

9.1 Easy Tier overview

Easy Tier is an optional licensed feature of IBM Storwize V5000 that brings enterprise storage functions to the midrange segment. It enables automated subvolume data placement throughout different storage tiers to intelligently align the system with current workload requirements and to optimize the usage of SSDs. This functionality includes the ability to automatically and non-disruptively relocate data (at the extent level) from one tier to another tier in either direction to achieve the best available storage performance for your workload in your environment.

Easy Tier reduces the I/O latency for hot spots, but it does not replace storage cache. Easy Tier and storage cache solve a similar access latency workload problem, but these methods weigh differently in the algorithmic construction that is based on “locality of reference”, recency, and frequency. Because Easy Tier monitors I/O performance from the device end (after cache), it can pick up the performance issues that cache cannot solve and complement the overall storage system performance.

In general, the storage environments I/O is monitored on volumes and the entire volume is always placed inside one appropriate storage tier. Determining the amount of I/O is too complex for monitoring I/O statistics on single extents, moving them manually to an appropriate storage tier, and reacting to workload changes.

Easy Tier is a performance optimization function that overcomes this issue because it automatically migrates (or moves) extents that belong to a volume between different storage tiers, as shown in Figure 9-1. Because this migration works at the extent level, it is often referred to as *sub-LUN migration*.

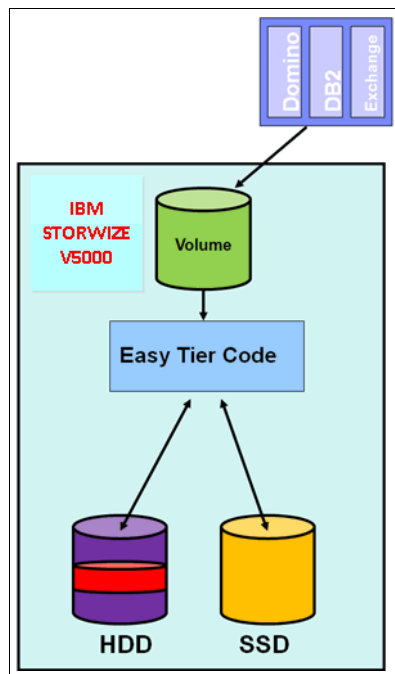


Figure 9-1 Easy Tier

You can enable Easy Tier for storage on a volume basis. It monitors the I/O activity and latency of the extents on all Easy Tier enabled volumes over a 24-hour period. Based on the performance log, it creates an extent migration plan and dynamically moves high activity or hot extents to a higher disk tier within the same storage pool. It also moves extents whose activity dropped off (or cooled) from higher disk tier MDisks back to a lower tier MDisk.

To enable this migration between MDisks with different tier levels, the target storage pool must consist of different characteristic MDisks. These pools are named *multitiered storage pools*. IBM Storwize V5000 Easy Tier is optimized to boost the performance of storage pools that contain HDDs and SSDs.

To identify the potential benefits of Easy Tier in your environment before actually installing higher MDisk tiers (such as, SSDs), it is possible to enable the Easy Tier monitoring volumes in single-tiered storage pools. Although the Easy Tier extent migration is not possible within a single-tiered pool, the Easy Tier statistical measurement function is possible. Enabling Easy Tier on a single-tiered storage pool starts the monitoring process and logs the activity of the volume extents. In this case, Easy Tier creates a migration plan file that can then be used to show a report on the number of extents that is appropriate for migration to higher level MDisk tiers, such as, SSDs.

The STAT is a no-cost tool that helps you to analyze this data. If you do not have an IBM Storwize V5000, use Disk Magic to get a better idea about the required number of SSDs that is appropriate for your workload. If you do not have any workload performance data, a good starting point can be to add approximately 5% of net capacity of SSDs to your configuration. However, this ratio is heuristics-based and changes according to different applications or different disk tier performance in each configuration. For database transactions, a ratio of fast SAS or Fibre Channel (FC) drives to SSD is about 6:1 to achieve the optimal performance, but this ratio depends on the environment on which it is implemented.

Easy Tier is available for IBM Storwize V5000 internal volumes and volumes on external virtualized storage subsystems because the SSDs can be internal or external drives. However, from the fabric point of view, it is a best practice to use SSDs inside the IBM Storwize V5000 (even if the lower tiered disk pool is on external storage) because this configuration reduces the traffic that is traversing the SAN environment.

9.2 Easy Tier for IBM Storwize V5000

This section describes the terms and gives an example implementation of Easy Tier on the IBM Storwize V5000. After reading this section, you should understand the Easy Tier concept as it relates to the IBM Storwize V5000.

9.2.1 Disk tiers

It is likely that IBM Storwize V5000 internal disks and external disks have different performance attributes. As described in Chapter 7, “Storage pools” on page 295, without Easy Tier, it is a best practice to place drives with the same attributes (the number of revolutions per minute, size, and type) in the same storage pool, and not to intermix different drives with different attributes. This configuration is also valid for external MDisks that are grouped into storage pools. All internal HDDs and external MDisks are initially put into the `generic_hdd` tier by default. An internal SSD is identified as a high-performance tier MDisk by IBM Storwize V5000 and all external SSD MDisks must be changed to the high-performance tier, as described in Chapter 7, “Storage pools” on page 295.

9.2.2 Tiered storage pools

With IBM Storwize V5000, we must differentiate between the following types of storage pools:

- ▶ Single-tiered storage pools
- ▶ Multitiered storage pools

As shown in Figure 9-2, single-tiered storage pools include one type for disk tier attribute. Each disk should have the same size and performance characteristics. Multitiered storage pools are populated with two different disk tier attributes, which means high-performance tier SSDs and generic HDDs. A volume migration is when the complete volume is migrated from one storage pool to another storage pool. An Easy Tier data migration moves only extents inside the storage pool to different performance attributes.

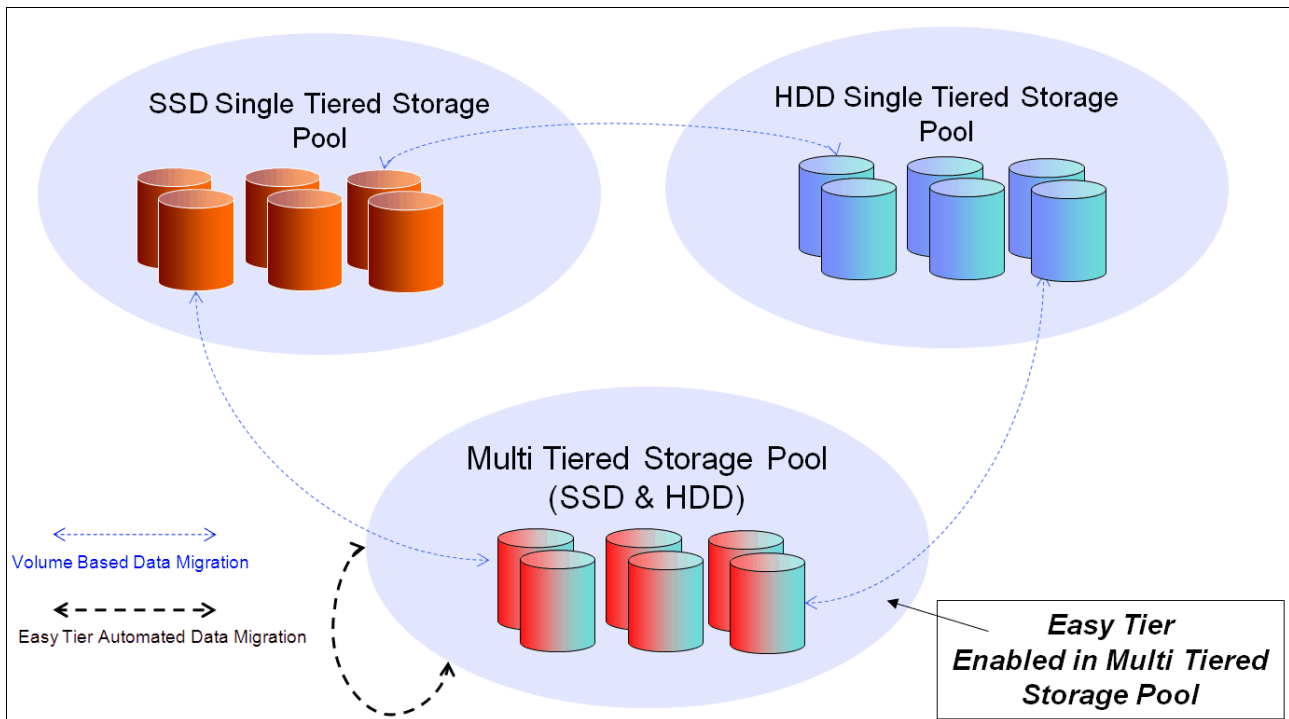


Figure 9-2 Tiered storage pools

9.3 Easy Tier process

The Easy Tier feature consists of four main processes. Figure 9-3 shows the flow between these processes. These processes ensure that the extent allocation in multitiered storage pools is optimized for the best performance that was monitored on your workload in the last 24 hours. Statistics about extent usage are collected at five-minute intervals. A heat map is created every 24 hours that is used by the internal algorithms to generate a migration plan and a summary report. This migration plan contains information about which extents to promote to the upper tier or to demote to the lower tier, and the summary report is used by STAT. For more information, see 9.6, “IBM Storage Tier Advisor Tool” on page 434.

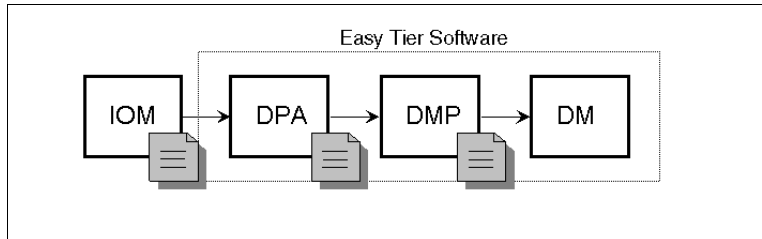


Figure 9-3 Easy Tier process flow

Easy Tier is based on an algorithm with a threshold to evaluate if an extent is cold or hot. If an extent activity is below this threshold, it is not considered by the algorithm to be moved to the SSD tier.

The four main processes and the flow between them are described in the following sections.

9.3.1 I/O Monitoring

The I/O Monitoring (IOM) process operates continuously and monitors host volumes for I/O activity. It collects performance statistics for each extent and derives averages for a rolling 24-hour period of I/O activity.

Easy Tier makes allowances for large block I/Os and thus considers only I/Os of up to 64 KB as migration candidates.

This process is an efficient process and adds negligible processing impact to the IBM Storwize V5000 node canisters.

9.3.2 Data Placement Advisor

The Data Placement Advisor (DPA) uses workload statistics to make a cost benefit decision about which extents should be candidates for migration to a higher performance (SSD) tier.

This process also identifies extents that must be migrated back to a lower (HDD) tier.

9.3.3 Data Migration Planner

By using the previously identified extents, the Data Migration Planner (DMP) process builds the extent migration plan for the storage pool.

9.3.4 Data Migrator

The Data Migrator (DM) process involves scheduling and the actual movement or migration of the volume's extents up to, or down from, the high disk tier. The extent migration rate is capped to a maximum of up to 15 MBps. This rate equates to around 2 TB a day that is migrated between disk tiers, as shown in Figure 9-4.

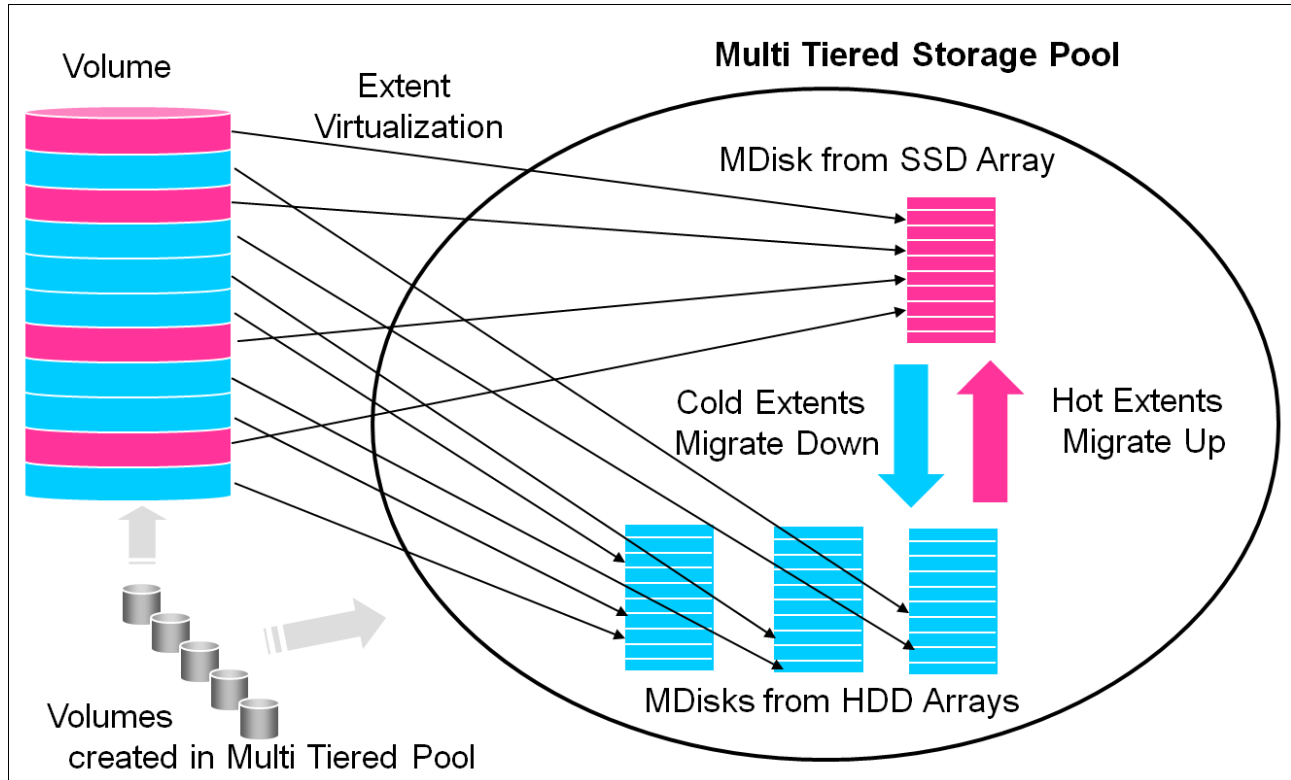


Figure 9-4 Easy Tier Data Migrator

9.3.5 Easy Tier operating modes

IBM Storwize V5000 offers the following operating modes for Easy Tier:

- ▶ Easy Tier: OFF

Easy Tier can be turned off. No statistics are recorded and no extents are moved.

- ▶ Evaluation Mode

If you turn on Easy Tier in a single-tiered storage pool, it runs in Evaluation Mode, which means it measures the I/O activity for all extents. A statistic summary file is created and can be offloaded from the IBM Storwize V5000. This file can be analyzed with the IBM Storage Tier Advisory Tool, as described in 9.6, "IBM Storage Tier Advisor Tool" on page 434. This analysis shows the benefits for your workload if you were to add SSDs to your pool before any hardware is acquired.

- ▶ Auto Data Placement Mode

This operating mode is enabled by default if you create a multitiered storage pool. Easy Tier is also set to On to all volumes inside the multitiered storage pool. The extents are migrated dynamically by the Easy Tier processes to achieve the best performance. The movement is not apparent to the host server and applications, and it provides increased performance only.

If you do want to disable Auto Data Placement Mode for single volumes that are inside a multitiered storage pool, it is possible to turn off the mode at the volume level. This action excludes the volume from Auto Data Placement Mode and measures the I/O statistics only.

The statistic summary file can be offloaded for input to the advisor tool. The tool produces a report about the extents that are moved to SSD and a prediction of performance improvement that can be gained if more SSD was available.

9.3.6 Easy Tier rules

The following operating rules apply when IBM System Storage Easy Tier is used on the IBM Storwize V5000:

- ▶ Automatic data placement and extent I/O activity monitors are supported on each copy of a mirrored volume. Easy Tier works with each copy independently of the other copy.

Volume mirroring: Volume mirroring can have different workload characteristics on each copy of the data because reads are normally directed to the primary copy and writes occur to both. Thus, the number of extents that Easy Tier migrates to SSD tier probably is different for each copy.

- ▶ Easy Tier works with all striped volumes, which include the following volumes:
 - Generic volumes
 - Thin-provisioned volumes
 - Mirrored volumes
 - Thin-mirrored volumes
 - Global and Metro Mirror sources and targets
- ▶ Easy Tier automatic data placement is not supported for image mode or sequential volumes. I/O monitoring for such volumes is supported, but you cannot migrate extents on such volumes unless you convert image or sequential volume copies to striped volumes.
- ▶ If possible, IBM Storwize V5000 creates volumes or volume expansions by using extents from MDisks from the HDD tier. Extents from MDisks from the SSD tier are used if no HDD space is available.
- ▶ When a volume is migrated out of a storage pool that is managed with Easy Tier, Automatic Data Placement Mode is no longer active on that volume. Automatic Data Placement is also turned off while a volume is migrated, even if it is between pools that both have Easy Tier Automatic Data Placement enabled. Automatic Data Placement for the volume is re-enabled when the migration is complete.
- ▶ SSD performance is dependent on block sizes (small blocks perform much better than larger blocks). Because Easy Tier is optimized to work with SSD, it decides whether an extent is hot by measuring I/O smaller than 64 KB, but it migrates the entire extent to the appropriate disk tier.
- ▶ As extents are migrated, the use of smaller extents makes Easy Tier more efficient.
- ▶ The first migration of hot data to SSD starts about one hour after Automatic Data Placement Mode is enabled. It takes up to 24 hours to achieve optimal performance.
- ▶ In the current IBM Storwize V5000 Easy Tier implementation, it takes about two days before hot spots are considered moved from SSDs, which prevents hot spots from being moved from SSDs if the workload changes over a weekend.
- ▶ If you run an unusual workload over a longer period, Automatic Data Placement can be turned off and on online to avoid data movement.

Depending on which storage pool and which Easy Tier configuration is set, a volume copy can have the Easy Tier states that are shown in Table 9-1.

Table 9-1 Easy Tier states

Storage pool	Single-tiered or multitiered storage pool	Volume copy Easy Tier setting	Easy Tier status
Off	Single-tiered	Off	Inactive ^a
Off	Single-tiered	On	Inactive ^a
Off	Multitiered	Off	Inactive ^a
Off	Multitiered	On	Inactive ^a
Auto ^b	Single-tiered	Off	Inactive ^a
Auto ^b	Single-tiered	On	Inactive ^a
Auto ^b	Multitiered	Off	Measured ^c
Auto ^b	Multitiered	On	Active ^{d e}
On	Single-tiered	Off	Measured ^c
On	Single-tiered	On	Measured ^c
On	Multitiered	Off	Measured ^c
On	Multitiered	On	Active ^d

- a. When the volume copy status is inactive, no Easy Tier functions are enabled for that volume copy.
- b. The default Easy Tier setting for a storage pool is Auto, and the default Easy Tier setting for a volume copy is On. This scenario means that Easy Tier functions are disabled for storage pools with a single tier, and that automatic data placement mode is enabled for all striped volume copies in a storage pool with two tiers.
- c. When the volume copy status is measured, the Easy Tier function collects usage statistics for the volume, but automatic data placement is not active.
- d. If the volume copy is in image or sequential mode or is being migrated, the volume copy Easy Tier status is measured instead of active.
- e. When the volume copy status is active, the Easy Tier function operates in automatic data placement mode for that volume.

9.4 Easy Tier configuration by using the GUI

This section describes how to activate Easy Tier by using the IBM Storwize V5000 GUI.

9.4.1 Creating multitiered pools: Enable Easy Tier

In this section, we describe how to create multitiered storage pools by using the GUI. When a storage pool changes from single-tiered to multitiered, Easy Tier is enabled by default for the pool and on all volume copies inside this pool.

To create multitiered pools, complete the following steps:

1. Click **Volumes** → **Volumes by Pool**. Figure 9-5 shows that the V5000_Pool_1 storage pool exists and Easy Tier is inactive in our example. An SSD must be added to the V5000_Pool_2 storage pool to enable Easy Tier.



Figure 9-5 Single-tiered pool

2. Click **Pools** → **Internal Storage**. Figure 9-6 shows that one internal SSD is available and it is in the Unused status. Internal SSDs are assigned the generic_ssd high performance tier automatically by the IBM Storwize V5000.

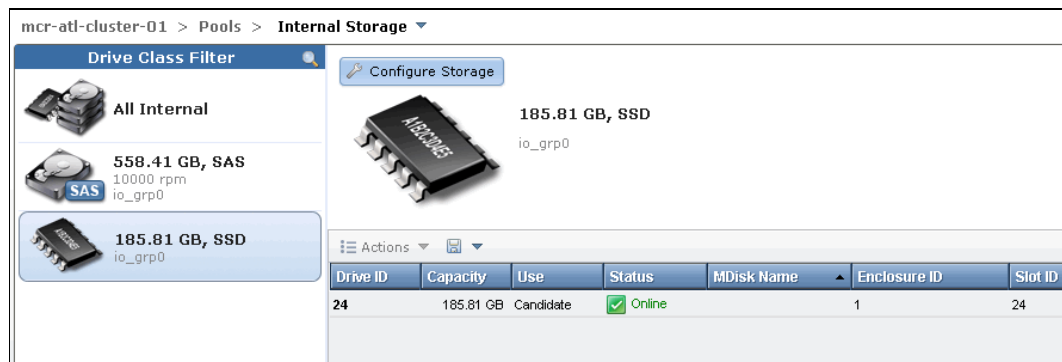


Figure 9-6 Internal SSDs

3. Click **Configure Storage** and the Storage Configuration wizard opens. Because the SSD is in Unused status, an information message is displayed, as shown in Figure 9-7. Click **Yes** to proceed to the next window.

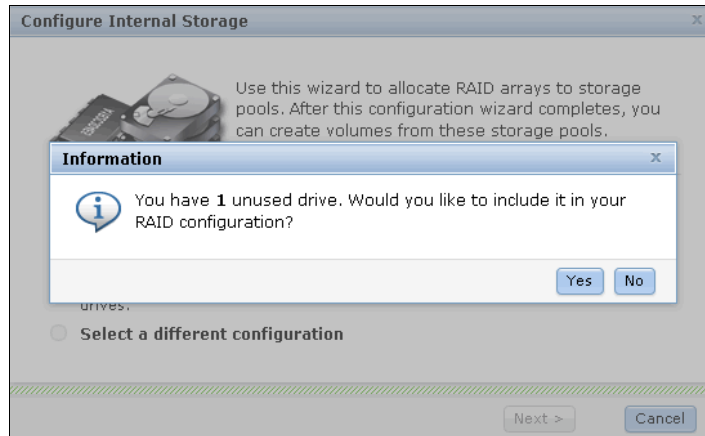


Figure 9-7 Unused drives information message

4. The IBM Storwize V5000 configures the newly detected drive and changes its status to Candidate after the task is completed, as shown in Figure 9-8. Click **Close** to start the SSD configuration.

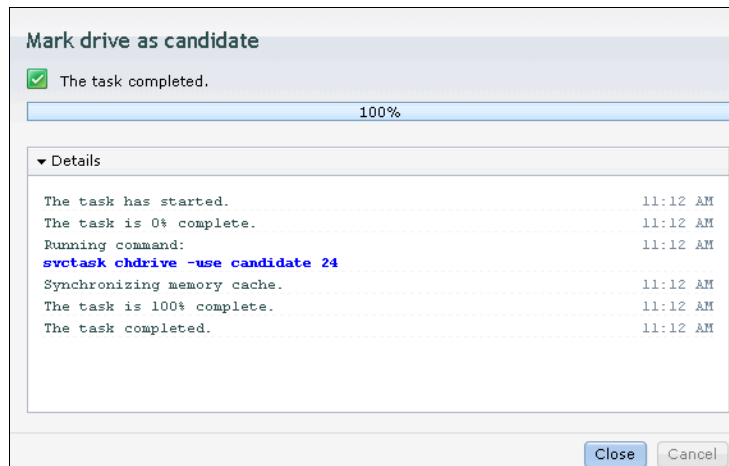


Figure 9-8 Drive configuration output

Figure 9-9 shows the first step of the Configuration wizard.

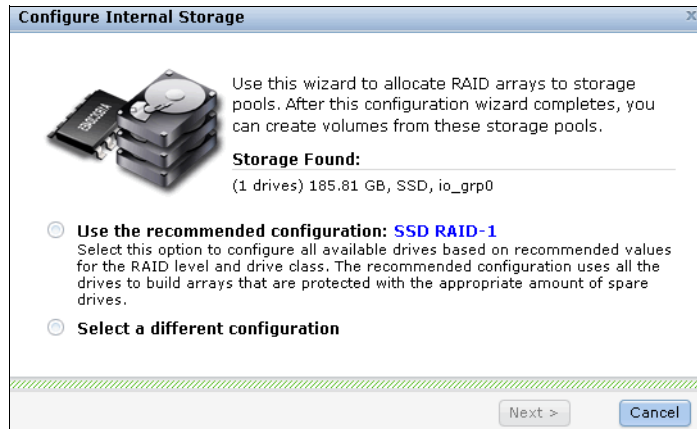


Figure 9-9 Configure Internal Storage window

The wizard recommends the use of the SSD to enable Easy Tier. If you select **Use recommended configuration**, it selects the recommended RAID level and hot spare coverage for your system automatically, as shown in Figure 9-10.

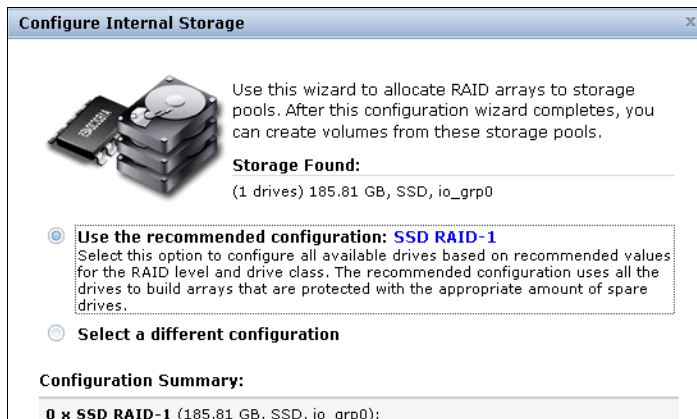


Figure 9-10 Recommended configuration

If you select **Select a different configuration** (as shown in Figure 9-11), you can select the preset.

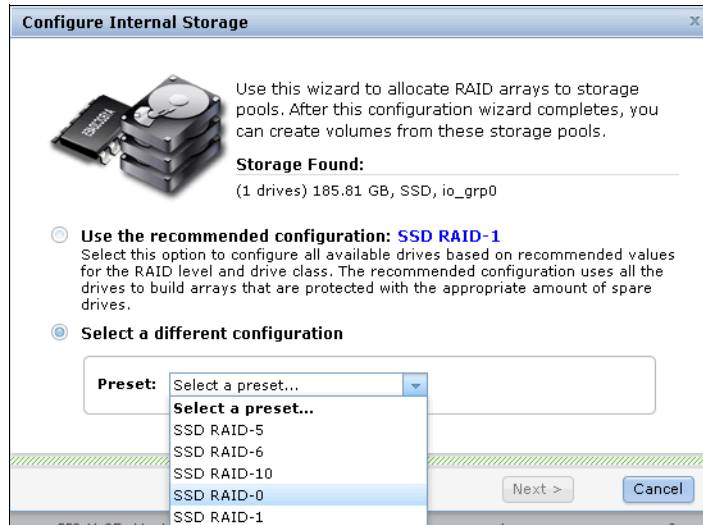


Figure 9-11 Select a preset menu

5. Choose a custom RAID level, or you can also select the SSD Easy Tier preset to review and modify the recommended configuration. Because we do not have enough drives in our configuration, the SSD Easy Tier preset is not available from the preset selection. When it is available, this preset configures a RAID 10 array with a spare goal of one drive. In this example, we create a RAID 0 array, although this is not best practice and is not used in a production environment. Because there are not enough drives, an error message is displayed, as shown in Figure 9-12 on page 423.



Figure 9-12 Select RAID 0 preset

This error message can be avoided if the Automatically configure spares option is not chosen, as shown in Figure 9-13. A RAID 0 array with one drive and zero spares is created.

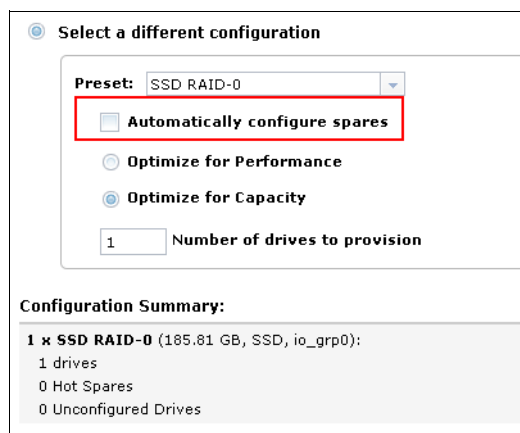


Figure 9-13 Array creation configuration summary

- To create a multitiered storage pool, the SSD must be added to an existing generic HDD pool. Select **Expand an existing pool** (as shown in Figure 9-14) and select the pool that you want to change to a multitiered storage pool. In our example, V5000_Pool_2 is selected. Click **Finish**.

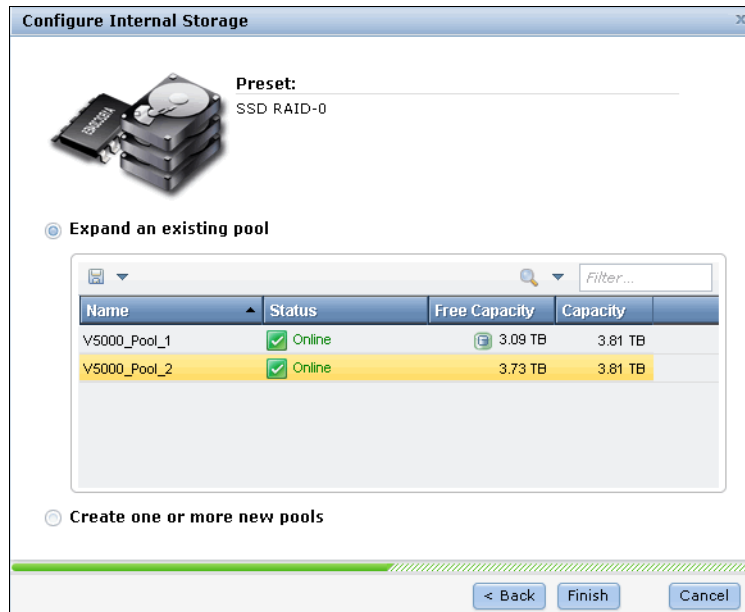


Figure 9-14 Expand an existing pool

- Now the array is configured on the SSDs and added to the selected storage pool. Click **Close** after the task completes, as shown in Figure 9-15.

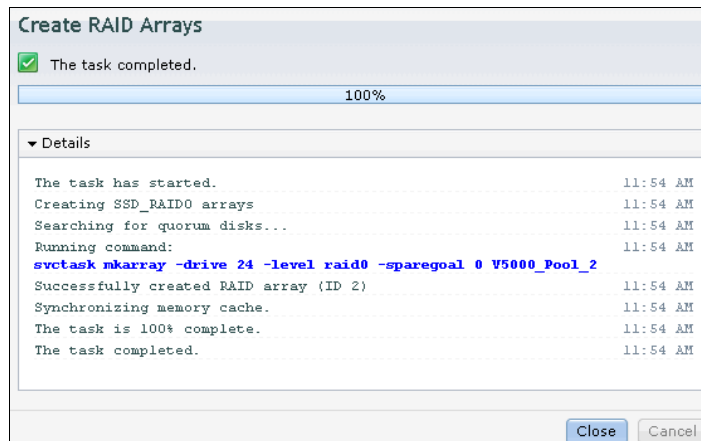


Figure 9-15 Array creation completed task

Figure 9-16 shows that the internal SSDs usage changed to Member and that the wizard created an MDisk that is named mdisk2.

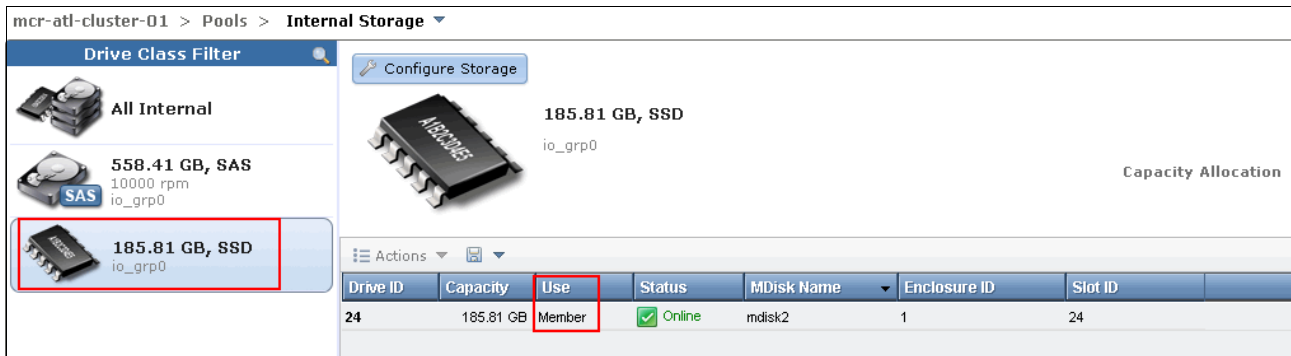


Figure 9-16 SSD usage is changed

In Figure 9-17, you see that the new MDisk is now part of the V5000_Poo1_2 storage pool and that the status of the Easy Tier changed to Active. In this pool, Automatic Data Placement Mode is started and the Easy Tier processes start to work.

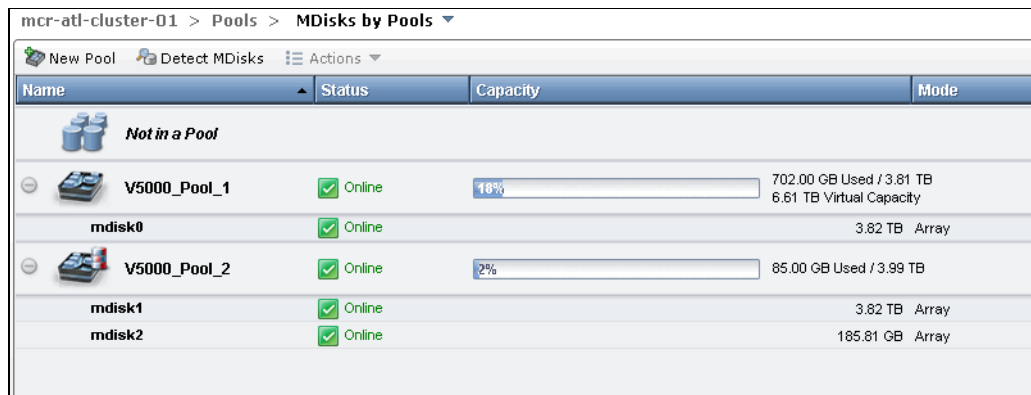



Figure 9-17 Easy Tier active

The storage pool was successfully changed to a multitiered storage pool (as indicated by the icon: ) and Easy Tier was activated by default. To reflect this change, we renamed the storage pool and changed the function icon, as described in Chapter 7, “Storage pools” on page 295 and as shown in Figure 9-18.

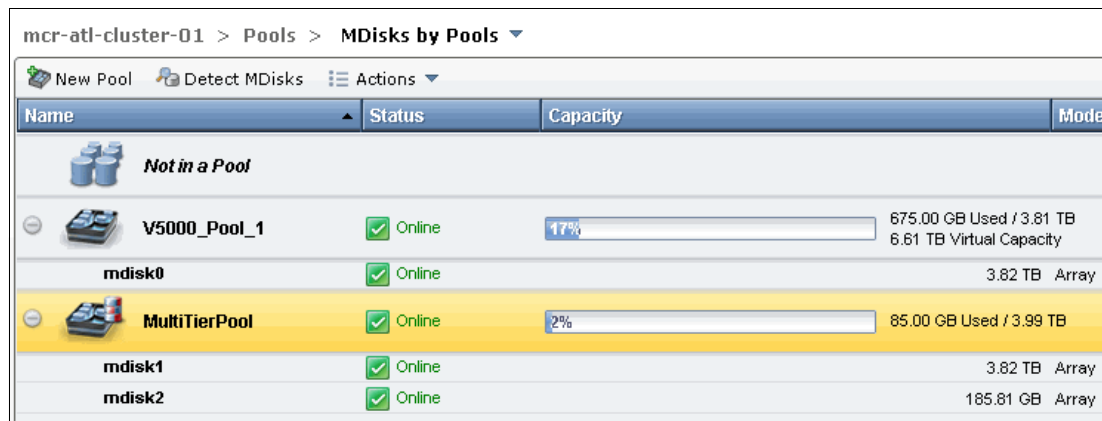


Figure 9-18 Multitiered storage pool

By default, Easy Tier is now active in this storage pool and all its volumes. Figure 9-19 shows three volumes on the multitiered storage pool.

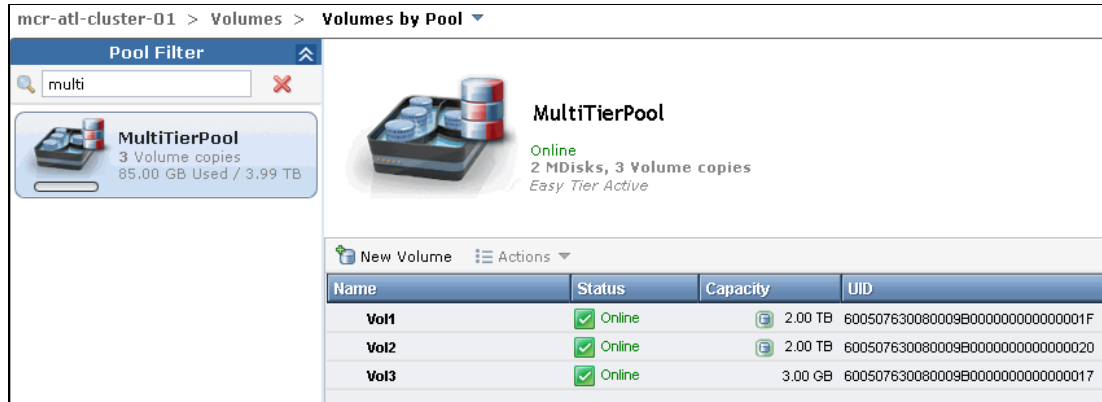


Figure 9-19 Volumes by Pool

If you open the properties of a volume by clicking **Actions** → **Properties**, you can also see that Easy Tier is enabled on the volume by default, as shown in Figure 9-20.

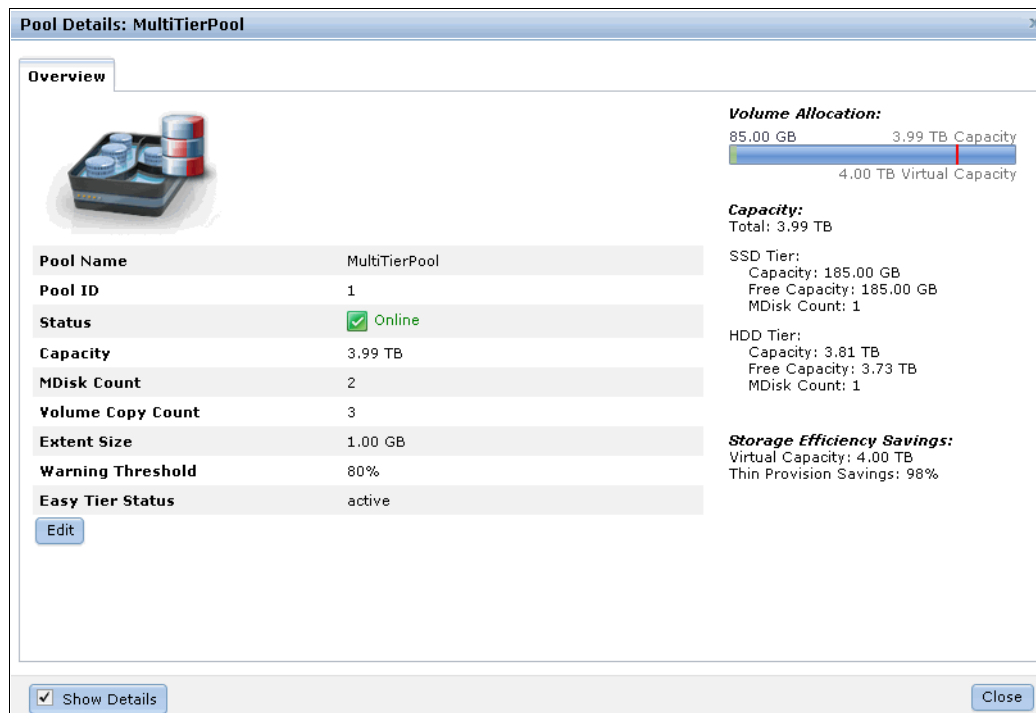


Figure 9-20 Easy Tier enabled volume

If a volume has more than one copy, Easy Tier can be enabled and disabled on each copy separately. This action depends on the storage pool where the volume copy is defined. You can see a volume with two copies that are stored in two different storage pools, as shown in Figure 9-21.

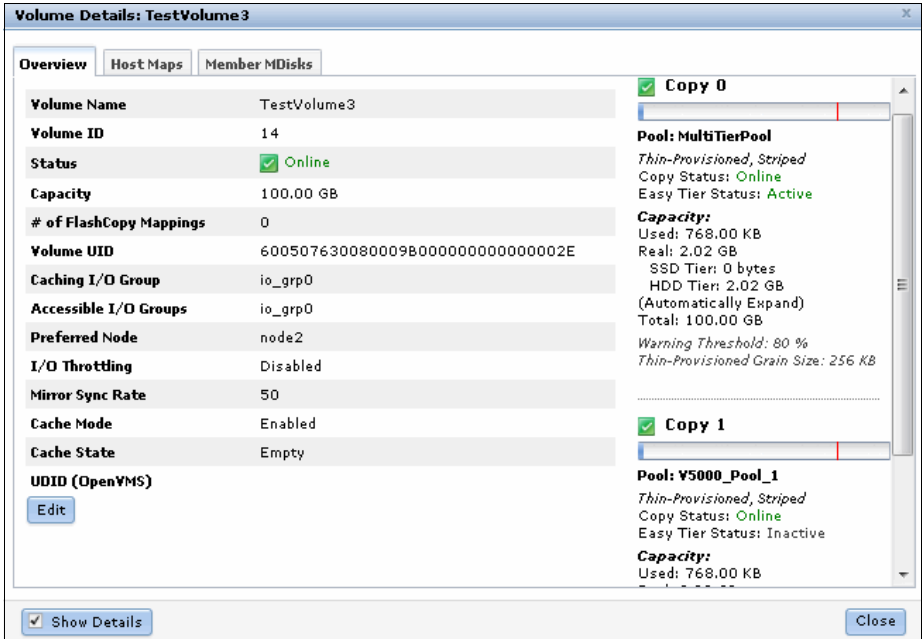


Figure 9-21 Easy Tier by Copy

If you want to enable Easy Tier on the second copy, change the storage pool of the second copy to a multitiered storage pool by repeating these steps.

If an external SSD is used, you must select the tier manually, and then add the external SSD MDisk to a storage pool, as described in as described in Chapter 7, “Storage pools” on page 295. This action also changes the storage pools to multitiered storage pools and enables Easy Tier on the pool and the volumes.

9.4.2 Downloading Easy Tier I/O measurements

Easy Tier is now enabled and Automatic Data Placement Mode is active. Extents are automatically migrated to, or from, high -performance disk tiers, and the statistic summary collection is now active. The statistics log file can be downloaded to analyze how many extents were migrated, and to monitor if it makes sense to add more SSDs to the multitiered storage pool.

To download the statistics file, complete the following steps:

1. Click **Settings** → **Support**, as shown in Figure 9-22.

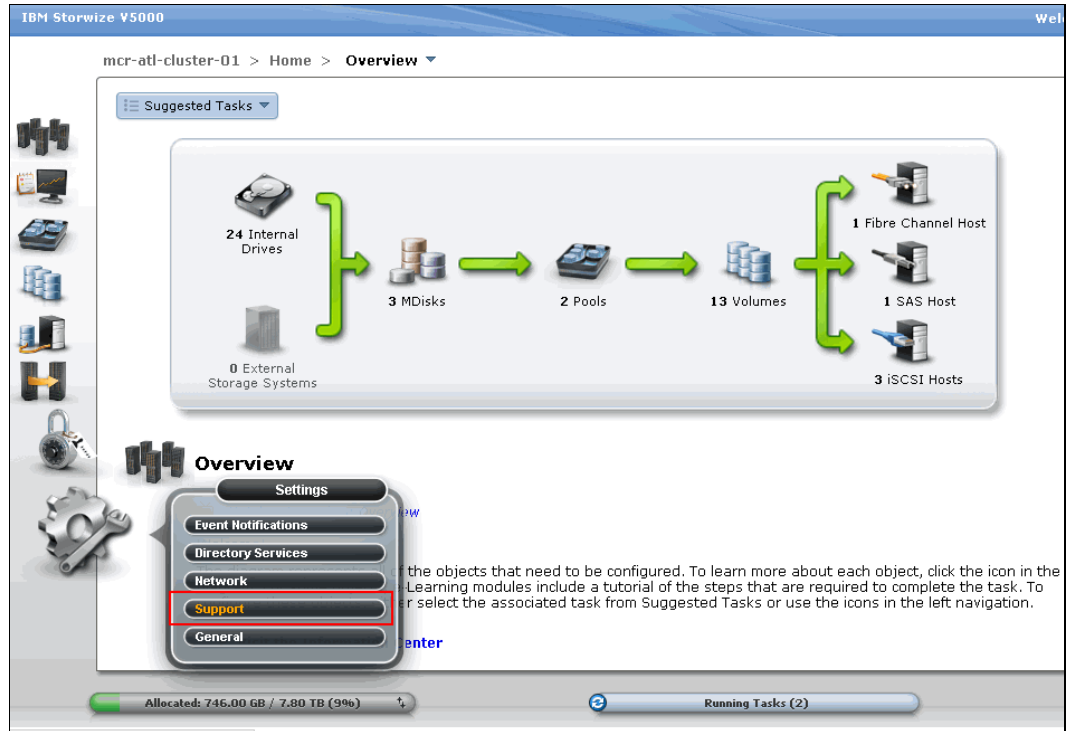


Figure 9-22 Settings menu

2. Click **Show full log listing**, as shown in Figure 9-23.

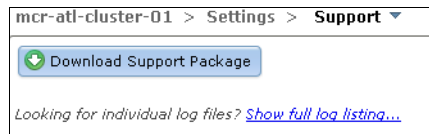


Figure 9-23 Download files menu

This action lists all the log files available to download, as shown in Figure 9-24. The Easy Tier log files are always named `dpa_heat.canister_name_date.time.data`.

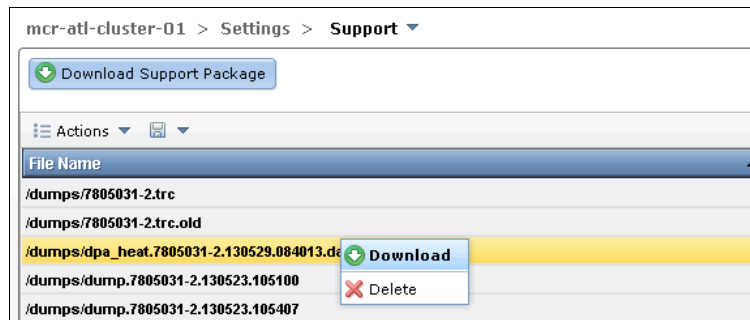


Figure 9-24 Download dpa_heat file

Log file creation: Depending on your workload and configuration, it can take up to 24 hours until a new Easy Tier log file is created.

If you run Easy Tier for a longer period, it generates a heat file at least every 24 hours.

The time and date of the file creation is included in the file name. The heat log file always includes the measured I/O activity of the last 24 hours.

3. Right-click the `dpa_heat.canister_name_date.time.data` file and click **Download**. Select the file for Easy Tier measurement for the most representative period.

You can also use the search field that is on the right to filter your search, as shown in Figure 9-25.

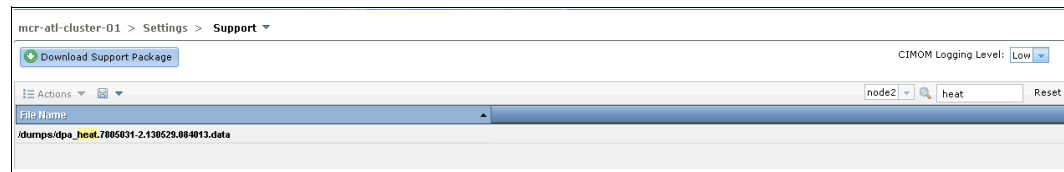


Figure 9-25 Filter your search

Depending on your browser settings, the file is downloaded to your default location, or you are prompted to save it to your computer. This file can be analyzed as described in 9.6, “IBM Storage Tier Advisor Tool” on page 434.

9.5 Easy Tier configuration by using the command-line interface

The process that is used to enable IBM Storwize V5000 Easy Tier by using the GUI is described in 9.4, “Easy Tier configuration by using the GUI” on page 419. Easy Tier can also be configured by using the command-line interface (CLI). For the advanced user, this method offers several more options for Easy Tier configuration.

Before you use the CLI, you must configure CLI access, as described in Appendix A, “Command-line interface setup and SAN Boot” on page 609.

Readability: In most examples that are shown in this section, many lines were deleted in the command output or responses so that we can concentrate only on the information that is related to Easy Tier.

9.5.1 Enabling Easy Tier evaluation mode

If you want to enable Easy Tier in evaluation mode, you must enable Easy Tier on a single-tiered storage pool. Connect to your IBM Storwize V5000 by using the CLI and run the `lsmdiskgrp` command, as shown in Example 9-1 on page 430. This command shows an overview about all configured storage pools and the Easy Tier status of the pool. In our example, there are two storage pools listed: `mdiskgrp0` with Easy Tier inactive, and `Multi_Tier_Pool` with Easy Tier enabled.

Example 9-1 List storage pools

```
IBM_2078:admin>lsmdiskgrp
id name                status mdisk_count ... easy_tier easy_tier_status
0 mdiskgrp0            online 3                ... auto     inactive
1 Multi_Tier_Pool     online 3                ... auto     active
```

To get a more detailed view of the single-tiered storage pool, run the **lsmdiskgrp storage pool name** command, as shown in Example 9-2.

Example 9-2 Storage Pools details: Easy Tier inactive

```
IBM_2078:admin>lsmdiskgrp mdiskgrp0
id 0
name mdiskgrp0
status online
mdisk_count 3
...
easy_tier auto
easy_tier_status inactive
tier generic_ssd
tier_mdisk_count 0
...
tier generic_hdd
tier_mdisk_count 3
...
```

To enable Easy Tier on a single-tiered storage pool, run the **chmdiskgrp -easytier on storage pool name** command, as shown in Example 9-3. Because this storage pool does not have any SSD MDisks, it is not a multitiered storage pool; only measuring is available.

Example 9-3 Enable Easy Tier on a single-tiered storage pool

```
IBM_2078:admin>chmdiskgrp -easytier on mdiskgrp0
IBM_2078:admin>
```

Check the status of the storage pool again by running the **lsmdiskgrp storage pool name** command again, as shown in Example 9-4.

Example 9-4 Storage pool details: Easy Tier ON

```
IBM_2078:admin>lsmdiskgrp mdiskgrp0
id 0
name mdiskgrp0
status online
mdisk_count 3
vdisk_count 7
...
easy_tier on
easy_tier_status active
tier generic_ssd
tier_mdisk_count 0
...
tier generic_hdd
tier_mdisk_count 3
...
```

Run the `svcinfolsmdiskgrp` command again, as shown in Example 9-5. You see that Easy Tier is turned on the storage pool now, but Automatic Data Placement Mode is not active on the multitiered storage pool.

Example 9-5 Storage pool list

```
IBM_2078:admin>lsmdiskgrp
id name          status mdisk_count vdisk_count ... easy_tier easy_tier_status
0 mdiskgrp0      online 3          7          ... on       active
1 Multi_Tier_Pool online 3          0          ... auto    active
```

To get the list of all the volumes defined, run the `lsvdisk` command, as shown in Example 9-6. For this example, we are only interested in the `redhat1` volume.

Example 9-6 All volumes list

```
IBM_2078:admin>lsvdisk
id name      IO_group_id IO_group_name status      mdisk_grp_id
5 redhat1 0          io_grp0      online many    many
...
```

To get a more detailed view of a volume, run the `lsvdisk volume name` command, as shown in Example 9-7. This output shows two copies of a volume: Copy 0 is in a multitiered storage pool and Automatic Data Placement is active, Copy 1 is in the single-tiered storage pool, and Easy Tier evaluation mode is active, as indicated by the `easy_tier_status` measured line.

Example 9-7 Volume details

```
IBM_2078:admin>lsvdisk redhat1
id 5
name redhat1
IO_group_id 0
IO_group_name io_grp0
status online
mdisk_grp_id many
mdisk_grp_name many
capacity 10.00GB
...

copy_id 0
status online
sync yes
primary yes
mdisk_grp_id 1
mdisk_grp_name Multi_Tier_Pool
...
easy_tier on
easy_tier_status active
tier generic_ssd
tier_capacity 0.00MB
tier generic_hdd
tier_capacity 10.00GB
...

copy_id 1
status online
sync yes
```

```

primary no
mdisk_grp_id 0
mdisk_grp_name mdiskgrp0
....
easy_tier on
easy_tier_status measured
tier generic_ssd
tier_capacity 0.00MB
tier generic_hdd
tier_capacity 10.00GB
...

```

These changes are also reflected in the GUI, as shown in Figure 9-26. Select the **Show Details** option to view the details of the Easy Tier for each of the volume copies.

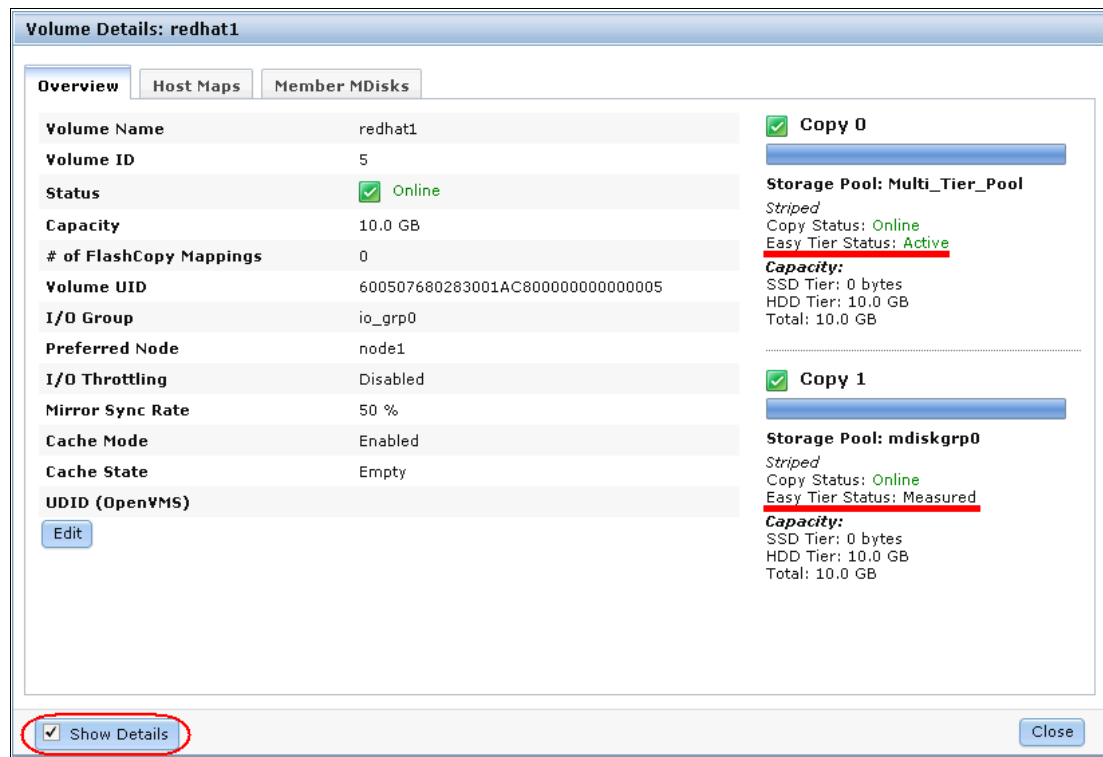


Figure 9-26 Easy Tier status by volume

Easy Tier evaluation mode is now active on the single-tiered storage pool (mdiskgrp0), but only for measurement. For more information about downloading and analyzing the I/O statistics, see 9.4.2, “Downloading Easy Tier I/O measurements” on page 427.

9.5.2 Enabling or disabling Easy Tier on single volumes

If you enable Easy Tier on a storage pool, all volume copies inside the Easy Tier pools also have Easy Tier enabled by default. This setting applies to multitiered and single-tiered storage pools. It is also possible to turn Easy Tier on and off for single volume copies.

To disable Easy Tier on single volumes, run the `chvdisk -easytier off volume name` command, as shown in Example 9-8.

Example 9-8 Disable Easy Tier on a single volume

```
IBM_2078:admin>chvdisk -easytier off redhat1
IBM_2078:admin>
```

This command disables Easy Tier on all copies of this volume. Example 9-9 shows that the Easy Tier status of the copies did change, even if Easy Tier is still enabled on the storage pool.

Example 9-9 Easy Tier disabled

```
IBM_2078:admin>lsvdisk redhat1
id 5
name redhat1
IO_group_id 0
IO_group_name io_grp0
status online
mdisk_grp_id many
mdisk_grp_name many
capacity 10.00GB
...

copy_id 0
status online
sync yes
primary yes
mdisk_grp_id 1
mdisk_grp_name Multi_Tier_Pool
...
easy_tier off
easy_tier_status measured
tier generic_ssd
tier_capacity 0.00MB
tier generic_hdd
tier_capacity 10.00GB
...

copy_id 1
status online
sync yes
primary no
mdisk_grp_id 0
mdisk_grp_name mdiskgrp0
....
easy_tier off
easy_tier_status measured
tier generic_ssd
tier_capacity 0.00MB
tier generic_hdd
tier_capacity 10.00GB
...
```

To enable Easy Tier on a volume, run the `chvdisk -easytier on volume name` command (as shown in Example 9-10), and the Easy Tier Status changes back to enabled (as shown in Example 9-7 on page 431).

Example 9-10 Easy Tier enabled

```
IBM_2078:admin>chvdisk -easytier on redhat1
IBM_2078:admin>
```

9.6 IBM Storage Tier Advisor Tool

The STAT is a Windows console tool. If you run Easy Tier in evaluation mode, the tool analyzes the extents and estimates how much benefit you derive if you implement Easy Tier Automatic Data Placement with SSD MDisks. If Automatic Data Placement Mode is already active, the analysis also includes an overview of migrated hot data and recommendations about whether you can derive any benefit by adding more SSD drives. The output provides a graphical representation of the performance data that is collected by Easy Tier over a 24-hour operational cycle.

9.6.1 Creating graphical reports

STAT takes input from the `dpa_heat` log file and produces an HTML file that contains the report. Download the `heat_log` file, as described in 9.4.2, “Downloading Easy Tier I/O measurements” on page 427, and save it to the HDD of a Windows system.

For more information about the tool and to download it, see this website:

<http://www-01.ibm.com/support/docview.wss?uid=ssg1S4000935>

Click **Start** → **Run**, enter `cmd`, and then click **OK** to open a command prompt.

Typically, the tool is installed in the `C:\Program Files\IBM\STAT` directory. Enter the command to generate the report, as shown in Example 9-11.

```
C:\Program Files\IBM\STAT>STAT.exe -o c:\directory_where_you_want_the_output_to_go
c:\location_of_dpa_heat_data_file
```

If you do not specify `-o c:\directory_where_you_want_the_output_to_go`, the output goes to the directory where the `STAT.exe` file is located.

Example 9-11 Generate HTML file

```
C:\EasyTier>STAT.exe -o C:\EasyTier C:\StorwizeV5000_Logs\dpa_heat.31G00KV-1.101
209.131801.data
```

```
CMUA00019I The STAT.exe command has completed.
```

```
C:\EasyTier>
```

Browse the directory where you directed the output file, and there is a file named `index.html`. Open the file by using your browser to view the report.

9.6.2 STAT reports

If you open the `index.html` file of an IBM Storwize V5000 system that is in Easy Tier evaluation mode, a window opens that gives you an estimate of the benefit if you were to add SSDs, as shown in Figure 9-27. The report shows a heading of IBM Storwize V7000. However, you can ignore this heading because this tool was originally available for IBM Storwize V7000 but also works with IBM Storwize V5000.

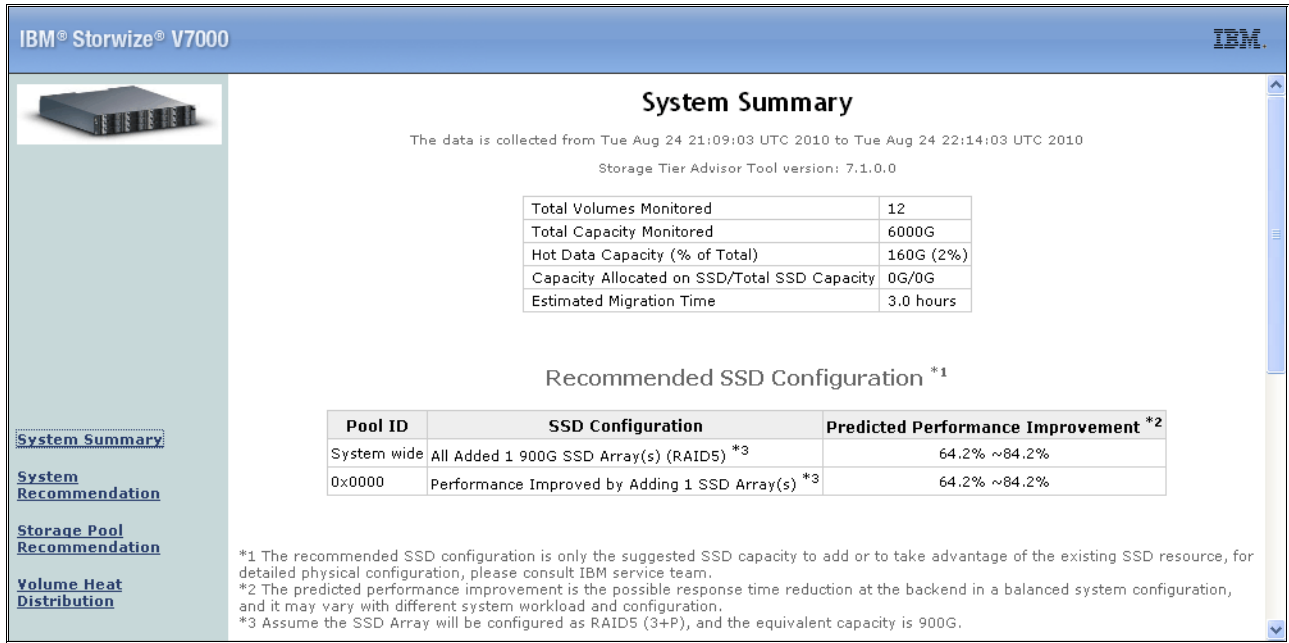


Figure 9-27 STAT report: System Summary

Important: Because this tool was originally available for SAN Volume Controller and IBM Storwize V7000, you can ignore the fact that it is showing IBM Storwize V7000 in the report banner.

The System Summary window provides the most important numbers. In Figure 9-27, we see that 12 volumes were monitored with a total capacity of 6000 GB. The result of the analysis of the hot extents is that about 160 GB (which means 2%) should be migrated to the high-performance disk tier.

It also recommends that one SSD RAID 5 array should be added as a high-performance tier that consists of four SSD drives (3+P). This predicted performance improvement is the possible response time reduction at the back end in a balanced system is 64% - 84%.

Click **Volume Heat Distribution** to change to a more detailed view, as shown in Figure 9-28.

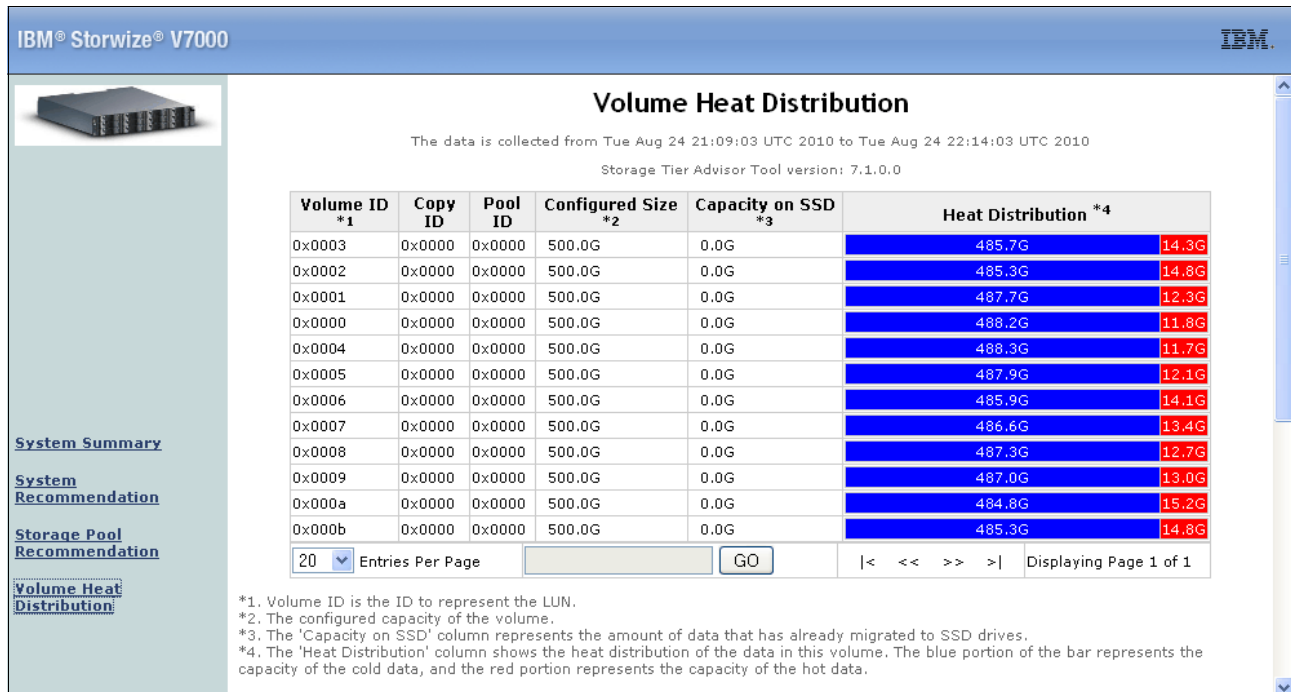


Figure 9-28 Volume Heat Distribution window

The table that is shown in Figure 9-28 gives you a more detailed view as to how the hot extents are distributed across your system. It contains the following information:

- ▶ Volume ID: The unique ID of each volume on the IBM Storwize V5000.
- ▶ Copy ID: If a volume owns more than one copy, the data is measured for each copy.
- ▶ Pool ID: The unique ID of each pool that is configured on the IBM Storwize V5000.
- ▶ Configured Size: The configured size of each volume that is represented to the host.
- ▶ Capacity on SSD: Capacity of the volumes on high-performance disk tier (even in evaluation mode, volumes can be on high performance disk tiers if they were moved there before).
- ▶ Heat Distribution: Shows the heat distribution of the data in this volume. The blue portion of the bar represents the capacity of the cold extents and the red portion represents the capacity of the hot extents. The red hot data is a candidate to be moved to the high-performance disk tier.

9.7 Tivoli Storage Productivity Center

The IBM Tivoli Storage Productivity Center provides a set of policy-driven automated tools for managing storage capacity, availability, events, performance, and assets in your enterprise environment. Tivoli Storage Productivity Center provides storage management from the host and application to the target storage device. It also provides disk and tape subsystem configuration and management, Performance Management, SAN fabric management and configuration, and usage reporting and monitoring. In this section, we describe how to use Tivoli Storage Productivity Center to get usage reporting and to monitor performance data.

Tivoli Storage Productivity Center can help you to identify, evaluate, control, and predict your enterprise storage management assets. Because it is policy-based, it can detect potential problems and automatically make adjustments that are based on the policies and actions that you define. For example, it can notify you when your system is running out of disk space or warn you of an impending storage hardware failure. By alerting you to these and other issues that are related to your stored data, you can prevent unnecessary system and application downtime.

9.7.1 Tivoli Storage Productivity Center benefits

Tivoli Storage Productivity Center includes the following benefits:

- ▶ Simplifies the management of storage infrastructures
- ▶ Manages, configures, and provisions SAN-attached storage
- ▶ Monitors and tracks performance of SAN-attached devices
- ▶ Monitors, manages and controls (through zones) SAN fabric components
- ▶ Manages the capacity usage and availability of the file systems and databases
- ▶ Offers performance monitoring and reporting
- ▶ Reports can be viewed by using a web-based GUI

9.7.2 Adding IBM Storwize V5000 in Tivoli Storage Productivity Center

After the Tivoli Storage Productivity Center is installed, it is ready to connect to the IBM Storwize V5000 system.

Complete the following steps to connect Tivoli Storage Productivity Center to the IBM Storwize V5000 system:

1. Open your browser and use the following link to start Tivoli Storage Productivity Center, as show in Figure 9-29:

http://TPC_system_Hostname:9550/ITSRM/app/en_US/index.html

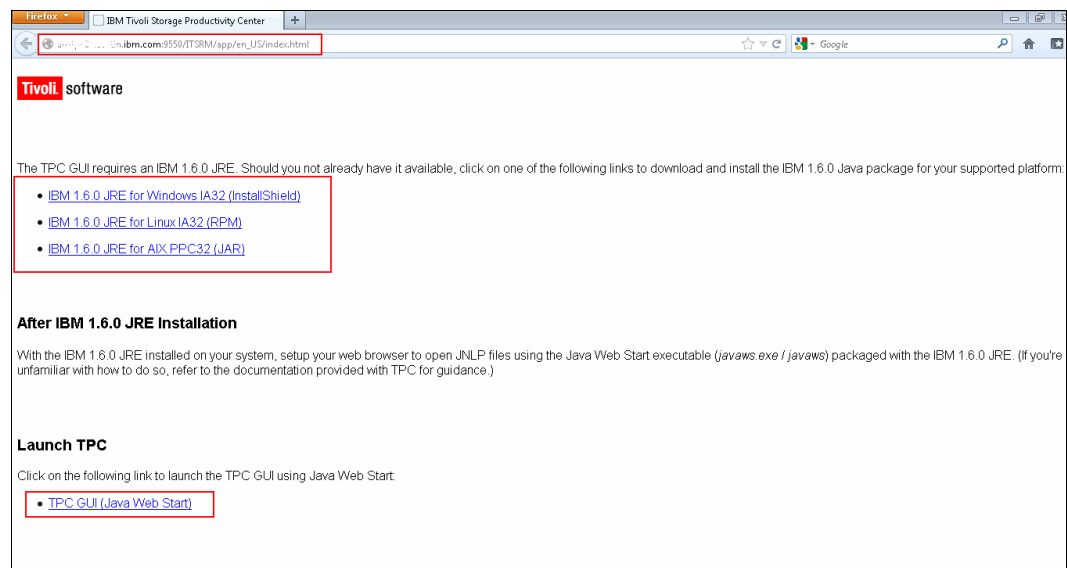


Figure 9-29 Tivoli Storage Productivity Center front page

You also can find a link on the webpage to download IBM Java, if required. To start Tivoli Storage Productivity Center console, click the Tivoli Storage Productivity Center GUI (Java Web Start).

Tivoli Storage Productivity Center starts an application download, as shown in Figure 9-30. If this is the first time you logged in, it takes time to install the required Java packages to the local system.

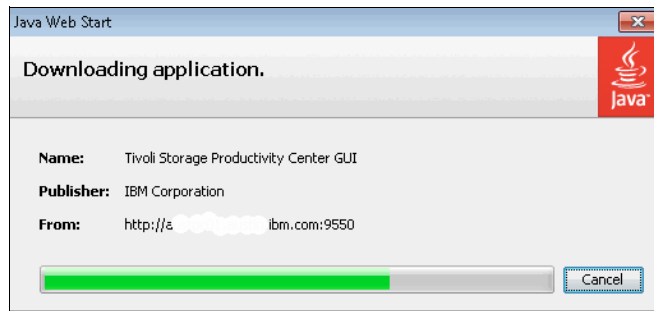


Figure 9-30 Downloading Tivoli Storage Productivity Center application

2. Use your login credentials to access Tivoli Storage Productivity Center, as shown in Figure 9-31.

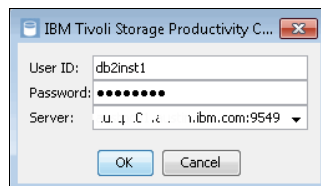


Figure 9-31 Tivoli Storage Productivity Center login access

3. After successfully logging in, you are ready to add storage devices to Tivoli Storage Productivity Center, as shown in Figure 9-32.

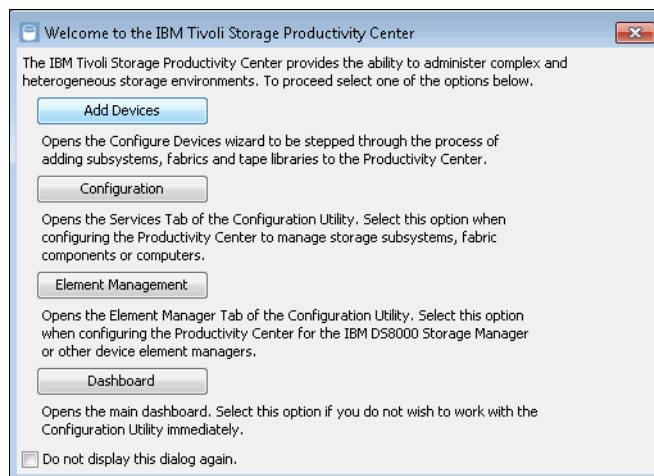


Figure 9-32 Add Devices window

4. Enter the details of your IBM Storwize V5000 in Tivoli Storage Productivity Center, as shown in Figure 9-33.

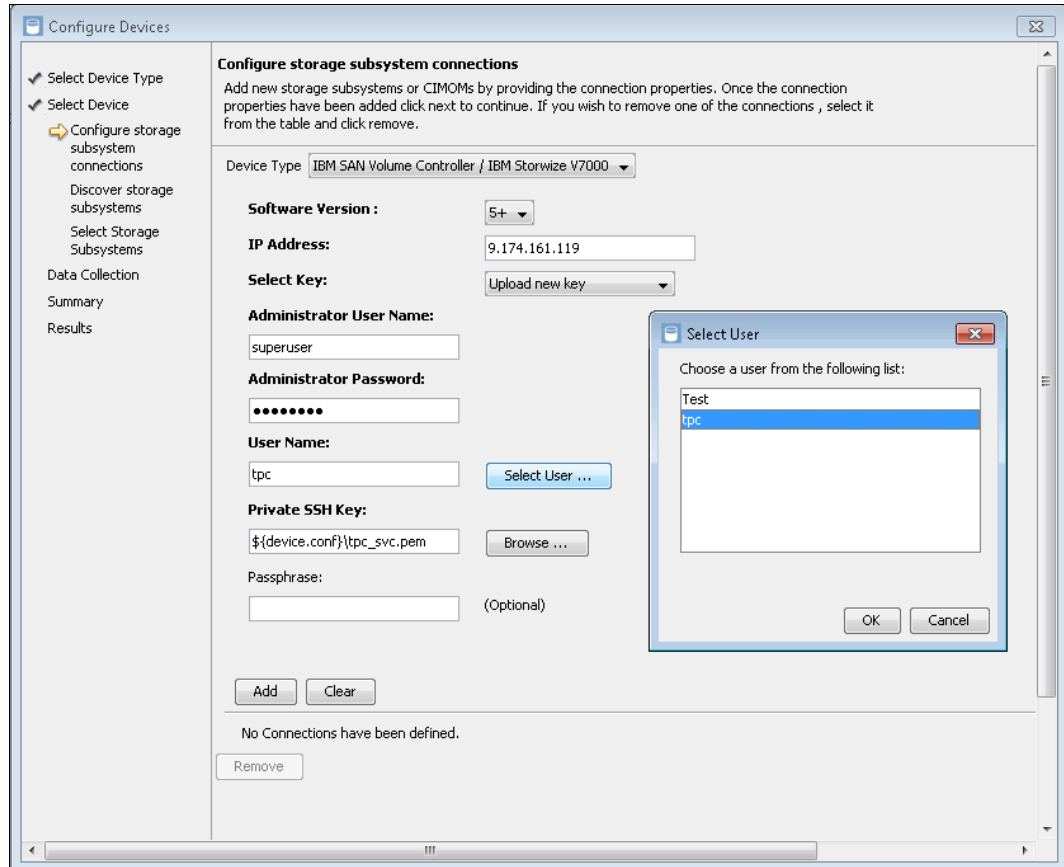


Figure 9-33 Configure device in Tivoli Storage Productivity Center

Continue to following the wizard after you complete all the required fields. After the wizard is completed, Tivoli Storage Productivity Center collects information from IBM Storwize V5000. A summary of details is shown at the end of discovery process.

9.8 Administering and reporting an IBM Storwize V5000 system through Tivoli Storage Productivity Center

This section shows examples of how to use Tivoli Storage Productivity Center to administer, configure, and generate reports for IBM Storwize V5000 system. A detailed description about Tivoli Storage Productivity Center reporting is beyond the intended scope of this book.

9.8.1 Basic configuration and administration

By using Tivoli Storage Productivity Center, you can administer, monitor, and configure your IBM Storwize V5000 system. However, not all of the options that are normally associated with the IBM Storwize V5000 GUI or CLI are available.

After successfully adding your IBM Storwize V5000 system, click **Disk Manager** → **Storage Subsystems** to view your configured devices, as shown in Figure 9-34.

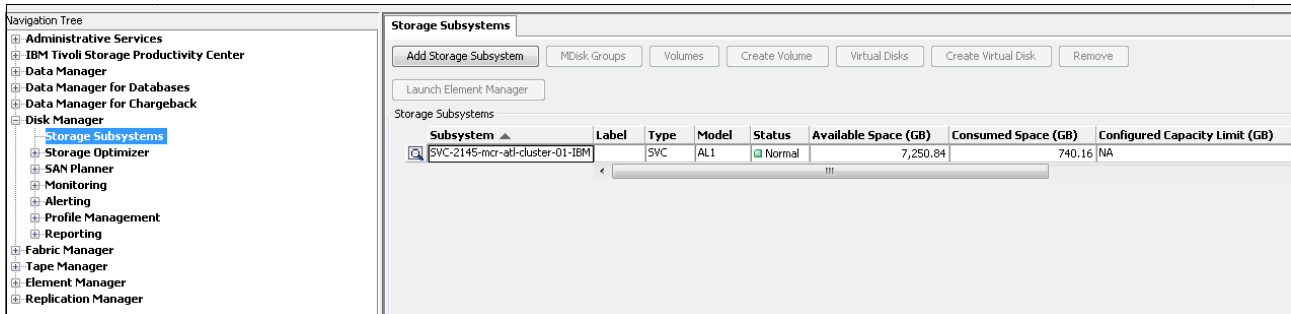


Figure 9-34 Storage Subsystem view

When you highlight the IBM Storwize V5000 system, action buttons become available that you can use to view the device configuration or create virtual disks, as shown in Figure 9-35.

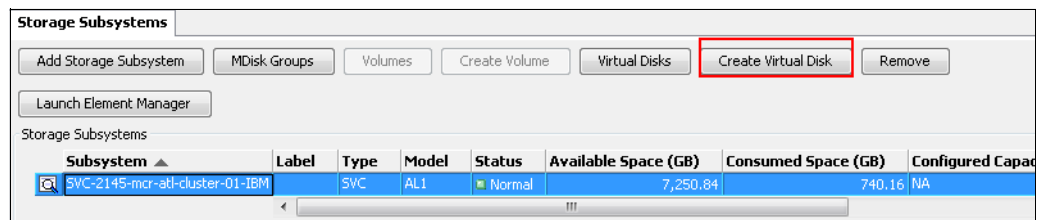


Figure 9-35 Action buttons

The MDisk Groups option provides you with a detailed list of the configured MDisk groups including, pool space, available space, configured space, and Easy Tier Configuration.

The Virtual Disks option lists all the configured disks with the added option to filter them by MDisk Group. The list includes several attributes, such as, capacity, volume type, and type.

Important: Tivoli Storage Productivity Center and SAN Volume Controller use the following terms:

- ▶ Virtual Disk: The equivalent of a Volume on a Storwize device
- ▶ MDisk Group: The equivalent of a Storage Pool on a Storwize device.

If you click **Create Virtual Disk**, the Create Virtual Disk wizard window opens, as shown in Figure 9-36 on page 441. Use this window to create volumes and specify several options, such as, size, name, thin provisioning, and add MDisks to an MDisk group.

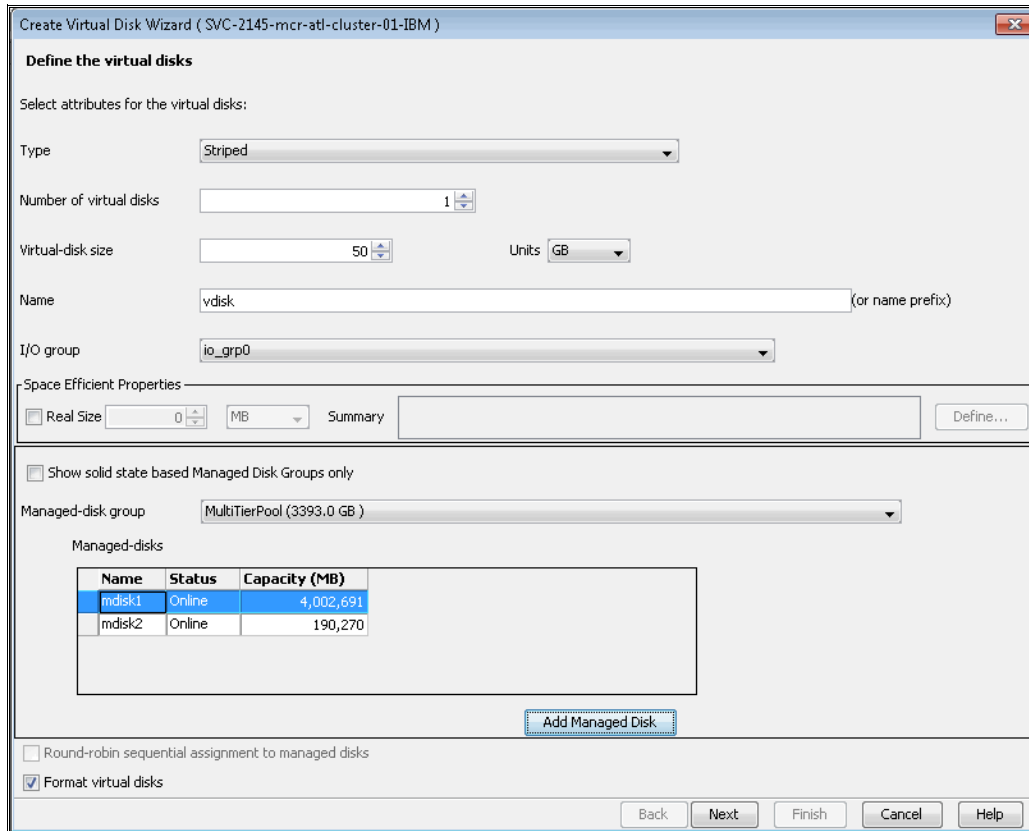


Figure 9-36 Virtual Disk wizard Creation

9.8.2 Generating reports by using Java GUI

In this section, we describe how to generate sample reports by using the GUI. We also create a probe to collect information from IBM Storwize V5000, as shown in Figure 9-37.

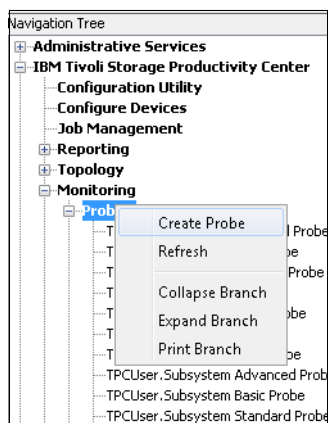


Figure 9-37 Create Probe option

Add the IBM Storwize V5000 probe for collecting information, as shown in Figure 9-38.

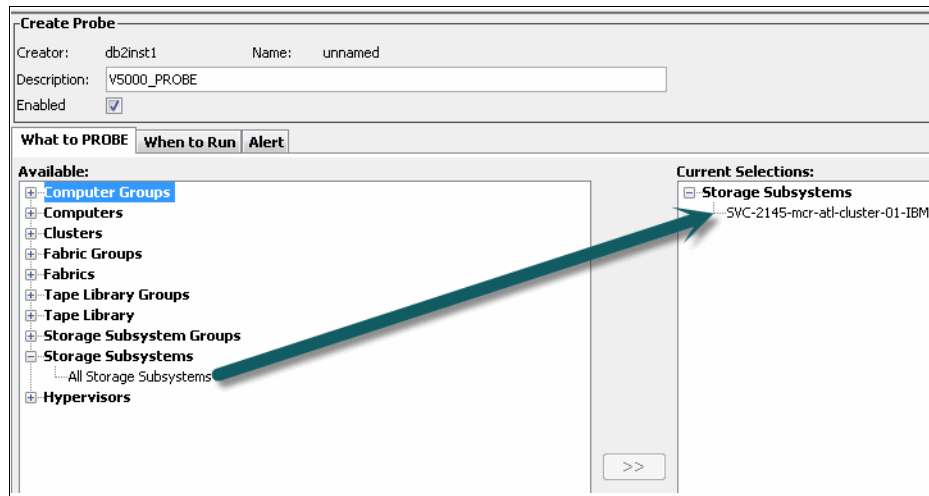


Figure 9-38 Adding IBM Storwize V5000 in probe

After you create a probe, you can click **Create Subsystem Performance Monitor**, as shown in Figure 9-39.

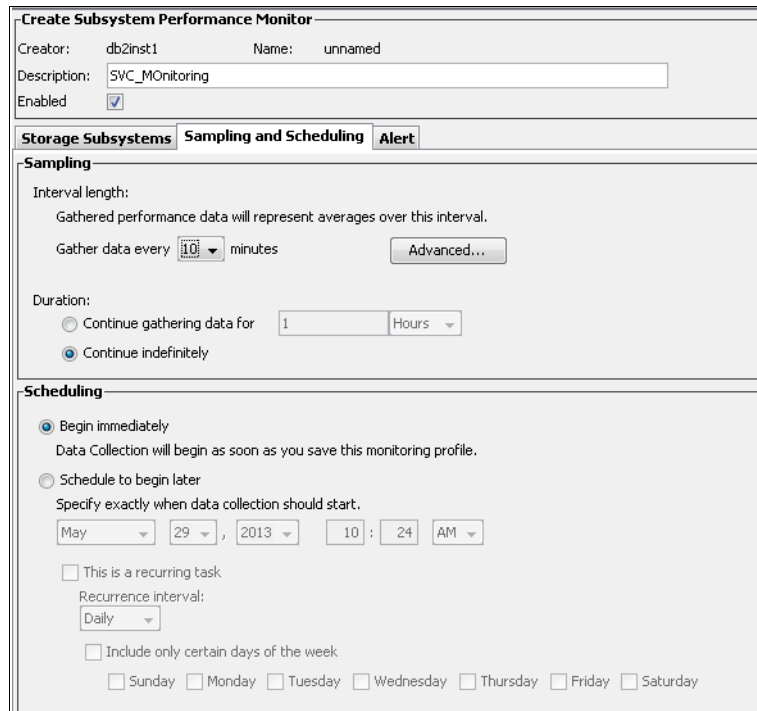


Figure 9-39 Create subsystem performance monitor

To check the MDisk performance, click **Disk Manager** → **Reporting** → **Storage Subsystem Performance** → **By Managed Disk**. You see many options to include in the wizard to check MDisk performance, as shown in Figure 9-40.

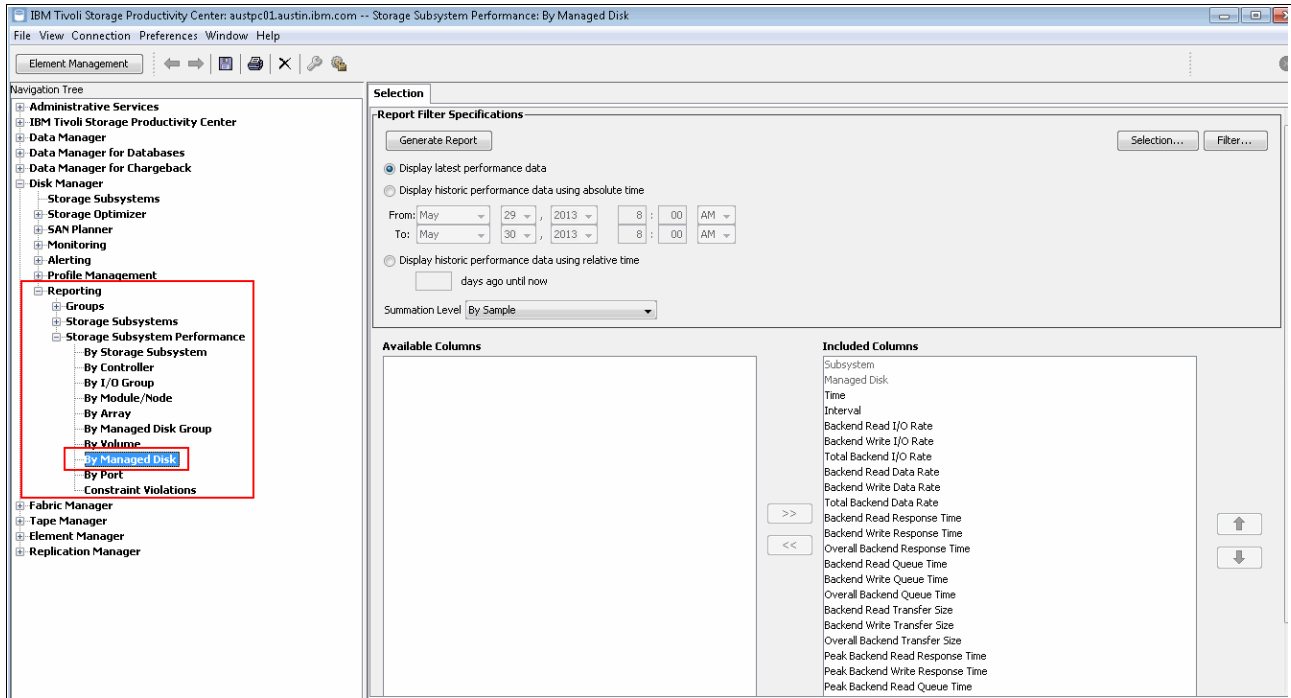


Figure 9-40 Managed disk performance report filter specification

Click **Generate Report** to see a report, as shown in Figure 9-41.

Selection		Managed Disks						
Storage Subsystem Performance: By Managed Disk								
Number of Rows: 3								
Subsystem	Managed Disk	Time	Interval	Backend Read I/O Rate	Backend Write I/O Rate	Total Backend I/O Rate		
SVC-2145-mcr-atl-cluster-01-IBM	mdisk0	May 30, 2013 7:50:20 AM	600 s	8 ops/s	0 ops/s	8 o		
SVC-2145-mcr-atl-cluster-01-IBM	mdisk1	May 30, 2013 7:50:20 AM	600 s	0 ops/s	8.01 ops/s	8.01 o		
SVC-2145-mcr-atl-cluster-01-IBM	mdisk2	May 30, 2013 7:50:20 AM	600 s	0 ops/s	0.37 ops/s	0.37 o		

Figure 9-41 MDisk performance report

Click the upper left icon to see a history chart report of the selected MDisk, as shown in Figure 9-42 on page 444.

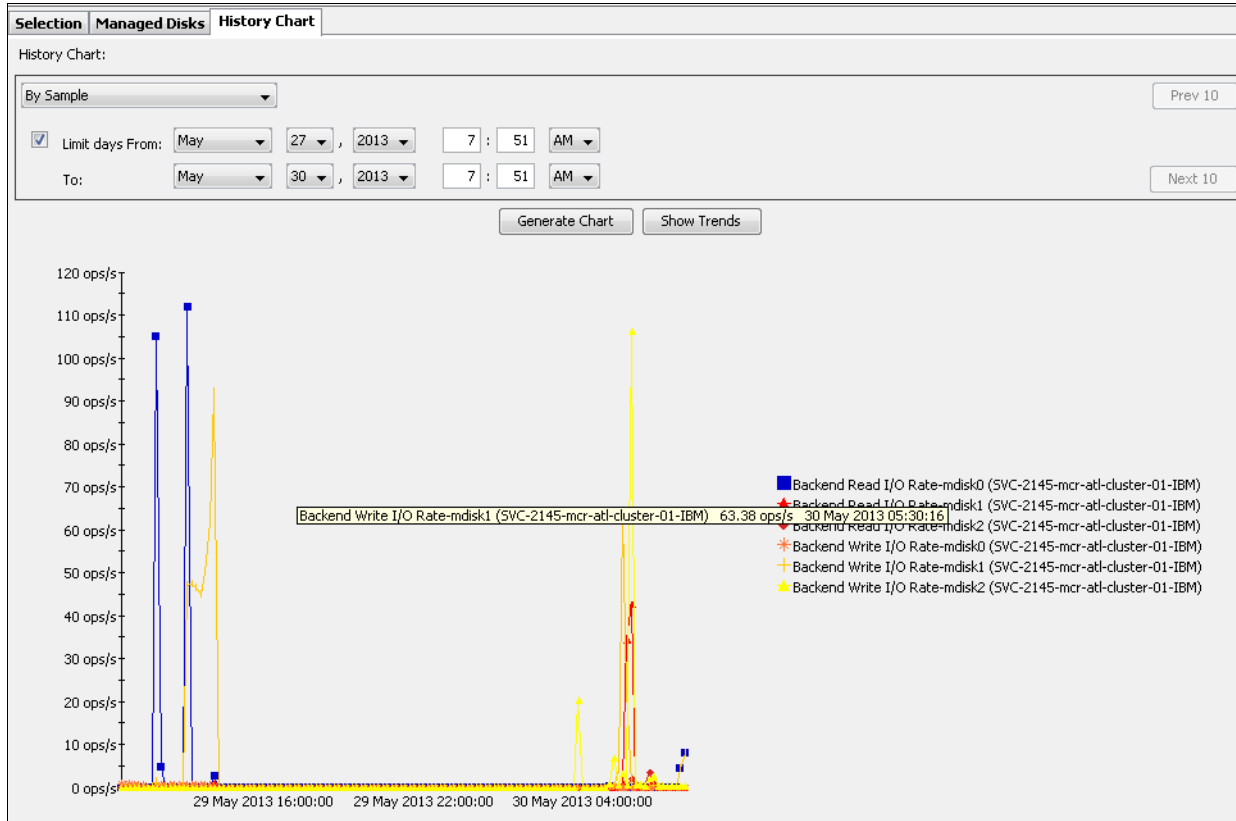


Figure 9-42 MDisk history chart

9.8.3 Generating reports by using Tivoli Storage Productivity Center web console

In this section, we describe how to generate reports by using the Tivoli Storage Productivity Center web console.

To connect to the web page, browse to the following URL:

<https://tpchostname.com:9569/srm/>

You see a login panel (as shown in Figure 9-43) and log in by using your Tivoli Storage Productivity Center credentials.



Figure 9-43 Tivoli Storage Productivity Center login panel

After you log in, you see the Tivoli Storage Productivity Center web dashboard, as shown in Figure 9-44. The Tivoli Storage Productivity Center web-based GUI is used to show information about the storage resources in your environment. It contains predefined and custom reports about performance and storage tiering.



Figure 9-44 Tivoli Storage Productivity Center Dashboard

You can use IBM Tivoli Common Reporting to view predefined reports and create custom reports from the web-based GUI. Predefined reports are also included, as shown in Figure 9-45.



Figure 9-45 Tivoli Storage Productivity Center web-based reporting

Figure 9-46 shows how to select predefined Storage Tiering reporting.

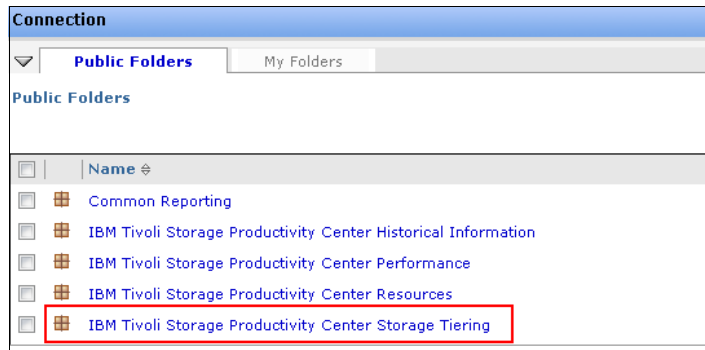


Figure 9-46 Tivoli Storage Productivity Center Storage tiering reporting

Figure 9-47 shows the different report options for Storage Tiering.



Figure 9-47 Details reports

Figure 9-48 shows the output from VDisk Details report.

IBM® Tivoli®																	
VDisk Details Report																	
Storage Virtualizer				SVC-2145-mcr-atl-cluster-01-IBM						VDisk				win			
Date filter				All													
MDisk Group	VDisk	VDisk Report Date	VDisk Capacity (GB)	VDisk Thin Provisioning Capacity Status	VDisk Peak Utilization (%)	VDisk Read I/O Rate (ops/s)	VDisk Max Read I/O Rate (ops/s)	VDisk Average Read Response Time (ms/op)	VDisk Max Read Response Time (ms/op)	VDisk Max Read Response Time Threshold of {70 %} Exceeded By (hr.:min.)	VDisk Write I/O Rate (ops/s)	VDisk Average Write Response Time (ms/op)	VDisk Max Write Response Time (ms/op)	VDisk Max Write Response Time Threshold of {90 %} Exceeded By (hr.:min.)	VDisk Max Read Density Rate (ops/s per GB)	VDisk I/O Rate Share of Total VDisk Read I/O Rate (%)	VDisk Capacity Share of Total VDisk Capacity (%)
MultiTierPool	win_2k8-iscsi	May	5.00	OK	0.06	0.00	0.06	8.25	8.36	00:00	0.00	1.34	1.34	00:00	0.01	0.08	0

Figure 9-48 VDisk Details report

Figure 9-49 shows the Report Overview in pie-chart format.

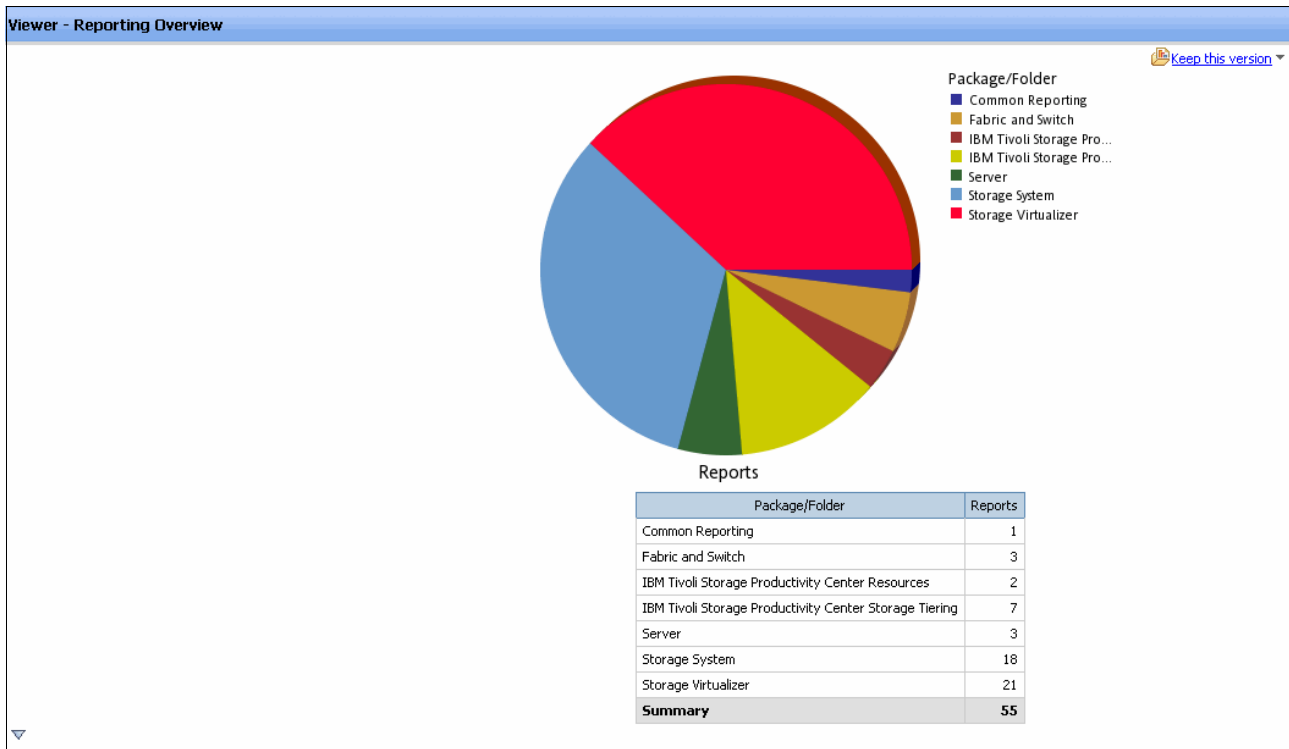


Figure 9-49 Reporting Overview

Figure 9-50 shows the Easy Tier usage for volumes. To open this report in Tivoli Storage Productivity Center, click **Storage Resources** → **Volumes**.

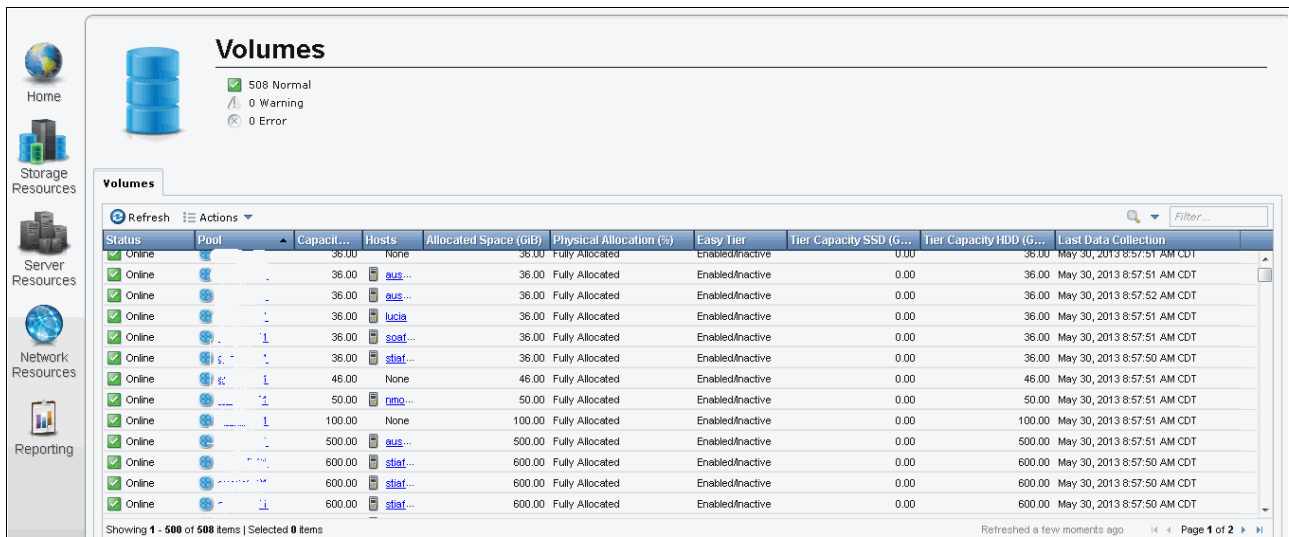


Figure 9-50 Volume Easy Tier usage

Figure 9-51 shows a detailed list of storage pools.

Name	Physical Allocation (%)	Shortfall (%)	Solid St...	Easy Tier	Tier Capacity SSD (...)	Tier Capacity HDD (G...	Tier Available Space SSD (G...	Last Data Collection
	34%	0%	No	Auto/Inactive	0.00	25,098.75	0.00	May 30, 2013 8:57:26 AM CDT
	34%	0%	No	Auto/Inactive	0.00	25,098.75	0.00	May 30, 2013 8:57:26 AM CDT
	77%	0%	No	Auto/Active	3,339.75	102,710.75	0.00	May 30, 2013 8:57:26 AM CDT
	70%	0%	No	Auto/Active	3,339.75	102,712.25	0.00	May 30, 2013 8:57:26 AM CDT
	97%	0%	No	Auto/Active	3,339.75	102,718.00	0.00	May 30, 2013 8:57:26 AM CDT
	98%	0%	No	Auto/Active	3,339.75	102,712.25	0.00	May 30, 2013 8:57:26 AM CDT
MultiTierPool	99%	100%	No	N/A				May 30, 2013 9:01:01 AM CDT
V5000_Pool_1	6%	12%	No	N/A				May 30, 2013 9:01:01 AM CDT
V5000_Pool_3	1%	0%	No	N/A				May 30, 2013 9:01:01 AM CDT

Figure 9-51 Pool Easy Tier information

Figure 9-52 shows Storage Virtualized Pool details in graph format.

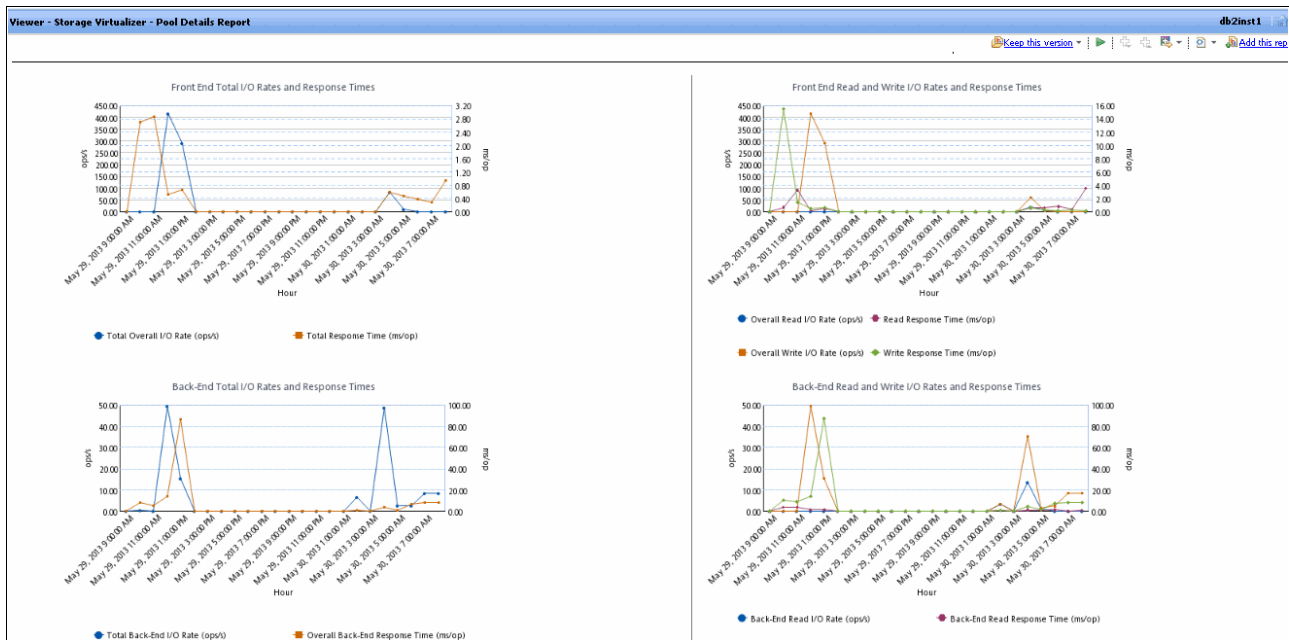


Figure 9-52 Pool details



Copy services

In this chapter, we describe the copy services functions that are provided by the IBM Storwize V5000 storage system, including FlashCopy and Remote Copy. Copy services functions are useful for making data copies for backup, application test, recovery, and so on. The IBM Storwize V5000 system makes it easy to apply these functions to your environment through its intuitive GUI.

This chapter includes the following topics:

- ▶ FlashCopy
- ▶ Remote Copy
- ▶ Troubleshooting Remote Copy
- ▶ Managing Remote Copy by using the GUI

10.1 FlashCopy

By using the FlashCopy function of the IBM Storwize V5000 storage system, you can create a *point-in-time copy* of one or more volumes. In this section, we describe the structure of FlashCopy and provide details about its configuration and use.

You can use FlashCopy to solve critical and challenging business needs that require the duplication of data on your source volume. Volumes can remain online and active while you create consistent copies of the data sets. Because the copy is performed at the block level, it operates below the host operating system and cache and, therefore, is not apparent to the host.

Flushing: Because FlashCopy operates at the block level below the host operating system and cache, those levels do need to be flushed for consistent FlashCopy copies.

While the FlashCopy operation is performed, I/O to the source volume is frozen briefly to initialize the FlashCopy bitmap and then is allowed to resume. Although several FlashCopy options require the data to be copied from the source to the target in the background (which can take time to complete), the resulting data on the target volume copy appears to complete immediately. This task is accomplished by using a bitmap (or bit array) that tracks changes to the data after the FlashCopy is started, and an indirection layer, which allows data to be read from the source volume transparently.

10.1.1 Business requirements for FlashCopy

When you are deciding whether FlashCopy addresses your needs, you must adopt a combined business and technical view of the problems you must solve. Determine your needs from a business perspective, and then determine whether FlashCopy fulfills the technical needs of those business requirements.

With an immediately available copy of the data, FlashCopy can be used in the following business scenarios:

- ▶ Rapidly creating consistent backups of dynamically changing data

FlashCopy can be used to create backups through periodic running; the FlashCopy target volumes can be used to complete a rapid restore of individual files or the entire volume through Reverse FlashCopy (by using the `-restore` option).

The target volumes that are created by FlashCopy can also be used for backup to tape. By attaching them to another server and performing backups from there, it allows the production server to continue largely unaffected. After the copy to tape completes, the target volumes can be discarded or kept as a rapid restore copy of the data.

- ▶ Rapidly creating consistent copies of production data to facilitate data movement or migration between hosts

FlashCopy can be used to facilitate the movement or migration of data between hosts while it minimizes downtime for applications. FlashCopy allows application data to be copied from source volumes to new target volumes while applications remain online. After the volumes are fully copied and synchronized, the application can be stopped and then immediately started on the new server that is accessing the new FlashCopy target volumes. This mode of migration is faster than other migration methods that are available through the IBM Storwize V5000 because the size and the speed of the migration is not as limited.

- ▶ Rapidly creating copies of production data sets for application development and testing
Under normal circumstances to perform application development and testing, data must be restored from traditional backup media, such as, tape. Depending on the amount of data and the technology in use, this process can easily take a day or more. With FlashCopy, a copy can be created and be online for use in just a few minutes. The time varies based on the application and the data set size.
- ▶ Rapidly creating copies of production data sets for auditing purposes and data mining
Auditing or data mining normally require the usage of the production applications. This situation can cause high loads for databases track inventories or similar data. With FlashCopy, you can create copies for your reporting and data mining activities. This feature reduces the load on your production systems, which increases their performance.
- ▶ Rapidly creating copies of production data sets for quality assurance
Quality assurance is an interesting case for FlashCopy. Because traditional methods involve so much time and labor, the refresh cycle is typically extended. This reduction in time that is required allows much more frequent refreshes of the quality assurance database.

10.1.2 FlashCopy functional overview

FlashCopy occurs between a source volume and a target volume. The source and target volumes must be the same size. Multiple FlashCopy mappings (source-to-target relationships) can be defined, and point-in-time consistency can be maintained across multiple point-in-time mappings by using consistency groups. For more information about FlashCopy consistency groups, see “FlashCopy consistency groups” on page 456.

The minimum granularity that IBM Storwize V5000 storage system supports for FlashCopy is an entire volume; it is not possible to use FlashCopy to copy only part of a volume. Additionally, the source and target volumes must belong to the same IBM Storwize V5000 storage system, but they do not have to be in the same storage pool.

Before you start a FlashCopy (regardless of the type and options that are specified), the IBM Storwize V5000 must put the cache into write-through mode, which flushes the I/O that is bound for the source volume. If you are scripting FlashCopy operations from the CLI, you must run the `prestartfcmap` or `prestartfcconsistgrp` command. However, this step is managed for you and carried out automatically by the GUI. This is not the same as flushing the host cache, which is not required. After FlashCopy is started, an effective copy of a source volume to a target volume is created. The content of the source volume is immediately presented on the target volume and the original content of the target volume is lost. This FlashCopy operation is also referred to as a *time-zero copy* (T_0).

Immediately following the FlashCopy operation, the source and target volumes are available for use. The FlashCopy operation creates a bitmap that is referenced and maintained to direct I/O requests within the source and target relationship. This bitmap is updated to reflect the active block locations as data is copied in the background from the source to target and updates are made to the source.

Figure 10-1 shows the redirection of the host I/O toward the source volume and the target volume.

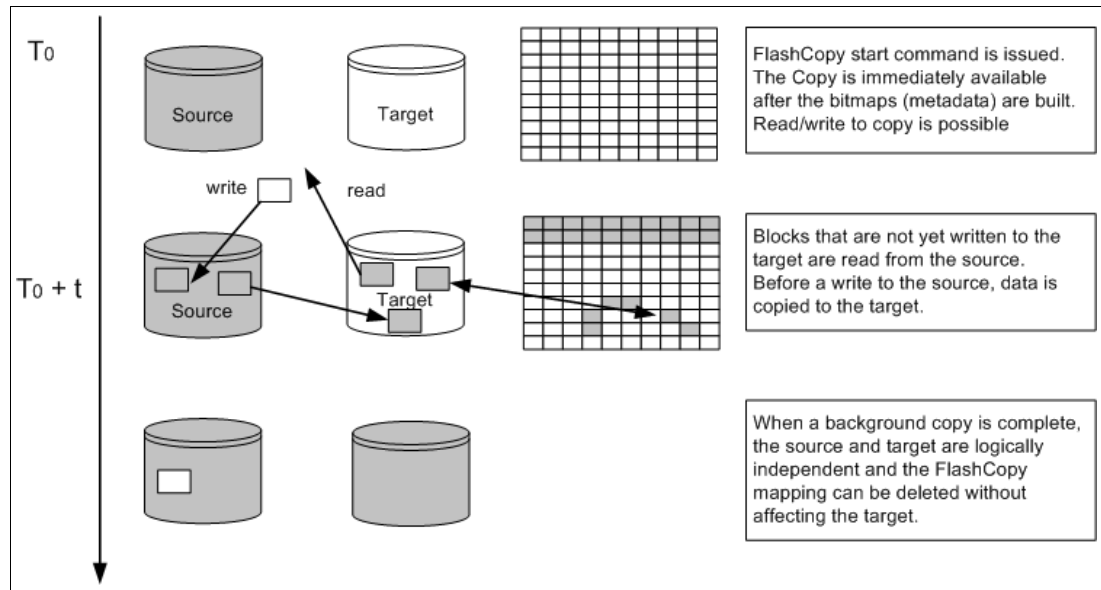


Figure 10-1 Redirection of host I/O

When data is copied between volumes, it is copied in units of address space known as *grains*. Grains are units of data that are grouped to optimize the use of the bitmap that track changes to the data between the source and target volume. You have the option of using 64 KB or 256 KB grain sizes (256 KB is the default). The FlashCopy bitmap contains 1 bit for each grain and is used to track whether the source grain was copied to the target. The 64 KB grain size uses bitmap space at a rate of four times the default 256 KB size.

The FlashCopy bitmap dictates the following read and write behaviors for the source and target volumes:

- ▶ Read I/O request to source: Reads are performed from the source volume the same as for non-FlashCopy volumes.
- ▶ Write I/O request to source: Writes to the source cause the grains of the source volume to be copied to the target if they were not already and then the write is performed to the source.
- ▶ Read I/O request to target: Reads are performed from the target if the grains already were copied; otherwise, the read is performed from the source.
- ▶ Write I/O request to target: Writes to the target cause the grain to be copied from the source to the target first, unless the entire grain is being written and then the write completes to the target only.

FlashCopy mappings

A FlashCopy mapping defines the relationship between a source volume and a target volume. FlashCopy mappings can be stand-alone mappings or a member of a consistency group, as described in “FlashCopy consistency groups” on page 456.

Incremental FlashCopy mappings

In an incremental FlashCopy, the initial mapping copies all of the data from the source volume to the target volume. Subsequent FlashCopy mappings copy only data that was modified since the initial FlashCopy mapping. This action reduces the amount of time that it takes to re-create an independent FlashCopy image. You can define a FlashCopy mapping as incremental only when you create the FlashCopy mapping.

Multiple target FlashCopy mappings

You can copy up to 256 target volumes from a single source volume. Each relationship between a source and target volume is managed by a unique mapping such that a single volume can be the source volume for up to 256 mappings.

Each of the mappings from a single source can be started and stopped independently. If multiple mappings from the same source are active (in the copying or stopping states), a dependency exists between these mappings.

If a single source volume has multiple target FlashCopy volumes, the write to the source volume does not cause its data to be copied to all of the targets. Instead, it is copied to the newest target volume only. The older targets refer to new targets first before they refer to the source. A dependency relationship exists between a particular target and all newer targets that share a source until all data is copied to this target and all older targets.

Cascaded FlashCopy mappings

The cascaded FlashCopy function allows a FlashCopy target volume to be the source volume of another FlashCopy mapping. Up to 256 mappings can exist in a cascade. If cascaded mappings and multiple target mappings are used, a tree of up to 256 mappings can be created.

Cascaded mappings differ from multiple target FlashCopy mappings in depth. Cascaded mappings have an association in the manner of $A > B > C$, while multiple target FlashCopy has an association in the manner $A > B1$ and $A > B2$.

Background copy

The background copy rate is a property of a FlashCopy mapping that is defined as a value of 0 - 100. The background copy rate can be defined and dynamically changed for individual FlashCopy mappings. A value of 0 disables background copy. This option is also called the *no-copy option*, which provides pointer-based images for limited lifetime uses.

With FlashCopy background copy, the source volume data is copied to the corresponding target volume in the FlashCopy mapping. If the background copy rate is set to 0 (which means disable the FlashCopy background copy), only data that changed on the source volume is copied to the target volume. The benefit of using a FlashCopy mapping with background copy enabled is that the target volume becomes a real independent clone of the FlashCopy mapping source volume after the copy is complete. When the background copy is disabled, only the target volume is a valid copy of the source data while the FlashCopy mapping remains in place. Copying only the changed data saves your storage capacity (assuming it is thin provisioned and the `-rsize` option was correctly setup.)

The relationship of the background copy rate value to the amount of data that is copied per second is shown in Table 10-1 on page 454.

Table 10-1 Background copy rate

Value	Data that is copied per second	Grains per second (256 KB grain)	Grains per second (64 KB grain)
1 - 10	128 KB	0.5	2
11 - 20	256 KB	1	4
21 - 30	512 KB	2	8
31 - 40	1 MB	4	16
41 - 50	2 MB	8	32
51 - 60	4 MB	16	64
61 - 70	8 MB	32	128
71 - 80	16 MB	64	256
81 - 90	32 MB	128	512
91 - 100	64 MB	256	1024

Data copy rate: The data copy rate remains the same regardless of the FlashCopy grain size. The difference is the number of grains that are copied per second. The grain size can be 64 KB or 256 KB. The smaller size uses more bitmap space and thus limits the total amount of FlashCopy space possible. However, it might be more efficient regarding the amount of data that is moved, depending on your environment.

Cleaning rate

The cleaning rate provides a method for FlashCopy copies with dependant mappings (multiple target or cascaded) to complete their background copies before their source goes offline or is deleted after a stop is issued.

When you create or modify a FlashCopy mapping, you can specify a cleaning rate for the FlashCopy mapping that is independent of the background copy rate. The cleaning rate is also defined as a value of 0 - 100, which has the same relationship to data copied per second as the backup copy rate (see Table 10-1).

The cleaning rate controls the rate at which the cleaning process operates. The purpose of the cleaning process is to copy (or flush) data from FlashCopy source volumes upon which there are dependent mappings. For cascaded and multiple target FlashCopy, the source might be a target for another FlashCopy or a source for a chain (cascade) of FlashCopy mappings. The cleaning process must complete before the FlashCopy mapping can go to the stopped state. This feature and the distinction between *stopping* and *stopped* states was added to prevent data access interruption for dependent mappings when their source is issued a stop.

FlashCopy mapping states

A mapping is in one of the following states at any point:

- Idle or Copied

The source and target volumes act as independent volumes even if a mapping exists between the two. Read and write caching is enabled for the source and the target volumes.

If the mapping is incremental and the background copy is complete, the mapping records only the differences between the source and target volumes. The source and target volumes go offline if the connection to both nodes in the IBM Storwize V5000 storage system that the mapping is assigned to is lost.

► Copying

The copy is in progress. Read and write caching is enabled on the source and the target volumes.

► Prepared

The mapping is ready to start. The target volume is online, but is not accessible. The target volume cannot perform read or write caching. Read and write caching is failed by the SCSI front end as a hardware error. If the mapping is incremental and a previous mapping completed, the mapping records only the differences between the source and target volumes. The source and target volumes go offline if the connection to both nodes in the IBM Storwize V5000 storage system that the mapping is assigned to is lost.

► Preparing

The target volume is online, but not accessible. The target volume cannot perform read or write caching. Read and write caching is failed by the SCSI front end as a hardware error. Any changed write data for the source volume is flushed from the cache. Any read or write data for the target volume is discarded from the cache. If the mapping is incremental and a previous mapping completed, the mapping records only the differences between the source and target volumes. The source and target volumes go offline if the connection to both nodes in the IBM Storwize V5000 storage system that the mapping is assigned to is lost.

► Stopped

The mapping is stopped because you issued a stop command or an I/O error occurred. The target volume is offline and its data is lost. To access the target volume, you must restart or delete the mapping. The source volume is accessible and the read and write cache is enabled. If the mapping is incremental, the mapping is recording write operations to the source volume. The source and target volumes go offline if the connection to both nodes in the IBM Storwize V5000 storage system that the mapping is assigned to is lost.

► Stopping

The mapping is copying data to another mapping. If the background copy process is complete, the target volume is online while the stopping copy process completes. If the background copy process did not complete, data is discarded from the target volume cache. The target volume is offline while the stopping copy process runs. The source volume is accessible for I/O operations.

► Suspended

The mapping did start, but it did not complete. Access to the metadata is lost, which causes the source and target volume to go offline. When access to the metadata is restored, the mapping returns to the copying or stopping state and the source and target volumes return online. The background copy process resumes.

Any data that was not flushed and was written to the source or target volume before the suspension is in cache until the mapping leaves the suspended state.

FlashCopy consistency groups

Consistency groups address the requirement to preserve point-in-time data consistency across multiple volumes for applications that include related data that spans them. For these volumes, consistency groups maintain the integrity of the FlashCopy by ensuring that dependent writes are run in the application's intended sequence. For more information about dependent writes, see "Dependent writes" on page 456.

When consistency groups are used, the FlashCopy commands are issued to the FlashCopy consistency group, which performs the operation on all FlashCopy mappings that are contained within the consistency group.

Figure 10-2 shows a consistency group that consists of two FlashCopy mappings.

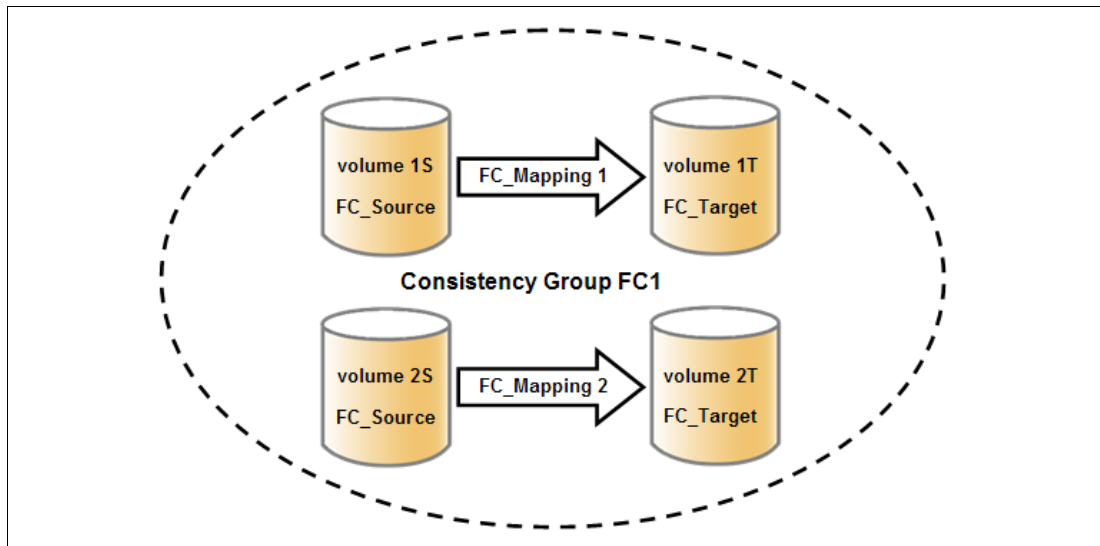


Figure 10-2 FlashCopy consistency group

FlashCopy mapping management: After an individual FlashCopy mapping was added to a consistency group, it can be managed only as part of the group. Operations such as, start and stop are no longer allowed on the individual mapping.

Dependent writes

To show why it is crucial to use consistency groups when a data set spans multiple volumes, consider the following typical sequence of writes for a database update transaction:

1. A write is run to update the database log, which indicates that a database update is about to be performed.
2. A second write is run to complete the actual update to the database.
3. A third write is run to update the database log, which indicates that the database update completed successfully.

The database ensures the correct ordering of these writes by waiting for each step to complete before it starts the next step. However, if the database log (updates 1 and 3) and the database (update 2) are on separate volumes, it is possible for the FlashCopy of the database volume to occur before the FlashCopy of the database log. This situation can result in the target volumes seeing writes (1) and (3) but not (2) because the FlashCopy of the database volume occurred before the write was completed.

In this case, if the database was restarted by using the backup that was made from the FlashCopy target volumes, the database log indicates that the transaction completed successfully when, in fact, it had not. This situation occurs because the FlashCopy of the volume with the database file was started (bitmap was created) before the write completed to the volume. Therefore, the transaction is lost and the integrity of the database is in question.

To overcome the issue of dependent writes across volumes and to create a consistent image of the client data, it is necessary to perform a FlashCopy operation on multiple volumes as an atomic operation by using consistency groups.

A FlashCopy consistency group can contain up to 512 FlashCopy mappings. The more mappings that you have, the more time it takes to prepare the consistency group. FlashCopy commands can then be issued to the FlashCopy consistency group and simultaneously for all of the FlashCopy mappings that are defined in the consistency group. For example, when the FlashCopy for the consistency group is started, all FlashCopy mappings in the consistency group are started at the same time, which results in a point-in-time copy that is consistent across all FlashCopy mappings that are contained in the consistency group.

A consistency group aggregates FlashCopy mappings, not volumes. Thus, where a source volume has multiple FlashCopy mappings, they can be in the same or separate consistency groups. If a particular volume is the source volume for multiple FlashCopy mappings, you might want to create separate consistency groups to separate each mapping of the same source volume. Regardless of whether the source volume with multiple target volumes is in the same consistency group or in separate consistency groups, the resulting FlashCopy produces multiple identical copies of the source data.

The consistency group can be specified when the mapping is created. You can also add the FlashCopy mapping to a consistency group or change the consistency group of a FlashCopy mapping later.

Important: Do not place stand-alone mappings into a consistency group because they become controlled as part of that consistency group.

FlashCopy consistency group states

A FlashCopy consistency group is in one of the following states at any point:

- ▶ Idle or Copied
All FlashCopy Mappings in this consistency group are in the Idle or Copied state.
- ▶ Preparing
At least one FlashCopy mapping in this consistency group is in the Preparing state.
- ▶ Prepared
The consistency group is ready to start. While in this state, the target volumes of all FlashCopy mappings in this consistency group are not accessible.
- ▶ Copying
At least one FlashCopy mapping in the consistency group is in the Copying state and no FlashCopy mappings are in the Suspended state.
- ▶ Stopping
At least one FlashCopy mapping in the consistency group is in the Stopping state and no FlashCopy mappings are in the Copying or Suspended state.

- ▶ Stopped

The consistency group is stopped because you issued a command or an I/O error occurred.
- ▶ Suspended

At least one FlashCopy mapping in the consistency group is in the Suspended state.
- ▶ Empty

The consistency group does not have any FlashCopy mappings.

Reverse FlashCopy

Reverse FlashCopy enables FlashCopy targets to become restore points for the source without breaking the FlashCopy relationship and without waiting for the original copy operation to complete. It supports multiple targets and multiple rollback points.

A key advantage of Reverse FlashCopy is that it does not delete the original target, thus allowing processes that use the target, such as, a tape backup, to continue uninterrupted.

You can also create an optional copy of the source volume that is made before the reverse copy operation is started. This copy restores the original source data, which can be useful for diagnostic purposes.

Figure 10-3 shows an example of the reverse FlashCopy scenario.

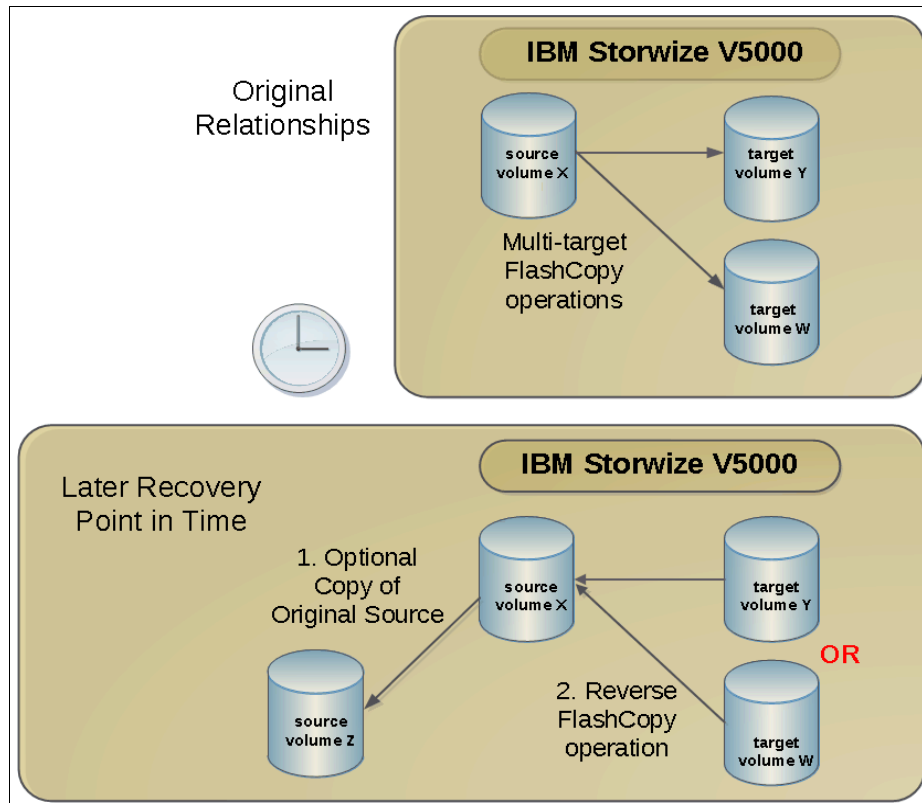


Figure 10-3 Reverse FlashCopy scenario

To restore from a FlashCopy backup by using the GUI, complete the following steps:

1. (Optional) Create a target volume (volume Z) and run FlashCopy on the production volume (volume X) to copy data on to the new target for later problem analysis.
2. Create a FlashCopy map with the backup to be restored (volume Y) or (volume W) as the source volume and volume X as the target volume.
3. Start the FlashCopy map (volume Y to volume X).

The -restore option: In the CLI, you must add the **-restore** option to the command **svctask startfcmap** manually. For more information about using the CLI, see Appendix A, “Command-line interface setup and SAN Boot” on page 609.

Regardless of whether the initial FlashCopy map (volume X to volume Y) is incremental, the Reverse FlashCopy operation copies only the modified data.

Consistency groups are reversed by creating a set of new “reverse” FlashCopy maps and adding them to a new “reverse” consistency group. Consistency groups cannot contain more than one FlashCopy map with the same target volume.

10.1.3 Planning for FlashCopy

There are several items that must be considered before a FlashCopy is performed, which are described in this section.

Guidelines for FlashCopy implementation

Consider the following guidelines for FlashCopy implementation:

- ▶ The source and target volumes must be on the same IBM Storwize V5000 storage system.
- ▶ The source and target volumes do not need to be in the same storage pool.
- ▶ The FlashCopy source and target volumes can be thin-provisioned.
- ▶ The source and target volumes must be the same size. The size of the source and target volumes cannot be altered (increased or decreased) while a FlashCopy mapping is defined.
- ▶ FlashCopy operations perform in direct proportion to the performance of the source and target disks. If you have a fast source disk and slow target disk, the performance of the source disk is reduced because it must wait for the write operation to occur at the target before it can write to the source.

Maximum configurations for FlashCopy

Table 10-2 shows some of the FlashCopy maximum configurations. For more information about the latest values, see the IBM Storwize V5000 Information Center at this website:

http://pic.dhe.ibm.com/infocenter/storwize/v5000_ic/index.jsp

Table 10-2 FlashCopy maximum configurations

FlashCopy property	Maximum
FlashCopy targets per source	256
FlashCopy mappings per cluster	4,096
FlashCopy consistency groups per cluster	127
FlashCopy mappings per consistency group	512

FlashCopy presets

The IBM Storwize V5000 storage system provides three FlashCopy presets (Snapshot, Clone, and Backup) to simplify the more common FlashCopy operations, as shown in Table 10-3.

Table 10-3 FlashCopy presets

Preset	Purpose
Snapshot	Creates a point-in-time view of the production data. The snapshot is not intended to be an independent copy. Instead, it is used to maintain a view of the production data at the time the snapshot is created. This preset automatically creates a thin-provisioned target volume with none of the capacity that is allocated at the time of creation. The preset uses a FlashCopy mapping with none of the background copy so that only data written to the source or target is copied to the target volume.
Clone	Creates an exact replica of the volume, which can be changed without affecting the original volume. After the copy operation completes, the mapping that was created by the preset is automatically deleted. This preset automatically creates a volume with the same properties as the source volume and creates a FlashCopy mapping with a background copy rate of 50. The FlashCopy mapping is configured to automatically delete when the FlashCopy mapping reaches 100% completion.
Backup	Creates a point-in-time replica of the production data. After the copy completes, the backup view can be refreshed from the production data, with minimal copying of data from the production volume to the backup volume. This preset automatically creates a volume with the same properties as the source volume. The preset creates an incremental FlashCopy mapping with a background copy rate of 50.

Presets: All of the presets can be adjusted by using the Advanced Settings expandable section in the GUI.

10.1.4 Managing FlashCopy by using the GUI

The IBM Storwize V5000 storage system provides a separate function icon to access copy service management. The following windows are available for managing FlashCopy under the Copy Services function icon:

- ▶ FlashCopy
- ▶ Consistency Groups
- ▶ FlashCopy Mappings

The Copy Services function icon is shown in Figure 10-4.

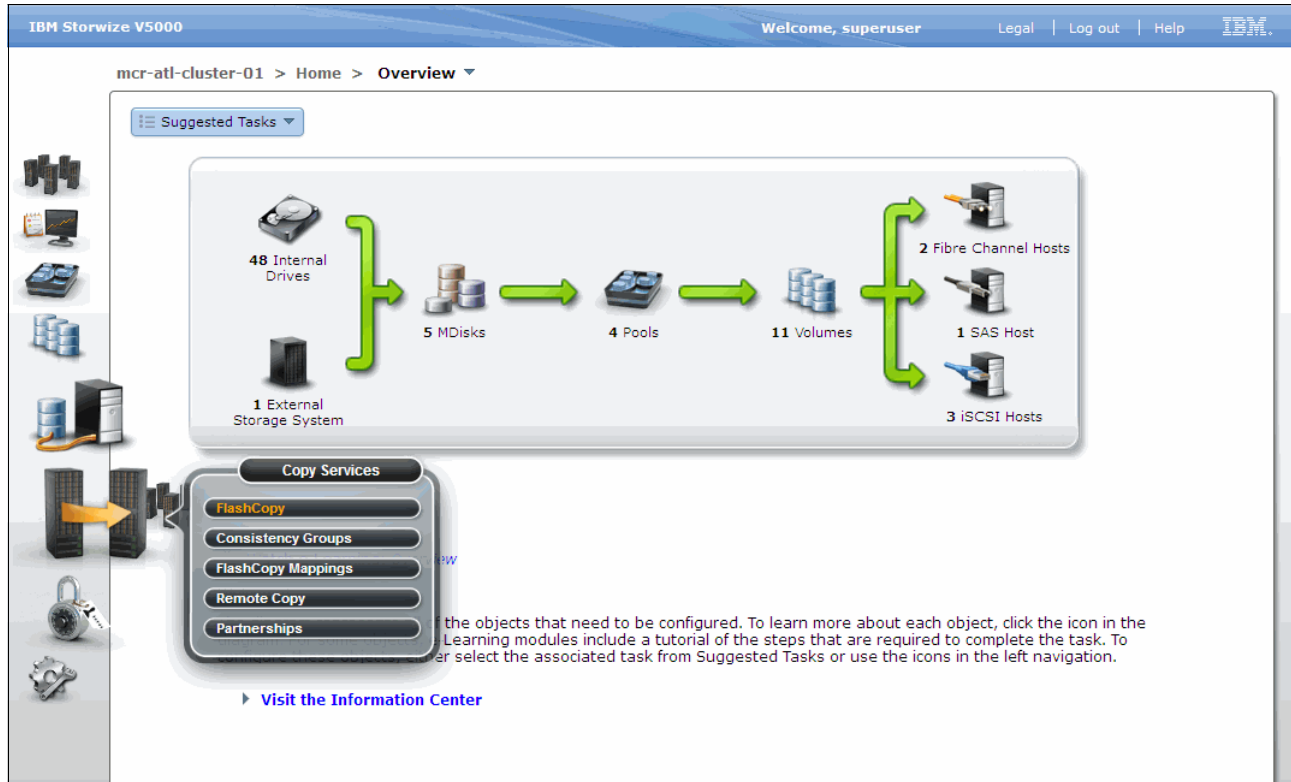


Figure 10-4 Copy Services function icon

Most of the actions to manage the FlashCopy mapping can be done in the FlashCopy window or the FlashCopy Mappings windows, although the quick path to create FlashCopy presets can be found only in the FlashCopy window.

Click **FlashCopy** in the Copy Services function icon menu and the FlashCopy window opens, as shown in Figure 10-5 on page 462. In the FlashCopy window, the FlashCopy mappings are organized by volumes.

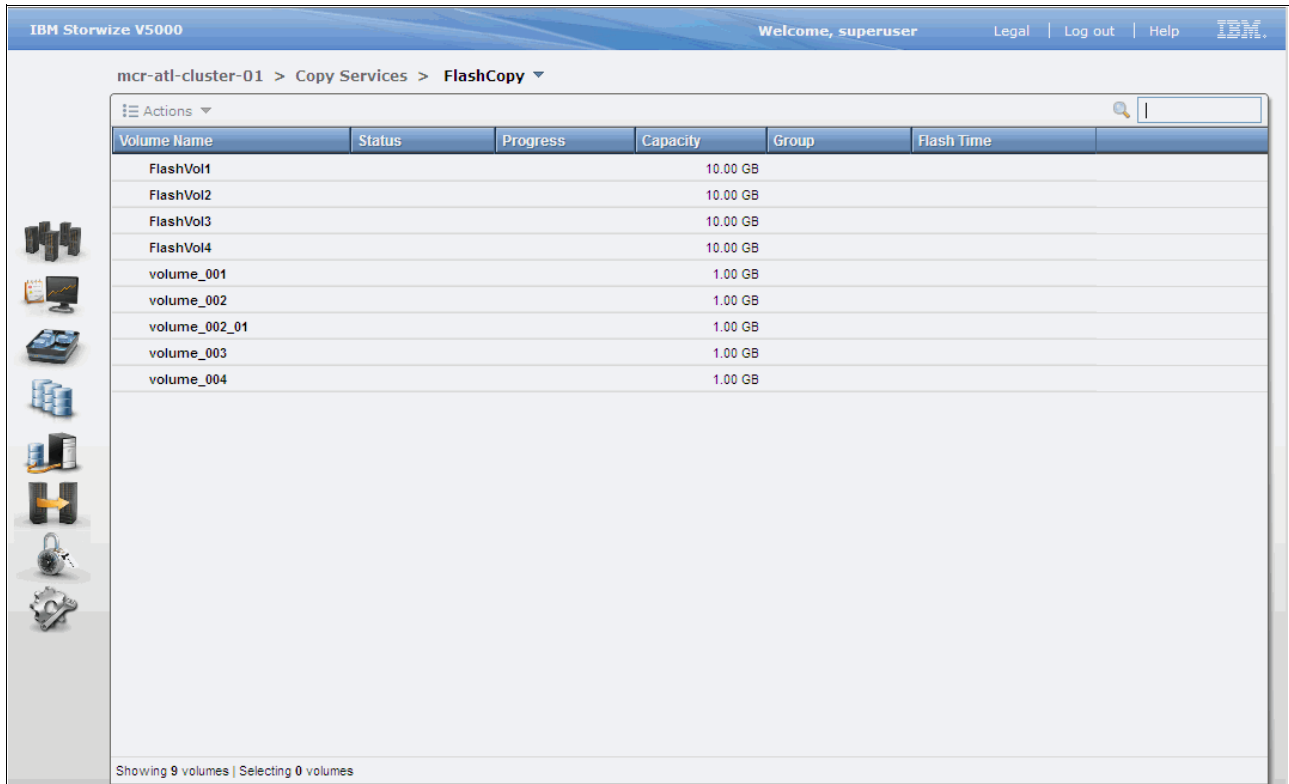


Figure 10-5 FlashCopy window

Click **FlashCopy Mappings** in the Copy Services function icon menu and the FlashCopy Mappings window opens, as shown in Figure 10-6. In the FlashCopy Mappings window, the FlashCopy mappings are listed individually.



Figure 10-6 FlashCopy Mappings window

The Consistency Groups window is used to manage the consistency groups for FlashCopy mappings. Click **Consistency Groups** in the Copy Services function icon menu and the Consistency Groups window opens, as shown in Figure 10-7.

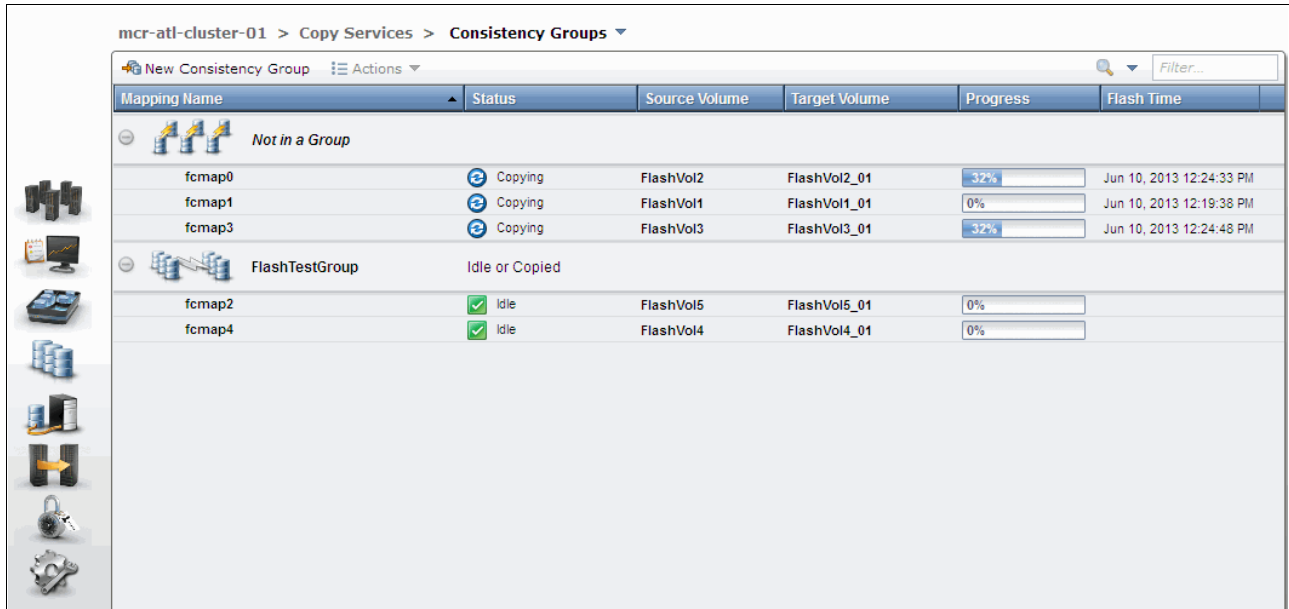


Figure 10-7 Consistency Groups window

Quick path to create FlashCopy presets

It is easy to create a FlashCopy by using the presets in the FlashCopy window.

Creating a snapshot

In the FlashCopy window, choose a volume and click **New Snapshot** from the Actions drop-down menu, as shown in Figure 10-8. Alternatively, you can highlight your chosen volume and right-click to access the Actions menu.

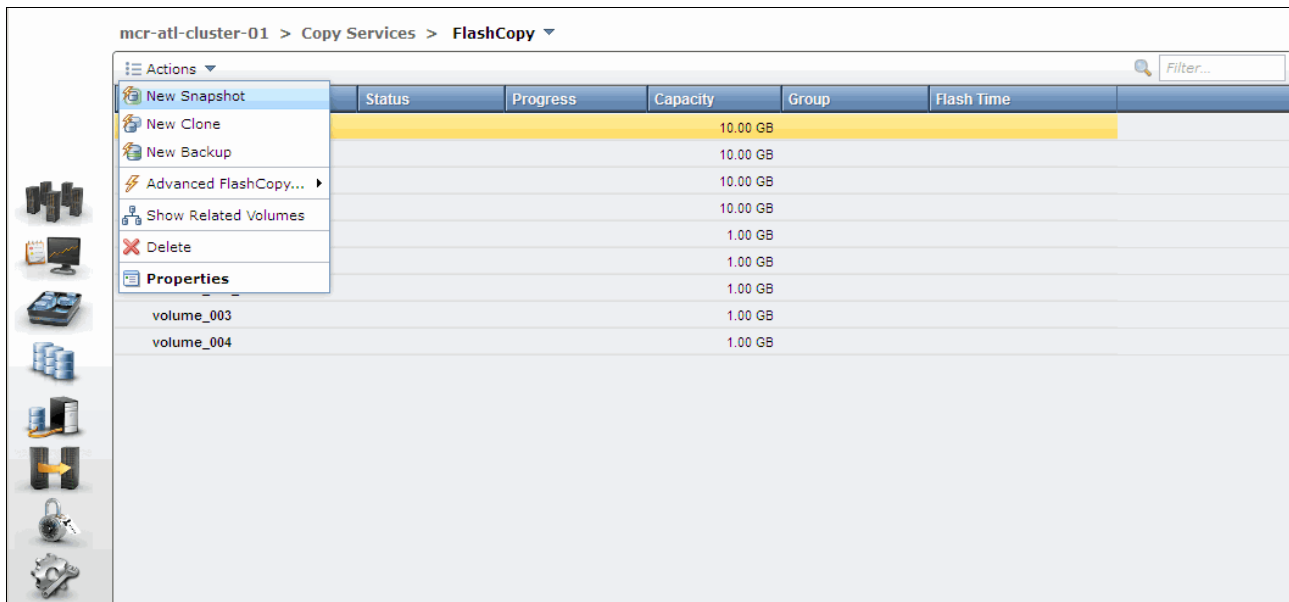


Figure 10-8 Create a snapshot by using the preset

You now have a snapshot volume for the volume you selected.

Creating a clone

In the FlashCopy window, choose a volume and click **New Clone** from the Actions drop-down menu, as shown in Figure 10-9. Alternatively, highlight your chosen volume and right-click to access the Actions menu.

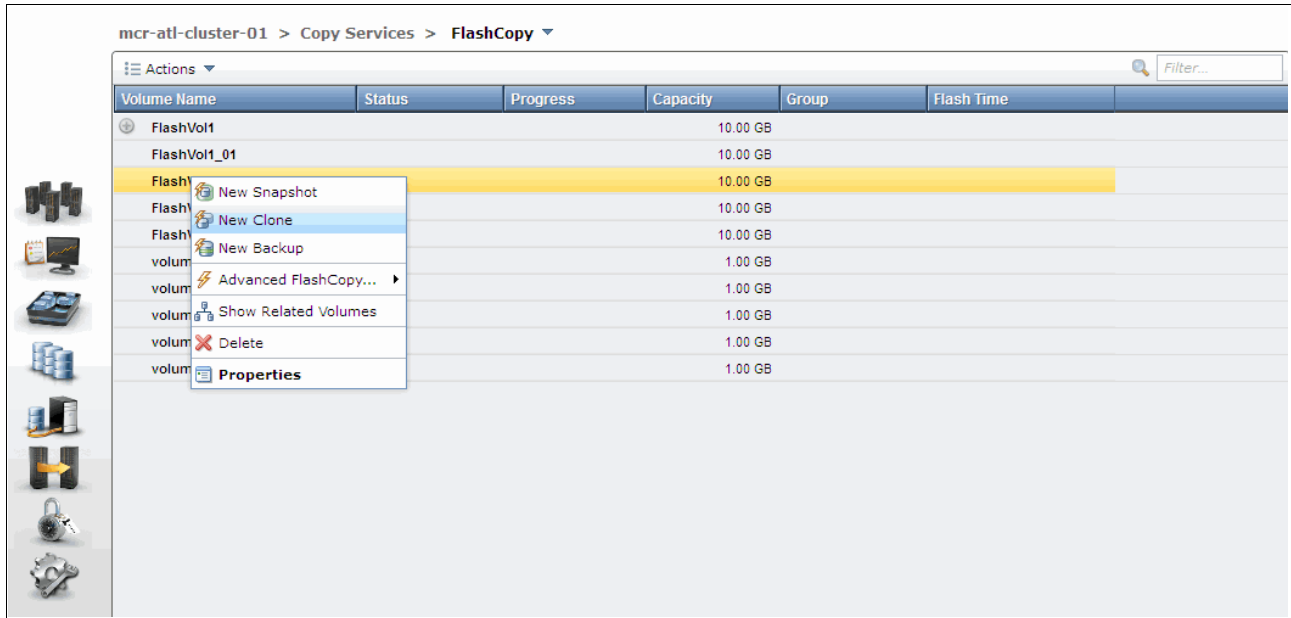


Figure 10-9 Create a clone from the preset

You now have a clone volume for the volume you selected.

Creating a backup

In the FlashCopy window, choose a volume and click **New Backup** from the Actions drop-down menu, as shown in Figure 10-10 on page 465. Alternatively, highlight your chosen volume and right-click to access the Actions menu.

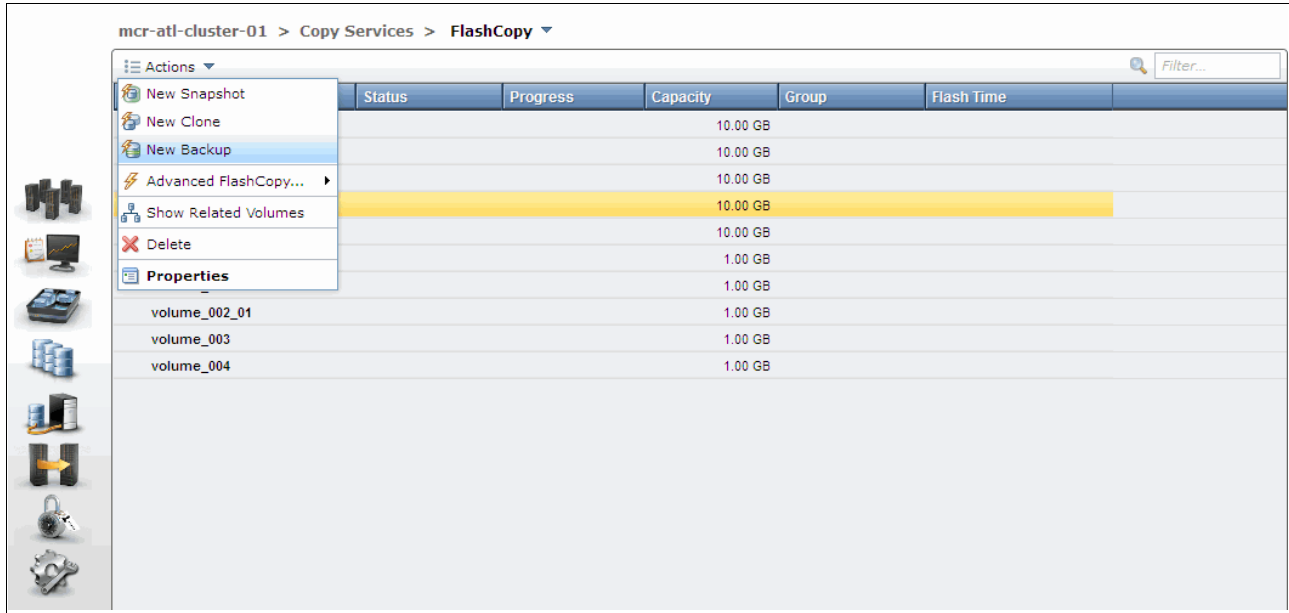


Figure 10-10 Create a backup from the preset

You now have a backup volume for the volume you selected.

In the FlashCopy window and in the FlashCopy Mappings window, you can monitor the progress of the running FlashCopy operations, as shown in Figure 10-11. The progress bars for each target volume indicate the copy progress in percentage. The copy progress remains 0% for snapshots (there is no change until data is written to the target volume). The copy progresses for clone and backup continues to increase until the copy process completes.

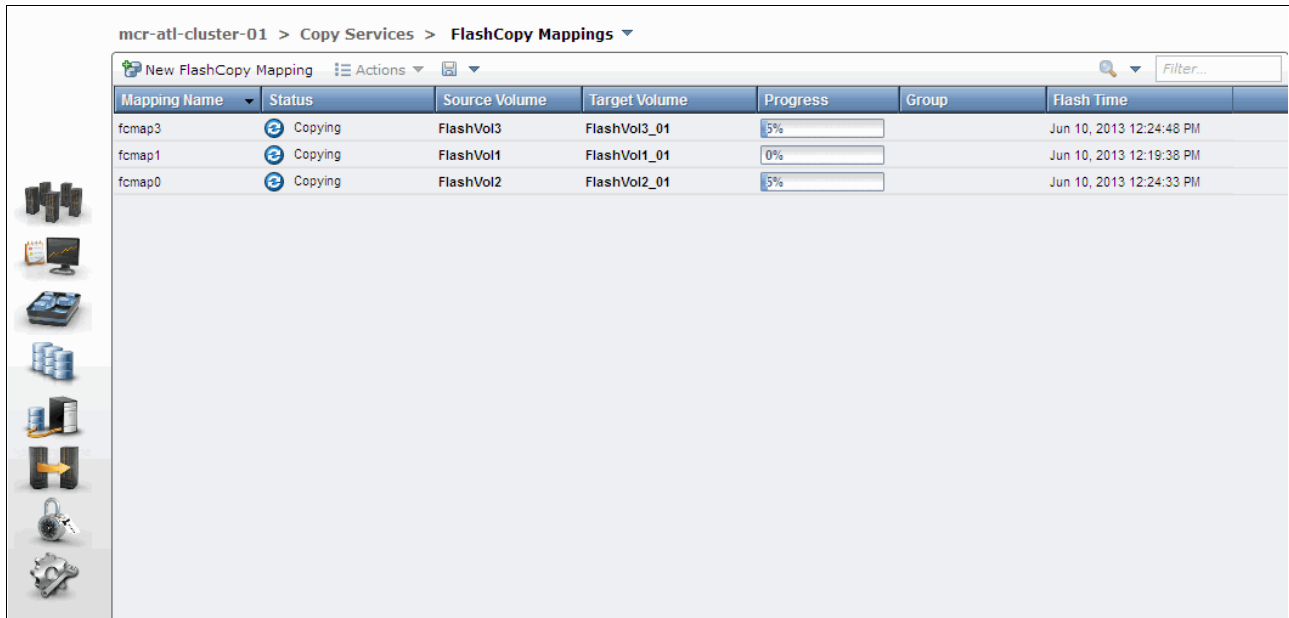


Figure 10-11 FlashCopy in progress that is viewed in the FlashCopy Mappings window

The copy progress can be also found under the Running Tasks status indicator, as shown in Figure 10-12.

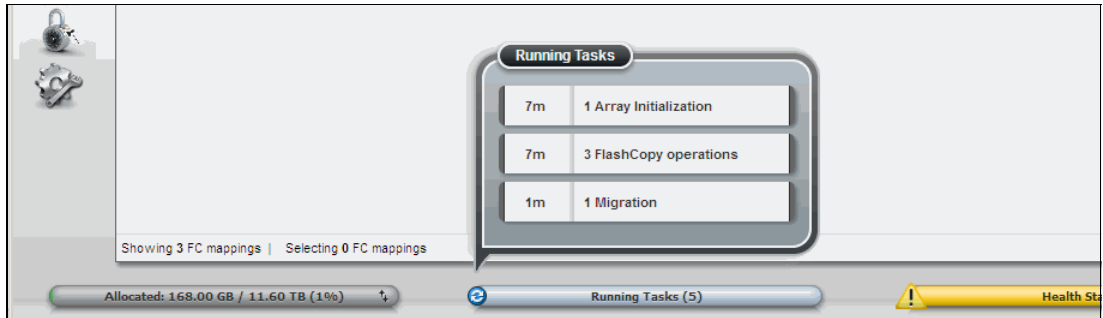


Figure 10-12 Running Tasks bar: FlashCopy operations

This view is slightly different from the FlashCopy and FlashCopy Mappings windows, as shown in Figure 10-13.

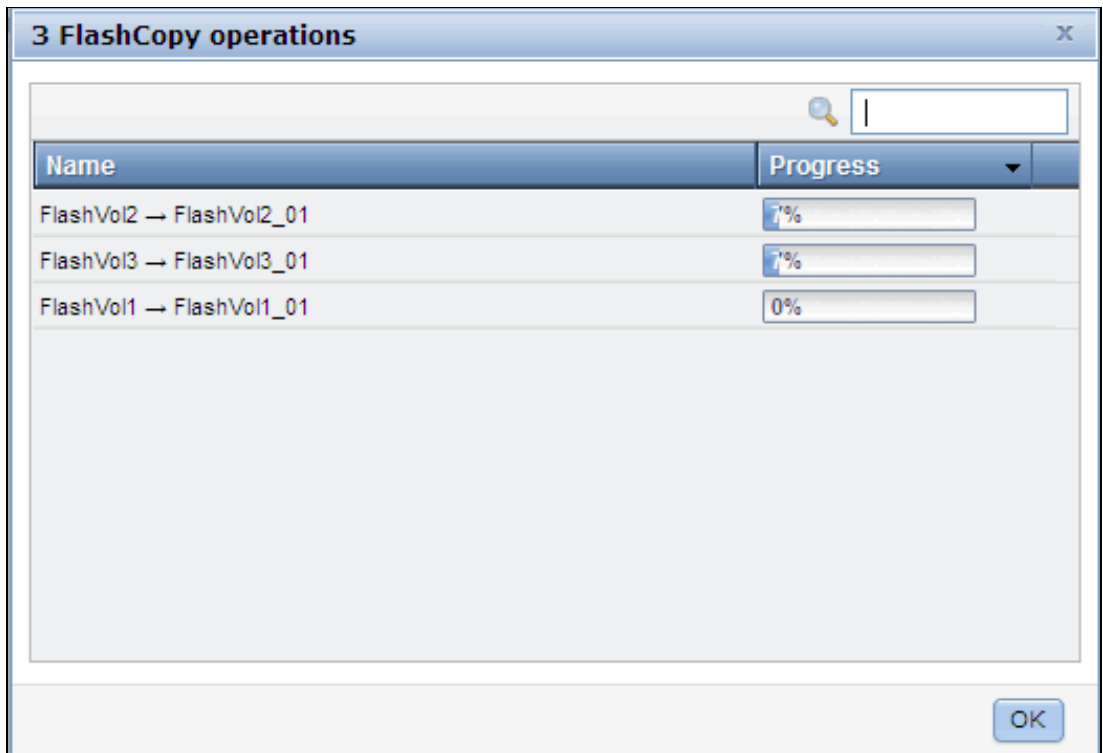


Figure 10-13 FlashCopy operations that are shown through Running Tasks

After the copy processes complete, you find the FlashCopy mapping with the clone preset (FlashVol12 in our example) was deleted automatically, as shown in Figure 10-14. There are now two identical volumes that are independent of each other.

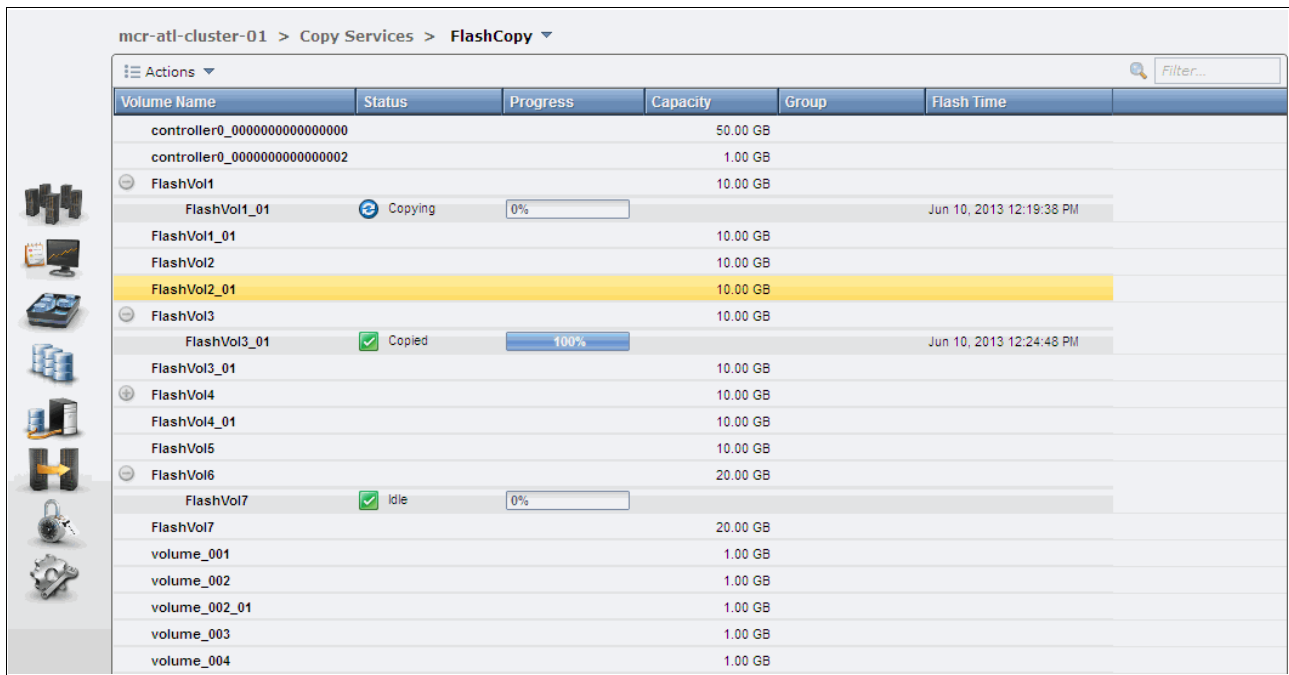


Figure 10-14 FlashCopy progresses complete

10.1.5 Managing FlashCopy mappings

The FlashCopy presets cover the most frequently used FlashCopy configurations for general situations. However, customized FlashCopy mappings are still necessary in some complicated scenarios.

Creating FlashCopy mappings

You can create FlashCopy mappings through the FlashCopy window. Select the volume that you want to be the source volume for the FlashCopy mapping and click **Advanced FlashCopy...** from the Actions drop-down menu, as shown in Figure 10-15 on page 468. Alternatively, select the volume and right-click.

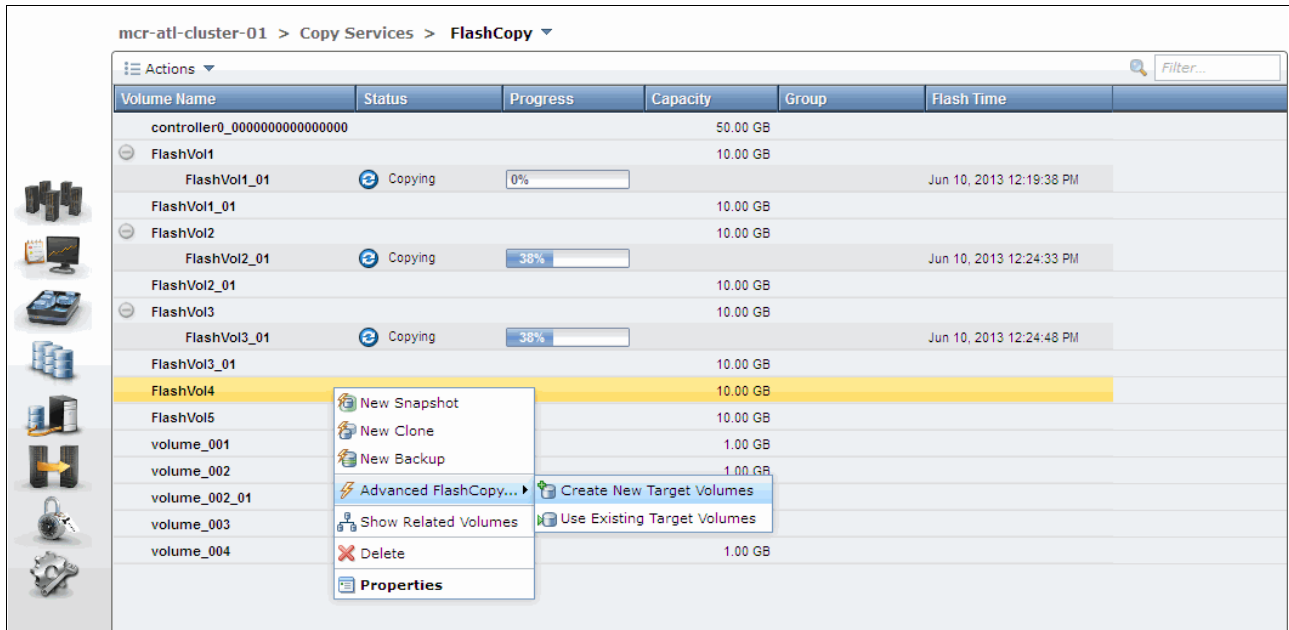


Figure 10-15 Create advanced FlashCopy

You can Create New Target Volumes as part of the mapping process or Use Existing Target Volumes. We describe creating volumes next. To use existing volumes, see “Using existing target volumes” on page 474.

Creating target volumes

Complete the following steps to create target volumes:

1. Click **Create new target volumes** if you have not yet created the target volume.
2. The wizard guides you to choose a preset, as shown in Figure 10-16. Choose one preset that has the most similar configuration to the one that is required and click **Advanced Settings** to make any appropriate adjustments to the advanced settings.

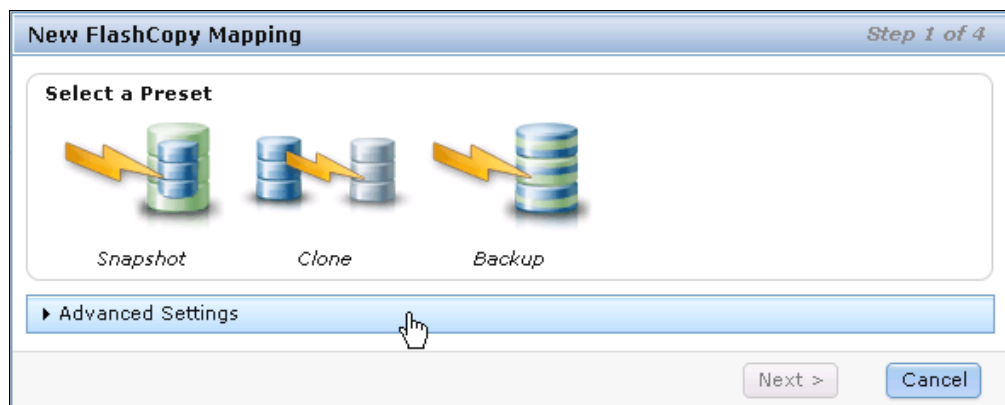


Figure 10-16 Choose a preset most similar to your requirement

The following default advanced settings for the snapshot preset are shown in Figure 10-17:

- Background Copy: 0
- Incremental: No
- Auto Delete after completion: No
- Cleaning Rate: 0

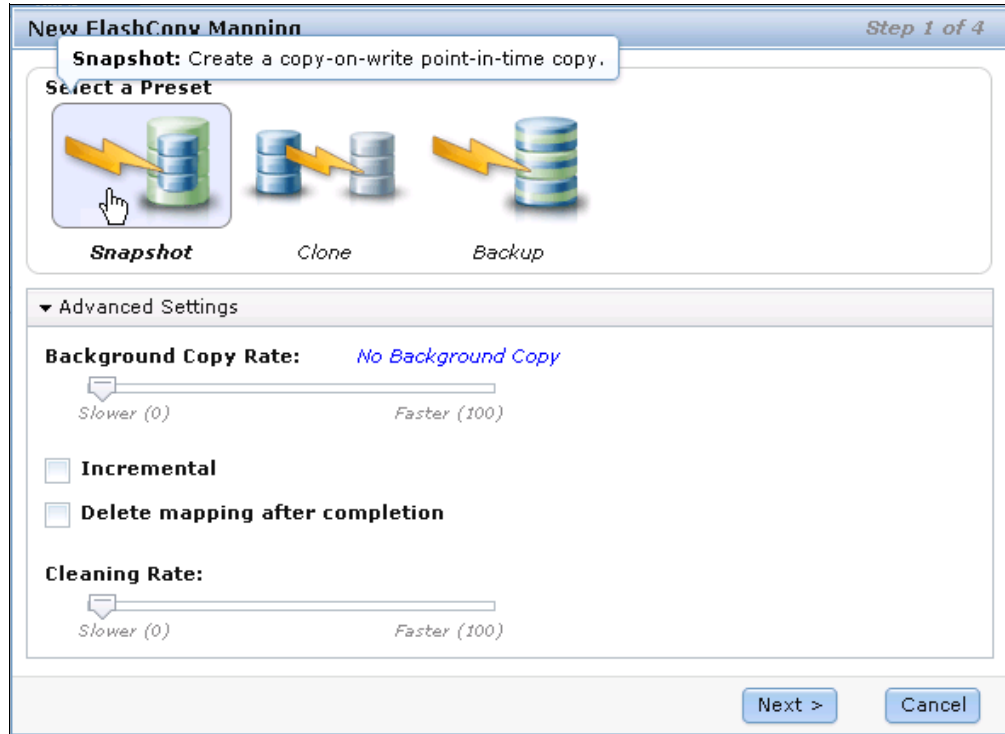


Figure 10-17 Default setting for the snapshot preset

The following Advanced Settings for the Clone Preset are shown in Figure 10-18:

- Background Copy: 50
- Incremental: No
- Auto Delete after completion: Yes
- Cleaning Rate: 50

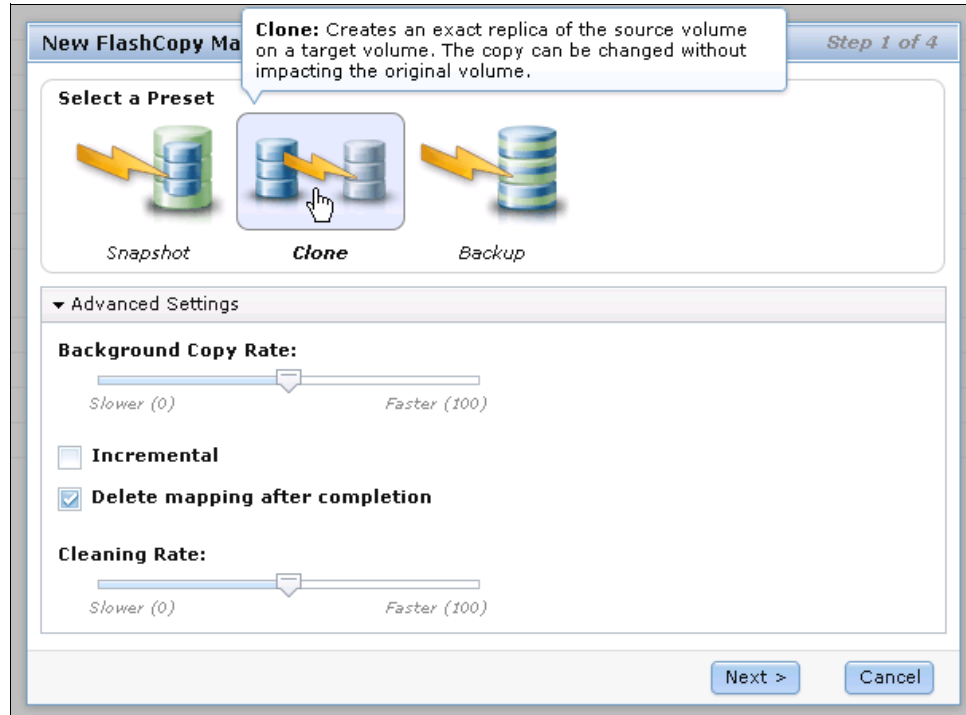


Figure 10-18 Default settings for the clone preset

Figure 10-19 shows the following Advanced Settings for the Backup preset:

- Background Copy: 50
- Incremental: Yes
- Auto Delete after completion: No
- Cleaning Rate: 50

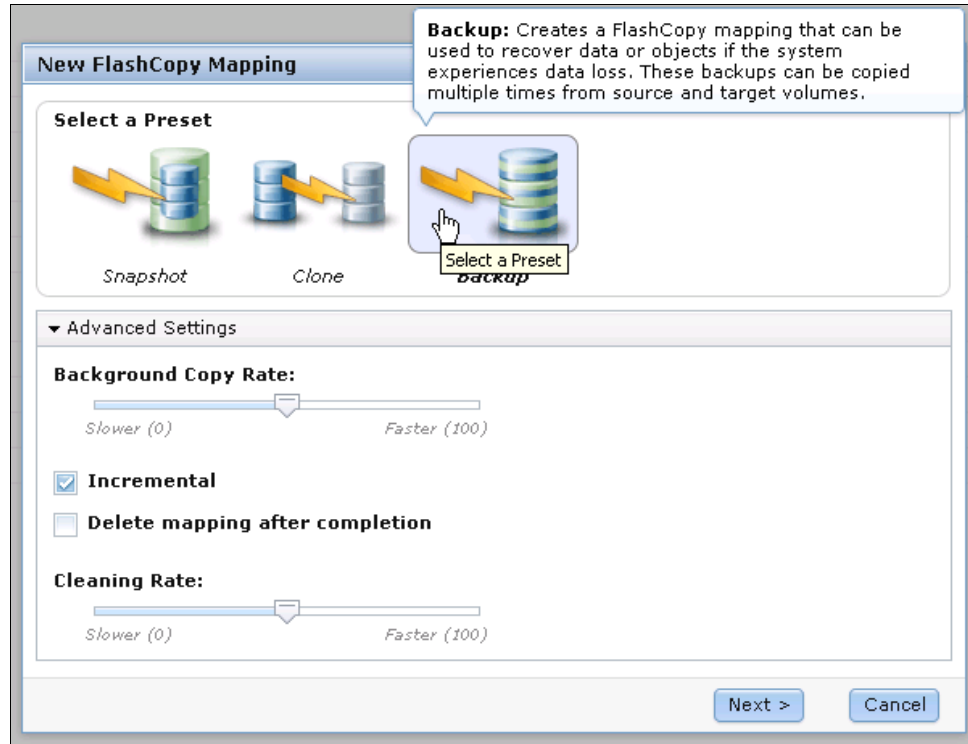


Figure 10-19 Default settings for the backup preset

3. Change the settings of the FlashCopy mapping according to your requirements and click **Next**.
4. In the next step, you can add your FlashCopy mapping to a consistency group, as shown in Figure 10-20. If the consistency group is not ready, the FlashCopy mapping can be added to the consistency group afterward. Click **Next** to continue.

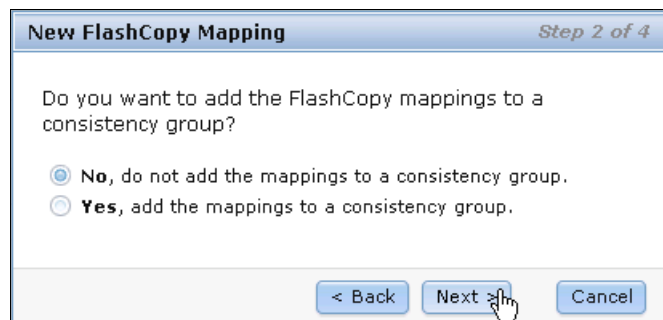


Figure 10-20 Add FlashCopy mapping to a consistency group

5. You can choose from which storage pool you want to create your target volume. As shown in Figure 10-21, you can select the same storage pool that is used by the source volume or a different pool. Click **Next** to continue.

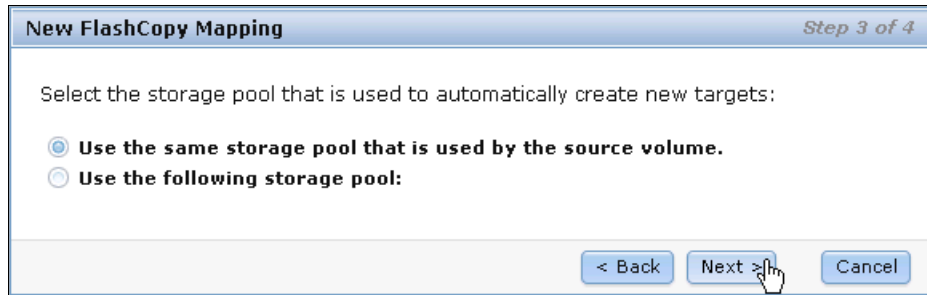


Figure 10-21 Using the same storage pool with the source volume

6. You can define how the new target volumes manage capacity. The Create a generic volume option is your default choice if you selected Clone or Backup as your basic preset. If you select a thin-provisioned volume, more options are available, as shown in Figure 10-22.

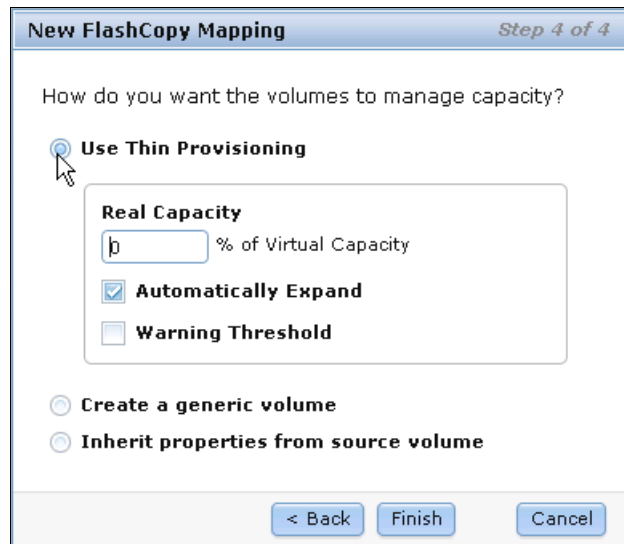


Figure 10-22 Creating a thin provisioned target volume

- Click **Finish** when you make your decision and the mappings and volume are created, as shown in Figure 10-23.

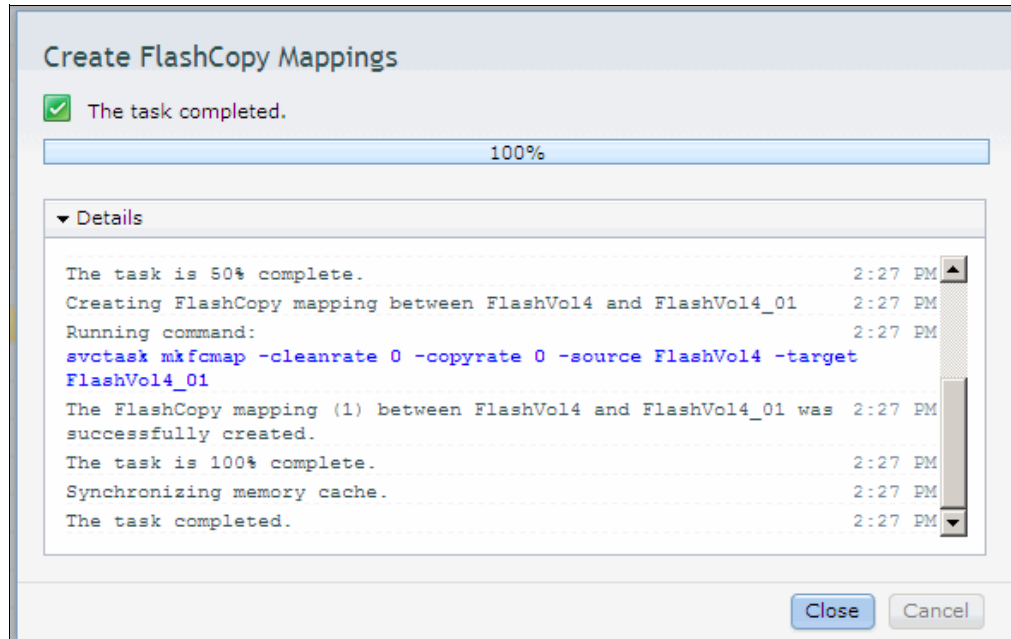


Figure 10-23 Advanced FlashCopy create task complete

- Close the window to see the FlashCopy mapping that is created on your volume with a new target, as shown in Figure 10-24. The status of the newly created FlashCopy mapping is Idle; it can be started, as described in “Starting a FlashCopy mapping” on page 477.

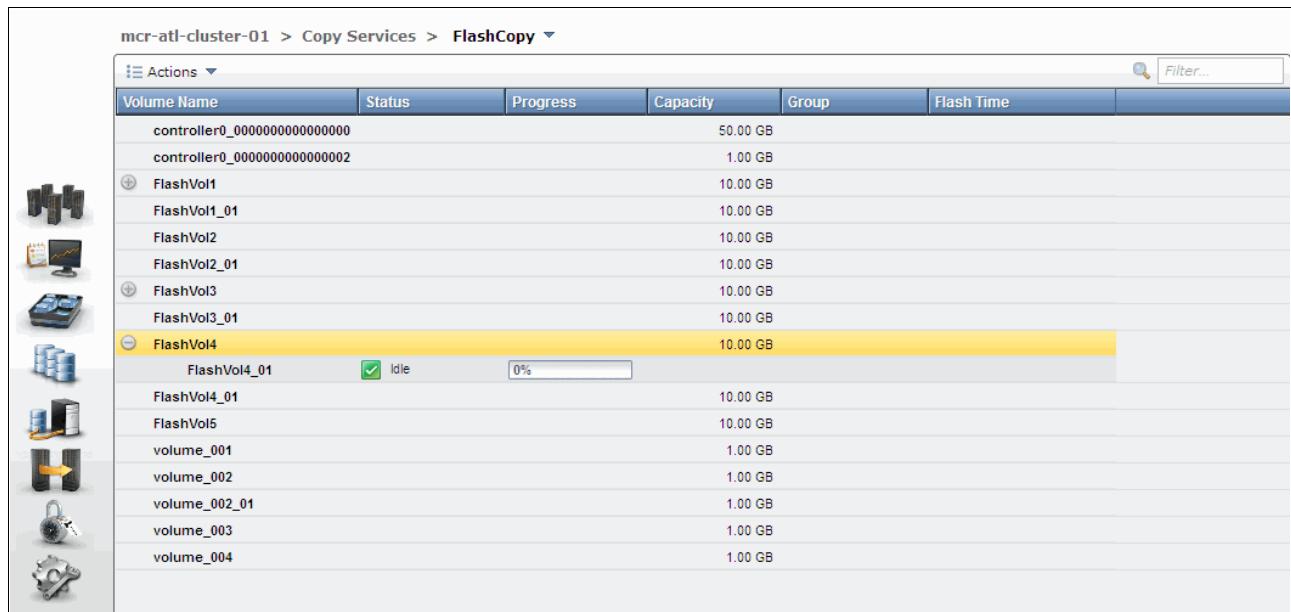


Figure 10-24 New FlashCopy mapping is created with a new target

Using existing target volumes

Complete the following steps to use existing target volumes:

1. If you already have candidate target volumes, select **Use Existing Target Volumes** in the Advanced FlashCopy menu, as shown in Figure 10-25.

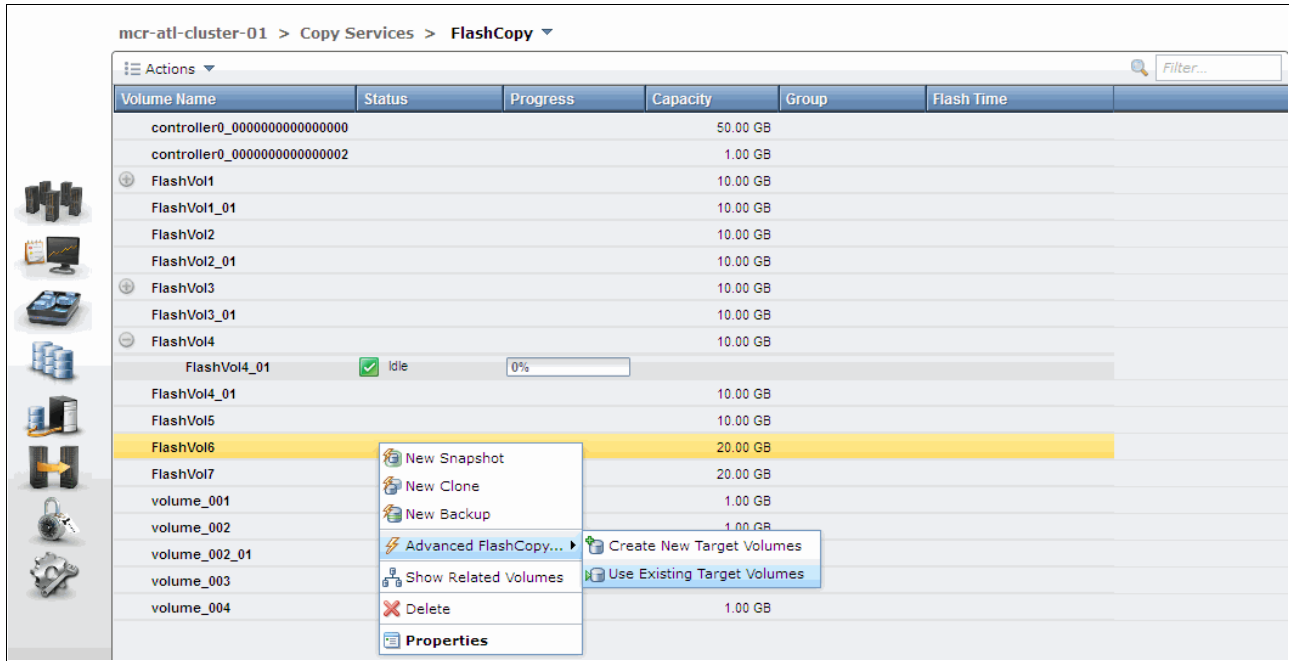


Figure 10-25 Create FlashCopy mapping by using existing target volume

2. You must choose the target volume for the source volume that you selected. Select the target volume from the drop-down menu in the right pane of the window and click **Add**, as shown in Figure 10-26.

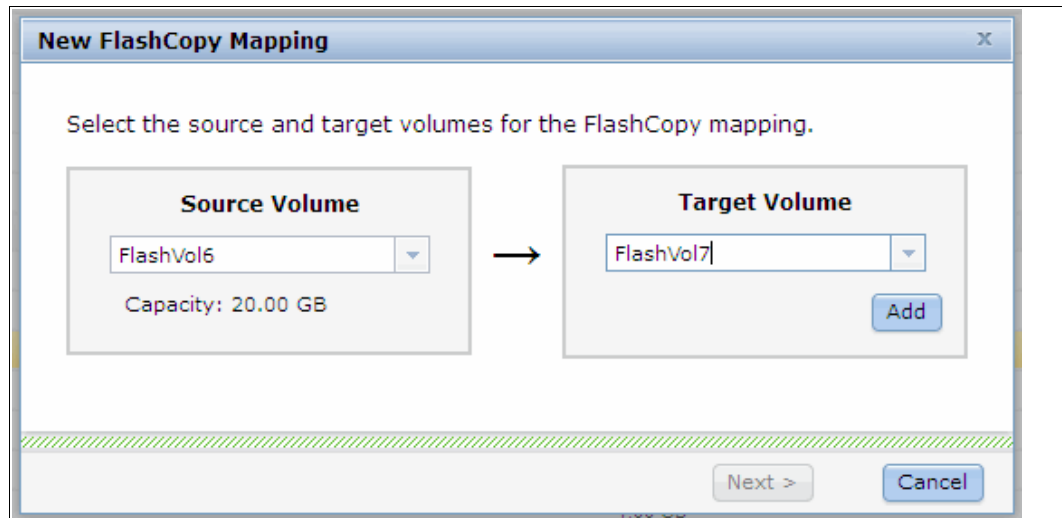


Figure 10-26 Select the target volume

3. After you click **Add**, the FlashCopy mapping is listed, as shown in Figure 10-27 on page 475. Click the red X if the FlashCopy mapping is not the one you want to create. If the FlashCopy mapping is what you want, click **Next** to continue.

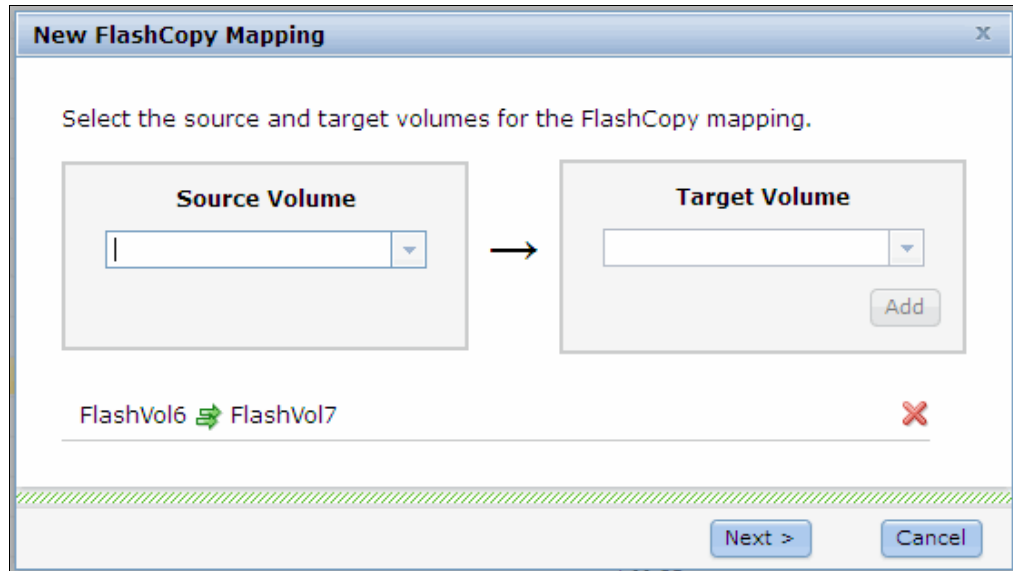


Figure 10-27 Add FlashCopy mapping

4. Select the preset and (if necessary) adjust the settings by using the Advanced Settings section as shown in Figure 10-28. (For more information about the advanced setting, see “Creating target volumes” on page 468.) Confirm that the settings meet your requirements and then click **Next**.

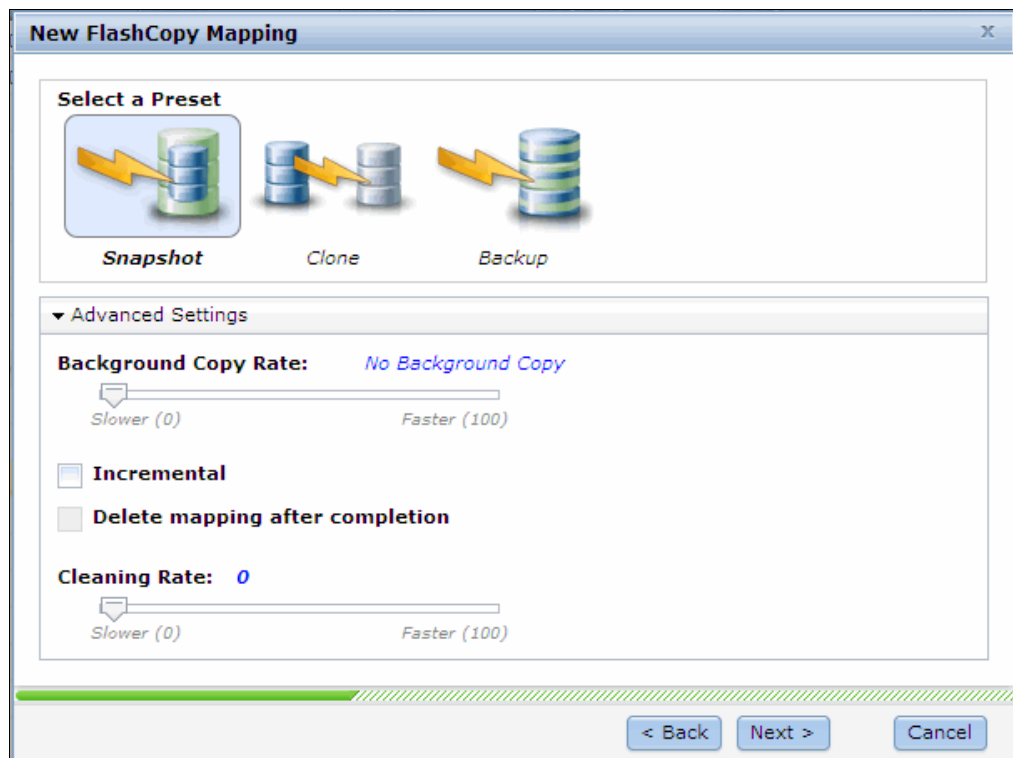


Figure 10-28 Select a preset and make your adjustments

- You can now add the FlashCopy mapping to a consistency group (if necessary), as shown in Figure 10-29. Selecting **Yes** shows a drop-down menu from which you can select a consistency group. Click **Finish** and the FlashCopy mapping is created with the status of Idle, as shown in Figure 10-24 on page 473.

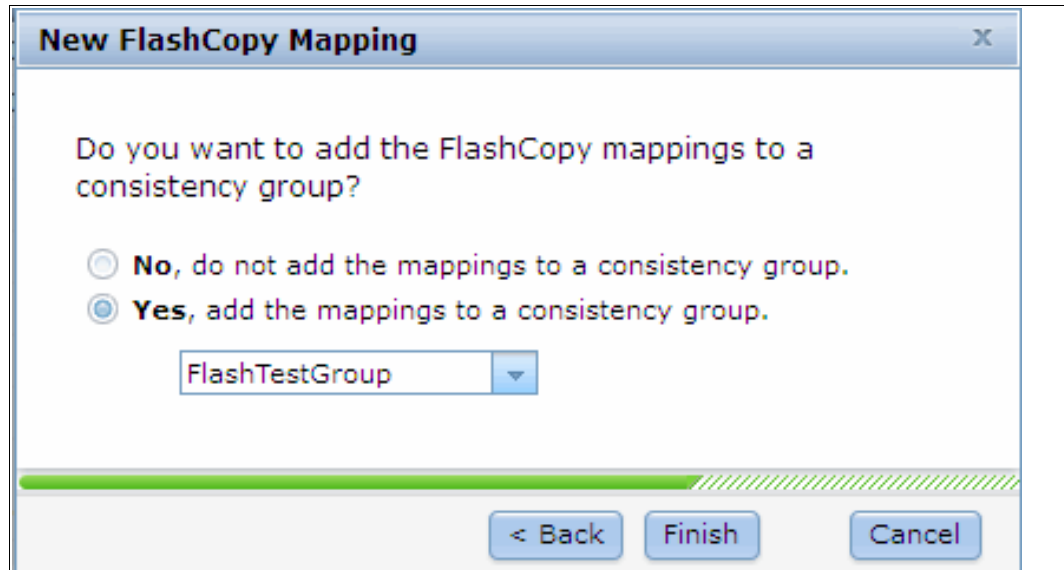


Figure 10-29 Select a consistency group to add the FlashCopy mapping

Creating new FlashCopy mappings

You can also create FlashCopy mappings in the FlashCopy Mappings window by clicking **New FlashCopy Mapping** at the upper left, as shown in Figure 10-30.

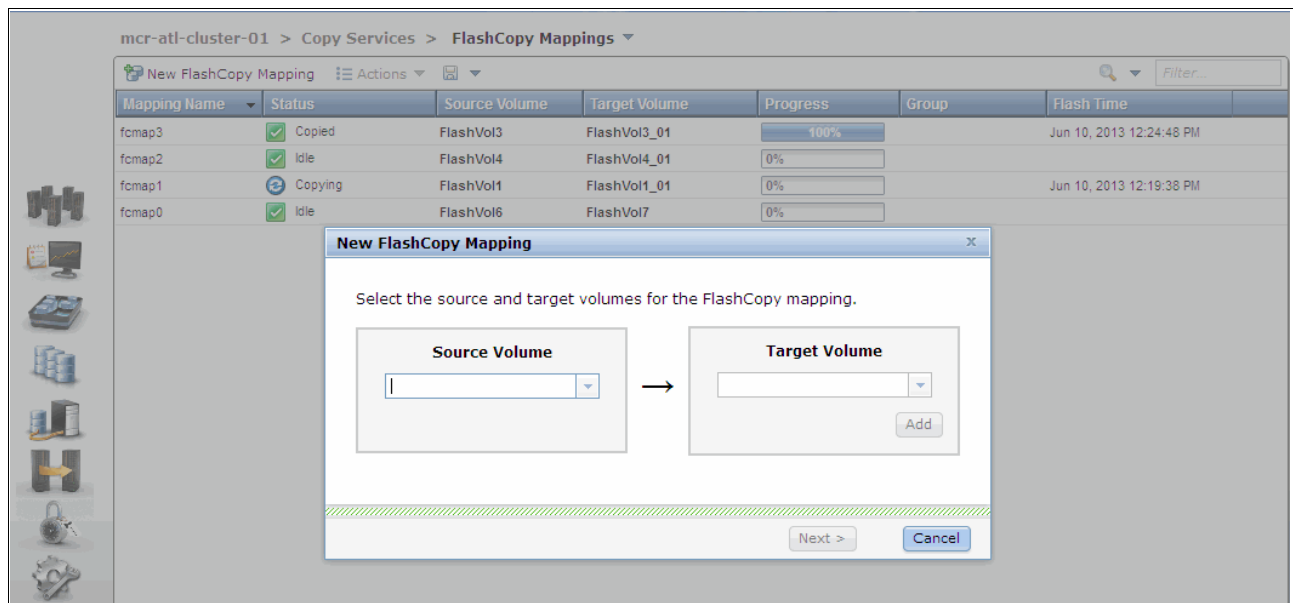


Figure 10-30 Create a FlashCopy mapping in the FlashCopy Mappings window

A wizard guides you through the process to create a FlashCopy mapping. The steps are the same as creating an Advanced FlashCopy mapping by using Existing Target Volumes, as described in “Using existing target volumes” on page 474.

Starting a FlashCopy mapping

Most of the FlashCopy mapping actions can be performed in the FlashCopy window or the FlashCopy Mapping window. For the actions that are available in both windows, we show in the following sections the steps in the FlashCopy window, although the steps are the same if you were to use the FlashCopy Mapping window.

You can start the mapping by selecting the FlashCopy target volume in the FlashCopy window and selecting the **Start** option from the Actions drop-down menu (as shown in Figure 10-31) or by selecting the volume and right-clicking. The status of the FlashCopy mapping changes from Idle to Copying.

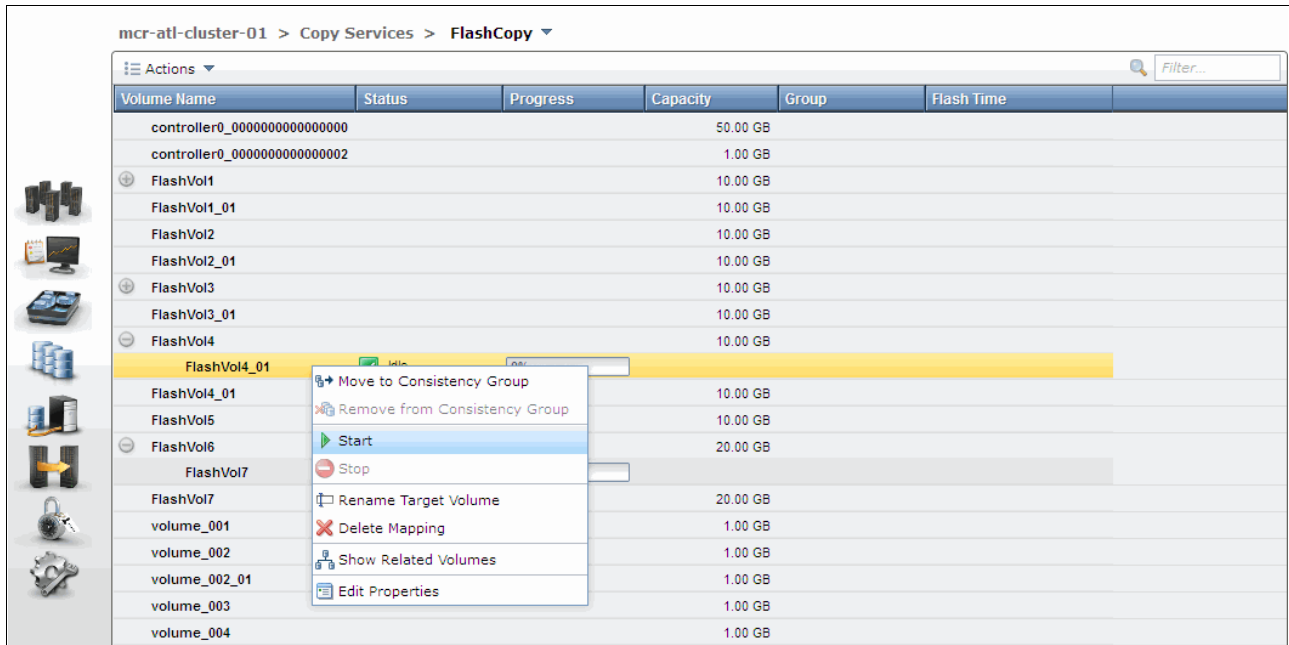


Figure 10-31 Start FlashCopy mapping

Stopping a FlashCopy mapping

The FlashCopy mapping can be stopped by selecting the FlashCopy target volume in the FlashCopy window and clicking the **Stop** option from the Actions drop-down menu, as shown in Figure 10-32 on page 478. After the stopping process completes, the status of the FlashCopy mapping is changed to Stopped.

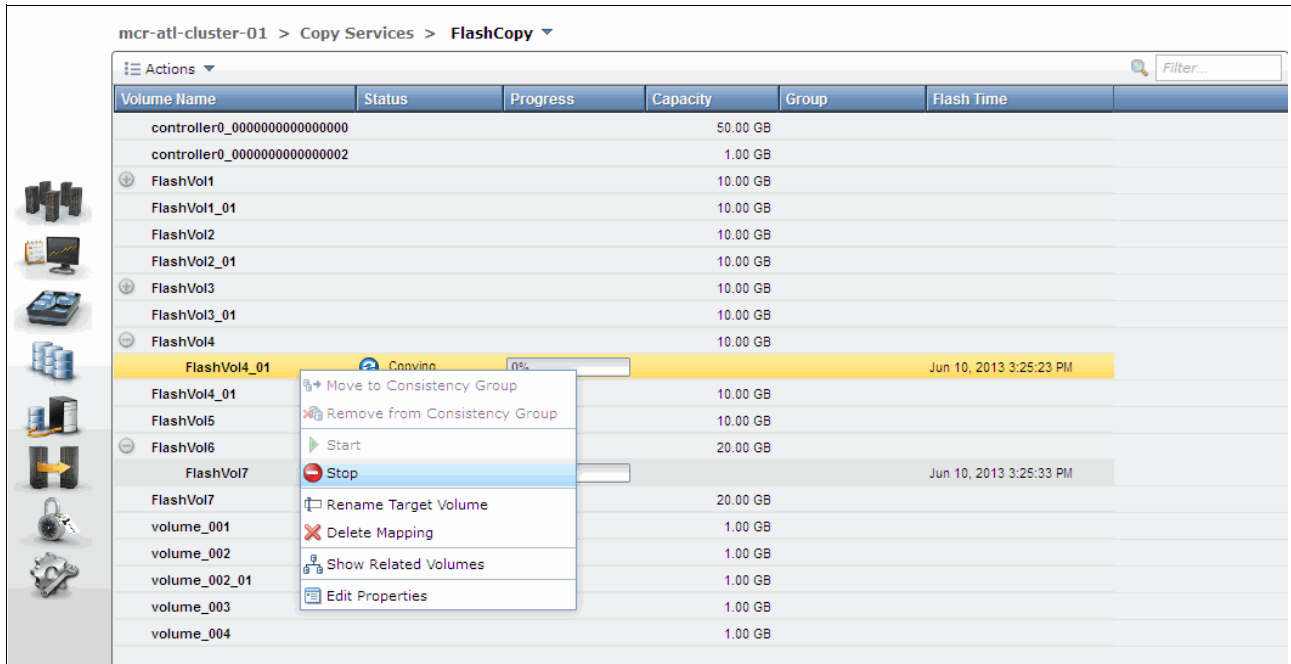


Figure 10-32 Stopping a FlashCopy mapping

Renaming the target volume

If the FlashCopy target volumes were created automatically by IBM Storwize V5000 storage system, the name of the target volume is the source volume name plus a suffix that includes numbers. The name of the target volumes can be changed to be more meaningful in your environment.

To change the name of the target volume, select the FlashCopy target volume in the FlashCopy window and click the **Rename Target Volume** option from the Actions drop-down menu (as shown in Figure 10-33) or right-click the selected volume.

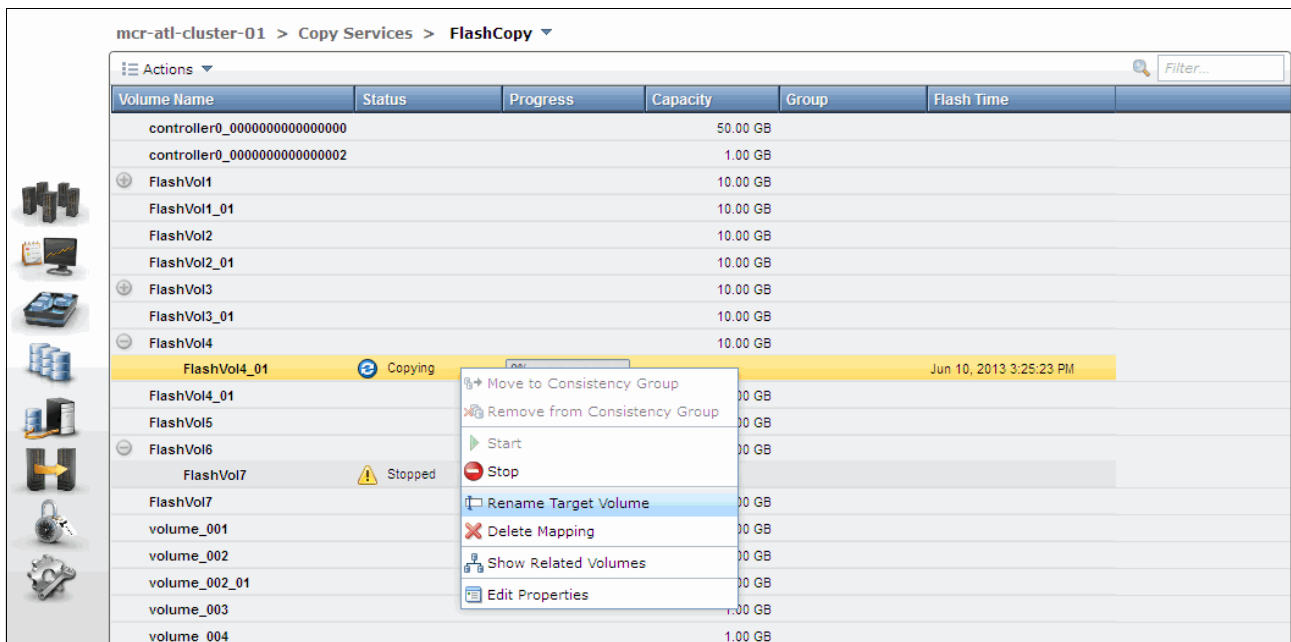


Figure 10-33 Rename a target volume

Enter your new name for the target volume, as shown in Figure 10-34. Click **Rename** to finish.



Figure 10-34 Rename a target volume

Renaming a FlashCopy mapping

The FlashCopy mappings are created with names that begin with fcmmap. The name of FlashCopy mappings can be changed to be more meaningful to you.

To change the name of a FlashCopy mapping, select the FlashCopy mapping in the FlashCopy Mappings window and click **Rename Mapping** in the Actions drop-down menu, as shown in Figure 10-35.

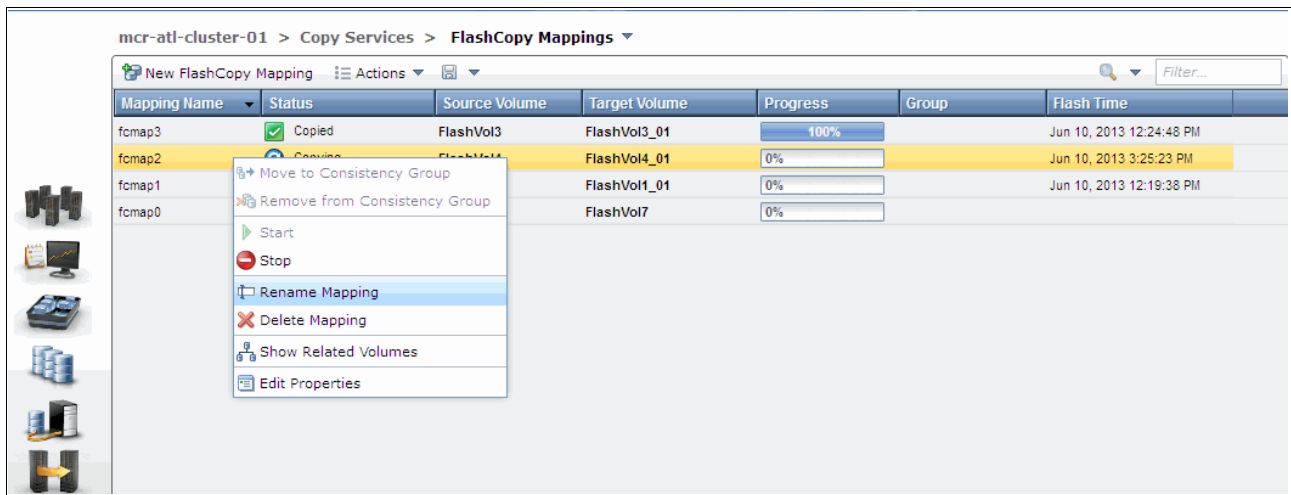


Figure 10-35 Rename a FlashCopy mapping

You must enter the new name for the FlashCopy mapping, as shown in Figure 10-36. Click **Rename** to finish.

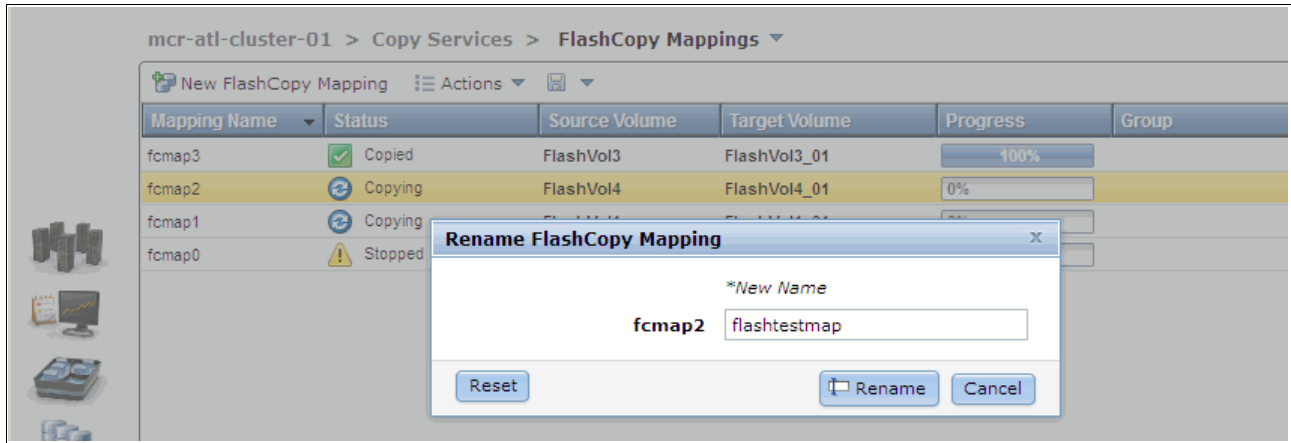


Figure 10-36 Enter a new name for the FlashCopy mapping

Deleting a FlashCopy mapping

The FlashCopy mapping can be deleted by selecting the FlashCopy target volume in the FlashCopy window and clicking **Delete Mapping** in the Actions drop-down menu (as shown in Figure 10-37) or by right-clicking the selected volume.

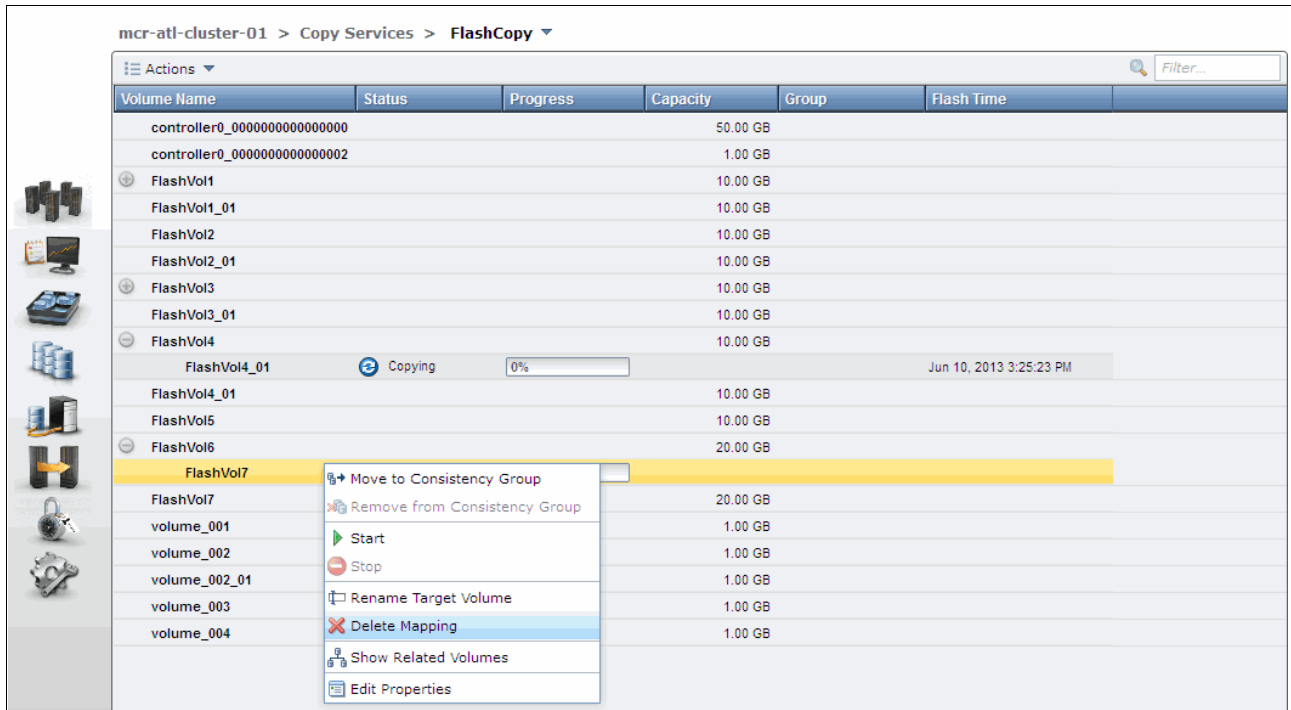


Figure 10-37 Select Delete Mapping

FlashCopy Mapping state: If the FlashCopy mapping is in the Copying state, it must be stopped before it is deleted.

You must confirm your action to delete FlashCopy mappings in the window that opens, as shown in Figure 10-38. Verify the number of FlashCopy mappings that you must delete. If you want to delete the FlashCopy mappings while the data on the target volume is inconsistent with the source volume, select the option to do so. Click **Delete** and your FlashCopy mapping is removed.

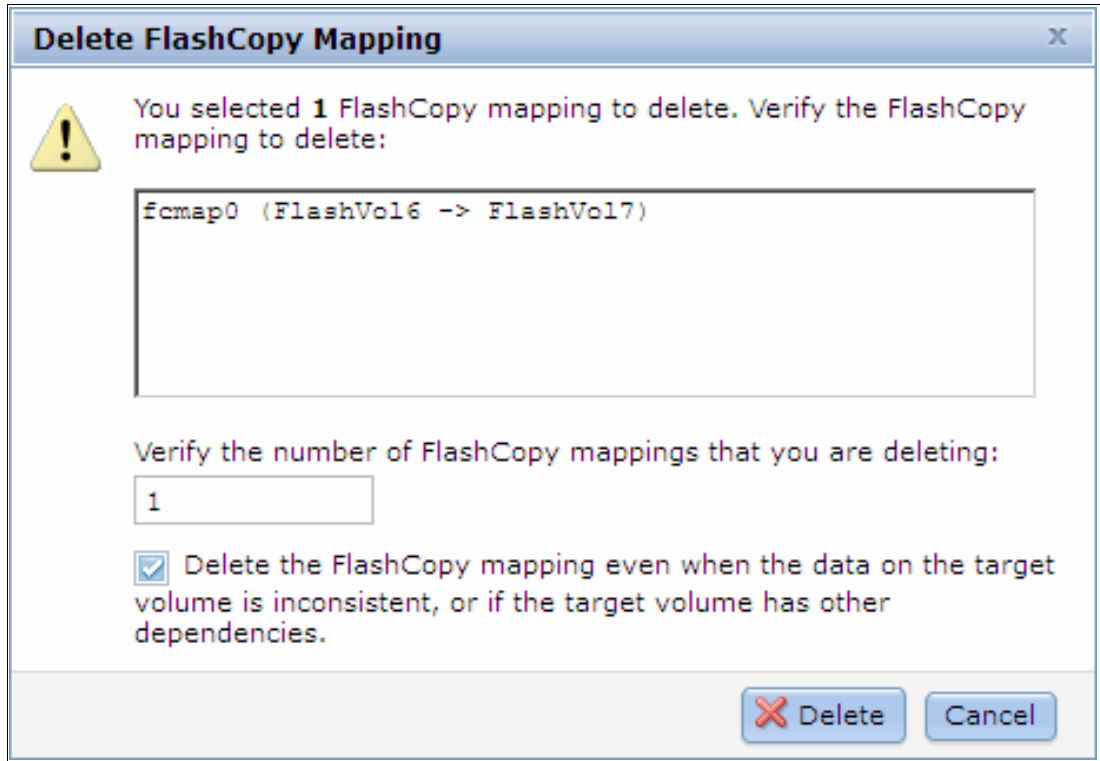


Figure 10-38 Confirm the deletion of FlashCopy mappings

Deleting FlashCopy mapping: Deleting the FlashCopy mapping does not delete the target volume. If you must reclaim the storage space that is occupied by the target volume, you must delete the target volume manually.

Showing related volumes

You can show the FlashCopy mapping dependencies by selecting a target or source volume in the FlashCopy window and clicking **Show Related Volumes** in the Actions drop-down menu (as shown in Figure 10-39) or right-clicking the selected volume.

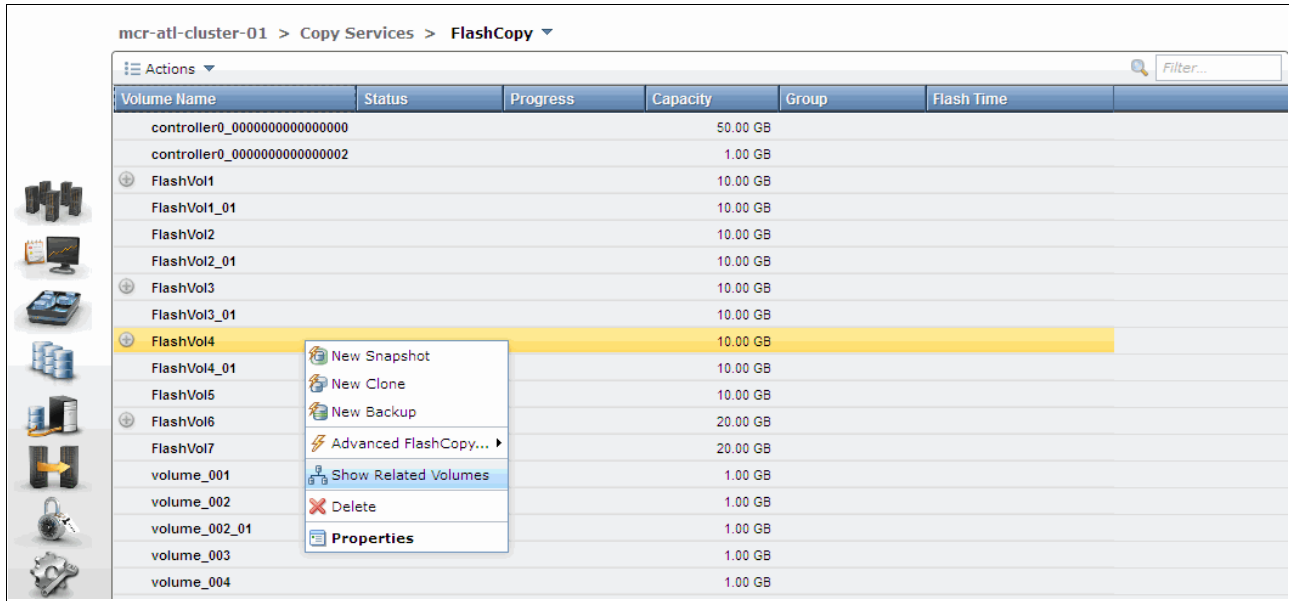


Figure 10-39 Show Related Volumes menu

The FlashCopy mapping dependency tree opens, as shown in Figure 10-40.

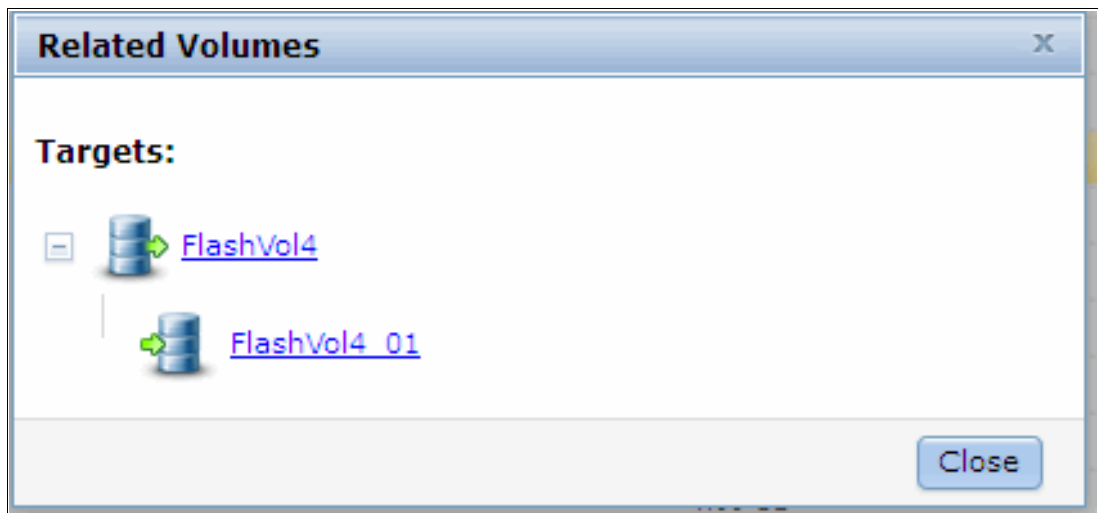


Figure 10-40 FlashCopy mapping dependency

Clicking either volume shows the properties of the volume, as shown in Figure 10-41.

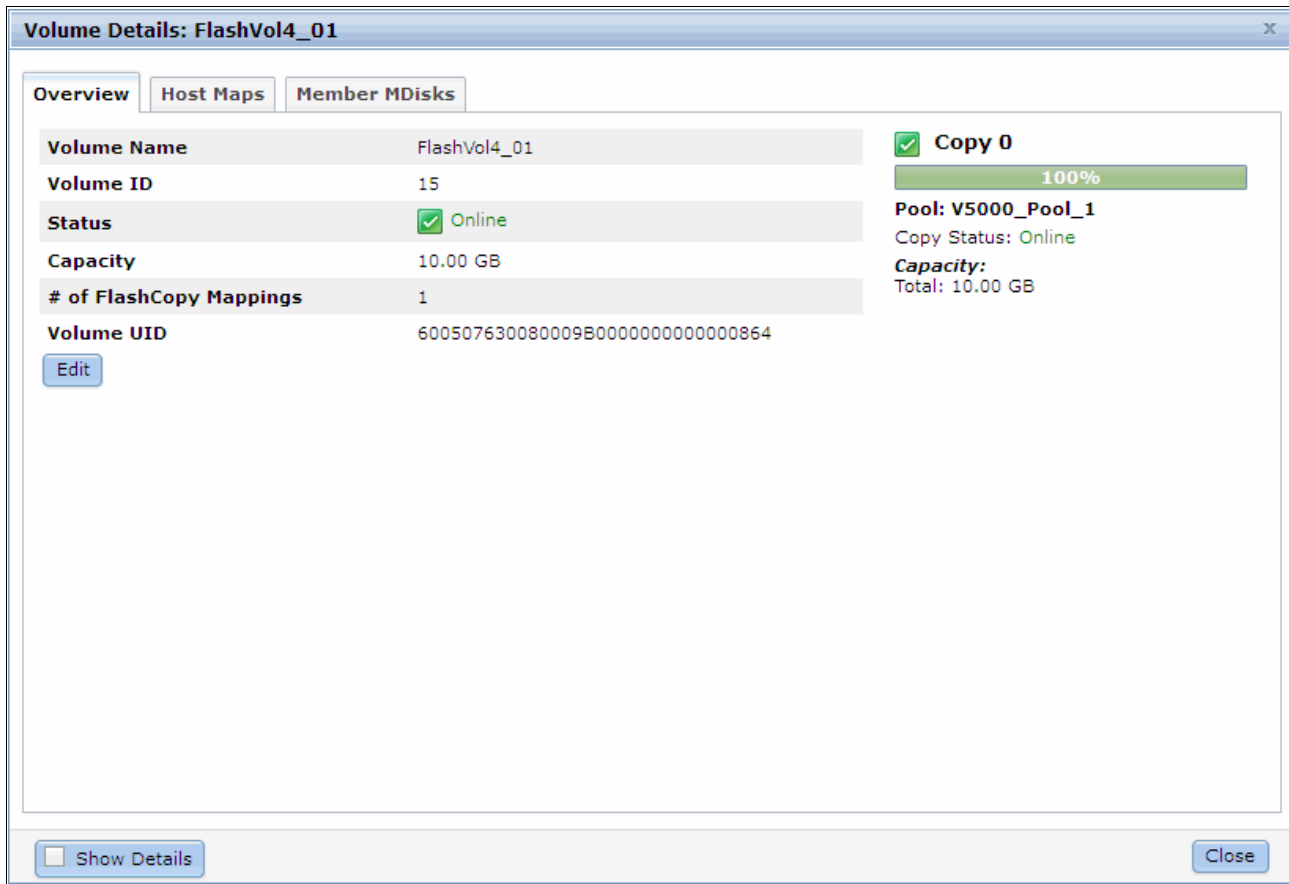


Figure 10-41 Target FlashCopy Volume details

Editing properties

The background copy rate and cleaning rate can be changed after the FlashCopy mapping is created. Select the FlashCopy target mapping in the FlashCopy window and click **Edit Properties** in the Actions drop-down menu (as shown in Figure 10-42) or right-click.

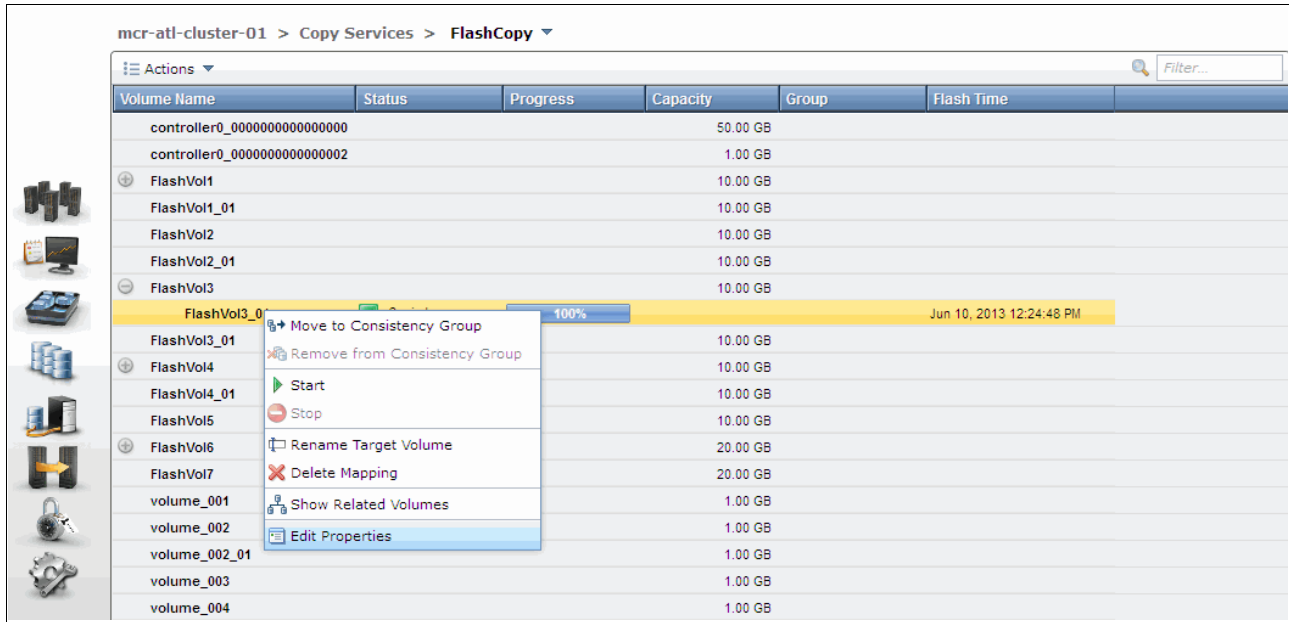


Figure 10-42 Edit Properties menu

You can then modify the value of the background copy rate and cleaning rate by moving the pointers on the bars, as shown in Figure 10-43. Click **Save** to save changes.

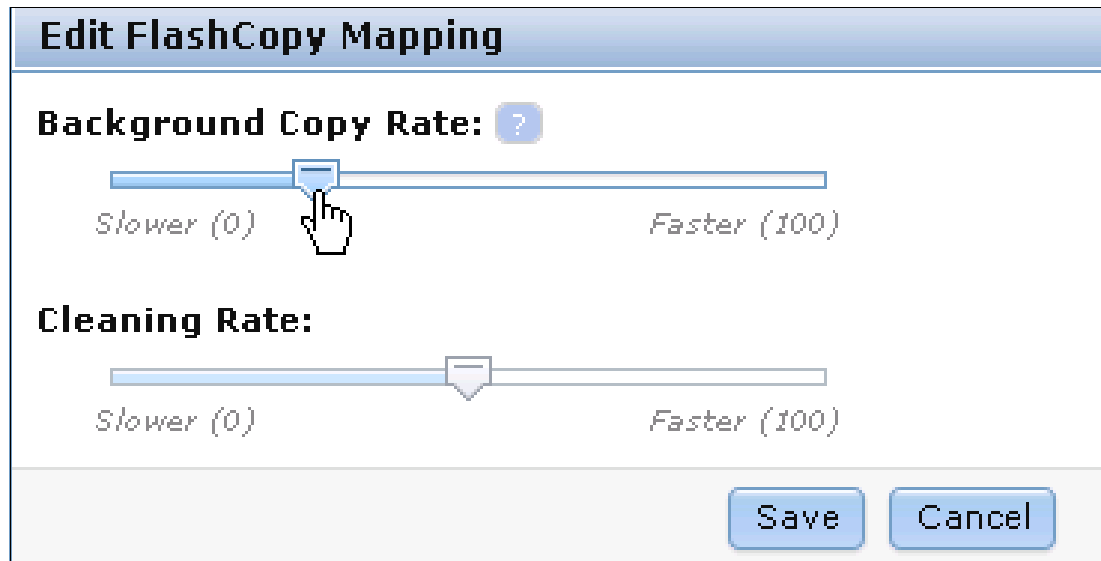


Figure 10-43 Change the copy rate

Restoring from a FlashCopy

Complete the following steps to manipulate FlashCopy target volumes to restore a source volume to a previous known state:

1. Identify the FlashCopy relationship that you want to restore. In our example, we want to restore FlashVol1, as shown in Figure 10-44.

Volume Name	Status	Progress	Capacity	Group	Flash Time
controller0_0000000000000002			1.00 GB		
DS3400_0000000000000002			1.00 GB		
FlashVol1			10.00 GB		
FlashVol1_01	Copying	0%		FlashTestGroup	Jun 12, 2013 1:19:18 PM
FlashVol1_01			10.00 GB		
FlashVol2			10.00 GB		
FlashVol2_01			10.00 GB		
FlashVol2_02			10.00 GB		
FlashVol3			10.00 GB		
FlashVol3_01			10.00 GB		
FlashVol4			10.00 GB		
FlashVol4_01	Copying	90%			Jun 12, 2013 10:55:34 AM
FlashVol4_01			10.00 GB		
FlashVol5			10.00 GB		
FlashVol5_01			10.00 GB		
FlashVol6			20.00 GB		
FlashVol7	Stopped	0%			
FlashVol7			20.00 GB		
FlashVol6	Copied	100%			Jun 12, 2013 1:23:28 PM
Source_FC			1.00 GB		
Target_FC	Copied	100%			Jun 12, 2013 12:33:32 PM
Source_FC_01	Copied	100%			Jun 12, 2013 12:49:22 PM
Source_FC_02	Idle	0%			

Showing 23 volumes | Selecting 0 volumes

Figure 10-44 Starting FlashCopy restore

2. Create a mapping by using the target volume of the mapping to be restored. In our example, it is FlashVol1_01, as shown in Figure 10-45. Select **Advanced FlashCopy** → **Use Existing Target Volumes**.

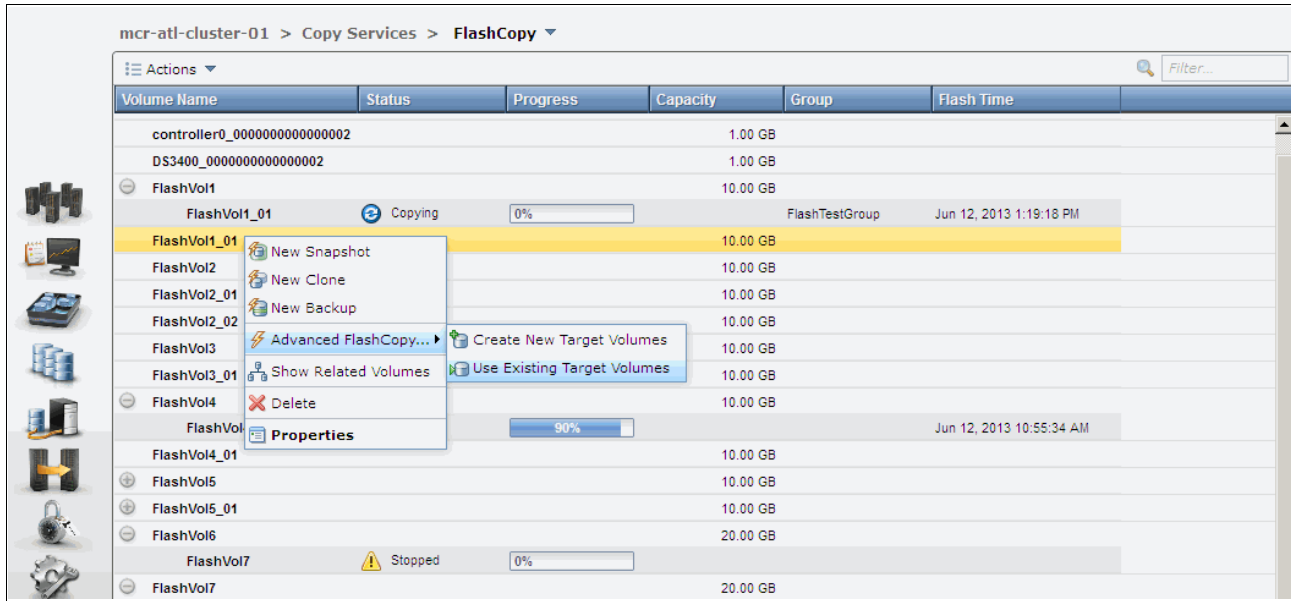


Figure 10-45 Create reverse mapping

3. The Source Volume is preselected with the target volume that was selected in the previous step. Select the Target Volume from the drop-down menu (you select the source volume that you want to restore). In our example, we select FlashVol1, as shown in Figure 10-46.

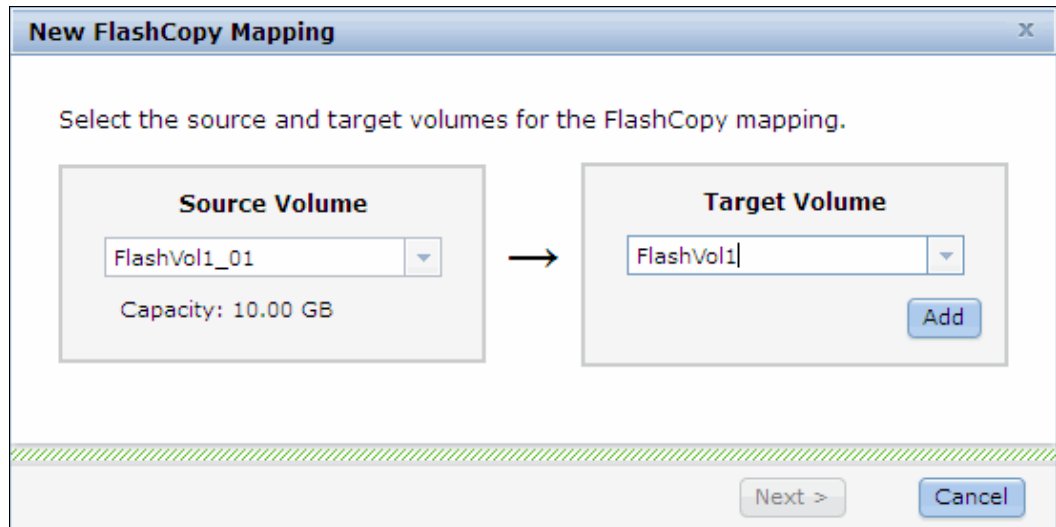


Figure 10-46 Select target volume

- Click **Add**. A warning message appears, as shown in Figure 10-47. Click **Close**. This message is shown because we are using a source as a target.

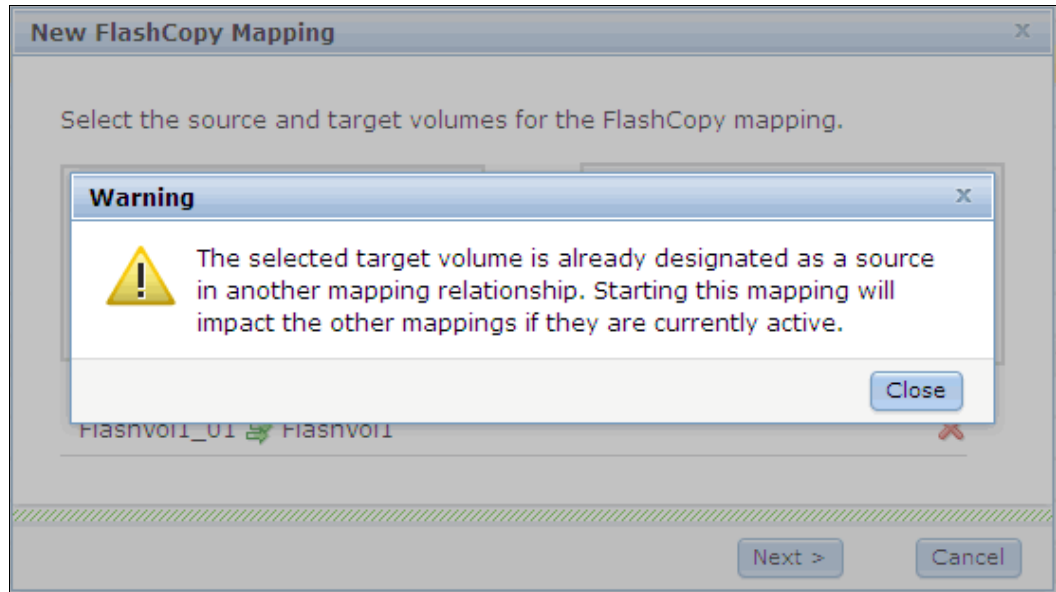


Figure 10-47 Flash restore warning

- Click **Next** and you see a snapshot preset choice, as shown in Figure 10-48.

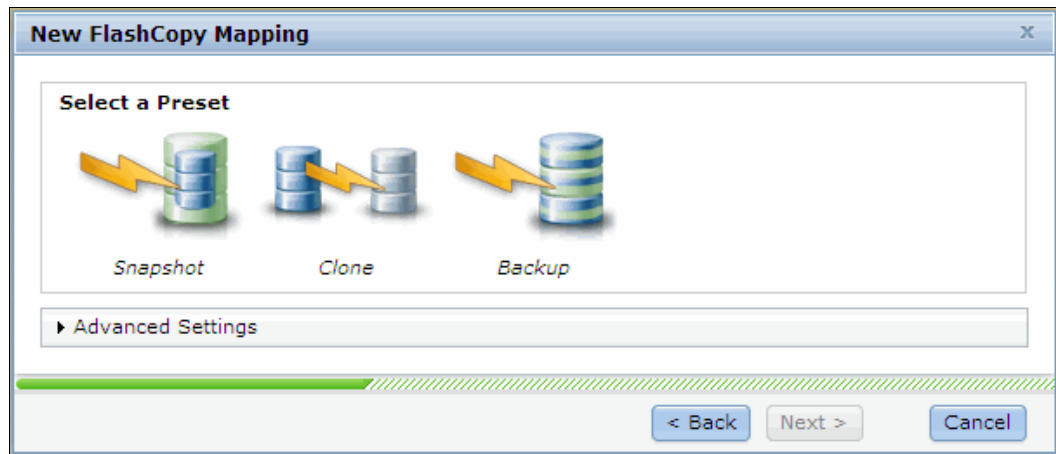


Figure 10-48 Choose snapshot preset

Select **Snapshot** and click **Next**

- In the next window, you are asked if the new mapping is to be part of a consistency group, as shown in Figure 10-49. In our example, the new mapping is not part of a consistency group, so we click **No** and then **Finish** to create the mapping.

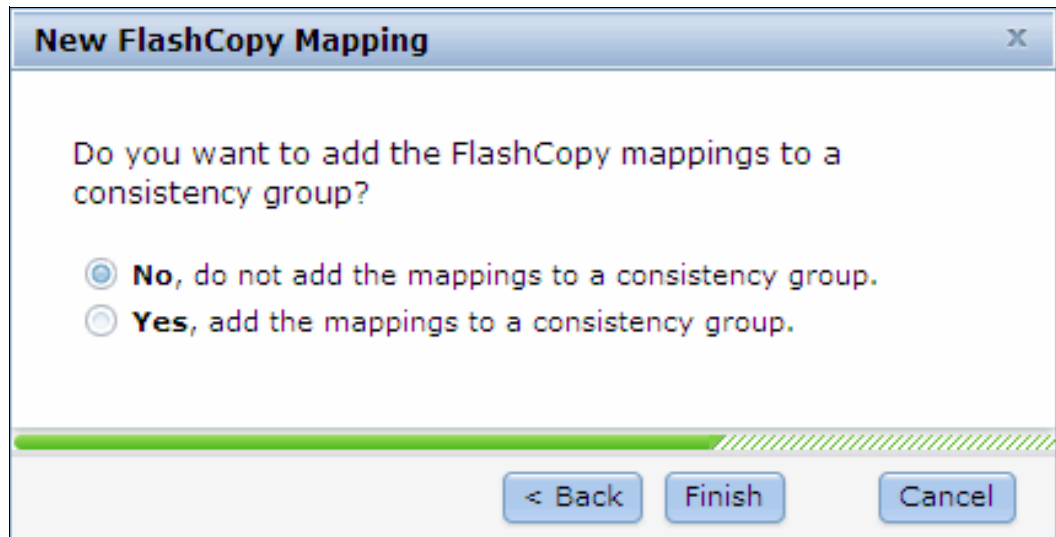


Figure 10-49 Add new mapping to consistency group

- The new reverse mapping is now created and shown in the Idle state, as shown in Figure 10-50.

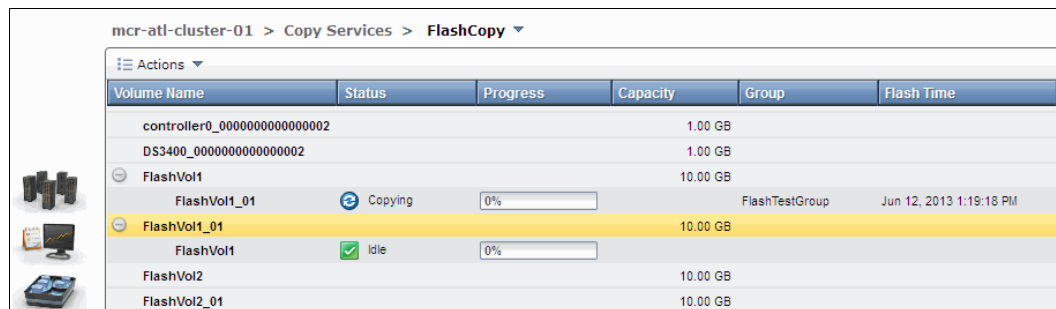


Figure 10-50 New reverse mapping

- To restore the original source volume FlashVol1 with the snapshot we took (FlashVol1_01), we select the new mapping and right-click to open the Actions menu, as shown in Figure 10-51.

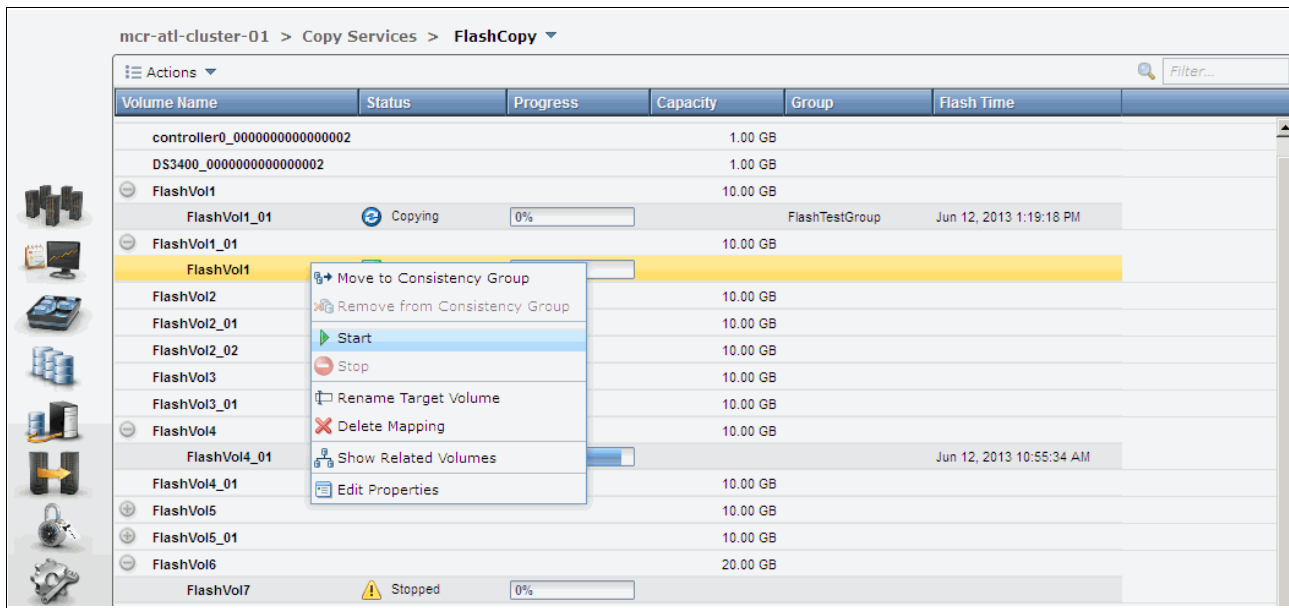


Figure 10-51 Starting the reverse mapping

- Click **Start** to write over FlashVol1 with the original bitmap data that was saved in the FlashCopy FlashVol1_01. The command then completes, as shown in Figure 10-52.

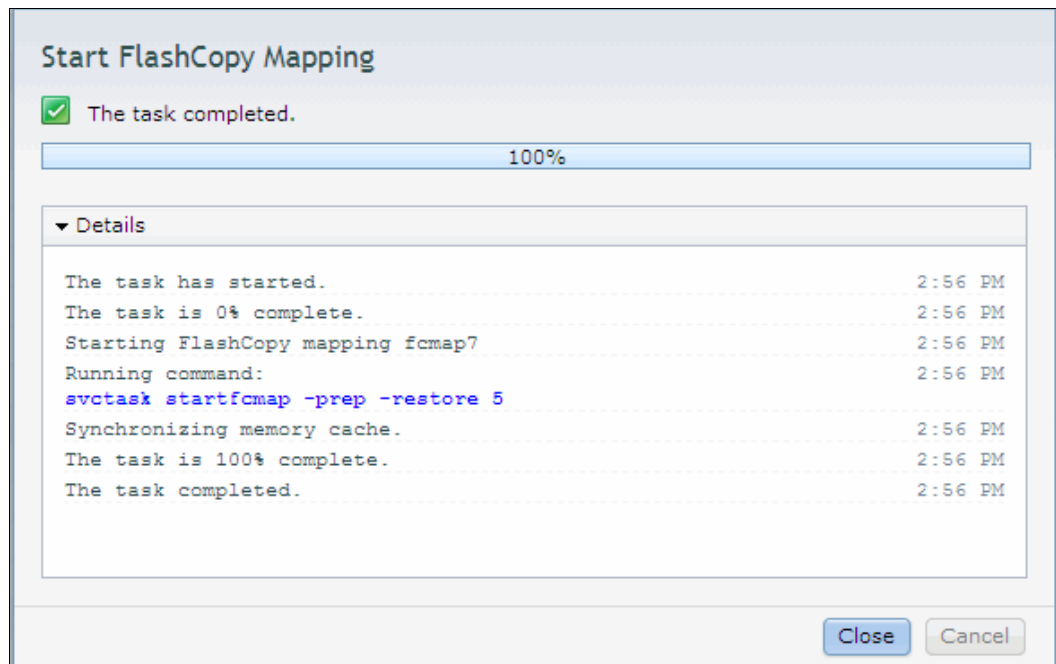


Figure 10-52 Flash Restore command

Important: The underlying command that is run by the IBM Storwize V5000 appends the **-restore** option automatically.

10.The reverse mapping now shows as 100% copied, as shown in Figure 10-53.

Volume Name	Status	Progress	Capacity	Group	Flash Time
controller0_0000000000000002			1.00 GB		
DS3400_0000000000000002			1.00 GB		
FlashVol1			10.00 GB		
FlashVol1_01	Copying	0%		FlashTestGroup	Jun 12, 2013 1:19:18 PM
FlashVol1_01	Copied	100%	10.00 GB		Jun 12, 2013 2:53:30 PM
FlashVol2			10.00 GB		
FlashVol2_01			10.00 GB		
FlashVol2_02			10.00 GB		
FlashVol3			10.00 GB		

Figure 10-53 Source volume restore complete

10.1.6 Managing a FlashCopy consistency group

FlashCopy consistency groups can be managed by clicking **Consistency Groups** under the Copy Services function icon, as shown in Figure 10-54.

IBM Storwize V5000 Welcome, superuser Legal | Log out | Help IBM.

mcr-atl-cluster-01 > Home > Overview

Suggested Tasks

48 Internal Drives
1 External Storage System
6 MDisks
3 Pools
18 Volumes
2 Fibre Channel Hosts
1 SAS Host
3 iSCSI Hosts

Copy Services

- FlashCopy
- Consistency Groups**
- FlashCopy Mappings
- Remote Copy
- Partnerships

the objects that need to be configured. To learn more about each object, click the icon in the Learning modules include a tutorial of the steps that are required to complete the task. To configure these objects, either select the associated task from Suggested Tasks or use the icons in the left navigation.

▶ Visit the Information Center

Figure 10-54 Access to the Consistency Groups window

As shown in Figure 10-55, the Consistency Groups window is where you can manage consistency groups and FlashCopy mappings.

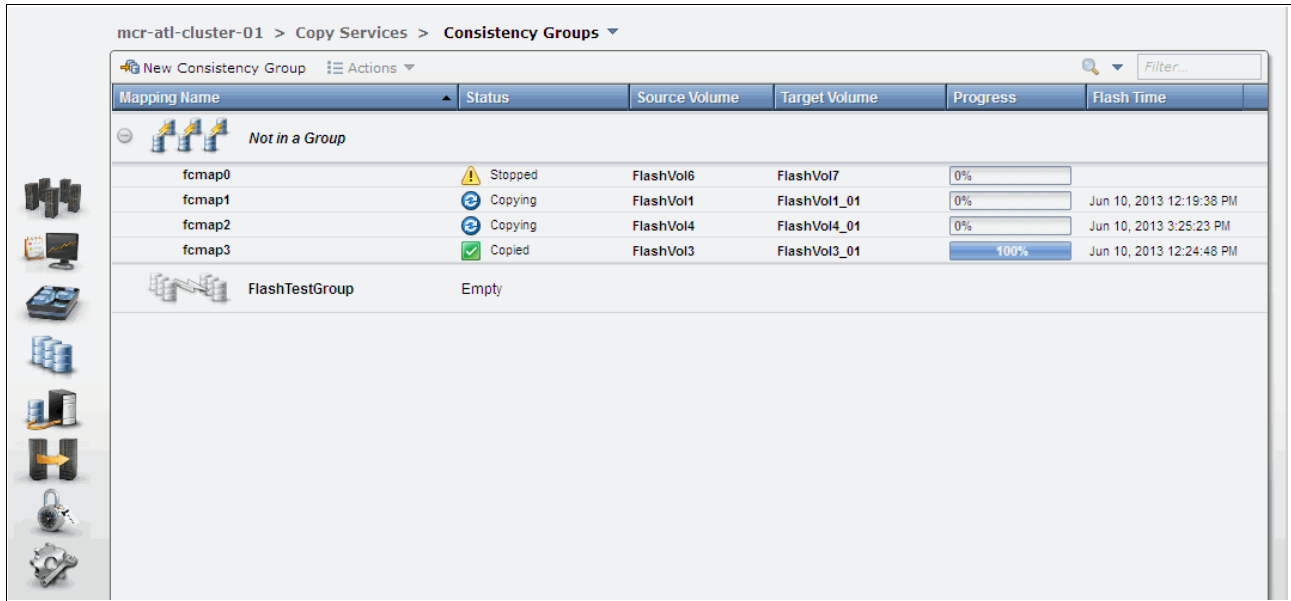


Figure 10-55 Consistency Groups window

In the left pane of the Consistency Groups window, you can list the consistency groups that you need. Click **Not in a Group**, and then expand your selection by clicking the plus (+) icon next to it. All the FlashCopy mappings that are not in any consistency groups are displayed underneath.

In the lower pane of the Consistency Groups window, you can discover the properties of a consistency group and the FlashCopy mappings in it. You can also take action on any consistency groups and FlashCopy mappings within the Consistency Groups window, as allowed by their state. For more information, see 10.1.5, “Managing FlashCopy mappings” on page 467.

Creating a FlashCopy consistency group

To create a FlashCopy consistency group, click **New Consistency Group** at the top of the Consistency Groups window, as shown in Figure 10-56.

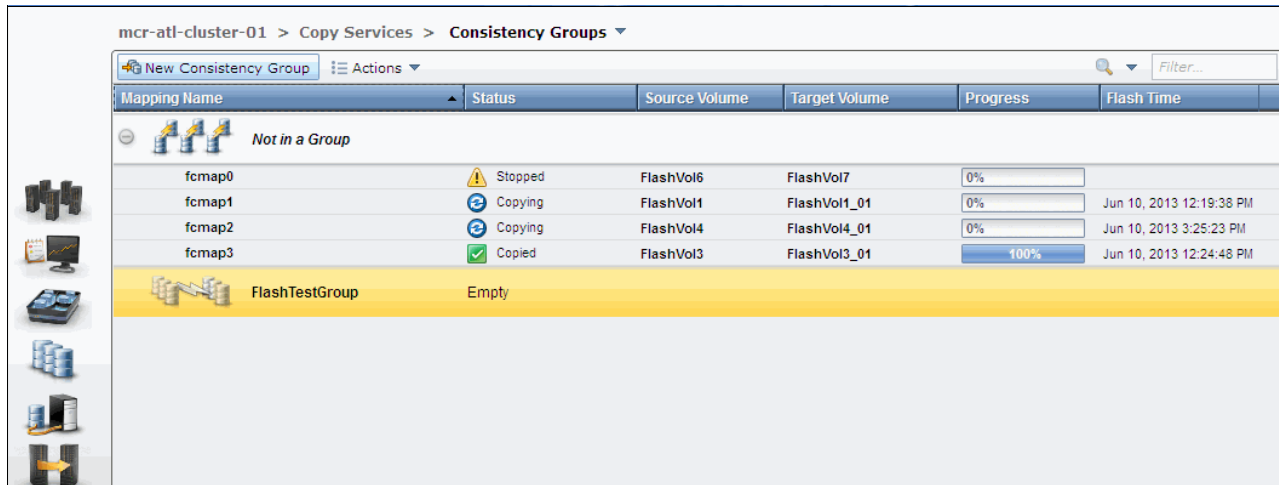


Figure 10-56 New Consistency Group option

You are prompted to enter the name of the new consistency group, as shown in Figure 10-57. Following your naming conventions, enter the name of the new consistency group in the name field and click **Create**.

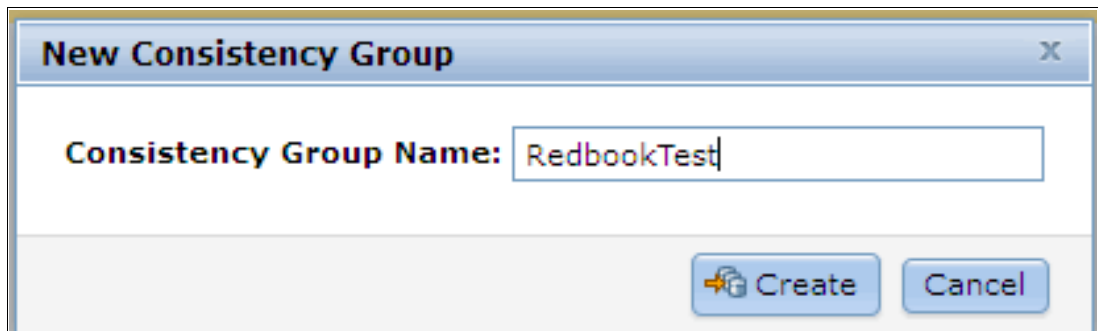


Figure 10-57 Entering the name for the consistency group

After the creation process completes, you find a new consistency group, as shown in Figure 10-58.

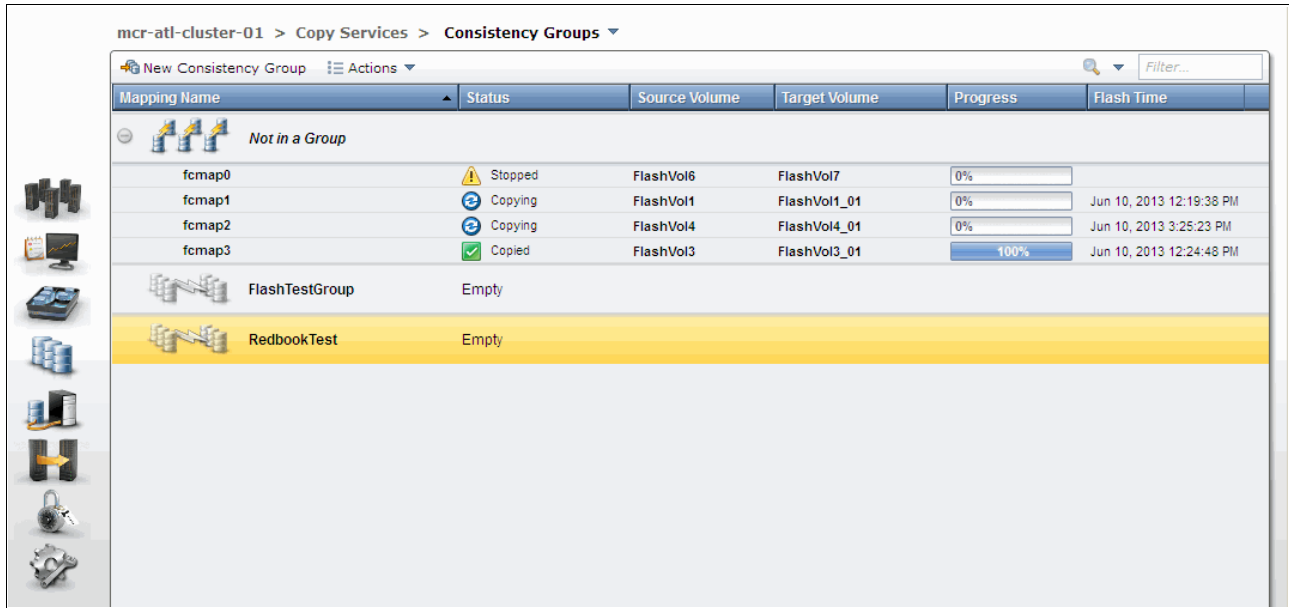


Figure 10-58 New consistency group

You can rename the Consistency Group by selecting it and then right-clicking or by using the Actions drop-down menu. Select **Rename** and enter the new name, as shown in Figure 10-59. Next to the name of the consistency group, the state shows that it is now an empty consistency group with no FlashCopy mapping in it.

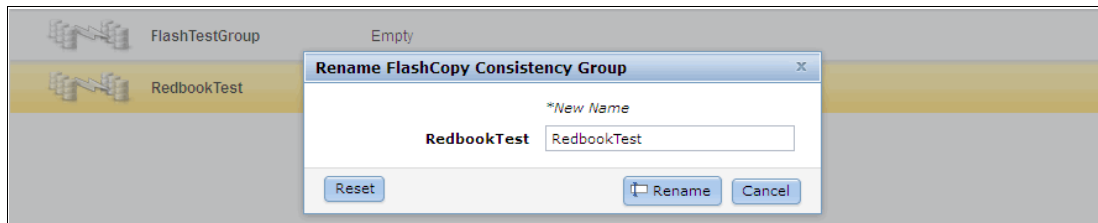


Figure 10-59 Renaming a consistency group

Adding FlashCopy mappings to a consistency group

Click **Not in a Group** to list all the FlashCopy mappings with no Consistency Group. You can add FlashCopy mappings to a Consistency Group by selecting them and clicking the **Move to Consistency Group** option from the Actions drop-down menu, as shown in Figure 10-60 on page 494.

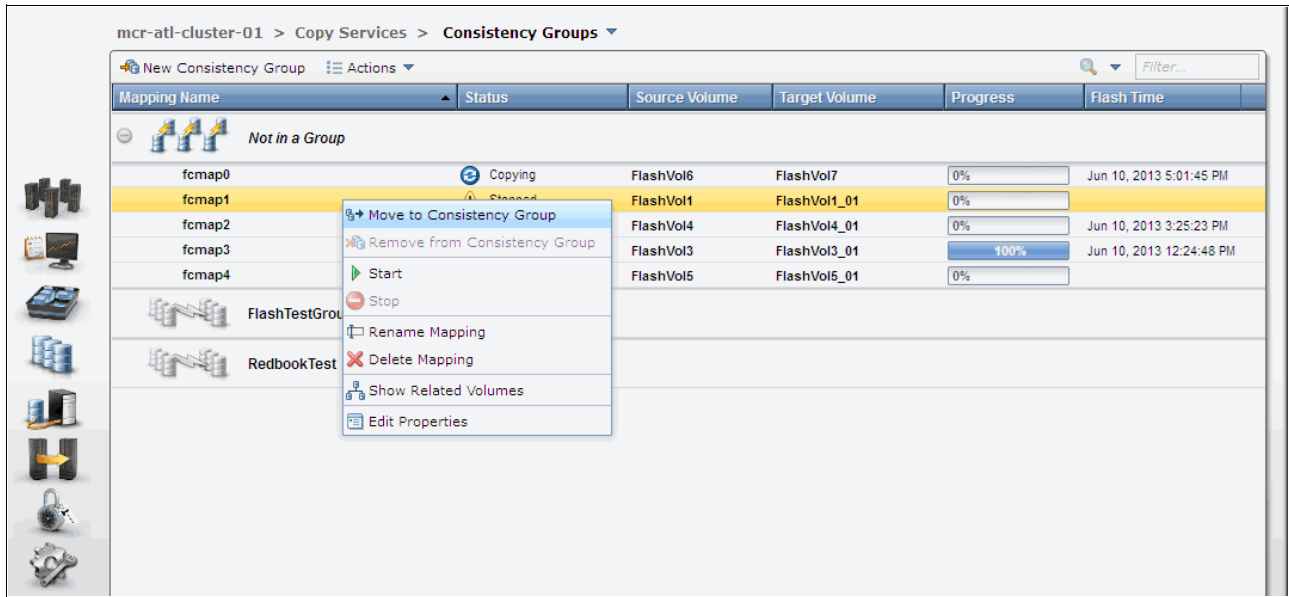


Figure 10-60 Select the FlashCopy mappings to add to a consistency group

Important: You cannot move mappings that are copying. Selecting a snapshot that is already running results in the Move to Consistency Group option being disabled.

Selections of a range are performed by highlighting a mapping, pressing and holding the Shift key, and clicking the last item in the range. Multiple selections can be made by pressing and holding the Ctrl key and clicking each mapping individually. The option is also available by right-clicking individual mappings.

You are prompted to specify which consistency group you want to move the FlashCopy mapping into, as shown in Figure 10-61. Choose from the list in the drop-down menu. Click **Move to Consistency Group** to continue.

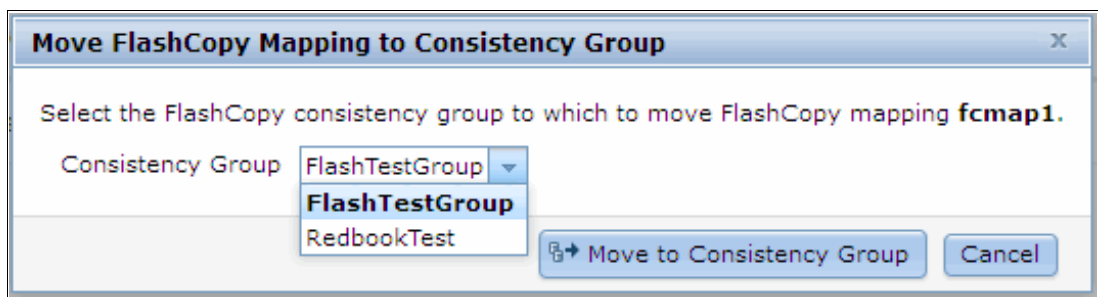


Figure 10-61 Select consistency group

After the action completes, you find that the FlashCopy mappings you selected were removed from the Not In a Group list to the consistency group you chose.

Starting a consistency group

To start a consistency group, highlight the required group and click **Start** from the Actions drop-down menu or right-click the consistency group, as shown in Figure 10-62.

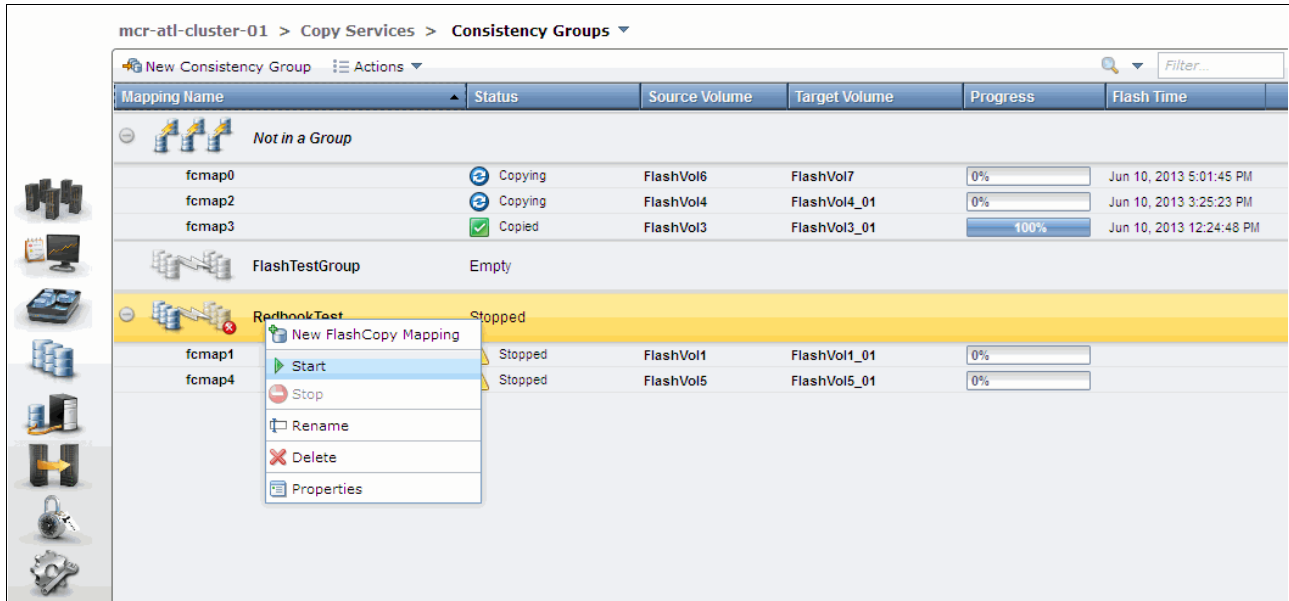


Figure 10-62 Start a consistency group

After you start the consistency group, all the FlashCopy mappings in the consistency group start at the same time. The state of consistency group and all the underlying mappings changes to Copying, as shown in Figure 10-63.

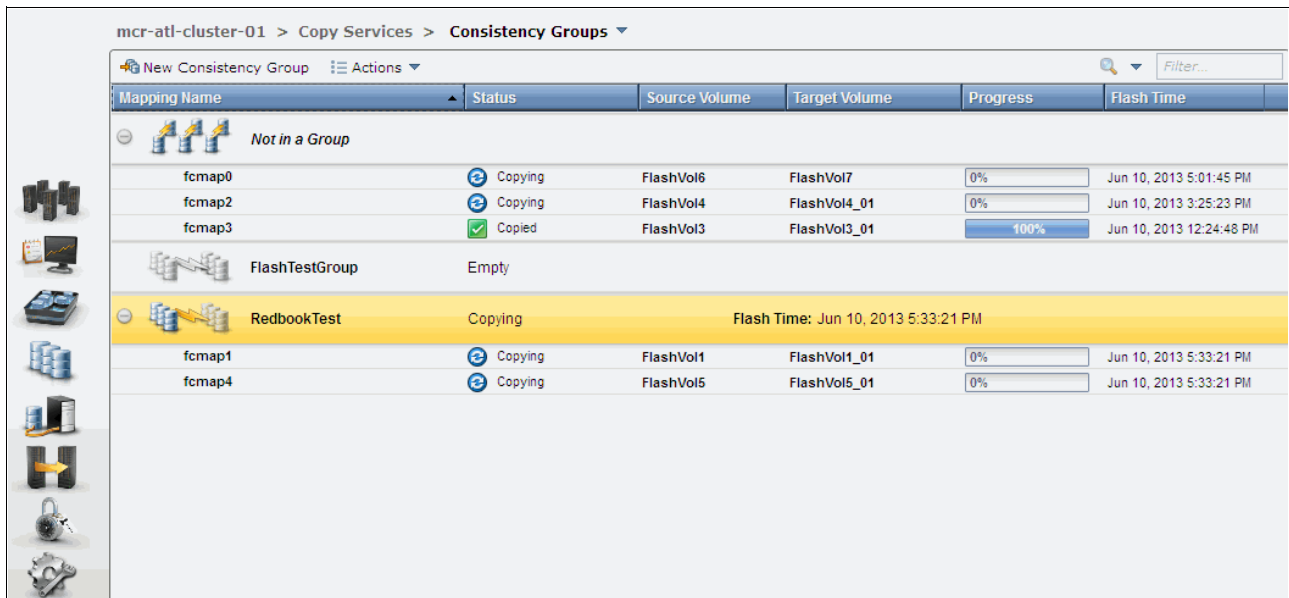


Figure 10-63 Consistency group start completes

Stopping a consistency group

The consistency group can be stopped by selecting **Stop** from the Actions drop-down menu or right-clicking, as shown in Figure 10-64.

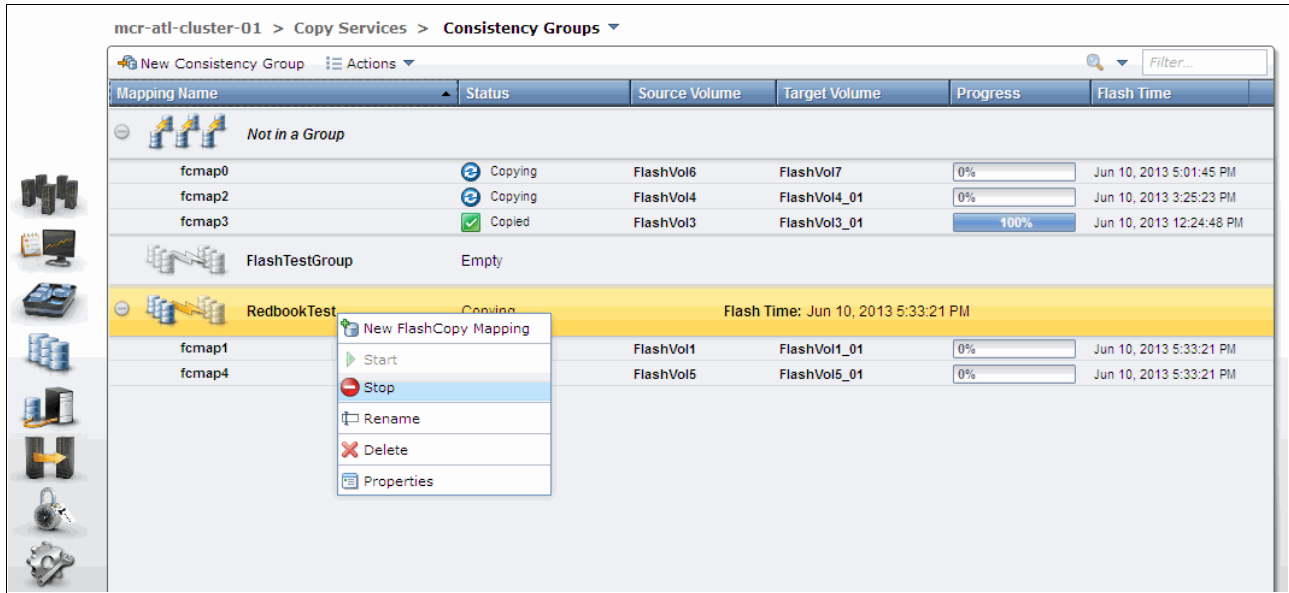


Figure 10-64 Stop a consistency group

After the stop process completes, the FlashCopy mappings in the consistency group are in the Stopped state and a red X icon appears on the function icon of this consistency group to indicate an alert, as shown in Figure 10-65.

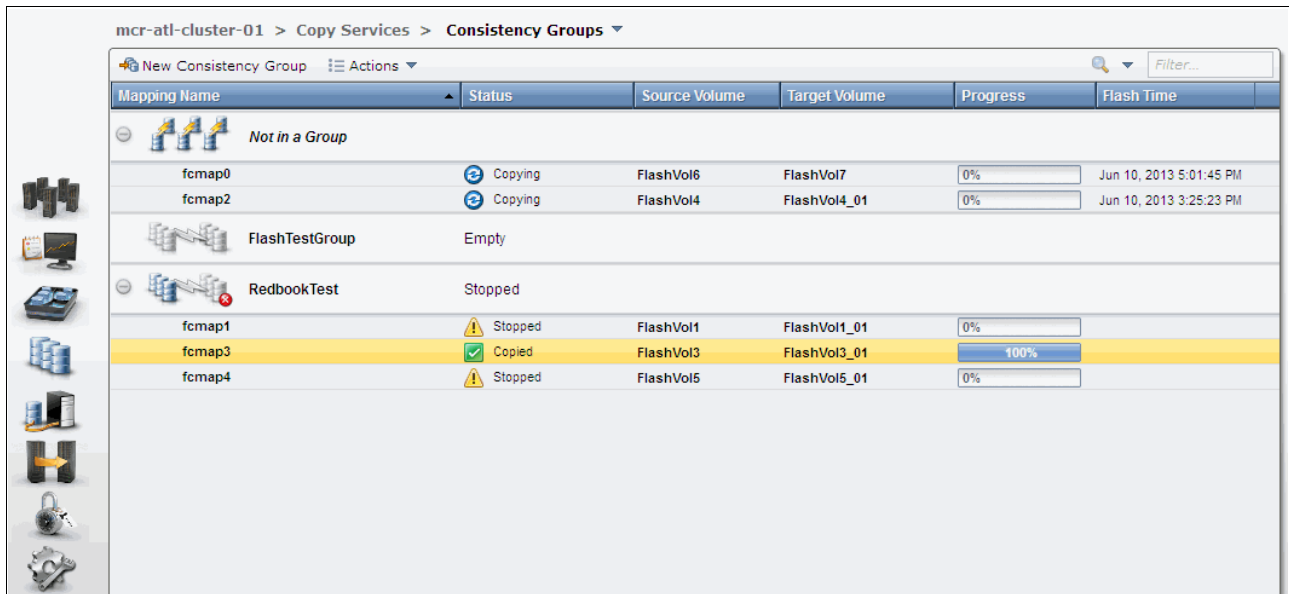


Figure 10-65 Consistency group stop completes

Previously copied relationships that were added to a consistency group that was later stopped before all members of the consistency group completed synchronization remain in the Copied state.

Removing FlashCopy mappings from a consistency group

FlashCopy mappings can be removed from a consistency group by selecting the FlashCopy mappings and clicking **Remove from Consistency Group** from the Actions drop-down menu of the FlashCopy mapping or right-clicking, as shown in Figure 10-66.

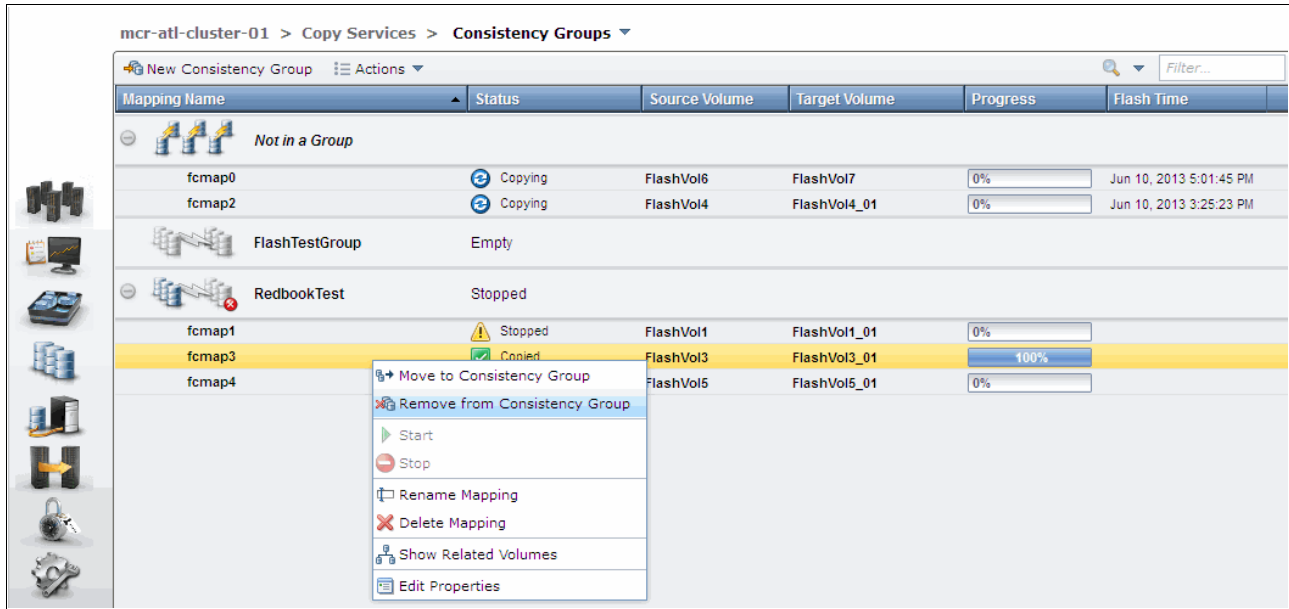


Figure 10-66 Remove from consistency group

The FlashCopy mappings are returned to the Not in a Group list after they are removed from the consistency group.

Deleting a consistency group

A consistency group can be deleted by clicking **Delete** from the Actions drop-down menu or right-clicking the selected group, as shown in Figure 10-67.

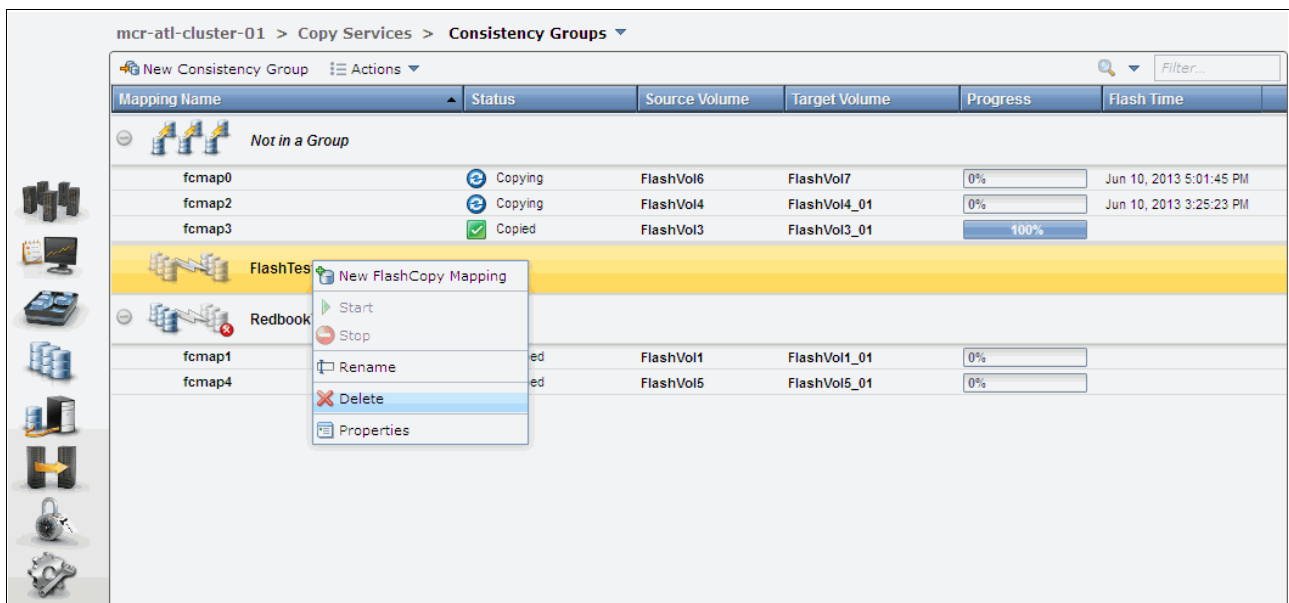


Figure 10-67 Delete a consistency group

Restoring from a FlashCopy Consistency Group

It is possible to manipulate FlashCopy mappings that were captured as part of a consistency group to restore the source volumes of those mappings to the state they were all in at the time the FlashCopy was taken.

To restore a consistency group from a FlashCopy, we must create a reverse mapping of all the individual volumes that are contained within the original consistency group. In our example, we have two FlashCopy mappings (fcmap1 and fcmap4) in a consistency group that is known as FlashTestGroup, as shown in Figure 10-68.

Mapping Name	Status	Source Volume	Target Volume	Progress	Flash Time
Not in a Group					
FlashTestGroup Copying Flash Time: Jun 12, 2013 1:19:18 PM					
fcmap1	Copying	FlashVol1	FlashVol1_01	0%	Jun 12, 2013 1:19:18 PM
fcmap4	Copying	FlashVol5	FlashVol5_01	0%	Jun 12, 2013 1:19:18 PM
RedbookTest Idle or Copied Flash Time: Jun 12, 2013 1:29:03 PM					

Figure 10-68 Creating FlashCopy reverse mapping

Complete the following steps:

1. Click **New Consistency Group** in the upper left corner (as shown in Figure 10-68) and create a consistency group. In our example, we created a group called RedBookTest.
2. Follow the procedure that is described in “Restoring from a FlashCopy” on page 485 to create reverse mappings for each of the mappings that exist in the source consistency group (FlashTestGroup). When prompted to add to a consistency group (as shown in Figure 10-49 on page 488), select **Yes** and from the drop-down menu and then select the new “reverse” consistency group that you created in step 2. In our example, this group is RedBookTest. The result should be similar to what is shown in Figure 10-69.

Mapping Name	Status	Source Volume	Target Volume	Progress	Flash Time
Not in a Group					
FlashTestGroup Copying Flash Time: Jun 12, 2013 1:19:18 PM					
fcmap1	Copying	FlashVol1	FlashVol1_01	0%	Jun 12, 2013 1:19:18 PM
fcmap4	Copying	FlashVol5	FlashVol5_01	0%	Jun 12, 2013 1:19:18 PM
RedbookTest Idle or Copied Flash Time: Jun 12, 2013 1:29:03 PM					
fcmap3	Copied	FlashVol5_01	FlashVol5	100%	Jun 12, 2013 1:29:03 PM
fcmap7	Copied	FlashVol1_01	FlashVol1	100%	

Figure 10-69 Reverse Consistency group populated.

- To restore the consistency group, highlight the reverse consistency group and click **Start**, as shown in Figure 10-70.

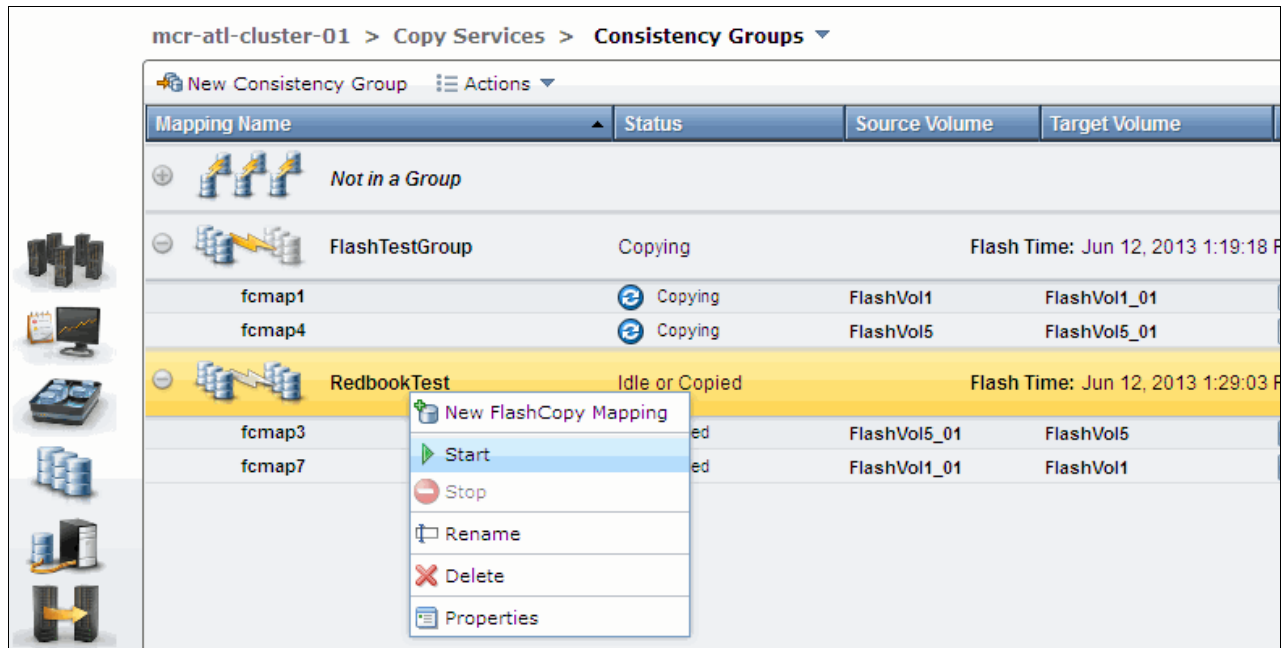


Figure 10-70 Starting Consistency group restore

- Click **Start** to overwrite FlashVol1 and FlashVol5 with the original bitmap data that was saved in the FlashTestGroup FlashCopy consistency group mapping. The command completes, as shown in Figure 10-71.

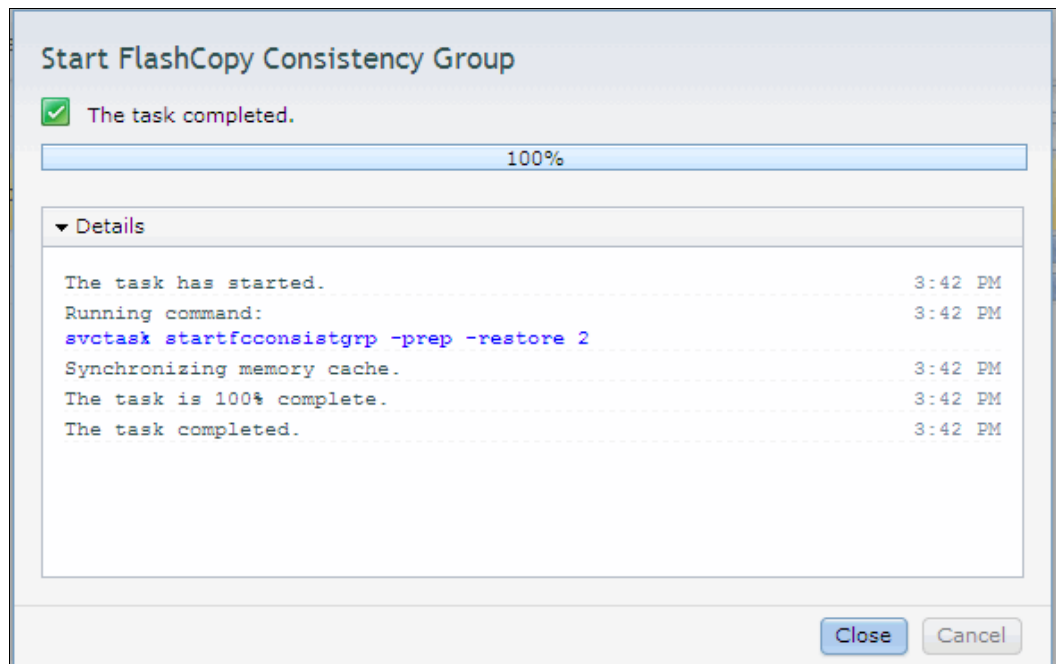
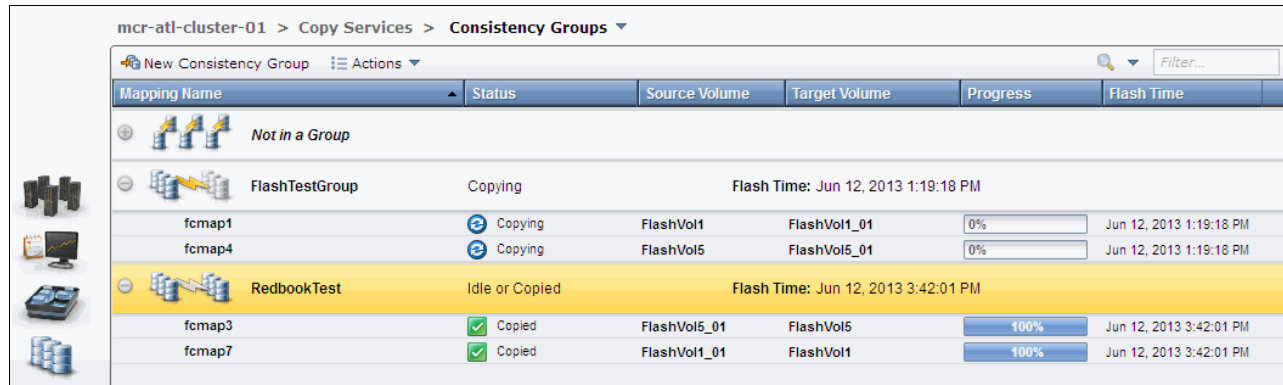


Figure 10-71 Consistency Group restore command

Important: The IBM Storwize V5000 automatically appends the **-restore** option to the command.

5. Click **Close** and the command panel returns to the Consistency Group window. The reverse consistency group now shows as a 100% copied and all volumes in the original FlashTestGroup were restored, as shown in Figure 10-72.



Mapping Name	Status	Source Volume	Target Volume	Progress	Flash Time
Not in a Group					
FlashTestGroup Copying Flash Time: Jun 12, 2013 1:19:18 PM					
fcmap1	Copying	FlashVol1	FlashVol1_01	0%	Jun 12, 2013 1:19:18 PM
fcmap4	Copying	FlashVol5	FlashVol5_01	0%	Jun 12, 2013 1:19:18 PM
RedbookTest Idle or Copied Flash Time: Jun 12, 2013 3:42:01 PM					
fcmap3	Copied	FlashVol5_01	FlashVol5	100%	Jun 12, 2013 3:42:01 PM
fcmap7	Copied	FlashVol1_01	FlashVol1	100%	Jun 12, 2013 3:42:01 PM

Figure 10-72 Consistency Group restored

10.2 Remote Copy

In this section, we describe how the Remote Copy function works in IBM Storwize V5000. We also provide the implementation steps for Remote Copy configuration and management by using the GUI.

Remote Copy consists of three methods for copying: Metro Mirror, Global Mirror, and Global Mirror with Change Volumes. Metro Mirror is designed for metropolitan distances with a synchronous copy requirement. Global Mirror is designed for longer distances without requiring the hosts to wait for the full round-trip delay of the long-distance link through asynchronous methodology. Global Mirror with Change Volumes is an added piece of functionality for Global Mirror that is designed to attain consistency on lower-quality network links.

Metro Mirror and Global Mirror are IBM branded terms for the functions Synchronous Remote Copy and Asynchronous Remote Copy. Throughout this book, the term “Remote Copy” is used to refer to both functions where the text applies to each term equally.

10.2.1 Remote Copy concepts

Remote Copy concepts are described in this section.

Partnership

When a partnership is created, we connect two separate IBM Storwize V5000 systems or an IBM SAN Volume Controller, Storwize V3700, or Storwize V7000, and an IBM Storwize V5000. After the partnership creation is configured on both systems, further communication between the node canisters in each of the storage systems is established and maintained by the SAN. All inter-cluster communication goes through the Fibre Channel network.

The partnership must be defined on both IBM Storwize V5000 or on the IBM Storwize V5000 and the other IBM SAN Volume Controller, Storwize V3700, or Storwize V7000 storage system to make the partnership fully functional.

Interconnection: Interconnects between IBM Storwize products were introduced in Version 6.3.0. Because IBM Storwize V5000 supports only version 7.10 or higher, there is no problem with support for this functionality. However, any other Storwize product must be at a minimum level of 6.3.0 to connect to the IBM Storwize V5000 and the IBM Storwize V5000 must set the replication layer by using the `svctask chsystem -layer replication` limitations that are described next.

Introduction to layers

IBM Storwize V5000 implements the concept of *layers*. Layers determine how the IBM Storwize portfolio interacts with the IBM SAN Volume Controller. Currently, there are two layers: replication and storage.

The replication layer is used when you want to use the IBM Storwize V5000 with one or more IBM SAN Volume Controllers as a Remote Copy partner. The storage layer is the default mode of operation for the IBM Storwize V5000, and is used when you want to use the IBM Storwize V5000 to present storage to an IBM SAN Volume Controller.

The layer for the IBM Storwize V5000 can be switched by running `svctask chsystem -layer replication`. Generally, switch the layer while your IBM Storwize V5000 system is not in production. This situation prevents potential disruptions because layer changes are not I/O-tolerant.

Figure 10-73 shows the effect of layers on IBM SAN Volume Controller and IBM Storwize V5000 partnerships.

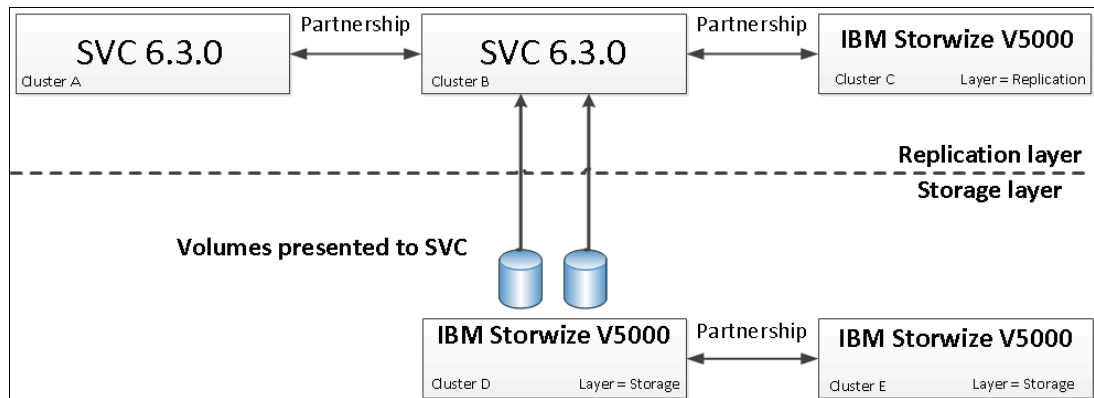


Figure 10-73 IBM Storwize V5000 virtualization layers

The replication layer allows an IBM Storwize V5000 system to be a Remote Copy partner with an IBM SAN Volume Controller. The storage layer allows an IBM Storwize V5000 system to function as back-end storage for an IBM SAN Volume Controller. An IBM Storwize V5000 system cannot be in both layers at the same time.

Limitations on the SAN Volume Controller and Storwize V5000 partnership

IBM SAN Volume Controller and IBM Storwize V5000 systems can be partners in a Remote Copy partnership. However, the following limitations apply:

- ▶ The layer for the V5000 must be set to replication. The default is storage.
- ▶ If any other SAN Volume Controller or IBM Storwize V5000 ports are visible on the SAN (aside from the ones on the cluster where you are making the changes), you cannot change the layer.
- ▶ If any host object is defined to an IBM SAN Volume Controller or IBM Storwize V5000 system, you cannot change the layer.
- ▶ If any MDisks from an IBM Storwize V5000 other than the one you are making the layer change on are visible, you cannot change the layer.
- ▶ If any cluster partnership is defined, you cannot change the layer.

Partnership topologies

A partnership between up to four IBM Storwize V5000 systems is allowed.

The following typical partnership topologies between multiple IBM Storwize V5000s are available:

- ▶ Daisy-chain topology, as shown in Figure 10-74.



Figure 10-74 Daisy chain partnership topology for IBM Storwize V5000

- ▶ Triangle topology, as shown in Figure 10-75.

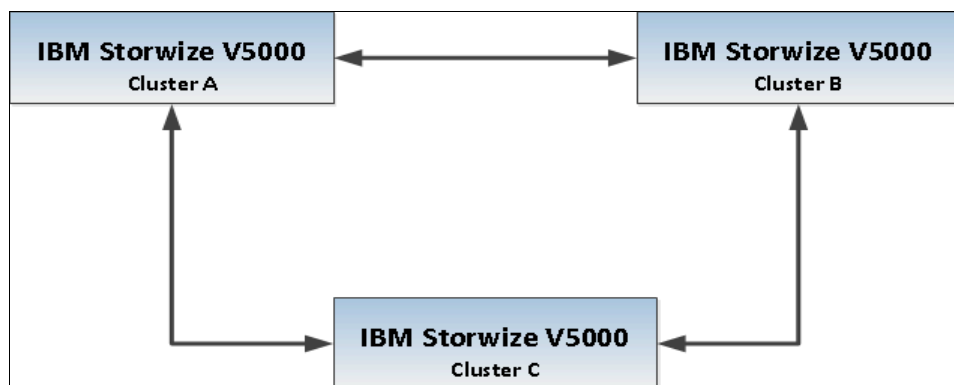


Figure 10-75 Triangle partnership topology for IBM Storwize V5000

- ▶ Star topology, as shown in Figure 10-76.

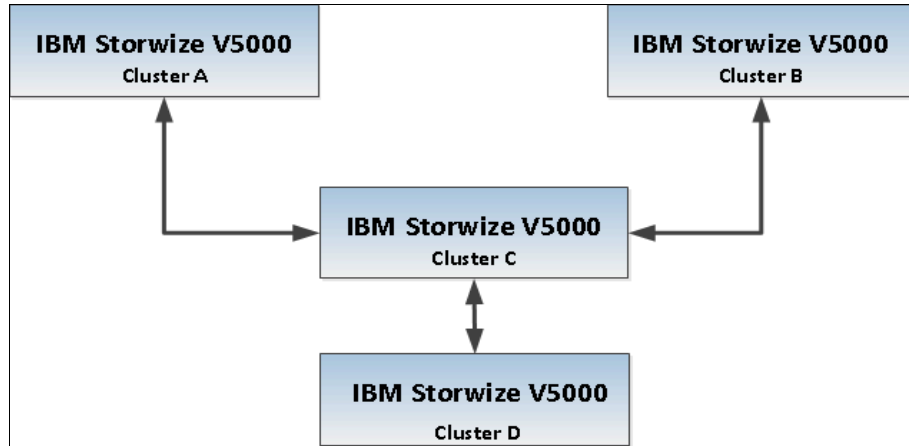


Figure 10-76 Star topology for IBM Storwize V5000

- ▶ Full-meshed topology, as shown in Figure 10-77.

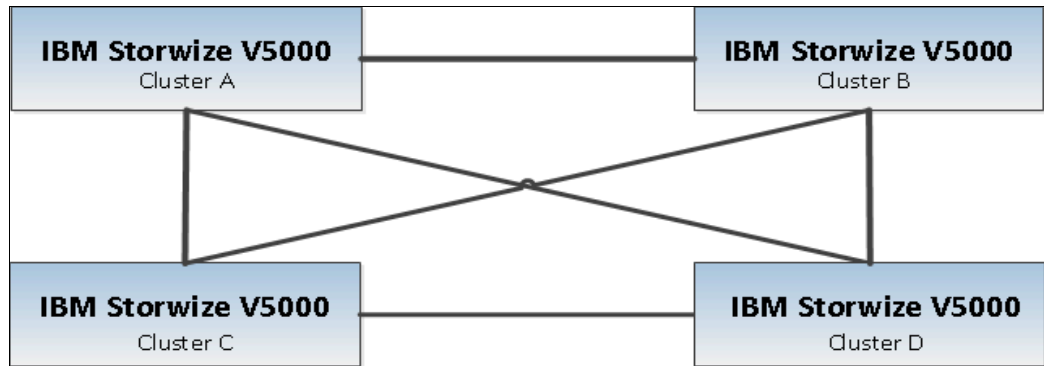


Figure 10-77 Full meshed IBM Storwize V5000

Partnerships: These partnerships are valid for configurations with SAN Volume Controllers and IBM Storwize V5000 systems if the IBM Storwize V5000 systems are using the replication layer. They are also valid for Storwize V3700 and V7000 products.

Partnership states

A partnership has the following states:

- ▶ Partially Configured

Indicates that only one cluster partner is defined from a local or remote cluster to the displayed cluster and is started. For the displayed cluster to be configured fully and to complete the partnership, you must define the cluster partnership from the cluster that is displayed to the corresponding local or remote cluster.

- ▶ Fully Configured

Indicates that the partnership is defined on the local and remote clusters and is started.

- ▶ Remote Not Present

Indicates that the remote cluster is not present for the partnership.

- ▶ **Partially Configured (Local Stopped)**
Indicates that the local cluster is only defined to the remote cluster and the local cluster is stopped.
- ▶ **Fully Configured (Local Stopped)**
Indicates that a partnership is defined on the local and remote clusters and the remote cluster is present, but the local cluster is stopped.
- ▶ **Fully Configured (Remote Stopped)**
Indicates that a partnership is defined on the local and remote clusters and the remote cluster is present, but the remote cluster is stopped.
- ▶ **Fully Configured (Local Excluded)**
Indicates that a partnership is defined between a local and remote cluster; however, the local cluster was excluded. This state can occur when the fabric link between the two clusters was compromised by too many fabric errors or slow response times of the cluster partnership.
- ▶ **Fully Configured (Remote Excluded)**
Indicates that a partnership is defined between a local and remote cluster; however, the remote cluster was excluded. This state can occur when the fabric link between the two clusters was compromised by too many fabric errors or slow response times of the cluster partnership.
- ▶ **Fully Configured (Remote Exceeded)**
Indicates that a partnership is defined between a local and remote cluster and the remote is available; however, the remote cluster exceeds the number of allowed clusters within a cluster network. The maximum of four clusters can be defined in a network. If the number of clusters exceeds that limit, the IBM Storwize V5000 system determines the inactive cluster or clusters by sorting all the clusters by their unique identifier in numerical order. The inactive cluster partner that is not in the top four of the cluster-unique identifiers shows Fully Configured (Remote Exceeded).

Remote Copy relationships

A Remote Copy relationship is a relationship between two individual volumes of the same size. These volumes are called a *master (source) volume* and an *auxiliary (target) volume*.

Typically, the master volume contains the production copy of the data and is the volume that the application normally accesses. The auxiliary volume often contains a backup copy of the data and is used for disaster recovery.

The master and auxiliary volumes are defined when the relationship is created, and these attributes never change. However, either volume can operate in the primary or secondary role as necessary. The primary volume contains a valid copy of the application data and receives updates from the host application, which is analogous to a source volume. The secondary volume receives a copy of any updates to the primary volume because these updates are all transmitted across the mirror link. Therefore, the secondary volume is analogous to a continuously updated target volume. When a relationship is created, the master volume is assigned the role of primary volume and the auxiliary volume is assigned the role of secondary volume. The initial copying direction is from master to auxiliary. When the relationship is in a consistent state, you can reverse the copy direction.

The two volumes in a relationship must be the same size. The Remote Copy relationship can be established on the volumes within one IBM Storwize V5000 storage system, which is called an *intra-cluster relationship*. The relationship can also be established in different IBM Storwize V5000 storage systems or between an IBM Storwize V5000 storage system and an IBM SAN Volume Controller, IBM Storwize V3700, or IBM Storwize V7000, which are called *inter-cluster relationships*.

Important: The use of Remote Copy target volumes as Remote Copy source volumes is not allowed. A FlashCopy target volume can be used as Remote Copy source volume and also as a Remote Copy target volume.

Metro Mirror

Metro Mirror is a type of Remote Copy that creates a synchronous copy of data from a master volume to an auxiliary volume. With synchronous copies, host applications write to the master volume but do not receive confirmation that the write operation completed until the data is written to the auxiliary volume. This action ensures that both volumes have identical data when the copy completes. After the initial copy completes, the Metro Mirror function always maintains a fully synchronized copy of the source data at the target site.

Figure 10-78 shows how a write to the master volume is mirrored to the cache of the auxiliary volume before an acknowledgement of the write is sent back to the host that issued the write. This process ensures that the auxiliary is synchronized in real time if it is needed in a failover situation.

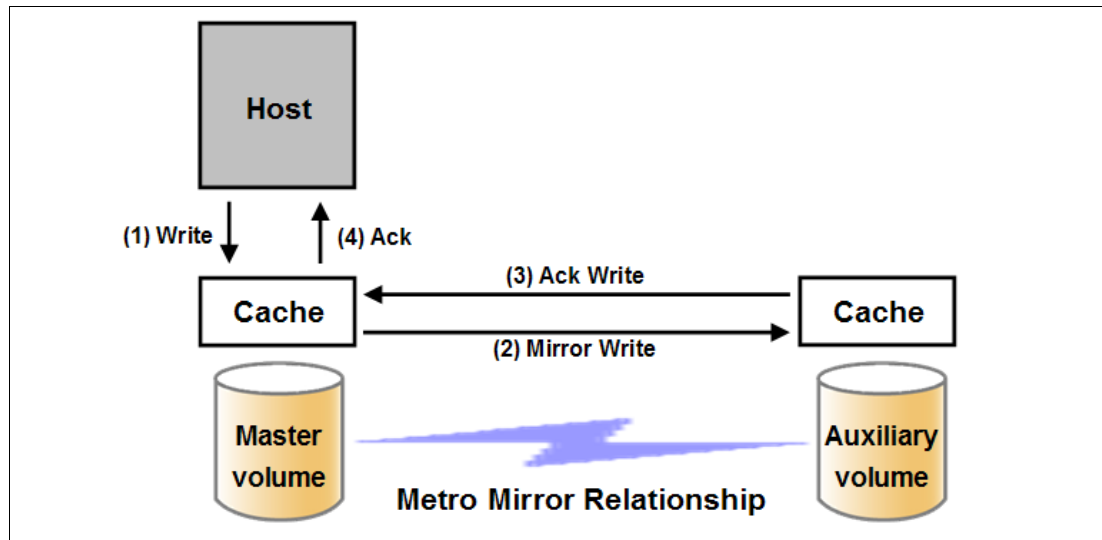


Figure 10-78 Write on volume in a Metro Mirror relationship

The Metro Mirror function supports copy operations between volumes that are separated by distances up to 300 km. For disaster recovery purposes, Metro Mirror provides the simplest way to maintain an identical copy on the primary and secondary volumes. However, as with all synchronous copies over remote distances, there can be a performance impact to host applications. This performance impact is related to the distance between primary and secondary volumes and, depending on application requirements, its use might be limited based on the distance between sites.

Global Mirror

Global Mirror provides an asynchronous copy, which means that the secondary volume is not an exact match of the primary volume at every point. The Global Mirror function provides the same function as Metro Mirror Remote Copy without requiring the hosts to wait for the full round-trip delay of the long-distance link; however, some delay can be seen on the hosts in congested or overloaded environments. Make sure that you closely monitor and understand your workload.

In a synchronous Remote Copy (which Global Mirror provides), write operations are completed on the primary site and the write acknowledgement is sent to the host before it is received at the secondary site. An update of this write operation is sent to the secondary site at a later stage, which provides the capability to perform Remote Copy over distances that exceed the limitations of synchronous Remote Copy.

The distance of Global Mirror replication is limited primarily by the latency of the WAN link that is provided. Global Mirror has a requirement of 80 ms round-trip-time for data that is sent to the remote location. The propagation delay is roughly 8.2 μ s per mile or 5 μ s per kilometer for Fibre Channel connections. Each device in the path adds more delay of about 25 μ s. Devices that use software (such as, some compression devices) adds much more time. The time that is added by software-assisted devices is highly variable and should be measured directly. Be sure to include these times when you are planning your Global Mirror design.

You should also measure application performance that is based on the expected delays before Global Mirror is fully implemented. The IBM Storwize V5000 storage system provides you with an advanced feature of Global Mirror that permits you to test performance implications before Global Mirror is deployed and a long-distance link is obtained. This advanced feature is enabled by modifying the IBM Storwize V5000 storage system parameters **gmintradelaysimulation** and **gminterdelaysimulation**. These parameters can be used to simulate the write delay to the secondary volume. The delay simulation can be enabled separately for each intra-cluster or inter-cluster Global Mirror. You can use this feature to test an application before the full deployment of the Global Mirror feature. For more information about how to enable the CLI feature, see Appendix A, “Command-line interface setup and SAN Boot” on page 609.

Figure 10-79 on page 507 shows that a write operation to the master volume is acknowledged back to the host that is issuing the write before the write operation is mirrored to the cache for the auxiliary volume.

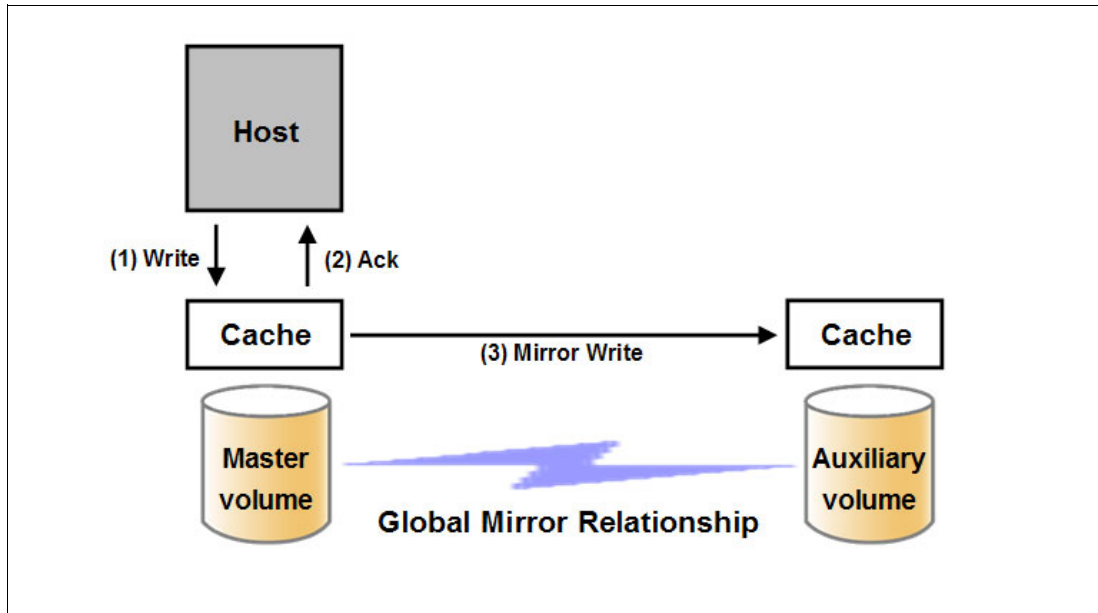


Figure 10-79 Global Mirror write sequence

The Global Mirror algorithms always maintain a consistent image on the auxiliary volume. They achieve this consistent image by identifying sets of I/Os that are active concurrently at the master, assigning an order to those sets, and applying those sets of I/Os in the assigned order at the secondary.

In a failover scenario where the secondary site must become the master source of data (depending on the workload pattern and the bandwidth and distance between local and remote site), certain updates might be missing at the secondary site. Therefore, any applications that use this data must have an external mechanism for recovering the missing updates and reapplying them; for example, a transaction log replay.

10.2.2 Global Mirror with Change Volumes

Global Mirror within the IBM Storwize V5000 is designed to achieve a recovery point objective (RPO) as low as possible so that data is as up-to-date as possible. This capability places some strict requirements on your infrastructure and in certain situations (with low network link quality or congested or overloaded hosts), you might be affected by multiple 1920 (congestion) errors.

Congestion errors happen in the following primary situations:

- ▶ Congestion at the source site through the host or network.
- ▶ Congestion in the network link or network path.
- ▶ Congestion at the target site through the host or network.

Global Mirror includes functionality that is designed to address the following conditions that negatively affect some Global Mirror implementations:

- ▶ Estimation of bandwidth requirements tends to be complex.
- ▶ It is often difficult to ensure that the latency and bandwidth requirements can be met.
- ▶ Congested hosts on the source or target site can cause disruption.
- ▶ Congested network links can cause disruption with only intermittent peaks.

To address these issues, *Change Volumes* were added as an option for Global Mirror relationships. Change Volumes use the FlashCopy functionality but cannot be manipulated as FlashCopy volumes because they are special-purpose only. Change Volumes replicate point-in-time images on a cycling period (the default is 300 seconds). This situation means that your change rate must include only the condition of the data at the point-in-time the image was taken instead of all the updates during the period. This situation can provide significant reductions in replication volume.

Figure 10-80 shows a basic Global Mirror relationship without Change Volumes.

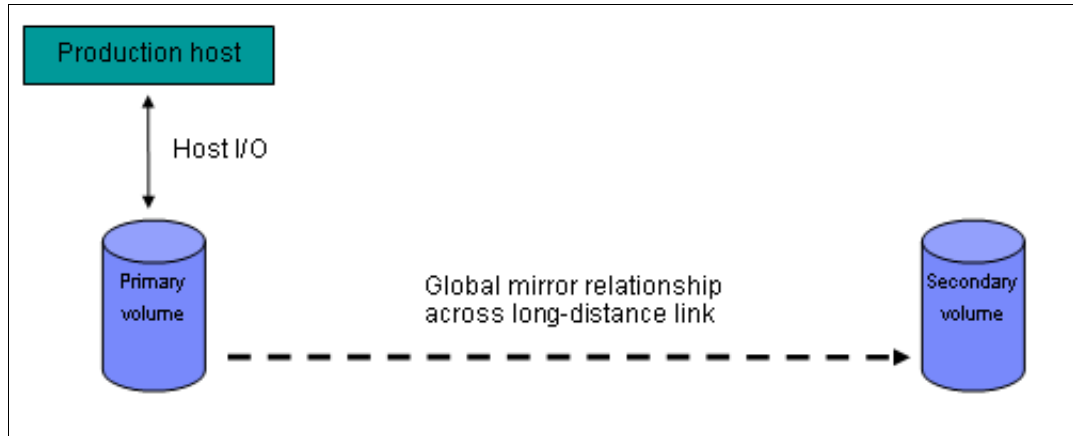


Figure 10-80 Global Mirror without Change Volumes

Figure 10-81 shows a relationship with the Change Volumes.

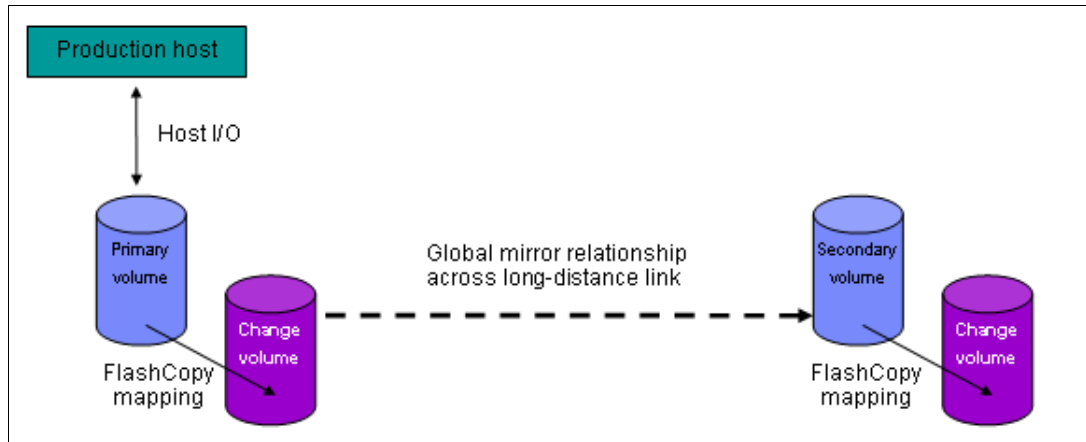


Figure 10-81 Global Mirror with Change Volumes

With Change Volumes, a FlashCopy mapping exists between the primary volume and the primary Change Volume. The mapping is updated during a cycling period (every 60 seconds to one day). The primary Change Volume is then replicated to the secondary Global Mirror volume at the target site, which is then captured in another change volume on the target site. This situation provides a consistent image at the target site and protects your data from being inconsistent during resynchronization.

Figure 10-82 shows a number of I/Os on the source volume, the same number on the target volume, and in the same order. Assuming that this set is the same set of data that is updated over and over, these updates are wasted network traffic and the I/O can be completed much more efficiently, as shown in Figure 10-83.

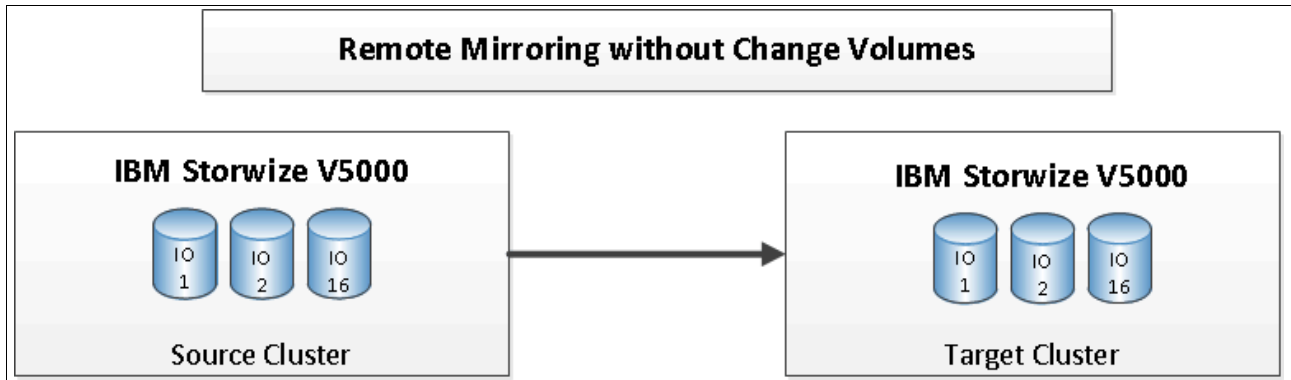


Figure 10-82 Global Mirror I/O replication without Change Volumes

In Figure 10-83, the same data is being updated repeatedly, so Change Volumes demonstrate significant I/O transmission savings because you must send only I/O number 16, which was the last I/O before the cycling period.

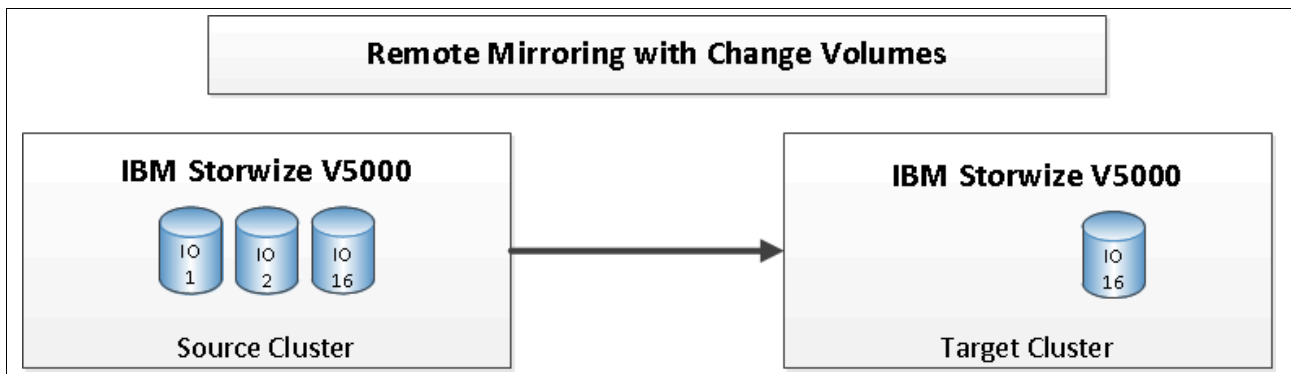


Figure 10-83 Global Mirror I/O replication with Change Volumes

The cycling period can be adjusted by running `chrcrelationship -cycleperiodseconds <60-86400>`. If a copy does not complete in the cycle period, the next cycle does not start until the prior cycle completes. It is for this reason that the use of Change Volumes gives you the following possibilities for RPO:

- ▶ If your replication completes in the cycling period, your RPO is twice the cycling period.
- ▶ If your replication does not complete within the cycling period, your RPO is twice the completion time. The next cycling period starts immediately after the prior period is finished.

Careful consideration should be put into balancing your business requirements with the performance of Global Mirror with Change Volumes. Global Mirror with Change Volumes increases the inter-cluster traffic for more frequent cycling periods, so going as short as possible is not always the answer. In most cases, the default should meet your requirements and perform reasonably well.

Important: When Global Mirror volumes with Change Volumes are used, make sure that you remember to select the Change Volume on the auxiliary (target) site. Failure to do so leaves you exposed during a resynchronization operation.

The GUI automatically creates Change Volumes for you. However, it is a limitation of this initial release that they are fully provisioned volumes. To save space, you should create thin-provisioned volumes in advance and use the existing volume option to select your change volumes.

Remote Copy consistency groups

A consistency group is a logical entity that groups copy relationships. By grouping the relationships, you can ensure that these relationships are managed in unison and the data within the group is in a consistent state. For more information about the necessity of consistency groups, see 10.1.6, “Managing a FlashCopy consistency group” on page 490.

Remote Copy commands can be issued to a Remote Copy consistency group, and, therefore, simultaneously for all Metro Mirror relationships that are defined within that consistency group, or to a single Metro Mirror relationship that is not part of a Metro Mirror consistency group.

Figure 10-84 shows the concept of Remote Copy consistency groups. Because the RC_Relationships 1 and 2 are part of the consistency group, they can be handled as one entity, while the stand-alone RC_Relationship 3 is handled separately.

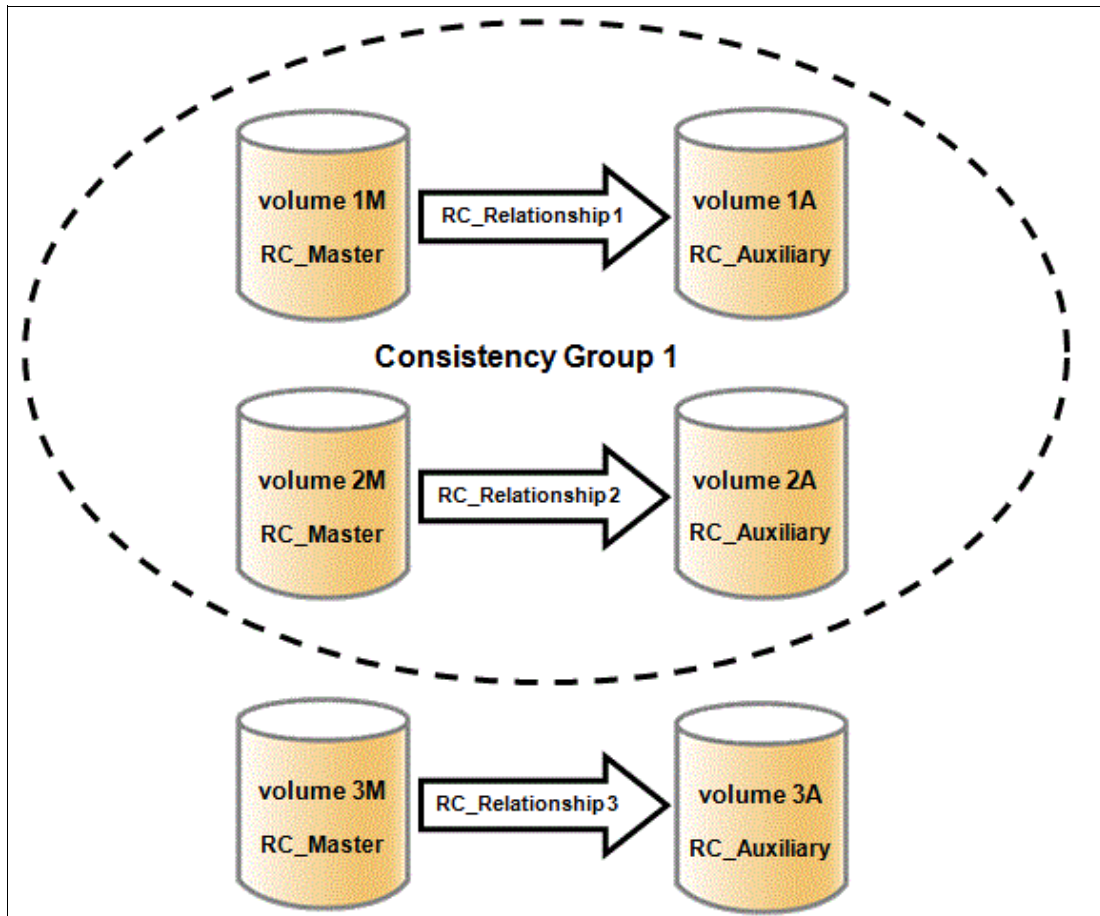


Figure 10-84 Remote Copy consistency group

Remote Copy relationships can belong only to one consistency group, but they do not have to belong to a consistency group. Relationships that are not part of a consistency group are called *stand-alone relationships*. A consistency group can contain zero or more relationships. All relationships in a consistency group must have matching primary and secondary clusters, which are sometimes referred to as *master clusters* and *auxiliary clusters*. All relationships in a consistency group must also have the same copy direction and state.

Metro Mirror and Global Mirror relationships cannot belong to the same consistency group. A copy type is automatically assigned to a consistency group when the first relationship is added to the consistency group. After the consistency group is assigned a copy type, only relationships of that copy type can be added to the consistency group.

Remote Copy and consistency group states

Stand-alone Remote Copy relationships and consistency groups share a common configuration and state model. All of the relationships in a non-empty consistency group have the same state as the consistency group.

The following states apply to the relationships and the consistency groups, except for the Empty state, which is only for consistency groups:

- ▶ **InconsistentStopped**

The primary volumes are accessible for read and write I/O operations, but the secondary volumes are not accessible for either one. A copy process must be started to make the secondary volumes consistent.
- ▶ **InconsistentCopying**

The primary volumes are accessible for read and write I/O operations, but the secondary volumes are not accessible for either one. This state indicates that a copy process is ongoing from the primary to the secondary volume.
- ▶ **ConsistentStopped**

The secondary volumes contain a consistent image, but it might be outdated about the primary volumes. This state can occur when a relationship was in the ConsistentSynchronized state and experiences an error that forces a freeze of the consistency group or the Remote Copy relationship.
- ▶ **ConsistentSynchronized**

The primary volumes are accessible for read and write I/O operations. The secondary volumes are accessible for read-only I/O operations.
- ▶ **Idling**

The primary volumes and the secondary volumes are operating in the primary role. Therefore, the volumes are accessible for write I/O operations.
- ▶ **IdlingDisconnected**

The volumes in this half of the consistency group are all operating in the primary role and can accept read or write I/O operations.
- ▶ **InconsistentDisconnected**

The volumes in this half of the consistency group are all operating in the secondary role and cannot accept read or write I/O operations.

- ▶ ConsistentDisconnected

The volumes in this half of the consistency group are all operating in the secondary role and can accept read I/O operations but not write I/O operations.

- ▶ Empty

The consistency group does not contain any relationships.

10.2.3 Remote Copy planning

Before you use Remote Copy, you must plan for its usage.

General guidelines for Remote Copy

General guidelines for Remote Copy include the following considerations:

- ▶ Partnerships between up to four IBM Storwize V5000 storage systems, IBM SAN Volume Controller systems, IBM Storwize V7000, or IBM Storwize V3700 is allowed. The partnership must be defined on any partnered IBM Storwize storage systems or IBM SAN Volume Controller systems to make it fully functional.
- ▶ The two volumes in a relationship must be the same size.
- ▶ The Remote Copy relationship can be established on the volumes within one IBM Storwize V5000 storage system or in different IBM Storwize V5000 storage systems. When the two volumes are in the same cluster, they must be in the same I/O group.
- ▶ You cannot use Remote Copy target volumes as Remote Copy source volumes. However, a FlashCopy target volume can be used as Remote Copy source volume. Other restrictions are outlined in Table 10-5 on page 514.
- ▶ The Metro Mirror function supports copy operations between volumes that are separated by distances up to 300 km.
- ▶ One Remote Copy relationship can belong only to one consistency group.
- ▶ All relationships in a consistency group must have matching primary and secondary clusters, (master clusters and auxiliary clusters). All relationships in a consistency group must also have the same copy direction and state.
- ▶ Metro Mirror and Global Mirror relationships cannot belong to the same consistency group.
- ▶ To manage multiple Remote Copy relationships as one entity, relationships can be made part of a Remote Copy consistency group, which ensures data consistency across multiple Remote Copy relationships and provides ease of management.
- ▶ An IBM Storwize V5000 storage system implements flexible resynchronization support, which enables it to resynchronize volume pairs that experienced write I/Os to both disks and to resynchronize only those regions that are known to changed.
- ▶ Global Mirror with Change Volumes should have Change Volumes that are defined for the master and auxiliary volumes.

Remote Copy configuration limits

Table 10-4 lists the Remote Copy configuration limits.

Table 10-4 Remote Copy configuration limits

Parameter	Value
Number of Remote Copy consistency groups per cluster	256
Number of Remote Copy relationships per consistency group	8,192
Number of Remote Copy relationships per I/O Group	2,048
Total Remote Copy volume capacity per I/O Group	1024 TB (This limit is the total capacity for all master and auxiliary volumes in the I/O group.)

SAN planning for Remote Copy

In this section, we describe some guidelines that can be used for planning for a SAN for Remote Copy.

Zoning recommendation

Node canister ports on each IBM Storwize V5000 must communicate with each other so that the partnership can be created. These ports must be visible to each other on your SAN. Proper switch zoning is critical to facilitating inter-cluster communication.

The following SAN zoning recommendation should be considered:

- ▶ For each node canister, exactly two Fibre Channel ports should be zoned to exactly two Fibre Channel ports from each node canister in the partner cluster.
- ▶ If dual-redundant inter-switch links (ISLs) are available, the two ports from each node should be split evenly between the two ISLs; that is, exactly one port from each node canister should be zoned across each ISL. For more information, see this website:
<http://www-01.ibm.com/support/docview.wss?uid=ssg1S1003634&myns=s033&mynp=famil yind5329743&mync=E>
- ▶ All local zoning rules should be followed. A properly configured SAN fabric is key to not only local SAN performance, but Remote Copy. For more information about these rules, see this website:
http://publib.boulder.ibm.com/infocenter/storwize/ic/index.jsp?topic=%2Fcom.ibm.storwize.V5000.doc%2Fsvc_configrulessummary_02171530.html

Fabrics: When a local fabric and a remote fabric are connected for Remote Copy purposes, the ISL hop count between a local node and a remote node cannot exceed seven.

Remote Copy link requirements

The following link requirements are valid for Metro Mirror and Global Mirror:

- ▶ Round-trip latency
The total round-trip latency must be less than 80 ms and less than 40 ms in each direction. Latency simulations should be performed with your applications before any network links are put in place to see whether the applications perform at an acceptable level while they meet the round-trip latency requirement.

► **Bandwidth**

The bandwidth must satisfy the following requirements:

- If you are not using Change Volumes, be able to sustain peak write load for all mirrored volumes and background copy traffic.
- If you are using Change Volumes with Global Mirror, be able to sustain change rate of Source Change Volumes and background copy traffic.
- Other background copy rate (the best practice is 10% to 20% of maximum peak load) for initial synchronization and resynchronization.
- Remote Copy internal communication at idle with or without Change Volumes is approximately 2.6 Mbps. This amount is the minimum amount.

Redundancy: If the link between two sites is configured with redundancy so that it can tolerate single failures, the link must be sized so that the bandwidth and latency requirement can be met during single failure conditions.

Interaction between Remote Copy and FlashCopy

Table 10-5 lists which combinations of FlashCopy and Remote Copy are supported.

Table 10-5 FlashCopy and Remote Copy interaction

Component	Remote Copy primary	Remote Copy secondary
FlashCopy source	Supported	Supported. When the FlashCopy relationship is in the Preparing and Prepared states, the cache at the Remote Copy secondary site operates in write-through mode. This process adds more latency to the already latent Remote Copy relationship.
FlashCopy target	This combination is supported and has the following restrictions: <ul style="list-style-type: none"> ► Running stop -force might cause the Remote Copy relationship to fully resynchronize. ► The I/O group must be the same. 	This combination is supported by the restriction that the FlashCopy mapping cannot be copying, stopping, or suspended. Otherwise, the restrictions are the same as at the Remote Copy primary site.

If you are not using Global Mirror with Change Volumes, for disaster recovery purposes, you can use the FlashCopy feature to create a consistent copy of an image before you restart a Global Mirror relationship.

When a consistent relationship is stopped, the relationship enters the consistent_stopped state. While in this state, I/O operations at the primary site continue to run. However, updates are not copied to the secondary site. When the relationship is restarted, the synchronization process for new data is started. During this process, the relationship is in the inconsistent_copying state.

The secondary volume for the relationship cannot be used until the copy process completes and the relationship returns to the consistent state. When this situation occurs, start a FlashCopy operation for the secondary volume before you restart the relationship. While the relationship is in the Copying state, the FlashCopy feature can provide a consistent copy of the data. If the relationship does not reach the synchronized state, you can use the FlashCopy target volume at the secondary site.

10.3 Troubleshooting Remote Copy

Remote Copy (Global Mirror and Metro Mirror) has the following primary error codes:

- ▶ 1920: This error is a congestion error that means that the source, the link between source and target, or the target cannot keep up with the rate of demand.
- ▶ 1720: This error is a heartbeat or cluster partnership communication error. This error tends to be more serious because failing communication between your cluster partners involves some extended diagnostic time.

10.3.1 1920 error

A 1920 error (event ID 050010) can have several triggers. The following official probable cause projections are available:

- ▶ Primary cluster or SAN fabric problem (10%)
- ▶ Primary cluster or SAN fabric configuration (10%)
- ▶ Secondary cluster or SAN fabric problem (15%)
- ▶ Secondary cluster or SAN fabric configuration (25%)
- ▶ Inter-cluster link problem (15%)
- ▶ Inter-cluster link configuration (25%)

In practice, the error that is most often overlooked is latency. Global Mirror has a round-trip-time tolerance limit of 80 ms. A message that is sent from your source cluster to your target cluster and the accompanying acknowledgement must have a total time of 80 ms (or 40 ms each way).

The primary component of your round-trip time is the physical distance between sites. For every 1,000 km (621.36 miles), there is a 5 ms delay. This delay does not include the time that is added by equipment in the path. Every device adds a varying amount of time depending on the device, but you can expect about 25 μ s for pure hardware devices. For software-based functions (such as, compression that is implemented in software), the added delay tends to be much higher (usually in the millisecond-plus range).

Consider this example. Company A has a production site that is 1,900 km from their recovery site. Their network service provider uses five devices to connect the two sites. In addition to those devices, Company A uses a SAN Fibre Channel Router at each site to provide FCIP to encapsulate the Fibre Channel traffic between sites. There are now seven devices, and 1,900 km of distance delay. All the devices add 200 μ s of delay each way. The distance adds 9.5 ms each way, for a total of 19 ms. Combined with the device latency that is 19.4 ms of physical latency at a minimum. This latency is under the 80 ms limit of Global Mirror, but this number is the best case number. Link quality and bandwidth play a significant role here. Your network provider likely ensures a latency maximum on your network link; be sure to stay below the Global Mirror RTT (Round Trip Time) limit. You can easily double or triple the expected physical latency with a lower quality or lower bandwidth network link. As a result, you are suddenly within range of exceeding the limit the moment a large flood of I/O happens that exceeds the bandwidth capacity you have in place.

When you get a 1920 error, always check the latency first. The FCIP routing layer can introduce latency if it is not properly configured. If your network provider reports a much lower latency, this report can be an indication of a problem at your FCIP Routing layer. Most FCIP Routing devices have built-in tools that you can use to check the round-trip delay time (RTT). When you are checking latency, remember that TCP/IP routing devices (including FCIP routers) report RTT by using standard 64-byte ping packets.

Figure 10-85 shows why the effective transit time should be measured only by using packets large enough to hold a Fibre Channel frame. This packet size is 2148 bytes (2112 bytes of payload and 36 bytes of header) and you should allow more capacity to be safe because different switching vendors have optional features that might increase this size.

Packet Size	Link Size	Serialization Delay (Time Required to Send Data)	Unit
64	256 Kbps	2.0E+03	microseconds
64	1.5 Mbps	3.4E+02	microseconds
64	100 Mbps	5.1E+00	microseconds
64	155 Mbps	3.3E+00	microseconds
64	622 Mbps	8.2E-01	microseconds
64	1 Gbps	5.1E-04	microseconds
64	10 Gbps	5.1E-05	microseconds
1500	256 Kbps	4.7E+04	microseconds
1500	1.5 Mbps	8.0E+03	microseconds
1500	100 Mbps	1.2E+02	microseconds
1500	155 Mbps	7.7E+01	microseconds
1500	622 Mbps	1.9E+01	microseconds
1500	1 Gbps	1.2E+01	microseconds
1500	10 Gbps	1.2E+00	microseconds
2148	256 Kbps	6.7E+04	microseconds
2148	1.5 Mbps	1.1E+04	microseconds
2148	100 Mbps	1.7E+02	microseconds
2148	155 Mbps	1.1E+02	microseconds
2148	622 Mbps	2.8E+01	microseconds
2148	1 Gbps	1.7E+01	microseconds
2148	10 Gbps	1.7E+03	microseconds

Figure 10-85 Effect of packet size (in bytes) versus link size

Before you proceed, take a quick look at the second largest component of your round-trip-time; that is, serialization delay. Serialization delay is the amount of time that is required to move a packet of data of a specific size across a network link of a bandwidth. This delay is based on a simple concept that the time that is required to move a specific amount of data decreases as the data transmission rate increases.

In Figure 10-85, there are orders of magnitude of difference between the different link bandwidths. It is easy to see how 1920 errors can arise when your bandwidth is insufficient and why you should never use a TCP/IP ping to measure RTT for FCIP traffic.

Figure 10-85 compares the amount of time in microseconds that is required to transmit a packet across network links of varying bandwidth capacity. The following packet sizes are used:

- ▶ 64 bytes: The size of the common ping packet
- ▶ 1500 bytes: The size of the standard TCP/IP packet
- ▶ 2148 bytes: The size of a Fibre Channel frame

Your path MTU affects the delay that is incurred in getting a packet from one location to another, when it causes fragmentation, or is too large and causes too many retransmits when a packet is lost. After you verified your latency by using the correct packet size, proceed with normal hardware troubleshooting.

10.3.2 1720 error

The 1720 error (event ID 050020) is the other primary error code of Remote Copy. Because the term “System Partnership” implies that all involved virtualization systems are partners, they must communicate with each other. When a partner on either side stops communicating, you see a 1720 error appear in your error log. According to official documentation, there are no likely field replaceable unit breakages or other causes.

In practice, the source of this error is most often a fabric problem or a problem in the network path between your partners. When you receive this error, if your fabric has more than 64 HBA zoned ports, you should check your fabric configuration for zoning of more than one HBA port for each node per I/O group. One port for each node per I/O group that is associated with the host is the recommended zoning configuration for fabrics. For those fabrics with 64 or more host ports, this recommendation becomes a rule. You must follow this zoning rule or the configuration is technically unsupported.

Improper zoning leads to SAN congestion, which can inhibit remote link communication intermittently. Checking the zero buffer credit timer through IBM Tivoli Storage Productivity Center and comparing its value against your sample interval might reveal potential SAN congestion. When a zero buffer credit timer is above 2% of the total time of the sample interval, it is likely to cause problems.

Next, always ask your network provider to check the status of the link. If the link is okay, watch for repetition of this error. It is possible in a normal and functional network setup to have occasional 1720 errors, but multiple occurrences indicate a larger problem.

If you receive multiple 1720 errors, recheck your network connection and then check the IBM Storwize V5000 partnership information to verify their status and settings. Perform diagnostic tests for every piece of equipment in the path between the two systems. It often helps to have a diagram that shows the path of your replication from logical and physical configuration viewpoints.

If your investigation fails to resolve your Remote Copy problems, you should contact your IBM support representative for a complete analysis.

10.4 Managing Remote Copy by using the GUI

The IBM Storwize V5000 storage system provides a separate function icon for copy service management. The following windows are available for managing Remote Copy, which are accessed through the Copy Services function icon:

- ▶ Remote Copy
- ▶ Partnerships

As the name implies, these windows are used to manage Remote Copy and the partnership.

10.4.1 Managing cluster partnerships

The Partnership window is used to manage a partnership between clusters. To access the Partnership window, click the **Copy Services** function icon and then click **Partnerships**, as shown in Figure 10-86.

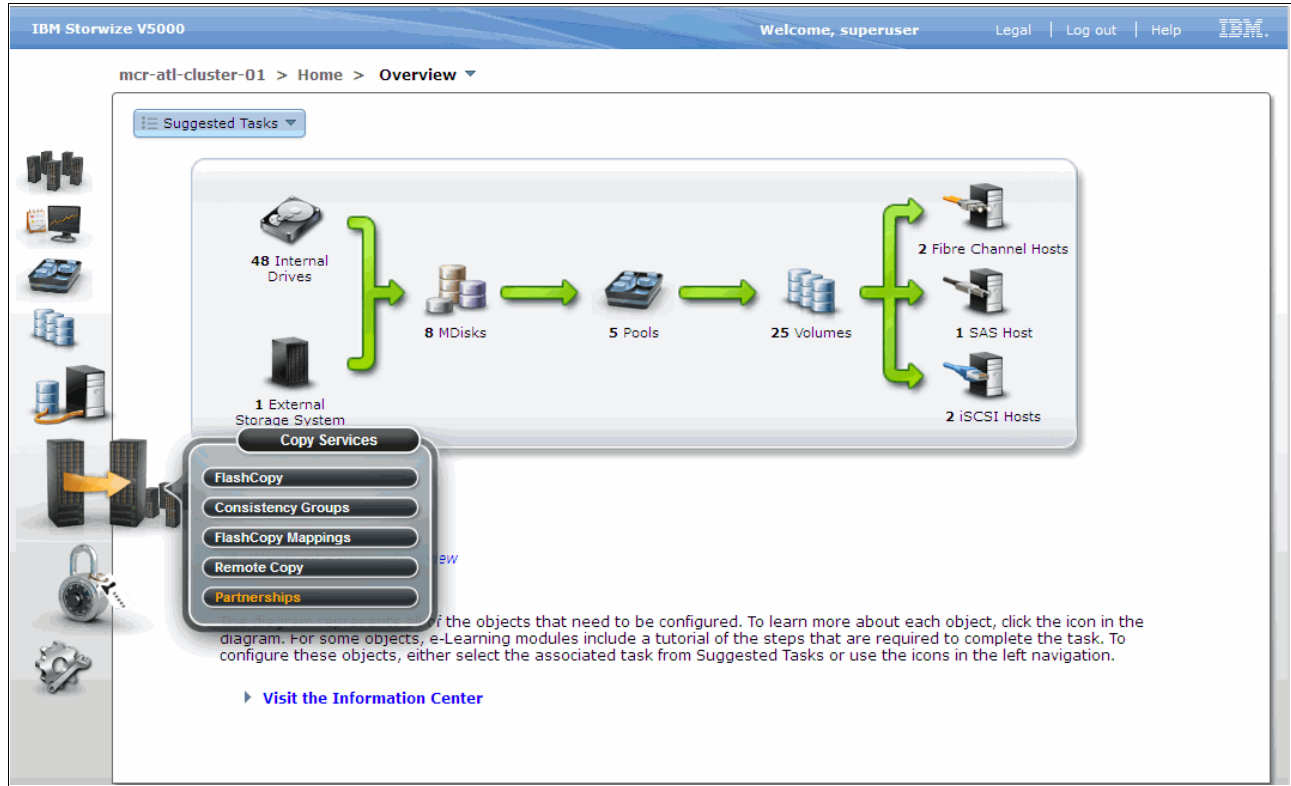


Figure 10-86 Partnership window

Creating a partnership

No partnership is defined in our example (see Figure 10-87), so you must create a partnership between the IBM Storwize V5000 systems. Click **New Partnership** in the Partnership window.

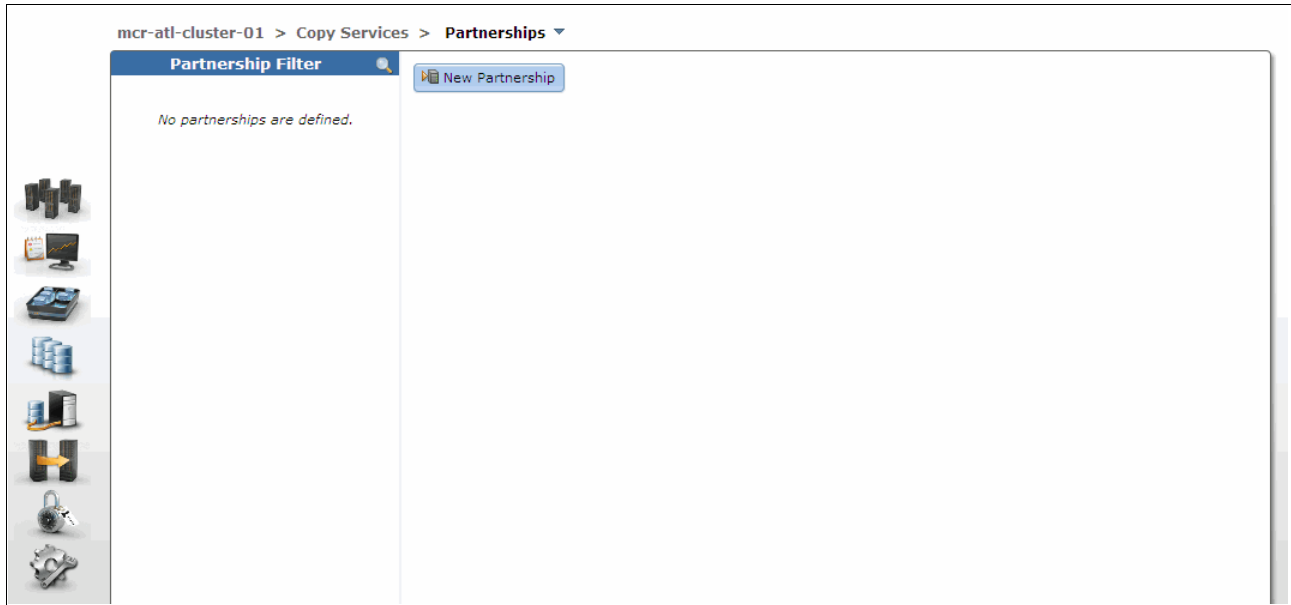


Figure 10-87 Create a cluster partnership

If there is no partnership candidate, an error window opens, as shown in Figure 10-88.

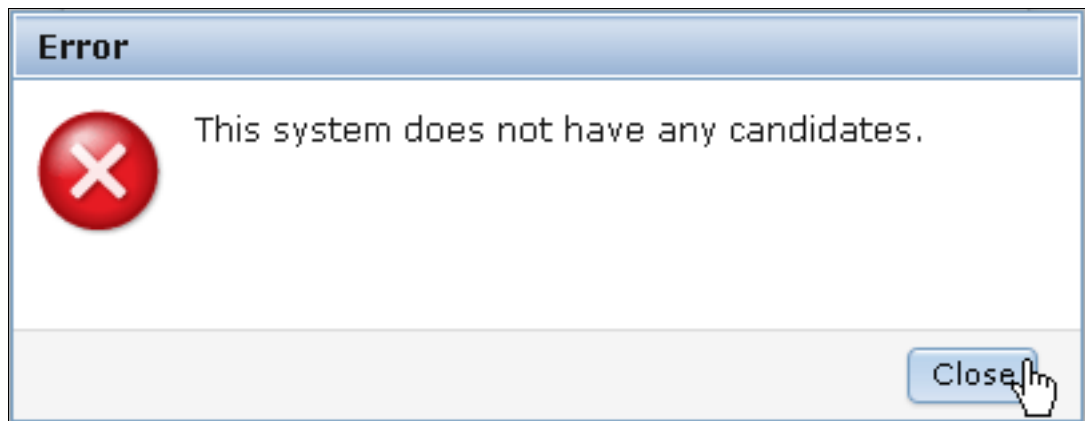


Figure 10-88 No candidates are available to create a partnership

Check the zoning and the system status and make sure that the clusters can see each other. Then, you can create your partnership by selecting the appropriate remote storage system (as shown in Figure 10-89 on page 520), and defining the available bandwidth between both systems.

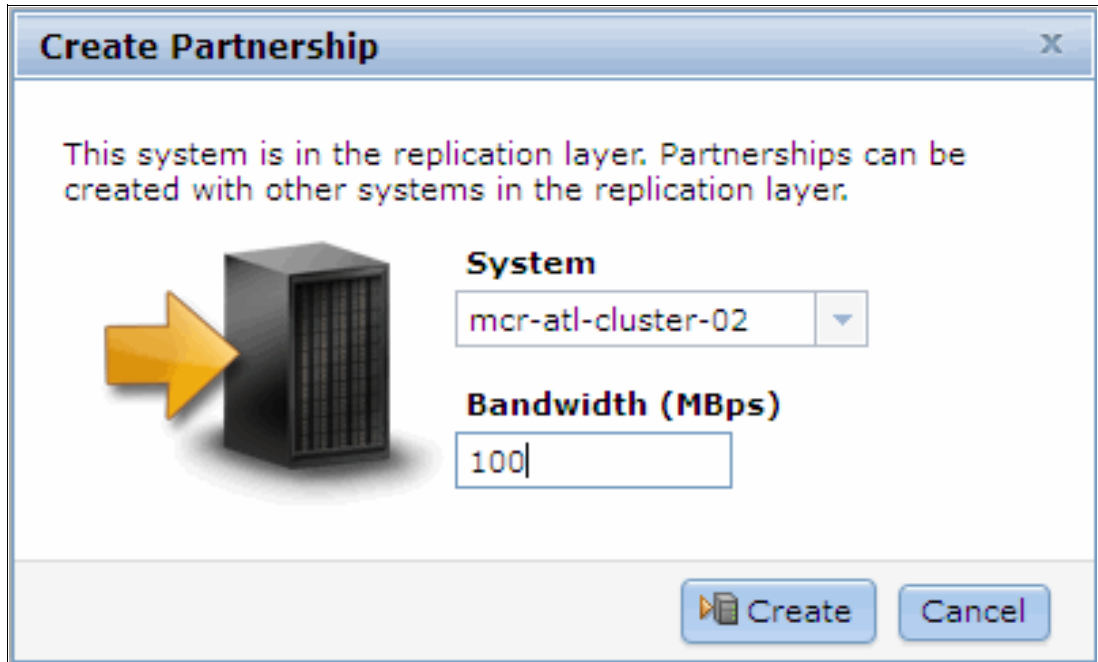


Figure 10-89 Select the remote IBM Storwize storage system for a new partnership

The bandwidth that you must enter here is used by the background copy process between the clusters in the partnership. To set the background copy bandwidth optimally, make sure that you consider all three resources (primary storage, inter-cluster link bandwidth, and auxiliary storage) to avoid overloading them, which affects the foreground I/O latency.

Click **Create** and the partnership definition is complete on the first IBM Storwize V5000 system. You can find the partnership that is listed in the left pane of the Partnership window. If you select the partnership, more information for this partnership is displayed on the right, as shown in Figure 10-90.

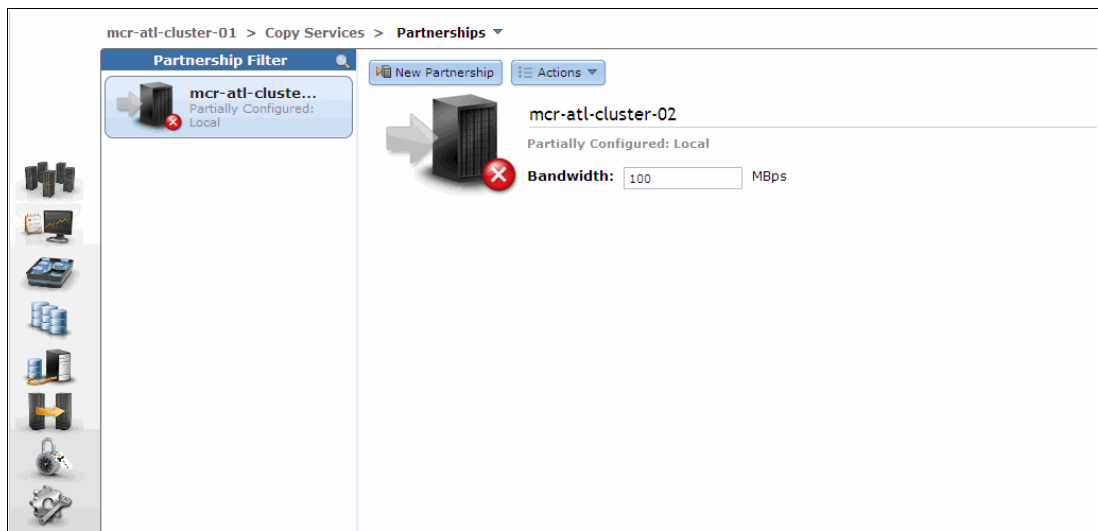


Figure 10-90 Partially configured partnership

Important: The partnership is in the “Partially Configured: Local” state because we did not yet define it on the other IBM Storwize V5000. For more information about partnership states, see “Remote Copy and consistency group states” on page 511.

Complete the same steps on the second storage system for the partnership to become fully configured. The Remote Copy partnership is now implemented between the two IBM Storwize V5000 systems and both systems are ready for further configuration of Remote Copy relationships, as shown in Figure 10-91.

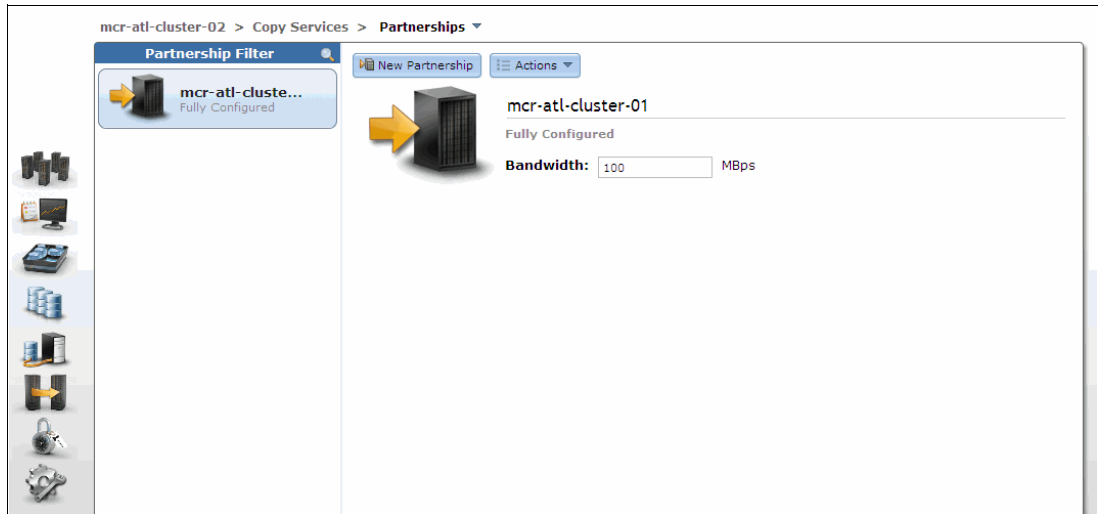


Figure 10-91 Fully configured partnership

You can also change the bandwidth setting for the partnership in the Partnerships window. Click **Apply Changes** to confirm your modification.

Stopping and starting a partnership

You can stop the partnership by clicking **Stop Partnership** from the Actions drop-down menu, as shown in Figure 10-92. If you stop the partnership, the relationship that uses this partnership is disconnected.

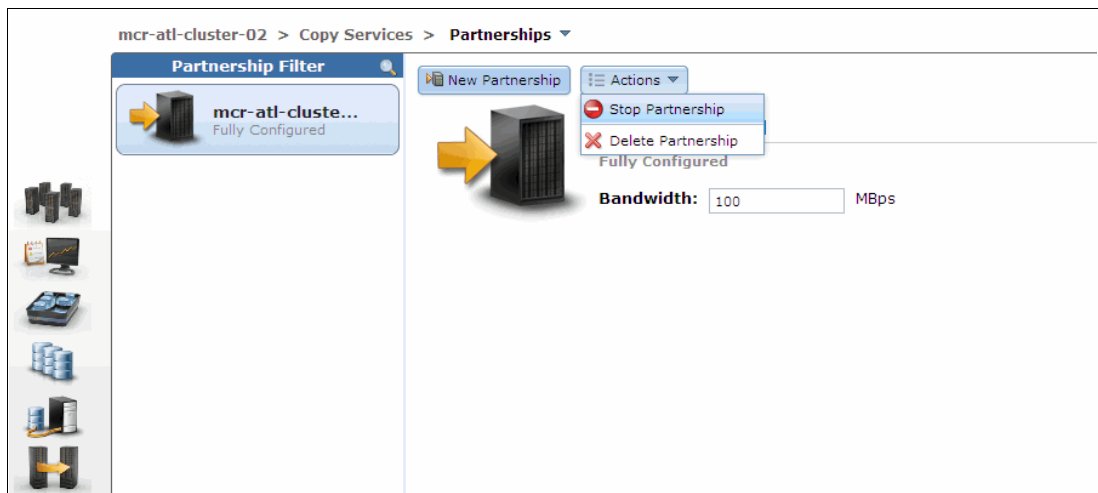


Figure 10-92 Stop the partnership

After you stop the partnership, your partnership is listed as Fully Configured: Stopped, as shown in Figure 10-93.

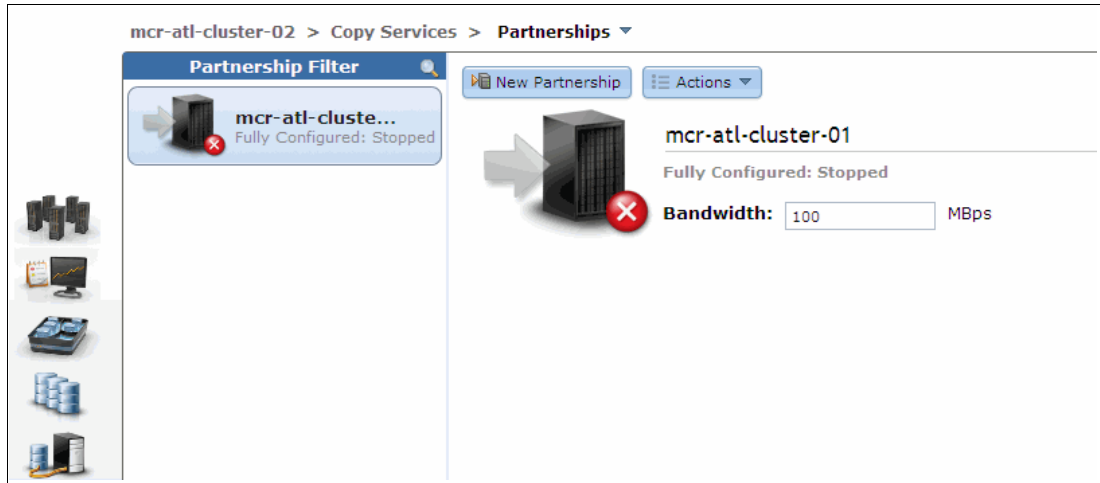


Figure 10-93 Fully configured partnership in Stopped state

You can restart a stopped partnership by clicking **Start Partnership** from the Actions drop-down menu.

The partnership returns to the fully configured status when it is restarted.

Deleting a partnership

You can delete a partnership by selecting **Delete Partnership** from the Actions drop-down menu, as shown in Figure 10-92 on page 521.

10.4.2 Managing stand-alone Remote Copy relationships

A Remote Copy relationship can be defined between two volumes where one is the master (source) and the other one is the auxiliary (target) volume. Use of Remote Copy auxiliary volumes as Remote Copy master volumes is not allowed. Open the Remote Copy window to manage Remote Copy by clicking the **Copy Services** icon and then clicking **Remote Copy**, as shown in Figure 10-94 on page 523.

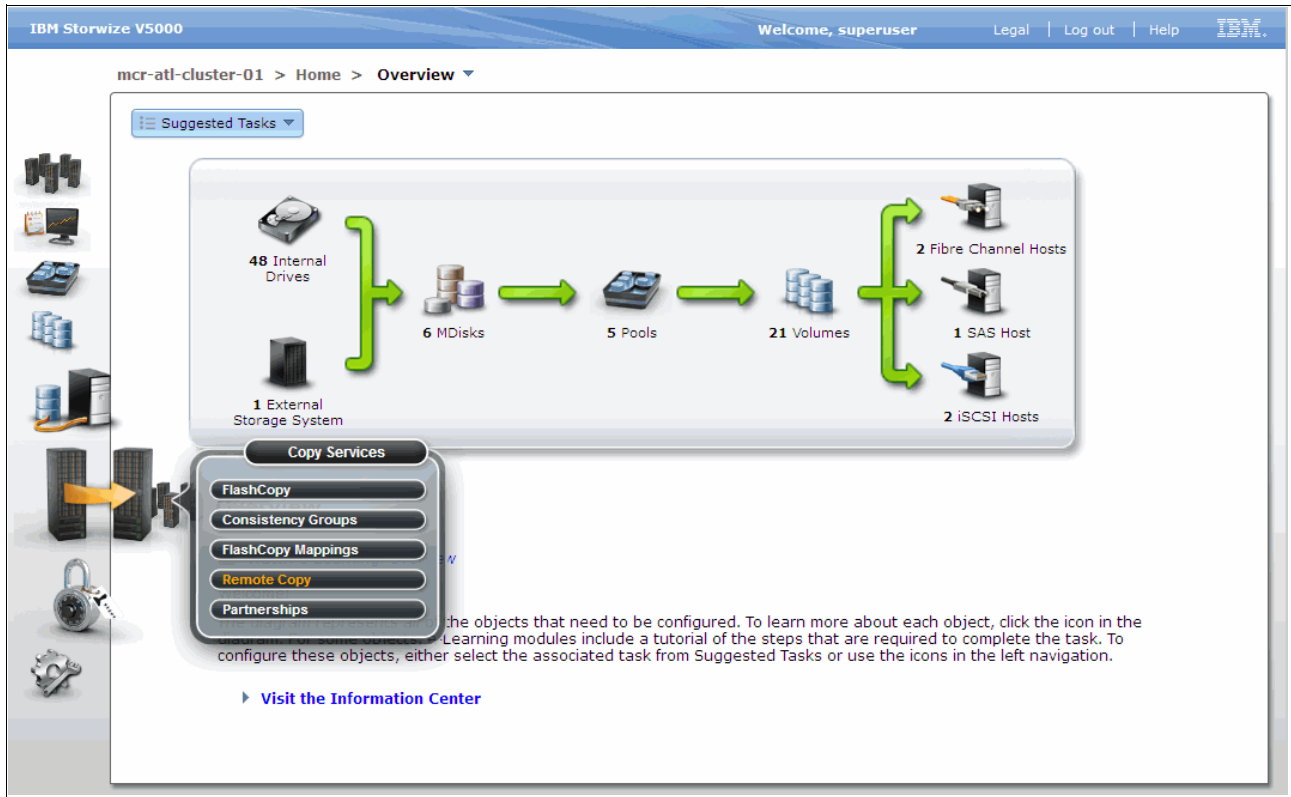


Figure 10-94 Open Remote Copy window

As shown in Figure 10-95, the Remote Copy window is where you can manage Remote Copy relationships and Remote Copy consistency groups.

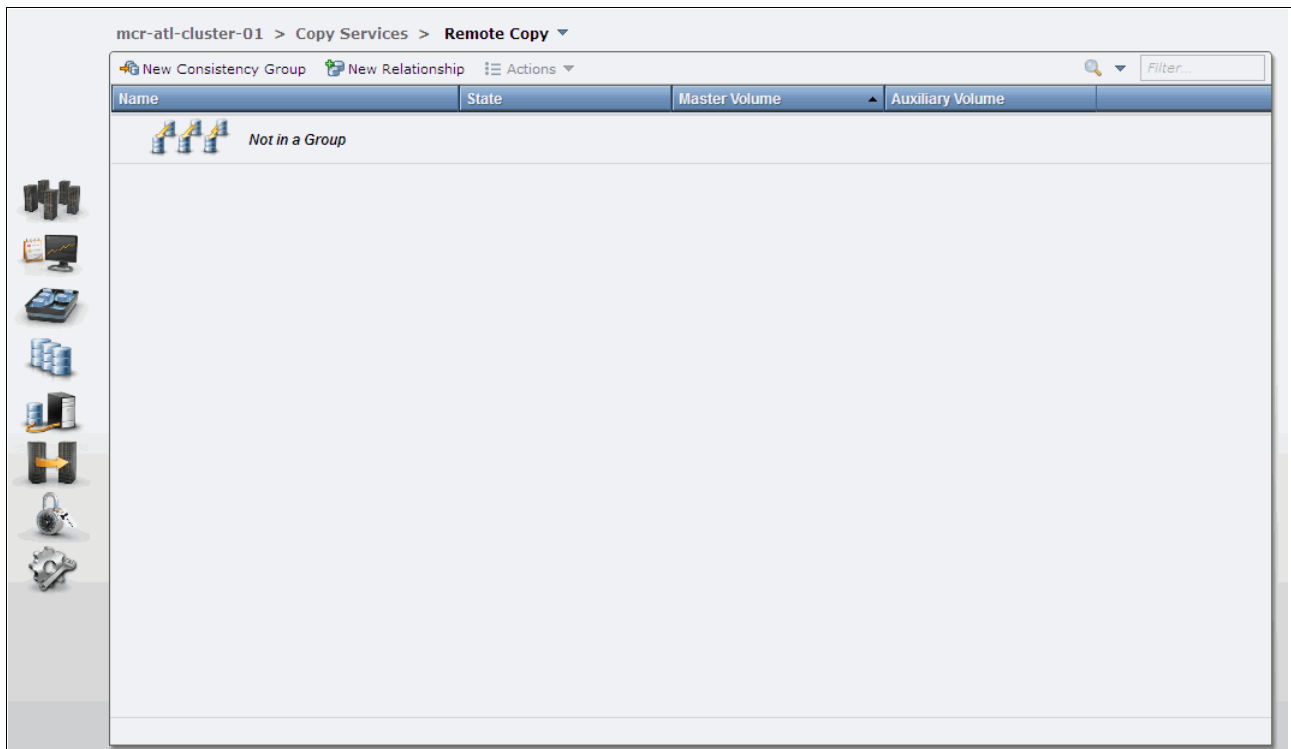


Figure 10-95 Remote Copy window

The Remote Copy window displays a list of Remote Copy consistency groups. You can also take actions on the Remote Copy relationships and Remote Copy consistency groups. Click **Not in a Group** and all the Remote Copy relationships that are not in any Remote Copy consistency groups are displayed. To customize the blue column heading bar and select different attributes of Remote copy relationships, right-click anywhere in the blue bar.

Creating stand-alone Remote Copy relationships

Important: Before a remote copy relationship is created, target volumes that are the same size as the source volumes that you want to mirror must be created. For more information about creating volumes, see Chapter 5, “I/O Group basic volume configuration” on page 161.

To create a Remote Copy relationship, click **New Relationship** at the top of the Remote Copy window, as shown in Figure 10-95 on page 523. A wizard opens and guides you through the Remote Copy relationship creation process.

As shown in Figure 10-96, you must set the Remote Copy relationship type first. Based on your requirements, you can select Metro Mirror (synchronous replication) or Global Mirror (asynchronous replication). Select the appropriate replication type and click **Next**.

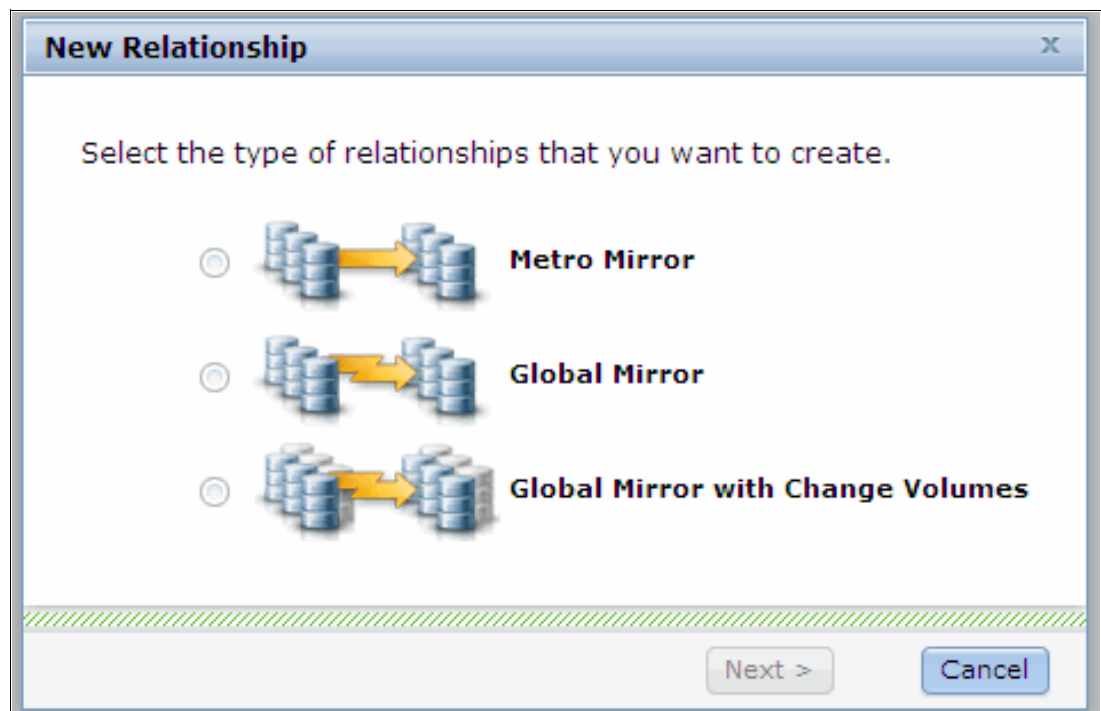


Figure 10-96 Select the appropriate Remote Copy type

You must select where your auxiliary (target) volumes are: the local system or the already defined second storage system. In our example (as shown in Figure 10-97), choose another system to build an inter-cluster relationship. Click **Next** to continue.

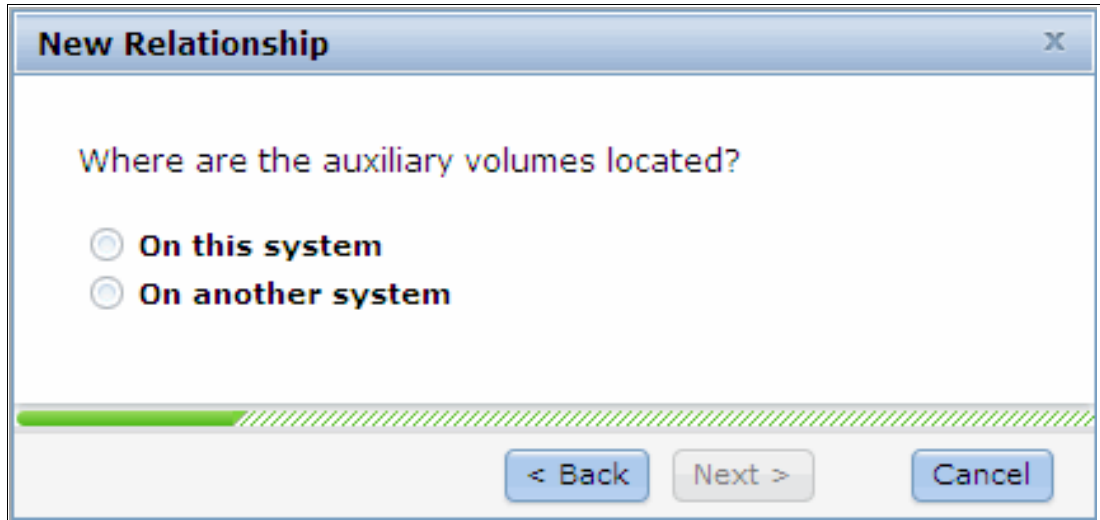


Figure 10-97 Select Remote Copy partner

The Remote Copy master and auxiliary volume must be specified. Both volumes must have the same size. As shown in Figure 10-98, the system offers only appropriate auxiliary candidates with the same volume size as the selected master volume. After you select the volumes that are based on your requirement, click **Add**.

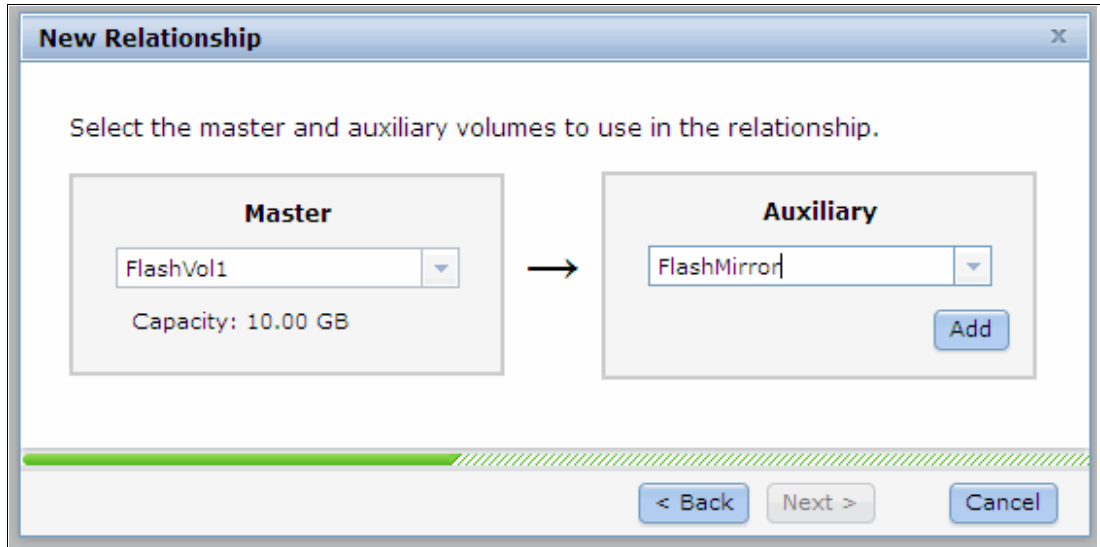


Figure 10-98 Select the master and auxiliary volume

You can define multiple and independent relationships by clicking **Add**. You can remove a relationship by clicking the red cross. In our example, we create two independent Remote Copy relationships, as shown in Figure 10-99.

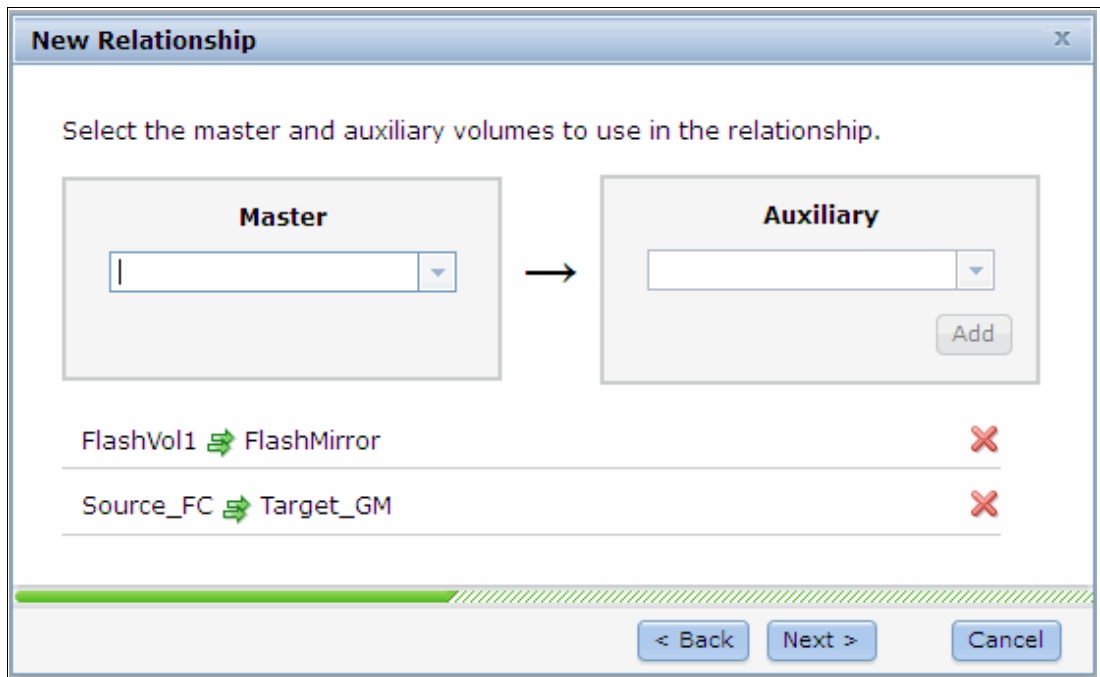


Figure 10-99 Define multiple independent relationships

A window opens and prompts you to select if the volumes in the relationship are already synchronized. In most situations, the data on the master volume and on the auxiliary volume are not identical, so click **No** and then click **Next** to enable an initial copy, as shown in Figure 10-100.

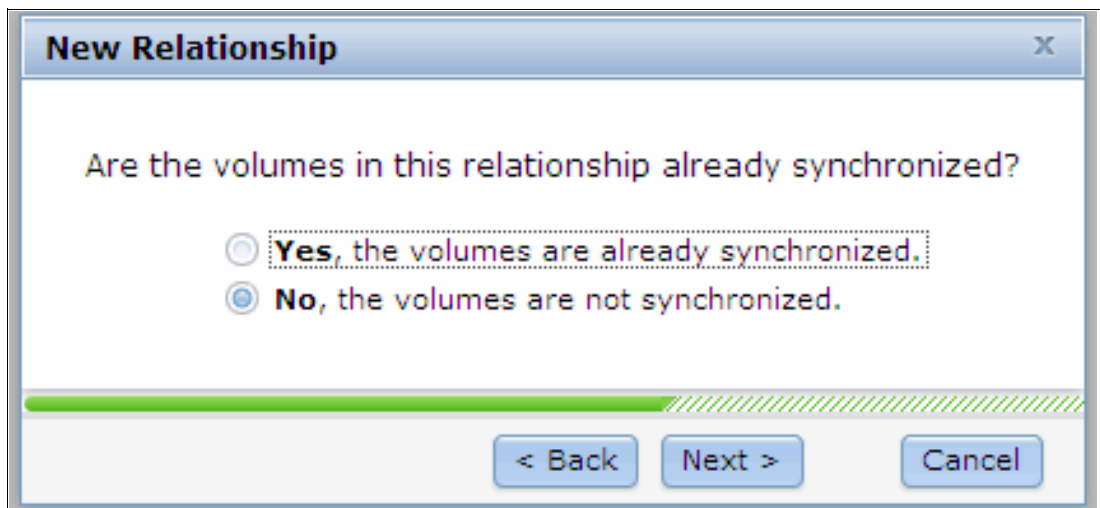


Figure 10-100 Activate initial data copy

If you select **Yes, the volumes are already synchronized** in this step, a warning message opens, as shown in Figure 10-101. Confirm that the volumes are truly identical, and then click **OK** to continue.

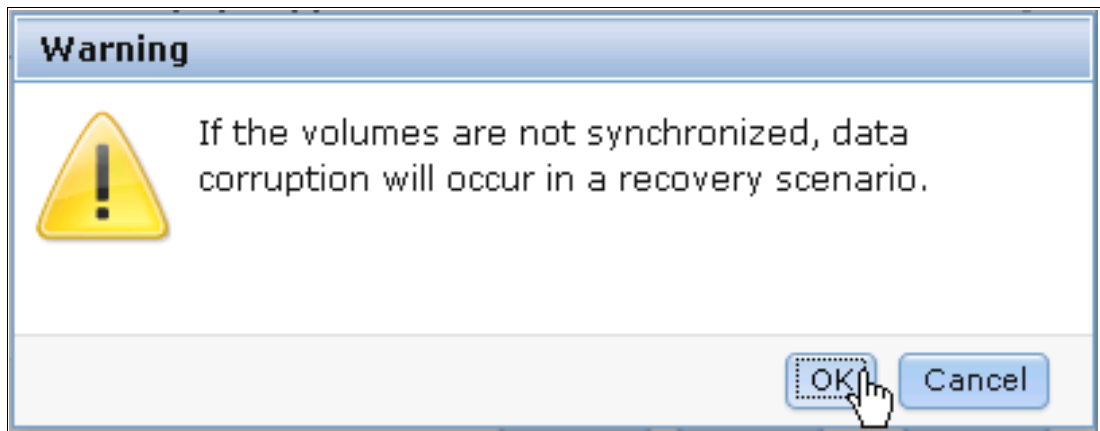


Figure 10-101 Warning message to make sure that the volumes are synchronized

You can choose to start the initial copying progress now or wait to start it later. In our example, select **Yes, start copying now** and click **Finish**, as shown in Figure 10-102.

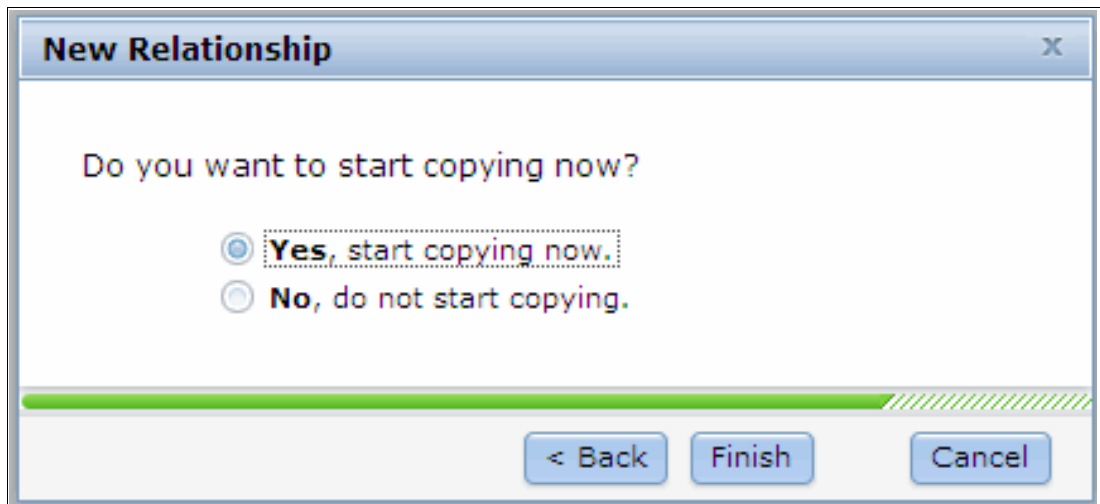


Figure 10-102 Choose if you want to start copying now or later

After the Remote Copy relationships creation completes, two independent Remote Copy relationships are defined and displayed in the Not in a Group list, as shown in Figure 10-103.

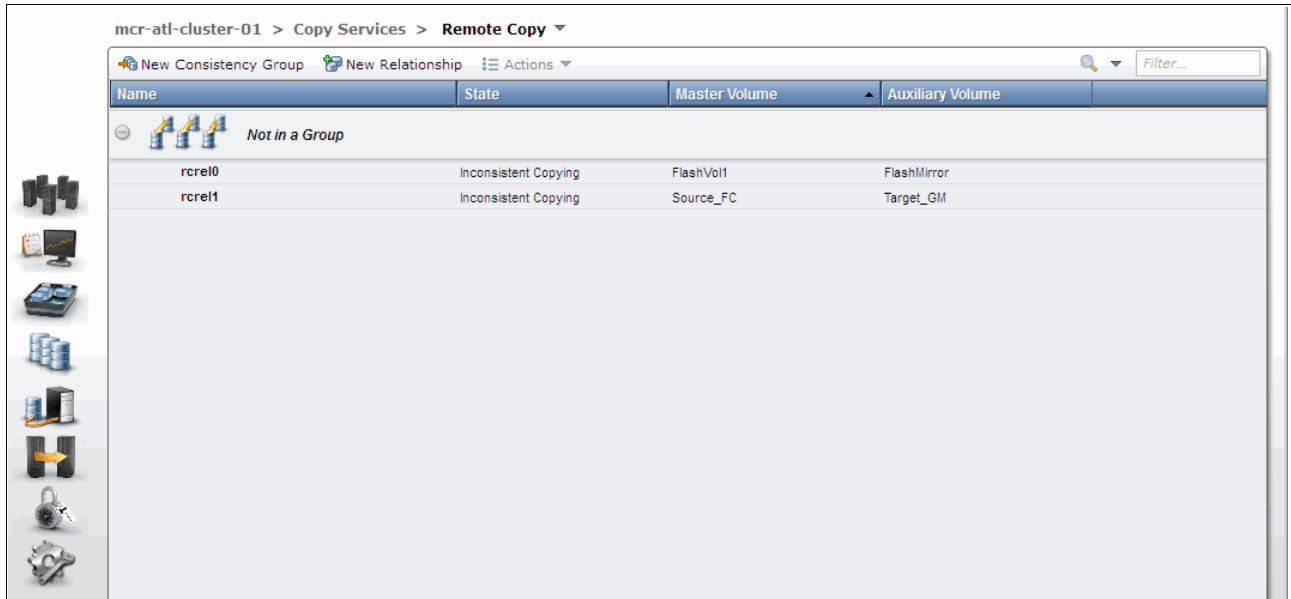


Figure 10-103 Creating a Remote Copy relationship process completes

Optionally, you can monitor the ongoing initial synchronization in the Running Tasks status indicator, as shown in Figure 10-104. Highlight one of the operations and click to see the progress.



Figure 10-104 Remote copy initialization progress through Running Tasks

Stopping a stand-alone Remote Copy relationship

The Remote Copy relationship can be stopped by selecting the relationship and clicking **Stop** from the Actions drop-down menu, as shown in Figure 10-105.

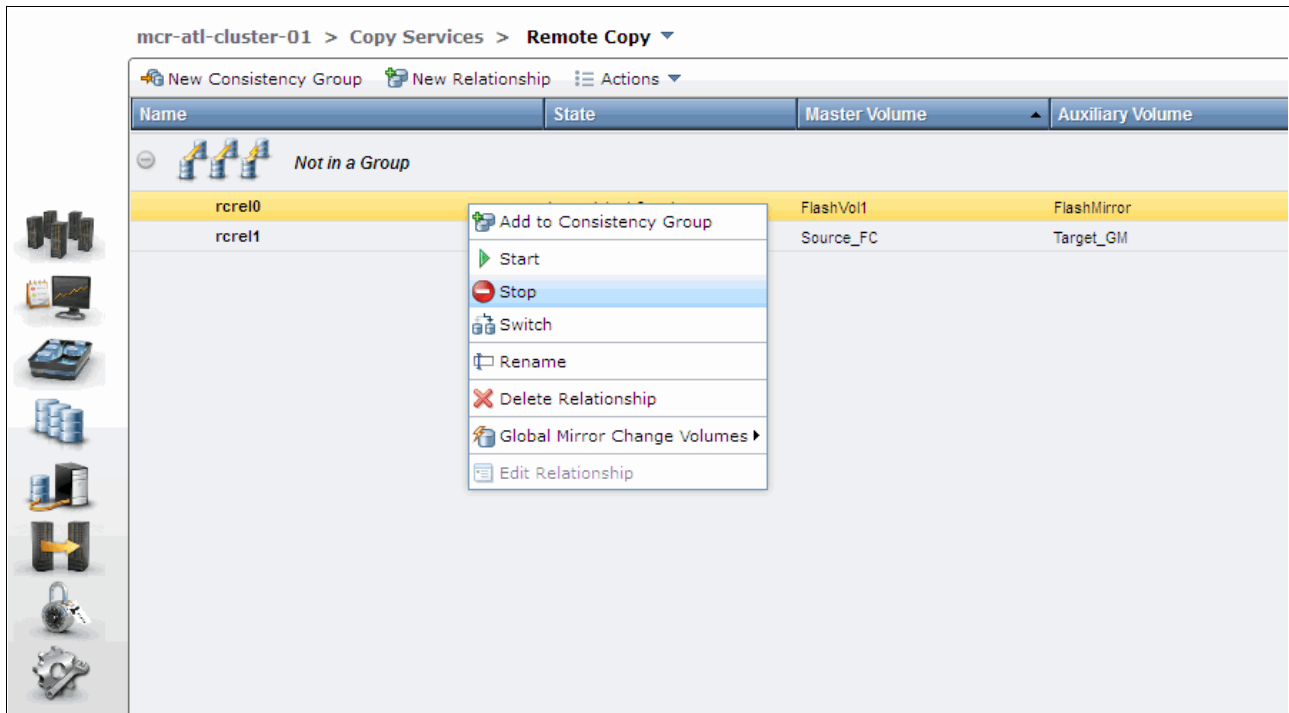


Figure 10-105 Stop Remote Copy relationship

A prompt appears. Click to allow secondary read/write access, if required, and then click **Stop Relationship**, as shown in Figure 10-106.

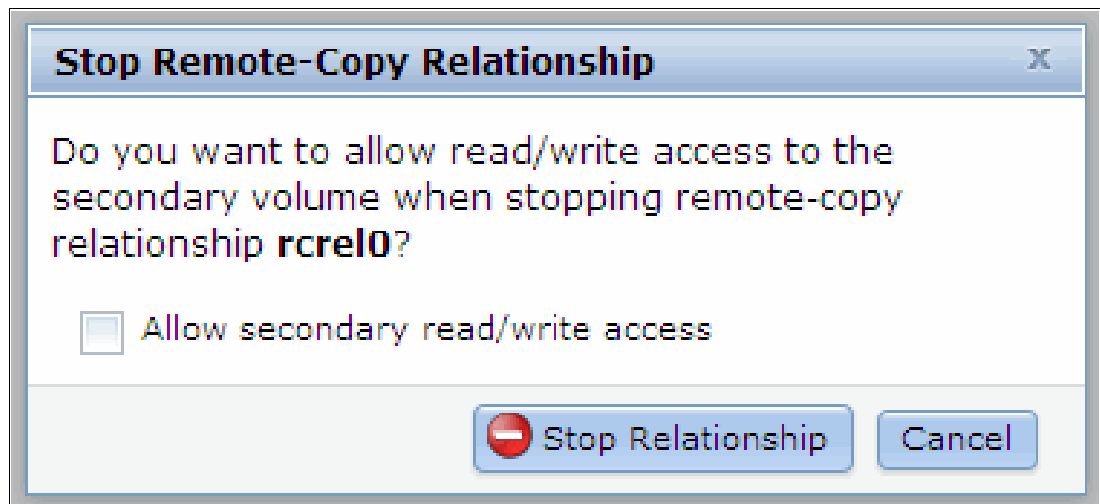


Figure 10-106 Option to allow secondary read/write access

After the stop completes, the state of the Remote Copy relationship is changed from Consistent Synchronized to Idling, as shown in Figure 10-107. Read/write access to both volumes is now allowed unless you selected otherwise.

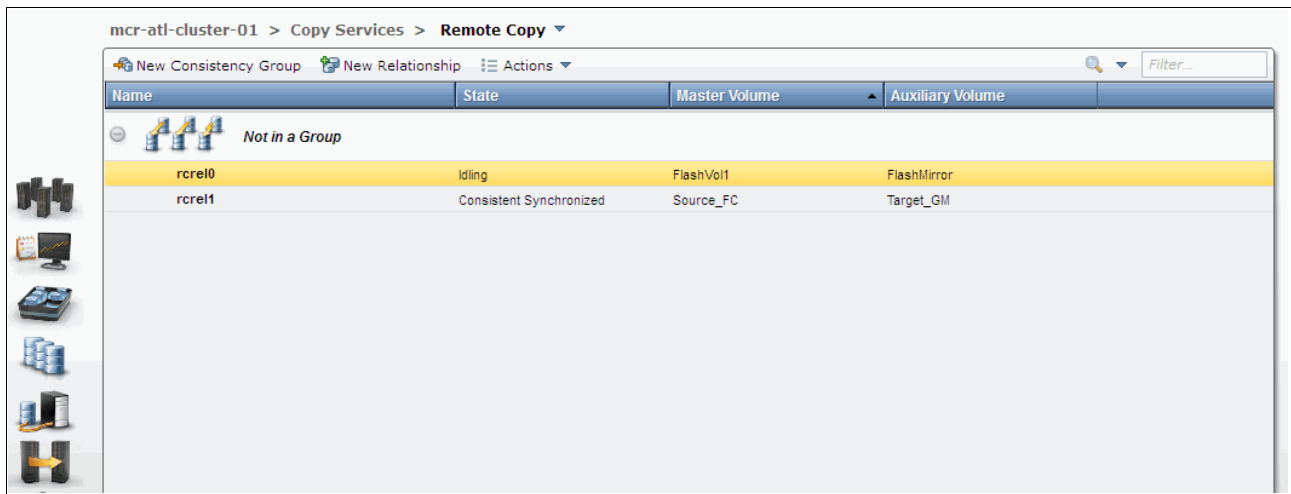


Figure 10-107 Remote Copy relationship stop completes

Starting a stand-alone Remote Copy relationship

The Remote Copy relationship can be started by selecting the relationship and clicking **Start** from the Actions drop-down menu, as shown in Figure 10-108.

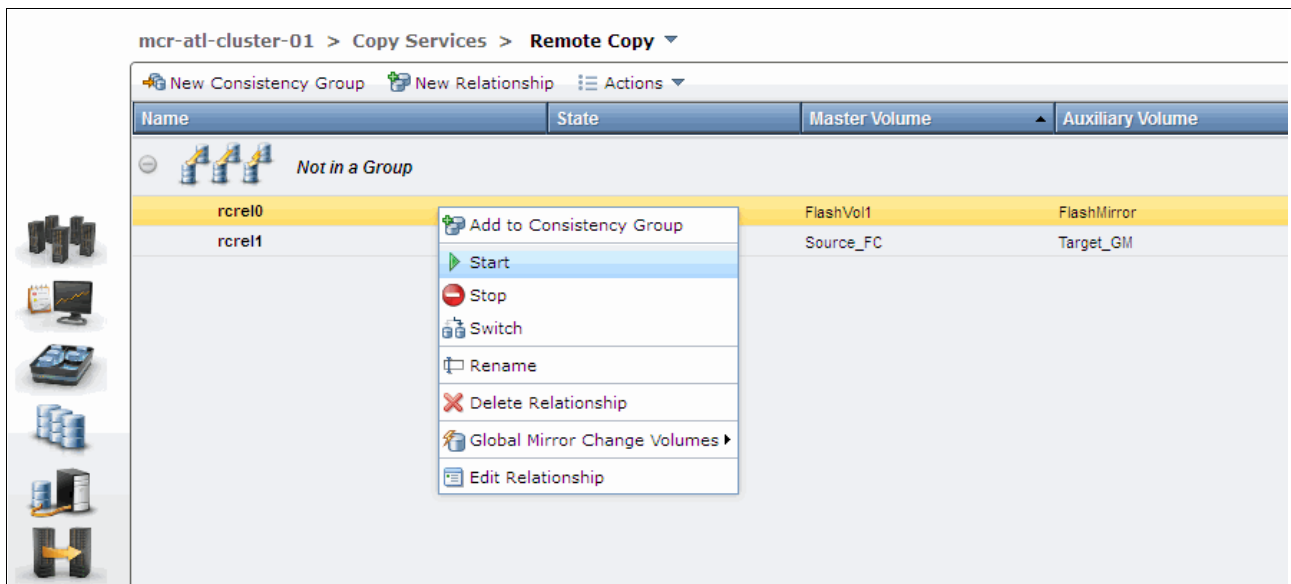


Figure 10-108 Start a Remote Copy relationship

When a Remote Copy relationship is started, the most important item is selecting the copy direction. Both master and auxiliary volumes can be the primary. Make your decision that is based on your requirements and click **Start Relationship**. In our example, choose the master volume to be the primary, as shown in Figure 10-109.

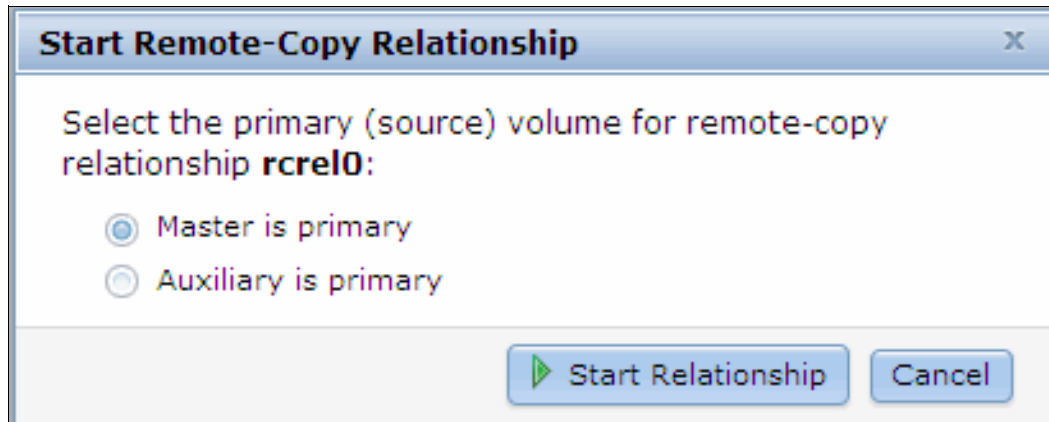


Figure 10-109 Choose the copy direction

Switching the direction of a stand-alone Remote Copy relationship

The copy direction of the Remote Copy relationship can be switched by selecting the relationship and clicking **Switch** from the Actions drop-down menu, as shown in Figure 10-110.

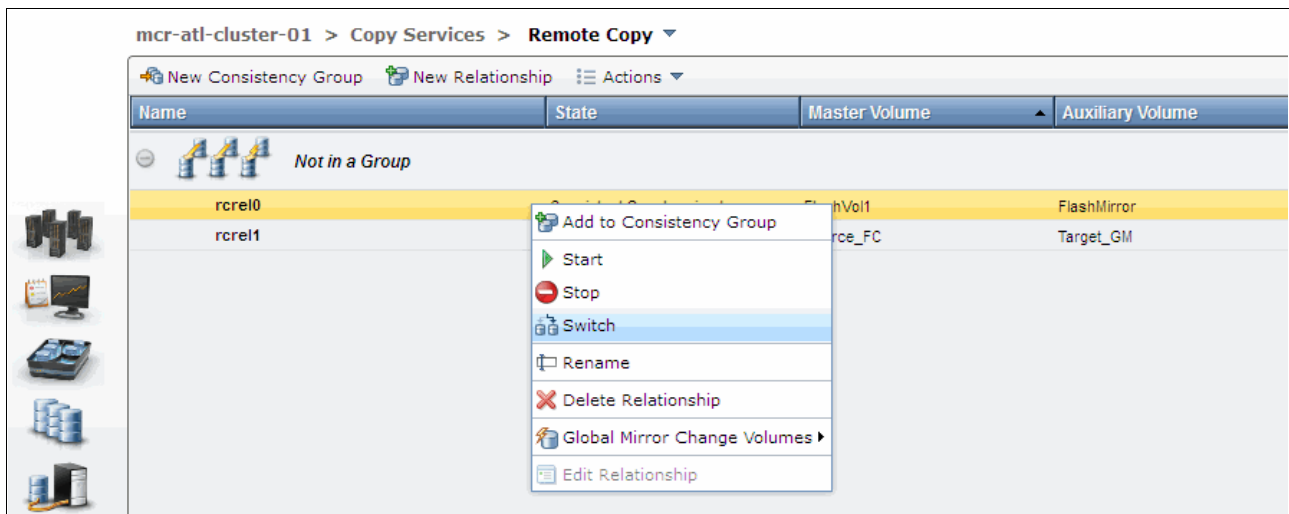


Figure 10-110 Switch Remote Copy relationship

A warning message opens and shows you the consequences of this action, as shown in Figure 10-111 on page 532. If you switch the Remote Copy relationship, the copy direction of the relationship becomes the opposite; that is, the current primary volume becomes the secondary while the current secondary volume becomes the primary. Write access to the current primary volume is lost and write access to the current secondary volume is enabled. If it is not a disaster recovery situation, you must stop your host I/O to the current primary volume in advance. Make sure that you are prepared for the consequences. If so, click **OK** to continue.

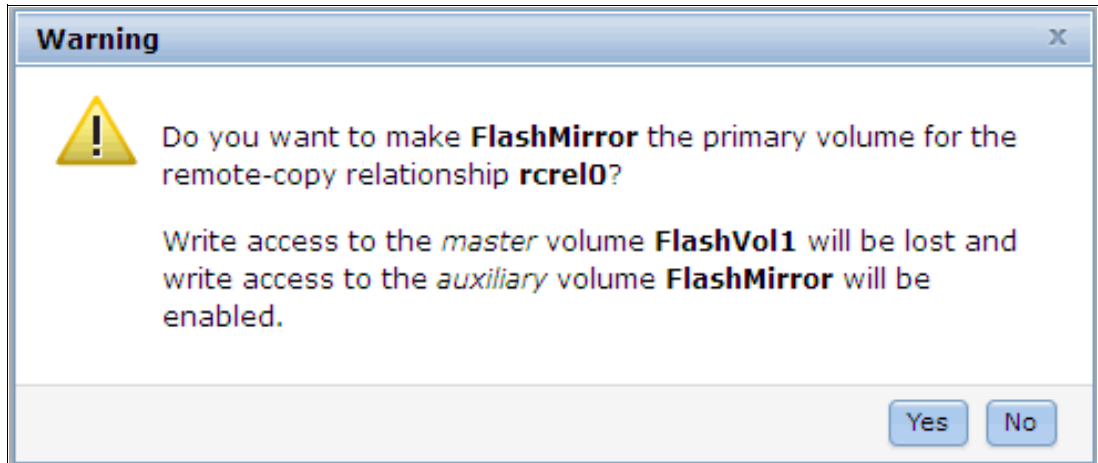


Figure 10-111 Warning message for switching direction of a Remote Copy relationship

After the switch completes, your Remote Copy relationship is tagged (as shown in Figure 10-112), and shows you that the primary volume in this relationship was changed.

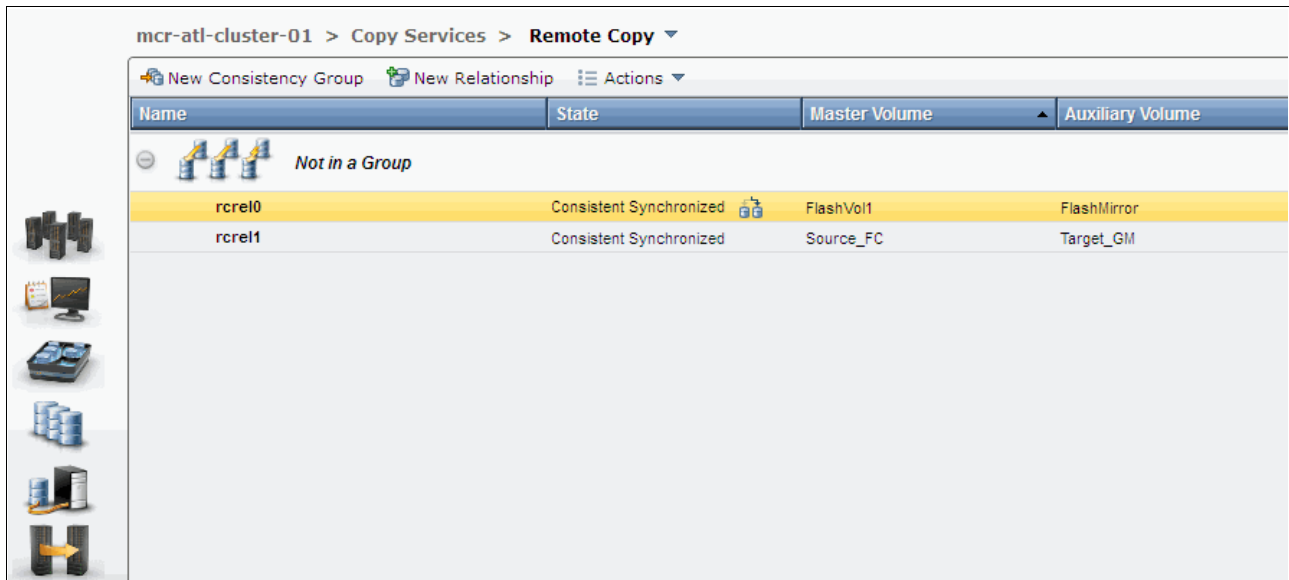


Figure 10-112 Switch icon on the state of the relationship

Renaming a stand-alone Remote Copy relationship

The Remote Copy relationship can be renamed by selecting the relationship and clicking **Rename** from the Actions drop-down menu, as shown in Figure 10-113 on page 533.

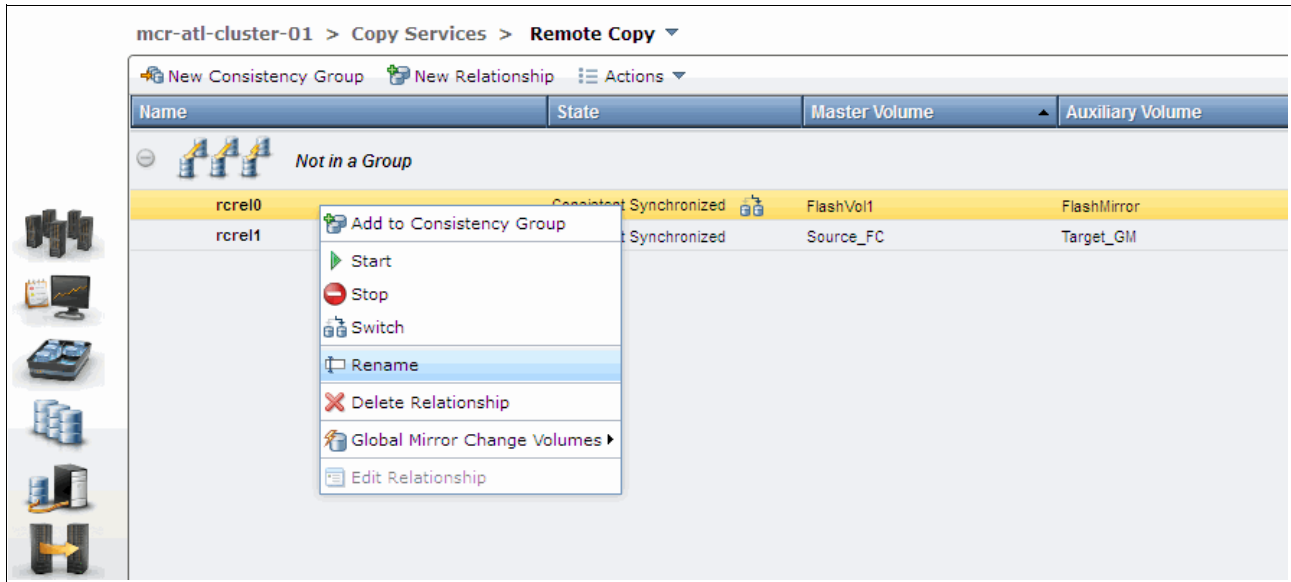


Figure 10-113 Rename the Remote Copy relationship

Enter the new name for the Remote Copy relationship and click **Rename**.

Deleting a stand-alone Remote Copy relationship

The Remote Copy relationship can be deleted by selecting the relationship and clicking **Delete Relationship** from the Actions drop-down menu, as shown in Figure 10-114.

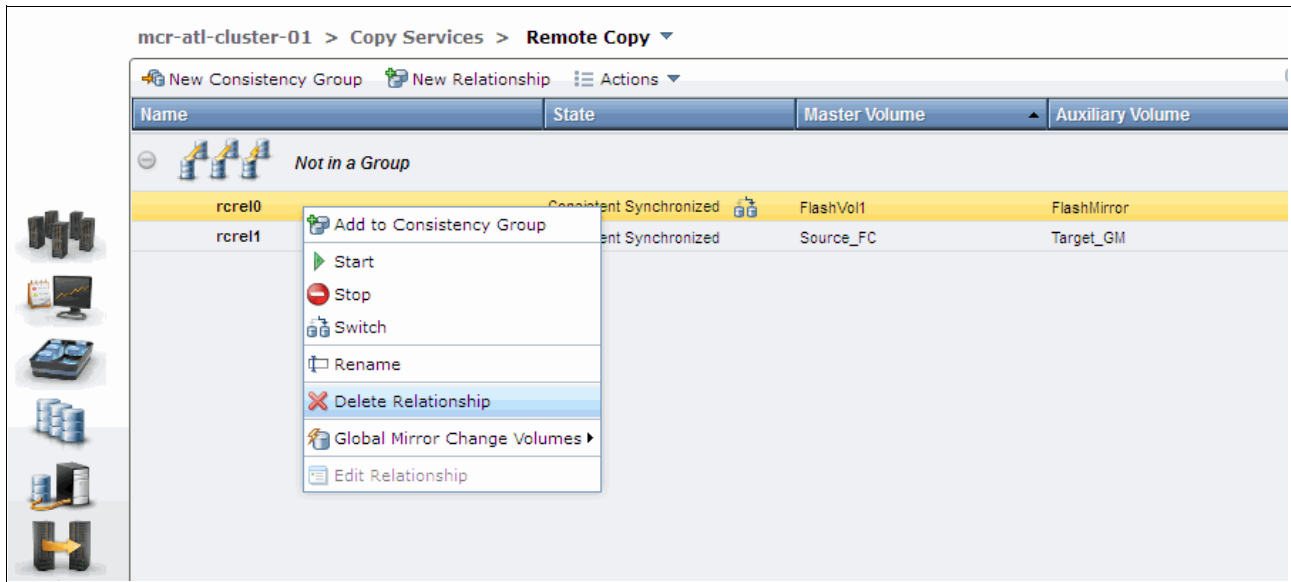


Figure 10-114 Delete a Remote Copy relationship

You must confirm this deletion by verifying the number of relationships to be deleted, as shown in Figure 10-115 on page 534. Click **Delete** to proceed.

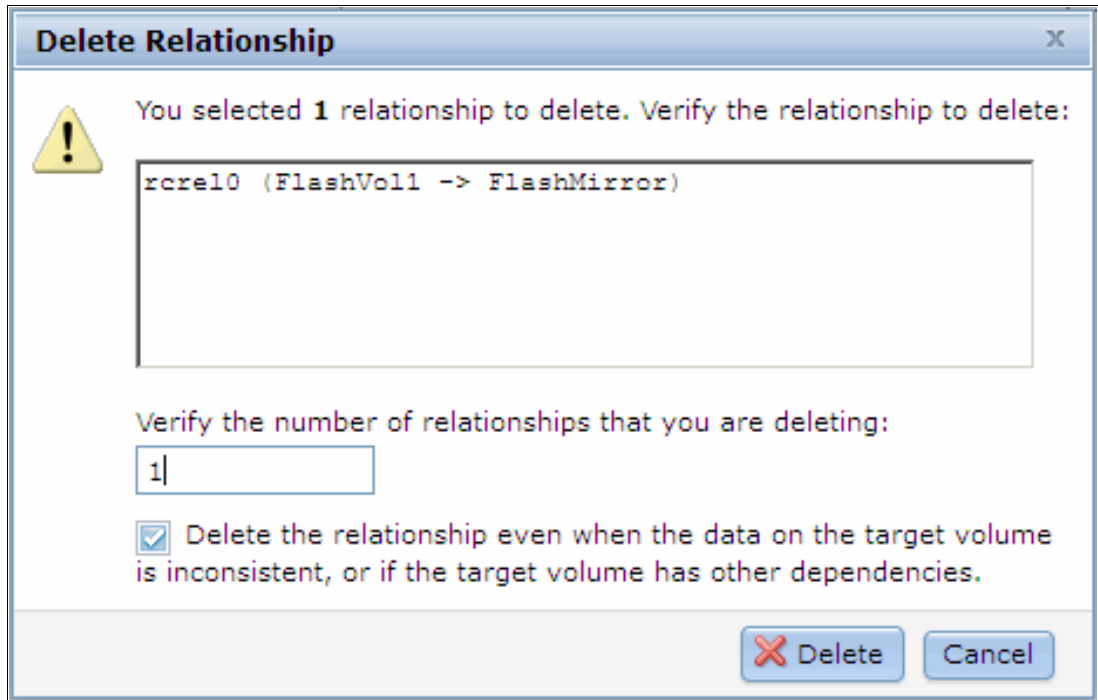


Figure 10-115 Confirm the relationship deletion

10.4.3 Managing a Remote Copy consistency group

A Remote Copy consistency group can be managed from the Remote Copy window as well.

Creating a Remote Copy consistency group

To create a Remote Copy consistency group, click **New Consistency Group**, as shown in Figure 10-116.

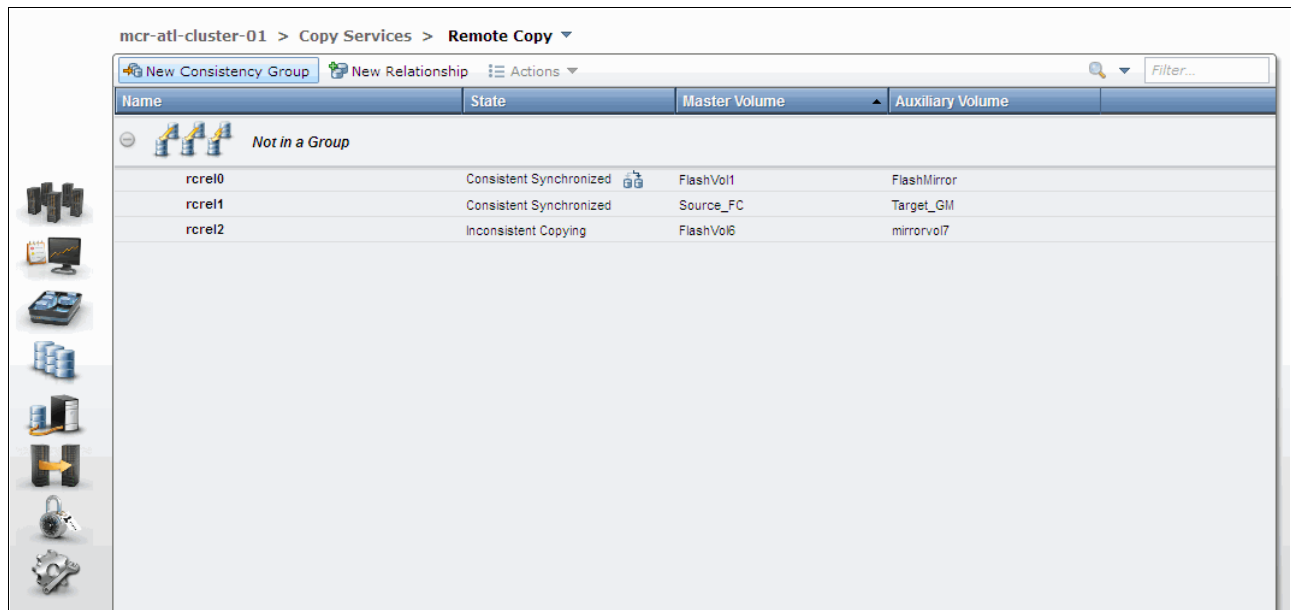
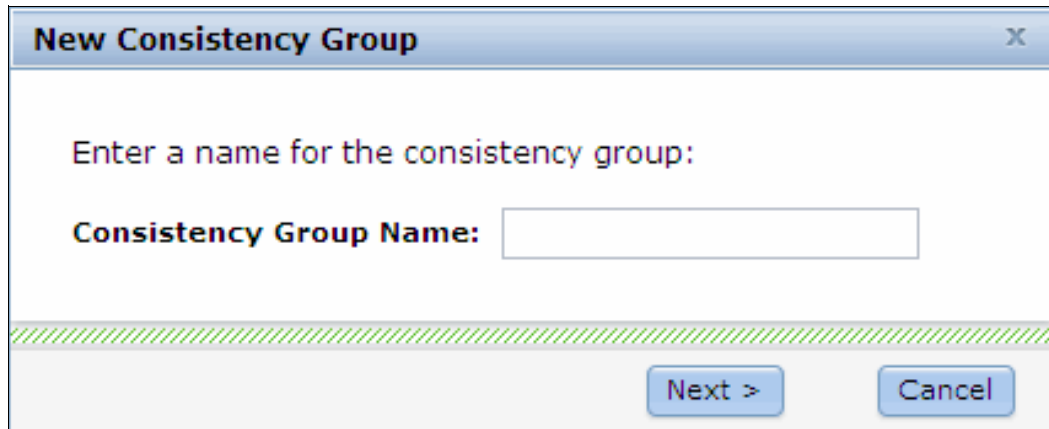


Figure 10-116 Create a Remote Copy consistency group

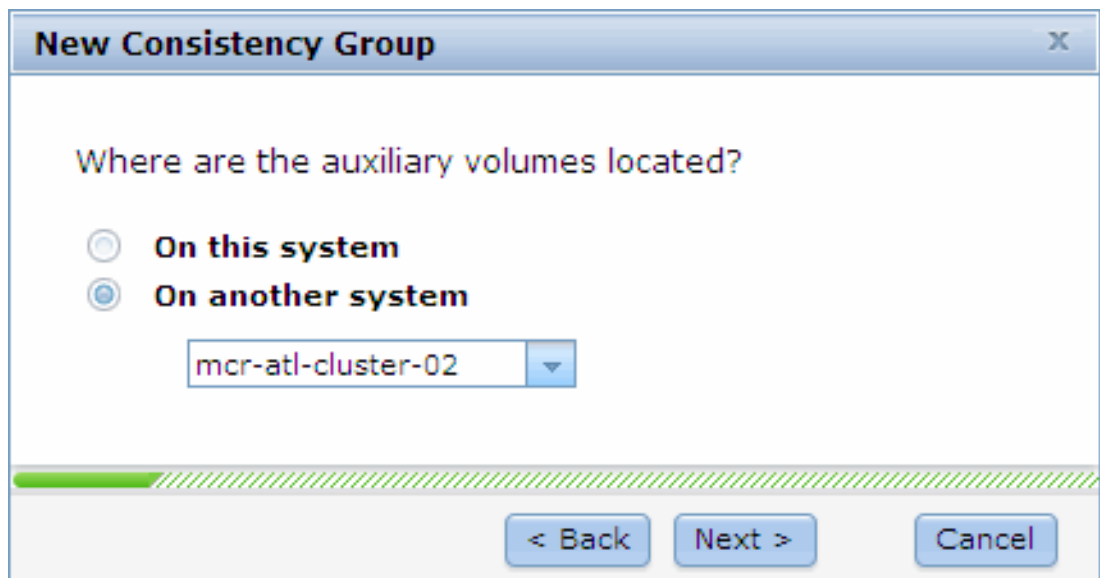
You must enter a name for your new consistency group, as shown in Figure 10-117.



The dialog box titled "New Consistency Group" has a close button (X) in the top right corner. The main text reads "Enter a name for the consistency group:". Below this is a label "Consistency Group Name:" followed by an empty text input field. At the bottom of the dialog, there are two buttons: "Next >" and "Cancel".

Figure 10-117 Enter a name for the new consistency group

You are prompted for the location of auxiliary volumes, as shown in Figure 10-118. In our example, these volumes are on another system. Select the relevant options and from the drop-down menu, select the correct remote system. In our example, we have only one remote system defined. Click **Next** to continue.



The dialog box titled "New Consistency Group" has a close button (X) in the top right corner. The main text reads "Where are the auxiliary volumes located?". Below this are two radio button options: "On this system" (unselected) and "On another system" (selected). Under the "On another system" option is a drop-down menu with the text "mcr-atl-cluster-02" and a downward arrow. At the bottom of the dialog, there are three buttons: "< Back", "Next >", and "Cancel".

Figure 10-118 Remote Copy consistency group auxiliary volume location window

You are then prompted to create an empty consistency group or add relationships to it, as shown in Figure 10-119.

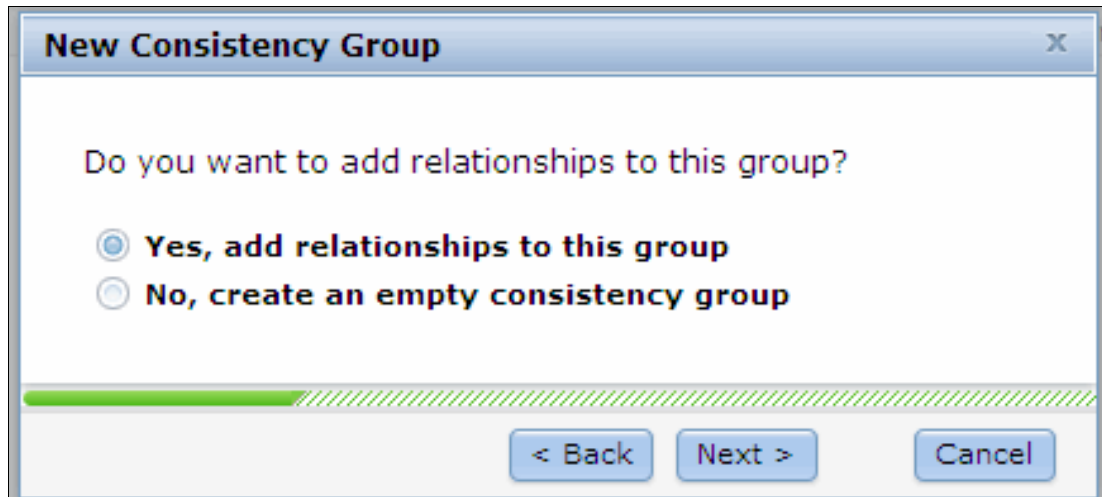


Figure 10-119 Creating an empty consistency group

If you select **No** and click **Finish**, the wizard completes and creates an empty Remote Copy Consistency Group. Selecting **Yes** prompts for the type of copy to create, as shown in Figure 10-120.

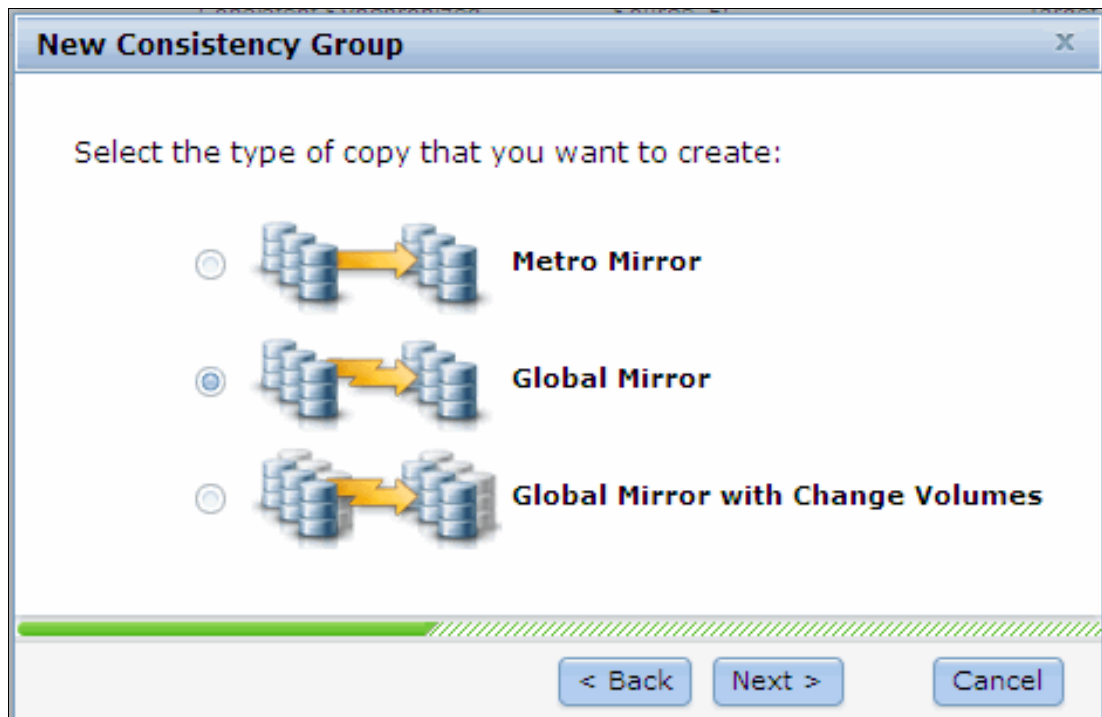


Figure 10-120 Remote Copy consistency group copy type

Choose the relevant copy type and click **Next**. In the following window, you can choose existing relationships to add the new consistency group. This step is optional. Use the Ctrl and Shift keys to select multiple relationships to add. If you decide that you do not want to use any of these relationships but you do want to create other relationships, click **Next**.

However, if you already highlighted a relationship and then decide you do not want any of these relationships, you cannot remove the relationship. You must stop the wizard and start again, as shown in Figure 10-121.

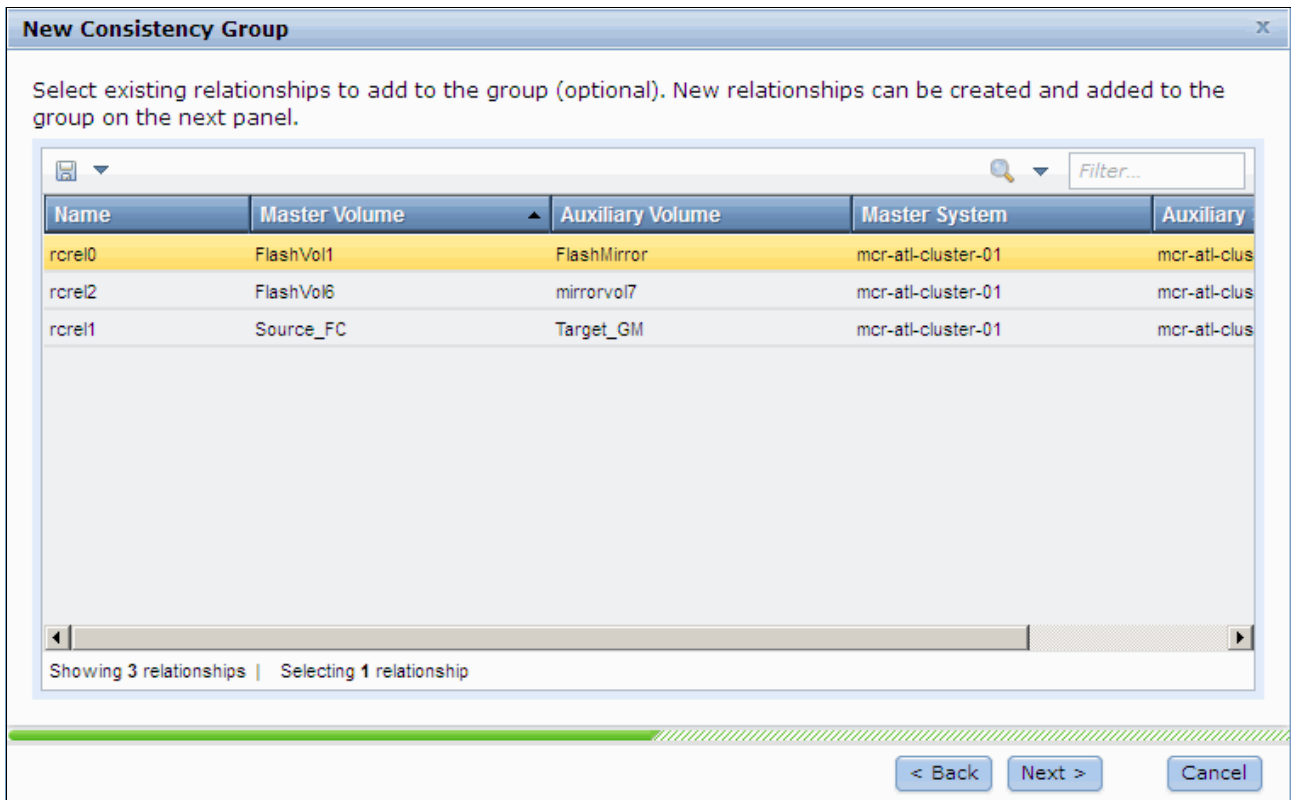


Figure 10-121 Selecting existing relationships

The next window is optional and gives the option to create relationships to add to the consistency group, as shown in Figure 10-122.

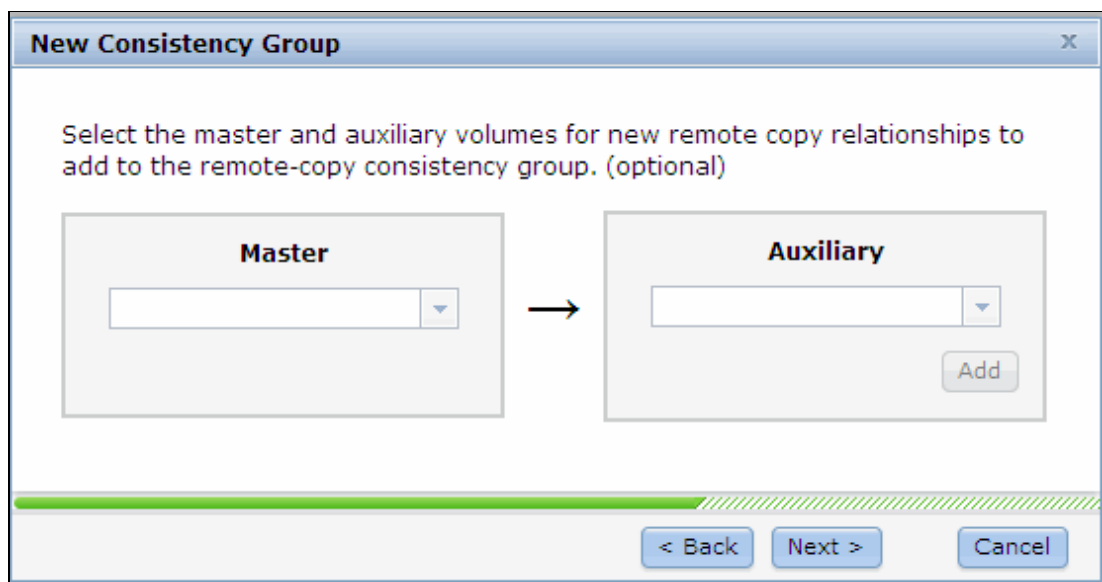


Figure 10-122 Creating relationships for Remote Copy consistency group

Select the relevant Master and Auxiliary volumes for the relationship you want to create and click **Add**. Multiple relationships can be defined by selecting another Master and Auxiliary volume and clicking **Add** again. When you finish, click **Next**. The next window prompts for whether the relationships are synchronized, as shown in Figure 10-123.

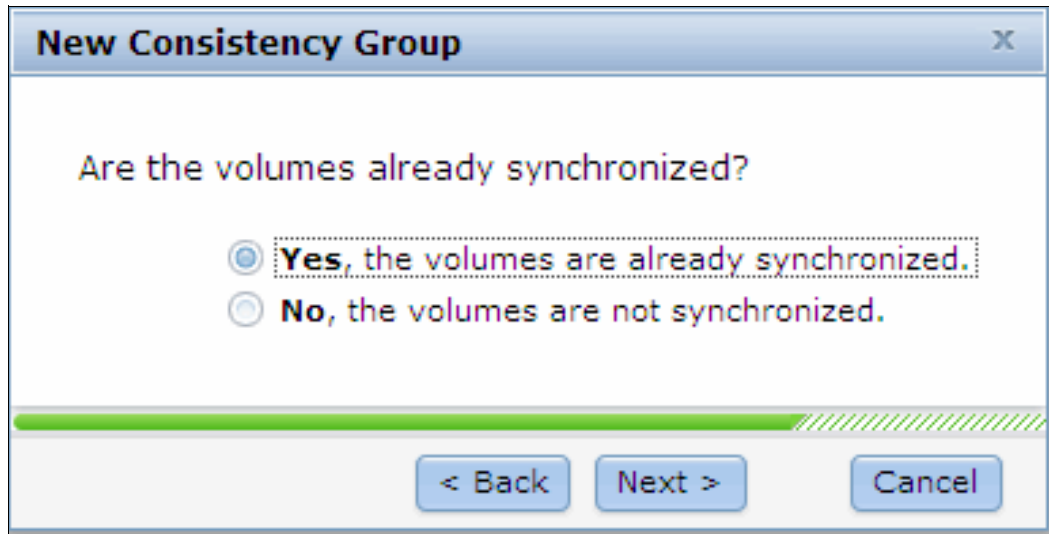


Figure 10-123 Volume synchronization

In the next window, you are asked whether you want to start copying the volumes now, as shown in Figure 10-124.

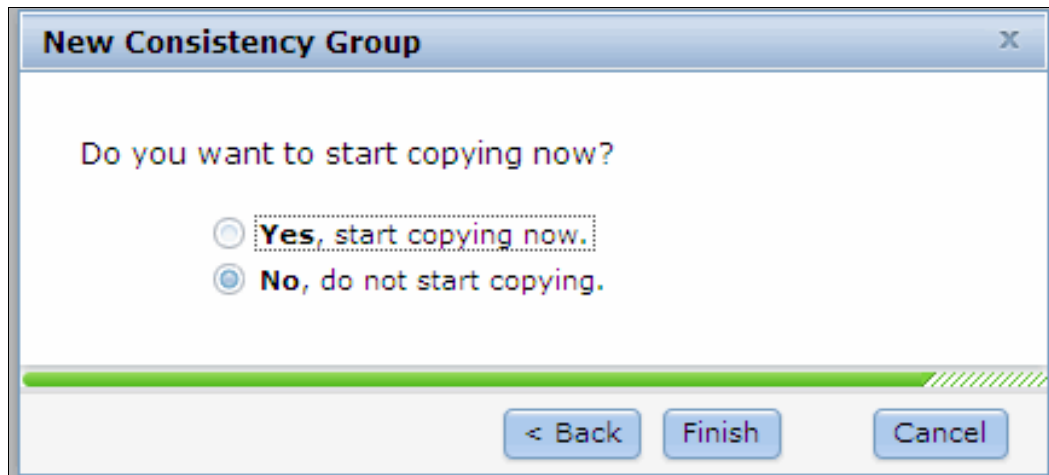


Figure 10-124 Remote Consistency group start copying option

After you select this option, click **Finish** to create the Remote Copy Consistency Group. Click **Close** to close the task window and the new consistency group is now shown in the GUI, as shown in Figure 10-125 on page 539.

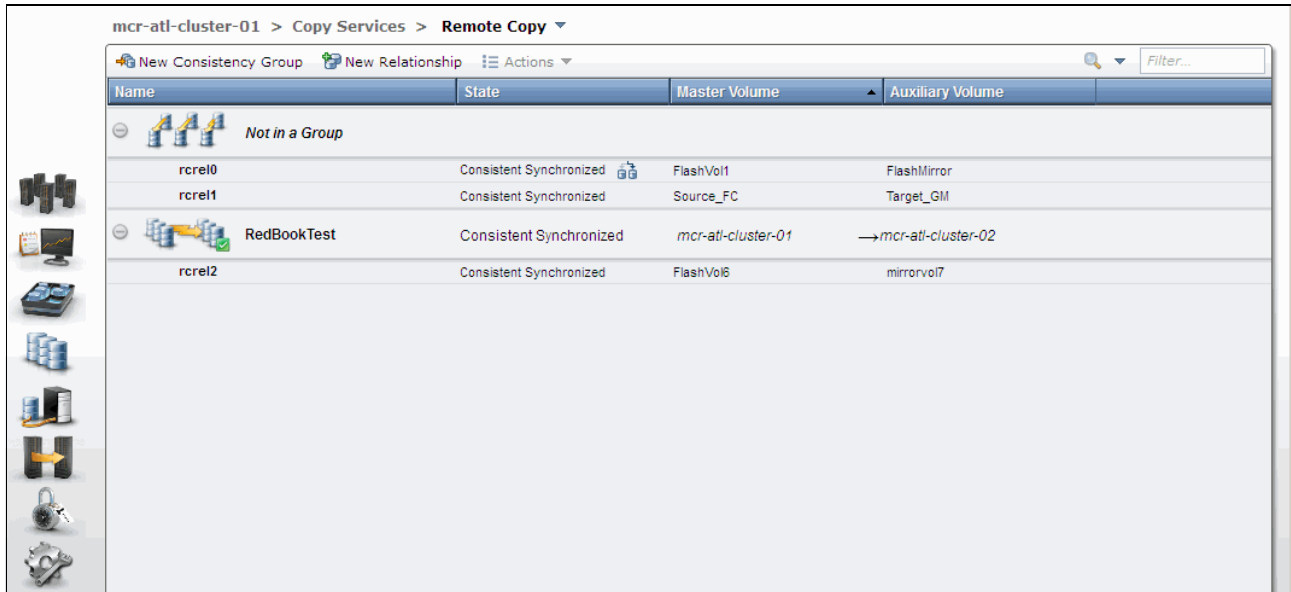


Figure 10-125 New Remote Consistency group created

In our example, we created a consistency group with a single relationship. Other Remote Copy relationships are added to the consistency group later.

You can find the name and the status of the consistency group beside the Relationship function icon. It is easy to change the name of consistency group by right-clicking the name, selecting **Rename** and then entering a new name. Alternatively, highlight the consistency group and select **Rename** from the Actions drop-down menu. Similarly, below the Relationship function icon is the Remote Copy relationships in this consistency group. The actions on the Remote Copy relationships can be applied here by using the Actions drop-down menu or right-clicking the relationships, as shown in Figure 10-126.

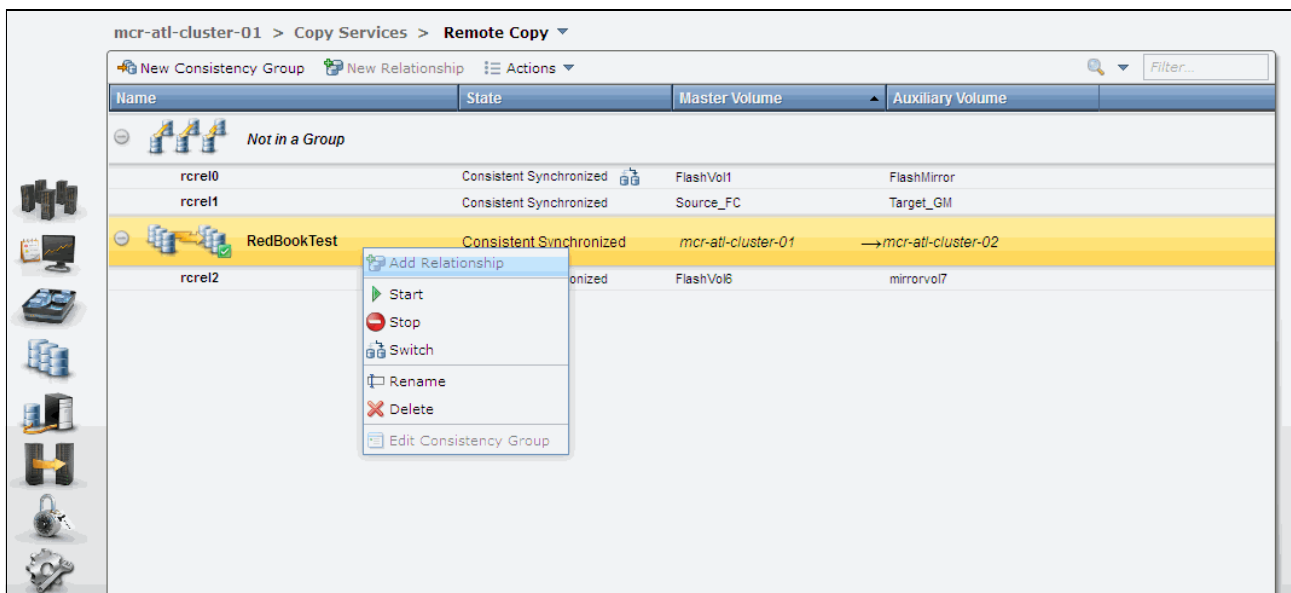


Figure 10-126 Drop-down menu options

Adding Remote Copy to a consistency group

The Remote Copy relationships in the Not in a Group list can be added to a consistency group by selecting the volumes and clicking **Add to Consistency Group** from the Actions drop-down menu, as shown in Figure 10-127.

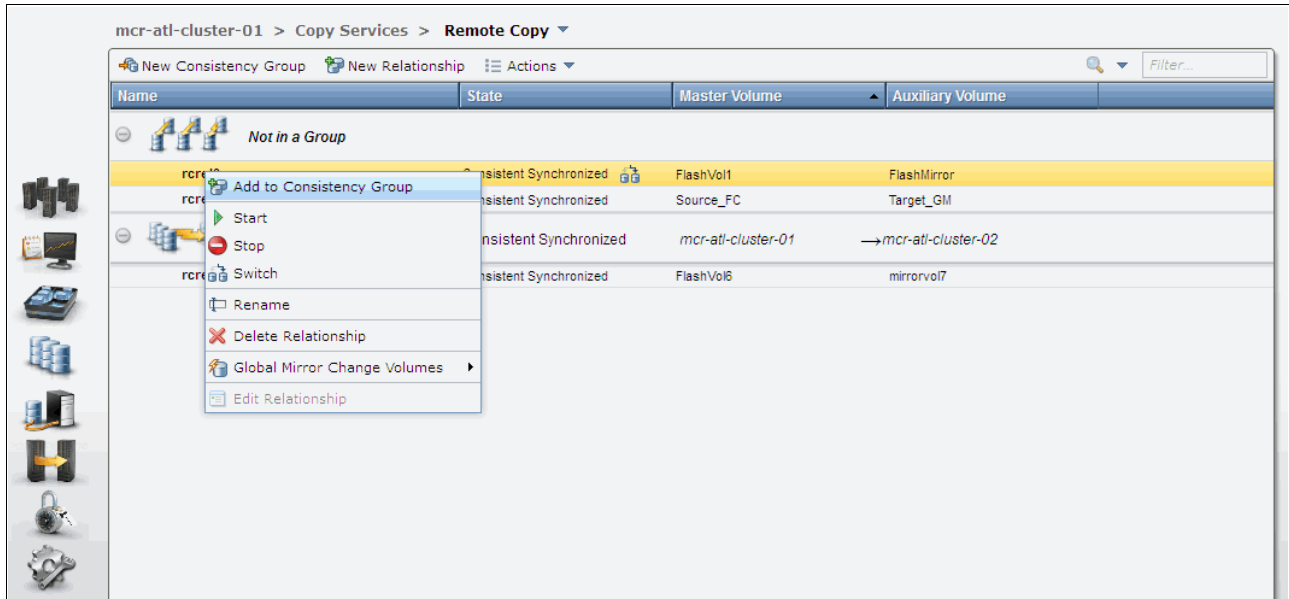


Figure 10-127 Add Remote Copy relationships to a consistency group

You must choose the consistency group to which to add the Remote Copy relationships. Based on your requirements, select the appropriate consistency group and click **Add to Consistency Group**, as shown in Figure 10-128.

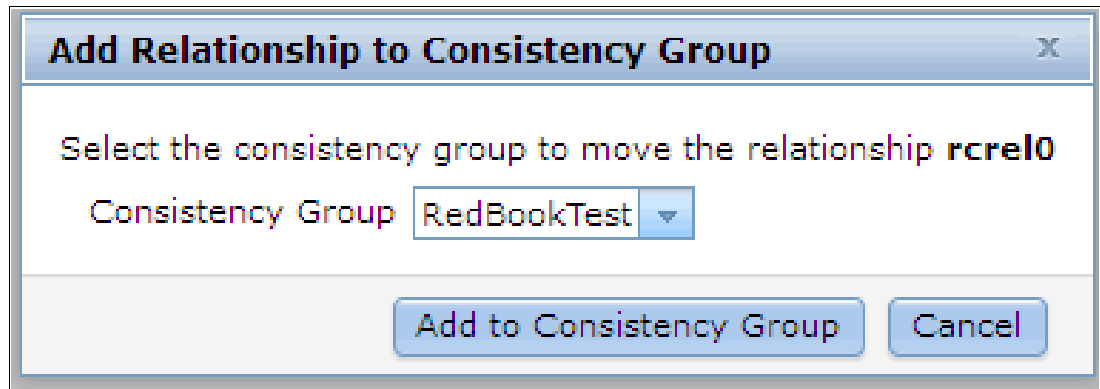


Figure 10-128 Choose the consistency group to add the remote copies

Your Remote Copy relationships are now in the consistency group that you selected.

Starting a consistency group

The Remote Copy relationship can be started by clicking **Start** from the Actions drop-down menu, as shown in Figure 10-129 on page 541.

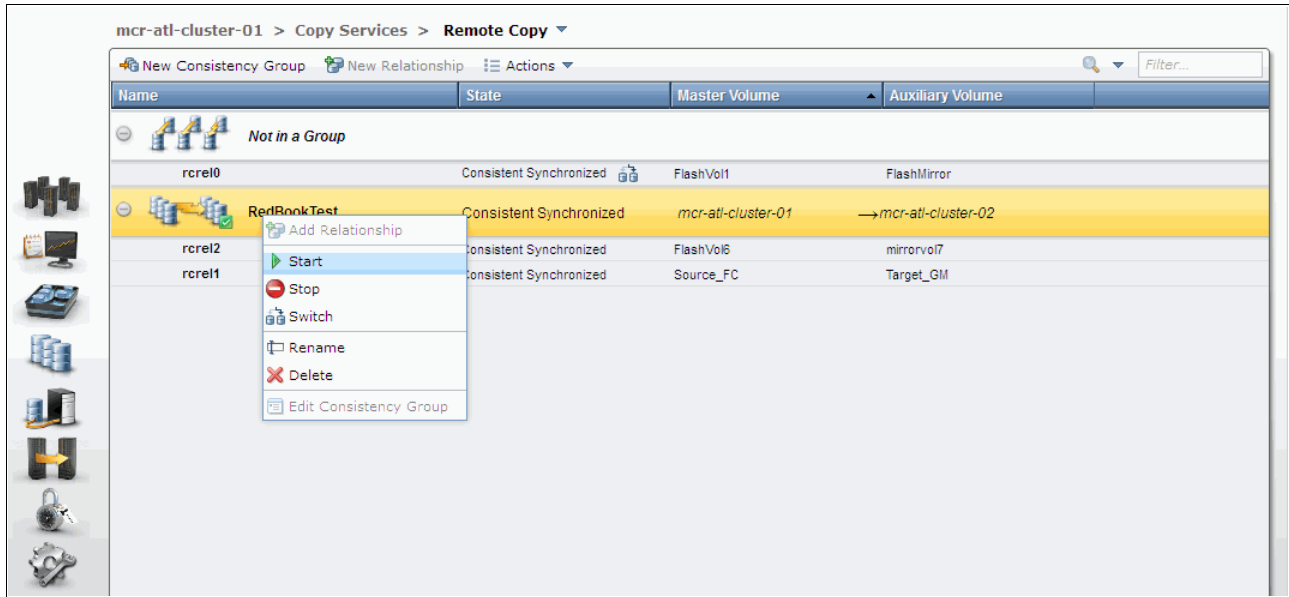


Figure 10-129 Start the consistency group

The consistency group starts copying data from the primary to the secondary.

Stopping a consistency group

The Remote Copy relationship can be stopped by clicking **Stop** in the Actions drop-down menu, as shown in Figure 10-130.

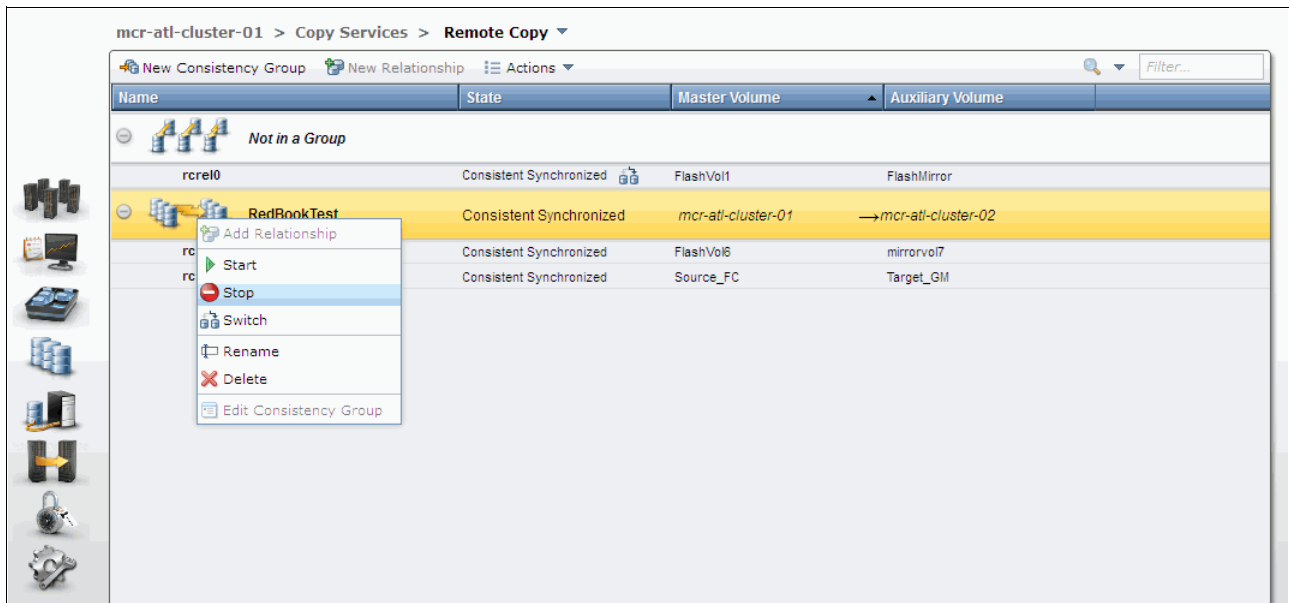


Figure 10-130 Stop the consistency group

You can allow read/write access to secondary volumes by selecting the option (as shown in Figure 10-131) and clicking **Stop Consistency Group**.

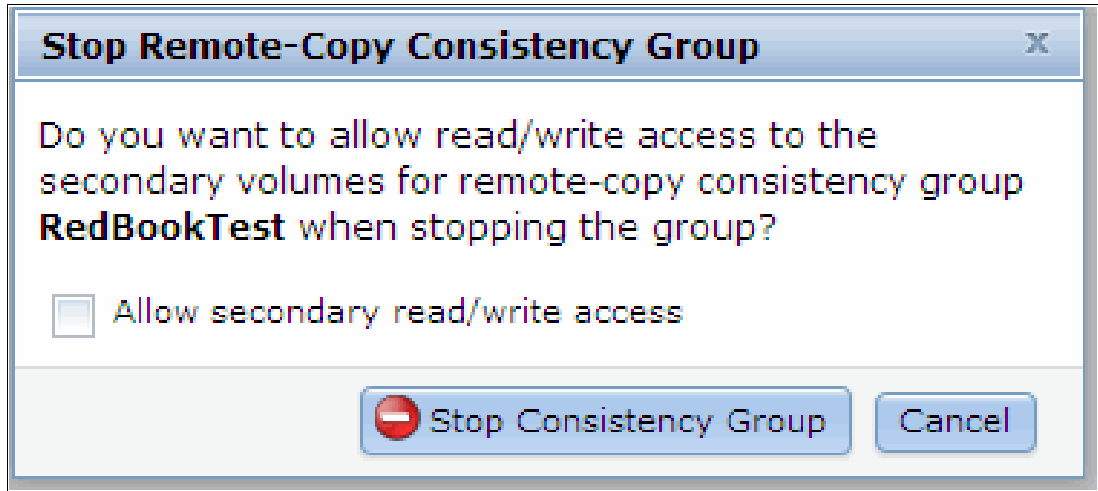


Figure 10-131 Confirm consistency group stop and allow secondary read/write access

Switching a consistency group

As with the switch action on the Remote Copy relationship, you can switch the copy direction of the consistency group. To switch the copy direction of the consistency group, click **Switch** from the Actions drop-down menu, as shown in Figure 10-132.

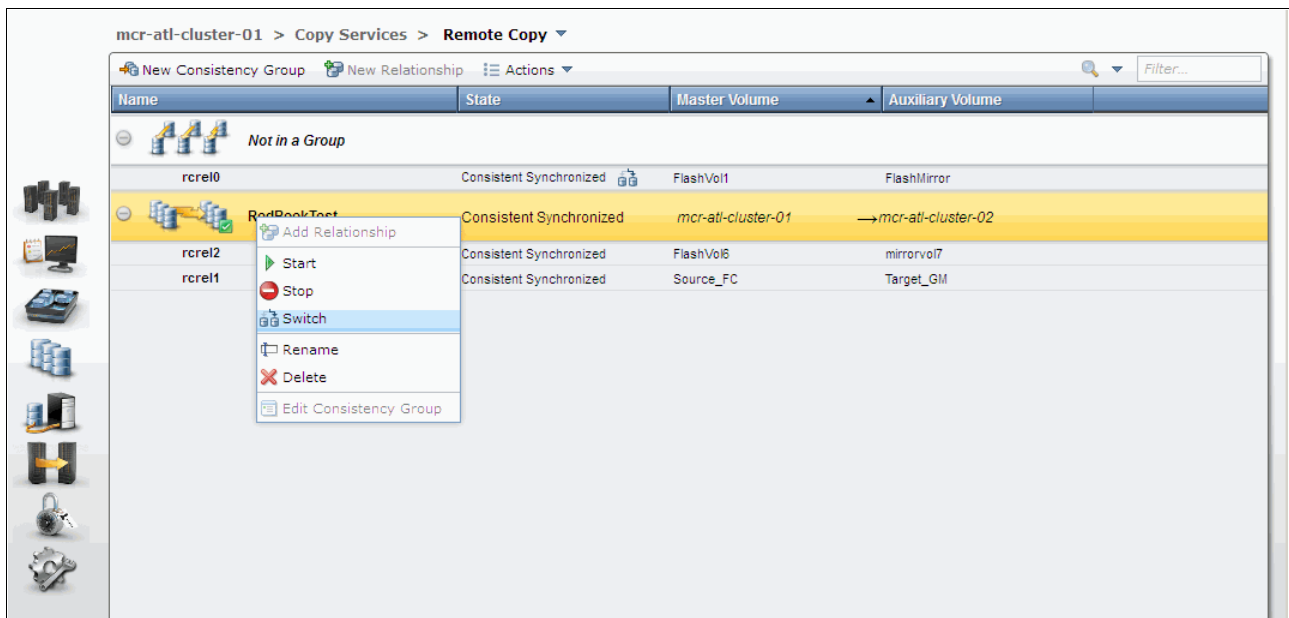


Figure 10-132 Switch the copy direction of a consistency group

A warning message opens, as shown in Figure 10-133. After the switch, the primary cluster in the consistency group changes. Write access to current master volumes is lost, while write access to the current auxiliary volumes is enabled. This change affects host access, so make sure that these settings are what you need, and if so, click **OK** to continue.

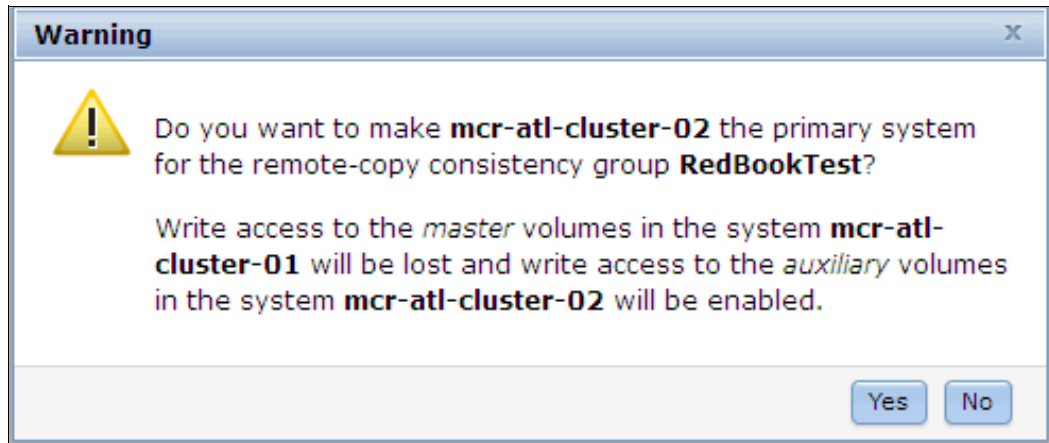


Figure 10-133 Warning message to confirm the switch

Removing Remote Copy relationships from a consistency group

The Remote Copy relationships can be removed from the consistency group by selecting the Remote Copy relationships and clicking **Remove from Consistency Group** from the Actions drop-down menu, as shown in Figure 10-134.

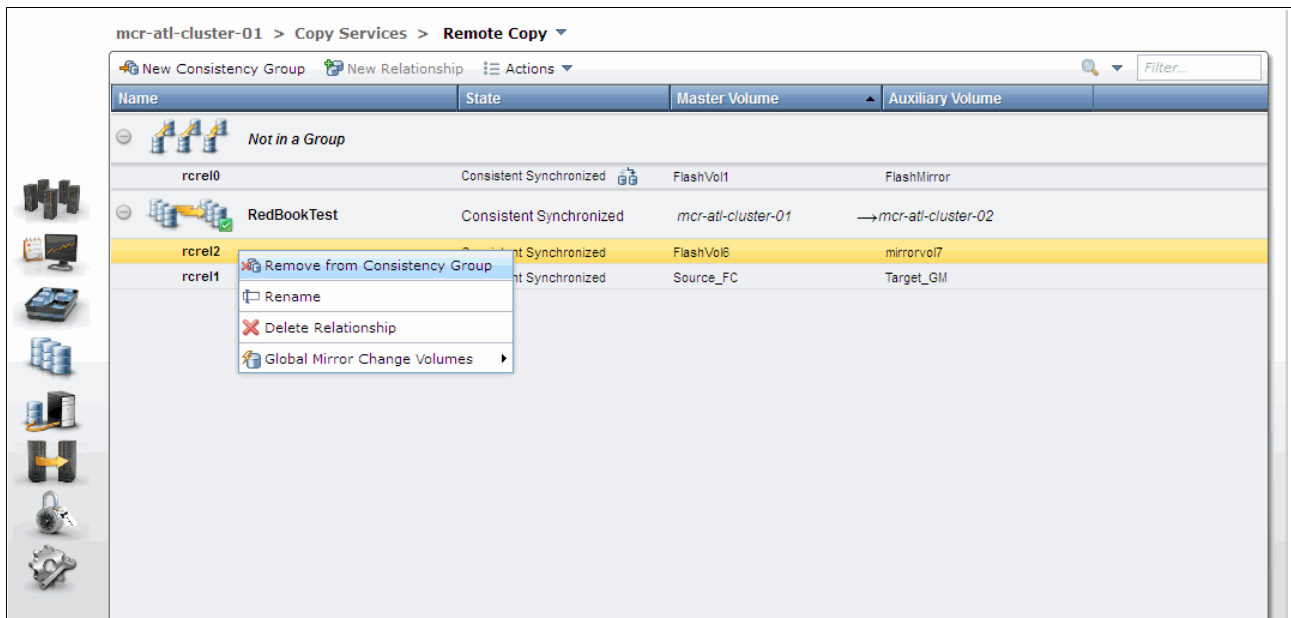


Figure 10-134 Remove Remote Copy relationships from a consistency group

You are prompted to confirm the Remote Copy relationships you want to delete from the consistency group, as shown in Figure 10-135. Make sure the Remote Copy relationships that are shown in the field are the ones that you must remove from the consistency group. Click **Remove** to proceed.

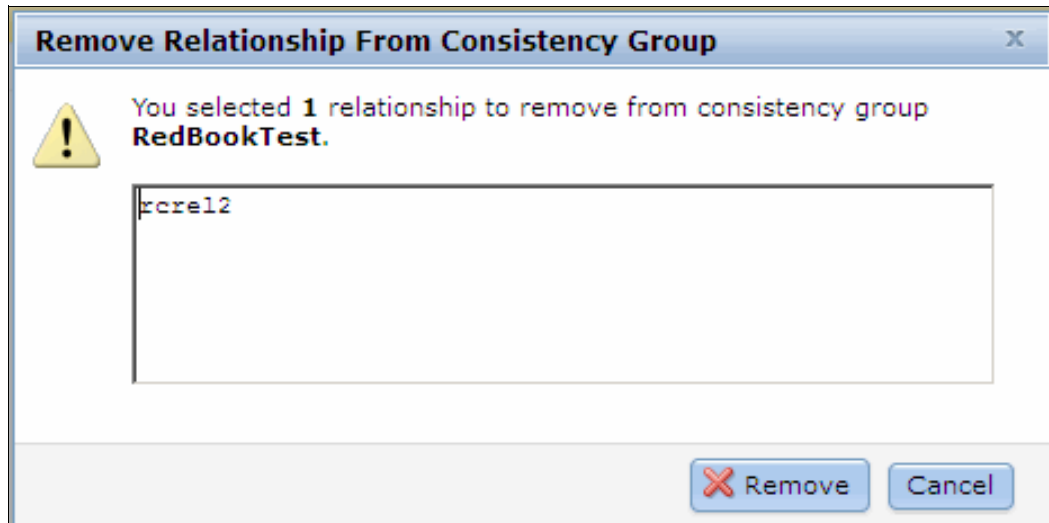


Figure 10-135 Confirm the relationships to remove from the Remote Copy consistency group

After the removal process completes, the Remote Copy relationships are deleted from the consistency group and displayed in the Not in a Group list.

Deleting a consistency group

The consistency group can be deleted by selecting **Delete** from the Actions drop-down menu, as shown in Figure 10-136.

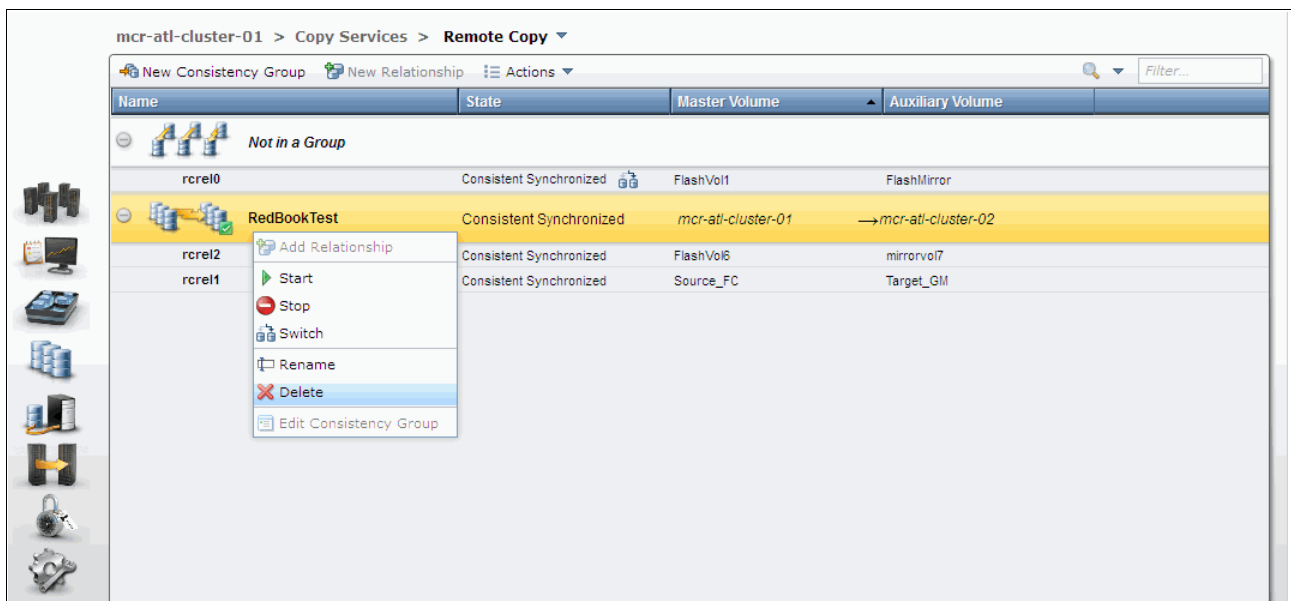


Figure 10-136 Delete a consistency group

You must confirm the deletion of the consistency group, as shown in Figure 10-137. Click **OK** if you are sure that this consistency group should be deleted.

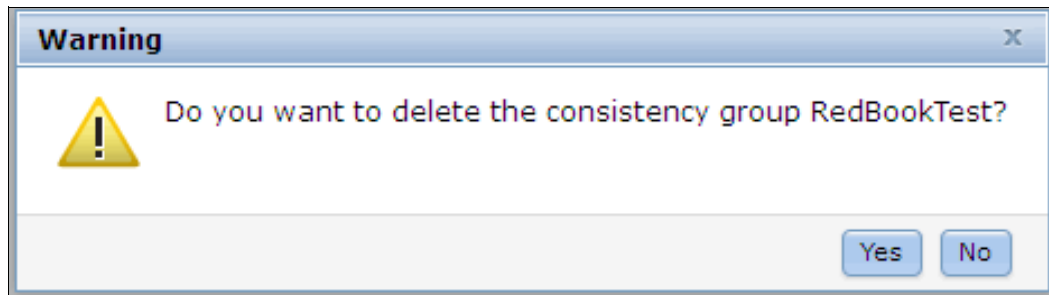


Figure 10-137 Warning to confirm deletion of the consistency group

The consistency group is deleted. Any relationships that were part of the consistency group are returned to the Not in a Group list.



External storage virtualization

In this chapter, we describe how to incorporate external storage systems into the virtualized world of the IBM Storwize V5000. A key feature of IBM Storwize V5000 is its ability to consolidate disk controllers from various vendors into pools of storage. In this way, the storage administrator from a single user interface can manage and provision storage to applications, and use a common set of advanced functions across all the storage systems under the control of the IBM Storwize V5000.

This chapter includes the following topics:

- ▶ Planning for external storage virtualization
- ▶ Working with external storage

11.1 Planning for external storage virtualization

In this section, we describe how to plan for virtualizing external storage with IBM Storwize V5000. Virtualizing the storage infrastructure with IBM Storwize V5000 makes your storage environment more flexible, cost-effective, and easy to manage. The combination of IBM Storwize V5000 and an external storage system allows more storage capacity benefits from the powerful software function within the IBM Storwize V5000.

The external storage systems that are incorporated into the IBM Storwize V5000 environment can be new systems or existing systems. The data on the existing storage systems can be easily migrated to the IBM Storwize V5000 managed environment, as described in Chapter 6, “Storage migration wizard” on page 237, and Chapter 7, “Storage pools” on page 295.

11.1.1 License for external storage virtualization

From a licensing standpoint, when external storage systems are virtualized by IBM Storwize V5000, a per-enclosure External Virtualization license is required. For more information, contact your IBM account team or IBM Business Partner for further assistance.

Migration: If the IBM Storwize V5000 is used as a general migration tool, the appropriate External Virtualization licenses must be ordered. The only exception is if you want to migrate existing data from external storage systems to IBM Storwize V5000 internal storage because you can temporarily configure your External Storage license within 45 days. For a more than 45 day migration requirement from external storage to IBM Storwize V5000 internal storage, an appropriate External Virtualization license must be ordered.

You can configure the IBM Storwize V5000 licenses by clicking the **Settings** icon and then clicking **General** → **Licensing**, as shown in Figure 11-1.

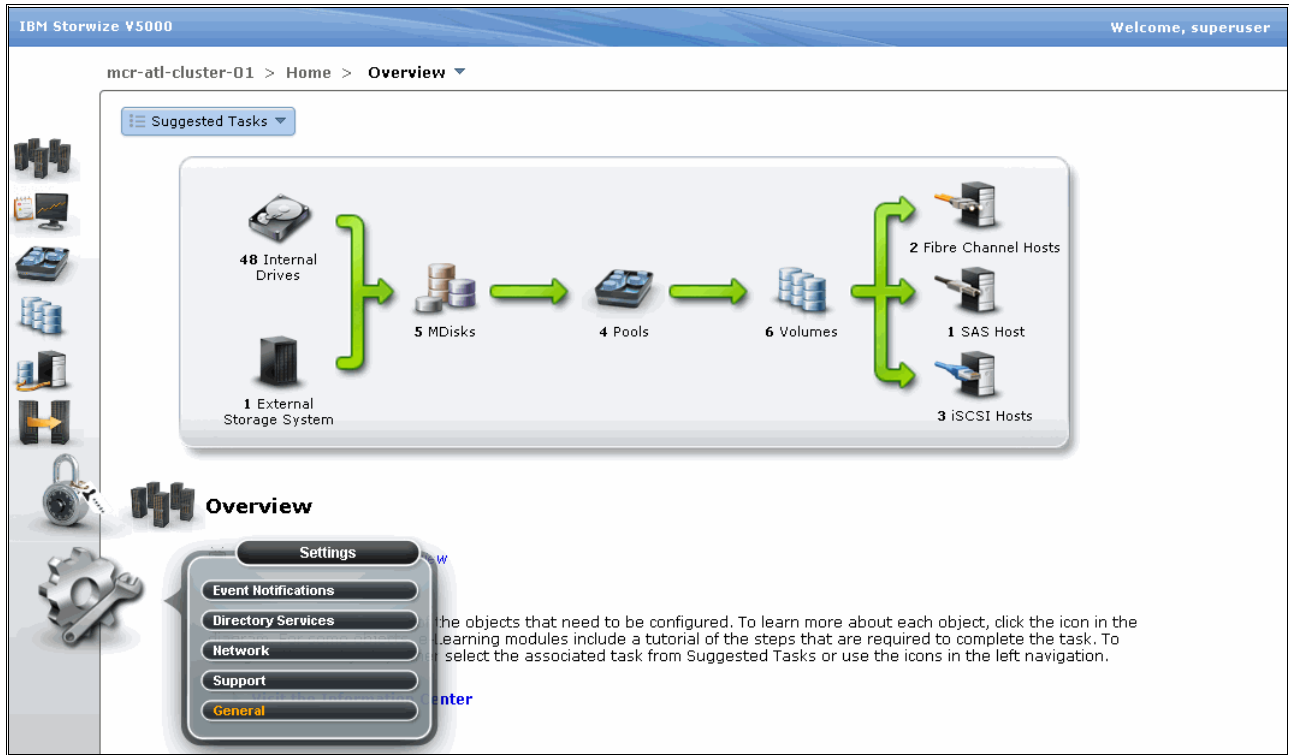


Figure 11-1 General option

In the Advanced window, click **Licensing** and the Update License view opens in the right pane, as shown in Figure 11-2.

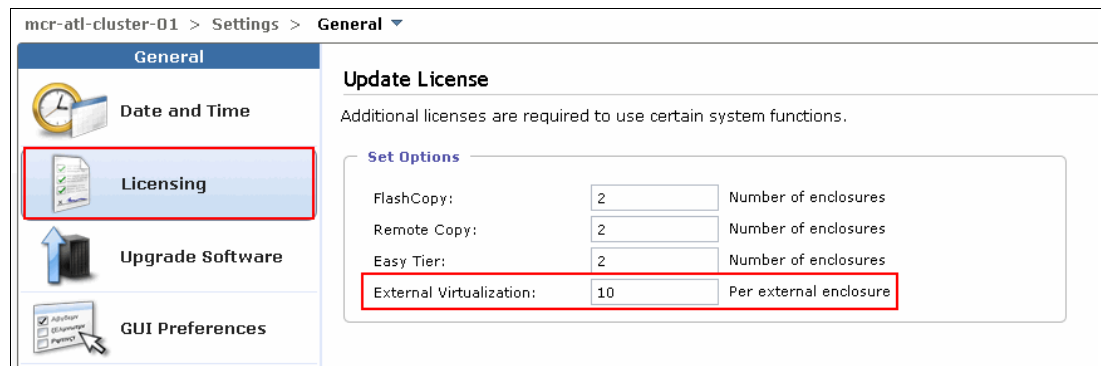


Figure 11-2 Update License window

In the Update License pane, there are two license options you can set: External Virtualization Limit and Remote-Copy Limit. Set these license options to the limit you obtained from IBM.

For assistance with licensing questions or to purchase an External Virtualization or Remote Copy license, contact your IBM account team or IBM Business Partner.

11.1.2 SAN configuration planning

External storage controllers that are virtualized by IBM Storwize V5000 must be connected through SAN switches. A direct connection between the IBM Storwize V5000 and storage controllers or hosts ports is not supported.

Make sure that the switches or directors are at the firmware levels that are supported by the IBM Storwize V5000 and that the IBM Storwize V5000 port login maximums that are listed in the restriction document are not exceeded. The configuration restrictions can be found on the Support home page, which is available at this website:

<http://www-947.ibm.com/support/entry/portal/Overview>

The recommended SAN configuration is composed of a minimum of two fabrics. The ports on external storage systems are virtualized by the IBM Storwize V5000 and the IBM Storwize V5000 ports and are evenly split between the two fabrics to provide redundancy if one of the fabrics goes offline.

After the IBM Storwize V5000 and external storage systems are connected to the SAN fabrics, zoning must be implemented. In each fabric, create a zone with the four IBM Storwize V5000 worldwide port names (WWPNs), two from each node canister with up to a maximum of eight WWPNs from each external storage system.

Ports: IBM Storwize V5000 supports a maximum of 16 ports or WWPNs from an external storage system that is virtualized.

Figure 11-3 shows an example of how to cable devices to the SAN. Refer to this example as we describe the zoning.

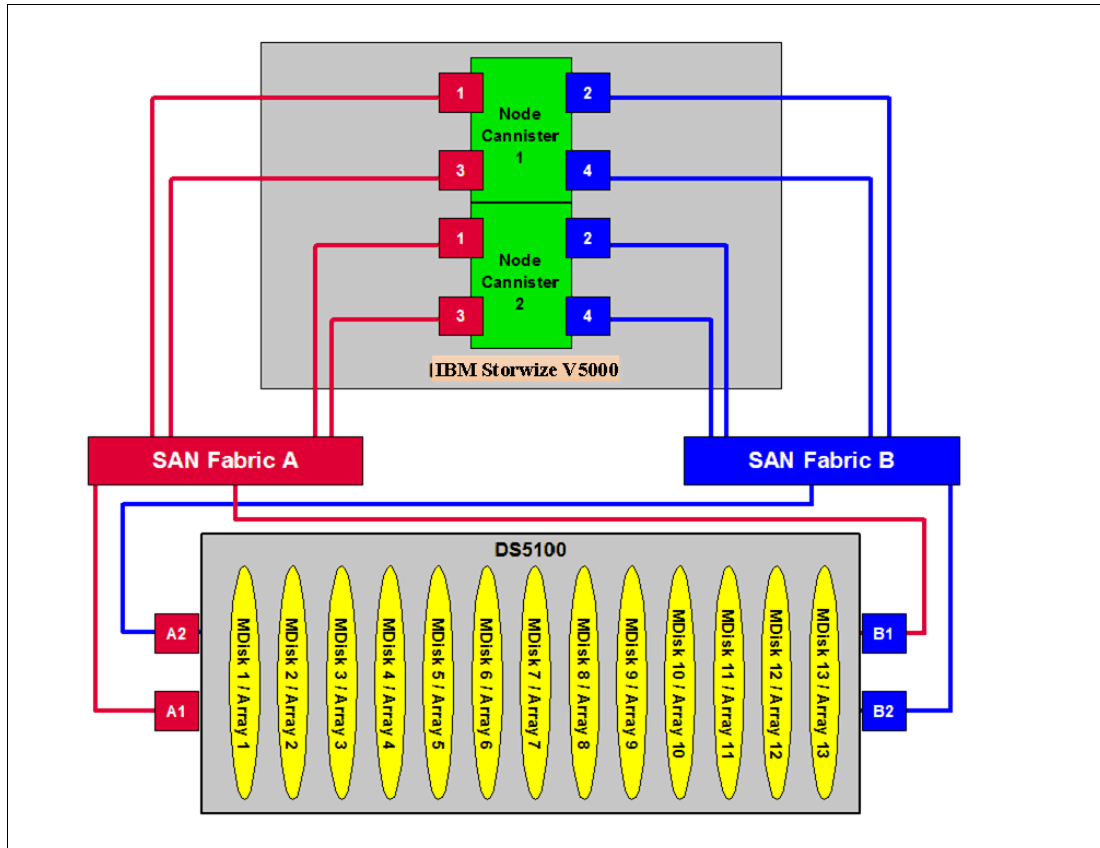


Figure 11-3 SAN cabling and zoning example diagram

Create an IBM Storwize V5000/external storage zone for each storage system to be virtualized, as shown in the following examples:

- ▶ Zone DS5100 controller ports A1 and B1 with all node ports 1 and 3 in the RED fabric
- ▶ Zone DS5100 controller ports A2 and B2 with all node ports 2 and 4 in the BLUE fabric

11.1.3 External storage configuration planning

External storage systems provide redundancy through various RAID levels, which prevents a single physical disk failure from causing an MDisk, storage pool, or associated host volume from going offline. To minimize the risk of data loss, virtualize storage systems only where logical unit numbers (LUNs) are configured by using a RAID level other than RAID 0 (for example RAID 1, RAID 10, RAID 0+1, RAID 5, or RAID 6).

Verify that the storage controllers to be virtualized by IBM Storwize V5000 meet the requirements. The configuration restrictions can be found on the Support home page, which is available at this website:

<http://www-947.ibm.com/support/entry/portal/Overview>

Make sure that the firmware or microcode levels of the storage controllers to be virtualized are supported by IBM Storwize V5000.

IBM Storwize V5000 must have exclusive access to the LUNs from the external storage system that are mapped to it. LUNs cannot be shared between IBM Storwize V5000s or between an IBM Storwize V5000 and other storage virtualization platforms or between an IBM Storwize V5000 and hosts. However, different LUNs can be mapped from one external storage system to an IBM Storwize V5000 and other hosts in the SAN through different storage ports.

Make sure to configure the storage subsystem LUN masking settings to map all LUNs to all the WWPNs in the IBM Storwize V5000 storage system.

Ensure that you see the IBM Storwize V5000 Information Center and review the “Configuring and servicing external storage system” topic before you prepare the external storage systems for discovery by the IBM Storwize V5000 system. This Information Center can be found at this website:

<http://publib.boulder.ibm.com/infocenter/storwize/ic/index.jsp>

11.1.4 Guidelines for virtualizing external storage

When external storage is virtualized by using the IBM Storwize V5000, the following guidelines must be followed:

- ▶ Avoid splitting arrays into multiple LUNs at the external storage system level. When possible, create a single LUN per array for mapping to the IBM Storwize V5000.
- ▶ Except for Easy Tier, do not mix MDisks that vary in performance or reliability in the same storage pool. Always put similarly sized MDisks into one storage pool. For more information about Easy Tier, see Chapter 9, “Easy Tier” on page 411.
- ▶ Do not leave volumes in image mode. Use image mode only to import or export existing data into or out of the IBM Storwize V5000. Migrate such data from image mode MDisks to other storage pools to benefit from storage virtualization.
- ▶ The use of the copy services in Storwize V5000 gives you a unified method to manage data integrity across heterogeneous storage systems.
- ▶ The Easy Tier function is included with the IBM Storwize V5000 system. The external storage system can benefit from this powerful storage tiering function to remove hot spots and improve overall performance.

11.2 Working with external storage

In this section, we describe how to manage external storage by using an IBM Storwize V5000.

The basic concepts of managing external storage system are the same as internal storage. IBM Storwize V5000 discovers LUNs from the external storage system as one or more MDisks. These MDisks are added to a storage pool in which volumes are created and mapped to hosts, as needed.

11.2.1 Adding external storage

To add new external storage systems to the IBM Storwize V5000 virtualized environment, complete the following steps:

1. Zone a minimum of two and a maximum of 16 Fibre Channel ports from the external storage system with all eight Fibre Channel ports on the IBM Storwize V5000 system. As a best practice, have two fabrics for redundancy in the SAN. Then, in each fabric, zone two ports from each node canister in the IBM Storwize V5000 system with half the ports from the external system. As the IBM Storwize V5000 is virtualizing your storage, hosts should be zoned with the V5000 controllers WWPNs.
2. By using the storage partitioning or LUN masking feature of the external storage system, create a group that includes all eight IBM Storwize V5000 WWPNs.
3. Create equal size arrays on the external system by using any RAID level except zero.
4. Create a single LUN per RAID array.
5. Map the LUNs to all eight Fibre Channel ports on the IBM Storwize V5000 system by assigning them to the group that was created in step 2.
6. Verify that IBM Storwize V5000 discovered the LUNs as unmanaged MDisks. If they do not show up automatically, click **Detect MDisk** from the MDisk window of the GUI, as described in Chapter 7, “Storage pools” on page 295. You should see the MDisks mapped to the IBM Storwize V5000 under the respective Storage system.
7. Select the storage tier for the MDisks.
8. Create a storage pool.
9. Add the MDisks to the pool.
10. Create volumes and map them to hosts, as needed.

If the external storage systems are not new systems (that is, there is existing data on the LUNs that must be kept after virtualization), complete the steps that are described in Chapter 6, “Storage migration wizard” on page 237 to prepare the environment. You can then migrate the existing data with or without the use of the wizard to IBM Storwize V5000 internal storage or some other external storage system.

Chapter 6, “Storage migration wizard” on page 237 shows how to manually import MDisks and migrate the data to other storage pools. Whether you migrate the data with the wizard, you can select your destination storage pools to be internal storage pools or external storage pools.

11.2.2 Managing external storage

The IBM Storwize V5000 provides an individual external window for managing external storage systems.

You can access the external window by opening the Getting Started window and clicking the External Storage System function icon. Extended help information for external storage appears. Click **Physical Storage** and the external window opens.

Figure 11-4 shows how to access the External Storage window from the Getting Started window.

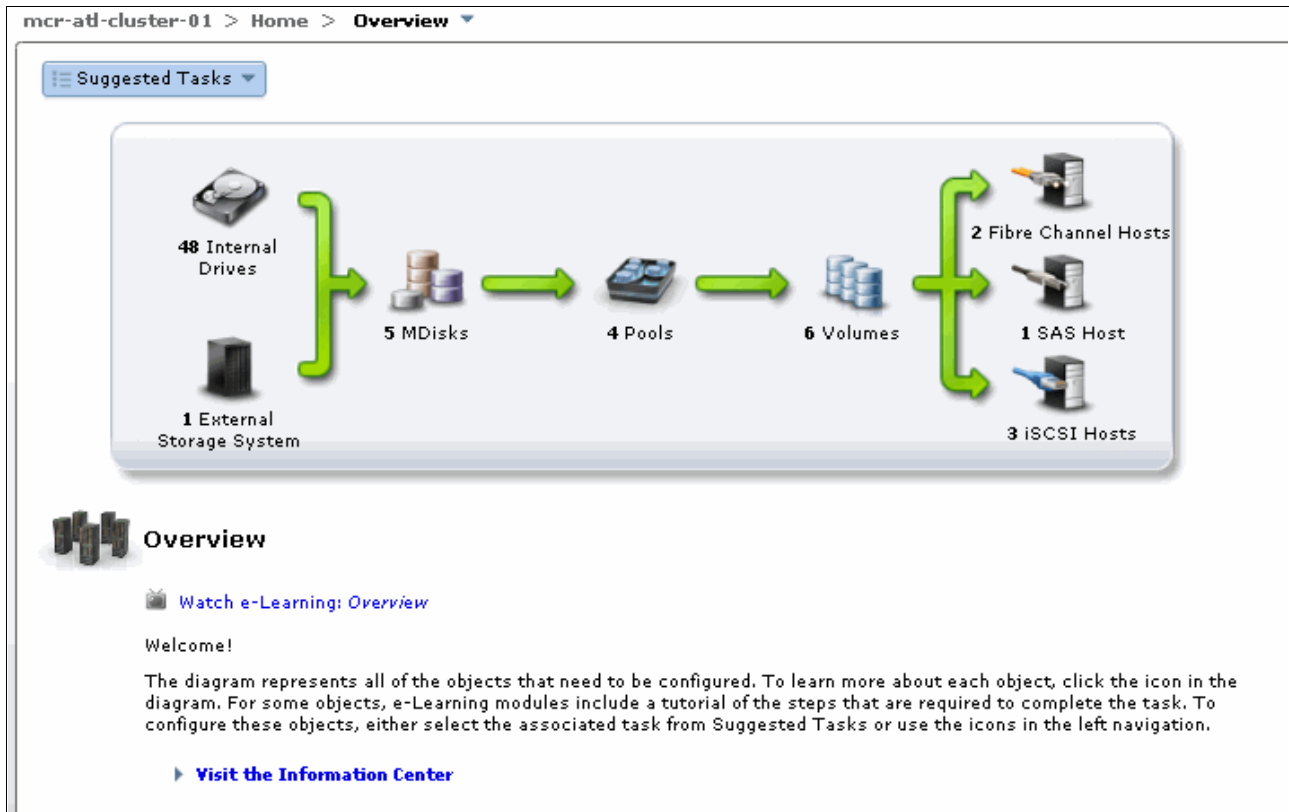


Figure 11-4 Access the External Storage window from the Getting Started window

The other method to access the external window is to use the Physical Storage function icons that are shown in the left pane, as shown in Figure 11-5.

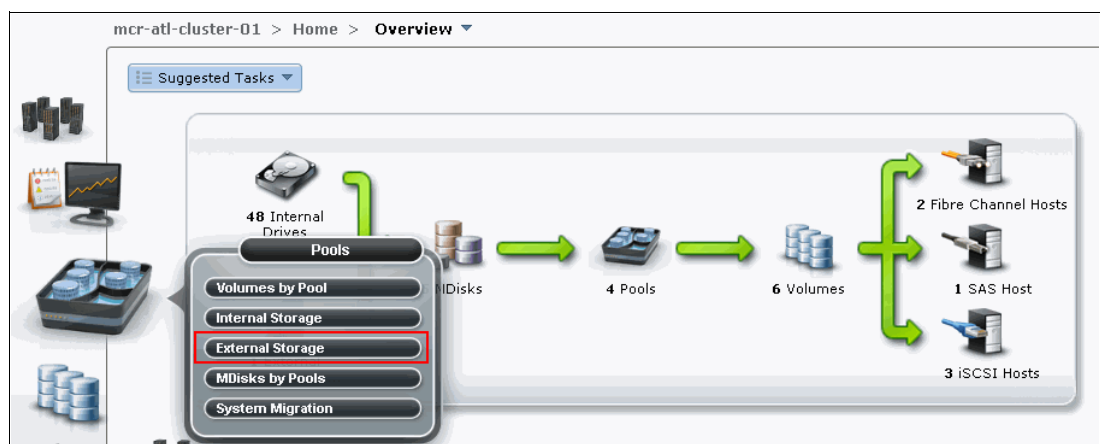


Figure 11-5 Access the External window from the Getting Started window

The External window (see Figure 11-6) gives you an overview of all your external storage systems. There is a list of the external storage systems on the left side of the window. With the help of the filter, you can show only the external storage systems on which you must act. If you click and highlight the external storage system, detailed information is shown in the right pane, including all the MDisks that are provided by it.

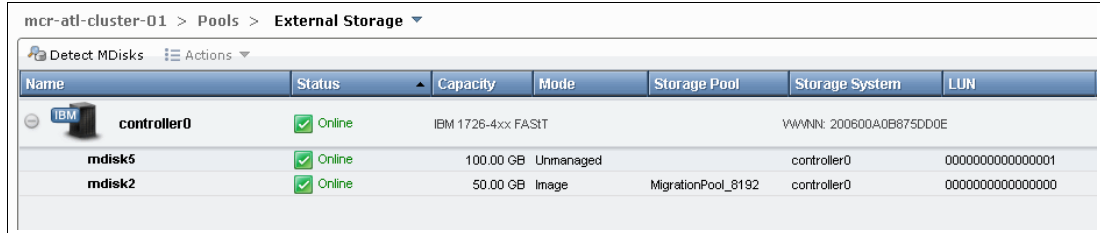


Figure 11-6 External Storage window

On the right side of the window, you can change the name of external storage system by clicking the name beside the picture of the external storage box. The status of the external storage system and its WWPN can also be found under the name.

From the Actions drop-down list (which is found at the top of the name of external storage on the right part of the External window), you can find the Show Dependent Volumes option, as shown in Figure 11-7.

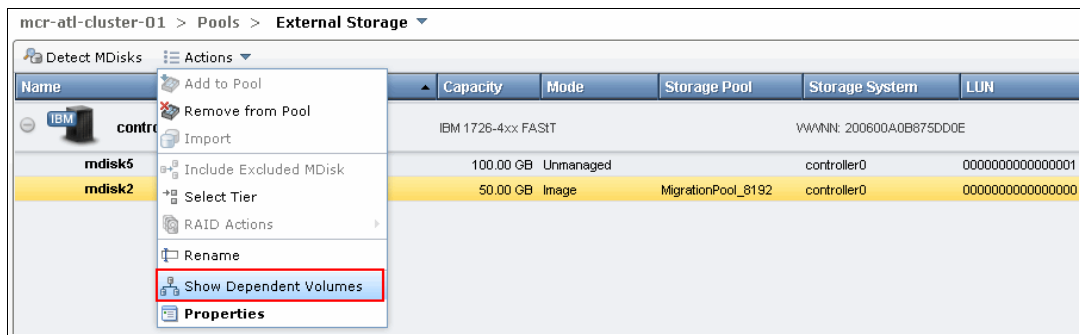


Figure 11-7 Show Dependent Volumes option in the Actions drop-down menu

Clicking the **Show Dependent Volumes** option shows you the volumes in this external storage system, as shown in Figure 11-8.

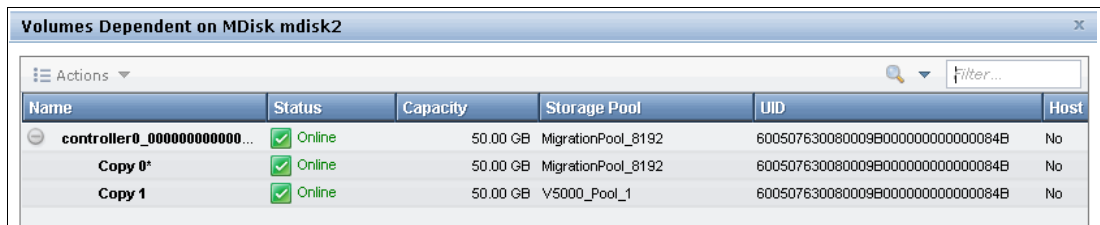


Figure 11-8 Volumes dependent on external storage

In the window that is shown in Figure 11-8 on page 555, you can take volume actions, including Map to Host, Shrink, Expand, Migrate to Another Pool, and Volume Copy Actions, as shown in Figure 11-9.

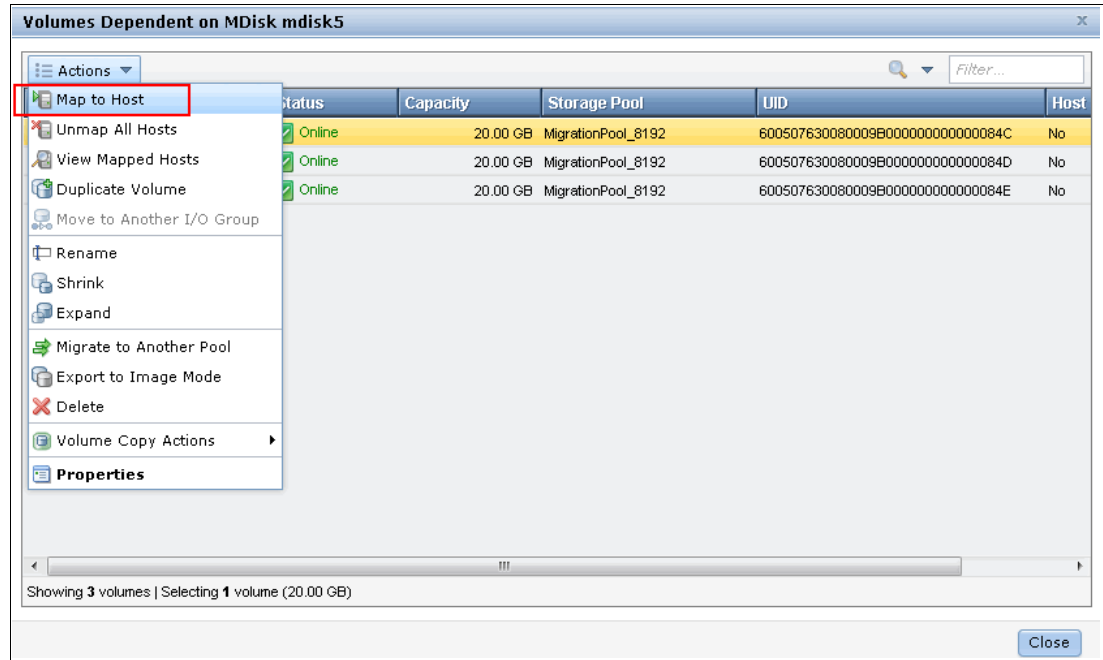


Figure 11-9 Actions that you can take with volumes

One of the features of the IBM Storwize V5000 storage system is that it can be used as a data migration tool. In the IBM Storwize V5000 virtualization environment, you can migrate your application data nondisruptively from one internal or external storage system to another storage system, which makes storage management much simpler with less risk.

Volume copy is another key feature that you can benefit from by using IBM Storwize V5000 virtualization. Two copies can be applied to your data to enhance the availability for a critical application. A volume copy can be also used to generate test data or data migration.

For more information about the volume actions of the IBM Storwize V5000 storage system, see Chapter 8, “Advanced host and volume administration” on page 349.

Returning to the External window, you discover an MDisk menu on the right, including an MDisk list that shows the MDisks that are provided by this external storage system. You can find the name of an MDisk, its capacity, the storage pool, and the storage system it belongs to in the list. The actions on MDisks can also be made through the menu, including Detect MDisks, Add to Pool, and Import. This menu is the same as the one in the MDisks window.

Figure 11-10 shows the MDisk menu for the external storage window.

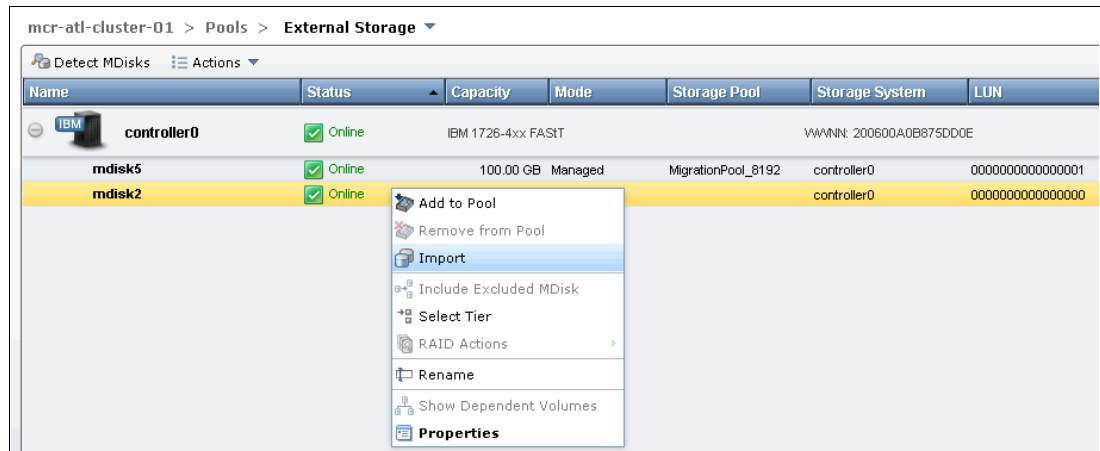


Figure 11-10 MDisk menu in the External window

11.2.3 Removing external storage

If you want to remove the external storage systems from the IBM Storwize V5000 virtualized environment, you have the following options:

- ▶ If you want to remove the external storage systems and discard the data on it, complete the following steps:
 - a. Stop any host I/O on the volumes.
 - b. Remove the volumes from the host file system, logical volume, or volume group, and remove the volumes from the host device inventory.
 - c. Remove the host mapping of volumes and the volumes on IBM Storwize V5000.
 - d. Remove the storage pools to which the external storage systems belong, or you can keep the storage pool and remove the MDisks of the external storage from the storage pools.
 - e. Unzone and disconnect the external storage systems from the IBM Storwize V5000.
 - f. Click **Detect MDisks** to make IBM Storwize V5000 discover the removal of the external storage systems.
- ▶ If you want to remove the external storage systems and keep the volumes and their data on the IBM Storwize V5000, complete the following steps:
 - a. Migrate volumes and their data to the other storage pools that are on IBM Storwize V5000 internal storage or other external storage systems.
 - b. Remove the storage pools to which the external storage systems belong, or you can keep the storage pools and remove the MDisks of the external storage from the storage pools.
 - c. Unzone and disconnect the external storage systems from the IBM Storwize V5000.
 - d. Click **Detect MDisks** to make IBM Storwize V5000 discover the removal of the external storage systems.

- ▶ If you want to remove the external storage systems from IBM Storwize V5000 control and keep the volumes and their data on external storage systems, complete the following steps:
 - a. Migrate volumes and their data to the other storage pools that are on IBM Storwize V5000 internal storage or other external storage systems, as described in Chapter 6, “Storage migration wizard” on page 237.
 - b. Remove the storage pools to which the external storage systems belong, or you can keep the storage pools and remove the MDisks of the external storage from the storage pools.
 - c. Export volumes to image mode with the MDisks on the external storage systems. For more information about the restrictions and prerequisites for migration, see Chapter 6, “Storage migration wizard” on page 237.

You also must record pre-migration information; for example, the original SCSI IDs the volumes used to be mapped to hosts. Some operating systems do not support changing the SCSI ID during the migration. For more information about migration, see the IBM Storwize V5000 Information Center at this website:

<http://publib.boulder.ibm.com/infocenter/storwize/ic/index.jsp>

- d. Unzone and disconnect the external storage systems from the IBM Storwize V5000.
 - a. Click **Detect MDisks** to make IBM Storwize V5000 discover the removal of the external storage systems.



RAS, monitoring, and troubleshooting

There are various ways to monitor and troubleshoot the IBM Storwize V5000. In this chapter, we show the ways in which the IBM Storwize V5000 can be administered from a monitoring and troubleshooting point of view.

This chapter includes the following topics:

- ▶ Reliability, availability, and serviceability on the IBM Storwize V5000
- ▶ IBM Storwize V5000 components
- ▶ Configuration backup procedure
- ▶ Upgrading software
- ▶ Event log
- ▶ Collecting support information
- ▶ Powering on and shutting down IBM Storwize V5000

12.1 Reliability, availability, and serviceability on the IBM Storwize V5000

This section describes the Reliability, Availability, and Serviceability (RAS) features of IBM Storwize V5000 monitoring and troubleshooting. RAS features are important concepts in the design of the IBM Storwize V5000. Hardware and software features, design considerations, and operational guidelines all contribute to make the IBM Storwize V5000 reliable.

Fault tolerance and a high level of availability are achieved by the following features:

- ▶ The RAID capabilities of the underlying disk subsystems.
- ▶ The compass architecture that is used by the IBM Storwize V5000 nodes.
- ▶ Auto-restart of nodes that are hung.
- ▶ Battery units to provide cache memory protection in the event of a site power failure.
- ▶ Host system multipathing and failover support.

High levels of serviceability are achieved by providing the following benefits:

- ▶ Cluster error logging
- ▶ Asynchronous error notification
- ▶ Dump capabilities to capture software detected failures
- ▶ Concurrent diagnostic procedures
- ▶ Directed maintenance procedures
- ▶ Concurrent log analysis and memory dump data recovery tools
- ▶ Concurrent maintenance of all IBM Storwize V5000 components
- ▶ Concurrent upgrade of IBM Storwize V5000 Software and microcode
- ▶ Concurrent addition or deletion of a node canister in a cluster
- ▶ Software recovery through the Service Assistant Tool
- ▶ Automatic software version correction when a node is replaced
- ▶ Detailed status and error conditions that are displayed via the Service Assistant Tool
- ▶ Error and event notification through Simple Network Management Protocol (SNMP), syslog, and email
- ▶ Node canister support package gathering via USB, in case of network connection problem

At the heart of the IBM Storwize V5000 is a redundant pair of *node canisters*. The two canisters share the data transmitting and receiving load between the attached hosts and the disk arrays.

12.2 IBM Storwize V5000 components

This section describes each of the components that make up the IBM Storwize V5000 system. Components are described in terms of location, function, and serviceability.

12.2.1 Enclosure midplane assembly

The enclosure midplane assembly is the unit that contains the node or expansion canisters and the power supply units. The enclosure midplane assembly initially is generic and configured as a control enclosure midplane or an expansion enclosure midplane. During the basic system configuration, Vital Product Data (VPD) is written to the enclosure midplane assembly, which decides whether the unit is a control enclosure midplane or an expansion enclosure midplane.

Control enclosure midplane

The control enclosure midplane holds node canisters and the power supply units. The control enclosure midplane assembly has specific VPD, such as, WWNN 1, WWNN 2, machine type and model, machine part number, and serial number. The control enclosure midplane must be replaced only by a trained service provider. After a generic enclosure midplane assembly is configured as a control enclosure midplane, it is no longer interchangeable with an expansion enclosure midplane.

Expansion enclosure midplane

The expansion enclosure midplane holds expansion canisters and the power supply units. The expansion enclosure midplane assembly also has specific VPD, such as, machine type and model, machine part number, and serial number. After a generic enclosure midplane assembly is configured as an expansion enclosure midplane, it is no longer interchangeable with a control enclosure midplane. The expansion enclosure midplane must be replaced only by a trained service provider.

Figure 12-1 shows back of the Enclosure Midplane Assembly.

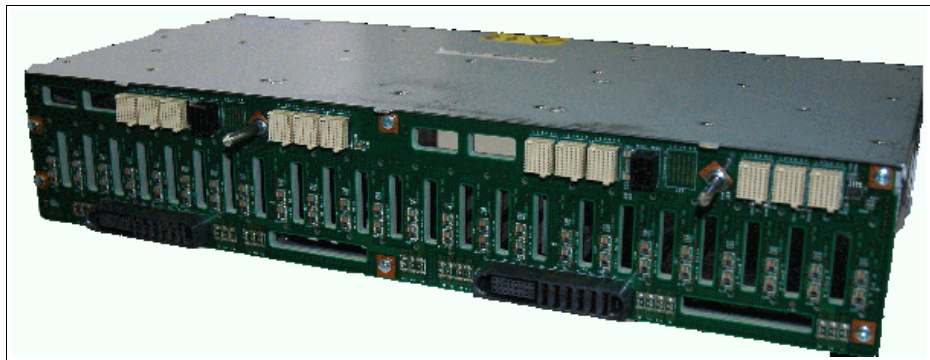


Figure 12-1 Rear view of Enclosure Midplane Assembly

For more information about replacing the control or expansion enclosure midplane, see the IBM Storwize V5000 Information Center at this website:

http://pic.dhe.ibm.com/infocenter/storwize/V5000_ic/index.jsp

12.2.2 Node canisters: Ports and LED

There are two node canister slots along the top of the unit. The left slot is canister 1 and the right slot is canister 2.

Figure 12-2 shows the back of a fully equipped node enclosure.



Figure 12-2 Node canister

USB ports

There are two USB connectors side-by-side and they are numbered as 1 on the left and as 2 on the right. There are no indicators that are associated with the USB ports. Figure 12-3 shows the USB ports.

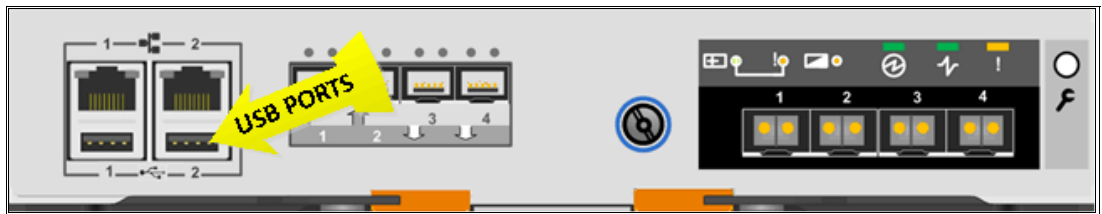


Figure 12-3 Node Canister USB ports

Ethernet ports

There are two 10/100/1000 Mbps Ethernet ports side-by-side on the canister and they are numbered 1 on the left and 2 on the right. Port 1 is required and port 2 optional. The ports are shown in Figure 12-4.



Figure 12-4 Node canister Ethernet ports

Each port has two LEDs and their status is shown in Table 12-1.

Table 12-1 Ethernet LEDs status

LED	Color	Meaning
Link state	Green	On: There is an Ethernet link.
Activity	Yellow	Flashing: There is activity on the link.

SAS ports

There are four 6-Gbps Serial Attached SCSI (SAS) ports side-by-side on the canister. They are numbered 1 on the left to 4 on the right. IBM Storwize V5000 uses port 1 and 2 for host connectivity and ports 3 and 4 to connect optional expansion enclosure. The ports are shown in Figure 12-5.

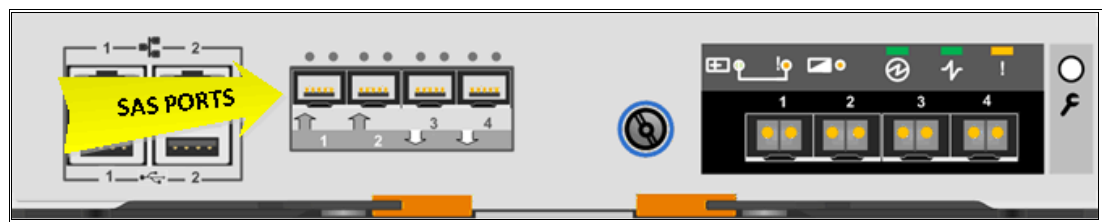


Figure 12-5 Node canister SAS ports

The SAS LED status meanings are described in Table 12-2.

Table 12-2 SAS LED Status

State	Meaning
green	Indicates at least one of the SAS lanes on this connector are operational. If the light is off when it is connected, there is a problem with the connection.
amber	If the light is on, one of the following errors occurred: <ul style="list-style-type: none"> ▶ One or more (but not all) of the four lanes are up for this connector (if none of the lanes are up, the activity light is off) ▶ One or more of the up lanes are running at a different speed to the others ▶ One or more of the up lanes are attached to a different address to the others

IBM Storwize V5000 uses SFF-8644 mini-SAS HD connector cable to connect enclosures, as shown in Figure 12-6.

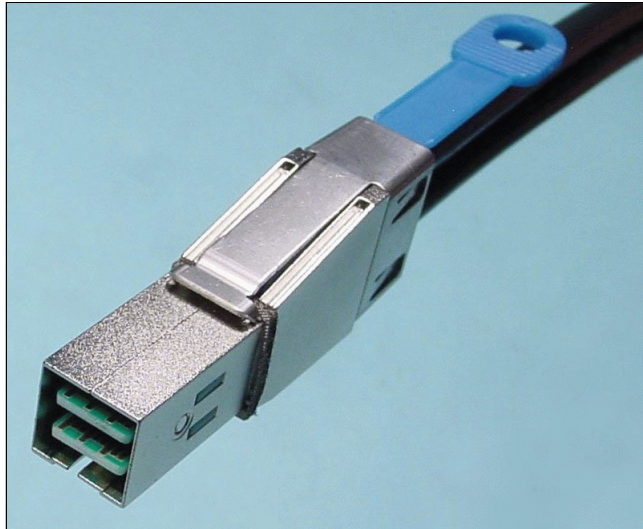


Figure 12-6 Mini-SAS HD SFF 8644 connector

Battery status

Each node canister stores a battery, the status of which is displayed on three LED on the back of the unit, as shown in Figure 12-7.

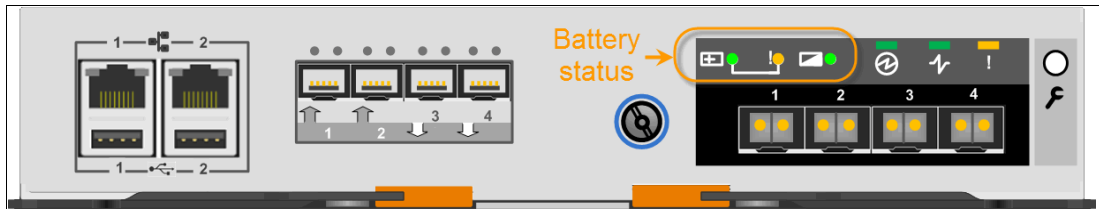


Figure 12-7 Node canister battery status

The battery indicator status meanings are described in Table 12-3.

Table 12-3 Battery indicator on Node canister

Color	Name	Definition
Green (left)	Battery Status	<ul style="list-style-type: none"> ▶ Fast flash: Indicates that the battery is charging and has insufficient charge to complete a single memory dump. ▶ Flashing: Indicates that the battery has sufficient charge to complete a single memory dump only. ▶ Solid: Indicates that the battery is fully charged and has sufficient charge to complete two memory dumps.
Amber	Fault	Indicates a fault with the battery.
Green (right)	Battery in use	Indicates that hardened or critical data is writing to disk.

Canister status

The status of each canister is displayed by three LEDs, as shown in Figure 12-8.

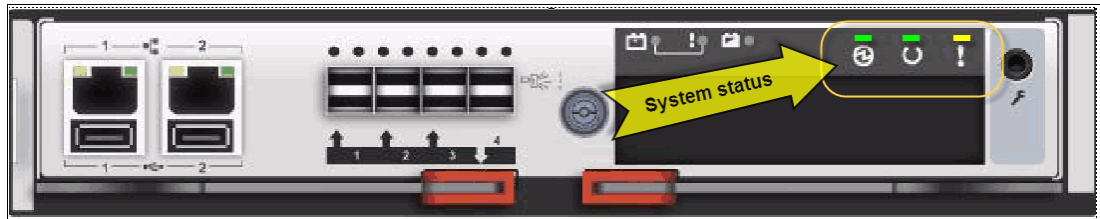


Figure 12-8 System status indicator

The system status LED meanings are described in Table 12-4.

Table 12-4 System status indicator

Color	Name	Definition
Green (left)	System Power	<ul style="list-style-type: none"> ▶ Flashing: The canister is in standby mode in which case IBM Storwize V5000 is not running. ▶ Fast flashing: The cannister is running a self test. ▶ On: The cannister is powered up and the IBM Storwize V5000 code is running.
Green (mid)	System Status	<ul style="list-style-type: none"> ▶ Off: There is no power to the canister, the canister is in standby mode, Power On SelfTest (POST) is running on the canister, or the operating system is loading. ▶ Flashing: The node is in candidate or service state; it cannot perform I/O. It is safe to remove the node. ▶ Fast flash: A code upgrade is running. ▶ On: The node is part of a cluster.
Amber	Fault	<ul style="list-style-type: none"> ▶ Off: The node is in candidate or active state. This state does not mean that there is no hardware error on the node. Any error that is detected is not severe enough to stop the node from participating in a cluster (or there is no power). ▶ Flashing: Identifies the canister. ▶ On: The node is in service state, or there is an error that is stopping the software from starting.

The host interface card replacement is shown in Figure 12-10.

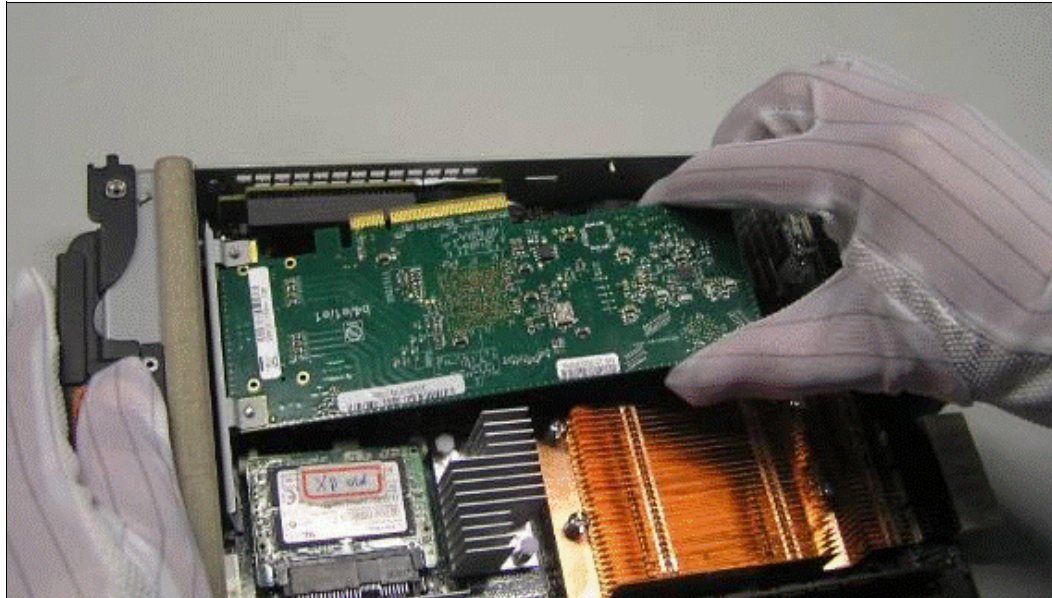


Figure 12-10 Host Interface card replacement

Memory replacement

For more information about the memory replacement process, see the IBM Storwize V5000 Information Center at this website:

http://pic.dhe.ibm.com/infocenter/storwize/V5000_ic/topic/com.ibm.storwize.V5000.641.doc/V5000_rplc_nodcan_dimm.html

At the website, browse to **Troubleshooting Removing and replacing parts** → **Replacing the node canister memory (2x 4 GB DIMM)**.

Figure 12-11 shows the memory location.

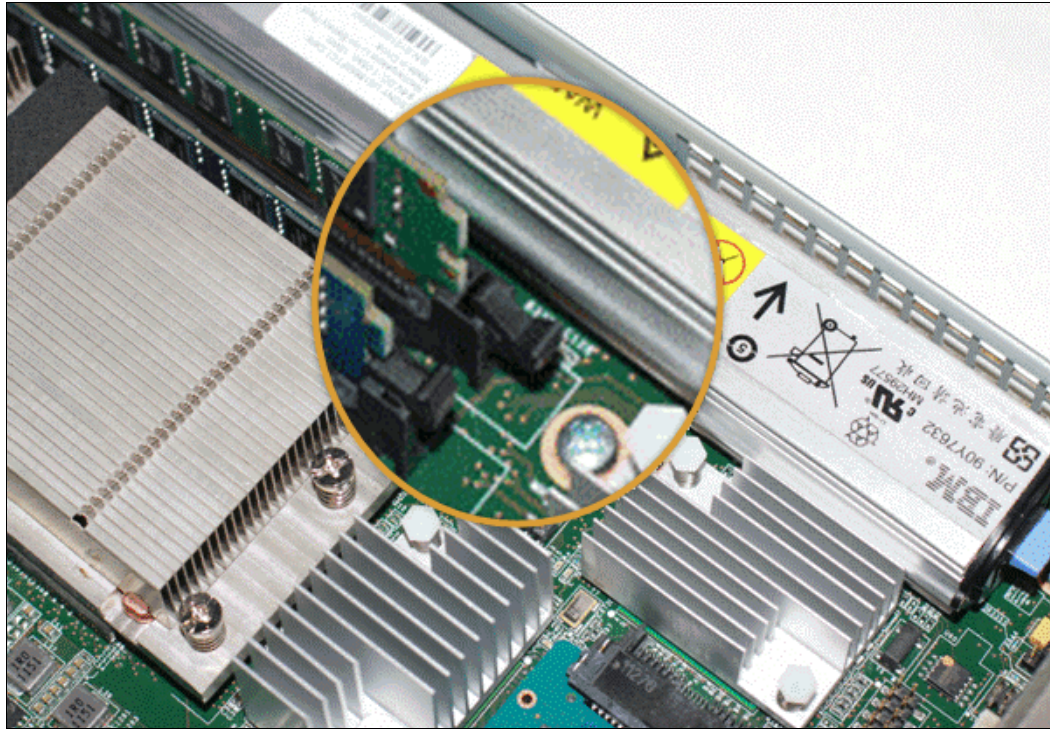


Figure 12-11 Memory replacement

Battery Backup Unit replacement

Caution: The battery is a lithium ion battery. To avoid possible explosion, do not incinerate the battery. Exchange the battery only with the part that is approved by IBM.

Because the Battery Backup Unit (BBU) is seated in the node canister, the BBU replacement leads to a redundancy loss until the replacement is completed. Therefore, it is recommended to replace the BBU only when advised to do so. It is also recommended to follow the Directed Maintenance Procedures (DMP).

For more information about how to replace the BBU, see the Information Center at this website:

http://pic.dhe.ibm.com/infocenter/storwize/V5000_ic/topic/com.ibm.storwize.V5000.641.doc/V5000_rplc_batt_nodecan.html

At the website, browse to **Troubleshooting Removing and replacing parts** → **Replacing battery in a node canister**.

Complete the following steps to replace the BBU:

1. Grasp the blue touch points on each end of the battery, as shown in Figure 12-12.

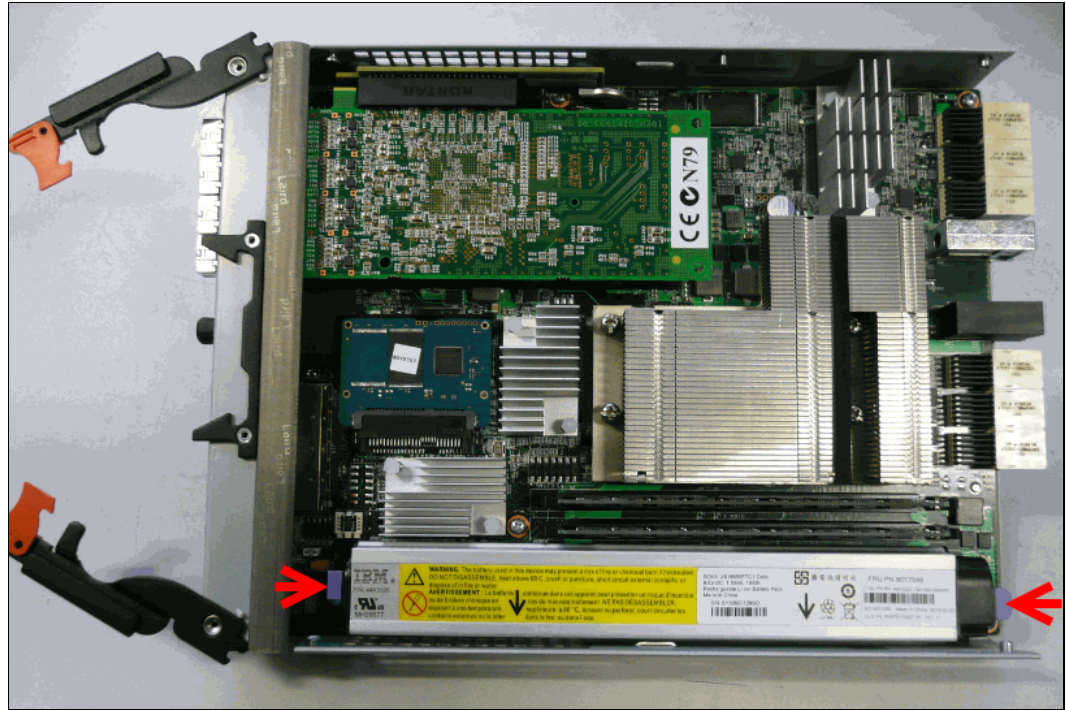


Figure 12-12 BBU replacement: Step 1

2. Lift the battery vertically upwards until the connectors disconnect.

Important: During a BBU change, the battery must be kept parallel to the canister system board while it is removed or replaced, as shown in Figure 12-13. Keep equal force, or pressure, on each end.



Figure 12-13 BBU replacement: Step 2

12.2.4 Expansion canister: Ports and LED

There are two expansion canister slots along with top of the unit.

SAS ports

SAS ports are used to connect the expansion canister to the node canister or to an extra expansion in the chain. Figure 12-14 shows the SAS ports that are on the expansion canister.



Figure 12-14 Expansion canister SAS ports

The meaning of the SAS port LEDs is described in Table 12-5.

Table 12-5 SAS LED status meaning

State	Meaning
Green	Indicates at least one of the SAS lanes on these connectors are operational. If the light is off when connected, there is a problem with the connection.
Amber	If the light is on, one of the following errors occurred: <ul style="list-style-type: none">▶ One or more (but not all) of the four lanes are up for this connector (if none of the lanes are up, the activity light is off).▶ One or more of the up lanes are running at a different speed to the others.▶ One or more of the up lanes are attached to a different address to the others.

Canister status

Each expansion canister has its status displayed by three LEDs, as shown in Figure 12-15.



Figure 12-15 Enclosure canister status

The LED status is described in Table 12-6.

Table 12-6 Enclosure canister status

Color	Name	Definition
Green (left)	Power	Indicates that the canister is receiving power.
Green (mid)	Status	If the light is on, the canister is running normally. If the light is flashing, there is an error communicating with the enclosure.
Amber	Fault	If the light is solid, there is an error logged against the canister or the firmware is not running.

12.2.5 Disk subsystem

The IBM Storwize V5000 disk subsystem is made up of control and expansion enclosures. The system can have one or two control enclosures, with each control enclosure attaching to up to six expansion enclosures. Each enclosure contains the drives that are based on the enclosure type.

This section describes the parts of the disk subsystem.

SAS cabling

Expansion enclosures are attached to control enclosures by using SAS cables. There are two supported SAS chains and up to three expansion enclosures can be attached to each chain. The node canister uses SAS ports 3 and 4 for enclosures while ports 1 and 2 for host connectivity.

Important: When an SAS cable is inserted, ensure that the connector is oriented correctly by confirming that the following conditions are met:

- ▶ The pull tab must be below the connector.
- ▶ Insert the connector gently until it clicks into place. If you feel resistance, the connector is probably oriented the wrong way. Do not force it.
- ▶ When inserted correctly, the connector can be removed only by pulling the tab.

The expansion canister has SAS port 1 for channel input and SAS port 2 for output to connect another expansion enclosure.

The SAS cabling is shown in Figure 12-16.

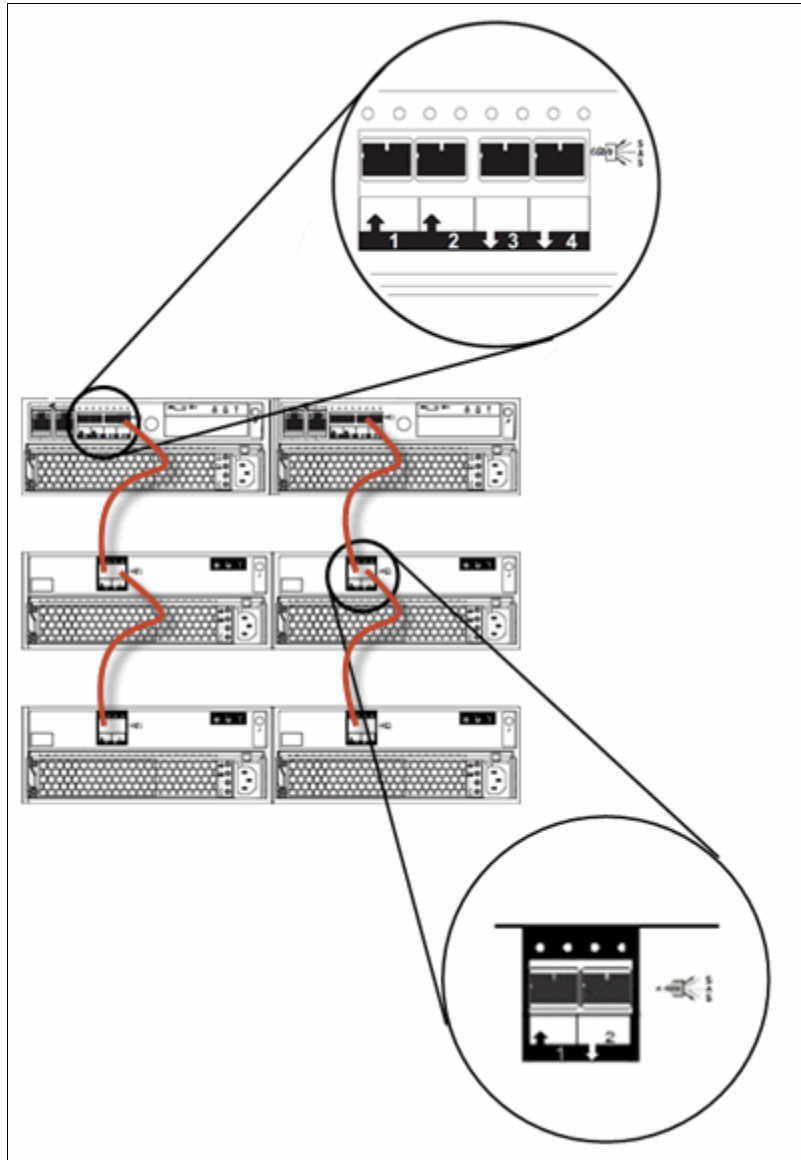


Figure 12-16 SAS cabling for single I/O Group)



A strand starts with an SAS initiator chip inside an IBM Storwize V5000 node canister and progresses through SAS expanders, which connect disk drives. Each canister contains an expander. Each drive has two ports, each of which is connected to a different expander and strand. This configuration means both nodes directly access each drive and there is no single point of failure.

At system initialization when devices are added to or removed from strands (and at other times), the IBM Storwize V5000 Software performs a discovery process to update the state of the drive and enclosure objects.

Slot numbers in enclosures

The IBM Storwize V5000 is made up of enclosures. There are four types of enclosures, as described in Table 12-7.

Table 12-7 Enclosure slot numbering

Enclosure type	Number of slots
Enclosure 12x 3.5-inch drives: <ul style="list-style-type: none"> ▶ Control enclosure 2077-12C ▶ Expansion enclosure 2077-12E 	Enclosure with 12 slots. 
Enclosure 24x 2.5-inch drives: <ul style="list-style-type: none"> ▶ Control enclosure 2077-24C ▶ Expansion enclosure 2077-24E 	Enclosure with 24 slots. 

Array goal

Each array has a set of goals that describe the wanted location and performance of each array member. A sequence of drive failures and hot spare takeovers can leave an array unbalanced; that is, with members that do not match these goals. The system automatically rebalances such arrays when appropriate drives are available.

RAID level

An IBM Storwize V5000 supports the RAID 0, RAID 1, RAID 5, RAID 6, or RAID 10. Each RAID level is described in Table 12-8.

Table 12-8 RAID levels that are supported by an IBM Storwize V5000

RAID level	Where data is striped	Drive count (Min - Max)
0	Arrays have no redundancy and do not support hot-spare takeover.	1 - 8
1	Provides disk mirroring, which duplicates data between two drives. A RAID 1 array is internally identical to a two-member RAID 10 array.	2
5	Arrays stripe data over the member drives with one parity strip on every stripe. RAID 5 arrays have single redundancy with higher space efficiency than RAID 10 arrays, but with some performance penalty. RAID 5 arrays can tolerate no more than one member drive failure.	3 - 16
6	Arrays stripe data over the member drives with two parity strips on every stripe. A RAID 6 array can tolerate any two concurrent member drive failures.	5 - 16

RAID level	Where data is striped	Drive count (Min - Max)
10	Arrays stripe data over mirrored pairs of drives. RAID 10 arrays have single redundancy. The mirrored pairs rebuild independently. One member out of every pair can be rebuilding or missing at the same time. RAID 10 combines the features of RAID 0 and RAID 1.	2 - 16

Disk scrubbing

The scrub process runs when arrays do not have any other background processes. The process checks that the drive logical block addresses (LBAs) are readable and array parity is synchronized. Arrays are scrubbed independently and each array is entirely scrubbed every seven days.

Solid-state drives

Solid-state drives (SSDs) are treated no differently by an IBM Storwize V5000 than hard disk drives (HDDs) concerning RAID arrays or MDisks. The SSDs in the storage that are managed by the IBM Storwize V5000 are combined into an array, usually in RAID 10 or RAID 5 format. It is unlikely that RAID 6 SSD arrays are used because of the double parity impact, with two SSD logical drives that are used for parity only.

12.2.6 Power supply unit

All enclosures require two power supply units (PSUs) for normal operation. A single PSU can power the entire enclosure for redundancy.

Figure 12-17 shows the power supplies.

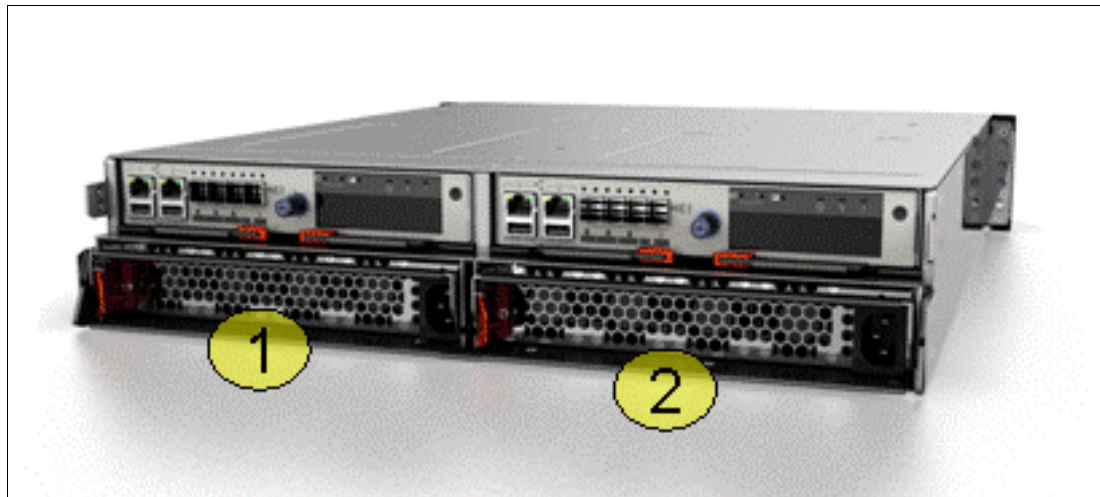


Figure 12-17 Power supply

The left side PSU is numbered 1 and the right side PSU is numbered 2.

PSU LED indicator

The indicators are the same for the control and expansion unit.

Figure 12-18 shows the PSU LED Indicators.



Figure 12-18 PSU LED Indicators

Table 12-9 shows the colors and meaning of the LEDs.

Table 12-9 PSU LED definitions

Position	Color	Marking	Name	Definition
1	Green	In	AC Status	Main power is delivered
2	Green	DC	DC Status	DC power is available
3	Amber	Fault exclamation mark	Fault	Fault on PSU
4	Blue	OK	Service action that is allowed	N/A

12.3 Configuration backup procedure

If there is a serious failure that requires the system configuration must be restored, the configuration backup file must be used. The file contains configuration data such as, arrays, pools, and volumes (but no customer applications data). The backup file is updated by the cluster every day.

Even so, it is important to save the file after you change your system configuration, which requires a command-line interface (CLI) connection to start manual backup.

Regularly saving a configuration backup file on the IBM Storwize V5000 is important and it must be done manually. Download this file regularly to your management workstation to protect the configuration data (a best practice is to automate this download procedure by using a script and saving it daily on a remote system).

12.3.1 Generating a configuration backup by using the CLI

To generate a configuration backup by using the CLI, run the **svconfig backup** command, as shown in Example 12-1.

Example 12-1 Example for backup CLI command

svconfig backup

The progress of the command is detected by dots, as shown in Example 12-2.

Example 12-2 Backup CLI command progress and output

.....
.....
.....
CMMVC6155I SVCCONFIG processing completed successfully

The **svconfig backup** command creates three files that provide information about the backup process and cluster configuration. These files are created in the /tmp directory on the configuration node and are listed on the support view.

The three files that are created by the backup process are described Table 12-10.

Table 12-10 File names that are created by the backup process

File name	Description
svc.config.backup.xml	This file contains your cluster configuration data.
svc.config.backup.sh	This file contains the names of the commands that were issued to create the backup of the cluster.
svc.config.backup.log	This file contains details about the backup, including any error information that might be reported.

12.3.2 Downloading a configuration backup by using the GUI

To download a configuration backup file by using the GUI, complete the following steps:

1. Click the **Settings** icon and then click **Support**, as shown in Figure 12-19.



Figure 12-19 Configuration backup open support view

2. Select the configuration node on the support view, as shown in Figure 12-20.



Figure 12-20 Configuration backup select configuration node

3. Select the **Show full log listing...** option (as shown in Figure 12-21) to list all of the available log files that are stored on the configuration node.

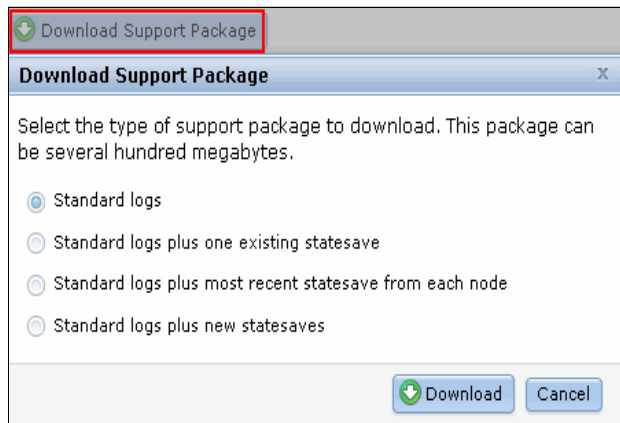


Figure 12-21 Support package selection

4. Search for a file named `/dumps/svc.config.backup.xml_*`, as shown in Figure 12-22. Select the file, right-click it, and then select **Download**.

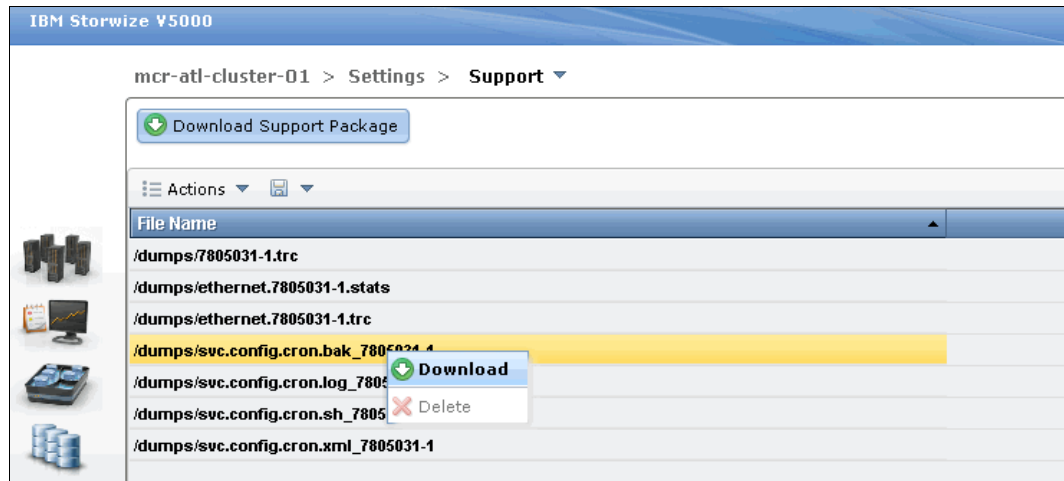


Figure 12-22 Configuration backup start download

5. Save the configuration backup file on your management workstation where it can be found easily, as shown in Figure 12-23.

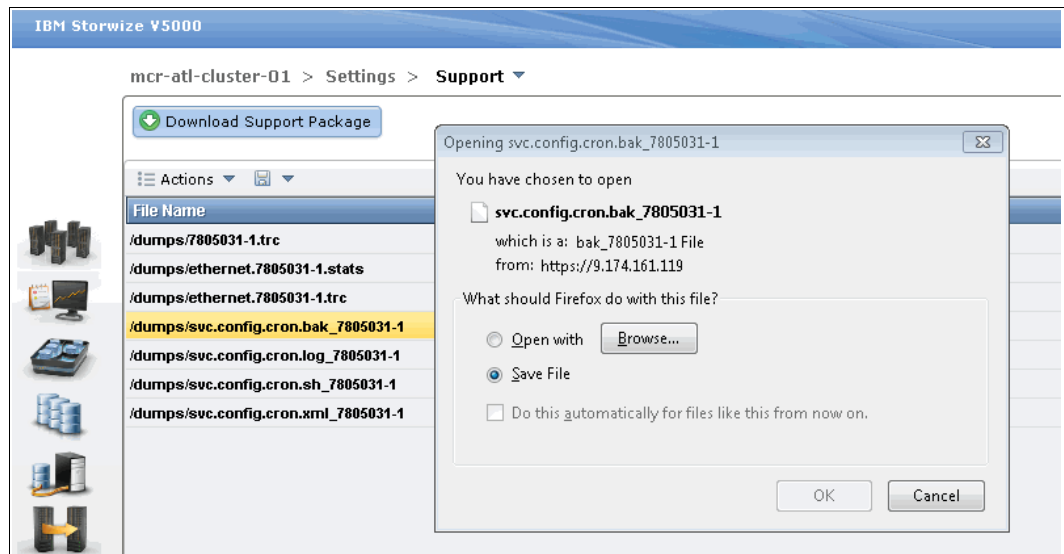


Figure 12-23 Configuration backup save file

Even if the configuration backup file is updated automatically, it might be of interest to verify the time stamp of the actual file. Therefore, the /dumps/svc.config.backup.xml_xx file must be opened with an editor, such as, WordPad, as shown in Figure 12-24.

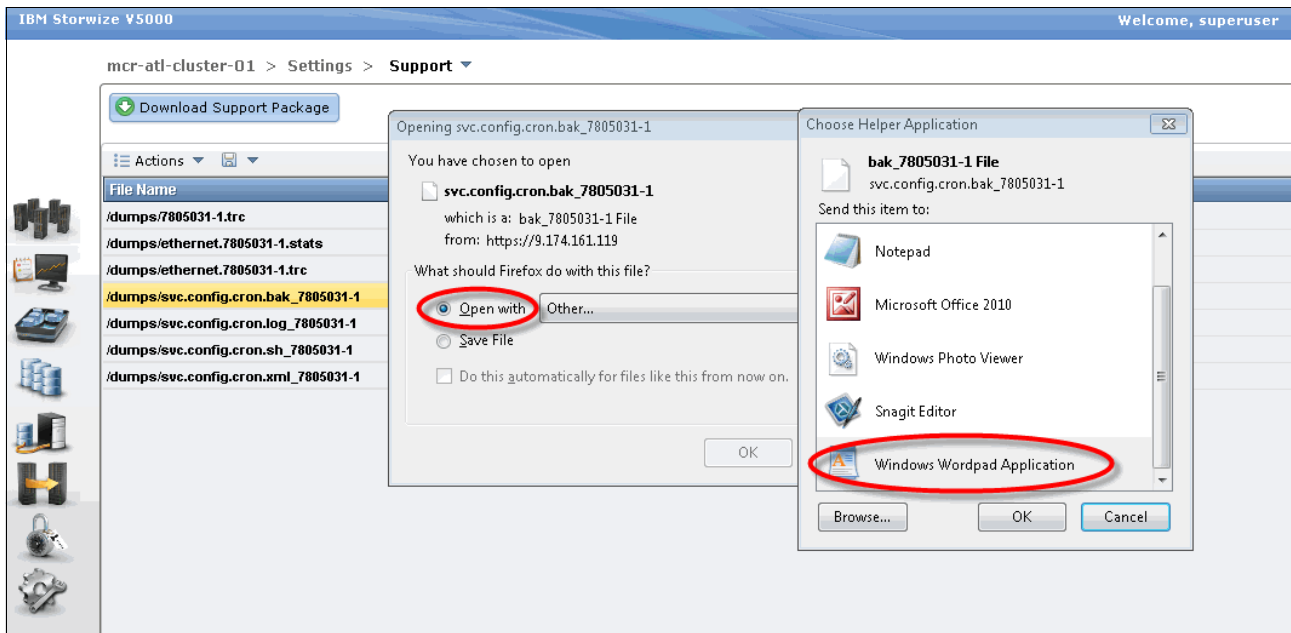


Figure 12-24 Open backup XML file with WordPad

Open the /dumps/svc.config.backup.xml_xx file with an editor (we used WordPad) and search for the string timestamp=, which is found near of the top of the file. Figure 12-25 shows the file that is opened and the time stamp information in it.

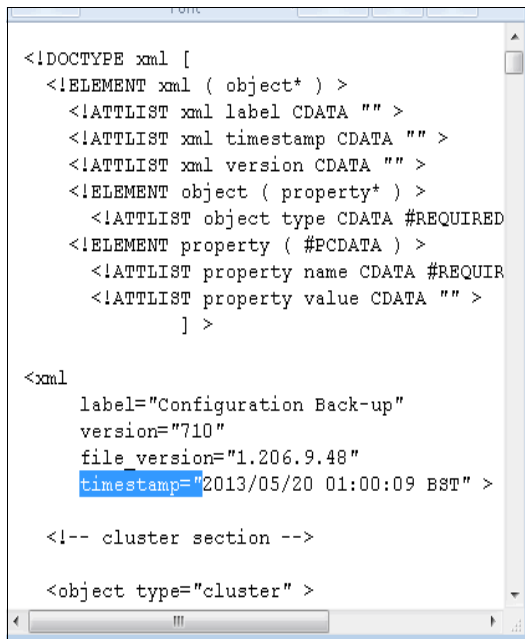


Figure 12-25 Timestamp in backup XML file

12.4 Upgrading software

The system upgrade process involves the upgrading of your entire IBM Storwize V5000 environment.

Allow sufficient time to plan your tasks, review your preparatory upgrade tasks, and complete the upgrade of the IBM Storwize V5000 environment. The upgrade procedures can be divided into these general processes. Table 12-11 shows the software upgrade tasks.

Table 12-11 Software upgrade tasks

Sequence	Upgrade tasks
1	Decide whether you want to upgrade automatically or manually. During an automatic upgrade procedure, the clustered system upgrades each of the nodes systematically. The automatic method is the preferred procedure for upgrading software on nodes. However, you can upgrade each node manually.
2	Ensure that CIM object manager (CIMOM) clients are working correctly. When necessary, upgrade these clients so that they can support the new version of IBM Storwize V5000 code.
3	Ensure that multipathing drivers in the environment are fully redundant. If you experience failover issues with multipathing driver support, resolve these issues before you start normal operations.
4	Upgrade other devices in the IBM Storwize V5000 environment. Examples might include upgrading hosts and switches to the correct levels.
5	Upgrade your IBM Storwize V5000.

Important: The amount of time it takes to perform an upgrade can vary depending on the amount of preparation work that is required and the size of the environment. Generally, allow more than two hours for an upgrade if you have two I/O Groups.

Some code levels support upgrades only from specific previous levels. If you upgrade to more than one level above your current level, you might be required to install an intermediate level.

Important: Ensure that you have no unfixed errors in the log and that the system date and time are correctly set. Start the fix procedures, and ensure that you have fix any outstanding errors before you attempt to concurrently upgrade the code.

12.4.1 Upgrading software automatically

During the automatic upgrade process, each node in the system upgrades individually and the new code is staged on the nodes. While each node restarts, there might be some degradation in the maximum I/O rate that can be sustained by the system. After all of the nodes in the system are successfully restarted with the new code level, the new level is automatically committed.

The upgraded node is temporarily unavailable and all I/O operations fail to that node. As a result, the I/O error counts increase and the failed I/O operations are directed to the partner node of the working pair. Applications do not see any I/O failures. When new nodes are added to the system, the upgrade package is automatically downloaded to the new nodes from the IBM Storwize V5000 system.

The upgrade can be performed concurrently with normal user I/O operations. However, there is a possibility that performance might be affected.

Multipathing requirement

Before you upgrade, ensure that the multipathing driver is fully redundant with every path that is available and online. You might see errors that are related to the paths, which go away (failover) and the error count that increases during the upgrade. When the paths to the nodes return, the nodes fall back to become a fully redundant system. After an approximate 30-minute delay, the paths to the other node fail.

12.4.2 GUI upgrade process

The automatic upgrade process is started in the GUI by starting the Upgrade wizard, as shown in Figure 12-26. Browse to **Settings** → **General** → **Upgrade Software** → **Launch Upgrade wizard**.

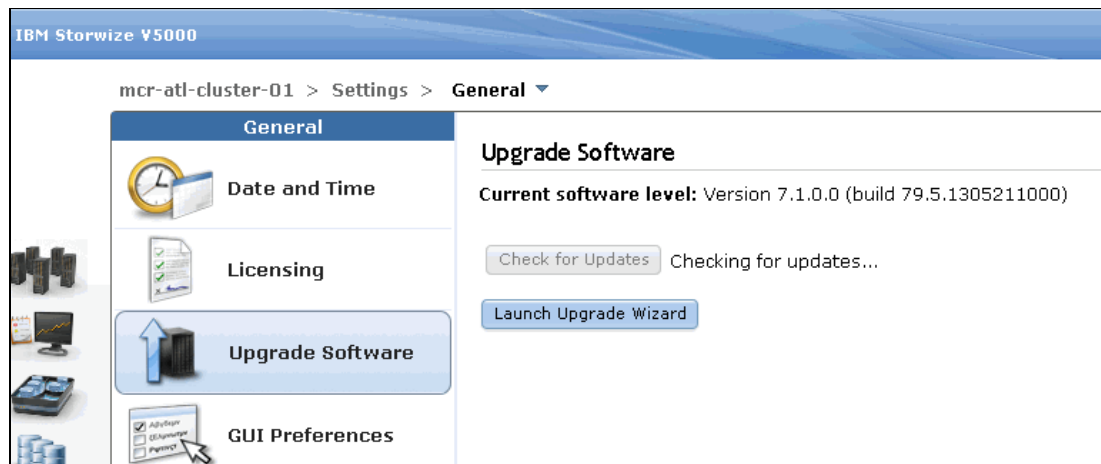


Figure 12-26 Start Upgrade wizard

As a first step, the Upgrade test utility must be downloaded from the Internet (the link is provided within the panel). If the tool was downloaded and stored on the management station, it can be uploaded, as shown in Figure 12-27.

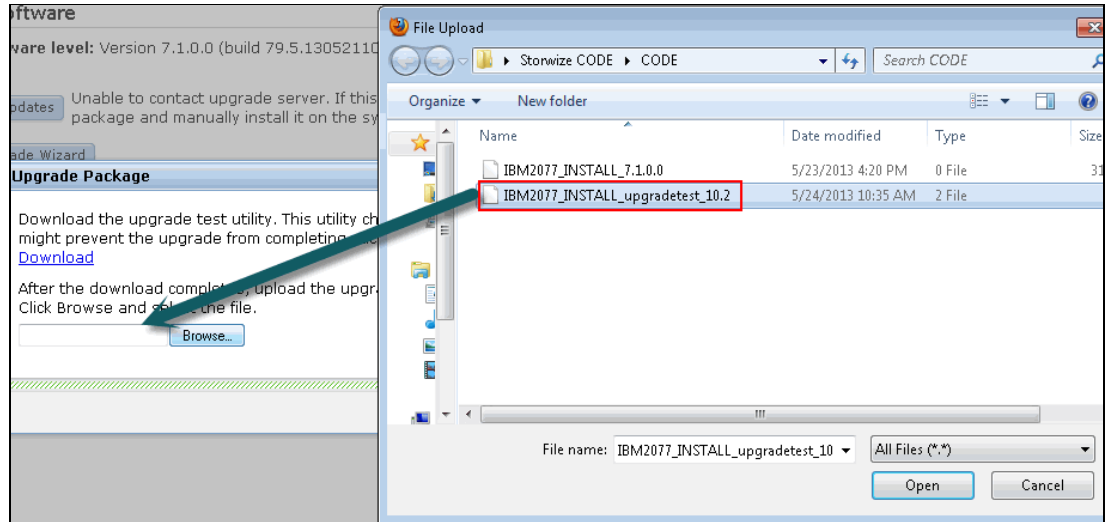


Figure 12-27 Download Upgrade test utility

A confirmation panel opens, as shown in Figure 12-28.

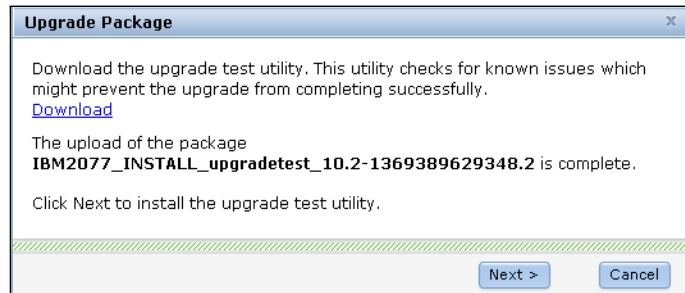


Figure 12-28 Upload test utility completed

The version to which the system should be upgraded must be entered in step 2 of the wizard. By default, the latest code level (at the time of writing) is shown, as shown in Figure 12-29.

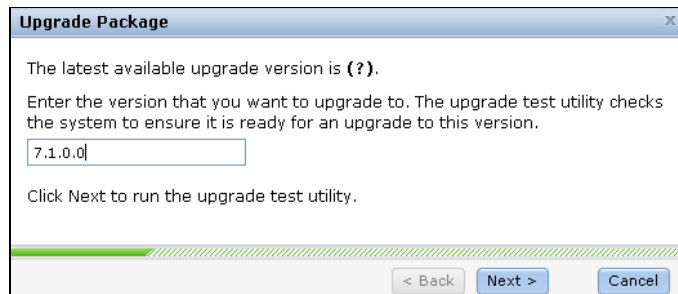


Figure 12-29 Enter version to be checked by tool

Important: You must choose the correct code level because you cannot recheck this information later. The version that is selected is used throughout the rest of the process.

Figure 12-30 shows the panel that indicates the background test task is running.

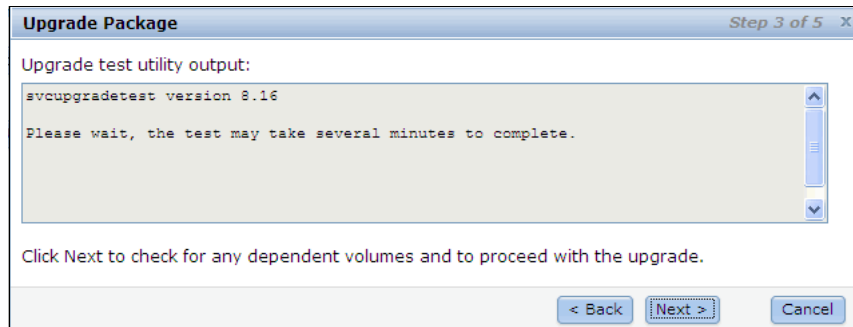


Figure 12-30 Wait utility to complete

The utility can be run as many times as necessary on the same system to perform a readiness check in preparation for a software upgrade.

Next, the code must be downloaded. If the code was downloaded to the management station, it can be directly uploaded, as shown in Figure 12-31. Verify that the correct code file is used.

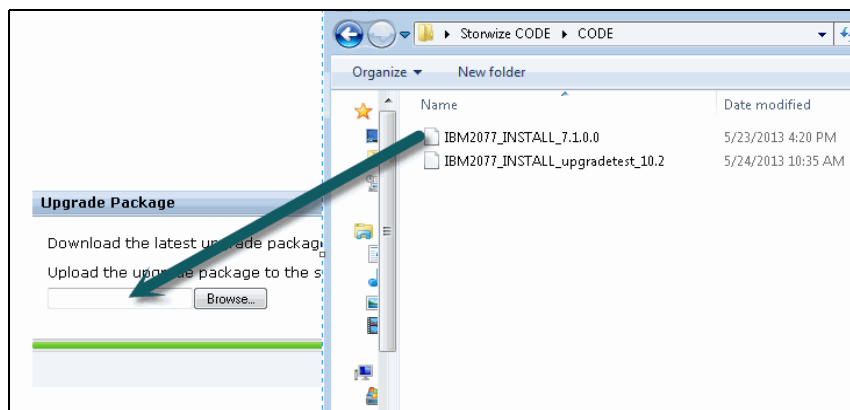


Figure 12-31 Download code

As shown in Figure 12-32, a confirmation window opens.

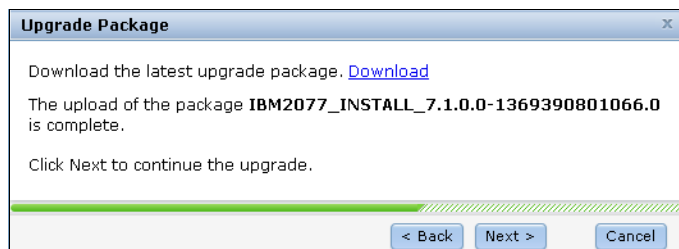


Figure 12-32 Code upload that is completed

The automated code upgrade can be started when the Automatic upgrade option is selected in the panel, as shown in Figure 12-33 (this is the default choice). If the upgrade is done manually for any reason, the selection must be made; however, an automatic upgrade is recommended.

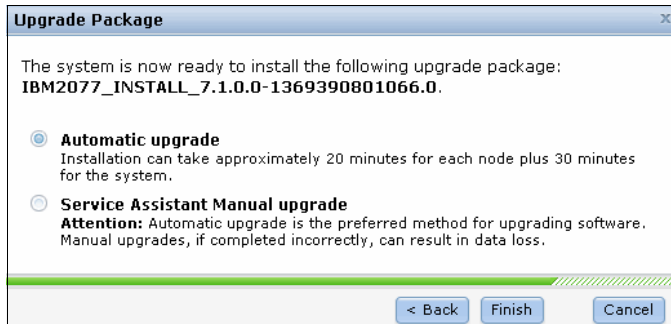


Figure 12-33 Upload mode decision

If you choose to select the **Service Assistant Manual** upgrade option, see 12.4.3, “Upgrading software manually” on page 584.

Select **Finish** to start the upgrade process on the nodes. Messages inform you when the nodes are upgraded. When all nodes are rebooted, the upgrade process is complete. It can take up to two hours to finish this process.

12.4.3 Upgrading software manually

Important: It is highly recommended to upgrade the IBM Storwize V5000 automatically by following the Upgrade wizard. If a manual upgrade is used, make sure that you do not skip any step.

The steps for manual upgrade are shown on the Service Assistant Manual Upgrade panel.

Complete the following steps to manually upgrade the software:

1. In the management GUI, click **Settings** → **General** → **Upgrade Software** and run the Upgrade wizard. In step 5 of the wizard, select **Service Assistant Manual upgrade**, as shown in Figure 12-34.

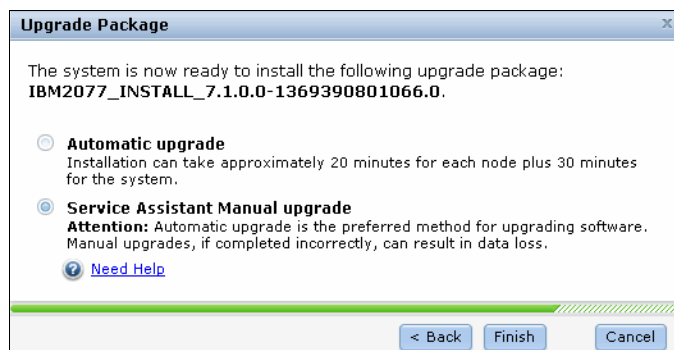


Figure 12-34 Select manual upgrade mode

After you select manual upgrade, a warning appears, as shown in Figure 12-35.

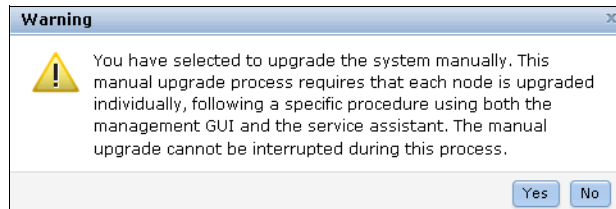


Figure 12-35 Manual upgrade warning

Both nodes are set to “Waiting for Upgrade” status in the Upgrade Machine Code panel, as shown in Figure 12-36.

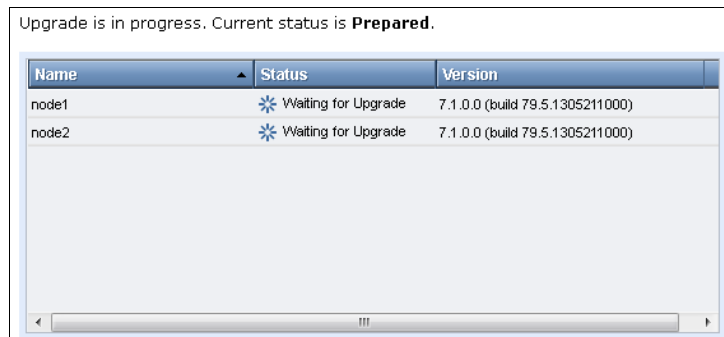


Figure 12-36 Node status to waiting for upgrade

2. In the management GUI, select **System Details** and select the canister that contains the node you want to upgrade next. As shown in Figure 12-37, select **Remove Node** in the Action menu, which shows you a Health Status alert.

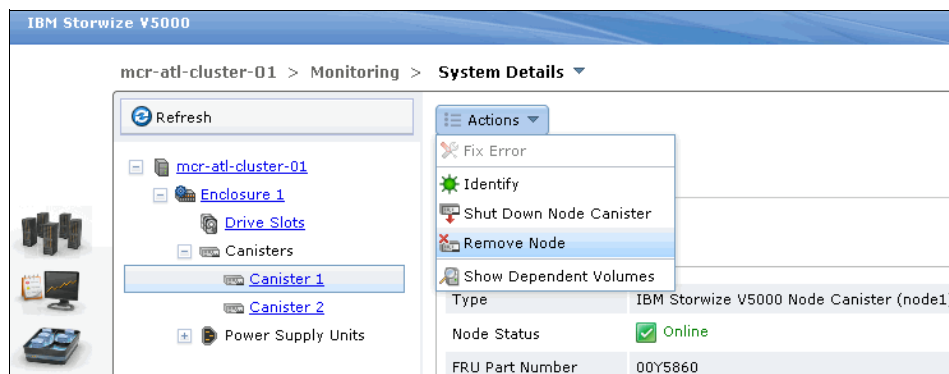


Figure 12-37 Remove the non-config node from cluster

Important: Make sure that you select the non-config node first.

A warning message appears, as shown in Figure 12-38.

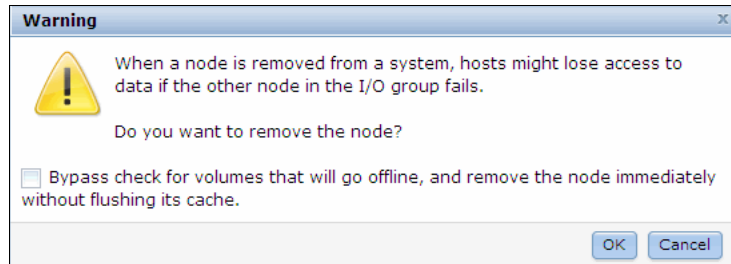


Figure 12-38 Remove node warning message

The non-configuration node is removed from GUI Upgrade Machine Code panel, as shown in Figure 12-39.

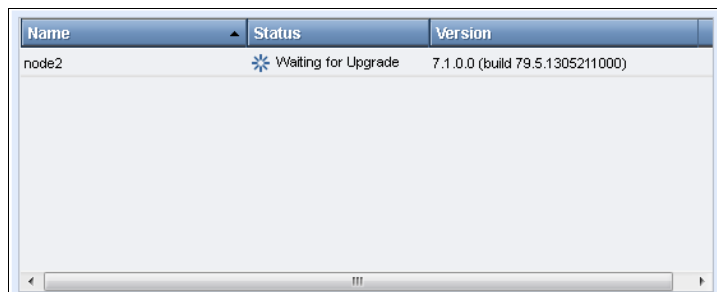


Figure 12-39 Non-configuration node was removed

In the System Details panel, the node is shown as Unconfigured, as shown in Figure 12-40.

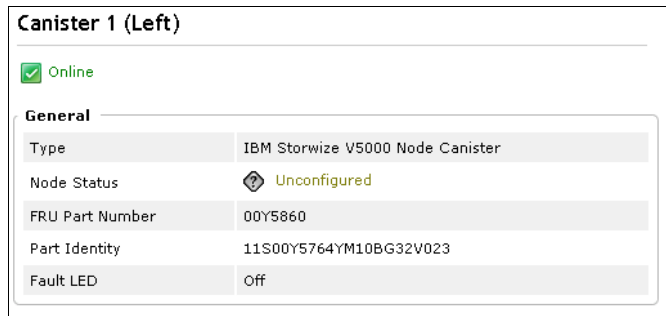


Figure 12-40 Node status shows unconfigured

3. In the Service Assistant panel, the node that is ready for upgrade must be selected. Select the node that shows Node status as service mode and has no available cluster information, as shown in Figure 12-41.

Change Node						
Node Name	Node Status	Error	Panel	System	Relationship	
node2	Active	715 1 1 0	01-2	mcr-atl-cluster-01	Local	
	Service	690	7805031-1		Partner	

Figure 12-41 Select node in service mode for upgrade

- In the Service Assistant panel, select **Upgrade Manually** and select the machine code version that you want to upgrade on selected node, as shown in Figure 12-42.

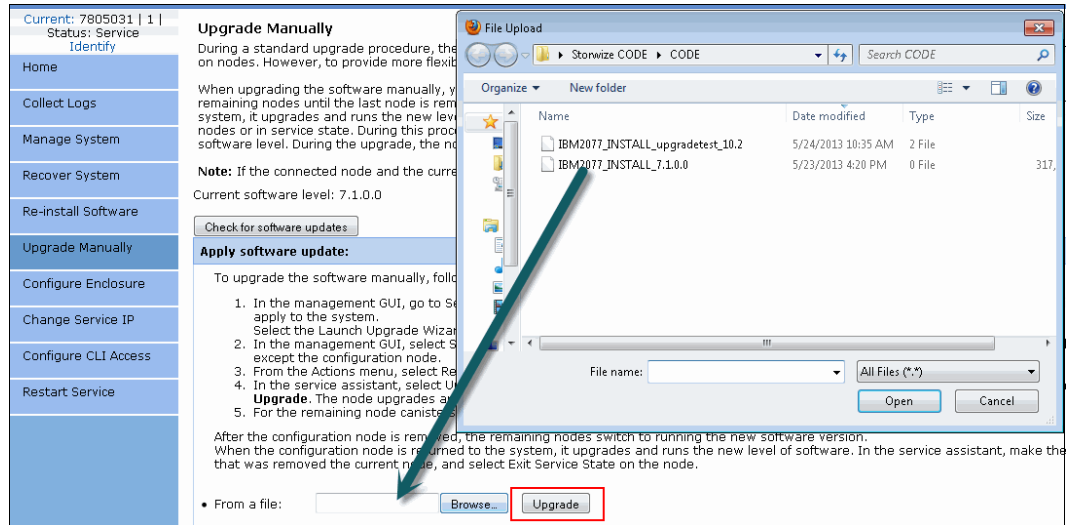


Figure 12-42 Select machine code file for upgrade

- Click **Upgrade** to start the upgrade process on the first node.

The node is added automatically into the system after upgrade. Upgrading and adding the node again can take up to 30 minutes, as shown in Figure 12-43.

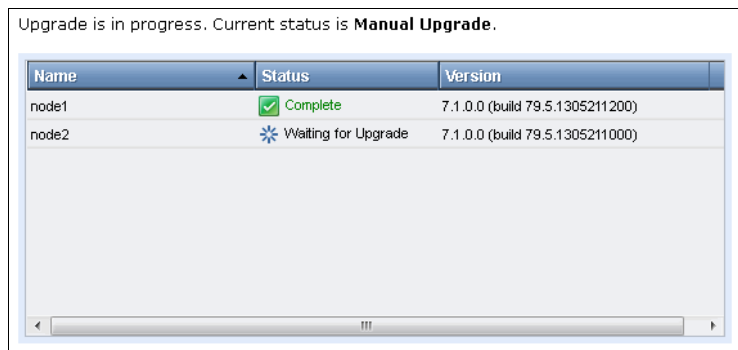


Figure 12-43 Non-config node completed upgrade

- Repeat steps 2 - 4 for the remaining node (or nodes).

After you remove the configuration node from the cluster for upgrade, a warning appears, as shown in Figure 12-44.

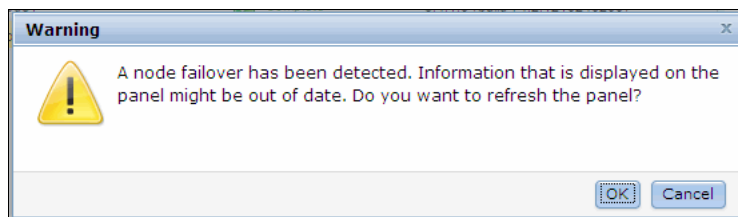


Figure 12-44 Configuration node failover warning

Important: The configuration node remains in Service State when it is added again to the cluster. Therefore, exit Service State manually.

- To exit from service state, browse to the home panel of the Service Assistant and open the Action menu. Select **Exit Service State**, as shown in Figure 12-45.



Figure 12-45 Exit service state to add node back in cluster

Both the nodes are now back in the cluster (as shown in Figure 12-46) and the system is running on the new code level.

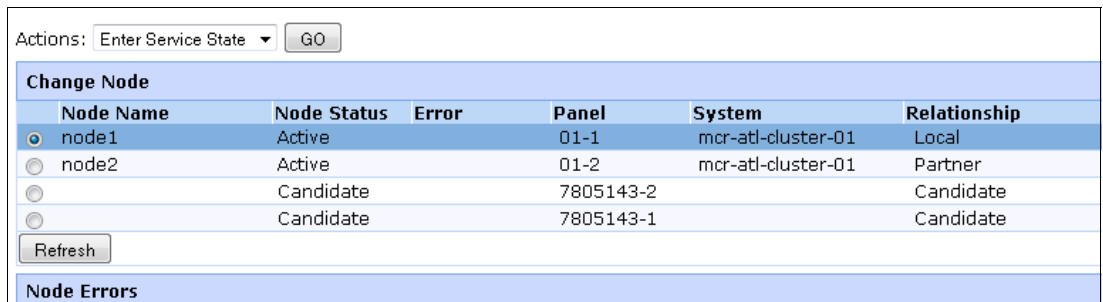


Figure 12-46 Cluster is active again and running new code level

12.5 Event log

Whenever a significant change in the status of IBM Storwize V5000 is detected, an event is submitted to the event log.

All events are classified as *alerts* or *messages*.

An alert is logged when the event requires some action. Some alerts have an associated error code that defines the service action that is required. The service actions are automated through the fix procedures. If the alert does not have an error code, the alert represents an unexpected change in state. This situation must be investigated to see whether it is expected or represents a failure. Investigate an alert and resolve it when it is reported.

A message is logged when a change that is expected is reported; for instance, an IBM FlashCopy operation completes.

The event log panel can be opened via the GUI by clicking **Monitoring** → **Events**, as shown in Figure 12-47.



Figure 12-47 Open eventlog panel

Figure 12-48 shows the event log.

A screenshot of the IBM Storwize V5000 GUI showing the 'Events' view. The breadcrumb navigation is 'mcr-atl-cluster-01 > Monitoring > Events'. A 'Next Recommended Action' box is visible, containing the text 'SAS host ports not operational', 'Error Code: 1667, Event ID: 073715', and '84 minutes ago', with a 'Run This Fix Procedure' button. Below this is a table of events.

Error Code	Last Time Stamp	Status	Description	Object Type	Object ID	Object Name
	5/24/13 8:13:47 AM	Message	SAS discovery occurred, configuration changes complete	io_grp	0	io_grp0
	5/24/13 8:13:42 AM	Message	One node in the IO group is unavailable	node	6	
	5/24/13 8:13:42 AM	Message	SAS discovery occurred, configuration changes pending	io_grp	0	io_grp0
	5/24/13 8:09:42 AM	Message	SAS discovery occurred, configuration changes complete	io_grp	0	io_grp0

Figure 12-48 The event log view

12.5.1 Managing the event log

The event log features a size limit. After it is full, newer entries replace the older entries, which are not required.

To avoid a repeated event that fills the event log, some records in the event log refer to multiple occurrences of the same event. When event log entries are coalesced in this way, the time stamp of the first occurrence and the last occurrence of the problem is saved in the log entry. A count of the number of times that the error condition occurred also is saved in the log entry. Other data refers to the last occurrence of the event.

Event log panel columns

Right-clicking in any column header opens the option menu in which you can select columns that are shown or hidden.

Figure 12-49 shows all of the possible columns that can be displayed in the error log view.

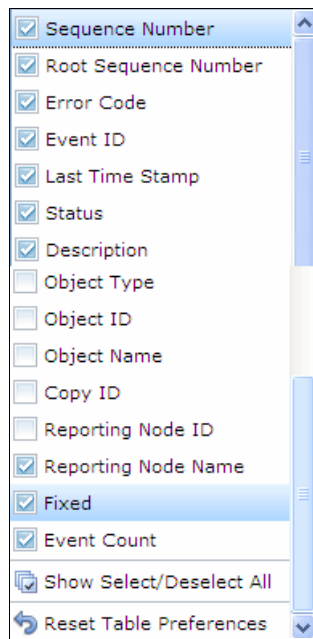


Figure 12-49 Possible event log columns

The following available fields are recommended at a minimum to assist you in diagnosing problems:

- ▶ Event ID

This number precisely identifies the reason why the event was logged.

- ▶ Error code

This number describes the service action that should be followed to resolve an error condition. Not all events have error codes that are associated with them. Many event IDs can have the same error code because the service action is the same for all of the events.

- ▶ Sequence number

A number that identifies the event.

- ▶ **Event count**
The number of events that are coalesced into this event log record.
- ▶ **Fixed**
When an alert is shown for an error condition, it indicates whether the reason for the event was resolved. In many cases, the system automatically marks the events that are fixed when appropriate. There are some events that must be manually marked as fixed. If the event is a message, this field indicates that you read and performed the action. The message must be marked as read.
- ▶ **Last time**
The time when the last instance of this error event was recorded in the log.
- ▶ **Root sequence number**
If set, this number is the sequence number of an event that represents an error that probably caused this event to be reported. Resolve the root event first.

Event log panel options

Figure 12-50 shows the main Event log panel options, which should be used to handle system events.

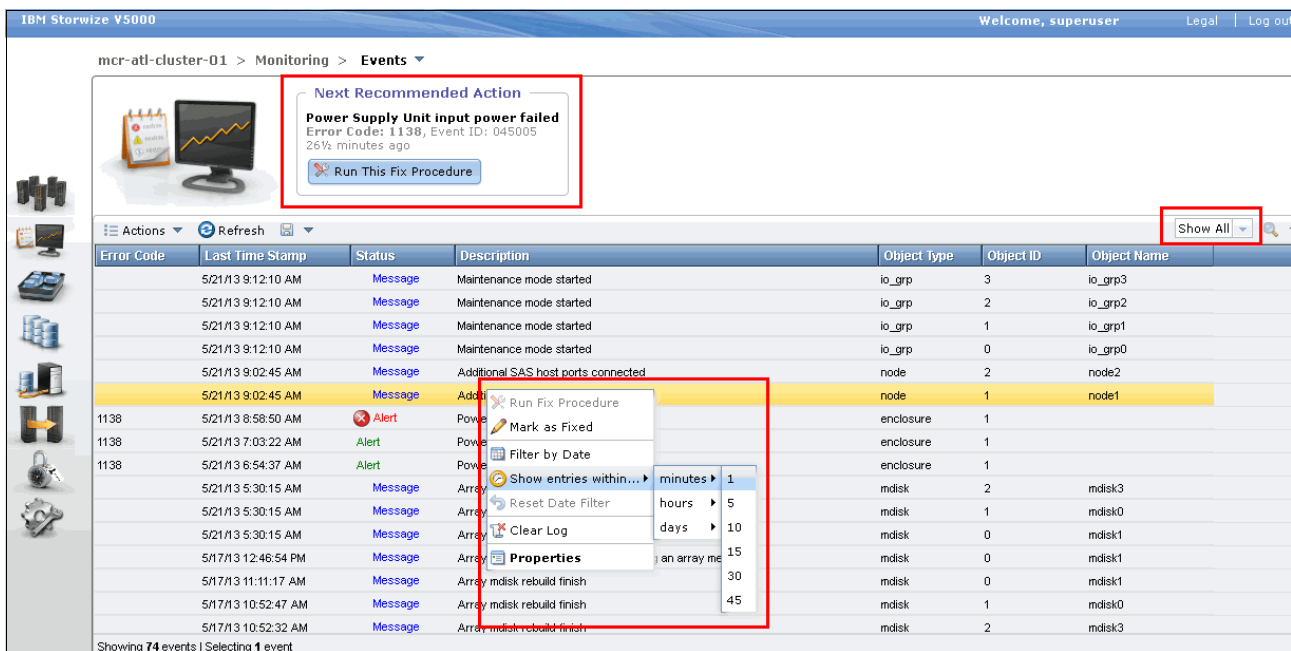


Figure 12-50 Eventlog Panel

Event log filter options

The following log filter options are available:

- ▶ **Show all**
This option lists all available events.
- ▶ **Unfixed Messages and Alerts**
This option lists unfixed events. This option is useful to find events that must be handled but no actions are required or recommended.

- ▶ Recommended Actions (default)

Only events with recommended actions (Status Alert) are displayed.

Important: Check for this filter option if no event is listed. There might be events that are not associated to recommended actions.

Figure 9-51 shows an event log with no items found, which does not necessarily mean that the event log is clear. To check whether the log is clear, use the filter option Show all.

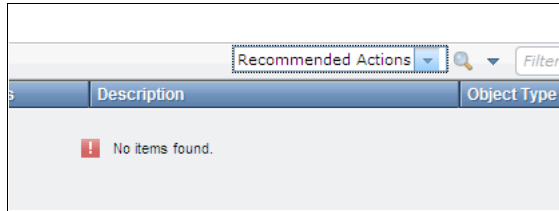


Figure 12-51 No items found in event log

Actions on single event

Right-clicking a single event gives the following options that might be used for that specific event:

- ▶ Mark as Fixed

It is possible to start the Fix Procedure on this specific event, even if it is not the recommended next action.

Some events, such as, messages, must be set to Mark as Fixed.

- ▶ Show entries within... minutes/hours/days

This option is to limit the error log list to a specific date or a time slot. The following selectable values are available:

- Minutes: 1, 5, 10, 15, 30, and 45
- Hours: 1, 2, 5, and 12
- Days: 1, 4, 7, 15, and 30

- ▶ Clear Log

This option clears the complete error log, even if only one event was selected.

Important: These actions cannot be undone and might prevent the system from being analyzed when severe problems occur.

- ▶ Properties

This option provides more sense data for the selected event that is shown in the list.

Recommended Actions

A fix procedure is a wizard that helps you to troubleshoot and correct the cause of an error. Some fix procedures reconfigure the system that is based on your responses, ensure that actions are carried out in the correct sequence, and prevent or mitigate loss of data. For this reason, you always must run the fix procedure to fix an error, even if the fix might seem obvious.

To run the fix procedure for the error with the highest priority, go to the Recommended Action panel at the top of the Event page and click **Run This Fix Procedure**. When you fix higher priority events first, the system often can automatically mark lower priority events as fixed.

For more information about how to run a DMP, see 12.5.2, “Alert handling and recommended actions” on page 593.

12.5.2 Alert handling and recommended actions

All events in Alert status require attention. Alerts are listed in priority order and should be fixed sequentially by using the available fix procedures.

Example: SAS cable fault

For this example, we created an error on one SAS cable connection between two expansions by removing the cable from one port.

The following example shows how faults are represented in the error log, how information about the fault can be gathered, and the recommended action (DMP) can be used to fix the error:

► Detect an alert

The Health Status indicator, which is permanently present on most of the GUI panel (for more information, see Chapter 3, “Graphical user interface overview” on page 75) is showing a yellow alert. Click the indicator to retrieve the specific information, as shown in Figure 12-52.

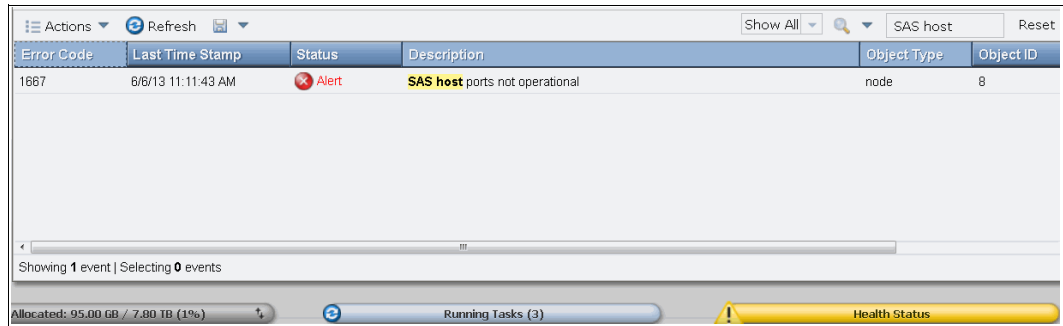


Figure 12-52 Health check shows degraded system status

Review the event log for more information.

► Find alert in event log

The default filter in the error log view is Recommended actions. This option lists the alert event only. Figure 12-53 shows the Next Recommended Action list.

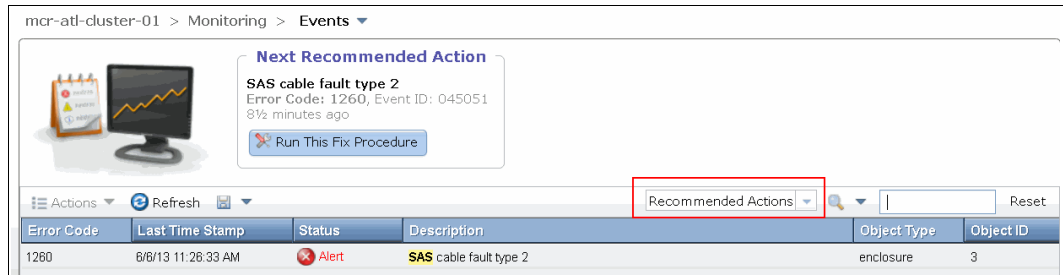


Figure 12-53 Next Recommended Action list

- Gather additional information: Show all

Find the events that are logged around the alert to understand what happened or find more information for better understanding and to find the original problem. Use the **Show all** filter to see all of the logged events, as shown in Figure 12-54.

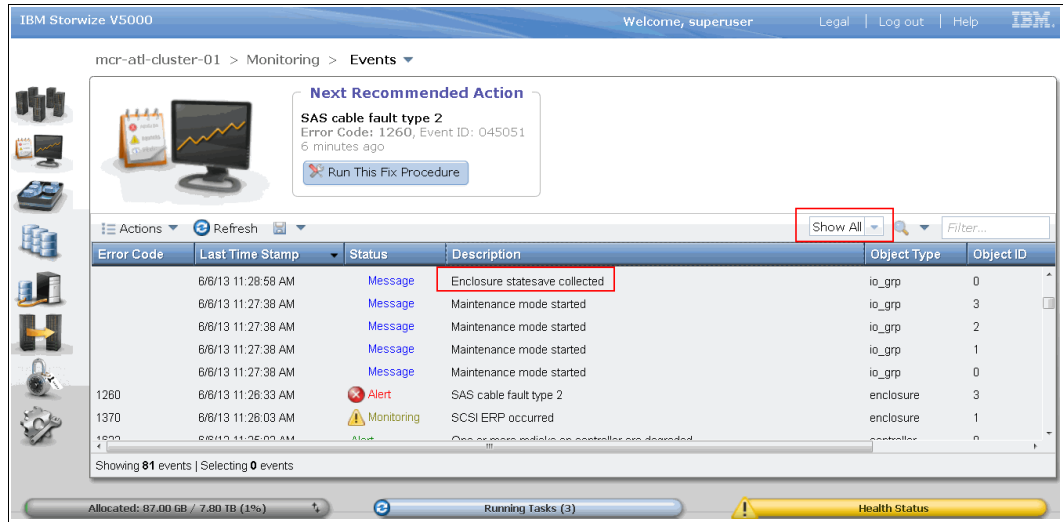


Figure 12-54 Show all events

- Gather additional information: Alert properties

More details about the event (for example, enclosure ID and canister ID) can be found in the properties option, as shown in Figure 12-55 on page 595. This information might be of interest for problem fixing or for root cause analysis.

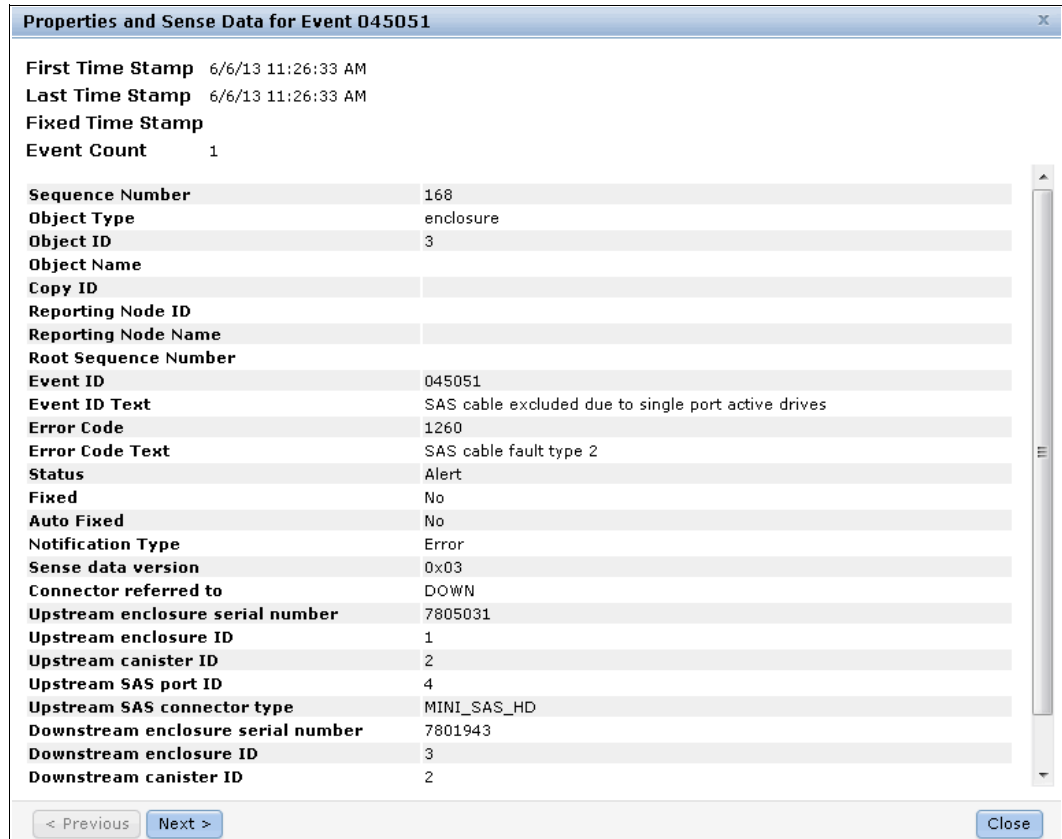


Figure 12-55 Alert properties

► Run recommended action (DMP)

It is highly recommended to fix alerts under the guidance of the recommended action by using the DMP. There are running tasks in the background that might be missed when the DMP is bypassed. Not all alerts have DMPs available.

To start the DMP, right-click the alert record or click **Run this fix procedure** at the top of the window.

The steps and panels of DMP are specific to the error that must be fixed. The following figures represent the recommended action (DMP) for the SAS cable event example.

Figure 12-56 shows step 1 of the DMP SAS cable event.

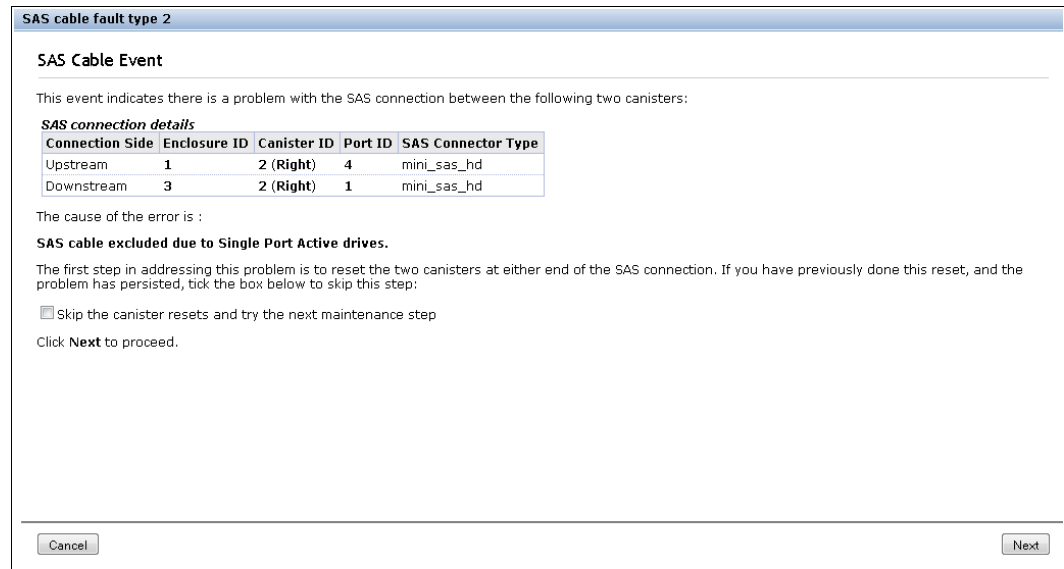


Figure 12-56 SAS cable Recommended action DMP step 1

Figure 12-57 shows step 2 of the DMP SAS cable event.

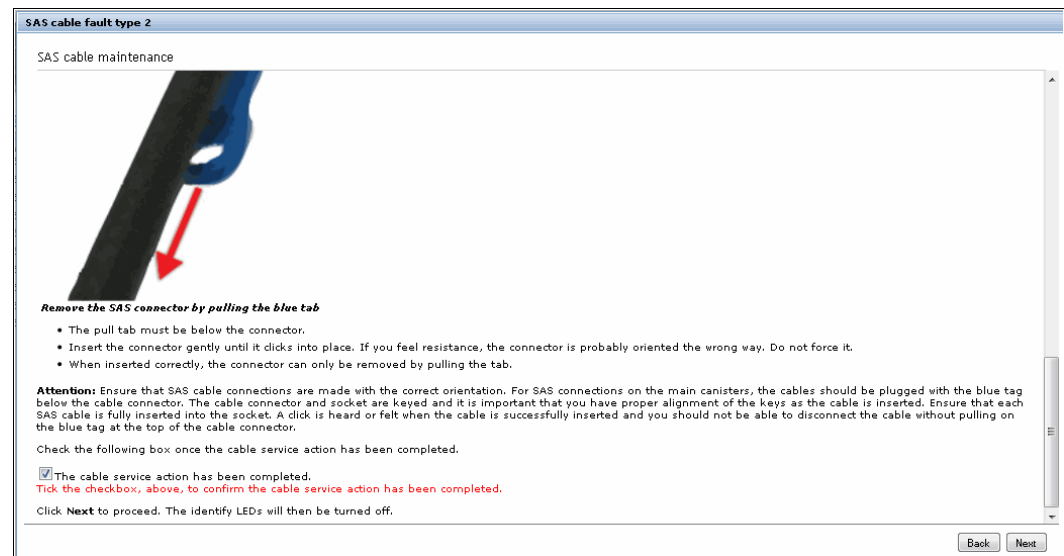


Figure 12-57 SAS cable Recommended action DMP step 2

Figure 12-58 shows step 3 of the DMP SAS cable event.

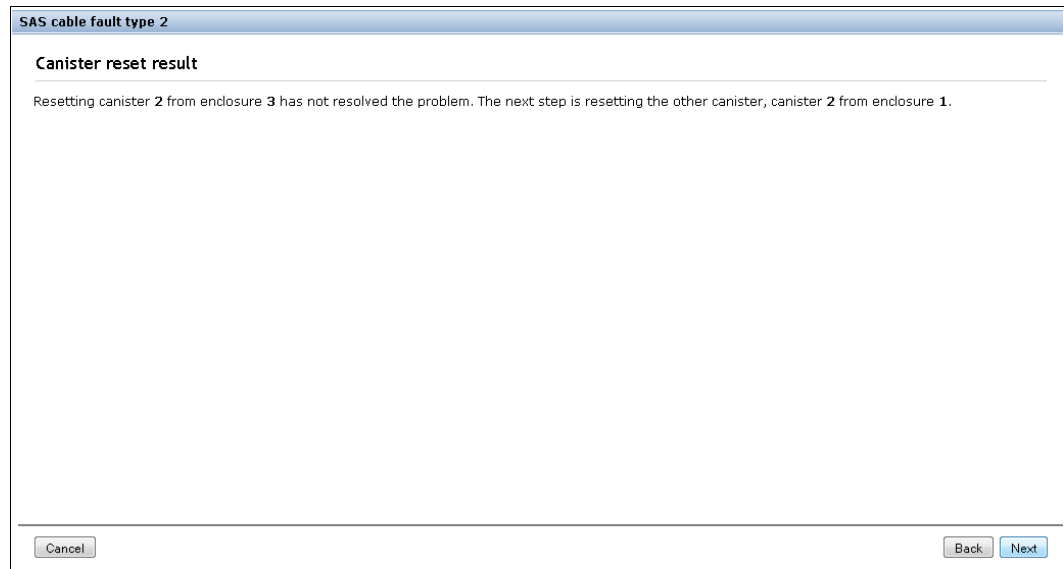


Figure 12-58 SAS cable Recommended action DMP step 3

Figure 12-59 shows step 4 of the DMP SAS cable event.

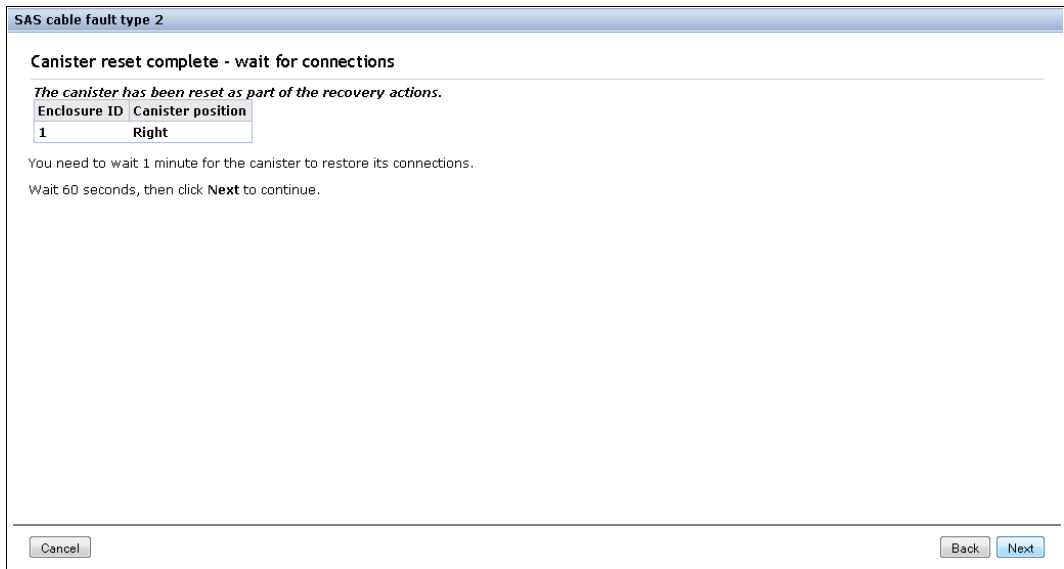


Figure 12-59 SAS cable Recommended action DMP step 4

Figure 12-60 shows step 5 of the DMP SAS cable event.

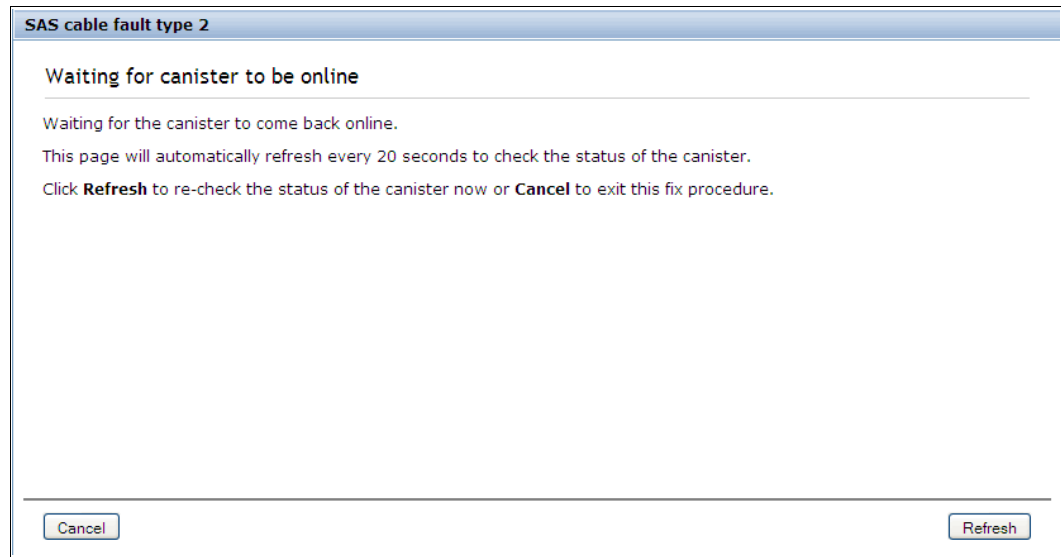


Figure 12-60 SAS cable Recommended action DMP step 5

Figure 12-61 shows step 6 of the DMP SAS cable event.

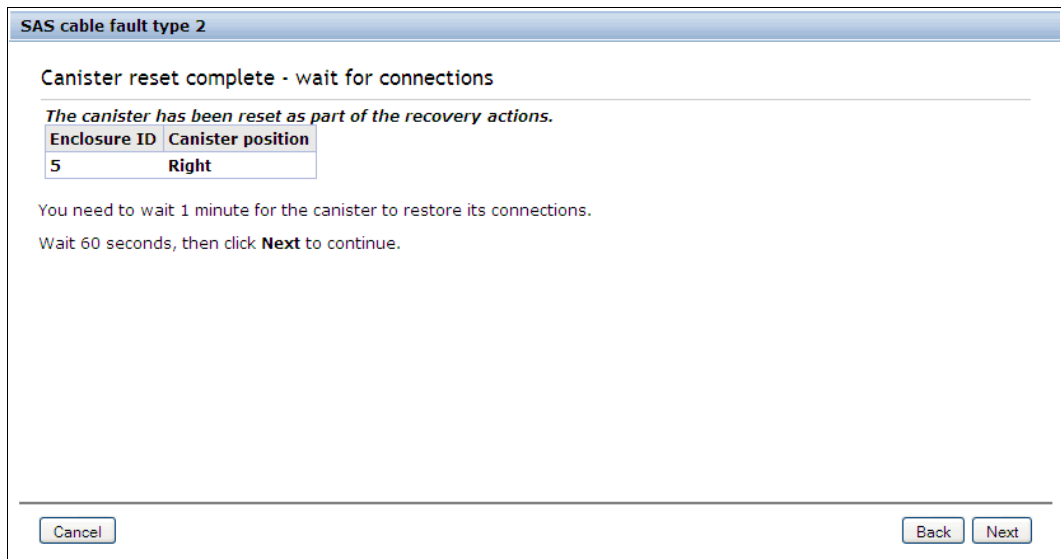


Figure 12-61 SAS cable Recommended action DMP step 6

Figure 12-62 shows step 7 of the DMP SAS cable event.

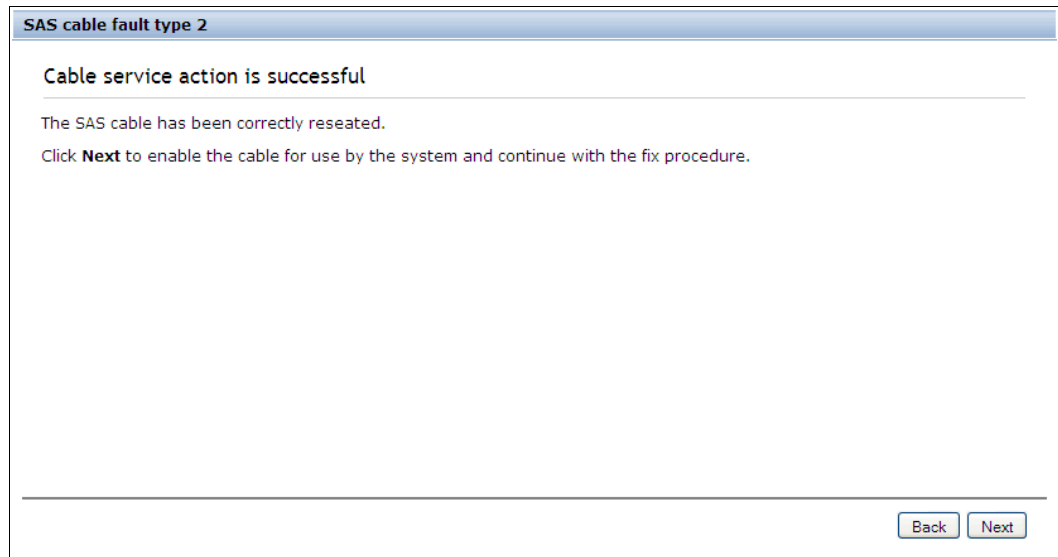


Figure 12-62 SAS cable Recommended action DMP step 7

Figure 12-63 shows step 8 of the DMP SAS cable event.

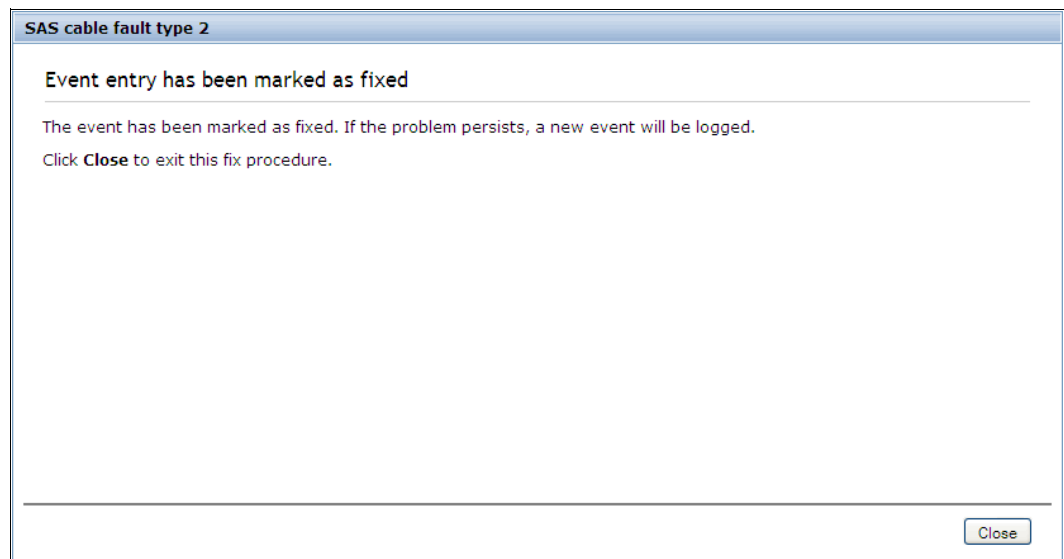


Figure 12-63 SAS cable Recommended action DMP step 8

When all of the steps of the DMP are processed successfully, the recommended action is complete and the problem should be fixed. Figure 12-64 on page 600 shows the red color of the event status changed to green. The system health status is green and there are no other that must be addressed.

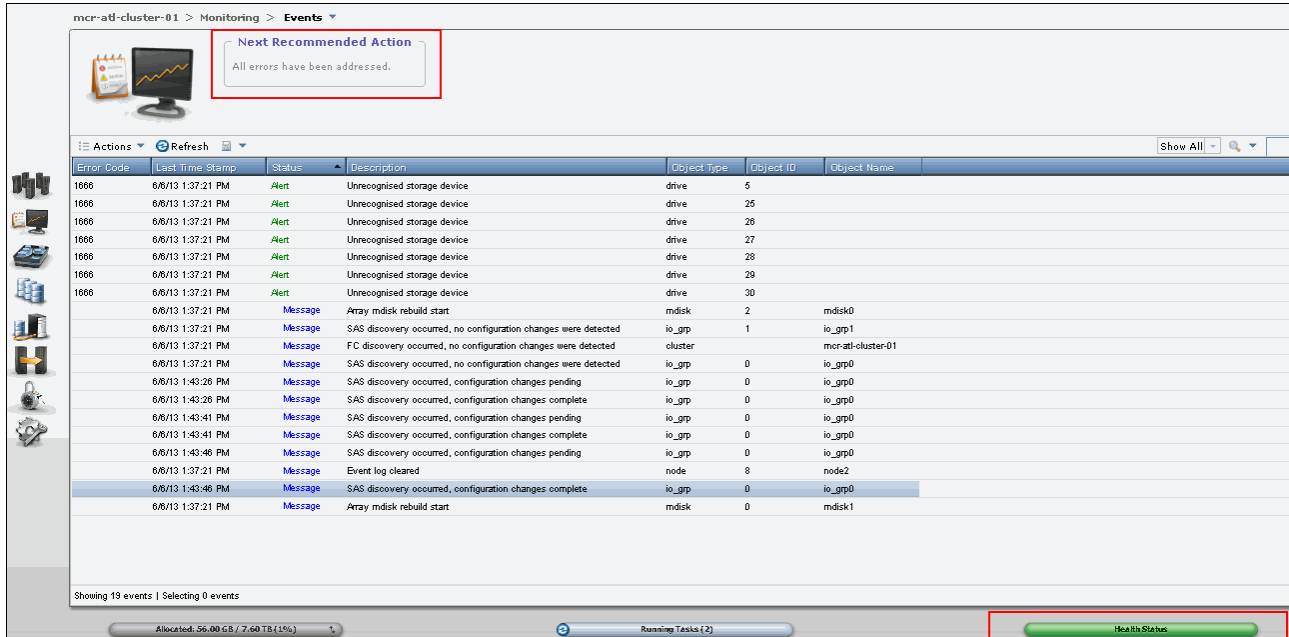


Figure 12-64 Recommended action that is completed

Handling multiple alerts

If there are multiple alerts that are logged, the IBM Storwize V5000 recommends a next action to fix the problem (or problems).

Figure 12-65 shows the event log that displays multiple alert.

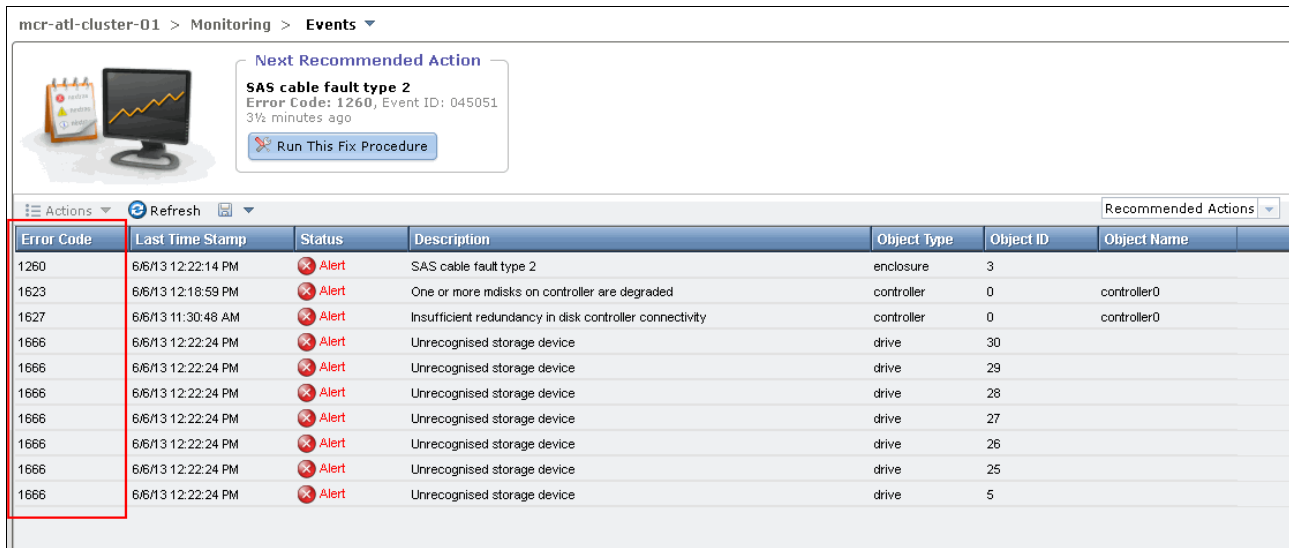


Figure 12-65 Multiple alert events that are displayed in the event log

The Next Recommended Action function orders the alerts by severity and displays the events with the highest severity first. If multiple events have the same severity, they are ordered by date and the oldest event is displayed first.

The following order of severity starts with the most severe condition:

- ▶ Unfixed alerts (sorted by error code; the lowest error code has the highest severity)
- ▶ Unfixed messages
- ▶ Monitoring events (sorted by error code; the lowest error code has the highest severity)
- ▶ Expired events
- ▶ Fixed alerts and messages

Faults are often fixed with the fixture of the most severe fault.

12.6 Collecting support information

If you have a problem and call the IBM Support Center, you might be asked to provide support data, as described in the next section.

12.6.1 Support information via GUI

Complete the following steps to collect support information by using the GUI:

1. Click **Settings** and then the **Support** tab (as shown in Figure 12-66) to begin the procedure of collecting support data.



Figure 12-66 Support Files VIA® GUI

2. Click **Download Support Package**, as shown in Figure 12-67.

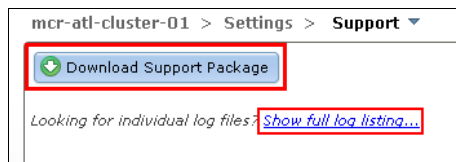


Figure 12-67 Download Support Package

The panel that is shown in Figure 12-68 opens and you can select one of four different versions of the svc_snap support package.

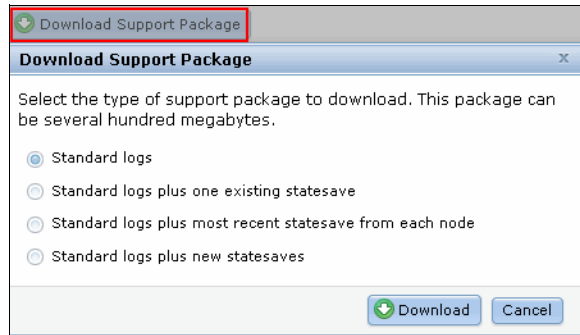


Figure 12-68 Support Package Selection

The version that you download that depends on the event that you are investigating. For example, if you noticed in the event log that a node was restarted, capture the snap with the latest existing statesaves.

The following components are included in the support package:

- ▶ Standard logs
Contains the most recent logs that were collected from the system. These logs are most commonly used by Support to diagnose and solve problems.
- ▶ Standard logs plus one existing statesave
Contains the standard logs from the system and the most recent statesave from any of the nodes in the system. Statesaves are also known as *memory dumps* or *live memory dumps*.
- ▶ Standard logs plus most recent statesave from each node
This option is used most often by the support team for problem analysis. They contain the standard logs from system and the most recent statesave from each node in the system.
- ▶ Standard logs plus new statesave
This option might be requested by the Support team for problem determination. It generates a new statesave (livedump) for all of the nodes and packages them with the most recent logs.

Save the resulting snap file in a directory for later use or to upload to IBM support.

12.6.2 Support information via Service Assistant

The IBM Storwize V5000 management GUI collects information from all the components in the system. The Service Assistant collects information from all node canisters. The snap file is the information that is collected and packaged in a single file.

If the package is collected by using the Service Assistant, ensure that the node from which the logs are collected is the current node, as shown in Figure 12-69.

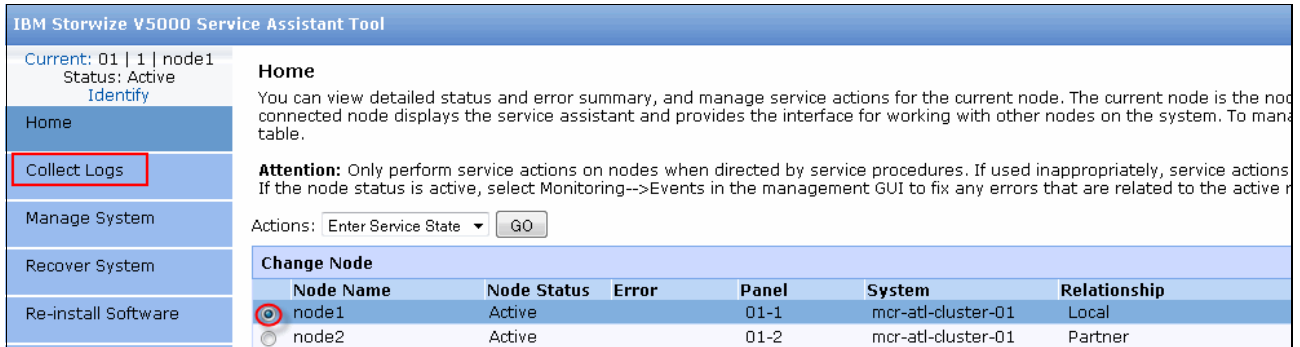


Figure 12-69 Collect logs with Service Assistance

Support information can be downloaded with or without the latest statesave, as shown in Figure 12-70.

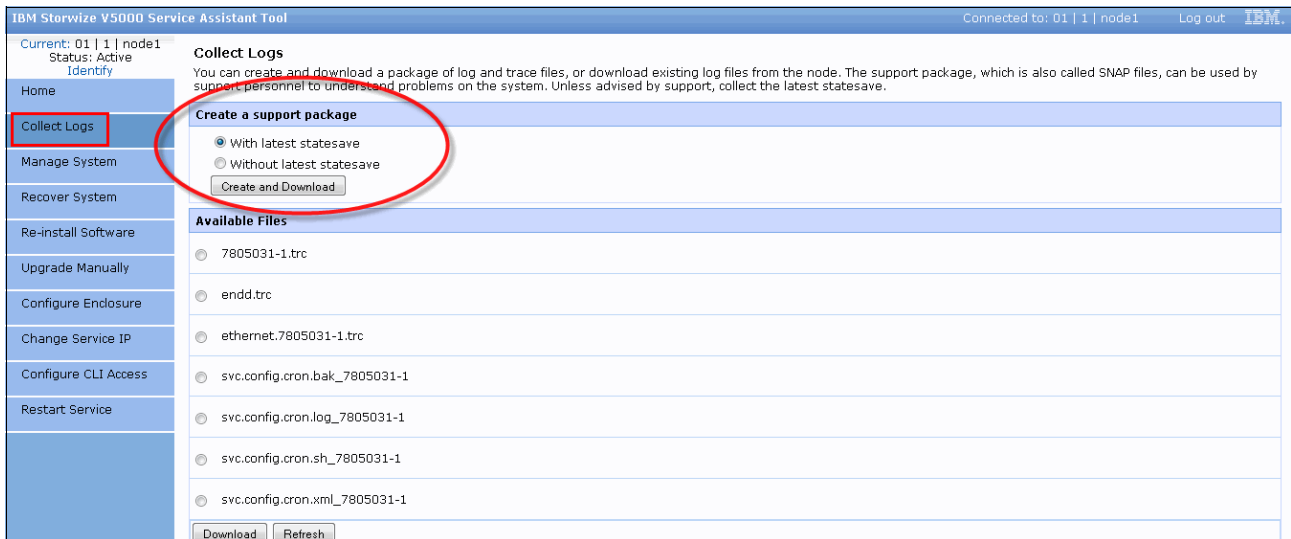


Figure 12-70 Download support file via Service Assistant

12.6.3 Support Information onto USB stick

Whenever GUI, Service Assistant, or a remote connection is unavailable, snaps can be collected from each single node by using the USB stick.

Complete the following steps to collect snaps by using the USB stick:

1. Create a text file that includes the following command:


```
satask snap -dump
```
2. Save the file as `satask.txt` in the root directory of the USB stick.
3. Insert the USB stick in the USB port of the node from which the support data should be collected.
4. Wait until no write activities are recognized (this process can take 10 minutes or more).

- Remove the USB stick and check the results, as shown in Figure 12-71.

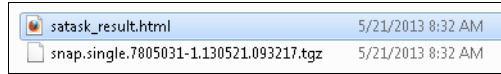


Figure 12-71 Single snap result files on USB stick

satask_result file

The `satask_result.html` file is the general response to the command that is issued via the USB stick. If the command did not run successfully, it is noticed in this file. Otherwise, any general system information is stored here, as shown in Figure 12-72.

```

satask snap -dump

Tue May 21 09:32:12 BST 2013

System Status

sainfo Isservicenodes

panel_name cluster_id cluster_name node_id node_name relation node_status error_data
01-1 000000C0200026C0 mcr-atl-cluster-01 1 node1 local Active
01-2 000000C0200026C0 mcr-atl-cluster-01 2 node2 partner Active
  
```

Figure 12-72 `satask_result.txt` on USB stick (header only)

Snap memory dump on USB

A complete statesave of the node where the USB was attached is stored in a `.zip` file. The name of the file includes the node name and the time stamp. The content of the `.zip` file is shown in Figure 12-73.

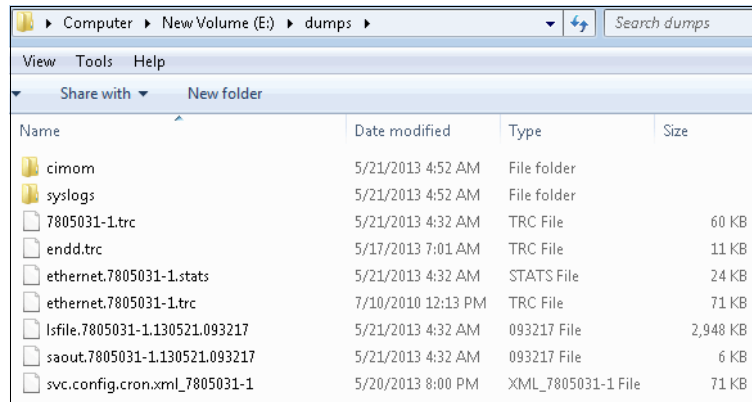


Figure 12-73 Single snap memory dump on USB stick

12.7 Powering on and shutting down IBM Storwize V5000

In the following sections, we describe the process to power on and shut down the IBM Storwize V5000 system by using the GUI and the CLI.

12.7.1 Shutting down the system

In this section, we show how to shut down the IBM Storwize V5000 system by using the GUI and CLI.

Important: You should never shut down your IBM Storwize V5000 by powering off the PSUs, removing both PSUs, or removing both power cables from a running system.

Powering down by using the GUI

You can shut down only one node canister or the entire cluster. When you shut down only one node canister, all of the activities remain active. When you shut down the entire cluster, you must power on locally to restart the system.

To shut down by using the GUI, complete the following steps:

1. Browse to the Monitoring function icon (as shown in Figure 12-74) and click **System Details**.

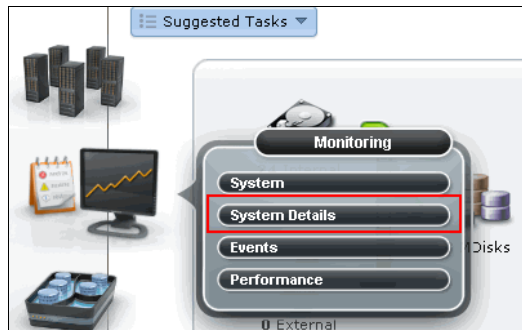


Figure 12-74 Power down via system details

2. Select the root level of the system detail tree, click **Actions**, and then select **Shut Down System**, as shown in Figure 12-75.

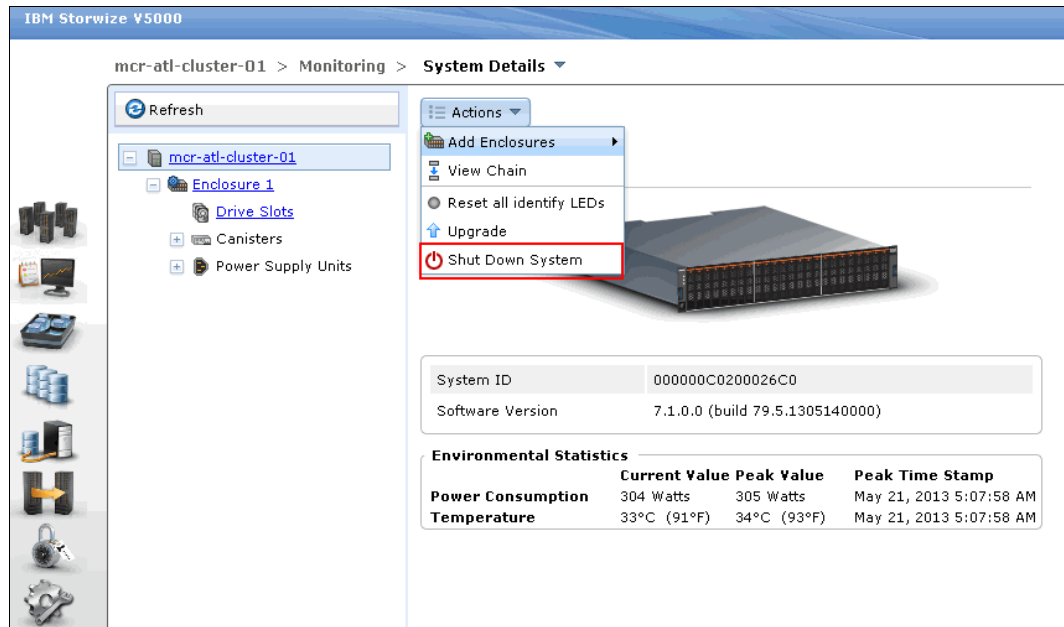


Figure 12-75 Power Down System option

The following process can be used as an alternative to steps 1 and 2, as shown in Figure 12-75:

- a. Browse to the Monitoring navigator and open the System view.
- b. Click the System that is under the system display.
An information panel opens.
- c. Click the **Manage** tab.
- d. Click **Shut Down System** to shut down, as shown in Figure 12-76.

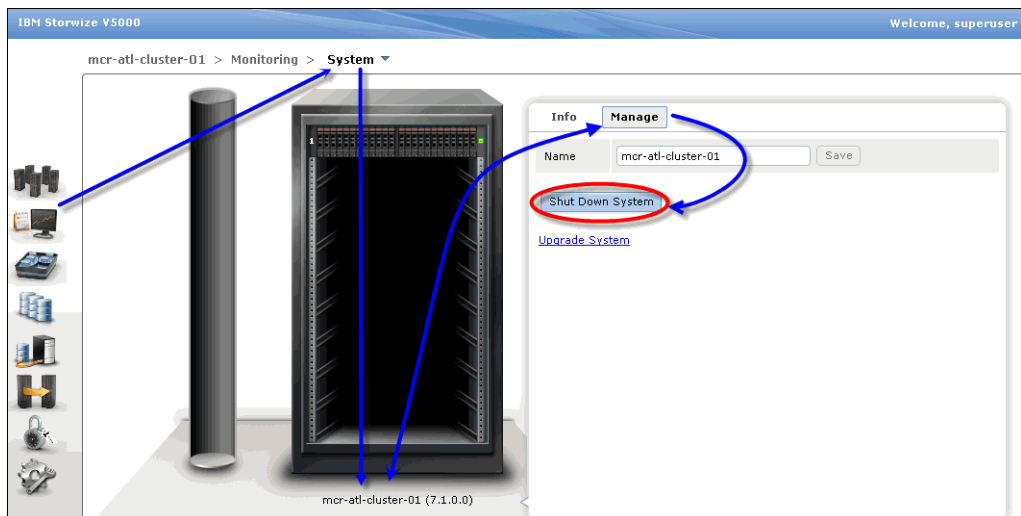


Figure 12-76 Shut down system via Monitoring system GUI

3. The Confirm System Shutdown window opens. A message opens and prompts you to confirm whether you want to shut down the cluster. Ensure that you stopped all FlashCopy mappings, data migration operations, and forced deletions before you continue. Enter Yes and click **OK** to begin the shutdown process, as shown in Figure 12-77.

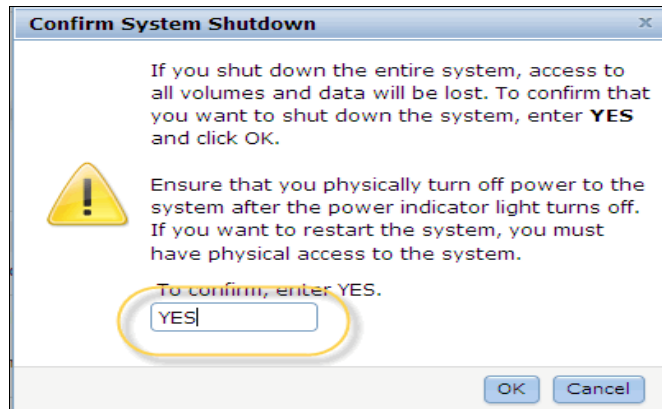


Figure 12-77 Shut Down confirmation

4. Wait for the power LED on both node canisters in the control enclosure to flash at 1 Hz, which indicates that the shutdown operation completed (1 Hz is half as fast as the drive indicator LED).

Tip: When you shut down an IBM Storwize V5000, it does not automatically restart. You must manually restart the system.

Shutting down by using the CLI

The CLI is the other option that can be used to shut down an IBM Storwize V5000. The CLI is accessed via the PuTTY utility.

Warning: If you are shutting down the entire system, you lose access to all volumes that are provided by this system. Shutting down the system also shuts down all IBM Storwize V5000 nodes. This shutdown causes the hardened data to be dumped to the internal HDD.

Run the **stopsystem** command to shut down a clustered system, as shown in Example 12-3.

Example 12-3 Shut down

```
stopsystem
```

```
Are you sure that you want to continue with the shut down?
```

```
# Type y to shut down the entire clustered system.
```

12.7.2 Powering on

Complete the following steps to power on the system:

Important: This process assumes that all power is removed from the enclosure. If the control enclosure is shut down but the power is not removed, the power LED on all node canisters flash at a rate of half of one second on, half of one second off. In this case, remove the power cords from both power supplies and then reinsert them.

1. Ensure that any network switches that are connected to the system are powered on.
2. Power on any expansion enclosures by connecting the power cord to both power supplies in the rear of the enclosure or turning on the power circuit.
3. Power on the control enclosure by connecting the power cords to both power supplies in the rear of the enclosure and turning on the power circuits.

The system starts. The system starts successfully when all node canisters in the control enclosure have their status LED permanently on, which should take no longer than 10 minutes.

4. Start the host applications.



Command-line interface setup and SAN Boot

This appendix describes the setup of the command-line interface (CLI) and provides more information about the SAN Boot function.

This appendix includes the following sections:

- ▶ Command-line interface
- ▶ SAN Boot

Command-line interface

The IBM Storwize V5000 system has a powerful CLI, which offers even more functions than the GUI. This section is not intended to be a detailed guide to the CLI because that topic is beyond the scope of this book. The basic configuration of the IBM Storwize V5000 CLI and some example commands are covered. However, the CLI commands are the same as in the SAN Volume Controller. In addition, there are more commands that are available to manage internal storage. If a task is completed in the GUI, the CLI command always is displayed in the details, as shown throughout this book.

Detailed CLI information is available in the IBM Storwize V5000 Information Center under the Command Line section, which can be found at this website:

http://pic.dhe.ibm.com/infocenter/storwize/V5000_ic/index.jsp?topic=%2Fcom.ibm.storwize.V5000.641.doc%2Fsvc_clicommandscontainer_229g0r.html

Implementing the IBM Storwize V7000 V6.3, SG24-7938, also has information about the use of the CLI. The commands in that book also apply to the IBM Storwize V5000 system because it is part of the Storwize family.

Basic setup

In the IBM Storwize V5000 GUI, authentication is done by using a user name and a password. The CLI uses a Secure Shell (SSH) to connect from the host to the IBM Storwize V5000 system. A private and public key pair or user name and password is necessary. The following steps are required to enable CLI access with SSH keys:

1. A public key and private key are generated as a pair.
2. A public key is uploaded to the IBM Storwize V5000 system by using the GUI.
3. A client SSH tool is configured to authenticate with the private key.
4. A secure connection is established between the client and IBM Storwize V5000 system.

Secure Shell is the communication vehicle that is used between the management workstation and the IBM Storwize V5000 system. The SSH client provides a secure environment from which to connect to a remote machine. It uses the principles of public and private keys for authentication.

SSH keys are generated by the SSH client software. The SSH keys include a public key, which is uploaded and maintained by the clustered system, and a private key, which is kept private on the workstation that is running the SSH client. These keys authorize specific users to access the administration and service functions on the system. Each key pair is associated with a user-defined ID string that can consist of up to 40 characters. Up to 100 keys can be stored on the system. New IDs and keys can be added, and unwanted IDs and keys can be deleted. To use the CLI, an SSH client must be installed on that system, the SSH key pair must be generated on the client system, and the client's SSH public key must be stored on the IBM Storwize V5000 systems.

The SSH client that is used in this book is PuTTY. There also is a PuTTY key generator that can be used to generate the private and public key pair. The PuTTY client can be downloaded at no cost at the following website:

<http://www.chiark.greenend.org.uk>

The following tools should be downloaded:

- ▶ PuTTY SSH client: `putty.exe`
- ▶ PuTTY key generator: `puttygen.exe`

Generating a public and private key pair

To generate a public and private key pair, complete the following steps:

1. Start the PuTTY key generator to generate the public and private key pair, as shown in Figure A-1.

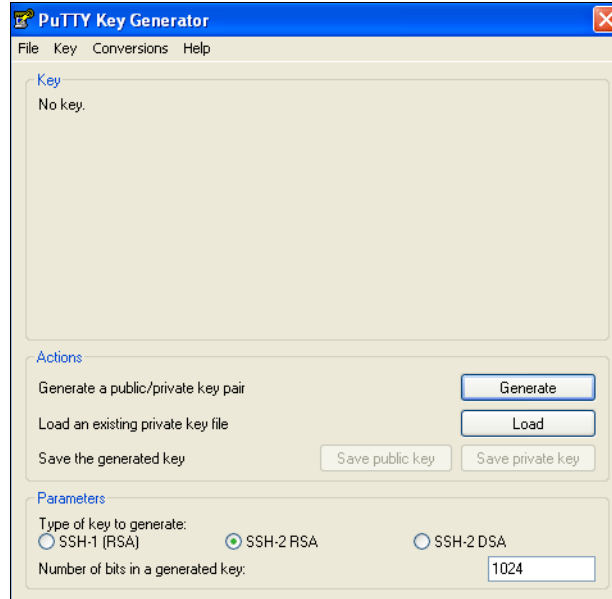


Figure A-1 PuTTY key generator

Make sure that the following options are selected:

- Type of key to generate: SSH2 RSA
- Number of bits in a generated key: 1024

2. Click **Generate** and move the cursor over the blank area to generate the keys, as shown in Figure A-2.

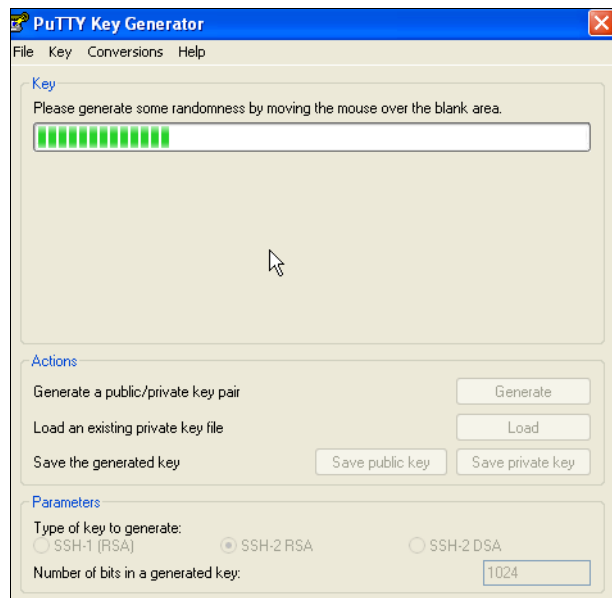


Figure A-2 Generate keys

Generating keys: The blank area that is indicated by the message is the large blank rectangle on the GUI inside the section of the GUI labeled Key. Continue to move the mouse pointer over the blank area until the progress bar reaches the far right side. This action generates random characters to create a unique key pair.

3. After the keys are generated, save them for later use. Click **Save public key**, as shown in Figure A-3.

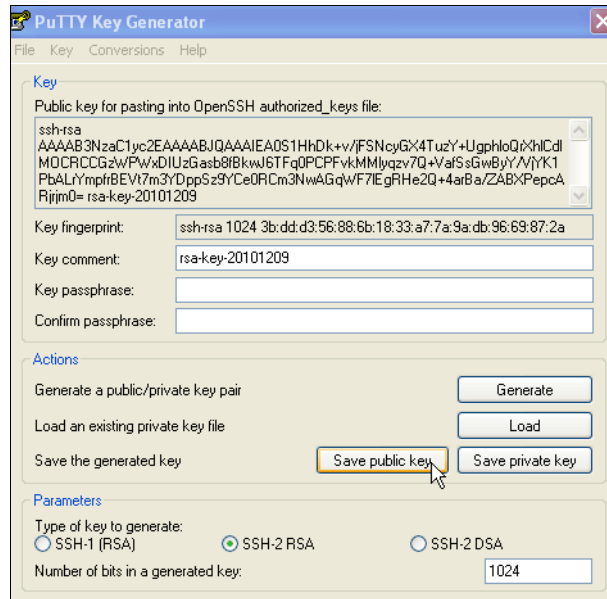


Figure A-3 Save public key

4. You are prompted for a name (for example, pubkey) and a location for the public key (for example, C:\Support Utils\PuTTY). Click **Save**.

Be sure to record the name and location of this SSH public key because this information must be specified later.

Public key extension: By default, the PuTTY key generator saves the public key with no extension. Use the string pub for naming the public key; for example, pubkey, to differentiate the SSH public key from the SSH private key.

5. Click **Save private key**, as shown in Figure A-4.



Figure A-4 Save private key

6. You receive a warning message, as shown in Figure A-5. Click **Yes** to save the private key without a passphrase.

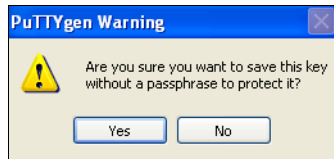


Figure A-5 Confirm the security warning

7. When prompted, enter a name (for example, i cat), select a secure place as the location, and click **Save**.

Key generator: The PuTTY key generator saves the private key with the PPK extension.

8. Close the PuTTY key generator.

Uploading the SSH public key to the IBM Storwize V5000

After you create your SSH key pair, you must upload your SSH public key onto the SAN Volume Controller system. Complete the following steps to upload the key:

1. Open the user section, as shown in Figure A-6.

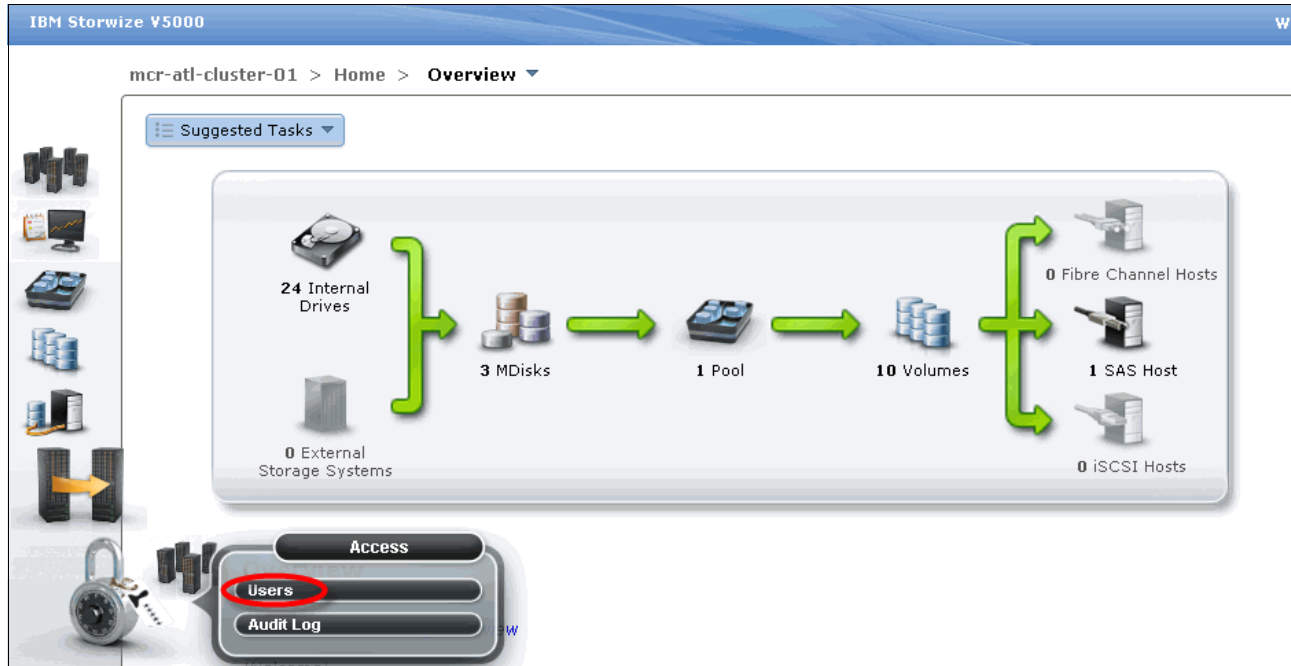


Figure A-6 Open user section

2. Right-click the user for which you want to upload the key and click **Properties**, as shown in Figure A-7.

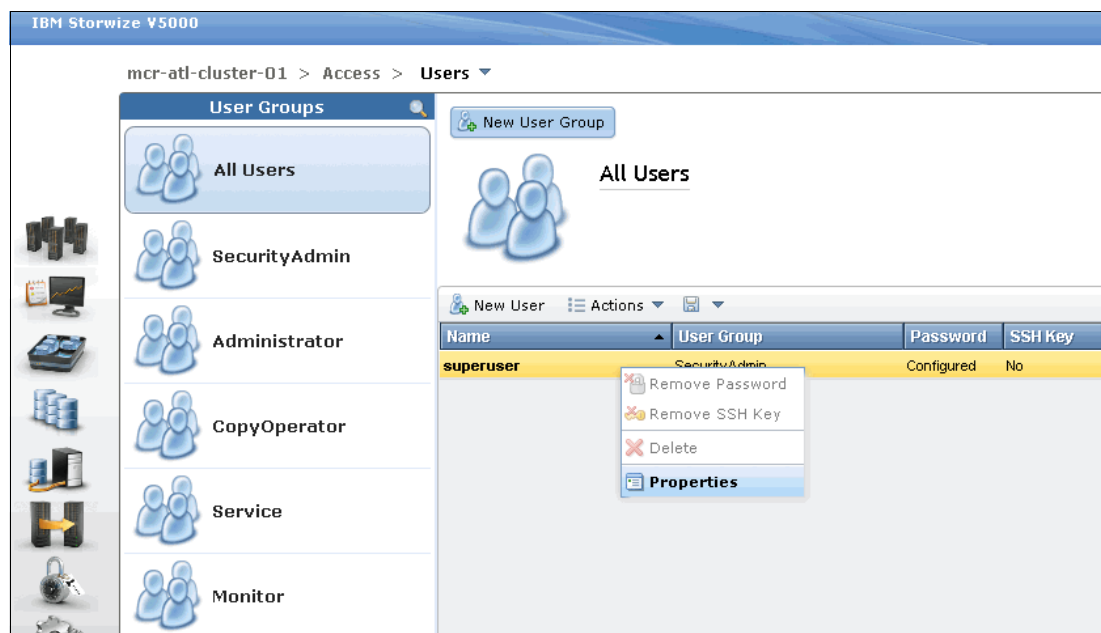


Figure A-7 Superuser properties

- To upload the public key, click **Browse**, select your public key, and click **OK**, as shown in Figure A-8.

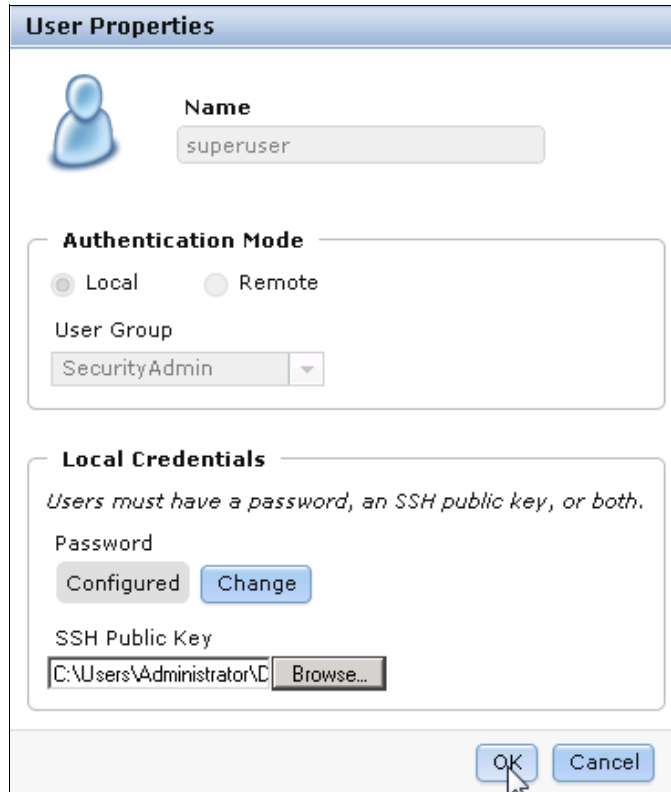


Figure A-8 Select public key

- Click **OK** and the key is uploaded, as shown in Figure A-9.

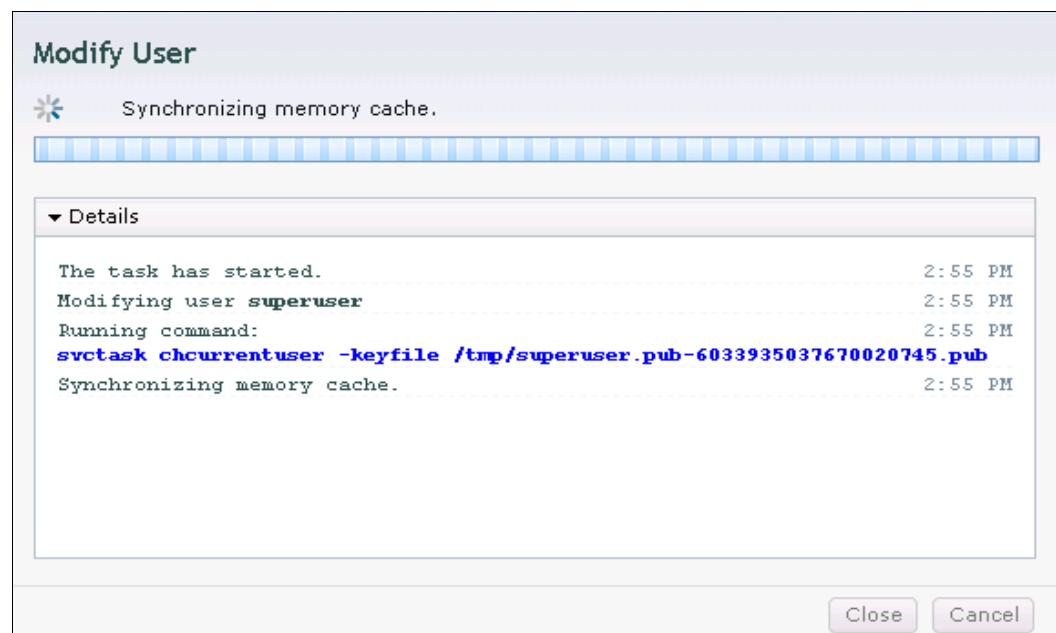


Figure A-9 Public key upload complete

5. Click **Close** to return to the GUI.

Configuring the SSH client

Before the CLI can be used, the SSH client must be configured. Complete the following steps to configure the client:

1. Start PuTTY, as shown in Figure A-10.

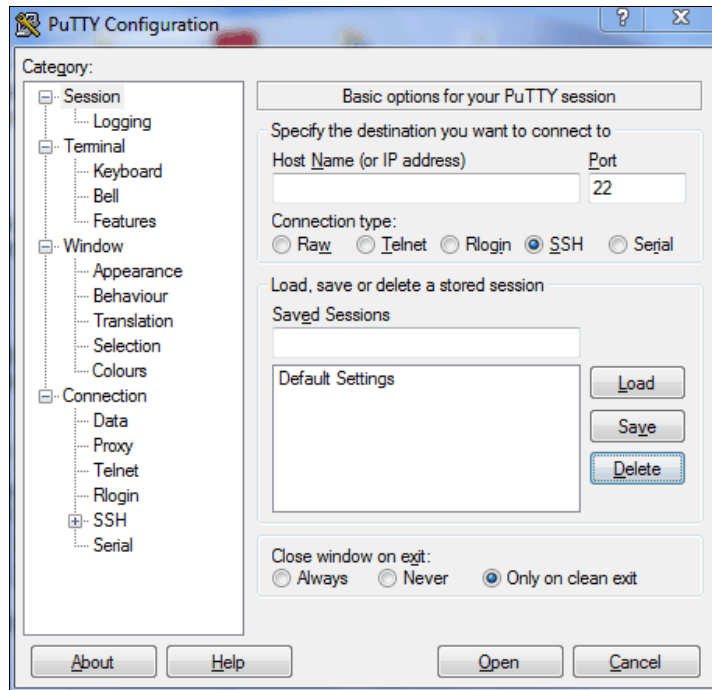


Figure A-10 PuTTY

In the right side pane under the “Specify the destination you want to connect to” section, select **SSH**. Under the “Close window on exit” section, select **Only on clean exit**, which ensures that if there are any connection errors, they are displayed in the user’s window.

- From the Category pane on the left side of the PuTTY Configuration window, click **Connection** → **SSH** to open the PuTTY SSH Configuration window, as shown in Figure A-11.

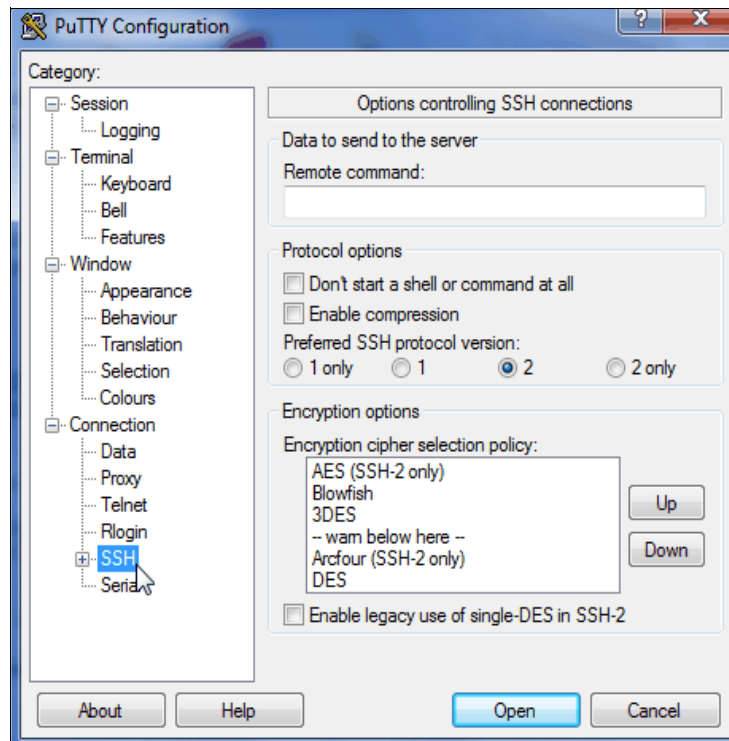


Figure A-11 SSH protocol version 2

- In the right side pane in the “Preferred SSH protocol version” section, select 2.

- From the Category pane on the left side of the PuTTY Configuration window, click **Connection** → **SSH** → **Auth**. As shown in Figure A-12, in the right side pane in the “Private key file for authentication:” field under the Authentication Parameters section, browse to or manually enter the fully qualified directory path and file name of the SSH client private key file that was created earlier (for example, C:\Support Utils\PuTTY\icat.PPK).

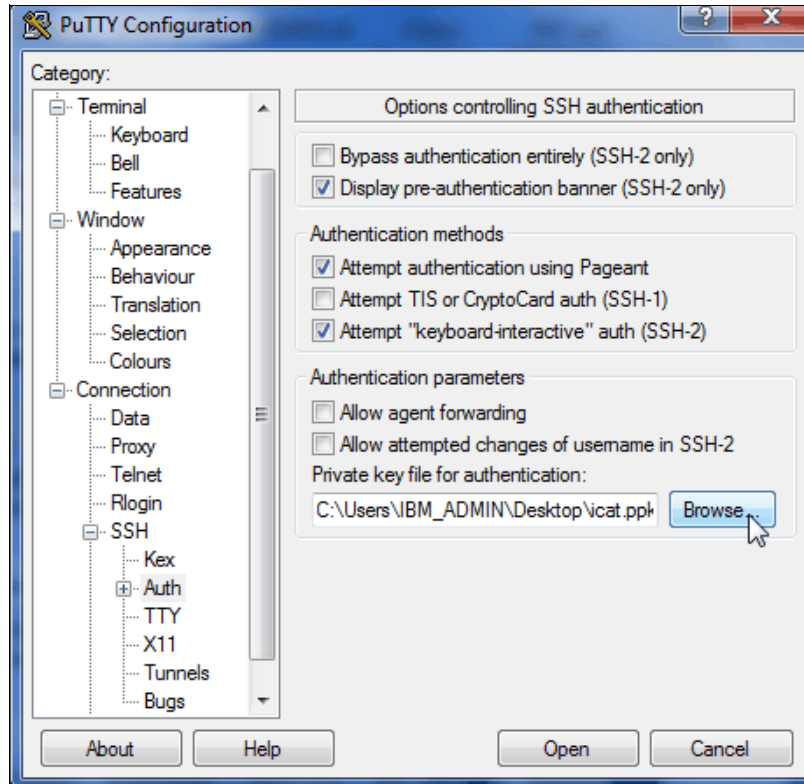


Figure A-12 SSH authentication

- From the Category pane on the left side of the PuTTY Configuration window, click **Session** to return to the Session view, as shown in Figure A-10 on page 616.

6. In the right side pane, enter the host name or system IP address of the IBM Storwize V5000 clustered system in the Host Name field. Enter a session name in the Saved Sessions field, as shown in Figure A-13.

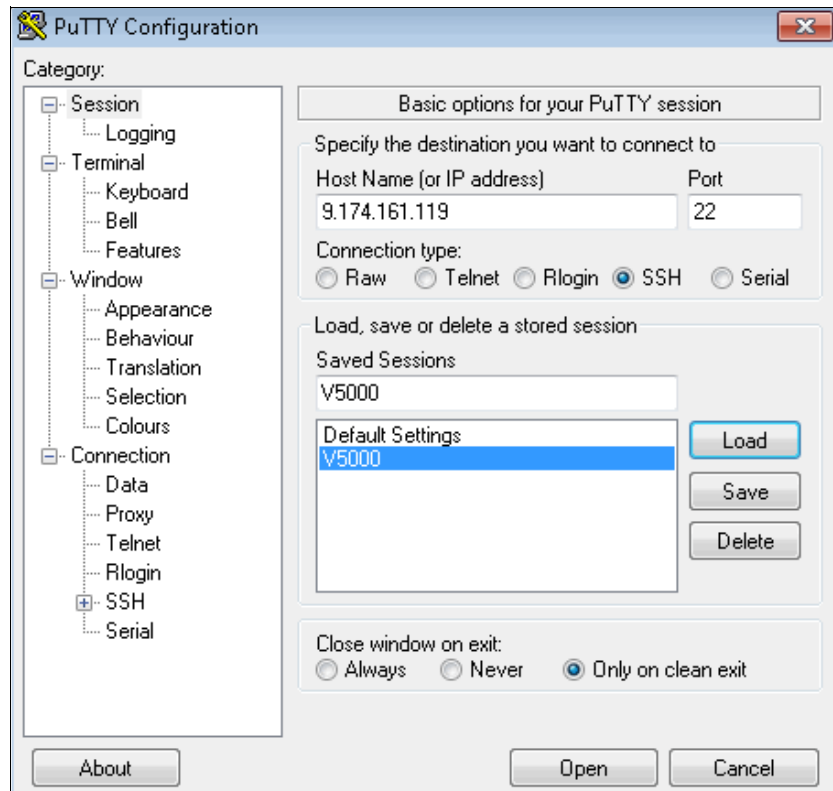


Figure A-13 Enter session information

7. Click **Save** to save the new session, as shown in Figure A-14.

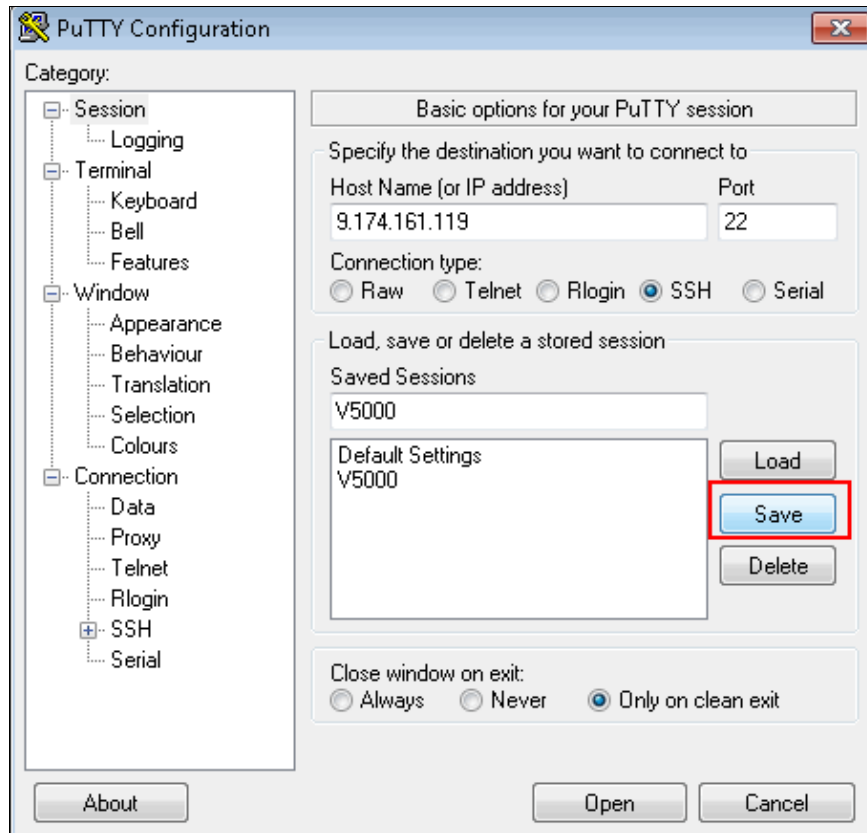


Figure A-14 Save Session

8. Highlight the new session and click **Open** to connect to the IBM Storwize V5000 system.
9. PuTTY now connects to the system and prompts you for a user name. Enter `admin` as the user name and press Enter (see Example A-1).

Example: A-1 Enter user name

```
login as: superuser
Authenticating with public key "rsa-key-20130521"
Last login: Tue May 21 15:21:55 2013 from 9.174.219.143
IBM_Storwize:mcr-atl-cluster-01:superuser>
```

The CLI is now configured for IBM Storwize V5000 administration.

Example commands

A detailed description about all of the available commands is beyond the intended scope of this book. In this section, sample commands that we referenced in this book are presented.

The `svcin` and `svctask` prefixes are no longer needed in IBM Storwize V5000. If you have scripts that use this prefix, they run without problems. If you enter `svcin` or `svctask` and press the Tab key twice, all of the available subcommands are listed. Pressing the Tab key twice also auto-completes commands if the input is valid and unique to the system.

Enter **lsvdisk** (as shown in Example A-2) to list all configured volumes on the system. The example shows that six volumes are configured.

Example: A-2 List all volumes

```

IBM_Storwize:mcr-atl-cluster-01:superuser>lsvdisk
id name          IO_group_id IO_group_name status mdisk_grp_id mdisk_grp_name
capacity type    FC_id FC_name RC_id RC_name vdisk_UID  opy_count
fast_write_state se_copy_count RC_change compressed_copy_count
0 V5000_Vol1 0          io_grp0      online 0          V5000_Pool
20.00GB striped                6005076300800 empty
1          no          0
1 V5000_Vol2 0          io_grp0      online 0          V5000_Pool
2.00GB striped                6005076300800 empty
1          no          0
2 V5000_Vol3 0          io_grp0      online 0          V5000_Pool
2.00GB striped                6005076300800 empty
1          no          0
3 V5000_Vol4 0          io_grp0      online 0          V5000_Pool
2.00GB striped                6005076300800 empty
1          no          0
4 V5000_Vol5 0          io_grp0      online 0          V5000_Pool
2.00GB striped                6005076300800 empty
1          no          0
5 V5000_Vol6 0          io_grp0      online 0          V5000_Pool
2.00GB striped                6005076300800 empty
1          no          0

```

Enter **lshost** to see a list of all configured hosts on the system, as shown in Example A-3.

Example: A-3 List hosts

```

IBM_Storwize:mcr-atl-cluster-01:superuser>lshost
id name          port_count iogrp_count status
0 windows2008r2 2          4          online

```

To map the volume to the hosts, enter **mkvdiskhostmap**, as shown in Example A-4.

Example: A-4 Map volumes to host

```

IBM_Storwize:mcr-atl-cluster-01:superuser>mkvdiskhostmap -host ESXi-1 -scsi 0 -force
ESXi-Redbooks
Virtual Disk to Host map, id [0], successfully created

```

To verify the host mapping, enter **lsvdiskhostmap**, as shown in Example A-5.

Example: A-5 List all hosts that are mapped to a volume

```

IBM_Storwize:mcr-atl-cluster-01:superuser>lsvdiskhostmap ESXi-1
id name          SCSI_id vdisk_id vdisk_name  vdisk_UID
4 ESXi-1 0          2          ESXi-Redbooks 600507680185853FF000000000000011

```

In the CLI, there are more options available than in the GUI. All advanced settings can be set; for example, I/O throttling. To enable I/O throttling, change the properties of a volume by using the **changevdisk** command, as shown in Example A-6. To verify the changes, run the **lsvdisk** command.

Example: A-6 Enable advanced properties: I/O throttling

```
IBM_Storwize:mcr-atl-cluster-01:superuser>chvdisk -rate 1200 -unit mb
ESXi-Redbooks
IBM_Storwize:mcr-atl-cluster-01:superuser>
IBM_Storwize:mcr-atl-cluster-01:superuser>lsvdisk ESXi-Redbooks

id 2
name ESXi-Redbooks
.
.
vdisk_UID 600507680185853FF000000000000011
virtual_disk_throttling (MB) 1200
preferred_node_id 2
.
.
IBM_Storwize:mcr-atl-cluster-01:superuser>
```

Command output: The **lsvdisk** command lists all available properties of a volume and its copies. To make it easier to read, lines in Example A-6 were deleted.

If you do not specify the unit parameter, the throttling is based on I/Os instead of throughput, as shown in Example A-7.

Example: A-7 Throttling based on I/O

```
IBM_Storwize:mcr-atl-cluster-01:superuser>chvdisk -rate 4000 ESXi-Redbooks
IBM_Storwize:mcr-atl-cluster-01:superuser>lsvdisk ESXi-Redbooks
id 2
name ESXi-Redbooks
.
.
vdisk_UID 600507680185853FF000000000000011
throttling 4000
preferred_node_id 2
.
.
IBM_Storwize:mcr-atl-cluster-01:superuser>
```

To disable I/O throttling, set the I/O rate to 0, as shown in Example A-8.

Example: A-8 Disable I/O Throttling

```
IBM_Storwize:mcr-atl-cluster-01:superuser>chvdisk -rate 0 ESXi-Redbooks
IBM_Storwize:mcr-atl-cluster-01:superuser>lsvdisk ESXi-Redbooks
id 2
.
.
vdisk_UID 600507680185853FF000000000000011
throttling 0
preferred_node_id 2
IBM_Storwize:mcr-atl-cluster-01:superuser>
```

SAN Boot

IBM Storwize V5000 supports SAN Boot for Windows, VMware, and many other operating systems. SAN Boot support can change, so regularly check the IBM Storwize V5000 interoperability matrix at this website:

<http://www-01.ibm.com/support/docview.wss?uid=ssg1S1004111>

The IBM Storwize V5000 Information Center has more information about SAN Boot for different operating systems, which is available at this website:

http://pic.dhe.ibm.com/infocenter/storwize/V5000_ic/index.jsp?topic=%2Fcom.ibm.storwize.V5000.641.doc%2Fsvc_hostattachmentmain.html

For more information about SAN Boot, see the *IBM System Storage Multipath Subsystem Device Driver User's Guide*, GC52- 1309-03, which is available at this website:

ftp://ftp.software.ibm.com/storage/subsystem/UG/1.8--3.0/SDD_1.8--3.0_User_Guide_English_version.pdf

Enabling SAN Boot for Windows

Complete the following steps to install the Windows host by using SAN Boot:

1. Configure the IBM Storwize V5000 system so that only the boot volume is mapped to the host.
2. Configure the Fibre Channel SAN so that the host sees only one IBM Storwize V5000 system node port. Multiple paths during installation are not supported.
3. Configure and enable the host bus adapter (HBA) BIOS.
4. Install the operating system by using the normal procedure and select the volume as the partition on which to install.

HBAs: You might need to load another HBA device driver during installation, depending on your Windows version and the HBA type.

5. Install SDDDSM after the installation completes.
6. Modify your SAN zoning to allow multiple paths.

7. Check your host to see whether all paths are available.
8. Set redundant boot devices in the HBA BIOS to allow the host to boot when its original path fails.

Enabling SAN Boot for VMware

Complete the following steps to install a VMware ESXhost by using SAN Boot:

1. Configure the IBM Storwize V5000 system so that only the boot volume is mapped to the host.
2. Configure the Fibre Channel SAN so that the host sees only one IBM Storwize V5000 system node port. Multiple paths during installation are not supported.
3. Configure and enable the HBA BIOS.
4. Install the operating system by using the normal procedure and select the volume as the partition on which to install.

HBAs: You might need to load another I HBA device driver during installation, depending on your ESX level and the HBA type.

5. Modify your SAN zoning to allow multiple paths.
6. Check your host if all paths are available and modify the multipath policy, if required.

Windows SAN Boot migration

If you have a host that runs a Windows 2000 Server, Windows Server 2003, or Windows Server 2008 operating system and have existing SAN Boot images that are controlled by storage controllers, you can migrate these images to image-mode volumes that are controlled by the IBM Storwize V5000 system.

SAN Boot procedures: For more information about SAN Boot procedures for other operating systems, see the IBM Storwize V5000 Information Center at this website:

http://pic.dhe.ibm.com/infocenter/storwize/V5000_ic/index.jsp?topic=%2Fcom.ibm.storwize.V5000.641.doc%2FV5000_ichome_641.html

Complete the following steps to migrate your existing SAN Boot images:

1. If the existing SAN Boot images are controlled by an IBM storage controller that uses the IBM Subsystem Device Driver (SDD) as the multipathing driver, you must use SDD V1.6 or higher. Run the SDD **datapath set bootdiskmigrate 2076** command to prepare the host for image migration. See the Multipath SDD matrix to download packages at this website:
http://www-01.ibm.com/support/docview.wss?rs=540&context=ST52G7&dc=DA400&uid=ssg1S7001350&loc=en_US&cs=utf-8&lang=en#WindowsSDD
2. Shut down the host.
3. Complete the following configuration changes on the storage controller:
 - a. Write down the SCSI LUN ID each volume is using (for example, boot LUN SCSI ID 0, Swap LUN SCSI ID 1, and Database LUN SCSI ID 2).
 - b. Remove all of the image-to-host mappings from the storage controller.

- c. Map the existing SAN Boot image and any other disks to the IBM Storwize V5000 system.
4. Change the zoning so that the host can see the IBM Storwize V5000 I/O group for the target image mode volume.
5. Complete the following configuration changes on the IBM Storwize V5000 system:
 - a. Create an image mode volume for the managed disk (MDisk) that contains the SAN Boot image. Use the MDisk unique identifier to specify the correct MDisk.
 - b. Create a host object and assign the host HBA ports.
 - c. Map the image mode volume to the host by using the same SCSI ID as before. For example, you might map the boot disk to the host with SCSI LUN ID 0.
 - d. Map the swap disk to the host, if required. For example, you might map the swap disk to the host with SCSI LUN ID 1.
6. Change the boot address of the host by completing the following steps:
 - a. Restart the host and open the HBA BIOS utility of the host during the booting process.
 - b. Set the BIOS settings on the host to find the boot image at the worldwide port name (WWPN) of the node that is zoned to the HBA port.
7. If SDD V1.6 or higher is installed and you ran the **bootdiskmigrate** command in step 1 on page 624, reboot your host, update SDDDSM to the latest level, and go to step 14. If SDD V1.6 is not installed, go to step 8.
8. Modify the SAN Zoning so that the host sees only one path to the IBM Storwize V5000.
9. Boot the host in single-path mode.
10. Uninstall any multipathing driver that is not supported for IBM Storwize V5000 system hosts that run the applicable Windows Server operating system.
11. Install SDDDSM.
12. Restart the host in single-path mode and ensure that SDDDSM was properly installed.
13. Modify the SAN Zoning to enable multipathing.
14. Rescan drives on your host and check that all paths are available.
15. Reboot your host and enter the HBA BIOS.
16. Configure the HBA settings on the host. Ensure that all HBA ports are boot-enabled and can see both nodes in the I/O group that contains the SAN Boot image. Configure the HBA ports for redundant paths.
17. Exit the BIOS utility and finish starting the host.
18. Map any other volumes to the host, as required.

Related publications and information

The publications and information that is listed in this section are considered particularly suitable for a more detailed discussion of the topics that are covered in this book.

IBM Redbooks

The following IBM Redbooks publications provide more information about the topics in this book. Some publications that are referenced in the following list might be available in softcopy only:

- ▶ *Implementing the IBM System Storage SAN Volume Controller V6.3*, SG24-7933
- ▶ *Implementing the IBM Storwize V7000 V6.3*, SG24-7938
- ▶ *SAN Volume Controller: Best Practices and Performance Guidelines*, SG24-7521
- ▶ *Implementing an IBM/Brocade SAN with 8 Gbps Directors and Switches*, SG24-6116

You can search for, view, download, or order these documents and other Redbooks, Redpapers, Web Docs, draft, and other materials at this website:

<http://www.ibm.com/redbooks>

IBM Storwize V5000 publications

Storwize V5000 publications are available at this website:

<https://ibm.biz/BdxyDL>

IBM Storwize V5000 support

Storwize V5000 support is available at this website:

<https://ibm.biz/BdxyD9>

Help from IBM

IBM Support and downloads

<https://www.ibm.com/support>

IBM Global Services

<https://www.ibm.com/services>



Redbooks

Implementing the IBM Storwize V5000

(1.0" spine)

0.875" x 1.498"

460 <-> 788 pages



Implementing the IBM Storwize V5000



Redbooks®

Easily manage and deploy systems with embedded GUI

Experience rapid and flexible provisioning

Protect data with remote mirroring

Organizations of all sizes are faced with the challenge of managing massive volumes of increasingly valuable data. But storing this data can be costly, and extracting value from the data is becoming more difficult. IT organizations have limited resources but must stay responsive to dynamic environments and act quickly to consolidate, simplify, and optimize their IT infrastructures. The IBM Storwize V5000 system provides a smarter solution that is affordable, easy to use, and self-optimizing, which enables organizations to overcome these storage challenges.

Storwize V5000 delivers efficient, entry-level configurations that are specifically designed to meet the needs of small and midsize businesses. Designed to provide organizations with the ability to consolidate and share data at an affordable price, Storwize V5000 offers advanced software capabilities that are usually found in more expensive systems.

This IBM Redbooks publication is intended for pre-sales and post-sales technical support professionals and storage administrators.

The concepts in this book also relate to the IBM Storwize V3700.

This book was written at a software level of Version 7 Release 1.

INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION

BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

For more information:
ibm.com/redbooks

SG24-8162-00

ISBN 0738438766