



[AWS Black Belt Online Seminar]

Amazon S3 / Amazon S3 Glacier

サービスカットシリーズ

アマゾンウェブサービスジャパン株式会社

ソリューションアーキテクト

焼尾 徹

2019/2/20

自己紹介

焼尾 徹

技術統括本部 レディネス&テックソリューション本部
ソリューション アーキテクト



普段の業務

個別相談会のお客様を技術的にサポート
場所にとらわれない働き方の模索

好きなAWSの取り組み

Amazon Wind Farm



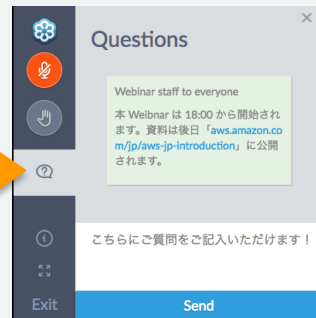
AWS Black Belt Online Seminar とは

「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2019年02月19日時点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が東京リージョンを使用する場合、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.

本日の内容

- Amazon S3の位置付け
- Amazon S3の概要
- Amazon S3へのアクセス
- Amazon S3のデータ保護
- Amazon S3の管理
- Amazon S3パフォーマンス最適化
- Amazon S3の料金
- まとめ

Amazon S3の位置付け

Amazon S3の概要

Amazon S3へのアクセス

Amazon S3のデータ保護

Amazon S3のデータ管理

Amazon S3パフォーマンス最適化

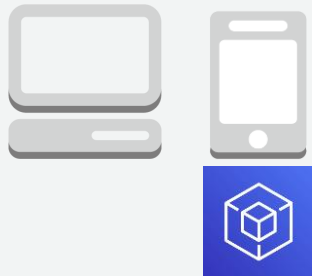
Amazon S3の料金

Amazon S3の位置付け

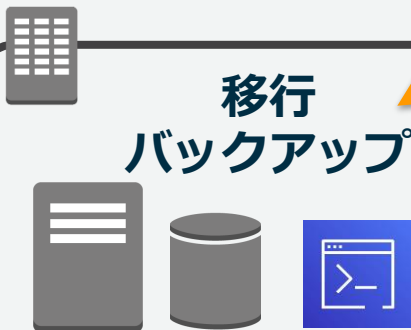
Amazon S3の位置付け



Web / モバイル
アプリケーション



移行
バックアップ

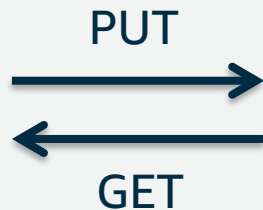


Amazon S3とは

Amazon Simple Storage Service (S3)は、ユーザがデータを安全に、容量制限なく、データ保存が可能な、クラウド時代のオブジェクトストレージです。



S3 API

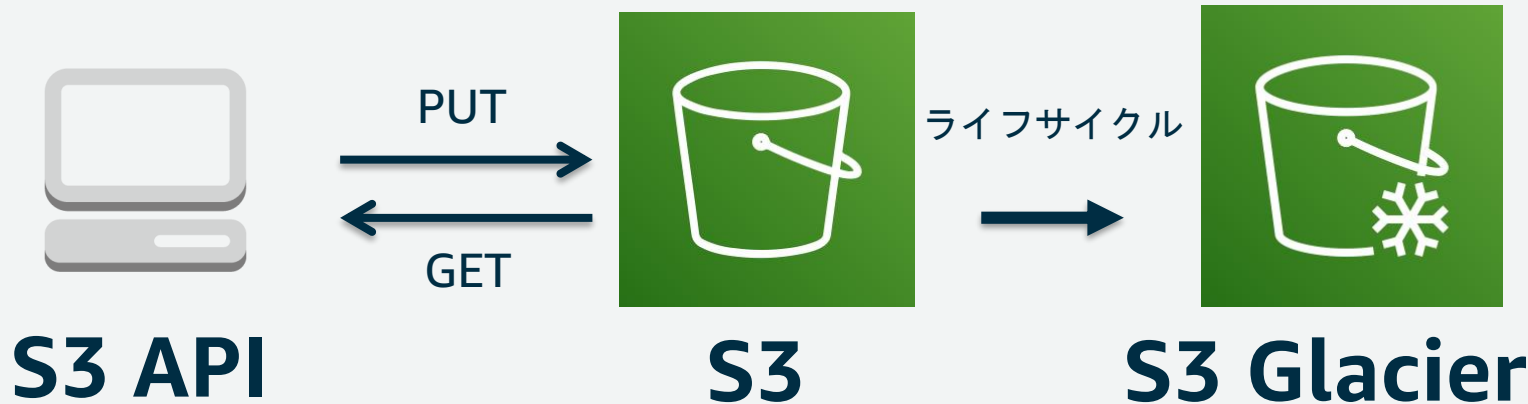


S3

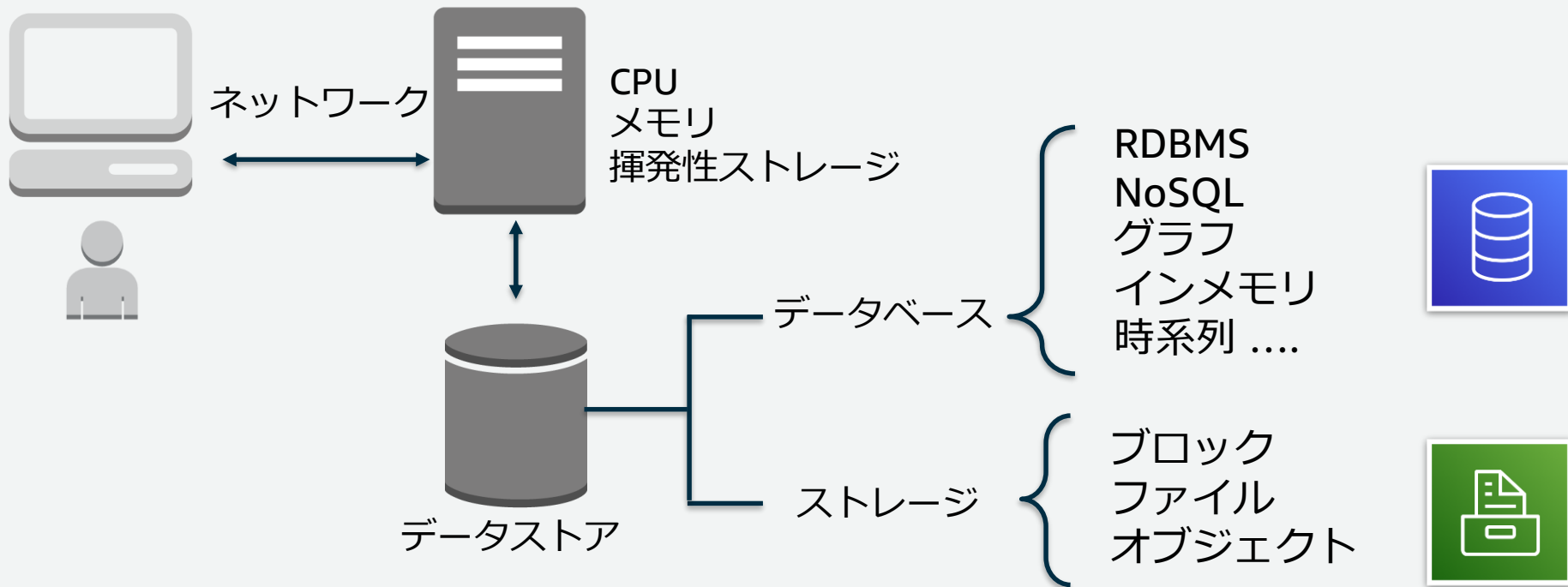
Amazon S3 及び Amazon S3 Glacier

Amazon Simple Storage Service (S3)は、ユーザがデータを安全に、容量制限なく、データ保存が可能な、クラウド時代のオブジェクトストレージです。

Amazon S3 Glacier は、安全性とコスト効率を重視したアーカイブ向けストレージです。



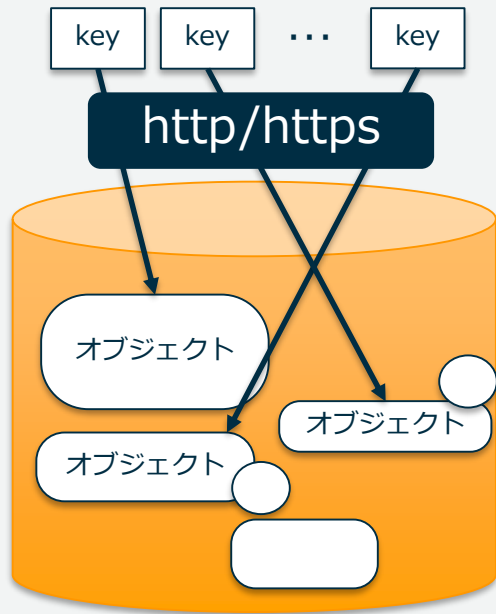
データストアの選択シーン



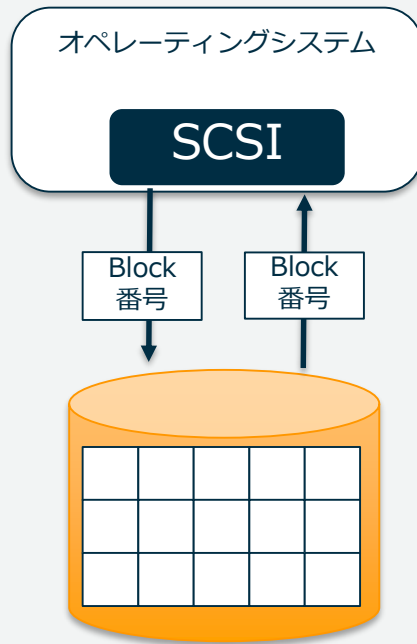
オブジェクトストレージの特徴

オブジェクトストレージ

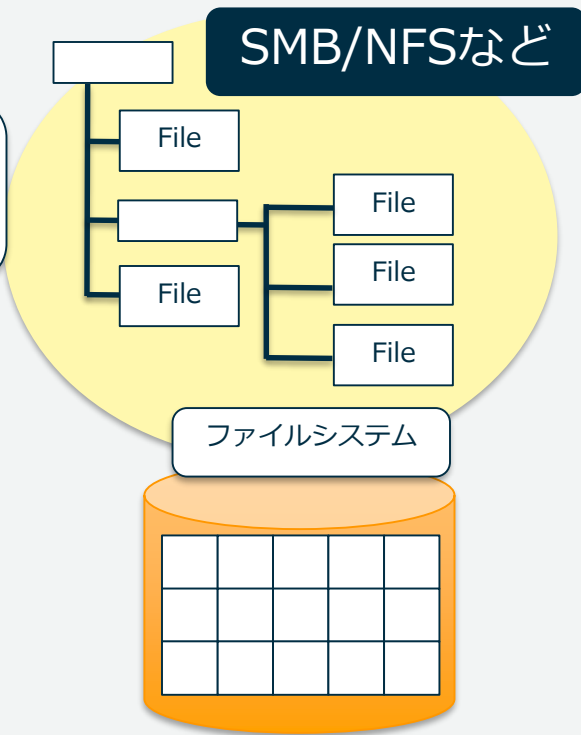
オブジェクト、それに付随するメタデータ、そのオブジェクトにアクセスするためのユニークなIDで構成されるデータの倉庫



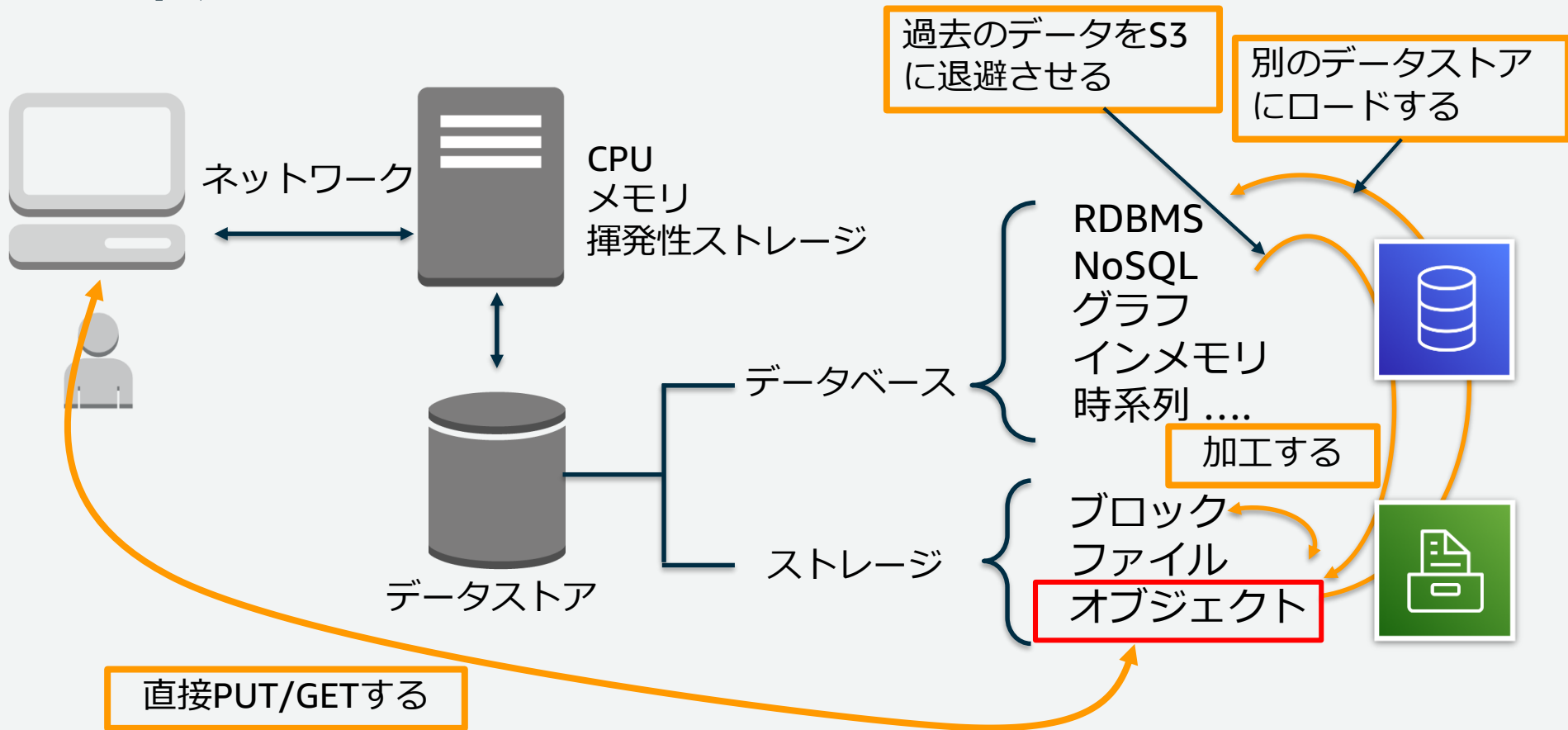
ブロックストレージ



ファイルストレージ



S3 利用シーン



ユースケース <https://aws.amazon.com/jp/blogs/news/webinar-bb-s3-usecase-2018/>

Amazon S3の位置付け

Amazon S3の概要

Amazon S3へのアクセス

Amazon S3のデータ保護

Amazon S3のデータ管理

Amazon S3パフォーマンス最適化

Amazon S3の料金

Amazon S3の概要

Amazon S3 特徴

- **容量無制限**

- 1ファイル最大5TBまで

- **高い耐久性**

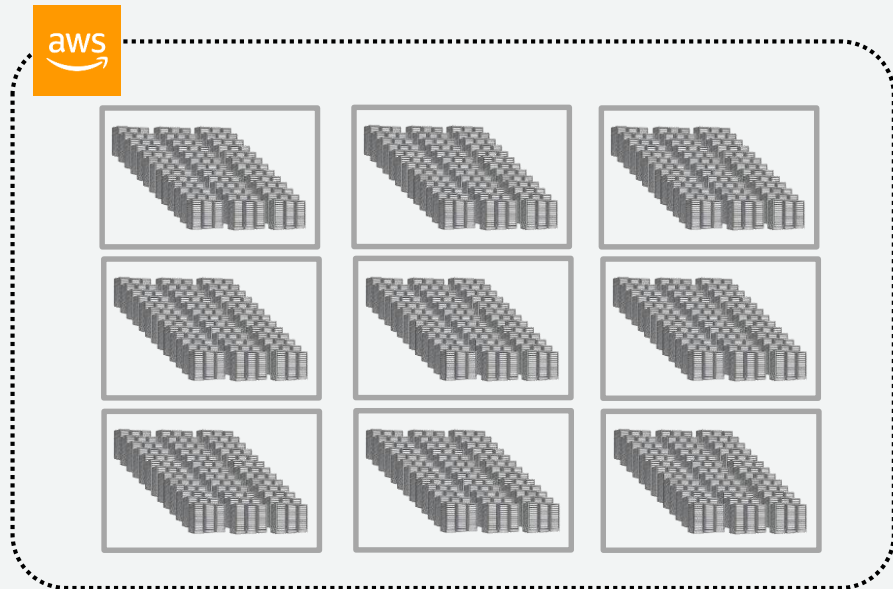
- 99.9999999999%

- **安価なストレージ**

- 容量単価:月額1GB / 約3円※

- **スケラブルで安定した性能**

- データ容量に依存しない性能（ユーザが、サーバ台数、媒体本数やRAID、RAIDコントローラを考慮する必要がない）



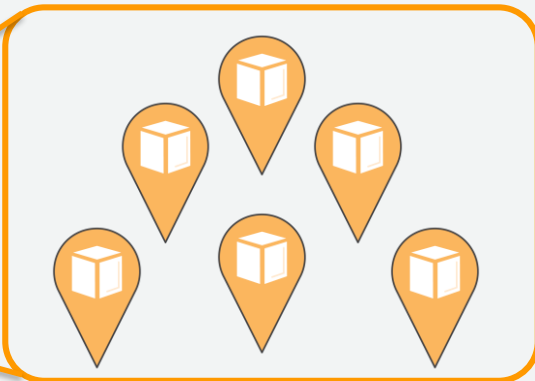
※2019年2月 <https://aws.amazon.com/jp/s3/pricing/>
東京リージョンにおけるスタンダードが、US\$0.025/GB

S3 とリージョン、アベイラビリティゾーン(AZ)



S3 標準は少なくとも3つのAvailability Zones(AZs)にデータを格納する

物理的に離れている - つまり、万一災害が起きても、1つのAZへの影響しかない



1つのAZは最大8つのデータセンターで構成

データセンター間、AZ間は低遅延のプライベートネットワークで接続されている



1つのデータセンターのダウン、または、1つのAZのダウンは、S3としての可用性に影響しない

Amazon S3の耐久性 99.999999999%

Amazon S3 用語



S3 バケット

バケット

- オブジェクトの保存場所。各AWSアカウントにてデフォルト100個まで作成可能。**名前はグローバルでユニークな必要あり**。上限緩和申請で100以上も利用可能に。

オブジェクト

- データ本体。S3に格納されるファイルでURLが付与される。バケット内オブジェクト数は無制限。1オブジェクトサイズは0から5TBまで(1つのPUTでアップロード可能なオブジェクトの最大サイズは5GB)。

キー

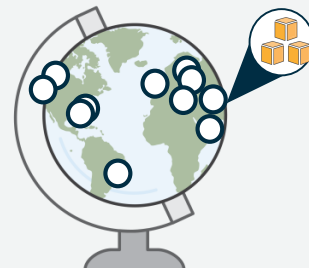
- オブジェクトの格納URLパス。「バケット+キー+バージョン」が必ず一意になる。

メタデータ

- オブジェクトに付随する属性の情報。システム定義メタデータ、ユーザ定義メタデータがある。

リージョン

- バケットを配置するAWSのロケーション。目的のアプリケーションと同じリージョンであると有利。

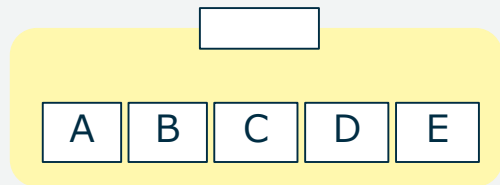


Amazon S3 の概要 - オブジェクトのネーミングスキーマ

オブジェクトはバケット内にフラットに格納される。
キーのパス指定でフォルダ階層のように表示も可能。「/」を区切り記号として、マネジメントコンソールでは、フォルダ構造を表現する。

バケット名

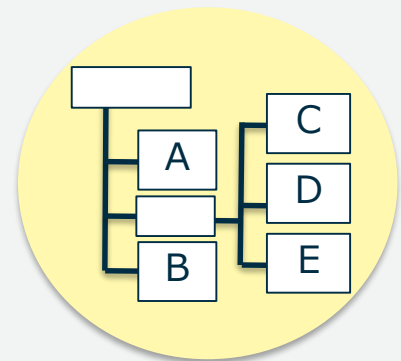
オブジェクトのキー名



S3: //ExampleAWSbucket/Logistics/packing-list.pdf

プレフィックス(Prefix)

オブジェクト名



(*) 2018.3月 バケット命名規則について、米国東部リージョンでも適用されるようになり、これで全てのリージョンにてDNS命名規則に沿って、命名する必要があります。

Amazon S3 の概要 - Data Consistency モデル

Amazon S3はデータを複数の場所に複製することで高い可用性を実現するため、データの更新・削除にはEventual Consistency Readモデル（結果整合性）が採用されている。

オペレーション	Consistencyモデル	挙動
新規登録 (New PUTs)	Consistency Read(*)	登録後、即時データが参照できる
更新 (Overwrite PUTs)	Eventual Consistency Read(結果整合性)	更新直後は、以前のデータが参照される可能性がある
削除 (DELETE)	Eventual Consistency Read (結果整合性)	削除直後は、削除前のデータが参照される可能性がある

- 同じオブジェクトへの複数同時書き込み制御のためのロック処理は行われず、最新のタイムスタンプのリクエストが優先される。
- （ロック処理があるような仕組みと比べて）読み込みの待ち時間が小さくなるのがメリット

(*) 2015.8月 new putについて、read-after-write consistencyがUS Standard regionでもサポートされるようになり、全てのリージョンにてread-after-write-consistencyとなりました。

Amazon S3 の概要 – ストレージクラス

用途に応じて、オブジェクトを格納するS3の場所の使い分け

ストレージクラス	特徴	耐久性（設計上）	可用性（設計上）
STANDARD (スタンダード)	複数AZにデータを複製。デフォルトのストレージクラス。	99.999999999%	99.99%
STANDARD-IA (標準低頻度アクセスストレージ)	スタンダードに比べ格納コストが安価。いつでもアクセス可能だが、データの読み出し容量に対して課金。IAはInfrequent Accessの略。	99.999999999%	99.9%
INTELLIGENT_TIERING	アクセス頻度が高いオブジェクトと低いオブジェクトを自動的に最適化するストレージクラス	99.999999999%	99.9%
ONEZONE_IA (1ゾーン-低頻度アクセスストレージ)	Single AZにデータを格納するが、複製の考え方はスタンダード、STANDARD-IAと同じ。ただし、地震や洪水などの大災害による1アベイラビリティゾーンの物理的な損失には耐性はありません。	99.999999999%	99.5%
S3 Glacier (アーカイブ)	低コストだが、データの取り出しにコストと時間を要する。ライフサイクルマネジメントにて指定する。	99.999999999%	99.99% Object復元後
S3 Glacier Deep Archive (予定)	もっとも低コストなコールドストレージ。取り出しには半日から2-3日がかかる。	99.999999999%	(未定)
低冗長化ストレージ(RRS)	RRS はReduced Redundancy Storageの略。Glacierから取り出したデータの置き場所として利用。	99.99%	99.99%

New 2018.11月
New 2018.4月



Amazon S3 の操作



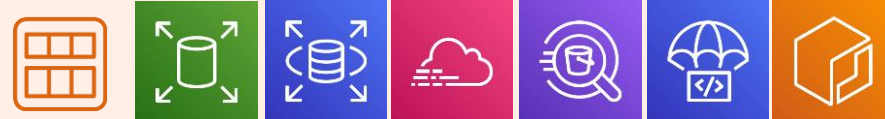
オペレーション	処理	特徴
GET	S3から任意のファイルをダウンロード	<ul style="list-style-type: none">RANGE GETに対応。S3 Glacierにアーカイブされ、RestoreされていないオブジェクトへのGETリクエストはエラー
PUT	S3に対してファイルをアップロード (新規・更新)	<ul style="list-style-type: none">シングルPUTオペレーションでは最大5GBまで、Multipart Uploadを利用すると5TBまで格納可能。
LIST	S3バケット内のオブジェクト一覧を取得	<ul style="list-style-type: none">Prefixによるパス指定での取得一覧のフィルタリングが可能。1回のリクエストでは1,000オブジェクトまで情報を取得可能。それ以上の場合は再帰的にリクエストを実施する必要がある
COPY	S3内でオブジェクトの複製を作成	<ul style="list-style-type: none">シングルCOPYオペレーションでは最大5GBまで、Multipart Uploadを利用すると5TBまでのファイルの複製が可能
DELETE	S3から任意のファイルを削除	<ul style="list-style-type: none">シングルDELETEオペレーションで最大1,000個のオブジェクトを削除可能MFA(Multi Factor Authentication)と連携した削除制御が可能
HEAD	オブジェクトのメタデータを取得	<ul style="list-style-type: none">オブジェクトそのものをGETオペレーションで取得しなくても、メタデータだけを取得可能
RESTORE	アーカイブされたオブジェクトを一時的にS3に取り出す。またはアーカイブされたオブジェクトへSelectクエリが可能	<ul style="list-style-type: none">S3 Glacierからのデータの取り出し低冗長化ストレージに指定期間オブジェクトがコピーされ、その指定期間中、ダウンロードが可能になるGlacier Select によるデータの部分的な取り出し
SELECT	ファイルへのSelect クエリをかけられる	<ul style="list-style-type: none">S3 Selectによるデータの部分的な取り出し

New 2017.11月

New 2018.4月



S3 へのアクセス方法

操作		利用イメージ
アプリケーション開発	AWS SDK	<pre>PutObjectRequest putObjectRequest = new PutObjectRequest(bucketName, Key, file); PutObjectResult result = this.client.putObject(putObjectRequest)</pre> 
コマンドラインやシェル	AWS CLI	<pre>\$ aws s3 cp xxxx.mp4 s3://bucketName/ \$ aws s3api get-object --bucket-name <bucket-name> --key <prefix/file-name></pre>
手動、人間の操作	Management Console 3rdパーティツール	 <p>AWS Management Console 3rd Party Tools</p>
アプリケーションやAWSサービスでのS3利用	HTTPS AWS SDK	<p>そのアプリケーションやAWSサービスが透過的にS3を活用する</p>  <p>AMI EBS RDS CloudTrail Athena CodeDeploy ECR</p>

Amazon S3の位置付け

Amazon S3の概要

Amazon S3へのアクセス

Amazon S3のデータ保護

Amazon S3のデータ管理

Amazon S3パフォーマンス最適化

Amazon S3の料金

Amazon S3へのアクセス

アクセス管理

きめ細やかなバケットもしくはオブジェクトへのアクセス権の設定

デフォルトでは、S3のバケットやオブジェクトなどは全てプライベートアクセス権限 (Owner:作成したAWSアカウント)のみに設定

IAMユーザ、クロスアカウントユーザ、匿名アクセスなどバケット/オブジェクト単位で指定可能

- **ユーザポリシー**

- IAM Userに対して、S3やバケットへのアクセス権限を設定
- 複数バケットやS3以外のものも含めて一元的にユーザ権限を指定する場合など

- **バケットポリシー**

- S3バケット毎に、アクセス権限を指定
- クロスアカウントでのS3バケットアクセス権を付与する場合など

- **ACL**

- 各バケットおよびオブジェクトのアクセス権限を指定
- バケット単位やオブジェクト単位で簡易的に権限を付与する場合など

アクセス管理(続き) - ユーザーポリシー

ユーザポリシーサンプル

```
{
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:ListAllMyBuckets"
      ],
      "Resource": "arn:aws:s3:::*"
    },
    {
      "Effect": "Allow",
      "Action": [
        "s3:ListBucket", "s3:GetBucketLocation"
      ],
      "Resource": "arn:aws:s3:::examplebucket"
    },
    {
      "Effect": "Allow",
      "Action": [
        "s3:PutObject", "s3:GetObject", "s3:DeleteObject"
      ],
      "Resource": "arn:aws:s3:::examplebucket/*"
    }
  ]
}
```



AWS Identity and Access Management (IAM)

「AWSにおいて、このユーザは何ができるか？」

- IAMのアイデンティティベースのポリシー
- IAMの環境において、何らかの制御を行う目的
- 全てのAWSサービスに言えることでS3に限らない

ユーザポリシーを利用して、IAMユーザに対して任意のバケットへのアクセス権限を付与


その他サンプルは下記URLを参照

http://docs.aws.amazon.com/ja_jp/AmazonS3/latest/dev/example-policies-s3.html

アクセス管理(続き) - バケットポリシー


バケットポリシーサンプル

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AddPerm",
      "Effect": "Allow",
      "Principal": "*",
      "Action": ["s3:GetObject"],
      "Resource": ["arn:aws:s3:::examplebucket/*"]
    }
  ]
}
```



S3 バケットポリシー

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "IPAllow",
      "Effect": "Allow",
      "Principal": "*",
      "Action": "s3:*",
      "Resource": "arn:aws:s3:::examplebucket/*",
      "Condition": {
        "IpAddress": {"aws:SourceIp": "54.240.143.0/24"}
      }
    }
  ]
}
```



S3 バケットポリシー

「このS3リソースには誰がアクセスできるのか？」

- IAMのリソースベースのポリシー
- S3環境において、何らかの制御を行う目的
- Conditionを利用してIAM User、クロスアカウント、IPアドレス制限、HTTP Referrer制限、CloudFront, MFA制限なども指定可能

バケットポリシーを利用して、全てのユーザーに対して、任意のバケットへのGETリクエストを許可

バケットポリシーを利用して、任意のIPアドレスレンジからバケットへのアクセスを許可

その他サンプルは下記URLを参照

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

アクセス管理(続き) - アクセスコントロールリスト(ACL)

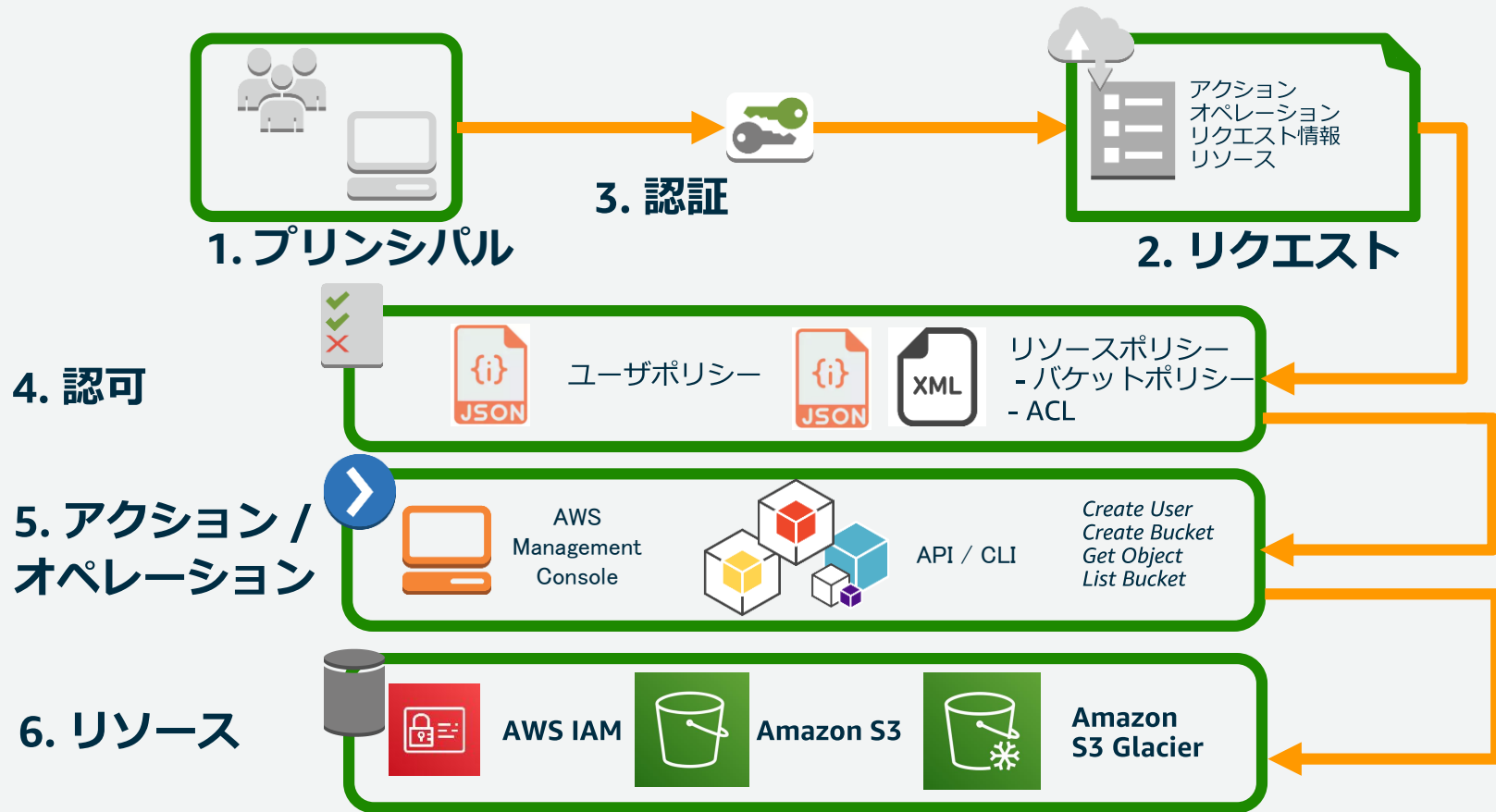
ACLはバケット単位のACLとオブジェクト単位のACLが存在

- バケットACLはバケット内のオブジェクトにも影響を与えるが、オブジェクトが個別にACLを設定している場合、オブジェクトACLが優先させる
- ACLよりも、**ユーザポリシーやバケットポリシーが優先**される
- 例えば、違うアカウントが所有するオブジェクトのアクセス許可を管理する場合に、オブジェクトACLが有用
- バケットACLを利用するのは、Amazon S3 のログ配信グループに、バケットのアクセスログオブジェクトの書き込みアクセス許可を付与する場合のみ



=> 通常は、バケットポリシーを用いましょう

S3へのアクセス、ここまでの整理



意図せずバケットがパブリックアクセスになるのを抑制する

- アクセス許可チェック、S3コンソールで公開アクセスが許可されたバケットが容易に分かるようなインジケータを表示する

2017.11月

バケット名	アクセス許可	リージョン	公開日時
aws-athena-query-results-ap-northeast-	オブジェクトは公開可能	アジアパシフィック (東京)	7月 20, 2018 9:44:57 午後 GMT+0900
aws-athena-query-results-us-east-1-9	オブジェクトは公開可能	米国東部 (バージニア北部)	10月 23, 2017 3:02:55 午後 GMT+0900
aws-glue-scripts ap-northeast-1	オブジェクトは公開可能	アジアパシフィック (東京)	10月 3, 2018 6:50:03 午後 GMT+0900
aws-glue-temporary- >northeast-1	オブジェクトは公開可能	アジアパシフィック (東京)	10月 3, 2018 6:50:03 午後 GMT+0900
aws-toryakio-logs	オブジェクトは公開可能 公開	アジアパシフィック (東京)	11月 20, 2017 6:07:40 午後 GMT+0900
backup-bucket-20170904	オブジェクトは公開可能	アジアパシフィック (東京)	1月 5, 2017 2:22:17 午後 GMT+0900
cf-templates-319f4yn57gef-ap-northeast-1	オブジェクトは公開可能	アジアパシフィック (東京)	2月 15, 2017 8:55:58 午前 GMT+0900
cf-templates-319f4yn57gef-ap-northeast-2	オブジェクトは公開可能	アジアパシフィック (ソウル)	2月 15, 2017 8:55:58 午前 GMT+0900
cf-templates-319f4yn57gef-us-east-1	オブジェクトは公開可能	米国東部 (バージニア北部)	11月 10, 2017 10:54:59 午前 GMT+0900
cloudtrail-awslog -isengard-do-not-delete	オブジェクトは公開可能	米国東部 (バージニア北部)	12月 2, 2016 5:22:25 午後 GMT+0900
cloudwatch-logs-export-toryakio	オブジェクトは公開可能	アジアパシフィック (東京)	11月 20, 2017 6:07:40 午後 GMT+0900
codebuild-ap-northeast-1-toryakio-input-bucket	オブジェクトは公開可能	アジアパシフィック (東京)	8月 18, 2017 1:15:55 午後 GMT+0900
codebuild-ap-northeast-1-toryakio-output-bucket	オブジェクトは公開可能	アジアパシフィック (東京)	8月 18, 2017 1:16:52 午後 GMT+0900

公開されている・
されていないがすぐ
わかる

↓
間違いがあれば
すぐ気付くことが
できる

- AWS Configにて、S3バケットがPublicにreadできたり、誰でも書き込みめるようになっていないかをチェックするマネージドルールを提供

2017.8月

<https://aws.amazon.com/jp/blogs/news/aws-config-update-new-managed-rules-to-secure-s3-buckets/>

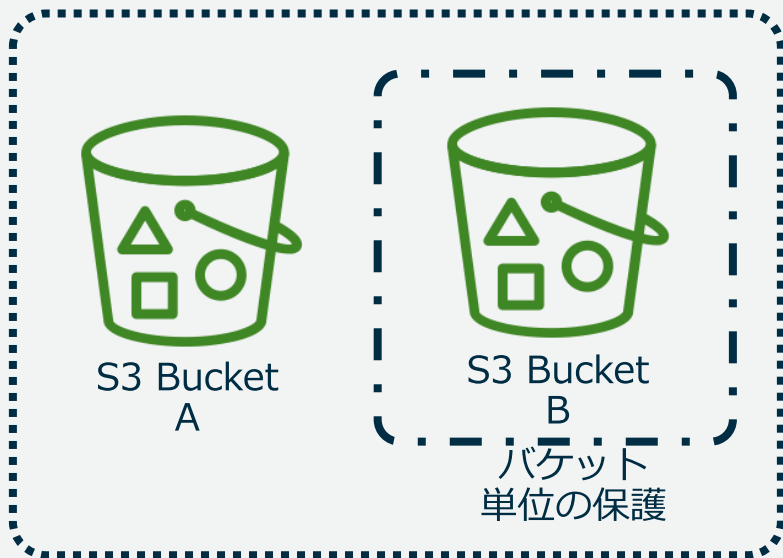


Amazon S3 Block Public Access

New 2018.11月

アカウントレベル、もしくはバケットレベルで「あらかじめ」意図せずバケットがパブリックアクセスになるのを抑制する

アカウント単位の保護



パブリックなアクセスを許すバケットポリシー

```
{  
  "Version": "2012-10-17",  
  "Id": "Policy1537...5299",  
  "Statement": [  
    {  
      "Sid": "Allow get object by any",  
      "Effect": "Allow",  
      "Principal": "*",  
      "Action": "s3:*",  
      "Resource": "arn:aws:s3:::ex-bucket/*"  
    }  
  ]  
}
```

- 新規で作成されるバケット、新規で作成されるアカウントにはデフォルトで適用（安全側に）

Amazon S3 Block Public Access (続き)

設定	Trueとした場合の効果	備考
BlockPublicAcls	パブリックなACL設定、パブリックなオブジェクトのアップロードをさせない	
IgnorePublicAcls	パブリックなACLの設定をしていても、それを無力化する	
BlockPublicPolicy	パブリックなバケットポリシーの設定をさせない	アカウントレベルで有効にするのが効果的。 AWS Organizations など
RestrictPublicBuckets	パブリックなバケットポリシー設定を持つバケットに対して、パブリックなアクセス、クロスアカウントでのアクセスを無力化する	パブリックなバケットポリシー設定を持っていなければ、そのバケットへのアクセスの影響はない



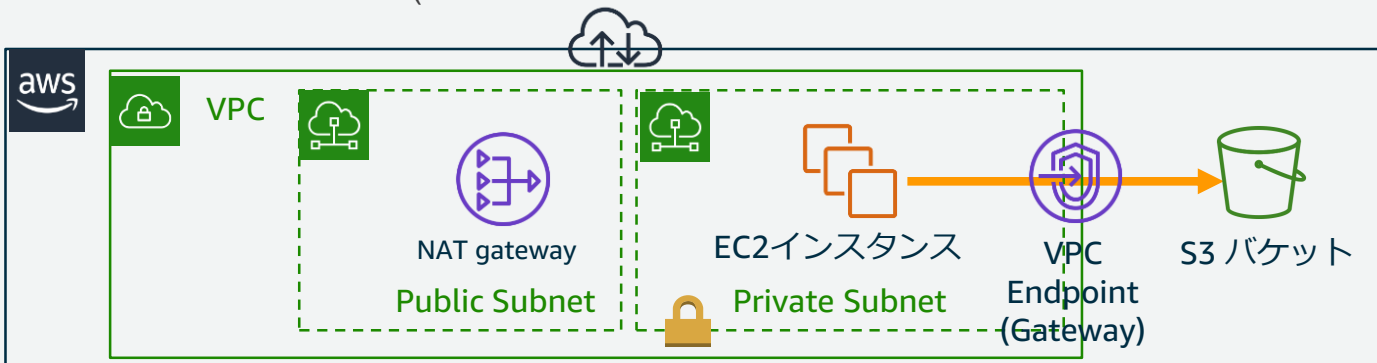
- 「パブリック」の意味= どなたにもアクセスしうる状況
 - ACLで、All Users Authenticated Usersへの許可
 - 誰でもアクセスできるようなバケットポリシー
 - https://docs.aws.amazon.com/ja_jp/AmazonS3/latest/dev/access-control-block-public-access.html
- ブロックパブリックアクセス設定は既存のポリシーまたはACL **設定内容を変更するものではない**
 - ブロックパブリックアクセス設定を削除すると、パブリックポリシーまたはACLを持つバケットまたはオブジェクトは再びパブリックにアクセス可能になる

VPC Endpoint

(*) 2015.5月より

VPC内の**Private Subnet**上で稼働するサービスから、NAT Gatewayを経由せずに、直接S3とセキュアに通信させることが可能

- 通信可能なのは**同一リージョン**のS3のみ
- VPC管理画面のEndpointで作成し、S3と通信したいSubnetの**ルートテーブル**に追加
- Endpoint作成時にアクセスポリシーを定義し、通信可能なBucketや通信元のVPCの指定が可能 (バケットポリシーやIAMポリシーを利用したSource IPやVPC CIDRによる制限は利用不可)
- 別のVPCやSubnetを跨いだ直接のEndpointの利用は不可
- ゲートウェイエンドポイント(PrivateLinkベースのインターフェースエンドポイントではない)



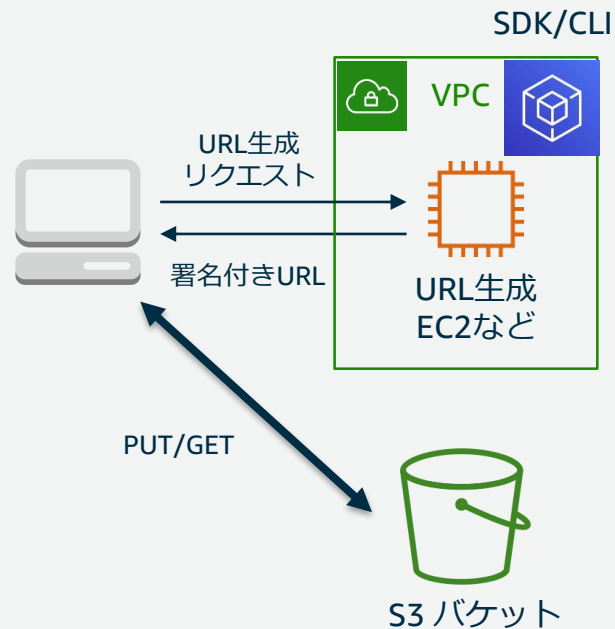
http://docs.aws.amazon.com/ja_jp/AmazonVPC/latest/UserGuide/vpc-endpoints.html

http://aws.typepad.com/aws_japan/2015/05/vpcendpointfors3.html

Pre-signed Object URL (署名付きURL)

AWS SDK (またはAWS CLI)を利用して生成される、署名されたURLを利用し、S3上のプライベートなオブジェクトに対して**一定時間のアクセス**を許可

- Pre-signed URLを利用することで、セキュアにS3とのデータのやり取りが可能
- GETとPUTオペレーションで利用可能
 - 任意のユーザへの一時的なオブジェクト共有
 - 任意のユーザからの一時的なS3へのオブジェクトアップロード権限の付与
- URLを生成したIAMユーザ/ロールの権限が用いられる
- バケット名、オブジェクトキー、HTTPメソッド(GETもしくははPUT)、Expire時間を指定する
- 生成されたURLはExpireする前までが有効
- 注意：そのURLで誰でもそのアクションを実行できる



Pre-signed Object URL (署名付きURL、続き)

署名URLの生成ソースサンプル (Python)

```
# Get the service client.
s3 = boto3.client('s3')

# Generate the URL to get 'key-name' from 'bucket-name'
url = s3.generate_presigned_url(
    ClientMethod='get_object',
    Params={
        'Bucket': 'sample-bucket-cf',
        'Key': 'contents/test.txt'
    },
    ExpiresIn=3600
)
```

```
# 以降でPUTもしくはGET処理を実装
:
```

署名付きURL生成

GET/PUTのいずれかの処理
を指定

対象バケットおよびオブジェ
クトの指定

URL有効期間の指定 (秒)



Webサイトホスティング

静的なWebサイトをS3のみでホスティング可能

- バケット単位で指定
 - Management Consoleで設定可能
 - パブリックアクセスを許可するため別途バケットポリシーで全ユーザーにGET権限を付与
- 独自ドメインの指定
 - ドメイン名をバケット名として指定(www.example.com)
 - 通常は `http://バケット名.s3-website-ap-northeast-1.amazonaws.com`
 - Route53のAlias設定でドメイン名とS3のバケット名を紐付けたレコードを登録
- リダイレクト機能
 - 任意のドメインにリダイレクト設定が可能
 - `x-amz-website-redirect-location`(メタデータの一つ)にセットされる

Webサイトホスティング（続き）

- CORS(Cross-origin Resource Sharing)の設定
 - AJAXなどを利用して、異なるドメインからのS3アクセス時に利用
 - Management Console の場合Bucket PropertiesのPermissionより設定

```
<CORSConfiguration>
<CORSRule>
  <AllowedOrigin>http://www.example.com</AllowedOrigin>
  <AllowedMethod>PUT</AllowedMethod>
  <AllowedMethod>POST</AllowedMethod>
  <AllowedMethod>DELETE</AllowedMethod>
  <AllowedHeader>*</AllowedHeader>
</CORSRule>
</CORSConfiguration>
```

設定例

クロスドメインがwww.example.comの場合、
全てのリクエストを許可

• CloudFrontとの連携

Amazon CloudFront



- WebサーバとしてS3を利用する場合は、**CloudFront経由で配信することを推奨**
- バケットポリシーを利用してCloudFrontからのHTTP/HTTPSリクエストのみを許可することも可能
 - バケットポリシーのPrincipalにCloudFrontのCanonicalUserを指定（CloudFrontの「Origin Access Identity」のコンフィグレーション）

Amazon S3の位置付け
Amazon S3の概要
Amazon S3へのアクセス
Amazon S3のデータ保護
Amazon S3のデータ管理
Amazon S3パフォーマンス最適化
Amazon S3の料金

Amazon S3のデータ保護

暗号化によるデータ保護

保管時(Amazon S3 データセンター内のディスクに格納されているとき)のデータを暗号化して保護するもの

サーバサイド暗号化

- AWSのサーバリソースを利用して格納データの暗号化処理を実施
- 暗号化種別
 - SSE-S3 : AWSが管理する鍵を利用して暗号化
 - SSE-KMS : Key Management Service(KMS)の鍵を利用して暗号化
 - SSE-C : ユーザーが提供した鍵を利用して暗号化(AWSで鍵は管理しない)

デフォルト暗号化

- バケットポリシーを定義することなく、バケットに格納するオブジェクトの暗号化を強制する

クライアントサイド暗号化

- 暗号化プロセスはユーザー管理
- クライアント側で暗号化したデータをS3にアップロード
- 暗号化種別
 - AWS KMSで管理されたカスタマーキーを利用して暗号化
 - クライアントが管理するマスターキーを利用して暗号化

2017.11月

New

暗号化

なし

AES-256
Amazon S3 では、サーバー側の暗号化でデータを暗号化します。

AWS-KMS

キャンセル 保存

デフォルト暗号化

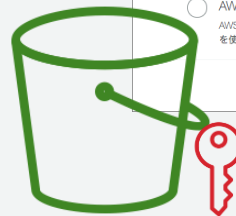
このプロパティは、バケットの既存のオブジェクトには影響しません。

なし

AES-256
Amazon S3 で管理されたキー (SSE-S3) によるサーバー側の暗号化を使用

AWS-KMS
AWS KMS で管理されたキー (SSE-KMS) によるサーバー側の暗号化を使用

キャンセル 保存



S3 Bucket



AWS KMS

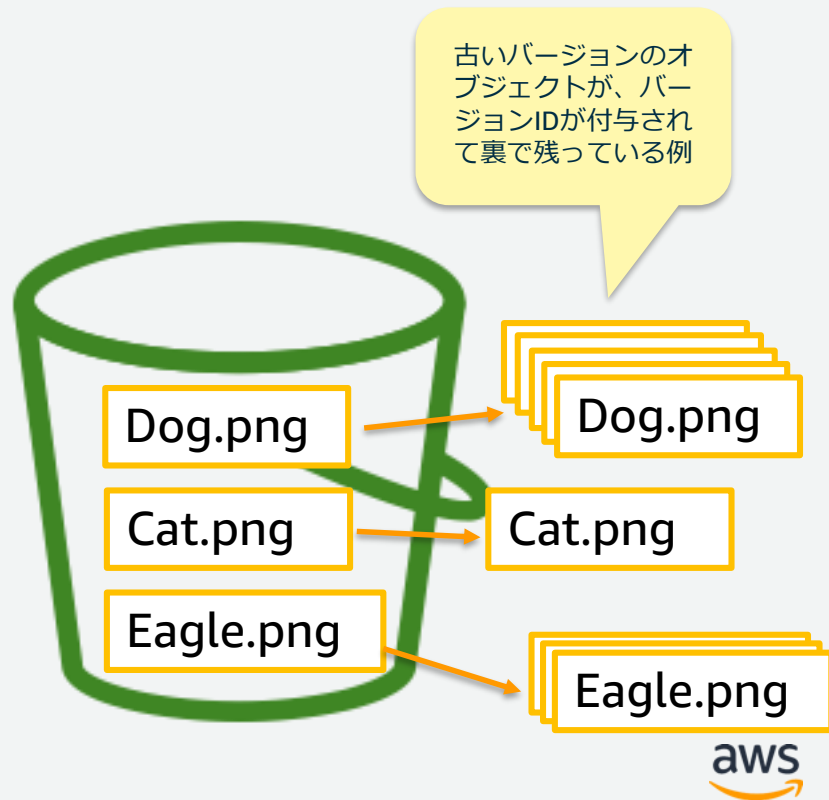


https://docs.aws.amazon.com/ja_jp/AmazonS3/latest/dev/UsingEncryption.html

バージョン管理機能 (Versioning)

ユーザやアプリケーションの誤操作による削除対策に有効

- バケットに対して設定(Enable/Disable)
- 同じキー名でアップロードすると前のバージョンが残る
- バージョン保管されている任意のオブジェクトを参照可能
- バージョニングにより保管されているオブジェクト分も課金
- ライフサイクル管理(後述)と連携し、保存期間(有効期限)も指定可能
- バケットを削除したい場合は、古いバージョンのオブジェクトも削除する
 - ここでも、ライフサイクル管理が便利



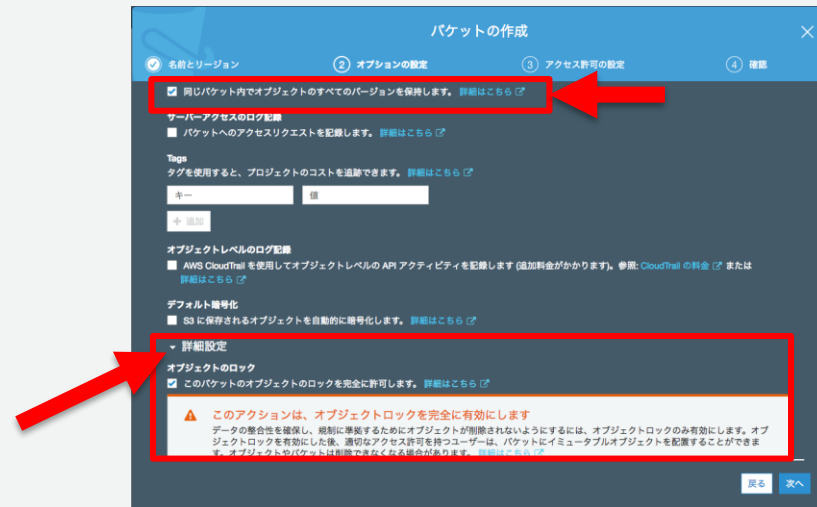
S3 Object Lock (WORM機能)

2018.11月

New

Write Once Read Many(WORM)モデルでのオブジェクト保存を提供する。そのオブジェクトに対する一定期間の上書き、または、削除ができないようロックする。

- Object Lockの有効化は、バケット単位で設定する（新規バケットのみ）
- 保護モード(Retention Mode)及び、保持期間(Retention Period)はバケット単位（デフォルトの設定になる）、または、オブジェクト格納時に明示指定する
- 保持期間とは、このロック（=削除できない状態）が有効な期間のこと
- もしくはリーガルホールド(Legal Hold)のON/OFFが可能
- バージョンングを併用するので、「見た目上の」削除や上書きの動きは妨げられない



例) 30日のRetention Period
(そのバケットに格納されるオブジェクトのデフォルトのロック保持期間が30日になる)

例) オブジェクト単位で、
60日間のロック保持期間



S3 Object Lock (WORM機能)(続き)

2種類の保護モード(Retention Mode)がある

Retention Mode	特徴
コンプライアンスモード (Compliance)	「コンプライアンス」の目的 rootアカウントですら削除ができない、また無効化ができない Cohasset Associates (*1)によるSEC 17a-4アセスメント済み
ガバナンスモード (Compliance)	ガバナンスの効いた「データ保護」 特別な権限(*2)(*3)で WORM保護されたオブジェクトの削除が可能 コンプライアンスモードに変更可能

(*1) <https://d1.awsstatic.com/r2018/b/S3-Object-Lock/Amazon-S3-Compliance-Assessment.pdf>

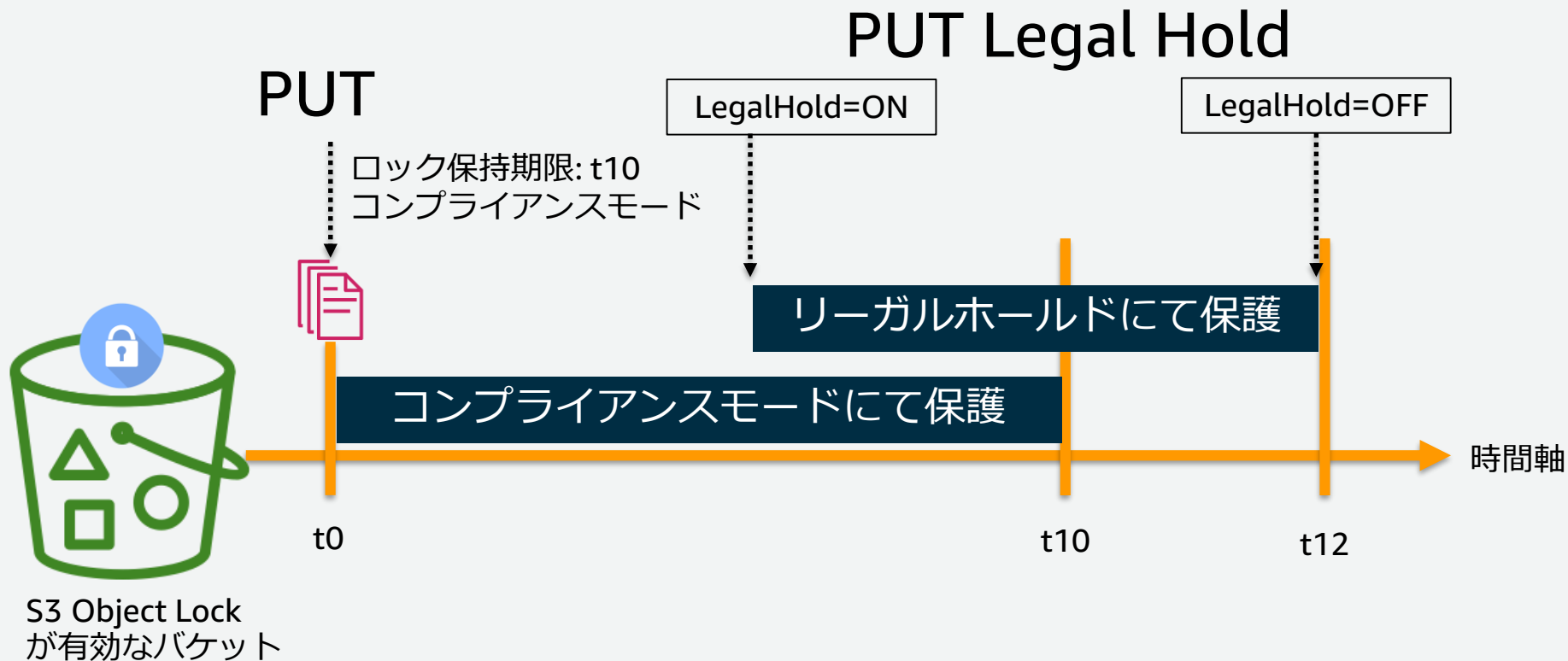
(*2) S3 Object Lockのアクセス許可

オペレーション	必要なアクセス許可
オブジェクトバージョンのRetention Modeや期間を作成、変更	s3:PutObjectRetention
オブジェクトバージョンに対して、Legal Holdを作成、変更	s3:PutObjectLegalHold
オブジェクトバージョンのRetention Modeや期間を取得	s3:GetObjectRetention
オブジェクトバージョンのLegal Holdの状態を取得	s3:GetObjectLegalHold
ガバナンスモードをバイパスする	s3:BypassGovernanceRetention
バケットのObject Lockの設定情報を取得	s3:GetBucketObjectLockConfiguration
バケットのObject Lockの設定を作成、変更	s3:PutBucketObjectLockConfiguration

マネジメントコンソールでの削除操作には x-amz-bypass-governance-retention:trueのリクエストヘッダがつく



S3 Object Lock (WORM機能)(続き)



クロスリージョンレプリケーション

異なるリージョン間のS3バケットオブジェクトのレプリケーションを実施

- バケットに対するオブジェクトの作成、更新、削除をトリガーに非同期でレプリケーションを実行
 - 対象元バケットは**バージョンング**の機能を有効にする必要がある
 - バケットはそれぞれ異なるリージョンでなければならない
 - レプリケーション時は、リージョン間データ転送費用が発生
 - バケット、プレフィックス、オブジェクト単位でのレプリケーション ← *New*
 - レプリケーション元、レプリケーション先でのストレージクラスの指定 ← 2018.11月
 - レプリケーション元でのObject Lockは利用できない
 - マルチアカウントでの利用（レプリケーション先でのオブジェクトオーナーの変更）

東京リージョン
S3 Standard



非同期レプリケーション →



北米リージョン
S3 Glacier

Amazon S3の位置付け
Amazon S3の概要
Amazon S3へのアクセス
Amazon S3のデータ保護
Amazon S3のデータ管理
Amazon S3パフォーマンス最適化
Amazon S3の料金

Amazon S3のデータ管理

ストレージクラス

New

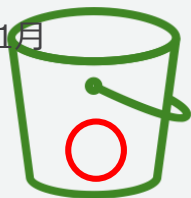
2019年予定



S3 Standard



S3 Intelligent-Tiering



S3 Standard-IA



S3 One Zone-IA



S3 Glacier



S3 Glacier Deep Archive

2018.11月

- アクティブ、頻繁にアクセスするデータ
- ミリ秒アクセス
- ≥ 3 AZ
- \$0.0210/GB~

- 変化するアクセスパターンのデータ
- ミリ秒アクセス
- ≥ 3 AZ
- \$0.0210~\$0.0125/GB
- オブジェクト毎の管理料金
- 最低保持期限

- 低頻度アクセスデータ
- ミリ秒アクセス
- ≥ 3 AZ
- \$0.0125/GB~
- GB毎の取り出し料金
- 最低保持期限
- 最小オブジェクトサイズ

- 再作成可能な低頻度アクセスデータ
- ミリ秒アクセス
- 1 AZ
- \$0.0100/GB~
- GB毎の取り出し料金
- 最低保持期限
- 最小オブジェクトサイズ

- アーカイブデータ
- 分~時間アクセス
- ≥ 3 AZ
- \$0.0040/GB~
- GB毎の取り出し料金
- 最低保持期限

- アーカイブデータ
- 時間アクセス
- ≥ 3 AZ
- \$0.00099/GB~
- GB毎の取り出し料金
- 最低保持期限

30日以上

30日以上、128KB以上

90日以上

いずれもできるだけサイズの「大きな」オブジェクトでの利用が良い



ライフサイクル管理

バケット内のオブジェクトに対して、ストレージクラスの変更や、削除処理に関する自動化

- バケット全体もしくはPrefixに対して、オブジェクトの更新日をベースに日単位での指定が可能
- 最大1,000までLifecycleのルールを設定可能
- 毎日0:00UTCに処理がキューイングされ順次実行
- Lifecycleを利用してIAに移動できるのは128KB以上のオブジェクトのみでそれ以外はIAに移動されない
- STANDARD-IA・アーカイブおよび削除の日程をそれぞれ指定した組み合わせも可能
- マルチアップロード処理で完了せず残った分割ファイルの削除にも対応
- MFA delete が有効なバケットにはライフサイクル設定は不可



The screenshot shows the 'Lifecycle Rule' configuration page in the Amazon S3 console. The page title is 'ライフサイクルルール' (Lifecycle Rule). The navigation tabs are: 1. 名前とスコープ (Name and Scope), 2. 移行 (Transition) (selected), 3. 有効期限 (Expiration), and 4. 確認 (Review). The main content area is titled 'ストレージクラスの移行' (Storage Class Transition). Below the title, there is a description: 'ライフサイクルの設定にルールを追加して、別のストレージクラスにオブジェクトを移行するように Amazon S3 に指定できます。 詳細はこちら' (Add a rule to the lifecycle configuration to specify Amazon S3 to transition objects to another storage class. See details here). There are two radio buttons: '現行バージョン' (Current version) (checked) and '以前のバージョン' (Previous version) (unchecked). Below this, there is a section for 'オブジェクトの現行バージョン' (Current version of objects) with a '+ 移行を追加する' (Add transition) button. The table below shows the transition rules:

オブジェクト作成	オブジェクト作成からの日数
標準-IA への移行の期限	30 X
標準-IA への移行の期限	
インテリジェントへの移行の期限	
ゾーン - IA への移行の期限	
Amazon Glacier への移行の期限	

At the bottom right, there are two buttons: '戻る' (Back) and '次へ' (Next).

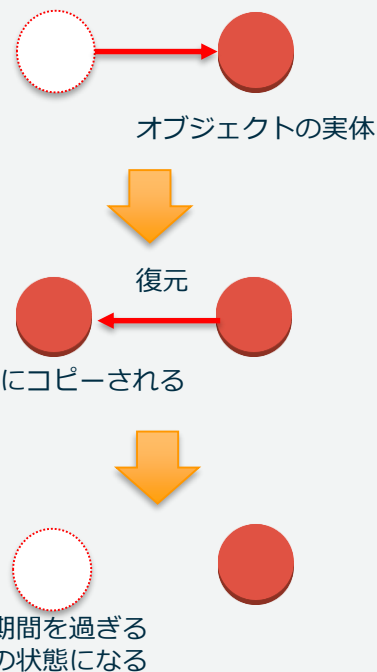
S3 Glacierへのアーカイブと復元

アーカイブ

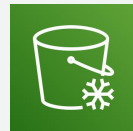
- オブジェクトのデータはS3 Glacierに移動(アーカイブ後、マスターはS3 Glacierとなる)
- **オブジェクトを S3 Glacierに直接PUTが可能** *New* 2018.11月
- S3上のデータを削除することで、S3 Glacier側のデータも削除される
- S3には8KBのオブジェクト名とメタデータのみが保管
- S3 Glacierには32KBのインデックスおよび関連メタデータが追加で保管
- アーカイブしたオブジェクトを90日以内に削除しても、90日間アーカイブされたのと同じ課金対象

オブジェクトの復元(restore)

- オブジェクト毎に復元
- **オブジェクト復元時(復元開始と復元完了)のNotification** *New* 2018.11月
- データは一時的にS3の低冗長化ストレージに指定日数間複製される
- 復元後の、S3上での保持期間の変更も可能
- 復元にかかる時間について、3種類から選択可能
- 復元期間中は、S3の低冗長化ストレージとS3 Glacier双方で課金



S3 Glacierへのアーカイブと復元（続き）



S3 Glacier

- 復元リクエスト時に指定できる3つの選択肢
 - **Expedited**: 少ない数のファイルについて、緊急のアクセスを要する場合の取得
 - **Standard**: 3-5時間の間にファイルを取得する標準的な取得
 - **Bulk**: 5-12時間の間にファイルを取得する最も低価格で、大量のデータを取得
- 復元リクエストのアップグレード *New* 2018.11月

	迅速(Expedited)	標準(Standard)	大容量(Bulk)
データアクセス時間	1~5分	3~5時間	5~12時間
データ復元容量	\$0.033 / GB	\$0.011 / GB	\$0.00275 / GB
復元リクエスト	オンデマンド: \$0.011 リクエストごと プロビジョンド: \$110プロビジョンド キャパシティユニットごと(*)	\$0.0571: 1,000 リクエストあたり	\$0.0275: 1,000 リクエストあたり

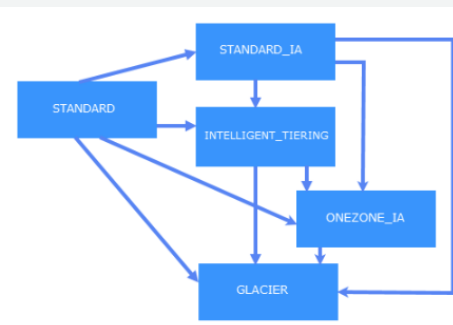
(*)プロビジョンド=あらかじめデータを取り出すリソースを購入できる考え方
1プロビジョンドキャパシティユニット=5分間に、3回までのExpedited復元リクエスト、かつ、復元時スループットが150MB/sec以内

東京リージョン<https://aws.amazon.com/jp/glacier/pricing/>



ストレージクラス間のオブジェクト移動の整理

ライフサイクル管理で指定する経過
時間による移行



STANDARD
(標準)

STANDARD-IA
(低頻度アクセス)

S3 Glacier

Glacier からオブジェクトを復元

選択: 1 オブジェクト, 0 フォルダ 合計サイズ: 3.5 MB オブジェクトの合計: 1

復元されたコピーを使用可能な日数
低冗長化ストレージ (RFS) 内の復元されたコピーは、指定した日数が経過すると自動的に削除されます。

日間

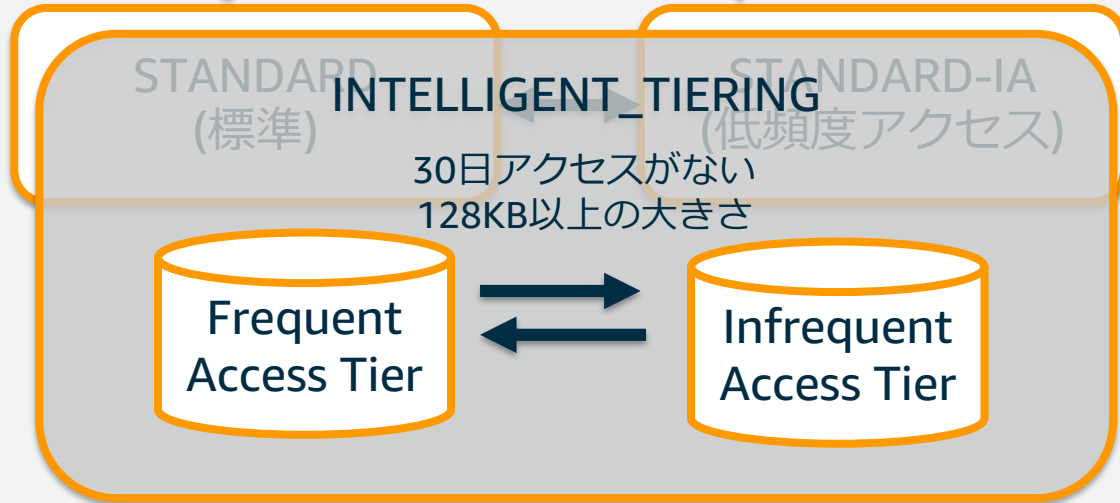
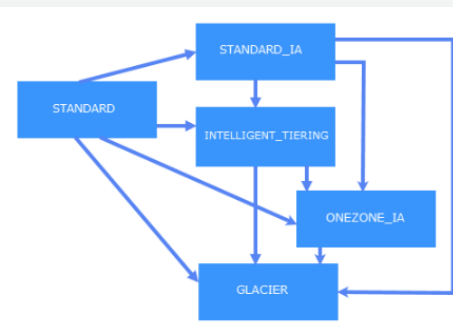
復元の階層
Glacier ではリクエスト料金と GB 単位の取り出し料金がかかります。この料金は選択された階層によって異なります。参照: [S3 の料金](#)

- 一括取得
通常は 5~12 時間以内
- 標準取り出し
通常は 3~5 時間以内
- 迅速取り出し
250 MB 以下の取り出しの場合、通常 1~5 分以内

REDUCED_REDUNDANCY
(低冗長化ストレージ)

ストレージクラス間のオブジェクト移動の整理

ライフサイクル管理で指定する経過
時間による移行



S3 Analytics

「STANDARD-IAとS3 Glacierどちらににいつ移動すればいいのだろうか？」この疑問に応える、データのアクセスパターンの簡易可視化

開始方法

- 目的のバケットに対して、分析フィルターを定義する
- 結果が出るまで、フィルター作成してから24~48時間ほど待つ

CSVでも結果を出力する場合→



http://docs.aws.amazon.com/ja_jp/AmazonS3/latest/dev/analytics-storage-class.html

S3 Analytics (続き)

青が格納量、紫がどれだけそのデータが読まれたか？

この例の場合は、

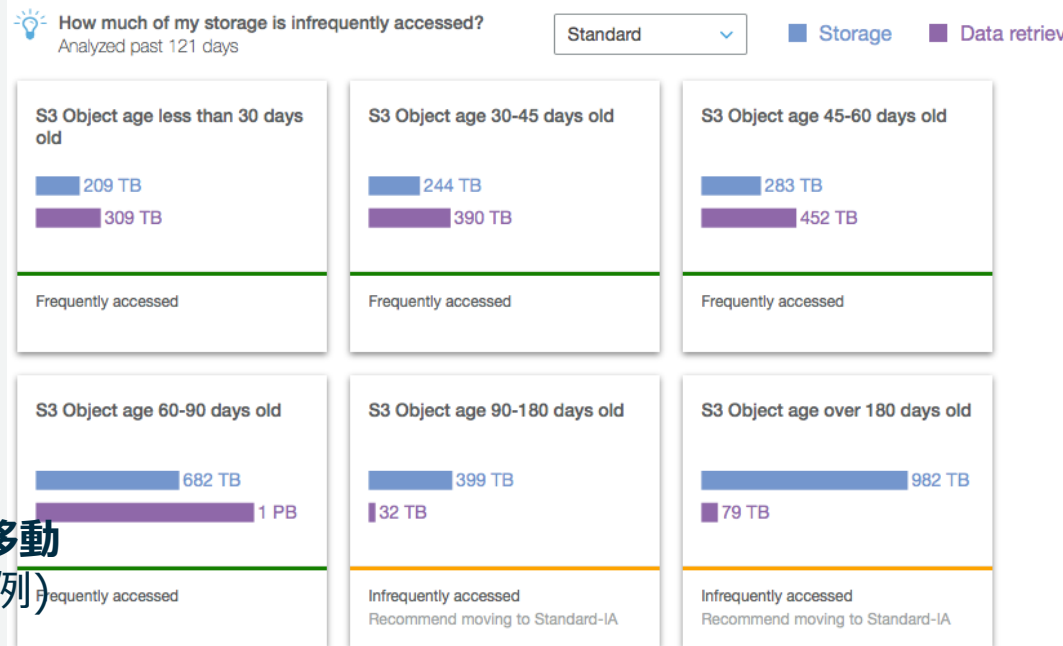
- 90日までのオブジェクトは、そこそこアクセスがある。
- 90日以降のオブジェクトのニーズが急に減っている。
- 90日以降でも、全くアクセスがないわけではない。
- 他のコンプライアンス要件などを加味したとして、...



90日経過したデータをSTANDARD-IAへ移動

365日経過したデータをS3 Glacierへ移動(例)

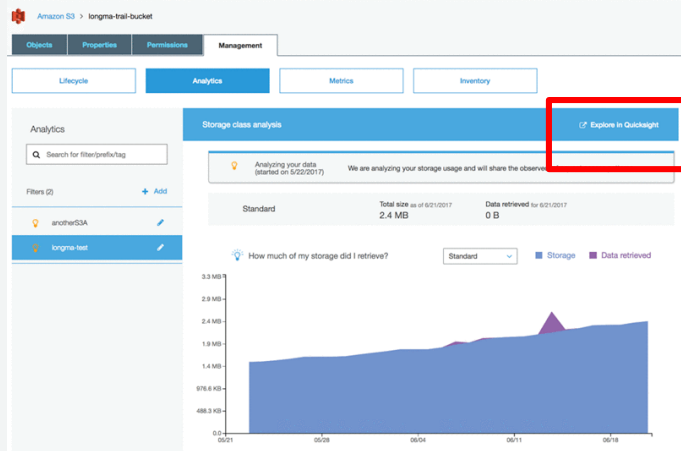
5年経過したデータは削除(例)



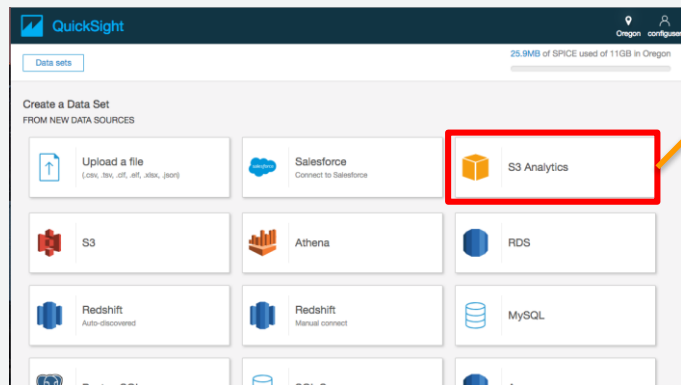
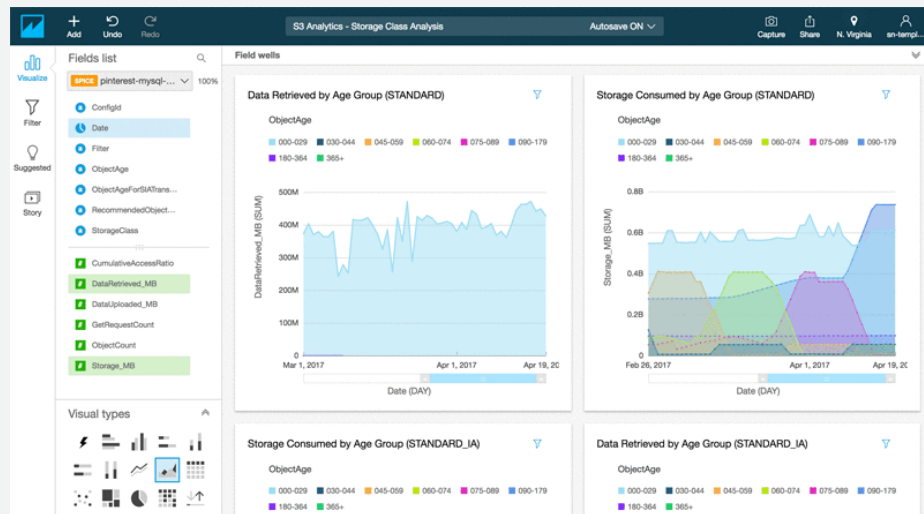
S3 Analytics を QuickSightでみる



Amazon QuickSight



S3管理コンソールの 管理->分析からたどる



もしくは、QuickSightのデータセット作成時 (New Dataset)でS3 Analyticsを選択

<https://aws.amazon.com/blogs/big-data/visualize-amazon-s3-analytics-data-with-amazon-quicksight/>



S3 インベントリ

S3に入っているオブジェクトのリストを、一気にCSVまたはORCファイルで取得する

- オブジェクトのリストを取得するにあたって、List Bucketの処理に時間や手間がかかる場合に有益
- スケジュール化（日単位・週1回）してレポートを取得
- 初回の結果が出るまで、48時間待つ
- ある時点のsnapshotとしてのPUT/DELETE（結果整合性）結果のインベントリリストとなる



http://docs.aws.amazon.com/ja_jp/AmazonS3/latest/dev/storage-inventory.html

S3インベントリ (続き)

例) インベントリを取得したいバケット: sample-bucket-analytics-oregon

インベントリ名	フィルター	送信先バケット	送信先プレフィックス	頻度
<input type="text" value="sample-inventory"/>	<input type="text" value="プレフィックスでフィルター (省略)"/>	<input type="text" value="redshift-bucket-toruyakio"/>	<input type="text" value="s3inventory"/>	<input type="text" value="週1回"/>

マニフェストファイルの吐き出し先

`destination-prefix/source-bucket/config-ID/YYYY-MM-DDTHH-MMZ/manifest.json`
`destination-prefix/source-bucket/config-ID/YYYY-MM-DDTHH-MMZ/manifest.checksum`

この例の場合:

`s3://redshift-bucket-toruyakio/s3inventory/sample-bucket-analytics-oregon/sample-inventory/2017-02-14T15-02Z/manifest.json`
`s3://redshift-bucket-toruyakio/s3inventory/sample-bucket-analytics-oregon/sample-inventory/2017-02-14T15-02Z/manifest.checksum`

インベントリリストの吐き出し先

`destination-prefix/source-bucket/data/example-file-name.csv.gz`

この例の場合

`s3://redshift-bucket-toruyakio/s3inventory/sample-bucket-analytics-oregon/sample-inventory/data/0042fc70-0dee-4e0a-9fb5-92c639d1d93c.csv.gz`

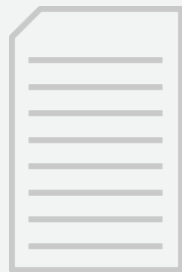
Bucket	Key	VersionID	IsLatest	IsDeleteMaker	Size	Last modified date	Etag	Storage Class	Replication Status
sample-bucket-analytics-oregon	bad_keys/00000002/2017-	d2fPieFQm7	TRUE	FALSE	1024	2017-02-15T00:23:43.000Z	0f343b09311	STANDARD	COMPLETED
sample-bucket-analytics-oregon	bad_keys/00000002/2017-	fnS18W.tD4H	TRUE	FALSE	1024	2017-02-15T00:24:05.000Z	0f343b09311	STANDARD	COMPLETED
sample-bucket-analytics-oregon	bad_keys/00000002/2017-	sTSM3kb7E5	TRUE	FALSE	1024	2017-02-15T00:24:15.000Z	0f343b09311	STANDARD	COMPLETED
sample-bucket-analytics-oregon	bad_keys/00000002/2017-	UZkrmkdrqZH	TRUE	FALSE	1024	2017-02-15T00:23:54.000Z	0f343b09311	STANDARD	COMPLETED
sample-bucket-analytics-oregon	bad_keys/00000002/2017-	kkxyNpUDpl	TRUE	FALSE	1024	2017-02-15T00:25:53.000Z	0f343b09311	STANDARD	COMPLETED
sample-bucket-analytics-oregon	bad_keys/00000002/2017-	Dhe37pgyHs	TRUE	FALSE	1024	2017-02-15T00:24:55.000Z	0f343b09311	STANDARD	COMPLETED
sample-bucket-analytics-oregon	bad_keys/00000002/2017-	Oo8UJgBwC	TRUE	FALSE	1024	2017-02-15T00:25:09.000Z	0f343b09311	STANDARD	COMPLETED
sample-bucket-analytics-oregon	bad_keys/00000002/2017-	ShIIN9n_agC	TRUE	FALSE	1024	2017-02-15T00:25:53.000Z	0f343b09311	STANDARD	COMPLETED
sample-bucket-analytics-oregon	bad_keys/00000002/2017-	zxO8Q9dLe	TRUE	FALSE	1024	2017-02-15T00:25:11.000Z	0f343b09311	STANDARD	COMPLETED
sample-bucket-analytics-oregon	bad_keys/00000002/2017-	WaaQPOaJq	TRUE	FALSE	1024	2017-02-15T00:26:17.000Z	0f343b09311	STANDARD	COMPLETED

S3 バッチオペレーション(Preview)

New

数千、数百万、数十億のオブジェクトに対するAPIアクションを一括実行

オブジェクトの選択



マニフェストファイル
- S3 インベントリ (CSV)
- CSVファイル

オペレーションの選択



- COPY (PUT Object Copy)
- S3 Glacierからのリストア
- PUT ObjectACL
- PUT Object Tagging
- Lambda関数の呼び出し



ジョブの作成



ジョブの進捗
ジョブの通知
ジョブの状態
完了レポート

S3 イベント通知



SNS



SQS



AWS Lambda

バケットにてイベントが発生した際に、Amazon SNS, SQS, Lambdaに対して通知することでシームレスなシステム連携が可能

イベントタイプ	概要
s3:ObjectCreated:*	S3バケットにオブジェクト作成された時 (PUT/POST/COPYのAPIがコールされた時)
s3:ObjectCreated:Put	
s3:ObjectCreated:Post	
s3:ObjectCreated:Copy	
s3:ObjectCreated:CompleteMultipartUpload	
s3:ObjectRemoved:*	S3バケットから、オブジェクトが削除された時 Delete = バージョニングされていないオブジェクトの削除、またはバージョンングされているバケットのオブジェクトの完全な削除 DeleteMarkerCreated = バージョニングされているオブジェクトの削除マーカ作成
s3:ObjectRemoved:Delete	
s3:ObjectRemoved:DeleteMarkerCreated	
s3:ObjectRestore:Post	S3 Glacierストレージクラスから復元の開始、完了した時
s3:ObjectRestore:Completed	
s3:ReducedRedundancyLostObject	低冗長化ストレージにてデータロストが発生した時

- Amazon SNS: メール送信, HTTP POST, モバイルPushなどのTopics実行
- Amazon SQS: キューメッセージの登録
- Amazon Lambda: 指定Lambda Functionの実行



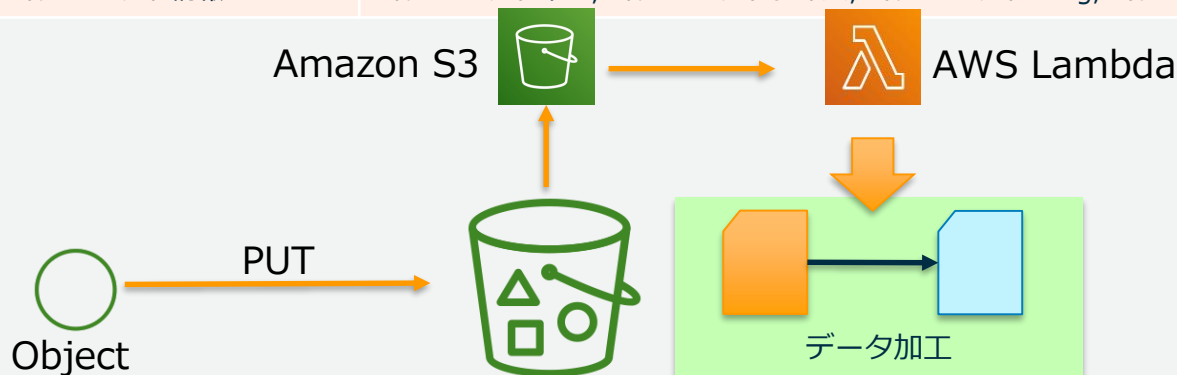
S3 イベント通知（続き）

S3からの実行権限の付与

- SNSおよびSQSはそれぞれのPolicyに対してS3からの実行権限を付与
- Lambdaに関しては、Lambdaが利用するIAM RoleにS3の権限を付与

イベントでの通知内容

	通知項目
共通情報	リージョン, タイムスタンプ, Event Type
リクエスト情報	Request Actor Principal ID, Request Source IP, Request ID, Host ID
バケット情報	Notification Configuration Destination ID, バケット名, バケットARN, バケットOwner Principal ID
オブジェクト情報	オブジェクトキー, オブジェクトサイズ, オブジェクトETag, オブジェクトバージョンID



Amazon CloudWatchによるメトリクス管理



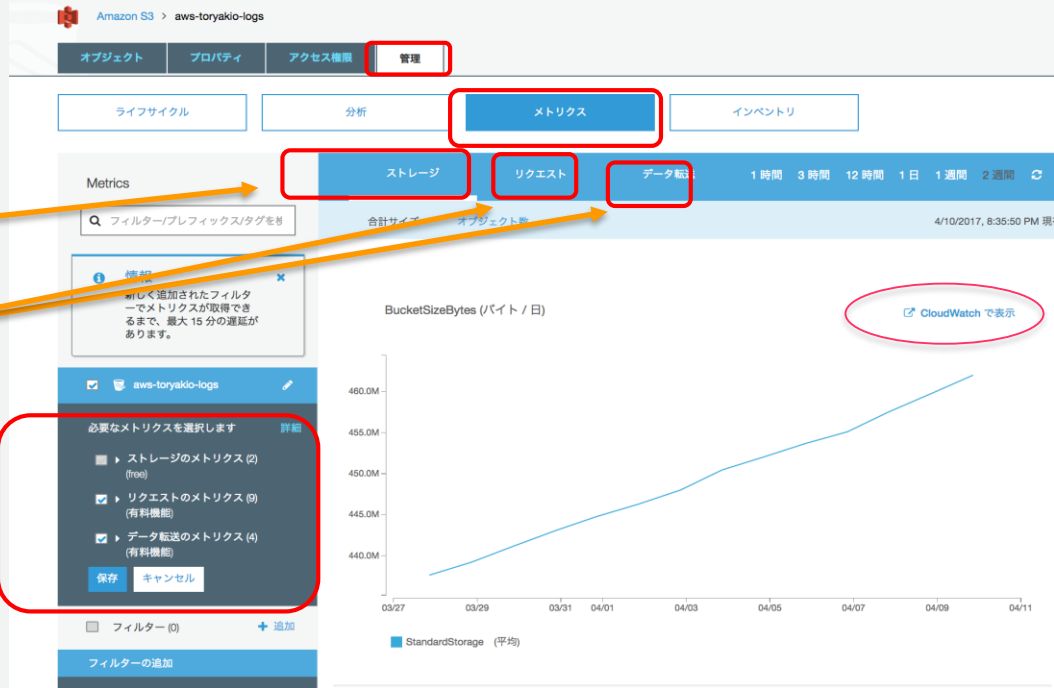
Amazon CloudWatch

1. バケットに対するストレージメトリクス → 日単位
2. オブジェクトに対するリクエストメトリクス → 分単位

ストレージメトリクス

リクエストメトリクス

そのバケットで、リクエストメトリクスを利用する際に設定→



https://docs.aws.amazon.com/ja_jp/AmazonS3/latest/user-guide/configure-metrics.html

Amazon CloudWatchによるメトリクス管理（続き）

New

S3 Glacier

ストレージメトリクス

- バケット単位および、Storage Typeごとにメトリクスを把握する
- 1日間隔でのレポート、状況把握（追加料金なし）

メトリクス	詳細
BucketSizeBytes	標準ストレージクラス、INTELLIGENT_TIERING、STANDARD IAストレージクラス、OneZone-IA、Glacier、または低冗長化ストレージ (RRS) クラスのバケットに保存されたバイト単位のデータ量
NumberOfObjects	すべてのストレージクラスのバケットに保存されたオブジェクトの総数

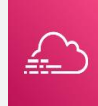
リクエストメトリクス

- タグやプレフィックスの指定にて細かい粒度での把握も可能
- 1分間隔でのメトリクスで、通常のCloudWatchの料金

New
S3 Select

メトリクス	単位	メトリクス	単位	メトリクス	単位
AllRequests	Count	SelectRequests	Count	BytesDownloaded	MB
PutRequests	Count	SelectScannedBytes	Bytes	BytesUploaded	MB
GetRequests	Count	SelectReturnedBytes	Bytes	4xxErrors	Count
ListRequests	Count			5xxErrors	Count
DeleteRequests	Count			FirstByteLatency	ms
HeadRequests	Count			TotalRequestLatency	ms
PostRequests	Count				

AWS CloudTrailによるAPI管理



AWS CloudTrail

CloudTrailを有効にすることでS3への操作ログ(API Call)を収集することが可能

いつ、どこから、誰がS3の操作を行ったか、コンプライアンスの目的で把握可能(S3 イベント通知との使い分けを意識)

CloudTrailでのイベント	操作
データイベント	GetObject, DeleteObject, PutObjectなどのS3のオブジェクトに対するAPI操作
管理イベント	S3のバケット操作はもちろん、その他のすべてのAPI操作

監査対象とは別のS3バケットを用意することを推奨

100,000イベントごとに、\$0.1の料金

ログに記録されるS3オペレーションは下記を参照

http://docs.aws.amazon.com/ja_jp/AmazonS3/latest/dev/cloudtrail-logging.html

http://docs.aws.amazon.com/ja_jp/awscloudtrail/latest/userguide/logging-management-and-data-events-with-cloudtrail.html



その他のモニタリングや管理に有効な機能

Logging

- バケット単位でバケットに対するアクセスログの出力設定が可能
- 出力先としてS3バケットを指定
- ログフォーマットは下記を参照

http://docs.aws.amazon.com/ja_jp/AmazonS3/latest/dev/LogFormat.html

Tag管理

- バケット、オブジェクトに対してタグの指定が可能
- タグ指定によりリソースグループにて関連するAWSサービスとの紐付けが可能
- **オブジェクトに対してのタグ**は、ここまで紹介したライフサイクル、分析、モニタリング、クロスリージョンレプリケーション機能で活用可能

Amazon S3の位置付け
Amazon S3の概要
Amazon S3へのアクセス
Amazon S3のデータ保護
Amazon S3のデータ管理
Amazon S3パフォーマンス最適化
Amazon S3の料金

Amazon S3 パフォーマンス最適化



大きなサイズのファイルを快適に、ダウンロード、アップロード

GETリクエストについて、**RANGE GETを活用**することで、マルチスレッド環境では高速にダウンロードが可能

- マルチパートアップロード時と同じチャンクサイズを利用する



マルチパートアップロードの活用によるアップロード(PUT)オペレーションの高速化

- チャンクサイズと並列コネクション数のバランスが重要
 - 帯域が太い場合は**20MB-50MB**チャンクサイズから調整
 - モバイルや帯域が細い場合は**10MB程度**から調整

パフォーマンスの最適化（続き）



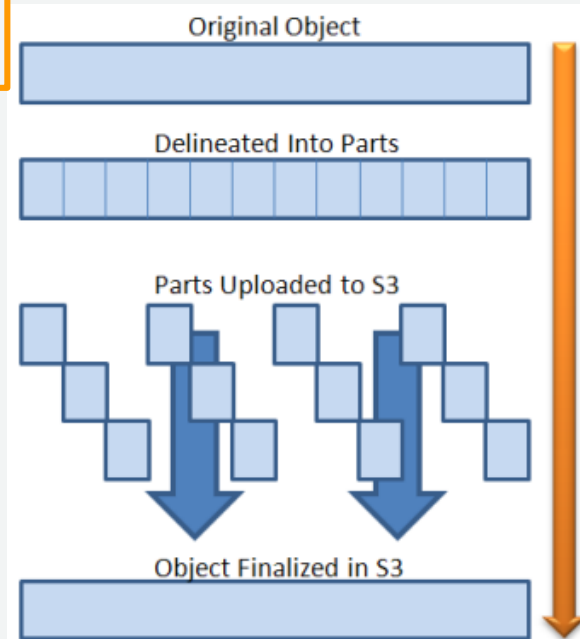
主にメディア

目安100MB以上のファイルのアップロードを快適にしたい場合のマルチパートアップロード機能

S3にアップロードする際に、ファイルを複数のチャンクに分割して並列アップロードを実施

- ファイルが100MBを超える場合、利用することを推奨
- 各チャンクは5GB以下に設定(5MB-5GB)
- 全てのチャンクがアップロードされるとS3側で結合
- Multipart Uploadを利用することで単一オブジェクトで5TBまで格納可能

各SDKにてMultipart Uploadの機能は実装済みAWS CLIの場合は、ファイルサイズを元に自動的に判別PUT処理を並列化することでのスループット向上を期待→広帯域ネットワークが重要



http://docs.aws.amazon.com/ja_jp/AmazonS3/latest/dev/mpuoverview.html

<http://docs.aws.amazon.com/cli/latest/topic/s3-config.html>

S3 Transfer Acceleration

AWSのマネージドバックボーンネットワークを活用した高速ファイル転送手法

全世界149箇所(*)にあるAWSのエッジネットワークから、最適化されたAWSのネットワークを經由して、高速にAmazon S3とのデータ転送を実現

- 利用者は自動的に最短のエッジネットワークに誘導

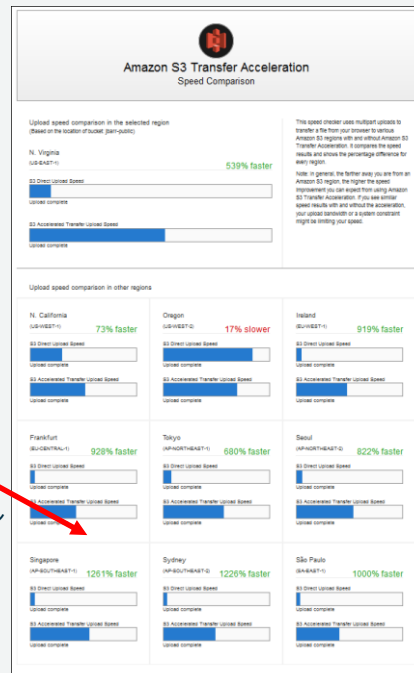
S3 Bucketに対してAccelerationを有効化

- S3へのアクセスエンドポイントを変更するだけで利用可能
- Acceleration有効後、転送速度が高速化されるまでに最大30分かかる場合がある
- バケット名はピリオド"."が含まれない名前にする必要がある
- IPv6 (dualstack)エンドポイントも指定可能

利用している端末からの無料スピード測定ツールも提供



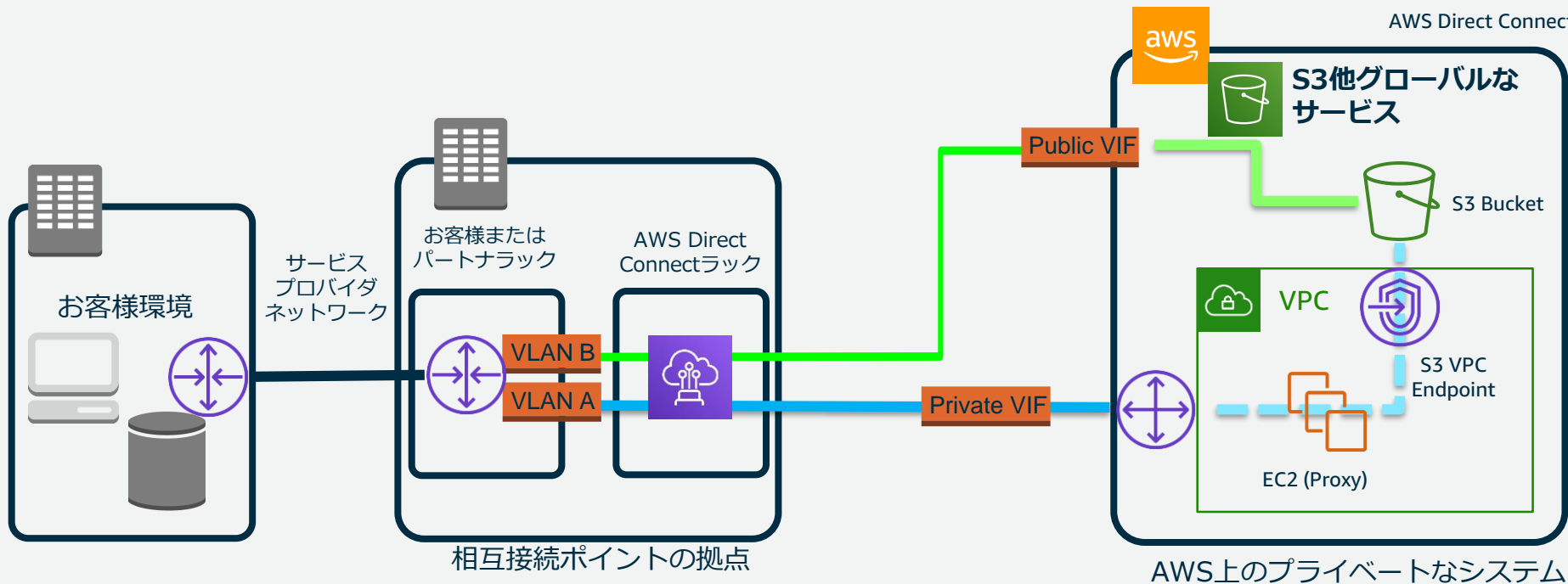
マネージメントコンソールからも起動可能



S3 と Direct Connect



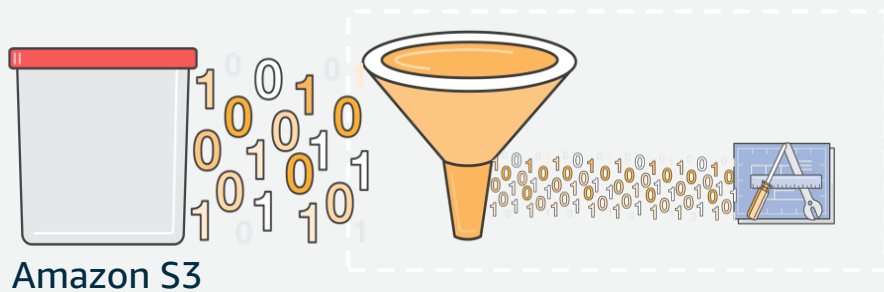
AWS Direct Connect



- 1) お客様環境と VPC への専用線による接続 (Private接続)
- 2) AWS のグローバルなサービスとの専用線による接続(Public接続)

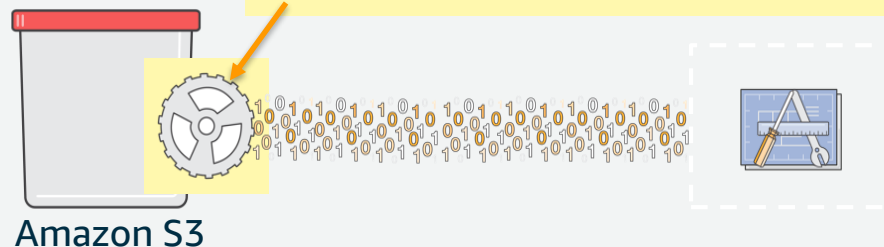
S3 Select

Amazon S3 に格納されているオブジェクトに対して、SQL式にて、部分のみを抽出できる



オブジェクト全体を取得して、アプリケーションにて、抽出する

S3 Selectによりフィルタリング



価格、速度でのメリット

アプリケーションが、S3 Select を利用して、オブジェクトの一部のみを取得する

S3 Select (続き)

- Input : フォーマットはCSV, JSON、圧縮(GZIP,BZIP3), 暗号化(SSE)
- Output: CSV, JSON
- SDK: Java, Python(boto3), Ruby, Go, .NET, JavaScript

句	データタイプ	オペレータ	関数
Select	String	Conditional	String
From	Integer, Float, Decimal	Math	Cast
Where	Timestamp	Logical	Math
	Boolean	String (Like,)	Aggregate

制約など: <https://docs.aws.amazon.com/AmazonS3/latest/dev/selecting-content-from-objects.html>

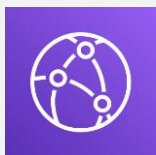
SQLリファレンス : <https://docs.aws.amazon.com/AmazonS3/latest/dev/s3-glacier-select-sql-reference.html>

リクエストレート

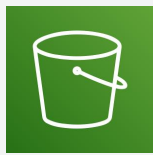
Amazon S3 は、自動的にスケールするよう設計されています。データの追加操作で最大 3,500 リクエスト/秒の、データの取得操作で最大 5,500 リクエスト/秒をサポートできるようにパフォーマンスを向上させています。

重要

大量のGETリクエストが発生するワークロードの場合は、Amazon CloudFrontを併用することを推奨



Amazon CloudFront



Amazon S3

定常的にS3バケットへのPUT/LIST/DELETEリクエストが3,500 RPSを超える、もしくはGETリクエストが5,500RPSを超える場合、キー名先頭部分の文字列をランダムにすることでレート向上が期待できるが、その必要性があるワークロードかどうかはよく見極める

```
examplebucket/2013-26-05-15-00-00/cust1234234/photo1.jpg  
examplebucket/2013-26-05-15-00-00/cust3857422/photo2.jpg  
...
```

```
examplebucket/2013-26-05-15-00-01/cust1248473/photo4.jpg  
examplebucket/2013-26-05-15-00-01/cust1248473/photo5.jpg
```

ほとんどのユースケースでプレフィックスをハッシュする必要はない

```
examplebucket/232a-2013-26-05-15-00-00/cust1234234/photo1.jpg  
examplebucket/7b54-2013-26-05-15-00-00/cust3857422/photo2.jpg  
...  
examplebucket/9810-2013-26-05-15-00-01/cust1248473/photo4.jpg  
examplebucket/c34a-2013-26-05-15-00-01/cust1248473/photo5.jpg  
...
```

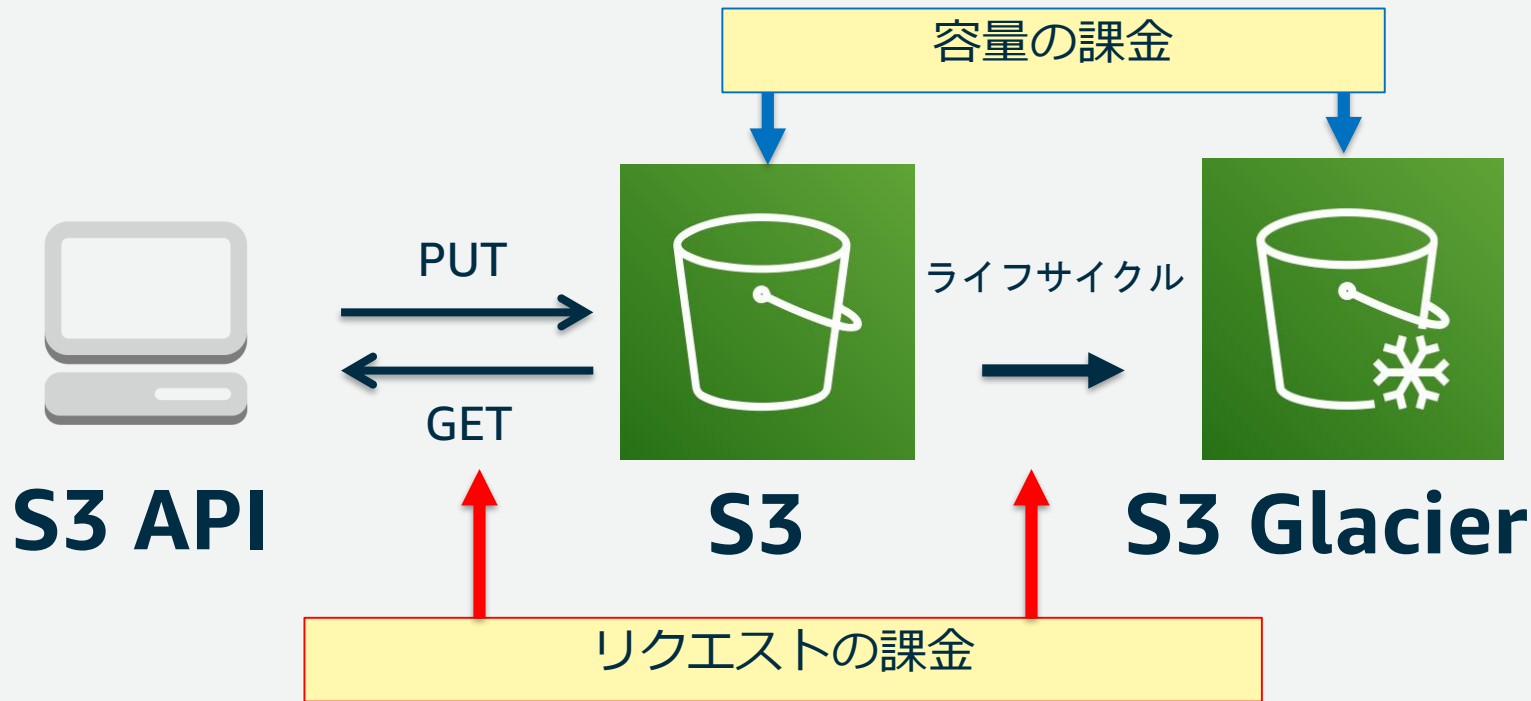
<https://aws.amazon.com/jp/about-aws/whats-new/2018/07/amazon-s3-announces-increased-request-rate-performance/>



Amazon S3の位置付け
Amazon S3の概要
Amazon S3へのアクセス
Amazon S3のデータ保護
Amazon S3のデータ管理
Amazon S3パフォーマンス最適化
Amazon S3の料金

Amazon S3の料金

主に、容量の料金とオペレーションの料金



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。

Amazon S3の料金

ストレージ料金

	スタンダード	STANDARD-IA(*)	S3 OneZone -IA	S3 Glacier
最初の50TB/月	\$0.025 / GB	\$0.019 / GB	\$0.0152 / GB	\$0.005 / GB
次の450TB/月	\$0.024 / GB	\$0.019 / GB	\$0.0152 / GB	\$0.005 / GB
500TB月以上	\$0.023 / GB	\$0.019 / GB	\$0.0152 / GB	\$0.005 / GB

(*) STANDARD-IAの請求対象となる最小オブジェクトサイズは 128 KB です。128 KB より小さいサイズのオブジェクトは、128 KBとして課金されます。

リクエスト料金

	スタンダード	INTELLIGENT_TIERING	STANDARD-IA	S3 OneZone -IA	S3 Glacier
PUT、COPY、POST、または LIST リクエスト	\$0.0047 : 1,000 リクエストあたり	\$0.0047 : 1,000 リクエストあたり	\$0.01 : 1,000 リクエストあたり	\$0.01 : 1,000 リクエストあたり	-
GET、SELECTおよび他のすべてのリクエスト	\$0.00037 : 1,000 リクエストあたり	\$0.00037 : 1,000 リクエストあたり	\$0.001 : 1,000 リクエストあたり	\$0.001 : 1,000 リクエストあたり	-
S3 Selectによって返されたデータ	\$0.0008 / GB	\$0.0008 / GB	\$0.01 / GB	\$0.01 / GB	Glacier Select
S3 Selectによってスキャンされたデータ	\$0.00225 / GB	\$0.00225 / GB	\$0.00225 / GB	\$0.00225 / GB	Glacier Select
ライフサイクル移行リクエスト	-	\$0.01 : 1,000 リクエストあたり	\$0.01 : 1,000 リクエストあたり	\$0.01 : 1,000 リクエストあたり	\$0.0571 : 1,000 リクエストあたり
取り出し (容量)			\$0.01 / GB	\$0.01 / GB	Glacier取り出し料金 (slide 47)

Amazon S3 の料金(続き)

ストレージマネジメント料金

管理	料金
S3 Inventory	リストされるオブジェクト 100 万個ごとに \$0.0028
S3 Analytics Storage Class Analysis	モニターされるオブジェクト 100 万個ごとに月あたり \$ 0.10
S3 Object Tagging	10,000 タグごとに月あたりUS\$0.01
CloudWatch リクエストメトリクス	CloudWatch 料金
CloudTrail データイベント	100,000 件のイベントあたり \$0.1
S3 Intelligent-Tiering モニタリング、オート メーション	オブジェクト1,000件ごとに \$0.0025

データ転送料金

New 2018.9月

転送方向		価格
IN	全てのデータ転送「IN」	\$0.000/GB
OUT (AWS Network)	同じリージョンのAmazon EC2	\$0.000/GB
	別のAWSリージョン	\$0.090/GB
	Amazon CloudFront	\$0.000/GB
OUT (Internet)	最初の1GB/月	\$0.000/GB
	10TBまで/月	\$0.114/GB
	次の40TB/月	\$0.089/GB
	次の100TB/月	\$0.086/GB
	次の350TB/月	\$0.084/GB
	350TB/月以上	お問い合わせ

2019年2月時点の東京リージョン料金表
<http://aws.amazon.com/jp/s3/pricing/>



Amazon S3 の料金(続き)

S3 Transfer Acceleration料金

転送方向		価格
S3へのデータIN	米国、欧州、日本のエッジロケーションによる高速化	\$0.04/GB
	その他の国のエッジロケーションによる高速化	\$0.08/GB
S3からのデータOUT (Internet)	エッジロケーションによる高速化	\$0.04/GB
S3と別のAWSリージョン間	エッジロケーションによる高速化	\$0.04/GB

S3 Transfer Accelerationの費用は、S3のデータ転送コストとは別に加算されることに注意

S3 Transfer Accelerationを利用してデータをやり取りする場合、通常のS3との転送よりも高速であるかを確認します。通常の転送に比べTransfer Accelerationが高速でないと判断した場合は、Transfer Accelerationの料金は請求されず、Transfer Accelerationシステムをバイパスする可能性があります。

S3 無料枠(1年)

- 標準ストレージ 5GB
- 20,000 GETリクエスト / 2,000 PUTリクエスト

まとめ

まとめ

Amazon Simple Storage Service (S3)は、ユーザがデータを安全に、容量制限なく、データ保存が可能な、クラウド時代のオブジェクトストレージです。

- Amazon S3の位置付け
- Amazon S3の概要
- Amazon S3へのアクセス
- Amazon S3のデータ保護
- Amazon S3のデータ管理
- Amazon S3パフォーマンス最適化
- Amazon S3の料金

様々なAWSサービスと連携し、利用者のAWS利用を支えてくれるストレージ・データストア

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



ご視聴ありがとうございました