る仕組みを備えています。このため、AZの障害時には影響を受けますが、単一のディスク障害は回避できます。

EBSは、以下の機能を備えています。

●複数のボリュームタイプ

EBSには複数のボリュームタイプが用意されており、ユーザーは使用容量、コスト、パフォーマンスの観点から選択することができます。

以下に、主なボリュームタイプを説明します。EBSのボリュームタイプの詳細は「4-4　ストレージサービスにおけるパフォーマンス」を参照してください。

● 汎用 SSD（General Purpose SSD：gp2 および gp3）

デフォルトで提供されるボリュームタイプで、SSDを安価に利用することができます。確保した容量に応じてIOPSが設定されており、OSのルート領域や、高性能なI/Oを要求されないデータ領域で利用されます。

● プロビジョンド IOPS SSD（PIOPS SSD：io1 および io2）

高パフォーマンスを実現できるボリュームタイプで、SSDをベースにユーザーが自由にIOPSを設定して利用できます。汎用SSDより高いパフォーマンスが求められる場合に利用されます。ストレージ容量に加えて、指定したIOPSに対しても課金されます。

● スループット最適化 HDD（st1）

HDDタイプのボリュームタイプで、大容量のストレージを安価に利用できます。

● コールド HDD（sc1）

アクセス頻度が低い大量データの保存に適したHDDです。同じHDDタイプのst1より安価に利用できます。

● マグネティック

旧世代のボリュームタイプです。アクセス頻度が低い用途に適したHDDです。

試験対策

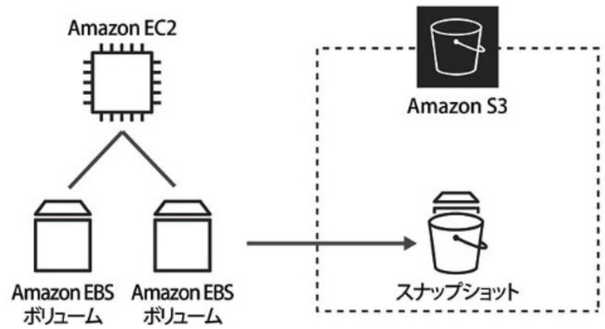
Amazon EBSは、使用容量、コスト、パフォーマンスの観点から適切なボリュームタイプを選択することが重要です。

●スナップショットによるバックアップ

EBSでは、**スナップショット**を用いてバックアップを取得することができます。スナップショットは自動的にS3へ保管されるため、高い耐久性を備えています。

オンプレミス環境のストレージでは、バックアップを取得する際に取得完了まで待つ必要がありますが、EBSスナップショットの場合は取得を実行した時点のものが保管されるため、完了を待つ必要はありません。また、スナップショットは、初回はフルバックアップを行いますが、以後は増分で取得します。

【EBSスナップショット】



●保管データのセキュリティ

EBSは、主にAmazon EC2で利用するため、VPCやセキュリティグループなどのネットワークアクセス制御を行うことで、データのアクセス制御を行います。

また、EBSボリュームはKMSを利用して暗号化することが可能で、ボリューム内のデータや暗号化したボリュームから作成したスナップショットを暗号化することができます。なお、EBSボリュームは作成時に暗号化する必要があり、作成後には暗号化できません。

作成済みEBSボリュームを暗号化する場合は、以下のいずれかの方法で暗号化することができます。

・一度スナップショットを作成し、スナップショットからEBS復元時に暗号化を指定する

・一度スナップショットを作成し、スナップショットを暗号化してコピーしたあと、暗号化したスナップショットからEBSを復元する

3 Amazon Elastic File System（EFS）

Amazon EFSはスケーラブルな共有ストレージサービスで、複数のEC2インスタンスやECSコンテナなどの共有ファイルストレージとして利用することができます。

以下に、EFSの特徴を説明します。

●高可用性

EFSは複数のAZに冗長化してデータを保管します。そのため、AZを横断してファイルにアクセスでき、1つのAZに障害が発生してもアクセス不可にはなりません。

●自動スケーリング

EFSはマネージド型のサービスで、ストレージ容量やパフォーマンスを自動的にスケーリングする機能を備えています。

ファイル削除時は自動で縮小されるため、EBSより柔軟に利用することができます。

●標準的なファイルシステム

EFSは、標準的なファイルシステムであるNFS[※16]によってアクセスします。EC2インスタンスだけでなくオンプレミス環境のサーバーからも利用できます。

●ストレージクラス

Amazon S3と同様に、EFSにもストレージクラスが用意されており、用途によって使い分けることでコストを削減することができます。

● 標準（スタンダード）

標準ストレージクラスでは、複数のAZにデータを冗長化して保存するため、高可用性を実現することができます。

※16 【NFS】Network File System：LinuxなどUNIX系のOSで利用されるファイル共有システムのこと。

● 標準 - 低頻度アクセス（Standard-Infrequent Access）

標準ストレージクラスと同じ耐久性を備えたストレージクラスで、スタンダードと比較して低コストで利用可能です。ただし、データの読み取りに対して課金されるため、長期保管やバックアップ用に利用されます。

● 1ゾーン

EFSの標準ストレージクラスは複数のAZにデータを冗長化して保存しますが、1ゾーンのみで管理することで低コストで利用できます。耐久性を要求されないものの頻繁にアクセスが発生する一時ファイルなどの利用に適しています。

● 1ゾーン - 低頻度アクセス（One Zone-Infrequent Access）

標準低頻度アクセスほど耐久性を必要としない場合に利用するストレージクラスです。1カ所のAZのみにデータを保存するため、標準低頻度アクセスよりさらに安価で利用することができます。

試験対策

Amazon EFSはNFSでアクセスして利用する、という点が重要です。

4 Amazon FSx

Amazon FSxは、コスト効率のよいファイルシステムを提供するサービスです。

通常、ファイルサーバーなどのファイルシステムを導入する場合は、サイジングや冗長性、ソフトウェア管理やバックアップなど多岐にわたる事項を検討する必要があります。FSxはマネージドサービスであるため、これらの検討事項を簡素化するのに有効です。

以下に、FSxの特徴を説明します。

●高可用性と高耐久性

FSxはマルチAZに対応しているため、高い可用性で利用できます。また、Amazon S3へのバックアップも可能なため、より高い耐久性を実現することができます。

●高速で柔軟なパフォーマンス

データ保管用ストレージとしてHDD(ハードディスクドライブ)とSSD(ソリッドステートドライブ)に対応しており、システム要件に応じて最適なパフォーマンスを選択することができます。

●セキュリティ保護

FSxに保管されたデータや通信中のデータは、すべて暗号化されます。また、IAMポリシーによるAPIコールのアクセス制御や、セキュリティグループを使ったアクセス制御を行うことも可能です。

●コスト最適化

FSxは、ストレージタイプとしてHDDとSSDを選択できるため、用途に応じてコストとパフォーマンスを最適化することができます。また、ユーザー単位や部門単位などのクォータを設定してストレージの使用量を制限することができ、適切な予算で運用できます。

●FSxのファイルシステム

FSxでは、以下に説明する2種類のファイルシステムが多く利用されています。

● FSx for Windows File Server

主にWindowsサーバーで使われているSMB[※17]プロトコルを利用して、データへのアクセスが可能なフルマネージドファイルサーバーです。Windowsのファイルサーバーとして利用されることが前提のため、Active Directoryとの連携やWindowsのACL(Access Control List)によるアクセス制御が可能です。

● FSx for Lustre

機械学習やHPC(High Performance Computing)など、処理速度を重視するシステムに利用されるLinuxベースのファイルシステムです。データ処理用に最適化されているため、非常に高いパフォーマンスを発揮します。また、Linuxのコマンドで実行できるため、Linuxベースのアプリケーションに変更を加えることなく利用できることも特徴です。

※17 【SMB】Server Message Block：主にWindows OSのネットワークでファイル共有やプリンタ共有などに使用される通信プロトコルのこと。

FSxには、主にSMBプロトコルで利用するFSx for Windows File Serverと機械学習やHPCで利用されるFSx for Lustreがあることを覚えておきましょう。なかでも、FSx for Windows File ServerがActive Directoryと連携できるという点は重要です。

FSxでは、本文で説明したファイルシステム以外にも、FSx for NetApp ONTAPとFSx for OpenZFSと呼ばれるファイルシステムが提供されています。
FSx for NetApp ONTAPは、ネットアップ社のデータ管理ソフトウェアであるNetApp ONTAP上に構築されたストレージで、オンプレミス環境からスムーズな移行を可能にしたり、災害対策のバックアップ／レプリケーションなどに利用できます。
FSx for OpenZFSは、NFSで最高クラスの低レイテンシー ストレージサービスで、機械学習や高度なデータ分析などの用途に向いています。

5 その他のストレージサービス

●インスタンスストア(エフェメラルディスク)

Amazon EC2の特定のインスタンスタイプで利用できる無料のストレージサービスです。ホストコンピュータのローカル領域を使用しているため高パフォーマンスですが、揮発性のディスクであるため、EC2インスタンスの停止時にはデータが消失します。そのため、重要なデータの場合は保管方法を別途検討する必要があります。

●AWS Storage Gateway(SGW)

NFSやSMB、iSCSI[※18]といった標準プロトコルでAmazon S3やAmazon FSxにアクセスできるようにするサービスで、主にオンプレミス環境からAWSのストレージを使いたい場合に利用します。Storage Gatewayは、EC2インスタンスだけでなくオンプレミスサーバーにS3をマウントして利用することも可能です。

SGWのゲートウェイタイプには、次の種類があります。

※18 【iSCSI】Internet Small Computer System Interface：データ転送にTCP/IPを使用する仕組み。ネットワーク上のストレージに直接接続できるメリットがある。

● キャッシュ型ボリュームゲートウェイ

データはS3に格納しますが、アクセス頻度の高いデータはローカルにキャッシュすることで高速のアクセスを実現します。インターフェイスにはiSCSIを使用します。

● 保管型ボリュームゲートウェイ

データをローカルに格納し、非同期でデータ全体をスナップショットとしてS3に格納します。インターフェイスにはiSCSIを使用します。

● テープゲートウェイ

物理テープ装置の代替としてデータをS3(Glacier)に格納します。インターフェイスにはiSCSIを使用します。

● S3 ファイルゲートウェイ

データをオブジェクトとして、直接S3に格納します。インターフェイスにはNFSおよびSMBを使用します。

● FSx ファイルゲートウェイ

データをFSxに格納します。インターフェイスにはSMBを使用します。

> **試験対策**
> Storage Gatewayは、主にオンプレミス環境からAWSのストレージを使いたい場合に利用し、5種類のゲートウェイタイプから選択することを覚えておきましょう。

●AWS Snow Family

AWS Snow Familyは、大容量のデータをコスト効率よく移行できるAWSサービス群で、デバイスをNFSプロトコルで利用し、データを暗号化して保管できます。主に、AWSから物理搬送される耐久性の高い筐体に大容量のデータを保管し、AWS内のストレージへ転送してデータを移行します。

物理デバイスの受領からデータの移行、物理デバイスの返送まで数週間のリードタイムは発生しますが、ネットワーク回線による転送では時間がかかりすぎる場合や、ネットワーク回線のないロケーションからデータを移行する場合などに利用されます。

Snow Familyには、以下のサービスモデルがあります。

● AWS Snowcone

Snowconeは、コンパクトで可搬性が高いデバイスです。HDDの場合は8TB、SSDの場合は14TBのストレージが利用可能で、移動デバイスであるIoTや車載データの保管に向いています。AWSへデバイスを返送してデータを転送したり、AWS DataSyncを利用してオンラインでデータ転送することができます。

● AWS Snowball

Snowballは、Edge Storage OptimizedとEdge Compute Optimizedの2種類のオプションを選択できるデバイスです。

Edge Storage Optimizedは、最大80TBのストレージが利用可能で、大容量のローカルストレージや大規模なデータ移行に適しています。

Edge Compute Optimizedは、最大42TBのストレージが利用可能で、高度な機械学習データの収集や、遠隔地にある軍事基地からのデータ収集に適しています。

● AWS Snowmobile

Snowmobileは、Snowballを超える大容量データをAWSへ移行するケースに利用するサービスです。最大100PBのストレージが用意されており、データは、デバイスではなく大規模な専用トレーラーを利用して移行されます。

> **試験対策**
> Snow Familyは、主に大容量のデータ移行に利用しますが、物理搬送が必要になります。そのため、データ移行には時間を要することに注意しましょう。

●AWS Transfer Family

AWS Transfer Familyは、SFTPやFTPS、FTPでAmazon S3やAmazon EFSと通信して、直接ファイルをやり取りできるようにするサービス群です。

フルマネージドサービスであるため、通信に必要なインフラストラクチャやストレージを用意する必要がなく、高い可用性と拡張性を備えています。

用意されているプロトコルは次のとおりです。

- AWS Transfer for SFTP
- AWS Transfer for FTPS
- AWS Transfer for FTP

AWS Backup

AWS Backupは、AWS内におけるデータのバックアップを一元化／自動化するサービスです。バックアップポリシーを設定することで、たとえば、EBSボリュームやEC2インスタンス、RDSデータベースなどのバックアップを定期的に実行したり、バックアップ状況を監視することができます。

AWS Backupでサポートされている AWSサービスは、次のとおりです。

- Amazon EC2
- Amazon EBS
- Amazon S3
- Amazon RDS
- Amazon DynamoDB
- Amazon Neptune
- Amazon DocumentDB
- Amazon EFS
- Amazon FSx
- AWS Storage Gateway
- Amazon Outposts（オンプレミス環境）

Q 演習問題

1 **あるシステムのバックアップファイルをAmazon S3に保管して運用していたところ、チームメンバーが誤ってS3のデータを削除してしまいました。今後誤った削除を防止するためには、次のうちどの設定を行うことが適切ですか。**

A　ストレージクラスを変更する

B　ライフサイクルポリシーを有効化する

C　オブジェクトロック機能を有効化する

D　署名付きURLを発行する

2 **ある会社では、データストレージにNFSプロトコルを利用したアプリケーションをAWSへ移行したいと考えています。次のうち、移行先のストレージとして適切なサービスはどれですか。**

A　Amazon S3

B　Amazon EFS

C　Amazon FSx for Windows File Server

D　AWS Snowball

3 **あるアプリケーションは、参照する動画データをAmazon S3に保管しています。動画データは、年に数回しかアプリケーションから参照されませんが、すぐに取り出して利用できる必要があります。データの可用性を損なうことなくコストを削減するためには、次のうちどの方法が適切ですか。**

A　S3バケットのストレージクラスを、標準低頻度アクセスに変更する

B　S3バケットのストレージクラスを、Glacier Deep Archive に変更する

C　S3バケットにライフサイクルポリシーを設定する

D　S3バケットのバージョニング機能を有効化する

A 解答

1 C

Aのストレージクラスは、データの耐久性やアクセス頻度に応じてS3のコストを最適化する機能です。
Bのライフサイクルポリシーは、設定した期間が経過したあと、オブジェクトを自動で削除したりアーカイブしたりすることができる機能です。
Cのオブジェクトロック機能は、オブジェクト操作を制限し、意図しないオブジェクトの上書きまたは削除の操作を防止することができる機能です。
Dの署名付きURLは、S3上のデータに対して一定時間だけアクセスを許可するためのURLを発行する機能です。
したがって、**C**が正解です。

2 B

AのAmazon S3は、高耐久・大容量のオブジェクトストレージサービスで、マネジメントコンソールやSDKから利用するため、NFSプロトコルのストレージとして利用できません。
BのAmazon EFSは、スケーラブルな共有ストレージサービスで、NFSプロトコルを利用してアクセスします。
CのAmazon FSx for Windows File Serverは、主にWindowsサーバーで使われているSMBプロトコルを利用してアクセスします。
DのAWS Snowballは、大容量のデータをコスト効率よく移行することができるAWSサービスで、アプリケーションのストレージとして利用するものではありません。
したがって、**B**が正解です。

3 A

Aの標準低頻度アクセスのストレージクラスは、スタンダードと同じ耐久性を備えているストレージクラスで、スタンダードと比較して低コストで利用できます。
BのGlacier Deep Archiveは、年に1、2回程度しかアクセスされないようなデータに最適なアーカイブストレージクラスですが、データの取り出しには数時間以上かかります。

Cのライフサイクルポリシーは、設定した期間が経過したあと、オブジェクトを自動で削除したりアーカイブできる機能です。
Dのバージョニング機能は、オブジェクトの世代管理を提供する機能で、世代管理を行う分だけコストが発生します。
したがって、**A**が正解です。

1-6 データベースサービス

AWSでは非常に多くのデータベースサービスが提供されており、各サービスの特徴や扱うことができるデータ構造の違いを理解することが重要です。
本節では、AWSで提供されているデータベースサービスの概要を説明します。

1 AWSのデータベースの特徴

AWSの特徴の1つとして、マネージドサービスがあげられます。AWSにおけるマネージドサービスとは、AWS上のインフラストラクチャ環境やミドルウェア環境をAWSが管理するサービスの形態を指します。AWSで提供されているデータベースサービスは、すべてマネージドサービス型のデータベースです。

本来、データベースはデータを格納するために使用するものであり、データベースの設計に携わるエンジニアは、データ格納用のテーブルやカラムの設計にできるだけ時間を費やしたいと考えるものです。

しかし、通常業務で利用するデータベースサーバーを設計・構築・運用する際には、以下のような作業も必要になります。

- データベースサーバー用ハードウェアに関係する要件定義
- ハードウェアの調達、サーバーへのデータベースのインストール
- 構築時や運用開始後のパッチ適用
- データベースのバックアップ／リストアの設計および実際の作業
- DR対策、など

AWSのデータベースサービスでは、前述のような作業をAWSが提供する機能に任せることができます。たとえば、次のような特徴があります。

- 構築時点からデータベースがインストールされており、さらに初期設定も完了している状態で提供されるため、すぐにデータベースの機能を利用できる
- データベースサーバーが動作しているOSや、データベースソフトウェアへのパッチ適用をAWSが行う
- 自動バックアップが可能
- データベースのスケールアップやスケールアウトは、AWSの機能を利用して行うことができる

これによって、データベースのテーブル設計など、本来の業務に多くの時間を充てることができます。

AWSには、さまざまな目的や用途で利用できるように、複数のデータベースサービスが用意されています。次項から、AWSで提供されているデータベースサービスについて説明します。

2 Amazon Relational Database Service（RDS）

Amazon RDSは、マネージドサービス型のリレーショナルデータベースサービスです。

リレーショナルデータベースにおける一連のデータ処理をトランザクションといい、すべてのトランザクションは、データの完全性を確保するため、次に示す4つの特性（ACID特性）を備えていなければなりません。

ACIDとは、Atomicity（原子性）、Consistency（一貫性）、Isolation（独立性）、Durability（耐久性）の4つの単語の頭文字を取った略語で、「アシッド」と読みます。

【ACID特性】

各特性を表す単語	説　明
Atomicity（原子性）	トランザクションはすべて正常に実行されるか、無効になるかのどちらかであること。トランザクションの一部が失敗した場合は、トランザクション全体が無効になる
Consistency（一貫性）	データベースに書き込まれたデータが、定義されているすべてのルールと制約（制限、カスケード、トリガーなど）に従っていること
Isolation（独立性）	各トランザクションがそれぞれ独立し、並行処理を実現すること
Durability（耐久性）	トランザクションの正常完了後、データベースに加えられたすべての変更が永続的な状態になること

RDSは、以下のデータベースエンジンをサポートしています。（2022年12月時点）

- Amazon Aurora
- PostgreSQL
- MySQL
- MariaDB
- Oracle Database
- Microsoft SQL Server

Aurora以外の5つのデータベースエンジンは、AWS以外の環境でも利用されている一般的なデータベースです。それぞれの特徴については各データベースエンジンの公式マニュアルなどを参照してください。Auroraは、AWS固有のデータベースサービスです。

●パッチ適用の管理負荷軽減

RDSでは、曜日や時間帯をメンテナンスウィンドウで指定するだけで、AWSが自動でパッチを適用します。通常は数カ月に1回のペースでパッチがリリースされますが、パッチのリリース後は、最初にメンテナンスウィンドウで設定した曜日や時間帯になった時点で自動適用されます。

マルチAZ（95ページを参照）を利用している場合は、まずスタンバイデータベースのメンテナンスを実施し、そのスタンバイデータベースをプライマリ

データベースとして昇格させます。この一連の処理が終了すると、メンテナンス前のプライマリデータベースのメンテナンスを実施し、これをスタンバイデータベースとして維持します。

このようにスタンバイデータベースを利用することで、ダウンタイムのないメンテナンスを実施することができます。

セキュリティに関わるパッチがリリースされた場合は、早急に対応する必要があります。
緊急でパッチを適用する場合は、メンテナンスウィンドウで設定した時間とは関係なく、ユーザーのシステム利用を調整するなどの対策が必要です。

パッチ適用以外に、DBインスタンス ストレージの変更やDBインスタンスタイプの変更なども、マルチAZを利用することで、ダウンタイムを短縮することができます。ただし、DBエンジンのバージョンアップグレードは、スタンバイデータベースとプライマリデータベースを同時に更新するため、ダウンタイムが発生します。

●バックアップとリストア

標準設定の場合、1日1回自動バックアップが実施され、データベースとトランザクションログが取得されます。

データベース（データベースインスタンス）のバックアップにはストレージボリュームのスナップショットを作成し、データベースインスタンス全体を取得します。

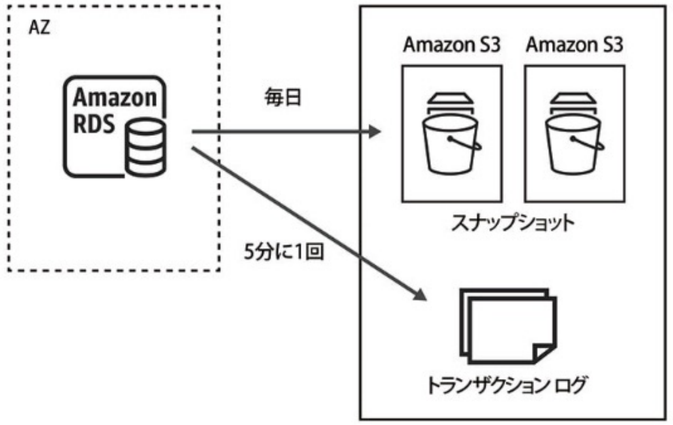
保存期間を設定することが可能で、最大で35日分を保存することができます。また、自動ではなく手動で任意の時間にバックアップを取得することも可能です。

RDSの復元方法には、主に次の2つがあります。

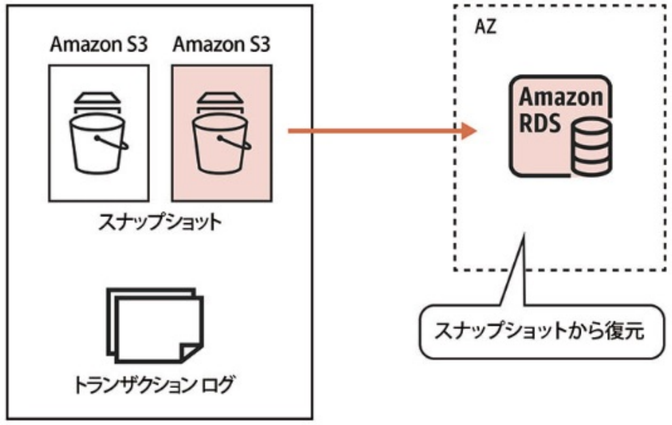
- 通常のリストア：スナップショットからデータベースインスタンスを復元
- ポイントインタイムリカバリ：バックアップのデータベースとトランザクションログが残っている範囲内の特定の日時時点にデータを復元

RDSでは、データベースインスタンスのトランザクションログを5分ごとにS3に保存しています。したがって、最新のデータを復元すると過去5分以内の状態にデータベースインスタンスを戻すことができます。

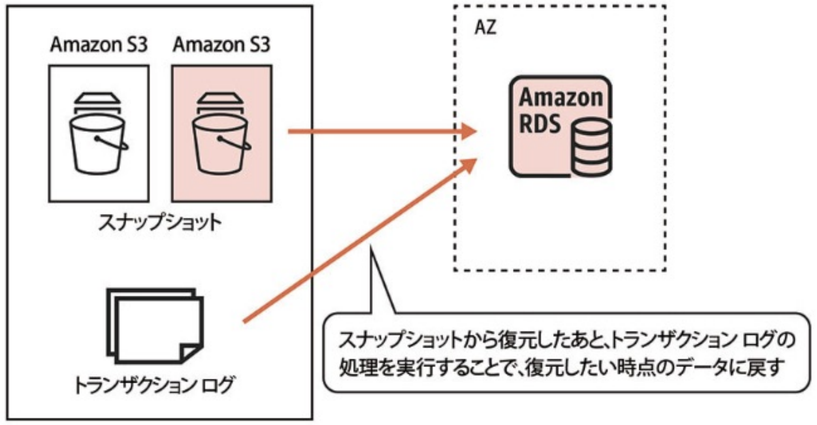
【RDSのバックアップイメージ】



【RDSの復元イメージ(通常のリストア)】



【RDSの復元イメージ(ポイントインタイムリカバリ)】



データベースを過去の状態に戻すことになった場合は上述の方法で復元します。このとき、新しくデータベースインスタンスを作成・起動することになるため、アプリケーション側は新しいエンドポイントに接続するように修正する必要があります。

バックアップのスナップショットはリージョン間でコピーすることができるため、DR対策として利用することもできます。

> 参考
> RDSの各インスタンスには、「エンドポイント」と呼ばれるホスト名が設定されます。アプリケーションからRDSのインスタンスへ接続するには、このエンドポイントを指定して接続します。

> 参考
> RDSではデータベースエンジンごとにオプショングループがあり、データベースの追加機能を設定できます。また、ほかにもパラメータグループがあり、データベースパラメータを設定することで、割り当てるメモリ容量などを指定することができます。
> ただ、これらの設定は自動バックアップ対象には含まれていないため、RDSの自動バックアップ設計を行う際は注意が必要です。

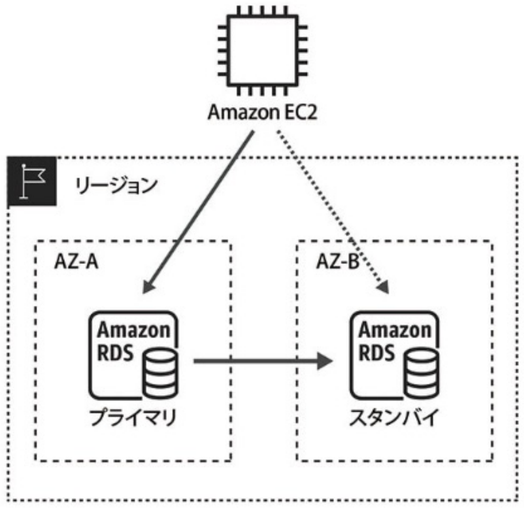
● マルチAZ

RDSのマルチAZとは、RDSを複数のAZに構築することを指します。

Aurora以外のRDSデータベースでは、マルチAZのデータベースインスタンスを作成すると、RDSはプライマリのデータベースインスタンスを自動的に作成し、それと同時に異なるAZにあるスタンバイのデータベースインスタンスにもデータを複製します。

マルチAZにすることで、耐久性と可用性が向上します。

【マルチAZのデータ複製のイメージ】



シングルAZでは、データベースに障害が発生した場合にバックアップを利用して復元する必要があります。障害発生直前までデータを戻したくても、データは直近5分以内にしか戻せないため、それよりも新しいデータは復元できません。また、バックアップからの復元に数時間を要することがあります。

マルチAZではデータベースに障害が発生した場合、最新のスタンバイデータベースへのフェイルオーバー(Auroraの場合はレプリカへのフェイルオーバー)が自動で開始されるため、データの欠損はほとんど起こらず、データベースの切り替えもすぐに完了します。

ただし、シングルAZよりマルチAZのほうがリソースを消費するためコストがかかります。

参考

マルチAZ構成と似た仕組みとして、マルチAZ DBクラスターと呼ばれる構成も存在します。マルチAZ構成はデータベースの耐久性と可用性の向上を目的として構築しますが、マルチAZ DBクラスター構成にすると、1つの書き込みDBインスタンスと2つの読み取りDBインスタンスが配置されるため、より高い可用性と読み取りワークロードの向上、低レイテンシーが実現できます。マルチAZ DBクラスター構成は、MySQLエンジンとPostgreSQLエンジンのみサポートされています。(2022年12月時点)

試験対策

Amazon Auroraで障害が発生した場合は、レプリカをプライマリインスタンスに昇格させてフェイルオーバーしますが、どのレプリカを昇格させるかはレプリカに設定された優先度をもとに判断されます。優先度が同じ場合は、インスタンスサイズが最大のレプリカを昇格させ、サイズも同一の場合は任意のレプリカが昇格されます。AuroraとAurora以外のデータベースのフェイルオーバーの違いと仕組みを理解しておきましょう。

● リードレプリカ

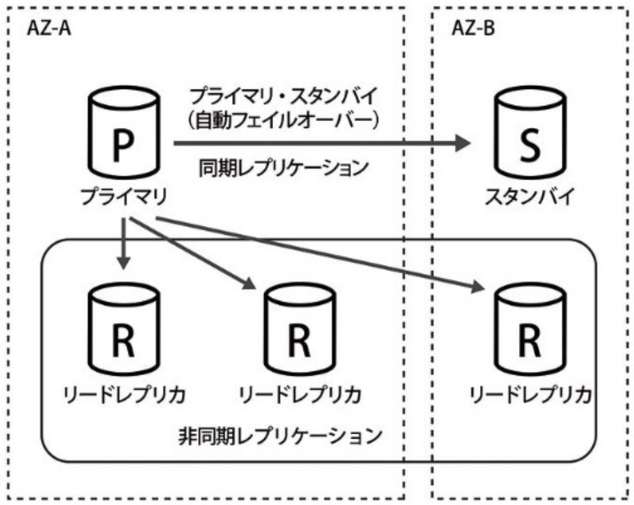
プライマリデータベースを複製し、読み取り専用として構築したデータベースを**リードレプリカ**といいます。リードレプリカを利用することにより、読み取り頻度の高いデータベースを増設できるため、読み取りスループットを高めてパフォーマンスを向上させることができます。

リードレプリカはMySQL、MariaDB、PostgreSQL、Oracle Database、Microsoft SQL Server、Auroraで利用できます。

リードレプリカは、マルチAZや異なるリージョン間で構築できます。また、プライマリデータベースと異なるインスタンスタイプやストレージで構築することもできます。

プライマリデータベースとの同期処理は非同期に実行されます。

【リードレプリカのイメージ】



試験対策

プライマリデータベースとリードレプリカの構成にすることで、読み取り要求と書き込み要求の処理を分離することができ、パフォーマンスと耐久性が向上します。

クロスリージョンリードレプリカ

Amazon RDSでは、DBインスタンスのリードレプリカを異なるリージョンに作成できます。異なるリージョンにレプリカを作成することによって、災害対策やユーザーに近いリージョンでのDBアクセスが可能になります。

クロスリージョンでリードレプリカを作成できるデータベースエンジンはMariaDB、MySQL、Oracle Database、PostgreSQLの4種だけで、PostgreSQL以外はソースDBインスタンス削除時に自動的にプライマリDBインスタンスに昇格します。PostgreSQLは自動的に昇格しませんが、手動で昇格させることができます。

クロスリージョンレプリカを使用することで、リージョンの障害時に迅速に別リージョンに切り替えて利用できるようになります。

ただし、その分だけリージョン間の転送コストが発生したり、プライマリ昇格までに数分を要するためダウンタイムが発生する、などの制約事項もあります。

暗号化

Amazon RDSは、データを保管するDBインスタンスや自動バックアップ、リードレプリカ、スナップショットなどを、AES-256暗号化アルゴリズムを使用してサーバー側で暗号化できます。暗号化にはAWS Key Management Service(AWS KMS)のキーだけでなく、AWSマネージドキーやカスタマーマネージドキーも利用できます。

ただし、RDSの暗号化はDBインスタンスの作成時にのみ設定でき、作成後は暗号化できないことが注意点としてあげられます。そのため、DBインスタンス作成後に暗号化するには、次に示す手順で行う必要があります。

① 非暗号化状態のRDS DBインスタンスのスナップショットを取得する
② スナップショットのコピーを取得し、その際に暗号化を設定する
③ 暗号化設定したスナップショットからDBインスタンスを起動する

試験対策

Amazon RDSのDBインスタンス作成後の暗号化手順を覚えておきましょう。

SSL接続

AWSでは、SSLを使用してアプリケーションからRDS(MySQL、MariaDB、Microsoft SQL Server、Oracle Database、PostgreSQL)のDBインスタンスに接続することができます。

RDS Proxy

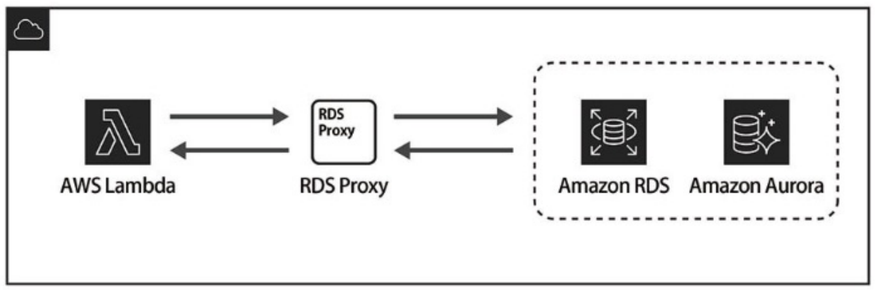
Lambdaなど複数のクライアントアプリケーションからRDSやAuroraに接続する場合、データベースへのコネクションをその都度作成していると最大接続数の上限に達してしまい、接続エラーが発生することがあります。

そのようなエラーを解消する方法としてRDS Proxyがあります。RDS Proxyは、RDSやAuroraデータベースとの接続の開閉や保持を、アプリケーションに代わって実行することで制御し、接続に伴うメモリやCPUのオーバーヘッド、TLS/SSLハンドシェイクなどもRDS Proxyインフラストラクチャ内で処理するため、データベースサーバーのオーバーヘッドを減らすことができます。

PDS Proxyの機能と特徴については、以下を参照してください。

https://docs.aws.amazon.com/ja_jp/AmazonRDS/latest/AuroraUserGuide/rds-proxy.howitworks.html

【RDS Proxyのイメージ】



試験対策

RDS Proxyの機能と特徴について覚えておきましょう。

Amazon Aurora

Amazon Auroraは、AWSが開発したリレーショナルデータベースサービスです。PostgreSQLやMySQLと互換性があるため、これらのデータベースを利用していた場合は、アプリケーションをほとんど変更することなくAuroraにデータを移行することができます。

性能は、PostgreSQLやMySQLよりもAuroraのほうが優れており、MySQLやPostgreSQL以上のスループットを実現できるとされています。

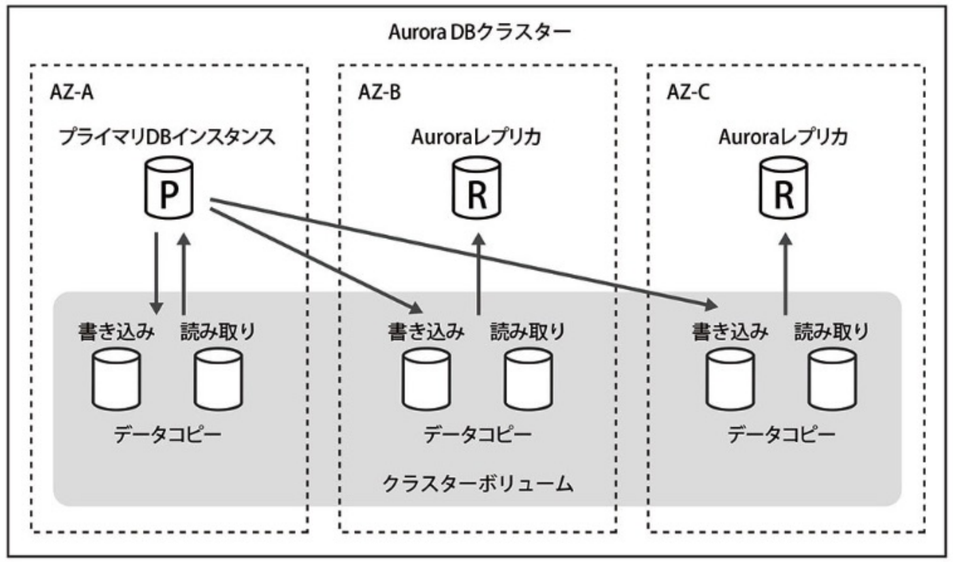
Auroraは、SQL処理を行う「データベースインスタンス」と、データを格納するストレージ部分である「クラスターボリューム」に分かれています。

データが格納されるクラスターボリュームは、3カ所のアベイラビリティーゾーン(AZ)に存在します。3つのAZのうち1つのデータベースインスタンスは、「プライマリデータベース(DB)インスタンス」と呼ばれるインスタンスを持ちます。このプライマリデータベース(DB)インスタンスは、データの読み取り・書き込み処理を行います。

そのほかの2つのデータベースインスタンスは「Auroraレプリカ」と呼ばれ、読み取り処理を行います。1つのAZにつき2カ所のディスクに書き込まれるため、合計で3つのAZの6カ所に保存されます。このようにデータが格納されることから、障害発生時にダウンタイムが発生することはほとんどありません。

次の図はAuroraの構成を表したイメージです。

【Auroraの構成イメージ】



以降では、Auroraを含むRDSサービスの主な特徴について説明します。

試験対策

Amazon Auroraの特徴は重要ですので、すべて覚えておきましょう。

Aurora マルチマスタークラスター

前段で、プライマリデータベースインスタンスはデータの読み取り／書き込み処理を行い、Auroraレプリカは読み取り処理を行うと説明しました。このデータの読み取り／書き込み処理を、すべてのデータベースインスタンスで行う構成が**Auroraマルチマスタークラスター**です。

この構成では、読み取り専用のデータベースが存在しないため、特定のデータベースインスタンスが利用できなくなっても、フェイルオーバーは発生しません。

しかし、アプリケーションからデータベースへの書き込み処理時に、データ書き込みの競合が発生する可能性があるため、書き込み処理は特定のエンドポイントに対して行うことが推奨されています。

読み取り専用のデータベースはありませんが、カスタムエンドポイントを使用していくつかのデータベースインスタンスをグループ化し、カスタムエンドポイントに対して読み取るように設定することもできます。

試験対策

Amazon Auroraのマルチマスタークラスターで利用可能なエンドポイントの種類を覚えておきましょう。

●Amazon Aurora Global Database

Amazon Aurora Global Databaseを使用すると、複数リージョンにAuroraクラスターをデプロイすることができます。Aurora Global Databaseでは、特定の1つのリージョンにプライマリDBクラスターが配置され、それ以外の異なるリージョンに最大5つまでセカンダリDBクラスターを配置することができます。

データの書き込みはプライマリDBクラスターのみで行われ、セカンダリDBクラスターにはデータが保存されたクラスターボリュームがレプリケートされます。

DBインスタンス自体をレプリケートしないため、低レイテンシーでのレプリケーションを実現します。また、レプリケートされたセカンダリDBクラスターは、読み取り処理のみサポートされています。

Aurora Global Databaseの詳細については、次のWebページを参照してください。

https://docs.aws.amazon.com/ja_jp/ja_jp/AmazonRDS/latest/AuroraUserGuide/aurora-global-database.html#aurora-globaldatabase.limitations

試験対策

Amazon Aurora Global Databaseは高い可用性と耐障害性を持つ一方で、Aurora MySQLとAurora PostgreSQLの特定バージョンだけしか対応していない、などの制約事項もあります。どのような制約があるか確認しておきましょう。

●Aurora Serverless

データベースの用途によっては、必要なときだけ起動して処理できればよい場合があります。また、処理に対する負荷によっては、必要に応じてスケールしてデータを処理したい場合などもあります。

前述したように、データベースインスタンスとクラスターボリュームを持つAuroraというデータベースサービスが提供されていますが、このサービスにはAurora Serverlessというフルマネージドサービスもあります。

Aurora ServerlessはAuroraと異なり、通常はデータベースインスタンスが起動しておらず、データが格納されているクラスターボリュームだけが存在します。SQLのリクエストを受け取って初めてデータベースインスタンスが起動し、SQLの処理を行います。

処理の負荷が高いときは、負荷に応じてデータベースインスタンスを自動でスケールして適切に処理することができます。

Aurora Serverlessのユースケースについては、次のWebページを参照してください。

https://docs.aws.amazon.com/ja_jp/AmazonRDS/latest/AuroraUserGuide/aurora-serverless-v2.html

試験対策

Aurora Serverlessのユースケースを覚えておきましょう。このサービスは、利用頻度があまり高くないアプリケーションや、負荷が一定ではないアプリケーションのデータ格納に適しています。

3 Amazon DynamoDB

Amazon DynamoDBは、マネージド型のNoSQLデータベースサービスです。一般的なリレーショナルデータベース管理システム（RDBMS）で扱うことができるデータを構造化データといいますが、**NoSQL**は非RDBMSの総称で、構造化データ以外のデータ（半構造化／非構造化データ）を扱うこともできます。

【データ構造の種類】

データ構造の種類			説　明	データおよびファイルの例
リレーショナルデータ	構造化データ	データを管理する構造を決め、その構造に合わせて格納したデータ	企業の基幹システムなどで管理されているデータなど	RDB、Excel
非リレーショナルデータ	半構造化データ	非構造化データを管理できる柔軟性のある構造を用意し、その構造に格納されたデータ	システムのログ、センサー（IoT）データ、位置情報、SNSデータ、認定ファイルなど	LOG、JSON、XML
	非構造化データ	構造が定義されておらず、データの関係をモデルに当てはめることができないデータ	テキストファイル（メモ書きなど）、音声データなど	TXT、MPEG

●キーバリュー型データモデル

NoSQLにはいくつかのデータモデルがあります。その代表的なモデルに**キーバリュー型**があります。このデータモデルは非常にシンプルです。1つのデータはキーとバリュー(値)から構成され、データとして書き込む際には、このセットで保存されます。一方、データを呼び出す際には、キーを指定してバリュー(値)を取り出します。

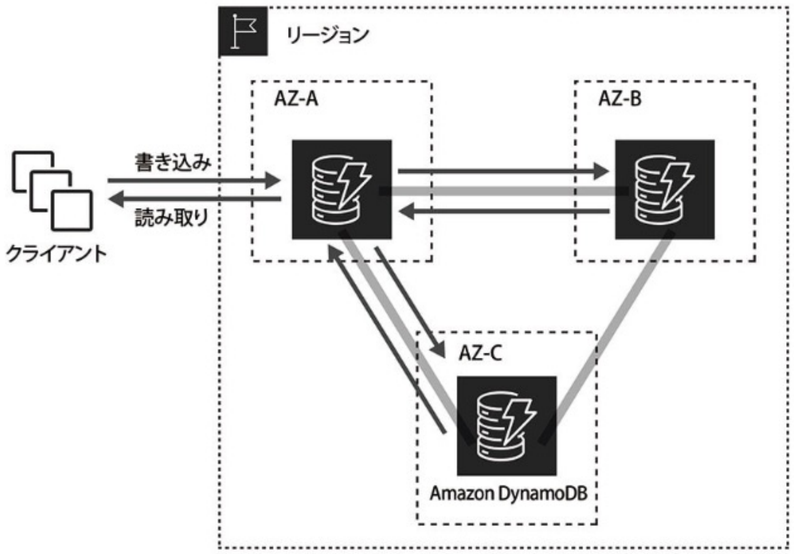
DynamoDBは、キーバリュー型のデータベースです。RDBのように「数値」「文字列」といったシンプルなデータだけではなく、JSONのような半構造化データもバリュー(値)として格納することができます。

> **試験対策**
> Amazon DynamoDBのデータ構造や、格納できるデータの種類を覚えておきましょう。

●マルチAZ

DynamoDBでは、データは自動的に**3カ所のAZ**に保存されます。

【DynamoDBの書き込み／読み取りイメージ】



DynamoDBには特徴的な機能があるため、マルチAZだけではなく次に説明する機能もあわせて理解しておく必要があります。

●結果整合性モデル

DynamoDBを利用するうえで考慮すべき点として、**結果整合性モデル**があります。このモデルでは、書き込んだデータは時間が経てば正しく反映されます(結果的には整合性が保証される)が、データを読み取るタイミングによっては書き込んだデータが反映されない状態になります。

従来のRDBMSでは、トランザクション処理のように厳密な一貫性を保証することができます。前段のマルチAZで説明したとおり、DynamoDBに書き込み処理されたデータは3カ所のAZに分散・保存されますが、この3カ所への書き込みが完了するには時間差があります。そのため、データの読み取りと書き込みが多数発生する場合、タイミングによっては最新の情報が反映されていないAZからデータを取得してしまうことがあります。

このように、「正しいデータが取得できないことがあるのでは機能要件上利用することが難しい」という場合の対応策として、「**強力な整合性**」を備えた読み込みという選択肢が提供されています。これを利用することで、書き込みが反映された最新データを確実に読み取ることができるようになります。

> **試験対策**
> Amazon DynamoDBのデータの保持方法と、整合性の違いを理解しておきましょう。
> **結果整合性**は、処理は高速ですが最新の結果が反映されていないことがあります。
> **強力な整合性**は、確実に最新のデータにアクセスできますが、応答は結果整合性よりも遅くなります。また、ネットワーク遅延が発生した場合はDynamoDBのテーブルにアクセスできなくなる、などの制約条件もあります。

> **参考**
> DynamoDBの3つのAZにデータを書き込む場合、2つのAZに正常に書き込みが完了した時点で、DynamoDBに対するデータの書き込みが正常に完了したと判断し、成功した旨の応答を戻します。
> この「2つのAZへの書き込みが完了し、残り1つのAZはデータ書き込み中」の状態で、まだ書き込みが完了していないAZで、次のデータ読み取り処理が実行されると、最新のデータが取得できない状態になります。

●有効期限

DynamoDB の有効期限を設定すると、レコードの保存期間を設定できます。有効期限はTTL(Time to LIVE)機能を利用して設定し、この有効期限を過ぎるとデータは自動削除されます。

● ポイントインタイムリカバリ

RDSと同様に、DynamoDBもポイントインタイムリカバリの機能を装備しています。ポイントインタイムリカバリを利用することで、人為的なミスによるデータ書き換えや削除などからデータを保護でき、過去35日間の任意の時点にデータを戻すことができます。

●DynamoDBの課金

DynamoDBは、テーブル内データの読み書きや、有効にしているオプションが課金対象になります。

この課金される金額は「**キャパシティモード**」と呼ばれる設定によって異なりますが、モードの選択はユースケースによって判断することになります。

次の表に、2種類のキャパシティモードの特徴を示します。

【キャパシティモード】

キャパシティモード	説　明	適したユースケース
オンデマンド	課金はDBのテーブルに実行したデータの読み書きに対して発生。ワークロードの拡大や縮小に即座に対応できるため、事前のパフォーマンスの予測や設定は不要	●未知のワークロードで新規テーブルを作成し利用する ●DBにアクセスするトラフィックが予測できない ●利用した分だけ支払うほうがよい
プロビジョニング済み	事前に1秒あたりの読み書きの回数を予測し、Auto Scalingを指定すれば、指定した利用率に応じてテーブルのキャパシティが自動的に調整可能	●DBにアクセスするトラフィックが予測できる ●DBにアクセスするトラフィックが一定、または、少しずつ変化する

試験対策

オンデマンドキャパシティモードの特徴とユースケースを覚えておきましょう。

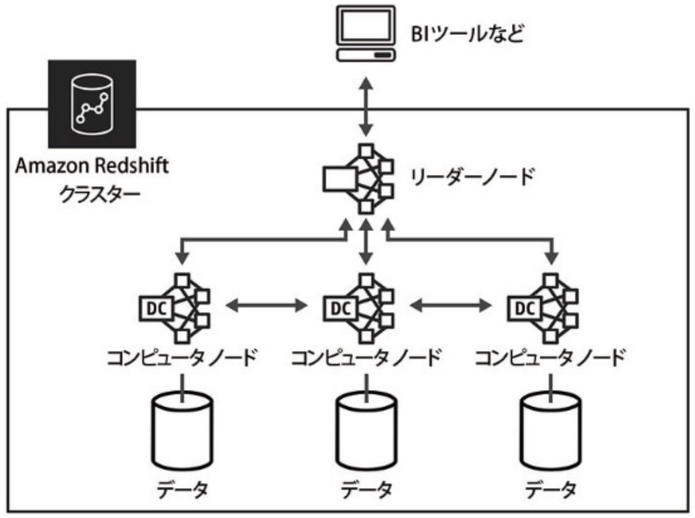
4 Amazon Redshift

Amazon Redshiftは、ペタバイトクラスのデータを扱うことができるマネージドサービス型のデータウェアハウスサービスで、BI(Business Intelligence)ツールなどを利用した大量データの集計・分析に向いています。

データウェアハウスとは、企業活動などの中で蓄積される膨大なデータに対して、ある特定の内容を分析するために利用するシステム、もしくはそのために利用する大量のデータを指します。

Redshiftは、「**ノード**」と呼ばれるコンピューティングリソースの集まりで構成されています。Redshiftに対して処理をする場合は、「**リーダーノード**」と呼ばれるノードが処理を受けます。このリーダーノードは、「**コンピュータノード**」と呼ばれるノードに対してそれぞれ処理を依頼します。これらのノードの集まりを**クラスター**といいます。

【Redshiftの構成イメージ】



●スナップショット

Redshiftは、スナップショットを作成することでバックアップできます。スナップショットは、クラスターのポイントインタイムバックアップです。

スナップショットには自動と手動の2つのタイプがあり、自動スナップショッ

トの設定を有効にしている場合は8時間ごと、またはノードあたり5GBのデータ変更があった場合に自動で取得します。

自動で取得したスナップショットは、設定した保存期間が過ぎるとRedshiftによって削除されるため、永続的にバックアップを保存したい場合は手動でスナップショットを取得します。

スナップショットの格納先は、標準ではRedshiftのクラスターが配置されているリージョンになりますが、別のリージョンにもコピーするように設定できます。これを**クロスリージョンスナップショット**といいます。

スナップショットのコピーを別のリージョンに保存しておくと、リージョンに障害が発生してクラスターやスナップショットが利用できなくなった場合でも、別リージョンのスナップショットからクラスターを復元できます。

> **参考**
> Redshiftの仕様はPostgreSQL 8.0.2に準拠しています。そのため、PostgreSQLを操作するためのpsqlコマンドを利用してRedshiftを操作することができ、一般的なRDBの操作経験者には扱いやすいのですが、両者には大きな違いもあります。
> RedshiftとPostgreSQLの違いについては、以下のURLを参照してください。
>
> https://docs.aws.amazon.com/ja_jp/redshift/latest/dg/c_redshift-and-postgressql.html

●クロスアカウントデータ共有

Redshiftのデータを異なるAWSアカウント間で共有する場合、クロスアカウントデータ共有が利用できます。

データを共有する側をプロデューサー、データを共有される側をコンシューマーと呼び、それぞれのAWSアカウントでデータ共有設定が必要になります。2つのAWSアカウント間、同一リージョンでデータを共有する場合は、次のような設定が必要です。

プロデューサー側

- 共有するデータベース、スキーマ、テーブルの指定
- コンシューマーアカウントの指定、承認

コンシューマー側

- プロデューサーから共有されたデータ情報を参照
- 共有されたデータからコンシューマーアカウント内に新規でデータベースを作成

> **試験対策**
> RedshiftのデータをAWSアカウント間で共有する方法を覚えておきましょう。

> **参考**
> 異なるリージョン間でデータを共有するには、2022年2月17日に一般提供が開始（GA：General Availability）された「クロスリージョンデータ共有」が利用できます。ただし、クロスリージョンデータ共有はRedshift RA3ノードタイプでしかサポートされていないため、注意が必要です。

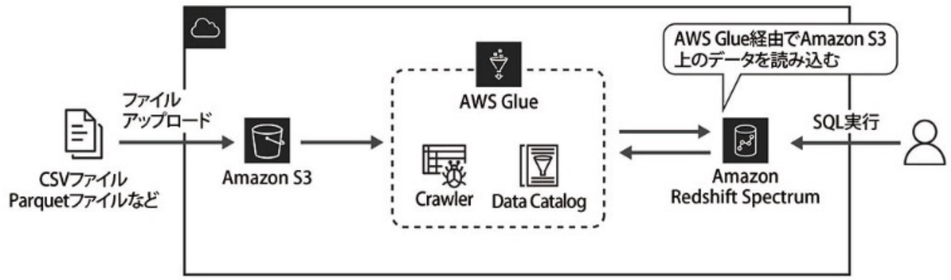
●Redshift Spectrum

Redshiftにデータをロードすると、データを効率よくクエリすることができますが、データが増え続けるとその分だけストレージコストも増加します。そのため、Redshiftのストレージコストの増加を抑えながらもデータをクエリする機能として**Redshift Spectrum**があります。

Redshift Spectrumでは、データの大部分を構造化または半構造化したファイルとしてAmazon S3に格納し、AWS Glueなどの外部データカタログを通してSQLでデータを取得することができます。Redshift Spectrumでクエリされるデータは、Redshift内部にロードしたデータとも結合できるため、高頻度で分析が必要なデータはRedshift内部に保持し、低頻度で必要なデータはS3に保持するといった構成も可能です。

次の図はS3、Glue、Redshift Spectrumを使用した構成例です。

【Redshift Spectrumを利用したシステム構成例】



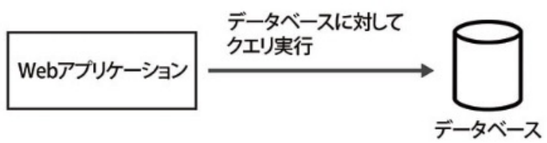
Redshiftの機能をフルマネージドで利用可能なRedshift Serverlessの一般提供が、2022年6月12日に開始されました。これまでは、構築時に必要なノードタイプやノード数などの設定を手動で行う必要がありましたが、Redshift Serverlessを利用することで、それらの設定がAWSで自動的に管理されます。スケーリングなども自動で行われるため、必要なときに必要な数だけスケーリングされ、コストも利用したワークロード分のみに最適化されます。

5 Amazon ElastiCache

Amazon ElastiCacheは、マネージド型のインメモリデータベースです。メモリ上で処理を実行するため、高スループットかつ低レイテンシーな処理を実現できます。

キャッシュを利用しない一般的なWebデータベースアプリケーションの場合は、次のような構成になります。

【一般的なWebデータベースアプリケーションの構成イメージ】



このシステムの場合、データへのアクセス時に常にデータベースに対するクエリが発生し、負荷がかかります。

AWS上でElastiCacheを利用して、次に示す図のような構成にすることで、データベースサーバーの負荷軽減とパフォーマンス向上を実現できます。

【ElastiCacheのアーキテクチャ】

①EC2インスタンスからリクエストされたクエリ結果がElastiCacheに存在すれば、ElastiCacheからEC2インスタンスに結果を返す

②ElastiCacheにクエリ結果が存在しなければ、RDSに対してクエリを実行し、結果を取得する

6 その他のデータベースサービス

Amazon Neptune

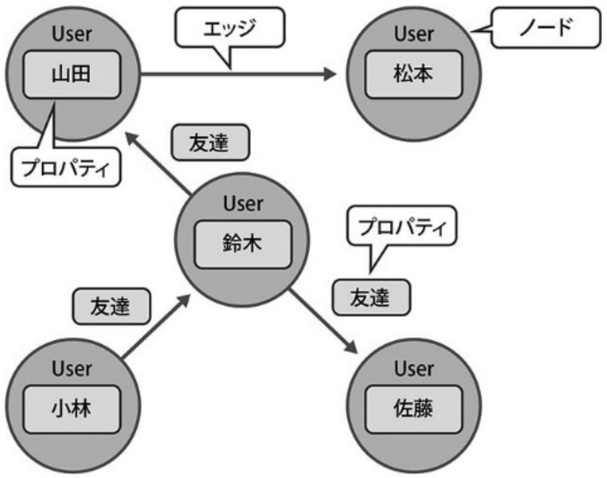
Amazon Neptuneは、高速かつ信頼性の高いフルマネージドなグラフデータベースサービスです。

グラフデータベースは、NoSQLに分類されるデータベースサービスの1つで、DynamoDBのようなキーバリュー型とは異なるデータモデルです。「グラフ」というと、円グラフや棒グラフのようなものを想像しがちですが、ここではモノ同士のつながりを表現するネットワーク状のデータ構造を意味します。

グラフでは、次の3つの要素でデータを定義します。

- ノード：データの要素
- エッジ：ノード間の関係
- プロパティ：ノードとエッジの属性情報

【グラフデータベースの要素と例】



グラフデータベースは、人と人とのつながりや、鉄道や車などの移動時に利用する最短経路検索など、ノード間の関係性の表現や計算に適しています。

このようなデータをリレーショナルデータベースで処理しようとしても、結合データが多すぎるために、気の遠くなるような時間を要することがありますが、グラフデータベースを利用すると短時間で処理が完了します。

グラフデータベースは近年、注目されており、人気の高まっているデータベースです。

●Amazon Athena

Amazon Athenaは、サーバーレスのクエリサービスです。設定、ソフトウェアの更新をはじめとするインフラストラクチャの管理は自動で行われます。

Athenaでは、S3をデータストアとして使用しており、標準SQLを使用してS3内のデータを分析することができます。

SQLを使用してクエリを発行するだけで実行できるので、操作は非常に簡単です。

【S3に格納されたファイルをAthenaで検索する場合の構成イメージ】

●Amazon QuickSight

Amazon QuickSightは、フルマネージド型のBIサービスです。Amazon RDSやAmazon Redshiftといったデータベースサービスをはじめ、ストレージサービスのAmazon S3やクエリサービスのAmazon Athenaなどと連携して、データをダッシュボードで可視化することができます。AWS以外の外部データベースや、ファイルデータソースなどのデータ可視化も可能です。

●AWS Database Migration Service（AWS DMS）

AWS Database Migration Serviceは、データベース間でデータを移行するためのサービスです。リレーショナルデータベース、NoSQLデータベース、データウェアハウスなど複数種類のデータベースに対応しており、オンプレミス環境のデータベースサーバーからAmazon RDSへの移行やRDSからRDSへの移行などに使用できます。

また、AWS Schema Conversion Tool（AWS SCT）と組み合わせて使用することで、異なるデータベース間の移行やインデックス、ビューの作成も可能です。

DMSは複数種類のデータベースの移行に対応していますが、移行元に指定可能なデータベースと移行先に指定可能なデータベースの種類は異なっており、移行先に指定するデータベースのほうが移行元よりも多くの種類がサポートされています。どのようなデータベースが移行先と移行元でサポートされているのか確認しておきましょう。

AWS DMSがサポートするデータベース（移行元と移行先）は、以下を参照してください。
https://docs.aws.amazon.com/dms/latest/userguide/CHAP_Source.html
https://docs.aws.amazon.com/dms/latest/userguide/CHAP_Target.html

●AWS DataSync

AWS DataSyncは、NFSやSMBなどで構築されたストレージシステムとAWSのストレージサービス間のデータ転送サービスです。

オンプレミス環境のファイルサーバーからAmazon EFSにデータを移行したり、Amazon S3上のファイルシステムに転送するなど、サポートされているサービス間でデータを柔軟にやり取りできます。

データ転送の相手先がAWSサービスであれば、IAM権限やセキュリティグループでの許可が必要で、AWSサービス以外であればエージェントインストールが必要になります。

データ転送の具体的な手順は、次のとおりです。

① DataSyncエージェントデプロイ
② 転送ソース設定
③ 転送ターゲット設定
④ データ転送タスク設定・作成
⑤ データ転送タスク開始

データ転送タスクでは、転送する際のオプションルールも設定することができます。ストレージシステムで管理されているファイルのグループIDや、ユーザーIDなどのメタデータをまとめてコピーするオプションがあることを理解しておきましょう。

AWS DataSyncによるデータ転送の詳細は、次のWebページを参照してください。
https://docs.aws.amazon.com/datasync/latest/userguide/create-task.html

AWS DataSyncと似たようなサービスでAWS Snowballという転送サービスも提供されています。AWS DataSyncは主にオンラインでのデータ転送に適しており、ネットワーク的にも余裕のあるケースで利用されることが多いです。
これに対して、AWS Snowballはインターネットから隔離された環境などから、大量のデータをオフラインで転送するような場合に利用されます。
このように、似たような機能を備えたサービスでも要件が異なる場合があるため、そのサービスがどのようなユースケースで利用されるか知っておくとよいでしょう。

●Amazon DocumentDB

Amazon DocumentDBは、マネージドなドキュメントデータベースサービスです。
JSONやXMLのような半構造データを保存し、検索・利用する場合に活用できます。代表的なNoSQLデータベースであるMongoDBと互換性があります。

●Amazon Keyspaces

Amazon Keyspaces（**Apache Cassandra向け**）は、Apache Cassandra互換のマネージドデータベースサービスです。サーバーレスのため、ユーザーの管理も不要です。

Cassandraは、Amazon DynamoDBと同じようにNoSQLデータベースの1つで、キーバリュー型のデータベースです。

●Amazon Timestream

Amazon Timestreamは、サーバーレスの時系列データベースサービスです。IoTのセンサーデータなど、大量に発生するデータを時系列かつ効率的に保存し、分析に利用することができます。

Timestreamを時系列データの分析用途で利用する場合は、従来のリレーショナルデータベースを利用するよりもコストを抑えられます。

●Amazon Quantum Ledger Database（QLDB）

Amazon Quantum Ledger Database（QLDB）は、フルマネージド型の台帳データベースサービスです。

金融取引など、データの整合性や過去履歴の真正性が求められる用途では、ブロックチェーンの仕組みを利用することがあります。この環境を自前で準備する場合は、システム開発や運用に相応のコストが必要になりますが、Amazon QLDBを利用すると低コストで容易に実現できます。

Q 演習問題

1 **企業では、Amazon EC2とAmazon RDSで稼働するシンプルなアプリケーションが稼働しています。ところが、最近になって稼働中のRDSが暗号化されていないことがわかりました。担当者は早急に対策を取りたいと考えています。実施するべき手順を次のなかから1つ選択してください。**

A. 稼働中のRDSの暗号化設定を有効化する

B. 稼働中のRDSのスナップショットを取得し、そのスナップショットから新しいRDSを起動する。起動時の設定で暗号化を有効化する

C. AWSサポートに問い合わせて、暗号化設定を依頼する

D. 稼働中のRDSのスナップショットを取得し、暗号化設定を有効化してスナップショットをコピーする。コピーしたスナップショットから新しいRDSを起動する

2 **ある企業では、数年分のデータをAmazon Redshiftに保存しており、そのデータをBIツールで分析しています。あるとき、分析チームの担当者は、企業データの一部をグループ会社の分析部門に在籍するユーザーと連携することを求められました。グループ会社でもAWSは利用されていますが、担当者は連携対象以外のデータは見せないようにしたいと考えています。また、管理対象となるユーザーの数も、極力増やしたくないと考えています。最小限の作業で迅速にこれらの要件を満たす手順は、次のうちどれですか。**

A. Redshiftのスナップショットを取得し、グループ会社のAWSアカウントと共有する

B. グループ会社ユーザーのアクセス用にRedshiftのユーザーを払い出して、必要な権限を付与する

C. クロスアカウントデータ共有を使用して、グループ会社のAWSアカウントと共有する

D. Redshiftから必要なデータだけファイルエクスポートし、そのファイルをグループ会社のユーザーと共有する

3 **Amazon RDSの特徴として正しくない項目は、次のうちどれですか。**

A. SSL接続が可能

B. リードレプリカを利用可能

C. マルチAZを利用可能

D. データキャッシュが可能

A 解答

1 D

RDSの暗号化の知識を問う問題です。稼働中の非暗号化RDSを暗号化するには、一度スナップショットを取得し、スナップショットコピーする際に暗号化を有効化する必要があります。AとCはスナップショットを使用していないため正解ではありません。したがって、**D**が正解です。

2 C

Redshiftデータの共有方法の知識を問う問題です。問題文に管理対象のユーザーを増やしたくないとあるため、Bのユーザーを払い出すという選択肢は除外されます。A、C、Dはいずれも実現可能ですが、数年分のデータを最小限の作業かつ迅速に連携するには、クロスアカウントデータ共有が適切です。したがって、**C**が正解です。

3 D

選択肢A、B、CはすべてAmazon RDSの特徴を正しく説明しています。データをキャッシュできるサービスはAmazon ElastiCacheです。

1-7 データ処理・分析サービス

AWSでは、アプリケーションがシームレスにデータ処理できるようにするために、通知やデータ連携、分析に関わるさまざまなサービスが提供されています。
本節では、データ処理・分析に関連するサービスの概要を説明します。

1 データ処理・分析サービスの概要

AWSでは、データ処理・連携を支援する各種サービスを提供しています。これらのサービスを組み合わせることにより、アプリケーションで自ら実装することなく、あらかじめ用意されているAWS上のサービスだけでデータの連携や処理、分析を簡単に行うことができます。

たとえば、Eメール送信を実現する「Amazon Simple Email Service(SES)」や、さまざまなプロトコルでサービスにデータを配信する「Amazon Simple Notification Service(SNS)」、ストリーミング処理を行う「Amazon Kinesis」など、連携するサービスやデータの特徴によってさまざまなサービスが提供されており、ユースケースに応じた利用方法を理解することが重要です。

2 メッセージキューイング処理サービス

メッセージをキューイングすることで、バッチのような非同期かつ並列の分散処理が可能になります。そうしたメッセージのキューイングは、AWSでは、**Amazon Simple Queue Service(SQS)**というマネージドサービスで実現できます(詳細は「5-3 需要と供給のマッチングによるコスト最適化」を参照)。たとえば、処理負荷の大きいビデオのエンコーディングなどのバッチ処理を、複数のEC2インスタンスに振り分けて分散処理する場合などに、SQSによるキューイング処理を利用できます。

このほかにも、オープンソースのApache ActiveMQやRabbitMQ向けのマ

ネージド型サービスであるAmazon MQが提供されています。

3 通知処理サービス

アプリケーションからメッセージやデータを通知したい場合に利用する代表的な通知処理サービスには、次の2つがあります。

●Amazon Simple Email Service（SES）

Amazon SESは、AWSで提供されるEメールサービスです。たとえば、WebアプリケーションからSESで提供されるAPIを実行することで、メールを利用した通知を簡単に行うことができます。

●Amazon Simple Notification Service（SNS）

Amazon SNSは、AWSで提供されるメッセージ通知サービスです。任意のメッセージを、HTTPSなどのさまざまなプロトコルでアプリケーションから簡単に送信できます。また、Amazon CloudWatchやAmazon SQS、AWS Lambdaをはじめとした AWSの各種サービスへの通知・連携も可能です。

Amazon SNSは、次の要素で構成されます。

● トピック

メッセージを送信し、通知を受信するためのアクセスポイントのことです。トピックを作成し、購読者がこのトピックを購読（サブスクライブ）することで、通知が送受信できるようになります。

● 購読者（サブスクライバ）

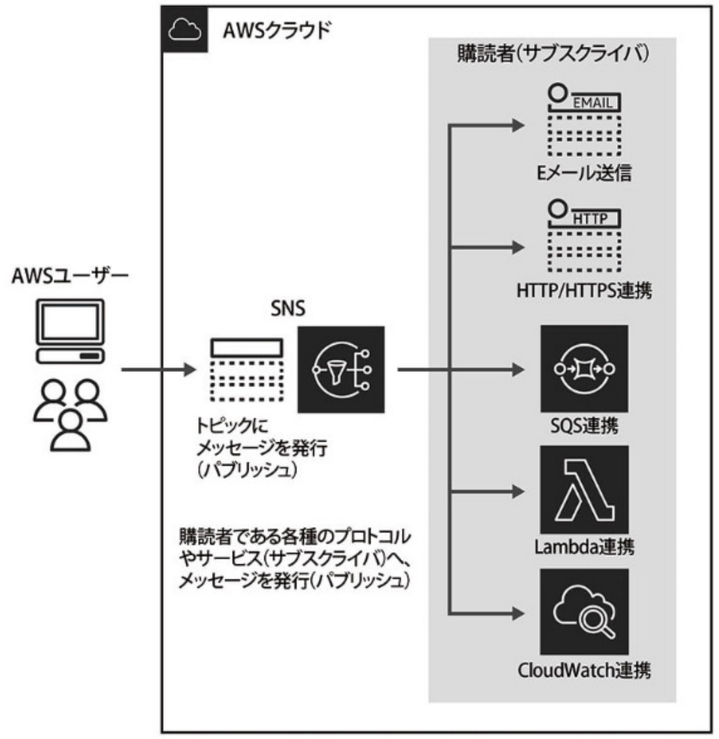
対象となるトピックから発信されるメッセージの購読者（サブスクライバ）を設定します。サブスクライバには、次のようなプロトコルやAWSサービスを選択できます。

- HTTP/HTTPS
- Email
- Amazon SQS
- AWS Lambda

● メッセージの発行（パブリッシュ）

作成したトピックに対して、アプリケーションなどからメッセージを発行（パブリッシュ）します。トピックに発行されたメッセージは、SNSサービスを通じて、登録されている購読者（サブスクライバ）に配信されます。たとえば、購読者がLambdaの場合、トピックに発行されたメッセージがSNSを通じてLambdaに配信され、そのメッセージがLambda側で処理されるといった連携も実行できます。

【SNSによるメッセージ通知・連携の構成例】



参考

Amazon SNSのように、メッセージを配信（パブリッシュ）／購読（サブスクライブ）する形式による通知の仕組みは、「Pub／Sub」と呼ばれています。

4 パイプライン処理サービス

データベースからのデータの抽出・変換・保存(ETL：Extract Transform Load)など、データの順次処理(パイプライン処理)が求められるケースがあります。AWSでは、**AWS Data Pipeline**サービスを利用して、ETLなどのデータの順次処理を実行することができます。

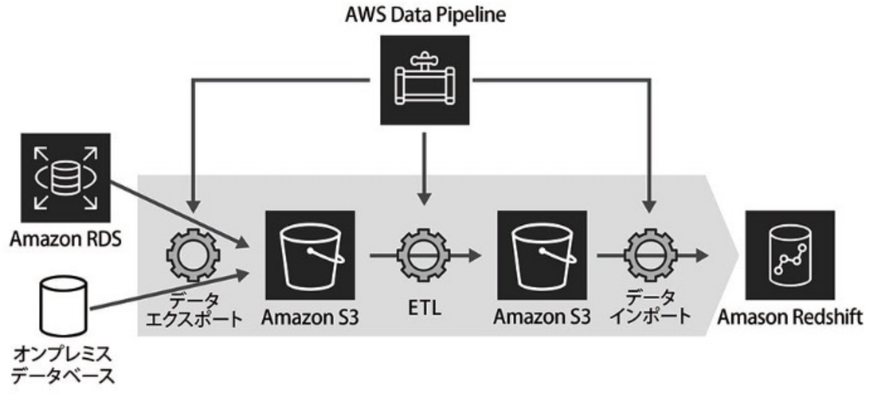
Data Pipelineの主な特徴は、次のとおりです。

- データのETL処理をスケジュール機能で自動化
- データの順次処理をワークフロー形式で定義
- 処理の成功・失敗などのイベントの通知が可能
- オンプレミス環境との連携が可能

たとえば、以下のようなユースケースでData Pipelineを利用します。

- 1日1回定期的にAmazon DynamoDBなどのデータベースからデータを取り出し、Amazon S3にバックアップしたあとで、取り出し元のテーブルをDynamoDBから削除する
- 複数拠点に分散したWebサーバーからアクセスログをAmazon S3に定期的に保存・加工し、S3からAmazon Redshiftへデータをロードする

【Data Pipelineによる順次処理の例】



また、**AWS Glue**というサービスも提供されています。

AWS GlueはData Pipelineと同様に、データの検出、加工や変換、結合などのETL処理を、ワークフローなどでパイプライン制御できるデータ統合サービスです。

たとえば、さまざまなソースからデータを検出・抽出したうえで、データのクリーニングや正規化、結合などのETL処理を行い、結果のデータをデータベースにロードする、といった一連の処理を、GUIやジョブ・ワークフロー実行などで簡易に行うことができます。

さらに、このようなデータの順序処理には**AWS Step Functions**も利用可能です。Step Functionsは、データのETL用途に限らず、汎用的に依存関係のある複数タスクやサービス実行を制御可能なサービスです。

> 参考
> AWS Data PipelineやAWS Glue、AWS Step Function以外にも、データのパイプライン処理はさまざまなサービスで実現が可能です。実際に処理を実行する際には詳細な要件や運用性などの観点からサービスを使い分ける必要がありますが、試験に向けては、細かな機能差異よりもデータ処理サービスとしての全体概要を押さえましょう。なお、Data Pipelineは2023年2月28日にサービス廃止の予定とアナウンスされています。当面、試験にはキーワードとして登場する可能性もあるため、サービス名称は押さえておきましょう。

5 ストリーミング処理サービス

リアルタイムに流れるデータの処理(ストリーミング処理)には、**Amazon Kinesis**を利用します。たとえば、IoTなどのデバイスからデータをリアルタイムに受信して分析する場合に利用します。

Kinesisは、主に次の3つのサービスから構成されます。

- Kinesis Data Streams：ストリーミングデータの収集
- Kinesis Data Firehose：ストリーミングデータの保存
- Kinesis Data Analytics：ストリーミングデータの分析

各サービスの詳細については、「5-3　需要と供給のマッチングによるコスト最適化」で説明します。

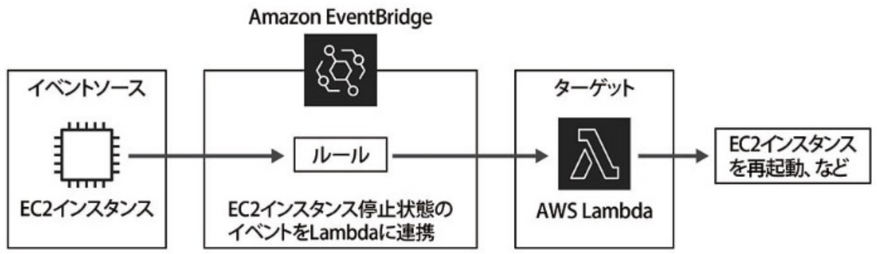
参考

Kinesisにはもう1つ、動画のストリーミング処理に使われる「Kinesis Video Streams」というサービスもあります。また、オープンソースのApache Kafkaをベースとしたストリーミングデータの処理・分析サービスとして、Amazon Managed Streaming for Apache Kafka（MSK）も提供されています。

6　イベント連携処理サービス

Amazon EventBridgeは、AWSの各種サービスや外部のSaaS(Software as a Service)などのイベントソースで発生するイベントを、あらかじめ設定したルールに基づいて、他サービスの各種ターゲットにリアリタイムで連携できるハブのようなサービスです。

【EventBridgeの連携イメージ】



7　その他のデータ連携・分析処理サービス

そのほかにも、AWSにはデータを連携し、分析処理を行える有用なサービスが用意されています。

●Amazon AppFlow

Amazon AppFlowは、SalesforceやSAP、SlackなどのSaaS製品と、Amazon S3やAmazon RedshiftなどのAWSサービスとの間で、データを安全に転送できるデータ連携サービスです。

●AWS Lake Formation

AWS Lake Formationは、データ処理・分析に必要なデータの蓄積場所であるデータレイクを簡易に構築・管理できるサービスです。

●Amazon EMR

Amazon EMRは、Apache SparkやApache Hive、Prestoなどのオープンソースを使用した、ペタバイト・クラスのデータを処理、分析できるスケーラブルな処理サービスです。

●Amazon OpenSearch Service（Amazon Elasticsearch Service）

Amazon OpenSearch Serviceは、オープンソースであるElasticsearchをベースとしたインタラクティブなログ分析・検索サービスです。キーワードをもとにしたウェブサイト検索などを簡単に実行できます。

●Amazon QuickSight

Amazon QuickSightは、分析されたデータを可視化する統合的なダッシュボードサービスです。

●AWS Data Exchange

AWS Data Exchangeは、世界中に存在する各種のサードパーティーデータ(例：ロイターなど)を、AWSがサブスクリプション方式で提供するサービスです。

●AWSにおける多種多様なAI関連サービス

これらのほかにもAWSでは、AIなどのさまざまなアルゴリズムに基づくデータ分析サービスが提供されています。「AIの民主化」と呼ばれるように、これらのサービスは、AWSが蓄積してきた高品質なアルゴリズムを一般的なユーザーでも簡易に利用できるような形で提供されています。

【AI関連サービス】

サービス名	概　要
Amazon Comprehend	機械学習により、テキストからインサイトや関係性を発見するための自然言語処理（NLP）サービス
Amazon Forecast	機械学習（ML）をベースにした時系列予測サービス
Amazon Fraud Detector	機械学習により、潜在的なオンライン不正行為の特定・発見や不正検出モデルを構築できるサービス
Amazon Kendra	機械学習（ML）を利用したインテリジェント検索サービス
Amazon Lex	チャットボットなどの会話型インターフェイスを備えたフルマネージド型人工知能（AI）
Amazon Polly	文章をリアルな音声に変換するサービス
Amazon Rekognition	画像と動画から情報とインサイトを抽出
Amazon SageMaker	機械学習（ML）モデルの統合的な開発プラットフォーム
Amazon Textract	スキャンしたドキュメントからテキストや手書き文字、データを自動的に抽出する機械学習（ML）サービス
Amazon Transcribe	音声をテキストに自動的に変換
Amazon Translate	カスタマイズ可能な言語翻訳

試験対策

これらのデータ連携・分析処理サービスの内容をすべて覚える必要はありませんが、どのような種類のAIサービスが提供されているか、概要は押さえておきましょう。

Q 演習問題

1 処理負荷の高いビデオのエンコーディングを、バッチ処理として複数のEC2インスタンスに振り分けたいと考えています。適切なAWSサービスは、次のうちどれですか。

A　Amazon SQS

B　Amazon SNS

C　Amazon SES

D　AWS STS

2 AWS Lambdaにメッセージを通知・連携したいときに利用するサービスとして、適切なサービスは次のうちどれですか。

A.　SMS

B.　Amazon SNS

C.　Amazon SES

D.　AWS STS

A 解答

1 A

メッセージのキューイングは、Amazon SQSというマネージドサービスで実現できます。この場合、ビデオのエンコーディングの依頼をメッセージとしてSQSに送信し、複数のEC2インスタンスがSQSからメッセージを取得して、並行してエンコーディングのバッチ処理を開始します。
Amazon SESはマネージド型のEメールサービス、AWS STSはAWSリソースへのアクセスを制御できる一時的なセキュリティ認証サービスです。

2 B

Amazon SNSはAWSで提供されるメッセージ通知サービスで、AWS Lambdaをはじめとした各種AWSサービスへの通知が可能です。

1-8 構成管理サービス・開発サービス

AWSには、システムの開発・リリースを短時間かつ高品質に行うための構成管理サービスや、開発者・運用者の作業を支援するサービスが用意されています。
本節では、AWSで提供されている構成管理サービスおよび開発者・運用者向けのサービスについて説明します。

1 構成管理に関する重要用語

AWSでは、ユーザーの環境構築に関わる作業を自動化するためのサービスが提供されています。このようなサービスを利用し、従来は手動で行っていた作業を自動化することで、作業時間の短縮や人為的な作業ミスの軽減が期待できます。ここでは、その基本となるサービスについて説明していきます。
まず、構成管理の説明でよく登場する2つの用語について説明します。

●プロビジョニング

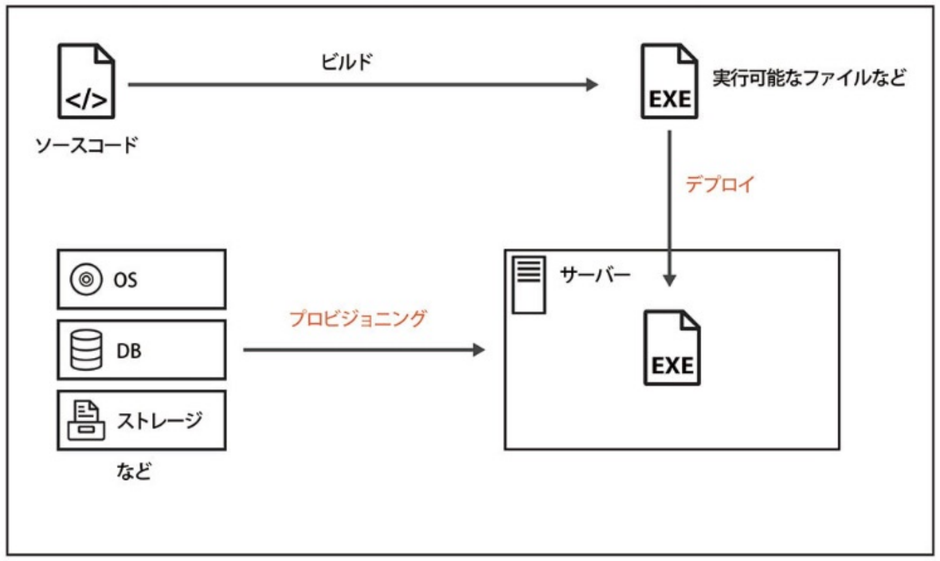
プロビジョニングは、「provision(供給)」という単語から派生した用語です。一般的には、システム利用の需要を予測し、必要に応じて設備やサービスなどのリソースを提供できるように準備することをプロビジョニングといいます。
クラウド環境における**プロビジョニング**とは、多くの場合は複数のITリソース(サーバー、ネットワーク、データベース、ストレージなど)を、システムの利用状況や障害発生の状況に応じて動的に利用したり割り当てることを指します。

●デプロイ

一般的には、アプリケーションを実行するためのファイル(バイナリファイルやソースコード)や、アセット(バイナリファイルやソースコードだけでなく、その他ドキュメント類をまとめたもの)を配布することをデプロイといいます。
クラウド環境における**デプロイ**とは、多くの場合はサーバーにファイルやアセットを配置し、それが利用できる状況にすることを指します。

これ以外に、ビルドという用語もよく使用されます。ソースコードから実行可能なファイルを生成することをビルドといいます。これらの機能イメージを示すと次の図のようになります。

【プロビジョニング・デプロイ・ビルドのイメージ】



2 コードによるインフラストラクチャ構成管理

ハードウェアを自社で調達してインフラストラクチャを構築する場合は、一般的にインフラストラクチャ担当のエンジニアが環境構築手順書を作成し、環境をセットアップします。その際の課題として、次に示すようなことがあげられます。

- 作業中にミスが発生する可能性がある
- 導入後に構成変更があることを考慮して、インフラストラクチャ管理(設計書やサーバーのパラメータなど)が必要

アプリケーション開発においても、以前はプログラムの変更やリリース作業で同様の問題が起きていました。しかし現在では、サーバーへのデプロイ作業が自動化できるようになるなど、作業の効率化に有効な手段が一般化しています。

このように、アプリケーション開発で実践されているコード管理・作業自動化という手法をインフラストラクチャ構築にも導入し、インフラストラクチャ構築の作業内容をコードで記述して管理できるようにしたInfrastructure as Code(IaC)という考え方が登場しています。

AWSでは、定義ファイルをコーディングし、コードを実行することでリソースのプロビジョニングが可能です。インフラストラクチャ環境の運用をコードで管理しているため、まったく同じ作業であれば以前に実行したコードをそのまま流用できます。構築作業に変更が生じた場合は、該当箇所のみを修正してコードを実行することで、修正版のインフラストラクチャ環境を構築することができます。

このように、従来では煩雑であったインフラストラクチャ構築手順のバージョン管理も容易になりました。

> 参考
> バージョン管理とは、主にアプリケーション開発やドキュメント作成時などの履歴を管理することです。たとえば、最初に開発したアプリケーションのコードやドキュメントをバージョン1.0として保存し、次に機能追加したアプリケーションとドキュメントをバージョン1.1として保存するといった作業になります。

次項では、AWSの代表的な構成管理サービスと開発者・運用者向けサービスについて説明します。

3 構成管理サービスと開発者・運用者向けサービス

●AWS CloudFormation

AWS CloudFormationは、AWS内のすべてのインフラストラクチャリソースを自動でプロビジョニングできるサービスです。

以下に、CloudFormationを理解するための2つの重要な用語について説明します。

● テンプレート

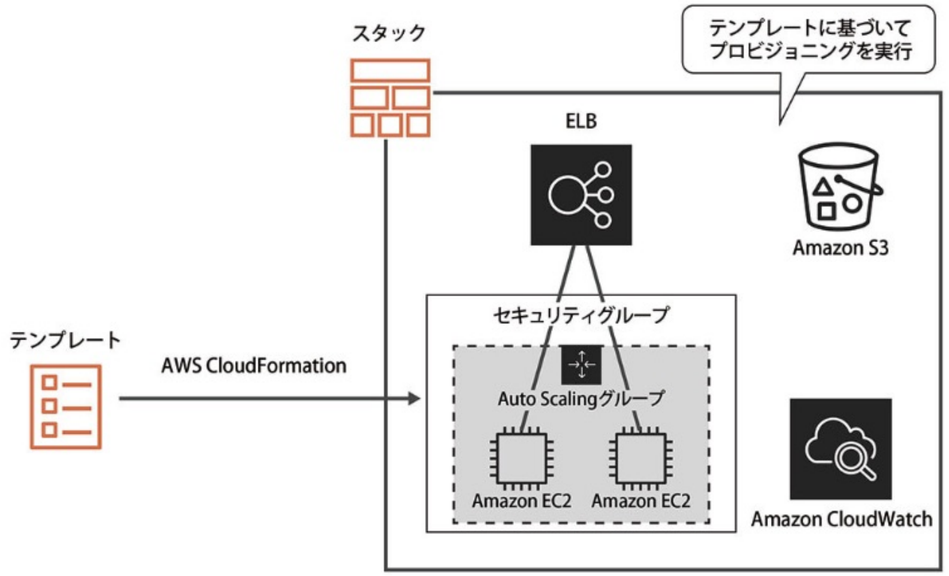
CloudFormationの設定ファイルを指し、このファイルにプロビジョニングしたいAWSリソースを記述します。

● **スタック**

テンプレートに従ってCloudFormationでプロビジョニングされるAWSリソースの集合を指します。

つまり、テンプレートに記載された内容に基づいて、プロビジョニングしてスタックを作成するのがCloudFormationの機能です。

【CloudFormationによるプロビジョニングの構成例】



参考

AWS CloudFormationの設定について、詳細な書式はAWSのWebページで確認できます。たとえば、テンプレートの基礎に関しては、以下に記載されています。
https://docs.aws.amazon.com/ja_jp/AWSCloudFormation/latest/UserGuide/gettingstarted.templatebasics.html

●AWS Control Tower

AWS Control Towerは、AWSのマルチアカウントを管理するサービスです。クラウド環境を利用している企業では、アカウントを複数保有していることが多くあります。

たとえば、次に示す単位でAWS環境を複数保有しているとします。

・システムごと
・開発・テスト・本番環境ごと
・部署やチームごと

AWSアカウントの数が少なければ、これらの環境を個別に管理することもできますが、数が増えるとセキュリティ管理やアカウント管理、リソース管理などの観点から考慮すべきことが多く、複数のアカウントを安全に運用することが困難になります。Control Towerは、そうしたマルチアカウント管理の課題に対応できるサービスです。

●AWS License Manager

AWS License Managerは、各ベンダーの提供するライセンスが管理できるサービスです。

たとえば、データベースとしてOracle Databaseを、ERPツールとしてSAPを導入するなど、現在ではクラウドとオンプレミスの混在環境でさまざまなベンダーの製品・サービスを利用することが一般的です。

License Managerを利用することで、AWSとオンプレミス環境でライセンスを簡単に管理でき、契約で規定されている以上のライセンスを使用するといったライセンス違反を防ぐこともできます。

●AWS Service Catalog

AWS Service Catalogは、AWSでの使用が承認されたITサービスのカタログを作成・管理できるサービスです。

企業や組織でクラウド環境を利用する場合、ユーザーが利用したいサービスや好みのインスタンスタイプを自由に選択して、環境を構築することはありません。一般的には、管理者がセキュリティやコストを検討したうえで構築されるケースがほとんどです。また、ユーザーからの要望にすべて対応していては、管理者の負荷が増大し、作業のスピード感も低下してしまいます。

Service Catalogを利用することで、管理者は事前に『利用者が利用・構築できる環境』をサービスカタログとして定義することができます。ユーザーはこのサービスカタログを使用して、管理者が許可したAWSリソースをユーザー自身で作成できます。

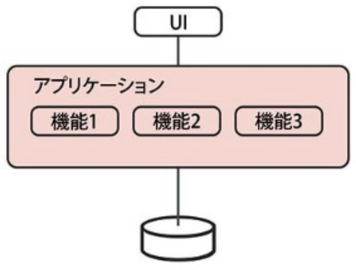
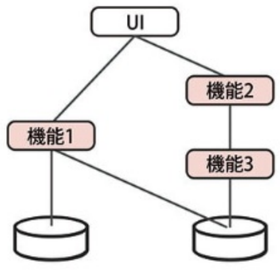
●AWS X-Ray

AWS X-Rayは、本番環境や分散アプリケーションを分析およびデバッグできるサービスです。

このサービスを解説するにあたり、まずアプリケーション開発モデルについて説明します。よく比較されるアプリケーション開発モデルとして、従来のアプリケーション開発で使用されていた**モノリシックアーキテクチャ**と、近年使用されることが多くなった**マイクロサービスアーキテクチャ**があります。

次の表は、この2つのアーキテクチャを比較したものです。

【アプリケーション開発モデルの比較】

	モノリシックアーキテクチャ	マイクロサービスアーキテクチャ
特　徴	複数の機能を1つのアプリケーションとして動作させる	個別の機能を独立したサービスとして動作させる
メリット	●規模が小さければ開発が容易 ●全体の動作を確認するためのテストが容易 ●デバッグが容易	●修正が必要な場合、機能単位での変更が容易 ●サーバーレスやコンテナとの相性がよい
デメリット	●大規模なアプリケーションは開発工数がかかる ●少し変更したい場合でも、すべてをデプロイし直す必要がある ●機能単位でのスケーリングが不可能	●機能が増えるほど、デバッグが複雑になる ●ボトルネックがわかりにくくため、パフォーマンス分析が難しい
イメージ		

マイクロサービスアーキテクチャは、モノリシックアーキテクチャのデメリットを補うために利用されることが多くなりましたが、一方で、デバッグしにくいという特徴もあります。X-Rayは、そのアーキテクチャのデメリットを補いながら、分散アプリケーションのデバッグやパフォーマンス分析に利用できるサービスです。

●AWS Proton

AWS Protonは、コンテナとサーバーレス環境のためのアプリケーションデプロイサービスです。

マイクロサービスアーキテクチャが普及してきたことで、システムを運用する際には、従来よりも数多くのコンポーネントを扱う必要が出てきました。また、それらを従来の方法で管理することも難しくなりました。Protonを利用すると、多数のコンポーネントを効率よくデプロイできるようになります。

参考

新試験（SAA-C03）では対象外ですが、このほかにも構成管理サービスとして次のサービスが提供されており、ここに参考情報として紹介します。

- **AWS OpsWorks**

  構築手順どおりにサーバー構築作業を自動化する構成管理サービスで、デプロイ、プロビジョニング、リリース後の監視の機能を備えています。ChefやPuppetといったサーバーの構成を自動化するツールを利用できます。

- **AWS CodeDeploy**

  アプリケーションのデプロイを自動化するサービスです。開発したアプリケーションのリリースを迅速化でき、また、リリース時の作業を自動化できるため、作業ミスの削減にもつながります。

- **AWS CodePipeline**

  アプリケーションのビルド、テスト、デプロイまでの処理手順を定義し、実行します。

- **AWS CodeBuild**

  ビルド・テストに利用します。

- **AWS CodeCommit**

  ソースコードやドキュメントなどの成果物を、AWS上の「Gitリポジトリ」で管理するサービスです。

ソフトウェア開発・運用時の作業と関係するAWSサービスは、次に示すとおりです。

「運用時のモニタ」に属する2つのサービスのうち、X-Rayはアプリケーションのパフォーマンス分析やデバッグに適しているのに対して、CloudWatchはインフラのパフォーマンス分析やアプリ・AWSサービスなどの監視に適しています。

Q 演習問題

1 「テンプレート」と呼ばれる定義ファイルに従って、AWSリソースをプロビジョニングするサービスはどれですか。

A AWS CloudFormation

B AWS License Manager

C AWS Service Catalog

D AWS Control Tower

2 AWS X-Rayサービスの特徴は、次のうちどれですか。

A 複数の機能を1つのアプリケーションとして動作させている

B サービスカタログを利用し、管理者の許可したAWSリソースを利用者自身が作成できる

C マイクロサービス アーキテクチャのデバッグに適している

D コンテナとサーバーレスのためのアプリケーションデプロイが可能

A 解答

1 A

テンプレートを使用してAWSリソースをプロビジョニングするサービスは、AWS CloudFormationです。
AWS License Managerは、ベンダー各社が提供するライセンスを管理できるサービスです。
AWS Service Catalogは、AWSでの使用が承認されたITサービスのカタログを作成および管理できるサービスです。
AWS Control Towerは、AWSのマルチアカウントを管理するサービスです。
したがって、**A**が正解です。

2 C

AWS X-Rayサービスは、本番環境や分散アプリケーションを分析およびデバッグできるサービスで、マイクロサービス アーキテクチャのデバッグに適しています。
複数の機能を1つのアプリケーションとして動作させているのは、モノリシック アーキテクチャの特徴です。
サービスカタログを利用し、管理者の許可したAWSリソースを利用者自身が作成できるサービスは、AWS Service Catalogです。
コンテナとサーバーレスのためのアプリケーションデプロイが可能なサービスは、AWS Protonです。
したがって、**C**が正解です。

1-9 運用管理サービス

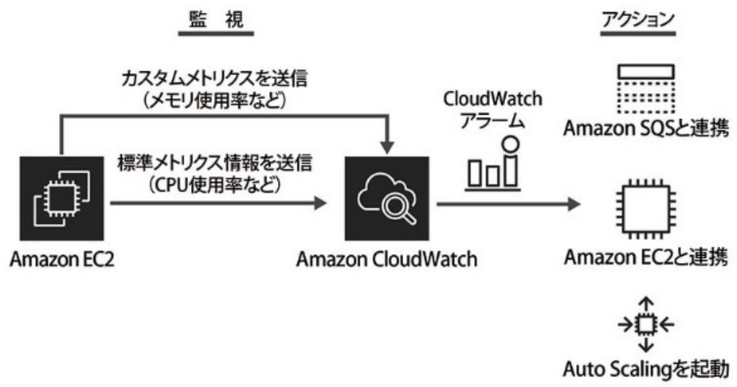
ここまでの節では、コンピューティングやストレージ、データベース、データ処理・分析サービスなど、AWS上でシステムを構築する際の重要なサービス群を紹介してきました。
本節では、AWSの多様なサービスを組み合わせて構築されたシステムやアーキテクチャを、安定的に運用する際に中核となる運用管理サービスを紹介します。

1 Amazon CloudWatchの概要

Amazon CloudWatchは、AWS上で稼働するさまざまなシステムやAWSのリソースの情報を収集・監視・可視化するサービスです。

たとえば、Amazon EC2インスタンスのCPU使用率を収集・監視し、ある一定以上の使用率になった場合に通知(アラートを送信)するといったアクションのほか、CPU使用率を時系列でグラフ化して監視できます。EC2だけでなく、Amazon RDSやAmazon S3、Amazon DynamoDBのような代表的なAWSサービスの使用状況、AWS利用料なども、CloudWatchで収集・監視できます。

【CloudWatchの動作イメージ】



●EC2のメトリクス

CloudWatchで取得・監視する項目のことを「**メトリクス**」と呼びます。EC2では、デフォルトでAWSから提供されている**標準メトリクス**と**カスタムメトリクス**の2種類があります。

● 標準メトリクス

AWSから提供されている標準の監視項目です。デフォルトでEC2からCloudWatchに対して、さまざまな監視データを送信できます。

たとえばEC2では、次のようなメトリクスがあらかじめ用意されています。

【標準メトリクス】

メトリクス	説　明
CPUUtilization	CPU使用率
DiskReadOps	インスタンスストアボリュームからの読み取り回数(指定された期間)
DiskReadBytes	インスタンスストアボリュームから読み取られたバイト数
NetworkIn	すべてのネットワークインターフェイスから受信されたバイト数

● カスタムメトリクス

標準メトリクス以外の独自のメトリクスを定義して収集したい場合は、**カスタムメトリクス**を使用します。

たとえば、メモリ使用率やディスク使用率、スワップ使用率など、標準メトリクスでは提供されていない監視項目は、ユーザーがEC2インスタンス上で監視スクリプトを作成・設定し、CloudWatchに送信して収集することができます。また、CloudWatchエージェントを通じて、カスタムメトリクスをCloudWatchに送信することができます。

EC2のメモリ使用率やディスク使用率、スワップ使用率を監視するには、カスタムメトリクスの設定が必要です。

●CloudWatchの監視間隔と保存期間

CloudWatchでEC2を監視する場合、基本モニタリングと詳細モニタリングの2つのプランが提供されています。

【CloudWatchのプラン】

プラン	課　金	監視間隔	収集データの保存期間
基本モニタリング	無料	5分間隔	最大15カ月
詳細モニタリング	追加料金が必要	1分間隔	最大15カ月

試験対策

基本モニタリングと詳細モニタリングでは、収集したデータの監視間隔が異なります。

監視間隔はAWSサービスによって異なります。たとえば、Amazon EBSやELB、Amazon RDSでは、1分間隔のデータが無料で閲覧できます。

●CloudWatchアラームによるアラームとアクション設定

CloudWatchアラームでは、CloudWatchで監視している項目が、ある一定の値に達した場合にアラームとアクションを起動することができます。たとえば、「EC2インスタンスのCPU使用率が80％を5分以上超えた場合は、アラームを送信してSNSでメール通知し、もう1台のEC2インスタンスをAuto Scalingポリシーで起動する」などの一連のアクションが自動化できます。

CloudWatchには、次の図に示すように3つの状態があります。

【CloudWatchの状態】

状　態	説　明
OK	定義された監視しきい値を下回っている（正常）状態
ALARM	定義された監視しきい値を上回っている（異常）状態
INSUFFICIENT_DATA	CloudWatchに送信されるデータが不足しているため、正常か異常かを判定できない状態

たとえば「CPU使用率のしきい値を、使用率80％以上で5分間持続した場合にアラームを送信」と定義した場合、80％が5分以上継続すると、OKからALARMへとCloudWatchの状態が遷移し、事前に定義したアクション（たとえば、Auto Scalingの起動）を実行します。これに対して、CPU使用率が80％以下を5分間継続した場合は、CloudWatchの状態がALARMからOKへと戻ります。

●CloudWatch Eventsによるイベント駆動型監視とアクション設定

CloudWatch Eventsは、AWS上のリソースの状態を監視し、あるイベントをトリガーにアクションを実行する機能です。たとえば、EC2インスタンスの状態変化を監視し、状態変化を検知したらLambdaやSNSを実行する、などの連携が可能になります。

前項で説明したCloudWatchはメトリクスが監視対象ですが、CloudWatch Eventsは状態変化などのイベントをトリガーにしてアクションを実行します。

状態監視の対象は**イベントソース**と呼び、「スケジュール」とAWSリソースに対する「イベント」の2種類があります。

● スケジュール

分、時、日など時間や期間を定義できます。たとえば、「日曜の朝6時」や「2日間隔」などが指定できます。

● AWS リソースのイベント

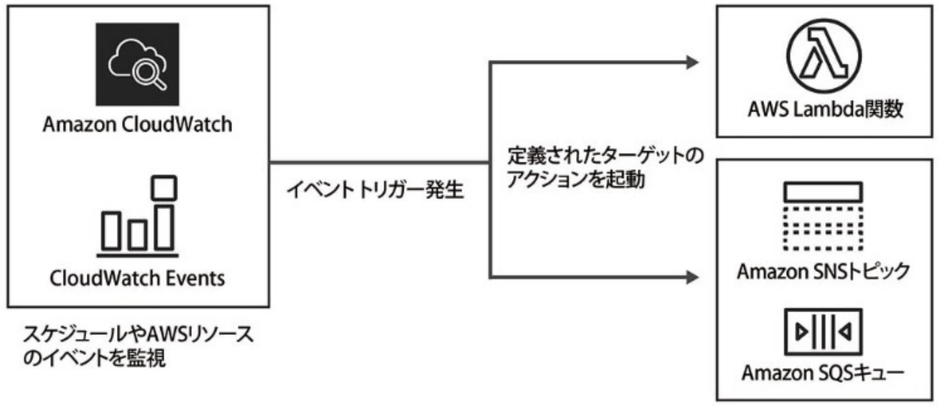
EC2インスタンスの状態変化（pending、running、terminatedなどの状態）や、Auto Scalingの状態変化（起動成功、起動失敗などの状態監視）など、AWSリソースのイベントをトリガーにできます。たとえば、Auto Scalingで「インスタンスが増加したら」などが指定できます。

監視後のアクションとして定義する対象を、**ターゲット**と呼びます。ターゲットには、次に代表されるAWSリソースを1つ、または複数指定できます。

- Lambda関数
- SNSトピック
- Kinesisストリーム
- SQSキュー

これらのイベントソースとターゲットを組み合わせると、たとえば、スポットインスタンスが強制停止する状態変化を監視し、検知したらLambdaを実行して新たなスポットインスタンスを起動する、などの自動化が可能になります。

【CloudWatch Eventsの動作イメージ】



なお、類似サービスとして、CloudWatch Eventsが機能強化されたEventBridgeというサービスがあります。EventBridgeは、CloudWatch Eventsと同じAPIが利用でき、将来的にはEventBridgeに統合される予定です。

> 参考
> 今後、Amazon CloudWatch EventsがAmazon EventBridgeへ置き換わる予定であることがAWSからアナウンスされています。試験では、EventBridge（CloudWatch Events）と補足される場合がありますが、どちらもイベント駆動型のサービスであるという点を押さえておきましょう。

●CloudWatch Logs

AWS CloudTrailやAmazon VPCフローログなど、さまざまなAWSのログを統合的に収集するサービスです。

EC2インスタンスからログを収集する場合は、独自のエージェントをインストールすることで、EC2インスタンスの各種ログをCloudWatch Logsで収集できるようになります。

また、収集したログに対しては、アラームを設定して監視することもできます。たとえば、「ERROR」から始まる行がログにあった場合にSNSに通知する、といったアラーム設定が可能です。

【CloudWatch Logsの動作イメージ】

> 試験対策
> Amazon CloudWatch、CloudWatchアラーム、CloudWatch Events、CloudWatch Logsの各々の特徴と違いを押さえておきましょう。
>
> **CloudWatch**
> CPU使用率などの各AWSリソースの状態・メトリクスを定期的に取得する機能
>
> **CloudWatch アラーム**
> CPU使用率のしきい値などのAWSリソースを監視し、状態の変化を警報（またはアクション）する機能
>
> **CloudWatch Events**
> 独自のイベントをトリガー（イベントソース）として定義し、後続のアクション（ターゲット）の組み合わせを定義する機能
>
> **CloudWatch Logs**
> EC2インスタンス等のAWSリソースのログを収集・監視する機能

2　その他の運用管理サービス

AWSには、Amazon CloudWatch以外にも、さまざまな運用管理サービスが用意されています。

●CloudTrail

AWSアカウントで利用された操作（イベントやAPIコール）をログとして記録するサービスです。

AWSマネジメントコンソールへのログインやEC2インスタンスの作成といっ

た「管理イベント」と、Amazon S3バケット上のデータ操作やLambda関数の実行といった「データイベント」の2種類を取得できます。

これらの操作ログを蓄積し、AWSへの不審なアクセスや操作がなされていないか、意図せぬ設定変更が行われていないかなどが追跡できます。また、CloudWatch Logsと組み合わせて不正な操作ログの文字列をアラームとして設定しておくことで、さまざまなセキュリティ監視や監査に活用できます。

●VPCフローログ

VPC内のネットワークインターフェイス間で行き来する通信の内容をキャプチャする機能です。CloudWatch Logsと組み合わせて、意図しない通信が行われていないかなどの監視・監査に利用します。

●AWS Config

AWSのサービスで管理されているリソースの構成変更を追跡するサービスです。たとえば、EC2インスタンスの作成や削除などの構成変更や、S3バケットに対する設定変更などの履歴をログに取得できます。

●AWS Systems Manager

AWS内のさまざまなリソースの運用情報を統合的に可視化・制御するサービスです。Systems Managerはいくつかの機能で構成されています。たとえば、パラメータストア機能では、Amazon RDSのデータベース文字列やパスワードなどの秘密データといった、AWS内で利用するさまざまなパラメータを一元管理することができます。また、オートメーション機能や実行コマンド機能では、EC2インスタンスなどのAWSリソースに対する運用タスクの制御や自動化も行うことができます。

●AWS Trusted Advisor

AWSのベストプラクティスに基づいて、「コスト最適化」「セキュリティ」「耐障害性」「パフォーマンス」「サービスの制限」の5項目で、ユーザーのAWS利用状況をチェックし、改善すべき事項を推奨するサービスです。

●AWS Managed Grafana / AWS Managed Service for Prometheus

オープンソースのGrafanaとPrometheusを、AWSがマネージドサービスとして提供しているサービス群です。

GrafanaとPrometheusはどちらも、統合的な運用管理ツールとして、さまざまなメトリクスやログの収集、分析・可視化を行うことができます。特にPrometheusは、コンテナ化されたアプリケーションおよびインフラのモニタリングを強みとしています。

試験対策

運用管理に関連するサービス群とそれらの役割の違いを押さえておきましょう。
特に、AWS CloudTrailによる操作ログの取得と、CloudWatch Logsと連携した不正監視の組み合わせは重要です。

Q 演習問題

1 **EC2インスタンスに対してCloudWatchのカスタムメトリクスを使用する必要がある監視項目は、次のうちどれですか。**

A　メモリ使用率

B　ディスク読み取り回数

C　CPU使用率

D　ネットワーク受信バイト数

2 **VPC内のネットワーク インターフェイス間のトラフィックを監視するサービスとして適切なサービスはどれですか。**

A　AWS CloudTrail

B　VPCフローログ

C　AWS Config

D　AWS Trusted Advisor

3 **あなたの会社では、WebアプリケーションをEC2インスタンス上で実行しています。EC2インスタンスのCPU使用率を監視し、あるしきい値に基づいてオートスケーリングしたい場合、どのサービスで監視するのが適切ですか。**

A　CloudWatch Logs

B　AWS Systems Manager

C　VPCフローログ

D　CloudWatchアラーム

A 解答

1 A

標準メトリクス以外の独自のメトリクスを定義して収集したい場合は、カスタムメトリクスを使用します。メモリ使用率は標準メトリクスでは提供されていない監視項目です。

2 B

VPCフローログは、VPC内のネットワークインターフェイス間で行き来する通信の内容をキャプチャする機能です。意図しない通信が行われていないかなどの監視・監査に利用します。

3 D

CloudWatchアラームは、CloudWatchで監視している項目がある一定の値になった場合に、アラームとアクションを起動できます。たとえば、EC2インスタンスのCPU使用率が80％を5分以上超えた場合にアラームを送信し、EC2インスタンスをAuto Scalingポリシーで起動するなどの一連のアクションが自動化できます。