
AWS
re:Invent

STG218

Accelerate your migration to Amazon S3

Avi Drabkin

Senior Storage Specialist SA

AWS

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

Introduction

Data migration challenges

AWS Storage Gateway and AWS DataSync introduction

Customer use cases

Additional resources

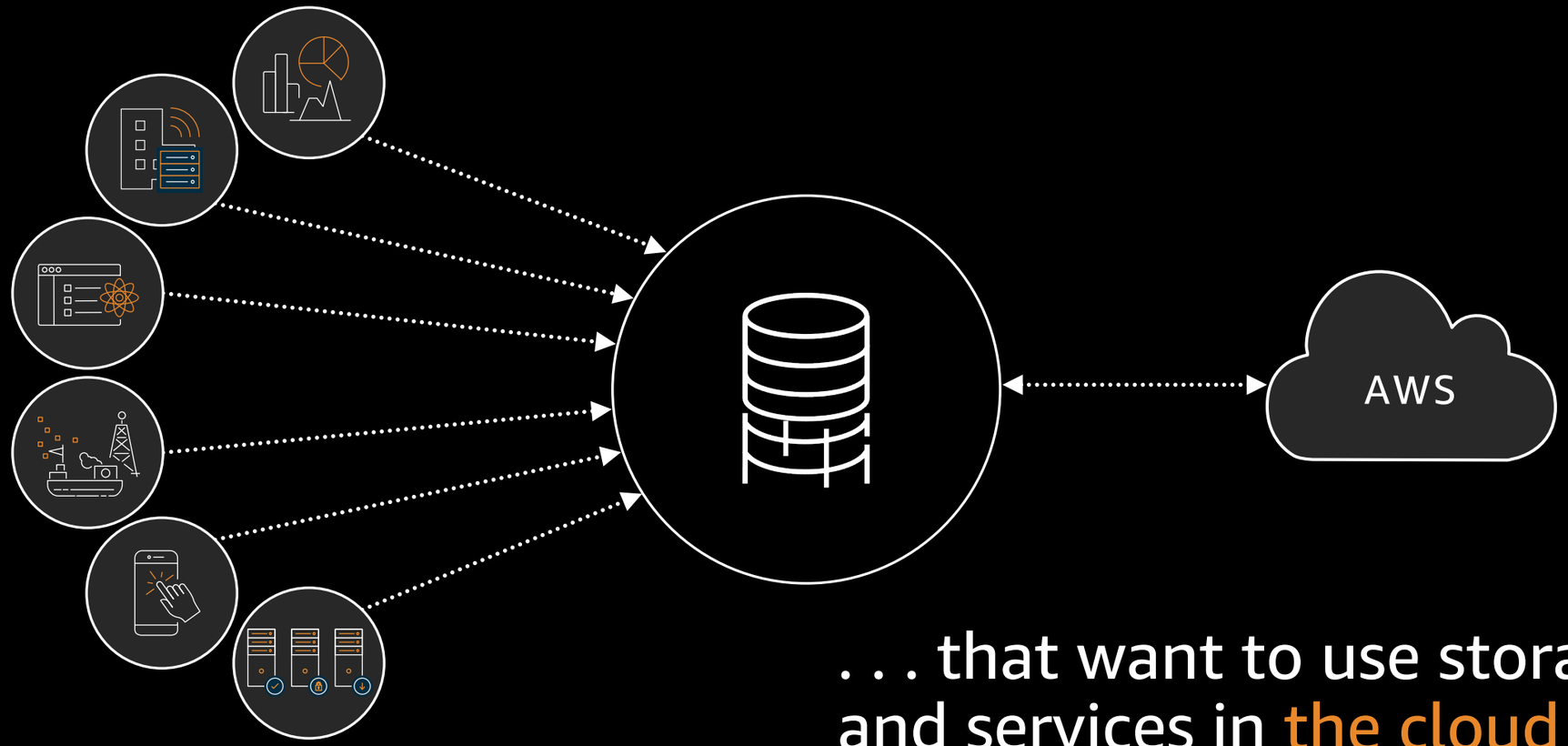


© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

You have on-premises data and applications . . .

Existing applications and data archives

New data sources



. . . that want to use storage and services in the cloud

https://www.idc.com/getdoc.jsp?containerId=prUS46286020

Data migration challenges

Migrating and copying data at scale can be complex and time consuming

DIY tools are out there but typically require a lot of work to automate

Using AWS DataSync along with AWS Storage Gateway can help simplify and accelerate your migration to Amazon S3

Introduction to AWS Storage Gateway and AWS DataSync



AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Storage Gateway overview



File gateway

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

File-based applications work without change



Volume gateway

Block storage on premises backed by cloud storage with local caching, Amazon EBS snapshots, and clones, integrated with AWS Backup

SAN-like with cloud recovery



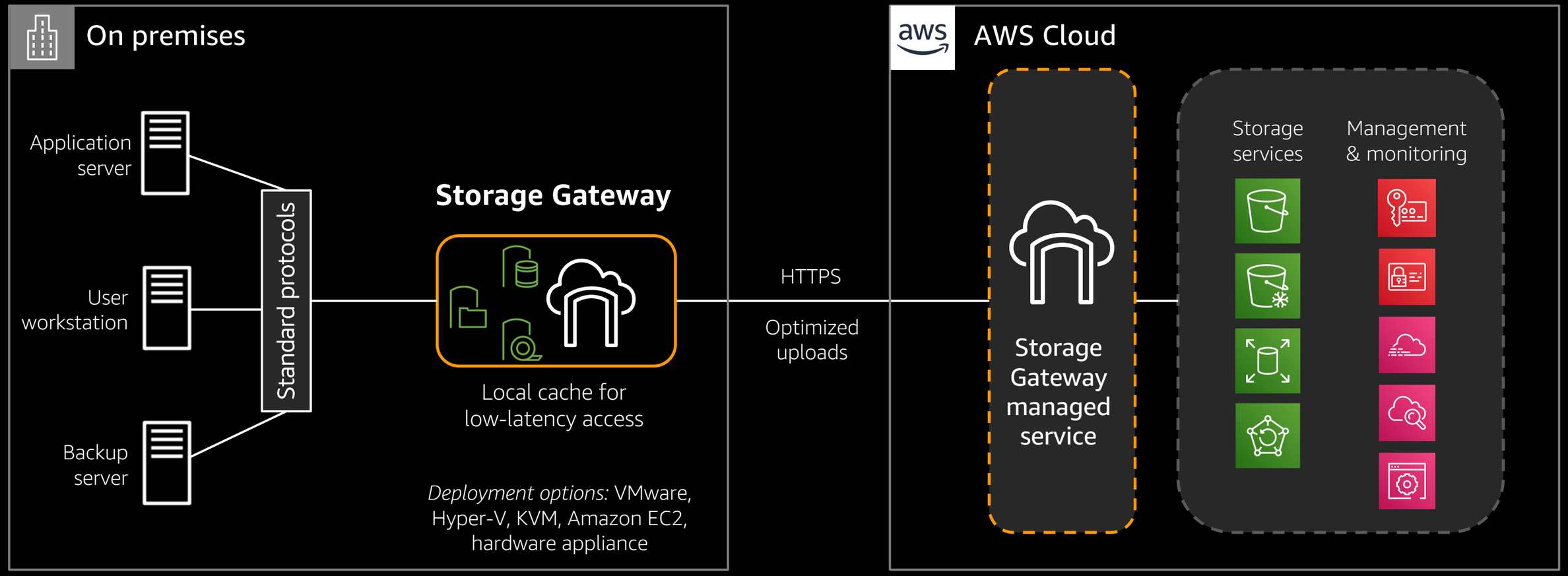
Tape gateway

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

Easily switch tape backups to AWS

How does AWS Storage Gateway work?

ON-PREMISES ACCESS TO VIRTUALLY UNLIMITED CLOUD STORAGE



Storage Gateway is available on a variety of form factors

	Virtual machines (VMs)	Amazon EC2	Hardware appliance
	VMware ESXi, Microsoft HyperV, Linux KVM		
Where	Corporate data centers	In AWS	Remote locations, branch offices
When	Leverage existing VM infrastructure	In-cloud data retrieval / access / DR	Out-of-the-box, simplified experience

AWS DataSync



Online data transfer service that simplifies, automates, and accelerates moving data to and from AWS Storage Gateway



Fast data transfer



Easy to use



Secure and reliable

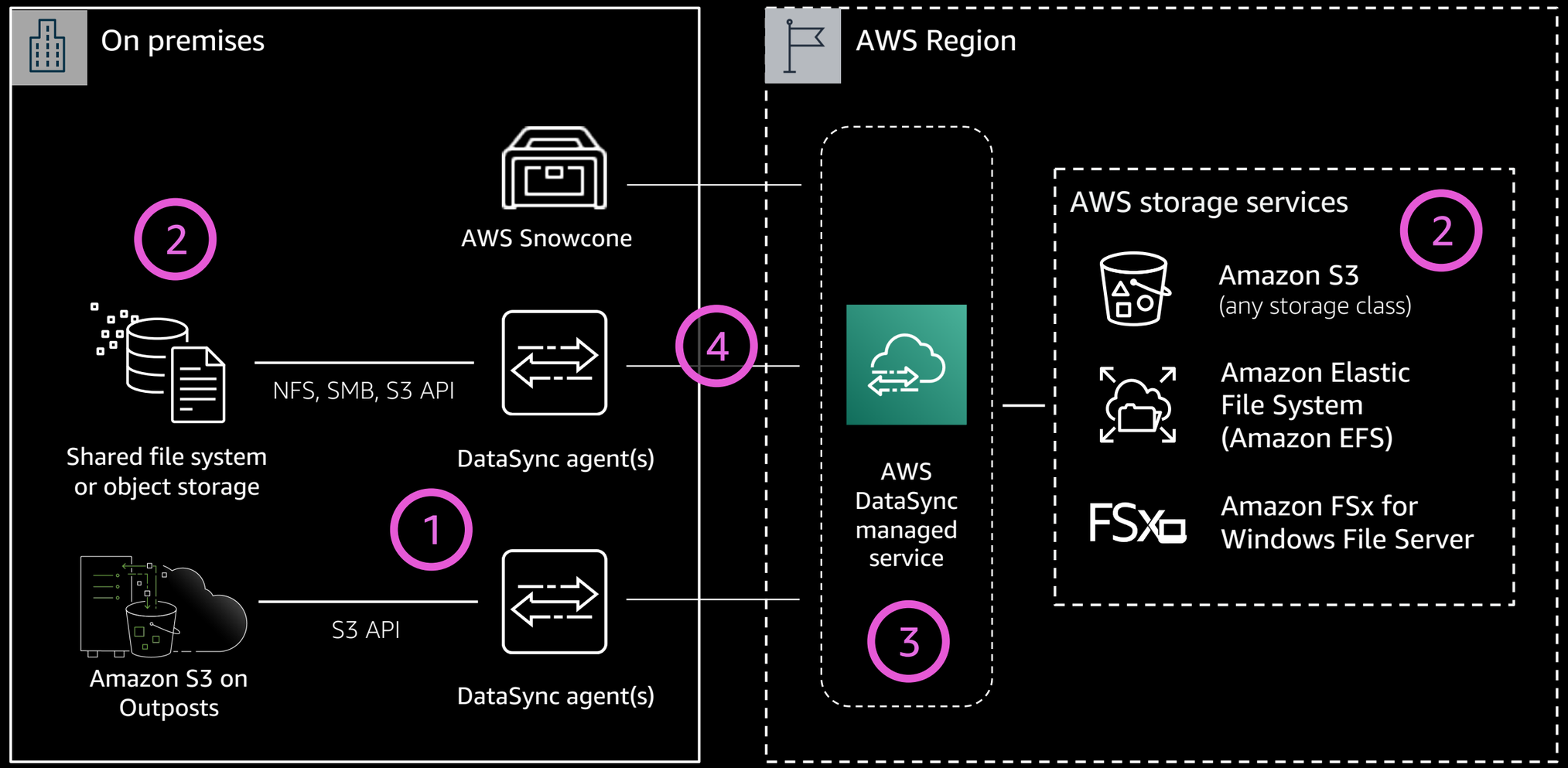


Fully managed



Cost-effective

How does AWS DataSync work?



1. Agents are deployed to connect to on-premises storage
2. Locations control how DataSync connects to storage
3. DataSync managed service connects to AWS storage and coordinates the transfer
4. For transfers between on premises and AWS, internet, AWS Direct Connect, and VPN are supported

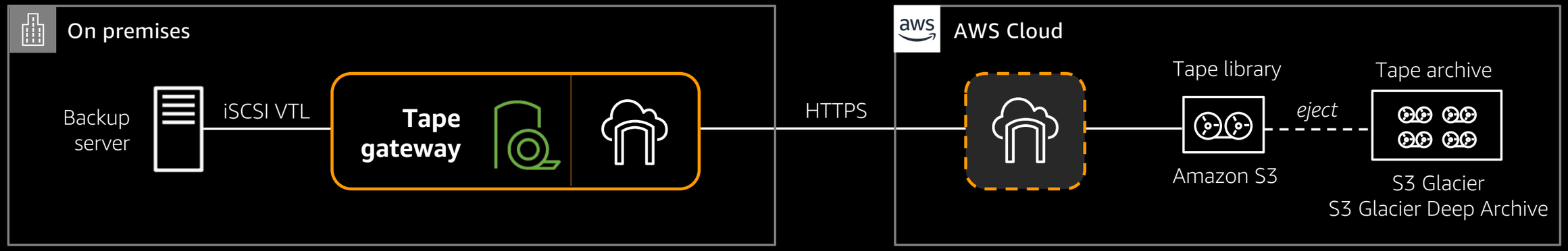
Customer use case scenarios



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Customer use case: Frictionless backup to Amazon S3 Glacier

MODERNIZE TAPE WORKFLOWS AND LEVERAGE S3 GLACIER AS A BACKUP ARCHIVE



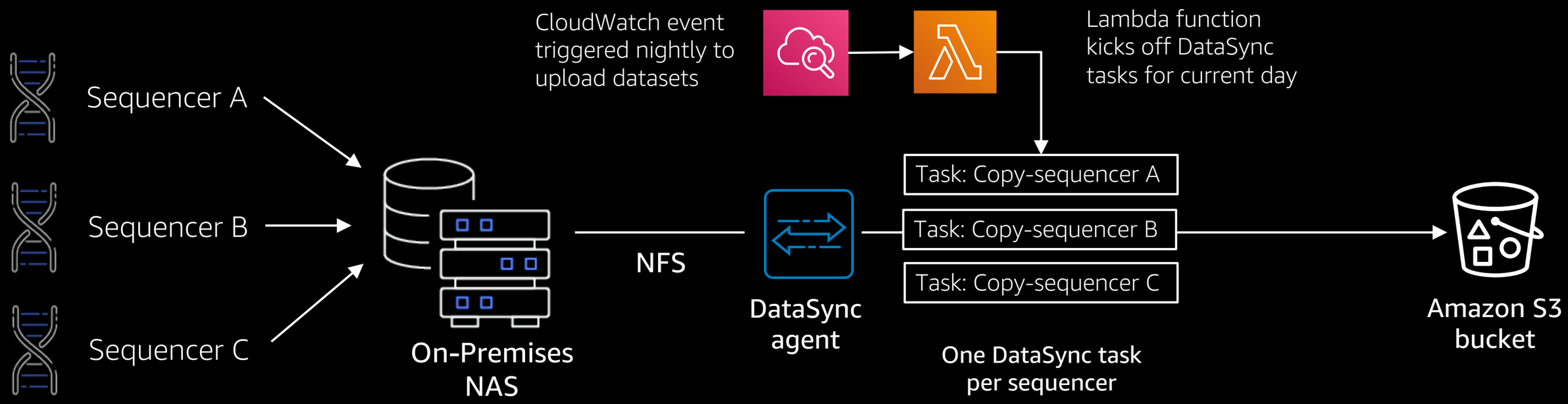
Goal – Eliminate the pain of managing physical tape libraries, including tape inventory management, offsite tape storage, and recovering from damaged tapes

Challenge – Seamless integration with existing backup software

Solution – Tape gateway to store backup data in S3 Glacier or S3 Glacier Deep Archive

Customer use case: Use AWS DataSync to scale on-premises workloads

ENABLING GENOMICS PROCESSING AND ANALYSIS IN THE CLOUD WITH AWS DATASYNC



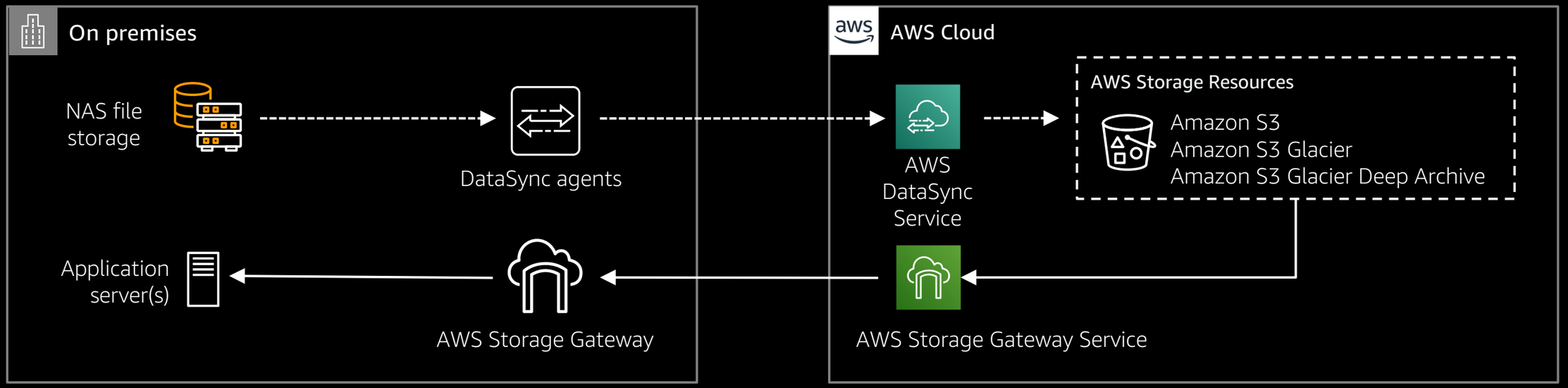
Goal – Send output of daily sequencer jobs to AWS for processing

Challenge – Keeping operational costs in line with new data coming in every day

Solution – AWS DataSync to transfer from on-premises NFS to S3

Customer use case: Migrate large datasets to Amazon S3

REPLACE AGING ON-PREMISES EQUIPMENT WITH A HYBRID CLOUD SOLUTION



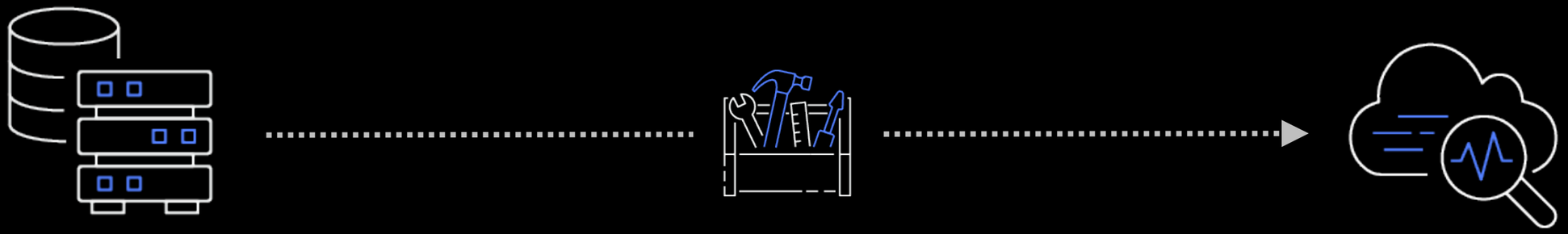
Goal – Replace aging NAS storage holding petabytes of archive data; some application servers still need to access data infrequently

Challenge – Complicated folder structure on source NAS; destination layout to be different than the source layout; need to get data off NAS as fast as possible

Solution – AWS DataSync to migrate data from on-premises NFS to S3; use AWS Storage Gateway to provide local cache for data stored in S3; deprecate old NAS storage

Can I use a free / open-source data transfer tool?

DATA TRANSFER AT SCALE CAN QUICKLY BECOME COMPLEX AND TIME-CONSUMING



With DIY transfers, you own:

- Management of performance and throughput
- Network and system complexity
- Security and data validation
- Scaling out multiple transfers
- Staff time and resources

COX AUTOMOTIVE™

"To be able to say, 'My data synchronization utility is too fast!' is fantastic! It buys you that agility to be able to move faster and not have your projects drag on for months and months."

Dan Boisvert
Lead Software Engineer

Problem

- Stacks of disk arrays on premises were expensive and required a lot of space
- Complex architecture and cache hierarchy
- Many readers via NFS

Solution

- Use AWS DataSync to move data to S3
- Active / active multi-region and versioning with lifecycles
- Storage Gateway for local access to cloud data

Outcome

- Saved ~85% on storage per location
- Storage engineers focused on high-value, strategic activities

Santos Limited

"Prior to learning about AWS DataSync, I had spent over a month creating, testing, and iterating on new scripts to get our backup files into Amazon S3. These scripts needed to be deployed on all of our servers and were difficult to centrally manage. By transitioning to DataSync, I was able to simplify and automate my backup management. Apart from its simplicity, DataSync also provided additional functionality, such as monitoring and error checking, that was quite valuable. It took less than one hour of my time to set up and get going, and we use DataSync every night to transfer 5–50 TB of data, which is roughly 3,000 files. For the task that I needed it for, DataSync was absolutely spot on."

Gavin Boyce
Cloud Solution Architect

Problem

- Lack of consistent backup of business-critical databases
- Complex set of scripts presented too much management overhead and lacked proper error checking

Solution

- Use AWS DataSync to simplify data transfer to S3
- Remove overhead of managing complex backup scripts

Outcome

- Saved months of development
- Transferred up to 50 TB of data per night
- Provided monitoring and verification of data transfer

Where can I get more information?



AWS Storage Gateway
aws.amazon.com/storagegateway



AWS DataSync
aws.amazon.com/datasync

Check out the following AWS storage blogs:
https://aws.amazon.com/blogs/storage/escaping-the-pain-of-physical-tape-management-with-aws-storage-gateway/
https://aws.amazon.com/blogs/storage/migrating-and-managing-large-datasets-on-amazon-s3/
https://aws.amazon.com/blogs/storage/protect-your-file-and-backup-archives-using-aws-datasync-and-amazon-s3-glacier/

Dive deeper into AWS DataSync and AWS Storage Gateway with a hands-on workshop:
https://github.com/aws-samples/aws-datasync-migration-workshop

Thank you!



AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Please complete the session survey

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.