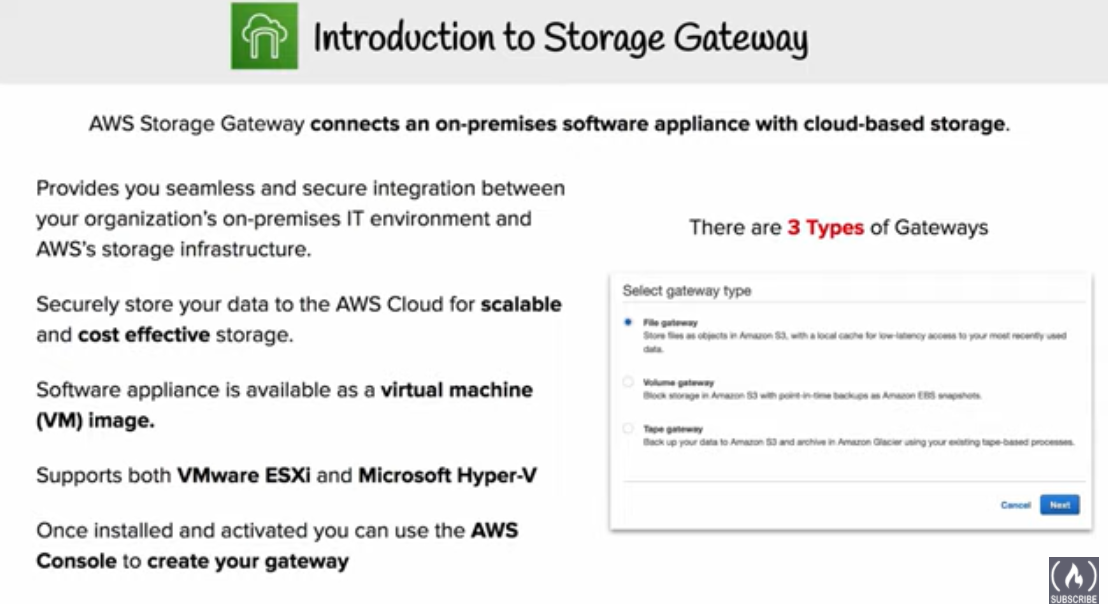
Storage Gateway Overview



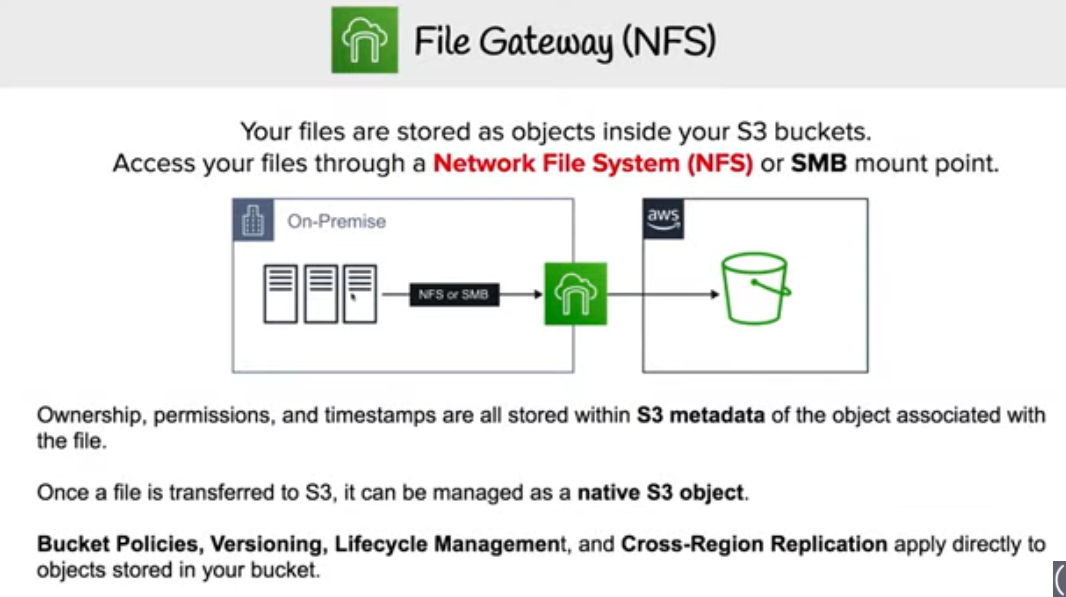


* AWS Storage Gateway connects on-premises software appliance with cloud-based storage to provide seamless integration with data security features between the on-premises IT environment and the AWS storage infrastructure.
* Storage Gateway allows storage of data in the AWS cloud for scalable and cost-effective storage while maintaining data security.
* Exposes compatible iSCSI interface on the front end to easily integrate with existing backup applications and represents an other disk drive
* AWS Storage Gateway backs up the data in Amazon Storage as incremental EBS snapshots
* AWS Storage Gateway can run either on-premises, as a virtual machine (VM) appliance, or in AWS, as an EC2 instance. So if the on-premises data center goes offline and there is no available host, the gateway can be deployed on an EC2 instance.
* Gateways hosted on EC2 instances can be used for disaster recovery, data mirroring, and providing storage for applications hosted on EC2
* **AWS Storage Gateway, by default, uploads data using SSL and provides data encryption at rest when stored in S3 or Glacier using AES-256**
* **AWS Storage Gateway performs compression of data in-transit and at-rest**

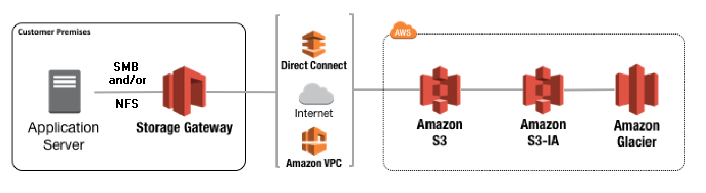
## Storage Gateway Types

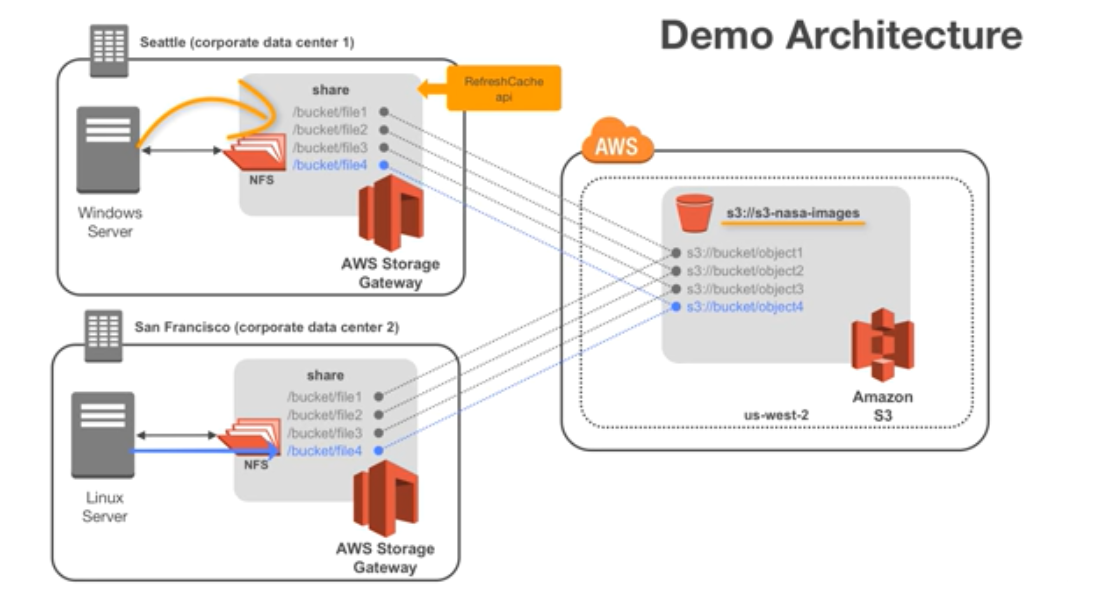
AWS Storage Gateway offers file-based, volume-based, and tape-based storage solutions

### File Gateways



* supports a file interface into S3 and combines a service and a virtual software appliance.
* allows storing and retrieving of objects in S3 using industry-standard file protocols such as **Network File System (NFS)** and **Server Message Block (SMB).**
* Software appliance, or gateway, is deployed into on-premises environment as a **virtual machine (VM) running on VMware ESXi** or **Microsoft Hyper-V hypervisor.**
* **provides access to objects in S3 as files or file share mount points. It can be considered as a file system mount on S3.**
* **provides a cost-effective alternative to on-premises storage.**
* **provides low-latency access to data through transparent local caching.**
* manages data transfer to and from AWS, buffers applications from network congestion, optimizes and streams data in parallel, and manages bandwidth consumption.
* **easily integrates with AWS services like IAM, KMS, CloudWatch etc.**
* File Gateway allows you to
  + store and retrieve files directly using the NFS version 3 or 4.1 protocol.
  + store and retrieve files directly using the SMB file system version, 2 and 3 protocol.
  + access the data directly in S3 from any AWS Cloud application or service.
  + manage S3 data using lifecycle policies, cross-region replication, and versioning.





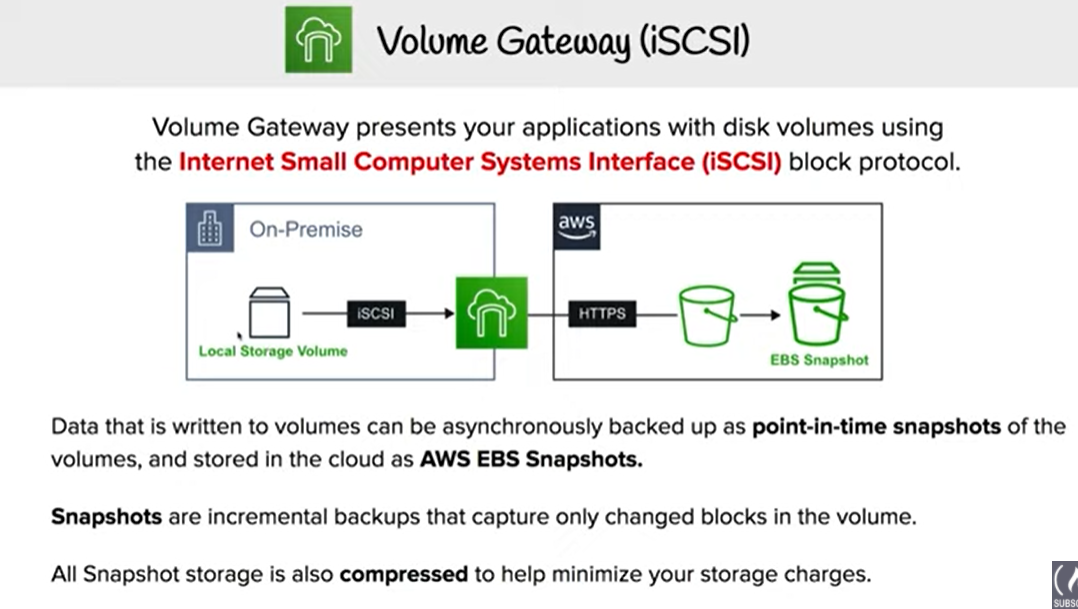
Graphical user interface, application

Description automatically generated

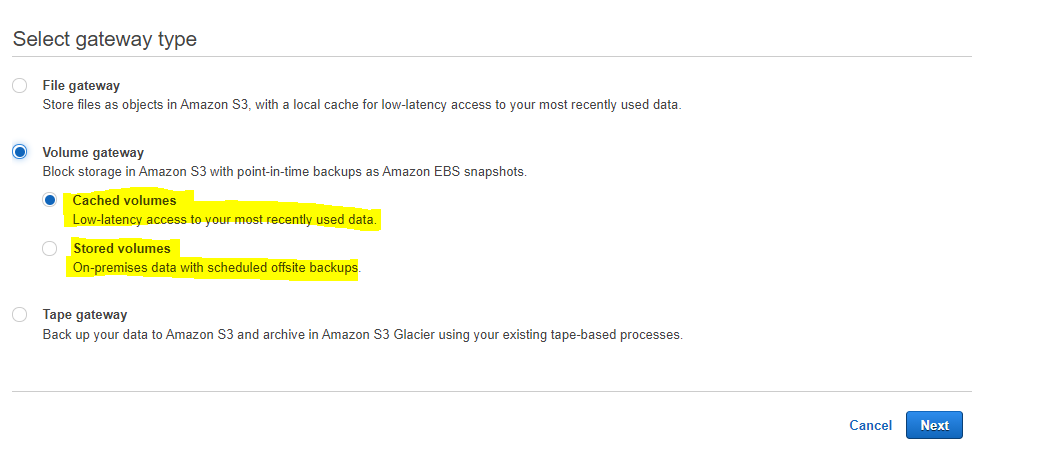
Diagram

Description automatically generated

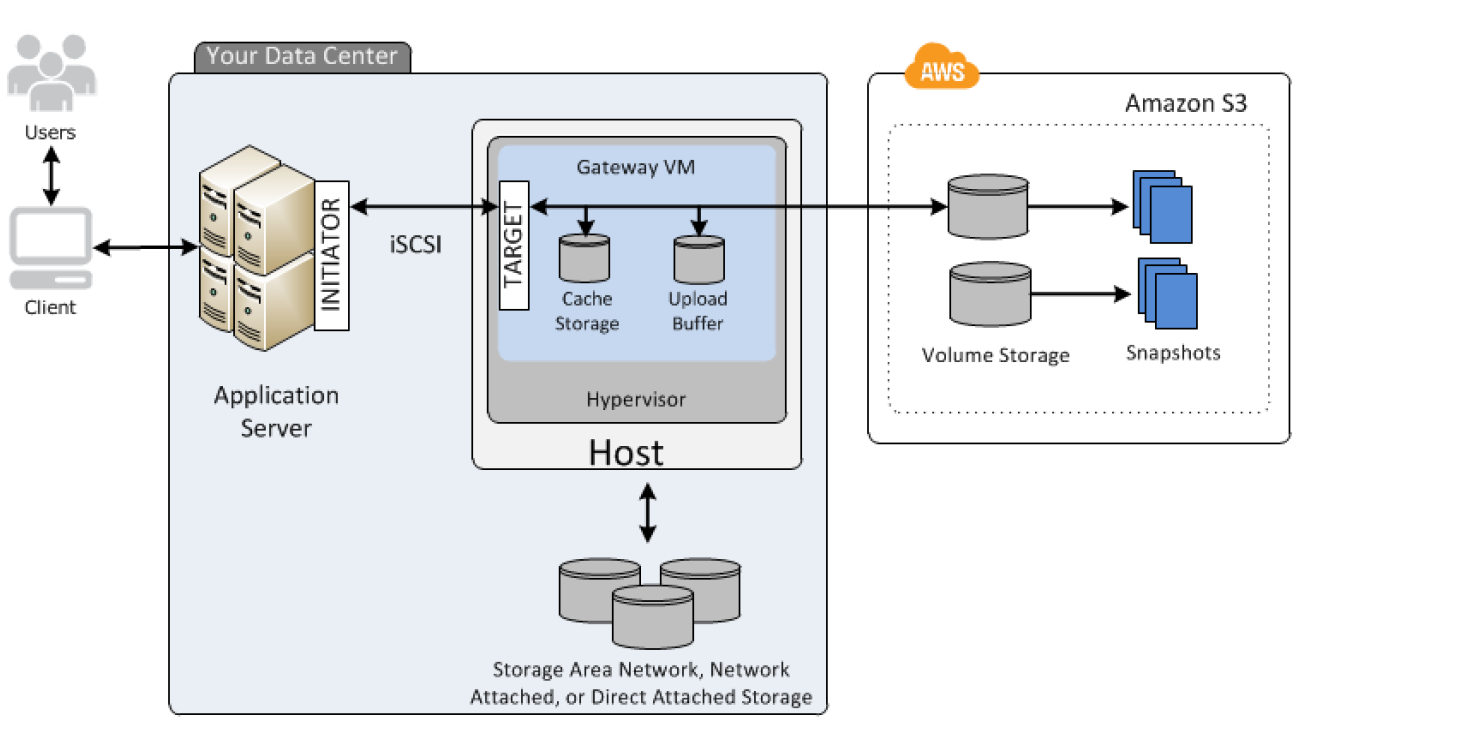
### Volume gateways

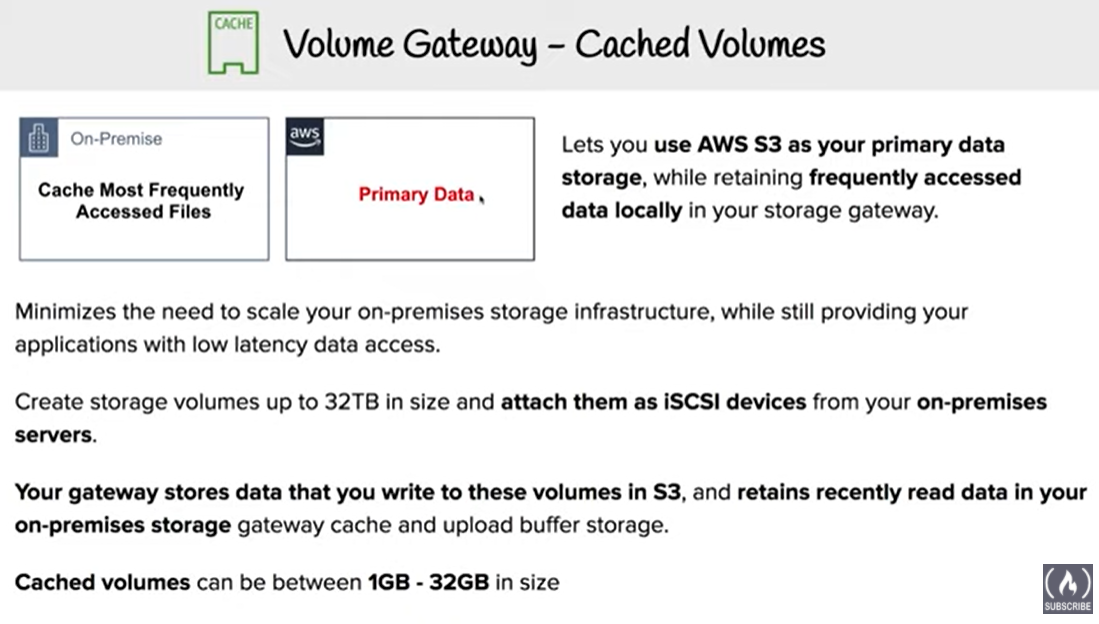


Volume gateways provide cloud-backed storage volumes that you can mount as Internet Small Computer System Interface (iSCSI) devices from your on-premises application servers. For Volume gateways all data is securely stored in AWS, the approach differs with how much data is stored on-premises



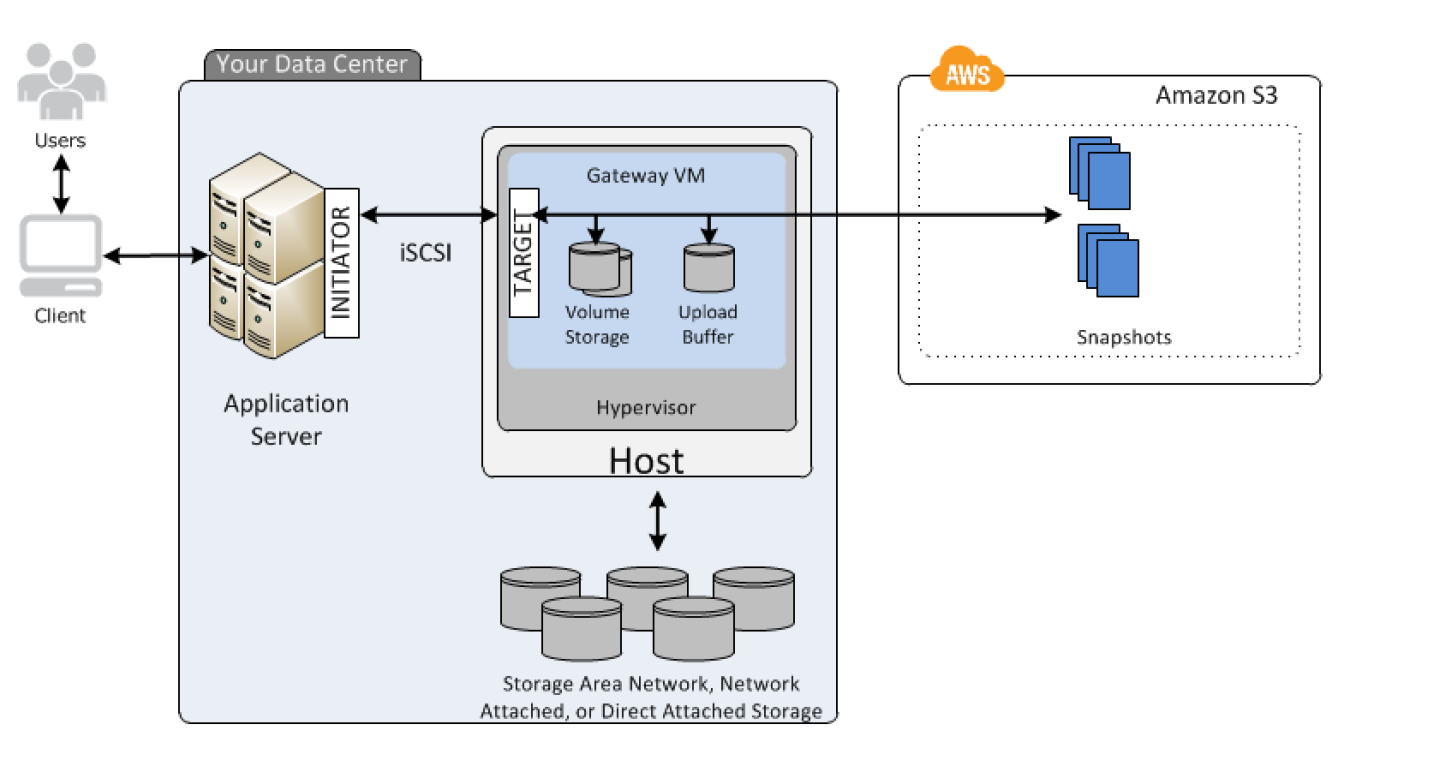
#### **Gateway-cached volumes**

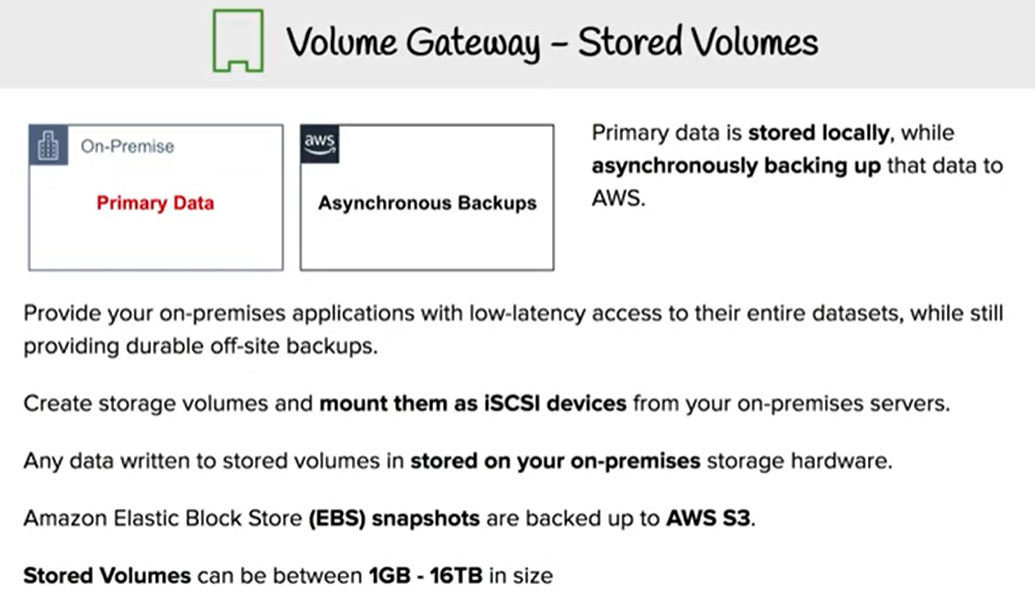




* Data is stored in S3 and acts as a Primary data storage
* Gateway retains a copy of recently read data locally for low latency access to the frequently accessed data
* Gateway-cached volumes offer a substantial cost savings on primary storage and minimize the need to scale your storage on-premises.
* Each gateway configured for gateway-cached volumes can support up to 32 volumes, with each volume ranging from 1GiB to 32TiB, for a total maximum storage volume of 1,024 TiB (1 PiB).
* Gateway-cached volumes can be attached as iSCSI devices from on-premises application servers
* Gateway-cached volumes can be backed up incrementally by taking snapshots which are stored as EBS snapshots in Amazon S3. These snapshots can be restored as gateway storage volume or used to create EBS volumes (if < 16TiB) to attached to an EC2 instance
* All gateway-cached volume data and snapshot data is stored in Amazon S3 encrypted at rest using server-side encryption (SSE) and it cannot be accessed with S3 API or any other tools
* Gateway VM can be allocated disks
  + Cache storage
    - Cache storage, acts as the on-premises durable storage, stores the data before uploading it to Amazon S3
    - Cache storage also stores recently read data for low-latency access
  + Upload buffer
    - Upload buffer acts as a staging area, before the data is uploaded to S3
    - Gateway uploads data over an encrypted Secure Sockets Layer (SSL) connection to AWS, where it is stored encrypted in Amazon S3

#### **Gateway-stored volumes**

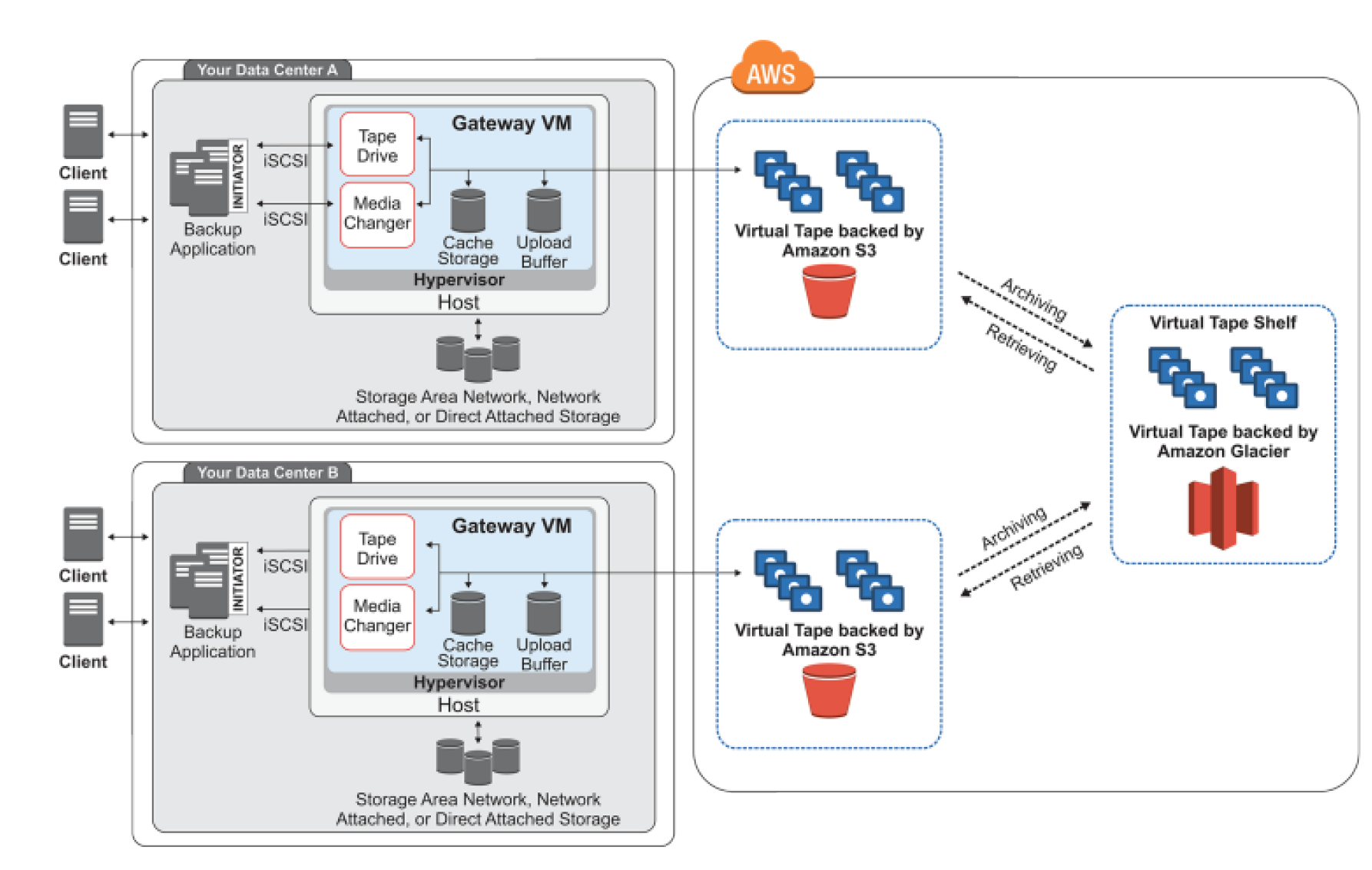


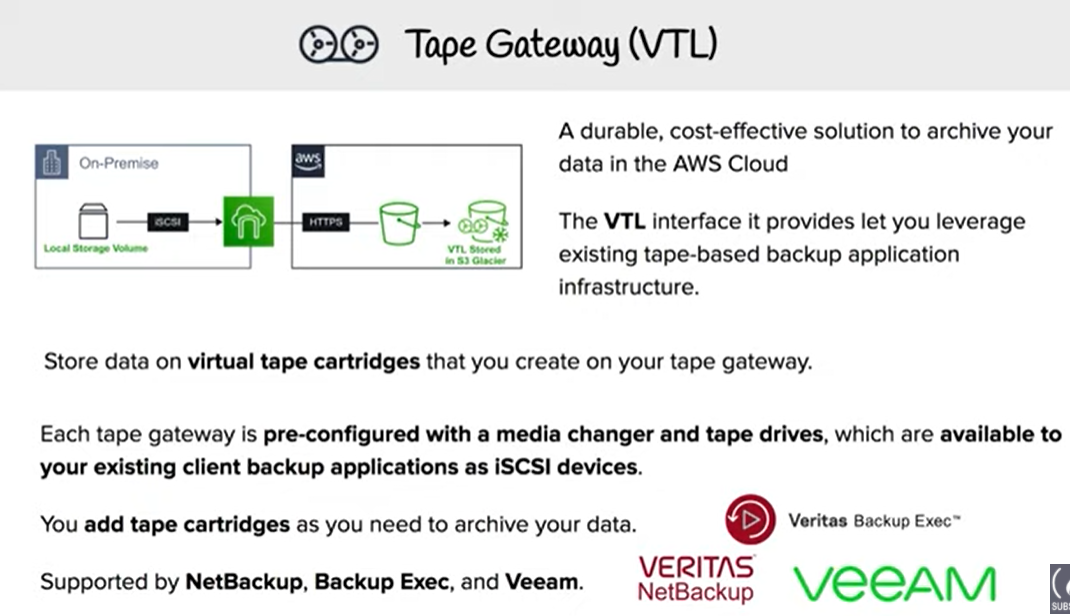


* Gateway-stored volumes maintain the entire data set locally to provide low latency access
* Gateway asynchronously backs up point-in-time snapshots (in the form of EBS snapshots) of the data to S3 which provides durable off-site backups
* Gateway-stored volumes can be attached as iSCSI devices from on-premises application servers
* Each gateway configured for gateway-stored volumes can support up to  32 volumes, ranging from 1GiB to 16TiB, and a total volume storage of 192 TiB
* Gateway-stored volume configuration provides durable and inexpensive off-site backups that you can recover to your local data center or Amazon EC2. For example, if you need replacement capacity for disaster recovery, you can recover the backups to Amazon EC2.
* Gateway-stored volumes can be backed up incrementally by taking snapshots which are stored as EBS snapshots in Amazon S3. These snapshots can be restored as gateway storage volume or used to create EBS volumes (if < 16TiB) to attached to an EC2 instance
* Gateway VM can be allocate disks
  + Volume Storage
    - For storing the actual data
    - Can be mapped to on-premises direct-attached storage (DAS) or storage area network (SAN) disks
  + Upload buffer
    - Upload buffer acts as a staging area, before the data is uploaded to S3
    - Gateway uploads data over an encrypted Secure Sockets Layer (SSL) connection to AWS, where it is stored encrypted in Amazon S3

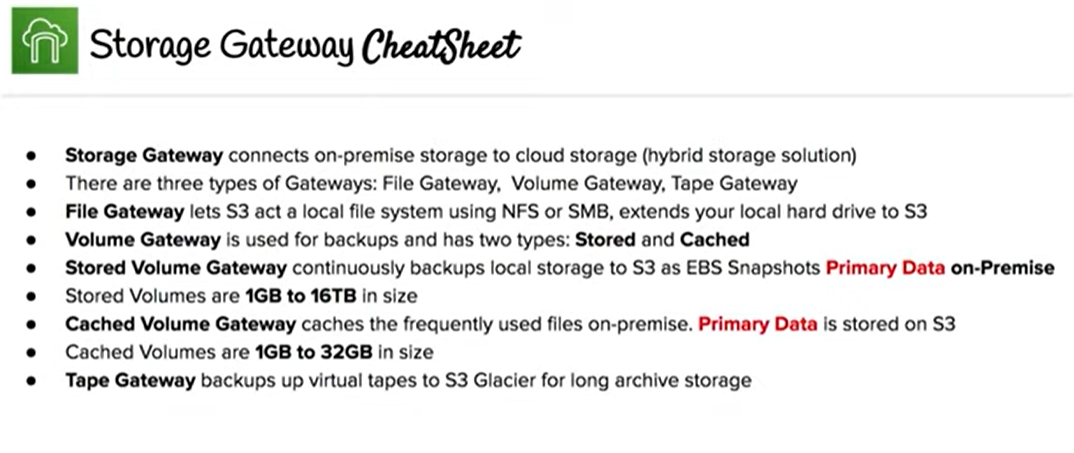
### Tape-based storage solutions

#### **Gateway–virtual tape library (VTL)**





* Gateway-VTL provides cost-effective and durable archival of backup data in Amazon Glacier.
* Gateway-VTL provides a virtual tape infrastructure that scales seamlessly with the business needs and eliminates the operational burden of provisioning, scaling, and maintaining a physical tape infrastructure.
* VTL interface lets you leverage your existing tape-based backup application infrastructure to store data on virtual tape cartridges that you create on your gateway-VTL.
* Each gateway-VTL is preconfigured with a media changer and tape drives, which are available to the existing client backup applications as iSCSI devices. Tape cartridges can be added as needed to archive your data.
* Gateway VTL has the following components :-
  + Virtual tape
    - Virtual tape are similar to the physical tape cartridge, except that the data is stored in the AWS storage solution
    - Each gateway can contain 1500 tapes or up to 150 TiB of total tape data, with each tape ranging from 100 GiB to 2.5 TiB
  + **Virtual tape library**
    - Virtual tape library is similar to the physical tape library with tape drives (replaced with VTL tape drive) and robotic arms (replaced with Media changer)
    - Tapes in the Virtual tape library are backup in Amazon **S3**
    - Backup software writes data to the gateway, the gateway stores data locally and then asynchronously uploads it to virtual tapes in Amazon S3.
  + **Virtual tape shelf**
    - Virtual tape shelf is similar to the offsite tape holding facility
    - Tapes in the Virtual tape library are backup in Amazon **Glacier**providing an extremely low-cost storage service for data archiving and backup
    - VTS is located in the same region where the gateway was created and every region would have a single VTS irrespective of the number of gateways
    - Archiving tapes
      * When the backup software ejects a tape, the gateway moves the tape to the VTS for long term storage
    - Retrieving tapes
      * Tape can be retrieved from VTS only by first retrieving the tapes first to VTL and would be available in the VTL in about 24 hours
* Gateway VM can be allocate disks
  + Cache storage
    - Cache storage, acts as the on-premises durable storage, stores the data before uploading it to Amazon S3
    - Cache storage also stores recently read data for low-latency access
  + Upload buffer
    - Upload buffer acts as a staging area, before the data is uploaded to the Virtual tape
    - Gateway uploads data over an encrypted Secure Sockets Layer (SSL) connection to AWS, where it is stored encrypted in Amazon S3



## AWS Certification Exam Practice Questions

1. Which of the following services natively encrypts data at rest within an AWS region?Choose 2 answers
   1. **AWS Storage Gateway**
   2. Amazon DynamoDB
   3. Amazon CloudFront
   4. **Amazon Glacier**
   5. Amazon Simple Queue Service
2. What does the AWS Storage Gateway provide?
   1. **It allows to integrate on-premises IT environments with Cloud Storage**
   2. A direct encrypted connection to Amazon S3.
   3. It’s a backup solution that provides an on-premises Cloud storage.
   4. It provides an encrypted SSL endpoint for backups in the Cloud.
3. You’re running an application on-premises due to its dependency on non-x86 hardware and want to use AWS for data backup. Your backup application is only able to write to POSIX-compatible block-based storage. You have 140TB of data and would like to mount it as a single folder on your file server. Users must be able to access portions of this data while the backups are taking place. What backup solution would be most appropriate for this use case?
   1. Use Storage Gateway and configure it to use Gateway Cached volumes.
   2. Configure your backup software to use S3 as the target for your data backups.
   3. Configure your backup software to use Glacier as the target for your data backups
   4. **Use Storage Gateway and configure it to use Gateway Stored volumes** (Data is hosted on the On-premise server as well. The requirement for 140TB is for file server On-Premise more to confuse and not in AWS. Just need a backup solution hence stored instead of cached volumes)
4. A customer has a single 3-TB volume on-premises that is used to hold a large repository of images and print layout files. This repository is growing at 500 GB a year and must be presented as a single logical volume. The customer is becoming increasingly constrained with their local storage capacity and wants an off-site backup of this data, while maintaining low-latency access to their frequently accessed data. Which AWS Storage Gateway configuration meets the customer requirements?
   1. **Gateway-Cached volumes with snapshots scheduled to Amazon S3**
   2. Gateway-Stored volumes with snapshots scheduled to Amazon S3
   3. Gateway-Virtual Tape Library with snapshots to Amazon S3
   4. Gateway-Virtual Tape Library with snapshots to Amazon Glacier
5. You have a proprietary data store on-premises that must be backed up daily by dumping the data store contents to a single compressed 50GB file and sending the file to AWS. Your SLAs state that any dump file backed up within the past 7 days can be retrieved within 2 hours. Your compliance department has stated that all data must be held indefinitely. The time required to restore the data store from a backup is approximately 1 hour. Your on-premise network connection is capable of sustaining 1gbps to AWS. Which backup methods to AWS would be most cost-effective while still meeting all of your requirements?
   1. Send the daily backup files to Glacier immediately after being generated (will not meet the RTO)
   2. Transfer the daily backup files to an EBS volume in AWS and take daily snapshots of the volume (Not cost effective)
   3. Transfer the daily backup files to S3 and use appropriate bucket lifecycle policies to send to Glacier (Store in S3 for seven days and then archive to Glacier)
   4. Host the backup files on a Storage Gateway with Gateway-Cached Volumes and take daily snapshots (Not Cost effective as local storage as well as S3 storage)
6. A customer implemented AWS Storage Gateway with a gateway-cached volume at their main office. An event takes the link between the main and branch office offline. Which methods will enable the branch office to access their data? Choose 3 answers
   1. Use a HTTPS GET to the Amazon S3 bucket where the files are located (gateway volumes are only accessible from the AWS Storage Gateway and cannot be directly accessed using Amazon S3 APIs)
   2. Restore by implementing a lifecycle policy on the Amazon S3 bucket.
   3. Make an Amazon Glacier Restore API call to load the files into another Amazon S3 bucket within four to six hours.
   4. **Launch a new AWS Storage Gateway instance AMI in Amazon EC2, and restore from a gateway snapshot**
   5. **Create an Amazon EBS volume from a gateway snapshot, and mount it to an Amazon EC2 instance.**
   6. **Launch an AWS Storage Gateway virtual iSCSI device at the branch office, and restore from a gateway snapshot**