
AWS
re:Invent
NOV. 28 – DEC. 2, 2022 | LAS VEGAS, NV

STG210

Best practices for moving data to AWS, online and offline

Randy Boutin
General Manager, DataSync
AWS

Erik Durand
Snow Product Leader
AWS

Meriem Belhadj
Sr. Storage Specialist Solutions Architect
AWS

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

Data migration overview

Online vs. offline

AWS DataSync overview

AWS Snow Family overview

Using DataSync and Snow together

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Large-scale data transfers are challenging



- Building and deploying scripts
- Encrypting and verifying data
- Recovering from errors
- Handling network availability
- Assuring performance

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Streamline data and application migrations

AWS PROVIDES A RANGE OF DATA MIGRATION SERVICES MATCHED TO YOUR MIGRATION NEEDS



AWS DataSync



AWS Snow Family



AWS Application Migration Service



AWS Database Migration Service



AWS Storage Gateway

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Streamline data and application migrations

AWS PROVIDES A RANGE OF DATA MIGRATION SERVICES MATCHED TO YOUR MIGRATION NEEDS



AWS DataSync



AWS Snow Family

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Online vs. offline cloud data migration



Online

- Available bandwidth
- High bandwidth WAN or AWS Direct Connect
- Firewall configurations may be needed
- Data available immediately in AWS after it’s transferred



Offline

- Limited bandwidth
  Low bandwidth, bandwidth constrained, heavily utilized WAN, air-gapped
- Data center capacity for devices
- Device management
- Allow time for data import

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Large (500 TB+) data migration challenges

Time to transfer online

	1 Gbps	2 Gbps	5 Gbps	10 Gbps
500 TB	58 days	29 days	12 days	6 days
5 PB	2 years	1 year	116 days	58 days
10 PB	4 years	2 year	232 days	116 days

Use AWS Snow Family

- For offline data migrations
- Bandwidth limitations
- Flexible timeline

Use AWS DataSync

- For online data migrations
- Available bandwidth
- Strict timelines

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS DataSync
Overview

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What is AWS DataSync?

Online data migration service
that simplifies, automates, and accelerates
copying file and object data to and from AWS storage



Fast data transfer



Easy to use



Secure and reliable



Fully managed



Cost effective

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS DataSync use cases



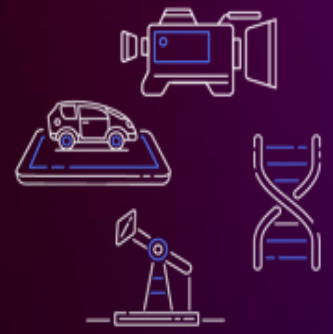
Migrate
application data
to AWS



Archive
on-premises
storage to free
up capacity



Replicate
for data protection
and recovery



Transfer
data for timely
in-cloud processing

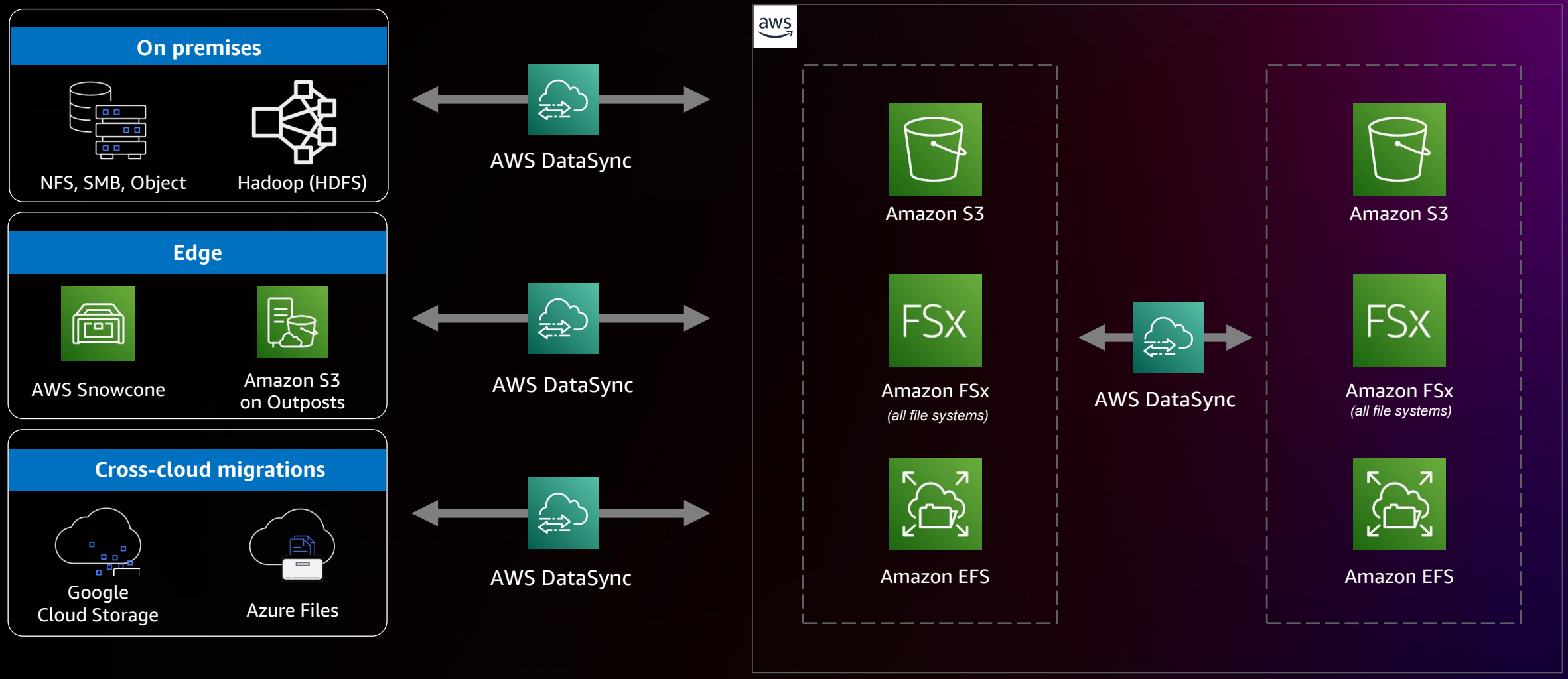
aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What can you do with AWS DataSync

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What can you do with AWS DataSync?

QUICKLY, SIMPLY, AND SECURELY MIGRATE YOUR DATA TO AND FROM AWS STORAGE SERVICES



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS DataSync Discovery (Preview)

Simplify and accelerate your migration to AWS



Gain insights into storage utilization

- ✓ Automated data collection
- ✓ Dashboards for aggregated views of data
- ✓ Find underutilized resources



Receive recommendations for AWS Storage services

- ✓ Select the right storage for your use case
- ✓ Optimize your AWS storage configuration
- ✓ Meet your performance needs while minimizing costs



Simplify migration planning

- ✓ Minimize time, effort, and costs
- ✓ Use estimated costs to inform your budget
- ✓ Validate assumptions before migrating

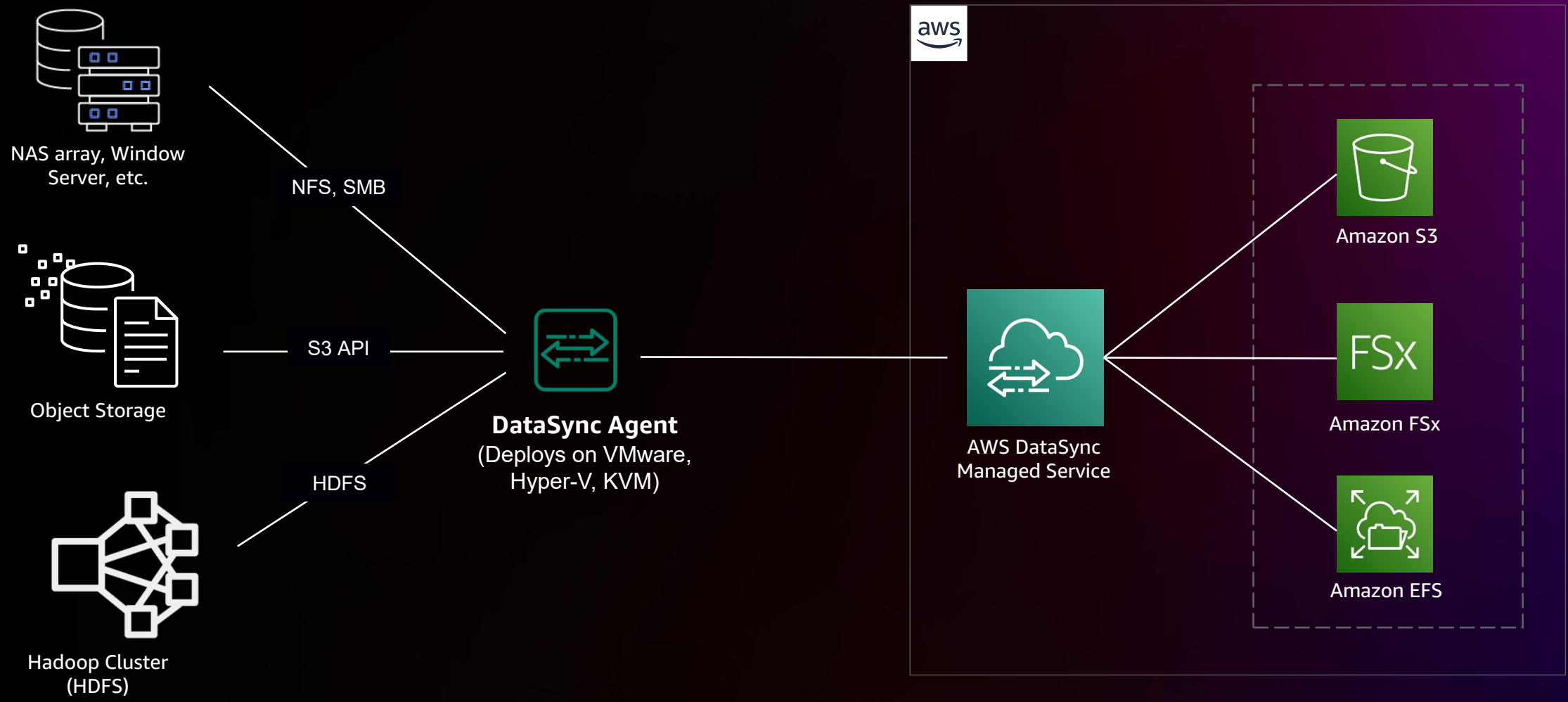
Learn more at
aws.amazon.com/datasync/discovery

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

On-premises transfers

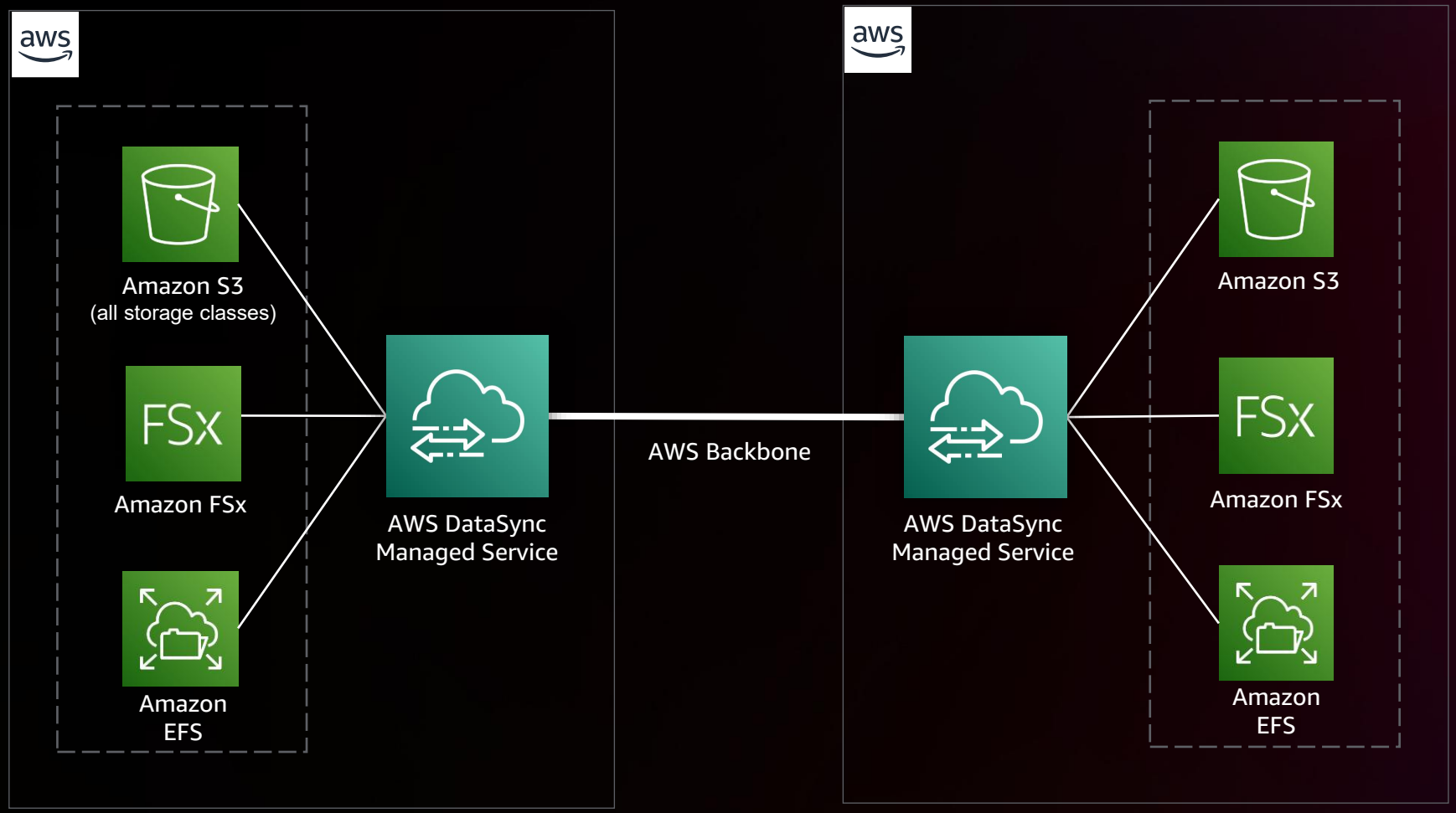
TRANSFER DATA BETWEEN ON-PREMISES STORAGE SYSTEMS AND AWS STORAGE SERVICES



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

In-cloud transfers

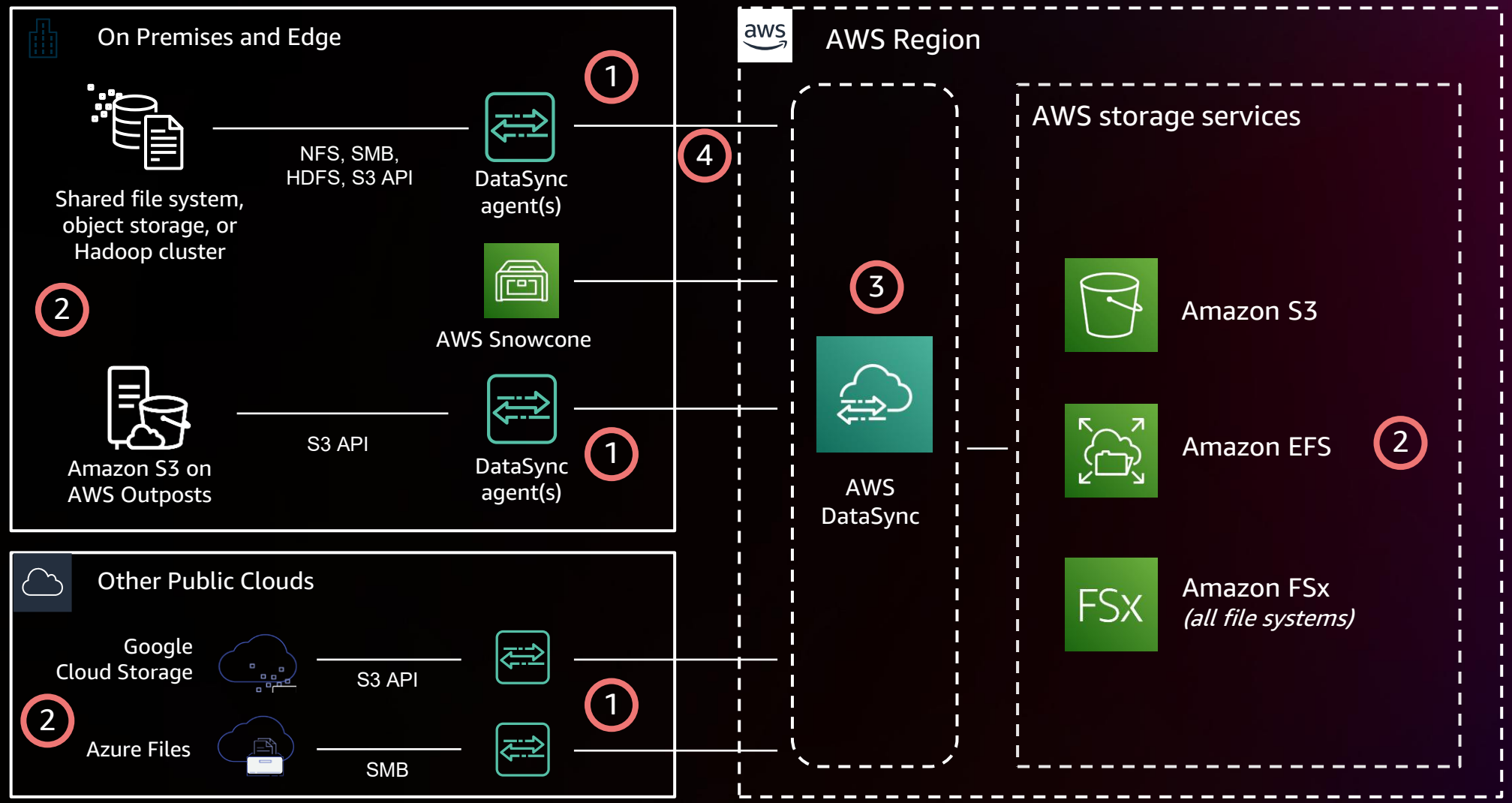
TRANSFER DATA BETWEEN AWS STORAGE SERVICES QUICKLY, EASILY, AND SECURELY



- ✓ No infrastructure to deploy or manage
- ✓ Copy data between any supported AWS Storage services
- ✓ Copy data within the same region or across regions
- ✓ All traffic stays within the AWS network
- ✓ All traffic encrypted in flight using TLS 1.2

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

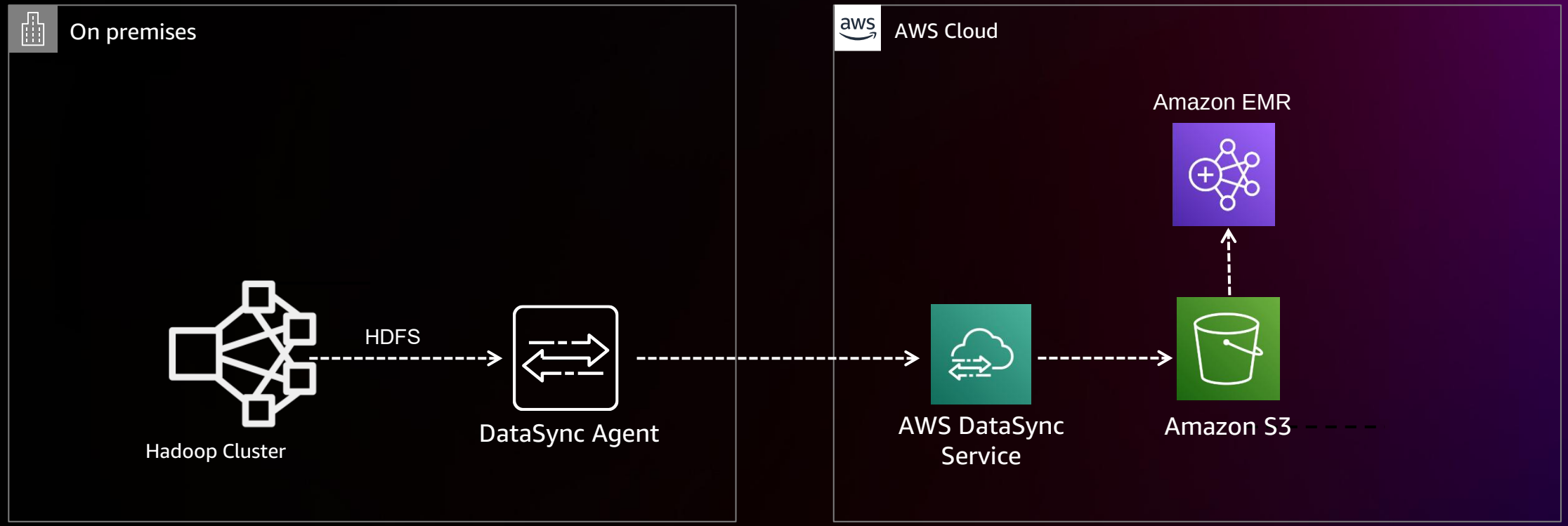
AWS DataSync: How it works



1. Agents are deployed to connect to on-premises storage, edge locations, or storage in other clouds
2. Create location resources for source and destination storage
3. AWS DataSync managed service connects to AWS storage and coordinates the transfer
4. For transfers between on premises and AWS, internet, AWS Direct Connect, and AWS Virtual Private Network (VPN) are supported

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Move your data to the cloud to get more insights



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

TMAP Mobility transferred 2.4 PB of Hadoop data using AWS DataSync



"With AWS DataSync, TMAP Mobility Co., Ltd. was able to quickly and reliably transfer PB scale Cloudera HDFS data within 2 months. AWS DataSync's native support for HDFS helped us meet our tight project launching schedule."

-Eun Hee Kim
Data Intelligence Group Leader

Problem

TMAP Mobility (S. Korea) needed to transfer 2.4 PB of on-premises Cloudera HDFS to a new data lake in Amazon S3 within 2 months. The new data lake would support an existing mobile application with more than 5 million daily active users.

The new data lake project would leverage real-time and batch processing with an AWS managed services–based analytics pipeline.

Solution

TMAP Mobility selected AWS DataSync as their petabyte-scale data movement solution for its performance, reliability, and ease-of-use. Integrated capabilities such as task monitoring, data encryption, data integrity validation, and native support for HDFS, further reduced project risk and simplified operations.

Outcome

Using AWS DataSync, TMAP Mobility was able to meet their timeline for migrating their data into their new Amazon S3 data lake. As a fully managed service, AWS DataSync made it easy for operators, with no prior experience with the service, to configure and perform the migration within a short 2 month time window.

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Snow Family
Overview

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What is AWS Snow Family?

Move petabytes of data to and from AWS,
or process data at the edge.



AWS
Snowcone



AWS
Snowcone SSD



AWS
Snowball

<1TB ——— 10s PB

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The AWS Snow Family purpose

Offline Data Movement

Providing secure and fast data collection, processing, and data movement to AWS cloud, primarily offline but also online

&

Rugged Edge Compute

Bringing AWS services and capabilities to the disconnected or partially connected edge



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The AWS Snow Family key value propositions



Offline Data Movement

Snow Family devices help migrate data when bandwidth is limited or non-existent



Secure

Snowball Edge devices use tamper-evident enclosures, 256-bit encryption, and industry-standard Trusted Platform Modules (TPM)



Run partially or fully disconnected

Snow devices and capabilities fully operate without network connectivity



Scalable

Add or remove Snow devices based on your needs



AWS compute services and capabilities at the edge

Snow offers core AWS services like EC2, EBS, S3, and IAM and AWS AI/ML capabilities like Greengrass and Lambda



Cloud compatibility

Use the same AWS service APIs that are used in the cloud

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Snowball Edge for Rugged and Tactical Edge

BRINGS AWS TO DDIL RUGGED EDGE AND LOCATIONS WITH NO OR LIMITED NETWORK CONNECTIVTY

Edge Compute and Edge Storage use cases:
32 - 104 vCPU, 208 - 416 GB RAM
28 TB NVMe or 80TB HDD storage
Up to 100G Networking

Tamper-resistant
Tamper-evident

- Handles up to **200G impact**
- Airdrop-able
- MIL-S-901D
- Meets FISMA High, FedRAMP ITAR, CJIS, and DoD SRG Impact Level 6 requirements



Operates on unconditioned power and in non-datacenter environments (0° to 45°C)

- Data secured with 256-bit encryption
- Designed to be operated by non-technical personnel

Weighs 49 lbs, can be handled by a single person

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Snowcone for mobile workloads

SMALL, ULTRA-PORTABLE, RUGGED, AND SECURE DDIL EDGE COMPUTING AND DATA TRANSFER DEVICE

Capabilities:
2 vCPU, 4 GB RAM
8 TB HDD or 14 TB SSD
1 or 10G Networking

Tamper-resistant
Tamper-evident



Operates on unconditioned power and in non-datacenter environments (0° to 45°C)

Handles vibration up to 50G RMS

- Data secured with 256-bit encryption
- Designed to be operated by non-technical personnel

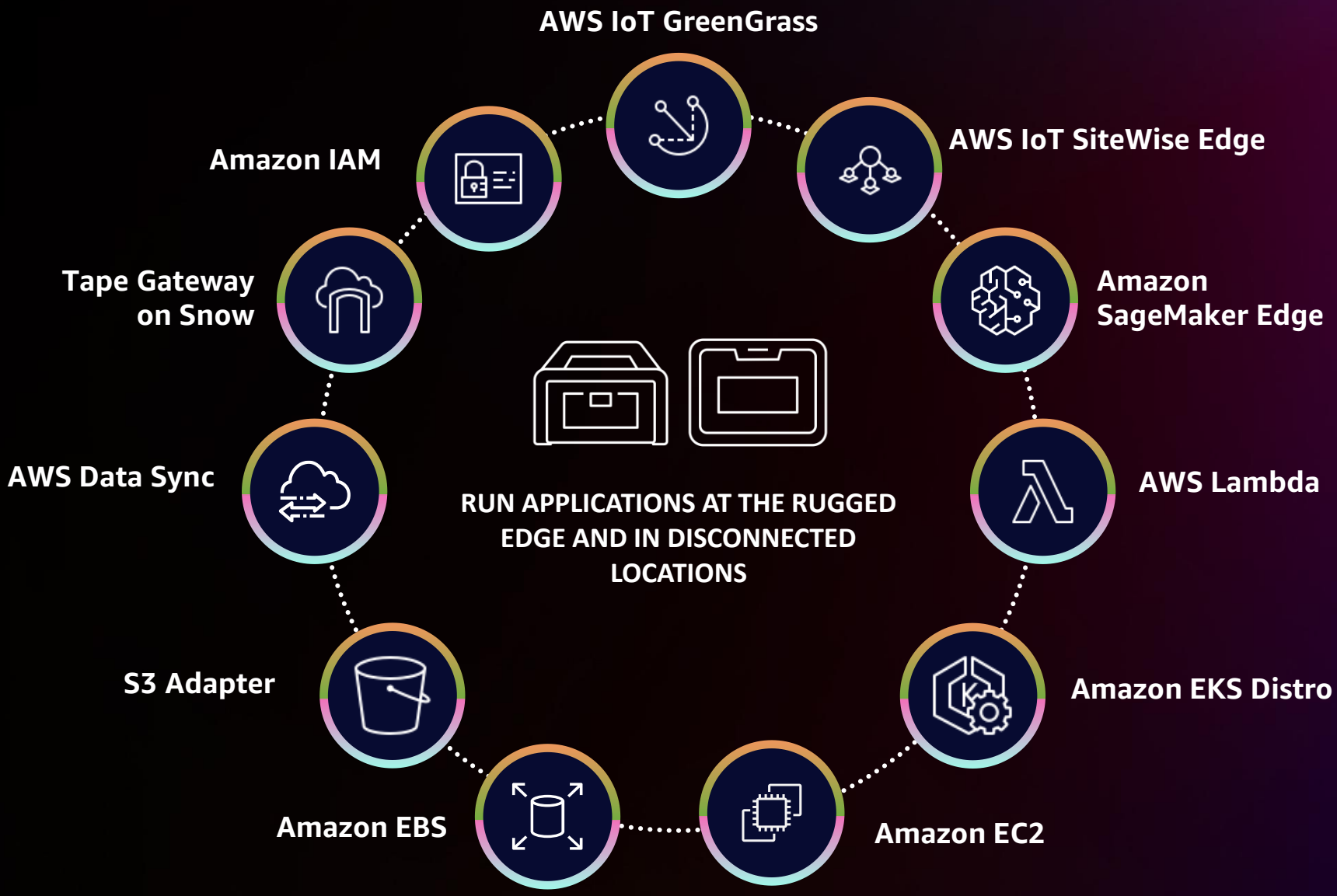
Weighs 4 lbs and can be carried in a backpack

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Use cases: M&E, Industrial IoT & Automotive use

AWS Services on Snow Family

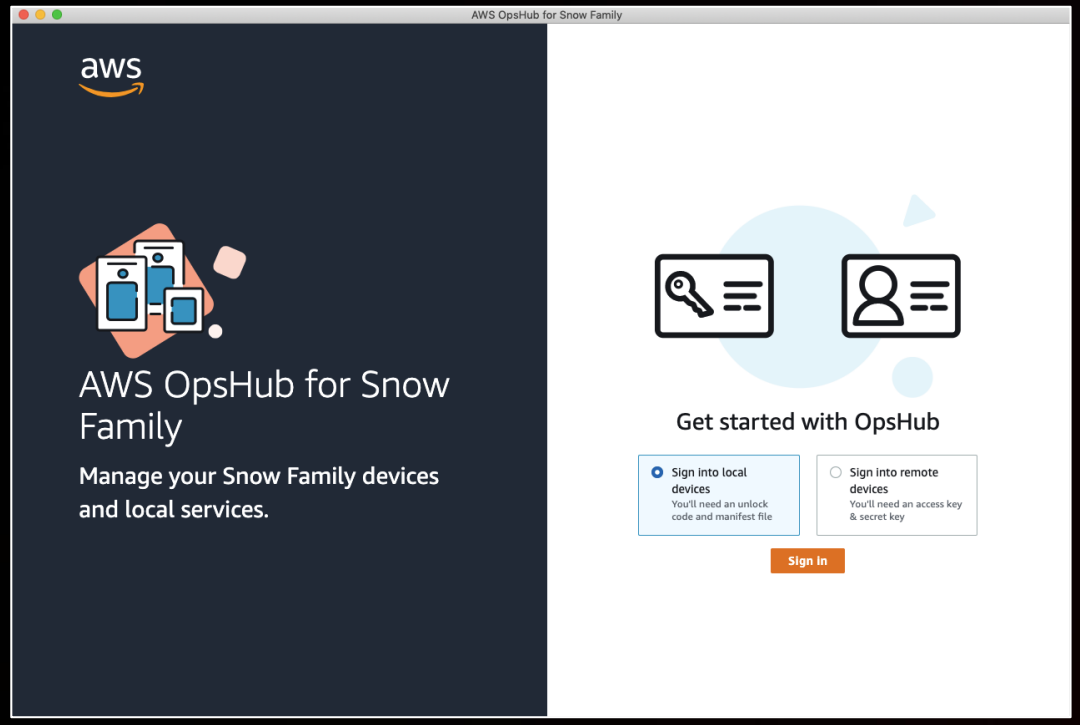
FAR EDGE

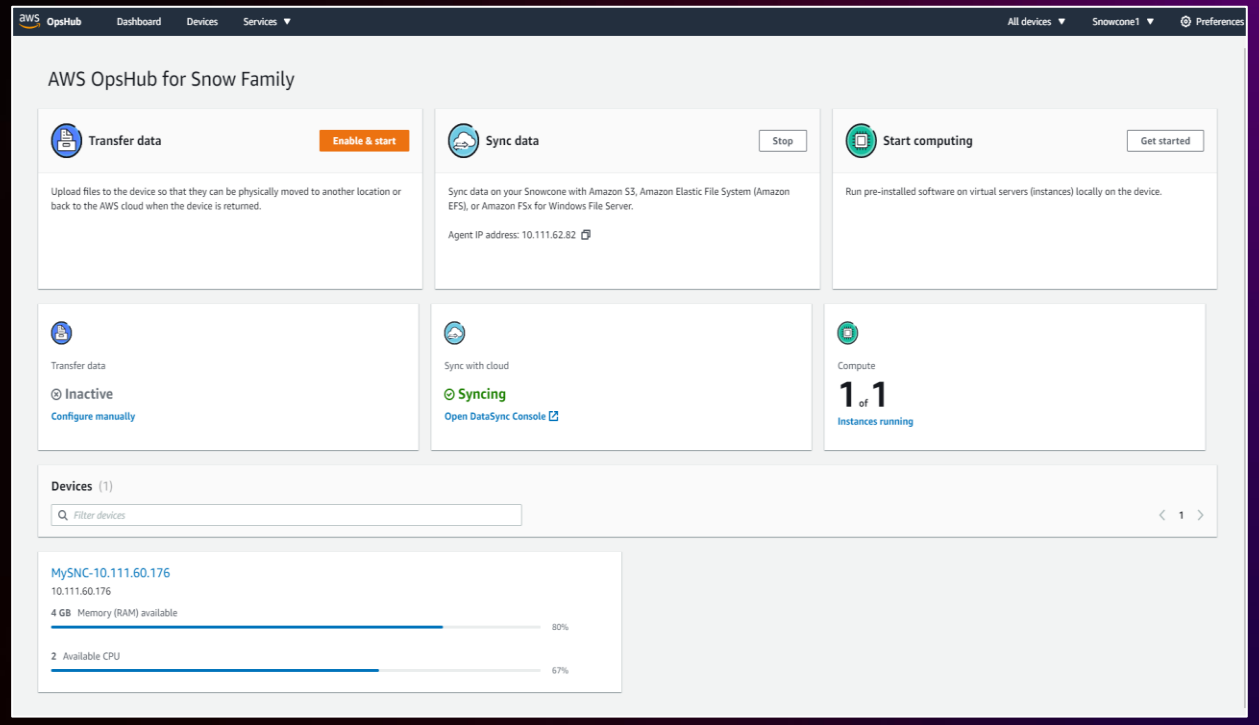


© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS OpsHub for Snow Family

GUI APP FOR CUSTOMERS TO MANAGE AWS SNOW FAMILY DEVICES





Visit the AWS Snowcone resources page to learn more about AWS OpsHub - https://aws.amazon.com/snowcone/resources

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Snow Family for offline data migration

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Snow Family data migration workflow



AWS Snow Family management console

Qualify SBE V3S then Create SBE V3S job

Job created & provisioned

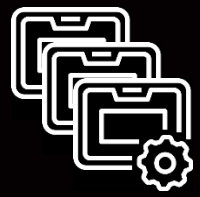
1-3 days



Courier transport service

Device shipped

1-2⁺ days



Use AWS OpsHub or AWS CLI to manage AWS Snow Family devices

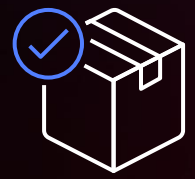
Device arrives



Copy data via >1GB/s local network

Data copied

1-N* days



Courier transport service

Device sent to AWS

1-2⁺ days

Device arrives at AWS



Your data begins ingesting into your Amazon S3 bucket

1-N days



Secure erasure using NIST 800-88 standards

Job Complete

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

* variable file transfer speeds | + business days

Joyn migrates 3.4PB to Amazon S3 with AWS Snowball Edge

joyn

Challenge

- High maintenance costs
- 3 month migration goal not attainable with online migration
- Need for increased flexibility and availability

Solution

- Chose to use multiple AWS Snowball Edge Storage Optimized concurrently
- Leveraged Amazon S3 Intelligent-Tiering optimized storage expenses.

Results

- Moved 3.4 petabytes (PB) in 90 days
- $500K+ savings per year by closing on-premises data center
- No increase in total cost of ownership (TCO)

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

joyn

“Moving our archive using AWS Snowball Edge was fast, simple and secure, but we’re also impressed by how cost-effective it was.”

Stefan Haufe

Media Engineer, Joyn

aws

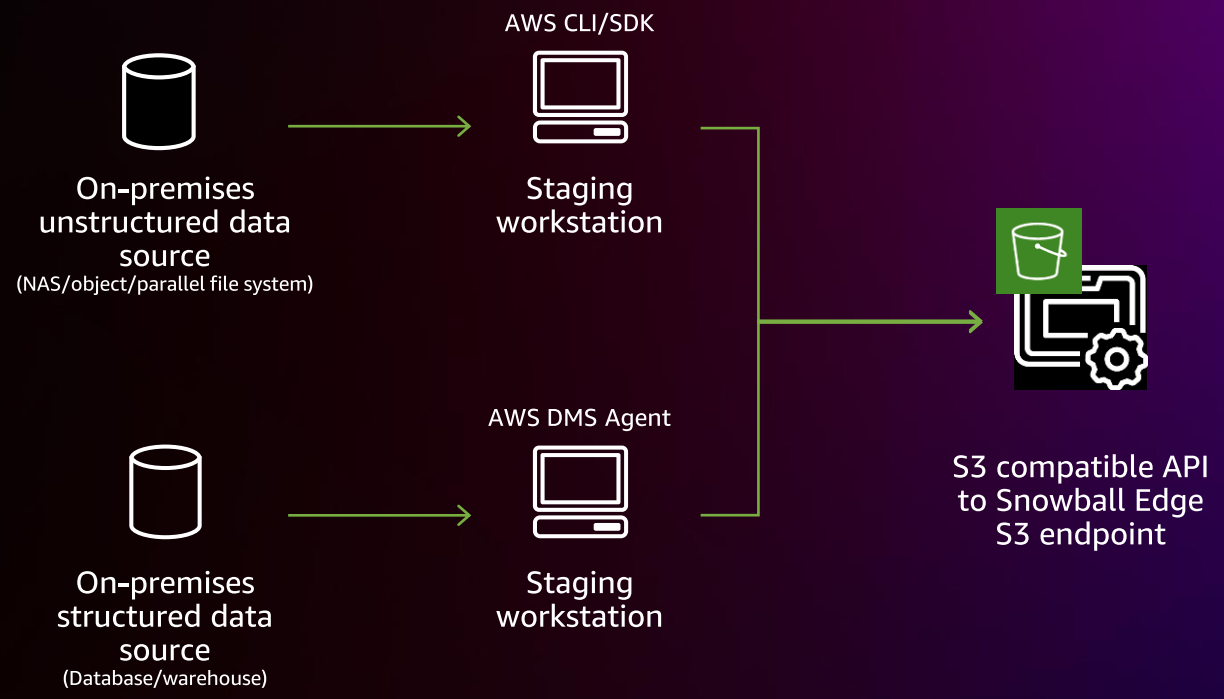
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Large-scale migrations with AWS Snowball

LEARNED FROM MULTI-PETABYTE CUSTOMER MIGRATIONS

Run a proof-of-concept

- Discover and remediate environmental issues early
- Deploy staging workstations
- Batch small files (< 1 MB) and ensure larger files (>5 MB)
- Benchmark and optimize data transfer (target 300 MB/s – 500 MB/s)



Plan devices and scheduling with your account team / TAM before ordering jobs

- White paper: AWS Snowball Edge Data Migration Guide
- Blog: Data Migration Best Practices with Snowball Edge

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Snow Large Data Migration Manager

Plan, execute, and monitor your large data migration projects



Snow Large Data Migration Manager

Provides an easy way to plan data migration projects that require multiple Snow devices

Plan multiple Snowball and Snowcone jobs based on your data transfer needs

Gain real-time status and insights into Snow jobs, logistics, and data ingestion progress

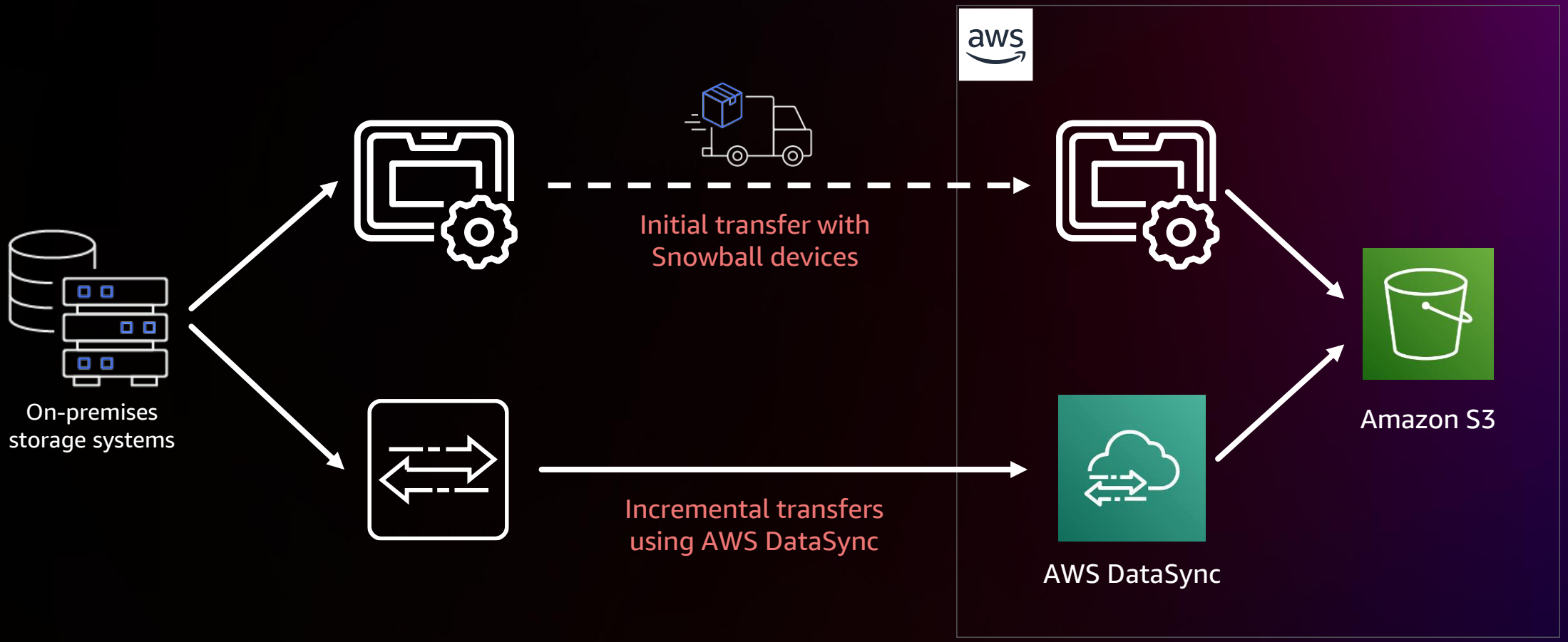
aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Using DataSync and Snow together

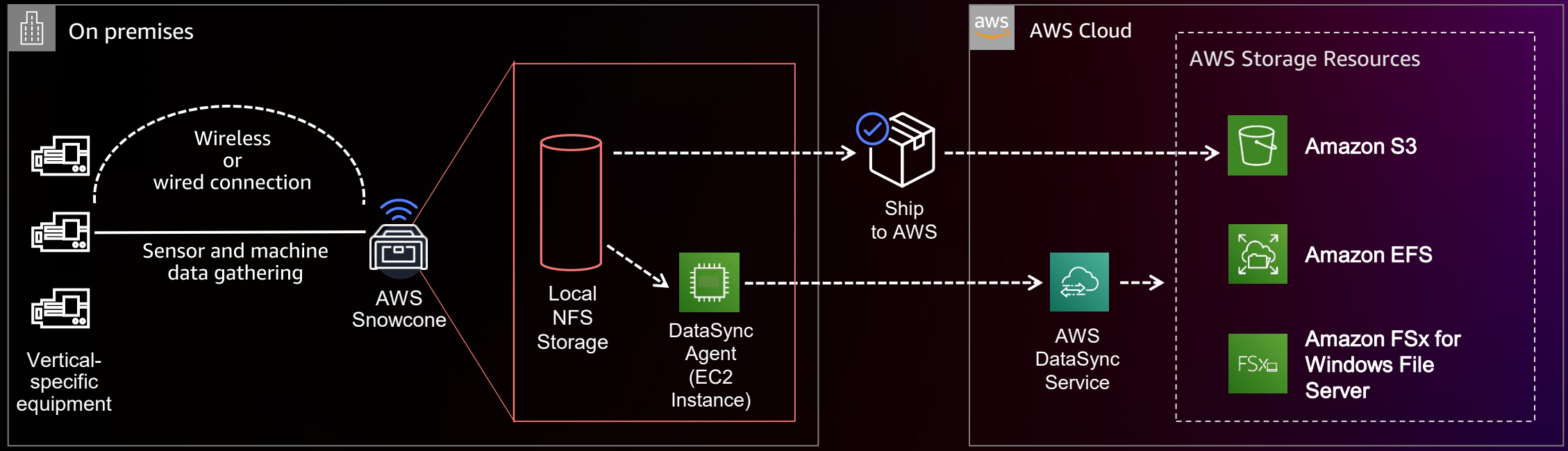
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Large migrations with AWS Snowball and AWS DataSync



aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Ongoing workflows with AWS Snowcone and AWS DataSync



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Wrap up

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Key takeaways

- Evaluate goals and resources for your cloud migration strategy
- Consider your migration mix between online versus offline
- Separate data by archival versus production to streamline migration project
- Plan your migration using available tools and your AWS team

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Resources



Get started with DataSync
aws.amazon.com/datasync



Get started with Snow Family
aws.amazon.com/snowball

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

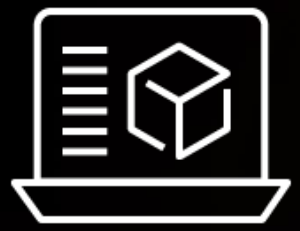
Additional Sessions

- STG211 - Hybrid cloud storage & edge compute: AWS, where you need It
- STG214 - Accelerate secure data migrations at scale with AWS DataSync (Replay)

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Continue your AWS Storage learning

Build a learning plan



Set your AWS Storage Learning Plans via **AWS Skill Builder**

Increase your knowledge



Use our **Ramp-Up Guides** to build your storage knowledge

Earn AWS Storage badges



Demonstrate your knowledge by achieving **digital badges**

aws.training/storage

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

Randy Boutin
rboutin@amazon.com

Meriem Belhadj
mebelhad@amazon.com

Erik Durand
eddurand@amazon.com



Please complete the session survey in the **mobile app**

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.