

AWS-57

# Amazon S3 による 効果的なデータ保存・保護方法

焼尾 徹

ストレージスペシャリスト ソリューションアーキテクト  
アマゾン ウェブ サービス ジャパン合同会社



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

# このセッションについて

- 本セッションの対象となる方

- AWSの活用を開始したものの、Amazon Simple Storage Service (S3) をもっと有効に活用したいと考えている方
- すでに AWS は活用しているが、Amazon S3 に格納しているデータはしっかり保護したいと考えている方
- こういった課題があるお客様にシステムの提案をする立場の方

- ねらい

- Amazon S3 データ保護方法を整理し、データを効果的に活用する

# 内容

- 増加するデータがもたらす課題とAmazon S3 の歴史
- Amazon S3 を支える基盤
- Amazon S3 での暗号化
- Amazon S3 でのレプリケーション
- Amazon S3 での誤削除対策
- Amazon S3 へのデータアクセスパス
- Amazon S3 ストレージクラス
- Amazon S3 を軸にしたデータアーキテクチャ
- まとめ

- ➡ 増加するデータがもたらす課題とAmazon S3 の歴史
- Amazon S3 を支える基盤
- Amazon S3 での暗号化
- Amazon S3 でのレプリケーション
- Amazon S3 での誤削除対策
- Amazon S3 へのデータアクセスパス
- Amazon S3 ストレージクラス
- Amazon S3 を軸にしたデータアーキテクチャ
- まとめ

# 増加するデータがもたらす課題と Amazon S3 の歴史

# 急激なデータ増加がもたらす課題



## チャレンジ

データがサイロ化する、拡張が困難になる、コストの増加傾向という悩み



## リスク

管理が複雑になることで生じるデータ保護やビジネス継続におけるリスク



## 活用

せっかく**収集**したデータをうまく**利用**するには、相応のイノベーションを必要とする

# AWS ストレージは クラウドにおける データアーキテクチャの 基礎となる

必要な時、  
必要な場所で利用する

常に保護する

うまく活用  
する



コストの  
最適化

# AWS ストレージ上でデータアーキテクチャを構築

## データは常に保護されている

データはセキュアに扱われ、バックアップを取得し、災害・障害対策のためにレプリケーションが可能です。誤った削除からデータの復旧ができるようにします。

## データは必要なときに必要なところから利用できる

データはどこからでも必要なときに利用でき、業界をリードする性能でワークロードを支えます。

## データはコスト最適化されている

低コストなアーカイブストレージの活用や、柔軟なストレージクラス、自動階層化により、コスト削減が可能です。

## データはうまく活用することができる

豊富な分析サービスや機械学習サービスと連携可能なデータレイクを構築することができます。



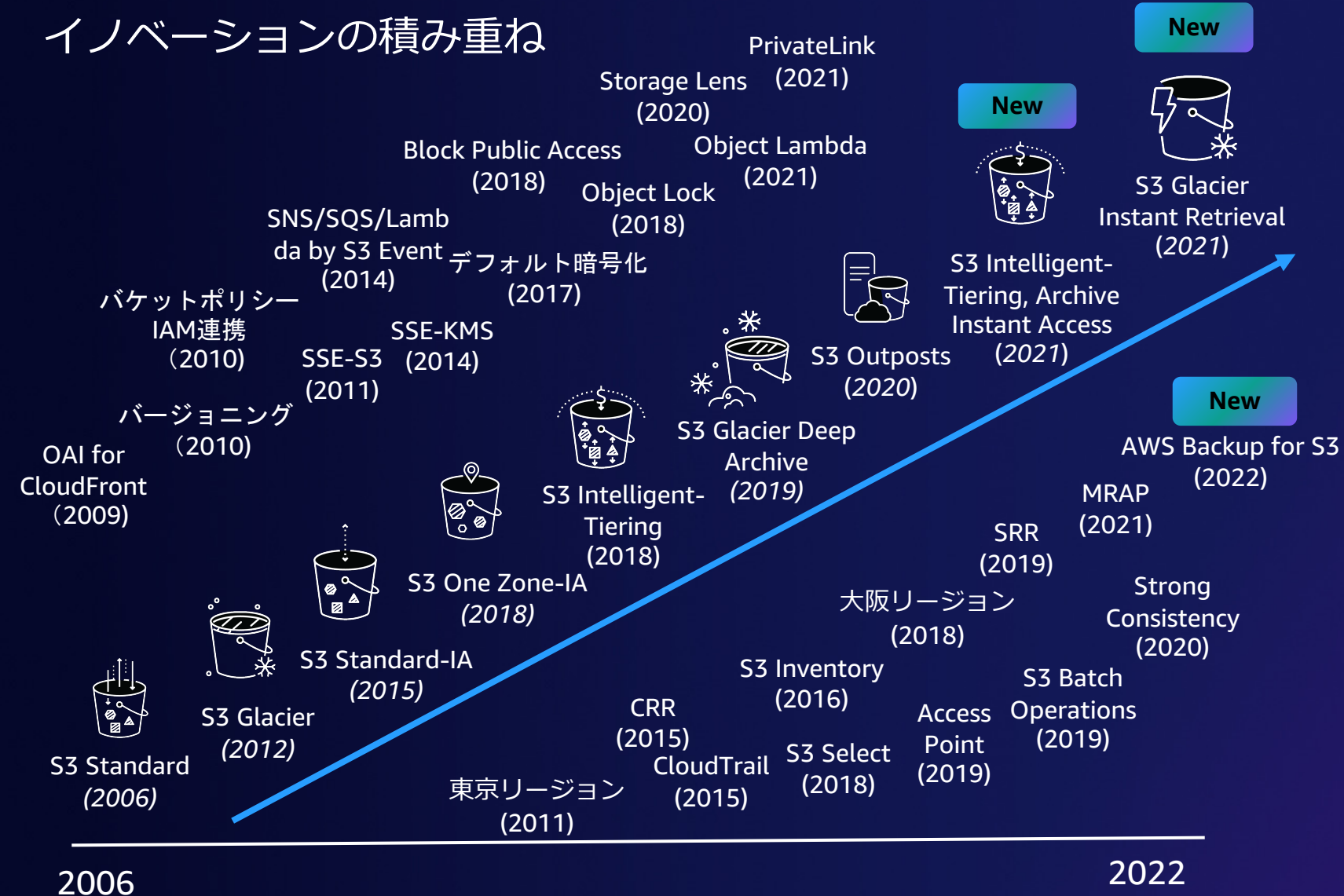
## Amazon S3

- ✓ 実質的に無制限のスケーラビリティ
- ✓ 最も優れた性能
- ✓ 最も安全性の高い
- ✓ 管理しやすい
- ✓ アーカイブデータを低コストで保存
- ✓ 16年にわたるイノベーションの積み重ね
- ✓ **New** アーカイブ用途でありながら、即時取り出しを実現

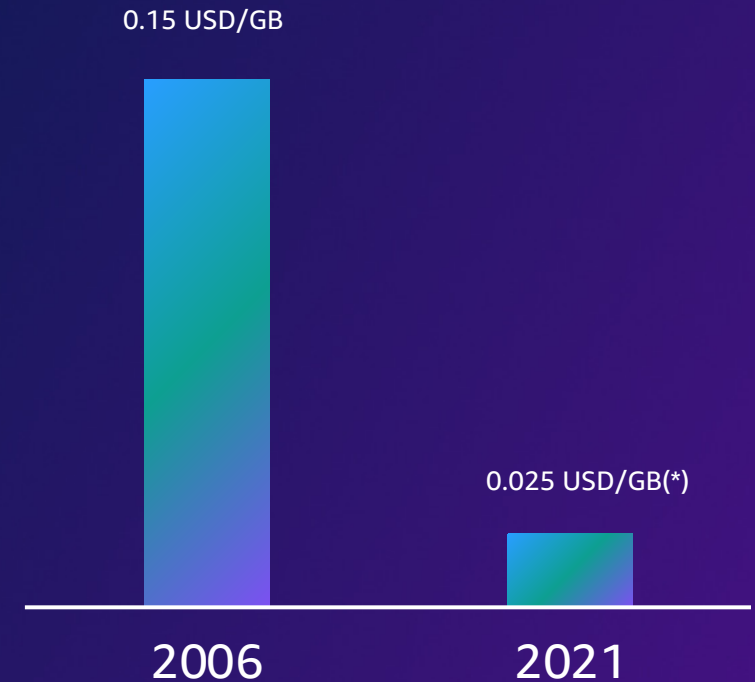


# Amazon S3 の歴史

イノベーションの積み重ね



容量単価の削減



増加するデータがもたらす課題とAmazon S3 の歴史  
→ Amazon S3 を支える基盤  
Amazon S3 での暗号化  
Amazon S3 でのレプリケーション  
Amazon S3 での誤削除対策  
Amazon S3 へのデータアクセスパス  
Amazon S3 ストレージクラス  
Amazon S3 を軸にしたデータアーキテクチャ  
まとめ

データは常に保護されている

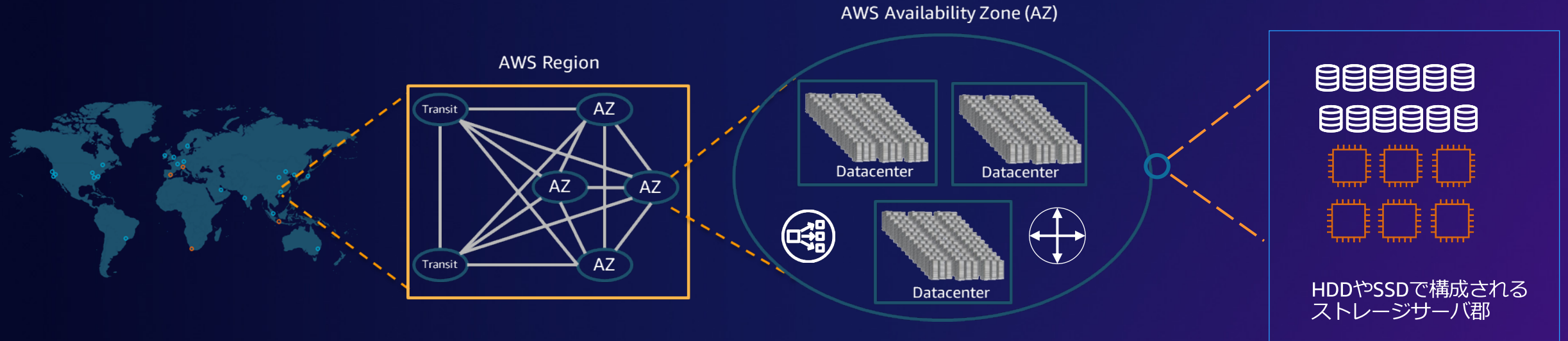
# Amazon S3 を支える基盤

# Amazon S3 のユーザからの見え方



Amazon Simple Storage Service (S3) バケット

# Amazon S3 を支える基盤



99.9999999999% (11 9s) のデータ耐久性

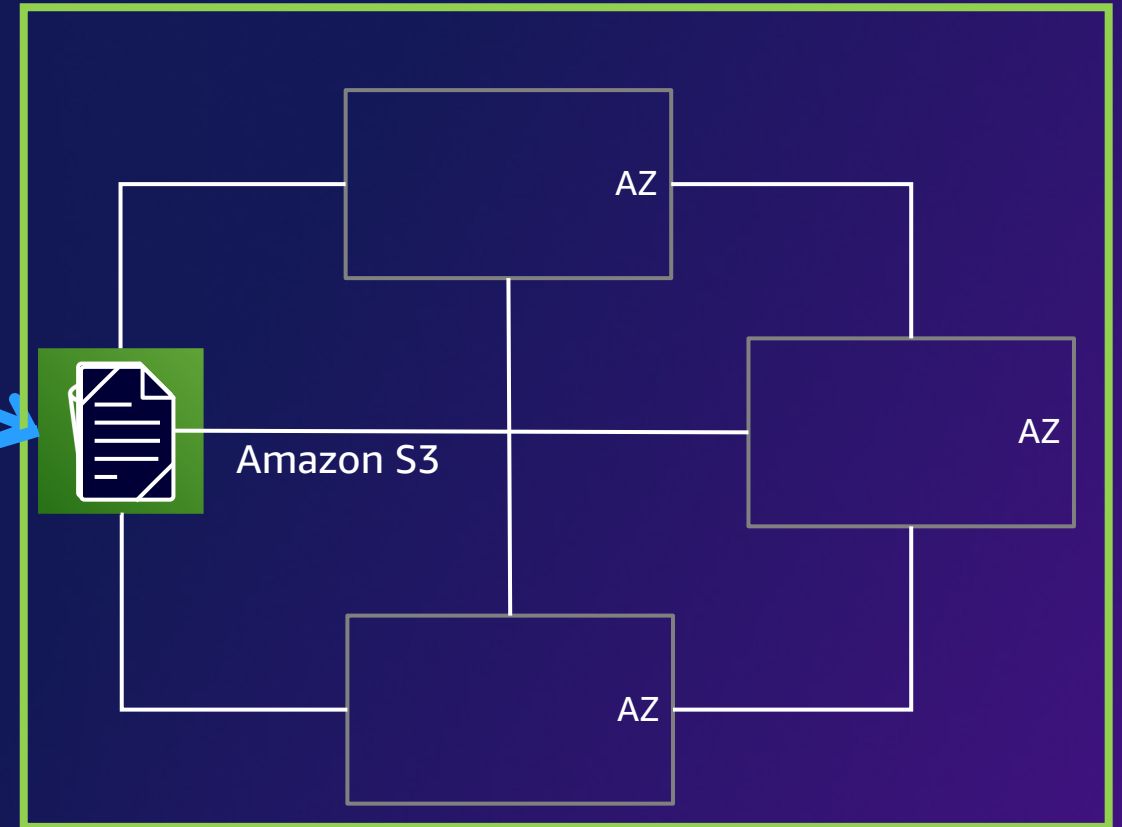
# Amazon S3 データ整合性モデル

強力な整合性 (Strong Consistency)

データをS3に格納する



データを利用する

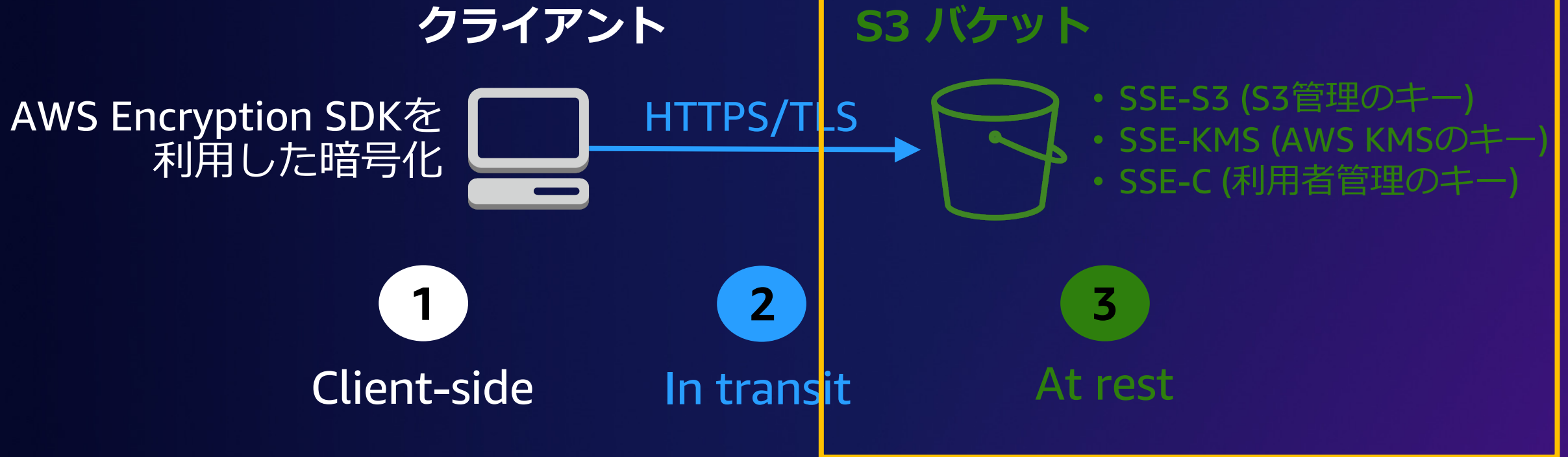


増加するデータがもたらす課題とAmazon S3 の歴史  
Amazon S3 を支える基盤  
→ Amazon S3 での暗号化  
Amazon S3 でのレプリケーション  
Amazon S3 での誤削除対策  
Amazon S3 へのデータアクセスパス  
Amazon S3 ストレージクラス  
Amazon S3 を軸にしたデータアーキテクチャ  
まとめ

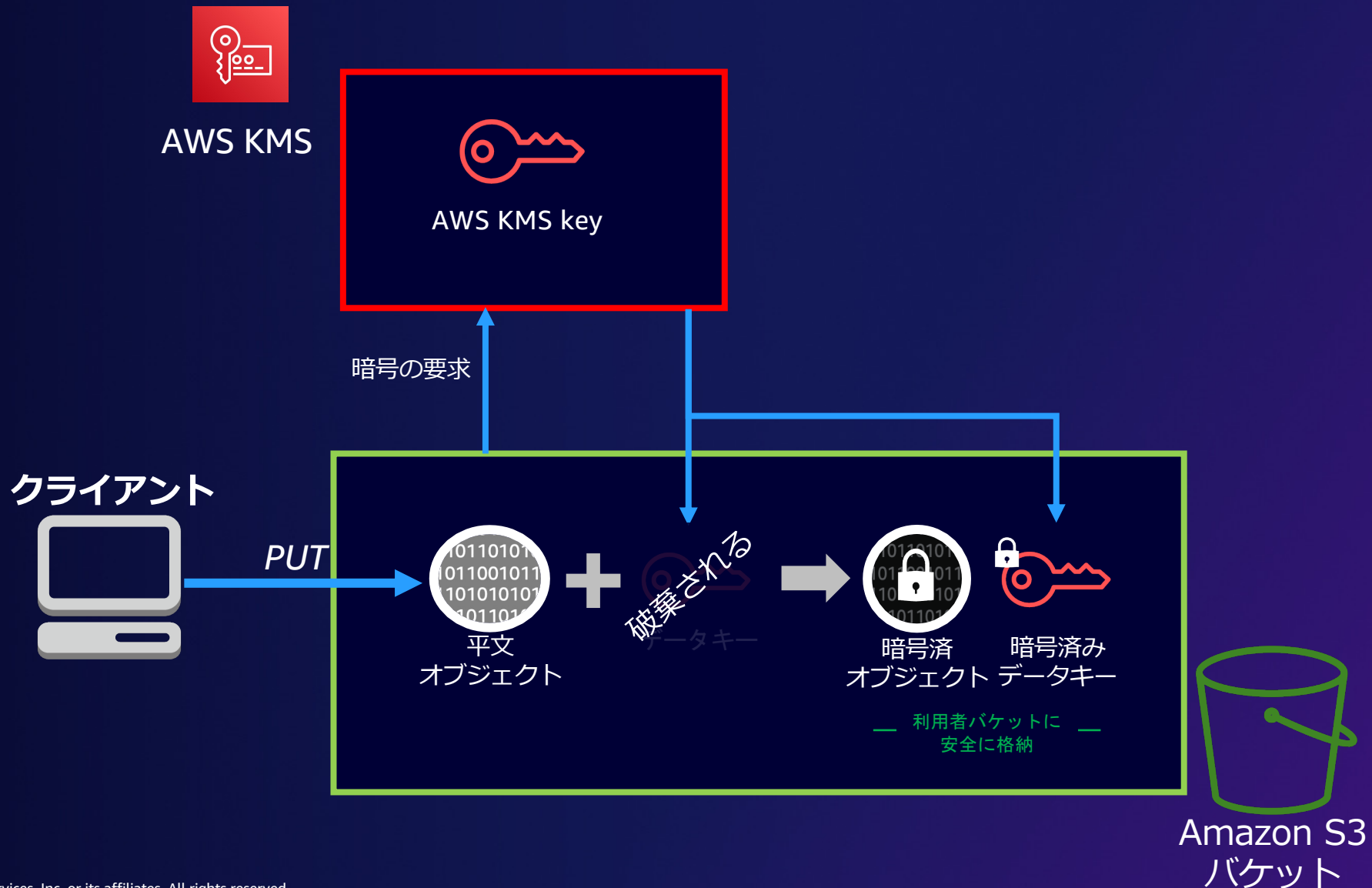
データは常に保護されている

# Amazon S3 での暗号化

# Amazon S3 における暗号化



# AWS KMS を利用して暗号化する流れ





# Amazon S3 デフォルト暗号化の活用

デフォルトの暗号化を編集 [情報](#)

**デフォルトの暗号化**  
このバケットに保存された新しいオブジェクトを自動的に暗号化します。 [詳細](#)

**サーバー側の暗号化**

☐ 無効にする

☒ 有効にする

**暗号化キータイプ**  
お客様が用意した暗号化キー (SSE-C) を使用してオブジェクトをアップロードするには、AWS CLI、AWS SDK、または Amazon S3 REST API を使用します。

☐ Amazon S3-マネージドキー (SSE-S3)  
Amazon S3 が作成、管理、使用する暗号化キー。 [詳細](#)

☒ AWS Key Management Service キー (SSE-KMS)  
AWS Key Management Service (AWS KMS) で保護されている暗号化キー。 [詳細](#)

**AWS KMS キー**

☐ AWS 管理キー (aws/s3)  
arn:aws:kms:ap-northeast-1:911111111111:alias/aws/s3

☒ AWS KMS キーから選択する

☐ AWS KMS キー ARN を入力する

**AWS KMS キー**

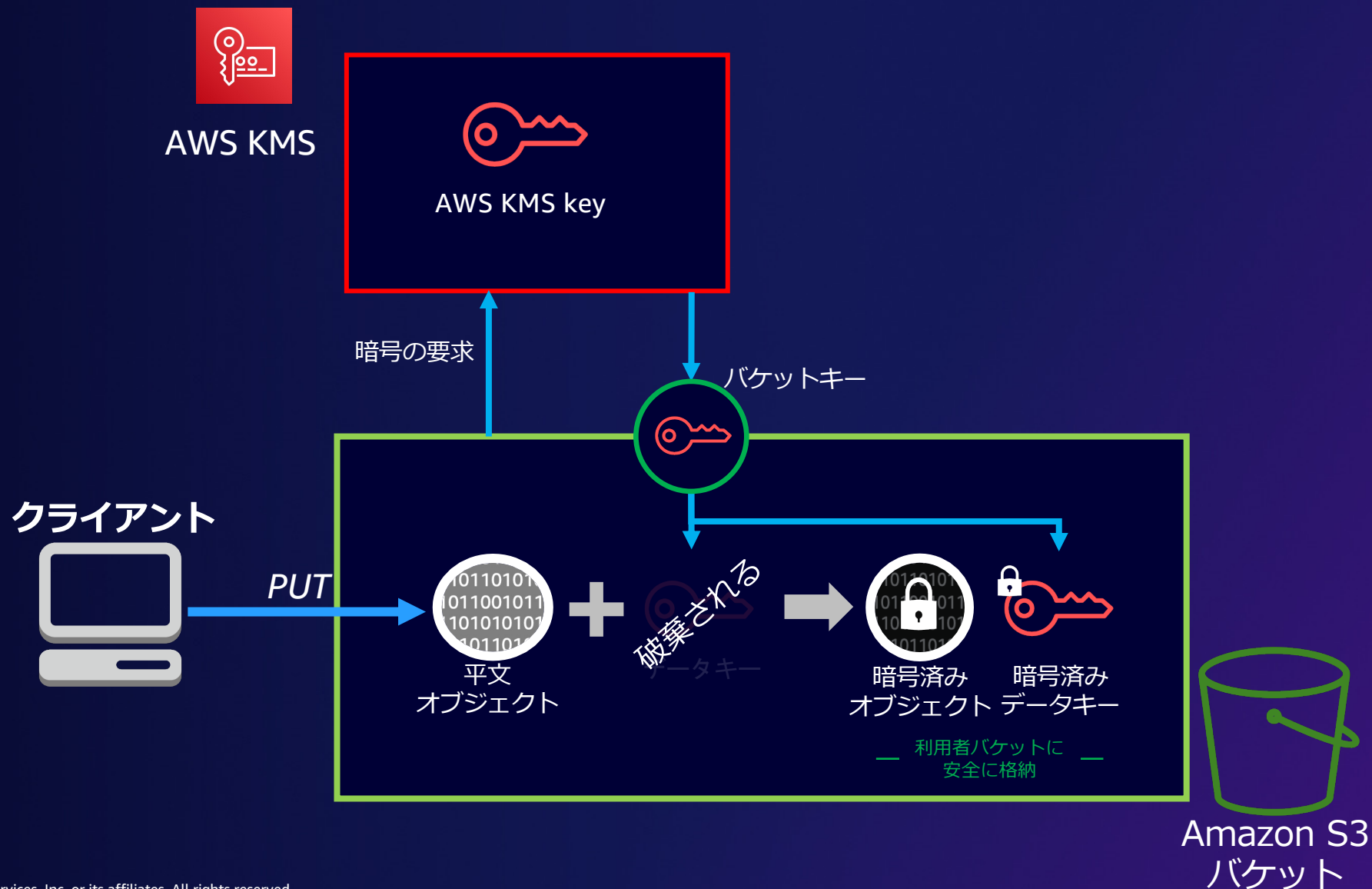
**バケットキー**  
このバケットの新しいオブジェクトに対する AWS KMS への呼び出しを減らすことで暗号化コストを削減します。オブジェクトの Bucket Key 設定を指定するには、AWS CLI、AWS SDK、または Amazon S3 Rest API を使用します。 [詳細はこちら](#)

☐ 無効にする

☒ 有効にする

- バケットレベルのコンフィグレーション
- すべての新しいオブジェクトの暗号化を自動化
- キー管理方法をここで指定
- 利用者が管理するか、Amazon S3に委ねるか
- AWS KMSの場合、キーを指定する
- 自己所有キーとするか、サービス管理か
- バケットキー活用で、AWS KMS のコスト削減
- AWS KMS へのリクエストを最大99%削減

# Amazon S3 バケットキー活用 (最初のリクエスト)



# Amazon S3 バケットキー活用(2回目以降リクエスト)



AWS KMS

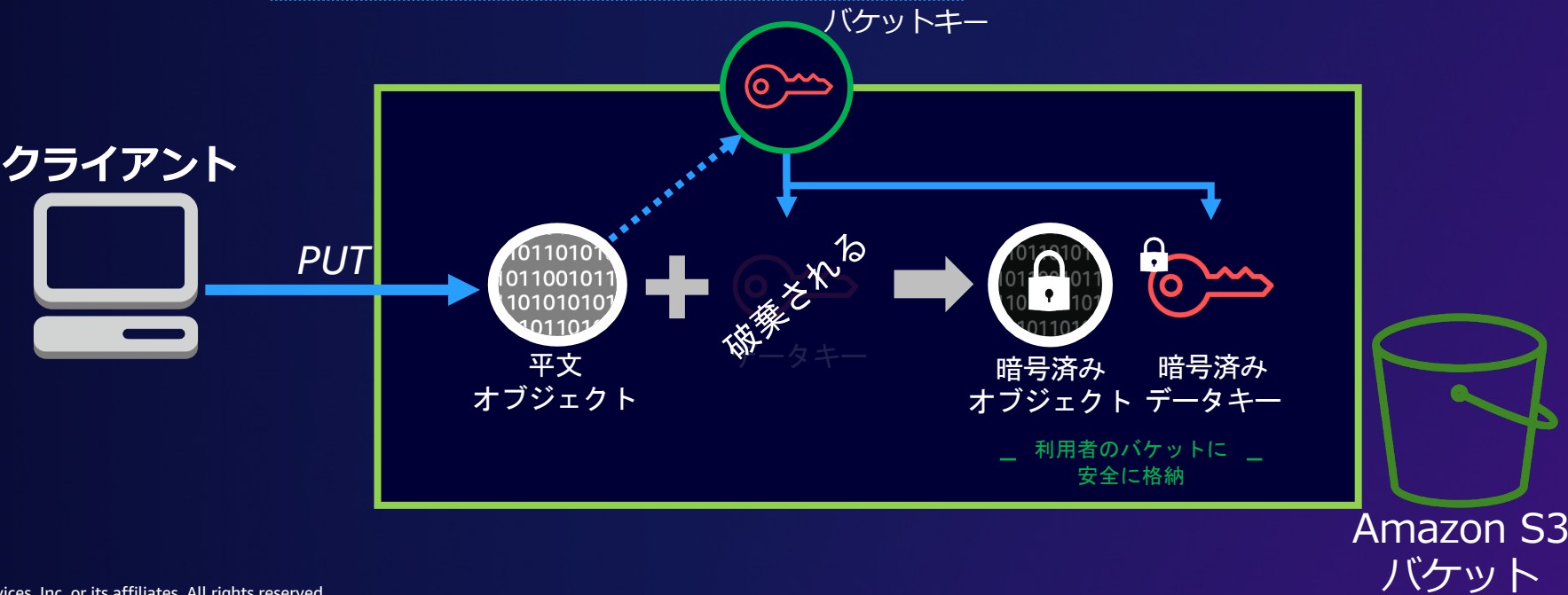



AWS KMS key



バケットキー  
はAmazon S3  
で一定時間利  
用される

AWS KMSにはリクエストされない

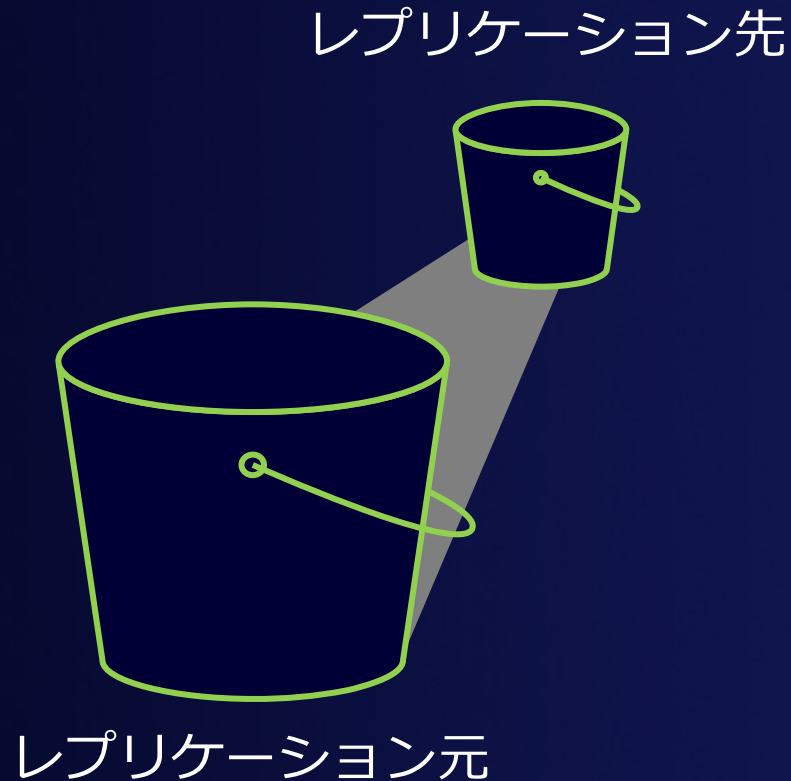


増加するデータがもたらす課題とAmazon S3 の歴史  
Amazon S3 を支える基盤  
Amazon S3 での暗号化  
 Amazon S3 でのレプリケーション  
Amazon S3 での誤削除対策  
Amazon S3 へのデータアクセスパス  
Amazon S3 ストレージクラス  
Amazon S3 を軸にしたデータアーキテクチャ  
まとめ

データは常に保護されている

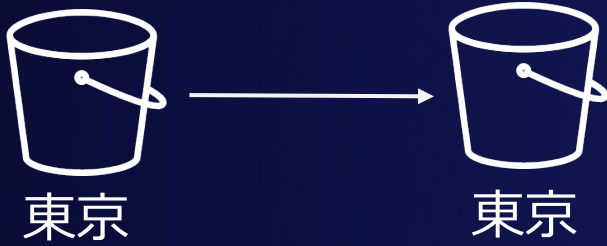
# Amazon S3 でのレプリケーション

# Amazon S3 レプリケーションとは



- S3レプリケーションは、バケット間でオブジェクトを複製する、フルマネージドな低価格の機能です
- S3レプリケーションは、レプリケーションを行う場所と方法を設定できる豊富な機能により、柔軟性を提供します
- レプリケーションルールが設定されると、S3 レプリケーションは自動的にソースバケットのオブジェクトとメタデータをレプリケートします

# 柔軟なレプリケーションシナリオ



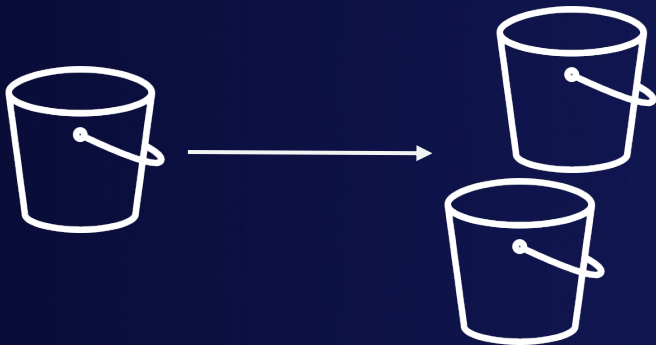
## Same-Region

同一リージョン内でのレプリケーション



## Cross-Region

異なるリージョン間でのレプリケーション



## Multiple destinations

複数のレプリケーション先



## Cross-account

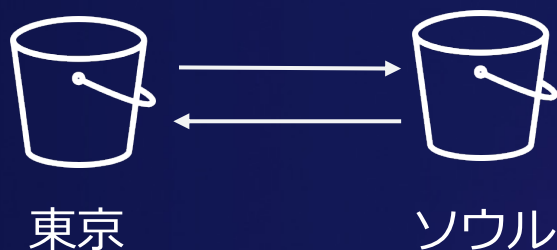
異なるAWS アカウント間でのレプリケーション

# 柔軟なカスタマイズ

オブジェクトを異なるストレージクラスにレプリケーションできる



Replica modification syncを使用することで、双方向レプリケーションが可能



メトリックスと通知を有効にしてレプリケーションの進捗状況の確認ができる






# S3 Batch Replicationの活用

既存のオブジェクトをレプリケーションする



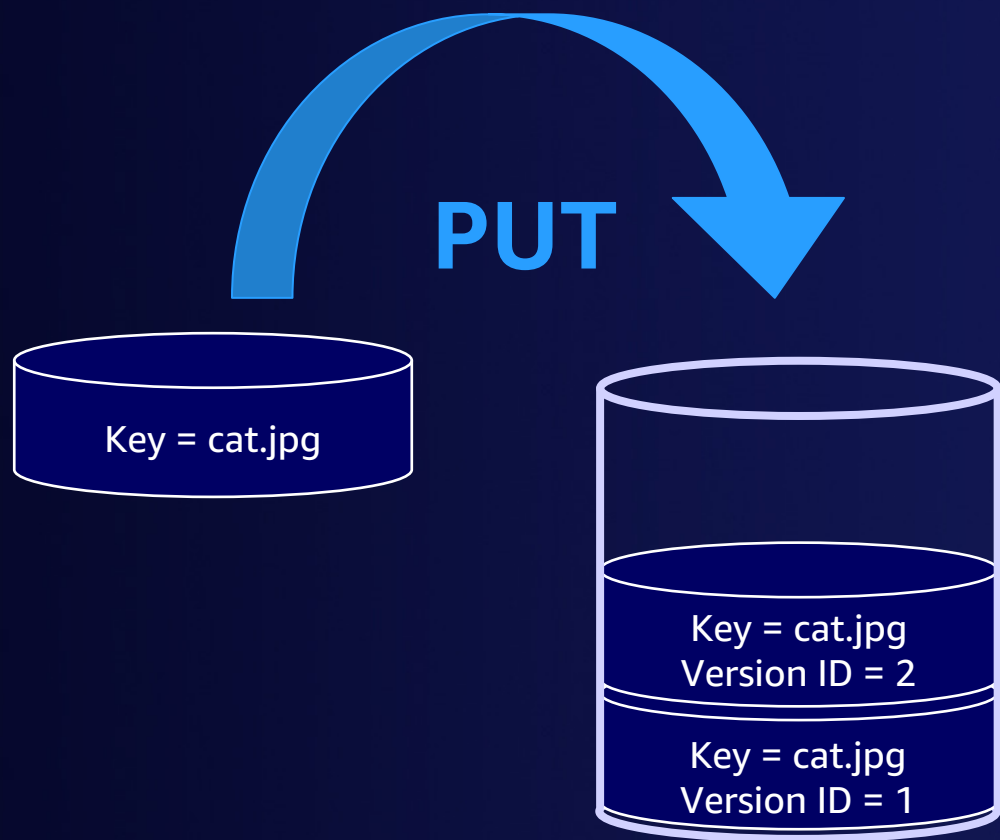


増加するデータがもたらす課題とAmazon S3 の歴史  
Amazon S3 を支える基盤  
Amazon S3 での暗号化  
Amazon S3 でのレプリケーション  
 Amazon S3 での誤削除対策  
Amazon S3 へのデータアクセスパス  
Amazon S3 ストレージクラス  
Amazon S3 を軸にしたデータアーキテクチャ  
まとめ

データは常に保護されている

# Amazon S3 での誤削除対策

# Amazon S3 バケットへのバージョニング設定



バージョニング = 有効

アップロードごとに新バージョンを作成  
以前のバージョンは上書きされずに保持される

ユーザーの意図しない削除を防ぐことができる。  
つまり、バージョンIDのない削除リクエストでは、オブジェクトへのアクセスは削除されるが、データは裏で保持される

ライフサイクルによる旧バージョンの管理が可能。  
オブジェクトが現在のバージョンでなくなった後、指定された日数で移行または失効させることができる

# Amazon S3 Object Lock (改ざん防止機能)

## S3 オブジェクトを変更できないようにする

いわゆる WORM 機能(Write Once Read Many)  
オブジェクトまたはバケットに対して適用



## リテンション指定

改ざんを防止するロックの期間の定義  
リーガルホールド

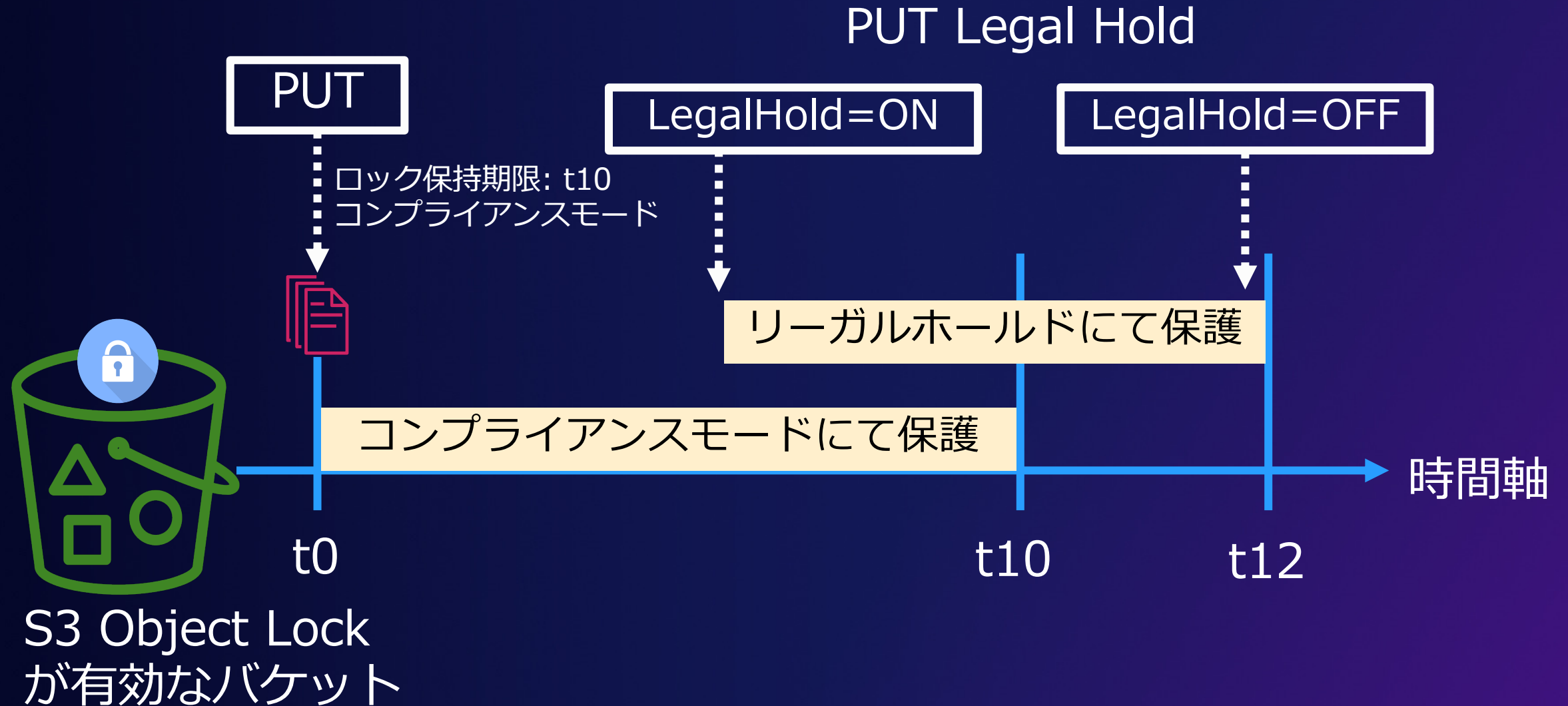
## データ保護とコンプライアンス

第三者機関によるSEC 17a-4アセスメント済み  
意図しない削除からの保護

## 二つのモード

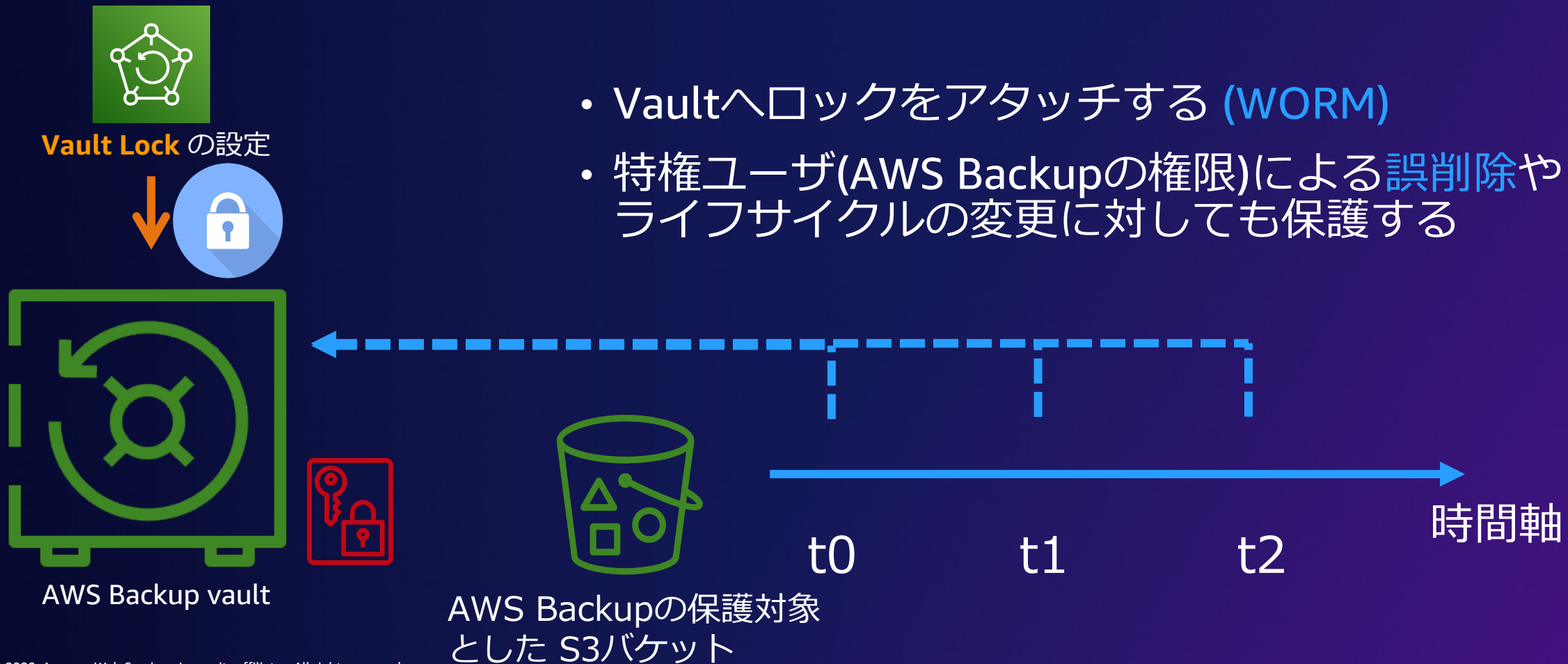
コンプライアンスモードとガバナンスモード


# Amazon S3 Object Lock 利用例



# AWS Backup Vault Lock の活用

**Vault Lock** はバックアップを **変更不可能** にし、不注意やランサムウェアのような悪意のある行為からバックアップを保護します



増加するデータがもたらす課題とAmazon S3 の歴史  
Amazon S3 を支える基盤  
Amazon S3 での暗号化  
Amazon S3 でのレプリケーション  
Amazon S3 での誤削除対策  
 Amazon S3 へのデータアクセスパス  
Amazon S3 ストレージクラス  
Amazon S3 を軸にしたデータアーキテクチャ  
まとめ

データは必要なときに必要なところから利用できる

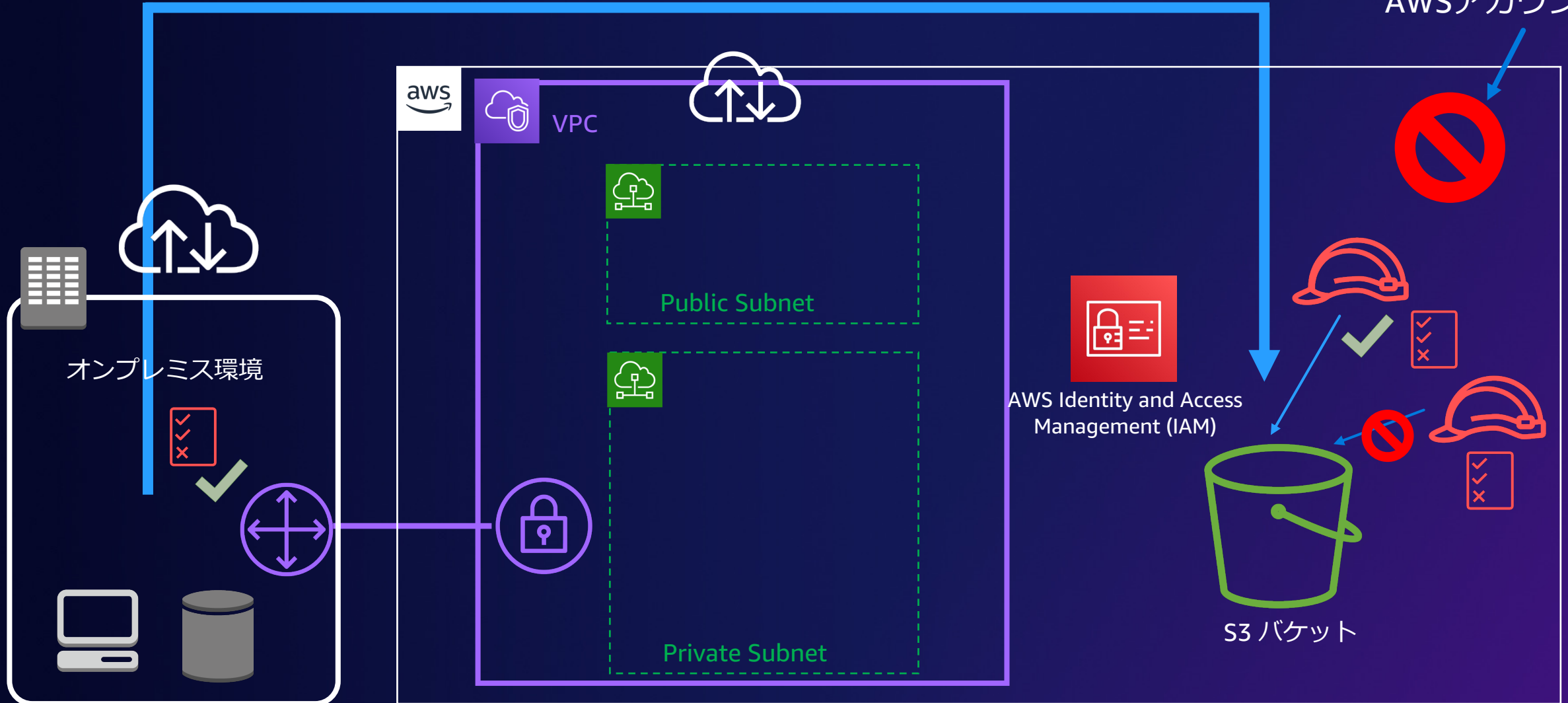
# Amazon S3 へのデータアクセスパス

# Amazon S3へのアクセスパス

Amazon S3バケットは  
デフォルトでプライベート  
なリソースである



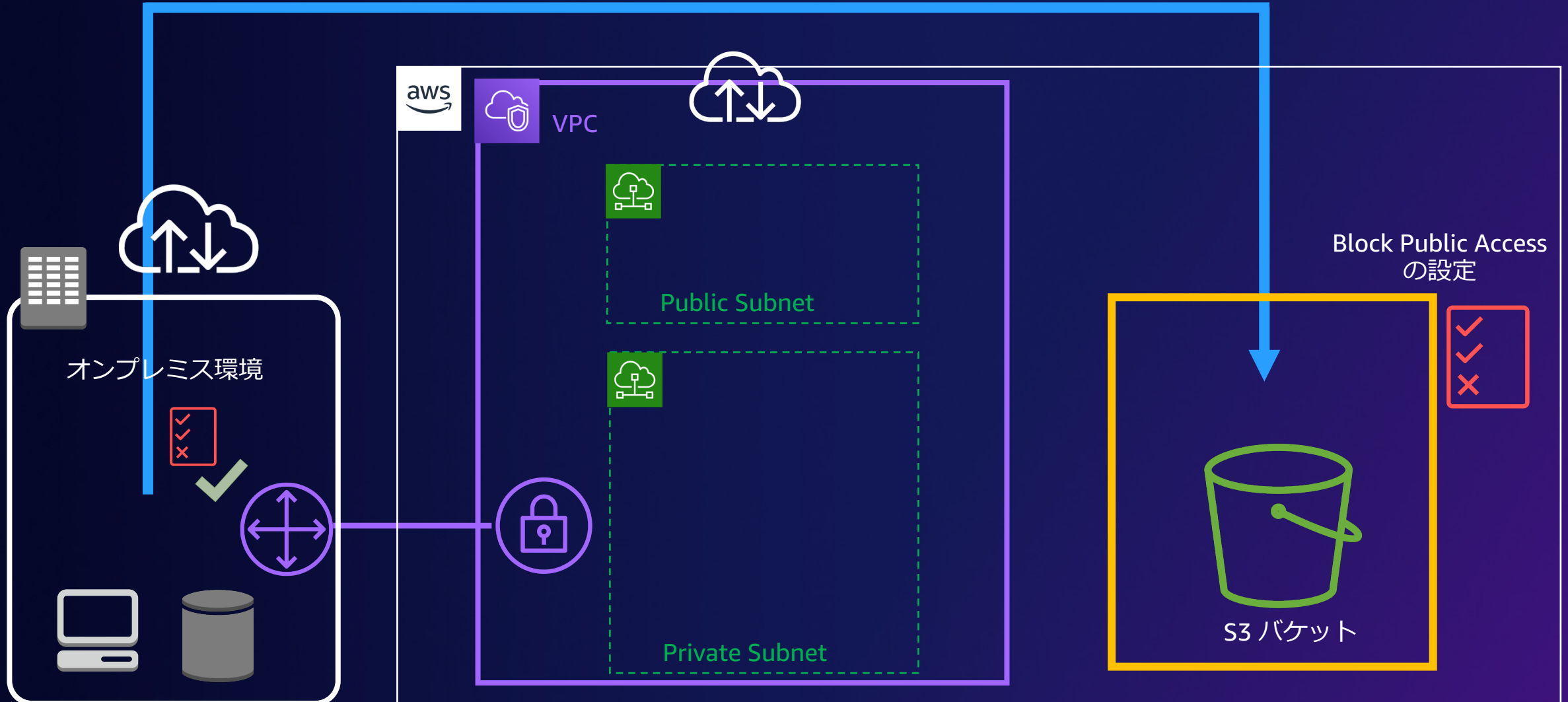
AWSアカウント外





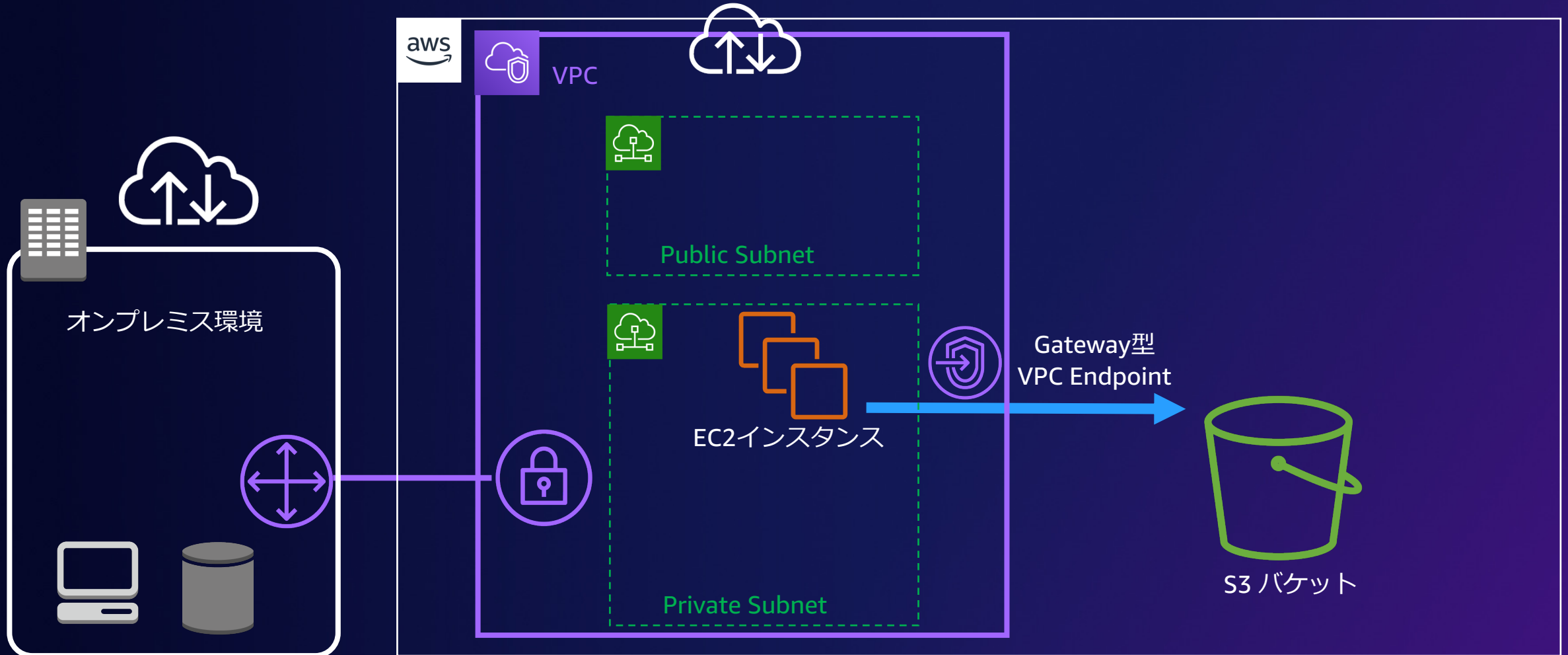
# Block Public Accessの活用

パブリックアクセス、つまり「誰でもアクセスしうる条件」設定できない、また、そのような設定しても無効化する事が可能

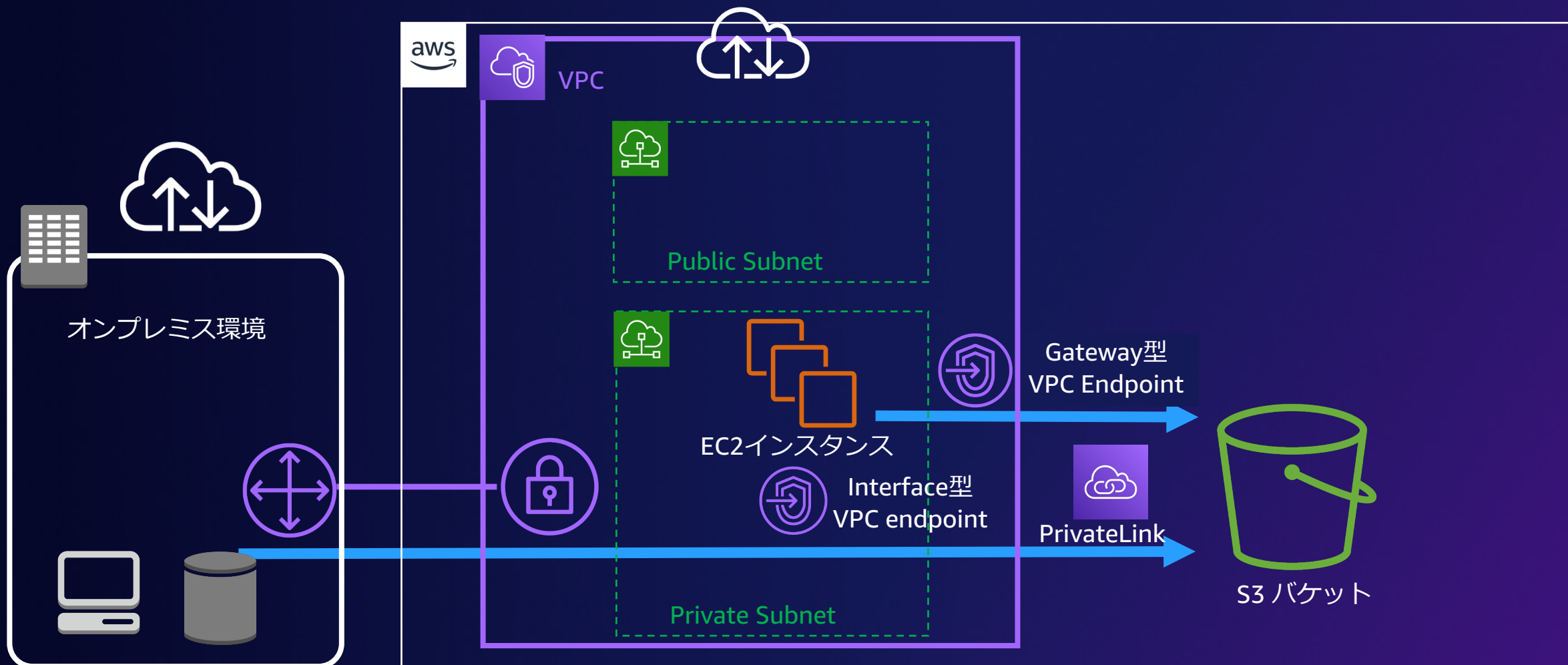




# Gateway型 VPC Endpoint の活用

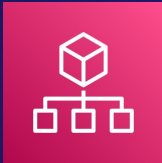


# Amazon S3 へのアクセスパス(PrivateLink)

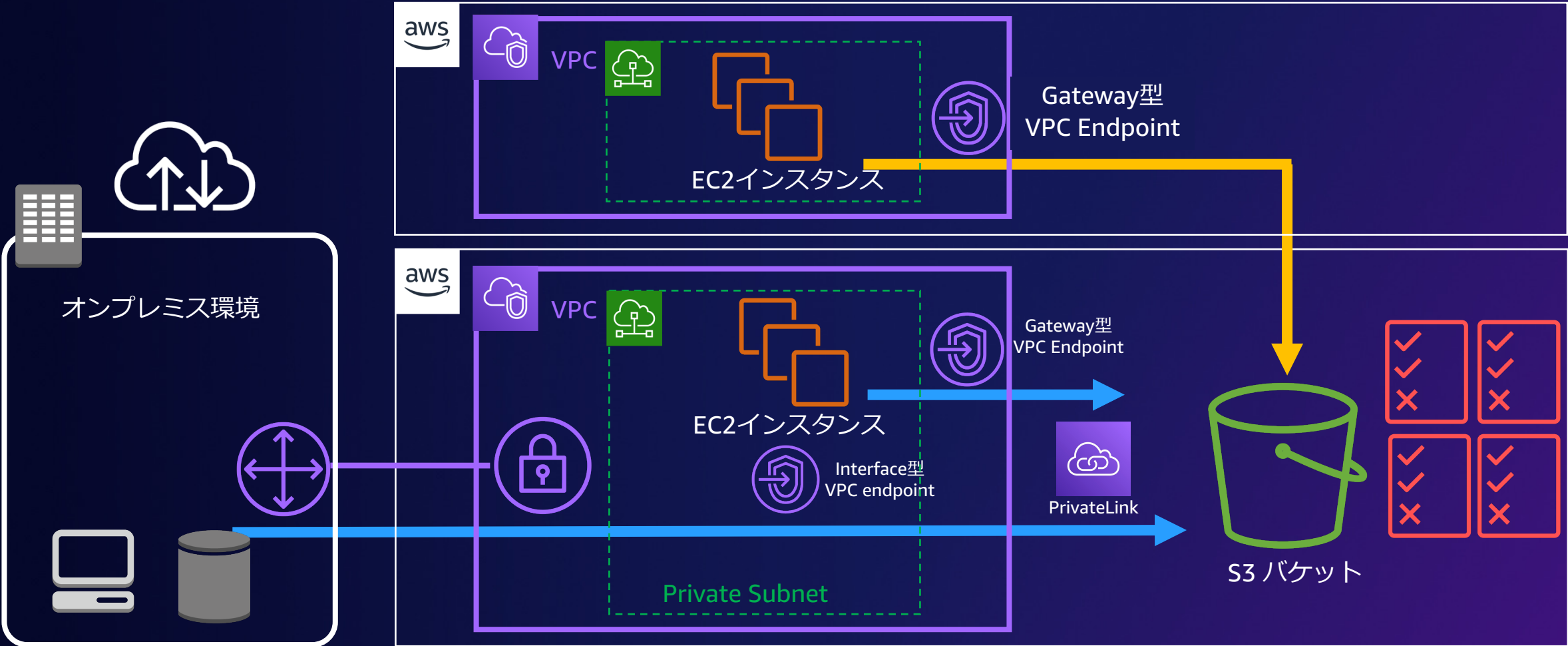


# 複数AWSアカウントからの活用

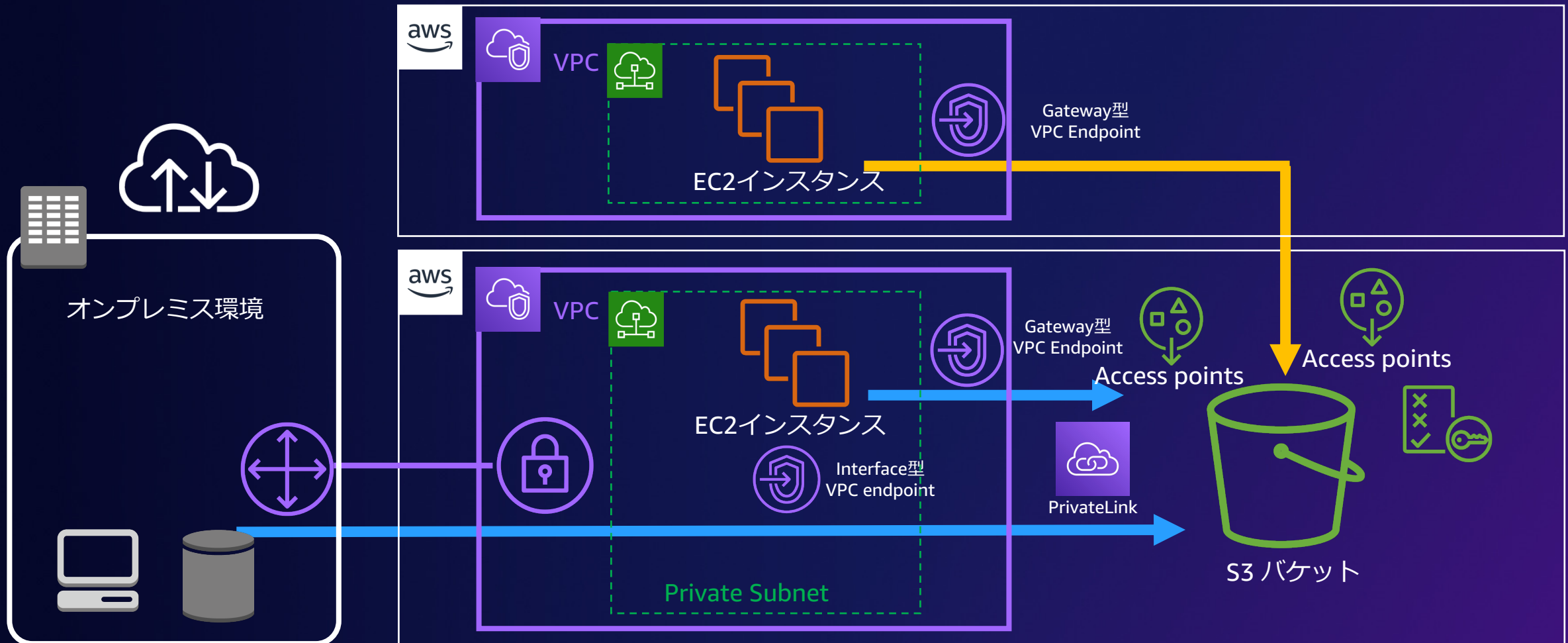
AWS Organizationsを活用することで、複数AWS アカウントへのアクセス設定を簡素化することが可能



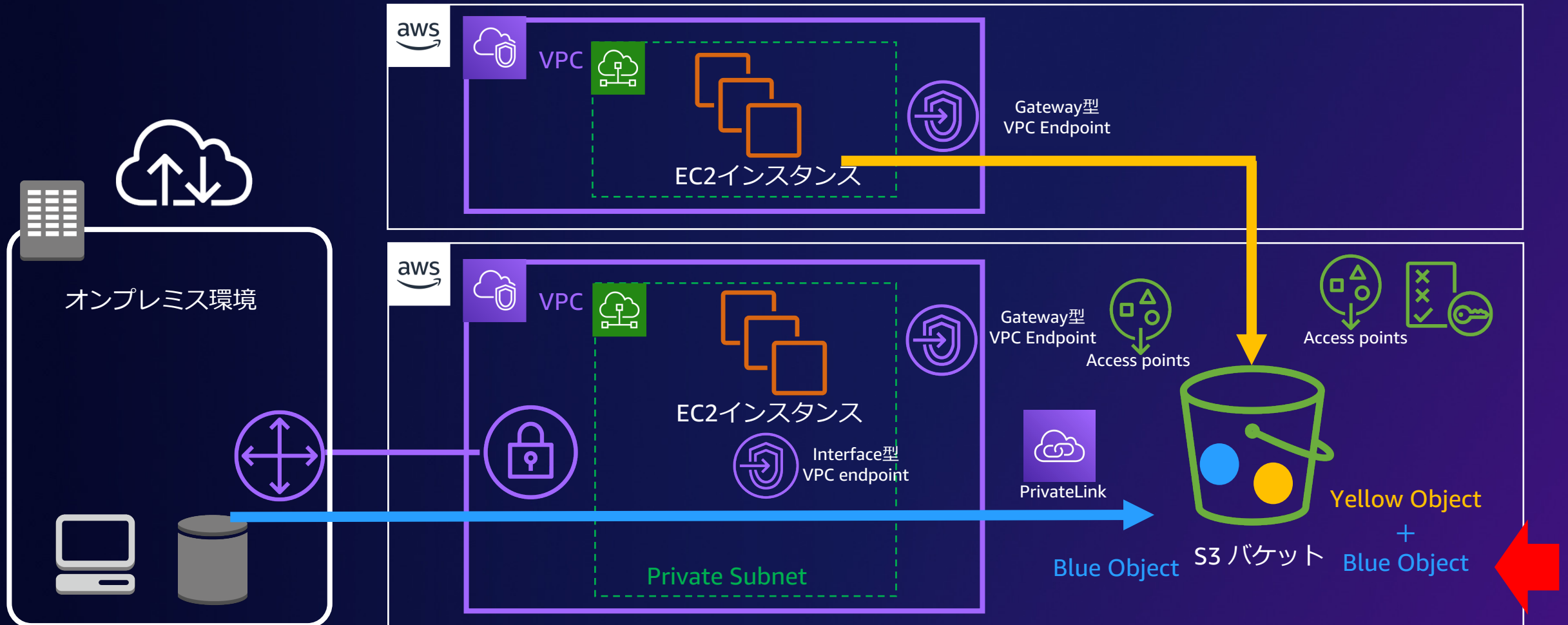
AWS Organizations



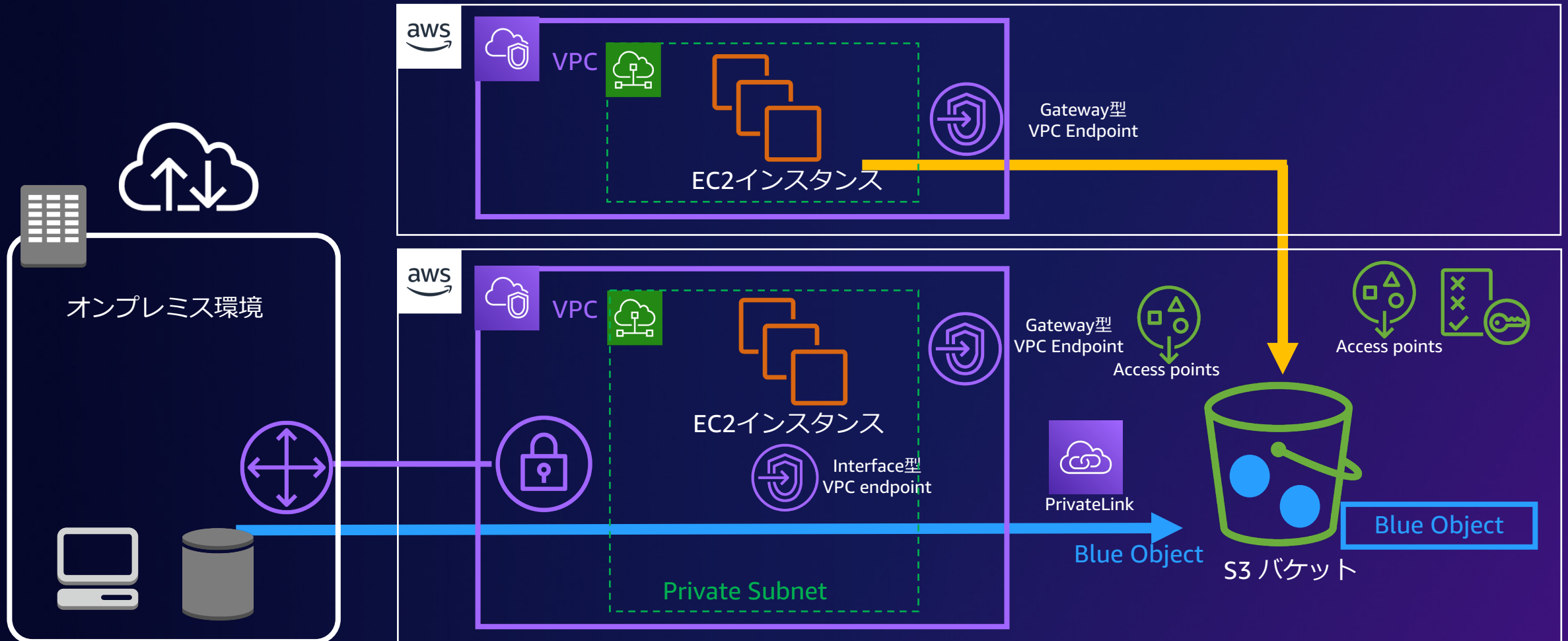
# 複数AWSアカウントからの活用(Access Points)



# ACLでのObject Ownership設定



# ACL無効設定(Object Ownership enforced)



# ACL無効設定



## オブジェクト所有者 情報

他の AWS アカウントからこのバケットに書き込まれ、アクセスコントロールリスト (ACL) を使用して付与されたオブジェクトの所有権を管理します。オブジェクトの所有権は、オブジェクトへのアクセスを指定できるユーザーを決定します。

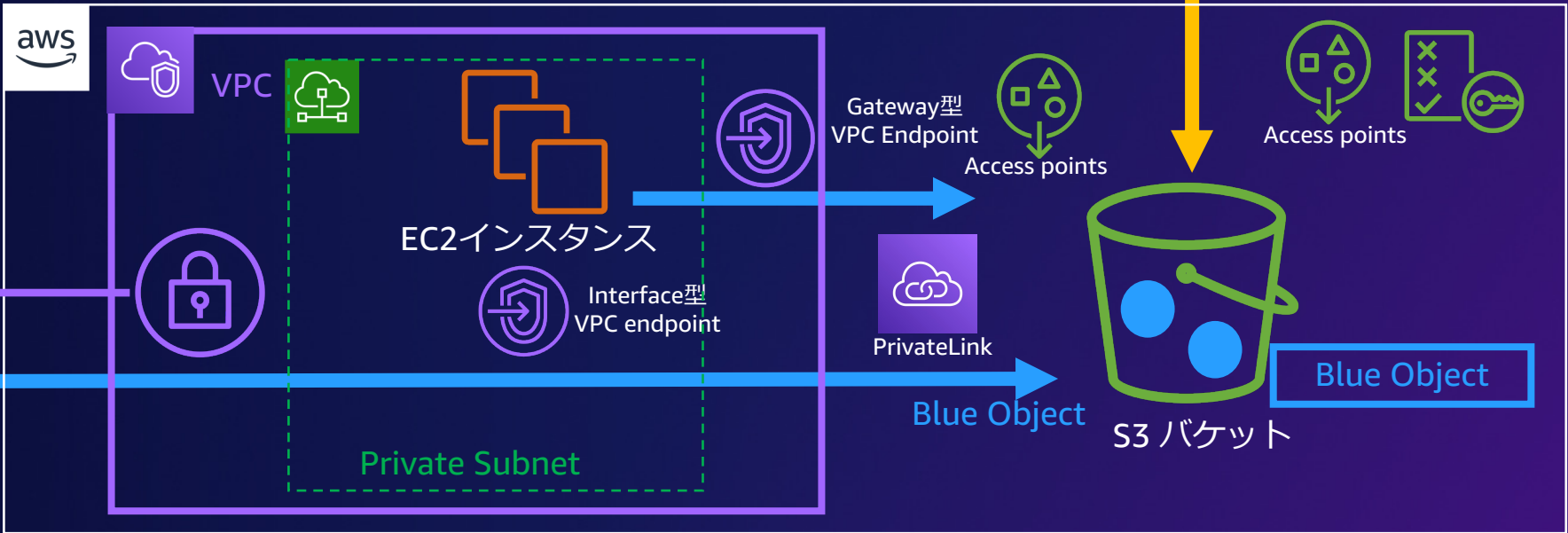
### ☒ ACL 無効 (推奨)

このバケット内のすべてのオブジェクトは、このアカウントによって所有されます。このバケットとそのオブジェクトへのアクセスは、ポリシーのみを使用して指定されます。


### ☐ ACL 有効

他の AWS アカウントがこのバケット内のオブジェクトの所有者となることができます。このバケットとそのオブジェクトへのアクセスは、ACL を使用して指定できます。

オブジェクト所有者  
バケット所有者の強制





増加するデータがもたらす課題とAmazon S3 の歴史  
Amazon S3 を支える基盤  
Amazon S3 での暗号化  
Amazon S3 でのレプリケーション  
Amazon S3 での誤削除対策  
Amazon S3 へのデータアクセスパス  
 Amazon S3 ストレージクラス  
Amazon S3 を軸にしたデータアーキテクチャ  
まとめ

データはコスト最適化されている

# Amazon S3 ストレージクラス



# S3の費用を決める要素



細かい多数のファイルを活用するユースケースは要注意  
使用頻度が低いファイルは束ねる、など。

# ストレージクラス

S3 Intelligent-Tiering

S3 Standard (S3 標準)

S3 Standard-IA (S3 標準-IA)

S3 Glacier Instant Retrieval

S3 Glacier Flexible Retrieval

S3 Glacier Deep Archive

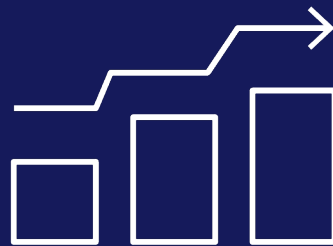
S3 One Zone-IA (S3 1ゾーン-IA)

AWSグローバル インフラストラクチャ	3つ以上のアベイラビリティゾーン (AZ)						1つのAZ
	アクセスパターンが 変化するデータ	頻繁にアクセスさ れるアクティブ データ	アクセス頻度が低 いデータ	めったにアクセス されないデータ	アーカイブデータ	長期保存のアーカ イブデータ	再生可能でアクセ ス頻度が低いデー タ
想定されるデータタイプ							
設計上の耐久性	99.999999999%	99.999999999%	99.999999999%	99.999999999%	99.999999999%	99.999999999%	99.999999999%
設計上の可用性	99.9%	99.99%	99.9%	99.9%	99.99%	99.99%	99.5%
可用性(SLA)	99%	99.9%	99%	99%	99.9%	99.9%	99%
レイテンシー	ミリ秒単位の アクセス	ミリ秒単位の アクセス	ミリ秒単位の アクセス	ミリ秒単位の アクセス	分から時間単位の 復元 (数分～12時間)	時間単位の 復元 (12～48時間)	ミリ秒単位の アクセス
取り出し料金	なし	なし	GBあたり	GBあたり	GBあたり	GBあたり	GBあたり
最低保存期間	—	—	30日	90日	90日	180日	30日
最小オブジェクトサイズ	—	—	128KB	128KB	40KB	40KB	128KB
ストレージ価格 *	0.025 ～ 0.002 USD/GB 月	0.025 ～ 0.023 USD/GB 月	0.0138 USD/GB 月	0.005 USD/GB 月	0.0045 USD/GB 月	0.002 USD/GB 月	0.011 USD/GB 月

# 取り出し速度や利便性とコストのバランスの課題



費用対効果の高い  
データ活用



データを価値ある  
ものに変える必要  
がある



データレイク、バック  
アップ、アーカイブなど  
のワークロードに対応

# ストレージクラスの活用例1

予測可能なアクセスパターンか？

Yes

ライフサイクルの活用

オブジェクト作成または  
更新日に基づく移動



PUT

GET

作成して  
0日



S3 Standard

90日経過



S3 Glacier  
Instant Retrieval

365日経過



S3 Glacier  
Flexible Retrieval

S3 Glacier Instant Retrieval はミリ秒単位でアクセスできる

一度、S3 Standard にRestoreで戻してからアクセスする

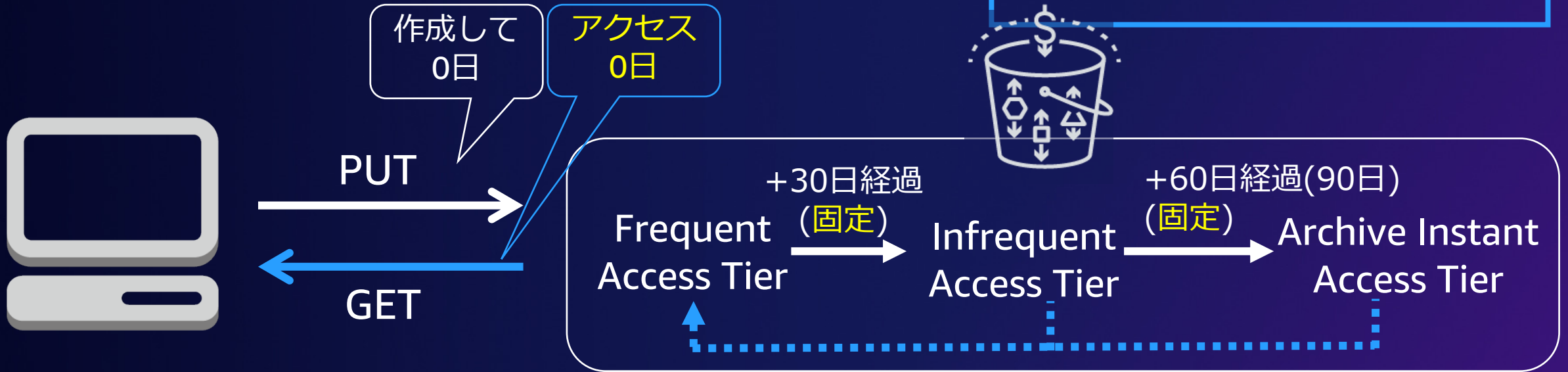


# ストレージクラスの活用例2

予測可能なアクセスパターンか？

No

S3 Intelligent-Tiering  
クラスの活用  
アクセス日に基づく移動



アクセス階層であればどのTierでも即時アクセスが可能

# ストレージクラスの活用例3

予測可能なアクセスパターンか？

No

S3 Intelligent-Tiering  
クラスの活用

アーカイブ層も併用(Opt-in)



PUT

作成して  
0日

アクセス  
0日

GET

+30日経過

Frequent  
Access Tier

(固定)

Infrequent  
Access Tier

+60日経過(90日)

(固定)

90~730日経過(設定)

Archive Access  
Tier

180~730日経過(設定)

Deep Archive  
Access Tier

アーカイブ階層からは一度  
Frequent Access Tierに戻る  
(Restoreオペレーション)

増加するデータがもたらす課題とAmazon S3 の歴史  
Amazon S3 を支える基盤  
Amazon S3 での暗号化  
Amazon S3 でのレプリケーション  
Amazon S3 での誤削除対策  
Amazon S3 へのデータアクセスパス  
Amazon S3 ストレージクラス  
➡ Amazon S3 を軸にしたデータアーキテクチャ  
まとめ

データはうまく活用することができる

# Amazon S3 を軸にしたデータアーキテクチャ



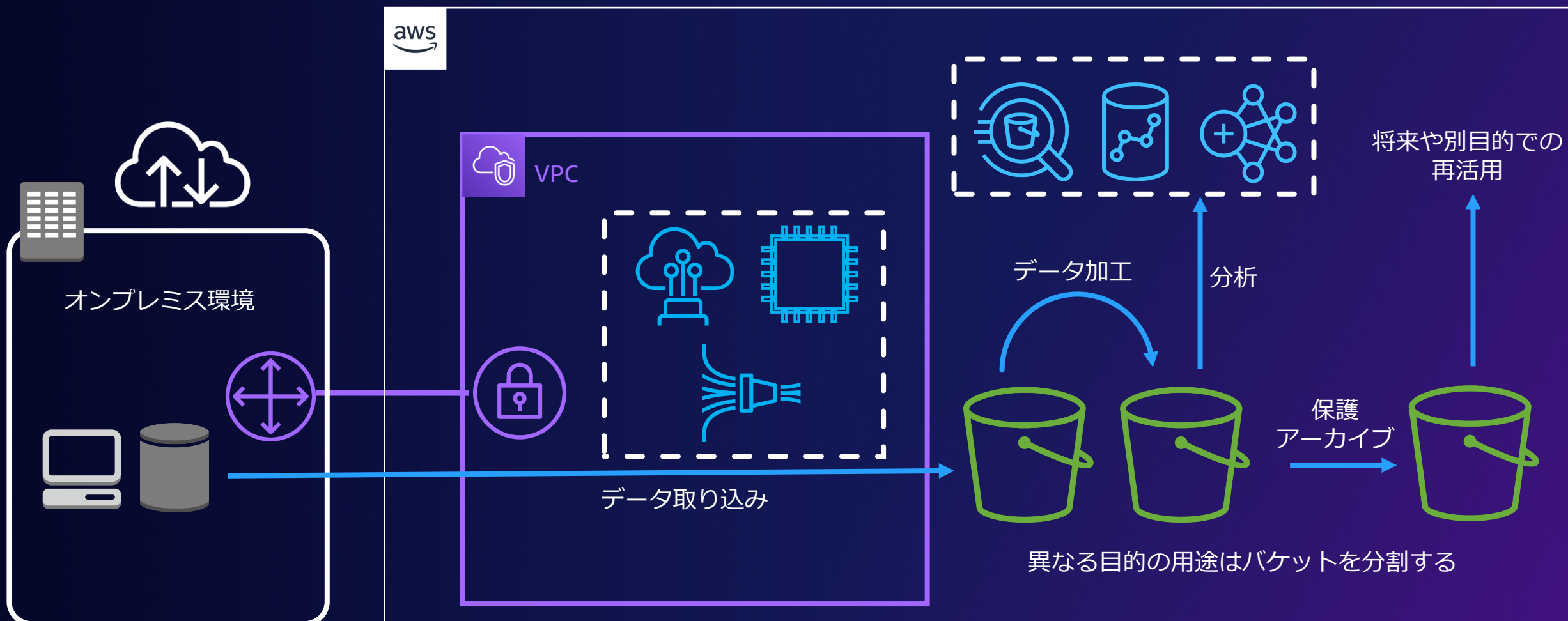
# Amazon S3 によるデータ保護、データセキュリティの強化、コストの最適化

## 得られる利点

- ✓ 暗号化/レプリケーション/誤削除対策に有効な機能でのデータ保護
- ✓ セキュアなデータアクセス環境の提供
- ✓ 様々なワークロードに対応するコスト効率の高いストレージクラス



# Amazon S3を活用するデータレイク



増加するデータがもたらす課題とAmazon S3 の歴史  
Amazon S3 を支える基盤  
Amazon S3 での暗号化  
Amazon S3 でのレプリケーション  
Amazon S3 での誤削除対策  
Amazon S3 へのデータアクセスパス  
Amazon S3 ストレージクラス  
Amazon S3 を軸にしたデータアーキテクチャ  
→ まとめ

# まとめ

# AWS ストレージ上でデータアーキテクチャを構築

## データは常に保護されている

データはセキュアに扱われ、バックアップを取得し、災害・障害対策のためにレプリケーションが可能です。誤った削除からデータの復旧ができるようにします。

## データは必要なときに必要なところから利用できる

データはどこからでも必要なときに利用でき、業界をリードする性能でワークロードを支えます。

## データはコスト最適化されている

低コストなアーカイブストレージの活用や、柔軟なストレージクラス、自動階層化により、コスト削減が可能です。

## データはうまく活用することができる

豊富な分析サービスや機械学習サービスと連携可能なデータレイクを構築することができます。

# Thank you!

