

Amazon Web Services - Storage Gateway

AWS Storage Gateway provides integration between the on-premises IT environment and the AWS storage infrastructure. The user can store data in the AWS cloud for scalable, data security features and cost-efficient storage.

AWS Gateway offers two types of storage, i.e. volume based and tape based.

AWS Gateway for type unknown

Volume Gateways

This storage type provides cloud-backed storage volumes which can be mount as Internet Small Computer System Interface (iSCSI) devices from on-premises application servers.

Gateway-cached Volumes

AWS Storage Gateway stores all the on-premises application data in a storage volume in Amazon S3. Its storage volume ranges from 1GB to 32 TB and up to 20 volumes with a total storage of 150TB. We can attach these volumes with iSCSI devices from on-premises application servers. It is of two categories –

Cache Storage Disk

Every application requires storage volumes to store their data. This storage type is used to initially store data when it is to be written to the storage volumes in AWS. The data from the cache storage disk is waiting to be uploaded to Amazon S3 from the upload buffer. The cache storage disk keeps the most recently accessed data for low-latency access. When the application needs data, the cache storage disk is first checked before checking Amazon S3.

There are few guidelines to determine the amount of disk space to be allocated for cache storage. We should allocate at least 20% of the existing file store size as cache storage. It should be more than the upload buffer.

Upload buffer disk – This type of storage disk is used to store the data before it is uploaded to Amazon S3 over SSL connection. The storage gateway uploads the data from the upload buffer over an SSL connection to AWS.

Snapshots – Sometimes we need to back up storage volumes in Amazon S3. These backups are incremental and are known as **snapshots**. The snapshots are stored in Amazon S3 as Amazon EBS snapshots. Incremental backup means that a new snapshot is backing up only the data that has changed since the last snapshot. We can take snapshots either at a scheduled interval or as per the requirement.

Gateway-stored Volumes

When the Virtual Machine (VM) is activated, gateway volumes are created and mapped to the on-premises direct-attached storage disks. Hence, when the applications write/read the data from the gateway storage volumes, it reads and writes the data from the mapped on-premises disk.

A gateway-stored volume allows to store primary data locally and provides on-premises applications with low-latency access to entire datasets. We can mount them as iSCSI devices to the on-premises application servers. It ranges from 1 GB to 16 TB in size and supports up to 12 volumes per gateway with a maximum storage of 192 TB.

Gateway-Virtual Tape Library (VTL)

This storage type provides a virtual tape infrastructure that scales seamlessly with your business needs and eliminates the operational burden of provisioning, scaling, and maintaining a physical tape infrastructure. Each gateway-VTL is preconfigured with media changer and tape drives, that are available with the existing client backup applications as iSCSI devices. Tape cartridges can be added later as required to archive the data.

Few terms used in Architecture are explained below.

Virtual Tape – Virtual tape is similar to a physical tape cartridge. It is stored in the AWS cloud. We can create virtual tapes in two ways: by using AWS Storage Gateway console or by using AWS Storage Gateway API. The size of each virtual tape is from 100 GB to 2.5 TB. The size of one gateway is up to 150 TB and can have maximum 1500 tapes at a time.

Virtual Tape Library (VTL) – Each gateway-VTL comes with one VTL. VTL is similar to a physical tape library available on-premises with tape drives. The gateway first stores data locally, then asynchronously uploads it to virtual tapes of VTL.

Tape Drive – A VTL tape drive is similar to a physical tape drive that can perform I/O operations on tape. Each VTL consists of 10 tape drives that are used for backup applications as iSCSI devices.

Media Changer – A VTL media changer is similar to a robot that moves tapes around in a physical tape library's storage slots and tape drives. Each VTL comes with one media changer that is used for backup applications as iSCSI device.

Virtual Tape Shelf (VTS) – A VTS is used to archive tapes from gateway VTL to VTS and vice-versa.

Archiving Tapes – When the backup software ejects a tape, the gateway moves the tape to the VTS for storage. It is used data archiving and backups.

Retrieving Tapes – Tapes archived to the VTS cannot be read directly so to read an archived tape, we need to retrieve the tape from gateway VTL either by using the AWS Storage Gateway console or by using the AWS Storage Gateway API.

Revision #1

Created 16 December 2022 10:58:07 by Admin

Updated 16 December 2022 10:58:39 by Admin