



# [AWS Black Belt Online Seminar]

## AWS DataSync

Senior Storage Specialist SA 川端 真

2021/3/16

<https://amzn.to/JPWebinar>



<https://amzn.to/JPArchive>



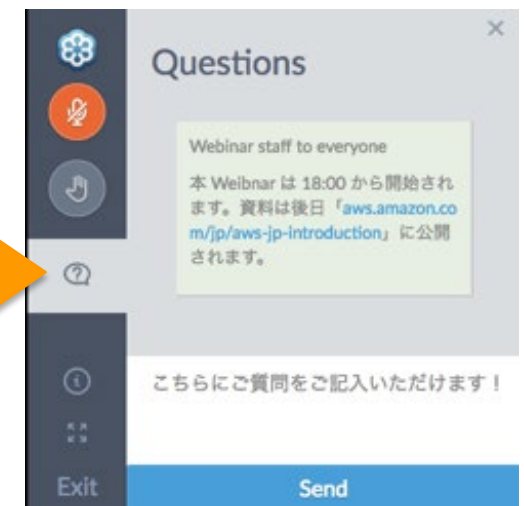
# AWS Black Belt Online Seminar とは

「サービス別」「ソリューション別」「業種別」のそれぞれのテーマに分かれて、アマゾンウェブサービス ジャパン株式会社が主催するオンラインセミナーシリーズです。

## 質問を投げることができます！

- 書き込んだ質問は、主催者にしか見えません
- 今後のロードマップに関するご質問はお答えできませんのでご了承下さい

- ① 吹き出しをクリック
- ② 質問を入力
- ③ Sendをクリック



 Twitter ハッシュタグは以下をご利用ください  
#awsblackbelt

# 内容についての注意点

- 本資料では2021年3月16日現在のサービス内容および価格についてご説明しています。最新の情報はAWS公式ウェブサイト(<http://aws.amazon.com>)にてご確認ください。
- 資料作成には十分注意しておりますが、資料内の価格とAWS公式ウェブサイト記載の価格に相違があった場合、AWS公式ウェブサイトの価格を優先とさせていただきます。
- 価格は税抜表記となっております。日本居住者のお客様には別途消費税をご請求させていただきます。
- AWS does not offer binding price quotes. AWS pricing is publicly available and is subject to change in accordance with the AWS Customer Agreement available at <http://aws.amazon.com/agreement/>. Any pricing information included in this document is provided only as an estimate of usage charges for AWS services based on certain information that you have provided. Monthly charges will be based on your actual use of AWS services, and may vary from the estimates provided.

# 本セミナーの概要

## □ 本セミナーで学習できること

- ❖ データ転送の課題と解決策
- ❖ AWS DataSyncの概要と詳細
- ❖ AWS DataSyncのユースケースとTIPS

## □ 対象者

- ❖ オンプレミスデータのAWS移行を検討されている方
- ❖ AWSサービス間のデータ移動、または複製を検討されている方
- ❖ リージョン間レプリケーションを検討されている方
- ❖ 次の AWS のサービスの概要レベルの知識が前提になります

Amazon s3 / Amazon EFS / Amazon FSx for Windows などの  
AWSストレージ関連サービス

# 自己紹介

川端 真 (Mac Kawabata)

Senior Storage Specialist Solution Architect

- ストレージサービス全般  
AWS DataSync  
Snow Family  
Transfer Family  
Storage Gatewayを担当
- 好きなAWSのサービス  
Amazon FSx for Windows  
AWS DataSync

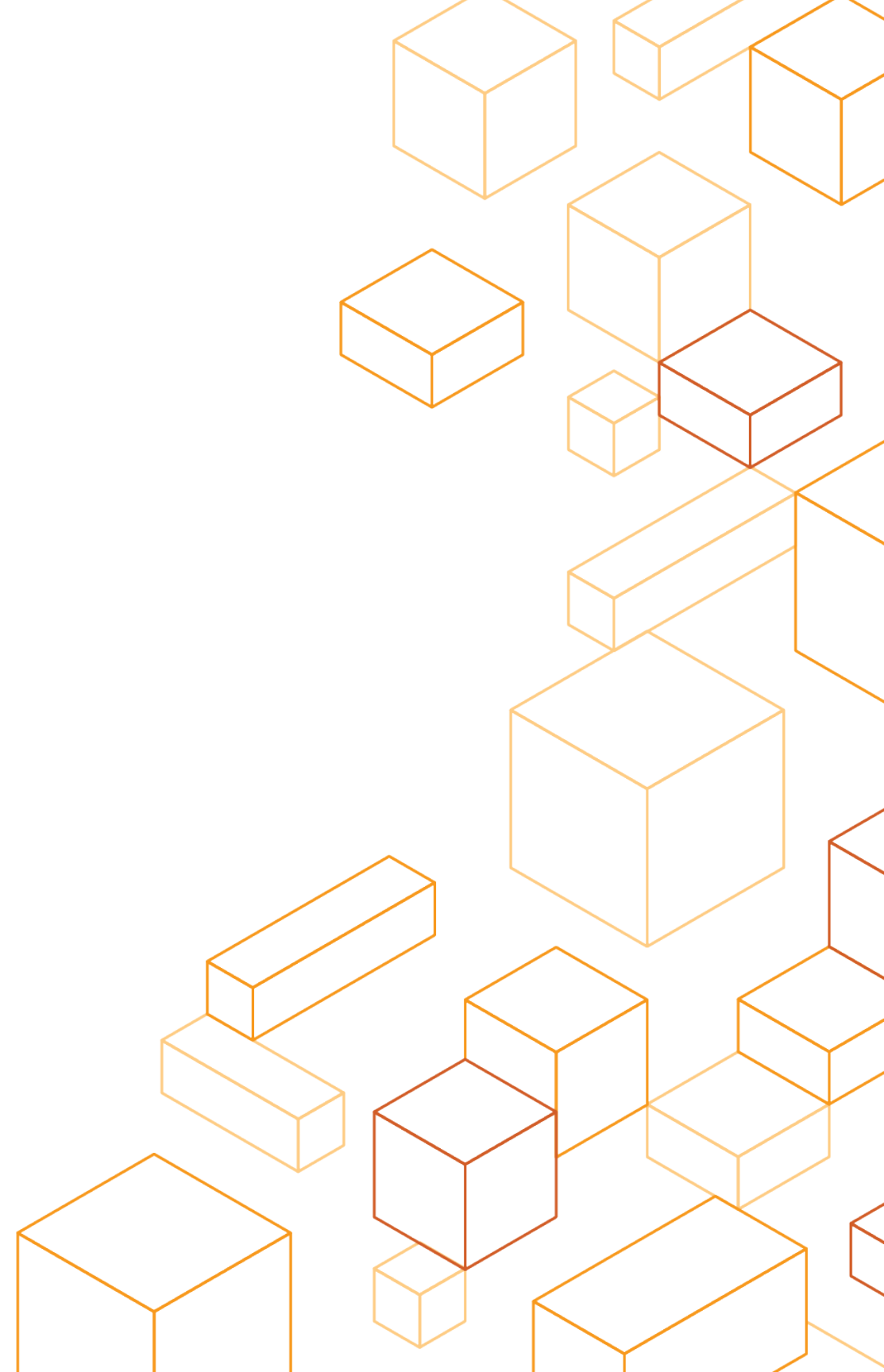


# 本日のアジェンダ

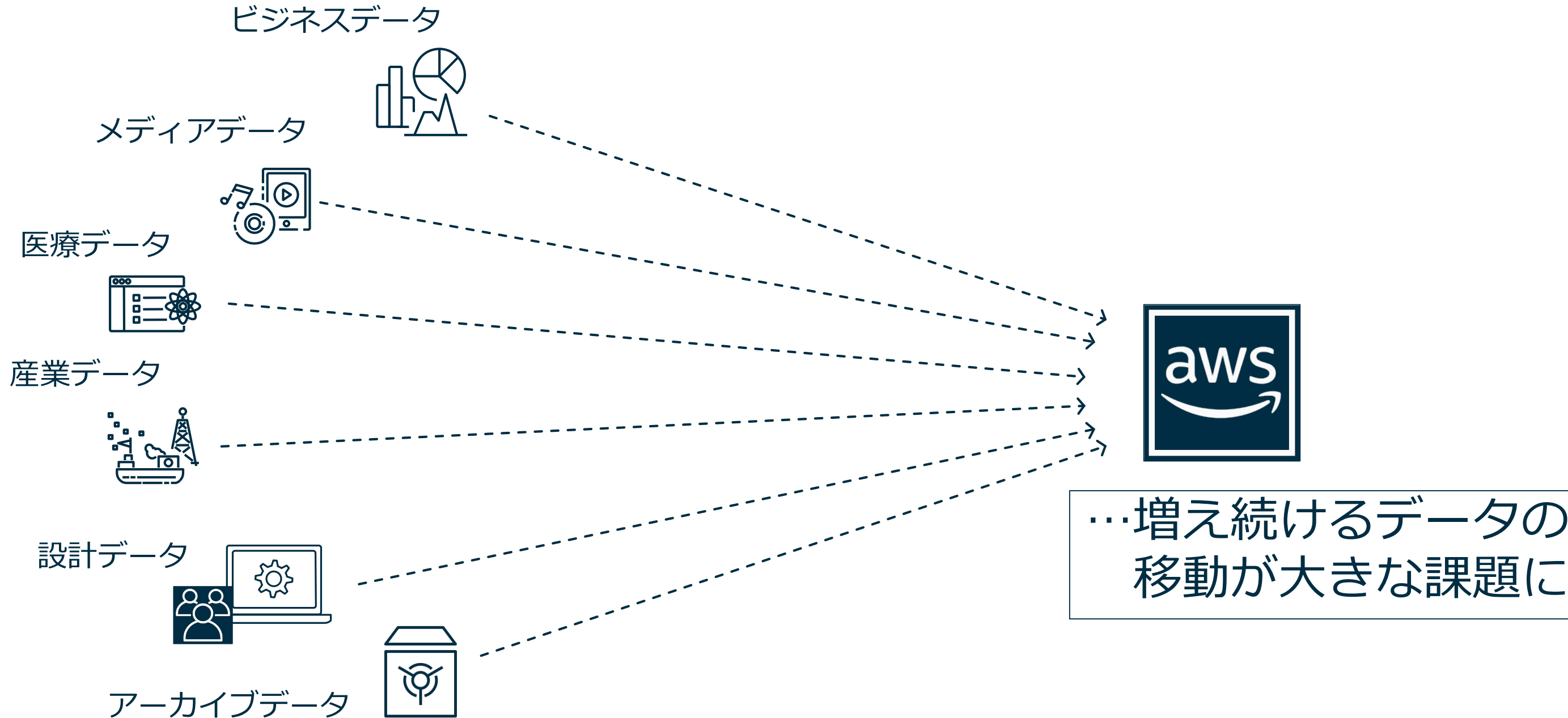
- データ転送の課題
- AWS DataSyncの概要
- AWS DataSyncのユースケース
- AWS DataSyncの詳細
- ベストプラクティス & TIPS
- 顧客事例



# データ転送の課題

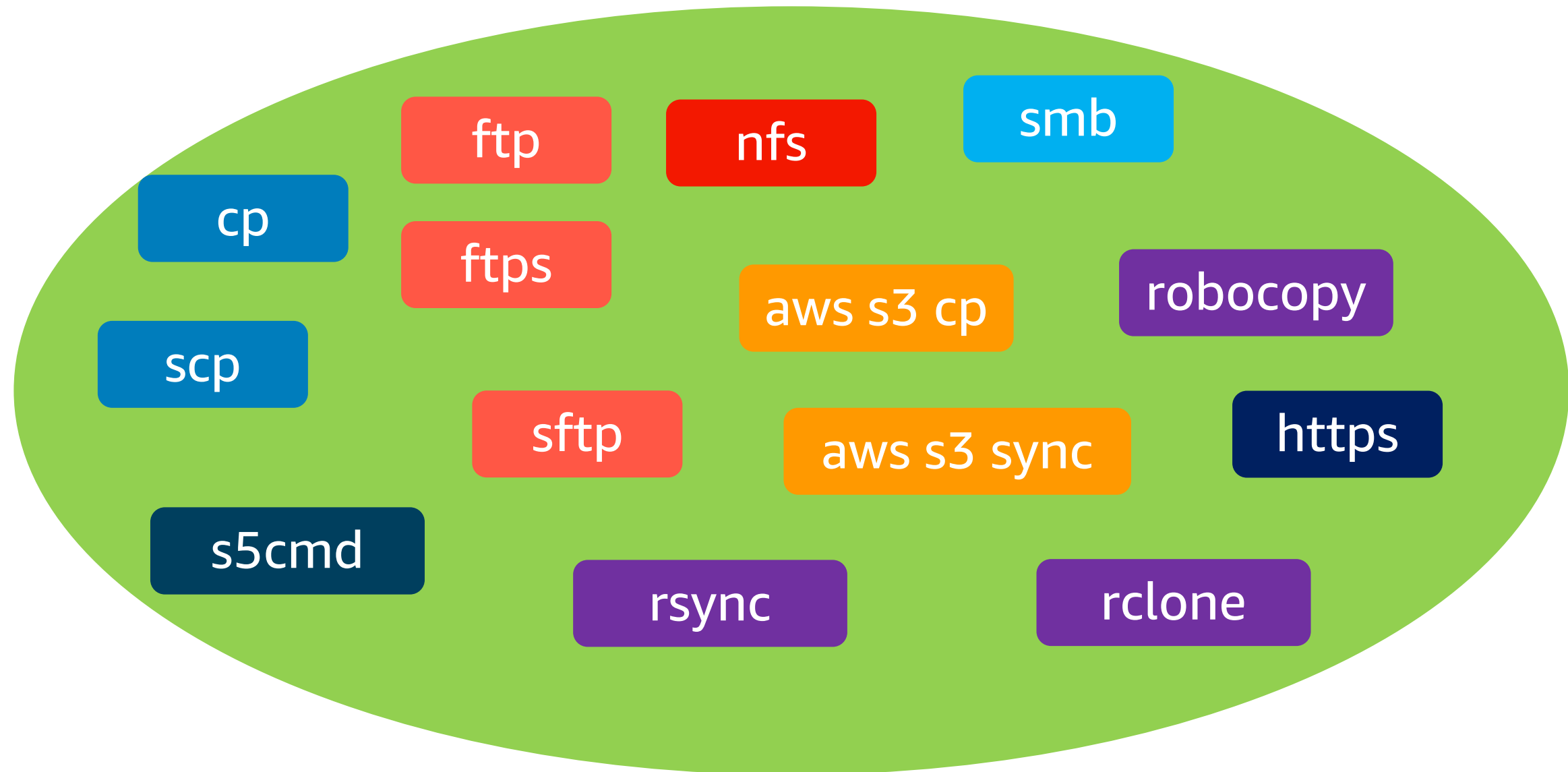


# 重要なワークロードがクラウドに移行するにつれて…



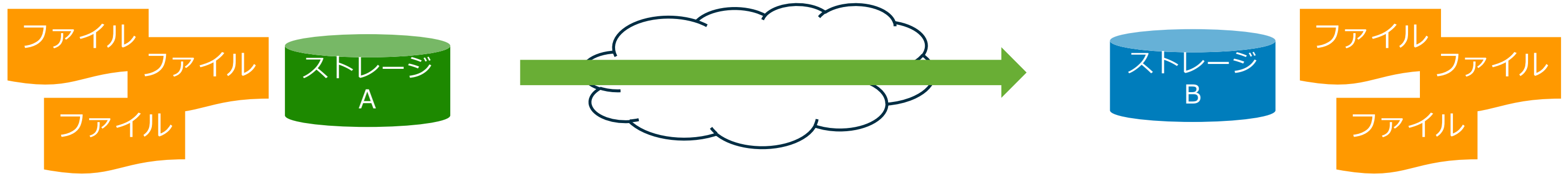


# 世の中には様々なデータ転送手法が存在



# データ転送の課題

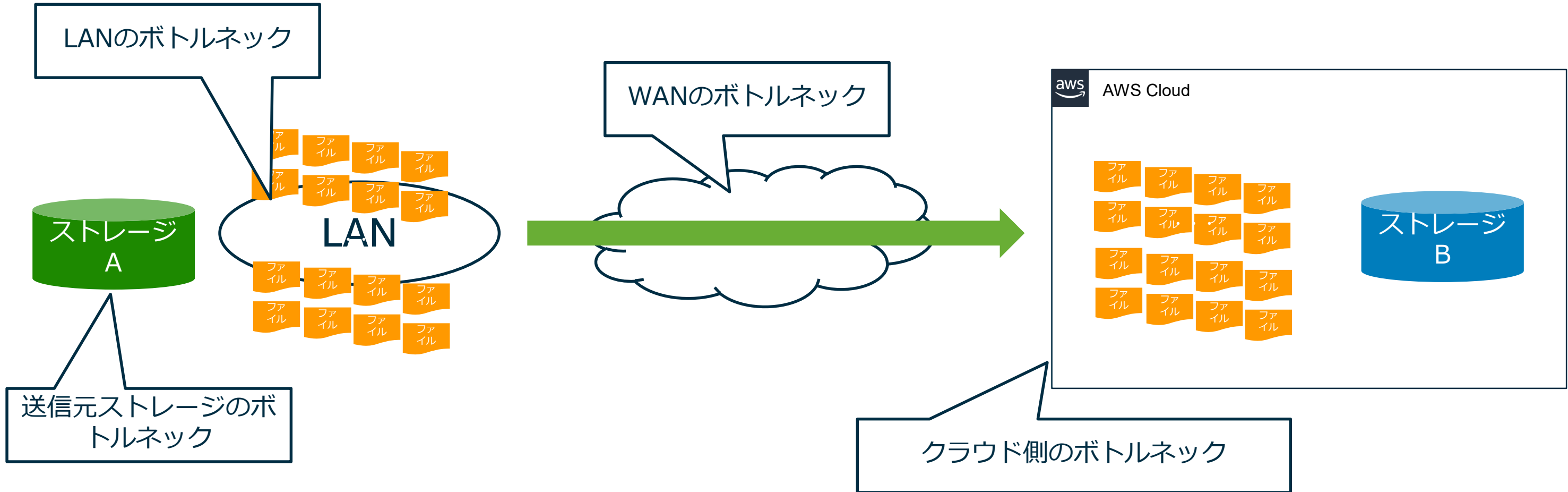
小規模のデータ転送なら気にする事はほとんどありませんが？



目視やツールでなんとかなるレベル

# データ転送の課題

数千万、数百万ファイルになると？



エラーハンドリングの考慮

差分比較時間の考慮

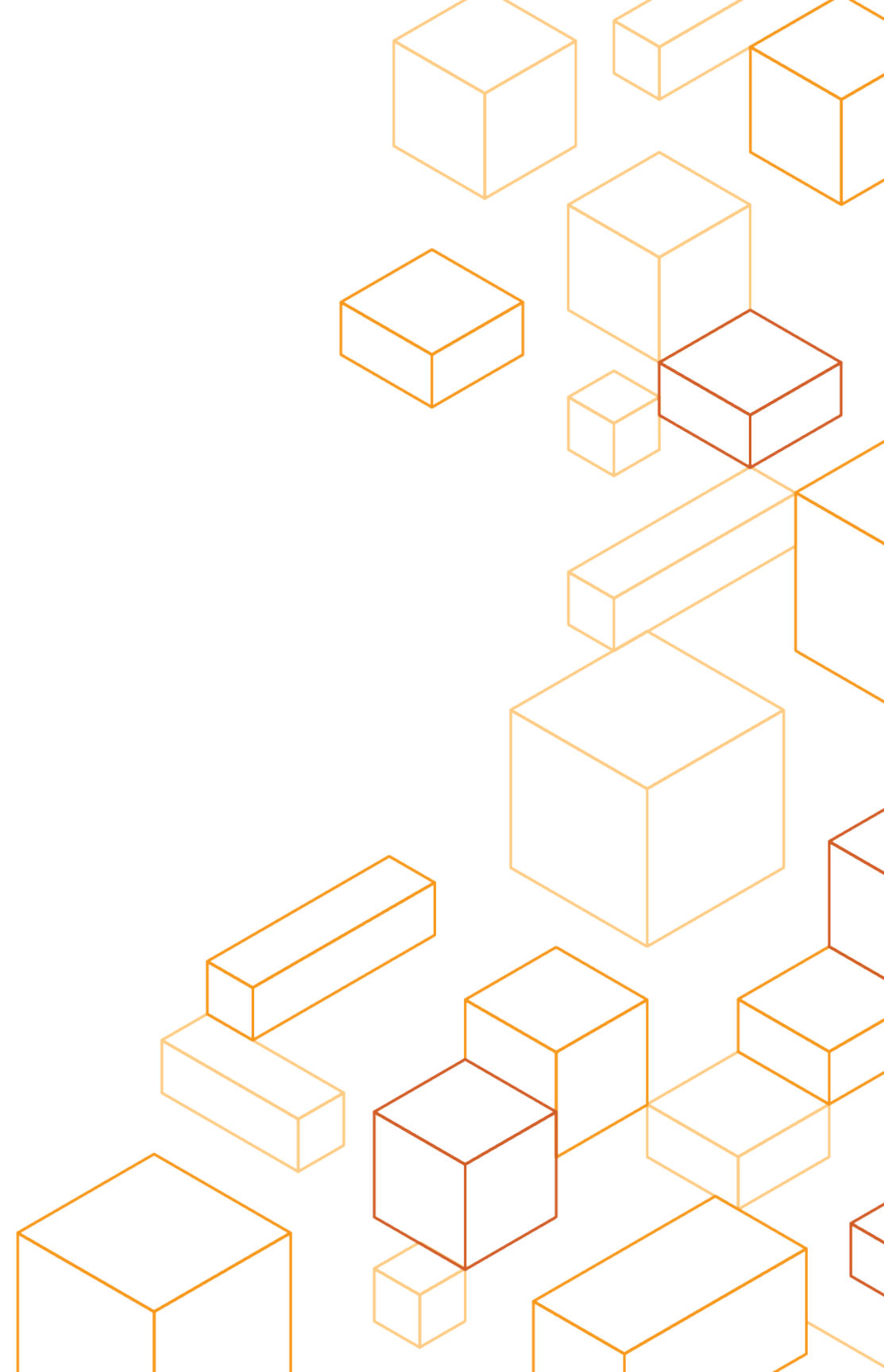
データ整合性の考慮

セキュリティの考慮

メタデータの考慮



# AWS DataSyncの概要



# AWS DataSyncとは

オンプレミスとAWSストレージサービス、又はAWSストレージサービス間のデータの移動を、高速・簡単・セキュア・低コストに行えるフルマネージドサービスです



高速データ転送



簡易な操作性



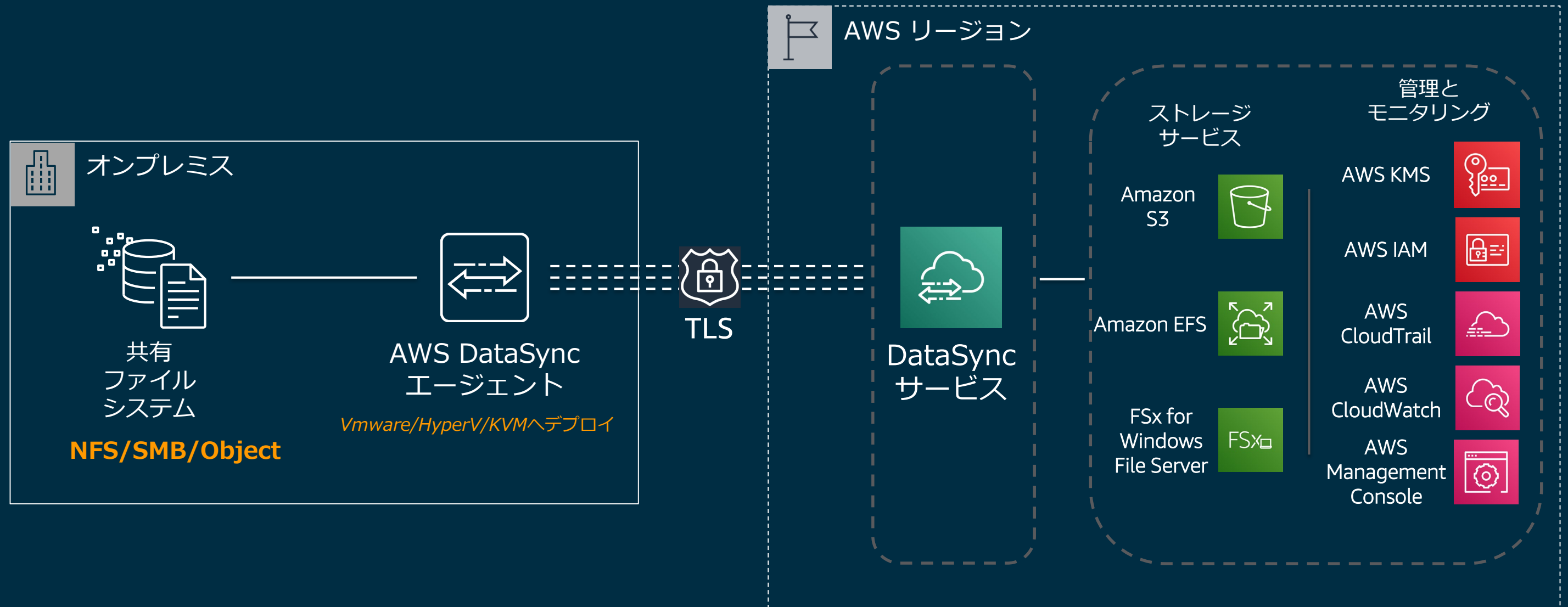
セキュアで高信頼



低コスト

# AWS DataSync動作概要

オンプレミスからAWS、又はAWSストレージ間のデータを簡単に同期



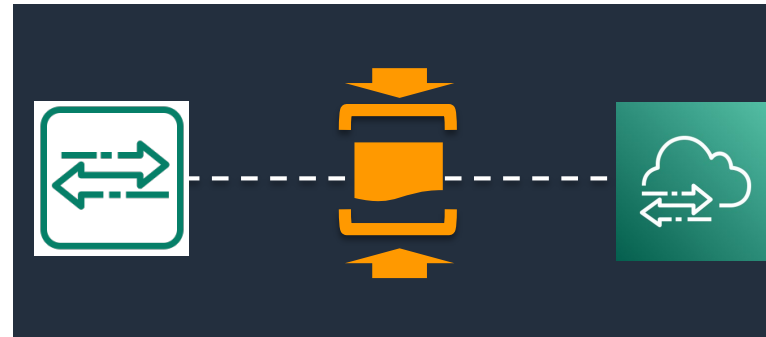
# 高速データ転送

## マルチセッション



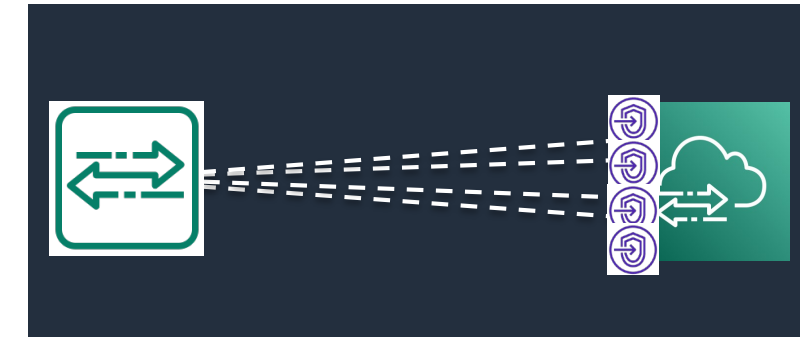
エージェントとサービス間で、同時に複数のTCPセッションを生成し並列転送

## 圧縮



エージェントとサービス間はデータを圧縮して効率化

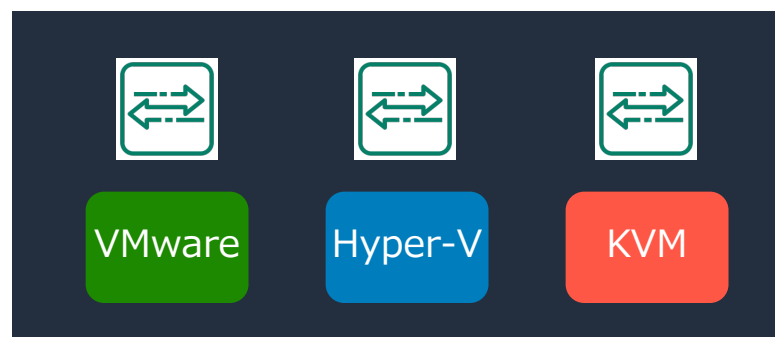
## ロードバランス



サービス側に複数のエンドポイントを自動生成しロードバランス

# 簡易な操作性

## エージェント



エージェントは容易に  
セットアップ可能な  
仮想アプライアンス  
として提供

## 3ステップの設定



簡単な3つのステップで  
設定出来、面倒な手順を  
憶えなくても使い始める  
事が出来ます

## 自動化



組み込みのスケジュール  
機能でデータの同期を  
柔軟に自動化



# セキュアで高信頼

## 伝送路暗号化



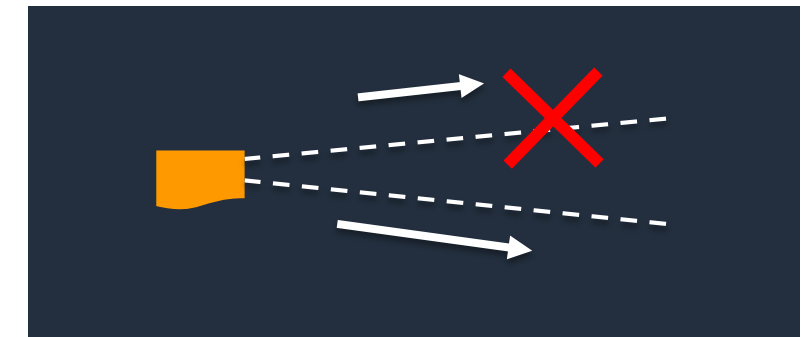
伝送路上のデータはTLSで暗号化されます

## 保存データ暗号化



S3、EFSへ暗号化してデータを保存する事が出来ます (FSx for Windowsはサービスのバックエンドで暗号化出来ます)

## 自動再送



ネットワークの切断等で転送に失敗した場合でも自動的に再送を試みます

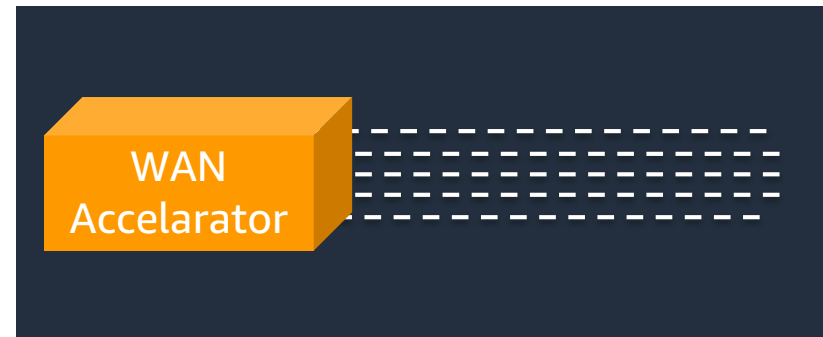
# 低コスト

## シンプルで安価な料金



DataSyncの料金は転送したデータ量のみで課金されます  
東京リージョンではGiBあたり0.0125USDです  
\*2021/3/16時点

## アクセラレータ不要



ネットワーク転送を高速化する専用機器は必要ありません

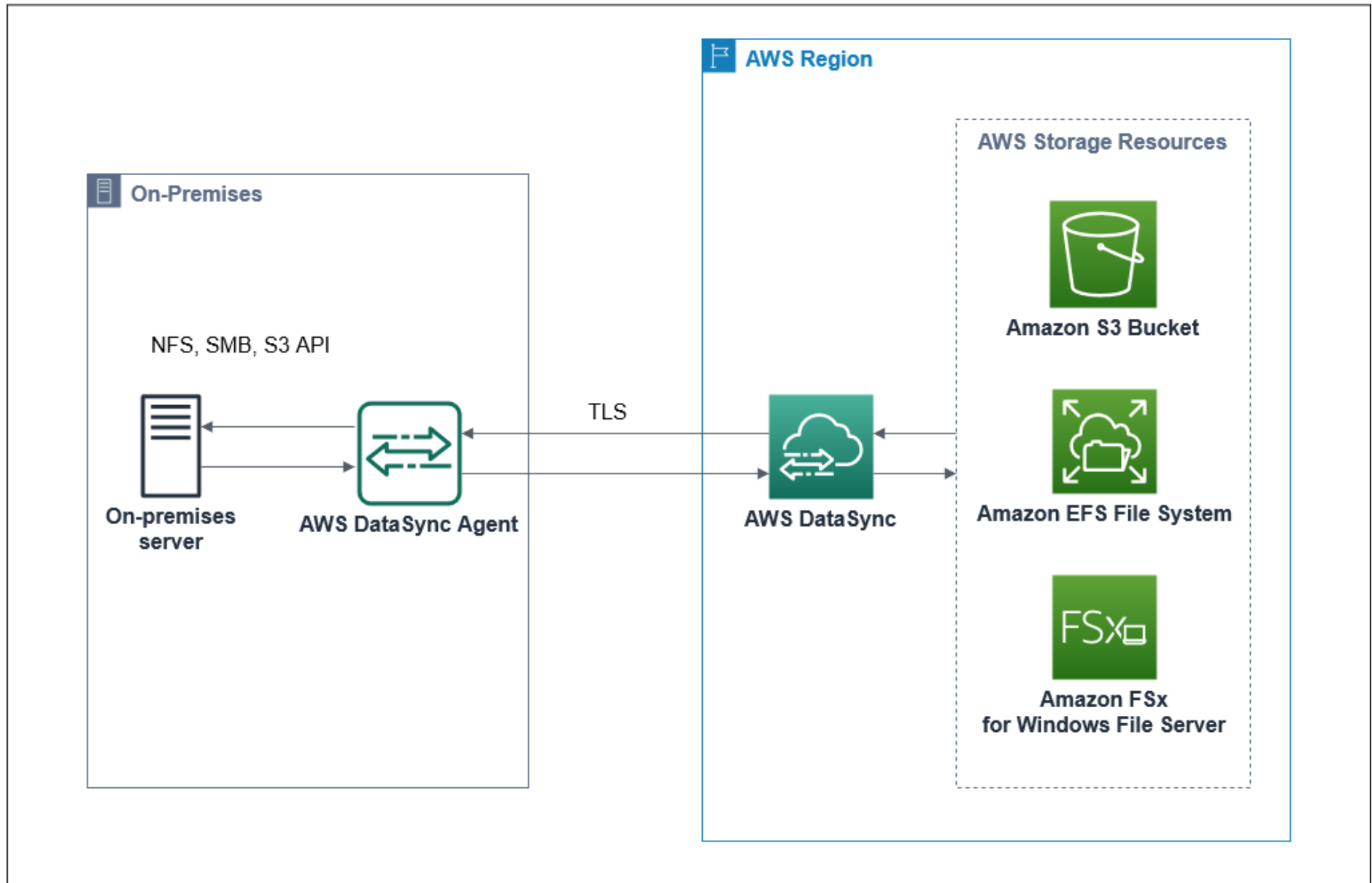
## 運用コスト削減



転送状態の管理や、転送後のデータ整合性検証等の面倒なオペレーション削減により、運用コストの削減が期待出来ます

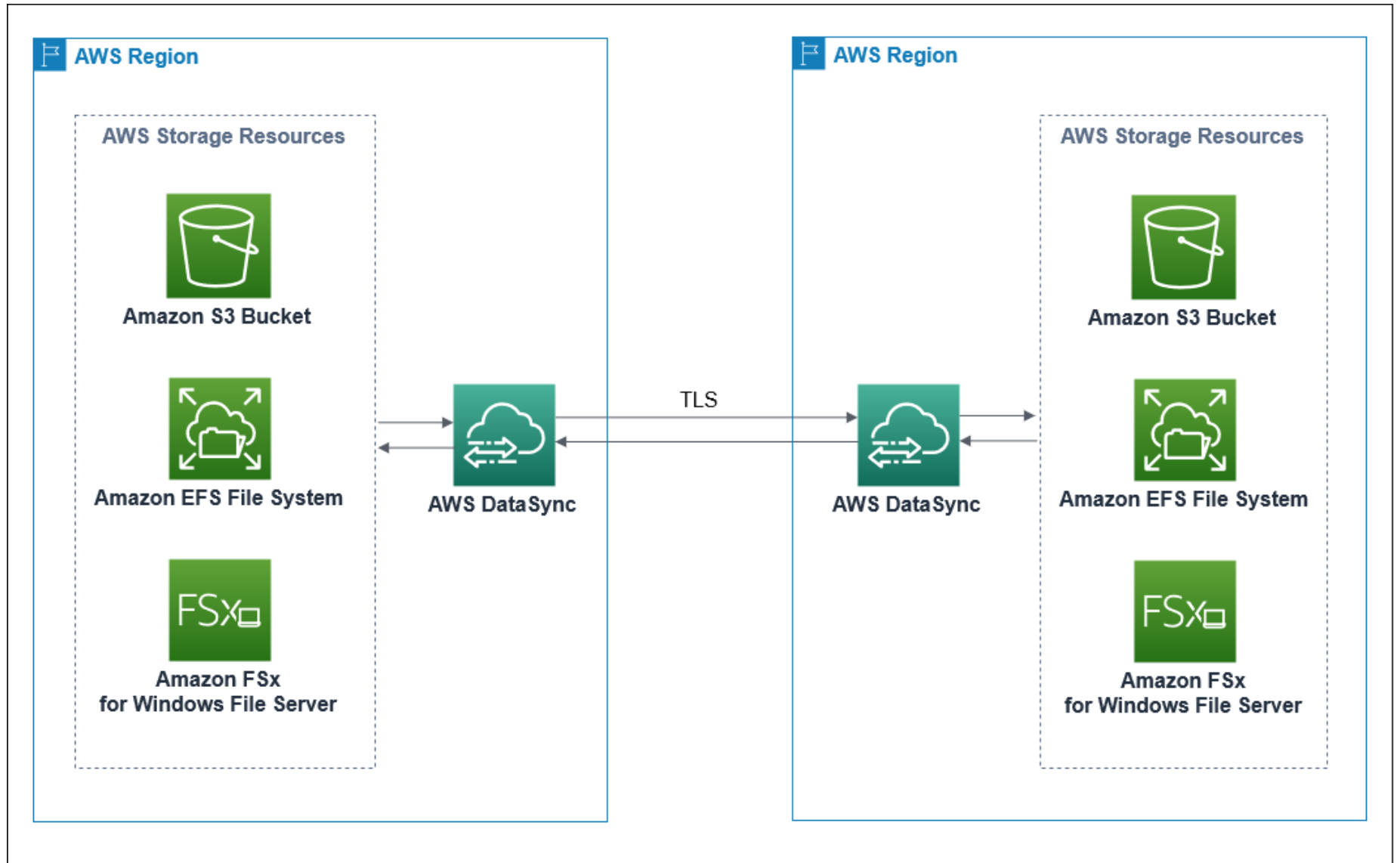
# AWS DataSyncの使用形態①： オンプレミスとAWSストレージサービス間の転送

オンプレミスからAWS  
ストレージサービスへ転送  
する場合、オンプレミス側  
にAWS DataSync Agent  
という仮想アプリケーション  
が必要です  
(vSphere, Hyper-V,  
KVMに対応しています)



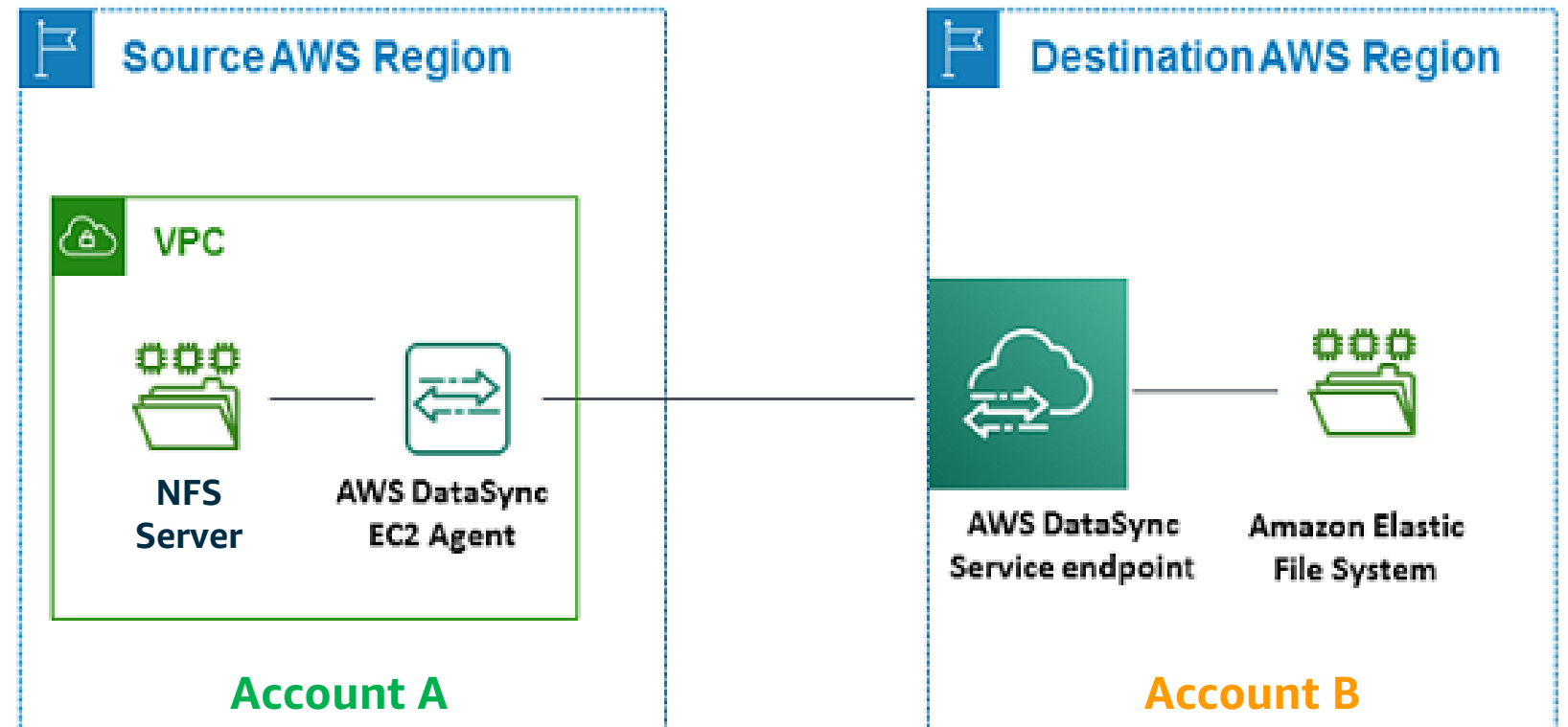
# AWS DataSyncの使用形態②： 自己所有AWSストレージサービス間の転送

同一アカウントが所有するAWSストレージサービス間で転送を行う場合、Agentの設置は不要です。リージョンと対象サービス、対象ファイルシステム又はバケットを指定する事で容易にタスクを構成出来ます。



# AWS DataSyncの使用形態③： クロスアカウント転送

EC2 Agentを異なるアカウントのVPCに配置する事で、クロスアカウント転送する事も出来ます



# サポートしている転送の組み合わせ

送信元	送信先
自己管理ストレージ (NFS、SMB、オブジェクト、Snowcone上のNFS)	Amazon S3 (AWSリージョン内)、Amazon EFS、Amazon FSx for Windows File Server
Amazon S3 (AWSリージョン内)、Amazon EFS、Amazon FSx for Windows File Server	自己管理ストレージ (NFS、SMB、オブジェクト、Snowcone上のNFS)
Amazon S3 (AWSリージョン内)、Amazon EFS、Amazon FSx for Windows File Server	Amazon S3 (AWSリージョン内)、Amazon EFS、Amazon FSx for Windows File Server
Amazon S3 (AWSリージョン内)	Amazon S3 on AWS Outposts
Amazon S3 on AWS Outposts	Amazon S3 (AWSリージョン内)

# サポートしている転送の組み合わせ（クロスアカウント転送）

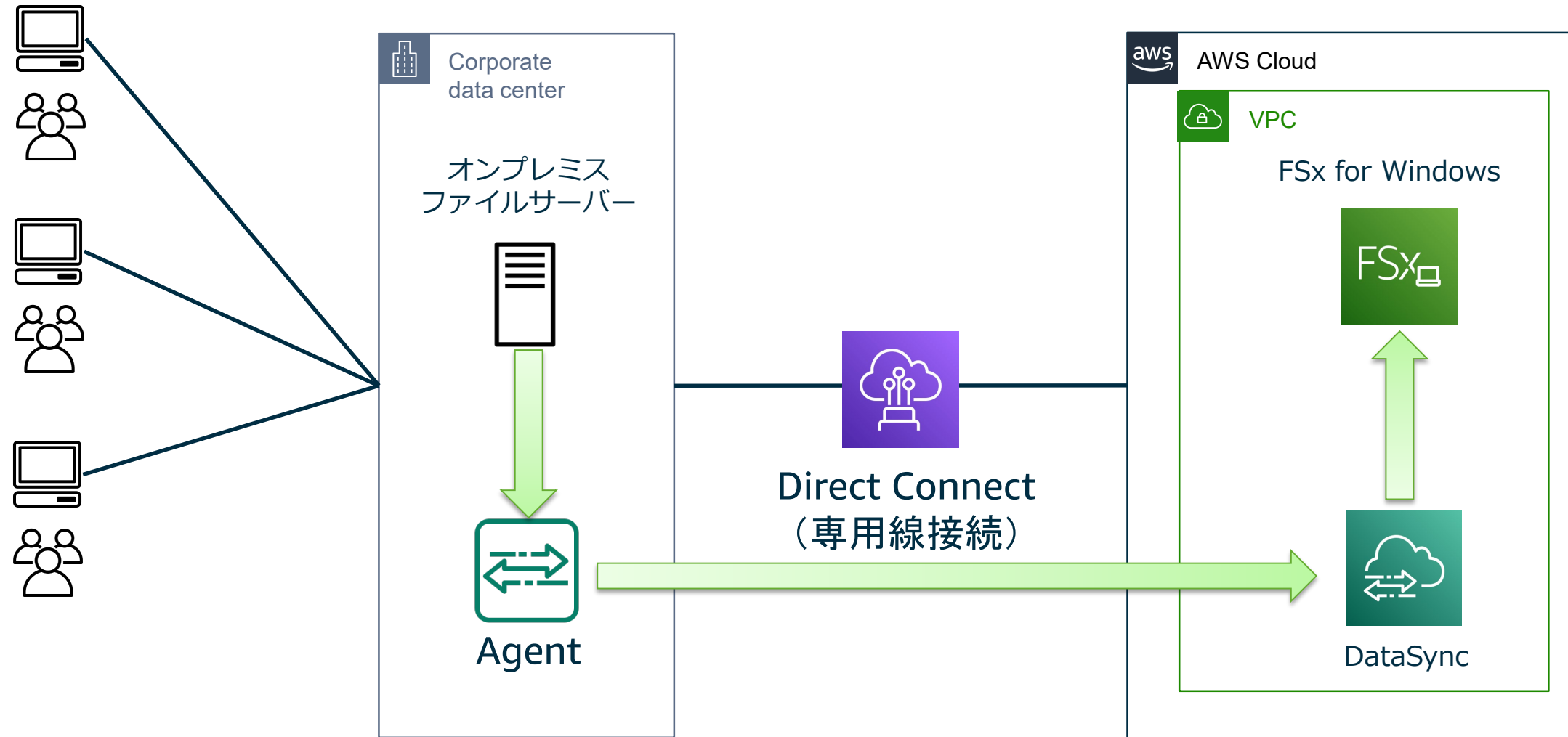
送信元	送信先
Amazon EFS（NFSマウントとして設定）、 Amazon FSx for Windows File Server（SMBマウントとして設定）	Amazon S3（AWSリージョン内）、Amazon EFS、 Amazon FSx for Windows File Server
Amazon S3（AWSリージョン内）	Amazon EFS（NFSマウントとして設定）、 Amazon FSx for Windows File Server（SMBマウントとして設定）



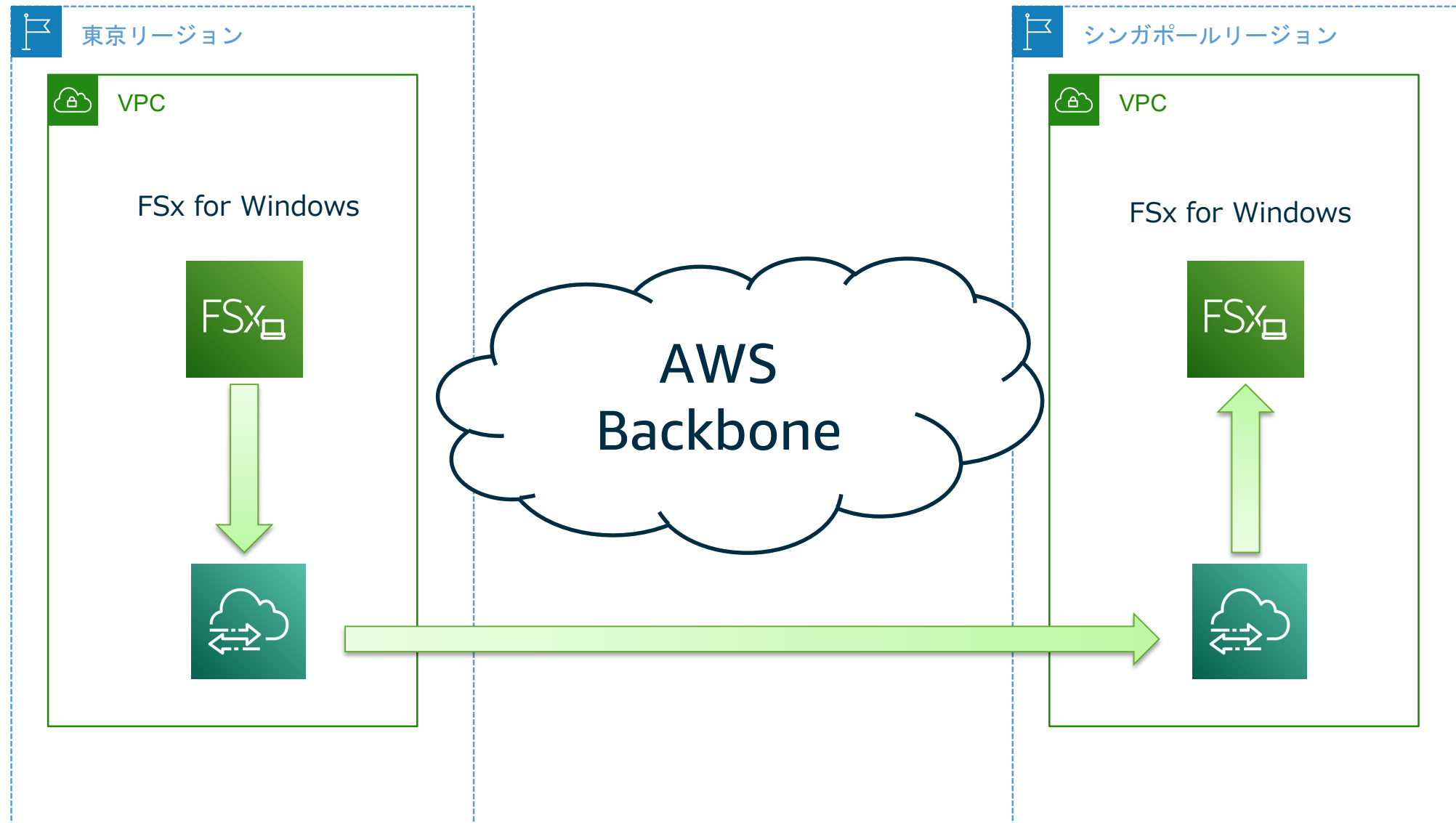
# AWS DataSyncのユースケース



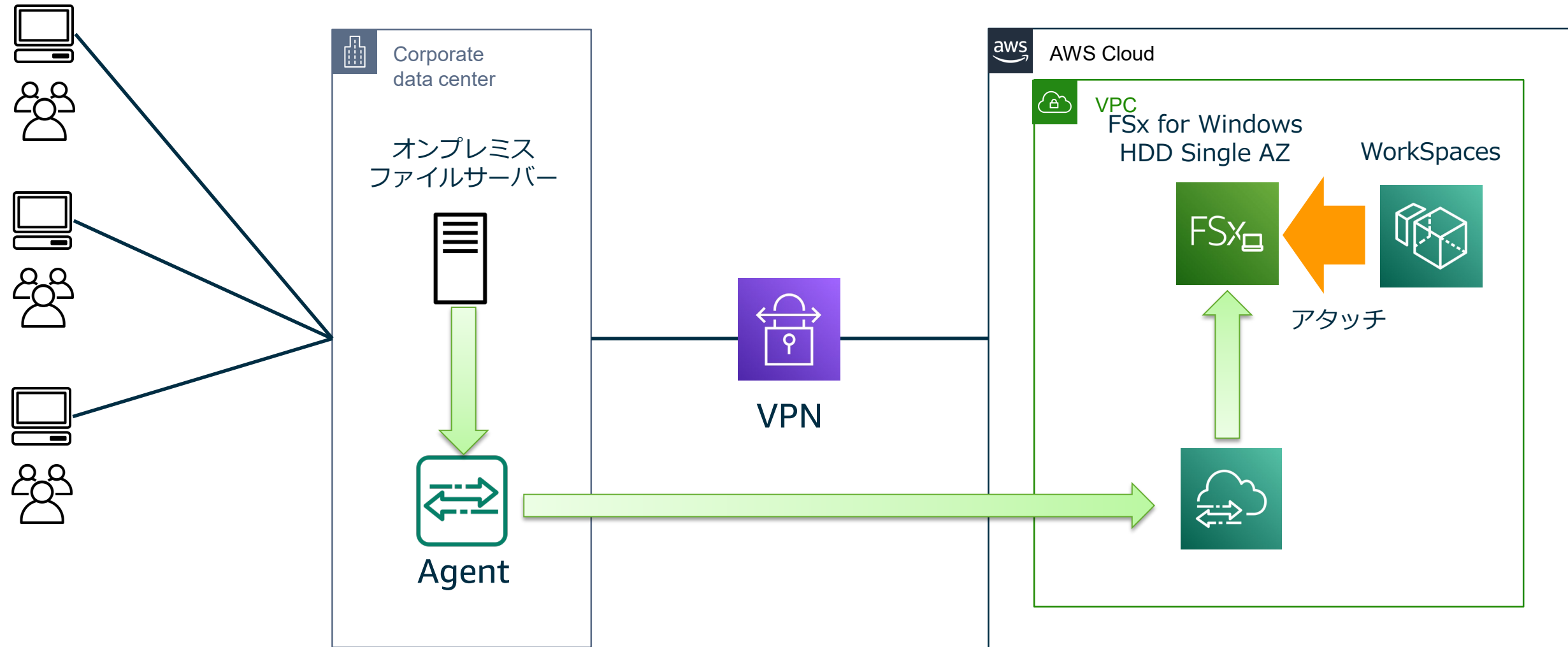
# オンプレミスからのデータ移行



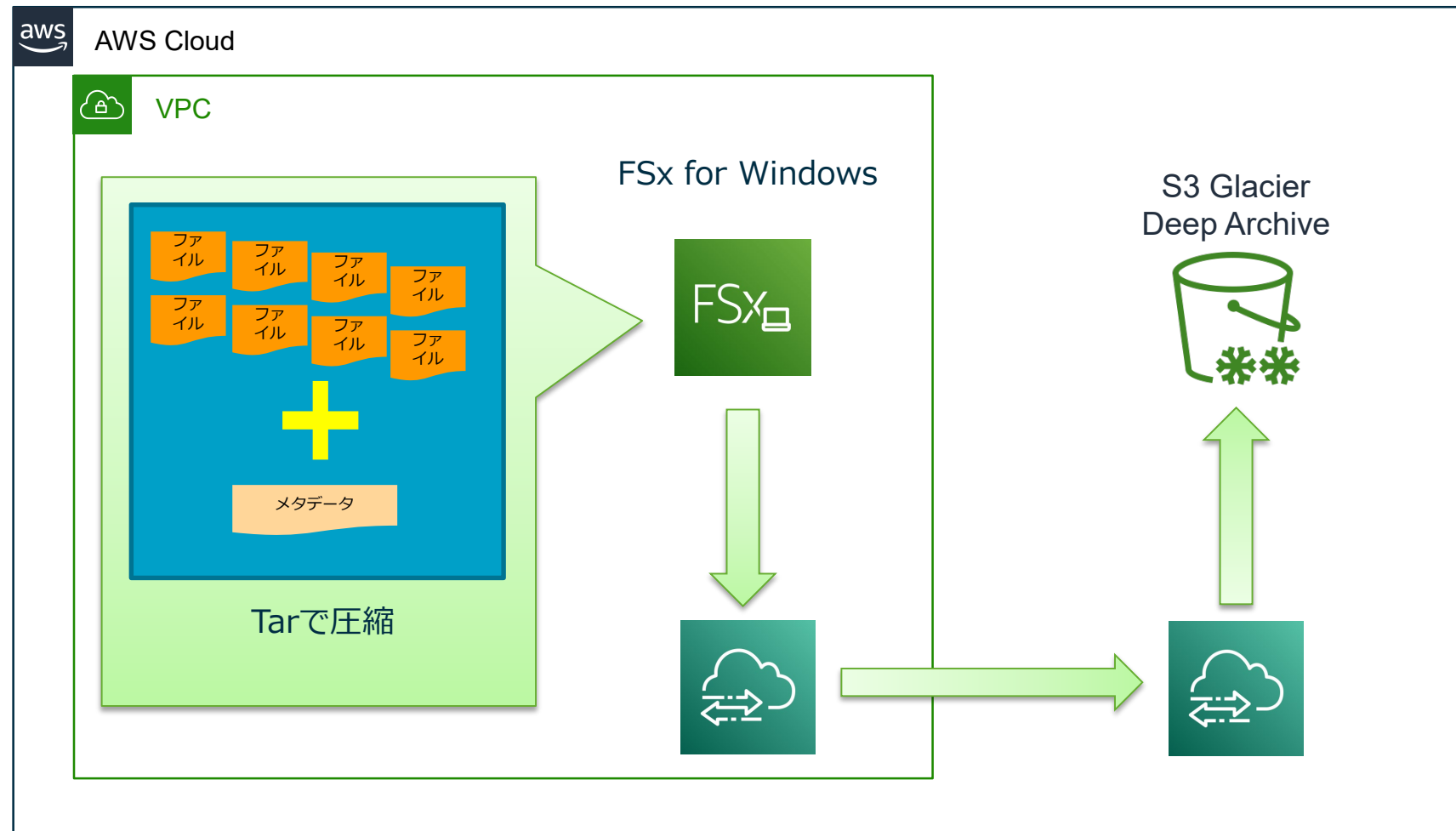
# ストレージレプリケーションによるDR



# SMBデータのHotバックアップ

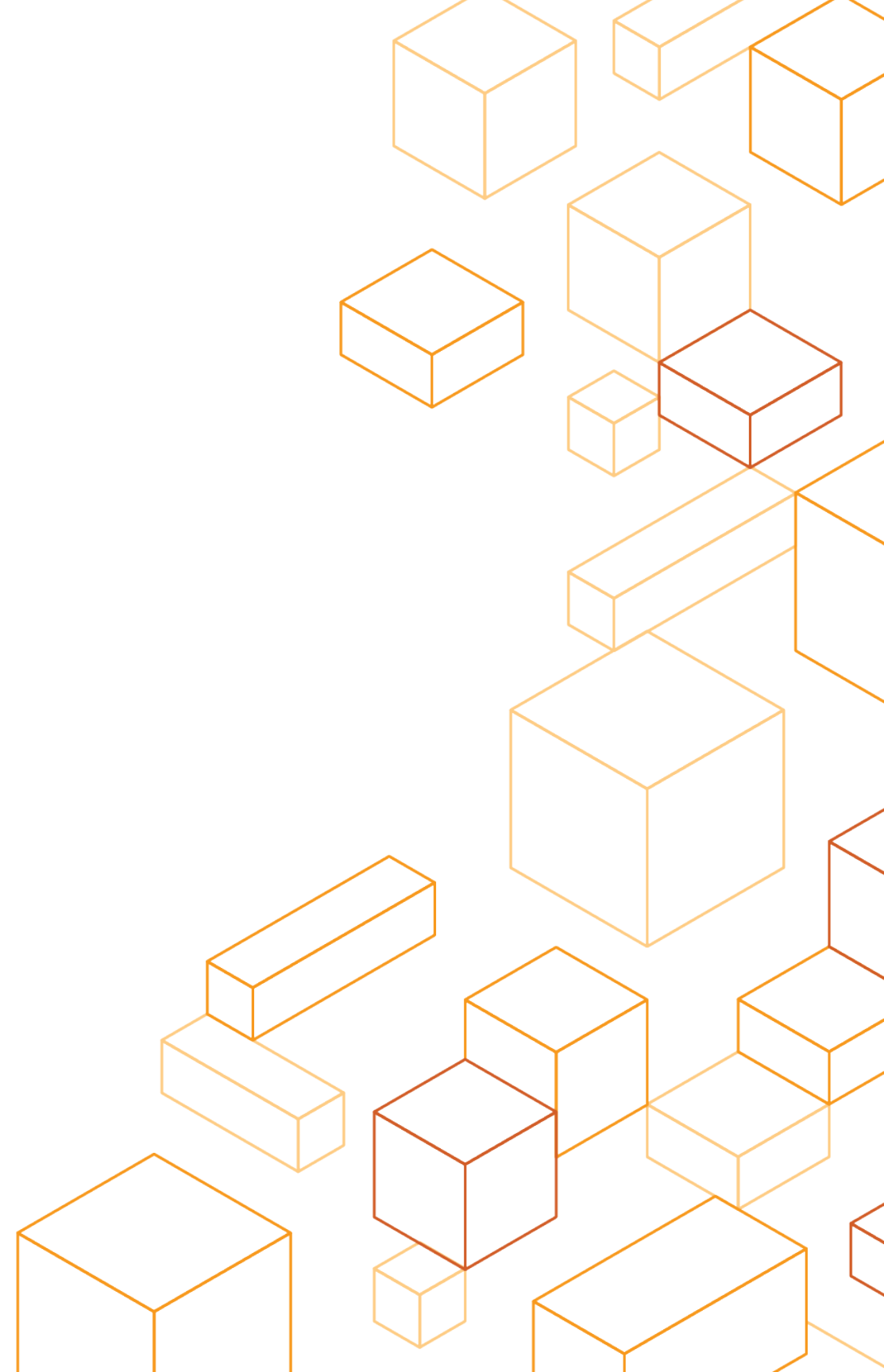


# データのアーカイブ





# AWS DataSyncの詳細

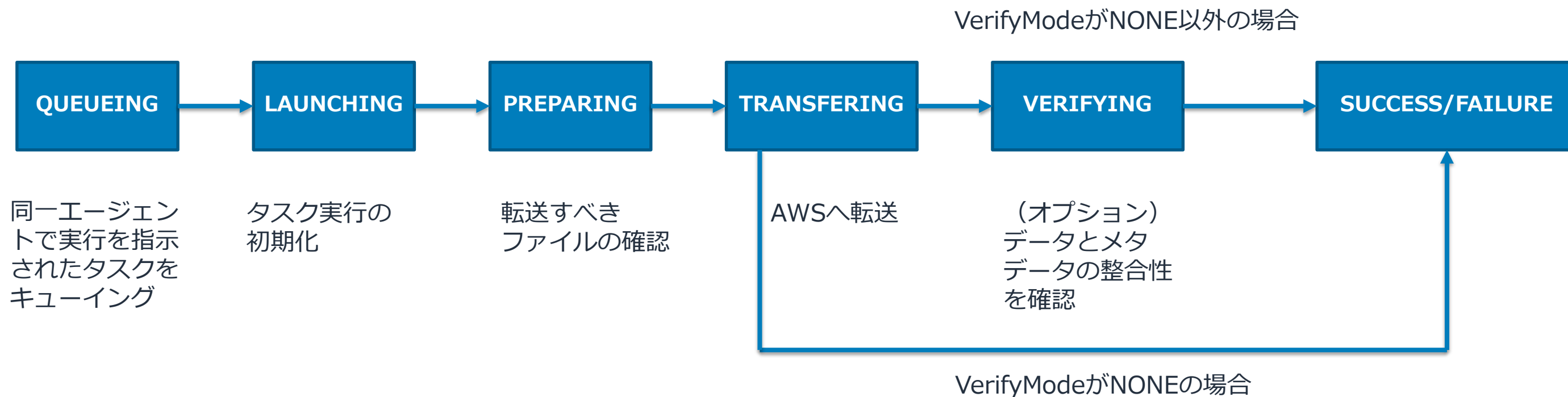


# AWS DataSyncのコンポーネントと用語

名前	説明
エージェント	自己管理ストレージとのデータの読み書きに使用する仮想マシンです。同一アカウントのAWSストレージサービス間でデータを転送する場合は、エージェントは必要ありません。
ロケーション	データ転送に使用する送信元または送信先のストレージです。(例: Amazon S3, Amazon EFS, Amazon FSx for Windows File Server, NFS, SMB, 又は自己管理オブジェクトストレージ)。
タスク	送信元ロケーションと送信先ロケーション、データ転送に関する各種設定から構成されます。常に送信元から送信先の方向で転送が行われます。各種設定にはスケジュールやネットワーク帯域制御等が含まれます。AWS DataSyncではタスクによりどのようにデータ転送するかが決定されます。
タスク実行	実際に実行されたタスクです。タスク実行には開始時間、終了時間、書き込まれたバイト数やタスクの状態等が含まれます。

# タスクの状態遷移

タスクは以下のような状態遷移を経ながら実行されます



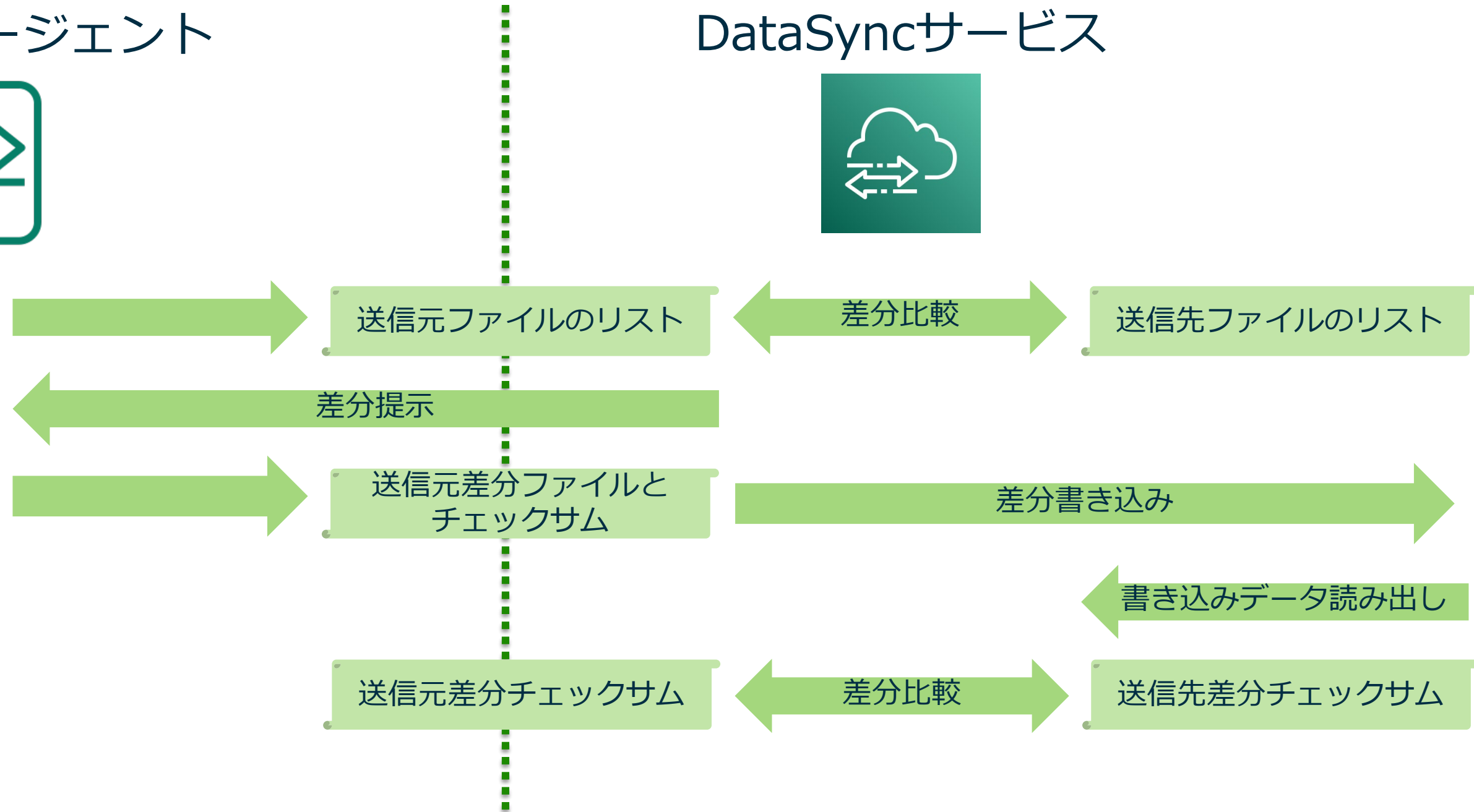
参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/how-datasync-works.html>

# データの整合性検証

DataSyncエージェント



DataSyncサービス



参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/how-datasync-works.html>



# 開いているファイル、ロックされているファイルの扱い

通常、DataSyncは特に制限なく開いているファイルを転送できます。

但し開いているファイルに書き込みが行われた場合、DataSyncはデータ整合性検証の段階で整合性の不一致を検出します。このため、DataSyncでの転送中は送信元ファイルへの書き込みへの注意が必要です。

もしファイルがロックされており、DataSyncはファイルを開けない場合、DataSyncはそのファイルの転送をスキップし、TRANSFERRINGの段階でエラーを出力し、整合性検証のエラーも出力されます。

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/how-datasync-works.html>

# エージェントの要求事項

ハイパーバイザーの種類	要件
VMware ESXi	バージョン6.5又は6.7（フリーバージョンの利用も可能） * VMwareのEnd of Supportポリシーに従います
Microsoft Hyper-V	バージョン2012 R2又は2016（フリーバージョンの利用も可能） * DataSync AgentはGeneration 1仮想マシンを使用
Linux KVM	Linux Kernel 2.6.20以上 * EC2 Linux上のKVMでのDataSync Agentは未サポート

ハードウェア	要件
仮想プロセッサ	4
ストレージ	80GB
メモリ	32GB：2000万ファイル未満 64GB：2000万ファイル以上

EC2	要件
インスタンスタイプ	m5.2xlarge：2000万ファイル未満 m5.4xlarge：2000万ファイル以上

参考：<https://docs.aws.amazon.com/datasync/latest/userguide/agent-requirements.html>

# ネットワークの要求事項

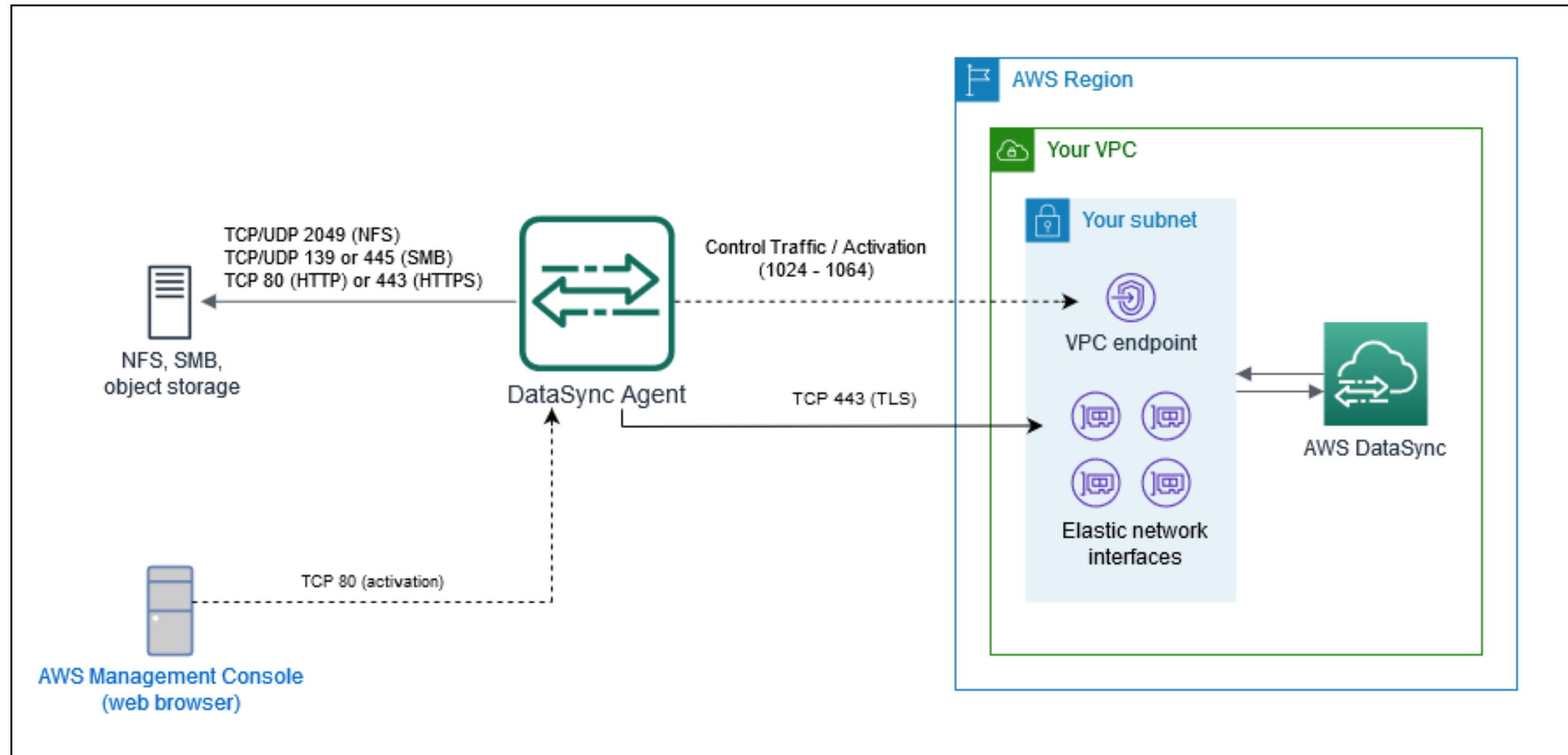
## 自己管理ストレージへのアクセス

送信元	送信先	プロトコル	ポート	用途
エージェント	NFSサーバー	TCP/UDP	2049 (NFS)	エージェントによるNFSマウントに使用 NFS3.x、NFS4.0、NFS4.1をサポート
エージェント	SMBサーバー	TCP/UDP	139 (SMB) or 445 (SMB)	エージェントによるSMBマウントに使用 SMB2.1、SMB3をサポート
エージェント	自己管理オブジェクト ストレージ	TCP	443 (HTTPS) or 80 (HTTP)	自己管理オブジェクトストレージへのアクセス に使用

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/datasync-network.html>

# ネットワークの要求事項

## プライベートエンドポイント使用时



参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/datasync-network.html>

# ネットワークの要求事項

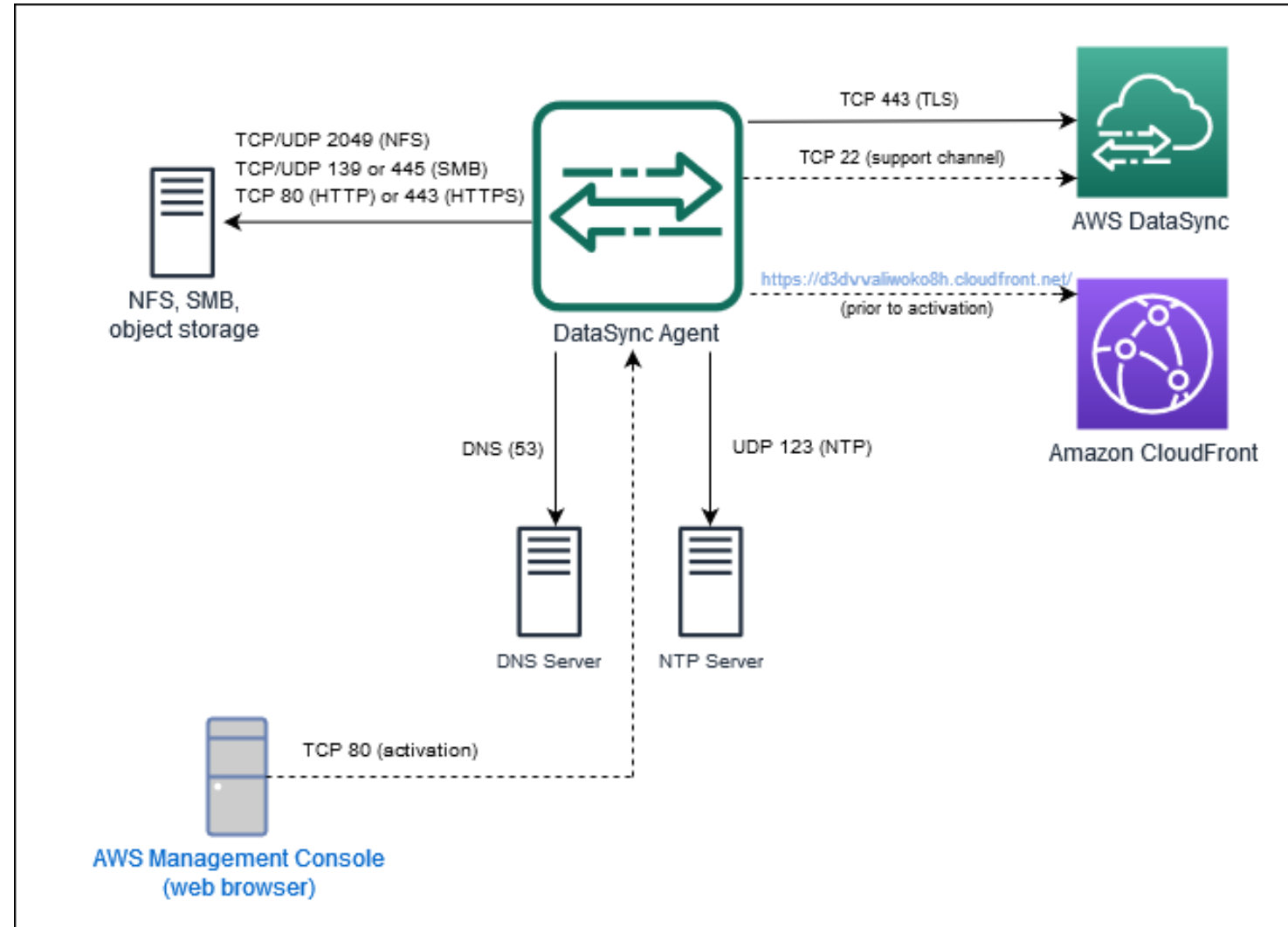
## 通信の詳細（プライベートエンドポイント使用時）

送信元	送信先	プロトコル	ポート	用途
コンソール端末のブラウザ	DataSync agent	TCP	80 (HTTP)	アクティベーションKey取り出しに使用 アクティベーション完了後はポートが閉じられる
エージェント	DataSync VPC endpoint	TCP	1024-1064	エージェントとサービスの間の制御通信用
エージェント	タスクが使用するENI	TCP	443 (HTTPS)	エージェントとサービスの間のデータ通信用
エージェント	DataSync VPC endpoint	TCP	22 (Support channel)	サポートによるリモート解析用

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/datasync-network.html>

# ネットワークの要求事項

パブリック又は\*FIPSエンドポイント使用  
時



\* FIPS (Federal Information Processing Standard) : 連邦情報処理規格140-2

# ネットワークの要求事項

## 通信の詳細 (パブリック又はFIPSエンドポイント使用時)

送信元	送信先	プロトコル	ポート	用途	エージェントがアクセスするエンドポイント
コンソール 端末の ブラウザ	DataSync agent	TCP	80 (HTTP)	アクティベーションKey取り出し に使用 アクティベーション完了後はポート が閉じられる	
エージェント	AWS	TCP	443 (HTTPS)	AWSアカウントと紐付けてアク ティベーションを実施するのに使 用アクティベーション完了後はブ ロックしてもOK	パブリックエンドポイントのアクティベーション用 activation.datasync. <i>\$region</i> .amazonaws.com FIPSエンドポイントのアクティベーション用 activation.datasync-fips. <i>\$region</i> .amazonaws.com
エージェント	AWS	TCP	443 (HTTPS)	エージェントとサービスの間の制 御及びデータ通信用	APIエンドポイント: datasync. <i>\$region</i> .amazonaws.com データ転送エンドポイント <i>\$taskId</i> .datasync-dp. <i>\$region</i> .amazonaws.com cp.datasync. <i>\$region</i> .amazonaws.com データ転送エンドポイント (FIPSエンドポイント用) cp.datasync-fips. <i>\$region</i> .amazonaws.com エージェントアップデート用 *ここでCloudFrontを使用 repo. <i>\$region</i> .amazonaws.com repo.default.amazonaws.com packages. <i>\$region</i> .amazonaws.com

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/datasync-network.html>

# ネットワークの要求事項（続き）

## 通信の詳細（パブリック又はFIPSエンドポイント使用時）

送信元	送信先	プロトコル	ポート	用途	エージェントがアクセスするエンドポイント
エージェント	DNSサーバー	TCP/UDP	53 (DNS)	エージェントとDNSサーバー間の通信用	
エージェント	AWS	TCP	22 (Support channel)	サポートによるリモート解析用	サポートチャネル 54.201.223.107
エージェント	NTPサーバー	UDP	123 (NTP)	ハイパーバイザーホストの時刻同期用	NTP 0.amazon.pool.ntp.org 1.amazon.pool.ntp.org 2.amazon.pool.ntp.org 3.amazon.pool.ntp.org * エージェントのローカルコンソールから変更可能

参考： <https://docs.aws.amazon.com/datasync/latest/userguide/datasync-network.html>



# 設定の流れ

## STEP① エージェントを準備

仮想アプライアンスを構成



マネジメントコンソールから  
エージェントをアクティベート

DataSync > Agents

Agents (1)  
Agents read data from or write to your on-premises locations.

Search

Agent ID	Name	Service endpoint	Status	Tasks
agent-0e9d2da3e194d0a4f	Agent1	Public	Online	2

## STEP② ロケーションを作成

自己管理ストレージの場合  
ロケーションをエージェントを紐付け



AWSストレージサービスの場合  
ロケーション作成のみ

## STEP③ タスクを作成

STEP②で作成したロケーションを2つ使用してタスクを作成



# AWS DataSyncの設定手順 – STEP①

## エージェントの作成

ハイパーバイザー種別を選択。  
選択したハイパーバイザーに応じて  
仮想マシンのダウンロードリンクが  
表示されます。  
EC2を選択した場合、設定手順への  
リンクが表示されます。

DataSync > Agents > Create agent

### Create agent

**Deploy agent**  
To transfer data between self-managed storage and AWS Storage services, ensure you have deployed and started the agent VM. Your agent can run on either an on-premises hypervisor or Amazon EC2.

Hypervisor

- VMware ESXi ▲
- VMware ESXi ▼
- Kernel-based Virtual Machine (KVM)
- Microsoft Hyper-V
- Amazon EC2

your on-premises VMware ESXi hypervisor.

**Service endpoints**

Choose the service endpoints that this agent will connect to. [Learn more](#)

Public service endpoints in Asia Pacific (Tokyo) ▼

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/activating-agent.html>

# AWS DataSyncの設定手順 – STEP①

## エージェントの作成

Activation Key設定では、エージェントの登録のために、エージェント固有のアクティベーションキーをAWS DataSyncサービスに登録する必要があります。作成したエージェントヘドメイン名やIPアドレスでネットワークアクセス可能であれば、ブラウザでエージェントからキーを取り出して、そのまま登録する事が出来ます。

もし、エージェントがネットワーク的に切り離された環境にある場合、エージェントのローカルコンソールからキーを取り出し、手動で入力する事も可能です。

参考：

<https://docs.aws.amazon.com/datasync/latest/userguide/activating-agent.html>

コンソールからエージェントとAWSコンソール双方にアクセス可能な場合

### Activation key

Activation securely associates the agent that you deployed with your AWS account. [Learn more](#)

Choose a method for obtaining your agent's activation key. [Learn more](#)

**Automatically get the activation key from your agent**  
Your browser will connect to your agent over HTTP to get a unique activation key.

**Manually enter your agent's activation key**  
Choose this option if you already have an activation key and need to enter it manually.

**Agent address**  
When your DataSync agent has been deployed, enter its domain name or IP address. When you click **Get key**, your browser will connect to this address to get a unique activation key.

http://

Make sure this address is accessible from your browser.

エージェントのローカルコンソールからアクティベーションキーを取り出す場合

### Activation key

Activation securely associates the agent that you deployed with your AWS account. [Learn more](#)

Choose a method for obtaining your agent's activation key. [Learn more](#)

**Automatically get the activation key from your agent**  
Your browser will connect to your agent over HTTP to get a unique activation key.

**Manually enter your agent's activation key**  
Choose this option if you already have an activation key and need to enter it manually.

**Activation key**  
Enter the activation key obtained from your agent.

# AWS DataSyncの設定手順 – STEP②

## □ケーションの作成

□ケーション作成段階では、送信元、送信先といった概念はまだありません。これは次のステップのタスクの作成時に決定されます。

6つの□ケーションタイプから選択します。このうちAmazonから始まるものがAWSのストレージサービスで、その他は自己管理ストレージです。エージェントを使った転送の場合、送信元と送信先の組み合わせは必ず自己管理ストレージとAWSストレージサービスの組み合わせになります。

**Create location**

**Configuration**

Location type

Choose the location type ▲

- Amazon EFS file system
- Amazon FSx for Windows File Server
- Amazon S3
- Network File System (NFS)
- Object storage
- Server Message Block (SMB)

□ケーションタイプは6種類

- AWSストレージサービス : Amazon S3, Amazon EFS, Amazon FSx for Windows File Server
- 自己管理ストレージ : NFS, SMB, Object

# AWS DataSyncの設定手順 – STEP②

## □ケーションの作成 – Amazon FSx for Windowsの例

□ケーション種別を選択したら、各種別毎に必要な詳細設定を入力します。

右はFSx for Windowsの例です。

### Create location

#### Configuration

Location type  
Amazon FSx for Windows File Server ▼

Region  
Asia Pacific (Tokyo)

FSx file system  
test filesystem (fs-0854d5f2d335cf4ac) ▼

Share name  
Mount path exported by the FSx server or a subdirectory of that path, delimited with forward slashes (for example, /my-folder/sub-folder).

▼ **Additional settings**

Security groups  
Security group(s) that can access the specified file system.  
Choose security groups ▼

default (sg-0594efbfb17d9c377) ✕

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/create-fsx-location.html>

# AWS DataSyncの設定手順 – STEP③

## タスクの作成

まず送信元ロケーションを選択します。

### Configure source location

#### Source location options

Create or select a source location

Choose the location that you want to transfer data from

Create a new location

Choose an existing location

#### Choose existing location

Region

Asia Pacific (Tokyo)

Existing locations

Choose existing source location

Q |

smb://10.11.0.104/share1/  
Public - Asia Pacific (Tokyo) admin

loc-0748e8bd606e4f59c

fsxw://ap-northeast-1.fs-03df29ba8d3e5a153/share/  
admin

loc-0686ece0d077d462e

Cancel

Next

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/creating-task.html>

# AWS DataSyncの設定手順 – STEP③

## タスクの作成

次に送信先ロケーションを選択します。

The screenshot shows the 'Configure destination location' step in the AWS DataSync console. It is divided into two main sections: 'Destination location options' and 'Choose existing location'.

**Destination location options**  
Create or select a destination location

Choose the location that you want to transfer data from

Create a new location  Choose an existing location

**Choose existing location**

Region  
Asia Pacific (Tokyo)  
Transfers from Server Message Block (SMB) must stay in the current region.

Existing locations

Choose existing destination location

fsxw://ap-northeast-1.fs-03df29ba8d3e5a153/share/admin loc-0686ece0d077d462e

Previous Next

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/creating-task.html>

# AWS DataSyncの設定手順 – STEP③

## タスクの作成

最後にタスクの詳細設定を行います。  
後々管理し易いので、タスクの名前を付けておく事をお勧めします。

### Configure settings

#### Task Name

From-SMB-To-FSxWin

#### Task execution configuration

##### Verify data

Verify the data in your destination after transfer. DataSync always performs data integrity checks while transferring and writing data.

Verify only the data transferred

##### Set bandwidth limit

Allocate maximum bandwidth to be utilized by this task

- Use available
- Set bandwidth limit (MiB/s)

##### ▼ Additional settings

##### Queueing

- Enabled  
Allow this task to be started and queued until one of its agents becomes available

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/creating-task.html>



# メタデータの取り扱い

メタデータは、送信元と送信先のメタデータが同種の場合（例：NFSとEFS）、送信元のメタデータが送信先にコピーされます。送信元と送信先のメタデータの種別が異なる場合（例：WindowsとEFS）、送信先がNFSかオブジェクトの場合はデフォルトPOSIXパーミッションがセットされ、送信先がSMBの場合はSMBへアクセスするユーザーに応じたデフォルトパーミッションがセットされます。

\* 詳細は次頁以降を参照して下さい。

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/special-files.html>

# メタデータの取り扱い

送信元	送信先	メタデータの扱い
SMB	EFS	デフォルトPOSIXメタデータを設定
SMB	S3	デフォルトPOSIXメタデータをユーザーメタデータとして設定
SMB	FSx for Windows	タイムスタンプ（作成日時、最終アクセス日時、最終更新日時）、オーナー、セキュリティーID、標準属性、NTFS DACLをコピー
NFS	EFS	タイムスタンプ（最終更新時間）、ユーザーID、グループID、POSIXパーミッションをコピー
NFS	S3	タイムスタンプ（最終更新時間）、ユーザーID、グループID、POSIXパーミッションをユーザーメタデータとして設定
NFS	FSx for Windows	ロケーションで指定したユーザーがオーナーとなり、そのオーナーのデフォルトWindowsパーミッションを設定
Object	EFS	デフォルトPOSIXメタデータを設定
Object	S3	オブジェクトメタデータとタグをコピー
Object	FSx for Windows	ロケーションで指定したユーザーがオーナーとなり、そのオーナーのデフォルトWindowsパーミッションを設定

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/special-files.html>

# メタデータの取り扱い

送信元	送信先	メタデータの扱い
EFS	EFS	タイムスタンプ（最終更新時間）、ユーザーID、グループID、POSIXパーミッションをコピー
EFS	S3	タイムスタンプ（最終更新時間）、ユーザーID、グループID、POSIXパーミッションをユーザーメタデータとして設定
EFS	FSx for Windows	□ケーションで指定したユーザーがオーナーとなり、そのオーナーのデフォルトWindowsパーミッションを設定
S3	EFS	ユーザーメタデータがある場合：タイムスタンプ（最終更新時間）、ユーザーID、グループID、POSIXパーミッションをコピー ユーザーメタデータがない場合：デフォルトPOSIXメタデータを設定
S3	S3	オブジェクトメタデータとタグをコピー
S3	FSx for Windows	□ケーションで指定したユーザーがオーナーとなり、そのオーナーのデフォルトWindowsパーミッションを設定
FSx for Windows	EFS	デフォルトPOSIXメタデータを設定
FSx for Windows	S3	デフォルトPOSIXメタデータをユーザーメタデータとして設定
FSx for Windows	FSx for Windows	タイムスタンプ（作成日時、最終アクセス日時、最終更新日時）、オーナー、セキュリティID、標準属性、NTFS DACLをコピー

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/special-files.html>

# メタデータの取り扱い

## デフォルトPOSIXパーミッション

パーミッション	値
UID	65534
GID	65534
フォルダーパーミッション	0755
ファイルパーミッション	0644

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/special-files.html>

# AWS DataSyncのセキュリティ

## データ保護について

- AWS DataSyncは転送中のデータをTLSによる暗号化で保護しています。
- S3のマネージド暗号鍵（SSE-S3）を使用したS3上のデータ暗号化、EFS上のデータ暗号化にも対応しています。

## アクセス制御及び監視

- IAMによりリソースへのアクセスを細かに制御する事が出来ます。
- CloudTrailでAPIへのアクセスを追跡する事が出来ます。

## コンプライアンス

- SOC, PCI, FedRAMP, HIPAA等に対応しています。  
詳細は[AWS services in scope by compliance program](#)を参照して下さい。

## 耐障害性

- 他のAWSサービス同様、AWS global infrastructure上に構成されており、複数のAZで動作するように配備されています。

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/security.html>

# AWS DataSyncのクォータ

## タスクのクォータ

リソース	クォータ
アカウント毎の最大タスク数	100
自己管理ストレージからAWSストレージサービスへデータを転送する際のタスク毎の最大ファイル数	5000万ファイル <b>重要：2000万ファイルを超える場合はエージェントのメモリを64GB以上割り当てて下さい</b>
AWSストレージサービス間でデータ転送する際タスク毎の最大ファイル数	2500万ファイル
Snowcone上のDataSyncで転送する際のタスク毎の最大ファイル数	20万ファイル
タスク毎の最大スループット	10Gbps

[AWS Support Center](#)経由で上限緩和申請する事が可能です

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/datasync-limits.html>

# AWS DataSyncのクォータ

## タスク実行のクォータ

リソース	クォータ
タスク実行履歴の保存日数	30日

## フィルターのクォータ

リソース	クォータ
フィルター設定の最大文字数	10万文字

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/datasync-limits.html>

# AWS Snowcone

## AWS OpsHubによる簡単な操作

Amazon S3へのオフラインデータインポート、エクスポート

Amazon S3, Amazon EFS, Amazon FSx for Windows File ServerへのAWS DataSyncを使ったオンラインデータインポート、エクスポート

コンピュータパワー:  
2 CPU X86 at 2.2 GHz,  
4 GB RAM

EC2 コンピューティング  
又はAWS IoT  
Greengrass ファンク  
ション

手軽に持ち運べる重量  
4.5 lbs (2.1 kg)  
(バックパック、ドローン等)

バッテリーによる移動中および移動先での使用

© 2020, Amazon Web Services, Inc. or its Affiliates.



802.1x Wi-Fi connectivity



8 TiBのストレージ容量  
(EthernetとWi-Fi経由のNFSデータコピー)



ミリタリーグレードの暗号化によるセキュリティ

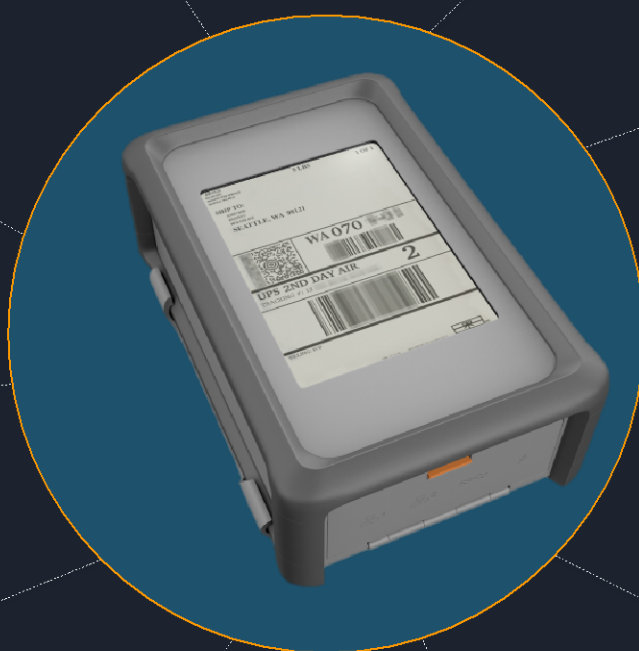
耐タンパー性を持った筐体



凍結温度(0°C/32°F)から砂漠地帯温度(38°C/100°F)まで、困難な環境でも使用可能



防塵性、防水性、防風性を備えた堅牢な筐体



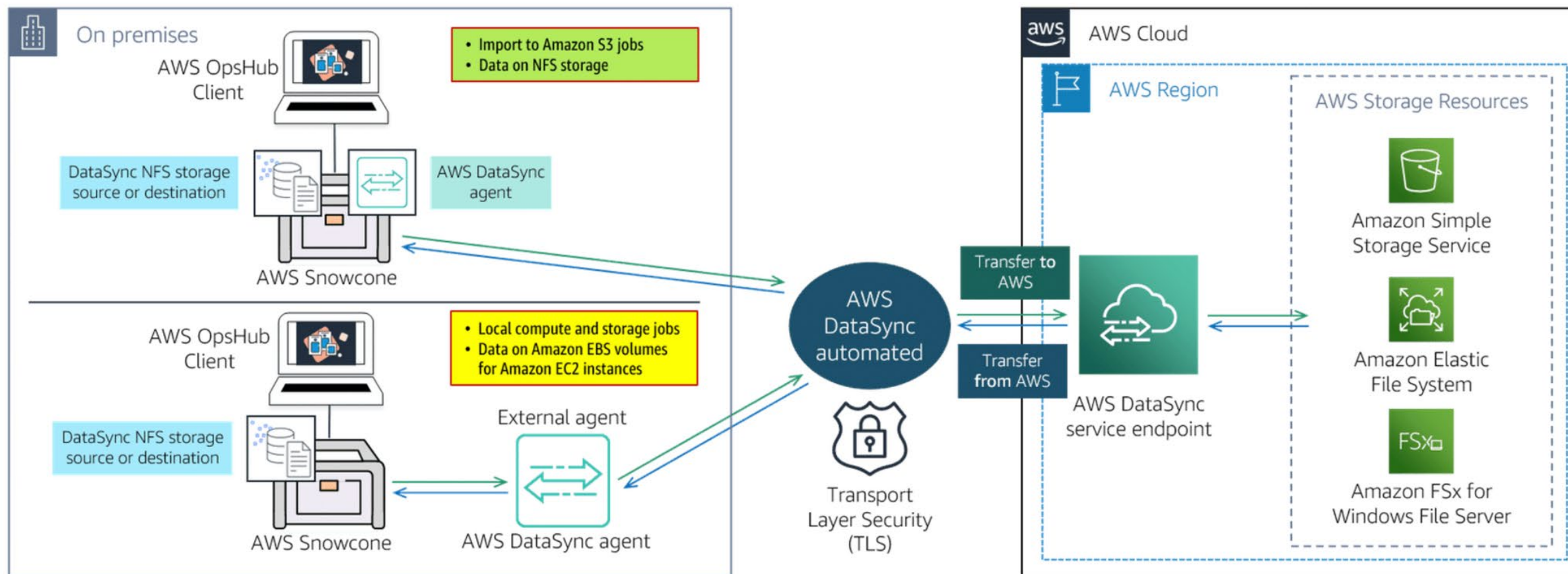
aws



# AWS Snowconeと AWS DataSync

AWS SnowconeにはDataSync Agentが同梱されています

## Online data transfers



# AWS Snowconeと AWS DataSync

## SnowconeでのDataSync Agentの要件

Snowcone instance type	Equivalent EC2 instance type	Compute and memory resources
snc1.micro	t2.micro	1 vCPU and 1 GB of memory
snc1.small	t2.small	1 vCPU and 2 GB of memory
snc1.medium	t2.medium	2 vCPU and 4 GB of memory
Preloaded DataSync client instance	t2.medium	2 vCPU and 4 GB of memory
Preloaded NFS client	t2.micro	1 vCPU and 1 GB of memory

# AWS DataSyncの料金

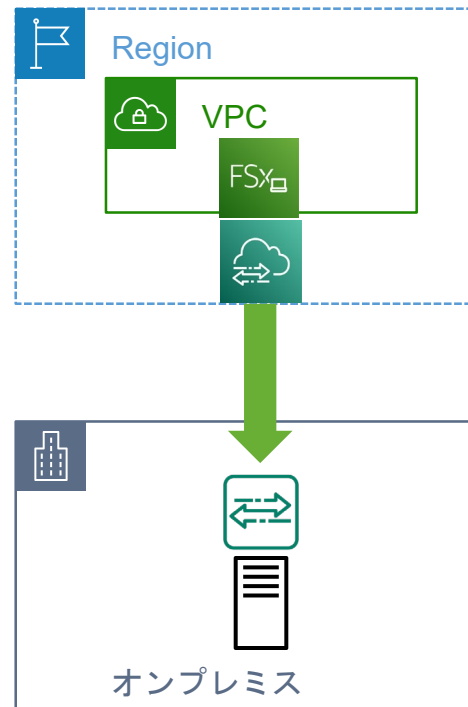
AWS DataSyncの料金はコピーされたデータ量に応じて、GB単位で課金されます

東京 (ap-northeast-1) リージョンの料金 \*2021/3/16時点

コピーされたデータ量当たりの料金	0.0125USD/GiB
------------------	---------------

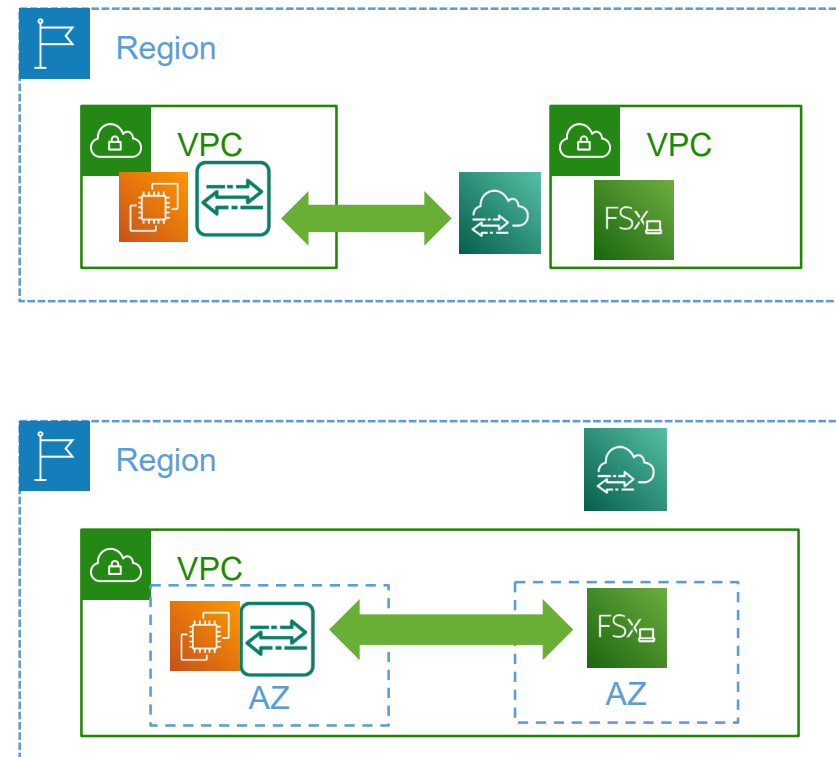
# その他追加料金

AWSからオンプレミス  
への転送の場合



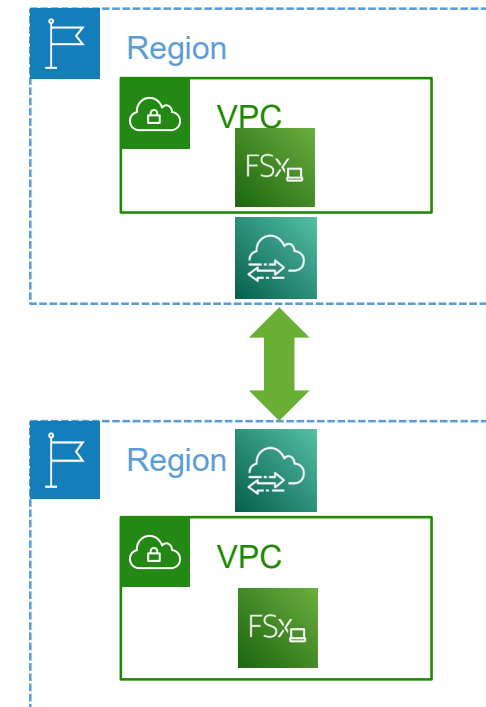
データ転送（送信）の  
料金が別途必要

同一リージョン内  
VPC間又はAZ間転送の場合



リージョン内データ転送  
料金が別途必要

リージョン間  
転送の場合



リージョン間データ転送  
料金が別途必要

参考 : <https://aws.amazon.com/jp/datasync/pricing/>

# AWS DataSyncの料金例① : S3へのアーカイブ

16 MBのファイルで構成された50 TBのデータを東京リージョンのAmazon S3標準へ移行する場合

- AWS DataSync料金

$$50 \text{ TB} * 1,024 \text{ GB} * 0.0125 \text{ USD/GB} = 640 \text{ USD}$$

Amazon S3リクエスト料金

$$(1 \text{ S3 LIST リクエスト} * 0.0047 \text{ USD/1000}) +$$

$$(50 \text{ TB}/16 \text{ MB S3 PUT リクエスト} * 0.0047 \text{ USD/1000}) = 14.69 \text{ USD}$$

合計  $640 + 14.69 = 654.69 \text{ USD}$

参考 : <https://aws.amazon.com/jp/datasync/pricing/>

# AWS DataSyncの料金例②： Amazon FSx for Windowsへのファイルサーバー移行

オンプレミスの Windows File Server から Amazon FSx に 100 TB を移行する場合

- AWS DataSync料金  
 $100 \text{ TB} * 1,024 \text{ GB} * 0.0125 \text{ USD/GB} = 1,280 \text{ USD}$

参考：<https://aws.amazon.com/jp/datasync/pricing/>

# AWS DataSyncの料金例③： Amazon EFSへの継続的なデータ取り込み

Amazon EFS に毎日 1 TB を転送した時の月額（31 日間）費用

- AWS DataSync料金  
 $1 \text{ TB} * 1,024 \text{ GB} * 0.0125 \text{ USD/GB} * 31 = 396.8 \text{ USD}$

参考：<https://aws.amazon.com/jp/datasync/pricing/>

# AWS DataSyncの料金例④： Disaster Recoveryのためのレプリケーション

10 TB のオンプレミスWindows File Serverデータをレプリケーションし、その後日々1 TBの差分データを更新する場合

- AWS DataSync料金
  - 10 TB \* 1,024 GB \* 0.0125 USD/GB = 128 USD (初回転送費用)
  - 1 TB \* 1,024 GB \* 0.0125 USD/GB \* 31 = 396.8 USD/月 (増分月額費用)

参考：<https://aws.amazon.com/jp/datasync/pricing/>

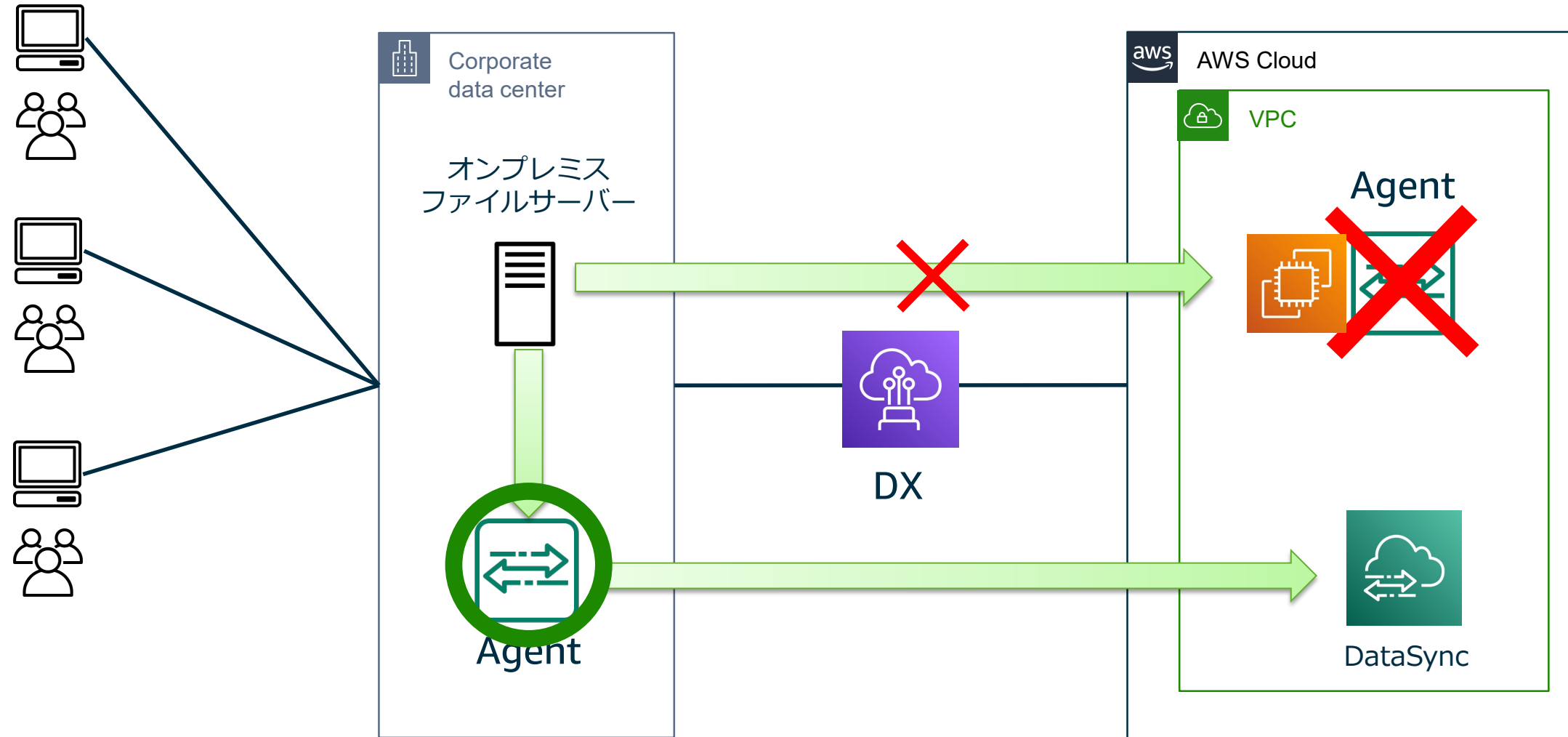




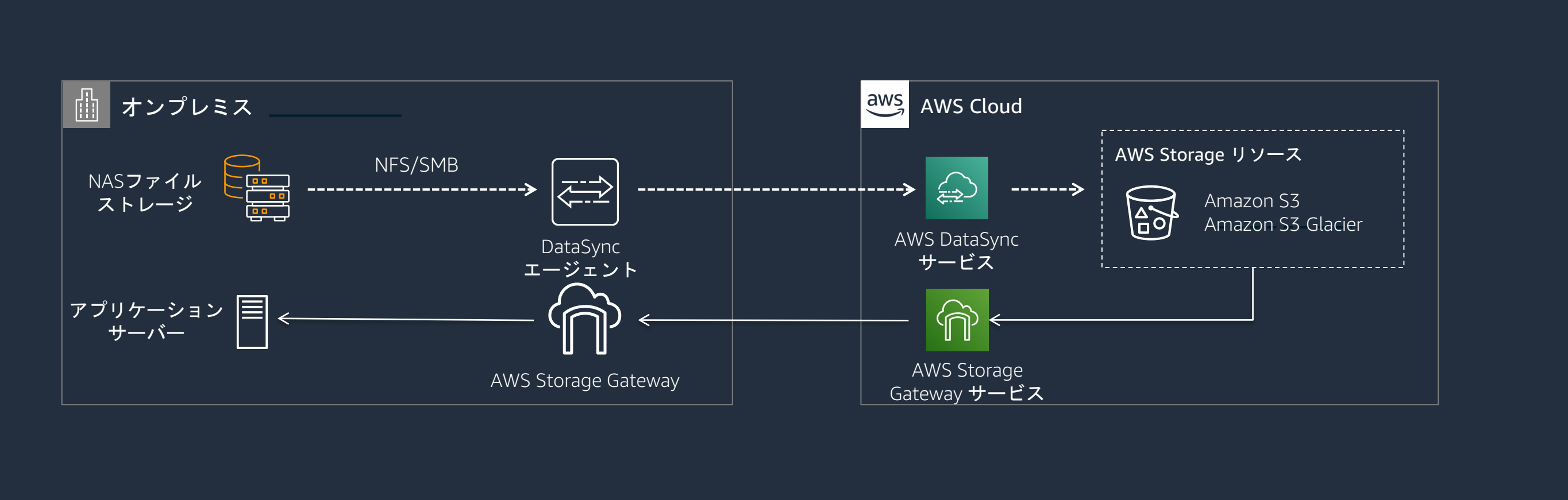
# ベストプラクティス&TIPS



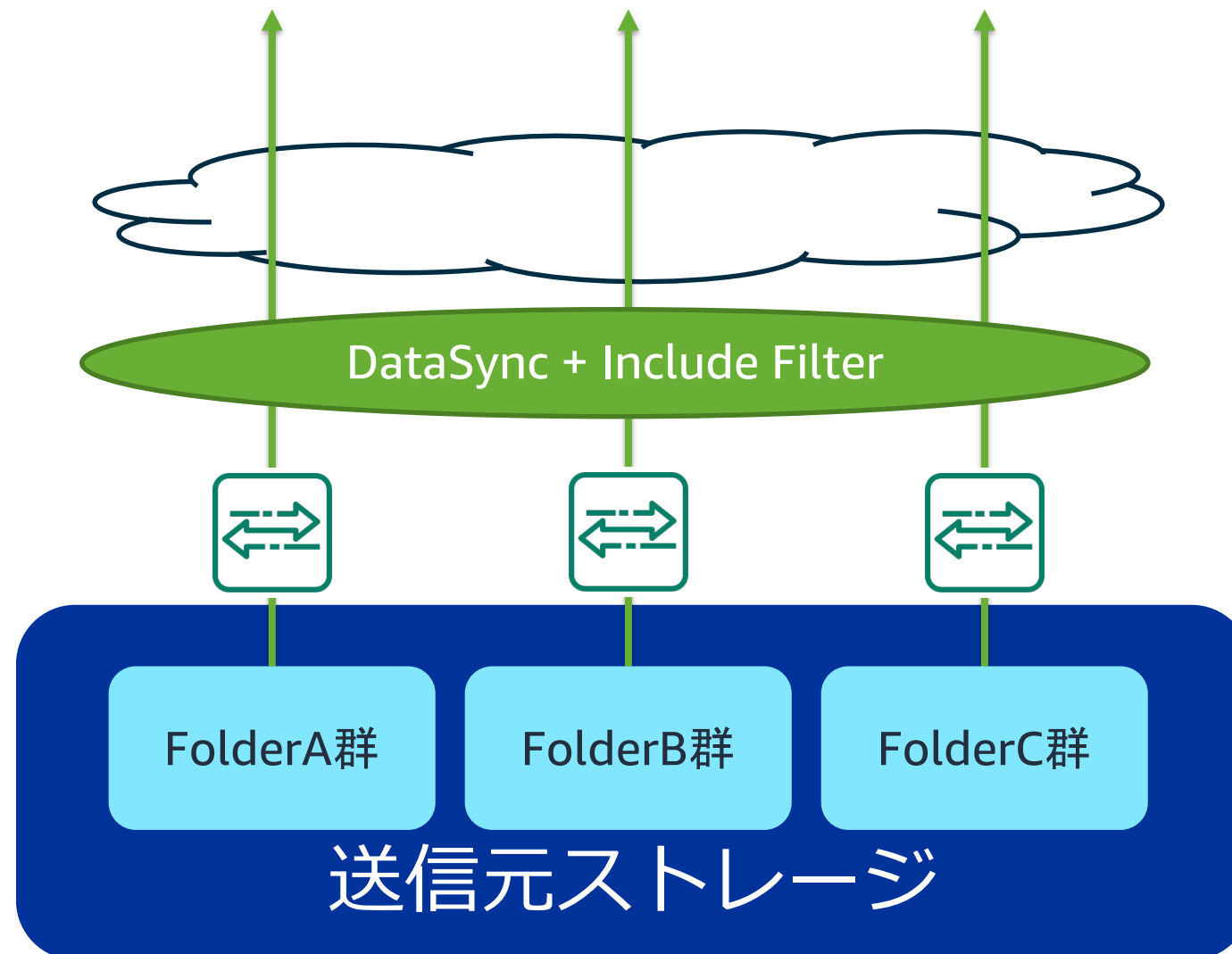
# エージェントは送信元の近くに置きましょう



# 高速に大量ファイルを移動するのがDataSync クラウドストレージへのアクセスを提供するのがStorage Gateway



# パフォーマンスが足りない場合はエージェントを並列化



参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/multiple-agents.html>

# ロケーションに関するTIPS

## EFSロケーションの作成時の考慮点

- 専用テナンシーを持つVPCに存在するEFSはサポート対象外です。
- バーストスループットモードのEFSを使用している場合、DataSyncの転送がバーストクレジットを消費し、パフォーマンスに影響を与える場合があります。アクティブなワークロードが動作しているファイルシステムにDataSyncで転送する場合は、プロビジョンドスループットの使用も検討して下さい。
- 汎用パフォーマンスモードのEFSには35,000 ファイルシステムオペレーション毎秒の制限があります。この制限がDataSyncの最大スループットに影響します。データやメタデータの読み取りには1 ファイルシステムオペレーションを消費し、データの書き込みやメタデータの更新は5 ファイルシステムオペレーションを消費します。つまり、35,000読み取りオペレーション毎秒、又は7,000書き込みオペレーション毎秒、又はそれぞれの混在したパフォーマンス値が最大となります。ファイルオペレーションは全てのクライアントからのアクセスが累計されます。

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/create-efs-location.html>

# ロケーションに関するTIPS

## FSx for Windowsロケーションの作成時の考慮点

- 専用テナンシーを持つVPCに存在するFSx for Windowsはサポート対象外です。
- 共有名に日本語は使えません。\*2021/3/16現在
- FSx for WindowsのスループットやIOPSがボトルネックにならないよう、DataSyncによる移行時の一時的なスループットキャパシティの増加やSSDの使用を検討して下さい。  
(<https://docs.aws.amazon.com/fsx/latest/WindowsGuide/performance.html>)

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/create-fsx-location.html>

# ロケーションに関するTIPS

## S3ロケーションの作成時の考慮点

- S3-Intelligent-Tiering、S3-Standard-IA、S3-OneZone-IAでは、128KB以下のオブジェクトはStandardとして扱われます  
既存オブジェクトの変更、削除、取り出しにより追加のコストが発生する事もあります
- S3 Glacier、S3 Glacier Deep Archiveでは、40KB以下のオブジェクトはStandardとして扱われます  
既存オブジェクトの変更、削除、取り出しにより追加のコストが発生する事もあります  
また、整合性検証モードで“Verify all data in the destination”を選択する事は出来ません

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/create-s3-location.html>

# ロケーションに関するTIPS

## NFSロケーションの作成時の考慮点

- 全てのデータを転送するために、DataSyncエージェントが全てのファイルへの読み出し権限を持つ必要が有ります。これを実現するために、no\_root\_squashオプションを付けてExportするか、全てのユーザーに読み出し権限を与えて下さい。また、エージェントがディレクトリにアクセスするには別途実行権限も必要です。
  - Snowcone内部のNFSサーバーからDataSyncでデータを転送するには以下のような設定が必要です
    - エージェント：Snowconeで起動したEC2エージェントを指定します
    - NFSサーバー：Snowcone上のNFSサーバーにアタッチされた仮想IPアドレスを指定します
    - マウントパス：S3バケットを想定したパス構成でフォーマットされており、次のようなパスで指定します。/buckets/<bucket-name>
- \* Snowconeに関する詳細はSnowconeのドキュメントを参照して下さい。  
<https://docs.aws.amazon.com/snowball/latest/snowcone-guide/use-data-sync.html>

参考：<https://docs.aws.amazon.com/datasync/latest/userguide/create-nfs-location.html>



# ロケーションに関するTIPS

## Objectロケーションの作成時の考慮点

- Objectロケーションを使用するには、対象のオブジェクトストレージが以下のAmazon S3 APIに対応している必要があります
  - AbortMultipartUpload
  - DeleteObject
  - GetBucketLocation
  - GetObject
  - GetObjectTagging
  - HeadBucket
  - ListBucket
  - ListBucketMultipartUploads
  - ListMultipartUploadParts
  - PutObject
  - PutObjectTagging

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/create-object-location.html>

# ロケーションに関するTIPS

## Objectロケーションの作成時の考慮点

- Objectロケーションを使用するには、対象のオブジェクトストレージが認証時にAWS signature version 4に対応している必要があります  
AWS signature version 2はサポート対象外です

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/create-object-location.html>

# タスクに関するTIPS

Verify dataの項目では3つの整合性検証方式から選択出来ます。デフォルトは差分データのみ整合性検証を行うようになっています。

Set bandwidth limitにより、転送帯域に制限をかけて、他の業務通信に影響を与えないようにする事も可能です。この値はタスクの実行中にダイナミックに変更する事も出来ます。

Queueingは特に理由が無ければデフォルトのまま使用して下さい。エージェントが同時実行出来るのは1タスクまでです。

## Task execution configuration

### Verify data

Verify the data in your destination after transfer. DataSync always performs data integrity checks while transferring and writing data.

Verify only the data transferred ▼

### Set bandwidth limit

Allocate maximum bandwidth to be utilized by this task

- Use available
- Set bandwidth limit (MiB/s)

### ▼ Additional settings

### Queueing

- Enabled  
Allow this task to be started and queued until one of its agents becomes available

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/creating-task.html>

# タスクに関するTIPS

Transfer modeでは差分転送を行うか、全てのデータを転送するかを選択できます。デフォルトは差分転送です。DataSyncは差分比較のため、データ転送前に送信元と送信先、双方のファイルをスキャンして、ファイルのリスト、ファイルのメタデータの読み込みを行います。ファイル数が多い場合、このスキャンに時間を要する事も有ります。データを1-Shotで転送する場合や、初回の同期作業の場合、差分比較する必要が無いいため、全量同期を選ぶ事によってスキャン時間を短縮出来ます。

その他、送信元で削除済みのデータを送信先で残しておくかどうか、同じ名前のファイルを上書きするかどうかを選択する事も出来ます。

**Data transfer configuration**

Transfer mode

Transfer only data that has changed

- Keep deleted files  
Keep files in destination even when deleted from source
- Overwrite files  
Overwrite files in destination with files from source

**Data transfer configuration**

Transfer mode

Transfer only data that has changed

Transfer only data that has changed  
Copy only data and metadata that differ between the source and the destination.

Transfer all data  
Copy all source content to the destination, without comparing to existing content on the destination.

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/creating-task.html>

# タスクに関するTIPS

転送するフォルダーやファイルに対してフィルターを設定する事も出来ます。

フィルターにはExcludeフィルターとIncludeフィルターの2種類が存在します。

2021年3月16日現在の仕様では、タスクの設定で使用可能なのはExcludeフィルター、タスクの実行時に使用可能なのはIncludeフィルターです。

フィルターの書き方のルールには制限が有りますので、こちらの公式ドキュメントを参照して下さい。

<https://docs.aws.amazon.com/datasync/latest/userguide/filtering.html>

### Filtering configuration

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](#)

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/creating-task.html>

# タスクに関するTIPS

スケジュールを設定してタスクを定期的に自動実行させる事が出来ます。1時間毎、1日毎、1週間毎、曜日指定、カスタムから選ぶ事が出来ます。  
カスタムの場合はCronの表現方式を使って指定する事が出来ます。

## 1日毎指定の例

**Schedule**  
Execute this task on a schedule by specifying the time and frequency.

Frequency  
Daily ▼

at  
HH:MM (UTC)

## カスタムの例

**Schedule**  
Execute this task on a schedule by specifying the time and frequency.

Frequency  
Custom ▼

Cron expression  
Use a cron expression for complex schedules. [Learn more and see syntax examples](#) ↗

0/15 \* \* \* ? \*

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/creating-task.html>

# タスクに関するTIPS

タスクの実行結果のログをCloudWatch Logに出力出来ます。Autogenerateボタンを押す事でCloudWatch log groupを自動生成する事も出来ます。ログを保存しない、エラーだけ保存する、全てのログを保存するの3つから選択出来ます。デフォルトはエラーだけ保存するです。



### Task logging

**Log level**  
Defines the level of detail that is logged to CloudWatch Logs. CloudWatch Logs pricing applies for logs generated by DataSync.

Log basic information such as transfer errors ▼

**CloudWatch log group**  
Log task execution activity such as error or other failures.

DataSyncLogs-MigrationWorkshop ▼

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/creating-task.html>

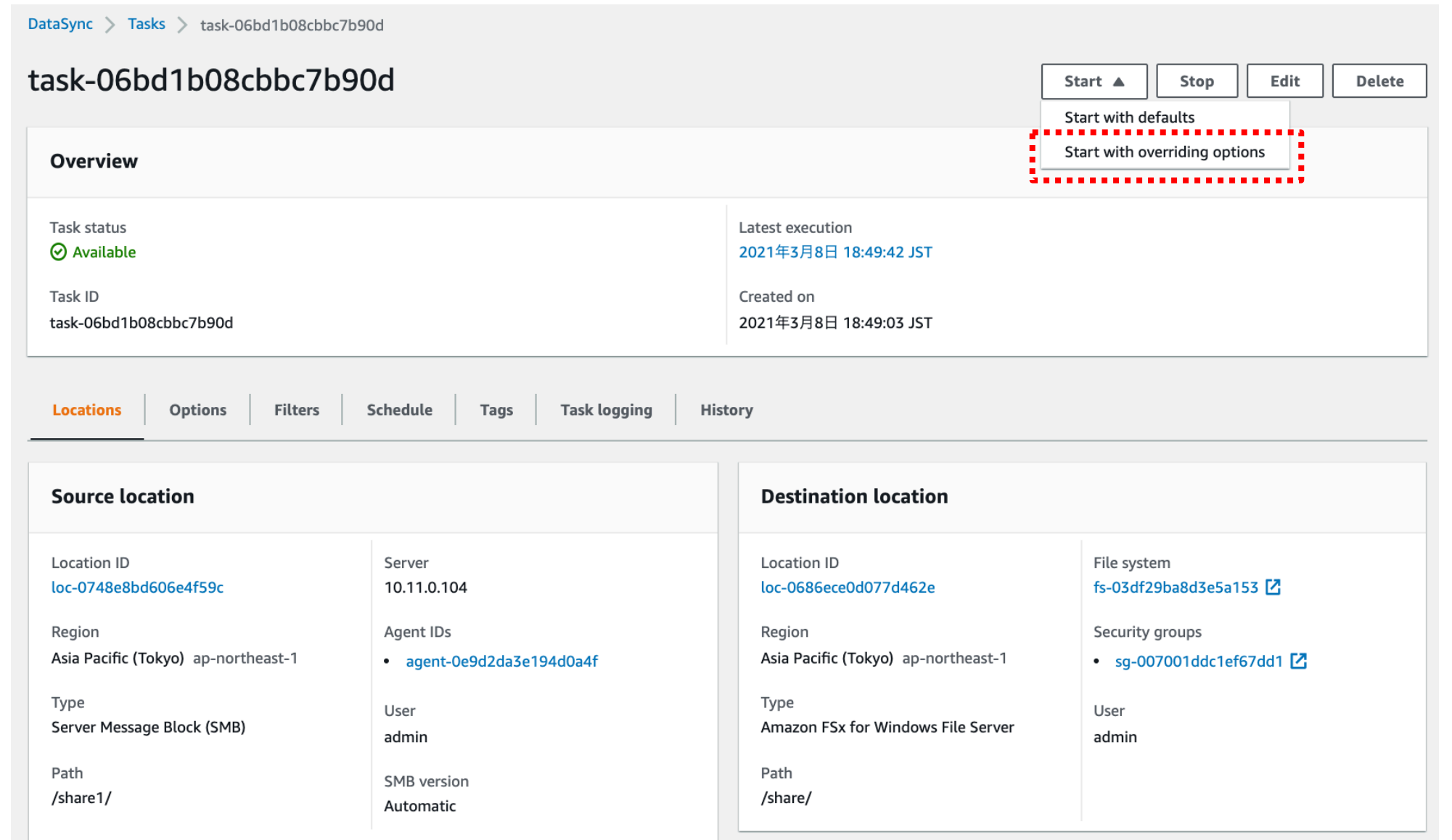
# タスクに関するTIPS

タスクを作成し、ステータスが Availableになるとタスクの実行が可能になります。

タスクを選択し、Startから手動でタスク実行する事が出来ます。

この時、Start with overriding optionを選ぶ事で、実行前にタスク設定を変更して実行する事が出来ます。

Includeフィルターを使えるのもこのタイミングです。



DataSync > Tasks > task-06bd1b08cbbc7b90d

## task-06bd1b08cbbc7b90d

Start ▲ Stop Edit Delete

Start with defaults  
Start with overriding options

### Overview

Task status Available	Latest execution 2021年3月8日 18:49:42 JST
Task ID task-06bd1b08cbbc7b90d	Created on 2021年3月8日 18:49:03 JST

Locations | Options | Filters | Schedule | Tags | Task logging | History

### Source location

Location ID loc-0748e8bd606e4f59c	Server 10.11.0.104
Region Asia Pacific (Tokyo) ap-northeast-1	Agent IDs • agent-0e9d2da3e194d0a4f
Type Server Message Block (SMB)	User admin
Path /share1/	SMB version Automatic

### Destination location

Location ID loc-0686ece0d077d462e	File system fs-03df29ba8d3e5a153
Region Asia Pacific (Tokyo) ap-northeast-1	Security groups • sg-007001ddc1ef67dd1
Type Amazon FSx for Windows File Server	User admin
Path /share/	



# リンクとディレクトリの取り扱い

## ハードリンク

- NFSサーバーとEFSの間ではハードリンクはコピーされます。
- ハードリンクをS3へコピーする場合、ハードリンクは一度だけコピーされ、ファイルとは別オブジェクトとして保存されます。  
リンクに変更が無ければNFSサーバーやEFSへ戻した時に復元も可能です。
- SMBファイル共有に対してはハードリンクはサポートしていません。  
仮に検出した場合は処理がスキップされ、CloudWatch logに記録されます。

## シンボリックリンク

- NFSサーバーとEFSの間ではシンボリックリンクはコピーされます。
- シンボリックリンクをS3へコピーする場合、ターゲットパスの情報がオブジェクトに保存されます。  
NFSサーバーやEFSへ戻した時に復元も可能です。
- SMBファイル共有に対してはシンボリックリンクはサポートしていません。  
仮に検出した場合は処理がスキップされ、CloudWatch logに記録されます。

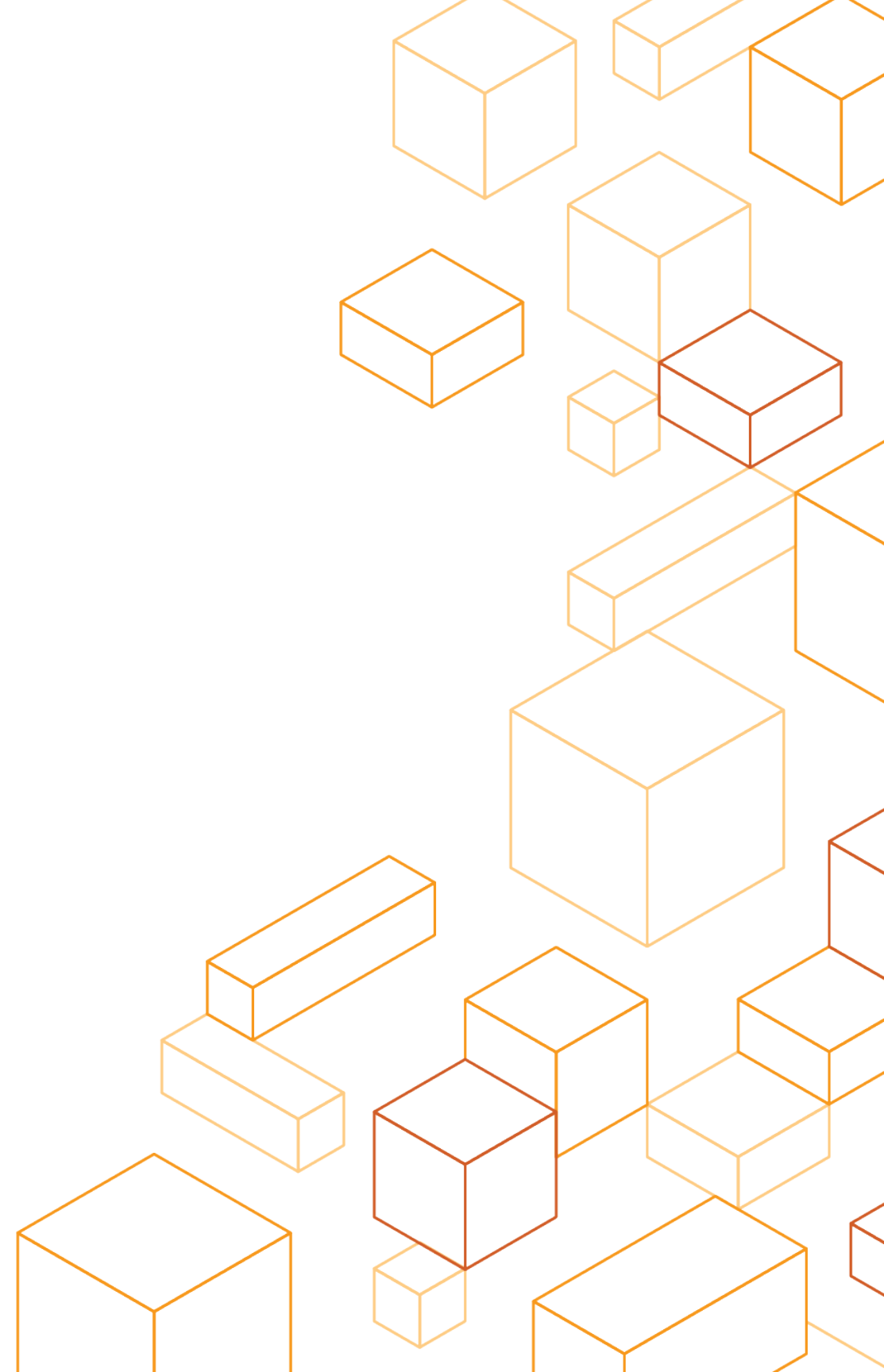
## ディレクトリ

- S3へディレクトリをコピーした場合、ディレクトリはパスの末尾が"/"で終わる空のオブジェクトとして登録されます。

参考 : <https://docs.aws.amazon.com/datasync/latest/userguide/special-files.html>



# 顧客事例



# オンプレストレージのS3へのマイグレーション



## 抱えていた問題

オンプレミスのストレージ機器は高価で、多くのスペースが必要

複雑なアーキテクチャとキャッシュ階層

多数のNFSクライアントからのリードアクセス

「『データ同期ツールが速すぎる!』と言えるなんて、素晴らしいです!

AWS DataSyncにより我々は俊敏性を手に入れ、プロジェクトを数ヶ月間引きずる必要がなくなりました。

## ソリューション

S3へのデータ移行にDataSyncを使用

ライフサイクルポリシーを使用した、アクティブ-アクティブマルチリージョン構成とバージョニングの使用

移行後のクラウドデータへのアクセスにファイルゲートウェイを使用

*Dan Boisvert*  
*Lead Software Engineer*

## 効果

ロケーション毎に85%のストレージコスト削減

ストレージエンジニアがより価値の高い、戦略的な仕事に注力可能に

# 数百TBのオンプレバックアップデータをS3へ移行



## 抱えていた問題

数ペタバイトのオンプレミスData Domainストレージからの脱却

複数年に渡るデータ保存の必要性

低コストで従量課金のAmazon S3をバージョンニングと共に活用

S3への移行にDataSyncを活用

全てのデータをバイト毎に比較して検証しながら、成功裏にS3へ移行

オンプレミスのData Domainストレージを廃棄

「オンプレミスから AWS へのペタバイト規模のデータ移行は、AWS DataSyncを使う事で最小限の労力で自分達の手で迅速に達成出来ました。このソリューションはゲームチェンジャーです!

## ソリューション

## 効果

Satish Kumar

Infrastructure Engineer

「AWS DataSync を使用して数百 TB のデータを Amazon S3 に移行」

# Q&A

お答えできなかったご質問については

AWS Japan Blog <https://aws.amazon.com/jp/blogs/news/> に  
後日掲載します。

# AWS の日本語資料の場所「AWS 資料」で検索



The screenshot shows the AWS Japanese website header with the AWS logo on the left. On the right, there are links for '日本担当チームへお問い合わせ' (Contact our Japanese team), 'サポート' (Support), '日本語' (Japanese), and 'アカウント' (Account), along with a 'コンソールにサインイン' (Sign in to the console) button. Below the header is a navigation menu with links for '製品' (Products), 'ソリューション' (Solutions), '料金' (Pricing), 'ドキュメント' (Documentation), '学習' (Learning), 'パートナー' (Partners), 'AWS Marketplace', and 'その他' (Other), followed by a search icon. The main content area features a large heading 'AWS クラウドサービス活用資料集トップ' (AWS Cloud Service Usage Resource Collection Top). Below this is a paragraph of text: 'アマゾン ウェブ サービス (AWS) は安全なクラウドサービスプラットフォームで、ビジネスのスケールと成長をサポートする処理能力、データベースストレージ、およびその他多種多様な機能を提供します。お客様は必要なサービスを選択し、必要な分だけご利用いただけます。それらを活用するために役立つ日本語資料、動画コンテンツを多数ご提供しております。(本サイトは主に、AWS Webinar で使用した資料およびオンデマンドセミナー情報を掲載しています。)' At the bottom of the main content area, there are four buttons: 'AWS Webinar お申込' (AWS Webinar Registration), 'AWS 初心者向け' (AWS for Beginners), '業種・ソリューション別資料' (Resources by Industry/Solution), and 'サービス別資料' (Resources by Service).

aws

日本担当チームへお問い合わせ サポート 日本語 ▼ アカウント ▼ コンソールにサインイン

製品 ソリューション 料金 ドキュメント 学習 パートナー AWS Marketplace その他 🔍

## AWS クラウドサービス活用資料集トップ

アマゾン ウェブ サービス (AWS) は安全なクラウドサービスプラットフォームで、ビジネスのスケールと成長をサポートする処理能力、データベースストレージ、およびその他多種多様な機能を提供します。お客様は必要なサービスを選択し、必要な分だけご利用いただけます。それらを活用するために役立つ日本語資料、動画コンテンツを多数ご提供しております。(本サイトは主に、AWS Webinar で使用した資料およびオンデマンドセミナー情報を掲載しています。)

AWS Webinar お申込 » AWS 初心者向け » 業種・ソリューション別資料 » サービス別資料 »

<https://amzn.to/JPArchive>

# AWS Well-Architected 個別技術相談会

毎週“W-A個別技術相談会”を実施中

- AWSのソリューションアーキテクト(SA)に  
対策などを相談することも可能

- **申込みはイベント告知サイトから**  
(<https://aws.amazon.com/jp/about-aws/events/>)

**AWS イベント**

**で[検索]**

AWS Well-Architected



# 4月以降のBlack Belt Online Seminarについて

ライブ配信によるBlack Belt Online Seminarは3月一杯で終了し、  
今後はオンデマンドによる定期配信に変更いたします。

今後もコンテンツを拡充して行きますので、楽しみにお待ちください。

オンデマンドでの配信スケジュールは、AWS Blog, AWSニュースレターでお知らせいたします（5月17日週に再開を予定しています）





# ご視聴ありがとうございました

<https://amzn.to/JPWebinar>



<https://amzn.to/JPArchive>

