Redbooks
ibm.com/redbooks

IBM Storage Virtualize 3-Site Replication

For IBM Storage Virtualize 3-Site Orchestrator 4.1

Dr. Pradip Ashok Waykos
Vasfi Gucer
Prateek Sudin Mandge
Ajinkya Nanavati

Storage

IBM

Redbooks

IBM Redbooks

IBM Storage Virtualize 3-Site Replication

January 2024

SG24-8504-01

Note: Before using this information and the product it supports, read the information in “Notices” on page vii.

Second Edition (January 2024)

This edition applies to IBM Storage Virtualize Version 8.6.

Contents

Notices

This information was developed for products and services offered in the US. This material might be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, MD-NC119, Armonk, NY 10504-1785, US

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you provide in any way it believes appropriate without incurring any obligation to you.

The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to actual people or business enterprises is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at “Copyright and trademark information” at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks or registered trademarks of International Business Machines Corporation, and might also be trademarks or registered trademarks in other countries.

FlashCopy®	IBM FlashSystem®	Redbooks (logo) ®
HyperSwap®	IBM Spectrum®	
IBM®	Redbooks®	

The following terms are trademarks of other companies:

ITIL is a Registered Trade Mark of AXELOS Limited.

The registered trademark Linux® is used pursuant to a sublicense from the Linux Foundation, the exclusive licensee of Linus Torvalds, owner of the mark on a worldwide basis.

Red Hat, are trademarks or registered trademarks of Red Hat, Inc. or its subsidiaries in the United States and other countries.

Other company, product, or service names may be trademarks or service marks of others.

Preface

IBM® Storage Virtualize software running on IBM Storage products provides clients an opportunity to maintain two independent copies of a single data volume across two sites by using the IBM Storage Virtualize Metro Mirror feature.

Starting with Storage Virtualize V8.3.1.1, two site replication can be extended to a third site. The use of a three-site setup means data can be replicated asynchronously to an independent system in a third site. This three-site setup is called IBM Storage Virtualize 3-Site Replication. This IBM Redbooks publication includes a description of IBM Storage Virtualize 3-Site Replication and its implementation, and highlight configuration guidelines, preferred practices, and troubleshooting procedures.

This IBM Redbooks publication is intended for experienced storage, SAN, and IBM FlashSystem® Storage administrators and technicians. Understanding this book requires advanced knowledge of these environments.

Authors

This book was produced by a team of specialists from around the world.

Dr. Pradip Ashok Waykos is an Advisory Software Engineer at IBM India Systems Development Lab, Pune. He carries 13 years of diverse experience of enterprise software development in systems, storage, and information security domain and is proficient in storage virtualization, data replication, and REST. Leading the efforts for IBM Storage Virtualize 3-Site replication development, he has significantly contributed towards Tech Sales enablement, L3 support enablement, and beta programs for 3-Site configurations. He has authored multiple technical white papers, blogs, and collateral on various different Storage Virtualize features. Pradip holds a Masters in International Business Management and is pursuing his doctorate program as a Blockchain Research Scholar.

Prateek Sudin Mandge is a Storage Engineer working as the QA/Test Developer at IBM India Systems Development Lab, Pune. He possesses two years of professional experience and has been actively involved in the IBM Storage FlashSystem/SAN Volume Controller project since 2021. During his tenure, he successfully contributed to the testing and implementation of various features, notably the 3-Site Replication (Orchestrator High Availability) and IPsec Host attach.

Ajinkya Nanavati is an Advisory Software Engineer at IBM India Systems Development Lab, Pune. He carries 8 years of diverse experience of enterprise software development in systems, storage, and capital market domain. He is proficient in storage virtualization, data replication and different plugins, with active contribution in plugin development of embedded VASA provider for IBM Storage Virtualize. He has significantly contributed towards L3 support for 3site replication and plugins. He has authored multiple technical white papers, blogs, and collateral on various different Storage Virtualize features. Ajinkya holds a Masters degree in Information Technology from Pune University, India

Vasfi Gucer works as the Storage Team Leader on the IBM Redbooks Team. He has more than 30 years of experience in the areas of systems management, networking hardware, and software. He writes extensively and teaches IBM classes worldwide about IBM products. His focus has been primarily on storage, cloud computing, and cloud storage technologies for the

last 8 years. Vasfi is also an IBM Certified Senior IT Specialist, Project Management Professional (PMP), IT Infrastructure Library (ITIL) V2 Manager, and ITIL V3 Expert.

The team also would like to thank the authors of the previous edition of this publication *IBM Spectrum Virtualize 3-Site Replication*, SG84-8504:

Denis Olshanskiy, Detlef Helmbrecht, Jackson Shea, Jasmin Rapp, Rene Oehme, Sergey Kubin, Sushil Sharma and **Thomas Vogel**

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an IBM® Redbooks® residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our books to be as helpful as possible. Send us your comments about this book or other IBM Redbooks publications in one of the following ways:

- Use the online **Contact us** review Redbooks form found at:

  ibm.com/redbooks

- Send your comments in an email to:

  redbooks@us.ibm.com

- Mail your comments to:

  IBM Corporation, IBM Redbooks
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806

- Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm

- Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

1

Introduction

This chapter provides a general overview of the 3-Site Replication feature in IBM Storage Virtualize, supported architectures, and existing replication and copy services that are used to implement this feature and current updates.

This chapter includes the following topics:

1.1 IBM Storage Virtualize 3-Site Replication overview

This section describes the IBM Storage Virtualize 3-Site Replication feature.

With IBM storage systems with IBM Storage Virtualize software you can maintain two independent copies of a single data volume between two systems at two separate sites by using its Metro Mirror synchronous replication.

The data is replicated synchronously from an IBM Storage Virtualize cluster at a first site to another IBM Storage Virtualize cluster in a second site. These two sites are independent from each other and ideally, in separate failure domains. If one site fails, the other site still has the same data at the surviving site.

Starting with IBM Storage Virtualize V8.3.1.1, this two-site replication can be extended to a third site for added protection. The use of a three-site setup means that data can be replicated *asynchronously* to an independent system in a third site. The third site can be far away from the first and second site geographically. Therefore, the replication uses an asynchronous replication technique. This three-site setup is called *IBM Storage Virtualize 3-Site Replication*.

3-Site Replication is managed, operated, and maintained by an external application that is called the *Orchestrator*. The Orchestrator runs on a Linux system that is independent from the IBM Storage Virtualize software, ideally not by using the IBM Storage Virtualize storage that it is managing.

3-Site Replication features the following properties:

- Three independent IBM Storage Virtualize clusters. An outage of one cluster does not affect the accessibility of another cluster.
- Two sites (Master site and AuxNear site) replicate volume data synchronously by using Metro Mirror communication.
- Volume data is replicated asynchronously to the third site.

  Note: In this document, the third site is referred to interchangeably as both AuxFar and auxiliary-far.

- When using IBM HyperSwap topology, two independent clusters with three independent sites: One HyperSwap cluster at the Master and AuxNear sites as sites 1 and 2 and another independent cluster as the AuxFar site. An outage of one site does not affect data accessibility on the other sites.
- Dedicated Orchestrator software assures operation of 3-Site Replication independently from the status and accessibility of any 3-Site Replication cluster.

An overview of a 3-Site Replication setup is shown in Figure 1-1, which includes the following sites:

- Site 1: Master site
- Site 2: AuxNear site
- Site 3: AuxFar site

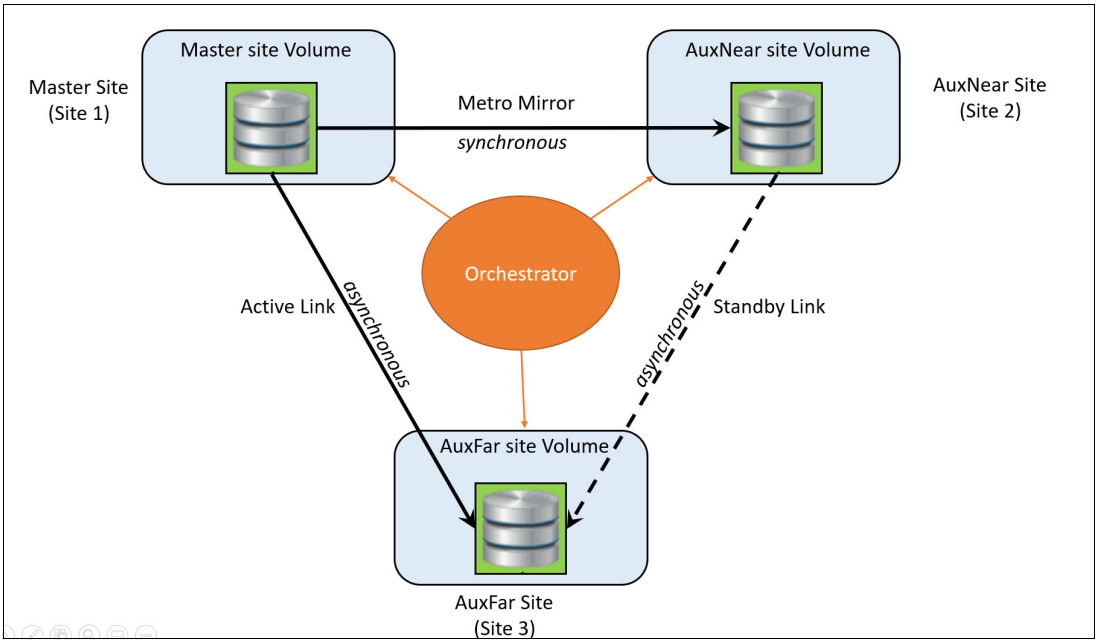

Figure 1-1 3-Site Replication overview, volume, and site point of view

Replication between one of the Near Sites and the AuxFar site uses the active link. The active and the standby link can be switched, as discussed in Chapter 5, “3-Site Replication monitoring and maintenance” on page 99. In short, the *active* periodic source determines the *active* link and the *inactive* periodic source determines the *standby* link.

The example in Figure 1-1provides a simplified overview to demonstrate the 3-site concept, and shows only one volume of a consistency group.

Note: Remote Copy consistency groups are a key component of 3-Site Replication. They are often used to group multiple volumes and relationships for a single consistent point-in-time action and are not necessary for single volumes. However, for 3-Site Replication, all actions are performed on consistency groups. Therefore, even if only one volume is in a single set of relationships, it must be in a consistency group.

Each of the three sites has a dedicated function. The AuxFar site always receives data from the Master site *or* AuxNear site. In normal operation, data is not replicated from the AuxFar site to any other site. This configuration might change in a failure scenario. For more information about failure scenarios, see in Chapter 6, “Failure protection and recovery procedures for 3-Site Replication” on page 125).

The Orchestrator does not have any requirements for where it is located, but the location must have reliable network connectivity to the systems at all three locations. In this example, the AuxFar Site (Site 3) is the suggested location. For HyperSwap configurations, Orchestrator can run on the same server on which the IP quorum is located.

Terminology

IBM Storage Virtualize 3-Site Replication uses the terminology that is listed in Table 1-1 to determine the sites in a 3-Site setup.

Table 1-1 3-Site replication terminology

Term	Definition
Near Sites	Set of two sites that are originally configured in Metro Mirror remote copy between two storage systems or HyperSwap topology within a single system. These sites can be the same physical site with two sufficiently isolated fault domains. This term does not exist in any commands or GUI references, but is used frequently in this document.
Master Site	The Near Site that is specified as the first site when configuring 3-Site Replication with a Metro Mirror configuration. This Master Site is often where the production workload exists. For HyperSwap, Site 1 is always Master. See Master Site (Site 1) in Figure 1-1 on page 7.
AuxNear Site	The Near Site that is specified as the second site when configuring 3-Site Replication with a Metro Mirror configuration. This AuxNear Site often is where the original recovery site exists. For HyperSwap, Site 2 is always AuxNear site; see AuxNear (Site 2) in Figure 1-1 on page 7.
AuxFar or auxiliary-far site or DR Site	The geographically separated site to which the Near Sites replicate a tertiary copy of the data on a periodic basis; see AuxFar (Site 3) in Figure 1-1 on page 7.
AccessPoint	Dedicated IBM FlashCopy® volumes that are used for replication between the Master and AuxFar sites or between the AuxNear and AuxFar sites by way of an asynchronous link.
3-Site remote copy (RC) consistency group (tscg)	This concept was introduced by 3-Site Replication. It is not a real IBM Storage Virtualize configuration object, but consists of the following independent remote copy consistency groups (RCCGs): ► RCCG between Master and AuxNear ► RCCG between Master and AuxFar ► RCCG between AuxNear and AuxFar It is created from an existing two-site RCCG through a conversion process where the two extra RCCGs are created along with the AccessPoints. The administrator can see an aggregated view of a `tscg` by using the Orchestrator or GUI (if running IBM Storage Virtualize 8.4.0 or later code). The state of a `tscg` is derived from the independent states of the three different RCCGs. The `tscg` does not have a unique object ID associated with it.
3-Site RC Relationship	This concept was introduced by 3-Site Replication. It is not an IBM Storage Virtualize configuration object, but consists of the following independent RCRELs between three different volumes at each of the three sites: ► RCREL between Master and AuxNear ► RCREL between Master and AuxFar ► RCREL between AuxNear and AuxFar It is created from an existing two-site RCRELs through a conversion process where the two extra RCRELs are created along with the AccessPoints. The administrator can see an aggregated view of a 3-site RC relationship by using the Orchestrator. Each 3-Site RC relationship is identified by a unique administrator-provided name. 3-Site RC relationships do not have a unique object ID associated with them.
3-Site configuration file	This configuration file contains the basic configuration attributes of the Master, AuxNear, and AuxFar sites, such as their IP addresses, cluster names, and access credential information.
freeze_time	The freeze_time attribute is identical in concept used in Global Mirror with Change Volumes (GMCV). It is the timestamp of the last successful transfer of all data within a periodic cycle that is needed for the AuxFar site to have a crash consistent copy of the Near Site volumes. Similar to GMCV, the cycle period might need to be increased to allow more time to replicate the changed data if it is observed that the freeze time is not current.

Primary Site	The production site where application writes are occurring. This site is the source of synchronous replication between the Near Sites and is used to determine the direction of replication in standard Metro Mirror configurations. This replication is dynamically and automatically handled by IBM Storage Virtualize in HyperSwap® topologies. In a 3-Site RC consistency group, only one of the two Near Sites can be made primary.
Periodic Source	One of the two Near Sites that is actively replicating data to the AuxFar site on a periodic basis. A key advantage of the IBM Storage Virtualize 3-Site Replication is the ability to dynamically switch periodic source.
Active Link	Link that is actively used for asynchronous or periodic replication to the AuxFar site (the Master-AuxFar or AuxNear-AuxFar).
Standby/ Inactive Link	Link that is not actively used for asynchronous or periodic replication to the AuxFar site (the Master-AuxFar or AuxNear-AuxFar).
3-Site User Group	New 3SiteAdmin user group that was introduced to enable access to the 3-site capabilities of IBM Storage Virtualize. The user that is created for 3-Site Replication management must belong to the role of 3SiteAdmin.
3-Site user	User with role 3SiteAdmin that is required for configuring 3-Site Replication across three IBM Storage Virtualize systems. All three sites must have this user who is created with the same name.

1.2 New features of 3-Site Orchestrator 4.1

3-Site Orchestrator 4.1 supports Metro Mirror or HyperSwap, GUI support, adding volumes with data from an existing 3-Site consistency group, removing volumes with data from an existing 3-Site consistency group, which is also supported on Orchestrator 4.0, and Orchestrator high availability.

3-Site Orchestrator 4.1 provides the following new features:

- Adding volumes with data from an existing 3-Site consistency group.
- Removing volumes with data from an existing 3-Site consistency group.
- Orchestrator high availability.

1.3 Copy Services used by 3-Site Replication

This section describes the components in IBM Storage Virtualize that are used to provide the 3-Site Replication feature:

- Metro Mirror
- FlashCopy
- Topology. Standard versus HyperSwap; Star versus Cascade
- Consistency groups
- Putting it all together

1.3.1 Metro Mirror

IBM 3-Site Replication uses Metro Mirror replicate data between sites. Remote copy consistency groups are used to ensure consistency among dependent writes and volumes.

Metro Mirror establishes a synchronous relationship between two volumes of equal size. The volumes in a Metro Mirror relationship are referred to as the *master* volume and the *auxiliary*

volumes. Traditional Metro Mirror is primarily used in a metropolitan area, or geographical area, up to a maximum distance of 300 Km (186.4 miles) to provide synchronous replication of data.

The volume properties of master and auxiliary are defined when the MM relationship between two volumes is configured. The master volume is considered as a production volume and the updates to this volume are mirrored to the auxiliary volume. However, the MM relationship direction copy can be changed. For example, if an event occurs at the master site and the application is brought up at the auxiliary site, you might want ti set the auxiliary volume to be the primary volume by using the **`-primary`** option to **`switchrcconsistgrp`** command and replicate back to the master.

The master volume serves I/O for a host, and writes are mirrored to the auxiliary volume. When Metro Mirror relationships are active, auxiliary volumes are not accessible for host Write I/O and are accessible in a read-only state without the possibility of mounting it to the host. However, an auxiliary volume can be used as a boot volume to expedite recovery.

To get the full access to the auxiliary volume, the Metro Mirror relationship must be stopped by using the **`-access`** parameter. The direction of the relationship can also be switched so that the auxiliary volume becomes the source of replication. In this case, be cautious because the writes to the auxiliary volume are mirrored back to the master volume and thus overwrites it.

With synchronous copies, host applications write to the master volume and the cache of the master volume is mirrored to the cache of the auxiliary volume before acknowledgment of the write is sent back to the host. This action ensures that both of the volumes feature identical data when the write completes.

After the initial copy completes, the Metro Mirror relationship always maintains a fully synchronized copy of the master site data at the auxiliary (AuxNear) site.

Consistency groups are used to maintain data integrity during dependent writes to multiple volumes.

3-Site Replication expands the concept of Metro Mirror consistency groups with a copy at a far third site.

For more information about Metro Mirror, see the following publications:

- *Implementation Guide for IBM Storage FlashSystem and IBM SAN Volume Controller: Updated for IBM Storage Virtualize Version 8.6*, SG24-8542
- *Performance and Best Practices Guide for IBM Storage FlashSystem and IBM SAN Volume Controller: Updated for IBM Storage Virtualize Version 8.6,* SG24-8543

1.3.2 IBM Storage Virtualize FlashCopy

The IBM Storage Virtualize FlashCopy capability provides usable Point-in-Time (PiT) copies of one or more volumes, even during synchronization. Because the copy occurs at the block level, it does not affect the application and host layer.

The following GUI presets for FlashCopy describe the three principal use cases:

- Snapshot: Sometimes referred to as *no-copy*, only the changed blocks of the source volume are copied to maintain the point in time. The target volume cannot be used without the source volume. This type is used for checkpoints and uses the least amount of space.

- Clone: Referred to as a *one-time fullcopy*. After the copy operation completes and the synchronization target volume becomes a fully independent volume, the mapping that was created by the preset is automatically deleted.
- Backup: Referred as *incremental backup*, consists of PiT full copy of the data, retains the mapping, and tracks changes for future periodic incremental refreshes. As the name implies, this preset is frequently used for backups. The full copy decouples heavy reads from the primary volume.

PiT copies with some additional features for tracking increments between iterations are a key component of IBM Storage Virtualize 3-Site Replication.

FlashCopy is used to create the AccessPoint volumes to achieve 3-Site Replication because usual 2-Site remote replication, such as Metro Mirror, is limited. In 2-site replication, the source volume, which becomes the master volume, can be in only one relationship with a remote volume, which becomes the auxiliary volume. The AccessPoint helps overcome this limitation and provides the necessary protection to the replicating data.

Creation and configuration of AccessPoints is controlled by the Orchestrator and is a part of the 3-Site configuration process.

Detailed information about IBM FlashCopy service functions and configuration is available in the following publications:

- *Implementation Guide for IBM Storage FlashSystem and IBM SAN Volume Controller: Updated for IBM Storage Virtualize Version 8.6*, SG24-8542
- *Performance and Best Practices Guide for IBM Storage FlashSystem and IBM SAN Volume Controller: Updated for IBM Storage Virtualize Version 8.6,* SG24-8543

1.3.3 Topology

When comparing a Resiliency Topology to a Replication Topology, the term *topology* is applied differently:

- *Resiliency topology*: Standard versus HyperSwap
- *Replication topology*: Star versus Cascade

Resiliency topology

Currently, 3-Site Replication is possible with following resiliency solution topologies:

- Classic topology: The Production site (Active site) is synchronous with the Disaster Recovery site (Standby site) and one or the other NearSite is asynchronously replicated to the Third site. Each of the sites is also represented as independent IBM Storage Virtualize clusters (for example, SAN Volume Clusters or Flash System clusters).
- HyperSwap topology: Where three independent sites with two independent clusters exist. The highly available Production cluster exists in the two Near Sites and the Disaster Recovery cluster exists in the third or Far site.

Figure 1-2 and Figure 1-3 show the solution topologies. Notice that three independent Storage clusters are used, one for each site.

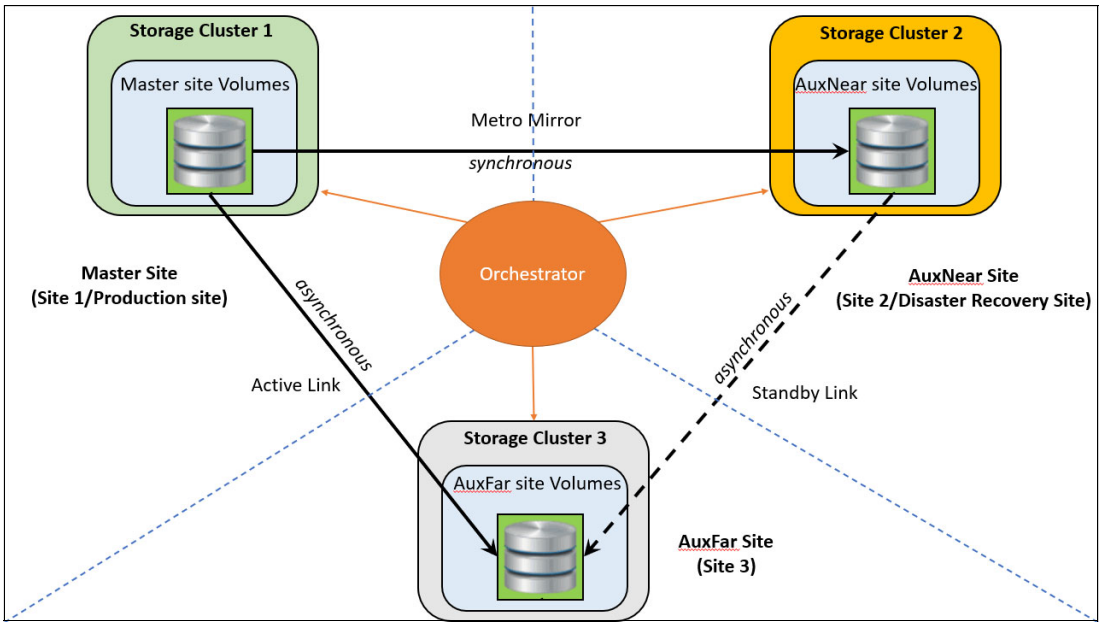

Figure 1-2 Production site and Disaster Recovery site with asynchronous replication to AuxFar site

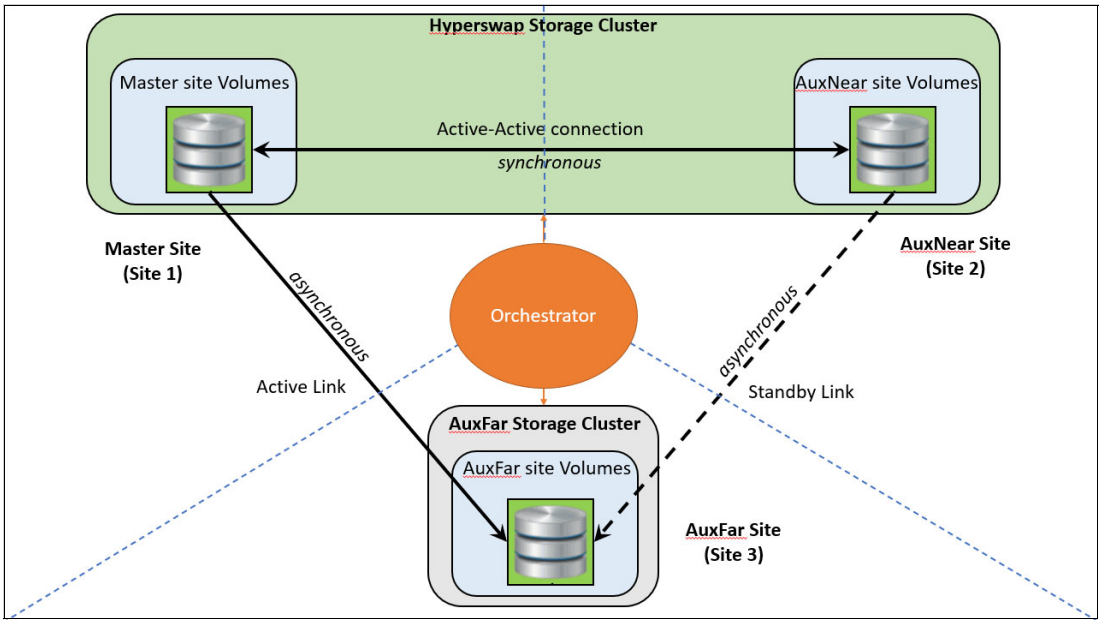

Figure 1-3 Master site and AuxNear site and asynchronous mirroring with AuxFar site

In Figure 1-3, notice that two independent Storage clusters exist and the Production site and Disaster Recovery site are not divided in HyperSwap topology.

Replication topology

For data replication, 3-Site Replication uses Metro Mirror with the remote copy consistency groups feature to maintain data consistency.

The data of the volumes in the consistency group is replicated by using one of the following replication topologies:

- Star topology:
  - The Master site volumes synchronously mirror data to the AuxNear site volumes.
  - The Master site volumes asynchronously or periodically replicate data to the AuxFar site volumes.

  The Master site volumes are the source of asynchronous/periodic replication to the AuxFar site and AuxNear site (see Figure 1-4).

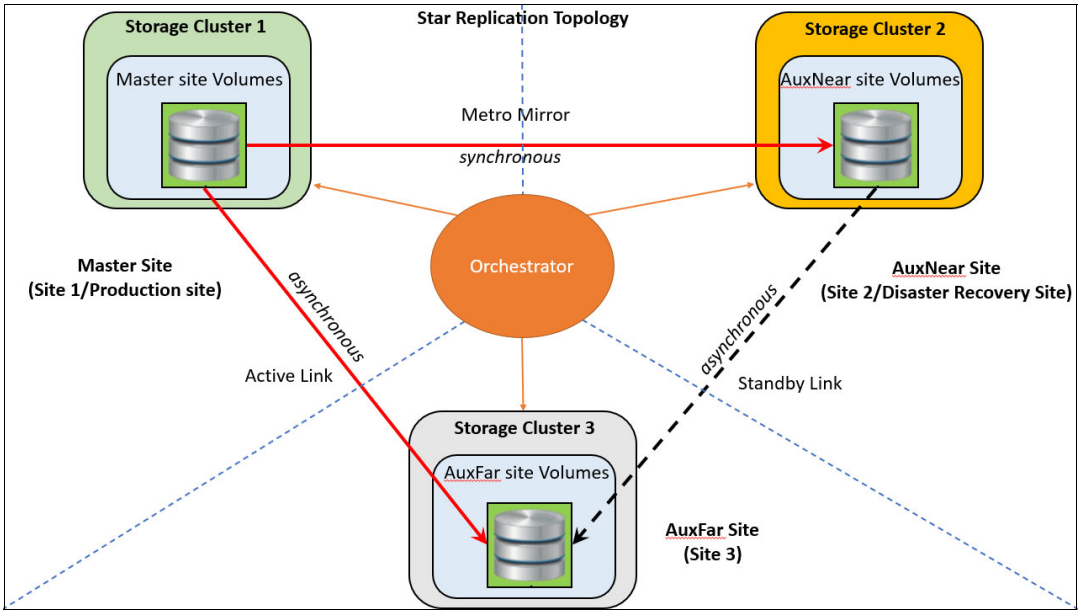

Figure 1-4 Star topology. Notice red arrows showing the replication direction

- Cascade topology:
  - The Master site volumes synchronously mirror data to the AuxNear site volumes.
  - The AuxNear site volumes asynchronously replicate data to the AuxFar site volumes.

The AuxNear site volumes are the source of asynchronous or periodic replication to the AuxFar site (see Figure 1-5).

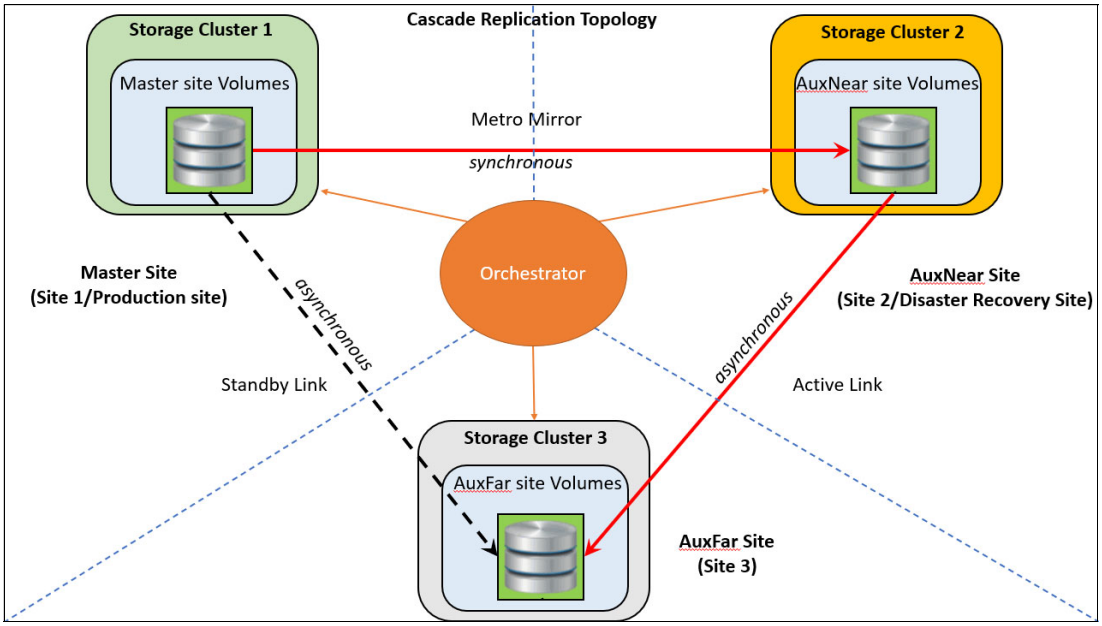

Figure 1-5 Cascade Replication topology

The source of a data replication to the AuxFar site for volumes in the consistency group is the Master site or AuxNear site. Only one of the sites is considered as the source site for replication to the AuxFar site.

An asynchronous replication link is called *active* if it is used for replicating data from the source site to the AuxFar site.

A replication asynchronous periodic link is called *standby* if it is *not* used for replication from the source site to the AuxFar site.

A key benefit of the IBM Storage Virtualize 3-Site Replication is its ability to dynamically choose and switch between Star and Cascade topology by switching the *periodic source*. This feature provides a tremendous savings in bandwidth and space.

Alternative IBM Storage Virtualize solutions for achieving a similar level of protection require concurrent replications from both sites, which use twice the amount of storage at the AuxFar site.0

For example, Cascade topology can be chosen for performance considerations. You might not want to load your Master or Production site with asynchronous replication traffic in addition to daily operation. You can move this specific role to your auxiliary-near (AuxNear) site.

1.3.4 3-Site Replication: Consistency groups

IBM Storage Virtualize 3-Site Replication uses Metro Mirror RCRELs between every two sites and groups them in RCCGs. However, a notion exists of a two-site versus a three-site remote copy consistency group.

After a two-site RCCG is converted to a three-site RCCG, it creates another pair of RCCGs for the asynchronous replication to the AuxFar site. In the Standard resiliency topology, one RCCG is on the Master cluster and another RCCG is on the AuxNear cluster.

In a HyperSwap resiliency topology, because the Site 1 and Site 2 systems are in a single cluster, both of the new consistency groups are visible from that cluster.

The naming for an automatically generated consistency group adheres to the following rules, as shown in the following example:

`tscg[lr]_<letters and numbers>`

- `tscg` is the prefix for all consistency groups that are created by Orchestrator
- `l` (left) or `r` (right) denotes the site of the consistency group: `l`eft for the Master to AuxFar consistency group, and `r`ight for the AuxNear to AuxFar consistency group.
- *letters and numbers* is a unique suffix and identical for the left and right consistency group of the Master to AuxNear consistency group

Example 1-1 shows the details about the underlying RCCGs that are revealed by using **`lscconsistgrp`**.

Example 1-1 Underlying RCCGs on the Master cluster revealed by lscconsistgrp

```
IBM_2145:Master:superuser>lsrcconsistgrp

id name            master_cluster_id master_cluster_name aux_cluster_id   aux_cluster_name
0  CG0             0000020321E0ECE0  Master              000002042040B714 AuxNear
1  tscgl_tzqiti467 0000020321E0ECE0  Master              0000020421601F10 AuxFar
```

Example 1-2 shows the output from **`lscconsistgrp`**, which shows the `HS-3site-CG1` with the associated RCCGs for replicating the “left” (site1) or “right” (site2) copy to the AuxFar cluster (`SVC_SA2`).

Example 1-2 Underlying RCCGs on a HyperSwap cluster revealed by lscconsistgrp

```
IBM_FlashSystem:Master:superuser>lsrcconsistgrp

id name            master_cluster_id master_cluster_name aux_cluster_id   aux_cluster_name
4  HS-3site-CG1    0000010024E00206  Master              0000010024E00206 Master
5  tscgl_ejrbtn163 0000010024E00206  Master              0000020420603BB2 SVC_SA2
6  tscgr_ejrbtn163 0000010024E00206  Master              0000020420603BB2 SVC_SA2
```

1.3.5 3-Site Replication: Putting it all together

3-Site Replication uses and enhances IBM Storage Virtualize copy services and techniques to provide a disaster recovery capability to add to a highly available or synchronously mirrored environment.

The following Copy Services and resources are used for 3-Site Replication:

- Metro Mirror
- FlashCopy
- Consistency groups
- AccessPoint volumes

As described in , “Terminology” on page 8, AccessPoint volumes are dedicated FlashCopy target volumes for asynchronous data replication to the AuxFar site.

This section explains the concepts of data replication and data consistency in an example with one volume. The 3-Site Replication consistency group concept is discussed in 1.3.4, “3-Site Replication: Consistency groups” on page 15.

This example uses Star Replication topology. Therefore, the Master site volume is the source for the replication to AuxNear and AuxFar site. In HyperSwap terms, Site 1 is the periodic source to the Remote site. Also, the Master site volume is the volume that the hypothetical application uses for production. The Cascade topology is explained later in this section, but it is where the synchronous site is used as the periodic source.

With IBM Storage Virtualize Metro Mirror replication, a volume can participate in one remote copy relationship only.

This relationship features the following roles:

- On the Master site, the production volume is often referred as the *Master volume*.
- On the target volume for replication at the remote site, AuxNear is referred to as the *Auxiliary volume*.
- Under normal conditions, the replication direction is from the Master (primary data) to the Auxiliary (replicated data). However, the auxiliary can be set as the primary when the *replication direction* is reversed.

Note: When a relationship is listed, Primary is set to `Yes` in the settings for the replication source volume. Therefore, the target volume includes a primary flag that is set to `No` and can be referred to as Secondary.

In our example, the Master volume is in a Metro Mirror relationship with the auxiliary volume on the AuxNear Site. Because data is replicated from the Master volume, it is a primary volume in the relationship and the auxiliary volume is referred to as secondary.

For the 3-Site solution, the following replications are needed:

- Synchronous data replication to AuxNear Site
- Periodic, asynchronous data replication to AuxFar site

The first replication is done by the Metro Mirror relationship between the Master site volume and the AuxNear site volume and it is synchronous in nature. HyperSwap uses intracluster Metro Mirror.

Because the Master site volume is in a Metro Mirror copy relationship with AuxNear volume, it cannot be used for replication to AuxFar site.

To overcome this limitation, use a snapshot PiT copy of the Near Site volumes and create Metro Mirror relationships between those volumes to the AuxFar tertiary volume. This process is done to achieve asynchronous or periodic replication to the AuxFar site.

Each volume that is participating in 3-Site Replication has an AccessPoint on each site. For example, the Master site volume (Primary) has one AccessPoint, the AuxNear site volume (Secondary) has one AccessPoint, and the AuxFar volume (Tertiary) has two AccessPoints.

AccessPoints at the Master site are also called Left AccessPoints. AccessPoints at the AuxNear site are called right AccessPoints. AuxFar volumes have Left and Right AccessPoints.

Figure 1-6 shows the primary volume and its corresponding AccessPoint at the Master site.

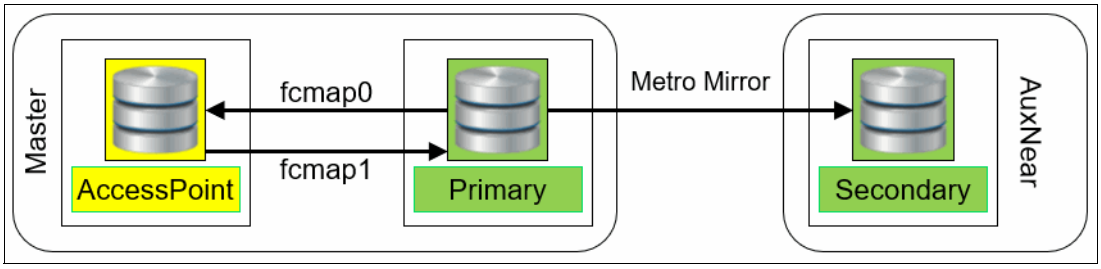

Figure 1-6 Primary volume with its AccessPoint at Master site

AccessPoints are accessible for the 3-Site Replication only and *not* for any other uses. Their behavior is different from standard volumes and normal IBM FlashCopy snapshots.

FlashCopy (FC) mappings are used to ensure consistency of the AccessPoints. Before data is replicated to the AuxFar site, a FlashCopy is taken from Primary volume. Normal FlashCopy behavior ensures that the split bitmap is used to read the data from the AccessPoint volume if it was split because of changed data on the primary volume, or through a redirected read to the source volume if not split.

Figure 1-7 on page 18 shows the Metro Mirror asynchronous relationship between the two left AccessPoints: the left AccessPoint of the Primary volume at the Master site, and the left AccessPoints of the third volume (Tertiary volume) at the AuxFar site. After synchronization is achieved for a periodic cycle and freeze time is established, the relationship is stopped and not continuously maintained.

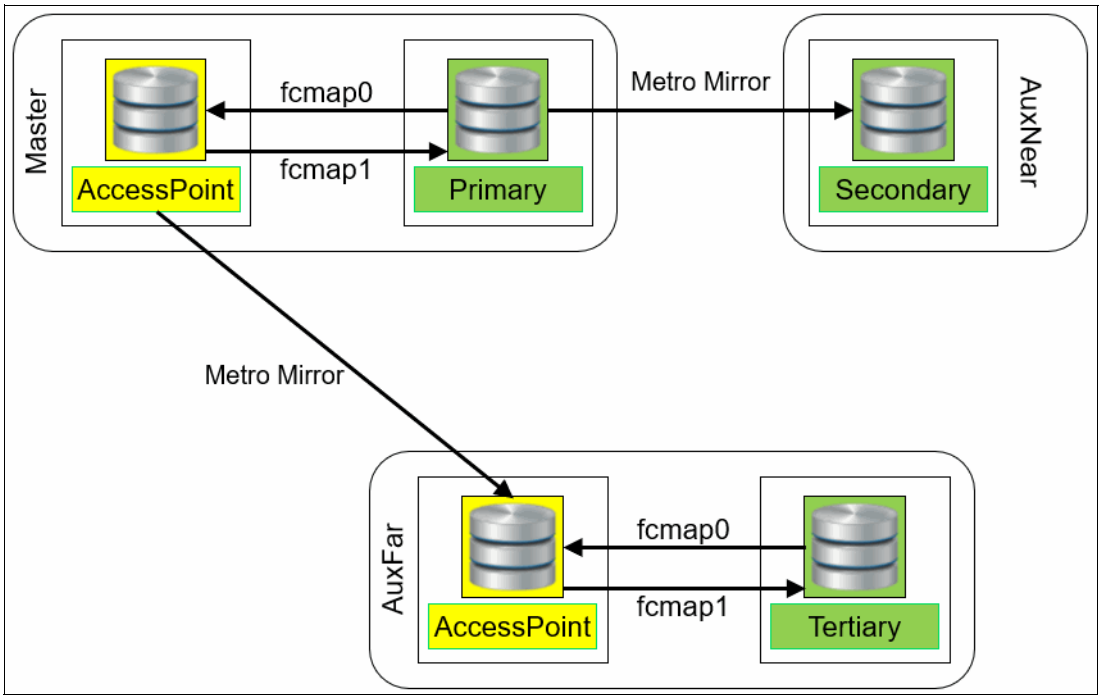

Figure 1-7 Left AccessPoints and their Metro Mirror relationship

Note: The replication between the two AccessPoints is implemented by using Metro Mirror replication. The Orchestrator is managing the remote replication between the Access Point in a cyclic or periodic manner instead of continuous, which is what makes the Metro Mirror relationship to the tertiary copy asynchronous, even though Metro Mirror is used. Typically, Metro Mirror is the synchronous replication option for remote copy and Global Mirror with Change Volumes is the asynchronous option.

Data is replicated periodically from the left AccessPoint to the AuxFar site. The initial replication fully backs up the data from the primary volume to the AuxFar site. Then, only the changes are periodically replicated to the AuxFar site.

The replicated data is crash consistent; that is, the write order is maintained. First, a FlashCopy snapshot of the primary volume is taken, which ensures that all writes before this point in time are included on the AccessPoint volume. Then, under Orchestrator's management, the Metro Mirror relationship between the AccessPoint of the primary volume and the AccessPoint of the tertiary volume transfers the data to the AuxFar site.

The Metro Mirror relationship is used for data transportation only. Data consistency is maintained by the FlashCopy techniques on the Master and AuxFar site. Therefore, the maximum Round Trip Time (RTT) between a Near Site and the AuxFar site is 250 ms, the same as for Global Mirror.

Data at the AuxFar site is not written to the AccessPoint volume but redirected to the Tertiary volume. This behavior is a specialized 3-Site Replication feature and has the following advantages:

- The changed blocks of the Tertiary volume are copied to its AccessPoint, which assures data consistency if the transfer is broken.
- A FlashCopy does *not* need to be reversed. A reversal is necessary only if the data is written to the AccessPoint.

If connectivity is lost between the Master and AuxFar site, you can switch to the Cascade replication topology. Then, AuxNear volumes become the periodic source to the AuxFar site. A full view of 3-Site Replication feature topology with AccessPoints for the left (Master) and the right (AuxNear) site is shown in Figure 1-8.

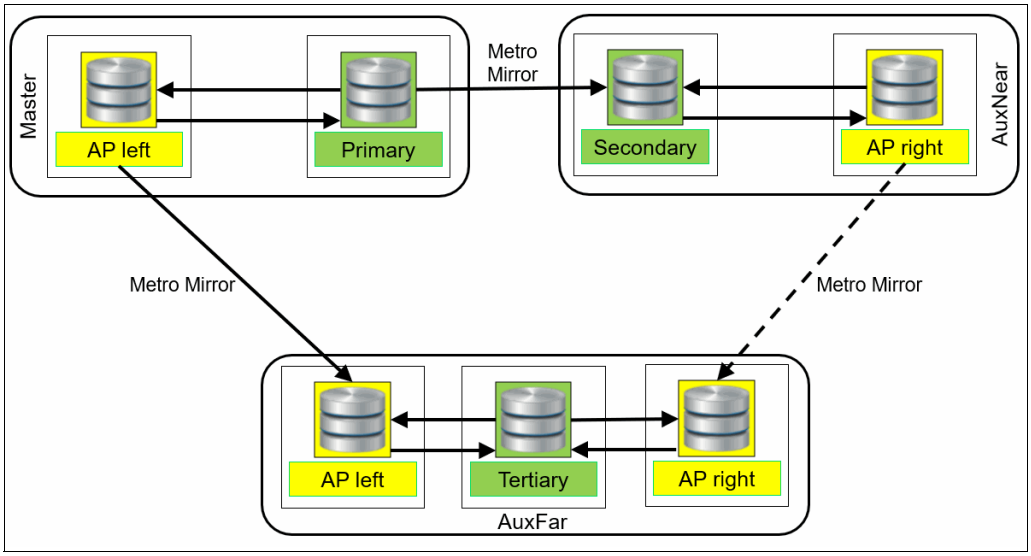

Figure 1-8 3-Site Replication volumes, AccessPoints, and FlashCopy mappings

The dashed arrow on the right of Figure 1-8 shows Standby link that can be activated if connectivity is lost between Master site and AuxFar site.

Figure 1-8 also shows the Metro Mirror connections between the three sites, the three volumes of the 3-Site Replication, and their AccessPoints. For more information about these relationships, see 1.3.4, “3-Site Replication: Consistency groups” on page 15.

For a specific consistency group and all its remote copy relationships, only one AuxFar remote copy is active to a Near Site. The active Near Site is called the *periodic source*. The left Metro Mirror is shown as a solid arrow Figure 1-8 to indicate that the active remote copy relationships are on the left site, and the Master site is the *periodic source*.

1.4 Data consistency

The Metro Mirror relationship between a Near Site and the AuxFar site is used for transporting data. Data consistency is provided by using FlashCopy snapshots on the Near Site and AuxFar site.

Coordinated FlashCopy snapshots are taken of the primary volume and the secondary volumes at the Master and AuxNear sites at the same point to ensure that Flash Copies of both Near Sites contain the same data. Therefore, these FlashCopy AccessPoints of the primary and secondary volumes contain identical data and are crash consistent; that is, the write order is maintained and the dynamic switching of periodic source is enabled.

New and changed data is periodically mirrored from a Near Site AccessPoint volume to the tertiary volume by using Metro Mirror. The FlashCopy bitmap is used to read the data from the AccessPoint volume after the FlashCopy is established while data is changing on the primary or secondary volumes.

Copy-on-write maintains the data of the previous cycle in the tertiary volume with that volume's associated AccessPoint. If needed, a roll-back to the state before the mirroring started is always possible.

When all data from a Near Site is replicated, the Metro Mirror is stopped and the FlashCopy from the tertiary volume to its AccessPoint is refreshed to establish the 3-Site freeze time. The process and concepts are similar to Global Mirror with Change Volumes, although the details, underlying mechanisms, and some of the nomenclature differs.

Note: The Metro Mirror is defined between the Near Site AccessPoint and the AccessPoint of the tertiary volume. However, data is directly replicated from the Near Site AccessPoint volume to the tertiary volume. No other reverse copying is needed. This dedicated feature is for the 3-Site Replication. FlashCopy mappings of the tertiary volumes ensure data consistency.

Figure 1-9 shows a high-level example of the Metro Mirror periodic data replication cycle from a Near Site to AuxFar.

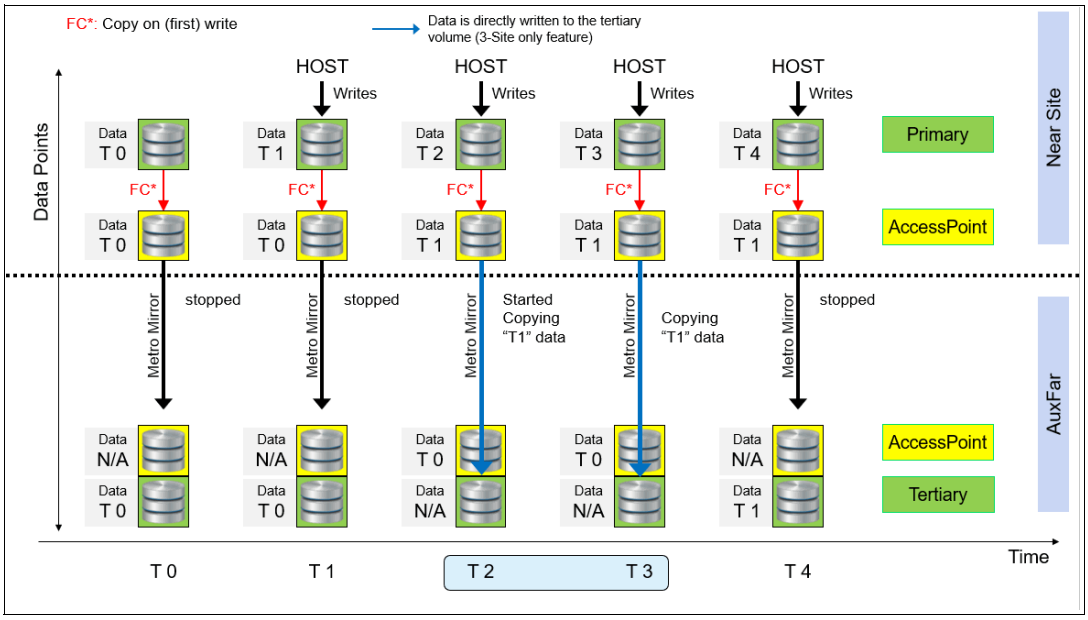

Figure 1-9 The 3-Site Replication periodic cycle

This example assumes that an initial cycle was completed and at least one freeze timestamps T2 and T3. The example also assumes that Metro Mirror is active and transferring the data, and the AuxFar AccessPoint provides consistency protection during the transfer and until a new freeze time is established.

Figure 1-9 also shows the data content of the primary and tertiary volume at different points. The date and its timestamps are shown to the left to the volume icon.

The volumes and status of the Metro Mirror at the different points include the following content:

- Time stamp T0:
  - All data was transferred to the third site by using Metro Mirror and freeze time achieved.
  - The FlashCopy "Primary → AccessPoint" is running (status "copying").

- Metro Mirror is stopped (state “consistent_stopped “) and the changed AccessPoint blocks are recorded, which must be transferred to the third site in the next cycle.

- Time stamp T1:
  - The host writes data to the primary volume.
  - The AccessPoint keeps the changed blocks by using “copy on (first) write”.
  - Metro Mirror keeps the information on the changed blocks.
- Time stamp T2:
  - A new primary volume FlashCopy is taken at T2, and the Metro Mirror to AuxFar is started. The FlashCopy behavior assures data consistency by way of the AuxFar AccessPoints. Because of two AccessPoints at the AuxFar site, the *left* for the Master or Site1 and the *right* for AuxNear or Site2, the standby maintains consistency and the active mediates the data transfer.
  - The Metro Mirror replication that has started to AuxFar is in the state inconsistent_copying, and is transferring the Point-in-Time (PiT) delta between T0 and T2 to the third site. Standard FlashCopy bitmap behavior ensures that the data is read from the primary volume or from the data from the AccessPoint if it changed.

    Although the Metro Mirror relationship is defined between the two AccessPoints, the data is directly written to the tertiary volume, and the changed blocks are saved in the AccessPoint volume for data consistency. feature of 3-Site Replication includes this dedicated feature.
- Time stamp T3:
  - The host writes data to the primary volume.
  - AccessPoint contains changed blocks to preserve PiT (by way of “copy-on-write”).
- Time stamp T4:
  - Metro Mirror transferred all T1 PiT data to the third site, a new freeze-time is recorded, and the AuxFar AccessPoint now holds a consistent copy of T1.
  - Metro Mirror is stopped with a state of `consistent_stopped`.
  - This T4 is equivalent to timestamp T0 where a freeze time was reached. The volumes are ready for the next periodic cycle.

1.5 3-Site user

The IBM Storage Virtualize 3-Site Replication feature relies on a dedicated user role, such as security administrator, monitor, and so on. The 3-Site Orchestrator requires a user that has the 3-Site administrator role to connect to the clusters participating in 3-Site Replication, and to run commands to manage those functions. Creating this user is described in Chapter 3, “Configuring a 3-Site Replication solution” on page 37.

Command-line interface: The IBM Storage Virtualize command-line interface (CLI) shows the details of the involved volumes, FlashCopy mappings, remote copy relationships, and remote copy consistency groups.

For more information about shows every IBM Storage Virtualize cluster, see Appendix A, “IBM Storage Virtualize 3-Site Replication components” on page 199.

2

Planning for 3-Site Replication

This chapter describes the steps that are required to plan the installation and configuration of 3-Site Replication in your environment. It includes inter-site link and partnership requirements, 3-Site Orchestrator requirements, partnership planning, capacity planning for access points, general recommendations, and restrictions.

Note: Make sure that the planned configuration is reviewed by IBM or an IBM Business Partner before implementation. Such a review can increase the quality of the final solution and prevent configuration errors that can affect solution delivery.

Consider planning as an interactive process so that the initial statement of goals and requirements can be modified to take account of factors such as costs, availability of personnel and skills, technology features and limitations, and opportunities that arise.

This chapter includes the following topics:

2.1 Functional requirements

A *functional requirement* defines a function of a system or its components; a *function* is described as a specification of behavior between outputs and inputs.

Start from the requirements and then consider features that might help you to meet those requirements, rather than starting from the product features. An example of a functional requirement that might affect your implementation choices might be that the transaction latency on production writes never exceed 5 ms because of replication.

Non-functional requirements

A non-functional requirement specifies criteria that can be used to judge the operation of a system, rather than specific behaviors.

The following general areas are considered to fall non-functional requirements:

- Performance
- Capacity
- Availability
- Usability
- Security
- Privacy
- Maintainability
- Manageability
- Flexibility

An example of a non-functional requirement that might affect your implementation choices might be that the transaction latency on production writes never affects application performance because of replication.

Depending on the environment, a requirement like this can drive a choice for a 3-Site solution that is based on Metro Mirror, HyperSwap, or Enhanced Stretched Cluster. Each customer environment features its own set of non-functional requirements.

2.2 Business requirements

Business continuance requirements maps to the copy services data protection capabilities:

- High Availability (HA)
- Disaster Recovery (DR)

During the planning process, you must ask a series of questions for each application data set. For example, you might decide that you need to protect your core transactional applications, but not your data warehouse data sets, because they are easily regenerated. Whenever data protection is discussed, think in terms of Recovery Point Objective (RPO) and Recovery Time Objective (RTO).

Question to ask for each application: How much data loss is acceptable?

Recovery Point Objective

RPO is defined as the amount of data loss, expressed in time units. Thus, an RPO of 1 minute means that when you recover from an outage, you lost the data that was updated in the last minute before the problem struck.

Question to ask for each application: How long can you afford to be offline?

Recovery Time Objective

RTO is defined as the amount of time that it takes you to get back up and running. Understand, though, that having the storage system ready to go again is not the entire picture. It might take considerable extra time after the storage system is accessible again to recover network and server access and restore application integrity, such as running a file system check and replaying log files.

2.2.1 High availability

An HA solution is designed to avoid loss of access to data on your site; that is, RPO=0 and RTO=0. If the HA mechanism acts quickly enough, no application outage occurs. The IBM copy services technology that maps most closely to this requirement is Volume Mirroring (formerly called VDisk mirroring), IBM SAN Volume Controller Stretched Cluster, and HyperSwap.

Question to ask for each application: Aside from “normal” HA, what other risks do you need to protect against, if any?

2.2.2 Disaster Recovery

Maintaining multiple sites with data copies ready to go is the usual way of ensuring rapid recovery from a large-scale failure.

The IBM Copy Services technologies that map most closely to a DR requirement are Metro Mirror, Global Mirror, and Global Mirror with Change Volumes. IBM SAN Volume Controller Enhanced Stretched Cluster and HyperSwap are also an option in some situations.

Along with RPO and RTO, consider budgets when undertaking DR planning. For example, the cost of network bandwidth can vary significantly by country or region. This difference might affect your choice. Licensing differences can also be a factor.

Question to ask for each application: Which technology choice provides the most cost-effective way to meet the recovery requirements?

Although the copies of data that are produced are crash-consistent, recovery routines, such as file system checks and log replays, might need to be run on them before applications can be used.

2.2.3 HA and DR site locations

When creating a multi-site solution, one of the key considerations is the geographical separation between the sites.

In 2003, the US Securities and Exchange Commission (SEC) suggested that 320 km (200 miles) is suitable between the primary and secondary sites for a US financial institution.

However, many commercial and governmental companies are content with distances of 30 km (18.64 miles) or even less as a compromise between performance, cost, convenience, and protection because most outages have relatively localized causes.

Questions to ask:

- What kind of disaster do I need to protect against?
- What performance effects can I tolerate?

2.3 3-Site Solution design: Metro Mirror or HyperSwap

Any 3-Site solution consists of a synchronous mirror between both near sites and an asynchronous mirror to the AuxFar site. The synchronous mirror between both near sites can be Metro Mirror or HyperSwap. Both options require a similar infrastructure and sizing, but they provide a fundamental difference in protection and system management.

2.3.1 Metro Mirror

Metro Mirror is a proven synchronous mirror between two independent storage systems. The 3-Site mirror solution that is based on Metro Mirror provides a data mirror between three independent storage systems.

This solution is based on a DR solution between the two near sites, supplemented with a third, asynchronous copy at the AuxFar sites. A synchronous mirror is always used between both near-site storage systems.

Mirroring to the third, independent site is always an asynchronous process. The Remote Copy activities must be managed between all three sites; no automatic failover or failback exists.

2.3.2 HyperSwap

HyperSwap is a proven synchronous mirror between two storage systems that are operating within a single cluster. The 3-Site mirror solution that is based on HyperSwap provides a data mirror between two independent storage systems, which are distributed across three sites.

This solution is based on a distributed HA solution that is in the two near sites, which is supplemented with a third, asynchronous copy at the AuxFar sites. A synchronous volume mirror is always used between both near-site storage systems. Mirroring to the third, independent site is always an asynchronous process. The Remote Copy activities must be managed for all copy operations to the AuxFar site. Failover and failback operations between both near sites are managed by HyperSwap.

2.4 Remote Mirroring concepts

This section describes the IBM Storage Virtualize relationship and consistency group concepts and the management capabilities in a 3-Site environment.

2.4.1 Remote Copy Relationship

Remote-copy relationships define the relationship between two volumes: a master volume and an auxiliary volume. Metro Mirror relationships are defined manually. HyperSwap relationships are created automatically when a HyperSwap volume is created in the management GUI or the command-line interface on systems that support HyperSwap topology.

Typically, a volume contains the master copy of the data and is the volume that the application normally accesses. The auxiliary volume typically contains a backup copy of the data and is used for DR.

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.

HyperSwap changes the copy direction automatically, depending on the amount of workload per site. The copy direction cannot be switched manually.

Metro Mirror copy directions can be changed manually only after the application is stopped and the primary volumes are unmapped.

The two volumes in a relationship must be the same size. The volume size cannot be changed for volumes in a 3-Site relationship.

For more information about relationships and relationship states, see Independent relationships.

2.4.2 Consistency groups

Consistency groups simplify managing related remote-copy relationships because copy actions are applied simultaneously to all relationships in the group. Consistency groups can be used to maintain consistent copies when data has dependencies across host applications or volumes.

Relationships that are not part of a consistency group are called *independent relationships*. All relationships in a consistency group must have matching primary (master) and secondary (auxiliary) systems or sites. All relationships in a consistency group must also have the same copy direction and state.

A 3-Site relationship can be assigned to a consistency group only; independent relationships cannot be used with 3-Site Mirroring.

For more information about consistency groups, see Consistency groups.

2.4.3 Managing copy relationships and consistency groups

A 3-Site environment features the following requirements and limitations:

- An independent relationship cannot be used for 3-Site mirroring. All 3-Site relationships must be assigned to a consistency group.
- A maximum of 16 consistency groups are supported.
- Up to 256 relationships are supported per consistency group.
- Up to 1250 (IBM SAN Volume Controller, FS9200, FS7200) and 1024 (FS5200) 3-Site Remote Copy relationships are supported.
- A relationship with existing data can be added to a 2-Site, but not to a 3-Site consistency group:
  - A relationship that is added to a 3-Site consistency group must contain empty volumes.
  - Adding a relationship with data to a 3-Site consistency group can cause inconsistent data at the AuxFar site and an unsuccessful DR.
- Relationships with active data can be added to a 3-Site consistency group by:
  - Converting the 3-Site consistency group to 2-Site
  - Adding of the 2-Site relationship to the 2-Site consistency group
  - Converting the consistency group to 3-Site (full data synchronization is required)
- A merge of multiple consistency groups is not possible. The situation can be managed by:
  - Converting both 3-Site consistency groups to 2-Site consistency groups
  - Removal of all relationships within one consistency group
  - Adding the 2-Site relationships to the 2-Site consistency group
  - Converting the consistency group to 3-Site (full data synchronization is required)
- A single relationship cannot be removed from a 3-Site consistency group. The situation can be managed by:
  - Converting the 3-Site consistency group to 2-Site
  - Removing the relationship
  - Converting the consistency group to 3-Site (full data synchronization is required)

Important: Adding a relationship with data to a 3-Site consistency group can cause inconsistent data at the volume at AuxFar site. For data, you must add these relationships to a 2-Site consistency group and convert these relationships to a 3-Site consistency group.

Based on those restrictions, careful planning is required. Consider the following points:

- Only volumes that are used from hosts from the same site are consolidated in the same consistency group.
- Consolidate all relationships in one or more 2-Site consistency groups before converting them to 3-Site and avoid changes later.
- Provide enough bandwidth from both near sites to the AuxFar site for data synchronization processes.

For more information, see Adding a Metro Mirror relationship to an existing 3-site consistency group.

2.4.4 Other requirements for 3-Site replication

This section provides the following general recommendations for your 3-Site replication environment planning:

- Ensure that you have the time zone configured correctly on all three systems and the 3-Site Orchestrator server. This configuration prevents timestamp inconsistencies and issues in your environment.
- Have a Network Time Protocol (NTP) server configured on all three systems and the 3-Site Orchestrator server. This configuration prevents timestamp inconsistencies and issues in your environment.
- Use only 3-Site Orchestrator (GUI or CLI) to manage your 3-Site objects.

2.4.5 Limitations and restrictions

When planning for your 3-Site Replication, consider the following restrictions and limits in place at the time of this writing:

- Standard Metro Mirror restrictions apply for the near site.
- Standard Global Mirror restrictions apply for the far site.
- Only one 3-Site Orchestrator host instance can be active at a time.
- A maximum of 16 consistency groups is supported by up to 256 relationships in each consistency group.
- Stand-alone relationships are not supported in 3-Site.
- The minimum cycle period for far site replication is 5 minutes. Based on this period, the minimum RPO is 10 minutes.
- Volume resizing is not supported.
- VVOLs are not supported on 3-Site Replication as VVOLs.
- Metro Mirror consistency protection feature is not supported in 3-Site Replication configurations.

Consider the following limitations:

- The 3-Site Orchestrator requires access to all three IBM Spectrum® Virtualize systems during 3-Site Orchestrator start or restart to detect the configuration. If one of those systems is not responding, the 3-Site Orchestrator cannot manage the current environment.
- The 3-Site Orchestrator is sometimes busy by running configuration change commands in the background (that is, conversion from 3-Site to 2-Site or creation of an AuxFar volume) for a long time.
- Native IP replication from a HyperSwap from both AuxNear sites to the AuxFar site can be considered.

Note: For more information about the IBM Storage Virtualize 8.6.0 limitation for native IP replication, see Release Note for systems built with IBM Storage Virtualize.

2.5 3-Site Orchestrator

The IBM Storage Virtualize 3-Site Orchestrator is available to configure and manage Metro Mirror configurations that span three sites. In addition, on system models that support HyperSwap configurations, you can use IBM Storage Virtualize 3-Site Orchestrator to extend the configuration to a third site. Both configurations require that you set up specific objects on each system and 3-Site Orchestrator settings.

The IBM Storage Virtualize 3-Site Orchestrator is based on a stateless and fault tolerant architecture for managing 3-Site Replication.

The 3-Site Orchestrator gathers and monitors the status from the three systems by:

- Running view commands
- Caching the results
- Being notified of changes

The 3-Site Orchestrator detects the environment by using SSH connections to the Storage systems. The SSH configuration and SSH security keys are stored on the Orchestrator.

No configuration or status information is permanently stored at the 3-Site Orchestrator. The Orchestrator is gathering the required information during start or restart of the system.

After initial gathering, the Orchestrator is permanently updating the configuration and status of the system by using SSH connections. With release 8.4.0.2 or higher, all three storage systems must be accessible during this process of initial gathering of the configuration, which enables 3-Site administration.

If a permanent loss of the 3-Site Orchestrator occurs, the recovery process is like a new Orchestrator installation and configuration, including gathering the configuration.

The 3-Site Orchestrator maintains the following main tasks:

- Configure and manage 3-Site configuration
- Issue consistency group cycle commands asynchronous replication.

The following commands are included in the Audit Log:

- Commands are issued to create and destroy 3-Site consistency groups, add volumes, remove volumes, switch periodic sources, and so on.

The commands to start each cycle of periodic data replication are *not* included in the Audit Log.

Important: Only one active instance of IBM Storage Virtualize 3-Site Orchestrator is supported. Multiple instances can result in unexpected data inconsistency and inaccurate statuses of the 3-Site configuration and 3-Site consistency groups.

To use the management GUI to configure and manage both 3-Site configurations, all systems must be running the 8.4.0.2 or later release.

2.5.1 3-Site Orchestrator requirements

When planning for the 3-Site Orchestrator installation, deploy a server that meets the following hardware and software requirements:

- Red Hat Enterprise Linux (RHEL) version 7.5 or later. Then, install the correct version of IBM Storage Virtualize 3-Site Orchestrator on an x64-bit architecture server.
- The server can be physical or virtual.
- The server has a minimum of 4 GB random access memory (RAM) and 4-core microprocessors.
- The server ha 300 MB free storage space for RPM file and installation.
- IP network access to all three sites is available.
- The OpenSSH utility is installed.
- Passwordless SSH login is available for all three sites from the 3-Site Orchestrator host.
- The physical location of the host can be at any of the three sites or at a fourth site.
- All three Storage Virtualize systems fulfill remote copy license requirements.
- 3-Site Orchestrator with Metro Mirror:
  - All systems are running the 8.3.1or later release.
  - All three Storage Virtualize systems include Fibre Channel remote copy partnerships.
  - Firewall port 7003 is open for inter-process communication between the Orchestrator Daemon and the Event Monitor Daemon.
  - Three firewall ports are open for events and SSH tunneling.
- 3-Site Orchestrator with HyperSwap:
  - All systems are running the 8.4.0 or later release.
  - Firewall port 7003 is open for inter-process communication between the Orchestrator Daemon and the Event Monitor Daemon.
  - Two firewall ports are open for events and SSH tunneling.

2.5.2 3-Site Orchestrator recommendations

As 3-Site Orchestrator is becoming a key component in your environment that manages 3-Site Mirroring and starts all cycling commands for asynchronous replication. Consider the following points:

- Consider HA for the 3-Site Orchestrator. The Orchestrator can survive a single site failure without restart or reinstallation.
- Be prepared against an Orchestrator failure by configuring and maintaining a second, independent Orchestrator instance, including suitable SSH keys and installation images.
- Running multiple instances can result in unexpected data inconsistency and inaccurate statuses of both the 3-Site configuration and 3-Site consistency groups.

- There are several Orchestrator releases are available. Only Orchestrator 3.0 or higher supports Metro Mirror, HyperSwap, and GUI. See Table 2-1. For more information, see IBM Storage Virtualize 3-Site Orchestrator Tool.

Table 2-1 Supported versions of the installation packages

Version of installation package	Supported 3-Site configurations	Supported management interfaces	Supported IBM Storage Virtualize releases
Orchestrator 1.0	Metro Mirror only	CLI	8.3.1, 8.4.0
Orchestrator 2.0	HyperSwap only	CLI	8.4.0
Orchestrator 3.0	Metro Mirror and HyperSwap configurations	GUI and CLI	8.4.0.2 or later
Orchestrator 3.1	Metro Mirror and HyperSwap configurations	GUI and CLI	8.4.0.2 or later
Orchestrator 4.0	Metro Mirror and HyperSwap configurations	GUI and CLI	8.4.0.2 or later
Orchestrator 4.1	Metro Mirror and HyperSwap configurations	GUI and CLI	8.4.0.2 or later

2.6 Inter-site link and partnership planning and requirements

Planning link and partnership requirements can be distinguished by the following parts:

- Planning for Synchronous Mirror between both Near sites (Metro Mirror/HyperSwap).
- Planning for Asynchronous Mirror to the AuxFar site.

Note: Because of an issue in IP Replication, customers that use this feature should not upgrade to v8.4.0.0 or later. For more information, see Release Note for systems built with IBM Spectrum® Virtualize.

2.6.1 Planning for Synchronous Mirror between both Near sites

Synchronous replication (Metro Mirror or HyperSwap) can be separated by up to 300 km (186.4 miles) in some Fibre Channel configurations. Remember that a 100 km (62.1 mile) separation adds approximately 1 ms transaction latency for each round trip. Because some operations require multiple round-trip transactions, the overall latency might increase significantly.

Often, a Fibre Channel-based environment with synchronous mirroring is established with a distance of less than 10 km (6.21 miles) up to 30 km (18.6 miles). Larger distances might incur a significant higher response time compared to up-to-date SSD-based or FCM-based storage systems.

IP connections can be used for replication in the following ways:

- FCIP router can be used for protocol conversion from Fibre Channel to Ethernet.
- Native IP connections can be used sometimes.

In both cases, Link bandwidth and quality significantly affect the overall capabilities and response time.

Inter-site bandwidth calculation

The two major parameters of a link are its bandwidth and latency. Latency might limit the maximum bandwidth that is available over IP links, depending on the details of the technology that is used (such as the circuit type, SAN hardware, and connections).

The synchronous replication between both near sites (Metro Mirror and HyperSwap) requires bandwidth for a Storage System Intersystem heartbeat. Table 2-2 lists the amount of heartbeat traffic (in megabits per second) that is generated by various sizes of clustered systems.

Table 2-2 Intersystem heartbeat traffic in Mbps

SAN Volume Controller System 1	**SAN Volume Controller System 2**			
	2 nodes	**4 nodes**	**6 nodes**	**8 nodes**
2 nodes	5	6	6	6
4 nodes	6	10	11	12
6 nodes	6	11	16	17
8 nodes	6	12	17	21

Note: Table 2-2 uses the IBM SAN Volume Controller as an example. You can use the same values based on your FlashSystem controllers amount.

These numbers estimate the amount of traffic between the two clustered systems when no I/O is occurring on mirrored volumes. Half of the data is sent by each of the systems. The traffic is divided evenly over all available inter-cluster links. Therefore, if you have two redundant links, half of this traffic is sent over each link.

The bandwidth between sites must be sized to meet the peak workload requirements. You can estimate the peak workload requirement by measuring the maximum write workload that is averaged over a period of 1 minute or less and adding the heartbeat bandwidth. Statistics must be gathered over a typical application I/O workload cycle, which might be days, weeks, or months, depending on the environment on which the system is used.

When planning the inter-site link, also consider the initial sync and any future resync workloads. It might be worthwhile to secure more link bandwidth for the initial data synchronization.

If the link between the sites is configured with redundancy so that it can tolerate single failures, you must size the link so that the bandwidth and latency requirements are met even during single failure conditions.

Therefore, link bandwidth sizing for synchronous replication is based on the following factors:

- Peak Write workload
- Initial synchronization or resynchronization workload
- Buffer for future growth
- Storage System Inter-system heartbeat traffic

- Bandwidth for writes to the secondary site until the switch of the replication directions removes this effect (HyperSwap only)

When planning the inter-site link, carefully note whether it is dedicated to the inter-cluster traffic or if it is going to be used to carry any other data. Sharing the link with other traffic (for example, cross-site IP traffic) might reduce the cost of creating the inter-site connection and improve link usage. However, doing so might also affect the links' ability to provide the required bandwidth for data replication and might result in unacceptable network latency

Note: Contact your IBM Sales representative or IBM Business Partner to perform these calculations.

2.6.2 Planning for asynchronous mirror to the AuxFar site

To plan for asynchronous replication in a 3-Site environment, consider that Fibre Channel is fully supported.

Asynchronous replication allows data replication over long distances, depending on the round-trip latency between sites. Up to 80 ms round-trip delay time is supported in a Fibre Channel infrastructure. Up to 250 ms round-trip delay is supported, if you meet the following requirements:

- All nodes that are used for replication are of a supported model.
- A Fibre Channel partnership exists between the systems, not an IP partnership.
- The remote copy buffer size setting is 512 MB on each system in the partnership. This setting can be accomplished by running the **`chsystem -rcbuffersize 512`** command on each system.
- Two Fibre Channel ports on each node that is used for replication are dedicated for replication traffic by using SAN zoning and port masking.
- SAN zoning is applied to provide separate inter-system zones for each local-remote I/O group pair that is used for replication.

Link bandwidth sizing for asynchronous replication depends on the following parameters:

- Amount of Write workload
- Write workload distribution across grains
- Replication cycle time
- Buffer for future growth
- Storage System Inter-system heartbeat traffic:
  - A detailed link planning is not possible for such an environment because workload distribution across grains can change permanently. It is considered a best practice to size the environment for a short cycle time, such as 300 seconds, by using the guidelines for synchronous mirroring as discussed in "Inter-site bandwidth calculation" on page 33.
  - If the requirements cannot be fulfilled because of the infrastructure, consider increasing the bandwidth between both sites or increase the cycle time until all modified data can be replicated to AuxFar.
  - Increase bandwidth between both sites or reduce the cycle time if more data must be replicated to AuxFar.

2.6.3 Capacity planning

As part of 3-Site replication configuration, the 3-Site Orchestrator creates dedicated volumes that manage asynchronous data replication among the three sites. These volumes are referred to as *accesspoints*. Although not all accesspoint volumes are intended to be used at the same time, they are used in different ways in different scenarios.

For capacity planning purposes, consider the following points:

- 3-Site Orchestrator creates:
  - All accesspoint volumes; you cannot create them manually.
  - The accesspoint volume as compressed if it is in a data reduction pool (DRP) that contains arrays with compressing drives.
  - The accesspoint volume as thin with auto-expand on if in a DRP containing arrays with regular drives.
  - The accesspoint volume as thin with auto-expand on if in a regular pool.
- 3-Site Orchestrator uses:
  - The same storage pool of a volume to create its accesspoint volume.
  - The same pool as the primary copy of the volume if the volume includes a volume copy.
- It is not possible to change the accesspoint volume settings.

Note: The accesspoint volumes normally uses capacity that is equal to the initially specified rsize of 2%. During periodic cycling, the accesspoint volumes grow as they retain the changed data; therefore, the capacity of the accesspoint volumes depends on how actively your data is overwritten. If your data changes 30%, the accesspoint volumes grow 30%. After cycling, the accesspoint volume shrinks back to their initially specified rsize of 2%.

3

Configuring a 3-Site Replication solution

This chapter provides about the initial configuration of a 3-Site Replication solution. It includes installing and setting up the 3-Site Orchestrator, preparing IBM Storage Virtualize clusters, and creating a 3-Site configuration, which can be configured by using a 3-Site configuration with Metro Mirror or a 3-Site configuration with HyperSwap.

A management GUI is available to configure and manage both 3-Site configuration options. All systems must be running 8.4.0.2 release or higher.

The steps that are provided in this chapter assume that IBM Storage Virtualize systems are configured, hardware is installed, and a basic configuration that uses the Setup Wizard was performed.

This chapter requires the reader to be knowledgeable of IBM Storage Virtualize copy services. For more information about how they operate, see the following publications:

- *Implementation Guide for IBM Storage FlashSystem and IBM SAN Volume Controller: Updated for IBM Storage Virtualize Version 8.6*, SG24-8542
- *Performance and Best Practices Guide for IBM Storage FlashSystem and IBM SAN Volume Controller: Updated for IBM Storage Virtualize Version 8.6,* SG24-8543

This chapter includes the following topics:

3.1 Establishing partnerships for a 3-Site configuration

The management GUI provides an overview about the necessary steps to set up a 3-Site partnership, which is available by selecting **CopyServices** → **Remote Copy** → **Create Partnership**.

In the next menu, select **3-Site partnership** and then select **Create**. On the partnerships page, an overview of the necessary steps is displayed, as shown in Figure 3-1. The sections in this chapter follow those steps.

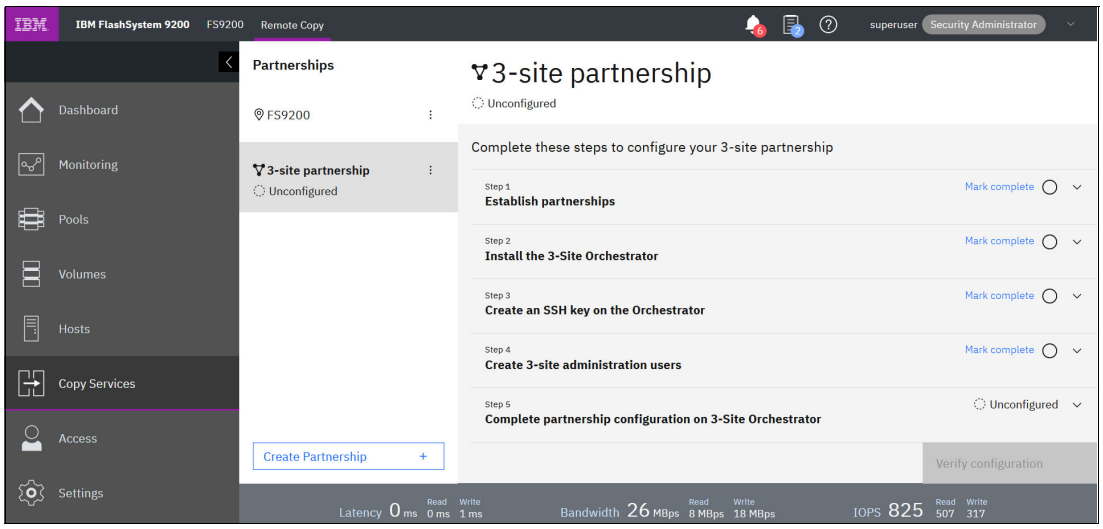

Figure 3-1 Overview about setup steps for a 3-Site partnership

Complete the following steps to correctly set up a 3-Site configuration:

1. Establish partnerships between clusters.
2. Install 3-Site Orchestrator.
3. Create an SSH key on the Orchestrator.
4. Create 3-Site administration users on the clusters.
5. Complete the partnership configuration on the Orchestrator.

Depending on your type of 3-Site Replication configuration, you must create a fully configured partnership between the systems on each site. The number of necessary partnerships depends on whether you use a Metro Mirror or a HyperSwap configuration.

For Metro Mirror, three clusters are part of the 3-Site configuration; therefore, three partnerships must be established. For HyperSwap configurations, only two clusters are part of the configuration; therefore, only a single partnership must be set up. For more information, see Chapter 2, “Planning for 3-Site Replication” on page 23.

Important: IP partnerships are supported for 3-Site replication in a HyperSwap setup. Check the current limits and restrictions of your current firmware.

3.1.1 Preparing Storage Virtualize clusters

Complete the following steps to prepare the Storage Virtualize clusters at Master, AuxNear, and AuxFar sites to work in a 3-Site Replication solution:

1. Verify that all clusters are configured with unique system names. The system name is set during system initialization. It is displayed in the top line of a GUI page, and is shown in a CLI prompt.

   If required, the system name can be changed by using the **`chsystem -name <new_system_name>`** command.

   > **Note:** The iSCSI IQN for each node is generated by using the system and node names. If you use the iSCSI protocol, changing either name also changes the IQN of all of the nodes in the system and might require reconfiguration of all iSCSI-attached hosts.

   System names must remain different through the life of the 3-Site configuration and should never be changed in this release.

2. Check that the Remote Mirroring feature is licensed on all three clusters by browsing to **Settings** → **System** → **Licensed Functions** in the GUI or use the **`lslicense`** command. If needed, correct the licensed capacity (or the number of enclosures) to match it to the purchased licensed capacity.

3.1.2 Establishing partnership for Metro Mirror configurations

As the next step, you need to create replication partnerships between the systems. There must be a partnership that is configured from each system to the other two, so all of them are inter-connected, as illustrated in Figure 3-2.

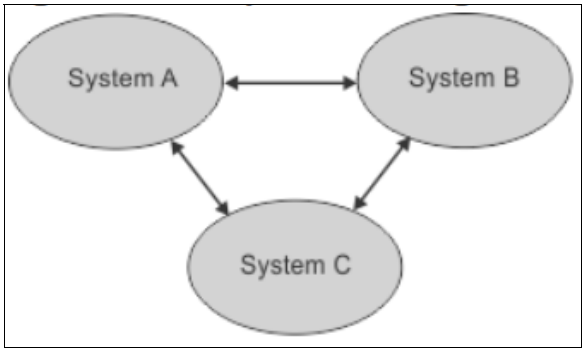

Figure 3-2 Partnership scheme

For this configuration, the following fully configured partnerships must be configured; one partnership between the system at the:

- Master site and the auxiliary-near site
- Auxiliary-near site and the auxiliary-far site
- Master site and the auxiliary-far site

To create partnerships, use the CLI or the GUI. These sections provide only a brief overview of the process. For more information, see the following publications:

- *Implementation Guide for IBM Storage FlashSystem and IBM SAN Volume Controller: Updated for IBM Storage Virtualize Version 8.6*, SG24-8542
- *Performance and Best Practices Guide for IBM Storage FlashSystem and IBM SAN Volume Controller: Updated for IBM Storage Virtualize Version 8.6,* SG24-8543

To configure partnerships between clusters by using the CLI, complete the following steps:

1. Connect to the CLI of system A (see Figure 3-2 on page 39); for example, Master site. Run the **`lspartnershipcandidate`** command to list the clustered systems that are available for setting up a partnership with the local system. Run the **`mkfcpartnership`** command to create partnerships with systems B and C, as shown in Example 3-1. Specify the available bandwidth for replication in megabits per second (Mbps) and the maximum bandwidth percentage that can be used for background copy.

Example 3-1 Listing available partners and creating partnerships

```
IBM_Spectrum_Virtualize:ReplMaster:superuser>lspartnershipcandidate
id                 configured  name
0000001C6EE00020 no          ReplAuxNear
0000001C65200018 no          ReplAuxFar

IBM_Spectrum_Virtualize:ReplMaster:superuser>
mkfcpartnership -linkbandwidthmbits 32000 -backgroundcopyrate 50 ReplAuxNear

IBM_Spectrum_Virtualize:ReplMaster:superuser>
mkfcpartnership -linkbandwidthmbits 32000 -backgroundcopyrate 50 ReplAuxFar

IBM_Spectrum_Virtualize:ReplMaster:superuser>lspartnership
id                 name        location partnership                type
0000001C61200002 ReplMaster  local
0000001C6EE00020 ReplAuxNear remote   partially_configured_local fc
0000001C65200018 ReplAuxFar  remote   partially_configured_local fc
```

2. The partnership is created in `partially_configured_local` state. To get it to `fully_configured`, **`mkfcpartnerhsip`** must be run on system B and system C (ReplAuxNear and ReplAuxFar) against a cluster ID or the name of system A, as shown in Example 3-2.

Example 3-2 Making a partnership fully_configured

```
IBM_Spectrum_Virtualize:ReplAuxNear:superuser>lspartnershipcandidate
id                 configured  name
0000001C61200002 yes         ReplMaster
0000001C65200018 no          ReplAuxFar

IBM_Spectrum_Virtualize:ReplAuxNear:superuser>
mkfcpartnership -linkbandwidthmbits 32000 -backgroundcopyrate 50 ReplMaster

IBM_Spectrum_Virtualize:ReplAuxNear:superuser>lspartnership
id                 name        location partnership      type
0000001C6EE00020 ReplAuxNear local
0000001C61200002 ReplMaster  remote   fully_configured fc
```

3. Create a partnership between system B and system C in the same way by running **`mkfcpartnership`** on both systems.

4. Verify that **`lspartnership`** on each system shows two remote clusters with a partnership in `fully_configured` state.

Complete the following steps to configure partnerships by using the GUI:

1. Connect to the GUI of system A (see Figure 3-3); for example, Master site and then browse to **Copy Services** → **Remote Copy** and click **Create Partnership**. After you select the FC partnership type, the dialog that is shown in Figure 3-3 appears. In the drop-down menu, select the partner system (for example, on a AuxNear site), specify the link bandwidth and background copy rate, and click **Create**.

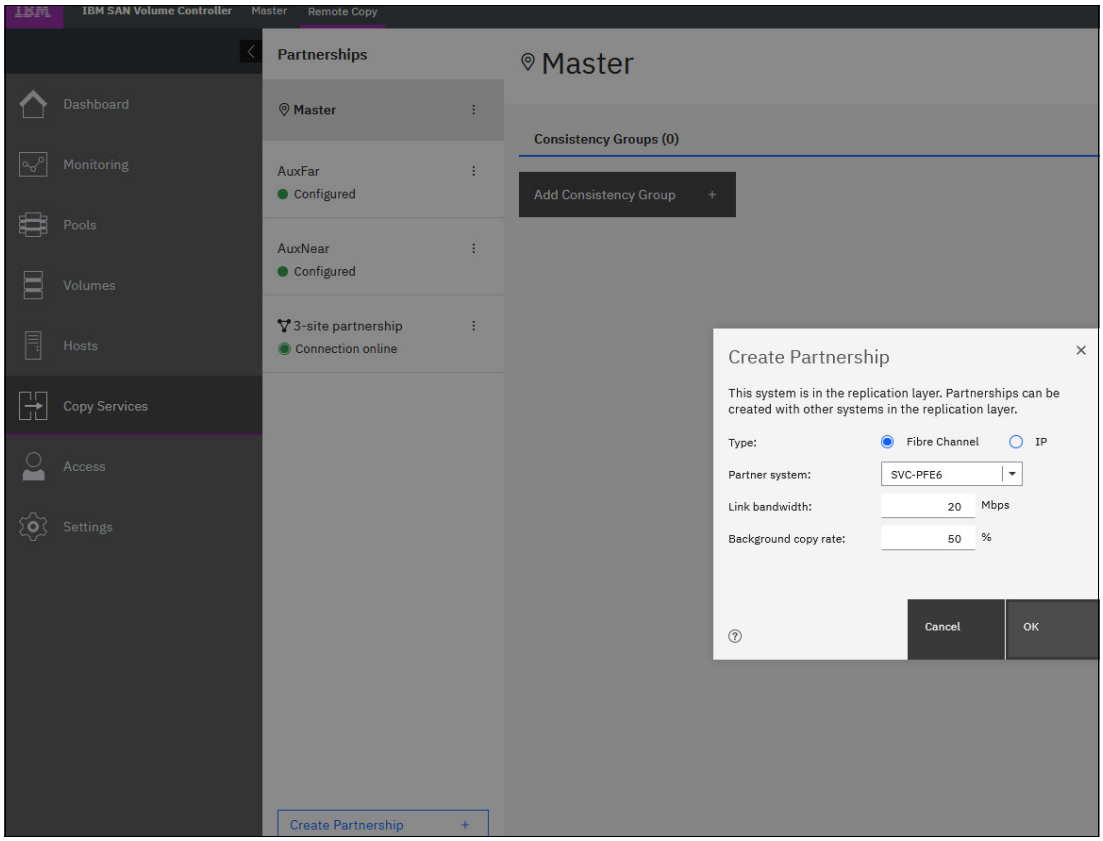

Figure 3-3 Create Partnership dialog

2. Repeat the previous step to create a partnership with the AuxFar system. The GUI shows two partnerships in `partially_configured_local` state.
3. To get the partnerships to a `fully_configured` state, go to the GUI of each Aux system and complete the same steps to create a partnership with the Master cluster.
4. Create the partnership between system B and system C (AuxNear and AuxFar).

As the result, each pair of sites is in a `fully_configured` partnership.

3.1.3 Establishing partnership for HyperSwap configurations

For this configuration, set up only one fully configured partnership between the HyperSwap cluster,(which includes systems on Master and AuxNear site, and the system on the AuxFar site (see Figure 3-4). To be fully configured, a partnership must be configured from both systems.

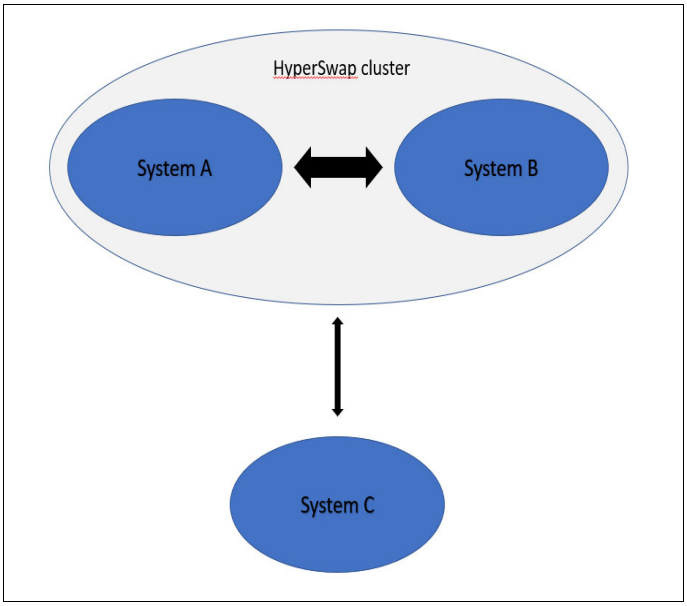

Figure 3-4 Partnership scheme for HyperSwap configuration

To create partnerships, use the CLI or the GUI. The next sections provide only a brief overview of the process. For more information, see the following resources:

- *Implementation Guide for IBM Storage FlashSystem and IBM SAN Volume Controller: Updated for IBM Storage Virtualize Version 8.6*, SG24-8542
- *Performance and Best Practices Guide for IBM Storage FlashSystem and IBM SAN Volume Controller: Updated for IBM Storage Virtualize Version 8.6,* SG24-8543

Complete the following steps to configure partnerships between clusters by using the CLI:

1. Connect to the CLI of the HyperSwap cluster (see Figure 3-4); for example, Master site. Run the **`lspartnershipcandidate`** command to list the clustered systems that are available for setting up a partnership with the local system. Run the **`mkfcpartnership`** command to create a partnership with system C, as shown in Example 3-3 on page 43. Specify the available bandwidth for replication in Mbps and the maximum bandwidth percentage that can be used for background copy.

Example 3-3 Listing available partners and creating partnerships

```
IBM_Spectrum_Virtualize:ReplMaster:superuser>lspartnershipcandidate
id               configured  name
0000001C65200018 no          ReplAuxFar

IBM_Spectrum_Virtualize:ReplMaster:superuser>
mkfcpartnership -linkbandwidthmbits 32000 -backgroundcopyrate 50 ReplAuxFar

IBM_Spectrum_Virtualize:ReplMaster:superuser>lspartnership
id               name        location partnership                type
0000001C61200002 ReplMaster  local
0000001C65200018 ReplAuxFar  remote   partially_configured_local fc
```

2. The partnership is created in `partially_configured_local` state. To get it to `fully_configured`, **`mkfcpartnerhsip`** must be run on system C (ReplAuxFar) against a cluster ID or name of system A, as shown in Example 3-4.

Example 3-4 Making a partnership fully_configured

```
IBM_Spectrum_Virtualize:ReplAuxNear:superuser>lspartnershipcandidate
id               configured  name
0000001C61200002 yes         ReplMaster
0000001C65200018 no          ReplAuxFar

IBM_Spectrum_Virtualize:ReplAuxNear:superuser>
mkfcpartnership -linkbandwidthmbits 32000 -backgroundcopyrate 50 ReplMaster

IBM_Spectrum_Virtualize:ReplAuxNear:superuser>lspartnership
id               name        location partnership      type
0000001C6EE00020 ReplAuxNear local
0000001C61200002 ReplMaster  remote   fully_configured fc
```

3. Verify that **`lspartnership`** on each system shows a partnership in `fully_configured` state.

Complete the following steps to configure partnerships by using the GUI:

1. Connect to the GUI of the HyperSwap cluster (see Figure 3-4 on page 42); for example, Master site. Then, browse to **Copy Services** → **Copy Services** and click **Create Partnership**.

   After you select the FC partnership type, the dialog that is shown in Figure 3-5 on page 44 appears. In the drop-down menu, select the partner system (for example, on a AuxNear site), specify the link bandwidth and background copy rate, and click **Create**.

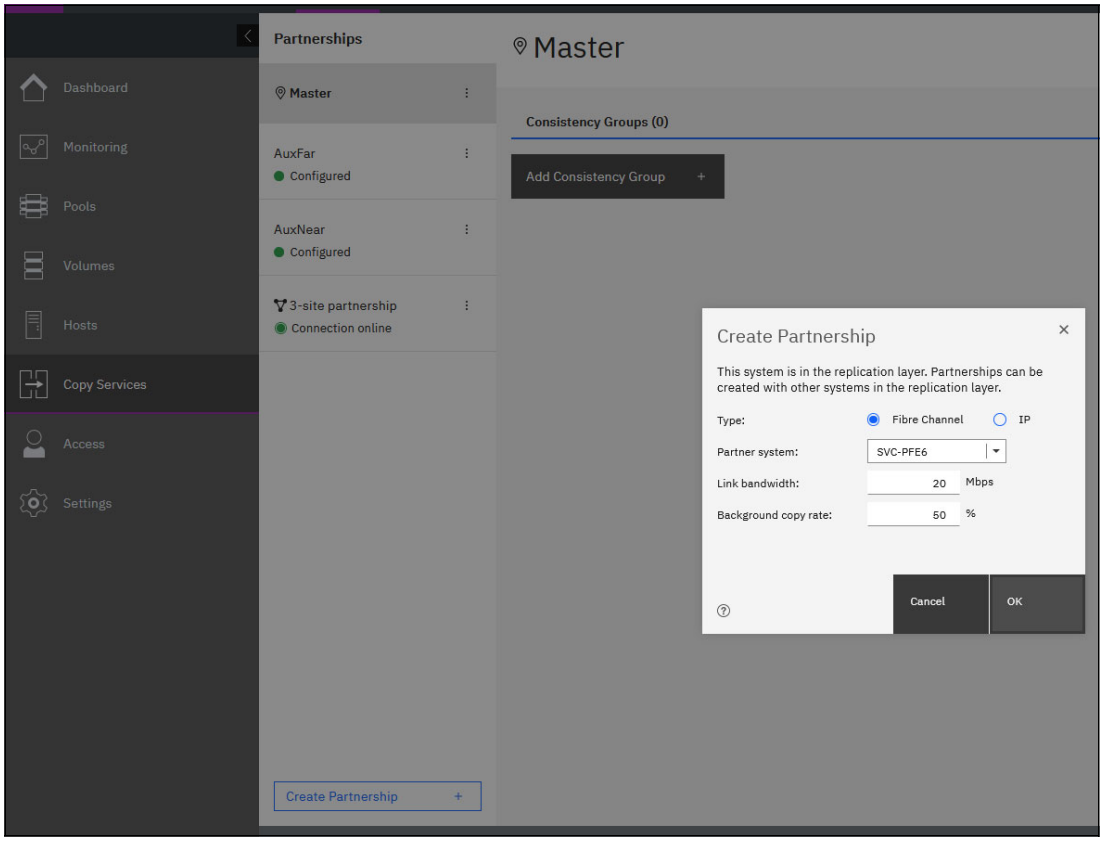

Figure 3-5 Create Partnership dialog

2. To get the partnership to a `fully_configured` state, go to the GUI of the AuxFar system and perform the same steps to create a partnership with the Master cluster.

3.2 Preparing the 3-Site Orchestrator host

Completing the steps that are described in this section prepares a Linux host for operation as a 3-Site Orchestrator.

- Certain versions of the Orchestrator software exist; that is, some support only one of the configuration options (Metro Mirror or HyperSwap). Choose a suitable version according to the supported functions that are listed in Table 3-1. For more information, see IBM Storage Virtualize 3-Site Orchestrator Tool.

Table 3-1 Supported versions of the installation packages

Version of installation package	Supported 3-Site configurations	Supported management interfaces	Supported IBM Storage Virtualize releases
Orchestrator 1.0	Metro Mirror only	CLI only	8.3.1 8.4.0
Orchestrator 2.0	HyperSwap only	CLI only	8.4.0

Version of installation package	Supported 3-Site configurations	Supported management interfaces	Supported IBM Storage Virtualize releases
Orchestrator 3.0	Metro Mirror and HyperSwap configurations	GUI and CLI	8.4.0.2 or later
Orchestrator 3.1	Metro Mirror and HyperSwap configurations	GUI and CLI	8.4.0.2 or later
Orchestrator 4.0	Metro Mirror and HyperSwap configurations	GUI and CLI	8.4.0.2 or later
Orchestrator 4.1	Metro Mirror and HyperSwap configurations	GUI and CLI	8.4.0.2 or later

Complete the following steps:

1. Install Orchestrator software:
   a. To obtain the 3-Site Orchestrator `.rpm` file from IBM, see IBM Support: IBM Storage Virtualize 3-Site Orchestrator Tool. Upload it to your Linux host.
   b. Install it by using the **`rpm`** command, as shown in Example 3-5.

Example 3-5 Installing orchestrator

```
[root@localhost ~]# rpm -i ibm-SpectrumVirtualize-rc3site-orchestrator-1...rpm
[root@localhost ~]# rpm -qa | grep 3site
ibm-SpectrumVirtualize-rc3site-orchestrator-1...x86_64
```

3-Site Orchestrator components are installed to the `/opt/ibm/SpectrumVirtualize/rc3site/` directory. Also, the following services are installed:

- `/etc/systemd/system/rc3site.eventd.service`: Communicates with all three sites and monitors important attributes of the sites and configured consistency groups.
- `/etc/systemd/system/rc3site.tsmiscservicesd.service`: Maintains an audit log of all 3-Site Orchestrator commands.
- `/etc/systemd/system/rc3site.tstaskd.service`: Controls replication on configured 3-Site Orchestrator consistency groups.

The host administrator can use the **`systemctl`** command to ensure that the service is running, as shown in Example 3-6.

Example 3-6 TaskDaemon status check

```
[root@localhost ~]# systemctl status rc3site.tstaskd.service
? rc3site.tstaskd.service - rc3site tstaskd service. Orchestrator VERSION =
3.0.201209113224
   Loaded: loaded (/etc/systemd/system/rc3site.tstaskd.service; enabled; vendor preset:
disabled)
   Active: active (running) since Mon 2021-03-15 11:43:57 CET; 2 weeks 3 days ago
  Process: 2942 ExecStart=/usr/bin/tstaskd start (code=exited, status=0/SUCCESS)
 Main PID: 2957 (tstaskd)
   CGroup: /system.slice/rc3site.tstaskd.service
           ??  513 ssh -fN -L 8000:localhost:50100 -i /root/.ssh/3sitehsadmin...
```

```
           ??  515 ssh -fN -R 8000:localhost:22 -i /root/.ssh/3sitehsadmin -p...
           ?? 2957 /usr/bin/tstaskd start
           ??32238 ssh -fN -R 8001:localhost:22 -i /root/.ssh/3sitehsadmin -p...
           ??32239 ssh -fN -L 8001:localhost:50100 -i /root/.ssh/3sitehsadmin...

Mar 15 11:43:57 localhost.localdomain systemd[1]: Starting rc3site tstaskd se...
Mar 15 11:43:57 localhost.localdomain systemd[1]: Started rc3site tstaskd ser...
Hint: Some lines were ellipsized, use -l to show in full.

[root@localhost ~]# systemctl status rc3site.tsmiscservicesd.service
? rc3site.tsmiscservicesd.service - rc3site tsmiscservicesd service. Orchestrator
VERSION = 3.0.201209113224
   Loaded: loaded (/etc/systemd/system/rc3site.tsmiscservicesd.service; enabled; vendor
preset: disabled)
   Active: active (running) since Mon 2021-03-15 11:43:57 CET; 2 weeks 3 days ago
  Process: 2934 ExecStop=/usr/bin/tsmiscservicesd stop (code=exited, status=0/SUCCESS)
  Process: 2941 ExecStart=/usr/bin/tsmiscservicesd start (code=exited, status=0/SUCCESS)
 Main PID: 2946 (tsmiscservicesd)
   CGroup: /system.slice/rc3site.tsmiscservicesd.service
           ??2946 /usr/bin/tsmiscservicesd start

Mar 15 11:43:57 localhost.localdomain systemd[1]: Starting rc3site tsmiscserv...
Mar 15 11:43:57 localhost.localdomain systemd[1]: Started rc3site tsmiscservi...
Hint: Some lines were ellipsized, use -l to show in full.

[root@localhost ~]# systemctl status rc3site.eventd.service
? rc3site.eventd.service - rc3site eventd service. Orchestrator VERSION =
3.0.201209113224
   Loaded: loaded (/etc/systemd/system/rc3site.eventd.service; enabled; vendor preset:
disabled)
   Active: active (running) since Mon 2021-03-15 11:43:57 CET; 2 weeks 3 days ago
 Main PID: 2940 (eventd)
   CGroup: /system.slice/rc3site.eventd.service
           ??2940 /usr/bin/eventd 7003

Mar 15 11:43:57 localhost.localdomain systemd[1]: Started rc3site eventd serv...
Hint: Some lines were ellipsized, use -l to show in full.
```

After 3-Site Orchestrator is installed, 3-Site configuration commands become available.

2. Prepare passwordless communication with Storage Virtualize systems. This step is needed to verify that the Orchestrator can log in to each system without a password prompt:

   a. Verify that SSH communication on port 22 is possible from your 3-Site Orchestrator Linux host to the Storage Virtualize systems on all three replication sites.

   b. Verify that RSA signatures for the cluster IP addresses of all three Storage Virtualize systems are present in `/root/.ssh/known_hosts`, or add them there, as shown in Example 3-7.

   Example 3-7 Adding signatures to known_hosts

```
[root@localhost ~]# ssh-keyscan <ReplMasterIP> >> /root/.ssh/known_hosts
[root@localhost ~]# ssh-keyscan <ReplAuxNearIP> >> /root/.ssh/known_hosts
[root@localhost ~]# ssh-keyscan <ReplAuxFarIP>  >> /root/.ssh/known_hosts
```

   c. Generate the SSH key pair for password less communication by using the **ssh-keygen** command, as shown in Example 3-8 on page 47. If no parameters are provided, **ssh-keygen** generates 2048-bit RSA keys. Do *not* specify a passphrase.

Example 3-8 Generating key pair for password less communication

```
[root@localhost ~]# ssh-keygen
Generating public/private rsa key pair.
Enter file in which to save the key (/root/.ssh/id_rsa): /root/.ssh/3site
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /root/.ssh/3site.
Your public key has been saved in /root/.ssh/3site.pub.
The key fingerprint is:
SHA256:765KxaRLJ3SYQQIcbLqapsuFrpm6xv2kSmqyUuqKXDO root@localhost
The key's randomart image is:
+---[RSA 2048]----+
...
+----[SHA256]-----+
```

d. To ensure that the management GUI can access 3-Site Orchestrator objects, add the public SSH key to the authorized keys by entering the following command on the 3-Site Orchestrator host: `ssh-copy-id -i ~/.ssh/<id_rsa.pub> root@<ORCH-HOST>`.

e. Copy the public SSH key to the `/tmp` directory on each Storage Virtualize system with scp, as shown in Example 3-9. If you plan to use the GUI to set up a user ID on Storage Virtualize systems, download keys from the Orchestrator host to your PC.

   More steps that are needed to configure passwordless authentication are performed from Storage Virtualize cluster GUI or CLI and are described next.

Example 3-9 Copying public key

```
[root@localhost ~]# scp /root/.ssh/3site.pub superuser@<ReplMasterIP>:/tmp/
Password:
3site.pub                                     100%  392     1.3MB/s   00:00
[root@localhost ~]# scp /root/.ssh/3site.pub superuser@<ReplAuxNearIP>:/tmp/
Password:
3site.pub                                     100%  392     1.3MB/s   00:00
[root@localhost ~]# scp /root/.ssh/3site.pub superuser@<ReplAuxFarIP>:/tmp/
Password:
3site.pub                                     100%  392     1.3MB/s   00:00
```

3.3 Creating 3-Site administration users

This section includes how to configure a user with a 3-Site Administrator role on each of the Storage Virtualize systems. It is used by 3-Site Orchestrator to track and control replication processes.

The 3-Site administrator user must belong to the 3-Site Administrator user group, which does not exist on the system by default, and must be created.

To configure a 3-Site Administrator user by using the CLI, complete the following steps:

1. Create a user group with a descriptive name and `3SiteAdmin` role by using the **`mkusergrp`** command, as shown in Example 3-10.

Example 3-10 Creating a user group

```
IBM_Spectrum_Virtualize:ReplMaster:superuser>mkusergrp -name 3SiteAdminGroup -role
3SiteAdmin
User Group, id [6], successfully created
```

2. Create a local user record for 3-Site Orchestrator, and specify the SSH key file that was copied to the system as described in step d on page 47. When you run the command, the SSH key is copied into system state and activated for the user, and the input file is deleted. Example 3-11 shows how the **`mkuser`** command is used.

Example 3-11 Creating 3site orchestrator user

```
IBM_Spectrum_Virtualize:ReplMaster:superuser>mkuser -name 3SiteOrch -usergrp
3SiteAdminGroup -keyfile /tmp/3site.pub
User, id [1], successfully created
```

3. Repeat steps 1 and 2 on other clusters. The user name that is specified on all clusters must match and use the same SSH public key.
4. If you prefer configuring users with the GUI, complete the following steps:
   i. Browse to **Access** → **Users by group** and click **Create User Group**, as shown in Figure 3-6.

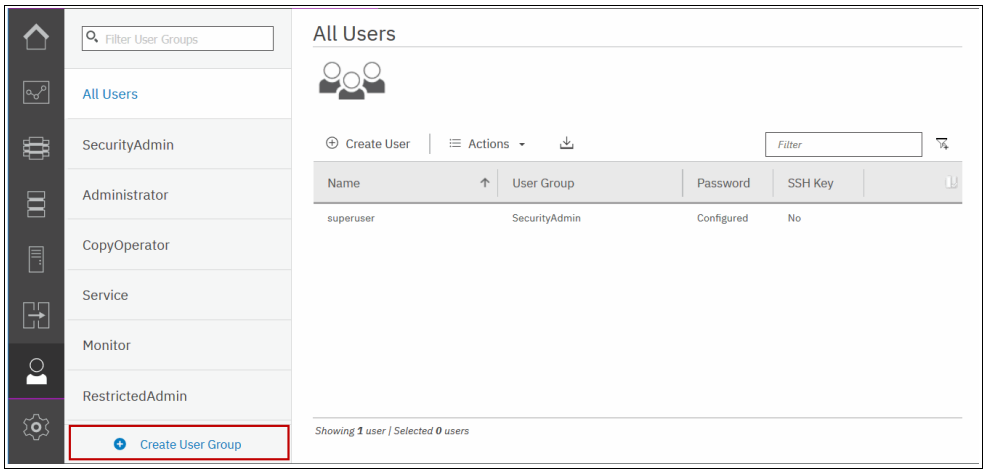

Figure 3-6 Create user group button

ii. In a Create User Group dialog, specify the group's name and assign the 3-Site Administrator role, as shown in Figure 3-7.

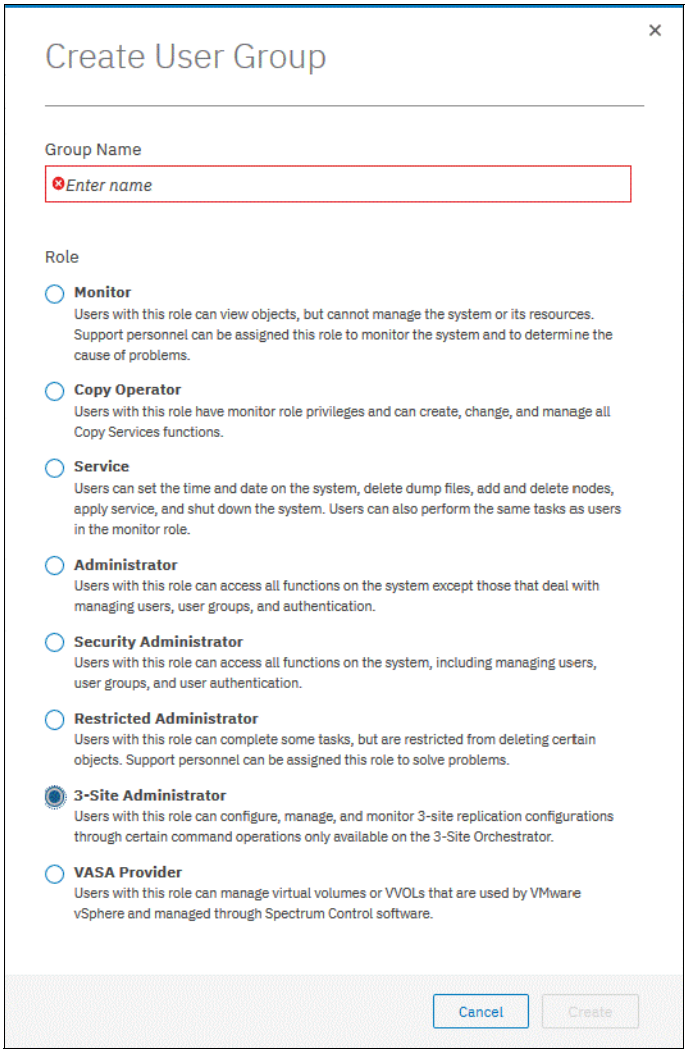

Figure 3-7 Create User Group dialog

iii. After the group is created, click **Create User** to see the dialog that is shown in Figure 3-8. Set the user name that 3-Site Orchestrator uses, and ensure that the User Group is set to one that was created with the previous step. Click **Browse** to select and upload the SSH public key that was created on a 3-Site Orchestrator system as described in c on page 46.

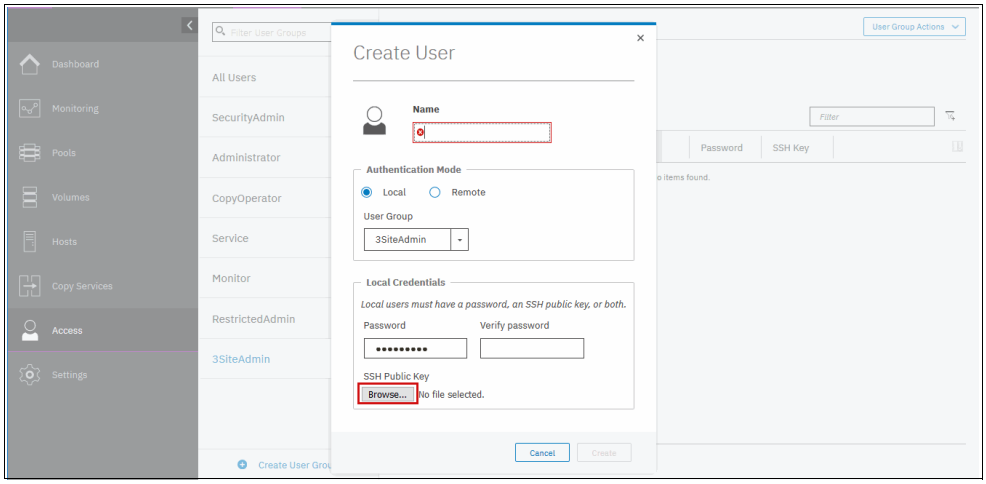

Figure 3-8 Create user dialog

iv. Repeat steps i on page 48, - iii on the other clusters. The user name that is specified on all three clusters must match and use the same SSH public key.

3.4 Completing the configuration on 3-Site Orchestrator

After creating the 3-Site administrator, users on each system, and the list of know hosts on the 3-Site Orchestrator, you must finish the configuration process before you can configure a 3-Site Replication.

Complete the following steps on the 3-Site Orchestrator:

1. Run the **`mk3siteconfig`** command to set up the configuration, as shown in Example 3-12 and Example on page 51. Include the parameters that are listed in Table 3-2.

Example 3-12 mk3siteconfig command syntax

```
>> mk3siteconfig -ip1 IP1 [-ip2 IP2] -username USERNAME
-keyfile KEYFILE -port PORT [-sitealias SITEALIAS] [-topology TOPOLOGY]
```

Table 3-2 Parameters used for mk3siteconfig

Parameter	Required / optional	Description
ip1	required	Comma-separated list of Storage Virtualize cluster IPv4 addresses for the sites. List is in master, auxnear, auxfar order (for Metro Mirror) or HyperSwap cluster, and auxfar for HyperSwap configurations.
ip2	optional	Comma-separated list of extra Storage Virtualize cluster IPv4 addresses for the sites. List is in master, auxnear, auxfar order (for Metro Mirror) or HyperSwap cluster, and auxfar for HyperSwap configurations.

Parameter	Required / optional	Description
username	required	A 3-Site administrator user name that is defined on Storage Virtualize clusters on all three sites.
keyfile	required	Path for private SSH key file for 3-Site administrator user.
port	required	Starting address for three local ports to be used for connection with three sites over ssh tunnel. Orchestrator uses three consecutive ports, starting with the specified address.
sitealias	optional	A colon-separated list of site alias for three sites. List must be in master, auxnear, or auxfar order.
topology	optional	Specify topology of the 3-Site configuration.

Note: At the time of the writing, the site aliases must match the cluster names. This limitation of the current GUI implementation is scheduled to be changed in a future release.

The **mk3siteconfig** command for a HyperSwap configuration is shown in Example 3-13.

Example 3-13 mk3siteconfig command for a HyperSwap configuration

```
[root@linuxsrv1 ~]# mk3siteconfig -ip1 10.10.10.125,10.10.10.198
-keyfile '/root/.ssh/id_rsa' -username '3SiteOrch'  -port 6000 -sitealias
master,auxfar -topology hyperswap
```

The **mk3siteconfig** command for a Metro Mirror configuration is shown in Example 3-14.

Example 3-14 mk3siteconfig command for a Metro Mirror configuration

```
[root@linuxsrv1 ~]# mk3siteconfig -ip1 10.10.10.125,10.10.10.11,10.10.10.198
-keyfile '/root/.ssh/id_rsa' -username '3SiteOrch'  -port 6000 -sitealias
master,auxnear,auxfar
```

2. Verify the configuration by using the **ls3siteconfig** command, as shown in Example 3-15.

Example 3-15 Verify the configuration with command ls3siteconfig

```
[root@localhost ~]# ls3siteconfig
master_name Master
master_site_name
master_site_status online
master_ip1 10.10.10.125
master_ip1_status reachable
master_ip2
master_ip2_status not_configured
master_port 6000
auxnear_name AuxNear
auxnear_site_name
auxnear_site_status online
auxnear_ip1 10.10.10.11
auxnear_ip1_status reachable
```

```
auxnear_ip2
auxnear_ip2_status not_configured
auxnear_port 6001
auxfar_name AuxFar
auxfar_site_status online
auxfar_ip1 10.10.10.198
auxfar_ip1_status reachable
auxfar_ip2
auxfar_ip2_status not_configured
auxfar_port 6002
username 3SiteOrch
keyfile /root/.ssh/id_rsa
topology metromirror
[root@linuxsrv1 ~]# [root@linuxsrv1 ~]#
```

The 3-Site Orchestrator configuration is saved on an Orchestrator Linux host in an XML file at ‘/opt/ibm/SpectrumVirtualize/rc3site/config.xml‘.

When a 3-Site configuration is created by 3-Site Orchestrator, no event exists that indicates it in the Storage Virtualize cluster event log or audit log.

The management GUI provides a verification process for the correct setup; you must return to the step-by-step 3-Site partnership page that is shown in Figure 3-1 on page 38. Check that all steps as complete and click **Verify configuration**, in Figure 3-9.

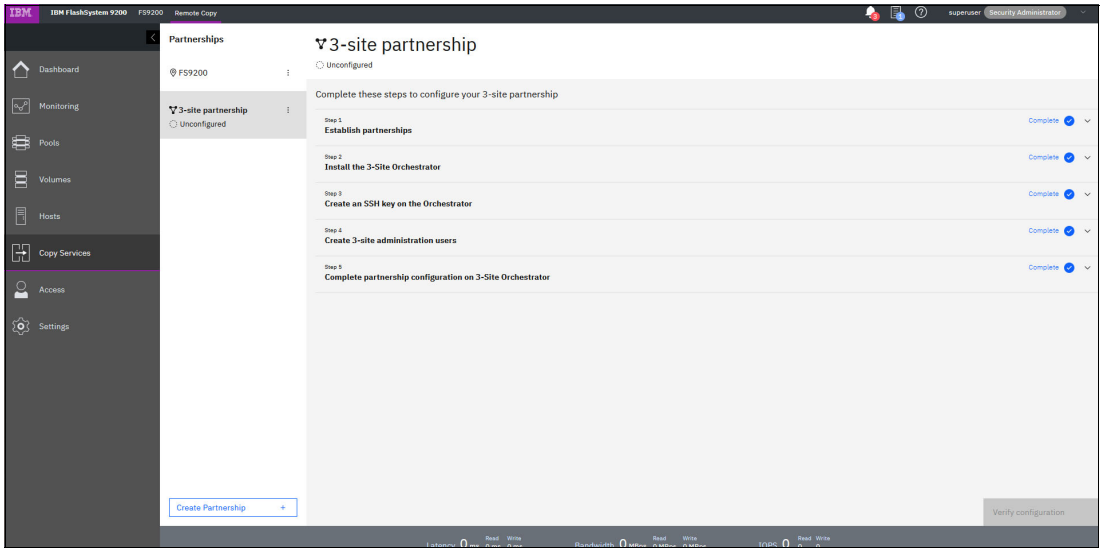

Figure 3-9 Verify configuration after finishing all steps

After the 3-Site configuration is verified, you can convert a Metro Mirror or a HyperSwap consistency group to a 3-Site consistency group.

3.5 Configuring and using the 3-Site standby Orchestrator

The 3-Site standby Orchestrator is used to automatically detect any failure between two Orchestrators by establishing connectivity. The active Orchestrator is the existing instance, and the new instance is the standby Orchestrator. The active Orchestrator is responsible for running 3-Site data cycling and monitoring the status of the standby Orchestrator. The

standby Orchestrator acts as a backup if there is detection of the active Orchestrator's failure. If there is no response from the active Orchestrator after approximately 5 minutes, the standby Orchestrator takes over the role of the active Orchestrator and performs 3-Site data cycling.

You can configure and manage the standby Orchestrator through the command-line interface (CLI).

This function can be used to ensure continuity of 3-Site replication during:

- Orchestrator maintenance and disasters
- Orchestrator to storage link failures
- Orchestrator migration

Configuring a 3-Site standby Orchestrator solution

This topic focuses on the initial setup of a 3-Site standby orchestrator solution, covering the installation and configuration of the 3-Site standby Orchestrator.

This can be achieved by using either a 3-Site configuration with Metro Mirror or a 3-Site configuration with HyperSwap.

Preparing the 3-Site Orchestrator host

Perform the following steps for preparing the 3-Site Orchestrator host:

1. For configuring standby orchestrator, install the same version of the Orchestrator (v4.1 or later) on both active and standby Orchestrators.
2. Verify that the active and standby Orchestrator have access to all 3-Sites.
3. Create 3-Site Admin users on the target sites.
4. Establish a passwordless SSH connection between the host and target site by using the same SSH keys (private and public keys). The path to the SSH key files must be the same on both active and standby Orchestrator.
5. Enable the required ports for communication on both active and standby Orchestrator. A 3-Site configuration must exist on the Orchestrator

Completing the configuration of 3-Site standby Orchestrator

Using the command-line interface (CLI), the active and standby roles can be assigned to Orchestrator instances.

To configure standby Orchestrator, you must update the installation version to Orchestrator 4.1.

Note: Configuration and management commands in standby Orchestrator require root privileges, and this feature is available for both the Metro Mirror and the HyperSwap configurations.

Configuring active and standby Orchestrators by using the command-line interface (CLI)

Notes:

- By default, after installing 3-Site orchestrator HA service will not be started.
- After 3-Site configuration is completed using **mk3siteconfig** command, Admin may run **ch3siteconfig** command with option **-startha**. This starts the HA service and sets the role of orchestrator as an active orchestrator.
- 3-Site configuration must be present to set up the HA functionality.

Perform the following steps:

1. Log in to the Orchestrator that already has a 3-Site configuration to configure the active Orchestrator.
2. Enter the **ls3siteconfig c**ommand to validate if the `Orchestrator_ha_status` shows value. See Example 3-16.

Example 3-16 ls3siteconfig command

```
[root@localhost]# ls3siteconfig
master_name Master
master_site_name
master_site_status online
master_ip1 x.xxx.xxx.xx
master_ip1_status reachable
master_ip2
master_ip2_status not_configured
master_port 5000
auxnear_name Aux_Near
auxnear_site_name
auxnear_site_status online
auxnear_ip1 x.xxx.xxx.xx
auxnear_ip1_status reachable
auxnear_ip2
auxnear_ip2_status not_configured
auxnear_port 5001
auxfar_name Aux_Far
auxfar_site_status online
auxfar_ip1 x.xxx.xxx.xx
auxfar_ip1_status reachable
auxfar_ip2
auxfar_ip2_status not_configured
auxfar_port 5002
username testuser
keyfile /root/.ssh/id_rsa
topology metromirror
orchestrator_version 4.1.221116141559
orchestrator_ha_status
```

3. To start the active Orchestrator, enter the following command.

   ```
   [root@localhost home]# ch3siteconfig  -startha
   ```

   The optional parameter `-startha` sets the role of Orchestrator as an active Orchestrator.

Note: The **`ch3siteconfig`** command generates a .tar file at `/opt/ibm/SpectrumVirtualize/rc3site/config.tar`.

4. Copy the `config.tar` file from active Orchestrator to standby Orchestrator to configure the standby Orchestrator.

   ```
   [root@localhost home]# scp  -r /opt/ibm/SpectrumVirtualize/rc3site/config.tar
   root@ x.xxx.xxx.xx:/home/
   password:
   config.tar 100% 10KB 9.5MB/s 00:00
   ```

5. Switch to standby Orchestrator and enter the following command to configure standby Orchestrator.

   ```
   [root@rhel7host108 home]# ch3siteconfig  -startha  -standby  -config
   /home/config.tar
   ```

6. Verify the Orchestrator status by entering the **`ls3siteconfig`** command on the active Orchestrator. Your output should be similar to the following, as shown in Example 3-17.

Example 3-17 ls3siteconfig command

```
[root@localhost]# ls3siteconfig
master_name Master
master_site_name
master_site_status online
master_ip1 x.xxx.xxx.xx
master_ip1_status reachable
master_ip2
master_ip2_status not_configured
master_port 5000
auxnear_name Aux_Near
auxnear_site_name
auxnear_site_status online
auxnear_ip1 x.xxx.xxx.xx
auxnear_ip1_status reachable
auxnear_ip2
auxnear_ip2_status not_configured
auxnear_port 5001
auxfar_name Aux_Far
auxfar_site_status online
auxfar_ip1 x.xxx.xxx.xx
auxfar_ip1_status reachable
auxfar_ip2
auxfar_ip2_status not_configured
auxfar_port 5002
username testuser
keyfile /root/.ssh/id_rsa
topology metromirror
orchestrator_version 4.1.221116141559
orchestrator_ha_status online
```

4

Creating and managing 3-Site Replication relationships

This chapter describes the process to start replicating data to the third site. It is assumed that all preparation steps that are described in Chapter 3, “Configuring a 3-Site Replication solution” on page 37 are complete.

Instructions are provided to configure 3-Site consistency groups and 3-Site replication relationships in Metro Mirror and HyperSwap configurations.

This chapter also contains instructions about converting 3-Site replication relationships back to 2-Site configuration.

This chapter includes the following topics:

4.1 Replication configuration workflow

To start replicating data over three sites in a configuration where volumes on Master and AuxNear sites are replicated with Metro Mirror, the following tasks must be completed:

- 3-Site configuration is created, as described in Chapter 3, “Configuring a 3-Site Replication solution” on page 37.
- On Master and AuxNear sites, Metro Mirror replication relationships are created, and Master and AuxNear volumes must be synchronized (therefore, replication is running in 2-Site mode).
- Relationships are added to a consistency group. A stand-alone relationship cannot be converted to 3-Site mode.
- By using the Storage Virtualize GUI of the system on Master or AuxNear sites, a consistency group was converted from 2-Site to 3-Site mode.
- After the consistency group is converted, 3-Site Orchestrator started data cycling to AuxFar site automatically.

If a 3-Site consistency group exists, you can add stand-alone 2-Site replication relationships to it. By doing so, the relationship is converted to 3-Site mode.

In a HyperSwap configuration, the following tasks must be completed to start 3-Site Replication:

- 3-Site configuration is created, as described in Chapter 3, “Configuring a 3-Site Replication solution” on page 37.
- If you are creating HyperSwap volumes, they can be created with the third copy from the start. In the GUI, the Create Volume dialog suggests that you can create a 3-Site consistency group or add a volume to a 3-Site consistency group. By following this dialog, a new HyperSwap volume is created with a third copy on the AuxFar site.
- If you are adding a third copy to HyperSwap volumes, you must create a 2-Site consistency group where you add the volumes then you convert the 2-Site consistency group to a 3-Site.
- After the consistency group is converted, the 3-Site Orchestrator starts data cycling to the AuxFar site automatically.

You can also add HyperSwap volumes to the 3-Site consistency group.

In both types of configuration, 3-Site with Metro Mirror and 3-Site with HyperSwap, you can convert 3-Site consistency group back to 2-Site Metro Mirror consistency group or to a consistency group of HyperSwap volumes with copies on two sites.

Important: Do not delete any volumes or 2-Site remote copy (RC) relationships by using the management GUI or CLI if they participate in 3-Site configuration. If a volume must be deleted, remove it from the 3-Site configuration first.

4.2 Creating 3-Site relationships in a Metro Mirror configuration

Before a replication can become 3-Site, it must be created as a 2-Site Metro Mirror replication. This section provides a brief overview of the process. For more information, see the following publications:

- *Implementation Guide for IBM Storage FlashSystem and IBM SAN Volume Controller: Updated for IBM Storage Virtualize Version 8.6*, SG24-8542
- *Performance and Best Practices Guide for IBM Storage FlashSystem and IBM SAN Volume Controller: Updated for IBM Storage Virtualize Version 8.6,* SG24-8543

If a 2-Site Metro Mirror solution is running and you must add a third site into the configuration, see 4.2.2, “Converting 2-Site Metro Mirror relationships to 3-Site” on page 62.

4.2.1 Creating 2-Site Metro Mirror relationships

Complete the following steps to create a 2-Site Metro Mirror replication:

1. On the AuxNear system, create a volume that contains a replica of the Master volume. The provisioned (available to host) capacity of a volume must be the same as the provisioned capacity of the Master volume. The volume type can be different. For example, the Master volume can be Fully Allocated, though the AuxNear volume is compressed.
2. On the Master system, create a Metro Mirror consistency group and relationships:
   a. In the GUI, browse to **Copy Services** → **Remote Copy** and click **Add Consistency Group**, as shown in Figure 4-1. It can be done on any configured partnership, as the system that contains replication target volumes is selected in the next step.

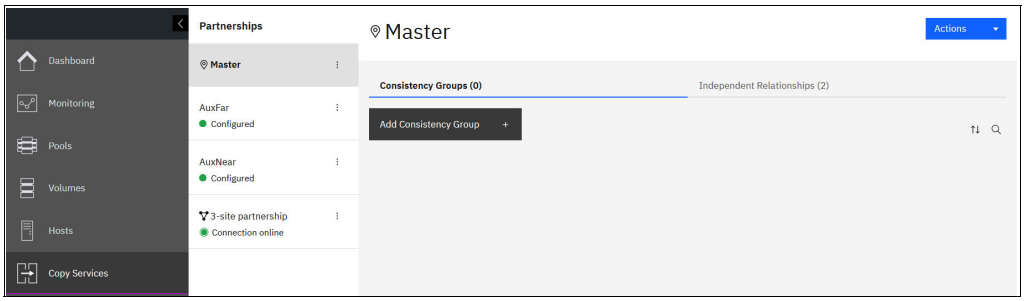

Figure 4-1 Creating a consistency group

   b. Specify a consistency group name, select the AuxNear system that stores auxiliary volumes, as shown in Figure 4-2 on page 60. Click **Add** to create an empty consistency group.

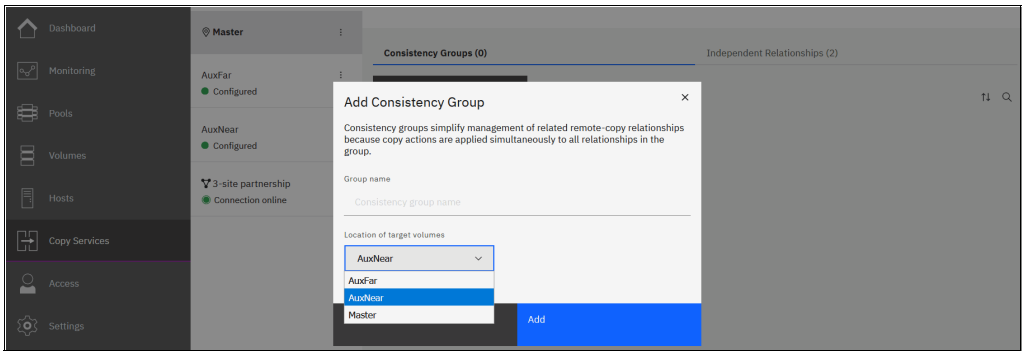

Figure 4-2 Auxiliary volumes location

c. Click the system that is the Metro Mirror auxiliary in the **Partnerships** column to find the created consistency group. You see a list of 2-Site consistency groups that include replication relationships between Master and AuxNear sites, as shown in Figure 4-3.

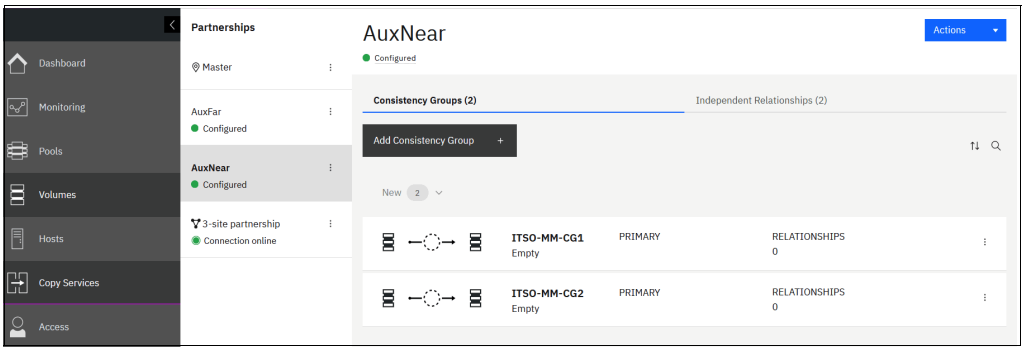

Figure 4-3 Listing 2-Site consistency groups

d. To add existing Metro Mirror relationships, click **Independent Relationships** to select and add relationships. To create Metro Mirror relationship, select a consistency group in the list and click **Create Relationship**, as shown in Figure 4-4.

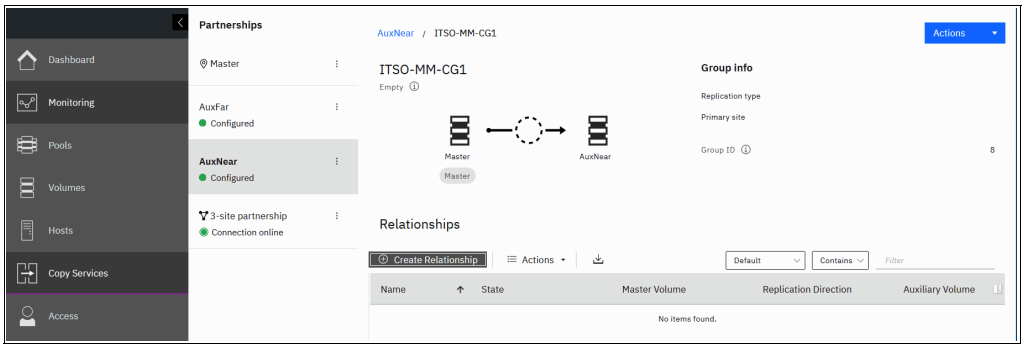

Figure 4-4 Listing and creating relationships in consistency group

e. Select **Metro Mirror** as the copy type. Consistency protection feature for Metro Mirror is supported in with 3-Site replication starting from Storage Virtualize code v8.4.0.0 and is not supported for older code versions. An example is shown in Figure 4-5 on page 61. Click **Next**.

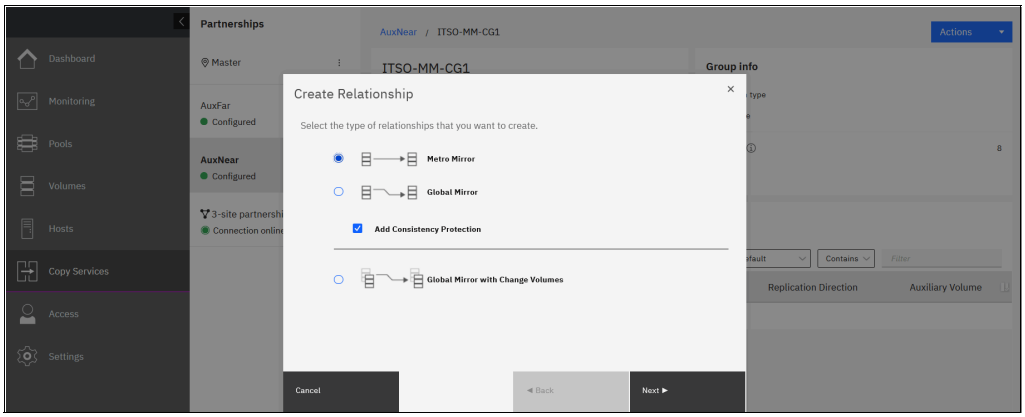

Figure 4-5 Selecting copy type

f. Select the Master and Auxiliary volumes that are to be in a relationship, as shown in Figure 4-6.

> **Note:** After you select a Master volume, only suitable candidates are displayed that are available for selection as an Auxiliary. If your wanted auxiliary volume is not in the list, verify that it has the same provisioned capacity as the Master volume.

You can create multiple relationships by clicking **Add**. If you enabled the consistency protection feature, it is suggested that you create a Master Change Volume or select an existing volume for this role.

When done, click **Next**.

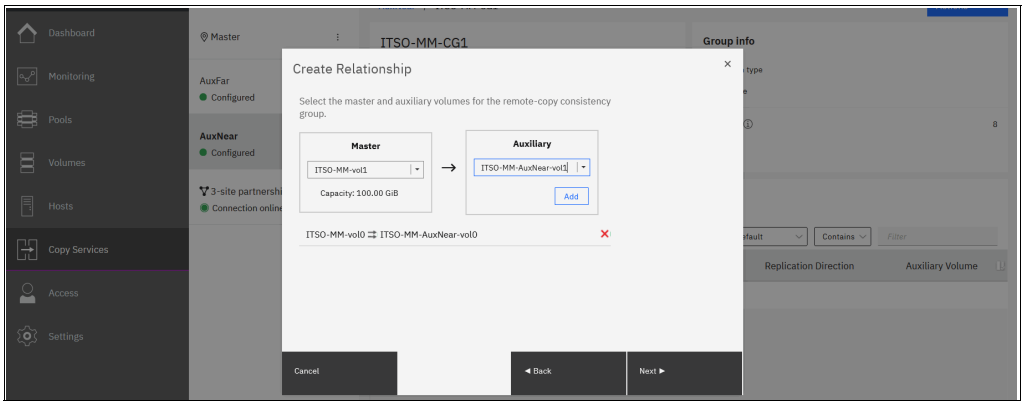

Figure 4-6 Selecting volumes for a relationship

g. In the next window, choose whether synchronization is required. For example, if both Master and Auxiliary volumes are newly created and formatted, no synchronization is needed. If your Master volume contains application data and Auxiliary is new, synchronization is required. If both volumes were created and have no data on them, no synchronization is needed.

3. The consistency group is in the stopped state. Click **Actions** → **Start Group** to start 2-Site C, as shown in Figure 4-7 on page 62.

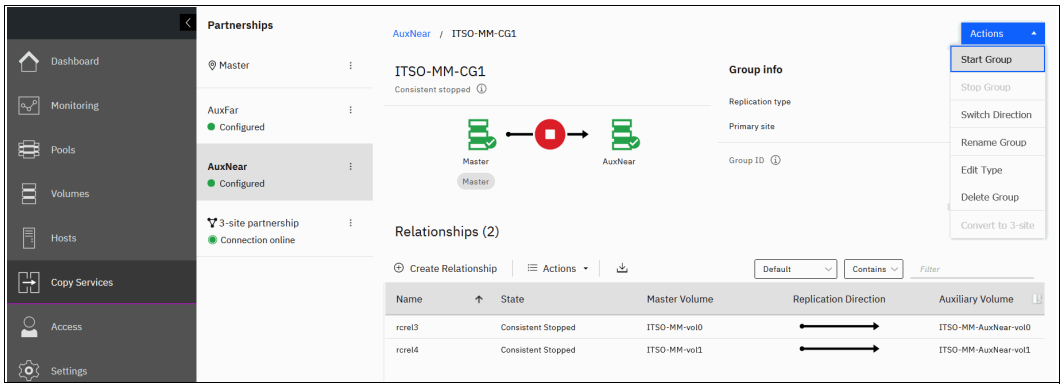

Figure 4-7 Starting consistency group

If your relationships need synchronization, the consistency group is in an `Inconsistent_Copying` state until synchronization is complete.

When the consistency group is in a `Consistent_Syncronized` state, it can be converted to 3-Site mode.

4.2.2 Converting 2-Site Metro Mirror relationships to 3-Site

Before the third volume replicas can be added, determine which storage pool on the AuxFar system is to be used to store them. Create Standard or Data Reduction pools (if required) and note their IDs because they are needed later.

To determine the pool ID, run the **`lsmdiskgrp`** CLI command in AuxFar system CLI, or select **Pools** → **Pools** in the AuxFar system GUI and enable displaying pool IDs, as shown in Figure 4-8.

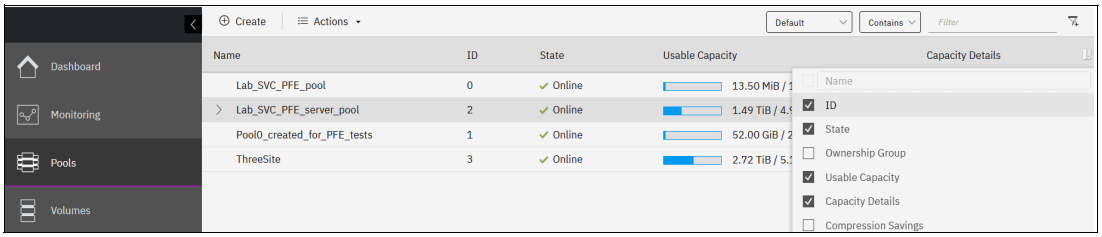

Figure 4-8 Displaying pool ID on AuxFar system

To convert a 2-Site Metro Mirror consistency group to 3-Site mode, you can use the GUI of a Master or AuxNear system. You cannot perform this operation from the GUI of the AuxFar system, as it is not aware of the 2-Site replication relationships between Master and AuxNear.

As the example in Figure 4-9 on page 63 shows, convert a 2-Site consistency group by navigating to **Copy Services** → **Remote Copy**. Then, select **3-Site partnership** in the Partnerships column. Click **Convert 2-Site consistency group**.

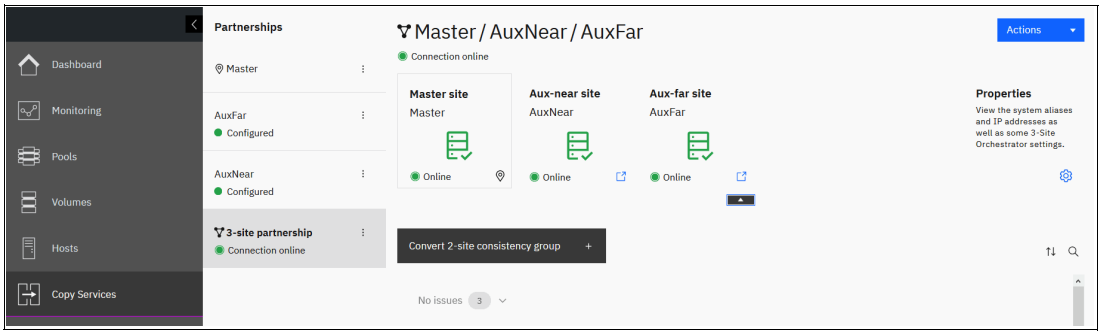

Figure 4-9 Convert 2-Site consistency group button

A dialog opens that shows a list of configured 2-Site consistency groups that are available for conversion, as shown in Figure 4-10. For conversion to be possible, a 2-Site consistency group must be in `Consistent_Syncronized` state. Select the consistency group to convert and click **Next**.

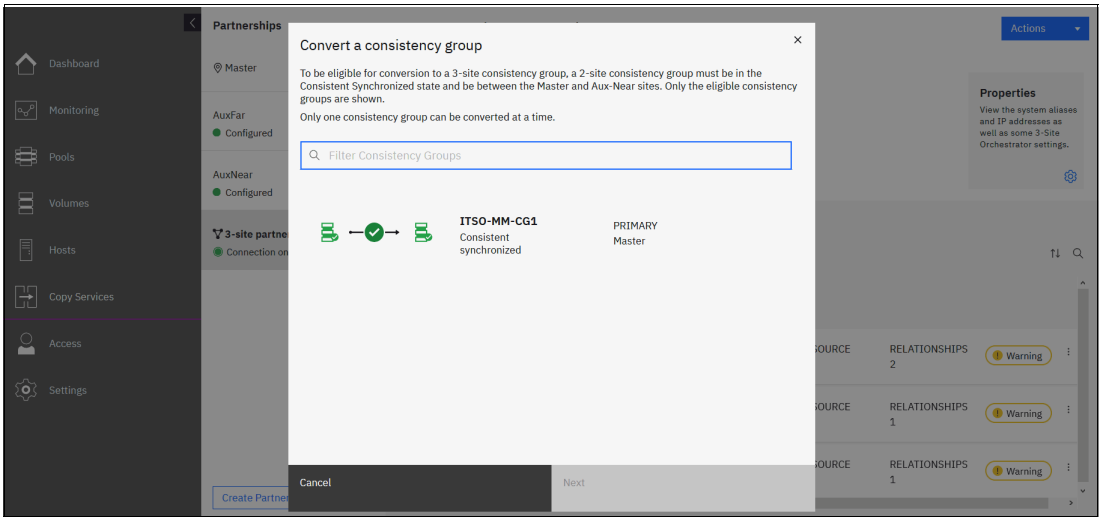

Figure 4-10 List of 2-Site consistency groups that are ready to be converted

Choose the **cycling period and 3-Site topology** parameters of the 3-Site consistency group, as shown in Figure 4-11 on page 64, then click **Next**.

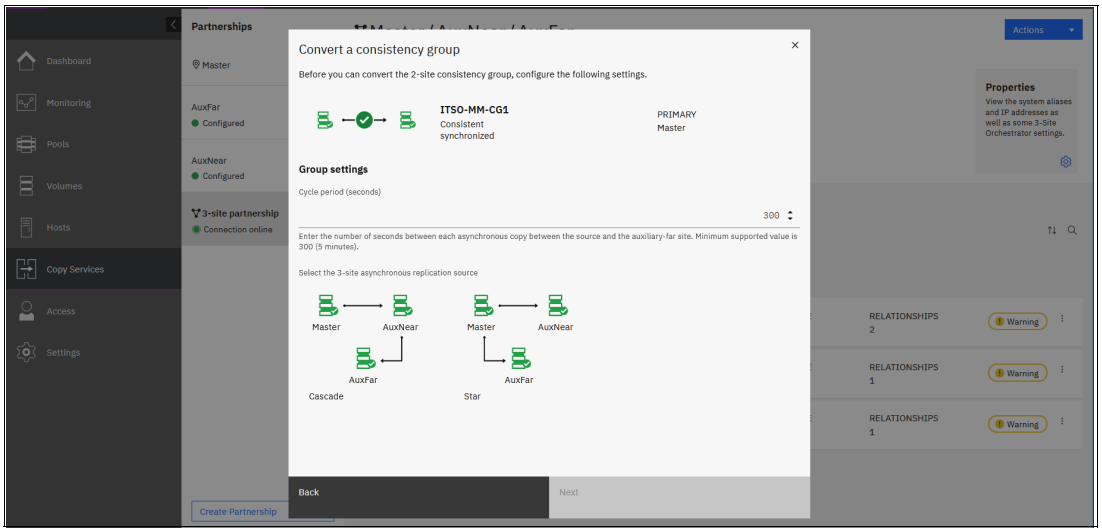

Figure 4-11 Selecting 3-Site consistency group parameters

On the next page, enter the pool ID of a storage pool on the AuxFar site on which the volumes are automatically created. How to determine pool ID on the AuxFar system was explained in the beginning of this section on page 62.

Set the pool parameters and capacity saving options for volumes on the third site. The conversion dialog that is run on Master or AuxNear systems does not perform any checks on the AuxFar system. If the specified pool does not exist or the specified capacity saving options are not supported by the AuxFar system, the conversion dialog returns an error and must be restarted.

Note: Verify that the correct input is provided to this dialog.

Spectrum Virtualize code version 8.4.0.2 has an issue that results in mis-configured 3-Site consistency groups if incorrect parameters are given in this step (for example, a non-existent storage pool is specified).

If the conversion wizard returns an error, do not resume configuring the 3-Site consistency by using the GUI; instead, revert the incomplete 3-Site consistency group to 2-Site, and restart the conversion.

On the same dialog page, you can specify the name template for the volumes on the third site, as shown in Figure 4-12 on page 65.

When all of the parameters are set, click **Convert**.

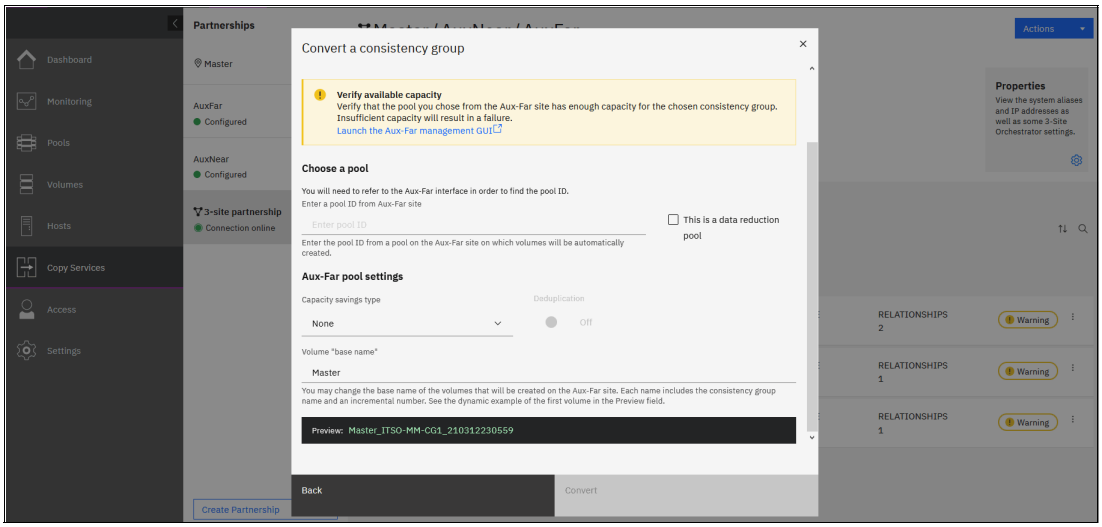

Figure 4-12 Selecting AuxFar volume parameters

The 2-Site replications in a consistency group are converted into 3-Site mode individually. When all of them are converted, the GUI shows the 3-Site consistency group in `Initializing` state, as shown in Figure 4-13. The progress of the initial copy from the source to the third site can be monitored in the Asynchronous Copy Progress column for each relationship.

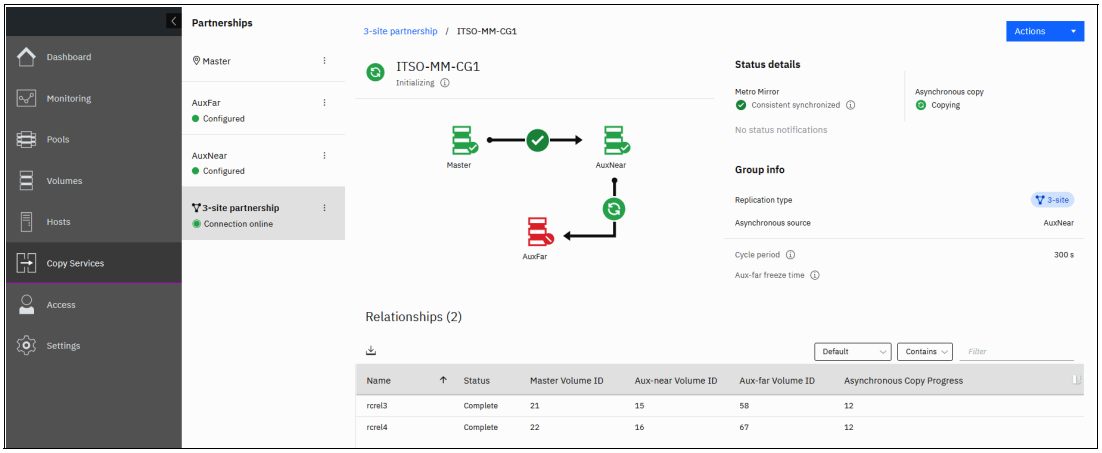

Figure 4-13 Converted 3-Site consistency group

After the initial copy is complete, the state of the 3-Site consistency group becomes `Consistent` and starts data cycling automatically.

4.3 Creating HyperSwap volumes with a copy on the third site

If you are deploying a new HyperSwap solution and your volumes were not yet created and loaded with host data, you can create HyperSwap volumes with a copy on the third site at the beginning of the process. This approach allows you to avoid initial synchronization of a third copy with HyperSwap volume.

To create a HyperSwap volume with a copy on the third site, prepare your configuration as described in Chapter 3, “Configuring a 3-Site Replication solution” on page 37. After the 3-Site configuration is created, use the GUI of a HyperSwap system (spread across Master

and AuxNear sites) to browse to **Volumes** → **Volumes** and click **Create Volumes.** In a dialog window, click the **HyperSwap** tab to switch to creating HyperSwap volumes, as shown in Figure 4-14.

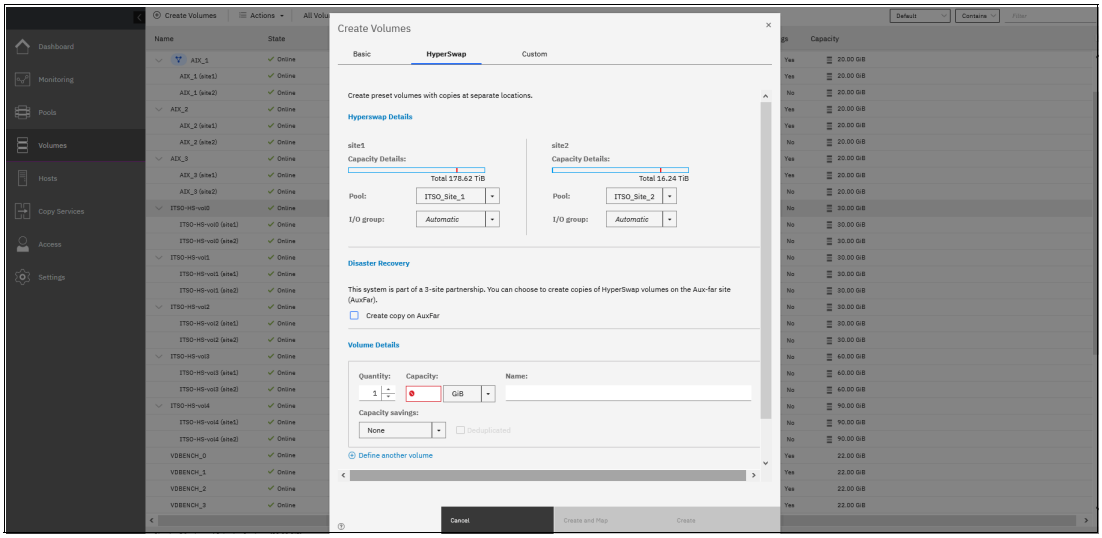

Figure 4-14 Create HyperSwap volumes dialog

If 3-Site configuration is created, the dialog includes a **Disaster Recovery** section. In the **Disaster Recovery** section is an option to create a third copy of a new HyperSwap volume on the system that is assigned to the third site. In this example, the name of the third site is AuxFar. Select the **Create copy** option. A button in lower right changes from Create to Next.

Enter the HyperSwap volume parameters and the quantity of volumes to be created, and click **Next**. The GUI runs commands to create HyperSwap volumes. Click **Continue** to start adding a third copy, as shown in Figure 4-15.

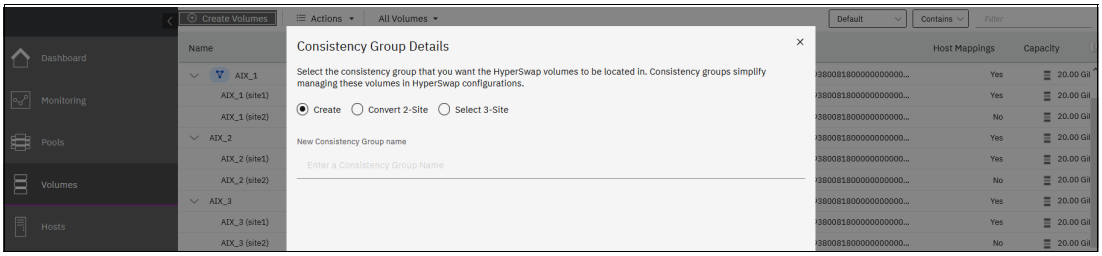

Figure 4-15 Adding HS volumes to a new or existing consistency group

A 3-Site configuration operates with consistency groups. You cannot run 3-Site Replication on a volume that is not a member of any consistency group. The dialog features the following options:

- Create

  If this option is selected, a new 2-Site consistency group is created. Volumes that were created are added and the consistency group is converted to 3-Site mode.

- Convert 2-Site

  If 2-Site consistency groups exists, you can select this option to add the newly created volumes to one of them. When this option is selected, a list of 2-Site consistency groups is shown that are suitable for conversion to 3-Site mode. Volumes are added to the selected group and then the group is converted.

- Select 3-Site

  You can also select the 3-Site consistency group and add new volumes. When newly created volumes join a 3-Site consistency group, a third copy is created.

After you set a name for the new consistency group or select an existing 2-Site consistency group to add new volumes, as the next step you need to set 3-Site consistency group settings. Enter the cycling period, as shown in Figure 4-16, and click **Next**.

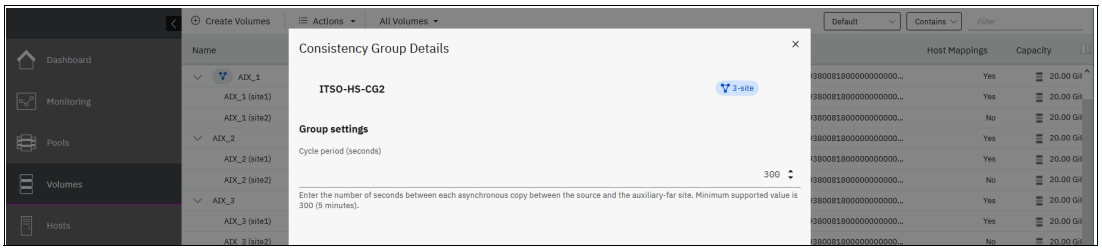

Figure 4-16 Configuring 3-Site consistency group cycling period

On the next page of the dialog, enter the pool ID of a storage pool on the AuxFar site, on which volumes are to be automatically created. How to determine pool ID on the AuxFar system was explained in 4.2.2, “Converting 2-Site Metro Mirror relationships to 3-Site” on page 62.

Set the pool parameters and capacity saving options for volumes on the third site. The conversion dialog that is run on Master or AuxNear systems does not perform any checks on the AuxFar system. If the specified pool does not exist or specified capacity saving options are not supported by the AuxFar system, the conversion dialog returns an error, and the 3-Site consistency group is created in a Partial state.

Note: Verify that correct input is provided to this dialog.

Spectrum Virtualize code version 8.4.0.2 has an issue that results in mis-configured 3-Site consistency groups if incorrect parameters are given in this step (for example, a non-existent storage pool is specified).

If the conversion wizard returns an error, do not resume configuring the 3-Site consistency by using the GUI. Instead, revert the incomplete 3-Site consistency group to 2-Site, and restart the conversion.

On the same dialog page, you can specify the name template for the volumes on the third side, as shown in Figure 4-17 on page 68.

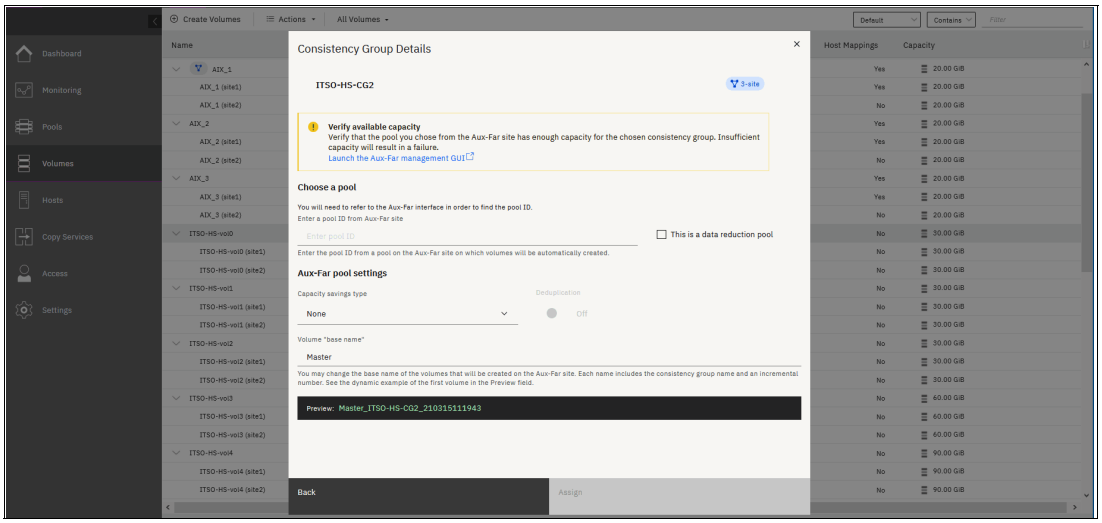

Figure 4-17 3-Site consistency group parameters

After you click **Assign**, the consistency group is created, new HyperSwap volumes are assigned to it, and the group is converted to a 3-Site configuration. You can verify the state of a 3-Site consistency group by browsing to **Copy Services** → **Remote Copy**, clicking **3-Site partnership** and then selecting the created consistency group in the list. See Figure 4-18.

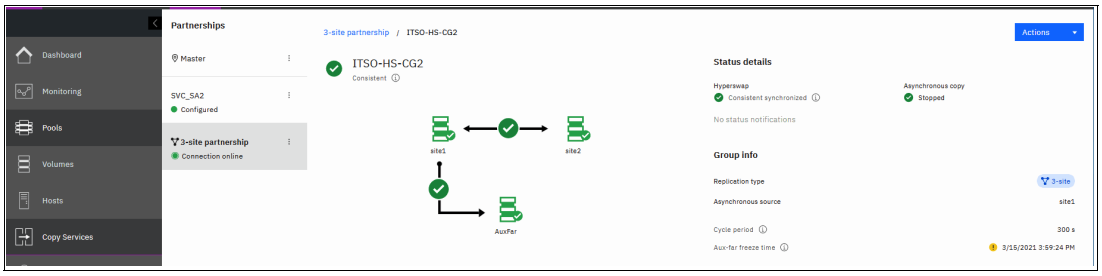

Figure 4-18 State of created 3-Site consistency group

4.4 Converting HyperSwap volumes to 3-Site configuration

Complete the following steps to convert existing synchronized HyperSwap volumes to 3-Site HyperSwap:

1. Use the HyperSwap cluster management GUI to browse to **Copy Services** → **Remote Copy**, and click **Add Consistency Group**. In the Location of target volumes drop-down list, specify the consistency group name and select your HyperSwap system, which is named “Master” in this example. Do *not* select the system where the third copy will be. See Figure 4-19 on page 69.

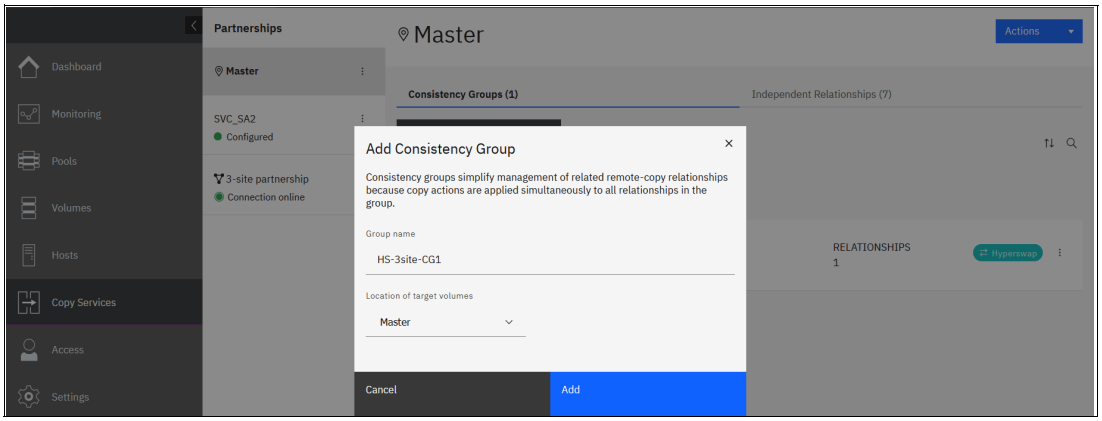

Figure 4-19 Creating consistency group

2. When a group is created, click **Independent Relationships**. Each HyperSwap volume is represented by an RC relationship, which keeps HyperSwap volume copies in-sync.
3. In the list, select one or more HyperSwap volumes to be converted, right-click them, and select **Add to Consistency Group**, as shown in Figure 4-20. A window opens in which you can select a 2-Site consistency group from the drop-down list. Click **Add to Consistency Group**.

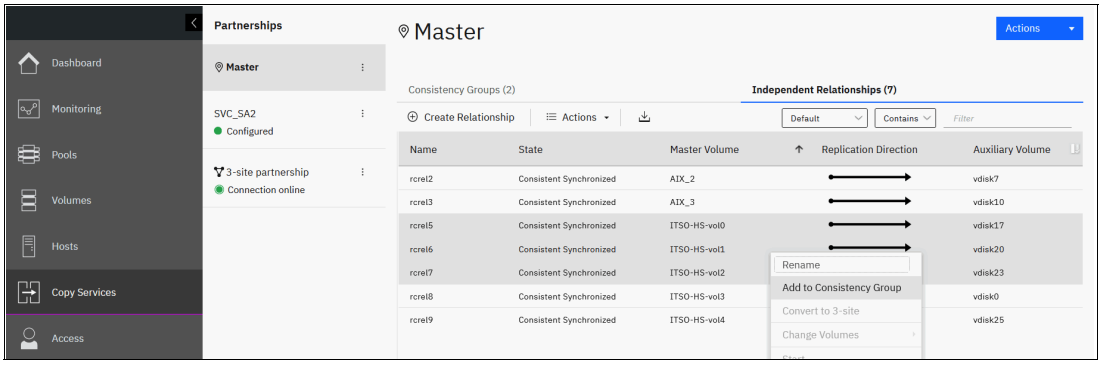

Figure 4-20 Adding HyperSwap volumes to consistency group

4. Return to the consistency groups list. Find the consistency group that must be converted, ensure that it is in the `consistent_synchronized` state, and click the options menu on the right, as shown in Figure 4-21. Select **Convert to 3-Site** to start the conversion wizard.

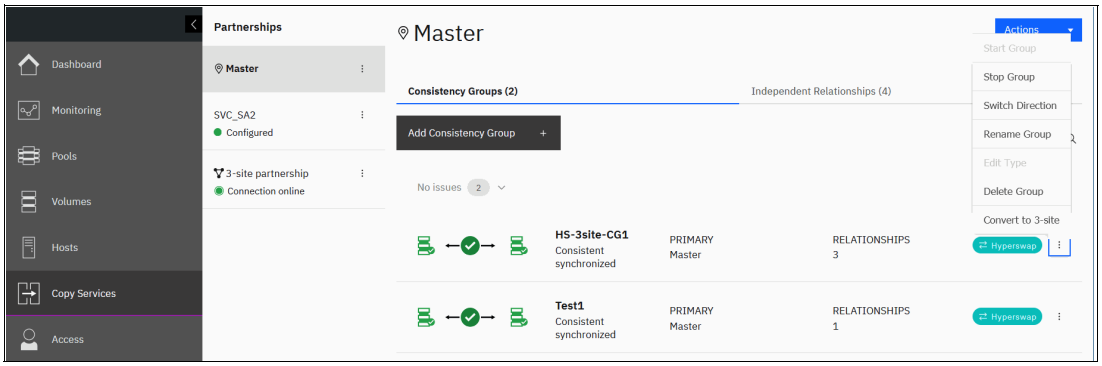

Figure 4-21 Starting conversion of a HyperSwap consistency group to 3-Site

5. On the first page of the wizard, set the cycling period for 3-Site asynchronous replication, as shown in Figure 4-22. Click **Next**.

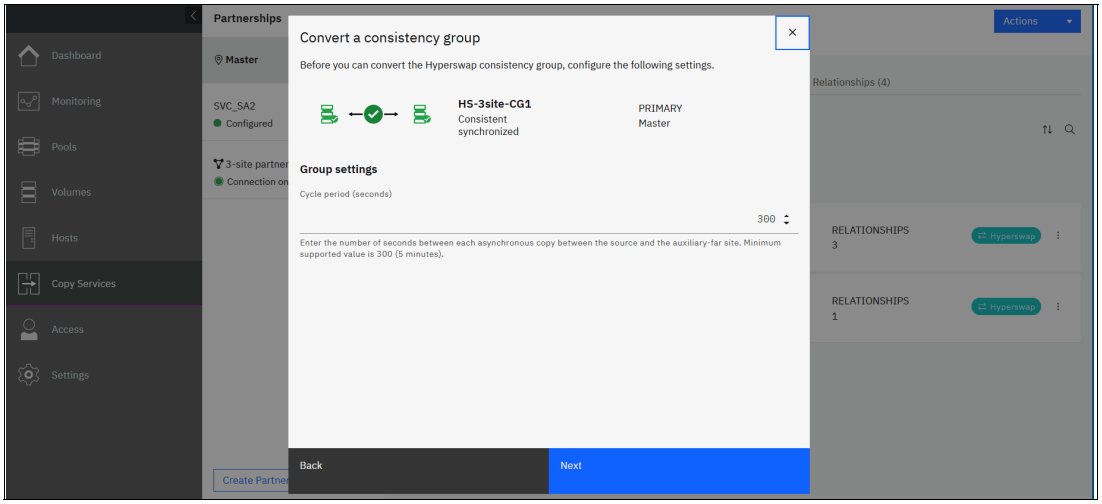

Figure 4-22 Setting cycle period

6. On the next page of the dialog, enter the pool ID of a storage pool on the AuxFar site on which volumes are automatically created. (How to determine pool ID on the AuxFar system was explained in 4.2.2, “Converting 2-Site Metro Mirror relationships to 3-Site” on page 62.)

   Set the pool parameters and capacity saving options for volumes on the third site. The conversion dialog that is run on Master or AuxNear systems does not perform any checks on the AuxFar system. If the specified pool does not exist or specified capacity saving options are not supported by the AuxFar system, the conversion dialog returns an error and the 3-Site consistency group is created in a Partial state.

7. On the same dialog page, specify the name template for the volumes on the third side, as shown in Figure 4-23.

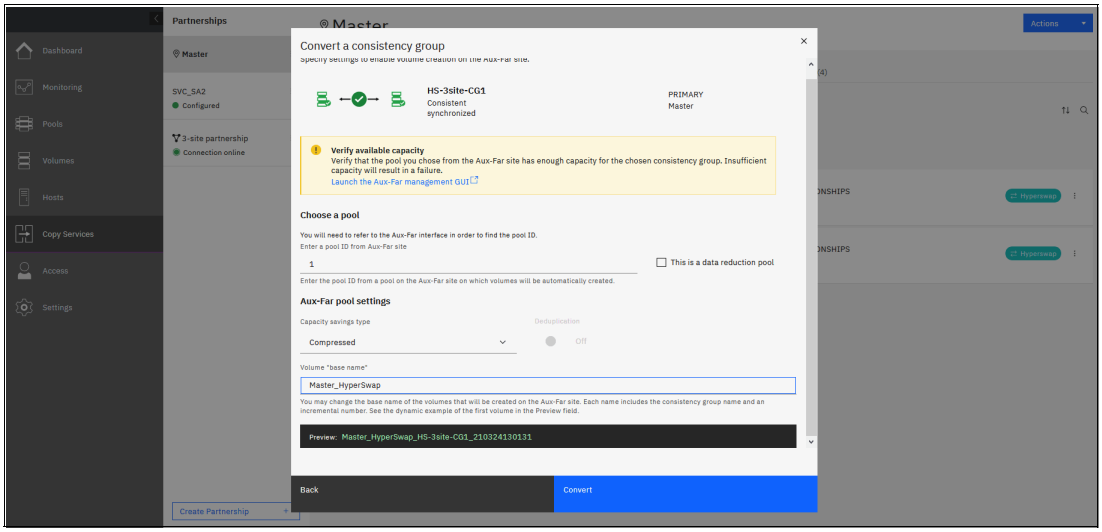

Figure 4-23 Setting third copy parameters

8. Click **Convert** and the consistency group is converted to a 3-Site configuration. You can verify the state of a 3-Site consistency group by browsing to **Copy Services** → **Remote Copy**, clicking **3-Site partnership**, and then, selecting the created consistency group in the list.

After the conversion, volumes on the third side begin initial synchronization. Until the process is complete, the GUI shows the 3-Site consistency group in the `Initializing` state, as shown in Figure 4-24. You can check synchronization progress for each volume in the Asynchronous Copy Progress column. When all relationships are synchronized, the 3-Site consistency group turns to the `Consistent` state.

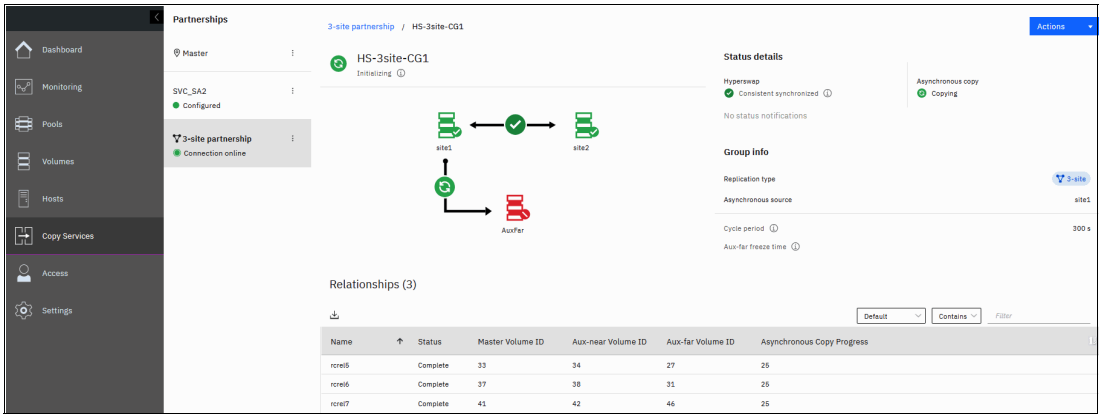

Figure 4-24 3-Site consistency group in Initializing state

4.5 Adding a relationship to a 3-Site consistency group

You can add 2-Site Metro Mirror RC relationships or HyperSwap volumes to existing and running 3-Site consistency groups.

To add an independent relationship to the 3-Site consistency group, the following conditions must be met:

- The 3-Site consistency group is in a consistent state.
- No host I/O operations are running on the independent relationship. The volumes that are used at the master and auxiliary-near sites cannot contain data.
- If a Metro Mirror relationship is added, all of its attributes, such as Master system, Aux system, and copy direction, must match the properties of the 3-Site consistency group to which it is being added.
- The state of the new Metro Mirror relationship must be consistently synchronized, and HyperSwap volume copies must be in-sync.

Note: The relationship that is being added must include empty volumes without any user data. This use case is for new relationships only. Adding relationships with data can result in inconsistent data at the AuxFar far site and prevent accurate disaster recovery.

If relationships with active data exist, you must add these relationships to a new 2-Site consistency group and then convert the group to a 3-Site consistency group. For more information, see 4.4, “Converting HyperSwap volumes to 3-Site configuration” on page 68

To add new HyperSwap volume to a 3-Site replication consistency group, you can use same procedure that is described in 4.3, “Creating HyperSwap volumes with a copy on the third site” on page 65.

The HyperSwap volume includes a creation dialog with which you can select the option to create a third copy. On the next page of the dialog, choose **Select 3-Site** to select a consistency group, as shown in Figure 4-25.

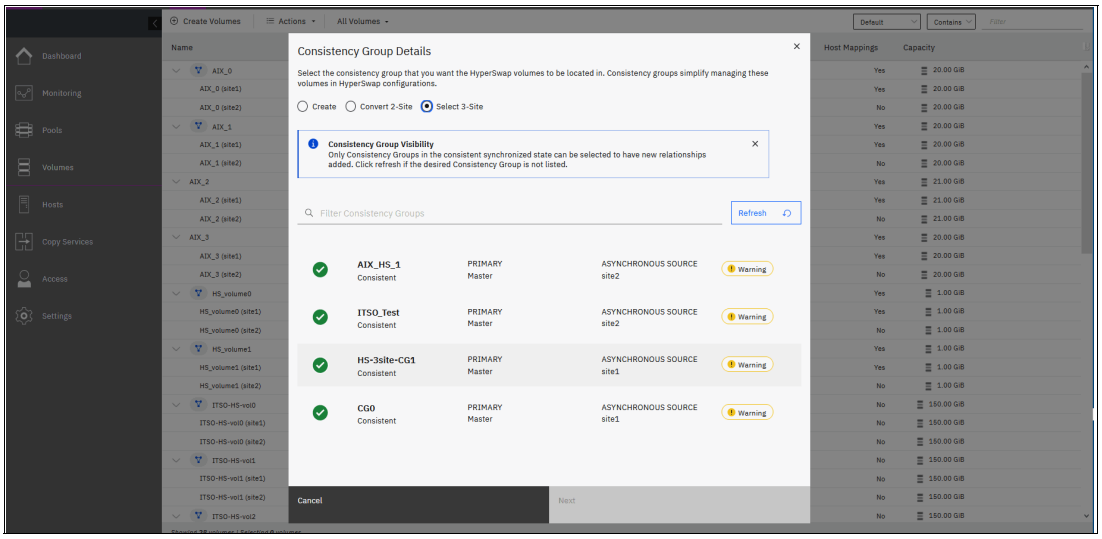

Figure 4-25 Adding new HyperSwap volume to a consistency group

To add new 2-Site Metro Mirror relationships or a HyperSwap volume with no data on it to an active 3-Site consistency group, browse to **Copy Services** → **Remote Copy** in the GUI of a

Metro Mirror Master or Metro Mirror AuxNear system or HyperSwap system. Ensure that the Master or AuxNear system is selected in the Partnerships column and switch to the Independent relationships list. Select one or multiple relationships that you need to add, right-click, and choose **Convert to 3-Site** in the menu, as shown in Figure 4-26.

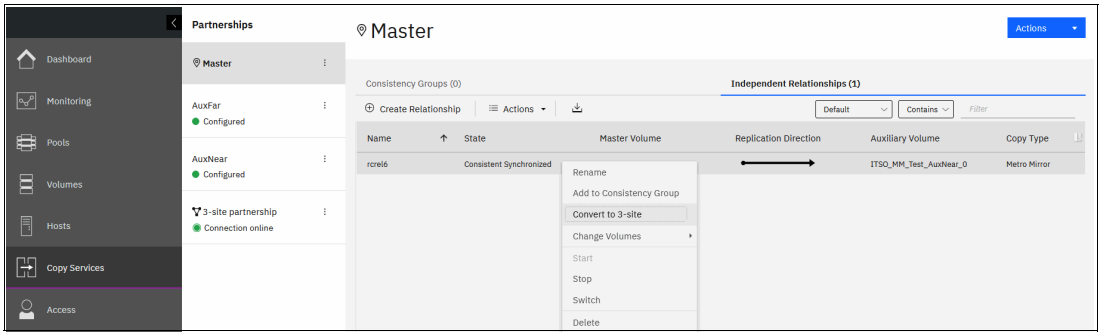

Figure 4-26 Converting independent relationship to 3-Site

A window opens that provides a list of all available 3-Site consistency groups. Only 3-Site consistency groups that are in `consistent` state are displayed. The list displays groups independent of their copy direction, but you can add relationships only if its Primary matches the consistency group's.

Select **3-Site Consistency Group** and click **Next**, as shown in Figure 4-27.

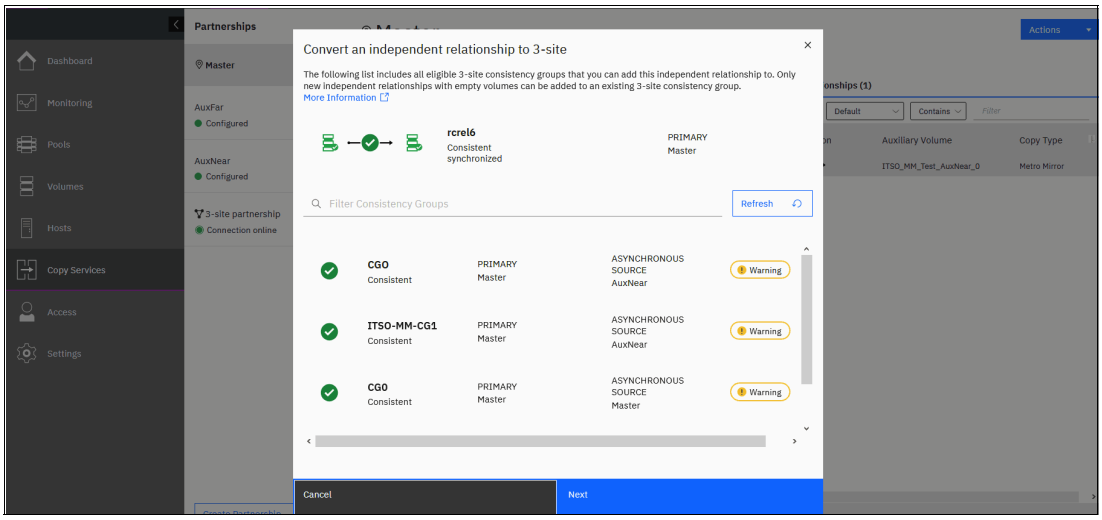

Figure 4-27 3-Site consistency groups where independent relationships can be added

In the next window, you must enter the parameters for the volume on the third side that becomes the asynchronous copy target. That is, enter the pool ID and capacity savings type on the AuxFar system. The parameters can be different than other relationships of the same 3-Site consistency group.

The conversion dialog that is run on Master or AuxNear systems does not perform any checks on the AuxFar system. If the specified pool does not exist or specified capacity saving options are not supported by the AuxFar system, the conversion dialog returns an error, and the relationship is not converted.

On the same dialog page, you can specify the name template for the volumes on the third side, as shown in Figure 4-28 on page 74.

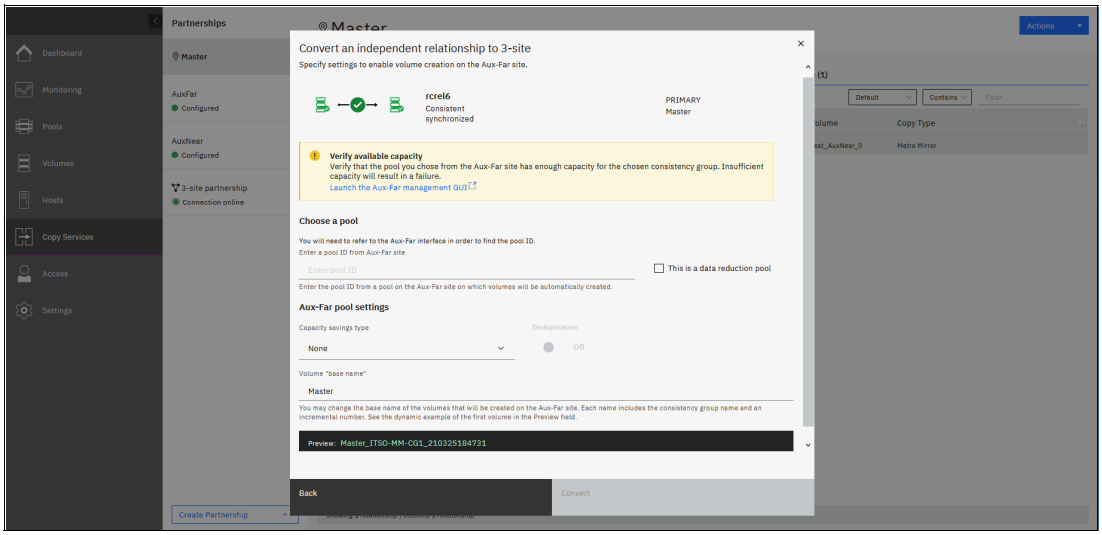

Figure 4-28 Third copy volume parameters

After you click **Convert**, the volumes on the third side are created, relationships are converted, and data cycling is started.

4.6 Reverting 3-Site consistency group to 2-Site configuration

If needed, the 3-Site consistency group can be converted back to a 2-Site replication running between the Master and AuxNear sites, or back to a 2-Site HyperSwap configuration. All relationships within the 3-Site consistency group are converted. Data cycling must be stopped before the conversion.

After the 3-Site consistency group is converted, all volumes on the AuxFar site remain and become accessible. All internal objects that were used for 3-Site replication are deleted.

To convert 3-Site consistency group in Metro Mirror or HyperSwap configuration, complete the following steps:

1. Stop the 3-Site consistency group to stop data cycling. By using the GUI of any of the three systems, browse to **Copy Services** → **Remote Copy** and select **3-Site partnership** in the Partnerships column. Click the menu button that is next to the 3-Site consistency group that you are going to revert, and select **Stop Group**. You are shown a warning that says that the third copy will become progressively out-of-date, click **Confirm**.

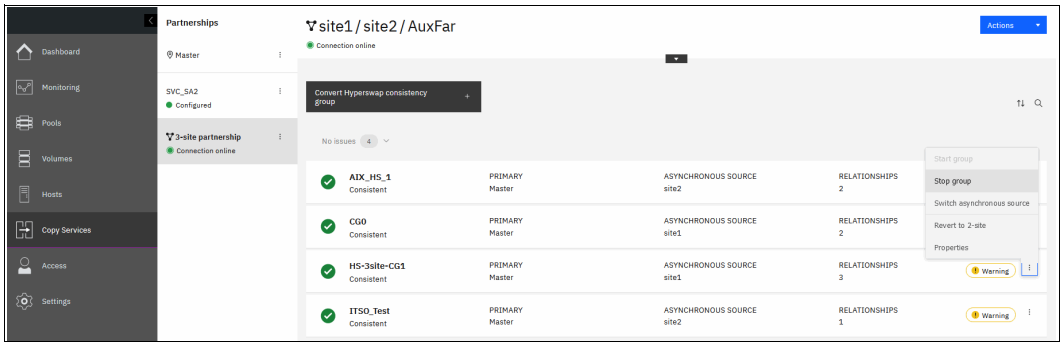

Figure 4-29 Stopping 3-Site consistency group

2. When the 3-Site consistency group is stopped, click again at the menu on the right and select **Revert to 2-Site**, as shown in Figure 4-30. A warning window indicates which of the sites has a Master copy of 2-Site relationship and which has an Aux copy. Click **Revert** to complete the operation.

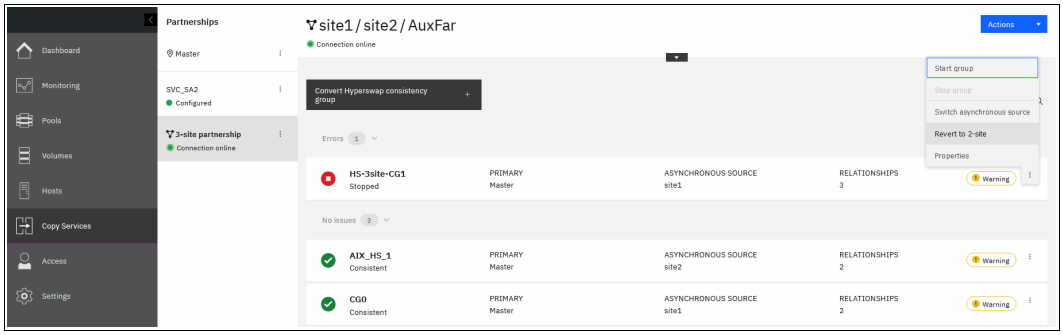

Figure 4-30 Reverting stopped 3-Site consistency group to 2-Site

The reverted consistency group continues to run in 2-Site configuration. If needed, you can stop the 2-Site Metro Mirror consistency group, or delete it. You can also add or remove 2-Site relationships or HyperSwap volumes, and convert them again to a 3-Site configuration.

4.7 Adding existing relationships with data to a 3-Site consistency group

To add a volume or Metro Mirror or HyperSwap stand-alone relationship with data to an existing 3-Site consistency group, you must have 3-Site Orchestrator version 4.0 or later. You can add one or more such objects to the 3-Site consistency group.

4.7.1 Prerequisites

Verify the following prerequisites are met:

- The auxiliary-far site must have sufficient capacity to accommodate the size of volumes being added to the 3-Site consistency group.
- IBM Storage Virtualize 3-Site Orchestrator version 4.0 or later.
- All the attributes of the independent relationship must match the attributes of the 3-Site consistency group that it is being added to.
- The state of the independent relationship must be consistent synchronized.
- Target 3-Site consistency group must be in stopped state.

4.7.2 Adding relationships

Use the GUI to add relationships to a 3-site consistency group:

1. Open the SVC GUI dashboard and click **Copy Services Partnerships** → **Remote Copy**. Under Partnerships, click **3-Site partnership**. See Figure 4-31.

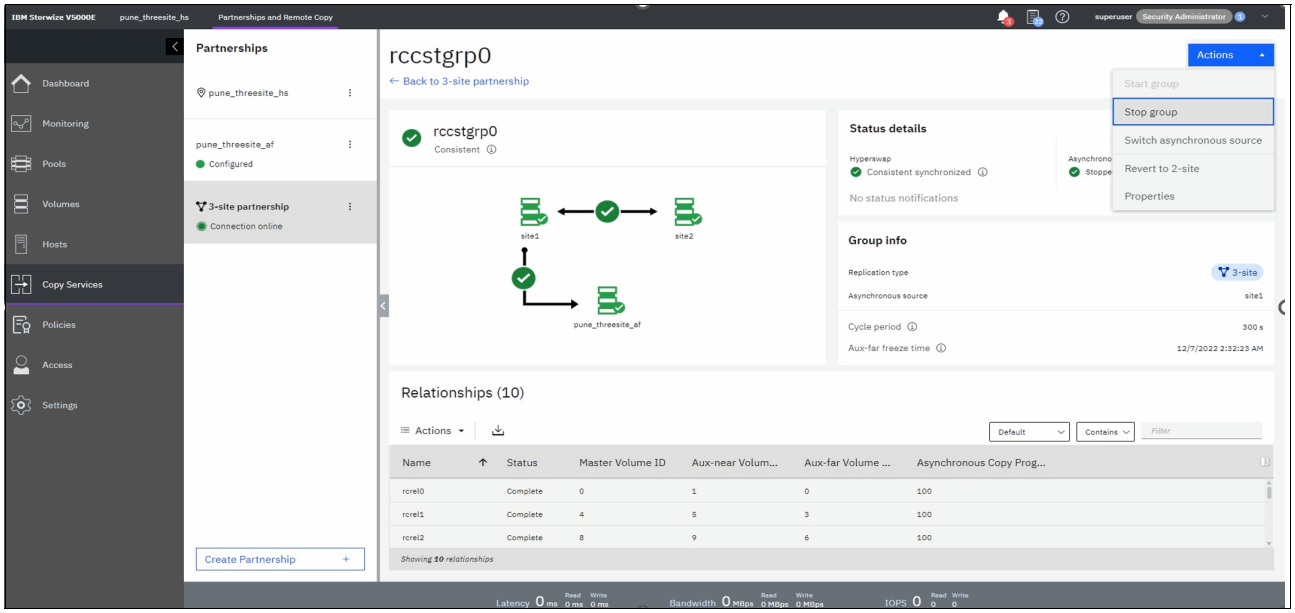

Figure 4-31 3-Site partnership

2. Ensure that the 3-Site consistency group to be added to the independent relationship is in stopped state. If the group is not in `stopped` state, you must stop it before adding the independent relationship. To stop the 3-Site consistency group, select the consistency group and then click **Actions** → **Stop group**. See Figure 4-32.

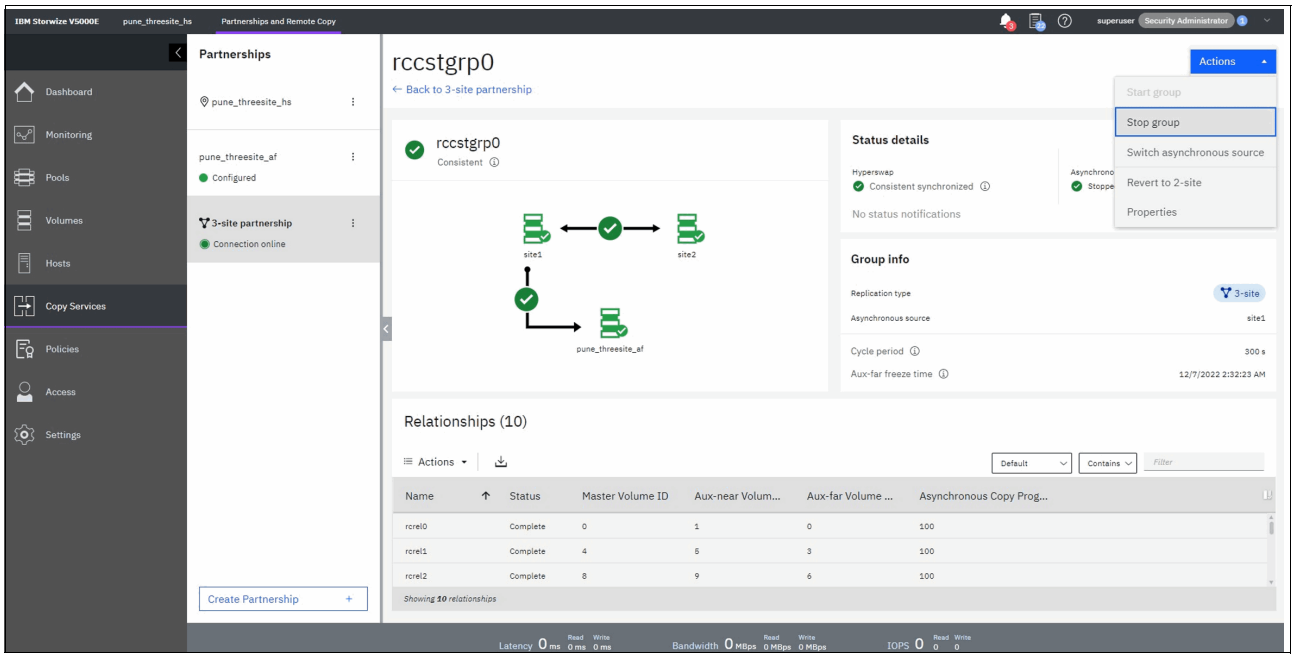

Figure 4-32 Performing stop operation to a 3-Site consistency group

3. Click **Confirm** to proceed. See Figure 4-33.

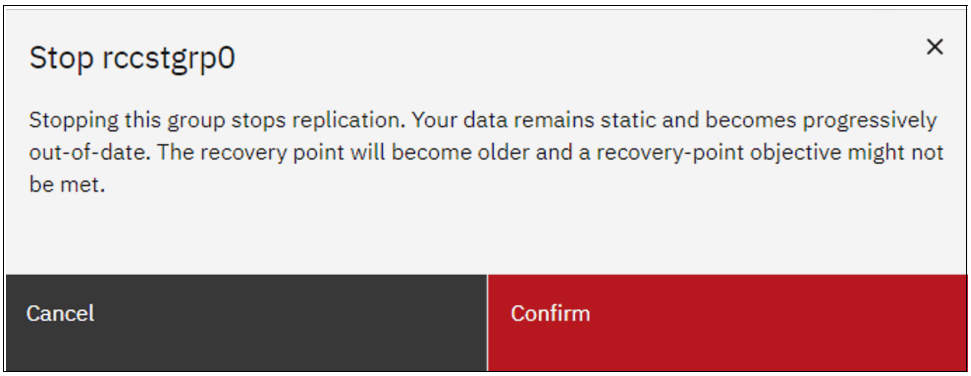

Figure 4-33 Confirmation for stopping 3-Site consistency group

4. Ensure that the consistency group stopped successfully before proceeding. See Figure 4-34.

Figure 4-34 Stop 3-Site consistency group operation completion

5. In the Partnerships panel, select the 2-Site partnership that contains the independent relationship that you want to add to the 3-Site consistency group. See Figure 4-35.

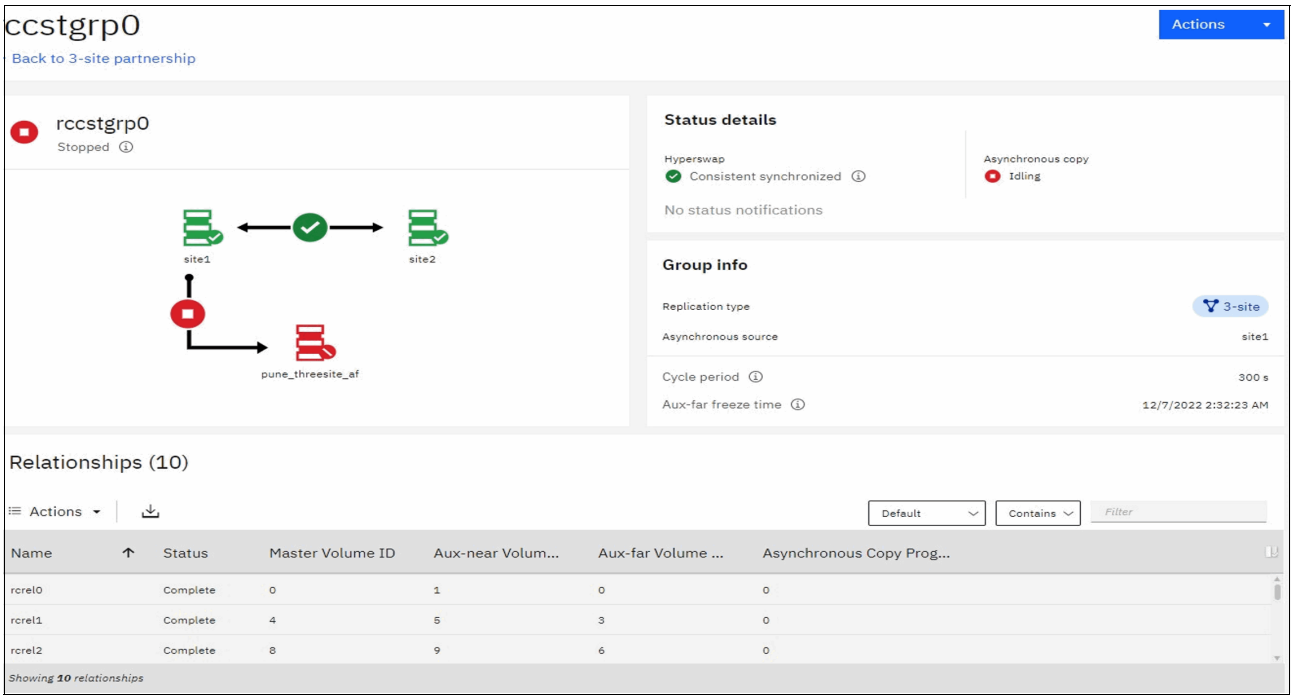

Figure 4-35 3-Site consistency group information

6. On the **Independent Relationships** tab, right-click the independent relationship, and click **Convert to 3-Site**. See Figure 4-36.

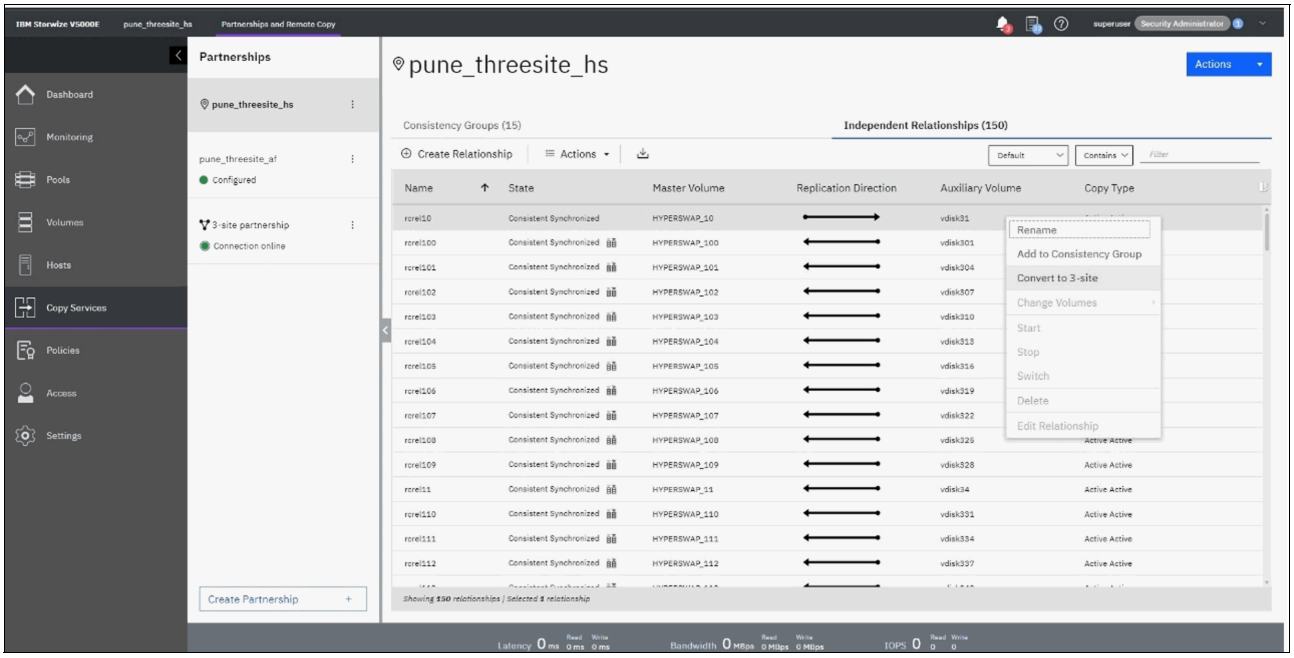

Figure 4-36 Independent relationships

7. In the resulting page, ensure that the **This relationship contains data** is selected. Select the 3-Site consistency group from the list and click **Next**. See Figure 4-37.

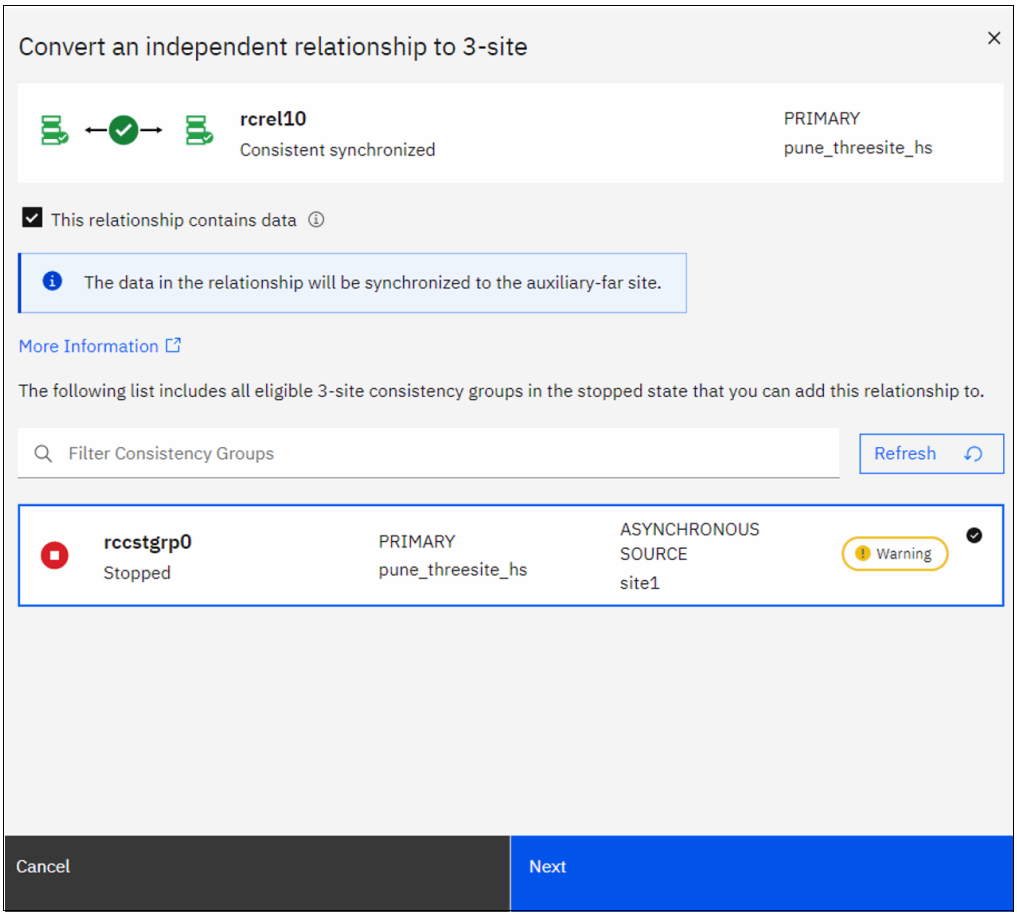

Figure 4-37 Select relationship with data checkbox

8. In the **Choose a pool** field, enter the numerical ID for the storage pool on the auxiliary-far site that provides capacity for the target volumes.

9. To verify these details, click **Launch the Aux-Far management GUI**, and then select → . If the selected pool is a data reduction pool, then select **This is a data reduction pool**. See Figure 4-38.

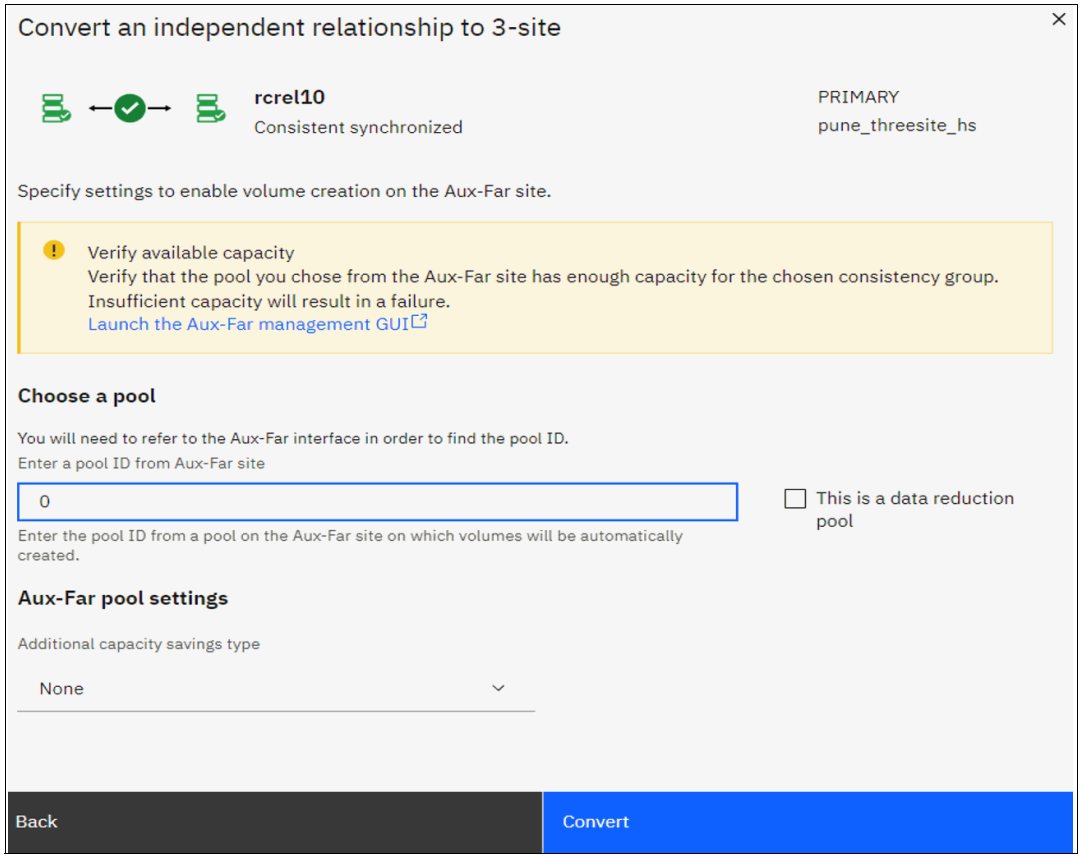

Figure 4-38 Selecting pool information

10. Under Aux-Far pool settings, select the method for saving capacity that this pool uses. Enter a naming convention that identifies the volume as part of the 3-Site consistency group in the **Volume base name** field.

Note: The names are automatically generated to include the specified base name, the name of the 3-Site consistency group, and an incremental number.

11. Click **Convert** and then click close on the following window to proceed. See Figure 4-39.

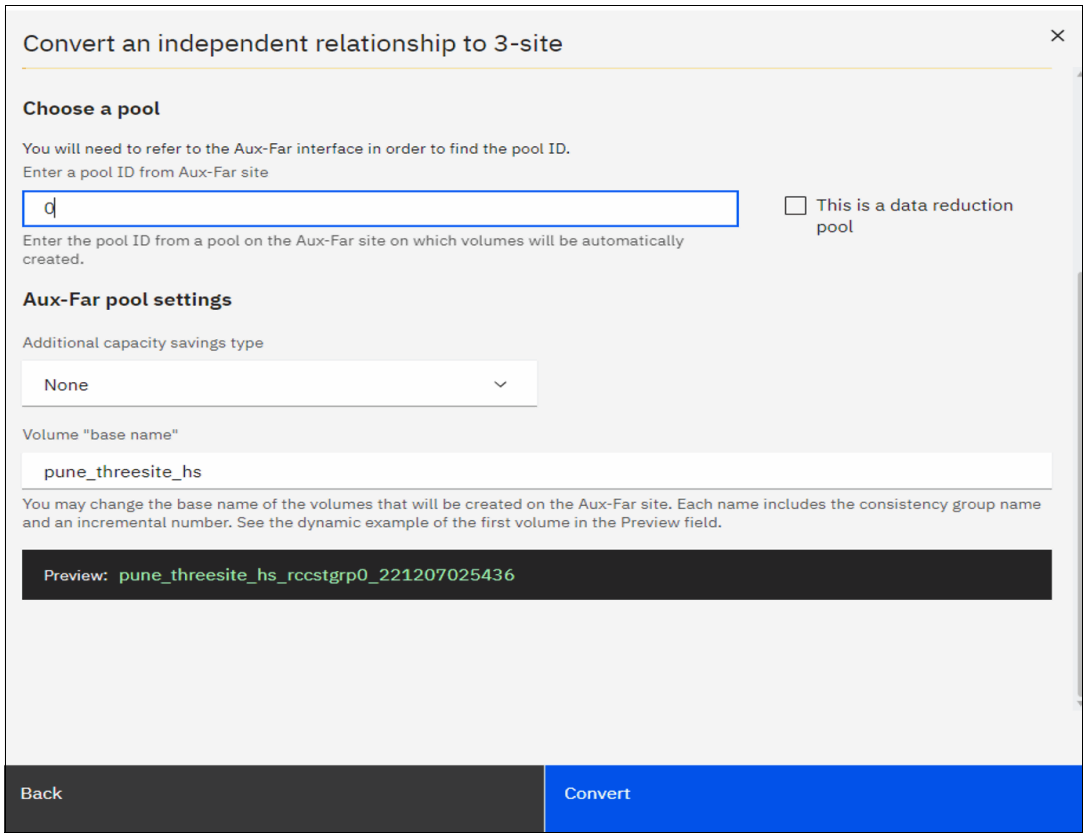

Figure 4-39 Required information for conversion of independent relationship to 3-Site relationship

Figure 4-40 shows the Add Relationship to 3-Site consistency Group task is successful.

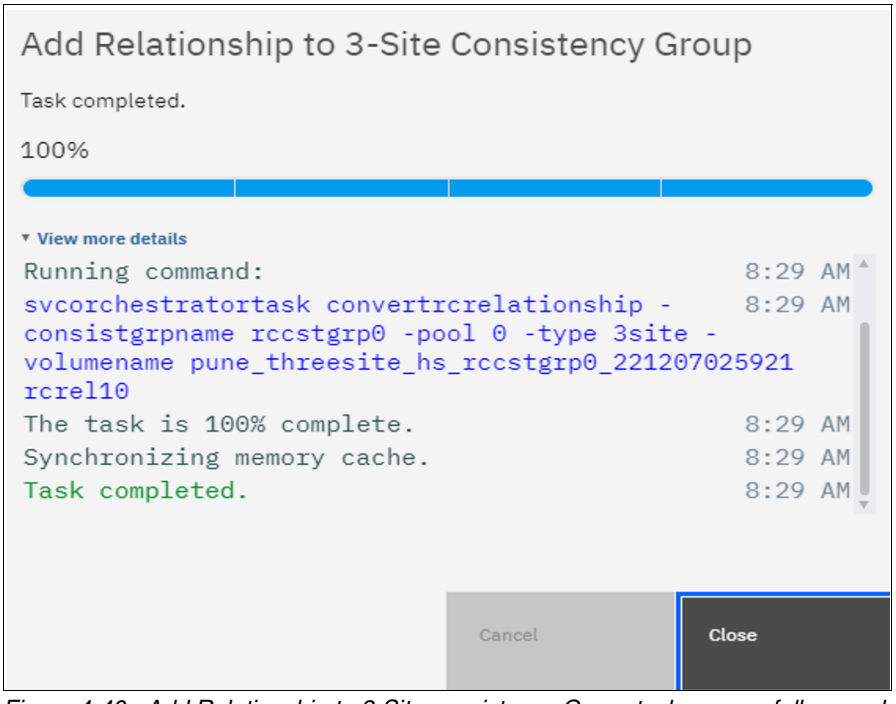

Figure 4-40 Add Relationship to 3-Site consistency Group task successfully completed

12. After the successful conversion, verify that the independent relationship is part of the 3-Site consistency group. Select the 3-Site consistency group and make sure that the new relationship is displayed in the list of relationships with Status as **New**.
13. Add any other independent relationships to the 3-Site consistency group. See Figure 4-41.

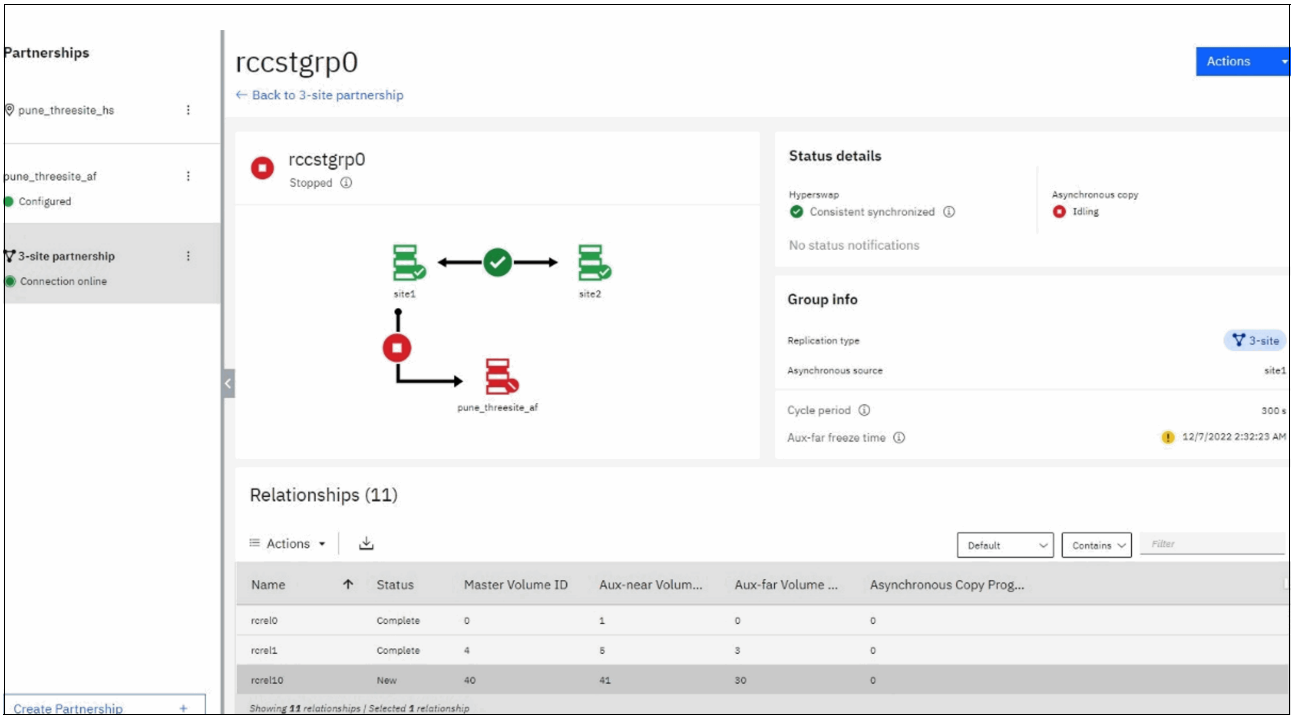

Figure 4-41 Detailed information for 3-Site consistency group

14. In the Partnerships panel, click **3-Site partnership**.

15. Click the **3-Site consistency group** and then click **Actions** → **Start group**. See Figure 4-42.

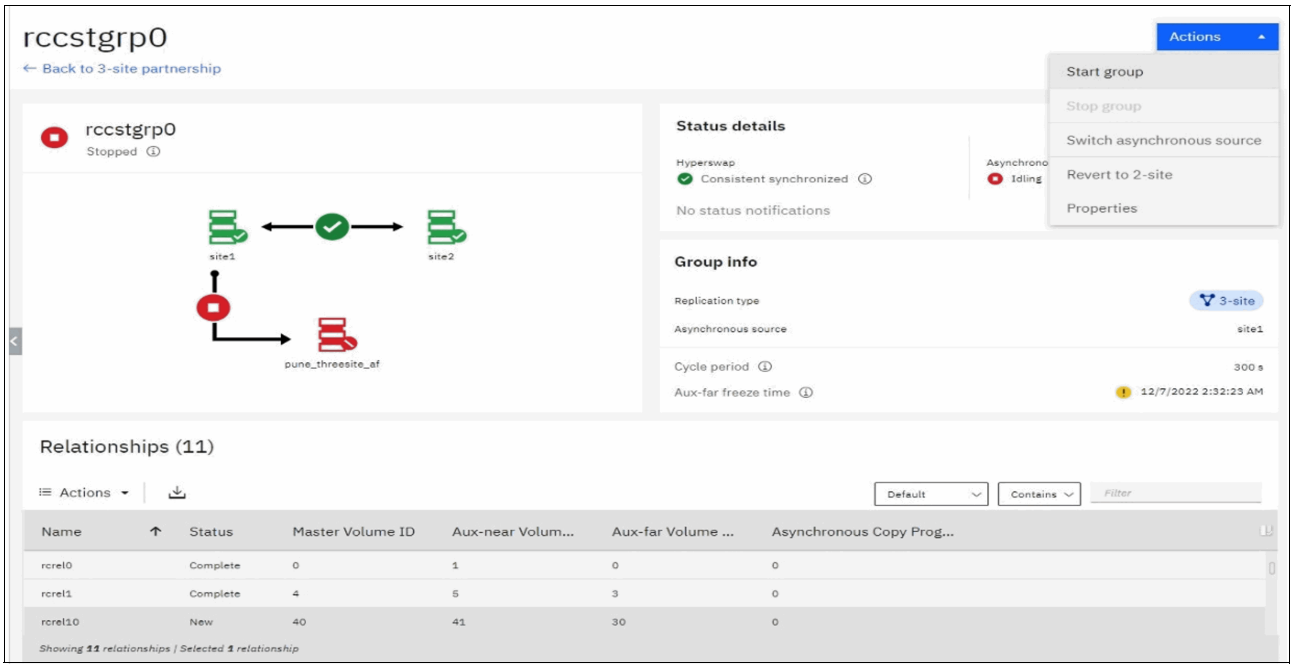

Figure 4-42 Select Start 3-Site consistency group

16. In the resulting pop-up click **Start**.See Figure 4-43. Wait for the task to complete. When the task is 100% completed, the 3-Site consistency group is added successfully.

Figure 4-43 Restart replication

4.8 Removing specific relationships from a 3-Site consistency group

Orchestrator version 4.0 or later is required to remove one or more 3-Site relationships from an existing 3-Site consistency. This operation is applicable for both Metro Mirror and HyperSwap 3-Site configurations. After completion of this operation, the standalone relationship is retained between the Master and Aux-near site, and a volume with data on the Aux-Far site.

4.8.1 Prerequisites and restrictions

Before you remove a 3-Site relationship from a 3-site consistency group, be aware of the following restrictions:

- The 3-Site consistency group must not have any errors, and all three sites must be reachable and online.
- The 3-Site consistency group must be in a stopped state.
- Only one relationship can be removed and reverted at a time.
- The partial state of a 3-Site consistency group is supported only if the **`addnewrelationship`** or **`convertrcrelationship`** commands fail or when removing a relationship from a 3-Site consistency group to a 2-Site consistency group.
- When the 3-Site consistency group is moved to partial state due to failure, the error must be fixed first. The removal will resume only after fixing the error.

4.8.2 Removing a 3-Site relationship from a 3-site consistency group

To remove a 3-Site relationship from a 3-Site consistency group, perform the following steps:

1. Click **Copy Services** → **Partnerships and Remote Copy**.
2. In the Partnerships panel, click **3-Site partnership**.
3. Ensure that the 3-Site consistency group that you want to revert to a 2-Site relationship is in a stopped state. If the consistency group is not in stopped state, you must stop it before adding the independent relationship. To stop the 3-Site consistency group, select the consistency group and then click **Actions** → **Stop group**. See Figure 4-44.

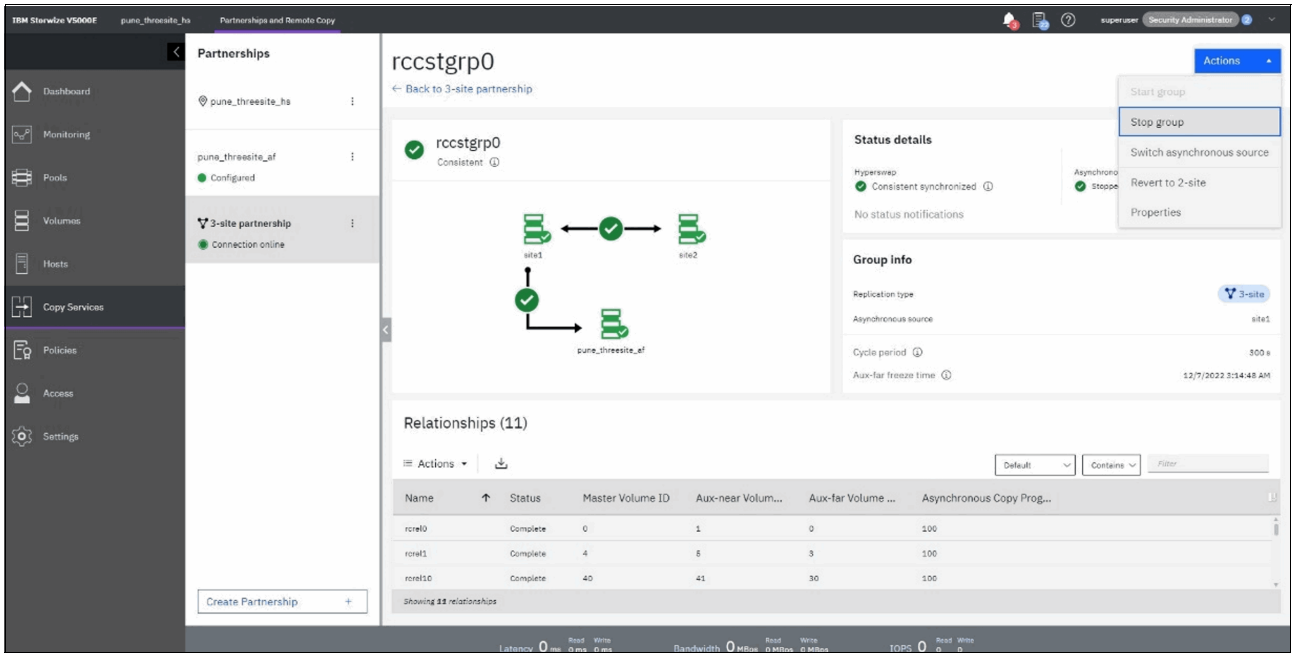

Figure 4-44 Select Stop 3-Site consistency group

4. Click **Confirm** to proceed. See Figure 4-45.

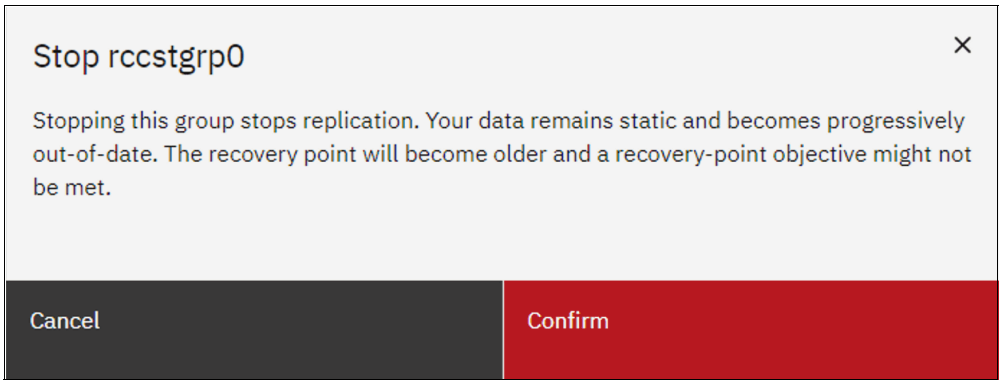

Figure 4-45 Confirmation for stopping 3-Site consistency group

5. Ensure that the state of the consistency group changed to `stopped` before proceeding. See Figure 4-46.

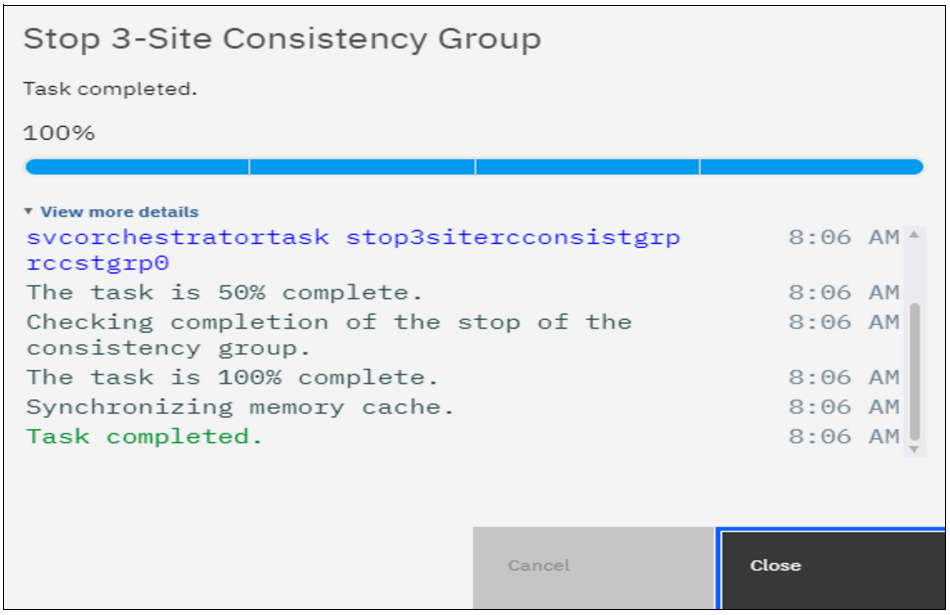

Figure 4-46 Stop 3-Site consistency group operation completion

6. Figure 4-47 shows the 3-Site consistency group information. On the Details page for the selected 3-Site consistency group, right-click the **3-Site relationship** from the list of relationships displayed under Relationships and then select **Remove and Revert to 2-Site**.

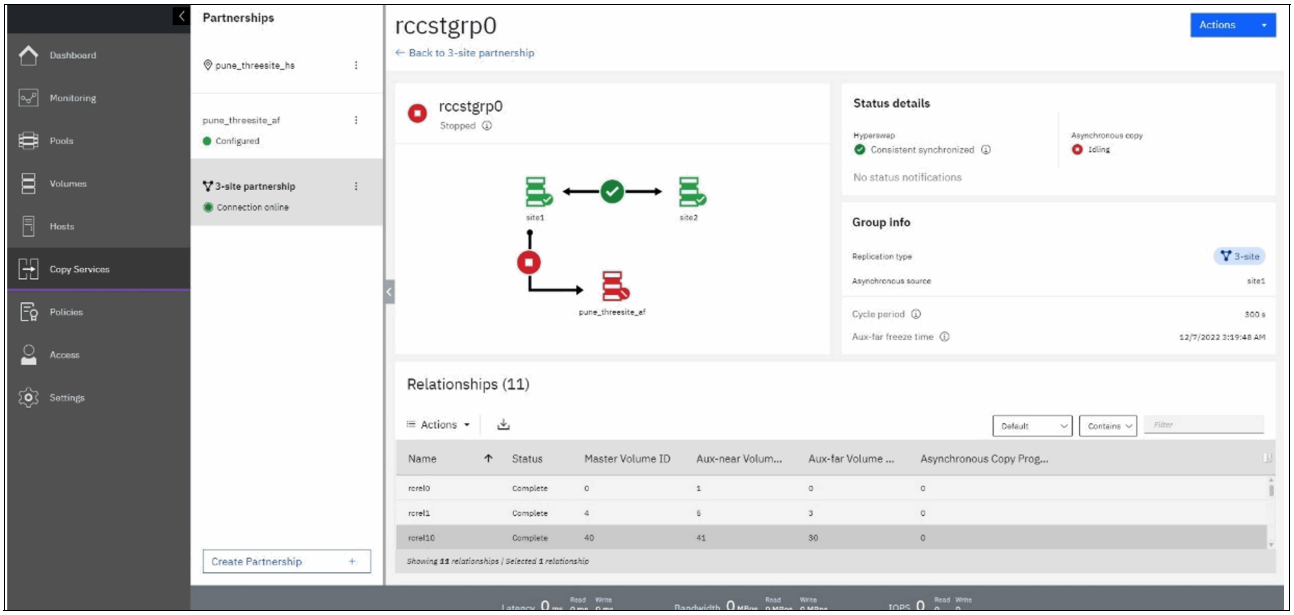

Figure 4-47 3-Site consistency group information

7. Click **Revert** and then click **Close** on the following window to proceed. See Figure 4-48.

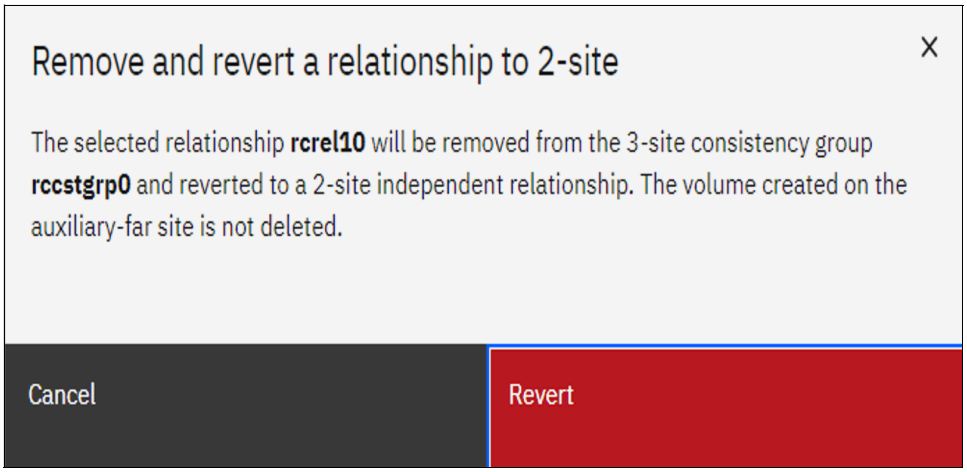

Figure 4-48 Remove and revert a relationship to 2-Site

8. Under Partnerships, click **2-Site partnership** and then click the **Independent Relationships** tab and make sure that the reverted relationship is displayed in the list of independent relationships. This indicates successful completion of the task. See Figure 4-49 on page 87.

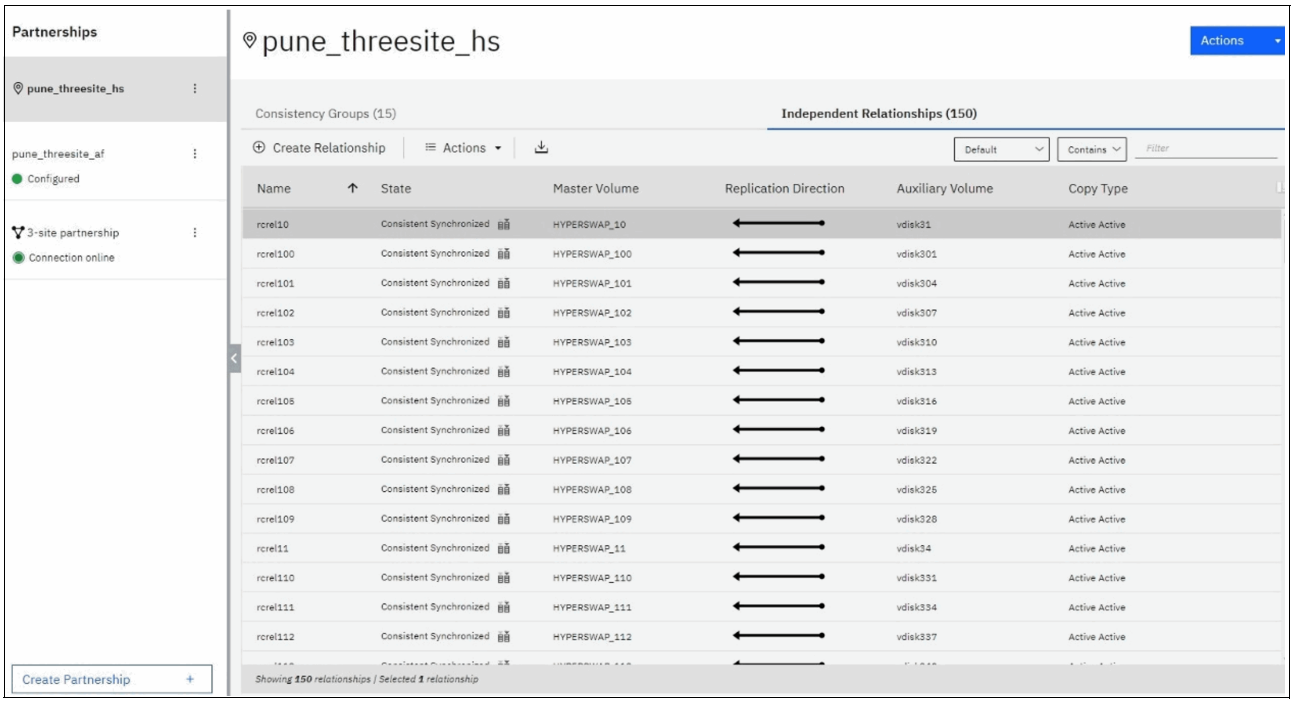

Figure 4-49 Independent relationships

4.9 Using the CLI to create and manage 3-Site relationships

For most of the 3-Site solution configuration tasks, the management GUI is used. However, some scenarios might require the use of 3-Site Orchestrator CLI. This section describes CLI commands that are used to convert 2-site replication relationships to 3-Site and CLI commands that are used to revert 3-Site consistency groups back to 2-site.

3-Site Orchestrator CLI commands can be issued from either the Orchestrator host system CLI or the management CLI of a Spectrum Virtualize system.

To run Orchestrator commands from a Spectrum Virtualize CLI, special prefixes **`svcorchestratorinfo`** and **`svcorchestratortask`** must be used.

The following Orchestrator commands are available with **`svcorchestratorinfo`** prefix:

- `ls3siteconfig`
- `ls3sitercconsistgrp`
- `ls3sitercrelationship`

The following commands require the **`svcorchestratortask`** prefix:

- `convertrcrelationship`
- `convertrcconsistgrp`
- `addnewrcrelationship`
- `ch3sitercconsistgrp`
- `start3sitercconsistgrp`
- `stop3sitercconsistgrp`

4.9.1 Listing 3-Site objects with CLI

Use the **ls3sitercconsistgrp** command to list the detailed status of all 3-Site remote copy consistency groups that exist in a 3-Site configuration. Based on the status of a 3-Site remote copy consistency group, the administrator can take suitable action, as necessary. If 3-Site Orchestrator cannot connect to any of the three sites, this command fails.

This command includes no arguments. Output for this command is listed in Table 4-1.

Table 4-1 ls3sitercconsistgrp command outputs

Attribute	Value
name	Name of a 3-Site remote copy consistency group.
state	State of a 3-Site consistency group. The following values are possible: ► partial ► stopped ► 3site_consistent ► 3site_inconsistent ► 2site_periodic
primary	Alias of a site or cluster name of a system on the primary site for the 3-Site consistency group.
sync_copy_state	State of the synchronous copy between Near sites or HyperSwap state. The following values are possible: ► consistent_synchronized ► idling_disconnected ► idling ► consistent_copying ► consistent_stopped
periodic_copy_state	State of the periodic copy to AuxFar site. The following values are possible: ► copying ► copied ► stopped ► idling
freeze_time	Timestamp of the last available snapshot at the AuxFar site in YY/MM/DD/HH/MM/SS format.
periodic_copy_source	Alias of a site or cluster name of a system that is the source for periodic remote copy.
cycle_period	3-Site replication cycle period in seconds, in the range 300 - 86400 (1 day).
status	3-Site data cycling status for the 3-Site consistency group. For more information about possible statuses, see Chapter 5, “3-Site Replication monitoring and maintenance” on page 99.

Command output example is shown in Example 4-1.

Example 4-1 ls3sitercconsistgrp command example

```
IBM_FlashSystem:Master:superuser>svcorchestratorinfo ls3sitercconsistgrp
name      state            primary periodic_copy_source sync_copy_state
periodic_copy_state freeze_time         cycle_period status
```

```
AIX_HS_1  3site_consistent Master  site2               consistent_synchronized
stopped             2021/03/26/06/09/42 300          online
ITSO_Test 3site_consistent Master  site2               consistent_synchronized
stopped             2021/03/26/06/12/17 300          online
Consistency Group0       3site_consistent Master  site1
consistent_synchronized stopped             2021/03/26/06/11/32 300
online
```

Use the **ls3sitercrelationship** command to list the status for all 3-Site remote copy relationships in a 3-Site configuration.

The copying progress can be different for every relationship in a 3-Site consistency group when the group state is `copying`. When all relationships in a consistency group are in the `consistent_synchronized` state, copying progress is blank. A relationship has a `partial` status if any of the 3-Site object creation fails for the relationship. If one of the three sites is not reachable from 3-Site Orchestrator, this command fails.

This command includes no arguments. Potential output for this command is listed in Table 4-2.

Table 4-2 ls3sitercrelationship command outputs

Attribute	Value
`name`	Name of the 3-Site remote copy relationship.
`consistgrp`	Name of the parent 3-Site remote copy consistency group.
`master_vdisk_id`	Master volume ID in the 3-Site remote copy relationship.
`auxnear_vdisk_id`	AuxNear volume ID in the 3-Site remote copy relationship.
`auxfar_vdisk_id`	AuxFar volume ID in the 3-Site remote copy relationship.
`sync_copy_progress`	Copying progress for synchronous copy (or HyperSwap volume) as a percentage.
`periodic_copy_progress`	Progress for periodic copy.
`status`	3-Site relationship status. The value is `complete` or `partial`.

Command output example is shown in Example 4-2.

Example 4-2 ls3sitercrelationship command example

```
IBM_FlashSystem:Master:superuser>svcorchestratorinfo ls3sitercrelationship
name    consistgrp master_vdisk_id auxnear_vdisk_id auxfar_vdisk_id sync_copy_progress
periodic_copy_progress status
rcrel0  AIX_HS_1   0            1             7                      100
complete
rcrel1  AIX_HS_1   4            5             10                     100
complete
rcrel14 ITSO_Test  69           70            18                     100
complete
rcrel12 CG0       61           62            42                     100
complete
rcrel11 CG0       57           58            37                     100
complete
```

4.9.2 Converting 2-Site replication relationships to 3-Site

Replication relationships are migrated individually. When the first relationship in a consistency group is converted, you must specify which system (Master or AuxNear) is the source for a replication to the third site, and a cycle period for periodic replication. All other converted relationships in the same consistency group inherit those parameters. After all consistency group members are migrated to 3-Site mode, 3-Site replication starts automatically.

Migration is done by using the **convertrcrelationship** 3-Site Orchestrator command. Its syntax is shown in Example 4-3, and the arguments are listed in Table 4-3.

Example 4-3 convertrcrelationship command syntax

```
>> convertrcrelationship -- -type 3site -- -periodicsrc periodic_source_alias ->

>--- -cycleperiod cycleperiod ---------------- -pool auxfar_pool_id ----------->

>----+----------------------------+-----+---------------------------------+--->
   +- -iogrp auxfar_iogrp_id ---+ +- -volumename auxfar_volume_name -+

>----+---------------------------------------+------ 2site_relationship_name
   +- -thin --------+--+------------------+-'
    '- -compressed --'  '- -deduplicated --'
```

Table 4-3 3-Site Replication

Parameters	Description
type	(Required parameter) Intended type of target RC relationship post conversion. Can only be 3site; no other types are available.
periodicsrc	Alias of a site or cluster name of a system that is the source for replication to the third site. Site aliases are defined when the 3-Site configuration is created, **ls3siteconfig** CLI can be used to list them. This parameter is required for the first Metro Mirror relationship being converted within a 2-Site consistency group. For HyperSwap topology, this parameter is optional. If a specific HyperSwap site must play the role of periodicsrc, site_name can be provided. If no value is provided, Orchestrator sets the master site as periodicsrc for the 3-Site consistency group
cycleperiod	Cycle period in seconds for periodic replication. The minimum supported value is 300 (5 minutes). This parameter is required for the first relationship being converted within a 2-Site consistency group.
pool	(Required parameter) Storage pool ID where the replica volume is created on the AuxFar site.
iogrp	(Optional parameter) Caching IO group ID for the new volume on the AuxFar site. If not specified, the I/O group 0 is assigned.
volumename	(Optional parameter) Specifies the name for the new volume on the AuxFar site. If not set, *volumeXX* name is used.

Parameters	Description
thin compressed deduplicated	(Optional parameter) Specifies the volume type for the AuxFar site. It can be thin, compressed, thin deduplicated, or compressed deduplicated. If none of the parameters are set, the volume is created as fully allocated.
relationship_name	(Required parameter) Relationship name of the 2-Site replication relationship that is being converted.

In this example, Consistency group HS-3site-CG1 is converted to 3-Site configuration. It contains three HyperSwap volumes and is linked with RC relationships rcrel5, rcrel6, and rcrel7. To convert the consistency group, complete the following steps:

1. Convert a consistency group to 3-Site mode by converting the first relationship. Use the **convertrcrelationship** command, as shown in Example 4-4.

   On the AuxFar site, 3-Site Orchestrator creates a thin-provisioned volume in Pool ID 1. The **periodicsrc** argument is not supplied because it is a HyperSwap solution.

   The output of the **ls3sitercrelationship** command shows that the 3-Site relationship has a status of complete, which means that the necessary components on all three clusters are created and ready to operate.

   The output of the **ls3siterccconsistgrp** command shows that the consistency group is in a partial state. It indicates that the consistency group, or any of the relationships in it, are incomplete. Also, the periodic_copy_state is none because the 3-Site consistency group does not start copying data to the third site before all 2-Site relationships are converted.

Example 4-4 Converting the first relationship

```
IBM_FlashSystem:Master:superuser>lsrcrelationship
id name            master_cluster_id master_cluster_name master_vdisk_id
master_vdisk_name aux_cluster_id   aux_cluster_name aux_vdisk_id aux_vdisk_name
primary consistency_group_id consistency_group_name state
bg_copy_priority progress copy_type    cycling_mode freeze_time
33 rcrel5          0000010024E00206  Master              33
ITSO-HS-vol0      0000010024E00206 Master           34           vdisk17
master 4            HS-3site-CG1       consistent_synchronized 50
activeactive
37 rcrel6          0000010024E00206  Master              37
ITSO-HS-vol1      0000010024E00206 Master           38           vdisk20
master 4            HS-3site-CG1       consistent_synchronized 50
activeactive
41 rcrel7          0000010024E00206  Master              41
ITSO-HS-vol2      0000010024E00206 Master           42           vdisk23
master 4            HS-3site-CG1       consistent_synchronized 50
activeactive
IBM_FlashSystem:Master:superuser>
IBM_FlashSystem:Master:superuser>svcorchestratortask convertrcrelationship
-type 3site -cycleperiod 300 -pool 1 -thin rcrel5
3-Site RC relationship created with volume id [28].
IBM_FlashSystem:Master:superuser>
IBM_FlashSystem:Master:superuser>svcorchestratorinfo ls3sitercrelationship
name    consistgrp   master_vdisk_id auxnear_vdisk_id auxfar_vdisk_id
sync_copy_progress periodic_copy_progress status
rcrel5 HS-3site-CG1 33         34          28
complete
IBM_FlashSystem:Master:superuser>
```

```
IBM_FlashSystem:Master:superuser>svcorchestratorinfo ls3sitercconsistgrp
name        state           primary periodic_copy_source sync_copy_state
periodic_copy_state freeze_time         cycle_period status
HS-3site-CG1 partial          Master  site1
consistent_synchronized                                      300
```

2. The remaining 2-Site relationships are converted by using the same **convertrcrelationship** command. They can be converted without specifying the **-cycleperiod** and **-periodicsrc** parameters. Example 4-5 shows the conversion commands and their result.

   When all four relationships of the 2-Site consistency group are migrated and shown by 3-Site Orchestrator in `complete` state, the 3-Site consistency group changes its status to `3site_inconsistent`, and the first data replication cycle starts. It can take a significant amount of time for initial synchronization with the third site.

Example 4-5 Completing consistency group conversion to 3-Site

```
IBM_FlashSystem:Master:superuser>svcorchestratortask convertrcrelationship -type 3site
-pool 1 -thin rcrel6
3-Site RC relationship created with volume id [34].
IBM_FlashSystem:Master:superuser>svcorchestratortask convertrcrelationship -type 3site
-pool 1 -thin rcrel7
3-Site RC relationship created with volume id [49].
IBM_FlashSystem:Master:superuser>
IBM_FlashSystem:Master:superuser>svcorchestratorinfo ls3sitercrelationship
name    consistgrp   master_vdisk_id auxnear_vdisk_id auxfar_vdisk_id sync_copy_progress
periodic_copy_progress status
rcrel5  HS-3site-CG1 33              34               28
0                    complete
rcrel6  HS-3site-CG1 37              38               34
0                    complete
rcrel7  HS-3site-CG1 41              42               49
0                    complete
IBM_FlashSystem:Master:superuser>
IBM_FlashSystem:Master:superuser>svcorchestratorinfo ls3sitercconsistgrp
name         state              primary periodic_copy_source sync_copy_state
periodic_copy_state freeze_time         cycle_period status
HS-3site-CG1 3site_inconsistent Master  site1                consistent_synchronized
stopped                                 300          online
```

3. After the first cycle completes successfully, the 3-Site consistency group becomes `3site_consistent`, as shown in Example 4-6. `Periodic_copy_state` remains stopped until the next replication cycle.

Example 4-6 Verifying that 3-Site consistency group is 3site_consistent

```
IBM_FlashSystem:Master:superuser>svcorchestratorinfo ls3sitercrelationship
name    consistgrp   master_vdisk_id auxnear_vdisk_id auxfar_vdisk_id sync_copy_progress
periodic_copy_progress status
rcrel5  HS-3site-CG1 33              34               28
100                  complete
rcrel6  HS-3site-CG1 37              38               34
100                  complete
rcrel7  HS-3site-CG1 41              42               49
100                  complete
IBM_FlashSystem:Master:superuser>
```

```
IBM_FlashSystem:Master:superuser>svcorchestratorinfo ls3sitercconsistgrp
name         state            primary periodic_copy_source sync_copy_state
periodic_copy_state freeze_time         cycle_period status
HS-3site-CG1 3site_consistent Master  site1               consistent_synchronized
stopped             2021/03/26/07/16/32 300          online
```

4.9.3 Adding independent relationships to a 3-Site consistency group

You can add a stand-alone 2-Site replication relationship or a HyperSwap volume to a 3-Site consistency group. The following prerequisites must be met:

- The 3-Site consistency group is in a `3site_consistent` state.
- The stand-alone relationship is created between the new empty volumes with no application data on them and the HyperSwap volume is empty with no user data.
- The state of the stand-alone relationship or HyperSwap relationship that is added is `consistent_synchronized`.
- All of the attributes of a 2-Site relationship that are added match with the target 3-Site consistency group.
- No I/O are operations running on the 2-Site relationship that is being added.

The command **`addnewrcrelationship`** is used for this task. See Example 4-7. Its parameters are listed in Table 4-4.

Example 4-7 Syntax of addnewrcrelationship command

```
>> addnewrcrelationship -- -consistgrpname 3siteCGname -- -pool auxfar_pool -->

>----+----------------------------+-----+---------------------------------+-->
   +- -iogrp auxfar_iogrp_id ---+ +- -volumename auxfar_volume_name -+

>----+-----------------------------------------+------ 2site_relationship_name
   +- -thin --------+--+-----------------+-'
    '- -compressed --'  '- -deduplicated --'
```

Table 4-4 CLI addnewrcrelationship parameters

Parameters	Description
`consistgrpname`	(Required parameter) The name of the 3-Site consistency group to which a relationship is added.
pool	(Required parameter) The Storage pool ID where the replica volume is created on the AuxFar site.
iogrp	(Optional parameter) The caching IO group ID for the new volume on the AuxFar site. If not specified, I/O group 0 is assigned.
volumename	(Optional parameter) Specifies the name for the new volume on the AuxFar site. If not set, *volumeXX* name is given.
thin compressed deduplicated	(Optional parameter) Specifies the volume type for the AuxFar site. It can be thin, compressed, thin deduplicated, or compressed deduplicated. If none of the parameters are set, the volume is created as fully allocated.
2site_relationship_name	(Required parameter) The relationship name of the 2-Site replication relationship that is added.

Example 4-8 shows how an independent HyperSwap volume is added to a consistency group. It also shows that the new HyperSwap volume is created and added to a 3-Site consistency group HS-3site-CG1, which was converted.

Example 4-8 Adding a new relationship

```
IBM_FlashSystem:Master:superuser>mkvolume -name ITSO-HS-new -size 100 -unit gb -pool 0:2
-thin
Volume, id [73], successfully created
IBM_FlashSystem:Master:superuser>lsrcrelationship | grep ITSO-HS-new
73 rcrel15         0000010024E00206  Master              73              ITSO-HS-new
0000010024E00206 Master        74        vdisk46     master
consistent_synchronized 50                        activeactive
IBM_FlashSystem:Master:superuser>
IBM_FlashSystem:Master:superuser>svcorchestratortask addnewrcrelationship -consistgrpname
HS-3site-CG1 -pool 1 -thin rcrel15
3-Site RC relationship created with volume id [52].
IBM_FlashSystem:Master:superuser>
IBM_FlashSystem:Master:superuser>svcorchestratorinfo ls3sitercrelationship
name    consistgrp   master_vdisk_id auxnear_vdisk_id auxfar_vdisk_id sync_copy_progress
periodic_copy_progress status
rcrel15 HS-3site-CG1 73              74               52
100                    complete
rcrel5  HS-3site-CG1 33              34               28
100                    complete
rcrel6  HS-3site-CG1 37              38               34
100                    complete
rcrel7  HS-3site-CG1 41              42               49
100                    complete
IBM_FlashSystem:Master:superuser>svcorchestratorinfo ls3sitercconsistgrp
name         state            primary periodic_copy_source sync_copy_state
periodic_copy_state freeze_time         cycle_period status
HS-3site-CG1 3site_consistent Master  site1          consistent_synchronized idling
2021/03/26/07/26/32 300          online
IBM_FlashSystem:Master:superuser>
IBM_FlashSystem:Master:superuser>svcorchestratorinfo ls3sitercconsistgrp
name         state            primary periodic_copy_source sync_copy_state
periodic_copy_state freeze_time         cycle_period status
HS-3site-CG1 3site_consistent Master  site1          consistent_synchronized stopped
2021/03/26/07/32/07 300          online
```

4.9.4 Reverting 3-Site consistency group to 2-Site mode

The command **convertrcconsistgrp** is used to revert a 3-Site consistency group to a 2-Site consistency group. The command uses the following syntax (its arguments are listed in Table 4-5):

```
>> convertrcconsistgrp -- -type -2site- ----- 3-Site_consistency_group_name
```

Table 4-5 CLI convertrcconsistgrp parameters

Parameter	Description
type	(Required parameter) The intended type of the target RC relationship post conversion. It can be 2site only; no other types are available.
3-Site_consistency_group_name	(Required parameter) The name of a 3-Site consistency group that is converted to 2-Site.

For conversion to 2-Site to be possible, the state of the 3-Site consistency group must be `stopped` or `partial`. To stop it, the 3-Site Orchestrator **`stop3sitercconsistgrp`** command must be issued. The only required argument for this command is the name of the consistency group that must be stopped.

Issue a **`stop3sitercconsistgrp`** command to switch the 3-Site consistency group to a `stopped` state, as shown in Example 4-9.

Example 4-9 Stopping 3-Site consistency group

```
IBM_FlashSystem:Master:superuser>svcorchestratorinfo ls3sitercconsistgrp
name         state            primary periodic_copy_source sync_copy_state
periodic_copy_state freeze_time         cycle_period status
HS-3site-CG1 3site_consistent Master  site1               consistent_synchronized
stopped             2021/03/26/07/32/07 300          online
IBM_FlashSystem:Master:superuser>
IBM_FlashSystem:Master:superuser>svcorchestratortask stop3sitercconsistgrp
HS-3site-CG1
IBM_FlashSystem:Master:superuser>
IBM_FlashSystem:Master:superuser>svcorchestratorinfo ls3sitercconsistgrp
name         state            primary periodic_copy_source sync_copy_state
periodic_copy_state freeze_time         cycle_period status
HS-3site-CG1 stopped          Master  site1               consistent_synchronized
idling              2021/03/26/07/32/07 300          online
```

Issue a **`convertrcconsistgrp`** command to revert stopped 3-Site consistency group to 2-Site, as shown in Example 4-10. The consistency group disappears from the `ls3sitercconsistgrp` output after it is reverted, but remains as 2-Site consistency group.

Example 4-10 Converting consistency group to 2-Site

```
IBM_FlashSystem:Master:superuser>svcorchestratortask convertrcconsistgrp -type
2site HS-3site-CG1
IBM_FlashSystem:Master:superuser>
IBM_FlashSystem:Master:superuser>svcorchestratorinfo ls3sitercconsistgrp
IBM_FlashSystem:Master:superuser>
IBM_FlashSystem:Master:superuser>lsrcconsistgrp | grep HS-3site-CG1
4 HS-3site-CG1   0000010024E00206 Master          0000010024E00206 Master
master  consistent_synchronized 4                 activeactive
```

4.9.5 Adding Metro Mirror or HyperSwap relationships with data to a 3-Site consistency group

To add an independent remote-copy relationship with data to a 3-Site consistency group, perform the following steps.

1. Stop the 3-Site consistency group by entering the following command:

   `stop3sitercconsistgrp` *`3site_consistency_group_name`*

   where *`3site_consistency_group_name`* is the name of the 3-Site consistency group that you are adding the remote-copy relationship with data to.

2. Verify that the 3-Site consistency group has stopped, by entering the following command:

   `ls3sitercconsistgrp`

Observe the output to ensure that the `sync_copy_state` is in `consistent_synchronized` and the state is in stopped state, which indicates a successful operation. See Example 4-11

Example 4-11 ls3sitercconsistgrp output

```
Name state primary periodic_copy_source sync_copy_state periodic_copy_state freeze_time cycle_period status
rccstgrp0 stopped Master Auxnear consistent_synchronized idling 2022/01/14/05/33/48 300 online
```

3. Verify the 2-Site relationship, by entering the following command:

   `lsrcrelationship` *`id_or_name`*

   where *`id_or_name`* is the ID or name of the 2-Site relationship. Observe the output to ensure that the state of the 2-Site relationship is `consistent_synchronized`.

4. Add the 2-Site relationship to the 3-Site consistency group, by entering the following command:

   `convertrcrelationship -type 3site -periodicsrc` *`periodic_source_site_name`* `-pool` *`auxfar_pool_id`* `-cycleperiod 500 - consistgrpname` *`name 2site_relationship_name`*

   The command includes the following values:

Parameter	Description
`periodicsrc`	The source site system name or alias for periodic replication
`pool`	The ID for the pool for the new volume on the auxiliary-far site
`cycleperiod`	(Optional) The cycle period in seconds for periodic replication. The value must be greater than or equal to 300.
`consistgrpname`	The name of the 3-Site consistency group to which you are adding the remote-copy relationship
`2site_relationship_name`	The name of the source 2-Site relationship that you want to add to the 3-Site consistency group

   For a Metro Mirror configuration, use the following example:

   ```
   convertrcrelationship -type 3site -periodicsrc Master -cycleperiod 500 -pool 0
   -consistgrpname rccstgrp0 rcrel1
   ```

   For a HyperSwap configuration, use the following example:

   ```
   convertrcrelationship -type 3site -periodicsrc site1 -cycleperiod 500 - pool 0
   -consistgrpname rccstgrp0 rcrel1
   ```

5. Verify that the 2-Site relationship was added successfully to the 3-Site consistency group, by entering the following command:

   `ls3sitercrelationship | grep -i` *`3site_consistency_group_name`*

   where *`3site_consistency_group_name`* is the name of the 3-Site consistency group.

   Observe the output to ensure that the state of the newly added 2-Site relationship is new. In the output in Example 4-12 on page 97, the name of the relationship is rcrel0.

Example 4-12 ls3sitercrelationship output

```
nameconsistgrp master_vdisk_id auxnear_vdisk_id auxfar_vdisk_id sync_copy_progress
periodic_copy_progress status
rcrel0         rccstgrp0       12               12              12
new
```

If the command status shows as `partial`, it indicates that the conversion process has failed. You need to identify and resolve the reported error before rerunning the command.

6. Repeat these steps to add any additional 2-Site relationships with data.
7. Start the 3-Site consistency group, by entering the following command:

   `start3sitercconsistgrp` *`3site_consistency_group_name`*

   where *`3site_consistency_group_name`* is the name of the 3-Site consistency group that the 2-Site relationship was added to.
8. Verify that the newly added 2-Site relationship is part of the 3-Site consistency group by entering the following command:

   `ls3siterelationship | grep -i` *`3site_consistency_group_name`*

   where *`3site_consistency_group_name`* is the name of the 3-Site consistency group. Observe the output to ensure that the state of the newly added 2-Site relationship is complete. In the output in Example 4-13, the name of the relationship is rcrel0.

Example 4-13 ls3siterelationship output

```
name consistgrp master_vdisk_id auxnear_vdisk_id auxfar_vdisk_id sync_copy_progress periodic_copy_progress status
rcrel0 rccstgrp0 12             12               12              100              complete
```

Listing 3-Site objects with CLI

Table 4-6 shows how to list 3-Site objects with CLI.

Table 4-6 Possible results of ls3sitercrelationship command

Result	Description of status
new	Indicates that this is the relationship with data added to existing 3-Site consistency group. Not consistent, to be synced to AuxFar Volume.
complete	Indicates that conversion to 3-Site relationship is successful.
partial	**I**ndicates that conversion from 2-Site relationship to 3- Site relationship has failed with some partial 3-Site object creation.

Resulting output of the **`ls3sitercrelationship`** command is shown in Example 4-14.

Example 4-14 ls3sitercrelationship output - complete

```
name    consistgrp  master_vdisk_id auxnear_vdisk_id auxfar_vdisk_id sync_copy_progress
periodic_copy_progres  status
rcrel0 cg0             0               0                0           70
complete
```

Resulting output of **`ls3sitercrelationship`** command the is shown in Example 4-15.

Example 4-15 s3sitercrelationship output - partial

```
name    consistgrp  master_vdisk_id auxnear_vdisk_id auxfar_vdisk_id sync_copy_progress
periodic_copy_progres  status
rcrel0  cg0         0
partial
```

Resulting output of the **`ls3sitercrelationship`** command is shown in Example 4-16.

Example 4-16 s3sitercrelationship output - new

```
name    consistgrp  master_vdisk_id auxnear_vdisk_id auxfar_vdisk_id sync_copy_progress
periodic_copy_progres  status
rcrel0 rccstgrp0      12               12                 12
new
```

Note: To remove and revert a relationship from 3-Site consistency group, see Removing and reverting relationship from 3-site consistency group.

5

3-Site Replication monitoring and maintenance

This chapter describes monitoring and maintaining 3-Site Replication and includes the following topics:

5.1 Monitoring 3-Site objects by using the 3-Site Orchestrator

The 3-Site Orchestrator features the following commands that are used to check the configuration of the Orchestrator and the status of the 3-Site relationships and consistency groups:

- The command **ls3siteconfig** lists the Orchestrator configuration.
- The command **ls3sitercconsistgrp** lists the detailed status for all 3-Site remote copy consistency groups that are present in a 3-Site configuration.
- The command **ls3sitercrelationship** lists the detailed status of all 3-Site remote copy relationships that are present in a 3-Site configuration.

> **Note:** The 3-Site Orchestrator service can run only one command at a time, as shown in Example 5-1.

Example 5-1 Orchestrator command error

```
[root@localhost ~]# ls3sitercrelationship
CMMVC9481E 3-Site orchestrator is busy processing another command.
```

If the error CMMVC9481E is shown, repeat the command until it succeeds.

5.1.1 Monitoring the 3-Site Orchestrator configuration

3-Site Orchestrator installs the following services on the host as part of its package deployment. These services are responsible for managing the 3-Site Replication:

- Task Daemon

  Task Daemon service controls the 3-Site Replication on the configured 3-Site setup. This service is responsible for running the data cycle and managing the 3-Site objects. The service is installed as part of a systemd operation and can be listed by using the unit file rc3site.tstaskd.service. The service is enabled by default at start and includes a watchdog timer of 30 seconds if failures occur. The status of the service can be checked by using the **systemctl status operation** command, as shown in Example 5-2.

Example 5-2 Checking the status of Task Daemon service

```
   "# systemctl status rc3site.tstaskd.service
   "? rc3site.tstaskd.service - rc3site tstaskd service. Orchestrator VERSION = 3.0.201209113224
   "Loaded: loaded (/etc/systemd/system/rc3site.tstaskd.service; enabled; vendor preset:
disabled)
   "Active: active (running) since Fri 2021-04-09 08:35:14 BST; 10min ago
   "Process: 8162 ExecStart=/usr/bin/tstaskd start (code=exited, status=0/SUCCESS)
   "Main PID: 8169 (tstaskd)
   "CGroup: /system.slice/rc3site.tstaskd.service
   "??8169 /usr/bin/tstaskd start
   "??8247 ssh -fN -L 8000:localhost:50100 -i /root/.ssh/id_rsa -p 22 tsuser@9.71.48.219 -o
ServerAliveInterval=10 -o ServerAliveCountMax=2
   "??8248 ssh -fN -R 8000:localhost:22 -i /root/.ssh/id_rsa -p 22 tsuser@9.71.48.219
   "??8255 ssh -fN -L 8001:localhost:50100 -i /root/.ssh/id_rsa -p 22 tsuser@9.71.48.154 -o
ServerAliveInterval=10 -o ServerAliveCountMax=2
   "??8256 ssh -fN -R 8001:localhost:22 -i /root/.ssh/id_rsa -p 22 tsuser@9.71.48.154
```

```
   "Apr 09 08:35:14 hostshared006-1 systemd[1]: Starting rc3site tstaskd service. Orchestrator
VERSION = 3.0.201209113224...
   "Apr 09 08:35:14 hostshared006-1 systemd[1]: Started rc3site tstaskd service. Orchestrator
VERSION = 3.0.201209113224.
```

- Event Daemon

  Event Daemon service is used by 3-Site Orchestrator to monitor IBM Storage Virtualize object change notifications and take any necessary actions. This service is responsible for notifying on data cycle completion events for various operations that are performed by Orchestrator.

  The service uses the SSH tunnel that is created to listen for IBM Storage Virtualize events. This service is configured to listen only for events on objects that are managed by the 3-Site Orchestrator. The service is configured by using **systemd** and is enabled by default at start with a watchdog timer of 30 seconds if failures occur. It also uses port 7003 to communicate with the Task Daemon service. The status of the service can be checked by using **systemctl status**, as shown in Example 5-3.

Example 5-3 Check the status of Event Daemon service

```
"# systemctl status rc3site.eventd.service
"? rc3site.eventd.service - rc3site eventd service. Orchestrator VERSION = 3.0.201209113224
"Loaded: loaded (/etc/systemd/system/rc3site.eventd.service; enabled; vendor preset: disabled)
"Active: active (running) since Fri 2021-04-09 08:35:14 BST; 17min ago
"Main PID: 8161 (eventd)
"CGroup: /system.slice/rc3site.eventd.service
"??8161 /usr/bin/eventd 7003

"Apr 09 08:35:14 hostshared006-1 systemd[1]: Started rc3site eventd service. Orchestrator
VERSION = 3.0.201209113224
```

- Miscellaneous Daemon

  Miscellaneous Daemon service is used to capture the audit log for 3-Site configuration commands. This service is installed during the installation of 3-Site Orchestrator and enabled by default at start with a watchdog timer of 30 seconds if failures occur. The status of the service can be checked by using **systemctl status**, as shown in Example 5-4.

Example 5-4 Detailed status of the current 3-Site configuration

```
"# ls3siteconfig
"master_name HSSite
"master_site_name site1
"master_site_status online
"master_ip1 9.71.48.219
"master_ip1_status reachable
"master_ip2
"master_ip2_status not_configured
"master_port 8000
"auxnear_name HSSite
"auxnear_site_name site2
"auxnear_site_status online
"auxnear_ip1 9.71.48.219
"auxnear_ip1_status reachable
"auxnear_ip2
"auxnear_ip2_status not_configured
"auxnear_port 8000
```

```
"auxfar_name FarSite
"auxfar_site_status online
"auxfar_ip1 9.71.48.154
"auxfar_ip1_status reachable
auxfar_ip2
auxfar_ip2_status not_configured
auxfar_port 8001
username tsuser
keyfile /root/.ssh/id_rsa
topology hyperswap
```

As shown in Example 5-4 on page 101, every cluster features only one configured IP address. Users can use the following fields to monitor the status of the 3-Site configuration:

- xxx_ip_status

  This field is populated for every site in a 3-Site configuration. It is used to distinguish the IP connectivity to the designated site from the 3-Site Orchestrator. The following values can be used for the field:

  - reachable. A specified IP address can be pinged.
  - unreachable. A specified IP address cannot be pinged.
  - not_configured. No IP address is configured in Orchestrator.
  - authentication_failed. The SSH keyfile that is used for authentication is invalid.

- xxx_site_status

  This field indicates the designated site status. It is used to distinguish the capability of the designated site to run 3-Site operations from the Orchestrator. The following values can be used for the field:

  - online

    Designated site is available for Orchestrator operations.

  - offline

    Designated site is not available for Orchestrator operation. The user must check whether an `authentication_failed` message exists in the `xxx_ip_status` field of the **`ls3siteconfig`** command-line interface (CLI) command. For T2 recovery, the status field reports as `offline` if the user did not enable the SAN Volume Controller CLI on the cluster.

- topology

  This field is used to indicate the topology provided by the user when the 3-Site configuration is created. This topology defines the behavior of 3-Site Orchestrator. This property cannot be changed for an existing configuration. To change the topology, the user must remove and create a 3-Site configuration.

5.1.2 Monitoring the 3-Site consistency groups

The command **`ls3sitercconsistgrp`** lists the detailed status of all 3-Site remote copy consistency groups that exist in a 3-Site configuration.

Note: The command fails if the 3-Site Orchestrator cannot connect to either of the sites. The user must check the **`ls3siteconfig`** operation to check the connectivity status between Orchestrator and the site that is reporting the connection failure.

Example 5-5 shows the output of the **ls3sitercconsistgrp** command.

Example 5-5 3-Site consistency group information

```
name  state              primary periodic_copy_source  sync_copy_state                 periodic_copy_state freeze_time  cycle_period status
hscg0 3site_consistent   HSSite  site1                 consistent_synchronized idling  2021/04/09/10/09/24 300          online
hscg1 3site_consistent   HSSite  site1                 consistent_synchronized stopped 2020/10/05/26/38 300             online
hscg2 3site_consistent   HSSite  site1                 consistent_synchronized idling  2020/12/10/05/26/43 300          online
hscg3 stopped            HSSite  site1                 consistent_synchronized idling  2020/12/10/05/26/43 300          online
hscg4 stopped            HSSite  site1                 consistent_synchronized idling  2020/12/10/05/26/43 300          online
```

The **ls3sitercconsistgrp** command features the following attributes:

- name

  Specifies the name of a 3-Site remote copy consistency group. The value is a string with a maximum length of 15 characters.

- state

  Specifies the state from the 3-Site Replication state machine:

  - 3site_consistent (3SC)

    Indicates that a consistent copy of data is available on the auxiliary-far site. The auxiliary-far site is within a maximum of two cycle periods of the application and meets the defined recovery point objective (RPO).

  - 3site_inconsistent (3SI)

    Indicates that data on the auxiliary-far site volume is inconsistent with data on the near sites. This issue might be the result of errors, or the first data cycling is still ongoing and did not yet complete.

  - 2site_periodic (2SP)

    Indicates that the non-primary near site is not part of the 3-Site configuration. The primary near site can replicate data to the auxiliary-far site, and a consistent copy of data is available on the auxiliary-far site. The auxiliary-far site is within a maximum of two cycle periods of the application.

  - partial

    The 3-Site Replication for at least one relationship of the consistency group is not fully configured. This issue can occur if an error occurred during the 3-Site configuration, such as the use of an unsuitable type for third site volume, or the near sites were not synchronized, or other errors.

    After the errors are fixed, the 3-Site consistency group configuration command can be repeated to finish the configuration.

  - stopped

    Indicates that 3-Site data cycling is stopped because of a link failure or because of the **stop3sitercconsistgrp** command. This command stops the periodic cycling of a consistency group and this status is shown. When 3-Site data cycling is stopped, the RPO of the auxiliary-far site copy of data is extended by the duration for which 3-Site data cycling is stopped.

- primary

  Specifies the primary site for the 3-Site remote copy consistency group. This value is the site alias or cluster name.

- sync_copy_state

  Specifies the state of the near-site synchronous copy. The following values are available:

  - consistent_synchronized
  - idling_disconnected

- idling
- consistent_copying
- consistent_stopped

- periodic_copy_state

  Specifies the state of the remote site periodic copy. The following values are available:

  - copying

    This state is a normal state during a periodic data cycling. The data from periodic source AccessPoint is transferred to the third site.

  - copied

    This state is a normal state after a periodic data cycling transfers all of the data.

  - stopped

    This state is a normal state after a periodic data cycling ends.

  - idling

    This state is a normal state between copied and stopped; for example, to set the correct freeze_time value.

- freeze_time

  Specifies the last available snapshot at the auxiliary-far site. The value is a string in YY/MM/DD/HH/MM/SS format.

- periodic_copy_source

  Specifies the source site for periodic remote copy site (Master or AuxNear). The value is the site's alias cluster name.

- cycle_period

  Specifies the 3-Site Replication cycle period in seconds in the range 300 - 86400 (1 day).

- status

  Specifies the 3-Site data cycling status for the 3-Site consistency group.

  Note: Status change can take up to one cycle to report change.

  The status online indicates that the 3-Site cycling is working correctly. Every other status indicates an error, which must be fixed. The following values are available:

  - online

    The 3-Site data cycling is working in 3site_consistent mode.

  - sync_link_offline

    The near site link is disconnected. The 3-Site relationship is in the stopped or 2site_periodic state. Fix the link to resume 3site_consistent mode. Alternately, switch to 2site_periodic mode by changing the periodic source site.

  - sync_link_stopped

    The near site link is reconnected, but it is in a stopped state. The 3-Site relationship is in a stopped or 2site_periodic state. Rejoin a non-primary near site to resume 3site_consistent mode.

  - active_link_offline

    This value indicates that the active link is down and 3-Site data cycling is in a stopped state. To resume 3site_consistent mode, fix the link manually or move the periodic source site to continue in degraded mode.

– inactive_link_offline

  This value indicates that an idle link is down. It does not affect the 3-Site data cycling. The user must manually fix the inactive link.

– primary_offline

  This value indicates that the primary site for this consistency group is down or is isolated from the partnership. (A network is not available and the site cannot be reached from the remaining two sites).

– periodic_src_offline

  This value indicates that the periodic site for this consistency group is down or is isolated from the partnership. (No network is available and the site cannot be reached from the remaining two sites).

– primary_and_periodic_src_offline

  This value indicates that the primary and periodic source sites for this consistency group are down or isolated from the partnership. The 3-Site consistency group is in the stopped state. Data cycling can be started again in 2site_periodic mode by changing the primary and periodic source sites.

– partner_offline

  This value indicates that the near partner site, which is neither a primary nor periodic source, is offline. Data cycling is in 2site_periodic mode and the user must fix the near partner site.

– auxfar_offline

  This value indicates that the auxiliary-far site (which is not a primary nor periodic source) is offline. Data cycling is stopped. User must fix the auxiliary-far site. Data cycling resumes automatically.

– primary_storage_offline

  The 3-Site relationship is in the stopped or 2site_periodic state. The primary site storage is offline. The user must fix the storage problems to change to 3site_consistent mode or change the primary to 2site_periodic mode.

– periodic_src_storage_offline

  The 3-Site relationship is in the stopped or 2site_periodic state. The periodic site storage is offline. The user must fix the storage problems to change to 3site_consistent mode or change the periodic source to 2site_periodic mode.

– partner_storage_offline

  This value indicates that at the near partner site storage is offline, which is neither a primary nor a periodic source. Data cycling is in 2site_periodic mode. The user must fix the master site.

– auxfar_storage_offline

  This value indicates that storage is offline at the auxiliary-far site, which is neither a primary nor a periodic source. Data cycling is stopped. The user must fix the auxiliary-near site. Data cycling resumes automatically.

– site_unreachable

  This value indicates one or more sites lost connection with 3-Site Orchestrator. The user must fix the network connection. Data cycling resumes automatically.

– maintenance_exclusion_mode

  This value indicates that a site is excluded for maintenance and its state is 2site_periodic or stopped.

– sync_link_not_synchronized

  Indicates that data on the synchronous link is not in sync. This issue can occur if the application is writing data at a higher rate than the replication link between near sites.

– hyperswap_and_auxfar_partnership_offline

  Indicates that the partnership between the HyperSwap system and the auxiliary-far system is offline.

– hyperswap_cluster_offline

  Indicates that the HyperSwap system is offline (applicable only for HyperSwap topology).

– partner_dead_domain

  Applicable for only HyperSwap topology. Indicates that one of the near sites in the HyperSwap system encountered a dead domain.

– master_and_auxnear_storage_offline

  Indicates that storage on the master site and auxiliary-near site is offline.

5.2 Monitoring the 3-Site remote objects by using IBM Storage Virtualize GUI

3-Site Orchestrator version 3.0 and newer supports integrating IBM Storage Virtualize GUI with the external Orchestrator. Users that are running Orchestrator version 3.0 or later can check the status of the 3-Site objects by using the GUI that is available on any of the sites in the 3-Site configuration.

Note: An SSH public key must be added on the host for the GUI to communicate with the external Orchestrator.

Complete the following steps:

1. In the GUI, click **Storage Virtualize GUI** → **Copy Service** → **Remote Copy** → **3 site partnership** → **3-Site partnership properties** to view the 3-Site configuration. See Figure 5-1.

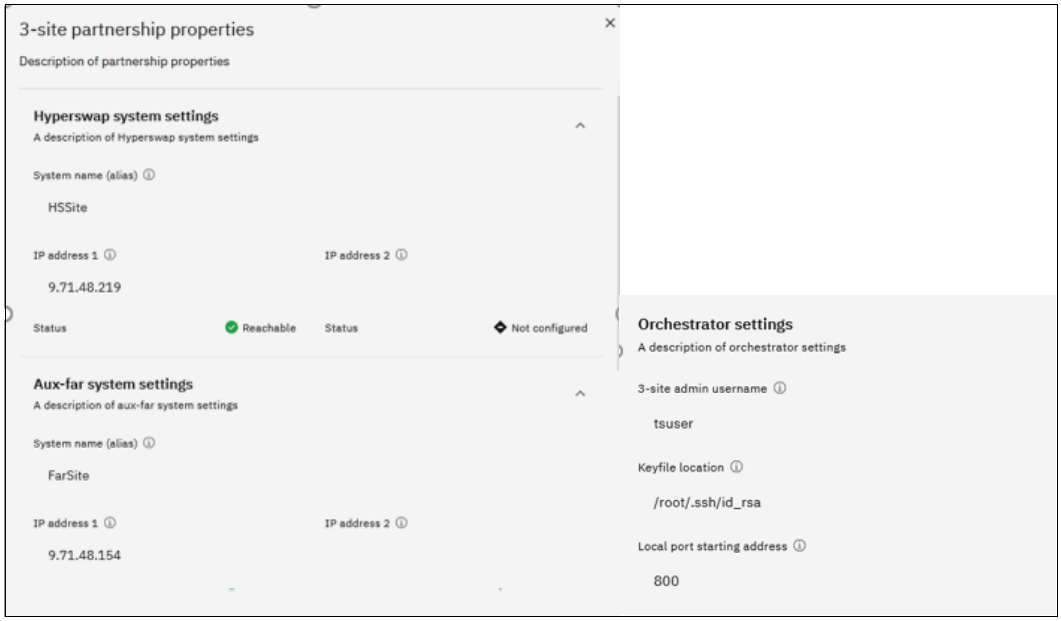

Figure 5-1 3-Site partnership properties window

2. View the 3-Site consistency groups by clicking **Storage Virtualize GUI** → **Copy Service** → **Remote Copy** → **3-Site Partnership**.

   Figure 5-2 shows the 3-Site consistency groups that are configured in the system. View the relationship by clicking the 3-Site consistency group.

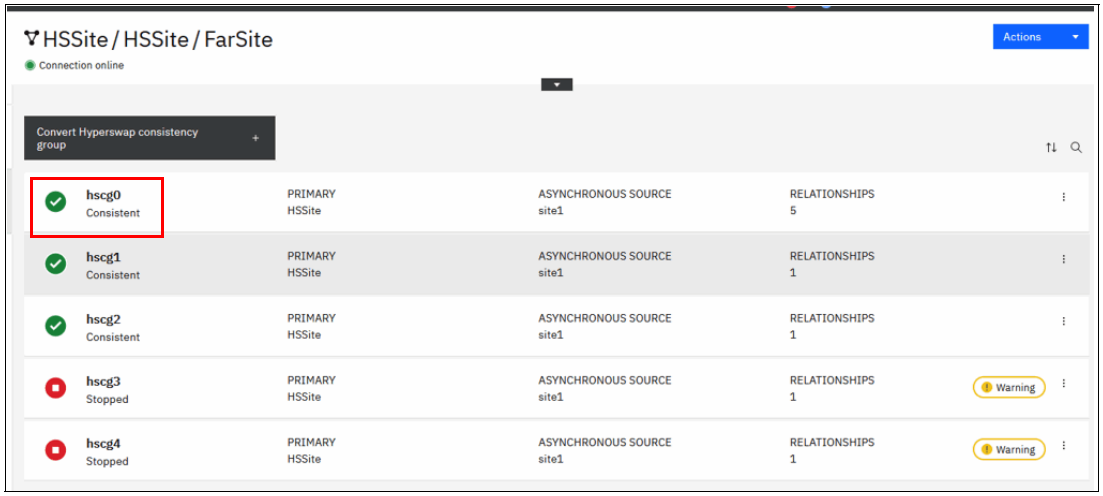

Figure 5-2 Configured 3-Site consistency groups

The list of relationships is shown for the consistency group. See Figure 5-3.

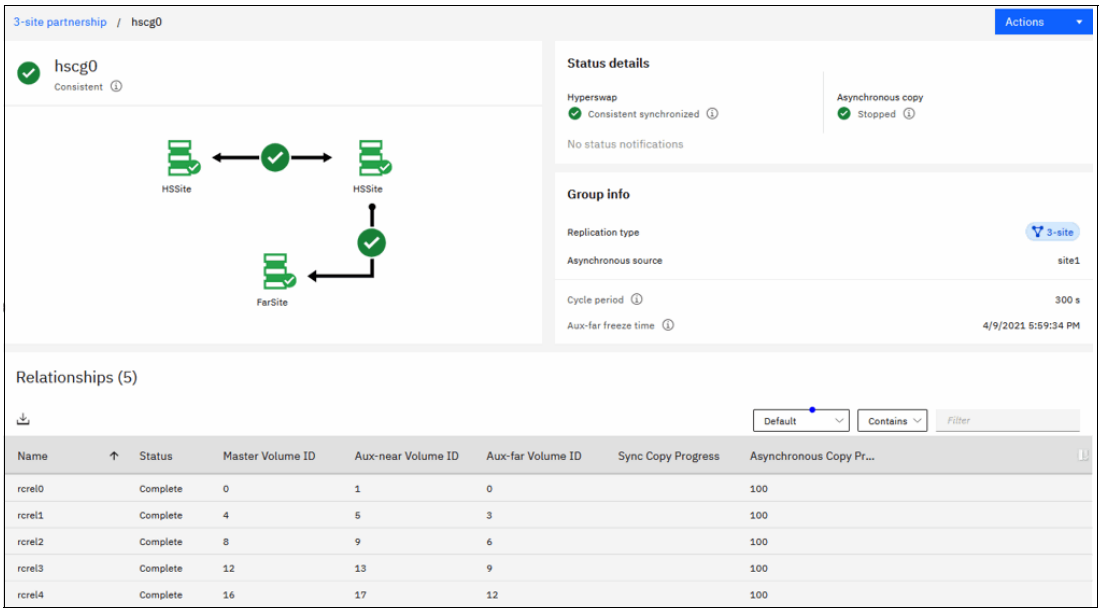

Figure 5-3 3-Site consistency group relationships

5.3 Orchestrator log files

The Orchestrator service that is running on the external host collects all logs on the host at the path `/opt/ibm/SpectrumVirtualize/rc3site/logs/`. The logs are backed up daily and these backups are kept for 15 days before being rotated. The following log files are generated by the Orchestrator:

- tstaskd.log

  This log is generated by the Task Daemon service and holds logs for all consistency groups and their data cycle. The log file contains the timestamp, thread-id, and log level information with log text. The log file can be sorted by using thread-id for checking operation on a specific cluster group. A sample log output is shown in Example 5-6.

Example 5-6 tstaskd.log

```
2021-04-09 09:35:50,244 | 140709816436544 | INFO | starting task daemon
2021-04-09 09:35:50,244 | 140709816436544 | INFO | Orchestrator version: 3.0.201209113224
2021-04-09 09:35:50,318 | 140709816436544 | INFO | Initiating Orchestrator Recovery..
```

- tsdiagnostic.log

  This log is generated by the Event Daemon Service and holds logs for events that are received and processed from Storage Virtualize. The log file contains the timestamp, thread-id, and log level information with log text. A sample log output is shown in Example 5-7.

Example 5-7 tsdiagnostic.log

```
2021-04-09 09:35:50,317 | 140709816436544 | INFO | Trace ['2021-04-09 09:35:50.317448', 1,
'GEN_TRACE_HISTORY', "('Connected to Eventd on', '127.0.0.1', ':', 7003)"]
2021-04-09 09:35:52,982 | 140708287604480 | INFO | Trace ['2021-04-09 09:35:52.982047', 1,
'DM_TO_EVENTD_HISTORY', 'Command:6,Topology:1']
```

```
2021-04-09 09:35:52,990 | 140708287604480 | INFO | Trace ['2021-04-09 09:35:52.990825', 1,
'GEN_TRACE_HISTORY', "('RELOGIN_SITE ', 0, '9.71.48.219')"]
```

- all_attempted_tscmd_audit

  This log is generated by Miscellaneous service and holds audit log information for all management operations that are performed by the user by using the Orchestrator CLI. A sample log output is shown in Example 5-8.

Example 5-8 .all_attempted_tscmd_audit

```
Histlog Entry Start
        Sequence_Num        : 2
        Src_Timestamp       : Wed Oct 14 13:46:28 2020
        Dest_Timestamp      : Wed Oct 14 13:46:28 2020
        Cgname              :
        Action_Cmd          : convertrcrelationship -type 3site -periodicsrc site1 -cycleperiod
300 -pool 0 -iogrp 0 -gui rcrel0
        Response            : 3-Site RC relationship created with volume id [0].
Histlog Entry End
```

Note: Timestamps at Orchestrator and IBM Storage Virtualize can differ if the time zones are not in sync. Therefore, it is a good practice to use Network Time Protocol to sync the timezone between all sites and the Orchestrator host.

Orchestrator event mechanism

The 3-Site Orchestrator reports five events on the IBM Storage Virtualize system that are used to manage or notify the user of changes in the 3-Site object state. The events are logged by the Task Daemon service when it detects a state change in any of the 3-Site consistency groups or if it detects a problem with configuration.

The events that are logged are available at every consistency group level. The events are logged on the auxiliary-far site by Orchestrator and can be viewed in the Event log. However, if the auxiliary-far site is not reachable to report the error, the same event is logged on the Master site in a 3-Site configuration.

If any of the site loses its connection with the Orchestrator, an event about the lost connection is reported on the site that lost the connection, as shown in the following example:

```
EventID: 050992 3-Site data cycling is stopped because SVC cluster has lost
connection with orchestrator
```

This event is the only event that is not logged at the consistency group level. Instead, it is logged at site level.

Table 5-1 lists the possible event IDs and their event text.

Table 5-1 Event status table

ID	Event	Type
050988	Configuration change requested by the last command completed successfully.	Informational
050990	Configuration change requested by the last command failed.	Warning
050991	The cycle time for 3-Site consistency group exceeded the cycle period that was set by the user.	Warning
050992	3-Site data cycling is stopped because the IBM SAN Volume Controller cluster lost its connection with 3-Site Orchestrator.	Warning
050993	3-Site consistency group is now in the stopped state.	Warning
050994	3-Site consistency group is now in 2-Site data cycling mode.	Warning

Note: Event IDs 050992, 050993, and 050994 include dump objects that are associated with them that guide the user to resolve the problem.

Figure 5-4 shows a sample event that was logged on the auxiliary-far site.

Error Code	Last Time Stamp	Status	Description	Object Type	Object ID	Object Name
1206	12/9/2020 6:15:17 PM	Alert	Drive format incorrect	drive	10	
1674	12/9/2020 6:15:17 PM	Alert	Drive unlock required	drive	10	
1970	12/10/2020 10:01:52 AM	Alert	3-site data cycling for a 3-site consistency group has stopped due to an ex...	3site_rc_consist_...	4	af_hscg4
1970	12/10/2020 10:01:47 AM	Alert	3-site data cycling for a 3-site consistency group has stopped due to an ex...	3site_rc_consist_...	3	af_hscg3

Figure 5-4 Sample event logged on the auxiliary-far site

You can click the event to open the details window. See Figure 5-5.

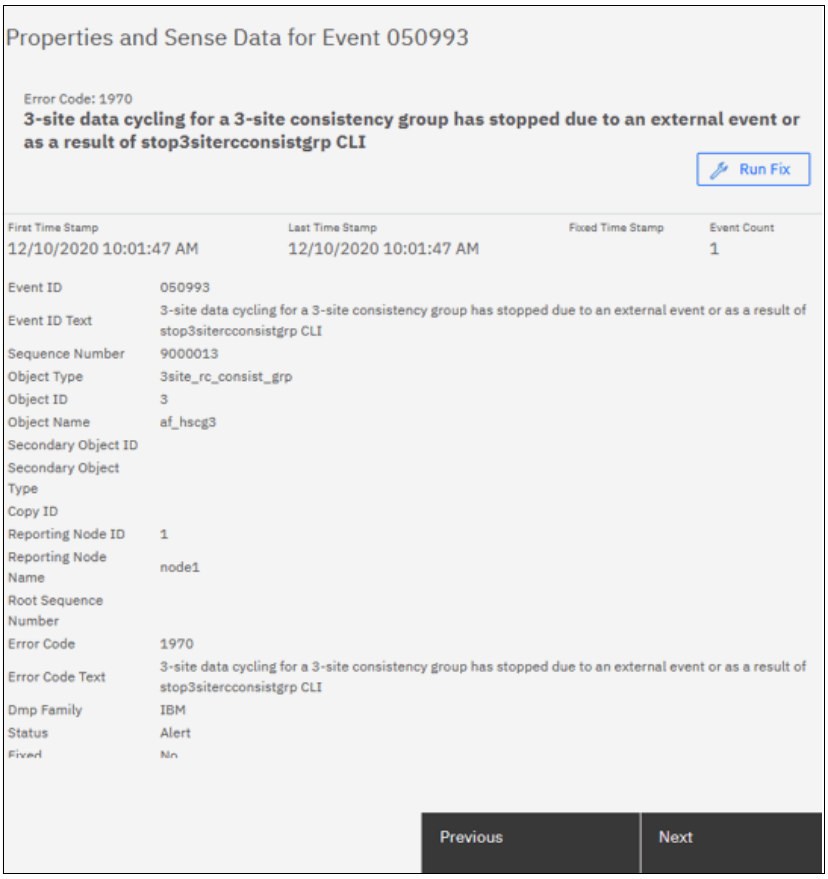

Figure 5-5 Event 050993

5.4 Maintenance

This section describes planned maintenance tasks. 3-Site Orchestrator provides the following commands that are used to alter the configuration of the Orchestrator and the 3-Site consistency groups:

- **`ch3siteconfig`**

  This command changes the current Orchestrator configuration.

- **`ch3sitercconsistgrp`**

  This command changes a specified 3-Site remote copy consistency group.

5.4.1 Changing the Orchestrator configuration

To change the Orchestrator configuration, complete the procedure that is described in this section. Run the **`ch3sitercconsistgrp`** *<attribute> <parameter>* command to change the Orchestrator configuration.

Note: The **`ch3sitercconsistgrp`** command fail except for the attribute **`-updatesystemname`** if all consistency groups are not in a stopped state.

Example 5-9 shows examples of changing the alias name of the master site and the user name.

Example 5-9 Changing the alias name of the master site and the user name

```
# # rename the master_name site alias
[root@localhost ~]# ch3siteconfig -site Master -sitealias Master_New
# # change the 3-Site username and keyfile
[root@localhost ~]# ch3siteconfig -username 3SiteAdmin
# # list the current configuration
```

Table 5-2 lists the available attributes and parameters.

Table 5-2 ch3siteconfig attributes and parameters

Attribute	Parameter	Description
-keyfile	*ssh_keyfile_path*	(Optional) Specifies the updated path for the private SSH key file for the 3-Site user on the host. This parameter is mutually exclusive with -site.
-username	*3site_username*	(Optional) Specifies the updated 3Ssite username for all three sites. This parameter is mutually exclusive with -site.
-port	local_port_offset	(Optional) Specifies the updated starting address for the three local ports that are used for connection with the three sites over the SSH tunnel.
-site	site_name	(Optional) Specifies the site to update.
-ip1	*site_ip1*	(Optional) Specifies the new IPv4 address for IP address 1. This parameter must be used with -site and is exclusive with -ip2.
-ip2	*site_ip2*	(Opti -sonal) Specifies the new IPv4 address for IP address21. This parameter must be used with -site and is mutually exclusive with -ip1.
-sitealias	*new_site_alias*	(Optional) Specifies the updated site alias name. This parameter must be used with -site.
-updatesystemname		(Optional) Specify this parameter to refresh the system name change of any of the three sites

Note: The **ch3siteconfig** operation is not supported from Storage Virtualize GUI.

5.4.2 Changing attributes of 3-Site consistency group

By using 3-Site Orchestrator, you can change the different attributes of a consistency group by using the **ch3sitercconsistgrp** CLI command. The operation is an asynchronous operation and logs events (ID 050988 and ID 050990) on the auxiliary-far site, as described in 5.4, “Maintenance” on page 111. These maintenance operations can also be completed by way of IBM Storage Virtualize GUI with Orchestrator version 3.0 or newer. The user can perform the following maintenance operations on a consistency group:

- Change:
  - Cycle period
  - Periodic replication source
  - Primary (applicable for Metro Mirror topology only)
- Exclude and join a site (applicable for Metro Mirror topology only)

Changing a cycle period

Run the **ch3sitercconsistgrp** CLI to change the cycle period for a specific consistency group. The value of the cycle period can be 300–86400 seconds (1 day). Example 5-10 shows changing the cycle period by using the **ch3sitercconsistgrp** command.

Example 5-10 Changing cycle period by using ch3sitercconsistgrp command

```
## Changing cycle period for hscg0 from 300 to 500
# ls3sitercconsistgrp
name  state            primary periodic_copy_source sync_copy_state         periodic_copy_state freeze_time         cycle_period status
hscg0 3site_consistent HSSite  site1                consistent_synchronized stopped             2021/04/09/19/24/44 300          online
hscg1 3site_consistent HSSite  site1                consistent_synchronized stopped             2021/04/09/19/24/59 300          online
hscg2 3site_consistent HSSite  site1                consistent_synchronized stopped             2021/04/09/19/24/59 300          online
hscg3 stopped          HSSite  site1                consistent_synchronized idling              2020/12/10/05/26/43 300          online
hscg4 stopped          HSSite  site1                consistent_synchronized idling              2020/12/10/05/26/43 300          online
## Changing value to 500
# ch3sitercconsistgrp -cycleperiod 500 hscg0
# ls3sitercconsistgrp
name  state            primary periodic_copy_source sync_copy_state         periodic_copy_state freeze_time         cycle_period status
hscg0 3site_consistent HSSite  site1                consistent_synchronized stopped             2021/04/09/19/28/29 500          online
hscg1 3site_consistent HSSite  site1                consistent_synchronized stopped             2021/04/09/19/24/59 300          online
hscg2 3site_consistent HSSite  site1                consistent_synchronized stopped             2021/04/09/19/24/59 300          online
hscg3 stopped          HSSite  site1                consistent_synchronized idling              2020/12/10/05/26/43 300          online
hscg4 stopped          HSSite  site1                consistent_synchronized idling              2020/12/10/05/26/43 300          online
```

The same operation can be done by using the GUI window. Browse to **Storage Virtualize GUI** → **Copy Service** → **Remote Copy** → **3 site partnership** → **ConsistencyGrp** → **Properties**, as shown in Figure 5-6.

Figure 5-6 Properties window

Event ID 050988 is posted on the auxiliary-far cluster after a successful execution. See Figure 5-7.

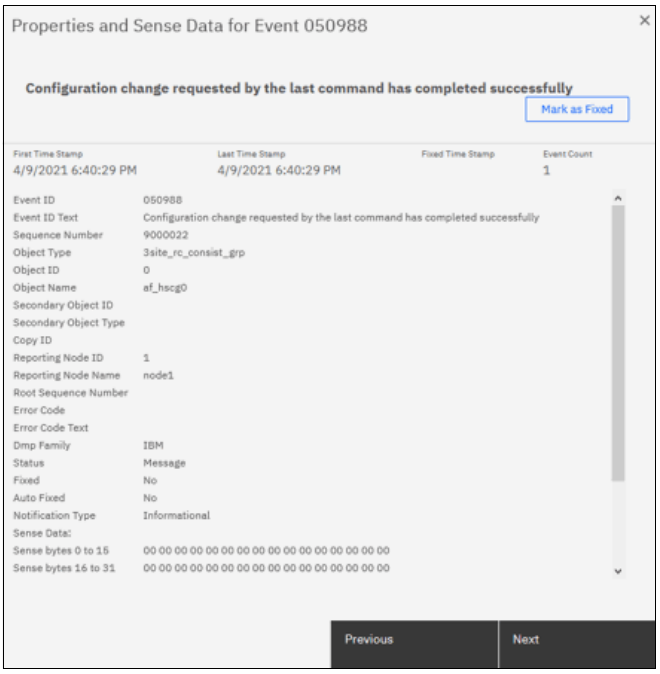

Figure 5-7 Event ID 050988 is logged

5.4.3 Changing the periodic replication source

Run the **ch3sitercconsistgrp** CLI command to change the periodic replication source for a specific consistency group. The periodic source value can be only the *site_alias* or *clustername* of either of the near sites. The operation is asynchronous and might take time to complete. The user must check the event (ID 050988) on the auxiliary-far site for successful completion of the operation.

Note: The auxiliary-far site cannot be set as the periodic replication source for a particular consistency group.

Example 5-11 shows the method that is used to change periodic copy source.

Example 5-11 Changing periodic copy source

```
20210409-14:21:19 root@hostshared006-1:/ # ls3sitercconsistgrp
name  state           primary periodic_copy_source sync_copy_state         periodic_copy_state freeze_time         cycle_period status
hscg0 3site_consistent HSSite  site1                consistent_synchronized stopped             2021/04/09/19/36/49 500          online
hscg1 3site_consistent HSSite  site1                consistent_synchronized stopped             2021/04/09/19/34/59 300          online
hscg2 3site_consistent HSSite  site1                consistent_synchronized stopped             2021/04/09/19/34/59 300          online
hscg3 stopped          HSSite  site1                consistent_synchronized idling              2020/12/10/05/26/43 300          online
hscg4 stopped          HSSite  site1                consistent_synchronized idling              2020/12/10/05/26/43 300          online
20210409-14:21:24 root@hostshared006-1:/ # ch3sitercconsistgrp -periodicsrc site2 hscg2
20210409-14:21:52 root@hostshared006-1:/ # ls3sitercconsistgrp
name  state           primary periodic_copy_source sync_copy_state         periodic_copy_state freeze_time         cycle_period status
hscg0 3site_consistent HSSite  site1                consistent_synchronized stopped             2021/04/09/19/36/49 500          online
hscg1 3site_consistent HSSite  site1                consistent_synchronized stopped             2021/04/09/19/34/59 300          online
hscg2 3site_consistent HSSite  site2                consistent_synchronized idling              2021/04/09/19/34/59 300          online
hscg3 stopped          HSSite  site1                consistent_synchronized idling              2020/12/10/05/26/43 300          online
hscg4 stopped          HSSite  site1                consistent_synchronized idling              2020/12/10/05/26/43 300          online
```

Figure 5-8 shows the event on the auxiliary-far site.

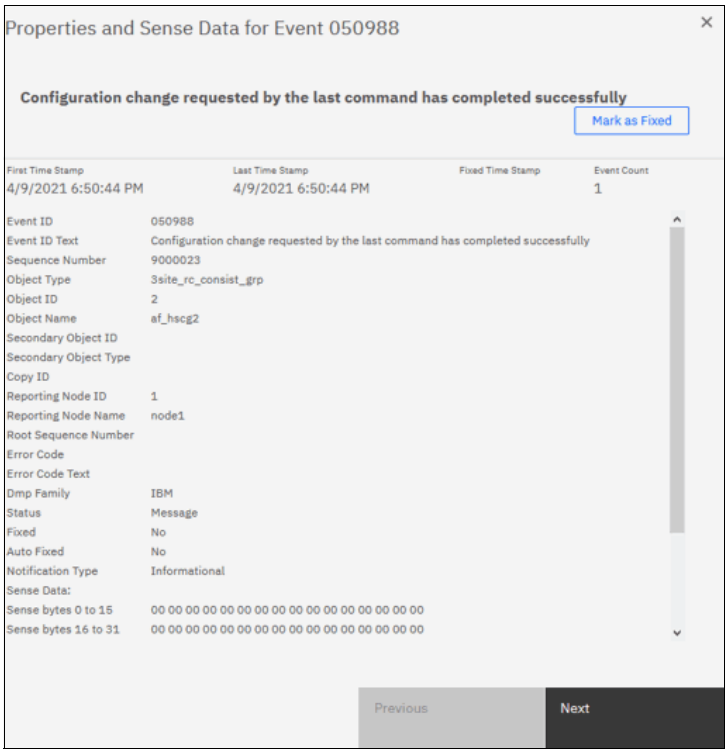

Figure 5-8 Event on the auxiliary-far site

The periodic replication source also can be changed by using the IBM Storage Virtualize GUI with Orchestrator version 3.0 or newer.

Select **Storage Virtualize GUI** → **Copy Service** → **Remote Copy** → **3 site partnership** → **ConsistencyGrp** → **Actions** → **Switch asynchronous source** and then select the wanted configuration (see Figure 5-9).

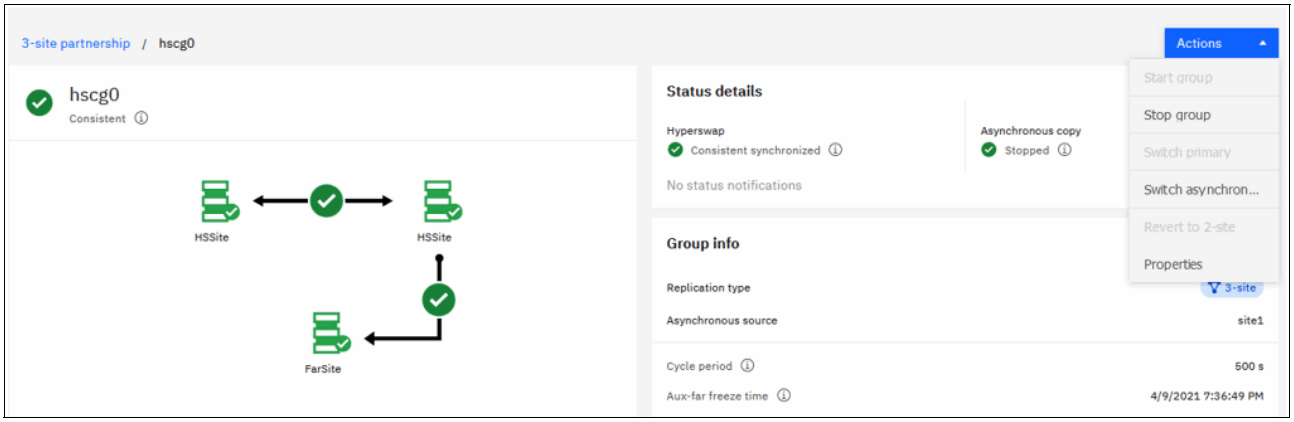

Figure 5-9 Changing periodic replication source by way of the GUI

5.4.4 Changing primary (applicable for Metro Mirror configuration only)

Run the **ch3sitercconsistgrp** CLI command to change the primary setting or value for a specific consistency group. The primary value can be only the `site_alias` or `clustername` of either of the near sites. The operation is asynchronous and might take time to complete.

The user must stop the application before switching the primary. The user must also check the event (ID 050988) on the auxiliary-far site to verify that the operation was successfully completed.

Note: The auxiliary-far site cannot be set as the primary for a specific consistency group.

Example 5-12 shows the method that was used to change primary.

Example 5-12 Changing primary

```
[root@localhost ~]# stop3sitercconsistgrp CG_0
[root@localhost ~]# ls3sitercconsistgrp
name state primary periodic_copy_source sync_copy_state
CG_0 stopped Master AuxNear consistent_synchronized
[root@localhost ~]# ch3sitercconsistgrp -primary AuxNear CG_0
[root@localhost ~]# start3sitercconsistgrp CG_0
[root@localhost ~]# ls3sitercconsistgrp
name state primary periodic_copy_source sync_copy_state
CG_0 3site_consistent AuxNear AuxNear consistent_synchronized
```

Event on the Auxiliary Far Site (by using **lseventlog**) shows the successful completion of the operation (see Example 5-13).

Example 5-13 Event on the auxiliary-far site

```
#lseventlog | tail -1
9000029 200518124045 3site_rc_consist_grp 0 CG_0 message no 050988
Configuration change requested by the last command has completed successfully
```

Changing the periodic replication source also can be performed by using IBM Storage Virtualize GUI with Orchestrator version 3.0 or newer. Click **Storage Virtualize GUI** → **Copy Service** → **Remote Copy** → **3 site partnership** → **ConsistencyGrp** → **Actions** → **Switch primary** and then select the wanted configuration (see Figure 5-10).

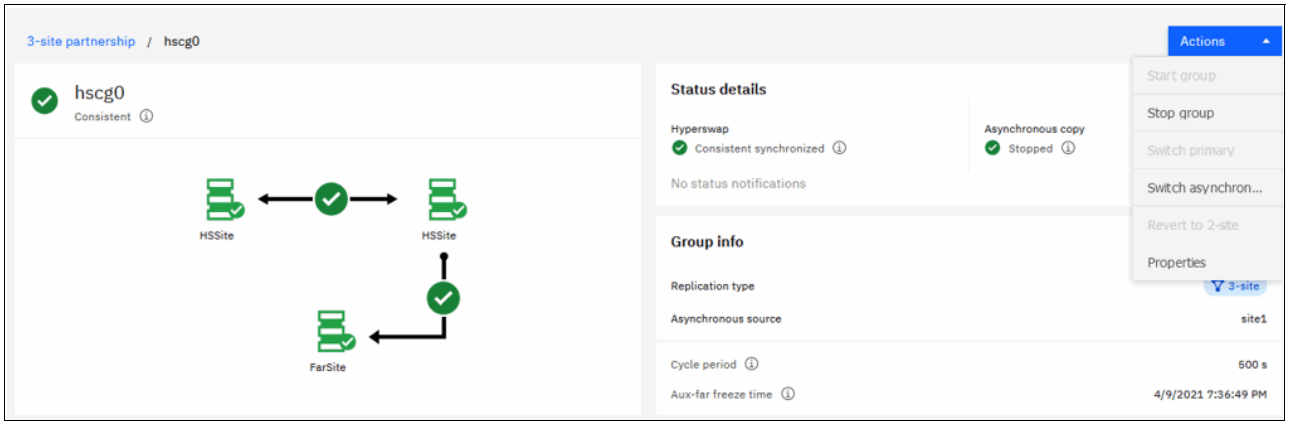

Figure 5-10 Changing primary by way of the GUI

5.4.5 Excluding and joining a site

Consider the following prerequisites when planning for a maintenance task (for example, a service action that takes one site offline):

- To exclude a site for maintenance, the site must not be the primary or periodic-source site.
- Exclusion of the site from data cycling stops the Metro Mirror consistency group between the two near sites.
- Every consistency group must be excluded.

The primary and periodic-source site is still replicating the data to the auxiliary-far site if the other near site is excluded from the data cycling. This configuration assures an asynchronous data copy when one near site is excluded.

To exclude a site from 3-Site Replication, complete the task that is described next for every 3-Site consistency group. Issue the **ch3sitercconsistgrp -exclude** command to exclude a consistency group, as shown in Example 5-14.

Example 5-14 To exclude a site from 3-Site Replication

```
[root@localhost ~]# ls3sitercconsistgrp
name state primary sync_copy_state periodic_copy_state status
CG_0 3site_consistent AuxNear consistent_synchronized stopped online
[root@localhost ~]# ch3sitercconsistgrp -exclude Master CG_0
[root@localhost ~]# ls3sitercconsistgrp
name state primary sync_copy_state periodic_copy_state status
CG_0 2site_periodic AuxNear consistent_stopped stopped sync_link_stopped
```

The state of the consistency group is now `2site_periodic` and the link status between the two near sites is stopped after the **ch3sitercconsistgrp -exclude** command successfully completes. The command notification event is shown Example 5-15.

Example 5-15 Command notification event

```
AuxFar >lseventlog | tail -1
9000031 200518125948 3site_rc_consist_grp 0 CG_0 alert no 050994
3-Site consistency group is now in 2-Site data cycling mode
```

The excluded site can be included after the site maintenance by using the **ch3sitercconsistgrp -join** command and parameter, as shown in Example 5-16.

Example 5-16 To include the site that was excluded

```
[root@localhost ~]# ch3sitercconsistgrp -join Master CG_0
[root@localhost ~]#
name state primary sync_copy_state periodic_copy_state status
CG_0 3site_consistent AuxNear consistent_synchronized stopped online
```

Note: The status field can take up to 1 cycle period time to reflect the correct state.

5.5 Upgrade support

3-Site Replication includes the following software components that must be upgraded:

- 3-Site Orchestrator package on host
- Storage Virtualize software on nodes

These components are described next.

5.5.1 3-Site Orchestrator package on host

Orchestrator RPM supports upgrade features with Orchestrator 3.0 and newer. Older versions of Orchestrator can be upgraded by using the **rpm upgrade** command. Example 5-17 shows the upgrade of Orchestrator RPM.

Example 5-17 rpm upgrade command

```
# rpm -Uvh /home/ibm-SpectrumVirtualize-rc3site-orchestrator-2.0.200303095938-1.el7_3.x86_64.rpm
   Preparing...                          ################################# [100%]
   Updating / installing...
   1:ibm-SpectrumVirtualize-rc3site-or################################# [ 50%]
   Cleaning up / removing...
   2:ibm-SpectrumVirtualize-rc3site-or################################# [100%]
```

5.5.2 Upgrading IBM Storage Virtualize cluster software

Before upgrading the storage system software in a 3-Site configuration of Storage Virtualize clusters, consider the following prerequisites:

- Only one Storage Virtualize cluster is upgraded at a time.
- Replication to the auxiliary-far site is stopped for all 3-Site consistency groups.

To upgrade the storage system, data cycling to the auxiliary-site must be stopped by using the **stop3sitercconsistgrp** command for every 3-Site consistency group and restarted after the upgrade. See Example 5-18.

Example 5-18 stop3sitercconsistgrp command

```
[root@localhost ~]# stop3sitercconsistgrp CG_0
[root@localhost ~]# ls3sitercconsistgrp
name state primary periodic_copy_source sync_copy_state
CG_0 stopped Master Master consistent_synchronized
[root@localhost ~]# # Upgrade of a spectrum virtualize storage system
[root@localhost ~]# start3sitercconsistgrp CG_0
[root@localhost ~]# ls3sitercconsistgrp
name state primary periodic_copy_source sync_copy_state
CG_0 3site_consistent Master Master consistent_synchronized
```

The procedure that is shown in Example 5-18 must be repeated for every cluster. This process assures replication to the auxiliary-site after every cluster software upgrade. It also minimizes the lack of the RPO that is caused by stopping the 3-Site Replication during a storage system upgrade.

5.6 Using the CLI to manage 3-Site HA services

Using the CLI, the active and standby Orchestrator can be managed.

Note: Configuration and management commands in standby Orchestrator require root privileges. This feature is available for both the Metro Mirror and the HyperSwap configurations.

To run Orchestrator commands from Spectrum Virtualize CLI, a prefix, **svcorchestratorinfo**, command must be used.

5.6.1 Listing 3-Site objects with CLI

The following Orchestrator commands are available with **svcorchestratorinfo** prefix:

- **ls3siteconfig**

The following Orchestrator commands are available on Orchestrator Host

- **ch3siteconfig**

The **ls3siteconfig** command now includes new attributes that administrators can use to verify the status of the standby Orchestrator in the 3-Site standby Orchestrator feature. See Table 5-3.

Table 5-3 ls3siteconfig command

Command	Description
ls3siterconfig	Validates the status of standby Orchestrator.
	New attribute `Orchestrator_ha_status` allows administrators to verify the state of standby Orchestrator. The following are the possible states: ► empty. If not configured. ► online. If configured and working without any issues. ► offline. If configured and HA service is not running or is at fault. ► standby_offline. If standby Orchestrator is not responding. ► disconnected. If communication failed with the auxiliary-far site.

Table 5-4 on page 120 shows the new attributes, functionality, and description of the **ch3siteconfig** command.

Table 5-4 The ch3siteconfig command new attributes functionality and description

Command	Description
ch3siteconfig -startha	Start HA services and set the role as Active orchestrator
ch3siteconfig -startha -standby -config *<path of config.tar>*	Configure standby orchestrator and start HA services to the auxiliary-far site and set role as Standby
ch3siteconfig -stopha -force	Stop HA service
ch3siteconfig -standby -force	Switching of orchestrator roles Active to Standby Vice versa
ch3siteconfig -update -config *<path of config.tar>*	Update config file on standby orchestrator

5.6.2 Switching from active Orchestrator to standby Orchestrator

To switch the instance from active Orchestrator to standby Orchestrator, perform the following steps.

1. View the status of active Orchestrator and standby Orchestrator by using the **ls3siteconfig** command, as shown in Example 5-19.

Example 5-19 View the status of active Orchestrator and standby Orchestrator

```
[root@localhost home]# ls3siteconfig
master_name Master
master_site_name
master_site_status online
master_ip1 x.xxx.xxx.xx
master_ip1_status reachable
master_ip2
master_ip2_status not_configured
master_port 17160
auxnear_name Aux_Near
auxnear_site_name
auxnear_site_status online
auxnear_ip1 x.xxx.xxx.xx
auxnear_ip1_status reachable
auxnear_ip2
auxnear_ip2_status not_configured
auxnear_port 17161
auxfar_name Auxfar_MM
auxfar_site_status online
auxfar_ip1 x.xxx.xxx.xxx
auxfar_ip1_status reachable
auxfar_ip2
auxfar_ip2_status not_configured
auxfar_port 17162
username tuser
keyfile /root/.ssh/pw_rsa
topology metromirror
Orchestrator_version 4.1.220922160427
Orchestrator_ha_status online
```

2. Ensure that both Orchestrators are up and running.

   To change standby Orchestrator to active Orchestrator and active Orchestrator to standby Orchestrator, enter the following command on active Orchestrator:

   ```
   ch3siteconfig -standby -force
   ```

Note: The command might take several minutes to change the standby Orchestrator to active and active to standby.

3. Validate the configuration by using the **ls3siteconfig** command. See Example 5-20.

Example 5-20 Validate the configuration

```
[root@localhost home]# ls3siteconfig
master_name Master
master_site_name
master_site_status online
master_ip1 x.xxx.xxx.xx
master_ip1_status reachable
master_ip2
master_ip2_status not_configured
master_port 17160
auxnear_name Aux_Near
auxnear_site_name
auxnear_site_status online
auxnear_ip1 x.xxx.xxx.xx
auxnear_ip1_status reachable
auxnear_ip2
auxnear_ip2_status not_configured
auxnear_port 17161
auxfar_name Auxfar_MM
auxfar_site_status online
auxfar_ip1 x.xxx.xxx.xxx
auxfar_ip1_status reachable
auxfar_ip2
auxfar_ip2_status not_configured
auxfar_port 17162
username tuser
keyfile /root/.ssh/pw_rsa
topology metromirror
Orchestrator_version 4.1.220922160427
Orchestrator_ha_status online
```

5.6.3 Updating configuration on standby

If the configuration on active Orchestrator is updated, the same configuration must be updated on standby Orchestrator by using the CLI.

To update configuration of IBM Spectrum Virtualize 3-Site Orchestrator and synchronize it to standby Orchestrator, follow these steps.

1. To change the port number on active Orchestrator:

   a. Validate the configuration on active Orchestrator by using the **ls3siteconfig** command. See Example 5-21 on page 122.

Example 5-21 Validate the configuration on active Orchestrator

```
[root@localhost home]# ls3siteconfig
master_name Master
master_site_name
master_site_status online
master_ip1 x.xxx.xxx.xx
master_ip1_status reachable
master_ip2
master_ip2_status not_configured
master_port 6000
auxnear_name Aux_Near
auxnear_site_name
auxnear_site_status online
auxnear_ip1 x.xxx.xxx.xx
auxnear_ip1_status reachable
auxnear_ip2
auxnear_ip2_status not_configured
auxnear_port 6001
auxfar_name Aux_Far
auxfar_site_status online
auxfar_ip1 x.xxx.xxx.xxx
auxfar_ip1_status reachable
auxfar_ip2
auxfar_ip2_status not_configured
auxfar_port 6002
username tuser
keyfile /root/.ssh/pw_rsa
topology metromirror
Orchestrator_version 4.1.220922160427
Orchestrator_ha_status online
```

b. Change the port number by using the **ch3siteconfig** command:

```
ch3siteconfig -port 7000
Port number is changed successfully.
```

c. Validate the configuration using the **ls3siteconfig** command. See Example 5-22.

Example 5-22 Validate the configuration

```
[root@localhost home]# ls3siteconfig
master_name Master
master_site_name
master_site_status online
master_ip1 x.xxx.xxx.xx
master_ip1_status reachable
master_ip2
master_ip2_status not_configured
master_port 7000
auxnear_name Aux_Near
auxnear_site_name
auxnear_site_status online
auxnear_ip1 x.xxx.xxx.xx
auxnear_ip1_status reachable
auxnear_ip2
auxnear_ip2_status not_configured
```

```
auxnear_port 7001
auxfar_name Aux_Far
auxfar_site_status online
auxfar_ip1 x.xxx.xxx.xxx
auxfar_ip1_status reachable
auxfar_ip2
auxfar_ip2_status not_configured
auxfar_port 7002
username tuser
keyfile /root/.ssh/pw_rsa
topology metromirror
Orchestrator_version 4.1.220922160427
Orchestrator_ha_status online
```

2. Copy the updated `config.tar` file at location `/opt/ibm/SpectrumVirtualize/rc3site/config.tar` from active Orchestrator to standby Orchestrator.

   ```
   [root@rhel7host108 home]# scp -r
   /opt/ibm/SpectrumVirtualize/rc3site/config.tar root@x.xxx.xx.xx:/home/
   root@x.xxx.xx.xx's password:
   config.tar100%10KB5.9MB/s00:00
   ```

3. Update the configuration file on standby Orchestrator, by entering the following command on standby Orchestrator:

   ```
   ch3siteconfig -update -config /home/config.tar
   Configuration file is successfully updated on standby Orchestrator.
   ```

5.6.4 Stopping standby Orchestrator

You can stop or remove standby Orchestrator by using the **`ch3siteconfig -stopha -force`** command on active or standby Orchestrator.

To stop the standby Orchestrator, perform the following steps:

1. Enter the following command on standby Orchestrator:

   ```
   ch3siteconfig -stopha -force
   ```

   Note: This command might take several minutes to reflect on the other Orchestrator.

2. Validate the status of standby Orchestrator by using the **`ls3siteconfig`** command on active Orchestrator. See Example 5-23.

Example 5-23 Validate the status of standby Orchestrator

```
[root@rhel7host108 home]# ls3siteconfig
master_name Master
master_site_name
master_site_status online
master_ip1 x.xxx.xxx.xx
master_ip1_status reachable
master_ip2
master_ip2_status not_configured
master_port 7000
auxnear_name Aux_Near
auxnear_site_name
```

```
auxnear_site_status online
auxnear_ip1 x.xxx.xxx.xx
auxnear_ip1_status reachable
auxnear_ip2
auxnear_ip2_status not_configured
auxnear_port 7001
auxfar_name Aux_Far
auxfar_site_status online
auxfar_ip1 x.xxx.xxx.xxx
auxfar_ip1_status reachable
auxfar_ip2
auxfar_ip2_status not_configured
auxfar_port 7002
username tuser
keyfile /root/.ssh/pw_rsa
topology metromirror
Orchestrator_version 4.1.220922160427
Orchestrator_ha_status
```

The `Orchestrato r_ha_status` field is empty, which indicates standby Orchestrator is not connected or stopped successfully.

6

Failure protection and recovery procedures for 3-Site Replication

This chapter discusses different failure scenarios for Metro Mirror and HyperSwap configurations and how to handle them by using the 3-Site capabilities. It includes the following scenarios:

- Link failures between sites
- Storage system failures
- Single and multiple site failures
- Connectivity loss to the 3-Site Orchestrator or Orchestrator crash

A loss of a site, the link between those sites, or a storage system can be temporary or permanent. In any case, a decision must be made whether a site change is worthwhile.

This chapter describes how to perform 3-Site recovery operations by using the 3-Site Orchestrator. The other required steps to enable host access and get those volumes back online are not described.

This chapter includes the following topics:

6.1 General failure handling

Maintaining a warm site with a copy of your data ready-to-go is the usual way of ensuring rapid recovery from a large-scale failure. Adding a third site that is often geographically further away provides more data protection against multiple site failures.

In any of those disaster scenarios, manual intervention is required for providing data access to the suitable hosts, changing copy directions, or declaring a site as temporarily or permanently lost. The required steps depend on the used topology and failure scenario. To provide consistency, perform the required 3-Site changes on the 3-Site Orchestrator.

Some recovery steps for 3-Site Replication are more complex than normal replication and susceptible to human errors; for example, to fail back from the third site (Auxiliary-far site or AuxFar) to production.

Other failure scenarios can involve a loss of metadata and need more assistance and special care. One example is *Recover System Procedures* (also known as *T3 Recoveries*), which recover from some double hardware failures and potentially involve some amount of data loss. Another example is a *Loss of Cache Data* as a result of a sudden loss of both nodes of an I/O group without saving their hardened data. An indication for this loss can be unrecoverable FlashCopy mappings or Remote Copy relationships and corrupted volumes.

Note: It is recommended to contact IBM support for 3-Site solution recovery assistance when recovery is not possible by way of the 3-Site Orchestrator Command Line Interface (CLI) or if a complex failure occurs.

6.2 Helpful recovery procedures

This section provides an overview of general procedures that are used to recover from any type of failure (link failure, storage failure, or site failure).

It includes the following information:

- How to temporarily change the topology
- How to rejoin a site to the 3-Site configuration
- What is an Automatic Cluster Recovery and which extra steps are required after it occurs
- What is the Recover System Procedure and what should be considered regarding the 3-Site configuration before and after performing this procedure

6.2.1 Changing the topology

The different topologies are described in 6.3.1, “Recovering from link failures in a Metro Mirror configuration” on page 133. To change the topology from Cascade to Star and vice versa, you must switch the Periodic Copy Source to the other site. A switch of the Periodic Source is required to move from the `stopped` state into `2site_periodic` and to resume data cycling.

Complete the following steps to temporarily change to a Star topology that is based on a near link failure in a Metro Mirror Cascade topology:

1. Move the consistency group into `2site_periodic` state by issuing the **`ch3sitercconsistgrp -periodicsrc`** command, which specifies the new Periodic Source site. Then, run the **`ls3sitercconsistgrp`** command to verify the state of the consistency group, as shown in Example 6-1.

   The consistency group details page in the GUI shows that the consistency group is no longer stopped but performing asynchronous copy only. The 050993 event on the auxiliary-far system indicates that the 3-Site data cycling stopped is automatically fixed. A new event (050994) indicates that the consistency group is in 2-Site data cycling mode was reported on the auxiliary-far system.

Example 6-1 Move the consistency group into 2site_periodic state

```
[root@localhost ~]# ch3sitercconsistgrp -periodicsrc Master CGO

[root@localhost ~]# ls3sitercconsistgrp
name state          primary periodic_copy_source sync_copy_state     periodic_copy_state
freeze_time         cycle_period status
CGO  2site_periodic Master  Master               idling_disconnected stopped
2021/03/04/12/20/24 300          sync_link_offline
```

Alternatively, use the corresponding option from the GUI by clicking **Copy Services** → **Remote Copy** → **3-Site partnership**. Then, select the 3-Site consistency group and choose **Actions** → **Switch asynchronous source**. A window opens. Verify the topology change from `Cascade` to `Star` and click **Switch**, as shown in Figure 6-1.

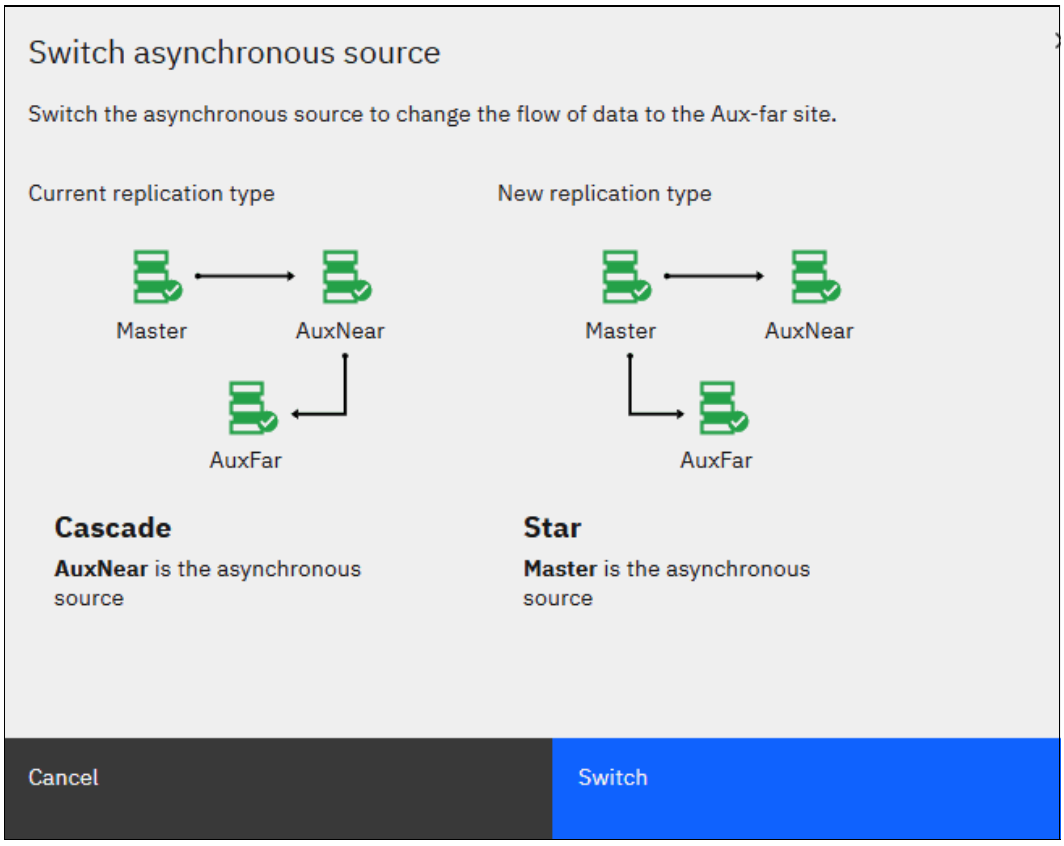

Figure 6-1 Temporarily switching to a Star topology

2. Ideally, a reconfiguration to the previous setup is performed when the link between both Near Sites returns to normal operation. This process can be done after data on the Near Site is synchronized. The resynchronization process can take some time. If the consistency group is in state `consistent_stopped`, a rejoin operation is required.

   For all other synchronous copy states, the resynchronization occurs automatically.

   Monitor the process until the next periodic cycling completes and returns to the `3site_consistent` state, as shown in Example 6-2. The state change might take some time and is reflected only after the `sync_copy_state` changes to `consistent_synchronized`. The difference to the previous command output is highlighted for better readability.

Example 6-2 Rejoin the AuxNear system to the 3-Site configuration after the link is repaired

```
[root@localhost ~]# ls3sitercconsistgrp
name state          primary periodic_copy_source sync_copy_state    periodic_copy_state
freeze_time         cycle_period status
CG0  2site_periodic Master  Master               consistent_stopped stopped
2021/03/04/12/27/14 300          sync_link_stopped

[root@localhost ~]# ch3sitercconsistgrp -join AuxNear CG0

[root@localhost ~]# ls3sitercconsistgrp
name state          primary periodic_copy_source sync_copy_state      periodic_copy_state
freeze_time         cycle_period status
CG0  2site_periodic Master  Master               inconsistent_copying stopped
2021/03/04/12/27/14 300          online

[root@localhost ~]# ls3sitercconsistgrp
name state          primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  2site_periodic Master  Master               consistent_synchronized stopped
2021/03/04/12/27/14 300          online

[root@localhost ~]# ls3sitercconsistgrp
name state            primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  3site_consistent Master  Master                 consistent_synchronized stopped
2021/03/04/12/31/26 300          online
```

3. In the GUI, select **Monitoring** → **Events**. Then, select the **Unfixed Alerts** view to check for any events that are not automatically fixed. Perform the recommended **Run Fix** procedure for any unfixed events, such as error code 1720 `Remote Copy - lost synchronization`, which is reported in the Event log of the Master system.
4. Change the setup back to a Cascade topology by using the **`ch3sitercconsistgrp -periodicsrc`** command or the corresponding action in the GUI. Then, verify the configuration, as shown in Example 6-3.

Example 6-3 Change the configuration back to Cascade topology and verify

```
[root@localhost ~]# ch3sitercconsistgrp -periodicsrc AuxNear CG0

[root@localhost ~]# ls3sitercconsistgrp
name state            primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
```

```
CGO  3site_consistent Master  AuxNear             consistent_synchronized copying
2021/03/04/12/36/26 300          online
```

6.2.2 Rejoining a site to the 3-Site configuration

If the `sync_copy_state` of the 3-Site consistency group is in a `consistent_stopped` state after repair, 3-Site data cycling does not continue automatically. A rejoin operation is required to resume 3-Site Replication.

Complete the following steps to rejoin a site after a near link failure in a Metro Mirror Cascade topology. Perform these steps after restoring the link between the Near Sites:

1. Issue the **`ls3sitercconsistgrp`** command to verify the state of the consistency group. In the resulting output, the state of the consistency group is `stopped`. The link is restored; therefore, the 3-Site consistency group status is now `sync_link_stopped` instead of `sync_link_offline`, as shown in Example 6-4.

Example 6-4 Verify consistency group state

```
[root@localhost ~]# ls3sitercconsistgrp
name state   primary periodic_copy_source sync_copy_state    periodic_copy_state
freeze_time         cycle_period status
CGO  stopped Master  AuxNear              consistent_stopped stopped
2021/03/04/12/27/14 300          sync_link_stopped
```

2. If the `sync_copy_state` is `consistent_stopped`, and the state of the 3-Site consistency group does not change to `3site_consistent`, complete the following steps:
   a. Issue the **`ch3sitercconsistgrp -join`** command to recover from the `stopped` state without moving into the `2site_periodic` state. Alternatively, you can find the corresponding option to join a site back in the 3-Site configuration in the GUI under **Copy Services** → **Remote Copy** → **3-Site partnership**. On the partnership page, select **Actions** → **Rejoin site**, if available.
   b. Use the **`ls3sitercconsistgrp`** command again to verify the state of the consistency group. The state change might take some time and during that process several different states are expected. The `3site_consistent` state is reflected after the `sync_copy_state` changes to `consistent_synchronized`, as shown in Example 6-5.

Example 6-5 Recover from stopped state without entering 2site_periodic

```
[root@localhost ~]# ch3sitercconsistgrp -join AuxNear CGO

[root@localhost ~]# ls3sitercconsistgrp
name state            primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CGO  3site_consistent Master  AuxNear              consistent_synchronized stopped
2021/03/04/12/31/26 300          online
```

   c. Check the event log for any unfixed events and perform the directed maintenance procedure (**Run Fix** procedure) to fix those issues.

6.2.3 Other required steps after an Automatic Cluster Recovery

An *Automatic Cluster Recovery* (or *Tier 2 [T2] Recovery*) recovers from software failures that occur on multiple nodes with loss of availability, but no loss of data.

If a T2 Recovery occurred (that is, both nodes of an I/O group are restarting at the same time), more steps are required before 3-Site Replication can resume. The GUI shows a warning (see Figure 6-2). An error code 1001 is reported in the event log to indicate that an Automatic Cluster Recovery ran.

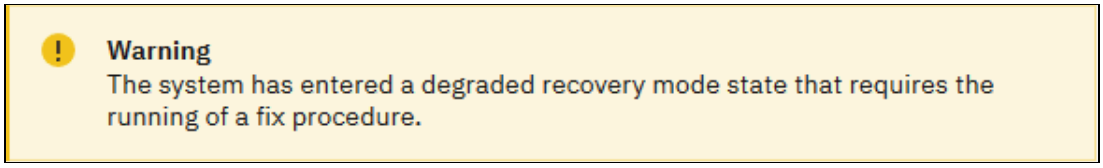

Figure 6-2 GUI indicates that an Automatic Cluster Recovery occurred

A T2 Recovery results in a short loss of I/O access for host systems until the recovery completes. After a successful T2 Recovery, the system continues operation; however, all configuration commands are blocked. Informational commands and data collection are still possible.

Complete the following steps:

1. Use the Service Assistant GUI (`https://<node_service_ip>/service`) to collect all necessary debug data, including the recovery dump. You need a single snap of each node. Under **Home**, select the node and go to **Collect Logs**. Under **Create a support package**, choose **With latest statesaves** and click **Create and Download**. The recovery dump can be found under **Available Files** (`/dumps/recover.*`).
2. Contact IBM Support to confirm that the recovery was successful. If software upgrade or downgrade was in progress, the Support team must ensure that all nodes are running the same code level as the cluster before allowing reconfiguration tasks.
3. After the state of the system is confirmed as safe, log in to the main cluster GUI:

   `https://<cluster_ip>/`

This login automatically starts the Fix Procedure (Directed Maintenance Procedure [DMP]) for the 1001 error, as shown in Figure 6-3.

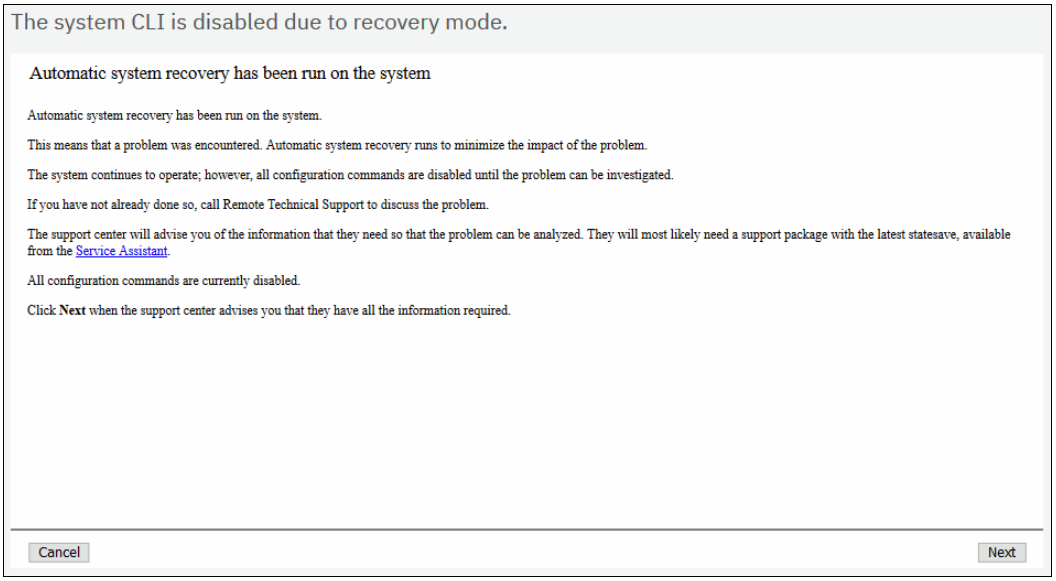

Figure 6-3 Fix Procedure (DMP) for the Automatic Cluster Recovery starts automatically

4. Follow the instructions that are presented. During this procedure, the GUI presents the option to re-enable the Command Line Interface (CLI). Select this option, as shown in Figure 6-4.

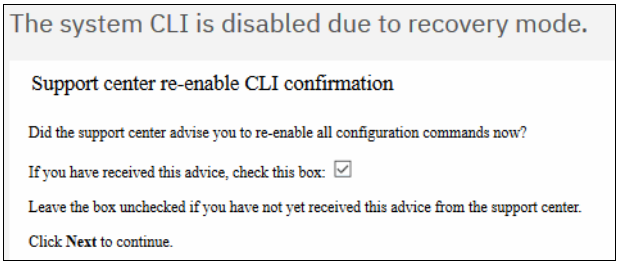

Figure 6-4 Tick the box to confirm to re-enable configuration commands

If a problem occurs when the GUI is used, you can unblock the configuration commands by using the system CLI. Enter the **svctask enablecli** command to enable 3-Site Orchestrator to access the storage system.

6.2.4 Considerations when the Recover System Procedure is required

The *Recover System Procedure* (also known as *Tier 3 [T3] Recovery*) is required if the system state is lost from all nodes and the system must be restored from the saved configuration. Although it recovers from some double hardware failures, it does potentially involve some loss of customer data. All nodes must be in candidate state. This process is not performed if any nodes are still active in the system.

How to perform the T3 Recovery is not covered in this book. For more information, see Recover system procedure.

On a HyperSwap system, some steps are required before a T3 Recovery can be started. T3 Recoveries for HyperSwap configurations always is performed with assistance from IBM Support.

Before performing the T3 Recovery procedure on an auxiliary-far system, data cycling must be stopped manually by using the **stop3sitercconsistgrp** command.

Before performing the Recover System Procedure on the HyperSwap system, the Auxiliary-far site volumes must be made accessible and the consistency groups must be converted from 3-Site to 2-Site.

After a T3 Recovery is performed, collect a snap with option 4 of the system, all `t3` files (from all nodes individually), the configuration file that was used for the recovery, and the recovery scripts (`svc.config.*`).

A T3 Recovery does not restore inter-system (that is Metro Mirror or Global Mirror) partnerships, consistency groups, and relationships. Also, FlashCopy mappings are not restored. Contact IBM Support for assistance with restoring the configuration, accessing the data, and replicating objects (such as partnerships, consistency groups, and relationships).

To recover the 3-Site Remote Copy configuration, more steps are required. IBM Support can assist with reconstructing the 3-Site Remote Copy configuration with a script that is run on the Orchestrator host.

For the script to be successful, the T3 Recovery first must run successfully on the system that suffered the disaster. The current version of the script automatically recovers only one of the Near Sites in a Metro Mirror configuration from a T3.

After a T3 Recovery at the auxiliary-far site or at the HyperSwap system, manual intervention is required. The script also is used by IBM Support to recover from a 3-Site consistency group status with `partner_dead_domain` in a HyperSwap configuration. No 3-Site data cycling or configuration changes should be performed by the Orchestrator host during the post-recovery process.

6.3 Recovering from link failures

Link failures between the sites can be because of temporary or permanent issues. If a permanent issue occurs, a topology change might be required.

One of the Near Sites is used as the source of asynchronous replication to the Auxiliary Far site. The site is called *Periodic Source* or *Periodic Copy Source*. This designation is especially important when differentiating between Star and Cascade topologies.

The Active Periodic Link describes the Active Link that is used for data replication from one Near Site to the far site. It is copying from the site that is defined as Periodic Source.

The Inactive Periodic Link describes the configured standby link to the auxiliary-far (AuxFar) site, which is normally not used for data replication. Both links are shown in Figure 6-5.

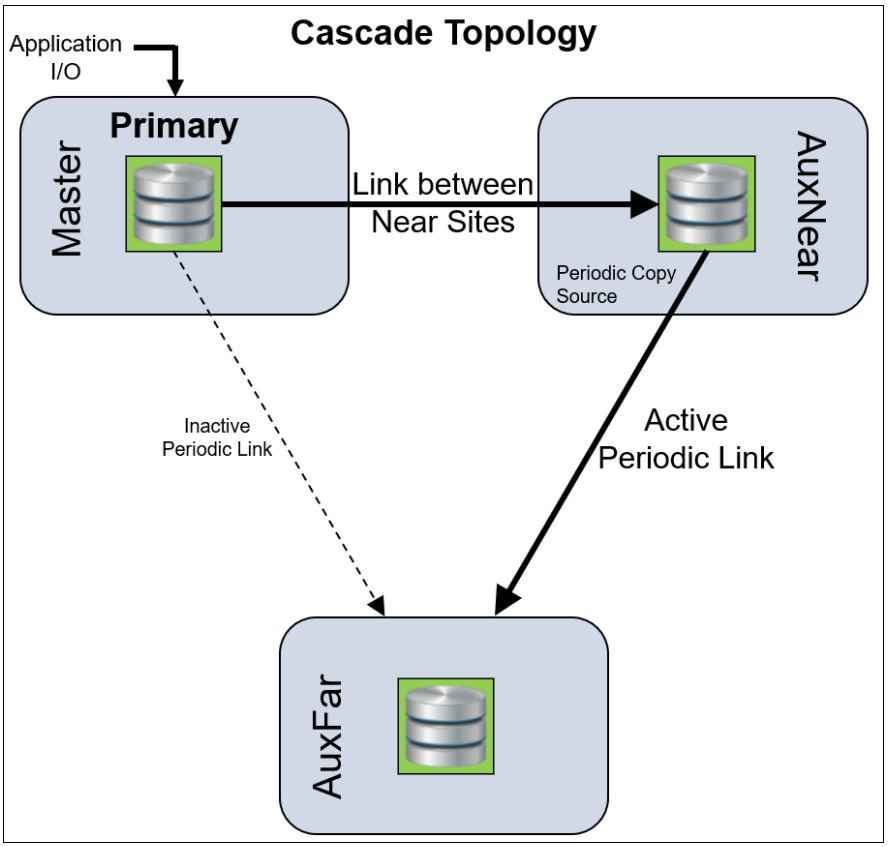

Figure 6-5 Link description between main sites

6.3.1 Recovering from link failures in a Metro Mirror configuration

This section describes how to recover from link failures that occur in a Metro Mirror configuration.

Link failure between Near Sites

If data replication between both Near Sites fails, you can assume that the auxiliary-near site (AuxNear) target no longer receives any updates and stays on the last consistent data copy before the link failure occurred.

In a Cascade topology, the auxiliary-far replication depends on the successful replication between both Near Sites. Therefore, you must differentiate between Star and Cascade topologies.

Star topology

Master is the asynchronous source site (that is, the Periodic Copy Source) to replicate data to AuxFar. The primary applications are hosted at the site that is labeled as Primary.

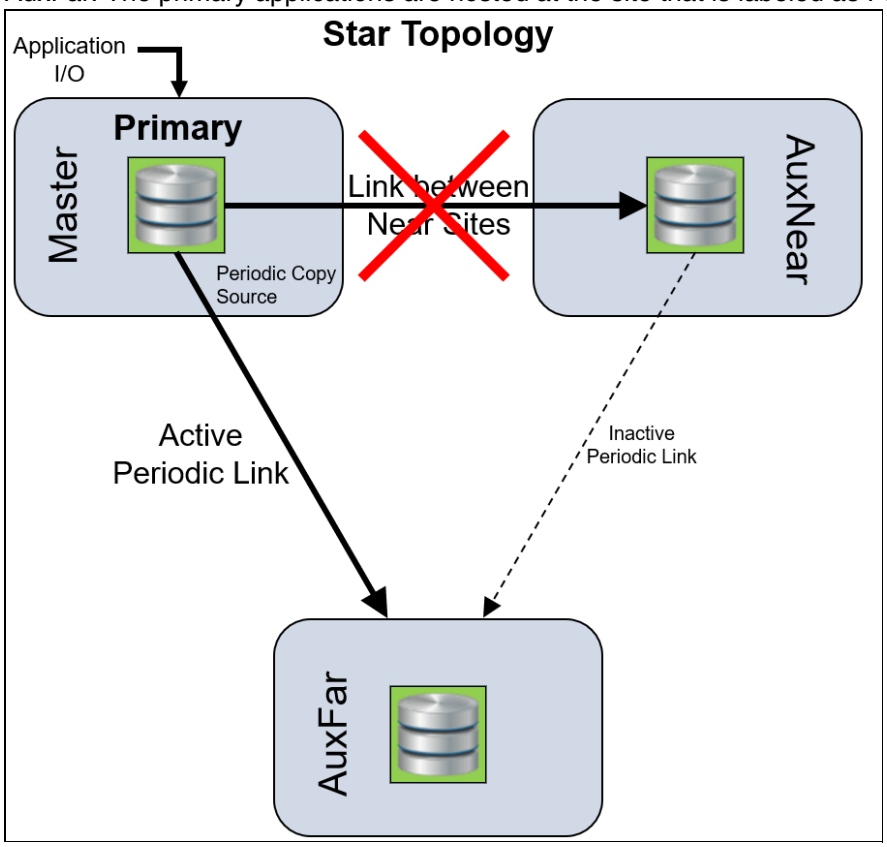

Figure 6-6 Link failure between Near Sites in a Star topology

To recover from a link failure between Near Sites in a Star topology as shown in Figure 6-6, complete the following steps:

1. Verify that the 3-Site consistency group is now in 2-Site data cycling mode. Verify the event notification by using the event log. The AuxFar system reports event ID 050994. The consistency group details in the GUI show that the synchronous link is offline and that the consistency group is performing asynchronous replication only.

   The Orchestrator sends the status change to the AuxFar system, as described in Chapter 5, “3-Site Replication monitoring and maintenance” on page 99. To verify, in the GUI of the AuxFar system, click **Monitoring** → **Events** to see the list of events. Then, search for the event that shows the event text 3-Site consistency `group is now in 2-Site data cycling mode`, as shown in Figure 6-7.

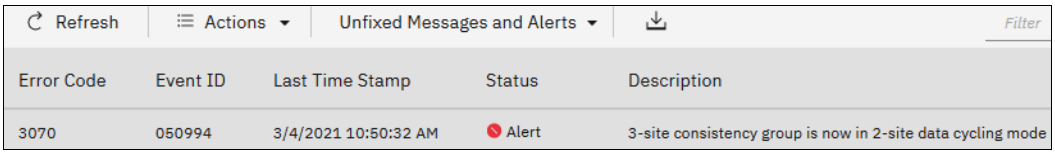

Figure 6-7 Event to indicate the change from 3-Site to 2-Site data cycling mode on AuxFar

2. Run the **ls3sitercconsistgrp** 3-Site Orchestrator CLI command or use the status details page in the GUI to verify the state and status of the 3-Site consistency group. In the resulting output, the state of the consistency group changes after some time to `2site_periodic`, as shown in Example 6-6. The status field displays `sync_link_offline`. This status confirms that the Near Site link is disconnected.

Example 6-6 Verify consistency group state

```
[root@localhost ~]# ls3sitercconsistgrp
name state            primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  3site_consistent Master  Master               consistent_synchronized stopped
2021/03/04/10/47/06 300          online

[root@localhost ~]# ls3sitercconsistgrp
name state          primary periodic_copy_source sync_copy_state     periodic_copy_state
freeze_time         cycle_period status
CG0  2site_periodic Master  Master               idling_disconnected stopped
2021/03/04/10/50/37 300          sync_link_offline
```

3. Repair the link between the Near Sites. The `sync_copy_state` changes from `idling_disconnected` to `consistent_stopped`. The status of the 3-Site consistency group switches from `sync_link_offline` to `sync_link_stopped`.
4. Run the **ls3sitercconsistgrp** command again to verify the state of the consistency group.

   If the `sync_copy_state` of a consistency group is `consistent_stopped`, issue the **ch3sitercconsistgrp -join** command, followed by the `site_alias` or `system name` and the consistency group to rejoin the site in 3-Site data cycling and verify the repair action. Data cycling moves to **3site_consistent** (see Example 6-7). The state can become `consistent_stopped` when the Near Site link is reconnected but in a stopped state, which is indicated by `sync_link_stopped`.

Example 6-7 Rejoin Near Site in 3-Site data cycling, status shows sync_link_stopped after the link is repaired

```
[root@localhost ~]# ls3sitercconsistgrp
name state          primary periodic_copy_source sync_copy_state    periodic_copy_state
freeze_time         cycle_period status
CG0  2site_periodic Master  Master               consistent_stopped stopped
2021/03/04/10/55/37 300          sync_link_stopped

[root@localhost ~]# ls3siteconfig  | grep auxnear_name
auxnear_name AuxNear
[root@localhost ~]# ch3sitercconsistgrp -join AuxNear CG0

[root@localhost ~]# ls3sitercconsistgrp
name state            primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  3site_consistent Master  Master               consistent_synchronized stopped
2021/03/04/11/00/38 300          online
```

   Alternatively, use the GUI and click **Copy Services** → **Remote Copy**. Then, on the **3-Site Partnership** page, choose **Actions** → **Re-join Site**, if available.
5. Perform the recommended actions (Run Fix procedures) for any unfixed events in the event log of the three systems. The event regarding 2-Site data cycling on the AuxFar system is fixed automatically.

Cascade topology

AuxNear is the asynchronous source site (that is, the Periodic Copy Source) to replicate data to AuxFar. The primary applications are hosted at the site that is labeled as Primary. In a Cascade topology, the status and RPO of the AuxFar replication depend on the successful replication between both Near Sites, as shown in Figure 6-8.

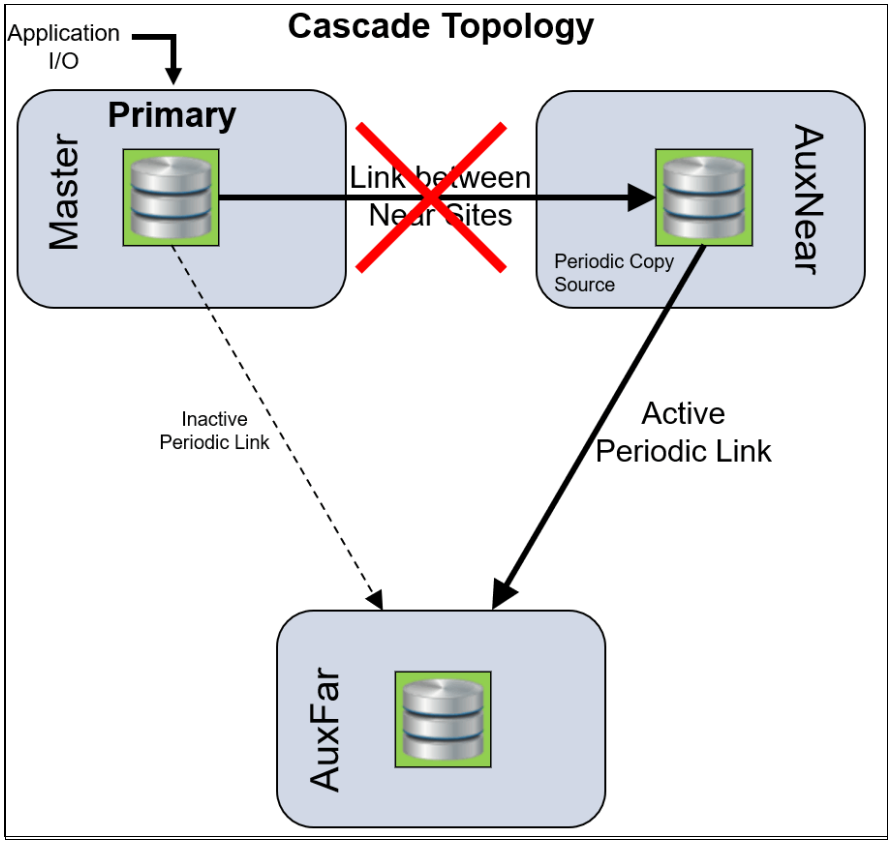

Figure 6-8 Link failure between Near Sites in a Cascade topology

Complete the following steps:

1. To confirm that a link failure in a Cascade topology occurred, verify the event notification by using the event log. Event ID 050993 is reported on the AuxFar system, which indicates that the 3-Site consistency group stopped replicating to the AuxFar site, as shown in Figure 6-9.

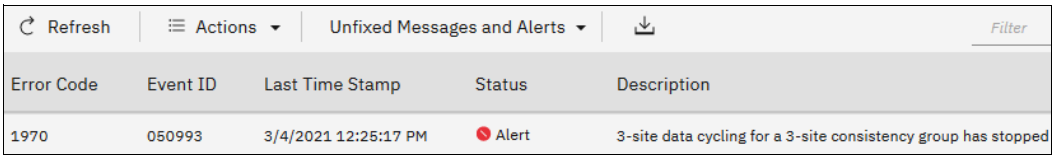

Figure 6-9 Stopped event for the 3-Site consistency group on the AuxFar system

2. Issue the **ls3sitercconsistgrp** 3-Site Orchestrator CLI command or use the status details page in the GUI to verify the state of the consistency group, which is now in `stopped` state. The status field displays `sync_link_offline`, as shown in Example 6-8. This status confirms that the Near Site link is disconnected.

Example 6-8 Verify consistency group state

```
[root@localhost ~]# ls3sitercconsistgrp
name state   primary periodic_copy_source sync_copy_state    periodic_copy_state freeze_time
cycle_period status
CGO  stopped Master  AuxNear              idling_disconnected stopped
2021/03/04/12/20/24 300          sync_link_offline
```

The following options are available as to how to proceed when the link between the Near Sites failed:

- Option 1: Temporarily change to a Star topology

  A temporary change to a Star topology enables replication between the Master and AuxFar sites. This option provides an up-to-date second (asynchronous) data copy, although the synchronous mirror between the Near Sites is not available.

  The 3-Site consistency group state changes to `2site_periodic`. A switch back to the former Cascade topology is possible when the link between both Near Sites is repaired and the synchronous copy state is back to `consistent_synchronized`.

  For more information, see 6.2.1, “Changing the topology” on page 126.

- Option 2: Rejoin the site to a 3-Site configuration

  Keep the current topology and reestablish the connection between both Near Sites when the link between those two sites returns to normal operation. The data copy on AuxNear and AuxFar cannot be updated until the link between both Near Sites returns to normal operation mode.

  When the consistency group is stopped, the data on the AuxFar site is consistent, based on the available freeze time (which specifies the last available snapshot at the Auxiliary-far site).

  The required steps to perform this operation are described in 6.2.2, “Rejoining a site to the 3-Site configuration” on page 129.

Link failure of the Active Periodic Link

The *Active Periodic Link* connects the AuxFar site with the Periodic Copy Source, as shown in a Star topology in Figure 6-10.To recover from an Active Periodic Link failure, complete the following steps:

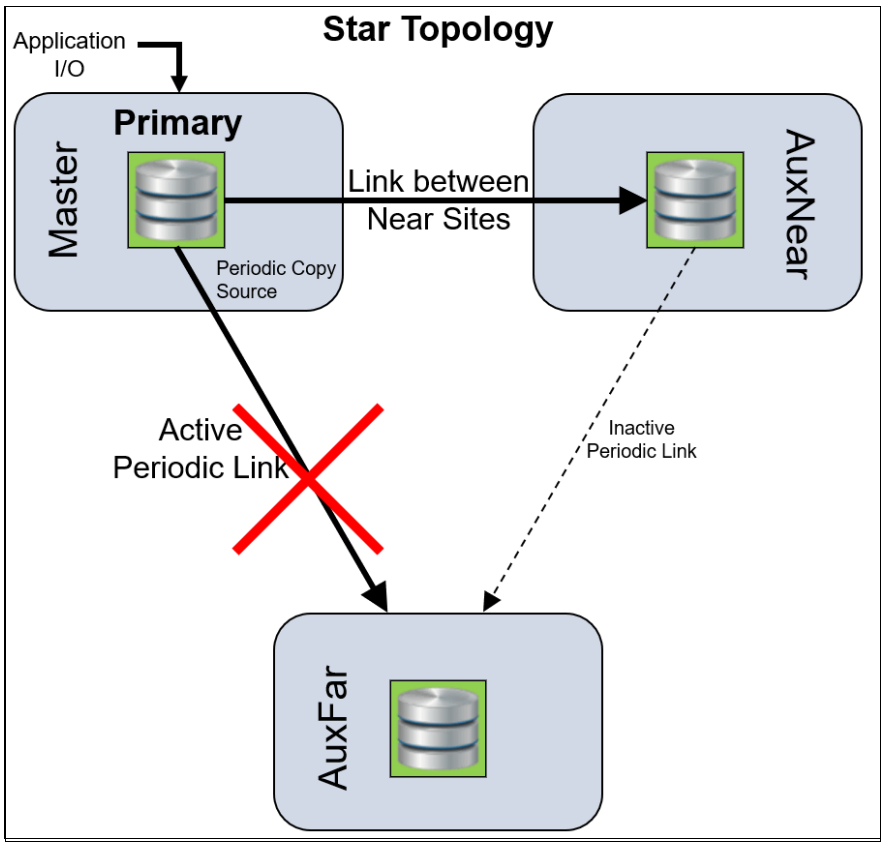

Figure 6-10 Active Periodic Link failure in a Star topology

1. Verify that the AuxFar system received the event log notification for event ID 05099, which indicates that the 3-Site consistency group is now in `stopped` state.
2. Run the **`ls3sitercconsistgrp`** command or use the status details page in the GUI to verify the state and status of the consistency group. In the resulting output, the state of the consistency group is `stopped`, as shown in Example 6-9. The status field displays `active_link_offline`.

Example 6-9 Verify the state of the consistency group

```
[root@localhost ~]# ls3sitercconsistgrp
name state   primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  stopped Master  Master               consistent_synchronized stopped
2021/03/04/11/59/41 300          active_link_offline
```

3. To complete the recovery from the link failure, use one of the following options:
   - Restore the failed link and data cycling resumes automatically. The event at the AuxFar system regarding the stopped 3-Site data cycling fixes automatically. The messages about the connection loss to the remote system on the Master and AuxFar systems should be fixed manually.
   - Switch the periodic replication source:
     - Use the **ch3sitercconsistgrp -periodicsrc** command to switch the Periodic Source to the other Near Site. Because the command completes asynchronously, check the events in the event log of the AuxFar system to confirm that the configuration change completed successfully. The state of the consistency group changes to `3site_consistent` and the status is now `inactive_link_offline`, as shown in Example 6-10. This status confirms that the other link (formerly standby link) is now used as the Active Periodic Link for data replication.

Example 6-10 Switch periodic replication source from Master to AuxNear

```
[root@localhost ~]# ch3sitercconsistgrp -periodicsrc AuxNear CGO

[root@localhost ~]# ls3sitercconsistgrp
name state            primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CGO  3site_consistent Master  AuxNear              consistent_synchronized stopped
2021/03/04/12/07/42 300          inactive_link_offline
```

     Alternatively, use the GUI to switch the asynchronous source under **Copy Services** → **Remote Copy** → **3-Site partnership**, select the 3-Site consistency group, and then choose **Actions** → **Switch asynchronous source**.
     - Restore the link and use the command **ch3sitercconsistgrp -periodicsrc** (as shown in Example 6-11) to switch the active periodic replication source back to the original site. The original link is used again as the Active Periodic Link for data replication.

Example 6-11 Switch periodic replication source back to Master

```
[root@localhost ~]# ch3sitercconsistgrp -periodicsrc Master CGO
```

If the Inactive Periodic Link also is not available (for example, because of a failure), changing the Periodic Source does not change the status. It still shows `active_link_offline`. In that case, the only option is to repair the Active Link to resume 3-Site Replication.

The status of the consistency group also shows `active_link_offline` if the Active Periodic Link and the synchronous link between Near Sites are offline at the same time.

Link failure of the Inactive Periodic Link

The *Inactive Periodic Link* connects the AuxFar site with the replication source site, as shown in a Star topology in Figure 6-11.

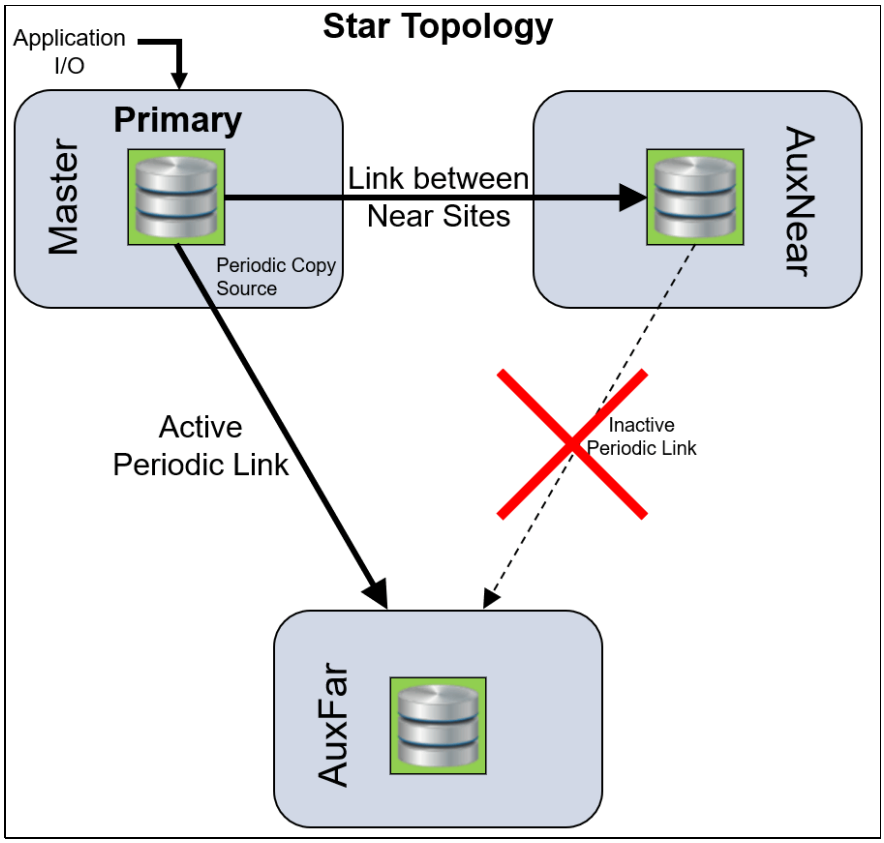

Figure 6-11 Inactive Periodic Link failure in a Star topology

To recover from an Inactive Periodic Link failure, complete the following steps:

1. Run the **`ls3sitercconsistgrp`** command or use the status details page in the GUI to verify the status of the consistency group. In the resulting output, the status is `inactive_link_offline`, as shown in Example 6-12. The Active Periodic Link is still used for data replication. A failure of the Inactive Periodic Link does not affect the 3-Site data cycling unless the Active Periodic Link also fails.

Example 6-12 Verify the state of the consistency group

```
[root@localhost ~]# ls3sitercconsistgrp
name state            primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  3site_consistent Master  Master               consistent_synchronized idling
2021/03/04/11/11/09 300          inactive_link_offline
```

2. Restore the link at the earliest opportunity. After the Inactive Periodic Link is back online, the error state automatically clears.

6.3.2 Recovering from link failures in a HyperSwap configuration

In a HyperSwap configuration, the site with ID 1 is designated as the Master site by the 3-Site Orchestrator. The site with ID 2 is designated as AuxNear.

If the link between sites in a HyperSwap configuration fails, you can use several recovery methods to fix the link failure. The near link is going down, or the far link fails.The near link is the link between Master and AuxNear that is used for the active-active relationship. The far link is the link between one of the Near Sites and AuxFar that is used for asynchronous data replication.

In a HyperSwap configuration, the link to the far site can be a Fibre Channel link or a native IP link, depending on the partnership type.

In an IP partnership, two inter-site links are configured by using two separate Remote Copy port groups. The port from one node of the I/O group is configured in the first remote-copy port group and the port from the other node is configured in the second remote-copy port group. Therefore, one port from each node actively participates in the IP partnership (run the **`svcinfo lsportip`** command to verify that both links show a `remote_copy_status` of `used`).

If a critical failure is observed on one node, the IP partnership continues over the other port on the other node. The use of external WAN optimization devices is not supported because IP Replication includes a built-in WAN accelerator software. Also, the use of an external WAN accelerator on the same link leads to contention, which causes problems with IP Replication.

A HyperSwap configuration ensures high availability during link failures because availability is maintained on a minimum of two copies of data across the three sites.

Different HyperSwap configurations are possible with Fibre Channel. It is important to distinguish between a HyperSwap configuration that uses inter-switch links (ISLs) and does not use ISLs.

When ISLs are used, two SAN fabrics are used, a private SAN fabric and a public SAN fabric. A private SAN fabric is dedicated for node-to-node communication. A public SAN fabric is dedicated for host, storage, and Remote Copy traffic to other clusters, such as participation in 3-Site Replication.

Lease renewal messages, cache mirroring and HyperSwap operations occur by way of the private SAN fabric by using the ports that are assigned for local node traffic. Those two redundant fabrics must be physically separated, which means that you must ensure that public and private fabrics traverse dedicated ISLs if virtual fabrics are used.

In a configuration that does not inter-switch links, each node is directly connected to the Fibre Channel switches in the local and remote production site. This configuration is often realized by using Longwave Small Form-factor Pluggables (SFPs) for the connections to the remote site (so that no ISLs are required) and Shortwave SFPs for the connections to the local site. It is possible to include powered components, such as dense wavelength division multiplexing (DWDM) Fibre Channel extenders, between the system and the switches.

Note: In a HyperSwap configuration, the 3-Site Orchestrator does not control the Primary. The `primary` attribute of a 2-Site consistency group indicates which volumes have the primary role (they are hosting the application primary and are accessible for write I/O). The direction of copying is from primary to secondary. Orchestrator manages only the Periodic Source, and HyperSwap manages the Primary.

Link failure between Near Sites

A link failure between Near Sites in a HyperSwap configuration is triggered by a failure of the private SAN fabric (which is used for node-to-node communication and HyperSwap synchronization traffic). In a configuration without ISLs, a failure of the Longwave connections that are used for internode communication to the remote site triggers the link failure. For example, if a DWDM to which all the Longwave links are connected fails, this results in all links becoming unavailable.

The connection between sites 1 and 2 (Master and AuxNear) is lost. The HyperSwap system continues to operate the site that wins the quorum race, and the nodes in the other site stop because of lease expiry, which triggers because the nodes are waiting for connectivity between sites 1 and 2 to be restored. The HyperSwap function automatically adjusts which volumes have the primary role (master or aux) for the 2-Site consistency group and switches direction to the site with the quorum, if required.

The failover of the Periodic Copy Source occurs automatically as well to the site that is serving as *Primary* for the 3-Site Remote Copy consistency group (that is, the site that is running the primary applications). 3-Site Replication takes the snapshot of the site that has the application primary and copies it to the far site.

Note: No automatic failback of the Periodic Copy Source occurs unless a second failure also occurs. A manual change is possible, but not required.

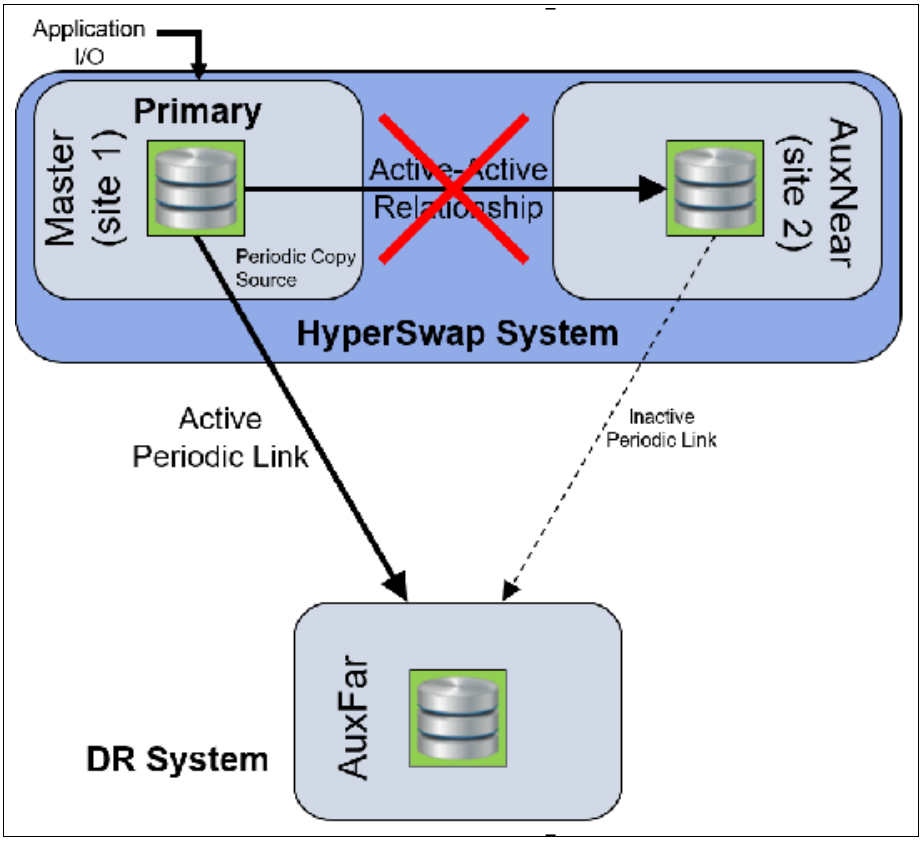

Figure 6-12 Link failure between Near Sites in a HyperSwap configuration

To recover from a link failure between Near Sites as shown in Figure 6-12 on page 142, complete the following steps:

1. Verify that the 3-Site consistency group is now in 2-Site data cycling mode. Verify the state of the consistency group, which is now `2site_periodic`, by using the **`ls3sitercconsistgrp`** Orchestrator CLI command or the status details page in the GUI. The status is `partner_offline`, as shown in Example 6-13.

Example 6-13 Consistency group details when a near link failure in a HyperSwap configuration

```
[root@localhost ~]# ls3sitercconsistgrp
CMMVC9507E SSH Connection failed with site [HyperSwap cluster IP].
[root@localhost ~]# ls3sitercconsistgrp
name state          primary periodic_copy_source sync_copy_state    periodic_copy_state
freeze_time         cycle_period status
CG0  2site_periodic Master  site1                consistent_copying stopped
2021/03/16/12/03/50 300          partner_offline
```

   If the configuration node changed because of the link failure, the Orchestrator might display a connection error to the HyperSwap system for approximately one minute until the management IP is active again.

   Event ID 050994 is reported on the AuxFar system, which indicates that the consistency group is in 2-Site data cycling mode. As opposed to Metro Mirror configurations, no event about lost connection to the remote cluster is reported if a near link failure occurs in a HyperSwap configuration.

   The GUI also reflects the consistency group status and indicates that the partner site is offline and that the consistency group is performing asynchronous replication only.

   On the Master system, error code `1940 HyperSwap volume or consistency group has lost synchronization between sites` is present and the synchronous copy state is `consistent_copying`.

2. Repair the link. The volume copies in the HyperSwap system start resynchronization. The events about the offline nodes and storage pool are automatically fixed.

3. Verify the state of the 3-Site consistency group by using the **`ls3sitercconsistgrp`** command or the status details page in the GUI. It automatically changes from `2site_periodic` to `3site_consistent` after some time. Then, the status changes to `online`. The event on the AuxFar system is fixed automatically.

4. After resynchronization completes, verify that the 1940 error on the Master system was marked as fixed, as shown in Figure 6-13. The synchronous copy state becomes `consistent_synchronized` after synchronization finishes. The 2030 errors that are related to the lease expiry are marked manually as fixed.

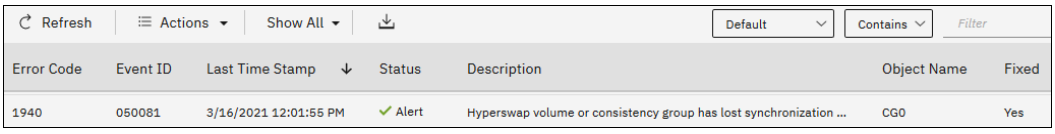

Figure 6-13 Lost HyperSwap synchronization error was fixed automatically

During and after the near link failure, data cycling to the far site continues and the freeze time is updated frequently.

Note: If the consistency group moved to `stopped` state instead of `2site_periodic`, you should contact IBM Support to recover from the link failure. In this case, the user lost the Primary (that is, the primary site for the 3-Site consistency group) and the use case is similar to a loss of the Primary site in the case of Metro Mirror.

Link failure between the Near Sites and AuxFar site

If any single Active or Inactive Link fails, data cycling continues by using the available link. Figure 6-14 shows an example of an Active Periodic Link failure in a HyperSwap configuration.

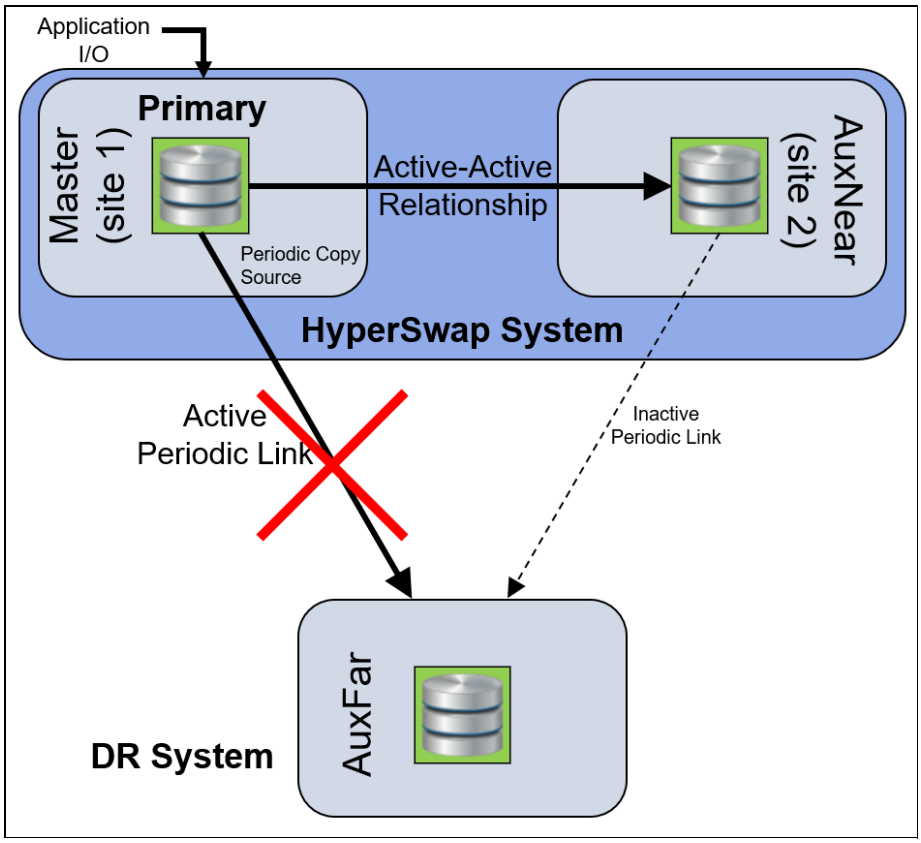

Figure 6-14 Active Periodic Link failure in a HyperSwap configuration

Because the system partnership is still the same, the 3-Site Replication is not affected. The 3-Site consistency group state remains in `3site_consistent` and the status displays as `online`. The Periodic Source does not need to be changed if the Active Link fails because HyperSwap manages access to the data by rerouting the traffic internally. Recover the link at your earliest convenience.

In an IP partnership, data cycling now uses a different remote-copy port group and node. No failure occurs in this scenario; however, the effective bandwidth is reduced to half. When the failure is corrected, the ports fail back and the IP partnership continues to operate over both links.

Only when the partnership between the HyperSwap system and the AuxFar system is stopped or fails, the state of the consistency group moves to the `stopped` state and an event ID 050993 is reported on the AuxFar system. This can occur if both Active and Inactive Links are not available. In such a scenario, the 3-Site consistency group status moves to `hyperswap_and_auxfar_partnership_offline`. The status details page in the GUI displays *Partnership Offline*, as shown in Figure 6-15.

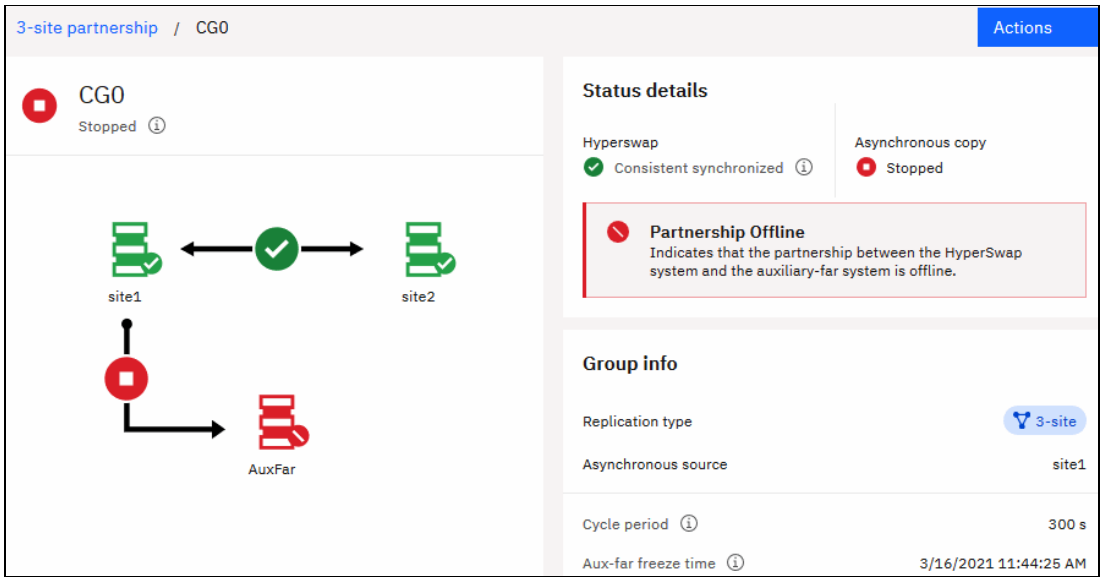

Figure 6-15 Consistency group status when none of the asynchronous copy links is available

Start the partnership again if it was stopped manually. Restore at least one of the two links to the far site to resume data cycling. The 050993 event on the AuxFar system is fixed automatically.

An error code `2020 IP Remote Copy link unavailable` is logged if the HyperSwap system cannot reach the AuxFar system over IP for the last 60 seconds. User intervention is not required for this auto fix event because the link is retried after connectivity is restored.

Note: Because data cycling continues normally, the 3-Site Orchestrator does not detect any single link failure between the Near Sites and the AuxFar site. Only if both Active and Inactive Link are broken does the data cycling stop, and a problem is detected.

6.4 Recovering from a storage failure

A storage system failure can be temporary or permanent. In either case, manual intervention is often required for recovery. As opposed to a complete site loss, a storage system failure affects only a subset of consistency groups (which is related to the failed storage) and not all consistency groups.

In a 3-Site Replication solution, a storage failure at one site implies a situation where a storage system is up and running. However, a failure occurs that takes a volume offline that is a member of the 3-Site Replication relationship.

The 3-Site Orchestrator can still access the site and detect both its partners over the replication links. An example of such a type of failure is a storage pool offline situation because of a lack of space or a SAN communication problem.

Other examples include a thin-provisioned volume copy out of space event, or a FlashCore Module (FCM) array that went write protected.

When unrecoverable relationships occur or FlashCopy mappings (event ID 050001 or 030020) or corrupt volumes (error code 1862) result because of lost metadata, contact IBM Support to assist with the recovery process.

Note: This section does not cover all of the steps that are required to fix the storage system or recover the host application operation. It focuses on only the recovery of the 3-Site Replication operations.

For more information about the storage system and host recovery, contact IBM Support or see the following publications:

- *Implementation Guide for IBM Storage FlashSystem and IBM SAN Volume Controller: Updated for IBM Storage Virtualize Version 8.6*, SG24-8542
- *Performance and Best Practices Guide for IBM Storage FlashSystem and IBM SAN Volume Controller: Updated for IBM Storage Virtualize Version 8.6,* SG24-8543

6.4.1 Recovering from a storage failure in a Metro Mirror configuration

Recovery steps in a Metro Mirror configuration depend on the topology (Star or Cascade) and the Primary site assignment of the 3-Site consistency group at the time of the disaster. You can use the **`ls3sitercconsistgrp`** command to determine the state and status of the consistency group, as shown in Example 6-14.

Example 6-14 Example output of a consistency group in online status

```
[root@localhost ~]# ls3sitercconsistgrp
name state            primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CGO  3site_consistent Master  AuxNear              consistent_synchronized copying
2021/03/12/12/41/33 300          online
```

The `primary` column indicates the site that is running the application primary. The `periodic_copy_source` shows which site is the source of asynchronous data replication to the far site. If `primary` and `periodic_copy_source` match, 3-Site Replication is running in a Star topology.

If a storage failure occurs at one of the sites, the consistency group status changes from `online` to one of the possible options that are listed in Table 6-1.

Table 6-1 Consistency group status when a storage failure occurs

3-Site data cycling status	Description
`primary_storage_offline`	Primary site for this consistency group is available, but one or more Primary site volumes in this consistency group are offline.
`periodic_src_storage_offline`	Non-primary Near Site (which is a Periodic Copy Source) is available, but one or more of its volumes that are part of this consistency group are offline. (Cascade topology)
`partner_storage_offline`	Near site (which is non-primary and not a Periodic Copy Source) is available, but one or more of its volumes that are part of this consistency group are offline. (Star topology)

3-Site data cycling status	Description
auxfar_storage_offline	AuxFar site (which is not a Primary a Periodic Copy Source) is available, but one or more of its volumes that are part of the consistency group are offline.
master_and_auxnear_storage_offline	The storage on both Near Sites (Master site and AuxNear site) is offline.

Table 6-2 lists the storage failure behavior in Metro Mirror configurations. It shows the state and status of the 3-Site consistency group that are possible if a storage failure occurs at one of the three sites.

Table 6-2 Storage failure behavior in a Metro Mirror configuration

	Topology	Failure at	Primary	Periodic Source	CG state	CG status
#1	Star	Master	Master	Master	stopped	primary_storage_offline
#2		AuxNear			2site_periodic	partner_storage_offline
#4		AuxFar			stopped	auxfar_storage_offline
#1	Cascade	Master		AuxNear	stopped	primary_storage_offline
#3		AuxNear			stopped	periodic_src_storage_offline
#4		AuxFar			stopped	auxfar_storage_offline

These scenarios are described next.

Scenario 1: Recovery from a storage failure at the Master site

Complete the following steps to recover from a storage failure at the Master site in a Star or Cascade topology:

1. Verify the state and status of the consistency group by using the **ls3sitercconsistgrp** command or the status details page in the GUI. The consistency group is in stopped state and the status is primary_storage_offline, as shown in Example 6-15. On the AuxFar system, the event log shows event ID 050993 to indicate that 3-Site data cycling stopped.

Example 6-15 Consistency group status during a primary storage failure in a Star topology

```
[root@localhost ~]# ls3sitercconsistgrp
name state   primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  stopped Master  Master               consistent_synchronized stopped
2021/03/05/10/15/56 300          primary_storage_offline
```

2. The recovery options depend on several factors, including the expected duration of the outage. The user has the following options:
   - Wait until the offline up-to-date copy is back online. During this time, no application I/O is performed on the surviving site and the consistency group remains in stopped state. After fixing the storage, the 3-Site consistency group recovers automatically and enters 3site_consistent.

This action can be taken if the value that is indicated in the `freeze_time` field is from too long ago, or if the date is before a recent vital update.

If data cycling does not continue automatically after the recovery, manually start the consistency group or rejoin the site, if required.

– Change the application primary and enable read and write access to the secondary volume to enter `2site_periodic`:

  i. Change the application primary by using the **`ch3sitercconsistgrp -primary`** command. The output of **`ls3sitercconsistgrp`** now shows that AuxNear is the Primary and the status changes to `periodic_src_storage_offline` in a Star topology.

  The status details page of the consistency group in the GUI shows `Asynchronous Source Storage Offline`, as shown in Figure 6-16. In a Cascade topology, Primary and Periodic Copy Source are now the same, and the new status is `partner_storage_offline`.

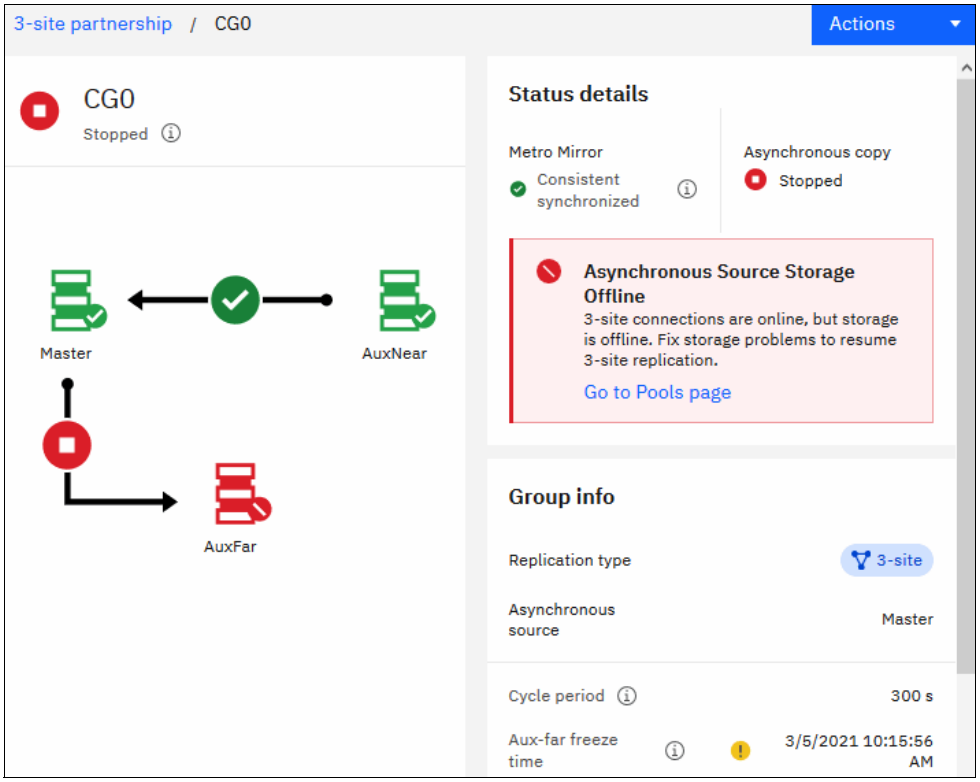

Figure 6-16 New consistency group status after switching primary in a Star topology

You can also choose the corresponding action from the GUI under **Copy Services** → **Remote Copy** → **3-Site partnership,** select the 3-Site consistency group, and on the details page, choose **Actions** → **Switch primary**. A window opens (see Figure 6-17).

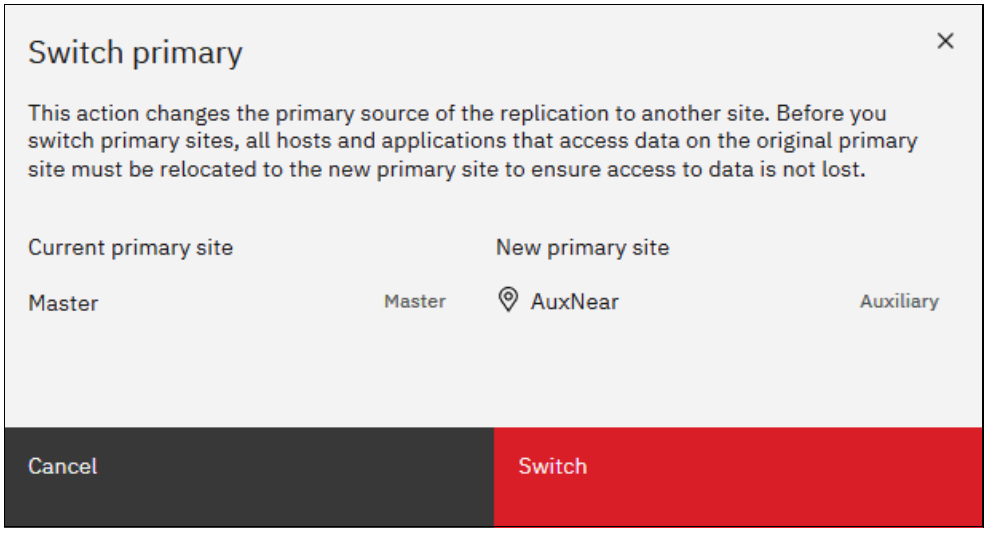

Figure 6-17 Switch primary dialog in a Star topology

ii. Enable read and write access to the secondary volume by using the system CLI of the AuxNear system. Open a secure shell (SSH) connection to the management IP address of the AuxNear system and log in as a user with CopyOperator or Administrator role.

The related CLI command, **`stoprcconsistgrp -access <consistency_group>`**, rewinds the state of the volume to the point in time (which is indicated by the `freeze_time`) of that stale copy. The Primary is set to `none`. Data that is presented to the hosts changes immediately to that of the stale copy.

In a Cascade topology, the 3-Site consistency group enters `2site_periodic`. If the consistency group was inconsistent, the command fails with no effect because you can enable only write access to a consistent secondary volume.

iii. Only in a Star topology after enabling read and write access as a next step it is required to change the Periodic Copy Source as well. Run the **`ch3sitercconsistgrp -periodicsrc`** command and specify the remaining AuxNear system. You can also choose the corresponding action in the GUI under **Copy Services** → **Remote Copy** → **3-Site partnership,** select the 3-Site consistency group, and on the details page, choose **Actions** → **Switch asynchronous source**. A new window opens as shown in Figure 6-18 on page 150.

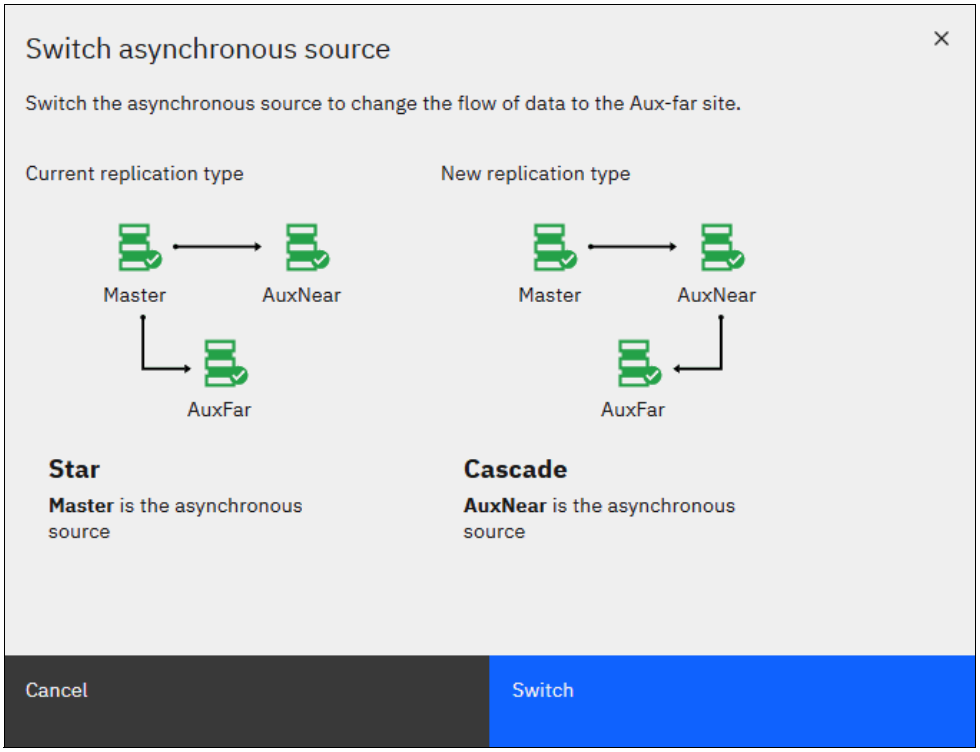

Figure 6-18 Switch asynchronous source to change into Cascade topology

iv. After the Master site is available again, the user must decide carefully if production resumes on the AuxNear site or if they want to switch back the applications. The decision is based on how much data was missing on the stale copy compared to the up-to-date copy and how much progress was made on the stale copy since access was enabled to it.

    Switching back again is a disruptive operation for the hosts. A switch back often is performed if the stale copy did not have useful data, or if the outage of the up-to-date copy was shorter than expected.

When access to the auxiliary volume is enabled for host operations, the host must be instructed to mount the volume before the application can be started, or perform a recovery process. Therefore, you must create the host mapping, perform a rescan from host side, and restart the host (if required).

For more information about how to enable read/write access to a Metro Mirror secondary volume, see *IBM System Storage SAN Volume Controller and Storwize V7000 Replication Family Services*, SG24-7574.

Note: Care must be taken with this procedure because running the **`stoprcconsistgrp -access`** command without precaution can crash your application and corrupt the stale copy. Do *not* perform this procedure if any data or state of the volume is still cached in the host systems. Ideally, the hosts that use the volume are shut down before this step is taken. Contact IBM Support for assistance with this procedure.

Scenario 2: Recovering from a storage failure at the non-primary AuxNear site in a Star topology

To recover from a storage failure at the non-primary AuxNear site in a Star topology, complete the following steps:

1. Verify the state of the consistency group by using the **ls3sitercconsistgrp** command or the status details page in the GUI. In the resulting output, the state is `2site_periodic` (as shown in Example 6-16) and the status shows `partner_storage_offline`, which indicates that the storage is offline at the non-primary site (which is also not a Periodic Copy Source). The AuxFar system reports event ID 050994 to indicate 2-Site data cycling mode.

Example 6-16 Consistency group output for storage failure at AuxNear in a Star topology

```
[root@localhost ~]# ls3sitercconsistgrp
name state          primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  2site_periodic Master  Master               consistent_synchronized stopped
2021/03/12/13/12/13 300          partner_storage_offline
```

2. Investigate the cause of the storage failure and perform the required recovery actions on the AuxNear storage to bring the volumes back online. After recovery of the storage is completed, the consistency group state switches back to `3site_consistent`.
3. If the failure state persists even when the Metro Mirror target volumes are back online, run the **ls3sitercconsistgrp** command to check the `sync_copy_state` and start the consistency group or perform rejoin action as required.
4. Confirm that the state of the consistency group is now `3site_consistent`.

At the time of this writing, an open issue still exists where a storage failure at the AuxNear site in a Star topology might not be reflected correctly in the GUI and CLI. The consistency group still shows the `3site_consistent` state and an `online` status, even though the volumes at the AuxNear site are offline. The status remains the same, even when the storage issue is fixed. Running the command **stop3sitercconsistgrp** followed by **start3sitercconsistgrp** resumes data cycling. It also is possible to stop and start the consistency group by using GUI.

Scenario 3: Recovering from a storage failure at the non-primary AuxNear site in a Cascade topology

Complete the following steps to recover from a storage failure at the non-primary AuxNear site in a Cascade topology:

1. Use the **ls3sitercconsistgrp** command or the status details page in the GUI to verify the state of the consistency group. In the resulting output, the state is `stopped` as shown in Example 6-17. The consistency group status shows `periodic_src_storage_offline`, which indicates that the storage is offline at the non-primary site that is acting as the Periodic Copy Source for the asynchronous data replication to the AuxFar system. The GUI shows `Asynchronous Source Storage Offline`. The AuxFar system reports the stopped event (ID 050993).

Example 6-17 Consistency group output for storage failure at AuxNear in a Cascade topology

```
[root@localhost ~]# ls3sitercconsistgrp
name state   primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  stopped Master  AuxNear              consistent_synchronized stopped
2021/03/12/13/11/37 300          periodic_src_storage_offline
```

2. Recover the AuxNear storage to bring the volumes back online. The state of the consistency group switches to `3site_consistent` after storage recovery. A rejoin is required if the `sync_copy_state` is `consistent_stopped` and data cycling does not resume automatically.

   Or, temporarily switch the Periodic Copy Source to change into 2-Site data cycling mode. Switching the Periodic Copy Source from the failed AuxNear site to the Master site changes the topology from Cascade to Star. This change can be reverted after the storage is available again.

   Changing the Periodic Copy might be an option if the storage cannot be repaired immediately and a requirement exists to provide an up-to-date second (asynchronous) data copy. Alternatively, you can leave the 3-Site data cycling inactive for a period when the storage at the AuxNear site experiences problems. If you decide to change into 2-Site data cycling, follow this step.

   To change the Periodic Source, use the **`ch3sitercconsistgrp -periodicsrc`** command and specify the remaining Near Site (which is the Master site). Or, choose **Switch asynchronous source** from the Actions menu in the GUI, as shown in Figure 6-19.

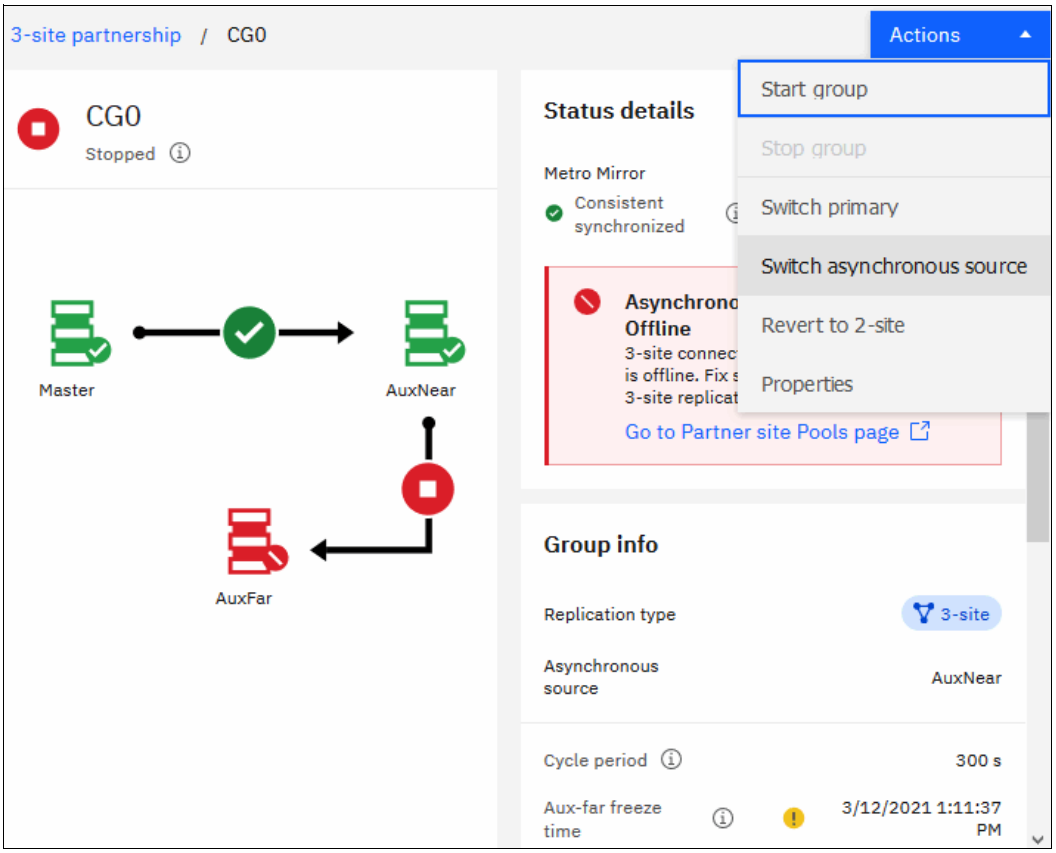

Figure 6-19 Switching Periodic Source during a storage failure at AuxNear in a Cascade topology

3. If you switched the Periodic Source as described in Step 3, you might want to change the configuration back to the original state by switching to the recovered AuxNear site as the Periodic Copy Source. Run the **`ch3sitercconsistgrp -periodicsrc`** command or use the corresponding action in the GUI.

At the time of this writing, an open issue exists that is under development investigation where after a switch of the Periodic Source during a storage failure at the AuxNear site in a Cascade topology, the status might not be reflected correctly in the GUI and CLI.

The consistency group correctly shows the status `partner_storage_offline`, but it still shows the `3site_consistent` state instead of `2site_periodic`. The status remains the same, even when the storage issue is fixed. Stop and Start the group to resume data cycling.

Scenario 4: Recovering from a storage failure at the AuxFar site

Complete the following steps:

1. Run the **`ls3sitercconsistgrp`** command or check the status details page in the GUI to verify the state of the consistency group. If a storage failure occurs at the AuxFar system, the state is `stopped` and the status is `auxfar_storage_offline` as shown in Example 6-18. The same status is reflected in the GUI. The AuxFar system reports the stopped event (ID 050993).

Example 6-18 Consistency group status during storage failure at far site

```
[root@localhost ~]# ls3sitercconsistgrp
name state            primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  stopped          Master  Master               consistent_synchronized copying
2021/03/09/11/25/11 300          auxfar_storage_offline
```

2. Perform the required steps to recover the storage at the AuxFar system.
3. After fixing the storage, the 3-Site Orchestrator automatically recovers and the data cycling continues in 3-Site data cycling mode. The stopped event on AuxFar is fixed automatically. You might still see an event ID `050991 The cycle time for 3-Site consistency group has exceeded cycle period set by user`, which automatically fixes after the next cycle. Figure 6-20 shows an example of an offline target volume for 3-Site Replication on the AuxFar system and the related events.

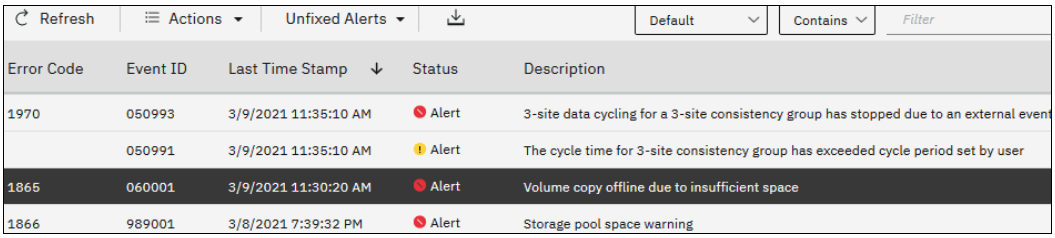

Figure 6-20 Events during a storage failure at AuxFar

4. If the storage failure persists even when the volumes are online again, manually start data cycling by using the **`start3sitercconsistgrp`** command or by way of the **Start group** action in the GUI.

6.4.2 Recovering from a storage failure in a HyperSwap configuration

In a HyperSwap configuration, the Primary and Periodic Copy Source are switched automatically. Therefore, you do not need to manually change the Primary or Periodic Copy Source if a storage failure occur. 3-Site data cycling resumes in 2-Site data cycling mode if a storage failure occurs at one of the Near Sites and automatically reverts to `3site_consistent` after the storage is recovered. If a storage failure occurs at the AuxFar system or if a storage failure occurs at both Near Sites, the 3-Site data cycling become `stopped`. It might be required to manually start data cycling after the storage is repaired.

Table 6-3 lists the migration of the Periodic Source and the consistency group `state` and `status`, depending on at which site the failed storage is located.

Table 6-3 Storage failure behavior in a HyperSwap configuration

Failure at	Automatic change of Periodic Source	CG state	CG status
Master (site 1)	Switches to AuxNear if Periodic Source was Master	2site_periodic	partner_storage_offline
AuxNear (site 2)	Switches to Master if Periodic Source was AuxNear	2site_periodic	partner_storage_offline
Both Near Sites	Not applicable	stopped	master_and_auxnear_ storage_offline
AuxFar	No change	stopped	auxfar_storage_offline

Recovering from a storage failure at a Near Site (Master or AuxNear)

To recover from a storage failure at a Near Site in a HyperSwap configuration, the storage must be fixed and then 3-Site Replication automatically switches back to `3site_consistent`. At the time of the failure, the HyperSwap Primary and the Periodic Source migrate to the partner site if required.

Complete the following steps:

1. Verify that the 3-Site Replication changed to 2-Site data cycling mode by running the **`ls3sitercconsistgrp`** command (as shown in Example 6-19) or checking the status details page in the GUI. In the resulting output, the state of the consistency group is `2site_periodic` and the status is `partner_storage_offline`, which indicates that the non-primary and not Periodic Source site storage failed.

   The Periodic Copy Source switched from site 2 to site 1 in the provided example and in the GUI, as shown in Figure 6-21 on page 155. On the AuxFar system, event ID 050994 is reported in the event log and event ID 050991, which indicates that the cycle time exceeded the configured cycle period. The HyperSwap system reports an event for the offline volume or storage pool.

Example 6-19 Consistency group status and switch of Periodic Source during storage failure at site 2

```
[root@localhost ~]# ls3sitercconsistgrp
name state            primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  3site_consistent Master  site2                consistent_synchronized stopped
2021/03/20/12/39/14 300          online

[root@localhost ~]# ls3sitercconsistgrp
name state            primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  3site_consistent Master  site2                consistent_copying      stopped
2021/03/20/12/44/14 300          online
[root@localhost ~]# ls3sitercconsistgrp
name state            primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  2site_periodic   Master  site1                consistent_copying      copying
2021/03/20/12/44/14 300          partner_storage_offline
```

```
[root@localhost ~]# ls3sitercconsistgrp
name state            primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  2site_periodic   Master  site1                consistent_copying      stopped
2021/03/20/12/54/39 300          partner_storage_offline
```

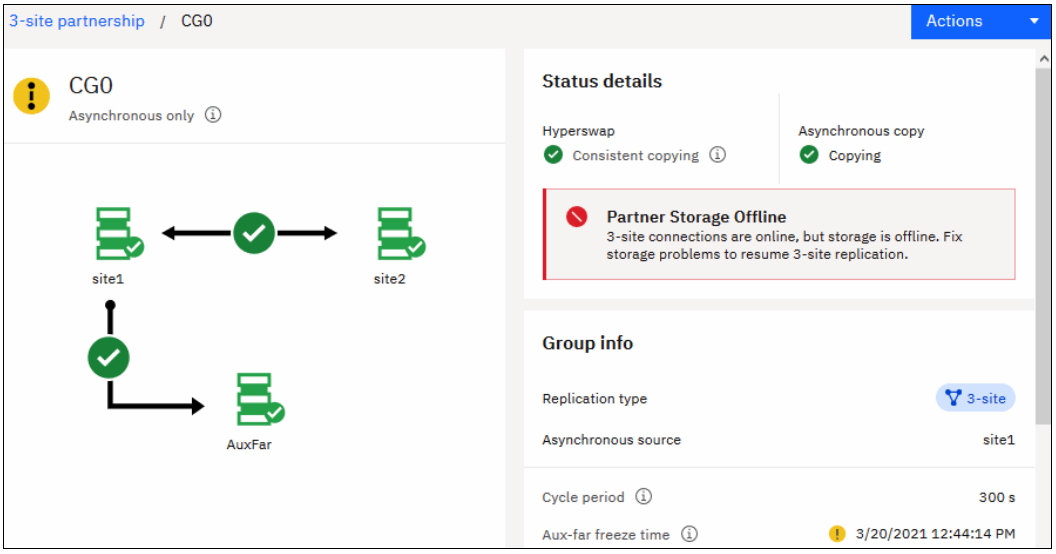

Figure 6-21 Periodic Source was switched to site1 and data cycling continues in 2-Site mode

Figure 6-22 shows the events for a storage failure at site 2 caused by a thin-provisioned volume copy, which is part of the 3-Site data cycling, going out of space.

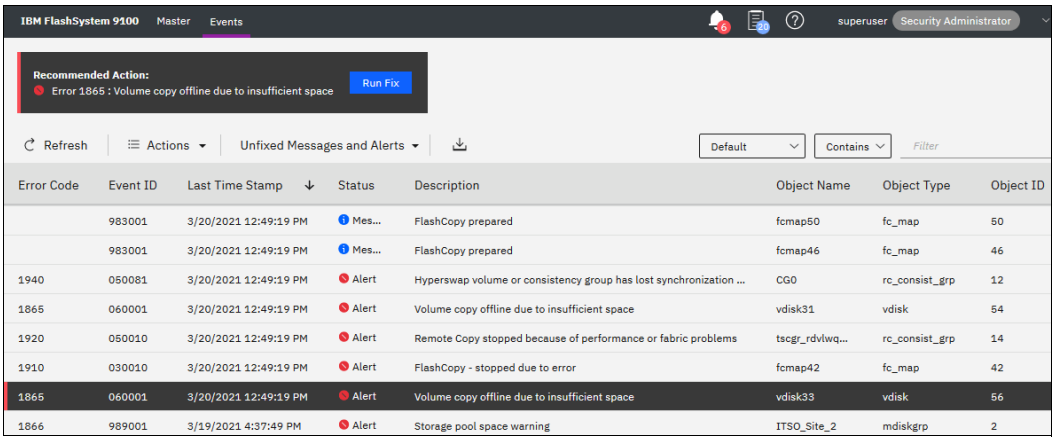

Figure 6-22 Volume copy offline at site 2 and related events on the HyperSwap system

2. Recover the storage at the failed Near Site (Master or AuxNear) to bring the HyperSwap volumes back online. The HyperSwap system clears the volume offline events after the storage is available again. Run the fix procedure (DMP) for any unfixed events; for example, `1910 Flashcopy stopped due to error` or `1920 Remote Copy stopped because of performance or fabric problems`. Contact IBM Support if the volumes do not come online automatically after the storage is repaired.
3. After the storage on the Master or AuxNear site is fixed, verify that the 3-Site Replication reverts to 3-Site data cycling mode. Run the **`ls3sitercconsistgrp`** command to verify the state of the consistency group or use the GUI to confirm the status.

In the resulting output, the state is `3site_consistent` and the status changed to `online`, as shown in Example 6-20. No failback exists of the Periodic Copy Source.

Example 6-20 Consistency group changes back to online and site 1 remains Periodic Source

```
[root@localhost ~]# ls3sitercconsistgrp
name state            primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  3site_consistent Master  site1                consistent_synchronized stopped
2021/03/20/13/19/49 300          online
```

The events on the AuxFar system are fixed automatically, as shown in Figure 6-23.

Error Code	Event ID	Last Time Stamp	Status	Description	Object Type	Object ID
3070	050994	3/20/2021 12:54:21 PM	✓ Alert	3-site consistency group is now in 2-site data cycling mode	3site_rc_consist_grp	3
	050991	3/20/2021 12:54:11 PM	✓ Alert	The cycle time for 3-site consistency group has exceeded cycle period set...	3site_rc_consist_grp	3

Figure 6-23 Events on AuxFar system are automatically marked as fixed

Recovering from a storage failure at the AuxFar site

Complete the following steps to recover from a storage failure at the AuxFar site:

1. Verify that the 3-Site Replication stopped by running the **`ls3sitercconsistgrp`** command, as shown in Example 6-21, or check the status details page for the consistency group in the GUI, as shown in Figure 6-24. The state of the consistency group indicates `stopped` and the status is `auxfar_storage_offline`. An event ID 050993 is reported in the event log of the AuxFar system.

Example 6-21 Consistency group details during storage failure at far site

```
[root@localhost ~]# ls3sitercconsistgrp
name state   primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  stopped Master  site2                consistent_synchronized copying
2021/03/19/16/00/09 300          auxfar_storage_offline
```

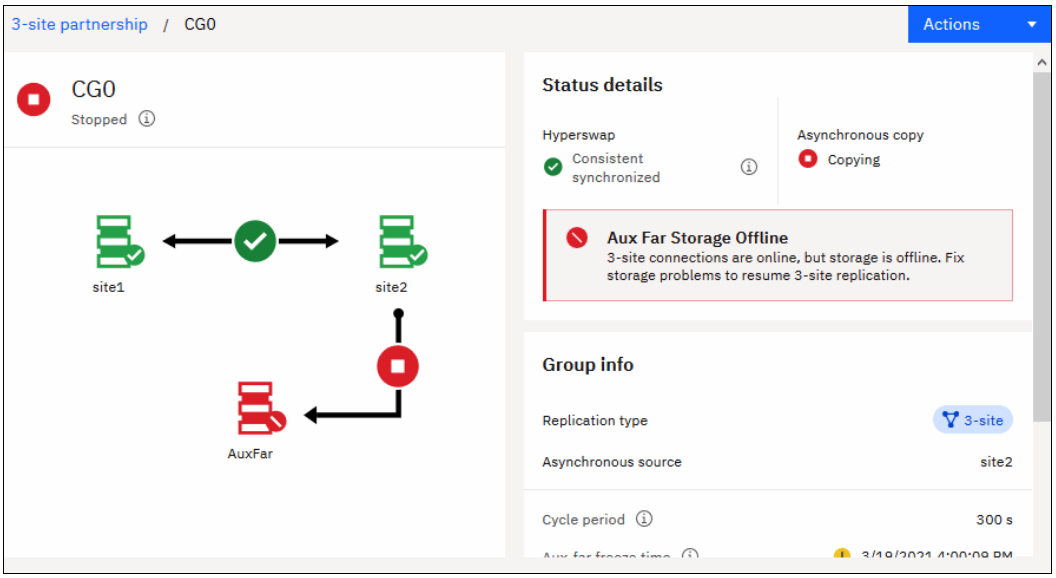

Figure 6-24 Consistency group stopped and status details indicate AuxFar Storage Offline

An event ID 050991 might also be present that indicates that the cycle time exceeded the configured cycle period, as shown in Figure 6-25.

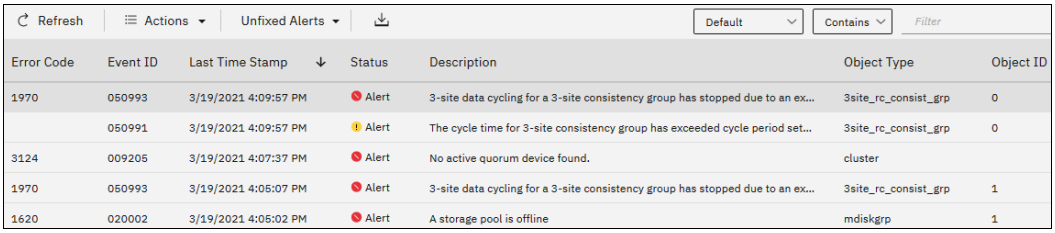

Figure 6-25 Example events during storage failure at the far site

2. Determine the cause of the storage failure and follow the Directed Maintenance Procedure (DMP) to bring the storage and volumes back online.
3. Verify the state of the consistency group by running the **`ls3sitercconsistgrp`** command or by using the GUI. The state is now `3site_consistent`, as shown in Example 6-22. Any events on the AuxFar system are fixed.

Example 6-22 Consistency group online after fixing the storage and 3-Site data cycling continues

```
[root@localhost ~]# ls3sitercconsistgrp
name state            primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  3site_consistent Master  site2                consistent_synchronized copied
2021/03/19/16/05/09 300          online
[root@localhost ~]# ls3sitercconsistgrp
name state            primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  3site_consistent Master  site2                consistent_synchronized stopped
2021/03/19/16/27/44 300          online
```

4. If the output still shows a storage failure state even after the volumes were brought back online, you can manually start data cycling by using the **`start3sitercconsistgrp`** command or the **Start group** action from the GUI.

6.5 Recovering from a site loss

A site loss can be temporary or permanent. In either case, manual intervention often is required for recovery. A site loss affects all consistency groups.

A site failure can be caused by a power outage that takes the entire site offline. Another example of a site failure is an auto-recovered cluster-wide disaster (Automatic Cluster Recovery or Tier 2 [T2] Recovery).

A site failure is possible if the site is still available, but all Fibre Channel and Ethernet connections are down. In that case, the site is not visible through synchronous and asynchronous replication links and the 3-Site Orchestrator SSH link and are considered as down.

A special case of a site failure is when the nodes lost their configuration data and the Recover System Procedure, also known as Tier 3 (T3) Recovery, must be performed. IBM Support must be involved to restore 3-Site Replication operations after a system on one of the sites has undergone a T3 Recovery. Post T3 Recovery steps are not covered in this book.

Note: In situations where the Recover System Procedure is required or the cache was lost, contact IBM Support before attempting to recover. Failure to follow the correct recovery procedures can result in data loss or unexpected system behavior. For Automatic Cluster Recoveries, check with IBM Support first to determine whether it is safe to proceed, especially if a software upgrade or downgrade was in progress.

At the time of this writing, an issue might prevent specific recovery actions (that is, T2 and T3 Recoveries) from completing successfully in systems where remote users (that is, LDAP users) are configured. Consider the following advice:

- If the system is not upgraded to 8.4.x code level but you want to upgrade:

  If available, upgrade to a 8.4.x code level, which contains the fix for APAR HU02325. A fix is targeted for the upcoming version 8.4.0.3. After the version with the fix is available, the APAR is listed in the release notes.

  As a workaround, remove all remote users *before* the upgrade and re-create them after. The Software Upgrade Test Utility blocks any systems from upgrading from pre-8.4 to affected 8.4 code levels if remote users are present.

- If the system was upgraded to 8.4.x code level:

  Action must be taken to work around the issue and make sure that system recovery works correctly in case it is required. Remove all remote users with a non-empty User Group ID and re-create them. Any remote users that were created on pre-8.4 code levels show a User Group ID of 256.

You can identify remote users by using the CLI command **`svcinfo lsuser -filtervalue remote=yes`** on the system CLI. In the GUI, click **Access** → **Users by Group** and review the Authentication column for the value `Remote`.

Note: If your system that is running on 8.4.0.x code level performed a Tier 2 (T2) or Tier 3 (T3) Recovery and is hitting asserts, contact IBM Support.

6.5.1 Determining the scope of the disaster and recovery plan

The recovery steps that are used depend on the 3-Site Replication topology and the Primary site assignment at the time of the disaster. You can use the **`ls3sitercconsistgrp`** command to determine 3-Site Replication state and status, as shown in Example 6-23.

Example 6-23 Example output for a consistency group in status online

```
[root@localhost ~]# ls3sitercconsistgrp
name state            primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  3site_consistent Master  AuxNear              consistent_synchronized copying
2021/03/09/12/40/51 300          online
```

The *primary* column indicates the site that is running the application primary, and `periodic_copy_source` shows which site is the source of asynchronous data replication to the far site. If `primary` and `periodic_copy_source` match, 3-Site Replication is running in a Star topology.

If there is a complete site failure, the consistency group status changes from `online` to one of the possible options that are listed in Table 6-4.

Table 6-4 Consistency group status if there is a site failure

3-Site data cycling status	Description
`primary_offline`	Primary site for this consistency group is down or it is isolated from the partnership (the network is unavailable and the site is unreachable from the remaining two sites).
`primary_and_periodic_src_offline`	The consistency group is configured as a Star topology (Primary and Periodic Source site match), and the Primary site is down or isolated.
`periodic_src_offline`	Non-primary Near Site that is a Periodic Source is down or isolated (Cascade topology).
`partner_offline`	Near site that is non-primary and not a Periodic Copy Source is down or isolated (Star topology and HyperSwap configuration).
`partner_dead_domain`	One of the near sites in the HyperSwap system suffered a complete loss of the I/O group, which resulted in loss of cache data and unrecoverable objects.
`auxfar_offline`	System on the AuxFar site is down or isolated.
`hyperswap_cluster_offline`	The complete HyperSwap system (site 1 and site 2) is down or isolated.

Table 6-5 summarizes the site failure behavior in a Metro Mirror configuration. It shows the `state` and `status` of the consistency group that are possible if a disaster occurs at one of the three sites.

Table 6-5 Site failure behavior in a Metro Mirror configuration

	Topology	Failure at	Primary	Periodic Source	CG state	CG status
#4	Star	Master	Master	Master	stopped	primary_and_periodic_src_offline
#2		AuxNear			2site_periodic	partner_offline
#1		AuxFar			stopped	auxfar_offline
#5	Cascade	Master		AuxNear	stopped	primary_offline
#3		AuxNear			stopped	periodic_src_offline
#1		AuxFar			stopped	auxfar_offline

Table 6-6 on page 160 summarizes the site failure behavior in a HyperSwap configuration. It shows the `state` and `status` of the consistency group that are possible if a disaster occurs at one of the three sites and at the entire HyperSwap system.

Table 6-6 Site failure behavior in a HyperSwap configuration

Failure at	Automatic change of Periodic Source	CG state	CG status
AuxFar	no change	stopped	auxfar_offline
Either Master (site 1) or AuxNear (site 2)	switches to remaining site if needed	2site_periodic	partner_offline
Whole HyperSwap system (site 1 and 2)	not applicable	stopped	hyperswap_cluster_offline

All of these scenarios are described in the following sections.

In addition to the consistency group state and status, the user can run the `ls3siteconfig` command to check the reachability of the affected sites whenever a site failure occurs.

6.5.2 Recovering from a site loss in a Metro Mirror configuration

This section includes scenarios for recovering from a site loss in a Metro Mirror configuration.

Scenario 1: Recovery from a disaster at the AuxFar site

Independent of the topology (Star or Cascade), recovery steps during a site failure of the far system in a Metro Mirror configuration remain the same. Figure 6-26 shows an example of a disaster that takes down the AuxFar site in a Cascade topology.

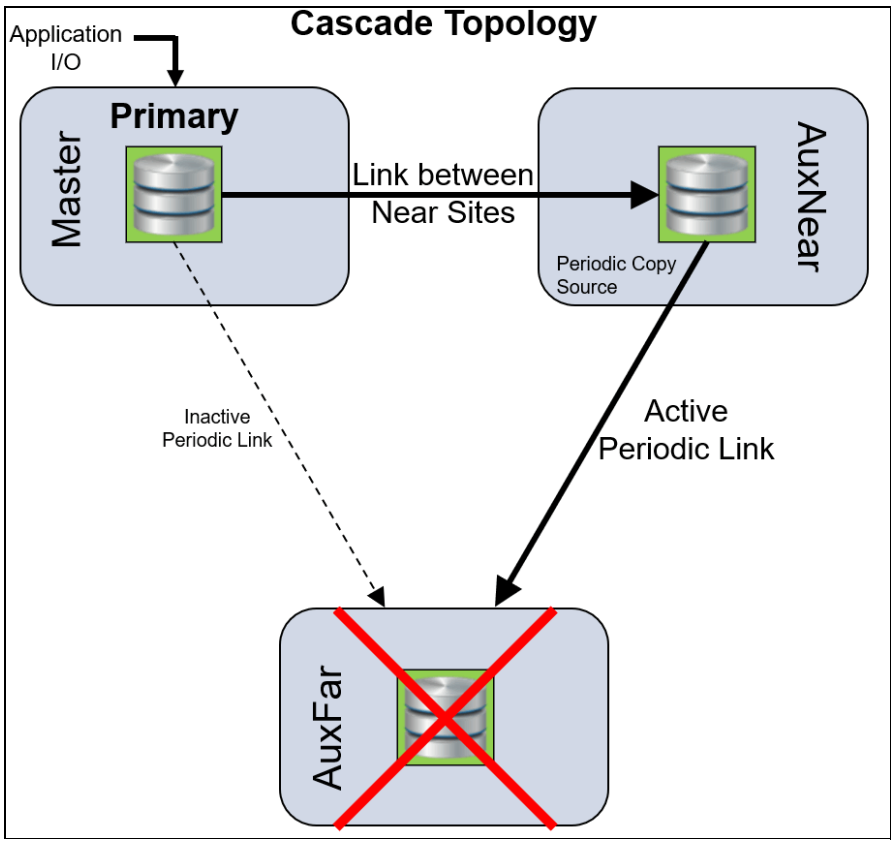

Figure 6-26 AuxFar site failure in a Cascade topology

Complete the following steps to recover:

1. Use the **ls3siteconfig** command to check the reachability of the site. It indicates an `auxfar_site_status` of `offline` and an `auxfar_ip1_status` of `unreachable`.
2. Run the **ls3sitercconsistgrp** command or use the status details page in the GUI to verify the state and status of the consistency group. If there is a failure at the AuxFar system, the state is `stopped` and the status is `auxfar_offline`, as shown in Example 6-24. An event ID 050993 is reported on the Master system, which indicates that the 3-Site data cycling stopped. Because the AuxFar site is unavailable, the event cannot be reported in the event log of the AuxFar system as usual.

Example 6-24 Consistency group output during offline AuxFar site

```
[root@localhost ~]# ls3sitercconsistgrp
name state   primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  stopped Master  AuxNear              consistent_synchronized stopped
2021/03/09/12/45/47 300          auxfar_offline
```

3. Perform the required disaster recovery procedures to recover the AuxFar system to a fully operational state. In the case that an Automatic Cluster Recovery ran, see 6.2.3, “Other required steps after an Automatic Cluster Recovery” on page 130 for more information about how to make the system accessible again from the Orchestrator host.
4. Verify the reachability of the site by using the **ls3siteconfig** command. It now shows an `auxfar_site_status` of `online` and an `auxfar_ip1_status` of `reachable`.
5. After the AuxFar site is recovered and 3-Site Orchestrator access is restored, data cycling starts automatically. If required, manually start cycling.

Scenario 2: Recovering from a disaster at non-primary AuxNear site in a Star topology

An example of a non-primary near site failure in a Star topology is shown in Figure 6-27.

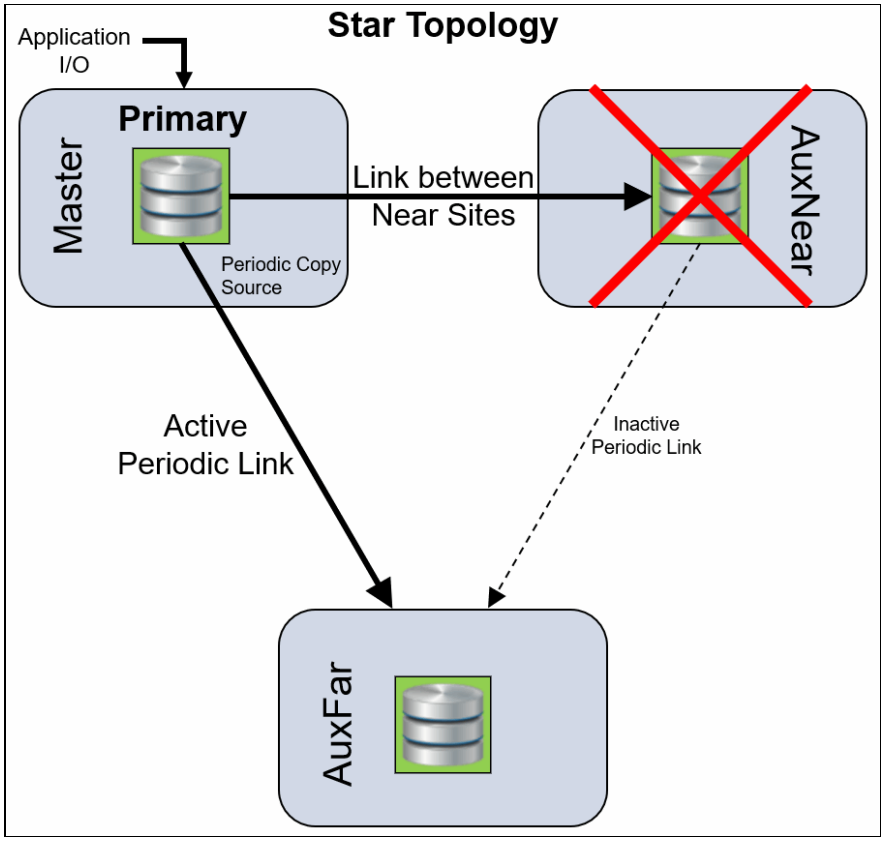

Figure 6-27 AuxNear site failure in a Star topology

To recover, complete the following steps:

1. Use the **`ls3siteconfig`** command to check the reachability of the site. It indicates an `auxnear_site_status` of `offline` and an `auxnear_ip1_status` of `unreachable`.
2. Issue the **`ls3sitercconsistgrp`** command or use the status details page in the GUI to verify the state and status of the consistency group. In the resulting output, the state of the consistency group is `2site_periodic`, as shown in Example 6-25. The status is `partner_offline` as the offline site is not the Primary nor the Periodic Copy Source. An event ID 050994 is reported on the AuxFar system to notify about the 2-Site data cycling mode of the consistency group. The GUI indicates that the consistency group is performing asynchronous replication only and that the partner site is offline.

Example 6-25 Verify consistency group information during non-primary AuxNear site failure in a Star topology

```
[root@localhost ~]# ls3sitercconsistgrp
name state          primary periodic_copy_source sync_copy_state     periodic_copy_state
freeze_time         cycle_period status
CG0  2site_periodic Master  Master               idling_disconnected stopped
2021/03/10/09/26/44 300          partner_offline
```

3. Perform recovery tasks on the AuxNear site to restore the site to a fully operational state. If an Automatic Cluster Recovery ran, see 6.2.3, “Other required steps after an Automatic Cluster Recovery” on page 130 for more information about how to make the system accessible again from the Orchestrator host.
4. To complete the recovery of the 3-Site setup, check the `sync_copy_state` in the **`ls3sitercconsistgrp`** output. Manually start the consistency group or rejoin the site as rejoin AuxNear into 3-Site data cycling as the `sync_copy_state` was `consistent_stopped`.

Example 6-26 Joining a site back into the consistency group

```
[root@localhost ~]# ch3sitercconsistgrp -join AuxNear CGO
```

5. Issue the **`ls3sitercconsistgrp`** command to verify that recovery is complete, as shown in Example 6-27. The event on the AuxFar system indicating that the consistency group is running in 2-Site data cycling mode is now fixed.

Example 6-27 Consistency group is back online and in 3site_consistent state

```
[root@localhost ~]# ls3sitercconsistgrp
name state            primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CGO  3site_consistent Master  Master               consistent_synchronized stopped
2021/03/10/09/53/42 300          online
```

Scenario 3: Recovering from a disaster at non-primary AuxNear site in a Cascade topology

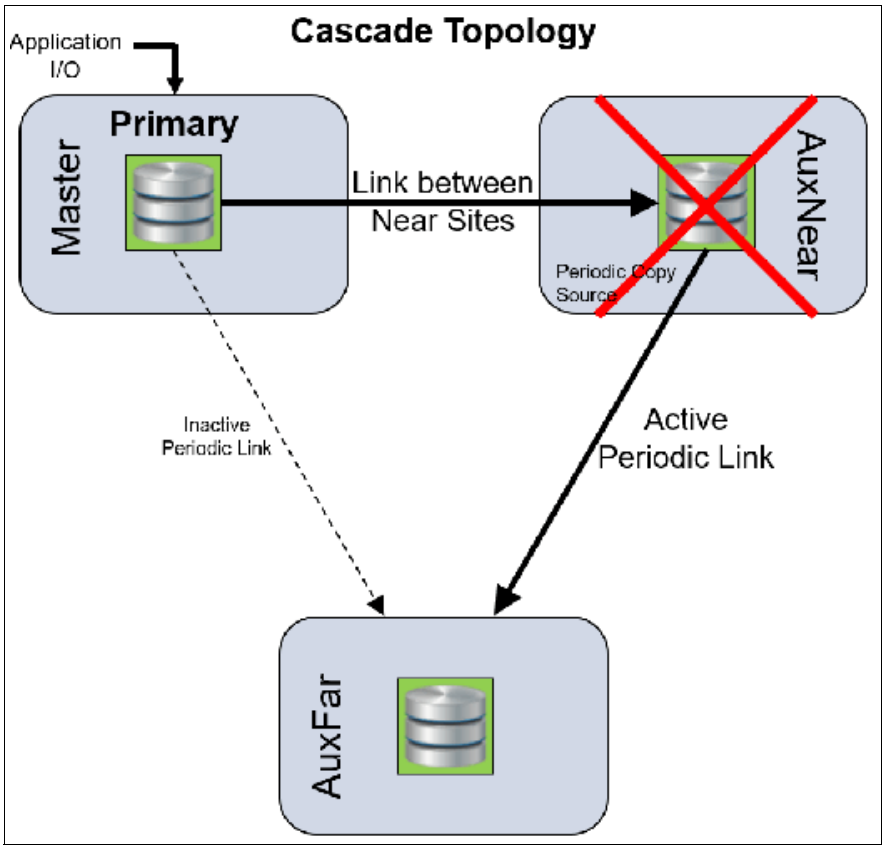

Figure 6-28 AuxNear site failure in a Cascade topology

An example of a non-primary near site failure in a Cascade topology is shown in Figure 6-28 on page 163. Complete the following steps to recover:

1. Use the **`ls3siteconfig`** command to check the reachability of the site. It indicates an `auxnear_site_status` of `offline` and an `auxnear_ip1_status` of `unreachable`.
2. Issue the **`ls3sitercconsistgrp`** command to verify the state and status of the consistency group, as shown in Example 6-28, or use the status details page in the GUI. In the resulting output, the state of the consistency group is `stopped` and the status is `periodic_src_offline`. The AuxFar system reports event ID 050993 to indicate the stopped state.

Example 6-28 Consistency group state and status during AuxNear site failure in a Cascade topology

```
[root@localhost ~]# ls3sitercconsistgrp
name state   primary periodic_copy_source sync_copy_state     periodic_copy_state
freeze_time         cycle_period status
CG0  stopped Master  AuxNear             idling_disconnected stopped
2021/03/10/10/39/44 300          periodic_src_offline
```

The GUI also indicates that the consistency group is stopped and that the asynchronous source is offline, as shown in Figure 6-29.

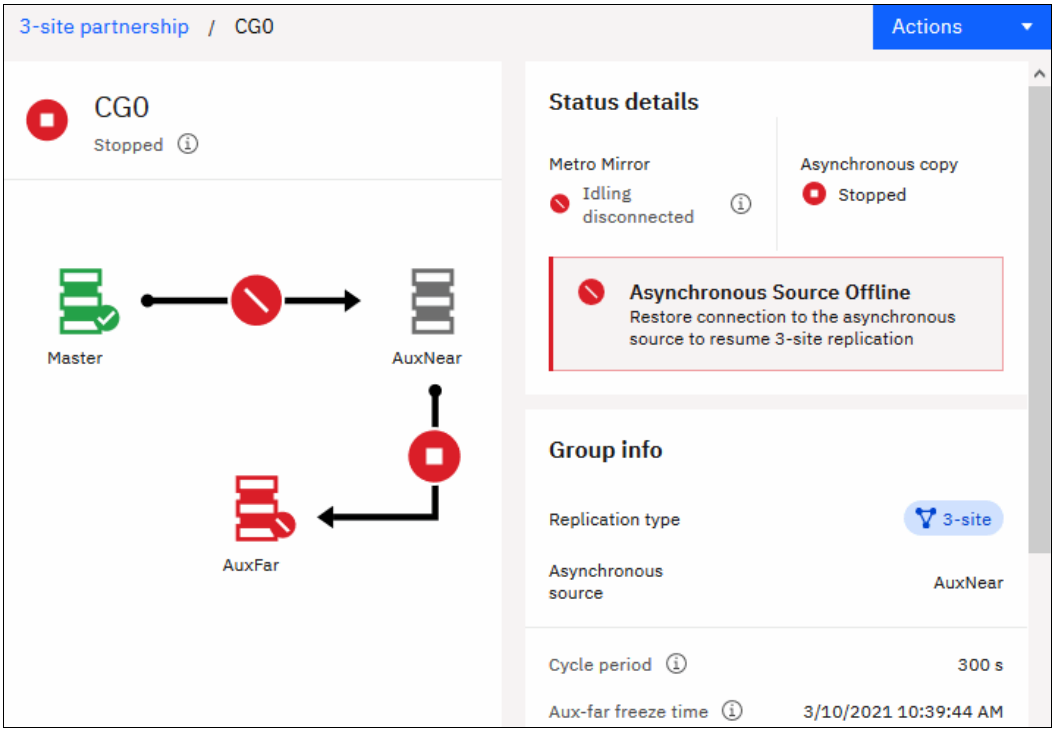

Figure 6-29 The GUI indicates that the Asynchronous Source site is offline

3. Temporarily change the Periodic Source from the failed AuxNear site to the remaining AuxNear site (that is, the Master). This option provides an up-to-date second (asynchronous) data copy although the AuxNear site is not available. You can also choose to leave 3-Site data cycling inactive for some time during which the AuxNear site experiences problems. If you decide for the second option, skip this step and step 7.

   To change the Periodic Source and start data cycling in 2-Site data cycling mode, use the **`ch3sitercconsistgrp -periodicsrc`** command to switch into `2site_periodic` state, as shown in Example 6-29 on page 165.

Example 6-29 Change Periodic Source to start cycling

```
[root@localhost ~]# ch3sitercconsistgrp -periodicsrc Master CGO

[root@localhost ~]# ls3sitercconsistgrp
name state          primary periodic_copy_source sync_copy_state     periodic_copy_state
freeze_time         cycle_period status
CGO  2site_periodic Master  Master               idling_disconnected stopped
2021/03/10/10/39/44 300          partner_offline
```

4. Perform recovery activities on the AuxNear site to restore the site to a fully operational state. If Automatic Cluster Recovery was run, see 6.2.3, “Other required steps after an Automatic Cluster Recovery” on page 130 for more information about how to make the system accessible again from the Orchestrator host.
5. After the AuxNear site is recovered, check the `sync_copy_state` in the **`ls3sitercconsistgrp`** output. Manually start the consistency group or perform rejoin action, as shown in Example 6-30 if the `sync_copy_state` should be `consistent_stopped`.

Example 6-30 Include the site back into 3-Site data cycling

```
[root@localhost ~]# ch3sitercconsistgrp -join AuxNear CGO
```

6. Issue the **`ls3sitercconsistgrp`** command to verify the recovery, as shown in Example 6-31.

Example 6-31 Confirm that the consistency group is in state 3site_consistent and status online

```
[root@localhost ~]# ls3sitercconsistgrp
name state            primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CGO  3site_consistent Master  Master               consistent_synchronized stopped
2021/03/10/10/53/51 300          online
```

7. (Optional) Change the setup back to a Cascade topology by switching the Periodic Source, as shown in Example 6-32.

Example 6-32 Change the configuration back to Cascade topology and verify

```
[root@localhost ~]# ch3sitercconsistgrp -periodicsrc AuxNear CGO
[root@localhost ~]# ls3sitercconsistgrp
name state            primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CGO  3site_consistent Master  AuxNear              consistent_synchronized idling
2021/03/10/10/53/51 300          online
```

Scenarios 4 and 5: Recovering from a disaster at the Primary site

If the Primary site system fails and read/write access to application data is required, you must use the system CLI of the surviving AuxNear site to enable access to the Metro Mirror secondary copy. Care must be taken with this procedure because if performed incorrectly, data might be lost. Consult with IBM Support before taking actions.

In both 3-Site topologies (Star and Cascade), recovery is similar. An example for a Master site failure in a Star topology is shown in Figure 6-30. For a Cascade topology, the same steps must be followed, except for the Periodic Source switch.

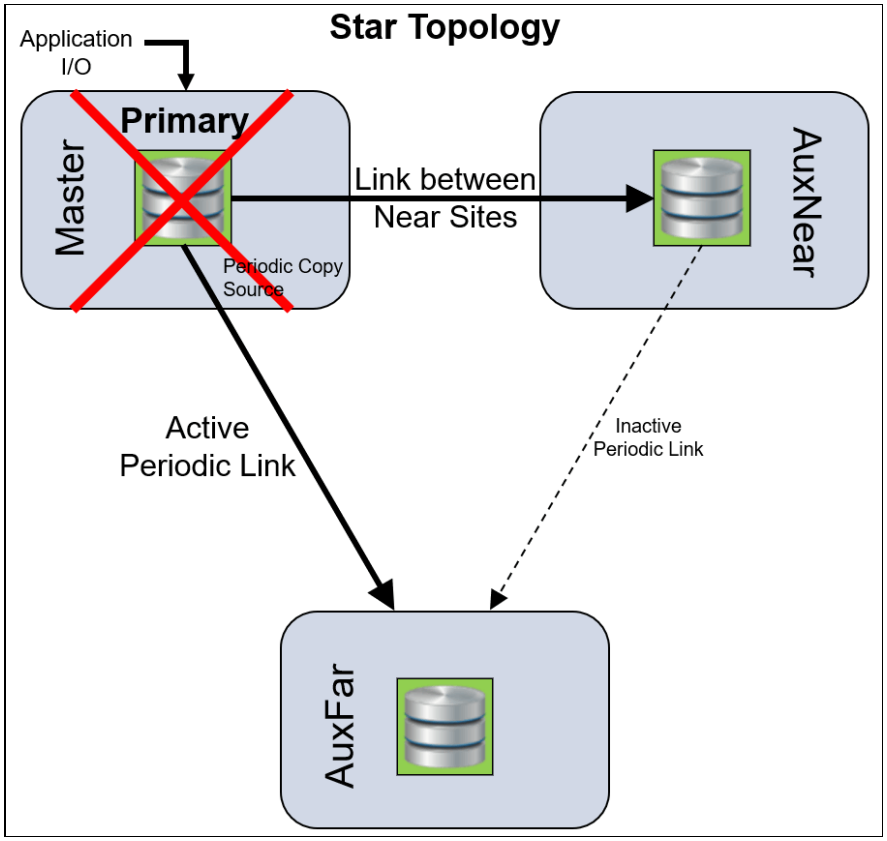

Figure 6-30 Master site failure in a Star topology

To recover from a disaster at the Master site, complete the following steps:

1. Use the **`ls3siteconfig`** command to check the reachability of the site. It indicates a `master_site_status` of `offline` and a `master_ip1_status` of `unreachable`.
2. Use the **`ls3sitercconsistgrp`** command to verify the state of the consistency group. In the resulting output, the state of the consistency group is `stopped`. Example 6-33 shows the status `primary_and_periodic_src_offline` in a Star topology.

Example 6-33 Consistency group status during Master site failure in a Star topology

```
[root@localhost ~]# ls3sitercconsistgrp
name state   primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  stopped Master  Master                consistent_disconnected stopped
2021/03/24/10/50/05 300          primary_and_periodic_src_offline
```

Example 6-34 shows the status `primary_offline` in a Cascade topology.

Example 6-34 Consistency group status during Master site failure in a Cascade topology

```
[root@localhost ~]# ls3sitercconsistgrp
name state   primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CGO  stopped Master  AuxNear              consistent_disconnected stopped
2021/03/24/10/50/05 300          primary_offline
```

3. If the Primary site fails, access to application data is disrupted. Until the site is recovered, the following options are available:
   - Option A is to not run any application I/O on the surviving site and the consistency group remains in the `stopped` state.
   - Option B is to enable read/write access to volume copies on the surviving AuxNear site by using Metro Mirror commands in the system CLI of the surviving site. The related CLI command is **`stoprcconsistgrp -access`**. For more information, see *IBM System Storage SAN Volume Controller and Storwize V7000 Replication Family Services*, SG24-7574.

     Consult with IBM support *before* performing this type of recovery action. The Metro Mirror consistency group state switches from `consistent_disconnected` to `idling_disconnected`, and the `primary` attribute is empty, as shown in Example 6-35. If the consistency group was in the `consistent_synchronized` or `consistent_stopped` state, it changes into `idling`.

Example 6-35 Enabling read and write access to aux volumes

```
IBM_FlashSystem:AuxNear:superuser>lsrcconsistgrp -delim :
id:name:master_cluster_id:master_cluster_name:aux_cluster_id:aux_cluster_name:primary:state:rela
tionship_count:copy_type:cycling_mode:freeze_time
0:CGO:0000020321E0ECE0:Master:000002042040B714:AuxNear:master:consistent_disconnected:2:metro::2
021/03/24/10/53/18

IBM_FlashSystem:AuxNear:superuser>svctask stoprcconsistgrp -access CGO

IBM_FlashSystem:AuxNear:superuser>lsrcconsistgrp -delim :
id:name:master_cluster_id:master_cluster_name:aux_cluster_id:aux_cluster_name:primary:state:rela
tionship_count:copy_type:cycling_mode:freeze_time
0:CGO:0000020321E0ECE0:Master:000002042040B714:AuxNear::idling_disconnected:2:metro::
```

4. If option A was chosen, the 3-Site configuration recovers automatically if the `sync_copy_state` is in any other state than stopped (check with the **`ls3sitercconsistgrp`** command) after the Master site is recovered. If it is `consistent_stopped`, it requires rejoining the site by using the **`ch3sitercconsistgrp -join`** command or the **Re-join Site** action in the GUI.
5. If option B was chosen, the 3-Site consistency groups in a Cascade topology switches to `2site_periodic` state, as shown in Example 6-36 on page 168, after access to volumes on the surviving site is enabled. The AuxFar system reports event ID 050994. The `primary` field in the **`ls3sitercconsistgrp`** command output is empty.

Example 6-36 Consistency group state in a Cascade topology switches to 2site_periodic after enabling access

```
[root@localhost ~]# ls3sitercconsistgrp
name state          primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CGO  2site_periodic         AuxNear              idling_disconnected     stopped
2021/03/24/11/03/50 300          primary_offline
```

For consistency groups in a Star topology, the Periodic Source must be switched by using the **ch3sitercconsistgrp -periodicsrc** command to get into `2site_periodic` state, as shown in Example 6-37

Example 6-37 After switching Periodic Source in a Star topology the consistency group changes to 2site_periodic

```
[root@localhost ~]# ls3sitercconsistgrp
name state   primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CGO  stopped         Master               idling_disconnected     stopped
2021/03/24/10/50/05 300          primary_and_periodic_src_offline

[root@localhost ~]# ch3sitercconsistgrp -periodicsrc AuxNear CGO

[root@localhost ~]# ls3sitercconsistgrp
name state          primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CGO  2site_periodic         AuxNear              idling_disconnected     stopped
2021/03/24/11/19/32 300          primary_offline
```

After the Master site is fully recovered, you must decide carefully if you want to switch the Primary and the Periodic Source back to the original configuration (and use the data on the Master volumes from before the disaster) or if you want to proceed with the current setup (and the data on the AuxNear volumes with all the updates since the disaster, which also were copied over to AuxFar).

Start the Metro Mirror consistency group from the system CLI and indicate the Primary (the volumes that are made accessible for write I/O, the possible options are `master` or `aux`) that you want to use. Indicate `aux` if you want to keep the current setup and use the AuxNear volumes, or specify `master` if you want to revert the data to that on the original volumes at the Master site. You cannot recover the data after choosing the incorrect site because the copy on AuxFar features the same data. As shown in Example 6-38, you can choose to continue with the current setup and select AuxNear as the Primary. The `-force` flag is required.

Example 6-38 Starting the Metro Mirror consistency group with AuxNear volumes as Primary

```
IBM_FlashSystem:AuxNear:superuser>svctask startrcconsistgrp -force -primary aux CGO
```

After the `sync_copy_state` reaches `consistent_synchronized`, 3-Site data cycling continues automatically.

Recovering from dual site failures

A dual site failure of both Near Sites can be caused by different situations. Some situations are recoverable and data cycling restarts after the problem is fixed.

One example of a dual site failure is a 3-Site Orchestrator SSH communication problem with the Master and AuxNear system. In that case, you must fix the network connection and data cycling resumes automatically.

Another example is when the Master and the AuxNear system suffer from an Automatic Cluster Recovery (T2 Recovery) because of a software issue. After reenabling CLI access, the Orchestrator can run commands again and data cycling continues automatically.

Both sites might also suffer from a power outage without the ability to save their hardened data, and a Tier 3 Recovery is required.

This section covers the situation where Master and AuxNear site cannot recover. Use the data on the AuxFar site to rebuild the objects and data to resume applications.

To recover from dual disasters at both Near Sites (as shown in Figure 6-31), complete the following high-level steps:

Important: This procedure is performed only with assistance from IBM Support.

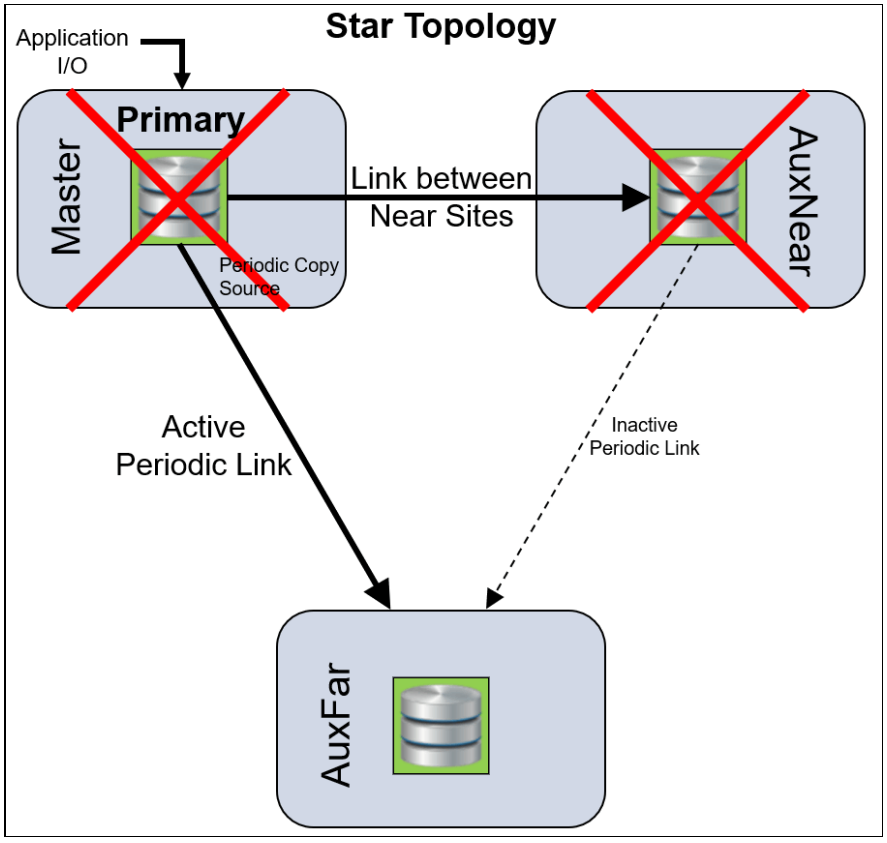

Figure 6-31 Two Near Sites are down at the same time

1. When that condition occurs, the 3-Site consistency group is in the `stopped` state. Verify this state by using the **`ls3sitercconsistgrp`** command. The AuxFar system reports an event ID 050993 in the event log to indicate the stopped state.

2. Use the **convertrcconsistgrp -type 2site** Orchestrator command followed by the name of the 3-Site consistency group to make all the AuxFar site volumes accessible and convert the 3-Site consistency group into a 2-Site consistency group.

   Alternatively, use the **Revert to 2-Site** option on the status details page of the consistency group in the GUI. 3-Site data cycling must be stopped before conversion. This command cleans up the 3-Site configuration where possible on the reachable sites. Any remaining objects that cannot be deleted by Orchestrator must be deleted manually with the assistance of IBM Support. The consistency group is now no longer listed in the output of **ls3sitercconsistgrp** and not shown in the GUI on the **3-Site partnership** page.
3. Recover both Near Sites (Master and AuxNear) by using the manual procedure.
4. Recover data from the AuxFar site default volume to a new volume on either of the two Near Sites by using a 2-Site Remote Copy service (that is, Metro Mirror or Global Mirror). Establish a Remote Copy relationship from the far site volume to a new volume on one of the Near Sites. Then, after the copy to the Near Site is complete, establish a Metro Mirror relationship with the alternative Near Site and wait until it is synchronized.
5. Optionally, you can again move to 3-Site data cycling by following the standard procedure of converting a 2-Site consistency group into a 3-Site consistency group by using the 3-Site Orchestrator. For this process, a new 3-Site relationship between the Near Site volume and an *empty* volume on the AuxFar site must be created. After a synchronized copy of data is available on the AuxFar site new volumes, delete the old volumes.

Important: Contact IBM Support *before* attempting this procedure, and before you attempt to recover from a dual disaster at any other two sites.

6.5.3 Recovering from a site loss in a HyperSwap configuration

During a site failure in a HyperSwap configuration, all nodes of an I/O group in a specific site (Master or AuxNear or AuxFar) are not available. For example, this issue can exist because of a sudden shutdown of both nodes if an input power loss occurs or an Automatic Cluster Recovery occurs during which all nodes restart at the same time because of software failures.

If one of the Near Sites in a HyperSwap system fails, the system continues to operate on the remaining site. Servers on the failed site without high availability (HA) functions stop. Servers with HA functions are restarted from the HA software. Servers on the remaining site continue to operate. Because the cache is fully independent on each site, the system has full performance even if one site is lost.

If the HyperSwap volume was fully synchronized at the time of the failure, the state of the active-active relationship (`sync_copy_state` in the **ls3sitercconsistgrp** 3-Site Orchestrator command output) changes to `consistent_copying`. If the volume was the Primary copy (running the primary applications), HyperSwap switches the copy direction to make the online copy the new Primary.

The *master* volume remains online and the *aux* volume remains offline, regardless of which copy is no longer accessible (the online copy is presented to the host with the unique ID of the master volume). After the offline copy is restored, a resynchronization occurs and the synchronous copy state again becomes `consistent_synchronized`.

After an Automatic Cluster Recovery, the system continues operation; however, it is required to re-enable CLI commands before 3-Site data cycling continues automatically. The 3-Site consistency group is in `stopped` state. For more information, see 6.2.3, “Other required steps after an Automatic Cluster Recovery” on page 130.

In a HyperSwap configuration, the state transitions of the 3-Site consistency groups during a site failure are similar to those transitions during link failures.

Recovery depends on the status of the nodes after the failure and if any data has been lost. In some situations, assistance from IBM Support is required to avoid a possible data loss when recovery is completed incorrectly.

Recovery actions that require assistance from IBM Support

If a Near Site dual disaster or a Tier 3 Recovery on the HyperSwap system occurs, the 3-Site consistency group changes to `stopped` state. Figure 6-32 shows a complete failure of the HyperSwap system. This failure results in both Near Sites (which are part of the 3-Site Replication) being unavailable.

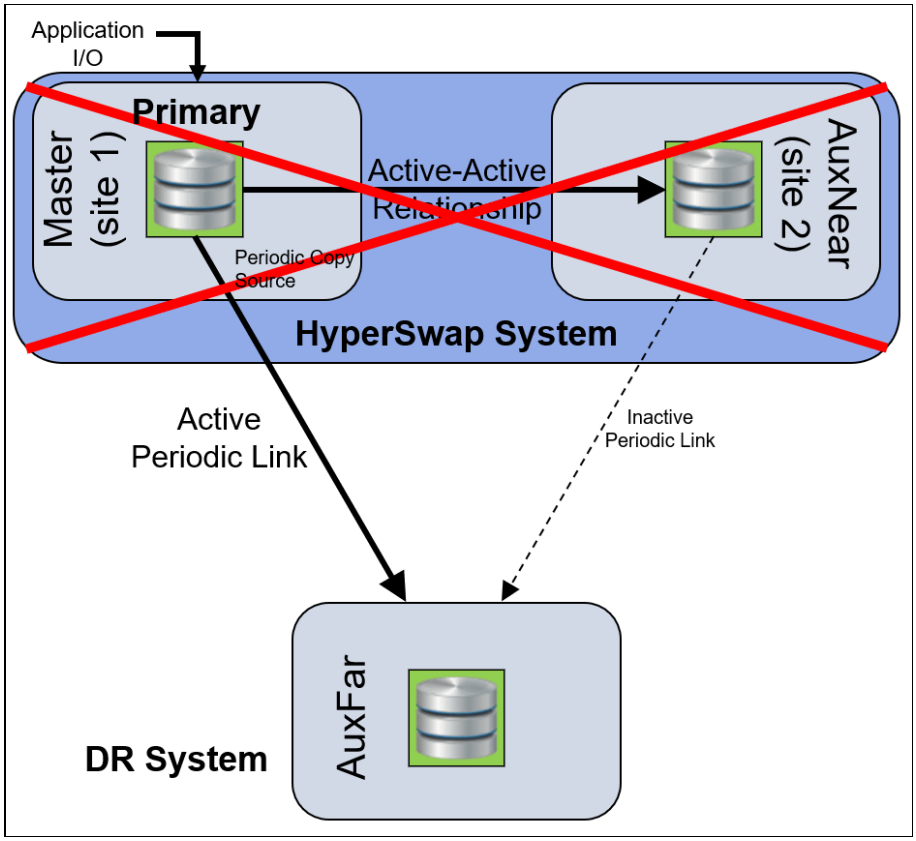

Figure 6-32 Complete failure of the HyperSwap system resulting in both Near Sites down

To recover from a persistent failure of the entire HyperSwap system (Master and AuxNear sites are lost), complete the following steps with assistance from IBM Support:

1. Use the **`ls3sitercconsistgrp`** command to verify that the consistency group is in `stopped` state.
2. Use the **`convertrcconsistgrp -type 2site`** Orchestrator command to make the AuxFar site volumes accessible and convert the 3-Site consistency group into a 2-Site consistency group.
3. Perform the Recover System Procedure. Intra-system Metro Mirror relationships are restored if all dependencies were successfully restored to their original I/O groups.
4. Restore the 3-Site configuration by using a script that is run on the Orchestrator host.

If a Tier 3 Recovery on the AuxFar system occurs, the 3-Site data cycling must be stopped before running the Recover System Procedure by using the **stop3sitercconsistgrp** command. After you verify that the consistency group is `stopped`, perform the Recover System Procedure. With the assistance from IBM Support, restore the configuration, access to the data, and replication objects (including partnerships, consistency groups, and relationships).

If unrecoverable objects exist because of a loss of cache data when a complete I/O group is lost (on any Near Site in a HyperSwap system or on the AuxFar system), IBM Support can assist with a script that is run on the 3-Site Orchestrator to restore the 3-Site configuration.

This situation can occur if both nodes in an I/O group lose their most up-to-date metadata. Both nodes might show node error 878 or one node shows this error and the other node does not have the most recent data as it shutdown first.

The system reports a 1940 event, which indicates that the HyperSwap volume lost synchronization, followed by error codes `1700 Unrecovered Metro Mirror or Global Mirror relationship` and `1895 Unrecovered FlashCopy mappings`. A recreation of the relationships and FlashCopy mappings is required to recover.

Another use case is when an offline volume copy exists because of corrupt metadata (error code 1862), which also requires the removal and recreation of the volume. Complete the following steps:

1. Verify the state and status of the consistency group by using the **ls3sitercconsistgrp** command. For a loss of an I/O group in a HyperSwap system, the 3-Site consistency group is in `2site_periodic` state and the status is `partner_dead_domain` instead of `partner_offline`. The consistency group state after a loss of an I/O group on the AuxFar system is `stopped`.
2. Get the current configuration backup and generate a snap with live dumps (option 4).
3. Contact IBM Support to recover the I/O group by removing and re-creating the related objects from the failed site. This process involves stopping any generic fcmaps that are not `stopped` or `idle_or_copied`. Then, the fcmaps are removed.

   The volumes come online, but the data on the target is unusable. Now, the volume copies in the failed site (including the relationship) are deleted by using the **rmvolumecopy -site** command. The 1700 and 1895 errors are automatically marked as fixed and the 1940 error disappears after the last relationship is deleted. Later, the relationships are re-created by using the **addvolumecopy** command, which completes asynchronously. The new volume copies synchronize automatically.
4. Restore the 3-Site configuration by using a script that is run on the Orchestrator host.

Recovering from a disaster at the AuxFar site

Figure 6-33 shows a failure of the far site in a HyperSwap configuration.

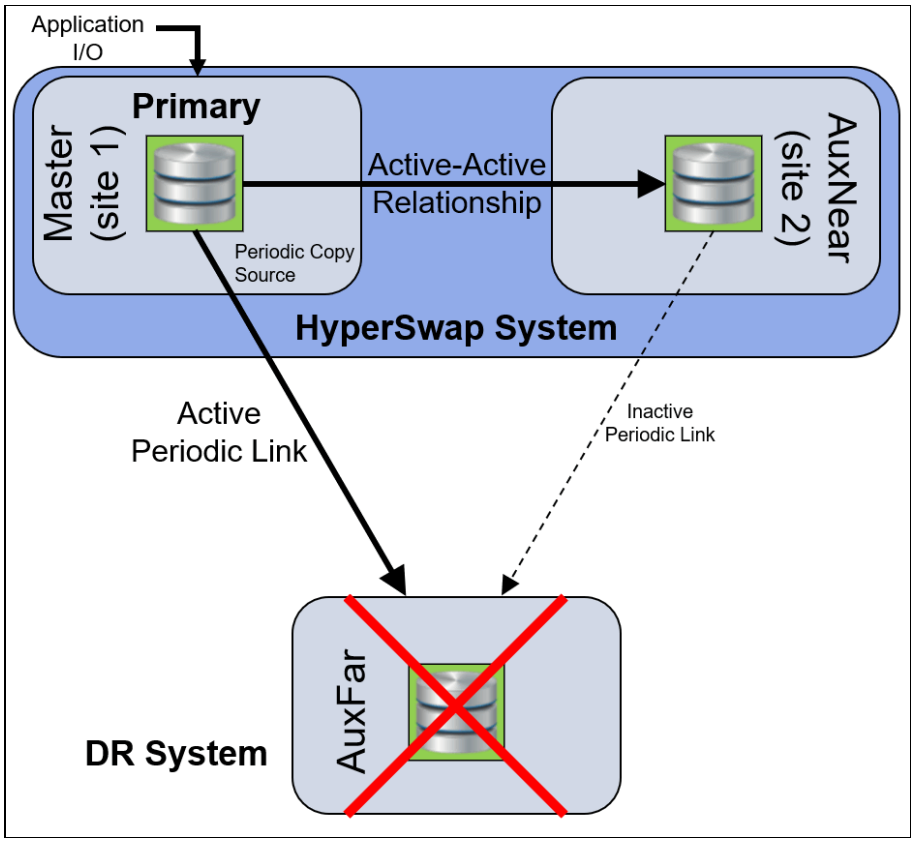

Figure 6-33 AuxFar site failure in a HyperSwap configuration

To recover from a disaster at the AuxFar site, complete the following steps:

1. Run the **ls3siteconfig** command to check the reachability of the site. It shows an `auxfar_site_status` of `offline` and an `auxfar_ip1_status` of `unreachable`.
2. Verify the state and status of the consistency group by using the **ls3sitercconsistgrp** command or check the status details page in the GUI. In the resulting output, the state is `stopped` and the status is `auxfar_offline`, as shown in Example 6-39. In the event log of the HyperSwap system (because the AuxFar system is unavailable), an event ID 050993 is reported, which indicates that 3-Site data cycling stopped.

Example 6-39 Consistency group state and status during disaster at AuxFar site

```
[root@localhost ~]# ls3sitercconsistgrp
name state   primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  stopped Master  site2                consistent_synchronized stopped
2021/03/16/13/38/53 300          auxfar_offline
```

3. Complete the required recovery actions to restore the site to a fully operational state. If an Automatic Cluster Recovery ran, see 6.2.3, “Other required steps after an Automatic Cluster Recovery” on page 130 for more information about how to make the system accessible again from the Orchestrator host.
4. After the site is recovered, use the **ls3siteconfig** command to verify that the `auxfar_site_status` is now `online` and the `auxfar_ip1_status` is `reachable`.

5. Verify that the site is automatically included in 3-Site data cycling. Use the **`ls3sitercconsistgrp`** command to confirm that the state of the consistency group is `3site_consistent` and the status is `online`, as shown in Example 6-40. The 050993 event on the HyperSwap system is automatically marked as fixed. Only the event regarding the connection loss to the remote cluster is fixed manually on the HyperSwap system and the AuxFar system.

Example 6-40 Consistency group online after recovery of AuxFar and data cycling continues

```
[root@localhost ~]# ls3sitercconsistgrp
name state            primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  3site_consistent Master  site2                consistent_synchronized stopped
2021/03/16/13/38/53 300          online
```

Recovering from a disaster at a Near Site (Master or AuxNear)

You can recover from a disaster at either of the Near Sites in a HyperSwap configuration. An example for a Near Site failure in a HyperSwap configuration is a failure of the Master site (site 1), as shown in Figure 6-34.

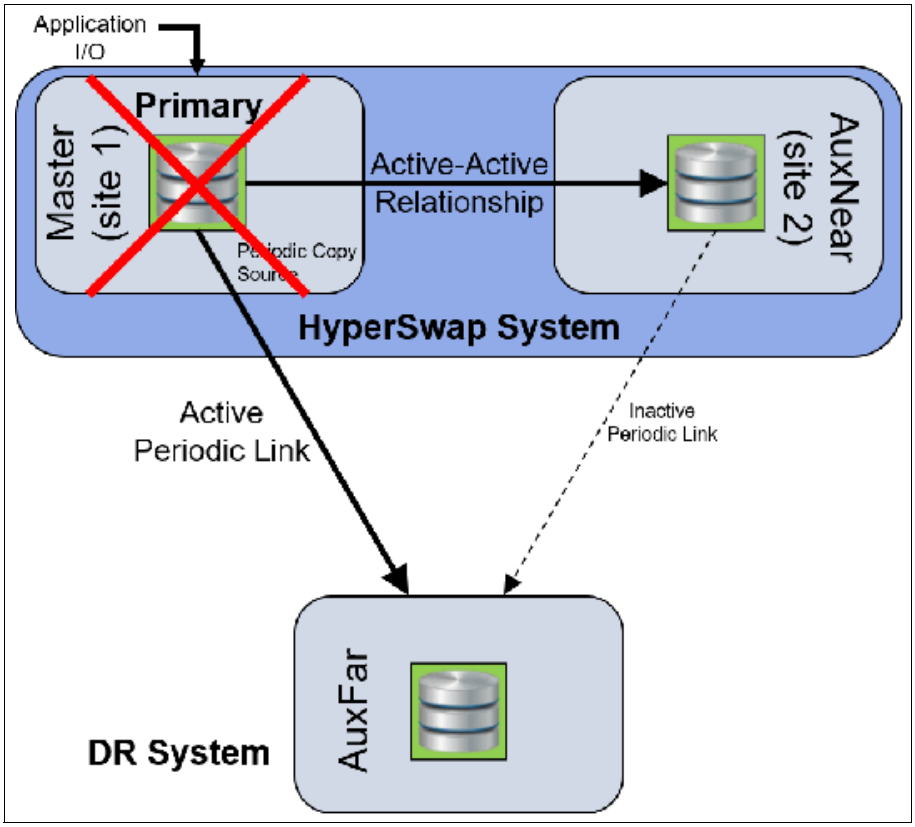

Figure 6-34 Near site failure in a HyperSwap configuration

Complete the following steps:

1. Confirm the reachability of the sites by using the **`ls3siteconfig`** command. Check the site status and the IP status for Master and AuxNear to identify which site failed. For example, if site 1 (Master) failed, the `master_site_status` is `offline` and the `master_ip1_status` is `unreachable`.

2. Verify the state and status of the consistency group by using the **ls3sitercconsistgrp** command or use the status details page in the GUI. In the resulting output, the state is `2site_periodic` and the status is `partner_offline`, as shown in Example 6-41.

Example 6-41 Periodic source automatically switches from site 1 to site 2 during a disaster at the Master site (site 1)

```
[root@localhost ~]# ls3sitercconsistgrp
name state            primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  3site_consistent Master  site1                consistent_synchronized stopped
2021/03/16/13/09/01 300          online

[root@localhost ~]# ls3sitercconsistgrp
name state          primary periodic_copy_source sync_copy_state    periodic_copy_state
freeze_time         cycle_period status
CG0  2site_periodic Master  site2                consistent_copying stopped
2021/03/16/13/13/43 300          partner_offline
```

The Periodic Copy Source changes to the current Primary site, if required. In this example, it changes from site 1 to site 2. The output of the **svcinfo lsrcconsistgrp** command that is run on the system CLI shows that the `primary` for the active-active relationship is now `aux`. The AuxFar system reports event ID 050994, which indicates that the consistency group is operating in 2-Site data cycling mode. The GUI shows that the partner site is offline and the consistency group is performing asynchronous replication only.

3. Follow troubleshooting procedures to bring the nodes back online. They are automatically added back to the system after all nodes for the site are available again.
4. Verify the state and status of the consistency group by using the **ls3sitercconsistgrp** command to confirm that the recovery completed. In the resulting output, the state is `3site_consistent` and the status is `online`, as shown in Example 6-42.

Example 6-42 Consistency group online after site 1 is available again, site2 remains Periodic Copy Source

```
[root@localhost ~]# ls3sitercconsistgrp
name state            primary periodic_copy_source sync_copy_state         periodic_copy_state
freeze_time         cycle_period status
CG0  3site_consistent Master  site2                consistent_synchronized stopped
2021/03/16/13/18/48 300          online
```

No switch back of the Periodic Copy Source occurred after the failed site is repaired. Data cycling continues automatically in 3-Site data cycling mode. The events about the offline nodes and storage pool are already marked as fixed. The event that indicates 2-Site data cycling mode on the AuxFar system is fixed automatically as well.

As opposed to Metro Mirror configurations, no event exists about lost connection to the remote cluster if a Near Site failure occurs in a HyperSwap configuration.

6.6 Recovering from a 3-Site Orchestrator failure

3-Site Orchestrator is a key component of the 3-Site Replication solution, which automatically coordinates the replication to the AuxFar site. It is used to configure and manage the replication configurations.

This section describes the required steps to recover from a connection loss between the 3-Site Orchestrator and any of the systems at the three sites. It also covers how to recover from a failure of the external system that is hosting the 3-Site Orchestrator.

With the release of Orchestrator version 3.0, which supports Metro Mirror and HyperSwap configurations, an installed Orchestrator package can now be updated to a newer version. Information in this section also describes how to recover the communication between the 3-Site Orchestrator and the three managed systems if a connection loss occurs after the Orchestrator is updated.

Sometimes you might observe that the Orchestrator is not responding to a command and is returning a command-line interface message that indicates that it is busy, as shown in Example 6-43.

Example 6-43 Orchestrator returning a message that it is busy

```
[root@localhost ~]# ls3sitercconsistgrp
CMMVC9481E 3-Site orchestrator is busy processing another command.
```

This issue can occur during snap collection, but also during normal operation. Run the command again or run later. This behavior can also be observed in the GUI. A task might be run several times before it completes successfully.

Figure 6-35 shows the task to convert a 2-Site consistency group to 3-Site with the information that Orchestrator is busy and that the task is to be retried.

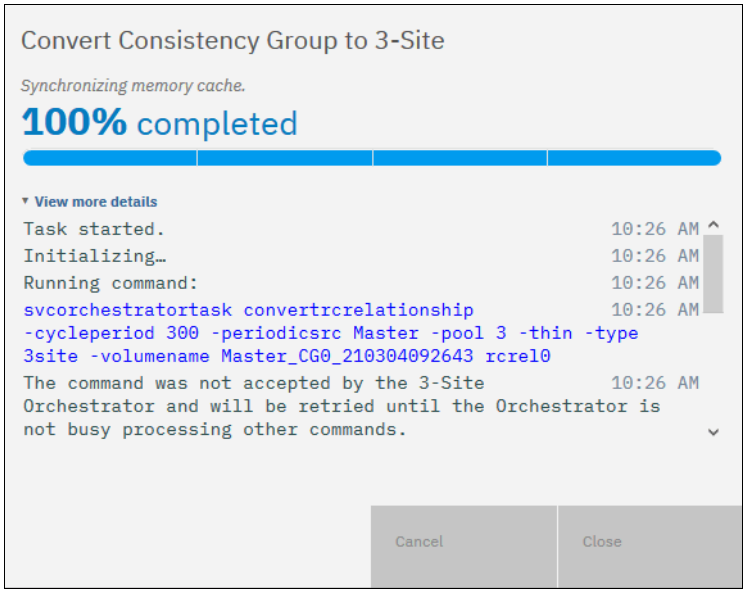

Figure 6-35 Task is to be retried because Orchestrator is busy

6.6.1 3-Site Orchestrator connectivity issues

A 3-Site Orchestrator connection failure might occur if the Orchestrator host machine crashed or failed, when the Management IP address of a system changed, or when IP connectivity issues exist between the Orchestrator host machine and the Management IP address. It can also occur when the 3-Site administrator user is removed or the key that is used for SSH connection is not available.

After the communication between the 3-Site Orchestrator and its managed systems is unavailable, 3-Site data cycling that is in progress completes the current cycle, but no new cycling is started after successful synchronization. 3-Site objects cannot be managed in the usual way. The GUI and CLI cannot be used to verify the state and status of the 3-Site consistency groups or to perform any changes except for a conversion from 3-Site to 2-Site. For this reason, the Orchestrator must be reactivated immediately.

If the 3-Site Orchestrator SSH connection to any site fails, complete the following steps:

1. Ensure that the 3-Site Orchestrator is functioning and that the 3-Site services are running correctly. Use the **`systemctl status`** Linux command followed by the service name (such as **`systemctl status rc3site.tstaskd.service`**) to check if any of the 3-Site services are stopped or inactive. If required, restart the Orchestrator services or the Orchestrator host.

   For more information about the 3-Site Orchestrator services, see Chapter 3, “Configuring a 3-Site Replication solution” on page 37.

2. Check the connectivity between the 3-Site Orchestrator and the sites by using the **`ls3siteconfig`** command, as shown in Example 6-44. Look for the IP status fields for Master, AuxNear and AuxFar. This information tells you whether the IP address is `unreachable` or in status `authentication_failed`. The expected status is `reachable`. In our example, the status is `unreachable` and the site is still online. If the site was down when the connection loss occurred, the output shows a site status of `offline`.

Example 6-44 Master IP address is unreachable

```
[root@localhost ~]# ls3siteconfig
master_name Master
master_site_name
master_site_status online
master_ip1 9.71.21.11
master_ip1_status unreachable
master_ip2
master_ip2_status not_configured
master_port 6000
auxnear_name AuxNear
auxnear_site_name
auxnear_site_status online
auxnear_ip1 9.71.21.21
auxnear_ip1_status reachable
auxnear_ip2
auxnear_ip2_status not_configured
auxnear_port 6001
auxfar_name AuxFar
auxfar_site_status online
auxfar_ip1 9.71.21.39
auxfar_ip1_status reachable
auxfar_ip2
auxfar_ip2_status not_configured
auxfar_port 6002
username 3SiteOrch
keyfile /root/.ssh/3site
topology metromirror
```

In the GUI, the 3-Site partnership shows an error that one or more sites are unreachable, as shown in Figure 6-36. Because various errors can occur between the management GUI and the 3-Site Orchestrator, the message can also indicate an SSH authentication or connection failure instead. You cannot see the 3-Site consistency groups.

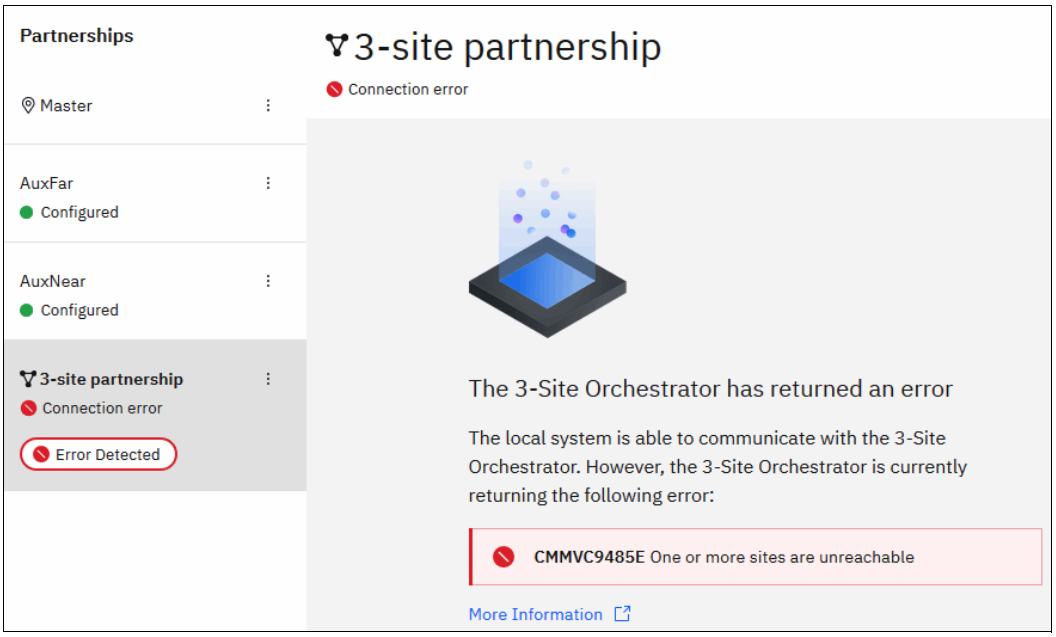

Figure 6-36 Orchestrator connectivity issue visible in the Remote Copy view of the GUI

3. If 3-Site consistency groups were created, all of the 3-Site Orchestrator commands, except for the **`convertrcconsistgrp`** command, return the error that is shown in Example 6-45.

Example 6-45 Error message during 3-Site Orchestrator disconnect

```
CMMVC9507E SSH Connection failed with site [Master_cluster_ip]
```

Because of the failed SSH connection, you cannot manage the 3-Site consistency groups.

4. In some special cases (for example, if the Management IP address of a system changed) an event ID 050992 is reported. This event indicates that the 3-Site data cycling stopped because the system lost connection with the 3-Site Orchestrator.

   The event is shown in the event log of all systems that lost connectivity with the Orchestrator host (if more than one system is affected). This event clears automatically when the SSH connection between the Orchestrator and the affected system is restored. Figure 6-37 shows an example of a fixed Orchestrator communication error.

Error Code	Event ID	Last Time Stamp	Status	Description	Object Name
1699	050992	3/8/2021 10:40:36 AM	Alert	3-site data cycling is stopped because the system has lost connection to the orchestrator	Master

Figure 6-37 Orchestrator connection was lost, event fixed automatically

As a result of this event, the 3-Site consistency groups move to the `stopped` state and an event ID 050993 is reported on the AuxFar system. Data cycling automatically restarts when the SSH connections between the affected sites are recovered.

5. Recover the SSH connections to any site. Verify the network connection between the 3-Site Orchestrator and the sites and ensure that the IP addresses of the managed systems are specified correctly. The connection might be lost because of an authentication failure or IP network connection problems, which are caused by system outages, physical networking issues, or firewall configuration errors.

To identify the cause of the issue, you can perform the following checks:

- Verify the network connectivity by using the **ping** command.
- To check that the SSH connection is working correctly from the Orchestrator host to the failed system, you can start an SSH session by using the private key file and the configured user name for 3-Site Replication, as shown in Example 6-46.

Example 6-46 Establishing an SSH session from the Orchestrator to the Master system

```
[root@localhost ~]# ssh -i /root/.ssh/3site 3SiteOrch@9.71.21.11
IBM_FlashSystem:Master:3SiteOrch>
```

If you receive a warning that the authenticity of the host cannot be established, you must use the **ssh-keyscan** command again to add the IP address to the list of known hosts, as shown in Example 6-47. This command is needed, for example, after a T3 Recovery occurred on a system, or if the Management IP address was changed intentionally.

Example 6-47 Adding the system to the list of known hosts during an SSH connection

```
[root@localhost ~]# ssh -i /root/.ssh/id_rsa 3SiteOrch@9.71.21.11
The authenticity of host '9.71.21.11 (9.71.21.11)' can't be established.
ECDSA key fingerprint is SHA256:/IJWplF+aJy5sA2YpqSS2l8DwbgxZjo5s/eCDzZbzJw.
ECDSA key fingerprint is MD5:fc:88:2c:f4:d4:5a:90:da:0e:38:87:f7:65:7d:0b:f2.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added '9.155.122.124' (ECDSA) to the list of known hosts.
IBM_FlashSystem:Master:3SiteOrch>

[root@localhost ~]# ssh-keyscan 9.71.21.11 >> /root/.ssh/known_hosts
# 9.71.21.11:22 SSH-2.0-OpenSSH_8.0
# 9.71.21.11:22 SSH-2.0-OpenSSH_8.0
# 9.71.21.11:22 SSH-2.0-OpenSSH_8.0
```

If the Management IP address was changed intentionally, you can specify the new address by using the **ch3siteconfig** command after adding the new IP address to the list of known hosts.

- From the system CLI of the failed site, you can run the **svcorchestratorinfo ls3siteconfig** command to see if the Orchestrator command is processed and provides the output or if it returns an error.

If the system IP address can be pinged from the Orchestrator host; if the correct SSH key is used; if the ports are open and the 3-Site Orchestrator services are running; and if the SSH connection from Orchestrator to the system is working in both directions; then the Orchestrator might still report a status of `unreachable`.

In most cases, a restart of the configuration node in the failed site fixes the problem and connectivity to the Orchestrator host is restored. The node restart re-initializes the network interfaces, ends all open SSH sessions, and restarts the services. This process recovers from most of the problems.

A similar behavior in the GUI and CLI can be observed after a code update of a Storage Virtualize system, which is part of the 3-Site configuration. The GUI indicates that the `connection to the 3-Site Orchestrator is offline`. If you recently updated your system to the 8.4.0.x release and a 3-Site configuration exists that used an earlier version of the Orchestrator installation package, you must update to Orchestrator version 3.0 to use the management GUI. After the connection is restored, verify that the error condition is no longer present.

6.6.2 Recovering from a 3-Site Orchestrator down situation

After an Orchestrator failure, the Orchestrator services automatically restart and reconstruct the configuration that is based on the existing configuration and the state of the 3-Site objects. However, you might want to perform a migration or complete reinstallation if the 3-Site Orchestrator cannot be recovered.

A 3-Site Orchestrator migration might be necessary if the host system fails and cannot be recovered, or if it becomes necessary to migrate the 3-Site Orchestrator instance to another host.

It might be required to reinstall 3-Site Orchestrator if the 3-Site Orchestrator installation is corrupted, the host fails, or a similar catastrophic event occurs.

Important: To successfully detect the current configuration, the 3-Site Orchestrator requires access to all three Storage Virtualize systems during 3-Site Orchestrator start or restart. If one of the systems is not responding, the 3-Site Orchestrator cannot detect or manage the current environment.

To migrate or reinstall 3-Site Orchestrator, complete the following steps:

1. Verify the installation ports for the 3-Site Orchestrator. You can use the **`ls3siteconfig`** command to view the ports configured for Master, AuxNear, and AuxFar.
2. Stop the Orchestrator services (or if required, uninstall the 3-Site Orchestrator package) to shut down the active 3-Site Orchestrator instance. The replication on each 3-Site consistency group can be stopped by using the **`stop3sitercconsistgrp`** command. Then, verify that the state is **`stopped`** by using the **`ls3sitercconsistgrp`** command.
3. Install the 3-Site Orchestrator package on the new host where migration of the host is expected. Ensure that you choose the correct installation package depending on the configuration you are running (Metro Mirror or HyperSwap) and the installed Storage Virtualize release.
4. Configure SSH key-based password-less authentication between the new host and the systems.
5. Because no `config.xml` exists on the new host, create the 3-Site Orchestrator configuration by using the **`mk3siteconfig`** command. All three storage systems must be available at this point.
6. After the configuration is complete, restart the 3-Site Orchestrator host or the Orchestrator services (eventd, tstaskd, and tsmiscservicesd). This restart recovers of all the 3-Site objects on the target 3-Site Orchestrator host and starts data cycling for all available 3-Site Orchestrator consistency groups.

A new 3-Site Orchestrator can be taken into operation only if all three sites are reachable by using SSH. This issue might not be the case in a DR situation. Therefore, a standby 3-Site Orchestrator can be planned and installed.

However, only one instance of the 3-Site Orchestrator to manage 3-Site configuration should be running at any time. Multiple instances can result in unexpected data consistency or an inaccurate state of the 3-Site configuration and 3-Site consistency groups.

Note: In the unexpected case where the Orchestrator fails to restore the 3-Site configuration after a restart or crash, it reports a CLI message `CMMVC9564E 3-Site Orchestrator crash recovery failed`. This issue can occur, for example, after a Tier 3 (T3) Recovery was performed and a conflict exists in of the available 3-Site objects. Contact IBM Support to identify the problem.

6.6.3 Recovering from a 3-Site Orchestrator update failure

The Orchestrator update includes removing the older version after the new package is installed. If the update fails, the RPM package rollbacks to the previous version. It is also possible to revert to an earlier version of the Orchestrator by using the `--force` or the `--oldpackage` option, if needed.

After an attempted update of the 3-Site Orchestrator, the connection that was established between the systems (Master, AuxNear, or AuxFar) might break.

Complete the following steps to recover from an Orchestrator update failure:

1. Verify that the IP status in the output of the **`ls3siteconfig`** command shows as unreachable. Verify that the **`ls3sitercconsistgrp`** command returns an error `CMMVC9507E SSH Connection failed with site [`*`Master_cluster_ip`*`]`.
2. To resolve the issue, enter the **`ch3siteconfig -port`** ***`XXXX`*** command where *`XXXX`* refers to any port number other than what was used. This command is successful only if no replication is running. Otherwise, it fails with a CMMVC9488E error, which indicates that the consistency groups are not in a stopped state.

   Because of the port change, the connection can be established again.

7

Troubleshooting

This chapter provides information about how to troubleshoot problems in a 3-Site Replication environment. It includes the following topics:

7.1 Installing and configuring a 3-Site Replication environment

This section includes some troubleshooting tips for installing and configuring a 3-Site Replication environment.

7.1.1 Installing on Red Hat Enterprise Linux version 8.0 or newer

RHEL 8.0 features the following extra steps that must be performed on the host to make Orchestrator RPM usable:

1. Run the **`update-crypto-policies --set LEGACY`** command.
2. Add `PubkeyAcceptedKeyTypes=+ssh-dss` to the end of `/etc/ssh/sshd_config`.
3. Restart **`sshd systemctl restart sshd`**.

7.1.2 Password-less SSH configuration

A three-site configuration requires password-less authentication that is configured between the orchestrator host and all the three sites to be part of the three-site configuration.

Consider the following points:

- Verify that the user that is specified in **`mk3siteconfig`** command has a role of 3SiteAdmin user group.
- Verify that password-less SSH is configured.
- Ensure that the keys scan to known_hosts is correctly done. You can clean up Ecdsa keys in known_hosts if authentication is not working. Only RSA keys must be preserved.

7.2 Site State Machine

Site State Machine features the following states:

- STOPPED

  When the 3-Site consistency group is in this state, it implies that 3-Site data cycling is stopped because of a link failure or by specific invocation of the **`stop3sitercconsistgrp`** command. A transition to this state generates a notification to the administrator.

  Instances exist of link failure when one or more 3-Site consistency groups can go to the STOPPED state. When 3-Site data cycling is stopped, the RPO of the AuxFar copy of data is extended by the duration for which 3-Site data cycling is stopped.

  Note: The STOPPED state indicates only the state of 3-Site data cycling. It does not indicate the state of the near site consistency group.

- 3 SITE INCONSISTENT (3SI)

  In this state, data on the far site volume is inconsistent. If a failure occurs on one of the near sites when 3-Site consistency groups are in this state, a complete resynchronization is necessary so that the two surviving sites can be DR partners of each other.

A 3-Site Configuration Group can be in this state before the first 3-Site data cycle is completed for a 3-Site Configuration Group. This state of a 3-Site Configuration Group can be, for example, upon creating 3-Site Configuration Group or when a far site that suffered a failure is rejoining a 3-Site configuration. A newly created 3-Site Configuration Group is in STOPPED state. When its data cycling is started, it enters a 3SI state until the first cycle completes.

Note: This state indicates the inconsistency of data of the volumes on the far site only. The relationship between the near sites continues uninterrupted assuming that the two near sites are functioning and have a working partnership.

- 3 SITE CONSISTENT (3SC)

  This state indicates that a consistent copy of data (that is guaranteed to be within two cycles of the near site) is available on the far site. When a 3-Site Configuration Group is in this state, if a failure occurs on one of the sites, the remaining two sites are guaranteed to become DR partners of each other within a maximum of 2 cycle periods of a 3-Site Configuration Group. Under normal circumstances when all three sites are functioning and the Orchestrator can reach each of the three sites, this state is the state of a 3-Site consistency group.

- 2 SITE PERIODIC (2SP)

  This state indicates that one of the near sites is not part of the 3-Site configuration. This state is possible if a failure occurs, the Metro Mirror link fails, or one of the near sites is excluded.

  In this state, the far site is within a maximum of 2 cycle periods of the application. After the excluded site rejoins the configuration, 3-Site cycling is automatically resumed after resync between the two near sites is completed. After 3-Site resync completes, after the first data cycle, the 3-Site Configuration Group transitions to the 3SC state.

- PARTIAL

  This state implies that the configuration of the 3-Site Configuration Group is incomplete. That is, the 3-Site consistency group is not fully formed or one of the relationships that was added to or removed from a 3-Site consistency group is incomplete. For example, a 3-Site consistency group can enter this state if an orchestrator crashes while a 3-Site configuration command is running.

  To exit this state, the administrator can choose to convert a *partial* 3-Site Configuration Group back to a 2-Site Configuration Group by using the **`convertrcconsistgrp`** CLI or retry the last failed operation of **`convertrcrelationship`**.

7.3 Understanding the 3-Site events at the Storage Virtualize cluster

For various alerts and failures that are identified by 3-Site Orchestrator, events are generated on the Storage Virtualize cluster. Events can be viewed at the AuxFar site. However, if an AuxFar site is unavailable, events are generated on the Master site.

Some of the more common events are described in this section.

7.3.1 WARNING: 3-Site consistency group is now in 2-Site data cycling mode

This event indicates that the 3-Site consistency group state was changed to 2Site_Periodic. It can be because of a near site membership change, a change in status of the near site partnership, or a near link or site disaster.

Complete the following steps to diagnose this problem:

- Use the **`ls3siteconfig`** command to find the status of each site and connectivity details.
- Use the **`ls3sitercconsistgrp`** command to check the status of near site copy state.

7.3.2 WARNING: 3-Site consistency group is now in STOPPED State

This event indicates that the 3-Site consistency group state was STOPPED. It can be caused by a near site membership change, or a change in the status of the near site partnership of a near or far site disaster.

The following conditions might exist when user sees this error:

- Near site link failed in Cascade topology
- Near site storage failure in Cascade topology
- Disaster of a periodic source near site
- Failure of active far site data cycling link
- Disaster of a Far site
- Partnership broken for HyperSwap configuration

7.3.3 WARNING: 3-Site data cycling is stopped because Storage Virtualize cluster lost connection with orchestrator

This event is generated when the Storage Virtualize cluster's connection to the orchestrator is lost. 3-Site data cycling is in STOPPED state. The administrator must verify the IP connectivity between the orchestrator and this site.

Connection can be lost because of authentication failure or IP network connection problems and the cycle time for 3-Site consistency group exceeded the cycle period set by the user. When the connection is restored, data cycling is started again automatically.

To determine the connectivity between each site and orchestrator, check the output of the **`ls3siteconfig`** command. In the command output, the site connection status is shown as `unreachable` or `authentication_failed`. 3-Site consistency group status can be verified by using the **`ls3siterconsistgrp`** command.

The administrator is expected to perform the suitable action to fix the connection error.

7.3.4 WARNING: The cycle time for 3-Site consistency group exceeded the cycle period set by user

3-Site consistency group cycle time passed the cycle period, and the Recovery Point Objective (RPO) threshold is exceeded. This issue can occur when a new 3-Site consistency group is synchronizing data to the AuxFar site volumes and excessive application writes are occurring at the primary site or issues exist on link that used for data cycling.

It is recommended that the determination of the cycle time include the number of writes that are occurring on the volumes. The cycle time exceeds the cycling period if it cannot replicate the data in the specified time.

The administrator can use the **ls3sitercconsistgrp** command to determine the cycle period for a 3-Site consistency group. The administrator can also review the `freeze_time` on the AuxFar site. This `freeze_time` shows the timestamp of the last consistent copy that was available at the AuxFar site. This information provides an indication of how much delay is observed at this moment.

7.3.5 WARNING: Configuration change that was requested by the last command failed

This event indicates that the last operation that requested a change in the configuration failed. This failure can occur during the following operations:

- Site rejoin, exclusion for Metro Mirror configuration.
- Change primary, periodic source, cycle period for Metro Mirror and HyperSwap configuration.
- Adding a relationship into 3-Site consistency group for Metro Mirror and HyperSwap configuration.

The administrator can use the **ls3siteconfig** or **ls3sitercconsistgrp** command to identify the issue.

7.3.6 INFO: Configuration change requested by the last command completed successfully

This event indicates that the last operation that requested the change in the configuration completed successfully. This event can happen during the following operations:

- Site rejoin, exclusion.
- Change primary, periodic source.
- Adding a relationship into a 3-Site consistency group.

7.4 3-Site auditing

3-Site Orchestrator features a dedicated service that makes every command that is run on 3-Site Orchestrator become part of the 3-Site audit trail.

The audit history can be found at:

`/opt/ibm/SpectrumVirtualize/rc3site/logs/all_attempted_tscmd_audit`

This file can expand to a maximum of 3 MB and can include a maximum of approximately 9,000–10,000 entries. After it reaches this limit, it is rotated and saved in the same directory with the `.old` suffix. A new file with the original name is opened for the audit trail. Therefore, you have approximately 18,000–20,000 entries of 3-Site commands at your disposal for audit purposes.

The second repository is a small in-memory database, which can store a maximum of 32 audit entries. A utility that is similar to **catauditlog**, called **tscatauditlog**, is available to conveniently dump this database on your terminal.

7.5 3-Site consistency group error codes

For more information about 3-Site consistency group error codes, see 5.1.2, “Monitoring the 3-Site consistency groups” on page 102.

7.6 3-Site Orchestrator logs

The 3-Site Orchestrator services and modules keep logging in to the logs directory of the installation directory. This information is useful for any debugging during the 3-Site Replication and management operations.

You can use the **`orchestrator_snap`** command to collect the snap that includes everything that is required for debugging purposes.

The 3-Site logs directory is available at `/opt/ibm/SpectrumVirtualize/rc3site/logs/`.

The following 3-Site Orchestrator logs are preserved for 14 days (data cycling logs are preserved):.

- tstaskd.log. This log file contains the datacycling logs for all the active 3-Site consistency group.
- tsdiagnostic.log. This log file maintains event history and diagnostic decisions that are made for data cycling.
- *_exec.logs. These logs are exception logs. Any code fault, exceptions, or asserts are maintained in *<module_name>*`_exce.logs`.
- all_attempted_commands.log. This file maintains audit logging for the command execution at orchestrator.

The `/dumps/3site_orchestrator.log` is on the Storage Virtualize system. This log maintains GUI helper logs. It is not part of debug data that is collected by the `orchestrator_snap` tool.

7.7 GUI stale entries

In specific conditions, the Storage Virtualize Management GUI might show the stale data of 3-Site configuration, which is not active and was removed.

Complete the following steps to delete this stale information on the GUI:

1. Delete 3-Site users from 3 Storage Virtualize systems.
2. Log in to the management GUI. Select **Settings** → **GUI Preferences** → **General** and then select **Refresh** under **Refresh GUI cache**.

7.8 Failure protection and recovery procedures for 3-Site Replication

In this section is a discussion of failure protection and recovery procedures for 3-Site Replication.

7.8.1 Orchestrator failure and recovery when adding 2-Site Metro Mirror or HyperSwap relationships with data

In case the IBM Spectrum Virtualize 3-Site Orchestrator encounters a failure during any phase of adding a 2-Site relationship with data, the following actions are required:

- If the Orchestrator fails before you add a new 2-Site relationship with data to a 3-Site consistency group, ensure that the following actions are taken:
  a. If the Orchestrator fails and recovers back in the stopped state, the periodic source is lost, and the 3-Site consistency group status can change to `periodic_src_offline`. If this occurs, the new 2-Site relationship with data cannot be added to the 3-Site consistency group. To add a relationship with data, you can set the periodic source back to the online state by restarting the 3-Site consistency group in the management GUI or use the **`start3siteconsistgrp`** command. This action returns the periodic source back to its original state. After verifying that the periodic source is correct, ensure that the 3-Site consistency group status is online.
  b. Run **`stop3siteconsistgrp`** command to stop the 3-Site consistency group.
  c. On the Orchestrator host, use the **`ls3sitercconsistgrp`** command and verify that the status of the 3-Site consistency group is online. In addition, verify that the original site for the periodic source is displayed in the **`periodic_copy_source`** parameter.
- If the Orchestrator fails and recovers during an attempt to add a new relationship with data, the 3-Site consistency group recovers in partial state. Run the **`convertrcrelationship`** command. After the conversion succeeds, the 2-Site relationship is added to the 3-Site consistency group and data cycling starts automatically.
- If the Orchestrator fails and recovers after you add a new relationship with data to the 3-Site consistency group, the new 2-Site relationship with a state of `new` cannot be recovered. It is assumed that the addition of the relationship was successful. The state of the 3-Site relationship after recovery is complete and the state or the 3-Site consistency group is either `3site_consistent` or `3site_inconsistent`.

7.8.2 Recovering from Orchestrator failure when removing and reverting relationships

In the case the IBM Spectrum Virtualize 3-Site Orchestrator encounters a failure when a 3-Site relationship is being removed from a 3-Site consistency group, the following actions are required:

- If the Orchestrator fails before removing a relationship from a 3-Site consistency:
  - If the Orchestrator fails and recovers to a stopped state, the periodic source is lost, and the 3-Site consistency group status can change to `periodic_src_offline`. If this occurs, you cannot remove the 2-Site relationship. To remove a relationship, you must set the periodic source back to the online state by restarting the 3-Site consistency group with the **`start3siteconsistgrp`** command. This action returns the periodic source back to its original state.

- After verifying that the periodic source is correct, ensure that the 3-Site consistency group status is online.
- After the status is online and orientation is set to default, you can remove the relationship from the 3-Site consistency group.

- If the Orchestrator recovery occurs during an attempt to remove a 3-Site relationship, the following will occur:
  - The Orchestrator recovery process removes the remaining 3-Site consistency group objects during recovery.
  - The remove relationship operation is automatically completed.
  - The state of the 3-Site consistency group recovers to the stopped state.
- If during Orchestrator recovery the auxiliary-far site is offline, the 3-Site consistency group recovers to the stopped state. After the auxiliary-far site is available, the recovery operation restarts, and the remove operation is automatically completed during recovery.

If the Orchestrator recovery occurs after the remove relationship operation, the 3-Site consistency group recovers its previous state before the recovery operation began.

7.9 3-Site Orchestrator high availability

This section describes troubleshooting the 3-Site Orchestrator high availability scenarios.

7.9.1 Handling AuxFar site and active Orchestrator failure and failback

If the active Orchestrator fails, you can shut down the active Orchestrator for standby Orchestrator to takeover. The following are the failure and failback scenarios of active or standby Orchestrator.

AuxFar site failure with both Orchestrators and failback

If the AuxFar site with Orchestrator fails, standby Orchestrator status appears as disconnected. See Example 7-1.

Example 7-1 Standby Orchestrator status appears as disconnected

```
[root@localhost home]# ls3siteconfig
master_name Master
master_site_name
master_site_status online
master_ip1 x.xxx.xxx.xx
master_ip1_status reachable
master_ip2
master_ip2_status not_configured
master_port 16000
auxnear_name Aux_Near auxnear_site_name
auxnear_site_status online
auxnear_ip1 x.xxx.xxx.xx
auxnear_ip1_status reachable
auxnear_ip2
auxnear_ip2_status not_configured
auxnear_port 16001
auxfar_name Aux_Far
auxfar_site_status online
```

```
auxfar_ip1 x.xxx.xxx.xxx
auxfar_ip1_status reachable
auxfar_ip2
auxfar_ip2_status not_configured
auxfar_port 16002
username tuser
keyfile /root/.ssh/pw_rsa
topology metromirror
Orchestrator_version 4.1.221028083514
Orchestrator_ha_status disconnected
```

After recovery of the AuxFar site, standby Orchestrator status appears as online. See Example 7-2.

Example 7-2 After recovery of the AuxFar site, standby Orchestrator status appears as online

```
[root@localhost home]# ls3siteconfig
master_name Master
master_site_name
master_site_status online
master_ip1 x.xxx.xxx.xx
master_ip1_status reachable
master_ip2
master_ip2_status not_configured
master_port 16000
auxnear_name Aux_Near auxnear_site_name
auxnear_site_status online
auxnear_ip1 x.xxx.xxx.xx
auxnear_ip1_status reachable
auxnear_ip2
auxnear_ip2_status not_configured
auxnear_port 16001
auxfar_name Aux_Far
auxfar_site_status online
auxfar_ip1 x.xxx.xxx.xxx
auxfar_ip1_status reachable
auxfar_ip2
auxfar_ip2_status not_configured
auxfar_port 16002
username tuser
keyfile /root/.ssh/pw_rsa
topology metromirror
Orchestrator_version 4.1.221028083514
Orchestrator_ha_status online
```

Active Orchestrator failure and failback

To validate the status of both Orchestrators, enter the **ls3siteconfig** command. See Example 7-3.

Example 7-3 Validating the status of both Orchestrators

```
[root@localhost home]# ls3siteconfig
master_name Master
master_site_name
master_site_status online
```

```
master_ip1 x.xxx.xxx.xx
master_ip1_status reachable
master_ip2
master_ip2_status not_configured
master_port 16000
auxnear_name Aux_Near auxnear_site_name
auxnear_site_status online
auxnear_ip1 x.xxx.xxx.xx
auxnear_ip1_status reachable
auxnear_ip2
auxnear_ip2_status not_configured
auxnear_port 16001
auxfar_name Aux_Far
auxfar_site_status online
auxfar_ip1 x.xxx.xxx.xxx
auxfar_ip1_status reachable
auxfar_ip2
auxfar_ip2_status not_configured
auxfar_port 16002
username tuser
keyfile /root/.ssh/pw_rsa
topology metromirror
Orchestrator_version 4.1.221028083514
Orchestrator_ha_status online
```

To demonstrate the failure of active Orchestrator, shutdown the active Orchestrator by using the **shutdown** command.

Standby Orchestrator becomes the active Orchestrator after trying to connect to the active Orchestrator five times. Standby Orchestrator becomes the active Orchestrator and starts data cycling. As there is only one active Orchestrator, standby Orchestrator status in the **ls3siteconfig** command output appears as `standby_offline`. See Example 7-4.

Example 7-4 Standby Orchestrator status in ls3siteconfig will appear as standby_offline.

```
[root@localhost home]# ls3siteconfig
master_name Master
master_site_name
master_site_status online
master_ip1 x.xxx.xxx.xx
master_ip1_status reachable
master_ip2
master_ip2_status not_configured
master_port 16000
auxnear_name Aux_Near auxnear_site_name
auxnear_site_status online
auxnear_ip1 x.xxx.xxx.xx
auxnear_ip1_status reachable
auxnear_ip2
auxnear_ip2_status not_configured
auxnear_port 16001
auxfar_name Aux_Far
auxfar_site_status online
auxfar_ip1 x.xxx.xxx.xxx
auxfar_ip1_status reachable
auxfar_ip2
```

```
auxfar_ip2_status not_configured
auxfar_port 16002
username tuser
keyfile /root/.ssh/pw_rsa
topology metromirror
Orchestrator_version 4.1.221028083514
Orchestrator_ha_status standby_offline
```

After the original active Orchestrator re-joins, it will recognize its role as standby Orchestrator, and `Orchestrator_ha_status` appears as online. See Example 7-5.

Example 7-5 "Orchestrator_ha_status" appears as online

```
[root@localhost home]# ls3siteconfig
master_name Master
master_site_name
master_site_status online
master_ip1 x.xxx.xxx.xx
master_ip1_status reachable
master_ip2
master_ip2_status not_configured
master_port 16000
auxnear_name Aux_Near auxnear_site_name
auxnear_site_status online
auxnear_ip1 x.xxx.xxx.xx
auxnear_ip1_status reachable
auxnear_ip2
auxnear_ip2_status not_configured
auxnear_port 16001
auxfar_name Aux_Far
auxfar_site_status online
auxfar_ip1 x.xxx.xxx.xxx
auxfar_ip1_status reachable
auxfar_ip2
auxfar_ip2_status not_configured
auxfar_port 16002
username tuser
keyfile /root/.ssh/pw_rsa
topology metromirror
Orchestrator_version 4.1.221028083514
Orchestrator_ha_status online
```

7.9.2 Near sites link failure with active Orchestrator

If the near sites become unreachable from the active Orchestrator, data cycling is halted, and the standby Orchestrator becomes the active Orchestrator if the near sites are reachable from it.

The following case illustrates near sites failure with active Orchestrator.

For example, if near sites are accessible from active Orchestrator, the status appears as reachable. See Example 7-6 on page 194.

Example 7-6 Status appears as reachable

```
[root@localhost home]# ls3siteconfig
master_name Master
master_site_name
master_site_status online
master_ip1 x.xxx.xxx.xx
master_ip1_status reachable
master_ip2
master_ip2_status not_configured
master_port 6222
auxnear_name Aux_Near auxnear_site_name
auxnear_site_status online
auxnear_ip1 x.xxx.xxx.xx
auxnear_ip1_status reachable
auxnear_ip2
auxnear_ip2_status not_configured
auxnear_port 6223
auxfar_name Aux_Far
auxfar_site_status online
auxfar_ip1 x.xxx.xxx.xxx
auxfar_ip1_status reachable
auxfar_ip2
auxfar_ip2_status not_configured
auxfar_port 6224
username tuser
keyfile /root/.ssh/pw_rsa
topology metromirror
Orchestrator_version 4.1.221028083514
Orchestrator_ha_status online
```

If the near sites are inaccessible from active Orchestrator, status appears as unreachable, and data cycling is stopped. Enter the **`ls3siteconfig`** command to validate the configuration. See Example 7-7.

Example 7-7 Validating the configuration - status appears as unreachable

```
[root@localhost home]# ls3siteconfig
master_name Master
master_site_name
master_site_status online
master_ip1 x.xxx.xxx.xx
master_ip1_status reachable
master_ip2
master_ip2_status not_configured
master_port 6222
auxnear_name Aux_Near auxnear_site_name
auxnear_site_status online
auxnear_ip1 x.xxx.xxx.xx
auxnear_ip1_status reachable
auxnear_ip2
auxnear_ip2_status not_configured
auxnear_port 6223
auxfar_name Aux_Far
auxfar_site_status online
auxfar_ip1 x.xxx.xxx.xxx
```

```
auxfar_ip1_status reachable
auxfar_ip2
auxfar_ip2_status not_configured
auxfar_port 6224
username tuser
keyfile /root/.ssh/pw_rsa
topology metromirror
Orchestrator_version 4.1.221028083514
Orchestrator_ha_status online
```

After five minutes, when the standby Orchestrator takes over the role of active Orchestrator, near sites become accessible and data cycling will resume. Enter the the**ls3siteconfig** command to validate the configuration. See Example 7-8.

Example 7-8 Validating the configuration - near sites become accessible

```
[root@localhost home]# ls3siteconfig
master_name Master
master_site_name
master_site_status online
master_ip1 x.xxx.xxx.xx
master_ip1_status reachable
master_ip2
master_ip2_status not_configured
master_port 31000
auxnear_name Aux_Near auxnear_site_name
auxnear_site_status online
auxnear_ip1 x.xxx.xxx.xx
auxnear_ip1_status reachable
auxnear_ip2
auxnear_ip2_status not_configured
auxnear_port 31001
auxfar_name Aux_Far
auxfar_site_status online
auxfar_ip1 x.xxx.xxx.xxx
auxfar_ip1_status reachable
auxfar_ip2
auxfar_ip2_status not_configured
auxfar_port 31002
username tuser
keyfile /root/.ssh/pw_rsa
topology metromirror
Orchestrator_version 4.1.221028083514
Orchestrator_ha_status online
```

Note: In a Metro Mirror configuration, if the near sites link fails with the active Orchestrator, only one of the near sites is accessible from the standby Orchestrator. In this case, it is not possible to change the role to active Orchestrator due to the following reasons:

- The site accessible from the standby Orchestrator can be the primary site or primary to only a partial number of consistency groups.
- The site accessible from the standby Orchestrator can be aux-near to all consistency groups.
- The site accessible from the standby Orchestrator might lose access to the accessible near sites.

7.10 When to contact IBM Support

Contact IBM Support if you encounter any of the following situations:

- A dual disaster of near sites or near and DR sites.
- 3-Site Orchestrator is not starting after an orchestrator or host crash.
- Stale objects are deleted from a site that is not an active participant in the 3-Site configuration.
- A dead domain occurs. This issue can be caused by dead removal of I/O groups by using the **`leavecluster`** and **`rmnodecanister`** commands.
- Recovering 3-Site configuration if a T3 disaster occurs at any of the three sites.

7.11 Frequently asked questions

In this section, the following frequently asked questions are presented:

- Where do you find the IBM Storage Virtualize 3-Site Orchestrator installer?

  To download the RPM installer, see IBM Storage Virtulize 3-Site Orchestrator Tool.

- Is a stand-by HA instance of an orchestrator possible?

  No. An HA instance of orchestrator is not recommended. Having two orchestrator configurations serving the same 3-Site configuration can result in data corruption or an unexpected state. However, the user can set up a dead instance of 3-Site Orchestrator without creating the 3-Site configuration, which is ready to quickly start from a **`mk3siteconfig`**.

- What happens if a 3-Site Orchestrator service or host crashes?

  3-Site Orchestrator services are configured to restart automatically upon failure of stopping. All services are interdependent.

- What happens to my configuration after 3-Site Orchestrator restarts?

  3-Site Orchestrator is a stateless and fault-tolerant architecture. It does not save any persistent configuration of 3-Site objects on the Storage Virtualize system. However, it does preserve the configuration information that is required to reestablish the connection with 3-Site.

  After a connection is created, 3-Site Orchestrator reads all the attributes on 3 Storage Virtualize systems and re-creates the configuration. That is, the 3-Site Orchestrator

restores the configuration after restart and resumes the data cycling. At this point, the state of the 3-Site consistency group depends on status of 3-Site.

- Is it possible to add 2-Site relationships to a 3-Site consistency group?

  Yes. Adding 2-Site relationships to a 3-Site consistency group is possible by using the **`addnewrcrelationship`** command. This 2-Site relationship must match all the attributes, such as the primary, aux roles, and replication type.

  Important prerequisite: The relationships to be added must be between empty volumes. Adding a relationship with data is not supported at the time of this writing.

- On which site does 3-Site Orchestrator need to be configured?

  3-Site Orchestrator is an external management software program. It must be installed on an RHEL Linux host version 7.5 or newer.

  The host can be configured at any of the three sites of the new fourth site. No suggestion exists for the site hosting location.

- Is the AuxFar site as the primary configuration possible?

  No. Only one of the near sites can play the role of primary sites. Auxfar as primary is *not* supported.

A

IBM Storage Virtualize 3-Site Replication components

The IBM Storage Virtualize command line interface (CLI) shows the details of the volumes that are involved with replication, FlashCopy mappings, remote copy relationships, and remote copy consistency groups.

This appendix provides information about the commands that are used in an IBM Storage Virtualize cluster. The details of a consistency group that is shown in the GUI can be listed by the CLI commands that are shown in the examples that are used in this appendix.

The output of the CLI commands and the header names are shortened for better readability. Also, only one volume of the consistency group is shown.

This appendix includes the following topics:

Master

The volume VOL_1_L includes the associated AccessPoint vdisk0, which can be checked by using the **lsfcmap** and **lsvdisk** command. The output is shown in Example A-1.

Example A-1 Master volumes and FlashCopy mappings

```
Master:monitorUser>lsvdisk -delim : | cut -d: -f2,5,13,29 | tr ":" "\t"
name    status  RC_name         function
VOL_1_L online  rcrel0          master
vdisk0  online  tsrell_ffyvmf72 master

Master:monitorUser>lsfcmap -delim : | cut -d: -f2,4,6,9,17 | tr ":" "\t"
name    source_vdisk_name target_vdisk_name status       start_time
fcmap0  VOL_1_L           vdisk0            copying      200323151004
fcmap1  vdisk0            VOL_1_L           idle_or_copied
```

The Configuration Group tscgl_nawfbf644 can be checked by using the **lsrcconsistgrp** command, as shown in Example A-2. The Metro Mirror remote copies that are used for the two volumes are listed by using the **lsrcrelationship** command, as shown in Example A-2.

Example A-2 Master RCs and Configuration Groups

```
Master:monitorUser>lsrcrelationship -delim : | cut -d: -f2,4,6,8,10,11,13
name            master_ master_ aux_    aux     primary consistency_group_name
                cluster vdisk   cluster vdisk
rcrel0          Master  VOL_1_L AuxNear VOL_1_R master  CG0
tsrell_ffyvmf72 Master  vdisk0  AuxFar  vdisk0  master  tscgl_nawfbf644

Master:monitorUser>lsrcconsistgrp -delim : | cut -d: -f2,4,6-8,10 | tr ":" "\t"
name            master_ aux_    primary state                   copy_type
                cluster cluster
CG0             Master  AuxNear master  consistent_synchronized metro
tscgl_nawfbf644 Master  AuxFar  master  consistent_stopped      metro
```

AuxNear

The volume VOL_1_R includes the associated AccessPoint vdisk0, which can be checked by using the **lsfcmap** and **lsvdisk** command output, as shown in Example A-3.

Example A-3 AuxNear RCs and Configuration Groups

```
AuxNear:monitorUser>lsvdisk -delim : | cut -d: -f2,5,13,29 | tr ":" "\t"
name    status  RC_name         function
VOL_1_R online  rcrel0          master
vdisk0  online  tsrelr_ffyvmf72 master

AuxNear:monitorUser>lsfcmap -delim : | cut -d: -f2,4,6,9,17 | tr ":" "\t"
name    source_vdisk_name target_vdisk_name status       start_time
fcmap0  VOL_1_R           vdisk0            copying      200323151004
fcmap1  vdisk0            VOL_1_R           idle_or_copied
```

The Configuration Group tscgr_nawfbf644, as shown in Example A-4 on page 201, can be viewed by using the **lsrcconsistgrp** command, as shown in Example A-4 on page 201. The

Metro Mirror relationships that are used for the two volumes are listed by using the **lsrcrelationship** command. See Example A-4. The remote copy name and the Configuration Group to the AuxFar site of the corresponding volumes on the AuxNear cluster feature the same prefix and suffix. They differ in the site notation only: `l` for the Master, and `r` for the AuxNear site.

Example A-4 AuxNear RCs and Configuration Groups

```
AuxNear:monitorUser>lsrcrelationship -delim : | cut -d: -f2,4,6,8,10,11,13
name            master_ master_ aux_    aux     primary consistency_group_name
                cluster vdisk   cluster vdisk
rcrel0          Master  VOL_1_L AuxNear VOL_1_R master  CG0
tsrelr_ffyvmf72 AuxNear vdisk0  AuxFar  vdisk1          tscgr_nawfbf644

AuxNear:monitorUser>lsrcconsistgrp -delim : | cut -d: -f2,4,6-8,10 | tr ":" "\t"
name            master_ aux_    primary state                   copy_type
                cluster cluster
CG0             Master  AuxNear master  consistent_synchronized metro
tscgr_nawfbf644 AuxNear AuxFar          idling                  metro
```

AuxFar

The volume `VOL_1_T` includes two associated AccessPoint: `vdisk0` and `vdisk1`. One is used for the remote copy to the Master cluster (on the left site), and the one for the RC to the AuxNear cluster (right site). Only one remote copy is active, depending on the periodic source.

The AuxFar volumes and FlashCopy mappings can be checked by using the **lsfcmap** and **lsvdisk** command output, as shown in Example A-5.

Example A-5 AuxFar volumes and FlashCopy mappings

```
AuxFar:monitorUser>lsvdisk -delim : | cut -d: -f2,5,13,29 | tr ":" "\t"
name    status  RC_name function
VOL_1_T online
vdisk0  online  tsrell_ffyvmf72 aux
vdisk1  offline tsrelr_ffyvmf72 aux

:AuxFar:monitorUser>lsfcmap -delim : | cut -d: -f2,4,6,9,17 | tr ":" "\t"
name    source_vdisk_name       target_vdisk_name       status  start_time
fcmap0  VOL_1_T vdisk0  copying 200323183024
fcmap1  vdisk0  VOL_1_T idle_or_copied
fcmap2  VOL_1_T vdisk1  idle_or_copied
fcmap3  vdisk1  VOL_1_T idle_or_copied

AuxFar:monitorUser>lsrcrelationship tsrell_ffyvmf72 | grep channel
channel left
AuxFar:monitorUser>lsrcrelationship tsrelr_ffyvmf72 | grep channel
channel right
```

If the remote copy of the left or right site is active, they can be listed by using the `channel` property of the remote copy that is associated with the active volume. As shown in Example A-5 on page 201, the volume `vdisk0` is online; therefore, the remote copy `tsrell_ffyvmf72` is active. The `channel` property of the remote copy `tsrell_ffyvmf72` is left.

The Configuration Groups `tscgl_nawfbf644` and `tscgr_nawfbf644` (left and right Configuration Groups), as shown in Example 1-2 on page 15, can be checked by using the **`lsrcconsistgrp`** command, as shown in Example A-6.

Example A-6 AuxFar RCs and Configuration Groups

```
Master:monitorUser>lsrcrelationship -delim : | cut -d: -f2,4,6,8,10,11,13
name            master_ master_ aux_    aux     primary consistency_group_name
                cluster vdisk   cluster vdisk
tsrell_ffyvmf72 Master  vdisk0  AuxFar  vdisk0  master  tscgl_nawfbf644
tsrelr_ffyvmf72 AuxNear vdisk0  AuxFar  vdisk1          tscgr_nawfbf644

Master:monitorUser>lsrcconsistgrp -delim : | cut -d: -f2,4,6-8,10 | tr ":" "\t"
name            master_ aux_    primary state                   copy_type
                cluster cluster
AuxFar:monitorUser>lsrcconsistgrp -delim : | cut -d: -f2,4,6-8,10 | tr ":" "\t"
name    master_cluster_name     aux_cluster_name        primary state   copy_type
tscgl_nawfbf644 Master  AuxFar  master  consistent_stopped      metro
tscgr_nawfbf644 AuxNear AuxFar          idling                  metro
```

The Metro Mirror remote copies that are used for the two volumes are listed by using the **`lsrcrelationship`** command, as shown in Example A-6. The remote copy name and the Configuration Group to the Near Sites of the corresponding volumes on the Master cluster feature the same prefix and suffix. They differ in the site donation only: `l` for the Master (left) and `r` for the AuxNear (right) site.

Related publications

The publications that are 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 topic in this document. Note that some publications that are referenced in this list might be available in softcopy only:

- *Implementation Guide for IBM Storage FlashSystem and IBM SAN Volume Controller: Updated for IBM Storage Virtualize Version 8.6*, SG24-8542
- *Performance and Best Practices Guide for IBM Storage FlashSystem and IBM SAN Volume Controller: Updated for IBM Storage Virtualize Version 8.6,* SG24-8543

You can search for, view, download, or order these documents and other Redbooks, Redpapers, Web Docs, draft, and additional materials, at the following website:

ibm.com/redbooks

Online resources

The IBM FlashSystem 9200 and 9100 8.6.x IBM Documentation website also is relevant as a further information source:

https://www.ibm.com/docs/en/flashsystem-9x00/8.6.x

Help from IBM

IBM Support and downloads:

ibm.com/support

IBM Global Services:

ibm.com/services

Redbooks

IBM Storage Virtualize 3-Site Replication

(0.2”spine)
0.17”<->0.473”
90<->249 pages

SG24-8504-01

ISBN 0738461431

Printed in U.S.A.