IBM Storage Protect
8.1.21

Introduction to Data Protection Solutions

Note:

Before you use this information and the product it supports, read the information in “Notices” on page 47.

Edition notice

This edition applies to version 8, release 1, modification 21 of IBM® Storage Protect (product numbers 5725-W98, 5725-W99, 5725-X15), and to all subsequent releases and modifications until otherwise indicated in new editions.

Contents

About this publication

This publication provides an overview of IBM Storage Protect concepts and data protection solutions that use best practices for IBM Storage Protect. A feature comparison chart helps you select the best solution for your organization's needs.

Who should read this guide

This guide is intended for anyone who is registered as an administrator for IBM Storage Protect. A single administrator can manage IBM Storage Protect, or several people can share administrative responsibilities.

You should be familiar with the operating system on which the server resides and the communication protocols required for the client/server environment. You also need to understand the storage management practices of your organization, such as how you are currently backing up workstation files and how you are using storage devices.

Publications

The IBM Storage Protect product family includes IBM Storage Protect Plus, IBM Storage Protect for Virtual Environments, IBM Storage Protect for Databases, and several other storage management products from IBM.

To view IBM product documentation, see IBM Documentation.

Part 1. IBM Storage Protect concepts

IBM Storage Protect provides a comprehensive data protection environment.

Chapter 1. IBM Storage Protect overview

IBM Storage Protect provides centralized, automated data protection that helps to reduce data loss and manage compliance with data retention and availability requirements.

Data protection components

The data protection solutions that IBM Storage Protect provides consist of a server, client systems and applications, and storage media. IBM Storage Protect provides management interfaces for monitoring and reporting the data protection status.

Server

Client systems send data to the server to be stored as backups or archived data. The server includes an *inventory*, which is a repository of information about client data.

The inventory includes the following components:

Database
Information about each file, logical volume, or database that the server archives, migrates or creates a backup of, is stored in the server database. The server database also contains information about the policy and schedules for data protection services.

Recovery log
Records of database transactions are kept in this log. The recovery log helps to maintain data consistency in the database.

Client systems and applications

Clients are applications, virtual machines, and systems that must be protected. The clients send data to the server, as shown in Figure 1 on page 3.

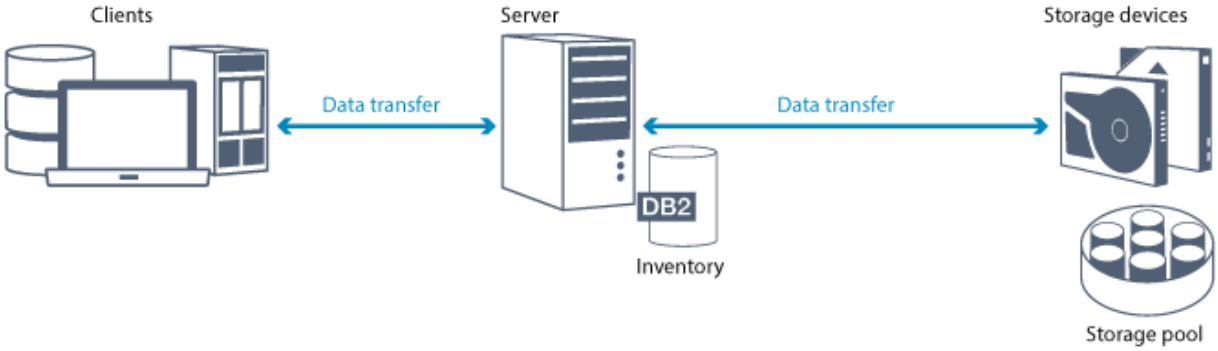


Figure 1. Components in the data protection solution

Client software
For IBM Storage Protect to protect client data, the appropriate software must be installed on the client system and the client must be registered with the server.

Client nodes
A *client node* is equivalent to a computer, virtual machine, or application, such as a backup-archive client that is installed on a workstation for file system backups. Each client node must be registered with the server. Multiple nodes can be registered on a single computer.

Storage media

The server stores client data to storage media. The following types of media are used:

Storage devices

The server can write data to hard disk drives, disk arrays and subsystems, stand-alone tape drives, tape libraries, and other types of random-access and sequential-access storage. Storage devices can be connected directly to the server or connected through a local area network (LAN) or a storage area network (SAN).

Storage pools

Storage devices that are connected to the server are grouped into *storage pools*. Each storage pool represents a set of storage devices of the same media type, such as disk or tape drives. IBM Storage Protect stores all of the client data in storage pools. You can organize storage pools into a *hierarchy*, so that data storage can be transferred from disk storage to lower-cost storage such as tape devices.

Cloud object storage

IBM Cloud Object Storage (COS) is a scalable cloud storage service that is designed for durability, resiliency, and security. The service helps to store, manage, and access your data through IBM's self-service portal and RESTful APIs. You can also use the COS service to connect applications directly to COS for using other IBM Cloud services with your data.

Data that is stored by using IBM Cloud Object Storage is encrypted and dispersed across multiple geographic locations and accessed over the Hypertext Transfer Protocol (HTTP) by using a REST API. The COS service makes use of the distributed storage technologies provided by IBM Cloud Object Storage System.

Data protection services

IBM Storage Protect provides data protection services to store and recover data from various types of clients. The data protection services are implemented through policies that are defined on the server. You can use client scheduling to automate the data protection services.

Types of data protection services

IBM Storage Protect provides services to store and recover client data as shown in Figure 2 on page 4.

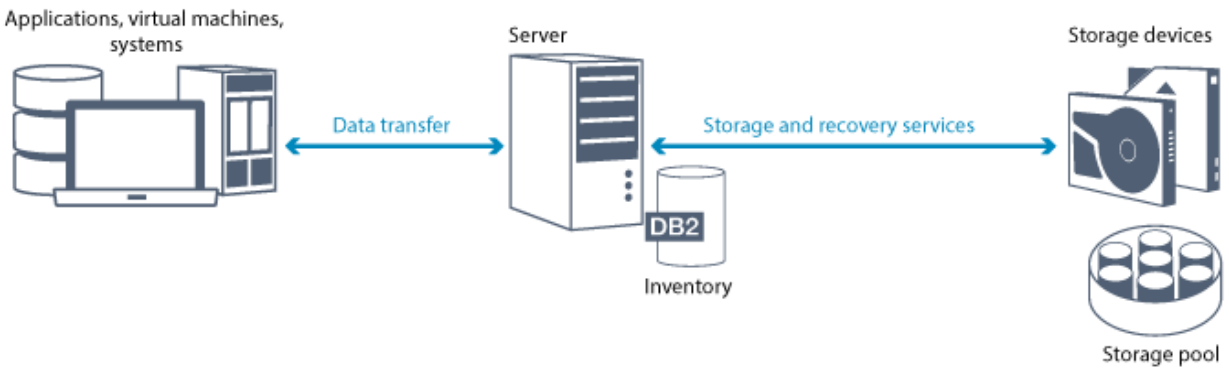


Figure 2. Data protection services

IBM Storage Protect provides the following types of data protection services:

Back up and restore services

You run a backup process to create a copy of a *data object* that can be used for recovery if the original data object is lost. A data object can be a file, a directory, or a user-defined data object, such as a database.

To minimize the use of system resources during the backup operation, IBM Storage Protect uses the *progressive incremental backup* method. For this backup method, a first full backup of all data objects is created and in subsequent backup operations only changed data is moved to storage. Compared to incremental and differential backup methods that require taking periodic full backups, the progressive incremental backup method provides the following benefits:

- Reduces data redundancy
- Uses less network bandwidth
- Requires less storage pool space

To further reduce storage capacity requirements and network bandwidth usage, IBM Storage Protect includes *data deduplication* for data backups. The data deduplication technique removes duplicate data extents from backups.

You run a restore process to copy an object from a storage pool to the client. You can restore a single file, all files in a directory, or all of the data on a computer.

Archive and retrieve services

You use the archive service to preserve data that must be stored for a long time, such as for regulatory compliance. The archive service provides the following features:

- When you archive data, you specify how long the data must be stored.
- You can request that files and directories are copied to long-term storage on media. For example, you might choose to store this data on a tape device, which can reduce the cost of storage.
- You can specify that the original files are erased from the client after the files are archived.

The retrieve service provides the following features:

- When you retrieve data, the data is copied from a storage pool to a client node.
- The retrieve operation does not affect the archive copy in the storage pool.

Migrate and recall services

You use migrate and recall services to manage space on client systems. The goal of space management is to maximize available media capacity for new data and to minimize access time to data. You can migrate data to server storage to maintain sufficient free storage space on a local file system. You can store migrated data in the following ways:

- On disk storage for long-term storage
- In a *virtual tape library* (VTL) for fast recall of files

You can recall files to the client node on demand, either automatically or selectively.

Types of client data that can be protected

You can protect data for the following types of clients with IBM Storage Protect:

Application clients

IBM Storage Protect can protect data for specific products or applications. These clients are called *application clients*. To protect the *structured data* for these clients, in other words the data in database fields, you must back up components that are specific to the application. IBM Storage Protect can protect the following applications:

- IBM Storage Protect for Enterprise Resource Planning clients:
  - Data Protection for SAP HANA
  - Data Protection for SAP for Db2®
  - Data Protection for SAP for Oracle
- IBM Storage Protect for Databases clients:
  - Data Protection for Microsoft SQL server
  - Data Protection for Oracle

- IBM Storage Protect for Mail clients:
  - Data Protection for HCL Domino®
  - Data Protection for Microsoft Exchange Server

Virtual machines
Virtual machines that are backed up by using application client software that is installed on the virtual machine. In the IBM Storage Protect environment, a virtual machine can be protected by the IBM Storage Protect for Virtual Environments.

System clients
The following IBM Storage Protect clients are called *system clients*:

- All clients that back up data in files and directories, in other words *unstructured data*, such as backup-archive clients and API clients that are installed on workstations.
- A server that is included in a server-to-server virtual volume configuration.
- A virtual machine that is backed up by using backup-archive client software that is installed on the virtual machine.

Processes for managing data protection with IBM Storage Protect

The IBM Storage Protect server inventory has a key role in the processes for data protection. You define policies in the IBM Storage Protect server inventory that the server uses to manage data storage.

Data management process

Figure 3 on page 6 shows the IBM Storage Protect data management process.

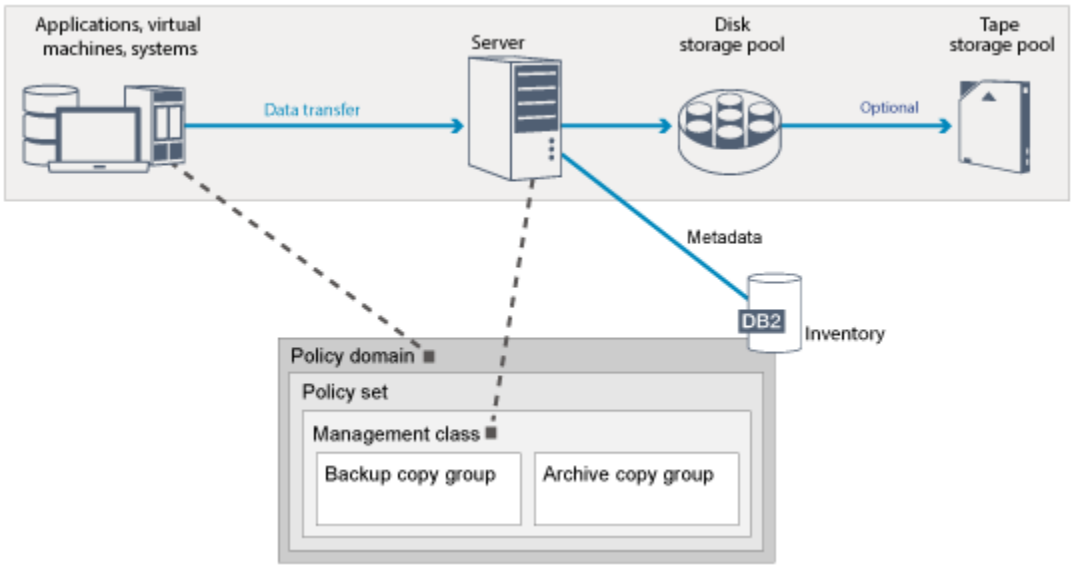


Figure 3. Data management process

IBM Storage Protect uses policies to control how the server stores and manages data objects on various types of storage devices and media. You associate a client with a policy domain that contains one active policy set. When a client creates a backup, archives, or migrates a file, the file is bound to a management class in the active policy set of the policy domain. The management class and the backup and archive copy groups specify where files are stored and how they are managed. If you set up server storage in a hierarchy, you can migrate files to different storage pools.

Inventory components

The following inventory components are key to the operation of the server:

Server database

The server database contains information about client data and server operations. The database stores information about client data, called *metadata*. Information about client data includes the file name, file size, file owner, management class, copy group, and location of the file in server storage. The database includes the following information that is necessary for the operation of the server:

- Definitions of client nodes and administrators
- Policies and schedules
- Server settings
- Records of server operations, such as activity logs and event records
- Intermediate results for administrative queries

Recovery log

The server records database transactions in the recovery log. The recovery log helps to ensure that a failure does not leave the database in an inconsistent state. The recovery log is also used to maintain consistency across server start operations. The recovery log consists of the following logs:

Active log

This log records current transactions on the server. This information is required to start the server and database after a disaster.

Log mirror (optional)

The active log mirror is a copy of the active log that can be used if the active log files cannot be read. All changes that are made to the active log are also written to a log mirror. You can set up one active log mirror.

Archive log

The archive log contains copies of closed log files that were in the active log. The archive log is included in database backups and is used for recovery of the server database. Archive log files that are included in a database backup are automatically pruned after a full database backup cycle is complete. The archive log must have enough space to store the log files for database backups.

Archive failover log (optional)

The archive failover log, also called a secondary archive log, is the directory that the server uses to store archive log files when the archive log directory is full.

Policy-based data management

In the IBM Storage Protect environment, a *policy* for data protection management contains rules that determine how client data is stored and managed. The primary purpose of a policy is to implement the following data management objectives:

- Control which storage pool is the client data initially stored in
- Define retention criteria that controls how many copies of objects are stored
- Define how long copies of objects are retained

Policy-based data management helps you to focus on the business requirements for protecting data rather than on managing storage devices and media. Administrators define policies and assign client nodes to a *policy domain*.

Depending on your business needs, you can have one policy or many. In a business organization, for example, different departments with different types of data can have customized storage management plans. Policies can be updated, and the updates can be applied to data that is already managed.

When you install IBM Storage Protect, a default policy that is named STANDARD is already defined. The STANDARD policy provides basic backup protection for user workstations. To provide different levels of service for different clients, you can add to the default policy or create a new policy.

You create policies by defining the following policy components:

Policy domain
The policy domain is the primary organizational method of grouping client nodes that share common rules for data management. Although a client node can be defined to more than one server, the client node can be defined to only one policy domain on each server.

Policy set
A *policy set* is a number of policies that are grouped so that the policy for the client nodes in the domain can be activated or deactivated as required. An administrator uses a policy set to implement different management classes based on business and user needs. A policy domain can contain multiple policy sets, but only one policy set can be active in the domain. Each policy set contains a default management class and any number of extra management classes.

Management class

A *management class* is a policy object that you can bind to each category of data to specify how the server manages the data. There can be one or more management classes. One management class is assigned to be the default management class that is used by clients unless they specifically override the default to use a specific management class.

The management class can contain a backup copy group, an archive copy group, and space management attributes. A copy group determines how the server manages backup versions or archived copies of the file. The space management attributes determine whether the file is eligible for migration by the space manager client to server storage, and under what conditions the file is migrated.

Copy group
A *copy group* is a set of attributes in a management class that controls the following factors:

- Where the server stores versions of backed up files or archive copies
- How long the server keeps versions of backed up files or archive copies
- How many versions of backup copies are retained
- What method to use to generate versions of backed up files or archive copies

Security management

IBM Storage Protect includes security features for registration of administrators and users. After administrators are registered, they must be granted authority by being assigned one or more administrative privilege classes. An administrator with system privilege can perform any server function. Administrators with policy, storage, operator, or node privileges can perform subsets of server functions. The server can be accessed by using the following methods, each controlled with a password:

- Administrator access to manage the server
- Client access to nodes to store and retrieve data

Also includes features that can help to ensure security when clients connect to the server with the following closed registration method:

Closed registration
Closed registration is the default method for client registration to the server. For this type of registration, an administrator registers all clients. The administrator can implement the following settings:

- Assign the node to any policy domain
- Determine whether the user can use compression or not, or if the user can choose
- Control whether the user can delete backed up files or archived files

You can add more protection for your data and passwords by using Secure Sockets Layer (SSL). SSL is the standard technology that you use to create encrypted sessions for servers and clients, and provides a secure channel to communicate over open communication paths. With SSL, the identity of the server is verified by using digital certificates. If you authenticate with a Lightweight Directory Access Protocol (LDAP) server, passwords between the server and the LDAP server are protected by Transport

Layer Security (TLS). The TLS protocol is the successor to the SSL protocol. When a server and client communicate, TLS ensures that third parties cannot intercept messages.

User interfaces for the IBM Storage Protect environment

For monitoring and configuration tasks, IBM Storage Protect provides various interfaces, including the Operations Center, a command-line interface, and an SQL administrative interface.

Interfaces for data storage management

The Operations Center is the primary interface for administrators to monitor and administer servers. A key benefit of the Operations Center is that you can monitor multiple servers, as shown in Figure 4 on page 9. You can also monitor and administer IBM Storage Protect from a command-line administrative interface.

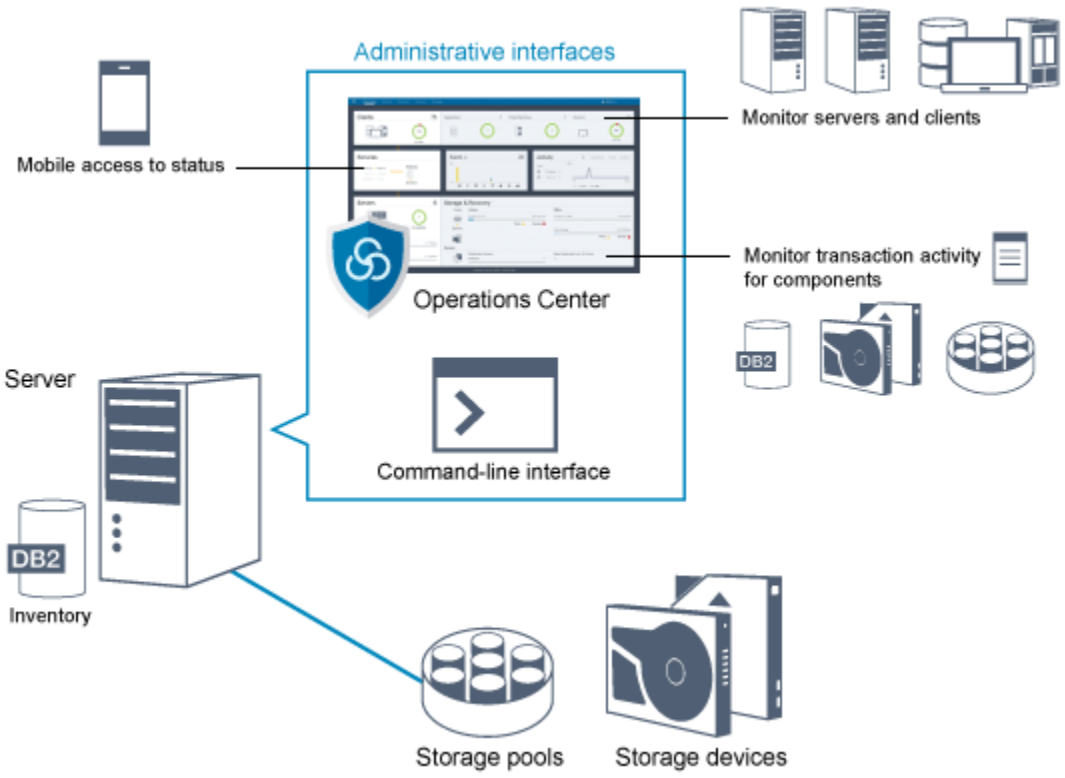


Figure 4. User interfaces for data storage management

You use the following interfaces to interact with IBM Storage Protect:

Operations Center

The Operations Center provides web and mobile access to status information about the IBM Storage Protect environment. You can use the Operations Center to complete monitoring and certain administration tasks, for example:

- You can monitor multiple servers and clients.
- You can monitor the transaction activity for specific components in the data path, such as the server database, the recovery log, storage devices, and storage pools.

Command-line interface

You can use a command-line interface to run administration tasks for servers. You can access the command-line interface through either the IBM Storage Protect administrative client or the Operations Center.

Access to information in the server database by using SQL statements

You can use SQL SELECT statements to query the server database and display the results. Third-party SQL tools are available to aid administrators in database management.

Interfaces for client activity management

IBM Storage Protect provides the following types of interfaces for managing client activity:

- An application programming interface (API)
- Graphical user interfaces for clients
- Browser interface for the backup-archive client
- Command-line interfaces for clients

Chapter 2. Data storage concepts in IBM Storage Protect

IBM Storage Protect provides functions to store data in a range of device and media storage.

To make storage devices available to the server, you must attach the storage devices and map storage pools to device classes, libraries, and drives.

Types of storage devices

You can use various storage devices with IBM Storage Protect to meet specific data protection goals.

Storage devices and storage objects

The IBM Storage Protect server can connect to a combination of manual and automated storage devices. You can connect the following types of storage devices to IBM Storage Protect:

- Disk devices that are directly attached, SAN-attached, or network attached
- Physical tape devices that are either manually operated or automated
- Virtual tape devices
- Cloud object storage

IBM Storage Protect represents physical storage devices and media with storage objects that you define in the server database. Storage objects classify available storage resources and manage migration from one storage pool to another. Table 1 on page 11 describes the storage objects in the server storage environment.

Table 1. Storage objects and representations

Storage object	What the object represents
Volume	A discrete unit of storage on disk, tape, or other storage media. Each volume is associated with a single storage pool.
Storage pool	A set of storage volumes or containers that is the destination that is used to store client data. IBM Storage Protect uses the following types of storage pool: • Directory-container storage pools • Cloud-container storage pools • Sequential-access storage pools that are associated with a device class • Random-access storage pools that are associated with a device class
Container	A data storage location, for example, a file, directory, or device.
Container storage pool	A storage pool that a server uses to store data. Data is stored in containers in file system directories or in cloud storage. Data is deduplicated, if necessary, as the server writes data to the storage pool.

Table 1. Storage objects and representations (continued)	
Storage object	**What the object represents**
Device class	The type of storage device that can use the volumes that are defined in a sequential-access or random-access storage pool. Each device class of removable media type is associated with a single library.
Library	A storage device. For example, a library can represent a stand-alone drive, a set of stand-alone drives, a multiple-drive automated device, or a set of drives that is controlled by a media manager.
Drive	An object of a tape library device that provides the capability to read and write data to tape library media. Each drive is associated with a single library.
Path	The specification of the data source and the device destination. Before a storage device can be used, a path must be defined between the device and the source server that is moving data.
Data mover	A SAN-attached device that is used to transfer client data. A data mover is used only in a data transfer where the server is not present, such as in a Network Data Management Protocol (NDMP) environment. Data movers transfer data between storage devices without using significant server, client, or network resources.
Server	A server that is managed by another IBM Storage Protect server.

The administrator defines the storage objects in the logical layer of the server, as illustrated in Figure 5 on page 13.

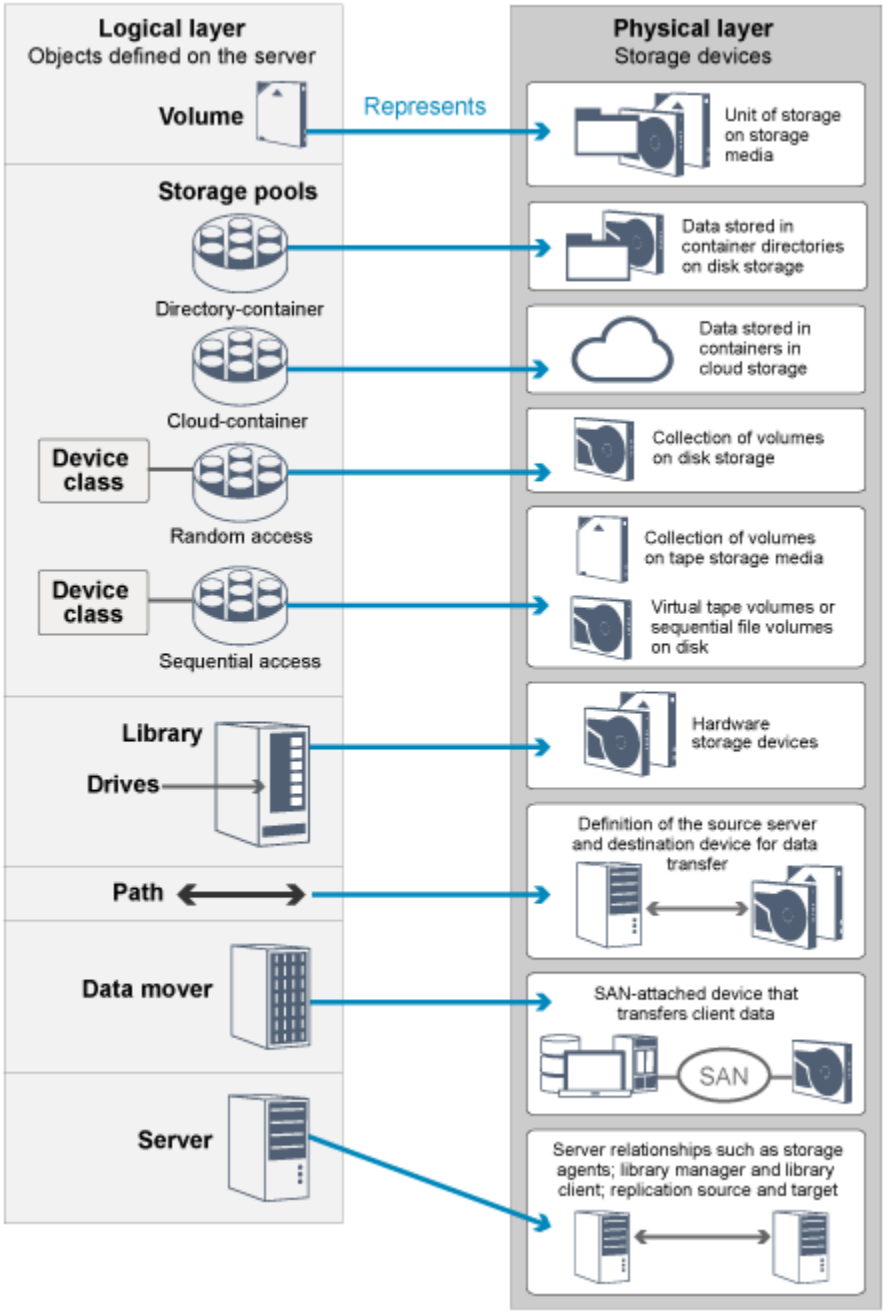


Figure 5. Storage objects

Disk devices

You can store client data on disk devices with the following types of volumes:

- Directories in directory-container storage pools
- Random-access volumes of device type DISK
- Sequential-access volumes of device type FILE

IBM Storage Protect offers the following features when you use directory-container storage pools for data storage:

- You can apply data deduplication and disk caching techniques to maximize data storage usage.
- You can retrieve data from disk much faster than you can retrieve data from tape storage.

Physical tape devices

In a physical tape library, the storage capacity is defined in terms of the total number of volumes in the library. Physical tape devices can be used for the following activities:

- Storing client data that is backed up, archived, or migrated from client nodes
- Storing database backups
- Exporting data to another server or offsite storage

Moving data to tape provides the following benefits:

- You can keep data for clients on a disk device at the same time that the data is moved to tape.
- You can improve tape drive performance by streaming the data migration from disk to tape.
- You can spread out the times when the drives are in use to improve the efficiency of the tape drives.
- You can move data on tape to off-site vaults.
- You can limit power consumption because tape devices do not consume power after data is written to tape.
- You can apply encryption that is provided by the tape drive hardware to protect the data on tape.

Compared to equivalent disk and virtual tape storage, the unit cost to store data tends to be much less for physical tape devices.

Virtual tape libraries

A virtual tape library (VTL) does not use physical tape media. When you use VTL storage, you emulate the access mechanisms of tape hardware. In a VTL, you can define volumes and drives to provide greater flexibility for the storage environment. The storage capacity of a VTL is defined in terms of total available disk space. You can increase or decrease the number and size of volumes on disk.

Defining a VTL to the IBM Storage Protect server can improve performance because the server handles mount point processing for VTLs differently than for real tape libraries. Although the logical limitations of tape devices are still present, the physical limitations for tape hardware are not applicable to a VTL thus affording better scalability. You can use the IBM Storage Protect VTL when the following conditions are met:

- Only one type and generation of drive and media is emulated in the VTL.
- Every server and storage agent with access to the VTL has paths that are defined for all drives in the library.

Data storage in storage pools

Logical storage pools are the principal components in the IBM Storage Protect model of data storage. You can optimize the usage of storage devices by manipulating the properties of storage pools and volumes.

Types of storage pools

The group of storage pools that you set up for the server is called *server storage*. You can define the following types of storage pools in server storage:

Primary storage pools
A named set of volumes that the server uses to store backup versions of files, archive copies of files, and files that are migrated from client nodes.

Copy storage pools
A named set of volumes that contain copies of files that are stored in primary storage pools. Copy storage pools are used only to back up the data that is stored in primary storage pools. A copy storage pool cannot be a destination for a backup copy group, an archive copy group, or a management class for space-managed files.

Container-copy storage pools
A named set of volumes that contain a copy of data extents that are stored in directory-container storage pools. Container-copy storage pools are used only to protect the data that is stored in directory-container storage pools.

Active-data storage pools
A named set of storage pool volumes that contain only active versions of client backup data.

Retention storage pools
A named set of volumes that the server uses to store copies of backup data that is to be retained for long-term storage. Retention storage pools are used only to store retained data. A retention storage pool cannot be a destination for a backup copy group, an archive copy group, or a management class for space-managed files.

Primary storage pools

When you restore, retrieve, recall, or export file data, the requested file is obtained from a primary storage pool. Depending on the type of primary storage pool, the storage pools can be onsite or offsite. You can arrange primary storage pools in a storage hierarchy so that data can be transferred from disk storage to lower-cost storage such as tape devices. Figure 6 on page 15 illustrates the concept of primary storage pools.

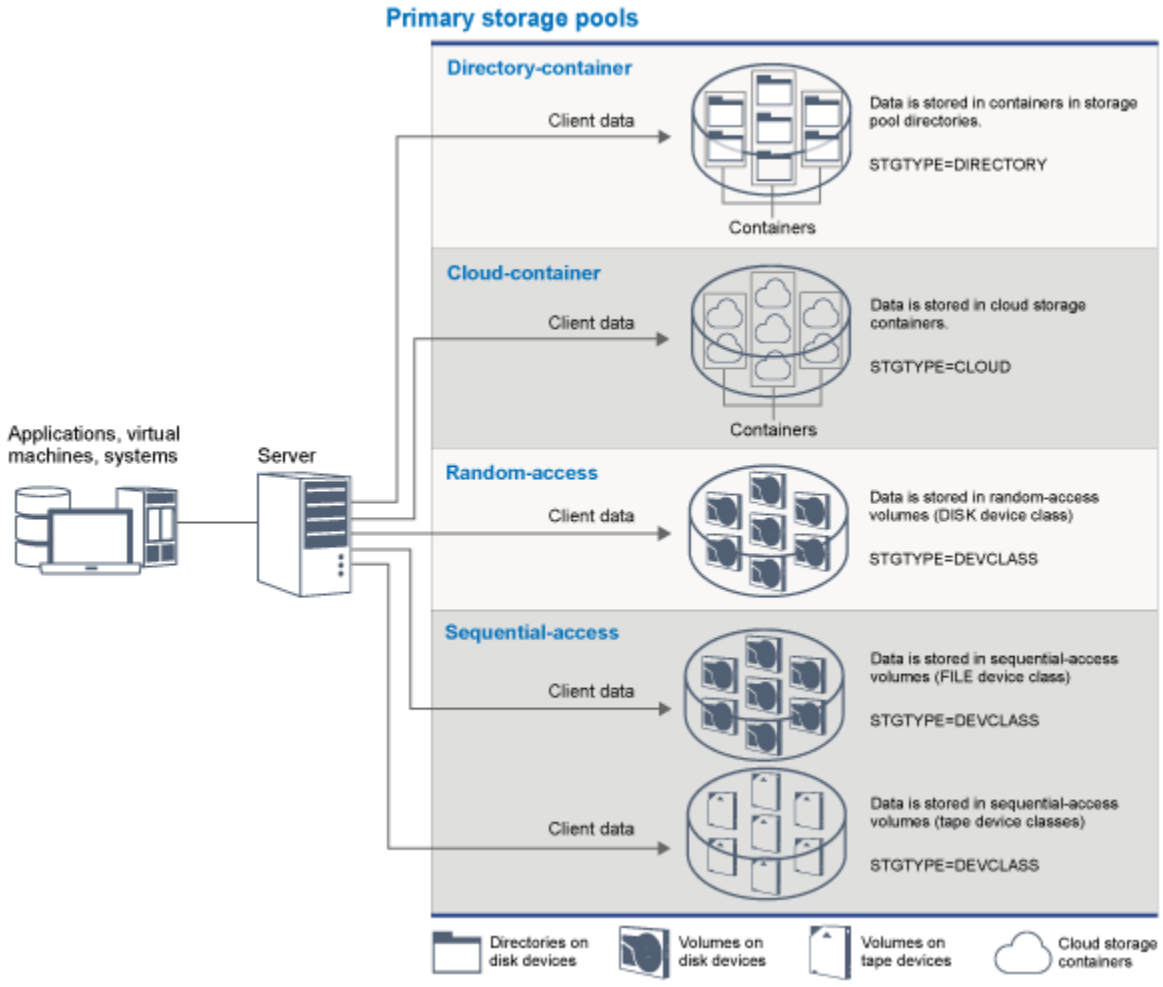


Figure 6. Primary storage pools

You can define the following types of primary storage pool:

Directory-container storage pools

A storage pool that the server uses to store data in containers in storage pool directories. Data that is stored in a directory-container storage pool can use either inline data deduplication, client-side data deduplication, inline compression, or client-side compression. Inline data deduplication or inline compression reduces data at the time it is stored.

Tip: Data that is compressed first cannot be deduplicated. However, deduplicated data can be compressed.

By using directory-container storage pools, you remove the need for volume reclamation, which improves server performance and reduces the cost of storage hardware. You can protect and repair data in directory-container storage pools at the level of the storage pool. You can tier data that is stored in a directory-container storage pool to a cloud-container storage pool.

Restriction: You cannot use any of the following functions with directory-container storage pools:

- Migration
- Reclamation
- Aggregation
- Colocation
- Simultaneous-write
- Storage pool backup
- Virtual volumes

Cloud-container storage pools

A storage pool that a server uses to store data in cloud storage. The cloud storage can be on premises or off premises. The cloud-container storage pools that are provided by IBM Storage Protect can store data to cloud storage that is object-based. By storing data in cloud-container storage pools, you can leverage the cost per unit advantages that clouds offer along with the scaling capabilities that cloud storage provides. You can use cloud tiering to lower costs by moving data from disk storage to a cloud-container storage pool. IBM Storage Protect manages the credentials, security, read and write I/Os, and the lifecycle for data that is stored to the cloud. When cloud-container storage pools are implemented on the server, you can write directly to the cloud by configuring a cloud-container storage pool with the cloud credentials. Data that is stored in a cloud-container storage pool uses both inline data deduplication and inline compression. The server writes deduplicated, compressed, and encrypted data directly to the cloud. You can back up and restore data or archive and retrieve data directly from the cloud-container storage pool.

You can define the following types of cloud-container storage pools:

On premises
You can use the on premises type of cloud-container storage pool to store data in a private cloud, for more security and maximum control over your data. The disadvantages of a private cloud are higher costs due to hardware requirements and onsite maintenance.

Off premises
You can use the off premises type of cloud-container storage pool to store data in a public cloud. The advantage of using a public cloud is that you can achieve lower costs than for a private cloud, for example by eliminating maintenance. However, you must balance this benefit against possible performance issues due to connection speeds and reduced control over your data.

Storage pools that are associated with device classes

You can define a primary storage pool to use the following types of storage devices:

DISK device class
In a DISK device type of storage pool, data is stored in random access disk blocks. You can use caching in DISK storage pools to increase client restore performance with some limitations on server processing. Space allocation and tracking by blocks uses more database storage space and requires more processing power than allocation and tracking by volume.

FILE device class
In a FILE device type of storage pool, files are stored in sequential volumes for better sequential performance than for storage in disk blocks. To the server, these files have the characteristics of a tape volume so that this type of storage pool is better suited for migration to tape. FILE volumes are useful for *electronic vaulting*, where data is transferred electronically to a remote site rather than by physical shipment of tape. In general, this type of storage pool is preferred over DISK storage pools.

The server uses the following default random-access primary storage pools:

ARCHIVEPOOL
In the STANDARD policy, this storage pool is the destination for files that are archived from client nodes.

BACKUPPOOL
In the STANDARD policy, this storage pool is the destination for files that are backed up from client nodes.

SPACEMGPOOL
This storage pool is for space-managed files that are migrated from IBM Storage Protect for Space Management client nodes.

Copy storage pools

Copy storage pools contain active and non-active versions of data that is backed up from primary storage pools. You cannot use a directory-container storage pool as a copy storage pool. In addition, data from a directory-container storage pool cannot be copied into a copy storage pool.

Figure 7 on page 17 illustrates the concept of copy storage pools.

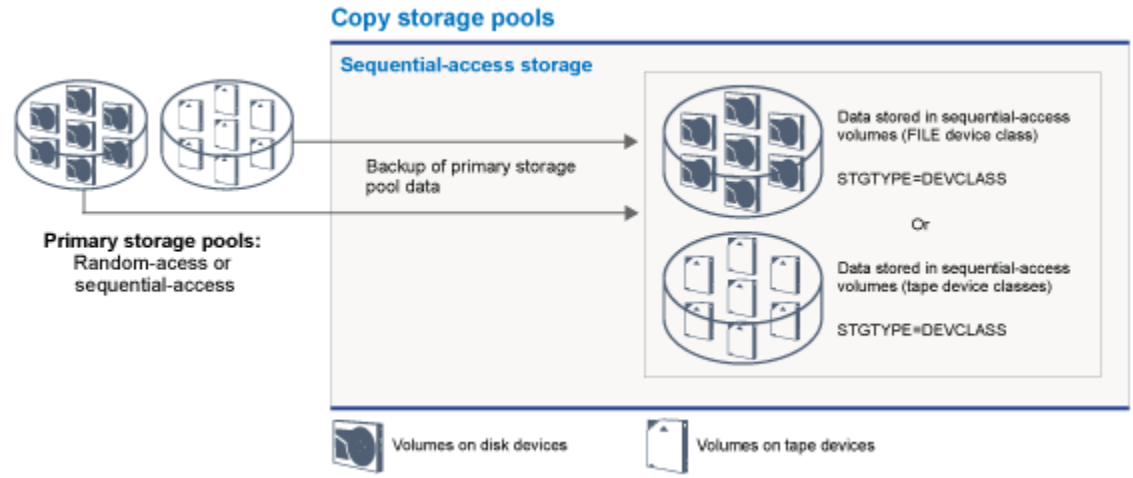


Figure 7. Copy storage pools

Copy storage pools provide a means of recovering from disasters or media failures. For example, when a client fails to retrieve a damaged file from the primary storage pool, the client can restore the data from the copy storage pool.

You can move the volumes of copy storage pools offsite and still have the server track the volumes. Moving these volumes offsite provides a means of recovering from an onsite disaster. A copy storage pool can use sequential-access storage only, such as a tape device class or FILE device class.

Container-copy storage pools

You can protect data in a directory-container storage pool by copying the data to container-copy storage pools, which are represented by tape volumes. The tape copy is used to repair damage to data in a directory-container storage pool. Data in container-copy storage pools is stored on tape volumes, which can be stored onsite or offsite. Damaged data in directory-container storage pools can be repaired by

using deduplicated extents in container-copy storage pools. Container-copy storage pools provide an alternative to using a replication server to protect data in a directory-container storage pool.

Restriction: If all server data is lost, container-copy storage pools alone do not provide the same level of protection as replication:

- With replication, you can directly restore client data from the target replication server if the source replication server is not available.
- With container-copy storage pools, you must first restore the server from a database backup and then repair directory-container storage pools from tape volumes.

Figure 8 on page 18 illustrates the concept of container-copy storage pools.

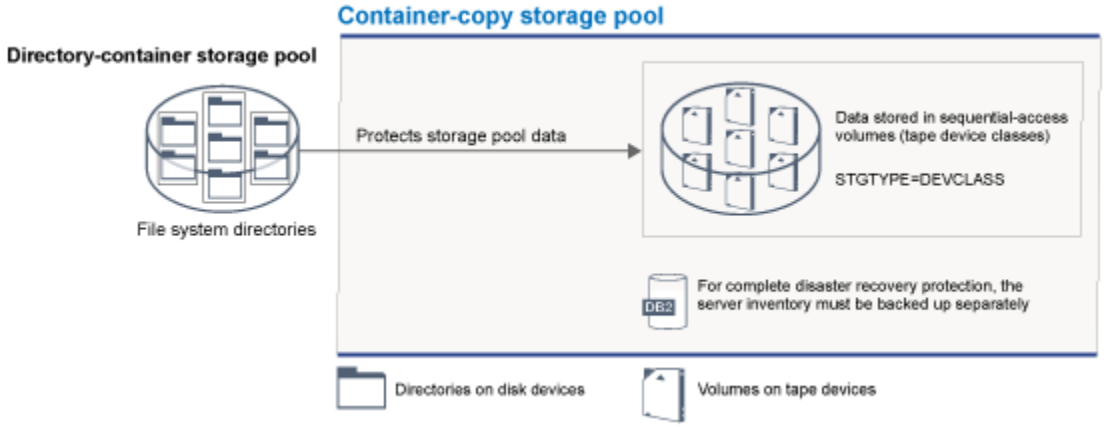


Figure 8. Container-copy storage pools

Depending on your system configuration, you can create protection schedules to simultaneously copy the directory-container storage pool data to onsite or offsite container-copy storage pools to meet your requirements:

- If replication is enabled, you can create one offsite container-copy pool. The offsite copy can be used to provide extra protection in a replicated environment.
- If replication is not enabled, you can create one onsite and one offsite container-copy storage pool.

Depending on the resources and requirements of your site, the ability to copy directory-container storage pools to tape has the following benefits:

- You avoid maintaining another server and more disk storage space.
- Data is copied to storage pools that are defined on the server. Performance is not dependent on, or affected by, the network connection between servers.
- You can satisfy regulatory and business requirements for offsite tape copies.

Active-data storage pools

An active-data pool contains only active versions of client backup data. In this case, the server does not have to position past non-active files that do not have to be restored. A directory-container storage pool cannot be used as an active-data storage pool. You use active-data pools to improve the efficiency of data storage and restore operations. For example, this type of storage pool can help you to achieve the following objectives:

- Increase the speed of client data restore operations.
- Reduce the number of onsite or offsite storage volumes.
- Reduce the amount of data that is transferred when you copy or restore files that are vaulted electronically in a remote location.

Data that is migrated by hierarchical storage management (HSM) clients and archive data are not permitted in active-data pools. As updated versions of backup data are stored in active-data pools,

older versions are removed as the remaining data is consolidated from many sequential-access volumes onto fewer, new sequential-access volumes. Figure 9 on page 19 illustrates the concept of active-data storage pools.

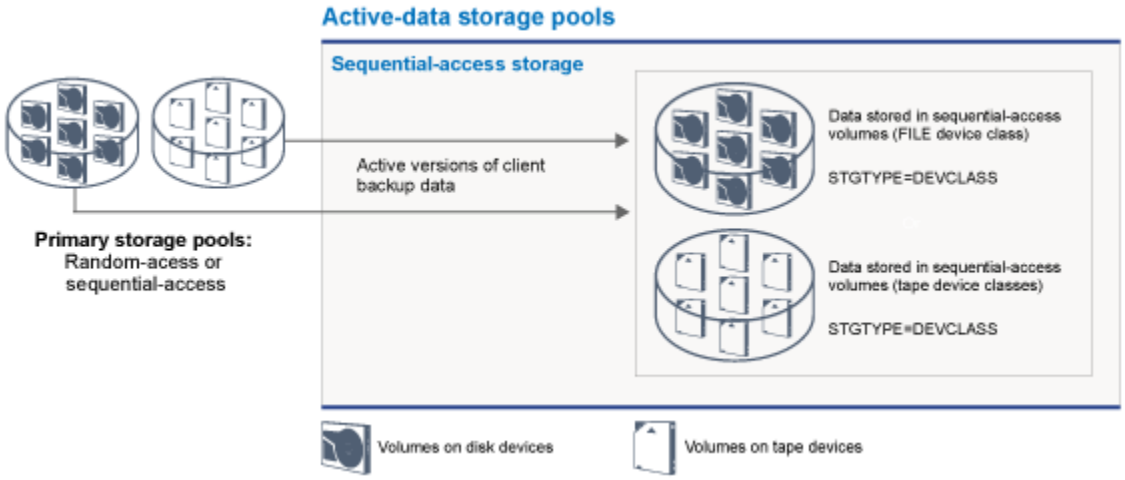


Figure 9. Active-data storage pools

Active-data pools can use any type of sequential-access storage. However, the benefits of an active-data pool depend on the device type that is associated with the pool. For example, active-data pools that are associated with a FILE device class are ideal for fast client restore operations because of the following reasons:

- FILE volumes do not have to be physically mounted.
- Client sessions that are restoring from FILE volumes in an active-data pool can access the volumes concurrently, which improves restore performance.

Retention storage pools

Retention storage pool is used to store copies of data that is retained by the server in retention sets. The data is copied from primary storage to a retention storage pool on tape or cloud object storage.

Figure 10 on page 19 illustrates the concept of retention storage pools.

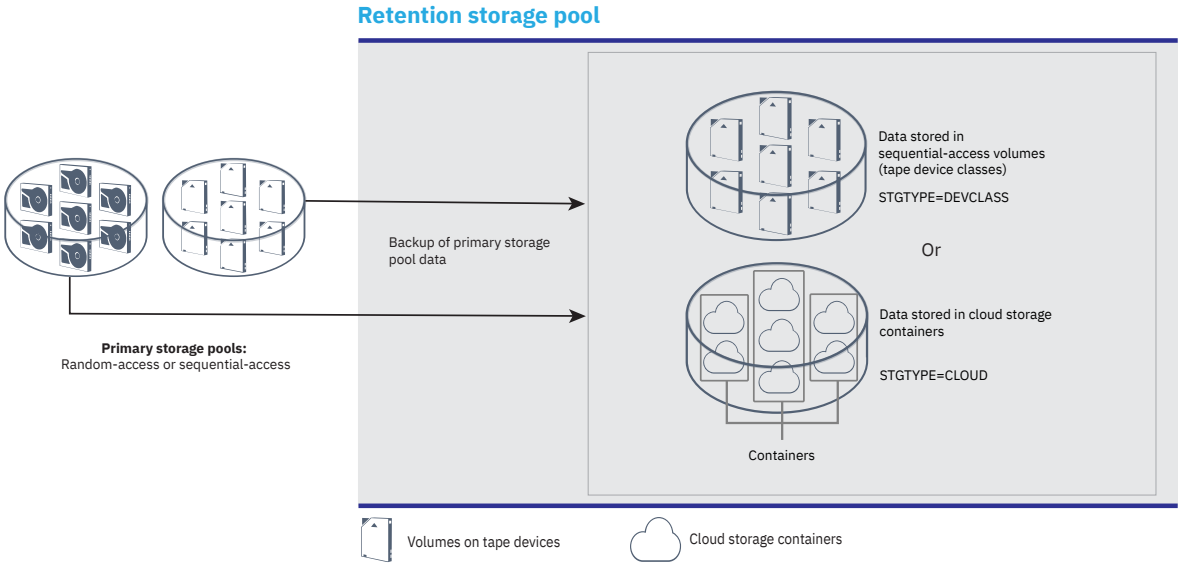


Figure 10. Retention storage pools

With retention storage pools, you can optimize the processes for storing retained data at an offsite location. You can separate long-term data from data that you want to keep on a short-term basis.

By using retention storage pools, you can reduce the need for other maintenance activities, such as reclamation. These activities are necessary when long-term data is stored together with short-term data. The short-term data is kept for operational recovery.

Tip: When you create a retention storage pool, a retention-copy storage rule with the same name is created automatically at the same time. The retention-copy storage rule runs once each day to copy retention set data from primary storage to the retention storage pool. However, you can run a retention-copy storage rule without waiting for the scheduled time.

A retention storage pool represents 3592 tape devices, LTO tape devices, Ecartridge devices, or CLOUD device classes. The retention storage pool can also use the tape library like other non-retention storage pools. Therefore, the storage devices such as tape drives and tape libraries defined with retention storage pool can be used to store all types of data.

Data transport to storage across networks

The IBM Storage Protect environment provides ways to securely move data to storage across various types of networks and configurations.

Network configurations for storage devices

IBM Storage Protect provides methods for configuring clients and servers on a local area network (LAN), on a storage area network (SAN), LAN-free data movement, and as network-attached storage.

Data backup operations over a LAN

Figure 11 on page 20 shows the data path for IBM Storage Protect backup operations over a LAN.

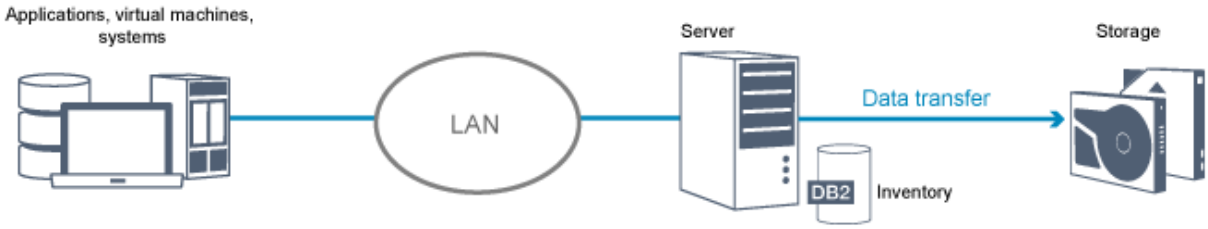


Figure 11. IBM Storage Protect backup operations over a LAN

In a LAN configuration, one or more tape libraries are associated with a single IBM Storage Protect server. In this type of configuration, client data, electronic mail, terminal connection, application program, and device control information must all be handled by the same network. Device control information and client backup and restore data flow across the LAN.

Data backup operations over a SAN

Figure 12 on page 20 shows the data path for IBM Storage Protect backup operations over a SAN.

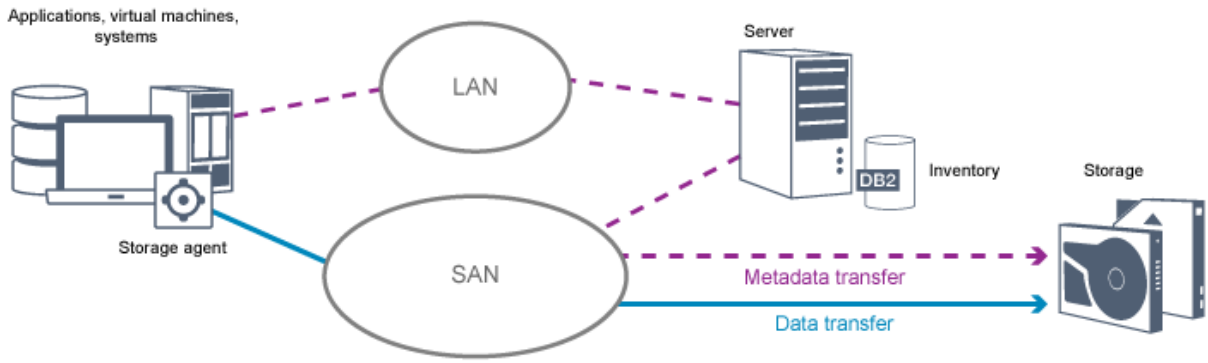


Figure 12. IBM Storage Protect backup operations over a SAN

A SAN is a dedicated storage network that can improve system performance. On a SAN, you can consolidate storage and relieve the distance, scalability, and bandwidth limitations of LANs and wide area networks (WANs). By using IBM Storage Protect in a SAN, you can take advantage of the following functions:

- Share storage devices among multiple IBM Storage Protect servers. Devices that use the GENERICTAPE device type are not included.
- Move data from a client system directly to storage devices without using the LAN. LAN-free data movement requires the installation of a storage agent on the client system. The storage agent is available with the IBM Storage Protect for SAN product.

  Through the storage agent, the client can directly back up and restore data to a tape library or shared file system such as . The IBM Storage Protect server maintains the server database and recovery log, and acts as the library manager to control device operations. The storage agent on the client handles the data transfer to the device on the SAN. This implementation frees bandwidth on the LAN that would otherwise be used for client data movement.
- Share tape drives and libraries that are supported by the IBM Storage Protect server.
- Consolidate multiple clients under a single client node name in a cluster.

Network-attached storage

Network-attached storage (NAS) file servers are dedicated storage servers whose operating systems are optimized for file-serving functions. NAS file servers typically interact with IBM Storage Protect through industry-standard network protocols, such as network data management protocol (NDMP) or as primary storage for random-access or sequential access storage pools. IBM Storage Protect provides the following basic types of configurations that use NDMP for backing up and managing NAS file servers:

- IBM Storage Protect backs up a NAS file server to a library device that is directly attached to the NAS file server. The NAS file server, which can be remote from the IBM Storage Protect server, transfers backup data directly to a drive in a SCSI-attached tape library. Data is stored in NDMP-formatted storage pools, which can be backed up to storage media that can be moved offsite for protection in case of an onsite disaster.
- IBM Storage Protect backs up a NAS file server over the LAN to a storage-pool hierarchy. In this type of configuration, you can store NAS data directly to disk, either random access or sequential access, and then migrate the data to tape. You can also use this type of configuration for system replication. Data can also be backed up to storage media that can be moved offsite. The advantage of this type of configuration is that you have all of the data management features associated with a storage pool hierarchy.
- The IBM Storage Protect client reads the data from the NAS system by using NFS or CIFS protocols and sends the data to the server to be stored.

Storage management

You manage the devices and media that are used to store client data through the IBM Storage Protect server. The server integrates storage management with the policies that you define for managing client data in the following areas:

Types of devices for server storage

With IBM Storage Protect, you can use directly attached devices and network-attached devices for server storage. IBM Storage Protect represents physical storage devices and media with administrator-defined storage objects.

Data migration through the storage hierarchy

For primary storage pools other than directory-container storage pools, you can organize the storage pools into one or more hierarchical structures. This storage hierarchy provides flexibility in a number of ways. For example, you can set a policy to back up data to disks for faster backup operations. The IBM Storage Protect server can then automatically migrate data from disk to tape.

Removal of expired data

The policy that you define controls when client data automatically expires from the IBM Storage Protect server. To remove data that is eligible for expiration, a server expiration process marks data as expired and deletes metadata for the expired data from the database. The space that is occupied by the expired data is then available for new data. You can control the frequency of the expiration process by using a server option.

Media reuse by reclamation

As server policies automatically expire data, the media where the data is stored accumulates unused space. For storage media other than directory-container storage pools or random disk storage pools, the IBM Storage Protect server implements *reclamation*, a process that frees media for reuse without traditional tape rotation. Reclamation automatically defragments media by consolidating unexpired data onto other media when the free space on media reaches a defined level. The reclaimed media can then be used again by the server. Reclamation allows media to be automatically circulated through the storage management process and minimize the number of media that are required.

Consolidating backed up client data

By grouping the client data that is backed up, you can minimize the number of media mounts required for client recovery. The IBM Storage Protect server provides the following methods for grouping client files on storage media other than directory-container storage pools:

Collocating client data

The IBM Storage Protect server can *collocate* client data, in other words store client data on a few volumes instead of spreading the data across many volumes. Collocation by client minimizes the number of volumes that are required to back up and restore client data. Data collocation might increase the number of volume mounts because each client might have a dedicated volume instead of data storage from several clients in the same volume.

You can set the server to collocate client data when the data is initially placed in server storage. In a storage hierarchy, you can collocate the data when the server migrates the data from the initial storage pool to the next storage pool in the storage hierarchy. You can collocate by client, by file space per client, or by a group of clients. Your selection depends on the size of the file spaces that are stored and restore requirements.

Associating active-data pools with various devices

Active-data pools are useful for fast restoration of client data. Benefits include a reduction in the number of onsite or offsite storage volumes, or reducing bandwidth when you copy or restore files that are vaulted electronically in a remote location. Active-data pools that use removable media, such as tape, offer similar benefits. Although tape devices must be mounted, the server does not have to position past inactive files. However, the primary benefit of using removable media in active-data pools is that the number of volumes that are used for onsite and offsite storage is reduced. If you store data to a remote location, you can minimize the amount of data that must be transferred by copying and restoring only active data.

Creating a backup set

A backup set contains all of the active backed-up files that exist for that client in server storage. The backup set is portable and is retained for the time that you specify. A backup set is in addition to the backups that are already stored and requires extra media.

Moving data for a client node

You can consolidate data for a client node by moving the data within server storage. You can move a backup set to different media, where the backup set is retained until the time that you specify. Consolidating data can help to improve efficiency during client restore or retrieve operations.

Chapter 3. Data protection strategies with IBM Storage Protect

IBM Storage Protect provides ways for you to implement various data protection strategies.

You can configure IBM Storage Protect to send data to storage devices that are on the local site or on a remote site. To maximize data protection, you can configure replication to a remote server.

Strategies to minimize the use of storage space for backups

To minimize the amount of storage space that is required, IBM Storage Protect backs up data by using the data deduplication and progressive incremental backup techniques.

Data deduplication

When the IBM Storage Protect server receives data from a client, the server identifies duplicate data extents and stores unique instances of the data extents in a directory-container storage pool. The data deduplication technique improves storage utilization and eliminates the need for a dedicated data deduplication appliance.

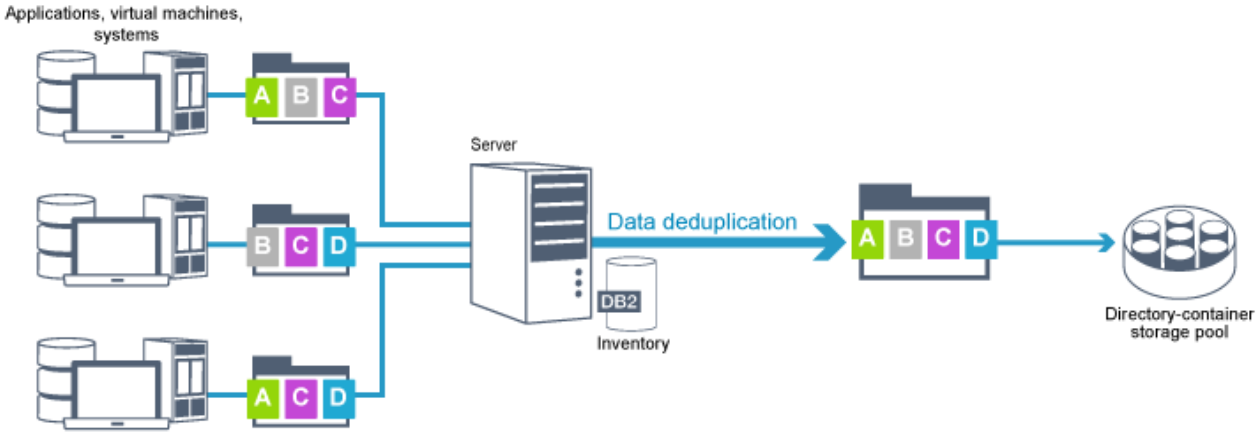


Figure 13. Data deduplication process

If the same byte pattern occurs many times, data deduplication greatly reduces the amount of data that must be stored or transferred. In addition to whole files, IBM Storage Protect can also deduplicate parts of files that are common with parts of other files.

IBM Storage Protect provides the following types of data deduplication:

Server-side data deduplication

The server identifies duplicate data extents and moves the data to a directory-container storage pool. The server-side process uses *inline data deduplication*, where data is deduplicated at the same time that the data is written to a directory-container storage pool. Deduplicated data can also be stored in other types of storage pools. Inline data deduplication on the server provides the following benefits:

- Eliminates the need for reclamation
- Reduces the space that is occupied by the stored data

Client-side data deduplication

With this method, processing is distributed between the server and the client during a backup process. The client and the server identify and remove duplicate data to save storage space on the server. In client-side data deduplication, only compressed, deduplicated data is sent to the server.

The server stores the data in the compressed format that is provided by the client. Client-side data deduplication provides the following benefits:

- Reduces the amount of data that is sent over the local area network (LAN)
- Eliminates extra processing power and time that is required to remove duplicate data on the server
- Improves database performance because the client-side data deduplication is also inline

You can combine both client-side and server-side data deduplication in the same production environment. The ability to deduplicate data on either the client or the server provides flexibility in terms of resource utilization, policy management, and data protection.

Compression

Use inline compression to reduce the amount of space that is stored in container storage pools. Data is compressed as it is written to the container storage pool.

Restriction: The IBM Storage Protect server cannot compress encrypted data.

Progressive incremental backup

In a progressive incremental backup process, the server monitors client activity and backs up any files that are new or have changed since the last backup. Entire files are backed up, so that the server does not need to reference base versions of the files. This backup technique eliminates the need for multiple full backups of client data thus saving network resources and storage space.

Strategies for disaster protection

IBM Storage Protect provides strategies to protect data if a disaster occurs. These strategies include node replication to a remote site, multi-site replication to two remote sites, storage pool protection, database backups, moving backup tapes offsite, and device replication to a standby server.

Replication to a remote site

Node replication is the process of incrementally copying data from one server to another server. The server from which client data is replicated is called a *source replication server*. The server to which client data is replicated is called a *target replication server*. For the purposes of disaster protection, the target replication server is on a remote site. A replication server can function as a source replication server, a target replication server, or both. You use replication processing to maintain the same level of files on the source and the target replication servers.

Node replication provides for immediate availability of data through failover. Although node replication protects most of the metadata, this approach does not provide adequate protection for database damage. You can provide more comprehensive protection by using storage pools to store data backups.

Advantages

- Failover so that data is available immediately if a disaster occurs.
- Incremental replication, which results in fast transmission of data.
- Electronic transfer
- Protects both data and most metadata

Disadvantages

- Both data and metadata must be recovered.
- Data on the source replication server must be replicated again from the remote site.

Figure 14 on page 25 shows the node replication process to a remote site.

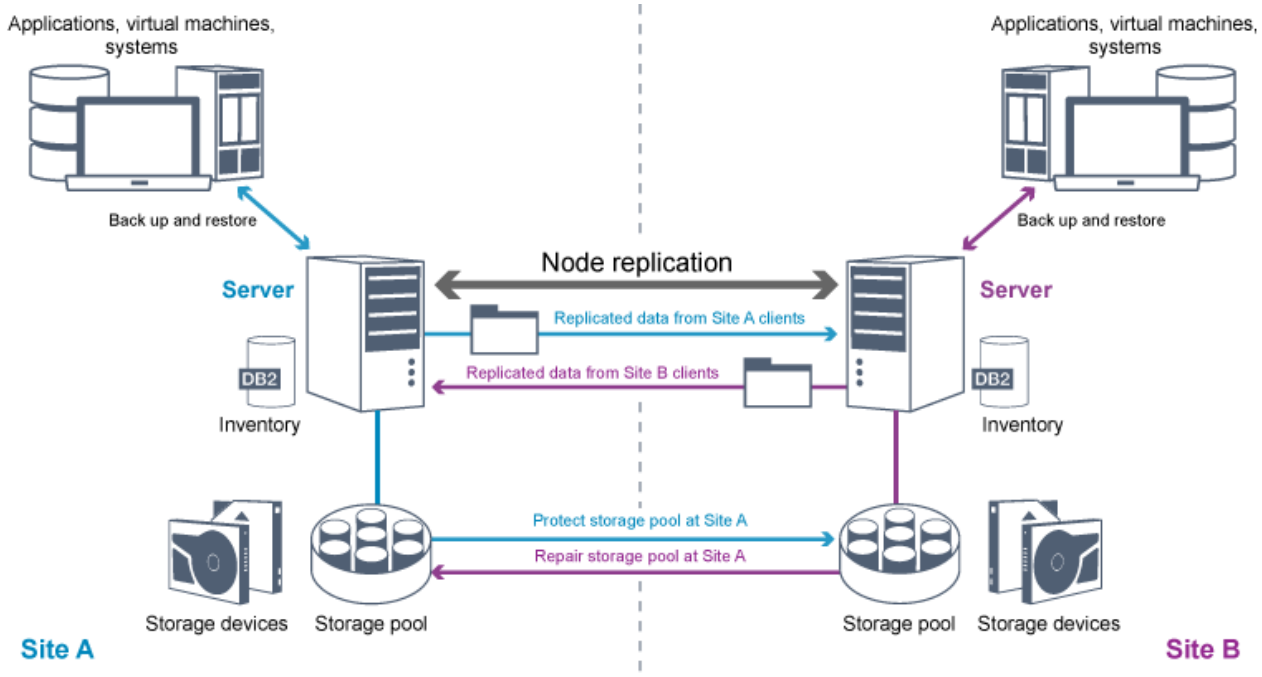


Figure 14. Node replication process

When client data is replicated, data that is not on the target replication server is copied to the target replication server. When replicated data exceeds the retention limit, the target replication server automatically removes the data from the source replication server. To maximize data protection, you synchronize the local server and the remote server; for example, Site B replicates data from Site A and Site A replicates data from Site B. As part of replication processing, client data that was deleted from the source replication server is also deleted from the target replication server.

IBM Storage Protect provides the following replication functions:

- You can define policies for the target replication server in the following ways:
  - Identical policies on the source replication server and target replication server
  - Different policies on the source replication server and target replication server to meet different business requirements.

  If a disaster occurs and the source replication server is not available, clients can recover data from the target replication server. If the source replication server cannot be recovered, you can direct clients to store data on the target replication server. When an outage occurs, the clients that are backed up to the source replication server can automatically fail over to restore their data from the target replication server.
- You can use replication processing to recover damaged files from storage pools. You must replicate the client data to the target replication server before the file damage occurs. Subsequent replication processes detect damaged files on the source replication server and replace the files with undamaged files from the target replication server.

Role of replication in disaster protection

If a disaster occurs, you can recover replicated data from the remote site and maintain the same level of files on the source and target replication servers. You use replication to achieve the following objectives:

- Control network throughput by scheduling node replication at specific times
- Recover data after a site loss.
- Recover damaged files on the source replication server.

Storage pool protection

As part of a disaster recovery strategy, ensure that a backup copy of data in storage pools is available at a remote site.

Advantages

- Fast recovery and rebuild of the source system.

Disadvantages

- Only data is protected; metadata is not protected.
- For each storage pool, you must define the storage medium.

You use different techniques to protect against the permanent loss of data that is stored in container storage pools and in FILE and DISK storage pools.

Directory-container storage pools

If you do not need to replicate all the data that is contained in a client node, you use container-copy storage pools to protect some directory-container storage pools. By protecting a directory-container storage pool, you do not use resources that replicate existing data and metadata, which improves server performance.

The preferred method is to protect the directory-container storage pool before you replicate the client node. When node replication is started, the data extents that are already replicated through storage pool protection are skipped, which reduces the replication processing time. If the data in a directory-container storage pool becomes damaged, you can repair the data from a copy in a container-copy storage pool.

Container-copy storage pools

You protect directory-container storage pools by copying the data in the directory-container storage pool to container-copy storage pools. Use container-copy storage pools to create up to two tape copies of a directory-container storage pool. The tape copies can be stored onsite or offsite. Damaged data in directory-container storage pools can be repaired by using container-copy storage pools. Container-copy storage pools provide an alternative to using a replication server to protect data in a directory-container storage pool.

Storage pools that are associated with FILE and DISK device classes

For storage pools that are associated with FILE and DISK device classes, you use node replication to maintain a node-consistent copy of the data at the target replication server. The data copy can be directly restored from the target replication server to the storage pools.

Database backups

You use database backups to recover your system following database damage. Also, database backup operations must be used to prevent Db2 from running out of archive log space. Database backup operations are not part of node replication. A database backup can be full, incremental, or snapshot. To provide for disaster recovery, a copy of the database backups must be stored offsite. To restore the database, you must have the backup volumes for the database. You can restore the database from backup volumes by either a point-in-time restore or a most current restore operation.

Point-in-time restore

Use point-in-time restore operations for situations such as disaster recovery or to remove the effects of errors that can cause inconsistencies in the database. Restore operations for the database that use snapshot backups are a form of point-in-time restore operation. The point-in-time restore operation includes the following actions:

- Removes and re-creates the active log directory and archive log directory that are specified in the `dsmserv.opt` file.
- Restores the database image from backup volumes to the database directories that are recorded in a database backup or to new directories.

- Restores archive logs from backup volumes to the overflow directory.
- Uses log information from the overflow directory up to a specified point in time.

Most current restore

If you want to recover the database to the time when the database was lost, recover the database to the most current state. The most current restore operation includes the following actions:

- Restores a database image from the backup volumes to the database directories that are recorded in a database backup or to new directories.
- Restores archive logs from backup volumes to the overflow directory.
- Uses log information from the overflow directory and archive logs from archive log directory.

The most current restore does not remove and re-create the active log directory or archive log directory.

Alternative methods for disaster protection

In addition to replication, storage pool protection, and database backups, you can also use the following methods to protect data and implement disaster recovery with IBM Storage Protect:

Sending backup tapes to a remote site

Data is backed up to tape at scheduled times by the source replication server. The tapes are sent to a remote site. If a disaster occurs, the tapes are returned to the site of the source replication server and the data is restored on the source clients. Offsite copies of data on backup tape can also help you to recover from ransomware attacks.

Multisite appliance replication to a standby server

In the multisite appliance configuration, the source appliance is replicated to a remote server in a SAN architecture. In this configuration, if the client hardware at the original site is damaged, the source device can be replicated from the standby server at the remote site. This configuration provides disk-based backup and restore operations.

Comparison of protection configuration strategies

Consider the following potential data-loss scenarios:

- Database data is damaged: protect against loss of data in the database by using onsite database backup.
- Storage pool data is damaged: protect against loss of data in storage pools by using onsite copy storage pools or node replication.
- Disaster scenario where both the onsite database and storage pools are lost: protect against a full disaster by using node replication and both off-site database backup and storage pool backup copies.

The following possible configurations address the most common data protection scenarios:

Configurations for damage protection only

- Implement database backup operations onsite with an optional container-copy storage pool onsite to protect data in directory-container storage pools.
- Implement database backup operations onsite and node replication onsite.

Configurations for disaster recovery and damage protection

- Implement database backup operations offsite with container-copy storage pools offsite to protect data in directory-container storage pools.
- Implement database backup operations onsite and node replication offsite with an optional container-copy storage pool onsite for faster recovery of damaged data.

Strategies for disaster recovery with IBM Storage Protect

IBM Storage Protect provides several ways to recover the server if the database or storage pools fail.

Automatic failover for disaster recovery

Automatic failover is an operation that switches to a standby system if a software, hardware, or network interruption occurs. Automatic failover is used with node replication to recover data after a system failure. Figure 15 on page 28 shows the IBM Storage Protect automatic failover process.

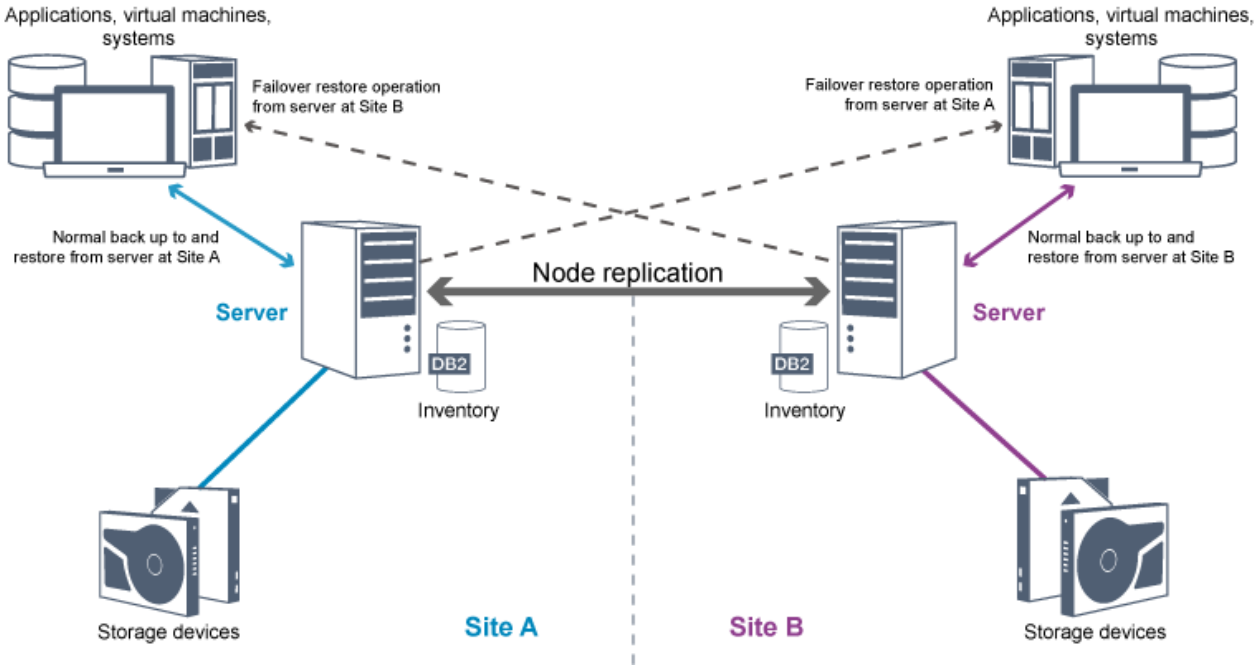


Figure 15. Automatic failover process

Automatic failover for data recovery occurs if the source replication server is unavailable because of a disaster or a system outage. During normal operations, when the client accesses a source replication server, the client receives connection information for the target replication server. The client node stores the failover connection information in the client options file.

During client restore operations, the server automatically changes clients from the source replication server to the target replication server and back again. Only one server per node can be used for failover protection at any time. When a new client operation is started, the client attempts to connect to the source replication server. The client resumes operations on the source server if the source replication server is available.

To use automatic failover for replicated client nodes, the source replication server, the target replication server, and the client must be at the V7.1 level or later. If any of the servers are at an earlier level, automatic failover is disabled and you must rely on a manual failover process.

Recovery of IBM Storage Protect components

The server database, recovery log, and storage pools are critical to the operation of IBM Storage Protect and must be protected. If the database is unusable, the entire server is unavailable and recovering data that is managed by the server might be difficult or impossible.

Even without the database, fragments of data or complete files might be read from storage pool volumes that are not encrypted and security can be compromised. Therefore, you must always back up the database. Also, always encrypt sensitive data by using the client or the storage device, unless the storage media is physically secured.

IBM Storage Protect provides several data protection methods, which include backing up storage pools and the database. For example, you can define schedules so that the following operations occur:

- After the initial full backup of your storage pools, incremental storage pool backups are run every night.
- Incremental database backups are run every night.
- Full database backups are run once a week.

For tape-based environments, you can use disaster recovery manager (DRM) to assist you in many of the tasks that are associated with protecting and recovering data. DRM is available with IBM Storage Protect Extended Edition.

Preventive actions for recovery

Recovery is based on the following preventive actions:

- Mirroring, by which the server maintains a copy of the active log
- Backing up the database
- Backing up the storage pools
- Auditing storage pools for damaged files and recovery of damaged files when necessary
- Backing up the device configuration and volume history files
- Validating the data in storage pools by using cyclic redundancy checking
- Storing the `cert.kdb` file in a safe place to ensure that the Secure Sockets Layer (SSL) is secure

If you are using tape for storage, you can also create a disaster recovery plan to guide you through the recovery process by using DRM. You can use the disaster recovery plan for audit purposes to certify the recoverability of the server. The disaster recovery methods of DRM are based on taking the following actions:

- Creating a disaster recovery plan file for the server
- Backing up server data to tape
- Sending the server backup data to a remote site or to another server
- Storing client system information
- Defining and tracking the storage media that is used for storing and recovering client data

If you take preventive actions for recovery, you should have backups for the following files and directories. The files and directories that are required to restore a server and its operations:

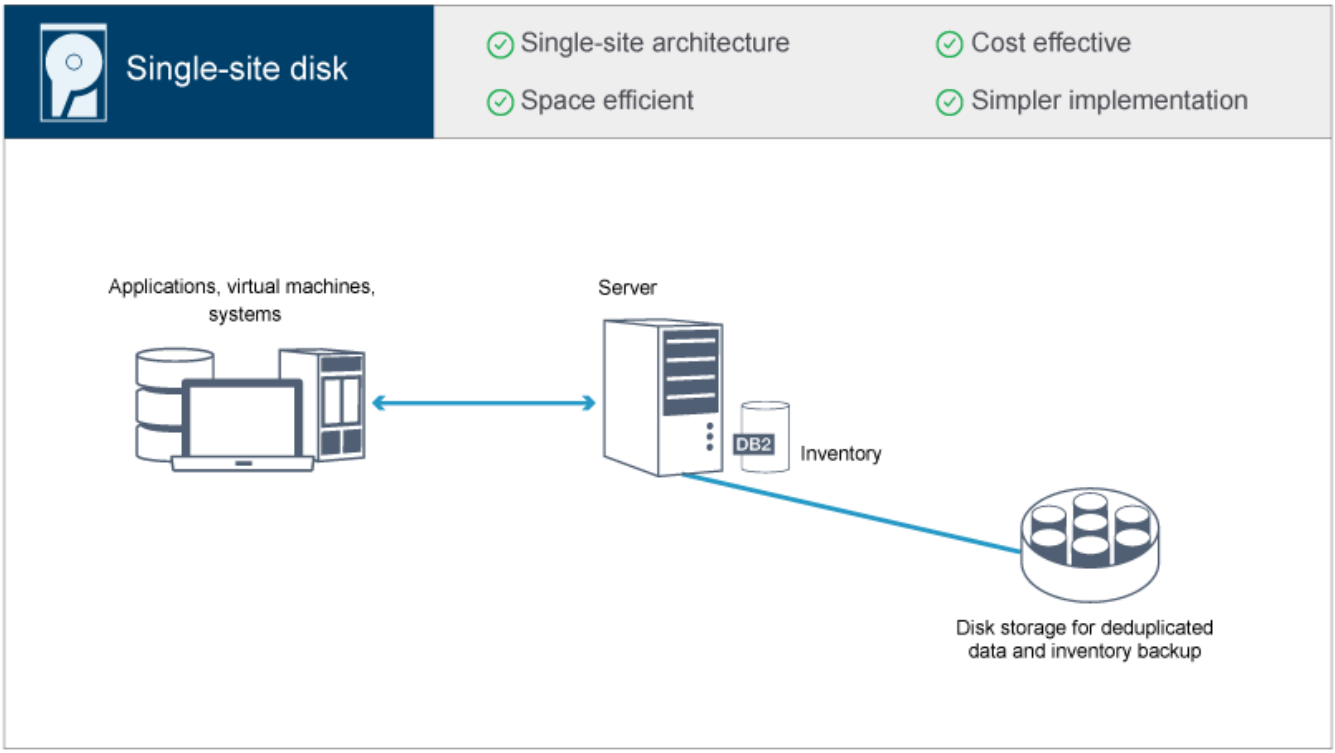
- Server options file (`dsmserv.opt`)
- Device configuration file (for example, `devconf.dat`)
- Volume history file (for example, `volhist.dat`)
- Master encryption key files (`dsmkeydb.kdb` or `dsmkeydb.sth`)
- Server certificate and private key files (`cert.kbd` or `cert.sth`)

Part 2. IBM Storage Protect solutions for data protection

To help you to deploy a data protection environment, review information about IBM Storage Protect configurations, and select the best solution for your business needs.

Chapter 4. Disk-based implementation of a data protection solution for a single site

This disk-based implementation of a data protection solution with IBM Storage Protect uses inline data deduplication and provides protection for data on a single site.



This data protection solution provides the following benefits:

- Server system and storage hardware at a single site
- Cost-effective use of storage through the data deduplication feature
- Space-efficient solution with minimal hardware setup
- Minimal implementation that requires installation and configuration for only one server and supporting storage hardware

In this solution, the client sends data to the IBM Storage Protect server, where the data is deduplicated and stored in a directory-container storage pool that is implemented in disk storage. Data from the inventory is also backed up to disk storage. This solution is suitable for entry-level environments for which a second copy of data is not required.

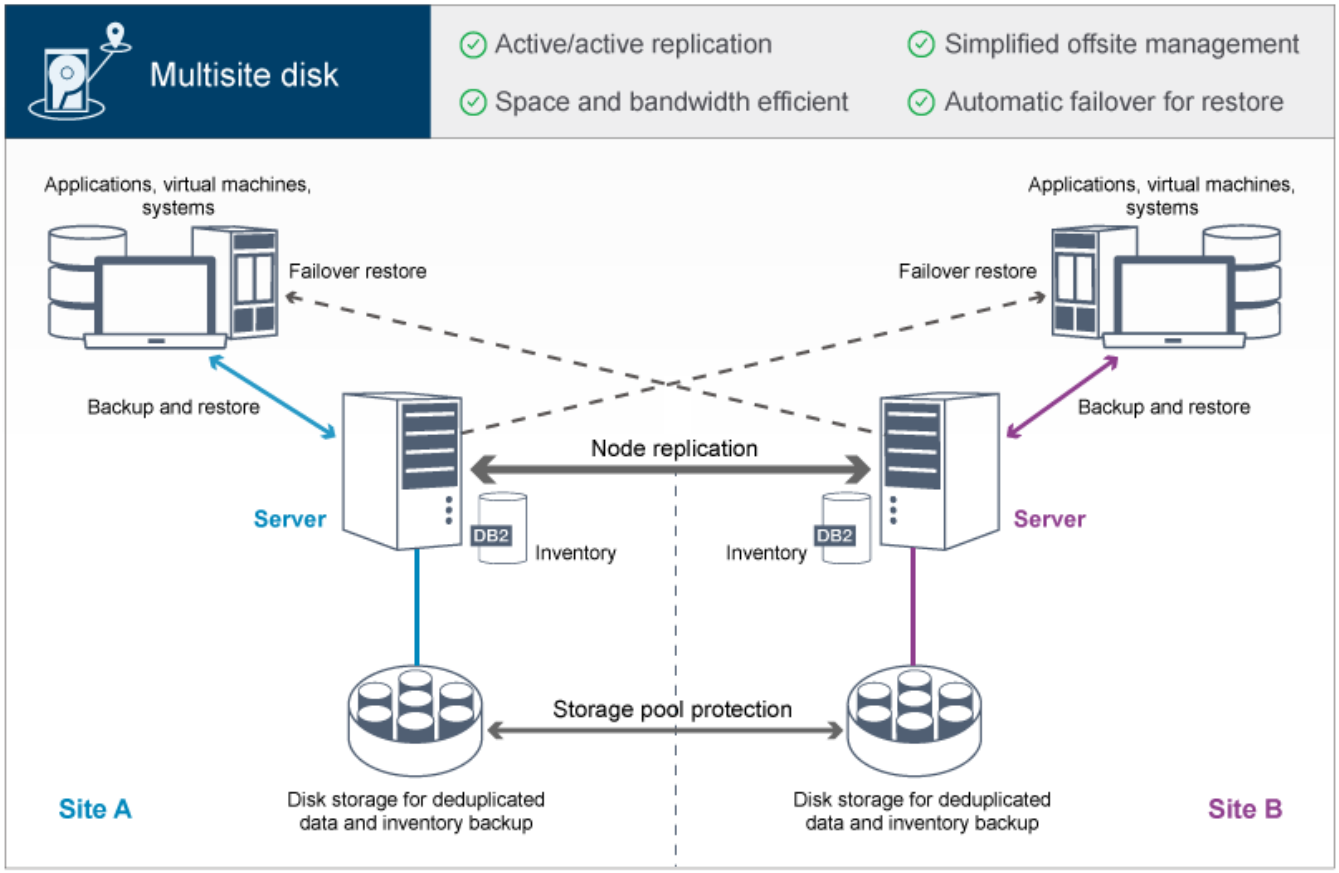
Related reference

Comparison of data protection solutions
Compare the key features for each IBM Storage Protect solution to determine which configuration best meets your data protection requirements. Then, review the available documentation to implement the solution.

Roadmap for implementing a data protection solution
Plan and implement the most suitable data protection solution for your business environment with IBM Storage Protect.

Chapter 5. Disk-based implementation of a data protection solution for multiple sites

This disk-based implementation of a data protection solution with IBM Storage Protect uses inline data deduplication and replication at two sites.



This data protection solution provides the following benefits:

- Replication can be configured at both sites so that each server protects data for the other site
- Offsite data storage for each location is simplified
- Bandwidth is used efficiently because only deduplicated data is replicated between the sites
- Clients can automatically fail over to a target replication server if the source replication server is unavailable

In this solution, clients send data to the source replication server, where the data is deduplicated and stored in a directory-container storage pool that is implemented in disk storage. The data is replicated to the storage pool on the target replication server for each site. This solution is suitable for environments that require disaster protection. If mutual replication is configured, clients at both sites can use failover recovery for continued backups and data recovery from the available server on the other site.

Related reference

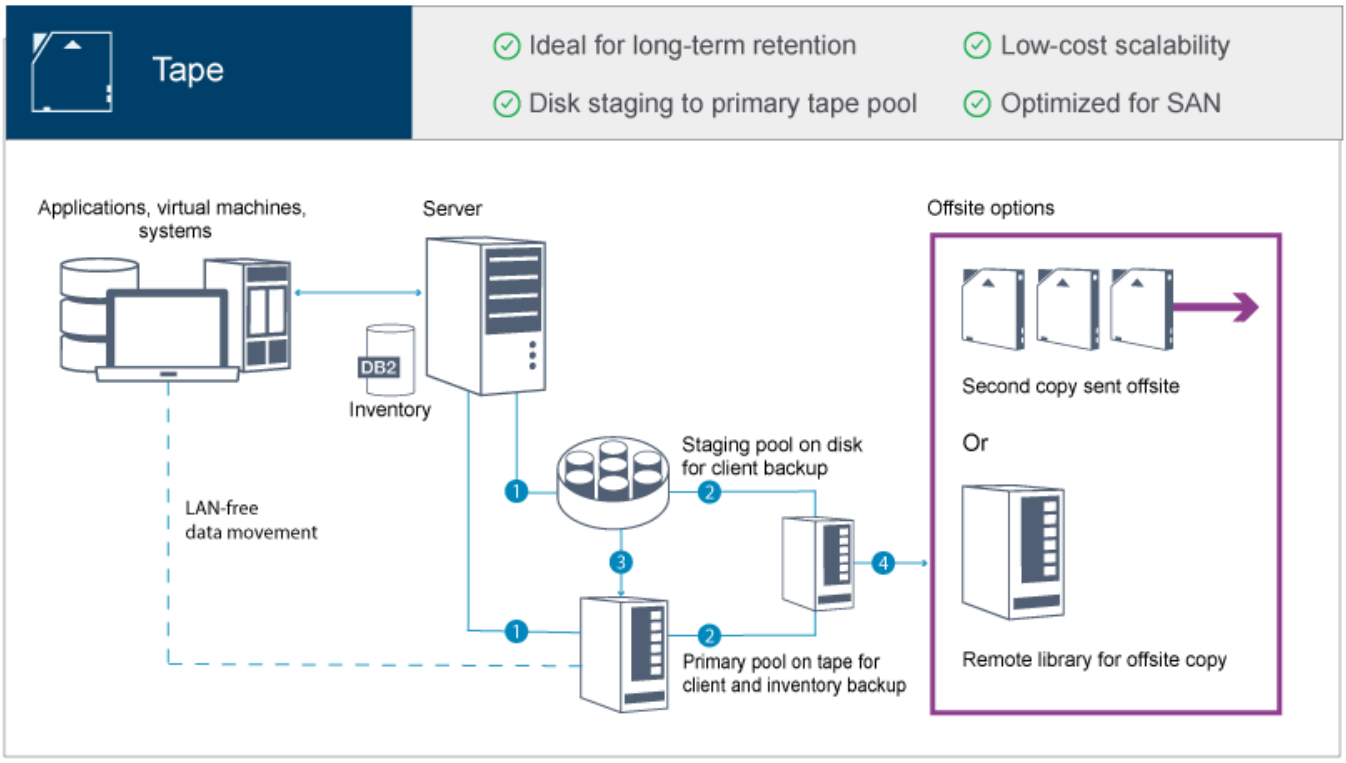
Comparison of data protection solutions
Compare the key features for each IBM Storage Protect solution to determine which configuration best meets your data protection requirements. Then, review the available documentation to implement the solution.

Roadmap for implementing a data protection solution

Plan and implement the most suitable data protection solution for your business environment with IBM Storage Protect.

Chapter 6. Tape-based implementation of a data protection solution

This implementation of a data protection solution with IBM Storage Protect uses one or more tape storage devices to back up data. Tape backup provides low-cost scalability that is optimized for long-term retention.



This data protection solution provides the following benefits:

- Performance is optimized for backup operations on high-speed storage area networks (SAN) directly to tape for large data types and for long-term retention of data.
- Data availability is optimized by storing copies of data at offsite locations for disaster recovery. If you enable the disaster recovery management (DRM) function and a disaster occurs, DRM helps to streamline the process of recovering your servers.
- Data security is optimized because copies of data are stored offsite on tape devices that are *not* connected to the internet. Ransomware attacks rely on internet connections; therefore, offsite storage can help to protect against such attacks.
- Low-cost scalability is achieved by reducing the need for additional disk hardware and lowering energy costs.

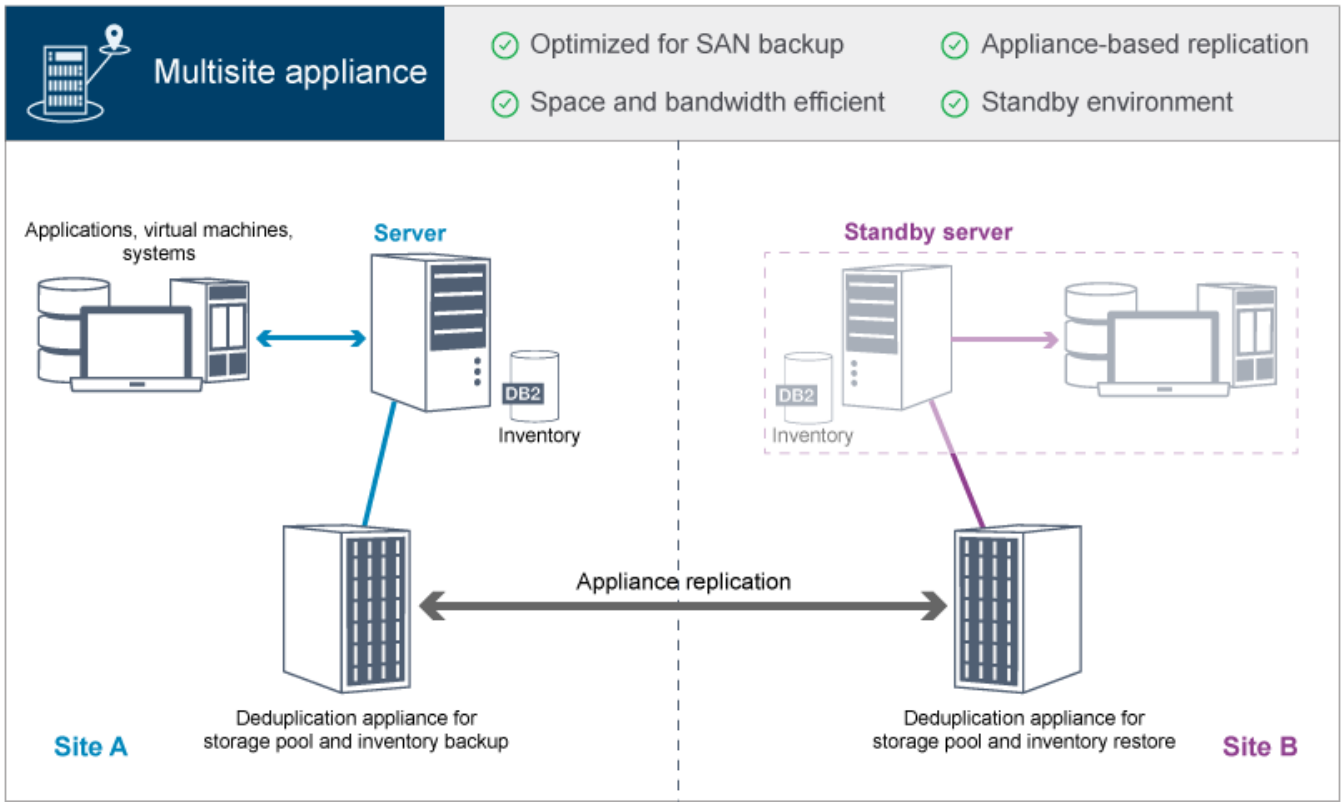
Related reference

Comparison of data protection solutions

Compare the key features for each IBM Storage Protect solution to determine which configuration best meets your data protection requirements. Then, review the available documentation to implement the solution.

Chapter 7. Appliance-based implementation of a data protection solution for multiple sites

This implementation of a multi-site IBM Storage Protect data protection solution uses appliance-based data deduplication and replication. A standby server is configured at a second site to recover data if the primary server is unavailable.



This data protection solution provides the following benefits:

- Performance is optimized for backups on high-speed storage area networks (SAN) and for use with IBM Storage Protect for SAN, when clients back up directly to SAN-attached virtual tape devices.
- Fast, appliance-based replication frees the server from having to track replication metadata in the server database.
- Bandwidth and storage space are used efficiently because only deduplicated data is replicated between the sites.
- A standby environment provides for disaster recovery, but does not require the amount of resources that are needed for a fully active site.

In this data protection configuration, the server uses hardware appliances to deduplicate and replicate data. The appliance at Site A deduplicates data and then replicates the data to the appliance at Site B for disaster protection. If a failure at Site A occurs, you make the standby server active by restoring the most recent database backup, and by activating the replicated copy of data.

For instructions about setting up a multisite appliance solution, see Chapter 9, “Roadmap for implementing a data protection solution,” on page 43.

For more information about configuring virtual tape libraries, see Configuring virtual tape libraries.

Related reference

Comparison of data protection solutions

Compare the key features for each IBM Storage Protect solution to determine which configuration best meets your data protection requirements. Then, review the available documentation to implement the solution.

Roadmap for implementing a data protection solution
Plan and implement the most suitable data protection solution for your business environment with IBM Storage Protect.

Chapter 8. Comparison of data protection solutions

Compare the key features for each IBM Storage Protect solution to determine which configuration best meets your data protection requirements. Then, review the available documentation to implement the solution.

	Single-site disk	**Multisite disk**	**Multisite appliance**	**Tape**
				
Highlights				
Cost	$	$$$	$$$$	$$
Protection level	One data copy	Two or more data copies	Two or more data copies	Two or more data copies
Disaster recovery	None	Active server	Standby server	Offsite copies
Key benefits				
Leading-edge data reduction	✓	✓	✓	✓
Fast and efficient disk-based backup and restore operations	✓	✓	✓	
Simplified offsite management		✓		
Data deduplication feature at no extra cost	✓	✓		
Replication processing included at no extra charge		✓		
Data deduplication at both the source and target replication server		✓		
Low-cost scalability and optimized for long-term retention				✓
Efficiency and cost				
Optimized for high-speed storage area network (SAN) backup operations			✓	✓
Optimized for high-speed local area network (LAN)	✓	✓	✓	
Global data deduplication across all data types and sources	✓	✓	✓	
Bandwidth-efficient replication		✓	✓	
Lower energy costs				✓
Option for a second copy without extra disk hardware				✓
Availability				

	Single-site disk	Multisite disk	Multisite appliance	Tape
				
Offsite copy capability		✓	✓	✓
Appliance-based replication			✓	
Client recovery from high-availability server		✓		
Replication target in the cloud		✓		
Independent management of retention policies for replication data; ability to keep more or less data at recovery site		✓		
Application-level replication; ability to choose which systems and applications are replicated		✓		
Scalability				
Global data deduplication across servers			✓	
SAN-optimized backup directly to tape for large data types				✓
Single-instance petabyte scalability				✓

What to do next

Review available documentation for the solutions in Chapter 9, “Roadmap for implementing a data protection solution,” on page 43.

Related reference

Disk-based implementation of a data protection solution for a single site
This disk-based implementation of a data protection solution with IBM Storage Protect uses inline data deduplication and provides protection for data on a single site.

Disk-based implementation of a data protection solution for multiple sites
This disk-based implementation of a data protection solution with IBM Storage Protect uses inline data deduplication and replication at two sites.

Appliance-based implementation of a data protection solution for multiple sites
This implementation of a multi-site IBM Storage Protect data protection solution uses appliance-based data deduplication and replication. A standby server is configured at a second site to recover data if the primary server is unavailable.

Tape-based implementation of a data protection solution
This implementation of a data protection solution with IBM Storage Protect uses one or more tape storage devices to back up data. Tape backup provides low-cost scalability that is optimized for long-term retention.

Chapter 9. Roadmap for implementing a data protection solution

Plan and implement the most suitable data protection solution for your business environment with IBM Storage Protect.

Single-site disk solution

For steps that describe how to plan for, implement, monitor, and operate a single-site disk solution, see Single-site disk solution.

Multisite disk solution

For steps that describe how to plan for, implement, monitor, and operate a multisite disk solution, see Multisite disk solution.

Tape solution

For steps that describe how to plan for, implement, monitor, and operate a tape device solution, see Tape solution.

Multisite appliance solution

For an overview of the tasks that are required to implement a multisite appliance solution, review the following steps:

1. Begin planning for the solution by reviewing information at the following links:
   - AIX: Capacity planning
   - Linux: Capacity planning
   - Windows: Capacity planning
2. Install the server and optionally, the Operations Center. Review information at the following links:
   - Installing and upgrading the server
   - Installing and upgrading the
3. Configure the server for storage in a virtual tape library.
   - Managing virtual tape libraries
   - Attaching tape devices for the server

   For guidance about improving system performance, see Configuration best practices.
4. Configure policies to protect your data. Review the information in Customizing policies.
5. Set up client schedules. Review the information in Scheduling backup and archive operations.
6. Install and configure clients. To determine the type of client software that you need, review the information in Adding clients for details.
7. Configure monitoring for your system. Review the information in Monitoring storage solutions.

Related reference

Comparison of data protection solutions
Compare the key features for each IBM Storage Protect solution to determine which configuration best meets your data protection requirements. Then, review the available documentation to implement the solution.

Disk-based implementation of a data protection solution for a single site

This disk-based implementation of a data protection solution with IBM Storage Protect uses inline data deduplication and provides protection for data on a single site.

Disk-based implementation of a data protection solution for multiple sites
This disk-based implementation of a data protection solution with IBM Storage Protect uses inline data deduplication and replication at two sites.

Appliance-based implementation of a data protection solution for multiple sites
This implementation of a multi-site IBM Storage Protect data protection solution uses appliance-based data deduplication and replication. A standby server is configured at a second site to recover data if the primary server is unavailable.

Tape-based implementation of a data protection solution
This implementation of a data protection solution with IBM Storage Protect uses one or more tape storage devices to back up data. Tape backup provides low-cost scalability that is optimized for long-term retention.

Appendix A. Accessibility features for the IBM Storage Protect product family

Accessibility features assist users who have a disability, such as restricted mobility or limited vision, to use information technology content successfully.

Overview

The IBM Storage Protect family of products includes the following major accessibility features:

- Keyboard-only operation
- Operations that use a screen reader

The IBM Storage Protect family of products uses the latest W3C Standard, WAI-ARIA 1.0 (www.w3.org/TR/wai-aria/), to ensure compliance with US Section 508 and Web Content Accessibility Guidelines (WCAG) 2.0 (www.w3.org/TR/WCAG20/). To take advantage of accessibility features, use the latest release of your screen reader and the latest web browser that is supported by the product.

The product documentation in IBM Documentation is enabled for accessibility.

Keyboard navigation

This product uses standard navigation keys.

Interface information

User interfaces do not have content that flashes 2 - 55 times per second.

Web user interfaces rely on cascading style sheets to render content properly and to provide a usable experience. The application provides an equivalent way for low-vision users to use system display settings, including high-contrast mode. You can control font size by using the device or web browser settings.

Web user interfaces include WAI-ARIA navigational landmarks that you can use to quickly navigate to functional areas in the application.

Vendor software

The IBM Storage Protect product family includes certain vendor software that is not covered under the IBM license agreement. IBM makes no representation about the accessibility features of these products. Contact the vendor for accessibility information about its products.

Related accessibility information

In addition to standard IBM help desk and support websites, IBM has a TTY telephone service for use by deaf or hard of hearing customers to access sales and support services:

TTY service
800-IBM-3383 (800-426-3383)
(within North America)

For more information about the commitment that IBM has to accessibility, see IBM Accessibility (www.ibm.com/able).

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

For license inquiries regarding double-byte character set (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan

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 supply in any way it believes appropriate without incurring any obligation to you.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

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

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

The performance data discussed herein is presented as derived under specific operating conditions. Actual results may vary.

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.

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 the names and addresses used by an actual business enterprise 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.

Each copy or any portion of these sample programs or any derivative work must include a copyright notice as follows: © (your company name) (year). Portions of this code are derived from IBM Corp. Sample Programs. © Copyright IBM Corp. _enter the year or years_.

Trademarks

IBM, the IBM logo, and ibm.com® are trademarks or registered trademarks of International Business Machines Corp., 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 www.ibm.com/legal/copytrade.shtml.

Adobe is a registered trademark of Adobe Systems Incorporated in the United States, and/or other countries.

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

Intel and Itanium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

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

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

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

UNIX is a registered trademark of The Open Group in the United States and other countries.

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

Terms and conditions for product documentation

Permissions for the use of these publications are granted subject to the following terms and conditions.

Applicability

These terms and conditions are in addition to any terms of use for the IBM website.

Personal use

You may reproduce these publications for your personal, noncommercial use provided that all proprietary notices are preserved. You may not distribute, display or make derivative work of these publications, or any portion thereof, without the express consent of IBM.

Commercial use

You may reproduce, distribute and display these publications solely within your enterprise provided that all proprietary notices are preserved. You may not make derivative works of these publications, or reproduce, distribute or display these publications or any portion thereof outside your enterprise, without the express consent of IBM.

Rights

Except as expressly granted in this permission, no other permissions, licenses or rights are granted, either express or implied, to the publications or any information, data, software or other intellectual property contained therein.

IBM reserves the right to withdraw the permissions granted herein whenever, in its discretion, the use of the publications is detrimental to its interest or, as determined by IBM, the above instructions are not being properly followed.

You may not download, export or re-export this information except in full compliance with all applicable laws and regulations, including all United States export laws and regulations.

IBM MAKES NO GUARANTEE ABOUT THE CONTENT OF THESE PUBLICATIONS. THE PUBLICATIONS ARE PROVIDED "AS-IS" AND WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, NON-INFRINGEMENT, AND FITNESS FOR A PARTICULAR PURPOSE.

Privacy policy considerations

IBM Software products, including software as a service solutions, (“Software Offerings”) may use cookies or other technologies to collect product usage information, to help improve the end user experience, to tailor interactions with the end user, or for other purposes. In many cases no personally identifiable information is collected by the Software Offerings. Some of our Software Offerings can help enable you to collect personally identifiable information. If this Software Offering uses cookies to collect personally identifiable information, specific information about this offering’s use of cookies is set forth below.

This Software Offering does not use cookies or other technologies to collect personally identifiable information.

If the configurations deployed for this Software Offering provide you as customer the ability to collect personally identifiable information from end users via cookies and other technologies, you should seek your own legal advice about any laws applicable to such data collection, including any requirements for notice and consent.

For more information about the use of various technologies, including cookies, for these purposes, see IBM’s Privacy Policy at http://www.ibm.com/privacy and IBM’s Online Privacy Statement at http://www.ibm.com/privacy/details in the section entitled “Cookies, Web Beacons and Other Technologies,” and the “IBM Software Products and Software-as-a-Service Privacy Statement” at http://www.ibm.com/software/info/product-privacy.

Glossary

A glossary is available with terms and definitions for the IBM Storage Protect family of products.

See the IBM Storage Protect glossary.

Index

M

N

O

P

R

S

T

V

W

Product Number: 5725-W98
5725-W99
5725-X15