



AWS Storage Gateway

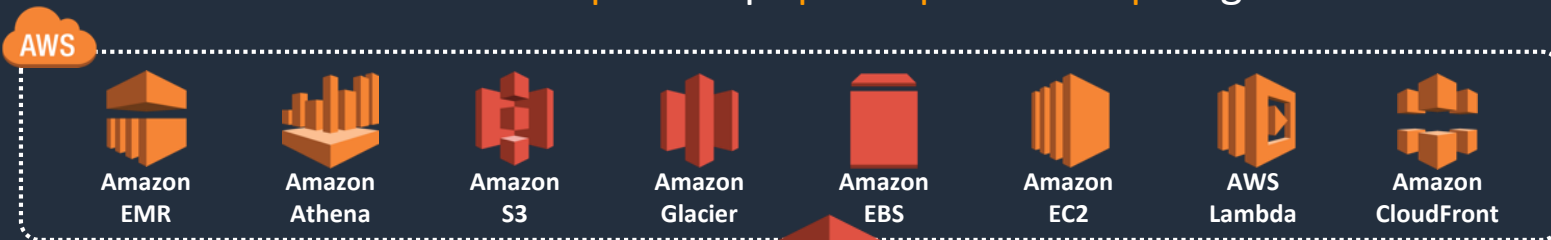
Fred Hutch - Partly Cloudy Conference

Jeff Bartley | Storage Solutions Architect | jeffbart@amazon.com

October 25th, 2018

Enabling a spectrum of hybrid use cases

Analytics | File Services | Production Tiering | Data Processing
Data Distribution | Backup | DR | Archive | Migration



AWS Storage Gateway



Enterprise Data Centers

Remote Offices

Research Sites

Small-Medium Businesses



Multimedia content



File servers



Databases



App. servers



Devices



Users



Backup servers

AWS Storage Gateway Family

Hybrid storage service enabling applications to seamlessly use AWS storage



File Gateway

Store and access objects in Amazon S3 from file-based applications with local caching



Volume Gateway

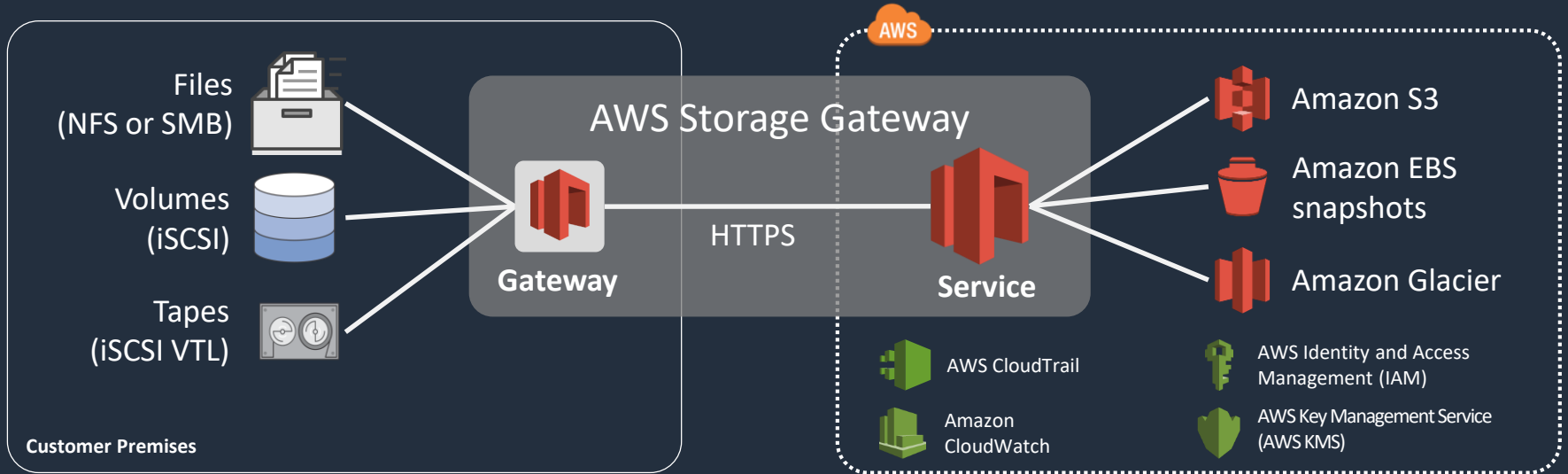
Block storage on-premises backed by cloud storage with local caching, EBS snapshots, and clones



Tape Gateway

Drop-in replacement for physical tape infrastructure backed by cloud storage with local caching

AWS Storage Gateway Family



Gateway provides applications

- Protocol conversion and device emulation
- Caching (read-through / write-back)
- Optimized data transfer

Native storage in AWS

- Objects in S3 (file)
- Snapshots in EBS (volume)
- Archives in Glacier (tape)

File Gateway

Store and access objects in Amazon S3 from file-based applications with local caching



Reduce on-premises storage infrastructure

Fully managed local cache for low latency access

Durability, scalability, and reliability of Amazon S3 storage

File Gateway support for SMB



Store and access objects in Amazon S3 buckets
from file-based Windows applications

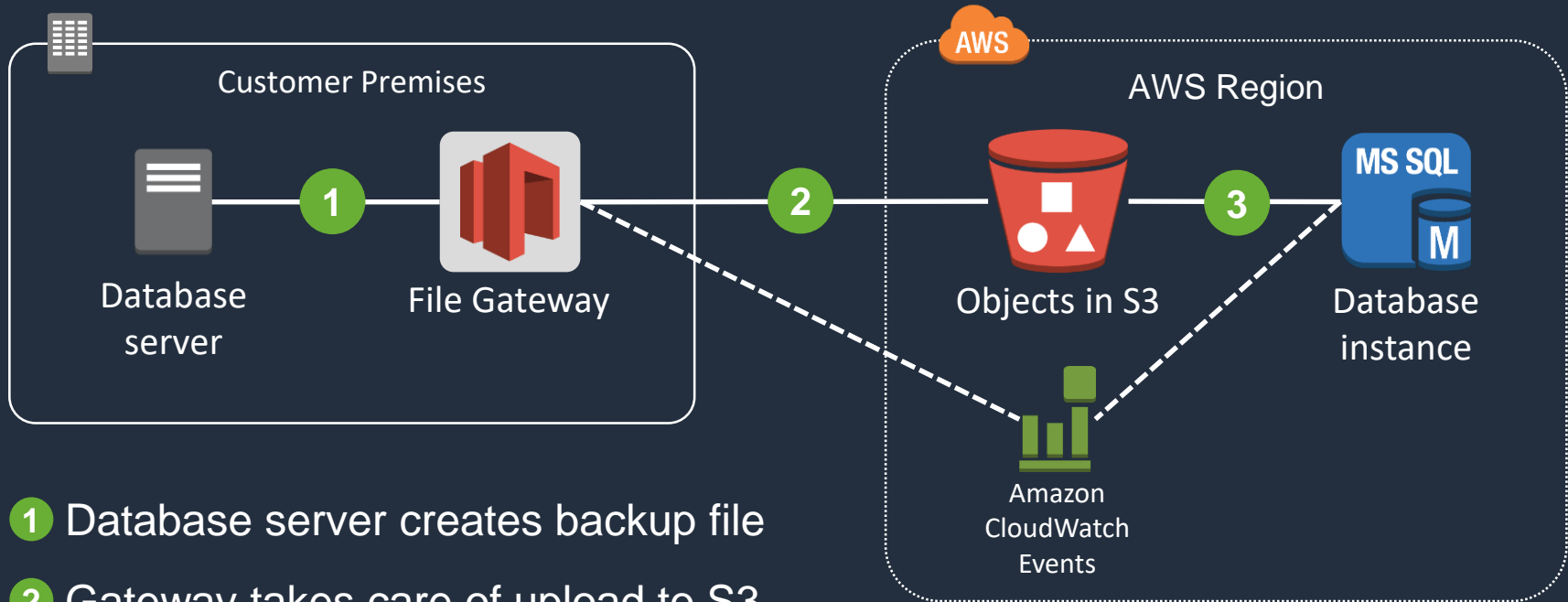
On-premises Active
Directory (AD) or AWS
Directory services

Share-level access can
be restricted to
configured users and
groups

POSIX ACLs
(compatible subset of
NTFS) for object-level
permissions

Objects created
directly in S3 inherit
ACLs from parent
folder

Hybrid File Use Case – Backup to AWS



- 1 Database server creates backup file
- 2 Gateway takes care of upload to S3
- 3 Automate restore or validation in-cloud or lifecycle to Glacier for archival

Hybrid File Use Case – Content Distribution



- 1 Application in San Diego writes files which are uploaded to S3 by gateway
- 2 After refresh cache, files are visible to applications in Sydney. Local cache improves access performance.

Hybrid File Use Case – Active Archive



- 1 Use Snowball to ship data from on-premises offline archives
- 2 Online access to all data through gateway, reduces on-premises storage cost, S3 lifecycle policies to lower costs further

Enabling Cloud Workloads

Easily migrate existing applications to the cloud

“We have applications that don’t support S3 natively, and **we can’t make changes to the applications**, so being able to use native protocols is very important to us.

Storage Gateway allows Celgene to get data to the cloud without making any changes to the application or environment.”

– Lance Smith, IT Director, Celgene



File Gateway Deep Dive

File Gateway – Control Over Data Storage and Access



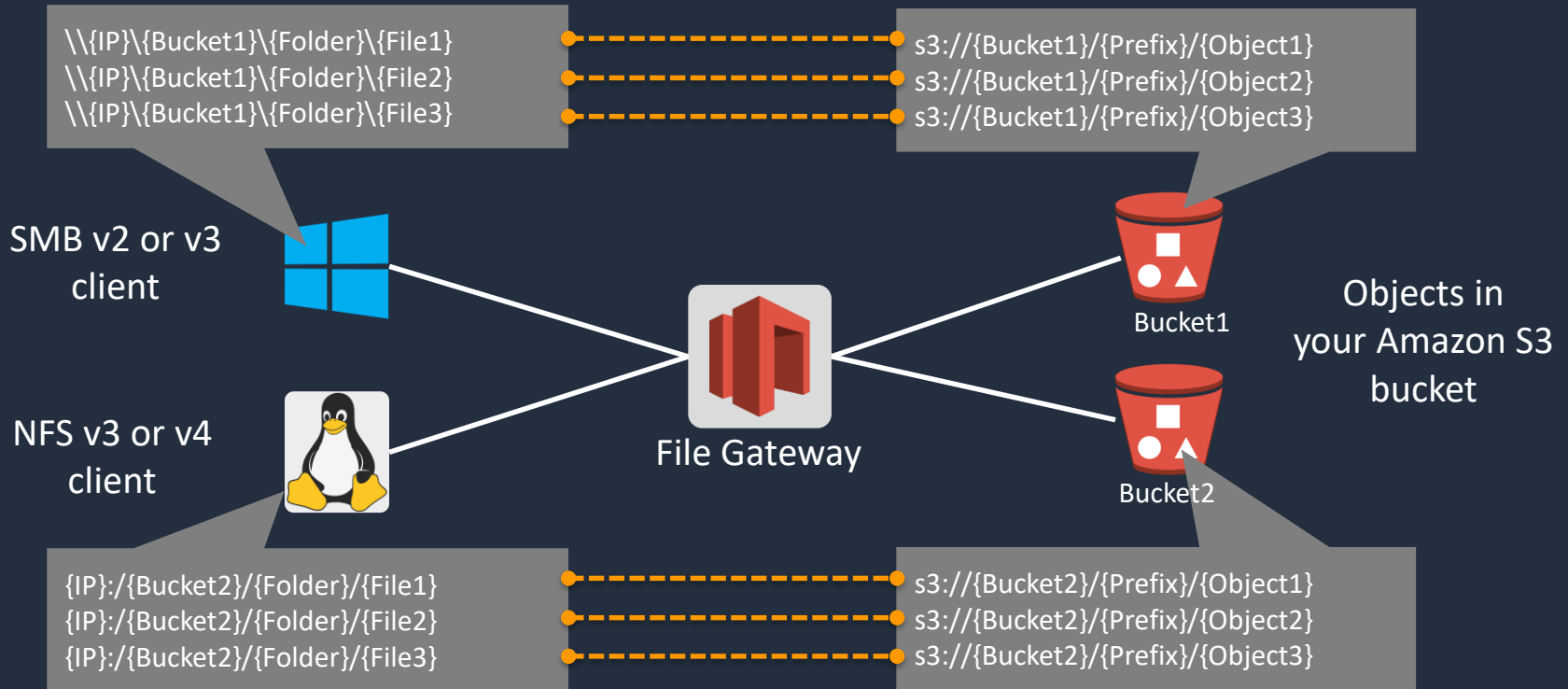
File share options

- Restrict access by IP (NFS) or AD (SMB)
- Read-only/read-write
- Default ownership and permissions
- User squashing (NFS)

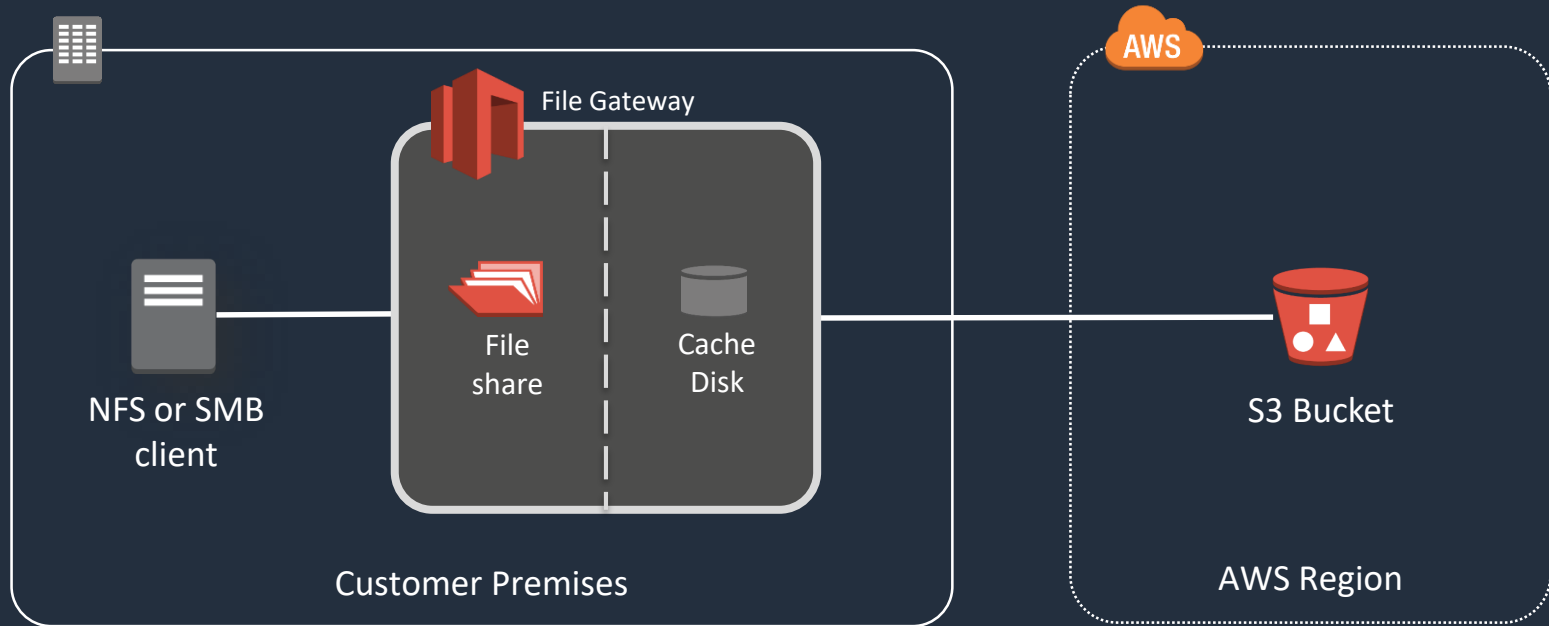
Amazon S3 options per bucket

- IAM role
- Storage class
- Encryption with AWS KMS
- Guess MIME type, requester pays, bucket owner ACL, ...

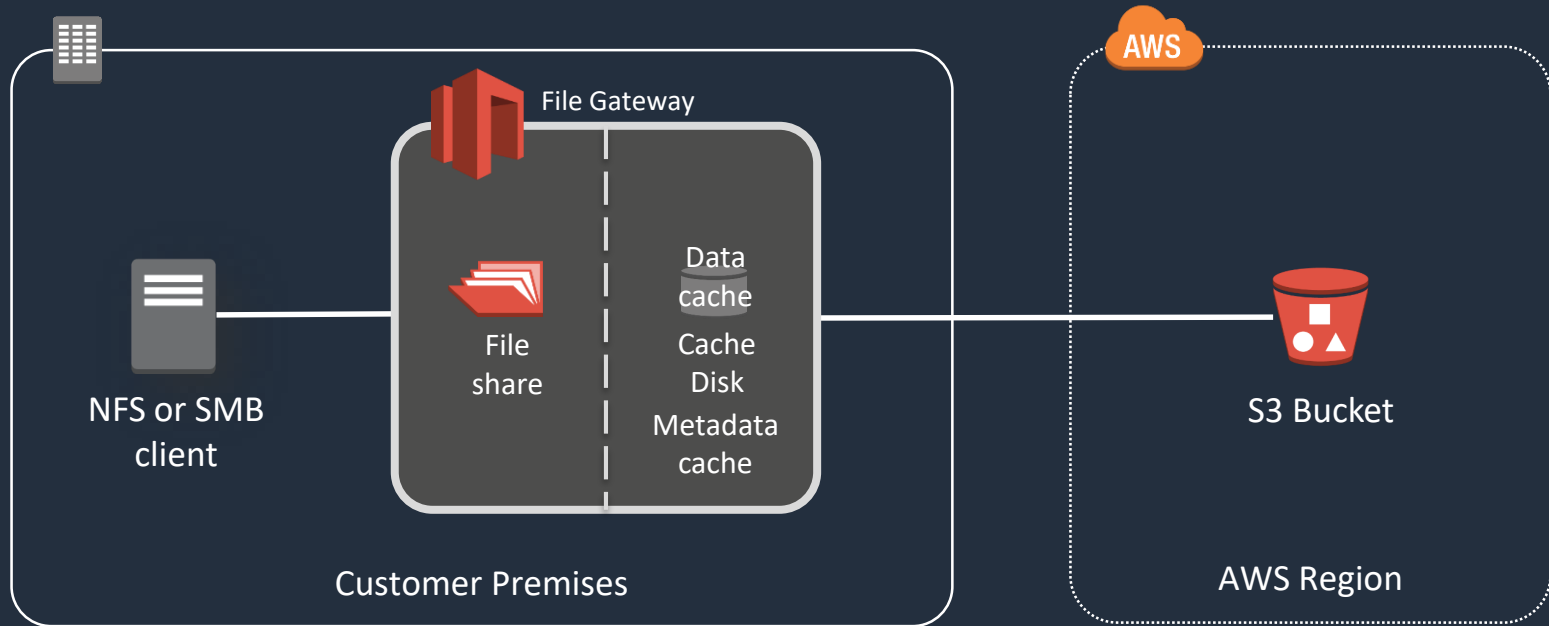
File Gateway – Mapping Files to Objects



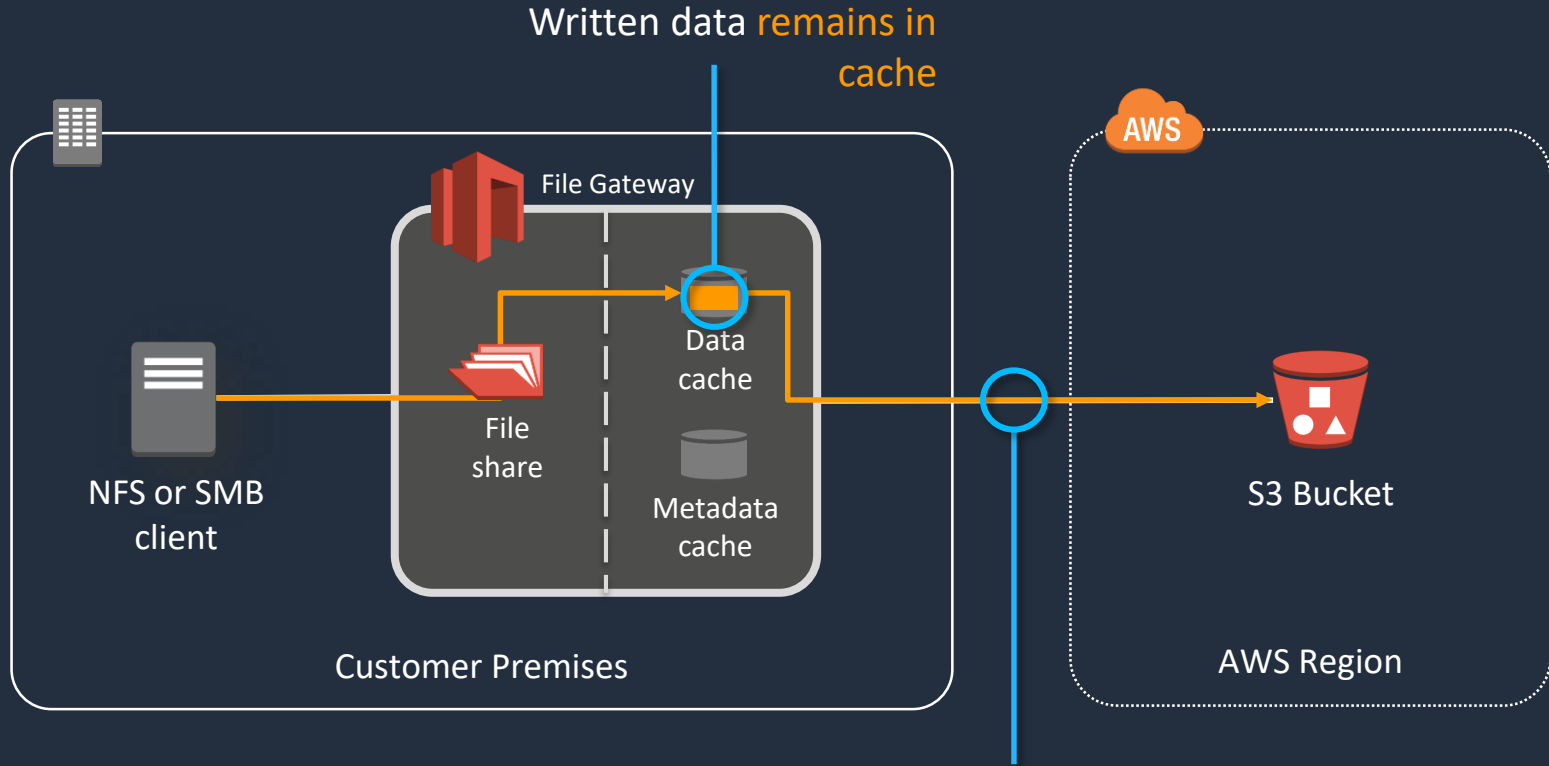
File Gateway – Storage Architecture



File Gateway – Storage Architecture

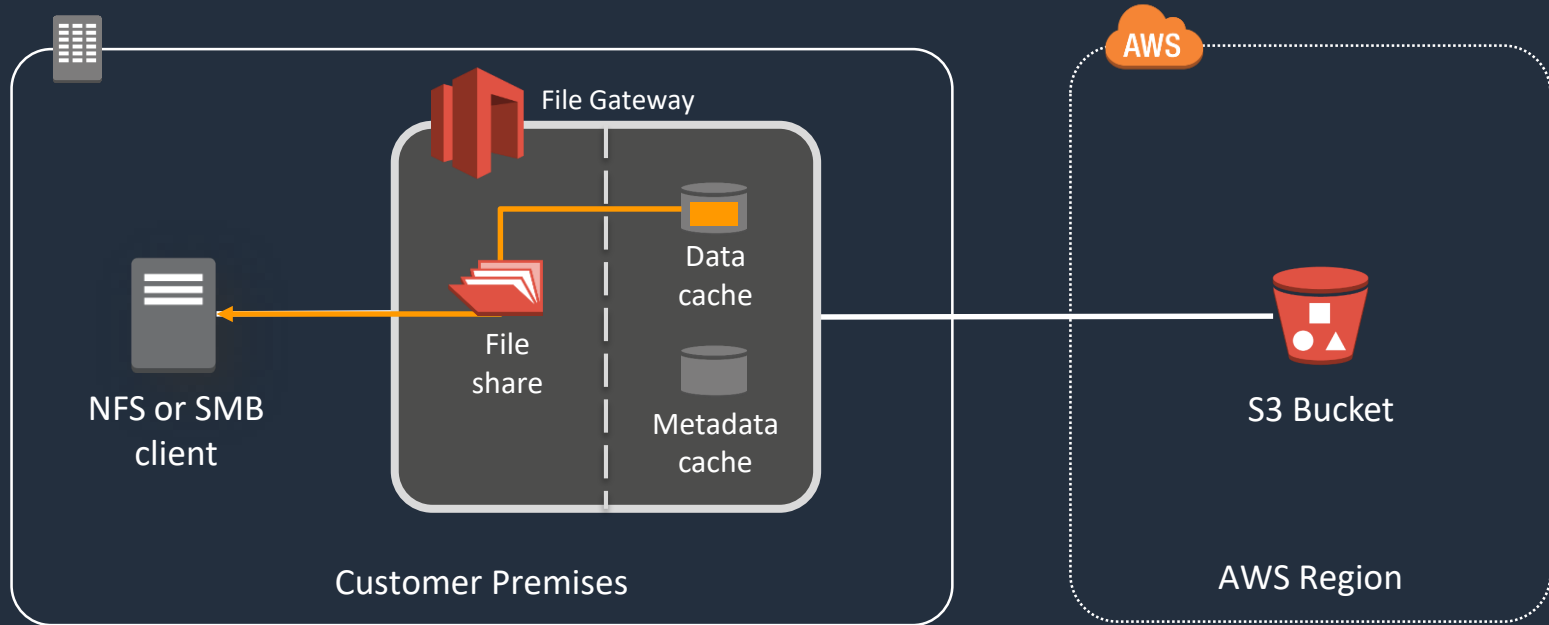


File Gateway – Write

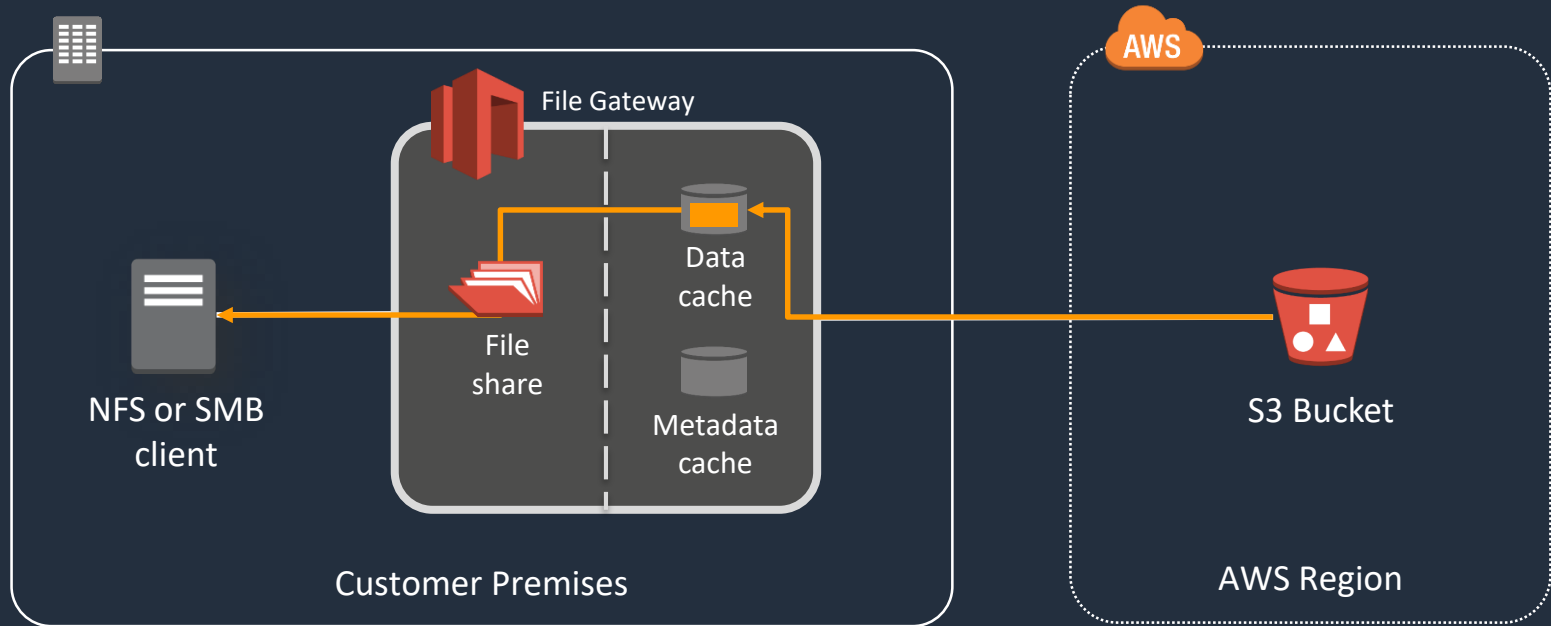


Writes to S3 are **optimized** and **encrypted**

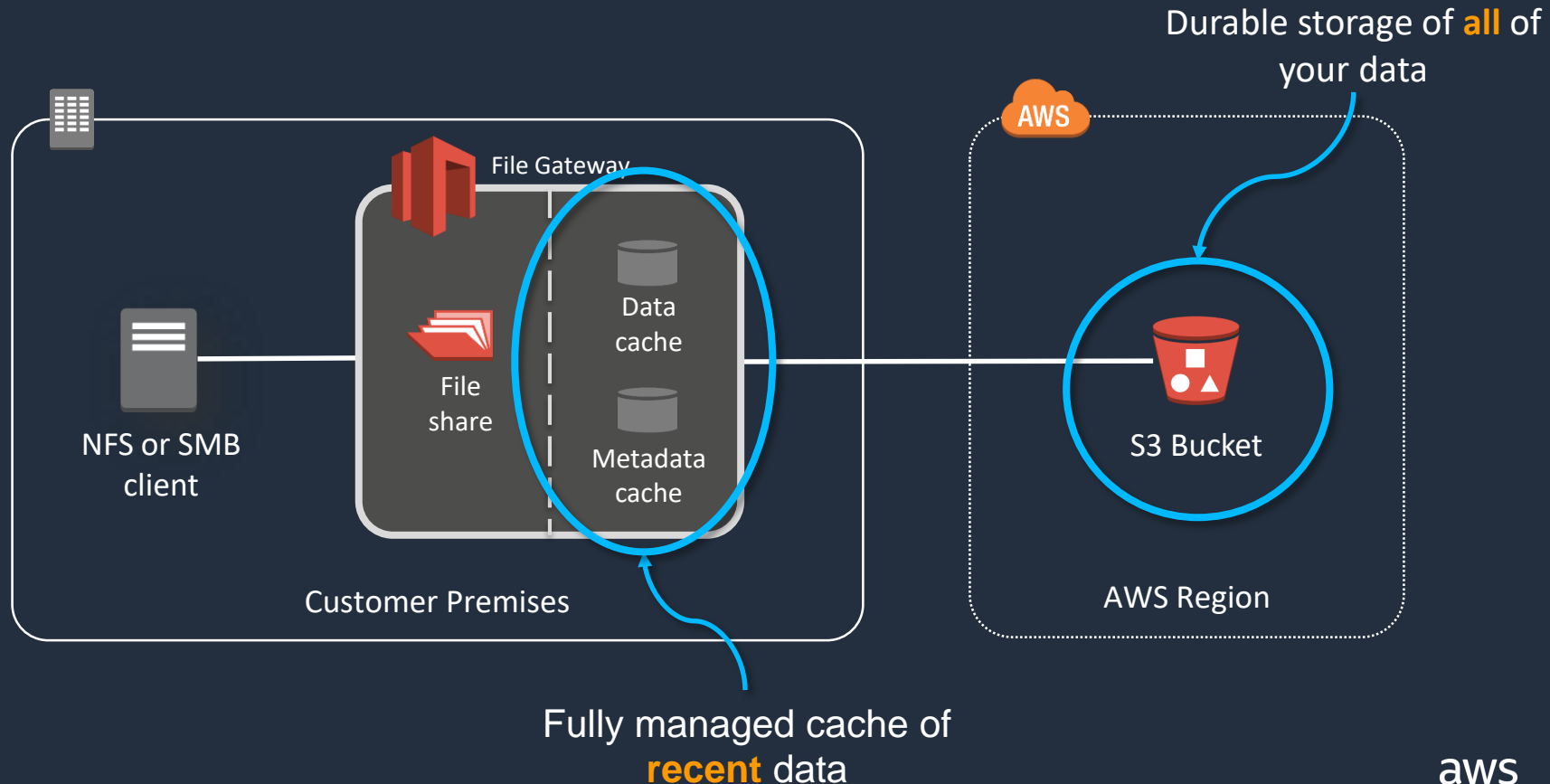
File Gateway – Read from Cache



File Gateway – Read from S3

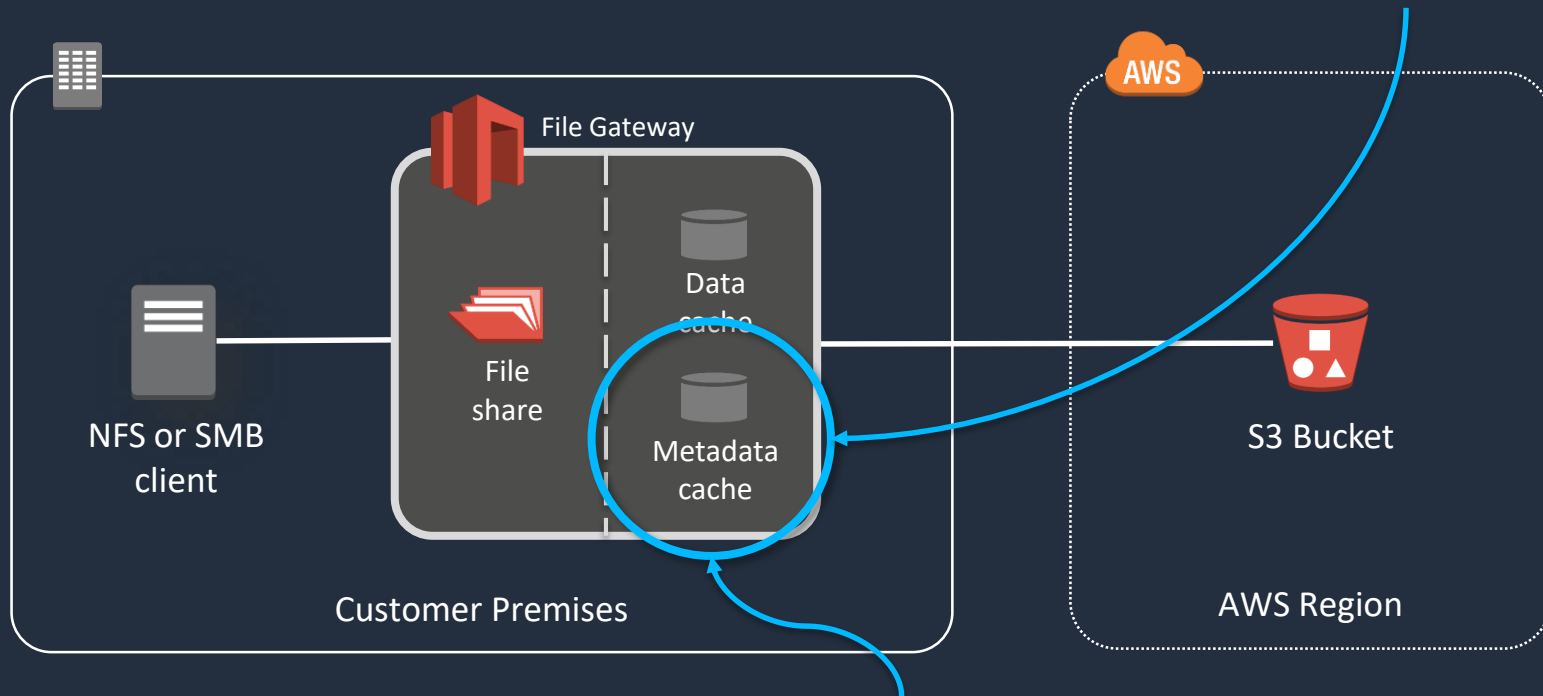


File Gateway – Low Latency Access to Active Data



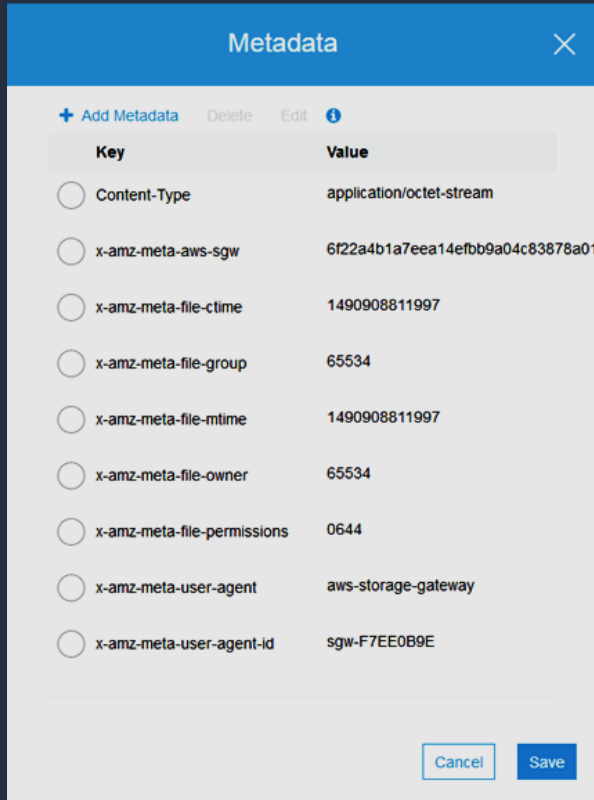
File Gateway – Metadata Cache

Bucket can have **billions of objects** in it, gateway discovers them on-demand



Metadata is **cached based on operations** performed by the file client

File Gateway - File System Metadata



Key	Value
<input type="radio"/> Content-Type	application/octet-stream
<input type="radio"/> x-amz-meta-aws-sgw	6f22a4b1a7eea14efbb9a04c83878a01
<input type="radio"/> x-amz-meta-file-ctime	1490908811997
<input type="radio"/> x-amz-meta-file-group	65534
<input type="radio"/> x-amz-meta-file-mtime	1490908811997
<input type="radio"/> x-amz-meta-file-owner	65534
<input type="radio"/> x-amz-meta-file-permissions	0644
<input type="radio"/> x-amz-meta-user-agent	aws-storage-gateway
<input type="radio"/> x-amz-meta-user-agent-id	sgw-F7EE0B9E

- File system metadata (permissions, ownership, ctime, etc.) **persisted in object user-metadata**
- **Configurable defaults** for objects that don't have this metadata (i.e. objects that were already in the bucket)
- Changing file metadata results in a copy-put operation

File Gateway - File system operations

directories

- Folder objects created in S3 using same pattern as S3 Console
- Walking the file system can be expensive (ls -R, find, etc).
- Caching of metadata locally reduces latency of directory level operations (such as “ls”).

delete

- Removes the object in S3

rename

- Atomic for clients connected to the same gateway
- Copy-put request on S3, so eventually consistent for S3 clients

links

- Hard links and symbolic links are not supported
- No analog in S3, so returns “Operation not supported”

sparse files

- Object is zero-filled for sparse ranges

Deployment

vmware®



On-Premises virtual environments

Run as virtual machine on VMware ESXi and Microsoft Hyper-V



Hardware Appliance

Pre-installed with Storage Gateway software. Ideal for branch offices, warehouses, and “outpost” offices that lack dedicated IT resources.



Amazon EC2

Run Storage Gateway in your AWS VPC environment

Deployment Best Practices

VM environments

- Four virtual processors assigned to the VM.
- 16 GiB of reserved RAM assigned to the VM.
- 80 GiB of disk space for installation of VM image and system data.

Cache Disk

- One or more disks - 150 GB to 16 TB total cache
- Use redundant storage

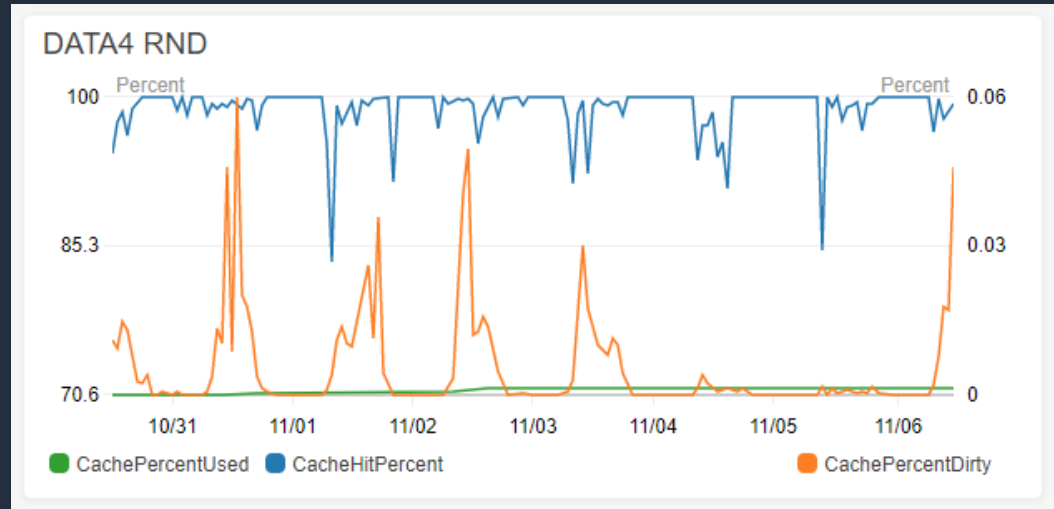
Amazon CloudWatch monitoring/alerting

Metrics

- Cache metrics
- S3 read/write bytes
- Share read/write bytes

Events

- Notify on file upload
- Cache refresh



To Learn More...

aws.amazon.com/storagegateway/

Menu [Contact Sales](#) [Products](#) [Solutions](#) [Pricing](#) [Getting Started](#) [Documentation](#) [More](#) [English](#) [My Account](#) [Create an AWS Account](#)

AWS Storage Gateway

Hybrid cloud storage with seamless local integration and optimized data transfer

[Get started with AWS Storage Gateway](#)

[Request more information](#)

[Product details](#) [Pricing](#) [FAQs](#) [Documentation](#) [Resources](#) [Customers](#)

AWS Storage Gateway is a hybrid storage service that enables your on-premises applications to seamlessly use AWS cloud storage. You can use the service for backup and archiving, disaster recovery, cloud data processing, storage tiering, and migration. Your applications connect to the service through a gateway appliance using standard storage protocols, such as NFS, SMB and iSCSI. The gateway connects to AWS storage services, such as Amazon S3, Amazon Glacier, and Amazon EBS, providing storage for **files**, **volumes**, and **virtual tapes** in AWS. The service includes a highly-optimized data transfer mechanism, with bandwidth management, automated network resilience, and efficient data transfer, along with a local cache for low-latency on-premises access to your most active data.

Storage Gateway hybrid storage solutions

Enables using standard storage protocols to access AWS storage services.

- Files
- Volumes
- Tapes
- AWS S3
- Amazon S3
- Amazon Glacier
- Amazon EBS
- Amazon EBS snapshots
- Amazon CloudWatch
- AWS Identity and Access Management (IAM)
- AWS CloudTrail
- AWS Config
- AWS IAM

Tech Chat: AWS Storage Gateway Deep Dive 33:06

- Whitepapers
- Webinars
- Demos & Videos
- Developer Resources
- Recorded AWS re:Invent Sessions
- Product Documentation
- Pricing
- FAQs
- Customer Case Studies

Thank you!