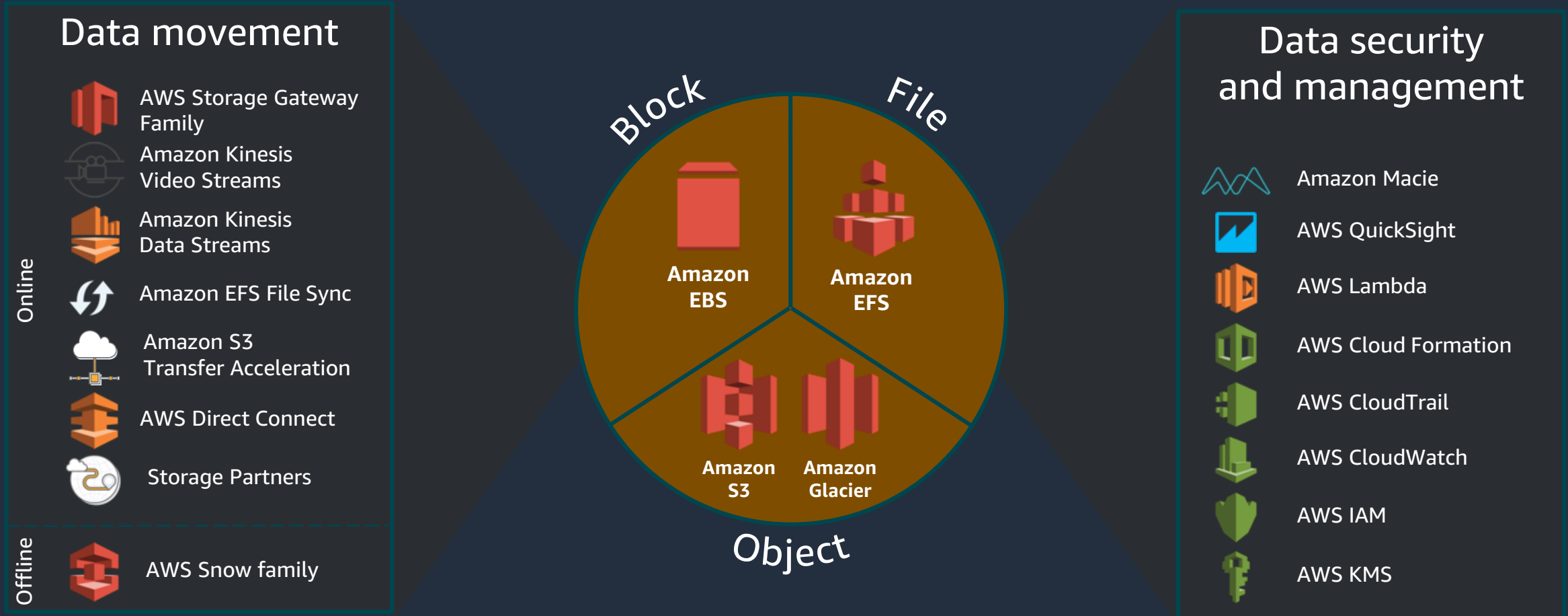# Migrate Large-Scale Datasets to AWS (Level 200)
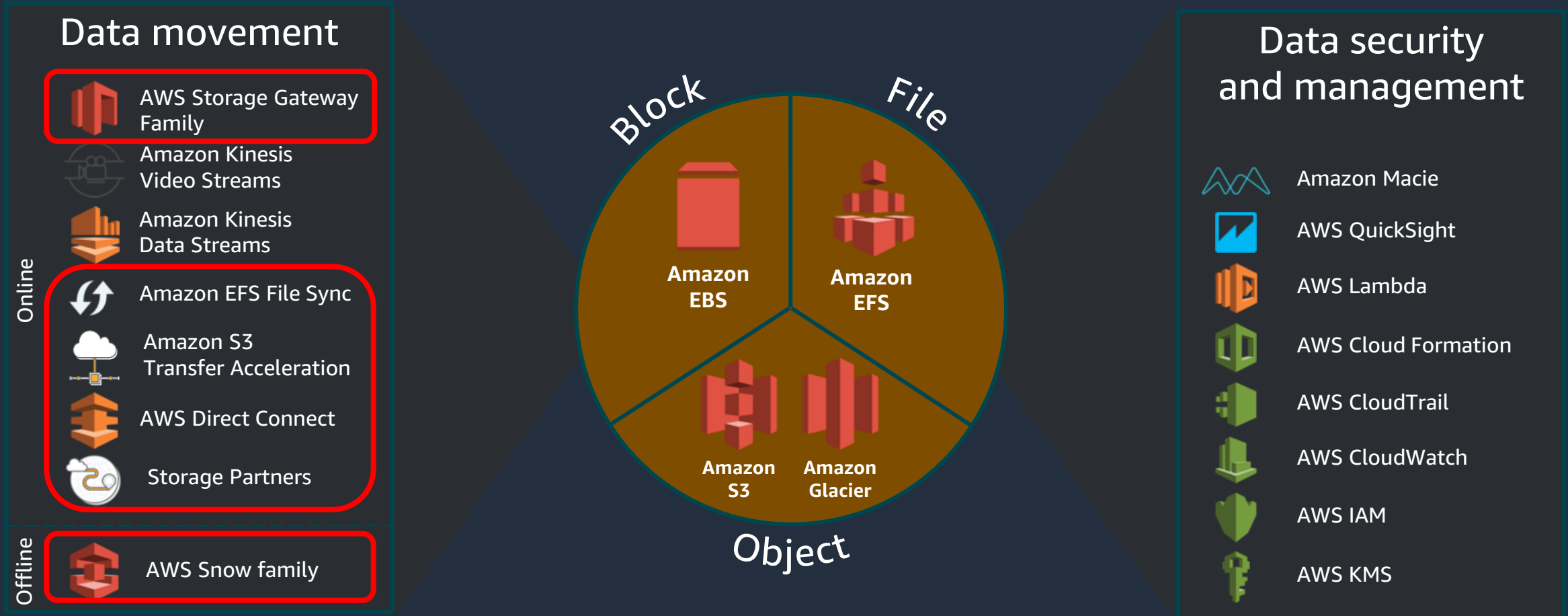
Luke Anderson, Head of Storage

Complete set of data building blocks

Data movement

Online

- AWS Storage Gateway Family
- Amazon Kinesis Video Streams
- Amazon Kinesis Data Streams
- Amazon EFS File Sync
- Amazon S3 Transfer Acceleration
- AWS Direct Connect
- Storage Partners

Offline

- AWS Snow family

Block
Amazon EBS

File
Amazon EFS

Object
Amazon S3 · Amazon Glacier

Data security and management

- Amazon Macie
- AWS QuickSight
- AWS Lambda
- AWS Cloud Formation
- AWS CloudTrail
- AWS CloudWatch
- AWS IAM
- AWS KMS

aws | intel

Complete set of data building blocks

Data movement

Online

- AWS Storage Gateway Family
- Amazon Kinesis Video Streams
- Amazon Kinesis Data Streams
- Amazon EFS File Sync
- Amazon S3 Transfer Acceleration
- AWS Direct Connect
- Storage Partners

Offline

- AWS Snow family

Block

Amazon EBS

File

Amazon EFS

Object

Amazon S3   Amazon Glacier

Data security and management

- Amazon Macie
- AWS QuickSight
- AWS Lambda
- AWS Cloud Formation
- AWS CloudTrail
- AWS CloudWatch
- AWS IAM
- AWS KMS

aws | intel

Moving to the Cloud

Common drivers of moving data to the cloud:

Cost effectiveness

Workforce productivity

Business agility

Operational resilience

Access to innovative services

Pain points of data migration:

While every cloud adoption journey is unique, all require a data migration step, which can be hindered by blockers such as:

Disruption

Speed

Costs

Channels

Cloud compatibility

AWS | intel

Migrating Data: 5 Key Questions

Migrating Data: 5 Key Questions

Amazon Glacier   IOT   RDS   Redshift

EFS   EBS   Aurora   EMR

EC2   S3

AWS

1) What kind of data is it, and where is it going?

Files   Block Volumes   Databases   IOT   Streams

aws | intel

Migrating Data: 5 Key Questions

Amazon Glacier   IOT   RDS   Redshift

EFS   EBS   Aurora   EMR

EC2   **AWS**   S3

1) What kind of data is it, and where is it going?

2) 1-time or continuous movement?

Files   Block Volumes   Databases   IOT   Streams

aws | intel

Migrating Data: 5 Key Questions

Amazon Glacier **IOT** **RDS** **Redshift**

EFS **EBS** **Aurora** **EMR**

EC2 **S3**

AWS

1) What kind of data is it, and where is it going?

2) 1-time or continuous movement?

3) 1-way or bi-directional?

Files

Block Volumes

Databases

IOT

Streams

aws | intel

Migrating Data: 5 Key Questions

Amazon Glacier   IOT   RDS   Redshift

EFS   EBS   Aurora   EMR

EC2   S3

AWS

1) What kind of data is it, and where is it going?

2) 1-time or continuous movement?

3) 1-way or bi-directional?

4) How much data & time do you have?

Files   Block Volumes   Databases   IOT   Streams

aws | intel

AWS Direct Connect

AWS Direct Connect

- Establish private connectivity between AWS and your data center
- Dedicated connection can be partitioned into multiple virtual interfaces
- Maintain network separation between public and private environments

BENEFITS
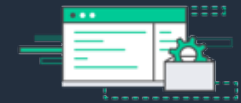
| Reduce Bandwidth Costs | Consistent Network Performance | Compatible with all AWS Services | Private Connectivity to VPC | Elastic | Simple |
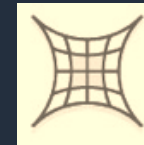
AWS Snow*

AWS Snow* family

AWS Snowball

- Petabyte-scale data migration
- Rugged 8.5 G impact case
- Rain and dust resistant
- Data encryption end-to-end
- 80 TB capacity/10 Gb network
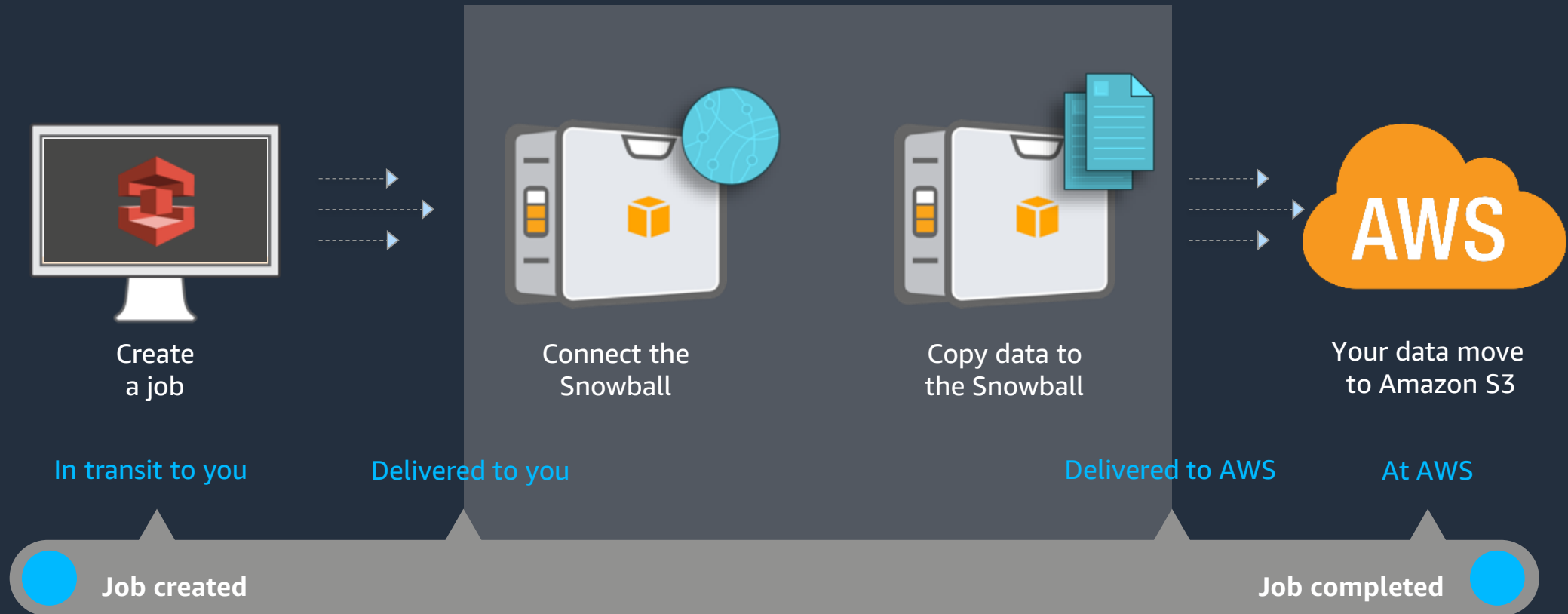- Move 250 TB in 1 Week!

AWS Snowball Edge

- Compute and storage for hybrid/edge workloads
- 100 TB local storage
- Local compute equivalent to an Amazon EC2 m4.4xlarge instance
- 10GBase-T, 10/25 Gb SFP28, and 40 Gb QSFP+ copper, and optical networking
- Ruggedized and rack-mountable

AWS Snowmobile

- Exabyte-scale data migration
- Up to 200 PB capacity
- Data encryption end-to-end
- Dedicated security personnel
- GPS tracking, alarm monitoring, 24/7 surveillance, and optional additional security

aws | intel

AWS Snow* family

Create a job

Connect the Snowball

Copy data to the Snowball

Your data move to Amazon S3

In transit to you

Delivered to you

Delivered to AWS

At AWS

Job created

Job completed

aws | (intel)

Amazon EFS File Sync

Amazon EFS File Sync
Sync data from existing file systems into EFS file systems

Simple
Set up and manage easily
from the AWS Console

Fast
Up to 5x faster than
standard Linux copy tools

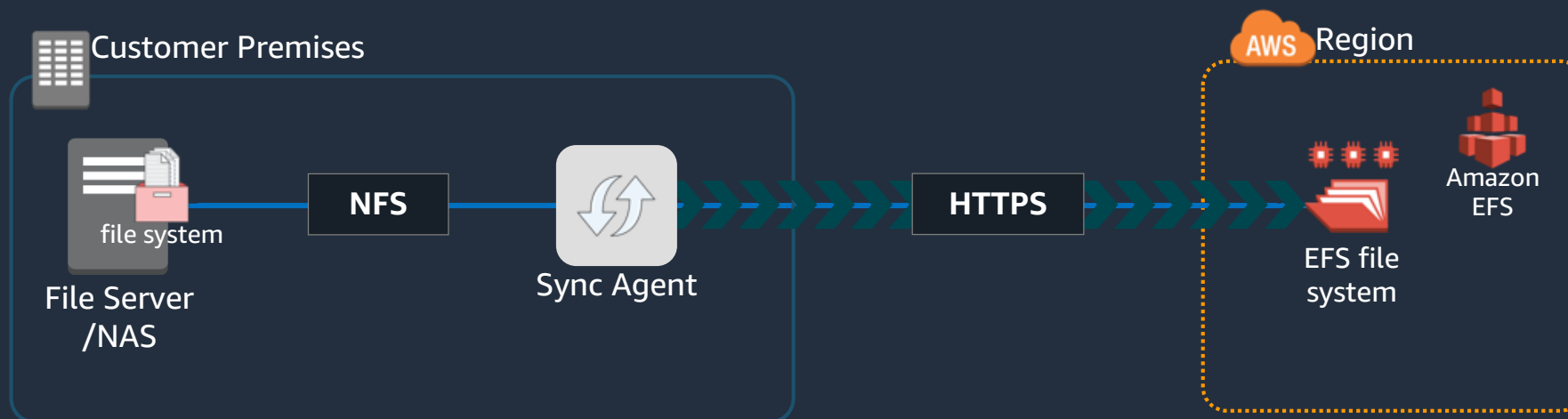### Secure
Encrypted parallel data
transfer to AWS

File systems from on-premises to EFS

Use EFS File Sync to copy…

DIY in-cloud file systems to EFS

EFS file systems between AWS Regions

aws | intel

Amazon EFS File Sync
Fast and simple way to copy file systems to Amazon EFS

Customer Premises

NFS

Sync Agent

File Server /NAS

file system

HTTPS

AWS Region

EFS file system

Amazon EFS

- Sync Agent deployed as a VM
- No data is stored on the Sync Agent
- Optimized TCP stack for high-throughput
- Data is compressed and encrypted across the network

aws | intel

Amazon S3 Transfer Acceleration
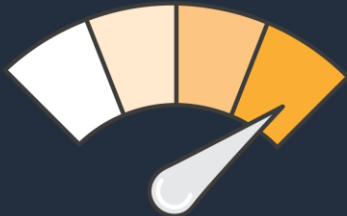
aws | intel

Amazon S3 Transfer Acceleration

Change your endpoint, not your code

Leverages 59 global edge locations

Optimized protocols

No firewall exceptions

No client software required

S3 Bucket
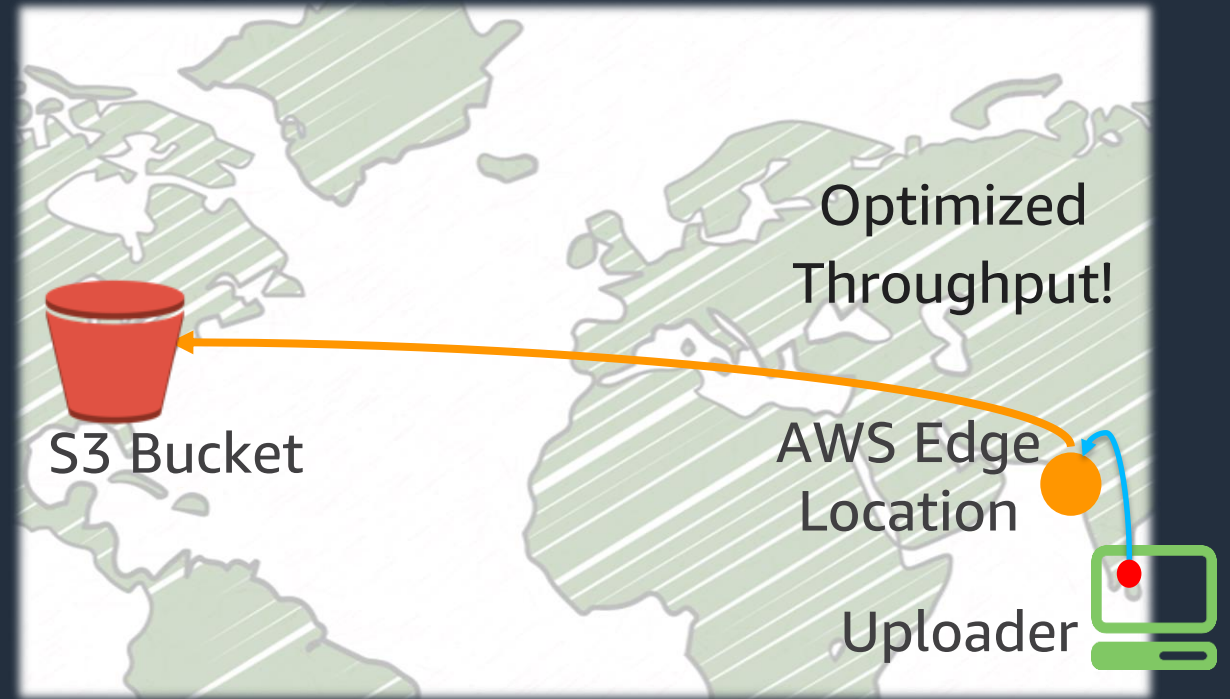
Uploader

aws | intel

Amazon S3 Transfer Acceleration

Change your endpoint, not your code

Leverages 59 global edge locations

Optimized protocols

No firewall exceptions

No client software required

S3 Bucket

Uploader

aws | intel

Amazon S3 Transfer Acceleration

Change your endpoint, not your code

Leverages 59 global edge locations

Optimized protocols

No firewall exceptions
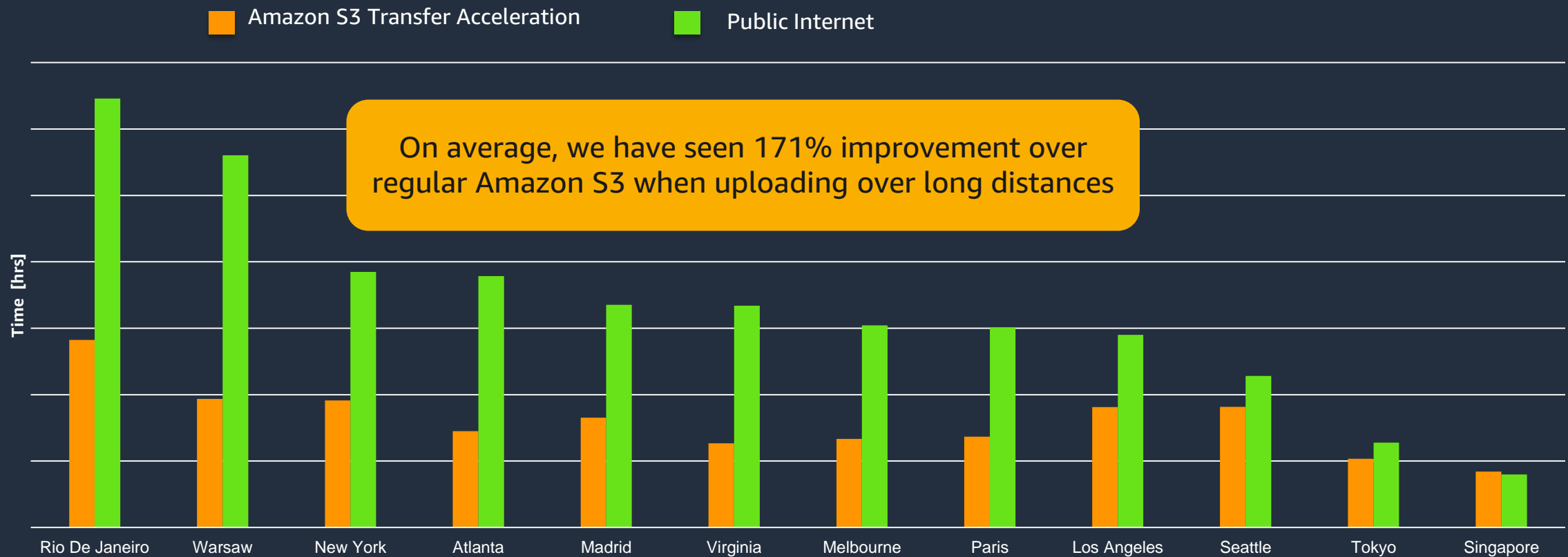
No client software required

S3 Bucket

AWS Edge Location

Uploader

Amazon S3 Transfer Acceleration

Change your endpoint, not your code

Leverages 59 global edge locations

Optimized protocols

No firewall exceptions

No client software required

S3 Bucket

Optimized Throughput!

AWS Edge Location

Uploader

Amazon S3 Transfer Acceleration

How fast is Amazon S3 Transfer Acceleration?

Legend: ■ Amazon S3 Transfer Acceleration (orange) ■ Public Internet (green)

On average, we have seen 171% improvement over regular Amazon S3 when uploading over long distances

Y-axis: Time [hrs]

X-axis categories: Rio De Janeiro, Warsaw, New York, Atlanta, Madrid, Virginia, Melbourne, Paris, Los Angeles, Seattle, Tokyo, Singapore

500 GB upload from these edge locations to a bucket in Singapore

Amazon S3 Transfer Acceleration
Case Study

Typical Friday during football season: More than 35 hours of video every minute is uploaded

S3-TA = >20% increase in upload and encoding speeds

" Amazon S3 transfer acceleration reduces the average time it takes for us to ingest videos from our global user base by almost half. This gives our customers the ability to edit and share videos sooner where speed is a critical factor. "
- Brian Kaiser, CTO

AWS Storage Gateway

AWS Storage Gateway hybrid storage solutions
Use standard storage protocols to access AWS storage services

On-premises
- Files
- Volumes
- Tapes

AWS Storage Gateway

AWS Cloud
- Amazon S3
- Amazon Glacier
- Amazon EBS snapshots

Amazon CloudWatch

AWS CloudTrail

AWS Identity and Access Management (IAM)

AWS Key Management Service (KMS)

File gateway
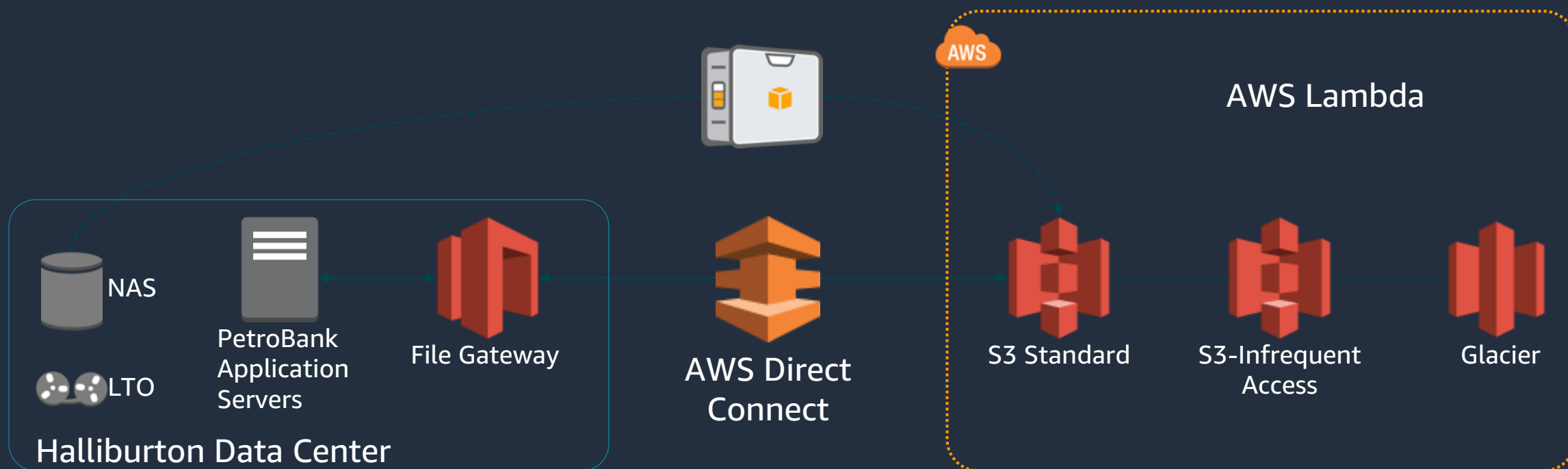## On-premises file storage maintained as objects in Amazon S3

Customer Premises

AWS

Application Server — NFS v3 / v4.1 — Storage Gateway VM — HTTPS — Storage Gateway — S3 Standard — S3 Standard-Infrequent Access — Amazon Glacier
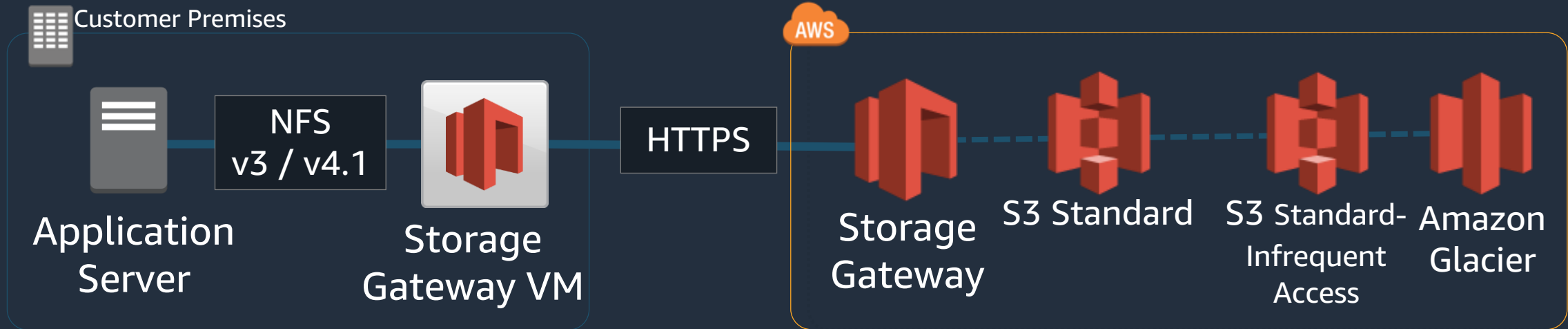
- Data stored and retrieved from your Amazon S3 buckets
- One-to-one mapping from files to objects
- File metadata stored in object metadata
- Bucket access managed by IAM role you own and manage
- Use Amazon S3 lifecycle policies, versioning, or CRR to manage data

Active archive migration from LTO
Cost-effective storage in AWS with local data access

Self service loading of data
Reduced time-to-data by days or weeks
Storage archive costs reduced by 90%

HALLIBURTON | Landmark

AWS Lambda

AWS

NAS

LTO

PetroBank Application Servers

File Gateway

AWS Direct Connect

S3 Standard

S3-Infrequent Access

Glacier

Halliburton Data Center

aws | intel

Tape Gateway
Virtual tape storage in Amazon S3 and Amazon Glacier with VTL

Customer Premises

Tape Gateway VM

Region

iSCSI

HTTPS

Backup Server

Media Changer

Tape Drive

Storage Gateway Service

Virtual Tapes stored in Amazon S3

Archived Tapes stored in Amazon Glacier
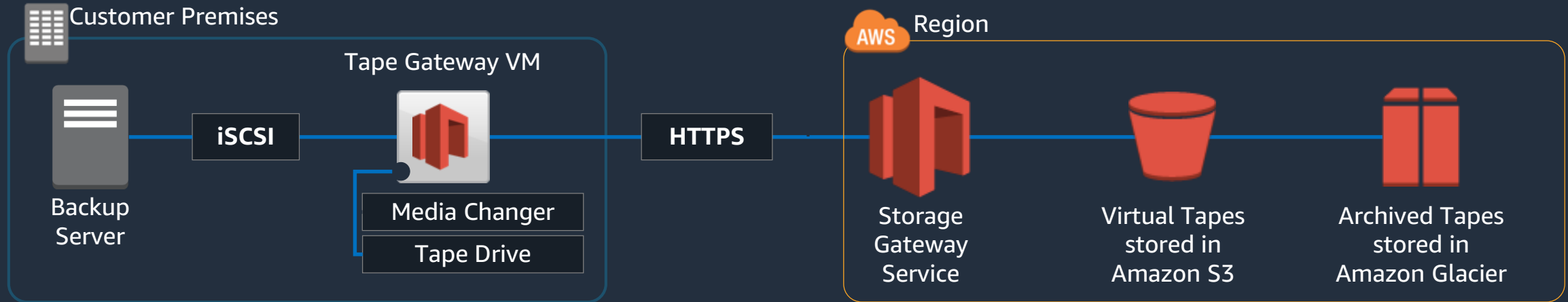
- Virtual tape storage in S3 and Glacier accessed via tape gateway
- Data compressed in-transit and at-rest
- Up to 1 PB total tape storage per gateway, unlimited archive capacity
- Supports leading backup applications
- **3-5 hour tape retrieval from Amazon Glacier

VERITAS™   DELLEMC   veeAM   COMMVAULT   Hewlett Packard Enterprise   Microsoft System Center Data Protection Manager 2012   arcserve

aws   intel

Move your tape archive
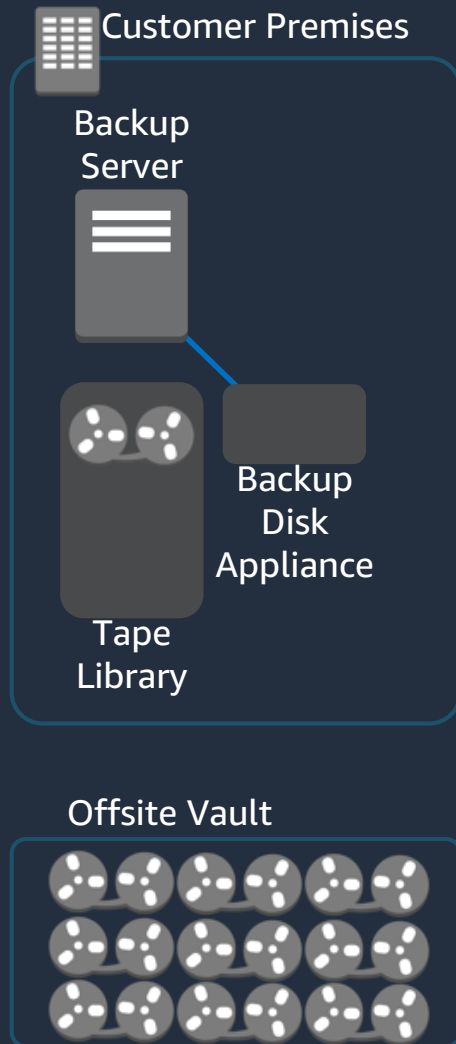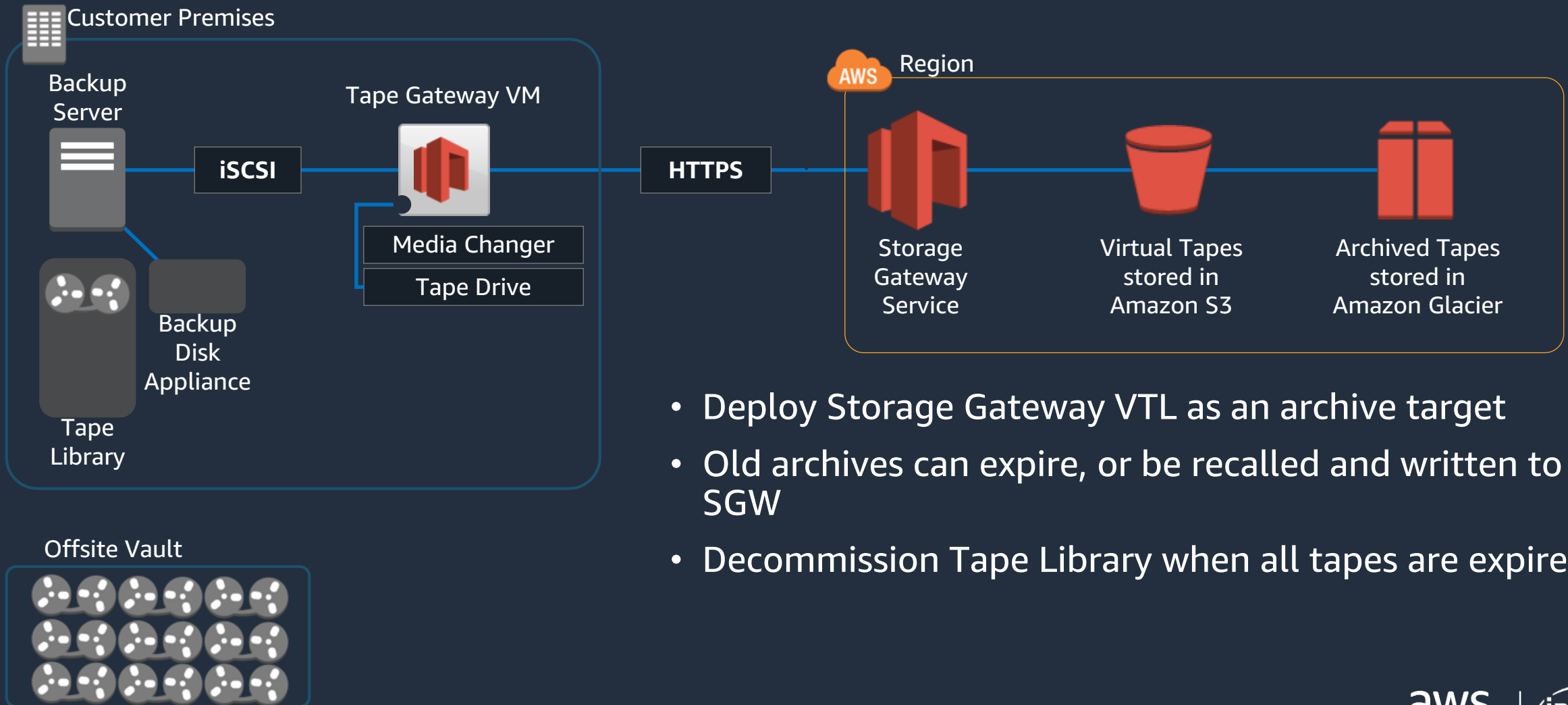
Customer Premises

Backup
Server
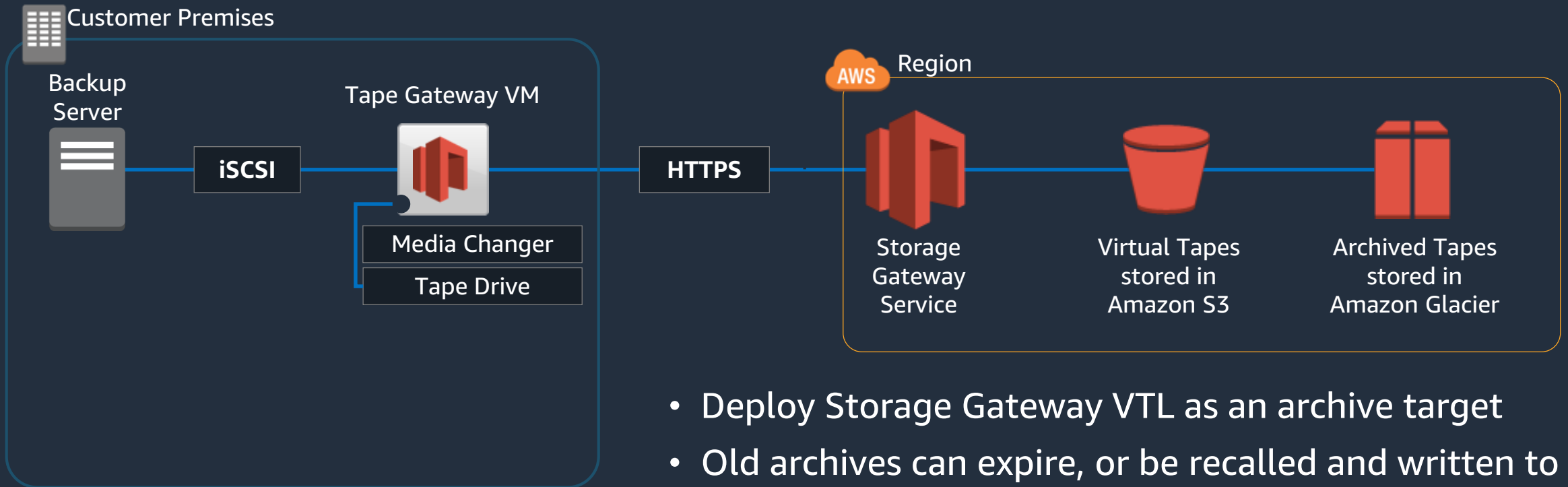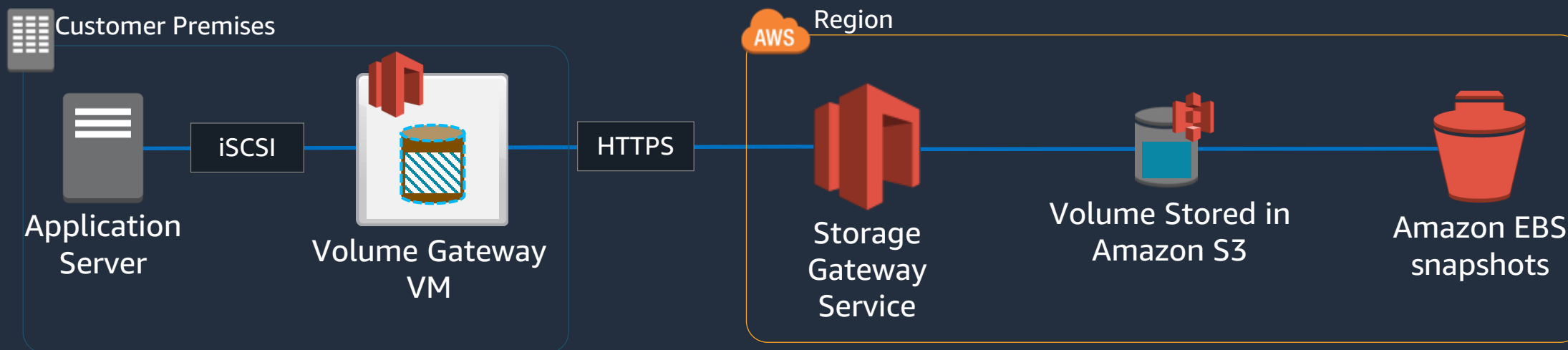
Backup
Disk
Appliance

Tape
Library

Offsite Vault

Move your tape archive

- Deploy Storage Gateway VTL as an archive target
- Old archives can expire, or be recalled and written to SGW
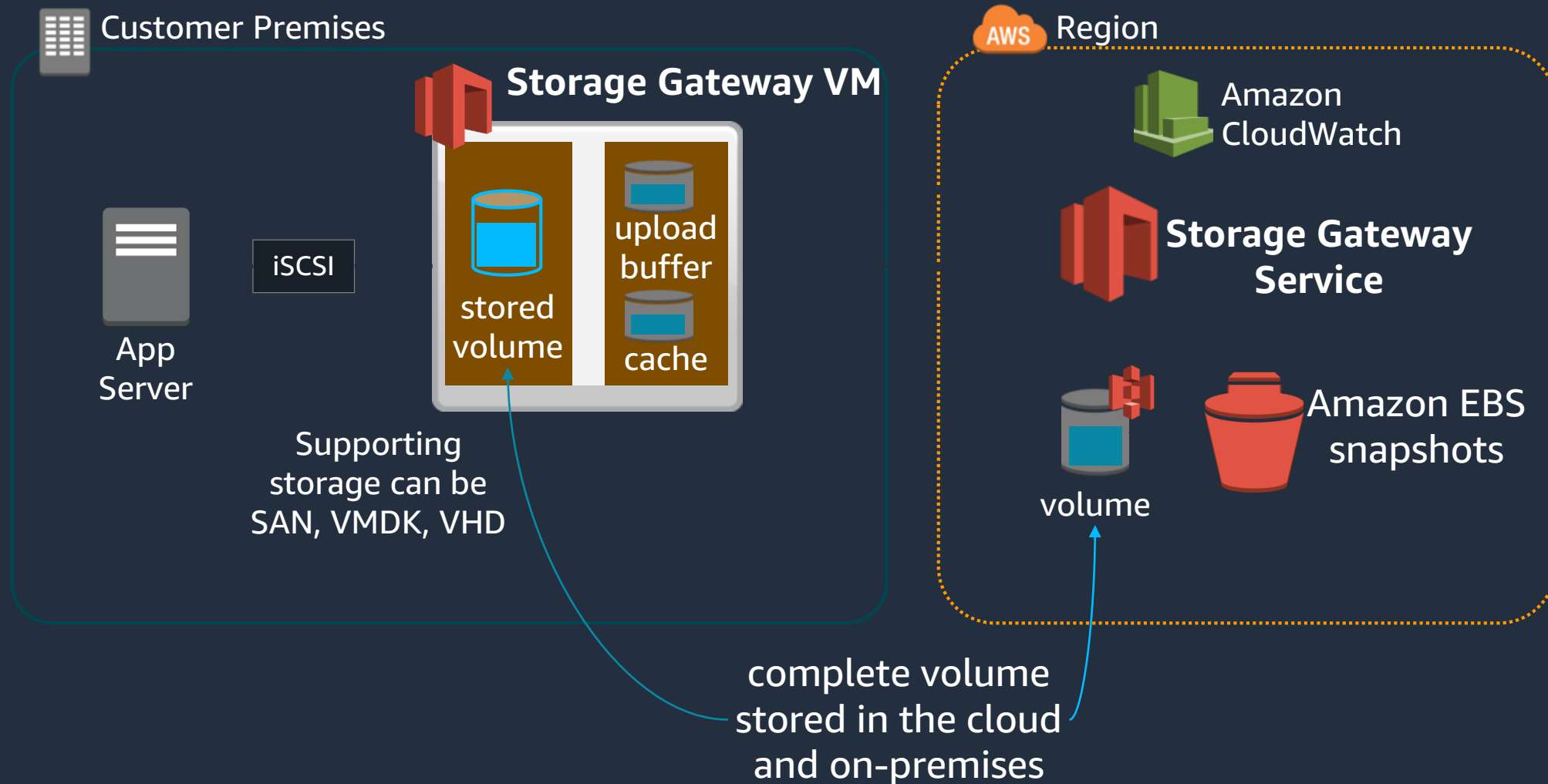- Decommission Tape Library when all tapes are expired

Move your tape archive

- Deploy Storage Gateway VTL as an archive target
- Old archives can expire, or be recalled and written to SGW
- Decommission Tape Library when all tapes are expired

Volume gateway
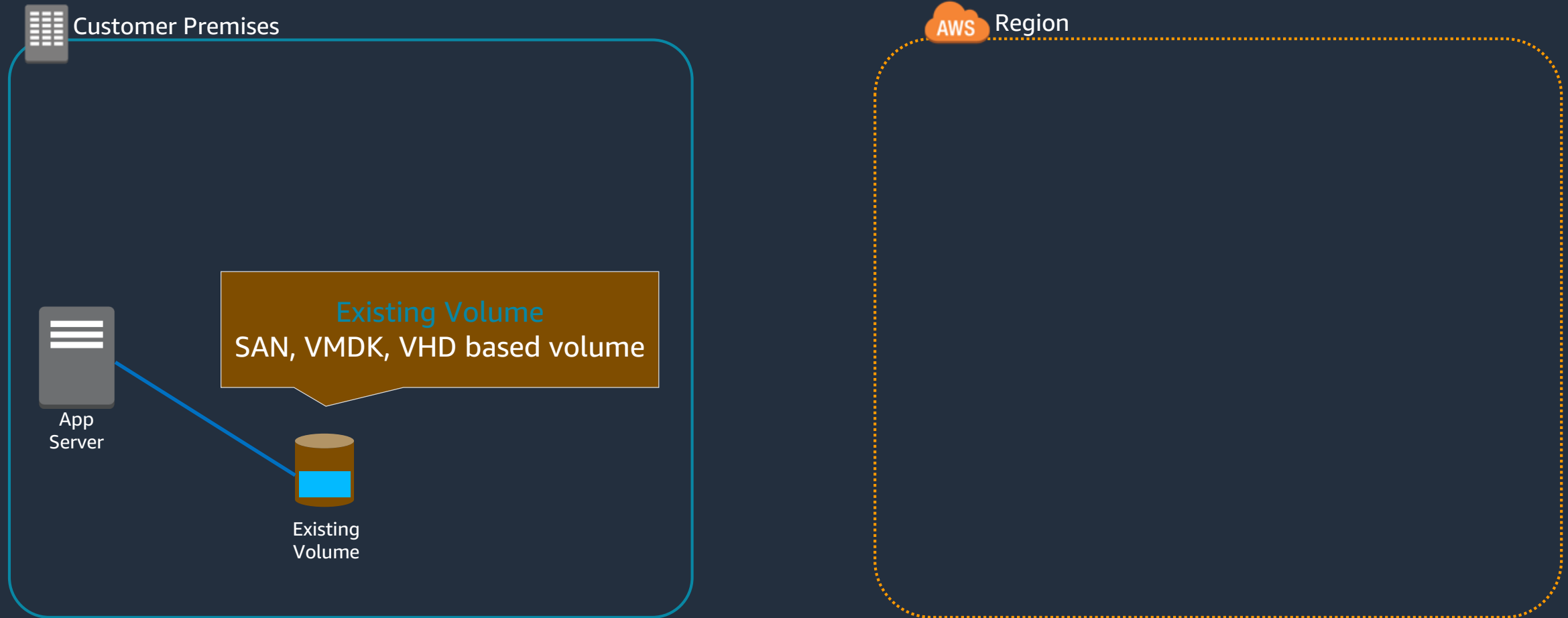## On-premises volume storage backed by Amazon S3 with EBS snapshots

Customer Premises

Application Server — iSCSI — Volume Gateway VM — HTTPS

AWS Region

Storage Gateway Service — Volume Stored in Amazon S3 — Amazon EBS snapshots

- Presents block storage on top of Amazon S3 accessed via the volume gateway
- Compression of data in transit and at rest
- Back up on-premises volumes to EBS snapshots
- Create on-premises volumes from EBS snapshots
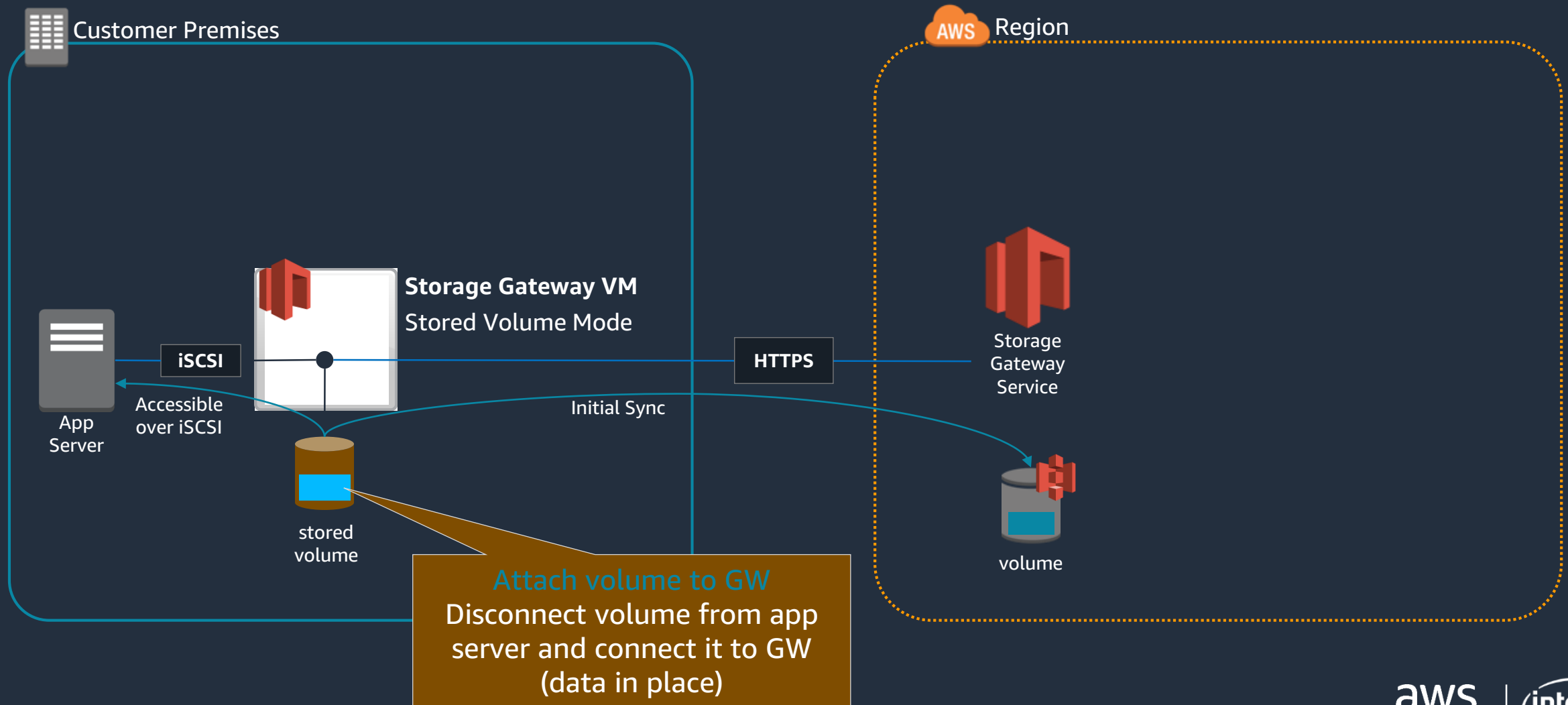- Up to 1 PB of total volume storage per gateway

aws | (intel)
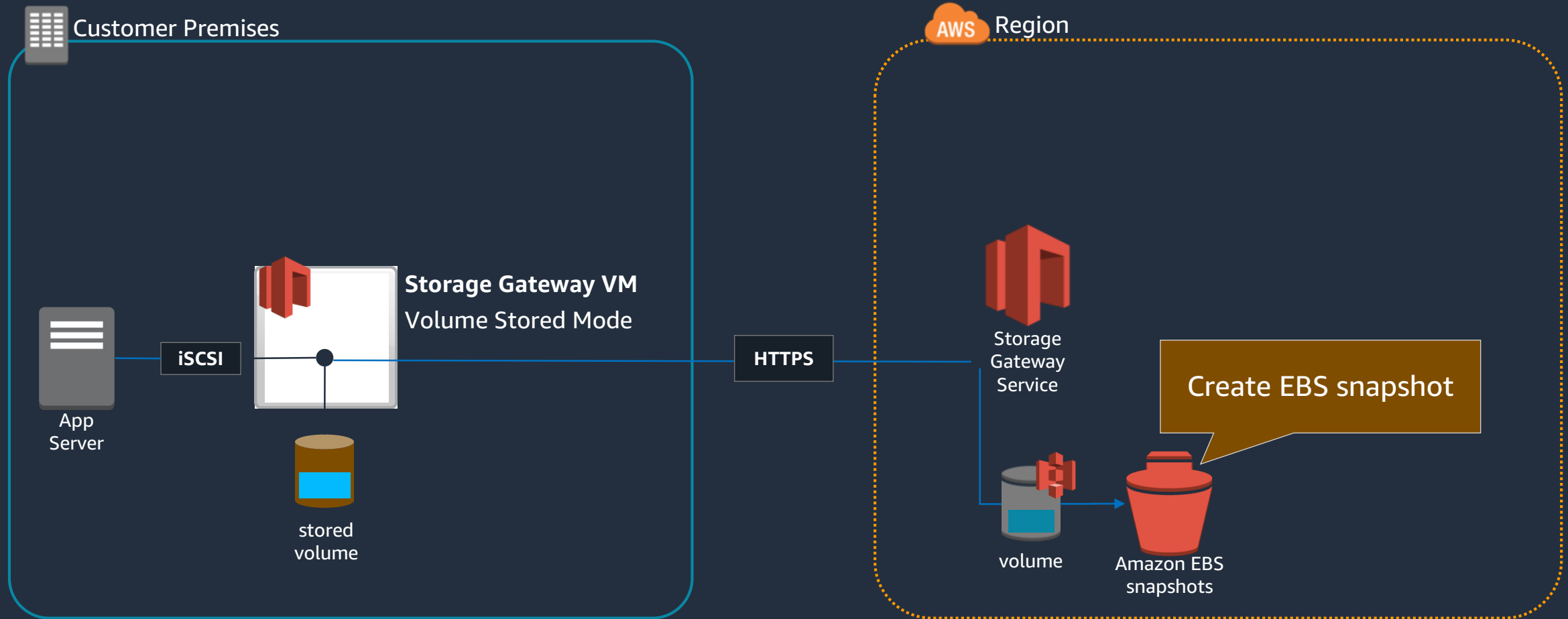
Volume stored

Customer Premises

Storage Gateway VM

App Server

iSCSI

stored volume

upload buffer

cache

Supporting storage can be SAN, VMDK, VHD

AWS Region

Amazon CloudWatch

Storage Gateway Service

volume

Amazon EBS snapshots

complete volume stored in the cloud and on-premises

aws | intel

Migration with Storage Gateway volumes

Migration with Storage Gateway volumes

Customer Premises

AWS Region

App Server

iSCSI

Accessible over iSCSI

Storage Gateway VM
Stored Volume Mode

HTTPS

Initial Sync

Storage Gateway Service

stored volume

volume

Attach volume to GW
Disconnect volume from app server and connect it to GW (data in place)

aws | intel

Migration with Storage Gateway volumes

Migration with Storage Gateway volumes

Migration with Storage Gateway volumes

Customer Premises

AWS Region

DATA

Windows Server

EC2 instance

AWS Third Party Integrations

The AWS Partner Network (APN) helps you seamlessly integrate AWS services into your existing storage environment

Storage Training

For Enterprise Storage Engineers

Learn how to architect and manage highly available solutions on AWS storage services

Advance toward AWS certifications

Help your organization migrate to the cloud faster

- Online at www.aws.training

- Access 100+ new digital training courses including advanced training on storage

- Deep Dives on S3, EFS, and EBS

- Migrating and Tiering Storage to AWS (Hybrid Solutions)

With deep expertise on AWS, APN Partners can help your organization at any stage of your Cloud Adoption Journey.

aws msp

AWS Managed Service Providers

APN Consulting Partners who are skilled at cloud infrastructure and application migration, and offer proactive management of their customer's environment.

aws competency

AWS Competency Partners

APN Partners who have demonstrated technical proficiency and proven customer success in specialized solution areas.

aws marketplace

AWS Marketplace

A digital catalog with thousands of software listings from independent software vendors that make it easy to find, test, buy, and deploy software that runs on AWS.

aws service delivery
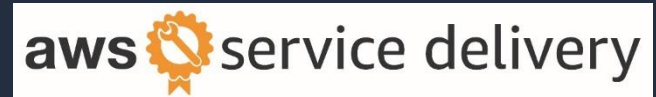
AWS Service Delivery Partners

APN Partners with a track record of delivering specific AWS services to customers.

Ready to get started with an APN Partner?
Find a partner: https://aws.amazon.com/partners/find/
Learn more at the AWS Partner Network Booth

aws | intel

Thank You for Attending AWS Innovate

We hope you found it interesting! A kind reminder to **complete the survey.**
Let us know what you thought of today's event and how we can improve the event experience for you in the future.

- aws-apac-marketing@amazon.com
- twitter.com/AWSCloud
- facebook.com/AmazonWebServices
- youtube.com/user/AmazonWebServices
- slideshare.net/AmazonWebServices
- twitch.tv/aws

aws | intel