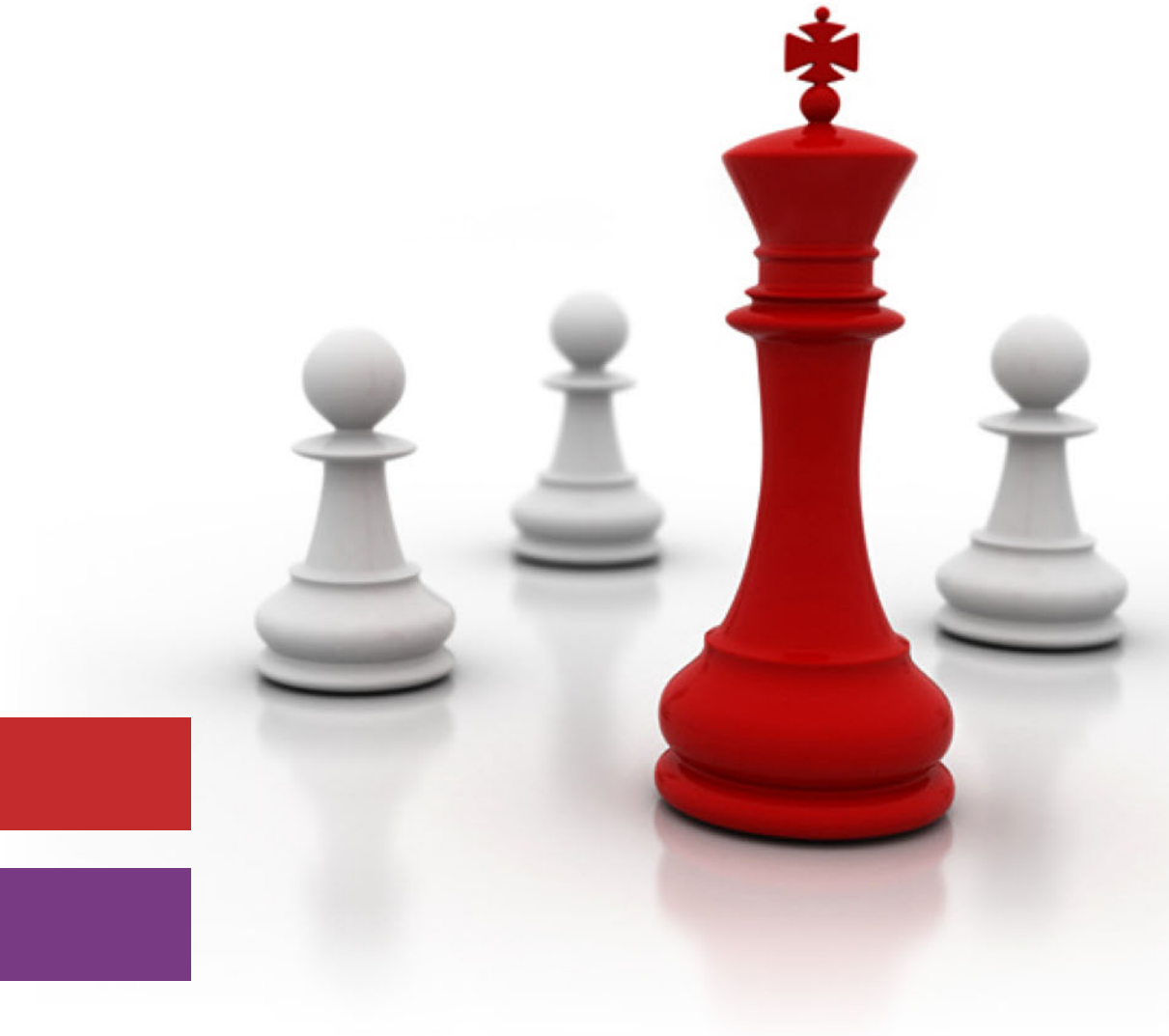
Redbooks
ibm.com/redbooks

IBM Spectrum Virtualize for Public Cloud on AWS Version 8.3.1 Implementation Guide

Jordan Fincher
Hemanand Gadgil
Vasfi Gucer
Jackson Shea



Cloud

Storage

IBM

Redpaper

IBM Redbooks

IBM Spectrum Virtualize for Public Cloud on AWS Version 8.3.1 Implementation Guide

July 2020

REDP-5588-00

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

First Edition (July 2020)

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®	HyperSwap®	Passport Advantage®
Aspera®	IBM®	Redbooks®
DB2®	IBM Cloud®	Redbooks (logo) ®
Db2®	IBM FlashSystem®	Storwize®
Easy Tier®	IBM Spectrum®	
FlashCopy®	IBM Spectrum Storage™	

The following terms are trademarks of other companies:

Intel, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

ITIL is a Registered Trade Mark of AXELOS Limited.

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

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

Java, and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

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

VMware, and the VMware logo are registered trademarks or trademarks of VMware, Inc. or its subsidiaries in the United States and/or other jurisdictions.

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

Preface

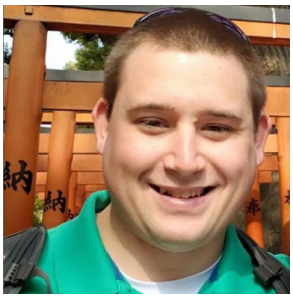
IBM® Spectrum Virtualize is a key member of the IBM Spectrum® Storage portfolio. It is a highly flexible storage solution that enables rapid deployment of block storage services for new and traditional workloads, whether on-premises, off-premises, or a combination of both.

The initial release of IBM Spectrum Virtualize for Public Cloud is now available on Amazon Web Services (AWS). This IBM Redpaper™ publication gives a broad understanding of the IBM Spectrum Virtualize for Public Cloud on AWS architecture. It also provides planning and implementation information about the common use cases for this new product.

This publication helps storage and networking administrators plan, implement, install, modify, and configure the IBM Spectrum Virtualize for Public Cloud on AWS offering Version 8.3.1. It also provides a detailed description of troubleshooting tips.

Authors

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



Jordan Fincher is a Product Field Engineer working in SVC and FlashSystems Level 3 Support at IBM. He received his Bachelor of Science degree in Information Security from Western Governor's University. Jordan started his IBM career in 2012 as a systems engineer for the IBM Business Partner e-TechServices, doing pre-sales consulting and implementation work for many IBM accounts in Florida. In 2015, Jordan started working in his current role as a Product Field Engineer for IBM Spectrum Virtualize storage products.

Hemanand Gadgil is an IBM Storage Solutions Architect working with various independent software vendors (ISVs) partners for storage solutions designs in Pune, India. He received his Bachelor of Engineering degree in Electronics from the University of Pune, India. His current interests are in cloud, hybrid multicloud, disaster recovery (DR), and business continuity solutions. Hemanand joined IBM in 2015 as a Storage Solutions Architect. His skills include various storage technologies, data center migration, and infrastructure and data center consulting.

Vasfi Gucer is a project leader with the IBM Systems WW Client Experience Center. He has more than 20 years of experience in the areas of systems management, networking hardware, and software. He writes extensively and teaches IBM classes worldwide about IBM products. His focus has been primarily on storage and cloud computing for the last 8 years. Vasfi is also an IBM Certified Senior IT Specialist, Project Management Professional (PMP), IT Infrastructure Library (ITIL) V2 Manager, and ITIL V3 Expert.

Jackson Shea is a Level 2 certified IBM Information Technology Specialist/Architect who performing design and implementation engagements through Lab Services. He has been with IBM since April 2010. He was a Lead Storage Administrator with a large health insurance consortium in the Pacific Northwest, and has been working with IBM equipment since 2002. He has over 12 years of experience with IBM Spectrum Virtualize (formerly known as the IBM SAN Volume Controller) and related technologies. Jackson is based out of Portland, Oregon. He received his Bachelor of Science degree in Philosophy with minors in Communications and Chemistry from Lewis & Clark College. Jackson's professional focus is IBM Spectrum Virtualize, but he is conversant with storage area network design, implementation, extension, and storage encryption.

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

Eva L Ho, Antonio Y Pacheco, Zalvay Perkins, Matt Puccini, Cliff Ray, Michelle Tidwell, Erica Wazewski, Derrod Williams
IBM US

Jon Tate
IBM UK

Long Wen Lan and Li Li Yang
IBM China

Thanks to the following authors of the previous version of this Redpaper:

Jimmy John, Kendall C Williams, Thomas Vogel

Now you can become a published author, too!

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

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

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our 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, IBM Redbooks
  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

Introduction

This chapter describes the IBM Spectrum Virtualize product that is implemented in a cloud environment, which is referred to as *IBM Spectrum Virtualize for Public Cloud*.

This chapter also provides a brief overview of the technology that is behind the product introduces the drivers and business values of the use of IBM Spectrum Virtualize in the context of public cloud services. Finally, it describes how the solution works from a high-level perspective.

This publication describes IBM Spectrum Virtualize for Public Cloud V8.3.1

This chapter includes the following topics:

1.1 Introduction to IBM Spectrum Virtualize for Public Cloud

Companies are undergoing a digital transformation and making architecture decisions that determine how their businesses are going to operate in the next couple of years. They recognize the value of delivering services by using the cloud, and many of them are already using public clouds to some degree. The role of the cloud is maturing and it is more often being considered as a platform for innovation and business value. The cloud is a key enabler to drive transformation and innovation for IT agility and new capabilities.

Nevertheless, one of the challenges for these organizations is how to integrate those public cloud capabilities with the existing backend. Organizations want to retain flexibility without introducing complexity or requiring significant new capital investment.

Cloud integration can occur between different endpoints (cloud-to-cloud, on-premises to off-premises, or cloud to non-cloud) and at different levels within the cloud stack: infrastructure layer, service layer and for example, at the application layer or at the management one. Within the infrastructure as a service (IaaS) domain, storage layer integration is often the most attractive approach for ease of migration and replication of heterogeneous resources and data consistency.

In this sense, coming from the IBM Spectrum Storage™ family, IBM Spectrum Virtualize for Public Cloud supports clients in their IT architectural transformation and migration towards the cloud service model. It enables hybrid cloud strategies or for a cloud-native workload, provides the benefits of familiar and sophisticated storage functions on public cloud data centers, which enhances the existing cloud offering.

Running on-premises, IBM Spectrum Virtualize software supports capacity that is built into storage systems, and capacity in over 400 different storage systems from IBM and other vendors. This wide range of storage support means that the solution can be used with almost any storage in a data center today and integrated with its counterpart IBM Spectrum Virtualize for Public Cloud, which supports Amazon Web Services (AWS) Elastic Block Store (EBS) and its various options. For more information, see Chapter 3, “Solution architecture” on page 27, and Chapter 4, “Planning and preparing for the IBM Spectrum Virtualize for Public Cloud on AWS deployment” on page 41.

IBM Storwize® rebranding: On February 11, 2020, IBM rebranded IBM Storwize storage systems as IBM FlashSystem®; for example, IBM Storwize V5030 is now called IBM FlashSystem 5030. Although this publication is updated to use the new terminology, the “Storwize” name might still be used in some of the hotlinks.

1.2 IBM Spectrum Virtualize for Public Cloud

Designed for SDS environments, IBM Spectrum Virtualize for Public Cloud represents a solution for public cloud implementations, and includes technologies that complement and enhance public cloud offering capabilities.

For example, traditional practices that provide data replication by copying storage at one facility to largely identical storage at another facility are not an option for public cloud. Also, the use of conventional software to replicate data imposes unnecessary loads on application servers. Use cases are analyzed in Chapter 2, “Typical use cases for IBM Spectrum Virtualize for Public Cloud” on page 11.

IBM Spectrum Virtualize for Public Cloud delivers a powerful solution for the deployment of IBM Spectrum Virtualize software in public clouds. This new capability provides a monthly license to deploy and use IBM Spectrum Virtualize for Public Cloud on AWS to enable hybrid cloud solutions, which offer the ability to transfer data between on-premises data centers by using any IBM Spectrum Virtualize-based appliance and multiple cloud environments.

With a deployment that is designed for the cloud, IBM Spectrum Virtualize for Public Cloud can be deployed in cloud data centers around the world where, after provisioning the infrastructure, an installation script automatically installs the software.

1.2.1 Primers of storage virtualization and software-defined storage

The term *virtualization* is used widely in IT and applied to many of the associated technologies. Its usage in storage products and solutions is no exception. IBM defines *storage virtualization* as a technology that makes one set of resources resemble another set of resources, preferably with more wanted characteristics.

It is a logical representation of resources that is not constrained by physical limitations and hides part of the complexity. It also adds or integrates new functions with existing services and can be nested or applied to multiple layers of a system.

The aggregation of volumes into storage pools enables you to better manage capacity, performance, and multiple tiers for the workloads. IBM Spectrum Virtualize for Public Cloud provides virtualization only at the disk layer (block-based) of the I/O stack, and for this reason it is referred to as *block-level virtualization*, or the block aggregation layer. For the sake of clarity, the block-level volumes that are provided by the cloud are exposed as target volumes, and are seen by IBM Spectrum Virtualize as a managed disk (MDisk).

These MDisks are then aggregated into a storage pool, which is sometimes referred to as a *managed disk group* (mdiskgrp). IBM Spectrum Virtualize then creates logical volumes (referred to as *volumes* or *VDisks*) that are striped across all of the MDisks inside of their assigned pool.

The virtualization terminology is included into the wider concept of SDS, which is an approach to data storage in which the programming that controls storage-related tasks is decoupled from the physical storage hardware. This separation allows SDS solutions to be placed over any storage systems or more generally, installed on any commodity x86 hardware and hypervisor.

Shifting to a higher level in the IT stack allows for a deeper integration and response to application requirements for storage performance and capabilities. SDS solutions offer a full suite of storage services (equivalent to traditional hardware systems) and federation of multiple persistent storage resources: internal disk, cloud, other external storage systems, or cloud and object platforms.

In general, SDS technology uses the following concepts:

- A shared-nothing architecture (or in some cases a partial or fully shared architecture) with no single point of failure and nondisruptive upgrades.
- Scale-up or scale-out mode: Add building blocks for a predictable increase in capacity, performance, and resiliency.
- Multiple classes of service: File-based, object-based, block-based, and auxiliary and storage support service. SDS solutions also can be integrated into a hybrid or composite SDS solution.

- High availability (HA) and disaster recovery (DR): Can tolerate levels of availability and durability as self-healing and adjusting.
- Lower total cost of ownership (TCO): Lower the TCO for those workloads that can use SDS.

1.2.2 IBM Spectrum Virtualize for Public Cloud benefits

IBM Spectrum Virtualize for Public Cloud offers a powerful value proposition for enterprise and cloud users who are searching for more flexible and agile ways to deploy block storage on cloud. By using standard Intel servers, IBM Spectrum Virtualize for Public Cloud can be easily added to cloud infrastructures to deliver more features and functions, which enhance the storage offering that is available on the public cloud catalog.

The benefits of deploying IBM Spectrum Virtualize for Public Cloud are two-fold:

- Public cloud storage offering enhancement: IBM Spectrum Virtualize for Public Cloud enhances the public cloud catalog by increasing standard storage, and offering capabilities and features that decrease specific limitations:
  - Snapshots: A volume's snapshots occur on high-tier storage with no options for a lower-end storage tier. By using IBM Spectrum Virtualize, the administrator has more granular control, which enables a production volume to have a snapshot that is stored on lower-end storage.
  - Volume size: Most cloud storage providers have a maximum volume size (typically a few terabytes) that can be provided by a few nodes. At the time of this writing, IBM Spectrum Virtualize allows for up to 320 TB and up to 20,000 host connections.
  - Native storage-based replication: Replication features are natively supported, but are typically limited to specific data center pairs and a predefined minimum recovery point objective (RPO). They are accessible only when the primary volume is down. IBM Spectrum Virtualize provides greater flexibility in storage replication to allow for user-defined RPO and replication between any other system that is running IBM Spectrum Virtualize.
- New features for public cloud storage offering: IBM Spectrum Virtualize for Public Cloud introduces to the public cloud catalog new storage capabilities. Those features are available on SAN Volume Controller and IBM Spectrum Virtualize, but are not available by default. The following extra features that are provided on public cloud are related to hybrid cloud scenarios and its support to foster all those solutions for improved hybrid architectures:
  - Replication or migration of data between on-premises storage and public cloud storage

    In a heterogeneous environment, replication consistency is achieved through storage-based replica peer cloud storage with primary storage on-premises. Because of standardization of the storage service model and inability to move its own storage to a cloud data center, the storage-based replica is achievable only by involving an SDS solution on-premises.

    In this sense, IBM Spectrum Virtualize for Public Cloud offers data replication between the FlashSystem family, SAN Volume Controller, or VersaStack and Public Cloud and extends replication to all types of supported virtualized storage on-premises. Working together, IBM Spectrum Virtualize and IBM Spectrum Virtualize for Public Cloud support synchronous and asynchronous mirroring between the cloud and on-premises for more than 400 different storage systems from various vendors. In addition, they support other services, such as IBM FlashCopy® and IBM Easy Tier®.

– DR strategies between on-premises and public cloud data centers as alternative DR solutions

  One of the reasons to replicate is to have a copy of the data from which to restart operations in case of an emergency. IBM Spectrum Virtualize for Public Cloud enables DR for virtual and physical environments, which adds new possibilities compared to the software replicators in use today that handle virtual infrastructure only.

– Benefit from familiar, sophisticated storage functions in the cloud to implement reverse mirroring

  IBM Spectrum Virtualize enables the possibility to reverse data replication to offload from a cloud provider back to on-premises or to another cloud provider.

On-premises and on-cloud IBM Spectrum Virtualize provides a data strategy that is independent of the choice of infrastructure, which delivers tightly integrated functions and consistent management across heterogeneous storage and cloud storage. The software layer that is provided by IBM Spectrum Virtualize on-premises or in the cloud can provide a significant business advantage by delivering more services faster and more efficiently, which enables real-time business insights and supports more customer interaction.

Capabilities, such as rapid, flexible provisioning; simplified configuration changes; nondisruptive movement of data among tiers of storage; and a single user interface helps make the storage infrastructure (and the hybrid cloud) simpler, more cost-effective, and easier to manage.

1.2.3 IBM Spectrum Virtualize for Public Cloud features

IBM Spectrum Virtualize for Public Cloud helps make cloud storage volumes (block-level) more effective by including functions that are not natively available on the public cloud catalogs and that are traditionally deployed within disk array systems in the on-premises environment. For this reason, IBM Spectrum Virtualize for Public Cloud improves and expands the capabilities of the cloud offering.

Table 1-1 lists the IBM Spectrum Virtualize for Public Cloud features and benefits.

Table 1-1 IBM Spectrum Virtualize for Public Cloud features and benefits

Feature	Benefits
Single point of control for cloud storage resources.	Designed to increased management efficiency and help support application availability.
Pools the capacity of multiple storage volumes.	► Helps overcome the volume size limitations. ► Helps manage storage as a resource to meet business requirements, and not just as a set of independent volumes. ► Helps an administrator to better deploy storage as required beyond traditional “islands”. ► Can help to increase the use of storage assets. ► Insulate applications from maintenance or changes to a storage volume offering.
Clustered pairs of servers that are configured as IBM Spectrum Virtualize for Public Cloud engines.	► Use of cloud-catalog Intel servers foundation. ► Designed to avoid single point of hardware failures.

Feature	Benefits
Manages tiered storage	► Helps to balance performance needs against infrastructures costs in a tiered storage environment. ► Automated policy-driven control to put data in the right place at the right time automatically among different storage tiers or classes.
Easy-to-use IBM FlashSystem family management interface	► A single interface for storage configuration, management, and service tasks regardless of the configuration that is available from the public cloud portal. ► Helps administrators use storage assets and volumes more efficiently. ► IBM Spectrum Control Insights and IBM Spectrum Protect provide more capabilities to manage capacity and performance.
Dynamic data migration	► Migrates data among volumes or LUNs without taking applications that use that data offline. ► Manages and scales storage capacity without disrupting applications.
Advanced network-based copy services	► Copies data across multiple storage systems with IBM FlashCopy. ► Copy data across metropolitan and global distances as needed to create high-availability storage solutions between multiple data centers.
Thin provisioning, data reduction pools, and snapshot replication	► Reduces volume requirements by using storage only when data changes. ► Improves storage administrator productivity through automated on-demand storage provisioning. ► Supports thin-provisioning in standard and data reduction pools, which reduces capacity requirements by using storage only when data changes. ► Supports data reduction pools with native data reduction features, such as host unmap and reclaiming usable capacity. ► Supports deduplicated and compressed volumes in data reduction pools for more capacity savings. Snapshots are available on lower-tier storage volumes.
IBM Spectrum Protect Snapshot application-aware snapshots	► Performs near-instant application-aware snapshot backups, with minimal performance impact for IBM DB2®, Oracle, SAP, Microsoft SQL Server, and Microsoft Exchange. ► Provides advanced, granular restoration of Microsoft Exchange data.
Native IP replication	► Embedded compress replication traffic for WAN optimization. ► Reduces network costs or speed replication cycles, which improves the accuracy of remote data.
IBM Spectrum Connect Cloud Storage Management	► Manages container storage in Kubernetes.

Note: The following features are *not* supported in the first IBM Spectrum Virtualize for Public Cloud release:

- Stretched cluster
- IBM HyperSwap®
- IBM Real-time Compression
- Encryption
- Cloud backup
- Hot spare node

Some of these features are planned for future releases and will be prioritized for implementation based on customer feedback.

1.3 IBM Spectrum Virtualize for Public Cloud on AWS

The initial release of IBM Spectrum Virtualize for Public Cloud is available on AWS. Block virtualization further uses public cloud infrastructure for various types of workload deployments whether it is new or traditional.

The following features are supported on the AWS infrastructure:

- Data replication with any IBM Spectrum Virtualize product and between public clouds
- FlashCopy snapshots in the cloud
- Common Management: IBM Spectrum Virtualize GUI
- Deployment in any AWS region
- Encryption at rest by using Amazon EBS encrypted volumes
- Data redundancy with volume mirroring
- Automated block-level storage tiering by using Easy Tier
- Scale on demand by thin provisioning volumes and paying for AWS storage as you grow
- Supports thin-provisioning in both standard and data reduction pools, which reduces capacity requirements by using storage only when data changes.
- Supports data reduction pools with native data reduction features, such as host unmap and reclaiming usable capacity.
- Supports deduplicated and compressed volumes in data reduction pools for more capacity savings

The AWS infrastructure is an established platform for today's computing needs. By deploying the IBM Spectrum Virtualize for Public Cloud platform, the features of IBM Spectrum Virtualize further enrich the capabilities of the cloud infrastructure.

Figure 1-1 shows the general layout of IBM Spectrum Virtualize for Public Cloud on AWS.

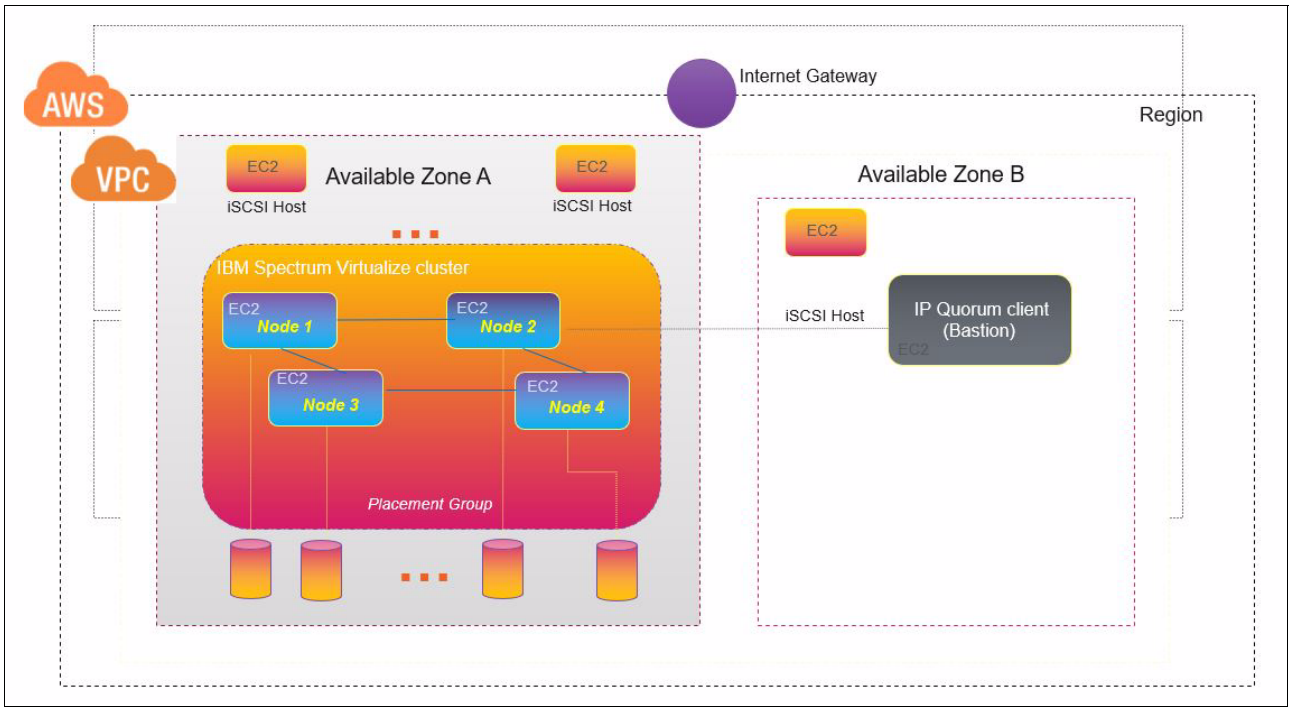


Figure 1-1 High-level architecture of IBM Spectrum Virtualize for Public Cloud on AWS

In AWS, the Amazon EBS storage is directly attached to the IBM Spectrum Virtualize node instances that compose a single node pair (or I/O group) that provides a shared storage pool that is used by IBM Spectrum Virtualize. IBM Spectrum Virtualize supports the following Amazon EBS types:

- General Purpose solid-state drive (SSD) (gp2)
- Provisioned IOPS SSD (io1)
- Throughput Optimized hard disk drive (HDD) (st1)

1.4 What is new in Version 8.3.1 with IBM Spectrum Virtualize for Public Cloud on AWS

Since the initial release of IBM Spectrum Virtualize for Public Cloud on AWS with the 8.3.0 code, much work was done to increase functionality. The following enhancements are available in this version:

- Support for up to four nodes (previously 2) in a cluster on AWS.
- Support for up to 20 Amazon EBS volumes (previously 16) provisioned to a cluster on AWS.
- Support for up to 320 TB of storage in a single cluster on AWS.
- Improved I/O performance compared to V8.3.0. on larger Amazon Elastic Compute Cloud. (Amazon EC2) instances as a result of more RAM and processor capacity.
- Data reduction pools with deduplication on AWS.
- Support for software-based compression on AWS.

Announcement letter: For more information, see the announcement letter *IBM Spectrum Virtualize for Public Cloud V8.3.1 expands hybrid multicloud support and offers additional ordering and licensing options*, which is available at this web page.

2

Typical use cases for IBM Spectrum Virtualize for Public Cloud

This chapter covers four use cases for IBM Spectrum Virtualize for Public Cloud and includes the following topics:

2.1 Deploying whole IT services in the public cloud

Companies are approaching and using public cloud services from multiple angles. Users that are rewriting and modernizing applications for cloud complement those users that are looking to move to cloud-only new services or to extend existing IT into a hybrid model to address quickly changing capacity and scalability requirements. Different delivery models are available for public cloud, such as software as a service (SaaS), platform as a service (PaaS), and infrastructure as a service (IaaS).

The workload deployment is composed of two major use cases, as shown in Figure 2-1:

- *Hybrid cloud*: The integration between the off-premises public cloud services with an existing on-premises IT environment.
- *Cloud-native*: The full application's stack is moved to cloud as SaaS, PaaS, IaaS, or as a combination of the three delivery models.

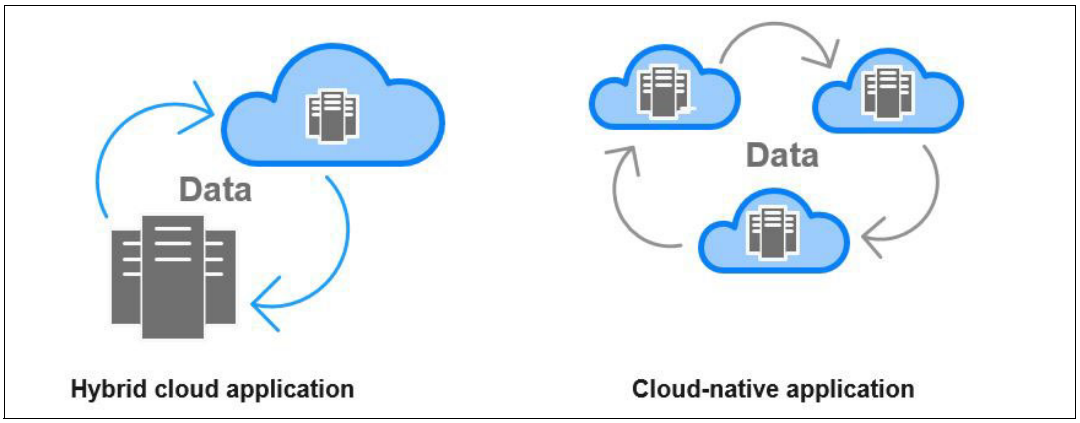


Figure 2-1 The two major deployment models for public cloud

Cloud-native implementations (that is, whole IT services that are deployed in the public cloud) are composed of several use cases, all with the lowest common denominator of having a full application deployment in the public cloud data centers. The technical details, final architecture, and roles and responsibilities depend on SaaS, PaaS, or IaaS usage.

Within the IaaS domain, the transparency of cloud services is the highest because the user's visibility (and responsibility) into the application stack is much deeper compared to the other delivery models. Conversely, the *burden* for its deployment is higher because all the components must be designed from the server up. At the time of this writing, IBM Spectrum Virtualize for Public Cloud is framed within only the IaaS cloud delivery model so that the user can interact with their storage environment as they did on-premises, which provides more granular control over performance.

2.1.1 Business justification

A workload or an application that is stand-alone, with few on-premises dependencies, relatively undemanding of I/O performance (low-latency and low response time and high IOPS), and that is not processing highly regulated data, represents a good fit for a cloud-native deployment. The drivers that motivate businesses towards cloud-native deployment span from capital expenditure and operating expenditure reduction, better resource management and controls against hidden or *shadow* IT resources, more flexibility and scalability, and improved flow in delivering IT service because of the global footprint of the cloud data centers.

At its core, the cloud environment is highly focused on standardization and automation. Therefore, the full spectrum of features and customization that are available in a typical on-premises or outsourcing deployment might not be natively available in the cloud catalog.

Nevertheless, the client does not lose performance and capabilities when deploying a cloud-native application. In this context, the storage virtualization with IBM Spectrum Virtualize for Public Cloud enables the IT staff to maintain the existing technical capabilities and skills to deploy, run, and manage highly available and highly reliable cloud-native applications in a public cloud. In this context, the IBM Spectrum Virtualize for Public Cloud acts as a bridge between the standardized cloud delivery model and the enterprise assets that the client uses in their traditional IT environment.

In a hybrid multicloud environment, the orchestration of the infrastructure requires multiple entities that are tightly integrated with each other and smartly respond to administrator or user needs. It is here where a software-defined environment (SDE) has an important role in the overall orchestration. Integration between service delivery, management, orchestration, automation, and hardware systems is becoming a requirement to support the emergence of SDEs. For SDEs to provide their benefits, they must understand and manage all the components of the infrastructure, including storage, and that makes software-defined storage (SDS) more relevant and important.

Automation is becoming an integral part of IT operations these days, with IT environments that are too complex and must scale up and down quickly for system administrators and developers to meet the business requirements. Ansible is becoming popular for orchestration automation because of the following reasons:

- It is simple to set up and use with no special skills required to create Ansible playbooks.
- You can orchestrate the environment and customize it based on your needs either at on-premises or in hybrid multicloud environments giving the agility and flexibility.
- No need to install any other software or agents on the client system that you want to automate.
- With Ansible and IBM Storage, clients can easily use cutting-edge technology by automating tasks, such as configuration management, provisioning, workflow orchestration, application deployment, and lifecycle management.
- By using Ansible and IBM Storage, clients can reduce system inconsistencies with the automation modules. Ansible can also be used to configure end-to-end infrastructure in an orchestrated fashion.
- Ansible provides a single pane of glass visibility to a multi-cluster, multicloud environment, which allows lines of business to use those playbooks to accomplish their goals without needing to understand the details of how the work is being done.

IBM is a Red Hat Certified Support Module Vendor that provides simple management for IBM Flash system and Spectrum Virtualize for Public Cloud. For more information, see this web page.

2.1.2 Highly available deployment models

The architecture is directly responsible for an application's reliability and availability if a component failure (hardware and software) occurs. When an application is fully hosted on cloud, the cloud data center becomes the primary site (production site).

Cloud deployment does not guarantee 100% uptime, that the backups are available by default, or that the application is automatically replicated between different sites. These security, availability, and recovery features are often not the client's responsibility if the service is delivered by the SaaS model, are partially the user's responsibility in the PaaS model, and are *entirely* the client's responsibility in the IaaS model.

Having reliable cloud deployments means that the service provider must meet the required service level agreement (SLA), which guarantees service availability and uptime. Companies that use a public cloud IaaS can meet required SLAs by implementing highly available solutions and duplicating the infrastructure in the same data center or in two data centers to maintain business continuity if failures occur.

If business continuity is not enough to reach the requirements of the SLA, disaster recovery (DR) implementations, which split the application among multiple cloud data centers (usually with a distance of at least 300 km [186.4 miles]) prevent failure in a major disaster in the organization's main campus. The following highly available deployment models are available for an application that is fully deployed on public cloud:

- Single primary site

  All of the solution's components are duplicated (or more) within the same data center. Although this solution continues to function because no single points of failure (SPOF) exist, it does *not* function if the data center is unavailable.

- Multi-site

  The architecture is split among multiple cloud data centers within the same campus to mitigate the failure of an entire data center or spread globally to recover the solution if a major disaster occurs that affects the campus.

Highly available cloud deployment on a single primary site

When fully moving an application to a cloud IaaS that is the primary site for service delivery, a reasonable approach is implementing at least a highly available architecture. Each component (servers, network components, and storage) is redundant to avoid SPOF.

Within the single primary site deployment, storage often is deployed as native cloud storage. By using the public cloud catalog storage, users can take advantage of the intrinsic availability (and SLAs) of the storage service, which is Amazon Elastic Block Store (EBS) volumes types in this case.

When IBM Spectrum Virtualize for Public Cloud is deployed as clustered pair of Elastic Compute Cloud (EC2) instances, it mediates between the cloud block storage and the workload hosts. In the specific context of single-site deployment, IBM Spectrum Virtualize for Public Cloud supports extra features that enhance the public cloud block-storage offering. At the storage level, IBM Spectrum Virtualize for Public Cloud resolves some limitations because of the standardized model of public cloud providers: a maximum number of LUNs per host, a maximum volume size, and poor granularity in the choice of tiers for storage snapshots.

IBM Spectrum Virtualize for Public Cloud also provides a new view for the storage management other than the cloud portal. It is a high-level view of the storage infrastructure and some limited specific operations at the volume level (such as volume size, IOPS tuning, and snapshot space increase). What is not provided is a holistic view of the storage from the application perspective.

The benefits of an IBM Spectrum Virtualize for Public Cloud single site deployment are listed in Table 2-1.

Table 2-1 Benefits of IBM Spectrum Virtualize for Public Cloud single site deployment

Feature	Benefits
Single point of control for cloud storage resources.	Designed to increase management efficiency and to help to support application availability.
Pools the capacity of multiple storage volumes.	► Helps to overcome volume size limitations. ► Helps to manage storage as a resource to meet business requirements, and not just as a set of independent volumes. ► Helps administrator to better deploy storage as required beyond traditional “islands”. ► Can help to increase the use of storage assets. ► Insulates applications from maintenance or changes to a storage volume offering.
Manages tiered storage.	► Helps to balance performance needs against infrastructures costs in a tiered storage environment. ► Features automated policy-driven control to put data in the right place at the right time automatically among different storage tiers and classes.
Easy-to-use IBM FlashSystem family management interface.	► Has a single interface for storage configuration, management, and service tasks, regardless of the configuration that is available from the public cloud portal. ► Helps administrators use storage assets and volumes more efficiently. ► Has IBM Spectrum Control Insights and IBM Spectrum Protect for extra capabilities to manage capacity and performance.
Dynamic data migration.	► Migrates data among volumes and LUNs without taking applications that use that data offline. ► Manages and scales storage capacity without disrupting applications.
Advanced network-based copy services.	► Copy data across multiple storage systems with IBM FlashCopy. ► Copy data across metropolitan and global distances as needed to create high-availability storage solutions between multiple data centers.

Feature	Benefits
Thin provisioning, data reduction pools, and snapshot replication.	► Reduces volume requirements by using storage only when data changes. ► Improves storage administrator productivity through automated on-demand storage provisioning. ► Supports thin-provisioning in standard and data reduction pools, which reduces capacity requirements by using storage only when data changes. ► Supports data reduction pools with native data reduction features, such as host unmap and reclaiming usable capacity. ► Supports deduplicated and compressed volumes in data reduction pools for more capacity savings. ► Snapshots are available on lower tier storage volumes.
IBM Spectrum Protect Snapshot application-aware snapshots.	► Perform near-instant and application-aware snapshot backups, with minimal performance impact for IBM Db2®, Oracle, SAP, VMware, Microsoft SQL Server, and Microsoft Exchange. ► Provide advanced and granular restoration of Microsoft Exchange data.
Third-party native integration	► Integration with VMware vRealize.

Highly available cloud deployment on multiple sites

When the application architecture spans over multiple data centers, it can tolerate the failure of the entire primary data center by switching to the secondary data center. The primary and secondary data centers can be deployed as:

- *Active-active*: The secondary site is always running and synchronously aligned with the primary site.
- *Active-passive:* The secondary site is always running but asynchronously replicated (with a specific recovery point objective [RPO]) or running for only specific situations, such as acting as a recovery site or test environment. Storage is always active and available for data replication.

The active-passive configuration often is the best fit for many cloud use cases, including DR, as described in 2.2, “Disaster recovery” on page 17. Provisioning compute resources on demand in a few minutes with only the storage that is provisioned and aligned with a specific RPO is a huge driver for a cost-effective DR infrastructure. It also lowers the total cost of ownership (TCO).

The replication among multiple cloud data centers is no different from the traditional approach except for the number of available tools in the cloud. Although existing solutions that are based on hypervisor or application-layer replication, such as VMware, Veeam, and Zerto, are available in the public cloud, storage-based replication is still the preferable approach if the environment is heterogeneous (virtual servers, bare metal servers, multiple hypervisors, and so on).

Active-passive asynchronous mirroring that uses Global Mirror with Change Volumes (GMCV) provides a minimum RPO of 2 minutes (the Change Volume [CV] cycle period ranges is 1 minute - 1 day, and a best practice is setting the cycle period to be half of the RPO), and can replicate a heterogeneous environment.

2.2 Disaster recovery

Since 2018, customers are harnessing and securing proliferating data in their environment and infrastructure workloads have the highest increase in the adoption of DR.

Technology is only one crucial piece of a DR solution, and not the one that dictates the overall approach.

The section describes DR approach and benefits of IBM Spectrum Virtualize for Public Cloud on AWS.

A DR strategy is the predominant aspect of an overall resiliency solution because it determines what classes of physical events the solution can address, sets the requirements in terms of distance, and sets constraints on technology.

2.2.1 Business justification

Table 2-2 lists the drivers and the challenges of having a DR solution on cloud and what capabilities IBM Spectrum Virtualize for Public Cloud provides in these areas.

Table 2-2 Drivers, challenges, and capabilities that are provided by IBM Spectrum Virtualize for Public Cloud

Adoption drivers	Challenges	IBM Spectrum Virtualize for IBM Public Cloud capabilities
The promise of reduced operational expenditures and capital expenditures	► Hidden costs. ► Availability of data when needed.	► Optimized for Cloud Block Storage ► IBM Easy Tier solution to optimize the most valuable storage usage, which maximizes Cloud Block Storage performance ► Thin provisioning to control the storage provisioning ► Snapshots feature for backup and DR solution ► HA clusters architecture
Bridging technologies from on-premises to cloud	Disparate Infrastructure: How can my on-premises production data be readily available in the cloud in a disaster?	► Any to any replication ► Supporting over 400 different storage devices (on-premises), including iSCSI on-premises and when deployed in cloud
Using the cloud for backup and DR	► Covering virtual and physical environments. ► Solutions to meet a range of RPO/RTO needs.	► A storage-based, serverless replication with options for low RPO/RTO: – Global Mirror for Asynchronous replication with an RPO close to “0” – Metro Mirror for Synchronous replication – GMCVs for Asynchronous replication with a tunable RPO

At the time of this writing, IBM Spectrum Virtualize for Public Cloud includes the following DR-related features:

► Can be implemented at any location in AWS Cloud and installed by using AWS Marketplace.

- Is deployed on an Amazon EC2 instance.
- Offers data replication with the FlashSystem family, V9000, SAN Volume Controller, or VersaStack and public cloud.
- Supports two node or four node clusters in AWS Cloud.
- Offers data services for Amazon EBS.
- Offers common management with the IBM Spectrum Virtualize GUI with full admin access and a dedicated instance.
- No incoming data transfer cost.
- Replicates between two AWS Cloud locations.
- Replicates between on-premises and AWS Cloud running IBM Spectrum Virtualize on-premises and IBM Spectrum Virtualize for Public Cloud on AWS.

2.2.2 Two common DR scenarios with IBM Spectrum Virtualize for Public Cloud

The following most common scenarios can be implemented with IBM Spectrum Virtualize for Public Cloud:

- IBM Spectrum Virtualize Hybrid Cloud DR for “Any to Any”
- IBM Spectrum Virtualize for Public Cloud solution on AWS Cloud DR, as shown in Figure 2-2

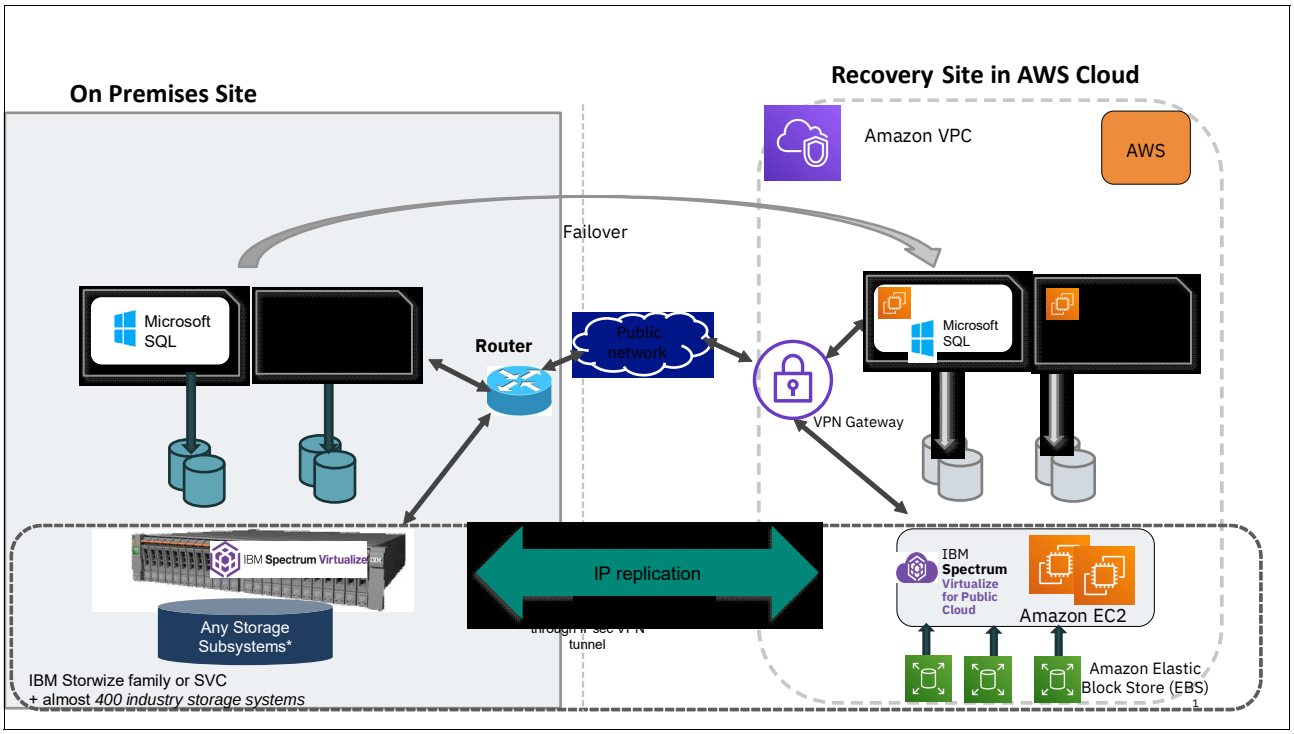


Figure 2-2 IBM Spectrum Virtualize for Public Cloud on AWS Cloud DR solution

As shown in Figure 2-2, a customer can deploy a storage replication infrastructure in a public cloud by using IBM Spectrum Virtualize for Public Cloud.

Consider the following points regarding this scenario:

- Primary storage is in the customer's physical data center. The customer has an on-premises IBM Spectrum Virtualize solution that is installed.
- Auxiliary storage sits on the DR site, which can be an IBM Spectrum Virtualize cluster running in the public cloud.
- The virtual IBM Spectrum Virtualize cluster manages the storage that is provided by an Amazon EBS volume.

A replication partnership that uses GMCVs is established between an on-premises IBM Spectrum Virtualize cluster or FlashSystem solution and the virtual IBM Spectrum Virtualize cluster to provide DR.

When talking about DR, understand that IBM Spectrum Virtualize for Public Cloud is an important piece of a more complex solution that has some prerequisites considerations and best practices that must be applied.

Note: To see an example of a simple implementation of a DR solution, including IBM FlashSystem and IBM Spectrum Virtualize for Public Cloud, see 5.5, "Shrinking the configuration from four nodes to two nodes in Amazon Web Services" on page 85.

2.3 IBM FlashCopy in the public cloud

The IBM FlashCopy function in IBM Spectrum Virtualize can perform a point-in-time (PiT) copy of one or more volumes. You can use FlashCopy to help you solve critical and challenging business needs that require duplication of data of your source volume. Volumes can remain online and active while you create consistent copies of the data sets. Because the copy is performed at the block level, it operates below the host operating system and its cache. Therefore, the copy is not apparent to the host unless it is mapped.

2.3.1 Business justification

The business applications for FlashCopy are wide-ranging. Common use cases for FlashCopy include, but are not limited to, the following examples:

- Rapidly creating consistent:
  - Backups of dynamically changing data
  - Copies of production data to facilitate data movement or migration between hosts
- Rapidly creating copies of:
  - Production data sets for application development and testing
  - Production data sets for auditing purposes and data mining
  - Production data sets for quality assurance
  - Replication targets for testing data integrity.

Regardless of your business needs, FlashCopy with IBM Spectrum Virtualize is flexible and offers a broad feature set, which makes it applicable to many scenarios.

2.3.2 FlashCopy mapping

The association between the source volume and the target volume is defined by a *FlashCopy mapping*. The FlashCopy mapping can have three different *types*, four *attributes*, and seven different *states*.

FlashCopy in the GUI can be one of the three types:

- Snapshot: Sometimes referred to as *nocopy*. A PiT copy of a volume without a background copy of the data from the source volume to the target. Only the changed blocks on the source volume are copied. The target copy cannot be used without an active link to the source, which is achieved by setting the copy and clean rate to zero.
- Clone: Sometimes referred to as *full copy*. A PiT copy of a volume with a background copy of the data from the source volume to the target. All blocks from the source volume are copied to the target volume. The target copy becomes a usable independent volume, which is achieved with a copy and clean rate greater than zero and an autodelete flag; therefore, no cleanup is necessary after the background copy is finished.
- Backup: Sometimes referred to as *incremental*. A backup FlashCopy mapping consists of a PiT full copy of a source volume, plus periodic increments or “deltas” of data that changed between two points in time. This mapping is where the copy and clean rates are greater than zero, no autodelete flag is set, and you use an incremental flag to preserve the bitmaps between activations so that only the deltas since the last “backup” must be copied.

The FlashCopy mapping has four property attributes (clean rate, copy rate, autodelete, and incremental) and seven different states. The actions users can perform on a FlashCopy mapping are:

- Create: Define a source and a target, and set the properties of the mapping.
- Prepare: The system must be prepared before a FlashCopy copy starts. It basically flushes the cache and makes it “transparent” for a short time so that no data is lost.
- Start: The FlashCopy mapping is started and the copy begins immediately. The target volume is immediately accessible.
- Stop: The FlashCopy mapping is stopped (by the system or user). Depending on the state of the mapping, the target volume is usable or not.
- Modify: Some properties of the FlashCopy mapping can be modified after creation.
- Delete: Delete the FlashCopy mapping. This does not delete any of the volumes (source or target) from the mapping.

The source and target volumes must be the same size. The minimum granularity that IBM Spectrum Virtualize supports for FlashCopy is an entire volume. It is not possible to use FlashCopy to copy only part of a volume.

Important: As with any PiT copy technology, you are bound by operating system and application requirements for interdependent data and the restriction to an entire volume.

The source and target volumes must belong to the same IBM Spectrum Virtualize system, but they do not have to be in the same I/O group or storage pool.

Volumes that are members of a FlashCopy mapping cannot have their sizes increased or decreased while they are members of the FlashCopy mapping.

All FlashCopy operations occur on FlashCopy mappings. FlashCopy does not alter source volumes. Multiple operations can occur at the same time on multiple FlashCopy mappings by using *consistency groups*.

2.3.3 Consistency groups

To overcome the issue of dependent writes across volumes and create a consistent image of the client data, perform a FlashCopy operation on multiple volumes as an atomic operation. To accomplish this task, IBM Spectrum Virtualize supports the concept of consistency groups. *Consistency groups* preserve PiT data consistency across multiple volumes for applications that include related data that spans multiple volumes. For these volumes, consistency groups maintain the integrity of the FlashCopy by ensuring that *dependent writes* are run in the application's intended sequence.

FlashCopy mappings can be part of a consistency group, even if only one mapping is in the consistency group. If a FlashCopy mapping is not part of any consistency group, it is referred to as *stand-alone*.

2.3.4 Crash-consistent copy and host considerations

FlashCopy consistency groups do not provide application consistency. They ensure only that volume points-in-time are consistent between volumes.

Because FlashCopy is at the block level, you must understand the interaction between your application and the host operating system. From a logical standpoint, it is easiest to think of these objects as "layers" that sit on top of one another. The application is the topmost layer, and beneath it is the operating system layer.

Both of these layers feature various levels and methods of caching data to provide better speed. Because the IBM SAN Volume Controller and FlashCopy sit below these layers, they are unaware of the cache at the application or operating system layers.

To ensure the integrity of the copy that is made, it is necessary to flush the host operating system and application cache for any outstanding reads or writes before the FlashCopy operation is performed. Failing to flush the host operating system and application cache produces what is referred to as a *crash-consistent* copy.

The resulting copy requires the same type of recovery procedure, such as log replay and file system checks, that is required following a host crash. FlashCopy copies that are crash-consistent often can be used after the file system and application recovery procedures.

Various operating systems and applications provide facilities to stop I/O operations and ensure that all data is flushed from the host cache. If these facilities are available, they can be used to prepare a FlashCopy operation. When this type of facility is unavailable, the host cache must be flushed manually by quiescing the application and unmounting the file system or drives.

The target volumes are overwritten with a complete image of the source volumes. Before the FlashCopy mappings are started, it is important that any data that is held on the host operating system (or application) caches for the target volumes is discarded. The easiest way to ensure that no data is held in these caches is to unmount the target volumes *before* the FlashCopy operation starts.

Best practice: From a practical standpoint, when you have an application that is backed by a database and you want to make a FlashCopy of that application's data, it is sufficient in most cases to use the write-suspend method that is available in most modern databases because the database maintains strict control over I/O.

This method is as opposed to flushing data from the application and backing database, which is always the suggested method because it is safer. However, this method can be used when facilities do not exist or your environment includes time sensitivity.

2.4 Workload relocation into the public cloud

In this section, a use case for IBM Spectrum Virtualize for Public Cloud is illustrated where an entire workload segment is migrated from a client's enterprise into the cloud. Although the process for relocating a workload into the cloud by using IBM Spectrum Virtualize can use only Remote Copy, other mechanisms are available that can accomplish this task.

2.4.1 Business justification

All the drivers that motivate businesses to use virtualization technologies makes deploying services into the cloud even more compelling because the cost of idle resources is further absorbed by the cloud provider. However, certain limitations in regulatory or process controls might prevent a business from moving all workloads and application services into the cloud.

An ideal case with regards to a hybrid cloud solution is the relocation of a specific segment of the environment that is well-suited, such as development. Another case might be a specific application group that does not require the regulatory isolation or low response time integration with on-premises applications.

Although performance might be a factor, do not assume that cloud deployments automatically create a diminished performance. Depending on the location of the cloud service data center and the intended audience for the migrated service, the performance conceivably can be superior to on-premises pre-migration.

In summary, moving a workload into the cloud might provide similar functions with better economies because of scaling of physical resources in the cloud provider. Moreover, the cost of services in the cloud is structured, measurable, and predictable.

2.4.2 Data migration

Several methods are available for performing data migrations to the cloud, including the following general approaches:

- IBM Spectrum Virtualize Remote Copy
- Host-side mirroring (Storage vMotion or IBM AIX® Logical Volume Manager mirroring)
- Appliance-based data transfer, such as IBM Aspera® or IBM Transparent Data Migration Facility

The first method was described in 2.3, “IBM FlashCopy in the public cloud” on page 19, and is essentially the same process as DR. The only difference is that instead of a persistent replication, after the initial synchronization is complete, the goal is to schedule the cutover of the application onto the compute nodes in the cloud environment that is attached to the IBM Spectrum Virtualize storage.

Host-side mirroring requires the server to have concurrent access to local and remote storage, which is not feasible. Also, because the object is to relocate the workload (both compute and storage) into the cloud environment, that task is more easily accomplished by replicating the storage and, after it is synchronized, bringing up the server in the cloud environment and making the appropriate adjustments to the server for use in the cloud.

The second method is largely impractical because it requires the host to access source and target simultaneously, and the practical impediments to creating an iSCSI (the only connection method that is available for IBM Spectrum Virtualize in the Public Cloud) connection from on-premises host systems into the cloud are beyond the scope of this use case. Traditional VMware Storage vMotion is similar, but again requires the target storage to be visible through iSCSI to the host.

The third method includes the use of third-party software or hardware to move the data from one environment to another one. The general idea is that the target system has an operating system and some empty storage that is provisioned to it that acts as a landing pad for data that is on the source system. More information about these methods is also outside the scope of this publication, but the process is no different between an on-premises to cloud migration as it is to an on-premises to on-premises migration.

Table 2-3 lists the migration methods.

Table 2-3 Migration methods

Method	Best-suited operating system	Pros versus cons
Remote Copy	Stand-alone Windows, Linux, or VMWare (any version)	Simple versus limited scope
Host Mirror	VMWare vSphere 5.1 or higher	Simple versus limited scope
Appliance	N/A	Flexible versus cost and complexity

2.4.3 Host provisioning

In addition to the replication of data, compute nodes and networking must be provisioned within the cloud provider upon which to run the relocated workload. Currently, in the AWS Cloud, the EC2 compute nodes are available with storage that is provisioned to the EC2 compute instance by using an iSCSI connection.

2.4.4 Implementation considerations

The workload relocation into the public cloud use case includes the following implementation considerations:

- Naming conventions

  This important consideration concerns the manageability of a standard on-premises IBM Spectrum Virtualize environment. However, because of the multiple layers of virtualization in a cloud implementation, maintaining a consistent and meaningful naming convention for all objects (managed disks (MDisks), volumes, FlashCopy mappings, Remote Copy relationships, hosts, and host clusters) is necessary.

- Monitoring integration

  Integration into IBM Spectrum Control or some other performance monitoring framework is useful for maintaining metrics for reporting or troubleshooting. IBM Spectrum Control is well-suited for managing IBM Spectrum Virtualize environments.

- Planning and scheduling

  Regardless of the method that is chosen, gather as much information ahead of time as possible, such as file system information, application custodians, and full impact analysis of related systems.

- Ensure a solid backout

  If inter-related systems or other circumstances require rolling back the application servers to on-premises, plan the migration to ensure as little difficulty as possible in the roll-back, which might mean keeping zoning in the library (even if it is not in the active configuration), and not destroying source volumes for a specific period.

2.5 IBM Spectrum Virtualize for Public Cloud with RedHat OpenShift Container Platform

Containerized environment can be deployed by using RedHat OpenShift Container Platform and IBM block CSI (Container Storage Interface) driver plug-in, for IBM Spectrum Virtualize on Public Cloud AWS cloud.

Red Hat OpenShift Container Platform provides developers and IT organizations with a hybrid cloud application platform for deploying new and existing applications on secure, scalable resources with minimal configuration and management overhead. OpenShift Container Platform supports a wide selection of programming languages and frameworks, such as Java, JavaScript, Python, Ruby, and PHP.

Built on Red Hat Enterprise Linux and Kubernetes, OpenShift Container Platform provides a more secure and scalable multi-tenant operating system for today's enterprise-class applications, while delivering integrated application runtimes and libraries. OpenShift Container Platform enables organizations to meet security, privacy, compliance, and governance requirements.

For more information about Red Hat OpenShift Container Platform and latest version, see this Red Hat web page.

The high-level architecture for deploying IBM Spectrum Virtualize for Public Cloud with RedHat OpenShift Container Platform and CSI drivers is shown in Figure 2-3.

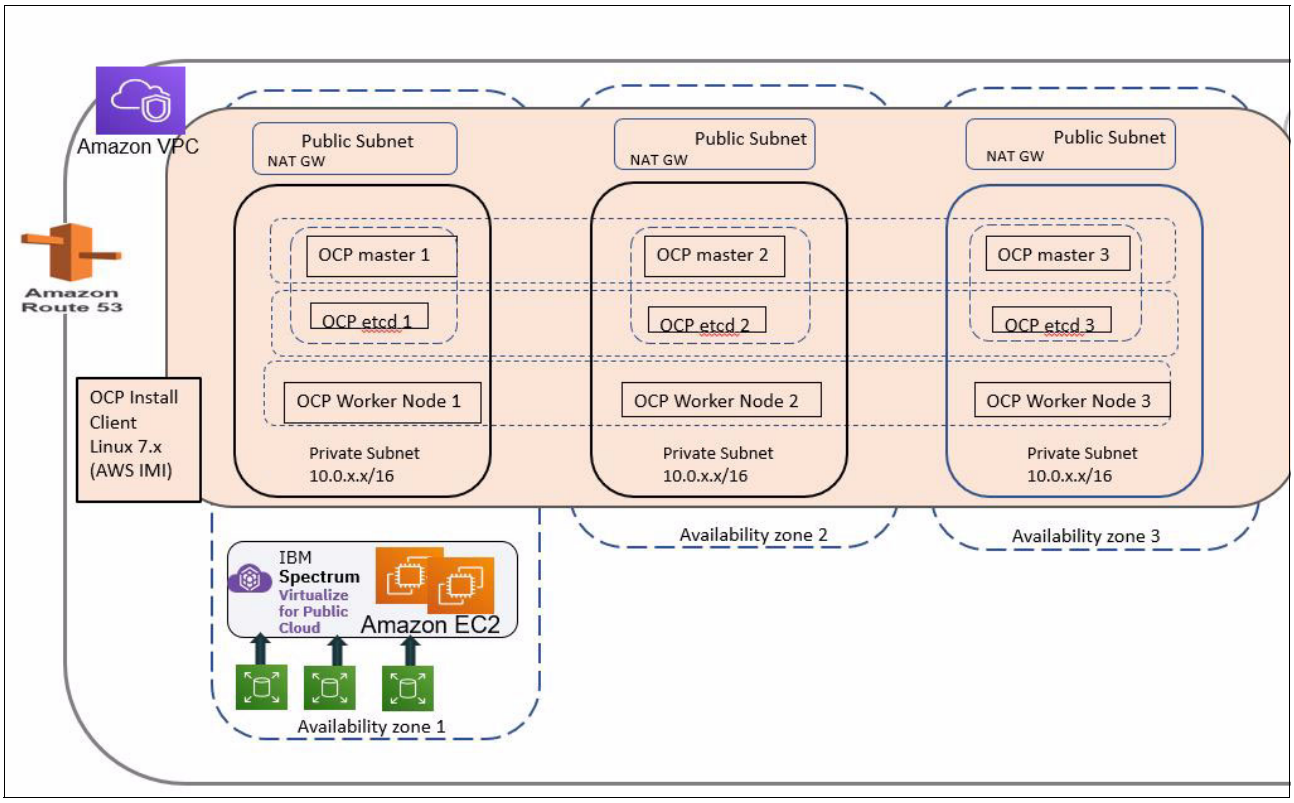


Figure 2-3 IBM Spectrum Virtualize for Public Cloud with RedHat OpenShift Container Platform and CSI drivers

IBM released its open source CSI driver, which allows dynamic provisioning storage for containers on Kubernetes and Red Hat OpenShift container platform by using IBM Storage subsystems.

IBM Spectrum Storage family and IBM Spectrum Virtualize for Public Cloud on AWS supports clients in their IT architectural transformation and migration towards the cloud service model. This support enables hybrid cloud strategies or for a cloud-native workload, provides the benefits of familiar and sophisticated storage functions on public cloud data centers, which enhances the cloud offering.

3

Solution architecture

This chapter provides a technical overview of the Amazon Web Services (AWS) environment regarding IBM Spectrum Virtualize for Public Cloud deployment. It also provides functional definitions of the solution components and how they interact and interrelate.

This chapter includes the following topics:

3.1 Amazon Web Services

AWS delivers infrastructure as a service (IaaS) in the form of virtual private clouds (VPCs) within which network, compute, and storage resources are housed. IBM Spectrum Virtualize nodes are built on EC2 instances and virtualize Amazon EBS volumes that are provisioned to those nodes, as shown in Figure 3-1. This configuration provides advanced capacity savings functions and replication and point-in-time (PiT) copy services over a block virtualization layer through a user interface that is familiar to IBM Spectrum Virtualize clients.

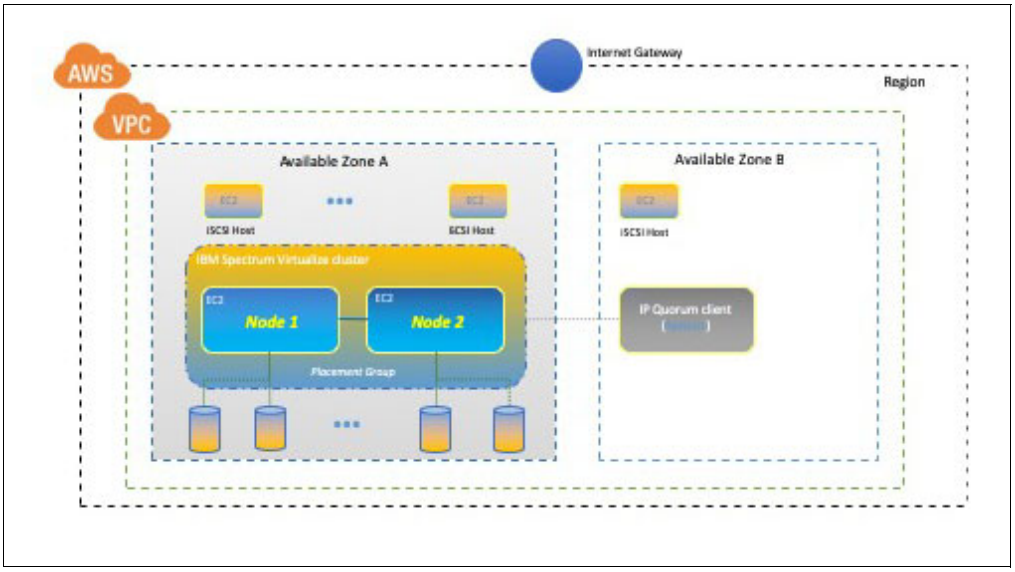


Figure 3-1 Architectural overview of high-level AWS components

For more information about the common components in AWS, see Table 3-1.

Table 3-1 Definition of terminology in AWS

Item	Definition
Elastic Compute Cloud (EC2)	A service that you can use to start virtual machine (VM) instances in various operating systems.
Elastic Block Store (EBS)	Persistent block storage volumes that are used with Amazon EC2 (as opposed to the more common Simple Storage Service [S3]).
Availability zones	Distinct locations that are insulation from failures.
Virtual private cloud (VPC)	Virtual network in your own logically isolated area within the AWS Cloud. It is populated by infrastructure, platform, and application services that share common security and interconnection.
CloudFormation Template	Creates and configures AWS resources and discovers dependencies.
Amazon Machine Images (AMI)	Template that contains a software configuration (for example, an operating system, an application server, and applications).
Simple Storage Service (S3)	Storage for the internet. You can use Amazon S3 to store and retrieve any amount of data at any time, from anywhere on the web.

3.2 IBM Spectrum Virtualize

IBM Spectrum Virtualize is a software-enabled storage virtualization engine that provides a single point of control for storage resources within the data centers. IBM Spectrum Virtualize is a core software engine of established and IBM storage virtualization solutions, such as IBM SAN Volume Controller and all versions of the IBM FlashSystem family of products. This technology is now available in AWS, which provides increased flexibility in data center infrastructure and cloud systems. This section describes the components of IBM Spectrum Virtualize as they are deployed in the cloud.

3.2.1 Nodes

IBM Spectrum Virtualize software is installed on EC2 instances that are provisioned in AWS. Each EC2 is called a *node*. The node provides the virtualization for a set of volumes, cache, and copy services functions. The nodes are deployed in pairs (*I/O groups*) and 1 - 4 pairs make up a *clustered system*. At the time of this writing, IBM Spectrum Virtualize for Public Cloud on AWS is limited to one or two I/O group, with plans to expand to four I/O groups.

One of the nodes within the system is assigned the role of the *configuration node*. The configuration node manages the configuration activity for the system and owns the cluster IP address that is used to access the management GUI and command-line interface (CLI) connections. If this node fails, the system chooses a new node to become the configuration node.

Because the active nodes are installed in pairs, each node maintains cache coherence with its partner to provide seamless failover functions and fault tolerance, which are described next.

3.2.2 I/O groups

A specific *volume* is always presented to a host server or cluster by a single I/O group in the system. When a host server performs I/O to one of its volumes, all the I/Os for a specific volume are directed to one specific I/O group in the system. Under normal conditions, the I/Os for that specific volume are always processed by the same node within the I/O group. This node is referred to as the *preferred node* for this specific volume.

When the preferred node receives a write into its cache, that write is mirrored to the partner node before the write is acknowledged back to the host. Reads are serviced by the preferred node. For more information, see 3.2.8, “Cache” on page 32.

Both nodes of an I/O group act as the preferred node for their own specific subset of the total number of volumes that the I/O group presents to the host servers. However, both nodes also act as failover nodes for their respective partner node within the I/O group. Therefore, a node takes over the I/O workload from its partner node, if required. For this reason, it is mandatory for servers that are connected to use multipath drivers to handle these failover situations.

If required, host servers can be mapped to more than one I/O group within the IBM Spectrum Virtualize system. Therefore, they can access volumes from separate I/O groups. You can move volumes between I/O groups to redistribute the load between the I/O groups. Modifying the I/O group that services the volume can be done concurrently with I/O operations if the host supports nondisruptive volume moves and is zoned to support access to the target I/O group.

It also requires a rescan at the host level to ensure that the multipathing driver is notified that the allocation of the preferred node changed, and the ports by which the volume is accessed changed. This modification can be done in the situation where one pair of nodes becomes overused.

3.2.3 System

The current IBM Spectrum Virtualize for Public Cloud on AWS system or clustered system consists of one or two I/O group with plans to support more. Certain configuration limitations are then set for the individual system. For example, at the time of this writing, the maximum number of basic volumes that is supported per system is 10000, and the maximum managed disks (MDisks) that are supported is ~28 PIB (pebibytes) or 32 PB (petabytes) per system.

The current AWS implementation is optimized around 20 Amazon EBS volumes and the largest single Amazon EBS volume on AWS is 16384 GiB. Because of those limitations, the practical limit of MDisks that are managed is 20x16 TB or 320 TB.

All configuration, monitoring, and service tasks are performed at the system level. Configuration settings are replicated to all nodes in the system. To facilitate these tasks, a management IP address is set for the system.

Note: The management IP is also referred to as the system or cluster IP and is active on the configuration node. Each node in the system is also assigned a service IP to allow for individually interacting with the node directly.

A process is provided to back up the system configuration data onto disk so that it can be restored if a disaster occurs. This method does not back up application data. Only the IBM Spectrum Virtualize system configuration information is backed up.

For the purposes of remote data mirroring, two or more systems must form a *partnership* before relationships between mirrored volumes are created. Currently, Spectrum Virtualize is limited to a single IP-based partnership.

For more information about the maximum configurations that apply to the system, I/O group, and nodes, see V8.3.1.x Configuration Limits and Restrictions for IBM Spectrum Virtualize for Public Cloud.

3.2.4 MDisks

IBM Spectrum Virtualize for Public Cloud on AWS views the Amazon EBS volumes that are presented to the EC2 instance nodes by AWS as several disks or LUNs, which are known as *MDisks*. Because IBM Spectrum Virtualize does not attempt to provide recovery from physical failures within the back-end controllers, an MDisk often is typically provisioned from a RAID array and you assume that the Amazon EBS volumes are suitably protected and redundant.

However, the application servers do not see the MDisks. Rather, they see several logical disks, which are known as *virtual disks* or *volumes*, which are presented by the I/O groups through the LAN (iSCSI) to the servers. The MDisks are placed into storage pools where they are divided into extents that are used to create the *virtual disks* or *volumes*.

For more information about the total storage capacity that is manageable per system regarding the selection of extents, see V8.3.1.x Configuration Limits and Restrictions for IBM Spectrum Virtualize for Public Cloud.

MDisks that are presented to IBM Spectrum Virtualize can have the following modes of operation:

- Unmanaged MDisk

  An MDisk is reported as unmanaged when it is not a member of any storage pool. An unmanaged MDisk is not associated with any volumes and has no metadata that is stored on it. IBM Spectrum Virtualize does not write to an MDisk that is in unmanaged mode, except when it attempts to change the mode of the MDisk to one of the other modes.

- Managed MDisk

  Managed MDisks are members of a storage pool and they contribute extents to the storage pool. This mode is the most common and normal mode for an MDisk.

3.2.5 Storage pool

A *storage pool* or *MDisk group* is a collection of MDisks that provides the pool of storage from which volumes are provisioned. The size of these pools can be changed (expanded or shrunk) nondisruptively by adding or removing MDisks without taking the storage pool or the volumes offline. At any point, an MDisk can be a member in one storage pool only.

Each MDisk in the storage pool is divided into extents. The size of the extent is selected by the administrator when the storage pool is created and cannot be changed later, although methods are available to address this issue with volume mirroring (see 3.2.7, “Volumes” on page 32). The size of the extent can be 16 MiB (mebibyte) - 8192 MiB, with the default being 1024 MiB unless a Data Reduction Pool (DRP), in which case the default extent size is 4096 MiB.

It is a best practice to use the same extent size for all storage pools in a system. This approach is a prerequisite for supporting volume migration between two storage pools. If the storage pool extent sizes are not the same, you must use volume mirroring to copy volumes between pools.

3.2.6 Data Reduction Pools

Spectrum Virtualize code introduces Data Reduction Pools (DRP) to Spectrum Virtualize for Public Cloud. This functionality provides important extra capacity savings.

To realize these capacity savings, a single Data Volume is available per I/O group to which data is written with each of the logical disks or volumes containing a set of pointers to the Data Volume. Multiple volumes can point to a single region of identical data, thus achieving deduplication.

All volumes in Spectrum Virtualize are constrained to a 128,000 extent limit. Therefore, the practical limit for a DRP becomes 128 TB per I/O group with the normal default extent size of 1024 MiB. To make this limitation less impactful, the default extent size for DRPs is now 4096 MiB so that a DRP can address per I/O group is 512 TB.

For more information about Data Reduction Pools, see *Introduction and Implementation of Data Reduction Pools and Deduplication*, SG24-8430.

3.2.7 Volumes

Volumes are logical disks that are presented to the host or application servers by IBM Spectrum Virtualize. The hosts cannot see the MDisks; they can see only the logical volumes that are created from combining extents from a storage pool.

An IBM Spectrum Virtualize for Public Cloud on AWS volume is allocated one extent in turn from each MDisk in the storage pool. This process continues until the space that is required for the volume is satisfied.

It is also possible to supply a specific list of MDISKs to use, although this is typically not done. Regardless of whether all MDISKs or a subset of MDISKs are used, the default volume type is striped, which rotates through the MDISKs in the pool. Other volume types are supported by the IBM Spectrum Virtualize product, but striped is the only type that is relevant for Spectrum Virtualize for Public Cloud on AWS implementations.

3.2.8 Cache

The primary benefit of storage cache is to improve I/O response time. Reads and writes to a magnetic disk drive experience seek and latency time at the drive level, which can result in 1 ms - 10 ms of response time (for an enterprise-class disk).

IBM Spectrum Virtualize provides a flexible cache model, and the node's memory can be used as read or write cache. The cache management algorithms allow for improved performance of many types of underlying disk technologies. The IBM Spectrum Virtualize capability to manage in the background the destaging operations that are incurred by writes (in addition to still supporting full data integrity) helps the IBM Spectrum Virtualize capability to achieve good database performance.

The cache is separated into two layers: upper cache and lower cache.

Figure 3-2 on page 33 shows the separation of the upper and lower cache.

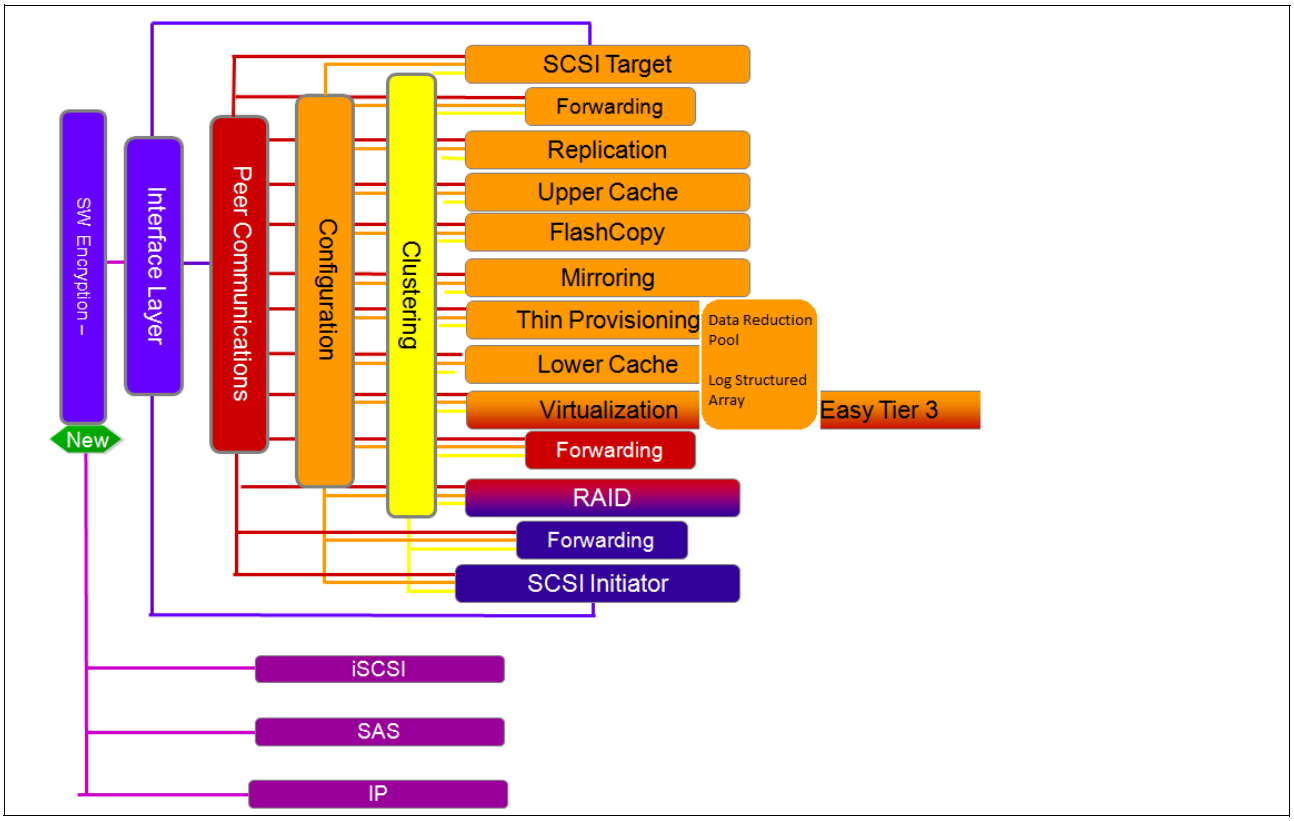


Figure 3-2 Separation of upper and lower cache

The upper cache delivers fast write response times to the host by being as high up in the I/O stack as possible. The lower cache works to help ensure that cache between nodes are in sync, pre-fetches data for an increased read cache hit ratio on sequential workloads, and optimizes the destaging of I/O to the backing storage controllers.

Combined, the two levels of cache also deliver the following functions:

- Pins data when the LUN goes offline.
- Provides enhanced statistics for IBM Spectrum Control or IBM Storage Insights, and maintains compatibility with an earlier version.
- Provides trace data for debugging.
- Reports media errors.
- Resynchronizes cache correctly and provides the atomic write function.
- Ensures that other partitions continue operation when one partition becomes 100% full of pinned data.
- Supports fast-write (two-way and one-way), flush-through, and write-through.
- Integrates with T3 recovery procedures.
- Supports two-way operations.
- Supports none, read-only, and read/write as user-exposed caching policies.
- Supports flush-when-idle.
- Supports expanding cache as more memory becomes available to the platform.

- Supports credit throttling to avoid I/O skew and offer fairness and balanced I/O between the two nodes of the I/O group.
- Enables switching of the preferred node without needing to move volumes between I/O groups.

3.2.9 IBM Easy Tier

IBM Easy Tier is a performance function that automatically migrates or moves extents of a volume to or from one MDisk storage tier to another MDisk storage tier. IBM Spectrum Virtualize code can support a three-tier implementation, although version 8.3.1 of the code now supports up to five different tier definitions (Storage Class Memory or SCM, tier0 flash, tier1 flash, enterprise, and nearline). However, only three tier types can be specified in a single pool: generally, a high-performance tier (SCM, T0 or T1 flash), midrange or enterprise tier, and a cold nearline tier.

Easy Tier monitors the host I/O activity and latency on the extents of all volumes with the Easy Tier function, which is turned on in a multitier storage pool over a 24-hour period.

Next, it creates an extent migration plan that is based on this activity. It then dynamically moves high-activity or hot extents to a higher disk tier within the storage pool. It also moves extents whose activity dropped off or cooled down from the high-tier MDisks back to a lower-tiered MDisk. The condition for hot extents is frequent small block (64 Kb or less) reads.

Easy Tier: The Easy Tier function can be turned on or off at the storage pool and volume level.

The automatic load-balancing (*auto-rebalance*) function is enabled by default on each volume and cannot be turned off by using the GUI. This load-balancing feature is not considered the same as the Easy Tier function, but it uses the same principles. Auto-balance evens the load for a pool across MDisks. Therefore, even the addition of new MDisks, or having MDisks of different sizes within a pool, does not adversely affect the performance.

The Easy Tier function can make it more suitable to use smaller storage pool extent sizes. The usage statistics file can be offloaded from the IBM Spectrum Virtualize nodes. Then, you can use the IBM Storage Advisor Tool (STAT) to create a summary report. STAT is available at no initial cost at this IBM Support web page. Starting with Spectrum Virtualize 8.3.1, the STAT tool is incorporated into the GUI as “EasyTier Reports.”

3.2.10 Hosts

Volumes can be mapped to a host to allow access for a specific server to a set of volumes. A host within the IBM Spectrum Virtualize is a collection of iSCSI-qualified names (IQNs) that are defined on the specific server.

The iSCSI software in IBM Spectrum Virtualize supports IP address failover when a node is shut down or restarted. As a result, a node failover (when a node is restarted) can be handled without having a multipath driver that is installed on the iSCSI-attached server. An iSCSI-attached server can reconnect after the node shutdown to the original target IP address, which is now presented by the partner node. However, to protect the server against link failures in the network, you need a multipath driver. As a result, it is a best practice to implement multipathing on all hosts that are attached to IBM Spectrum Virtualize systems.

3.2.11 Host cluster

A *host cluster* is a host object in IBM Spectrum Virtualize. A host cluster is a combination of two or more servers that is connected to IBM Spectrum Virtualize through an iSCSI connection. A host cluster object can see the same set of volumes, so volumes can be mapped to a host cluster to allow all hosts to have a common mapping.

3.2.12 iSCSI

The *iSCSI function* is a software function that is provided by the IBM Spectrum Virtualize code. IBM introduced software capabilities to allow the underlying virtualized storage to attach to IBM Spectrum Virtualize by using the iSCSI protocol.

The major functions of iSCSI include encapsulation and the reliable delivery of Command Descriptor Block (CDB) transactions between initiators and targets through the IP network, especially over a potentially unreliable IP network.

Every iSCSI node in the network must have an iSCSI name and address. Consider the following points:

- An *iSCSI name* is a location-independent, permanent identifier for an iSCSI node. An iSCSI node has one iSCSI name, which stays constant for the life of the node. The terms *initiator name* and *target name* also refer to an iSCSI name.
- An *iSCSI address* specifies the iSCSI name and location of an iSCSI node. The address consists of a host name or IP address, a TCP port number (for the target), and the iSCSI name of the node. An iSCSI node can have any number of addresses, which can change at any time, particularly if they are assigned by way of Dynamic Host Configuration Protocol (DHCP). An IBM Spectrum Virtualize node represents an iSCSI node and provides statically allocated IP addresses.

3.2.13 IP replication

IP replication allows data replication between IBM Spectrum Virtualize family members. IP replication uses IP-based ports of the cluster nodes.

The configuration of the system is straightforward and IBM FlashSystem family systems normally find each other in the network and can be selected from the GUI.

IP replication includes *Bridgeworks SANSlide* network optimization technology and is available at no additional charge. Remote Mirror is a chargeable option, but the price does not change with IP replication. Existing Remote Mirror users can access the function at no extra charge.

IP connections that are used for replication can have long latency (the time to transmit a signal from one end to the other), which can be caused by distance or by many “hops” between switches and other appliances in the network. Traditional replication solutions transmit data, wait for a response, and then transmit more data, which can result in network utilization as low as 20% (based on IBM measurements). In addition, this scenario gets worse the longer the latency.

Bridgeworks SANSlide technology, which is integrated with the IBM FlashSystem family, requires no separate appliances and so requires no extra cost or configuration steps. It uses

artificial intelligence (AI) technology to transmit multiple data streams in parallel, adjusting automatically to changing network environments and workloads.

Bridgeworks SANSlide improves network bandwidth utilization up to 3x. Therefore, customers can deploy a less costly network infrastructure, or take advantage of faster data transfer to speed replication cycles, improve remote data currency, and enjoy faster recovery.

IP replication is limited to a single partnership in the current code.

3.2.14 Synchronous or asynchronous remote copy

The general application of remote copy seeks to maintain two copies of data. Often, the two copies are separated by distance, but not always. The remote copy can be maintained in either synchronous or asynchronous modes. IBM Spectrum Virtualize, Metro Mirror, and Global Mirror are the IBM branded terms for the functions that are synchronous remote copy and asynchronous remote copy.

Synchronous remote copy ensures that updates are committed at the primary and secondary volumes before the application considers the updates complete. Therefore, the secondary volume is fully up to date if it is needed in a failover. However, the application is fully exposed to the latency and bandwidth limitations of the communication link to the secondary volume and the performance characteristics of the storage device on which the secondary volume resides. In a truly remote situation, this extra latency can have a significant adverse effect on application performance at the primary site.

Special configuration guidelines exist for SAN fabrics and IP networks that are used for data replication. Considerations must be taken regarding the distance and available bandwidth of the intersite links.

A function of Global Mirror for low bandwidth was introduced in IBM Spectrum Virtualize 6.3 code. It uses Change Volumes (CVs) that are associated with the primary and secondary volumes. These CVs are special system generated point-in-time copies or FlashCopies that are used to record changes to the primary volume that are transmitted to the remote volume at an interval that is specified by the cycle period.

When a successful transfer of changes from the master CV to the auxiliary volume is achieved within a cycle period, a snapshot is taken at the remote site from the auxiliary volume onto the auxiliary CV to preserve a consistent state and a freeze time is recorded. This function is enabled by setting the *Global Mirror cycling mode*. Figure 3-3 shows an example of this function where you can see the association between volumes and CVs.

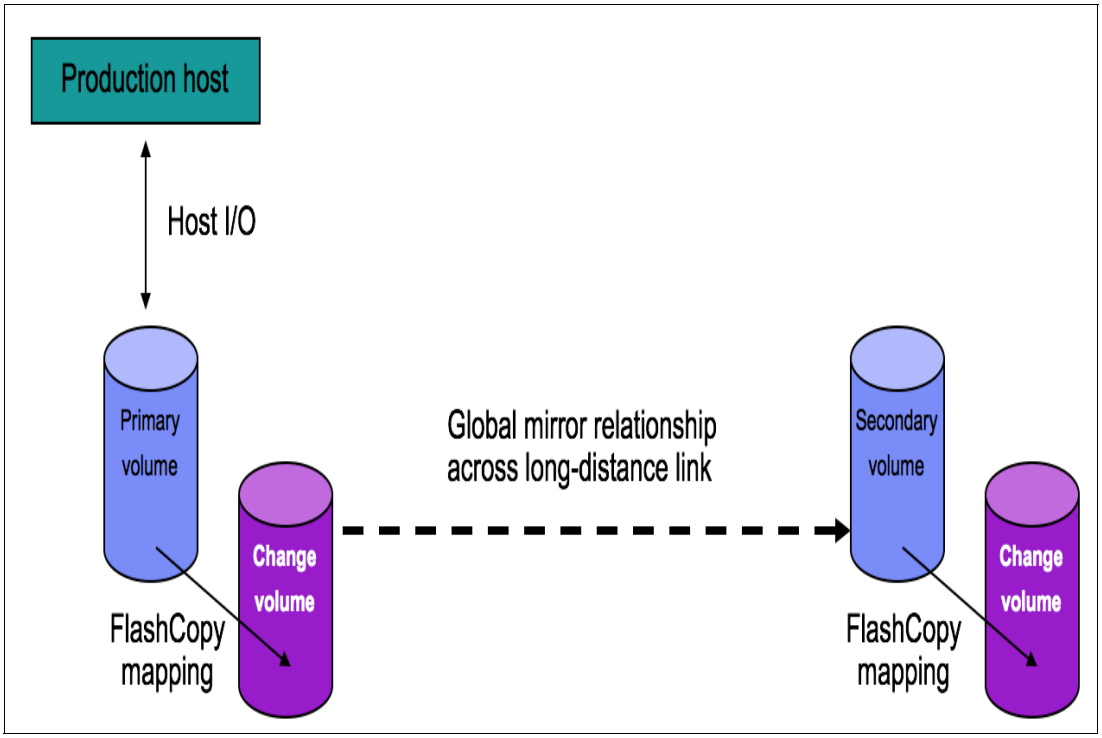


Figure 3-3 Global Mirror cycling mode

3.2.15 IBM FlashCopy

FlashCopy is sometimes described as an instance of a time-zero (T0) copy or a PiT (Point-in-Time) copy technology.

FlashCopy can be performed on multiple source and target volumes. FlashCopy permits the management operations to be coordinated so that a common single PiT is chosen for copying target volumes from their respective source volumes.

With IBM Spectrum Virtualize, multiple target volumes can undergo FlashCopy from the same source volume. This capability can be used to create images from separate PiTs for the source volume, and to create multiple images from a source volume at a common PiT. Source and target volumes can be thin-provisioned volumes.

Reverse FlashCopy enables target volumes to become restore points for the source volume without breaking the FlashCopy relationship, and without waiting for the original copy operation to complete. IBM Spectrum Virtualize supports multiple targets, and has multiple rollback points.

Most clients aim to integrate the FlashCopy feature for PiT copies and quick recovery of their applications and databases. An IBM solution is provided by IBM Spectrum Protect, which is described in What can IBM Spectrum Protect do for your business?

For a use case for using FlashCopy with IBM Spectrum Virtualize for Public Cloud, see 2.3, “IBM FlashCopy in the public cloud” on page 19.

3.3 Solution architecture

This section describes the solution architecture.

3.3.1 Overview

As shown in Figure 3-1 on page 28, the IBM Spectrum Virtualize for Public Cloud environment is contained in a set of networks in a Virtual Private Cloud (VPC), the IBM Spectrum Virtualize Cluster is on EC2 nodes, the Bastion and initial IP quorum are on another smaller EC2 node, and the second IP quorum is in another availability zone for redundancy. As shown in Figure 3-4, those components are placed into the larger context of the solution as built for this document.

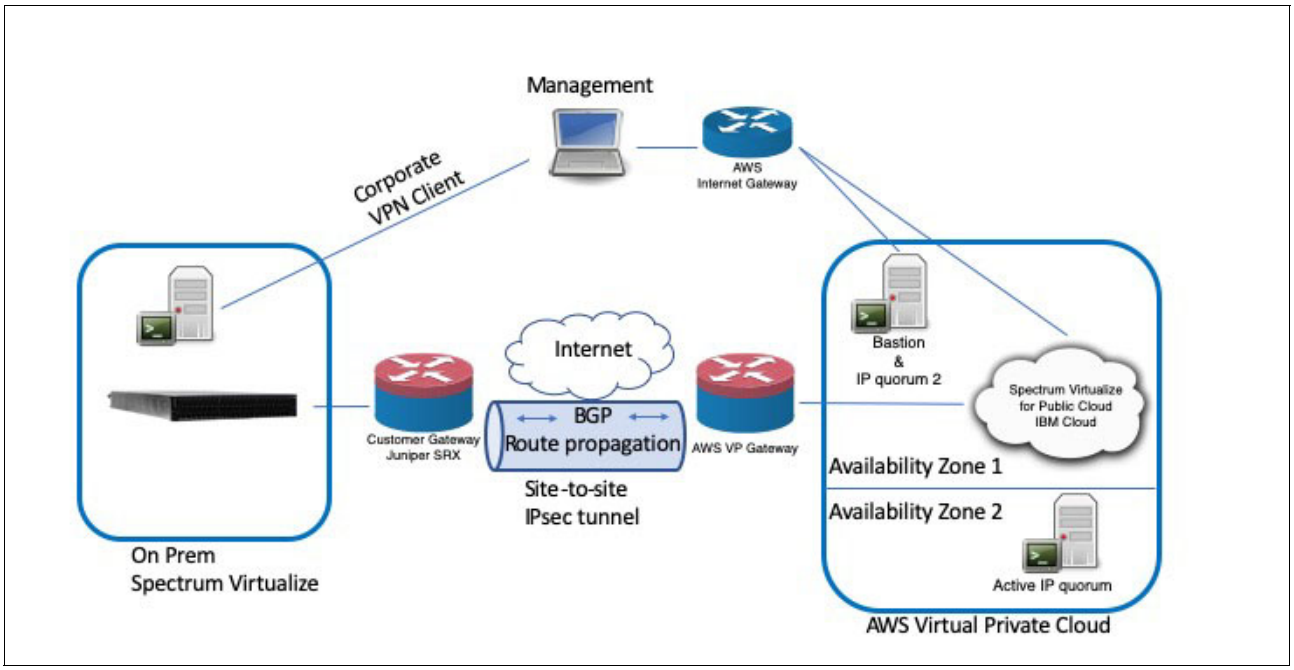


Figure 3-4 Connectivity between on-premises and Amazon AWS and management connections

In support of replication connectivity and the Transparent Cloud Tiering (TCT) function, internet access, and a site-to-site IPSec tunnel were added to the configuration. The site-to-site tunnel also provides an alternative method for managing the environment versus the use of Bastion as a port-forwarder for the IBM Spectrum Virtualize GUI (which is not recommended from a security standpoint) or the complex process of setting up a *client virtual private network (VPN) Endpoint* in AWS (which requires configuring a certificate authority, and then, transferring certificates to the AWS Certificate Manager or integrating Active Directory for authentication).

3.3.2 Objective

The design of the solution for this publication was intended to show two key features of IBM Spectrum Virtualize for Public Cloud: Easy Tier and replication.

Easy Tier

Easy Tier with thin provisioning provides a compelling business justification for the use of IBM Spectrum Virtualize in Public Cloud, especially for AWS. With thin provisioning, capacity and performance can be extended for Amazon EBS volumes, which allows applications to achieve the same level of performance, but with lower cost than without IBM Spectrum Virtualize. For the purposes of this example, two out of the four available Amazon EBS types are blended into a single Easy Tier pool. For the enterprise or nearline tier, *st1* is used and the flash tier is made up of *gp2* volumes. The four available types of Amazon EBS volumes are compared in Table 3-2.

Table 3-2 Amazon EBS storage types

Items	**Solid-state drives (SSDs)**		**Hard disk drives (HDDs)**	
Volume type	General-purpose SSD (gp2)*	Provisioned IOPS SSD (io1)	Throughput-optimized HDD (st1)	Cold HDD (sc1)
Description	General-purpose SSD volume that balances price and performance for various workloads.	Highest-performance SSD volume for mission-critical, low-latency, or high-throughput workloads.	Low-cost HDD volume that is designed for frequently accessed, throughput-intensive workloads.	Lowest cost HDD volume that is designed for less frequently accessed workloads.
Use cases	► Recommended for most workloads. ► System boot volumes. ► Virtual desktops. ► Low-latency interactive apps. ► Development and test environments.	► Critical business applications that require sustained IOPS performance, or more than 16,000 IOPS or 250 MiBps of throughput per volume. ► Large database workloads, such as: – MongoDB – Cassandra – Microsoft SQL Server	► Streaming workloads requiring consistent, fast throughput at a low price. ► Big data. ► Data warehouses. ► Log processing. ► Cannot be a boot volume.	► Throughput-oriented storage for large volumes of data that is infrequently accessed. ► Scenarios where the lowest storage cost is important. ► Cannot be a boot volume.
API name	gp2	io1	st1	sc1

Replication

Data mobility is firmly established as a foundational use case for block storage virtualization and the cornerstone of the IBM Spectrum Virtualize product. Replication between on-premises and AWS is the use case for the network diagram that is shown in Figure 3-4 on page 38.

3.3.3 Considerations

Consider the following important points when this solution is implemented:

- Use only the first three categories (and not the *sc1* class storage because of the high latency of that storage class).
- Because of the nature of Amazon EBS storage provisioning to EC2 instances, any single Amazon EBS volume is provisioned only to one of the two IBM Spectrum Virtualize nodes. However, the forwarding layer within the IBM Spectrum Virtualize software uses the custom tags that are attached to the Amazon EBS volumes to provide seamless handling of failover events.
- To ensure seamless handling of failover events, a 16 Amazon EBS volume limit per I/O group or IBM Spectrum Virtualize node pair is enforced.
- The Cloud Formation Template (CFT) that governs the installation of IBM Spectrum Virtualize for Public Cloud on AWS implements the Bastion host and IP quorum server in the same availability zone for installations that are requesting a new VPC, and installations into an existing VPC allows the installer to choose a different availability zone for the Bastion and IP quorum server. In installations that include the creation of a VPC, it is a best practice that an extra subnet is provisioned in a different availability zone and a secondary IP quorum server is started in that subnet and made the active quorum device.

4

Planning and preparing for the IBM Spectrum Virtualize for Public Cloud on AWS deployment

This chapter describes the planning considerations for implementing an IBM Spectrum Virtualize for Public Cloud on Amazon Web Services (AWS) deployment solution.

This chapter includes the following topics:

4.1 Introduction

This chapter describes the planning and preparation steps to provision network, server, and storage components on AWS. These components are required for installing IBM Spectrum Virtualize for Public Cloud on AWS.

Background information about the AWS cloud networking architecture and storage offerings is also described to help the reader who is unfamiliar with the AWS cloud plan for the IBM Spectrum Virtualize for Public Cloud placement into the larger context of an application environment.

4.2 General planning introduction

To realize the most benefit from IBM Spectrum Virtualize for Public Cloud on AWS, preinstallation planning must include several important steps, as described in this section.

4.2.1 Prerequisites for AWS

Before you install IBM Spectrum Virtualize for Public Cloud software from the AWS Marketplace, ensure that you complete the following tasks on the AWS site:

1. Sign up for AWS.
2. Create an AWS Identity and Access Management (IAM) administrator profile.
3. Assign the appropriate rules for installation and usage.
4. Create a key pair.

You can use the default AWS administrator profile to install the IBM Spectrum Virtualize for Public Cloud software or you can create an installer user profile that includes only the required permissions for deploying the software. Creating a second user for only monitoring is a best practice. To create those two extra users, complete the following steps:

1. Create a suitable user profile for the installer and the monitoring user.
2. Create one user for installation and one for monitoring, and assign the suitable user profile.

Creating an AWS user profile

To create an installer user profile, complete the following steps:

1. Log on to the AWS Management Console with the AWS default administrator profile.
2. Select **Services** in the upper left and click **IAM** to open the Identity and Access Management console.
3. In the **Navigation** pane, select **Policies** → **Create policy**.
4. Click the **JSON** tab and add the JSON content that is shown in Example 4-1.

Example 4-1 User profile for an installer user

```
{
   "Version": "2012-10-17",
   "Statement": [
       {
           "Sid": "VisualEditor0",
           "Effect": "Allow",
```

```
            "Action": [
                "aws-marketplace:ViewSubscriptions",
                "aws-marketplace:Unsubscribe",
                "aws-marketplace:StartBuild",
                "aws-marketplace:ListBuilds",
                "aws-marketplace:Subscribe",
                "iam:CreateInstanceProfile",
                "cloudformation:CreateUploadBucket",
                "sns:DeleteTopic",
                "iam:RemoveRoleFromInstanceProfile",
                "iam:CreateRole",
                "cloudformation:UpdateTerminationProtection",
                "s3:CreateBucket",
                "sns:ListTopics",
                "sns:Unsubscribe",
                "iam:PutRolePolicy",
                "iam:AddRoleToInstanceProfile",
                "iam:PassRole",
                "cloudformation:DescribeStackEvents",
                "ssm:DescribeParameters",
                "iam:DeleteRolePolicy",
                "cloudformation:UpdateStack",
                "sns:Subscribe",
                "s3:DeleteObject",
                "s3:DeleteBucket",
                "cloudformation:ListStackResources",
                "iam:DeleteInstanceProfile",
                "iam:GetRole",
                "cloudformation:ListStacks",
                "iam:GetInstanceProfile",
                "sns:GetTopicAttributes",
                "cloudformation:DescribeStackResources",
                "sns:CreateTopic",
                "iam:ListRoles",
                "iam:DeleteRole",
                "ssm:GetParameters",
                "iam:ListInstanceProfiles",
                "cloudformation:GetTemplateSummary",
                "cloudformation:DescribeStacks",
                "s3:PutObject",
                "s3:GetObject",
                "cloudformation:GetStackPolicy",
                "s3:ListAllMyBuckets",
                "cloudformation:CreateStack",
                "cloudformation:GetTemplate",
                "cloudformation:DeleteStack",
                "ec2:*",
                "cloudformation:ListChangeSets"
            ],
            "Resource": "*"
        }
    ]
}
```

5. Click **Review Policy** and add a name for the policy, such as `SV_install_policy`.

6. Click **Create Policy**.

To create a user profile with limited permissions, repeat steps 1 - 6 but use the JSON content that is shown in Example 4-2 when you create the customized policy.

Example 4-2 User profile for a monitoring user

```
{
   "Version": "2012-10-17",
   "Statement": [
       {
           "Sid": "VisualEditor0",
           "Effect": "Allow",
           "Action": [
               "ec2:RebootInstances",
               "iam:GetRole",
               "ec2:Describe*",
               "ec2:StartInstances",
               "iam:ListRoleTags",
               "iam:ListAttachedRolePolicies",
               "iam:ListRoles",
               "iam:ListPolicies",
               "ec2:StopInstances",
               "iam:ListRolePolicies",
               "iam:ListInstanceProfiles",
               "iam:GetRolePolicy",
               "ec2:Get*"
           ],
           "Resource": "*"
       }
   ]
}
```

7. Click **Review Policy** and add a name for the policy, such as `SV_monitor_policy`. Click **Create Policy**.

Creating an AWS user and assigning the appropriate profile

To create two users for installation and monitoring, complete the following steps twice:

1. Log on to the AWS Management Console with the AWS default administrator profile.
2. Select **Services** in the upper left and click IAM to open the Identity and Access Management console.
3. In the **Navigation** pane, select **Users** → **Add user**.
4. Enter a name and password, and ensure that you select AWS Management Console access for the **Access type**. Click **Next: Permissions**.
5. Select **Attach existing policies directly** and select the policy that you created in “Creating an AWS user profile” on page 42. Click **Next: Tags**.
6. Ensure that you add a tag that includes the email address for the installer user profile.

A successful login that uses the new user ID requires the login link that is provided by the AWS email, which must be manually sent during the creation process.

Because these steps might change, see IBM Knowledge Center for any new or changed information.

4.2.2 Prerequisites for IBM Spectrum Virtualize for Public Cloud

The IBM Spectrum Virtualize for Public Cloud on AWS software is a Buy Your Own License (BYOL) offering in AWS Marketplace. During the deployment, the installation template verifies the proof of entitlement that indicates that a valid license is purchased from IBM. If the proof of entitlement is not present, the installation fails. To obtain the license and proof of entitlement for the software, complete the following steps:

1. Go to the IBM Passport Advantage website to obtain a license and proof of entitlement for the software.
2. At the website, follow the directions to enter your IBM customer number and the maximum number of terabytes of virtual storage to provision with your systems.

4.3 Requirements and limitations

The installation is available on AWS Marketplace and uses AWS CloudFormation service to simplify provisioning of AWS resources. The IBM Spectrum Virtualize for Public Cloud installation provides two templates to provision and configure the required AWS services. One template installs the software on a new virtual private cloud (VPC) and the other template is used for installations on an existing VPC.

When the installation template is started from AWS Marketplace, the user is prompted to provide information, such as a customer ID for the entitlement check. For more information, see IBM Knowledge Center.

IBM Spectrum Virtualize for Public Cloud on AWS provides the following set of storage functions, which are similar to the functions that are provided by IBM Spectrum Virtualize for on-premises installations:

- IP-based Copy Services are available for the following types of replication:
  - Global Mirror
  - Metro Mirror
  - Global Mirror with Change Volumes (GMVC)
- Replication is possible between components:
  - On-premises SAN Volume Controller, IBM FlashSystem, or IBM Spectrum Virtualize as software only on Bare Metal Servers to AWS Cloud
  - IBM Spectrum Virtualize for Public Cloud on AWS instances that are deployed into two different availability zones
- FlashCopy, IBM Easy Tier, and thin provisioning are supported by IBM Spectrum Virtualize for Public Cloud on AWS.

Consider the following scalability-related limitations:

- Two or four nodes per cluster only.
- IPv4 only (no IPv6).
- A total of 20 Amazon Elastic Block Store (EBS) volumes per I/O group.
- The general maximum Amazon EBS Volume size is 16 TiB.

Multiple Availability Zones are supported through only a separate Global Mirror instance of IBM Spectrum Virtualize for Public Cloud.

The following features are unsupported in this release:

- Stretched Cluster
- HyperSwap
- Real-time Compression
- IBM Spectrum Virtualize native encryption
- Hot Spare Node (not applicable to cloud)
- DRAID and encrypted DRAID (not applicable to cloud)
- N_Port ID virtualization (NPIV) (not applicable to cloud)
- SCSI Unmap for host and back end

Note: IBM Spectrum Virtualize for Public Cloud on AWS is configured by the AWS time server by using underlying operating system methods. Changing the time server or setting a static time is not a best practice and might cause difficulties. For more information about the AWS time server, see Setting the Time for Your Linux Instance.

Consider the following points:

- Configuration requirement: The Amazon EBS volumes and the Elastic Compute Cloud (EC2) instance to which it attaches must be in the same Availability Zone.
- Configuration best practice: The EC2 instance that is acting as the IP quorum should be in another Availability Zone in the same VPC.

For more information, see IBM Knowledge Center.

4.4 Amazon Web Services resources

Multiple resources are required for IBM Spectrum Virtualize for Public Cloud on AWS. Each IBM Spectrum Virtualize node requires one EC2 server instance. A single EC2 instance is required for the IP quorum device. Amazon EBS storage devices are used as IBM Spectrum Virtualize managed disks (MDisks).

After the CloudFormation template completes the installation, the result is a fully configured 2-node cluster in a private network with two MDisks. In addition, the installation process performs the following tasks:

- Validate the entitlement for an IBM Spectrum Virtualize for Public Cloud purchase (the client provides a customer number).
- Configure all IPs (cluster IP, service IPs, node IPs, and iSCSI port IPs).
- Configure NTP and DNS with AWS internal servers.

The security rules are automatically configured according to AWS requirements:

- For access to services that are provided by the IBM Spectrum Virtualize for Public Cloud system (web GUI, SSH, and iSCSI)
- For IBM Spectrum Virtualize inter-node communication, including network failover (IP)
- For IBM Spectrum Virtualize for Public Cloud to manage Amazon EBS
- One IP quorum client is configured on a third EC2 instance (a Bastion host)

At least 11 IP addresses are required for a single IBM Spectrum Virtualize for Public Cloud on AWS installation in a VPC:

- Two node IP addresses per node
- Two port IP addresses per node
- One service IP address per node
- One IBM Spectrum Virtualize for Public Cloud on AWS cluster IP address

4.4.1 Amazon EC2 instances

The network bandwidth, the number of vCPUs, and the amount of memory are determined by instance type. The AWS instances (C5.4xlarge, C5.9xlarge, and C5.18xlarge) are the available options in the first release. The technical specifications are shown on Table 4-1. Dedicated Hosts mode is not supported for the first release.

Table 4-1 Amazon AWS EC2 on-demand resources

EC2 instance	vCPU	Memory (GiB)	Dedicated Amazon EBS bandwidth (Mbps)	Network performance (Gbps)
c5.4xlarge	16	32	3.500	Up to 10
c5.9xlarge	36	72	7.000	10
c5.18xlarge	72	144	14.000	25

Note: These specifications were valid at the time of writing (May 2019).

For more information about available Amazon EC2 instances and pricing, see the following resources:

- Amazon EC2 Instances Types
- Amazon EC2 Pricing

Note: Selection of the EC2 instances should be done carefully because no EC2 node exchange or upgrade process is in place. A mix of different EC2 instances in the same IBM Spectrum Virtualize cluster is not supported in the first release. Migration to a different hardware platform can be done by replication to a new cluster.

4.4.2 Amazon Elastic Block Store

All Amazon EBS volume types are designed for 99.999% availability. They fall into the following categories:

- Solid-state drive (SSD)-based volumes that are optimized for transactional workloads with a small I/O size. The dominant performance attribute is IOPS.
- Hard disk drive (HDD)-based volumes that are optimized for streaming workloads, which are measured in MiBps.

Different volume types are available for Amazon EBS. They differ in performance characteristics, as listed in Table 4-2.

Table 4-2 EBC volume types

Item	SSDs		HDDs	
Volume type	General-purpose SSD	Provisioned IOPS SSD	Throughput-optimized HDD	Cold HDD
API name	gp2	io1	st1	sc1
Max IOPS / volume	16,000	64,000	500	250
Max throughput / volume in MiBps	250	1,000	500	250

The following Amazon EBS volume types are recommended:

- General-purpose SDSD (gp2)
- Provisioned IOPS SSD (io1)
- Throughput-optimized(st1)

Note: The Cold HDD class of storage is *not* recommended for use with Spectrum Virtualize for Public Cloud on AWS.

All volume types appear as “Enterprise Disks” tier in IBM Spectrum Virtualize for Public Cloud and the tier level should be adapted afterward according to their capabilities.

General-purpose SSD (gp2) volumes are the default volume type for Amazon EBS volumes that are created from the console. The gp2 volumes have a throughput limit of 128 MiBps - 250 MiBps, depending on volume size. These volumes earn I/O credits at the baseline performance rate of 3 IOPS per GiB of volume size. For example, a 100 GiB gp2 volume has a baseline performance of 300 IOPS. When a volume below 1 TiB size requires more than the baseline performance I/O level, it draws on I/O credits in the credit balance to burst to the required performance level. For more about the Amazon EBS volume types, the I/O credits, and pricing, see the following resources:

- Amazon EBS Volume Types
- Amazon EBS Pricing

4.4.3 Amazon Web Services cost estimation

THe AWS cost depends on the following factors:

- Bandwidth
- Virtual CPUs and memory
- Storage capacity and performance
- Duration of usage

Amazon Total Cost of Ownership (TCO) helps with AWS cost estimation.

4.5 Network and security

AWS uses a *shared responsibility model* where AWS provides a global secure infrastructure and services. AWS customers are responsible for protecting the confidentiality, integrity, and availability of their data in the cloud, and fulfillment of specific business requirements. Careful planning is required for the network environment to fulfill future scalability and performance requirements.

Important: Involve the customer network architect and the AWS architect in the early phases of planning to ensure a successful implementation.

For more information about how AWS keeps your data safe and how AWS meets compliance requirements, see AWS Cloud Security.

For more information about Amazon Security best practices, see *AWS Security Best Practices*.

4.5.1 Data security

Data security and protection can be achieved by encrypting the following types of data:

- Data in motion
- Data in use
- Data at rest

Native encryption is not supported by IBM Spectrum Virtualize for Public Cloud on AWS in the first release. However, Amazon EBS volumes can be ordered as encrypted or non-encrypted. Use of unencrypted and Amazon EBS volumes that are encrypted with the `(default) aws/ebs` Master key are supported. Amazon EBS volumes that are encrypted with the AWS `Adminkey` are not supported. These `Adminkey` encrypted volumes generate an error and fail when a user attempts to add them to a pool.

The data-at-rest encryption occurs on the servers that host EC2 instances running the IBM Spectrum Virtualize nodes, which provides encryption of data as it moves between EC2 instances and the encrypted Amazon EBS volumes. For more information, see the following resources:

- Amazon EBS features
- AWS Identity and Access Management

AWS has a full range of security IaaS. Software installation and system management must be integrated under that infrastructure consistently and securely. Three different user types with the suitable IAM roles are available that are required for successful installation and system management, as described in the following sections.

Installer

To install the IBM Spectrum Virtualize for Public Cloud from AWS marketplace, an Installer user profile should be predefined manually in the AWS IAM service. Any user can be used as the Installer if the user meets the minimum privileges of the Installer profile. Six service-related permissions are needed to deploy an IBM Spectrum Virtualize for Public Cloud cluster on the AWS cloud.

User

A user profile can be defined based on your own IT security policy. However, it is a best practice to limit the permissions of these users to actions that they complete as part of their daily work.

IP quorum management

IP quorum management requires permissions to access quorum-related actions.

For more information about planning an installation on AWS, see IBM Knowledge Center.

4.6 Storage performance optimization

IBM Spectrum Virtualize for Public Cloud assigns workloads to MDisks according to their physical capabilities. Those capabilities must be set manually for external MDisks, such as Amazon EBS volumes.

Easy Tier is a solution that you can use to optimize the most valuable storage usage and maximize Cloud Block Storage performance. Those settings are used on Easy Tier for hot extent relocation and optimal performance. By selecting the suitable tier for the Amazon EBS volume, IBM Spectrum Virtualize can use the back-end MDisk according to its capabilities, and not underdrive or overdrive the volume.

By default, all MDisks appear in IBM Spectrum Virtualize for Public Cloud on AWS as “Enterprise” tier with the “default” **easytierload** (medium), as shown in Example 4-4.

Example 4-3 Verifying the current MDisk tier settings

```
IBM_IBM Spectrum_Virtualize:REDBOOKS-SV4PC:superuser>lsmdisk 0 | grep tier
tier tier_enterprise
easy_tier_load
IBM_IBM Spectrum_Virtualize:REDBOOKS-SV4PC:superuser>I
```

The assignment to the suitable Easy Tier level that is shown in Table 4-3 is a best practice and must be adjusted manually by using the **chmdisk -tier -easytierload** command.

Table 4-3 AWS assignment to Easy Tier level

Drive	AWS volume type	Easy Tier level	easytierload command flag
Provisioned IOPS SSD	io1	tier0_flash	**high**
General-purpose SSD	gp2	tier1_flash	**low**
Throughput-optimized HDD	st1	tier_nearline	**low**

As shown in Example 4-4, you assign an Amazon EBS gp2 volume to the suitable IBM Spectrum Virtualize for Public Cloud storage tier. Finally, you verify the tier level by running the **lsmdisk** command.

Example 4-4 Changing the MDisk tier level settings

```
IBM_IBM Spectrum_Virtualize:REDBOOKS-SV4PC:superuser>chmdisk -tier tier1_flash
-easytierload low 0
IBM_IBM Spectrum_Virtualize:REDBOOKS-SV4PC:superuser>lsmdisk 0 | grep tier
tier tier1_flash
easy_tier_load low
IBM_IBM Spectrum_Virtualize:REDBOOKS-SV4PC:superuser>
```

4.7 Data Reduction Pools

Consider the following points regarding the use of Data Reduction Pools (DRPs).

- Minimum pool size (see Table 4-4)
- Pool extent size: 4 GB to accommodate a 512 TB pool.
- Maximum utilization of pool: 85% to allow maximum efficiency in space reclamation.

Table 4-4 DRP extent size, minimum, and maximum

Extent Size (in gigabytes)	Minimum (In terabytes)	Maximum per I/O group
1 GB or smaller	1.1 TB	128 TB
2 GB	2.1 TB	256 TB
4 GB	4.2 TB	512 TB
8 GB	8.5 TB	1024 TB

The minimum size indicates the space that is needed for metadata management, the maximum is a reflection of the 128,000 extent per volume limit (for more information, see 3.3, “Solution architecture” on page 38).

Be especially aware of the minimum size requirement when creating a minimum base configuration that provisions 2x512GB Amazon EBS volumes at time of cluster creation. This requirement is below the minimum for a DRP with 1 GB extents let alone the default 4 GB extent.

5

Implementation

This chapter describes how to implement an IBM Spectrum Virtualize for Public Cloud on Amazon Web Services (AWS) environment. It includes the following topics:

5.1 Implementing IBM Spectrum Virtualize for Public Cloud on Amazon Web Services

This section describes implementing IBM Spectrum Virtualize for Public Cloud on AWS. The IBM Spectrum Virtualize for Public Cloud on AWS implementation starts from the assumption that the required IBM Spectrum Virtualize Licenses were purchased and you can access IBM Passport Advantage®.

Designed for software-defined environments (SDEs), IBM Spectrum Virtualize for Public Cloud on AWS represents a solution for public cloud implementations. It also includes technologies that complement and enhance public cloud offering capabilities.

IBM Spectrum Virtualize for Public Cloud on AWS provides for the deployment of IBM Spectrum Virtualize software in public clouds in Amazon Cloud. IBM Spectrum Virtualize for Public Cloud on AWS includes a monthly license to deploy and use IBM Spectrum Virtualize for Public Cloud on AWS to enable hybrid cloud solutions, offering the ability to have storage as service in a multicloud environment.

Table 5-1 lists the IBM Spectrum Virtualize for Public Cloud on AWS that is available on Amazon Cloud.

Table 5-1 IBM Spectrum Virtualize for Public Cloud on AWS at a glance

Items	On AWS
Storage supported	Amazon Elastic Block Store (EBS)
Licensing approach	Simple, flat cost per managed terabyte and monthly licensing
Platform	IBM Spectrum Virtualize for Public Cloud on AWS installed on an Elastic Compute Cloud (EC2) instance

5.1.1 Installing IBM Spectrum Virtualize for Public Cloud on AWS

The IBM Spectrum Virtualize for Public Cloud installation uses AWS CloudFormation templates that simplify provisioning and management on AWS. These templates are available on AWS Marketplace and simplify the provisioning and installation process.

Ensure that all prerequisites are complete before you install the IBM Spectrum Virtualize for Public Cloud software from AWS Marketplace. For more information, see the AWS Marketplace web page.

Note: Before installing IBM Spectrum Virtualize for Public Cloud on AWS, ensure that the key pair (`ssh_key`) is defined and accessible after the installation is complete. This key is used to access the Bastion host, any other EC2 instances that are created, and other key-based authentication.

To install the IBM Spectrum Virtualize for Public Cloud software, complete the following steps:

1. Go to the IBM Spectrum Virtualize for Public Cloud BYOL Marketplace Offering (or search for “IBM Spectrum Virtualize” on the AWS marketplace).

   Log in to with your AWS account, as shown in Figure 5-1.

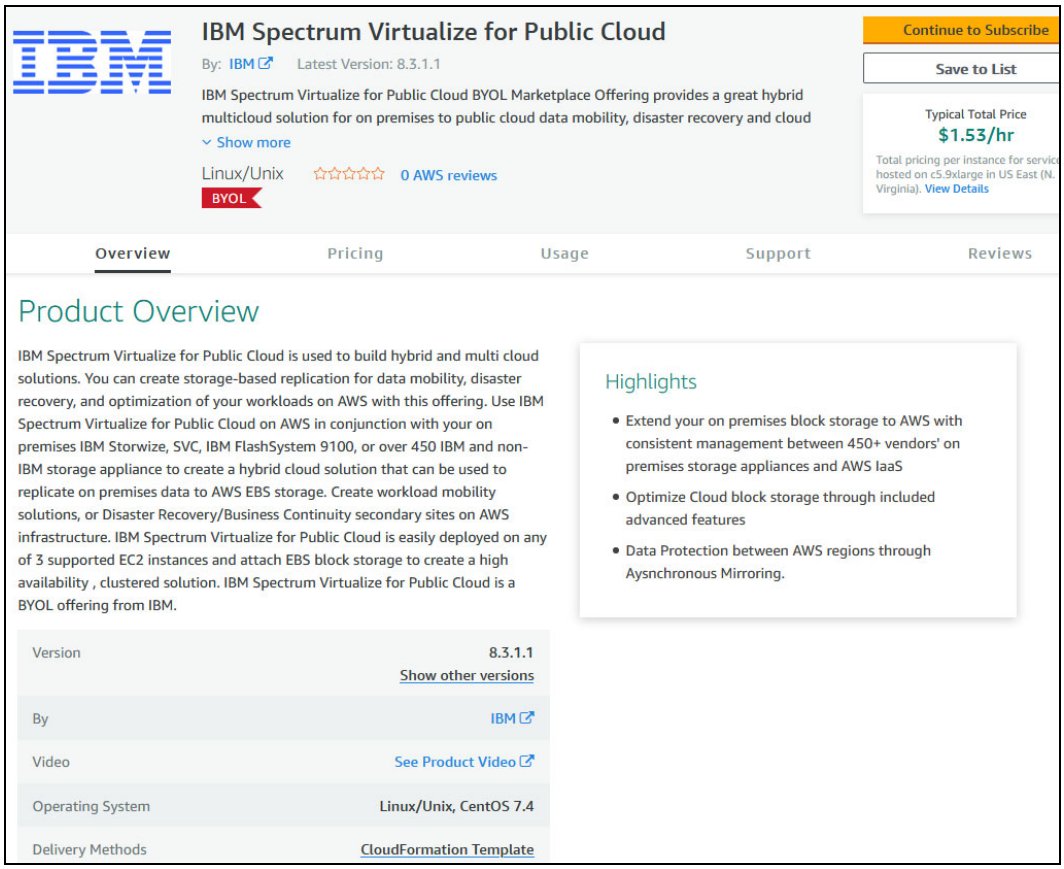


Figure 5-1 IBM Spectrum Virtualize for Public Cloud AWS Marketplace page

2. Scroll down from the **Overview** section to **Pricing** or select it from the menu items at the top of the Marketplace page. Enter or validate the following information for your installation:

   – Region
   – Fulfillment Option (by using a virtual private cloud [VPC] or a new VPC)
   – EC2 Instance type (the default is c5.9xlarge)

AWS Marketplace provides a dynamic pricing display that is based on your selections. If you are satisfied with your selections, click **Continue to Subscribe** in the upper right corner of the page and follow the instructions, as shown in Figure 5-2.

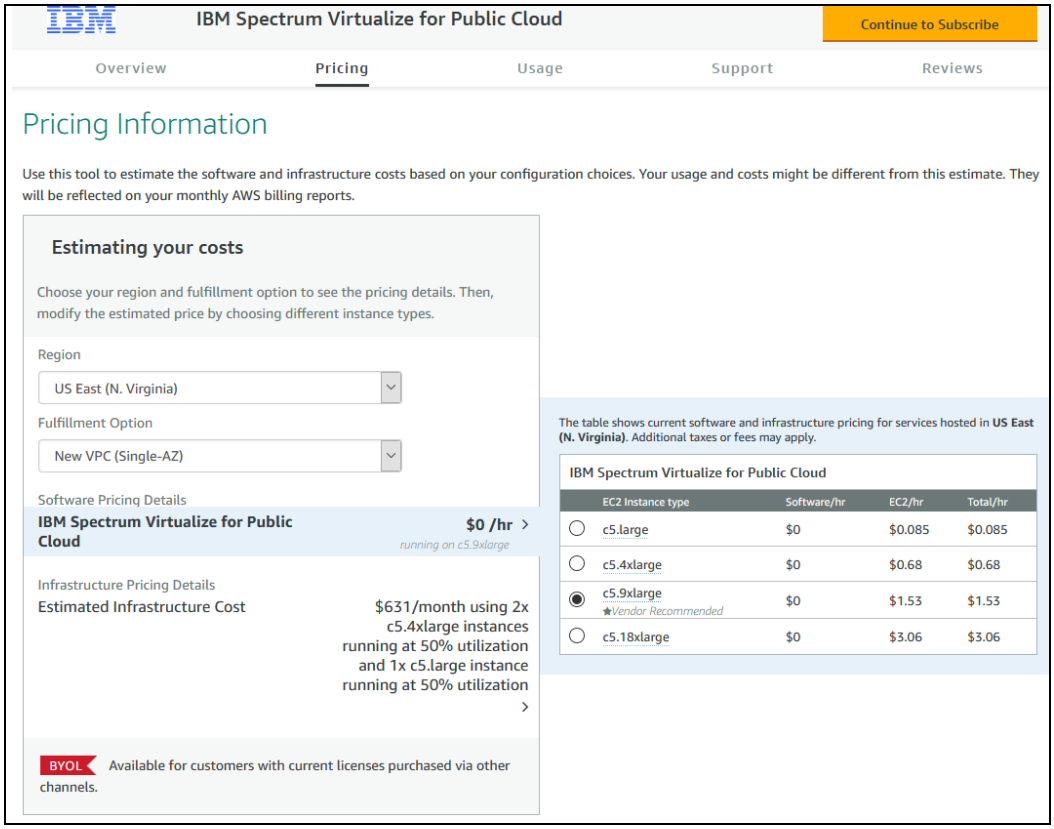


Figure 5-2 IBM Spectrum Virtualize for Public Cloud on AWS Marketplace Pricing Summary

3. The Terms and Conditions window opens and shows the Product information, as shown in Figure 5-3. After you are satisfied with the results, click **Continue to Configuration**.

Figure 5-3 IBM Spectrum Virtualize for Public Cloud AWS Marketplace Terms and Conditions

4. Specify whether a New VPC is requested or if you want to deploy into an existing VPC. If the deployment in New VPC option is selected, it carries the reminder that only a single Availability Zone is provided by default. Therefore, the Bastion host and the initial IP quorum are placed in the same Availability Zone as the IBM Spectrum Virtualize nodes.

   We elaborate on the significance of this setup and the steps to remediate it by adding another subnet in a different Availability Zone, starting a private network-only EC2 instance into that new Availability Zone, and installing the IP quorum application on that server. The software version also can be changed if others are available, and the region of AWS for the deployment.

   In our example, select the existing VPC option. After all of the options are finalized, click **Continue to Launch** (see Figure 5-4).

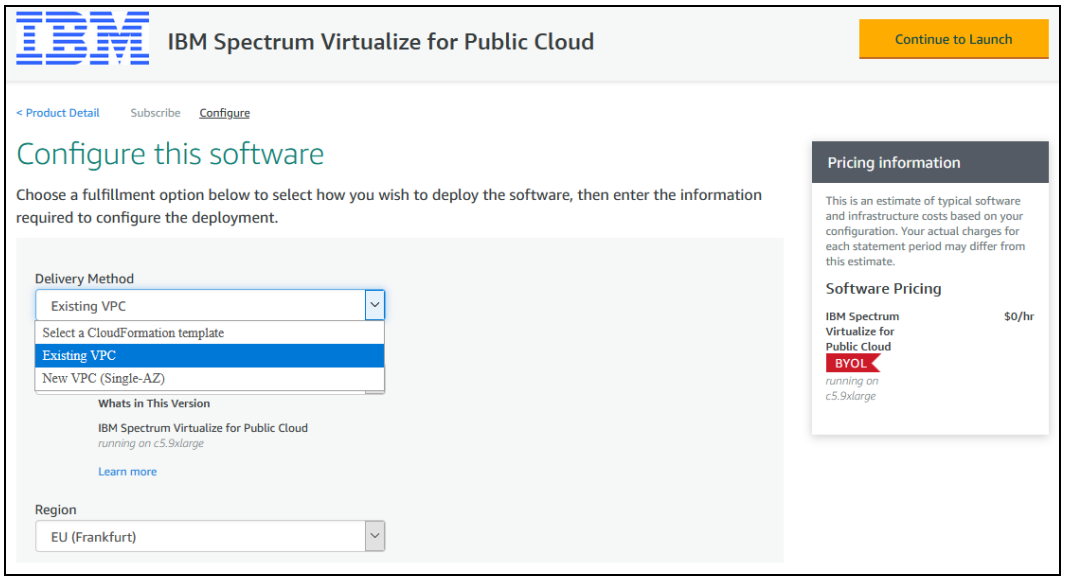


Figure 5-4 Continue to Launch

5. The CloudFormation template opens, which automates the rest of the installation process after some key parameters are entered. Because the default action is to launch the CloudFormation process, click **Launch** (see Figure 5-5).

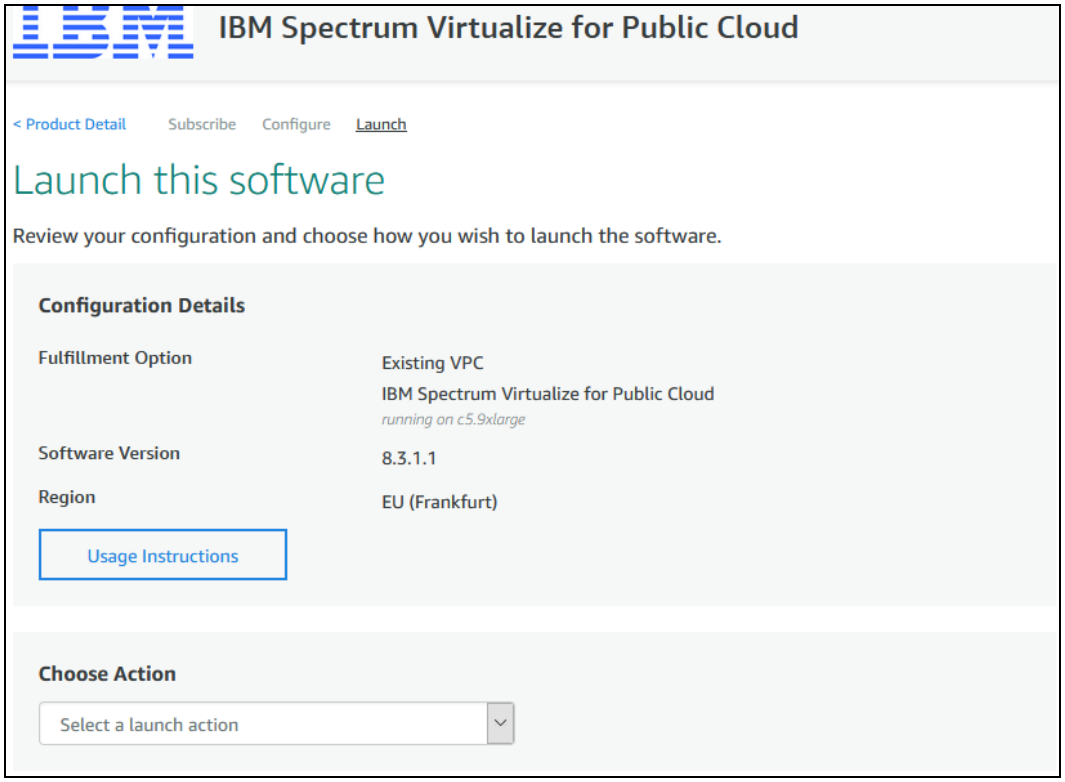


Figure 5-5 Launching the CloudFormation templates

6. The stack must be created next. Use the default settings and do not change the Amazon Simple Storage Service (S3) URL. This template location is provided by IBM Spectrum Virtualize for Public Cloud and contains critical information for installation automation. Click **Next** (see Figure 5-6).

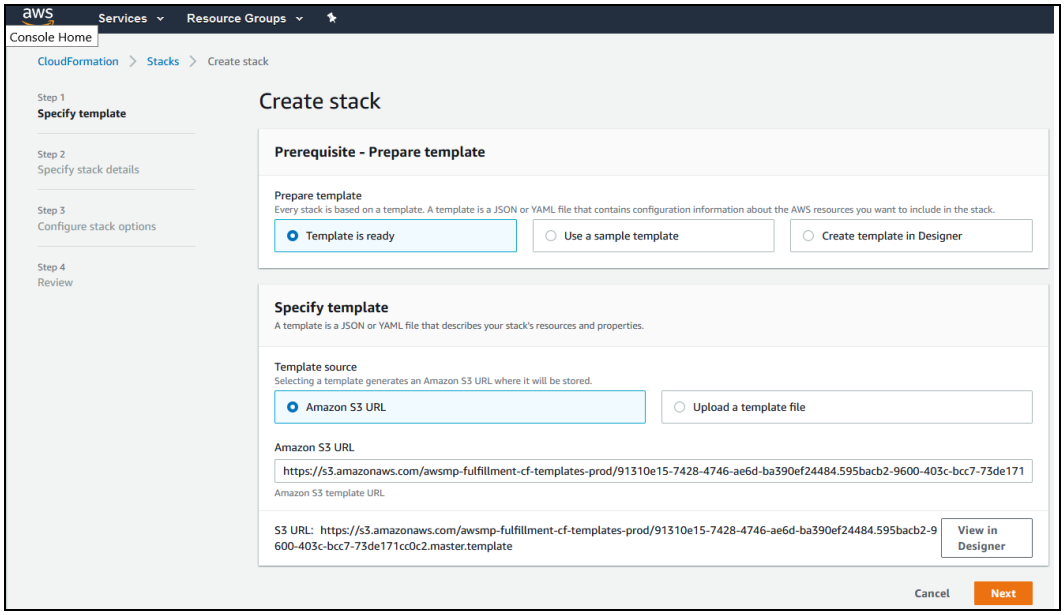


Figure 5-6 Starting the CloudFormation stack creation process

7. Enter the stack name that is the basis of the IBM Spectrum Virtualize cluster or system name. Specify the availability zone and modify the network parameters as needed. Select the existing VPC ID and CIDR block (in our example, it is `172.16.0.0/16` or `172.16.0.0-172.16.255.255`). The default settings are suitable for public and private clouds, especially in a new VPC (see Figure 5-7).

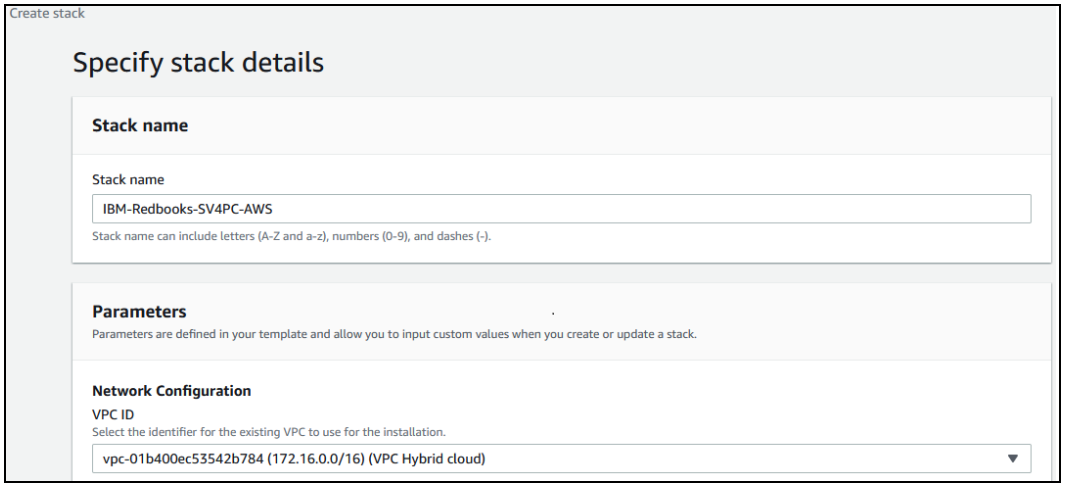


Figure 5-7 CloudFormation Stack: Name

8. Set the access filter as needed. For this example, we did not restrict which IPs were allowed to access `sv_cloud`. The EC2 server size for the IBM Spectrum Virtualize nodes and the Bastion host also can be changed if needed, as shown in Figure 5-8.

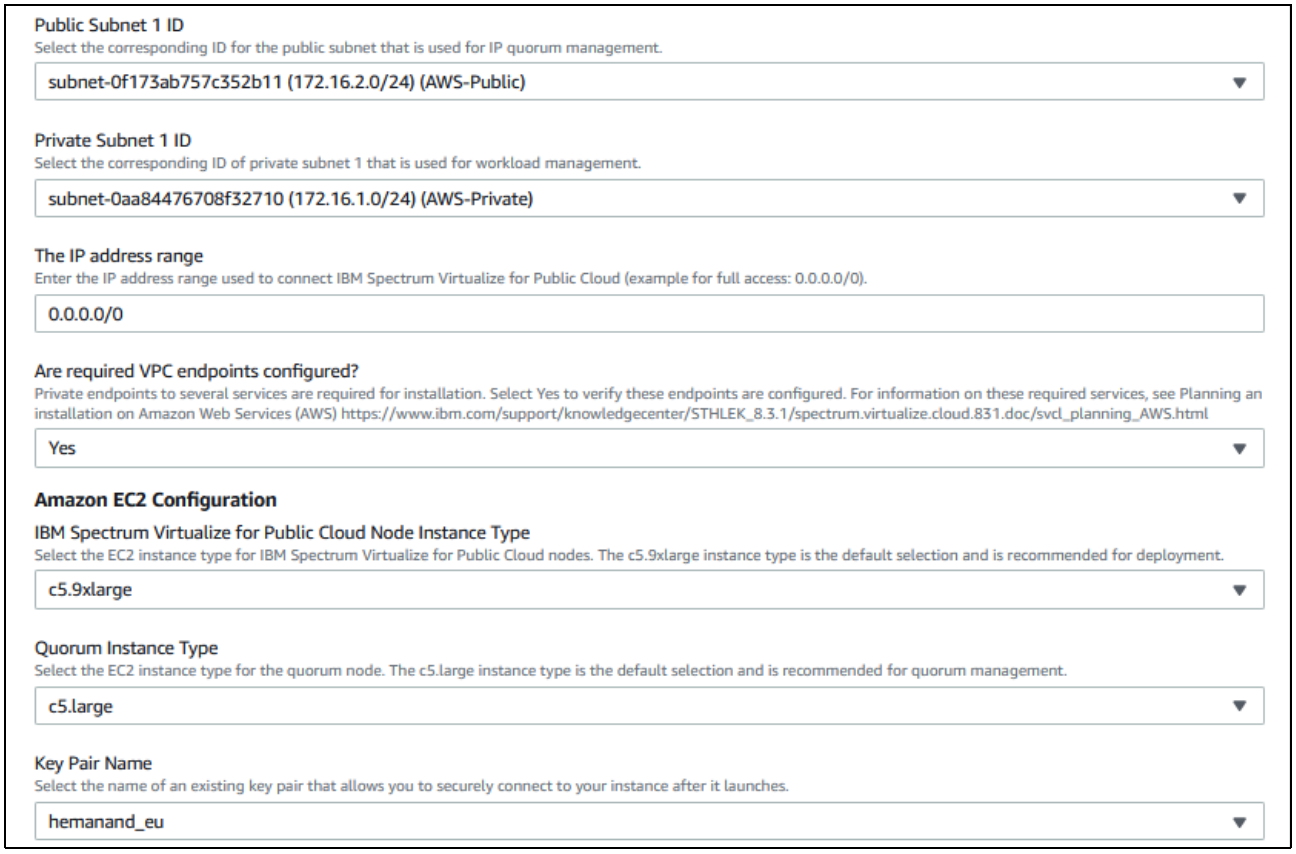


Figure 5-8 Access filter configuration

9. In the configuration page that is shown in Figure 5-9, specify the size of the two Amazon EBS gp2 volumes that are included in a pool as part of the cluster creation. Also, select the I/O group configuration, which depends on a 2-node or 4-node cluster.

Figure 5-9 CloudFormation Stack: Initialling Amazon EBS gp2 volume size

10. The summary and acknowledgment page is shown next. This page features two steps. Review your selections and edit them if necessary, as shown in Figure 5-10 and Figure 5-11.

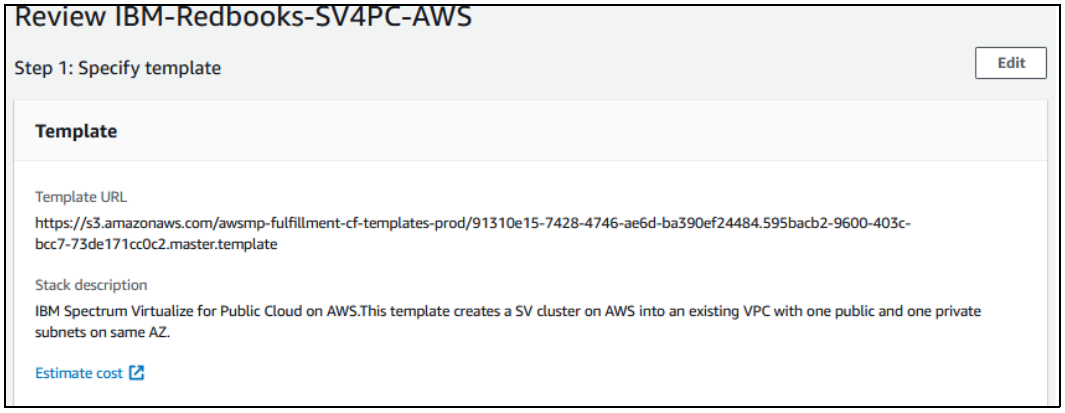


Figure 5-10 CloudFormation Stack: Use the default template

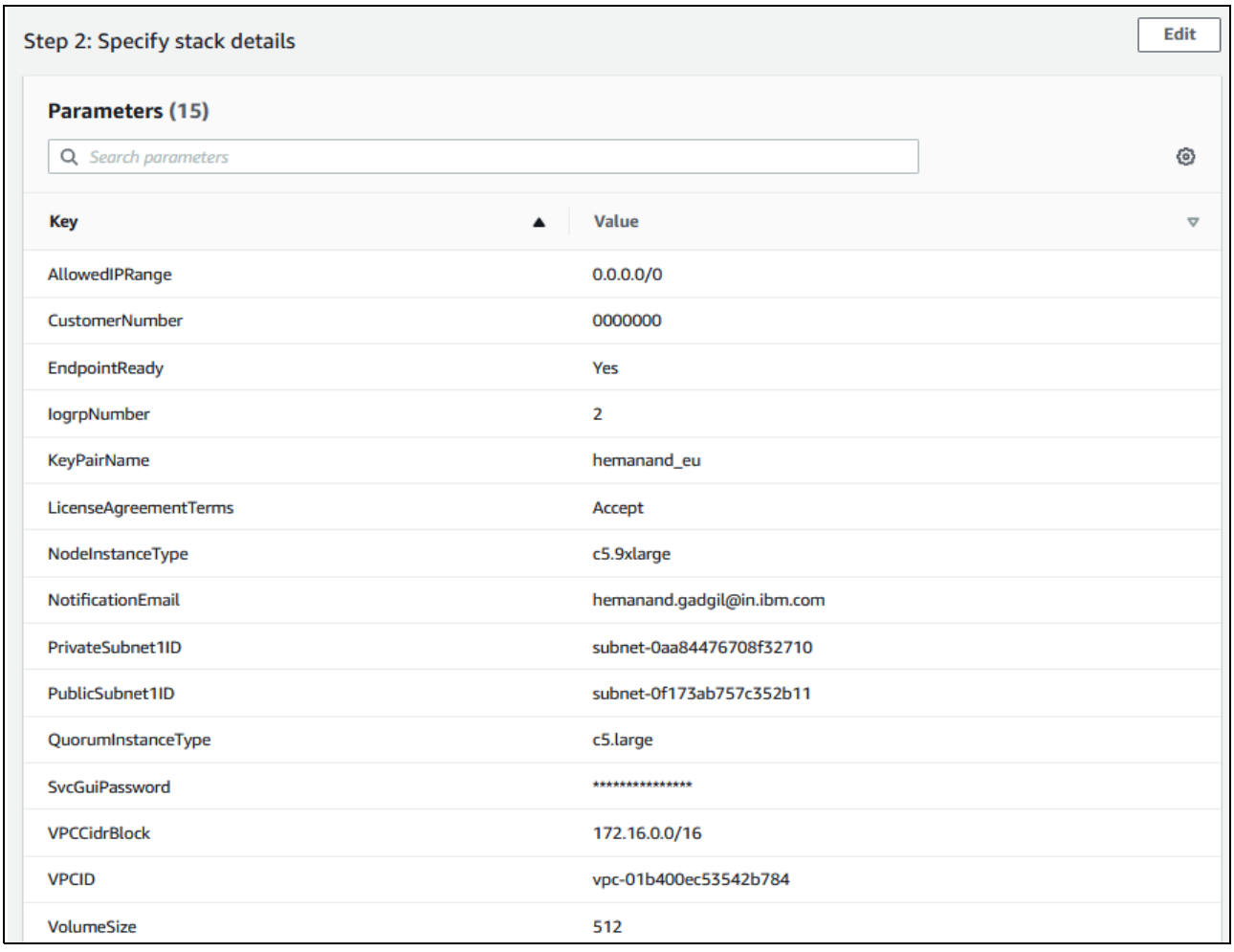


Figure 5-11 CloudFormation Stack: Review the selections

11. After you review your selections, click **Create Stack**, as shown in Figure 5-12.

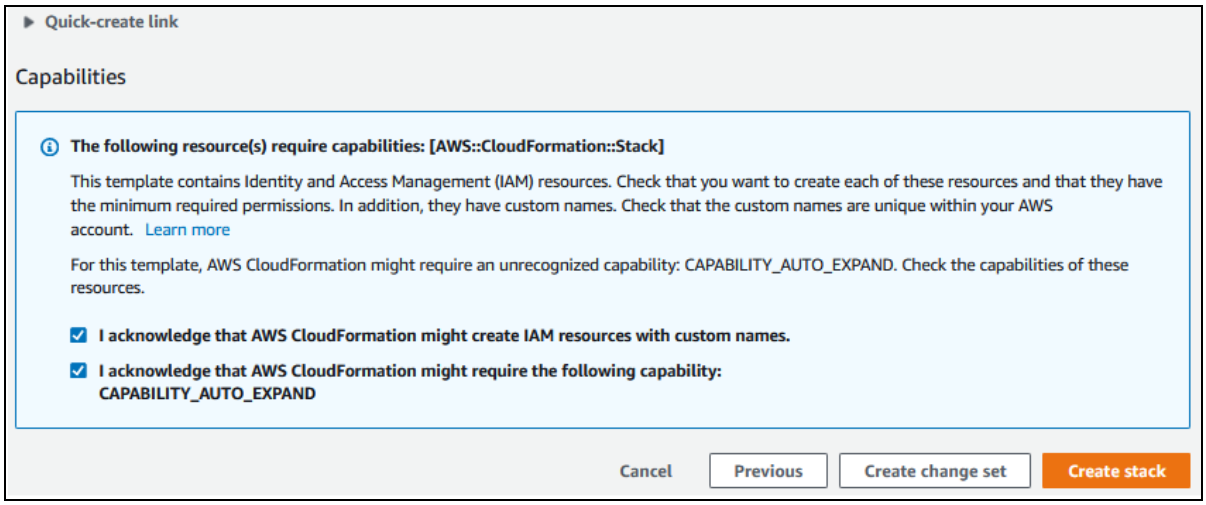


Figure 5-12 CloudFormation Stack: Create the stack

12. The stack creation process takes about 20 minutes for new VPCs and 15 minutes for existing VPCs. Progress can be monitored by going to the AWS console and selecting **CloudFormation** → **Stacks**, and then clicking the **Events** tab. After the stack and associated WorkloadStack reaches `CREATE_COMPLETE`, the environment is ready for interaction, as shown in Figure 5-13.

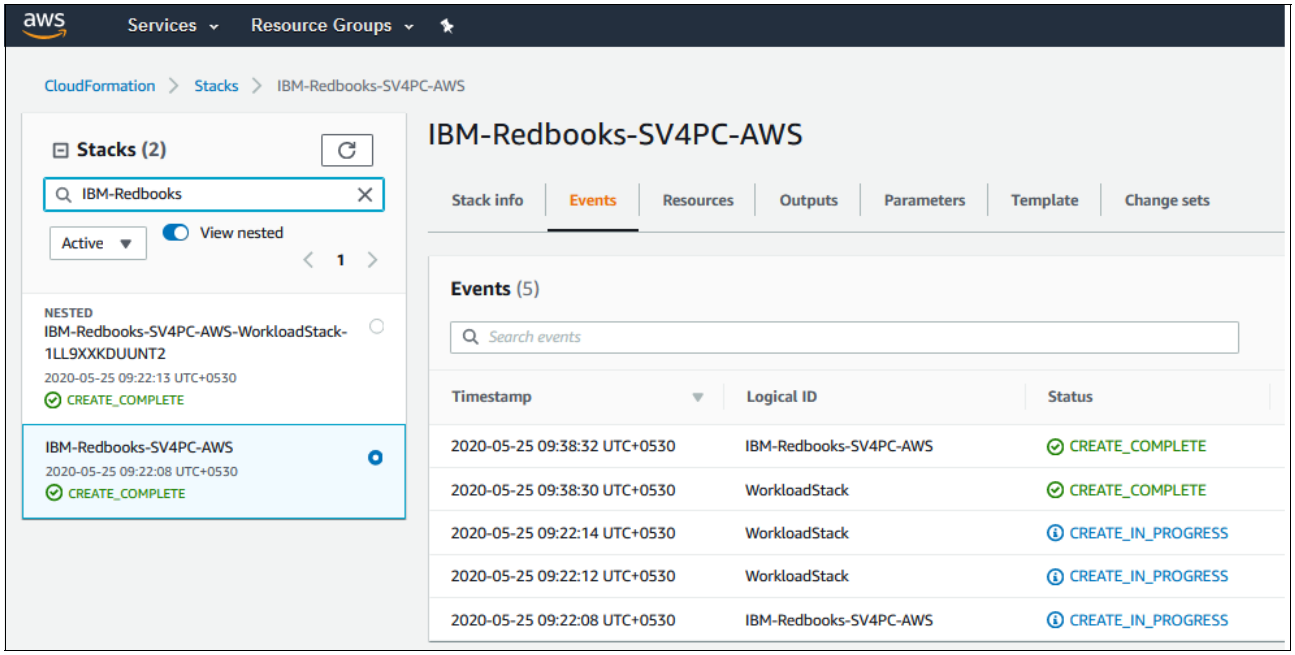


Figure 5-13 CloudFormation Stack: Creation complete

13. In this same view, you can view important IP address information by clicking the Outputs tab, as shown in Table 5-2.

Table 5-2 Output of CloudFormation auto-provisioning after AWS finishes stack creation

Names	IP address	Descriptions
IBMSVClusterIP	172.16.1.61	IBM Spectrum Virtualize Cloud Cluster IP
IBMSVNode1Port1NodeIP	172.16.1.104	IBM Spectrum Virtualize Node 1 Port 1
IBMSVNode1Port2NodeIP	172.16.1.134	IBM Spectrum Virtualize Node1 Port 2
IBMSVNode1PortIP1	172.16.1.91	IBM Spectrum Virtualize Node 1 Port IP 1
IBMSVNode1PortIP2	172.16.1.154	IBM Spectrum Virtualize Node 1 Port IP 2
IBMSVNode1ServiceIP	172.16.1.181	IBM Spectrum Virtualize Node1 Service IP
IBMSVNode2Port1NodeIP	172.16.1.241	IBM Spectrum Virtualize Node 2 Port 1
IBMSVNode2Port2NodeIP	172.16.1.40	IBM Spectrum Virtualize Node 2 Port 2
IBMSVNode2PortIP1	172.16.1.36	IBM Spectrum Virtualize Node 2 Port IP 1
IBMSVNode2PortIP2	172.16.1.236	IBM Spectrum Virtualize Node 2 Port IP 2
IBMSVNode2ServiceIP	172.16.1.193	IBM Spectrum Virtualize Node 2 Service
IBMSVNode3Port1NodeIP	172.16.1.20	IBM Spectrum Virtualize Node 3 Port 1
IBMSVNode3Port2NodeIP	172.16.1.73	IBM Spectrum Virtualize Node 3 Port 2
IBMSVNode3PortIP1	172.16.1.198	IBM Spectrum Virtualize Node 3 Port IP 1
IBMSVNode3PortIP2	172.16.1.77	IBM Spectrum Virtualize Node 3 Port IP 2
IBMSVNode3ServiceIP	172.16.1.211	IBM Spectrum Virtualize Node 3 Service
IBMSVNode4Port1NodeIP	172.16.1.59	IBM Spectrum Virtualize Node 4 Port 1
IBMSVNode4Port2NodeIP	172.16.1.173	IBM Spectrum Virtualize Node 4 Port 2
IBMSVNode4PortIP1	172.16.1.107	IBM Spectrum Virtualize Node 4 Port IP 1
IBMSVNode4PortIP2	172.16.1.94	IBM Spectrum Virtualize Node 4 Port IP 2
IBMSVNode4ServiceIP	172.16.1.119	IBM Spectrum Virtualize Node 4 Service
IBMSVQuorumClientEC2IP	172.16.2.42	IBM Spectrum Virtualize Quorum Client EC2 Private IP
IBMSVVersion	8.3.1.1	IBM SV Cloud version
StackUpdateTemplate	https://svcloud-beta.s3-us-west-2.amazonaws.com/8311/Beta/sv-cloud-node-	Template to add or remove IBM SV Cluster I/O group

5.2 Logging in to IBM Spectrum Virtualize for Public Cloud on AWS

After the CloudFormation is created, you can log in to IBM Spectrum Virtualize for Public Cloud to refine the configuration. Because this server is the only server with an externally exposed address, it features the following functions:

- SSH jump host
- GUI proxy
- Cloud Call Home gateway
- SMTP gateway (optionally)
- Remote Support Proxy (RSP) server (optionally)
- Storage Insights DataCollector host (optionally)

5.2.1 Using SSH to access the Bastion host

Use the AWS console to access the list of EC2 instances and look for an instance that starts with your stack name and ends in `IBM-SV-QuorumNode`. Above that instance are the four IBM Spectrum Virtualize nodes. Select the QuorumNode instance or Bastion host and look for the IPv4 Public IP in the Description tab, as shown in Figure 5-14.

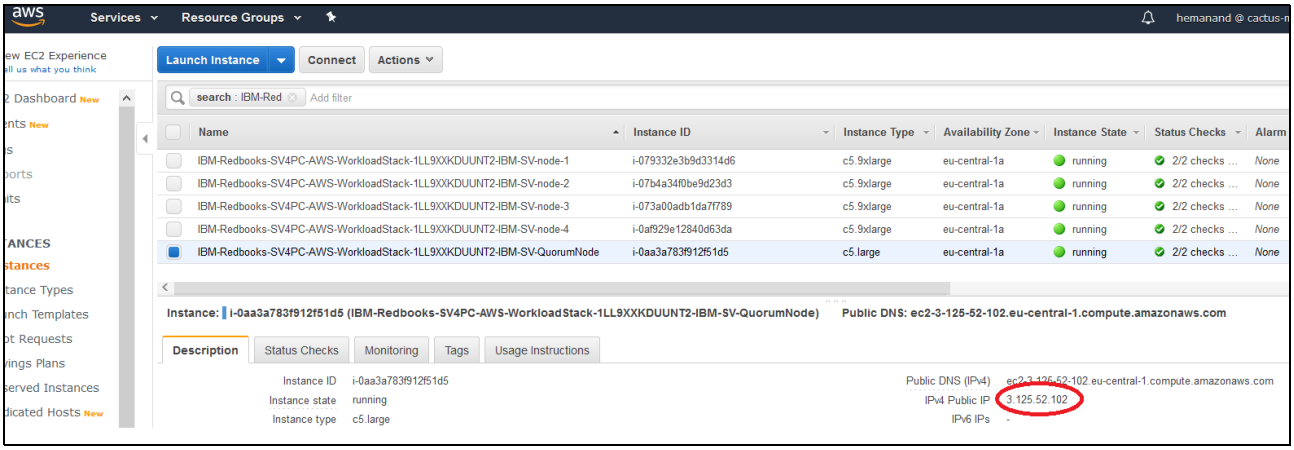


Figure 5-14 Public IP of QuorumNode (Bastion host)

By using the IP address, run **ssh** to access the Bastion host. The output is shown in Figure 5-15.

```
login as: centos
Authenticating with public key "imported-openssh-key"
Last login: Mon May 25 04:17:32 2020 from 59.95.87.202

######################## Welcome to IBM Spectrum Virtualize for Public Cloud world!
 ########################
# You can enable/disable access to IBM Spectrum Virtualize for Public Cloud mana
gement GUI by:
#  1. enable GUI: enable-sv-cloud-management-gui
#  2. disable GUI: disable-sv-cloud-management-gui
# To access IBM Spectrum Virtualize for Public Cloud management GUI, enter the a
ddress in a web browser:
#  https://3.125.52.102:8443
##############################################################################
 ########################
[centos@svpc-bastion ~]$
```

Figure 5-15 Running ssh to access the Bastion host

5.2.2 Configuring the Bastion host

To configure the Bastion host, complete the tasks that are described in the following sections.

Enabling GUI access

Run **ssh** to access the Bastion host by using the `ssh-key` that you specified during the installation, as shown in Example 5-1.

Example 5-1 SSH connection to the Bastion host to enable GUI access

```
[centos@svpc-bastion~]$ enable-sv-cloud-management-gui
```

Note: Port forwarding of port 8443, which is needed for GUI access, is disabled by default. Enable it for added security.

Configuring the Remote Support Proxy server

An *RSP* is a server that can be deployed to use the remote support assistance features that are offered in the IBM Spectrum Virtualize software. This section describes how to install the RSP server and configure the proxy in IBM Spectrum Virtualize to enable remote support connections into the cluster.

For the purposes of this publication, assume that a separate virtual server is created in the environment that can access the public network and the private network, including routes to the subnet in which IBM Spectrum Virtualize is running. Also, for this example, assume that the virtual server that is deployed is Red Hat Linux 7.x.

Complete the following steps:

1. Download the RSP software from your product support page. At the time of this writing, this code is under the Others category, as shown in Figure 5-16.

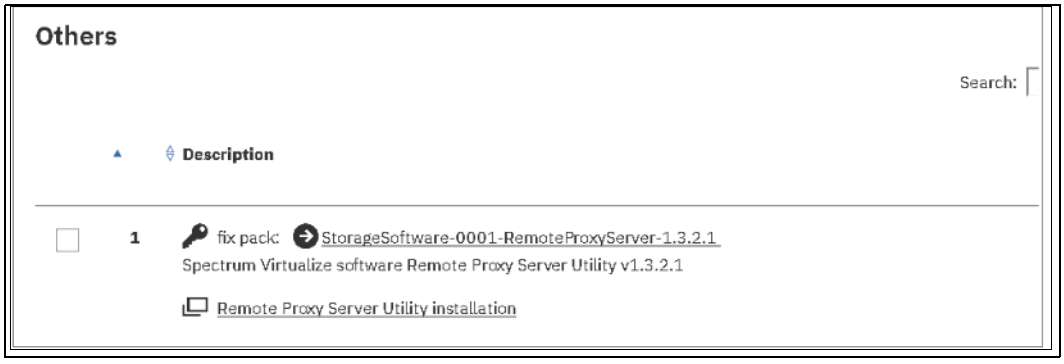


Figure 5-16 Downloading code from the product support page

2. After the code is downloaded to the administrator's notebook, you must upload the file to the server in which the proxy will be installed. To do so, run the **scp** command. You also must install the `redhat-lsb` package if it is not installed. When the file is uploaded to the server and all prerequisite packages are installed, you can proceed with the installation, as shown in Example 5-2 on page 68.

Example 5-2 Installing the Remote Support Proxy

```
[root@itso-dal10-sv-rsp ~]# chmod +x
supportcenter_proxy-installer-rpm-1.3.2.1-b1501.rhel7.x86_64.bin
[root@itso-dal10-sv-rsp ~]#
./supportcenter_proxy-installer-rpm-1.3.2.1-b1501.rhel7.x86_64.bin
Starting installer, please wait...
```

Tip: For the installation to succeed, ensure that the required packages are installed. On Red Hat systems, install the packages `redhat-lsb` and `bzip2`. On SUSE systems, install the package `insserv`.

3. When the installer is started, you see the International License Agreement for Non-Warranted Programs. To complete the installation, enter `1` to accept the license agreement and complete the installation.
4. When the installation completes, you must configure the proxy server to listen for connections. This configuration is done by editing the `supportcenter/proxy.conf` configuration file, which is in the `/etc` directory. The minimum modification that is required is to edit the fields `ListenInterface` and `ListenPort`. By default, the file has “?” as the value for both.
5. To complete the configuration, specify `ListenInterface` with the interface name in Linux that can access the IBM Spectrum Virtualize clusters. You can discover this name by running the **`ifconfig`** command and identifying the interface that accesses the AWS Cloud private network. Also, set `ListenPort` to the TCP port number to listen on for remote support requests. A sample configuration file is shown in Example 5-3.

Tip: Consider the internal address of the Bastion host (in Example 5-3, it is `10.0.32.86`). This address is available from the same AWS console view from where we retrieved the public IP, but it is useful to find it from `ifconfig` on the server. The internal IP is used for several configuration items on the IBM Spectrum Virtualize system. Also, make a note the port that is specified for `ListenPort` of the remote proxy because it is needed later in EasySetup for Support Proxy.

Example 5-3 Sample proxy configuration

```
[centos@svpc-bastion ~]$ ifconfig
eth0: flags=4163<UP,BROADCAST,RUNNING,MULTICAST>  mtu 1500
        inet 10.0.32.86 netmask 255.255.224.0  broadcast 10.93.4.127
        inet6 fe80::490:fbff:fed6:7120  prefixlen 64  scopeid 0x20<link>
        ether 06:90:fb:d6:71:20  txqueuelen 1000  (Ethernet)
        RX packets 58690  bytes 59492454 (56.7 MiB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 15492  bytes 2239603 (2.1 MiB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0

lo: flags=73<UP,LOOPBACK,RUNNING>  mtu 65536
        inet 127.0.0.1  netmask 255.0.0.0
        inet6 ::1  prefixlen 128  scopeid 0x10<host>
        loop  txqueuelen 1  (Local Loopback)
        RX packets 46  bytes 2693 (2.6 KiB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 46  bytes 2693 (2.6 KiB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0
```

```
[root@itso-dal10-sv-rsp ~]# cat /etc/supportcenter/proxy.conf
# Configuration file for remote support proxy 1.3

# Mandatory configuration

# Network interface and port that the storage system will connect to
ListenInterface eth0
ListenPort 8988

#Remote support for SVC and Storwize systems on the following front servers
ServerAddress1 129.33.206.139
ServerPort1 443
ServerAddress2 204.146.30.139
ServerPort2 443

# Optional configuration

# Network interface (lo for local) for status queries
# StatusInterface ?
# StatusPort ?

# HTTP proxy for connecting to the internet
# HTTPProxyHost ?
# HTTPProxyPort ?
# Optional authentication data for HTTP proxy
# HTTPProxyUser ?
# HTTPProxyPassword ?

# External logger (default is none)
# Logger /usr/share/supportcenter/syslog-logger

# Restricted user
# User nobody

# Log file
# LogFile /var/log/supportcenter_proxy.log

# Optional debug messages for troubleshooting
# DebugLog No

# Control IPv4/IPv6 usage
# UseIPv4 yes
# UseIPv6 yes
# UseIPv6LinkLocalAddress no
```

6. When the service is configured, you must start the service so that the server can start listening for requests. Optionally, you can also configure the service to start on system start. To start the service, run the **service** or **systemctl** command. To make the service start on system start, run the **chkconfig** command. Both of these processes are shown in Example 5-4 on page 70.

Example 5-4 Starting the service

```
[root@itso-dal10-sv-rsp ~]# service supportcenter_proxy start
Starting IBM remote support proxy:                         [  OK  ]
[root@itso-dal10-sv-rsp ~]# chkconfig supportcenter_proxy on
```

When the service starts, you are ready to configure IBM Spectrum Virtualize to use the proxy to start remote support requests.

5.2.3 Logging in to the IBM Spectrum Virtualize for Public Cloud cluster and completing the installation

When the installation is complete, you can log in to the IBM Spectrum Virtualize for Public Cloud on AWS cluster through the WebGUI (as shown in Figure 5-17) by using the Bastion public IP address as a proxy after the **`enable-sv-cloud-management-gui`** command is run on the Bastion host. Complete the following steps:

1. With the proxy enabled, open a browser to the Bastion public IP and append the port ID (:8443) to access the IBM Spectrum Virtualize WebGUI. In our example, it was `https://3.125.52.102:8443`.

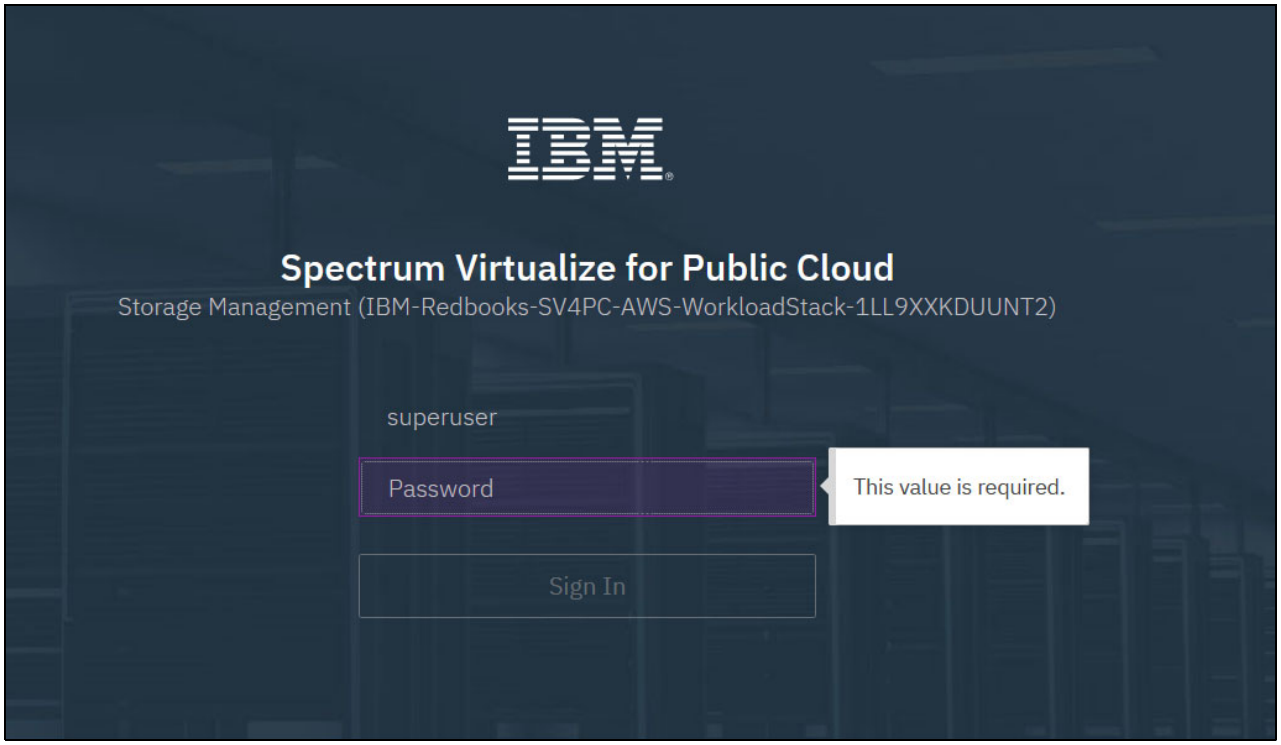


Figure 5-17 Logging in to WebGUI

2. You are redirected to the Welcome window. Click **Next** (see Figure 5-18).

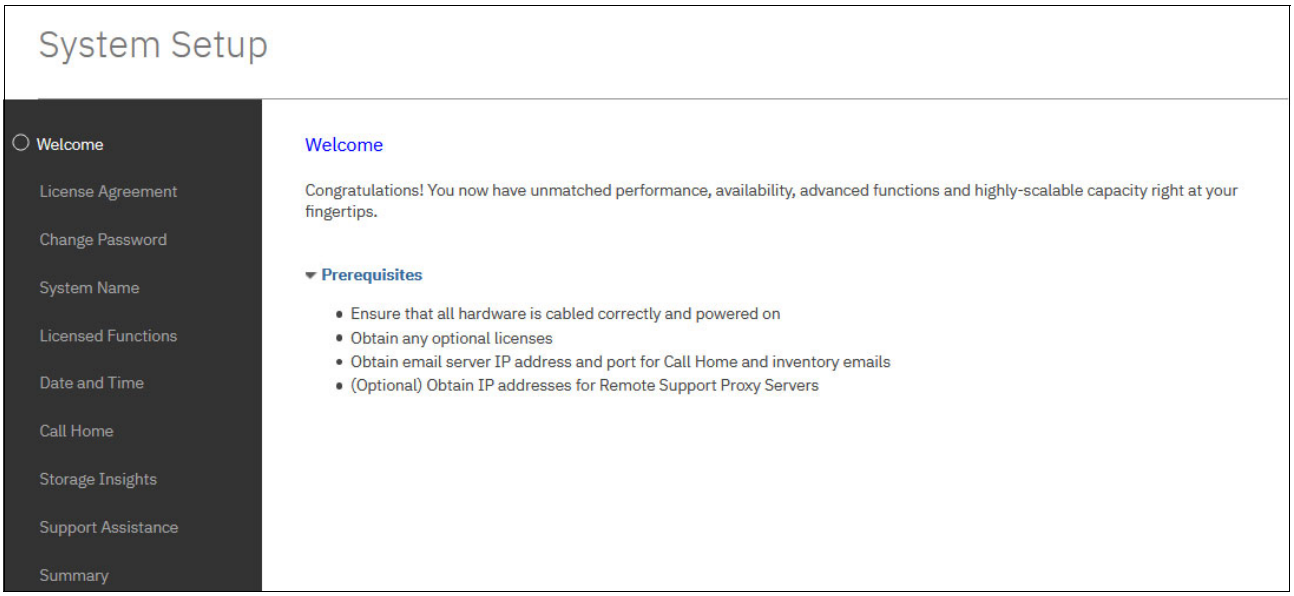


Figure 5-18 EasySetup: Welcome window

3. You are redirected to the Change Password window, as shown in Figure 5-19. Change your password, and then, click **Apply** and then, **Next** to open the next window.

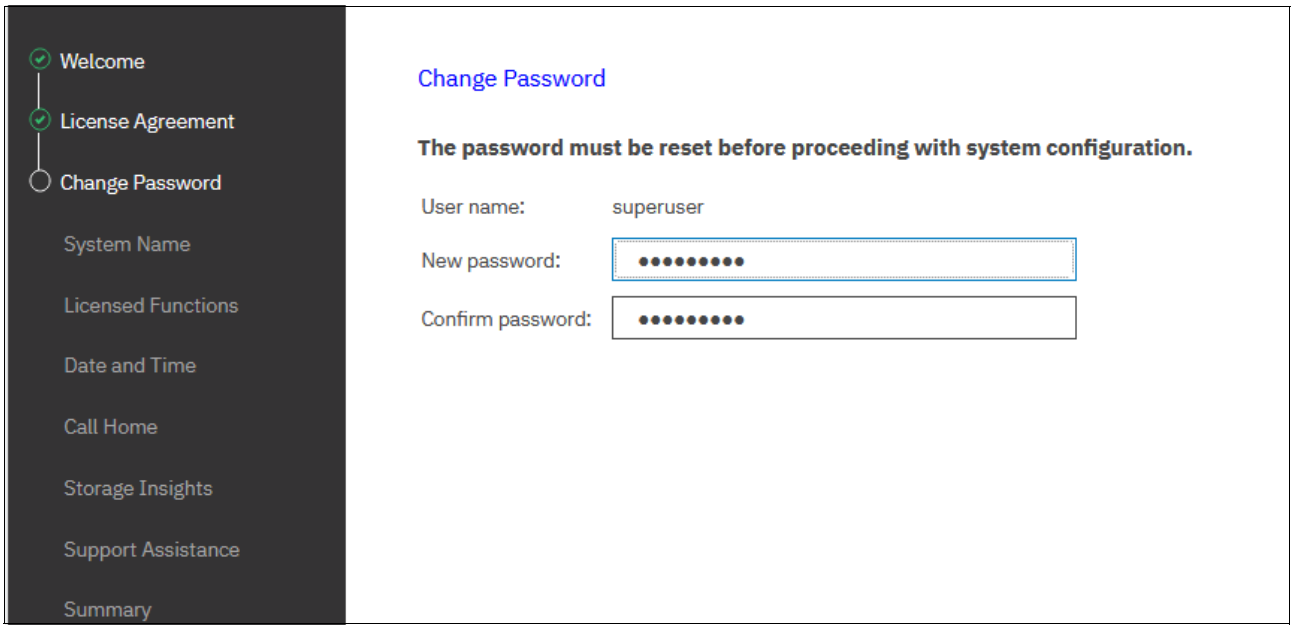


Figure 5-19 Easy Setup: Change Password window

4. You can change your cluster name, which defaults to the stack ID name and -WorkloadStack-{stack unique identifier}. As a best practice, trim the unique identifier at the end, as shown in Figure 5-20. Click **Apply** and **Next** to open the next one.

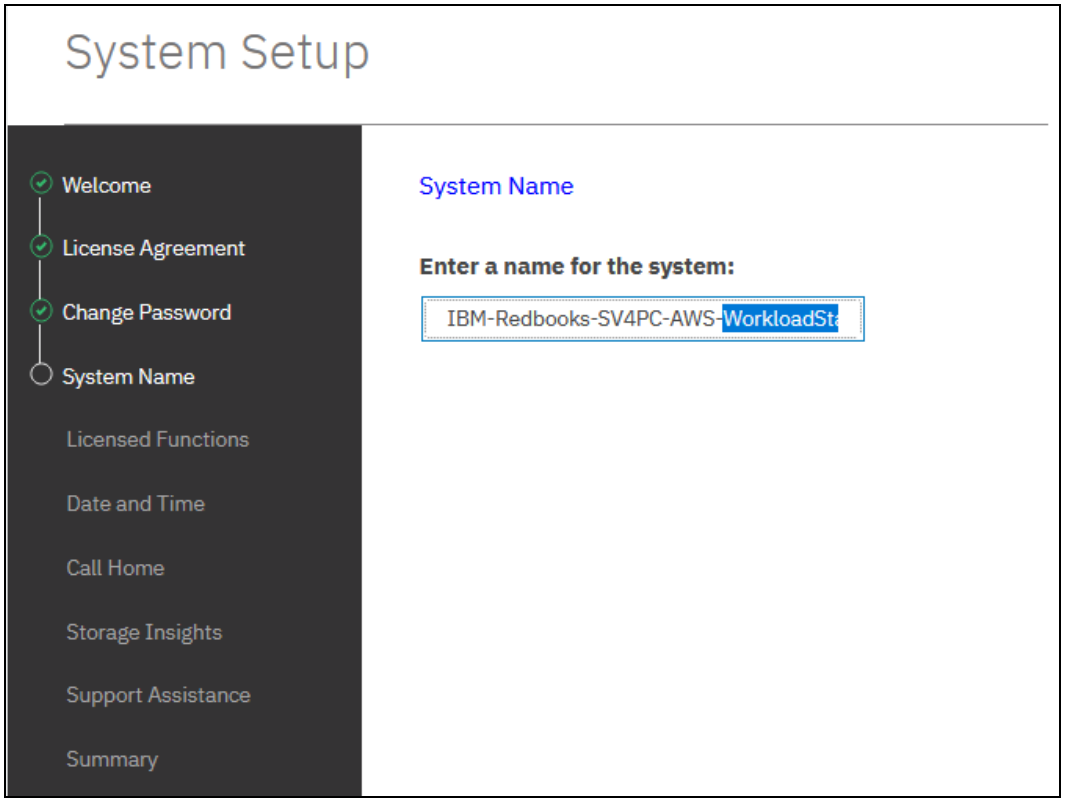


Figure 5-20 EasySetup: Trimming the system name

5. Enter your capacity license in accordance with your IBM agreement, as shown in Figure 5-21. Click **Apply** and **Next** to open the next window.

Note: An IBM Spectrum Virtualize for Public Cloud license uses simple TiB values instead of Storage Capacity Units. This feature keeps the licensing model simple and still realizes economic benefits through thin provisioning and IBM Easy Tier. It also allows for overallocation of the Amazon EBS volumes that are purchased and for the use of fewer expensive high-performance Amazon EBS volumes and cheaper low-performance volumes.

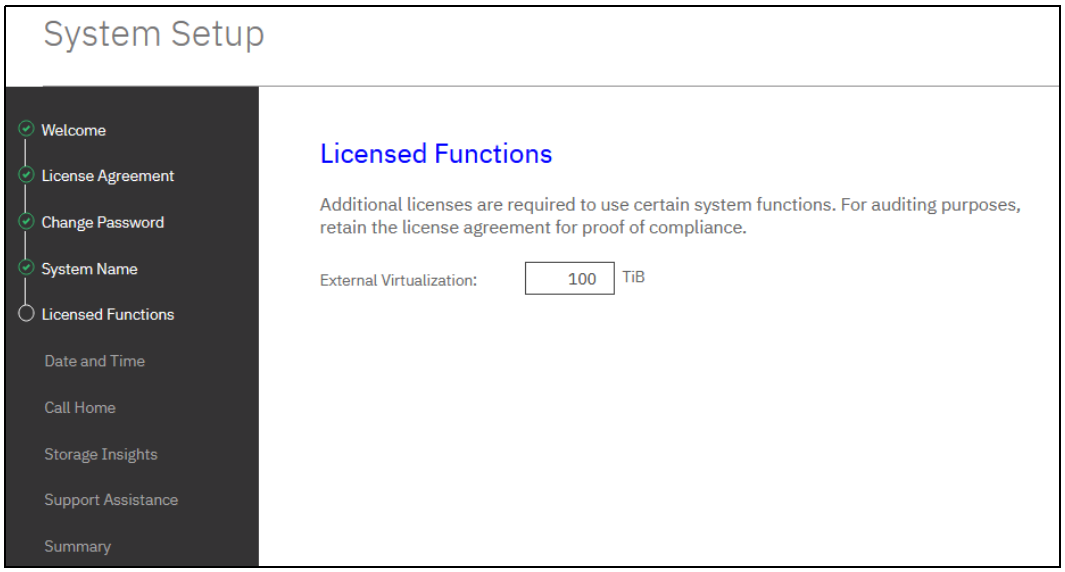


Figure 5-21 EasySetup: Licensed Functions

6. You do not need to set the Date and Time because it is controlled by AWS. IBM Spectrum Virtualize for Public Cloud on AWS is configured by the AWS time server by using underlying operating system methods.

Note: Changing the time server or setting a static time is not recommended and might cause difficulties.

For more information about the AWS time server, see Setting the Time for Your Linux Instance.

Ensure that the time zone is set. For ease of troubleshooting across multiple time zones, it is a best practice to use GMT or UTC+0, as shown in Figure 5-22.

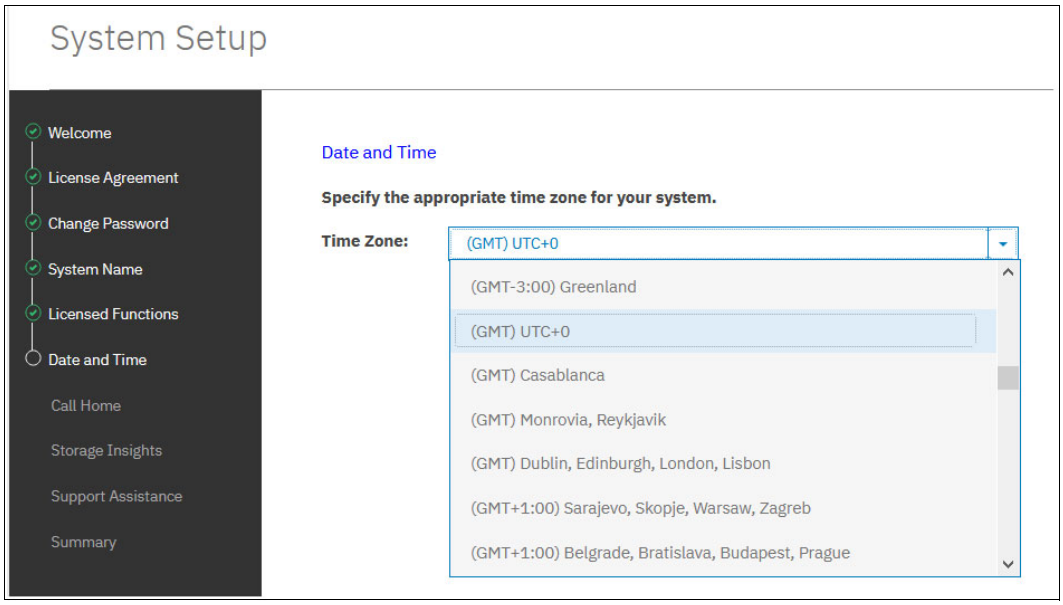


Figure 5-22 EasySetup: Time Zone

7. IBM Spectrum Virtualize for Public Cloud on AWS is preconfigured with Cloud Call Home that uses the Bastion host as a gateway. When the EasySetup process enters the Call Home configuration, Cloud Call Home verifies the connection to the support center, as shown in Figure 5-23.

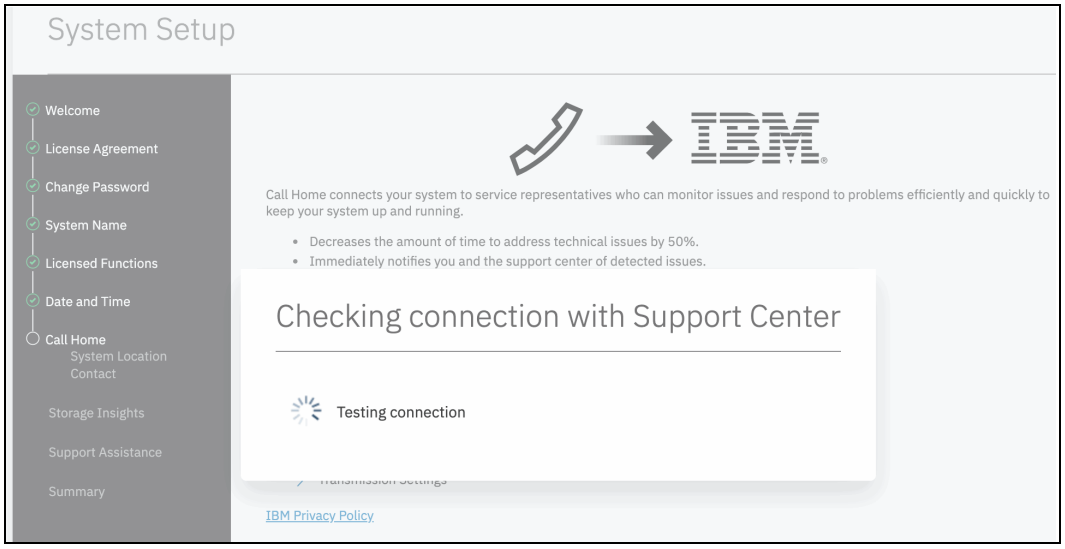


Figure 5-23 EasySetup: Cloud Call Home verification

This verification should succeed, as shown in Figure 5-24, which is the System Location window.

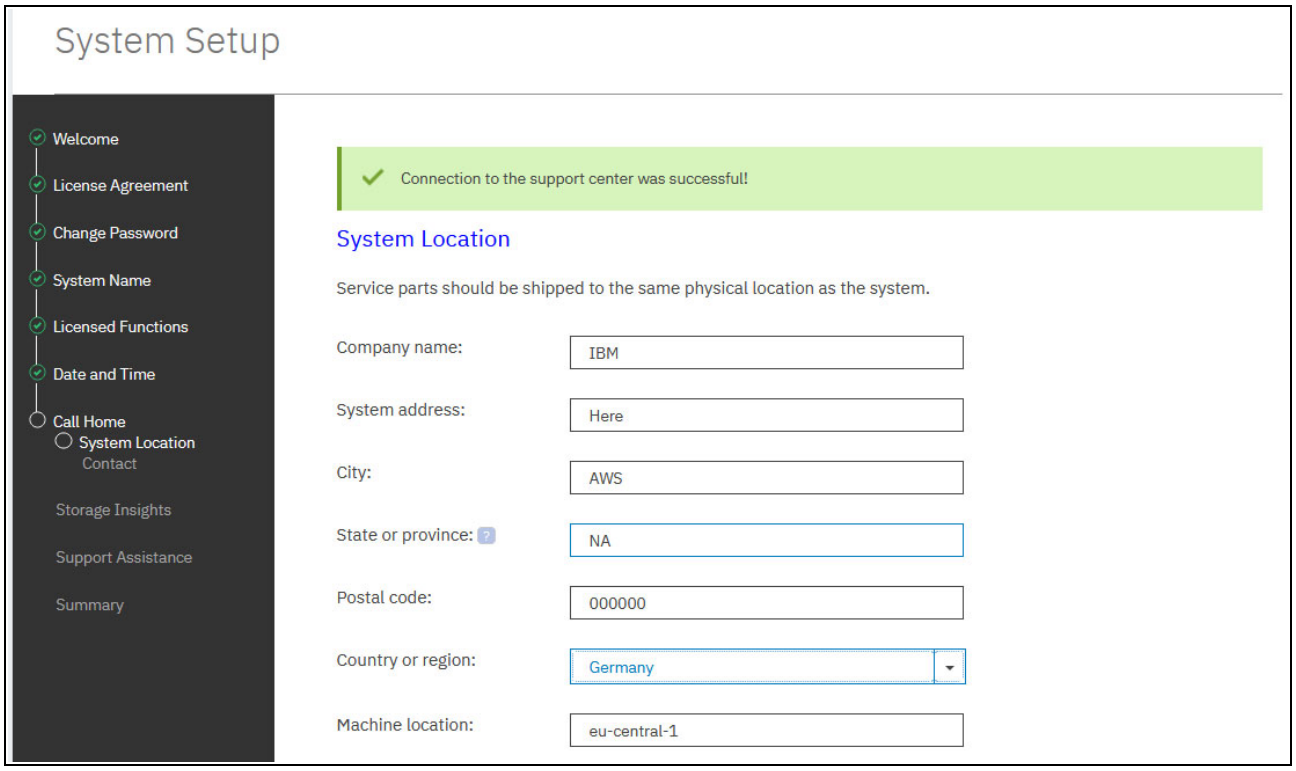


Figure 5-24 EasySetup: Successful Cloud Call Home and System Location information

8. Finish the Call Home configuration by entering the contact information, as shown in Figure 5-25.

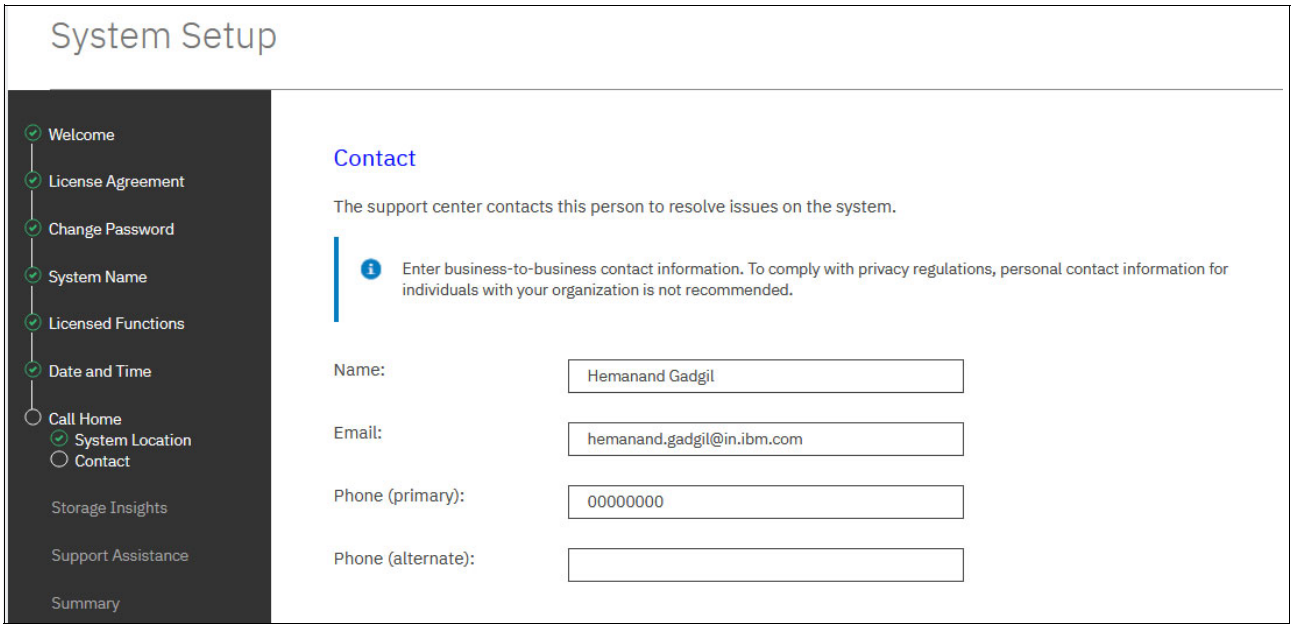


Figure 5-25 EasySetup: Contact information

9. The IBM Storage Insights configuration must be completed (this process is *not* done during EasySetup). The process requires registering for a no-charge account and installing a DataCollector, which can be installed on the Bastion host. Figure 5-26 shows the IBM Storage Insights configuration window. Skip this step for now.

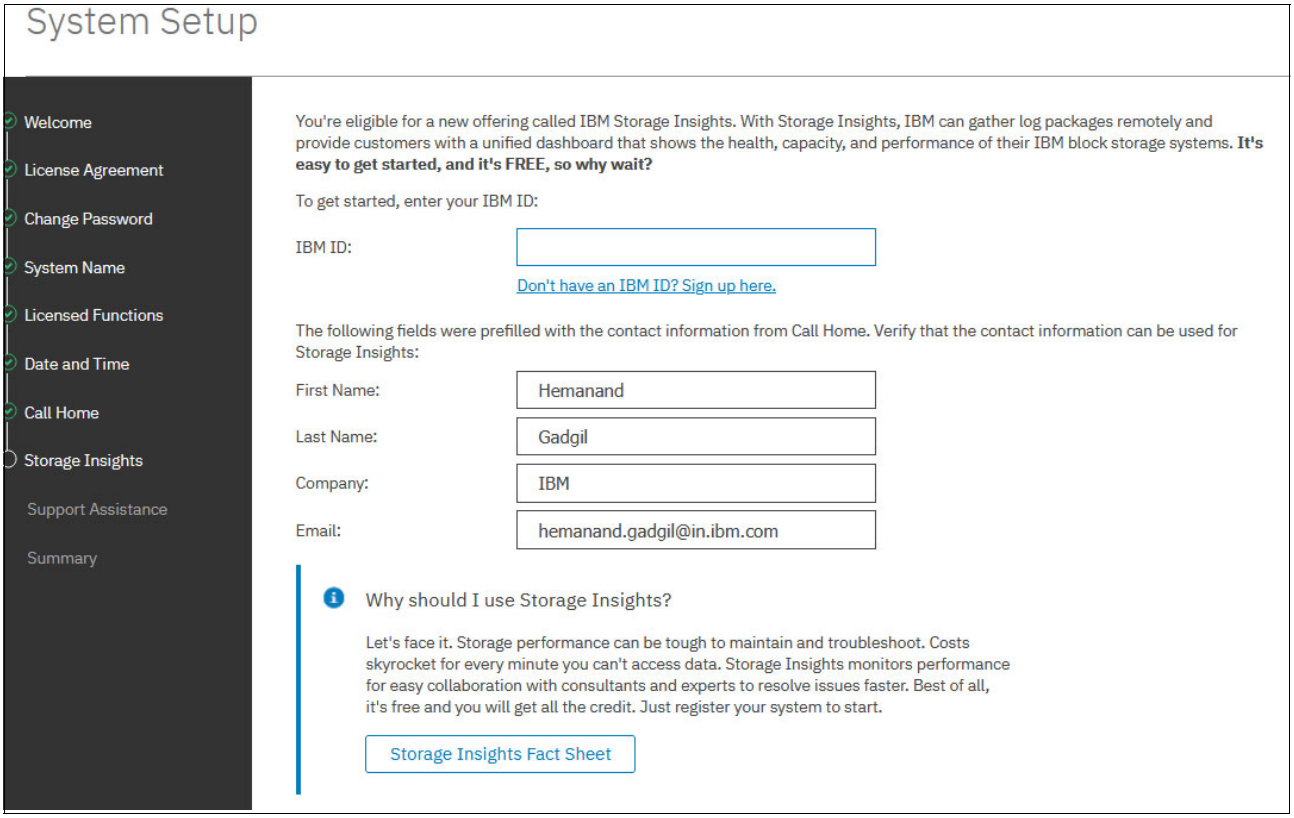


Figure 5-26 EasySetup: IBM Storage Insights

10. Configure your RSP, as shown in Figure 5-27.

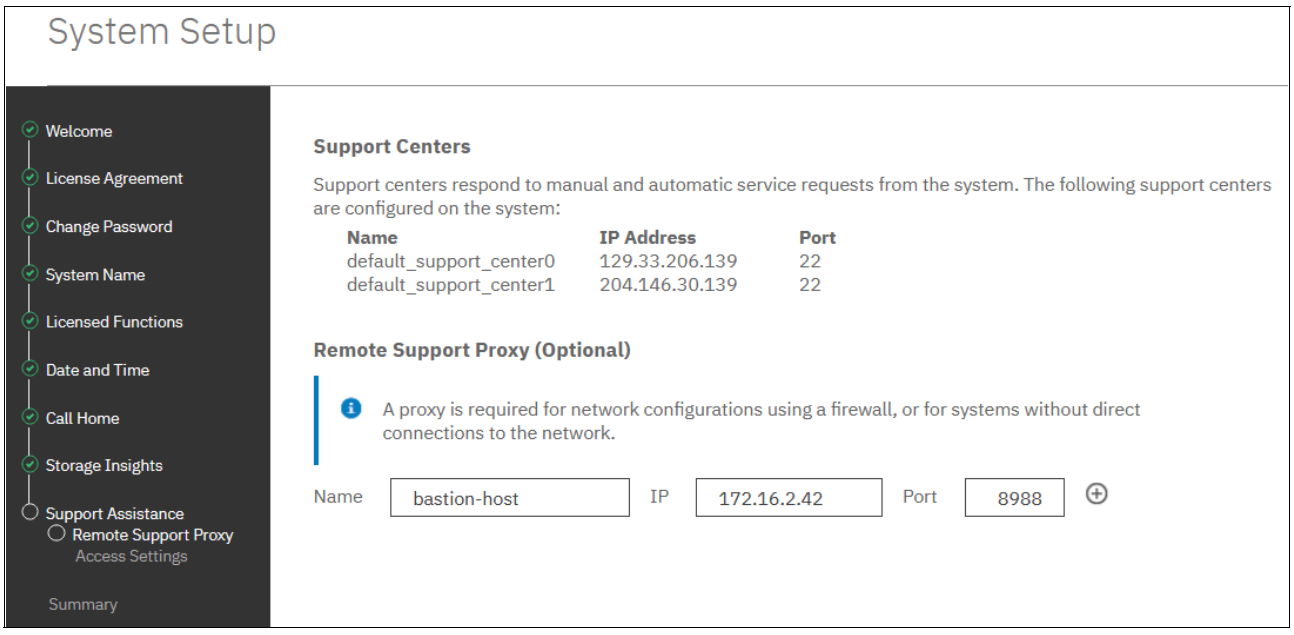


Figure 5-27 EasySetup: Remote Support Proxy

Note: This step assumes that you deployed an RSP. Again, the Bastion host is a logical choice. Note the internal IP address of the Bastion host and the `ListenPort` that was specified in 5.2.2, “Configuring the Bastion host” on page 67.

Figure 5-28 shows a summary of your configuration. Your cluster setup is complete.

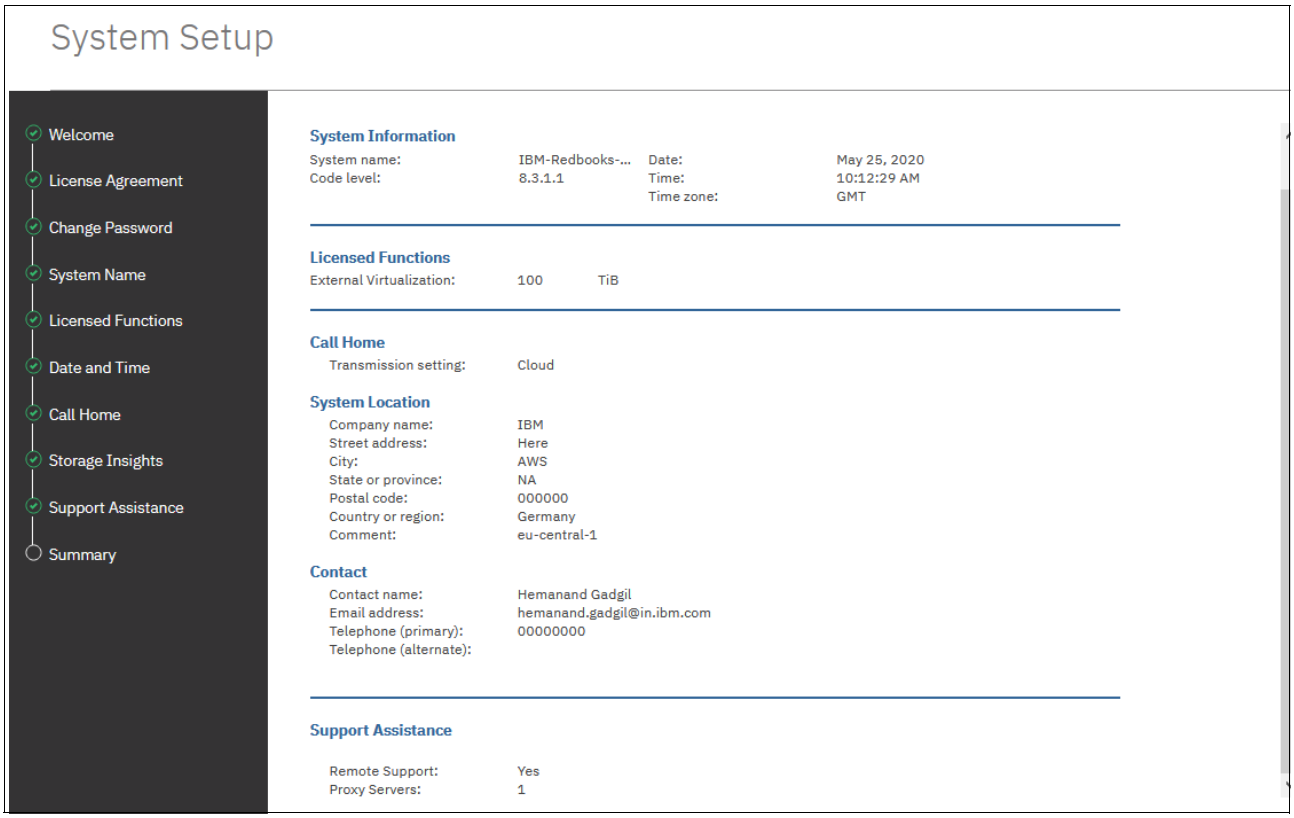


Figure 5-28 EasySetup: Summary

Note: Call Home is set up with Cloud Call Home. However, email notification is useful for event notification and can be set up after the EasySetup process is complete. The Bastion host runs an SMTP service and can be used as the email gateway.

5.3 Configuring the cloud quorum

IP quorum applications are used in Ethernet networks to resolve failure scenarios when half the nodes on the system become unavailable. These applications determine which nodes can continue processing host operations and avoids a split-brain scenario in which both halves attempt to service independently I/O, which causes corruption.

As part of the installation of IBM Spectrum Virtualize for Public Cloud on AWS, a Bastion host is provisioned and the IP quorum application is installed and configured on this instance. This Bastion host operates as the IP quorum and the network gateway for the configuration.

Note: An IP quorum is configured during the installation. You configure an extra IP quorum only if you want to enhance the fault tolerance by putting the active one in a different Availability Zone for installations into new VPCs.

Strict requirements for the IP network that uses IP quorum applications must be met. All IP quorum applications must be reconfigured and redeployed to hosts when specific aspects of the system configuration change. These aspects include adding or removing a node from the system or when node service IP addresses are changed.

Other examples include changing the system certificate or experiencing an Ethernet connectivity issue. Such a connectivity issue prevents an IP quorum application from accessing a node that is still online.

If an IP application is offline, it must be reconfigured because the system configuration changed.

To view the state of an IP quorum application in the management GUI, select **Settings** → **System** → **IP Quorum**, as shown Figure 5-29.

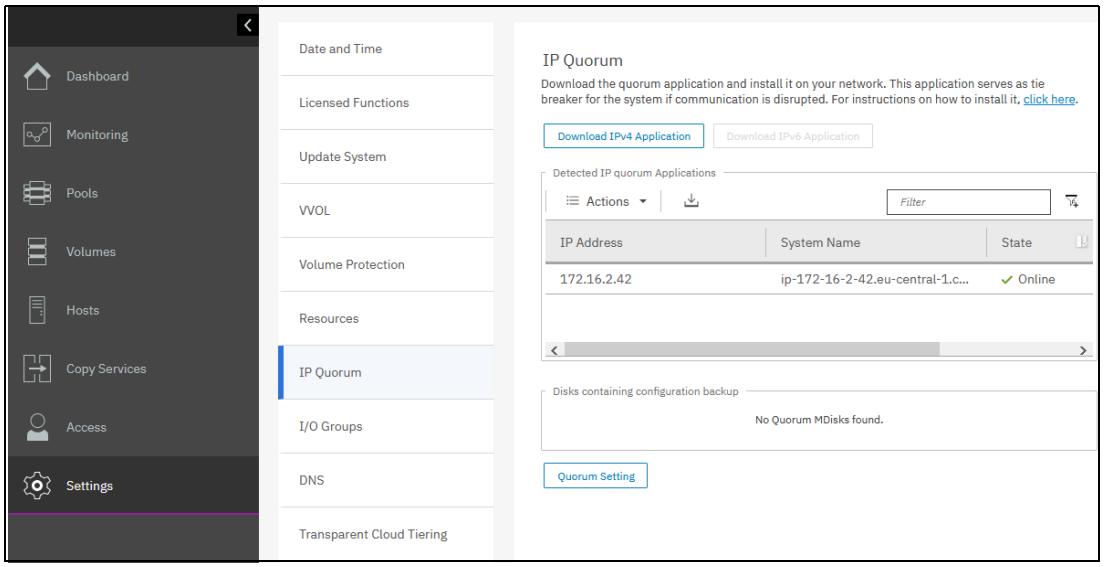


Figure 5-29 IP quorum example from the GUI

Even with IP quorum applications on an EC2 instance, quorum disks are required on each node in the system to contain backups of the configuration and recovery information. On EC2 instances where IBM Spectrum Virtualize connectivity with its nontraditional back-end storage connectivity, the quorum disks cannot be on external storage or internal disk as in SAN Controller Volume or FlashSystem systems. Therefore, they are automatically allocated on the EC2 instance boot device for each IBM Spectrum Virtualize node.

The IBM Spectrum Virtualize command **`lsquorum`** shows only the IP quorum.

The maximum number of IP quorum applications that can be deployed is five. Applications can be deployed on multiple hosts to provide redundancy.

For stable quorum resolutions, an IP network must meet the following requirements:

- Connectivity from the servers that are running an IP quorum application to the service IP addresses of all nodes.
- The network must also deal with the possible security implications of exposing the service IP addresses because this connectivity also can be used to access the service assistant interface if the IP network security is configured incorrectly.
- Port 1260 is used by IP quorum applications to communicate from the hosts to all nodes.
- The maximum round-trip delay must not exceed 80 milliseconds (ms), which means 40 ms each direction.
- A minimum bandwidth of 2 MBps is guaranteed for node-to-quorum traffic.

For more information about the IP quorum configuration, see IBM Knowledge Center.

Note: The current Cloud Formation Template (CFT) for new VPCs deploys the Bastion host (which houses the initial IP quorum device) into the same Availability Zone as the IBM Spectrum Virtualize nodes. If deploying into an existing VPC, it is possible to place that Bastion host on a subnet that is in a different Availability Zone from the IBM Spectrum Virtualize nodes.

However, if you are deploying into a new VPC that is created as part of the IBM Spectrum Virtualize installation process, it is a best practice that you create a new subnet in that VPC that belongs to a different Availability Zone. Then, start a new secure EC2 instance by using only a private interface in that new subnet with no direct access from the internet. Next, you deploy an IP quorum application on that server and restart the one on the Bastion host so that the secure, redundant IP quorum is the active quorum device.

In summary, deploying a second IP quorum server with a new VPC includes the following steps:

1. Create a subnet within the VPC in a different Availability Zone than the IBM Spectrum Virtualize nodes and Bastion host.
2. Start a new EC2 instance. You can use the Amazon Linux Amazon Machine Images (AMI) 2018.03.0 image from the quick start because it has Java preinstalled. The default type of t2.micro is suitable but do *not* select Review and Launch.
3. Click **Next: Configure Instance Details** and select the correct VPC and subnet that you created in step 1. Leave **Public IP** disabled for added security and use an existing security group (same as the Bastion host).
4. Click **Review and Launch** to review the configuration and then, click **Launch**.
5. Select the key pair that was used during the creation of the cluster because the key pair is needed to access the new EC2 instance.
6. After the instance is provisioned, run the **scp** command on the private key that is used to access the Bastion host over *to* the Bastion host.
7. Run **ssh** to access the Bastion host and run **scp** to transfer the `ip_quorum.jar` file from the Bastion host over to the new EC2 instance by using the private key:

   ```
   scp -i ~/.ssh/privkey.pem /usr/local/bin/ip_quorum.jar ec2-user@{new EC2 IP}:
   ```

8. Run **ssh** to access the new EC2 instance and test the ip_quorum service:

   ```
   java -jar ~/ip_quorum.jar
   ```

9. Set up the quorum as a service or install a cronjob to ensure that it is always running.
10. Exit the new EC2 instance and restart the `ip_quorum` service on the Bastion host:

    ```
    systemctl restart ip-quorum
    ```

5.4 Expanding from a 2-node to 4-node cluster in AWS

IBM Spectrum Virtualize for Public Cloud software in AWS supports 2-node and 4-node cluster configurations. You can expand a 2-node cluster to four nodes by adding nodes to a stack in AWS.

5.4.1 Prerequisites

Before you expand a 2-node cluster to a 4-node cluster, you must ensure that both the nodes that are added to the configuration and existing nodes in the cluster are updated to the latest version of the IBM Spectrum Virtualize for Public Cloud software. For information, see IBM Knowledge Center.

To expand a 2-node cluster to a 4-node cluster in AWS, complete the following steps:

1. Log on to the AWS Management Console with the AWS default administrator profile or the installer profile.
2. Select **CloudFormation** → **Stacks**. Select the existing 2-node cluster configuration. It is displayed as a nested workload with the following name format:

   `[stack-name]-workstack-{resource id}`

   The stack-name is specified when the cluster is created with the AWS CloudFormation template (see Figure 5-30).

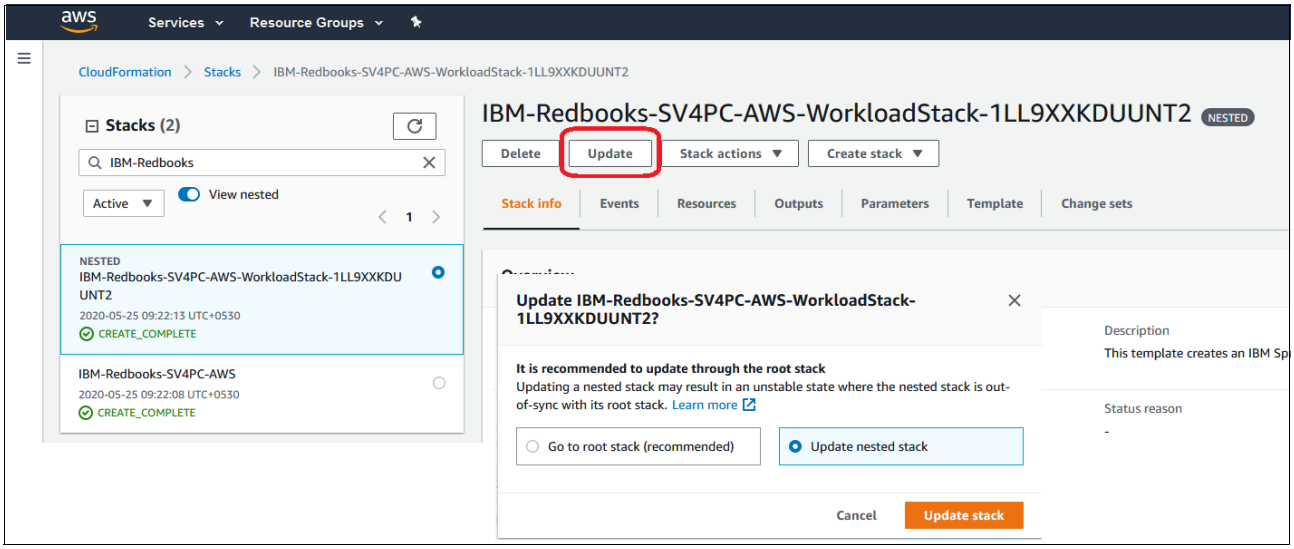


Figure 5-30 Expanding the 2 nodes cluster to 4 nodes cluster

3. Click **Update**.
4. Select **Updated nested stack** and click **Update stack**.
5. On the Update stack page, select the following options:
   a. In the Prerequisite-Prepare template section, select **Replace current template**.
   b. In the Specify template section, select **Amazon S3 URL**.
   c. In the Amazon S3 URL field, enter the URL that is displayed in the StackUpdateTemplate field.
   d. Click **Next**.

This information is included in the summary and email notification when the node instances were first installed in AWS. This information is also included on the Output tab when the node instances were first installed in AWS (see Figure 5-31).

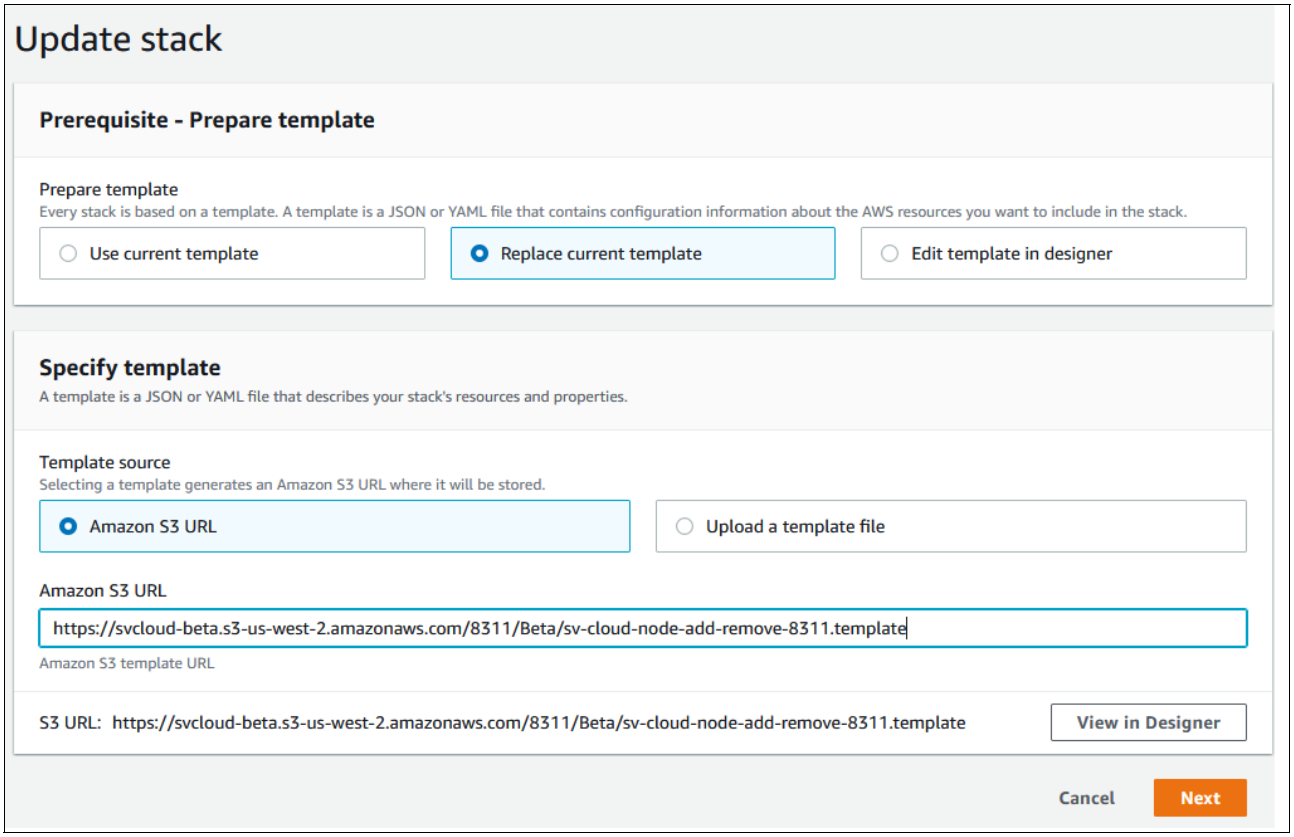


Figure 5-31 Update stack URL

6. Click **Next**.

7. On the Specify stack details page, keep the values that are configured for the existing configuration. Review the Amazon EC2 Configuration section and confirm the node instance type for the new I/O group is correct. Click **Next**, as shown in Figure 5-32.

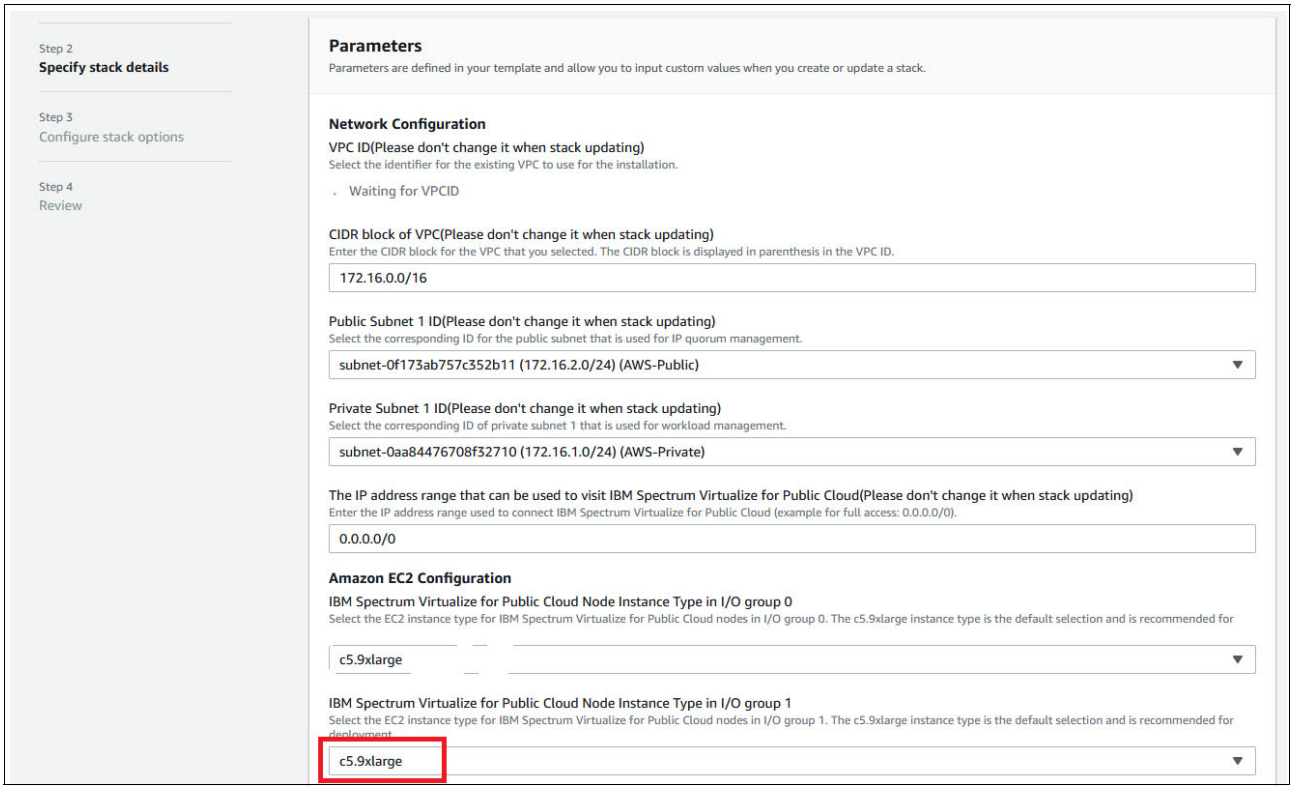


Figure 5-32 Selection of IO group configuration for expansion from 2 node to 4 node

8. On the Configure stack options page, keep the values that are configured. Click **Next**.
9. On the Review page, review the options. Click **Next**.
10. On the Change set preview page, review the changed resources. Several resources are modified and two more EC2 instances are added for the nodes. After you verify these changes, ensure that **I acknowledge that AWS CloudFormation might create IAM resource** is selected.
11. Click **Update stack**. Verify that the status of the nested stack changes to `Update_In_Progress`.

12. After the stack is listed as `UPDATE_COMPLETE`, review the details that are listed for the updated nested stack on the **CloudFormation** → **Stacks** page. It includes the configuration of the existing cluster and the new nodes (see Figure 5-33).

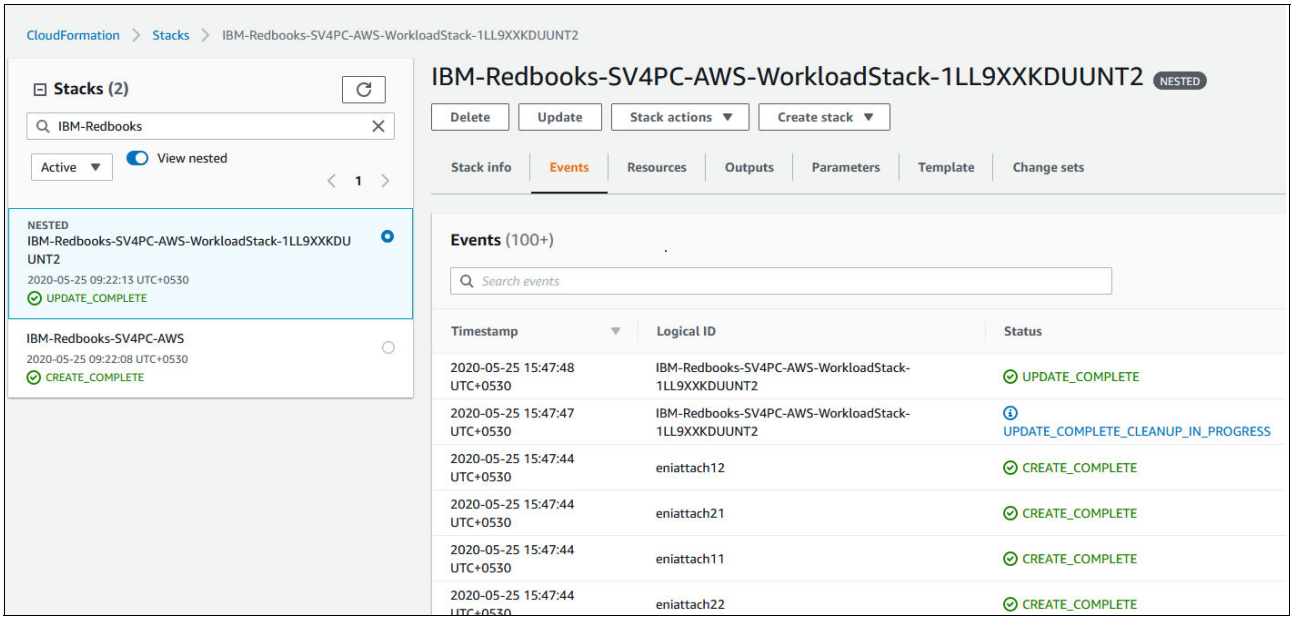


Figure 5-33 Completion of 2-node to 4-node update task

The output with all the IP addresses of newly deployed nodes is displayed in Output tab with node 3 and node 4 added, as shown in Figure 5-34 on page 85.

Names	IP address	Descriptions
IBMSVNode1Port1NodeIP	172.16.1.104	IBM Spectrum Virtualize Node 1 Port 1 Node IP
IBMSVNode1Port2NodeIP	172.16.1.134	IBM Spectrum Virtualize Node1 Port 2 Node IP
IBMSVNode1PortIP1	172.16.1.91	IBM Spectrum Virtualize Node 1 Port IP 1
IBMSVNode1PortIP2	172.16.1.154	IBM Spectrum Virtualize Node 1 Port IP 2
IBMSVNode1ServiceIP	172.16.1.181	IBM Spectrum Virtualize Node1 Service IP
IBMSVNode2Port1NodeIP	172.16.1.241	IBM Spectrum Virtualize Node 2 Port 1 Node IP
IBMSVNode2Port2NodeIP	172.16.1.40	IBM Spectrum Virtualize Node 2 Port 2 Node IP
IBMSVNode2PortIP1	172.16.1.36	IBM Spectrum Virtualize Node 2 Port IP 1
IBMSVNode2PortIP2	172.16.1.236	IBM Spectrum Virtualize Node 2 Port IP 2
IBMSVNode2ServiceIP	172.16.1.193	IBM Spectrum Virtualize Node 2 Service IP
IBMSVNode3Port1NodeIP	**172.16.1.20**	**IBM Spectrum Virtualize Node 3 Port 1 Node IP**
IBMSVNode3Port2NodeIP	**172.16.1.73**	**IBM Spectrum Virtualize Node 3 Port 2 Node IP**
IBMSVNode3PortIP1	**172.16.1.198**	**IBM Spectrum Virtualize Node 3 Port IP 1**
IBMSVNode3PortIP2	**172.16.1.77**	**IBM Spectrum Virtualize Node 3 Port IP 2**
IBMSVNode3ServiceIP	**172.16.1.211**	**IBM Spectrum Virtualize Node 3 Service IP**
IBMSVNode4Port1NodeIP	**172.16.1.59**	**IBM Spectrum Virtualize Node 4 Port 1 Node IP**
IBMSVNode4Port2NodeIP	**172.16.1.173**	**IBM Spectrum Virtualize Node 4 Port 2 Node IP**
IBMSVNode4PortIP1	**172.16.1.107**	**IBM Spectrum Virtualize Node 4 Port IP 1**
IBMSVNode4PortIP2	**172.16.1.94**	**IBM Spectrum Virtualize Node 4 Port IP 2**
IBMSVNode4ServiceIP	**172.16.1.119**	**IBM Spectrum Virtualize Node 4 Service IP**

Figure 5-34 IP addresses of newly deployed nodes

5.5 Shrinking the configuration from four nodes to two nodes in Amazon Web Services

IBM Spectrum Virtualize for Public Cloud software in AWS supports 2-node and 4-node cluster configurations. You can shrink a four-node cluster to two nodes by removing nodes in the stack in AWS.

For more information about prerequisites and restrictions, see IBM Knowledge Center.

Move the volumes to the remaining I/O group per the procedure that is described in IBM Knowledge Center.

After the prerequisites are completed, follow steps 1 - 6 as described in 5.4, “Expanding from a 2-node to 4-node cluster in AWS” on page 81. Then, complete the following steps:

1. On the Specify stack details page, keep the values that are configured for the configuration. Review the Amazon EC2 Configuration section and ensure that the I/O group that is being removed is set to **None** (see Figure 5-35 on page 86). For example, in this procedure, `iogrp0` is being removed and the IBM Spectrum Virtualize node instance type for I/O group 0 must be set to **None**. Click **Next**.

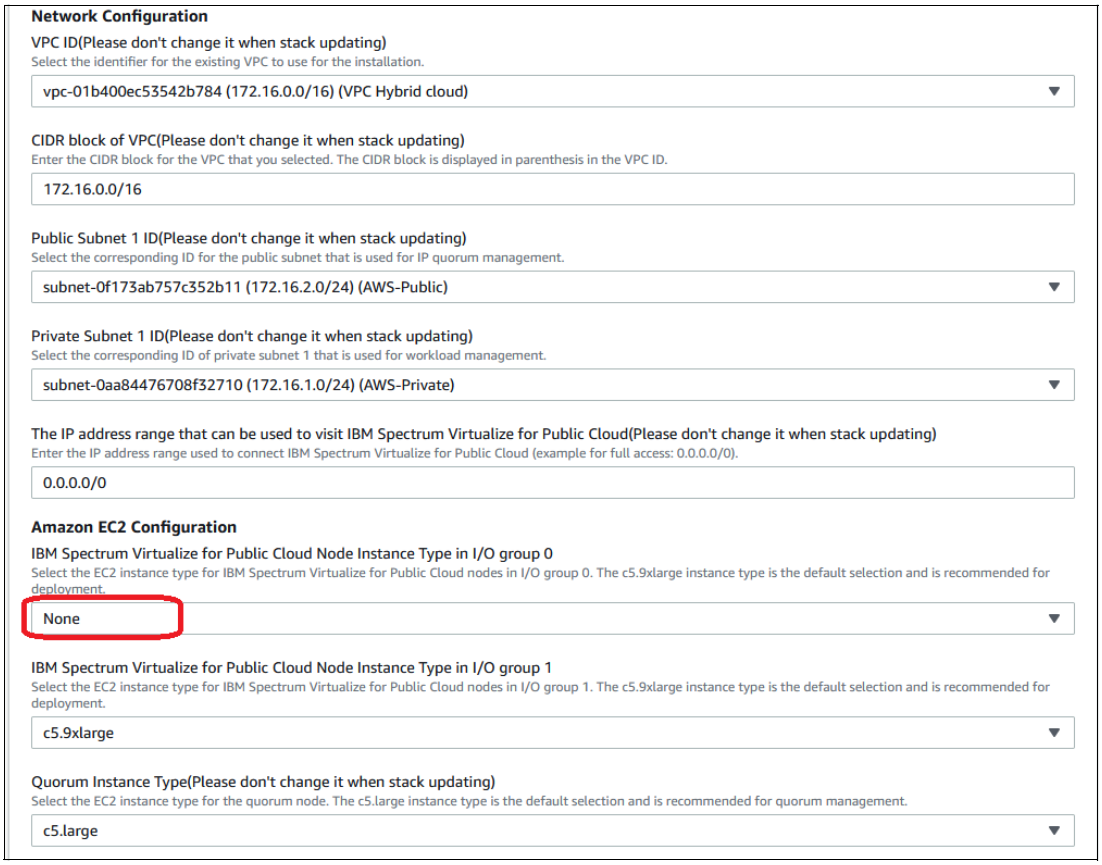


Figure 5-35 Selection of IO group none to shrink the cluster

2. On the Configure stack options page, keep the configured values. Click **Next**.
3. On the Review page, review the options. Click **Next**.
4. On the Change set preview page, review the changed resources. Several resources are modified and two EC2 instances are removed for the two nodes that are being deleted from the cluster. After verifying these changes, ensure that **I acknowledge that AWS CloudFormation might create IAM resource** is selected.
5. Click **Update stack**. Verify that the status of the nested stack changes to `Update_In_Progress`.
6. After the stack is listed as `UPDATE_COMPLETE`, review the details that are listed for the updated nested stack on the **CloudFormation** → **Stacks** page (see Figure 5-36 on page 87).

Names	IP address	Descriptions
IBMSVNode3Port1NodeIP	172.16.1.20	IBM Spectrum Virtualize Node 3 Port 1 Node IP
IBMSVNode3Port2NodeIP	172.16.1.73	IBM Spectrum Virtualize Node 3 Port 2 Node IP
IBMSVNode3PortIP1	172.16.1.198	IBM Spectrum Virtualize Node 3 Port IP 1
IBMSVNode3PortIP2	172.16.1.77	IBM Spectrum Virtualize Node 3 Port IP 2
IBMSVNode3ServiceIP	172.16.1.211	IBM Spectrum Virtualize Node 3 Service IP
IBMSVNode4Port1NodeIP	172.16.1.59	IBM Spectrum Virtualize Node 4 Port 1 Node IP
IBMSVNode4Port2NodeIP	172.16.1.173	IBM Spectrum Virtualize Node 4 Port 2 Node IP
IBMSVNode4PortIP1	172.16.1.107	IBM Spectrum Virtualize Node 4 Port IP 1
IBMSVNode4PortIP2	172.16.1.94	IBM Spectrum Virtualize Node 4 Port IP 2
IBMSVNode4ServiceIP	172.16.1.119	IBM Spectrum Virtualize Node 4 Service IP

Figure 5-36 Details that are listed for the updated nested stack

5.6 Configuring the back-end storage and pools

IBM Spectrum Virtualize for Public Cloud on AWS uses the back-end storage that is provided by Amazon Elastic Block Store (EBS) as external MDisks. As part of the initial default installation, two gp2 Amazon EBS volumes are allocated and put into a pool on the IBM Spectrum Virtualize cluster (see Figure 5-37).

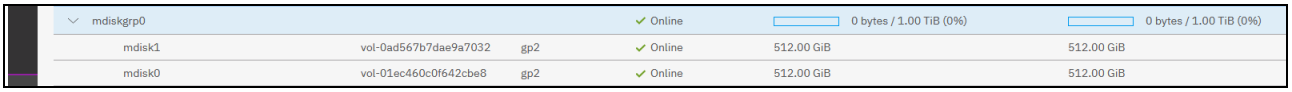


Figure 5-37 Default Amazon EBS gp2 volumes that are specified during CloudFormation template configuration

If more or even different storage is want, complete the following steps:

1. To order back-end storage, log in to the AWS Console.
2. Click **Services** in the upper left corner of the browser window. Then, click **EC2**.
3. Under Resources, click **Volumes**. In the window that opens, you can create volumes and view current volumes.

> **Note:** The AWS CloudFormation template provides two gp2 Amazon EBS volumes of a size that is specified during the CloudFormation template configuration for use with your IBM Spectrum Virtualize cluster.
>
> Either before adding Amazon EBS volumes to a storage pool or as a part of the assignment process, be sure to follow the recommendation for correctly aligning the Amazon EBS volume type to IBM Spectrum Virtualize performance expectations in accordance with Table 4-2 on page 48.

4. To create a volume, click **Create Volume** in upper left of the window.
5. Select the volume type and size of the volume that is required, as shown in Figure 5-38 on page 88.

> **Note:** When you create an Amazon EBS volume, ensure that you choose the same Availability Zone as the IBM Spectrum Virtualize for Public Cloud on AWS instance.

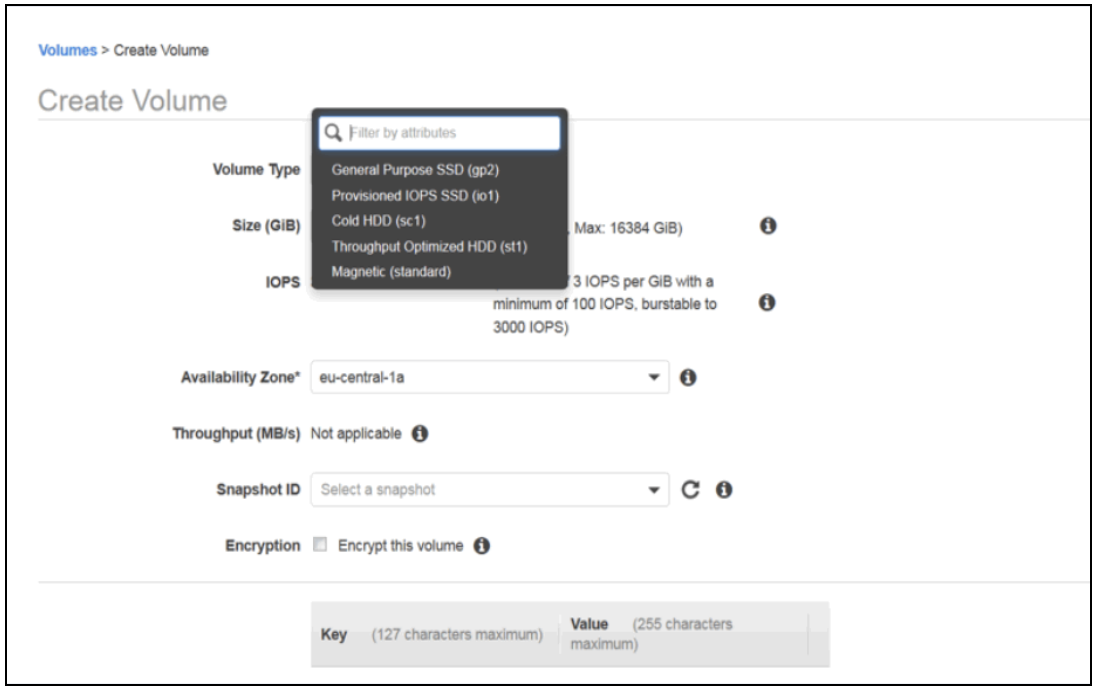


Figure 5-38 Amazon EBS: Create Volume on the AWS Console

Volumes that are created are viewable on the AWS Console in the **EBS volumes** section, and they should include a status of `Available`.

As shown in Figure 5-39, two pools are created on IBM Spectrum Virtualize for Public Cloud on AWS and each pool features one MDisk assigned, which is the Amazon EBS external storage that is purchased on AWS Cloud.

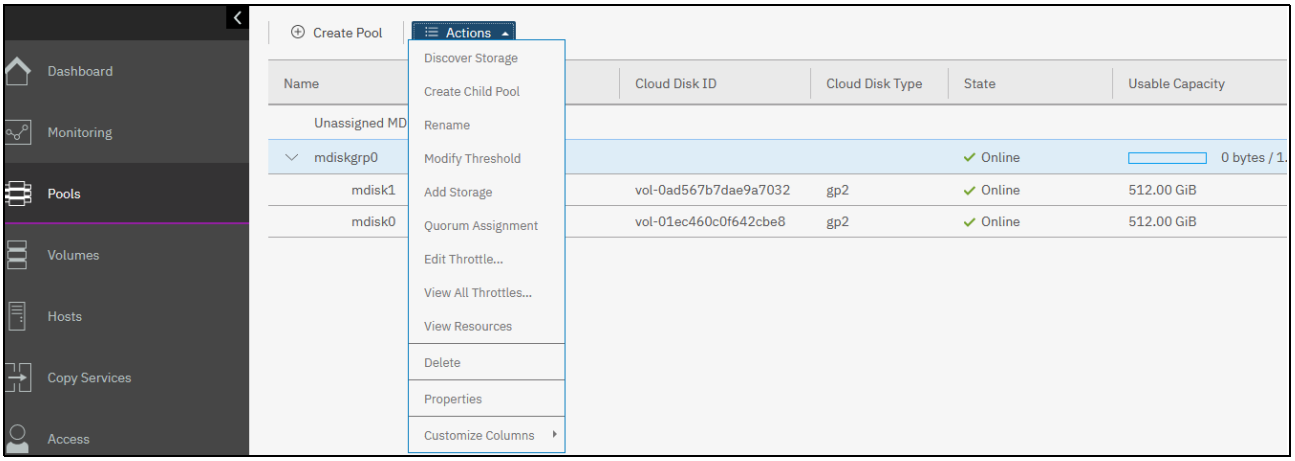


Figure 5-39 Pool creation

6. To create pool on IBM Spectrum Virtualize for Public Cloud on AWS, log in to the IBM Spectrum Virtualize for Public Cloud on AWS GUI and select **Pools** → **Create Pool**.

7. After the pool is created, select **Action** → **Discover Storage**. The Amazon EBS volumes that were purchased on AWS Cloud and are free and unused are visible under Unassigned MDisk. To cross-verify that the correct volume is added to the pool, check to see whether the Amazon EBS Volume ID is the same volume ID that is seen on the AWS Cloud console.
8. Add storage in the form of MDisks to the pool. Only 16 MDisks can be used per I/O group.
9. IBM Spectrum Virtualize for Public Cloud on AWS from version 8.3.1 supports data reduction pools. To use data reduction technologies on the system, you must create a data reduction pool, create volumes with the data reduction pool, and map these volumes to hosts that support SCSI **unmap** commands. For more information about data reduction pools, see IBM Knowledge Center.
10. Create a VDisk and assign the volume for host access by using iSCSI.

5.6.1 Configuring an IBM Spectrum Virtualize volume

In this section, you create a volume by using the pool that was created with the Amazon EBS volumes or MDisks. Volumes can be fully allocated or thinly provisioned (space-efficient). The default pre-allocation that is indicated by the command-line interface (CLI) in Example 5-5 is 2% (specified by the real size [rsize]). You have 98% of the capacity for the volumes that is available in the pool for other volumes until this volume claims it.

Example 5-5 Thinly provisioned (space-efficient) volume creation by using the CLI

```
svctask mkvdisk -autoexpand -grainsize 256 -mdiskgrp 2 -name thin-test -rsize 2%
-size 32212254720 -unit b -warning 80%
```

Figure 5-40 on page 90 shows thinly provisioned (space-efficient) volume creation by using the GUI.

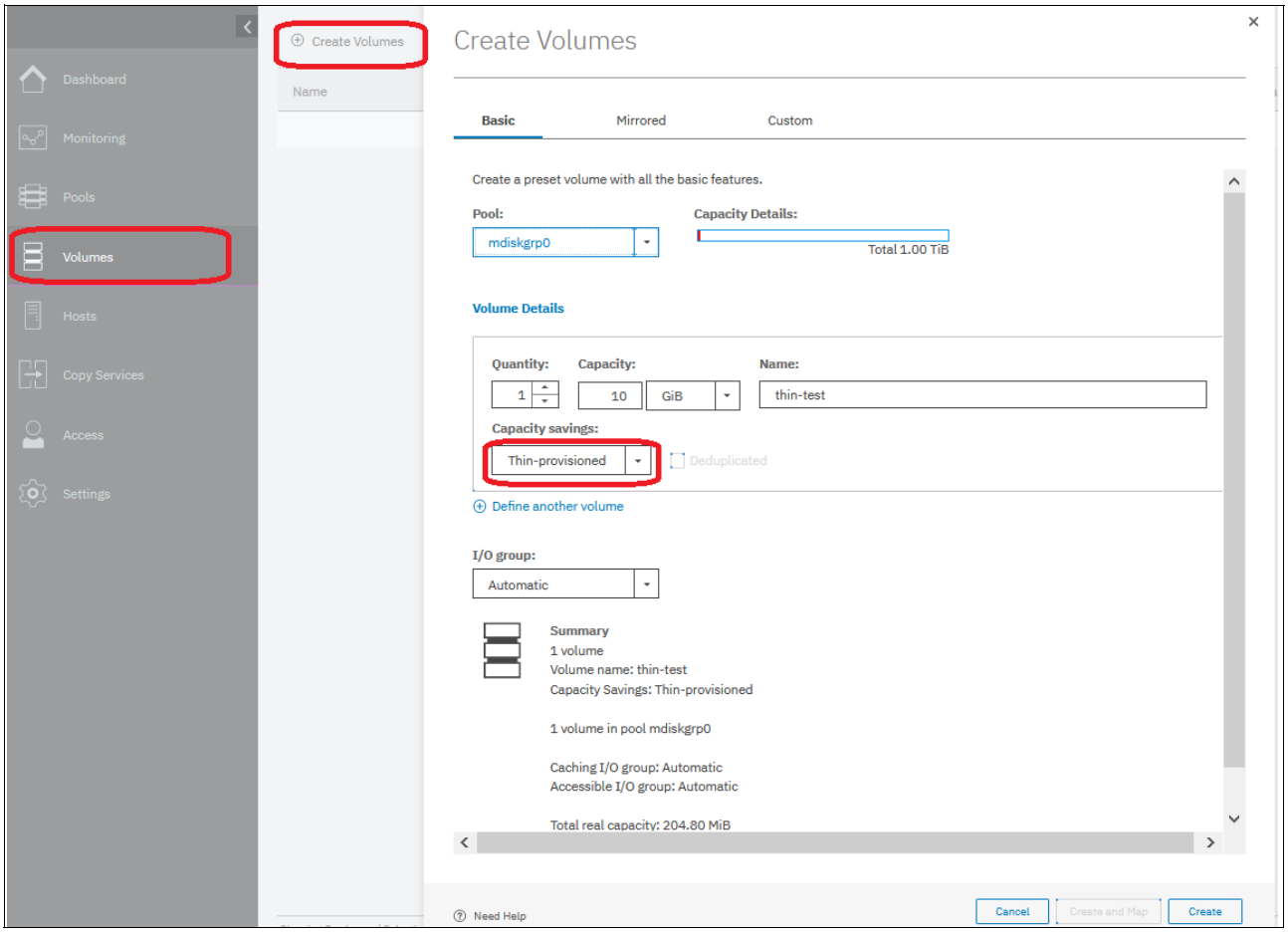


Figure 5-40 Thinly provisioned (space-efficient) volume creation with the GUI

Thinly provisioned volumes allow users to over-provision the Amazon EBS volumes and therefore reduce the overall operational cost in AWS.

5.6.2 Configuring the host and volume mapping

To use the volume that you created, you must map it to a host object. The host object represents a single Bare Metal Server on your cloud account and its iSCSI-qualified identifier (IQN), which is similar to a worldwide port name (WWPN) for an FC host.

To create a host object, you must collect its IQN. The place and procedure to collect the IQN from can vary with each operating system. For more information about the required steps for your operating system, see the operating system's documentation.

When you create your host object and map your volume, depending on what operating system you using, you must install the iSCSI initiator and run some specific operations to use your mapped volumes with the hosts.

Linux host

Install the Linux software iSCSI initiator. The initiator software on RHEL systems is packaged as `iscsi-initiator-utils`, and the suggested version is 6.2.0.873-35 or later. The initiator software on SUSE Linux Enterprise Server systems is packaged as `open-iscsi`, and the suggested version is 6.2.0.873-33.2 or later.

According to IBM Knowledge Center, set the IQN, target discovery, and authentication. Then, enable multipathing for the Linux hosts.

After creating the host object and mapping VDisks to it, on the IBM Spectrum Virtualize cluster, scan for the disks on the host by using the specific iSCSI command as with an anon-premises IBM Spectrum Virtualize Cluster.

Check the **multipath** output (run **multipath -ll**) to ensure that your VDisks are attached correctly through the multipath tool. A typical output of a VDisk is shown in Example 5-6.

Example 5-6 Linux multipath -ll output example

```
mpathch (3600507680181820bc800000000000009) dm-1 IBM     ,2145
size=500G features='1 queue_if_no_path' hwhandler='0' wp=rw
|-+- policy='round-robin 0' prio=50 status=active
| |- 26:0:0:5 sdf  8:80   active ready running
| `- 27:0:0:5 sdl  8:176  active ready running
`-+- policy='round-robin 0' prio=10 status=enabled
  |- 28:0:0:5 sdr  65:16  active ready running
  `- 29:0:0:5 sdx  65:112 active ready running
```

Windows host

The software iSCSI initiator is built in to the system on Windows 2008 and later. Access the iSCSI initiator from the Control Panel or search from the Start menu.

Discover the iSCSI target by using Send Targets or by using iSNS. For more information, see IBM Knowledge Center.

Connect to the discovered targets, as described in IBM Knowledge Center.

Now, the mapped volumes are visible to Windows disk management services. The system volumes can be initialized, formatted, and mounted. You can view the details of the discovered disks by using the Windows Command Prompt. An example output is shown in Example 5-7.

Example 5-7 Diskpart command example

```
DISKPART> list disk
 Disk ###  Status      Size     Free     Dyn  Gpt
 --------  ----------  -------  -------  ---  ---
 Disk 0    Online       149 GB    78 GB   *
 Disk 1    Online       149 GB    78 GB   *
 Disk 2    Online       565 MB   565 MB
 Disk 3    Online       337 MB   337 MB
DISKPART> select disk 2
Disk 2 is now the selected disk.
DISKPART> detail disk
IBM      2145             SCSI Disk Device
Disk ID: 00000000
Type   : iSCSI
Bus    : 0
Target : 2
LUN ID : 0
There are no volumes.
```

5.7 Configuring a site-to-site virtual private network IPSec tunnel for hybrid cloud connectivity in AWS Cloud

This section describes how to configure hybrid cloud connectivity between the AWS Cloud and the on-premises environment. This section also describes the lab setup and the steps to configure the site-to-site IPSec tunnel for communication between AWS Cloud and the on-premises site.

The virtual private network (VPN) IPSec site-to-site tunnel creates a secure communication network between the AWS Cloud infrastructure and on-premises infrastructure. Network communication between the private subnets is controlled by the access control list (ACL) that is populated when you create the VPN IPSec site-to-site tunnel.

AWS configuration for the VPN IPSec tunnel

Complete the following steps at the VPC level in AWS Cloud to establish the IPSec tunnel:

1. Create a customer gateway by logging in to the AWS console with resource provisioning privileges. Select **Services** at the upper left, and then, **VPC**. Select **Virtual Private Network (VPN)** in the pane on the left. Then, click the customer gateways and enter the required details.
2. Create the virtual private gateways by clicking the **Virtual private gateways** section in the VPC and configure the required details.
3. Attach a virtual private gateway to the VPC.
4. Create a site-to-site VPN connection in AWS Console by selecting the virtual private gateway and customer gateway parameters. Attach the virtual private gateway to the VPC in AWS.
5. After the site-to-site connection is complete, a configuration file is generated for the end-to-end point. This step creates two tunnels in the VPC. The same configuration file is used for the configuration at the other end of the tunnel.

5.8 Configuring replication from on-premises IBM Spectrum Virtualize to IBM Spectrum Virtualize for Public Cloud on AWS

This section describes how to configure replication from an on-premises solution that can be a FlashSystem or SAN Volume Controller system to an IBM Spectrum Virtualize for Public Cloud on AWS solution.

Our example uses a FlashSystem system in the on-premises data center and a 2-node IBM Spectrum Virtualize for Public Cloud on AWS as a DR storage solution.

This scenario uses IBM Spectrum Virtualize *Global Mirror with Change Volume* (GMCV) to replicate the data from the on-premises data center to AWS Cloud.

This implementation starts with the assumption that the IP connectivity between the on-premises data center and AWS Cloud is established through a *Multiprotocol Label Switching* (MPLS) or VPN connection. Because several methods are available to implement the IP connectivity, this section does not consider that specific configuration. For more information, contact your organizations's network technical specialist.

To configure the GMCV, complete the following steps:

1. Configure your IBM Spectrum Virtualize Private IP ports so that they are enabled for remote copy. This configuration is required on both sites, as shown in Figure 5-41.

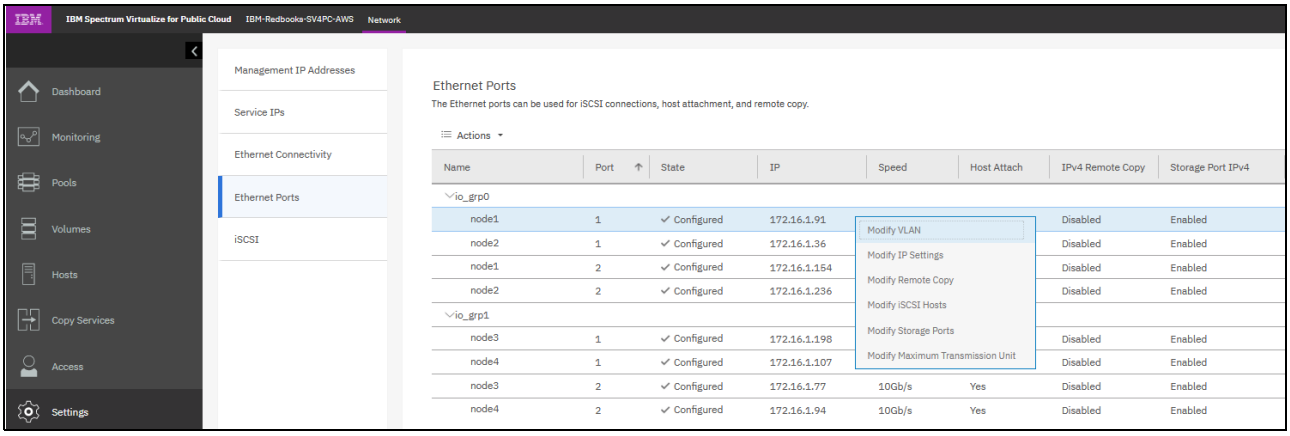


Figure 5-41 Remote copy IP port example

You are redirected to choose which copy group to use, as shown in Figure 5-42.

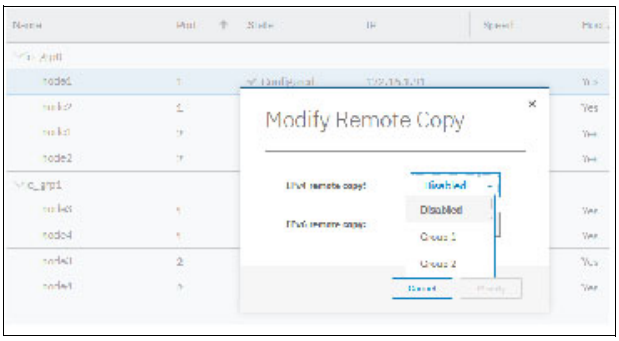


Figure 5-42 Group 1 configuration example

2. Repeat step 1 for all of the IP ports that you want to configure. A similar configuration is created, as shown in Figure 5-43.

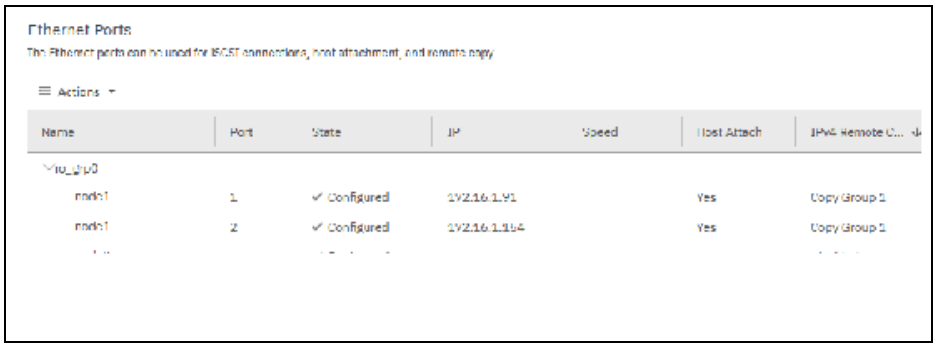


Figure 5-43 IBM Spectrum Virtualize configuration complete

3. Run the same configuration for the on-premises FlashSystem storage system or SAN Volume Controller, as shown in Figure 5-44 and Figure 5-45 on page 94.

> **Note:** It is important to understand what versions of IBM Spectrum Virtualize software are supported. For supported and interoperability versions, see IBM Spectrum Virtualize Family of Products Inter-System Metro Mirror and Global Mirror Compatibility Cross Reference.

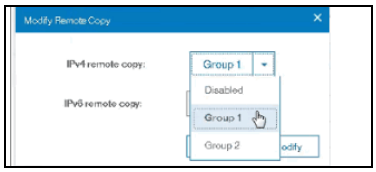


Figure 5-44 On-premises copy group example

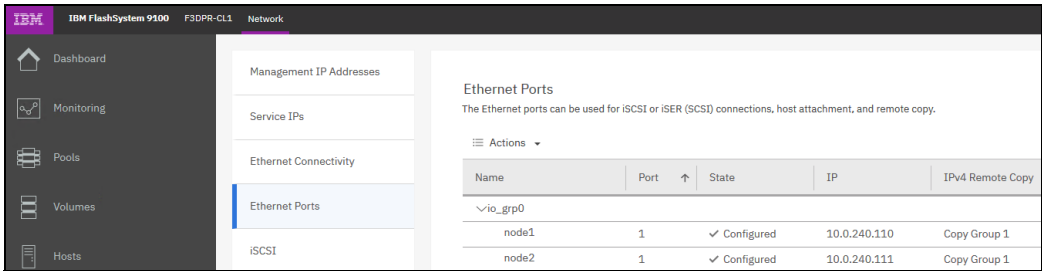


Figure 5-45 On-premises configuration completion example

4. Create a cluster partnership between the on-premises data center and IBM Spectrum Virtualize for Public Cloud on AWS from the on-premises GUI, as shown in Figure 5-46.

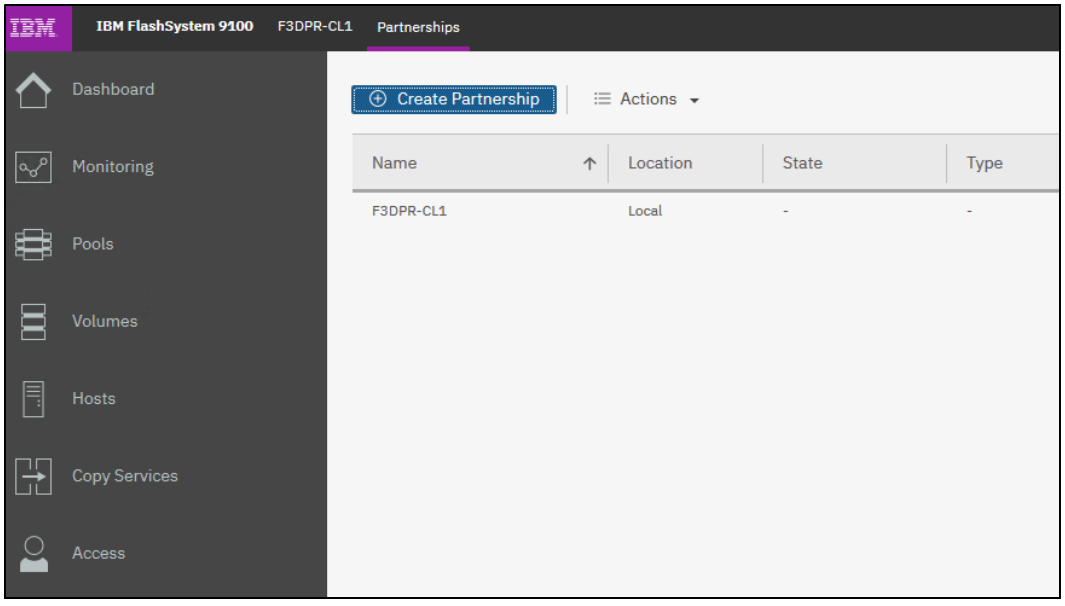


Figure 5-46 Create Partnership setup example

5. Complete the partnership creation from on-premises, as shown in Figure 5-47.

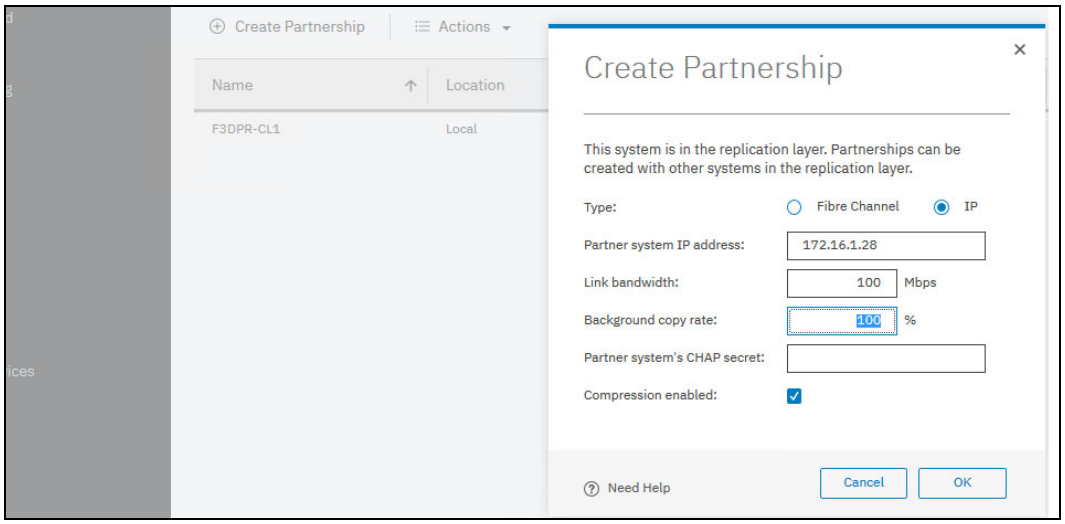


Figure 5-47 Inserting an IP address example

As shown in Figure 5-48, the partnership is partially complete. You must complete the partnership in the IBM Spectrum Virtualize on-premises GUI.

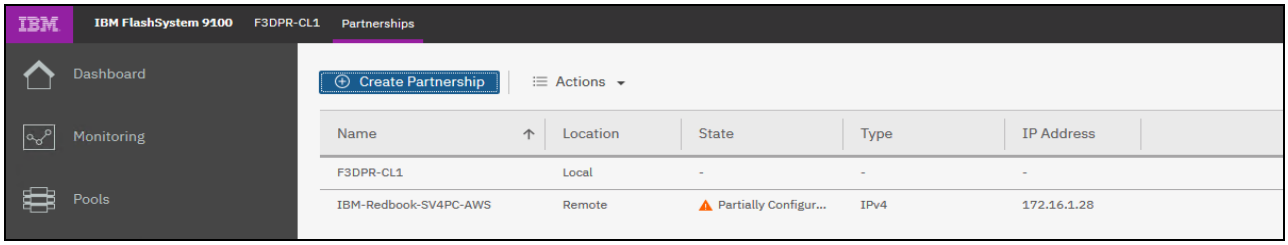


Figure 5-48 Partnership partially configured

6. Complete the partnership configuration in the IBM Spectrum Virtualize for Public Cloud on AWS side, as shown in Figure 5-46 on page 94 and Figure 5-47 on page 95 by providing on-premises cluster IP address.

   Now, your partnership is fully configured, as shown in Figure 5-49.

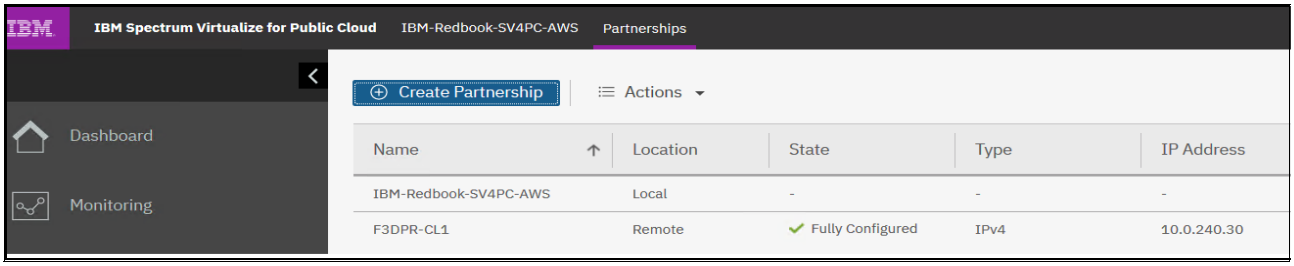


Figure 5-49 Fully configured example

Note: The connection might take a few seconds to synchronize, but double-clicking **Partnership** shows the confirmed status of the partnership quicker.

7. In our example, an on-premises 10 GiB volume is used with its Change Volume (CV) that must be replicated to a 10 GiB volume in the AWS Cloud instance that is defined in our IBM Spectrum Virtualize for Public Cloud installation. The on-premises volumes are thin-provisioned, but this use of thin-provisioned is not a specific requirement; rather, it is a choice. The CV can be thin-provisioned or fully provisioned, regardless of whether the master or auxiliary volume is thin-provisioned or space-efficient.

   The CV must store only the changes that accumulated during the cycle period. Therefore, it should use real capacity when possible, as shown in Figure 5-50.

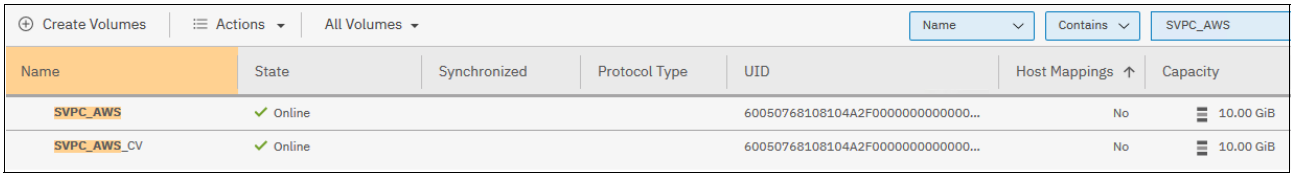


Figure 5-50 Volumes example

8. Create a volume remote copy relationship for a GMCV from the on-premises data center, as shown in Figure 5-51.

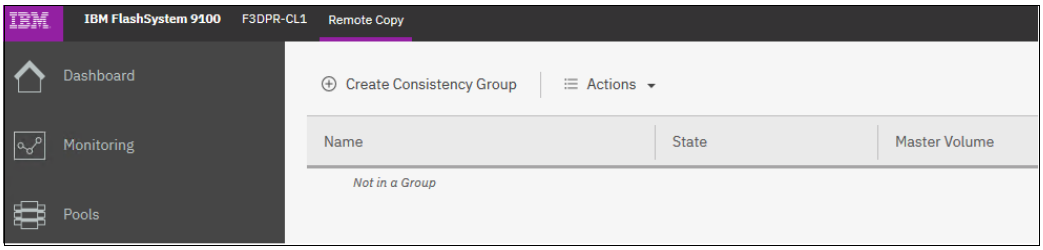


Figure 5-51 Creating a relationship

9. Select the type of relationship, as shown in Figure 5-52.

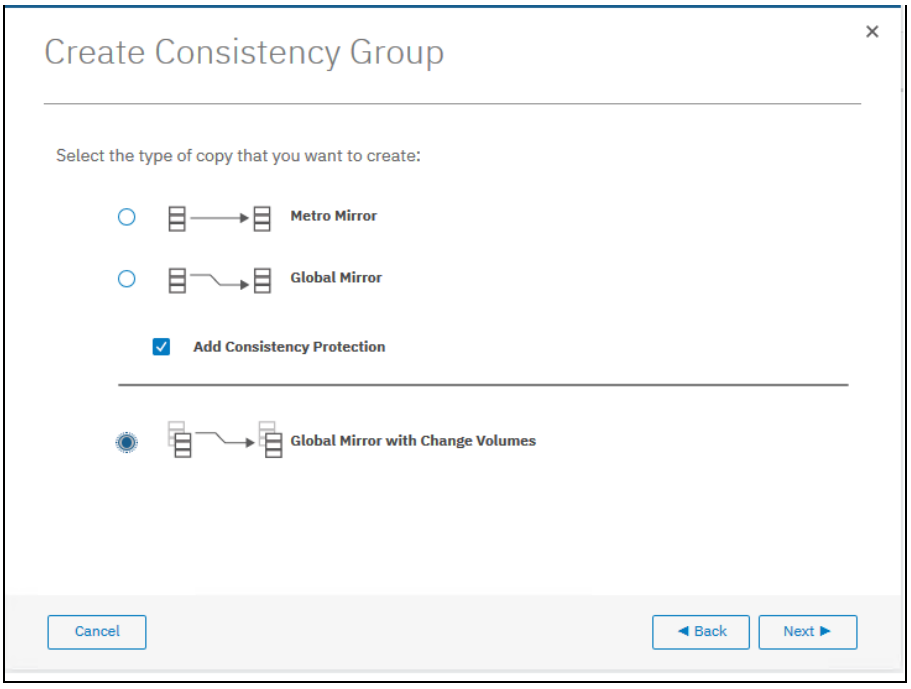


Figure 5-52 Global Mirroring with Change Volumes example

10. Select the remote system, as shown in Figure 5-53, and select the volumes that must be in the relationship, as shown in Figure 5-54.

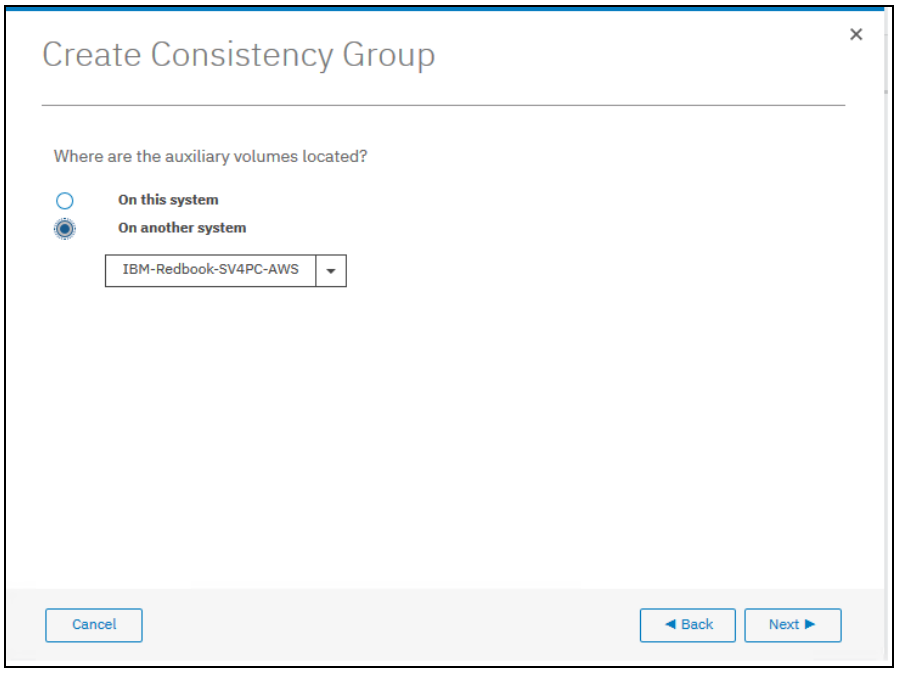


Figure 5-53 Remote system

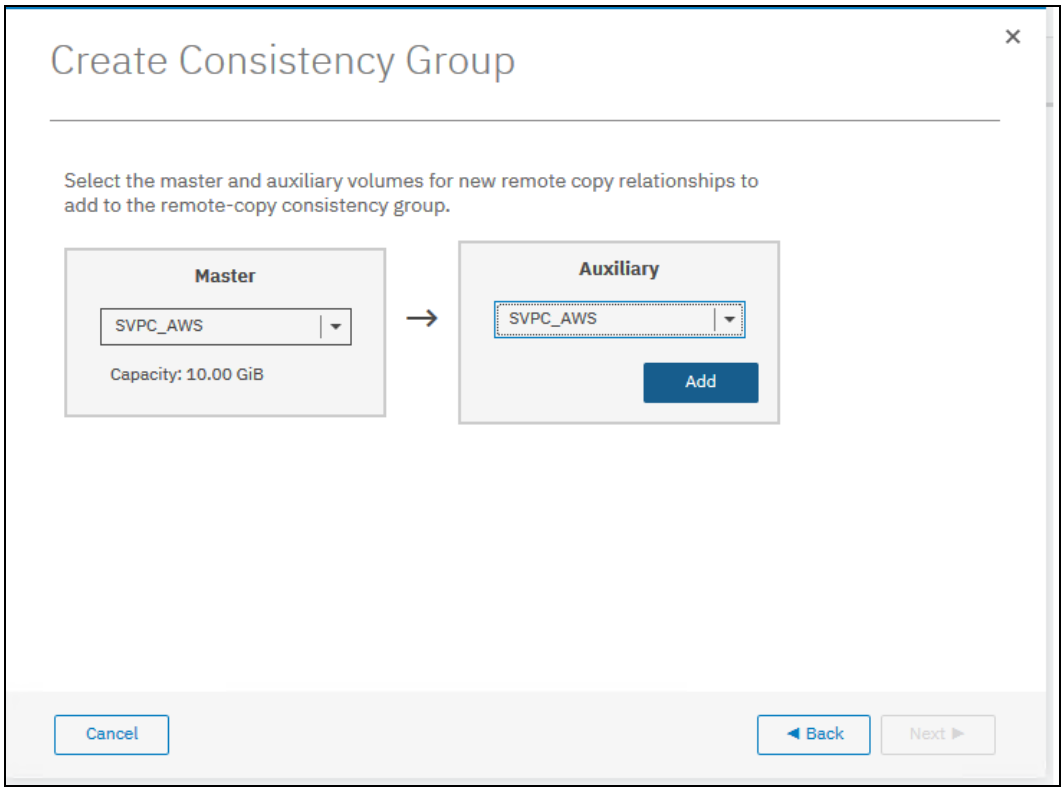


Figure 5-54 Master and auxiliary volumes example

In our example, we select **No, do not add a master change volume**, as shown in Figure 5-55. These volumes are added later.

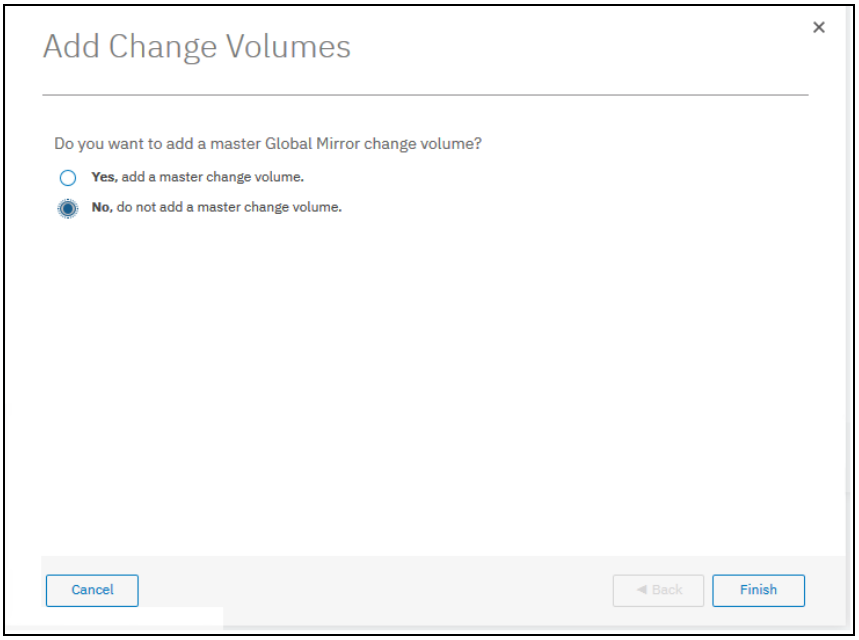


Figure 5-55 Do not add change volume example

We select **No, do not start copying**, as shown in Figure 5-56.

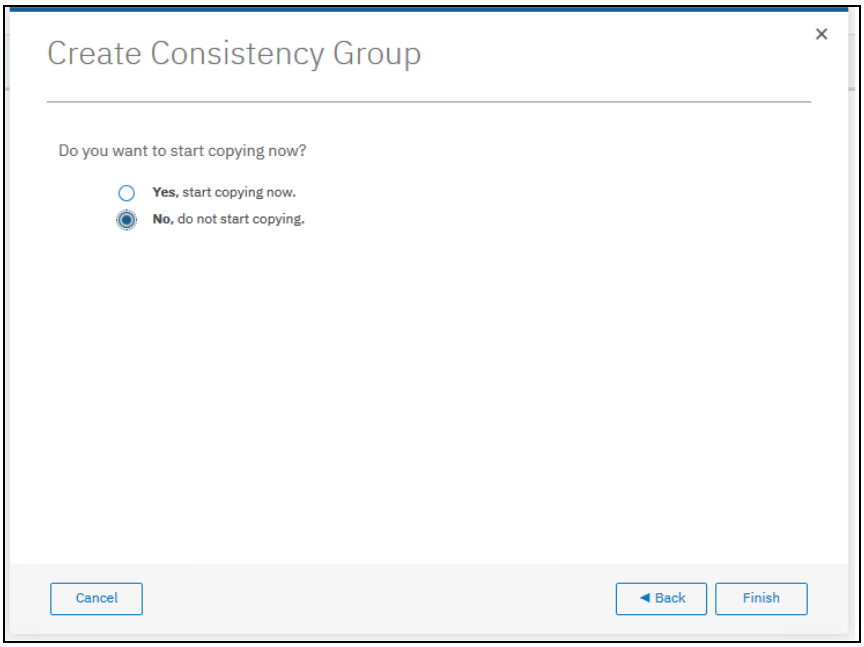


Figure 5-56 Do not start relationship example

11. Edit the relationship and set the **Cycling Mode** and **Cycling Period**.

12. Add the CV volumes to your relationship on both sides, as shown in Figure 5-57, Figure 5-58, and Figure 5-59 on page 100.

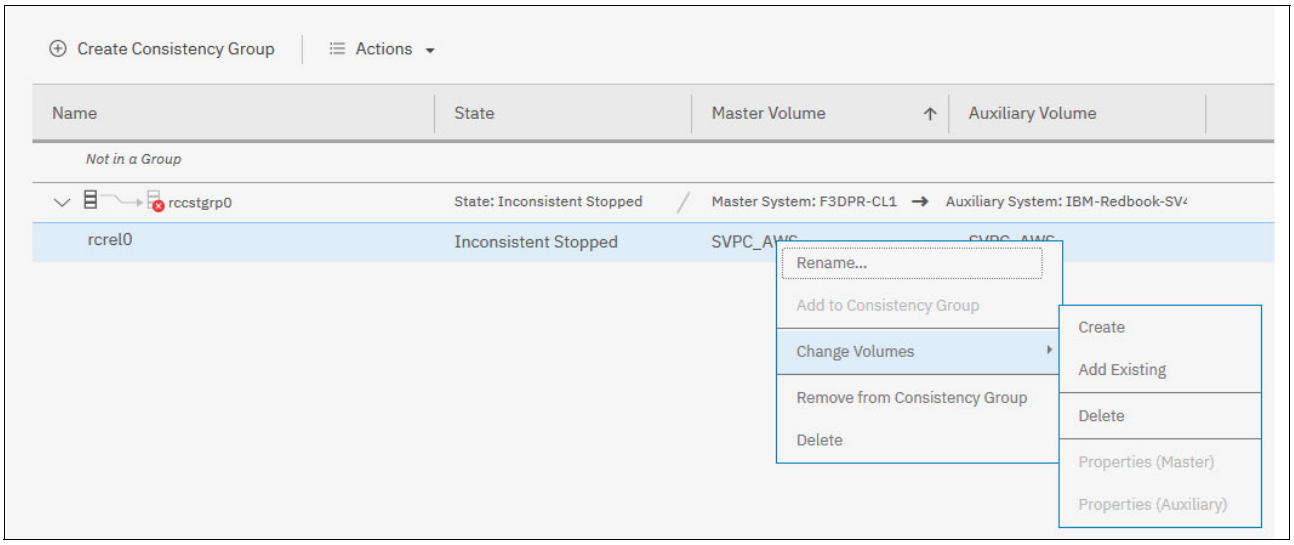


Figure 5-57 Adding a change volume from the on-premises site

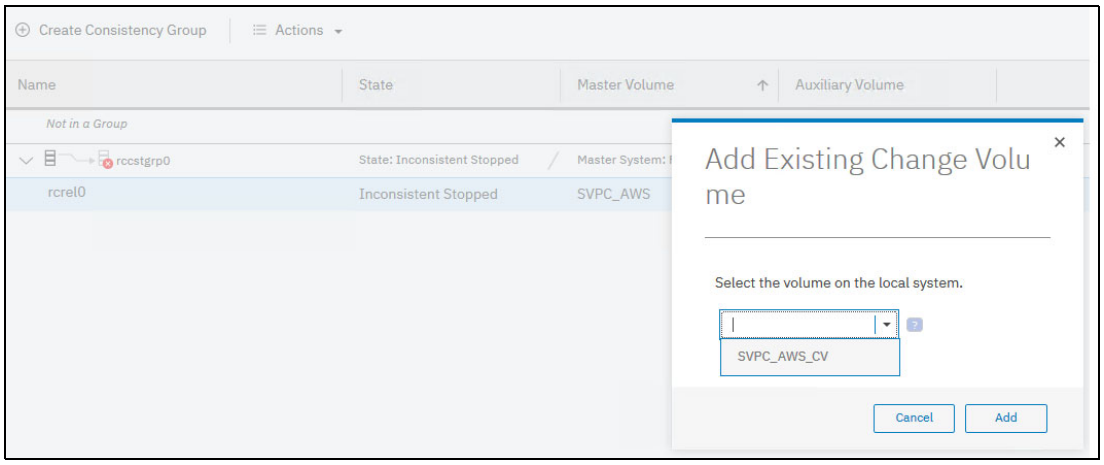


Figure 5-58 Selecting the change volume from the on-premises site

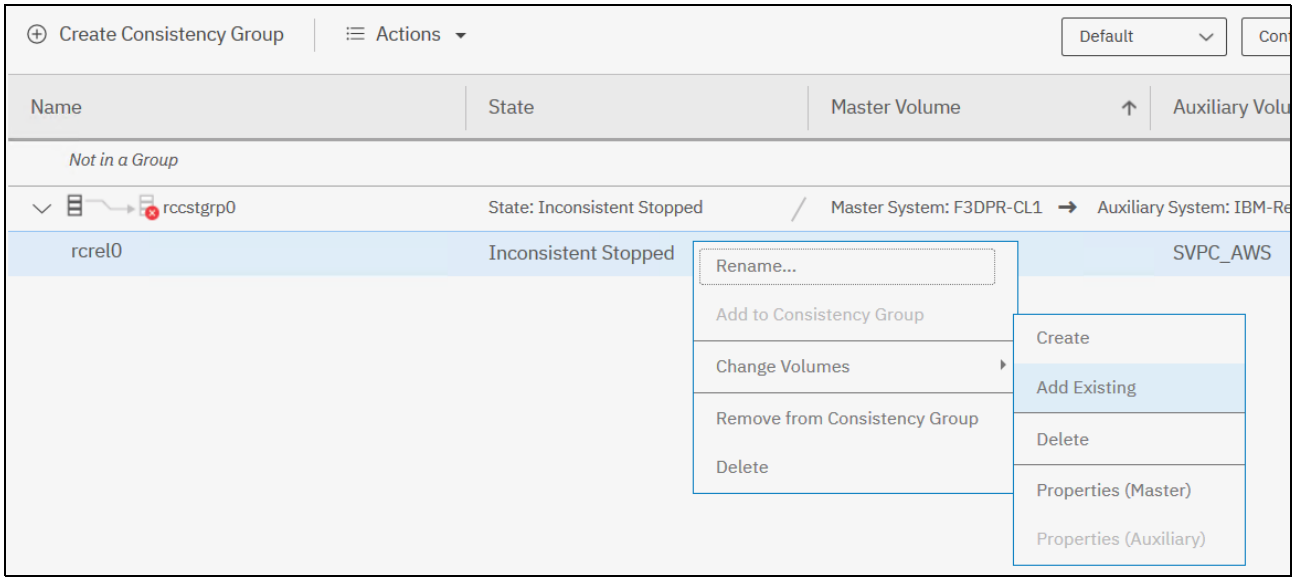


Figure 5-59 Adding the change volume to the AWS Cloud site

13. Start your relationship from the on-premises site, as shown in Figure 5-60.

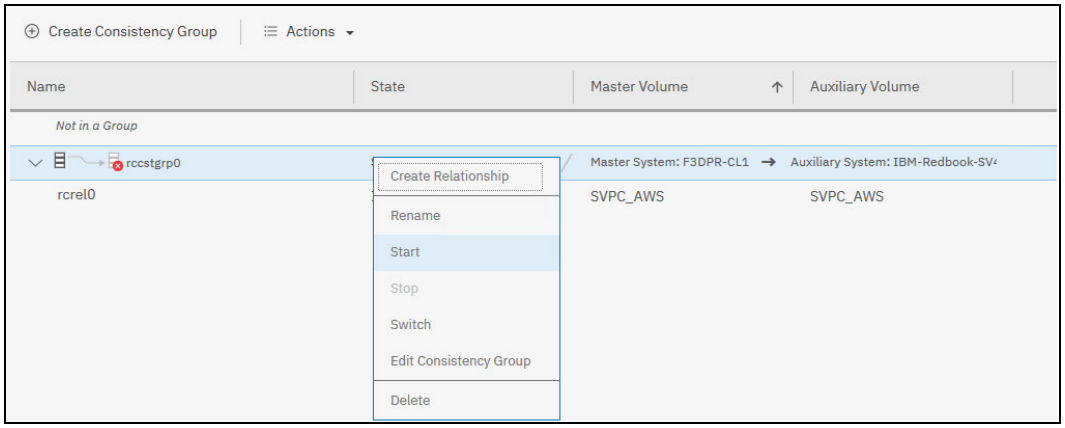


Figure 5-60 Starting the relationship

14. Create a GM consistency group and add your relationship to it, as shown in Figure 5-61 and Figure 5-62.

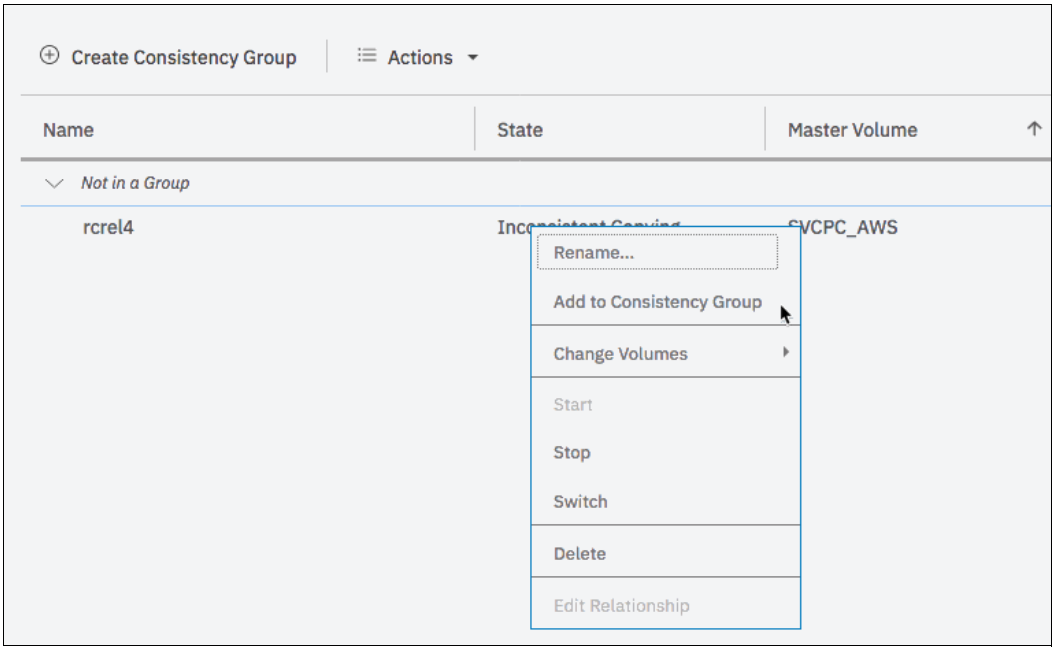


Figure 5-61 Adding a consistency group

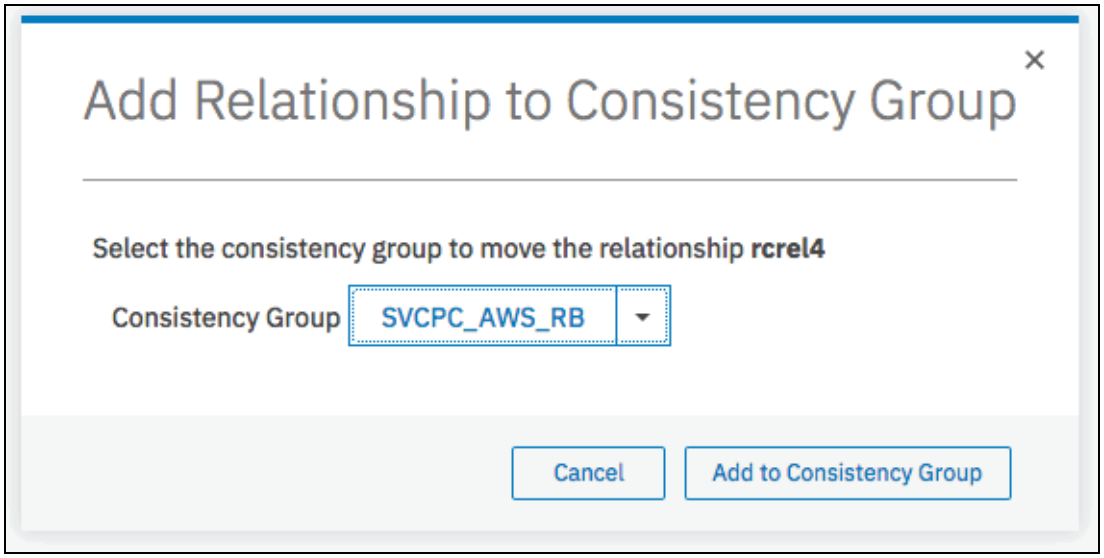


Figure 5-62 Add Relationship to Consistency Group

You can see the status of your consistency group, as shown in Figure 5-63 and Figure 5-64.

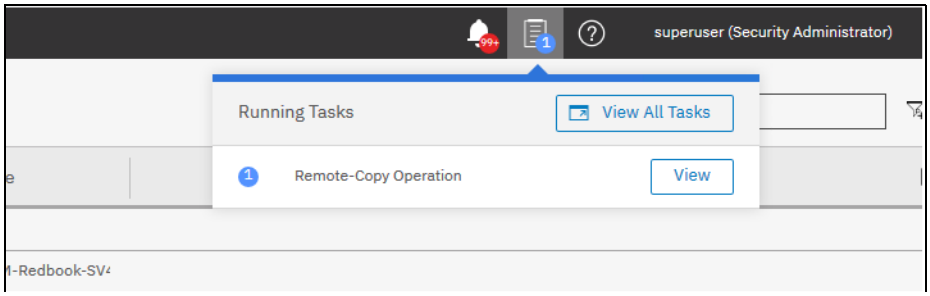


Figure 5-63 Consistency group status

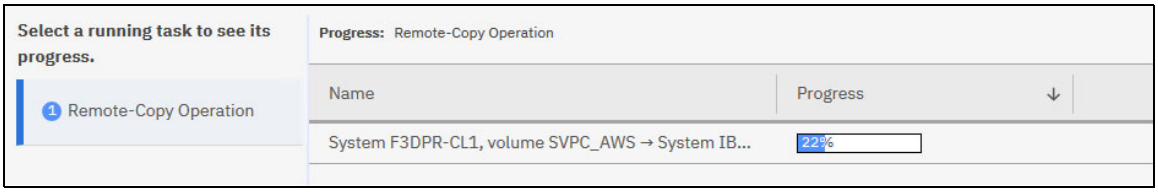


Figure 5-64 Copying status

In our example, we show the status by using the IBM Spectrum Virtualize for AWS Cloud GUI.

When the copy approaches completion, the CV algorithm starts to prepare a freeze time in accordance with the cycling windows. When your copy reaches 100%, a FlashCopy is taken from the auxiliary volume to the auxiliary-CV to be used if a real disaster or DR test occurs. At 100%, the status is `consistent copying`.

What we described is only an example of how to configure a GMCV relationship from an on-premises solution to an IBM Spectrum Virtualize for Public Cloud on AWS solution. It can be valuable to configure a snapshot (FlashCopy) of your GMCV auxiliary volume to be used for DR testing or other purposes.

These steps were completed by using the GUI, but they can also be done by using the CLI.

For more information about how to manage FlashSystem, IBM Spectrum Virtualize, or SAN Volume Controller copy functions, see the following publications:

- *Implementing the IBM Storwize V7000 with IBM Spectrum Virtualize V8.2.1*, SG24-7938
- *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521
- *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933

6

Supporting the solution

This chapter provides guidance about how support for this solution is built. This solution is made up of the following basic support segments:

- Amazon Web Services (AWS)
- IBM Storage support teams.

It is important to understand who to contact if a problem occurs.

This chapter includes the following topics:

6.1 Who to call for support

The IBM Spectrum Virtualize for Public Cloud on AWS solution is composed of several components, much like the traditional storage offerings. However, when deployed in the public cloud on AWS, IBM Spectrum Virtualize is an application running in a stack on AWS Cloud. Therefore, to ensure that your level of support matches the level that is required to support your application, see AWS Support for more information about the various levels of support that are available to an AWS user.

In this solution, the cloud provider is responsible for providing the infrastructure, network components, storage, support, and assistance for this solution. The cloud user and any involved third parties are responsible for deploying and configuring the solution from the network layer up to the operating system, and the software that is installed. IBM Systems Support is responsible for providing support and assistance with the IBM Spectrum Virtualize application.

The solution is composed of multiple parties with different roles and responsibilities. In situations where the cloud environment is having a problem, the client collects as much information about the problem that is known and opens a ticket with the cloud provider.

In the situations where the IBM Spectrum Virtualize is having a problem (and the cloud environment is not), the client collects as much information about the problem and diagnostic data surrounding the event and opens a Problem Management Record (PMR) with IBM.

If the origin of the problem is unclear, the client collects as much information about the problem and diagnostic data surrounding the event and opens a case with the cloud provider and IBM.

6.2 Working with AWS support

AWS is a service that provisions the infrastructure, network, operating systems, and back-end storage that is used in this solution. AWS support is responsible for helping resolve problems and answer questions for products and services that are acquired through the AWS Marketplace Console. AWS cloud users can open support cases in the support section of the AWS Marketplace Console.

6.3 Working with IBM Spectrum Virtualize Support

Support engagement for the IBM Spectrum Virtualize for Public Cloud component of the solution is the same as it is for all of the other IBM Spectrum Virtualize based solutions. IBM Support can be engaged by using one of the following methods:

- Visit the IBM Service requests and the PMRs web page to open a PMR
- By phone: 1-800-IBM-SERV
- IBM Call Home

After you receive a PMR or ticket number, you can begin working with support to troubleshoot the problem. You might be asked to collect diagnostic data or to open a remote support session for an IBM Support representative to dial in to the system and investigate.

6.3.1 Email notifications and the IBM Call Home function

The IBM Call Home function of IBM Spectrum Virtualize uses the email notification that is sent to the specific IBM Support center. Therefore, the configuration is similar to sending emails to the specific person or system owner.

Complete the following steps to configure email notifications and emphasizes what is specific to IBM Call Home:

1. Prepare your contact information that you want to use for the email notification and verify the accuracy of the data. From the left menu of the GUI, select **Settings** → **Notifications** (see Figure 6-1).

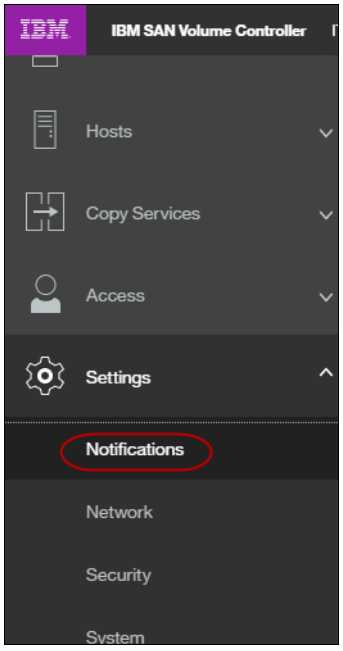


Figure 6-1 Notifications menu

2. Select **Email** and then, click **Enable Notifications** (see Figure 6-2).

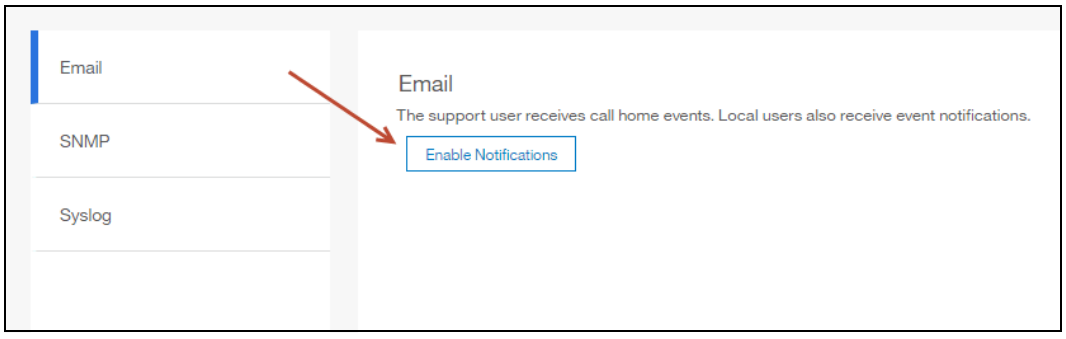


Figure 6-2 Configuration of email notifications

For the correct functionality of email notifications, ask your network administrator if Simple Mail Transfer Protocol (SMTP) is enabled in the network and is not blocked by firewalls or the foreign destination `@de.ibm.com` is not blocked.

Be sure to test the accessibility to the SMTP server by using the **`telnet`** command (port 25 for a non-secured connection, port 465 for Secure Sockets Layer (SSL)-encrypted communication) by using any server in the same network segment.

3. After clicking **Next** in the welcome window, enter the information about the location of the system (see Figure 6-3) and contact information of IBM Spectrum Virtualize administrator (see Figure 6-4 on page 107) to be contactable by IBM Support. *Always* keep this information current.

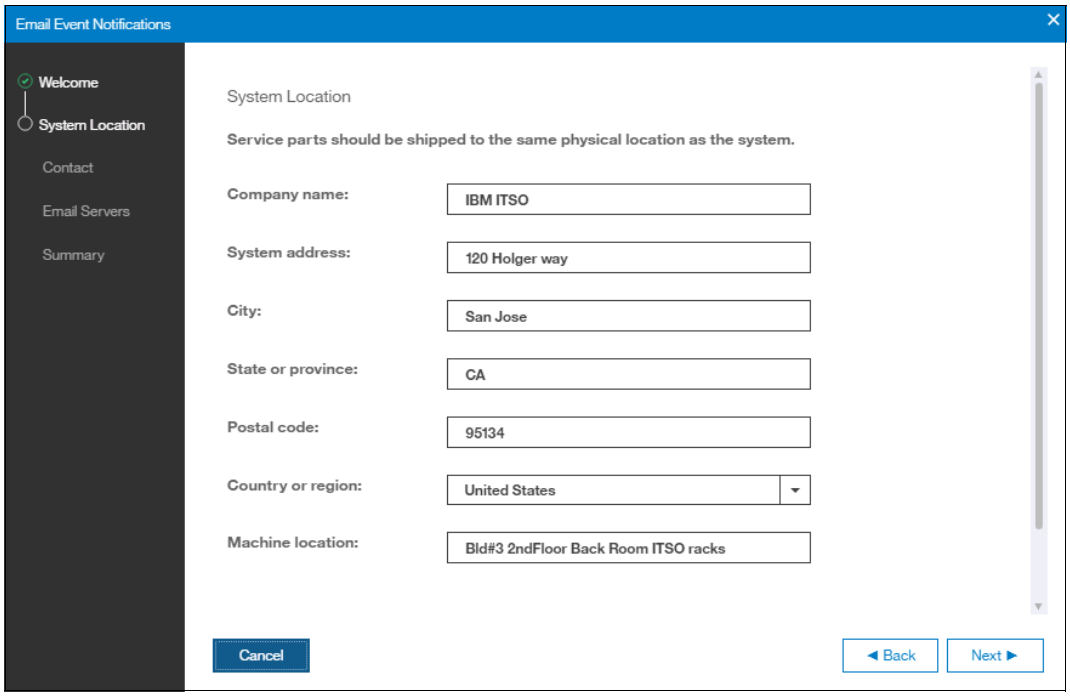


Figure 6-3 Location of the device

Figure 6-4 shows the contact information of the owner.

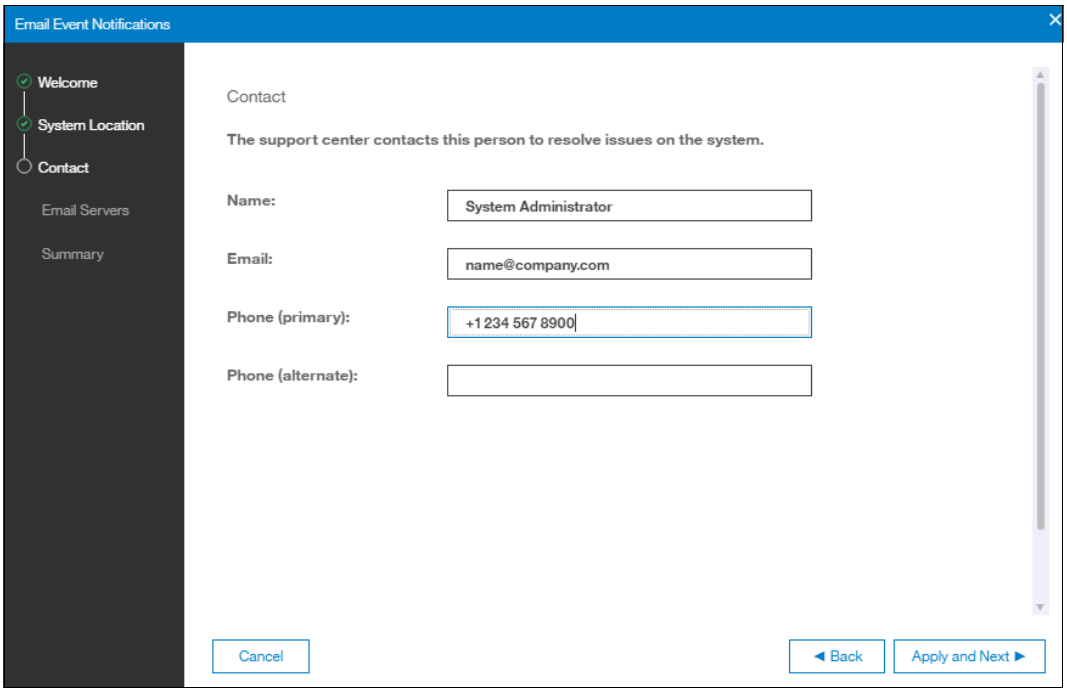


Figure 6-4 Contact information

4. Configure the IP address of your company SMTP server, as shown in Figure 6-5. When the correct SMTP server is provided, you can test the connectivity by using the **Ping** option to its IP address. You can configure more SMTP servers by clicking the **+** at the end of the entry line.

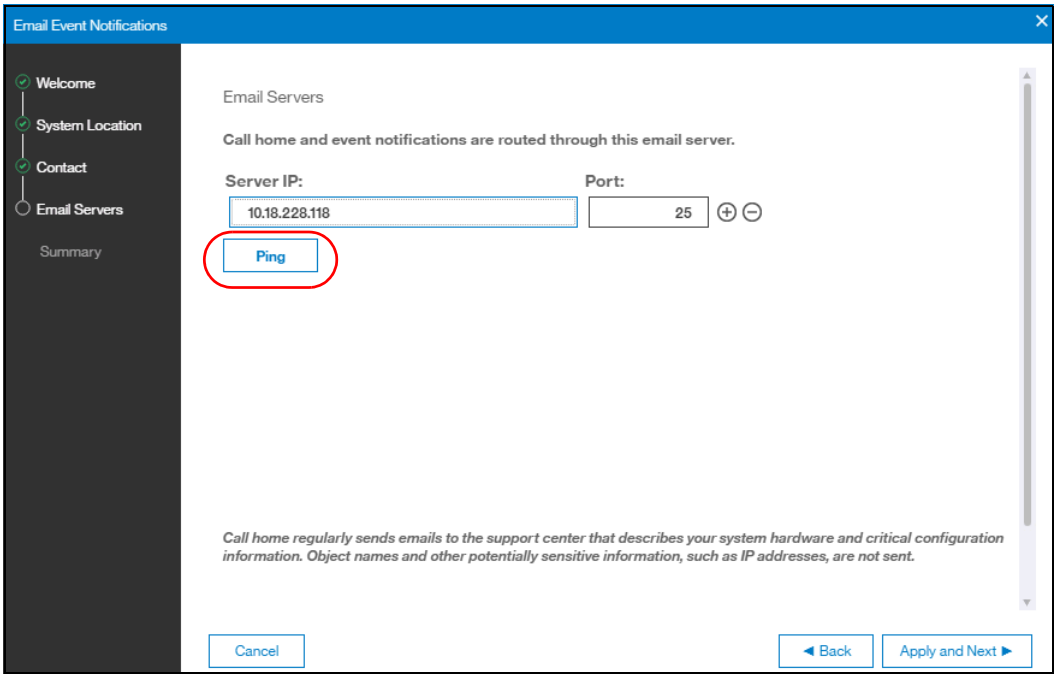


Figure 6-5 Configure email servers and inventory reporting

5. The summary window opens. Verify it and click **Finish**. You are returned to the Email Settings window in which you can verify email addresses of IBM Support (`callhome0@de.ibm.com`) and optionally add local users who also must receive notifications (see Figure 6-6).

   The default support email address `callhome0@de.ibm.com` is predefined by the system to receive Error Events and Inventory. We recommend not changing these settings.

   You can modify or add local users by using Edit mode after the initial configuration was saved.

   The **Inventory Reporting** function is enabled by default for IBM Call Home. Rather than reporting a problem, an email is sent to IBM that describes your system hardware and critical configuration information. Object names and other information, such as IP addresses, are not included. The inventory email is sent weekly by default, which allows an IBM Cloud® service to analyze and inform you if the hardware or software that you use requires an update because of any known issues.

   Figure 6-6 shows the configured email notification and IBM Call Home settings.

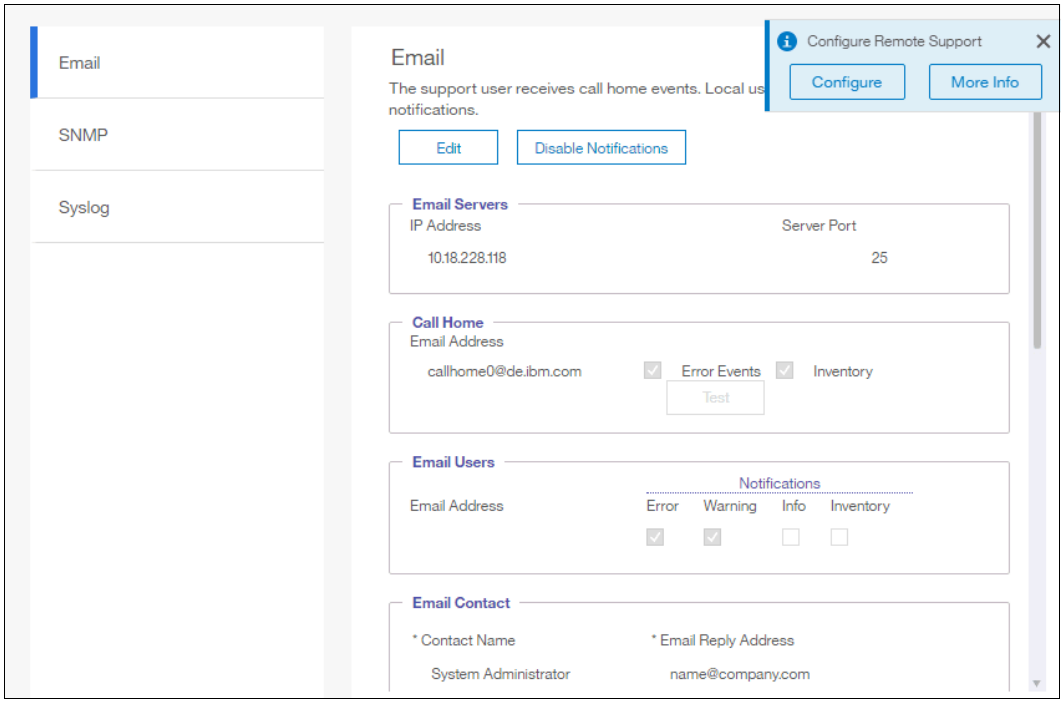


Figure 6-6 Setting email recipients and alert types

6. After completing the configuration wizard, we can test the email function. To do so, enter Edit mode, as shown in Figure 6-7. In the same window, email recipients can be defined or any contact and location information can be changed as needed.

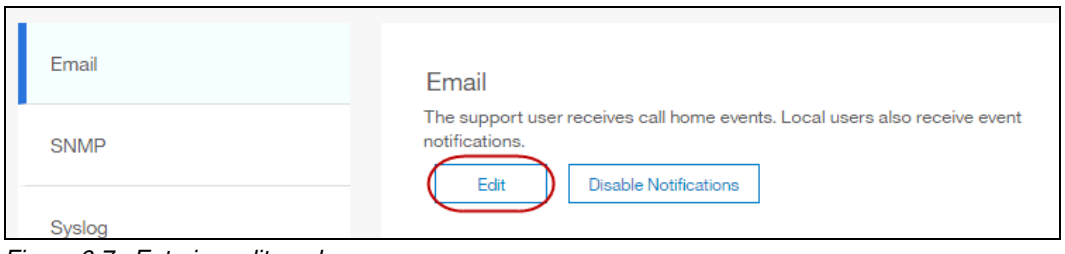


Figure 6-7 Entering edit mode

We strongly suggest that you keep the sending inventory option enabled to IBM Support; however, it might not be of interest to local users, although inventory content can serve as a basis for inventory and asset management.

7. In Edit mode, we can change any of the configured settings. After these parameters are edited, recipients are added, or the connection is tested, the configuration can be saved so that any changes take effect (see Figure 6-8).

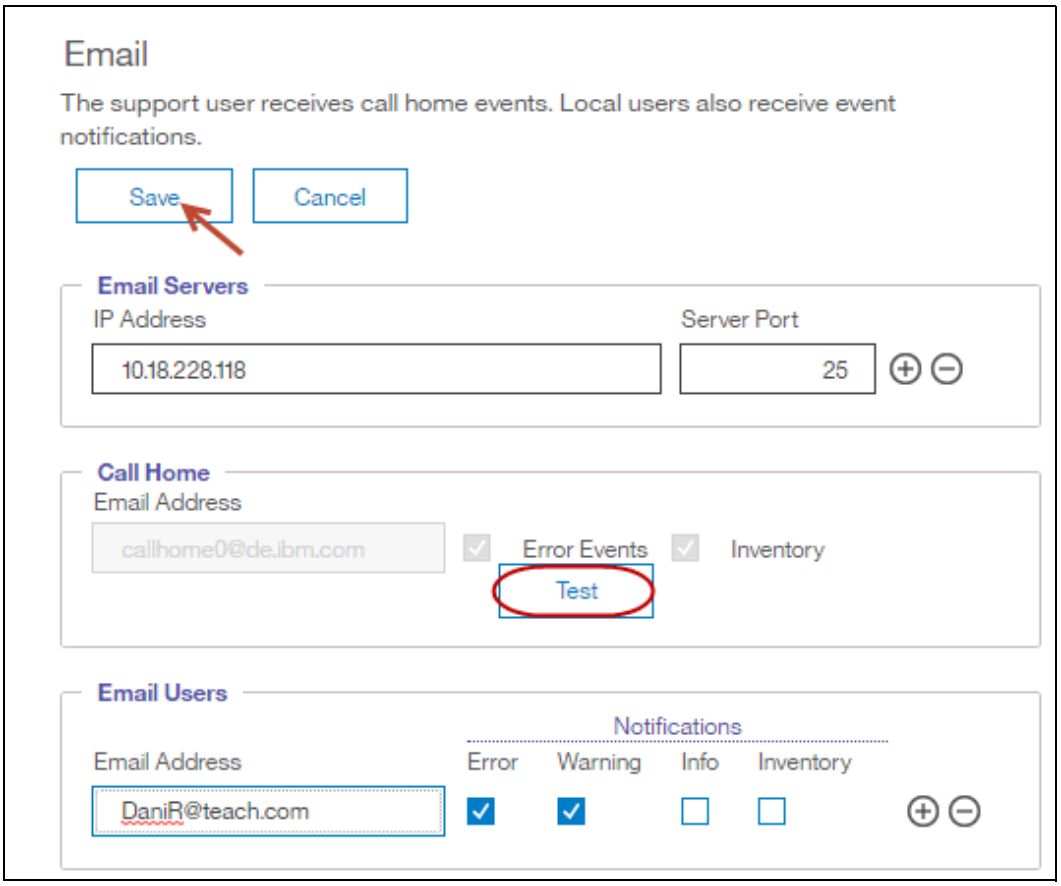


Figure 6-8 Saving modified configuration

Note: The Test option is available for new email users after first saving and then editing again.

6.3.2 Disabling and enabling notifications

At any time, you can temporarily or permanently disable email notifications, as shown in Figure 6-9 on page 110. This feature is good practice when performing activities in your environment that might generate expected errors on your IBM Spectrum Virtualize, such as SAN reconfiguration or replacement activities. After the planned activities, remember to re-enable the email notification function. The same results can be achieved by using the CLI **`svctask stopmail`** and **`svctask startmail`** commands.

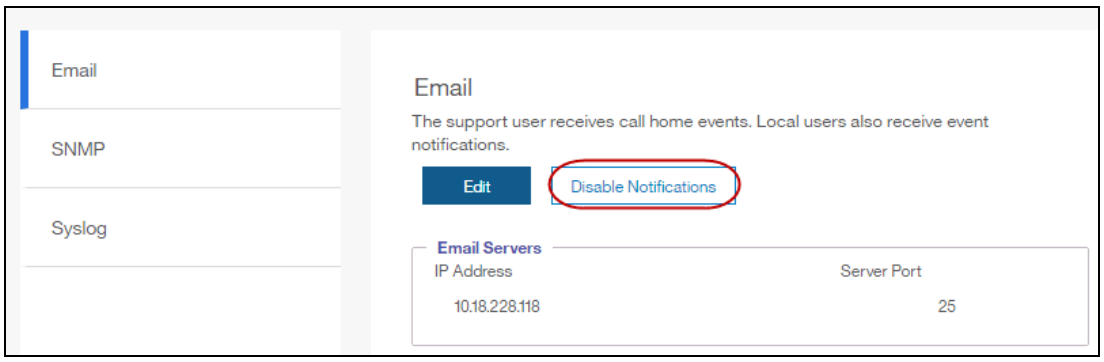


Figure 6-9 Disabling or enabling email notifications

6.3.3 Collecting Diagnostic Data for IBM Spectrum Virtualize

Occasionally, if a problem occurs and the IBM Support Center is contacted, they most likely ask you to provide the support package. You can collect and upload this package from the **Settings** → **Support** menu.

Collecting information by using the GUI

To collect information by using the GUI, complete the following steps:

1. Click **Settings** → **Support** and the **Support Package** tab (see Figure 6-10). Then, click **Upload Support Package**.

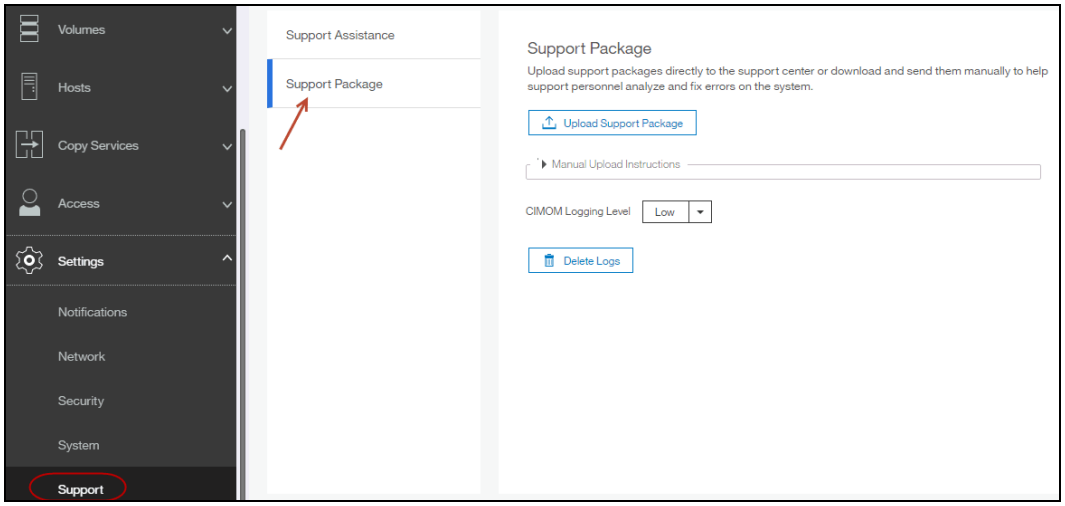


Figure 6-10 Support Package option

Assuming the problem that was encountered was an unexpected node restart that logged a 2030 error, we collect the default logs plus the most recent statesave from each node to capture the most relevant data for support.

Note: When a node unexpectedly restarts, it first dumps its current statesave information before it restarts to recover from an error condition. This statesave is critical for support to analyze what occurred. Collecting a snap type 4 creates statesaves at the time of the collection, which is not useful for understanding the restart event.

2. From the Upload Support Package window, four options are available for data collection. You were contacted by IBM Support because your system called home or you manually opened a call with IBM Support; therefore, you receive a PMR number. Enter that PMR number into PMR field and select the snap type, which is often referred to as an *option 1, 2, 3, 4 snap*, as requested by IBM Support (see Figure 6-11). In our example, we enter our PMR number, select **snap type 3** (option 3) because this choice automatically collects the statesave that is created at the time the node restarted. Click **Upload**.

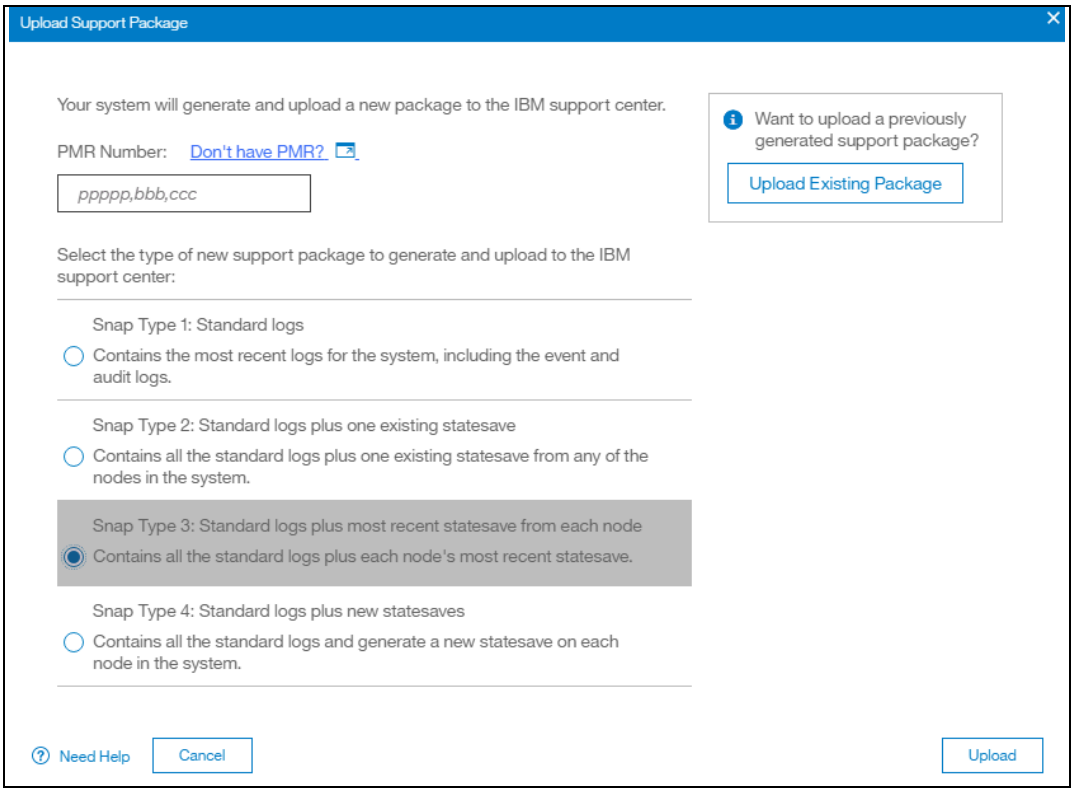


Figure 6-11 Upload Support Package window

The procedure to create the snap on an IBM Spectrum Virtualize system, including the latest statesave from each node, starts. This process might take a few minutes (see Figure 6-12).

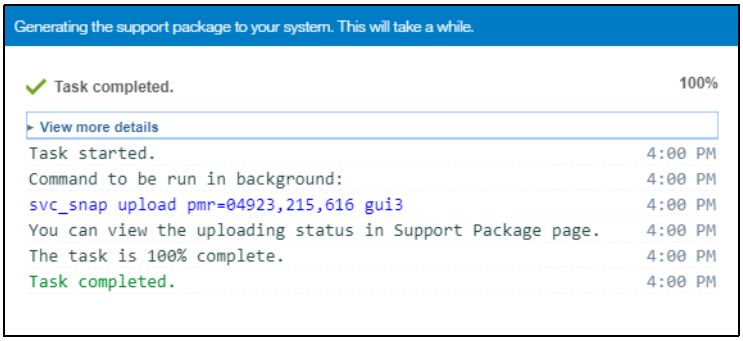


Figure 6-12 Task detail window

Collecting logs by using the CLI

Complete the following steps to use the CLI to collect and upload a support package as requested by IBM Support:

1. Log in to the CLI to run the **svc_snap** command that matches the type of snap that is requested by IBM Support:
   - Standard logs (type 1):

     ```
     svc_snap upload pmr=ppppp,bbb,ccc gui1
     ```

   - Standard logs plus one existing statesave (type 2):

     ```
     svc_snap upload pmr=ppppp,bbb,ccc gui2
     ```

   - Standard logs plus most recent statesave from each node (type 3):

     ```
     svc_snap upload pmr=ppppp,bbb,ccc gui3
     ```

   - Standard logs plus new statesaves:

     ```
     svc_livedump -nodes all -yes
     svc_snap upload pmr=ppppp,bbb,ccc gui3
     ```

2. We collect the type 3 (option 3) and automatically upload it to the PMR number that is provided by IBM Support, as shown in Example 6-1.

   Example 6-1 The svc_snap command

   ```
   ssh superuser@10.18.228.64
   Password:
   IBM_2145:ITSO DH8_B:superuser>>svc_snap upload pmr=04923,215,616 gui3
   ```

3. If you do not want to automatically upload the snap to IBM, do not specify the '`upload pmr=ppppp,bbb,ccc`' part of the commands. In this case, when the snap creation completes, it creates a file that is named in the following format:

   ```
   /dumps/snap.<panel_id>.YYMMDD.hhmmss.tgz
   ```

   It takes a few minutes for the snap file to complete (longer if statesaves are included).

4. The generated file can then be retrieved from the GUI under the **Settings** → **Support** → **Manual Upload Instructions** twisty → **Download Support Package**. Click **Download Existing Package**, as shown in Figure 6-13.

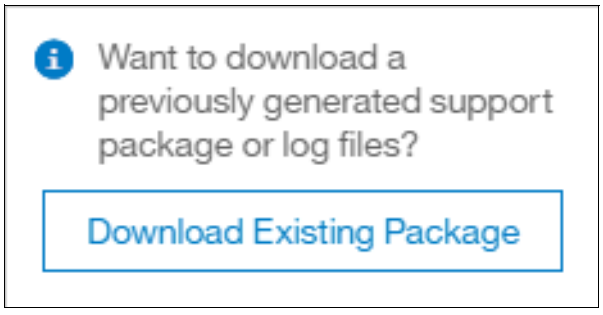


Figure 6-13 Downloaded Existing Package

5. A new window opens. Click in the Filter box and enter `snap`. Then, press Enter. A list of snap files is shown (see Figure 6-14). Locate the exact name of the snap that was generated by the **`svc_snap`** command that was issued earlier. Click to select that file and then, click **Download**.

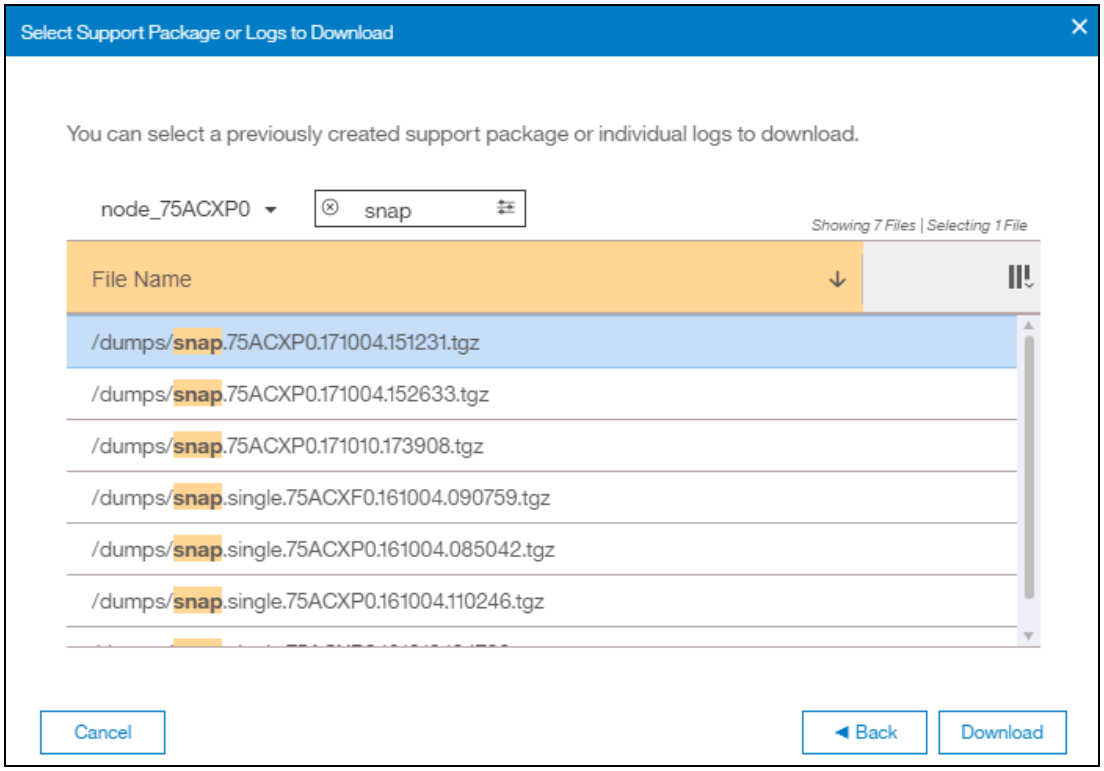


Figure 6-14 Filtering on snap to download

6. Save the file to a folder of your choice on your workstation.

6.3.4 Uploading files to the Support Center

If you chose not to have IBM Spectrum Virtualize upload the support package automatically, it might still be uploaded for analysis by using the Enhanced Customer Data Repository (ECuRep). Any uploads are associated with a specific PMR. The PMR is also known as a *service request* and is a mandatory requirement when uploading.

To upload information, complete the following steps:

1. By using a browser, navigate to the ECuRep Secure Upload web page (see Figure 6-15).

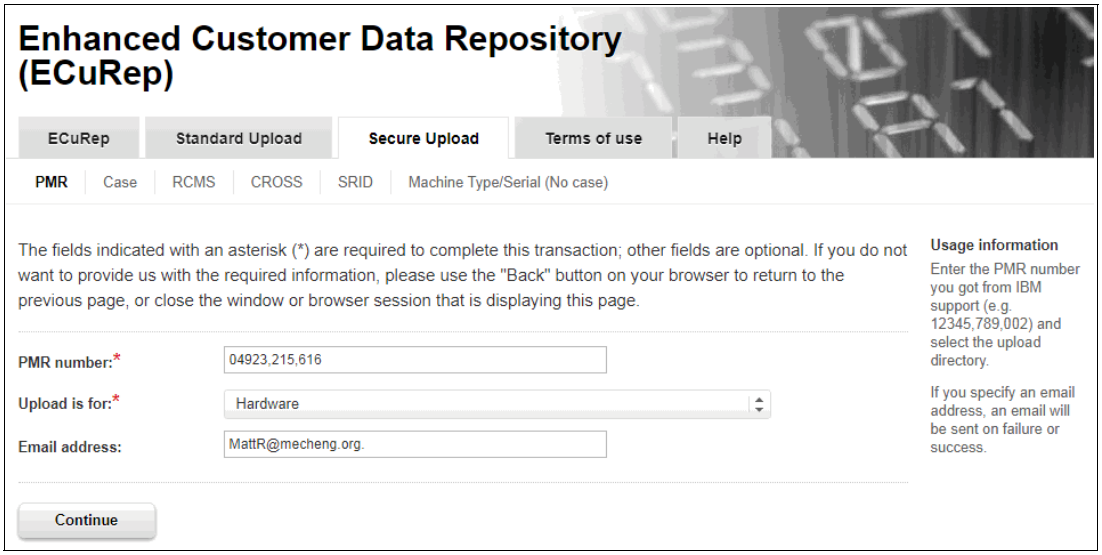


Figure 6-15 ECuRep details

2. Complete the following required fields:
   - PMR number as provided by IBM Support for your specific case. This number should be in the format of `ppppp,bbb,ccc` (for example, `04923,215,616`) using a comma (,) as a separator.
   - Upload is for. Select **Hardware** from the drop-down menu.

   Although completing the Email address field is not required, we suggest entering your email address to be automatically notified of a successful or unsuccessful upload.

3. When completed, click **Continue**. The Input window opens (see Figure 6-16).

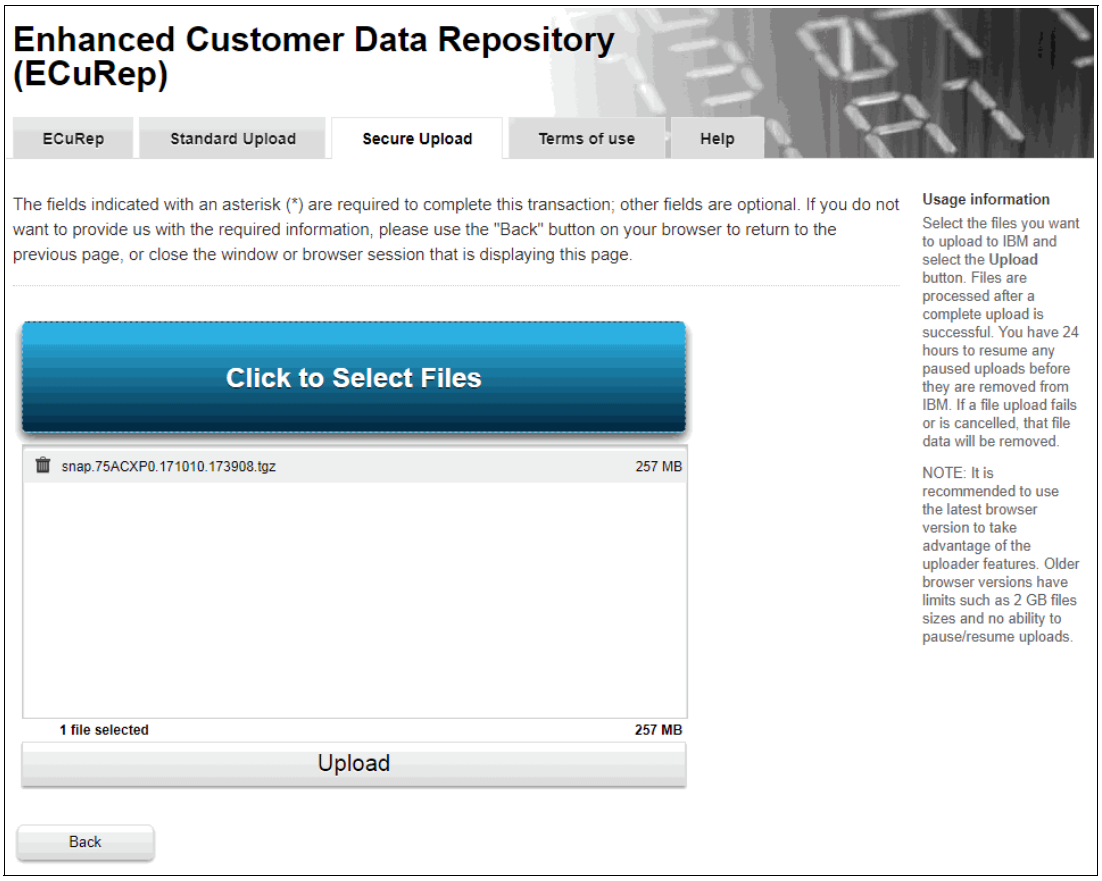


Figure 6-16 ECuRep File upload

4. After the files are selected, click **Upload** to continue, and follow the directions.

6.3.5 Service Assistant Tool

The Service Assistant Tool (SAT) is a web-based GUI that is used to service individual node canisters, primarily when a node has a fault and is in a service state. A node is not an active part of a clustered system while it is in service state.

Typically, an IBM Spectrum Virtualize cluster is initially configured with the following IP addresses:

- One service IP address for each IBM node.
- One cluster management IP address, which is set when the cluster is created.

The SAT is available even when the management GUI is not accessible. The following information can be obtained and tasks can be accomplished by using the Service Assistance Tool:

- Status information about the connections and the nodes
- Basic configuration information, such as configuring IP addresses
- Service tasks, such as restarting the Common Information Model (CIM) object manager (CIMOM) and updating the WWNN

- Details about node error codes
- Details about the hardware, such as IP address and Media Access Control (MAC) addresses.

The SAT GUI is available by using a service assistant IP address that is configured on each node. It also can be accessed through the cluster IP addresses by appending `/service` to the cluster management IP.

If the clustered system is down, the only method of communicating with the nodes is through the SAT IP address directly. Each node can have a single service IP address on Ethernet port 1 and should be configured for all nodes of the cluster, including any Hot Spare Nodes.

To open the SAT GUI, enter one of the following URLs into a web browser:

- `http(s)://<cluster IP address of your cluster>/service`
- `http(s)://<service IP address of a node>/service`

Complete the following steps to access the SAT:

1. When you are accessing SAT by using `<cluster IP address>/service`, the configuration node canister SAT GUI login window opens. Enter the Superuser Password, as shown in Figure 6-17.

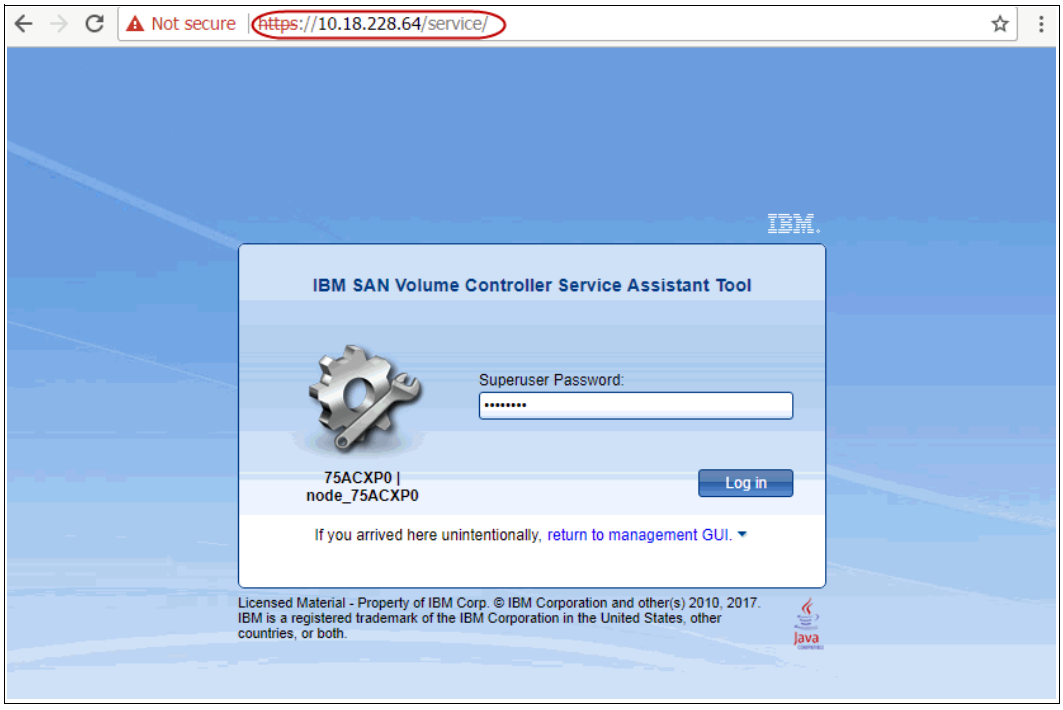


Figure 6-17 Service Assistant Tool Login GUI

2. After you are logged in, you see the Service Assistant Home window, as shown in Figure 6-18. The SAT can view the status and run service actions on other nodes, in addition to the node into which the user is logged.

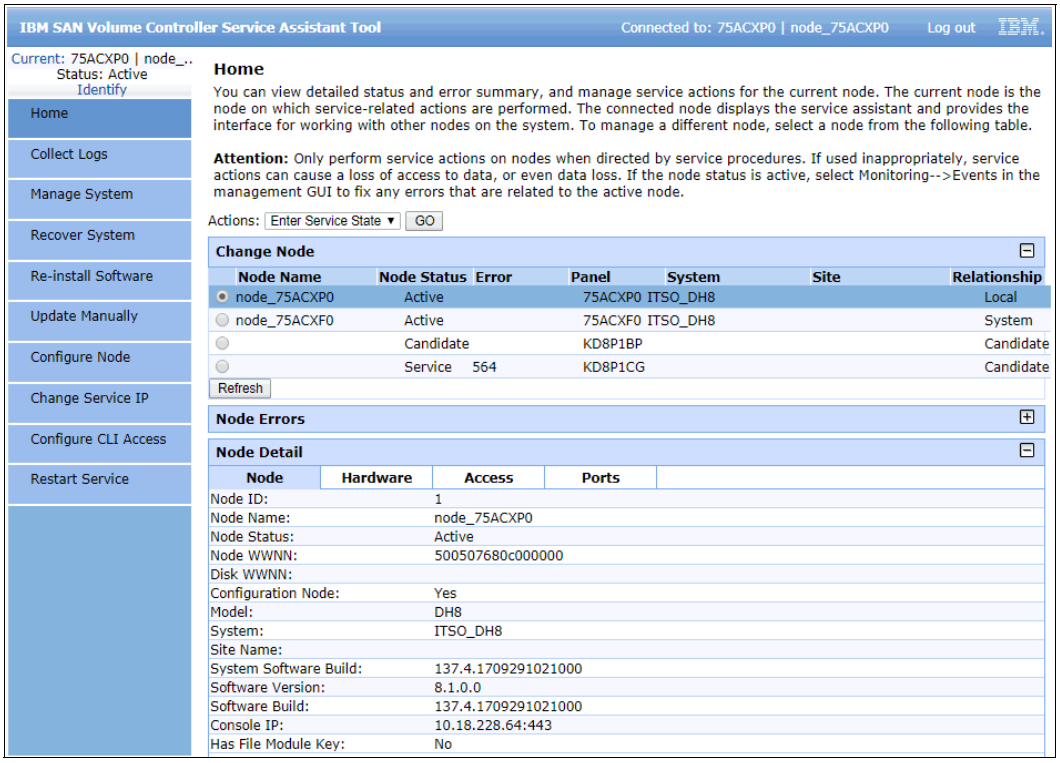


Figure 6-18 Service Assistant Tool GUI

3. The current selected Spectrum Virtualize node is displayed in the upper left corner of the GUI. As shown in Figure 6-18, this node is ID 1. Select the wanted node in the Change Node section of the window. You see the details in the upper left change to reflect the selected node.

Note: The SAT GUI provides access to service procedures and shows the status of the nodes. It is advised that these procedures are carried out only if directed to do so by IBM Support.

For more information about how to use the SA Tool, see hIBM Knowledge Center.r

6.3.6 Remote Support Assistance

Introduced with V8.1, Remote Support Assistance allows IBM Support to remotely connect to the Spectrum Virtualize by way of a secure tunnel to perform analysis, log collection, or software updates. The tunnel can be enabled ad hoc by the client or enable a permanent connection, if wanted.

Note: Clients who purchased Enterprise Class Support (ECS) are entitled to IBM Support that uses Remote Support Assistance to quickly connect and diagnose problems. However, because IBM Support might choose to use this feature on non-ECS systems at their discretion, we recommend configuring and testing the connection on all systems.

If you are enabling Remote Support Assistance, ensure that the following prerequisites are met:

- IBM Call Home is configured with a valid email server.
- A valid service IP address is configured on each node on the IBM Spectrum Virtualize.
- If your IBM Spectrum Virtualize is behind a firewall or if you want to route traffic from multiple storage systems to the same place, a Remote Support Proxy server is configured. Before you configure remote support assistance, the proxy server is installed and configured separately. During the setup process for support assistance, the IP address is specified and the port number for the proxy server on the remote support centers window.
- If you do not have firewall restrictions and the IBM Spectrum Virtualize nodes are directly connected to the internet, request your network administrator to allow connections to 129.33.206.139 and 204.146.30.139 on Port 22.
- Uploading support packages and downloading software require direct connections to the internet. A DNS server must be defined on your IBM Spectrum Virtualize for both of these functions to work.
- To ensure that support packages are uploaded correctly, configure the firewall to allow connections to the following IP addresses on port 443: 129.42.56.189, 129.42.54.189, and 129.42.60.189.
- To ensure that software is downloaded correctly, configure the firewall to allow connections to the following IP addresses on port 22:
  - 170.225.15.105
  - 170.225.15.104
  - 170.225.15.107
  - 129.35.224.105
  - 129.35.224.104
  - 129.35.224.107

Figure 6-19 shows a window that opens in the GUI after updating to V8.1 in which you are prompted to configure your IBM Spectrum Virtualize for Remote Support. You can choose not to enable it, open a tunnel when needed, or open a permanent tunnel to IBM.

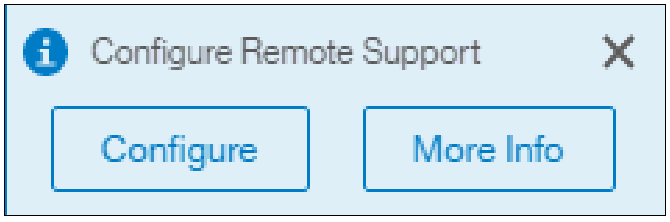


Figure 6-19 Prompt to configure Remote Support Assistance

From the prompt window, we can choose to configure or learn some more about the feature or close the window by clicking the **X**. Figure 6-20 shows how we can find the Setup Remote Support Assistance if the window is closed.

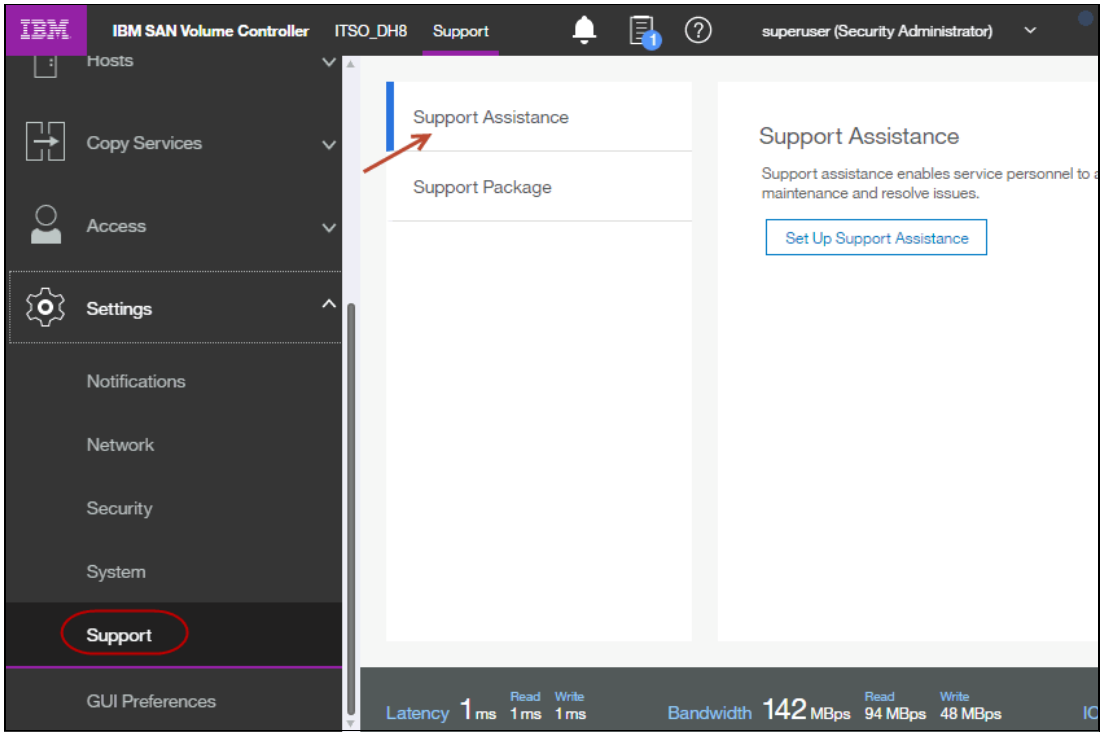


Figure 6-20 Remote Support Assistance menu

Choosing to set up support assistance opens a wizard to guide us through the configuration. Figure 6-21 on page 120 shows the first wizard window, where we can choose not to enable remote assistance by selecting **I want support personnel to work on-site only** or enable remote assistance by choosing **I want support personnel to access my system both on-site and remotely**. We chose to enable remote assistance and click **Next**.

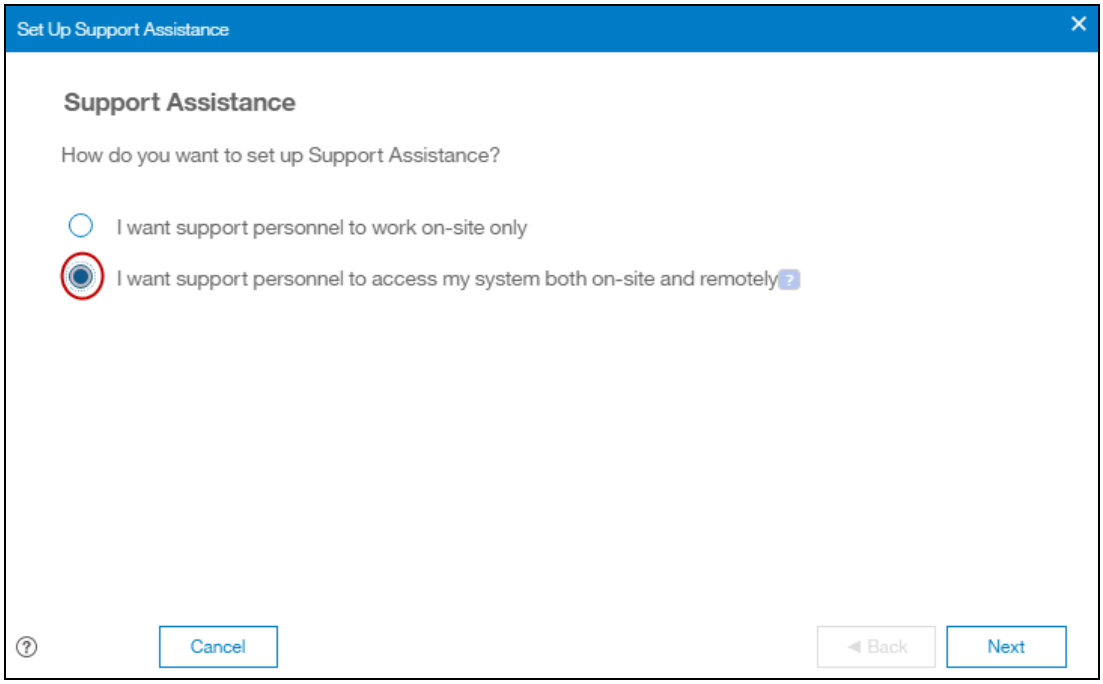


Figure 6-21 Remote Support wizard enable or disable

In the next window (see Figure 6-22), the IBM Support centers IP addresses and SSH port that must be open in your firewall are listed. Here, we can also define a Remote Support Assistance Proxy if multiple IBM Spectrum Virtualize clusters are in the same cloud, which enables firewall configurations being required for only the Proxy Server and not every storage system. Because we do not have a proxy server, and the field is left blank. Click **Next**.

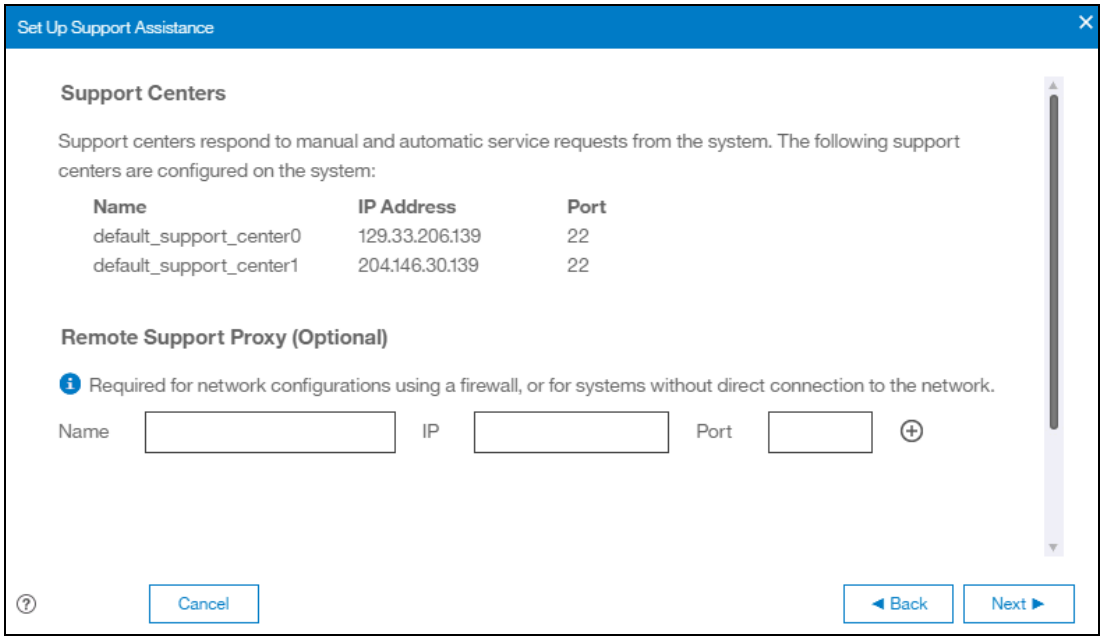


Figure 6-22 Remote Support wizard proxy setup

In the next window, we are prompted to decide whether we want to open a tunnel to IBM permanently, which allows IBM to connect to your IBM Spectrum Virtualize cluster at any time, or On Permission Only basis, which requires a storage administrator to log on to the GUI and enable the tunnel when required. We select this Permission option, as shown in Figure 6-23. Then, we click **Finish**.

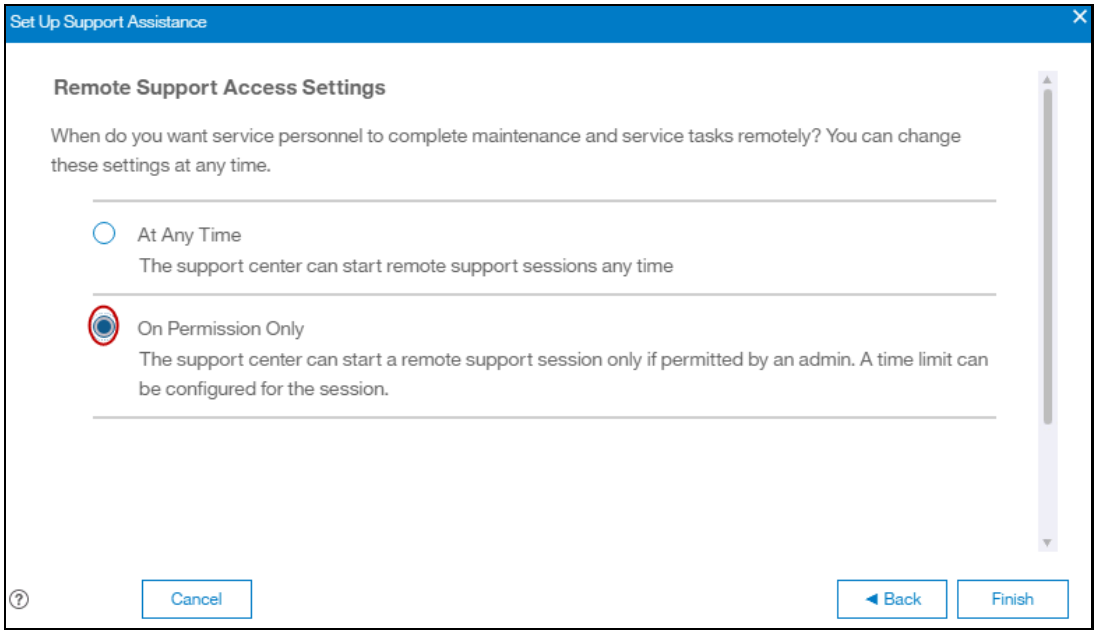


Figure 6-23 Remote Support wizard access choice

When we complete the remote support setup process, we can view the status of any remote connection, start a new session, test the connection to IBM, and reconfigure the setup. As shown in Figure 6-24, we successfully tested the connection. Now, we click **Start New Session** to open a tunnel for IBM Support to connect.

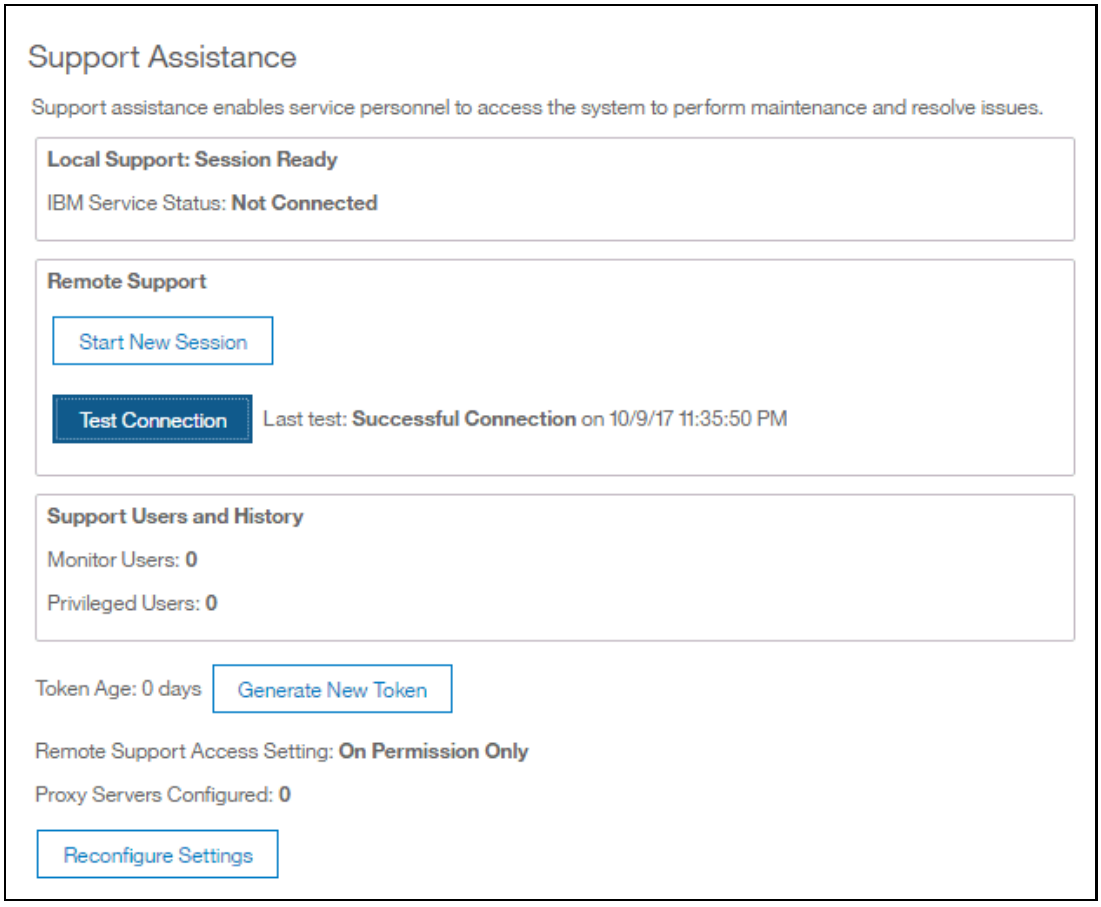


Figure 6-24 Remote Support Status and session management

A window opens in which we are prompted to indicate how long we want the tunnel to remain open if no activity occurs by setting a timeout value. Then, as shown in Figure 6-25, the connection establishes and is waiting for IBM Support to connect.

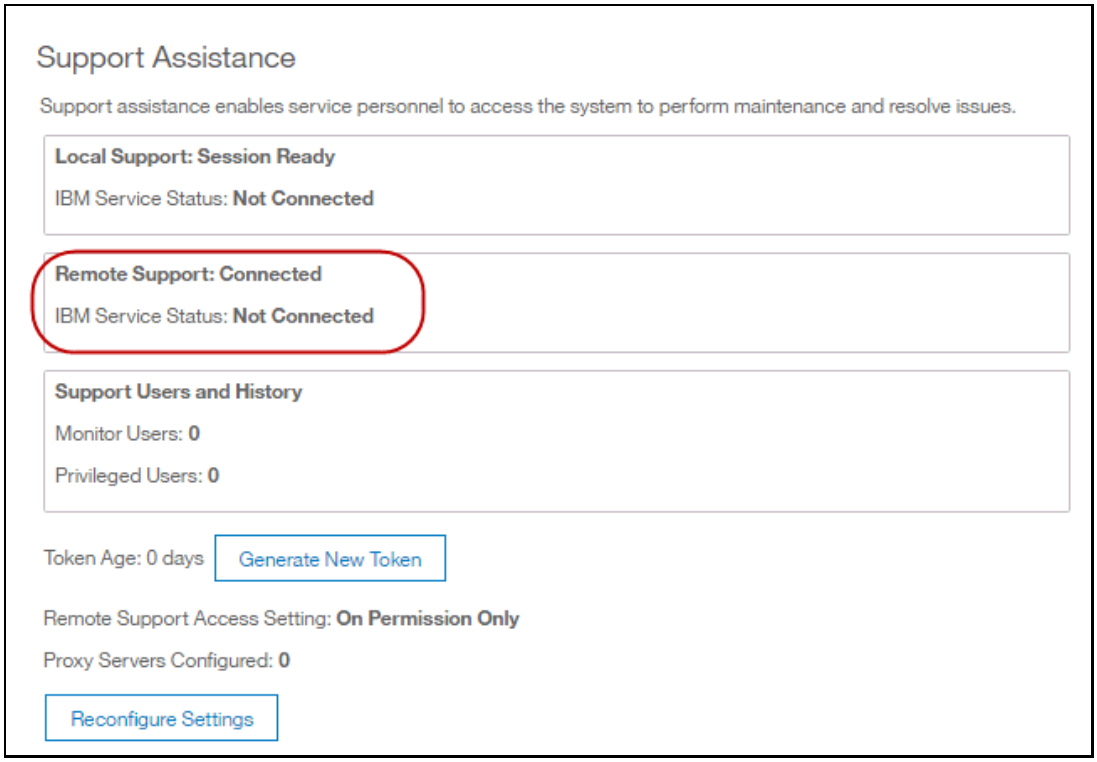


Figure 6-25 Remote Assistance tunnel connected

Related publications

The publications that are listed in this section are considered suitable for a more detailed description of the topics that are covered in this paper.

IBM Redbooks

The following IBM Redbooks publications provide more information about the topics in this document. Some publications that are referenced in this list might be available in softcopy only:

- *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521
- *Introduction and Implementation of Data Reduction Pools and Deduplication*, SG24-8430
- *Implementing IBM Spectrum Virtualize for Public Cloud Version 8.3*, REDP-5466
- *Implementing the IBM Storwize V7000 with IBM Spectrum Virtualize V8.2.1*, SG24-7938
- *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.1*, SG24-7933

You can search for, view, download, or order these documents and other Redbooks, Redpapers, web docs, draft s, and more materials, at the following website:

ibm.com/redbooks

Online resources

The following websites are also relevant as further information sources:

- Amazon Total Cost of Ownership (TCO) Calculator:

  https://aws.amazon.com/tco-calculator/
- Amazon Web Services (AWS) pricing:

  https://aws.amazon.com/ec2/pricing/on-demand/
- IBM FlashCopy solution:

  https://www.ibm.com/us-en/marketplace/data-protection-and-recovery
- IBM replication interoperability matrix:

  https://www-01.ibm.com/support/docview.wss?uid=ssg1S1003646
- IBM Spectrum Virtualize 8.2.1 configuration limits page:

  https://www-01.ibm.com/support/docview.wss?uid=ibm10741421
- IBM Storage Advisor Tool (STAT):

  http://www.ibm.com/support/docview.wss?uid=ssg1S4000935

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

REDP-5588-00

ISBN 0738458961

Printed in U.S.A.