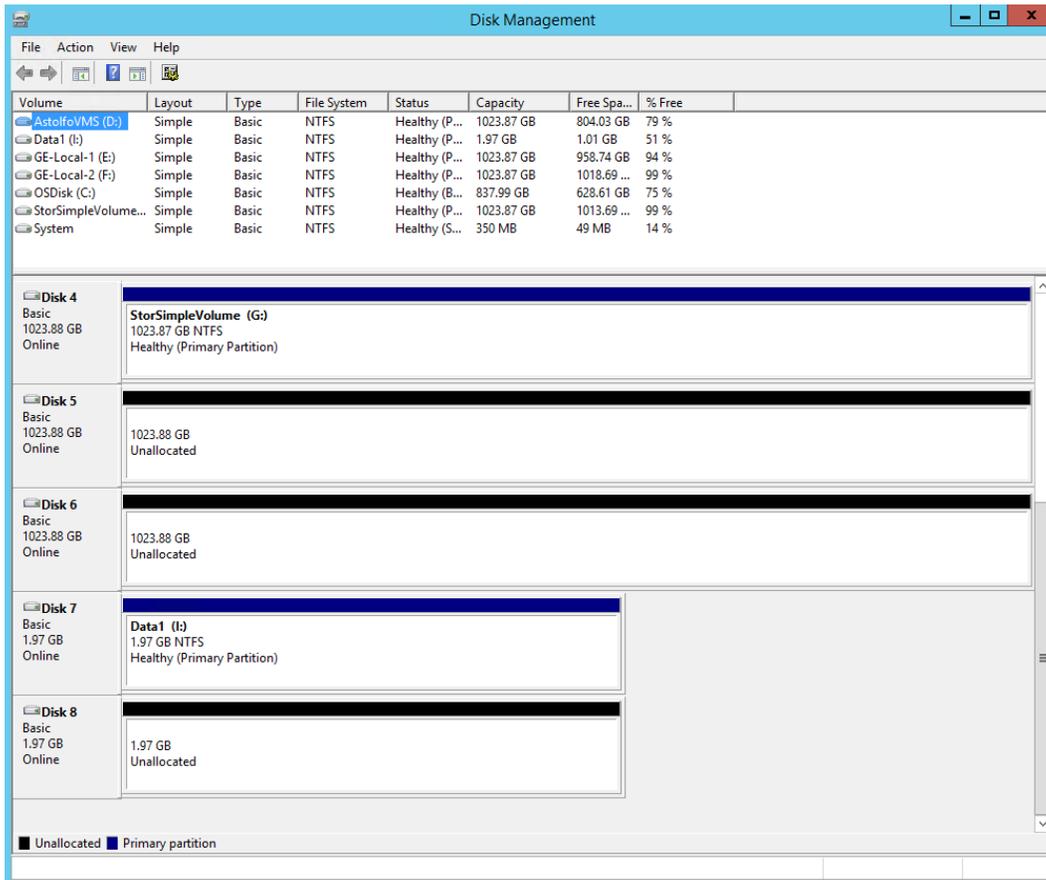
Creating a mirrored volume

The procedure in this topic allows you to create a disk mirror for continuous data availability, even if one of the disks fails. After the mirror is created and the synchronization between the mirrored disks is complete, the data will be present on both disks.

The following illustration shows the existing disk with labeled Data1 as it appears in the Disk Management Microsoft Management Console (MMC) snap-in. This disk will be mirrored with a new disk mounted from the StorSimple appliance.

To create the mirrored volume

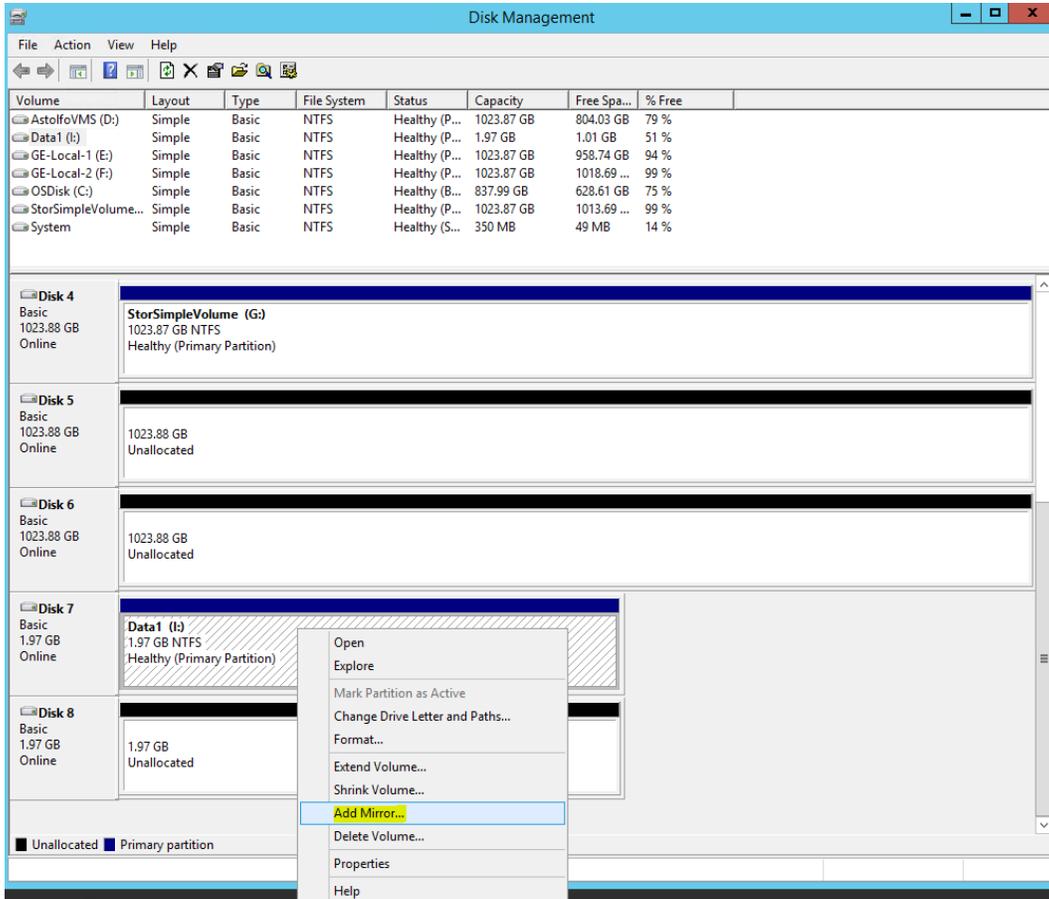
1. Open the Disk Management MMC snap-in: click **Start > Run > diskmgmt.msc**.



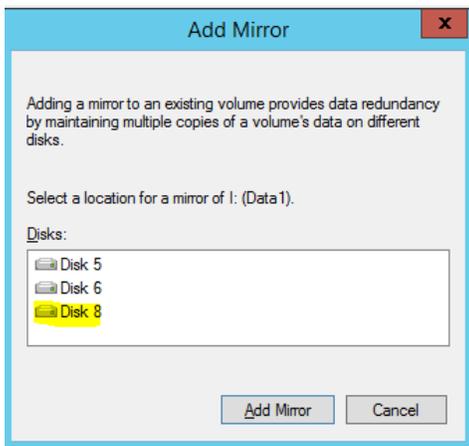
2. On the menu, click **Action > Rescan Disks** to discover the volume created on the StorSimple appliance.

Note: Please refer to StorSimple documentation for instructions when creating a volume and using iSCSI to connect the Windows server to the StorSimple appliance.

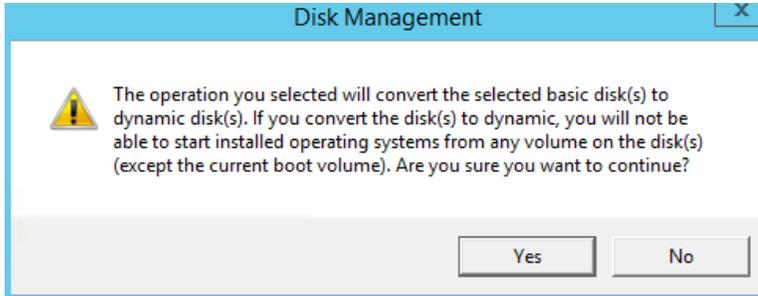
3. Right-click the new disk and bring the disk online.
4. Right-click the existing disk which contains data, and select **Add Mirror**.



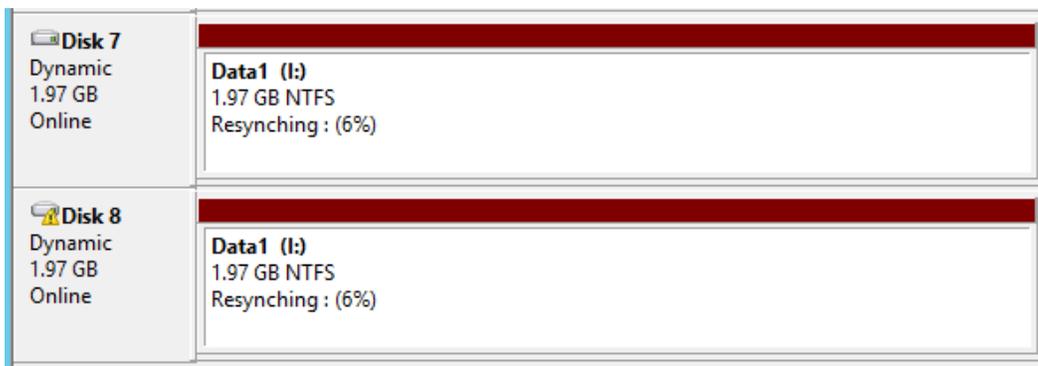
- In the **Add Mirror** dialog box, select the disk that is mounted from the StorSimple appliance.



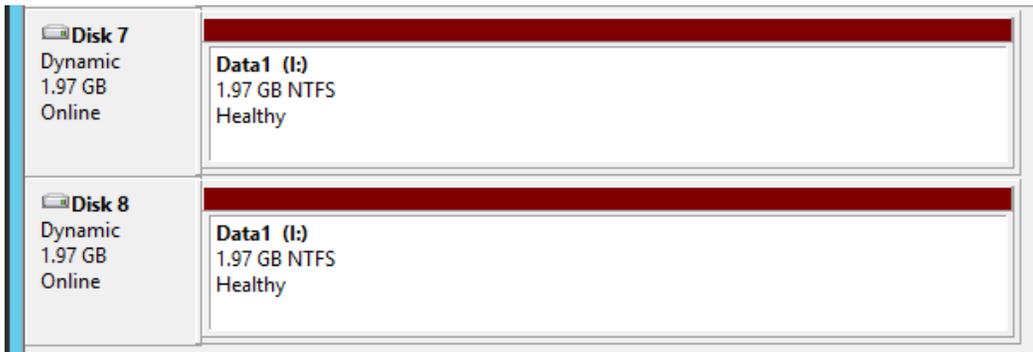
If the original data disk is a basic disk, it will be converted to a dynamic disk. The following confirmation message will be shown prior to conversion.



6. Click **Yes** to continue and create the mirror.
7. When the mirror is created, the resynching process starts. This process can take hours to days, depending on the total amount of data present on the original disk.



8. When the synchronization process is completed, the disks are shown as healthy in the Disk Management console.

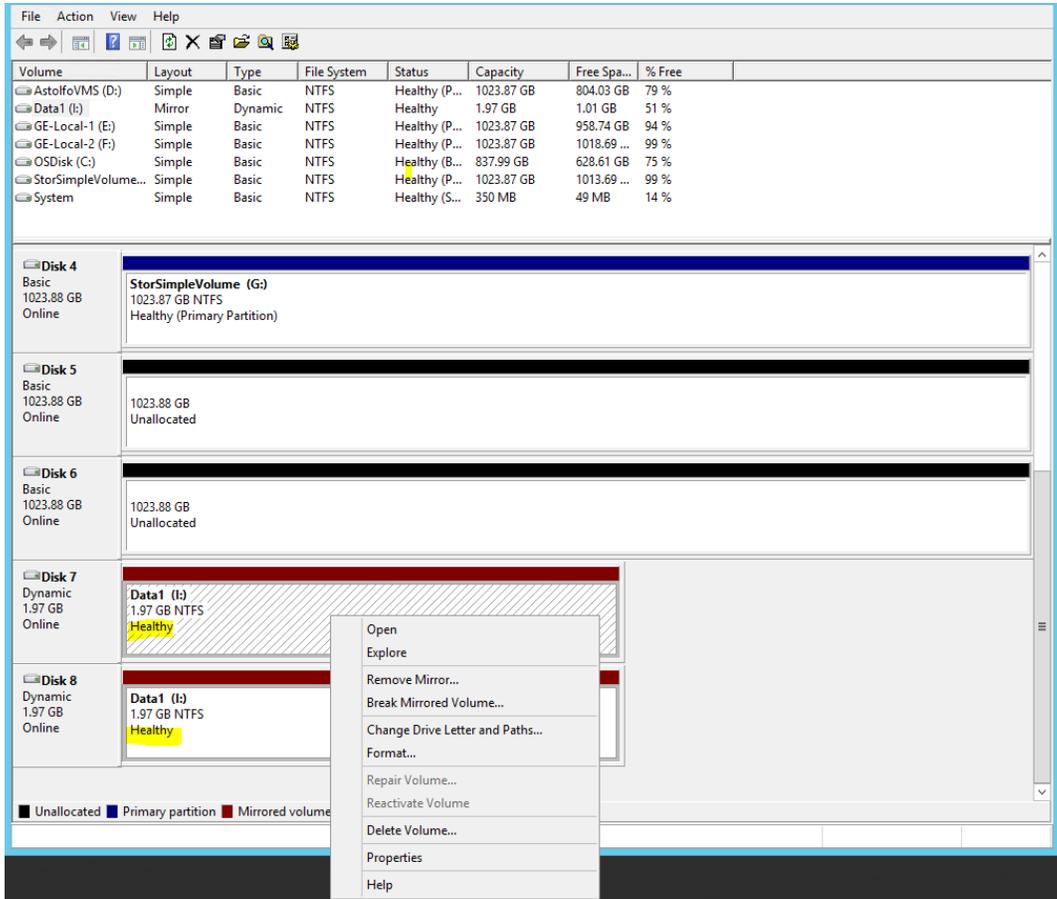


Removing the mirror

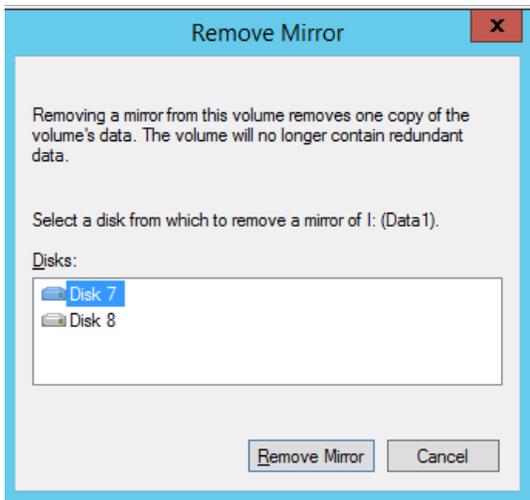
After the synchronization process is complete, the mirror can be removed, retaining only one disk of the mirrored set. The next section explains the procedure for removing the mirror. We recommend that you perform this process during a scheduled maintenance window.

To delete the mirror

1. In the Disk Management console, right-click one of the mirrored disks and select **Remove Mirror**.



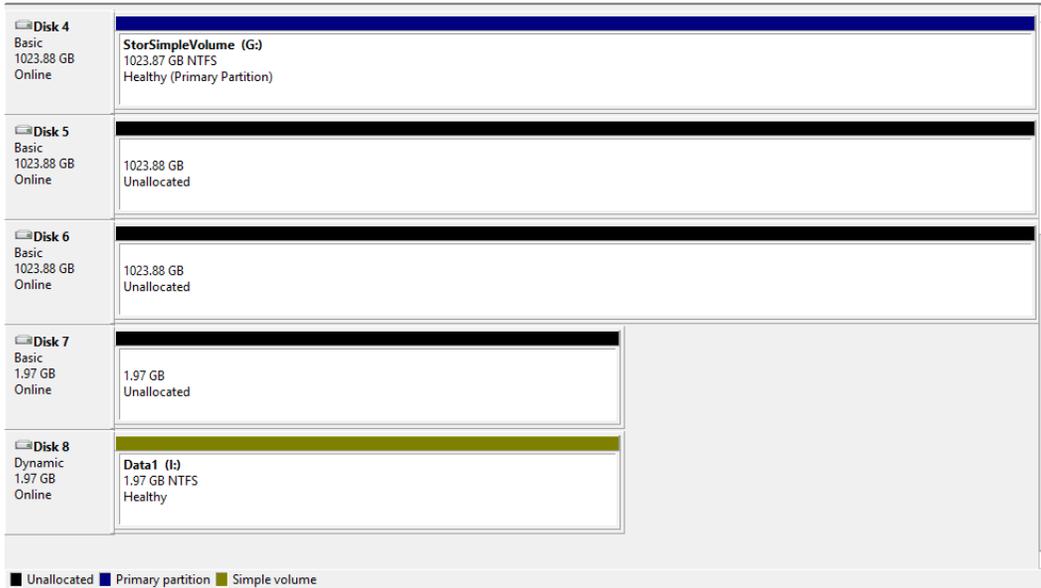
2. In the **Remove Mirror** dialog box, select the disk that you want to remove from the mirror, and retain the disk from the StorSimple appliance as the primary disk.



Note: Make sure that you select the disk that you want to remove from the mirror.

3. A warning message will appear. Click **Yes** to proceed.

4. After the disk mirror is removed, the new disk will be online and serving data.



Pros and cons of using disk mirroring

Pros	Cons
Disks are online and operational while mirroring process is in progress.	Mirrors cannot be created with CIFS shares mounted using a drive letter.
No changes to shares required during or after mirroring process thus ensuring users are not affected during the migration process.	Mirrors cannot be created on a NAS device. The mirroring process mirrors each and every block stored on the existing disk, including empty blocks.
Simple process to implement.	If the existing disk is a basic disk, it will be converted into a dynamic disk.
	Cloud snapshots taken during the sync process are not usable as they will contain incomplete data.
	NTFS AUS will be matched with original disk and may not be optimal for StorSimple.

References

[Microsoft Azure StorSimple for File Shares](#)

[PowerShell File Checksum Integrity Verifier \(PsFCIV\)](#)

[DiskBoss documentation](#)

[Windows Sysinternals](#)