



Quickly and securely move your files to AWS for migration or data protection using AWS DataSync

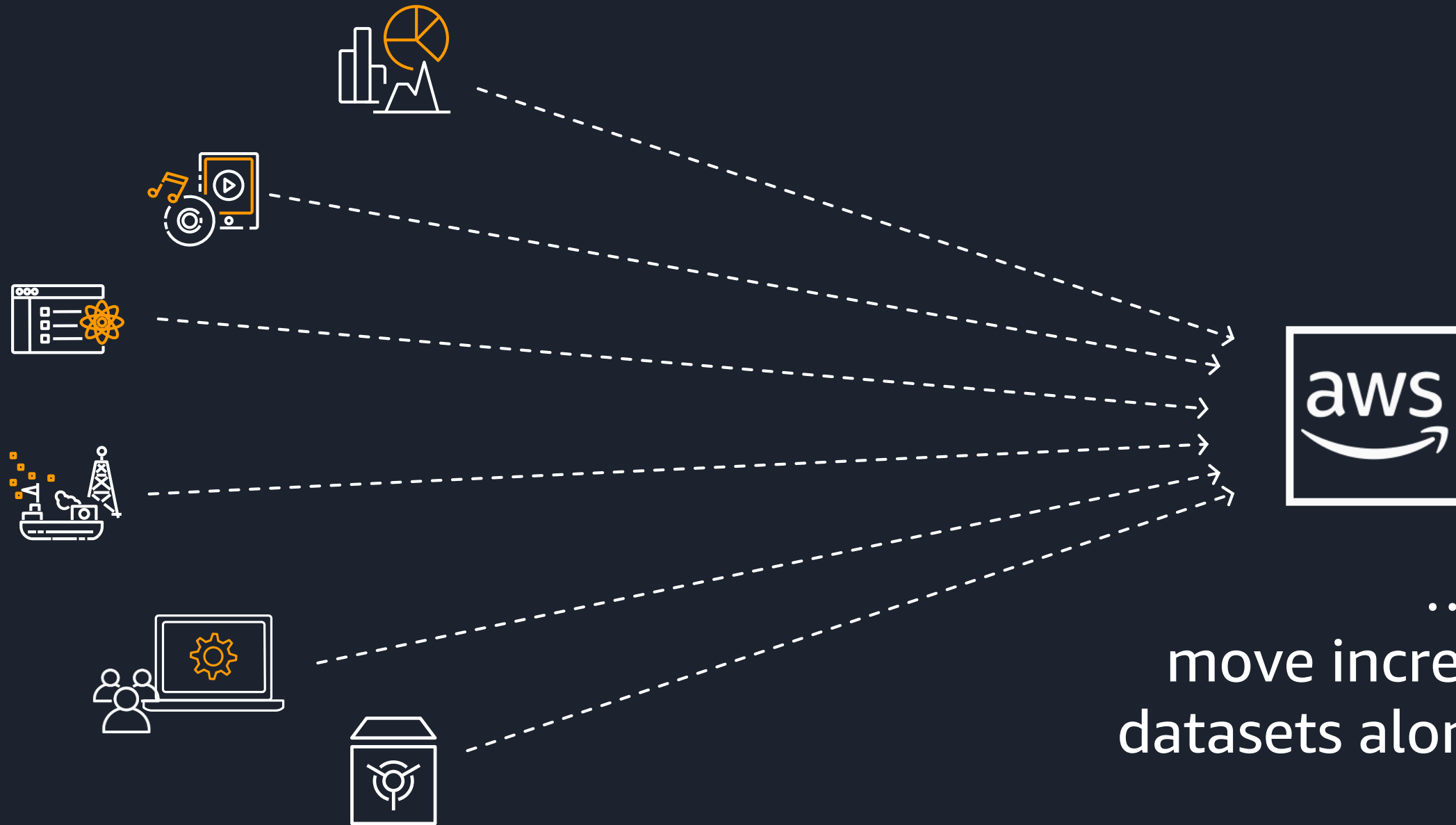


Jeff Bartley

Principal Solutions Architect

Hybrid Cloud Storage and Data Transfer Services

As more and more workloads move to the cloud...



...you need to
move increasingly large
datasets along with them

Challenges with DIY transfer tools

Data transfer at scale can quickly become complex and time-consuming



With DIY transfers, you need to take care of:

- Managing performance
- Recovering from network issues
- Data encryption and verification
- Incremental transfers
- Bandwidth consumption
- Staff time & resources

...and more

AWS DataSync

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



Fast data transfer



Easy to use



Secure and reliable



Fully managed



Cost-effective

Agenda

- Introduction
- Use cases
- Demo
- Best practices
- Get started

Introduction

AWS DataSync

Copy data between on premises and AWS or between AWS Storage services

On premises



Shared file system
or object storage

DataSync Agent
(VMware, Hyper-V, KVM)



AWS Snowcone



Amazon S3 on Outposts

DataSync Agent
(EC2 on Outposts)

AWS



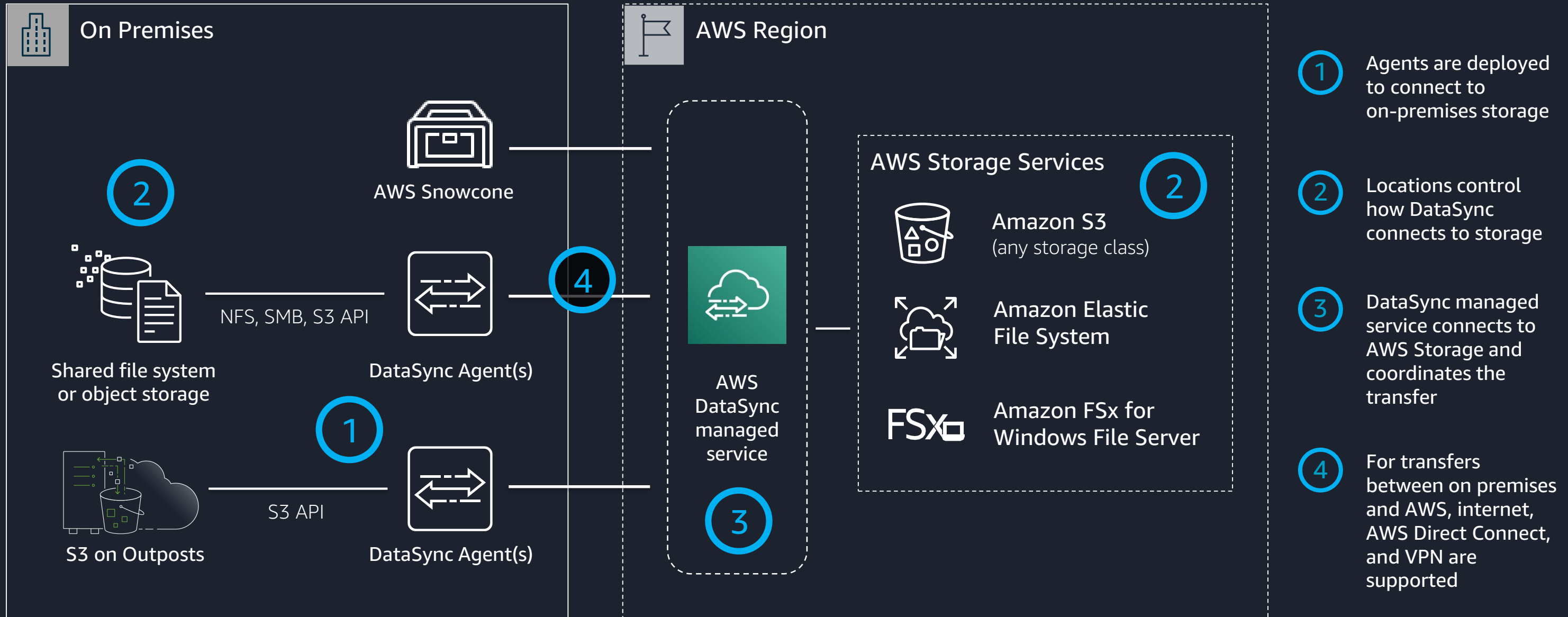
AWS Storage services
(S3, EFS, FSx for
Windows File Server)

AWS Storage services
(S3, EFS, FSx for
Windows File Server)



Amazon S3
(any S3 storage class)

AWS DataSync: How it works



AWS DataSync benefits



Fast data transfer

- Highly optimized, parallel network transfer
- Transfers only incremental changes
- Scale-out agents for increased performance



Easy to Use

- Schedule transfers
- Configure bandwidth limits
- Filter by file name patterns



Secure and Reliable

- Secure access to AWS Storage services
- End-to-end encryption
- End-to-end data verification
- VPC and FIPS endpoints



Fully managed

- Integrates with AWS management and monitoring services
- Direct transfer into all S3 storage classes

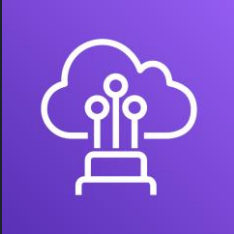


Cost-effective

- Pay only for data transferred
- \$0.0125/GB or \$12.50/TB

Deep integration with AWS services

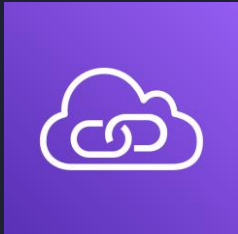
Networking



AWS Direct Connect



Amazon VPC



AWS PrivateLink

Monitoring and automation



Amazon CloudWatch



AWS CloudTrail

Security



Amazon IAM



Amazon S3



Amazon S3 Glacier



Amazon S3 on Outposts



Amazon Elastic File System



Amazon FSx for Windows File Server



AWS Snowcone

Storage services

Available in 22 AWS Regions

AMER

- US West (Oregon)
- US West (N. California)
- US East (N. Virginia)
- US East (Ohio)
- AWS GovCloud (US-West)
- AWS GovCloud (US-East)
- Canada (Central)
- South America (Sao Paulo)

EMEA

- Europe (Ireland)
- Europe (Frankfurt)
- Europe (London)
- Europe (Paris)
- Europe (Stockholm)
- Europe (Milan)
- Middle East (Bahrain)
- South Africa (Cape Town)



APAC

- Asia Pacific (Hong Kong)
- Asia Pacific (Mumbai)
- Asia Pacific (Seoul)
- Asia Pacific (Sydney)
- Asia Pacific (Singapore)
- Asia Pacific (Tokyo)

Use cases

AWS DataSync use cases



Migration of
active application
data



Archiving to offload
on-premises storage
capacity



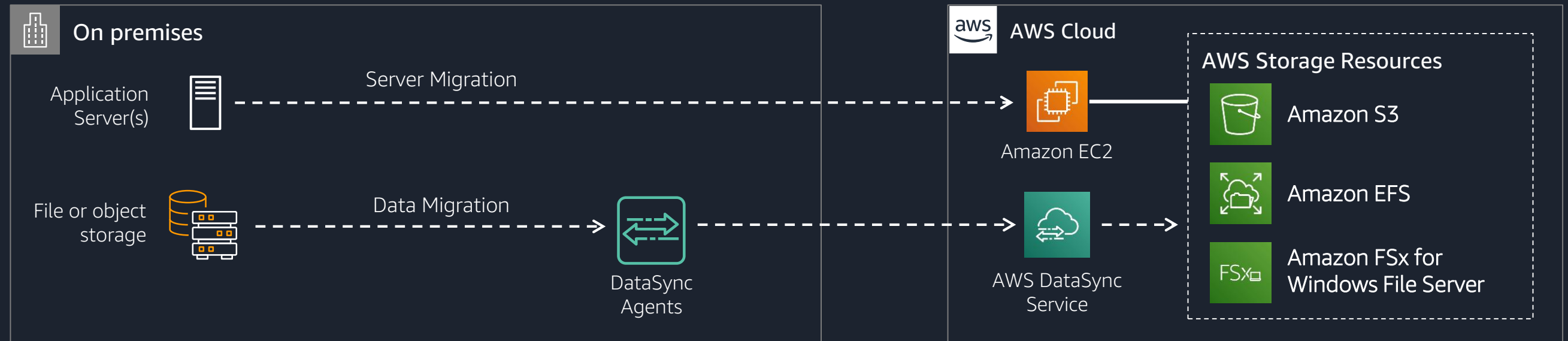
Replication for data
protection and
recovery



Transfers for
timely in-cloud
processing

Migrate active application data

Move your application data to the AWS cloud



Benefits

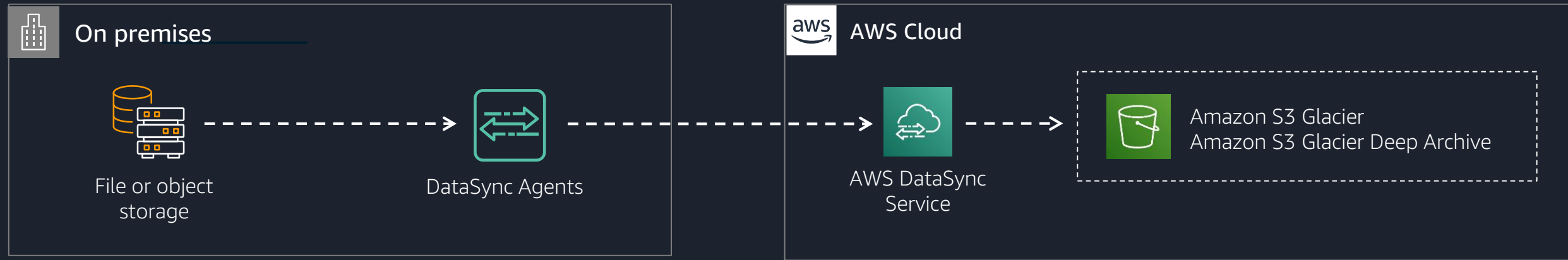
Move billions of files and PBs of data to AWS Storage quickly and securely

Data and metadata encrypted in-flight and verified byte for byte

Bandwidth throttling for efficient network utilization

Archive data to the cloud

Move cold data from expensive on-premises storage systems to durable, secure long-term storage



Benefits

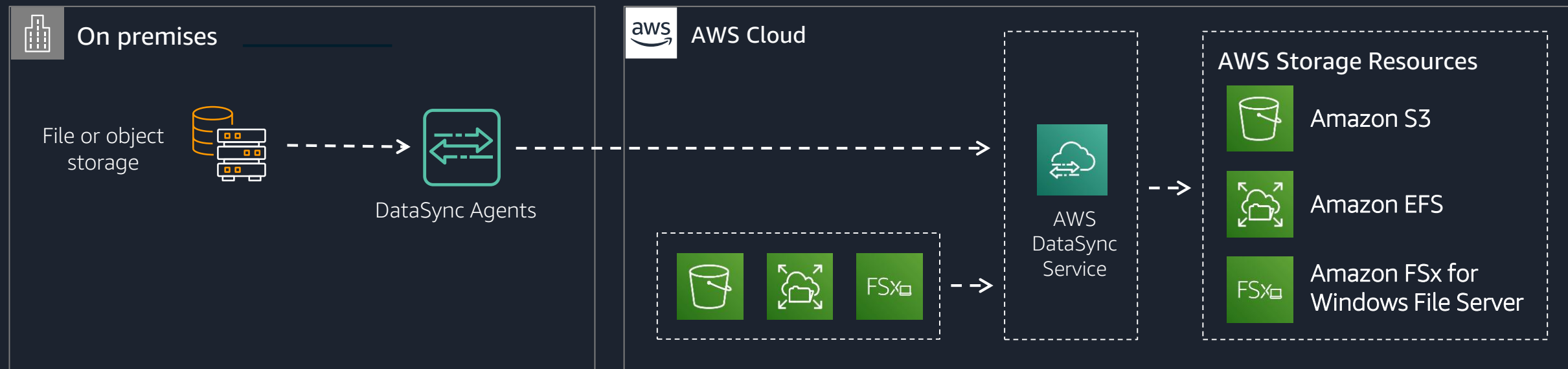
Transfer directly to the most cost-effective S3 storage class – no lifecycle required

Free-up space on expensive on-premises storage systems

Schedule regular transfers to meet compliance requirements

Replication for data protection

Make a copy of on-premises or in-cloud data to durable, secure, cost-effective storage



Benefits

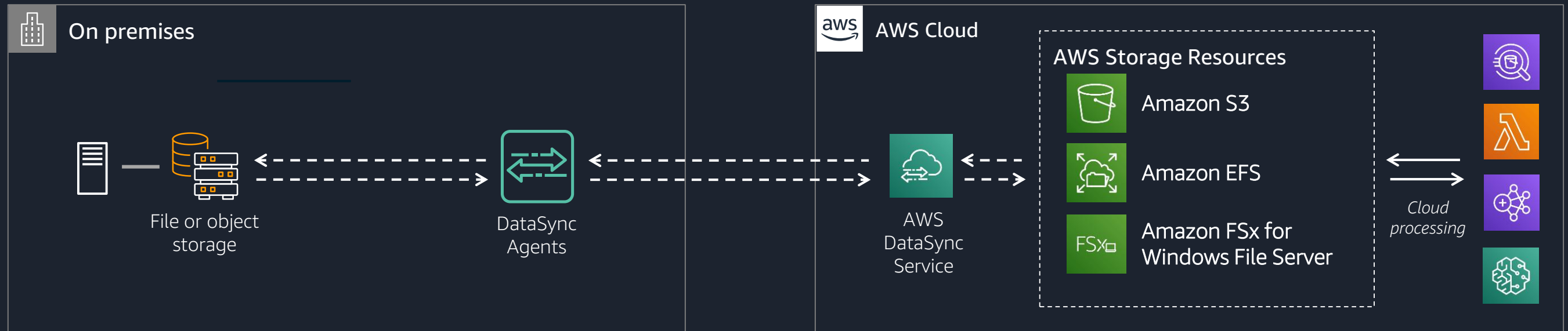
Copy data to S3 Glacier or S3 Glacier Deep Archive for low cost and high durability

Replicate EFS and FSx for Windows file systems across regions

Schedule tasks for ongoing replication

Transfer data for in-cloud processing

Accelerate hybrid cloud storage workflows



Benefits

Easily and securely move your data to where it is needed

Burst data to the cloud for analytics, machine learning, and further processing

Use filters to transfer only the files required for each workflow

Autodesk migrates their on-premises archive to Amazon S3 using AWS DataSync

Challenge

- Wanted to retire multi-petabyte on-premises Data Domain storage system
- Data retention policies required data to be retained for many years

Solution

- Used Amazon S3 for low cost, pay-as-you-go model as well as versioning support
- Used AWS DataSync to move data to S3

Benefits

- Successfully transferred dataset to S3 with full byte-for-byte verification
- Decommissioned on-premises Data Domain



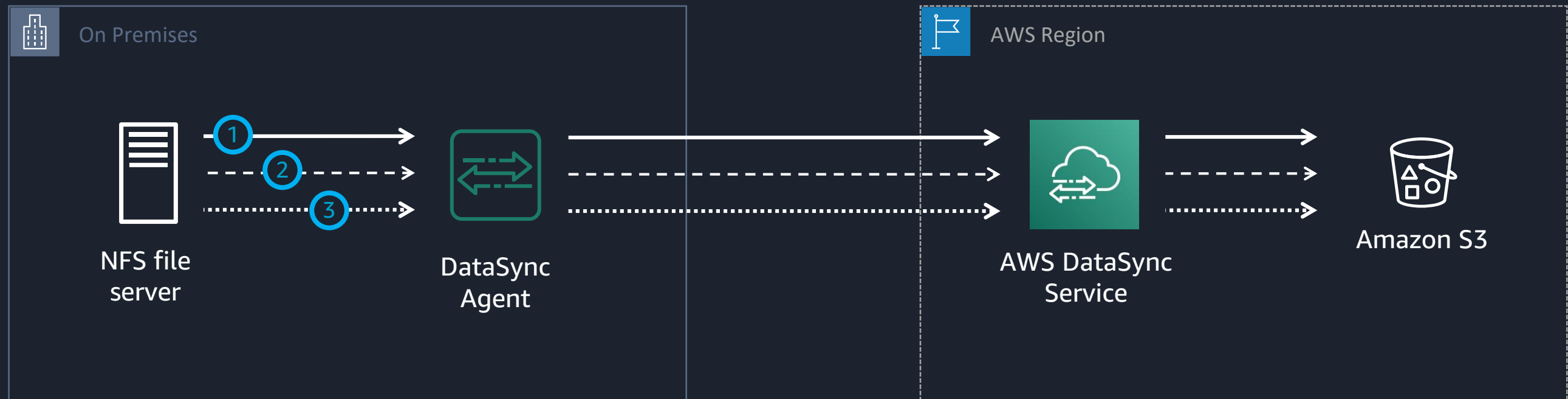
“ Our petabyte scale data migration journey from on-premises to AWS was accomplished swiftly with minimal effort and was completely self-managed with AWS DataSync. **This solution is a game changer!** ”

- Satish Kumar, Infrastructure Engineer, Autodesk

Read the AWS Storage Blog: [“Migrating hundreds of TB of data to Amazon S3 with AWS DataSync”](#)

Demo

Demo



- ① Initial migration of data from NFS to S3, with exclude filters
- ② Incremental copy of new data from NFS to S3
- ③ Using include filters to limit scope

Filtering to exclude files and folders

- Exclude files and folders that you never want a task to transfer
e.g. don't transfer files with .tmp or .temp extension

Filtering configuration - optional

When no filters are specified, the entire contents of the source location are transferred.

Exclude patterns

Files, folders, and objects with the specified patterns are excluded from the transfer. The pattern path is relative to the source location path. For example, /my-folder is a folder directly under the task's source location. Specifying /my-folder excludes the folder itself and all of its contents).

[Learn more and see syntax examples](#)

<input type="text" value="*.tmp"/>	<input type="button" value="Remove pattern"/>
<input type="text" value="*.temp"/>	<input type="button" value="Remove pattern"/>
<input type="button" value="Add pattern"/>	

Exclude filters are configured during task **creation**


Filtering to include files and folders

- + Include specific files/folders each time you execute the task
e.g. only transfer changed files under /new-files folder

▼ **Filtering configuration - optional**

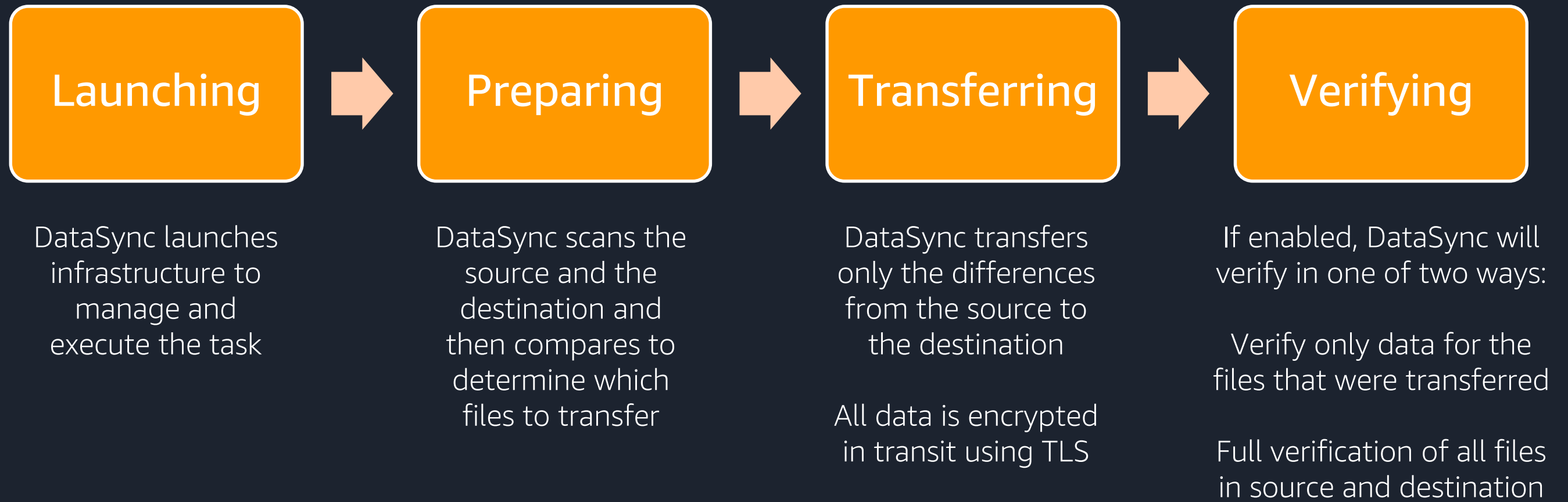
When no filters are specified, the entire contents of the source location are transferred.

Include patterns
Include patterns are applied for a specific task execution, in addition to the exclude patterns specified for a task. Only files, folders, and objects with the specified patterns are transferred. The pattern path is relative to the source location path. For example, /my-folder is a folder directly under the task's source location. Specifying /my-folder includes the folder itself and all of its contents).
Note: A wildcard (*) is only supported as the rightmost character of an include pattern.

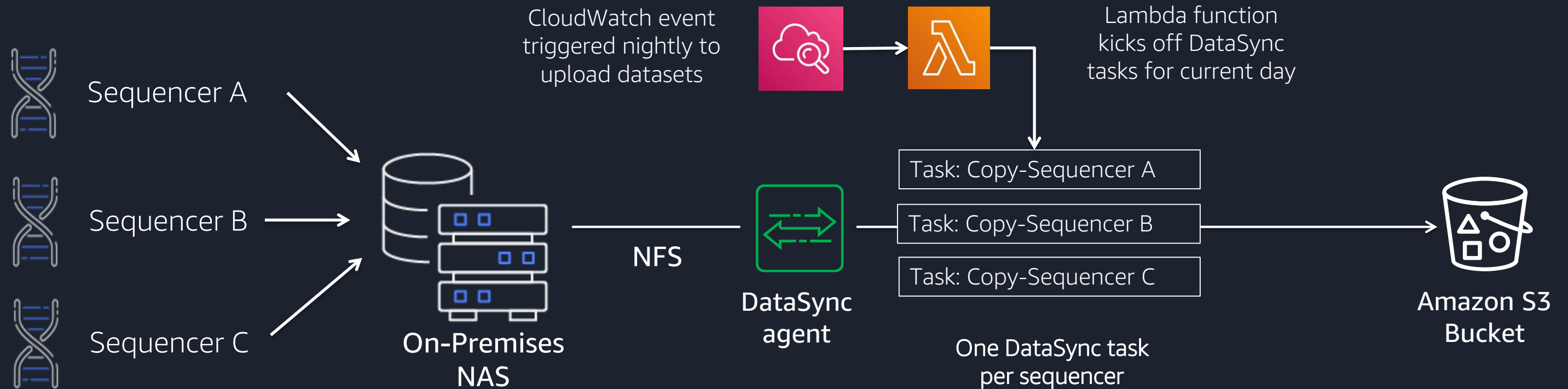
[Learn more and see syntax examples](#) 

Include filters are configured during task **execution**

Task stages



Using AWS DataSync to scale on-premises workloads



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

AWS DataSync best practices

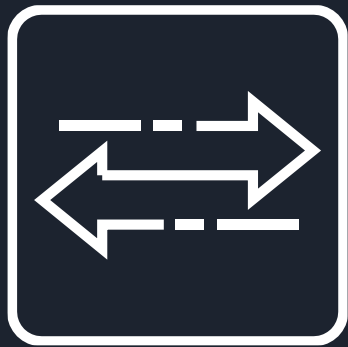
- ➔ Plan using available network bandwidth
- ➔ Always perform test runs before starting full transfer
- ➔ File size affects throughput
- ➔ Scale out performance with multiple agents
- ➔ When copying to EFS or FSx for Windows File Server, provision throughput appropriately

Get started

AWS DataSync - as simple as 1-2-3...

1

Download and deploy an agent and activate in your AWS Region



VMware, Hyper-V, KVM, EC2

2

Select source and destination locations and create a task



3

Run the task and monitor in CloudWatch or the DataSync console



Hands-on Workshops

github.com

[aws-samples/aws-datasync-migration-workshop](https://github.com/aws-samples/aws-datasync-migration-workshop)

[aws-samples/aws-datasync-fsx-migration](https://github.com/aws-samples/aws-datasync-fsx-migration)



aws.amazon.com/datasync/resources

Blogs, videos, demos, and more

Thank you