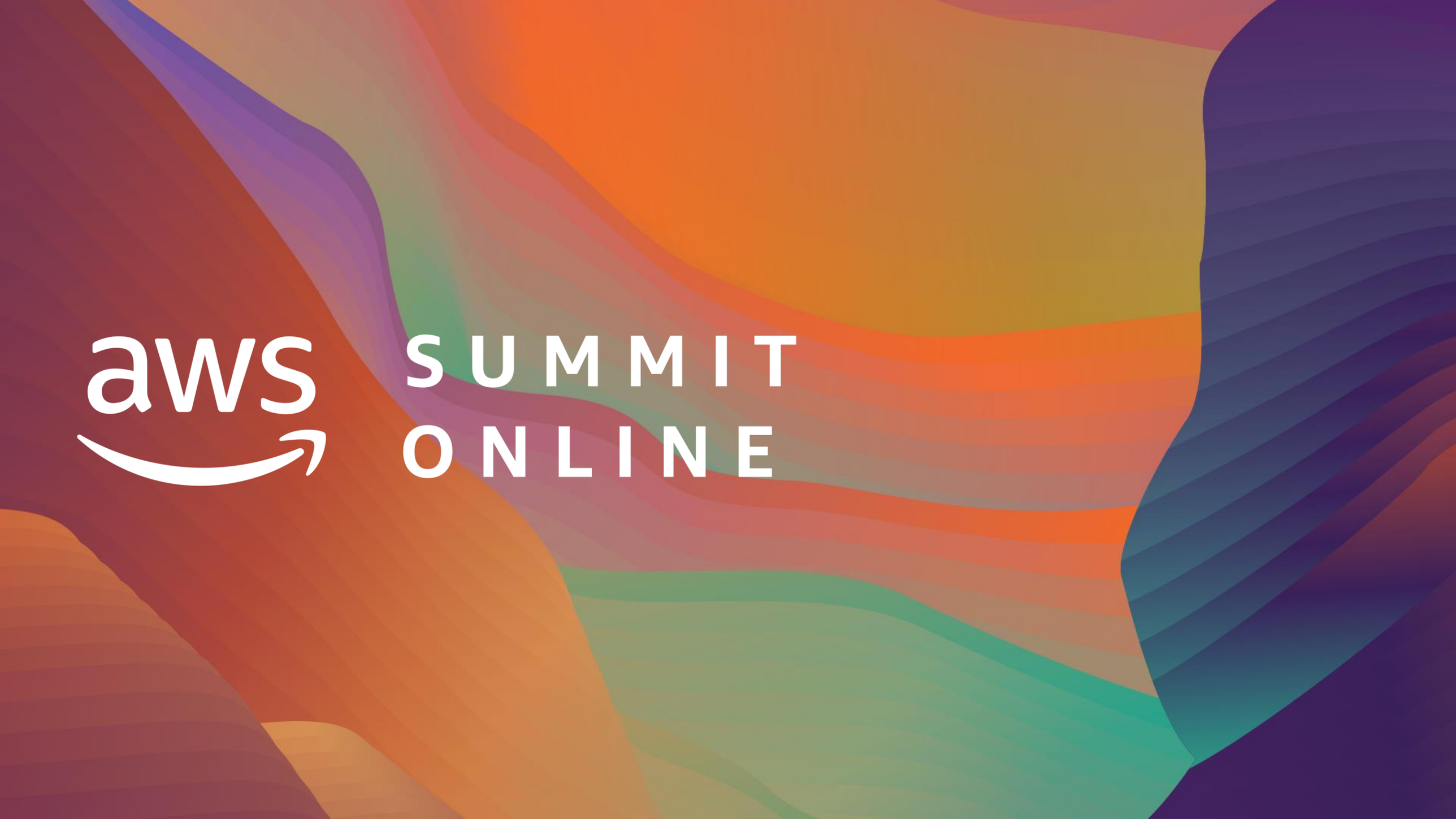

aws
SUMMIT
ONLINE

TRA10

Simplify and accelerate your data migration journey to AWS

Wali Akbari

Storage Solution Architect
Amazon Web Services

aws SUMMIT ONLINE

Data transfer use cases

Why do customers transfer data to the cloud

Application migration

Data lakes

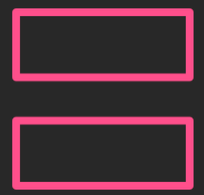
Sharing Data

Backups

Data management

Data transfer challenges

What are some challenges customers face when looking to transfer data



Mode

Speed

Time and effort

How long could it take?

	100Mbps	1Gbps	10Gbps
10TB	12 days	30 hours	3 hours
100TB	124 days	12 days	30 hours
1PB	3 years	124 days	12 days
10PB	34 years	3 years	124 days

~25% assumed network overhead

Orange: Online transfer

Light Blue: Offline transfer

White: Online or offline

How long could it take?

	100Mbps	1Gbps	10Gbps
10TB	12 days	30 hours	3 hours
100TB	124 days	12 days	30 hours
1PB	3 years	124 days	12 days
10PB	34 years	3 years	124 days

~25% assumed network overhead

Orange: Online transfer

Light Blue: Offline transfer

White: Online or offline

AWS Storage services portfolio

Block storage

Amazon EBS

File storage

Amazon EFS

Amazon FSx for Windows File Server

Amazon FSx for Lustre

Object storage

Amazon S3

Amazon S3 Glacier

Hybrid

AWS Storage Gateway Family

Transport & edge

AWS DataSync

AWS Transfer for SFTP

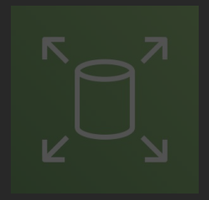
AWS Snow* Family

Backup

AWS Backup

AWS Storage services portfolio

Block storage



Amazon EBS

File storage

Amazon EFS

Amazon FSx for Windows File Server

Amazon FSx for Lustre

Object storage

Amazon S3

Amazon S3 Glacier

Hybrid

AWS Storage Gateway Family

Transport & edge

AWS DataSync

AWS Transfer for SFTP

AWS Snow* Family

Backup

AWS Backup

Online data migration

What is AWS DataSync?

Simplifies, automates, and accelerates your online data transfer

Migrate active application data

Transfer data for timely processing or Archiving

Replicate for data protection and recovery

Transfers up to 10 Gbps per agent

Simple data movement to S3, EFS, FSx for Windows

Secure and reliable transfers

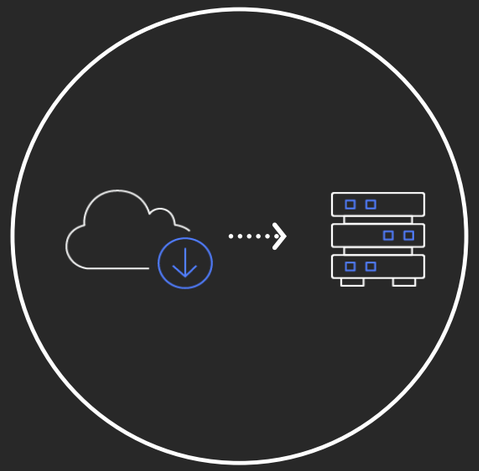
AWS integrated

Pay as you go

As simple as 1-2-3 …

Deploy DataSync agent

1



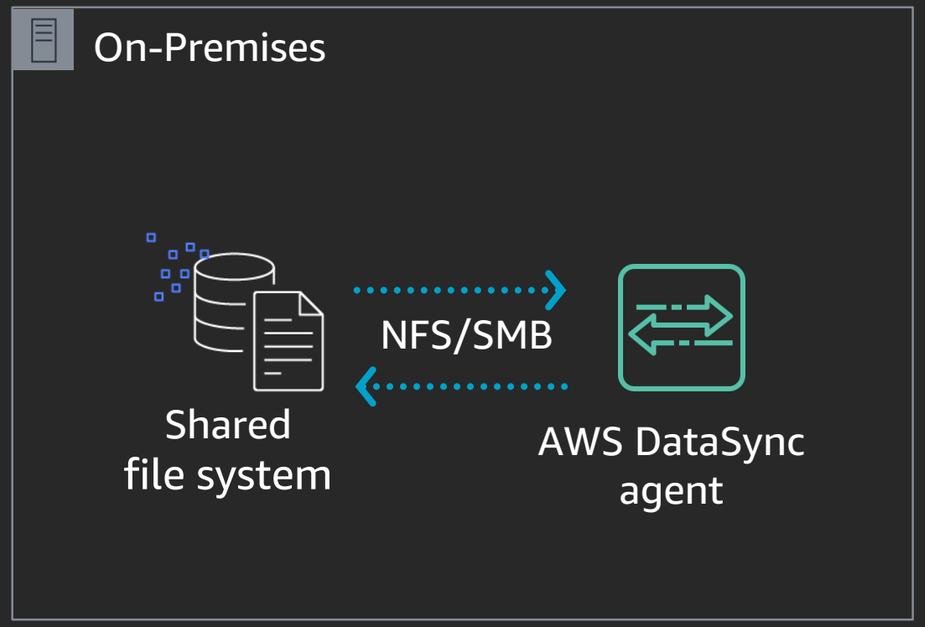
Create a task

2

Start the task and monitor

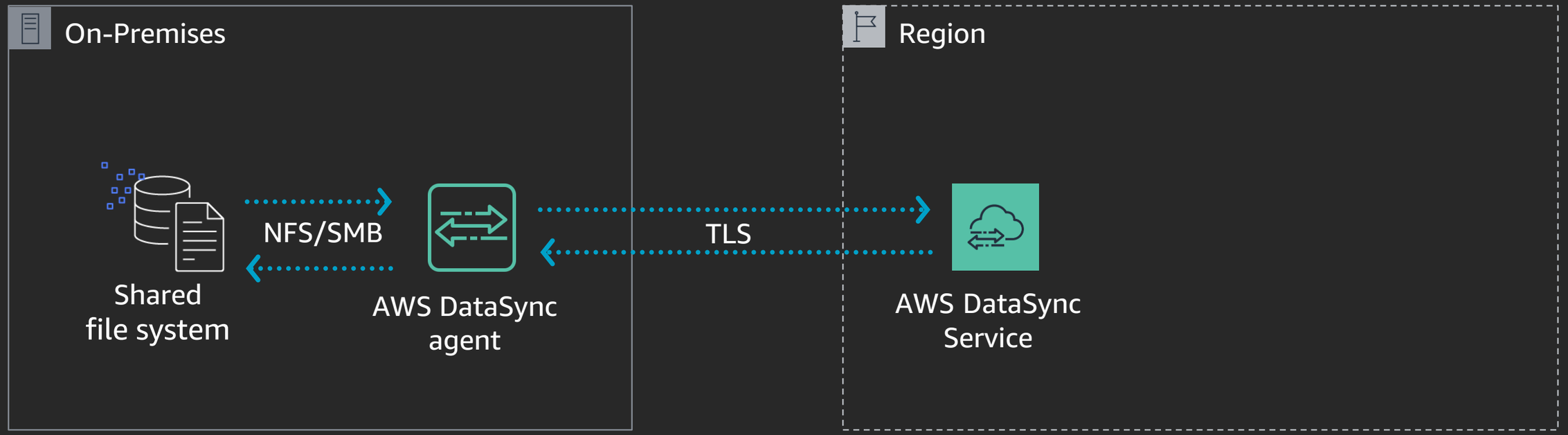
3

How AWS DataSync works



Highly parallel transfers using optimised network protocol

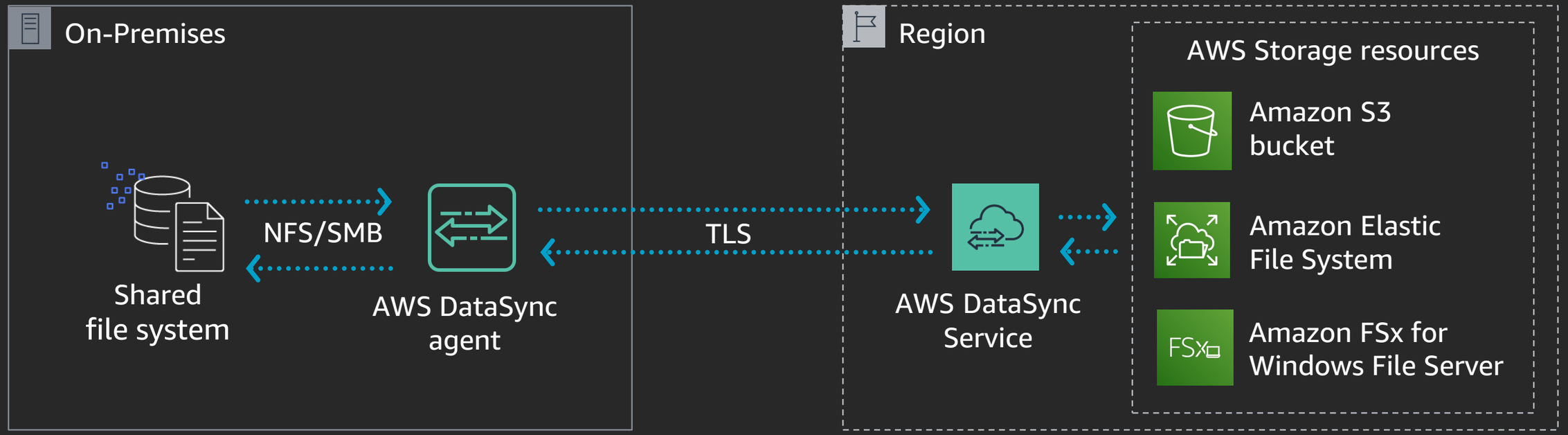
How AWS DataSync works



Highly parallel transfers using optimised network protocol

Fully-managed service scales to send/receive data from agent

How AWS DataSync works



Highly parallel transfers using optimised network protocol

Fully-managed service scales to send/receive data from agent

Optimised read & write to S3 & EFS using VPC Endpoints

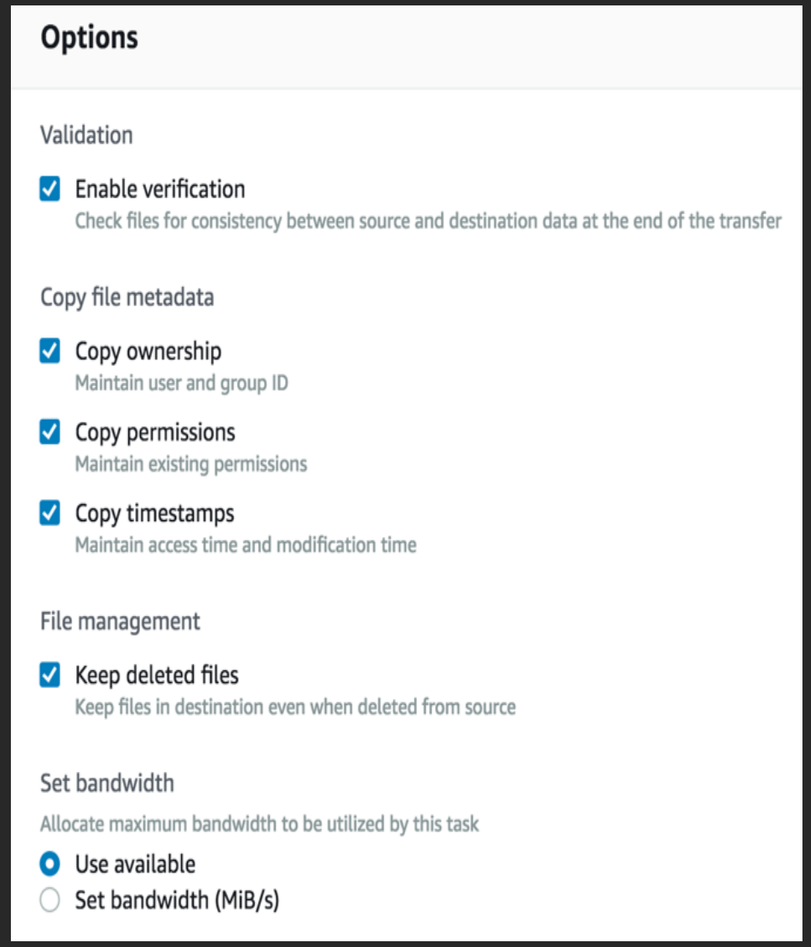
The speed and reliability of *network acceleration* software with the cost-effectiveness of *open source tools*

Task options

Invoke via schedule, API, or manually

File-level validation

Copy across file metadata



Data migration using AWS DataSync

Challenge

Multi-year retention of data, on-premises storage

Solution

Utilised AWS DataSync for fast & seamless migration to Amazon S3

Outcome

Successfully transferred & verified 700TB of data at a rate of approx. 500 GB/hour

AUTODESK.

Autodesk is a leader in 3D design, engineering, and entertainment software. Autodesk makes software for people who make things.

https://aws.amazon.com/blogs/storage/migrating-hundreds-of-tb-of-data-to-amazon-s3-with-aws-datasync/

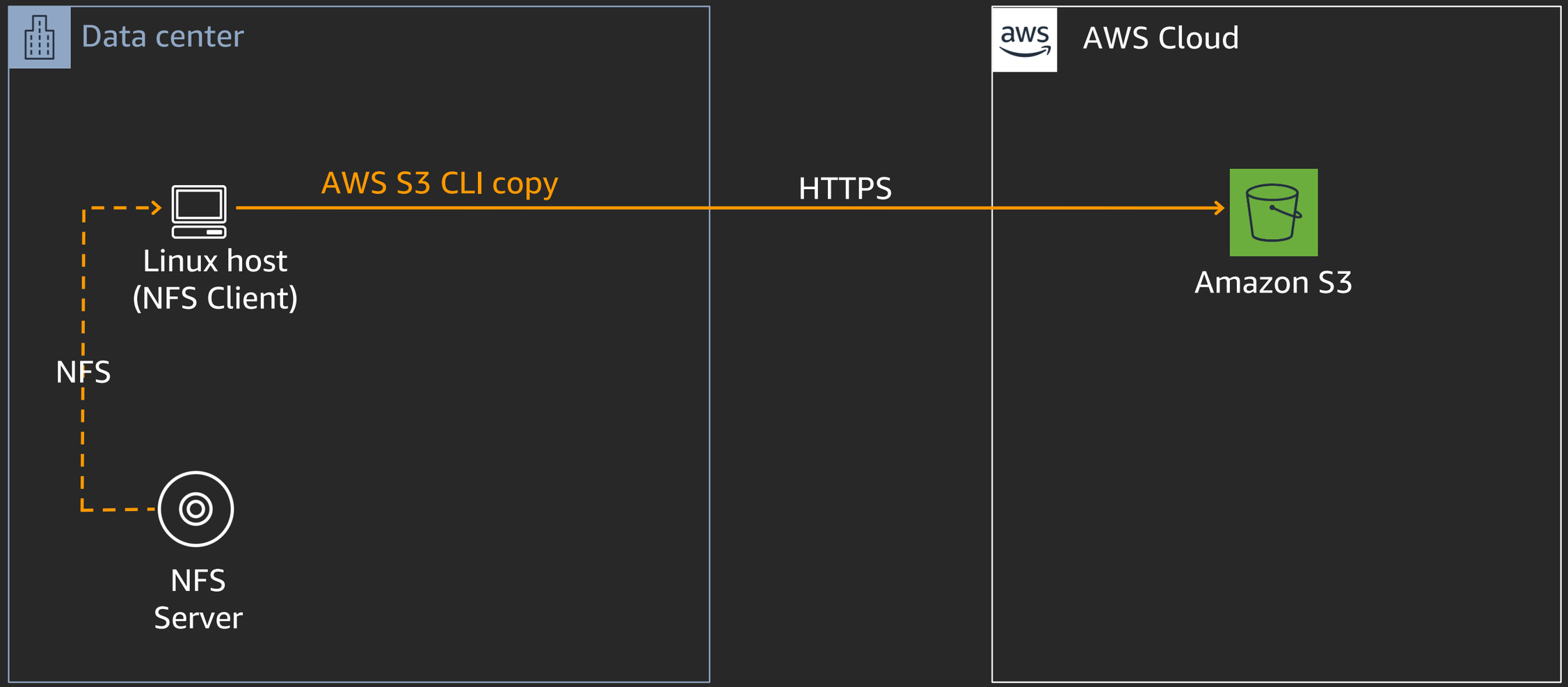
aws

Demo: Migration of 10,000 small files

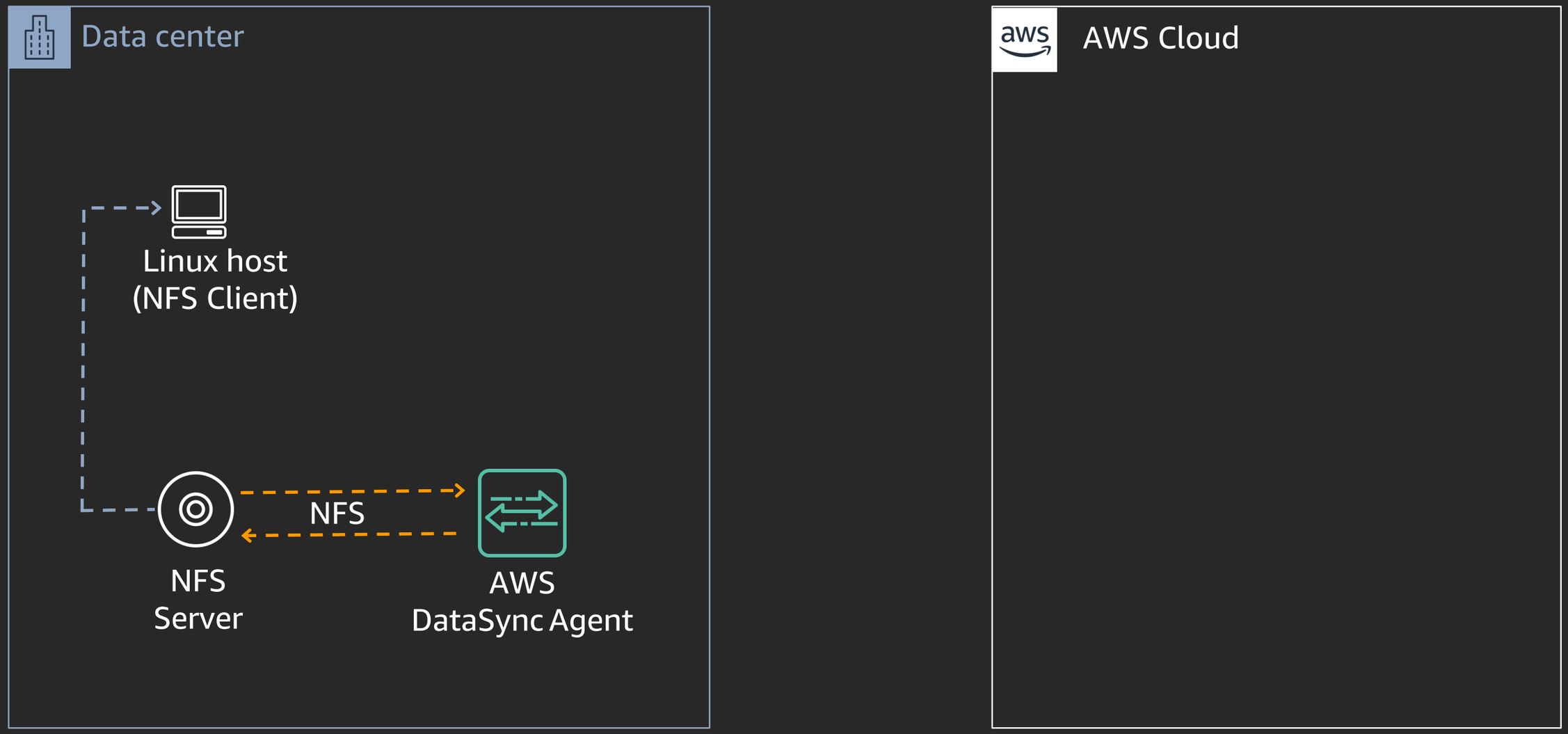
aws SUMMIT ONLINE

Demo 1: Using a copy script

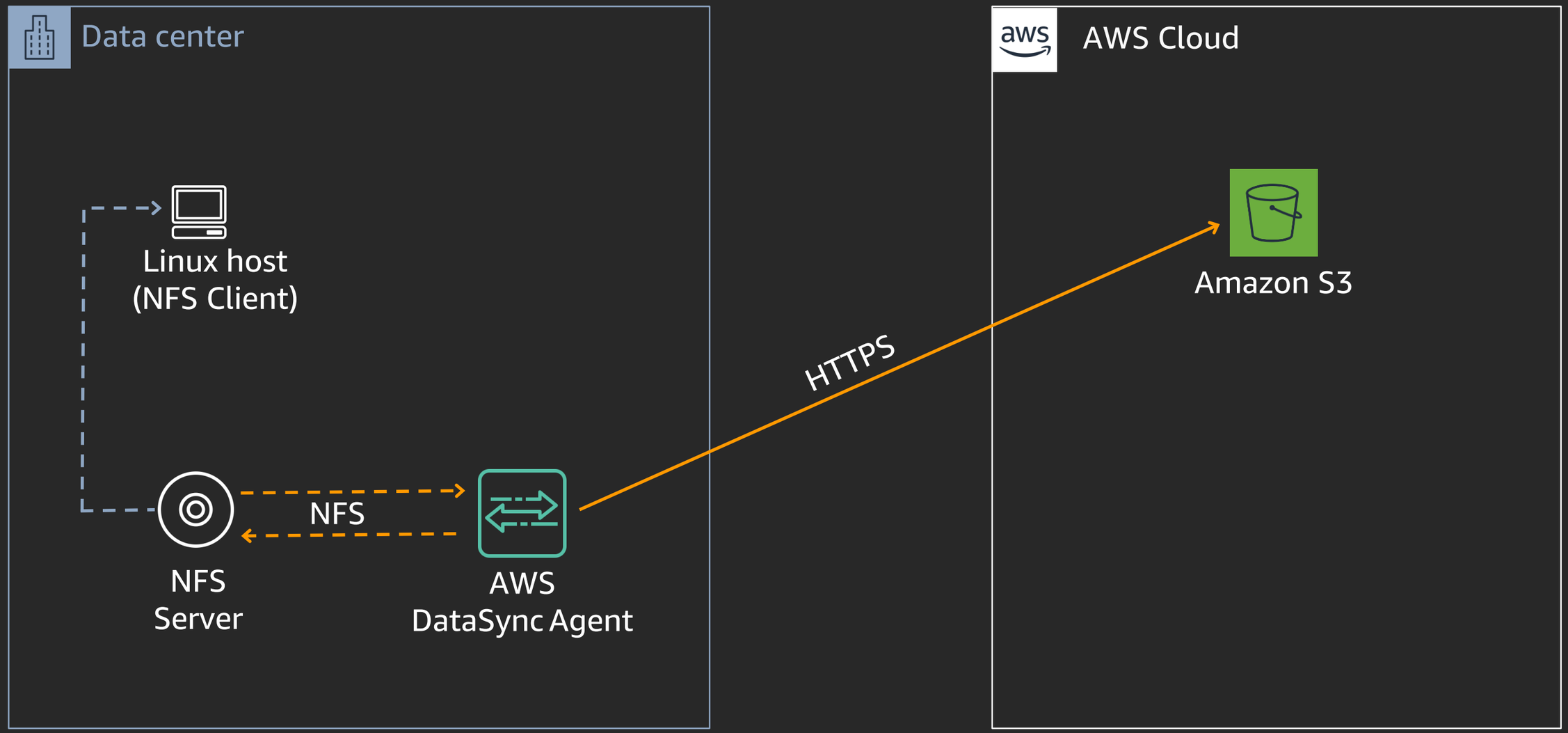
Demo 1: Using a copy script



Demo 2: Using AWS DataSync



Demo 2: Using AWS DataSync



What did I just see in this demo?

AWS DataSync accelerated the data transfer

Preserved file metadata such as timestamps and permissions

Provided data integrity verification

Think of this at scale, with simplicity in mind

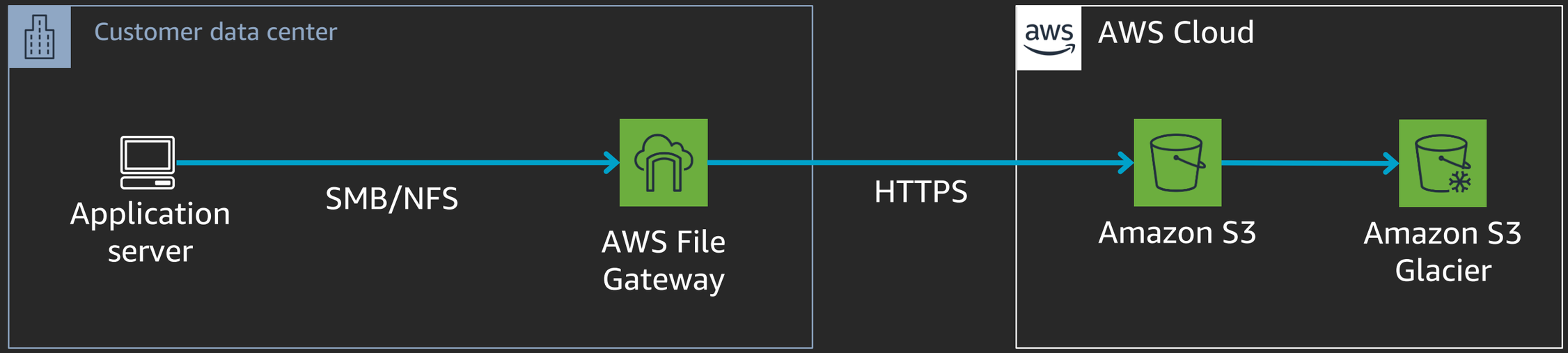
AWS DataSync usage scenarios

Data transfer → Amazon S3 storage classes

NFS share migration → Amazon EFS

SMB share migration → Amazon FSx for Windows File Server

What is AWS File Gateway?



- Allows low latency access to hot data via it’s local cache, all data is stored Amazon S3
- Utilise a network file system (NFS/SMB) to interface to Amazon S3 storage
- Stores file content and metadata in Amazon S3
- SMB shares can integrate with Microsoft Active Directory

AWS File Gateway usage scenarios

Migrate file shares → AWS File Gateway + Amazon S3 storage

Transfer data via File share → Amazon S3 storage

Amazon S3 data presented via → File shares

Tips for accelerating data transfers

Speed: Optimal disk cache configuration & network to AWS

Monitor: Amazon CloudWatch metrics - CloudBytesUploaded

Upload status: Amazon CloudWatch events - NotifyWhenUploaded

Offline data migration

AWS Snow Family Portfolio

AWS Snowball Edge

Data transfer & edge compute

- 42/100TB storage capacity (S3)
- 10/25/40GE networking
- Data encryption end-to-end
- Rugged 8.5 G impact case
- Chain of Custody, Tamper Detection
- Rain and dust resistant
- EC2/AMI support for edge computing
- NFSv4 Server

AWS Snowmobile

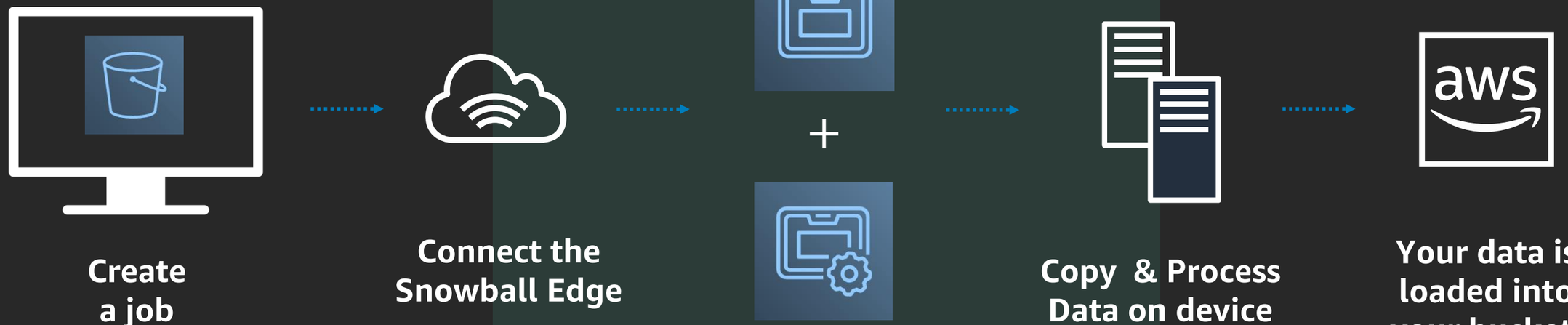
20+ PB data transfer

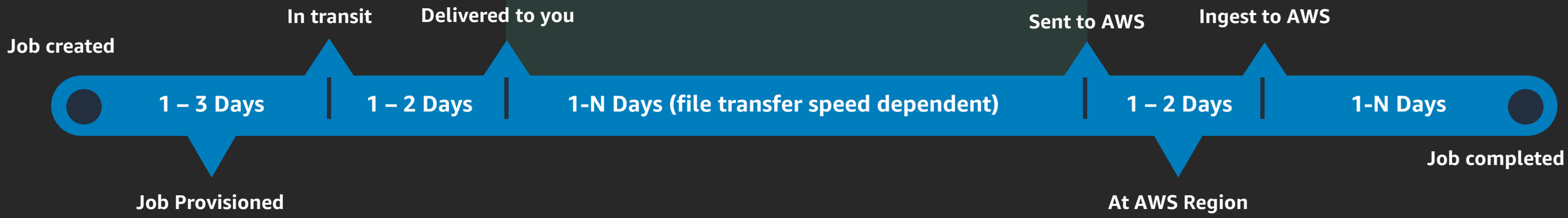
- Exabyte-scale storage in a 45ft container (90PB s3/Glacier/EBS)
- 10/25/40GE networking
- Data encryption end-to-end
- S3/Glacier Data import
- Dedicated security personnel
- GPS tracking, alarm monitoring, 24/7 surveillance, and optional additional security

AWS Snowball Edge usage scenarios

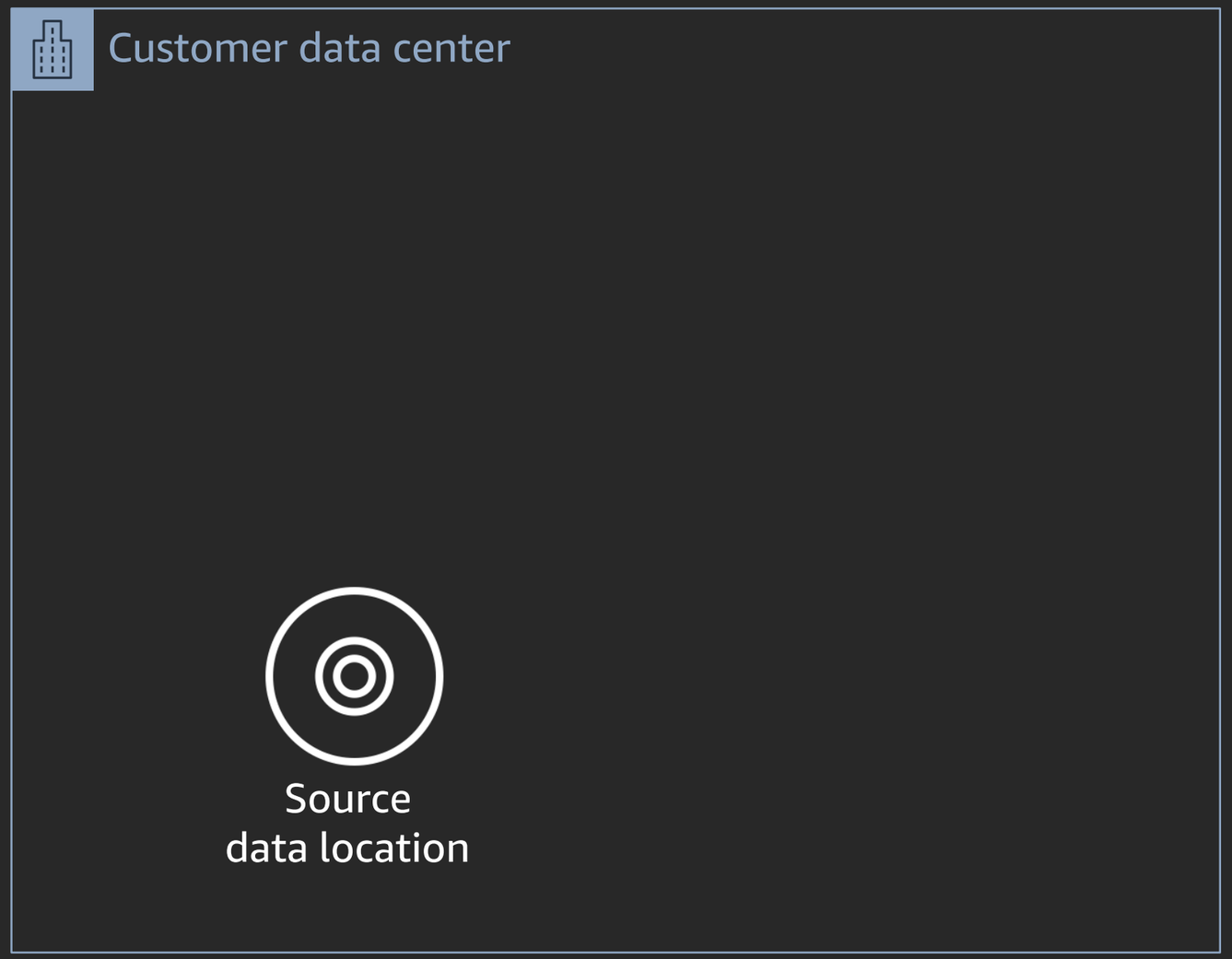
- Bulk data transfers - data not required immediately
- Bulk data transfers - limited network bandwidth
- Free up capacity on maxed-out on-premises storage
- Compute or data transformation at Edge

AWS Snowball Edge import workflow

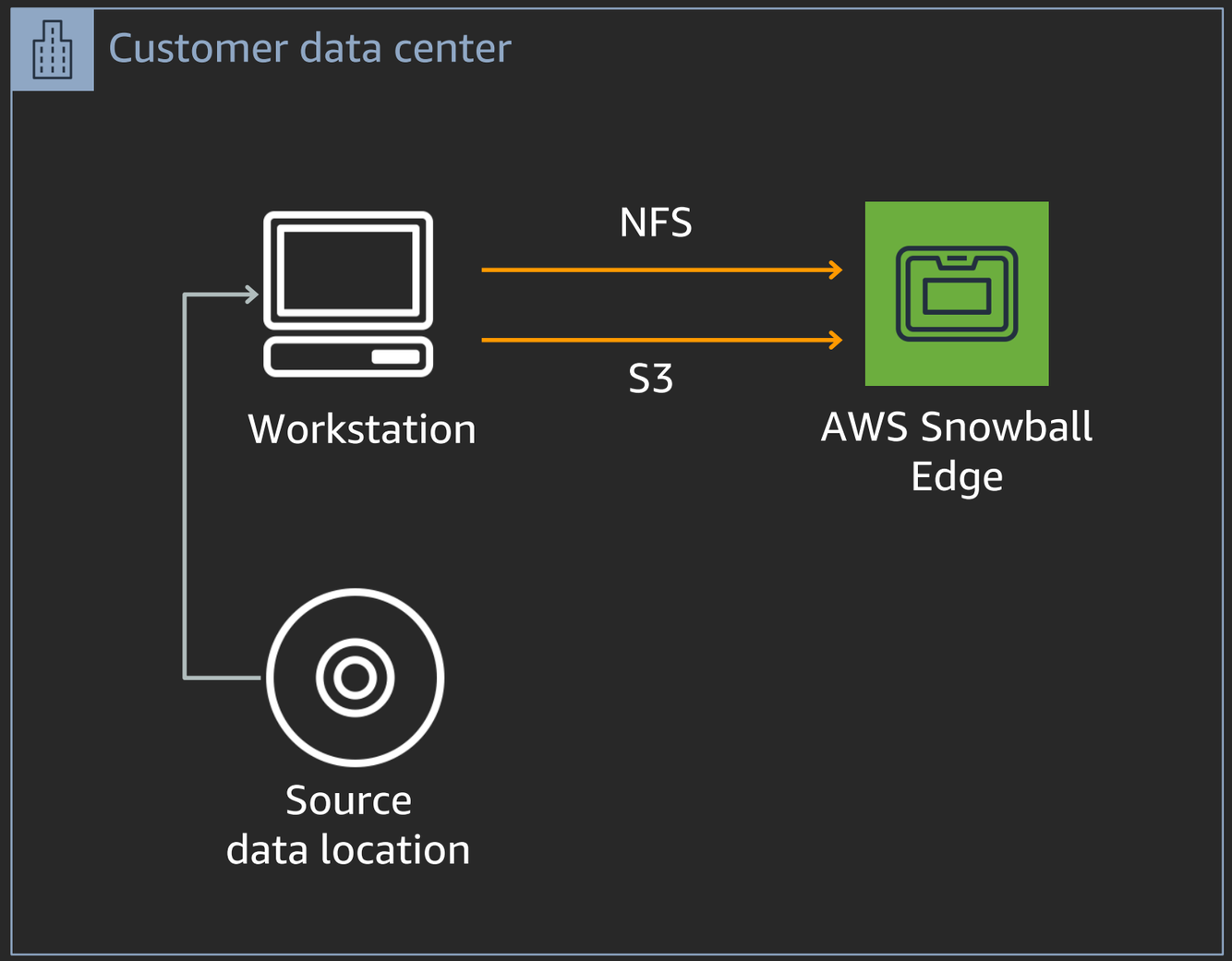




Ingesting data into AWS Snowball Edge



Ingesting data into AWS Snowball Edge



Tips for accelerating data transfer

Get up and running with AWS Snowball Edge Data Migration Guide

Define your tools and scripts

Understand your dataset characteristics

Aggregate small files into larger files

Utilise concurrent transfers sessions

When to use AWS DataSync

- Migration speed is a priority and network bandwidth is available
- Preserve file metadata and data integrity
- Integrate data transfer into existing workflows

When to use AWS File Gateway

- Seamlessly migrate data into Amazon S3 using file share interfaces
- Retain low latency access to cached hot-data
- Visibility to data stored in Amazon S3 through a file share

When to use AWS Snowball Edge

- Limited or no network bandwidth
- Bulk data transfers
- Compute at or data transformation at the edge

Thank you!

aws SUMMIT ONLINE