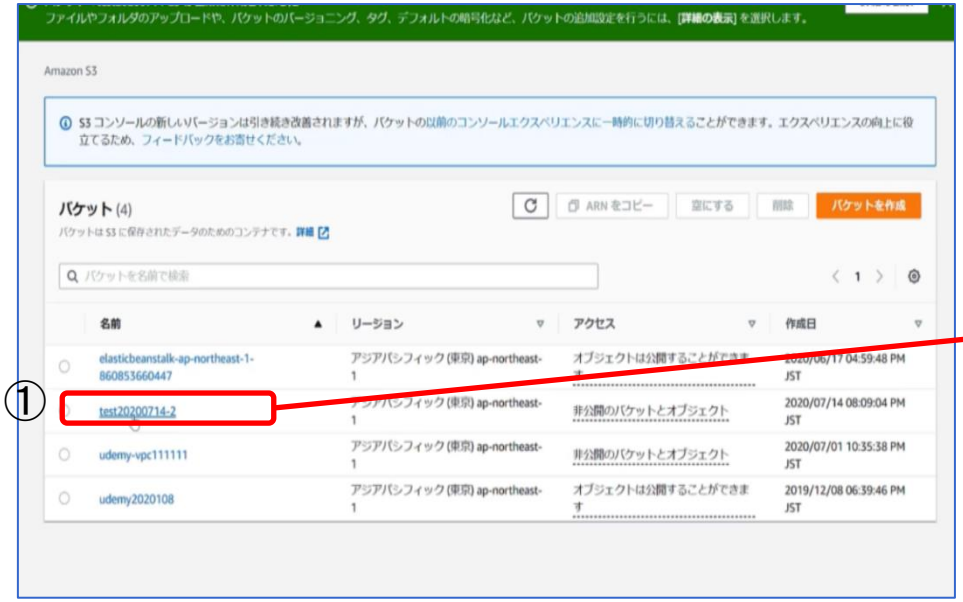
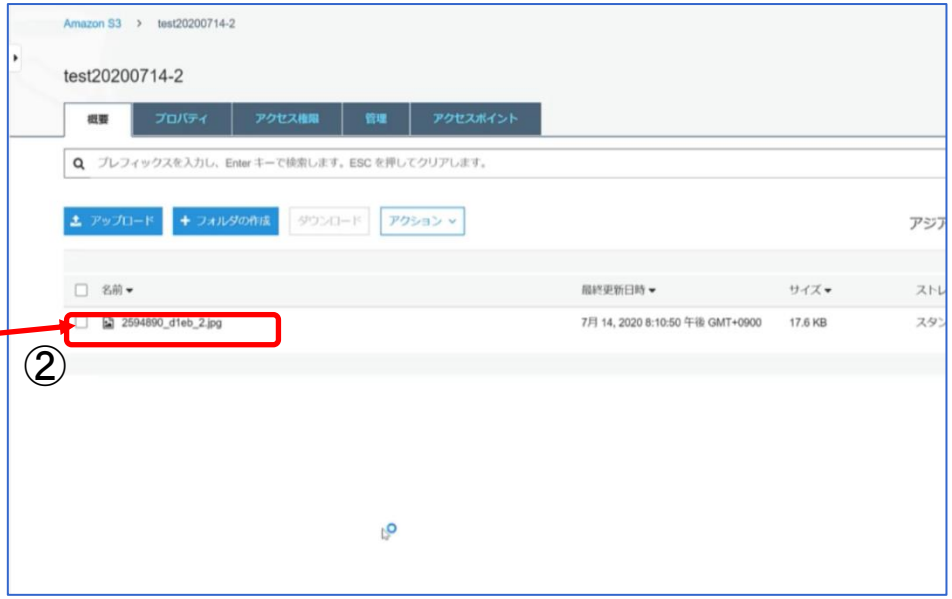
# S3の概要

# S3とは何か？

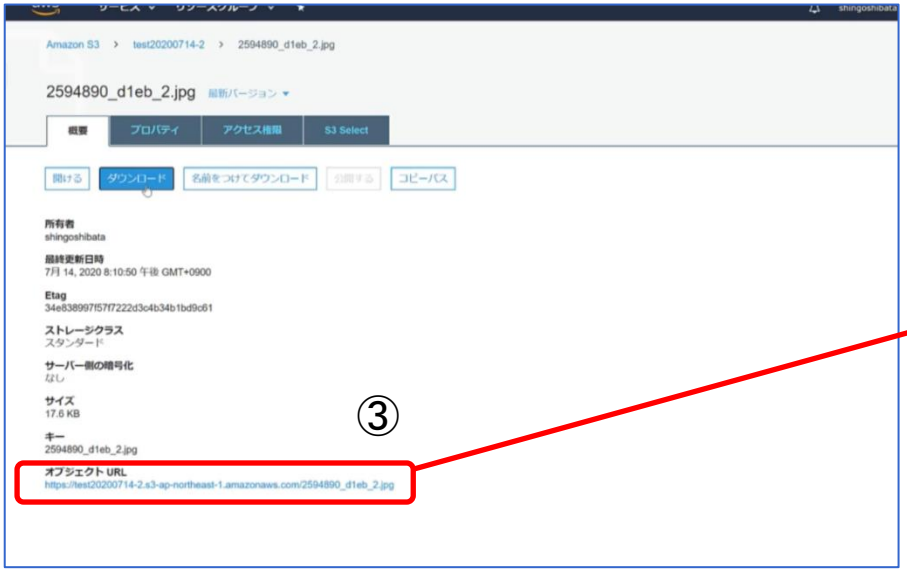
S3は耐久性と可用性が非常に高くデータの中長期保存に最適なストレージ



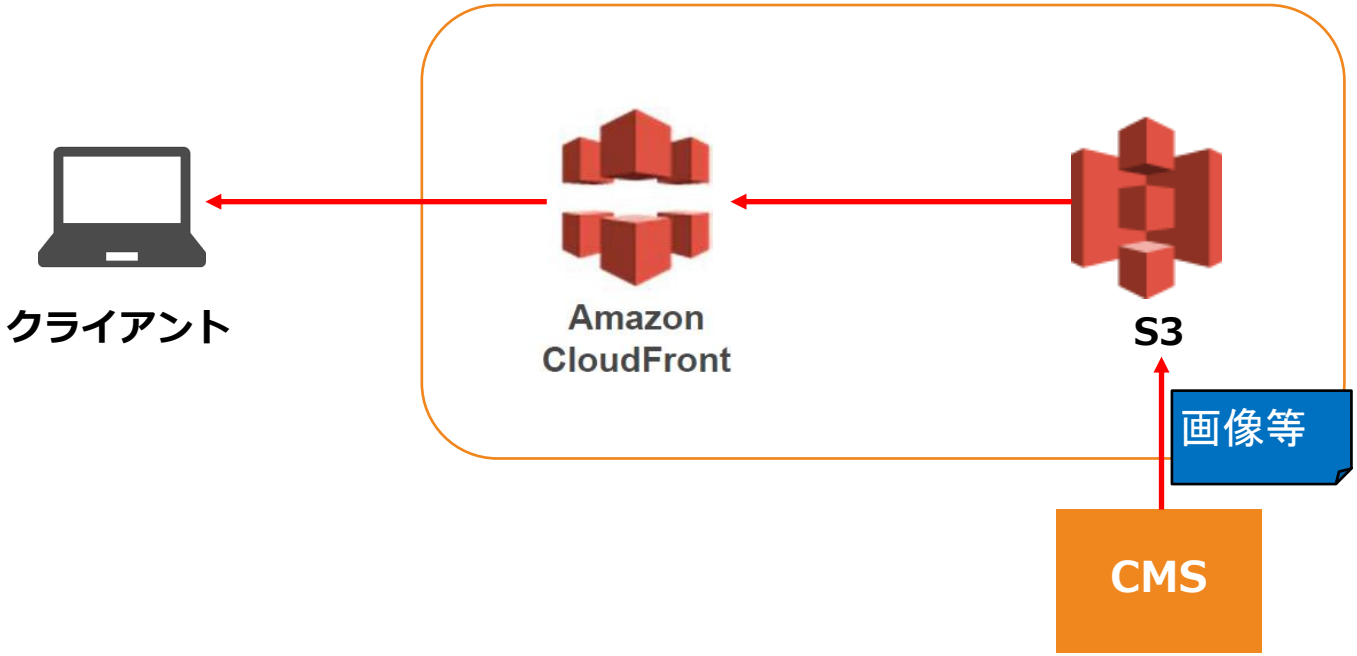
# S3とは何か？

S3は耐久性と可用性が非常に高くデータの中長期保存に最適なストレージ




# S3のユースケース

コンテンツ配信用の画像データなどをS3に保存して、
CloudFrontを利用して配信する。



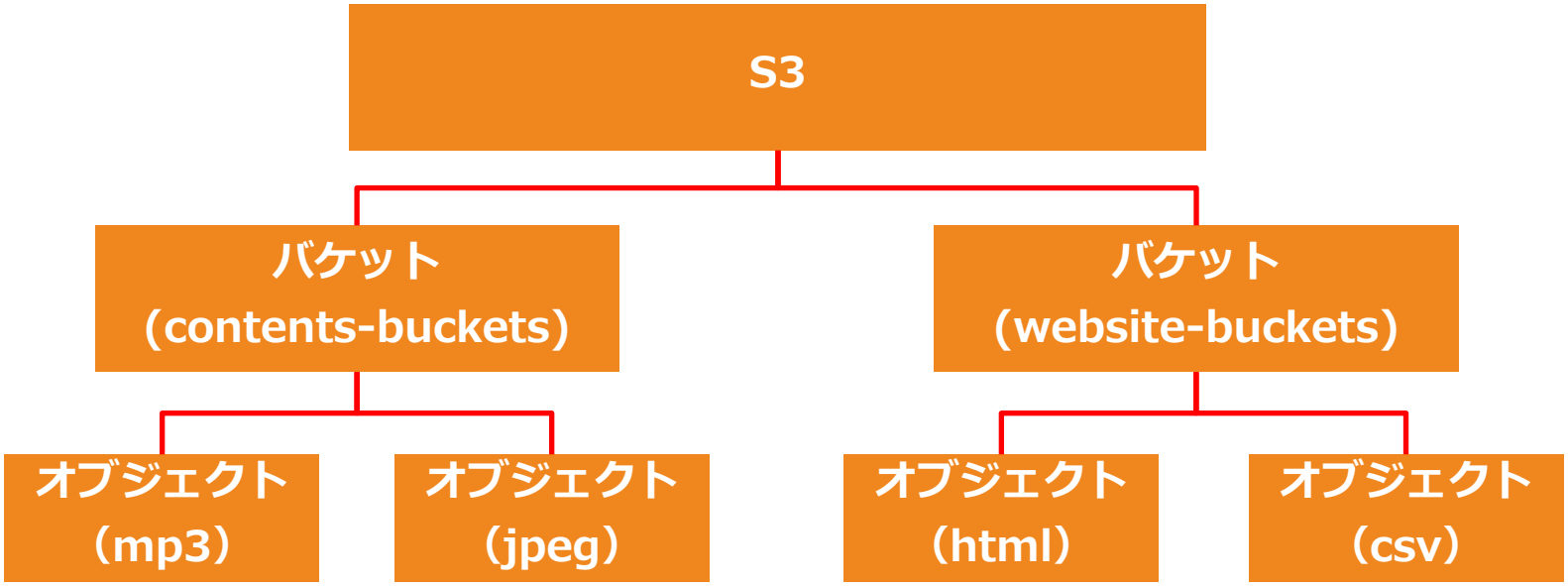
# S3ストレージの特徴

## AWSは３つの形式のストレージサービスを提供

| | |
|---|---|
| **ブロックストレージ** | ✓ EC2にアタッチして活用するディスクサービス<br>✓ ブロック形式でデータを保存<br>✓ 高速・広帯域幅<br>✓ 例：EBS、インスタンスストア |
| **オブジェクトストレージ** | ✓ 安価かつ高い耐久性をもつオンラインストレージ<br>✓ オブジェクト形式でデータを保存<br>✓ デフォルトで複数AZに冗長化されている。<br>✓ 例：S3、Glacier |
| **ファイルストレージ** | ✓ 複数のEC2インスタンスから同時にアタッチ可能な共有ストレージサービス<br>✓ ファイル形式でデータを保存<br>✓ 例：EFS |

# S3ストレージの特徴

S3はバケット単位で保存スペースを区分し、オブジェクトでデータを格納する



# バケット

## ユーザーが利用する１つのストレージ単位をバケットとして作成する。

- ✓ S3を利用する際に１つのバケットを作成して、そこにオブジェクト（ファイル）を格納する。
- ✓ バケットはリージョンに設定する。AZやVPCの範囲外。
- ✓ バケットにはグローバルに一意の名前を設定することが必要。つまり、全世界のＡＷＳユーザーで異なる名前を設定する。
- ✓ 命名規則を守る必要がある。

# バケットの命名規則

## S3バケット名は以下の命名規則に合致させる必要がある。

- ✓ バケット名は 3 (最少)～63 (最大) 文字の長さが必要
- ✓ バケット名は、小文字、数字、ドット (.)、およびハイフン (-) のみで構成できます。
- ✓ バケット名は、文字または数字で開始および終了する。
- ✓ IP アドレスの形式 (192.168.5.4 など)は利用できない。
- ✓ プレフィックスは xn—から始められない。
- ✓ バケット名のサフィックスは -s3alias で終わってはいけない。
- ✓ バケット名は、パーティション内のすべての AWS リージョン のすべての AWS アカウント にわたって一意である必要がある。
- ✓ バケットが削除されるまで、バケット名を同じパーティション内の別の AWS アカウント で使用できない。
- ✓ Amazon S3 Transfer Acceleration で使用されるバケットの名前にドット (.) を付けられない。

# オブジェクト

## バケットに保存されるデータの単位をオブジェクトと呼ぶ。オブジェクトは以下の要素で構成されている

■**Key**
オブジェクトの名前であり、バケット内のオブジェクトを一意に識別

■**Value**
データそのものであり、バイト値で構成される

■**バージョンID**
バージョン管理に用いるID

■**メタデータ**
オブジェクトに付随する属性の情報

■**サブリソース**
バケット構成情報を保存および管理するためのサポートを提供

例：アクセスコントロールリスト（ACL）

# プレフィックスの利用

プレフィックスはオブジェクトキー名の先頭にある文字列。オブジェクトはプレフィックスを利用して整理して保存する。

バケット (udemy-test) にsampleというjpgファイルを整理した例

udemy-test/**photos/2006/January/sample.jpg**
udemy-test/**photos/2006/February/sample2.jpg**
udemy-test/**photos/2006/February/sample3.jpg**
udemy-test/**photos/2006/February/sample4.jpg**

# オブジェクトキー

## オブジェクトキーはプレフィックス＋オブジェクト名（ファイル名）で構成される

バケット (udemy-test) に以下の２つのオブジェクトキーを持つオブジェクトがある場合

■udemy-test/Development/Projects.xls

■udemy-test/ s3-dg.pdf

# オブジェクトキー

## オブジェクトキーはプレフィックス＋オブジェクト名（ファイル名）で構成される

バケット (udemy-test) に以下の２つのオブジェクトキーを持つオブジェクトがある場合

■udemy-test/Development/Projects.xls

⇒**Developmentというプリフィックス＋オブジェクト名（ファイル名） ＝キー**

■udemy-test/ s3-dg.pdf

# オブジェクトキー

## オブジェクトキーはプレフィックス＋オブジェクト名（ファイル名）で構成される

バケット (udemy-test) に以下の２つのオブジェクトキーを持つオブジェクトがある場合

■udemy-test/Development/Projects.xls

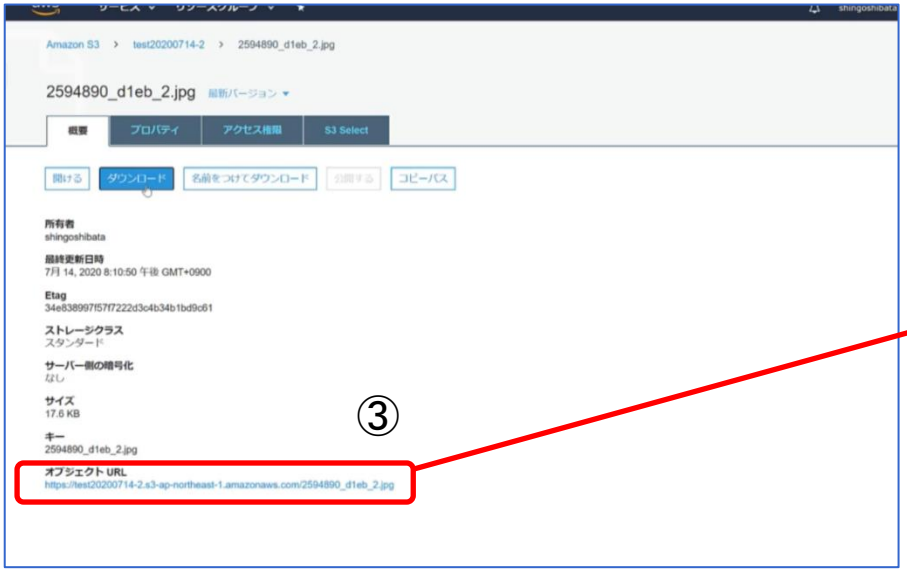
⇒**Developmentというプリフィックス＋オブジェクト名（ファイル名） ＝キー**

■udemy-test/ s3-dg.pdf

⇒**オブジェクト名（ファイル名）＝キー**

# オブジェクトURL

オブジェクトにアクセスする際はオブジェクトURLという固有のURLからインターネット経由でアクセスできる。




# S3のデータ容量制限

## S3のストレージ容量は無制限であり、 0KBから5TBまでのデータを保存可能

**S3のデータ容量制限**

**■バケット**
オブジェクトの保存場所。リージョンに設置されるため、名前はグローバルでユニークにする。データ保存容量は無制限であり、自動でストレージ容量が拡張される。

**■オブジェクト**
S3に格納されるファイル形式で、オブジェクトに対してURLが付与される。バケット内に保存可能なオブジェクト数は無制限

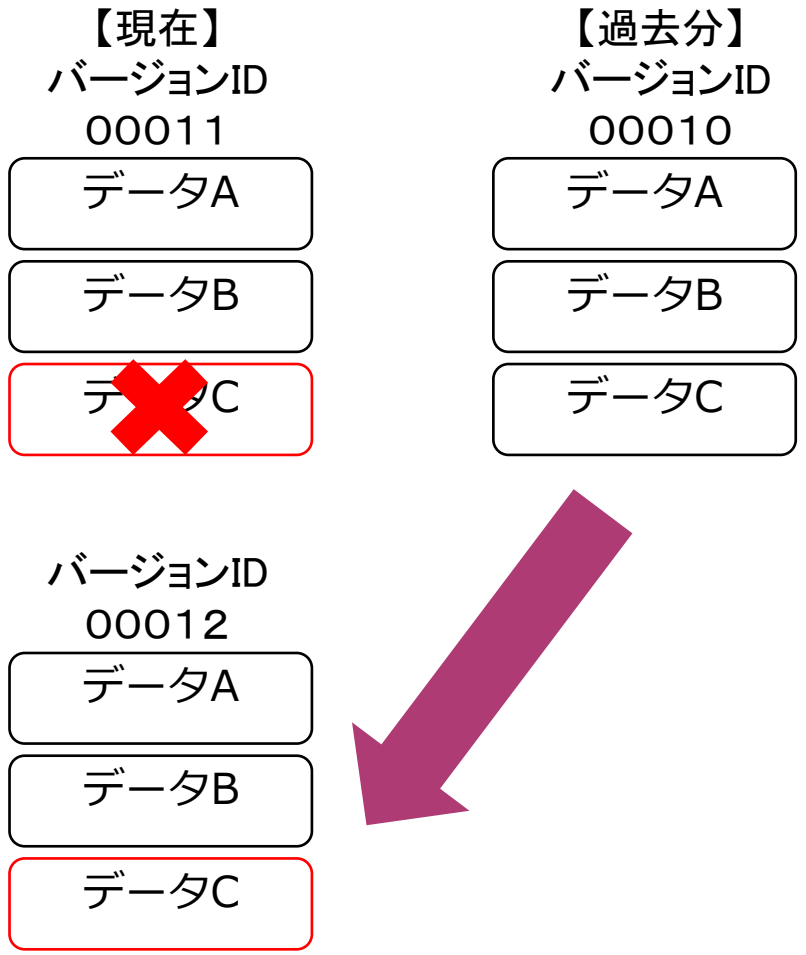
**■保存可能なオブジェクトサイズの制限**
オブジェクトあたりのデータサイズは0KBから5TBまで保存可能

# バージョン管理

## ユーザーによる誤操作でデータ削除などが発生してもバージョンから復元できる

**バージョニングの基本**

- バケット単位でオブジェクトのバージョンを管理する
- バージョンごとにオブジェクトが保管される。
- ライフサイクルルールによってバージョンが保存される期間を設定できる。
- バージョニングが有効になる前のバージョンはnullとなる。



# ストレージクラスの選択

## S3の用途に応じてストレージタイプを選択する

| タイプ | 特徴 | 性能 | 追加課金 |
|---|---|---|---|
| **STANDARD** | ✓ 複数個所にデータを複製するため耐久性が非常に高く、頻繁に利用するデータを大量に保存するのに向いている。<br>✓ データは3AZ以上で分散保存される。 | ■耐久性<br>99.999999999%<br>■可用性<br>99.99% | ■最低利用料金<br>なし<br>■データ取得料<br>なし |
| **STANDARD-IA** | ✓ IAはInfrequency Accessの略であり、低頻度アクセスデータ用のストレージ。One Zone-IAより重要なマスターデータ向け。データ取得は早い<br>✓ Standard に比べて安価だが、One Zone-IAよりは高い。 | ■耐久性<br>99.999999999%<br>■可用性<br>99.9% | ■最低利用料金<br>30日<br>■データ取得料<br>GB当たり取得料 |
| **One Zone-IA** | ✓ 低頻度アクセス用のストレージだが、マルチAZ分散されていないため可用性が低く、重要ではないデータ向け。その分Standard IAよりも値段が安い | ■耐久性<br>99.999999999%<br>■可用性<br>99.5% = 1AZ | ■最低利用料金<br>30日分<br>■データ取得料<br>GB当たり取得料 |
| **S3 Intelligent Tiering** | ✓ 高頻度と低頻度という2つのアクセス階層を利用し、アクセスがあるファイルは高頻度（標準クラス）に維持しつつ、アクセスがないファイルは低頻度（標準IAクラス）に自動で移動する。<br>✓ アクセスパターンがわからない場合に利用 | ■耐久性<br>99.999999999%<br>■可用性<br>99.9% | ■最低利用料金<br>30日<br>■データ取得料<br>なし |

# ストレージクラスの選択

## Glacierでは３つのストレージタイプから選択する。

| タイプ | 特徴 | 性能 | 追加課金 |
|---|---|---|---|
| **S3 Glacier Flexible Retrieval（通常のGlacier）** | ✓ 1 年に 1～2 回アクセスされ、非同期で取り出されるアーカイブデータ向け<br>✓ 通常のデータ検索で（3~5時間）を要する<br>✓ 迅速取り出しで(2～5分)で取り出し可能<br>✓ 一括検索で（5～12時間）で無料<br>✓ ライフサイクルマネジメントで指定<br>✓ ボールトロック機能でデータを保持 | ■耐久性<br>99.999999999%<br>■可用性<br>99.99% | ■最低利用料金<br>90日<br>■データ取得料<br>GB当たり取得料 |
| **S3 Glacier Instant Retrieval** | ✓ アクセスされることがほとんどなく、ミリ秒単位の取り出しが必要な長期間有効なデータ向け<br>✓ 医用画像やニュースメディアなど<br>✓ S3 Standard と同じパフォーマンスのミリ秒単位でのデータの取り出し | ■耐久性<br>99.999999999%<br>■可用性<br>99.9% | ■最低利用料金<br>90日<br>■データ取得料<br>GB当たり取得料 |
| **Amazon Glacier Deep Archive** | ✓ 最安のアーカイブ用ストレージ<br>✓ 7～10 年以上保持される長期間使用されるものの、めったにアクセスされないデータ向け<br>✓ 標準の取り出し速度で12 時間以内にデータを取得<br>✓ 大容量取り出しで48時間以内にデータを取得<br>✓ ライフサイクル管理で指定 | ■耐久性<br>99.999999999%<br>■可用性<br>99.99% | ■最低利用日数<br>180日<br>■データ取得料<br>GB当たり取得料 |

# S3の利用コスト

## ストレージのコストを比較するとインスタンスストアを除けば、最も値段が安いのはGlacier deep archive

| | |
|---|---|
| S3のデータ容量<br>に応じたコスト | ✓ 標準：1 GB あたり 0.025USD／月<br>✓ S3 Intelligent Tiering:標準と標準IAの組合せ<br>✓ 標準IA： 1 GB あたり 0.019USD／月<br>✓ One Zone IA： 1 GB あたり 0.0152USD／月<br>✓ Glacier：1 GB あたり 0.005USD／月<br>✓ Glacier deep archive：1 GB あたり 0.002USD／月 |
| EBSの汎用<br>ストレージのコスト | ✓ 汎用：1 GB あたり 0.12USD／月<br>✓ コールドHDD:1 GB あたり 0.03USD／月 |
| EFS<br>ストレージのコスト | ✓ 標準：1 GB あたり 0.36USD／月<br>✓ 低頻度アクセス：0.0272USD／月 |
| インスタンスストア | ✓ EC2インスタンスに含まれる。 |

https://aws.amazon.com/jp/s3/pricing/

# S3の利用コスト

## S3はデータ量とリクエストとデータ転送に対して料金が発生

| | |
|---|---|
| リージョン | ✓ リージョン：リージョン毎に価格が異なる。 |
| データ容量 | ✓ データ容量：データ量と保存期間に応じて料金がかかる。（GBあたり）<br>✓ S3 Intelligent Tiering、 IAストレージには、最低 30 日間の料金 |
| リクエストとデータ取得 | ✓ データに対するリクエストに応じて料金がかかる。（1000リクエストあたり）<br>✓ データを取得した量に応じて料金がかかる（GBあたり） |
| データ転送 | ✓ データ転送イン：無料<br>✓ インターネットへのデータ転送アウト（GBあたり）<br>✓ S3からAWS内でのデータ転送アウト（GBあたり） |

https://aws.amazon.com/jp/s3/pricing/

# S3の利用コスト

## S3はボリュームディスカウントの価格帯が設定されている

| | ストレージ料金表 |
|---|---|
| **S3 標準** - 頻繁にアクセスするデータに一般的に使用される、あらゆるタイプのデータの汎用ストレージ | |
| 最初の 50 TB/月 | 0.025USD/GB |
| 次の 450 TB/月 | 0.024USD/GB |
| 500 TB/月以上 | 0.023USD/GB |
| **S3 Intelligent - Tiering** * - アクセスパターンが不明または変化するデータの自動コスト削減 | |
| 高頻度アクセスティア、最初の 50 TB/月 | 0.025USD/GB |
| 高頻度アクセスティア、次の 450 TB/月 | 0.024USD/GB |
| 高頻度アクセスティア、500 TB/月を超える | 0.023USD/GB |
| 低頻度アクセスティア、すべてのストレージ/月 | 0.019USD/GB |
| モニタリングおよびオートメーション、すべてのストレージ/月 | オブジェクト 1,000 件あたり 0.0025USD |
| **S3 標準 - 低頻度アクセス** * - ミリ秒単位のアクセスが必要な、長期保管だがアクセス頻度の低いデータの場合 | |
| すべてのストレージ/月 | 0.019USD/GB |
| **S3 1 ゾーン - 低頻度アクセス** * - ミリ秒単位のアクセスが必要な、再作成可能なアクセス頻度の低いデータの場合 | |
| すべてのストレージ/月 | 0.0152USD/GB |
| **S3 Glacier** ** - 1 分から 12 時間の取り出しオプションを使用した長期バックアップとアーカイブの場合 | |
| すべてのストレージ/月 | 0.005USD/GB |
| **S3 Glacier Deep Archive** ** - 1 年に 1～2 回アクセスされ、12 時間以内に復元できる長期のデータアーカイブの場合 | |
| すべてのストレージ/月 | 0.002USD/GB |

https://aws.amazon.com/jp/s3/pricing/

# リクエスタ支払い

## S3バケットはデータ取得の際にも料金が発生する場合があるため、リクエスタ支払いはデータ取得したアカウントに課金する。

| | データ保存コスト | データ通信コスト |
|---|---|---|
| リクエスタ支払い無効 | ✓ バケットの所有者がデータの保存コストを支払う。 | ✓ **バケットの所有者**がデータのダウンロードコスト（通信料）を支払う。 |
| リクエスタ支払い有効 | ✓ バケットの所有者がデータの保存コストを支払う。 | ✓ **データダウンロードをリクエストしたアカウント**がデータのダウンロードコスト（通信料）を支払う。 |

https://aws.amazon.com/jp/s3/pricing/

# S3のアクセス管理

# S3のアクセス管理

## S3のアクセス管理にはユーザーベースのIAMとリソースベースのバケットポリシーとACLを主に利用する。

| 管理方式 | 特徴 |
| --- | --- |
| **IAM ユーザーポリシー** | ✓ IAMユーザーに対してAWSリソースとしてのS3へのアクセス権限を設定<br>✓ 内部のIAMユーザーやAWSリソースへの権限管理 |
| **バケットポリシー** | ✓ バケットのアクセス権をJSONで設定する。１つのバケットに対して１つだけ設定可能。<br>✓ 外部ユーザーやアプリケーションなども管理可能 |
| **ACL** | ✓ バケット／オブジェクト単位でのアクセス権限をXMLで設定することができる<br>✓ オブジェクトに個別に設定可能 |
| **アクセスポイント** | ✓ S3バケットにアクセスポリシーを設定する。<br>✓ バケットのアクセス権をJSONで設定する。１つのバケットに対して複数設定可能<br>✓ 外部ユーザーやアプリケーションなども管理可能 |

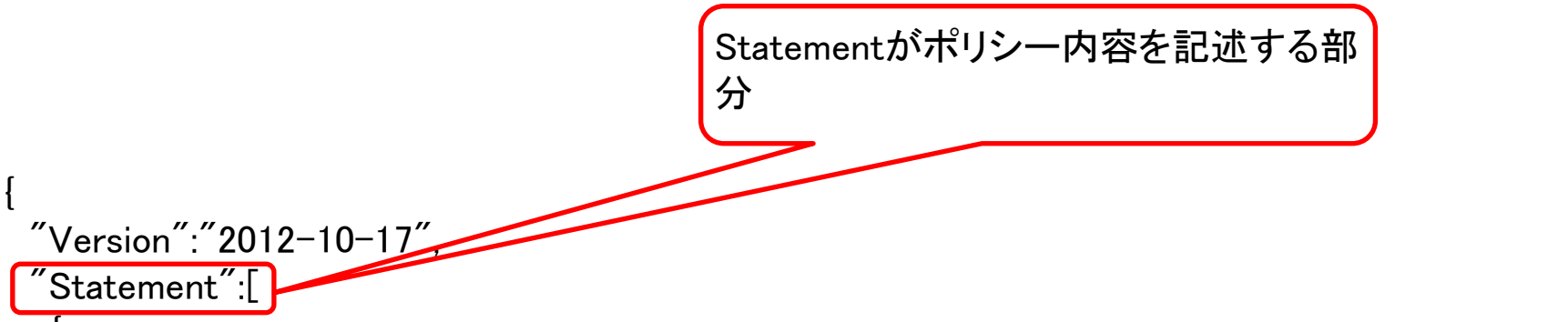
# S3バケットポリシー

ポリシーのバージョン。
必ず先頭に記載する。

```
{
  "Version":"2012-10-17",
  "Statement":[
    {
      "Sid":"AddCannedAcl",
      "Effect":"Allow",
      "Principal": {"AWS": ["arn:aws:iam::111122223333:root","arn:aws:iam::444455556666:root"]},
      "Action":["s3:PutObject","s3:PutObjectAcl"],
      "Resource":"arn:aws:s3:::awsexamplebucket1/*",
      "Condition":{"StringEquals":{"s3:x-amz-acl":["public-read"]}}
    }
  ]
}
```

Reference: https://docs.aws.amazon.com/ja_jp/AmazonS3/latest/dev/example-bucket-policies.html

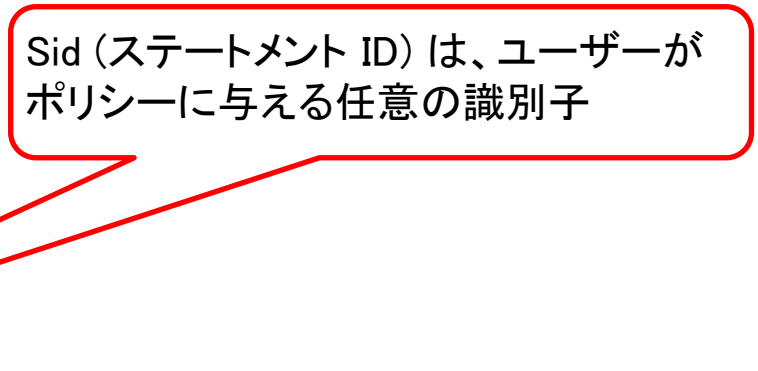
# S3バケットポリシー




```
{
  "Version":"2012-10-17",
  "Statement":[
    {
      "Sid":"AddCannedAcl",
      "Effect":"Allow",
      "Principal": {"AWS": ["arn:aws:iam::111122223333:root","arn:aws:iam::444455556666:root"]},
      "Action":["s3:PutObject","s3:PutObjectAcl"],
      "Resource":"arn:aws:s3:::awsexamplebucket1/*",
      "Condition":{"StringEquals":{"s3:x-amz-acl":["public-read"]}}
    }
  ]
}
```

Reference: https://docs.aws.amazon.com/ja_jp/AmazonS3/latest/dev/example-bucket-policies.html

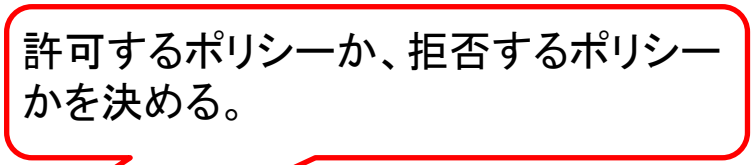
# S3バケットポリシー




```
{
  "Version":"2012-10-17",
  "Statement":[
    {
      "Sid":"AddCannedAcl",
      "Effect":"Allow",
      "Principal": {"AWS": ["arn:aws:iam::111122223333:root","arn:aws:iam::444455556666:root"]},
      "Action":["s3:PutObject","s3:PutObjectAcl"],
      "Resource":"arn:aws:s3:::awsexamplebucket1/*",
      "Condition":{"StringEquals":{"s3:x-amz-acl":["public-read"]}}
    }
  ]
}
```

Reference: https://docs.aws.amazon.com/ja_jp/AmazonS3/latest/dev/example-bucket-policies.html

# S3バケットポリシー




```
{
  "Version":"2012-10-17",
  "Statement":[
    {
      "Sid":"AddCannedAcl",
      "Effect":"Allow",
      "Principal": {"AWS": ["arn:aws:iam::111122223333:root","arn:aws:iam::444455556666:root"]},
      "Action":["s3:PutObject","s3:PutObjectAcl"],
      "Resource":"arn:aws:s3:::awsexamplebucket1/*",
      "Condition":{"StringEquals":{"s3:x-amz-acl":["public-read"]}}
    }
  ]
}
```

Reference: https://docs.aws.amazon.com/ja_jp/AmazonS3/latest/dev/example-bucket-policies.html

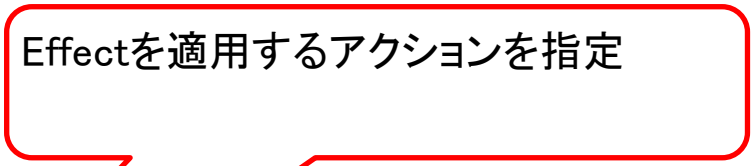
# S3バケットポリシー

対象となるプリンシパル（IAMユーザーやルートアカウントなど）を指定

```
{
  "Version":"2012-10-17",
  "Statement":[
    {
      "Sid":"AddCannedAcl",
      "Effect":"Allow",
      "Principal": {"AWS": ["arn:aws:iam::111122223333:root","arn:aws:iam::444455556666:root"]},
      "Action":["s3:PutObject","s3:PutObjectAcl"],
      "Resource":"arn:aws:s3:::awsexamplebucket1/*",
      "Condition":{"StringEquals":{"s3:x-amz-acl":["public-read"]}}
    }
  ]
}
```

Reference: https://docs.aws.amazon.com/ja_jp/AmazonS3/latest/dev/example-bucket-policies.html

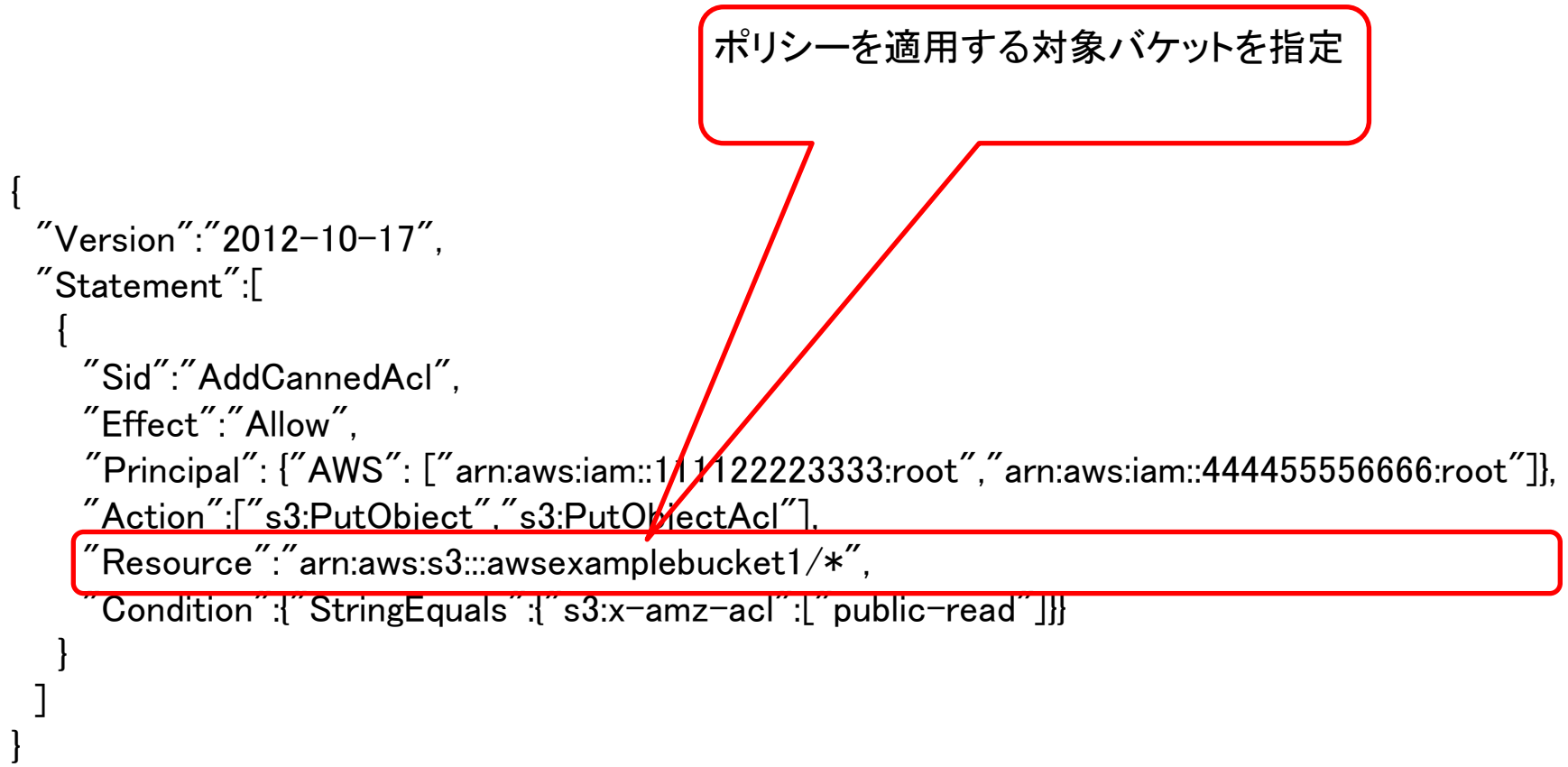
# S3バケットポリシー




```
{
  "Version":"2012-10-17",
  "Statement":[
    {
      "Sid":"AddCannedAcl",
      "Effect":"Allow",
      "Principal": {"AWS": ["arn:aws:iam::111122223333:root","arn:aws:iam::444455556666:root"]},
      "Action":["s3:PutObject","s3:PutObjectAcl"],
      "Resource":"arn:aws:s3:::awsexamplebucket1/*",
      "Condition":{"StringEquals":{"s3:x-amz-acl":["public-read"]}}
    }
  ]
}
```

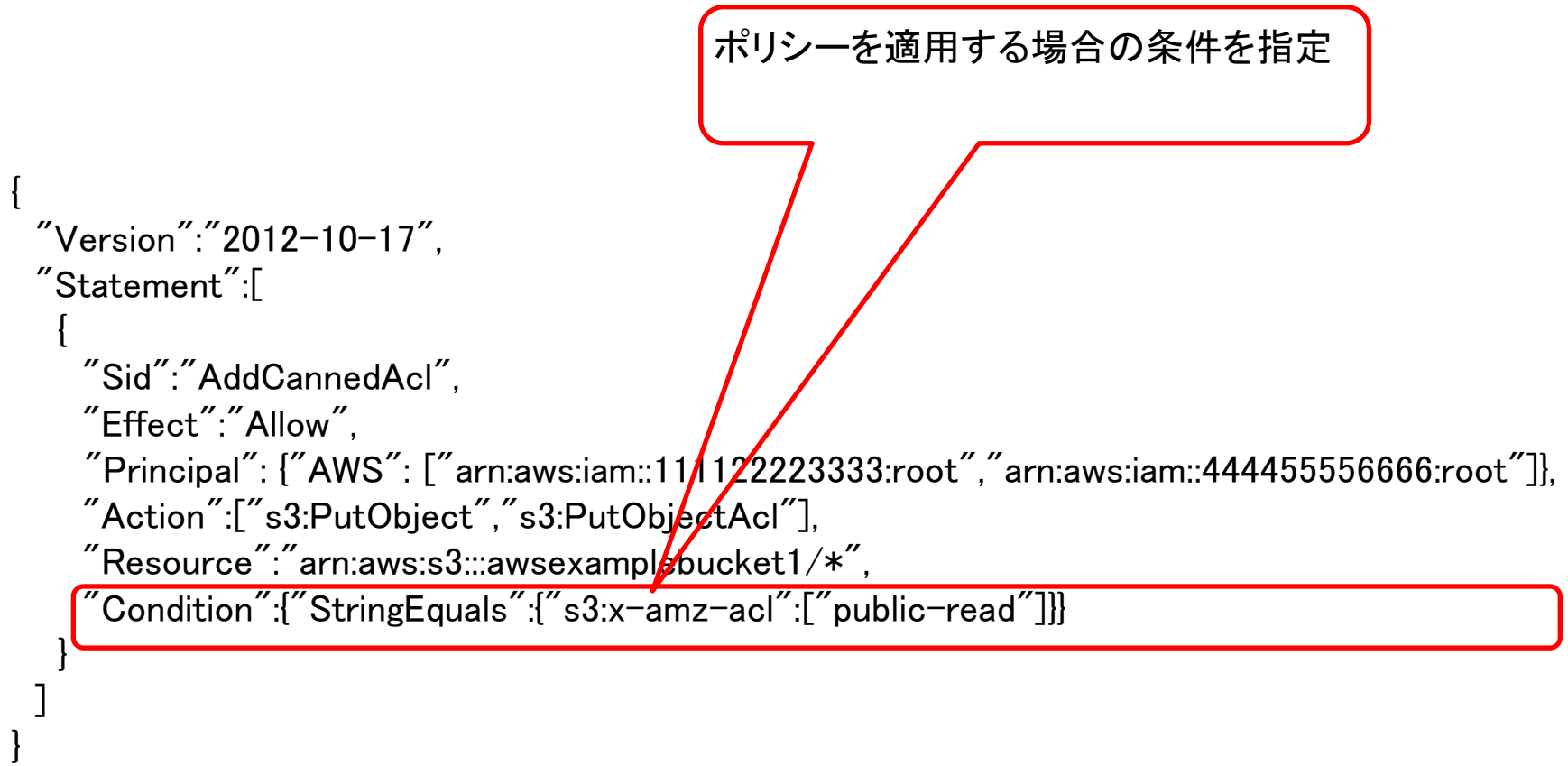
Reference: https://docs.aws.amazon.com/ja_jp/AmazonS3/latest/dev/example-bucket-policies.html

# S3バケットポリシー




Reference: https://docs.aws.amazon.com/ja_jp/AmazonS3/latest/dev/example-bucket-policies.html

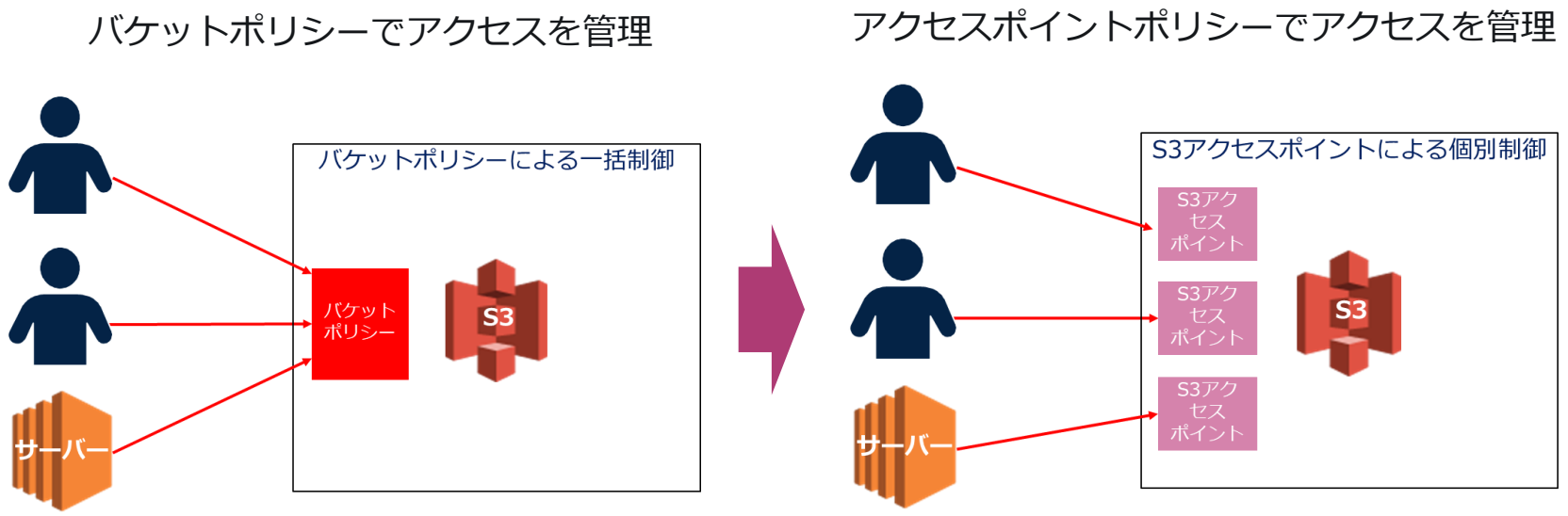
# S3バケットポリシー




Reference: https://docs.aws.amazon.com/ja_jp/AmazonS3/latest/dev/example-bucket-policies.html

# S3アクセスポイント

アクセス先に応じてアクセスポイントを作成して、ポリシーを適用してアクセス設定が可能になる。

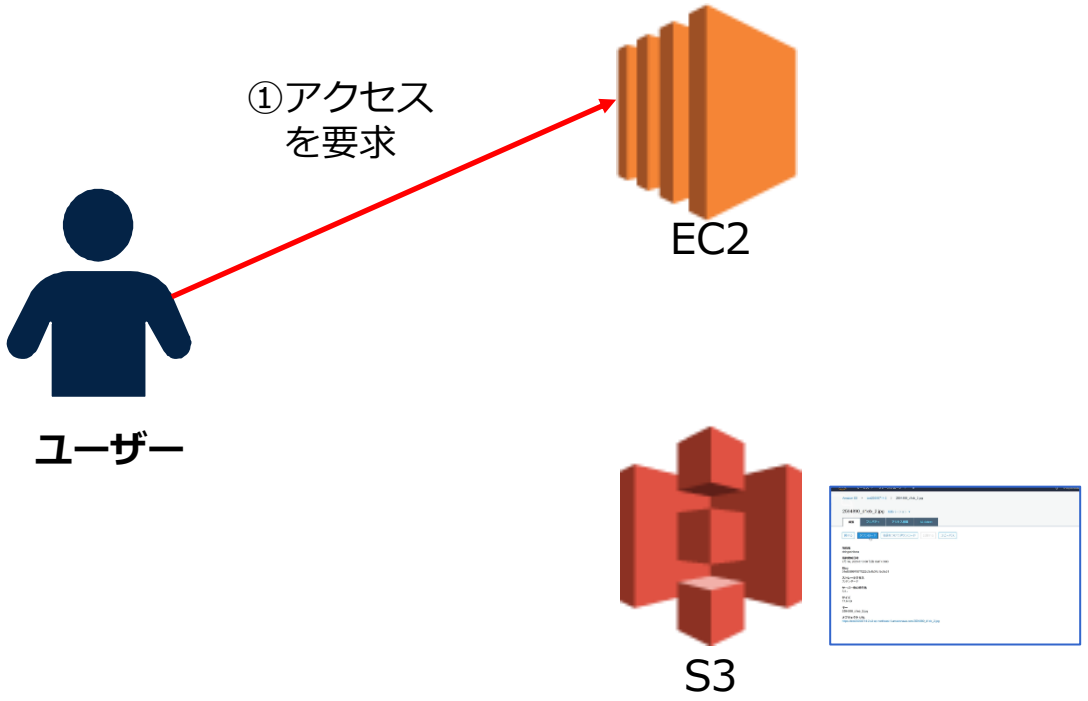


# ブロックパブリックアクセス

## インターネットからのアクセスをブロックする機能で、バケット作成時に初期設定で有効化されている。

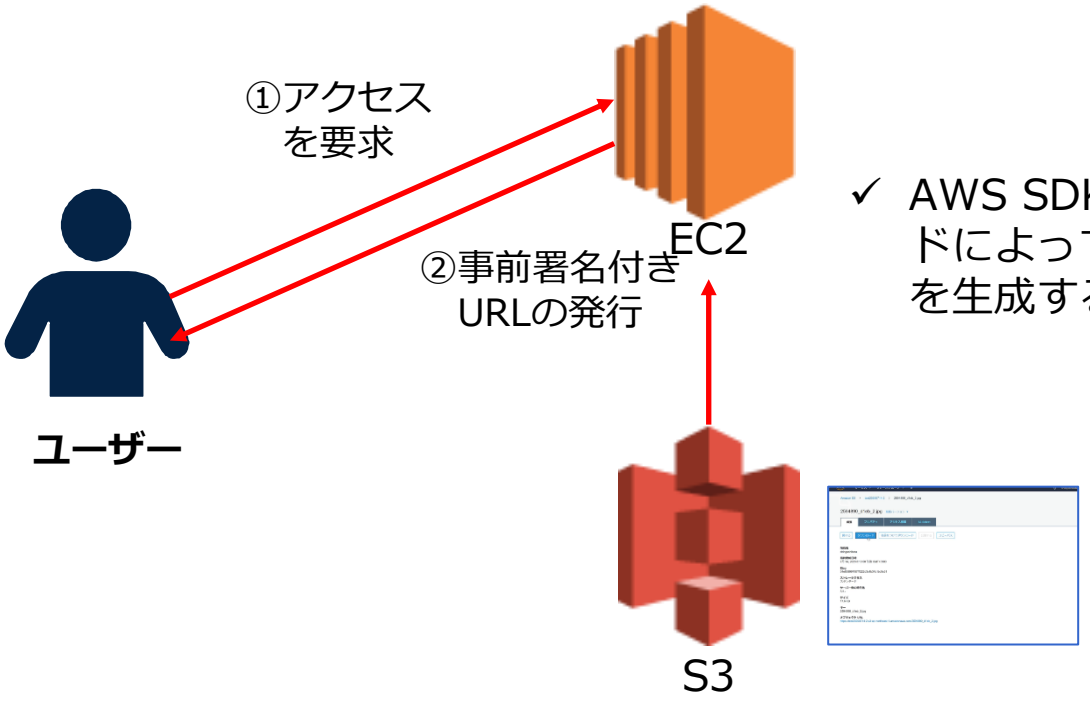
# 事前署名付きURL

事前署名付きURLにより、特定のユーザーのみがアクセスできる特別なURLが利用可能になる。



# 事前署名付きURL

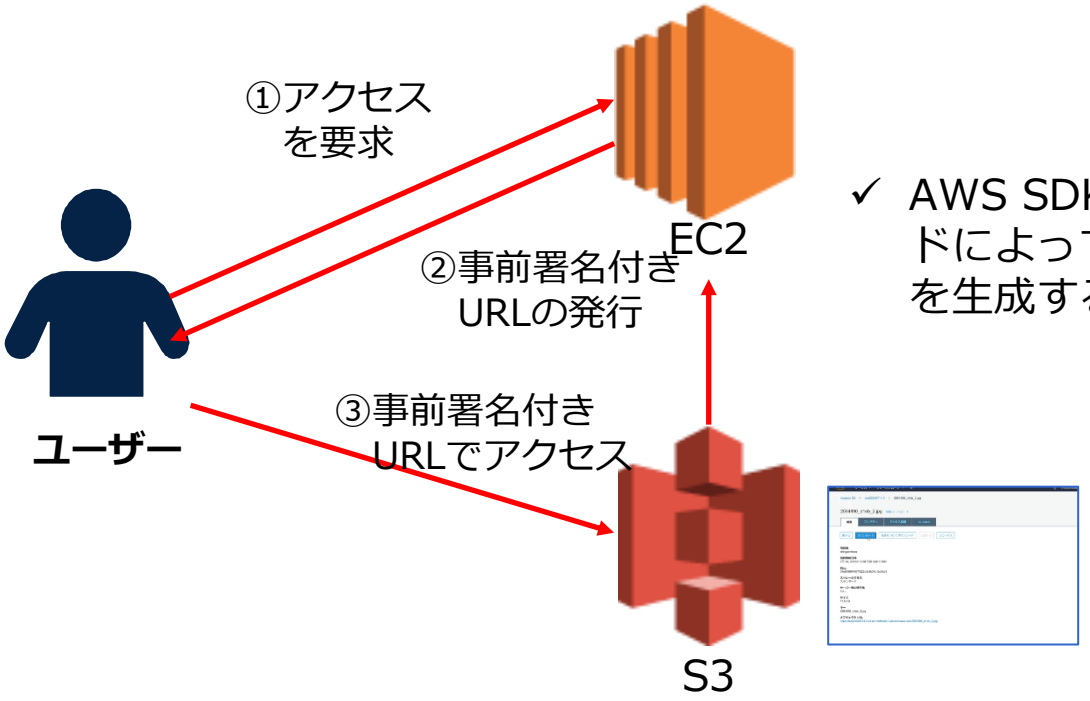
事前署名付きURLにより、特定のユーザーのみがアクセスできる特別なURLが利用可能になる。




- ✓ AWS SDKまたはCLIコマンドによって事前署名付きURLを生成することができる。

# 事前署名付きURL

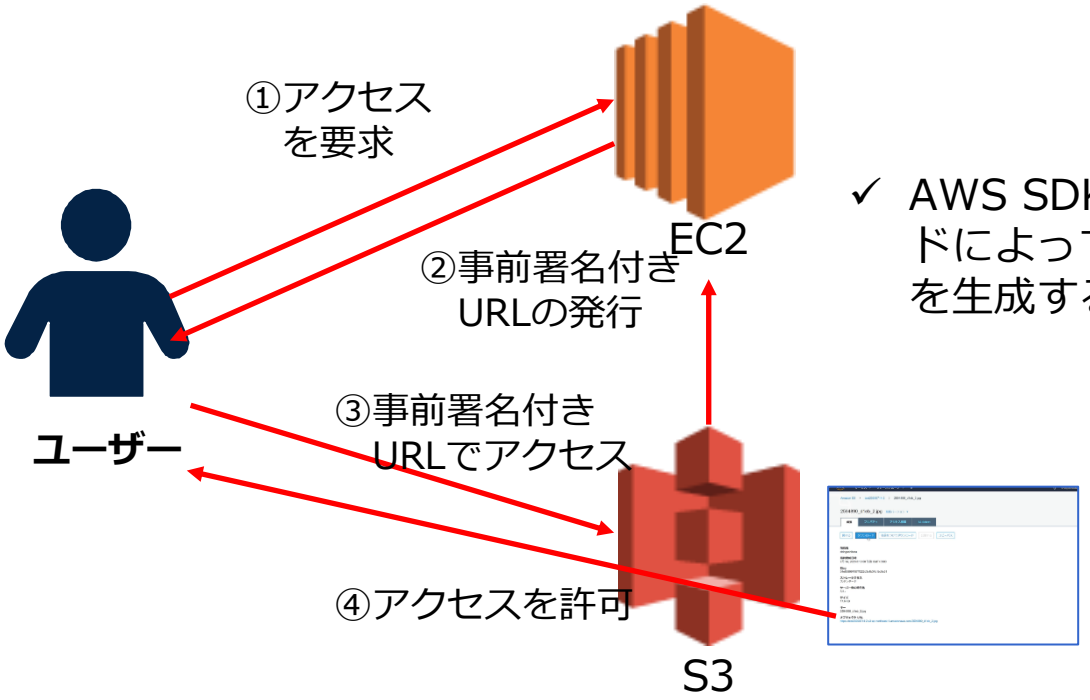
事前署名付きURLにより、特定のユーザーのみがアクセスできる特別なURLが利用可能になる。




- ✓ AWS SDKまたはCLIコマンドによって事前署名付きURLを生成することができる。

# 事前署名付きURL

事前署名付きURLにより、特定のユーザーのみがアクセスできる特別なURLが利用可能になる。




- ✓ AWS SDKまたはCLIコマンドによって事前署名付きURLを生成することができる。

事前署名付きURLは一定時間が経過すると利用できなくなる。
デフォルトでは3600秒間

# S3のデータ保護

# S3の通信暗号化

S3の公開エンドポイントはデフォルトでHTTPSが利用されており、HTTPSによる自動的にSSL/TLS通信が実施される。

| | |
|---|---|
| **HTTPエンドポイント** | ✓ 非暗号化されたURLを利用したアクセス<br>✓ 選択可能だが、デフォルトでは設定されておらず、非推奨となっている。 |
| **HTTPSエンドポイント** | ✓ 暗号化された通信経路でSSL/TLS通信が自動で適用される。<br>✓ ユーザー側で証明書などの設定は必要ない |

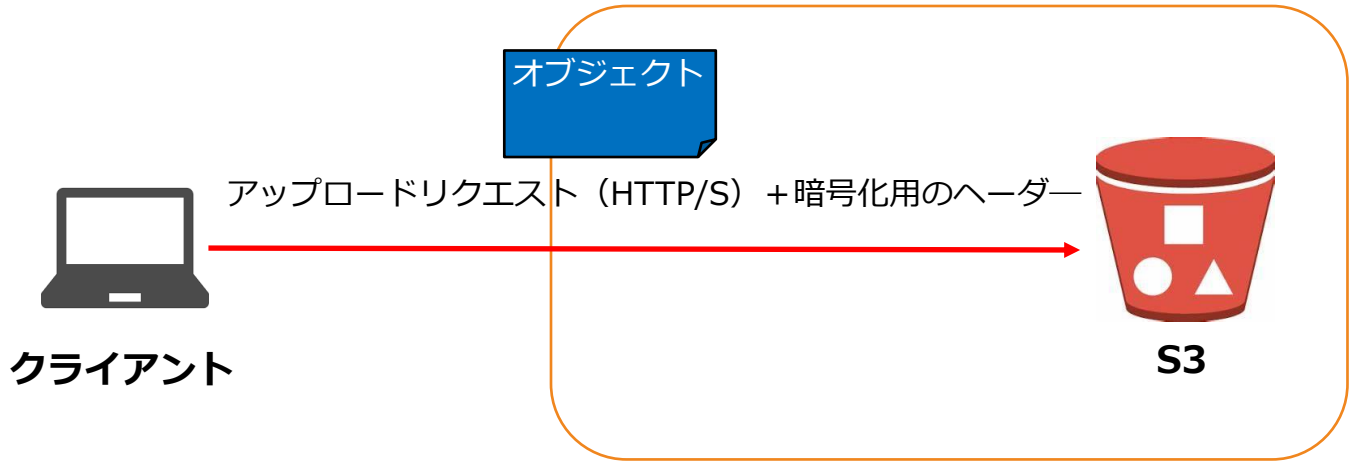
# S3の保管データの暗号化

## S３へのデータ保管時に暗号化形式として以下の４つの形式から選択する

| 暗号化方式 | 特徴 |
| --- | --- |
| **SSE-S3** | ✓ S3の標準暗号化方式で簡易に利用可能<br>✓ 暗号化用のマネージドキーの作成・管理をS3側で自動で実施<br>✓ ブロック暗号の 1 つである 256 ビットの Advanced Encryption Standard (AES-256) を使用してデータを暗号化 |
| **SSE-KMS** | ✓ AWS KMSに設定したキーを利用した暗号化を実施<br>✓ ユーザー側でAWS KMSを利用して暗号化用のマネージドキーを作成・管理することが可能<br>✓ AES-256を利用 |
| **SSE-C** | ✓ ユーザーが指定した暗号化用のマネージドキーをデータと共に送付して、サーバー暗号化 (SSE-C) を実施する<br>✓ 利用設定や管理が煩雑になるのがデメリット |
| **クライアントサイド暗号化（CSE）** | ✓ クライアント側の暗号化では、Amazon S3 に送信する前にデータを暗号化する方式<br>✓ アプリケーションに保存したマスターキーを使用 |

# 暗号化リクエスト

オブジェクトのアップロードリクエスト処理の際に暗号化用のヘーダ―が付与されて暗号化が実施される。

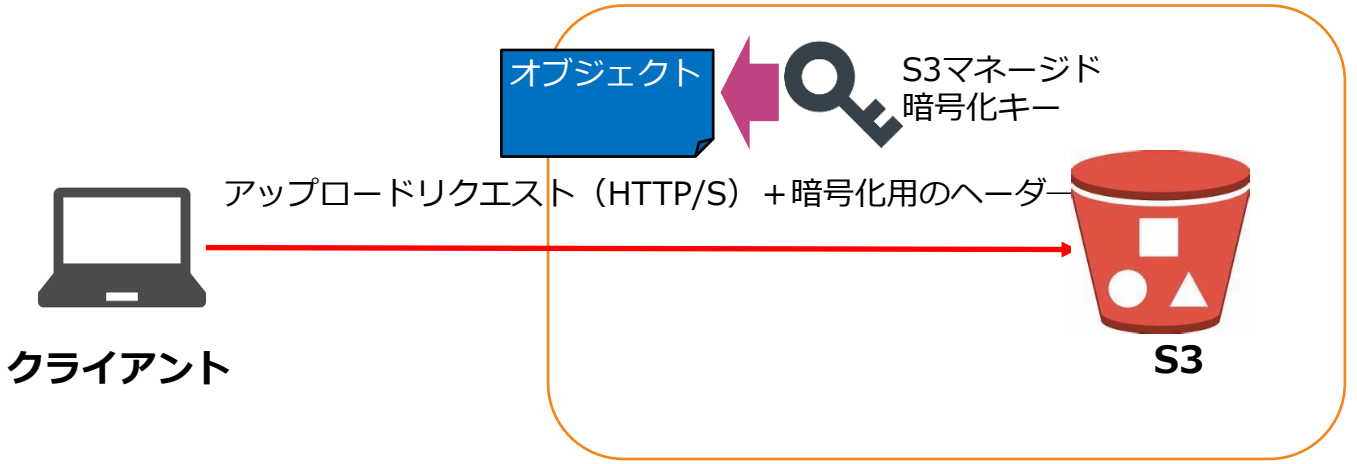


# SSE-S3

## SSE-S3ではS3がマネージドサービスとして管理する暗号化用のマネージドキーを利用してデータを暗号化する。

**SSE-S3**

- ✓ S3の標準暗号化方式で簡易に利用可能
- ✓ 暗号化用のマネージドキーの作成・管理をS3側で自動で実施
- ✓ ブロック暗号の 1 つである 256 ビットの Advanced Encryption Standard (AES-256) を使用してデータを暗号化

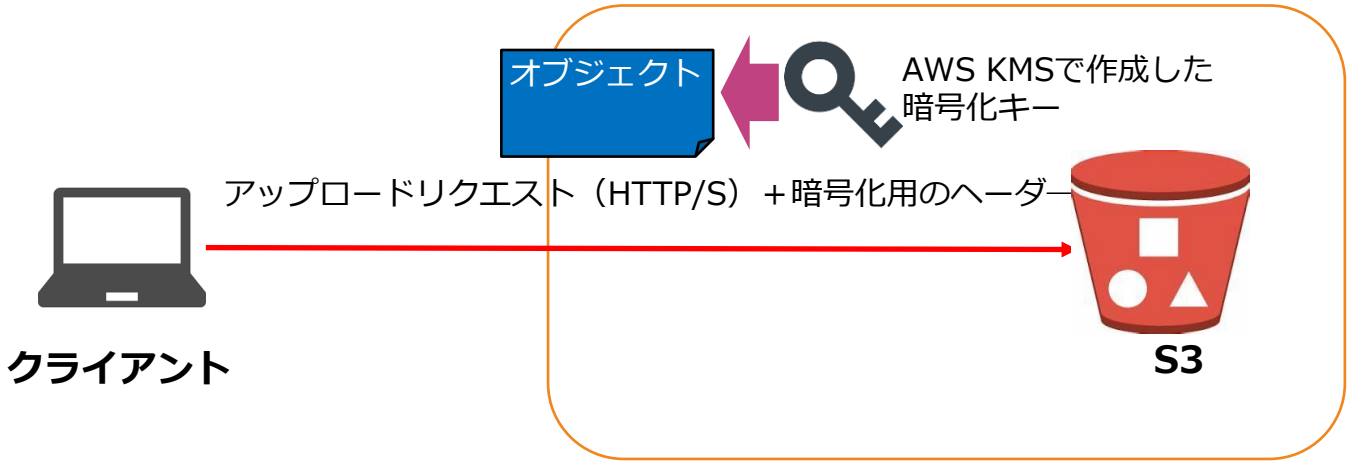


# SSE-KMS

## AWS KMSで暗号化用のマネージドキーを事前に作成して、ユーザー側でキーを指定して暗号化を実施する。

**SSE-KMS**

- ✓ AWS KMSに設定した暗号化用のマネージドキーを利用した暗号化を実施
- ✓ ユーザー側でAWS KMSを利用して暗号化用のマネージドキーを作成・管理する。

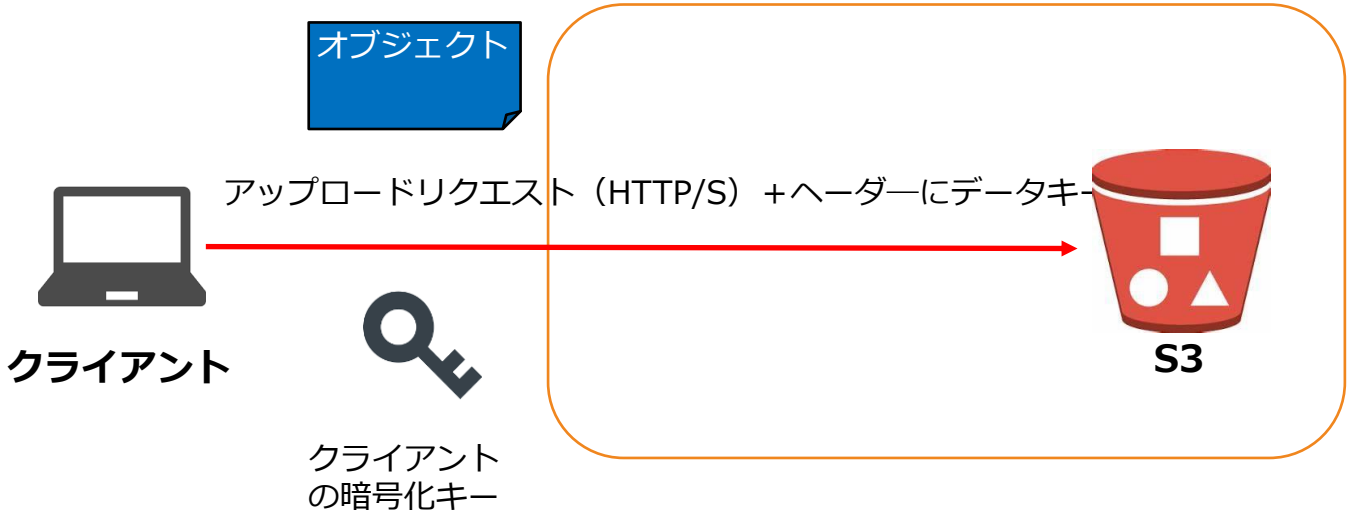


# SSE-C

## ユーザー側で作成した暗号化用のマネージドキーをデータと共に送付して、暗号化をサーバーサイドで実施する。

**SSE-C**

- ✓ ユーザーが指定した暗号化用のマネージドキーによるサーバー側の暗号化 (SSE-C) を使用することが可能
- ✓ 利用設定や管理が煩雑になるのがデメリット

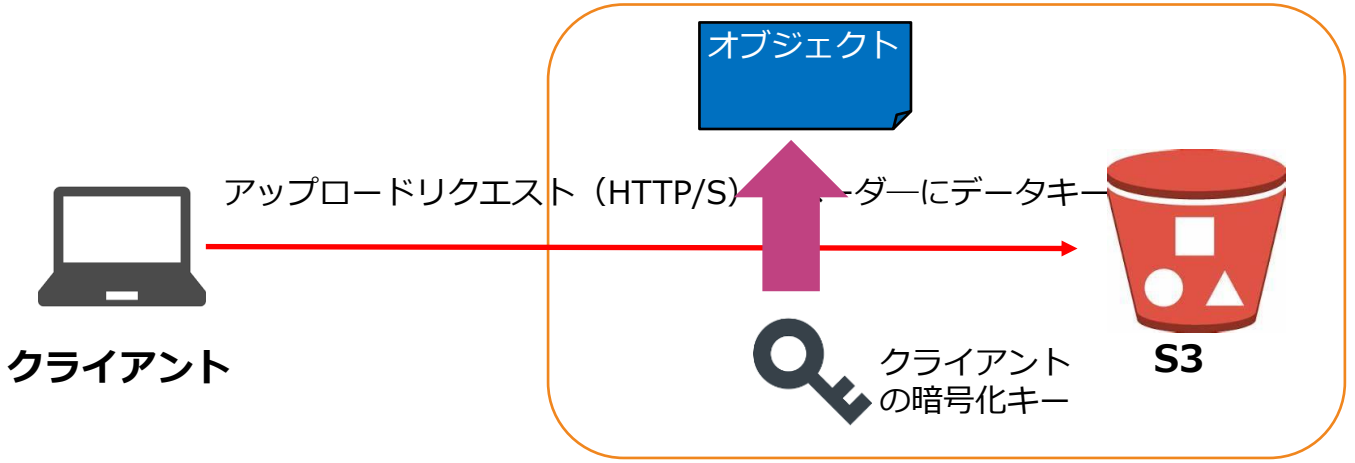


# SSE-C

## ユーザー側で作成した暗号化用のマネージドキーをデータと共に送付して、暗号化をサーバーサイドで実施する。

**SSE-C**

- ✓ ユーザーが指定した暗号化用のマネージドキーをデータと共に送付して、サーバー暗号化 (SSE-C) を実施する
- ✓ 利用設定や管理が煩雑になるのがデメリット

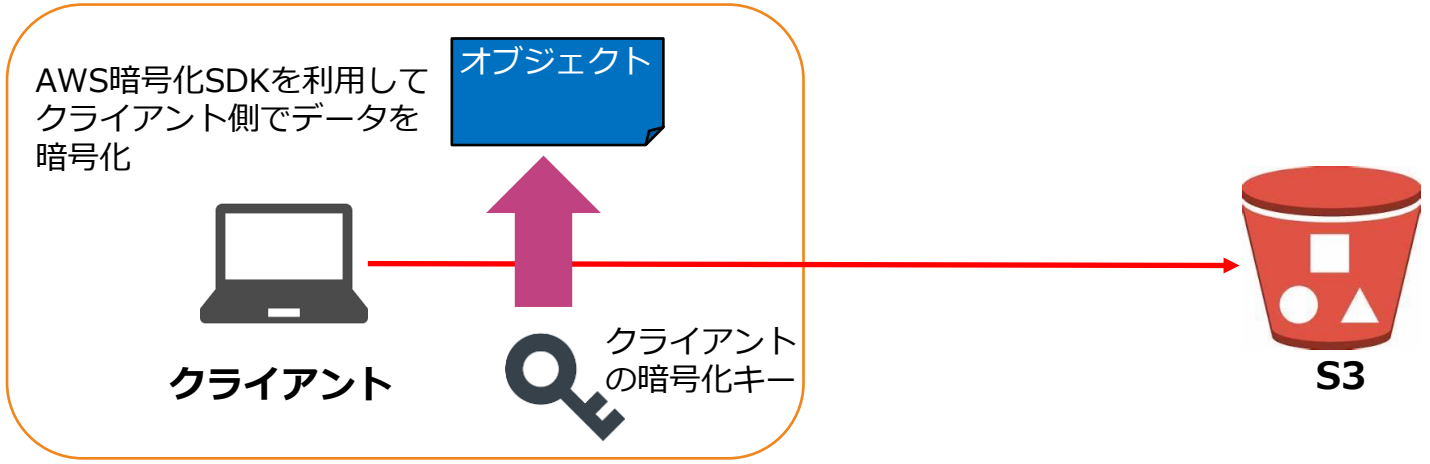


# CSE

## S３へのデータ保管時に暗号化形式として以下の４つの形式から選択する

**CSE**

- ✓ クライアント側の暗号化では、Amazon S3 に送信する前にデータを暗号化する方式
- ✓ アプリケーションに保存したマスターキーを使用



# S3 MFA Delete

## バージョニング機能のオプションとして、オブジェクト削除時にMFA認証を必須にできる。


- ✓ バージョニングを有効化したときのみ利用可能
- ✓ バージョニングファイルを削除する際にMFAが必要となる。
- ✓ バージョニングを停止する際にMFAが必要となる。
- ✓ バケットの所有者のみがMFA Deleteを設定可能

# オブジェクトロック

## アップロードされたデータを更新と削除をできないようにする機能。データが変更されないことを保証する。

- ✓ リテンションモード
  - ガバナンスモード：特別なアクセス許可なしに、ユーザーはオブジェクトのバージョンの上書きや削除、ロック設定を変更することはできない。
  - コンプライアンスモード：ルートユーザーを含め、ユーザーは保護されたオブジェクトのバージョンを上書きまたは削除することはできない。リテンションモードを変更することはできず、保持期間を短縮することはできない。保持期間中にオブジェクトのバージョンを上書きまたは削除できない。
- ✓ オブジェクトロックの有効期間
  - 期間指定：一定期間の間オブジェクトが削除されないようにする。
  - リーガルホールド：永続的にオブジェクトが削除されないようにする。リーガルホールドには関連する保持期間はなく、削除するまで有効となる。

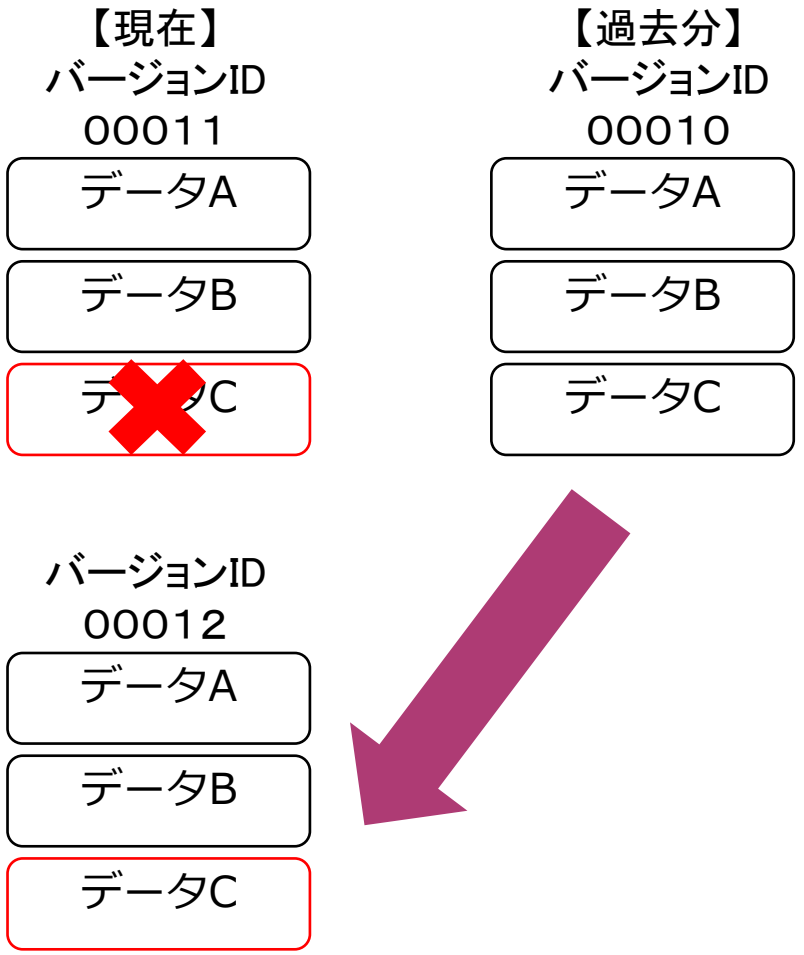
# S3のバージョン管理

# バージョン管理

## ユーザーによる誤操作でデータ削除などが発生してもバージョンから復元できる

**設定**

- バケット全体をバージョン管理する
- バージョンごとにオブジェクトが保管される。
- ライフサイクル管理によりバージョンが保存される期間を設定
- オブジェクトとは別に古いバージョン削除を実施する必要がある。



# S3 MFA Delete

バージョニング機能のオプションとして、オブジェクト削除時にMFA認証を必須にできる。

# S3 Intelligence Tiering の活用

# ストレージクラスの選択

## S3の用途に応じてストレージタイプを選択する

| タイプ | 特徴 | 性能 | 追加課金 |
|---|---|---|---|
| **STANDARD** | ✓ 複数個所にデータを複製するため耐久性が非常に高く、頻繁に利用するデータを大量に保存するのに向いている。<br>✓ データは3AZ以上で分散保存される。 | ■耐久性<br>99.999999999%<br>■可用性<br>99.99% | ■最低利用料金<br>なし<br>■データ取得料<br>なし |
| **STANDARD-IA** | ✓ IAはInfrequency Accessの略であり、低頻度アクセスデータ用のストレージ。One Zone-IAより重要なマスターデータ向け。データ取得は早い<br>✓ Standard に比べて安価だが、One Zone-IAよりは高い。 | ■耐久性<br>99.999999999%<br>■可用性<br>99.9% | ■最低利用料金<br>30日<br>■データ取得料<br>GB当たり取得料 |
| **One Zone-IA** | ✓ 低頻度アクセス用のストレージだが、マルチAZ分散されていないため可用性が低く、重要ではないデータ向け。その分Standard IAよりも値段が安い | ■耐久性<br>99.999999999%<br>■可用性<br>99.5% = 1AZ | ■最低利用料金<br>30日分<br>■データ取得料<br>GB当たり取得料 |
| **S3 Intelligent Tiering** | ✓ 高頻度と低頻度という2つのアクセス階層を利用し、アクセスがあるファイルは高頻度（標準クラス）に維持しつつ、アクセスがないファイルは低頻度（標準IAクラス）に自動で移動する。<br>✓ アクセスパターンがわからない場合に利用 | ■耐久性<br>99.999999999%<br>■可用性<br>99.9% | ■最低利用料金<br>30日<br>■データ取得料<br>なし |

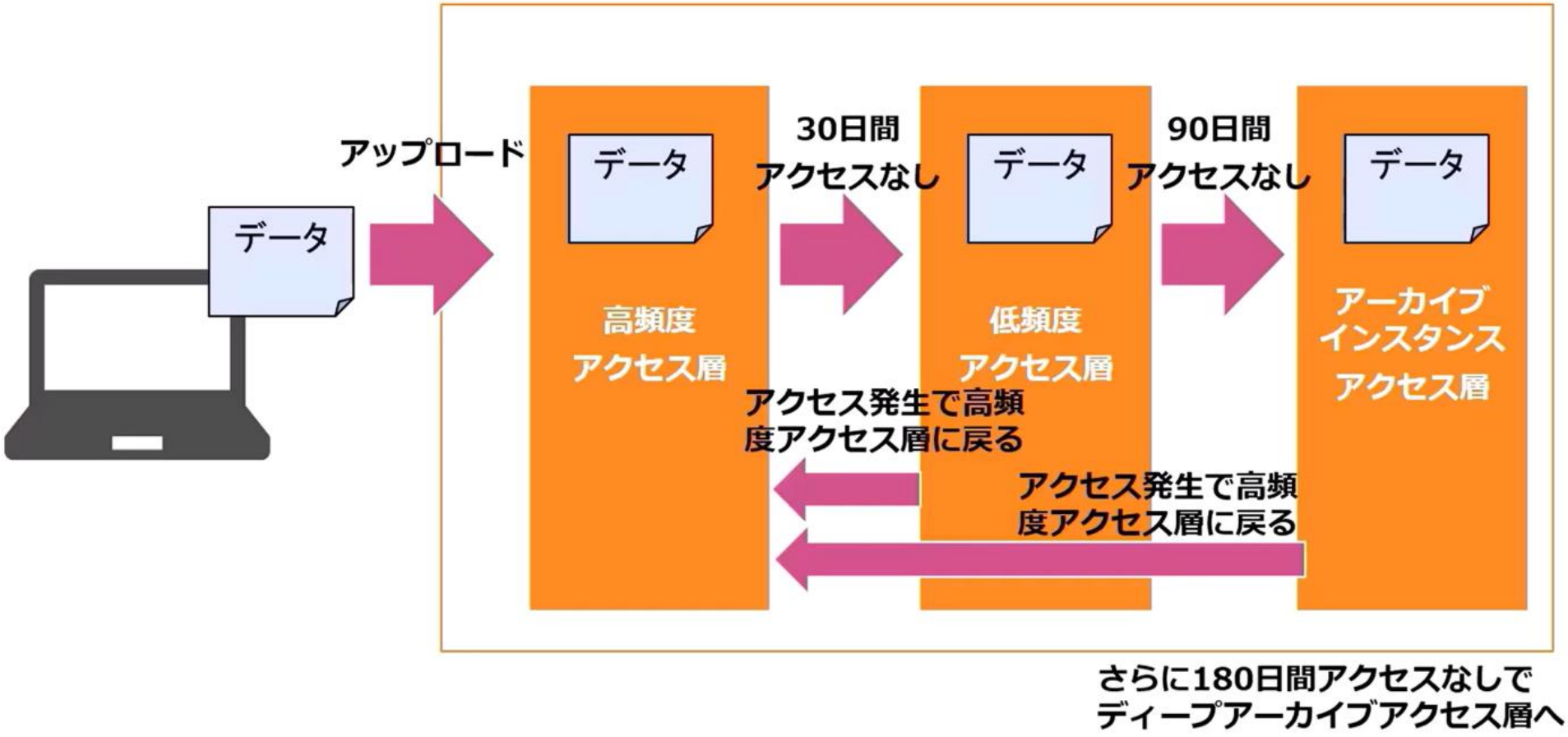
# ストレージクラスの選択

## Glacierでは３つのストレージタイプから選択する。

| タイプ | 特徴 | 性能 | 追加課金 |
|---|---|---|---|
| **S3 Glacier Flexible Retrieval（通常のGlacier）** | ✓ 1 年に 1～2 回アクセスされ、非同期で取り出されるアーカイブデータ向け<br>✓ 通常のデータ検索で（3~5時間）を要する<br>✓ 迅速取り出しで(2～5分)で取り出し可能<br>✓ 一括検索で（5～12時間）で無料<br>✓ ライフサイクルマネジメントで指定<br>✓ ボールトロック機能でデータを保持 | ■耐久性<br>99.999999999%<br>■可用性<br>99.99% | ■最低利用料金<br>90日<br>■データ取得料<br>GB当たり取得料 |
| **S3 Glacier Instant Retrieval** | ✓ アクセスされることがほとんどなく、ミリ秒単位の取り出しが必要な長期間有効なデータ向け<br>✓ 医用画像やニュースメディアなど<br>✓ S3 Standard と同じパフォーマンスのミリ秒単位でのデータの取り出し | ■耐久性<br>99.999999999%<br>■可用性<br>99.9% | ■最低利用料金<br>90日<br>■データ取得料<br>GB当たり取得料 |
| **Amazon Glacier Deep Archive** | ✓ 最安のアーカイブ用ストレージ<br>✓ 7～10 年以上保持される長期間使用されるものの、めったにアクセスされないデータ向け<br>✓ 標準の取り出し速度で12 時間以内にデータを取得<br>✓ 大容量取り出しで48時間以内にデータを取得<br>✓ ライフサイクル管理で指定 | ■耐久性<br>99.999999999%<br>■可用性<br>99.99% | ■最低利用日数<br>180日<br>■データ取得料<br>GB当たり取得料 |

# S3 Intelligent-Tiering

アクセス頻度に応じてオブジェクトを自動的に低コストのアクセス層に移動することでコストを削減する。



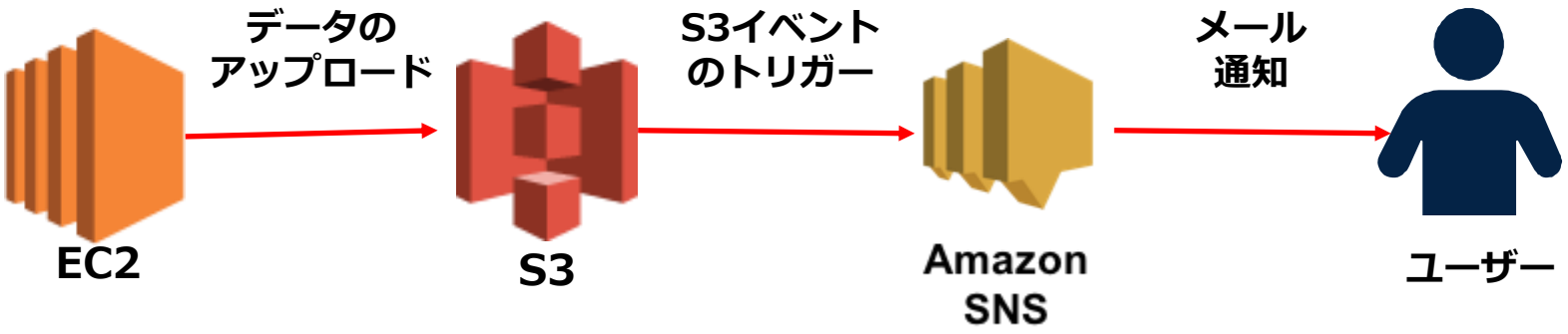
# イベント通知の設定

# S3イベント

## S3オブジェクト操作と連動したシステム連携処理を実現

**S3のイベント通知**

- バケット内イベントの発生をトリガーにして、SNS／SQS／Lambda/Amazon EventBridgeに通知設定が可能
- S3オブジェクト操作と連動したシームレスなシステム連携処理を実現
  －S3へのデータアップロードをSNSでメッセージ通知
  －S3オブジェクトのアップロードをトリガーにLambda関数を実行



# S3のパフォーマンス

# S3のパフォーマンス

## S３はプレフィックスに応じてリクエスト処理が可能なパフォーマンスが調整できる。

- ✓ プレフィックスごとに 1 秒あたり 3,500 回以上の PUT/COPY/POST/DELETE リクエストまたは 5,500 回以上の GET/HEAD リクエストが可能。
- ✓ 発生するリクエスト率が高い場合 (オブジェクトに対して 1 秒あたり 5,000 リクエストを超える率が持続される)に HTTP 503 slowdown レスポンスが発生する。
- ✓ データ転送時には約 100～200 ミリ秒の一定のレイテンシーを実現できる。

# パフォーマンスの向上

## カスタムプレフィックスを利用して並列処理をすることでパフォーマンスを向上させる

- パフォーマンスを最適化するためにカスタムプレフィックスを設定する。例えば、日付ベースの順次命名を使用する。
- カスタムプレフィックスを作成してデータ処理の並列化が可能となる。例えば、10個のプレフィックスで読み取りを並列化すると、 1 秒あたり 55,000 回の読み取りが可能となる。
- 複数の接続で データを同時に GET または PUT するアプリケーションを使用することで高スループット転送が可能となる。

# S3の整合性モデル

## Ｓ３はデータ登録・更新・削除などの処理時に強い整合性モデルを採用している。

| データ処理 | 整合性モデル |
| --- | --- |
| **新規登録** | ✓ Consistency Read<br>✓ 登録後即時にデータが反映される |
| **更新** | ✓ 2020年12月より強い整合性モデルに変更された。そのため、齟齬は発生しない。 |
| **削除** | ✓ 2020年12月より強い整合性モデルに変更された。そのため、齟齬は発生しない。 |

# アップロード時のデータ整合性確認

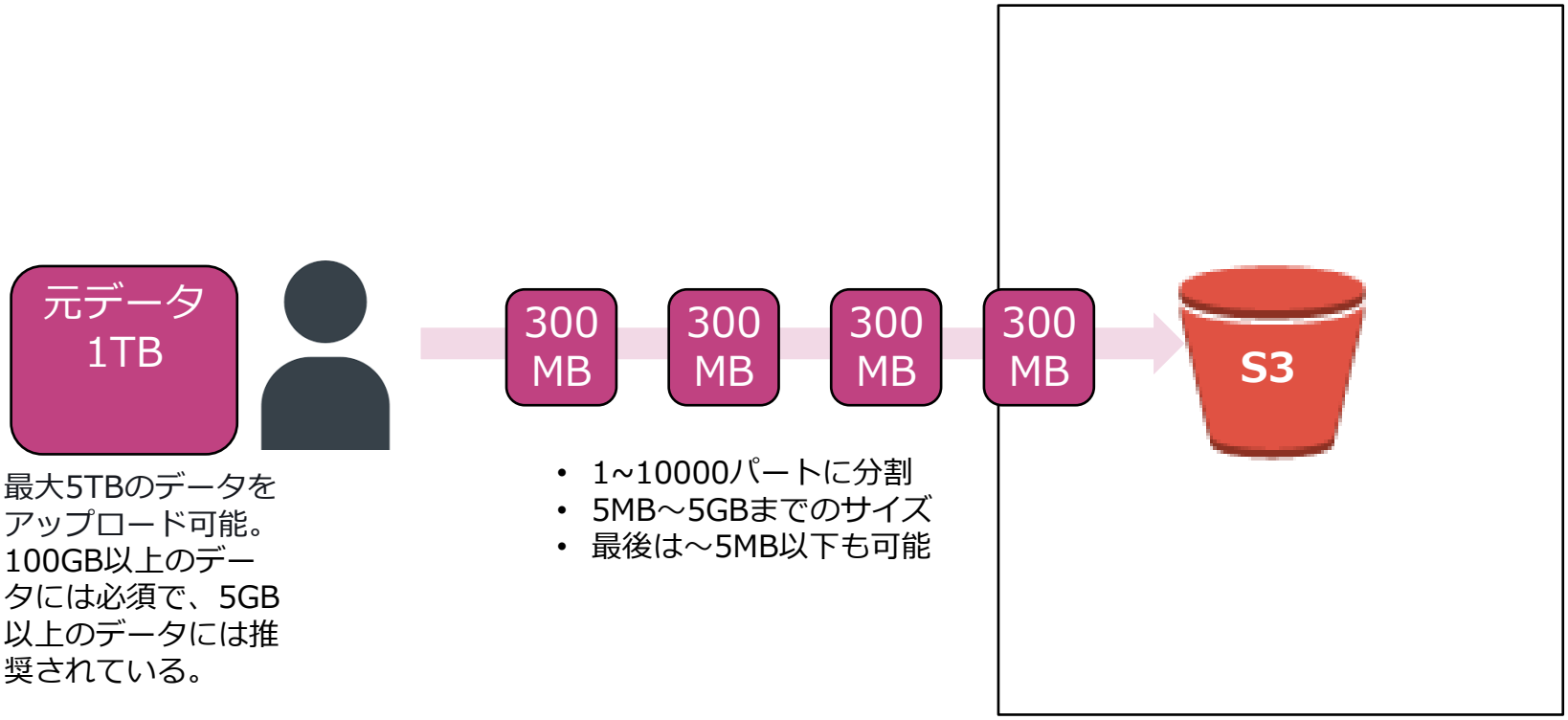
## Content-MD5 ヘッダーを使用してアップロードされたオブジェクトの整合性を確認することができる。

1. オブジェクトの base64 でエンコードされた MD5 チェックサム値を取得します。
2. アップロード中のオブジェクトの整合性を確認します。

   ただし、アップロードが AWS 署名バージョン 4 で署名されている場合、代わりに x-amz-content-sha256 ヘッダーを使用する必要があります。

# マルチパートアップロード

## 大容量オブジェクトをいくつかに分けてアップロードする機能




最大5TBのデータをアップロード可能。100GB以上のデータには必須で、5GB以上のデータには推奨されている。

- 1~10000パートに分割
- 5MB～5GBまでのサイズ
- 最後は～5MB以下も可能

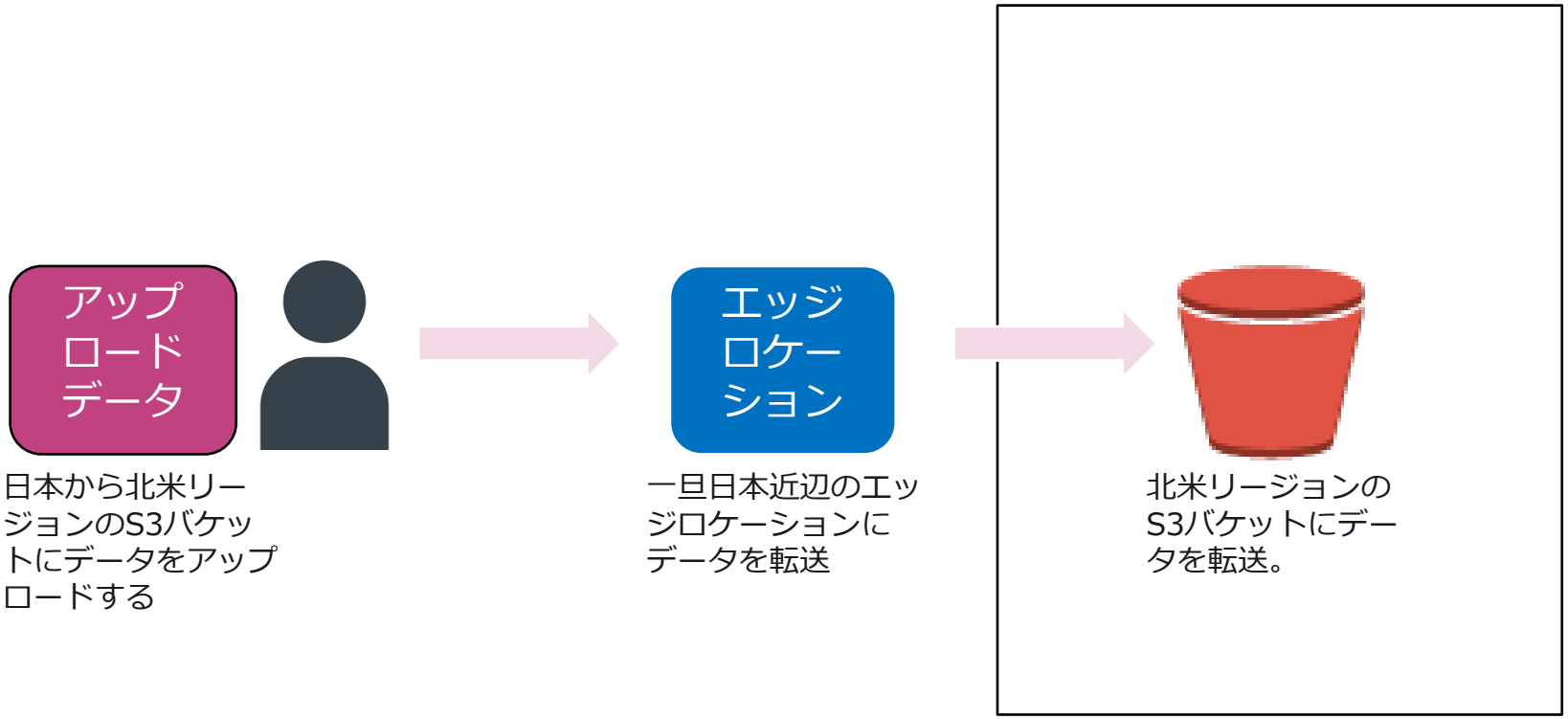
【失敗した場合】

■ アップロードを中止するとパートデータが残る

■ ライフサイクル管理でクリーンアップ設定が可能

# S3 Transfer Acceleration

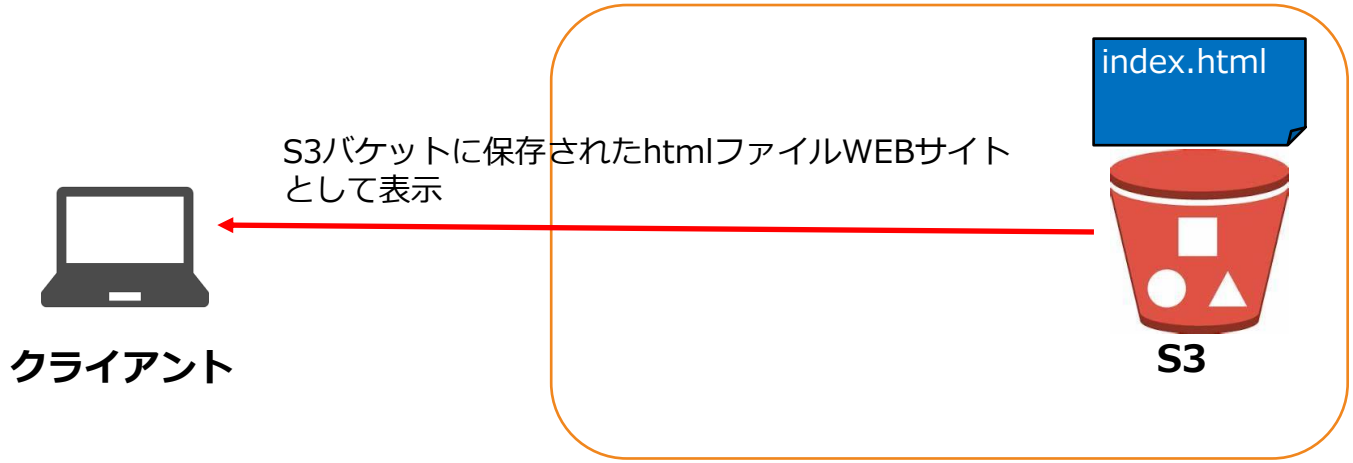
データ転送元から地理的に一番近いエッジロケーションを利用して高速にデータアップロードを実施する。



# 静的WEBホスティング
# の活用

# 静的WEBホスティング

## S3を利用して簡易な静的WEBサイトを作ることができる機能




リージョンに応じてAmazon S3 ウェブサイトエンドポイントは以下の 2 つの形式のいづれかになる。

s3-website ダッシュ (-) リージョン - http://bucket-name.s3-website-Region.amazonaws.com

s3-website ドット (.) リージョン -http://bucket-name.s3-website.Region.amazonaws.com

# 静的WEBホスティング

## 静的サイトを構築する場合は、静的WEBホスティングによる安価なWEBページを構築可能

| | |
|---|---|
| 静的WEBホスティング<br>メリット | • サーバーなしにWEBサイトをホスティング可能。<br>• サーバーが必要ないため値段が安い。<br>• マルチAZの冗長化を勝手にしてくれており、運用いらず<br>• Route53で独自ドメインを設定可能<br>• CloudFrontによる配信可能 |
| 静的WEBホスティング<br>デメリット | • サーバーサイドスクリプト言語を実行するなどの動的サイト不可<br>• 単独ではSSLが利用できず、SSL設定にはCloudFrontが必要 |
| WEBサイト<br>エンドポイント | 使用しているリージョンに応じて、Amazon S3 ウェブサイトエンドポイントは以下の 2 つの形式のいずれかになる。<br>✓ http://bucket-name.s3-website-Region.amazonaws.com<br>✓ http://bucket-name.s3-website.Region.amazonaws.com |

# 静的WEBホスティング

## 静的サイトを構築する場合は、静的WEBホスティングによる安価なWEBページを構築可能

| | |
|---|---|
| 静的WEBホスティング<br>メリット | • サーバーなしにWEBサイトをホスティング可能。<br>• サーバーが必要ないため値段が安い。<br>• マルチAZの冗長化を勝手にしてくれており、運用いらず<br>• Route53で独自ドメインを設定可能<br>• CloudFrontによる配信可能 |
| 静的WEBホスティング<br>デメリット | • サーバーサイドスクリプト言語を実行するなどの動的サイト不可<br>• 単独ではSSLが利用できず、SSL設定にはCloudFrontが必要 |
| WEBサイト<br>エンドポイント | 使用しているリージョンに応じて、Amazon S3 ウェブサイトエンドポイントは以下の 2 つの形式のいずれかになる。<br>✓ http://bucket-name.s3-website-Region.amazonaws.com<br>✓ http://bucket-name.s3-website.Region.amazonaws.com |

# 静的WEBホスティング

## 静的サイトを構築する場合は、静的WEBホスティングによる安価なWEBページを構築可能

# Route53によるドメイン設定

## S3の静的WEBホスティングのサイトにドメインを設定できる。

- トラフィック先としてS3 Webサイトエンドポイントへのエイリアス[Region(地域)]を選択します。
- レコードタイプとしてエイリアスレコードのA レコード（IPv4）タイプを利用してドメインを設定する。
- ターゲットの正常性の評価にはデフォルト値を設定する。
- バケット名とドメイン名またはサブドメイン名と同じにすることが必要

# S3のライフサイクル管理

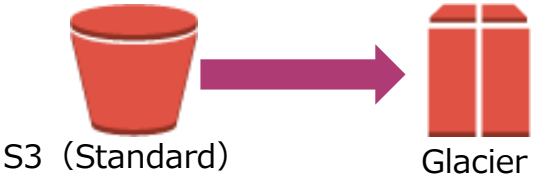
# ライフサイクル管理

## 時間に応じてオブジェクトのストレージクラスの変更や削除を自動的に行うルールを設定できる。

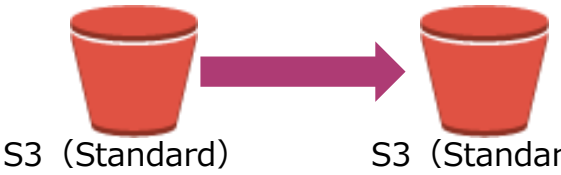
**設定方法**

- バケット全体やプレフィックスに設定
- オブジェクト更新日を基準にして日単位で指定し、毎日0:00UTCにキューを実行
- 最大1000までのライフサイクルルールを利用可能
- IAに移動できるのは128KB以上のオブジェクト
- MFA Deleteが有効だと設定不可

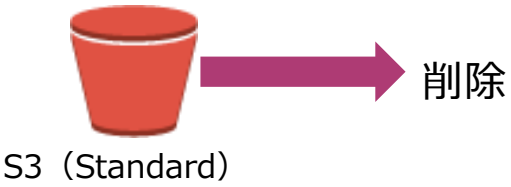
**一定期間で自動アーカイブ**




**一定期間で自動で安価な保存場所へ**

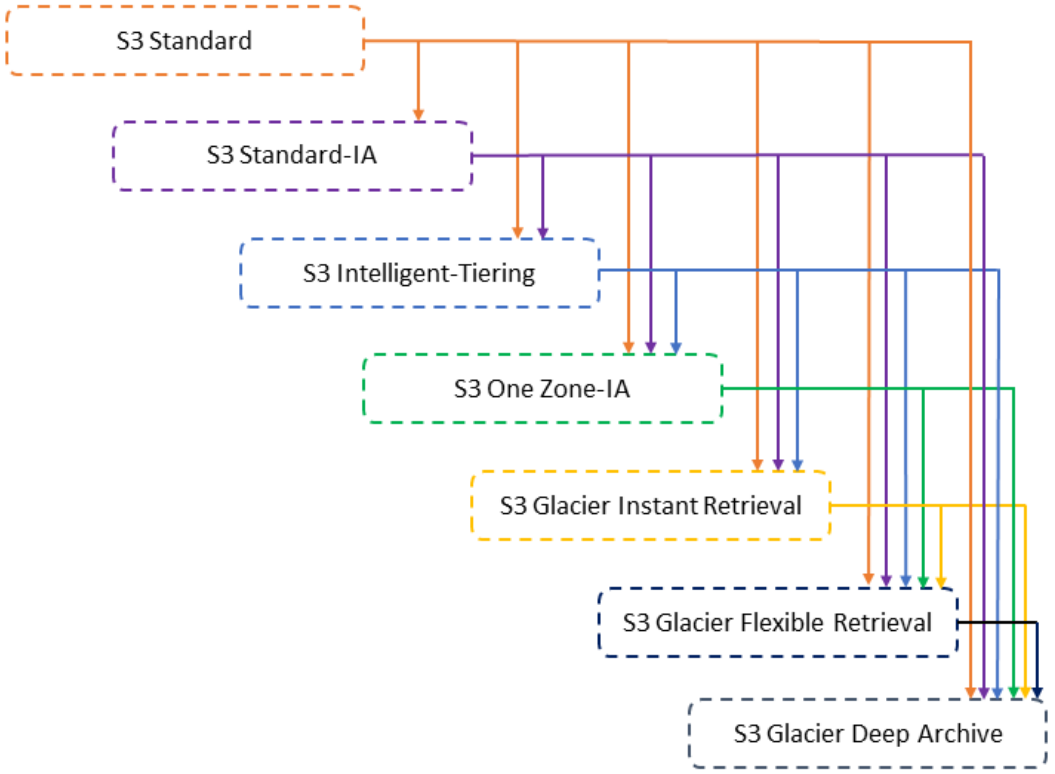



**一定期間で自動で削除**



# ライフサイクル管理

## ライフサイクルポリシーを設定可能なパスは以下の通り




Reference: https://docs.aws.amazon.com/ja_jp/AmazonS3/latest/dev/lifecycle-transition-general-considerations.html

# リージョン間を跨いだ
# レプリケーション

# S3のレプリケーション

S3バケットはレプリケーションによって別のバケットにデータを複製することができる。

| 同一リージョンレプリケーション（SRR） | 別リージョンのレプリケーション（CRR） |
| --- | --- |
| ✓ 同一リージョンのバケットに同じデータを保存する。<br>✓ 本番用のデータとテスト用のデータと分ける際などにレプリケーションしてバケットを複数設定 | ✓ 別リージョンのバケットに同じデータを保存する。<br>✓ 災害対応としてデータを別リージョンに保存<br>✓ リージョン別のアクセスを低レイテンシーにする。 |

# クロスリージョンレプリケーション

## S3はリージョン間を跨ぐクロスリージョンレプリケーションにより耐障害性を高める

### レプリケーションのトリガー

- ✓ レプリケーションを有効後に、バケットにおけるオブジェクトの作成・更新・削除をトリガーにレプリケーションを実行する。
- ✓ 有効前のデータはレプリケーションされない。

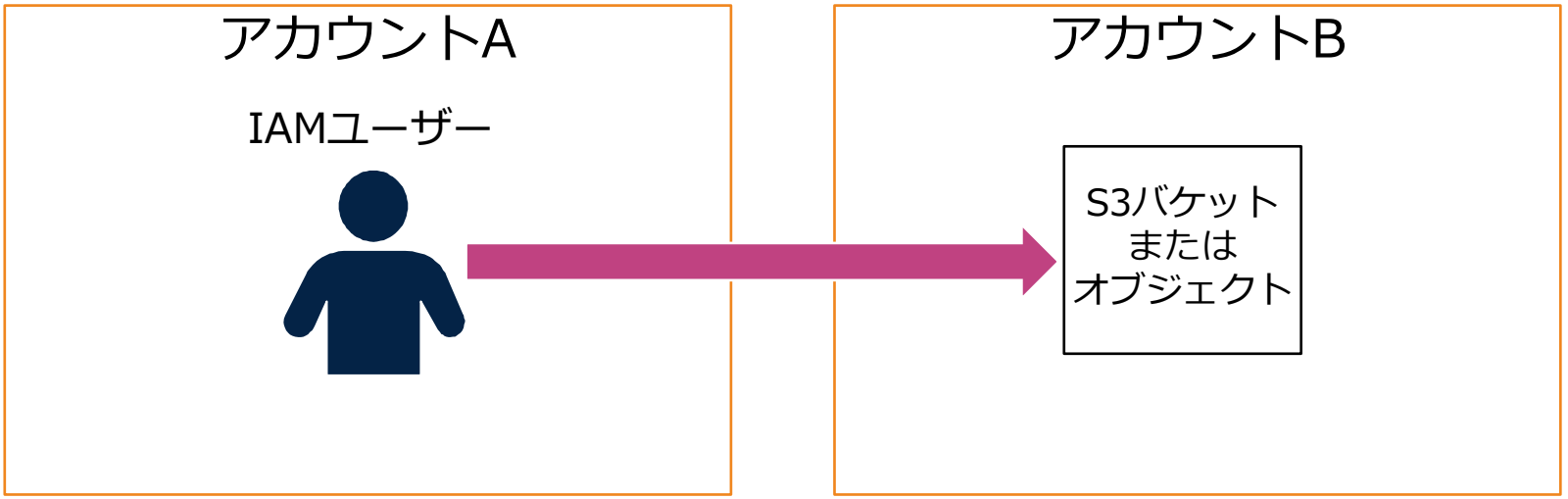
### 設定

- ✓ 事前にバージョニング機能を有効にする必要がある。
- ✓ レプリケーション先となるバケットは別リージョンに設置
- ✓ 双方向レプリケーションも可能
- ✓ データ転送費用が発生
- ✓ レプリケーションは3つ目のバケットには連鎖しない。

# S3のクロスアカウント
# アクセス設定

# S3のクロスアカウントアクセス

## 別のアカウントのIAMユーザーやロールからのアクセスを許可する設定



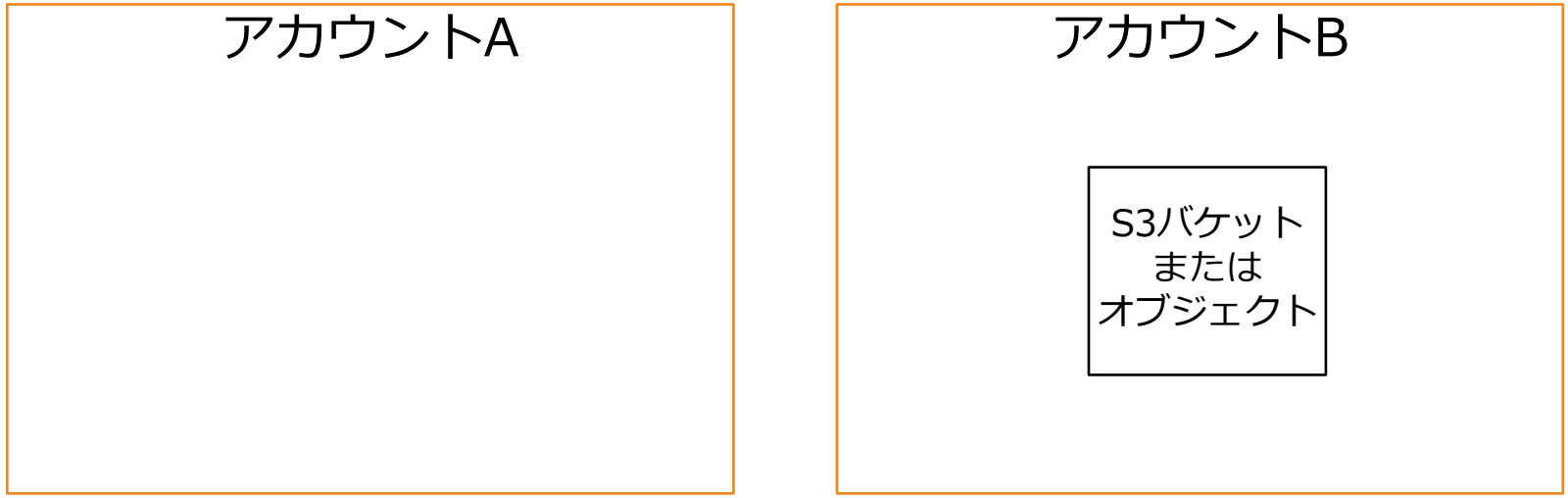
# クロスアカウントアクセス

## クロスアカウントアクセスを許可する設定は3つの方式がある

| 設定方式 | 詳細 |
| --- | --- |
| **バケットポリシーとIAMポリシーによる許可** | ✓ S3バケットへのアクセスを許可するIAMポリシーを設定する。<br>✓ IAMユーザーとロールにIAMポリシーを設定する。<br>✓ S3バケットへのクロスアカウントアクセスを許可する場合はバケットポリシーでアカウントを指定して許可を行う。 |
| **ACLとIAMポリシーによる許可** | ✓ S3バケットへのアクセスを許可するIAMポリシーを設定する。<br>✓ IAMユーザーとロールにIAMポリシーを設定する。<br>✓ S3バケットの特定オブジェクトへのクロスアカウントアクセスを許可する場合はACLでアカウントを指定して許可を設定 |
| **IAMロールによる許可** | ✓ IAMロールの権限移譲を利用して、S3バケット／オブジェクトへのプログラムによるアクセスまたはコンソールアクセス用のクロスアカウントの IAM ロールを設定する。<br>✓ AssumeRoleの実行を許可したロールにより別アカウントのユーザーに権限を委譲する |

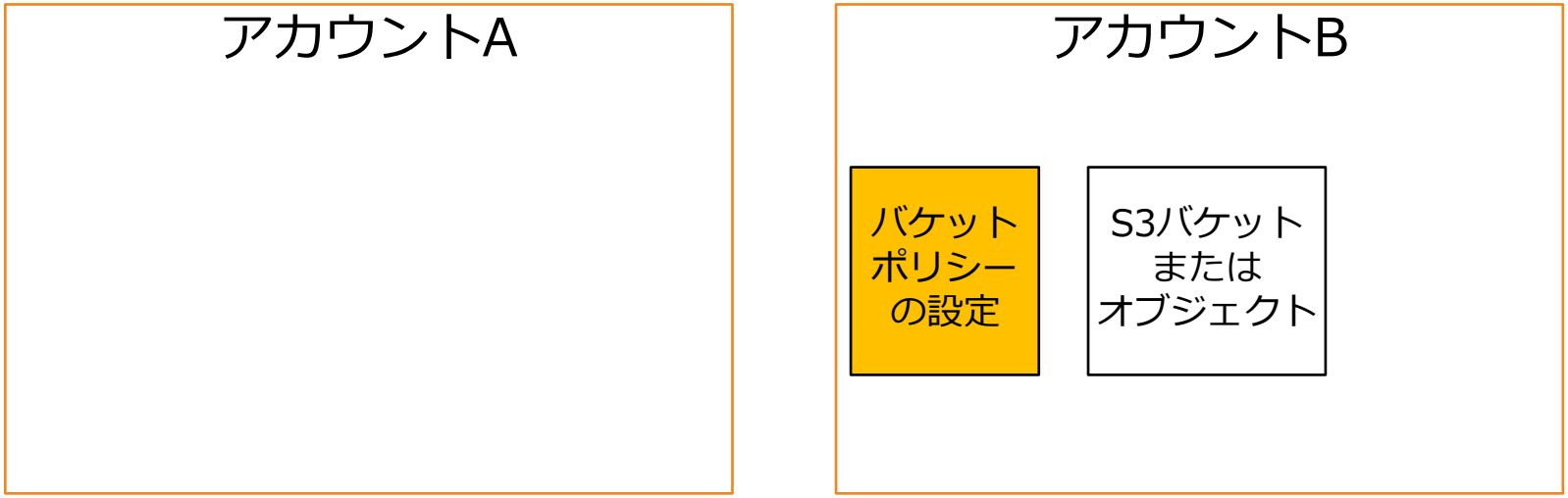
# このハンズオンで実施する内容

1. 現在利用している**アカウントではない**アカウントB
   において新規にS3バケットを作成する。



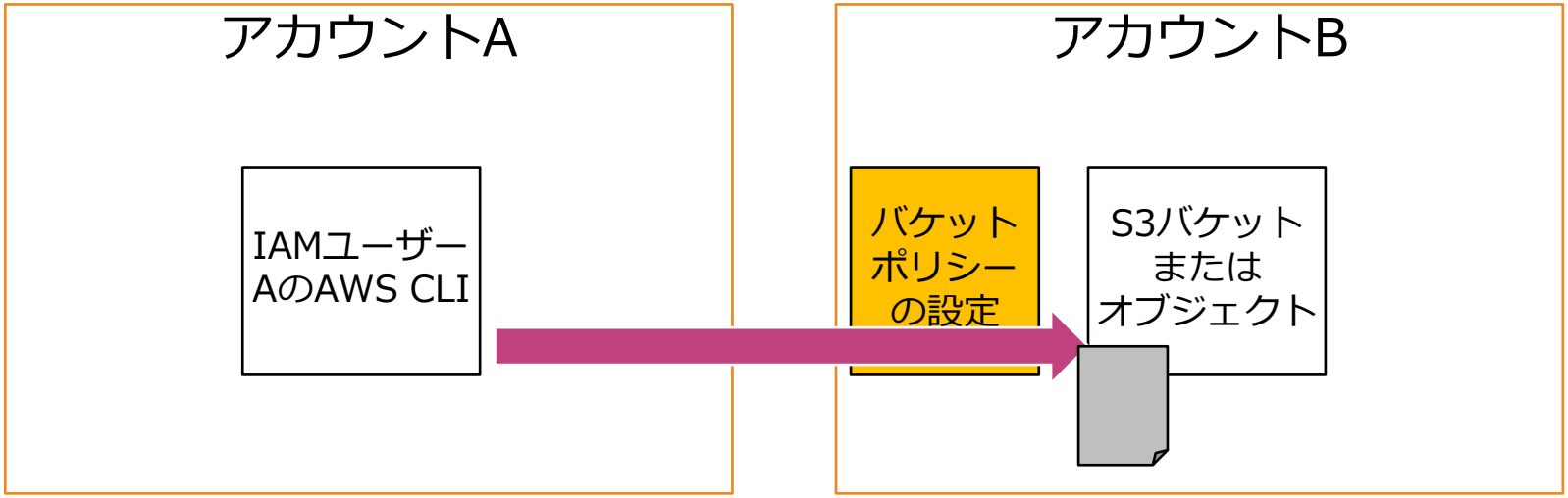
# このハンズオンで実施する内容

1. 現在利用している**アカウントではない**アカウントB
   において新規にS3バケットを作成する。
2. そのバケットに対して、アカウントAのIAMユー
   ザーを許可するバケットポリシーを設定する



# このハンズオンで実施する内容

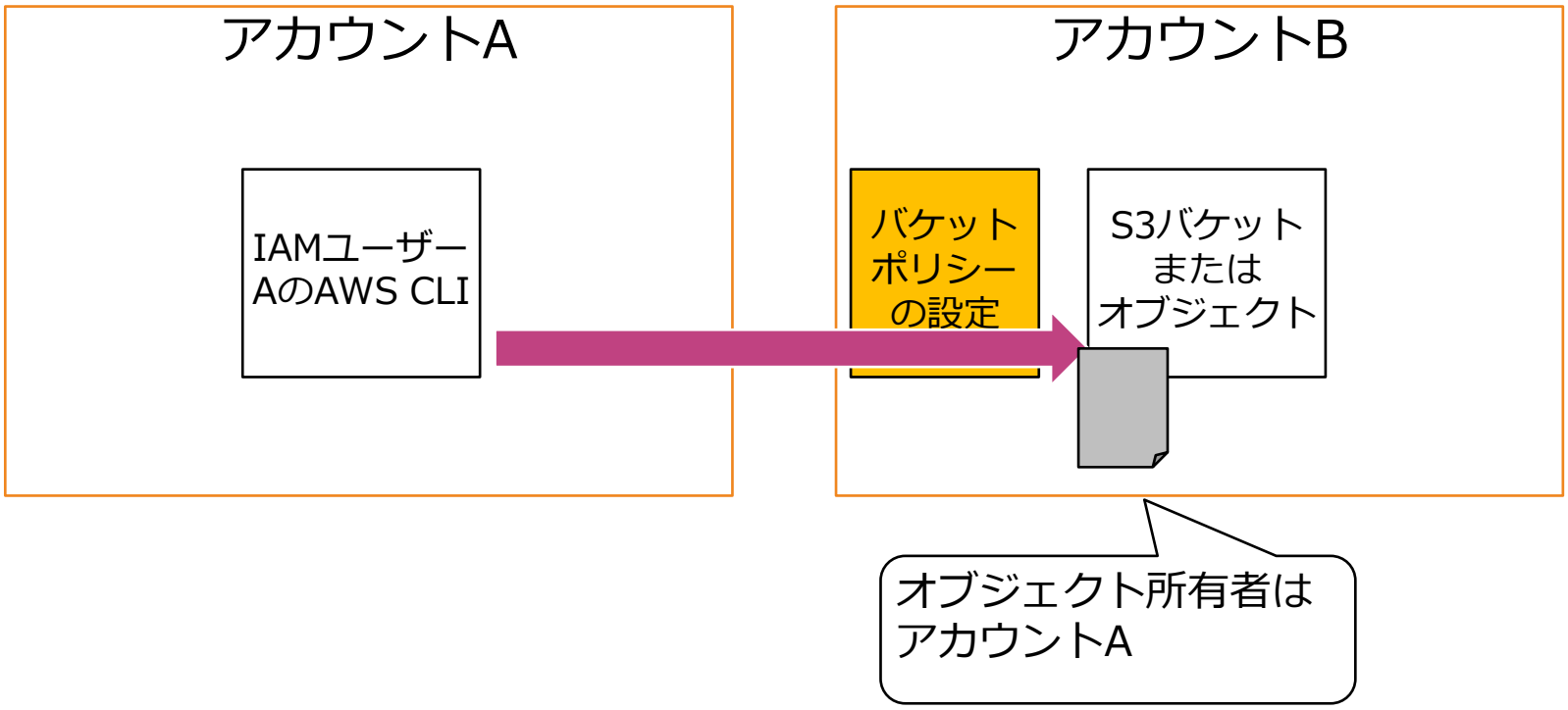
1. 現在利用している**アカウントではない**アカウントBにおいて新規にS3バケットを作成する。
2. そのバケットに対して、アカウントAのIAMユーザーAを許可するバケットポリシーを設定する
3. アカウントAのIAMユーザーAが設定されているAWS CLIからアカウントBのバケットにアクセスして、オブジェクトをアップロードする。



# オブジェクト所有者の変更

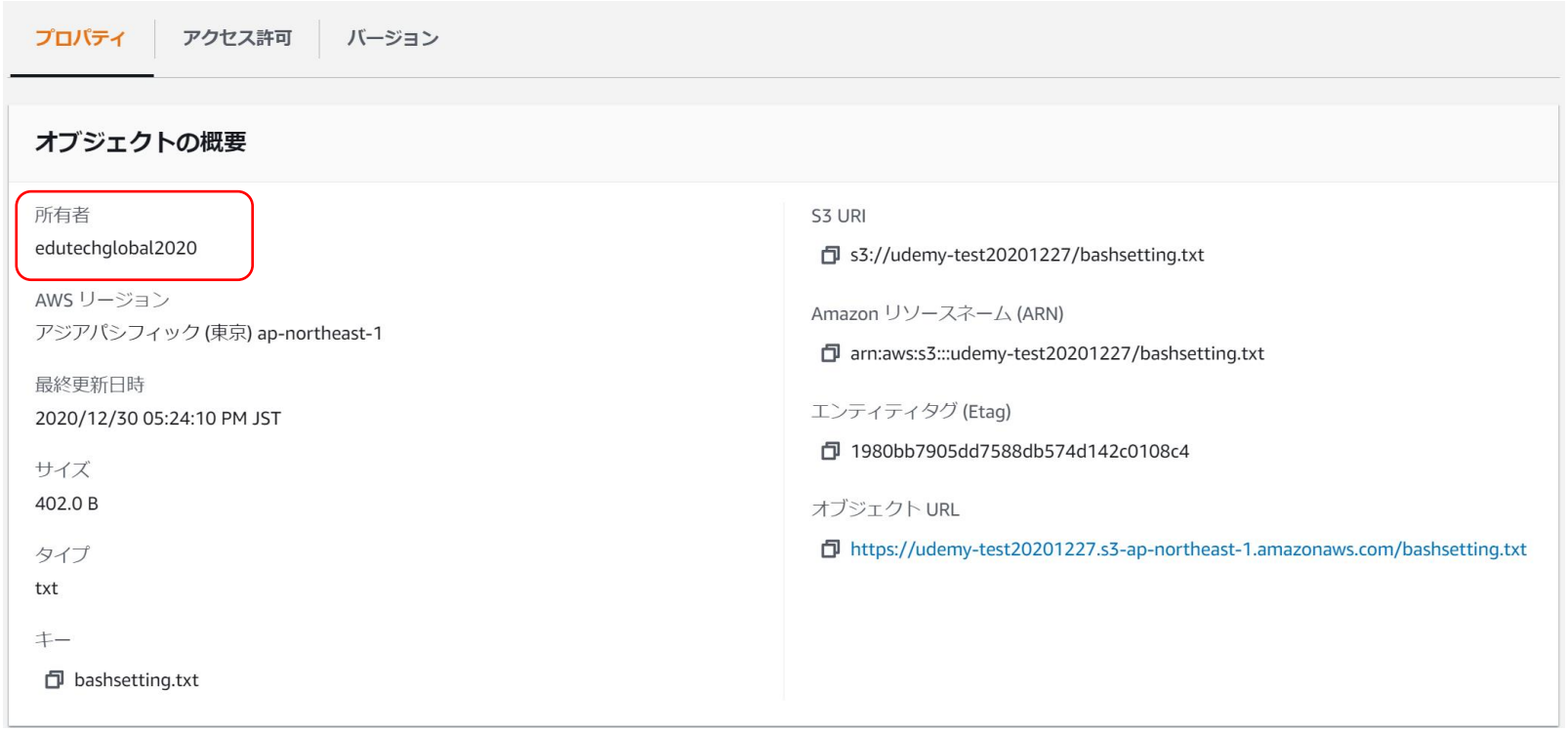
# オブジェクト所有者

## オブジェクトをアップロードしたユーザーがオブジェクト所有者になる



# オブジェクト所有者

## オブジェクトをアップロードしたユーザーがオブジェクト所有者になる



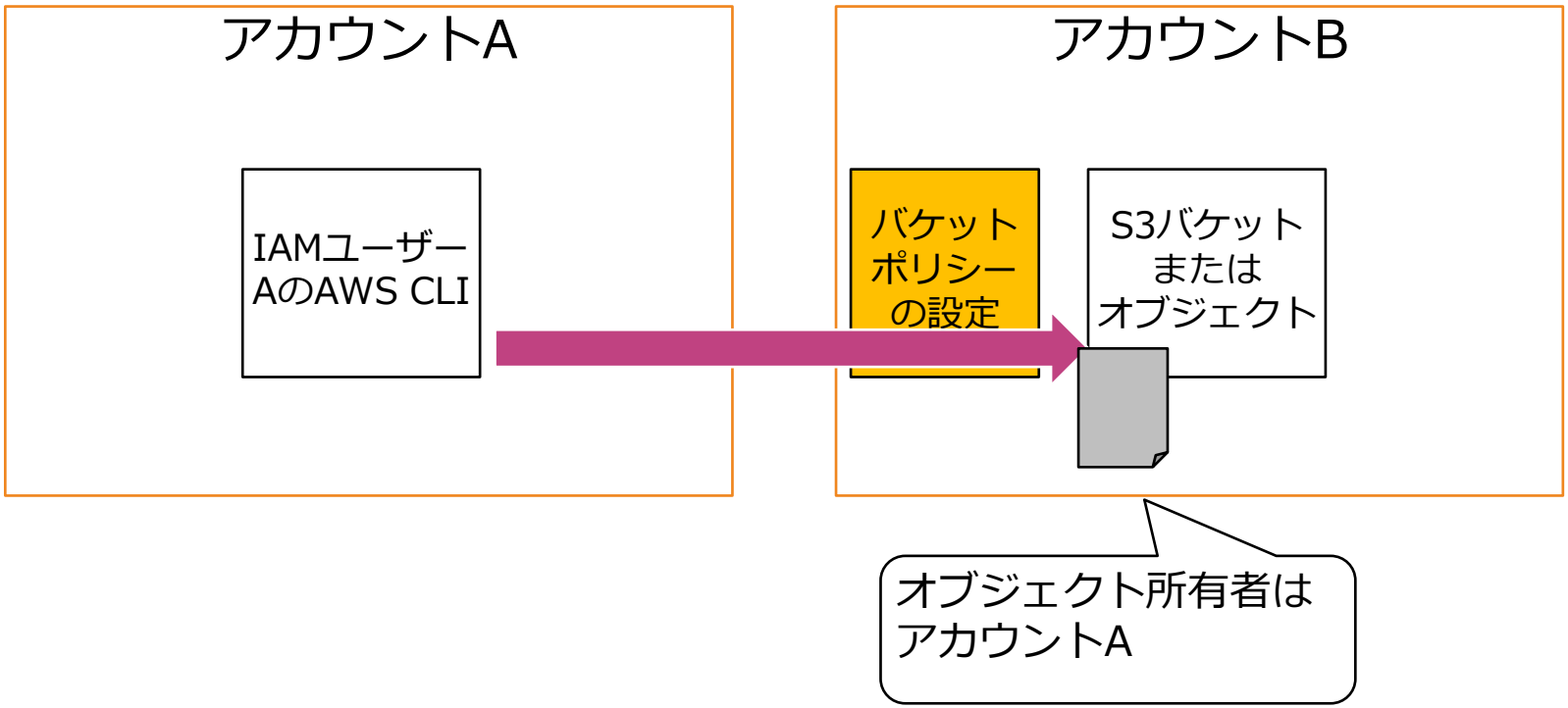
# オブジェクト所有者

## オブジェクト所有者にそのオブジェクトの権限管理を実施する権限がある

- アップロードしたユーザーがオブジェクト所有者になる。
- デフォルト設定では、オブジェクトにはオブジェクト所有者のみがアクセスできる
- オブジェクトの所有者は、アクセス制御で他のユーザにアクセスできるように変更する権限を有する。
- 署名付き URLを作成し、一時的な認証情報を付加することにより、ユーザに期限つきのURLを発行できる

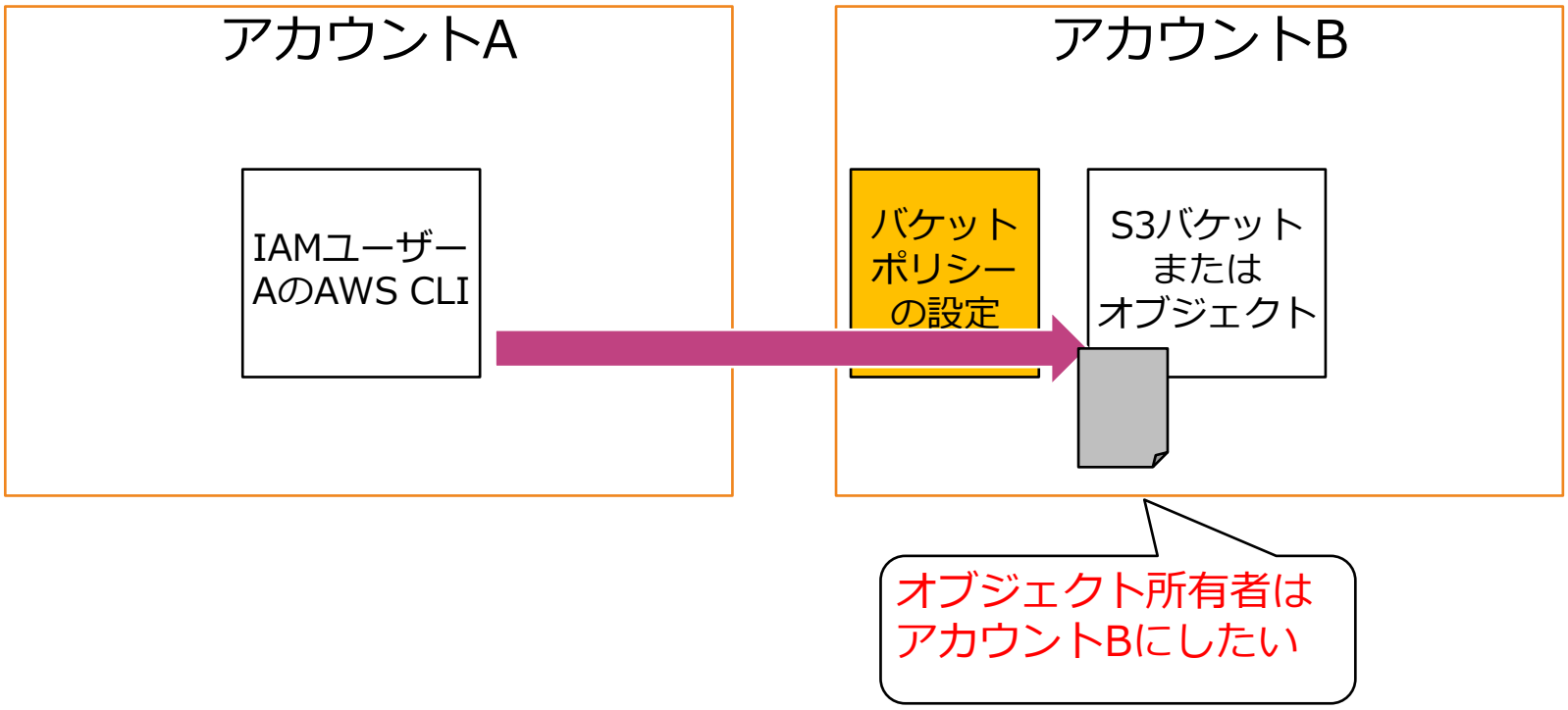
# オブジェクト所有者

アカウントBのバケットにお関わらず、アカウントAがアップロードしたオブジェクトに権限がないことになる。



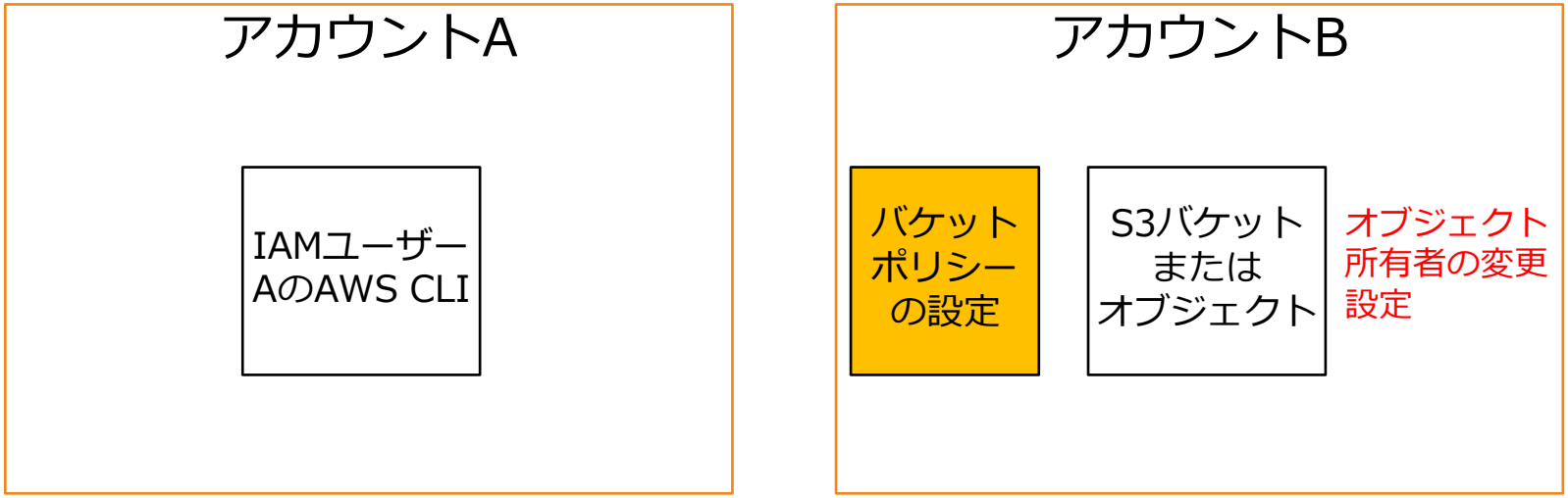
# オブジェクト所有者

アカウントBのバケットにお関わらず、アカウントAがアップロードしたオブジェクトに権限がないことになる。



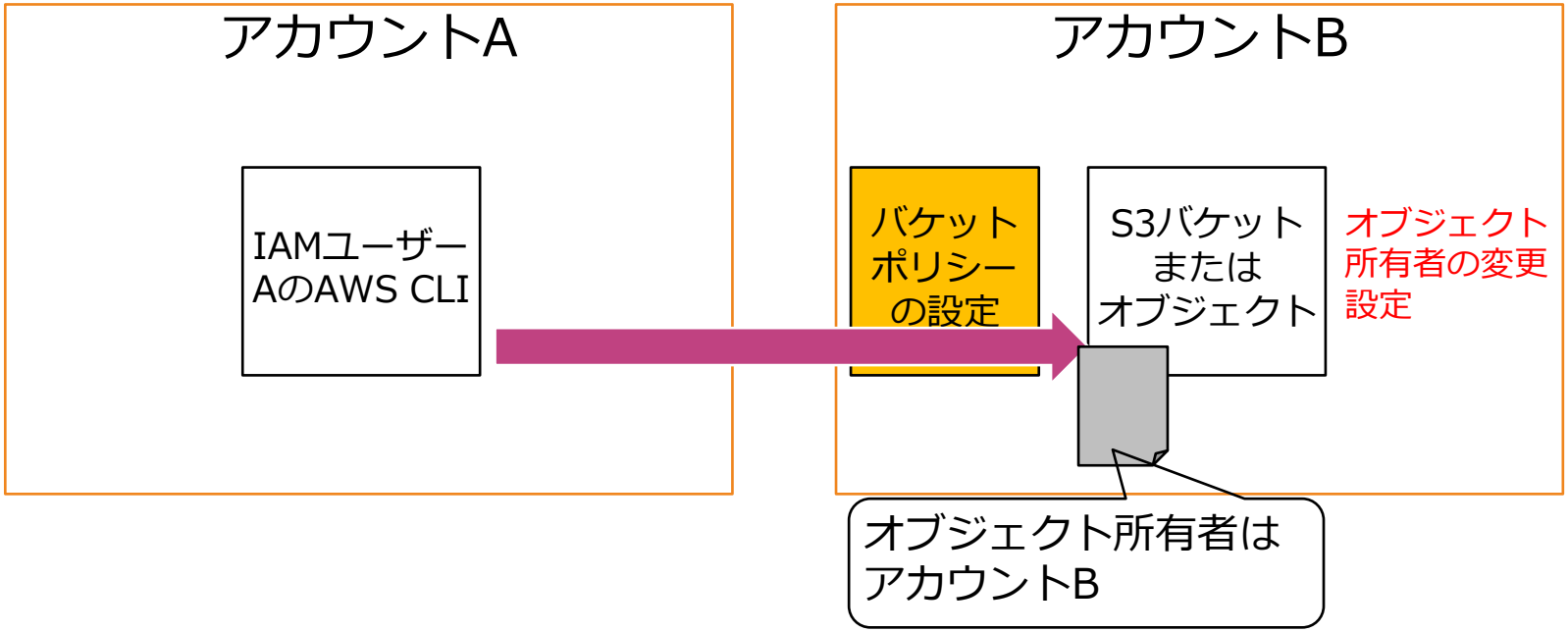
# このハンズオンで実施する内容

1. オブジェクト所有者をアカウントBに変更する



# このハンズオンで実施する内容

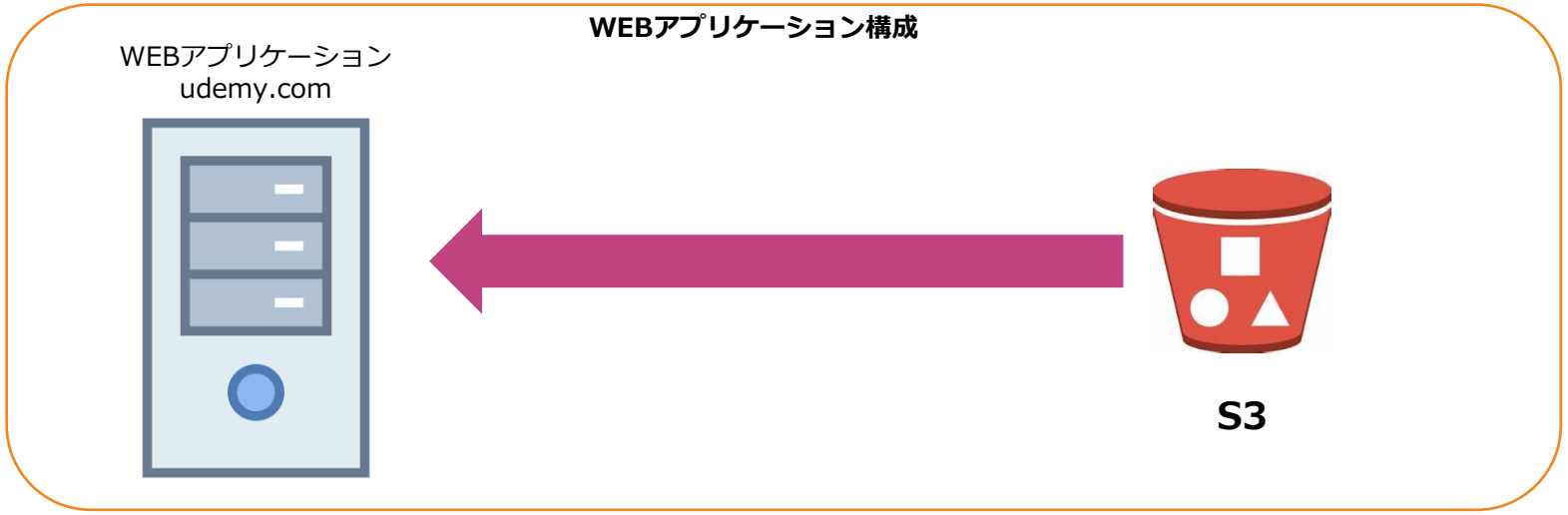
1. オブジェクト所有者をアカウントBに変更する
2. オブジェクト所有権を変更するACLを付与しながら、
   アカウントAのIAMユーザーAからオブジェクトの
   アップロードを実施する。



# クロスオリジンリソースシェアリング（CORS）の設定

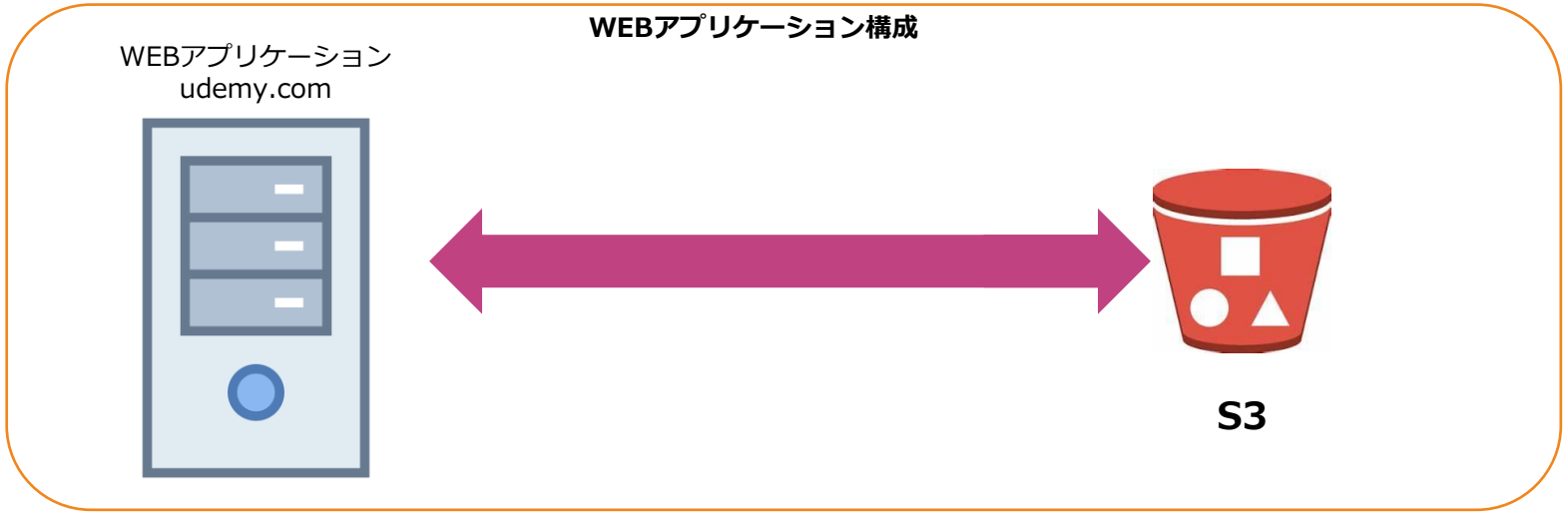
# クロスオリジンリソースシェアリング(CORS)

１つのS3を利用したWEBアプリケーションのドメインから、別のドメインが利用するS3リソースを相互利用する機能



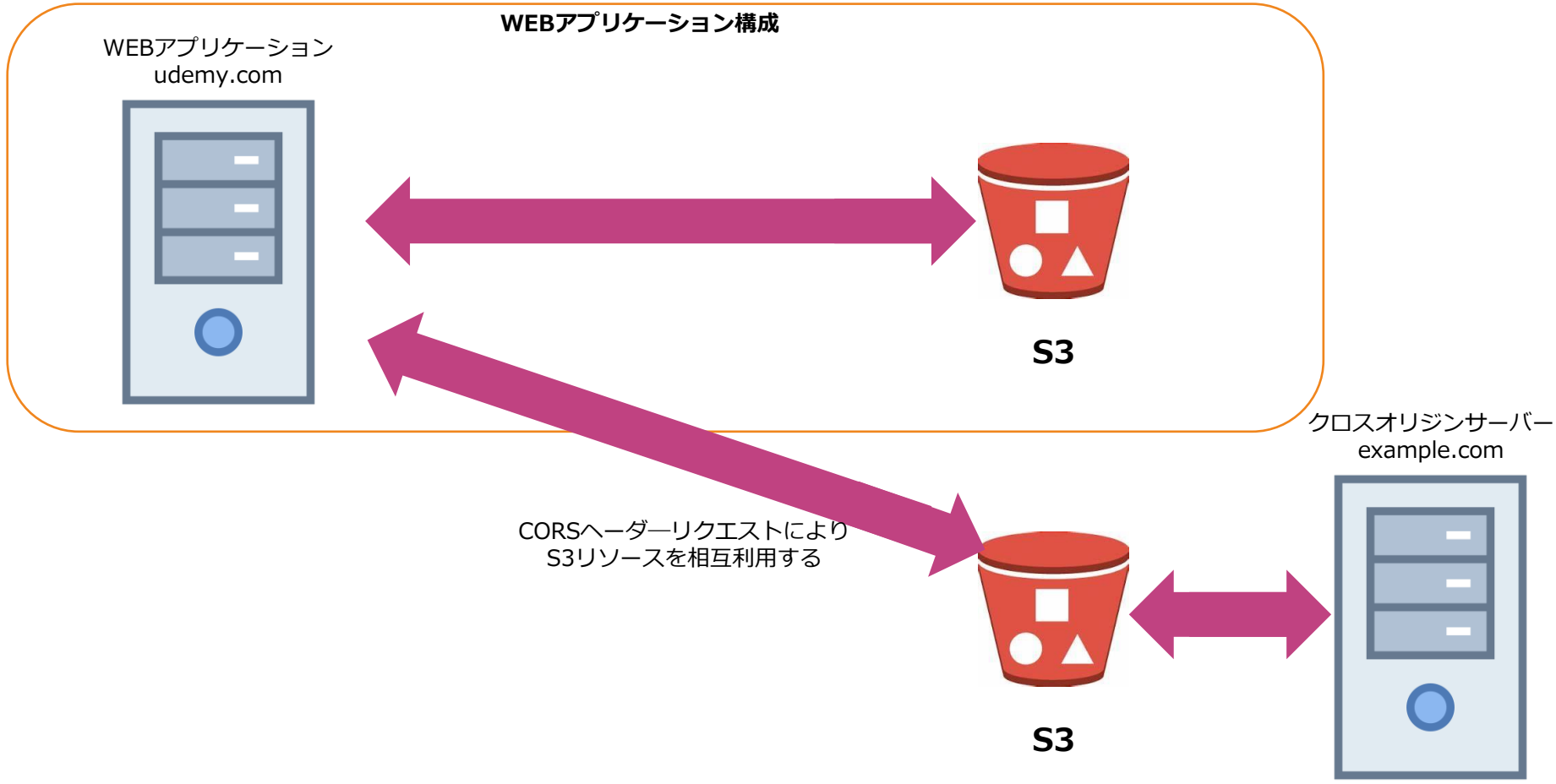
# クロスオリジンリソースシェアリング(CORS)

１つのS3を利用したWEBアプリケーションのドメインから、別のドメインが利用するS3リソースを相互利用する機能



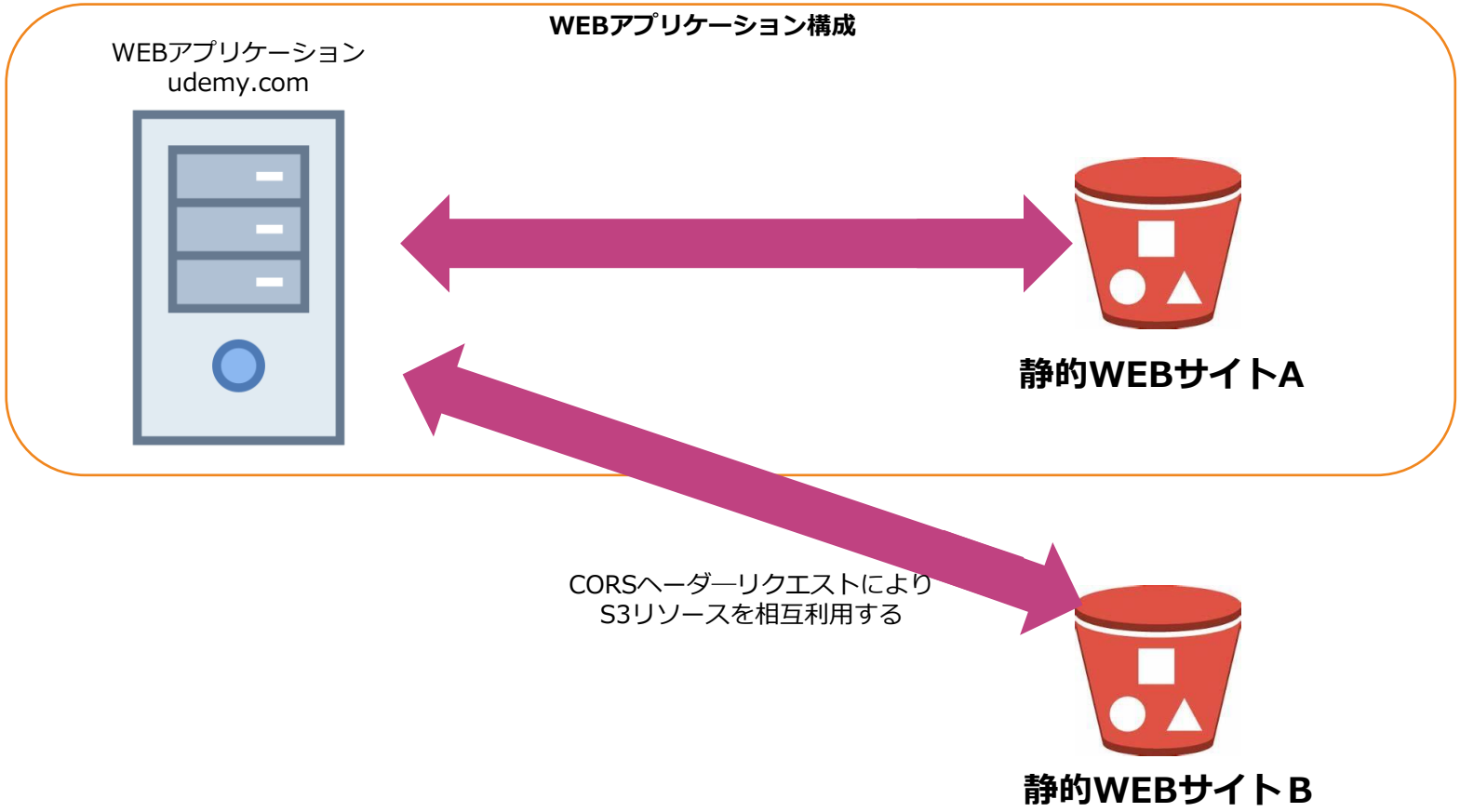

# クロスオリジンリソースシェアリング(CORS)

１つのS3を利用したWEBアプリケーションのドメインから、別のドメインが利用するS3リソースを相互利用する機能



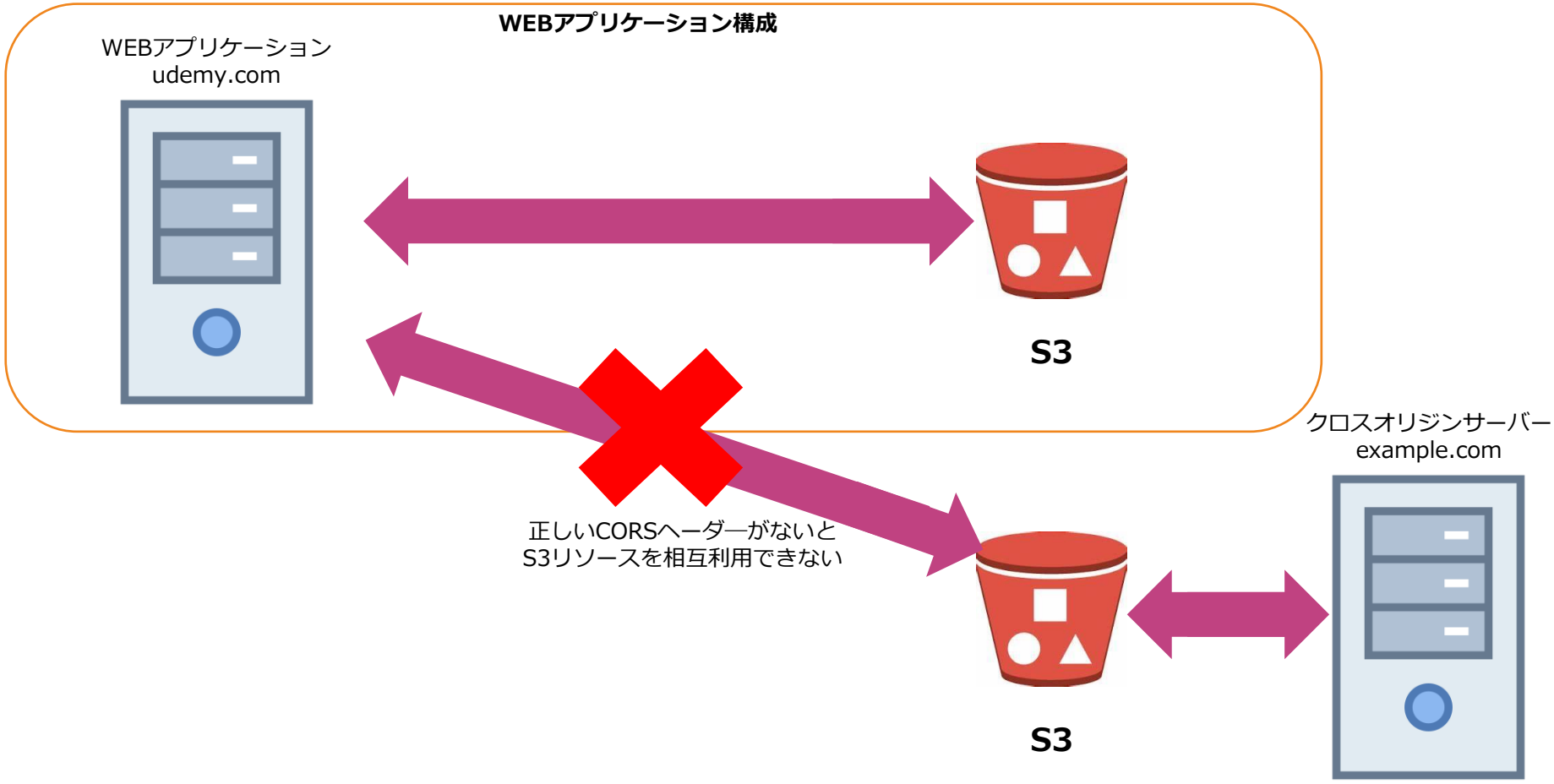
# クロスオリジンリソースシェアリング(CORS)

例えば、２つの静的WEBホスティングされたS3バケットリソースを１つのWEBサイトとしてオリジンサーバーで表示させる



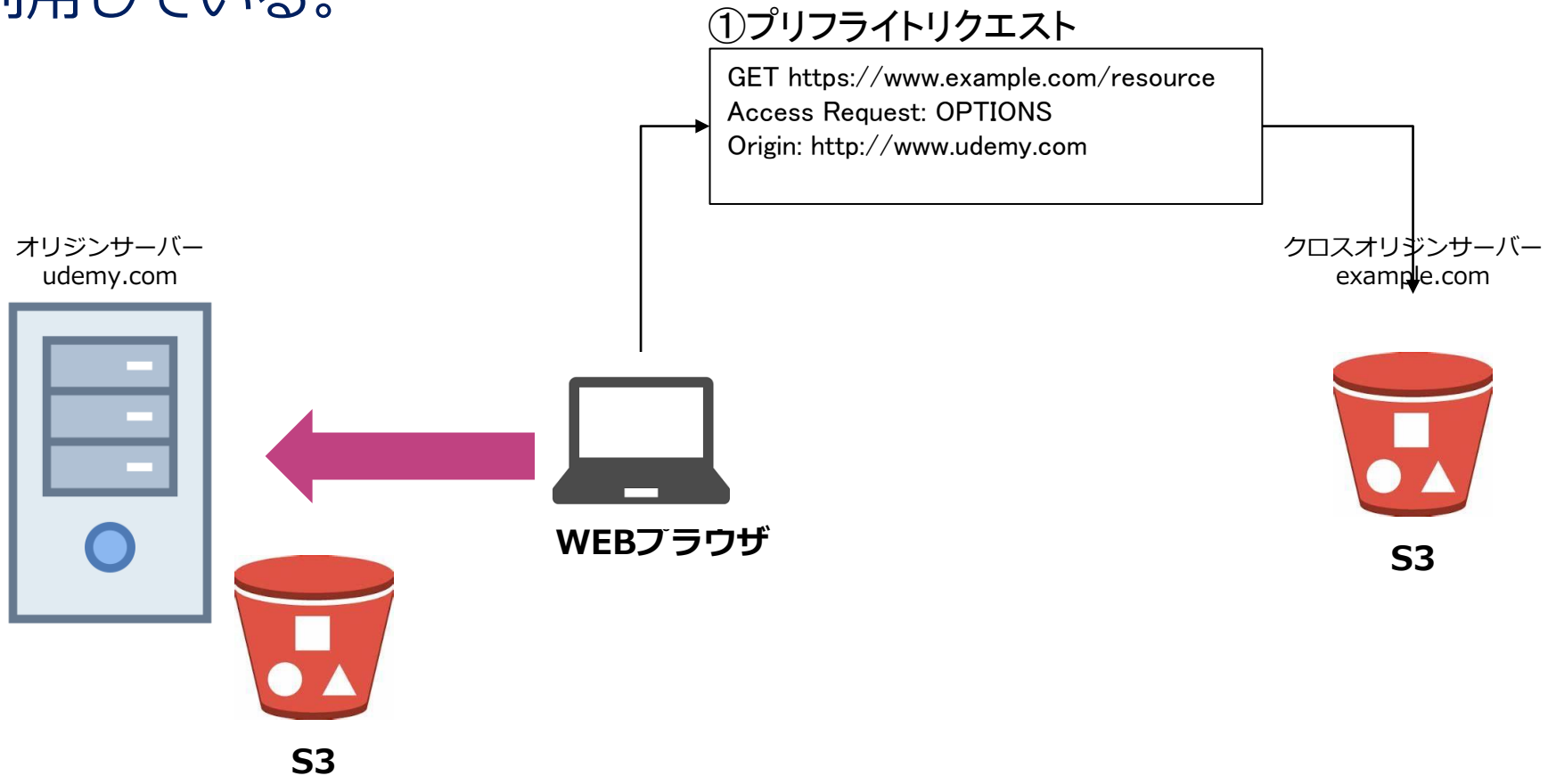
# クロスオリジンリソースシェアリング(CORS)

## 正しいCORSヘーダーを利用しない他のオリジンからのリクエストは実行されない。



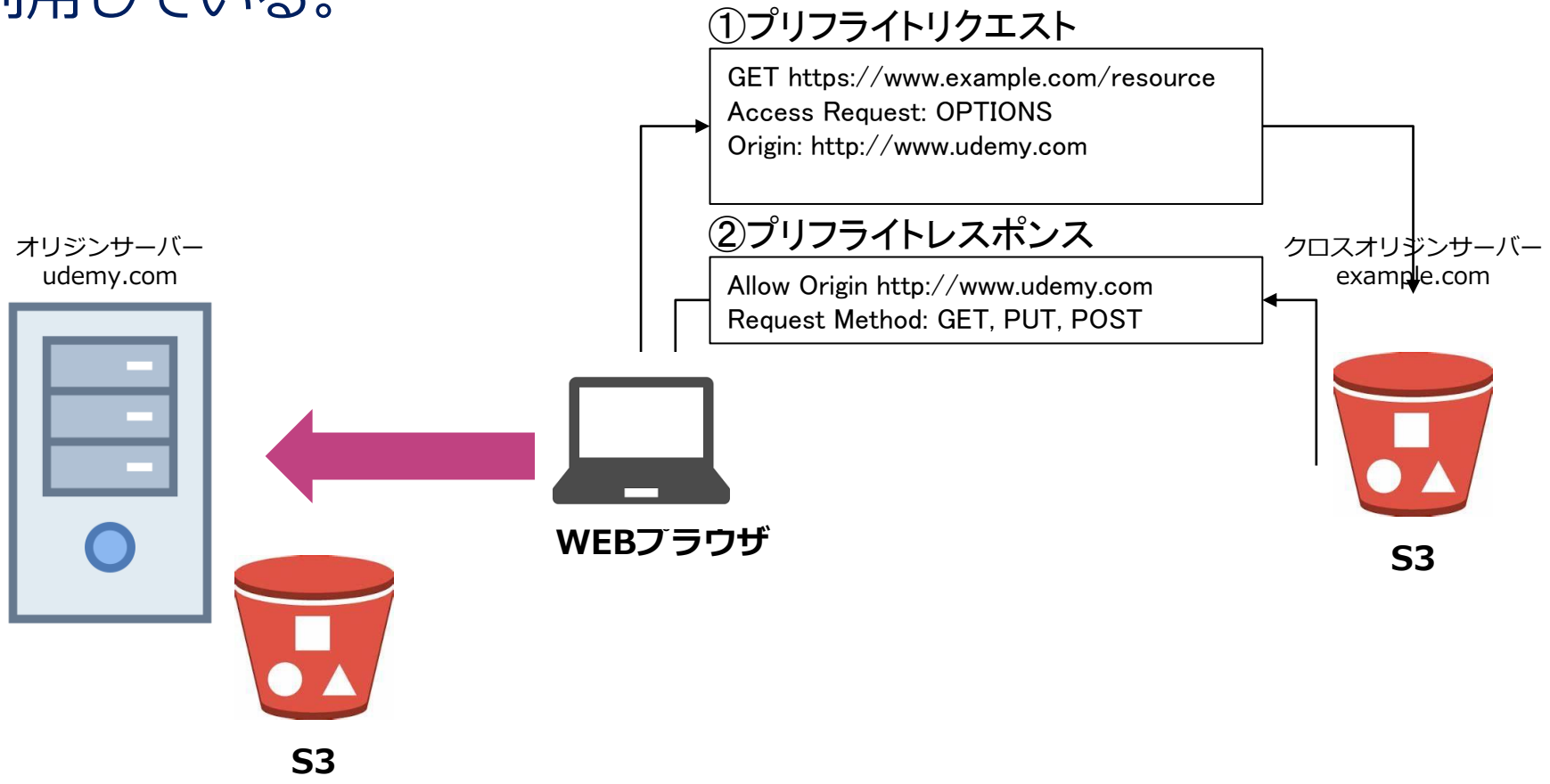
# クロスオリジンリソースシェアリング(CORS)

udemy.comが静的WEBサイトとしてS3バケットのリソースを利用している。



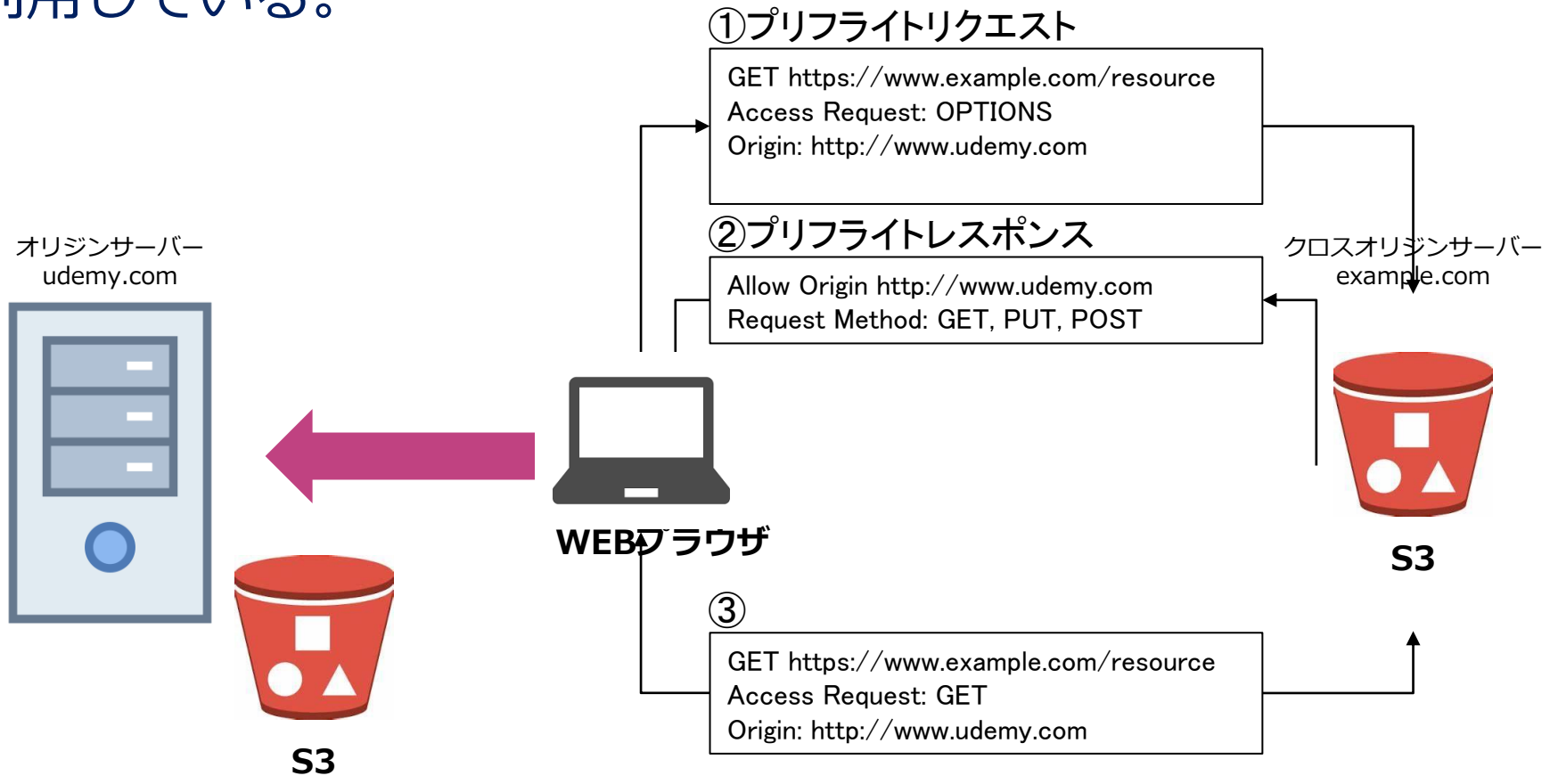
# クロスオリジンリソースシェアリング(CORS)

udemy.comが静的WEBサイトとしてS3バケットのリソースを利用している。



# クロスオリジンリソースシェアリング(CORS)

udemy.comが静的WEBサイトとしてS3バケットのリソースを利用している。



# S3の利用状況分析

# S3バケットの利用状況分析

## S3バケットの利用状況を分析することで、効率的な利用方法や安全な利用がされているかを支援する。

| 分析機能 | 特徴 |
| --- | --- |
| **S3サーバーアクセスログ** | ✓ バケットに対するリクエストの詳細が記録される。<br>✓ アクセスのログ情報は、セキュリティやアクセスの監査に利用できる。 |
| **S3アクセスアナライザー** | ✓ S3 バケットに対して任意のユーザーや他の AWS アカウント (組織外の AWS アカウントを含む) にアクセスを許可が適切になされているかを評価・警告する。 |
| **S3ストレージクラス分析** | ✓ データのアクセス頻度の分析<br>✓ バケット内の一定期間のストレージアクセスパターンを分析し、標準ストレージに格納されたデータを標準IAストレージクラスに移行すべきかを判断できる。 |
| **Amazon S3 Storage Lens** | ✓ データの使用状況とアクティビティの分析<br>✓ バケットの使用状況とアクティビティのメトリクスを集計することができる。 |

# サーバーアクセスログ

S3にアクセスした際のログを取得することが可能。バケットとプレフィックスをターゲットに設定する。

# S3アクセスアナライザー

## S3のアクセス状況がアクセスポリシーに沿っているか確認し、不正なアクセスの有無を監視する

**S3アクセスアナライザーの特徴**

- ✓ IAM アクセスアナライザーと連動したS3向けの機能
- ✓ バケットポリシー／ACLに沿ってポリシー違反がないかをモニタリング
- ✓ パブリックバケットまたは共有バケットアクセスを解析して、その解析結果を表示する
- ✓ バケットアクセスのソースを検証する場合は、列の情報を使用して、迅速で正確な是正措置を実行する

# ストレージクラス分析

## S3バケットの利用状況を確認して最適な利用設定を支援する

### S3のストレージクラス分析の特徴

- ✓ バケット内のデータのアクセスパターンを確認することができる
- ✓ レポートは日次でCSV形式で出力される。
- ✓ 最初のレポートの抽出には24時間～48時間を要する。
- ✓ アクセス頻度の低いデータや保存期間を確認して、ライフサイクルルール設定に活かすことができる。
- ✓ 主に標準ストレージから標準IAに移行するべきオブジェクトを確認することが可能となる。

# S3 Storage Lens

## バケットの使用状況とアクティビティのメトリクスを集計することができる。

**S3 Storage Lensの特徴**

- ✓ バケットの使用状況とアクティビティのメトリクスを集計することができる。
- ✓ 組織全体でストレージがどれくらい利用できるのか、最も急成長しているバケットとプレフィックスが何であるか、などのインサイトを抽出できる。
- ✓ Amazon S3 コンソールバケットページのダッシュボードではインサイトと傾向を可視化できる。
- ✓ CSV または Parquet 形式でレポートをダウンロードできる。
- ✓ ストレージコストの最適化や、データ保護のベストプラクティスに関するレコメンデーション事項が提示される。

# S3データの分析

# S3データの解析

## S3内のデータ検索・解析には用途に応じて複数サービスから選択が可能

| 分析サービス | 特徴 |
| --- | --- |
| **S3 Select（Glacier Select）** | ✓ S3の内部機能として有している検索機能で、S3内で直接にクエリを実行し、データをフィルタリングして取得できる<br>✓ GZIP圧縮データやCSVやJSONに対して実行可能 |
| **Amazon Athena** | ✓ Amazon S3 内のデータを直接、簡単に分析できるようにするインタラクティブなクエリサービス<br>✓ Athena SQL クエリで SageMaker 機械学習モデルを呼び出し、機械学習による推論も実行可能 |
| **Amazon Macie** | ✓ 機械学習によりAmazon S3 の機密データを検出、分類、保護する、フルマネージド型サービス<br>✓ 機密データ検出や調査を実施する |
| **Amazon Redshift Spectrum** | ✓ Amazon S3の格納データに対して、Amazon Redshiftから直接クエリを実行出来る機能<br>✓ Redshiftクラスターが起動されている前提であるため、Redshiftを利用している場合にお勧め |

# S3 Select

## SQLクエリを実行してS3バケット内のデータをフィルタリングして、抽出することができる。

Python

```python
import boto3

s3 = boto3.client('s3')

resp = s3.select_object_content(
    Bucket='s3select-demo',
    Key='sample_data.csv',
    ExpressionType='SQL',
    Expression="SELECT * FROM s3object s where s.\"Name\" = 'Jane'",
    InputSerialization = {'CSV': {"FileHeaderInfo": "Use"}, 'CompressionType': 'NONE'},
    OutputSerialization = {'CSV': {}},
)

for event in resp['Payload']:
    if 'Records' in event:
        records = event['Records']['Payload'].decode('utf-8')
        print(records)
    elif 'Stats' in event:
        statsDetails = event['Stats']['Details']
        print("Stats details bytesScanned: ")
        print(statsDetails['BytesScanned'])
        print("Stats details bytesProcessed: ")
        print(statsDetails['BytesProcessed'])
        print("Stats details bytesReturned: ")
        print(statsDetails['BytesReturned'])
```

Bash

```
python jane.py
```

以下の出力が得られます。

```
Jane,(949) 555-6704,Chicago,Developer

Stats details bytesScanned:
326
Stats details bytesProcessed:
326
Stats details BytesReturned:
38
```

【参照】https://aws.amazon.com/jp/blogs/news/querying-data-without-servers-or-databases-using-amazon-s3-select/

# Amazon Athena

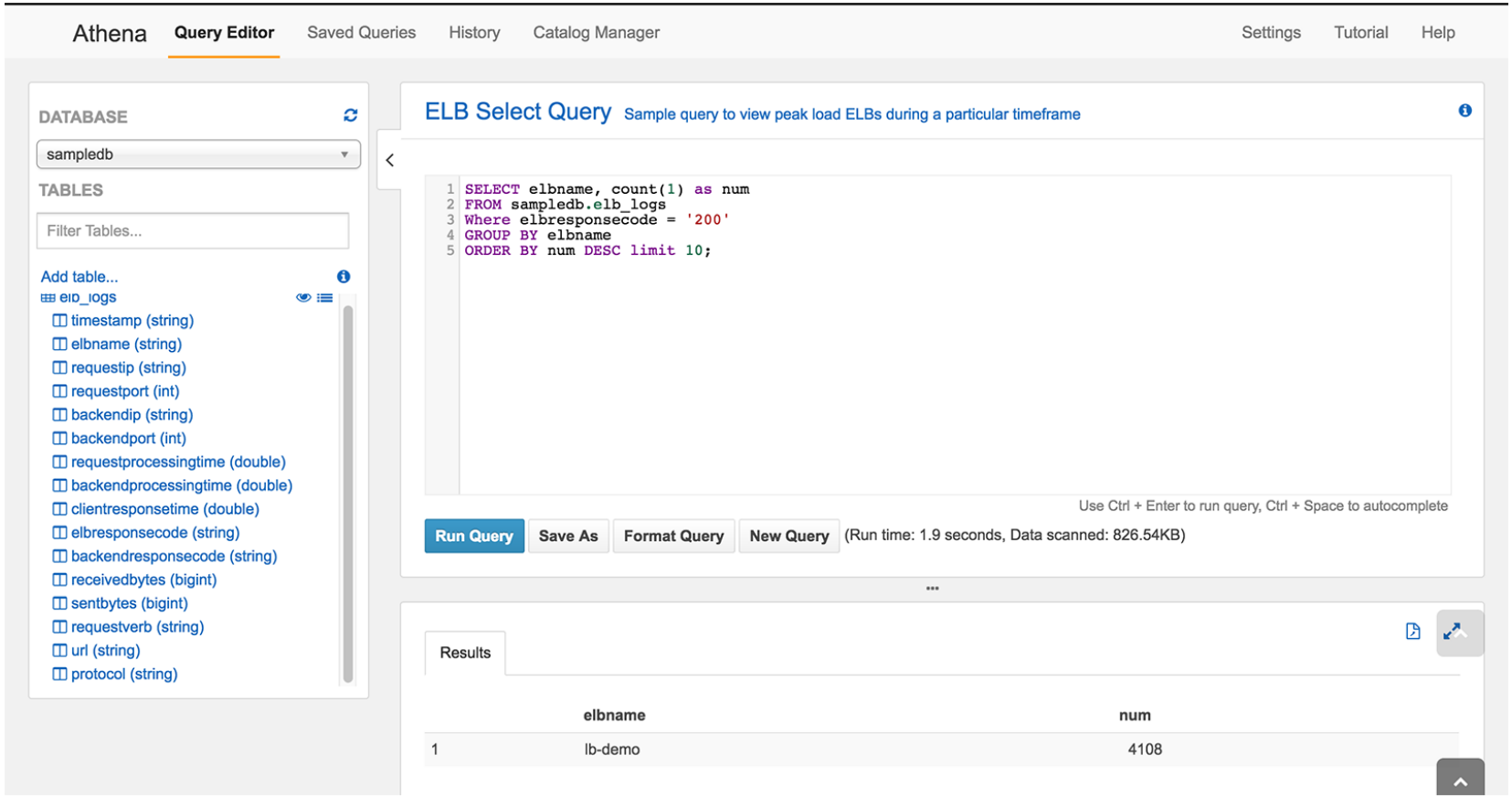
## Amazon S3データを標準 SQL を使用して簡単に分析するサーバレスサービス

**Athenaの特徴**

- ✓ S3バケット内のデータを指定し、スキーマを定義し、標準的な SQL を使用してクエリの実行することができる。
- ✓ 実行したクエリに対してのみ料金が発生する。
- ✓ AWS Glueデータカタログと連携して、さまざまなサービスにわたるメタデータの統合リポジトリを作成できる。
- ✓ データソースのクロールとスキーマの解析、新規および修正したテーブル定義とパーティション定義のカタログへの入力、スキーマのバージョニング保持が可能

# Amazon Athena

## Amazon S3データを標準 SQL を使用して簡単に分析するサーバレスサービス



# Amazon Macie

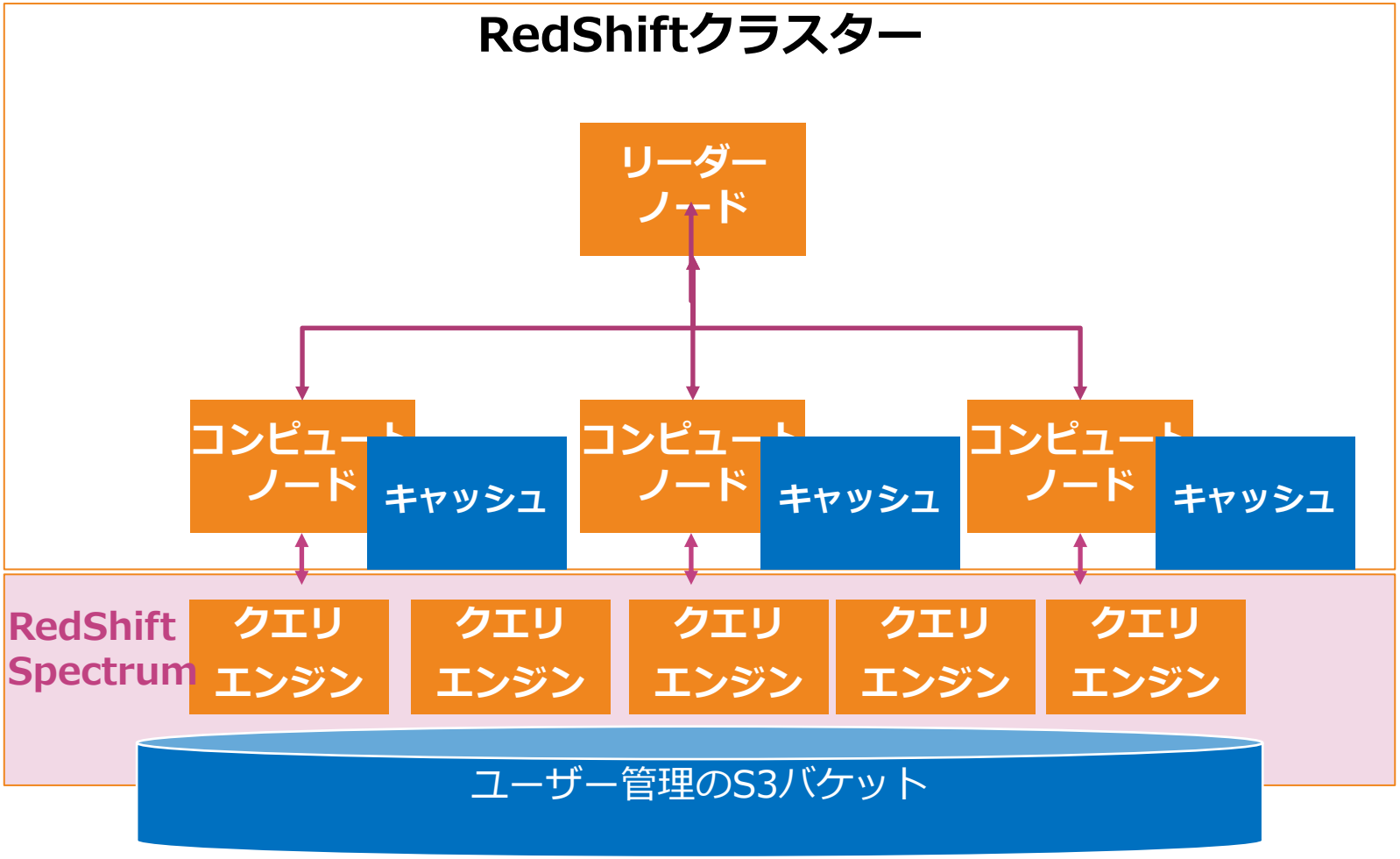
## 機械学習とパターンマッチングを使用して S3バケット内の機密データを検出して保護する

### Amazon Macieの特徴

- ✓ 大規模なデータセット内にある機密データの検出を自動化し、データ保護のコストを削減する。
- ✓ AWS アカウントと共有されているバケットのリストを含む、Amazon S3 バケットのインベントリが自動的に提供される。
- ✓ S3バケットに機械学習とパターンマッチング手法を適用して、個人識別情報 (PII) などの機密データを特定してアラートを発信する。

# Redshift Spectrum

RedShift Spectrumにより、ユーザーが管理するS3バケットに対して直接データ解析を実行可能



# Glacierの概要

# Amazon S3 Glacier

バックアップなど中長期保存用のS3よりも安価なストレージ

**S3と同じ耐久性で**
**値段が安い！**

**データ取得などの**
**迅速性がない！**

# Glacierの特徴

## バックアップなど中長期保存用のS3よりも安価なストレージ

- ✓ Amazon S3 Glacier では、データは「アーカイブ」に保存される
- ✓ 1 つのアーカイブの最大サイズは 40 TB
- ✓ 保存可能なアーカイブ数とデータ量に制限なし
- ✓ 各アーカイブには作成時に一意のアーカイブ ID が割り当てられ、作成後はアーカイブを更新できない。
- ✓ アーカイブを保存するためのコンテナとして「ボールト」を使用（1 つの AWS アカウントでは、最大 1,000 個のボールトを使用）
- ✓ Amazon S3 のライフサイクルルールと連携させることにより、Amazon S3 データのアーカイブを自動化し、全体的なストレージコストを削減
- ✓ Advanced Encryption Standard (AES) 256 ビット対称鍵を使用してデフォルトで自動的に暗号化
- ✓ S3と違って直接データをアップロード・取得という処理ができないため、S3ライフサイクル管理からか、プログラム処理によるアップロード／ダウンロードが必要
- ✓ Glacierの最低保持期間は90日

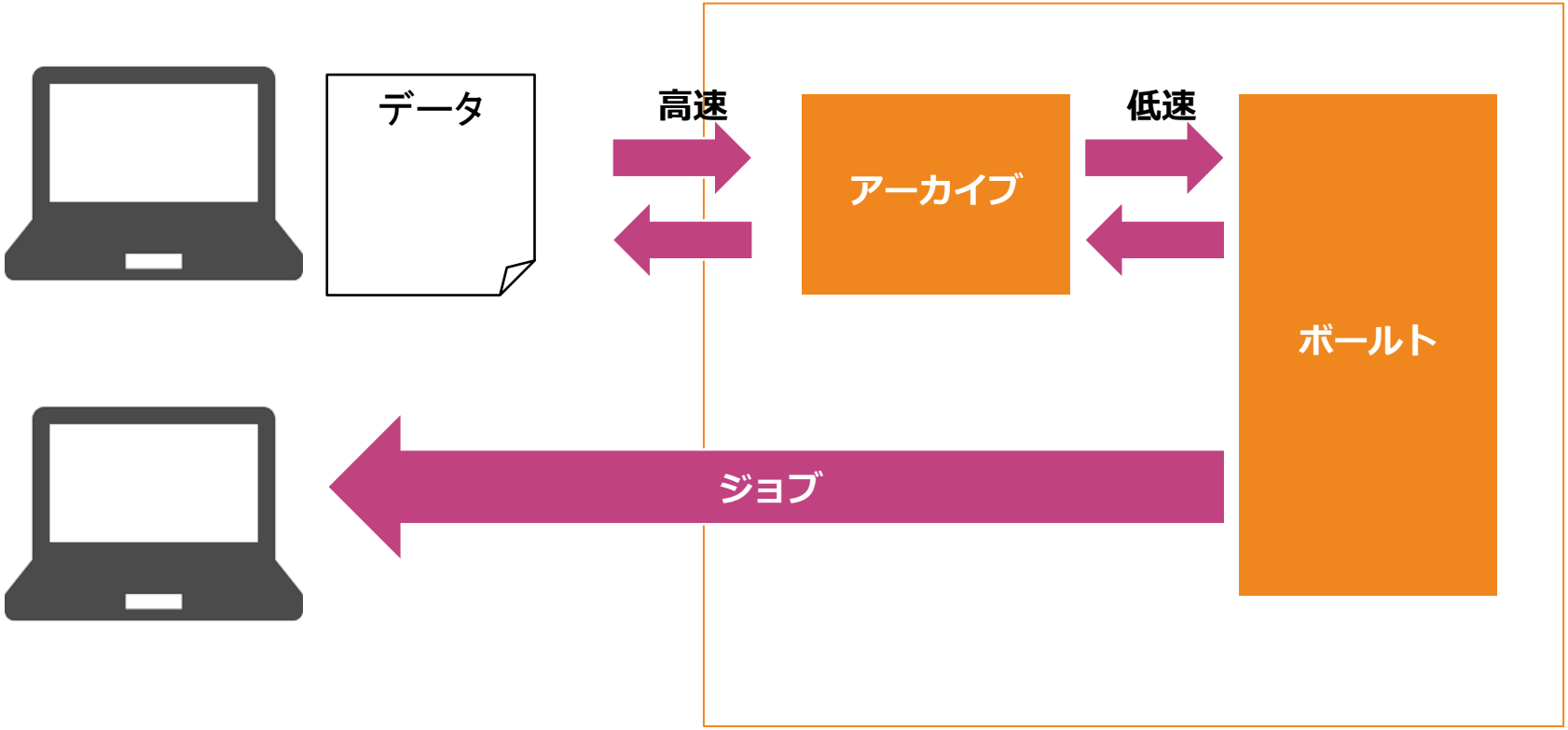
# Glacierの仕組み

## S3と異なり、ボールトとアーカイブという単位でデータを保存

| 管理方式 | 特徴 |
| --- | --- |
| **ボールト** | ✓ ボールトはアーカイブを格納するコンテナ<br>✓ ボールトはリージョンに作成 |
| **アーカイブ** | ✓ アーカイブは、写真、動画、ドキュメントなどの任意のデータで、S3 Glacier でのストレージの基本単位<br>✓ 各アーカイブは一意のアドレスを持ちます。 |
| **ジョブ** | ✓ アーカイブに SELECT クエリを実行したり、アーカイブを取得したり、ボールトのインベントリを取得したりする実行単位 |
| **通知設定** | ✓ ジョブの完了には時間がかかるため、ジョブの完了時にSNSと連携した通知設定が可能 |

# Glacierの仕組み

アーカイブに一時的にデータをアーカイブ処理して、ボールトに長期保存するという仕組み



# Glacierのデータ取出タイプ

## Glacierのデータ取得タイプの設定に応じてデータ取得時間と取得時の料金が変わる。

| タイプ | 特徴 |
| --- | --- |
| **迅速** | ✓ 迅速取り出しでは、アーカイブのサブセットが迅速に必要になった場合にデータにすばやくアクセスするモード。通常 1～5 分以内で使用可能になる |
| **プロビジョニングキャパシティ** | ✓ プロビジョンドキャパシティーは、迅速取り出しの取得容量を必要なときに利用できることを保証する仕組み |
| **標準** | ✓ 標準取り出しでは、数時間以内にすべてのアーカイブにアクセスできるデフォルト設定。通常、標準取り出しは 3～5 時間で完了 |
| **大容量** | ✓ 大容量取り出しは、最も安価な取り出しオプションであり、大量のデータ (ペタバイトのデータを含む) を 1 日以内に低コストで取得できます。通常、大容量取り出しは 5～12 時間で完了 |

# Glacierのデータ取出タイプ

## Glacierのデータ取得タイプの設定に応じてデータ取得時間と取得時の料金が変わる。

| ストレージクラスまたは階層 | 迅速 | Standard | 大容量 |
| --- | --- | --- | --- |
| S3 Glacier Flexible Retrieval または S3 Intelligent-Tiering Archive Access。 | 1～5分 | 3～5 時間 | 5～12時間 |
| S3 Glacier Deep Archive または S3 Intelligent-Tiering Archive アクセス | 利用不可 | 12 時間以内 | 48 時間以内 |

# ストレージクラスの選択

## Glacierでは３つのストレージタイプから選択する。

| タイプ | 特徴 | 性能 |
| --- | --- | --- |
| **S3 Glacier Flexible Retrieval（通常のGlacier）** | ✓ 1 年に 1～2 回アクセスされ、非同期で取り出されるアーカイブデータ向け<br>✓ 通常のデータ検索で（3~5時間）を要する<br>✓ 迅速取り出しで(2～5分)で取り出し可能<br>✓ 一括検索で（5～12時間）で無料でデータ取り出し<br>✓ ライフサイクルマネジメントで指定<br>✓ ボールトロック機能でデータを保持 | ■耐久性<br>99.999999999%<br>■可用性<br>99.99% |
| **S3 Glacier Instant Retrieval** | ✓ アクセスされることがほとんどなく、ミリ秒単位の取り出しが必要な長期間有効なデータ向け<br>✓ 医用画像やニュースメディアなど<br>✓ S3 Standard と同じパフォーマンスのミリ秒単位でのデータの取り出し | ■耐久性<br>99.999999999%<br>■可用性<br>99.9% |
| **Amazon Glacier Deep Archive** | ✓ 最安のアーカイブ用ストレージ<br>✓ 7～10 年以上保持される長期間使用されるものの、めったにアクセスされないデータ向け<br>✓ 標準の取り出し速度で12 時間以内にデータを取得<br>✓ 大容量取り出しで48時間以内にデータを取得<br>✓ ライフサイクル管理で指定 | ■耐久性<br>99.999999999%<br>■可用性<br>99.99% |

# アクセス管理

## Glacierのアクセス管理は用途に応じて方式を使い分ける

| 管理方式 | 特徴 |
| --- | --- |
| **IAMポリシー** | ✓ IAMユーザーやリソースに対してS3サービスへのアクセス権限を設定する<br>✓ 一元的にリソースへのアクセス権限を管理 |
| **ボールトポリシー** | ✓ ボールトで直接アクセスポリシーを定義して、組織内のユーザーや社外ユーザーに対してもボールトへのアクセス権を付与 |
| **データ取り出しポリシー** | ✓ データ取り出しに関する制限を定義<br>✓ [無料利用枠のみ] に制限。または無料利用枠を超える量を取り出したい場合は、[最大取得率] を指定すると、取り出し速度を制限して、取り出しコストの上限を設定 |
| **ボールトロックポリシー** | ✓ ロックによって変更を禁止することにより、コンプライアンス管理を強力に実施することが可能 |
| **署名** | ✓ 認証保護のために、全リクエストに署名が必要 |

# Amazon Glacierの料金

## バックアップなど中長期保存用のS3よりも安価なストレージ

| | |
|---|---|
| **容量当たりの料金** | GB/月 あたり 0.005USD（0.5円ほど）<br>→S3は標準で0.025USD／One zoneで0.0152USD/GB |
| **データ取り出し料金** | 迅速：0.033USD/GB<br>標準：0.011USD/GB<br>大容量：0.00275USD/GB |
| **データ取り出しリクエスト料金** | 迅速：11.00USD/リクエスト 1,000 件<br>標準：0.0571USD/リクエスト 1,000 件<br>大容量：0.0275USD/リクエスト 1,000 件 |
| **プロビジョニングされた迅速取り出し** | 110.00USD/プロビジョンド容量単位 |
| **データ転送料金** | データ転送 (イン)は無料<br>インターネットへのデータ転送 (アウト)は1 GB/月まで無料。それ以上は有料 |

※2020年7月あたりのお値段です。値段は変動する可能性があります。