

Cyber Resiliency with IBM Storage Sentinel and IBM Storage Safeguarded Copy

Nezih Boyacioglu

Gerd Franke

Thomas Gerisch

David Green

Vasfi Gucer

Guillaume Legmar

Markus Standau

Daniel Thompson

Christopher Vollmar

Axel Westphal



Storage



IBM Redbooks

**Cyber Resiliency with IBM Storage Sentinel and IBM
Storage Safeguarded Copy**

October 2023

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

First Edition (October 2023)

This edition applies to:

IBM Storage Copy Data Manager Version 15 2.2.19.0
IBM Storage Sentinel Version 1.1.2
Security Scan Engine Version 8.1.0 - Build 1.4
InterSystems IRIS Version 2022.1.1.374.0

© Copyright International Business Machines Corporation 2023. All rights reserved.

Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Contents

Notices	vii
Trademarks	viii
Preface	ix
Authors	ix
Now you can become a published author, too!	xii
Comments welcome	xii
Stay connected to IBM Redbooks	xii
Chapter 1. Introduction	1
1.1 Overview of cyber resiliency	2
1.1.1 Cybersecurity versus cyber resiliency	2
1.2 Approaches to data resiliency	4
1.2.1 Considering the restoration of static and dynamic data	5
1.2.2 Time to Recover	6
1.2.3 Secondary workload cyber resiliency	8
1.3 IBM Storage Sentinel overview	9
1.3.1 Supported applications	10
1.3.2 Use cases for Storage Sentinel	13
1.3.3 IBM Storage Sentinel workflow	13
1.3.4 IBM Storage Sentinel components	15
Chapter 2. Configuring the IBM Safeguarded Copy feature	17
2.1 Safeguarded snapshot with internal scheduler	18
2.2 Configuring Storage Sentinel with IBM Storage Copy Data Management	21
2.2.1 Registering providers	21
2.2.2 Configuring SLA policies	24
2.2.3 Creating backup jobs	26
2.2.4 Restore and recovery jobs	27
2.2.5 Prescript and postscript	28
Chapter 3. Protecting Epic cache and IRIS databases with IBM Safeguarded Copy and IBM Storage Sentinel	29
3.1 Introduction	30
3.2 Supported configurations for IBM Storage Copy Data Management and IBM Storage Sentinel for Epic databases	30
3.3 IBM Storage Sentinel server platform choice	32
3.3.1 Supported storage configurations for virtual Epic database servers	32
3.3.2 Supported storage configurations for physical Epic database servers	36
3.4 Setting up a CDM and Storage Sentinel environment to scan Epic databases	37
3.5 Performing a restore of an Epic database backup	46
Chapter 4. Configuring IBM Storage Sentinel for SAP HANA	49
4.1 SAP HANA integration into IBM Storage Copy Data Management	50
4.2 SAP HANA and data persistence	50
4.2.1 SAP HANA volumes	50
4.3 SAP HANA workflows and IBM Storage Copy Data Management	51
4.3.1 SAP HANA data backup workflow	51
4.3.2 SAP HANA restore workflow	55

4.3.3	SAP HANA requirements	56
4.4	IBM Storage Copy Data Management setup	57
4.4.1	Required user roles	57
4.4.2	Service Level Agreement (SLA) policies	58
4.5	Running SAP HANA backup and restore operations	59
4.5.1	Running an SAP HANA backup job	60
4.5.2	SAP HANA restore job	61
4.6	Daily operations, best practices and maintenance	64
4.6.1	Adding capacity to the SAP HANA data area	64
4.6.2	Combining backups not recommended	65
4.6.3	Backup of the IBM Storage Copy Data Management catalog	65
Chapter 5.	Scanning engine and its technology	67
5.1	Storage Sentinel architecture	68
5.2	Technology of the IBM Storage Sentinel scanning engine	69
5.3	The advantage of anomaly scanning versus signature scanning	69
5.3.1	The scanning process	69
5.3.2	Scanning process for databases	70
5.3.3	Machine learning	70
5.3.4	Scanning encrypted data	70
5.4	How to recognize and handle alerts	71
5.4.1	After alert workflow	72
5.4.2	What to do when the scanning engine finds an issue	72
5.4.3	Dealing with false positives	73
5.5	Scanning Engine planning considerations	73
5.5.1	Sizing considerations	73
5.5.2	Scaling of scan workloads	73
5.5.3	Virtual versus physical servers	73
5.6	Administration	74
5.6.1	Monitoring the scanning engine	74
5.6.2	Backing up and restoring the scanning engine components	75
5.6.3	Adding new applications	75
5.6.4	Adding new scanning engines	76
Chapter 6.	IBM Cyber Vault setup: Putting it all together	77
6.1	Introduction to IBM Cyber Vault	78
6.1.1	The four steps to IBM Cyber Vault	79
6.2	IBM Cyber Vault planning considerations	80
6.2.1	Definition of the Minimum Viable Company (MVC)	80
6.2.2	Establishing immutable copies of critical data	81
6.2.3	Crash consistency or application consistency?	81
6.2.4	Proactive monitoring	82
6.2.5	RPO, RTO, and data validation	82
6.2.6	Recovery planning	83
6.2.7	Further considerations	83
Chapter 7.	Supported patterns	85
7.1	Safeguarded Copy on a single system	86
7.2	Safeguarded Copy in a Metro Mirror or Global Mirror relationship	86
7.3	Safeguarded Copy in an IBM HyperSwap environment	88
Related publications		89
IBM Redbooks		89
Stay connected to IBM Redbooks		89

Help from IBM	89
---------------------	----

Notices

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

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

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing, IBM Corporation, North Castle Drive, MD-NC119, Armonk, NY 10504-1785, US

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

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

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

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

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

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

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

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

COPYRIGHT LICENSE:

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


Trademarks

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

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

AIX®
DS8000®
FlashCopy®
HyperSwap®
IBM®

IBM FlashSystem®
IBM Spectrum®
PowerVM®
QRadar®
Redbooks®

Redbooks (logo) ®
Storwize®
Tivoli®
XIV®

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.

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

VMware, VMware vSphere, 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 Storage Sentinel is a cyber resiliency solution for SAP HANA, Oracle, and Epic healthcare systems, designed to help organizations enhance ransomware detection and incident recovery. IBM Storage Sentinel automates the creation of immutable backup copies of your data, then uses machine learning to detect signs of possible corruption and generate forensic reports that help you quickly diagnose and identify the source of the attack. Because IBM Storage Sentinel can intelligently isolate infected backups, your organization can identify the most recent verified and validated backup copies, greatly accelerating your time to recovery.

This IBM Redbooks publication explains how to implement a cyber resiliency solution for SAP HANA, Oracle, and Epic healthcare systems using IBM Storage Sentinel and IBM Storage Safeguarded Copy. The target audience of this document is cybersecurity and storage specialists.

Authors

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



Nezih Boyacioglu has 20 years of experience as an SAN Storage specialist and currently works for IBM Premier business partner Istanbul Pazarlama in Turkey. He has over 20 years in the IT arena. His IBM storage journey starts with Tivoli Storage Manager and tape systems and his main focus for last 10 years has been on IBM Storage Virtualize family (IBM SAN Volume Controller, Storwize®, and IBM FlashSystem®), and Storage Area Networks. He is an IBM Certified Specialist for Enterprise Storage Technical Support, Flash Technical Solutions, Virtualized Storage, and Storage Spectrum software.



Gerd Franke is a Senior IT Architect in the EMEA Storage team of IBM Client Engineering. He is known as an expert for Storage strategy and architecture, focusing on Cybersecurity and Resiliency, Hybrid Cloud Storage and Modern Data Protection. He often leads large and complex client projects and is a regular speaker at IBM conferences. Gerd has worked for IBM in various national and international roles for more than 30 years. He holds an engineer's degree in Electrical and Communications Technology.



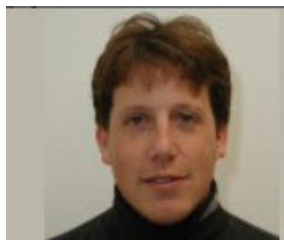
Thomas Gerisch is an IT Specialist working at Client Engineering Storage EMEA, focused on helping customers to build valuable solutions using their IBM Storage assets. When he joined IBM in 1999, he started as an instructor and specialist for open systems, before he became a storage expert at the IBM EMEA Storage Competency Center (ESCC) in Frankfurt, Germany. He has expertise in IBM storage development and testing. His current job role includes being the technical focal point for IBM customers running SAP HANA on IBM Storage. Thomas has authored several IBM White Papers about SAP HANA and IBM storage environment.



David Green works with the IBM SAN Central team troubleshooting performance and other problems on storage networks. He has authored, or contributed to, several IBM Redbooks publications. He is a regular speaker at IBM Technical University. You can find his blog at *Inside IBM Storage Networking* at <https://www.insidestoragenetworking.com/> where he writes about all things related to Storage Networking and IBM Storage Insights.



Vasfi Gucer works as the Storage Team Leader on the IBM Redbooks Team. He has more than 30 years of experience in the areas of systems management, networking hardware, and software. He writes extensively and teaches IBM classes worldwide about IBM products. His focus has been primarily on storage, cloud computing, and cloud storage technologies for the last 8 years. Vasfi is also an IBM Certified Senior IT Specialist, Project Management Professional (PMP), IT Infrastructure Library (ITIL) V2 Manager, and ITIL V3 Expert.



Guillaume Legmar has been involved with IBM Storage solutions for more than 20 years. Currently, he is a part of the Montpellier Garage to develop demonstrations about IBM FlashSystem and cyber resiliency. He is also a member of the IBM FlashSystem Beta and IBM Storage Virtualize Beta teams.



Markus Standau works for IBM Germany. He has more than 20 years of experience in the storage field in roles such as services, technical sales, and worldwide product management. He currently works in the storage sales acceleration team as the offering leader for Storage Virtualize, FlashSystem, and the Storage Control family in DACH. In his current role Markus organizes various business partner and customer events in DACH, such as IBM Storage Strategy Days, the SVC/FlashSystem user group and more. He holds a degree in Computer Science from Baden-Wuerttemberg Cooperative State University. Markus is the co-author of several IBM Redbooks on IBM Spectrum® Control and its predecessor, Tivoli® Storage Productivity Center (TPC).



Daniel Thompson has been working in IT for more than 40 years. His specialty is data protection (Backup and Restore, Disaster Recovery, Business Continuity and Cyber Resiliency). He currently works in the Advanced Technology Group (ATG), IBM Technology, Americas.



Christopher Vollmar is an IBM Certified Consulting IT Specialist (Level 3 Thought Leader) and Storage Architect who is based in Toronto, Ontario, Canada with the IBM Systems Group. Christopher is focused on helping customers build storage solutions by using the IBM Spectrum Storage Software-Defined Storage (SDS) family. He is also focused on helping customers develop private and hybrid storage cloud solutions by using the IBM Storage Virtualize family and Converged Infrastructure solutions. Christopher has worked for IBM for more than 20 years across many different areas of IBM, and has spent the past 12 years working with IBM System Storage. Christopher holds an honours degree in political science from York University.



Axel Westphal is an IT Specialist for IBM Storage Systems at the IBM European Storage Competence Center (ESCC) in Mainz, Germany. He joined IBM in 1996, working for Global Services as a System Engineer. His areas of expertise include setup and demonstration of IBM System Storage products and solutions in various environments. Since 2004, Alex has been responsible for storage solutions and Proof of Concepts conducted at the ESSC with IBM DS8000®, SAN Volume Controller, and IBM® XIV®. He has been a contributing author to several XIV and DS8000-related IBM Redbooks® publications.

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

James Munro, Martin Purkis
IBM UK

Brian Sherman, Pepe Lam, Dan Zehnpfennig, Matt Key, John Bernatz
IBM US

Michael Frankenberg
IBM Germany

Joseph Hand, Jim McGann
Index Engines

The IBM Advanced Technology Group (ATG) provided the environment and technical support used by the authors of this IBM Redbooks. Additionally one of the authors is a member of the ATG.

The ATG supports IBM Clients and Partners by providing demonstration environments for a variety of IBM solutions, hosting Accelerate with IBM web events that cover a variety of technical topics of interest to our clients and workshops that allow clients to work with technical specialists on a variety of topics. For more information, see [Advanced Technology Group](#).

Now you can become a published author, too!

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

Find out more about the residency program, browse the residency index, and apply online at: ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

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

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

ibm.com/redbooks

- ▶ Send your comments in an email to:

redbooks@us.ibm.com

- ▶ Mail your comments to:

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

Stay connected to IBM Redbooks

- ▶ Find 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>



Introduction

Topics in this chapter introduce the concepts of cyber resiliency and explain how cybersecurity is different from cyber resiliency. Other topics include an overview of IBM Storage Sentinel, a description of the currently supported applications, an introduction to some included use case, and a description of the IBM Storage Sentinel workflow.

This chapter has the following sections:

- ▶ “Overview of cyber resiliency” on page 2
- ▶ “Approaches to data resiliency” on page 4
- ▶ “IBM Storage Sentinel overview” on page 9

IBM Storage rebranding: In January 2023, IBM announced that the IBM Spectrum software-defined storage products are renamed to IBM Storage products. For example, IBM Spectrum Copy Data Management and IBM Spectrum Sentinel are renamed as IBM Storage Copy Data Management and IBM Storage Sentinel. At the time of writing, documentation that uses both Spectrum and Storage still exists, and documentation might refer to either or both naming conventions.

1.1 Overview of cyber resiliency

All businesses can be subjected to cyberattacks. These attacks often target applications that are critical to a business. Data or applications can be encrypted, stolen or both. Historically, disaster recovery and business continuity efforts focused on software and hardware failures. Businesses design redundancy into systems and storage, and use technologies such as backups and data replication to try to prevent the loss of data.

Many businesses are not prepared for, or are unaware of, the extent of the damage that a cyberattack can cause. They are also unaware of the costs of recovery from a cyberattack. Many of the businesses that do take steps to guard against cyberattacks focus efforts on prevention and not how to recover quickly from an incident.

Cyber resiliency is a measure of how well the applications and other infrastructure of a business can withstand events such as cyberattacks and natural disasters and still deliver business operations at a normal level. Cyber resiliency is critical for business continuity. It helps reduce financial losses and sometimes reduces damage to the business's reputation. *A cyber-resilient company has a competitive advantage because of efficient and effective operations and can maintain or even grow business during a crisis if its competitors cannot.*

1.1.1 Cybersecurity versus cyber resiliency

A Cybersecurity framework (CSF), such as that defined by the National Institute of Standards and Technology (NIST) is broader than recovery. Recovery in the event of a ransomware attack can be challenging, as key service data is increasingly fragmented across different data stores, both within and external to the organization. This complexity makes it even more difficult to identify the ransomware attack variant, identify the entry point of a corruption or encryption event, and then eradicate the risk of reinfection. Multiple data stores can also increase the complexity of recovering and testing data to ensure parity and synchronicity across data stores, which is required to ensure that the service is recovered in a safe and correct sequence.

Organizations that are able to focus on both cybersecurity and cyber resiliency improve their ability to recover from corruption or encryption-based events and ransomware style attacks. The two principals working in concert provide not only the capability of mitigating the attempts to disrupt the business by bad actors, but also provide for the ability to quickly recover the data that supports the organization. Cyber resiliency can be seen as the way for an organization to recover, test, restore environments and in the face of a ransomware, data corruption or encryption event.

Figure 1-1 shows the differences between cybersecurity and cyber resiliency.

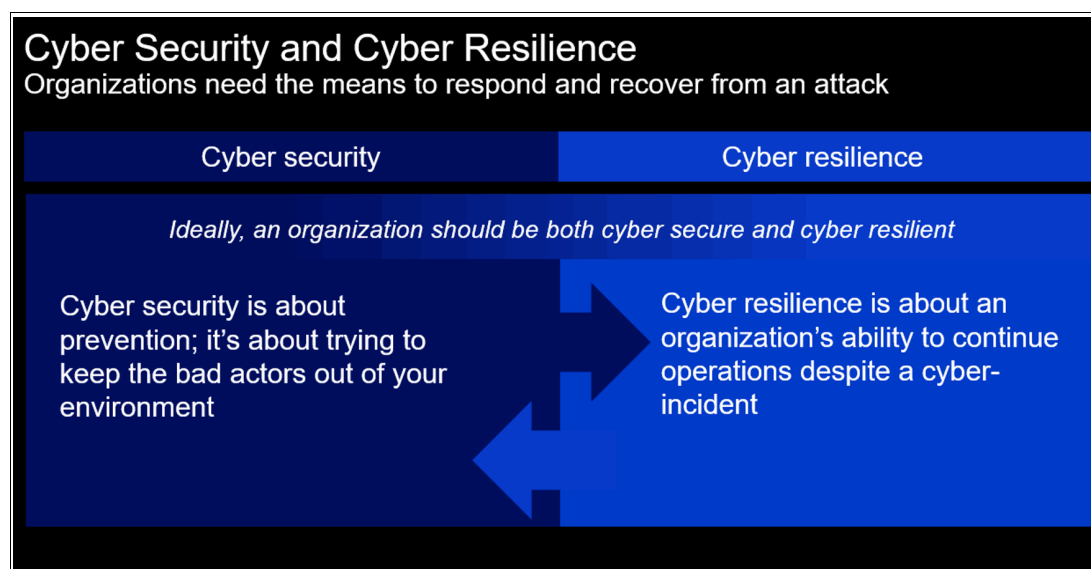


Figure 1-1 Cyber security and cyber resilience

Cybersecurity

Cybersecurity is the methods or practices that an organization uses to protect its systems and critical information from digital attacks. It is also known as Information Technology (IT) security. Cybersecurity measures are designed to combat threats against applications and networked applications. These threats can come from both inside and outside of organizations.

Cybersecurity includes IT intrusion detection and prevention, data loss or theft prevention, ransomware protection and protections for services that are running in the cloud. As organizations go mobile and employees work from mobile devices, mobile security is also a growing concern and must be included in a robust cybersecurity policy. Other concerns include minimizing potential damage from internal threat actors who already have privileged access to critical IT infrastructure.

Cybersecurity attacks often result in data theft, ransomware attacks or both. This data can include company secrets such as source code, or valuable customer and employee personal data.

Cyber resiliency

Where cybersecurity is focused solely on intrusion detection and data loss prevention, cyber resiliency is a measure of an organization's systems to survive a disaster or cyberattack and still allow the organization to function. Cyber resilience includes cybersecurity as a component. Cyber resilient organizations will be more secure than an organization that is not cyber resilient, but cyber resiliency is a measure of how well the organization functions during a cyber event and how quickly it can recover after an event (or a disaster event) occurs.

Cyber resiliency begins with a strategy or plan. This strategy identifies the critical assets that matter most to the organization and its stakeholders. Assets include the data or information that must be protected and is critical to the function of the organization, and the systems and services that matter most. The strategy must also include identifying the vulnerabilities and risks an organization faces.

The next part of cyber resiliency is the design. Design work chooses the controls, procedures, and training that are appropriate to prevent harm to critical assets. However, the design must be practical. An impractical design that cannot be implemented is not an effective one. The design work should also identify who has what authority to make decisions and act on them.

After the design is complete, the organization progresses to a test operational state. Where it is possible, resiliency is tested. Also, the customer closely monitors critical assets where it is not possible to test beforehand. The monitoring identifies when critical assets from the design phase are impacted by internal or external action. The design can be refined based on testing results.

After testing is complete, the organization moves to an operational state. In this phase, the design is deployed. Testing continues and uses controls to ensure that the operational state is effective and consistent

From the operational state, an organization with a mature cyber resilient design moves to evolution. Environments are constantly changing with new threats and new technologies. Organizations learn from incidents and how they recover from them. They need to modify procedures, training, and even strategy as they learn.

IBM Storage Sentinel can be part of an organization's cyber resilience strategy.

1.2 Approaches to data resiliency

The ability to recover to a prior point in time relies on the availability of backups - either point-in-time, array-based snapshots (such as copies made of Primary Workloads) or written to backup applications and their repositories such as disk, tape, or virtual tape library (VTL) or cloud (sometimes called Secondary Workloads). These recovery options require the availability of a system to use for recovery, and require the data being restored is not compromised in some way. The assumptions of system availability and uncorrupted data are often false when faced with recovering from a ransomware level event where the system or data is locked, corrupted or encrypted.

Both the Primary and Secondary copies that are available for use in recovery scenarios must be free of contamination to remove the risk of a repeated attack and reinfection. Also, where the backup is written to immutable storage, the backup must be validated as being free of infection before supporting recovery. Restoration from backups is traditionally limited to a single system, single files, or relatively small volumes of data. Restoration from backups is not designed to restore mass volumes of data to multiple systems in a short space of time. Recovery from tape-based systems (or Virtual Tape Libraries) is limited by seek time, retrieval times and network bandwidth, restricting the ability to recover at scale and speed.

Even when the Disaster Recovery systems, and restored data, are available, confidence in running business services by using the secondary site is typically low, either because it is not tested sufficiently, or because of differences between the primary and secondary configurations, or their integration to other interfaced components. Testing is often not representative of realistic failure scenarios and is typically based around site switching of single systems or single applications, which are quiesced and stopped gracefully first at the active site before being initialized at the recovery site. Often, the test at the secondary site does not include processing transactions, or if they are run, transactions run for a short period before switching back to the primary.

Clients can use data from traditional Disaster Recovery (DR) testing such as the priority and sequencing of applications, and data and infrastructure dependencies, in support of the recovery of business-critical systems and applications and elements of Service Management. More mature traditional DR testing might consider the following aspects:

- ▶ Recovery from loss of data
- ▶ Restarting systems
- ▶ Recovering and restarting applications
- ▶ Synchronous or asynchronous data mirroring
- ▶ Recovery point objective that is greater than zero (RPO > 0)
- ▶ Restarting business services to an earlier point-in-time
- ▶ Complexities associated with microservices
- ▶ Distributed systems and the synchronizing of data across system boundaries
- ▶ Interfaces with third parties

Existing business impact assessments (BIA) that include potential loss scenarios, prioritization, and service dependencies can be used to help determine appropriate recovery strategies for business services.

1.2.1 Considering the restoration of static and dynamic data

Some data is transactional and volatile whereas other data is relatively static. However, both types of data are important and support the organization in different ways. For example, in the energy and utilities sector, trading applications generate highly active transactional data but seismic records are static. However, both data types are regulated and might easily constitute the Minimum Viable Business of the organization.

Both of those data types can be protected but in different ways. Figure 1-2 on page 6 highlights that not all data needs as much protection because of factors, such as the frequency that the data is accessed. An organization can protect its data in various ways to support both its importance and its activity.

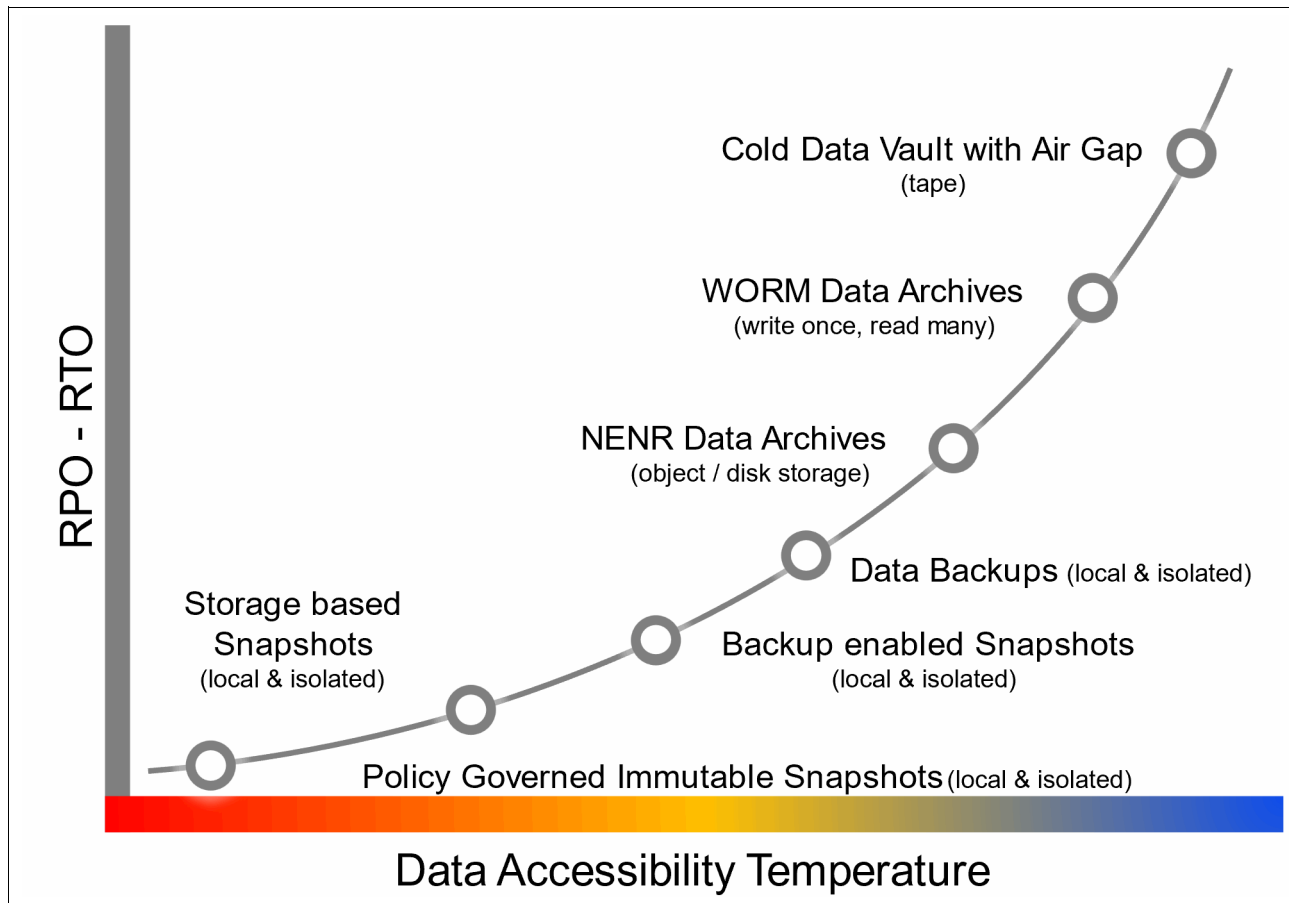


Figure 1-2 IBM's Point of View: Levels of data resilience

The most critical and highly transactional workloads generally need the lowest recovery time objective (RTO) and RPO and need the quickest accessibility to the cyber resilient copies by using something like policy-based immutable snapshots. Whereas, static data like seismic data or long-term records, that might have regulatory retention requirements, still need to be protected, but given their nature and size might need a different level of protection because they are not changing. This might be achieved by using something such as cold data storage, which might be using tape-based media.

1.2.2 Time to Recover

The ability of the organization to restore business services in a timely manner is key to success and in remaining within Impact Tolerance or in meeting the *Maximum Tolerable Period of Disruption (MTPD)*. This section identifies key considerations, which contribute to the elapsed duration of outage and recovery time.

The identification of the extent of the attack, including point of infection, nature of the attack, and extent of the damage is critical. All these factors affect the time that is needed to recover. Understanding these elements helps to determine the scope of recovery and what specific recovery actions are required. It helps to determine what services are impacted and what services can continue as they are.

The number and nature of infected systems directly affects the time needed to recover. Retrieval of data and restoration of data at a large scale can be limited by network bandwidth and the I/O capability of the source data repository, for example, the backup system. Whether recovery involves repairing a single file, data set or the complete restoration of a data volume, complete system, or server farm influences the time that is spent on recovery. The bandwidth available from the source repository, its media, and the network impacts how quickly data can be transported. Recovering large volumes of data from magnetic tape takes more time than recovering a FlashCopy of a disk subsystem volume.

Prevent reintroduction of contaminated data to the system after a ransomware attack. Before you restore data and restart services on the affected systems, verify that all traces of the ransomware is removed. Recovery of services on an alternate, isolated system can be done concurrently during recovery of the primary system.

As traditional HADR solutions are not content aware, additional checks, validation, and technical solutions are required to restore service. Ideally, these validations and checks are made before any need for recovery so that the integrity and usability of the stored data that is used for recovery is already known. However, if the state of the stored data is not known, then validation of data might be required during recovery, which can increase the recovery time.

Where data is restored to an earlier point-in-time ($RPO > 0$), the client might need to roll forward or replay transactions that occurred between the last known good backup of data and the time of the attack. The roll forward ensures that the system returns to a transactionally consistent state. Before you restart business services, verify the restored data is valid and free of malware. Reconciliation of data across multiple, interfaced systems and 3rd parties is also a key consideration and can add to the recovery time. Reconciliation of data might be required on applications that are hosted on the same platform (intra-system), on disparate systems (inter-system), and on 3rd-party systems and across the network in which the recovered system operates.

Technology solutions, such as IBM Cyber Vault for the protection of mainframe and open systems data, provide an ability to regularly backup data to immutable storage on the primary storage systems at production and DR sites. The solutions provide validation of data in an air-gapped system. Validated copies of data and an environment to restore, inspect, and enact recovery, significantly reduces the time to execute recovery and greatly improves the chance of a successful recovery than otherwise would be available. See Figure 1-3 on page 8.

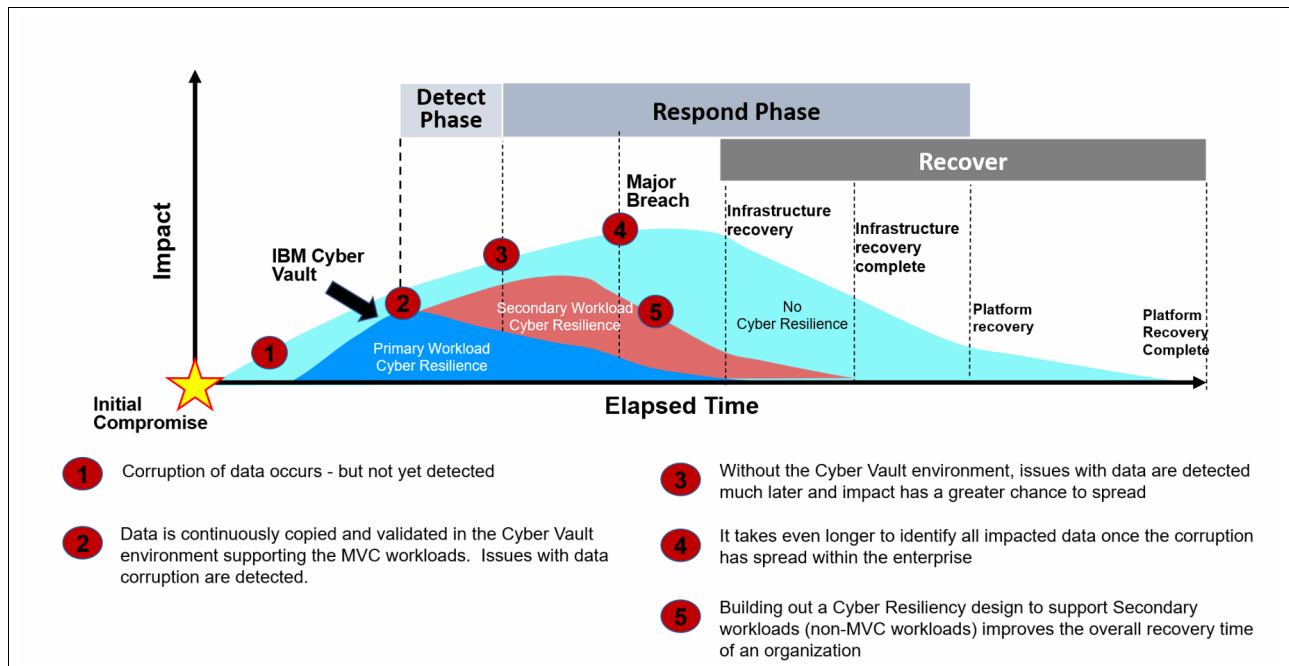


Figure 1-3 Recovery from ransomware attack timeline

1.2.3 Secondary workload cyber resiliency

For a more rapid recovery, an organization can prioritize its *minimum viable business* (MVB) assets from the typically large collections of systems and data in an enterprise. Depending on the spread of the corruption or encryption, the organization might need to address existing applications. Including a data resiliency strategy for existing applications can reduce recovery time for the remaining workloads. Part of a secondary workload strategy can include the following steps:

- ▶ Further prioritization of workloads for testing, recovery, and validation procedures
- ▶ Regular backup copies being moved to alternate immutable data platforms
- ▶ Moving copies from primary systems to secondary environments such as off the array
- ▶ Including a random sampling of workload for full recovery and validation

By separating MVB recovery from secondary workload recovery, both can be recovered using different strategies. See Figure 1-4 on page 9. The separate strategy provides for the ability to use different testing and validation methods. Both types of recovery can include some automation. Where MVB workloads might have targeted testing and application-specific tools, the secondary workloads can take a more mass scale approach using a more common set of generic tools. Proactively building an approach to support both primary and secondary recovery can mean accelerated recovery for both, which reduces overall organization disruption time.

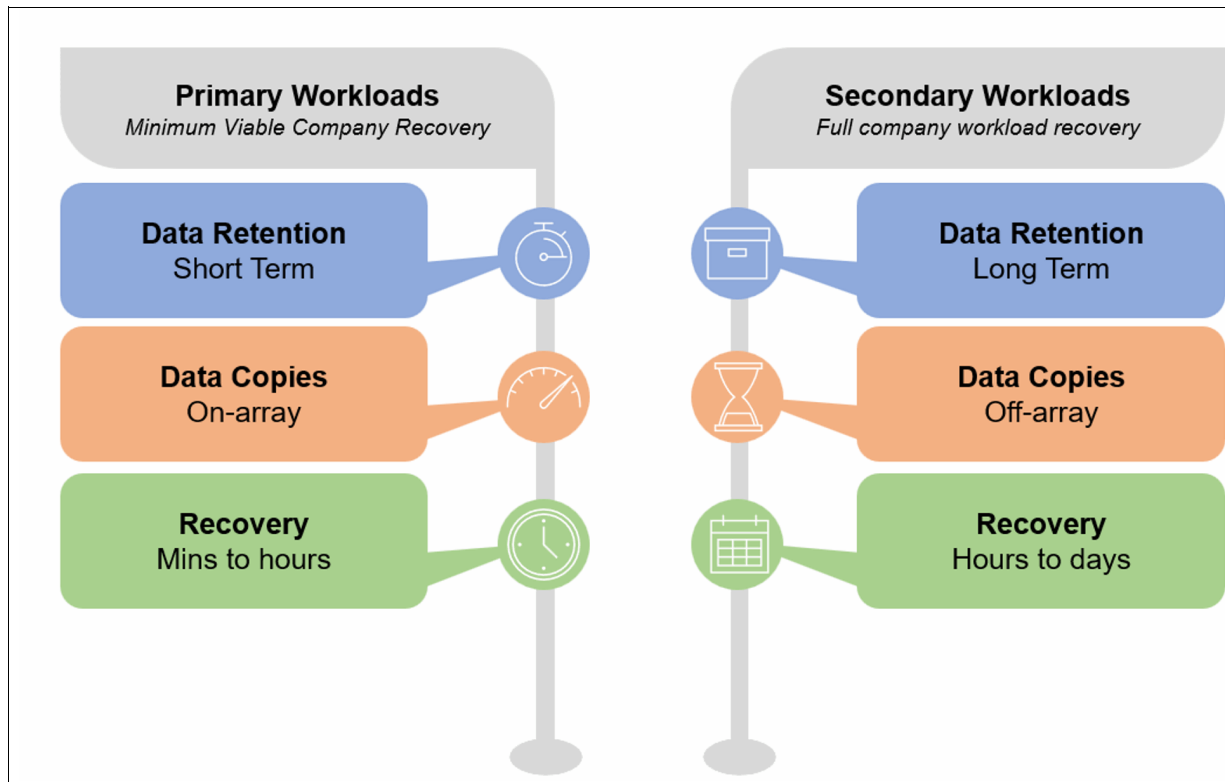


Figure 1-4 Segmentation of workloads

1.3 IBM Storage Sentinel overview

Organizations of all sizes in every industry are threatened by increasingly malevolent ransomware and other cyber threats. Despite the strongest defensive measures, an organization faces a risk that threats will get past every barrier and penetrate its information supply chain. In addition to the direct financial costs of lost business and recovery, these attacks can severely damage a company's brand, especially in industries where reputation is critical, such as financial services or healthcare.

Many organizations use some variant of a 30/60/90 policy for data backup and retention. Snapshots are captured hourly and/or daily with full backups generated every 30,60 and 90 days. Bad actors will exploit this by penetrating an organization's cyber defenses, install malware and then leave it dormant for long enough to fully infect the business's production and backup data. When the malware is triggered, it encrypts both production data plus all of the backup copies. In these situations the victim's only alternative is to pay the ransom.

You can find a full report on cyberattacks and their effects on companies here:

[IBM Security X-Force Threat Intelligence Index 2023](#)

IBM has also released a study on the impact of data breaches here;

[Cost of a Data Breach Report 2023](#)

Ransomware: Ransomware is an online attack that is perpetrated by cyber criminals or nation state-sponsored groups who demand a monetary ransom to release their hold on encrypted or stolen data.

A ransomware infection can be costly and disruptive if the only solution to return to normal business operations is to pay the cyber criminals' ransom. Statistics show, that only 50% of ransomware victims get back access to their data, even if the ransom was paid. One alarming trend is that cyber criminals now install malware and leave it dormant for 100 days or more before springing the trap. At that point, the malicious code has infected not only the target's production data systems and snapshots, but all of their backup copies, even if they use a *30 – 60 – 90 backup policy*. The victims have little choice but to pay.

Ransomware attacks can use several methods to infect a device or network. Some of the most prominent malware infection methods include:

- ▶ **Phishing emails and other social engineering attacks:** Phishing emails manipulate users into downloading and running a malicious attachment (which contains the ransomware that is disguised as a harmless looking .pdf, Microsoft Word document, or another file), or into visiting a malicious website that passes the ransomware through the user's web browser.
- ▶ **Operating system and software vulnerabilities:** Cyber criminals often exploit existing vulnerabilities to inject malicious code into a device or network.

IBM Storage Sentinel is a solution that is designed to help organizations detect ransomware and recover from cybersecurity incidents. It automatically creates immutable copies of data, then will use machine learning to detect possible corruption. It can generate forensic reports to help diagnose problems and find the source of an attack. Because it can isolate infected backups, you can identify which backup copies are verified and which are the most recent ones. This accelerates your time to recovery.

1.3.1 Supported applications

IBM Storage Sentinel support includes the several applications.

SAP HANA

SAP HANA is a database that stores data in system memory instead of on disk. This enables processing data at speeds that are magnitudes faster than disk-based systems. This allows for advanced, real-time analytics to be performed on the data. SAP HANA can be used on premises, in the cloud or in a hybrid cloud and deployed in both locations. SAP HANA can apply machine learning and AI to data from multiple areas of a business. For example, it can integrate data from several sources:

- ▶ Traditional documents - spreadsheets, contracts, and so forth
- ▶ Emails, website forms and other user experience documents
- ▶ Internet of Things (IoT) - such as data from sensors in warehouses or trucks, security sensors, RFID tags and the many types of sensors in all aspects of a business
- ▶ Mobile - data from the mobile devices of customers and employees

SAP HANA can integrate and analyze the vast amounts of data that sits in data warehouses and provide no value unless it is analyzed to provide more customer value and increase business impact.

SAP HANA can interact with other backup and restore functions.

Backup

The following types of backups are available for SAP HANA:

- ▶ IBM FlashCopy® NoCopy
- ▶ FlashCopy Incremental
- ▶ Global Mirror with Change volumes
- ▶ Safeguarded Copy

There is also an option to back up the log file.

Restore

Copy Data Services Manager creates a temporary volume from a backup, then mounts it to the original server for recovery.

Refer to Chapter 3, “Protecting Epic cache and IRIS databases with IBM Safeguarded Copy and IBM Storage Sentinel” on page 29 for more information.

Epic

EPIC is electronic healthcare record (EHR) software that covers all functions of healthcare operations. This includes patient records, patient engagement, billing, mobile, clinical data from medical tests, interoperability, specialist care, and even government regulations. EPIC uses two database technologies:

- ▶ An operational database that handles online transactions. This database runs Cache’ from Intersystems Corp.
- ▶ An analytical database that can run on either Microsoft SQL Server or Oracle.

Note: IBM Storage Sentinel currently does not support Microsoft SQL Server.

Refer to Chapter 4, “Configuring IBM Storage Sentinel for SAP HANA” on page 49 for more information.

Oracle

In June 2023 (with IBM Storage Sentinel Version V1.1.4), IBM announced support for Oracle DB running on both Linux and IBM AIX®.

Further information, see [What’s new in IBM Storage Sentinel anomaly scan software 1.1.4](#).

An example of an Oracle Backup job is shown in Figure 1-5 on page 12.

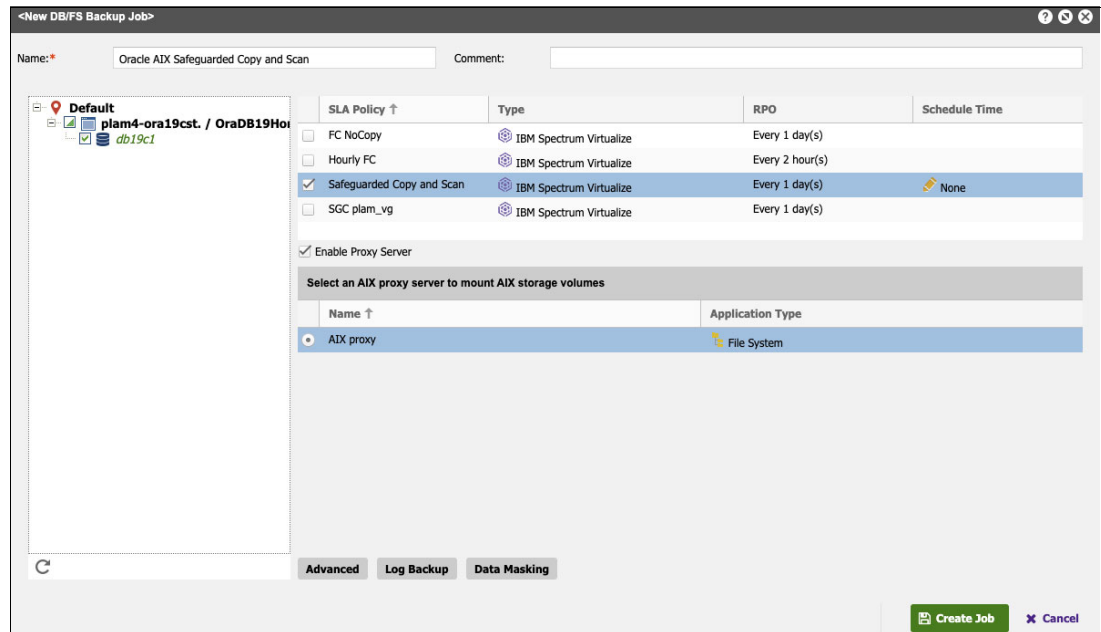


Figure 1-5 An example Oracle Backup job

The joblog shown in Figure 1-6 shows an anomaly detected by IBM Storage Sentinel.

Log			
Type	Time ↑	Task...	Message
			4ee56f638608
i	Apr 12 03:01:56 2023	2	Guest tools on 9.11.43.19 already at latest version: 2.15.4
i	Apr 12 03:01:57 2023	2	[9.11.43.19] Unix Agent 2.12.0.5 running as cdmsadmin for nfs mount (task ID: 678da049-85a0-4795-84c6-4ee56f638608)
i	Apr 12 03:01:57 2023	2	[9.11.43.19] Hostname: index-sle15-9 / Operating System: x86_64
i	Apr 12 03:01:57 2023	2	[9.11.43.19] Collecting list of NFS shares exported from the NFS Server: 9.11.60.88
i	Apr 12 03:01:57 2023	2	[9.11.43.19] Mounting NFS share: /tmp/mounts/10_11_59_121/1107/9_11_62_111/ec5b03cc629cccd6b800d89eec8f
i	Apr 12 03:01:57 2023	2	[9.11.43.19] NFS share mounted successfully
i	Apr 12 03:01:57 2023	2	[9.11.43.19] Completed mount operation in 0s, 1 NFS share(s) mounted successfully and 0 NFS share(s) failed
i	Apr 12 03:01:59 2023	2	Security Scanning of protected databases
i	Apr 12 03:01:59 2023	2	Starting Index job on mount path /tmp/mounts/10_11_59_121/1107 with job name 1107...
i	Apr 12 03:02:05 2023	2	Index job (121) created.
i	Apr 12 03:02:06 2023	2	Index job (121) started.
i	Apr 12 03:09:01 2023	2	Security Scan finished with state: Done. Previous threat detected: false. Number of new threats detected: 1.
i	Apr 12 03:09:01 2023	2	Unmounting database snapshot copies after Security Scanning
i	Apr 12 03:09:01 2023	2	ECX log dir=/data/log/ecxdeployer/2023-04-12/656b7731-c5e6-4608-81c8-7a4654bbbfed
i	Apr 12 03:09:04 2023	2	Guest tools on 9.11.43.19 already at latest version: 2.15.4
i	Apr 12 03:09:05 2023	2	[9.11.43.19] Unix Agent 2.12.0.5 running as cdmsadmin for cleanup (task ID: 656b7731-c5e6-4608-81c8-7a4654bbbfed)
i	Apr 12 03:09:05 2023	2	[9.11.43.19] Hostname: index-sle15-9 / Operating System: x86_64

Figure 1-6 An example of a Backup joblog containing a detected anomaly

Figure 1-7 shows an example of a Backup joblog with no detected anomaly.

ID	Type	Duration	Status	Message	Type	Time ↑	Task...	Message
1	Resolve	0h 0m 0s	CO...	COMPLETED	1	May 2 06:50:20 2023	2	Guest tools on 9.11.43.19 already at latest version: 2.15.4
2	Protection (Oracle)	0h 13m 33s	CO...	COMPLETED	1	May 2 06:50:21 2023	2	[9.11.43.19] Unix Agent 2.12.0.5 running as cdmdadmin for nfs mount (task ID: ebf77315-27c5-40ad-a639-f8489daa7949)
	Finding databases to protect:	Done (Total:1)			1	May 2 06:50:21 2023	2	[9.11.43.19] Hostname: index-sle15-9 / Operating System: x86_64
	Finding data and log disks of databases :	Done (Total:1)			1	May 2 06:50:21 2023	2	[9.11.43.19] Collecting list of NFS shares exported from the NFS Server: 9.11.60.88
	Resolving database disks on IBMSVC storage:	Done (Total:1)			1	May 2 06:50:21 2023	2	[9.11.43.19] Mounting NFS share: /tmp/mounts/10_11_59_121/1107/9_11_62_111/ec5b03cc629cccd6b800d89ec8fb
	Performing pre snapshot operations:	Done (Total:1)			1	May 2 06:50:21 2023	2	[9.11.43.19] NFS share mounted successfully
	Creating safeguard copies of volumes:	Done (Total:1)			1	May 2 06:50:21 2023	2	[9.11.43.19] Completed mount operation in 0s, 1 NFS share(s) mounted successfully and 0 NFS share(s) failed
	Performing post snapshot operations:	Done (Total:1)			1	May 2 06:50:21 2023	2	Security Scanning of protected databases
	Total databases protected:	1			1	May 2 06:50:23 2023	2	Starting Index job on mount path /tmp/mounts/10_11_59_121/1107 with job name 1107...
	Total databases not protected:	0			1	May 2 06:50:23 2023	2	Index job (131) created.
	Load storage data:	Done (Total:1)			1	May 2 06:50:23 2023	2	Index job (131) started.
	Load host data:	Done (Total:1)			1	May 2 06:50:28 2023	2	Security Scan finished with state: Done. No threats detected.
	Mount snapshot copies:	Done (Total:1)			1	May 2 06:50:29 2023	2	Unmounting database snapshot copies after Security Scanning
	Map LUNs:	Done (Total:1)			1	May 2 06:58:10 2023	2	ECX log dir=/data/log/ecxdeployer/2023-05-02/e1feb6ce-d556-49d2-9705-0ed216eef4d
	Security Scanning of protected databases:	Done			1	May 2 06:58:10 2023	2	Guest tools on 9.11.43.19 already at latest version: 2.15.4
	Dismount snapshot copies:	Done (Total:1)			1	May 2 06:58:13 2023	2	[9.11.43.19] Unix Agent 2.12.0.5 running as cdmdadmin for cleanup (task ID: e1feb6ce-d556-49d2-9705-0ed216eef4d)
	Cataloging objects:	Done (Total:17)			1	May 2 06:58:15 2023	2	[9.11.43.19] Hostname: index-sle15-9 / Operating System: x86_64
	Condensing catalog:	Done			1	May 2 06:58:15 2023	2	[9.11.43.19] Cleaning up mounted volumes

Figure 1-7 An example of a Backup joblog with no detected anomaly

1.3.2 Use cases for Storage Sentinel

IBM Storage Sentinel can be used to detect ransomware in snapshots, protect your data from corruption, and help recover after an attack.

Threat actors will often wait weeks or even months after ransomware is deployed to ensure that it infects all of a business's backups. IBM Storage Sentinel can detect ransomware in snapshots and other backups. It detects known patterns and uses machine learning to discover new patterns of infection to add to the known patterns.

Storage Sentinel can help protect data from corruption. It is designed to predict corruption and generate known good snapshots of data before they become corrupted or infected with malware.

Storage Sentinel helps recover after a ransomware attack. It can automatically generate reports listing the files or snapshots that were affected. This helps your organization identify clean copies of data that can be used to restore from.

1.3.3 IBM Storage Sentinel workflow

Figure 1-8 on page 14 shows where Storage Sentinel fits in an overall cyber resilience strategy. The IBM *Cyber Vault Blueprint* identifies the two main phases of cybersecurity. The first phase is to protect your data. The second phase is recovering from a cyberattack. Storage Sentinel spans both the protect and recover phases. It automates many of the tasks required for protecting data and can automatically identify safe recover points.

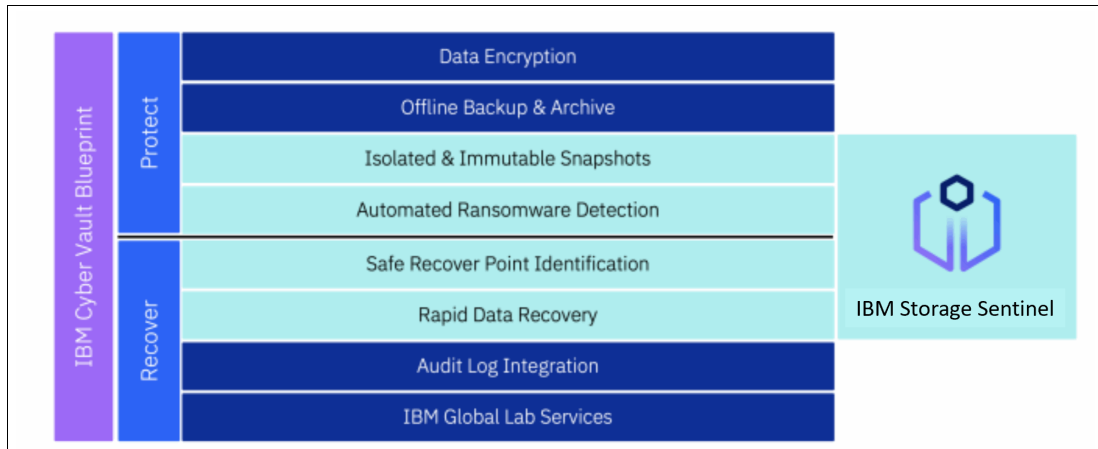


Figure 1-8 Storage Sentinel in Cyber Vault Blueprint

Figure 1-9 shows the key points on the Storage Sentinel timeline for responding to an attack:

1. Before an attack begins, IBM Safeguarded Copy creates a series of immutable snapshots, which are proactively scanned for malware by IBM Storage Sentinel.
2. Ransomware begins infecting production files, databases, and systems with Safeguarded Copy, and those snapshots are protected and scanned proactively by Storage Sentinel.
3. Even as the attack is taking place, Storage Sentinel scans snapshots and analyzes changes, file extension mismatch, and other signs of data corruption.
4. Snapshots can now be assessed to verify the snapshots that are free of malware and data corruption. This is the integrity review phase.
5. After the integrity review is complete, Storage Sentinel the most viable snapshot to restore from is identified.
6. The most recent, uncorrupted snapshot is used to restore data to the production environment so the organization can resume normal business operations.



Figure 1-9 Storage Sentinel attack timeline

1.3.4 IBM Storage Sentinel components

IBM Storage Sentinel has the following components.

IBM Safeguarded Copy

IBM Safeguarded Copy is a feature of the DS8000, IBM Storage FlashSystem, and IBM SVC storage systems that creates immutable snapshots of data to help protect against cyberattacks, malware, acts of disgruntled employees, and other data corruption. Safeguarded Copy uses the FlashCopy feature available on IBM Storage systems to create special FlashCopies that cannot be accessed by hosts. They cannot be mounted by or attached directly to a host. Instead, if recovery from a Safeguarded copy is required, another FlashCopy is created and that copy is presented to the host.

The IBM solution includes IBM Storage Sentinel. It complements IBM Safeguarded Copy by automatically scanning the copies that are created regularly by Safeguarded Copy and looking for signs of data corruption introduced by malware or ransomware.

IBM Copy Data Management (CDM)

IBM Copy Data Management is a tool available from IBM that can manage Safeguarded Copy snapshots. It also has additional capabilities to catalog the existing copy data environment, including storage, virtual machines and applications. After it is deployed, it can significantly improve the IT team's ability to deliver key functions and dramatically reduce the cost of infrastructure and ongoing operations.

Anomaly scanning and detection

Cyber protection solutions are designed to protect from an attack in real-time. However, these solutions are not 100% effective. Scanning data for anomalies adds additional protection to these solutions. It detects and locates corruption that occurs when a successful attack makes it into the data center. Early detection of issues enables IBM Copy Data Management software to start fast application recovery. This minimizes downtime and flattens the data resiliency curve.

The scanning software uses statistics about files on the host to identify corrupted files by using a machine learning model (MLM). The MLM is trained using real world malicious codes. The software identifies malicious code attacks and checks the integrity of databases to detect corruption of the internal database. This corruption might occur because of an attacker; data corruption due to logical or physical causes; damage at the disk or volume level; as a flaw in the process to create a snapshot; or as a flaw in the process to back up the database.

The scanning software examines existing database pages and allocation tables, if they exist, to ensure that all the allocated database pages are present and located in their correct position. Sometimes a data signature is available or enabled by the database administrator, such as a checksum or CRC. The scanning software recalculates the signature based on the current page contents and verifies it against the value found in the page header. Other ancillary fields are also verified within each page, depending upon the database application. The MLM of the scanning software is designed to tolerate a small amount of database corruption that is commonly observed in production database systems to avoid excessive false-positive alerts.

Figure 1-10 shows the five steps to cyber resilience.

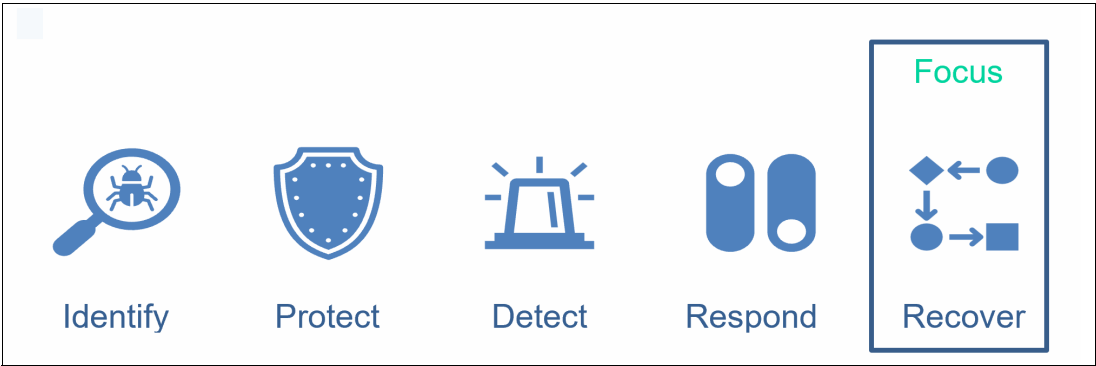


Figure 1-10 Five steps to cyber resilience



Configuring the IBM Safeguarded Copy feature

This chapter describes the IBM Safeguarded Copy feature, an integrated solution that is designed to create a logical air gap and protect critical data in the IBM FlashSystem product family. Safeguarded Copy creates immutable copies of the volumes and protect those copies to comply with a defined service level agreement (SLA) policy.

There are three methods to configure the IBM Safeguarded Copy process:

1. Safeguarded snapshot with internal scheduler
2. Orchestration with IBM Storage Copy Data Management (CDM)
3. Orchestration with IBM Copy Services Manager

You can use both item numbers 2 and 3 on the same storage system. Because CDM is licensed on a TB basis, you can use it for volumes that require application-aware copies and require anomaly scanning with IBM Storage Sentinel. The internal scheduler can be used for volumes where crash-consistent copies are sufficient.

This chapter has the following sections:

- ▶ “Safeguarded snapshot with internal scheduler” on page 18
- ▶ “Configuring Storage Sentinel with IBM Storage Copy Data Management” on page 21

Important: When the term *Safeguarded Copy* is mentioned in this book, it refers to IBM FlashSystem Safeguarded Copy only. *The DS8000 Safeguarded Copy function is not supported by IBM Storage Sentinel or IBM Storage Copy Data Management.*

2.1 Safeguarded snapshot with internal scheduler

Figure 2-1 shows the IBM Safeguarded Copy architecture.

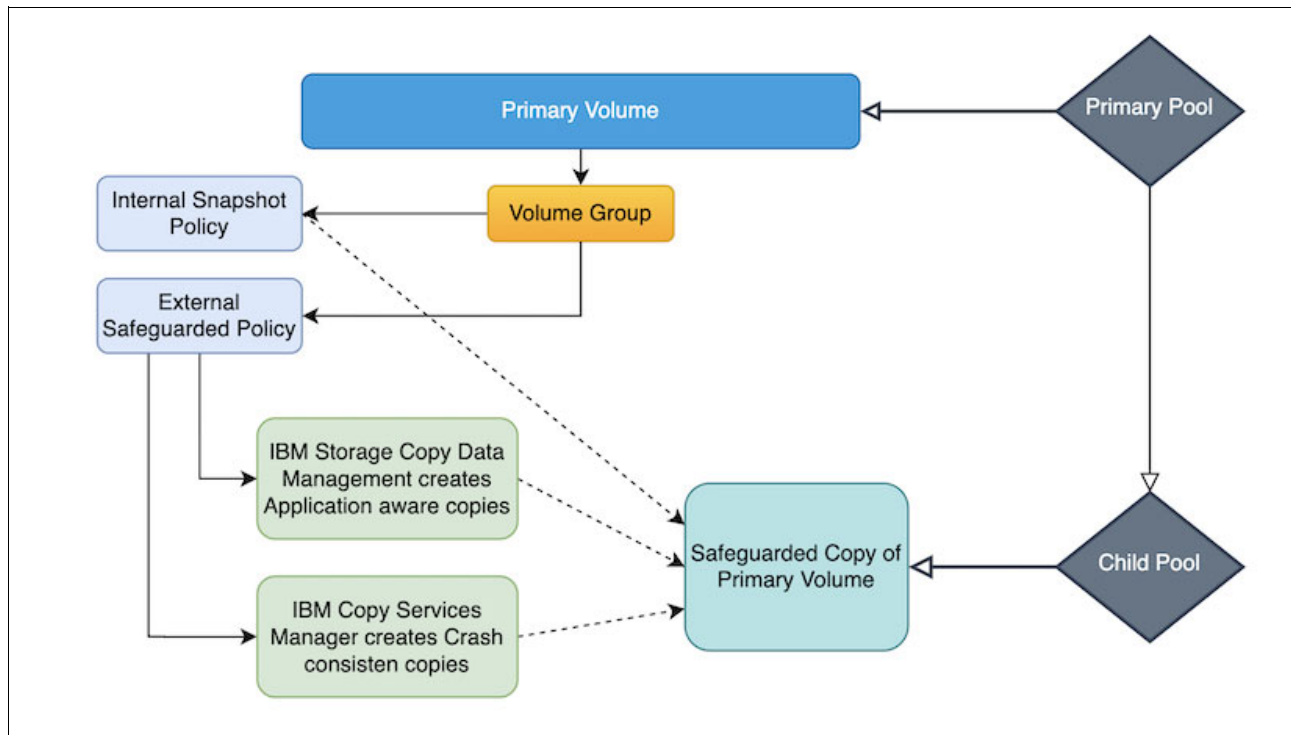


Figure 2-1 IBM Safeguarded Copy architecture

Safeguarded snapshot operates similar to snapshots and can use an internal scheduler. A new volume group can be created and pre-populated from another volume group's snapshot or Safeguarded snapshot. It inherits data, volumes, and volume groups from the snapshot.

Safeguarded copies that are managed by the internal scheduler, create crash-consistent copies of primary volumes. For the application aware copies of primary volumes, use IBM Storage Copy Data Management.

To create safeguarded copies by using the internal scheduler, consider the following requirements:

- ▶ Safeguarded child pool
A pool under the primary pool designated as Safeguarded.
- ▶ Volume
Volume must be in the primary pool that contains the Safeguarded child pool.
- ▶ Volume group
Assign the volumes to this group to create Safeguarded Copies.
- ▶ Safeguarded policy
Policy for how often copies are created and how many days they are retained.

Figure 2-2 shows the volume group policy selection page.

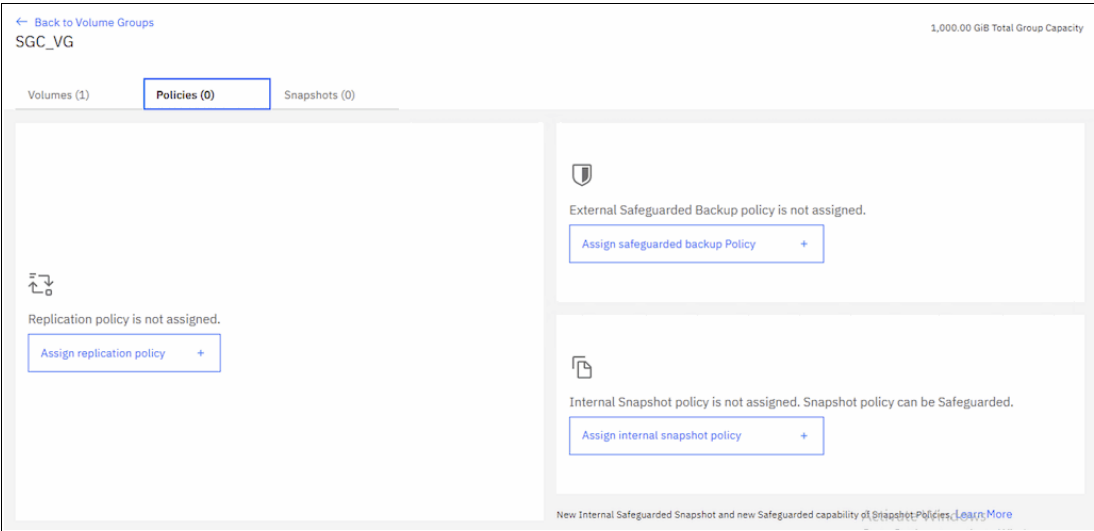


Figure 2-2 Volume group policy selection

Figure 2-3 shows how to assign internal snapshot policy to the volume group.

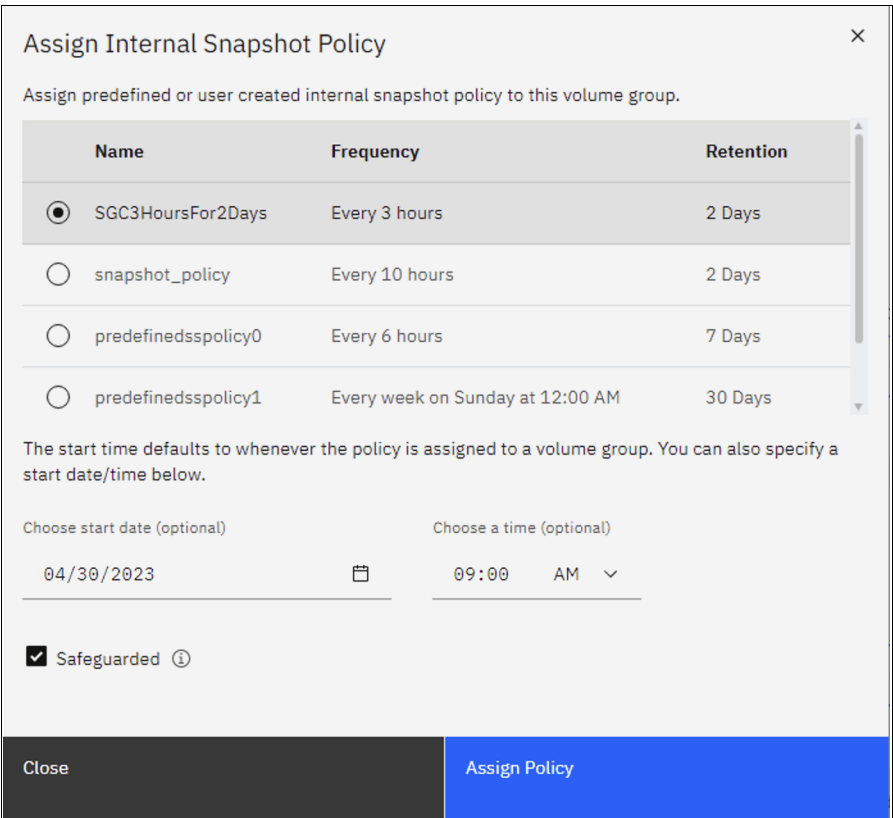


Figure 2-3 Assign internal snapshot policy to the volume group

After a Safeguarded snapshot policy is assigned to the volume group, then at the specified times, the internal scheduler automatically runs copy jobs and safeguarded copies that are listed on volume group details. See Figure 2-4 on page 20 and Figure 2-5 on page 20.

Safeguarded

Safeguarded Snapshot Policy

Volumes (1)

Policies (1)

Snapshots (9)

Capacity for Snapshots

Search

ID	Snapshot	State	Safeguarded
0	snapshot0	Active	Yes
1	snapshot1	Active	Yes
2	snapshot2	Active	Yes
3	snapshot3	Active	Yes
4	snapshot4	Active	Yes
5	snapshot5	Active	Yes

Figure 2-4 Safeguarded snapshots

Safeguarded

Safeguarded Snapshot Policy

Volumes (1)

Policies (1)

Snapshots (9)

Capacity for Snapshots

3.11 GiB / 8.79 TiB (0.03%)

Show details

Search

Take Snapshot

ID	Snapshot	State	Safeguarded	Time Created	Expiration Time	
0	snapshot0	Active	Yes	4/30/2023 4:17 PM	5/2/2023 4:17 PM	⋮
1	snapshot1	Active	Yes	4/30/2023 5:55 PM	5/2/2023 5:55 PM	Create Thin Clone
2	snapshot2	Active	Yes	4/30/2023 8:55 PM	5/2/2023 8:55 PM	Create Clone
3	snapshot3	Active	Yes	4/30/2023 11:56 PM	5/2/2023 11:56 PM	⋮
4	snapshot4	Active	Yes	5/1/2023 2:56 AM	5/3/2023 2:56 AM	⋮
5	snapshot5	Active	Yes	5/1/2023 5:55 AM	5/3/2023 5:55 AM	⋮
6	snapshot6	Active	Yes	5/1/2023 8:55 AM	5/3/2023 8:55 AM	⋮
7	snapshot7	Active	Yes	5/1/2023 11:55 AM	5/3/2023 11:55 AM	⋮
8	snapshot8	Active	Yes	5/1/2023 2:56 PM	5/3/2023 2:56 PM	⋮

Items per page: 25

1-9 of 9 items

11 of 1 page

Figure 2-5 Recovering from safeguarded snapshots

To recover safeguarded copies, it is possible to create either thin clones or fully allocated clones. After a clone is generated from a safeguarded copy, manually map the clone volume to the host to initiate the recovery process.

2.2 Configuring Storage Sentinel with IBM Storage Copy Data Management

The IBM Storage Copy Data Management workflow includes registering a provider, cataloging data, searching for objects, generating reports, and copying and using data.

2.2.1 Registering providers

Add providers to the Inventory by registering them. Providers include the following items:

- ▶ IBM Storage Virtualize based FlashSystem
- ▶ IBM SAN Volume Controller
- ▶ application servers
 - SAP HANA
 - Oracle
 - VMware ESX

Before registering providers, create a site to assign to your provider. A site is a user-defined grouping of providers that is based on location or department.

Registering IBM FlashSystem storage

Complete the following steps to register the IBM FlashSystem storage:

1. Click the **Configure** tab. In the Views window, select **Sites & Providers** and then select the **Providers** tab.
2. In the Provider browser window, select **IBM Spectrum Virtualize**.
3. Right-click **IBM Spectrum Virtualize**. Then, click **Register**. The Register dialog window opens.
4. Complete the fields in the dialog window. Select **New** to add credentials if they are not yet added through identities. In this case, a superuser account is used to register IBM FlashSystem storage. See Figure 2-6.

Name	Username	Type	
FS5200	superuser	System	

Figure 2-6 Registering IBM FlashSystem storage

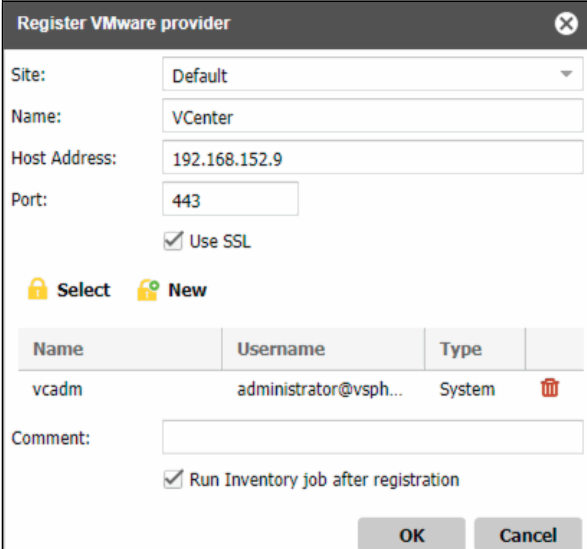
5. Click **OK**. IBM Storage Copy Data Management first confirms that a network connection exists and then adds the provider to the database.

Registering VMware vCenter

Complete the following steps to register the VMware vCenter:

1. Click the **Configure** tab. In the **Views** window, select **Sites & Providers** and then select the **Providers** tab.
2. In the Provider browser window, select **VMware**.
3. Right-click **VMware**. Then, click **Register**. The Register dialog window opens.
4. Complete the fields in the dialog window. Select **New** to add credentials if they are not yet added through identities. See Figure 2-7.

For the required VMware vSphere privileges, see [VMware vSphere Privileges](#).

The image shows a 'Register VMware provider' dialog box. It contains fields for Site (Default), Name (VCenter), Host Address (192.168.152.9), and Port (443). There is a checkbox for 'Use SSL' which is checked. Below these fields are two buttons: 'Select' (with a lock icon) and 'New' (with a plus icon). A table lists existing providers with columns for Name, Username, Type, and an action icon. One entry is shown: Name 'vcadm', Username 'administrator@vsph...', Type 'System'. Below the table is a 'Comment' field and a checkbox for 'Run Inventory job after registration' which is checked. At the bottom are 'OK' and 'Cancel' buttons.


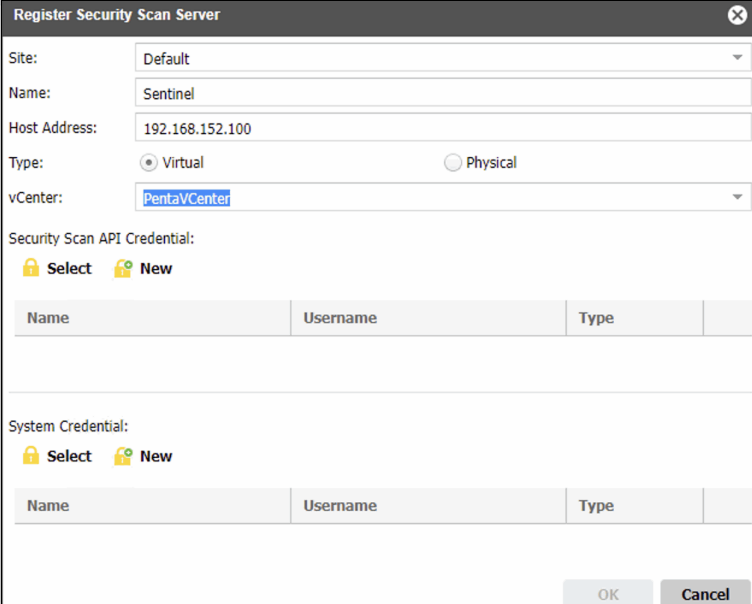
Name	Username	Type	
vcadm	administrator@vsph...	System	

Figure 2-7 Registering VMware vCenter

Registering Storage Sentinel Security Scan Server

Complete the following steps to register the Storage Sentinel security scan server:

1. Click the **Configure** tab. In the **Views** window, select **Sites & Providers** and then select the **Providers** tab.
2. In the Provider Browser window, select **Security Scan Server**.
3. Right-click **Security Scan Server**. Then, click **Register**. The Register dialog window opens.
4. Complete the fields in the dialog window. Select **New** to add credentials if they are not yet added through identities. If the security scan server is a virtual machine on VMware, select the pre-registered VMware VCenter and enter the credentials for the Sentinel server. See Figure 2-8 on page 23.



Register Security Scan Server

Site: Default

Name: Sentinel

Host Address: 192.168.152.100

Type: ☒ Virtual ☐ Physical

vCenter: PentaVCenter

Security Scan API Credential:

Select New

Name	Username	Type
------	----------	------

System Credential:

Select New

Name	Username	Type
------	----------	------

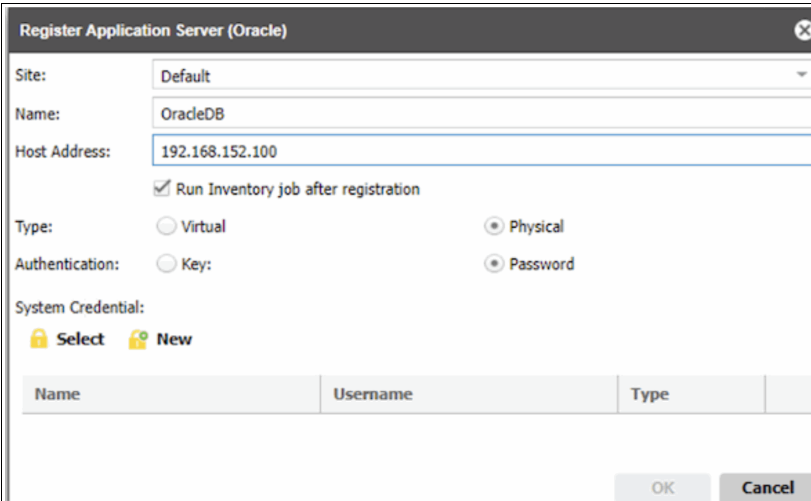
OK Cancel

Figure 2-8 Registering Sentinel Security Scan Server

Registering Oracle Database Server

Complete the following steps to register the Oracle database server:

1. Click the **Configure** tab. On the **Views** page, select **Sites & Providers**, then select the **Providers** tab.
2. In the Provider Browser window, select **Application Server**.
3. Right-click **Application Server**. Then click **Register**. The Register Application Server dialog opens.
4. Select **Oracle** as the Application Type and complete the fields in the dialog window. See Figure 2-9.



Register Application Server (Oracle)

Site: Default

Name: OracleDB

Host Address: 192.168.152.100

☒ Run Inventory job after registration

Type: ☐ Virtual ☒ Physical

Authentication: ☐ Key: ☒ Password

System Credential:

Select New

Name	Username	Type
------	----------	------

OK Cancel

Figure 2-9 Registering OracleDB Database Server

Registering SMTP Server

Complete the following steps to register the SMTP server:

1. Click the **Configure** tab. In the **Views** window, select **Sites & Providers** and then select the **Providers** tab.
2. In the Provider Browser window, select **SMTP**.
3. Right-click **SMTP**. Then, click **Register**. The Register window opens.
4. Complete the fields in the dialog window. See Figure 2-10.

Register SMTP provider

Name:

Host Address:

Port:

Select **New**

Name	Username	Type
------	----------	------

Comment:

▲ Email Options

From Address:

Subject Prefix:

Timeout (msec):

OK **Cancel**

Figure 2-10 Registering SMTP Server

2.2.2 Configuring SLA policies

The use of SLA policies allows for the customization of templates by administrators for the primary processes that are involved in the creation and use of Backup jobs. These policies configure copy types, destinations, and parameters that can be reused in future Backup jobs.

Note: For more information on SLA policies, see 4.4.2, “Service Level Agreement (SLA) policies” on page 58.

During the configuration of a Backup job, suitable SLA policies appear in the job creation wizard. The listed policies are tailored to the specific type of Backup job being created.

1. Click the **Configure** tab. On the **Views** window, select **SLA Policies**. The All SLA Policies page opens.
2. In the All SLA Policies page, click **New**. The New SLA Policies page opens.

3. Select a type of policy to create based on your storage provider. Select **IBM Spectrum Virtualize** to create an IBM Backup policy.

4. Add a sub-policy (SLA Policy) to an IBM Spectrum Virtualize SLA policy.

a) Select the source icon and define the recovery point objective to determine the minimum frequency and interval with which backups must be made. In the Frequency field, select Minutes, Hourly, Daily, Weekly, or Monthly, then set the interval in the **Interval** field. The smallest interval available is five minutes.

Note: If changes are made to the frequency and interval of an SLA Policy, those modifications impact all job schedules that are linked to it.

b) Click **Add Safeguarded Copy**.

c) In the Associated Safeguarded Volume Group page, expand the storage device and select the volume group that you want to back up by using Safeguarded Copy. Any volume that you want to back up by using Safeguarded Copy must belong to a volume group. If it is not a member of any of these groups, then it is not backed up as a Safeguarded Copy.

Note: The Associated Safeguarded Volume Group lists only those volume groups that have the Safeguarded Copy policy applied on the storage array side.

d) In the Options page, set the Safeguarded Copy sub-policy options.

Keep Snapshots

After a certain number of snapshot instances are created for a resource, older instances are purged from the storage controller. Enter the age of the snapshot instances to purge in the Days field.

Name

Enter an optional name to replace the default FlashCopy sub-policy name displayed in IBM Storage Copy Data Management. The default name is Safeguarded Copy0.

FlashCopy Volume Prefix

Enter an optional label to identify the FlashCopy. This label is added as a prefix to the FlashCopy name created by the job.

Tip: FlashCopy labels must contain only alphanumeric characters and underscores.

Perform Security Scan

Enable the scan and select your security scan servers, so you can scan for every backup number that you have specified.

e) Enter a name for the new sub-policy (SLA Policy).

f) Click **Finish**. See Figure 2-11 on page 26.

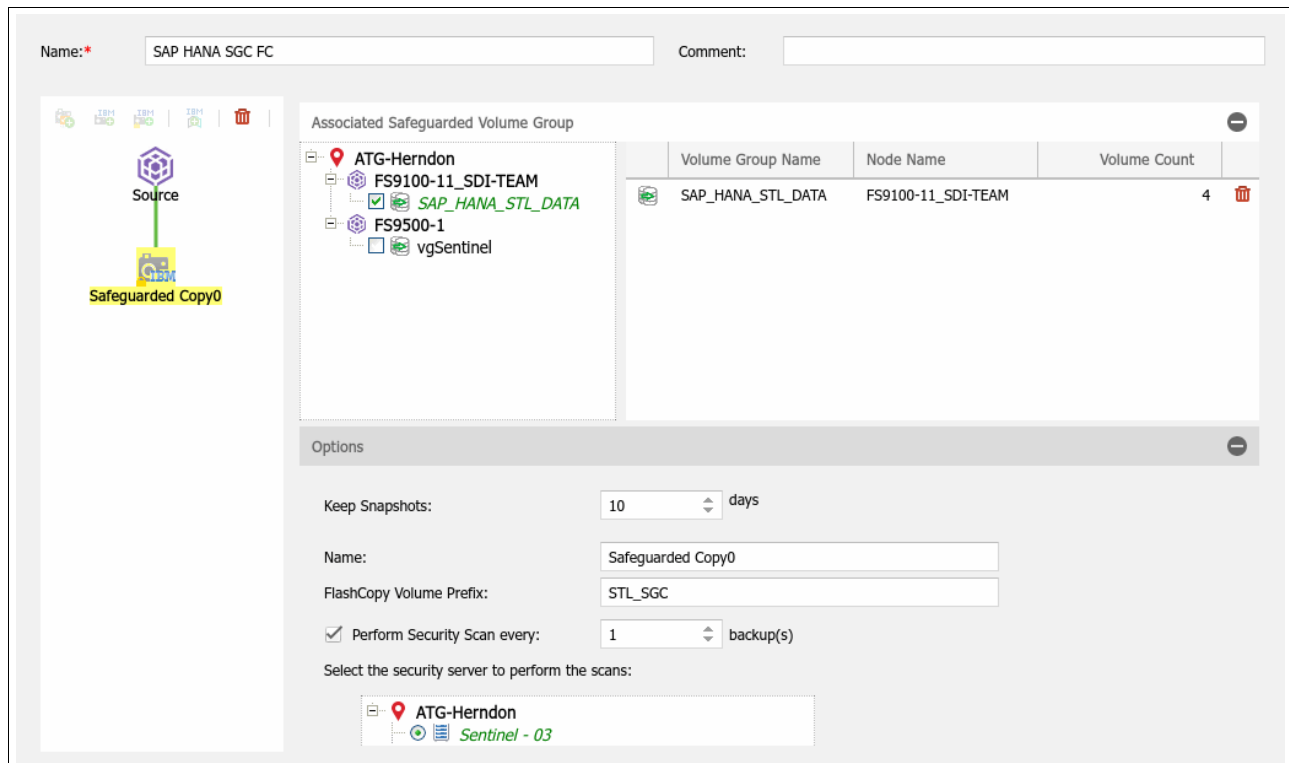


Figure 2-11 Configuring SLA policy and Storage Sentinel Scan frequency

2.2.3 Creating backup jobs

The backup jobs create a copy of your selected applications and dependent volumes, according to rules defined in an SLA policy. See Figure 2-12.

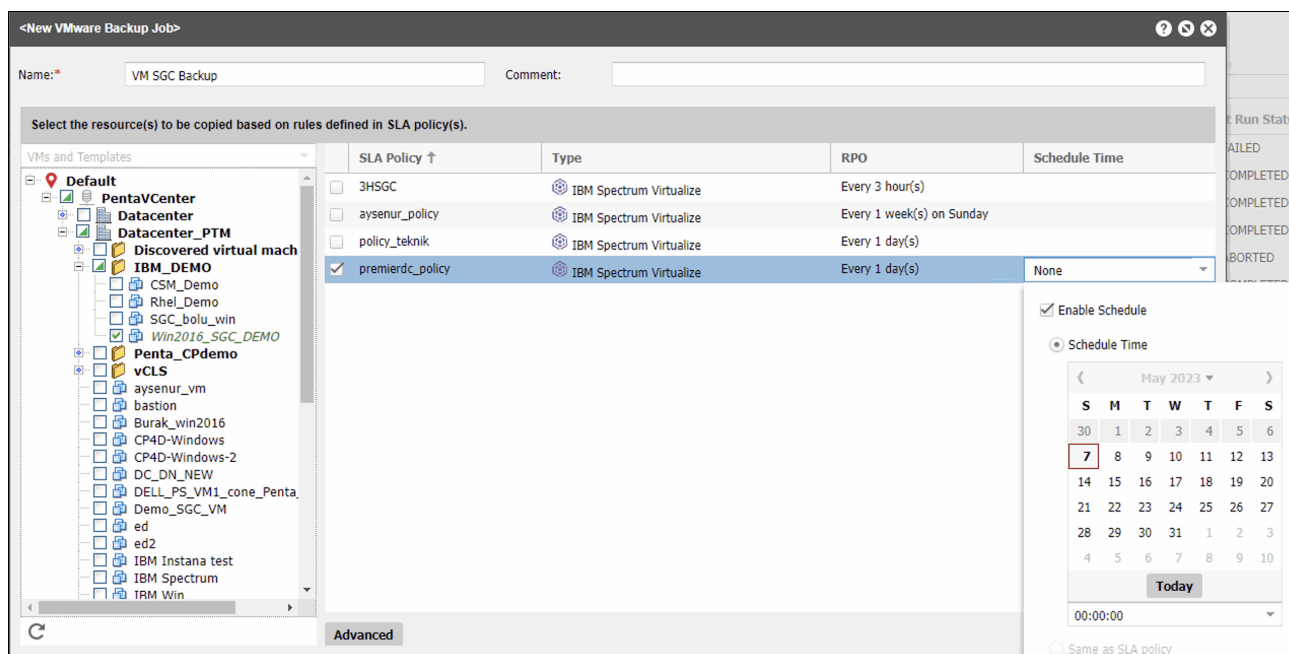


Figure 2-12 Creating a backup job with schedule

SAP HANA backup procedure is described in Chapter 4, “Configuring IBM Storage Sentinel for SAP HANA” on page 49.

2.2.4 Restore and recovery jobs

IBM Storage Copy Data Management uses Copy Data Management technology for testing and cloning use cases, instant recovery, and full disaster recovery.

There are multiple methods to regain access to your safeguarded copies for each application.

Note: Performing an SAP HANA Instant Database Restore job is described in “Performing an SAP HANA Instant Database Restore job” on page 62.

VMware restore options

For the VMware environment, you can choose **Instant Disk Restore**, **Instant VM Restore** or **Instant VM Restore (Long Distance)**. See Figure 2-13.

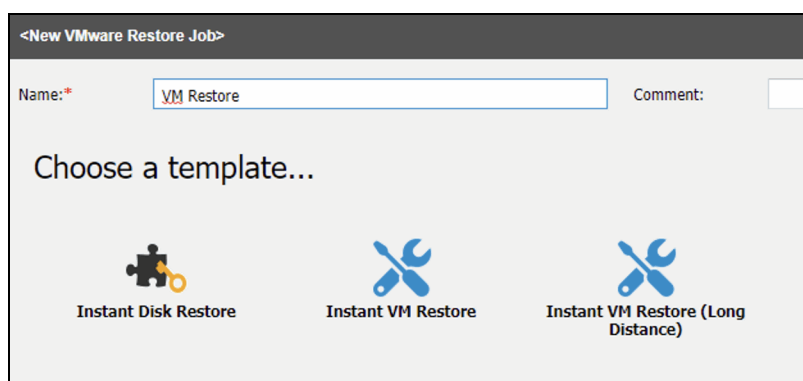


Figure 2-13 VMware restore options

Oracle database restore options

For Oracle databases, you can choose **DevOps**, **Instant Database Restore** and **Instant Disk Restore**. See Figure 2-14 on page 28.

For more information about Oracle databases with IBM Storage Copy Data Management, see [Oracle database support FAQ](#).

<New DB/FS Restore Job>

Name: Comment:

Choose a template...

Instant Disk Restore

Mount filesystems from storage copies containing application data.

Instant Database Restore

Recover database instances using non-masked storage copies and optionally roll forward to a specific point in time.

DevOps

Restore database instances using masked storage copies.

Figure 2-14 Oracle restore options

2.2.5 Prescript and postscript

Prescripts and postscripts are scripts that can be run before or after backup and restore jobs are run. Both at a job-level and before or after snapshots are captured. A script can consist of one or many commands, such as a shell script for Linux-based servers or Batch and PowerShell scripts for Windows-based servers.

Scripts can be created locally, uploaded to your environment through the Scripts page, then applied to job definitions. Volume Shadow Copy Service (VSS) was introduced in Windows 2003. If your application supports VSS and if you enable the option *Make these VMs application/file system consistent* when you create the VMware Backup job, then the Backup job triggers the VSS application quiesce logic. However, for applications that don't support VSS, or on Linux virtual machines, pre and post snapshot scripts can be used to quiesce your application for the snapshot backup.

Note: If adding a script to a Windows-based file system job definition, the user running the script must have the *Log on as a service* right enabled, which is required for running prescripts and postscripts.

Uploading a script

Perform the following steps for uploading a script.

1. Click the **Configure** tab. On the Views page, select **Scripts**.
2. Click **Upload**. The Upload Script dialog opens.
3. In the **Script** field, browse for a local script to upload, then click **Open**.
4. Enter an optional comment, then click **OK**. The script appears on the Scripts page and can be applied.



Protecting Epic cache and IRIS databases with IBM Safeguarded Copy and IBM Storage Sentinel

This chapter describes creating application-consistent, immutable snapshots of Intersystems Cache or IRIS databases and then having them scanned with IBM Storage Sentinel to detect possible malware corruption.

This chapter has the following sections:

- ▶ “Introduction” on page 30
- ▶ “Supported configurations for IBM Storage Copy Data Management and IBM Storage Sentinel for Epic databases” on page 30
- ▶ “IBM Storage Sentinel server platform choice” on page 32
- ▶ “Setting up a CDM and Storage Sentinel environment to scan Epic databases” on page 37
- ▶ “Performing a restore of an Epic database backup” on page 46

3.1 Introduction

This chapter describes creating application-consistent, immutable snapshots of InterSystems Cache or IRIS databases and then having them scanned with IBM Storage Sentinel to detect possible malware corruption.

Epic Systems databases often use InterSystems database technology to process health records. The Epic health records management solution is an industry leader in this space. Because cyber criminals often target healthcare organizations, the Epic databases were the first applications supported by IBM Storage Sentinel.

This chapter will describe the following:

- ▶ The currently supported configurations and how the backup, scanning and recovery flow will differ across those configurations.
- ▶ Registering storage, VMware, Storage Sentinel and Epic database components.
- ▶ Defining SLAs and running backups.
- ▶ Running a recovery of Epic databases.

Note: In this chapter, the details for each step of a job workflow are not included.

3.2 Supported configurations for IBM Storage Copy Data Management and IBM Storage Sentinel for Epic databases

Before implementation, verify your application, OS, platform and storage requirements against the most current product documentation for IBM Storage Copy Data Management and IBM Storage Sentinel. Requirements might have changed after this book was written. The current manuals and other documentation may be more up to date than this book due to the passage of time. As the time of writing, the listed supported configurations are for IBM Storage Copy Data Management v2.2.19 (CDM) and IBM Storage Sentinel 1.1.2.

Important: The product documentation is the official source of information. Always verify any design against the most current release of documentation before implementation.

The supported InterSystems database applications include the following releases:

- ▶ InterSystems Cache 2015, 2016, 2017, 2018 or later.
- ▶ InterSystems IRIS 2021, 2022 or later.

CDM and Storage Sentinel support the Epic databases being hosted in vSphere virtual machines or on physical hosts. Configuration and operation details are described in this chapter.

For virtual machines, vSphere 6.5 and 6.5.x levels, v6.7 and 6.7.x levels and, v7 and v7.0.x levels are supported.

InterSystems Databases running within virtual machines are supported on Red Hat Enterprise Linux (RHEL) 6.5 or later or CentOS 7.0 or later. InterSystems documentation states that CentOS is only supported for dev/test environments. See Figure 3-1 on page 31.

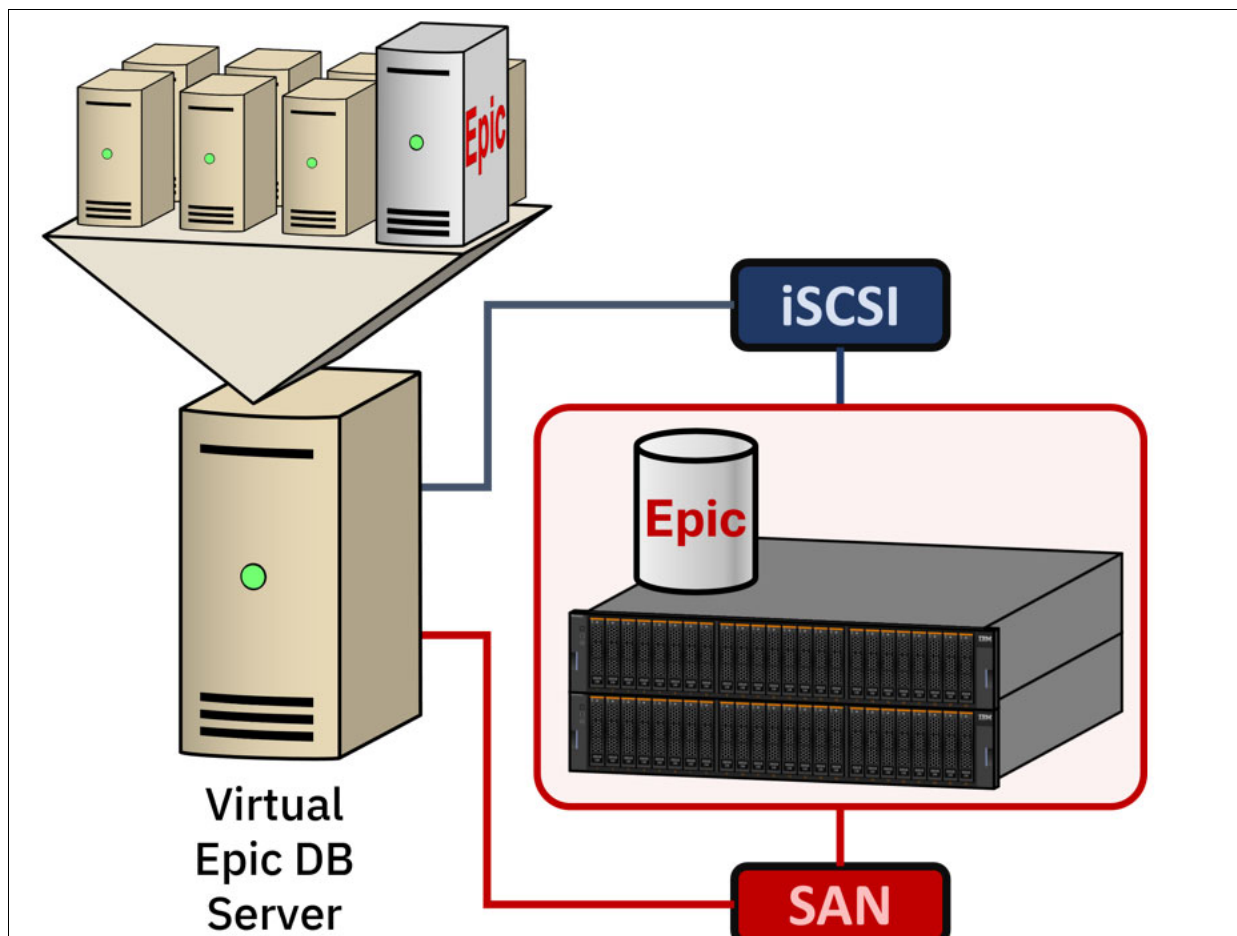


Figure 3-1 Virtual Epic DB Server

Physical machines on the Power platform (little-endian) are supported running AIX 6.1 TL9 or later, AIX 7.1 or later, AIX 7.2 or later and AIX 7.3 or later. *IRIS was tested only on the AIX 7.x operating system versions.*

Physical machines on the x64 platform are supported running RHEL 6.5 or later or CentOS 7 or later. It should be noted that Intersystems documentation states that CentOS is only supported for dev/test environments. For more information, see [InterSystems IRIS Data Platform 2023.2](#). See Figure 3-2 on page 32.

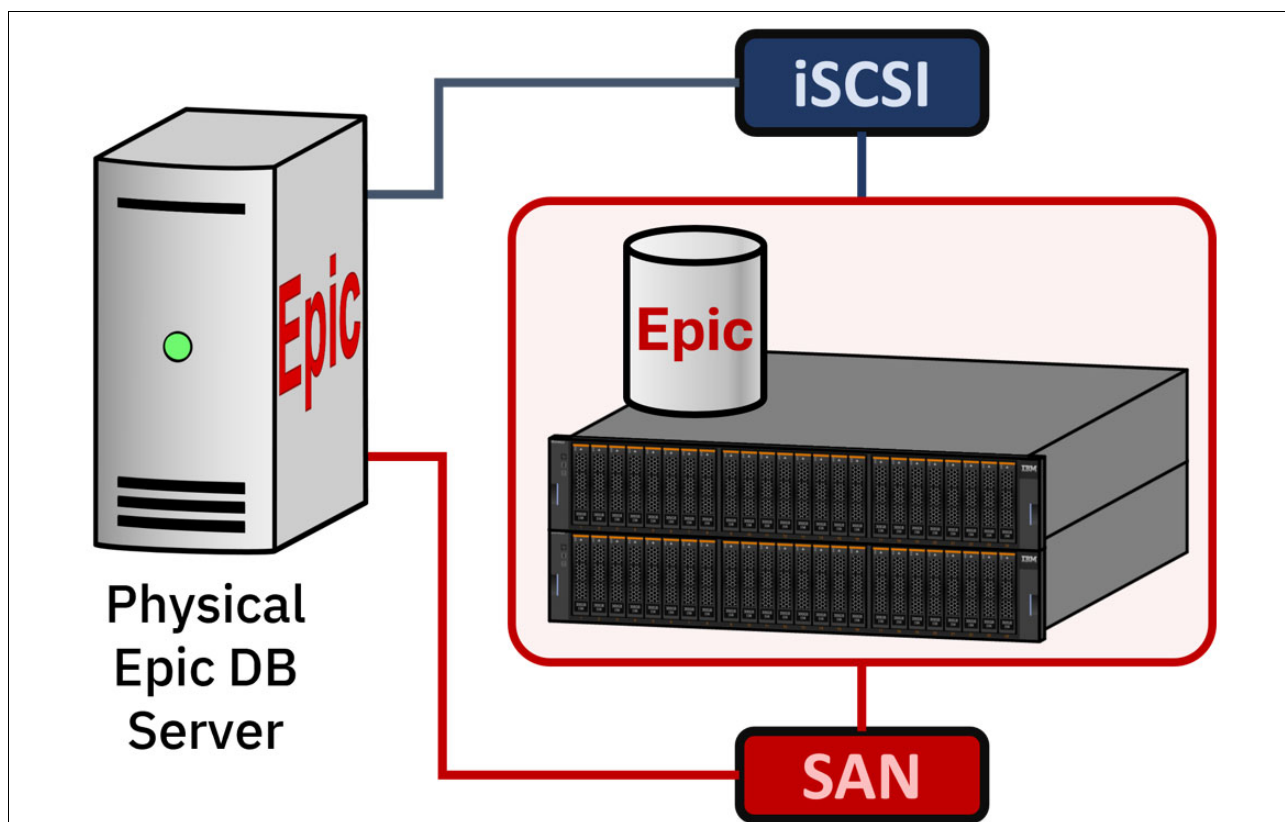


Figure 3-2 Physical Epic DB Server

3.3 IBM Storage Sentinel server platform choice

Storage Sentinel can be run either within a virtual machine or a physical machine. If Storage Sentinel is running on a physical machine, you can scan Safeguarded Copies taken of Epic databases also running on a physical machine only. If choosing to use a virtual machine to host Storage Sentinel, you can scan Safeguarded Copies of both physical and virtual machines.

3.3.1 Supported storage configurations for virtual Epic database servers

Although CDM supports both IBM and Pure storage for protecting Epic databases, this publication only covers IBM Storage Virtualize or IBM FlashSystem storage that provides Safeguarded Copy (SGC) functionality. In the current version, CDM supports SGC V1 but does not support SGC Volume Group Snapshots, also known as SGC V2.0).

There are 3 different supported configurations for taking Safeguarded Copies of Epic database volumes and then scanning them with Storage Sentinel. As mentioned in 3.2, “Supported configurations for IBM Storage Copy Data Management and IBM Storage Sentinel for Epic databases” on page 30, to scan the databases volumes of a virtual machine, Storage Sentinel needs to also be running as a virtual machine. If Storage Sentinel is running on a physical machine, you can scan Safeguarded Copies taken of Epic databases also running on a physical machine only.

Using virtual disks in a vSphere VMFS datastore

The first supported configuration is to configure the Epic VM to use virtual disks (VMDK) stored in a VMFS datastore. The storage must be a supported IBM Storage Virtualize or IBM Storage FlashSystem storage configuration. Volumes can be shared through a storage area network (SAN) or over iSCSI. When running in this configuration, the volumes containing the datastore are snapshotted in the storage controller. If you have other VMs sharing this datastore, there is no way to exclude them from the snapshots, as all volumes containing the VMFS datastore must be protected at the same time. It is recommended to limit how many other VMs share this datastore to reduce waste. Before the Storage Sentinel VM can scan the Epic DB volumes, CDM mounts a copy of the datastore snapshot and maps the VMDK files to the Storage Sentinel server. See Figure 3-3.

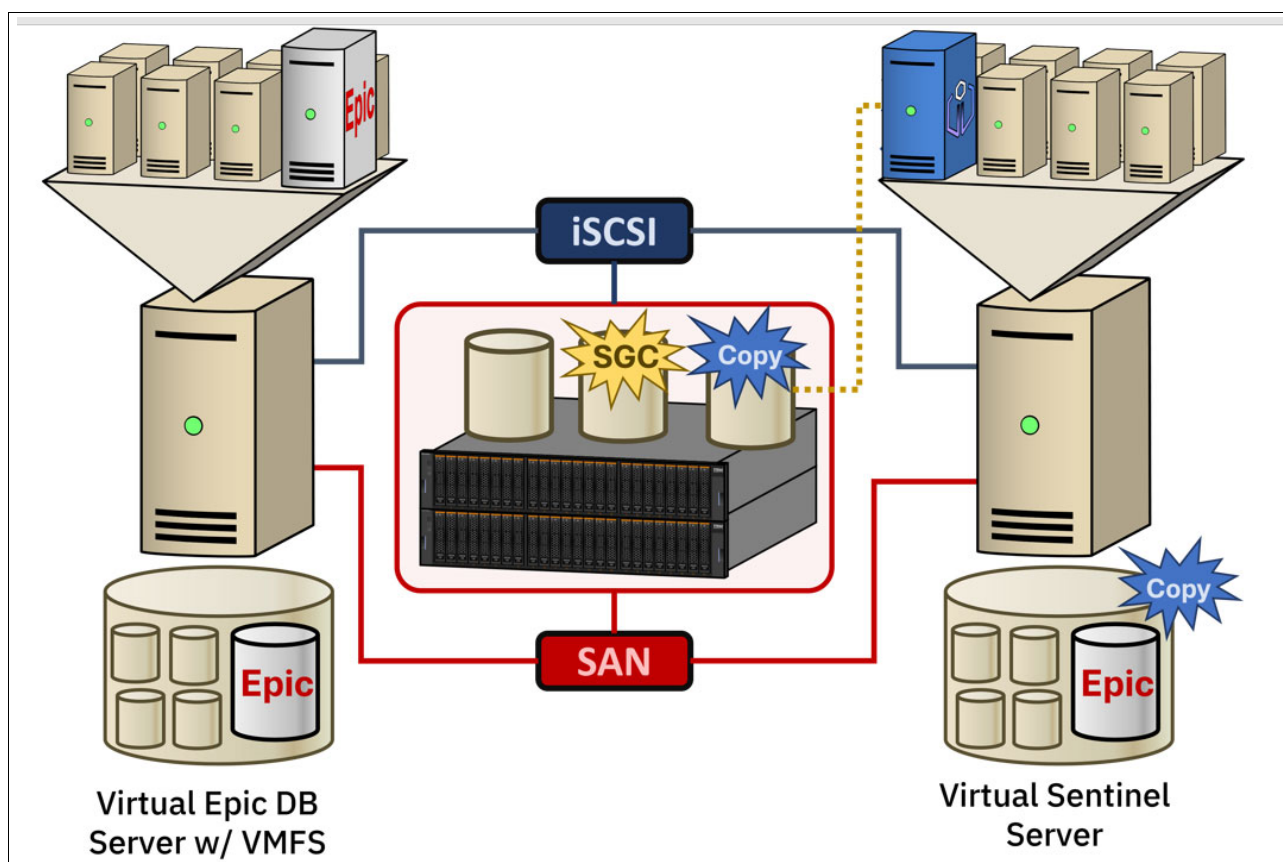


Figure 3-3 Virtual Epic DB server with virtual disks in a datastore

A backup job workflow will perform the following high-level tasks:

1. Quiesce the Epic database.
2. Take a Safeguarded Copy of the volumes containing the vSphere datastore.
3. Unquiesce the Epic database.
4. Create a copy of the Safeguarded copies.
5. Create a datastore on vSphere using these copies.
6. Modify the Storage Sentinel VM configuration and mount the virtual disks for scanning.
7. Scan the copy of the Epic DB.

7. Reverse the process to dismount the volumes, change the Storage Sentinel VM configuration and destroy the copies of the Safeguarded Copy volumes.
8. Flag the recovery point created by this backup job as having passed or failed the scan and raise an alert if it did fail.

Note: IBM does not currently support mixing VMDK disks and physical raw device mapped volumes with CDM or Storage Sentinel.

Using volumes shared over iSCSI directly to the virtual machine

This configuration refers to creating volumes and sharing volumes over iSCSI directly to the virtual machine, and does not refer to creating a VMFS datastore on iSCSI disk. In this configuration, the iSCSI volumes are not managed by the vSphere virtualization layer although the I/O is going over the vSphere virtual networks. When you register the Epic DB server to CDM, you register it as a physical machine, not as a virtual machine. This is the one configuration where a physical Storage Sentinel server can scan the Epic DB hosted in a VM.

A backup job workflow will perform the following high-level tasks:

1. Quiesce the Epic database.
2. Take a Safeguarded Copy of the iSCSI volumes.
3. Unquiesce the Epic DB.
4. Create a copy of the Safeguarded copies.
5. Mount the iSCSI volumes to the Storage Sentinel VM.
6. Scan the copy of the Epic DB.
7. Reverse the process to dismount the volumes and destroy the copies of the Safeguarded Copy volumes.
8. Flag the recovery point created by this backup job as having passed or failed the scan and raise an alert if it did fail.

Note: IBM does not currently support mixing iSCSI volumes with other configurations in a single machine.

3.3.2 Supported storage configurations for physical Epic database servers

CDM and Storage Sentinel support physical Epic database servers with volumes mapped from the Storage Virtualize and FlashSystem controllers over iSCSI or a SAN. These physical Epic DB servers can be scanned with either a physical Storage Sentinel server (Figure 3-5) or a virtual Storage Sentinel server (Figure 3-6).

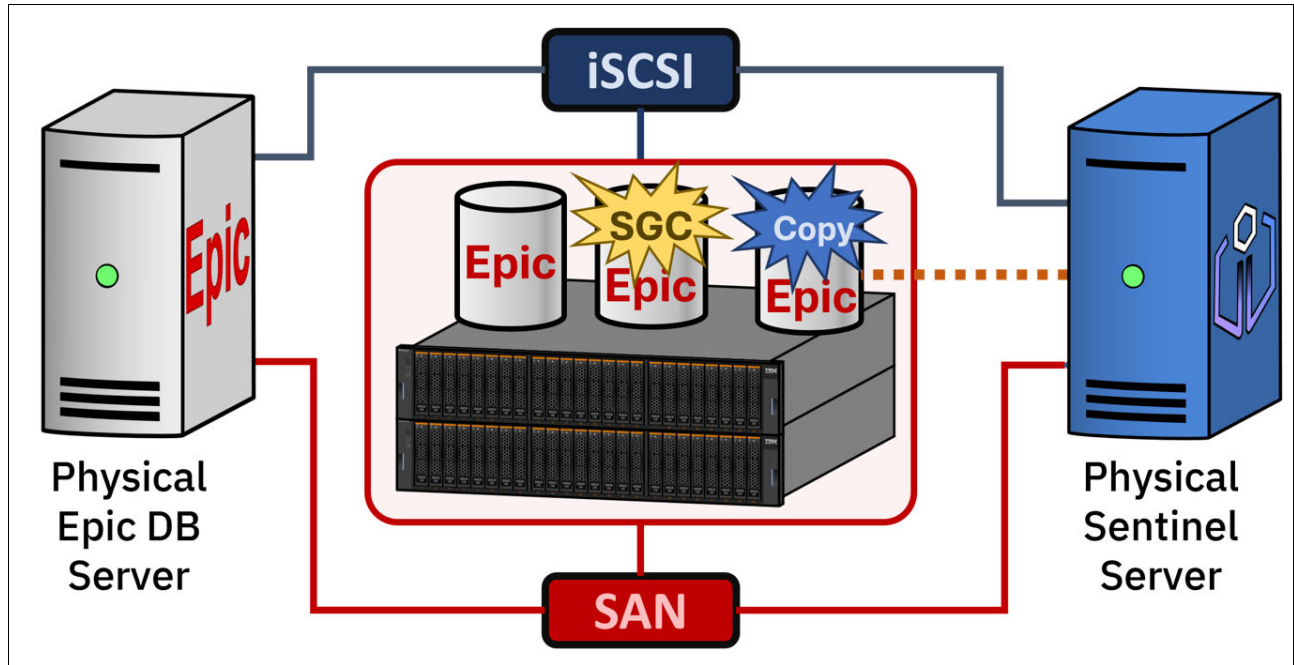


Figure 3-5 Physical Epic DB server and Physical Storage Sentinel Server

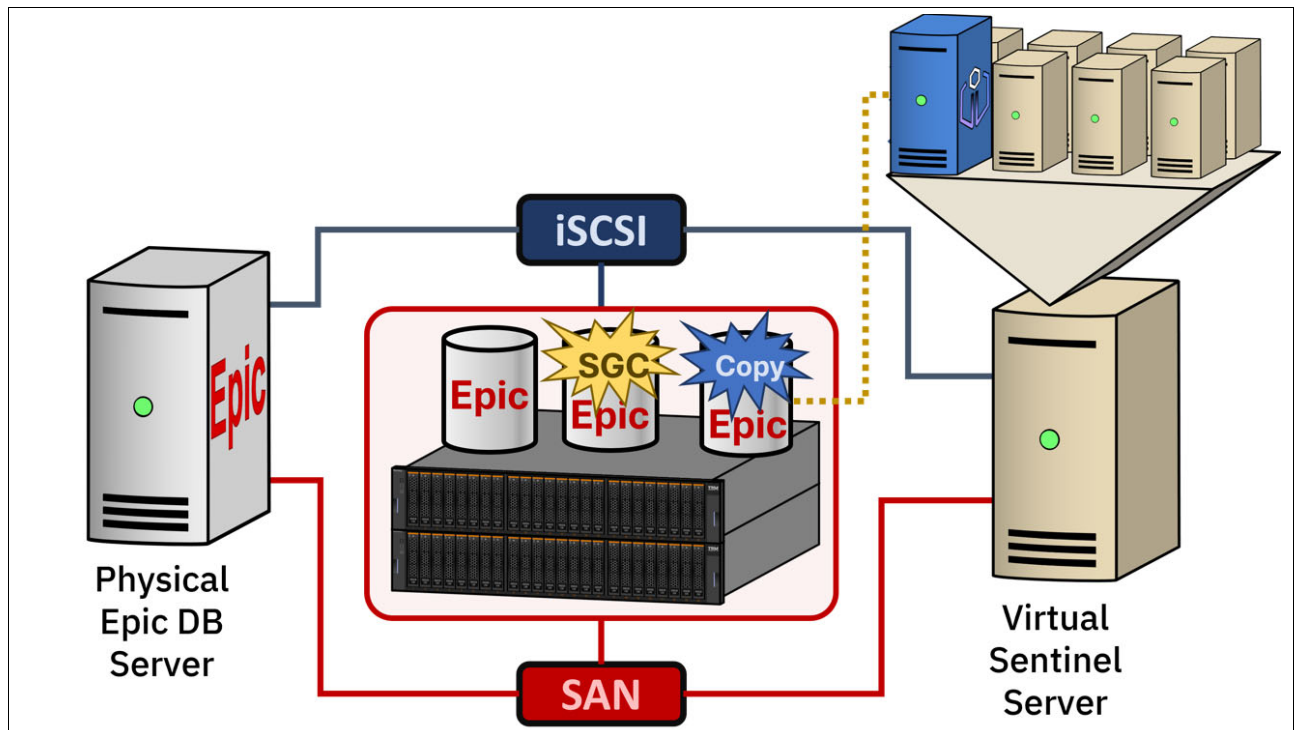


Figure 3-6 Physical Epic DB server and virtual Storage Sentinel server

A backup job workflow will perform the following high-level tasks:

1. Quiesce the Epic database.
2. Take a Safeguarded Copy of the SAN or iSCSI volumes.
3. Unquiesce the Epic DB.
4. Create a copy of the Safeguarded copies.
5. Share the volumes to the Storage Sentinel machine. If the volumes are served over iSCSI, the process will be identical when using a physical or virtual Storage Sentinel machine. For volumes shared over the SAN, the process will differ slightly between physical or virtual Storage Sentinel machines. If a physical Storage Sentinel host is used, the copy volumes will simply be mapped to that host. If Storage Sentinel is hosted in a VM, the copy volumes will be mapped to the appropriate ESXi host and then defined as pRDM volumes to the Storage Sentinel VM.
6. Scan the copy of the Epic DB.
7. Reverse the process to dismount the volumes and destroy the copies of the Safeguarded Copy volumes.
8. Flag the recovery point created by this backup job as having passed or failed the scan and raise an alert if it did fail.

Note: There are configuration steps omitted from this chapter for brevity. For example, before you can map volumes to either a physical machine or an ESXi host through a SAN, the SAN must be configured with host definitions and zones.

3.4 Setting up a CDM and Storage Sentinel environment to scan Epic databases

This section will describe the high-level steps needed to design and deploy a CDM and Storage Sentinel environment to protect and scan your Epic databases. Many of the tasks are described in more detail in other chapters in this book.

1. Plan and implement your supported server and storage deployment, as outlined earlier in this chapter. This will need to include predefining Safeguarded Copy volume groups.
2. Plan and implement the security settings and user accounts needed for creating the Safeguarded Copies, for integrating with vSphere, and for logging into CDM and Storage Sentinel, and any other necessary accounts. Decide to either use local accounts or use LDAP or Active Directory (AD). Decide the scope of authority for each account.
3. Plan and implement your Storage Sentinel configuration at the scale you need to be able to scan your Epic databases. Plan for other workloads that you plan to support with the Storage Sentinel configuration. Plan how to distribute your scanning workloads across the configuration, and plan how often to scan your application servers.
4. Deploy a new CDM virtual appliance, if needed. If using LDAP or AD for authentication, configure the security directory. To configure CDM to use LDAP or AD accounts, first register the LDAP server on the Provider Browser in the Configure Page (Figure 3-7 on page 38) and then import the LDAP group by going to the Access Control panel in the Configure Page, add a New User and Import the LDAP Group (Figure 3-8 on page 38).

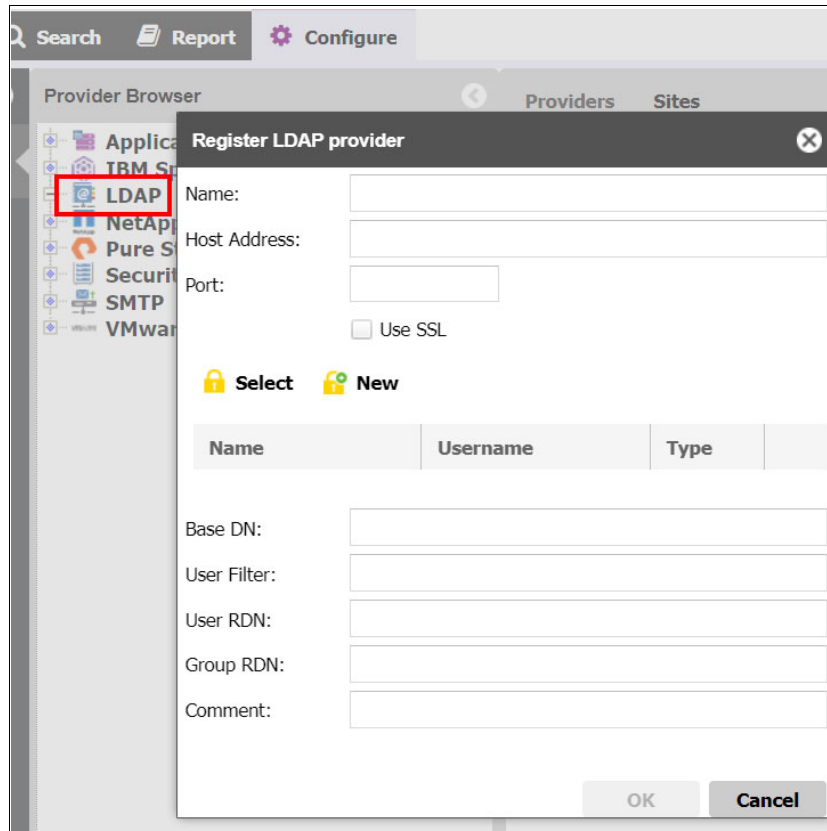


Figure 3-7 Register the LDAP server

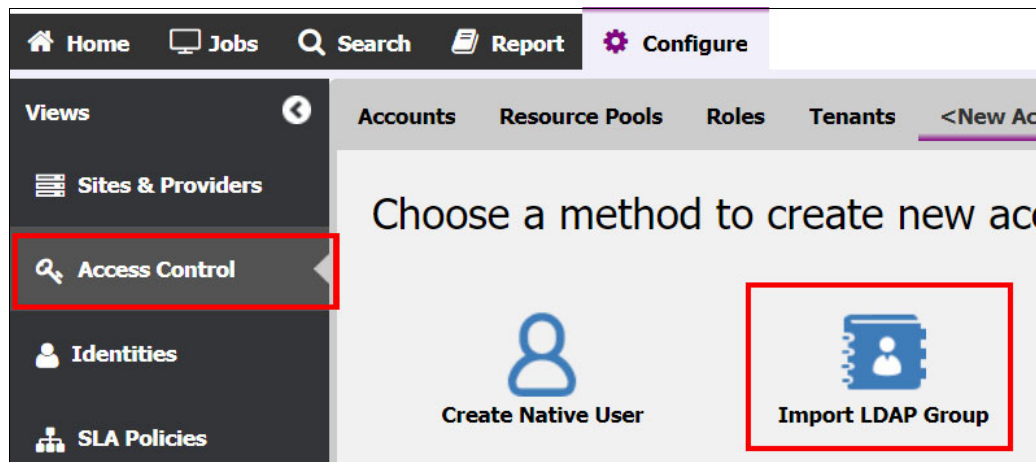


Figure 3-8 Import LDAP Group

5. Define a CDM Site that will contain your storage, vSphere, application server and Storage Sentinel components. You configure Sites on the Sites and Providers section of the Configuration tab. See Figure 3-9 on page 39.

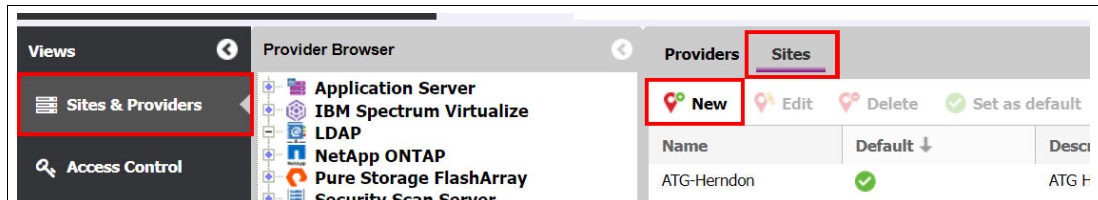


Figure 3-9 Configure Sites

6. Register your storage components, using the appropriate Storage Virtualize user accounts. This automatically adds the storage to the daily Storage Virtualize inventory job and starts an inventory of the newly registered storage. Register the storage under the IBM Spectrum Virtualize node in the Provider Browser on the Configuration page. See Figure 3-10.

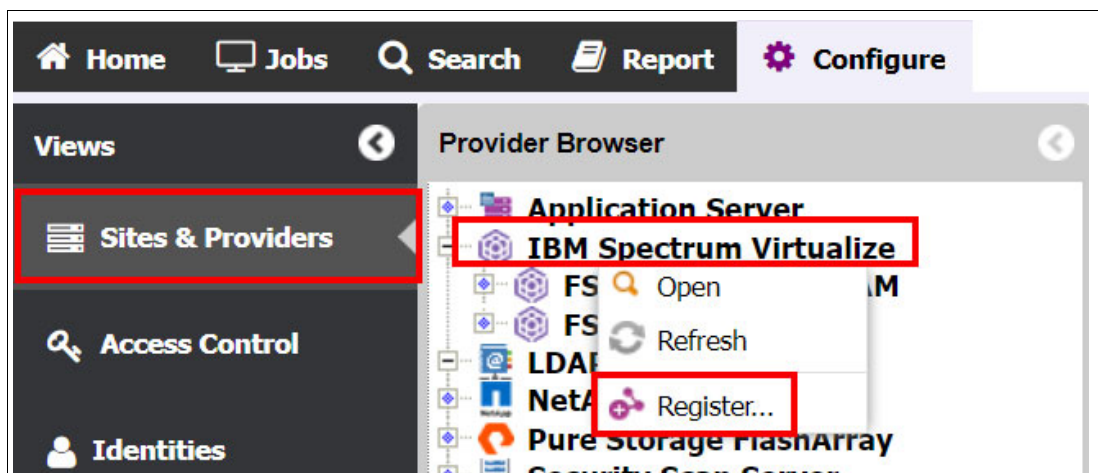


Figure 3-10 Register your storage components

7. If your application server or Storage Sentinel servers are VMs, register your vCenter server(s) to CDM using the appropriate account or certificate. This automatically adds the vSphere environment to the daily vSphere inventory job and starts an inventory of the newly registered components. Register any vCenter servers on the VMware node of the Provider Browser on the Configuration page. See Figure 3-11 on page 40.

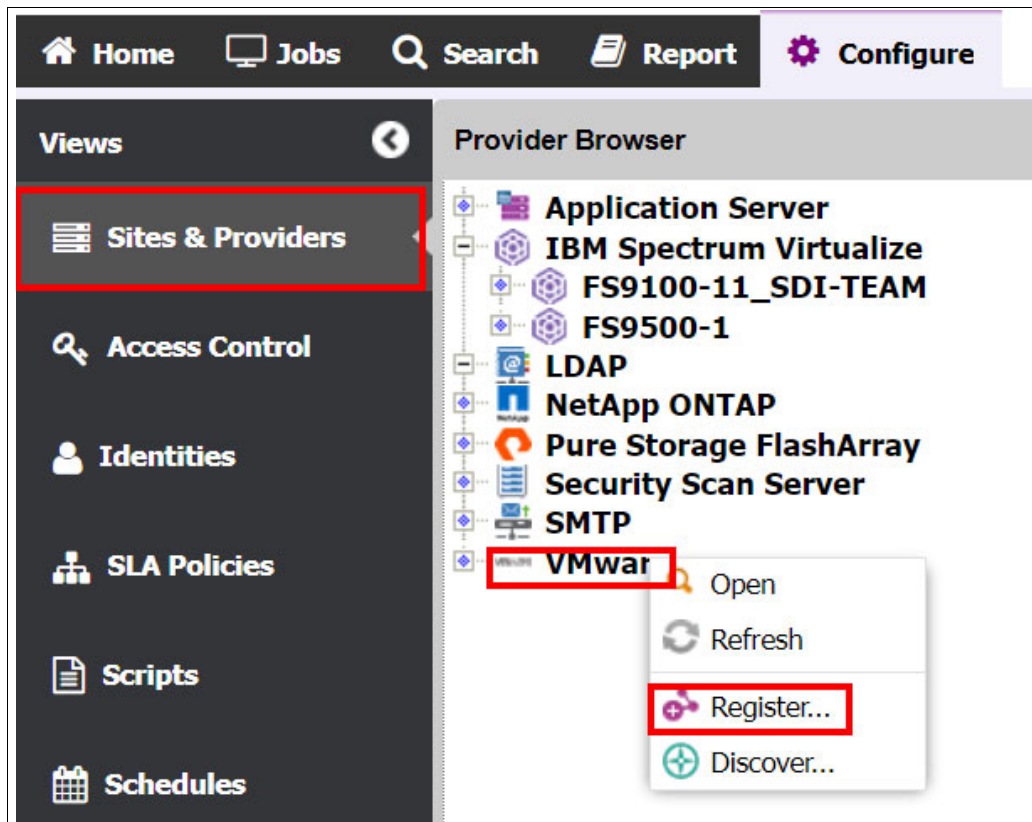


Figure 3-11 Register your vCenter server(s)

8. Register your Epic DB application servers. If your Epic DB servers are VMs, wait until the vCenter inventory finishes. See Figure 3-12.

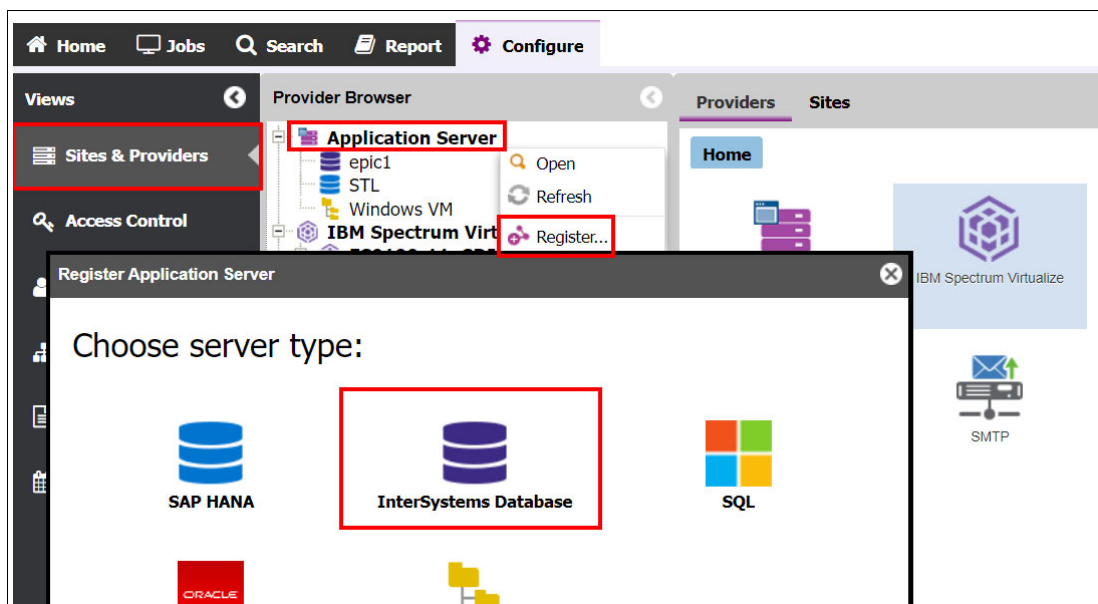


Figure 3-12 Register your Epic DB application servers

Validate the Epic VMs were found by expanding the navigation tree in the **Configuration** tab where you registered the vCenter server. You should see the ESXi

hosts and the VMs defined to vCenter in that tree, including the application servers you need to register. Register the Epic DB server in the Application Server node on the Provider Browser on the Configuration Page.

9. Register your Storage Sentinel server(s) using the Security Scanner node in the Configuration tree. If your Storage Sentinel servers are VMs, wait until the vCenter inventory has been completed and you have validated the Storage Sentinel VMs were found during the inventory by expanding the navigation tree in the Configuration tab where you registered the vCenter server. You should see the ESXi hosts and the VMs defined to vCenter in that tree, including the Storage Sentinel servers you need to register. You register the Storage Sentinel server under the Security Scan Server node on the Provider Browser in the Configuration page. See Figure 3-13.

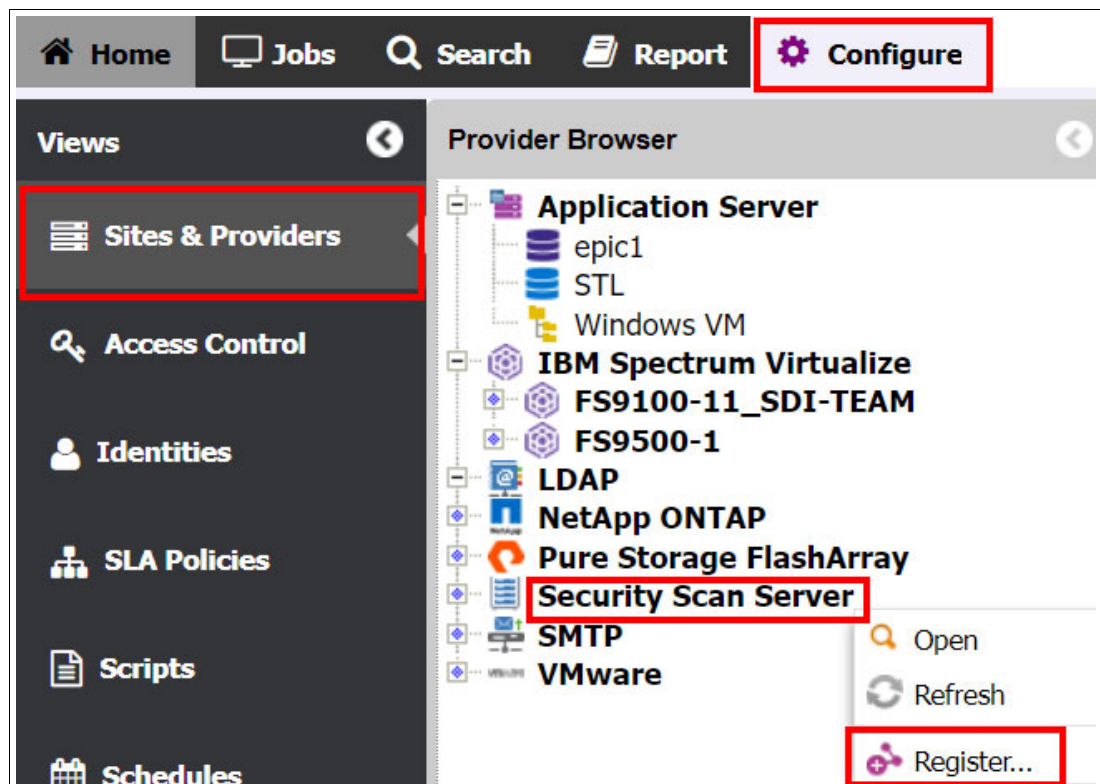


Figure 3-13 Register your Storage Sentinel server(s)

10. Validate each of your components have registered correctly. See Figure 3-14 on page 42
 - a. View the results of the Epic inventory job to validate that the Cache or IRIS instances were protected.
 - b. View the information on the application detected and the number of databases cataloged.
 - c. Navigate to the Jobs page, expand the application servers and select Intersystems.
 - d. Click on the **Default Inventory Job**.
 - e. On the bottom panel, click **History** and then click on the job log hyperlink to open the job log page.

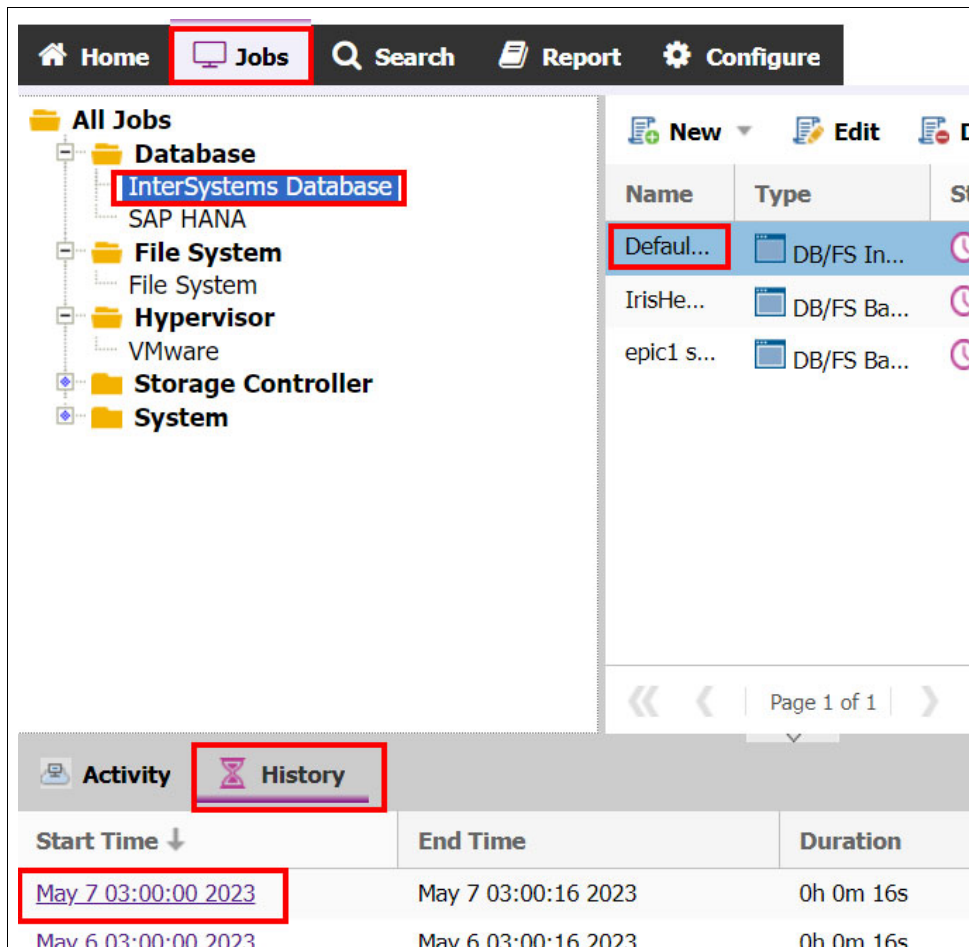


Figure 3-14 Click on the job log hyperlink to open the job log page

Figure 3-15 shows the job log page.

	May 7 03:00:16 2023	2	Cataloged 1 instance(s), 0 database group(s) 1 database(s), 8 disk(s)
--	---------------------	---	---

Figure 3-15 Job log page

11. Define an SLA for your Epic DB data protection. You will need to navigate to the SLA Policies section of the Configuration page, click **New** and then select **IBM Spectrum Virtualize**. See Figure 3-16 on page 43.

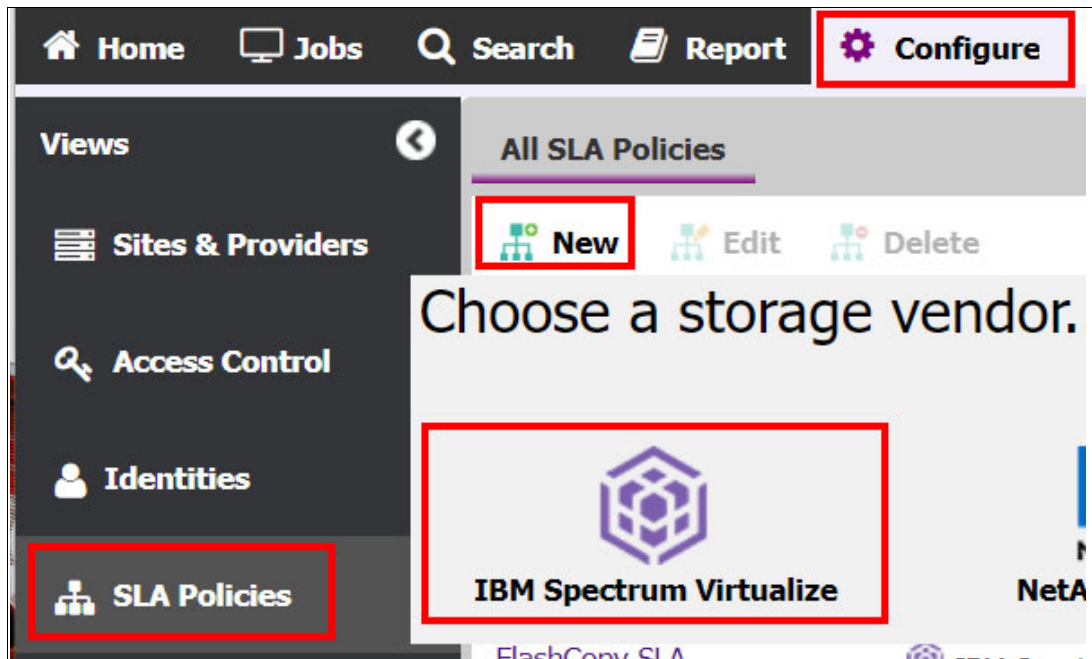


Figure 3-16 Define an SLA for your Epic DB data protection

- Give the SLA a unique name (ideally one that identifies the type of protection the SLA manages). Add a meaningful comment.
- Select the **Source** icon and enter the Frequency and Interval for this SLA.
- Right-click on **Source** and add **Safeguarded Copy**. See Figure 3-17.

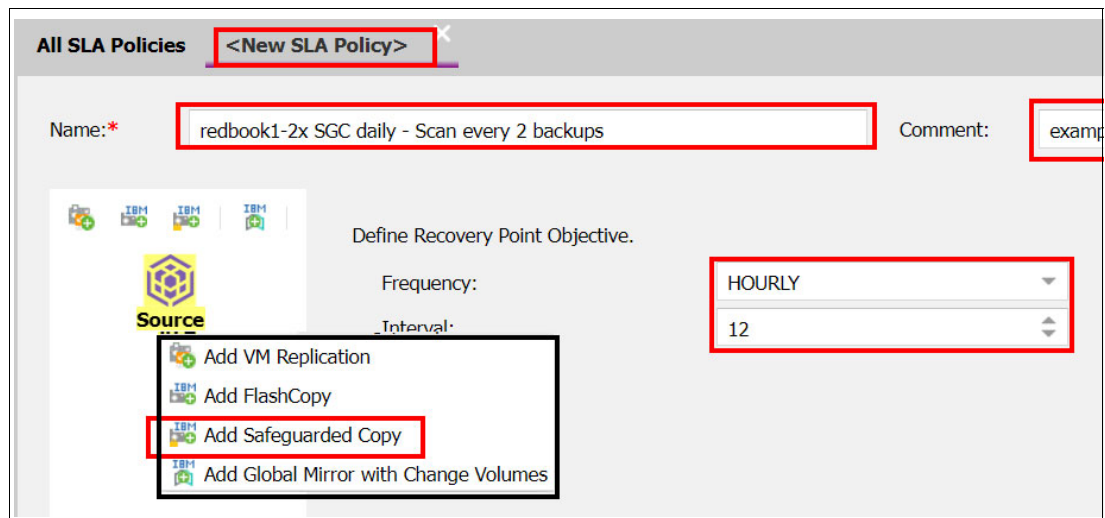


Figure 3-17 Add Safeguarded Copy

- Select the **Safeguarded Copy Volume Group** to associate with this SLA.
- Identify the number of days to retain the Safeguarded Copies.
- Give this set of Safeguarded Copies a unique and meaningful name.
- If desired, define a volume prefix for the Safeguarded Copies. It is recommended that you do use a unique and meaningful prefix to identify the CDM instance and SLA name, to help correlate Safeguarded Copies back to what created those copies.

- h. Select the checkbox to perform security scanning, identify how many backups between scans and select the Storage Sentinel instance you have previously registered for this workload. See Figure 3-18

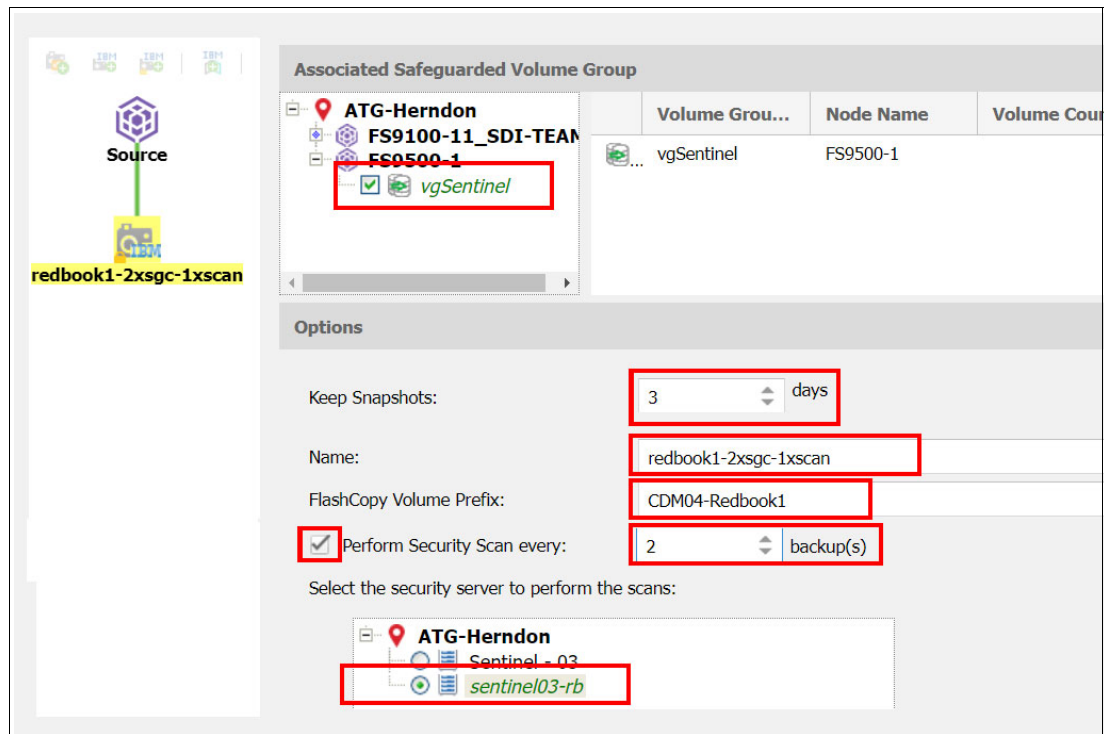


Figure 3-18 Completing the SLA configuration

- i. Save your SLA.
- 12) Define a backup job for your Epic application server(s).
 - a. Navigate to the Jobs page and expand the Database node and click on Intersystems Database.
 - b. Click on **New** and select **Backup**. See Figure 3-20 on page 45.

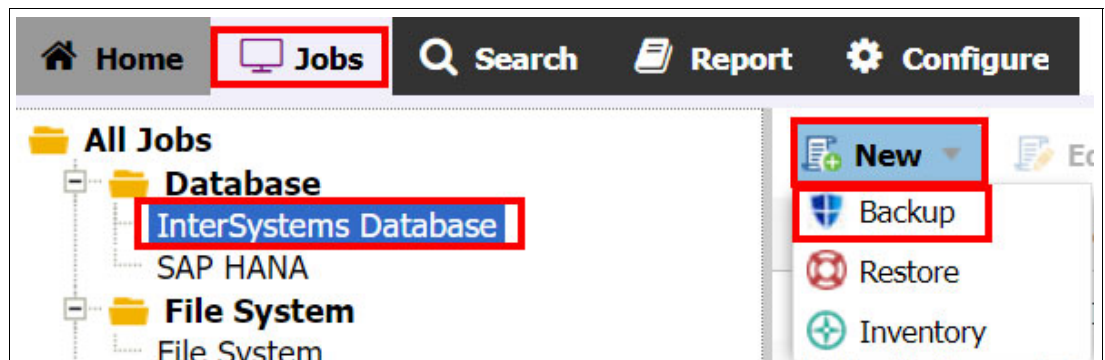


Figure 3-19 Defining a backup job

- c. Give this job a meaningful name and comment.
- d. Expand the node next to the site your Epic instance resides within and expand the Epic server's node. Check the box next to the instance or database, as you wish.
- e. Select your SLA. See Figure 3-20 on page 45.

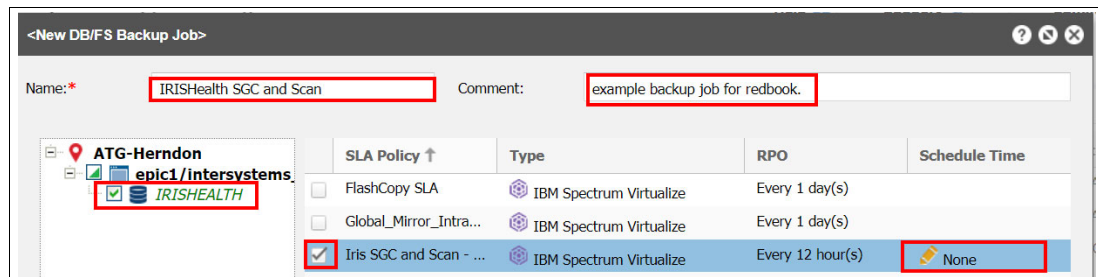


Figure 3-20 Select your SLA.

- f. Click on the **Schedule Time** column to define a schedule and set a start time. The SLA defines how often the job will run. For example, if you have defined the SLA to run every 12 hours and schedule the backup to start at midnight, it will run at midnight and at noon.
 - g. If you wish to define pre-job, post-job, pre-snapshot or post-snapshot scripts to run as part of this job, click the advanced button and identify the scripts you wish to run.
12. Monitor your data protection as shown in Figure 3-21. You can manually start the backup job at any time by right-clicking on the job definition, but you might affect your production Epic instance if you run the job at time of high activity.

Also, running a security scan might affect any running backups. After the job starts, review the job log for the backup job. Verify the instance was quiesced, a snapshot created, and the instance was unquiesced. If a security scan is scheduled to be run for this job, the job log lists activity showing copies made of the Safeguarded copies and then assigning and mounting those copies to the security scanner. When the job is finished, view the job log in the history panel. If the job executed correctly, you can define a restore job and select a recovery point, as outlined in 3.5, “Performing a restore of an Epic database backup” on page 46.

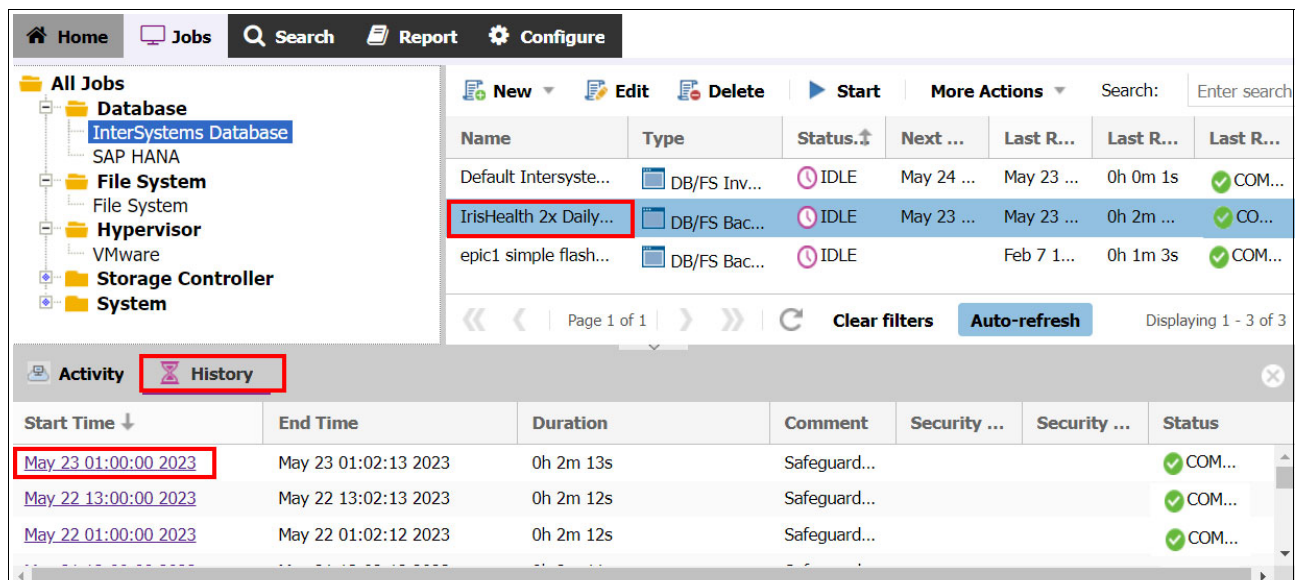


Figure 3-21 Monitor your data protection.

3.5 Performing a restore of an Epic database backup

After a scheduled job creates Safeguarded Copies of the Epic databases and scans them for malware corruption, you can perform a restore as needed.

Perform the following steps to define a Restore job.

1. On the Jobs page, expand the **Databases** node and select **InterSystems Database**. Click the **New** button and select **Restore**. See Figure 3-22.

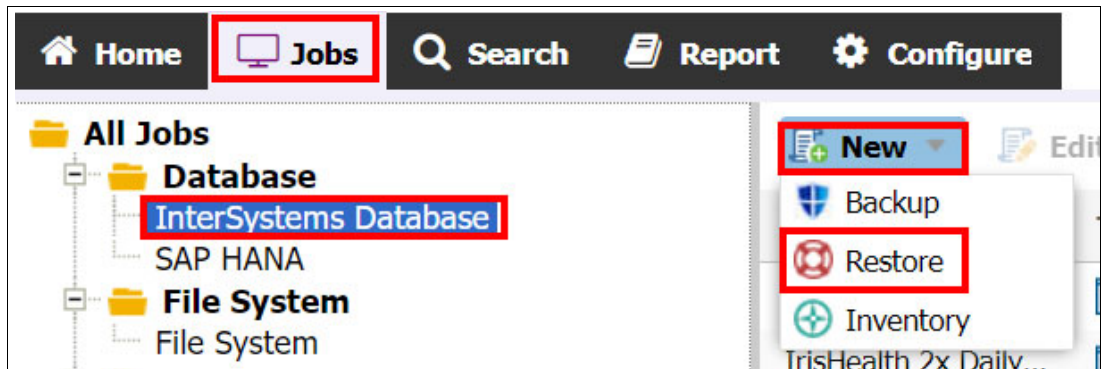


Figure 3-22 Select Restore

2. Enter a meaningful name and comment, select to perform an Instant Disk Restore or Instant Database Restore. An Instant Disk Restore mounts the file systems to the target host but does not define or start the database. An Instant Database restore defines and starts the database to an Epic database instance. For this example, select Instant Database Restore. See Figure 3-23.

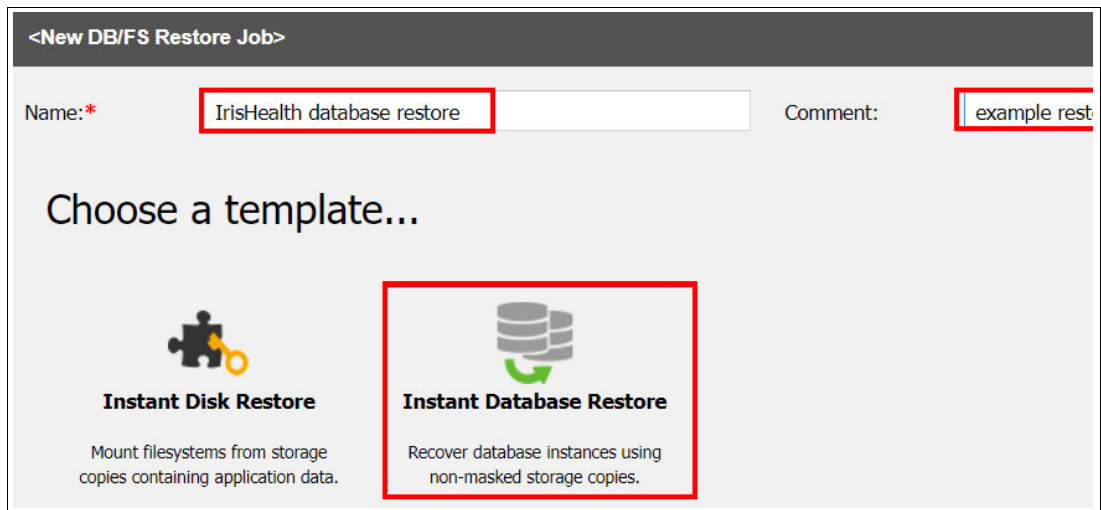


Figure 3-23 Instant Database Restore

3. The source icon will be automatically selected. Use the Application Browser tree to navigate and select the database to be restored. See Figure 3-24.



Figure 3-24 Select the database to be restored

4. Click the **Copy** icon. Click the **Select Specific Version** or **Use Latest Successful Scan** button. Here you can click the **Use Latest** or **Use Latest Successful Scan** buttons to select those recovery points. If you wish to select a specific version, click the **Use Latest** text in the Version column. See Figure 3-25.



Figure 3-25 Click the Copy icon

5. You can then select a specific version. See Figure 3-26.

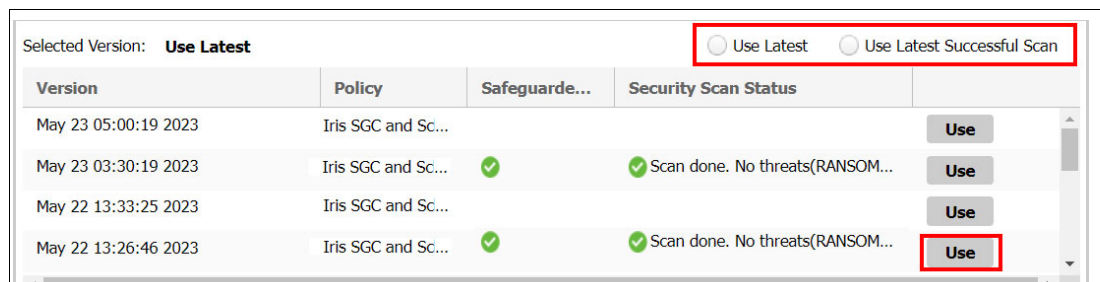


Figure 3-26 Select a specific version

6. Click the **Destination** icon. You can then select the database instance as the recovery target. Click **Create Job**. You can then start the job and monitor until completion.

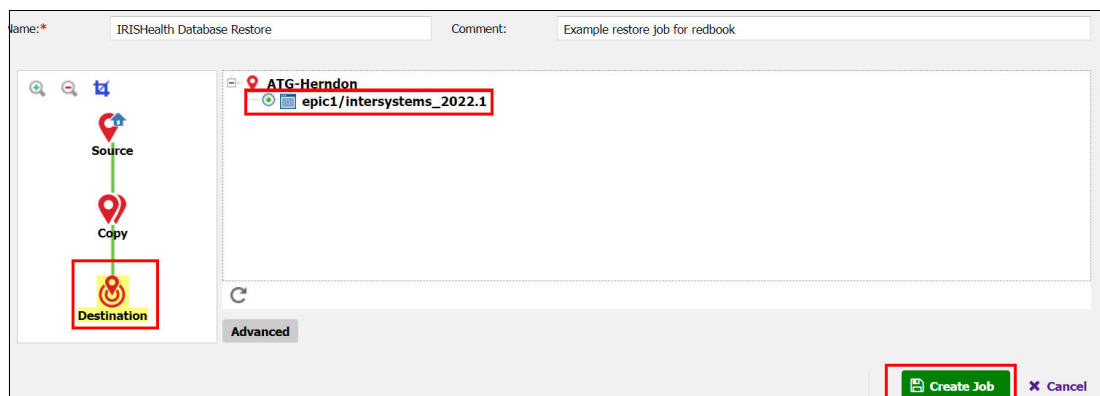


Figure 3-27 Create Job

When the job has completed mounting the volumes to the target, you will notice its status will change to Resource Active, as shown in Figure 3-28.



Figure 3-28 Resource Active

7. At this point right-click on the job hyperlink in the Activity panel and select either **Cancel Restore** or **Make Permanent**. Cancel Restore will shut down the DB, unmount the volumes and delete the temporary copies. Make Permanent will move those copies to a permanent status and complete the job.

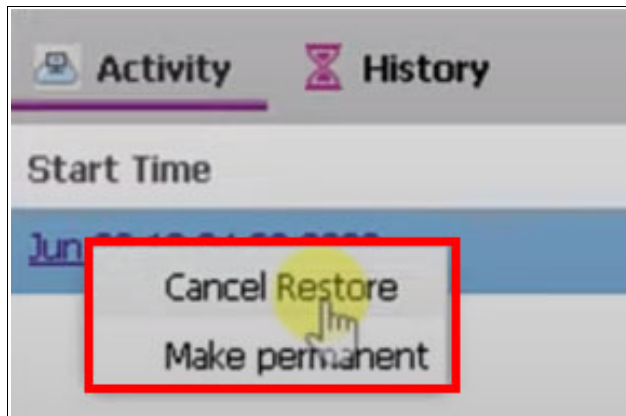


Figure 3-29 Cancel Restore and Make permanent options



Configuring IBM Storage Sentinel for SAP HANA

This chapter describes how to configure IBM Storage Sentinel to generate SAP HANA application consistent, storage-based backups and verify that the backups do not contain malware. Immutable snapshots of SAP HANA databases can be scanned, and potential corruption identified.

If an attack occurs, IBM Storage Sentinel helps to identify the best Safeguarded Copy to use for restoring the SAP HANA data. It also automates the process to restore SAP HANA data to online volumes. Because a restore uses the same snapshot technology as a backup, it is faster than the use of offline or cloud-based copies.

The main topics in this chapter are:

- ▶ “SAP HANA integration into IBM Storage Copy Data Management” on page 50
- ▶ “SAP HANA and data persistence” on page 50
- ▶ “SAP HANA workflows and IBM Storage Copy Data Management” on page 51
- ▶ “IBM Storage Copy Data Management setup” on page 57
- ▶ “Running SAP HANA backup and restore operations” on page 59
- ▶ “Daily operations, best practices and maintenance” on page 64

4.1 SAP HANA integration into IBM Storage Copy Data Management

SAP HANA (High-performance ANalytic Appliance) is a multi-model database that stores data in the server's main memory instead of keeping it on a disk. This results in data processing that is magnitudes faster compared to disk-based data systems and allows for advanced, real-time analytics. Serving as a platform for enterprise resource planning (ERP) software and other business applications.

Depending on the needs of an enterprise, SAP HANA can be deployed on premises, in the cloud, or as a hybrid system, blending the privacy and control of an on-premises system with the less cost, greater memory, and increased access of the cloud. Its ability to efficiently process enormous amounts of data makes it scalable to suit a growing business without sacrificing security or stability.

On the SAP HANA platform, developers can build their own tools and applications that integrate business logic, control logic, and the database layer with unprecedented performance.

IBM Storage Copy Data Management supports SAP HANA on the application level and includes integrated backup and restore features for the SAP HANA database. Therefore, IBM Storage Copy Data Management can run application consistent backups directly on the storage layer.

4.2 SAP HANA and data persistence

SAP HANA is in some respects different from other relational databases. While it is running, it stores all its data and its metadata in the server's main memory. Reads or updates of tables or columns are always done in memory. Usually, all the data is being read into memory when the database starts. Because data is kept in memory, SAP HANA maintains data persistence across a database shutdown. The following paragraphs describe how this works.

Note: SAP HANA can be configured in a way so that *warm data* is kept on disk. This feature is known as SAP HANA Native Storage Extension (NSE).

In the sample configuration, the underlying storage for both data and log file systems is provisioned by IBM SAN Volume Controller or IBM FlashSystem volumes. To help with designing the storage layout of SAP HANA, see: [IBM System Storage Architecture and configuration Guide for SAP HANA TDI](#).

4.2.1 SAP HANA volumes

The following list describes HANA volumes:

- Data Area

The data volumes are in a dedicated XFS file system, mounted at `/hana/data/<SID>`

- Log Area

Equivalent to the data volumes, the log volumes are stored on a different but dedicated XFS file system, mounted at `/hana/log/<SID>`

Transaction log disk operations

The transaction log saves all database transactions in chronological order. By using this log, database administrators can easily roll back to a defined point in time, or they can use the log to step forward to the most recent state of a database after restoring it from a backup. The transaction log is written continuously to disk and is always kept in sync with the actual database transactions. The log volume is overwritten when it fills up. Therefore, the log is backed up by SAP HANA frequently. Log backups can be maintained by third-party backup software like IBM Storage Protect for ERP. Using such a long-term backup solution, even older database backups can benefit from the log roll forward feature.

Transaction data disk operations

SAP HANA dumps its columns, tables, and metadata periodically to the data area. These dumps are also known as *savepoints*. Savepoints are usually started every five minutes, but the frequency can be configured in SAP HANA. When a savepoint is written, the former content is overwritten. Savepoints will be read when the database starts up. In addition to savepoints, SAP HANA backup operators and administrators can create database snapshots. A snapshot behaves like a regular backup. It contains a fixed time stamp and can be added to the database's backup catalog. Snapshots are written into the data area as regular files. Therefore, additional action is required to make a snapshot available to the backup catalog. The snapshot needs to be saved by either copying it to a safe location or by snapshotting the data area by using IBM FlashCopy. After saving it, the snapshot will be committed to the database and SAP HANA will remove it from the data area afterward.

4.3 SAP HANA workflows and IBM Storage Copy Data Management

This section includes discussions of the requirements and configuration of SAP HANA and IBM Storage Copy Data Management so that SAP HANA workflows can be run.

4.3.1 SAP HANA data backup workflow

The complete SAP HANA data backup workflow is controlled by IBM Storage Copy Data Management and includes several steps:

- ▶ Creation of the SAP HANA snapshot via SAP HANA SQL statements
- ▶ Freezing the XFS file system
- ▶ Taking a IBM FlashCopy of the underlying SAN volumes
- ▶ Committing the SAP HANA snapshot to the database.

The overall workflow is shown in Figure 4-1 on page 52.

SAP HANA – Data Volume Snapshot

Steps to create Storage Snapshot:

1. Create an internal data snapshot in SAP HANA using SQL command („prepare database“)
2. Freeze the DATA filesystem (xfsfreeze)
3. Create a FlashCopy of the whole data area (data vg)
4. Unfreeze the DATA filesystem
5. Confirm the snapshot as successful using SQL commands. This is necessary to include the snapshot in SAP HANA's backup catalog

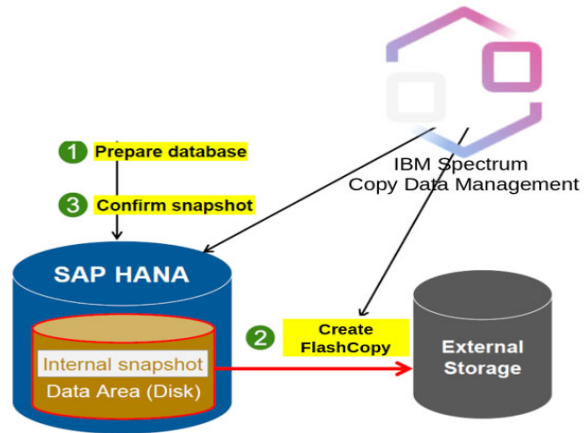


Figure 4-1 IBM Storage Copy Data Management - SAP HANA backup workflow

In a first step, the registered components are scanned by IBM Storage Copy Data Management. The information collected by this scan is stored in the internal database of IBM Storage Copy Data Management. This allows IBM Storage Copy Data Management to perform backup and restore jobs fast, without the need to re-scan the systems before and after running a backup job.

To backup an SAP HANA database, IBM Storage Copy Data Management needs to know the complete data path from the SAP HANA data volumes down to the storage volumes where this data is stored. For guidelines to configure the storage for SAP HANA properly, see: [IBM System Storage Architecture and configuration Guide for SAP HANA TDI](#).

A schematic view of the data path is shown in Figure 4-2 on page 53.

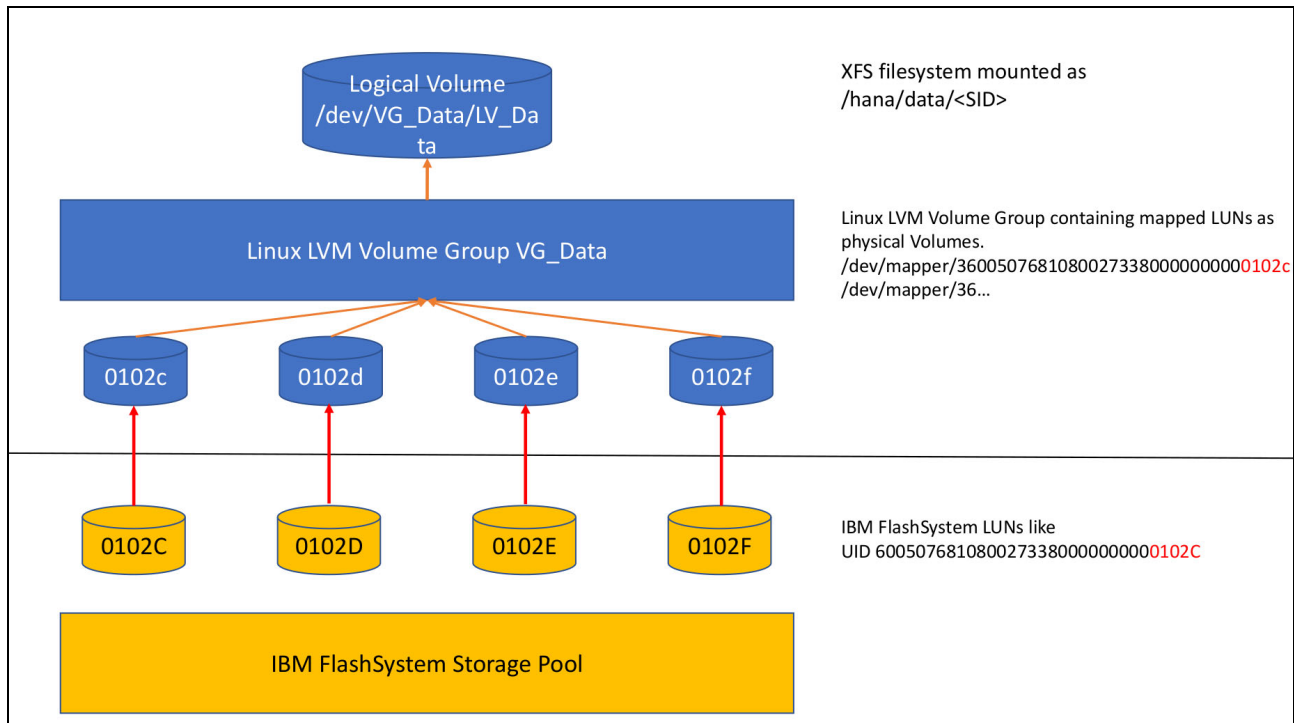


Figure 4-2 SAP HANA Data Volume: Data path in a SAN environment

In this example, IBM Storage Copy Data Management will identify the four Spectrum Virtualize volumes 0102C – 0102F as the LUNs holding the SAP data. These volumes will be backed up using the IBM Storage Virtualize FlashCopy feature.

When using FlashCopy, the storage system must ensure that all involved volumes are in a consistent state when the FlashCopy operation starts. I/O operations are not allowed during this time to ensure consistency of the copied data. Spectrum Virtualize uses consistency groups to guarantee time consistent flash copies.

With Safeguarded Copy, the write consistency is managed by the IBM Storage Virtualize volume group. However, at the time of writing, IBM Storage Copy Data Management supports FlashCopy with Consistency Groups or Spectrum Virtualize Snapshots when it flashes the volumes. Figure 4-3 on page 54 illustrates this behavior. The FlashCopy operation takes a few microseconds, so there is no measurable impact to the host IO performance. The supported version of Safeguarded Copy includes the following characteristics:

- Available as of Storage Virtualize 8.4.2
- Requires the creation of a child pool to store the immutable Safeguarded Copies
- Requires an external scheduler, such as IBM Storage Copy Data Management
- Can require management of FlashCopy mappings

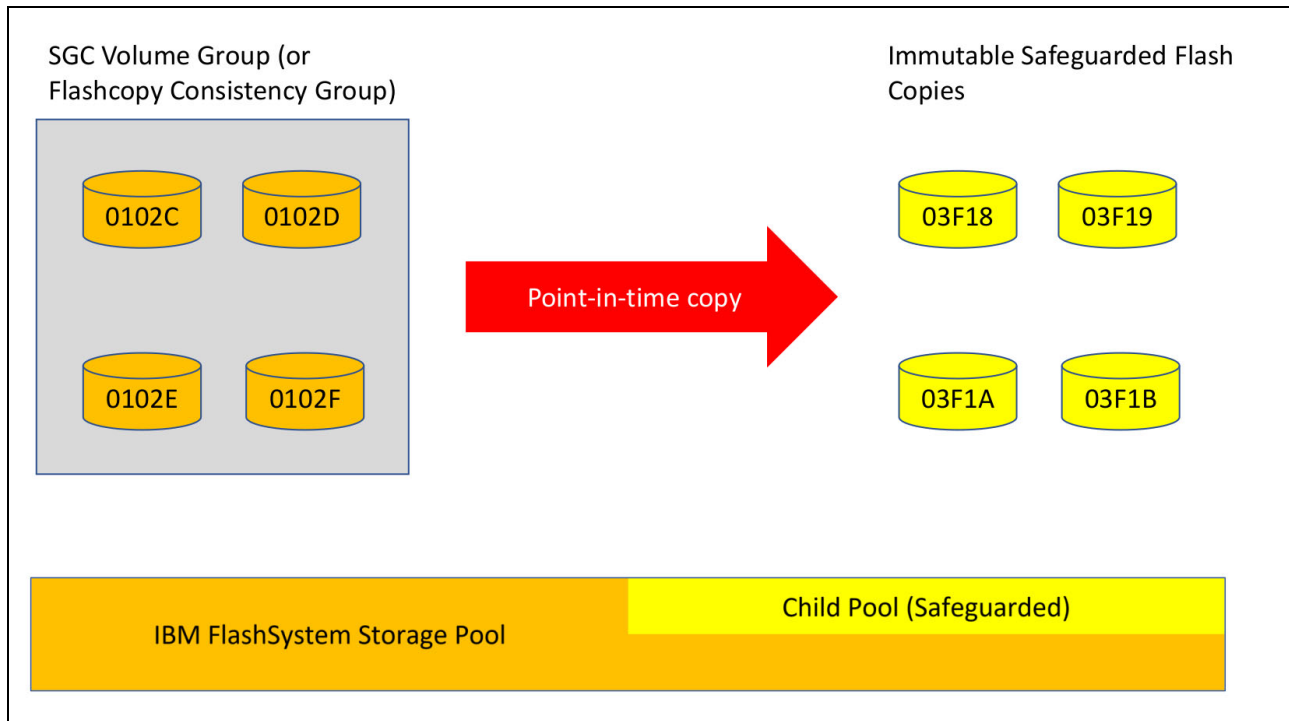


Figure 4-3 FlashCopy using Safeguarded volumes

Before creating FlashCopies of the data volumes, IBM Storage Copy Data Management prepares the SAP HANA database by running a set of SQL statements. The statements are run by an SAP HANA Database Backup Operator account. An SAP HANA snapshot, which is an online backup of the SAP HANA database, is written to the SAP HANA data volume. As soon as the snapshot has been created, I/O operations to the XFS file system are suspended, and a FlashCopy of the volumes is created. When the FlashCopy operation is finished, I/O operations are resumed. As a last step, additional SQL statements are generated to inform the database that the snapshot has now been saved or *committed*. SAP HANA stores the backup information into its backup catalog and removes the snapshot file from the data area.

With IBM Storage Sentinel, the Scanning Engine becomes an active part of the SAP HANA backup workflow. In previous versions of IBM Storage Copy Data Management, an SAP HANA backup job finished after it completed the steps shown in Example 4-1 on page 56. Storage Sentinel extends the workflow by mapping the Safeguarded Copy backup to the Scanning Engine, so that it can be analyzed there.

Because Safeguarded Copy volumes cannot be mapped to a host for scanning, IBM Storage Copy Data Management runs some additional steps:

1. Defines the Safeguarded Copy volumes as source volumes.
2. Creates additional FlashCopies from the Safeguarded Copy source volumes.
3. Mounts the additional FlashCopies to the Scanning Engine as working copies.
4. Runs a scan using the Scanning Engine
5. Reports the scan results to IBM Storage Copy Data Management. Depending on the security scan results, the backup will be marked as either successful or failed by IBM Storage Copy Data Management
6. Unmounts and removes the FlashCopy target volumes.

4.3.2 SAP HANA restore workflow

IBM Storage Copy Data Management offers two different SAP HANA restore scenarios:

Instant Disk Restore

Instant Disk Restore is a *crash consistent* restore of the SAP HANA data area. After the data area has been restored, the SAP HANA database administrator recovers the database. If the database transaction log is available, they can perform a roll forward operation to the most recent state of the database and reduce the Recovery Point Objective (RPO).

Instant Database Restore

Instant Database Restore is an *application consistent* restore of the SAP HANA database. After restoring the data area with, for example, a Disk Restore, IBM Storage Copy Data Management will recover the database using the SAP HANA snapshot, which is now available as a regular file in the data area. Any existing transaction log will be cleared by the database restore.

These scenarios apply to different use cases, which define what kind of restore is required. For example, in case of a crash, a Disk Restore should be chosen, to minimize the RPO and to give the SAP HANA administrators full control of the recovery process. However, for some use cases a Database Restore makes more sense compared to a Disk Restore, such as when restoring a test or development database to a previously defined point in time.

Important: When IBM Storage Copy Data Management Database Restore runs, the transaction log is deleted without further confirmation. The Database Restore is an automated SAP HANA recovery process which restores the database using the snapshot only. During the Database Restore, any transaction log that might allow a roll forward to a more recent state is deleted.

The behavior of a restore job, whether it is a Disk Restore or Database Restore, can be configured in the advanced restore job settings.

- Mountpoint name

A backup can be restored to either its original mountpoint or to a new one. A new mountpoint can be used for temporary restore operations, for example to check if the data is complete.

- Restore volume handling

IBM Storage Copy Data Management does not map the Safeguarded Copy volumes directly to the host. The default behavior is to create a set of additional recovery volumes and populate these by using any Safeguarded Copy as a source volume.

The mappings are started with the **nocopy** option, which means that the restore volumes stay dependent on the backup volumes. If you need independent volumes, this can be changed by making the restore volumes permanent. IBM Storage Virtualize will copy the content in the background.

Instead of creating a new set of recovery volumes, the existing SAP HANA production volumes, which are already mapped to the host can be overwritten with the content of the Safeguarded Copy volumes. This is called *reverse restore* and is the fastest way to restore data, and preferred for disaster recovery operations.

- Logical Volume Manager changes

If the file system containing the SAP HANA data area is organized in an LVM structure, IBM Storage Copy Data Management can rename the restored volume group as well as the physical volume IDs. This allows coexistence of both the original data and a cloned set of this data on the same host.

Note: Not all features are available for all environments. For example, the *reverse restore* feature has restrictions in VMware environments when the protected data is stored on VMware datastores.

4.3.3 SAP HANA requirements

This section explains the main requirements and configuration steps that have to be done on the SAP HANA application server before registering the SAP HANA server in IBM Storage Copy Data Management 2.2.19. IBM Storage Copy Data Management installs an agent on the SAP HANA application server when the SAP HANA server is registered.

Prerequisites on the SAP HANA application server:

- The SAP HANA data, log and operating system (OS) file systems must be on separate volumes. Ensure, that the SAP HANA data and log volume group contain volumes only from a single and supported storage system.
- The SAP HANA client must be installed on the SAP HANA application server.
- Create a symbolic link to the SAP HANA client installation directory through the following command: `ln -s /hana/shared/<SID>/hdbcli /opt/hana.`
- Edit the sudo settings for the HANA administrative user (for example, *stladm*, where *STL* is the SID) to allow the user to run all commands without having to enter a password. Always use the **visudo** command to edit the `/etc/sudoers` file, because **visudo** validates the sudoers file when it is saved. Add the lines shown in Example 4-1.

Example 4-1 visudo settings

```
stladm ALL=(ALL) NOPASSWD: ALL
Defaults:stladm !requiretty
```

- For SAP HANA 2.0 SPS 04 and SPS 05, the `hdbcli` module must be installed. Install the `hdbcli` module *after* the complete installation of the SAP HANA client. Compile and install the binary `hdbcli` python module, as shown in Example 4-2.

Example 4-2 Installing the hdbcli module

```
[root@hana-host ~]# cd /hana/shared/<SID>/hdbcli
[root@hana-host ~]# tar vfxz hdbcli-<VERSION>.tar.gz
[root@hana-host ~]# cd hdbcli-<VERSION>
[root@hana-host ~]# python3 setup.py install
```

- If SAP HANA is installed on a virtual machine in VMware that is using VMFS datastores for SAP HANA data and log volumes, UUID must be enabled to allow IBM Storage Copy Data Management to discover the LUN IDs of the volumes. To enable it, power off the virtual machine, then select the SAP HANA virtual machine in the vSphere Client and click **Edit Settings**. Select **VM Options**, then the **Advanced** section. Select **Edit Configuration**, then find the **disk.EnableUUID** parameter. If set to **FALSE**, change the value to **TRUE**. If the parameter is not available, add it by clicking **Add Row**, set the value to **TRUE**. Save the changes, then power on the virtual machine.

For further configuration details like supported storage systems or OS versions, see [Requirements Documents: IBM Storage Copy Data Management](#).

4.4 IBM Storage Copy Data Management setup

As outlined in 4.3, “SAP HANA workflows and IBM Storage Copy Data Management” on page 51, IBM Storage Copy Data Management requires management access to all components which are involved in backup or restore jobs including the following components:

- ▶ IBM Storage FlashSystem or IBM SAN Volume Controller
Configuration of IBM Storage FlashSystem is discussed in Chapter 2, “Configuring the IBM Safeguarded Copy feature” on page 17.
- ▶ SAP HANA database
Database access is required for several database queries, for running snapshots and for performing automated database recoveries.
- ▶ SAP HANA host
Host access is required for performing storage related operations, but also for OS related SAP HANA tasks.
- ▶ VMware vCenter (only required if SAP HANA runs in a virtualized environment)
Storage attachment and mapping is managed by VMware, requires a vCenter user with appropriate permissions.
- ▶ IBM PowerVM®
LPARs are treated as physical hosts.

4.4.1 Required user roles

When planning for IBM Storage Copy Data Management related users, a strategy of minimum required privilege should be mandated. For example, there is no need to add a vCenter administrator user to IBM Storage Copy Data Management. Instead, create a dedicated user which has access to the involved resources only, and which also has no other capabilities than the minimum required. To learn more about VMware user permissions, see [vSphere Permissions and User Management Tasks](#).

The SAP HANA host user used by IBM Storage Copy Data Management is typically the OS database administrator. This user is automatically created during the installation of SAP HANA. It follows a straight forward naming scheme, usually <SID>adm, where the SID is written in lowercase letters, for example “stladm” if the SAP HANA SID is STL. This user does not own further permissions or capabilities in the operating system. To keep the user administration in IBM Storage Copy Data Management as simple as possible, the database administrator is used by IBM Storage Copy Data Management for both database management and OS administration. Therefore, the user’s privileges will be enhanced by configuring the **sudo** utility as described by section 4.3.3, “SAP HANA requirements” on page 56.

For administering the SAP HANA database itself, a dedicated DB user should be created. A database user needs to be defined inside the database, it has no dependencies to the operating system. SAP HANA grants privileges by assigning users to previously defined roles. It is a best practice to follow a minimum required privilege strategy. For further information about SAP HANA backup users, refer to [Authorizations Needed for Backup and Recovery](#).

Note: The predefined “DATABASE BACKUP OPERATOR” role is sufficient for most SAP HANA tasks triggered by IBM Storage Copy Data Management, except IBM Storage Copy Data Management’s database recovery restore job. For database recovery, assign the “DATABASE RECOVERY OPERATOR” or “DATABASE ADMIN” role.

4.4.2 Service Level Agreement (SLA) policies

The SLA policy defines how to protect a database or file system application and uses specific features of the underlying storage system. Therefore, separate policies are predefined for different storage systems. Because many storage systems offer multiple backup features, additional policies may exist for a specific storage system. For IBM Storage Virtualize, four different types of SLA policies are available:

- ▶ VM Replication (VMware only feature)
- ▶ FlashCopy
- ▶ Safeguarded Copy
- ▶ Global Mirror with change volumes

The SLA policy also defines the RPO, the backup retention time, and the system which performs the backup. This grants the storage administrator full control of the backup process. For example, the SLA policy defines not only the target storage system but also the target pool and other storage related settings, such as volume prefixes. It enables strong separation of duties: the application backup operator does not configure anything on the storage system but uses predefined SLAs as backup policies for the application.

The objects for creating a Safeguarded Copy SLA policy are described in Chapter 2, “Configuring the IBM Safeguarded Copy feature” on page 17.

IBM Storage Copy Data Management has an integrated wizard for creating SLA policies. It requires a unique name for the policy and the desired RPO target. Next, the FlashSystem volume group for SAP HANA needs to be selected. Finally, the retention rules and the security scanning server of IBM Storage Sentinel need to be set, as shown in Figure 4-4 on page 59.

Name: * SAP HANA SGC FC

Comment:

Associated Safeguarded Volume Group

ATG-Herndon

- FS9100-11_SDI-TEAM
 - SAP_HANA_STL_DATA
 - FS9500-1
 - vgSentinel

Volume Group Name	Node Name	Volume Count
SAP_HANA_STL_DATA	FS9100-11_SDI-TEAM	4

Options

Keep Snapshots: 10 days

Name: Safeguarded Copy0

FlashCopy Volume Prefix: STL_SGC

☒ Perform Security Scan every: 1 backup(s)

Select the security server to perform the scans:

ATG-Herndon

Sentinel - 03

Figure 4-4 Define an SLA policy with Safeguarded Copy and security scan

There is one small but important change in the Safeguarded Copy SLA design compared to all former SLA definitions:

Traditionally, a SLA does not define backup sources. An SLA is a set of rules describing *how* to back up an application, but not *what* to back up. The latter is described by the job definition. One reason for this design is to achieve a proper separation of duty. For example, storage administrators should do storage system related tasks, and application admins are responsible for backing up the application itself.

Safeguarded Copy depends on a volume group in IBM Storage Virtualize. Volumes can be logically organized in such a volume group. The idea behind this feature is to provide an easier way to, for example, assign policies and create snapshots on a group of volumes. The volume group must exist before the creation of the SLA, and it will be associated with that SLA. This means that the SLA defines the backup source.

4.5 Running SAP HANA backup and restore operations

This section describes how to run a backup and restore job for SAP HANA in an IBM Storage Sentinel environment. In a first step the registered components are scanned by IBM Storage Copy Data Management. The information gathered by this scan is stored in the internal database of IBM Storage Copy Data Management. After a backup has been done, the backup can be configured to be automatically scanned by the anomaly scanning engine.

4.5.1 Running an SAP HANA backup job

After a backup job has been defined in IBM Storage Copy Data Management, it can be started either manually or by using the built-in scheduler. In this example we start the backup job manually.

1. Click the **Jobs** tab.
2. Select the SAP HANA backup job to run by clicking in the row containing the job name, as shown in Figure 4-5.
3. Click **Start**, or right-click the job name and select **Start**. A confirmation dialog box opens.
4. Click **Yes**. The job session runs.

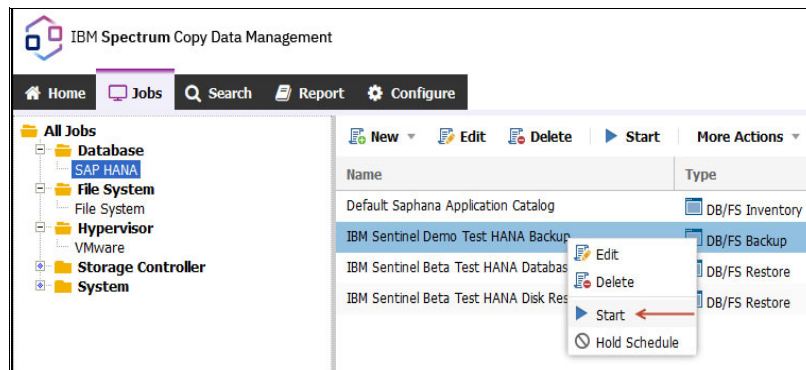


Figure 4-5 Start an SAP HANA backup job

When the backup is completed successfully, the backup job reflects the status as COMPLETED. If the backup has the Status FAILED, it could either be due to a setup error or the anomaly scanning engine has detected an infected backup. To check for the cause of a failed backup, view the backup activity log.

In the lower left of the IBM Storage Copy Data Management GUI is the Activity pane. Click the job name to view the activity log for the specific job, including the job session's start date and time, duration, description, status, and associated messages, as shown in Figure 4-6. In this backup log example, the scanning engine has detected a possible malware threat in the backup data. For this reason, the backup job has the status FAILED.

IBM Sentinel Demo Test HANA Backup@Nov 9 08:11:45 2022									
Tasks		Details		Log					
ID	Type	Duration	Status	Message	Type	Time ↑	Task...	Message	
1	Resolve	0h 0m 0s	CO...	COMPLETED	i	Nov 9 08:16:36 2022	2	[10.0.240.162] Resignaturing volume group VG_STS_DATA	
2	Protection (saphana)	0h 14m 1s	FAIL...	FAILED	i	Nov 9 08:16:36 2022	2	[10.0.240.162] Importing volume group with new name VG_STS_DATA physical volumes /dev/sde /dev/sdc /dev/sdb /dev/sdd	
	Finding data and log disks of databases :	Done (Total:4)			i	Nov 9 08:16:37 2022	2	[10.0.240.162] Mounting xfs volume /dev/mapper/VG_STS_DATA004 LV_STS_DATA to mount point /tmp/mounts/10_0_240_156/1005/ses	
	Performing pre snapshot operations:	Done (Total:1)			i	Nov 9 08:16:42 2022	2	[10.0.240.162] Completed mount operation in 3m 9s, 1 volume(s) mo	
	Performing post snapshot operations:	Done (Total:1)			i	Nov 9 08:16:44 2022	2	Successfully and 0 volume(s) failed	
	Cataloging objects:	Done (Total:46)			i	Nov 9 08:16:44 2022	2	Security Scanning of protected databases	
	Load storage data:	Done (Total:4)			i	Nov 9 08:16:44 2022	2	Starting Index job on mount path /tmp/mounts/10_0_240_156/1005/	
	Load VMware data:	Done (Total:3)			i	Nov 9 08:16:49 2022	2	name 1005...	
	Load host data:	Done (Total:1)			i	Nov 9 08:16:50 2022	2	Index job (98) created.	
	Mount snapshot copies:	Done (Total:4)			i	Nov 9 08:25:13 2022	2	Index job (98) started.	
	Map LUNs:	Done (Total:4)			i	Nov 9 08:25:13 2022	2	Security Scan finished with state: Done. Previous threat detected: fa	
	Locate LUNs:	Done (Total:4)			i	Nov 9 08:25:13 2022	2	Number of new threats detected: 1.	
	Mount VM disks:	Done (Total:4)			i	Nov 9 08:25:13 2022	2	Unmounting database snapshot copies after Security Scanning	
	Security Scanning of protected databases:	Done			i	Nov 9 08:25:13 2022	2	ECX log dir=/data/log/ecxdeployer/2022-11-09/61b28b91-fcaa-41c2-b	
	Unmount VM disks:	Done (Total:4)			i	Nov 9 08:25:16 2022	2	e42c1369e60d	
	Dismount snapshot copies:	Done (Total:1)			i	Nov 9 08:25:16 2022	2	Guest tools on 10.0.240.162 already at latest version: 2.17.21	
	Rescan storage:	Done							
	Condensing catalog:	Done							

Figure 4-6 Backup job log with error message for infected backup

Backup jobs can be scheduled by the built-in IBM Storage Copy Data Management scheduler, or you can use standard job automation tools to start IBM Storage Copy Data Management backup jobs using REST-API calls.

For email notifications about backup and restore jobs, at least one SMTP server must be configured in IBM Storage Copy Data Management, as shown in Figure 4-7. The SMTP server must be added to IBM Storage Copy Data Management, before defining a backup job.

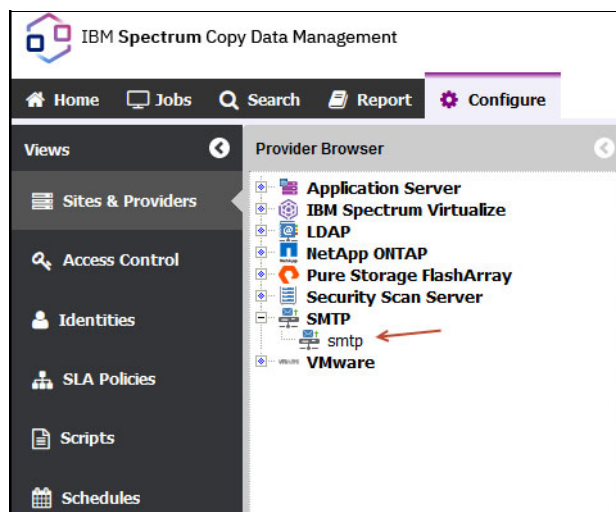


Figure 4-7 SMTP server registered in IBM Storage Copy Data Management.

IBM Storage Copy Data Management backup jobs can be configured to create status emails for every backup job. An example of an SAP HANA backup job status email is shown in Figure 4-8. This email also includes the job log.

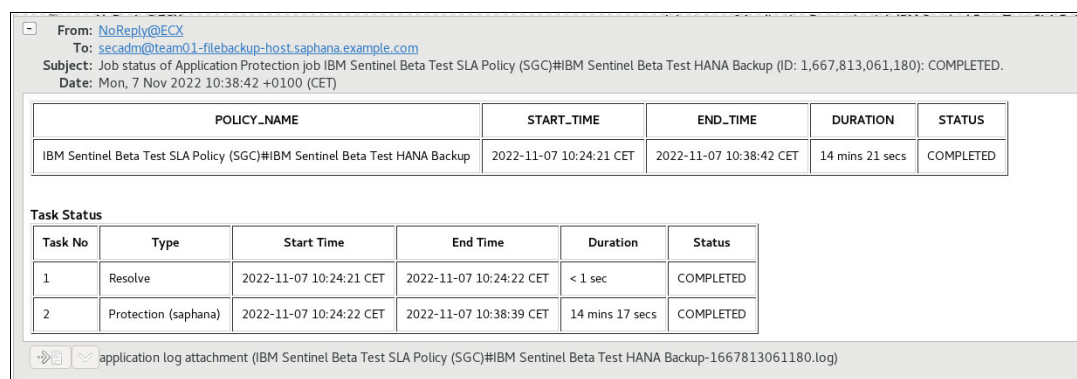


Figure 4-8 Backup job e-mail send by IBM Storage Copy Data Management

4.5.2 SAP HANA restore job

An IBM Storage Copy Data Management Database Restore performs a complete recovery of the SAP HANA data volumes using the selected FlashCopy backup on the IBM FlashSystem. There are two types of restore, Instant Disk Restore and Instant Database Restore, which are described in 4.3.2, “SAP HANA restore workflow” on page 55. In an IBM Storage Sentinel environment non-infected backups only are restored.

Performing an SAP HANA Instant Database Restore job

An IBM Storage Copy Data Management Instant Databases Restore job performs a complete recovery of the SAP HANA database using the selected FlashCopy backup on the IBM FlashSystem. When the database restore job has finished, the SAP HANA database is running and recovered to the point in time when the backup was made.

1. Login to the IBM Storage Copy Data Management GUI and click the **Jobs** tab. Expand the **Database** folder, then select **SAP HANA**. Click **New** and select **Restore**, as shown in Figure 4-9.

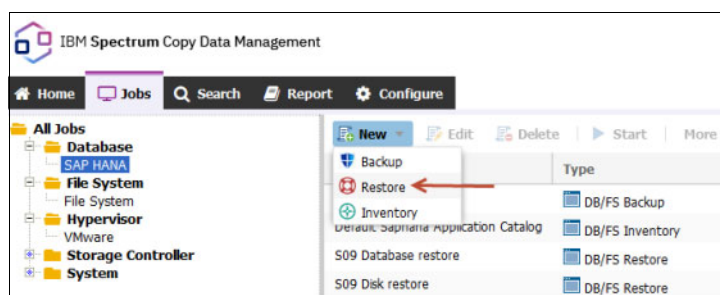


Figure 4-9 Create a HANA restore job

2. Enter a name for your job definition and a meaningful description. Select a template. Available options include **Instant Database Restore** and **Instant Disk Restore**. In this example, select **Instant Database Restore**, as shown in Figure 4-10.

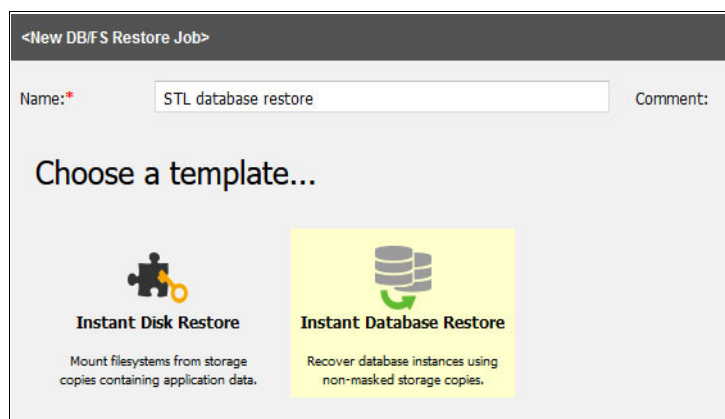


Figure 4-10 Choose the template "Instant Database Restore"

- Click on the **Source** icon. From the drop-down menu select the Application browser to select a source site and the registered SAP HANA application server to view available database recovery points.

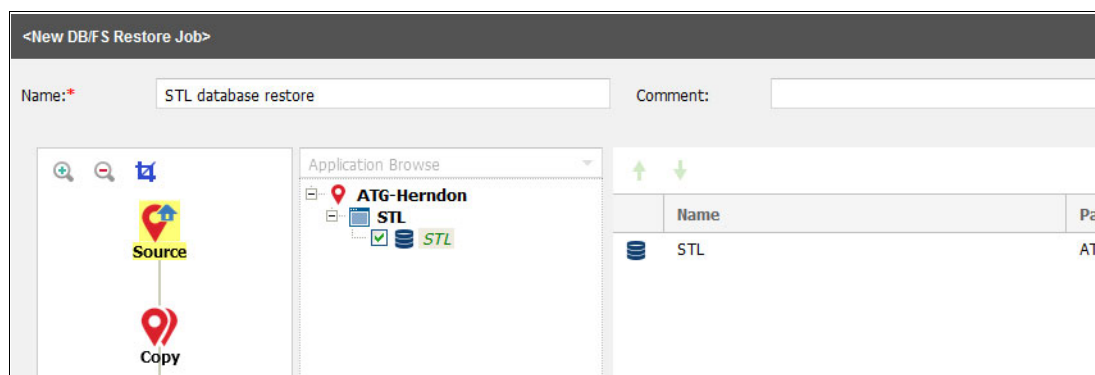


Figure 4-11 Choose the source (HANA application server) for the restore.

- Click **Copy**. The list of sites contain copies of the selected data. Select a site. By default, the latest copy of your data is used. To choose a specific version, select a site and click **Select Version**. Click the **Version** field to view specific copies and their associated job and completion time. See Figure 4-12.

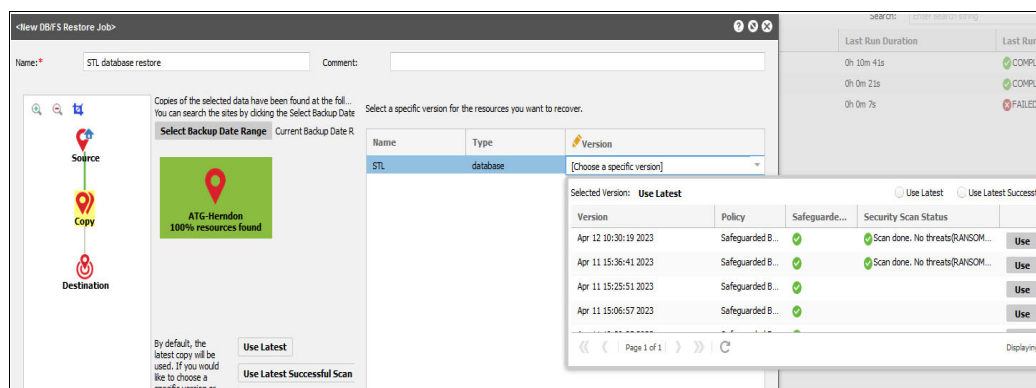


Figure 4-12 Choose a specific backup version

- Click **Destination**. Select a source site and an associated destination. The database restore can be performed on the original SAP HANA application server.
- Before creating the restore job definition, click the **Advanced** button. Set the job definition options with the following values:
 - Do not rename
Select this option to not rename mount points during recovery. IBM Storage Copy Data Management will mount them with the same path and name as the source.
 - Revert: Enabled
SAP HANA databases can be restored using the snapshot revert feature. Supported on PureStorage and IBM Storage. This decreases the restore time for large databases.
 - Protocol Priority: Fibre Channel
In this environment the protocol used for the attached volumes is fibre channel.
- When the restore job information is correct, click **Create Job**. The job runs as defined by a schedule, or can be directly started from the **Jobs** tab.

Performing an SAP HANA Instant Disk Restore job

An IBM Storage Copy Data Management Instant Disk Restore job performs a complete recovery of the SAP HANA database using the selected FlashCopy backup on the IBM FlashSystem.

An IBM Storage Copy Data Management instant disk restore job restores the data volume of the SAP HANA database. The SAP HANA database administrator then completes the database recovery and uses the transaction log to run a roll forward to restore the most recent state of the database.

The steps for creating an Instant Disk Restore are similar to the Instant Database Restore job described in “Performing an SAP HANA Instant Database Restore job”. The only difference is to choose the **Instant Disk Restore** job template (shown in Figure 4-10 on page 62), when creating the Instant Disk Restore job.

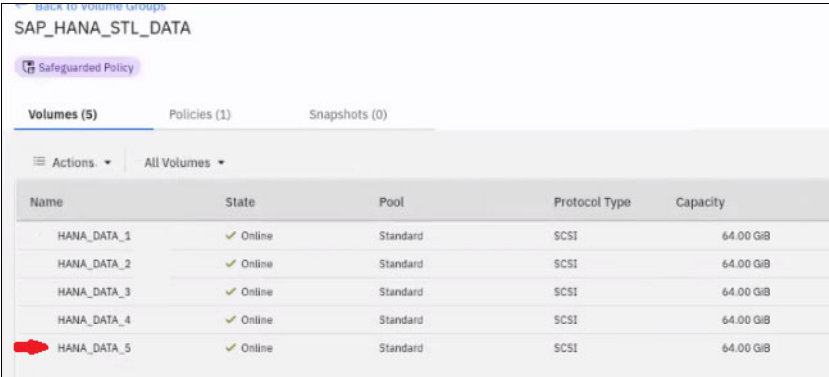
The job processing is similar to the database restore. However, the Instant Disk Restore requires a database recovery by the database administrator using, for example, SAP HANA Cockpit or SAP HANA Studio. This restore process allows the storage administrator to do a roll forward of the database to a specific point in time using the SAP HANA tools.

4.6 Daily operations, best practices and maintenance

In this section we discuss daily operations, best practices and maintenance for using IBM Storage Sentinel for SAP HANA.

4.6.1 Adding capacity to the SAP HANA data area

To increase the data area of the SAP HANA database, additional volumes are assigned to the SAP HANA data volume group. During a backup, new data volumes will be automatically discovered by IBM Storage Copy Data Management. The new volumes must reside in the same storage system and must belong to the same volume group as the existing data volumes, as shown in Figure 4-13.



SAP_HANA_STL_DATA				
Safeguarded Policy				
Volumes (5) Policies (1) Snapshots (0)				
Actions ▾ All Volumes ▾				
Name	State	Pool	Protocol Type	Capacity
HANA_DATA_1	✓ Online	Standard	SCSI	64.00 GiB
HANA_DATA_2	✓ Online	Standard	SCSI	64.00 GiB
HANA_DATA_3	✓ Online	Standard	SCSI	64.00 GiB
HANA_DATA_4	✓ Online	Standard	SCSI	64.00 GiB
HANA_DATA_5	✓ Online	Standard	SCSI	64.00 GiB

Figure 4-13 Creating a new volume for the HANA data volume group

In this example, a new volume *HANA_DATA_5* was created on the IBM FlashSystem to run in the *HANA* data volume group on the HANA server. The new volume was added online to the *HANA* data volume group *VG_STL_DATA*, as shown in Example 4-3 on page 65.

Example 4-3 Adding the block device /dev/sdn to the HANA vg VG_STL_DATA

```
pvccreate --dataalignment 1M /dev/sdn
vgextend VG_STL_DATA /dev/sdn
lvextend -i1 -l +63832 /dev/VG_STL_DATA/LV_STL_DATA
xfs_growfs /hana/data/STL
```

To detect the changed volume configuration before starting the next *HANA* backup, an SAP HANA inventory job must be executed. Now the IBM Storage Copy Data Management *HANA* backup can run and will back up all volumes of the *HANA* data volume group.

Note: It is a best practice to determine if the logical volume that should be extended is striped across multiple physical volumes *before* adding new physical volumes. To keep the LV extension striped, the number of physical volumes to add should be the same as the number of physical volumes used previously. For example, if the logical volume is striped across 4 physical volumes, another 4 physical volumes should be added. The extension can then be striped across those 4 new volumes, using the “-i 4” parameter in the **lvextend** command.

4.6.2 Combining backups not recommended

It is *not* recommended to mix Safeguarded Copy and FlashCopy copies for the same FlashSystem volumes. It can be configured by using two different SLA policies for a backup job. One policy is configured for Safeguarded Copy and the second policy for FlashCopy only. For each SAP HANA backup, one of the policies can be used.

4.6.3 Backup of the IBM Storage Copy Data Management catalog.

Because the catalog is a critical component of IBM Storage Copy Data Management, it is a best practice to regularly perform a catalog backup. To manage your IBM Storage Copy Data Management catalog login to the Administrative Console (<https://<HOSTNAME>:8090/>). In the **Menu** select **Catalog Manager and Backup Catalog**.



Scanning engine and its technology

This chapter discusses the technology and process of the malware scanning software of IBM Storage Sentinel. It describes the planning process, the different options for implementation, and other considerations.

This chapter has the following sections:

- ▶ “Storage Sentinel architecture” on page 68
- ▶ “Technology of the IBM Storage Sentinel scanning engine” on page 69
- ▶ “The advantage of anomaly scanning versus signature scanning” on page 69
- ▶ “The scanning process” on page 69
- ▶ “Scanning process for databases” on page 70
- ▶ “Machine learning” on page 70
- ▶ “Scanning encrypted data” on page 70
- ▶ “How to recognize and handle alerts” on page 71
- ▶ “Scanning Engine planning considerations” on page 73
- ▶ “Administration” on page 74

5.1 Storage Sentinel architecture

A data scan is an important part of data validation. A scan is meant to ensure that data in Safeguarded Copy volumes is not corrupted or altered by a cyberattack. A data scan can save time when it is needed most, such as in the following examples:

- ▶ You restore data that is potentially compromised into a clean room environment and start analyzing if it is safe to use. This process is done manually and requires skill and time if there is no automated and intelligent tooling.
- ▶ You restore data that is compromised and you need to restore one or more versions to find a clean one.
- ▶ Your environment is under attack, and as a precaution you restore data that you believe to be clean. However, the restored data has malware that is installed but is designed to activate at a later time to render the data inaccessible.

Figure 5-1 shows the overall idea of scanning the data. A Safeguarded Copy of production data is again copied into a recovery volume, which then is mounted to the scanning engine in a clean room environment. The scanning engine analyses the data found for anomalies or signs of malware behavior.

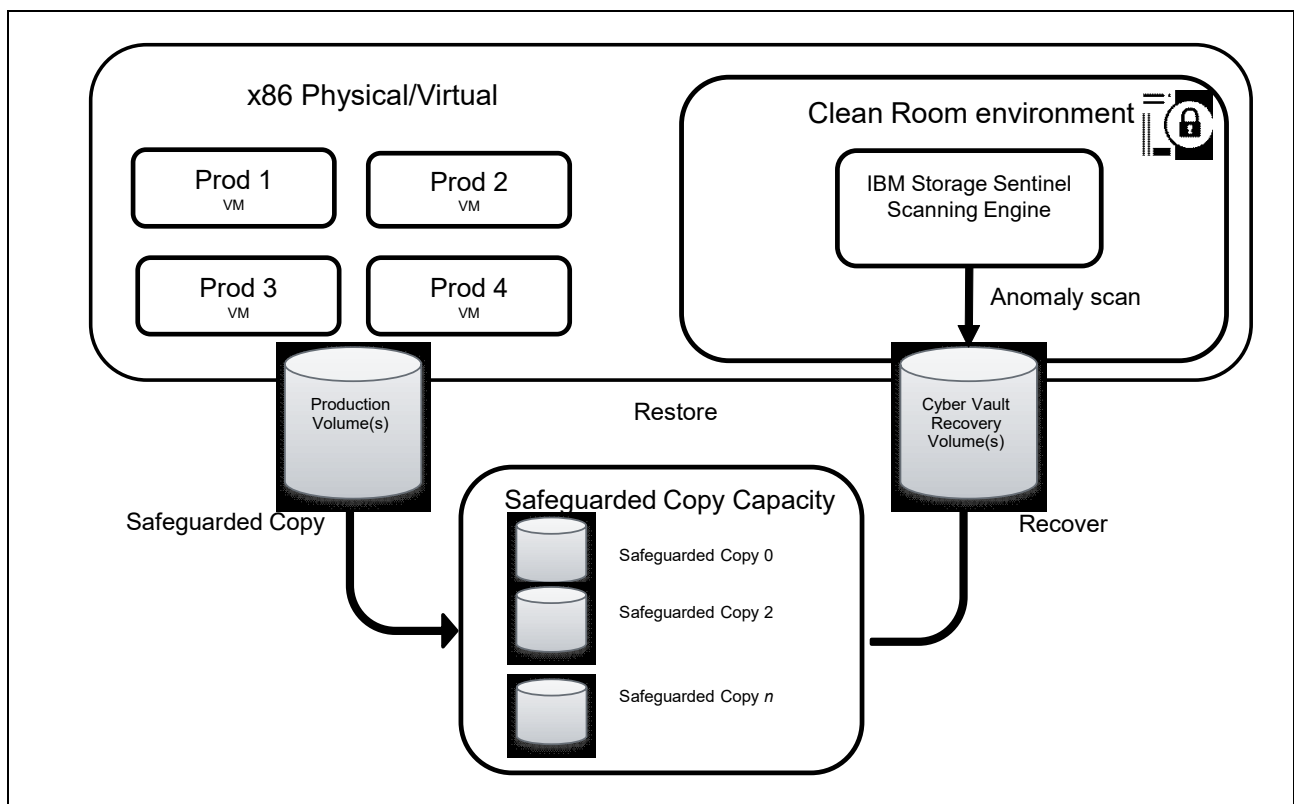


Figure 5-1 IBM Storage Sentinel Scanning workflow

5.2 Technology of the IBM Storage Sentinel scanning engine

This section includes discussion of the technology of the IBM Storage Sentinel scanning engine.

5.3 The advantage of anomaly scanning versus signature scanning

Antivirus software scans file systems for known signatures of viruses or other malware software. Because ransomware developers constantly change their code to avoid detection, the antivirus databases must be updated based on the discovery of the new variants. However, this type of scanning cannot discover zero-day exploits, which is new malicious software code that exploits previously unknown vulnerabilities. The signature of the new malware cannot yet be detected by antivirus software until the vendors update their signature databases and the antivirus software. Current advanced anomaly scanning techniques, like the scanning engine of IBM Storage Sentinel, looks for the effects of ransomware upon the data, which do not change at the same speed as the ransomware signatures itself. Encrypting data or modifying and manipulating the metadata of a file are the main effects of triggered ransomware. The anomaly scan engine in IBM Storage Sentinel is developed with the primary goal of discovering such data manipulations.

5.3.1 The scanning process

The scanning engine performs a full scan during the first time pass, and builds the metadata of what the data looks like now. Subsequent scans are incremental in comparing what the changed data looks like at that point to what it previously looked like. There are several checks that go into building the metadata and comparing the first scan versus subsequent scans.

The scanning engine uses machine learning technology to find anomalies, which are considered typical signs of malware activity. It analyzes more than 200 different data points of a file, and compares these data points with thousands of malware patterns, their effects, and previous data collected from tens of thousands of affected backups. This helps it to accurately understand if the files, or databases, are compromised by ransomware. One of those data points is entropy, which is one of the tell-tale signs of encryption. The deep level of analysis allows it to identify the most subtle attacks where bad actors are using partial encryption to accomplish two tasks. The primary goal of an attack is to avoid detection by changing a small portion of the file only, which is undetected by tools that are only basing their scans on metadata or exceeding thresholds. The second goal is to perform the attack as quickly as possible, and by only touching a portion of each file they compromise, move more quickly through a file system.

Unlike many standard antivirus software products, with the scanning process, the scanning engine builds an index containing historical data about previous scans. The comparison between current and older scan data helps it detect any suspicious changes between subsequent scans, improving the accuracy and sensitivity of the scan results.

5.3.2 Scanning process for databases

In active database applications, corruption can go unnoticed until the database is taken offline. It fails only as the database is brought online. Because the scanning engine can detect corruption in a database's snapshot or backup, whether it was from ransomware or from any other issues, even the most critical applications can be validated.

For databases, the scanning process differs from the one used for non-database files. The scanning engine looks for signs of corruption due to a ransomware attack. When it scans databases, the scanning engine first identifies the type of a file, verifying that the headers and metadata match the known format of the database file. Second, it begins examining the structure of the file, verifying that it is readable and has integrity. Lastly, it scans at the page level and verifies that the individual pages are intact and corruption free.

5.3.3 Machine learning

The Machine Learning Model of the scanning engine is configured by development. It is tuned and measured to meet a 95.5% recall rate with 0.002% false positive rate, by using more than 200 analytics that evaluate changes in the file content and metadata from one observation to the next. These analytics are fed to machine learning algorithms to make probabilistic decisions of data integrity. The training of the model consists of tens of millions of data sets including clean data, corrupted data, suspect data, and controlled detonation of live ransomware. Some data is obtained from the following sources:

- ▶ Public subscription services
- ▶ Attack simulation based on academia and global research data
- ▶ Anonymized customer data sets

After it is installed at a customer site, the machine learning engine is updated periodically, typically 3–4 times per year. Because it is not signature-based but looks for the effects of malware to data, it does not require weekly or daily updates to operate successfully.

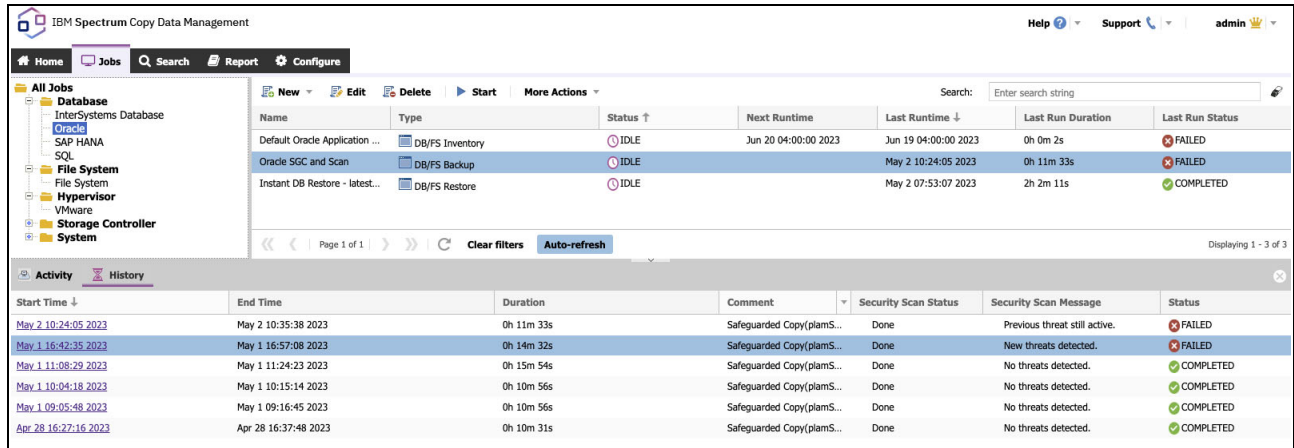
5.3.4 Scanning encrypted data

Whether encrypted data can be scanned depends on where the encryption occurred:

- ▶ Data that is encrypted at the volume level cannot be read and analyzed by the scanning engine, but fails to mount successfully and triggers an appropriate alert.
- ▶ If data is encrypted at file level, then changes to file name, file type, and file suffix can be discovered and analyzed if the change is suspected to be a malware effect. This type of corruption is more typical of a generic malware attack on a file system and will render a DB inaccessible, which will trigger an alert.
- ▶ If a database is encrypted for security purposes, the scanning engine cannot read the data from each page but can still see the structure and understands the concept of user-based application encryption. It can discover if the file contains corruption due to ransomware.

5.4 How to recognize and handle alerts

The first sign of an alert can be seen in the IBM Storage Copy Data Management GUI. In the job list, you see a failed scanning process with the message “*Threat detected*”. An example is shown in Figure 5-2.



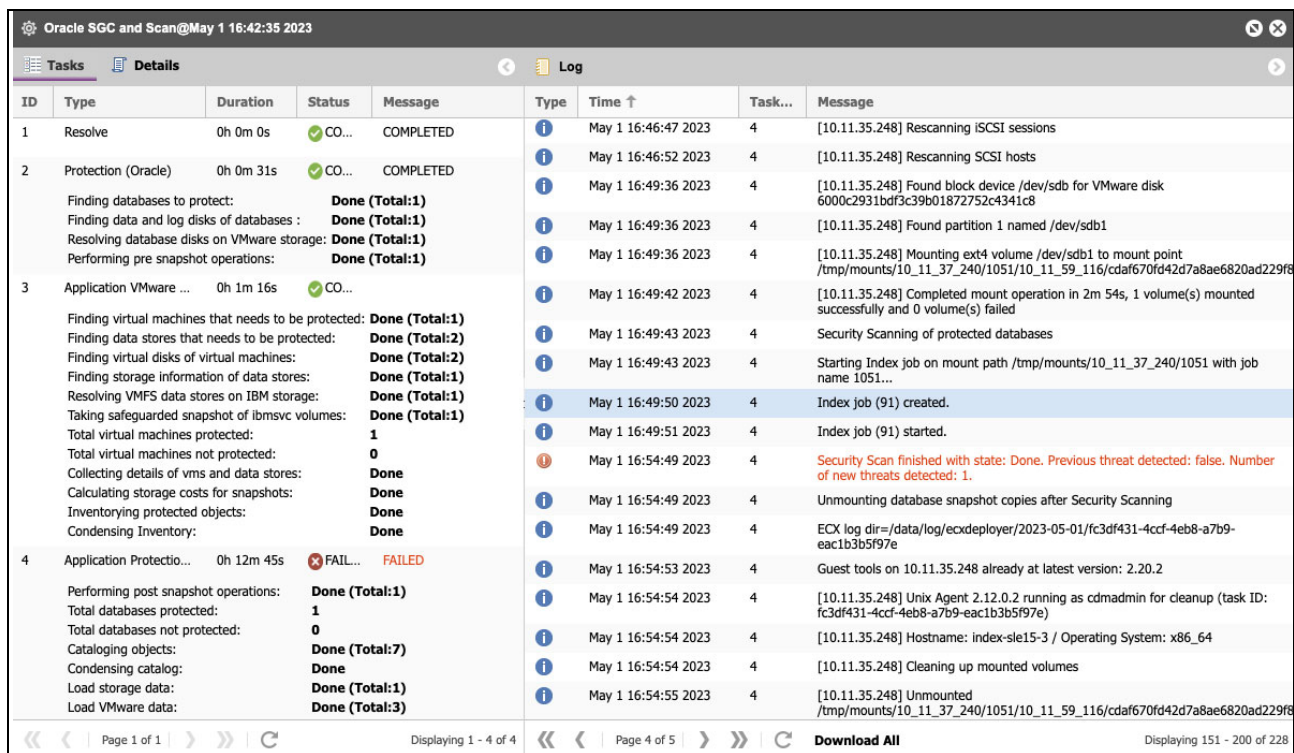
The screenshot shows the IBM Spectrum Copy Data Management GUI. The 'Jobs' tab is active, displaying a list of jobs. The job 'Oracle SGC and Scan' is highlighted, showing a status of 'FAILED' with the message 'Threat detected'. The 'Activity' tab is also visible, showing a detailed log of the job's execution, including the detection of a threat.

Name	Type	Status	Next Runtime	Last Runtime	Last Run Duration	Last Run Status
Default Oracle Application ...	DB/FS Inventory	IDLE	Jun 20 04:00:00 2023	Jun 19 04:00:00 2023	0h 0m 2s	FAILED
Oracle SGC and Scan	DB/FS Backup	IDLE	May 2 10:24:05 2023	May 2 10:24:05 2023	0h 11m 33s	FAILED
Instant DB Restore - latest...	DB/FS Restore	IDLE	May 2 07:53:07 2023	May 2 07:53:07 2023	2h 2m 11s	COMPLETED

Start Time	End Time	Duration	Comment	Security Scan Status	Security Scan Message	Status
May 2 10:24:05 2023	May 2 10:35:38 2023	0h 11m 33s	Safeguarded Copy(plamS...	Done	Previous threat still active.	FAILED
May 1 16:42:35 2023	May 1 16:57:08 2023	0h 14m 32s	Safeguarded Copy(plamS...	Done	New threats detected.	FAILED
May 1 11:08:29 2023	May 1 11:24:23 2023	0h 15m 54s	Safeguarded Copy(plamS...	Done	No threats detected.	COMPLETED
May 1 10:04:18 2023	May 1 10:15:14 2023	0h 10m 56s	Safeguarded Copy(plamS...	Done	No threats detected.	COMPLETED
May 1 09:05:48 2023	May 1 09:16:45 2023	0h 10m 56s	Safeguarded Copy(plamS...	Done	No threats detected.	COMPLETED
Apr 28 16:27:16 2023	Apr 28 16:37:48 2023	0h 10m 31s	Safeguarded Copy(plamS...	Done	No threats detected.	COMPLETED

Figure 5-2 Threat detected message

If you analyze the job log, you see exactly when and on which volume the threat was detected, as shown in Figure 5-3.



The screenshot shows the 'Log' tab for the job 'Oracle SGC and Scan' on May 1 16:42:35 2023. The log details the execution of the job, including the detection of a threat. The log is divided into two sections: 'Tasks' and 'Log'. The 'Tasks' section shows the overall progress of the job, and the 'Log' section shows the detailed execution steps and messages.

ID	Type	Duration	Status	Message	Type	Time	Task...	Message
1	Resolve	0h 0m 0s	CO...	COMPLETED	i	May 1 16:46:47 2023	4	[10.11.35.248] Rescanning iSCSI sessions
2	Protection (Oracle)	0h 0m 31s	CO...	COMPLETED	i	May 1 16:46:52 2023	4	[10.11.35.248] Rescanning SCSI hosts
	Finding databases to protect:		Done (Total:1)		i	May 1 16:49:36 2023	4	[10.11.35.248] Found block device /dev/sdb for VMware disk 6000c2931bdf3c39b01872752c4341c8
	Finding data and log disks of databases :		Done (Total:1)		i	May 1 16:49:36 2023	4	[10.11.35.248] Found partition 1 named /dev/sdb1
	Resolving database disks on VMware storage:		Done (Total:1)		i	May 1 16:49:36 2023	4	[10.11.35.248] Mounting ext4 volume /dev/sdb1 to mount point /tmp/mounts/10_11_37_240/1051/10_11_59_116/cdaf670fd42d7a8ae6820ad229f8
	Performing pre snapshot operations:		Done (Total:1)		i	May 1 16:49:36 2023	4	[10.11.35.248] Mounting ext4 volume /dev/sdb1 to mount point /tmp/mounts/10_11_37_240/1051/10_11_59_116/cdaf670fd42d7a8ae6820ad229f8
3	Application VMware ...	0h 1m 16s	CO...		i	May 1 16:49:42 2023	4	[10.11.35.248] Completed mount operation in 2m 54s, 1 volume(s) mounted successfully and 0 volume(s) failed
	Finding virtual machines that needs to be protected:		Done (Total:1)		i	May 1 16:49:43 2023	4	Security Scanning of protected databases
	Finding data stores that needs to be protected:		Done (Total:2)		i	May 1 16:49:43 2023	4	Starting Index job on mount path /tmp/mounts/10_11_37_240/1051 with job name 1051...
	Finding virtual disks of virtual machines:		Done (Total:1)		i	May 1 16:49:50 2023	4	Index job (91) created.
	Finding storage information of data stores:		Done (Total:1)		i	May 1 16:49:51 2023	4	Index job (91) started.
	Resolving VMFS data stores on IBM storage:		Done (Total:1)		i	May 1 16:54:49 2023	4	Security Scan finished with state: Done. Previous threat detected: false. Number of new threats detected: 1.
	Taking safeguarded snapshot of ibmsvc volumes:		Done (Total:1)		i	May 1 16:54:49 2023	4	Unmounting database snapshot copies after Security Scanning
	Total virtual machines protected:		1		i	May 1 16:54:49 2023	4	ECX log dir=/data/log/ecxdeployer/2023-05-01/fc3df431-4ccf-4eb8-a7b9-eac1b3b5f97e
	Total virtual machines not protected:		0		i	May 1 16:54:53 2023	4	Guest tools on 10.11.35.248 already at latest version: 2.20.2
	Collecting details of vms and data stores:		Done		i	May 1 16:54:54 2023	4	[10.11.35.248] Unix Agent 2.12.0.2 running as cdadmin for cleanup (task ID: fc3df431-4ccf-4eb8-a7b9-eac1b3b5f97e)
	Calculating storage costs for snapshots:		Done		i	May 1 16:54:54 2023	4	[10.11.35.248] Hostname: index-sle15-3 / Operating System: x86_64
	Inventorying protected objects:		Done		i	May 1 16:54:54 2023	4	[10.11.35.248] Cleaning up mounted volumes
	Condensing Inventory:		Done		i	May 1 16:54:55 2023	4	[10.11.35.248] Unmounted /tmp/mounts/10_11_37_240/1051/10_11_59_116/cdaf670fd42d7a8ae6820ad229f8
4	Application Protectio...	0h 12m 45s	FAIL...	FAILED				
	Performing post snapshot operations:		Done (Total:1)					
	Total databases protected:		1					
	Total databases not protected:		0					
	Cataloging objects:		Done (Total:7)					
	Condensing catalog:		Done					
	Load storage data:		Done (Total:1)					
	Load VMware data:		Done (Total:3)					

Figure 5-3 Job log showing details on the detected corruption

The IBM Storage Sentinel dashboard contains more details about the nature of the detected threat. See Figure 5-4.

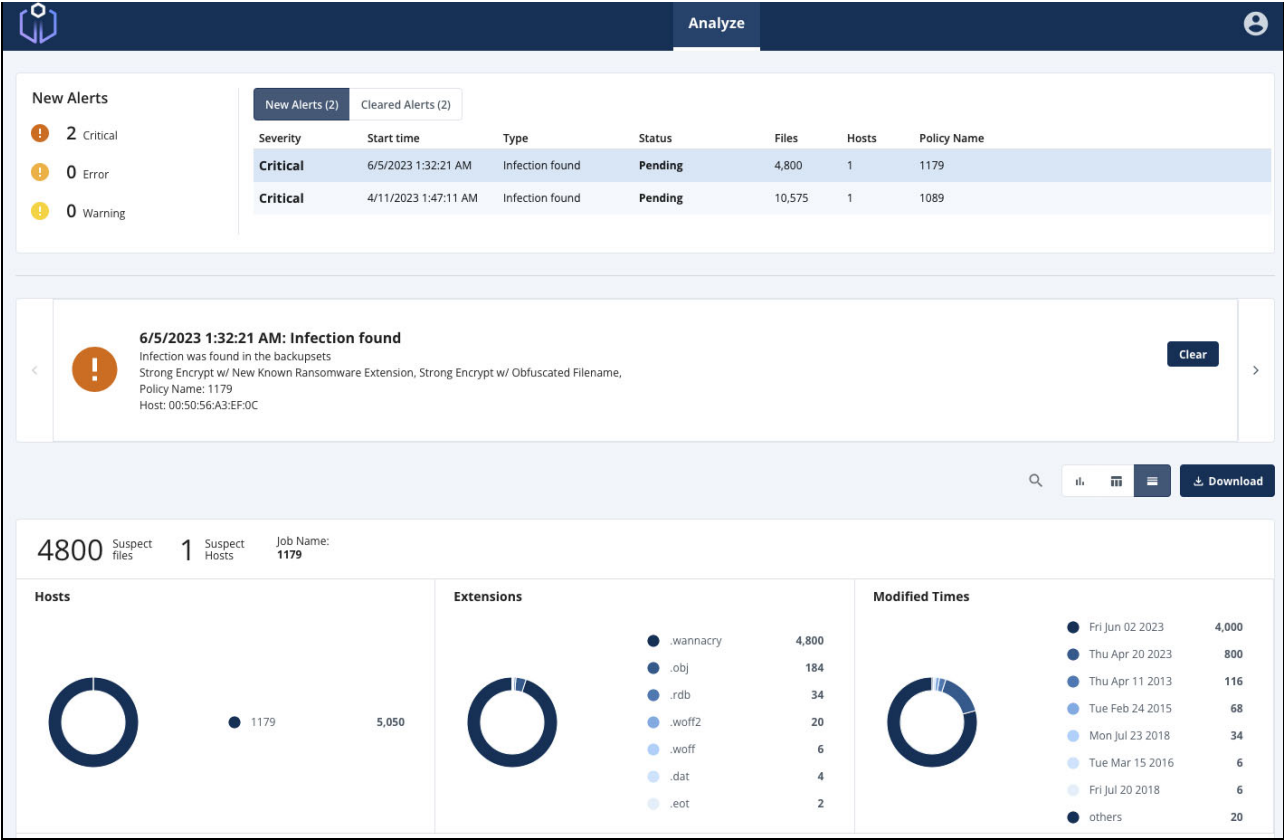


Figure 5-4 IBM Storage Sentinel Scanning Engine dashboard

5.4.1 After alert workflow

IBM Storage Copy Data Management raises an alert if a job fails due to Storage Sentinel detecting corruption. Also, the scanning engine itself can be configured to trigger email notifications and SYSLOG output. The latter can be scripted into any solution that can query logs, such as the typical SIEM and SOAR tools that are available.

5.4.2 What to do when the scanning engine finds an issue

If a backup is flagged for potential corruption due to ransomware, contact IBM support immediately. IBM support teams are trained to help customers and will involve further levels of support up to the development labs as needed. It is critical that a failed scan be analyzed by specialists to verify if an actual corruption is correctly detected, or if the failed scan is a rare false positive.

5.4.3 Dealing with false positives

When an anomaly found by the scanning engine is analyzed but is found to be an intended data change, it is called a false positive. In such a case, the administrator should make a note of what files were involved. If the false positive is caused by a file type that the scanning engine server does not recognize, the information should be submitted to IBM support describing the file type for inclusion in future builds.

5.5 Scanning Engine planning considerations

You can regard the scanning part of the solution mostly as a black box, but understanding some of the concepts is important.

5.5.1 Sizing considerations

For sizing recommendation of the scanning engine, see [Server requirements and recommendations](#).

5.5.2 Scaling of scan workloads

The history of scans is one of the things that makes the Storage Sentinel scanning so powerful. However, when you change the scanner that is used for a specific instance of your application, the history is not transferred, and the new scan engine creates a new history. If you are running a single server, the best practice is to stagger the workloads throughout the window of processing. If you are running multiple servers, balance the scanning jobs across the available scanning engines for optimum usage of the processing window. Moving a scanning job from one server to another requires a new initial scan, and you lose the scan history from the previous scans. Each scanning engine can run multiple jobs simultaneously. The scanning application is multi-threaded to improve performance.

The implementation of Storage Sentinel creates one job per snapshot. Multiple snapshots run as multiple jobs. If you have more jobs than a single server can handle, these should be split across two or more scan servers.

Storage Sentinel has a federated licensing scheme, so licenses of all scanning engines can be managed from a central instance. For more information, see [IBM Storage Sentinel anomaly scan software 1.1.4](#).

5.5.3 Virtual versus physical servers

If the scanning engine is installed on a physical server, it can scan volumes that are used by physical application servers only. If the scanning engine is installed on a virtual machine, it is able to scan volumes used by both physical or virtual application servers.

When scanning virtual application servers, there are times that you might want to limit the number of volumes that require a Safeguarded Copy. For example, if the application uses volumes that are virtual disks on a VMFS datastore, you can create dedicated datastores for the VMs that contain an instance of the application. More commonly, customers can place the protected applications on dedicated disks that are presented as physical raw device mapped (pRDM) volumes or by using iSCSI. At the time of writing, Storage Sentinel does not currently support vSphere vVols.

Storage Sentinel is deployed within a machine that is running x64 SUSE Linux Enterprise Server 15.

Note: At the time of writing, Index engines scanning software only supports SUSE Linux Enterprise Server 15.

This means that the file systems on volumes on application servers running x64 Linux can be directly mapped and mounted to the scanning engine. However, Storage Sentinel also supports protected applications running on an IBM Power server and AIX operating system, which cannot be mounted directly to the Storage Sentinel scanning engine. Copy Data Management requires an AIX proxy machine running on the Power platform, so that Copy Data Management can mount the volumes to be scanned to the proxy, and share the file systems over NFS for the scanning engine to access. See Figure 5-5.

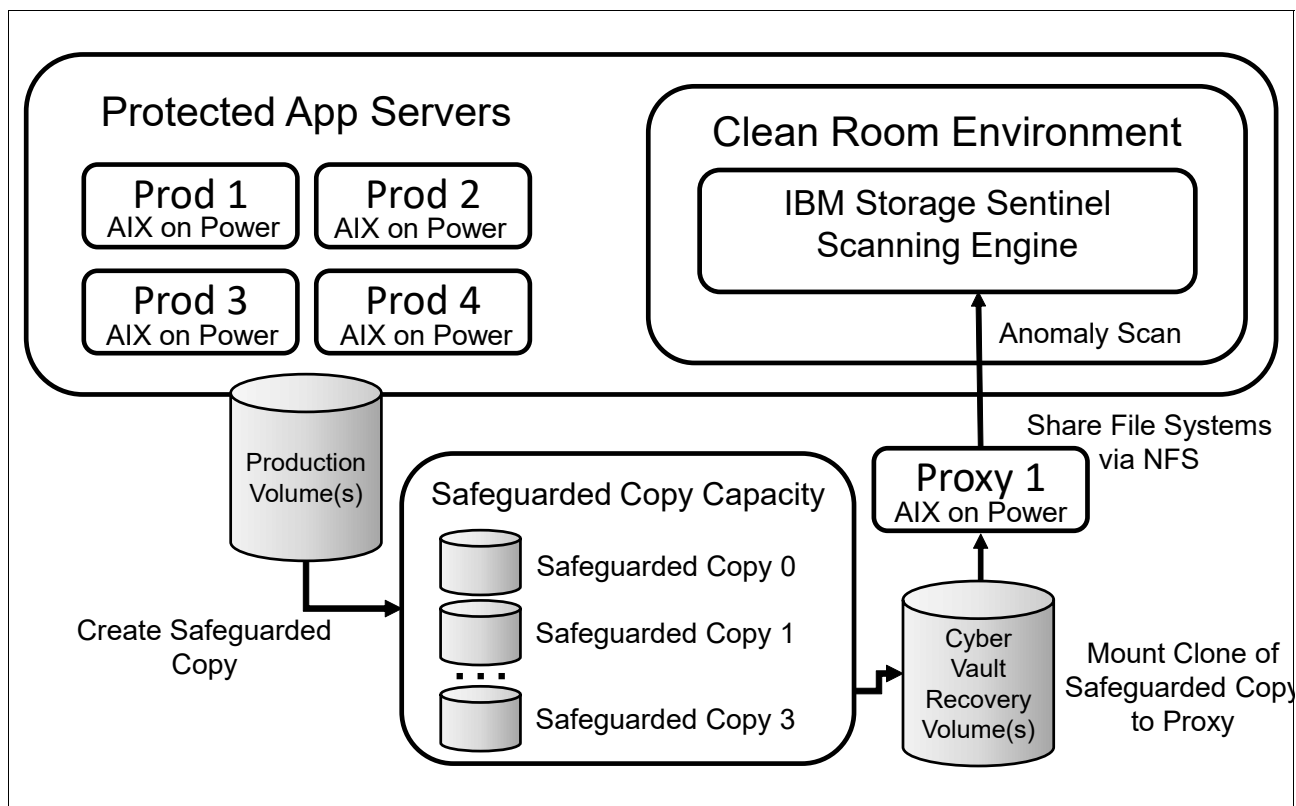


Figure 5-5 IBM Storage Sentinel configuration with an AIX proxy machine on the Power platform

5.6 Administration

This section includes discussion of the administration of the scanning engine.

5.6.1 Monitoring the scanning engine

The basic day to day monitoring task is checking the index size and the amount of front-end data that the system scans to make sure that the user is not exceeding the licensed amount. In addition, back up the configuration regularly.

5.6.2 Backing up and restoring the scanning engine components

Storage Sentinel includes utilities to allow backup software to take clean backups of the key application files. If you are running a federated Storage Sentinel environment, back up the member instances first, then the manager. At recovery time, you reverse the order and recover the manager first, then the members.

The default backup location of Storage Sentinel is `/opt/ie/backup`. However, you can specify a different location by modifying `/opt/ie/var/backup` location. The backup directory must reside in the same file system as `/opt/ie/var`.

For IBM Storage Protect (previously called Tivoli Storage Manager), follow these steps:

1. Add the following lines to `dsm.sys`:

```
PRESCHEDULECMD /opt/ie/bin/tsm_presched  
POSTSCHEDULECMD /opt/ie/bin/tsm_postsched
```

Note: The files `/opt/ie/bin/tsm_presched` and `/opt/ie/bin/tsm_postsched` already exist on the Index Engines system.

2. On the Storage Protect server, specify the copy group for the policy domain, policy set, and management class that the client belongs to and set Copy Serialization to Dynamic.
3. Define a scheduled job to back up the client directory `/opt/ie/backup` (or the alternate directory if you created one). Either an Incremental or Selective backup works. There are multiple ways to specify what you want to back up with Storage Protect. If you want to have a dedicated backup schedule for protecting the Storage Sentinel backup location, you can specify that location in the `objects` parameter in the schedule definition. You need to include any subdirectories, so specify **Subdir Yes** in the `dsm.sys` stanza or include it in the schedule settings in the `Options` parameter. For example:

```
UPDATE SCHEDULE examplePD BackupSentinel type=client action=incremental  
objects='/opt/ie/backup/*' options=-subdir=yes startdate=TODAY starttime=NOW  
dayofweek=any
```

If you are running a federated environment, be sure to back up the members before you back up the manager, so you need at least 2 schedules.

The Storage Sentinel interface is used to recover data from the current contents of the backup directory (Select **Administration** → **System** → **Recovery** → **Recover From Backup**).

If this is a pristine environment, you need to install Storage Sentinel, restore the backup directory from the backup software, and then use the Storage Sentinel GUI to rebuild the instance from the backup files. If you are running a federated Storage Sentinel environment, restore the manager first, then the members.

5.6.3 Adding new applications

Register new applications and define their scanning process in IBM Storage Copy Data Management. As you add workloads to be scanned, it is important to monitor your Storage Sentinel license information so that you do not exceed your licensed capacity. The number of protected applications or the number of scan servers is not monitored or restricted by the license.

5.6.4 Adding new scanning engines

Added scanning engines must be registered in the GUI of IBM Storage Copy Data Management. The scanning process for multiple applications should be spread across the scan servers for load balancing.



IBM Cyber Vault setup: Putting it all together

This chapter covers the IBM Cyber Vault architecture, a framework developed by IBM for highly automated protection of critical data. It discusses the planning process of IBM Cyber Vault, different options for implementation and additional considerations.

This chapter has the following sections:

- ▶ “Introduction to IBM Cyber Vault” on page 78
- ▶ “IBM Cyber Vault planning considerations” on page 80

6.1 Introduction to IBM Cyber Vault

The goal of the IBM Cyber Vault architecture is to establish a framework for automated protection of critical data. This helps ensure a faster business recovery after an incident. Immutable snapshots enhance the protection, especially after a successful cyberattack.

Many cyberattacks corrupt, encrypt, or even wipe application data, which usually causes affected applications to stop working. This architecture defines appropriate data protection measures along with environment monitoring, data validation, and recovery planning. The goal is to always have an uncorrupted copy of critical data, with the following characteristics:

- ▶ Cannot be compromised by a successful cyberattack
- ▶ Is regularly tested for being free of any signs of corruption or malware
- ▶ Can be recovered quickly and safely
- ▶ Can be used in a timely recovery of critical business services after a cyberattack

The IBM Cyber Vault architecture blueprint provides a detailed description of the overall concept. It can be downloaded at [IBM FlashSystem Cyber Vault](#).

In this context, it is important that the set of business-critical services and applications are identified, which must be protected, and which are first priority for recovery. These can be described as the *Minimum Viable Company (MVC)*. The application data that is required to recover the MVC are referred to as *primary workloads*; less critical data - that still are needed to recover the full company functionality - are the *secondary workloads*.

Recommendations for improving resiliency against cyberattacks, primary and secondary workloads are shown in Table 6-1. Note that the recommendations do not replace standard high availability and disaster recovery precautions. e

Table 6-1 Recommendations for workload protection¹

Property (feature / attribute) recommended for resilience against cyberattacks	Primary Workloads (Minimum viable company recovery)	Secondary Workloads (Full company workload recovery)
Data retention	short term (days)	long term (days - weeks)
Data copies	immutable (on-array)	immutable if feasible, off-array
Recovery time target	minutes to hours	hours to days
Air gap	logical	logical or physical

¹ Ian Shave, Roger Kasten: *The business Impacts of Cyber Attacks*, 17 January 2023, p. 7

Figure 6-1 shows an example of a complete architecture, which is built upon the IBM Cyber Vault blueprint, and includes data validation for primary and secondary workloads.

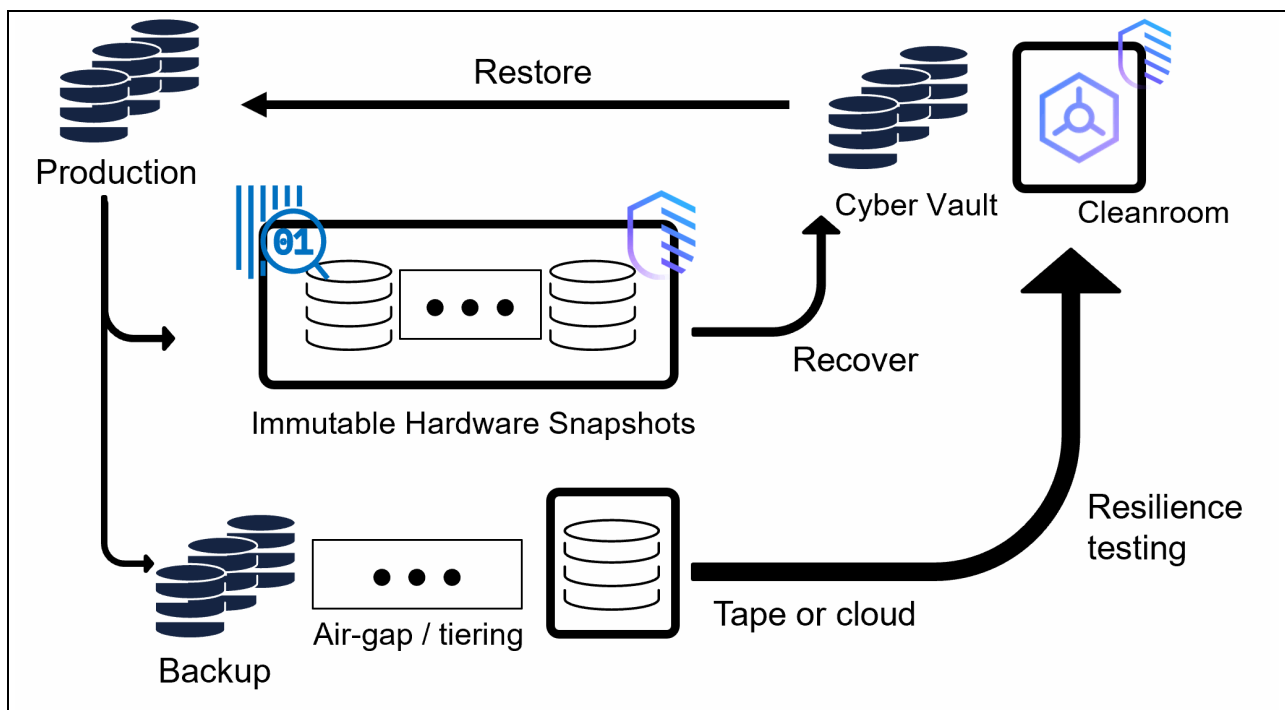


Figure 6-1 IBM Cyber Vault architecture example

6.1.1 The four steps to IBM Cyber Vault

To establish the IBM Cyber vault architecture, the following four steps can be followed, as shown in Figure 6-2 on page 80:

1. Protect

Protect critical data from cyberthreats by periodically creating immutable data copies. These data copies can be established as crash consistent or as application consistent copies. Application consistent copies are more complex to establish because, for example, a database application needs to be quiesced during snapshot creation. However, a snapshot of the database enables a faster restart without the need to do a manual database recovery.

2. Proactive monitoring

The IBM solution is QRadar, but the architecture also allows integration with other Security information and event management (SIEM) tools, such as Splunk. For more information, see [IBM Security QRadar Suite](#).

3. Data validation

Data validation depends on the applications that need to be protected. The CV architecture blueprint mentions some examples. Alternatively, or in addition, the client can define their own testing method. The testing method can be as simple as mounting a recovery volume to a VM with their application image and verifying it works as expected.

Recovery

The recovery processes are often application specific and well-defined, for example, target system for recovery, recovery process, level of automation, and so forth.

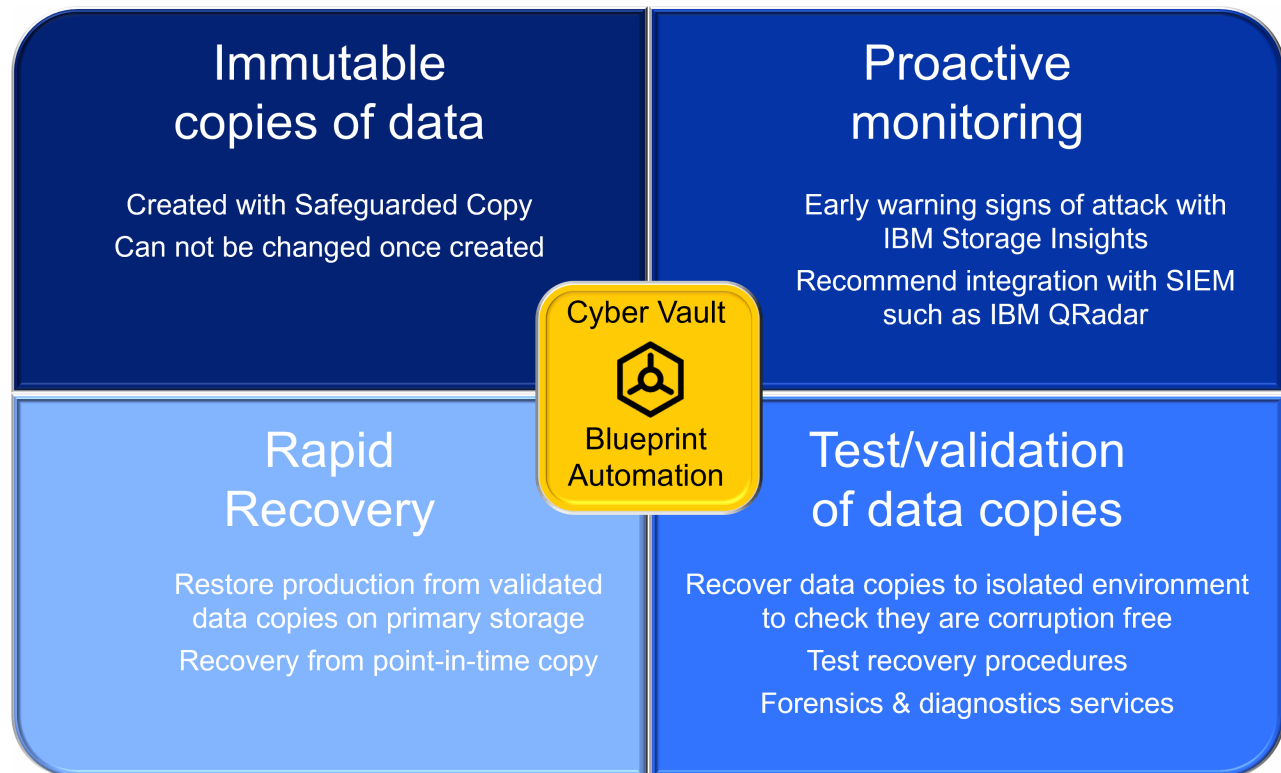


Figure 6-2 The four steps to IBM Cyber Vault

6.2 IBM Cyber Vault planning considerations

This section describes the required considerations for setting up a complete IBM Cyber Vault infrastructure to achieve an appropriate level of cyber resiliency while considering the business criticality of applications and data. After the parameters described in this section, such as the Recovery Time Objective (RTO) and Recovery Point Objective (RPO) for the most critical business services, have been defined, the infrastructure components can be planned and sized for the expected capacities and workloads.

6.2.1 Definition of the Minimum Viable Company (MVC)

Before infrastructure planning takes place, the critical set of business services and applications must be defined, which will be protected by the IBM Cyber Vault. These should include all services and applications that are required to restart the business in an emergency mode. It allows the business to survive until the full set of services can be recovered. Statistics show that without appropriate precautions for such attacks, a partial business recovery can take a few weeks or months. The goal is to shorten the recovery time by identifying the most critical services, and prioritizing their recovery.

6.2.2 Establishing immutable copies of critical data

Because cyberattacks often aim to corrupt, encrypt, or even completely wipe critical data, it is essential to keep data copies that are immune to these threats. Standard backups are not safe, as the backup infrastructure is also a target for cybercriminals. Immutable copies, however, are immune to these threats because they cannot be altered or deleted. They are the most efficient way to achieve data resiliency and can be stored within a storage array like IBM FlashSystem, or in an air-gapped medium such as tape. The frequency of taking immutable copies depends on the RPO (see also 6.2.3, “Crash consistency or application consistency?” on page 81). Ensure that the retention time of the copies are long enough to cover the duration of the process from the detection of an incident to the decision for recovery. Another consideration about retention time is the decreasing value of older data copies. If the business will not survive a longer outage anyway, there is no point to retain data longer than the assumed survival limit. Does it still make sense to recover data older than, for example, one week? Or would this amount of data loss (or such a long outage period) jeopardize the business as a whole? For even longer retention times, an in-array immutable copy can be migrated to an air-gapped copy on lower-cost media. Figure 6-3 shows typical backup frequencies and retention times that clients use.

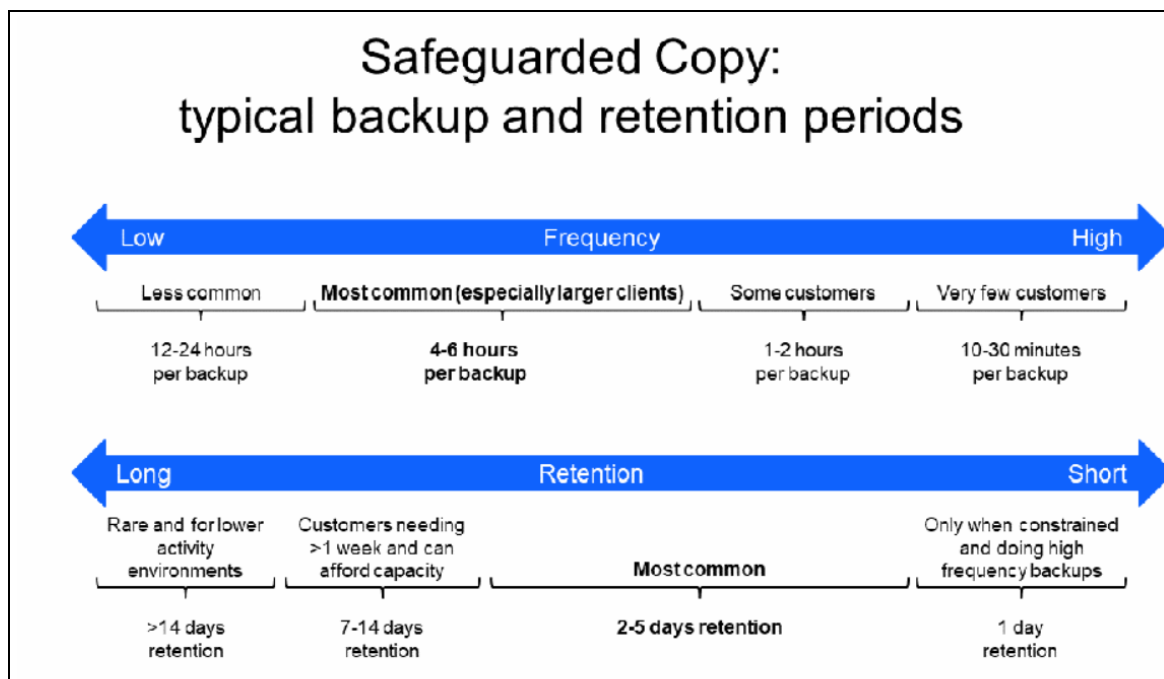


Figure 6-3 Safeguarded Copy: Typical backup and retention periods

6.2.3 Crash consistency or application consistency?

Just like any snapshot-based data copy, Safeguarded copies can be created in either crash-consistent or application-consistent mode.

Crash-consistent backups can be easier to create because the orchestration does not require knowledge about the application. However, the recovery process might take longer due to potential inconsistencies within the backup data. After a crash-consistent backup is restored, these inconsistencies need to be repaired on the application level before the service can be restarted.

Application-consistent backups are more complex to create because the file system needs to be in a consistent state before the snapshot is taken. Many applications use their own methods to establish data consistency, such as setting themselves in backup mode or quiescing the application. One example of a method is pausing all I/O processes and flushing the file system buffers so that there are no I/O operations left pending to the file system. After the snapshot is taken, the application can be resumed to normal operation. Because the snapshot now contains a consistent data set as seen from the application perspective, the recovery process is often quicker.

6.2.4 Proactive monitoring

As with other parts of the infrastructure, monitor the Cyber Vault infrastructure for any security-related incident. For integration with a *Security Information and Event Management (SIEM)* tool, usually message logs are forwarded to a SYSLOG server. Advanced SIEM tools, such as IBM QRadar®, correlate events from multiple sources and understand the relation between these which can help detect suspicious activities early to prevent the actual attack.

A specific monitoring use case that should be implemented for IBM Cyber Vault is to verify that the immutable copies are successfully created, according to the defined schedule. SIEM tools have the capability to check if, for example, a message sequence proving that the success of this process occurs regularly, and raise an incident if these messages are not logged after a specified time interval.

Some SIEM tools are able to run actions toward the Cyber Vault. One example is if an early warning sign of a cyberbreach is detected, then an immediate immutable copy is triggered. If this copy is not yet affected by the attack, it can help reduce the RPO. If the copy is corrupted, it can serve as a source for forensic analysis because it is very close to the time of the cyberattack.

6.2.5 RPO, RTO, and data validation

For the application data in the scope of the MVC, appropriate RPO and RTO targets must be defined. These targets can be derived from existing RPO and RTO targets for disaster recovery, but should be reviewed while focusing on the possibility of a cyberattack. The RPO target determines the frequency of immutable copies taken, and the RTO corresponds to the application recovery process.

In addition, it is important to assess when and how often a validation process of the data copy is initiated. IBM recommends running data validation as a periodic process that includes all copies that are taken. This ensures that the last good data copy is always known so that the recovery process can be started immediately after an incident. This strategy provides the optimum protection level along with the shortest expected recovery time.

Depending on the amount of data, it might not be feasible to validate every data copy. For example, if the validation process takes longer than the time between two data copies, then it might not be possible to validate all of the data copies. After an incident, some decisions must be made about which data copy to use for recovery. The decision involves a tradeoff between quicker recovery and minimizing the amount of data loss:

- ▶ Use the last validated copy or a more recent unvalidated copy
- ▶ Spend additional time to validate the most current data copy or attempt a quicker recovery

To minimize the time that is spent on data validation for multiple applications or volume groups, you can use more resources, such as memory or processing power, for the scanning engine, or you can use multiple scanning engines.

6.2.6 Recovery planning

The recovery of business services after a cyberattack follows similar principles like Disaster Recovery processes. The key difference is that, before a service can be recovered, the following factors are verified:

- ▶ The target infrastructure for recovery is not breached by an intruder or infected with malware. Depending on the expected RTO, it might be required to configure an available specific infrastructure setup that can be activated upon short notice.
- ▶ The source data is verified to be unaltered by the attack, and free of any signs of malware or other types of infection.

If IBM Storage Sentinel is used, the scan results of the available data copies can be examined, and the most recent, validated copy can be selected for recovery. For other data validation strategies, an appropriate selection process must be established.

For quick recovery, automate the processes. Automation helps to avoid human errors, but it also needs to be regularly tested to uncover infrastructure or software changes that might cause the automation process to fail. Also, include a human decision checkpoint, such as when you select the best data copy for recovery.

6.2.7 Further considerations

The following section discusses further considerations for Cyber Vault implementation.

Secondary workloads and tape technology

Treat data from secondary workloads similarly to primary data, while also considering the differences described in Table 6-1 on page 78. The full picture of an IBM Cyber Vault architecture including primary and secondary data is shown in Figure 6-1 on page 79.

Physical tape can be considered for storing copies of secondary data. It offers a physical air gap, has options for immutability, and is the lowest-cost storage medium available. It can also be used to store copies of primary data for longer retention times if required. However, recovery times can be longer compared to a snapshot-based recovery.

Tape is considered a secure storage medium because data on tape can be deleted only by overwriting tape cartridges. This process typically takes much longer than deleting data from a flash-based storage system. It also requires tape drives to be used. Because this activity likely deviates from normal tape workload behavior, it can be detected by monitoring tape drive and cartridge usage with a SIEM tool.

Preserving suspect data for law enforcement or other organizations

For some instances, an organization needs to preserve data that it believes is corrupted because of a cyberattack. This preserved data can be used for analysis of the attack or as evidence in a legal prosecution. The legal or contractual requirements for preserving this data vary by the location, the type of organization, and the contracts that might exist between the organization and its partners. An organization's policies on this matter should be established by its legal counsel and information security offices. It is especially important to confer with law enforcement and comply with rules that govern the creation and management of this type of evidence, such as chain of custody and securing the data against unwanted access.

It is often difficult to know exactly what data needs to be preserved after a cyberattack, but creating a clone of the suspect volumes with no expiration date is a good first step. An organization might want to create snapshots of the OS and application volumes for the server,

and any SIEM/SOAR logs that encompass the time frame of the attack. If the attacked host is being monitored, then retaining the reports for that same time frame can also be valuable.

In general, save anything that might be of value before the data expires. If the data is not needed, then it can be deleted when it is convenient.

It is recommended that an organization establish policies on what data is to be retained and in what format. To meet that policy, define a procedure that requires minimal human intervention and is, ideally, an automated procedure. Trying to figure this out after the fact is going to take valuable effort that is best spent recovering data and returning the key applications to service.



Supported patterns

This chapter describes supported patterns of IBM Storage Safeguarded Copy feature on Storage Virtualize based storage.

Safeguarded Copy can be deployed in a range of different topologies:

- ▶ “Safeguarded Copy on a single system” on page 86
- ▶ “Safeguarded Copy in a Metro Mirror or Global Mirror relationship” on page 86
- ▶ “Safeguarded Copy in an IBM HyperSwap environment” on page 88

Note: Configure and run the Storage Sentinel anomaly scan server at the same location as the safeguarded copies.

7.1 Safeguarded Copy on a single system

In single system environments, there are three main components:

1. The production or source volume, which refers to the data that is being copied
2. The safeguarded backup capacity, also known as the child pool, which is the storage location for the Safeguarded Copies and is inaccessible to the host
3. The recovery volume, which is where the safeguarded copies are restored for access

Safeguarded Copy enables a user to create up to 32,000 immutable copies of a source or production environment. These copies are immutable and cannot be accessed directly by any server.

To access the copies, immutable copies must be copied to a set of recovery volumes. This provides instant access to the data and can be accessed by a recovery system for various purposes, including data validation, forensic analysis, or to restore the data back to the production environment.

Safeguarded Copy is not a direct replacement for FlashCopy and both can be used as part of a cyber resilience solution. See Figure 7-1 on page 86.

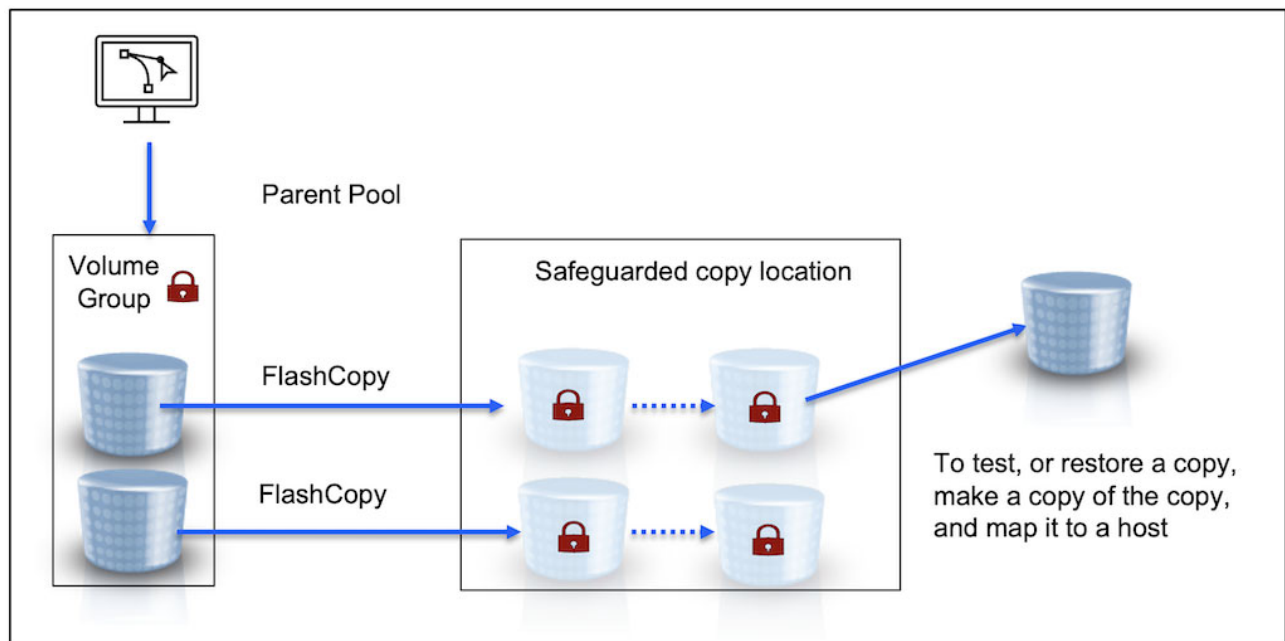


Figure 7-1 Safeguarded Copy on a single system

7.2 Safeguarded Copy in a Metro Mirror or Global Mirror relationship

In a Metro Mirror (MM), Global Mirror (GM), or Global Mirror with change volumes (GMCV) relationship, you can decide to create Safeguarded Copies of one site or both sites. The capacity limitations, recovery times, and recovery process must be considered, implemented, and tested.

Note: Safeguarded Copies are taken from the primary and the secondary but you cannot create Remote Copy relationships on Safeguarded Copy point-in-time copies.

Safeguarded Copy uses FlashCopy, so it has a system or cluster boundary for taking and restoring Safeguarded Copies. If the copies are only taken at the secondary site, you cannot use FlashCopy to flash back the Safeguarded Copies to the source volume because it is in a different cluster. You must replicate the data back to the primary site.

You can create safeguarded copies at the secondary site without pausing or suspending MM, GM, or GMCV. A crash-consistent point-in-time copy is taken. Some clients may prefer or require an application-consistent copy. To do this, they often use a script or cron job to pause the database, use the Copy Services Manager command line (CSMCLI) to take a Safeguarded Copy, and then resume database read and write activity. They might also use an external tool such as Storage CDM with CSM. See Figure 7-2 on page 87.

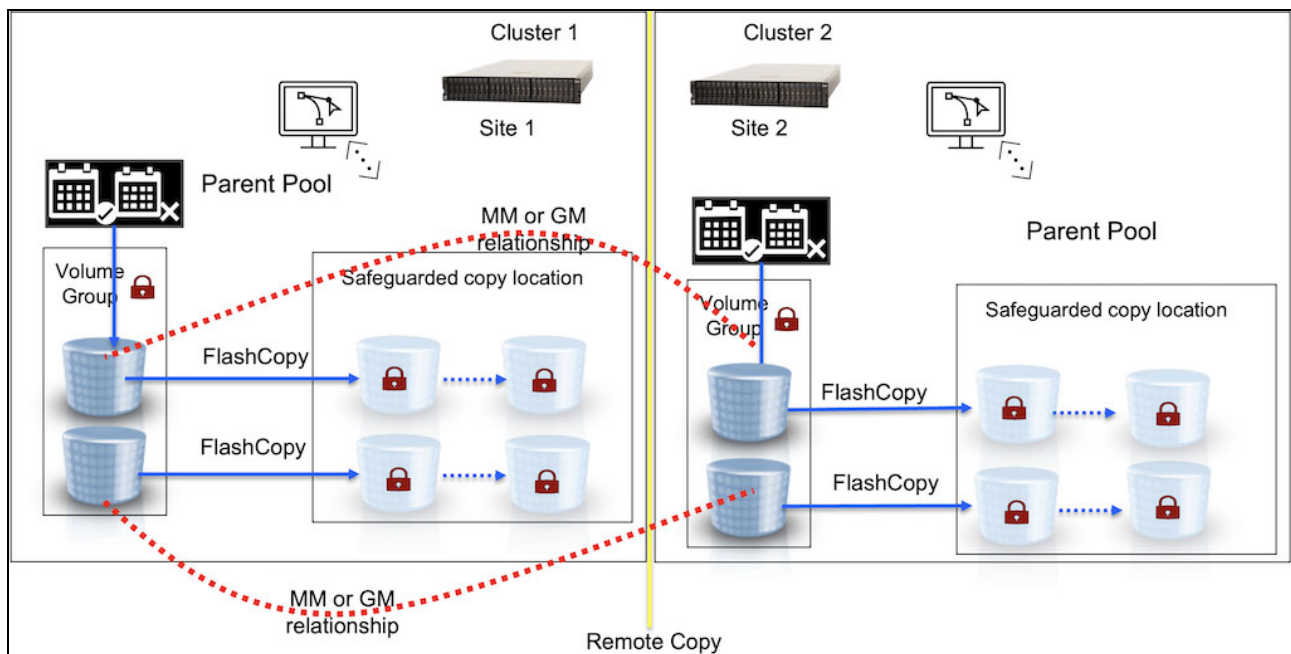


Figure 7-2 Safeguarded copy on Metro Mirror or Global Mirror relationship

7.3 Safeguarded Copy in an IBM HyperSwap environment

Likewise, in an IBM HyperSwap® environment, you can choose whether to make the Safeguarded Copies at the primary site, secondary site or both. See Figure 7-3.

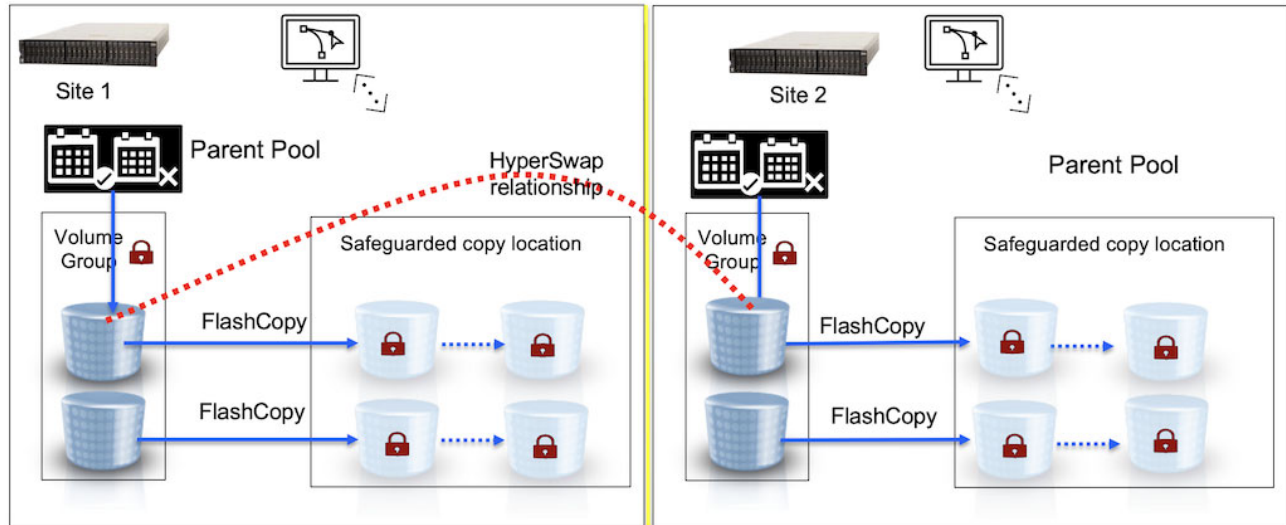


Figure 7-3 Safeguarded Copy in a HyperSwap environment

When you restore data in a HyperSwap environment, the most efficient method is to recover the data to the single-site copy, regardless of whether it is located on the primary or secondary site. After the data has been recovered and a new single-site copy has been created, mount the data to the host and verify the data. If necessary, a secondary copy can be created in a HyperSwap relationship.

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this book.

IBM Redbooks

The following IBM Redbooks publications provide additional information about the topic in this document. Note that some publications referenced in this list might be available in softcopy only.

- ▶ *Implementation Guide for IBM Storage FlashSystem and IBM SAN Volume Controller: Updated for IBM Storage Virtualize Version 8.6*, SG24-8542
- ▶ *Performance and Best Practices Guide for IBM Storage FlashSystem and IBM SAN Volume Controller: Updated for IBM Storage Virtualize Version 8.6*, SG24-8543
- ▶ *IBM FlashSystem Safeguarded Copy Implementation Guide*, REDP-5654

You can search for, view, download or order these documents and other Redbooks, Redpapers, Web Docs, draft and additional materials, at the following website:

ibm.com/redbooks

Stay connected to IBM Redbooks

- ▶ Look for us on LinkedIn:
<http://www.linkedin.com/groups?home=&gid=2130806>
- ▶ Explore new IBM 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>

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services



Cyber Resiliency with IBM Storage Sentinel and IBM

SG24-8541-00
ISBN DocISBN



(1.5" spine)
1.5" <-> 1.998"
789 <-> 1051 pages



Cyber Resiliency with IBM Storage Sentinel and IBM Storage

SG24-8541-00
ISBN DocISBN



(1.0" spine)
0.875" <-> 1.498"
460 <-> 788 pages



Cyber Resiliency with IBM Storage Sentinel and IBM Storage

SG24-8541-00
ISBN DocISBN



(0.5" spine)
0.475" <-> 0.873"
250 <-> 459 pages



Cyber Resiliency with IBM Storage Sentinel and IBM Storage Safeguarded

(0.2" spine)
0.17" <-> 0.473"
90 <-> 249 pages

(0.1" spine)
0.1" <-> 0.169"
53 <-> 89 pages



Cyber Resiliency with IBM Storage Sentinel and IBM Storage Safeguarded Copy

SG24-8541-00

ISBN DocISBN

(2.5" spine)
2.5" <-> nnn.n"
1315<-> nnnn pages



Cyber Resiliency with IBM Storage Sentinel and IBM Storage Safeguarded Copy

SG24-8541-00

ISBN DocISBN

(2.0" spine)
2.0" <-> 2.498"
1052 <-> 1314 pages





SG24-8541-00

ISBN 0738461350

Printed in U.S.A.

Get connected

