Redbooks
ibm.com/redbooks

IBM HyperSwap for IBM FlashSystem A9000 and A9000R

Francesco Anderloni

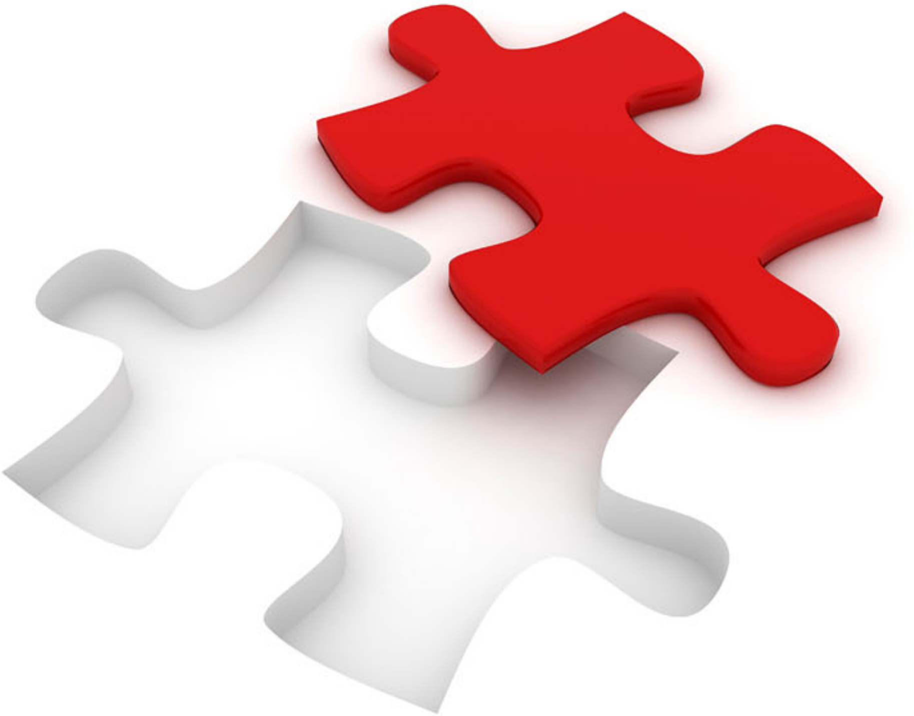
Bert Dufrasne

Roger Eriksson

Andrew Greenfield

Lisa Martinez

Stephen Solewin



Storage

Redpaper

International Technical Support Organization

IBM HyperSwap for IBM FlashSystem A9000 and A9000R

December 2018

REDP-5434-02

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

Thrid Edition (December 2018)

This edition applies to IBM FlashSystem A9000 and IBM FlashSystem A9000R with Sotware V12.3.
This document was created or updated on February 4, 2019.

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.

AIX®
DS8000®
FlashCopy®
Global Technology Services®
HyperSwap®
IBM®
IBM FlashSystem®
IBM Spectrum™
IBM Spectrum Accelerate™
IBM Spectrum Protect™
Redbooks®
Redpaper™
Redbooks (logo) ®
System Storage®
XIV®
z/VM®

The following terms are trademarks of other companies:

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Linear Tape-Open, LTO, the LTO Logo and the Ultrium logo are trademarks of HP, IBM Corp. and Quantum in the U.S. and other countries.

Microsoft, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

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

Preface

IBM® HyperSwap® is the high availability (HA) solution that provides continuous data availability in case of hardware failure, power failure, connectivity failure, or disasters. The HyperSwap capability is available for IBM FlashSystem® A9000 and IBM FlashSystem A9000R, starting with software version 12.2.1.

Version 12.3 introduces a function that combines HyperSwap and Asynchronous replication, which creates a solution that entails HA and Disaster Recovery (DR). One side of the HyperSwap pair has an active async link to the third system, and the other side has a standby link. Known as Multi-site HA/DR, this configuration provides HyperSwap active-active HA while keeping data mirrored to a third copy to ensure two levels of business continuity.

This IBM Redpaper™ publication gives a broad understanding of the architecture, design, and implementation of HyperSwap and Multi-site HA/DR solution. It also discusses and illustrates various use cases pertaining to their use and functionality.

This paper is intended for those users who want to deploy solutions that take advantage of HyperSwap and Multi-site HA/DR for FlashSystem A9000 and A9000R.

Authors

This paper was produced by a team of specialists from around the world working for the International Technical Support Organization, at the IBM European Storage Competence Center (ESCC) in Kelsterbach, Germany.

Francesco Anderloni is an Infrastructure Specialist working for the IBM Global Technology Services® organization in IBM Italy. Since he joined IBM in 2015, he has been working in the Storage Management area, delivering IBM Storage solutions to clients in Italy and Europe. He is a member of the Young Technical Expert Council in IBM Italy. Francesco holds a Bachelor's Degree in Information Engineering and a Master's Degree in Computer Science, both from the University of Padova, Italy.

Bert Dufrasne is an IBM Certified Consulting IT Specialist and Project Leader for IBM System Storage® disk products at the International Technical Support Organization (ITSO). He has worked at IBM in various IT areas. He has authored many IBM Redbooks® publications and also developed and taught technical workshops. Before joining the ITSO, he worked for IBM Global Services as an Application Architect. He holds a Master's degree in Electrical Engineering.

Roger Eriksson works at IBM Systems Lab Services Nordic, based in Stockholm, Sweden. He is a Senior Accredited IBM Product Service Professional. Roger has over 25 years of experience working on IBM servers and storage, including Enterprise and Midrange disk, NAS, SAN, IBM x86, and IBM Power. He has done consulting, proof of concepts, and education, mainly with the Spectrum Accelerate and Spectrum Virtualize product line, since December 2008. He also works with the Swedish technical sales team for storage solutions on a daily basis. He holds a Technical Collage Graduation in Mechanical Engineering.

Andrew Greenfield is an IBM Global XIV® and Flash Solution Engineer who is based in Phoenix, Arizona. He holds numerous technical certifications from Cisco, Microsoft, and IBM. He is also responsible for many of the photos and videos that are featured in this book and on the IBM website. Andrew brings over 24 years of data center experience with Fortune 100 companies to the team. He graduated Magna cum Laude, Honors, from the University of Michigan, Ann Arbor. Andrew has also written other IBM FlashSystem and XIV Gen3 IBM Redbooks publications

Lisa Martinez has been working in the Washington Systems Center—Storage Team as a storage specialist since January 2012. Her focus has been with pre-sales support for IBM DS8000® and the IBM Spectrum™ Accelerate family of storage products (FlashSystem A9000 and A9000R, XIV, and IBM Spectrum Accelerate™). She is also the lead instructor for FlashSystem A9000 and A9000R, XIV, and Spectrum Accelerate customer based workshops. Her experience includes roles as a storage architect in the Specialty Services Area in GTS and test architect in disk storage, focusing on system-level test for XiV for three years, and copy services for DS8000. Lisa holds degrees in Computer Science from New Mexico Highlands University and Electrical Engineering from the University of New Mexico.

Stephen Solewin is an IBM Corporate Solutions Architect based in Tucson, Arizona. He has over 20 years of experience in working on IBM storage, including enterprise and midrange disk, Linear Tape-Open (LTO) drives and libraries, SAN, storage virtualization, and storage software. Steve is a global resource, working with customers, IBM Business Partners, and fellow IBMers worldwide. He has been working on the XIV product line since 2008 and FlashSystem since 2013. He holds a BS degree in electrical engineering from the University of Arizona, where he graduated with honors.

Special thanks to **Yossi Siles**, **Rivka Matosevitch**, and **Iftach Landau** for their support and advice.

Thanks to the following people for their contributions to this project:

Anat Bar-Natan-Kigel
Kobi Beifus
Craig Gordon
Detlef Helmbrecht
John Hyams
Markus Oscheka
Marie Romero
Vadim Steckler
Max Vidgerovich
Ralf Wohlfart
IBM

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 ITSO 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 papers to be as helpful as possible. Send us your comments about this paper 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, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks

- Follow us on Twitter:

  http://twitter.com/ibmredbooks

- 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

HyperSwap for IBM FlashSystem A9000 and A9000R

This chapter provides a high-level, functional overview of the IBM HyperSwap feature that is available on IBM FlashSystem A9000 and A9000R systems, starting with system software version 12.1.

This chapter covers the following topics:

- 1.1, “HyperSwap feature overview” on page 2
- 1.2, “Architecture, design, and components” on page 5
- 1.3, “Independent QW” on page 8
- 1.4, “HyperSwap various topologies” on page 10

1.1 HyperSwap feature overview

HyperSwap is an IBM trademark that names the capability of active-active storage deployments with automatic failover. It is available with various IBM product lines, but its architecture and implementation are different in each case. This overview specifically applies to HyperSwap for FlashSystem A9000 and A9000R.

The HyperSwap feature of FlashSystem A9000 and A9000R, also referred to as *transparent failover*, delivers always-on, high availability (HA) storage service for storage volumes or Consistency Groups in a production environment. It is based on an active-active, cross-system, and cross-datacenter configuration and does not require additional licensing or special hardware.

HyperSwap builds upon the synchronous mirroring functionality that is already included with FlashSystem A9000 and A9000R, together with other advanced replication and stretched-cluster features.

HyperSwap relies on Asymmetrical Logical Unit Access (ALUA) support to indicate to the host the preferred paths to the storage system, and minimize latency.

HyperSwap volumes can autonomously and transparently switch between *primary* and *secondary* roles, based on the volume failover state. In effect, from a host perspective, the pair of mirrored volumes on both mirrored systems constitute a *single HyperSwap volume*, also referred to as *a stretched volume*.

1.1.1 Basic configuration

The basic HyperSwap configuration is shown in Figure 1-1.

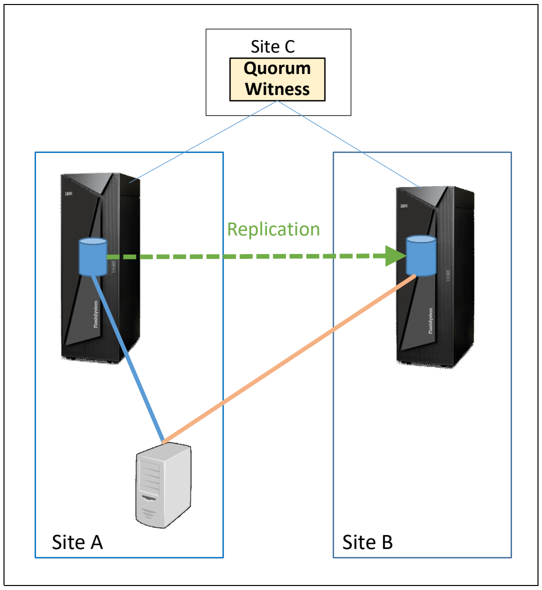


Figure 1-1 HyperSwap volume on peer systems serving host

The basic HyperSwap configuration consists at a minimum of the following components:

- *Two storage systems*: It can be any combination of FlashSystem A9000 or FlashSystem A9000R. The two systems must be able to interconnect for synchronous mirroring over Fibre Channel.

- *Application host*: The application host must have read and write access to both storage systems. Host I/Os can be over iSCSI or Fibre Channel.
- *Quorum Witness*: As in other high availability solutions, HyperSwap requires a quorum witness component. The Quorum Witness (QW) is included with the HyperSwap solution. Its role is to monitor both storage systems health and to arbitrate which storage system owns the primary volume. The QW is ideally located at a third, separate, physical site with IP connectivity to both storage systems.

Note: The basic configuration shown in Figure 1-1 on page 2 represents what is known as a *uniform* configuration where the host can access both the primary and secondary volumes. The uniform configuration is the preferred practice to protect a host from data access problems.

In a *non-uniform* configuration, the storage high availability relies on the server detection and failover of the application to the server with access to active storage. A non-uniform configuration can be used when the host is part of a cluster, can fail over to another host in the cluster, and the other host is connected to the peer system. This configuration is less costly from network perspective. However, it relies on the host failover, which in most cases would not be necessary in a uniform configuration.

1.1.2 HyperSwap as a critical feature for high availability

The automatic and transparent failover at the volume or Consistency Group level is critical to high availability in partial or complete system failures, as well as in site disaster scenarios, either in the same datacenter or between metro-distant datacenters serving single or clustered hosts.

HyperSwap protected workloads are not disrupted through most storage, network, server, application, and site failures and disasters. Enterprise IT systems can automatically fail over within seconds, without human intervention and in a straightforward way using a metro-distance stretch-cluster, to protect against loss of access to mission-critical data.

HyperSwap is, therefore, a critical feature for high availability, as further described in the following sections:

- Always-on, zero-downtime volume operation
- Simplicity
- Robustness
- Data protection
- Typical high availability use cases for HyperSwap

Always-on, zero-downtime volume operation

The ability to attain always-on flash storage systems is critical to applications (such as time-sensitive financial applications) that require constant operation with availability on a 24 by 7 basis. The active-active pairing, per volume or Consistency Group enables zero downtime operation.

Each individual FlashSystem A9000 or A9000R system has a rating of five nines high availability, without HyperSwap. The possibility that two totally separate systems will fail at once, is equivalent to a failure of a system with 10 nines availability rating. To get a sense of the rarity, a system with nine nines is expected to have a total of 31.5569 milliseconds downtime a year.

However, the overall high-availability (number of nines) of the solution also depends on the reliability of the applications, hosts, networks, power systems, cooling systems, and separation of sites. Generally, the high availability of a HyperSwap solution is as good as the weakest link in the chain. When all the components are redundant and appropriately separated among fault-domains, the overall availability from an application perspective can go above and beyond the five nines, and HyperSwap is made to enable such levels of reliability[1].

Simplicity

Using active-active systems is a way to greatly simplify failure management. Automation is provided mostly by the solution and does not require complex scripting or procedures. This aspect is especially important for smaller organizations that cannot deal with failover automation complexity.

Robustness

HyperSwap behavior is designed to address the impact of failures and disasters. The HyperSwap function can identify the following failures and recover from them automatically:

- Replication is down
- The QW is down
- The QW connectivity is down

Data protection

With HyperSwap on IBM FlashSystem A9000 and A9000R:

- Protected hosts and applications continue to benefit from non-disruptive storage service during any of the following failures:
  - Host connectivity to any one of the peer systems is down, for any reason.
  - One of the peer systems does not respond to the host, for any reason.
  - One of the peer systems is completely down, for any reason.
  - A host application or a whole host that is part of a server high availability cluster, fails-over to a remote host that is already connected to the peer HyperSwap volume.
- VMware virtual machines can be migrated non-disruptively using VMware Storage vMotion across vCenter servers and across data centers, supporting VMware vSphere Metro Storage Cluster (vMSC) configurations, as well as VMware Site Recovery Manager (SRM) 6.1 or later.

Typical high availability use cases for HyperSwap

HyperSwap on IBM FlashSystem A9000 and A9000R is mostly useful in the following circumstances:

- Critical applications must run on an all-flash storage system and require continuous operation.
- Requirements mandate active-active all-flash storage capability.
- When used with VMware Site Recovery Manager 6.1 (or later), HyperSwap enables zero-downtime with faster recovery for VMware ESXi clusters.

  Regulations or policies require system administrators to be off-site during weekends and holidays. Fully automated disaster recovery (DR) might be fast enough to compensate for administrator absence, but could be too complex or fragile for operation.

[1] The actual level of availability for applications that use HyperSwap in customer environments requires solution-specific expert assessment.

1.2 Architecture, design, and components

This section provides some insight into the architecture, design and components of the HyperSwap implementation for FlashSystem A9000 and A9000R.

1.2.1 HyperSwap volumes

HyperSwap high availability is based on dual storage systems with active-active pairing per volume, or per Consistency Group. It is important to note that a HyperSwap relationship is not defined at the storage system level. The HyperSwap function is between peer volumes or peer Consistency Groups on two separate FlashSystem A9000 or A9000R systems.

HyperSwap volumes have a copy on one storage system at one site and a copy on the other storage system at another site. Data that is written to the volume is automatically sent to both copies; if one site is no longer available, the other site can provide access to the volume.

Each volume or Consistency Group pair is using synchronous replication to keep both systems updated at all times. When certain conditions apply, an autonomous and completely transparent failover from a volume to its peer is performed, so that host applications experience no downtime.

As previously indicated, HyperSwap for FlashSystem A9000 and A9000R builds upon the synchronous mirroring functionality. As is the case with mirroring, one volume is designated as the *primary* and its peer is designated as the *secondary*. With HyperSwap, volumes can automatically and transparently switch between primary and secondary roles.

Primary and secondary volumes in a HyperSwap relationship are read and write enabled for host I/O. In effect, the pair of volumes on both mirrored systems have identical SCSI attributes and, as such, act as a single HyperSwap volume, also referred to as a *stretched volume*.

As shown in Figure 1-2, each HyperSwap-enabled volume exists simultaneously on two IBM FlashSystem A9000 or A9000R peer systems (depending on the specific deployment), in an active-active, synchronous replication relationship over Fibre Channel (FC).

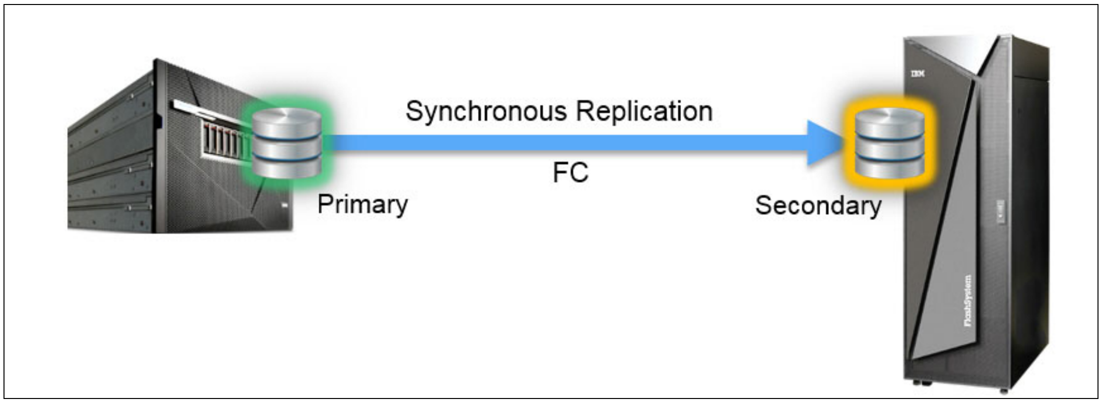


Figure 1-2 HyperSwap volume in synchronous replication over FC

The pair of volumes in a HyperSwap relationship have a unique an identical *SCSI identity* in Network Addressing Authority (NAA) format and the same characteristics relative to I/Os (size, lock state, SCSI reservations). Each storage system with volumes in a HyperSwap configuration has its own volume mapped to the application host, but the application sees them as one LUN. A HyperSwap stretch volume behaves as a single volume in all SCSI aspects. HyperSwap guarantees that the host can use the stretch volume as a single LUN.

Storage administrators can non-disruptively and without pausing synchronization, convert synchronously mirrored volumes to HyperSwap volumes, and vice versa. HyperSwap volumes can also be set up using offline initialization.

1.2.2 Host paths

For reliability host operating systems include a multipath storage driver that allows multiple paths to a storage volume. With recent operating systems, those paths are optimized using Asymmetric Logical Unit Access (ALUA) support from the multipath driver. ALUA, also known as *Target Port Groups Support* (TPGS), consists of SCSI concepts and commands that establish path preferences for storage devices.

In simple terms, ALUA allows a storage system to indicate to an operating system the state of the port group. The states include both priority (preferred or non-preferred) and functionality (active or unavailable). When everything is healthy, the port group state conveys only the preference, but in failure scenarios, it also tells the host which ports cannot be used for I/O.

The purpose of the primary secondary designation is to optimize latency: the primary volume must be co-located with the hosts that generate most of the I/O. The HyperSwap function in FlashSystem A9000 and A9000R marks paths to the system that currently owns the primary volume as *active/preferred*, while paths to the system that currently owns the secondary volume are marked as *active/non-preferred*. Consequently, most of the write I/Os will go directly to the primary volume because it is the preferred path.

Note: The host experiences a failover as a change in the state of its paths. The change is therefore handled by the host multipath driver with no impact to the application (except some delay) and no manual intervention.

Non-preferred paths are used by hosts to load-balance or mitigate disconnects. Occasional write I/Os going over the active, but not preferred path to the secondary volume are redirected to the primary. In this case the secondary volume acts as a proxy of the primary volume. Refer to Figure 1-3 for more information about the transaction.

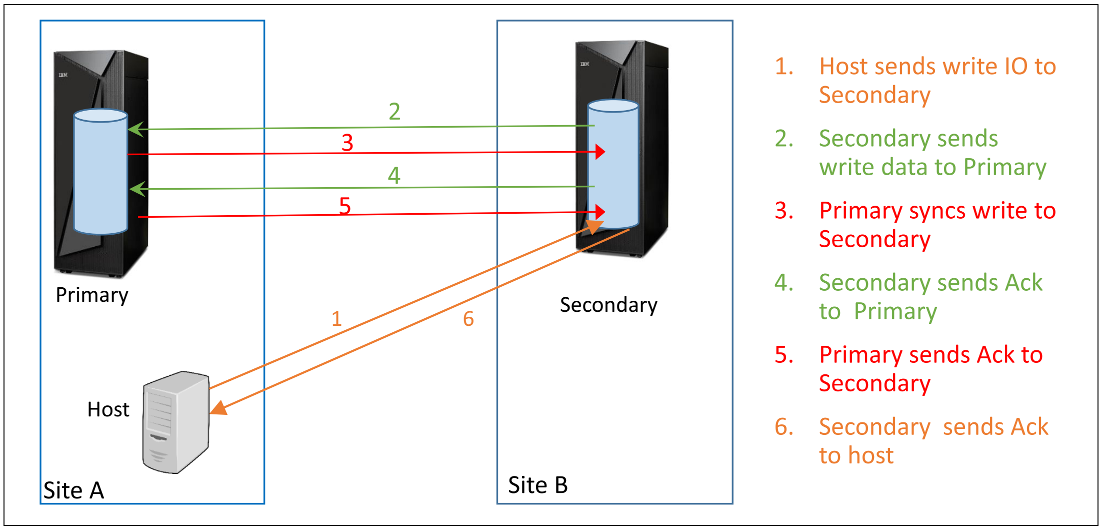

2
3
4
5
Primary
1
6
Secondary
Host
Site A
Site B
1. Host sends write IO to Secondary
2. Secondary sends write data to Primary
3. Primary syncs write to Secondary
4. Secondary sends Ack to Primary
5. Primary sends Ack to Secondary
6. Secondary sends Ack to host

Figure 1-3 Write I/O to secondary

Read I/Os requests on the non-preferred path are normally served directly by the secondary. However if the relationship was just activated and the secondary is not yet fully synchronized, read requests are also redirected to the primary volume until synchronization is complete.

1.2.3 Failure domains

For best protection against different types of possible failures, each of the two storage systems and the QW must be located in distinct failure domains. The failure domains must be defined by the failures that must be overcome. For example, three physical sites can be considered as three failure domains with regard to failures that affect the physical site but not in case of a major power failure when the three physical sites are connected to the same power grid.

Typically, you want to be protected from any type of disaster that might either affect both storage systems at the same time or that would simultaneously disable one of the storage systems and the quorum witness application.

However, the physical sites must still be within distances that allow synchronous mirroring between the storage systems, and acceptable latency. For more information, see Chapter 3, “Prerequisites” on page 17.

1.2.4 Connection with host stretch-cluster configurations

HyperSwap in stretch-cluster configurations include the following properties, as shown in Figure 1-4:

- Protected hosts are connected to both systems locally and remotely.
- I/O path priorities are assigned for minimum latency, by default. The two systems assign a preferred or non-preferred Asymmetric Logical Unit Access (ALUA) state to the port groups, which are translated by the host multipath driver to path priorities, resulting in optimized I/O latency.
- Multiple HyperSwap volumes can exist on each pair of peer systems.

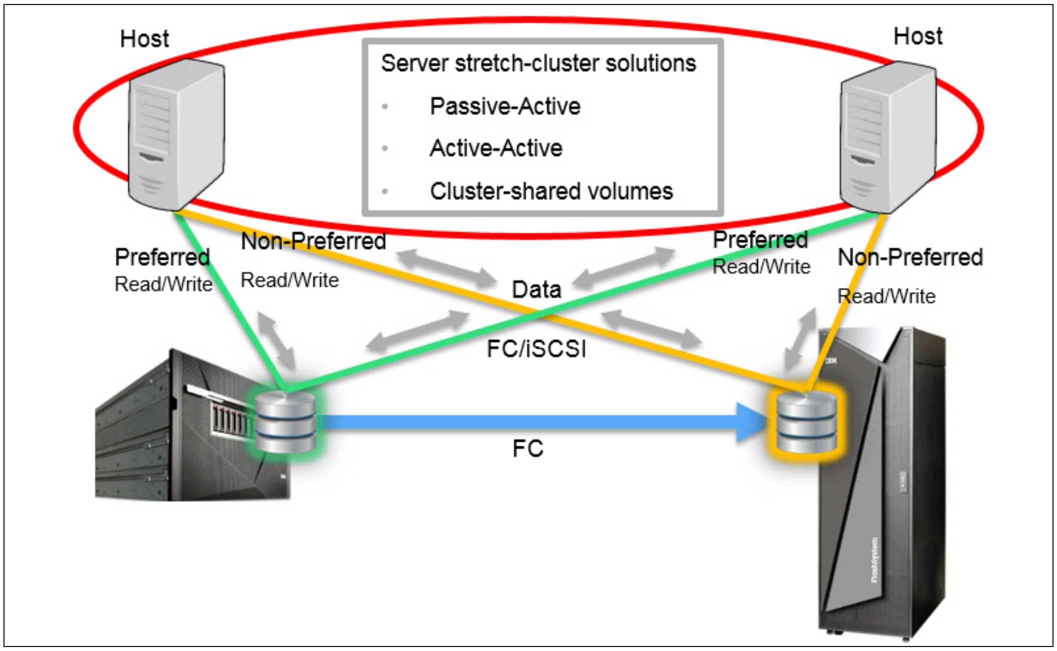


Figure 1-4 Stretch-cluster server configurations

1.3 Independent QW

An independent QW software component is used as part of the HyperSwap solution in order to enable transparent failover and facilitate coordination between two storage systems with volumes in a HyperSwap relationship. The QW role is to continuously monitor the systems health by using dedicated heartbeat messages and, when needed, to serve as a tie-breaker or ruler in possible split-brain scenarios.

The most fundamental function of the QW is to determine upon failure, per HyperSwap volume or Consistency Group (CG), which system should own the primary volume or CG and which system should own the secondary volume or CG.

Whenever a failure is detected at the system level, the QW is used immediately by the storage systems to resolve any potential contention and to determine which system should own each primary volume or CG. The QW component is, therefore, critical for the complete always-on HyperSwap functionality.

The two systems send keepalive messages to report their health to the QW, as shown in Figure 1-5. The systems also retrieve the information pertaining to the health of their peer systems from the QW. The communication with the QW is performed over Ethernet using a TCP/IP. A system can detect the failure of its peer system through the mirroring connectivity between them. If the peer failure is verified through the QW information, the system then takes over as owner of primary volumes or Consistency Groups, provided that it has a synchronized copy of the volume or Consistency Group data.

Some rare failure scenario might result in both systems attempting to take over the primary volumes. The QW is used to tie-break these cases and make sure a single system wins and becomes the primary for the volume or CG.

The automatic recovery ensures that, upon failure, the protected host can access at least one of the volume instances.

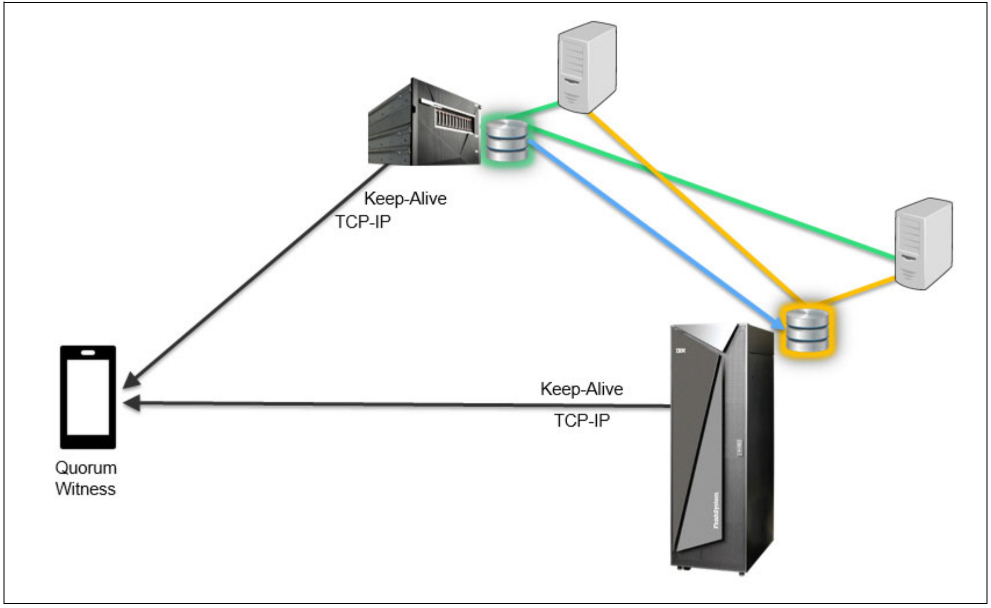


Figure 1-5 Using a keepalive connection with a QW node

The QW node consists of a small-footprint software application, typically deployed separately at a third site.

Tip: The QW is preferably deployed as a highly-available virtual machine on a VMware vSphere 6.1 (or later) cluster, using VMware HA.

Figure 1-6 shows a typical configuration of hosts, storage systems, and a QW node in a HyperSwap high availability solution. Typically, the hosts and the storage systems with primary volumes are located at the same site, and the QW is deployed at a separate third site.

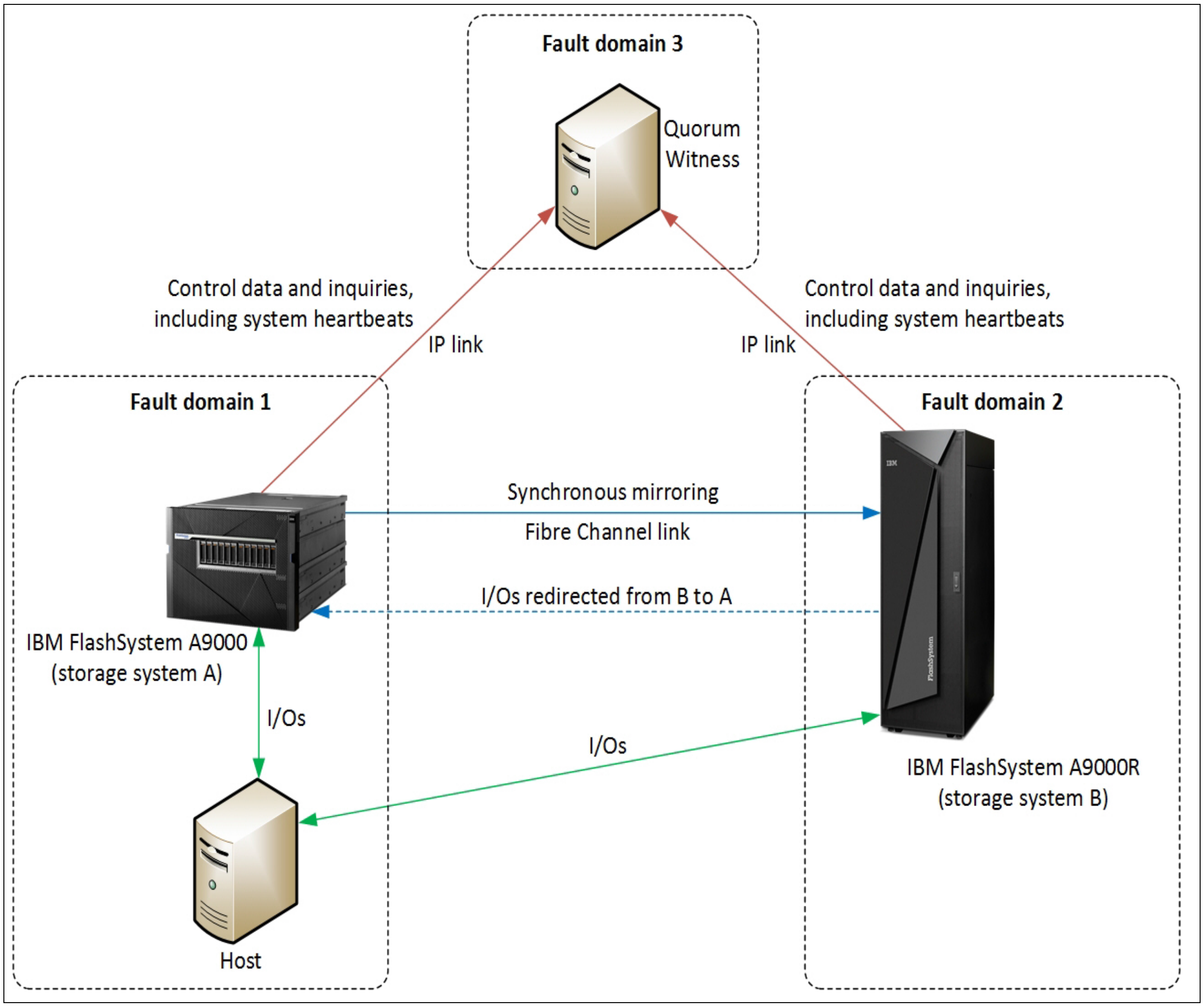

Fault domain 3
Quorum Witness
Control data and inquiries, including system heartbeats
IP link
IP link
Control data and inquiries, including system heartbeats
Fault domain 1
Fault domain 2
Synchronous mirroring
Fibre Channel link
I/Os redirected from B to A
IBM FlashSystem A9000 (storage system A)
I/Os
I/Os
IBM FlashSystem A9000R (storage system B)
Host

Figure 1-6 IBM FlashSystem A9000 and A9000R with QW deployment in separate fault domains

1.4 HyperSwap various topologies

HyperSwap on IBM FlashSystem A9000 and A9000R supports various topologies. We review in this section, the most typical ones.

1.4.1 Conventional topology

In a conventional topology, as shown in Figure 1-7, primary volumes or Consistency Groups are located in one of the systems, and the secondary volumes or Consistency Groups are located in the other system.

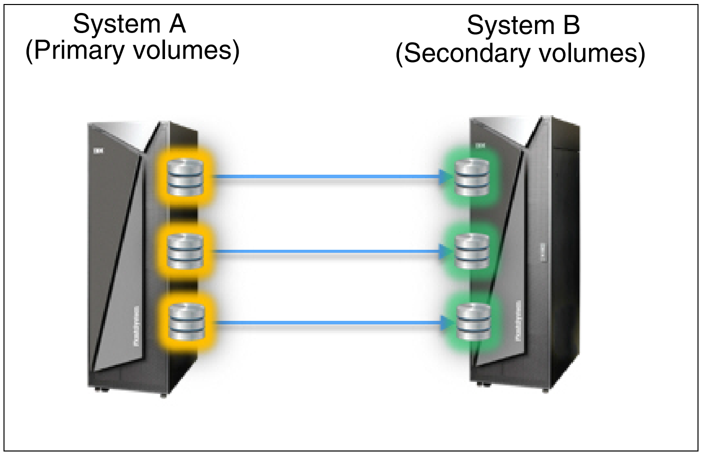


Figure 1-7 Conventional HyperSwap topology

1.4.2 Symmetrical topology

In a symmetrical topology, as shown in Figure 1-8, the systems have both primary and secondary volumes or Consistency Groups. This configuration can be used to split production workload over the two systems and thus balance the I/O utilization of the configuration.

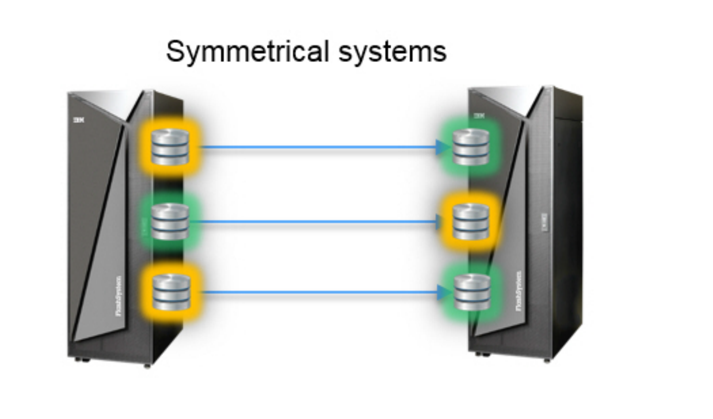


Figure 1-8 Symmetrical topology

1.4.3 Dedicated system topology

In a dedicated system topology, as shown in Figure 1-9, a single storage system at a third site is dedicated to serving two primary systems simultaneously, and is initially configured to contain only the *secondary volumes*.

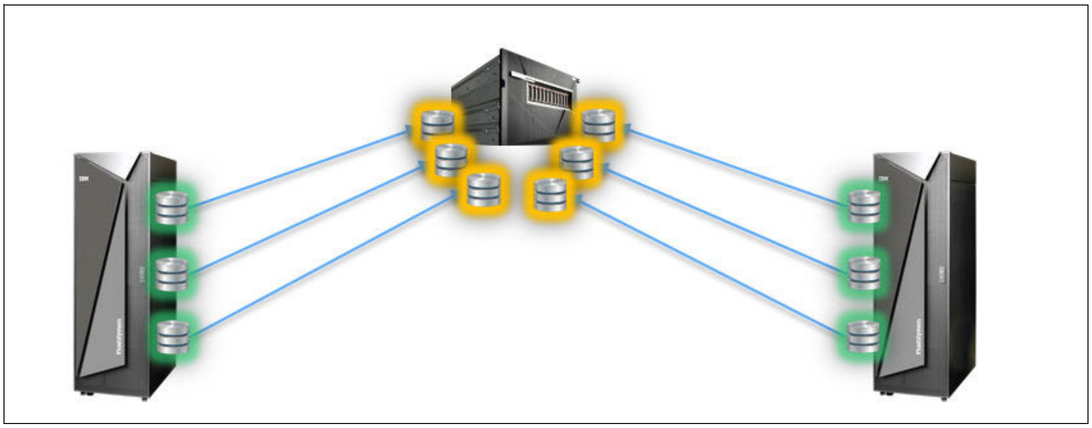


Figure 1-9 Dedicated HyperSwap system for secondary volumes topology

2

Multi-site HA/DR solution for FlashSystem A9000 and A9000R

This chapter gives a general overview of the Multi-site HA/DR feature and its architecture.

This chapter covers the following topics:

2.1 Multi-site HA/DR feature overview

Introduced with the IBM FlashSystem A9000 and A9000R storage system software version 12.3.0, the Multi-site HA/DR feature allows the deployment of well-established high availability (HA) and disaster recovery (DR) solutions over multiple sites, keeping three concurrent copies of data.

As shown in Figure 2-1, a Multi-site HA/DR configuration is composed of three IBM FlashSystem A9000 or A9000R systems in a multi-site topology. It comprises:

- One HyperSwap relationship between volumes on systems A and B
- One active asynchronous mirroring relationship between volumes on systems A and C
- Optionally, one standby asynchronous mirroring relationship between volumes on systems B and C

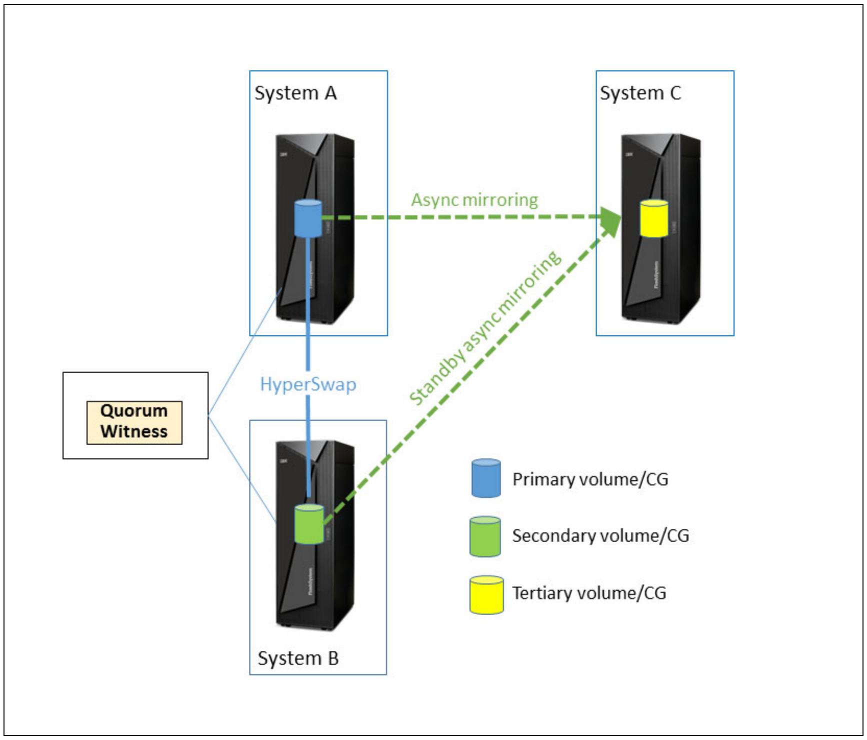


Figure 2-1   Basic Multi-site HA/DR configuration

This configuration provides HyperSwap active-active high availability, while keeping data mirrored to a third site to ensure two levels of business continuity.

System B holds the most recent snapshot and last replicated snapshots that were automatically created by the mirroring relationship with volumes on System A. Upon a HyperSwap failover, the asynchronous replication of volumes from System A to System C stops, and the standby asynchronous relation between corresponding volumes on System B and System C is automatically activated. At this point, the MRS and LRS on System B are used to start the asynchronous replication with System C from the same point where the System A to System C replication stopped.

The Multi-site HA/DR feature can operate with volumes and consistency groups.

Any storage system in a multi-site relation can have multiple multi-site relations with volumes and consistency groups (CG) in different roles.

2.2 Architecture, design, and components

A multi-site configuration is typically defined by extending a HyperSwap relationship or an asynchronous mirror. That is, defining the multi-site relationship assumes that at least a fully initialized two-way HyperSwap or asynchronous mirroring relationship exists.

The multi-site relationship runs on system A (see Figure 2-1 on page 14) by activating the A-B and A-C relationships. If one of these relationships is active, no further changes to that relationship are required.

Important: Running the same multi-site configuration on system B or C is not permitted.

In a multi-site configuration, the asynchronous replication can be over Fibre Channel (FC) or iSCSI; the HyperSwap connection must be FC only.

2.2.1 Roles in a multi-site relationship

The Multi-site HA/DR feature introduces some new terms in addition to the familiar HyperSwap and asynchronous mirroring terminology.

As shown in Figure 2-1 on page 14, the volume/CG roles feature the following definitions:

- Primary volume

  The primary volume in a HyperSwap relationship with the secondary volume, and in an asynchronous mirroring relationship with the tertiary volume. Normally, the primary volume is on system A.

- Secondary volume

  The secondary volume in a HyperSwap relationship between A and B. This volume also acts as the primary volume in the optional standby asynchronous mirroring relationship between B and C. Normally, the secondary volume resides on system B.

- Tertiary volume

  The secondary volume in two asynchronous mirroring relationships, between A and C (active) and B and C (standby). Normally, the tertiary volume is on system C.

In a normal, properly operating multi-site relationship, the role of each volume is as indicated here.

Important: As a counterpart to *Primary*, *Secondary*, and *Tertiary*, the input and output syntax of CLI commands use the established terminology of *Master*, *sMaster*, and *Slave*. This inconsistency is a necessary compromise that is required to avoid changes to older CLI commands that are used by customers, and to keep the CLI terminology consistent.

2.2.2 Standby asynchronous mirror

The standby asynchronous mirror between systems B and C can be defined in advance, at the time of the multi-site relationship assembly, or when needed for data recovery.

Defining the standby mirror in advance requires that the target connectivity between the systems B and C (or at least its definitions) be in place when the multi-site relationship is assembled. When defined, the B-C asynchronous mirror remains on standby under normal conditions and only becomes active (by request) when a disaster occurs on system A, which makes the primary volume inaccessible and triggers a failover in the A-B HyperSwap relationship.

The preferred practice is to define the standby asynchronous relationship in advance.

2.3 Multi-site HA/DR advantages

As implemented in IBM FlashSystem A9000 and A9000R, the Multi-site HA/DR solution offers the following advantages that are described next.

2.3.1 Flexibility

Any of the three peer systems that are participating in a multi-site relationship can be BM FlashSystem A9000 and A9000R. System models do not have to be identical. That is, a multi-site relationship peer system can be model 415 or 425.

The Multi-site HA/DR feature does not require separate licensing.

2.3.2 Non-disruptiveness

The Multi-site HA/DR solution is designed to maintain nondisruptive host data access and recover data in case of a disaster.

The administrator assembles it from existing relationships. Any existing two-way relationship (HyperSwap or asynchronous) can be extended to a multi-site relationship without disrupting the existing relation.

2.3.3 Ease of use

Multi-site HA/DR is easy to configure and monitor because its administration and management are based on existing features.

IBM Hyper-Scale Manager provides special management for the Multi-site HA/DR solution.

3

Prerequisites

This chapter covers the prerequisites and dependencies for implementing the IBM HyperSwap with the IBM FlashSystem A9000 and A9000R systems.

This chapter includes the following topics:

- 3.1, “Storage systems and Quorum Witness prerequisites” on page 18
- 3.2, “Software levels” on page 20
- 3.3, “Network requirement” on page 21
- 3.4, “Scalability” on page 22

3.1 Storage systems and Quorum Witness prerequisites

This section covers the requirements and prerequisites for FlashSystem A9000 and A9000R and the Quorum Witness (QW) to use the IBM HyperSwap feature.

3.1.1 FlashSystem A9000 and A9000R prerequisites

For any volume or Consistency Group, HyperSwap is implemented only between two FlashSystem A9000 and A9000R systems (any combination of the two). It includes the following requirements specific to these storage systems:

- Minimum software level to support HyperSwap is Version 12.1 and Hyper-Scale Manager V5.2 or later is required.
- For Multi-site HA/DR, any FlashSystem A9000 or A9000R participating in the multi-site solution must all be running at least Version 12.3 of the storage system software. To manage a multi-site environment, Hyper-Scale Manager V5.5 or later is required.
- When defining hosts, ensure that the correct type is selected. If you use Windows 2008, HP-UX, or IBM z/VM® host, you must change the type to match your host type. For all other hosts, such as IBM AIX®, Linux, Solaris, VMware, and Windows (except Windows 2008), the Default (or All Others) option is correct.

  Note: Starting with Host Attachment Kit (HAK) version 2.7.0, HP-UX and Solaris are no longer supported.

- PKCS12 certificates must be installed and enabled for the QW service, with the following different certificates of importance:
  - The QW certificate that is generated by the QW during QW installation must be copied to the local machine where you currently use Hyper-Scale Manager (see step 8 on page 28 in Chapter 4, “HyperSwap implementation and usage” on page 23). You also need to install this certificate on each FlashSystem A9000 or A9000R system using the corresponding QW (see 4.3, “Configuring the Quorum Witness on storage systems” on page 29).
  - On FlashSystem A9000 or A9000R systems where the HyperSwap volumes reside, the system certificate must be assigned to the QUORUM service:
    - For systems shipped with a minimum of Version 12.1 code, the certificate that was preinstalled by manufacturing is already assigned to the QUORUM service.
    - For older systems, you must assigned the certificate to the QUORUM service after the code is upgraded.

    In any case, use the XCLI **pki_list** command to check which services are already enabled as follows:

```
A9000R>>pki_list
Name Fingerprint                           Has signed certificate  Services
XIV  c5dfc4f842946b669314e544d189c6f7       yes                    XCLI,CIM,IPSEC
```

    Then, if QUORUM is not listed among the services, use the **pki_update** command to use the same certificate (XIV, in this example) for the QUORUM service as well:

```
A9000R>>pki_update name=XIV services=XCLI,CIM,IPSEC,QUORUM
Command executed successfully.
```

To check whether the QUORUM service was added, issue the **pki_list** command again:

```
A9000R>>pki_list
Name Fingerprint                         Has signed certificate  Services
XIV  c5dfc4f842946b669314e544d189c6f7    yes                     XCLI,CIM,IPSEC,QUORUM
```

- A minimum of four Fibre Channel ports must be used for replication between the two systems, set up as initiator and target on each system.
- A maximum of two QW servers per storage system.
- All management ports (three ports per system, thus a total of six) must be connected and configured.
- Management ports for the active QW must be connected and configured.

Note: Always refer to the latest storage system release notes to check for HyperSwap support. You can find release notes in IBM Knowledge Center.

3.1.2 Quorum Witness prerequisites

The QW is used to enable a non-disruptive, automatic failover in HyperSwap. Without a functioning QW, automatic failover cannot be performed. Its role in the solution is to track and report on the health of both storage systems and to serve as a tie-breaker in split-brain scenarios.

The QW application can be installed on a dedicated server or a virtual machine (VM) with the following minimum hardware requirements:

- RHEL 6.8 or 7.3, CentOS 6.8 or 7.3
- 64-bit dual-core CPU
- 4 GB RAM
- Hard drive requirements:
  - SAS/SATA interface
  - at least 7200 RPM
  - RAID protection
  - 40 GB free space
- 1 Gbps broadband network connection

Because the solution is implemented with an application for the QW, the user installing the application must have relevant permissions on the host (physical or virtual) to install the RPM and BIN files.

Installing the QW application on a highly available VM will ensure the QW will still be available in the event of a hardware problem on the server or VM. However, for the best performance, use a dedicated server.

A maximum of 12 storage systems can be configured per QW.

For more information about which ports to open and installing the application, see *IBM Spectrum Accelerate Family HyperSwap Quorum Witness User Guide*, SC27-4631, which is available at IBM Knowledge Center.

3.2 Software levels

This section covers the different software levels that are required to implement HyperSwap and Multi-site HA/DR.

3.2.1 Host Attachment Kit

The minimum Host Attachment Kit (HAK) version of IBM FlashSystem A9000 or A9000R for HyperSwap is 2.8.

For information about the HAK installation, review *IBM FlashSystem A9000, IBM FlashSystem A9000R, and IBM XIV Storage System: Host Attachment and Interoperability*, SG24-8368. In addition, for VM solutions, refer to *Using the IBM Spectrum Accelerate Family in VMware Environments: IBM XIV, IBM FlashSystem A9000 and IBM FlashSystem A9000R, and IBM Spectrum Accelerate*, REDP-5425.

You can download and install the HAK from Fix Central. For more information about the use of the HAK, see the IBM Storage Host Attachment Kit welcome page in IBM Knowledge Center.

3.2.2 SAN boot

By using SAN boot make sure that in the adapter BIOS targets (WWPNs) from both storage systems are defined. The adapter BIOS can be changed during boot time or by using tools from the HBA vendor running on the operating system.

3.2.3 Quorum Witness software levels

You can install QW software on a bare metal server using compatible versions of Red Hat Enterprise Linux (RHEL), CentOS, or as a VMware virtual machine.

Verify that the dependencies mentioned in *IBM Spectrum Accelerate Family HyperSwap Quorum Witness User Guide*, SC27-4631, are installed prior to installing the QW application itself.

These dependencies as mentioned previously are also supplied in the installation package together with the QW application BIN file and will be deployed prior to the application installation.

3.3 Network requirement

This section lists the configurations that have to be enabled or changed in both storage systems or in the network infrastructure to use a HyperSwap with FlashSystem A9000 and A9000R (any combination of the two).

3.3.1 Storage systems connectivity

The following storage systems connectivity requirements must be met, for HyperSwap:

- Fibre Channel connectivity of 4, 8, or 16 GB between the two FlashSystem A9000 and A9000R systems, for mirror connectivity.
- A minimum of four Fibre Channel ports for each primary and secondary system is required, for mirror connectivity.
- All Fibre Channel ports of each storage system that have volumes in HyperSwap relationships, must be in the same zone or zones.
- Users should have the ability to control the zone settings of the target and initiator FlashSystem A9000 and A9000R systems for failover situations.
- At least two host ports of either iSCSI or Fibre Channel on each target and initiator systems should be available for host application connectivity. In case of Fibre Channel connectivity, Fibre Channel ports on both storage systems should be set as target ports, each connected and zoned with the host application server.
- HyperSwap distances are limited by synchronous mirroring distances (typically up to 100 km apart) depending on the available bandwidth and impact of latency on the host. If the application is latency sensitive, you might want to further limit the distance to a maximum of 75 km.

  For more details, see *IBM FlashSystem A9000 and A9000R Business Continuity Solutions*, REDP-5401.

Consider the following points regarding a multi-site deployment:

- The distance between systems follow the same limitations that apply to HyperSwap and asynchronous mirroring.
- The asynchronous replication can be over Fibre Channel (FC) or iSCSI, while the HyperSwap connection must be FC only.

3.3.2 Quorum Witness network requirements

The QW network requirements are as follows:

- At least one network connection of one Gigabit configured and connected to be able to reach all management ports of both storage systems.
- It requires Secure Sockets Layer (SSL) certification to communicate with the storage systems.
- The connection must have ports TCP 8460 (for QW API communication), TCP 8461 (for retrieving logs), and TCP 8462 (for retrieving `nginx` statistics) unrestricted.
- The maximum latency that is allowed between the QW and FlashSystem A9000 or A9000R is 750 ms.

3.4 Scalability

This section covers the scalability of HyperSwap.

3.4.1 FlashSystem A9000 and A9000R scalability

As with replication in general for the IBM Spectrum Accelerate Family, a single FlashSystem A9000 or A9000R initiator system can have up to 10 target systems. These target systems can be used for mirroring, Hyper-Scale Mobility (FlashSystem A9000 only), and HyperSwap configurations. It is important to remember that HyperSwap can be configured only between two systems for any single volume or Consistency Group; however, a source system can have HyperSwap configured to multiple target systems using different volumes or Consistency Groups.

FlashSystem A9000 and A9000R systems have a maximum number of mirrors that can be configured. This number now includes pairs configured in a HyperSwap relationship due to using synchronous mirroring between the two systems. The maximum number of mirroring relations at the time of writing this document is *1536*. Refer to the latest document release notes for any updates to this number.

Important: Remember that each asynchronous mirror counts for three mirrors (primary volume snapshot, most recent snapshot and last replicated snapshot). If you have any combination of HyperSwap with asynchronous mirroring, be sure to plan accordingly. Refer to *IBM FlashSystem A9000 and A9000R Business Continuity Solutions*, REDP-5401 for more information about maximum mirrors.

3.4.2 Quorum Witness scalability

Although it is recommended to have the QW installed on a highly available virtual machine, there can be only a single QW defined as the active quorum for a system. This must also be the same QW defined on both storage systems configured for HyperSwap.

A single QW server can be defined on up to 12 FlashSystem A9000 and A9000R systems that will be running a HyperSwap configuration.

A second QW can be configured for each FlashSystem A9000 and A9000R that is part of a HyperSwap configuration; however, this configuration is used for redundancy purposes only as a backup system.

To maximize the protection, in a multi-site HA/DR deployment it is better that the three systems and the Quorum witness be deployed in four completely-separate failure domains (typically, separate sites).

4

HyperSwap implementation and usage

This chapter describes how to implement and use IBM HyperSwap for FlashSystem A9000 and FlashSystem A9000R. It covers the following topics:

Note: Most illustrations in this chapter are based on IBM Hyper-Scale Manager Version 5.5. Version 5.2 or later is required for system code level 12.1, which enables the HyperSwap feature.

Version 5.5 is required for system code level 12.3, which enables the Multi-site feature. For more information, see Chapter 6, “Multi-site HA/DR implementation and usage” on page 97.

4.1 HyperSwap implementation process overview

To ensure smooth implementation, make sure to first review Chapter 3, “Prerequisites” on page 17.

Implementing HyperSwap consists of the following steps:

1. IBM FlashSystem A9000 and A9000R HyperSwap initial configuration starts with the installation of the external Quorum Witness application. The Quorum Witness is installed on a dedicated server or as a VMware virtual machine, preferably at a separate site. A separate site is better because it creates another, isolated failure domain.
2. The Quorum Witness is connected to both IBM FlashSystem A9000 or A9000R with volumes to be used in the HyperSwap solution. Each storage system can be configured with one Quorum Witness instance, identified by its unique ID.
3. The volumes from both systems used in HyperSwap relationship must be set as HyperSwap volumes.
4. HyperSwap volumes on each storage system must be mapped to the same application host server.

4.2 Quorum Witness setup

The server on which Quorum Witness (QW) application will be installed must be configured with all the latest updated and prerequisites.

Use the following command to install the required Linux packages:

```
yum install –y <Explicit RPM name> […<Explicit RPM name>]
```

Note: The following example uses RHEL-7.3. For other versions of RHEL, see *IBM Spectrum Accelerate Family HyperSwap Quorum Witness User Guide*.

For RHEL-7, install the following RPMs:

- `erlang-18.1-1.el7.centos.x86_64.rpm`
- `jemalloc-3.6.0-1.el7.x86_64.rpm`
- `nginx-1.10.3-1.el7.ngx.x86_64.rpm`
- `uuid-1.6.2-26.el7.x86_64.rpm`
- `postgresql92-libs-9.2.14-1PGDG.rhel7.x86_64.rpm`
- `postgresql92-9.2.14-1PGDG.rhel7.x86_64.rpm`
- `postgresql92-contrib-9.2.14-1PGDG.rhel7.x86_64.rpm`
- `postgresql92-server-9.2.14-1PGDG.rhel7.x86_64.rpm`
- `rabbitmq-server-3.6.0-1.noarch.rpm`
- `redis-3.2.5-1.el7.x86_64.rpm`
- `libxslt-1.1.28-5.el7.x86_64 (Only for CentOS)`
- `bzip2`

Tip: RabbitMQ code can be downloaded from the RabbitMQ website.

4.2.1 Downloading and installing Quorum Witness

Follow these steps to install QW:

1. Open TCP ports 8460 and 8641, as shown in Example 4-1. For retrieving nginx statistics, you also need to open TCP port 8462.

   > **Note:** If you are using a different firewall software, refer to your firewall system documentation for specific instructions about how to open TCP ports.

   Example 4-1 Opening the firewall ports on RHEL 7.3

   ```
   [root@QW2 rhel7]# firewall-cmd --permanent --zone=trusted --add-interface=lo
   success
   [root@QW2 rhel7]# firewall-cmd --permanent --add-port=8460/tcp
   success
   [root@QW2 rhel7]# firewall-cmd --permanent --add-port=8461/tcp
   success
   [root@QW2 rhel7]# firewall-cmd --reload
   success
   ```

2. If the rabbitmq-server service is reported as not running, restart it, as shown in Example 4-2.

   Example 4-2 Restarting rabbitmq-server

   ```
   [root@QW2 ~]# systemctl restart rabbitmq-server
   [root@QW2 ~]# service rabbitmq-server status
   Status of node rabbit@QW2 ...
   [{pid,18910},
    {running_applications,[{rabbit,"RabbitMQ","3.6.0"},
                           {mnesia,"MNESIA  CXC 138 12","4.14.3"},
                           {os_mon,"CPO  CXC 138 46","2.4.2"},
                           {ranch,"Socket acceptor pool for TCP protocols.",
                                  "1.2.1"},
                           {xmerl,"XML parser","1.3.13"},
                           {rabbit_common,[],"3.6.0"},
                           {sasl,"SASL  CXC 138 11","3.0.3"},
                           {stdlib,"ERTS  CXC 138 10","3.3"},
                           {kernel,"ERTS  CXC 138 10","5.2"}]},
    {os,{unix,linux}},
    {erlang_version,"Erlang/OTP 19 [erts-8.3] [source] [64-bit] [async-threads:64]
   [hipe] [kernel-poll:true]\n"},
    {memory,[{total,41791760},
             {connection_readers,0},
             {connection_writers,0},
             {connection_channels,0},
             {connection_other,0},
             {queue_procs,2688},
             {queue_slave_procs,0},
             {plugins,0},
             {other_proc,19185560},
             {mnesia,57976},
             {mgmt_db,0},
             {msg_index,32376},
             {other_ets,863976},
   ```

```
          {binary,20592},
          {code,15489762},
          {atom,662433},
          {other_system,5476397}]},
 {alarms,[]},
 {listeners,[{clustering,25672,"::"},{amqp,5672,"::"}]},
 {vm_memory_high_watermark,0.4},
 {vm_memory_limit,771738828},
 {disk_free_limit,50000000},
 {disk_free,13061808128},
 {file_descriptors,[{total_limit,924},
                    {total_used,2},
                    {sockets_limit,829},
                    {sockets_used,0}]},
 {processes,[{limit,1048576},{used,128}]},
 {run_queue,0},
 {uptime,14},
 {kernel,{net_ticktime,60}}]
[root@QW2 ~]#
```

3. Download the installation package to a local folder on the Linux host that will be used as Quorum Witness server. You can download the installation package from IBM Fix Central.
4. Extract the installation package file using the following command:

   ```
   # tar -xzvf ibm_quorum_witness-1.0.0-<build number>-x86_64.tar.gz
   ```

5. Verify that you have relevant permissions on the host.
6. Enter `./ibm_quorum_witness-1.0.0--<build number>.bin` to start the installation, as shown in Example 4-3. Note that you might need to install `bzip2` if it is not already present on your system.

   Example 4-3 Installing by way of the .bin file

   ```
   [root@QW2 ~]# ./ibm_quorum_witness-1.0.0-1751-x86_64.bin
   Verifying dependencies...
   Starting installer, please wait...
   International License Agreement for Non-Warranted Programs

   Part 1 - General Terms
   ....
   ```

7. Review and accept the license agreement which is displayed after you run the installation file, as shown in Example 4-4.

   Example 4-4 Accepting the license agreement

   ```
   Press Enter to continue viewing the license agreement, or, Enter "1" to accept
   the agreement, "2" to decline it or "99" to go back to the previous screen, "3"
   Print, "4" Read non-IBM terms.
   1
   ```

   After installation, the Quorum Witness application starts automatically.

> **Note:** If you are using SELinux, the program activation can fail. In this case allow nginx service to bind to network interfaces and connect to the QW socket.
>
> ```
> setsebool -P nis_enabled 1 and semodule -i
> /opt/ibm/ibm_quorum_witness/conf.d/selinux/qw_rhel7.pp
> ```
>
> Verify that the policy has been applied correctly by running:
>
> ```
> cat /opt/ibm/ibm_quorum_witness/conf.d/selinux/qw_rhel7.te
> ```
>
> Restart the QW service:
>
> ```
> service ibm_quorum_witness restart
> ```

Refer to Example 4-5 for details.

Example 4-5 Troubleshooting ibm_quorum_witness restart problems

```
[root@QW2 ~]# systemctl status ibm_quorum_witness.service
? ibm_quorum_witness.service - Quorum Witness
   Loaded: loaded (/usr/lib/systemd/system/ibm_quorum_witness.service; enabled;
vendor preset: disabled)
   Active: failed (Result: exit-code) since Wed 2017-04-26 11:55:30 CEST; 1min
36s ago
  Process: 20225 ExecStart=/opt/ibm/ibm_quorum_witness/bin/ibm_quorum_witness
start (code=exited, status=4)

Apr 26 11:55:30 QW2.local systemd[1]: Starting Quorum Witness...
Apr 26 11:55:30 QW2.local ibm_quorum_witness[20225]: Error: nginx service is
not running.
Apr 26 11:55:30 QW2.local ibm_quorum_witness[20225]: Verify that SELinux is
configured to allow Quorum Witness to bind to the TCP ports 8460 8461
Apr 26 11:55:30 QW2.local ibm_quorum_witness[20225]: Refer to the user guide
for details.
Apr 26 11:55:30 QW2.local systemd[1]: ibm_quorum_witness.service: control
process exited, code=exited status=4
Apr 26 11:55:30 QW2.local systemd[1]: Failed to start Quorum Witness.
Apr 26 11:55:30 QW2.local systemd[1]: Unit ibm_quorum_witness.service entered
failed state.
Apr 26 11:55:30 QW2.local systemd[1]: ibm_quorum_witness.service failed.

[root@QW2 ~]# service nginx start
Redirecting to /bin/systemctl start  nginx.service
[root@QW2 ~]# service nginx status
Redirecting to /bin/systemctl status  nginx.service
? nginx.service - nginx - high performance web server
   Loaded: loaded (/usr/lib/systemd/system/nginx.service; enabled; vendor
preset: disabled)
   Active: active (running) since Wed 2017-04-26 12:04:55 CEST; 3s ago
     Docs: http://nginx.org/en/docs/
  Process: 20458 ExecStart=/usr/sbin/nginx -c /etc/nginx/nginx.conf
(code=exited, status=0/SUCCESS)
  Process: 20457 ExecStartPre=/usr/sbin/nginx -t -c /etc/nginx/nginx.conf
(code=exited, status=0/SUCCESS)
 Main PID: 20461 (nginx)
   CGroup: /system.slice/nginx.service
```

```
          ··20461 nginx: master process /usr/sbin/nginx -c
/etc/nginx/nginx.conf
          ··20462 nginx: worker process

Apr 26 12:04:55 QW2.local systemd[1]: Starting nginx - high performance web
server...
Apr 26 12:04:55 QW2.local nginx[20457]: nginx: the configuration file
/etc/nginx/nginx.conf syntax is ok
Apr 26 12:04:55 QW2.local nginx[20457]: nginx: configuration file
/etc/nginx/nginx.conf test is successful
Apr 26 12:04:55 QW2.local systemd[1]: Failed to read PID from file
/run/nginx.pid: Invalid argument
Apr 26 12:04:55 QW2.local systemd[1]: Started nginx - high performance web
server.
[root@QW2 ~]# service ibm_quorum_witness start
Starting ibm_quorum_witness (via systemctl):  Job for
ibm_quorum_witness.service failed because the control process exited with error
code. See "systemctl status ibm_quorum_witness.service" and "journalctl -xe"
for details.
                                                           [FAILED]
[root@QW2 ~]# systemctl status ibm_quorum_witness.service
? ibm_quorum_witness.service - Quorum Witness
   Loaded: loaded (/usr/lib/systemd/system/ibm_quorum_witness.service; enabled;
vendor preset: disabled)
   Active: failed (Result: exit-code) since Wed 2017-04-26 12:05:22 CEST; 30s
ago
  Process: 20507 ExecStart=/opt/ibm/ibm_quorum_witness/bin/ibm_quorum_witness
start (code=exited, status=4)

Apr 26 12:05:22 QW2.local systemd[1]: Starting Quorum Witness...
Apr 26 12:05:22 QW2.local ibm_quorum_witness[20507]: Error: redis service is
not running.
Apr 26 12:05:22 QW2.local ibm_quorum_witness[20507]: The required service may
have failed to start.
Apr 26 12:05:22 QW2.local ibm_quorum_witness[20507]: Refer to the user guide
for details.
Apr 26 12:05:22 QW2.local systemd[1]: ibm_quorum_witness.service: control
process exited, code=exited status=4
Apr 26 12:05:22 QW2.local systemd[1]: Failed to start Quorum Witness.
Apr 26 12:05:22 QW2.local systemd[1]: Unit ibm_quorum_witness.service entered
failed state.
Apr 26 12:05:22 QW2.local systemd[1]: ibm_quorum_witness.service failed.
[root@QW2 ~]# service redis start
Redirecting to /bin/systemctl start  redis.service
[root@QW2 ~]# service ibm_quorum_witness start
Starting ibm_quorum_witness (via systemctl):               [  OK  ]
[root@QW2 ~]#
```

8. Copy the certificate file (`qw.crt`) from the `/opt/ibm/ibm_quorum_witness/settings/ssl_cert` directory to the local workstation where you currently use Hyper-Scale Manager. That certificate will later be added to both A9000/A9000R systems participating in the HyperSwap relationship (see 4.3, “Configuring the Quorum Witness on storage systems” on page 29).

4.3 Configuring the Quorum Witness on storage systems

After you install the Quorum Witness service on the Linux server, you must define it on the FlashSystem A9000 and A9000R that will be involved in the HyperSwap configuration. You must define volumes as HyperSwap volumes and map them to the host.

4.3.1 Configuring FlashSystem A9000 and A9000R HyperSwap

Defining the Quorum Witness server on the storage systems can be done with the Hyper-Scale Manager GUI or using the XCLI.

Configuration that uses the Hyper-Scale Manager GUI

Complete the following steps for each storage system:

1. Navigate to the Systems view, as shown in Figure 4-1.

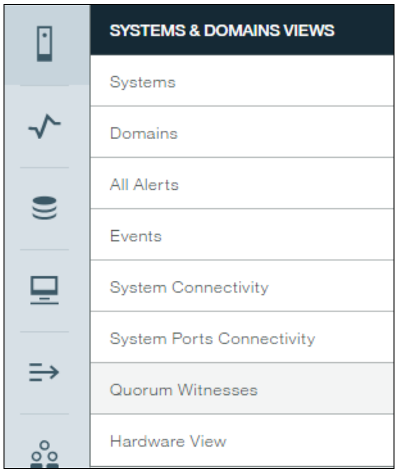


Figure 4-1 Navigating to the Systems view

2. Right-click the first FlashSystem A9000 or A9000R to be used in the HyperSwap relationship. From the drop-down menu, select **Quorum Witness** and then **Define Quorum Witness**, as shown in Figure 4-2.

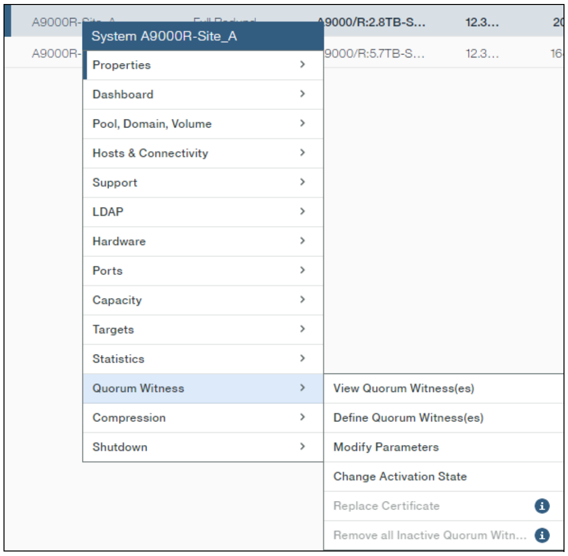
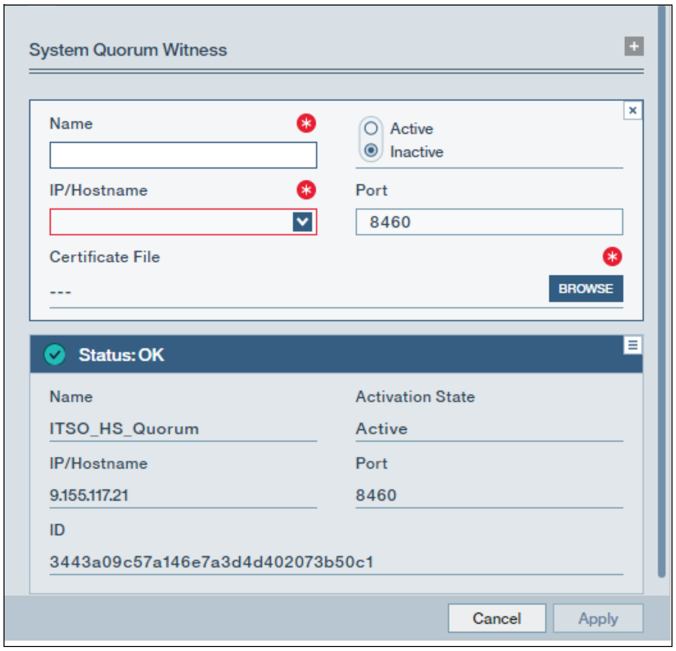


Figure 4-2 Define a Quorum Witness

3. The System Quorum Witness panel displays, as shown in Figure 4-3. You must enter the Quorum Witness name, the IP address of the Quorum Witness server, and the port number (default is 8460).



System Quorum Witness

Name

○ Active
● Inactive

IP/Hostname

Port 8460

Certificate File ---

BROWSE

Status: OK

Name: ITSO_HS_Quorum

Activation State: Active

IP/Hostname: 9.155.117.21

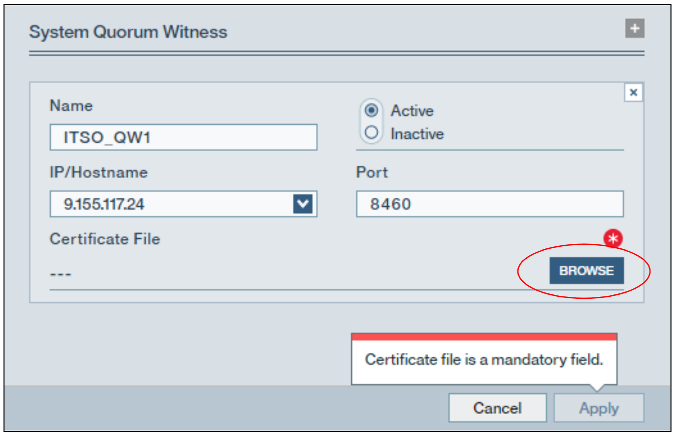
Port: 8460

ID: 3443a09c57a146e7a3d4d402073b50c1

Cancel Apply

Figure 4-3 Quorum Witness definition form

4. Select **Browse** to specify the mandatory Certificate file name, as shown in Figure 4-4.



Name: ITSO_QW1

● Active ○ Inactive

IP/Hostname: 9.155.117.24

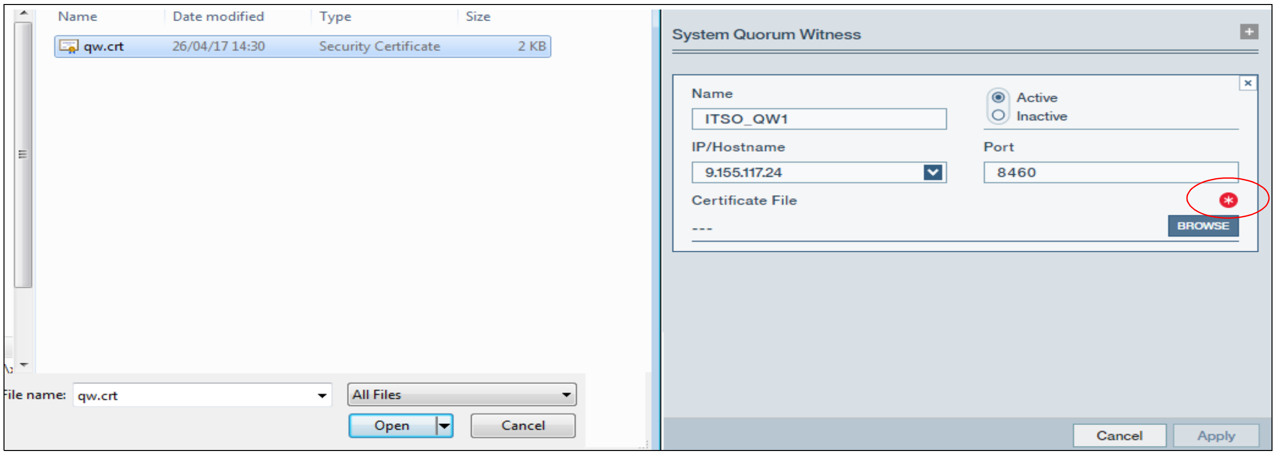
Port: 8460

Certificate File: ---

Certificate file is a mandatory field.

Figure 4-4 Inputting QW definitions

5. Point to the QW Certificate file, as shown in Figure 4-5, and click **Open**.



Name	Date modified	Type	Size
qw.crt	26/04/17 14:30	Security Certificate	2 KB

File name: qw.crt

Name: ITSO_QW1

● Active ○ Inactive

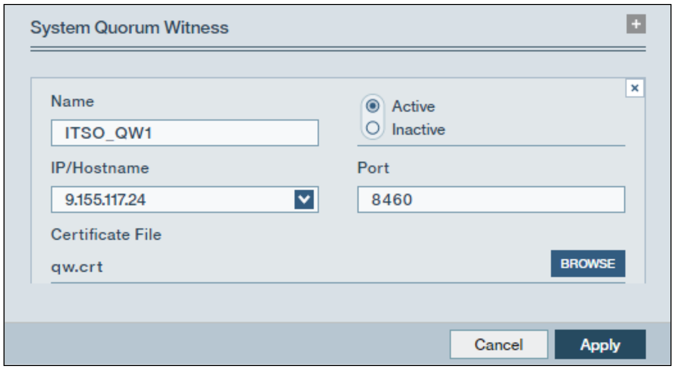
IP/Hostname: 9.155.117.24

Port: 8460

Certificate File: ---

Figure 4-5 Pointing to the qw.cert exported from the Quorum Witness

6. Click **Apply**, as shown in Figure 4-6.



Name: ITSO_QW1

● Active ○ Inactive

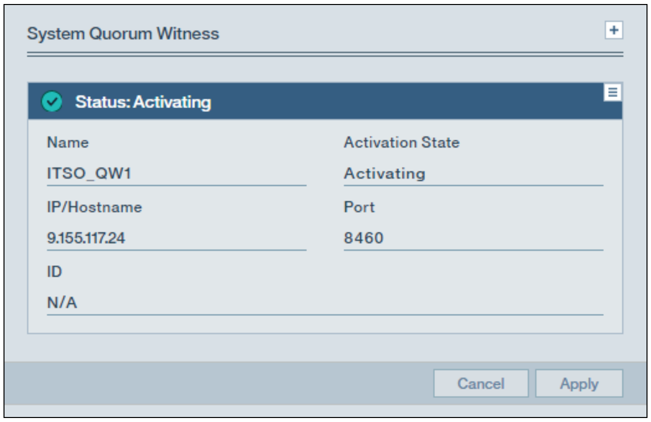
IP/Hostname: 9.155.117.24

Port: 8460

Certificate File: qw.crt

Figure 4-6 Adding the certificate to the Quorum Witness

7. The panel then shows the **Activating** status as illustrated in Figure 4-7.



System Quorum Witness

Status: Activating

Name: ITSO_QW1

Activation State: Activating

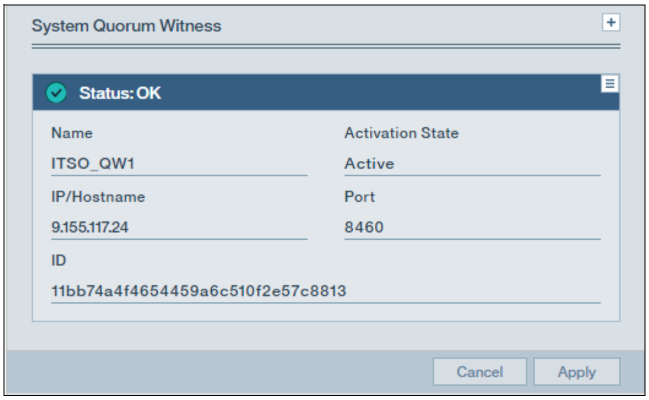
IP/Hostname: 9.155.117.24

Port: 8460

ID: N/A

Figure 4-7 Quorum Witness in Activating state

It finally goes to **Active** status, as shown in Figure 3-8.



System Quorum Witness

Status: OK

Name: ITSO_QW1

Activation State: Active

IP/Hostname: 9.155.117.24

Port: 8460

ID: 11bb74a4f4654459a6c510f2e57c8813

Figure 4-8 Quorum Witness in Active state

8. Repeat these steps on the second FlashSystem A9000 or A9000R.

Important: Make sure that you applied the FlashSystem A9000 or A9000R certificate to the QUORUM service. This process is normally the case on systems that shipped with software Version 12.1 or later. For systems that were upgraded from a lower version, use the **`pki_update`** command to add the QUORUM service. Refer to 3.1.1, "FlashSystem A9000 and A9000R prerequisites" on page 18 for more information.

For verification, follow these steps:

1. Go to the SYSTEMS & DOMAINS VIEWS, and then select **Quorum Witnesses,** as shown in Figure 4-9.

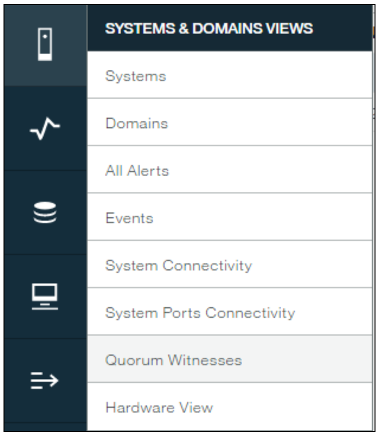
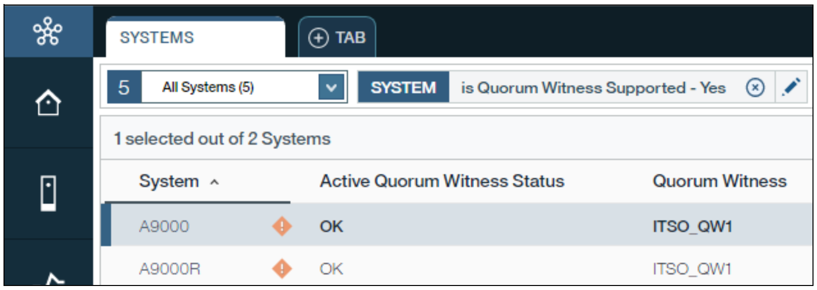


Figure 4-9 Navigating to the Quorum Witnesses

Active Quorum Witness Status should be **OK** and Quorum Witness should reflect the defined Quorum Witness Name, as shown in Figure 4-10.



1 selected out of 2 Systems

System ^	Active Quorum Witness Status	Quorum Witness
A9000	OK	ITSO_QW1
A9000R	OK	ITSO_QW1

Figure 4-10 Showing Quorum Witness status

To get the same information from the XCLI, use the `quorum_witness_list` command.

2. To attach target to QW, select the FlashSystem A9000 or A9000R to own the primary volume in the relationship. (You can change this option later by performing a Switch Role action.) On the right panel, select the Targets arm, as shown in Figure 4-11.

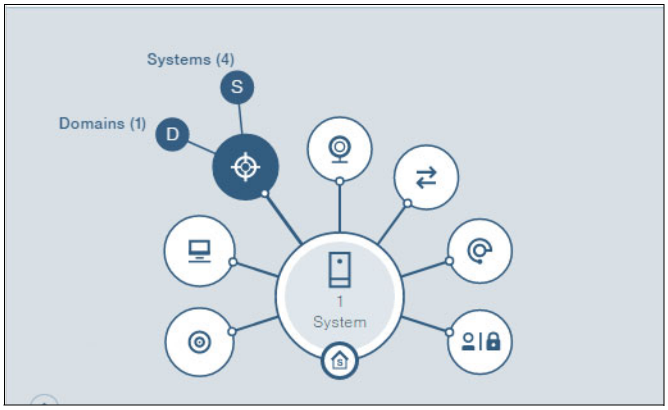
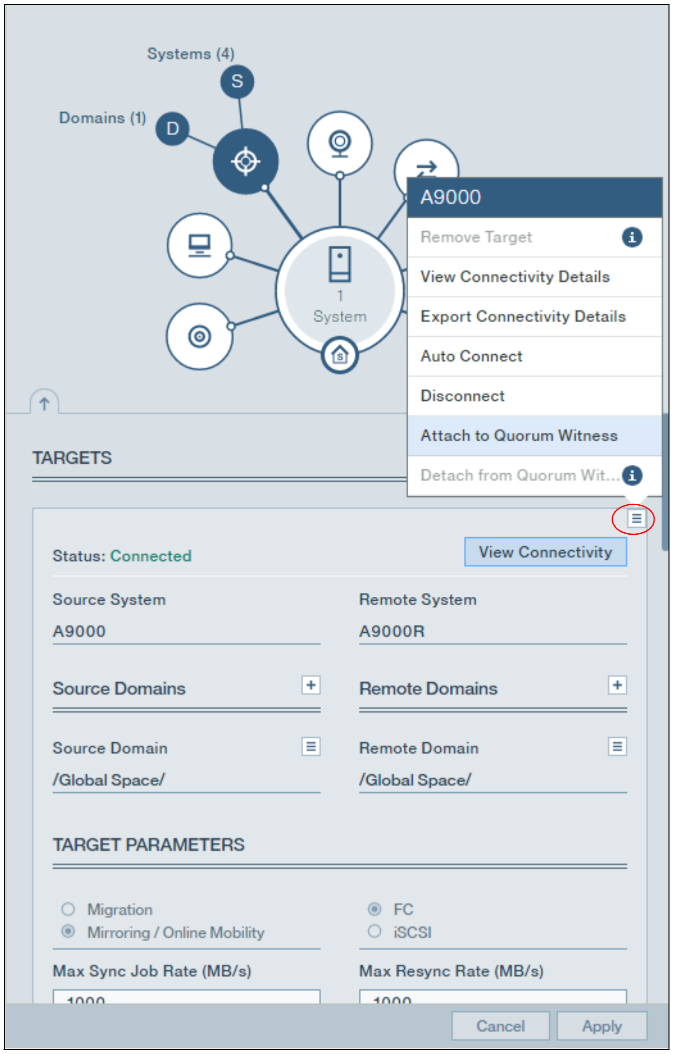


Figure 4-11 Selecting the Targets arm

3. In the target object window frame, select the menu box (the three small horizontal lines) in the top right corner, and select **Attach to Quorum Witness**, as shown in Figure 4-12.

> **Note:** It is necessary to run through the steps to attach the Quorum Witness.



Status: Connected

Source System: A9000

Remote System: A9000R

Source Domains

Remote Domains

Source Domain: /Global Space/

Remote Domain: /Global Space/

TARGET PARAMETERS

○ Migration
● Mirroring / Online Mobility

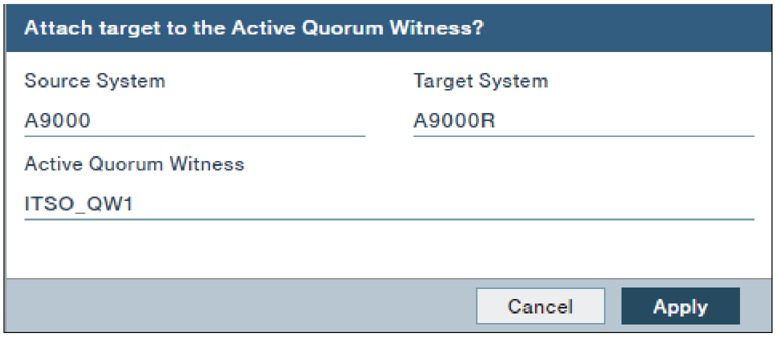
● FC
○ iSCSI

Max Sync Job Rate (MB/s): 1000

Max Resync Rate (MB/s): 1000

Figure 4-12 Attaching the Quorum Witness

4. Verify that the information presented in the panel is correct, and then, click **Apply,** as shown in Figure 4-13.



Attach target to the Active Quorum Witness?

Source System: A9000

Target System: A9000R

Active Quorum Witness: ITSO_QW1

Figure 4-13 Quorum Witness attachment verification

Configuration that uses the XCLI

How to configure the Quorum Witness by using the XCLI is shown in Example 4-6.

Example 4-6 XCLI Configuring the Quorum witness in XCLI

```
Primary System:
A9000>>quorum_witness_define address="9.155.117.24" port="8460" name="ITSO_QW1"
certificate="-----BEGIN
CERTIFICATE-----*MIIC6jCCAdKgAwIBAgIJAIw2396va8jRMAOGCSqGSIb3DQEBCwUAMCOxCzAJBgNV*
BAYTAlVTMQ4wDAYDVQQKDAVJQk1RVzEOMAwGA1UEAwwFSUJNUVcwHhcNMTcwNDI2*MDkOOTQ2WhcNMjcwN
DIOMDkOOTQ2WjAtMQswCQYDVQQGEwJVUzEOMAwGA1UECgwF*SUJNUVcxDjAMBgNVBAMMBUlCTVFXMIIBIj
ANBgkqhkiG9wOBAQEFAAOCAQ8AMIIB*CgKCAQEAzKs+WKCSnVD+tcac1ZdznrbKiZjwynrqMQuffLdIeUi
7G/D5MtJjWGpk*EYRyBpJHhp1EowPEHaizZ9rWX/zMmc2cJRhZMV4qz+3IHhDPOyYYQlaT7XOg4g6l*RWv
LpP/AVwAUdP3iib74ZhmFpaATsntrxqXh97hGLhnm9LuaEpG/Bvc1xJDCDR+u*F9aPr8PKf1L7w5e/J2Bs
7TzMFUhnXqt71MtdjlvV8nIKaLSIM7BV3wG3pB9CgrmR*MOCL7N23qpFpFvvMslmNCIyIZPJRrrKaOHcg1
f+Y5YA9Jl8c9tKsmmRck/sBuxJ7*gObFy3Is6kOdregV3bA5pVs/SeXOdwIDAQABowOwCzAJBgNVHRMEAj
AAMAOGCSqG*SIb3DQEBCwUAA4IBAQAw6cNed44ZDedfWdgy3QLMmEV5fWW2I3Bsrtmyx+OOaRKD*ibfOfi
snnWbfHmrd9aBniqdJOaXMShhwEpjkW1Vyoudy67E6rZH6f2brmF/BhsDE*j3uMSDoBYpU1eMOKAELMcgm
uYToJBgTAiZnnYpqn4mktWOOLhzvQNXt1mvsFdVuQ*vfeWOstX7tTA17Onw3u937zwZ8YCtb+U9uqVSn85
cFXK6il5sqk1633wpaB6odxN*8FTREYhMsZTqjPoqp1yu75w7dIND2UeSwy5EAh8Nth+s5j/vzgBrlY7wv
Q/y/Jsw*+eU2g22Mp8Up7ysNXRebSVu4/LLKZJ1wfofL4AxX*-----END CERTIFICATE-----*"
activate="yes"
Command executed successfully.

Secondary System:
A9000R>>quorum_witness_define address="9.155.117.24" port="8460" name="ITSO_QW1"
certificate="-----BEGIN
CERTIFICATE-----*MIIC6jCCAdKgAwIBAgIJAIw2396va8jRMAOGCSqGSIb3DQEBCwUAMCOxCzAJBgNV*
BAYTAlVTMQ4wDAYDVQQKDAVJQk1RVzEOMAwGA1UEAwwFSUJNUVcwHhcNMTcwNDI2*MDkOOTQ2WhcNMjcwN
DIOMDkOOTQ2WjAtMQswCQYDVQQGEwJVUzEOMAwGA1UECgwF*SUJNUVcxDjAMBgNVBAMMBUlCTVFXMIIBIj
ANBgkqhkiG9wOBAQEFAAOCAQ8AMIIB*CgKCAQEAzKs+WKCSnVD+tcac1ZdznrbKiZjwynrqMQuffLdIeUi
7G/D5MtJjWGpk*EYRyBpJHhp1EowPEHaizZ9rWX/zMmc2cJRhZMV4qz+3IHhDPOyYYQlaT7XOg4g6l*RWv
LpP/AVwAUdP3iib74ZhmFpaATsntrxqXh97hGLhnm9LuaEpG/Bvc1xJDCDR+u*F9aPr8PKf1L7w5e/J2Bs
7TzMFUhnXqt71MtdjlvV8nIKaLSIM7BV3wG3pB9CgrmR*MOCL7N23qpFpFvvMslmNCIyIZPJRrrKaOHcg1
f+Y5YA9Jl8c9tKsmmRck/sBuxJ7*gObFy3Is6kOdregV3bA5pVs/SeXOdwIDAQABowOwCzAJBgNVHRMEAj
AAMAOGCSqG*SIb3DQEBCwUAA4IBAQAw6cNed44ZDedfWdgy3QLMmEV5fWW2I3Bsrtmyx+OOaRKD*ibfOfi
snnWbfHmrd9aBniqdJOaXMShhwEpjkW1Vyoudy67E6rZH6f2brmF/BhsDE*j3uMSDoBYpU1eMOKAELMcgm
uYToJBgTAiZnnYpqn4mktWOOLhzvQNXt1mvsFdVuQ*vfeWOstX7tTA17Onw3u937zwZ8YCtb+U9uqVSn85
cFXK6il5sqk1633wpaB6odxN*8FTREYhMsZTqjPoqp1yu75w7dIND2UeSwy5EAh8Nth+s5j/vzgBrlY7wv
```

```
Q/y/Jsw*+eU2g22Mp8Up7ysNXRebSVu4/LLKZJ1wfofL4AxX*-----END CERTIFICATE-----*"
activate="yes"
Command executed successfully.
```

How to define Quorum Witness on the storage systems is shown in Example 4-7.

Example 4-7 Defining the Quorum Witness in XCLI

```
Primary System:
A9000>>target_add_quorum_witness quorum_witness="ITSO_QW1" target="A9000R"
Command executed successfully.

Secondary System:
A9000R>>target_add_quorum_witness quorum_witness="ITSO_QW1" target="A9000"
Command executed successfully.
```

4.4 HyperSwap volume and Consistency Groups

This section discusses the creation, activation, and other operations relevant to HyperSwap volumes and Consistency Groups, such as converting synchronously mirrored volumes to HyperSwap volumes or vice versa.

4.4.1 Creating and activating HyperSwap volume

You can create and activate HyperSwap volume from the GUI or the XCLI.

Using the GUI

To create a new HyperSwap volume, start by creating a new volume and assigning it to a pool, as usual. After you create the volume, complete the following steps to make it a HyperSwap volume:

> **Note:** In the following example, the volume `ITSO_HA_VOLUME` is created in the pool named *ITSO_HA*.

1. Navigate to the Volumes view in the Hyper-Scale Manager UI, for the FlashSystem A9000 or A9000R the new volume was created on.
2. Right-click the volume name and select **Replication** → **Define/View Replication relation,** as shown in Figure 4-14.

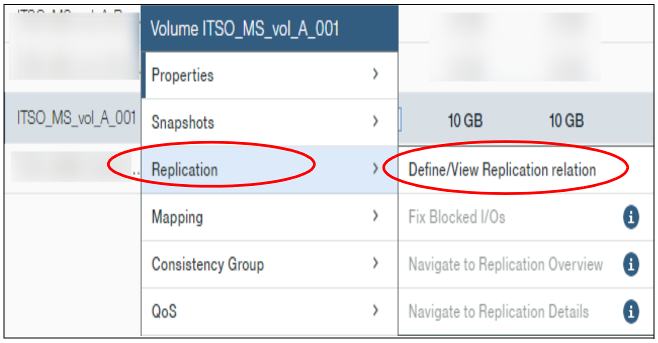
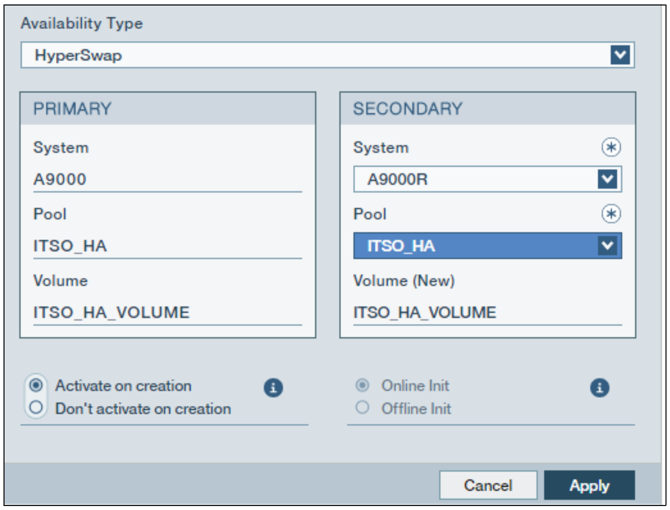


Figure 4-14 Selecting volume to define HyperSwap

3. Select the HyperSwap ADD button.
4. Select the system and the pool with the secondary volumes in that system, making sure to select **Activate on creation**, then click **Apply,** as shown in Figure 4-15.



Availability Type
HyperSwap

PRIMARY
System
A9000
Pool
ITSO_HA
Volume
ITSO_HA_VOLUME

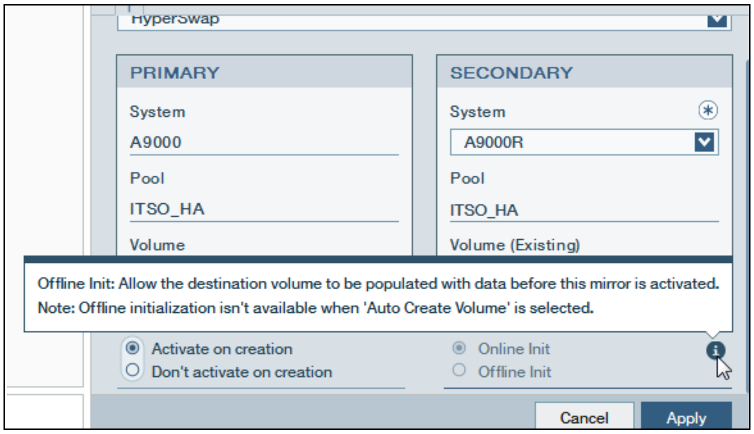
SECONDARY
System
A9000R
Pool
ITSO_HA
Volume (New)
ITSO_HA_VOLUME

● Activate on creation
○ Don't activate on creation

● Online Init
○ Offline Init

Figure 4-15 Volume HyperSwap definition

Alternatively, you can use *offline init* to populate the volume before the mirror is activated, as shown in Figure 4-16.



HyperSwap

PRIMARY
System
A9000
Pool
ITSO_HA
Volume

SECONDARY
System
A9000R
Pool
ITSO_HA
Volume (Existing)

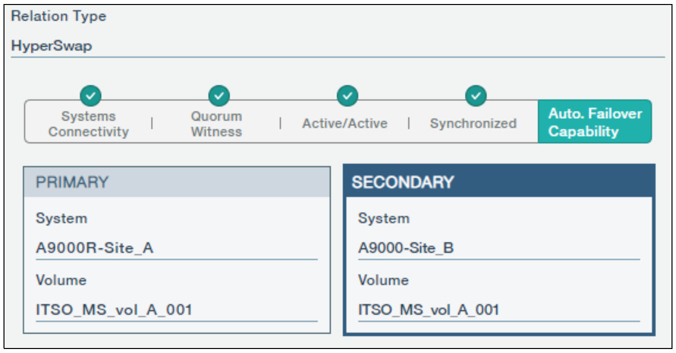
Offline Init: Allow the destination volume to be populated with data before this mirror is activated.
Note: Offline initialization isn't available when 'Auto Create Volume' is selected.

● Activate on creation
○ Don't activate on creation

● Online Init
○ Offline Init

Figure 4-16 Offline Init configuration

The GUI momentarily shows the I/O blocked on the system that owns the secondary volume. Shortly after, the volume shows as converted to HyperSwap, as shown in Figure 4-17.



Relation Type

HyperSwap

PRIMARY	SECONDARY
System	System
A9000R-Site_A	A9000-Site_B
Volume	Volume
ITSO_MS_vol_A_001	ITSO_MS_vol_A_001

Figure 4-17 Volume HyperSwap completion

Using XCLI configuration

Use the commands shown in Example 4-8 create and activate HyperSwap volumes, using the XCLI.

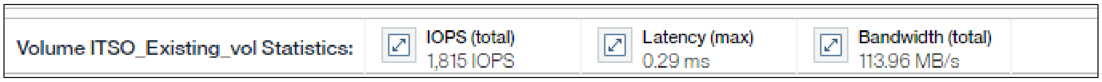
Example 4-8 XCLI commands for HyperSwap volume definition

```
A9000>>ha_create vol=ITSO_HA_VOLUME create_slave=yes remote_pool=ITSO_HA
target=A9000R

Command executed successfully.

A9000>>ha_activate vol=ITSO_HA_VOLUME

Command executed successfully.
```

4.4.2 Converting existing mapped volume to a HyperSwap volume

Note: The following illustrations are represented with a pre-existing volume (ITSO_Existing_vol).

The volume exists on one FlashSystem A9000 or A9000R system that is mapped to a host, and serving I/O, as shown in Figure 4-18.



Volume ITSO_Existing_vol Statistics:

IOPS (total) 1,815 IOPS

Latency (max) 0.29 ms

Bandwidth (total) 113.96 MB/s

Figure 4-18 Volume with active IOPS

The following steps are required to convert the volume to a HyperSwap volume:

1. Navigate to the Volumes view in the Hyper-Scale Manager GUI for the A9000 or A9000R on which the volume resides.
2. Right-click the volume name and select **HyperSwap** then **Define HyperSwap,** as shown in Figure 4-19.

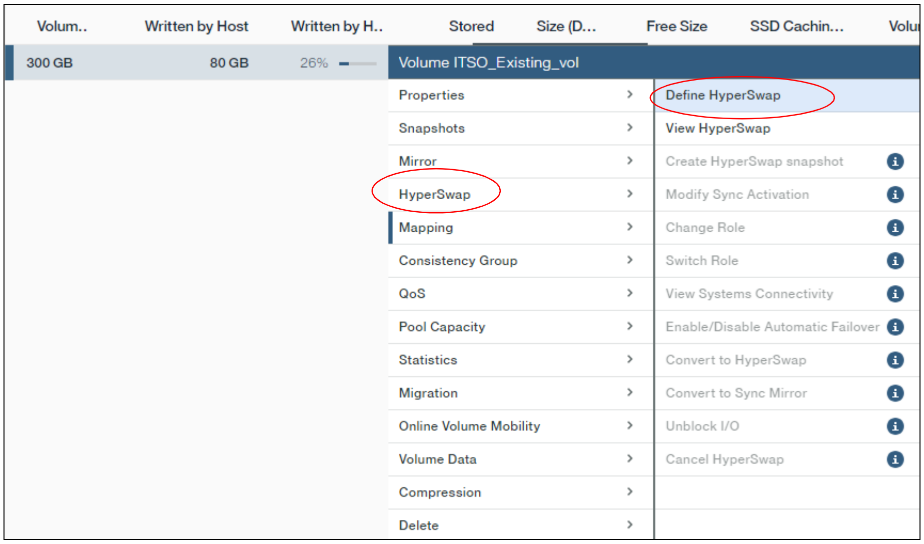
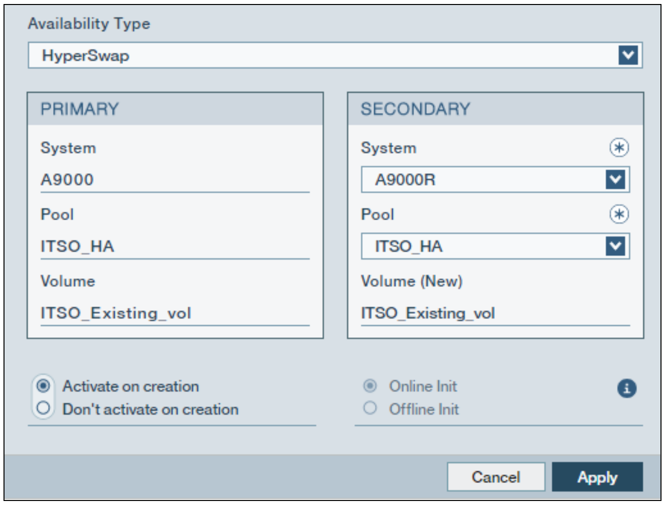


Figure 4-19 Defining HyperSwap

3. In the panel displayed, select the FlashSystem A9000 or A9000R system that owns the secondary volume and the corresponding pool on that system, making sure to select **Activate on creation** and then click **Apply,** as shown in Figure 4-20.



Availability Type

HyperSwap

PRIMARY

System

A9000

Pool

ITSO_HA

Volume

ITSO_Existing_vol

SECONDARY

System

A9000R

Pool

ITSO_HA

Volume (New)

ITSO_Existing_vol

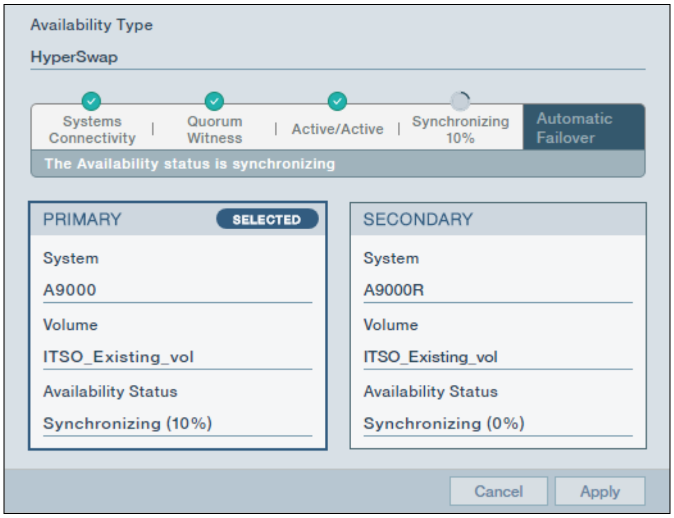
● Activate on creation
○ Don't activate on creation

● Online Init
○ Offline Init

Figure 4-20 Defining the secondary volume

> **Note:** The synchronizing state as depicted in Figure 4-21 will take time to complete depending on the amount of data, and the number and speed of the mirroring connections.

The Availability Status will change to Sychronizing, as shown in Figure 4-21.



Availability Type

HyperSwap

The Availability status is synchronizing

PRIMARY

System
A9000

Volume
ITSO_Existing_vol

Availability Status
Synchronizing (10%)

SECONDARY

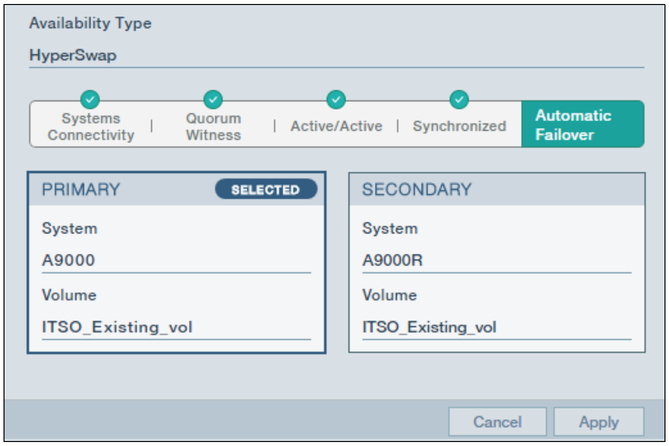
System
A9000R

Volume
ITSO_Existing_vol

Availability Status
Synchronizing (0%)

Figure 4-21 Volume in Synchronizing state

When completed, the volume will show as a HyperSwap volume, as shown in Figure 4-22.



Availability Type

HyperSwap

PRIMARY

System
A9000

Volume
ITSO_Existing_vol

SECONDARY

System
A9000R

Volume
ITSO_Existing_vol

Figure 4-22 Volume synchronization complete

4. Map the secondary volume created on the target system to the same host on which the primary data resides.

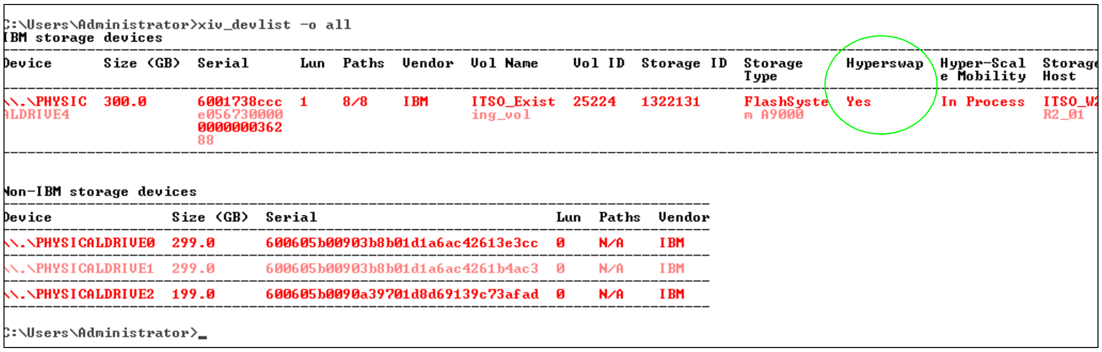
   Steps 1 through 4 can also be completed from the XCLI, as shown in Example 4-9.

Example 4-9 XCLI commands for these steps

```
A9000>>ha_create vol="ITSO_Existing_vol" remote_pool="ITSO_HA"
create_slave="yes" target="A9000R"

A9000>>ha_activate vol="ITSO_Existing_vol" target="A9000R"

A9000R>>map_vol vol="ITSO_Existing_vol" lun="1" host="ITSO_W2K12R2_01"
```

5. Use the **`xiv_fc_admin -R`** or **`xiv_iscsi-admin -R`** command, to perform a Fibre Channel or iSCSI rescan on the host to discover the additional new paths.

   Notice the HyperSwap column indicating the system identified the volume as a HyperSwap volume, as shown in Figure 4-23.



```
C:\Users\Administrator>xiv_devlist -o all
IBM storage devices
-----------------------------------------------------------------------------------------------------------------------------
Device       Size (GB)  Serial       Lun  Paths  Vendor  Vol Name    Vol ID  Storage ID  Storage       Hyperswap  Hyper-Scal    Storage
                                                                                         Type                     e Mobility    Host
-----------------------------------------------------------------------------------------------------------------------------
\\.\PHYSIC   300.0      6001738ccc   1    8/8    IBM     ITSO_Exist  25224   1322131     FlashSyste    Yes        In Process    ITSO_W2
ALDRIVE4                e056730000                       ing_vol                         m A9000                                R2_01
                        0000000362
                        88
-----------------------------------------------------------------------------------------------------------------------------


Non-IBM storage devices
----------------------------------------------------------------------------------
Device               Size (GB)  Serial                                Lun  Paths  Vendor
----------------------------------------------------------------------------------
\\.\PHYSICALDRIVE0   299.0      600605b00903b8b01d1a6ac42613e3cc      0    N/A    IBM
----------------------------------------------------------------------------------
\\.\PHYSICALDRIVE1   299.0      600605b00903b8b01d1a6ac4261b4ac3      0    N/A    IBM
----------------------------------------------------------------------------------
\\.\PHYSICALDRIVE2   199.0      600605b0090a39701d8d69139c73afad      0    N/A    IBM
----------------------------------------------------------------------------------

C:\Users\Administrator>_
```

Figure 4-23   Hypserswap indicator is set to Yes for the volume

6. Identify the connected paths and which are optimized and which are unoptimized using the **`mpclaim -s -d`** command.

   > **Note:** Example 4-10 on page 41 assumes a Windows host. Use the MPIO Disk number (xxx) in the **`mpclaim -s -d xxx`** command.

   Example 4-10   Determining optimized and unoptimized paths

```
C:\Users\Administrator>mpclaim -s -d
For more information about a particular disk, use 'mpclaim -s -d #' where # is
the MPIO disk number.

MPIO Disk     System Disk  LB Policy     DSM Name
-------------------------------------------------------------------------------
MPIO Disk3    Disk 4       RRWS          Microsoft DSM

C:\Users\Administrator>mpclaim -s -d 3

MPIO Disk3: 08 Paths, Round Robin with Subset, Implicit Only
Controlling DSM: Microsoft DSM
SN: 6001738CCCE05673000000000036288
Supported Load Balance Policies: FOO RRWS LQD WP LB
Path ID          State              SCSI Address      Weight
---------------------------------------------------------------
0000000077070003 Active/Unoptimized 007|000|003|001   0
TPG_State : Active/Unoptimized, TPG_Id: 1, : 4097

0000000077060003 Active/Unoptimized 006|000|003|001   0
TPG_State : Active/Unoptimized, TPG_Id: 1, : 4097

0000000077070002 Active/Optimized   007|000|002|001   0
* TPG_State : Active/Optimized  , TPG_Id: 0, : 768

0000000077070001 Active/Optimized   007|000|001|001   0
* TPG_State : Active/Optimized  , TPG_Id: 0, : 512
```

```
0000000077070000 Active/Optimized   007|000|000|001   0
* TPG_State : Active/Optimized  , TPG_Id: 0, : 256

0000000077060002 Active/Optimized   006|000|002|001   0
* TPG_State : Active/Optimized  , TPG_Id: 0, : 770

0000000077060001 Active/Optimized   006|000|001|001   0
* TPG_State : Active/Optimized  , TPG_Id: 0, : 514

0000000077060000 Active/Optimized   006|000|000|001   0
* TPG_State : Active/Optimized  , TPG_Id: 0, : 258

C:\Users\Administrator>
```

4.4.3 Converting a synchronously mirrored volume to HyperSwap volume

To convert a mirrored volume into a HyperSwap volume, the secondary volume of the mirror relationship has to be unmapped. After you unmap the secondary volume in the relationship, use the following steps to convert the mirrored volume to a HyperSwap volume:

> **Note:** The following example illustrations were completed using a volume (`ITSO_MIRROR_VOLUME`) mapped to a Windows 2012 R2 host with I/O running. *Only synchronously* mirrored volumes can be converted to a HyperSwap relationship.

1. Navigate to the Remote Views view and select **Replication Details**, as shown in Figure 4-24.

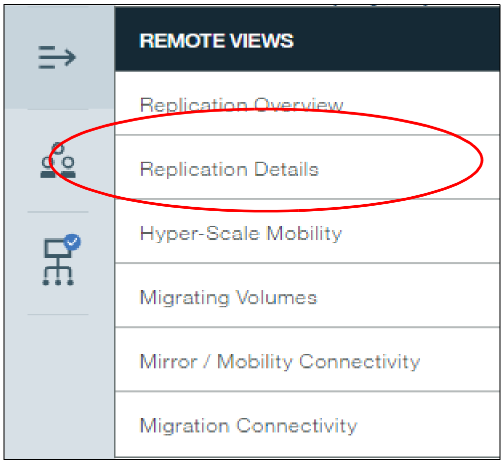
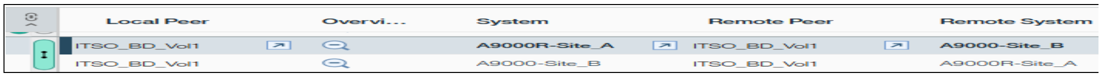


Figure 4-24 Navigating to the Replication Details view

2. Highlight the volume listed as primary in the mirror relationship, as shown in Figure 4-25.



Local Peer	Overvi...	System	Remote Peer	Remote System
ITSO_BD_Vol1		A9000R-Site_A	ITSO_BD_Vol1	A9000-Site_B
ITSO_BD_Vol1		A9000-Site_B	ITSO_BD_Vol1	A9000R-Site_A

Figure 4-25 Selecting the primary volume

3. Right click and select **Convert Type** → **Convert to HyperSwap**, as shown in Figure 4-26.

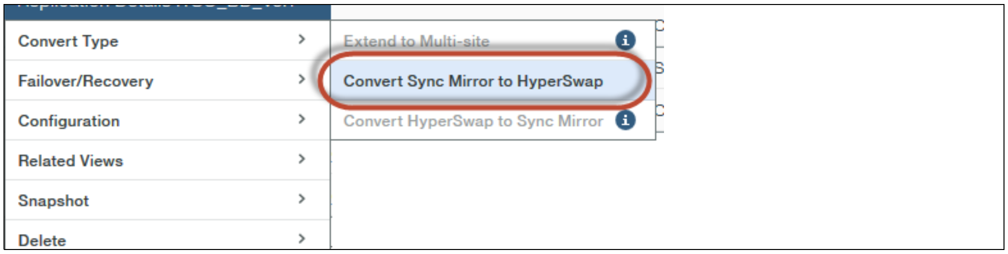
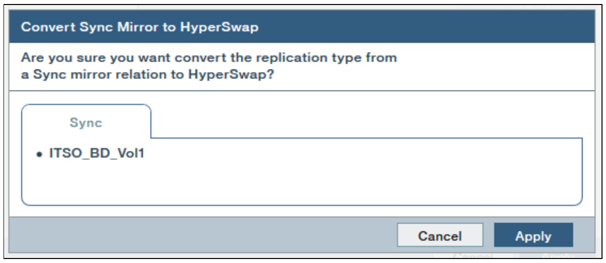


Figure 4-26 Converting to HyperSwap

4. Click **Apply,** as shown in Figure 4-27.


Convert Sync Mirror to HyperSwap

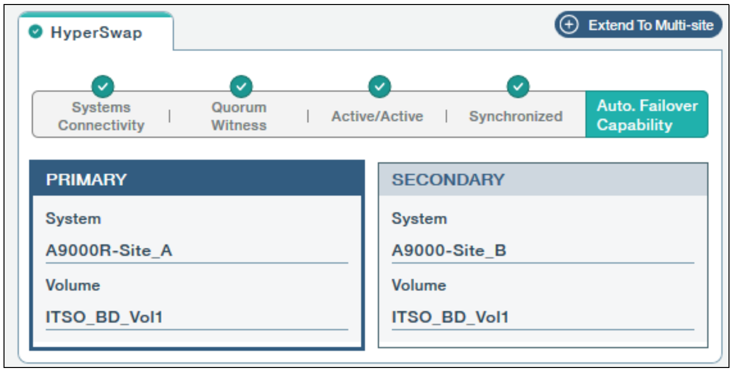
Are you sure you want convert the replication type from a Sync mirror relation to HyperSwap?

Sync

- ITSO_BD_Vol1

Figure 4-27 Applying the conversion

After the process completes successfully, the volumes are in a HyperSwap relationship, as shown in Figure 4-28.



PRIMARY

System

A9000R-Site_A

Volume

ITSO_BD_Vol1

SECONDARY

System

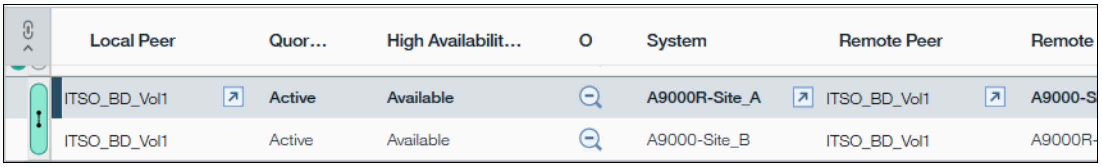
A9000-Site_B

Volume

ITSO_BD_Vol1

Figure 4-28 Replication updated to HyperSwap

Figure 4-29 shows the relationship displayed in the Replication Details view with HyperSwap indicated under the High Availability column.



Local Peer	Quor…	High Availabilit…	O	System	Remote Peer	Remote
ITSO_BD_Vol1	Active	Available		A9000R-Site_A	ITSO_BD_Vol1	A9000-S
ITSO_BD_Vol1	Active	Available		A9000-Site_B	ITSO_BD_Vol1	A9000R-

Figure 4-29 Volume view now shows HyperSwap relationship

For an example of how to convert a mirrored volume into a HyperSwap volume using the XCLI, see Example 4-11.

Example 4-11 XCLI to convert mirrored relation to HyperSwap

```
A9000>mirror_convert_into_ha vol="ITSO_MIRROR_VOLUME"

A9000R>>map_vol vol="ITSO_MIRROR_VOLUME" lun="2" host="ITSO_W2K12R2_01"
```

5. Map the secondary volume to the host and rescan, as shown in Example 4-12.

Example 4-12 Rescanning the host after conversion

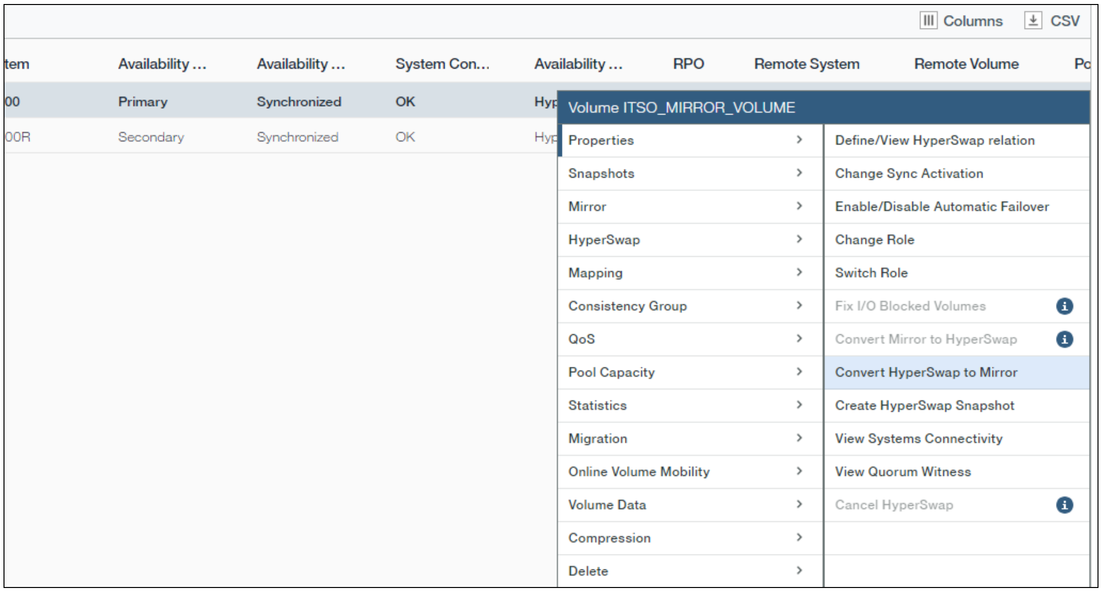
```
C:\Users\Administrator>xiv_fc_admin -R
```

4.4.4 Converting a HyperSwap relationship to a synchronous mirror relationship

A HyperSwap volume relationship can be converted to a synchronous mirror relationship. The following steps can be taken to make the conversion:

> **Note:** The following example illustrations were completed using a volume (ITSO_MIRROR_VOLUME) which was mapped as a HyperSwap volume to a Windows 2012 R2 host with I/O running.
>
> HyperSwap relationships can only be converted to a synchronous mirroring relationship.

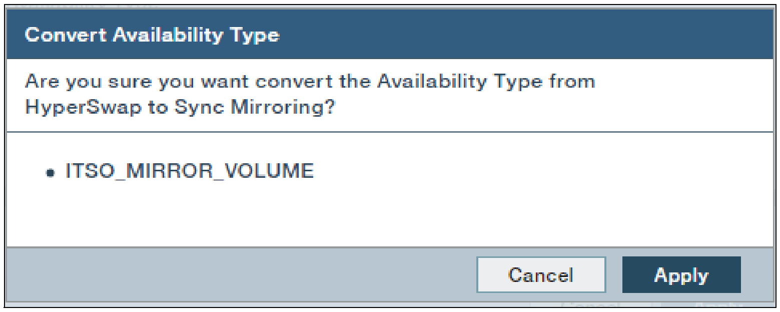
1. Unmap the secondary copy of the volume from the host.
2. In the Volumes view, select the primary copy of the volume, or the copy that is still mapped to the host.
3. Right-click the primary volume and select **HyperSwap** then select **Convert HyperSwap to Mirror**, as shown in Figure 4-30.



tem	Availability ...	Availability ...	System Con...	Availability ...	RPO	Remote System	Remote Volume	Pc
00	Primary	Synchronized	OK	Hyp				
00R	Secondary	Synchronized	OK	Hyp				

Figure 4-30 Selecting Convert HyperSwap to Mirror option

4. Click **Apply,** as shown in Figure 4-31.

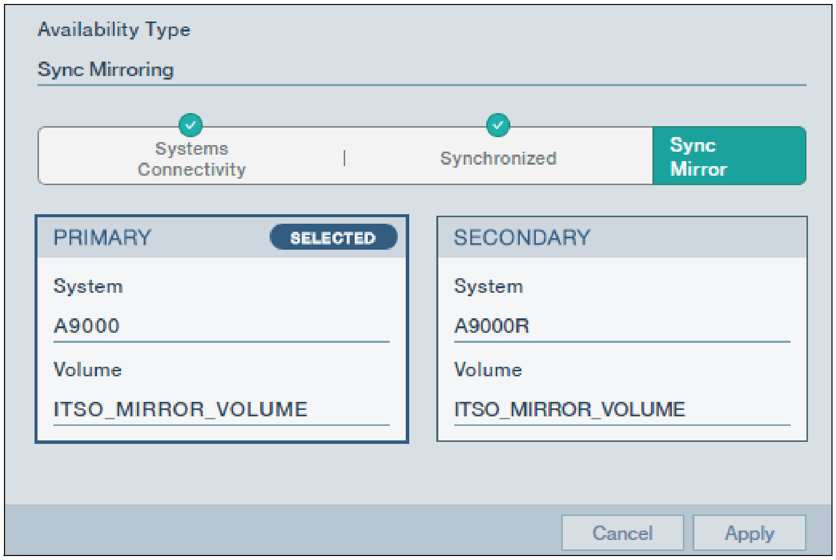

Convert Availability Type

Are you sure you want convert the Availability Type from HyperSwap to Sync Mirroring?

- ITSO_MIRROR_VOLUME

Figure 4-31 Applying the conversion

The Hub now reflects that the volume was successfully converted to a mirrored volume, as shown in Example 4-32.


Availability Type

Sync Mirroring

PRIMARY

System

A9000

Volume

ITSO_MIRROR_VOLUME

SECONDARY

System

A9000R

Volume

ITSO_MIRROR_VOLUME

Figure 4-32 Verification that volume is now a Sync Mirror

To convert a HyperSwap relationship to a synchronous mirror relationship in XCLI, see Example 4-13.

Example 4-13 XCLI commands to convert from a HyperSwap relation to a mirrored relation

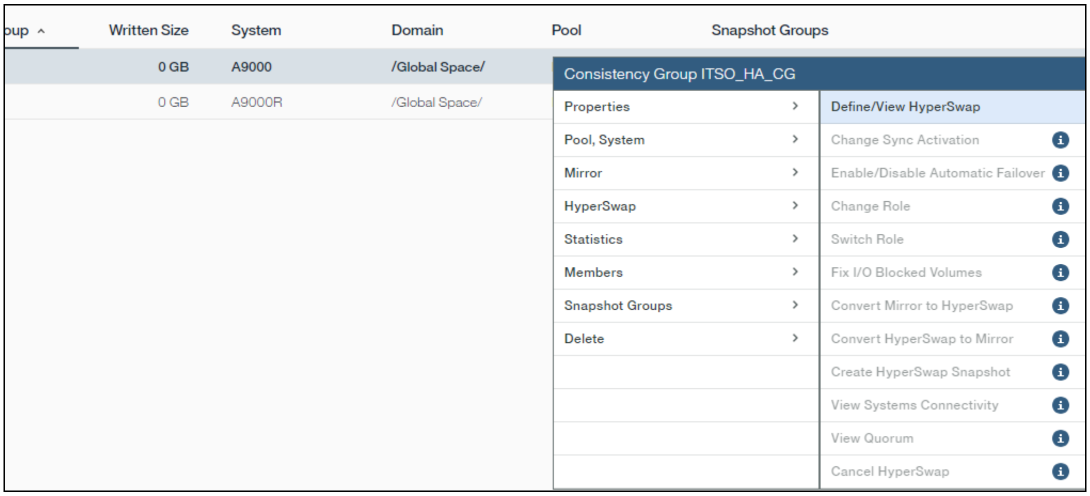
```
A9000R>>unmap_vol vol="ITSO_MIRROR_VOLUME" host="ITSO_W2K12R2_01"

A9000>>ha_convert_into_mirror vol="ITSO_MIRROR_VOLUME"
```

4.4.5 Defining a Consistency Group as a HyperSwap Consistency Group

HyperSwap Consistency Groups (CG) uses one CG on the source system and one CG on the destination (target) system. Both CGs must be empty, meaning they cannot contain any volume, before defining the CG as a HyperSwap CG.

The following steps can be performed to create a HyperSwap CG:

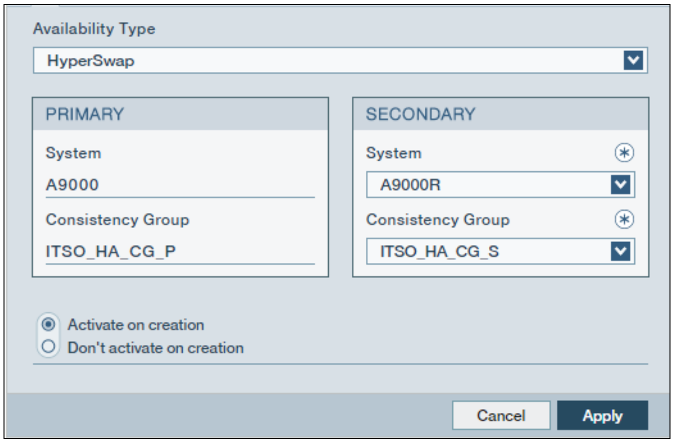
1. Create a CG on both storage systems.
2. Right-click the primary Consistency Group and select **HyperSwap** → **Define/View HyperSwap,** as shown in Figure 4-33.



oup ^	Written Size	System	Domain	Pool	Snapshot Groups
	0 GB	A9000	/Global Space/		
	0 GB	A9000R	/Global Space/		

Figure 4-33 Defining a new CG as a HyperSwap CG

3. Define the secondary Consistency Group on the other system and click **Apply,** as shown in Figure 4-34.



Availability Type
HyperSwap

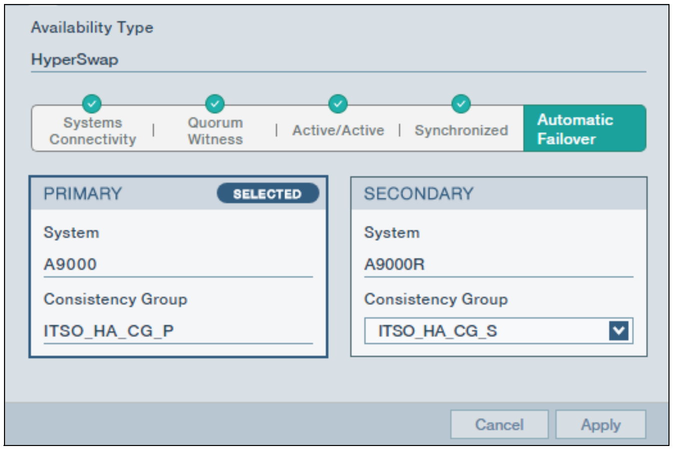
PRIMARY
System
A9000
Consistency Group
ITSO_HA_CG_P

SECONDARY
System
A9000R
Consistency Group
ITSO_HA_CG_S

● Activate on creation
○ Don't activate on creation

Figure 4-34 Defining the secondary system and CG

The Hub shows that the CG is now a HyperSwap CG as illustrated in Figure 4-35.



Availability Type

HyperSwap

PRIMARY

System

A9000

Consistency Group

ITSO_HA_CG_P

SECONDARY

System

A9000R

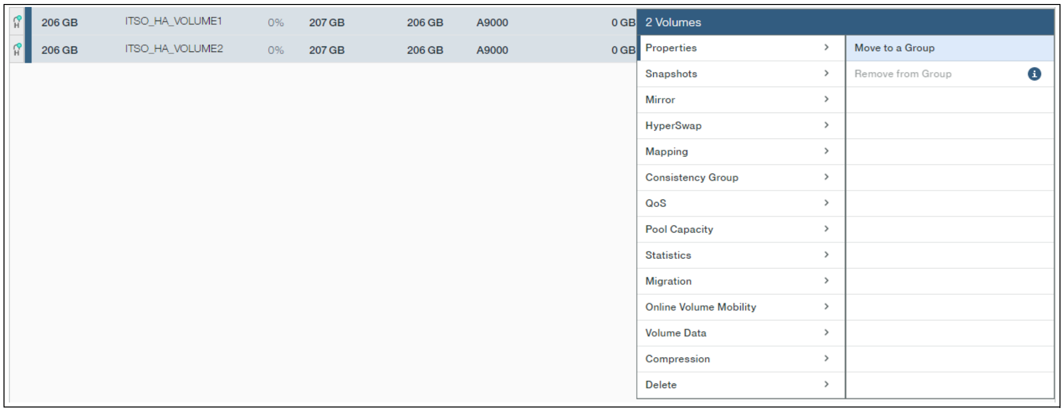
Consistency Group

ITSO_HA_CG_S

Figure 4-35 CG shows as a HyperSwap CG

4. Navigate to the Volumes View, select the volumes, and right-click the list of volumes to be added to the CG. Select **Consistency Group** → **Move to Group,** as shown in Figure 4-36.

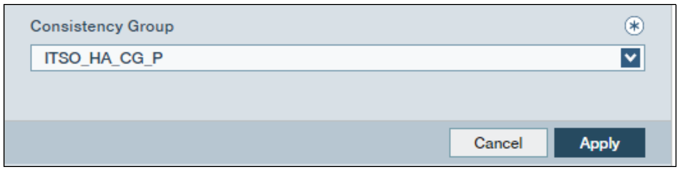
> **Note:** The volume or volumes to be added to the CG must already be a HyperSwap volume or volumes.



```
206 GB   ITSO_HA_VOLUME1   0%   207 GB   206 GB   A9000   0 GB
206 GB   ITSO_HA_VOLUME2   0%   207 GB   206 GB   A9000   0 GB
```

Figure 4-36 Selecting volumes to add to the HyperSwap CG

5. Select the appropriate CG, and click **Apply**, as shown in Figure 4-37.

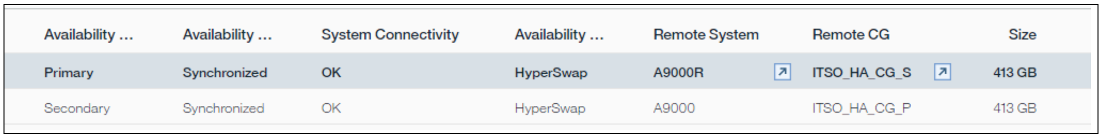


Consistency Group

ITSO_HA_CG_P

Figure 4-37 Selecting the CG to add the volumes to

The CG now shows as Synchronized and displays an Availability status of HyperSwap, as shown in Figure 4-38.



Availability ...	Availability ...	System Connectivity	Availability ...	Remote System	Remote CG	Size
Primary	Synchronized	OK	HyperSwap	A9000R ↗	ITSO_HA_CG_S ↗	413 GB
Secondary	Synchronized	OK	HyperSwap	A9000	ITSO_HA_CG_P	413 GB

Figure 4-38 CG shows Synchronized and HyperSwap

Configuring a HyperSwap CG in XCLI is shown in Example 4-14.

Example 4-14 XCLI commands for creating a HyperSwap Consistency Group

```
A9000>>ha_activate cg="ITSO_HA_CG_P" target="A9000R"
A9000>>ha_create slave_cg="ITSO_HA_CG_S" cg="ITSO_HA_CG_P" target="A9000R"

A9000>>ha_create vol="ITSO_HA_VOLUME1" remote_pool="ITSO_HA_POOL1"
create_slave="yes" target="A9000R"

A9000>>ha_activate vol="ITSO_HA_VOLUME1" target="A9000R"

A9000>>ha_create vol="ITSO_HA_VOLUME2" remote_pool="ITSO_HA_POOL1"
create_slave="yes" target="A9000R"

A9000>>ha_activate vol="ITSO_HA_VOLUME2" target="A9000R"

A9000>>cg_add_vol vol="ITSO_HA_VOLUME2" cg="ITSO_HA_CG_P"

A9000>>cg_add_vol vol="ITSO_HA_VOLUME1" cg="ITSO_HA_CG_P"
```

4.4.6 Converting a synchronous mirror CG to a HyperSwap CG

Only synchronously mirrored CGs can be converted to a HyperSwap CG.Complete the following steps to convert a Sync CG to a HyperSwap CG:

1. Navigate to the Remote Views view and select **Replication Details,** as shown in Figure 4-39.

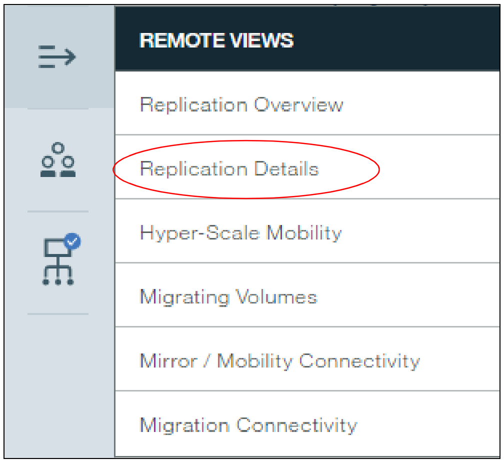
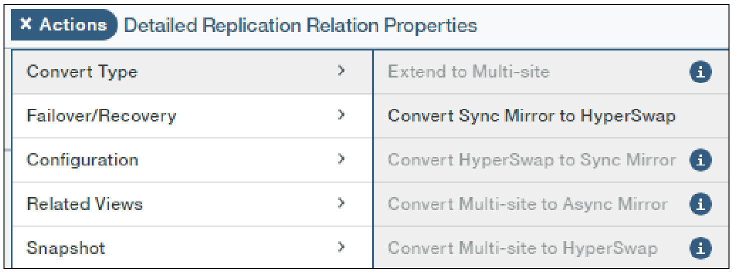


Figure 4-39 Navigating to the Replication Details view

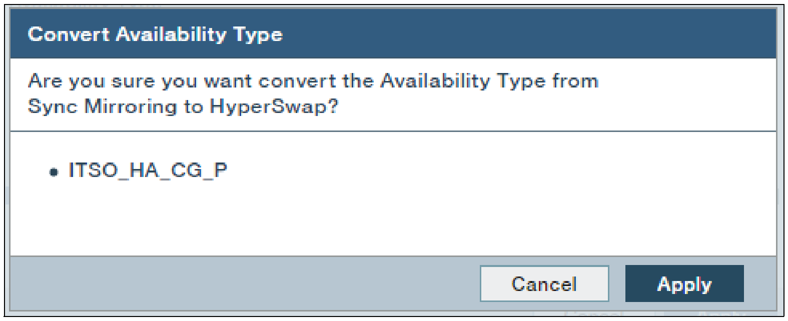
2. In the Replication Details view, select the primary CG. Right-click the primary CG (or under the Actions menu after the CG is selected) and select **Convert Type** → **Convert Sync Mirror to HyperSwap,** as shown in Figure 4-40.



✕ Actions
Detailed Replication Relation Properties
Convert Type
Failover/Recovery
Configuration
Related Views
Snapshot
Extend to Multi-site
Convert Sync Mirror to HyperSwap
Convert HyperSwap to Sync Mirror
Convert Multi-site to Async Mirror
Convert Multi-site to HyperSwap

Figure 4-40 Selecting convert to HyperSwap option

3. Click **Apply,** as shown in Figure 4-41.



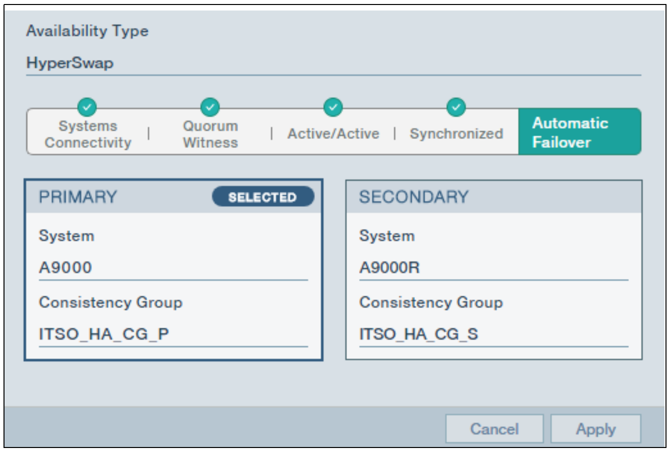
Convert Availability Type

Are you sure you want convert the Availability Type from Sync Mirroring to HyperSwap?

- ITSO_HA_CG_P

Figure 4-41 Applying the conversion of the CG to a HyperSwap CG

The Hub now shows that the CG was successfully converted to a HyperSwap CG, as illustrated in Figure 4-42.



Availability Type

HyperSwap

PRIMARY SELECTED

System

A9000

Consistency Group

ITSO_HA_CG_P

SECONDARY

System

A9000R

Consistency Group

ITSO_HA_CG_S

Figure 4-42 CG successfully converted

To convert a synchronous mirror CG to a HyperSwap CG in XCLI, see Example 4-15.

Example 4-15 XCLI commands to convert a Sync CG to a HyperSwap CG

```
A9000>>mirror_convert_into_ha cg="ITSO_HA_CG_P"
```

4.4.7 Converting a HyperSwap CG to a synchronous mirror CG

HyperSwap CGs can only be converted to a synchronously mirrored CG. The following steps can be user to perform the conversion:

1. Navigate to the Remote Views view and select **Replication Details,** as shown in Figure 4-43.

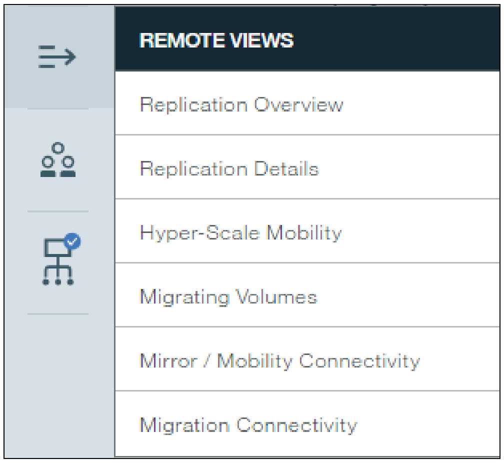


Figure 4-43 Navigating to the Replication Details view

2. In the Consistency Groups view, select the primary CG. Right-click the primary CG and select **HyperSwap** → **Convert HyperSwap to Sync Mirror,** as shown in Figure 4-44.

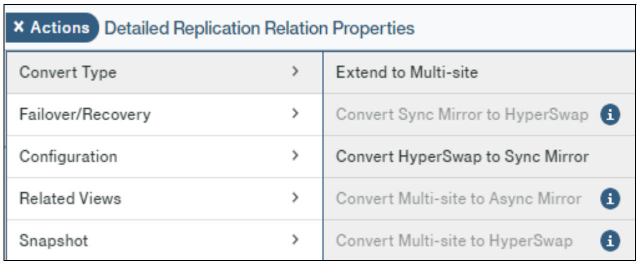
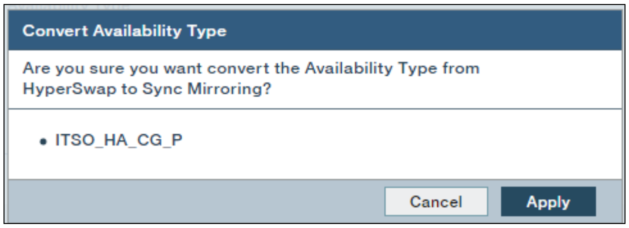

✕ Actions
Detailed Replication Relation Properties
Convert Type
Failover/Recovery
Configuration
Related Views
Snapshot
Extend to Multi-site
Convert Sync Mirror to HyperSwap
Convert HyperSwap to Sync Mirror
Convert Multi-site to Async Mirror
Convert Multi-site to HyperSwap

Figure 4-44 Selecting Convert HyperSwap to Sync Mirror

3. Click **Apply,** as shown in Figure 4-45.

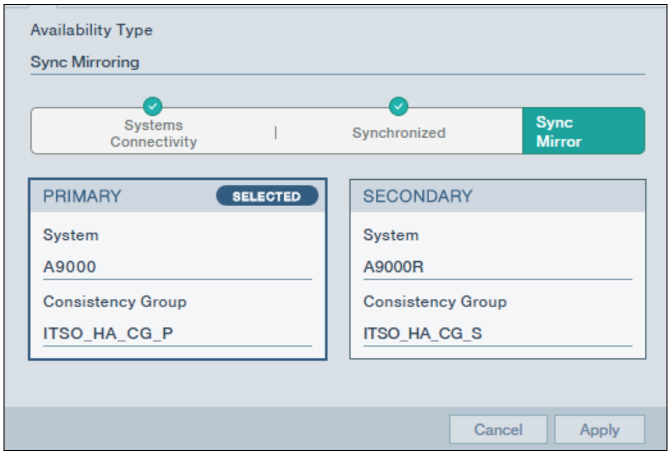

Convert Availability Type

Are you sure you want convert the Availability Type from HyperSwap to Sync Mirroring?

- ITSO_HA_CG_P

Figure 4-45 Applying the CG conversion

The Hub reflects that the HyperSwap CG was successfully converted to a Sync Mirrored CG, as shown in Figure 4-46.



Availability Type
Sync Mirroring

PRIMARY
System
A9000
Consistency Group
ITSO_HA_CG_P

SECONDARY
System
A9000R
Consistency Group
ITSO_HA_CG_S

Figure 4-46 CG successfully converted to Sync Mirror

To convert a HyperSwap CG to a synchronous mirror CG in XCLI, see Example 4-16.

Example 4-16 XCLI commands to convert a HyperSwap CG to a Sync CG

```
A9000>>ha_convert_into_mirror cg="ITSO_HA_CG_P"
```

4.5 Checking HyperSwap status

The current status of the QW, HyperSwap volumes, and CGs can be viewed by using the XCLI and the GUI.

To view the QW status by using the XCLI, use the **quorum_witness_list** command, as shown in Example 4-17.

Example 4-17 Viewing the QW list

```
Primary System:
A9000>>quorum_witness_list

Name       ID                                Address        Port   State      Connection   External Name
Db Health

ITSO_QW1   11bb74a4f4654459a6c510f2e57c8813   9.155.117.24   8460   Activated   Up          myquorum
OK

Secondary System:
A9000R>>quorum_witness_list

Name       ID                                Address        Port   State      Connection   External Name
Db Health

ITSO_QW1   11bb74a4f4654459a6c510f2e57c8813   9.155.117.24   8460   Activated   Up          myquorum
OK
```

To view the HypserSwap volumes and CGs by using the XCLI, use the **ha_list** command, as shown in Example 4-18.

Example 4-18 Viewing HyperSwap volumes and CGs

```
Primary System:
A9000>>ha_list
Name                 HA Object   Role     Remote System   Active   Status        Link Up   Automatic
Failover

ITSO_HA_VOLUME      Volume     Master   A9000R         yes     Synchronized   yes      N/A

ITSO_HA_AIX_DHE_001   Volume     Slave    A9000R         yes     Synchronized   yes      Active

ITSO_HA_AIX_DHE_002   Volume     Master   A9000R         yes     Synchronized   yes      N/A

ITSO_Existing_vol    Volume     Master   A9000R         yes     Synchronized   yes      N/A

ITSO_HA_VMware      Volume     Master   A9000R         yes     Synchronized   yes      N/A

ITSO_HA_CG_P        CG         Master   A9000R         yes     Synchronized   yes      N/A

ITSO_HA_VOLUME1      Volume     Master   A9000R         yes     Synchronized   yes      N/A

ITSO_HA_VOLUME2      Volume     Master   A9000R         yes     Synchronized   yes      N/A

Secondary System:
A9000R>>ha_list
Name                 HA Object   Role     Remote System   Active   Status        Link Up   Automatic
Failover

ITSO_HA_VOLUME      Volume     Slave    A9000          yes     Synchronized   yes      Active

ITSO_HA_AIX_DHE_001   Volume     Master   A9000          yes     Synchronized   yes      N/A

ITSO_HA_AIX_DHE_002   Volume     Slave    A9000          yes     Synchronized   yes      Active

ITSO_Existing_vol    Volume     Slave    A9000          yes     Synchronized   yes      Active

ITSO_HA_VMware      Volume     Slave    A9000          yes     Synchronized   yes      Active

ITSO_HA_CG_S        CG         Slave    A9000          yes     Synchronized   yes      Active

ITSO_HA_VOLUME1      Volume     Slave    A9000          yes     Synchronized   yes      N/A

ITSO_HA_VOLUME2       Volume      Slave    A9000           yes      Synchronized   yes       N/A
```

Using the Hyper-Scale Manager GUI, the volumes and CGs can be viewed by navigating to the Remote Views view and selecting **Replication Details,** as shown in Figure 4-47.

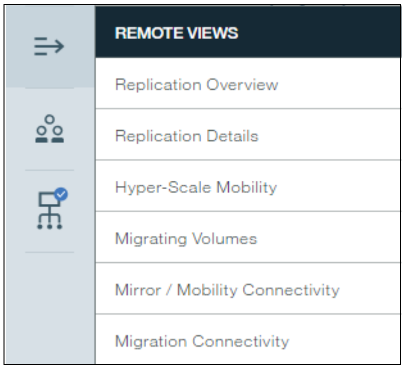


Figure 4-47 Remote Views menu

4.6 HyperSwap snapshots

As is the case with synchronous mirroring, there are three types of snapshots that can be created on a HyperSwap volume or Consistency Group: local, mirrored, and internal snapshot also designated as Last Consistent Snapshot (LCS).

Local snapshots on a HyperSwap volume behave exactly like they do for a regular volume. There are no special limitations; For example it is possible to create a snapshot on the secondary volume, even when that volume is not synchronized.

Mirrored snapshots for volumes in a HyperSwap relationship are created using the **`ha_create_snapshot`** command.This command is similar to **`mirror_create_snapshot`** and has similar behavior and limitations. Both commands create an equivalent snapshot on both peers at the same point-in-time, meaning that the contents of the two snapshots at the time of creation are identical. Those snapshots can be created only when the peers are connected and synchronized.

After you create the snapshots, they are independent and behave like local snapshots. On each peer, the snapshot can be unlocked, updated, deleted, and restored independently. Also snapshot deletion as a result of capacity exhaustion is independent for each peer.

The LCS is a special snapshot created on a consistent secondary volume before a resynchronization is started.

The advantage of creating mirrored snapshots for HyperSwap is that it allows you to restart a temporarily removed HyperSwap relationship, by restoring the HyperSwap volumes from their snapshots, by using as HyperSwap/Mirror with Offline Init option. The synchronization process will only need to transfer data that is different between primary and secondary, rather than initially transferring the whole primary volume contents.

Note: For more information about the general use of snapshots, see *IBM Hyper-Scale Manager for IBM Spectrum Accelerate Family: IBM XIV, IBM FlashSystem A9000 and A9000R, and IBM Spectrum Accelerate*, SG24-8376.

4.6.1 Creating HyperSwap snapshots

Complete the following steps to create HyperSwap snapshots by using the Hyper-Scale Manager GUI:

1. Navigate to the REMOTE VIEWS view and select **Replication Details,** as shown in Figure 4-48.

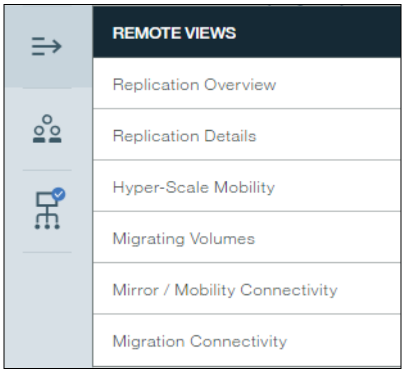


Figure 4-48 Navigation to Replication Details view

2. Right-click the primary HyperSwap volume and select **Snapshot** → **Create Replicated Snapshot,** as shown in Figure 4-49.

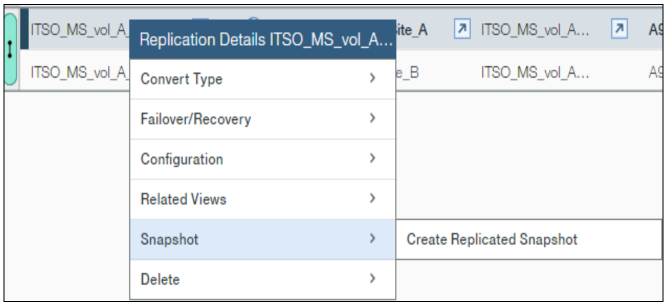
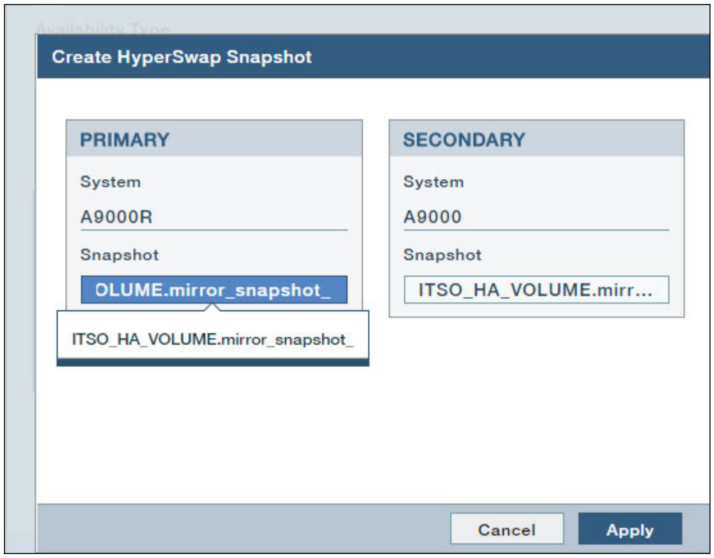


Figure 4-49 Create HyperSwap snapshot for HyperSwap volumes

3. Optionally, change the names of snapshots by pressing the snapshot name and edit, as shown in Figure 4-50, and then click **Apply**.

> **Note:** The default names of both snapshots are `<name_of_volume>.mirror_snapshot_`.



Create HyperSwap Snapshot

PRIMARY

System

A9000R

Snapshot

OLUME.mirror_snapshot_

ITSO_HA_VOLUME.mirror_snapshot_

SECONDARY

System

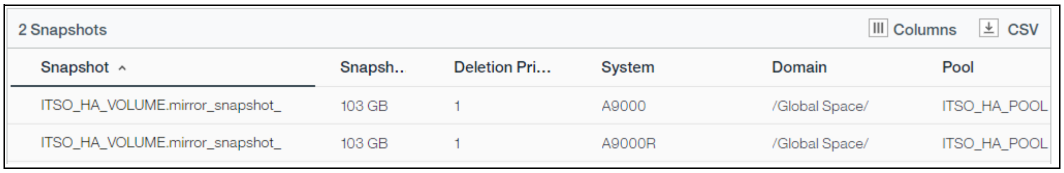
A9000

Snapshot

ITSO_HA_VOLUME.mirr...

Figure 4-50 Create HyperSwap snapshot box view

4. After the process completes successfully, the snapshots are shown in **POOLS & VOLUMES VIEWS** → **Snapshots** View, as shown in Figure 4-51.



2 Snapshots

Snapshot	Snapsh..	Deletion Pri...	System	Domain	Pool
ITSO_HA_VOLUME.mirror_snapshot_	103 GB	1	A9000	/Global Space/	ITSO_HA_POOL
ITSO_HA_VOLUME.mirror_snapshot_	103 GB	1	A9000R	/Global Space/	ITSO_HA_POOL

Figure 4-51 Snapshots view

When using XCLI, open an XCLI session for the system that contains the wanted HyperSwap primary volume and run the following command:

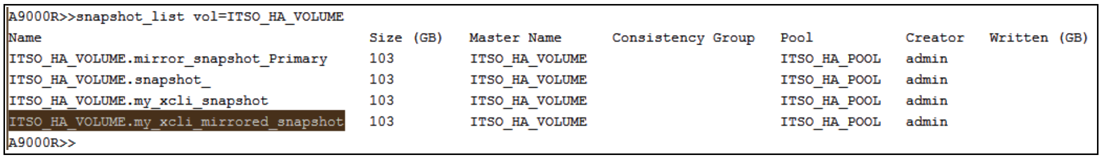
```
ha_create_snapshot vol=<volume_to_snapshot> delete_priority="<1 for Last deleted
until 4 for First>" name="<name_of_snapshot_for_primary>"
```

Example 4-19 shows creating HyperSwap snapshot in XCLI mode.

Example 4-19 Create HyperSwap snapshot in XCLI mode

```
A9000R>>ha_create_snapshot vol=ITSO_HA_VOLUME delete_priority="4"
name="ITSO_HA_VOLUME.my_xcli_mirrored_snapshot"
Command executed successfully.
A9000R>>
```

The snapshot can be shown using **`snapshot_list vol=<name_of_vol>`** as illustrated in Figure 4-52.



```
A9000R>>snapshot_list vol=ITSO_HA_VOLUME
Name                                    Size (GB)   Master Name      Consistency Group   Pool           Creator   Written (GB)
ITSO_HA_VOLUME.mirror_snapshot_Primary  103         ITSO_HA_VOLUME                       ITSO_HA_POOL   admin
ITSO_HA_VOLUME.snapshot_                103         ITSO_HA_VOLUME                       ITSO_HA_POOL   admin
ITSO_HA_VOLUME.my_xcli_snapshot         103         ITSO_HA_VOLUME                       ITSO_HA_POOL   admin
ITSO_HA_VOLUME.my_xcli_mirrored_snapshot  103       ITSO_HA_VOLUME                       ITSO_HA_POOL   admin
A9000R>>
```

Figure 4-52 Snapshot list in XCLI mode

> **Note:** For HyperSwap Consistency Group snapshots, perform the same operations for the desired Consistency Group, rather than for a volume.

Refer also to Chapter 10, “Volume Shadow Copy Service and HyperSwap Snapshots” on page 201 for a discussion on combining Microsoft Virtual Shadow coy Services (VSS) with Snapshots.

5

HyperSwap scenarios

This chapter describes different failure scenarios for IBM FlashSystem A9000 and A9000R with volumes or Consistency Groups in a HyperSwap relationship. The following topics are covered:

- 5.1, “Failures on the storage system serving the primary copy” on page 58
- 5.2, “Connection failure between host and storage systems” on page 61
- 5.3, “Quorum Witness failure scenarios” on page 69
- 5.4, “Failback scenarios” on page 86

The material in this chapter is based on an environment that uses one FlashSystem A9000 and one FlashSystem A9000R. The scenarios assume that the primary copy of a HyperSwap volume is on System A at Site A, as shown in Figure 5-1 on page 58.

5.1 Failures on the storage system serving the primary copy

This section discusses the following failure scenarios on the storage system serving I/Os for the primary volume:

- System A is down
- Data service failure on System A, such as no storage space left for writing on System A

Figure 5-1 shows System A (A9000) and System B (A9000R) with one volume in HyperSwap relationship. The host and the primary copy of the volume are at Site A.

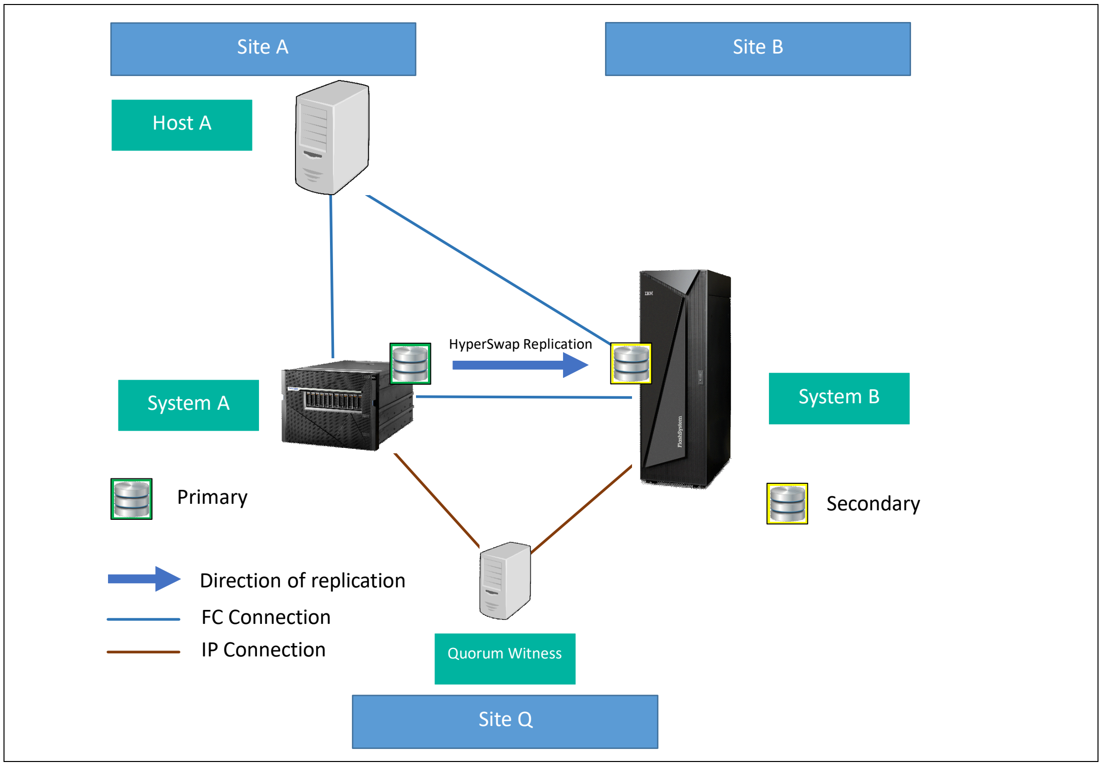


Figure 5-1   IBM FlashSystem HyperSwap Volume

5.1.1 System A failure scenario

Figure 5-2 shows a failure of System A at Site A, which can result from a loss of power affecting System A:

- Host A at Site A and System B (A9000R) at site B have lost their Fibre Channel (FC) connections to System A at Site A.
- The Quorum Witness at site Q lost its Ethernet connection to System A at Site A.

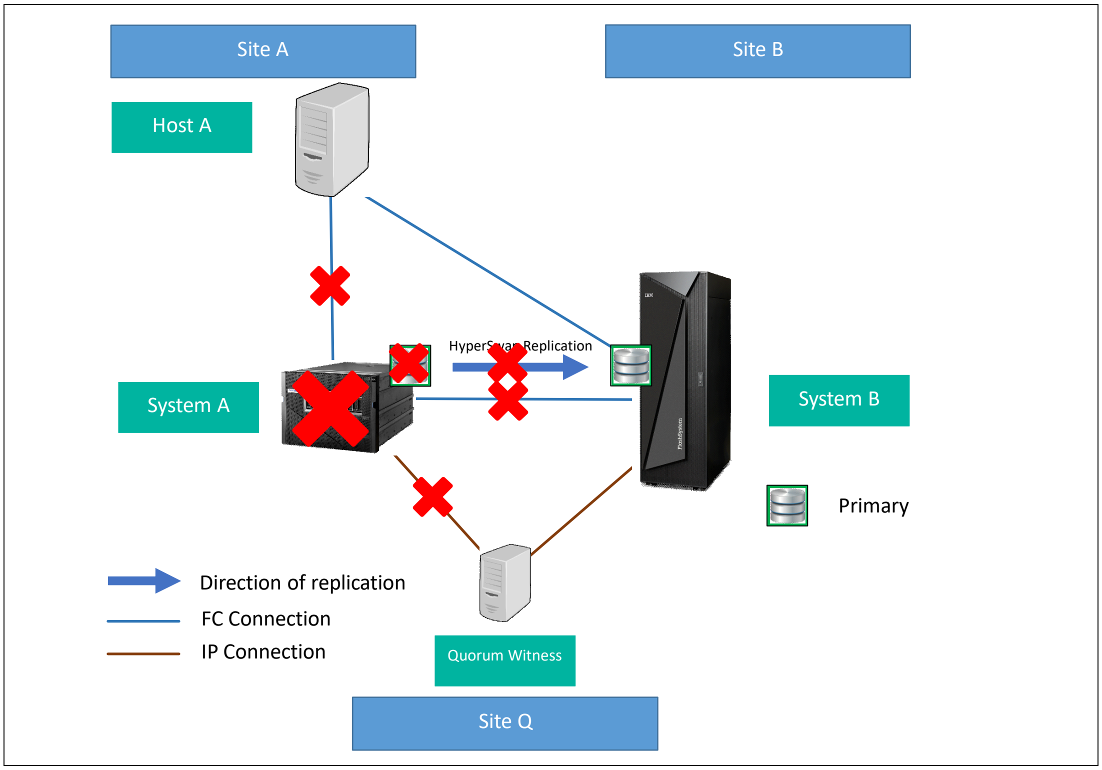
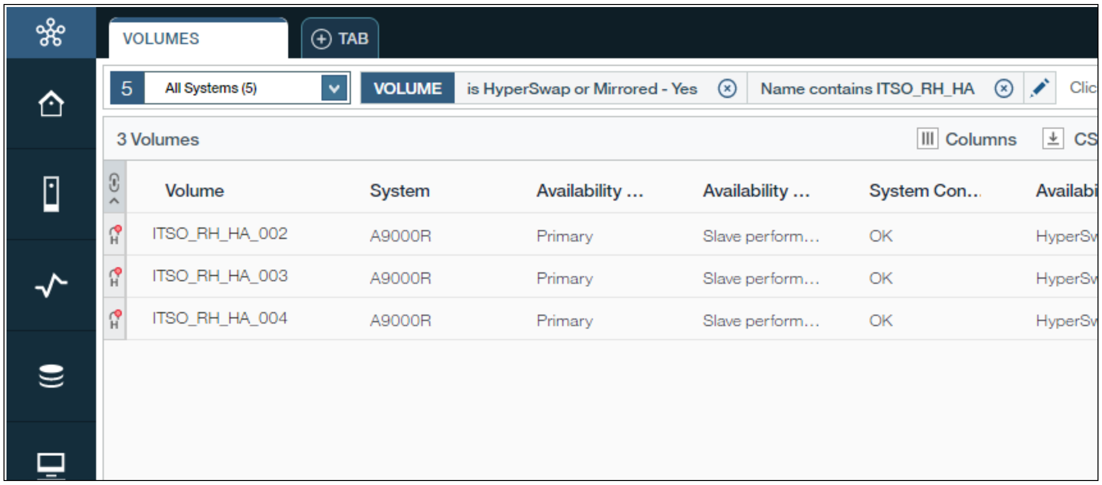


Figure 5-2 System A failure scenario

Host A has still access to the HyperSwap volume because all read and write I/Os can still be served from System B (A9000R). System B has no FC connection to System A and the Quorum Witness has no Ethernet connection to System A.

As a result of HyperSwap automatic failover, System B now serves read and write I/O to host A and volumes on System B take on the primary role, as can be seen from the Hyper-Scale Manager GUI that is shown in Figure 5-3.



3 Volumes

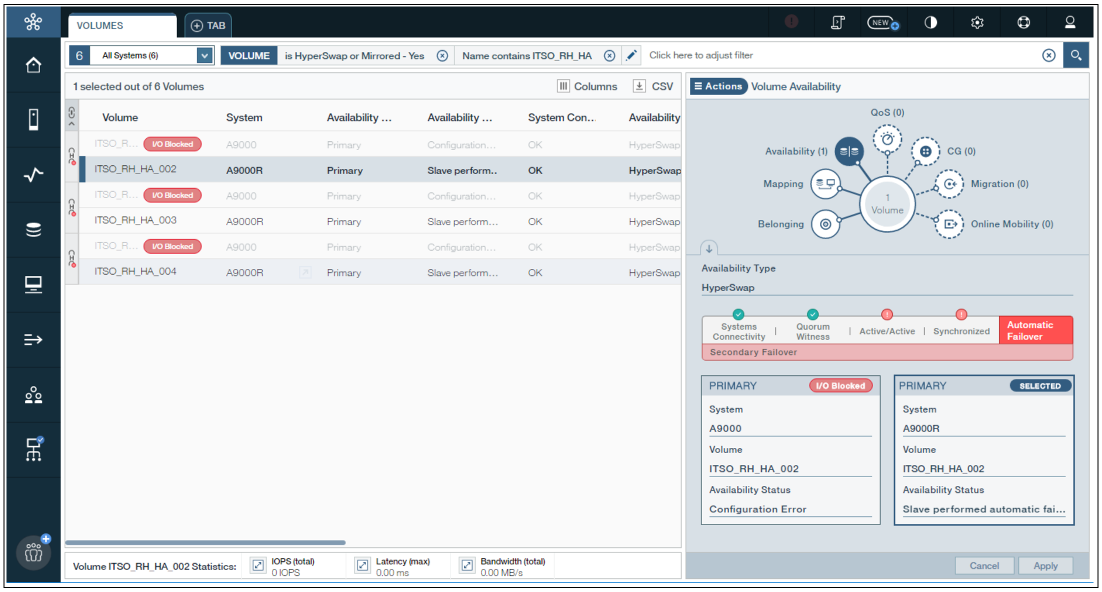
Volume	System	Availability ...	Availability ...	System Con...	Availabi
ITSO_RH_HA_002	A9000R	Primary	Slave perform...	OK	HyperSv
ITSO_RH_HA_003	A9000R	Primary	Slave perform...	OK	HyperSv
ITSO_RH_HA_004	A9000R	Primary	Slave perform...	OK	HyperSv

Figure 5-3 Volumes serving I/O after System A failure

After System A becomes operational again, all paths and connections are reestablished.

At that point, the HyperSwap volume is not synchronized. The HyperSwap volume ITSO_RH_HA_004 has its primary copy on System B (A9000R). I/O is served to the host from System B. The copy of this HyperSwap volume on System A (A9000) is in *I/O blocked* state because it is not synchronized with the copy of this HyperSwap volume on System B.

On System A, the copy of this HyperSwap volume is still in primary role because at the time System B started serving I/O, System A was down and did not get the information about the new role, as shown in Figure 5-4.



1 selected out of 6 Volumes

Volume	System	Availability ...	Availability ...	System Con...	Availability
ITSO_R... I/O Blocked	A9000	Primary	Configuration...	OK	HyperSwap
ITSO_RH_HA_002	A9000R	Primary	Slave perform..	OK	HyperSwap
ITSO_R... I/O Blocked	A9000	Primary	Configuration...	OK	HyperSwap
ITSO_RH_HA_003	A9000R	Primary	Slave perform...	OK	HyperSwap
ITSO_R... I/O Blocked	A9000	Primary	Configuration...	OK	HyperSwap
ITSO_RH_HA_004	A9000R	Primary	Slave perform...	OK	HyperSwap

Availability Type
HyperSwap

PRIMARY I/O Blocked
System
A9000
Volume
ITSO_RH_HA_002
Availability Status
Configuration Error

PRIMARY SELECTED
System
A9000R
Volume
ITSO_RH_HA_002
Availability Status
Slave performed automatic fai...

Volume ITSO_RH_HA_002 Statistics: IOPS (total) 0 IOPS Latency (max) 0.00 ms Bandwidth (total) 0.00 MB/s

Figure 5-4 HyperSwap volume state after System A is operational again

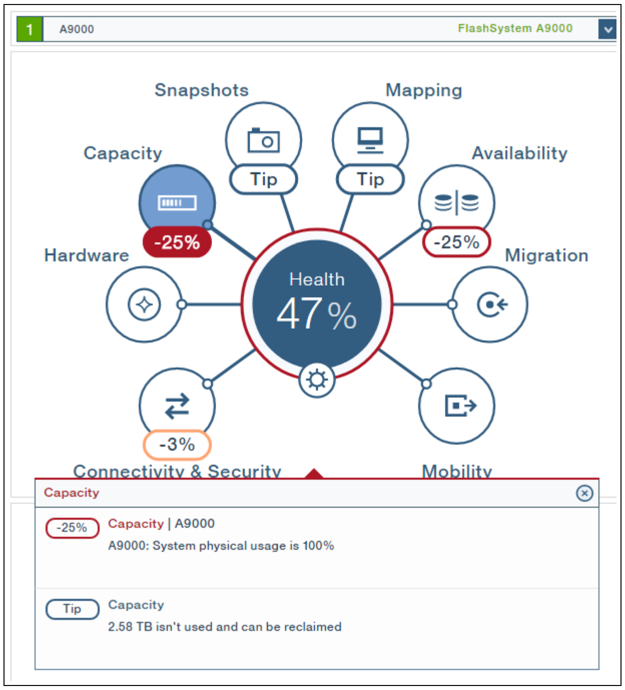
Section 5.4, “Failback scenarios” on page 86 describes the process to recover the I/O blocked volumes on System A and to get back to the **Automatic Failover** state.

5.1.2 Data service failure

This section describes the data services failure scenario. This type of failure can occur when no space is left on System A for writing a physical block. This situation is known as an out of partitions (OOP) situation.

In an OOP situation, the HyperSwap relationship will fail over to Site B and I/O will be served from System B at Site B as described in 5.1.1, “System A failure scenario” on page 59.

Figure 5-5 shows the dashboard information of System A in the OOP state.



Capacity

-25% Capacity | A9000
A9000: System physical usage is 100%

Tip Capacity
2.58 TB isn't used and can be reclaimed

Figure 5-5 Dashboard OOP information

5.2 Connection failure between host and storage systems

This section illustrates connection failure scenarios, either a Fibre Channel (FC) connection loss from System A or loss of the Ethernet connection to the Quorum Witness (QW). Figure 5-1 shows two IBM FlashSystem A9000 and A9000R with a volume in a HyperSwap relation. The hosts and the primary copy of the volume are at site A. The primary copy is serving I/O to host A.

The different scenarios are:

- Failure of Fibre Channel path between the host and the storage system
- Loss of the FC connection between System A and the host, then loss of the FC connection between System A and System B

5.2.1 FC path failure between host and storage system

This scenario considers a FC connectivity loss between the Host A and System A at site A as illustrated in Figure 5-6.

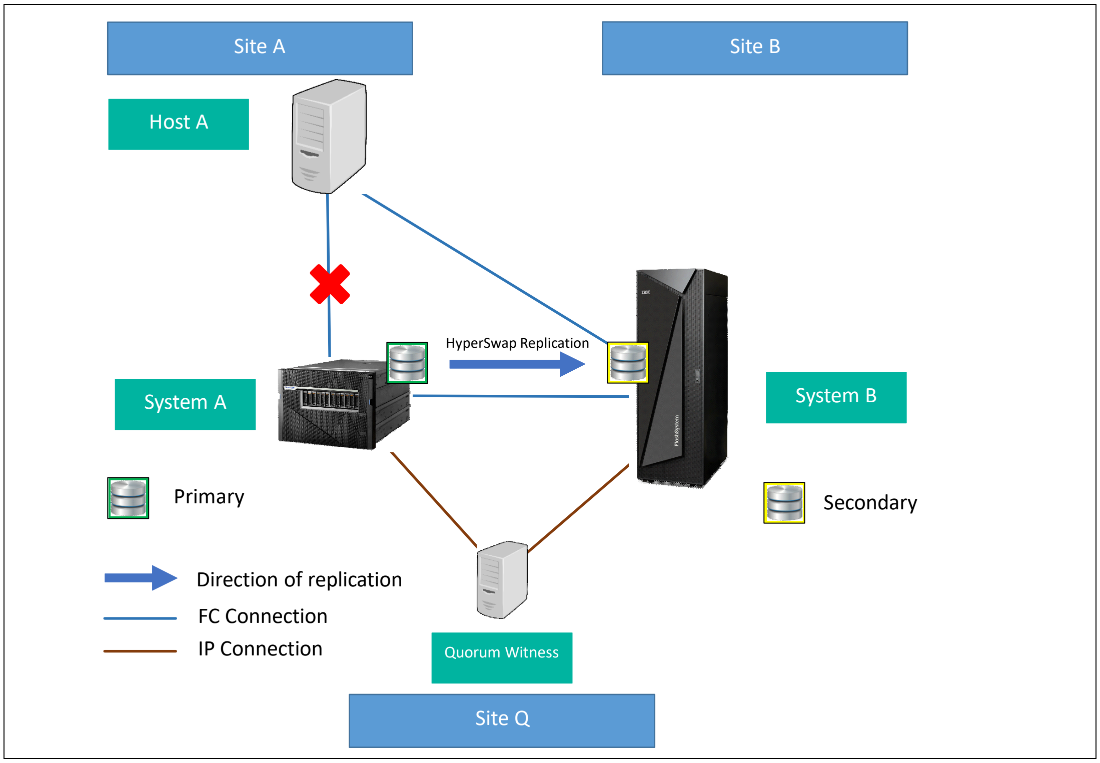


Figure 5-6 FC connection failure between Host A and System A

Figure 5-7, shows the Hyper-Scale Manager GUI representation of the host connection with both System A (A9000) and System B (A9000R). To get to this view, select **Hosts & Clusters Views** → **Hosts**, then selected the corresponding host and go to **Properties** → **Connectivity** → **View Connectivity Details**.

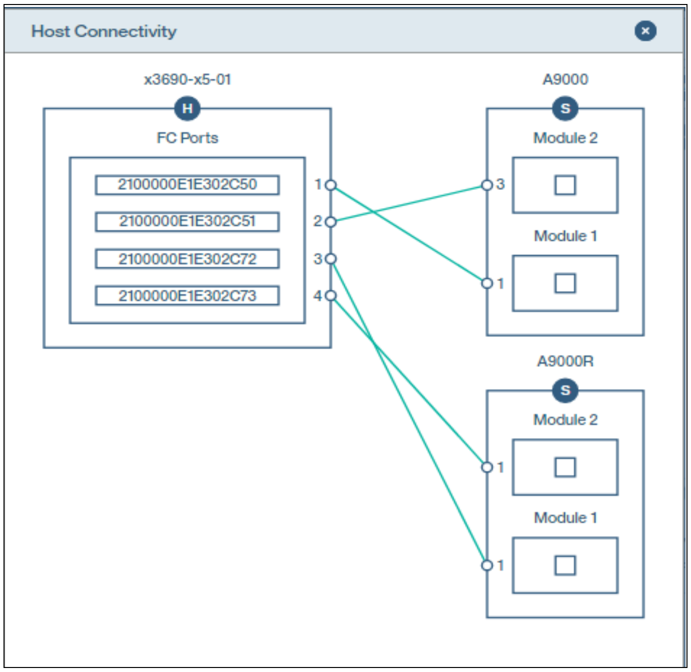


Figure 5-7 Host Connectivity when connected to both storage systems.

Figure 5-8, shows the Hyper-Scale Manager GUI view when connectivity between Host A and System A (A9000) is broken.

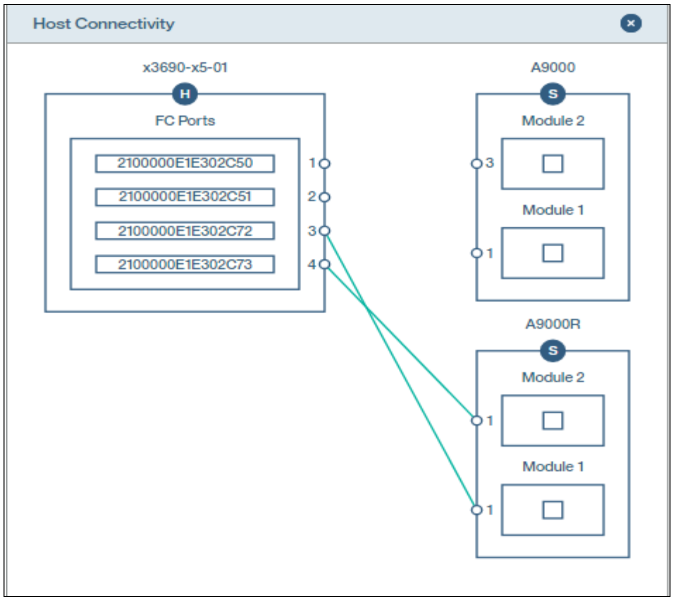


Figure 5-8 FC connection loss between Host and A9000

When FC connectivity between Host A and System A (owning the primary volume) is broken, Host A still has access to the HyperSwap volume because read and write I/Os can still be served from System B, as shown in Figure 5-9.

The data synchronization between System B and System A, before a destage of data to storage, occurs in the following sequence:

1. Host A sends write I/O to System B.
2. System B sends writes to System A.
3. System A caches the data and syncs the data with System B.
4. System B sends an acknowledgement of the writes to System A.
5. System A sends an acknowledgement to System B.
6. System B sends an acknowledgement to the host.

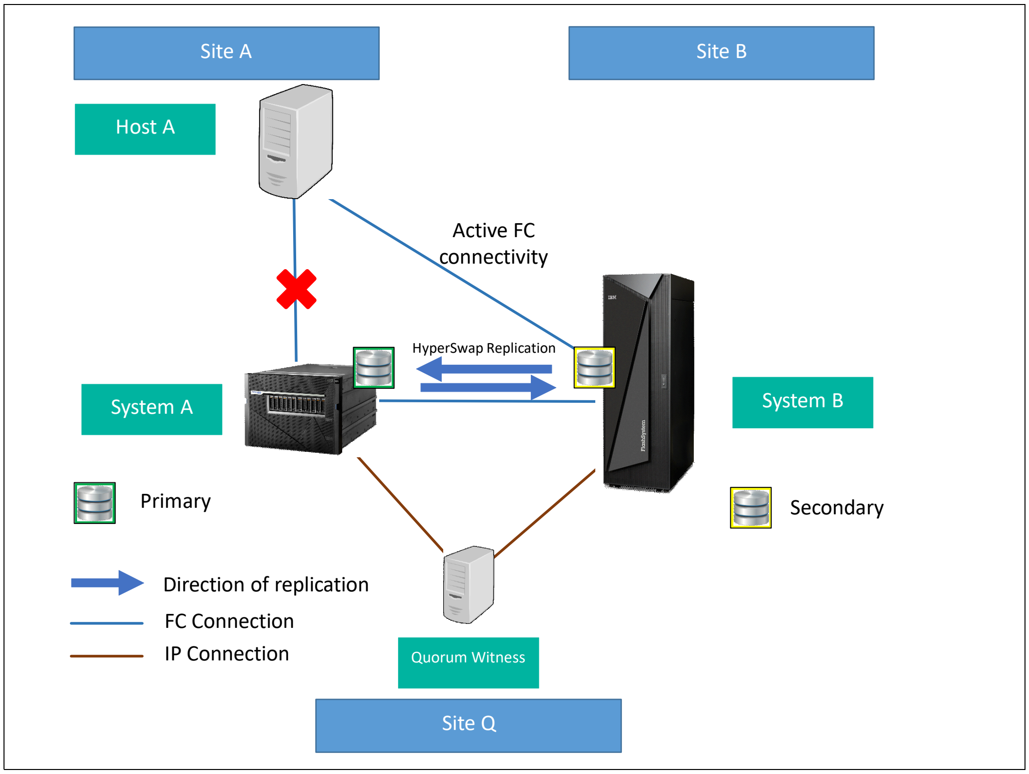


Figure 5-9   Host now reads or writes data to System B which syncs data with System A

When host connectivity with System A comes back, host I/Os are automatically served from System A without any additional process required to synchronize data between System A and System B.

5.2.2 FC connectivity loss between host and System A and between System A and System B

Figure 5-10 depicts a connectivity failure between Host A and System A (A9000), along with a connectivity failure between System A (A9000) and System B (A9000R), which also results in replication failure between System A and System B.

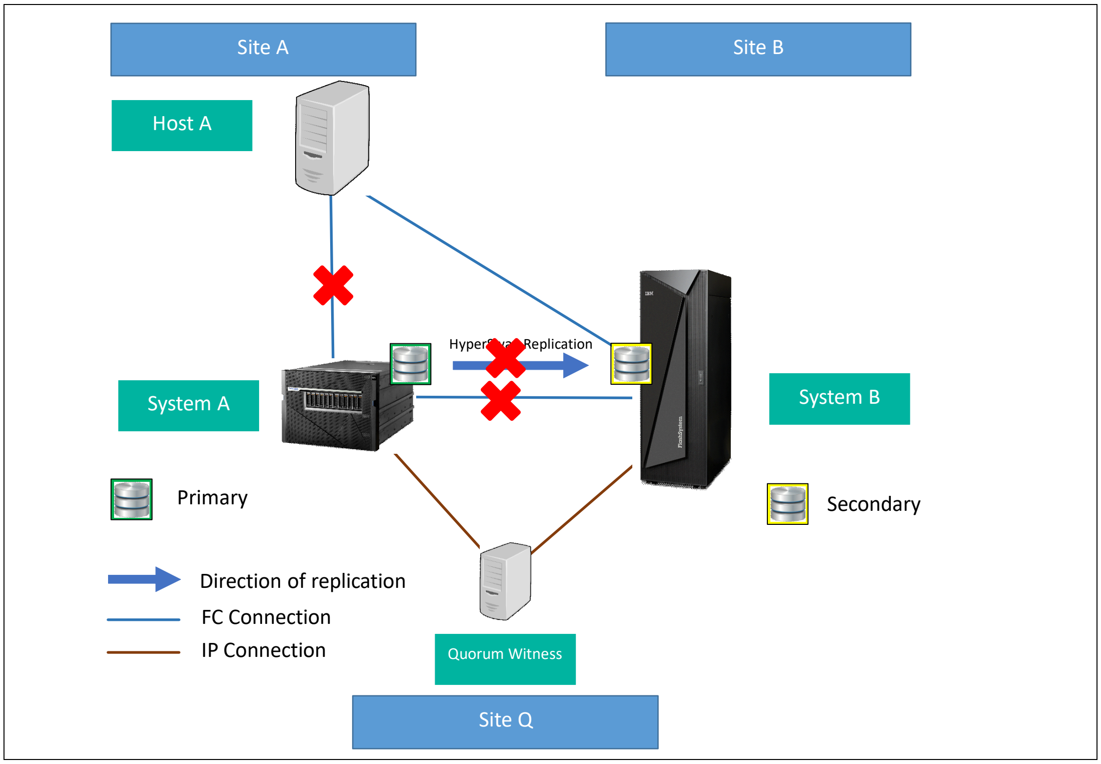


Figure 5-10 Connectivity failure between Host A and System A and between System A and System B

Figure 5-11 shows the Hyper-Scale Manager GUI representation of active FC connectivity between System A and System B. To get to this view, select **Systems & Domains Views,** then select the storage system and go to **System Connectivity** → **Targets** → **View Connectivity** → **Connectivity Details**.

The connectivity pane is displayed, as shown on the left side in Figure 5-11. The view on the right side of Figure 5-11 shows the GUI representation of inactive connectivity between System A and System B when all the links are down.

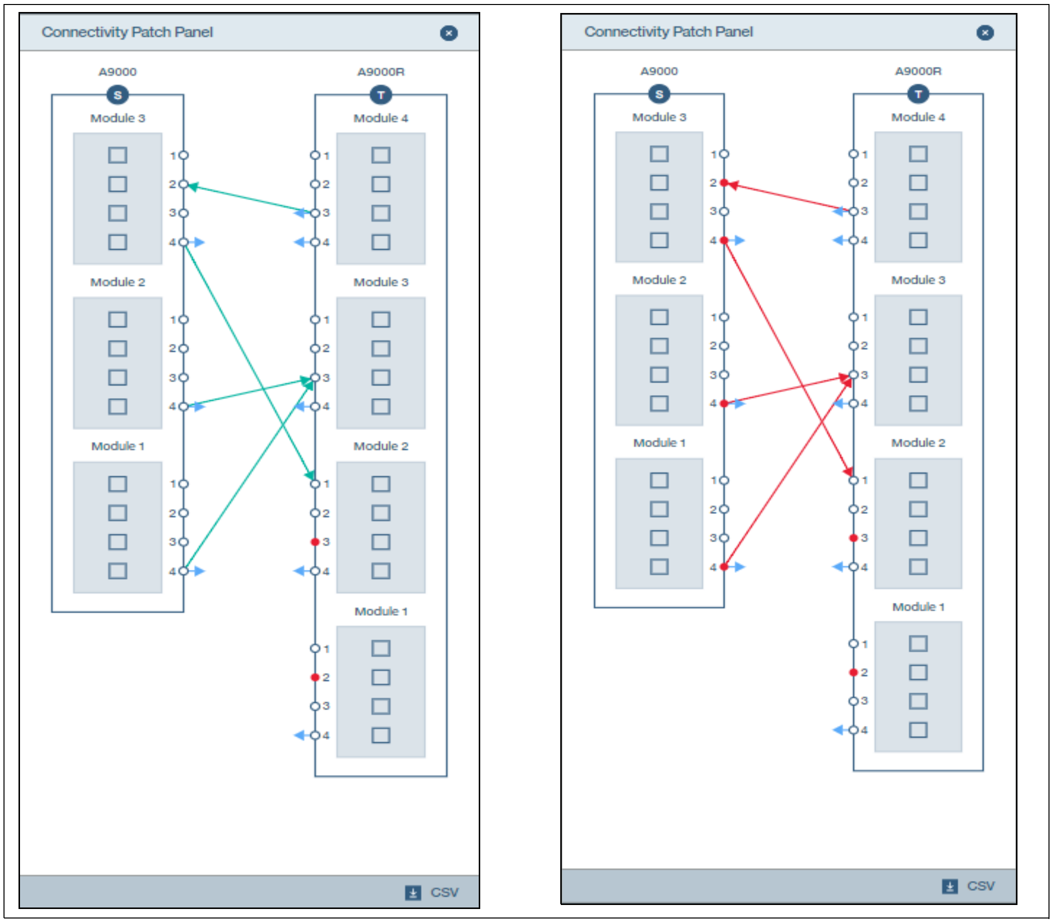
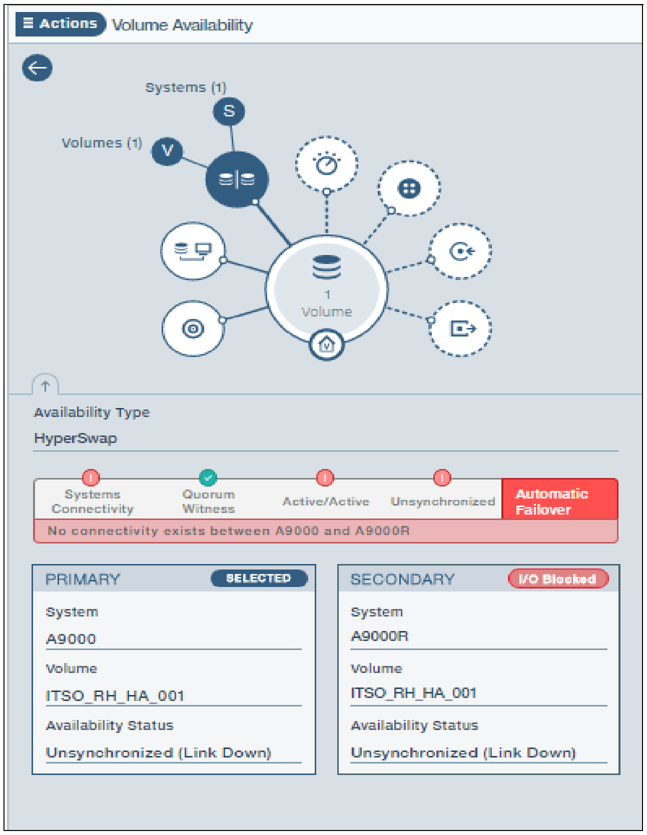


Figure 5-11 Active and Inactive connectivity between System A and System B

The following sequence of events occurs when connectivity between Host A and System A and connectivity between System A and System B fails:

1. Host A tries to divert I/O traffic to System B.
2. Quorum Witness is still up and running, so System A and System B poll the Quorum Witness to check if the other storage system (System B or System A) is up and running.

3. Because Quorum Witness can still talk to both System A and System B, it keeps the volume on System A in primary role and places System B in I/O blocked status, as shown in Figure 5-12.



Availability Type
HyperSwap

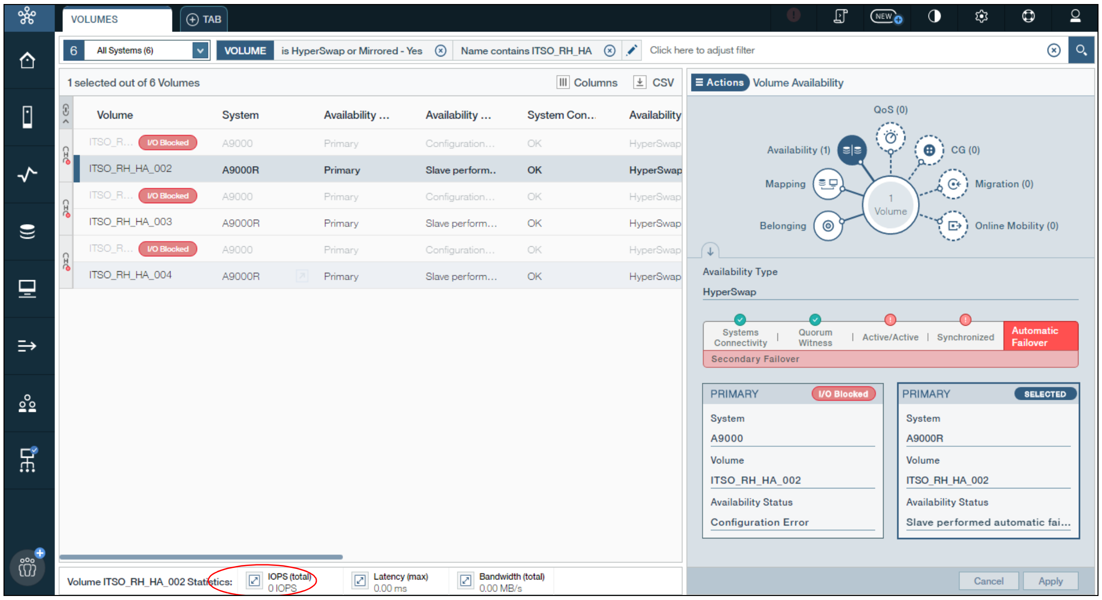
No connectivity exists between A9000 and A9000R

PRIMARY	SECONDARY
System	System
A9000	A9000R
Volume	Volume
ITSO_RH_HA_001	ITSO_RH_HA_001
Availability Status	Availability Status
Unsynchronized (Link Down)	Unsynchronized (Link Down)

Figure 5-12 System B I/O Blocked

Now when the connectivity between Host A and System A is restored, a manual failback can be done to System A and I/Os can be served again from System A.

Figure 5-13 shows the GUI view when the connectivity between Host A and System A is restored while the connectivity between System A and System B is still not restored.



1 selected out of 6 Volumes

Volume	System	Availability ...	Availability ...	System Con...	Availability
ITSO_R... I/O Blocked	A9000	Primary	Configuration...	OK	HyperSwap
ITSO_RH_HA_002	A9000R	Primary	Slave perform..	OK	HyperSwap
ITSO_R... I/O Blocked	A9000	Primary	Configuration...	OK	HyperSwap
ITSO_RH_HA_003	A9000R	Primary	Slave perform...	OK	HyperSwap
ITSO_R... I/O Blocked	A9000	Primary	Configuration...	OK	HyperSwap
ITSO_RH_HA_004	A9000R	Primary	Slave perform...	OK	HyperSwap

Availability Type
HyperSwap

Systems Connectivity ✓ | Quorum Witness ✓ | Active/Active ! | Synchronized ! | Automatic Failover

Secondary Failover

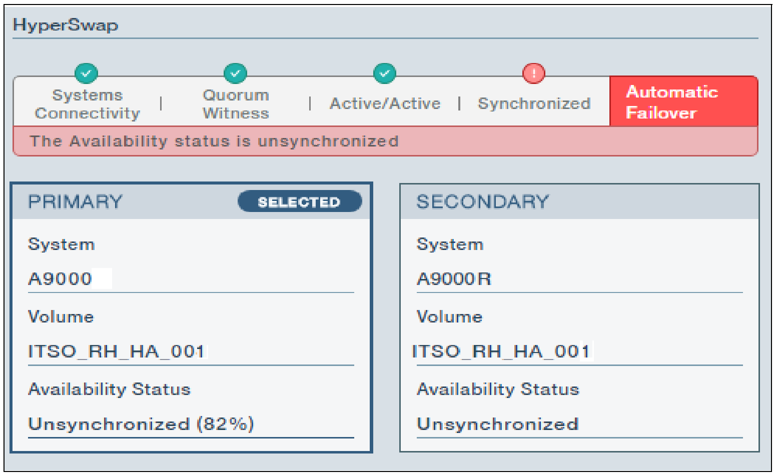
PRIMARY I/O Blocked	PRIMARY SELECTED
System	System
A9000	A9000R
Volume	Volume
ITSO_RH_HA_002	ITSO_RH_HA_002
Availability Status	Availability Status
Configuration Error	Slave performed automatic fai...

Volume ITSO_RH_HA_002 Statistics: IOPS (total) 0 IOPS | Latency (max) 0.00 ms | Bandwidth (total) 0.00 MB/s

Figure 5-13 Host Connectivity with System A restored: IOPS happening to System A

When the connectivity between System A and System B is restored, data synchronization between System A and System B occurs first and HyperSwap is then re-enabled.

Figure 5-14 shows the GUI view when the connectivity between System A and System B has been restored and data is being synchronized between both storage systems.



HyperSwap

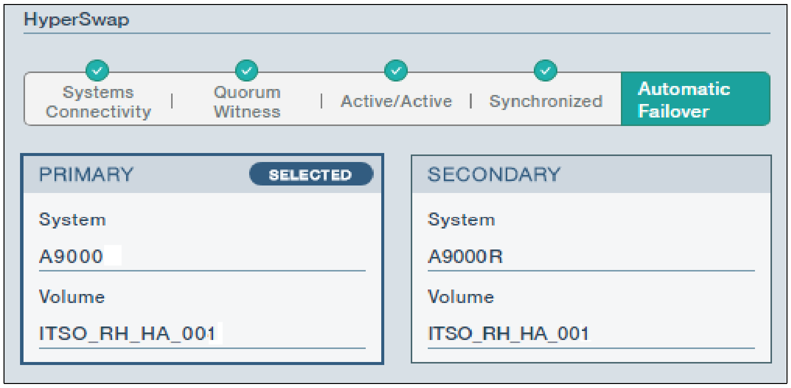
Systems Connectivity ✓ | Quorum Witness ✓ | Active/Active ✓ | Synchronized ! | Automatic Failover

The Availability status is unsynchronized

PRIMARY SELECTED	SECONDARY
System	System
A9000	A9000R
Volume	Volume
ITSO_RH_HA_001	ITSO_RH_HA_001
Availability Status	Availability Status
Unsynchronized (82%)	Unsynchronized

Figure 5-14 Connectivity between System A and System B is established

After the volume synchronization has completed, the GUI shows that the high availability is reestablished and an automatic failover is now possible again, as shown in Figure 5-15.



PRIMARY

System

A9000

Volume

ITSO_RH_HA_001

SECONDARY

System

A9000R

Volume

ITSO_RH_HA_001

Figure 5-15 Automatic failover re-established

5.3 Quorum Witness failure scenarios

This section covers the following failure scenarios that are related to the Quorum Witness (QW) availability:

- QW failure
- Loss of the Ethernet connection between System A and the QW
- QW connectivity failure to System A, followed by mirroring link failure
- QW down, restore from backup is possible
- QW down, new installation is needed

The host and the primary copy of the volume are at site A. The primary copy is serving I/O to host A. The QW on the third site is connected to System A on site A and System B on site B.

5.3.1 Quorum Witness failure

Figure 5-16 shows a QW failure scenario that might occur as a result of the QW server issue, or network problems, or even disaster at Site Q where the QW server resides.

When a QW is no longer available, then the HyperSwap automatic failover capability is disabled. However, because connectivity between System A and System B is normally still in place, replication between Site A and Site B is still operational and a manual failover (switch role) between System A and System B volumes remains possible.

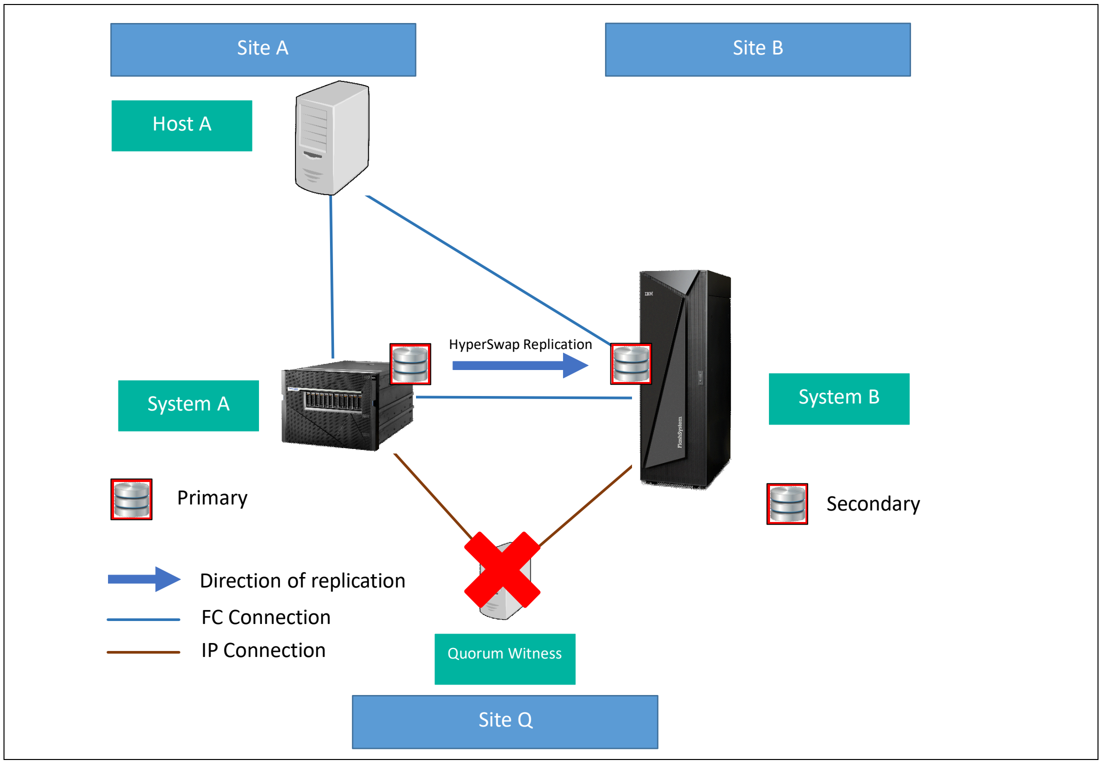
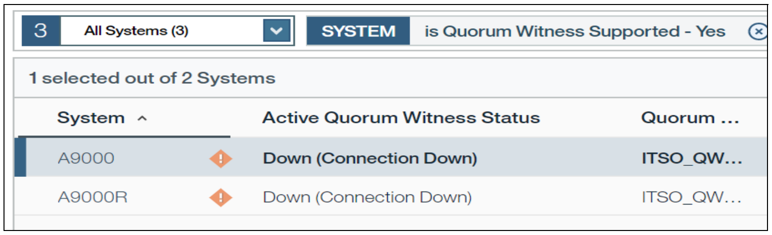


Figure 5-16 Quorum Witness failure

Figure 5-17 is the Hyper-Scale Manager GUI view indicating that QW connections to both System A (A9000) and System B (A9000R) are down.

To get to this view, select **Systems & Domains Views** → **Quorum Witnesses**.

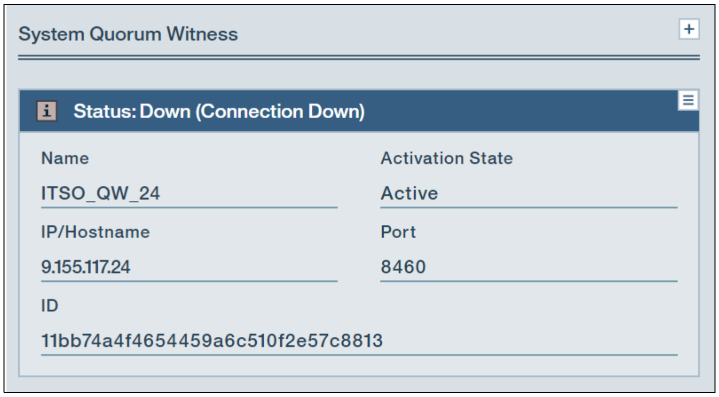

3 All Systems (3) | SYSTEM is Quorum Witness Supported - Yes

1 selected out of 2 Systems

System	Active Quorum Witness Status	Quorum ...
A9000	Down (Connection Down)	ITSO_QW...
A9000R	Down (Connection Down)	ITSO_QW...

Figure 5-17 Quorum Witness down

Figure 5-18 shows the Hyper-Scale manager GUI view of the Quorum witness connectivity status on System A.



System Quorum Witness

Status: Down (Connection Down)

Name: ITSO_QW_24

Activation State: Active

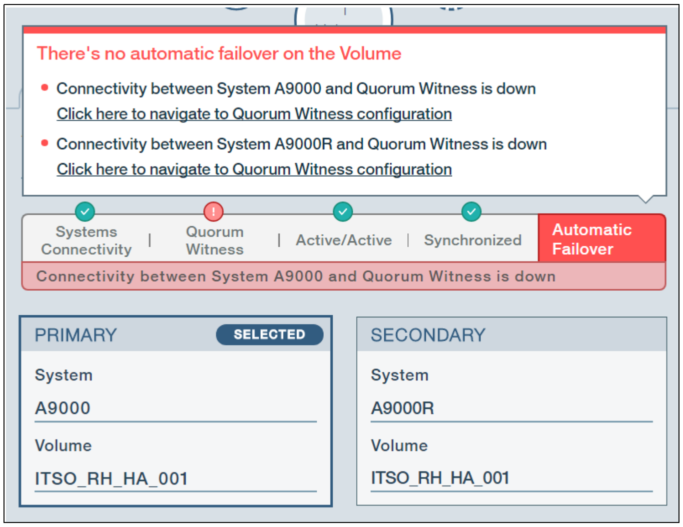
IP/Hostname: 9.155.117.24

Port: 8460

ID: 11bb74a4f4654459a6c510f2e57c8813

Figure 5-18 Connectivity between System A (A9000) and QW

Click **Remote Views** → **Mirrored/HyperSwap Volumes (Availability)** and select the concerned volume to view the **Volume Availability** details, as illustrated in Figure 5-19. Notice that it shows only QW as not available but System A and System B are still connected and replication between them is active and synchronized.



There's no automatic failover on the Volume

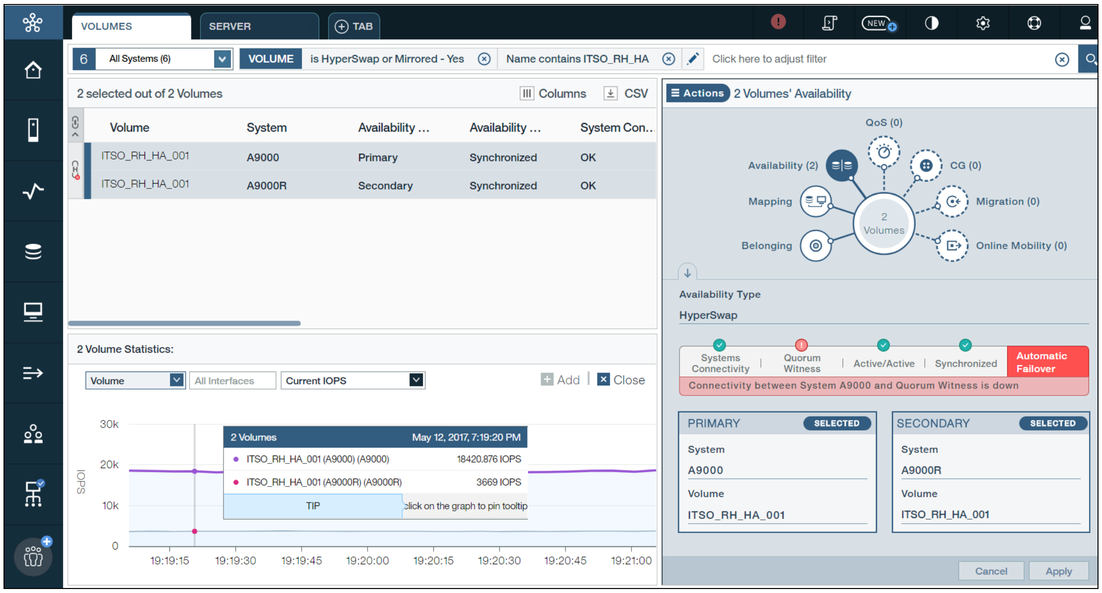
- Connectivity between System A9000 and Quorum Witness is down
  Click here to navigate to Quorum Witness configuration
- Connectivity between System A9000R and Quorum Witness is down
  Click here to navigate to Quorum Witness configuration

Connectivity between System A9000 and Quorum Witness is down

PRIMARY SELECTED	SECONDARY
System	System
A9000	A9000R
Volume	Volume
ITSO_RH_HA_001	ITSO_RH_HA_001

Figure 5-19 Volume availability view

As shown in Figure 5-20, I/O is served from System A (A9000) and the data is being replicated to System B (A9000R).



2 selected out of 2 Volumes

Volume	System	Availability ...	Availability ...	System Con..
ITSO_RH_HA_001	A9000	Primary	Synchronized	OK
ITSO_RH_HA_001	A9000R	Secondary	Synchronized	OK

2 Volumes May 12, 2017, 7:19:20 PM

ITSO_RH_HA_001 (A9000) (A9000)	18420.876 IOPS
ITSO_RH_HA_001 (A9000R) (A9000R)	3669 IOPS

TIP click on the graph to pin tooltip

Availability Type
HyperSwap

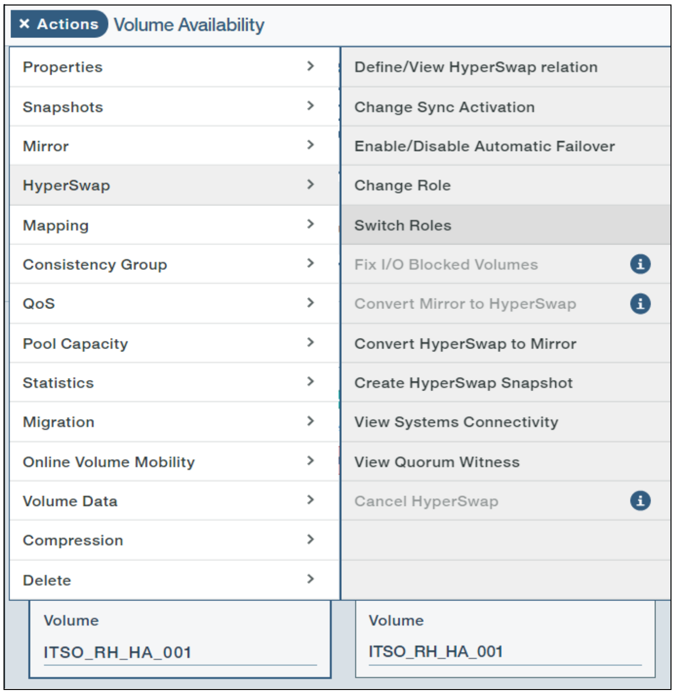
Connectivity between System A9000 and Quorum Witness is down

PRIMARY SELECTED
System
A9000
Volume
ITSO_RH_HA_001

SECONDARY SELECTED
System
A9000R
Volume
ITSO_RH_HA_001

Figure 5-20 Hyper-Scale Manager GUI view of I/Os being served and replicated

Although automatic failover is disabled in case of a QW failure, it is still possible to do a manual failover using the Switch Roles option, as shown in Figure 5-21.

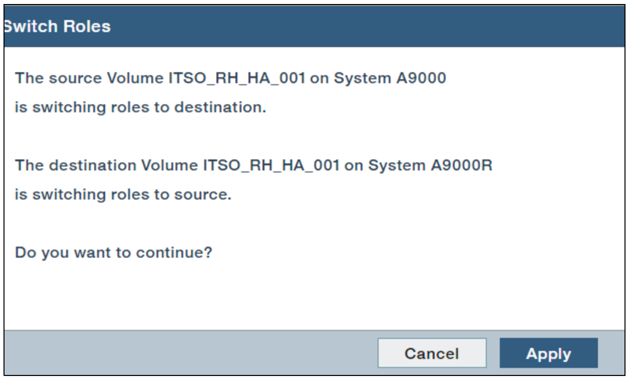


Volume
ITSO_RH_HA_001

Volume
ITSO_RH_HA_001

Figure 5-21 Switching roles of Storage systems in HyperSwap manually

After you select the **Switch Roles** option, confirm the action as shown in Figure 5-22. Then, click **Apply**.



Switch Roles

The source Volume ITSO_RH_HA_001 on System A9000

is switching roles to destination.

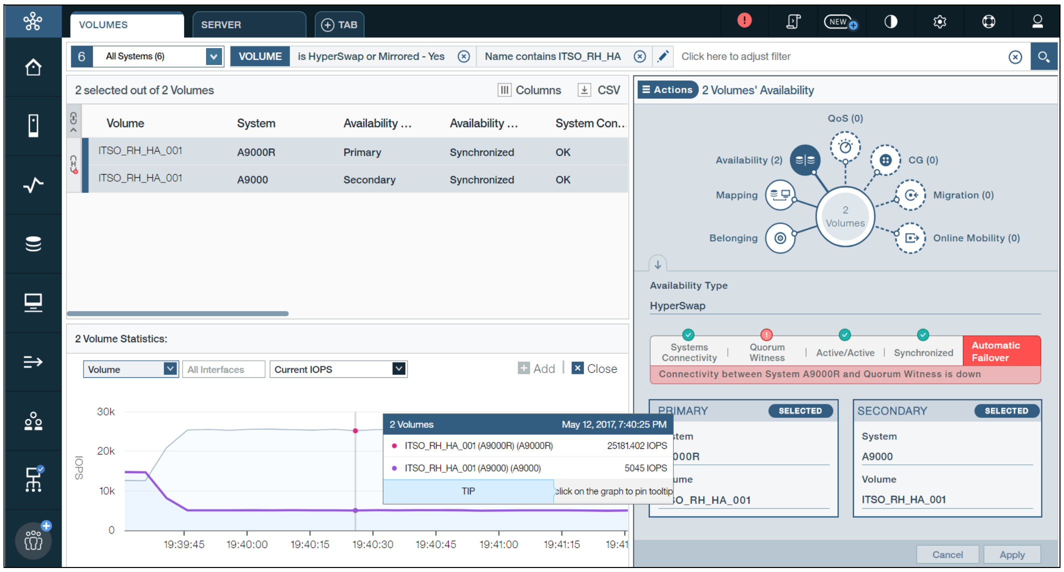
The destination Volume ITSO_RH_HA_001 on System A9000R

is switching roles to source.

Do you want to continue?

Figure 5-22 Prompt to confirm role switch for HyperSwap (Primary to Secondary)

When the switch role option is applied, system B (A9000R) is now serving I/Os to the host, as shown in Figure 5-23.



2 selected out of 2 Volumes

Volume	System	Availability ...	Availability ...	System Con...
ITSO_RH_HA_001	A9000R	Primary	Synchronized	OK
ITSO_RH_HA_001	A9000	Secondary	Synchronized	OK

2 Volume Statistics:

Volume | All Interfaces | Current IOPS

2 Volumes May 12, 2017, 7:40:25 PM

- ITSO_RH_HA_001 (A9000R) (A9000R) 25181.402 IOPS
- ITSO_RH_HA_001 (A9000) (A9000) 5045 IOPS

TIP click on the graph to pin tooltip

Availability Type

HyperSwap

Connectivity between System A9000R and Quorum Witness is down

PRIMARY SELECTED

tem

000R

ume

O_RH_HA_001

SECONDARY SELECTED

System

A9000

Volume

ITSO_RH_HA_001

Figure 5-23 Site B serving the I/O

Figure 5-24 shows that after the switch role is applied, I/Os from the host are served from System B and the data now is synchronously replicated to System A.

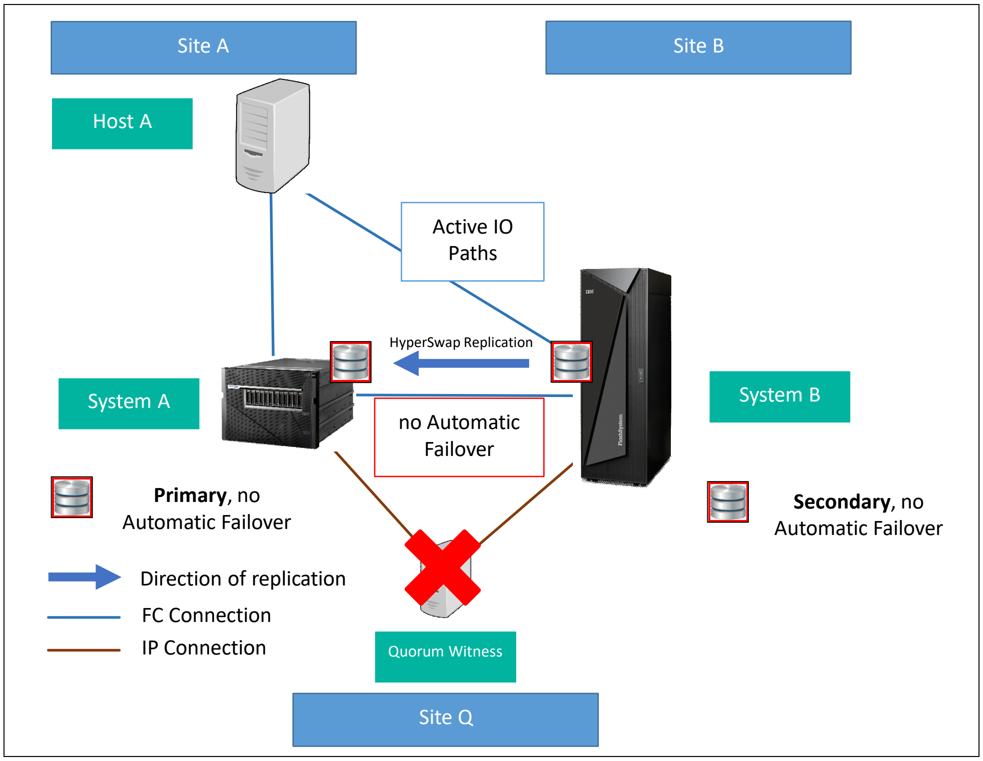


Figure 5-24 Host I/O being served from System B and are replicated to System A

When the QW starts working again and connects with the System A and System B, the automatic failover capability of the HyperSwap function is restored.

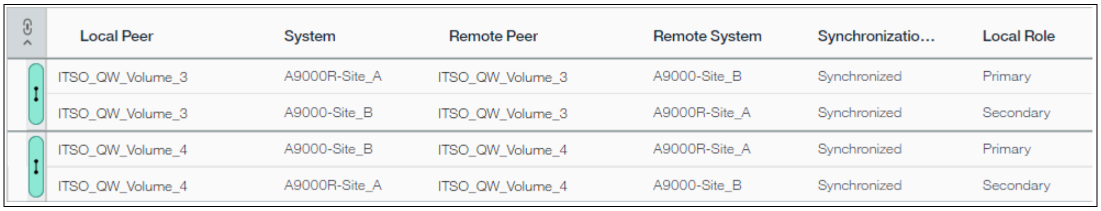
5.3.2 Ethernet connection loss between system A and QW

A loss of communication between a storage system and the Quorum witness affects the health of the HyperSwap relationships stemming from and to that system. A communication loss also compromises their normal operation in case of subsequent failures.

Automatic failover is compromised depending on the role of System A in the HyperSwap relationship. If A is primary, the HyperSwap relationship is in “orange” state, meaning that automatic failover is still possible if A fails or a loss of communication occurs between system A and system B (system B, holds the quorum at that point).

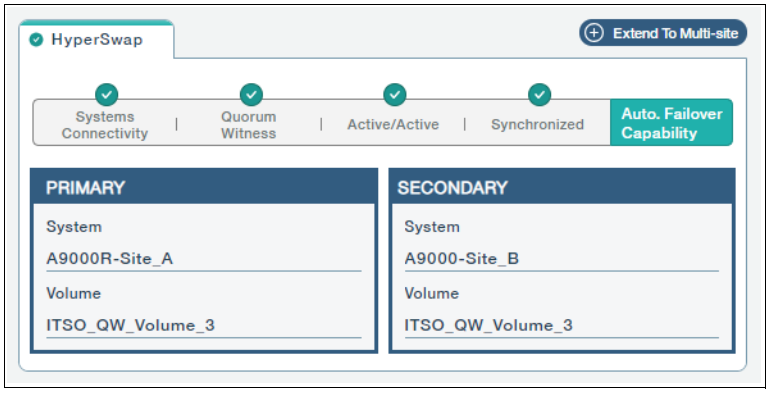
However, if A has the secondary role in the relationship, the HyperSwap state turns to “red”, which indicates that the automatic failover is not possible (if B were to fail, A does not hold the Quorum and therefore, cannot automatically assume the Primary role). Manual failover is still be possible.

In the two HyperSwap relationships shown in Figure 5-25, System A is designated respectively as Primary (Figure 5-26) and Secondary (Figure 5-27).



Local Peer	System	Remote Peer	Remote System	Synchronizatio...	Local Role
ITSO_QW_Volume_3	A9000R-Site_A	ITSO_QW_Volume_3	A9000-Site_B	Synchronized	Primary
ITSO_QW_Volume_3	A9000-Site_B	ITSO_QW_Volume_3	A9000R-Site_A	Synchronized	Secondary
ITSO_QW_Volume_4	A9000-Site_B	ITSO_QW_Volume_4	A9000R-Site_A	Synchronized	Primary
ITSO_QW_Volume_4	A9000R-Site_A	ITSO_QW_Volume_4	A9000-Site_B	Synchronized	Secondary

Figure 5-25 HyperSwap relationships in which System A perform different roles



PRIMARY

System
A9000R-Site_A

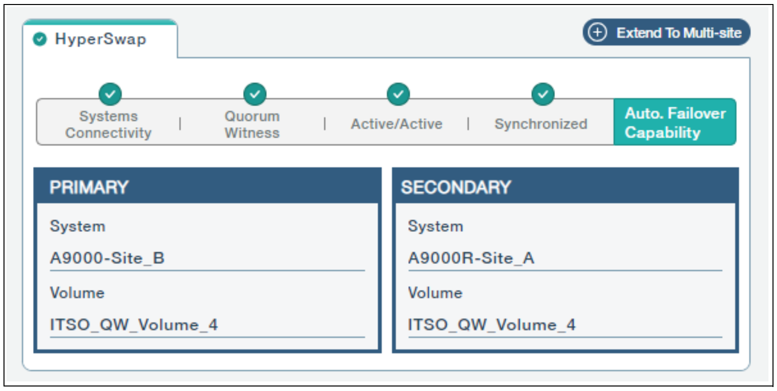
Volume
ITSO_QW_Volume_3

SECONDARY

System
A9000-Site_B

Volume
ITSO_QW_Volume_3

Figure 5-26 HyperSwap details for the relationship in which System A is Primary



PRIMARY

System
A9000-Site_B

Volume
ITSO_QW_Volume_4

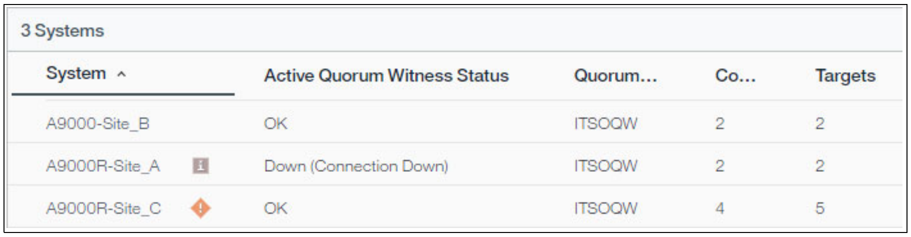
SECONDARY

System
A9000R-Site_A

Volume
ITSO_QW_Volume_4

Figure 5-27 HyperSwap details for the relationship in which System A is Secondary

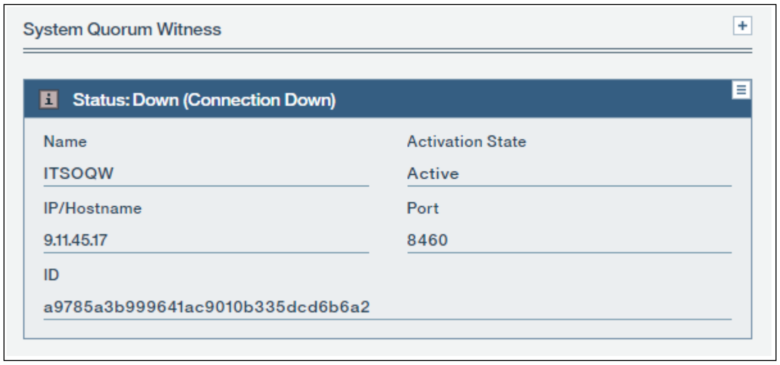
When System A loses connection to the Quorum Witness, its status is shown as Down in the Systems & Domains Views → Quorum Witnesses panel (see Figure 5-28) and information (see Figure 5-29 on page 76).



3 Systems

System ^	Active Quorum Witness Status	Quorum...	Co...	Targets
A9000-Site_B	OK	ITSOQW	2	2
A9000R-Site_A	Down (Connection Down)	ITSOQW	2	2
A9000R-Site_C	OK	ITSOQW	4	5

Figure 5-28 Down status for connection between Quorum Witness ITSOQW and System A



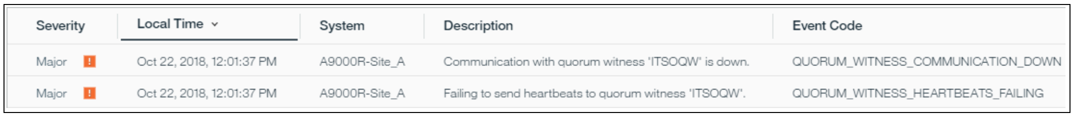
System Quorum Witness

Status: Down (Connection Down)

Name	Activation State
ITSOQW	Active
IP/Hostname	Port
9.11.45.17	8460
ID	
a9785a3b999641ac9010b335dcd6b6a2	

Figure 5-29 Quorum Witness connection details for System A

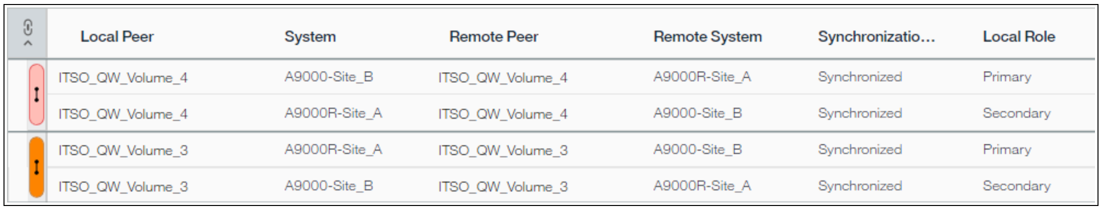
Corresponding entries are logged in the event list, which is reachable by selecting **Systems & Domains Views** → **Events**, as shown in Figure 5-30.



Severity	Local Time	System	Description	Event Code
Major !	Oct 22, 2018, 12:01:37 PM	A9000R-Site_A	Communication with quorum witness 'ITSOQW' is down.	QUORUM_WITNESS_COMMUNICATION_DOWN
Major !	Oct 22, 2018, 12:01:37 PM	A9000R-Site_A	Failing to send heartbeats to quorum witness 'ITSOQW'.	QUORUM_WITNESS_HEARTBEATS_FAILING

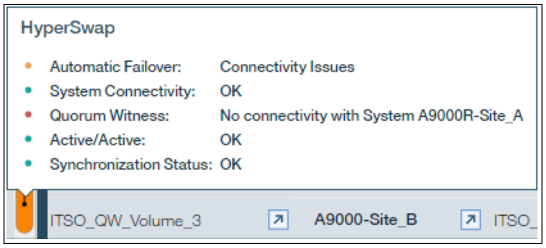
Figure 5-30 Loss of QW connections reported in the Events Log

The status of the HyperSwap relationships changes to reflect the failure, as shown in Figure 5-31. Notice the difference of behavior in the relationship where System A has the Primary role (see Figure 5-32 and Figure 5-33 on page 77) versus the one where it has Secondary role (see Figure 5-34 on page 77 and Figure 5-35 on page 77).



Local Peer	System	Remote Peer	Remote System	Synchronizatio...	Local Role
ITSO_QW_Volume_4	A9000-Site_B	ITSO_QW_Volume_4	A9000R-Site_A	Synchronized	Primary
ITSO_QW_Volume_4	A9000R-Site_A	ITSO_QW_Volume_4	A9000-Site_B	Synchronized	Secondary
ITSO_QW_Volume_3	A9000R-Site_A	ITSO_QW_Volume_3	A9000-Site_B	Synchronized	Primary
ITSO_QW_Volume_3	A9000-Site_B	ITSO_QW_Volume_3	A9000R-Site_A	Synchronized	Secondary

Figure 5-31 HyperSwap relationships status after connection loss between System A and the QW

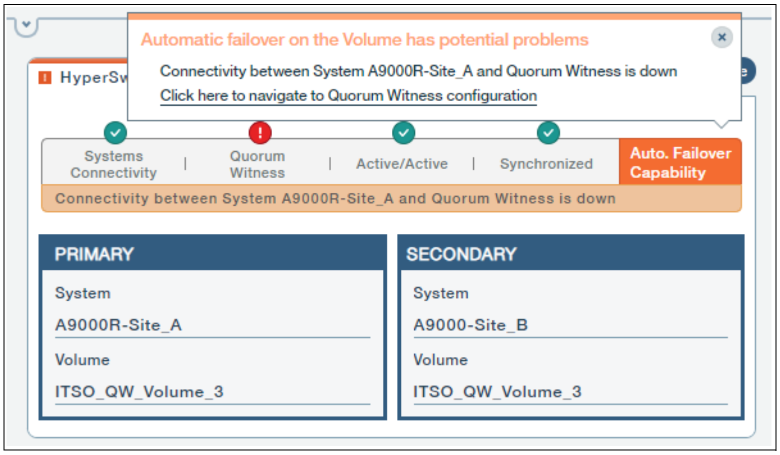


HyperSwap

● Automatic Failover:	Connectivity Issues
● System Connectivity:	OK
● Quorum Witness:	No connectivity with System A9000R-Site_A
● Active/Active:	OK
● Synchronization Status:	OK

ITSO_QW_Volume_3 A9000-Site_B ITSO_

Figure 5-32 Status of the relationship where System A has Primary role



Automatic failover on the Volume has potential problems

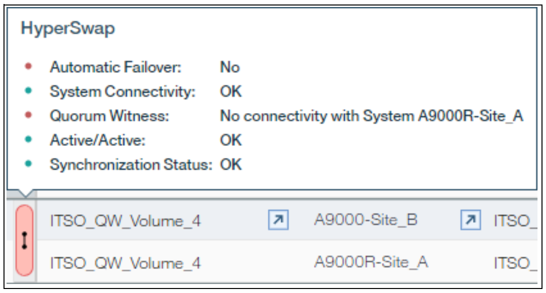
Connectivity between System A9000R-Site_A and Quorum Witness is down

Click here to navigate to Quorum Witness configuration

Connectivity between System A9000R-Site_A and Quorum Witness is down

PRIMARY	SECONDARY
System	System
A9000R-Site_A	A9000-Site_B
Volume	Volume
ITSO_QW_Volume_3	ITSO_QW_Volume_3

Figure 5-33 Automatic Failover reports potential problems, as shown in the relationship details

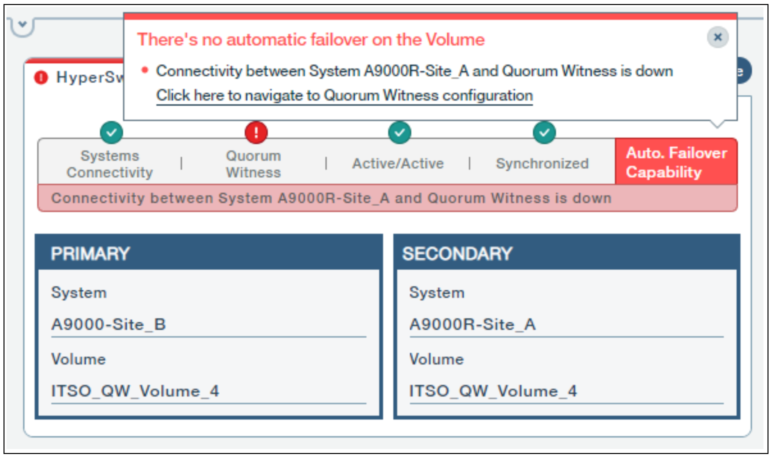


HyperSwap

Automatic Failover:	No
System Connectivity:	OK
Quorum Witness:	No connectivity with System A9000R-Site_A
Active/Active:	OK
Synchronization Status:	OK

ITSO_QW_Volume_4	A9000-Site_B	ITSO_
ITSO_QW_Volume_4	A9000R-Site_A	ITSO_

Figure 5-34 Status of the relationship where System A has Secondary role



There's no automatic failover on the Volume

- Connectivity between System A9000R-Site_A and Quorum Witness is down

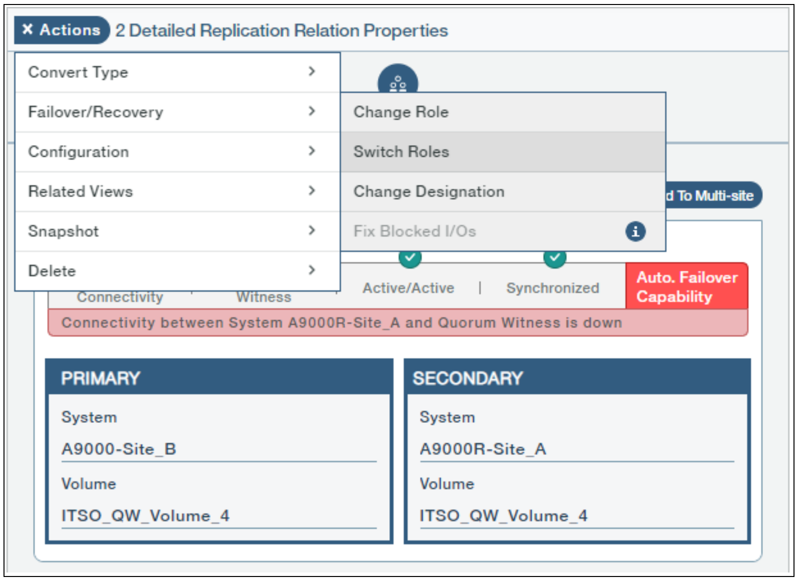
Click here to navigate to Quorum Witness configuration

Connectivity between System A9000R-Site_A and Quorum Witness is down

PRIMARY	SECONDARY
System	System
A9000-Site_B	A9000R-Site_A
Volume	Volume
ITSO_QW_Volume_4	ITSO_QW_Volume_4

Figure 5-35 In the relationship where System A has Secondary role, automatic failover is disabled

However, a manual failover in tis last case is still possible if needed, and can be performed by selecting **Failover/Recovery Switch Roles** in the relationship Actions menu, as shown in Figure 5-36.



Connectivity between System A9000R-Site_A and Quorum Witness is down

PRIMARY

System
A9000-Site_B

Volume
ITSO_QW_Volume_4

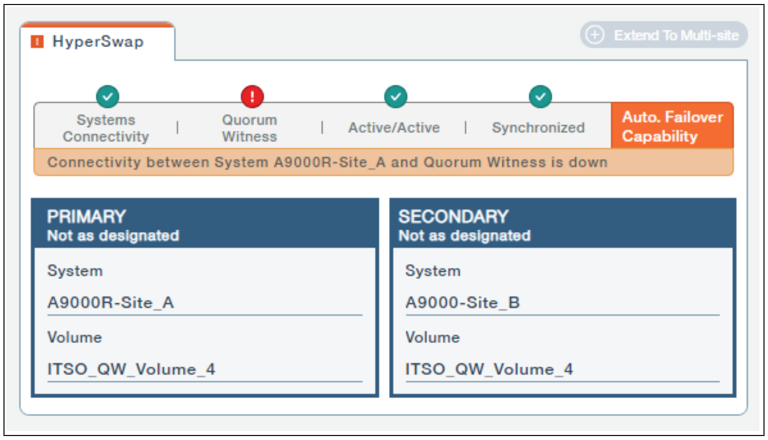
SECONDARY

System
A9000R-Site_A

Volume
ITSO_QW_Volume_4

Figure 5-36 Manually switching roles in the HyperSwap relationship

After confirming the role switch, the relationship enters the state that is shown in Figure 5-37, where Automatic Failover capability is now restored (System B is now Secondary and therefore can hold the QW and become Primary in case of failure on System A).



Connectivity between System A9000R-Site_A and Quorum Witness is down

PRIMARY
Not as designated

System
A9000R-Site_A

Volume
ITSO_QW_Volume_4

SECONDARY
Not as designated

System
A9000-Site_B

Volume
ITSO_QW_Volume_4

Figure 5-37 State of the relationship after the manual role switch

When the connection between System A and the Quorum Witness is restored, all of the HyperSwap relationships become healthy again, and both automatic and manual failover are possible.

5.3.3 QW connectivity failure to System A, followed by mirroring link failure

Figure 5-38 shows a connectivity failure between Quorum Witness and System A (A9000) along with a failure in the mirroring links between System A and System B. This scenario describes what happens when each failure occurs in succession (what is often described as a *rolling disaster*).

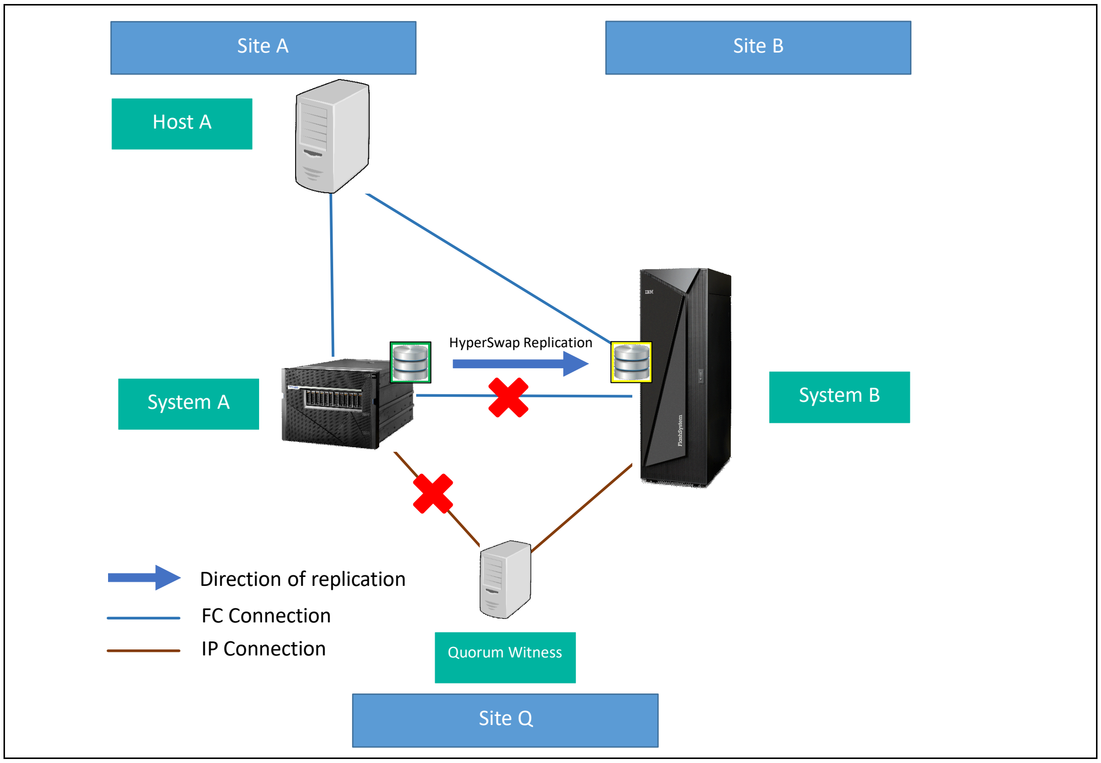
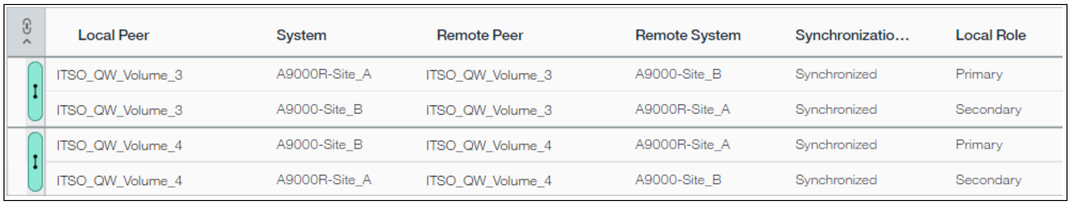


Figure 5-38 QW connectivity failure with failed mirroring links

We assume that the HyperSwap configuration was fine prior to the failures. The failure affecting the QW connectivity to the System A can occur because of a hardware failure on the server (the Ethernet card failed), a problem in the network preventing the QW from communicating with the system or potentially a failure on the system itself that affects the Ethernet connectivity.

This scenario does not affect host I/O to System A and System B because this connection is a Fibre Channel connection. However, depending on how the mirroring links failed, the host can also be affected with failed paths. In this example, we assume that the host paths are not affected to either system when the mirroring links fail.

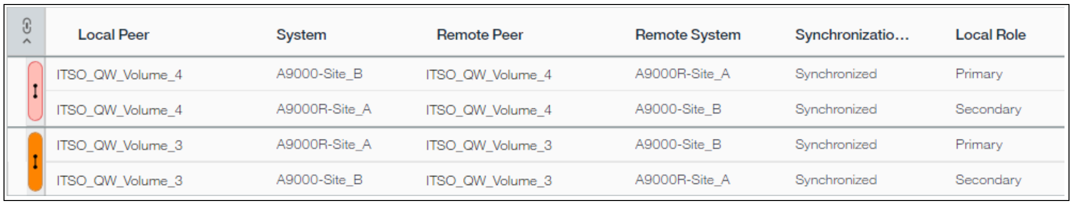
As described in 5.3.2, “Ethernet connection loss between system A and QW” on page 74, upon failure of the Quorum Witness connection, different behaviors occur in the HyperSwap relationship depending on the role of System A. Figure 5-39 shows the same examples, with two relationships in which System A has a different role.



Local Peer	System	Remote Peer	Remote System	Synchronizatio...	Local Role
ITSO_QW_Volume_3	A9000R-Site_A	ITSO_QW_Volume_3	A9000-Site_B	Synchronized	Primary
ITSO_QW_Volume_3	A9000-Site_B	ITSO_QW_Volume_3	A9000R-Site_A	Synchronized	Secondary
ITSO_QW_Volume_4	A9000-Site_B	ITSO_QW_Volume_4	A9000R-Site_A	Synchronized	Primary
ITSO_QW_Volume_4	A9000R-Site_A	ITSO_QW_Volume_4	A9000-Site_B	Synchronized	Secondary

Figure 5-39 Two HyperSwap relationships between System A and System B

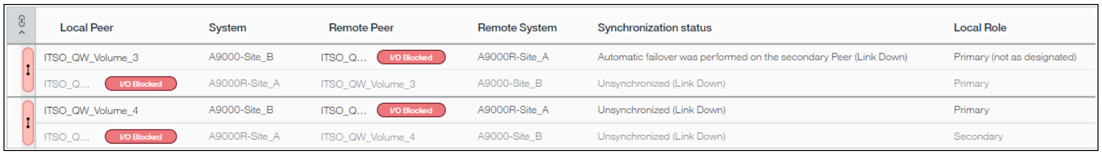
After System A loses its connection to the QW, the state of the two relationships changes, as shown in Figure 5-40.



Local Peer	System	Remote Peer	Remote System	Synchronizatio...	Local Role
ITSO_QW_Volume_4	A9000-Site_B	ITSO_QW_Volume_4	A9000R-Site_A	Synchronized	Primary
ITSO_QW_Volume_4	A9000R-Site_A	ITSO_QW_Volume_4	A9000-Site_B	Synchronized	Secondary
ITSO_QW_Volume_3	A9000R-Site_A	ITSO_QW_Volume_3	A9000-Site_B	Synchronized	Primary
ITSO_QW_Volume_3	A9000-Site_B	ITSO_QW_Volume_3	A9000R-Site_A	Synchronized	Secondary

Figure 5-40 State of the two HyperSwap relationships after the connectivity failure

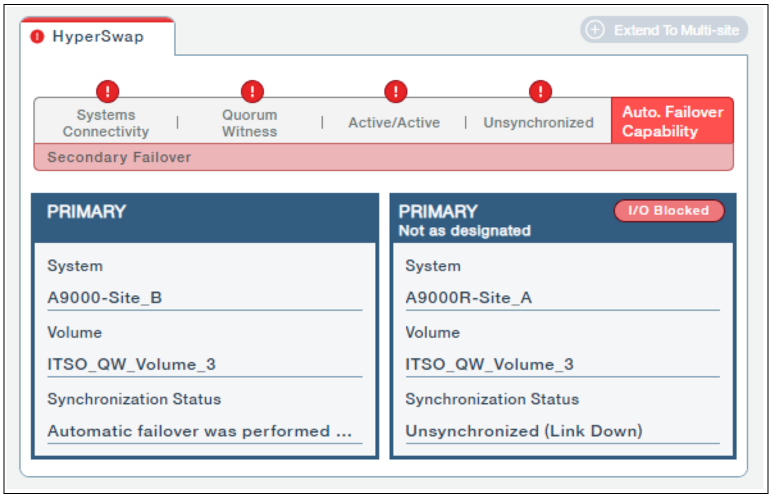
When the connectivity between System A and System B is interrupted, the situation that is shown in Figure 5-41 occurs.



Local Peer	System	Remote Peer	Remote System	Synchronization status	Local Role
ITSO_QW_Volume_3	A9000-Site_B	ITSO_Q... I/O Blocked	A9000R-Site_A	Automatic failover was performed on the secondary Peer (Link Down)	Primary (not as designated)
ITSO_Q... I/O Blocked	A9000R-Site_A	ITSO_QW_Volume_3	A9000-Site_B	Unsynchronized (Link Down)	Primary
ITSO_QW_Volume_4	A9000-Site_B	ITSO_Q... I/O Blocked	A9000R-Site_A	Unsynchronized (Link Down)	Primary
ITSO_Q... I/O Blocked	A9000R-Site_A	ITSO_QW_Volume_4	A9000-Site_B	Unsynchronized (Link Down)	Secondary

Figure 5-41 State of the two HyperSwap relationships after the connectivity failure

In the HyperSwap relationship where System A had the Primary role, an automatic failover was performed (because System B holding the QW). Missing connectivity with the QW and its partner system, System A was in compromised state and therefore blocked I/O on its copy of the volume. The HyperSwap relationship enters the compromised state, but host I/Os are still served through System B, as shown in Figure 5-42 on page 81.



Secondary Failover

PRIMARY

System

A9000-Site_B

Volume

ITSO_QW_Volume_3

Synchronization Status

Automatic failover was performed ...

PRIMARY
Not as designated

System

A9000R-Site_A

Volume

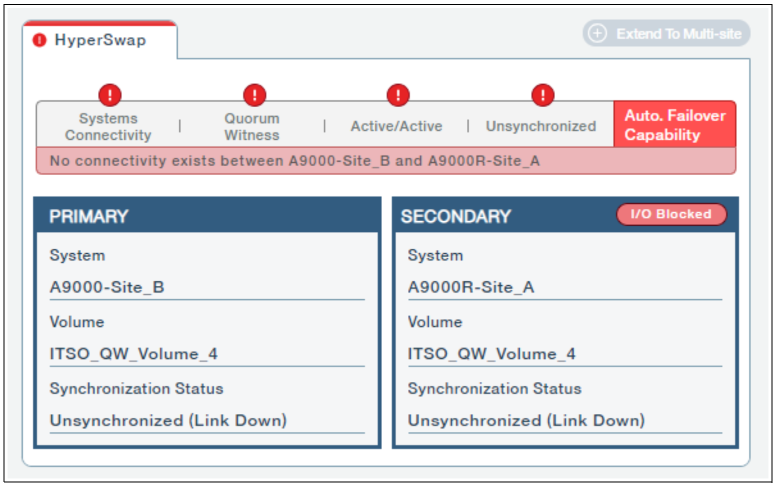
ITSO_QW_Volume_3

Synchronization Status

Unsynchronized (Link Down)

Figure 5-42   System B assumes Primary role after the Automatic Failover

In the HyperSwap relationship where System A had the Secondary role, System A was isolated from its partner system and the QW. Therefore, it blocked I/O on its copy of the volume to avoid possible data corruption. Recognizing the failure, System B held its Primary role and continued to serve host I/Os (see Figure 5-43).



No connectivity exists between A9000-Site_B and A9000R-Site_A

PRIMARY

System

A9000-Site_B

Volume

ITSO_QW_Volume_4

Synchronization Status

Unsynchronized (Link Down)

SECONDARY

System

A9000R-Site_A

Volume

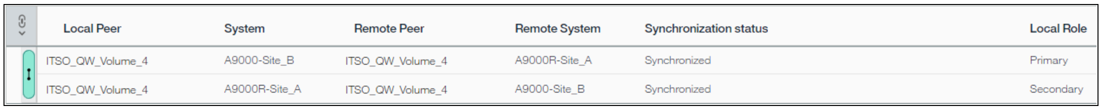
ITSO_QW_Volume_4

Synchronization Status

Unsynchronized (Link Down)

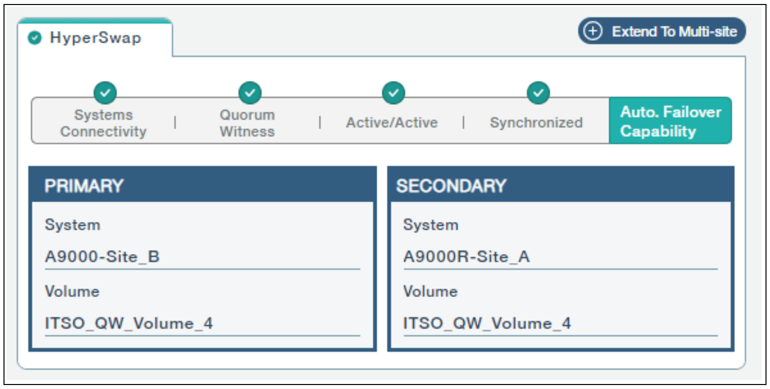
Figure 5-43   System B holds its Primary role, but I/Os are blocked on System A

When all failures are fixed, again different behaviors occur. The relationship in which System B was Primary and System A was Secondary automatically enters the Healthy state (after a brief resynchronization window), as shown in Figure 5-44 and Figure 5-45 on page 82.



Local Peer	System	Remote Peer	Remote System	Synchronization status	Local Role
ITSO_QW_Volume_4	A9000-Site_B	ITSO_QW_Volume_4	A9000R-Site_A	Synchronized	Primary
ITSO_QW_Volume_4	A9000R-Site_A	ITSO_QW_Volume_4	A9000-Site_B	Synchronized	Secondary

Figure 5-44   State of the HyperSwap relationship after the connectivity recovery



PRIMARY

System

A9000-Site_B

Volume

ITSO_QW_Volume_4

SECONDARY

System

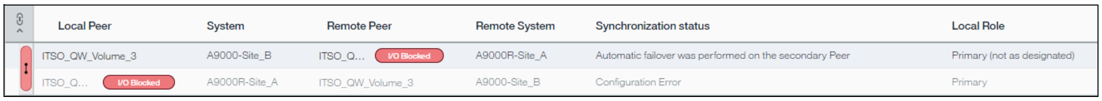
A9000R-Site_A

Volume

ITSO_QW_Volume_4

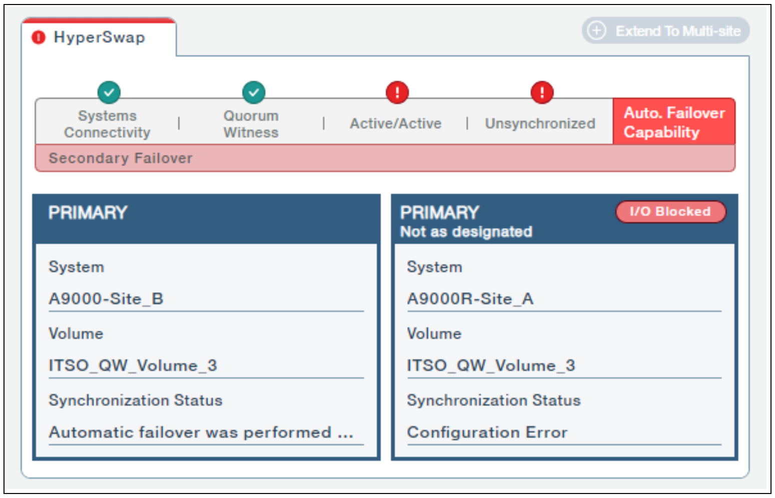
Figure 5-45 HyperSwap relationship details after the connectivity recovery

In the other relationship where System A was Designated Primary but System B took over the Primary role after the Automatic Failover, a role conflict now exists (see Figure 5-46 and Figure 5-47).



Local Peer	System	Remote Peer	Remote System	Synchronization status	Local Role
ITSO_QW_Volume_3	A9000-Site_B	ITSO_Q... I/O Blocked	A9000R-Site_A	Automatic failover was performed on the secondary Peer	Primary (not as designated)
ITSO_Q... I/O Blocked	A9000R-Site_A	ITSO_QW_Volume_3	A9000-Site_B	Configuration Error	Primary

Figure 5-46 Role conflict after the connectivity recovery



PRIMARY

System

A9000-Site_B

Volume

ITSO_QW_Volume_3

Synchronization Status

Automatic failover was performed ...

PRIMARY
Not as designated

System

A9000R-Site_A

Volume

ITSO_QW_Volume_3

Synchronization Status

Configuration Error

Figure 5-47 Relationship details shows the role conflict after the connectivity recovery

Note: Because I/Os are still blocked on the copy of the volume on System A, data is not compromised.

To return to the original state, we first need to resolve the role conflict. Because B served I/Os during the outage, we need to bring back the modified data to System A. Therefore, the first step is to temporarily change the role of System A to Secondary, by using the relationships Actions menu (click **Failover/Recovery** → **Change Role**), as shown in Figure 5-48.

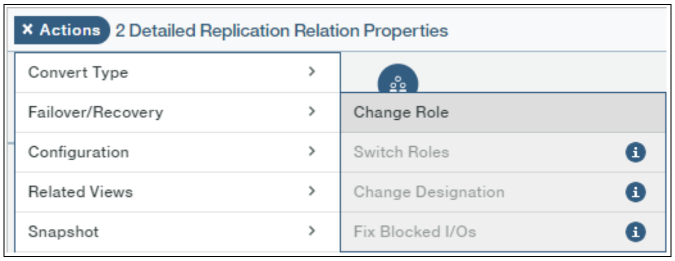
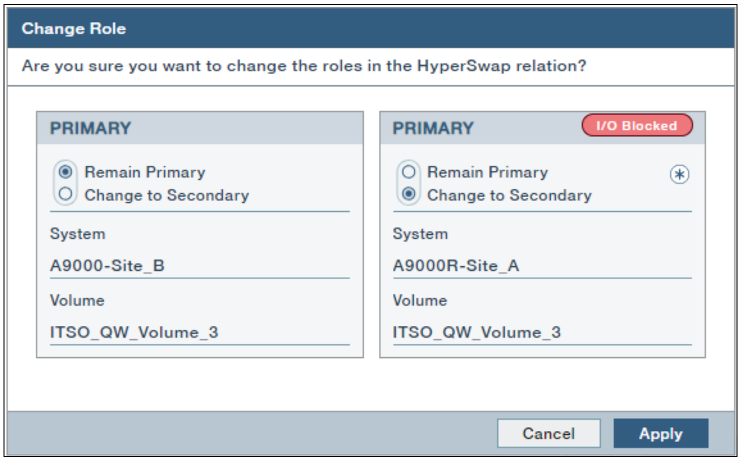


Figure 5-48 Performing a manual change role by way of the relationship's Actions menu

Confirm the role change (see Figure 5-49).



Change Role

Are you sure you want to change the roles in the HyperSwap relation?

PRIMARY	PRIMARY I/O Blocked
● Remain Primary ○ Change to Secondary	○ Remain Primary ● Change to Secondary
System	System
A9000-Site_B	A9000R-Site_A
Volume	Volume
ITSO_QW_Volume_3	ITSO_QW_Volume_3

Figure 5-49 Confirming the manual role change

Then, activate the relationship to start resynchronizing data from System B to System A (see Figure 5-50 on page 83).

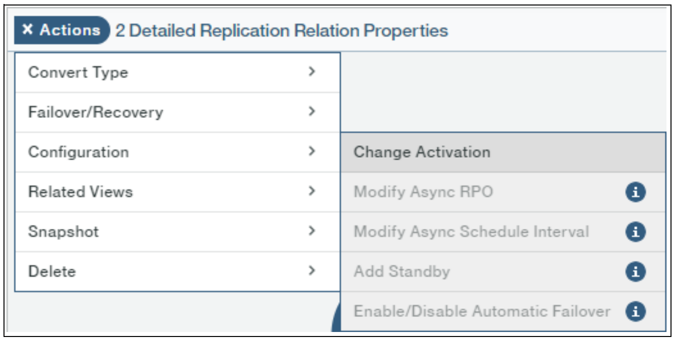
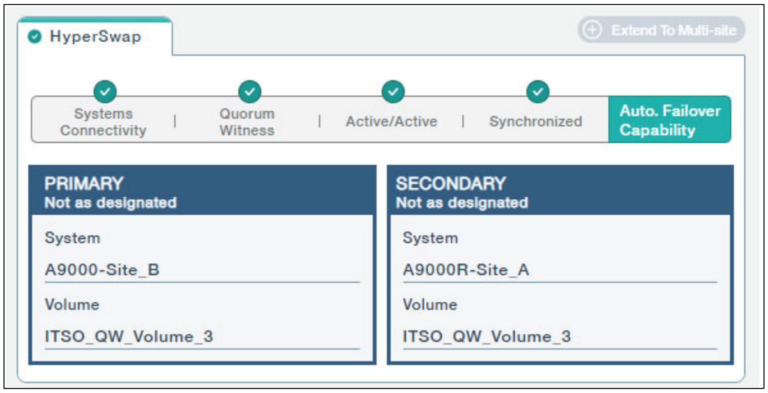


Figure 5-50 Activating the relationship

When the relationship is synchronized, it becomes healthy again (see Figure 5-51).



PRIMARY
Not as designated

System
A9000-Site_B

Volume
ITSO_QW_Volume_3

SECONDARY
Not as designated

System
A9000R-Site_A

Volume
ITSO_QW_Volume_3

Figure 5-51 Healthy state of the relationship, with both roles not as designated

At this point, we can perform a manual role switch to restore the initial configuration where System A is Primary and System B is Secondary. This switch is done by selecting **Configuration** → **Switch Roles** in the relationship Actions menu, as shown in Figure 5-52 on page 84.

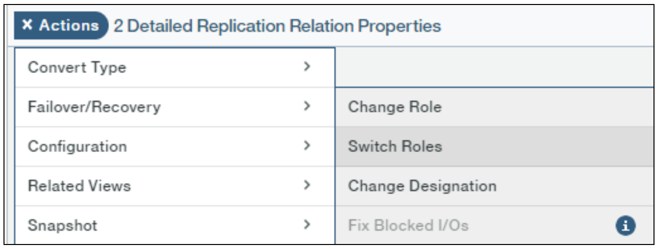
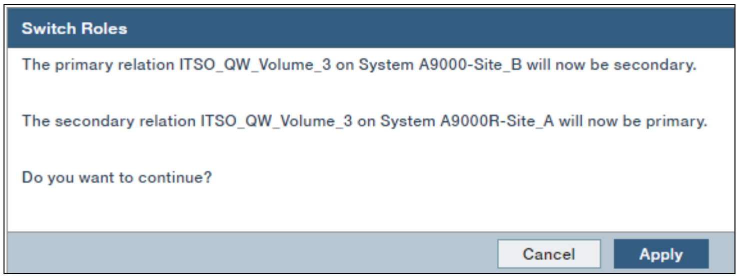


Figure 5-52 Performing a manual role switch

Confirm the switch (see Figure 5-53) to return to the initial state (see Figure 5-54 on page 85).



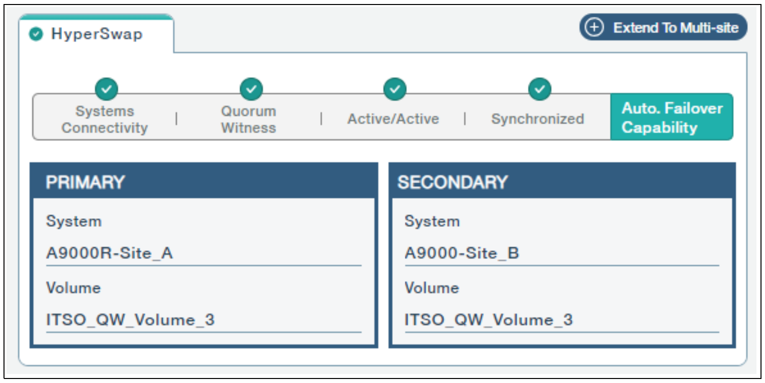
Switch Roles

The primary relation ITSO_QW_Volume_3 on System A9000-Site_B will now be secondary.

The secondary relation ITSO_QW_Volume_3 on System A9000R-Site_A will now be primary.

Do you want to continue?

Figure 5-53 Confirming the manual role switch



PRIMARY

System

A9000R-Site_A

Volume

ITSO_QW_Volume_3

SECONDARY

System

A9000-Site_B

Volume

ITSO_QW_Volume_3

Figure 5-54 Relationship finally returns its original configuration in Healthy state

5.3.4 Restoring or reinstalling the Quorum Witness

After a Quorum Witness (QW) failure you can restore the QW from a backup. Ensure that the following settings have not changed after the QW backup was taken:

- QW IP address
- QW Hostname
- QW certificate, in the qw.cert file

After restoring the QW the automatic failover capability becomes possible again.

Alternatively, you can reinstall using the QW saved configuration information. The following information is needed to reinstall the QW as it was:

- QW original IP address
- QW original Hostname
- QW original certificate files, the qw.cert and qw.key files

The qw.cert and the qw.key files are located in the `/opt/ibm/ibm_quorum_witness/settings/ssl_cert` directory on the QW server.

Follow these steps to re-install the QW:

1. Install the QW on a server with the original IP address and host name.
2. Copy the previously saved qw.cert and qw.key files to following folder on the QW server:

   ```
   opt/ibm/ibm_quorum_witness/settings/ssl_cert
   ```

3. Restart the **nginx** service, using the **`service nginx restart`** command.
4. Activate the Quorum Witness using the Hyper-Scale Manager.

After restoring the QW automatic failover capability becomes possible again.

5.4 Failback scenarios

After a failover, and when all the systems and connections have been reestablished, it is usually desirable to go back to the original configuration where the volume on System A at Site A has the primary role.

You also might need to recover from specific error situations. This section describes the following failback scenarios:

- Recovery from a failover where both volumes have primary role
- Recovery from an error during resynchronization

5.4.1 Recovery from a failover where both volumes have primary role

After a failover to the secondary volume of a HyperSwap relationship, the high availability of the HyperSwap volume has to be reestablished. This is only possible when the failed components are operational again. Because of the failover, the original primary volume on System A does not serve I/Os. It is in a blocked state, but still has the primary role.

The original secondary volume on System B is now serving the I/O and has the primary role. After the failed components are operational again the role of the original primary volume has to be manually changed to *secondary*.

Figure 5-55 shows the HyperSwap volume state after System A at site A is operational again. The HyperSwap peer volumes are not synchronized.

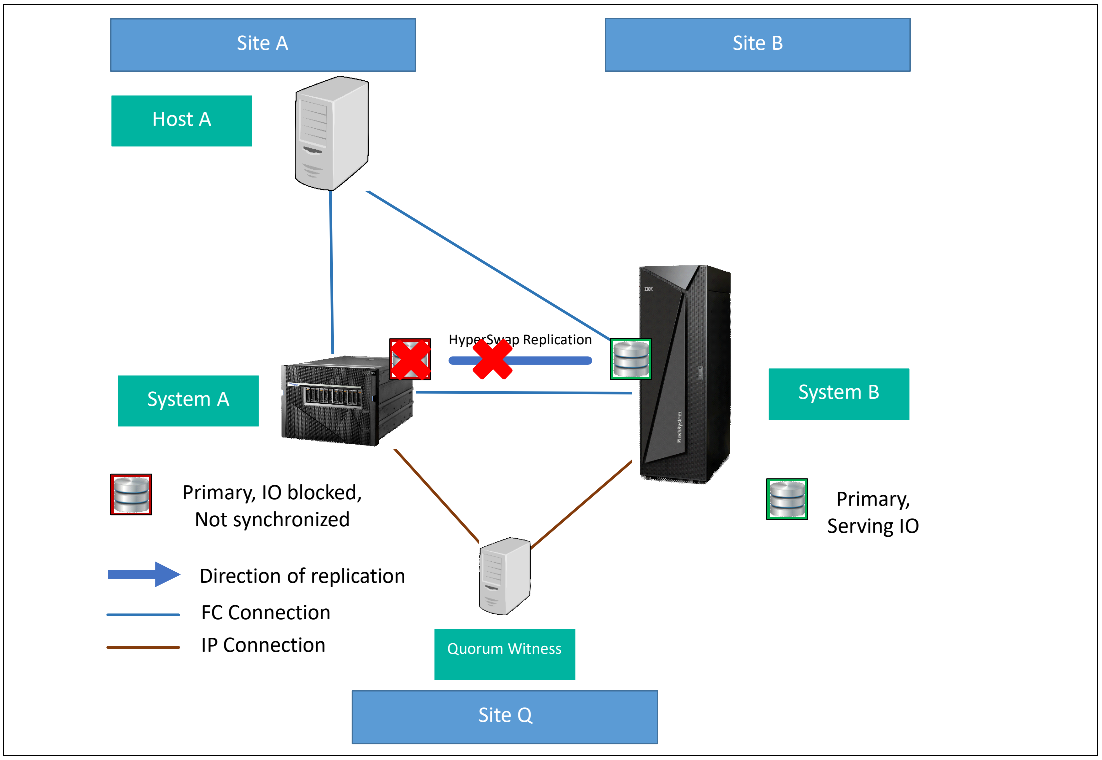


Figure 5-55 HyperSwap volume state after System A is operational again

The following steps are necessary to reestablish the high availability of the HyperSwap volume:

1. Set the correct role for the volume with blocked I/O.
2. Activate HyperSwap, which will automatically start the re-synchronization.

Optionally, change the **Primary** and **Secondary** roles so that System A serves read and write I/O to host A after a switch role on the HyperSwap volumes.

> **Note:** Synchronizing the volumes and activating HyperSwap is a manual process which has to be carefully planned and executed.

Reestablishing the high availability after a failover

The HyperSwap pane in the Hyper-Scale Manager GUI indicates the current state of a volume, as shown in Figure 5-4 on page 60. The detailed information shown in Figure 5-56 indicates that the system connectivity and the Quorum Witness are without any failure.

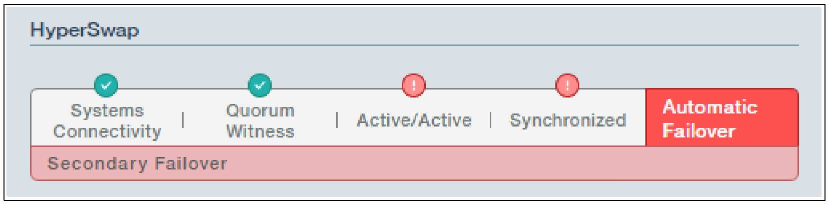
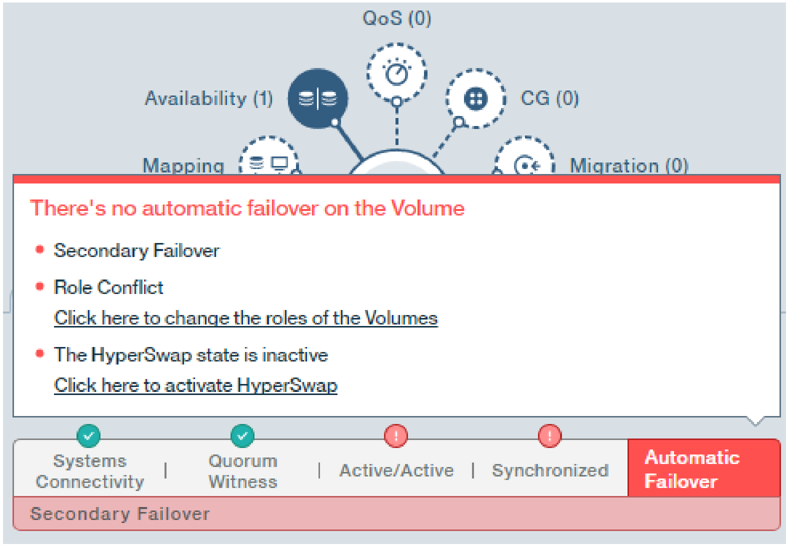


Figure 5-56 HyperSwap detailed availability status

Hovering over the failed components lists the necessary actions to restore the high availability of the volume, as shown in Figure 5-57.



There's no automatic failover on the Volume

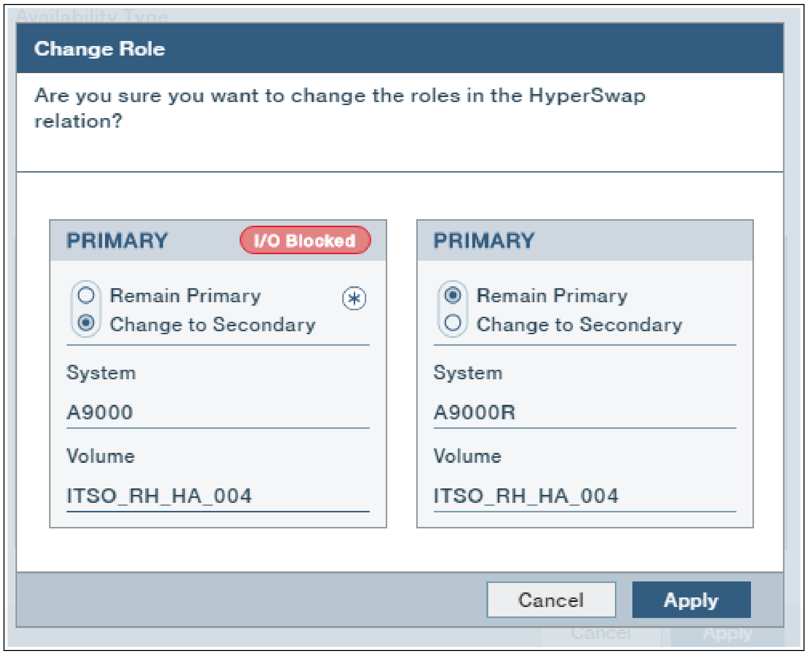
- Secondary Failover
- Role Conflict
  Click here to change the roles of the Volumes
- The HyperSwap state is inactive
  Click here to activate HyperSwap

Figure 5-57 Necessary actions to restore HyperSwap high availability

Select **Click here to change the roles of the volumes** to open the Change Role window.

The volume on System A at Site A has to become *secondary*, because the volume on System B at Site B is the current *primary*.

Select the **Change to Secondary** option and click **Apply**, as shown in Figure 5-58.



Change Role

Are you sure you want to change the roles in the HyperSwap relation?

PRIMARY I/O Blocked

○ Remain Primary
● Change to Secondary

System
A9000

Volume
ITSO_RH_HA_004

PRIMARY

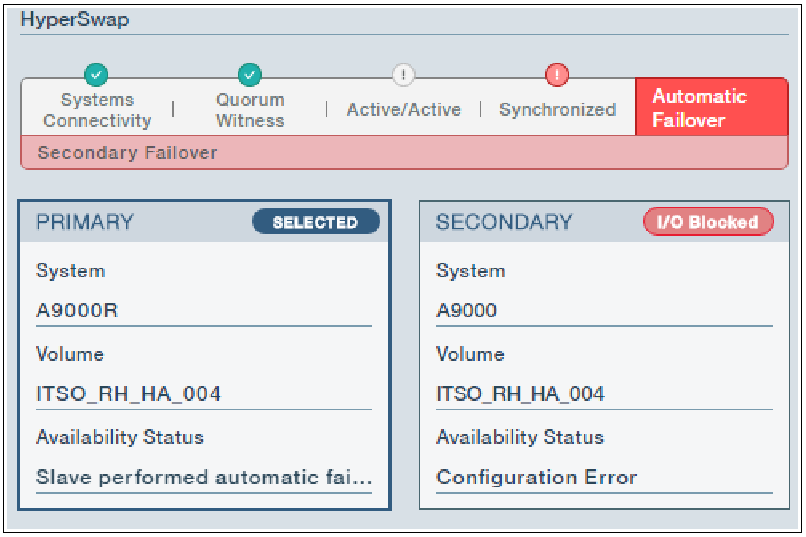
● Remain Primary
○ Change to Secondary

System
A9000R

Volume
ITSO_RH_HA_004

Figure 5-58 Changing the role of an I/O blocked volume

Figure 5-59 shows the result of changing the role.



HyperSwap

✓ Systems Connectivity | ✓ Quorum Witness | ! Active/Active | ! Synchronized | Automatic Failover

Secondary Failover

PRIMARY SELECTED

System
A9000R

Volume
ITSO_RH_HA_004

Availability Status
Slave performed automatic fai…

SECONDARY I/O Blocked

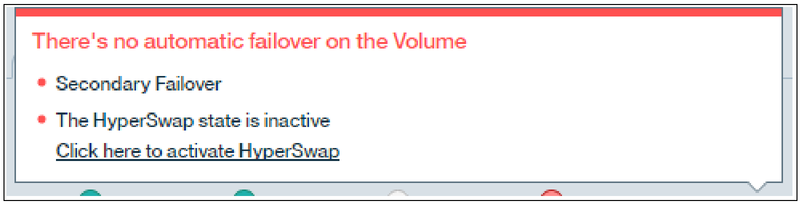
System
A9000

Volume
ITSO_RH_HA_004

Availability Status
Configuration Error

Figure 5-59 Correct role of an I/O blocked volume

Hovering over the failed components displays the second and last necessary actions to restore the high availability (automatic failover) of the volume, as shown in Figure 5-60.



There's no automatic failover on the Volume

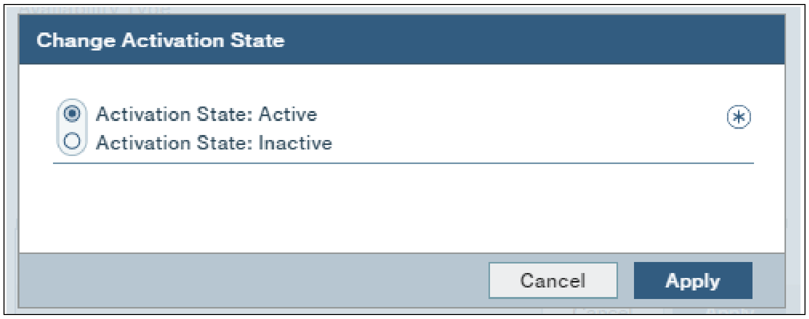
- Secondary Failover
- The HyperSwap state is inactive

Click here to activate HyperSwap

Figure 5-60 Last necessary action to restore HyperSwap volume high availability

Select **Click here to activate HyperSwap**, as shown in Figure 5-60 to open the Change Activation State dialog.

Select the **Activation State: Active** option and click **Apply,** as shown in Figure 5-61.

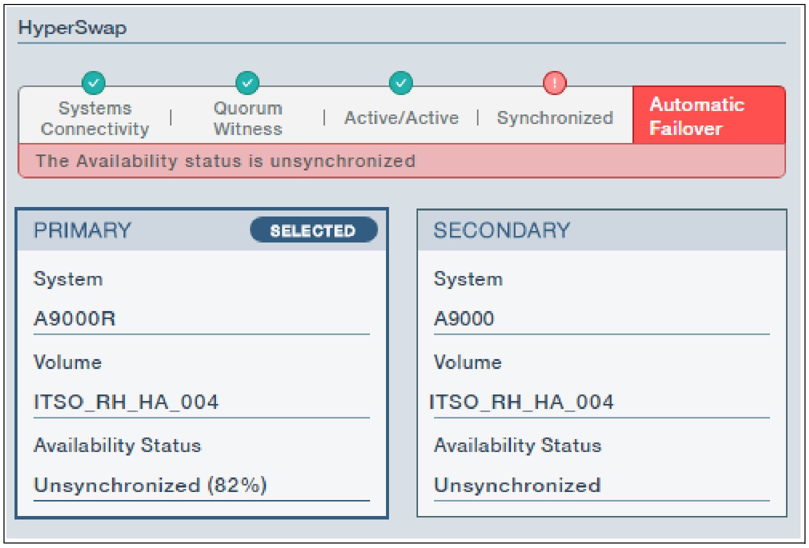


Change Activation State

● Activation State: Active
○ Activation State: Inactive

Figure 5-61 Activate window

Figure 5-62 shows the result of activating HyperSwap. The volumes on System A and System B will be automatically synchronized. The figure shows a value of 82%, which means that 82% of the HyperSwap volume are currently synchronized.



HyperSwap

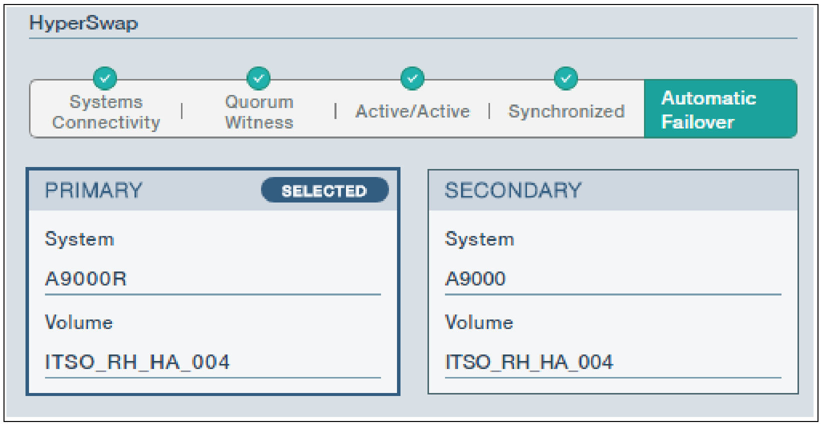
The Availability status is unsynchronized

PRIMARY SELECTED	SECONDARY
System	System
A9000R	A9000
Volume	Volume
ITSO_RH_HA_004	ITSO_RH_HA_004
Availability Status	Availability Status
Unsynchronized (82%)	Unsynchronized

Figure 5-62 Synchronizing the HyperSwap volume

A snapshot of the secondary volume is taken before starting synchronization from the primary volume to the secondary volume. The snapshot represents the last consistent copy of the secondary volume before it is synchronized. See 5.4.2, “Recovery from an error during resynchronization” on page 92 for an example of when that snapshot can be used. After successful synchronisation of the secondary volume, the snapshot is deleted.

After the volumes are synchronized, the GUI indicates that the high availability is reestablished and an automatic failover is now possible again, as shown in Example 5-63.



System

A9000R

Volume

ITSO_RH_HA_004

System

A9000

Volume

ITSO_RH_HA_004

Figure 5-63 Re-synchronized HyperSwap volume

The high availability of the HyperSwap volume is now reestablished but I/O is served from System B (A9000R) volume which is now the primary. Before the failover I/O was served from System A.

Switching the I/O back to the original primary (volume on System A)

Figure 5-63 on page 90 shows that the primary volume serving I/O is on System B (A9000R). This is not the original configuration that existed before the failover.

To fail back to the original configuration where the volume on System A was in primary role, use the **Switch Role** action for the HyperSwap volume.

> **Important:** The **Switch Role** action should be performed only *after* the synchronization process is complete.

Select the HyperSwap volume, right-click, and select **HyperSwap** → **Switch Role**, as shown in Figure 5-64 on page 91.

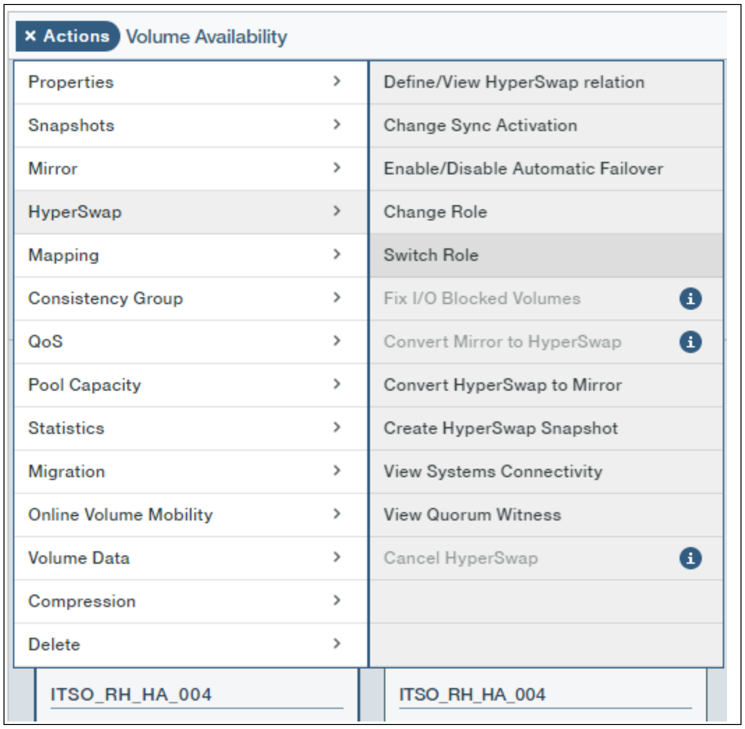
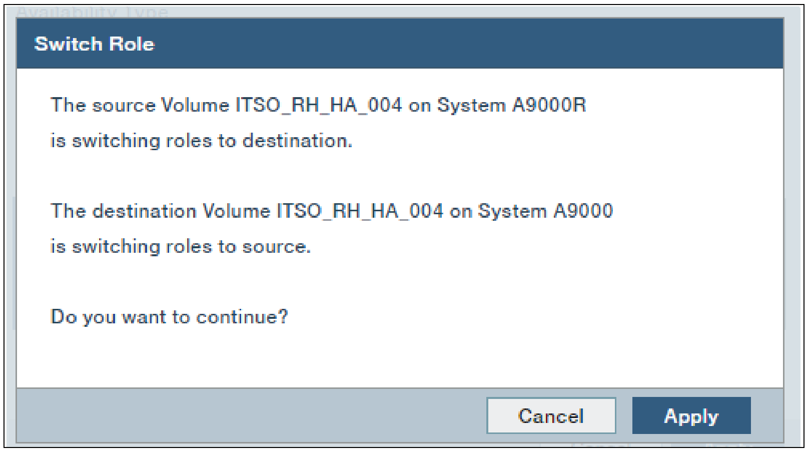


Figure 5-64 Action menu to select HyperSwap Switch Role

In the *Switch Role* dialog for the selected HyperSwap volume, click **Apply** to start the switch role process, as shown in Figure 5-65.



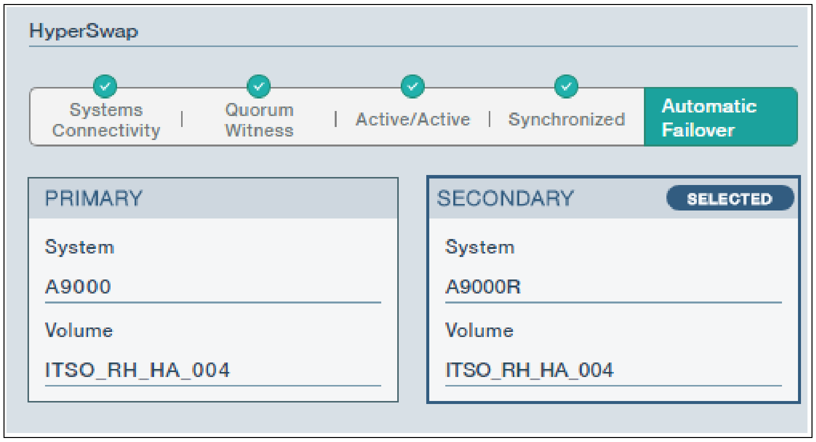
The source Volume ITSO_RH_HA_004 on System A9000R is switching roles to destination.

The destination Volume ITSO_RH_HA_004 on System A9000 is switching roles to source.

Do you want to continue?

Figure 5-65 HyperSwap volume switch role dialog

The primary HyperSwap volume is now back on System A, as shown in Figure 5-66. I/O is now served again from the original site, as it was before the failover occurred.



HyperSwap

Systems Connectivity | Quorum Witness | Active/Active | Synchronized | Automatic Failover

PRIMARY

System

A9000

Volume

ITSO_RH_HA_004

SECONDARY SELECTED

System

A9000R

Volume

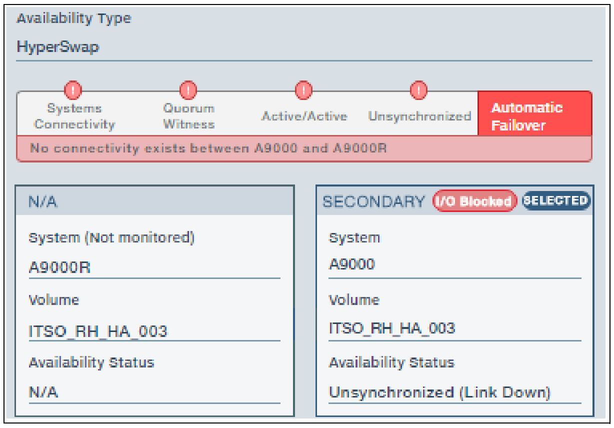
ITSO_RH_HA_004

Figure 5-66 Result of switching roles

5.4.2 Recovery from an error during resynchronization

This section describes the recovery from an error during the resynchronization of the HyperSwap volumes following a failover. The steps to start the re-synchronisation of the secondary volume are described in 5.4.1, “Recovery from a failover where both volumes have primary role” on page 86. The status during synchronisation of the secondary volume is shown in Figure 5-62 on page 89.

If System B with the current primary volume has an outage during the re-synchronisation, no I/O access from the host is possible. The secondary volume (currently on System A) is in **I/O blocked** state and the primary volume is not available, as shown in Figure 5-67.



Availability Type

HyperSwap

Systems Connectivity | Quorum Witness | Active/Active | Unsynchronized | Automatic Failover

No connectivity exists between A9000 and A9000R

N/A

System (Not monitored)

A9000R

Volume

ITSO_RH_HA_003

Availability Status

N/A

SECONDARY I/O Blocked SELECTED

System

A9000

Volume

ITSO_RH_HA_003

Availability Status

Unsynchronized (Link Down)

Figure 5-67 Volume availability after two failures

Because System B is not available, no information from the volume on System B can be displayed.

The secondary volume contains data from different times:

- Data from the point system A stopped
- Data partially updated during the synchronisation

Two scenarios are possible to get the volume in a storage consistent status:

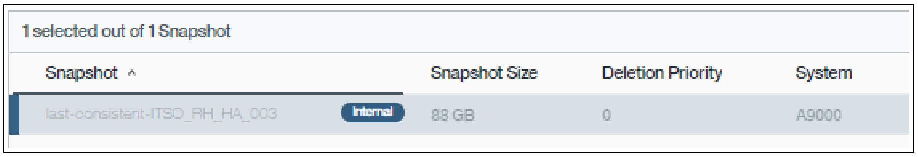
- Waiting until System B is available again
- Using the secondary volume for further host access

Waiting until System B is available again

The synchronisation of the secondary volume will start automatically after System B is available again. After synchronization, the automatic failover capability of the volume is re-established. The administrator can now use the *switch role* function to choose the site where the primary volume should reside.

Using the secondary volume for further host access

The last consistent status of the secondary volume is kept in the snapshot taken just before the synchronisation started. Reverting the volume to this snapshot will present a volume to the host, which is consistent from the storage side. Figure 5-68 shows this last consistent snapshot. This snapshot is automatically generated and marked as *internal*.

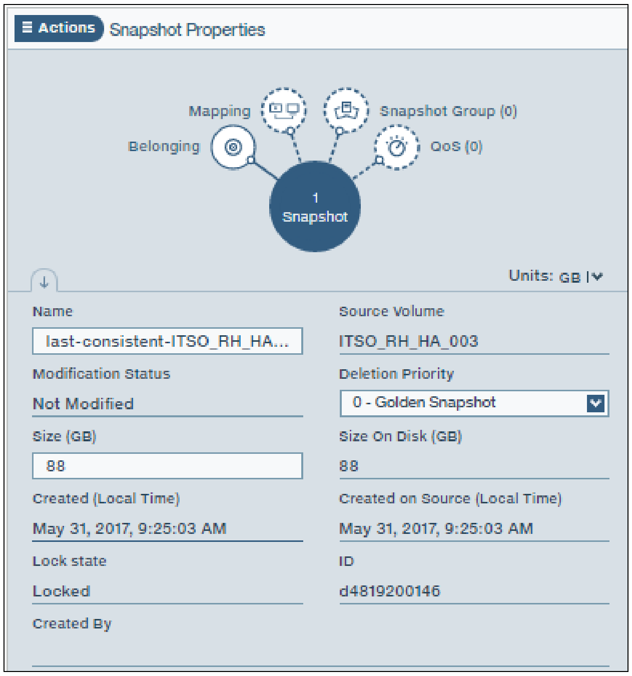


1 selected out of 1 Snapshot

Snapshot ^		Snapshot Size	Deletion Priority	System
last-consistent-ITSO_RH_HA_003	Internal	88 GB	0	A9000

Figure 5-68 Last consistent snapshot

Figure 5-69 shows the details of the internal last consistent snapshot.



Name last-consistent-ITSO_RH_HA...	Source Volume ITSO_RH_HA_003
Modification Status Not Modified	Deletion Priority 0 - Golden Snapshot
Size (GB) 88	Size On Disk (GB) 88
Created (Local Time) May 31, 2017, 9:25:03 AM	Created on Source (Local Time) May 31, 2017, 9:25:03 AM
Lock state Locked	ID d4819200146
Created By	

Figure 5-69 Details on the internal last consistent snapshot

The following steps are needed to revert the secondary volume to this last consistent snapshot and to serve I/O to the host:

1. Unmap the volume from the host.
2. Change role from secondary to primary.
3. Deactivate HyperSwap for this volume.

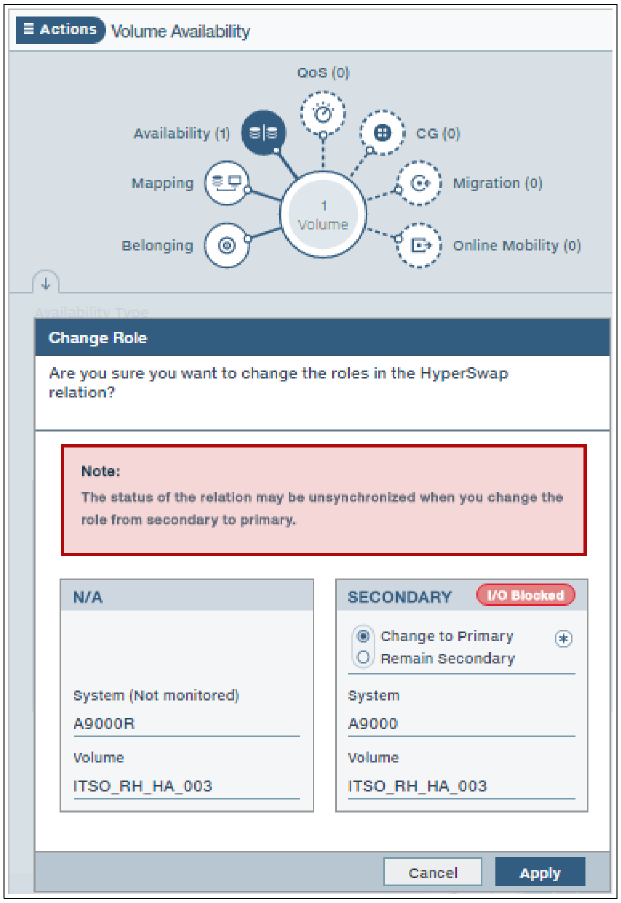
4. Restore the last consistent snapshot.
5. Map the volume from the host.
6. Delete snapshot.

These steps are described next.

The volume can be unmapped from the host using the Hyper-Scale Manager GUI.

Figure 5-70 shows how to change the role of the volume to primary. If the volume were still mapped to the host, the host can write and read inconsistent data on the volume. A note in the change role window is describing this situation. The volume should only be accessible by the host after the last consistent snapshot is restored.

> **Note:** The volume has to be unmapped from the host before switching its role to primary.



Change Role

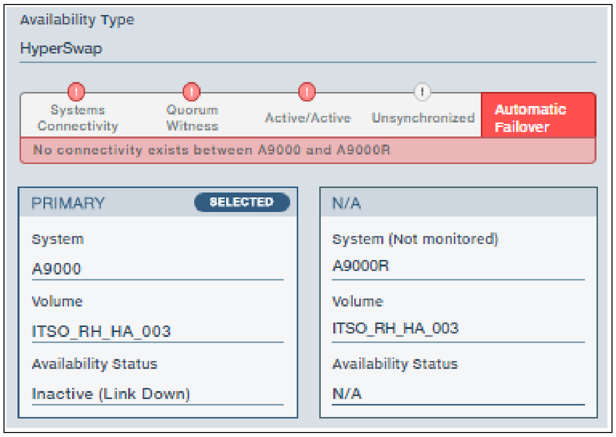
Are you sure you want to change the roles in the HyperSwap relation?

Note:
The status of the relation may be unsynchronized when you change the role from secondary to primary.

N/A	SECONDARY I/O Blocked
	● Change to Primary ○ Remain Secondary
System (Not monitored)	System
A9000R	A9000
Volume	Volume
ITSO_RH_HA_003	ITSO_RH_HA_003

Figure 5-70 Change role to primary

Figure 5-71 show the result of the **switch role** command.



No connectivity exists between A9000 and A9000R

PRIMARY SELECTED

System
A9000

Volume
ITSO_RH_HA_003

Availability Status
Inactive (Link Down)

N/A

System (Not monitored)
A9000R

Volume
ITSO_RH_HA_003

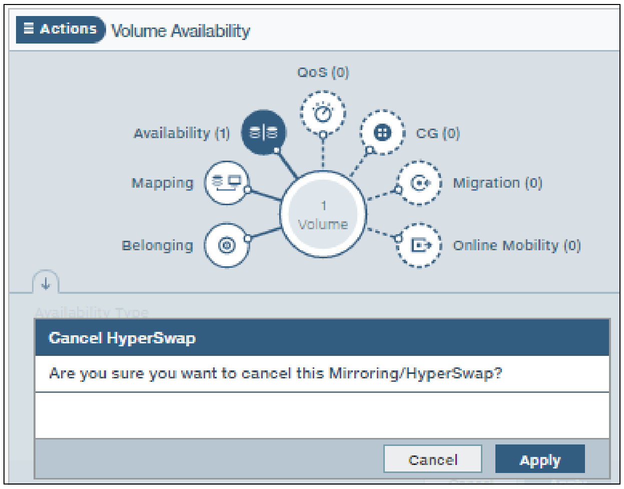
Availability Status
N/A

Figure 5-71 Result of the switch role command

The **switch role** command implies a renaming of the internal last consistent snapshot. It is not internal anymore and can be used by the administrator:

- Previous name: *last-consistent-<Volume Name>*
- New name: *external-last-consistent-<Volume Name>*

In our example *<Volume Name>* is *ITSO_RH_HA_003*. The HyperSwap relationship has to be cancelled, as shown in Figure 5-72, to be able to revert the volume to the last consistent snapshot.



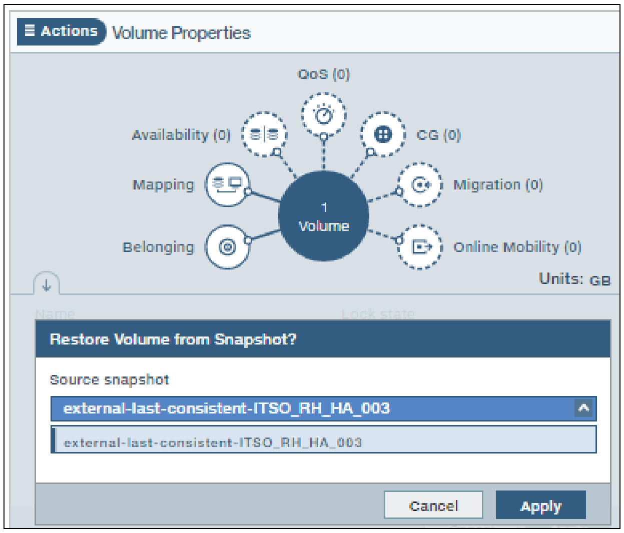
Cancel HyperSwap

Are you sure you want to cancel this Mirroring/HyperSwap?

Cancel   Apply

Figure 5-72 Deactivating HyperSwap

To restore the last consistent snapshot select **Snapshots** → **Restore from Snapshot** from the **Actions** Menu and choose the *external-last-consistent-<Volume Name>* snapshot, as shown in Figure 5-73.

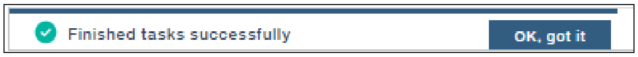


external-last-consistent-ITSO_RH_HA_003

external-last-consistent-ITSO_RH_HA_003

Figure 5-73 Restoring the last consistent snapshot

Click **Apply** to restore the snapshot. The Hyper-Scale Manager GUI indicates that the operation was successful, as shown in Figure 5-74.



Finished tasks successfully

Figure 5-74 Successful snapshot restore

The volume is now consistent from the storage side and can be mapped to a host. The last consistent snapshot can now be deleted.

6

Multi-site HA/DR implementation and usage

This chapter describes how to implement and use the Multi-site HA/DR feature for FlashSystem A9000 and FlashSystem A9000R.

The following topics are covered:

6.1 Multi-site HA/DR implementation process overview

To ensure smooth implementation, review Chapter 3, “Prerequisites” on page 17.

The scenarios are based on the configuration that is shown in Figure 6-1.

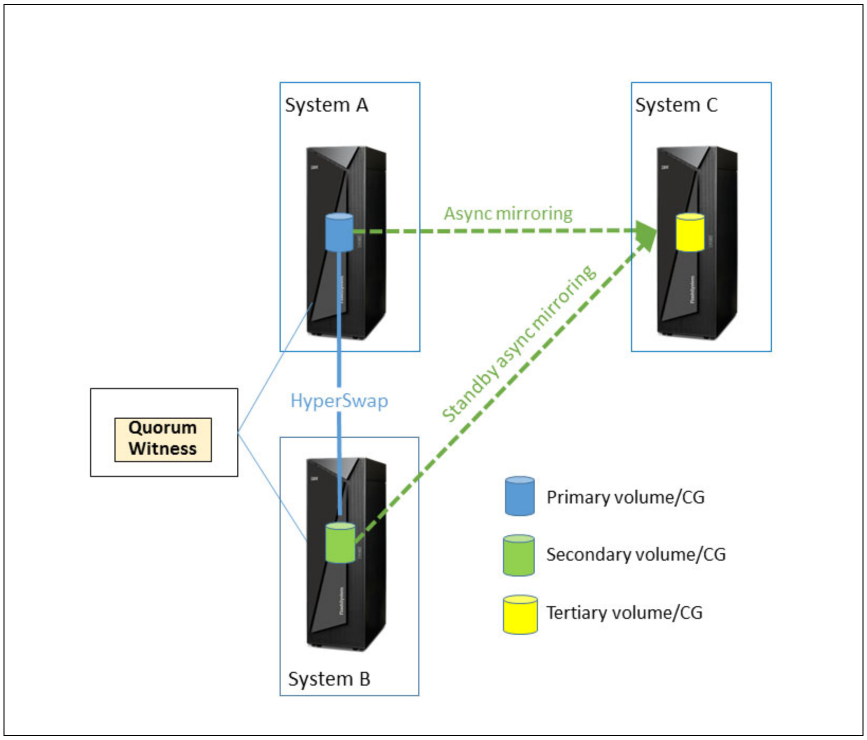


Figure 6-1 Basic Multi-site HA/DR configuration

Multi-site HA/DR CLI commands

The CLI commands that are used in the implementation scenarios are listed in Table 6-1.

Table 6-1 Multi-site HA/DR CLI commands

Command	Description
`multisite_define`	If a HyperSwap relation is established between the Primary and Secondary volumes, and an asynchronous mirror is defined between the Primary and Tertiary volumes, issuing this command on the Primary volume creates a named Multi-site relationship for a volume or consistency group. If the command completes successfully, the Multi-site relationship are recognized on all the involved systems.
`multisite_activate_async_mirror`	When issued on the Primary volume or Consistency Group, this command activates an asynchronous mirror as part of a Multi-site relationship.
`multisite_register_standby_mirror`	This command registers a Standby mirror in a multi-site relationship. When applied to a Consistency Group, this command registers the Standby relation on every volume in the Consistency Group.

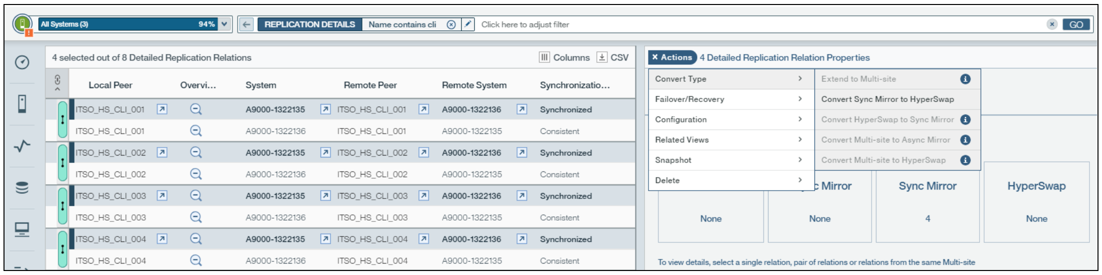
Command	Description
`multisite_list`	This command lists the configuration and status of multi-site relationships.
`multisite_change_role`	When issued for a non-operational relationship, this command changes the role of a multi-site relationship peer.
`multisite_switch_roles`	When issued on the Primary volume or Consistency Group, this command switches the roles between the Primary and the Secondary volumes or Consistency Groups.
`multisite_delete`	This command deletes a multi-site relationship. The two-way relationships that compose the multi-site relationship are not affected.

6.1.1 Assembly from scratch

Creating a Multi-site HA/DR configuration from scratch implies that you have a set of newly created or existing volumes or Consistency Groups that are not replicated to other systems. In this case, the suggested implementation strategy is to set up a HyperSwap relationship for these volumes or Consistency Groups (for more information, see 4.1, “HyperSwap implementation process overview” on page 24) and then extend it to a multi-site relationship, as described in 6.1.2, “Assembly from existing Synchronous relationship”.

6.1.2 Assembly from existing Synchronous relationship

Creating a Multi-site HA/DR relationship from an existing synchronous mirror is a relatively simple process. Select the volumes in the **Remote Views** → **Replication Details** table and select **Convert Type** → **Convert Sync Mirror to HyperSwap** in the Actions menu, as shown in Figure 6-2.



4 selected out of 8 Detailed Replication Relations

Local Peer	Overvi...	System	Remote Peer	Remote System	Synchronizatio...
ITSO_HS_CLI_001		A9000-1322135	ITSO_HS_CLI_001	A9000-1322136	Synchronized
ITSO_HS_CLI_001		A9000-1322136	ITSO_HS_CLI_001	A9000-1322135	Consistent
ITSO_HS_CLI_002		A9000-1322135	ITSO_HS_CLI_002	A9000-1322136	Synchronized
ITSO_HS_CLI_002		A9000-1322136	ITSO_HS_CLI_002	A9000-1322135	Consistent
ITSO_HS_CLI_003		A9000-1322135	ITSO_HS_CLI_003	A9000-1322136	Synchronized
ITSO_HS_CLI_003		A9000-1322136	ITSO_HS_CLI_003	A9000-1322135	Consistent
ITSO_HS_CLI_004		A9000-1322135	ITSO_HS_CLI_004	A9000-1322136	Synchronized
ITSO_HS_CLI_004		A9000-1322136	ITSO_HS_CLI_004	A9000-1322135	Consistent

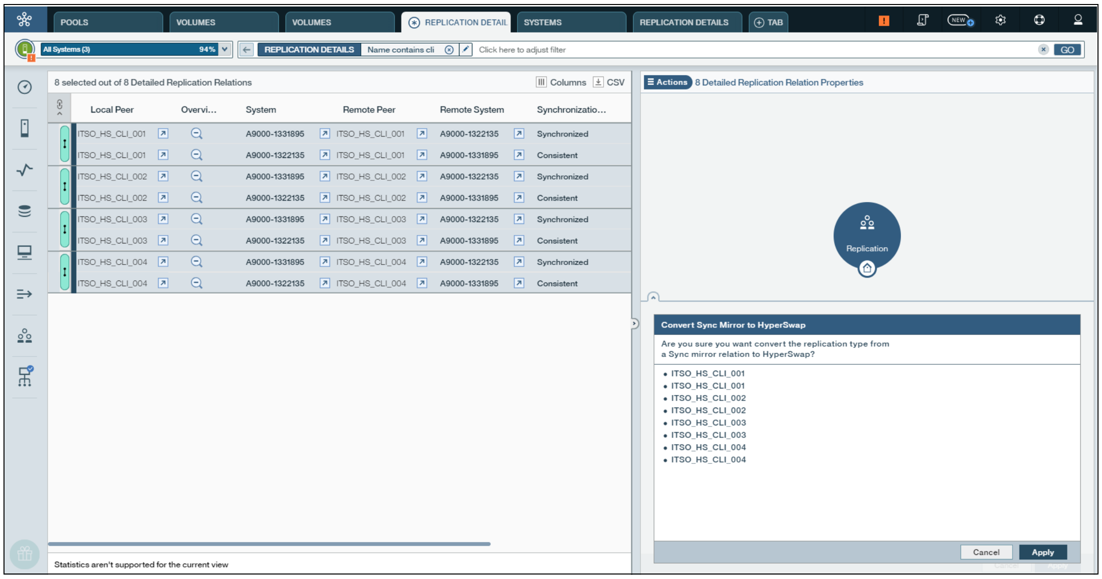
4 Detailed Replication Relation Properties

	c Mirror	Sync Mirror	HyperSwap
None	None	4	None

To view details, select a single relation, pair of relations or relations from the same Multi-site

Figure 6-2 Selecting volumes to convert to HyperSwap

Confirm the action as shown in Figure 6-3. This process is nondisruptive for I/O to the selected volumes.



8 selected out of 8 Detailed Replication Relations

Local Peer	Overvi...	System	Remote Peer	Remote System	Synchronizatio...
ITSO_HS_CLI_001		A9000-1331895	ITSO_HS_CLI_001	A9000-1322135	Synchronized
ITSO_HS_CLI_001		A9000-1322135	ITSO_HS_CLI_001	A9000-1331895	Consistent
ITSO_HS_CLI_002		A9000-1331895	ITSO_HS_CLI_002	A9000-1322135	Synchronized
ITSO_HS_CLI_002		A9000-1322135	ITSO_HS_CLI_002	A9000-1331895	Consistent
ITSO_HS_CLI_003		A9000-1331895	ITSO_HS_CLI_003	A9000-1322135	Synchronized
ITSO_HS_CLI_003		A9000-1322135	ITSO_HS_CLI_003	A9000-1331895	Consistent
ITSO_HS_CLI_004		A9000-1331895	ITSO_HS_CLI_004	A9000-1322135	Synchronized
ITSO_HS_CLI_004		A9000-1322135	ITSO_HS_CLI_004	A9000-1331895	Consistent

Statistics aren't supported for the current view

8 Detailed Replication Relation Properties

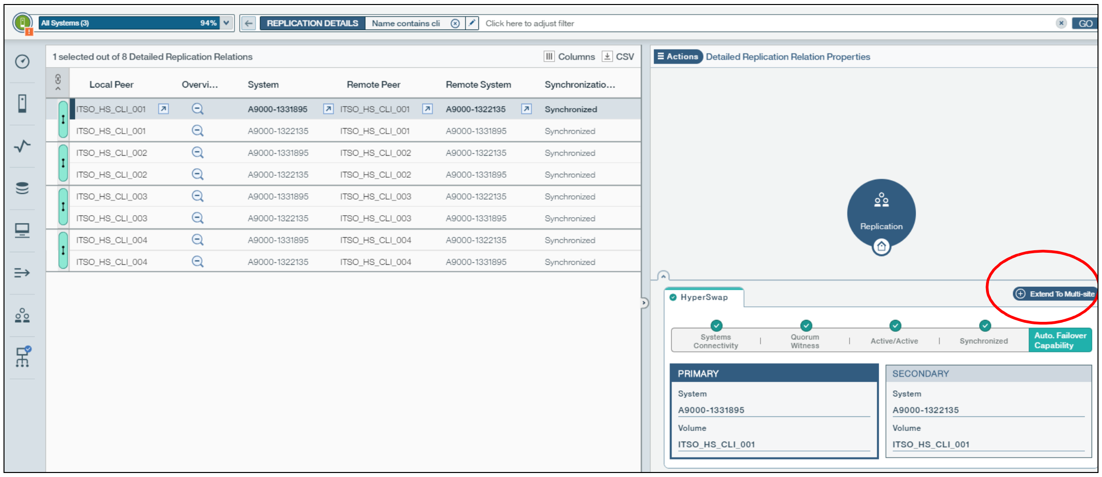
Convert Sync Mirror to HyperSwap

Are you sure you want convert the replication type from
a Sync mirror relation to HyperSwap?

- ITSO_HS_CLI_001
- ITSO_HS_CLI_001
- ITSO_HS_CLI_002
- ITSO_HS_CLI_002
- ITSO_HS_CLI_003
- ITSO_HS_CLI_003
- ITSO_HS_CLI_004
- ITSO_HS_CLI_004

Figure 6-3 Confirming the convert action

After the action is confirmed, the system runs through the conversion process. This process occurs quickly, as shown in Figure 6-4. For more information about the process, see Chapter 4, “HyperSwap implementation and usage” on page 23.



1 selected out of 8 Detailed Replication Relations

Local Peer	Overvi...	System	Remote Peer	Remote System	Synchronizatio...
ITSO_HS_CLI_001		A9000-1331895	ITSO_HS_CLI_001	A9000-1322135	Synchronized
ITSO_HS_CLI_001		A9000-1322135	ITSO_HS_CLI_001	A9000-1331895	Synchronized
ITSO_HS_CLI_002		A9000-1331895	ITSO_HS_CLI_002	A9000-1322135	Synchronized
ITSO_HS_CLI_002		A9000-1322135	ITSO_HS_CLI_002	A9000-1331895	Synchronized
ITSO_HS_CLI_003		A9000-1331895	ITSO_HS_CLI_003	A9000-1322135	Synchronized
ITSO_HS_CLI_003		A9000-1322135	ITSO_HS_CLI_003	A9000-1331895	Synchronized
ITSO_HS_CLI_004		A9000-1331895	ITSO_HS_CLI_004	A9000-1322135	Synchronized
ITSO_HS_CLI_004		A9000-1322135	ITSO_HS_CLI_004	A9000-1331895	Synchronized

Detailed Replication Relation Properties

PRIMARY	SECONDARY
System A9000-1331895	System A9000-1322135
Volume ITSO_HS_CLI_001	Volume ITSO_HS_CLI_001

Figure 6-4 Active HyperSwap relationship

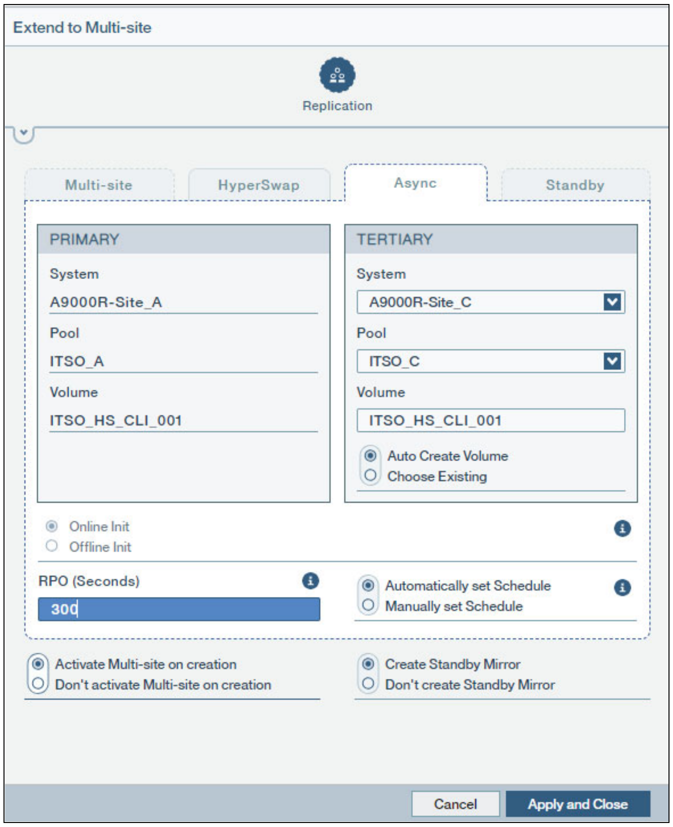
At this point, the conversion to Multi-site is possible, as described in 6.1.3, “Assembly from existing HyperSwap”.

6.1.3 Assembly from existing HyperSwap

If an existing HyperSwap relationship is the base for setting up a Multi-site configuration, the steps to extend are easily done by using the Hyper-Scale Manager (HSM) GUI or the command line.

Configuration by using the Hyper-Scale Manager GUI

With the HSM GUI, it is a matter of selecting the HyperSwap primary and clicking **Extend to Multi-site** (circled in red in Figure 6-4 on page 100). The window that is shown in Figure 6-5 opens and you must enter information about the tertiary, along with options to create the standby asynchronous relationship or not.



PRIMARY

System
A9000R-Site_A

Pool
ITSO_A

Volume
ITSO_HS_CLI_001

TERTIARY

System
A9000R-Site_C

Pool
ITSO_C

Volume
ITSO_HS_CLI_001

● Auto Create Volume
○ Choose Existing

● Online Init
○ Offline Init

RPO (Seconds)
300

● Automatically set Schedule
○ Manually set Schedule

● Activate Multi-site on creation
○ Don't activate Multi-site on creation

● Create Standby Mirror
○ Don't create Standby Mirror

Figure 6-5 Extend HyperSwap to Multi-site

Configuration by using the CLI

Creating a multi-site assembly from an existing HyperSwap configuration with the command line includes the following steps (see Example 6-1 on page 102):

1. On system A: Create an asynchronous mirror from A to C.
2. On system A: Issue the `multisite_define` command.
3. On system B: Create an asynchronous mirror from B to C.
4. On system A: Activate multi-site.
5. On system B: Register standby mirror for multi-site.

Example 6-1 Steps to extend existing HyperSwap to Multi-site by way of CLI

```
Step 1
A9000-Site-A>>mirror_create rpo=300 schedule=xiv_gui_schedule_1537991049849
vol=ITSO_HS_CLI_003 create_slave=yes part_of_multisite=yes
slave_vol=ITSO_HS_CLI_003 type=ASYNC_INTERVAL
remote_schedule=xiv_gui_schedule_1537991049849 init_type=ONLINE remote_rpo=300
target=A9000-Site-B remote_pool=ITSO_list2
Command executed successfully.

Step 2
A9000-Site-A>>multisite_define vol=ITSO_HS_CLI_003
Command executed successfully.

Step 3
A9000-Site-B>>mirror_create rpo=300 schedule=xiv_gui_schedule_1538677624024
vol=ITSO_HS_CLI_003 create_slave=no part_of_multisite=yes
slave_vol=ITSO_HS_CLI_003 type=ASYNC_INTERVAL
remote_schedule=xiv_gui_schedule_1537991049849 init_type=ONLINE remote_rpo=300
target=A9000-Site-C remote_pool=ITSO_list2
Command executed successfully.

Step 4
A9000-Site-A>>multisite_activate_async_mirror vol=ITSO_HS_CLI_003
Command executed successfully.

step 5
A9000-Site-B>>multisite_register_standby_mirror vol=ITSO_HS_CLI_003
Command executed successfully.
```

Tip: When extending to Multi-site by using the command line, you can use a schedule that meets the RPO and interval needed (which is found by issuing the **`schedule_list`** command) or creating a schedule by using the **`schedule_create`** CLI command.

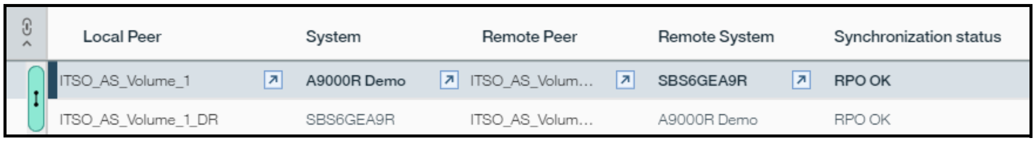
6.1.4 Assembly from existing Asynchronous relation

This section describes how to assemble a multi-site relationship from an existing asynchronous relationship between two volumes or consistency groups (CG). This task can be fulfilled with Hyper-Scale Manager (HSM) GUI or using the CLI.

Configuration by using the Hyper-Scale Manager GUI

Complete the following steps on the system that owns the Primary volume or CG in the Asynchronous relationship (system A):

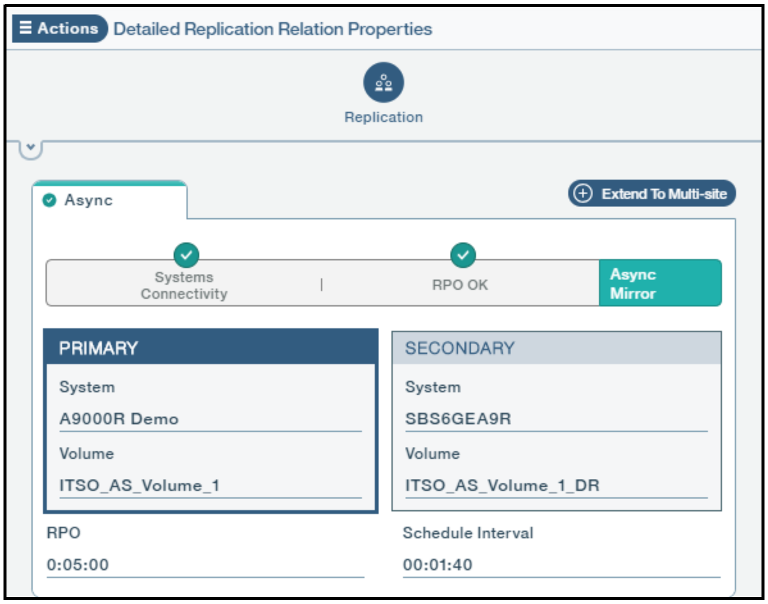
1. Navigate to the Volumes (or Consistency Groups) view and locate the required asynchronous relationship, as shown in Figure 6-6.



Local Peer	System	Remote Peer	Remote System	Synchronization status
ITSO_AS_Volume_1	A9000R Demo	ITSO_AS_Volum...	SBS6GEA9R	RPO OK
ITSO_AS_Volume_1_DR	SBS6GEA9R	ITSO_AS_Volum...	A9000R Demo	RPO OK

Figure 6-6 Select an Asynchronous relationship

2. Navigate to the selected relationship’s Replication Properties, as shown in Figure 6-7:



PRIMARY

System

A9000R Demo

Volume

ITSO_AS_Volume_1

SECONDARY

System

SBS6GEA9R

Volume

ITSO_AS_Volume_1_DR

RPO

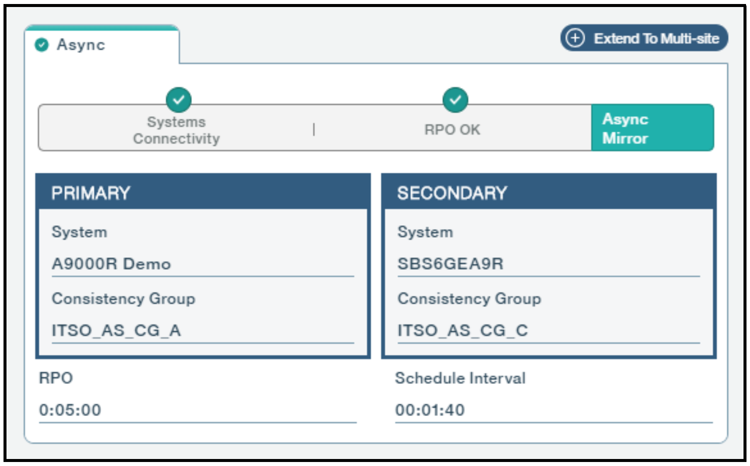
0:05:00

Schedule Interval

00:01:40

Figure 6-7 Replication properties of Asynchronous relationship between volumes

For consistency groups, this panel appears, as shown in Figure 6-8.



PRIMARY

System

A9000R Demo

Consistency Group

ITSO_AS_CG_A

SECONDARY

System

SBS6GEA9R

Consistency Group

ITSO_AS_CG_C

RPO

0:05:00

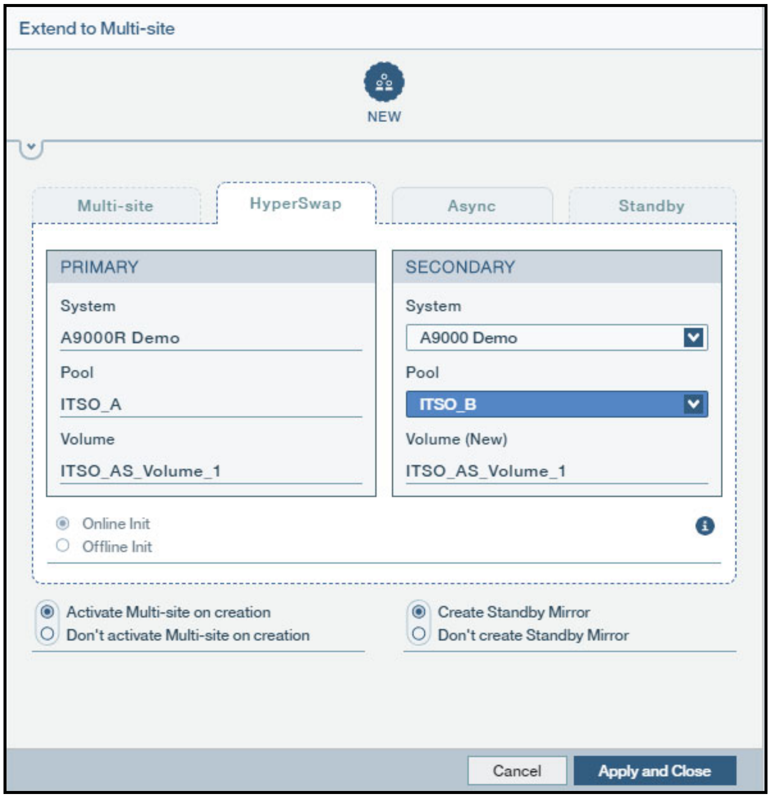
Schedule Interval

00:01:40

Figure 6-8 Replication properties of Asynchronous relationship between CGs

3. Select **Extend to Multi-site**. The Extend to Multi-site window opens, as shown in Figure 6-9 on page 104. You must create the secondary volume in the HyperSwap relationship that is part of the new Multi-site relationship. Select the system and the pool that owns the secondary volume. The volume name is identical to the primary volume’s name and cannot be changed.

   You can also decide whether to activate Multi-site on creation by selecting the appropriate option. In addition, you can choose whether to create a standby mirror as part of the new multi-site relationship. For the purposes of this example, both of these options are kept as default.



PRIMARY

System
A9000R Demo

Pool
ITSO_A

Volume
ITSO_AS_Volume_1

SECONDARY

System
A9000 Demo

Pool
ITSO_B

Volume (New)
ITSO_AS_Volume_1

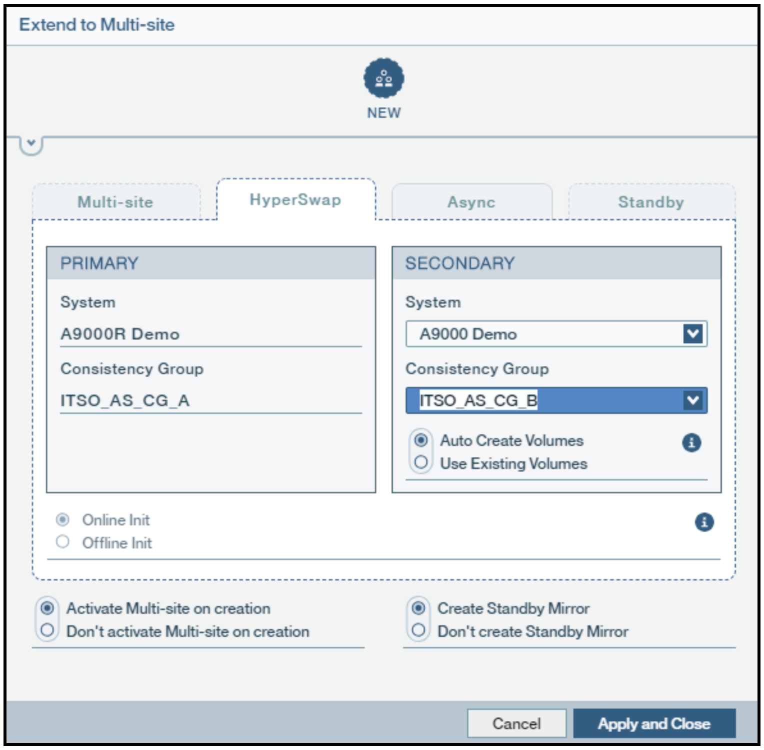
● Online Init
○ Offline Init

● Activate Multi-site on creation
○ Don't activate Multi-site on creation

● Create Standby Mirror
○ Don't create Standby Mirror

Figure 6-9 Extend to Multi-site panel for volumes

For consistency groups, the window that is shown in Figure 6-10 opens.



PRIMARY

System
A9000R Demo

Consistency Group
ITSO_AS_CG_A

SECONDARY

System
A9000 Demo

Consistency Group
ITSO_AS_CG_B

● Auto Create Volumes
○ Use Existing Volumes

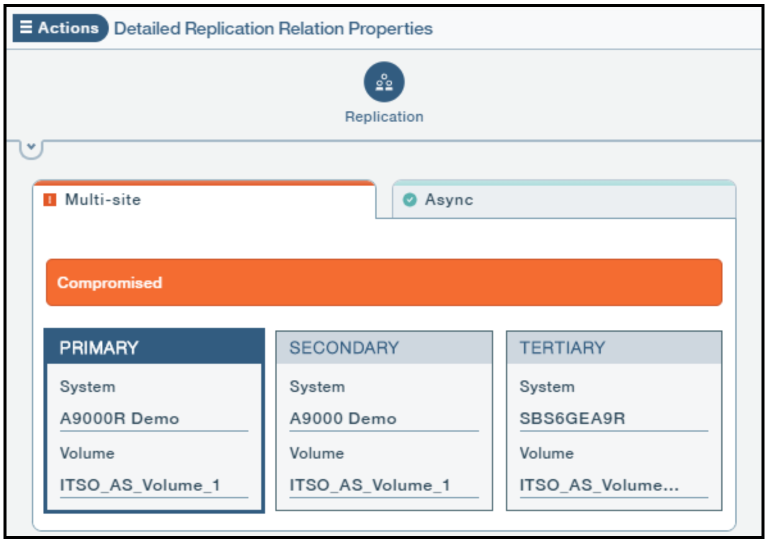
● Online Init
○ Offline Init

● Activate Multi-site on creation
○ Don't activate Multi-site on creation

● Create Standby Mirror
○ Don't create Standby Mirror

Figure 6-10 Extend to Multi-site panel for CGs

4. Select **Apply and Close** to start the multi-site relationship assembly process. The standby mirror creation is completed only after the secondary and tertiary volumes (or CGs) are fully synchronized. Until then, the multi-site relationship status remains Compromised, as shown in Figure 6-11.

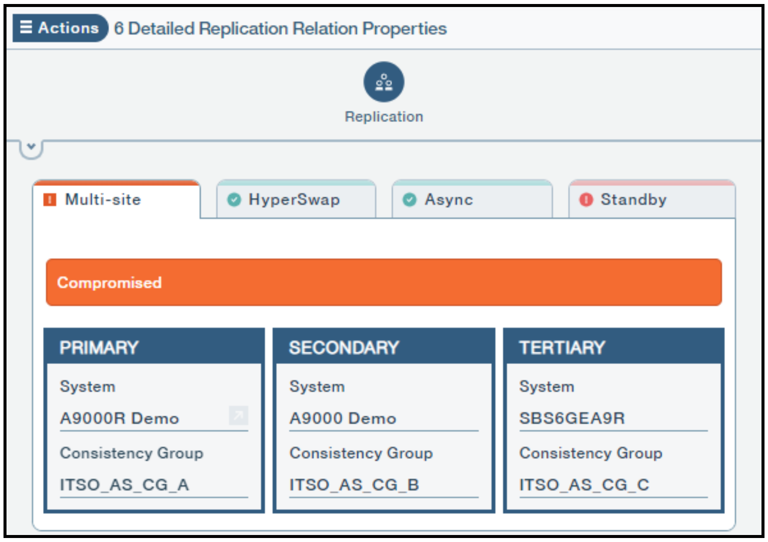


Compromised

PRIMARY	SECONDARY	TERTIARY
System	System	System
A9000R Demo	A9000 Demo	SBS6GEA9R
Volume	Volume	Volume
ITSO_AS_Volume_1	ITSO_AS_Volume_1	ITSO_AS_Volume...

Figure 6-11 Multi-site in Compromised status because of synchronization in progress (volumes)

For consistency groups, the window that is shown in Figure 6-12 opens.

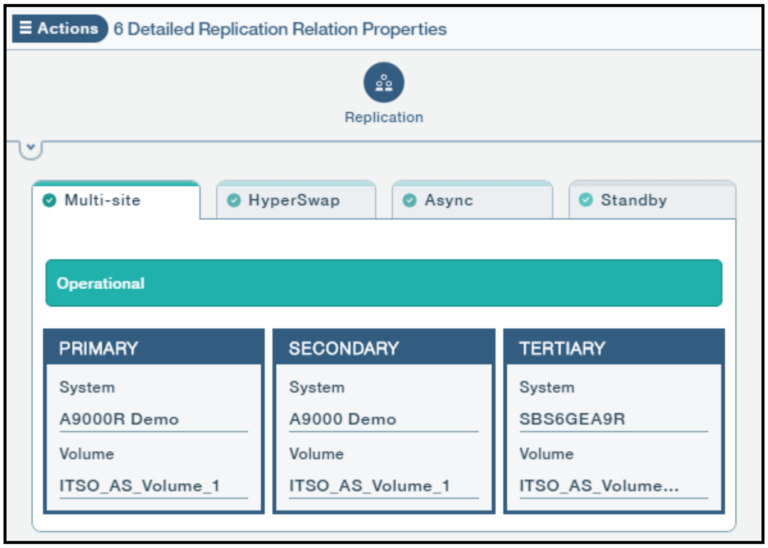


Compromised

PRIMARY	SECONDARY	TERTIARY
System	System	System
A9000R Demo	A9000 Demo	SBS6GEA9R
Consistency Group	Consistency Group	Consistency Group
ITSO_AS_CG_A	ITSO_AS_CG_B	ITSO_AS_CG_C

Figure 6-12 Multi-site in Compromised status because of synchronization in progress (CGs)

When the synchronization is complete, the multi-site relationship status changes to Operational, as shown in Figure 6-13.

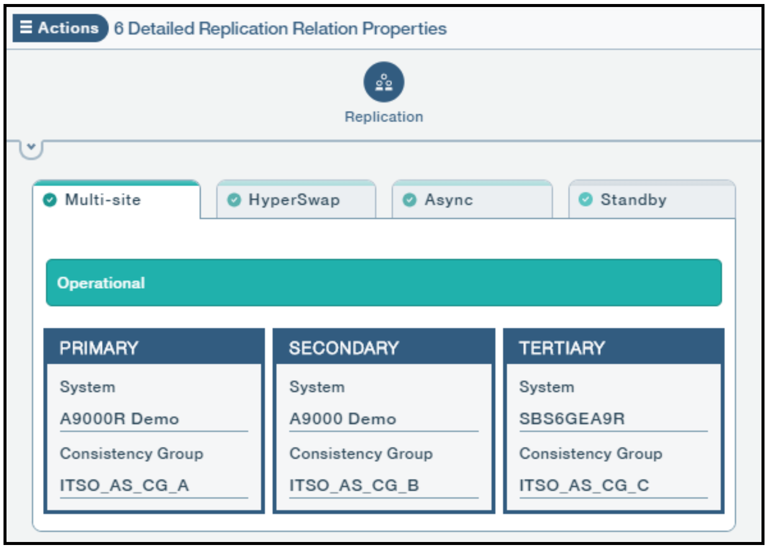


Operational

	PRIMARY	SECONDARY	TERTIARY
System	A9000R Demo	A9000 Demo	SBS6GEA9R
Volume	ITSO_AS_Volume_1	ITSO_AS_Volume_1	ITSO_AS_Volume...

Figure 6-13 Multi-site in Operational status (volumes)

For consistency groups, the window that is shown in Figure 6-14 opens.

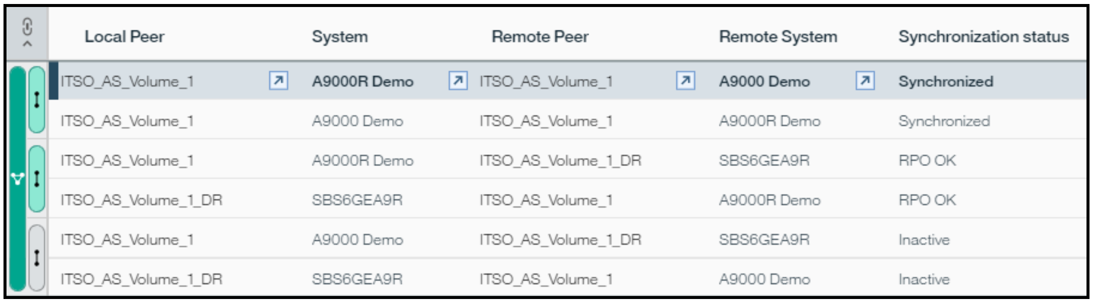


Operational

	PRIMARY	SECONDARY	TERTIARY
System	A9000R Demo	A9000 Demo	SBS6GEA9R
Consistency Group	ITSO_AS_CG_A	ITSO_AS_CG_B	ITSO_AS_CG_C

Figure 6-14 Multi-site in Operational status (CGs)

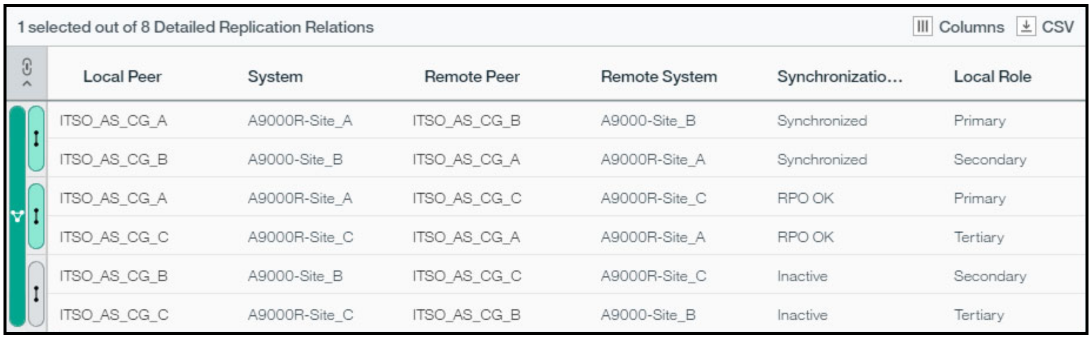
The newly created multi-site relationship appears in the Volumes view, as shown in Figure 6-15.



Local Peer	System	Remote Peer	Remote System	Synchronization status
ITSO_AS_Volume_1	A9000R Demo	ITSO_AS_Volume_1	A9000 Demo	Synchronized
ITSO_AS_Volume_1	A9000 Demo	ITSO_AS_Volume_1	A9000R Demo	Synchronized
ITSO_AS_Volume_1	A9000R Demo	ITSO_AS_Volume_1_DR	SBS6GEA9R	RPO OK
ITSO_AS_Volume_1_DR	SBS6GEA9R	ITSO_AS_Volume_1	A9000R Demo	RPO OK
ITSO_AS_Volume_1	A9000 Demo	ITSO_AS_Volume_1_DR	SBS6GEA9R	Inactive
ITSO_AS_Volume_1_DR	SBS6GEA9R	ITSO_AS_Volume_1	A9000 Demo	Inactive

Figure 6-15 New multi-site relationship in Volumes view

For CGs, the newly created multi-site relationship appears in the Consistency Groups view, as shown in Figure 6-16.



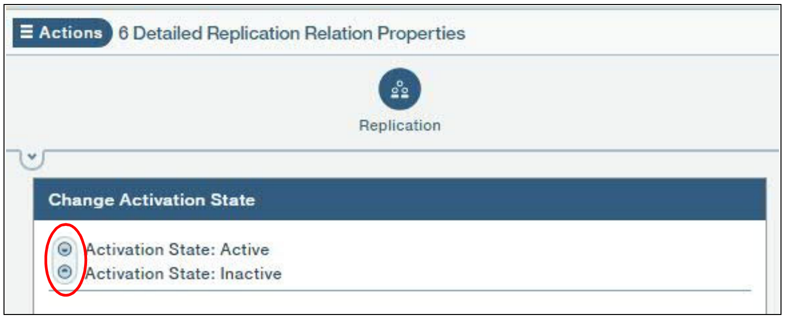
1 selected out of 8 Detailed Replication Relations

Local Peer	System	Remote Peer	Remote System	Synchronizatio...	Local Role
ITSO_AS_CG_A	A9000R-Site_A	ITSO_AS_CG_B	A9000-Site_B	Synchronized	Primary
ITSO_AS_CG_B	A9000-Site_B	ITSO_AS_CG_A	A9000R-Site_A	Synchronized	Secondary
ITSO_AS_CG_A	A9000R-Site_A	ITSO_AS_CG_C	A9000R-Site_C	RPO OK	Primary
ITSO_AS_CG_C	A9000R-Site_C	ITSO_AS_CG_A	A9000R-Site_A	RPO OK	Tertiary
ITSO_AS_CG_B	A9000-Site_B	ITSO_AS_CG_C	A9000R-Site_C	Inactive	Secondary
ITSO_AS_CG_C	A9000R-Site_C	ITSO_AS_CG_B	A9000-Site_B	Inactive	Tertiary

Figure 6-16 New multi-site relationship in Consistency Groups view

Deactivating Multi-site by using HSM GUI

When you want to deactivate a Multi-site relationship in the HSM GUI, the Change Activation State window initially displays half-filled radio buttons, as shown in Figure 6-17. You can select the **Activation State: Inactive** option to inactivate the Multi-site environment.



Change Activation State

◐ Activation State: Active
◐ Activation State: Inactive

Figure 6-17 Deactivate Multi-site

Assembly from existing asynchronous relationship by using the CLI

The assembly process to create a multi-site relationship starting from an existing asynchronous relationship by using the CLI is similar to the one process that is shown for the GUI. However, this process is broken down into more granular steps to create all of the required components that eventually form the multi-site relationship.

In this scenario, the starting point is an existing asynchronous relationship between two volumes or two Consistency Groups (holding multiple volumes) that you want to extend to multi-site by including a HyperSwap leg between systems A and B and a Standby Asynchronous leg between systems B and C.

We can check the mirroring details of the system that is hosting the Primary volumes (we assume this is system A in our example from now on, as shown in Figure 6-1 on page 98) by using the **`mirror_list`** command. Example 6-2 shows the output for a single relationship between volumes.

Example 6-2 Mirroring relationship details for a single volume

```
A9000R-Site_A>>mirror_list
Name             Mirror Type     Mirror Object  Remote System  Remote Peer
ITSO_AS_Volume_1 async_interval  Volume         A9000R-Site_C  ITSO_AS_Volume_1_DR
```

Example 6-3 shows mirroring between Consistency Groups across systems. In this example, the Primary volume `ITSO_AS_Volume_1` on system A is in an asynchronous mirroring relation with the Tertiary volume `ITSO_AS_Volume_1_DR` on system C.

Example 6-3 Mirroring relationship details for a Consistency Group

```
A9000R-Site_A>>mirror_list
Name               Mirror Type      Mirror Object   Remote System   Remote Peer
ITSO_AS_CG_A       async_interval   CG              A9000R-Site_C   ITSO_AS_CG_C
```

In this example, the Primary volumes in Consistency Group `ITSO_AS_CG_A` on System A are asynchronously mirrored to their Tertiary counterparts in Consistency Group `ITSO_AS_CG_C` on system C. Complete the following steps:

1. Create the Secondary volume(s) on system B that receives data from the HyperSwap leg of the multi-site relationship.

   For each volume you are operating on (in a single relationship or as member of a Consistency Group), create a volume on system B with the same name and size as the Primary volume. Example 6-4 shows the volume creation.

Example 6-4 Creating the Secondary volume

```
A9000-Site_B>>vol_create vol=ITSO_AS_Volume_1 pool=ITSO_B size=10
Command executed successfully.
```

   If you are working with Consistency Groups, you also must create a Consistency Group on the system that hosts the Secondary volumes in the HyperSwap level of the multi-site relationship, as shown in Example 6-5.

Example 6-5 Creating the Secondary Consistency Group

```
A9000-Site_B>>cg_create cg=ITSO_AS_CG_B pool=ITSO_B
Command executed successfully.
```

Next, you must create a HyperSwap relationship between the Consistency Group that contains the Primary volumes and the newly created one on system B that contains the Secondary volumes.

2. On System A, issue the **`ha_create`** command and specify the Primary (`cg`) and Secondary (`slave_cg`) Consistency Group names parameters, as shown in Example 6-6. If you are working with a single relationship, you can skip this step.

Example 6-6 Creating a HyperSwap relation with multi-site feature

```
A9000R-Site_A>>ha_create cg="ITSO_AS_CG_A" slave_cg="ITSO_AS_CG_B"
              part_of_multisite="yes" target="A9000-Site_B"
Command executed successfully.
```

3. Create the HyperSwap leg of the multi-site relationship.

   On the system that is hosting the Primary volumes, issue the **`ha_create`** command, as shown in Example 6-7, specifying the volume name, the target system, and the pool. Notice the use of the `part_of_multisite` parameter. If you are working with Consistency Groups, then you must perform this for each pair of Primary-Secondary volumes that are involved. The Secondary volumes automatically are added to the Secondary Consistency Group.

Example 6-7 Creating the HyperSwap leg of the multi-site relation

```
A9000R-Site_A>>ha_create vol="ITSO_AS_Volume_1" remote_pool="ITSO_B"
              create_slave="no" part_of_multisite="yes" init_type="ONLINE"
              target="A9000-Site_B"
Command executed successfully.
```

4. On system A, define the new multi-site relation by using the **`multisite_define`** command. Follow Example 6-8 for a single relationship and Example 6-9 for a Consistency Group.

Example 6-8 Defining the new multi-site relation between volumes

```
A9000R-Site_A>>multisite_define vol="ITSO_AS_Volume_1"
Command executed successfully.
```

Example 6-9 Defining the new multi-site relation between Consistency Groups

```
A9000R-Site_A>>multisite_define cg="ITSO_AS_CG_A"
Command executed successfully.
```

5. Create a sync job schedule on the system that is hosting the Secondary volumes (system B) and the system that is hosting the Tertiary volumes (system C). The schedules are used by the Standby Asynchronous relations that we are going to create next. Example 6-10 and Example 6-11 on page 110 show the relevant steps. The schedule name and interval must be the same across both systems.

Example 6-10 Creating a sync schedule for the Standby Asynchronous relation on the system hosting the Secondary volume

```
A9000-Site_B>>schedule_create schedule="ITSO_AS_Schedule_BC" interval="00:01:40"
Command executed successfully.
```

Example 6-11 Creating a sync schedule for the Standby Asynchronous relation on the system hosting the Secondary volume

```
A9000R-Site_C>>schedule_create schedule="ITSO_AS_Schedule_BC" interval="00:01:40"
Command executed successfully.
```

> **Note:** You can only choose between a few predetermined interval values. The interval value must be smaller than your wanted RPO. For more information, see *IBM FlashSystem A9000 Version 12.3.0 Command-Line Interface (CLI) Reference Guide*, SC27-8559.

6. On System B, establish the Standby Asynchronous relationships between the Consistency Group that is holding the Secondary Volumes and the one on System C that is holding the Tertiary volumes. The target Consistency Group must be the one that is in use in the Active Asynchronous relationship between System A and System C, as shown in Example 6-12. Specify your wanted RPO and make sure to use the schedules that are defined on both systems in the previous step.

   If you are working with a single relationship, you can skip this step.

Example 6-12 Creating the standby asynchronous relationship

```
A9000-Site_B>>mirror_create slave_cg="ITSO_AS_CG_C" rpo="300"
              schedule="ITSO_AS_Schedule_BC" cg="ITSO_AS_CG_B"
              part_of_multisite="yes" type="ASYNC_INTERVAL"
              remote_schedule="ITSO_AS_Schedule_BC" remote_rpo="300"
              target="A9000R-Site_C"
Command executed successfully.
```

7. The Standby Asynchronous relations between the Secondary and the Tertiary volumes can now be created. On the system that is hosting the Secondary volume, issue the **`mirror_create`** command and specify your wanted RPO, the local and remote sync schedule that was defined in step 5, the source volume that was defined in step 1 (Secondary) and the target volume (Tertiary). The full list of parameters that must be used in this step are shown in Example 6-13.

Example 6-13 Creating the Asynchronous mirror between the Secondary and Tertiary volumes

```
A9000-Site_B>>mirror_create rpo="300" schedule="ITSO_AS_Schedule_BC"
            vol="ITSO_AS_Volume_1" create_slave="no" part_of_multisite="yes"
            slave_vol="ITSO_AS_Volume_1_DR" type="ASYNC_INTERVAL"
            remote_schedule="ITSO_AS_Schedule_BC" init_type="ONLINE"
            remote_rpo="300" target="A9000R-Site_C"
Command executed successfully.
```

8. On the system that is hosting the Secondary volume, issue the **`multisite_register_standby_mirror`** and specify the name of the Secondary volume (for a single relation, see Example 6-14) or the name of the Secondary Consistency Group (see Example 6-15 on page 110).

Example 6-14 Register the multi-site Standby mirror volumes relationship

```
A9000-Site_B>>multisite_register_standby_mirror vol="ITSO_AS_Volume_1"
Command executed successfully.
```

Example 6-15 Register the multi-site Standby mirror Consistency Group relationship

```
A9000-Site_B>>multisite_register_standby_mirror cg="ITSO_AS_CG_B"
```

```
Command executed successfully.
```

9. Activate the HyperSwap relationships between the Primary and Secondary volumes (for a single volume, see Example 6-16) or the Primary and Secondary Consistency Groups (see Example 6-17) by issuing the **`ha_activate`** command on system A.

Example 6-16 Enabling the HyperSwap leg of the multi-site relationship for a single volume

```
A9000R-Site_A>>ha_activate vol="ITSO_AS_Volume_1" target="A9000 Demo"
Command executed successfully.
```

Example 6-17 Enabling the HyperSwap leg of the multi-site relationship for a Consistency Groups

```
A9000R-Site_A>>ha_activate cg="ITSO_AS_CG_A" target="A9000-Site_B"
Command executed successfully.
```

You can verify the state of the newly created multi-site relation by using the **`multisite_list`** command (see Example 6-18) for a single relationship and for a Consistency Group (in Example 6-19).

Example 6-18 Listing multi-site relations details for a single volume

```
A9000R-Site_A>>multisite_list
Name             State       MultisiteStandby  Master  SMaster      Slave
ITSO_AS_Volume_1 Operational Up                Local   A9000-Site_B A9000R-Site_C
```

Example 6-19 Listing multi-site relations details for Consistency Groups

```
A9000R Demo>>multisite_list
Name          State        MultisiteStandby  Master  SMaster       Slave
ITSO_AS_CG_A  Operational  Up                Local   A9000-Site_B  A9000R-Site_C
```

6.1.5 Assembly with Consistency Group

Creating a Multi-site HA/DR configuration can be done by using a new empty consistency group (CG). The process is similar to creating a mirrored consistency group for HyperSwap or replication and converting an existing mirrored or HyperSwap consistency group to Multi-site. This section reviews both options.

New CG

The steps that are involved in creating a Multi-site relationship for a CG are similar to the steps that are required for a new HyperSwap or Asynchronous relationship. You must start with an empty CG and configure an Asynchronous or HyperSwap relationship. After the empty CG Multi-site setup is operational, you can add Multi-site volumes to the Multi-site CG.

Existing CG in a HyperSwap relationship

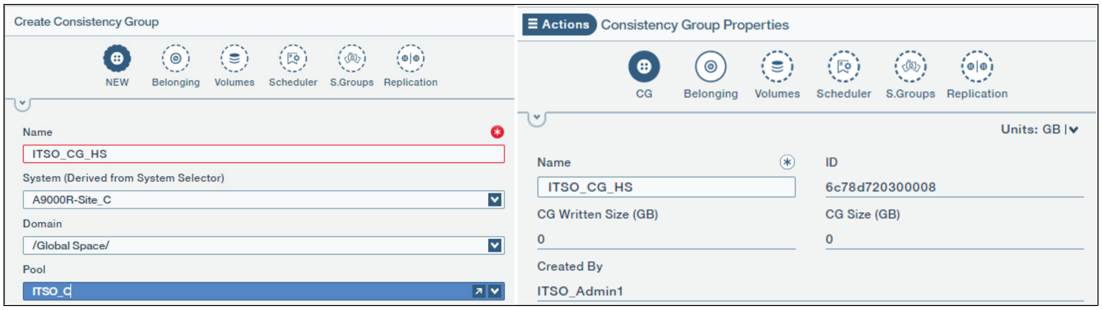
Creating a Multi-site relationship with an existing populated CG in a HyperSwap relationship includes some extra steps. In this case, Site A is the HyperSwap primary and Site B is the HyperSwap secondary. We are starting with a HyperSwap CG that is named ITSO_CG_HS.

> **Important:** Ensure that a Network Time Protocol (NTP) server is configured on all systems that are involved in Multi-site to avoid issues with potential failures because of time differences.

Extending to Multi-site in the Hyper-Scale Manager

Complete the following steps to extend this HyperSwap relationship into a Multi-site:

1. Create a local consistency group at Site C, as shown is Figure 6-18.



Name
ITSO_CG_HS

System (Derived from System Selector)
A9000R-Site_C

Domain
/Global Space/

Pool
ITSO_C

Name
ITSO_CG_HS

ID
6c78d720300008

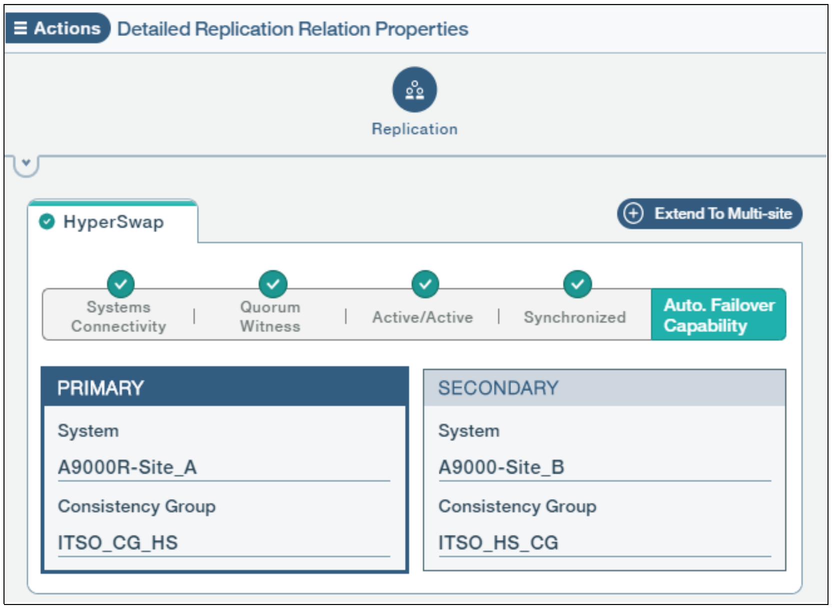
CG Written Size (GB)
0

CG Size (GB)
0

Created By
ITSO_Admin1

Figure 6-18 Create a new consistency group on C

2. Under the **Remote Views** → **Replication Details** view, select the HyperSwap consistency (in this case ITSO_CG_HS) and select **Extend to Multi-site**. This option is shown in Figure 6-19.



PRIMARY

System
A9000R-Site_A

Consistency Group
ITSO_CG_HS

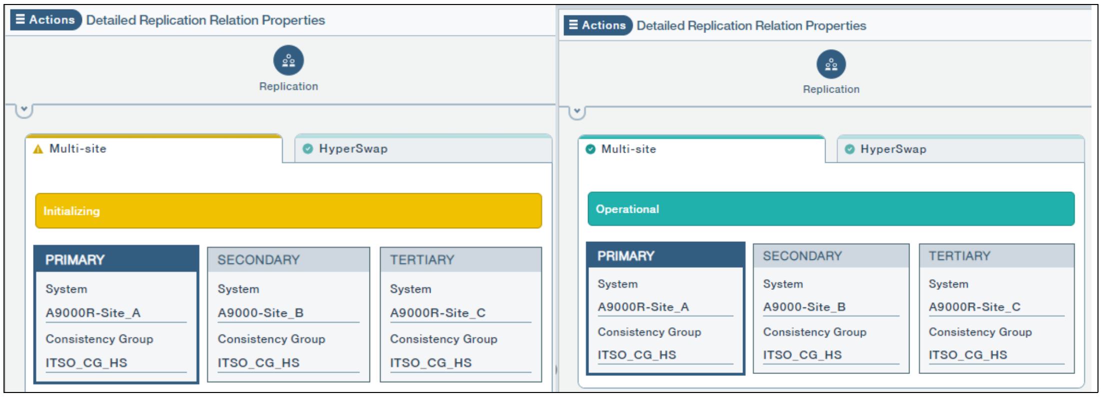
SECONDARY

System
A9000-Site_B

Consistency Group
ITSO_HS_CG

Figure 6-19 Extend to Multi-site option in the HSM GUI

3. At this point, the asynchronous active and standby replication links are created. The active asynchronous link starts with initialization and then Multi-site moves to an operational state (see Figure 6-20).



PRIMARY
System
A9000R-Site_A
Consistency Group
ITSO_CG_HS

SECONDARY
System
A9000-Site_B
Consistency Group
ITSO_CG_HS

TERTIARY
System
A9000R-Site_C
Consistency Group
ITSO_CG_HS

Figure 6-20 Multi-site changing states

Extending to Multi-site by using the CLI

The process for creating a Multi-site consistency group requires a few extra steps that are completed in the background when the HSM GUI is used. Complete the following steps:

1. Create a consistency group on site C.
2. Create a schedule, or find an existing schedule that meets the requirement on site A and B (the HyperSwap pair). In our example, we create the schedule.
3. Create the asynchronous consistency group mirror between sites A and C (this group becomes the active link).
4. Create the asynchronous mirror between the volumes that is in the consistency group on A and C.
5. Define the Multi-site relationship for the consistency group on A.
6. Create the standby asynchronous consistency group mirror between B and C.
7. Register the standby asynchronous mirror on A.
8. Activate the Multi-site asynchronous consistency group relationship on A.

After the empty Multi-site consistency group reaches an operational state, the Multi-site volumes can be added to the consistency group by using the command **cg_add_vol** (see Example 6-20).

Example 6-20 Setup Multi-site with command line

```
A9000R-Site_C>>cg_create cg=ITSO_HS_CLI pool=ITSO_C
Command executed successfully

A9000R-Site_A>>schedule_create schedule=xiv_gui_schedule_1538587261921
interval=00:01:40
Command executed successfully.

A9000-Site_B>>schedule_create schedule=xiv_gui_schedule_1538587261921
interval=00:01:40
Command executed successfully.

A9000R-Site_A>>mirror_create slave_cg=ITSO_HS_CLI rpo=300
schedule=xiv_gui_schedule_1538587261921 cg=ITSO_HS_CLI part_of_multisite=yes
```

```
type=ASYNC_INTERVAL  remote_rpo=300 remote_schedule=xiv_gui_schedule_1538587261921
target=A9000R-Site_C
Command executed successfully

A9000R-Site_A>>mirror_create rpo=300 schedule=xiv_gui_schedule_1538587261921
vol=ITSO_HS_CLI_003 remote_pool=ITSO_C create_slave=yes part_of_multisite=yes
slave_vol=ITSO_HS_CLI_003 type=ASYNC_INTERVAL
remote_schedule=xiv_gui_schedule_1538587261921 init_type=ONLINE remote_rpo=300
target=A9000R-Site_C
Command executed successfully

A9000R-Site_A>>multisite_define cg=ITSO_HS_CLI
Command executed successfully

A9000-Site_B>>mirror_create slave_cg=ITSO_HS_CLI rpo=300
schedule=xiv_gui_schedule_1538587261921 cg=ITSO_HS_CLI part_of_multisite=yes
type=ASYNC_INTERVAL  remote_rpo=300 remote_schedule=xiv_gui_schedule_1538587261921
target=A9000R-Site_C
Command executed successfully

A9000R-Site_A>>multisite_register_standby_mirror cg=ITSO_HS_CLI
Command executed successfully

A9000R-Site_A>>multisite_activate_async_mirror  cg=ITSO_HS_CLI
Command executed successfully

A9000R-Site_A>>cg_add_vol cg=ITSO_CG vol=ITSO_CG_MS_001
Command executed successfully
```

> **Important:** Without registering the standby mirror, no automatic activation occurs in the case of a HyperSwap failover.

When adding a mirrored volume to a mirrored consistency group by using the HSM GUI or command line, you might need to retry the operation if a sync job is in progress, which makes the last sync time different. If needed, retry the command. An example of this failure with the command line is shown in Example 6-21.

Example 6-21   Adding a mirrored volume to a CG might fail

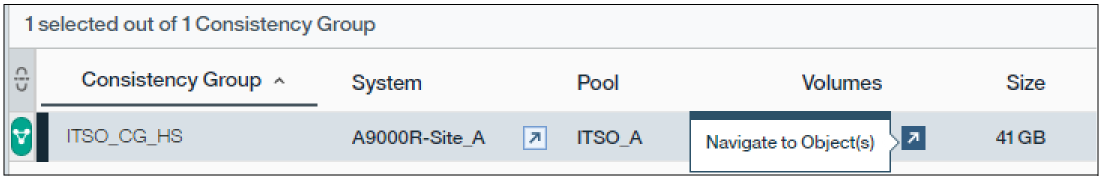
```
A9000R-Site_A>>cg_add_vol cg=ITSO_CG vol=ITSO_CG_MS_001
Command executed successfully
```

6.1.6 Leaving CG with active HyperSwap or asynchronous relation

It is possible to remove an active Multi-site volume from a Multi-site consistency group and maintain the active relationships.

Using the HSM GUI

Removing a Multi-site volume from a Multi-site consistency group is a simple process. Select the volumes in the Volumes view or link to them from the consistency group as shown in Figure 6-21 and then, select the volumes to remove.



1 selected out of 1 Consistency Group

Consistency Group ^	System	Pool	Volumes	Size
ITSO_CG_HS	A9000R-Site_A	ITSO_A	Navigate to Object(s)	41 GB

Figure 6-21 Navigate to the volumes in the consistency group

After the volume is selected, right-click (or open the actions menu) and select **Consistency Group** → **Remove from Group**, as shown in Figure 6-22.

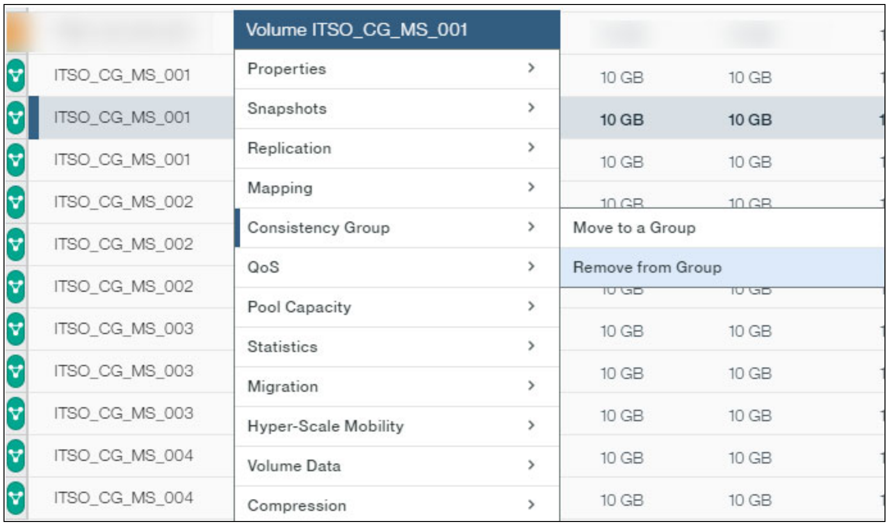
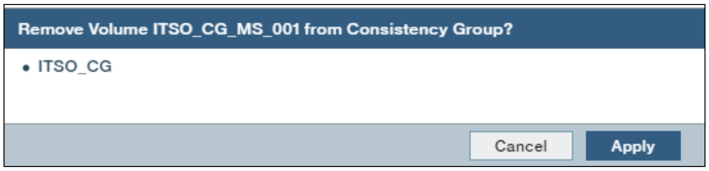


Figure 6-22 Select Remove from Group in the Actions menu

You are prompted to confirm the removal of the volumes from the consistency group. Click **Apply** to confirm, as shown in Figure 6-23. This process is the same process that is used for replicated or simplex volumes and consistency groups.

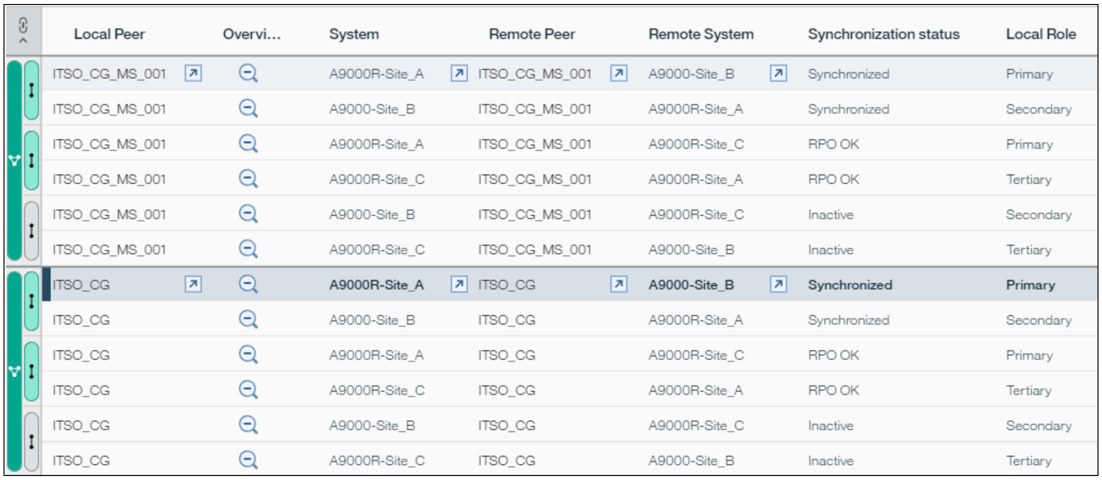


Remove Volume ITSO_CG_MS_001 from Consistency Group?

- ITSO_CG

Figure 6-23 Confirm removal from consistency group

After this process is complete, you can see the Multi-site consistency group and volumes in the Replication Details View, as shown in Figure 6-24.



Local Peer	Overvi...	System	Remote Peer	Remote System	Synchronization status	Local Role
ITSO_CG_MS_001	🔍	A9000R-Site_A	ITSO_CG_MS_001	A9000-Site_B	Synchronized	Primary
ITSO_CG_MS_001	🔍	A9000-Site_B	ITSO_CG_MS_001	A9000R-Site_A	Synchronized	Secondary
ITSO_CG_MS_001	🔍	A9000R-Site_A	ITSO_CG_MS_001	A9000R-Site_C	RPO OK	Primary
ITSO_CG_MS_001	🔍	A9000R-Site_C	ITSO_CG_MS_001	A9000R-Site_A	RPO OK	Tertiary
ITSO_CG_MS_001	🔍	A9000-Site_B	ITSO_CG_MS_001	A9000R-Site_C	Inactive	Secondary
ITSO_CG_MS_001	🔍	A9000R-Site_C	ITSO_CG_MS_001	A9000-Site_B	Inactive	Tertiary
ITSO_CG	🔍	A9000R-Site_A	ITSO_CG	A9000-Site_B	Synchronized	Primary
ITSO_CG	🔍	A9000-Site_B	ITSO_CG	A9000R-Site_A	Synchronized	Secondary
ITSO_CG	🔍	A9000R-Site_A	ITSO_CG	A9000R-Site_C	RPO OK	Primary
ITSO_CG	🔍	A9000R-Site_C	ITSO_CG	A9000R-Site_A	RPO OK	Tertiary
ITSO_CG	🔍	A9000-Site_B	ITSO_CG	A9000R-Site_C	Inactive	Secondary
ITSO_CG	🔍	A9000R-Site_C	ITSO_CG	A9000-Site_B	Inactive	Tertiary

Figure 6-24 Multi-site relationship persists for volume removed from CG

Using command line

The command for removing a Multi-site volume from a Multi-site consistency group is the same as removing a simplex volume from a consistency group (see Example 6-22).

Example 6-22 Removing a Multi-site volume from Multi-site consistency group

```
A9000R-Site_A>>cg_remove_vol vol=ITSO_CG_HS_004
Command executed successfully
```

6.2 Data migration

One possible scenario that is enabled by the use of Multi-site relationships is to perform Data Migration. Before the introduction of Multi-site, the existing Data Migration solutions for IBM FlashSystem A9000 and A9000R all involved a maximum of two systems; meaning that any existing replication relationship that involves the volumes that are to be migrated must be deleted in advance.

This situation is especially annoying in the case of the migration of a volume that has an Asynchronous mirroring relationship towards a DR system. Migration efforts require interrupting the DR mirroring for the entire process, which can take some time.

Multi-site relationships allow for three systems to replicate data from the same volume, which enables the migration of a volume from one system to another, while keeping data replicated asynchronously to a third system.

Figure 6-25 on page 117 shows the initial configuration. In this example, we want to migrate a volume from System A to System B, online (that is, without causing loss of access to the hosts that are using it) and retaining Asynchronous DR mirroring towards System C.

To proceed, the three systems that are involved must conform to the Multi-site implementation requirements. For more information, see Chapter 2, “Multi-site HA/DR solution for FlashSystem A9000 and A9000R” on page 13.

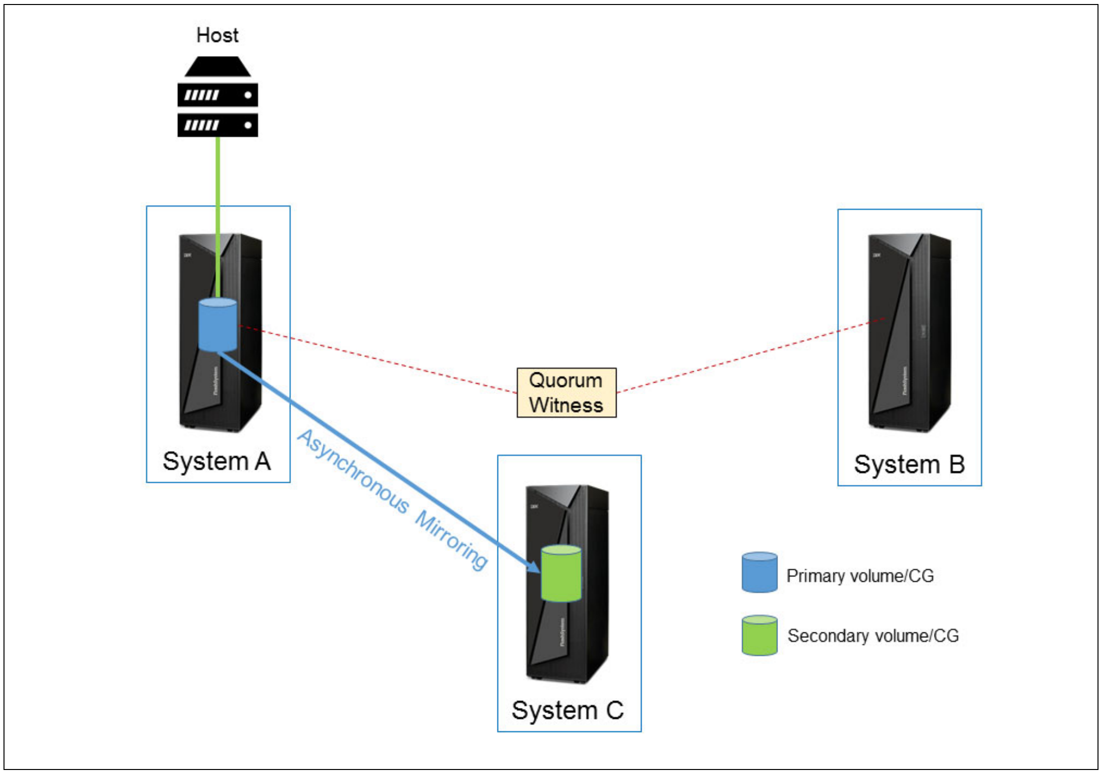


Figure 6-25 Initial configuration for the Data Migration scenario

The first step in the Data Migration procedure is to extend the existing Asynchronous mirroring relationship between System A and System C to a Multi-site relationship, including our target System, B. For more information about this process, see 6.1.4, “Assembly from existing Asynchronous relation” on page 102.

The Secondary volume in the Asynchronous relationship between A and C now becomes the Tertiary volume in the Multi-site relationship. The Primary volume retains its role, while a new Secondary volume is created on System B.

After the extension, data from the Primary volume is replicated to the Secondary by way of the HyperSwap leg of the Multi-site relationship, while a new Standby Asynchronous relationship is created between B and C. At the end of this step, the configuration resembles the configuration that shown in Figure 6-26 on page 118.

Host I/O is still served by the Primary volume on A. However, after mapping the Secondary volume in B to the hosts, more Unoptimized paths are available.

Important: When mapping the Secondary volumes, ensure to use the same LUN ID assignments that are used by the corresponding Primary volumes to avoid host access errors.

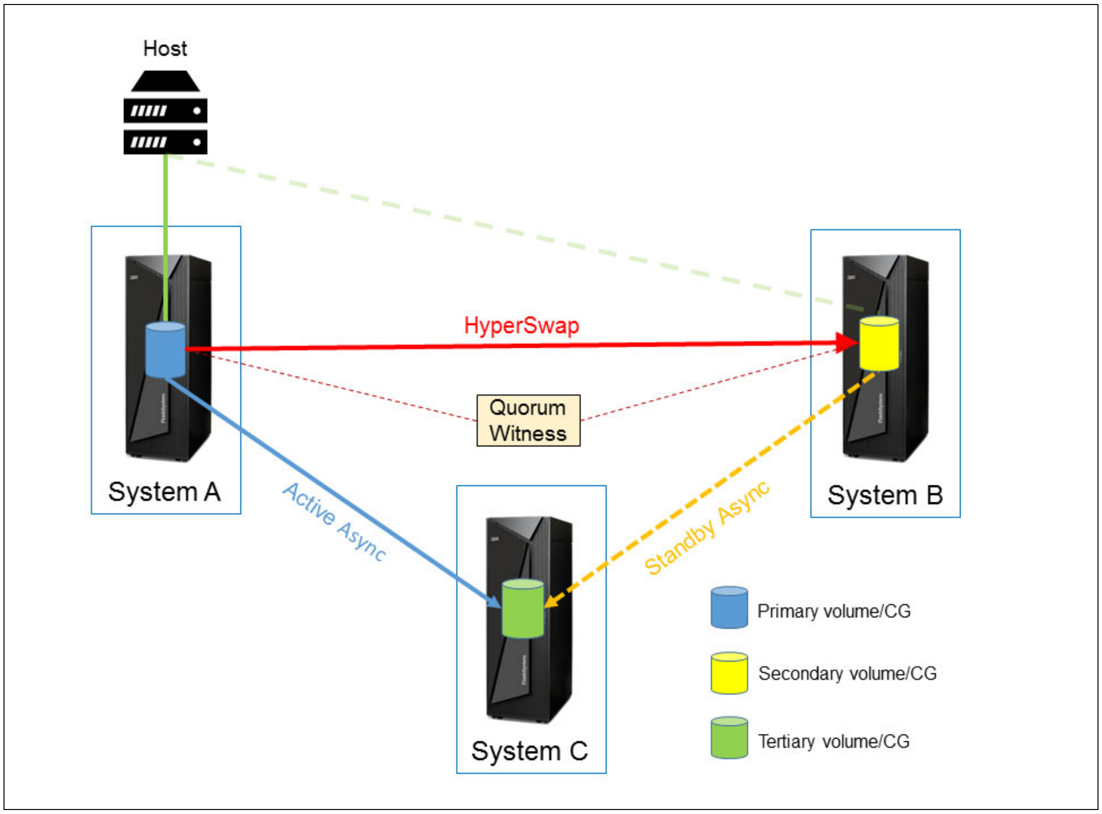
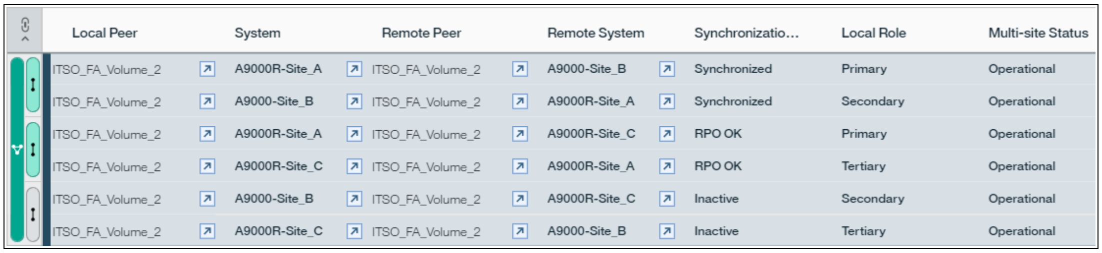


Figure 6-26 Configuration after the extension to a Multi-site relationship

Wait for the Multi-site relationship to be fully synchronized and operational, as shown in Figure 6-27.



Local Peer	System	Remote Peer	Remote System	Synchronizatio...	Local Role	Multi-site Status
ITSO_FA_Volume_2	A9000R-Site_A	ITSO_FA_Volume_2	A9000-Site_B	Synchronized	Primary	Operational
ITSO_FA_Volume_2	A9000-Site_B	ITSO_FA_Volume_2	A9000R-Site_A	Synchronized	Secondary	Operational
ITSO_FA_Volume_2	A9000R-Site_A	ITSO_FA_Volume_2	A9000R-Site_C	RPO OK	Primary	Operational
ITSO_FA_Volume_2	A9000R-Site_C	ITSO_FA_Volume_2	A9000R-Site_A	RPO OK	Tertiary	Operational
ITSO_FA_Volume_2	A9000-Site_B	ITSO_FA_Volume_2	A9000R-Site_C	Inactive	Secondary	Operational
ITSO_FA_Volume_2	A9000R-Site_C	ITSO_FA_Volume_2	A9000-Site_B	Inactive	Tertiary	Operational

Figure 6-27 Operational Multi-site relationship details

Primary and Secondary roles can now be switched to indicate to the hosts to start using paths towards System B to service I/Os, and activating the B-C Asynchronous relationship. To do so, start by deactivating the A-C (Active) Asynchronous relationship between A and C. From the Actions menu, click **Configuration** → **Change Activation** (see Figure 6-28).

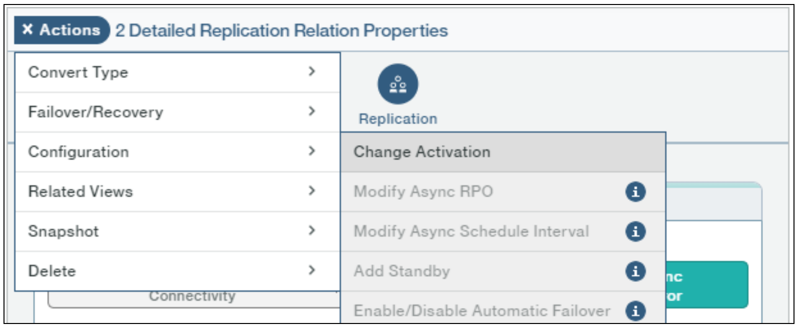
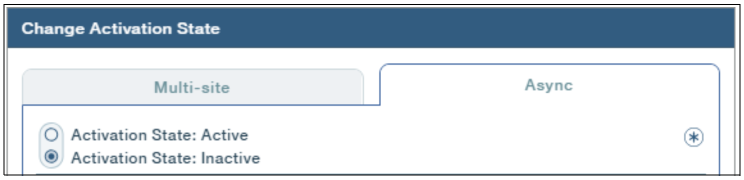


Figure 6-28 Action menu to deactivate the Active Asynchronous relationship

Select **Activation State: Inactive** and confirm by clicking **Apply**, as shown in Figure 6-29.



Change Activation State

Multi-site | Async

○ Activation State: Active
● Activation State: Inactive

Figure 6-29 Deactivating the Asynchronous relationship

Switch the Primary and Secondary roles by selecting from the Actions menu **Failover/Recovery** → **Switch Roles**, as shown in Figure 6-30.

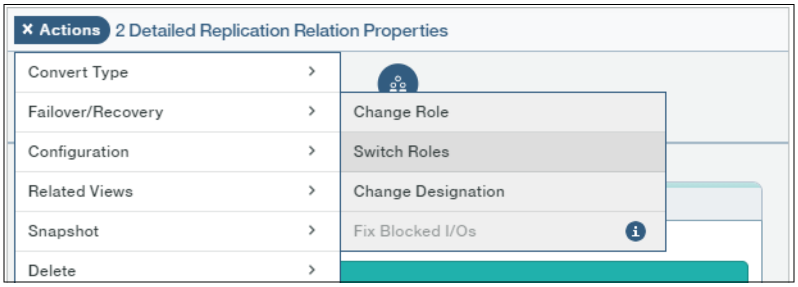
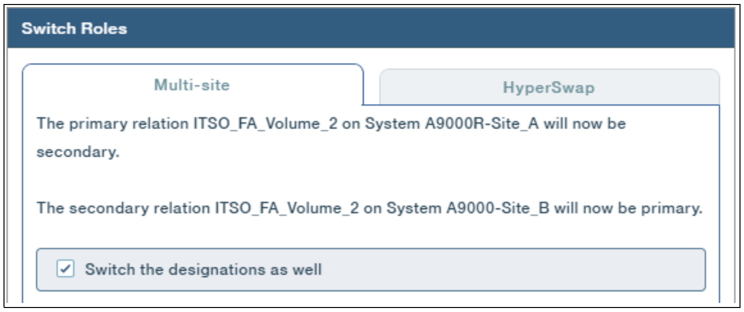


Figure 6-30 Switch Roles action menu

Ensure to select **Switch the designations as well** in the next window. Then, confirm by clicking **Apply** (see Figure 6-31).



Switch Roles

Multi-site | HyperSwap

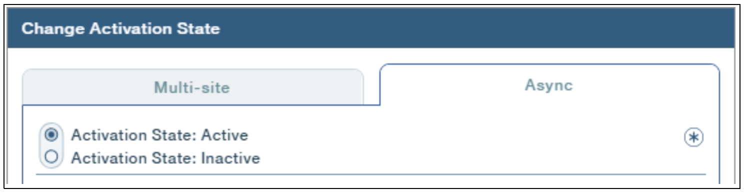
The primary relation ITSO_FA_Volume_2 on System A9000R-Site_A will now be secondary.

The secondary relation ITSO_FA_Volume_2 on System A9000-Site_B will now be primary.

☑ Switch the designations as well

Figure 6-31 Multi-site role switching dialog

Finally, activate the B to C Asynchronous relationship to bring the Multi-site relationship back to the Operational state (see Figure 6-32).


Change Activation State

● Activation State: Active
○ Activation State: Inactive

Figure 6-32 Activating the Asynchronous relationship between System B and System C

The new configuration is shown in Figure 6-33.

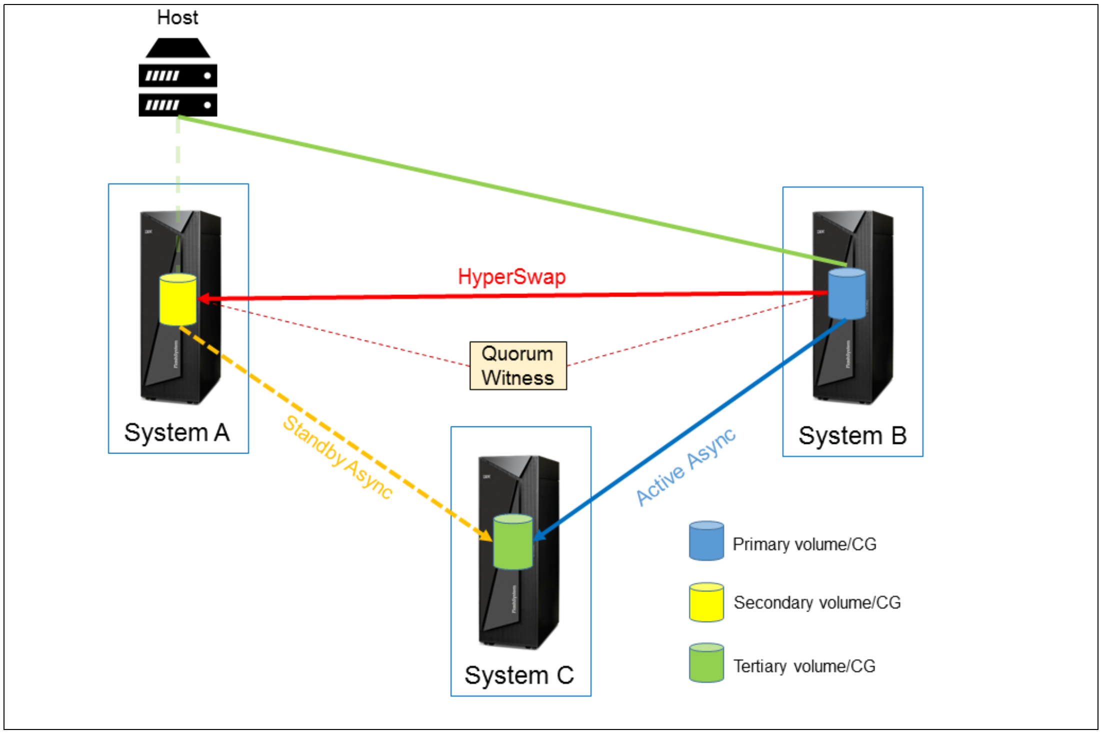
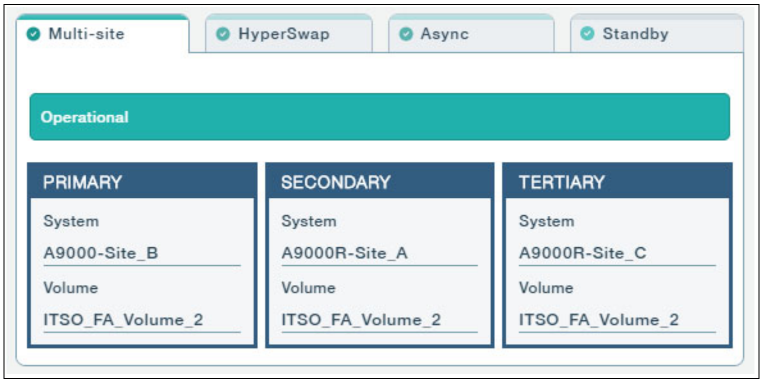


Figure 6-33 Multi-site configuration after the roles switch

Figure 6-34 shows the HSM GUI representation.



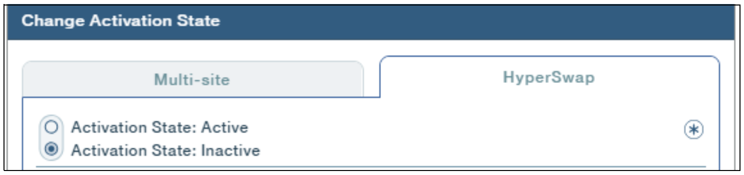
Operational

PRIMARY	SECONDARY	TERTIARY
System	System	System
A9000-Site_B	A9000R-Site_A	A9000R-Site_C
Volume	Volume	Volume
ITSO_FA_Volume_2	ITSO_FA_Volume_2	ITSO_FA_Volume_2

Figure 6-34 Multi-site configuration as shown in the Hyper-Scale Manager GUI

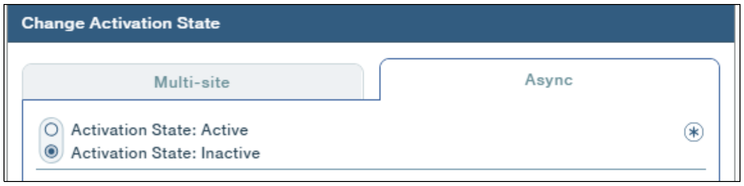
With the I/O now being serviced by the target System B, we can dismantle the Multi-site relationship to remove the volume from System A. Complete the following steps:

1. Deactivate the HyperSwap and the Active Asynchronous legs of the Multi-site relationship, as shown in Figure 6-35 and Figure 6-36.



○ Activation State: Active
● Activation State: Inactive

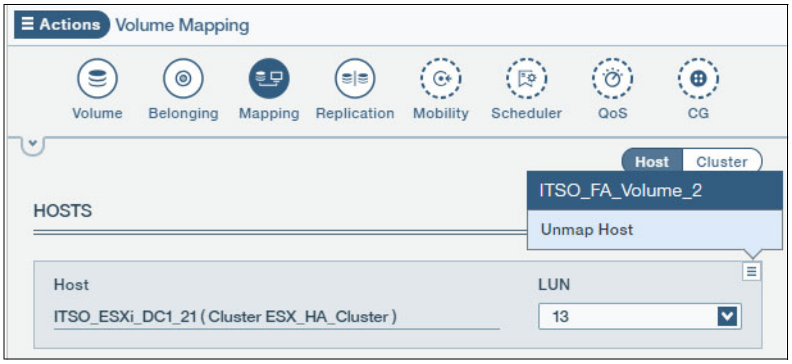
Figure 6-35 Deactivating the HyperSwap leg of the Multi-site relationship



○ Activation State: Active
● Activation State: Inactive

Figure 6-36 Deactivating the Active Asynchronous leg of the Multi-site relationship

2. Unmap the migrated volumes from the now designated Secondary system (A, in our example), as shown in Figure 6-37.



Host	LUN
ITSO_ESXi_DC1_21 (Cluster ESX_HA_Cluster)	13

Figure 6-37 Unmap migrated volumes from the Secondary system

3. The Multi-site relationship can now be dismantled. From the Actions menu, select **Delete** → **Convert Multi-site to Async Mirror,** as shown in Figure 6-38.

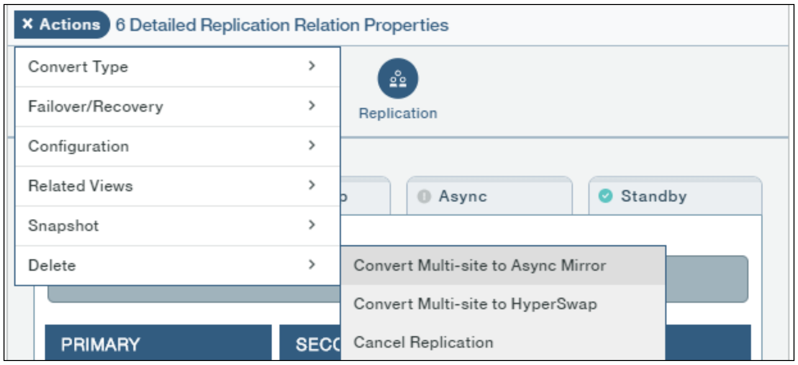
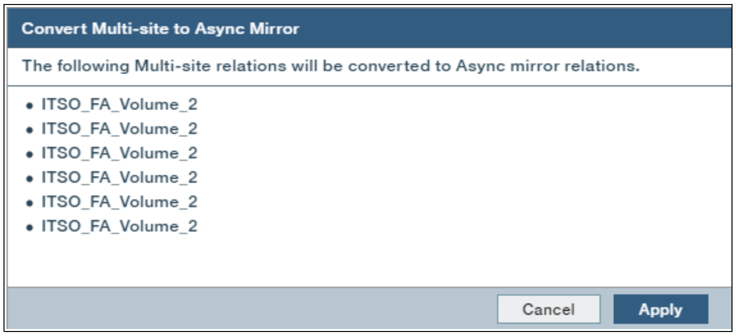


Figure 6-38 Action menu to convert the Multi-site relationship to Async Mirror

a. Review the summary and confirm by clicking **Apply** (see Figure 6-39).



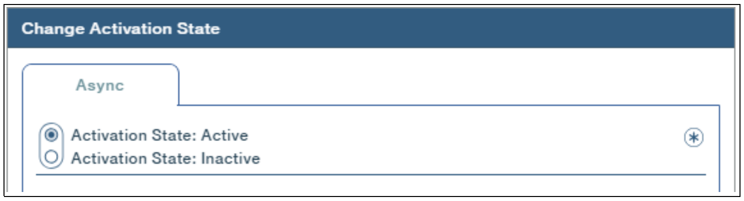
Convert Multi-site to Async Mirror

The following Multi-site relations will be converted to Async mirror relations.

- ITSO_FA_Volume_2
- ITSO_FA_Volume_2
- ITSO_FA_Volume_2
- ITSO_FA_Volume_2
- ITSO_FA_Volume_2
- ITSO_FA_Volume_2

Figure 6-39 Multi-site to Async Mirror convert summary

4. The HyperSwap and the Standby Asynchronous leg of the Multi-site relationship are now deleted, which leaves us with a stand-alone Asynchronous relationship between System B and System C. Activate this relationship as shown in Figure 6-40 to reach the final configuration, as shown in Figure 6-41.



Change Activation State

● Activation State: Active
○ Activation State: Inactive

Figure 6-40 Activate the Asynchronous relationship between B and C

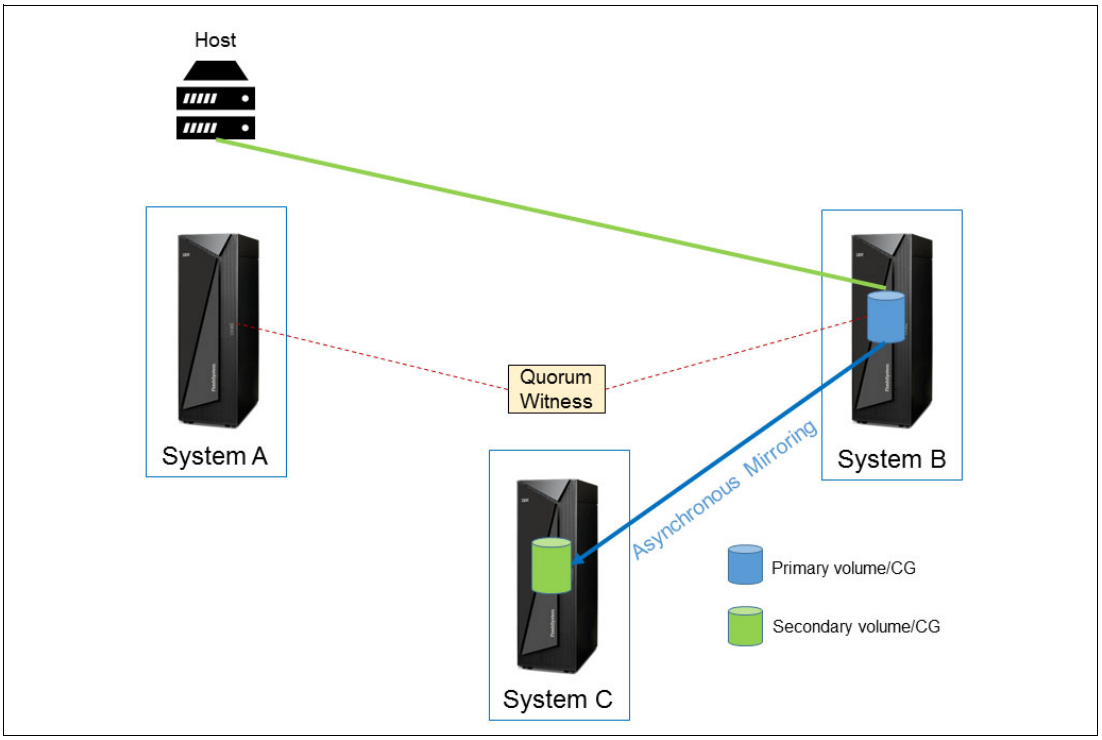
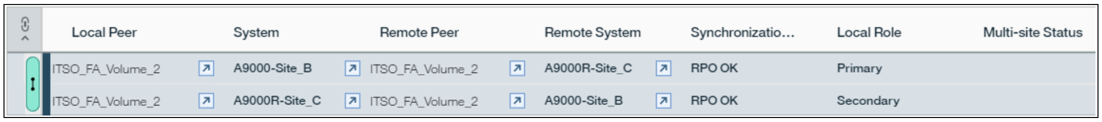


Figure 6-41 Final configuration

Figure 6-42 shows the HSM GUI representation.



Local Peer	System	Remote Peer	Remote System	Synchronizatio...	Local Role	Multi-site Status
ITSO_FA_Volume_2	A9000-Site_B	ITSO_FA_Volume_2	A9000R-Site_C	RPO OK	Primary	
ITSO_FA_Volume_2	A9000R-Site_C	ITSO_FA_Volume_2	A9000-Site_B	RPO OK	Secondary	

Figure 6-42 Final relationship details as shown in the Hyper-Scale Manager GUI

7

Multi-site scenarios

This chapter describes failure scenarios and recovery steps (failback). The following topics are covered:

7.1 Failure and recovery of the storage system that is serving the primary volume or CG

This section describes the failure of the storage system that is serving I/Os for the Primary volume or Consistency Group in a Multi-site relationship.

A Multi-site configuration with the Primary volume or Consistency Group on System A is shown in Figure 7-1.

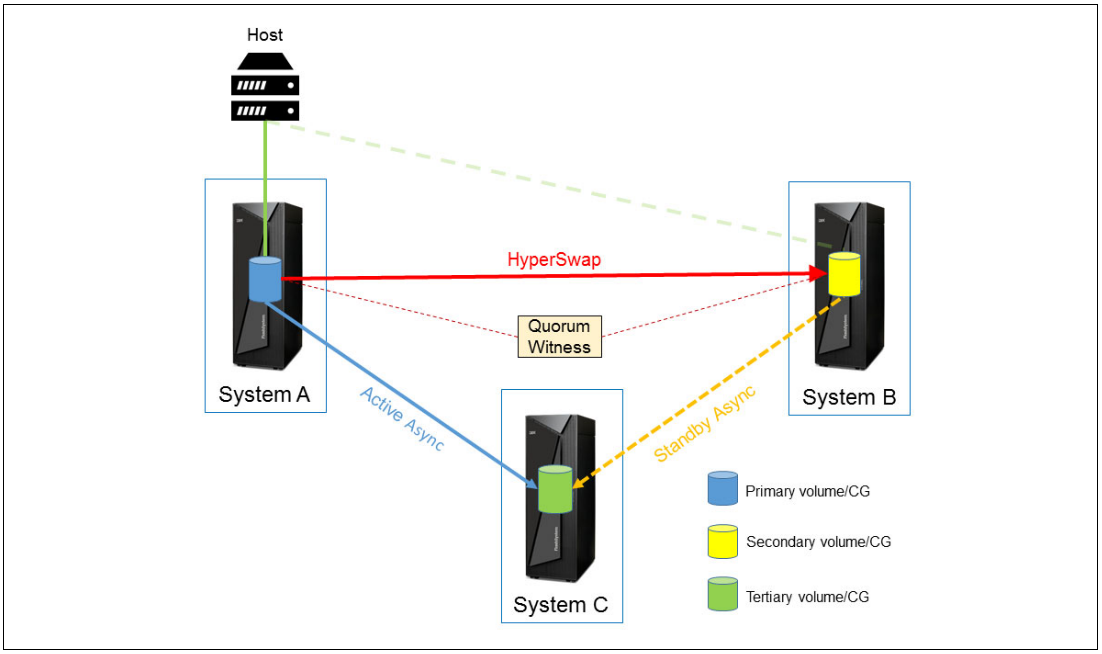
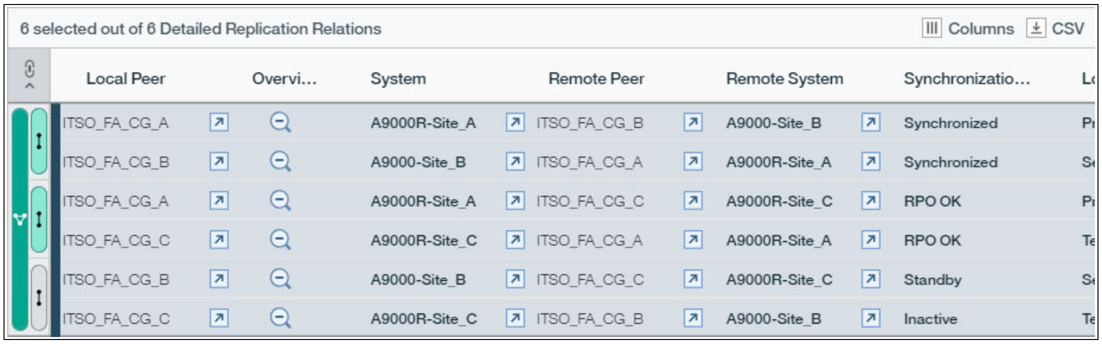


Figure 7-1 Multi-site relationship example

In this example, the host sees the volume through paths from System A and System B, but I/O is served only through the preferred paths from System A. Data is replicated synchronously from System A to System B through the HyperSwap leg of the Multi-Site relationship. It also is replicated asynchronously from A to C through the active Asynchronous leg. The Asynchronous B to C relationship is in Standby.

Figure 7-2 shows the replication details of a Consistency Group in a Multi-site relationship in a steady state.



6 selected out of 6 Detailed Replication Relations

Local Peer	Overvi...	System	Remote Peer	Remote System	Synchronizatio...	Lo
ITSO_FA_CG_A		A9000R-Site_A	ITSO_FA_CG_B	A9000-Site_B	Synchronized	Pr
ITSO_FA_CG_B		A9000-Site_B	ITSO_FA_CG_A	A9000R-Site_A	Synchronized	Se
ITSO_FA_CG_A		A9000R-Site_A	ITSO_FA_CG_C	A9000R-Site_C	RPO OK	Pr
ITSO_FA_CG_C		A9000R-Site_C	ITSO_FA_CG_A	A9000R-Site_A	RPO OK	Te
ITSO_FA_CG_B		A9000-Site_B	ITSO_FA_CG_C	A9000R-Site_C	Standby	Se
ITSO_FA_CG_C		A9000R-Site_C	ITSO_FA_CG_B	A9000-Site_B	Inactive	Te

Figure 7-2 Multi-site replication details

7.1.1 System A failure scenario

A failure of System A is shown in Figure 7-3. The following results occur:

- The host and System B lose connection to System A.
- The asynchronous replication from System A to System C is disrupted.

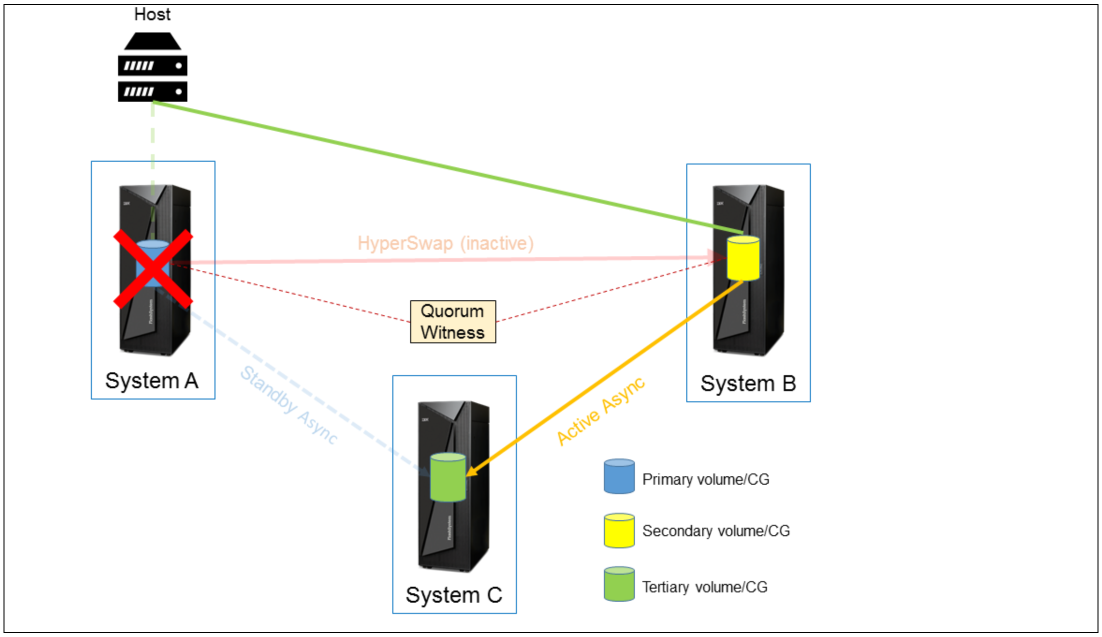
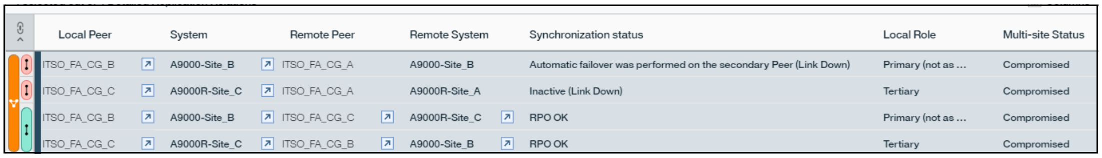


Figure 7-3 Multi-site configuration after the failure of System A

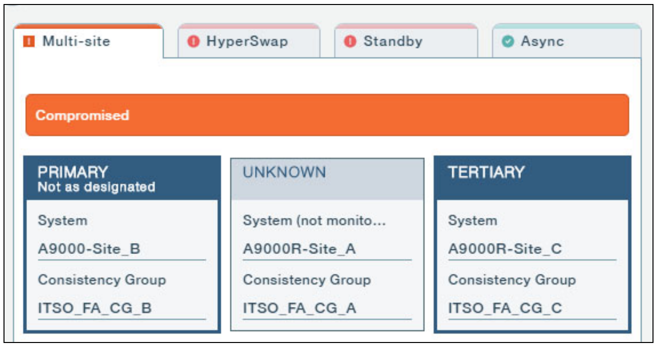
In the Hyper-Scale Manager GUI (see Figure 7-4), System A disappears from the replication relationship details and an automatic failover is performed to the Secondary volume on System B.



Local Peer	System	Remote Peer	Remote System	Synchronization status	Local Role	Multi-site Status
ITSO_FA_CG_B	A9000-Site_B	ITSO_FA_CG_A	A9000-Site_B	Automatic failover was performed on the secondary Peer (Link Down)	Primary (not as ...	Compromised
ITSO_FA_CG_C	A9000R-Site_C	ITSO_FA_CG_A	A9000R-Site_A	Inactive (Link Down)	Tertiary	Compromised
ITSO_FA_CG_B	A9000-Site_B	ITSO_FA_CG_C	A9000R-Site_C	RPO OK	Primary (not as ...	Compromised
ITSO_FA_CG_C	A9000R-Site_C	ITSO_FA_CG_B	A9000-Site_B	RPO OK	Tertiary	Compromised

Figure 7-4 Replication relations details after System A failure

Now the multi-site configuration status is compromised because System B is acting not as designated, and System A is temporarily not being monitored (see Figure 7-5).

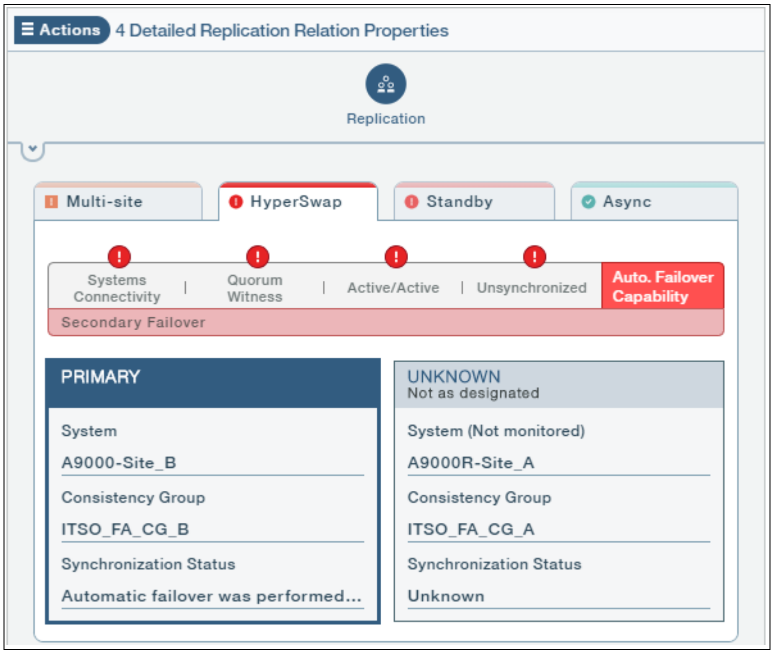


Compromised

PRIMARY Not as designated	UNKNOWN	TERTIARY
System	System (not monito...	System
A9000-Site_B	A9000R-Site_A	A9000R-Site_C
Consistency Group	Consistency Group	Consistency Group
ITSO_FA_CG_B	ITSO_FA_CG_A	ITSO_FA_CG_C

Figure 7-5 Multi-site relationship status after failover on B

The status of the HyperSwap relationship between the Primary and Secondary volumes is represented in greater detail in the HyperSwap tab of the replication relation properties, as shown in Figure 7-6.



PRIMARY

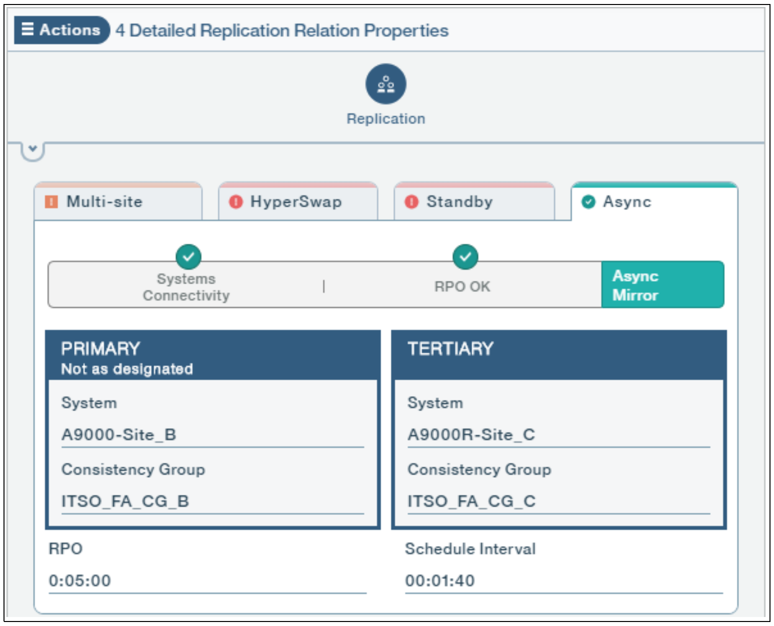
System
A9000-Site_B
Consistency Group
ITSO_FA_CG_B
Synchronization Status
Automatic failover was performed...

UNKNOWN
Not as designated

System (Not monitored)
A9000R-Site_A
Consistency Group
ITSO_FA_CG_A
Synchronization Status
Unknown

Figure 7-6 HyperSwap relationship status after failover on System B

Upon the failure of System A, the Standby asynchronous mirror between Systems B and C was automatically activated. Its details now appear in the Async tab of the replication relation properties, as shown in Figure 7-7.



PRIMARY
Not as designated

System
A9000-Site_B
Consistency Group
ITSO_FA_CG_B

TERTIARY

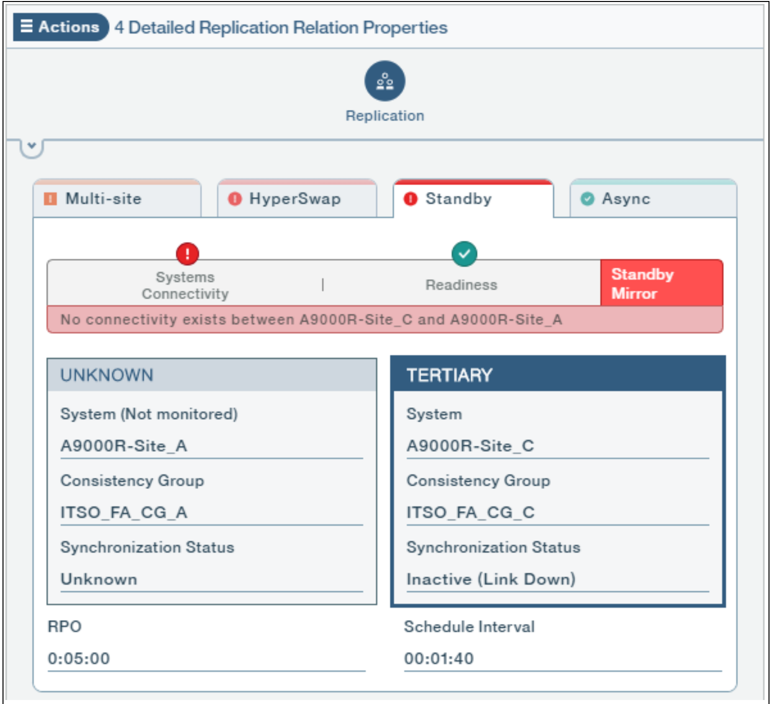
System
A9000R-Site_C
Consistency Group
ITSO_FA_CG_C

RPO
0:05:00

Schedule Interval
00:01:40

Figure 7-7 Systems B and C asynchronous mirror status after failover on System B

If System A were up, a Standby asynchronous mirror can be defined between it and System C. However, because System A is down, this option is not available, as shown in the Standby tab of the replication relation properties (see Figure 7-8).



4 Detailed Replication Relation Properties

No connectivity exists between A9000R-Site_C and A9000R-Site_A

UNKNOWN

System (Not monitored)
A9000R-Site_A

Consistency Group
ITSO_FA_CG_A

Synchronization Status
Unknown

TERTIARY

System
A9000R-Site_C

Consistency Group
ITSO_FA_CG_C

Synchronization Status
Inactive (Link Down)

RPO
0:05:00

Schedule Interval
00:01:40

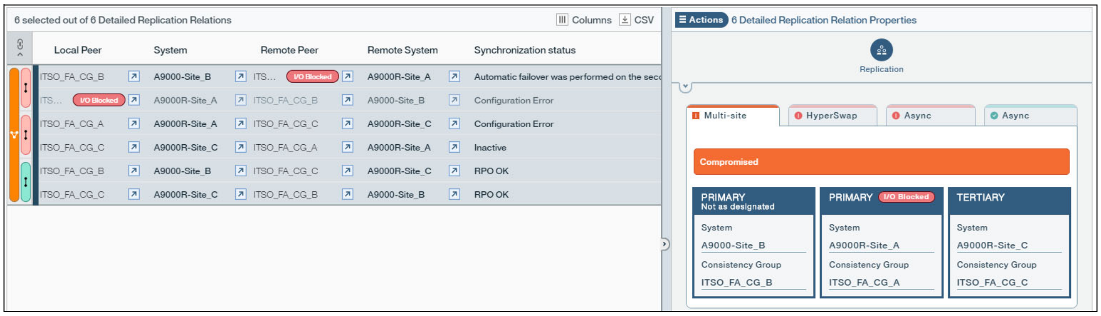
Figure 7-8 Standby asynchronous mirror status after failover on System B

7.1.2 Multi-site relationship recovery after System A failure

When System A is operational again, the original Primary volume is still not serving I/Os and is in a blocked state.

The original Secondary volume on System B is serving the I/O and has the primary role, as can be seen in the replication relations details.

The HyperSwap peer volumes are not synchronized. Therefore, the multi-site relationship is still compromised, as shown in the Multi-site tab (see Figure 7-9).



6 selected out of 6 Detailed Replication Relations

Local Peer	System	Remote Peer	Remote System	Synchronization status
ITSO_FA_CG_B	A9000-Site_B	ITS… I/O Blocked	A9000R-Site_A	Automatic failover was performed on the seco
ITS… I/O Blocked	A9000R-Site_A	ITSO_FA_CG_B	A9000-Site_B	Configuration Error
ITSO_FA_CG_A	A9000R-Site_A	ITSO_FA_CG_C	A9000R-Site_C	Configuration Error
ITSO_FA_CG_C	A9000R-Site_C	ITSO_FA_CG_A	A9000R-Site_A	Inactive
ITSO_FA_CG_B	A9000-Site_B	ITSO_FA_CG_C	A9000R-Site_C	RPO OK
ITSO_FA_CG_C	A9000R-Site_C	ITSO_FA_CG_B	A9000-Site_B	RPO OK

6 Detailed Replication Relation Properties

Compromised

PRIMARY
Not as designated

System
A9000-Site_B

Consistency Group
ITSO_FA_CG_B

PRIMARY I/O Blocked

System
A9000R-Site_A

Consistency Group
ITSO_FA_CG_A

TERTIARY

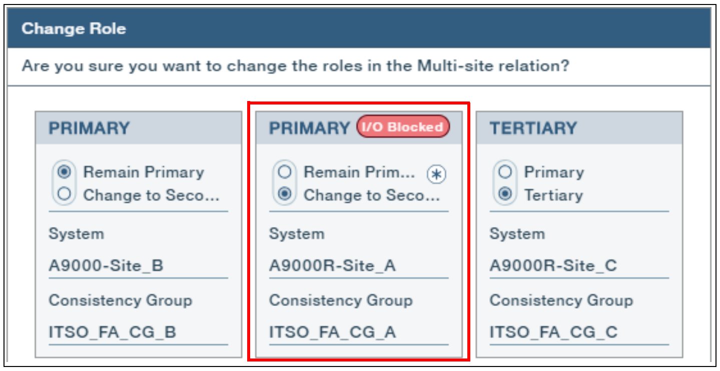
System
A9000R-Site_C

Consistency Group
ITSO_FA_CG_C

Figure 7-9 Replication relations details after restoring System A

Multi-site relationship recovery from the GUI

To re-establish the high availability, the role of the original Primary volume first must be manually changed to Secondary by using the replication Actions menu (select **Failover/Recovery** → **Change Role**) and then, selecting **Change to Secondary**, as shown in Figure 7-10.



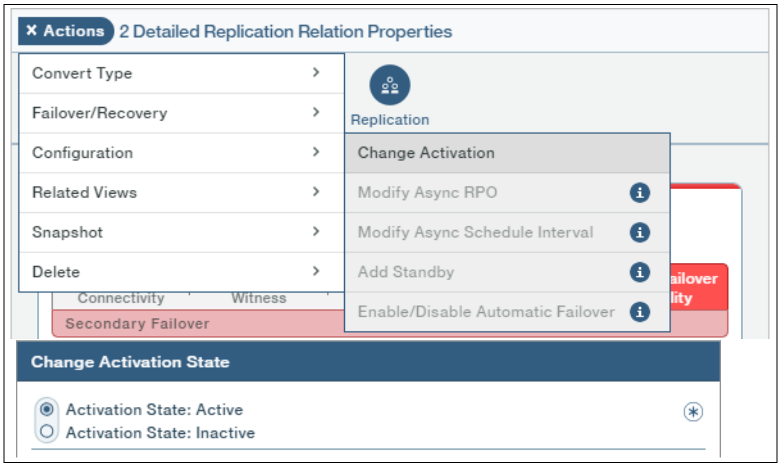
Change Role

Are you sure you want to change the roles in the Multi-site relation?

PRIMARY	PRIMARY I/O Blocked	TERTIARY
● Remain Primary ○ Change to Seco...	○ Remain Prim... ● Change to Seco...	○ Primary ● Tertiary
System	System	System
A9000-Site_B	A9000R-Site_A	A9000R-Site_C
Consistency Group	Consistency Group	Consistency Group
ITSO_FA_CG_B	ITSO_FA_CG_A	ITSO_FA_CG_C

Figure 7-10 Changing role of the original Primary volume to Secondary

Then, HyperSwap must be activated, as shown in Figure 7-11. Ensure that you select only the HyperSwap relationship to activate. If the entire Multi-site configuration is selected, you cannot activate HyperSwap only.

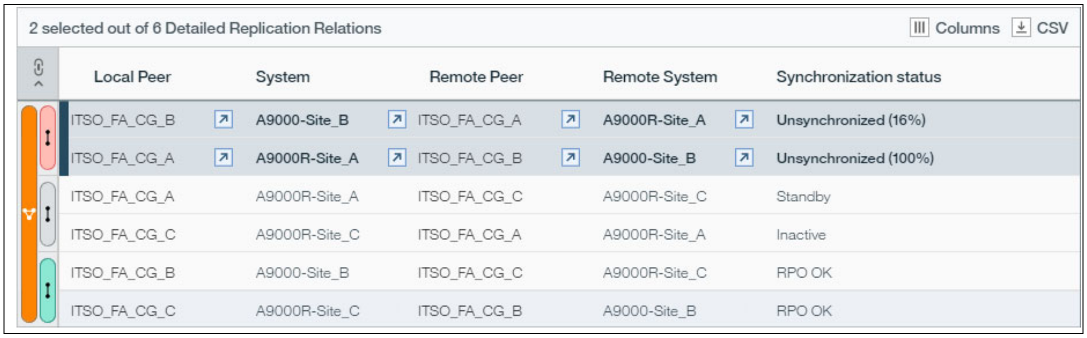


Change Activation State

● Activation State: Active
○ Activation State: Inactive

Figure 7-11 HyperSwap activation

Figure 7-12 shows the result of activating HyperSwap. The volumes on System A and System B are being automatically synchronized.

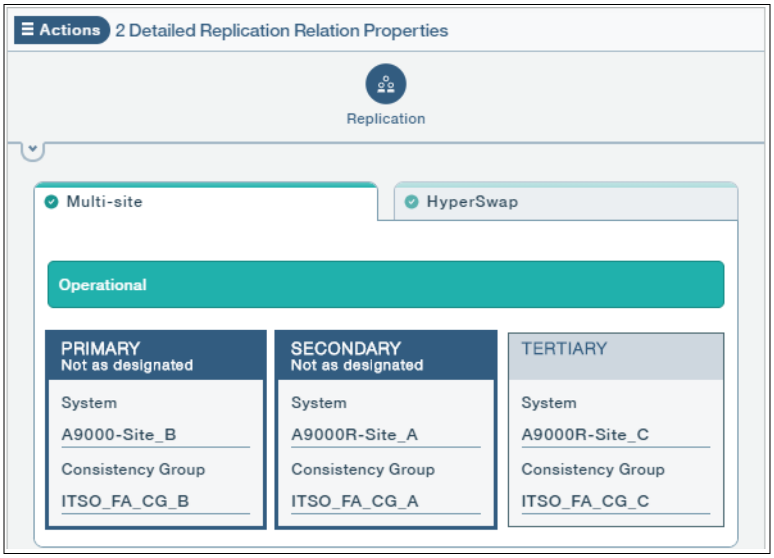


2 selected out of 6 Detailed Replication Relations

Local Peer	System	Remote Peer	Remote System	Synchronization status
ITSO_FA_CG_B	A9000-Site_B	ITSO_FA_CG_A	A9000R-Site_A	Unsynchronized (16%)
ITSO_FA_CG_A	A9000R-Site_A	ITSO_FA_CG_B	A9000-Site_B	Unsynchronized (100%)
ITSO_FA_CG_A	A9000R-Site_A	ITSO_FA_CG_C	A9000R-Site_C	Standby
ITSO_FA_CG_C	A9000R-Site_C	ITSO_FA_CG_A	A9000R-Site_A	Inactive
ITSO_FA_CG_B	A9000-Site_B	ITSO_FA_CG_C	A9000R-Site_C	RPO OK
ITSO_FA_CG_C	A9000R-Site_C	ITSO_FA_CG_B	A9000-Site_B	RPO OK

Figure 7-12 Synchronizing the HyperSwap volume

After the volumes are synchronized, the GUI indicates that the high availability is re-established. An automatic failover is possible, and the multi-site relationship is operational, as shown in Figure 7-13.



PRIMARY Not as designated	SECONDARY Not as designated	TERTIARY
System	System	System
A9000-Site_B	A9000R-Site_A	A9000R-Site_C
Consistency Group	Consistency Group	Consistency Group
ITSO_FA_CG_B	ITSO_FA_CG_A	ITSO_FA_CG_C

Figure 7-13 Re-synchronized HyperSwap volume

Important: Volumes in the CG are resynchronized individually, and you can observe in the GUI that is shown in Figure 7-12 on page 130 that the synchronization status shows reaching 100% several times as it cycles through all the volumes in the CG.

In addition, the Multi-site relationship does not immediately become operational. Instead, it remains in the Compromised state until the system completes all of the required background activity. You must wait until the status changes to Operational.

The high availability of the HyperSwap volume is now re-established, but I/O is served from System B, which now holds the Primary volume. The transitional state is shown in Figure 7-14 on page 132.

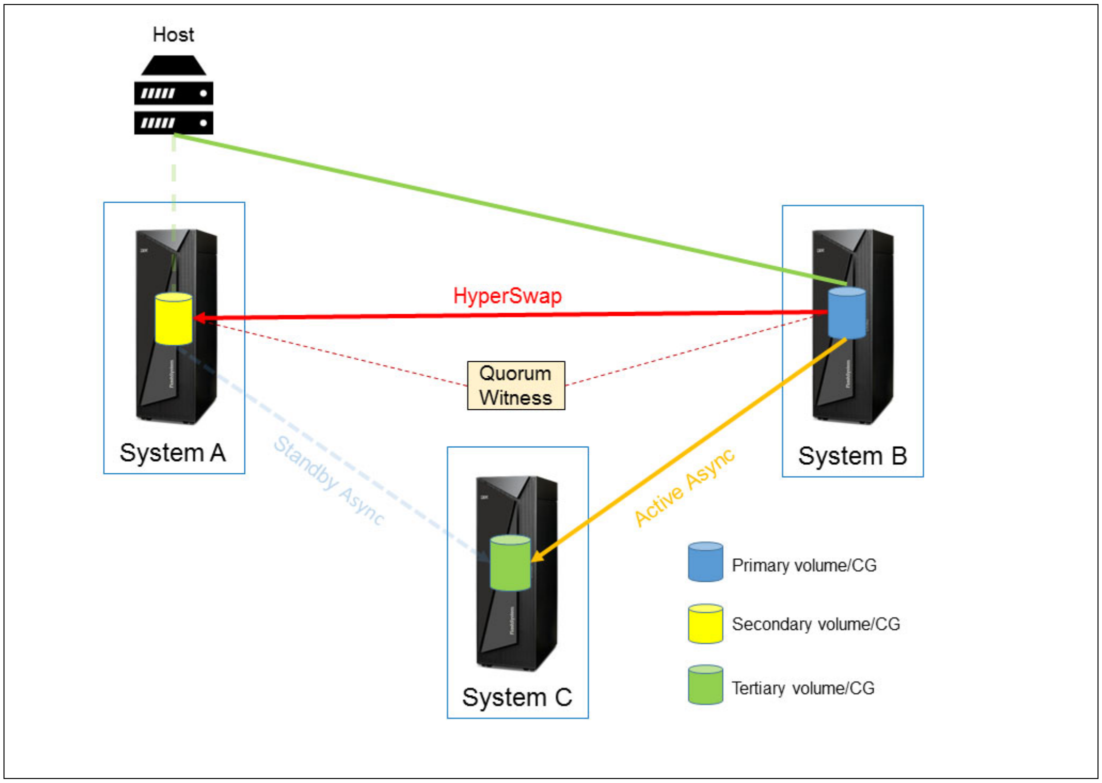
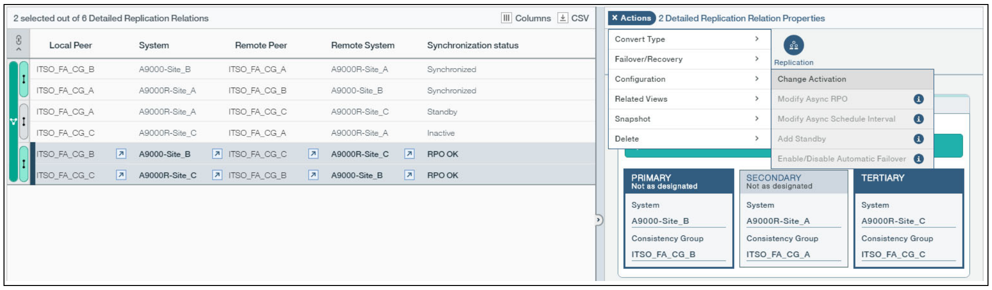


Figure 7-14 State of the Multi-site relationship after A is recovered and High Availability reinstated

To failback to the original configuration that existed before the failover, deactivate the asynchronous mirror between Systems B and C, as shown in Figure 7-15.



2 selected out of 6 Detailed Replication Relations

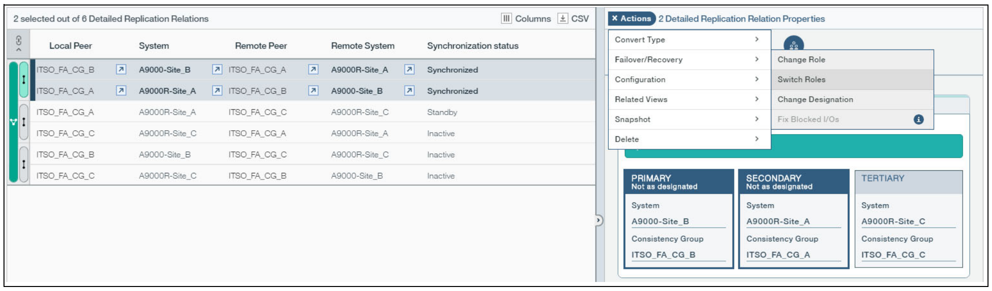
Local Peer	System	Remote Peer	Remote System	Synchronization status
ITSO_FA_CG_B	A9000-Site_B	ITSO_FA_CG_A	A9000R-Site_A	Synchronized
ITSO_FA_CG_A	A9000R-Site_A	ITSO_FA_CG_B	A9000-Site_B	Synchronized
ITSO_FA_CG_A	A9000R-Site_A	ITSO_FA_CG_C	A9000R-Site_C	Standby
ITSO_FA_CG_C	A9000R-Site_C	ITSO_FA_CG_A	A9000R-Site_A	Inactive
ITSO_FA_CG_B	A9000-Site_B	ITSO_FA_CG_C	A9000R-Site_C	RPO OK
ITSO_FA_CG_C	A9000R-Site_C	ITSO_FA_CG_B	A9000-Site_B	RPO OK

2 Detailed Replication Relation Properties

PRIMARY Not as designated	SECONDARY Not as designated	TERTIARY
System A9000-Site_B	System A9000R-Site_A	System A9000R-Site_C
Consistency Group ITSO_FA_CG_B	Consistency Group ITSO_FA_CG_A	Consistency Group ITSO_FA_CG_C

Figure 7-15 De-activating asynchronous mirror between Systems B and C

Then, switch the roles of the HyperSwap peers, as shown in Figure 7-16.



2 selected out of 6 Detailed Replication Relations

Local Peer	System	Remote Peer	Remote System	Synchronization status
ITSO_FA_CG_B	A9000-Site_B	ITSO_FA_CG_A	A9000R-Site_A	Synchronized
ITSO_FA_CG_A	A9000R-Site_A	ITSO_FA_CG_B	A9000-Site_B	Synchronized
ITSO_FA_CG_A	A9000R-Site_A	ITSO_FA_CG_C	A9000R-Site_C	Standby
ITSO_FA_CG_C	A9000R-Site_C	ITSO_FA_CG_A	A9000R-Site_A	Inactive
ITSO_FA_CG_B	A9000-Site_B	ITSO_FA_CG_C	A9000R-Site_C	Inactive
ITSO_FA_CG_C	A9000R-Site_C	ITSO_FA_CG_B	A9000-Site_B	Inactive

2 Detailed Replication Relation Properties

PRIMARY Not as designated	SECONDARY Not as designated	TERTIARY
System A9000-Site_B	System A9000R-Site_A	System A9000R-Site_C
Consistency Group ITSO_FA_CG_B	Consistency Group ITSO_FA_CG_A	Consistency Group ITSO_FA_CG_C

Figure 7-16 Switching the roles of the Primary and Secondary volumes

Finally, activate the asynchronous mirroring relationship between Systems A and C. After a brief period of RPO lagging, the Primary-Tertiary asynchronous mirroring is reactivated, and the multi-site relationship is fully restored.

Multi-site relationship recovery from the CLI

The recovery procedure can alternatively be carried out by issuing CLI commands to the three individual systems.

As with the GUI procedure, the first step is to re-establish the high availability by changing the role of the original Primary volume/Consistency Group to Secondary. This step can be done by issuing the **`multisite_change_role`** command on the system that is hosting the original Primary volume or Consistency Group, as shown in Example 7-1.

Example 7-1 Changing Multi-site roles by using the CLI

```
A9000R-Site_A>>multisite_change_role cg=ITSO_FA_CG_A new_role=SMaster

Warning:   Are you sure you want to change the Multi-site role? y/n: y
Command executed successfully.
```

To resynchronize the volume data between Systems A and B, the HyperSwap relationship must be reactived by using the **`ha_activate`** command on the system that is acting as primary (System B, in our case), as shown in Example 7-2.

Example 7-2 Activating the HyperSwap leg of the Multi-site relationship by using the CLI

```
A9000-Site_B>>ha_activate cg=ITSO_FA_CG_B
Command executed successfully.
```

Wait until the status of the HyperSwap relationship becomes “Synchronized” again. You can check the status by using the **`ha_list`** command, as shown in Example 7-3.

Example 7-3 Displaying HyperSwap relationship details by using the CLI

```
A9000-Site_B>>ha_list cg=ITSO_FA_CG_B
Name          HA Object  Role    Remote System  Active  Status
ITSO_FA_CG_B  CG         Master  A9000R-Site_A  no      Synchronized
```

Temporarily disable the Asynchronous relationship between the current Primary and the Tertiary volume or Consistency Group by issuing the **`mirror_deactivate`** command on System B, as shown in Example 7-4.

Example 7-4 Deactivating an Asynchronous relationship by using the CLI

```
A9000-Site_B>>mirror_deactivate cg=ITSO_FA_CG_B target=A9000R-Site_C

Warning:   Are you sure you want to deactivate mirroring? y/n: y
Command executed successfully.
```

At this point, it is possible to perform a switch role to reinstate the original Primary-Secondary roles. Run the **multisite_switch_roles** command on the current Primary system, as shown in Example 7-5.

Example 7-5 Switching roles in a Multi-site relationship by using the CLI

```
A9000-Site_B>>multisite_switch_roles  cg=ITSO_FA_CG_B

Warning:   Are you sure you want to switch the roles in this relation? y/n: y
Command executed successfully.
```

Activate the Asynchronous relationship between the systems now hosting the Primary and Tertiary volumes/Consistency Groups by using the **multisite_activate_async_mirror** command, as shown in Example 7-6.

Example 7-6 Activate the Active Asynchronous leg of a Multi-site relationship by using the CLI

```
A9000R-Site_A>>multisite_activate_async_mirror cg=ITSO_FA_CG_A
Command executed successfully.
```

Wait until the status of the Asynchronous relationship reports RPO OK. Use the **mirror_list** command to verify this status, as shown in Example 7-7.

Example 7-7 Displaying Asynchronous relationship details by using the CLI

```
A9000R-Site_A>>mirror_list cg=ITSO_FA_CG_A
Name          Mirror Type     Role    Remote System  Remote Peer   Status
ITSO_FA_CG_A  async_interval  Master  A9000R-Site_C  ITSO_FA_CG_C  RPO OK
```

Finally, check that the Multi-site relationship is in the Compromised state and reports as Operational by using the **multisite_list** command, as shown in Example 7-8.

Example 7-8 Displaying Multi-site relationship details by using the CLI

```
A9000R-Site_A>>multisite_list cg=ITSO_FA_CG_A
Name           Multisite Object   Multisite ID       Role     State
ITSO_FA_CG_A   CG                 0142C95000000009   SMaster  Operational
```

7.1.3 Host paths and I/O behavior during the outage and recovery

The Multi-site implementation uses HyperSwap to ensure high availability and prevent loss of access to volumes at the production sites. An Active Asynchronous mirror also remains active, even when the Primary or Secondary system is lost. As a result, data consistency at the DR site is also ensured.

Example 7-9 on page 135 shows the paths state on a VMware ESXi 6.0 hosts for a volume in a Consistency Group that is in a Multi-site relationship. As with a regular HyperSwap configuration, paths towards the Primary volumes (the ones in Target Port Group 0) are preferred or optimized (Array Priority: 1, TPG_state=AO, active optimized). Paths towards the Secondary system (Target Port Group 1) are available, but not preferred (Array Priority: 0, TPG_state=ANO, active not optimized).

Example 7-9 Path details for a volume in a Multi-site relationship on an ESXi host

```
[root@A9Demoesxdc1:~] esxcli storage nmp path list --device
naa.6001738c7c80539800000000000ecf70 | grep -E 'fc.|Prio|Storage'

fc.20000090faa30c8e:10000090faa30c8e-fc.5001738056750000:5001738056750112-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 0
   Storage Array Type Path Config:
{TPG_id=1,TPG_state=ANO,RTP_id=4129,RTP_health=UP}
fc.20000090faa30c8e:10000090faa30c8e-fc.5001738056750000:5001738056750132-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 0
   Storage Array Type Path Config:
{TPG_id=1,TPG_state=ANO,RTP_id=12321,RTP_health=UP}
fc.20000090faa30c8d:10000090faa30c8d-fc.5001738056750000:5001738056750110-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 0
   Storage Array Type Path Config:
{TPG_id=1,TPG_state=ANO,RTP_id=4097,RTP_health=UP}
fc.20000090faa30c8e:10000090faa30c8e-fc.5001738056750000:5001738056750122-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 0
   Storage Array Type Path Config:
{TPG_id=1,TPG_state=ANO,RTP_id=8225,RTP_health=UP}
fc.20000090faa30c8d:10000090faa30c8d-fc.5001738056750000:5001738056750120-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 0
   Storage Array Type Path Config:
{TPG_id=1,TPG_state=ANO,RTP_id=8193,RTP_health=UP}
fc.20000090faa30c8d:10000090faa30c8d-fc.5001738056750000:5001738056750130-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 0
   Storage Array Type Path Config:
{TPG_id=1,TPG_state=ANO,RTP_id=12289,RTP_health=UP}
fc.20000090faa30c8e:10000090faa30c8e-fc.5001738053980000:5001738053980132-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 1
   Storage Array Type Path Config:
{TPG_id=0,TPG_state=AO,RTP_id=770,RTP_health=UP}
fc.20000090faa30c8d:10000090faa30c8d-fc.5001738053980000:5001738053980120-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 1
   Storage Array Type Path Config:
{TPG_id=0,TPG_state=AO,RTP_id=512,RTP_health=UP}
fc.20000090faa30c8e:10000090faa30c8e-fc.5001738053980000:5001738053980142-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 1
   Storage Array Type Path Config:
{TPG_id=0,TPG_state=AO,RTP_id=1026,RTP_health=UP}
fc.20000090faa30c8d:10000090faa30c8d-fc.5001738053980000:5001738053980130-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 1
   Storage Array Type Path Config:
{TPG_id=0,TPG_state=AO,RTP_id=768,RTP_health=UP}
```

```
fc.20000090faa30c8e:10000090faa30c8e-fc.5001738053980000:5001738053980112-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 1
   Storage Array Type Path Config:
{TPG_id=0,TPG_state=AO,RTP_id=258,RTP_health=UP}
fc.20000090faa30c8d:10000090faa30c8d-fc.5001738053980000:5001738053980110-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 1
   Storage Array Type Path Config:
{TPG_id=0,TPG_state=AO,RTP_id=256,RTP_health=UP}
fc.20000090faa30c8d:10000090faa30c8d-fc.5001738053980000:5001738053980140-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 1
   Storage Array Type Path Config:
{TPG_id=0,TPG_state=AO,RTP_id=1024,RTP_health=UP}
fc.20000090faa30c8e:10000090faa30c8e-fc.5001738053980000:5001738053980122-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 1
   Storage Array Type Path Config:
{TPG_id=0,TPG_state=AO,RTP_id=514,RTP_health=UP}
```

In Example 7-9 on page 135, the system that is hosting the Primary volumes (A) in the Consistency Group is an A9000R zoned to host with eight ports across two fabrics (eight paths total), and the system that is hosting the Secondary volumes (B) is an A9000 with six ports that are zoned across two fabrics (six paths total). In a normal state, all of the paths are reported as UP.

When System A fails, the host starts reporting errors on the paths that lead to it and finally marks them as DOWN. It continues I/O on the paths towards System B, now selecting them as preferred. Example 7-10 (truncated for brevity) shows the paths state in this configuration.

Example 7-10 Path details after the failure of System A

```
[root@A9Demoesxdc1:~] esxcli storage nmp path list --device
naa.6001738c7c80539800000000000ecf70 | grep -E 'fc.|Prio|Storage'

fc.20000090faa30c8e:10000090faa30c8e-fc.5001738056750000:5001738056750112-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 1
   Storage Array Type Path Config:
{TPG_id=1,TPG_state=AO,RTP_id=4129,RTP_health=UP}
fc.20000090faa30c8d:10000090faa30c8d-fc.5001738053980000:5001738053980140-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 0
   Storage Array Type Path Config:
{TPG_id=0,TPG_state=UNAVAIL,RTP_id=1024,RTP_health=DOWN}
fc.20000090faa30c8e:10000090faa30c8e-fc.5001738056750000:5001738056750132-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 1
   Storage Array Type Path Config:
{TPG_id=1,TPG_state=AO,RTP_id=12321,RTP_health=UP}
fc.20000090faa30c8d:10000090faa30c8d-fc.5001738056750000:5001738056750110-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 1
   Storage Array Type Path Config:
{TPG_id=1,TPG_state=AO,RTP_id=4097,RTP_health=UP}
```

```
fc.20000090faa30c8e:10000090faa30c8e-fc.5001738056750000:5001738056750122-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 1
   Storage Array Type Path Config:
{TPG_id=1,TPG_state=AO,RTP_id=8225,RTP_health=UP}
fc.20000090faa30c8d:10000090faa30c8d-fc.5001738056750000:5001738056750120-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 1
   Storage Array Type Path Config:
{TPG_id=1,TPG_state=AO,RTP_id=8193,RTP_health=UP}
fc.20000090faa30c8e:10000090faa30c8e-fc.5001738053980000:5001738053980122-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 0
   Storage Array Type Path Config:
{TPG_id=0,TPG_state=UNAVAIL,RTP_id=514,RTP_health=DOWN}
...
```

When System A is available again, paths towards A are reported as UP again. However, the paths to System B remain preferred because its designation is still Primary (see Example 7-11).

Example 7-11 Paths state after A is recovered, with B still Primary

```
[root@A9Demoesxdc1:~] esxcli storage nmp path list --device
naa.6001738c7c80539800000000000ecf70 | grep -E 'fc.|Prio|Storage'

fc.20000090faa30c8e:10000090faa30c8e-fc.5001738056750000:5001738056750112-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 1
   Storage Array Type Path Config:
{TPG_id=1,TPG_state=AO,RTP_id=4129,RTP_health=UP}
fc.20000090faa30c8d:10000090faa30c8d-fc.5001738053980000:5001738053980140-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 0
   Storage Array Type Path Config:
{TPG_id=0,TPG_state=ANO,RTP_id=1024,RTP_health=UP}
fc.20000090faa30c8e:10000090faa30c8e-fc.5001738056750000:5001738056750132-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 1
   Storage Array Type Path Config:
{TPG_id=1,TPG_state=AO,RTP_id=12321,RTP_health=UP}
fc.20000090faa30c8d:10000090faa30c8d-fc.5001738056750000:5001738056750110-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 1
   Storage Array Type Path Config:
{TPG_id=1,TPG_state=AO,RTP_id=4097,RTP_health=UP}
fc.20000090faa30c8e:10000090faa30c8e-fc.5001738056750000:5001738056750122-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 1
   Storage Array Type Path Config:
{TPG_id=1,TPG_state=AO,RTP_id=8225,RTP_health=UP}
fc.20000090faa30c8d:10000090faa30c8d-fc.5001738056750000:5001738056750120-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 1
   Storage Array Type Path Config:
{TPG_id=1,TPG_state=AO,RTP_id=8193,RTP_health=UP}
```

```
fc.20000090faa30c8e:10000090faa30c8e-fc.5001738053980000:5001738053980122-naa.6001
738c7c80539800000000000ecf70
   Array Priority: 0
   Storage Array Type Path Config:
{TPG_id=0,TPG_state=ANO,RTP_id=514,RTP_health=UP}
...
```

Finally, when the Multi-site relationship is brought back to its original state (with A Primary and B Secondary) the path state returns to its original configuration, as shown in Example 7-9 on page 135.

The I/O activity on the systems that are involved in the Multi-site relationship is shown in Figure 7-17.

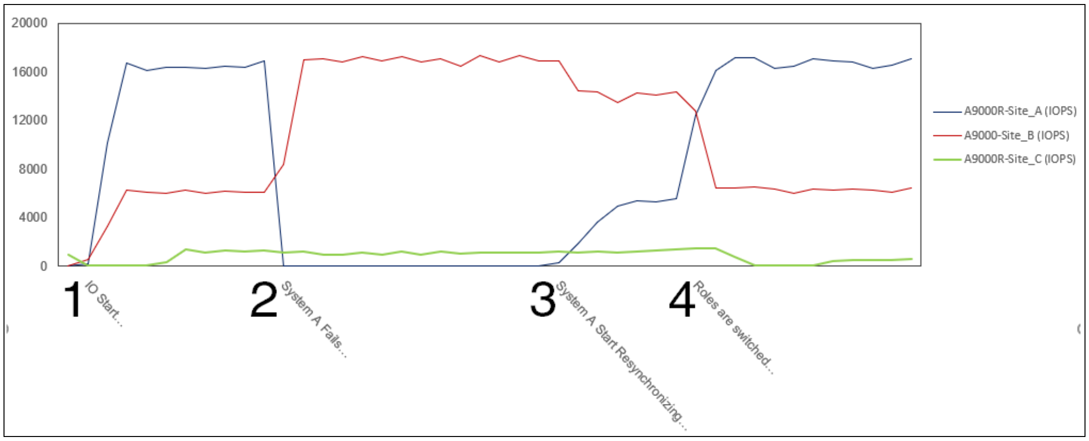


Figure 7-17 IO activity on the Systems involved in the Multi-site relationship

In the initial phase (steady state, between points 1 and 2 in Figure 7-17) System A (Primary) receives read and write operations from the host; System B receives the written data only as Secondary System in the HyperSwap leg of the Multi-site relationship. When System A fails (point 2), all I/Os from the host are diverted to System B, which continues to serve the host while A is down (time interval between point 2 and 3).

When System A is recovered and then designated as Secondary (point 3), the data that was modified on B during A's outage is resynchronized (as indicated by the increasing I/Os that are flowing from B to A in the interval between point 3 and 4).

Finally, when the roles are reverted as they originally were (point 4), the I/O distribution returns as it was in the initial steady state phase.

It is important to notice that, while all of these developments were involving System A and System B, System C had a steady flow of replication I/Os coming from either systems, because one Active Asynchronous relationship always existed from the current Primary System.

7.2 Primary volume or CG loses target

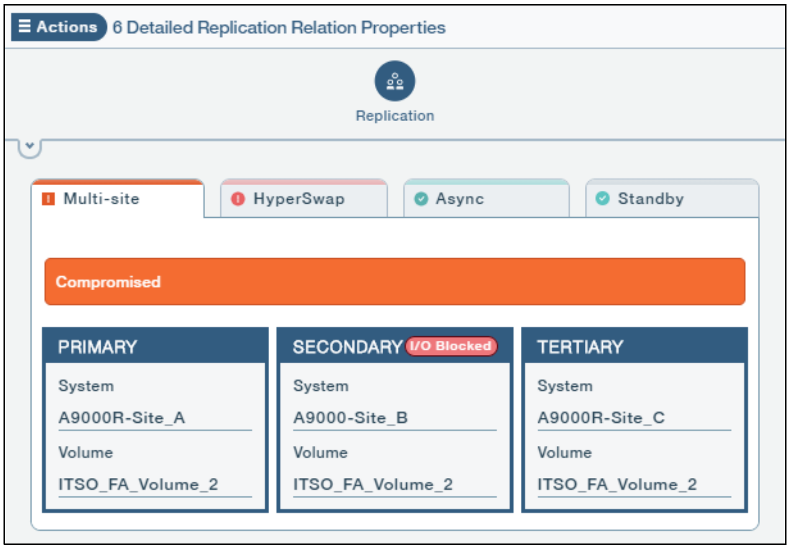
This section describes the scenarios in which System A loses either the Fibre Channel (FC) connection to System B, or the iSCSI connection to System C.

7.2.1 Losing Secondary volume or CG

This scenario considers an FC connectivity loss between Systems A and B, which as a result inhibits the automatic failover capability.

The following sequence of events occurs when connectivity between System A and System B fails:

1. Quorum Witness is up and running, so System A and System B poll the Quorum Witness to check if the other storage system (System B or System A) is up and running.
2. Because Quorum Witness can still talk to both System A and System B, it keeps the volume on System A in primary role and places System B in I/O blocked status, as shown in Figure 7-18.
3. The Multi-site relationship status changes to Compromised, as shown in Figure 7-18.



Compromised

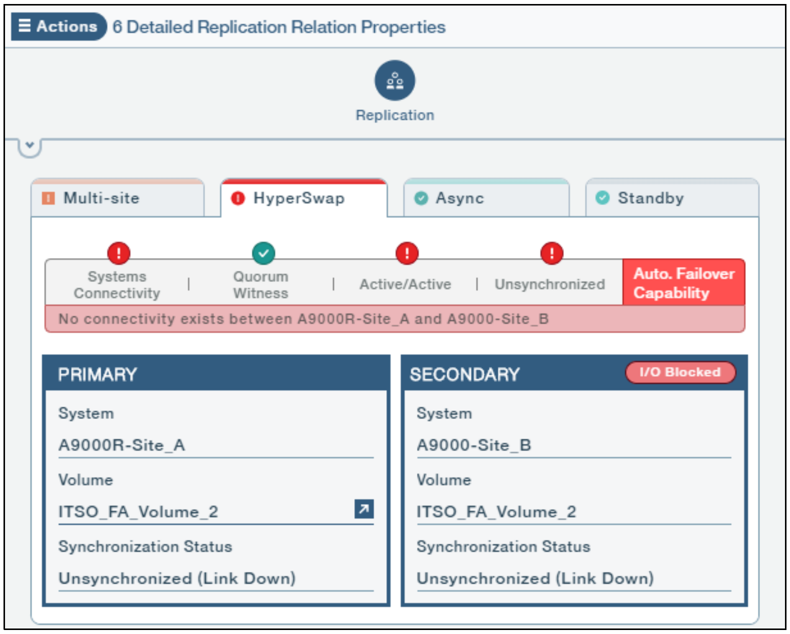
PRIMARY
System
A9000R-Site_A
Volume
ITSO_FA_Volume_2

SECONDARY I/O Blocked
System
A9000-Site_B
Volume
ITSO_FA_Volume_2

TERTIARY
System
A9000R-Site_C
Volume
ITSO_FA_Volume_2

Figure 7-18 System B I/O Blocked

The HyperSwap relationship details can be viewed in the HyperSwap tab of the Detailed Replication Relation Properties panel, as shown in Figure 7-19.



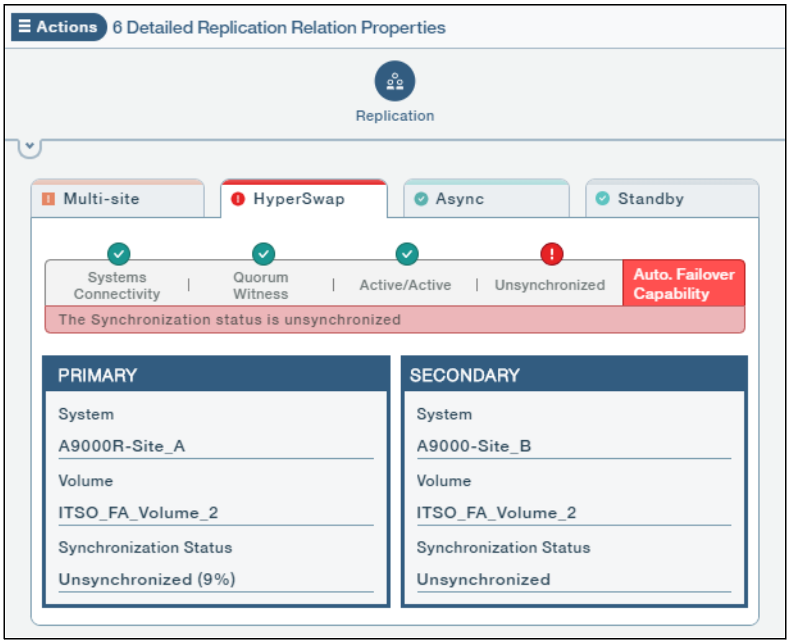
No connectivity exists between A9000R-Site_A and A9000-Site_B

PRIMARY	SECONDARY
System A9000R-Site_A	System A9000-Site_B
Volume ITSO_FA_Volume_2	Volume ITSO_FA_Volume_2
Synchronization Status Unsynchronized (Link Down)	Synchronization Status Unsynchronized (Link Down)

Figure 7-19 HyperSwap relationship details after Systems A and B connection loss

When the connectivity between System A and System B is restored, an automatic data synchronization between System A and System B occurs first and HyperSwap is then automatically re-enabled.

Figure 7-20 shows the GUI view when the connectivity between System A and System B is restored and data is being synchronized between both storage systems.

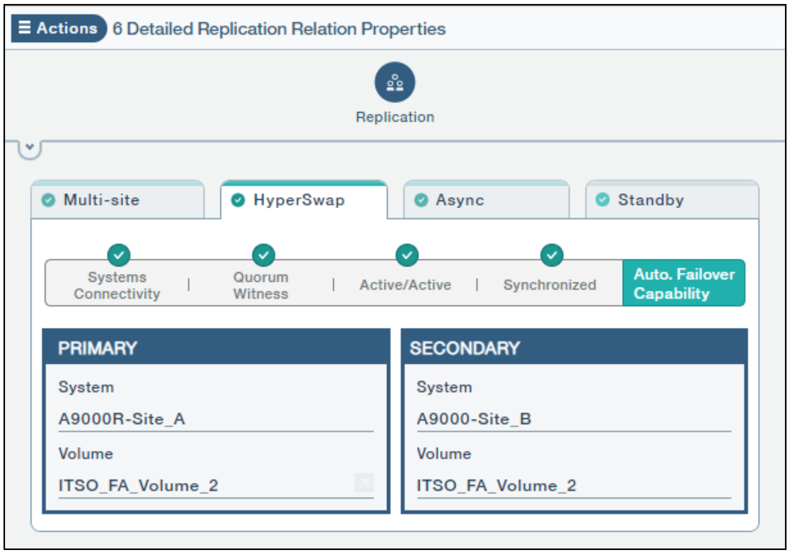


The Synchronization status is unsynchronized

PRIMARY	SECONDARY
System A9000R-Site_A	System A9000-Site_B
Volume ITSO_FA_Volume_2	Volume ITSO_FA_Volume_2
Synchronization Status Unsynchronized (9%)	Synchronization Status Unsynchronized

Figure 7-20 HyperSwap relationship details during A-B data synchronization

After the volume synchronization completes, the GUI shows that the high availability is reestablished and an automatic failover is now possible again, as shown in Figure 7-21.



PRIMARY

System

A9000R-Site_A

Volume

ITSO_FA_Volume_2

SECONDARY

System

A9000-Site_B

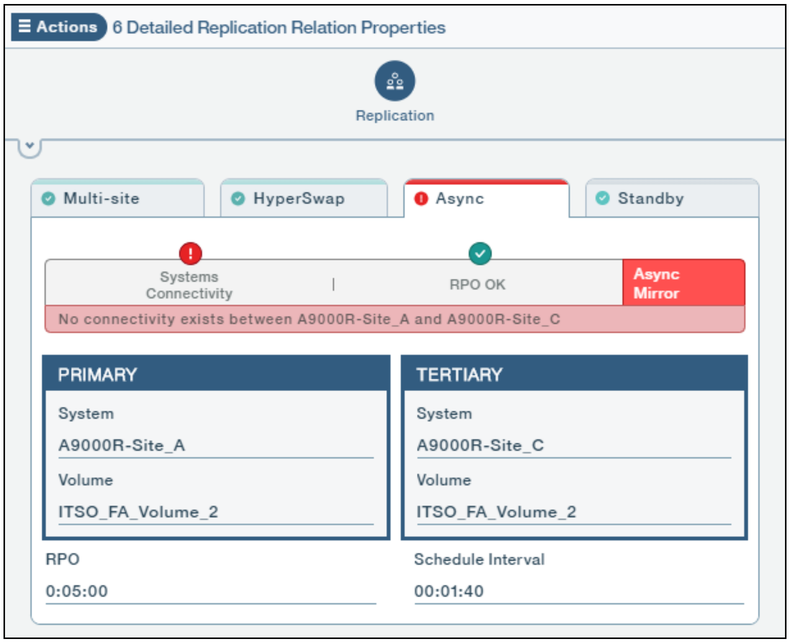
Volume

ITSO_FA_Volume_2

Figure 7-21 Automatic failover reenabled

7.2.2 Losing Tertiary volume or CG

This scenario considers an iSCSI connectivity loss between Systems A and C, which disables the active asynchronous mirroring relationship, as shown in Figure 7-22.



No connectivity exists between A9000R-Site_A and A9000R-Site_C

PRIMARY

System

A9000R-Site_A

Volume

ITSO_FA_Volume_2

TERTIARY

System

A9000R-Site_C

Volume

ITSO_FA_Volume_2

RPO

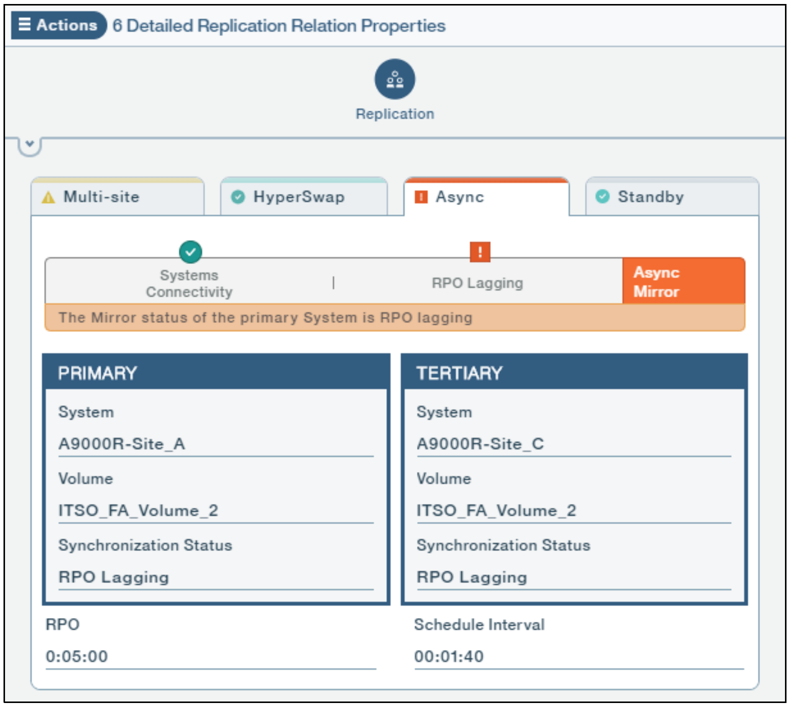
0:05:00

Schedule Interval

00:01:40

Figure 7-22 Asynchronous mirroring relationship details after Systems A and C connection loss

When the connectivity between System A and System C is reactivated, System C is automatically synchronized with the data changes that occurred at System A while the connection was down. Data changes are all contained in System A's most recent snapshots. Until synchronization from the most recent snapshot is completed, the asynchronous mirror remains in RPO Lagging state, as shown in Figure 7-23.



The Mirror status of the primary System is RPO lagging

PRIMARY

System
A9000R-Site_A

Volume
ITSO_FA_Volume_2

Synchronization Status
RPO Lagging

TERTIARY

System
A9000R-Site_C

Volume
ITSO_FA_Volume_2

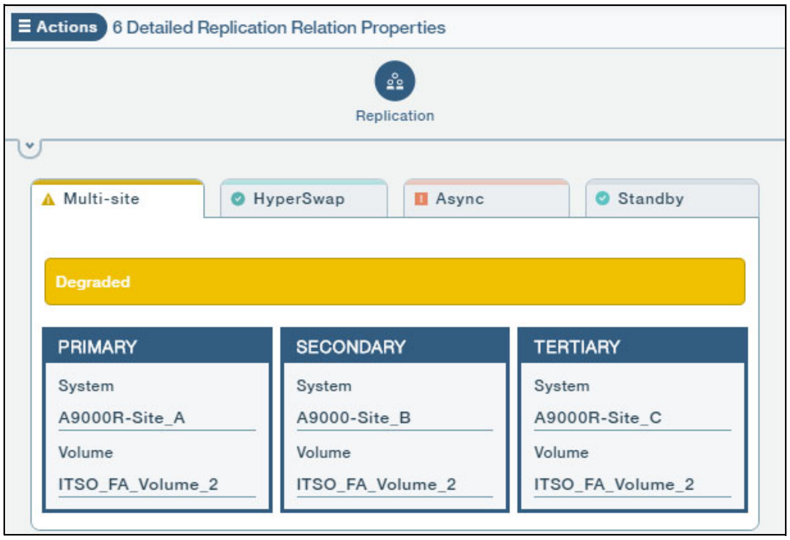
Synchronization Status
RPO Lagging

RPO
0:05:00

Schedule Interval
00:01:40

Figure 7-23 Asynchronous mirroring recovery in RPO Lagging state

During synchronization, the Multi-site relationship status is Degraded, as shown in Figure 7-24.



Degraded

PRIMARY

System
A9000R-Site_A

Volume
ITSO_FA_Volume_2

SECONDARY

System
A9000-Site_B

Volume
ITSO_FA_Volume_2

TERTIARY

System
A9000R-Site_C

Volume
ITSO_FA_Volume_2

Figure 7-24 Multi-site status during System C synchronization

After some time, synchronization is complete, and the asynchronous mirroring is reestablished.

7.3 DR test at secondary volume site

The procedure for testing the disaster recovery situation is similar whether a single volume or volumes that are part of a consistency group are used. In this section, we focus on the most likely scenario by using a consistency group.

Typically, the testing is run at the DR site, which is site C in the case of a Multi-site configuration.

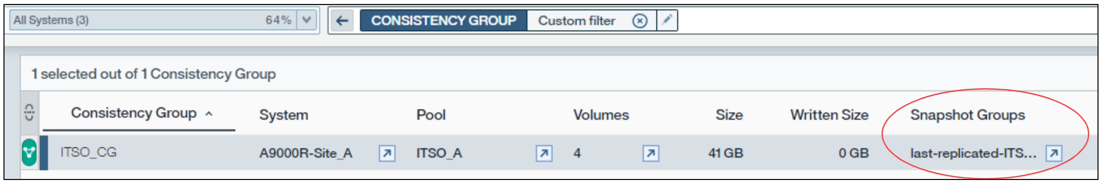
Asynchronous replication uses snapshots as a way to determine what data needs to be copied to the secondary system. After it is copied, the internal snapshot on the secondary is renamed as the *Last Replicated Snapshot (LRS)*. This snapshot is an internal snapshot and cannot be modified or deleted by the user. For more information about how snapshots are used for asynchronous replication, see *IBM FlashSystem A9000 and A9000R Business Continuity Solutions*, REDP-5401.

To perform the test based on the LRS, the single volume or the LRS snapshot group for a consistency group is duplicated. After the LRS volume or CG is duplicated, the user can unlock the snapshots that are part of a group, map to a host, and perform testing as needed.

Next, we describe those steps by using the Hyper-Scale Manager.

If you know which volumes are part of the consistency group, you can navigate to the Snapshots or Snapshot Groups view and select each snapshot.

Alternatively, you can start from the Consistency Group view that is found under the Pools and Volumes View. Select the consistency group and then, select the arrow that is under the Snapshot Groups column heading to navigate to the associated snapshot group, as shown in Figure 7-25.



1 selected out of 1 Consistency Group

Consistency Group	System	Pool	Volumes	Size	Written Size	Snapshot Groups
ITSO_CG	A9000R-Site_A	ITSO_A	4	41 GB	0 GB	last-replicated-ITS...

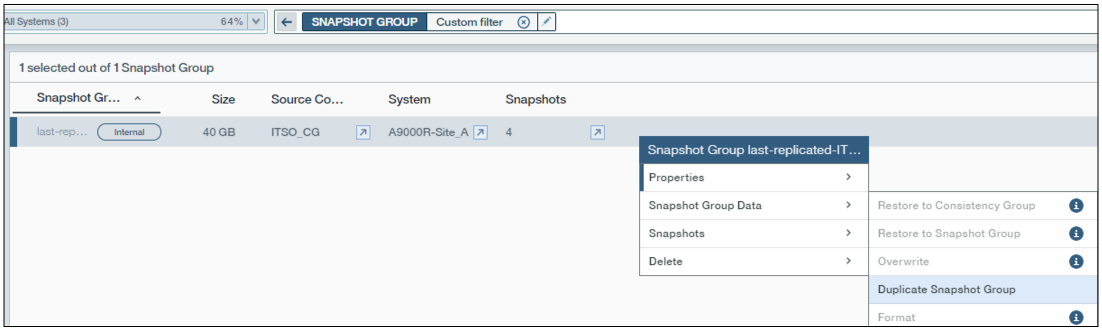
Figure 7-25 Navigating to the snapshot group

The selected snapshot group is the only snapshot that is displayed in the table. Notice that the name begins with “last-replicated” and includes an “internal” bubble next to the name.

That LRS is greyed out, which indicates that you cannot open, delete, or modify it in any way while there is an active asynchronous relationship associated with the corresponding consistency group.

However, if you right-click the snapshot group, or open the Actions menu, a few options are available. The list of actions includes duplicating and restoring, which is found under Snapshot Group Data. An example is shown in Figure 7-26 on page 144.

For the purposes of testing at the DR site, you select **Duplicate Snapshot Group**.



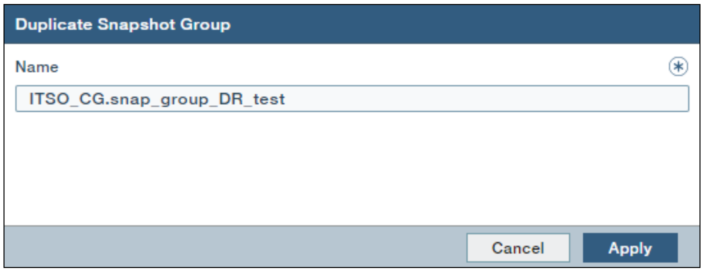
1 selected out of 1 Snapshot Group

Snapshot Gr…	Size	Source Co…	System	Snapshots
last-rep… Internal	40 GB	ITSO_CG	A9000R-Site_A	4

Figure 7-26 Duplicate snapshot group

A new window opens, in which you can change the default name of the duplicate before creating it, as shown in Figure 7-27. Names can also be changed later, if needed.

Click **Apply** to create the duplicate snapshot group.

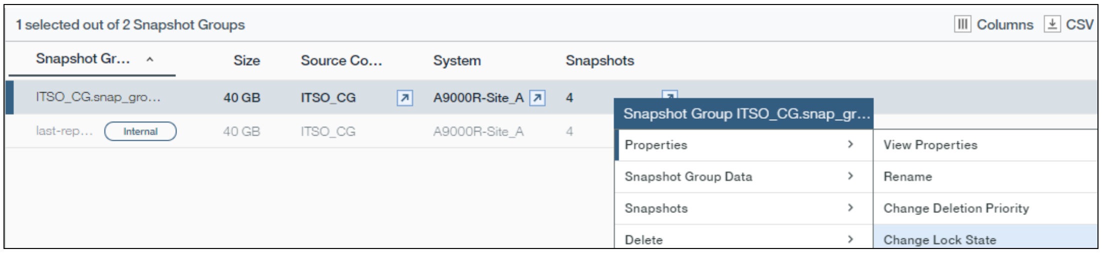


Name: ITSO_CG.snap_group_DR_test

Figure 7-27 Creating a duplicate

You now see the snapshot group you created and the internal LRS snapshot group in the table view. To use any snapshot within the group, you must first unlock it for writing (although a host can read from a locked snapshot, it cannot be changed if the snapshot remains locked).

To unlock the snapshot, right-click the duplicate or select the **Actions** menu followed by **Change Lock State**, as shown in Figure 7-28.

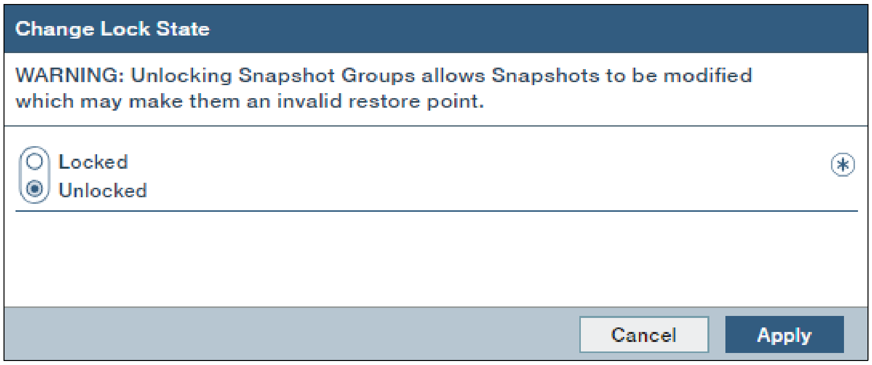


1 selected out of 2 Snapshot Groups

Snapshot Gr…	Size	Source Co…	System	Snapshots
ITSO_CG.snap_gro…	40 GB	ITSO_CG	A9000R-Site_A	4
last-rep… Internal	40 GB	ITSO_CG	A9000R-Site_A	4

Figure 7-28 Unlocking the snapshot group

You are prompted to confirm the lock state change, as shown in Figure 7-29.



Change Lock State

WARNING: Unlocking Snapshot Groups allows Snapshots to be modified which may make them an invalid restore point.

○ Locked
● Unlocked

Figure 7-29 Confirming unlocking snapshot group

After they are unlocked, the snapshots can be mapped to a host for testing. With the volumes selected, select **Mapping** → **View/Modify Mapping** under the Actions menu. Click ADD for host or cluster, depending on the environment (see Figure 7-30).

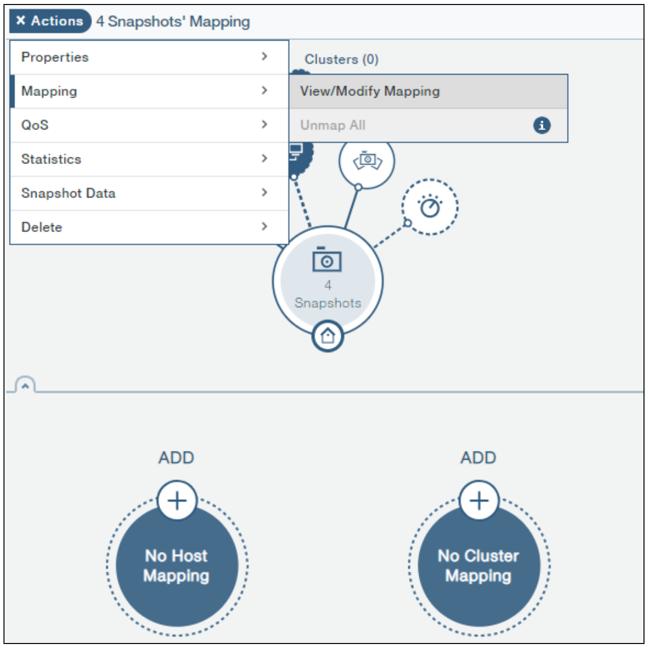
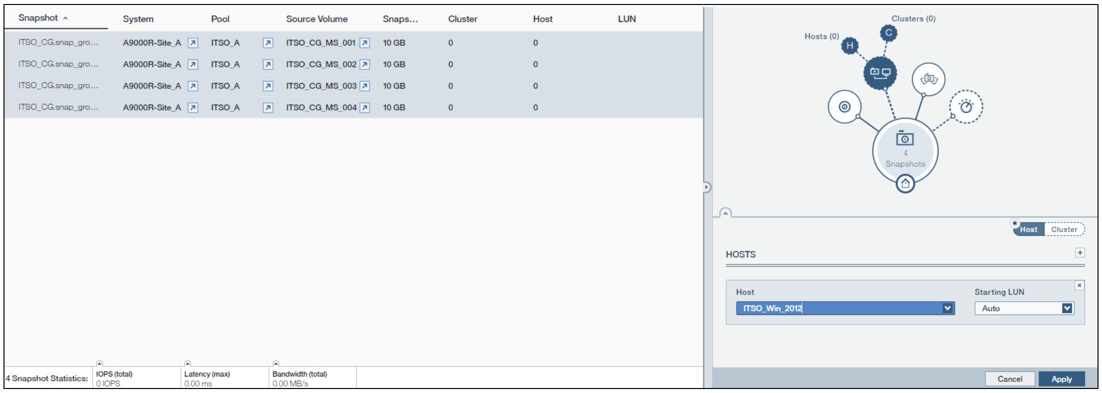


Figure 7-30 Map the snapshots

This selection allows you to map multiple volumes as the same time to a defined host, as shown in Figure 7-31.



Snapshot ^	System	Pool	Source Volume	Snaps...	Cluster	Host	LUN
ITSO_CG.snap_gro...	A9000R-Site_A	ITSO_A	ITSO_CG_MS_001	10 GB	0	0	
ITSO_CG.snap_gro...	A9000R-Site_A	ITSO_A	ITSO_CG_MS_002	10 GB	0	0	
ITSO_CG.snap_gro...	A9000R-Site_A	ITSO_A	ITSO_CG_MS_003	10 GB	0	0	
ITSO_CG.snap_gro...	A9000R-Site_A	ITSO_A	ITSO_CG_MS_004	10 GB	0	0	

4 Snapshot Statistics: IOPS (total) 0 IOPS; Latency (max) 0.00 ms; Bandwidth (total) 0.00 MB/s

HOSTS

Host: ITSO_Win_2012

Starting LUN: Auto

Figure 7-31   Host selection for mapping

If the host is not yet defined, you can create one at this point.

For the purposes of this example, we selected a host from the list and then clicked **Apply**.

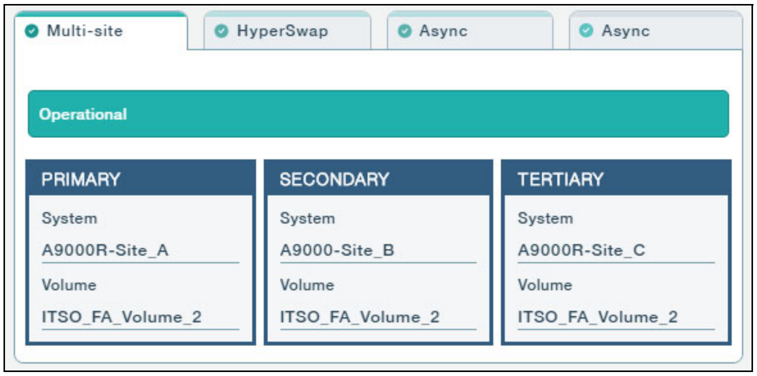
The testing of the DR copies can now be done to ensure that you recover from the secondary (or tertiary) site if the primary site fails (or the HyperSwap systems in the case of Multi-Site HA/DR).

7.4 Primary and Secondary volume or CG failure

This scenario considers a failure of System A, followed by a failure of System B.

7.4.1 Systems A and B failure scenario

We assume that the Multi-site configuration that was done before the failures was working and fully operational, with a HyperSwap connection between Systems A and B, an active asynchronous mirroring connection between Systems A and C, and a standby asynchronous mirroring connection between Systems B and C, as shown in Figure 7-32.

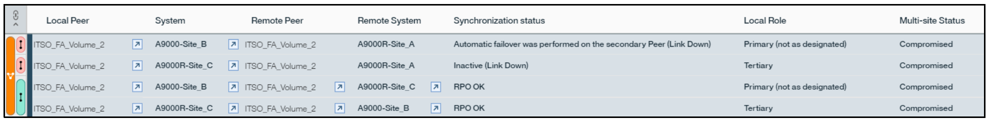


Operational

PRIMARY	SECONDARY	TERTIARY
System	System	System
A9000R-Site_A	A9000-Site_B	A9000R-Site_C
Volume	Volume	Volume
ITSO_FA_Volume_2	ITSO_FA_Volume_2	ITSO_FA_Volume_2

Figure 7-32   Multi-site status before failures

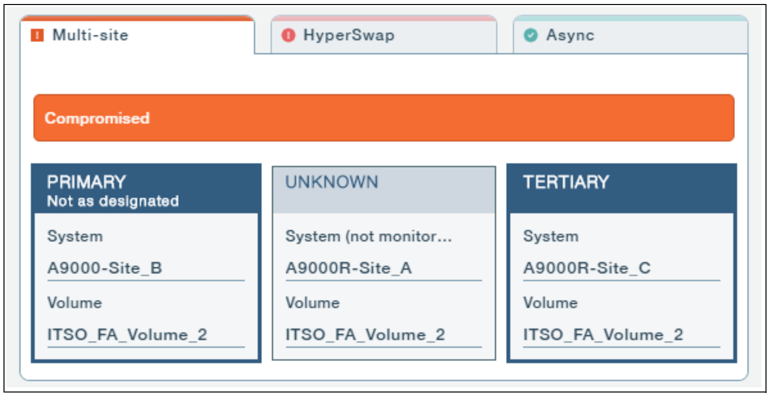
When System A fails, it disappears from the replication relationship details in the Hyper-Scale Manager HGUI (see Figure 7-33), and an automatic failover is performed to the Secondary volume on System B.



Local Peer	System	Remote Peer	Remote System	Synchronization status	Local Role	Multi-site Status
ITSO_FA_Volume_2	A9000-Site_B	ITSO_FA_Volume_2	A9000R-Site_A	Automatic failover was performed on the secondary Peer (Link Down)	Primary (not as designated)	Compromised
ITSO_FA_Volume_2	A9000R-Site_C	ITSO_FA_Volume_2	A9000R-Site_A	Inactive (Link Down)	Tertiary	Compromised
ITSO_FA_Volume_2	A9000-Site_B	ITSO_FA_Volume_2	A9000R-Site_C	RPO OK	Primary (not as designated)	Compromised
ITSO_FA_Volume_2	A9000R-Site_C	ITSO_FA_Volume_2	A9000-Site_B	RPO OK	Tertiary	Compromised

Figure 7-33 Replication relations details after System A failure

Now the multi-site configuration status is compromised, because System B is acting not as designated, and System A is temporarily not being monitored (see Figure 7-34).



Compromised

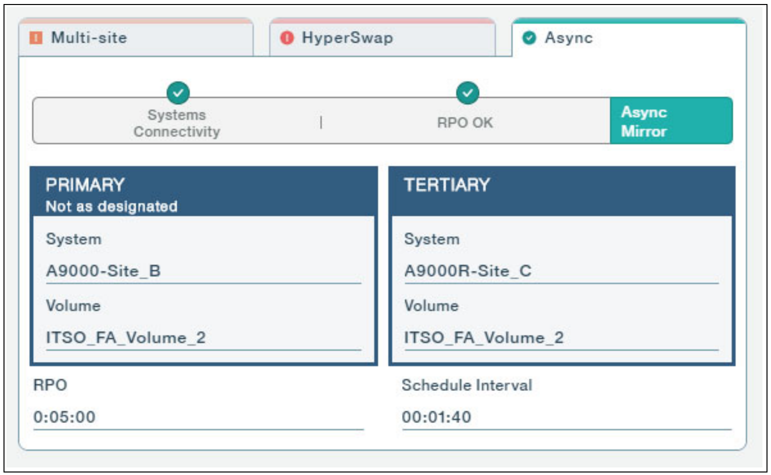
PRIMARY
Not as designated
System
A9000-Site_B
Volume
ITSO_FA_Volume_2

UNKNOWN
System (not monitor...
A9000R-Site_A
Volume
ITSO_FA_Volume_2

TERTIARY
System
A9000R-Site_C
Volume
ITSO_FA_Volume_2

Figure 7-34 Multi-site relationship status after failover on System B

In addition, upon the failure of System A, the Standby asynchronous mirror between Systems B and C is automatically activated. Its details now appear in the Async tab of the replication relation properties, as shown in Figure 7-35.



PRIMARY
Not as designated
System
A9000-Site_B
Volume
ITSO_FA_Volume_2

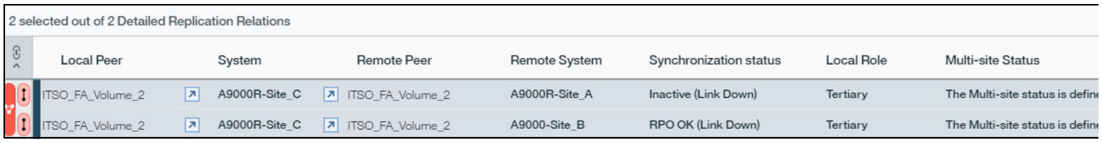
TERTIARY
System
A9000R-Site_C
Volume
ITSO_FA_Volume_2

RPO
0:05:00

Schedule Interval
00:01:40

Figure 7-35 Systems B and C asynchronous mirror status after failover on System B

Following the failure of System A, System B also fails. As shown in Figure 7-36, the only remaining storage system in the Multi-site configuration is System C.

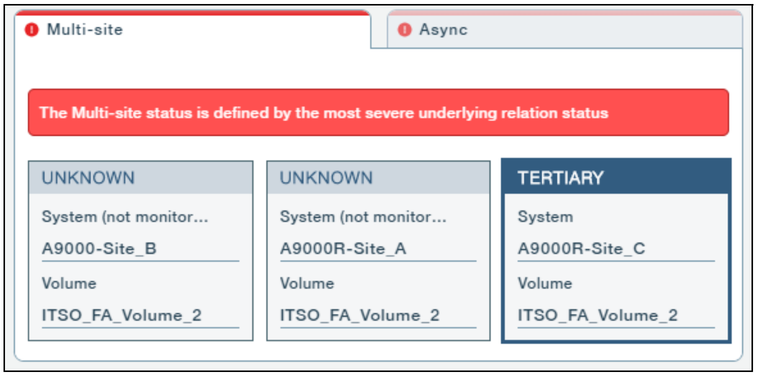


2 selected out of 2 Detailed Replication Relations

Local Peer	System	Remote Peer	Remote System	Synchronization status	Local Role	Multi-site Status
ITSO_FA_Volume_2	A9000R-Site_C	ITSO_FA_Volume_2	A9000R-Site_A	Inactive (Link Down)	Tertiary	The Multi-site status is define
ITSO_FA_Volume_2	A9000R-Site_C	ITSO_FA_Volume_2	A9000-Site_B	RPO OK (Link Down)	Tertiary	The Multi-site status is define

Figure 7-36 Replication relations details after Systems A and B failure

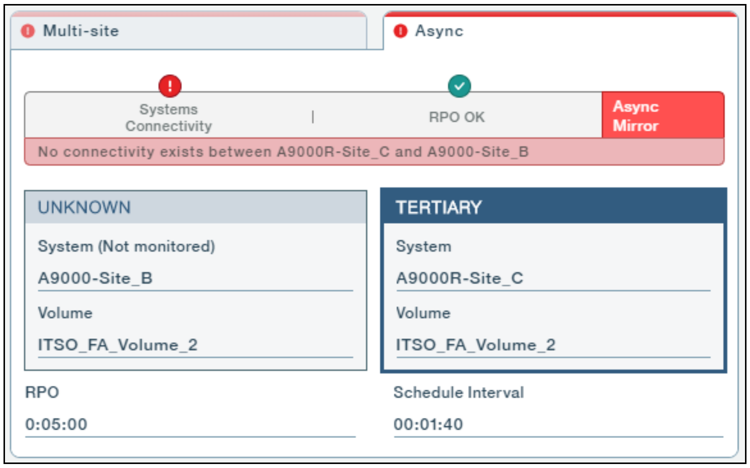
The Multi-site relationship is effectively destroyed (see Figure 7-37), and the asynchronous mirror between the Systems B and C is deactivated (see Figure 7-38).



The Multi-site status is defined by the most severe underlying relation status

UNKNOWN	UNKNOWN	TERTIARY
System (not monitor...	System (not monitor...	System
A9000-Site_B	A9000R-Site_A	A9000R-Site_C
Volume	Volume	Volume
ITSO_FA_Volume_2	ITSO_FA_Volume_2	ITSO_FA_Volume_2

Figure 7-37 Multi-site status after Systems A and B failure



No connectivity exists between A9000R-Site_C and A9000-Site_B

UNKNOWN	TERTIARY
System (Not monitored)	System
A9000-Site_B	A9000R-Site_C
Volume	Volume
ITSO_FA_Volume_2	ITSO_FA_Volume_2

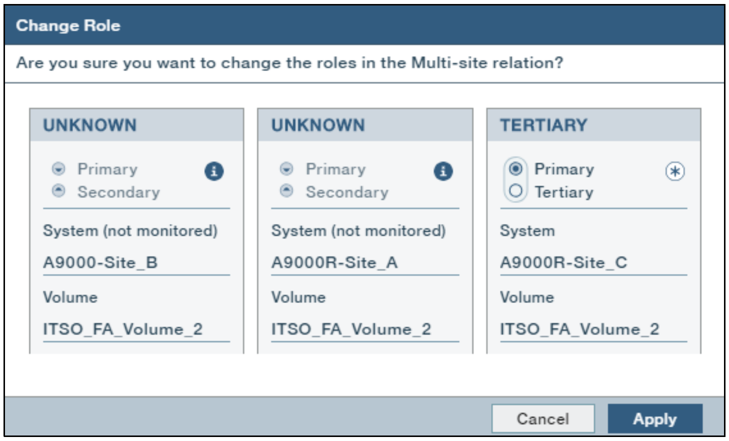
RPO: 0:05:00

Schedule Interval: 00:01:40

Figure 7-38 Systems B-C asynchronous mirror status after Systems A and B failure

7.4.2 Multi-site relationship recovery after Systems A and B failure

The recovery process of a Multi-site relationship after the failure of Systems A and B begins with changing the System C role from Tertiary to Primary (see Figure 7-39) by using the replication Actions menu (click **Failover/Recovery** → **Change Role**).



Change Role

Are you sure you want to change the roles in the Multi-site relation?

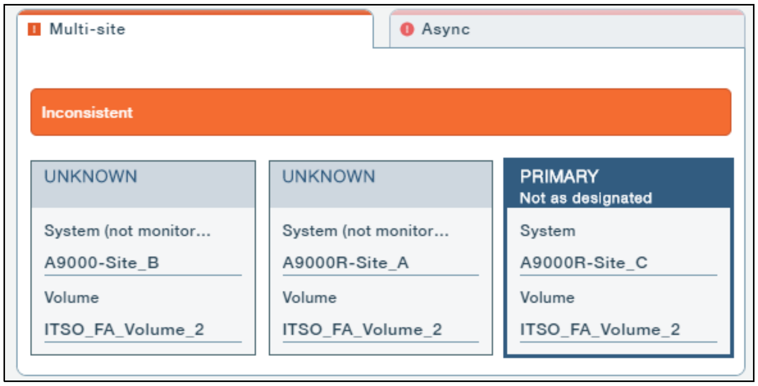
UNKNOWN
○ Primary
○ Secondary
System (not monitored)
A9000-Site_B
Volume
ITSO_FA_Volume_2

UNKNOWN
○ Primary
○ Secondary
System (not monitored)
A9000R-Site_A
Volume
ITSO_FA_Volume_2

TERTIARY
● Primary
○ Tertiary
System
A9000R-Site_C
Volume
ITSO_FA_Volume_2

Figure 7-39 Changing System C role to Primary

Because none of the Multi-site relationship peers are now performing their designated role, the Multi-site relationship status is inconsistent (see Figure 7-40).



Inconsistent

UNKNOWN
System (not monitor...
A9000-Site_B
Volume
ITSO_FA_Volume_2

UNKNOWN
System (not monitor...
A9000R-Site_A
Volume
ITSO_FA_Volume_2

PRIMARY
Not as designated
System
A9000R-Site_C
Volume
ITSO_FA_Volume_2

Figure 7-40 Multi-site status with the System C volume defined as Primary

At this point, the volume on System C must be mapped to start serving the I/O from the host.

When Systems A and B are back online, a role conflict occurs because each peer is now defined as Primary: System A because Primary is its designated role, System B because it became Primary after the System A failure, and System C because we changed its role to Primary (see Figure 7-41 on page 150).

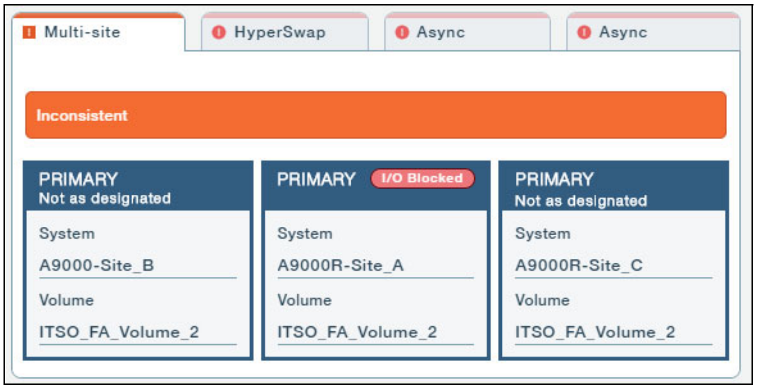


Figure 7-41 Role conflict with three Primary volumes

The volumes on Systems A and B (former Primary and Secondary volumes) are now out-of-date and must be synchronized with the data that is on System C. While the host workload is directed to System C, the volumes on Systems A and B can be safely unmapped and the blocked I/O can be fixed on System A (see Figure 7-42).

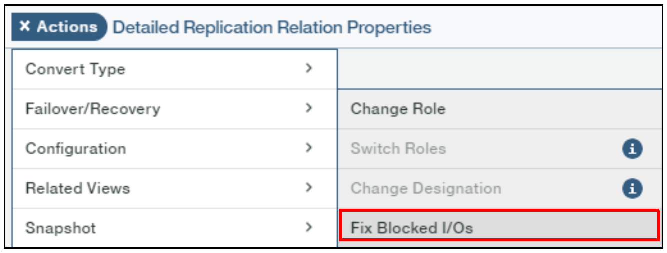


Figure 7-42 Fixing blocked I/O

The next step is to clean up the configuration to an original state by deleting the Multi-site relationships from each of the three systems. This procedure must be performed with CLI commands. Complete the following steps:

1. Delete the Multi-site relationship from System A by using the **multisite_delete** command (see Example 7-12).

Example 7-12 Deleting Multi-site on System A

```
A9000R-Site_A>>multisite_delete vol=ITSO_FA_Volume_2
Command executed successfully
```

2. Repeat the command on System B (see Example 7-13).

Example 7-13 Deleting Multi-site on System B

```
A9000-Site_B>>multisite_delete vol=ITSO_FA_Volume_2
Command executed successfully.
```

Take note of the use of the **force** flag on System C (see Example 7-14).

Example 7-14 Forcing the deletion of Multi-site on System C

```
A9000R-Site_C>>multisite_delete vol=ITSO_FA_Volume_2 force=yes
Command executed successfully.
```

3. Delete the HyperSwap relationship that is left over from the Multi-site on Systems A and B by using the **ha_delete** command (see Example 7-15).

Example 7-15 Deleting HyperSwap on System A and System B

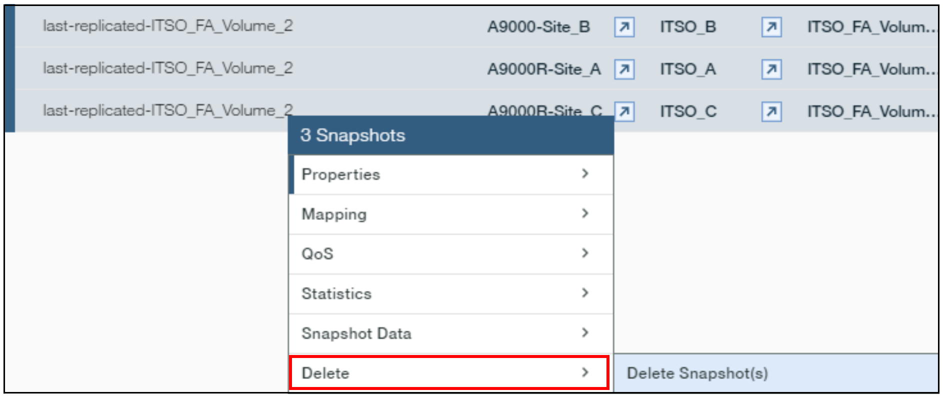
```
A9000R-Site_A>>ha_delete vol=ITSO_FA_Volume_2

Warning:   Are you sure you want to delete this HyperSwap relation? y/n: y
Command executed successfully.

A9000-Site_B>>ha_delete vol=ITSO_FA_Volume_2

Warning:   Are you sure you want to delete this HyperSwap relation? y/n: y
Command executed successfully.
```

4. Delete the left over Asynchronous mirroring relationships by issuing a **mirror_delete** command on all three Systems, as shown in Example 7-16.

Example 7-16

```
A9000R-Site_A>>mirror_delete vol=ITSO_FA_Volume_2 target=A9000R-Site_C

Warning:   Are you sure you want to delete this mirroring relation? y/n: y
Command executed successfully.

A9000-Site_B>>mirror_delete vol=ITSO_FA_Volume_2 target=A9000R-Site_C

Warning:   Are you sure you want to delete this mirroring relation? y/n: y
Command executed successfully.

A9000R-Site_C>>mirror_delete vol=ITSO_FA_Volume_2 target=A9000R-Site_A

Warning:   Are you sure you want to delete this mirroring relation? y/n: y
Command executed successfully.
A9000R-Site_C>>mirror_delete vol=ITSO_FA_Volume_2 target=A9000-Site_B

Warning:   Are you sure you want to delete this mirroring relation? y/n: y
Command executed successfully.
```

Delete the snapshots from each of the three systems, as shown in Figure 7-43.

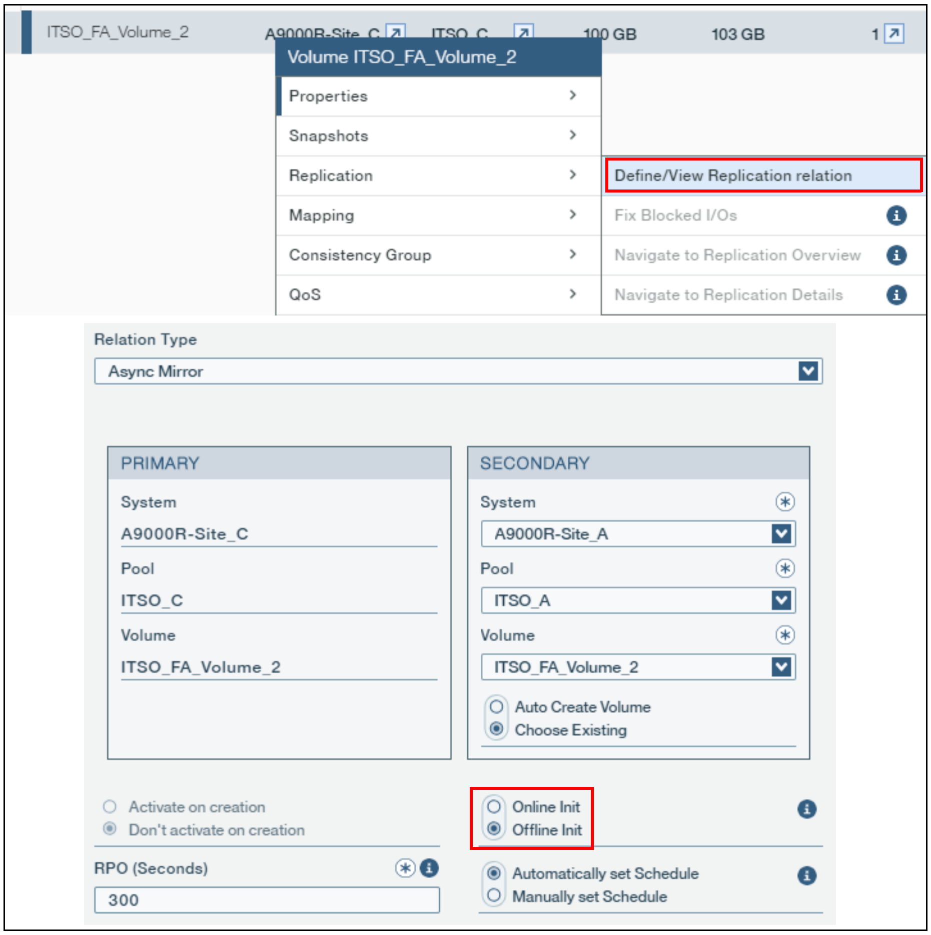


last-replicated-ITSO_FA_Volume_2	A9000-Site_B	ITSO_B	ITSO_FA_Volum..
last-replicated-ITSO_FA_Volume_2	A9000R-Site_A	ITSO_A	ITSO_FA_Volum..
last-replicated-ITSO_FA_Volume_2	A9000R-Site_C	ITSO_C	ITSO_FA_Volum..

Figure 7-43 Deleting snapshots

From this point on, we can restore the Multi-site relationship.

To restore the Multi-site relationship, start by creating an asynchronous mirroring relationship between C and A with offline initialization (as shown in Figure 7-44) to transfer data from System C to System A.



Relation Type: Async Mirror

PRIMARY	SECONDARY
System: A9000R-Site_C	System: A9000R-Site_A
Pool: ITSO_C	Pool: ITSO_A
Volume: ITSO_FA_Volume_2	Volume: ITSO_FA_Volume_2
	○ Auto Create Volume ● Choose Existing

○ Activate on creation
● Don't activate on creation

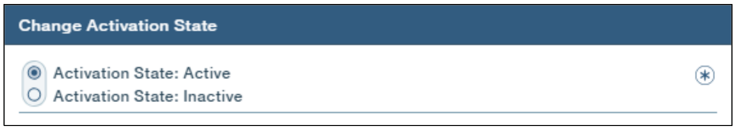
○ Online Init
● Offline Init

RPO (Seconds): 300

● Automatically set Schedule
○ Manually set Schedule

Figure 7-44 Creating an asynchronous mirroring relationship between A and C

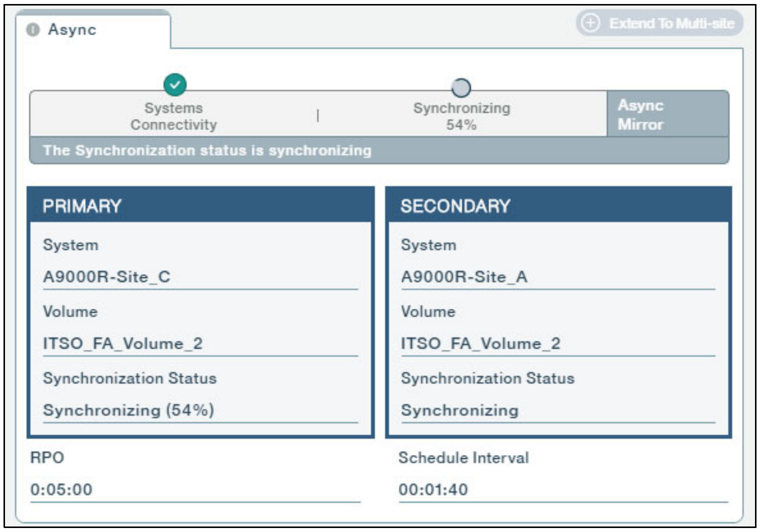
After the C-A asynchronous mirror is created, activate it (see Figure 7-45).


Change Activation State

● Activation State: Active
○ Activation State: Inactive

Figure 7-45 Activating the asynchronous mirror

Wait for the synchronization to complete (see Figure 7-46).



The Synchronization status is synchronizing

PRIMARY

System

A9000R-Site_C

Volume

ITSO_FA_Volume_2

Synchronization Status

Synchronizing (54%)

SECONDARY

System

A9000R-Site_A

Volume

ITSO_FA_Volume_2

Synchronization Status

Synchronizing

RPO

0:05:00

Schedule Interval

00:01:40

Figure 7-46 C-A synchronization in progress

The process to rebuild the Multi-site relationship begins at this point. Consider shutting down your disaster recovery hosts to suspend data changes until they are brought over to System A.

Switch roles between Systems A and C and change their designations (see Figure 7-47) to make System A Primary and System C Secondary (see Figure 7-48).

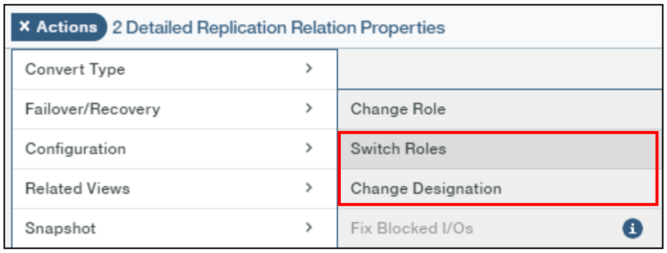
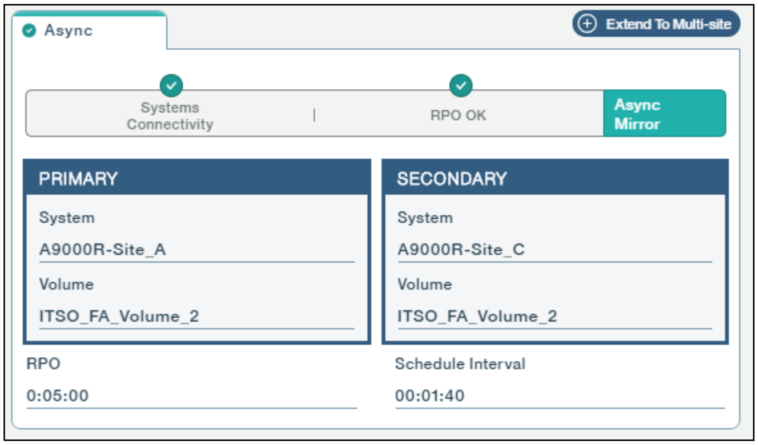


Figure 7-47 Switching roles and changing designations



PRIMARY

System

A9000R-Site_A

Volume

ITSO_FA_Volume_2

SECONDARY

System

A9000R-Site_C

Volume

ITSO_FA_Volume_2

RPO

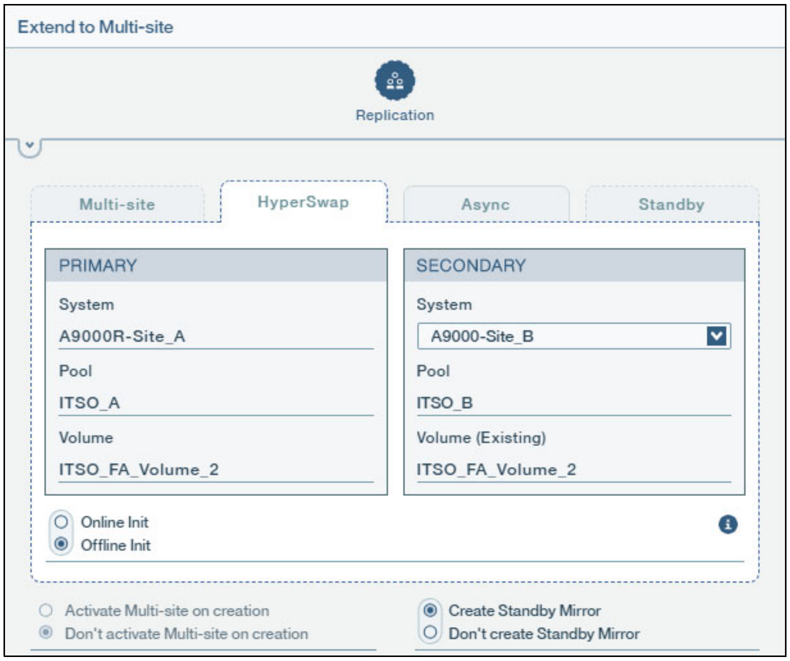
0:05:00

Schedule Interval

00:01:40

Figure 7-48 A-C asynchronous mirror status after swapping roles

Extend the A-C asynchronous mirroring relationship to Multi-site by using offline initialization, as shown in Figure 7-49.



PRIMARY

System

A9000R-Site_A

Pool

ITSO_A

Volume

ITSO_FA_Volume_2

SECONDARY

System

A9000-Site_B

Pool

ITSO_B

Volume (Existing)

ITSO_FA_Volume_2

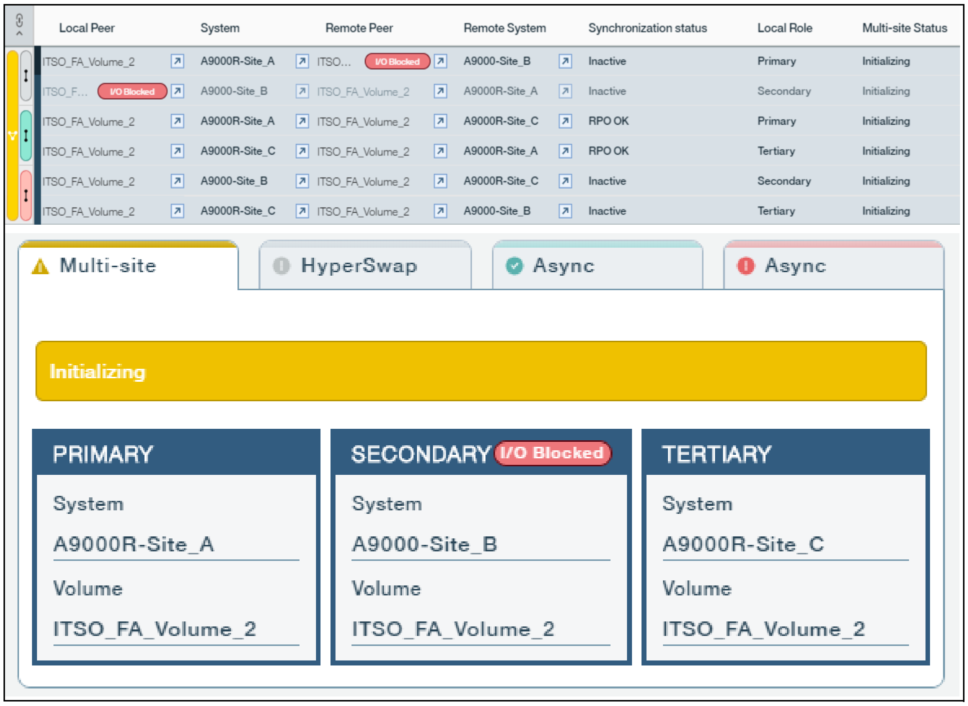
○ Online Init
● Offline Init

○ Activate Multi-site on creation
● Don't activate Multi-site on creation

● Create Standby Mirror
○ Don't create Standby Mirror

Figure 7-49 Extending the A-C asynchronous mirroring relationship to Multi-site

When the Multi-site relationship is created, its status is shown is Initializing, as shown in Figure 7-50.



Local Peer	System	Remote Peer	Remote System	Synchronization status	Local Role	Multi-site Status
ITSO_FA_Volume_2	A9000R-Site_A	ITSO… I/O Blocked	A9000-Site_B	Inactive	Primary	Initializing
ITSO_F… I/O Blocked	A9000-Site_B	ITSO_FA_Volume_2	A9000R-Site_A	Inactive	Secondary	Initializing
ITSO_FA_Volume_2	A9000R-Site_A	ITSO_FA_Volume_2	A9000R-Site_C	RPO OK	Primary	Initializing
ITSO_FA_Volume_2	A9000R-Site_C	ITSO_FA_Volume_2	A9000R-Site_A	RPO OK	Tertiary	Initializing
ITSO_FA_Volume_2	A9000-Site_B	ITSO_FA_Volume_2	A9000R-Site_C	Inactive	Secondary	Initializing
ITSO_FA_Volume_2	A9000R-Site_C	ITSO_FA_Volume_2	A9000-Site_B	Inactive	Tertiary	Initializing

Initializing

PRIMARY

System

A9000R-Site_A

Volume

ITSO_FA_Volume_2

SECONDARY I/O Blocked

System

A9000-Site_B

Volume

ITSO_FA_Volume_2

TERTIARY

System

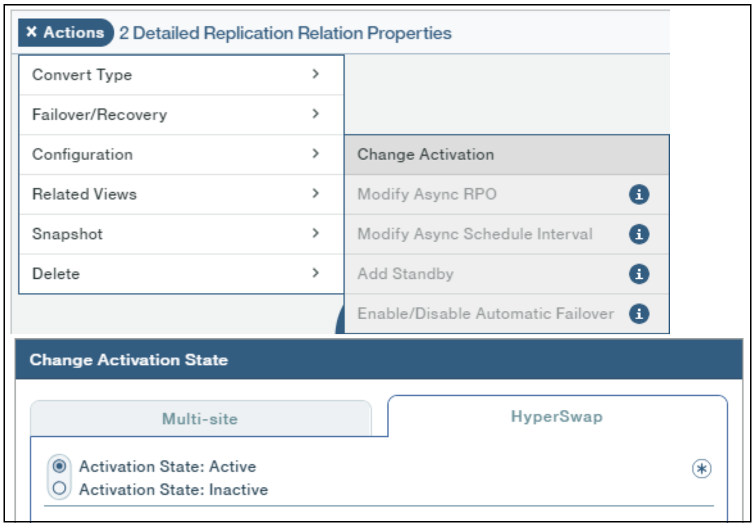
A9000R-Site_C

Volume

ITSO_FA_Volume_2

Figure 7-50 Multi-site relationship status after extending the asynchronous mirror

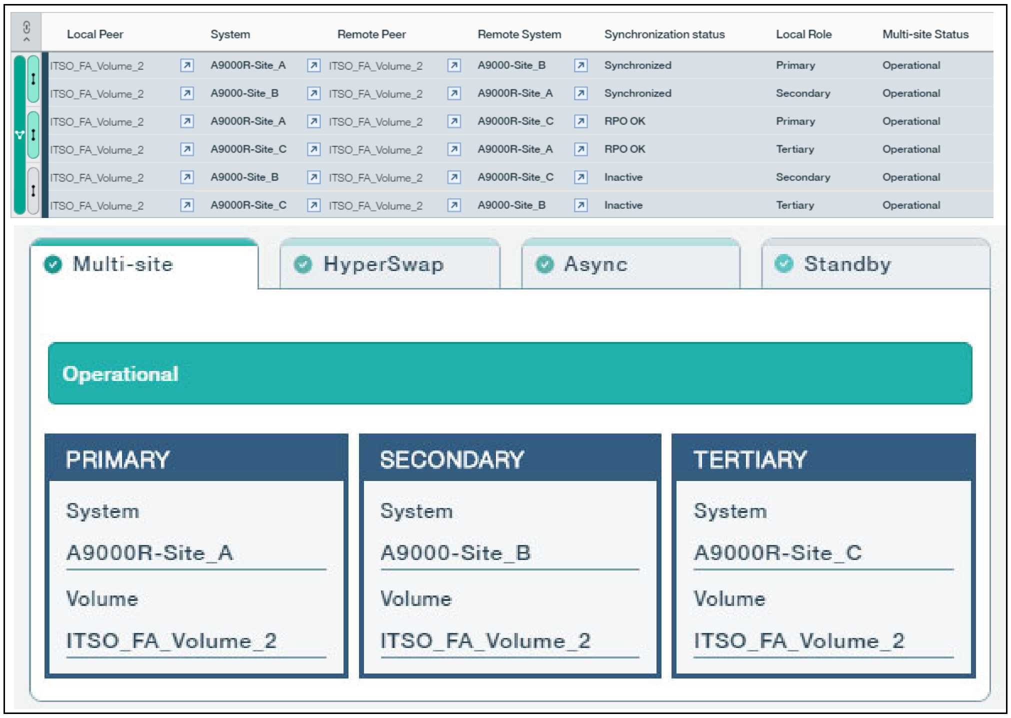
Finally, activate a HyperSwap relationship between A and B (see Figure 7-51).



● Activation State: Active
○ Activation State: Inactive

Figure 7-51 Activating the A-B HyperSwap relationship

Activating the HyperSwap relationship starts the synchronization of all Multi-site relationship components, including the Standby asynchronous relationship between Systems B and C. Upon completion of synchronization, the restored Multi-site relationship are fully operational, as shown in Figure 7-52.



Local Peer	System	Remote Peer	Remote System	Synchronization status	Local Role	Multi-site Status
ITSO_FA_Volume_2	A9000R-Site_A	ITSO_FA_Volume_2	A9000-Site_B	Synchronized	Primary	Operational
ITSO_FA_Volume_2	A9000-Site_B	ITSO_FA_Volume_2	A9000R-Site_A	Synchronized	Secondary	Operational
ITSO_FA_Volume_2	A9000R-Site_A	ITSO_FA_Volume_2	A9000R-Site_C	RPO OK	Primary	Operational
ITSO_FA_Volume_2	A9000R-Site_C	ITSO_FA_Volume_2	A9000R-Site_A	RPO OK	Tertiary	Operational
ITSO_FA_Volume_2	A9000-Site_B	ITSO_FA_Volume_2	A9000R-Site_C	Inactive	Secondary	Operational
ITSO_FA_Volume_2	A9000R-Site_C	ITSO_FA_Volume_2	A9000-Site_B	Inactive	Tertiary	Operational

Operational

PRIMARY	SECONDARY	TERTIARY
System	System	System
A9000R-Site_A	A9000-Site_B	A9000R-Site_C
Volume	Volume	Volume
ITSO_FA_Volume_2	ITSO_FA_Volume_2	ITSO_FA_Volume_2

Figure 7-52 Restored Multi-site relationship

7.5 Quorum Witness failures

Issues with Quorum Witness connectivity on one or more systems that are involved in a Multi-site relationship result in compromised high availability. This issue affects the HyperSwap leg of the relationship as described 5.3, “Quorum Witness failure scenarios” on page 69.

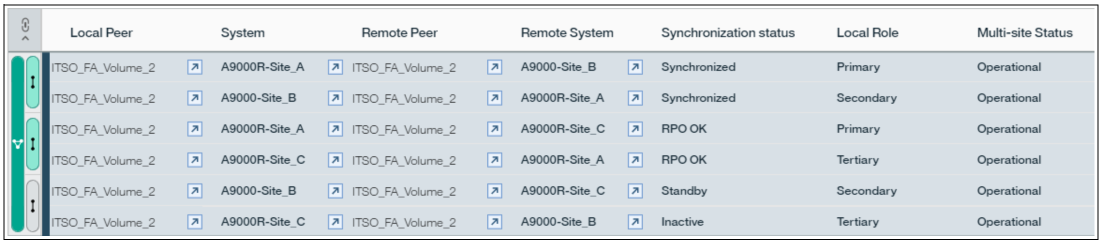
In this section, we describe two Quorum Witness failure scenarios that are similar to the scenarios that we saw for a simple HyperSwap relationship; in this case, the HyperSwap is part of a Multi-site relationship.

In the first scenario, only the Primary system loses connectivity to the Quorum Witness. This situation still allows for Automatic Failover in case of system failure, and recovery is straightforward.

In the second scenario, the Primary and the Secondary system lose connectivity to the Quorum Witness. This scenario is more complex and involves manual intervention in case of rolling disaster.

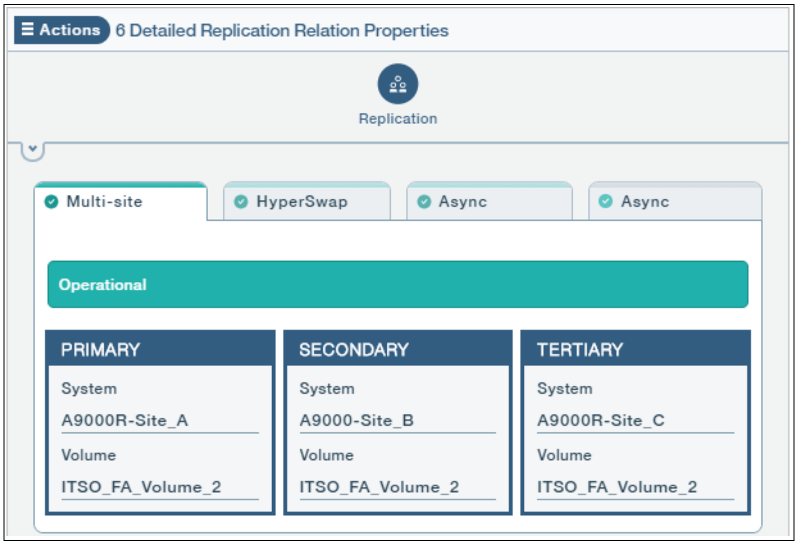
7.5.1 Loss of connection between the Primary system and the Quorum Witness

In this example, System A is the Primary in a Multi-site environment that involves System B (Secondary) and System C (Tertiary). The initial configuration is shown in Figure 7-53 and Figure 7-54.



Local Peer	System	Remote Peer	Remote System	Synchronization status	Local Role	Multi-site Status
ITSO_FA_Volume_2	A9000R-Site_A	ITSO_FA_Volume_2	A9000-Site_B	Synchronized	Primary	Operational
ITSO_FA_Volume_2	A9000-Site_B	ITSO_FA_Volume_2	A9000R-Site_A	Synchronized	Secondary	Operational
ITSO_FA_Volume_2	A9000R-Site_A	ITSO_FA_Volume_2	A9000R-Site_C	RPO OK	Primary	Operational
ITSO_FA_Volume_2	A9000R-Site_C	ITSO_FA_Volume_2	A9000R-Site_A	RPO OK	Tertiary	Operational
ITSO_FA_Volume_2	A9000-Site_B	ITSO_FA_Volume_2	A9000R-Site_C	Standby	Secondary	Operational
ITSO_FA_Volume_2	A9000R-Site_C	ITSO_FA_Volume_2	A9000-Site_B	Inactive	Tertiary	Operational

Figure 7-53 Multi-site relationship configuration for our example

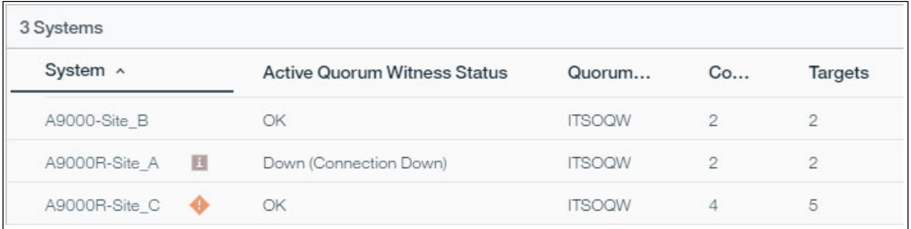


Operational

PRIMARY	SECONDARY	TERTIARY
System	System	System
A9000R-Site_A	A9000-Site_B	A9000R-Site_C
Volume	Volume	Volume
ITSO_FA_Volume_2	ITSO_FA_Volume_2	ITSO_FA_Volume_2

Figure 7-54 Multi-site relationship details for our example

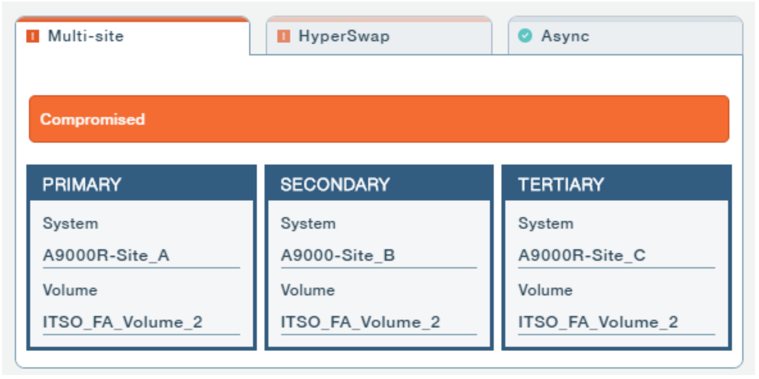
Upon loss of connectivity between System A and the Quorum Witness (see Figure 7-55), the Multi-site relationship enters the Compromised state (see Figure 7-56).



3 Systems

System ^		Active Quorum Witness Status	Quorum...	Co...	Targets
A9000-Site_B		OK	ITSOQW	2	2
A9000R-Site_A	ℹ	Down (Connection Down)	ITSOQW	2	2
A9000R-Site_C	◆	OK	ITSOQW	4	5

Figure 7-55 System A lost connectivity to the active Quorum Witness

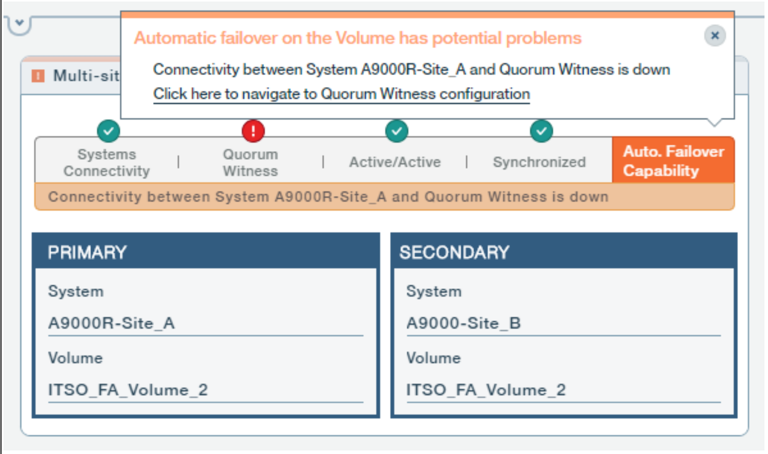


Compromised

PRIMARY	SECONDARY	TERTIARY
System	System	System
A9000R-Site_A	A9000-Site_B	A9000R-Site_C
Volume	Volume	Volume
ITSO_FA_Volume_2	ITSO_FA_Volume_2	ITSO_FA_Volume_2

Figure 7-56 The Multi-site relationship enters the Compromised state as a result

The details of the HyperSwap leg of the Multi-site relationship show potential problems (see Figure 7-57); however, Automatic Failover is still possible at this time.



Automatic failover on the Volume has potential problems

Connectivity between System A9000R-Site_A and Quorum Witness is down

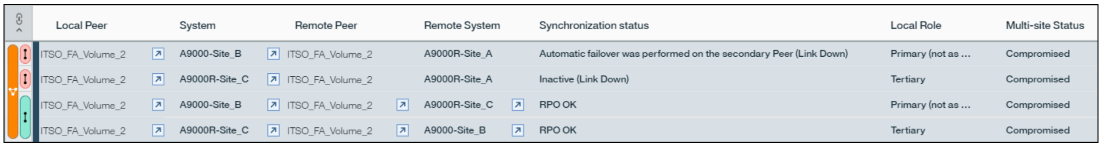
Click here to navigate to Quorum Witness configuration

Connectivity between System A9000R-Site_A and Quorum Witness is down

PRIMARY	SECONDARY
System	System
A9000R-Site_A	A9000-Site_B
Volume	Volume
ITSO_FA_Volume_2	ITSO_FA_Volume_2

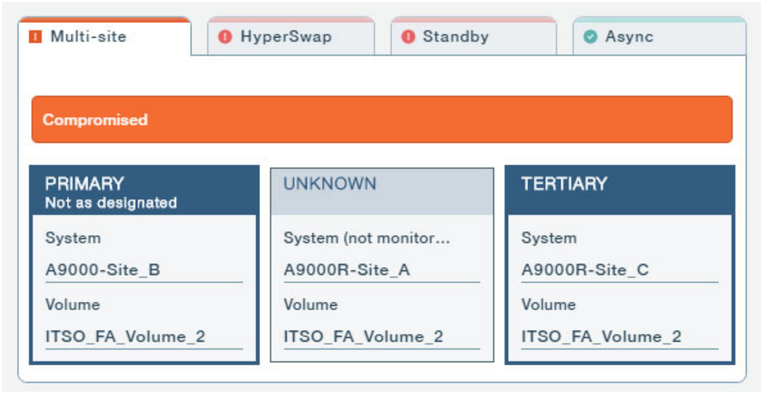
Figure 7-57 Hyperswap leg of the Multi-site relationship shows potential problems

If System A fails now, the Automatic Failover occurs as shown in Figure 7-58 and Figure 7-59.



Local Peer	System	Remote Peer	Remote System	Synchronization status	Local Role	Multi-site Status
ITSO_FA_Volume_2	A9000-Site_B	ITSO_FA_Volume_2	A9000R-Site_A	Automatic failover was performed on the secondary Peer (Link Down)	Primary (not as ...	Compromised
ITSO_FA_Volume_2	A9000R-Site_C	ITSO_FA_Volume_2	A9000R-Site_A	Inactive (Link Down)	Tertiary	Compromised
ITSO_FA_Volume_2	A9000-Site_B	ITSO_FA_Volume_2	A9000R-Site_C	RPO OK	Primary (not as ...	Compromised
ITSO_FA_Volume_2	A9000R-Site_C	ITSO_FA_Volume_2	A9000-Site_B	RPO OK	Tertiary	Compromised

Figure 7-58 Multi-site relationship configuration after System A failure

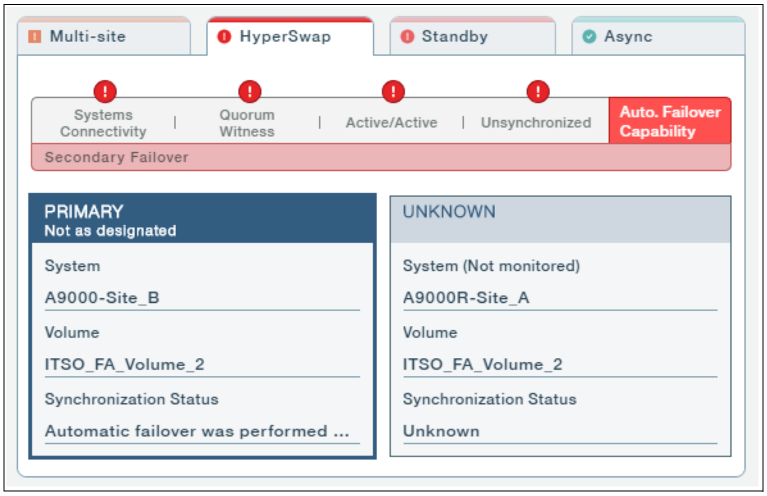


Compromised

PRIMARY Not as designated	UNKNOWN	TERTIARY
System	System (not monitor...	System
A9000-Site_B	A9000R-Site_A	A9000R-Site_C
Volume	Volume	Volume
ITSO_FA_Volume_2	ITSO_FA_Volume_2	ITSO_FA_Volume_2

Figure 7-59 Multi-site relationship details after System A failure

The HyperSwap relationship details window shows that System B assumed the Primary role while System A is down (see Figure 7-60).

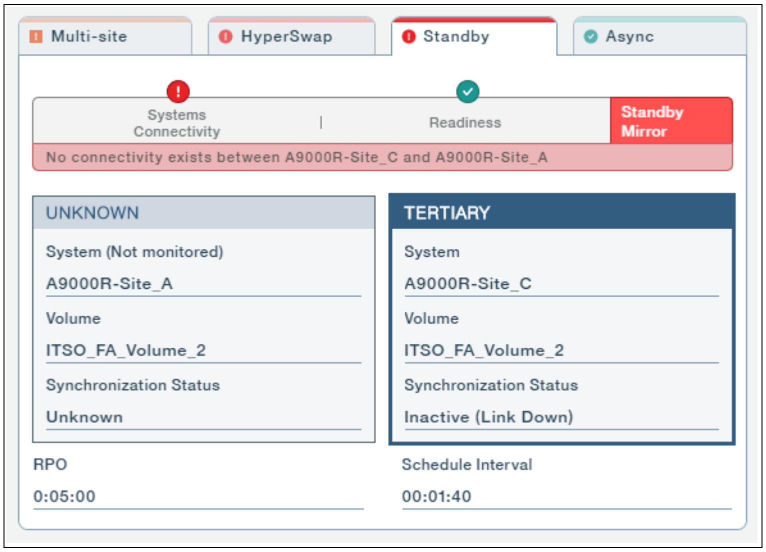


Secondary Failover

PRIMARY Not as designated	UNKNOWN
System	System (Not monitored)
A9000-Site_B	A9000R-Site_A
Volume	Volume
ITSO_FA_Volume_2	ITSO_FA_Volume_2
Synchronization Status	Synchronization Status
Automatic failover was performed ...	Unknown

Figure 7-60 HyperSwap relationship details after the Automatic Failover occurred

When the Automatic Failover is performed, the Active and Standby Asynchronous mirroring legs of the Multi-site relationship are switched; that is, the mirroring between System A and System C is deactivated and becomes Standby (see Figure 7-61), while the mirroring between System B (now temporarily Primary) and System C is activated.



No connectivity exists between A9000R-Site_C and A9000R-Site_A

UNKNOWN

System (Not monitored)
A9000R-Site_A

Volume
ITSO_FA_Volume_2

Synchronization Status
Unknown

TERTIARY

System
A9000R-Site_C

Volume
ITSO_FA_Volume_2

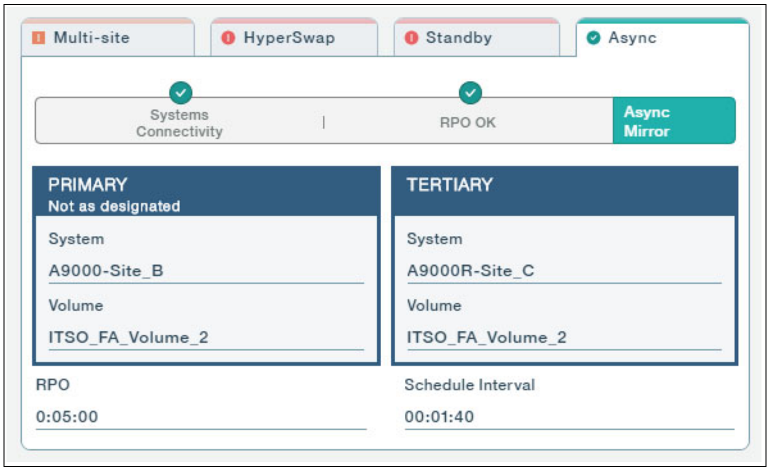
Synchronization Status
Inactive (Link Down)

RPO
0:05:00

Schedule Interval
00:01:40

Figure 7-61 Asynchronous relationship between System A and System C

This configuration ensures data replication to the DR site (see Figure 7-62).



PRIMARY
Not as designated

System
A9000-Site_B

Volume
ITSO_FA_Volume_2

TERTIARY

System
A9000R-Site_C

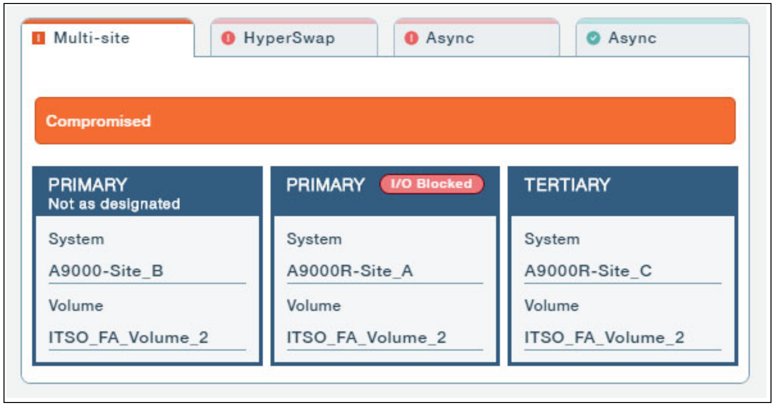
Volume
ITSO_FA_Volume_2

RPO
0:05:00

Schedule Interval
00:01:40

Figure 7-62 Asynchronous relationship between System B and System C

When the failure situation is resolved (System A comes back online and connectivity with System B, System C, and the Quorum Witness is restored) the Multi-site relationship enters a role conflict scenario (see Figure 7-63), which must be resolved as described in 7.1.2, “Multi-site relationship recovery after System A failure” on page 129.



Compromised

PRIMARY
Not as designated

System
A9000-Site_B

Volume
ITSO_FA_Volume_2

PRIMARY I/O Blocked

System
A9000R-Site_A

Volume
ITSO_FA_Volume_2

TERTIARY

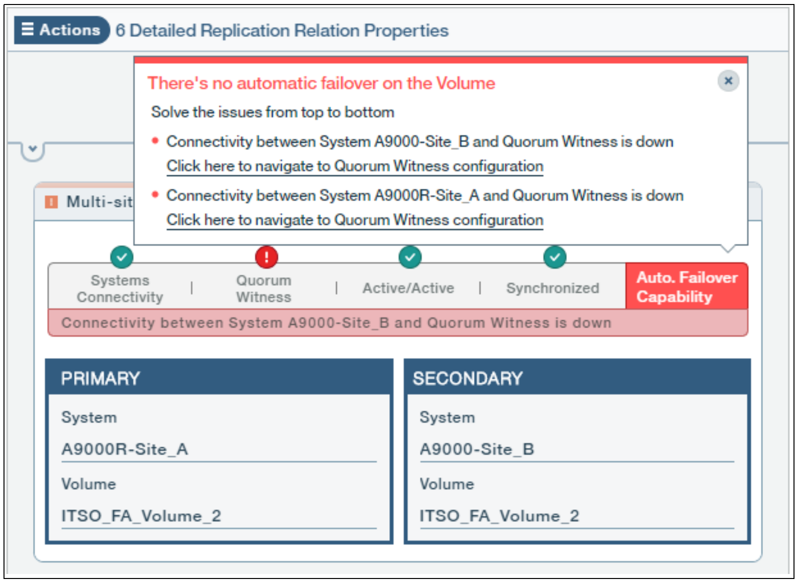
System
A9000R-Site_C

Volume
ITSO_FA_Volume_2

Figure 7-63 Role conflict after System A exits the failing state

7.5.2 Loss of connection between Primary and Secondary systems and the Quorum Witness

The initial configuration for this example is the same as the previous example, as shown in Figure 7-53 on page 156 and Figure 7-54 on page 156. However, in this case, System A (Primary) and System B (Secondary) lose connectivity with the Quorum Witness. This situation compromises the HyperSwap leg of the Multi-site relationship, which prevents Automatic Failure in case of subsequent failures (see Figure 7-64).



There's no automatic failover on the Volume

Solve the issues from top to bottom

- Connectivity between System A9000-Site_B and Quorum Witness is down
  Click here to navigate to Quorum Witness configuration
- Connectivity between System A9000R-Site_A and Quorum Witness is down
  Click here to navigate to Quorum Witness configuration

Connectivity between System A9000-Site_B and Quorum Witness is down

PRIMARY

System
A9000R-Site_A

Volume
ITSO_FA_Volume_2

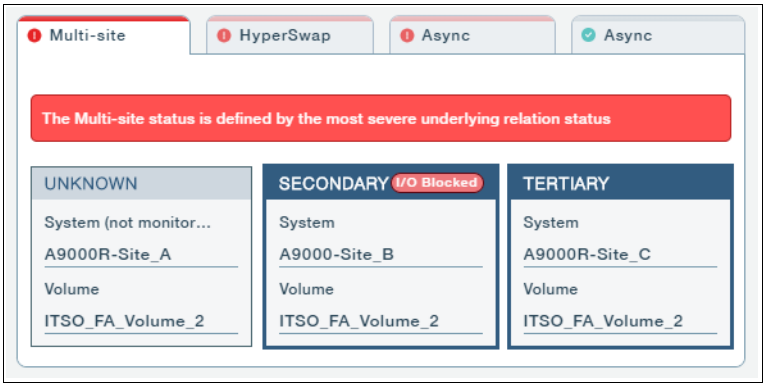
SECONDARY

System
A9000-Site_B

Volume
ITSO_FA_Volume_2

Figure 7-64 HyperSwap details after loss of Quorum Witness connectivity

If System A (Primary) fails, the Multi-site relationship enters a Compromised state (as shown in Figure 7-65 on page 161) in which host I/O is temporarily blocked (because A is down, while System B cannot assume the Primary role because it did not hold the quorum).

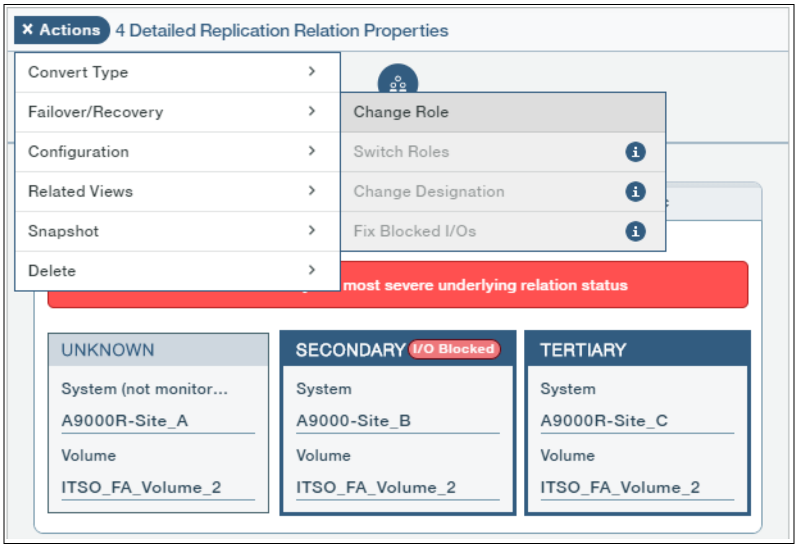


The Multi-site status is defined by the most severe underlying relation status

UNKNOWN	SECONDARY I/O Blocked	TERTIARY
System (not monitor...	System	System
A9000R-Site_A	A9000-Site_B	A9000R-Site_C
Volume	Volume	Volume
ITSO_FA_Volume_2	ITSO_FA_Volume_2	ITSO_FA_Volume_2

Figure 7-65   Multi-site relationship details after System A failure

To restore operations on the volume, a manual role change is needed. This change can be performed by selecting **Failover/Recovery** → **Change Role** in the relationship's Actions menu (see Figure 7-66).

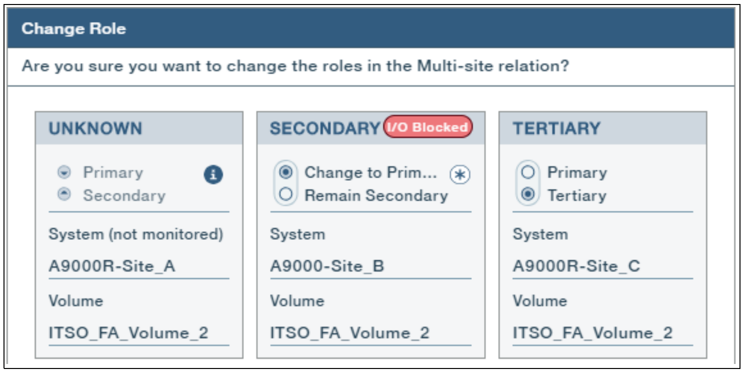


most severe underlying relation status

UNKNOWN	SECONDARY I/O Blocked	TERTIARY
System (not monitor...	System	System
A9000R-Site_A	A9000-Site_B	A9000R-Site_C
Volume	Volume	Volume
ITSO_FA_Volume_2	ITSO_FA_Volume_2	ITSO_FA_Volume_2

Figure 7-66   Performing a manual role change

System B can be given the role of Primary while System A is down (see Figure 7-67).

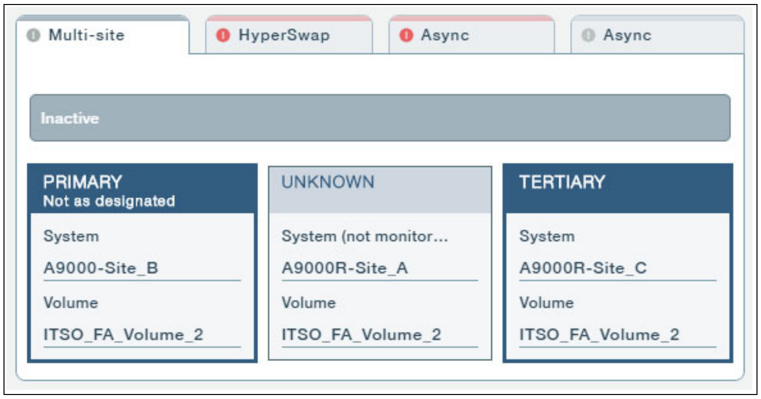

Change Role

Are you sure you want to change the roles in the Multi-site relation?

UNKNOWN	SECONDARY I/O Blocked	TERTIARY
○ Primary ○ Secondary	● Change to Prim... ○ Remain Secondary	○ Primary ● Tertiary
System (not monitored)	System	System
A9000R-Site_A	A9000-Site_B	A9000R-Site_C
Volume	Volume	Volume
ITSO_FA_Volume_2	ITSO_FA_Volume_2	ITSO_FA_Volume_2

Figure 7-67   System B is manually assigned the role of Primary while System A is down

Figure 7-68 shows the relationship details after the role change.



Inactive

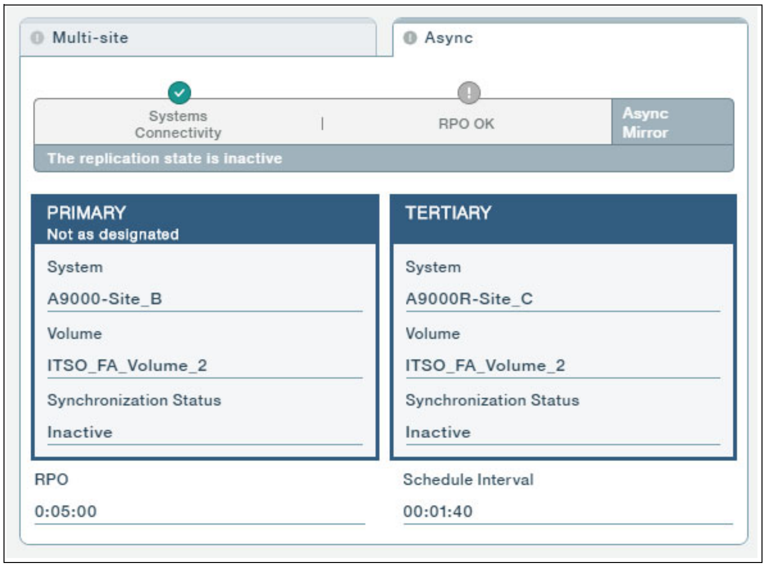
PRIMARY
Not as designated
System
A9000-Site_B
Volume
ITSO_FA_Volume_2

UNKNOWN
System (not monitor…
A9000R-Site_A
Volume
ITSO_FA_Volume_2

TERTIARY
System
A9000R-Site_C
Volume
ITSO_FA_Volume_2

Figure 7-68   Multi-site relationship details after the manual role change

In this case, because the Failover was manual and not automatic, the Asynchronous relationship between System B and System C was not automatically activated, as shown in Figure 7-69.



The replication state is inactive

PRIMARY
Not as designated
System
A9000-Site_B
Volume
ITSO_FA_Volume_2
Synchronization Status
Inactive

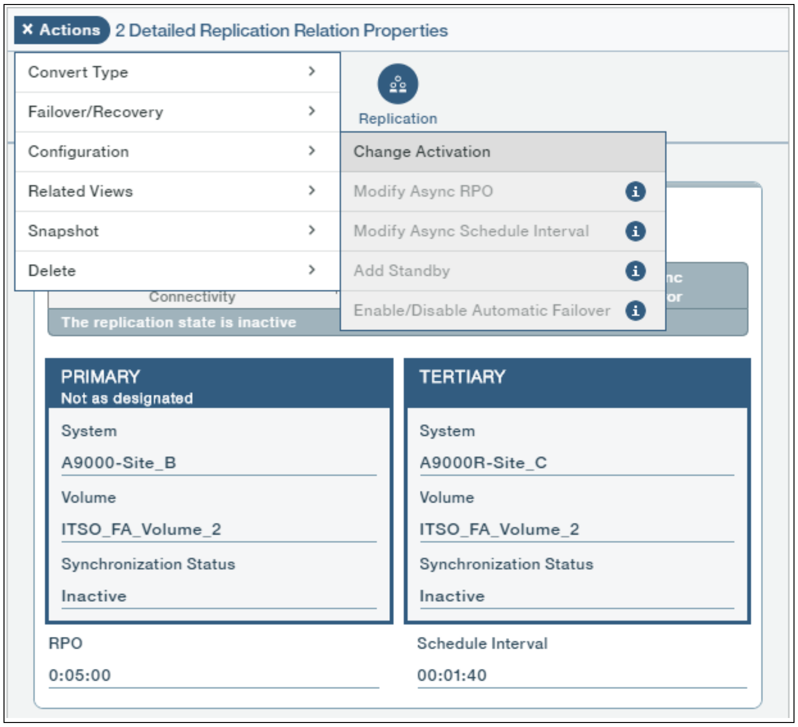
TERTIARY
System
A9000R-Site_C
Volume
ITSO_FA_Volume_2
Synchronization Status
Inactive

RPO
0:05:00

Schedule Interval
00:01:40

Figure 7-69   Details for the Asynchronous relationship between System B and System C

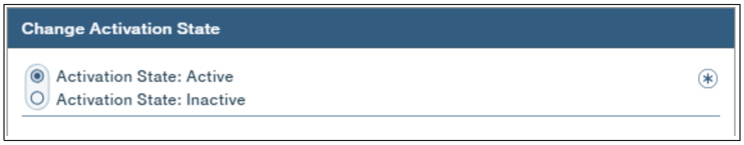
However, this activation can be done manually and after the role change by selecting **Configuration** → **Change Activation** in the relationship's Actions menu (see Figure 7-70 and Figure 7-71).



The replication state is inactive

PRIMARY Not as designated	TERTIARY
System	System
A9000-Site_B	A9000R-Site_C
Volume	Volume
ITSO_FA_Volume_2	ITSO_FA_Volume_2
Synchronization Status	Synchronization Status
Inactive	Inactive
RPO	Schedule Interval
0:05:00	00:01:40

Figure 7-70 Activating the Asynchronous relationship



Change Activation State

● Activation State: Active
○ Activation State: Inactive

Figure 7-71 Confirming the activation

The activation allows for data to be asynchronously replicated from System B to the DR site (System C) while System A is down.

When all failures are resolved, a manual role switch can be performed to return to the original configuration (System A - Primary, System B - Secondary, Active Asynchronous mirroring between System A and System C, and Standby Asynchronous mirroring between System B and System C).

8

Using VMware Site Recovery Manager and HyperSwap

This chapter explains how to use VMware Site Recovery Manager (SRM) with the HyperSwap feature for FlashSystem A9000 and A9000R. The solution is enabled by the IBM Spectrum Accelerate Family's Storage Replication Adapter (SRA) Version 3.0.0 or later.

> **Important:** VMware Cross vCenter vMotion is a feature that allows virtual machines (VMs) to switch from one vCenter Server instance to another to change the compute, network, storage, and management of the VM concurrently. This feature is a critical underlying component of SRM. Introduced in SRM version 6.1, this capability allows SRM to perform a "live migration" of VMs from the primary site to a secondary site with no outage or downtime.
>
> The following requirements must be met for a VM to migrate across vCenters:
>
> - The original and destination vCenter Server instances and ESXi servers are running vSphere version 6.0 or later.
> - Both vCenter Server instances are on the same single sign-on (SSO) domain.
>
> By using stretched storage between the sites (as in IBM HyperSwap technology), the VMs can be live migrated between these sites.

This chapter covers the following topics:

8.1 Overview

VMware uses IBM Spectrum Accelerate Family's Storage Replication Adapter (SRA) to perform a vMotion to migrate running VMs from the production site to the recovery site for planned failover. Version 3.0.0 added support for IBM HyperSwap, as introduced by the FlashSystem A9000 and A9000R storage systems (software Version 12.1 or later). During an unplanned failover where the primary site is down, VMware still uses similar technologies as with a planned outage failover.

The benefits of the use of IBM HyperSwap to VMware mirrors are an increased speed of recovery (including expedited sub-steps) and the SRA updates site preferences automatically. In practical terms, HyperSwap dramatically reduces failover times and makes the entire process far more streamlined.

Note: For the HyperSwap solution, a synchronous remote mirroring pairing must be used. For more information about installing and configuring SRM, see *Site Recovery Manager Installation and Configuration*, which is available at the VMware Site Recovery Manager Documentation website.

An overview of the solution is shown in Figure 8-1.

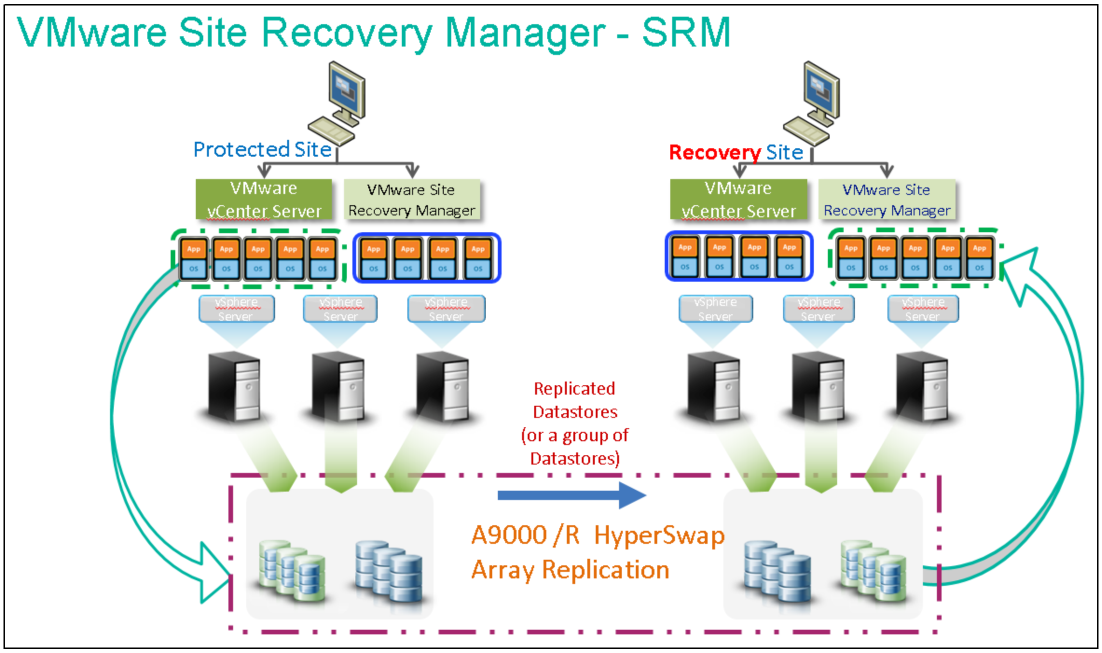

VMware Site Recovery Manager - SRM
Protected Site
VMware vCenter Server
VMware Site Recovery Manager
Recovery Site
VMware vCenter Server
VMware Site Recovery Manager
App
OS
vSphere Server
Replicated Datastores (or a group of Datastores)
A9000 /R HyperSwap Array Replication

Figure 8-1 SRM Overview with IBM FlashSystem A9000/R HyperSwap Replication

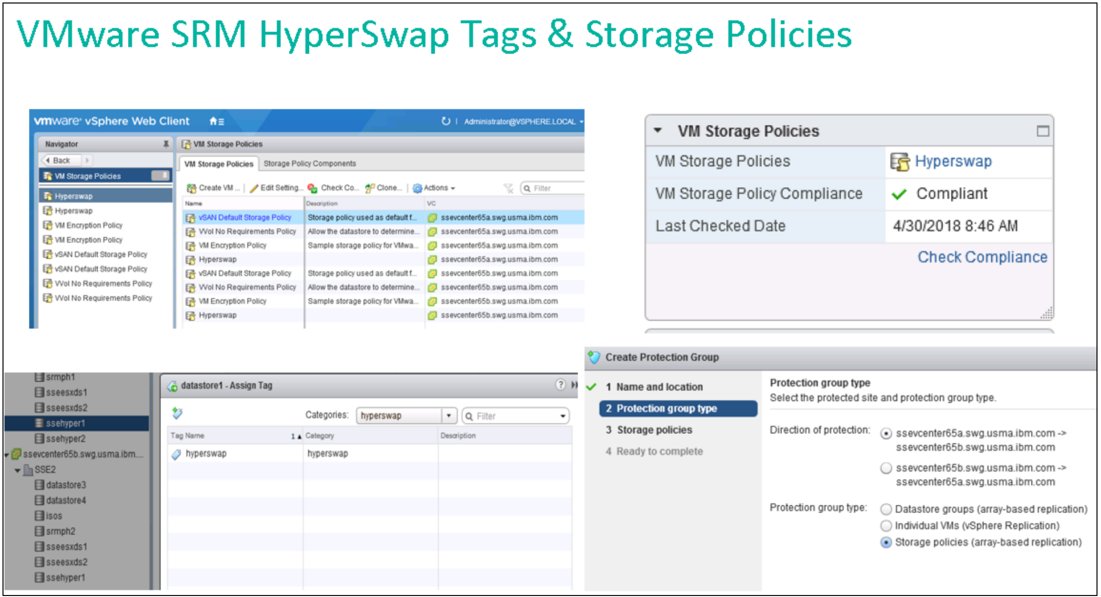
For HyperSwap, if no target connectivity exists between the individual peers and the recovery storage system has the volumes in secondary role, the IBM Spectrum Accelerate Family SRA cannot perform failover from the protected to recovery site. Ensure that none of the names of both storage systems were changed after the mirroring targets are defined; otherwise, both arrays do not pair in SRM.

Note: The name of the target system can be changed by using the Hyper-Scale Manager or the xcli command **`target_rename`**.

8.2 Prerequisites

The following prerequisites must be met:

- The source and destination vCenter Server instances and ESXi hosts run version 6.0 or later. For more information, see this article at VMware Knowledge Base.
- IBM SRA Version 3.0 or later, installed on each vCenter server.
- VMware SRM Version 6.1 or later, installed. Each vCenter server has its own SRM server.
- Both vCenter Server instances are on the same single sign-on domain. For more information about supported topologies, see this article at VMware Knowledge Base.
- VMware's Cluster Mapping is engaged by SRM to list all the hosts that are defined inside of the A9000 /R cluster. A 1:1 mapping between the A9000 /R Cluster (of volumes) and the Vsphere cluster is needed. Therefore, any host that is not part of the Vsphere cluster but is defined as part of the A9000/R cluster might interfere with the VMware's Cluster mapping, which might then interfere with SRM failover operations.
- Volumes at the protected site are marked as “Primary” inside the A9000/R Mirroring relationship. Volumes can also be part of A9000/R Consistency Groups.
- VMware Storage Tags and Policies are defined before SRM definitions are created. A brief overview is shown in Figure 8-2.



VMware SRM HyperSwap Tags & Storage Policies

Name	Description	VC
vSAN Default Storage Policy	Storage policy used as default f...	ssevcenter65a.swg.usma.ibm.com
VVol No Requirements Policy	Allow the datastore to determine...	ssevcenter65a.swg.usma.ibm.com
VM Encryption Policy	Sample storage policy for VMwa...	ssevcenter65a.swg.usma.ibm.com
Hyperswap		ssevcenter65a.swg.usma.ibm.com
vSAN Default Storage Policy	Storage policy used as default f...	ssevcenter65b.swg.usma.ibm.com
VVol No Requirements Policy	Allow the datastore to determine...	ssevcenter65b.swg.usma.ibm.com
VM Encryption Policy	Sample storage policy for VMwa...	ssevcenter65b.swg.usma.ibm.com
Hyperswap		ssevcenter65b.swg.usma.ibm.com

VM Storage Policies

VM Storage Policies	Hyperswap
VM Storage Policy Compliance	✓ Compliant
Last Checked Date	4/30/2018 8:46 AM

Categories: hyperswap

Tag Name	Category	Description
hyperswap	hyperswap	

Create Protection Group

✓ 1 Name and location
2 Protection group type
3 Storage policies
4 Ready to complete

Protection group type
Select the protected site and protection group type.

Direction of protection:
● ssevcenter65a.swg.usma.ibm.com -> ssevcenter65b.swg.usma.ibm.com
○ ssevcenter65b.swg.usma.ibm.com -> ssevcenter65a.swg.usma.ibm.com

Protection group type:
○ Datastore groups (array-based replication)
○ Individual VMs (vSphere Replication)
● Storage policies (array-based replication)

Figure 8-2 VMware Storage Policies and Tags are needed for before creating SRM definitions

8.3 Installing IBM Spectrum Accelerate Family SRA

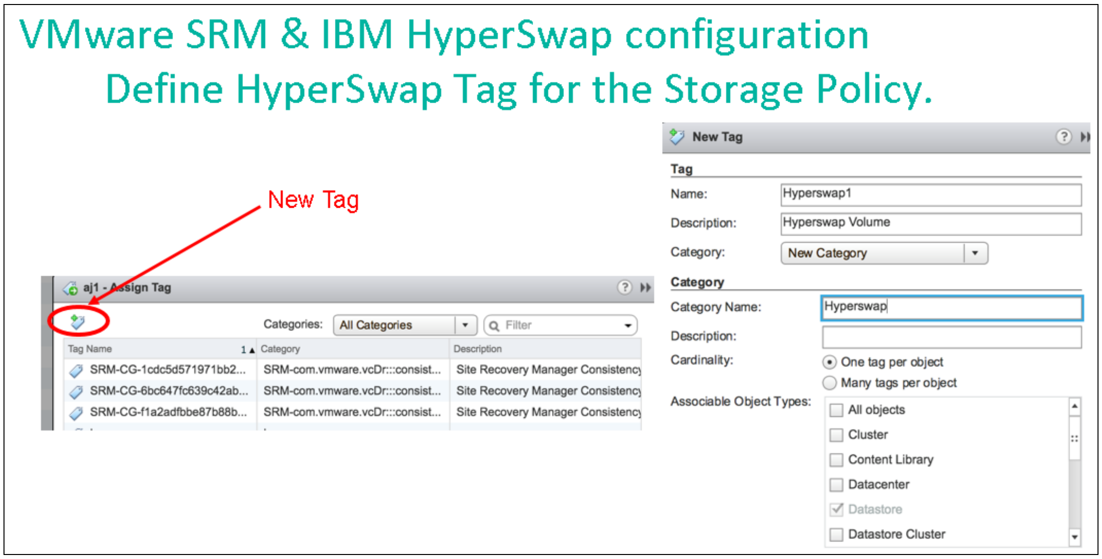
IBM Spectrum Accelerate Family SRA is a software add-on that integrates VMWare SRM to run failovers together with supported storage systems. The IBM Spectrum Accelerate Family SRA extends SRM capabilities and uses HyperSwap as part of the SRM comprehensive Disaster Recovery Planning (DRP) solution. VMWare administrators can automate the failover of HyperSwap volumes at the protected SRM site to a recovery SRM site. Immediately upon a failover, the VMWare ESX/ESXi servers at the recovery SRM site initiate the replicated datastores on the HyperSwap volumes.

The IBM Spectrum Accelerate Family SRA Version 3.0 and later is available for download at the VMware website.

IBM Spectrum Accelerate Family SRA Version 3.0 works with the SRM instance at the protected and recovery site. Therefore, it must be installed on the same server where the SRM instance is installed at the protected and recovery sites.

8.4 Configuring tags

Before SRM is configured, the VMware administrator must create and assign VMware tags to the resources that are configured in the SRM relationship. These tags are created before the Storage Policy is created. An overview is shown in Figure 8-3.



VMware SRM & IBM HyperSwap configuration

Define HyperSwap Tag for the Storage Policy.

New Tag

aj1 - Assign Tag

Categories: All Categories

Tag Name	Category	Description
SRM-CG-1cdc5d571971bb2...	SRM-com.vmware.vcDr:::consist...	Site Recovery Manager Consistency
SRM-CG-6bc647fc639c42ab...	SRM-com.vmware.vcDr:::consist...	Site Recovery Manager Consistency
SRM-CG-f1a2adfbbe87b88b...	SRM-com.vmware.vcDr:::consist...	Site Recovery Manager Consistency

New Tag

Tag

Name: Hyperswap1

Description: Hyperswap Volume

Category: New Category

Category

Category Name: Hyperswap

Description:

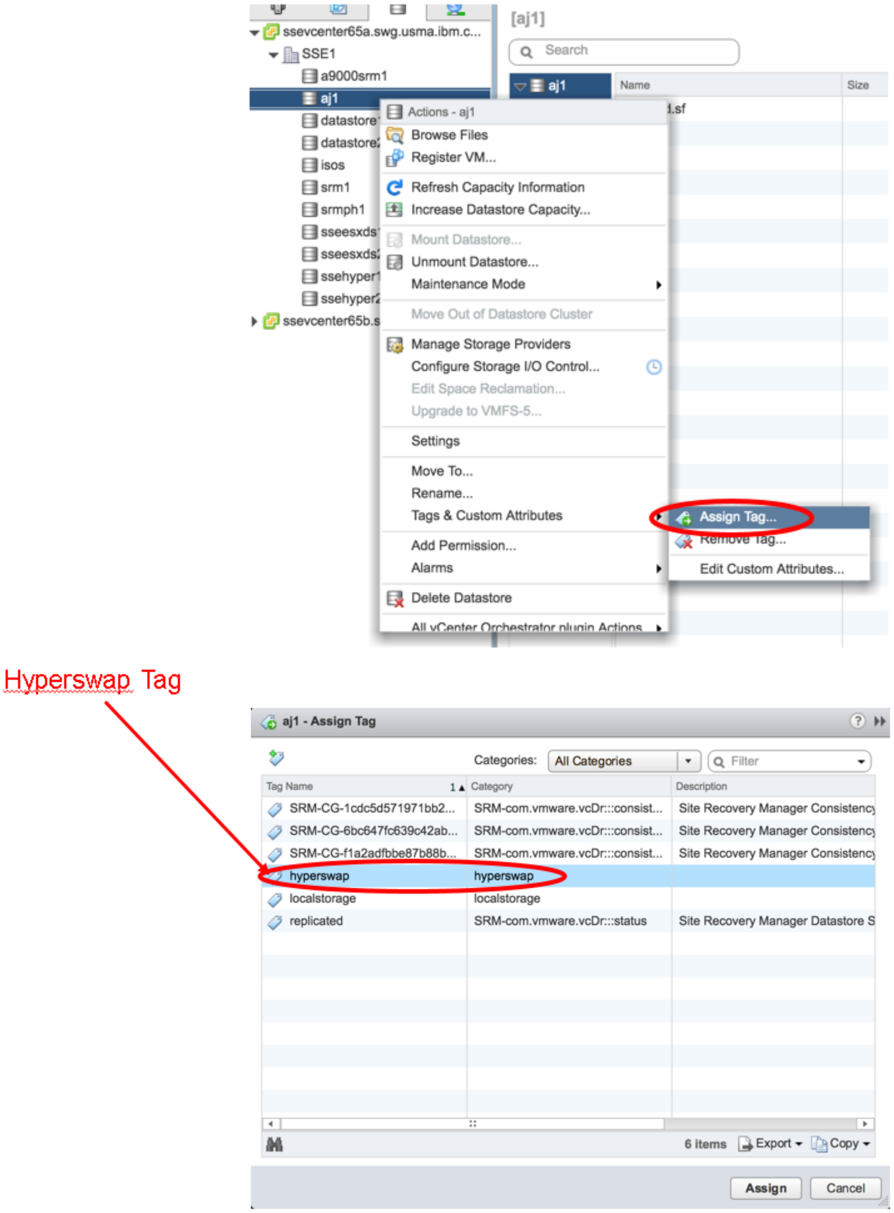
Cardinality: ● One tag per object ○ Many tags per object

Associable Object Types:
- □ All objects
- □ Cluster
- □ Content Library
- □ Datacenter
- ☑ Datastore
- □ Datastore Cluster

Figure 8-3 Initial Creation of VMware Vcenter Tag

After the tag is created, it must be assigned to the vCenter resources involved, as shown in Figure 8-4.

VMware SRM & IBM HyperSwap configuration: Define a HyperSwap Tag for the Storage Policy

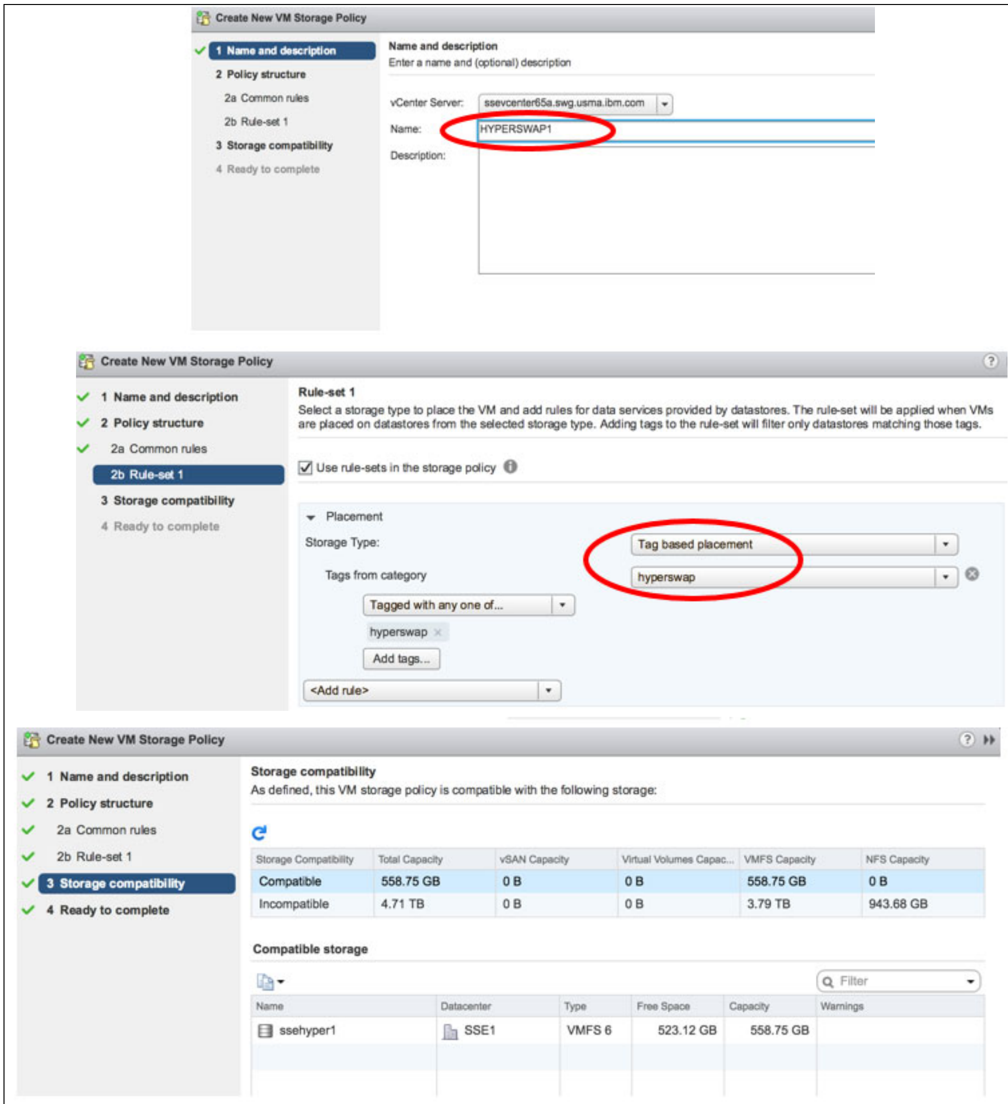


Tag Name	Category	Description
SRM-CG-1cdc5d571971bb2...	SRM-com.vmware.vcDr:::consist...	Site Recovery Manager Consistency
SRM-CG-6bc647fc639c42ab...	SRM-com.vmware.vcDr:::consist...	Site Recovery Manager Consistency
SRM-CG-f1a2adfbbe87b88b...	SRM-com.vmware.vcDr:::consist...	Site Recovery Manager Consistency
hyperswap	hyperswap	
localstorage	localstorage	
replicated	SRM-com.vmware.vcDr:::status	Site Recovery Manager Datastore S

Figure 8-4 Assigning the created tag, inside Vcenter

8.5 Configuring a VMware Storage Policy

After the necessary tags are created, the next step is to create a storage policy for SRM to use for its failover, and fail-back steps. From the VMware vSphere Web Client Home page, select VM Storage Policies. The window that is shown in Figure 8-5 opens.



Name and description

Enter a name and (optional) description

vCenter Server: ssevcenter65a.swg.usma.ibm.com

Name: HYPERSWAP1

Description:

Rule-set 1

Select a storage type to place the VM and add rules for data services provided by datastores. The rule-set will be applied when VMs are placed on datastores from the selected storage type. Adding tags to the rule-set will filter only datastores matching those tags.

☑ Use rule-sets in the storage policy

Placement

Storage Type: Tag based placement

Tags from category: hyperswap

Tagged with any one of...

hyperswap

Storage compatibility

As defined, this VM storage policy is compatible with the following storage:

Storage Compatibility	Total Capacity	vSAN Capacity	Virtual Volumes Capac...	VMFS Capacity	NFS Capacity
Compatible	558.75 GB	0 B	0 B	558.75 GB	0 B
Incompatible	4.71 TB	0 B	0 B	3.79 TB	943.68 GB

Compatible storage

Name	Datacenter	Type	Free Space	Capacity	Warnings
ssehyper1	SSE1	VMFS 6	523.12 GB	558.75 GB	

Figure 8-5 VMware Vcenter Storage Policy creation for IBM HyperSwap replication and protection

8.6 Configuring SRM

In this section, we assume that the SRM was configured to the point of assigning tags to datastores and Storage Policies and that the site to site pairing is configured.

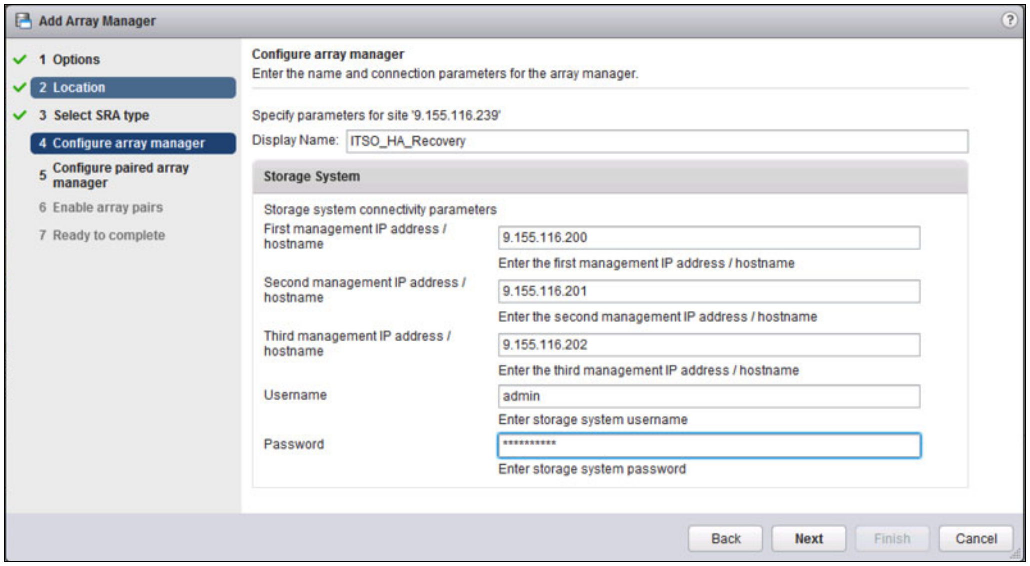
> **Important:** The solution does not work unless both vCenter Servers are configured with a single sign-on domain; that is, the same PSC (platform service controller).

For more information about setting up tags, Storage Policies, site pairing, and resource mappings, see the VMware publications *Site Recover Manager Installation and Configuration*, and *Site Recovery Manager Administration*, which are available at the VMware website.

After initial configuration is completed, the array manager must be configured so that SRM can discover the HyperSwap devices, compute datastores, and initiate storage operations.

To configure the Array Based Replication, complete the following steps:

1. In the vSphere web client, click the **Site Recovery** plug-in. Then click **Array Based Replication**, then **Add Array Manager**.
2. Select the **Add a pair of array managers** option and click **Next**.
3. Select the paired site relationship to use in the array manager wizard and click **Next**.
4. Select the IBM Spectrum Accelerate Family SRA displayed in the wizard for both the array managers and click **Next**.
5. On the **Configure array manager page**, specify a display name for the site that is displayed, in this example, the recovery site was the first site displayed in the wizard. Enter the management IP addresses for the displayed FlashSystemA9000 or A9000R site. Provide a username and password to access the FlashSystem A9000 or A9000R, and click **Next,** as shown in Figure 8-6.



Configure array manager

Enter the name and connection parameters for the array manager.

Specify parameters for site '9.155.116.239'

Display Name: ITSO_HA_Recovery

Storage System

Storage system connectivity parameters

First management IP address / hostname: 9.155.116.200

Enter the first management IP address / hostname

Second management IP address / hostname: 9.155.116.201

Enter the second management IP address / hostname

Third management IP address / hostname: 9.155.116.202

Enter the third management IP address / hostname

Username: admin

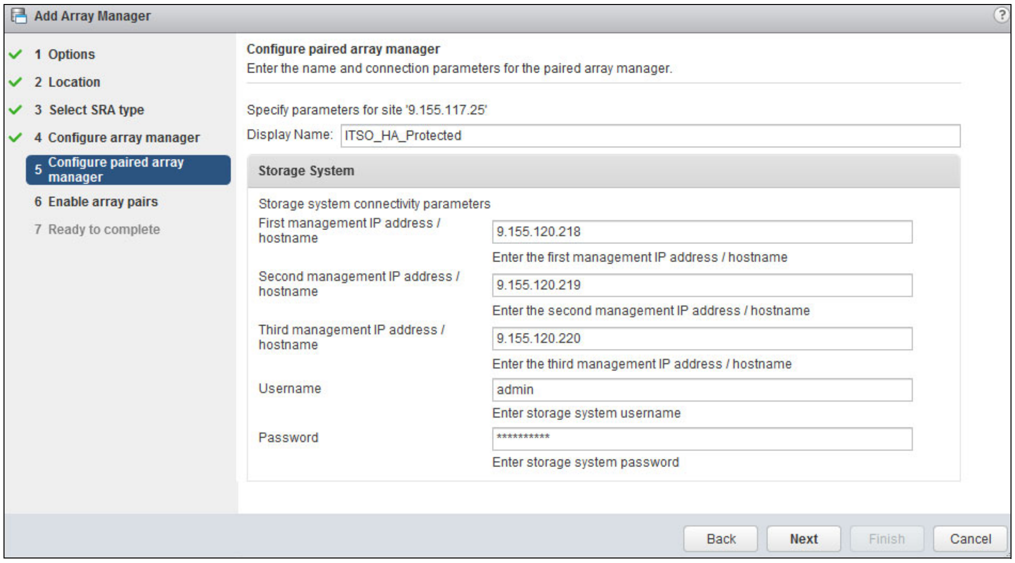
Enter storage system username

Password: **********

Enter storage system password

Figure 8-6 Defining recovery array manager

6. On the **Configure paired array manager**, specify a display name for the second displayed site. In this example the second site was the protected site. Enter the management IP addresses for the displayed site FlashSystem A9000 or A9000R. Provide a username and password to access the FlashSystem A9000 or A9000R, and click **Next,** as shown in Figure 8-7.



Configure paired array manager

Enter the name and connection parameters for the paired array manager.

Specify parameters for site '9.155.117.25'

Display Name: ITSO_HA_Protected

Storage System

Storage system connectivity parameters

First management IP address / hostname: 9.155.120.218

Enter the first management IP address / hostname

Second management IP address / hostname: 9.155.120.219

Enter the second management IP address / hostname

Third management IP address / hostname: 9.155.120.220

Enter the third management IP address / hostname

Username: admin

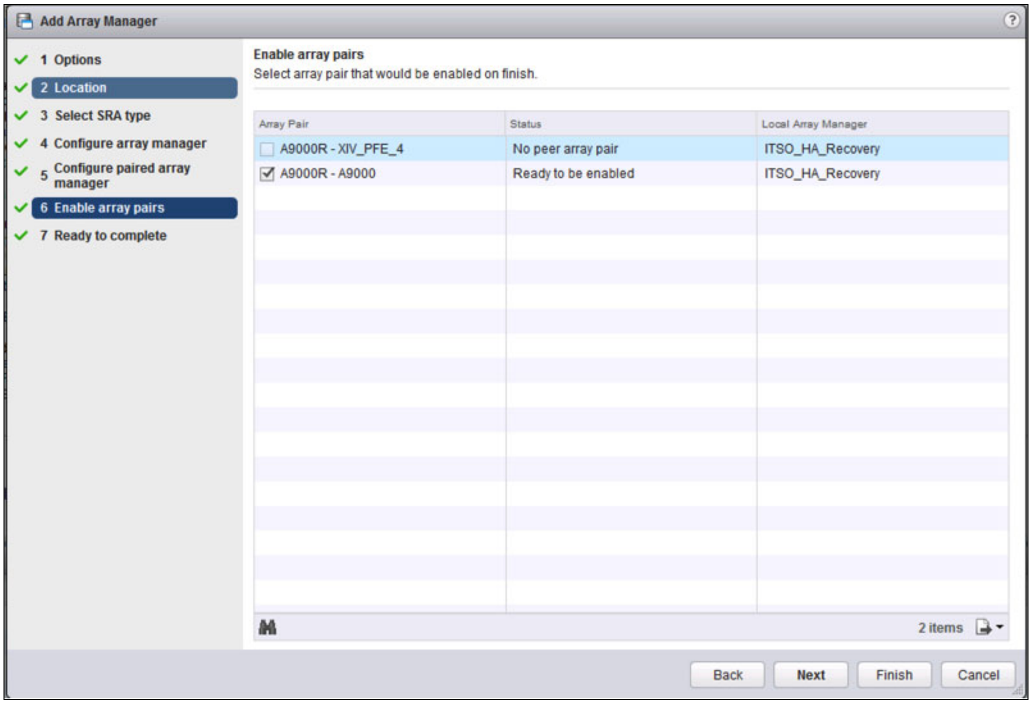
Enter storage system username

Password: **********

Enter storage system password

Figure 8-7 Defining the protected array manager

7. On the **Enable array pairs** page, select the FlashSystem A9000 or A9000R array pair to be enabled and click **Next**, as shown in Figure 8-8.



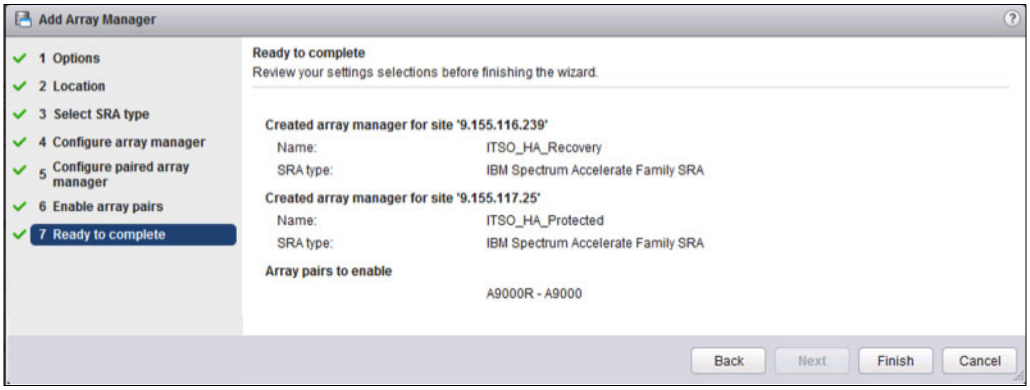
Enable array pairs

Select array pair that would be enabled on finish.

Array Pair	Status	Local Array Manager
☐ A9000R - XIV_PFE_4	No peer array pair	ITSO_HA_Recovery
☑ A9000R - A9000	Ready to be enabled	ITSO_HA_Recovery

Figure 8-8 Selecting the array pair

8. Review the content displayed in the wizard, and click **Finish**, as shown in Figure 8-9.



Ready to complete

Review your settings selections before finishing the wizard.

Created array manager for site '9.155.116.239'

Name: ITSO_HA_Recovery

SRA type: IBM Spectrum Accelerate Family SRA

Created array manager for site '9.155.117.25'

Name: ITSO_HA_Protected

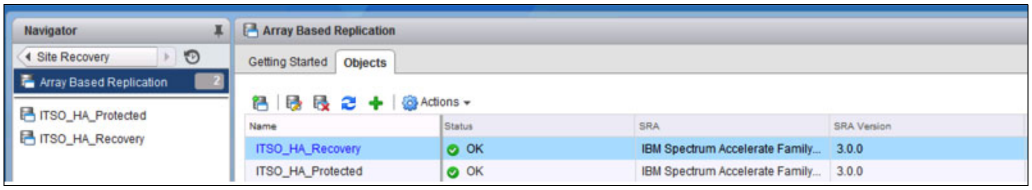
SRA type: IBM Spectrum Accelerate Family SRA

Array pairs to enable

A9000R - A9000

Figure 8-9 Verifying the array manager configuration

The Array Based Replication For both protected and recovery sites is displayed as OK, as shown in Figure 8-10.



Name	Status	SRA	SRA Version
ITSO_HA_Recovery	OK	IBM Spectrum Accelerate Family...	3.0.0
ITSO_HA_Protected	OK	IBM Spectrum Accelerate Family...	3.0.0

Figure 8-10 Array Based Replication status

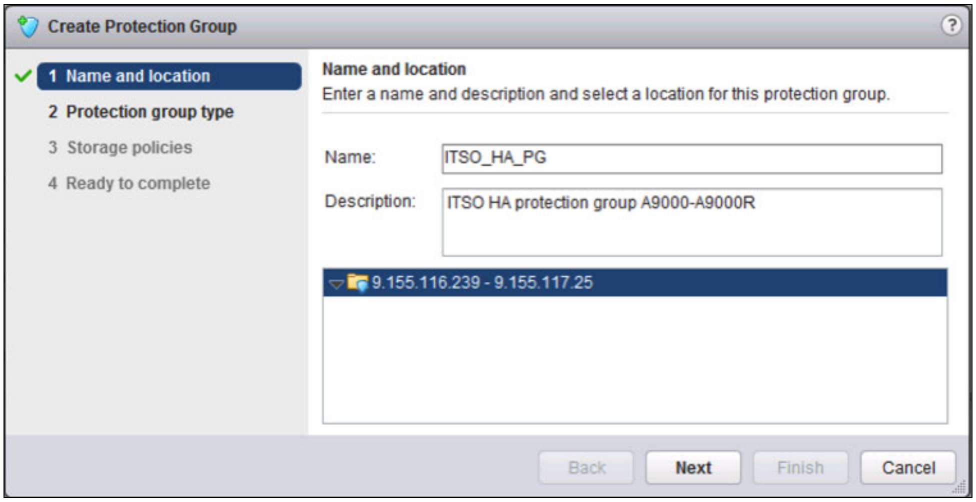
8.6.1 Creating storage policy-based protection group

A storage policy-based protection group enables automated protection of virtual machines that are associated with that storage policy. When a virtual machine is associated with or disassociated from a storage policy, SRM automatically protects or unprotects it.

To configure a protection group, complete the following steps:

1. In the vSphere web client, click the **Site Recovery** plug-in. Then, click **Protection Group** → **Create a Protection Group** to start a new protection group.

2. In the Create Protection Group wizard, enter a name description and location of the protection group then, click **OK**, as shown in Figure 8-11.



Create Protection Group

Name and location

Enter a name and description and select a location for this protection group.

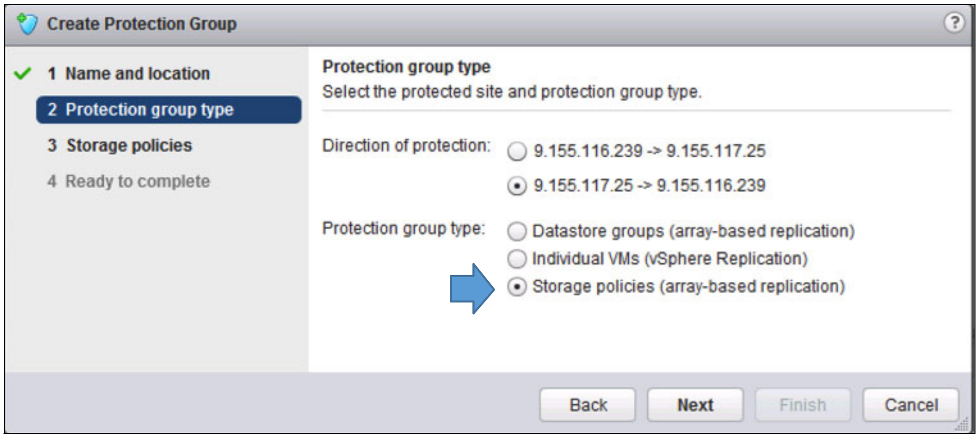
Name: ITSO_HA_PG

Description: ITSO HA protection group A9000-A9000R

9.155.116.239 - 9.155.117.25

Figure 8-11 Defining the protection group name

3. In the Protection group type window, specify the required direction of protection and ensure to select only **Storage policies (array-based replication)**, as shown in Figure 8-12.



Create Protection Group

Protection group type

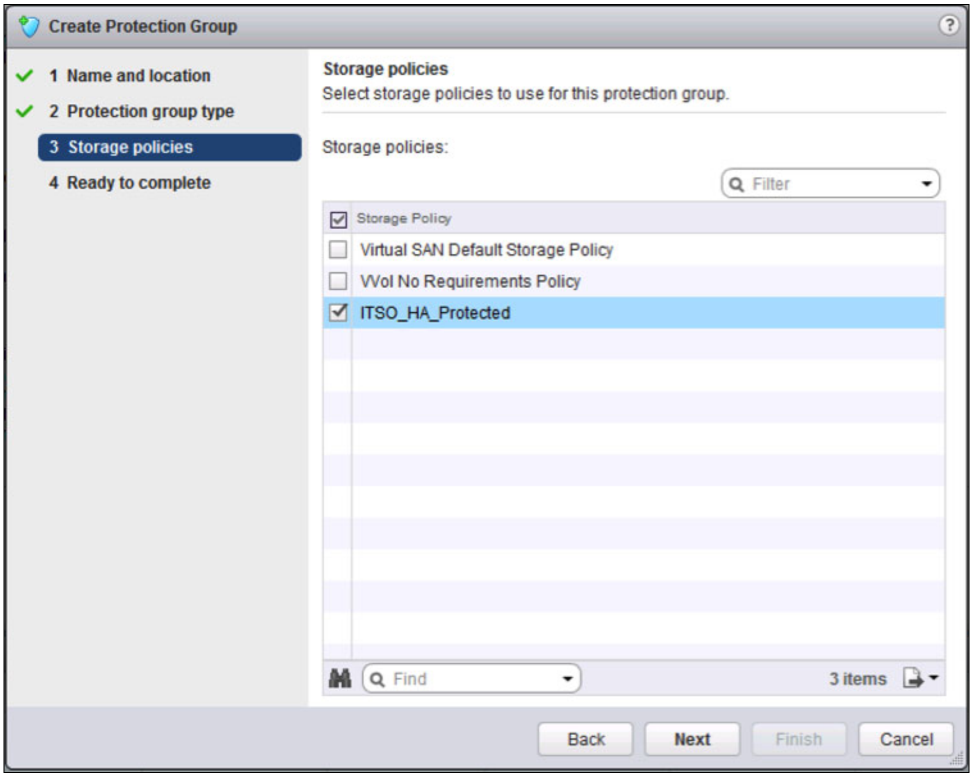
Select the protected site and protection group type.

Direction of protection:
○ 9.155.116.239 -> 9.155.117.25
● 9.155.117.25 -> 9.155.116.239

Protection group type:
○ Datastore groups (array-based replication)
○ Individual VMs (vSphere Replication)
● Storage policies (array-based replication)

Figure 8-12 Setting the direction and type of protection group

4. In the Storage policies window, select the storage policies that are to be used for the protection group and click **Next**, as shown in Figure 8-13.



Storage policies

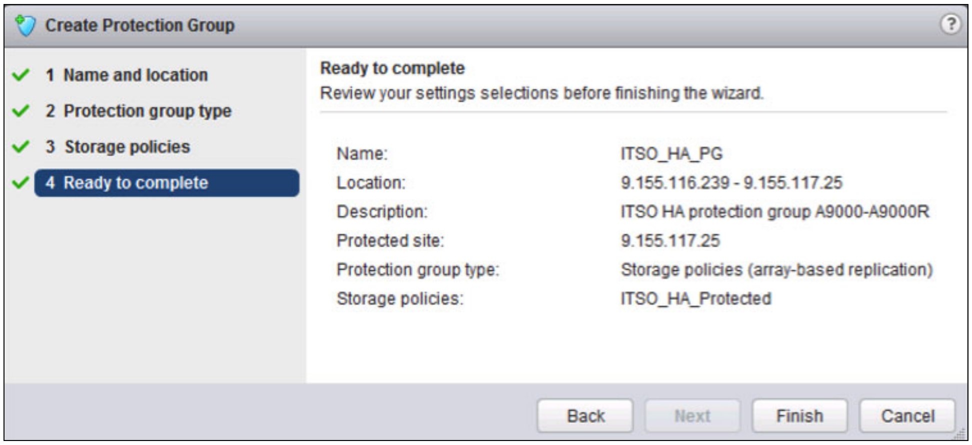
Select storage policies to use for this protection group.

Storage policies:

☑	Storage Policy
□	Virtual SAN Default Storage Policy
□	VVol No Requirements Policy
☑	ITSO_HA_Protected

Figure 8-13 Selecting the protection groups storage policies

5. Review the protection group settings and click **Finish** to create a storage policy-based protection group, as shown in Figure 8-14.



Ready to complete

Review your settings selections before finishing the wizard.

Name:	ITSO_HA_PG
Location:	9.155.116.239 - 9.155.117.25
Description:	ITSO HA protection group A9000-A9000R
Protected site:	9.155.117.25
Protection group type:	Storage policies (array-based replication)
Storage policies:	ITSO_HA_Protected

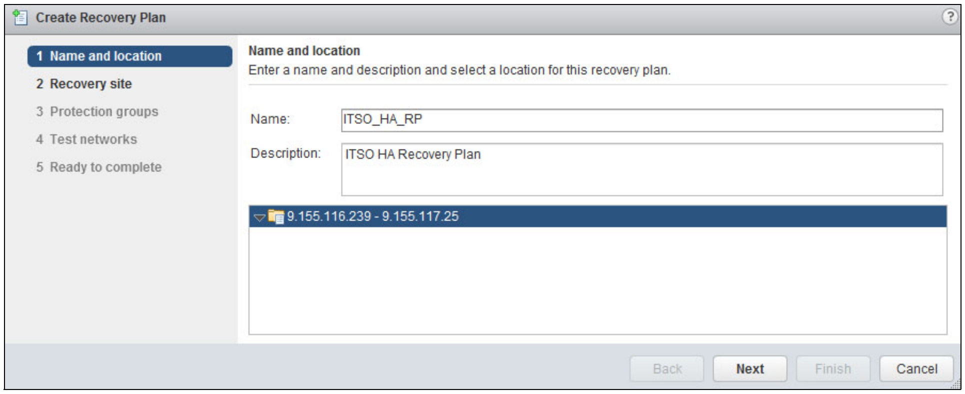
Figure 8-14 Verifying the protection group configuration

8.6.2 Creating a recovery plan

In VMware's Site Recovery Manager, a recovery plan is similar to an automated run book. It controls every step of the recovery process, including the order in which the SRM powers on and off the virtual machines, and the network addresses that the recovered virtual machines use, and so on. Recovery plans are flexible and customizable. A recovery plan includes one or more protection groups and a protection group can be included in multiple recovery plans. For example, one recovery plan can be created to handle a planned migration of services from the protected site to the recovery site, and another plan to handle an unplanned event such as a power failure or natural disaster.

After configuring a protection group, a recovery plan must be created and tested. Complete the following steps to create a recovery plan:

1. In the vSphere web client click the **Site Recovery** plug-in. Then, click **Recovery Plans**, and then **Create a Recovery Plan** to start creating a recovery plan.
2. In the Create Recovery Plan wizard, enter a name, description, and location for a recovery plan click **Next**, as shown in Figure 8-15.



Create Recovery Plan

1 Name and location
2 Recovery site
3 Protection groups
4 Test networks
5 Ready to complete

Name and location
Enter a name and description and select a location for this recovery plan.

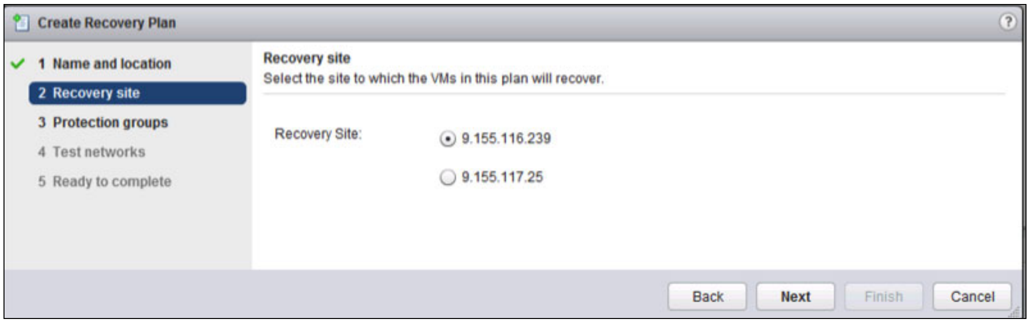
Name: ITSO_HA_RP

Description: ITSO HA Recovery Plan

9.155.116.239 - 9.155.117.25

Figure 8-15 Defining the recovery plan name

3. Select the site that the virtual machines in the recovery plans recover to and click **Next**, as shown in Figure 8-16.



Create Recovery Plan

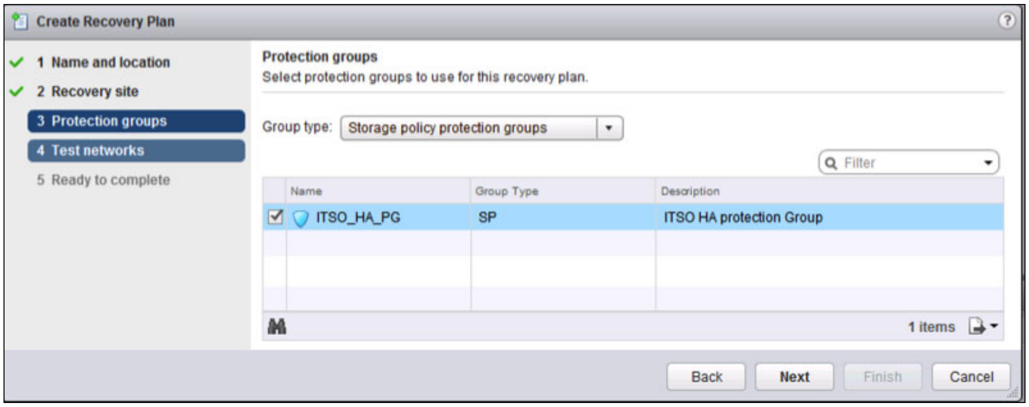
✓ 1 Name and location
2 Recovery site
3 Protection groups
4 Test networks
5 Ready to complete

Recovery site
Select the site to which the VMs in this plan will recover.

Recovery Site: ● 9.155.116.239
○ 9.155.117.25

Figure 8-16 Selecting the recovery site

4. In the Protection groups window, select **Storage policy protection groups** as the group type from the drop-down list. Select the protection group that was created earlier and click **Next**, as shown in Figure 8-17.



Protection groups

Select protection groups to use for this recovery plan.

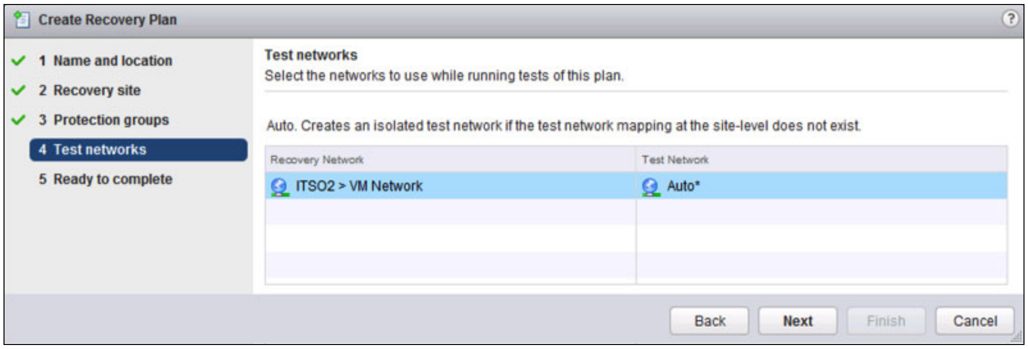
Group type: Storage policy protection groups

Name	Group Type	Description
☑ ITSO_HA_PG	SP	ITSO HA protection Group

1 items

Figure 8-17 Selecting the protection groups

5. Select the networks that is used while the recovery plan is running and click **Next**, as shown in Figure 8-18.



Test networks

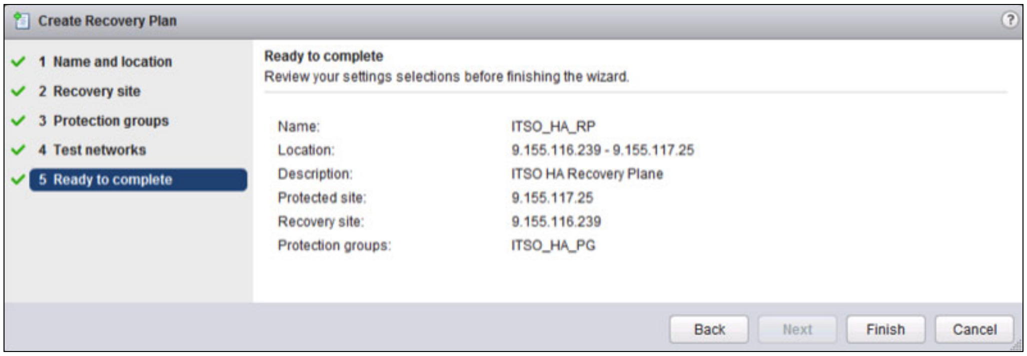
Select the networks to use while running tests of this plan.

Auto. Creates an isolated test network if the test network mapping at the site-level does not exist.

Recovery Network	Test Network
ITSO2 > VM Network	Auto*

Figure 8-18 Selecting the test networks

6. Verify the recovery plan settings and click **Finish** to create a recovery plan, as shown in Figure 8-19.



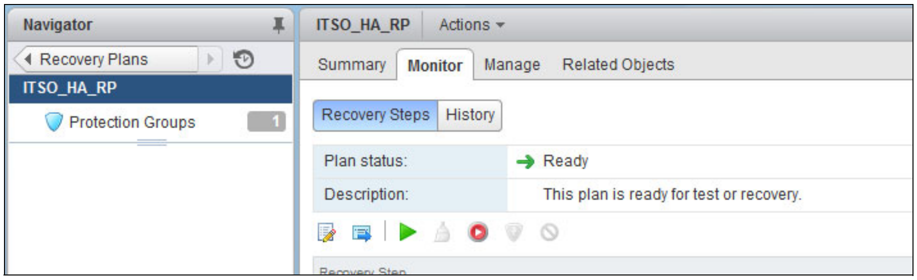
Ready to complete

Review your settings selections before finishing the wizard.

Name:	ITSO_HA_RP
Location:	9.155.116.239 - 9.155.117.25
Description:	ITSO HA Recovery Plane
Protected site:	9.155.117.25
Recovery site:	9.155.116.239
Protection groups:	ITSO_HA_PG

Figure 8-19 Verifying the recovery plan

7. After it successfully completes, the recovery plan shows a status of Ready, as shown in Figure 8-20.



Plan status: Ready

Description: This plan is ready for test or recovery.

Figure 8-20 Plan is in a Ready state

8.6.3 Testing a recovery plan

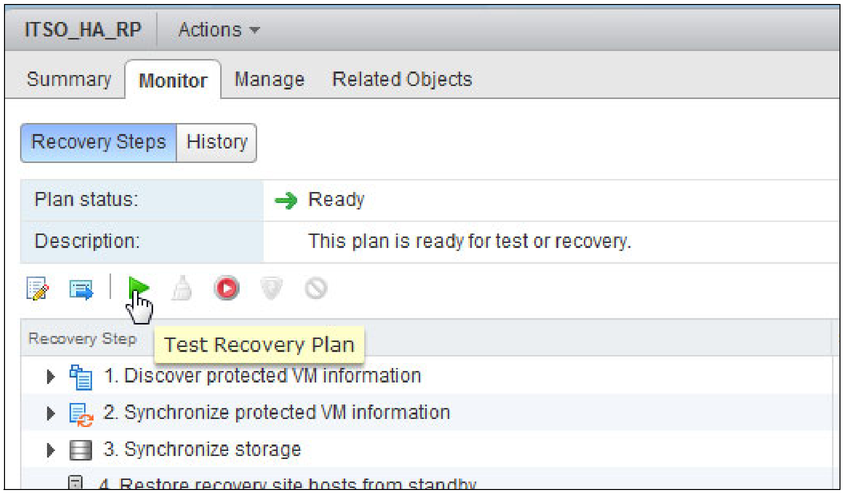
After a recovery plan is created, it must be tested to ensure that the virtual machines in the protection group are correctly recovered across the recovery sites. The tested recovery plans make the environment ready for disaster recovery situations. If the recovery plan is not tested, the actual virtual machines might not be recovered in case of a disaster, which can result in data loss. Testing a recovery plan runs almost every aspect of the recovery scenario.

In the array-based replication feature of SRM, a snapshot of the HyperSwap volume hosting the virtual machines disk file will be created while testing a recovery plan. Snapshots are removed upon cleanup.

Note: It is strongly advised to test all recovery plans for a planned migration and disaster recovery situation.

Complete the following steps to test a recovery plan:

1. In the vSphere web client, click the **Site Recovery** plug-in. Then, select a recovery plan and click the **Test Recovery Plan** icon, as shown in Figure 8-21.



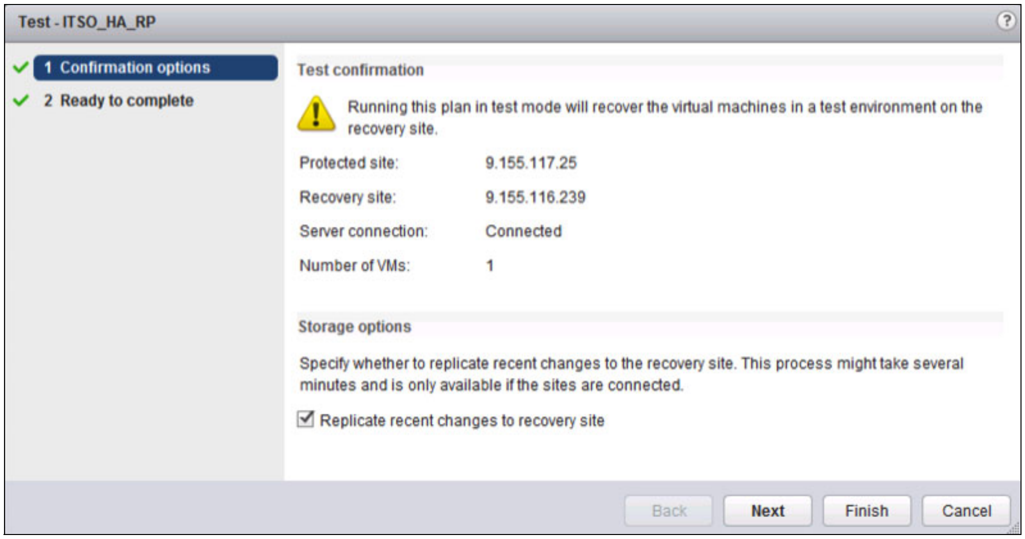
Plan status: Ready

Description: This plan is ready for test or recovery.

1. Discover protected VM information
2. Synchronize protected VM information
3. Synchronize storage
4. Restore recovery site hosts from standby

Figure 8-21 Clicking the Test Recovery Plan icon

2. Confirm the protected and recovery sites and click **Next**, as shown in Figure 8-22.



Test confirmation

Running this plan in test mode will recover the virtual machines in a test environment on the recovery site.

Protected site:	9.155.117.25
Recovery site:	9.155.116.239
Server connection:	Connected
Number of VMs:	1

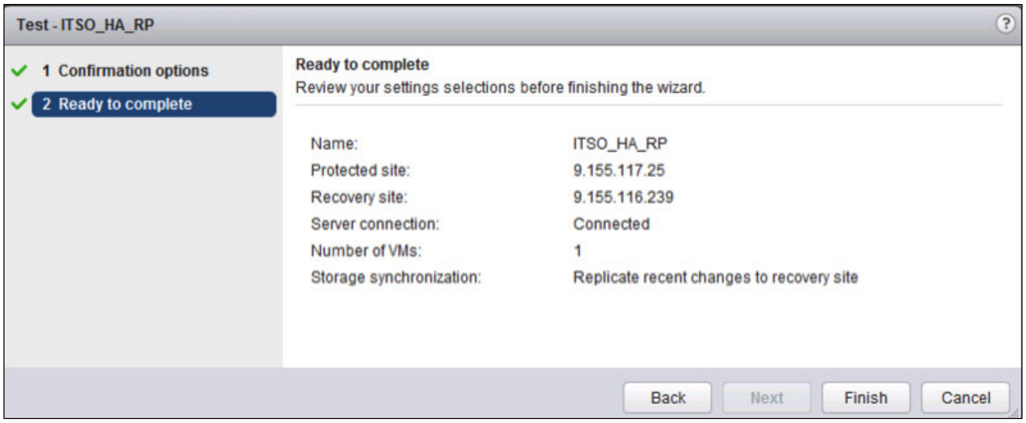
Storage options

Specify whether to replicate recent changes to the recovery site. This process might take several minutes and is only available if the sites are connected.

☑ Replicate recent changes to recovery site

Figure 8-22 Confirming the protected and recovery sites

3. Click **Finish** to start the recovery plan test, as shown in Figure 8-23.



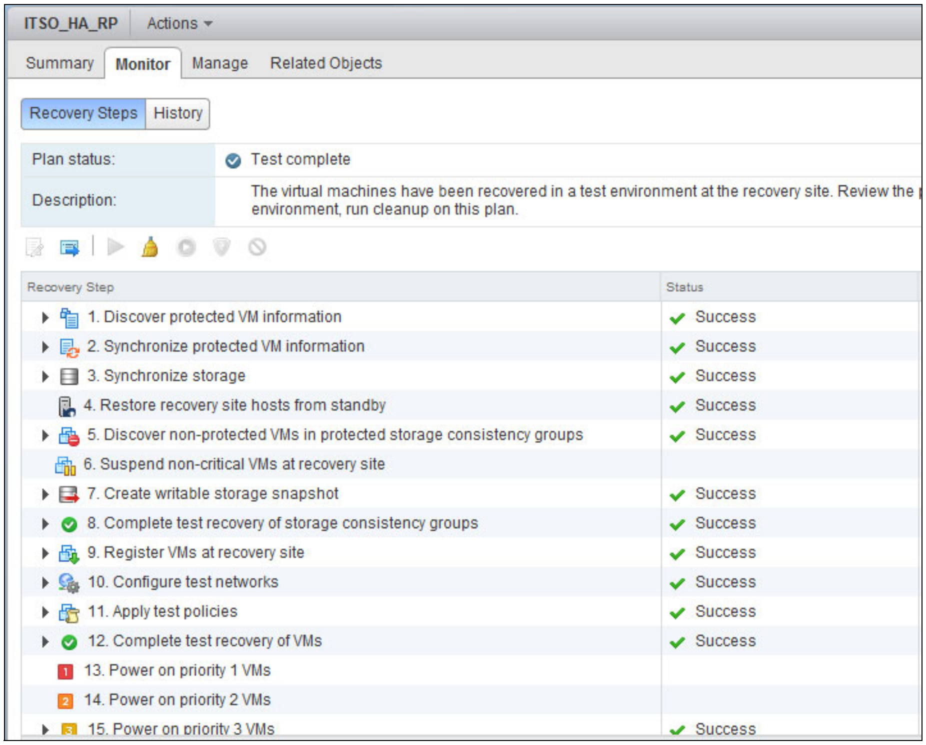
Ready to complete

Review your settings selections before finishing the wizard.

Name:	ITSO_HA_RP
Protected site:	9.155.117.25
Recovery site:	9.155.116.239
Server connection:	Connected
Number of VMs:	1
Storage synchronization:	Replicate recent changes to recovery site

Figure 8-23 Verifying the recovery plan and finishing

4. After completing the test, check the status of the operation for any failures and warnings on the Monitor tab. If the Test Recovery Plan returns Test Complete with no warnings, the recovery plan is ready for a planned migration or disaster recovery operation, as shown in Figure 8-24.

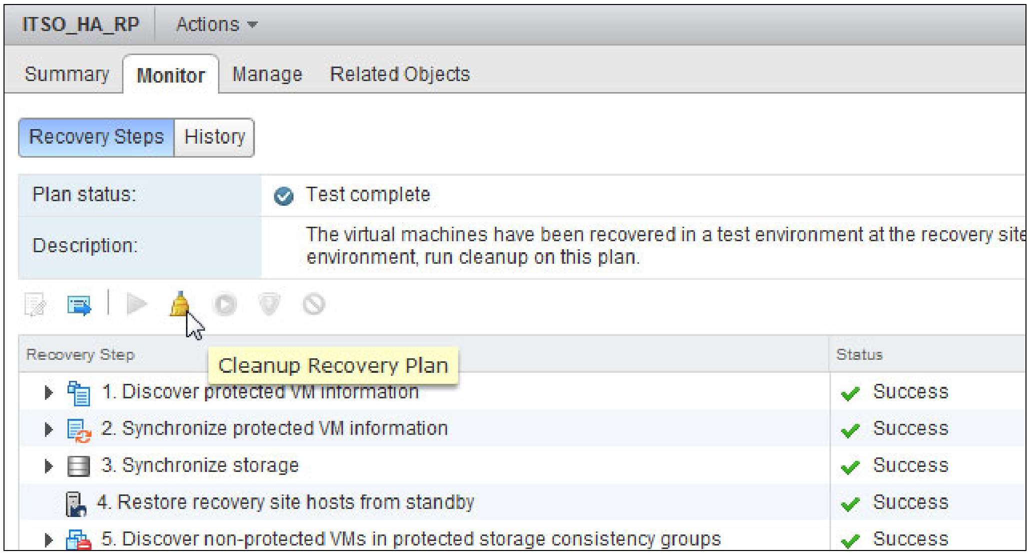


Plan status:	Test complete
Description:	The virtual machines have been recovered in a test environment at the recovery site. Review the environment, run cleanup on this plan.

Recovery Step	Status
1. Discover protected VM information	✓ Success
2. Synchronize protected VM information	✓ Success
3. Synchronize storage	✓ Success
4. Restore recovery site hosts from standby	✓ Success
5. Discover non-protected VMs in protected storage consistency groups	✓ Success
6. Suspend non-critical VMs at recovery site	
7. Create writable storage snapshot	✓ Success
8. Complete test recovery of storage consistency groups	✓ Success
9. Register VMs at recovery site	✓ Success
10. Configure test networks	✓ Success
11. Apply test policies	✓ Success
12. Complete test recovery of VMs	✓ Success
13. Power on priority 1 VMs	
14. Power on priority 2 VMs	
15. Power on priority 3 VMs	✓ Success

Figure 8-24 Successful test completion

5. After successfully running a recover plan test, click the **Cleanup Recovery Plan** icon to return the recovery plan to a Ready state. This cleanup operation is necessary before the recovery plan or a failover are run, as shown in Figure 8-25.

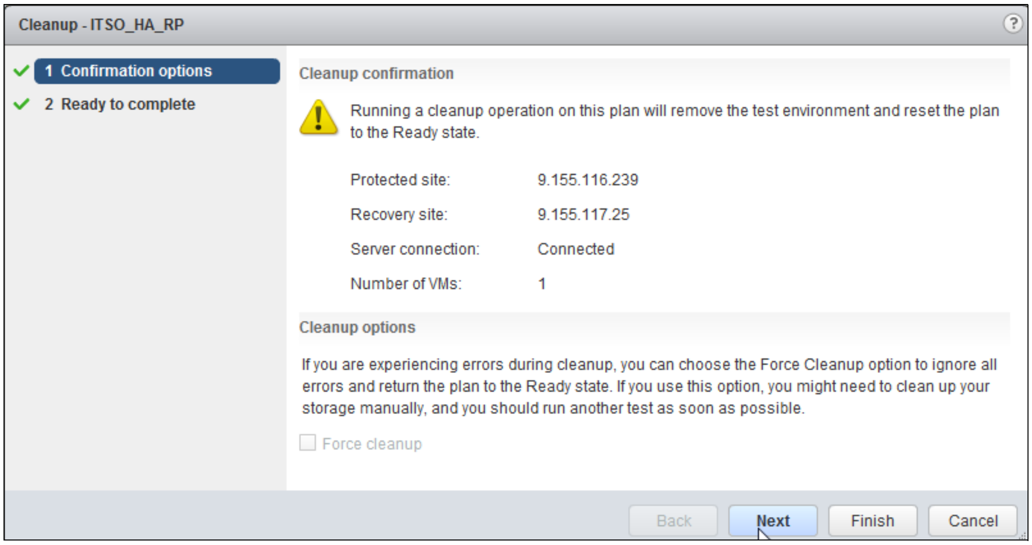


Plan status:	Test complete
Description:	The virtual machines have been recovered in a test environment at the recovery site environment, run cleanup on this plan.

Recovery Step	Status
1. Discover protected VM information	✓ Success
2. Synchronize protected VM information	✓ Success
3. Synchronize storage	✓ Success
4. Restore recovery site hosts from standby	✓ Success
5. Discover non-protected VMs in protected storage consistency groups	✓ Success

Figure 8-25 Clicking the Cleanup Recovery Plan icon

6. In the Confirm operations window click **Next**, as shown in Figure 8-26.



Cleanup confirmation

Running a cleanup operation on this plan will remove the test environment and reset the plan to the Ready state.

Protected site:	9.155.116.239
Recovery site:	9.155.117.25
Server connection:	Connected
Number of VMs:	1

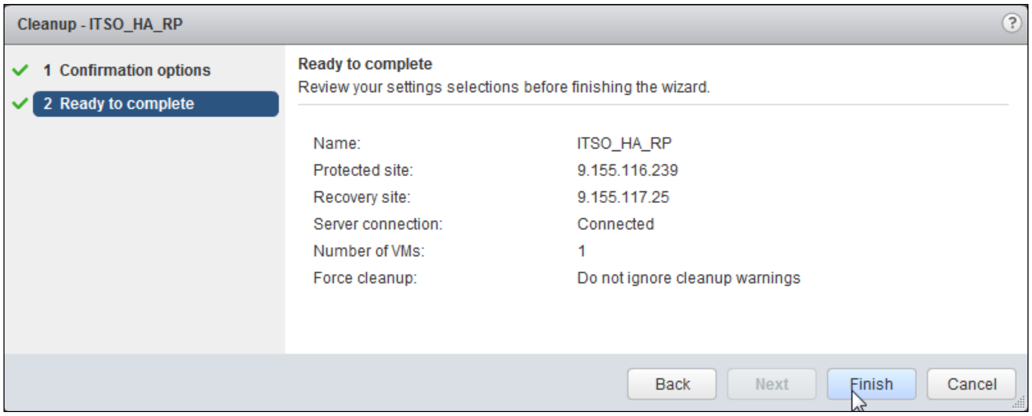
Cleanup options

If you are experiencing errors during cleanup, you can choose the Force Cleanup option to ignore all errors and return the plan to the Ready state. If you use this option, you might need to clean up your storage manually, and you should run another test as soon as possible.

□ Force cleanup

Figure 8-26 Confirming the cleanup operation

7. In the Ready to complete window, click **Finish,** as shown in Figure 8-27.



Ready to complete

Review your settings selections before finishing the wizard.

Name:	ITSO_HA_RP
Protected site:	9.155.116.239
Recovery site:	9.155.117.25
Server connection:	Connected
Number of VMs:	1
Force cleanup:	Do not ignore cleanup warnings

Figure 8-27 Confirming and starting the cleanup operation

8.6.4 Failing over a recovery plan

VMware Site Recovery Manager provides the following types of recovery options:

- In case of *planned migration*, the recovery plan recovers the virtual machines while both sites are running. If an error occurs at the protected site during recovery, planned migration will fail.
- In case of *disaster recovery*, the recovery plan recovers the virtual machines to the recovery site when the protected site encounters a problem. If an error occurs at the protected site, disaster recovery will continue and will not fail.

After testing the recovery plan successfully, the recovery plan is ready for planned migration or disaster recovery operations. Before running a failover, resolve any errors or warnings in the *Test Recovery Plan* operation.

Before a recovery plan is run, the virtual machine or machines that are to be failed must be at the protected site, as shown in Figure 8-28.

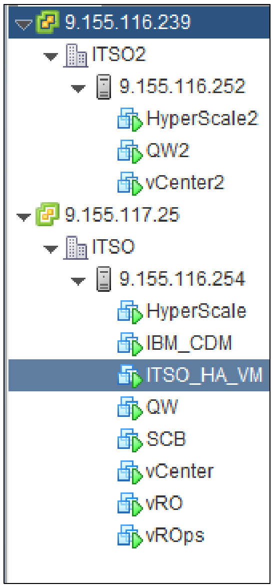


Figure 8-28 VM location prior to recovery operation

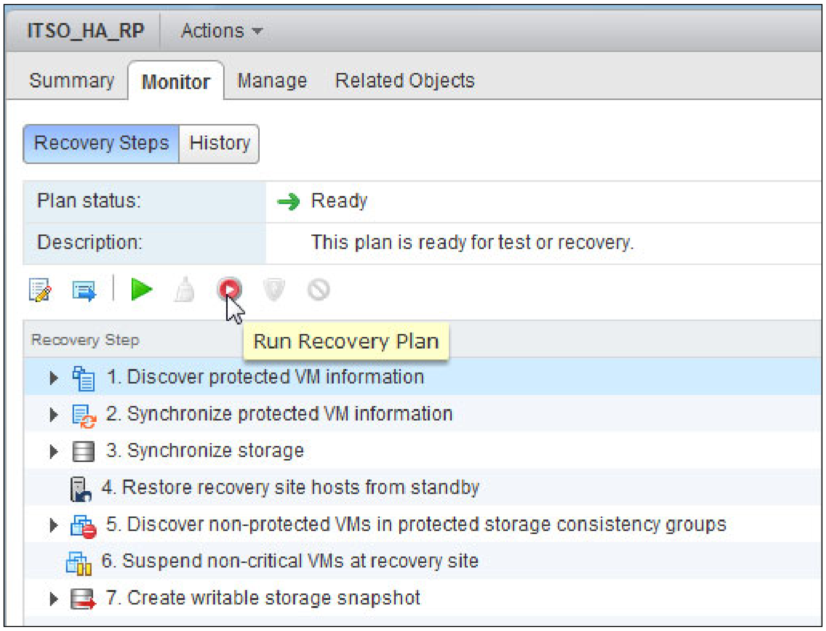
At this point, the HyperSwap volume is on FlashSystem A9000 designated as the secondary. When the migration is completed, it shows as the primary, as shown in Example 8-1.

Example 8-1 HyperSwap volume information about the FlashSystem A9000 prior to recovery operation

```
A9000>>ha_list vol=ITSO_HA_VOLUME3 -x

<XCLIRETURN STATUS="SUCCESS" COMMAND_LINE="ha_list vol=ITSO_HA_VOLUME3 -x">
   <OUTPUT>
      <ha id="5dc19b00006">
         <creator value="admin"/>
         <creator_category value="storageadmin"/>
         <id value="5dc19b00006"/>
         <domain_uid value="-1"/>
         <local_peer_id value="5881920005b"/>
         <local_peer_name value="ITSO_HA_VOLUME3"/>
         <designation value="Secondary"/>
         <current_role value="Slave"/>
         <remote_mirror_id value="6e61a800006"/>
         <remote_peer_name value="ITSO_HA_VOLUME3"/>
         <target_id value="59719000000"/>
         <target_name value="A9000R"/>
         <sync_state value="Synchronized"/>
         <active value="yes"/>
         <ha_connected value="yes"/>
         <operational value="yes"/>
         <sync_progress value="100"/>
         <size_to_synchronize value="-1"/>
         <estimated_sync_time value="0"/>
         <mirror_error value="No_Error"/>
         <ha_object value="Volume"/>
         <init_type value="online"/>
         <crash_consistent value="Consistent"/>
         <validate value="no"/>
         <ha_high_availability_state value="Enable"/>
         <ha_unavailable_reason value="N/A"/>
         <ha_sync_state value="Enable"/>
         <ha_object_state value="Unowned"/>
         <automatic_failover value="Active"/>
         <automatic_failover_reason value="N/A"/>
         <io_service value="Active"/>
      </ha>
   </OUTPUT>
</XCLIRETURN>
A9000>>
```

Complete the following steps:

1. In the vSphere web client, select the **Site Recovery** plug-in. Then, select the recovery plan and click the **Run Recovery Plan** icon, as shown in Figure 8-29.



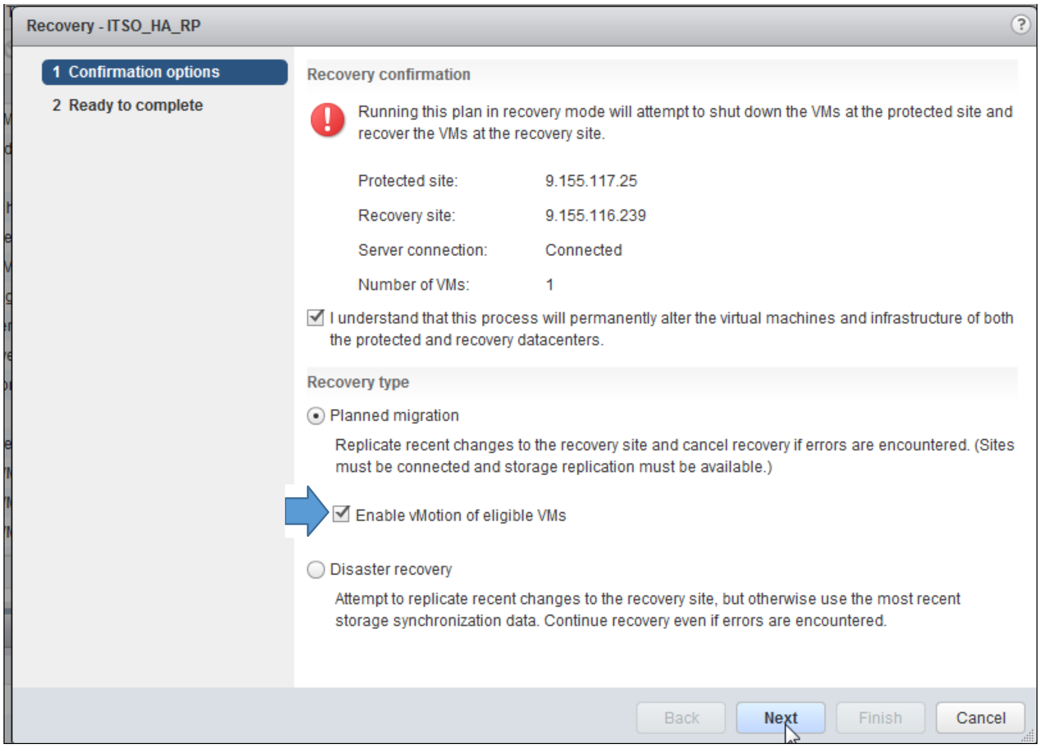
Plan status: Ready

Description: This plan is ready for test or recovery.

1. Discover protected VM information
2. Synchronize protected VM information
3. Synchronize storage
4. Restore recovery site hosts from standby
5. Discover non-protected VMs in protected storage consistency groups
6. Suspend non-critical VMs at recovery site
7. Create writable storage snapshot

Figure 8-29 Clicking the Run Recovery Plan icon

2. Select the option to provide recovery confirmation and select the recovery type. In this example, Planned migration is used. If planned migration is selected, ensure that Enable vMotion of eligible VMs is also selected, as shown in Figure 8-30.



Recovery - ITSO_HA_RP

1 Confirmation options

2 Ready to complete

Recovery confirmation

Running this plan in recovery mode will attempt to shut down the VMs at the protected site and recover the VMs at the recovery site.

Protected site: 9.155.117.25

Recovery site: 9.155.116.239

Server connection: Connected

Number of VMs: 1

☑ I understand that this process will permanently alter the virtual machines and infrastructure of both the protected and recovery datacenters.

Recovery type

● Planned migration

Replicate recent changes to the recovery site and cancel recovery if errors are encountered. (Sites must be connected and storage replication must be available.)

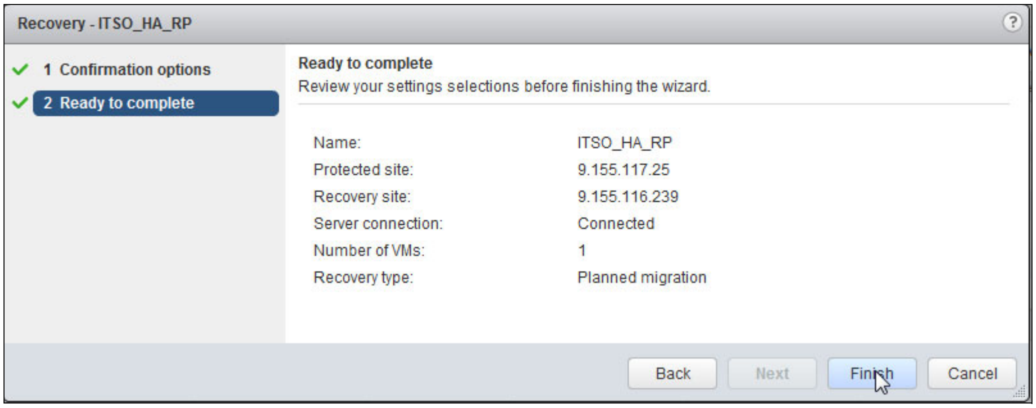
☑ Enable vMotion of eligible VMs

○ Disaster recovery

Attempt to replicate recent changes to the recovery site, but otherwise use the most recent storage synchronization data. Continue recovery even if errors are encountered.

Figure 8-30 Defining the recover options

3. In the Ready to complete window, verify all the settings and click **Finish** to run the recovery plan, as shown in Figure 8-31.



Recovery - ITSO_HA_RP

✓ 1 Confirmation options

✓ 2 Ready to complete

Ready to complete

Review your settings selections before finishing the wizard.

Name: ITSO_HA_RP

Protected site: 9.155.117.25

Recovery site: 9.155.116.239

Server connection: Connected

Number of VMs: 1

Recovery type: Planned migration

Figure 8-31 Verifying the recovery options

When the recovery plan completes successfully, the virtual machine or machines are now shown at the recovery site (see Figure 8-32).

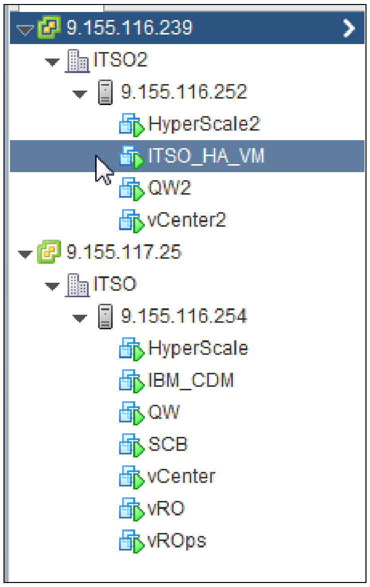


Figure 8-32 VM location after recovery operation

At this point, the HyperSwap volume appears on the A9000 as primary, as shown in Example 8-2.

Example 8-2 HyperSwap volume information about the FlashSystem A9000 after recovery

```
A9000>>ha_list vol=ITSO_HA_VOLUME3 -x
<XCLIRETURN STATUS="SUCCESS" COMMAND_LINE="ha_list vol=ITSO_HA_VOLUME3 -x">
   <OUTPUT>
      <ha id="5dc19b00006">
         <creator value="admin"/>
         <creator_category value="storageadmin"/>
         <id value="5dc19b00006"/>
         <domain_uid value="-1"/>
         <local_peer_id value="5881920005b"/>
         <local_peer_name value="ITSO_HA_VOLUME3"/>
         <designation value="Primary"/>
         <current_role value="Master"/>
         <remote_mirror_id value="6e61a800006"/>
         <remote_peer_name value="ITSO_HA_VOLUME3"/>
         <target_id value="59719000000"/>
         <target_name value="A9000R"/>
         <sync_state value="Synchronized"/>
         <active value="yes"/>
         <ha_connected value="yes"/>
         <operational value="yes"/>
         <sync_progress value="100"/>
         <size_to_synchronize value="-1"/>
         <estimated_sync_time value="0"/>
         <mirror_error value="No_Error"/>
         <ha_object value="Volume"/>
         <init_type value="online"/>
         <crash_consistent value="Consistent"/>
```

```
            <validate value="no"/>
            <ha_high_availability_state value="Enable"/>
            <ha_unavailable_reason value="N/A"/>
            <ha_sync_state value="Enable"/>
            <ha_object_state value="Unowned"/>
            <automatic_failover value="N/A"/>
            <automatic_failover_reason value="N/A"/>
            <io_service value="Active"/>
         </ha>
      </OUTPUT>
   </XCLIRETURN>
A9000>>
```

8.6.5 Re-protecting virtual machines after recovery

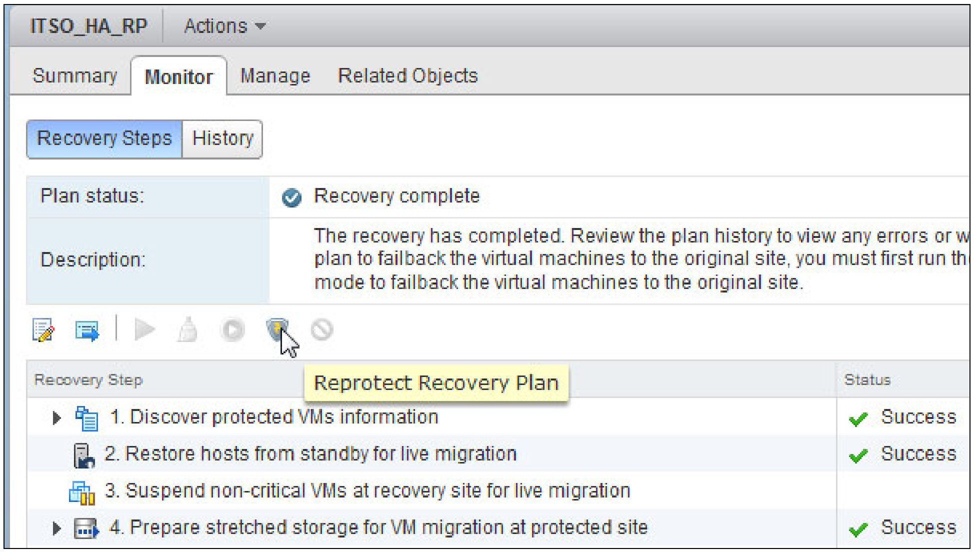
After recovery is complete, the recovery site becomes the primary site, but the virtual machines are not protected. Manually re-establishing protection in the opposite direction by re-creating all protection groups and recovery plans is time-consuming and prone to errors. SRM provides the reprotect function, which is an automated way to revert protection and automates the process.

By running *reprotect*, when the protected site comes back online, the direction of replication can be reversed to protect the recovered virtual machines on the recovery site back to the original protected site. The reprotect operation can be initiated only after successful recovery without any errors.

The reprotect operation reverses the direction of protection, and then forces the synchronization of virtual machines from the new protected site to the new recovery site.

To perform a reprotect operation, complete the following steps:

1. In the vSphere web client, click the **Site Recovery** plug-in. Then, select a recovery plan and click the **Reprotect Recovery Plan** icon, as shown in Figure 8-33.



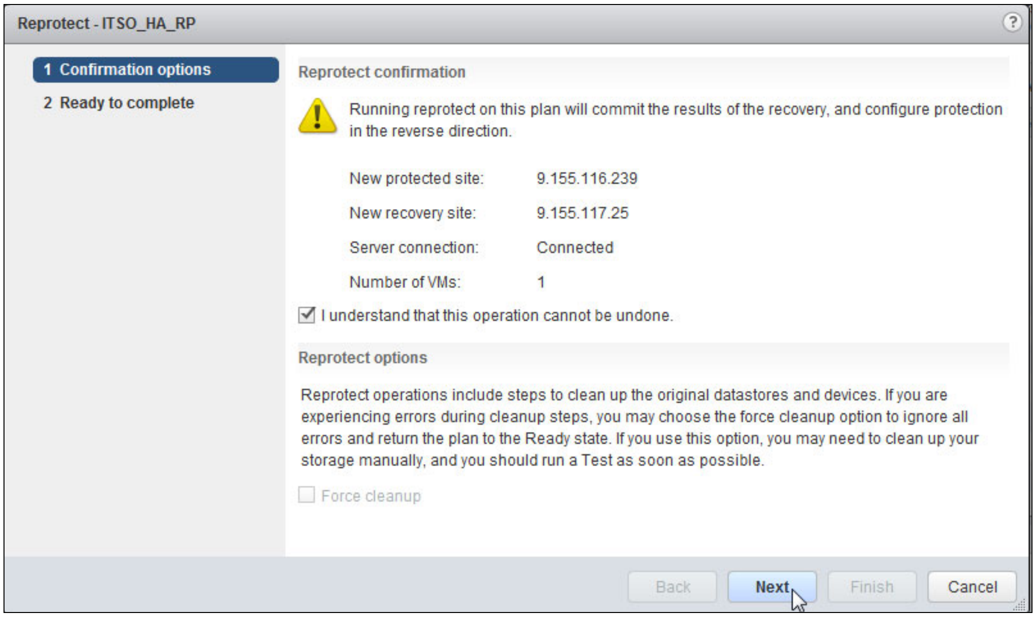
Plan status: Recovery complete

Description: The recovery has completed. Review the plan history to view any errors or w
plan to failback the virtual machines to the original site, you must first run th
mode to failback the virtual machines to the original site.

Recovery Step	Status
1. Discover protected VMs information	Success
2. Restore hosts from standby for live migration	Success
3. Suspend non-critical VMs at recovery site for live migration	
4. Prepare stretched storage for VM migration at protected site	Success

Figure 8-33   Clicking the Reprotect Recovery Plan icon

2. Provide the reprotect confirmation and click **Next**, as shown in Figure 8-34.



Reprotect confirmation

Running reprotect on this plan will commit the results of the recovery, and configure protection in the reverse direction.

New protected site:	9.155.116.239
New recovery site:	9.155.117.25
Server connection:	Connected
Number of VMs:	1

☑ I understand that this operation cannot be undone.

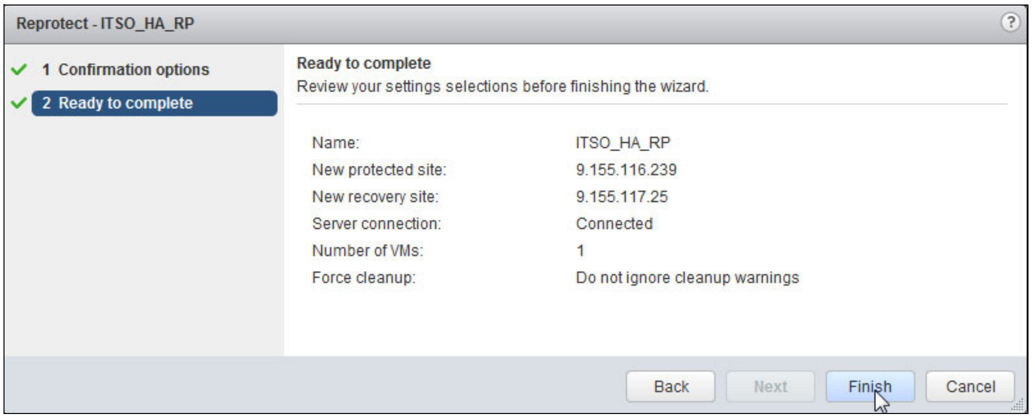
Reprotect options

Reprotect operations include steps to clean up the original datastores and devices. If you are experiencing errors during cleanup steps, you may choose the force cleanup option to ignore all errors and return the plan to the Ready state. If you use this option, you may need to clean up your storage manually, and you should run a Test as soon as possible.

□ Force cleanup

Figure 8-34 Confirming the reprotect operation options

3. Verify the information in the Ready to complete window and click **Finish** to reprotect the recovery plan, as shown in Figure 8-35.



Ready to complete

Review your settings selections before finishing the wizard.

Name:	ITSO_HA_RP
New protected site:	9.155.116.239
New recovery site:	9.155.117.25
Server connection:	Connected
Number of VMs:	1
Force cleanup:	Do not ignore cleanup warnings

Figure 8-35 Verifying the reprotect operation options

4. After the recovery operation is completed and the virtual machine or machines are reprotected, run the test recovery to ensure that the new configuration of protected and recovery sites completes successfully.

9

Microsoft Failover Clustering and HyperSwap

This chapter shows how to use Microsoft Failover Clustering between two sites that use HyperSwap for FlashSystem A9000 and A9000R. Do not confuse this solution with the IBM Multi-site HA/DR solution that was described in Chapter 6, “Multi-site HA/DR implementation and usage” on page 97.

This chapter covers the following topics:

- 9.1, “Microsoft Failover Clustering configuration” on page 190
- 9.2, “Planned and unplanned failovers” on page 193

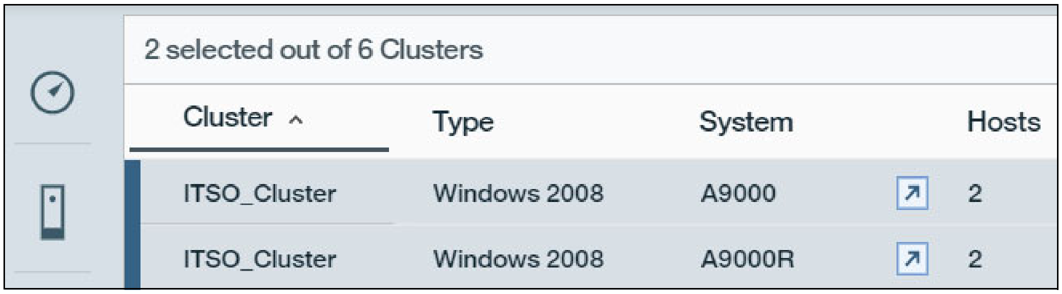
9.1 Microsoft Failover Clustering configuration

IBM FlashSystem A9000 and A9000R supports Microsoft Failover Clustering, starting with Windows Server version 2008. For more details about the supported versions, see the System Storage Interoperability Center (SSIC).

For more information about installing and configuring Microsoft Failover Clustering using Windows 2008 R2, see the Microsoft publication Failover Clustering Overview.

> **Note:** The following illustrations are based on Microsoft SQL Server 2012 Failover Cluster running on Windows 2008 R2, Hyper-Scale Manager 5.4, and FlashSystem A9000 and A9000R Version 12.2.1.

If the hosts on FlashSystem A9000 and A9000R are defined by Host Attachment Kit, the host type *Windows 2008* is automatically chosen. Make sure that for host and cluster type, *Windows 2008* on FlashSystem A9000 and A9000R is selected. In the Hyper-Scale Manager UI, select **HOST & CLUSTERS VIEWS** → **Clusters** to show the type**,** as shown in Figure 9-1. If another host and cluster type is chosen, it might result in a stop error on one Microsoft cluster node after a failover.



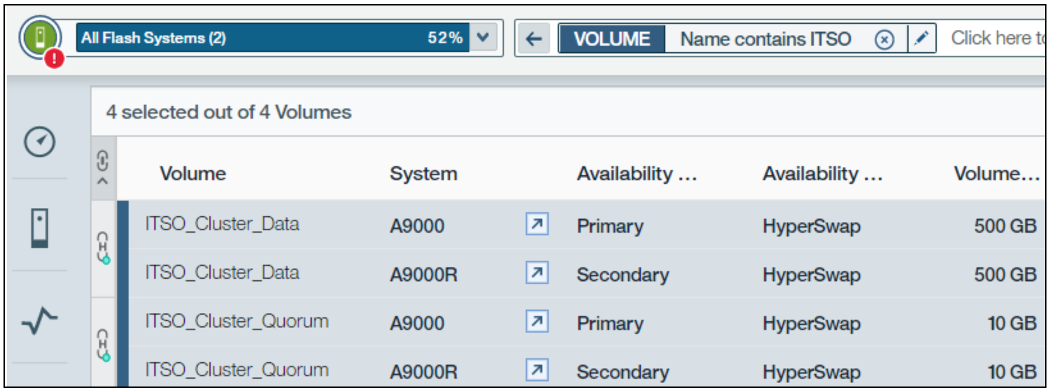
2 selected out of 6 Clusters

Cluster ^	Type	System		Hosts
ITSO_Cluster	Windows 2008	A9000	↗	2
ITSO_Cluster	Windows 2008	A9000R	↗	2

Figure 9-1 Cluster on FlashSystem A9000 and A9000R

The HyperSwap volumes are configured and mapped to the cluster, and an SQL Server 2012 Failover Cluster is running on one of these volumes.

Volume *ITSO_Cluster_Data* holds the SQL server data and *ITSO_Cluster_Quorum* is the Microsoft Cluster Quorum Disk Witness. Select **POOLS & VOLUMES VIEWS** → **Volumes** and filter for the volumes, as shown in Figure 9-2.



4 selected out of 4 Volumes

Volume	System		Availability ...	Availability ...	Volume...
ITSO_Cluster_Data	A9000	↗	Primary	HyperSwap	500 GB
ITSO_Cluster_Data	A9000R	↗	Secondary	HyperSwap	500 GB
ITSO_Cluster_Quorum	A9000	↗	Primary	HyperSwap	10 GB
ITSO_Cluster_Quorum	A9000R	↗	Secondary	HyperSwap	10 GB

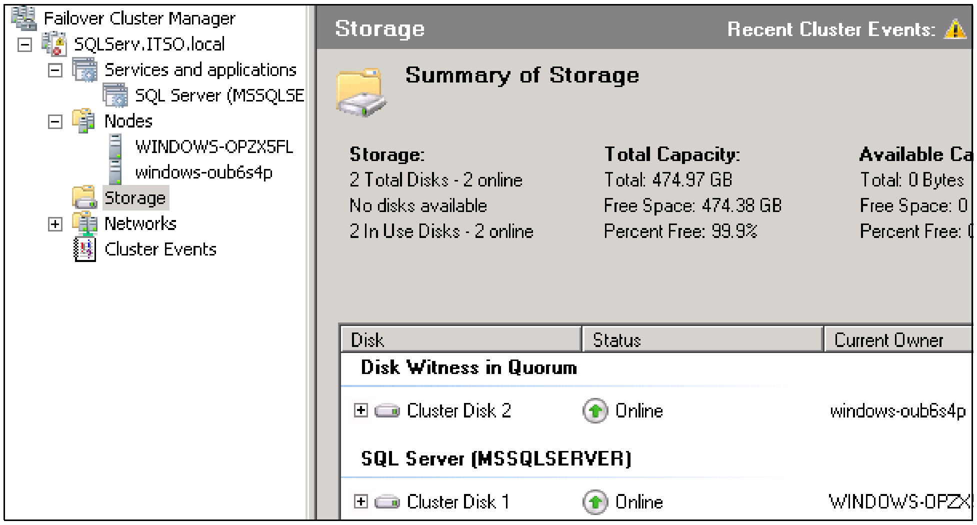
Figure 9-2 Failover Cluster HyperSwap volumes

To list the HyperSwap volumes on one of the cluster nodes, open a command prompt or a PowerShell window and run the **`xiv_devlist.exe`** command, as shown in Example 9-1.

Example 9-1 FlashSystem A9000 and A9000R volumes on Cluster node ITSO-WIN2

```
PS C:\Users\Administrator.ITSO> xiv_devlist
IBM storage devices
-------------------------------------------------------------------------------------------------
Device              Size (GB)  Paths  Vol Name             Vol ID  Storage ID  Storage Type
-------------------------------------------------------------------------------------------------
\\.\PHYSICALDRIVE1  500.0      8/8    ITSO_Cluster_Data    27130   1322131     FlashSystem A9000
-------------------------------------------------------------------------------------------------
\\.\PHYSICALDRIVE2  10.0       8/8    ITSO_Cluster_Quorum  27131   1322131     FlashSystem A9000
-------------------------------------------------------------------------------------------------


Non-IBM storage devices
------------------------------------
Device              Size (GB)  Paths
------------------------------------
\\.\PHYSICALDRIVE0  199.0      N/A
------------------------------------
```

Both disks are shown with their assignment in the Failover Cluster Manager, as shown in Figure 9-3.



Summary of Storage

Storage:
2 Total Disks - 2 online
No disks available
2 In Use Disks - 2 online

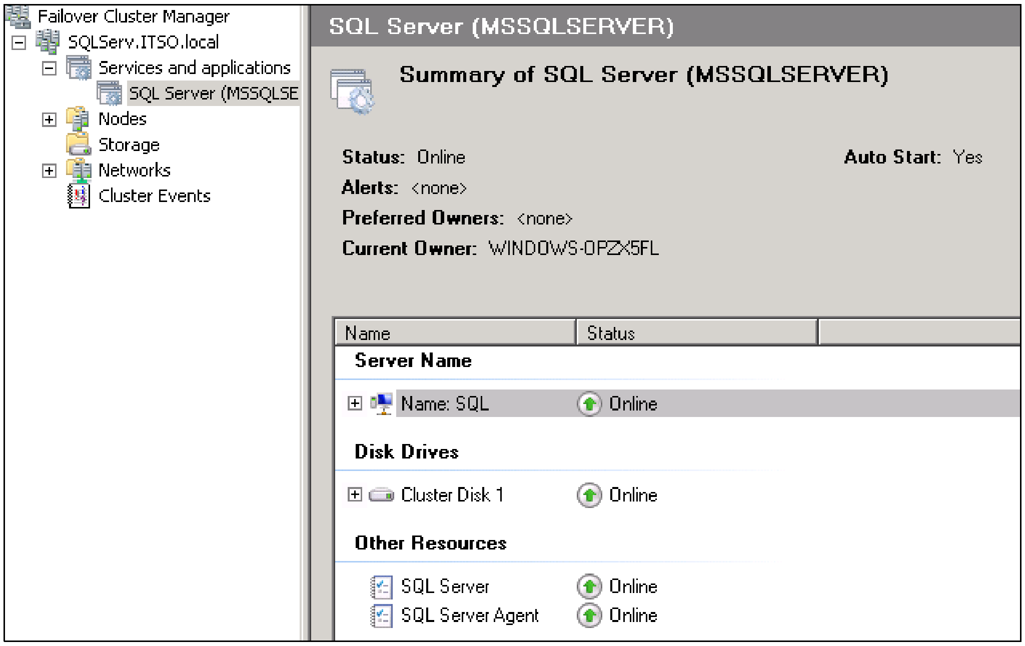
Total Capacity:
Total: 474.97 GB
Free Space: 474.38 GB
Percent Free: 99.9%

Available Ca
Total: 0 Bytes
Free Space: 0
Percent Free: (

Disk	Status	Current Owner
Disk Witness in Quorum		
Cluster Disk 2	Online	windows-oub6s4p
SQL Server (MSSQLSERVER)		
Cluster Disk 1	Online	WINDOWS-OPZX

Figure 9-3 Failover Cluster Manager Disks

The SQL Server is running on Node WINDOWS-0PZX5FL, as shown in Figure 9-4.



SQL Server (MSSQLSERVER)

Summary of SQL Server (MSSQLSERVER)

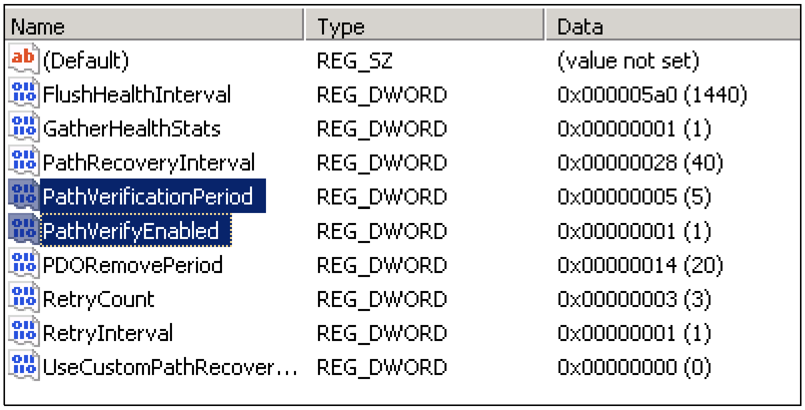
Status: Online **Auto Start:** Yes
Alerts: <none>
Preferred Owners: <none>
Current Owner: WINDOWS-0PZX5FL

Name	Status
Server Name	
Name: SQL	Online
Disk Drives	
Cluster Disk 1	Online
Other Resources	
SQL Server	Online
SQL Server Agent	Online

Figure 9-4 Failover Cluster Manager Roles

Tip: To speed up path failover, change two registry values on all Microsoft cluster nodes. In HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\services\mpio\Parameters, set the parameter PathVerifyEnabled to 1 and PathVerificationPeriod to 5, as shown in Figure 9-5. These two changes allow faster path failure detection, but are not mandatory.

For more information about these parameters, see this web page.



Name	Type	Data
(Default)	REG_SZ	(value not set)
FlushHealthInterval	REG_DWORD	0x000005a0 (1440)
GatherHealthStats	REG_DWORD	0x00000001 (1)
PathRecoveryInterval	REG_DWORD	0x00000028 (40)
PathVerificationPeriod	REG_DWORD	0x00000005 (5)
PathVerifyEnabled	REG_DWORD	0x00000001 (1)
PDORemovePeriod	REG_DWORD	0x00000014 (20)
RetryCount	REG_DWORD	0x00000003 (3)
RetryInterval	REG_DWORD	0x00000001 (1)
UseCustomPathRecover...	REG_DWORD	0x00000000 (0)

Figure 9-5 Registry parameters

9.1.1 File Share Witness or Disk Witness

Microsoft’s recommendation for failover clustering is to use *Node and File Share Majority* (file share witness) as quorum option, which requires a file share on a third site. But with FlashSystem A9000 and A9000R HyperSwap volumes, the use of *Node and Disk Majority* (disk witness) as quorum option is even better because the HyperSwap volumes are available to both sites. In split-brain situations with file share witness, that Microsoft cluster might decide that one site is the surviving site while FlashSystem A9000 and A9000R decide the other site is the surviving site. This situation can lead to an access loss. More details about the requirements and recommendations from Microsoft for a Failover Cluster can be found at Requirements and Recommendations for a Multi-Site Failover Cluster.

The scenarios in 9.2, “Planned and unplanned failovers” use a disk witness for the Microsoft failover cluster.

9.2 Planned and unplanned failovers

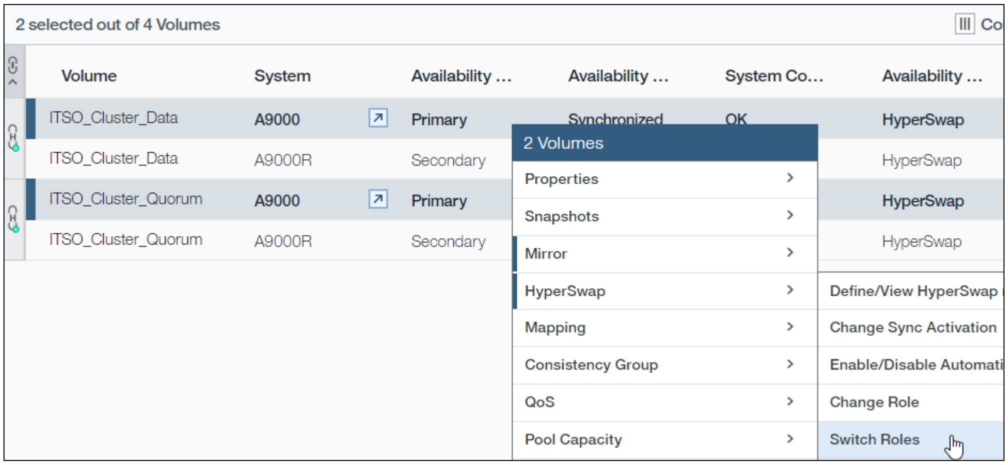
This section describes the following failover scenarios for IBM HyperSwap on FlashSystem A9000 and FlashSystem A9000R with Microsoft Multi-Site failover clustering:

- Planned failover
- All Fibre Channel connections lost on the owner node
- Site failure

9.2.1 Planned failover

You can use both IBM Hyper-Scale Manager and Microsoft Failover Cluster Manager to make it simple to move a clustered application to another site at metro distance. Complete the following steps:

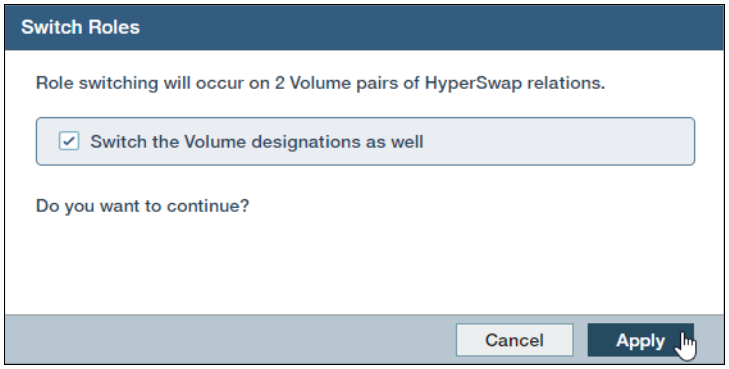
1. To fail over the HyperSwap volume from primary to secondary site, right-click the HyperSwap volumes and select **HyperSwap** → **Switch Roles**, as shown in Figure 9-6.


2 selected out of 4 Volumes

Volume	System	Availability ...	Availability ...	System Co...	Availability ...
ITSO_Cluster_Data	A9000	Primary	Synchronized	OK	HyperSwap
ITSO_Cluster_Data	A9000R	Secondary			HyperSwap
ITSO_Cluster_Quorum	A9000	Primary			HyperSwap
ITSO_Cluster_Quorum	A9000R	Secondary			HyperSwap

Figure 9-6 Switch Roles

2. Select **Switch the Volume designation as well** and click **Apply** to confirm the action, as shown in Figure 9-7.



Switch Roles

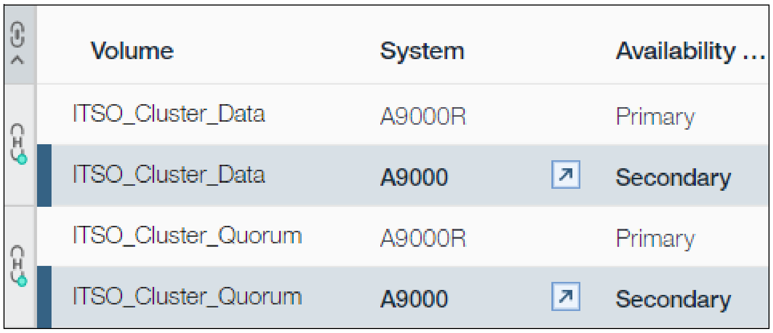
Role switching will occur on 2 Volume pairs of HyperSwap relations.

☑ Switch the Volume designations as well

Do you want to continue?

Figure 9-7 Prompt to confirm role switch for HyperSwap (Primary to Secondary)

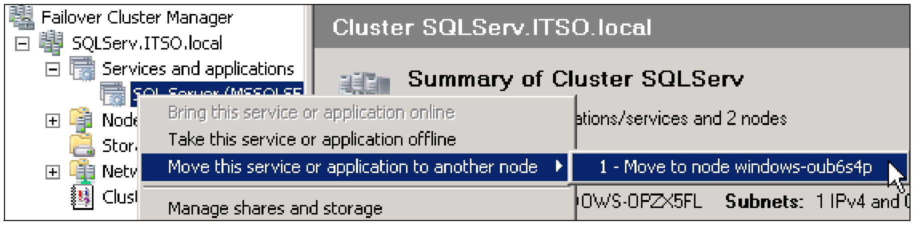
3. The HyperSwap volumes are switched, as shown in Figure 9-8.



Volume	System	Availability …
ITSO_Cluster_Data	A9000R	Primary
ITSO_Cluster_Data	A9000	Secondary
ITSO_Cluster_Quorum	A9000R	Primary
ITSO_Cluster_Quorum	A9000	Secondary

Figure 9-8 Roles switched

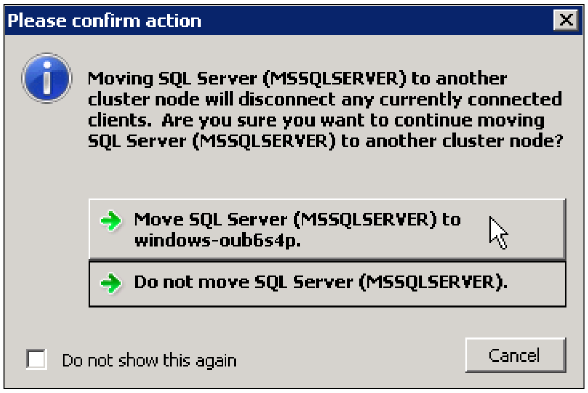
4. In the Failover Cluster Manager, right click the application and select **Move this service or application to another node** → **Move to node windows-oub6s4p**, as shown in Figure 9-9.



Failover Cluster Manager
SQLServ.ITSO.local
Services and applications
SQL Server (MSSQLSE
Node
Stor
Netv
Clus
Cluster SQLServ.ITSO.local
Summary of Cluster SQLServ
ations/services and 2 nodes
Bring this service or application online
Take this service or application offline
Move this service or application to another node
1 - Move to node windows-oub6s4p
Manage shares and storage
OWS-OPZX5FL
Subnets: 1 IPv4 and

Figure 9-9 Failover Cluster Manager: Failover to other node

5. Click **Move SQL Server (MSSQLSERVER) to windows-oub6s4p** as shown in Figure 9-10.



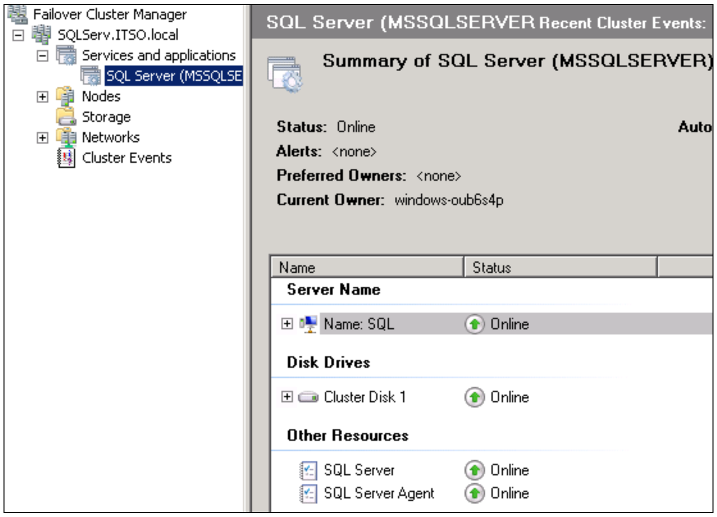
Please confirm action

Moving SQL Server (MSSQLSERVER) to another cluster node will disconnect any currently connected clients. Are you sure you want to continue moving SQL Server (MSSQLSERVER) to another cluster node?

Move SQL Server (MSSQLSERVER) to windows-oub6s4p.
Do not move SQL Server (MSSQLSERVER).
□ Do not show this again
Cancel

Figure 9-10 Failover Cluster Manager: Confirm Move of application

6. Application has moved to node `windows-oub6s4p`, as shown in Figure 9-11.



Summary of SQL Server (MSSQLSERVER)

Status: Online **Auto**

Alerts: <none>

Preferred Owners: <none>

Current Owner: windows-oub6s4p

Name	Status
Server Name	
Name: SQL	Online
Disk Drives	
Cluster Disk 1	Online
Other Resources	
SQL Server	Online
SQL Server Agent	Online

Figure 9-11 Failover Cluster Manager: Application moved

To return to the initial configuration, complete the same sequence of steps by using the other storage system and cluster node.

9.2.2 All Fibre Channel connections lost on the owner node

After the planned failover, the application resides on Site B. Assume that all Fibre Channel connections to the owner node (`windows-oub6s4p`) of the SQL Server Failover Cluster are lost, as shown in Figure 9-12.

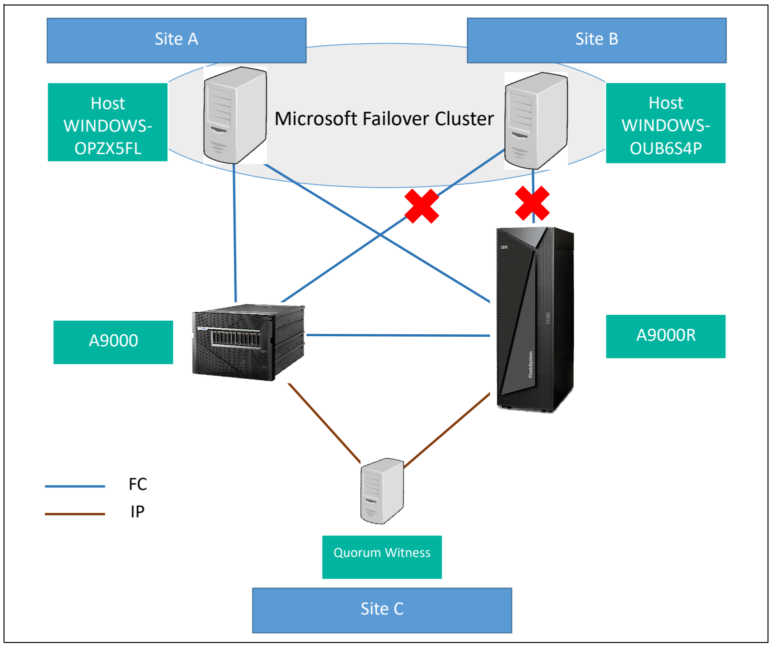


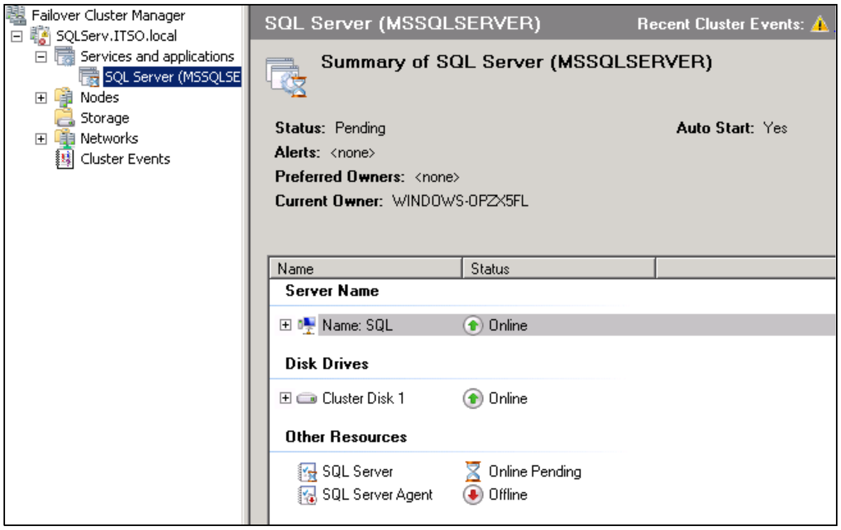
Figure 9-12 FC connection loss scenario

Using `xiv_devlist.exe` no longer lists the volumes, as shown in Example 9-2.

Example 9-2 FlashSystem A9000 and A9000R volumes on Cluster node windows-oub6s4p

```
PS C:\Users\Administrator.ITSO> xiv_devlist.exe
IBM storage devices
------------------------------------------------------------------------------------------
Device  Size (GB)  Paths  Vol Name  Vol ID  Storage ID  Storage Type  Hyper-Scale Mobility
------------------------------------------------------------------------------------------


Non-IBM storage devices
------------------------------------
Device              Size (GB)  Paths
------------------------------------
\\.\PHYSICALDRIVE0  200.0      N/A
------------------------------------
```

The SQL Server starts to fail over to the other cluster node `WINDOWS-OPZX5FL`, as shown in Figure 9-13.



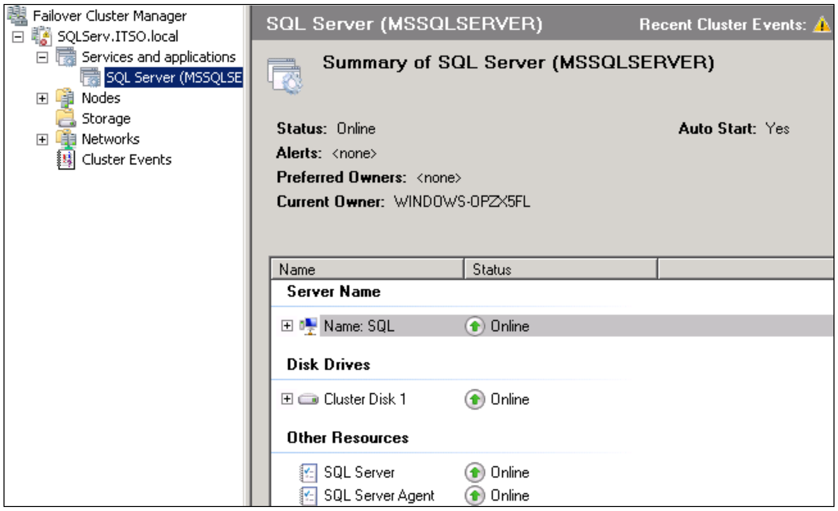
Summary of SQL Server (MSSQLSERVER)

Status: Pending **Auto Start:** Yes
Alerts: <none>
Preferred Owners: <none>
Current Owner: WINDOWS-OPZX5FL

Name	Status
Server Name	
Name: SQL	Online
Disk Drives	
Cluster Disk 1	Online
Other Resources	
SQL Server	Online Pending
SQL Server Agent	Offline

Figure 9-13 Failover Cluster Manager Roles: Failover starts

After the failover finishes, the SQL Server is running on `WINDOWS-OPZX5FL`, as shown in Figure 9-14.



Summary of SQL Server (MSSQLSERVER)

Status: Online **Auto Start:** Yes
Alerts: <none>
Preferred Owners: <none>
Current Owner: WINDOWS-OPZX5FL

Name	Status
Server Name	
Name: SQL	Online
Disk Drives	
Cluster Disk 1	Online
Other Resources	
SQL Server	Online
SQL Server Agent	Online

Figure 9-14 Failover Cluster Manager Roles: Failover finished

9.2.3 Site failure

Assume a site failure of the owner node of the SQL server (`WINDOWS-OPZX5FL`) and FlashSystem A9000 where the primary volume resides, as shown in Figure 9-15.

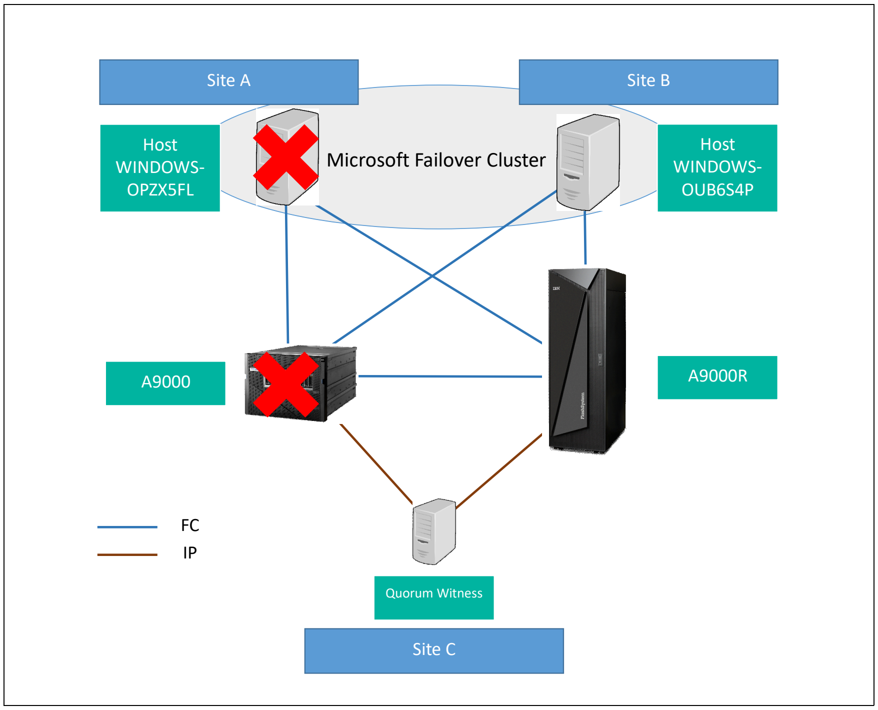


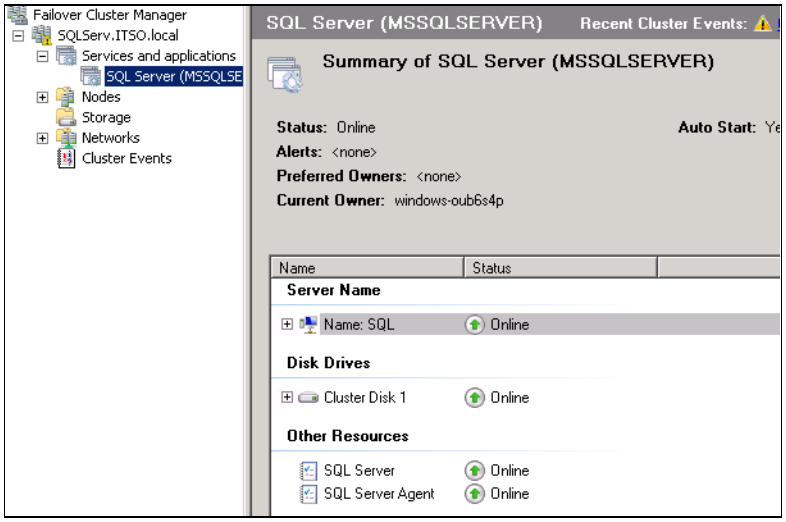
Figure 9-15 Site failure scenario

Only two paths connected to FlashSystem A9000R on the surviving node (`windows-oub6s4p`) are left, as shown in Example 9-3.

Example 9-3 FlashSystem A9000 volumes on Cluster node windows-oub6s4pafter site failure

```
PS C:\Users\Administrator.ITSO> xiv_devlist.exe
IBM storage devices
-------------------------------------------------------------------------------------------------
Device              Size (GB)  Paths  Vol Name             Vol ID  Storage ID  Storage Type
-------------------------------------------------------------------------------------------------
\\.\PHYSICALDRIVE1  500.0      2/2    ITSO_Cluster_Data    27130   1320902     FlashSystem A9000R
-------------------------------------------------------------------------------------------------
\\.\PHYSICALDRIVE2  10.0       2/2    ITSO_Cluster_Quorum  27131   1320902     FlashSystem A9000R
-------------------------------------------------------------------------------------------------

Non-IBM storage devices
------------------------------------
Device              Size (GB)  Paths
------------------------------------
\\.\PHYSICALDRIVE0  200.0      N/A
----------------------------------------
```

After the Microsoft cluster failover completes, the SQL Server is running on `windows-oub6s4p` as shown in Figure 9-16.



Failover Cluster Manager
- SQLServ.ITSO.local
  - Services and applications
    - SQL Server (MSSQLSE
  - Nodes
  - Storage
  - Networks
  - Cluster Events

SQL Server (MSSQLSERVER) Recent Cluster Events:

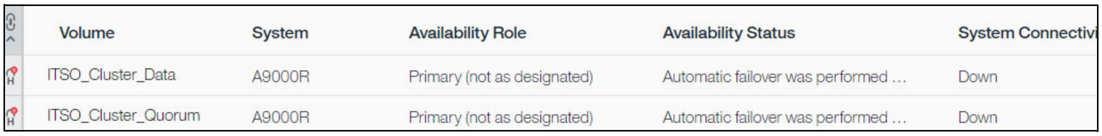
Summary of SQL Server (MSSQLSERVER)

Status: Online **Auto Start:** Ye
Alerts: <none>
Preferred Owners: <none>
Current Owner: windows-oub6s4p

Name	Status
Server Name	
Name: SQL	Online
Disk Drives	
Cluster Disk 1	Online
Other Resources	
SQL Server	Online
SQL Server Agent	Online

Figure 9-16 Failover Cluster Manager Roles: Failover finished after site failure

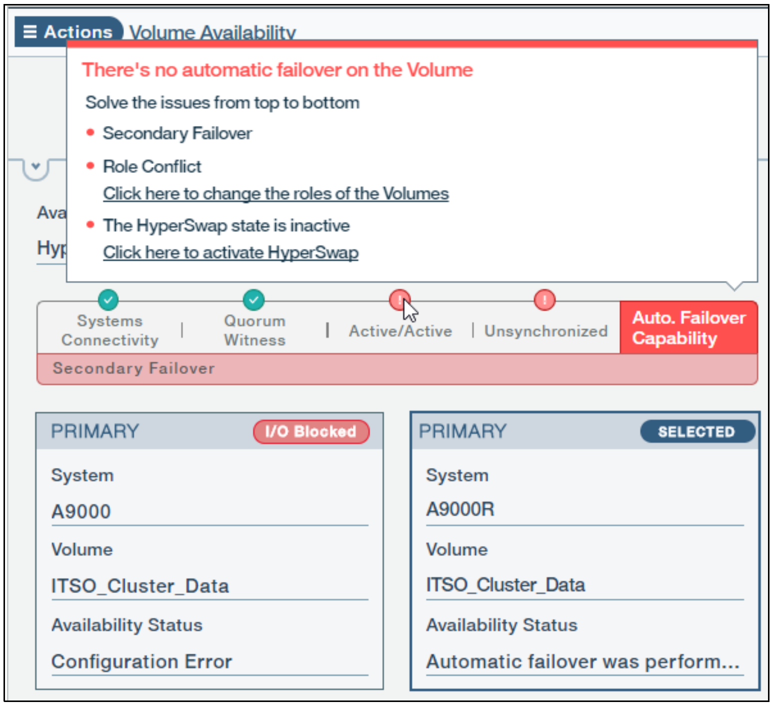
A9000 is down and the automatic failover for the HyperSwap volumes occurred, as shown in Figure 9-17.



Volume	System	Availability Role	Availability Status	System Connectivi
ITSO_Cluster_Data	A9000R	Primary (not as designated)	Automatic failover was performed ...	Down
ITSO_Cluster_Quorum	A9000R	Primary (not as designated)	Automatic failover was performed ...	Down

Figure 9-17 Automatic HyperSwap volumes failover

After System A comes up again, all paths and all connections are back, but the HyperSwap volumes are not synchronized. Select **REMOTE VIEWS** → **Mirrored/HyperSwap Volumes (Availability)**, as shown in Figure 9-18.



There's no automatic failover on the Volume

Solve the issues from top to bottom

- Secondary Failover
- Role Conflict
  Click here to change the roles of the Volumes
- The HyperSwap state is inactive
  Click here to activate HyperSwap

PRIMARY I/O Blocked

System

A9000

Volume

ITSO_Cluster_Data

Availability Status

Configuration Error

PRIMARY SELECTED

System

A9000R

Volume

ITSO_Cluster_Data

Availability Status

Automatic failover was perform...

Figure 9-18 HyperSwap volume after A9000 up again

The necessary steps for the recovery after a failure are described in 5.4, “Failback scenarios” on page 86.

10

Volume Shadow Copy Service and HyperSwap Snapshots

This chapter discusses how to combine the use of Windows Volume Shadow Copy Service (VSS) with FlashSystem A9000 and A9000R snapshots in a HyperSwap configuration.

Microsoft first introduced VSS in Windows 2003 Server and has included it in all subsequent releases. VSS provides a framework and the mechanisms to create consistent point-in-time copies (known as *shadow copies*) of databases and application data. It consists of a set of Microsoft COM APIs that enable volume-level snapshots to be performed while the applications that contain data on those volumes remain online and continue to write. This approach enables third-party software like IBM Spectrum Protect™ Snapshot or IBM Spectrum Copy Data Management to centrally manage the backup and restore operations.

For more information about VSS, see the Microsoft TechNet website.

For more information about Spectrum Protect Snapshot, see the IBM Spectrum Project website.

For more information about Spectrum Copy Data Manager, see *Using IBM Spectrum Copy Data Management with IBM FlashSystem A9000 or A9000R and SAP HANA*, REDP-5439.

This chapter covers the following topics:

10.1 VSS product and components

Without VSS, if you do not have an online backup solution implemented, you either must stop or quiesce applications during the backup process or live with the side effects of an online backup with only crash-consistent data and open files that cannot be backed up.

With VSS, you can produce application-consistent shadow copies by coordinating tasks with business applications, file system services, backup applications, fast recovery solutions, and storage hardware such as the FlashSystem A9000 or A9000R.

VSS enables you to perform online backup of applications, which otherwise is not possible, and is supported on the FlashSystem A9000 and A9000R. VSS accomplishes this backup by facilitating communications between the following entities:

- Requestor

  An application that requests that a volume shadow copy be taken. These are applications, such as backup (like Spectrum Protect Snapshot or Spectrum Copy Data Management) or storage management, that request a point-in-time copy of data or a shadow copy.

- Writer

  A component of an application that stores persistent information about one or more volumes that participate in shadow copy synchronization. Writers are software that is included in applications and services to help provide consistent shadow copies. Writers serve the following main purposes:

  – Responding to signals provided by VSS to interface with applications to prepare for shadow copy.

  – Providing information about the application name, icons, files, and a strategy to restore the files.

  Writers prevent data inconsistencies. A database application (such as SQL Server or Exchange Server) or a system service (such as Active Directory) can be a writer.

- Provider

  A component that creates and maintains the shadow copies. A provider can be implemented in the software or in the hardware. For FlashSystem A9000 and A9000R, you must install and configure the *IBM Spectrum Accelerate Family Provider for Microsoft Windows Volume Shadow Copy Service*.

Figure 10-1 shows the Microsoft VSS architecture and how the software provider and hardware provider interact through Volume Shadow Copy Services.

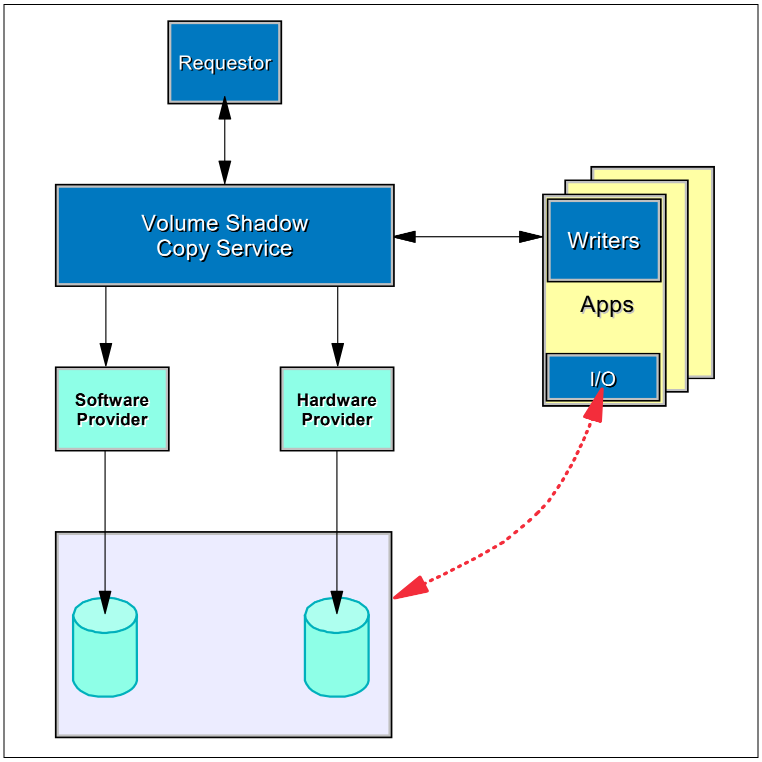


Figure 10-1 Microsoft VSS architecture

VSS uses the following terminology to characterize the nature of volumes participating in a shadow copy operation:

- Persistent

  A shadow copy that remains after the backup application completes its operations. This type of shadow copy also survives system reboots.

- Non-persistent

  A temporary shadow copy that remains only while the backup application needs it to copy the data to its backup repository.

- Transportable

  A shadow copy volume that is accessible from a secondary host so that the backup can be off-loaded. Transportable is a feature of hardware snapshot providers. You can mount a snapshot volume to another host.

- Source volume

  The volume that contains the data to be shadow copied. These volumes contain the application data.

- Target or snapshot volume

  The volume that retains the shadow-copied storage files. It is an exact copy of the source volume at the time of backup.

VSS supports the following shadow copy methods:

- Clone (full copy/split mirror)

  A clone is a shadow copy volume that is a full copy of the original data as it resides on a volume. The source volume continues to take application changes while the shadow copy volume remains an exact read-only copy of the original data at the point-in-time that it was created.

- Copy-on-write (differential copy)

  A copy-on-write shadow copy volume is a differential copy (rather than a full copy) of the original data as it resides on a volume. This method makes a copy of the original data before it is overwritten with new changes. Using the modified blocks and the unchanged blocks in the original volume, a shadow copy can be logically constructed that represents the point-in-time at which it was created.

- Redirect-on-write (differential copy)

  A redirect-on-write shadow copy volume is a differential copy (rather than a full copy) of the original data as it resides on a volume. This method is similar to copy-on-write, without the double-write penalty, and it offers storage-space- and performance-efficient snapshots. New writes to the original volume are redirected to another location set aside for the snapshot. The advantage of redirecting the write is that only one write takes place, whereas with copy-on-write, two writes occur (one to copy original data onto the storage space, the other to copy changed data). The XIV storage system supports redirect-on-write.

10.2 VSS function

VSS accomplishes the fast backup process when a backup application initiates a shadow copy backup. VSS coordinates with the VSS-aware writers to briefly hold writes on the databases, applications, or both. VSS flushes the file system buffers and requests a provider to initiate an IBM FlashCopy® of the data. When the FlashCopy is logically completed, VSS allows writes to resume and notifies the requestor that the backup has completed successfully. The volumes are mounted, hidden, and for read-only purposes, to be used when rapid restore is necessary. Alternatively, the volumes can be mounted on a different host and used for application testing or backup to tape.

The VSS FlashCopy process includes the following steps:

1. The requestor notifies VSS to prepare for shadow copy creation.
2. VSS notifies the application-specific writer to prepare its data for making a shadow copy.
3. The writer prepares the data for that application by completing all open transactions, flushing the cache, and writing in-memory data to disk.
4. When the data is prepared for shadow copy, the writer notifies the VSS, and it relays the message to the requestor to initiate the commit copy phase.
5. VSS temporarily quiesces application I/O write requests for a few seconds and the hardware provider performs the FlashCopy on the Storage Unit.
6. After the completion of FlashCopy, VSS releases the quiesce, and database writes resume.
7. VSS queries the writers to confirm that write I/Os were successfully held during Volume Shadow Copy.

10.3 IBM Spectrum Accelerate family VSS Provider (xProv)

A VSS hardware provider, such as the xProv, is used by third-party software to act as an interface between the hardware (storage system) and the operating system. The third-party application uses Spectrum Accelerate family VSS Provider to instruct FlashSystem A9000 or A9000R to perform a snapshot of a volume attached to the host system.

> **Note:** The following illustrations are based on Windows 2012 R2, Hyper-Scale Manager 5.2, and FlashSystem A9000 and A9000R Version 12.1.

10.3.1 Install the VSS Provider

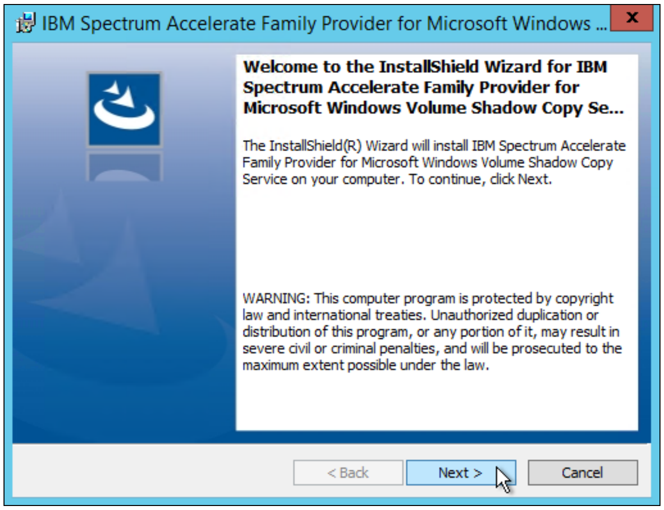
This section illustrates how to install the VSS Provider. At the time of writing, the IBM Spectrum Accelerate family VSS Provider Version 2.9.0 was available. To obtain the system requirements, see the *IBM Spectrum Accelerate Family Provider for Microsoft Windows Volume Shadow Copy Service Release Notes*, which includes a chapter about the system requirements. Download the Spectrum Accelerate family VSS Provider user guide and release notes from IBM Fix Central.

Verify that the prerequisites are met and the following packages are installed:

- Microsoft Visual C++ 2012 Redistributable-64
- Microsoft Visual C++ 2012 Redistributable-86

Installing the VSS Provider is a straightforward Windows application installation:

1. Locate the Spectrum Accelerate family VSS Provider installation file, also known as the *xProv* installation file. If the VSS Provider is downloaded from the Internet, the file name is `xProvSetup_2.9.0.0-b64_for_Windows-x64.exe`. Execute the file to start the installation.
2. Select the language for the installation, the Welcome window opens, as shown in Figure 10-2. Click **Next**.



IBM Spectrum Accelerate Family Provider for Microsoft Windows ...

Welcome to the InstallShield Wizard for IBM Spectrum Accelerate Family Provider for Microsoft Windows Volume Shadow Copy Se...

The InstallShield(R) Wizard will install IBM Spectrum Accelerate Family Provider for Microsoft Windows Volume Shadow Copy Service on your computer. To continue, click Next.

WARNING: This computer program is protected by copyright law and international treaties. Unauthorized duplication or distribution of this program, or any portion of it, may result in severe civil or criminal penalties, and will be prosecuted to the maximum extent possible under the law.

Figure 10-2 VSS provider installation: Welcome window

3. The License Agreement window also displays. To continue the installation you must accept the license agreement.

4. Specify the VSS Provider configuration file directory and the installation directory. Keep the default directory folder and installation folder or change it to meet your needs.
5. A dialog window for post-installation operations is opened, as shown in Figure 10-3. You can perform a post-installation configuration during the installation process or at a later time. When done, click **Next**.

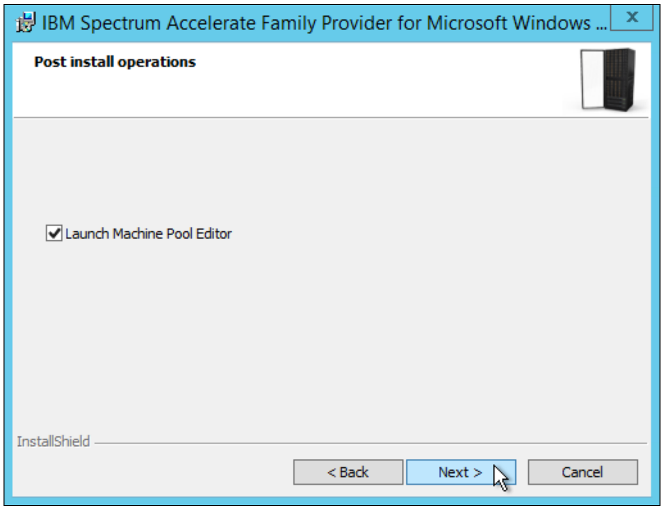

IBM Spectrum Accelerate Family Provider for Microsoft Windows ...
Post install operations
☑ Launch Machine Pool Editor
InstallShield
< Back
Next >
Cancel

Figure 10-3 Installation: Post-installation operation

6. A Confirm Installation window is displayed. You can go back to make changes if required, or confirm the installation by clicking **Install**.
7. Click **Close** to exit after the installation is complete.

10.3.2 Configure VSS Provider

Complete the following steps to configure the VSS Provider:

1. If the post installation option was selected during installation (Figure 10-3), the MachinePoolEditor window opens automatically.
2. Right-click the Machine Pool Editor window.
3. In the dialog that is shown in Figure 10-4, click **New System** to open the New System window.

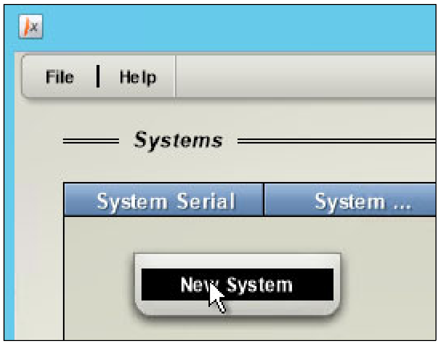
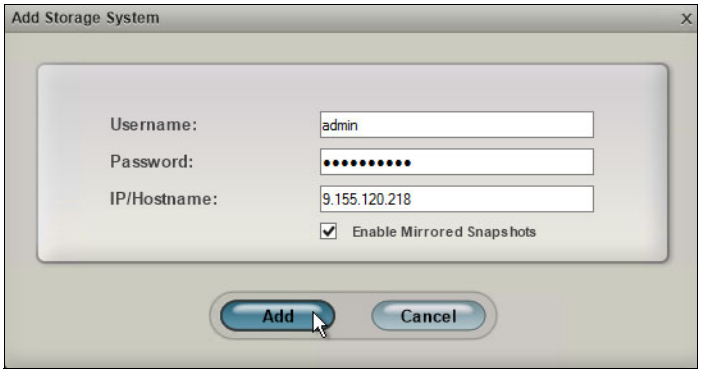


Figure 10-4 Configuration: Machine Pool Editor

4. The Add System Management window shown in Figure 10-5 displays. Enter the user name and password of a FlashSystem A9000 or A9000R user with administrator privileges (*storageadmin* role) and the primary IP address of the FlashSystem A9000 or A9000R. If the snapshot is taken of a volume that is in a mirror relationship and you want to have the snapshot on source and target system, then select **Enable Replicated Snapshots** and click **Add**.



Username: admin

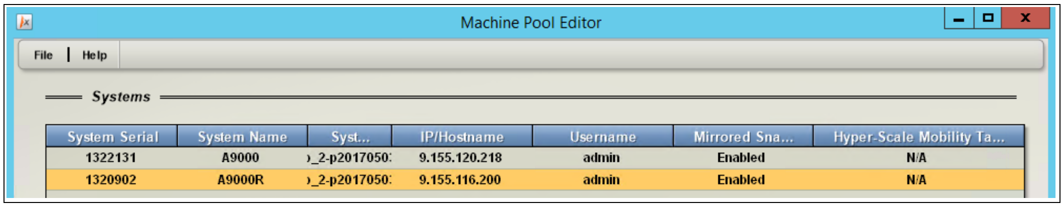
Password: ••••••••••

IP/Hostname: 9.155.120.218

☑ Enable Mirrored Snapshots

Figure 10-5 Configuration: Add Storage System

5. Repeat steps 3 and 4 to add additional FlashSystem A9000 or A9000R.
6. You are returned to the VSS Machine Pool Editor window. The VSS Provider collected additional information about the storage systems, as illustrated in Figure 10-6.



System Serial	System Name	Syst...	IP/Hostname	Username	Mirrored Sna...	Hyper-Scale Mobility Ta...
1322131	A9000	)_2-p2017050:	9.155.120.218	admin	Enabled	N/A
1320902	A9000R	)_2-p2017050:	9.155.116.200	admin	Enabled	N/A

Figure 10-6 Configuration: Machine Pool Editor

7. If the snapshot needs to be taken of a volume that is in a HyperSwap relationship and the snapshots are needed on both sites, open `C:\Program Files\IBM\IBM XIV Provider` for Microsoft Windows Volume Shadow Copy Service\etc\MachinePool.xml and set `create_ha_snapshots="true"`, as shown in Example 10-1.

Example 10-1 MachinePool.xml file

```
<?xml version="1.0" encoding="us-ascii"?>
<MachinePool><machine serial="1322131" username="admin"
create_mirror_snapshots="false" create_ha_snapshots="true"
password="LZnLBheVIybKTxpbCkMeDQ==">
  <aserver hostname="9.155.120.218" port="7778" />
  <aserver hostname="10.0.20.108" port="7778" />
  <aserver hostname="10.0.20.109" port="7778" />
</machine><machine serial="1320902" username="admin"
create_mirror_snapshots="false" create_ha_snapshots="true"
password="LZnLBheVIybKTxpbCkMeDQ==">
  <aserver hostname="9.155.116.200" port="7778" />
  <aserver hostname="9.155.116.201" port="7778" />
  <aserver hostname="9.155.116.202" port="7778" />
</machine></MachinePool>
```

8. At this point, the VSS Provider configuration is complete and you can close the Machine Pool Editor window.

After the VSS Provider has been configured, ensure that the operating system can recognize it. To do so, launch the **`vssadmin list providers`** command from the operating system command line.

Make sure that `IBM VSS HW Provider` is in the list of installed VSS providers returned by the **`vssadmin`** command, as shown in Example 10-2.

Example 10-2 Output of vssadmin command

```
PS C:\Users\Administrator.ITSO> vssadmin list providers
vssadmin 1.1 - Volume Shadow Copy Service administrative command-line tool
(C) Copyright 2001-2013 Microsoft Corp.

Provider name: 'Microsoft CSV Shadow Copy Helper Provider'
   Provider type: Software
   Provider Id: {26d02d81-6aac-4275-8504-b9c6edc5261d}
   Version: 1.0.0.1

Provider name: 'Microsoft CSV Shadow Copy Provider'
   Provider type: Software
   Provider Id: {400a2ff4-5eb1-44b0-8a05-1fcac0bcf9ff}
   Version: 1.0.0.1

Provider name: 'Microsoft File Share Shadow Copy provider'
   Provider type: Fileshare
   Provider Id: {89300202-3cec-4981-9171-19f59559e0f2}
   Version: 1.0.0.1

Provider name: 'Microsoft Software Shadow Copy provider 1.0'
   Provider type: System
   Provider Id: {b5946137-7b9f-4925-af80-51abd60b20d5}
   Version: 1.0.0.7

Provider name: 'IBM XIV VSS HW Provider'
   Provider type: Hardware
   Provider Id: {d51fe294-36c3-4ead-b837-1a6783844b1d}
   Version: 2.9.0.0
```

10.3.3 Diskshadow command line utility

All editions of Windows Server 2012 R2 contain a command line utility (`DiskShadow.exe`) for the creation, deletion, and restoration of shadow copies (snapshots). It is the first in-box VSS requestor that can create hardware shadow copies and one of many utilities for validating VSS operations. The tool is similar to vshadow (a tool included with the Volume Shadow Copy/VSS SDK), but has an interface similar to diskpart utility.

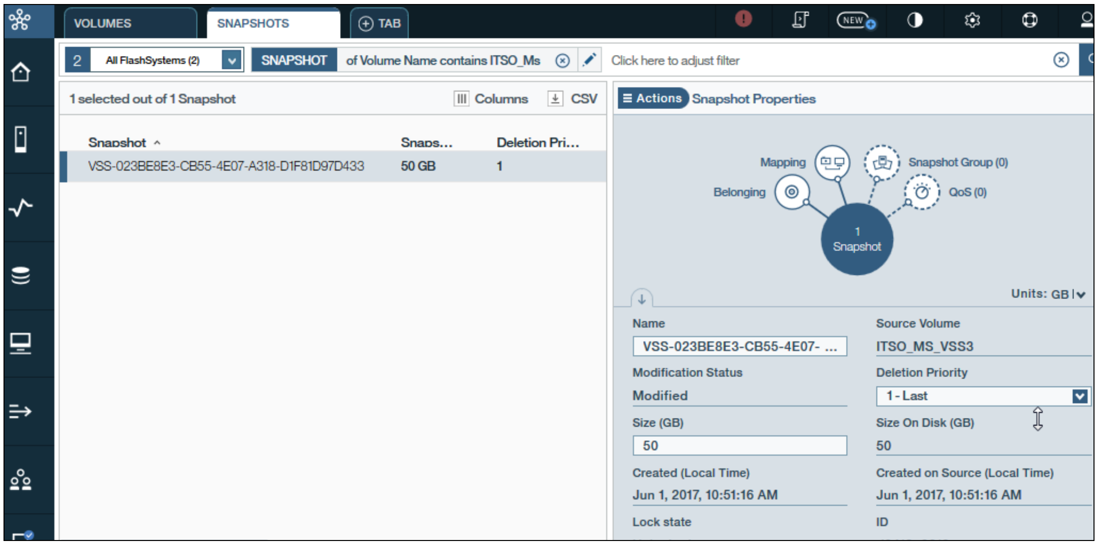
For more information about diskshadow, see the Microsoft TechNet website.

The steps to test the creation of a persistent snapshot of a basic disk on FlashSystem A9000 or A9000R by way of VSS are shown in Example 10-3. The snapshot will be automatically unlocked and mapped to the server. Assign a drive letter to the volume and access the data on the file system.

Example 10-3 Diskshadow snapshot creation

```
PS C:\Users\Administrator.ITSO> diskshadow
Microsoft DiskShadow version 1.0
Copyright (C) 2013 Microsoft Corporation
On computer:  ITSO-WIN1,  6/1/2017 10:49:45 AM
DISKSHADOW> set context persistent
DISKSHADOW> add volume y:
DISKSHADOW> create
COM call "lvssObject4->GetRootAndLogicalPrefixPaths" failed.
Alias VSS_SHADOW_1 for shadow ID {023be8e3-cb55-4e07-a318-d1f81d97d433} set as
environment variable.
Alias VSS_SHADOW_SET for shadow set ID {008cca0e-922a-46dd-8902-e764e3ca3d17} set
as environment variable.
Querying all shadow copies with the shadow copy set ID
{008cca0e-922a-46dd-8902-e764e3ca3d17}

        * Shadow copy ID = {023be8e3-cb55-4e07-a318-d1f81d97d433}
%VSS_SHADOW_1%
                - Shadow copy set: {008cca0e-922a-46dd-8902-e764e3ca3d17}
%VSS_SHADOW_SET%
                - Original count of shadow copies = 1
                - Original volume name:
\\?\Volume{1dc9da0e-45dd-11e7-80cb-6eae8b4b61ab}\ [Y:\]
                - Creation time: 6/1/2017 10:50:48 AM
                - Shadow copy device name:
\\?\Volume{1dc9da1c-45dd-11e7-80cb-6eae8b4b61ab}
                - Originating machine: ITSO-WIN1.ITSO.local
                - Service machine: ITSO-WIN1.ITSO.local
                - Not exposed
                - Provider ID: {d51fe294-36c3-4ead-b837-1a6783844b1d}
                - Attributes:  No_Auto_Release Persistent Hardware
Number of shadow copies listed: 1
```

Select **POOLS & VOLUMES VIEWS** → **Snapshots** and filter for the snapshot, the snapshot with his Shadow Copy ID is visible as shown in Figure 10-7.



1 selected out of 1 Snapshot

Snapshot	Snaps...	Deletion Pri...
VSS-023BE8E3-CB55-4E07-A318-D1F81D97D433	50 GB	1

Name: VSS-023BE8E3-CB55-4E07- ...
Source Volume: ITSO_MS_VSS3
Modification Status: Modified
Deletion Priority: 1 - Last
Size (GB): 50
Size On Disk (GB): 50
Created (Local Time): Jun 1, 2017, 10:51:16 AM
Created on Source (Local Time): Jun 1, 2017, 10:51:16 AM
Lock state
ID

Figure 10-7 VSS snapshot

10.3.4 Create HyperSwap VSS snapshot

Starting with xProv Version 2.9, you can create snapshots through VSS on a HyperSwap volume. Before you begin, the HyperSwap relationship must exist, be active and must be enabled in the `MachinePool.xml` file.

The VSS process will create a snapshot of the source volume and also the target volume. Complete the following steps to create the snapshot on both mirror sites:

1. Add both FlashSystem A9000 and A9000R to xProv and set `create_ha_snapshots="true"`.
2. Set up HyperSwap for the volumes you want to use for VSS snapshot replication.
3. Run a VSS create operation on the HyperSwap volume.

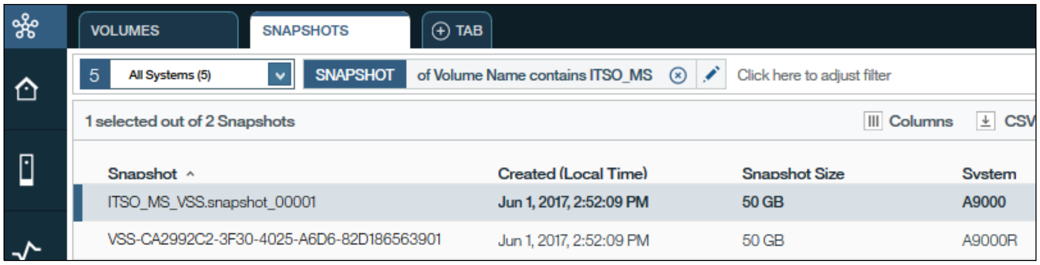
To be able to import the shadow copy to the same or a different computer the `transportable` option must be used and the `example1.cab` is needed, as shown in Example 10-4. The import works on basic disks only.

Example 10-4 VSS snapshot creation of a HyperSwap volume

```
PS C:\Users\Administrator.ITSO> diskshadow
Microsoft DiskShadow version 1.0
Copyright (C) 2013 Microsoft Corporation
On computer:  ITSO-WIN1,  6/1/2017 2:48:31 PM


DISKSHADOW> set context persistent

DISKSHADOW> set option transportable

DISKSHADOW> set metadata c:\Users\Administrator\example1.cab

DISKSHADOW> add volume z:

DISKSHADOW> create
```

```
COM call "lvssObject4->GetRootAndLogicalPrefixPaths" failed.
Alias VSS_SHADOW_1 for shadow ID {ca2992c2-3f30-4025-a6d6-82d186563901} set as environment variable.
Alias VSS_SHADOW_SET for shadow set ID {e7870a5b-9b2f-4c67-a1ed-7f7e589f7c83} set as environment variable.
```

The snapshots created by VSS on the source and target XIV storage system are depicted in Figure 10-8.



1 selected out of 2 Snapshots

Snapshot	Created (Local Time)	Snapshot Size	System
ITSO_MS_VSS.snapshot_00001	Jun 1, 2017, 2:52:09 PM	50 GB	A9000
VSS-CA2992C2-3F30-4025-A6D6-82D186563901	Jun 1, 2017, 2:52:09 PM	50 GB	A9000R

Figure 10-8 VSS mirrored snapshot on source and target

To test the import of the data afterwards to another server, copy the `example1.cab` file to that server. The host and its ports must be defined on the storage systems.

The commands to load the metadata and import the VSS snapshot to the server are shown in Example 10-5. Afterwards assign a drive letter to the volume and access the data on the file system.

Example 10-5 VSS import to another server

```
PS C:\Users\Administrator.ITSO> diskshadow
Microsoft DiskShadow version 1.0
Copyright (C) 2013 Microsoft Corporation
On computer:  ITSO-WIN2,  6/1/2017 3:31:02 PM

DISKSHADOW> load metadata example1.cab
Could not find metadata file.

DISKSHADOW> load metadata C:\Users\Administrator\example1.cab
Alias VSS_SHADOW_1 for value {ca2992c2-3f30-4025-a6d6-82d186563901} set as an
environment variable.
Alias VSS_SHADOW_SET for value {e7870a5b-9b2f-4c67-a1ed-7f7e589f7c83} set as an
environment variable.

DISKSHADOW> import

DISKSHADOW>
```

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 paper.

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:

- *IBM FlashSystem A9000 and IBM FlashSystem A9000R Architecture and Implementation*, SG24-8345
- *IBM Hyper-Scale Manager for IBM Spectrum Accelerate Family: IBM XIV, IBM FlashSystem A9000 and A9000R, and IBM Spectrum Accelerate, SG24-8376*
- *IBM FlashSystem A9000 and A9000R Business Continuity Solutions*, REDP-5401
- *IBM FlashSystem A9000, IBM FlashSystem A9000R, and IBM XIV Storage System: Host Attachment and Interoperability*, SG24-8368
- *Using the Using the IBM Spectrum Accelerate Family in VMware Environments: IBM XIV, IBM FlashSystem A9000 and IBM FlashSystem A9000R, and IBM Spectrum Accelerate*, REDP-5425

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

Other publications

The following publications are also relevant as further information sources:

- *IBM Spectrum Accelerate Family HyperSwap Quorum Witness User Guide*, SC27-4631
- IBM FlashSystem A9000 and A9000R product overview
- IBM FlashSystem A9000 and A9000R CLI reference guide
- IBM Hyper-Scale Manager user guide

Online resources

The following websites are also relevant as further information sources:

- IBM Fix Central:

  https://www.ibm.com/support/fixcentral

- IBM Knowledge Center:

  https://www.ibm.com/support/knowledgecenter/

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

REDP-5434-02

ISBN 0738457442

Printed in U.S.A.