AWS を使用した
バックアップと復元の手法

2016 年 6 月

注意

本文書は、情報提供の目的のみのために提供されるものです。本書の発行時点における AWS の現行製品と慣行を表したものであり、それらは予告なく変更されることがあります。お客様は本文書の情報および AWS 製品の使用について独自に評価する責任を負うものとします。これらの情報は、明示または黙示を問わずいかなる保証も伴うことなく、「現状のまま」提供されるものです。本書のいかなる内容も、AWS、その関係者、サプライヤ、またはライセンサーからの保証、表明、契約的責任、条件や確約を意味するものではありません。お客様に対する AWS の責任は、AWS 契約により規定されます。本書は、AWS とお客様の間で行われるいかなる契約の一部でもなく、そのような契約の内容を変更するものでもありません。

目次

要約

このホワイトペーパーは、企業 IT 環境でデータの保護を担当するエンタープライズソリューションアーキテクト、バックアップアーキテクト、および IT 管理者を対象としています。AWS を使用して実装し、バックアップおよび復旧ソリューションを補強または置き換えることができる本稼働ワークロードおよびアーキテクチャについて説明します。これらの手法により、目標復旧時間 (RTO)、目標復旧時点 (RPO)、コンプライアンスの要件を満たす低いコスト、高いスケーラビリティ、およびさらなる耐久性が提供されます。

はじめに

エンタープライズデータの増大が加速していく中で、それを保護するタスクの難易度が高くなっていきます。また、バックアップ方法の耐久性とスケーラビリティに関する質問もよく寄せられるようになります。たとえば、バックアップとアーカイブのニーズにクラウドはどのように対応するかという質問があります。

このペーパーでは、スケーラブルで信頼性が高いデータ保護ソリューションの構築に使用できるさまざまなバックアップアーキテクチャ (クラウドネイティブアプリケーション、ハイブリッド環境、オンプレミス環境) と、関連する AWS のサービスについて説明します。

AWS をデータ保護プラットフォームとして使用する理由

アマゾン ウェブ サービス (AWS) は、安全でパフォーマンスが高く、柔軟でコスト効果に優れた、使いやすいクラウドコンピューティングプラットフォームです。AWS はビジネスの差別化に直結しない運用負荷を処理し、スケーラブルなバックアップおよび復旧ソリューションの構築に使用できるツールとリソースを提供します。

データ保護戦略の一部として AWS を使用すると、多くの利点があります。

- **耐久性:** Amazon Simple Storage Service (Amazon S3) および Amazon Glacier は、それらに保存されたオブジェクトに対して 99.999999999% の耐久性を求めた設計になっています。両方のプラットフォームとも、バックアップデータ用の信頼できる場所を提供します。
- **セキュリティ:** AWS では、伝送時および保管時のデータのアクセス制御と暗号化用に数多くのオプションを用意しています。
- **グローバルインフラストラクチャ:** AWS のサービスは世界中で利用できるため、お客様のコンプライアンス要件に合ったリージョンでデータをバックアップおよび保存できます。
- **コンプライアンス:** AWS インフラストラクチャは、Service Organization Controls (SOC)、Statement on Standards for Attestation Engagements (SSAE) 16、International Organization for Standardization (ISO) 27001、Payment Card Industry Data Security Standard (PCI DSS)、Health

Insurance Portability and Accountability Act (HIPPA)、SEC[1]、Federal Risk and Authorization Management Program (FedRAMP) などの基準に準拠しているため、既存のコンプライアンス基準に簡単にバックアップソリューションを合せることができます。

- **スケーラビリティ:** AWS では、容量について心配する必要はありません。ニーズの変化に合わせて消費を拡大または縮小でき、管理のオーバーヘッドは必要ありません。
- **TCO の削減:** AWS オペレーションのスケールによりサービスのコストが下がり、ストレージの総所有コスト (TCO) の削減に役立ちます。AWS はこれらのコストの節約を料金の引き下げという形でお客様に還元します。
- **従量制の料金体系:** 必要なときに、使用する予定の期間だけ AWS のサービスを購入します。AWS の料金には、前払い料金、解約違約金、長期契約はありません。

データ保護用の AWS ストレージサービス

Amazon S3 および Amazon Glacier はバックアップやアーカイブに適したサービスです。両方とも耐久性があり、低コストのストレージプラットフォームです。さらに両方とも容量は無制限で、バックアップデータセットが大きくなってもボリュームまたはメディアの管理は不要です。従量制の料金モデルと GB/月あたりの低コストにより、これらのサービスはデータ保護のユースケースに最適です。

[1] https://aws.amazon.com/about-aws/whats-new/2015/09/amazon-glacier-receives-third-party-compliance-assessment-for-sec-rule-17a-4f-from-cohasset-associates-inc/

Amazon S3

Amazon S3 はセキュリティが高くスケーラブルなオブジェクトストレージを提供します。

Amazon S3 を使用すると、データの大きさにかかわらず、ウェブ上のどのような場所からでもいつでも保存、取得することができます。Amazon S3 は、オブジェクトとしてデータをバケットと呼ばれるリソースに保存します。AWS Storage Gateway および他の多くのサードパーティーのバックアップソリューションでは、お客様に代わって Amazon S3 オブジェクトを管理できます。バケットに任意の数のオブジェクトを保存し、バケット内のオブジェクトを書き込み、読み込み、削除できます。1 つのオブジェクトあたりのサイズは最大 5 TB です。

Amazon S3 では、さまざまなユースケースに合うように最適化された、広範なストレージクラスをご用意しています。具体的には次のとおりです。

- 頻繁にアクセスするデータ向けの汎用ストレージである **Amazon S3 Standard**。
- 存続時間が長いが、アクセス頻度の低いデータ向けの **Amazon S3 標準低頻度アクセス**。
- 長期アーカイブを目的とした **Amazon Glacier**。

Amazon S3 では、そのライフサイクルを通じてデータを管理するように設定できるライフサイクルポリシーもご用意しています。ポリシーが設定されると、データは適切なストレージクラスに移行され、アプリケーションへの変更はありません。詳細については、「S3 ストレージクラス」を参照してください。

Amazon Glacier

Amazon Glacier はきわめて低コストのクラウドアーカイブストレージサービスで、ストレージのセキュリティと耐久性を特徴としており、データのアーカイブやオンラインバックアップに適しています。コストを低く抑えるために、Amazon Glacier は、アクセス頻度の低いデータや、取り出しに数時間かかっても問題ないデータに合わせて最適化されています。お客様は Amazon Glacier を使って、1 ギガバイトあたり月額わずか 0.007 USD で、大量または少量のデータを確実に保存できます。オンプレミスのソリューションと比較して大幅にコストを削減することができます。Amazon Glacier は、保持期間が長期または無限、または長期のデータアーカイブ用のバックアップデータのストレージに最適です。詳細については、「Amazon Glacier」を参照してください。

AWS Storage Gateway

AWS Storage Gateway は、オンプレミスのソフトウェアアプライアンスをクラウドベースのストレージと接続することで、オンプレミスの IT 環境と AWS のストレージインフラストラクチャ間でシームレスかつ非常にセキュアな統合を実現します。詳細については、「AWS Storage Gateway」を参照してください。

AWS 転送サービス

サードパーティーのゲートウェイおよびコネクタに加えて、AWS Direct Connect、AWS Snowball、AWS Storage Gateway、Amazon S3 Transfer Acceleration などの AWS オプションを使って、データをすばやく転送することができます。詳細については、「クラウドデータの移行」を参照してください。

バックアップおよび復旧ソリューションの設計

データのバックアップと復旧の総合的な戦略を開発するときは、発生する可能性のある失敗または障害の状況と、それによるビジネスへの影響を識別する必要があります。業種によっては、データのセキュリティ、プライバシー、および記録保持の規制要件も考慮する必要があります。

以下を含む、ビジネスの RTO と RPO を満たすための適切なレベルの詳細度を提供するバックアッププロセスを実装します。

- ファイルレベルの復旧
- ボリュームレベルの復旧
- アプリケーションレベルの復旧 (データベースなど)
- イメージレベルの復旧

以下のセクションでは、インフラストラクチャの編成に応じたバックアップ、復旧、およびアーカイブの手法について説明します。IT インフラストラクチャは、大まかにクラウドネイティブ、オンプレミス、およびハイブリッドに分類できます。

クラウドネイティブなインフラストラクチャ

このシナリオでは、完全に AWS 上に存在するワークロード環境について説明します。次の図が示すように、これにはウェブサーバー、アプリケーションサーバー、モニタリングサーバー、データベース、および Active Directory が含まれます。

AWS からすべてのサービスを実行している場合、多くの組み込み機能を使ってデータの保護および復旧ニーズを満たすことができます。

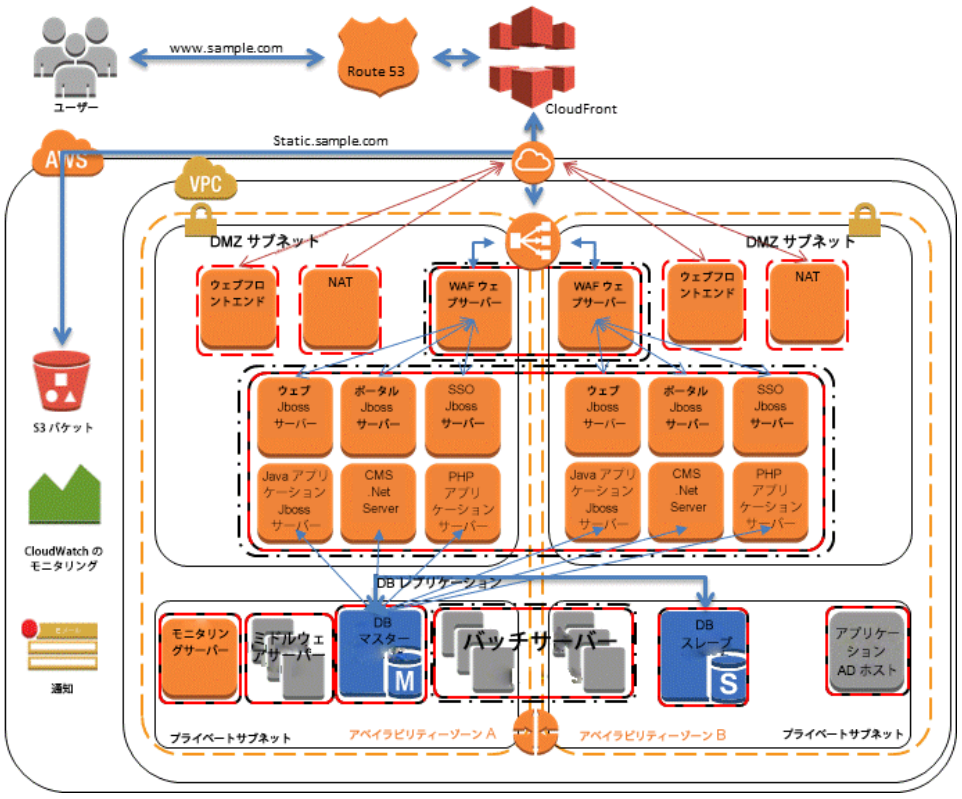

図 1: AWS ネイティブシナリオ

EBS スナップショットベースの保護

Amazon Elastic Compute Cloud[2] (Amazon EC2) でサービスを実行している場合、コンピューティングインスタンスは、Amazon Elastic Block Store (Amazon EBS) ボリュームを使ってプライマリデータを保存し、アクセスできます。このブロックストレージを、構造化データ (データベースなど) または非構造化データ (ボリュームのファイルシステムのファイルなど) に使用できます。

Amazon EBS には、任意の Amazon EBS ボリュームのスナップショット (バックアップ) を作成する機能があります。これはボリュームのコピーを作成し、それを Amazon S3 に配置します。そこで、複数のアベイラビリティーゾーンに冗長的に保存されます。最初のスナップショットはボリュームの完全なコピーです。それ以降のスナップショットでは増分のブロックレベルの変更のみが保存されます。

これは完全なボリュームデータを復元するための迅速で信頼性の高い方法です。部分的な復元のみが必要な場合は、別のデバイス名で実行中のインスタンスにボリュームをアタッチし、それをマウントしてから、オペレーティングシステムのコピーコマンドを使ってバックアップボリュームから本稼働ボリュームにデータをコピーします。

Amazon Elastic Cloud Compute ユーザーガイドで説明しているように、Amazon EBS スナップショットは、コンソールまたはコマンドラインから利用できる Amazon EBS スナップショットのコピー機能を使って AWS リージョン間でコピーできます[3]。この機能を使用して、基になるレプリケーションテクノロジーを管理することなく、別のリージョンにバックアップを保存できます。

---

[2] http://aws.amazon.com/ec2/

[3] http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ebs-copy-snapshot.html

EBS スナップショットの作成

スナップショットを作成すると、耐久性の高いディスクベースのストレージでデータを直接保護できます。Amazon EBS スナップショットは、AWS マネジメントコンソール、コマンドラインインターフェース (CLI)、または API を使って作成できます。

Amazon EC2 コンソールの [**Elastic Block Store Volumes**] ペインで [**Actions**] メニューの [**Create Snapshot**] を選択します。[**Create Snapshot**] ダイアログボックスで [**Create**] を選択して、Amazon S3 に保存されるスナップショットを作成します。

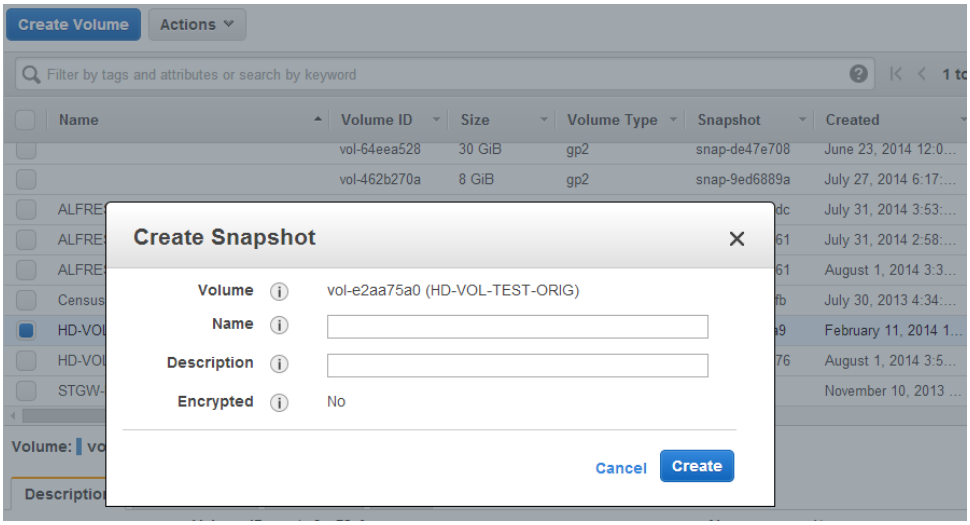

図 2: EC2 コンソールを使用したスナップショットの作成

CLI コマンドを使用してスナップショットを作成するには、次のコマンドを実行します。

```
➢  aws ec2 create-snapshot
```

定期的に `aws ec2 create-snapshot` コマンドをスケジュールおよび実行して、EBS データをバックアップできます。Amazon S3 の経済的な料金により、多くの世代のデータを維持することができます。そして、スナップショットはブロックベースであるため、最初のスナップショットの作成後に変更されたデータのスペースのみが消費されます。

EBS スナップショットからの復元

スナップショットからデータを復元するには、AWS マネジメントコンソール、CLI、または API を使用して、既存のスナップショットからボリュームを作成できます。

たとえば、以下のステップに従って、ボリュームを以前のポイントインタイムバックアップに復元します。

1. 次のコマンドを使用して、バックアップスナップショットからボリュームを作成します。

```
➢ aws ec2 create-volume --region us-west-1b --snapshot-
  id mysnapshot-id
```

2. Amazon EC2 インスタンスで、既存のボリュームをアンマウントします。

   Linux では `umount` を使用します。Windows では Logical Volume Manager (LVM) を使用します。

3. 次のコマンドを使用して、インスタンスから既存のボリュームをデタッチします。

```
➢ aws ec2 detach-volume --volume-id oldvolume-id –
  instance-id myec2instance-id
```

4. 次のコマンドを使用して、スナップショットから作成されたボリュームをアタッチします。

```
➢ aws ec2 attach-volume --volume-id newvolume-id
  --instance-id myec2instance-id --device /dev/sdf
```

5. 実行中のインスタンスでボリュームを再マウントします。

整合性のあるバックアップまたはホットバックアップの作成

バックアップを実行するときは、システムが I/O を実行していない状態にすることをお勧めします。マシンがトラフィックを受け入れていないことが理想的な状況ですが、24 時間休みない IT オペレーションが一般的になる中で、これはまれな状況になりつつあります。

このため、クリーンバックアップを実行するには、ファイルシステムまたはデータベースを休止させる必要があります。これを行う方法は、データベースまたはファイルシステムによって異なります。

データベース用の手順は次のとおりです。

- 可能な場合は、データベースをホットバックアップモードにします。
- Amazon EBS スナップショットのコマンドを実行します。
- データベースをホットバックアップモードから移行するか、リードレプリカを使っている場合は、リードレプリカインスタンスを削除します。

ファイルシステムの手順は類似していますが、オペレーティングシステムまたはファイルシステムの機能によって異なります。たとえば、XFS は一貫したバックアップ用にデータをフラッシュできるファイルシステムです。詳細については、「xfs_freeze」を参照してください[4]。

ファイルシステムがフリーズ機能をサポートしていない場合は、アンマウントし、スナップショットコマンドを発行してから、ファイルシステムを再マウントします。または、I/O のフリーズをサポートする論理ボリュームマネージャーを使用すると、このプロセスを容易に実行することができます。

スナップショットプロセスはバックグラウンドで続行され、スナップショットの作成は実行速度が速く特定時点をキャプチャするため、バックアップするボリュームは数秒アンマウントするだけで済みます。バックアップ時間は可能な限り短いため、停止時間は予測可能で、スケジュールすることができます。

[4] https://access.redhat.com/documentation/en-US/Red_Hat_Enterprise_Linux/6/html/Storage_Administration_Guide/xfsfreeze.html

複数ボリュームのバックアップの実行

場合によっては、論理ボリュームマネージャーを使ってスループットの可能性を高めることで、複数の Amazon EBS ボリューム間でデータをストライプすることができます。論理ボリュームマネージャー (mdadm や LVM など) を使用する場合、基になる EBS ボリュームからではなく、ボリュームマネージャーレイヤーからバックアップを実行することが重要です。これにより、すべてのメタデータの一貫性が確保され、サブコンポーネントのボリュームがわかりやすくなります。

これを完了するには、数多くの方法があります。たとえば、alestic.com[5] で作成されたスクリプトを使用できます。メモリバッファはディスクにフラッシュし、ディスクへのファイルシステムの I/O を停止して、RAID セットを構成するすべてのボリュームに対してスナップショットを同時に開始する必要があります。ボリュームのスナップショットを開始すると (通常は 1～2 秒後)、ファイルシステムはオペレーションを続行できます。スナップショットにはタグを付け、復元中にまとめて管理できるようにします。

また、論理ボリュームマネージャーやファイルシステムレベルからこれらのバックアップを実行することもできます。そのような場合は、従来のバックアップエージェントを使用すると、ネットワーク経由でデータをバックアップできます。インターネットおよび AWS Marketplace では、数多くのエージェントベースのバックアップソリューションを利用できます[6]。エージェントベースのバックアップソフトウェアでは、一貫性のあるサーバー名と IP アドレスが予期されることに注意してください。そのため、Amazon 仮想プライベートクラウド

---

[5] https://github.com/alestic/ec2-consistent-snapshot

[6] https://aws.amazon.com/marketplace/

(VPC)[7] でインスタンスをデプロイしてこれらのツールを使用することが、信頼性を確保するための最善の方法です。

別の手法は、1 つの大きなボリュームに存在するプライマリシステムボリュームのレプリカを作成することです。これによりバックアッププロセスが簡略化されます。これは 1 つの大きなボリュームのみをバックアップすればよく、バックアップはプライマリシステムでは実行されないためです。ただし、最初に 1 つのボリュームでバックアップ中に十分なパフォーマンスが得られること、および最大ボリュームサイズがアプリケーションに対して適切であることを確認する必要があります。

データベースのバックアップ手法

AWS には、データベース用の多くのオプションがあります。EC2 インスタンスで独自のデータベースを実行するか、Amazon Relational Database Service [8] (Amazon RDS) で提供されるいずれかのマネージド型サービスデータベースのオプションを使用できます。EC2 インスタンスで独自のデータベースを実行している場合は、ネイティブツール (MySQL[9]、Oracle[10]、MSSQL[11]、PostgreSQL[12] など) を使ってファイルにデータをバックアップするか、「EBS スナップショットベースの保護」で説明しているいずれかの方法を使ってデータを含むボリュームのスナップショットを作成できます。

---

[7] http://aws.amazon.com/vpc/

[8] https://aws.amazon.com/rds/

[9] http://dev.mysql.com/doc/refman/5.7/en/backup-and-recovery.html

[10] http://docs.oracle.com/cd/E11882_01/backup.112/e10642/rcmbckba.htm#BRADV8003

[11] http://msdn.microsoft.com/en-us/library/ms187510.aspx

[12] http://www.postgresql.org/docs/9.3/static/backup.html

データベースレプリカバックアップの使用

Amazon EBS ボリュームの RAID セットに構築されたデータベースの場合、データベースのリードレプリカを作成することで、プライマリデータベースでバックアップの負担をなくすことができます。これは別の Amazon EC2 インスタンスで実行されるデータベースの最新のコピーです。レプリカデータベースインスタンスは、ソースに類似した複数のディスクを使って作成するか、1 つの EBS ボリュームにデータを統合できます。次に、「EBS スナップショットベースの保護」で説明されているいずれかの手順を使って EBS ボリュームのスナップショットを作成できます。この手法は、24 時間休みなく実行される必要がある大きなデータベースでよく使用されます。その場合、必要なバックアップ時間が長すぎて、プロダクションデータベースをこのように長い時間停止することはできません。

バックアップのための Amazon RDS の使用

Amazon RDS には、データベースのバックアップを自動化するための機能が用意されています。Amazon RDS はデータベースインスタンスのストレージボリュームのスナップショットを作成し、個々のデータベースだけではなく、その DB インスタンス全体をバックアップします。

Amazon RDS には、DB インスタンスのバックアップと復旧を行うための 2 つの方法があります。

- **自動バックアップ**では、お客様の DB インスタンスのポイントインタイムリカバリが可能になります。自動バックアップは、新しい DB インスタンスを作成するとデフォルトでオンになります。Amazon RDS システムは、DB インスタンスを作成するときに定義する時間内に、データのフルバックアップ

を毎日実行します。自動バックアップの保持期間は、最大 35 日間まで設定できます。Amazon RDS が、トランザクションログと併せてこれらの定期データバックアップを使用することにより、最大で復元可能な最新時刻 (`LatestRestorableTime`) (一般的には 5 分前) まで、保持期間中の任意の時点の DB インスタンスを復元できます。DB インスタンスの最新の復旧可能時間を見つけるには、`DescribeDBInstances` API コールを使用するか、Amazon RDS コンソールでデータベースの [**Description**] タブを確認します。

ポイントインタイムリカバリを開始する際、DB インスタンスをリクエストされた時刻の状態に復元するために、最も適切なデイリーバックアップにトランザクションログを適用します。

- **DB スナップショット**はユーザーが開始するバックアップで、指定した頻度で既知の状態で DB インスタンスをバックアップすることができ、その後いつでもその状態に復旧することができます。DB スナップショットを作成するには、Amazon RDS コンソールまたは `CreateDBSnapshot` API コールを使用します。これらのスナップショットの保持期限はありません。コンソールまたは `DeleteDBSnapshot` API コールを使って明示的に削除するまで保持されます。

データベースを特定時点に復元するか、DB スナップショットから復元する場合、新しいエンドポイントを持つ新しいデータベースインスタンスが作成されます。このようにして、特定の DB スナップショットまたは特定時点から複数のデータベースインスタンスを作成できます。

AWS マネジメントコンソールまたは `DeleteDBInstance` 呼び出しを使用して古いデータベースインスタンスを削除できます。

AMI を使用した EC2 インスタンスのバックアップ

AWS は Amazon マシンイメージ (AMI) と呼ばれるシステムイメージを保存します。これらのイメージは、インスタンスの起動に必要なルートボリューム用のテンプレートで構成されます。AWS マネジメントコンソールまたは `aws ec2 create-image` CLI コマンドを使用して、AMI をルートボリュームとしてバックアップできます。

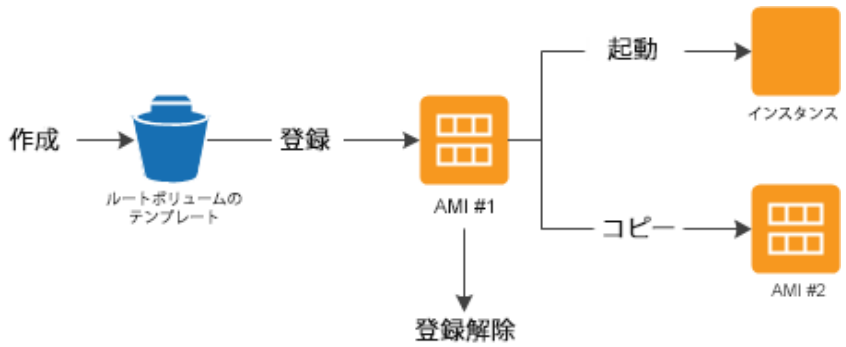

図 3: AMI を使用してインスタンスをバックアップし、起動する

AMI を登録すると、Amazon EBS スナップショットを使ってアカウントに保存されます。これらのスナップショットは Amazon S3 に置かれ、高い耐久性があります。

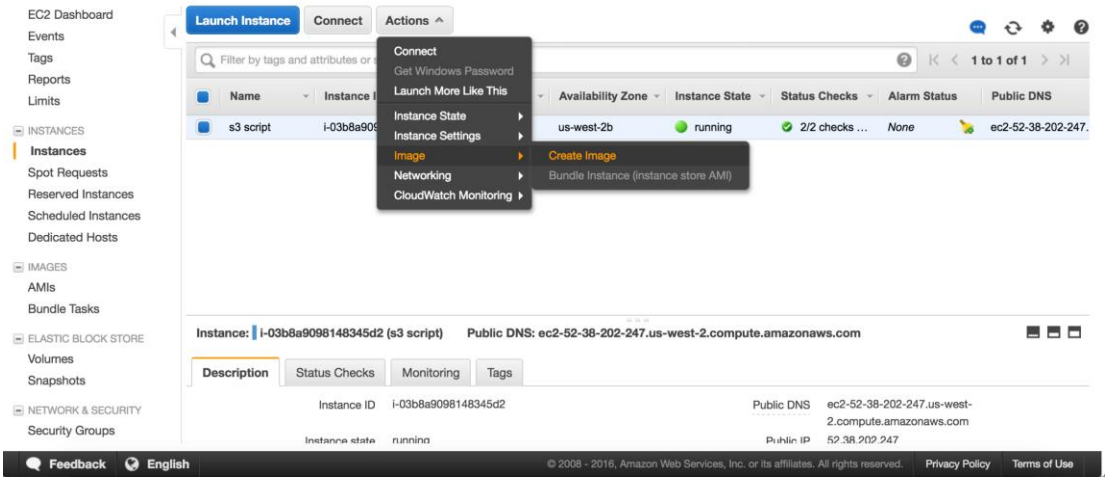

図 4: EC2 コンソールを使用してマシンイメージを作成する

Amazon EC2 インスタンスの AMI を作成したら、AMI を使用してインスタンスを再作成するか、インスタンスのより多くのコピーを起動することができます。また、アプリケーションの移行や災害対策用に、1 つのリージョンから別のリージョンに AMI をコピーすることもできます。

オンプレミスから AWS インフラストラクチャへ

このシナリオでは、クラウド上にコンポーネントがないワークロード環境について説明します。ウェブサーバー、アプリケーションサーバー、モニタリングサーバー、データベース、Active Directory などを含むすべてのリソースが、お客様のデータセンターまたはコロケーションを通じてホストされます。

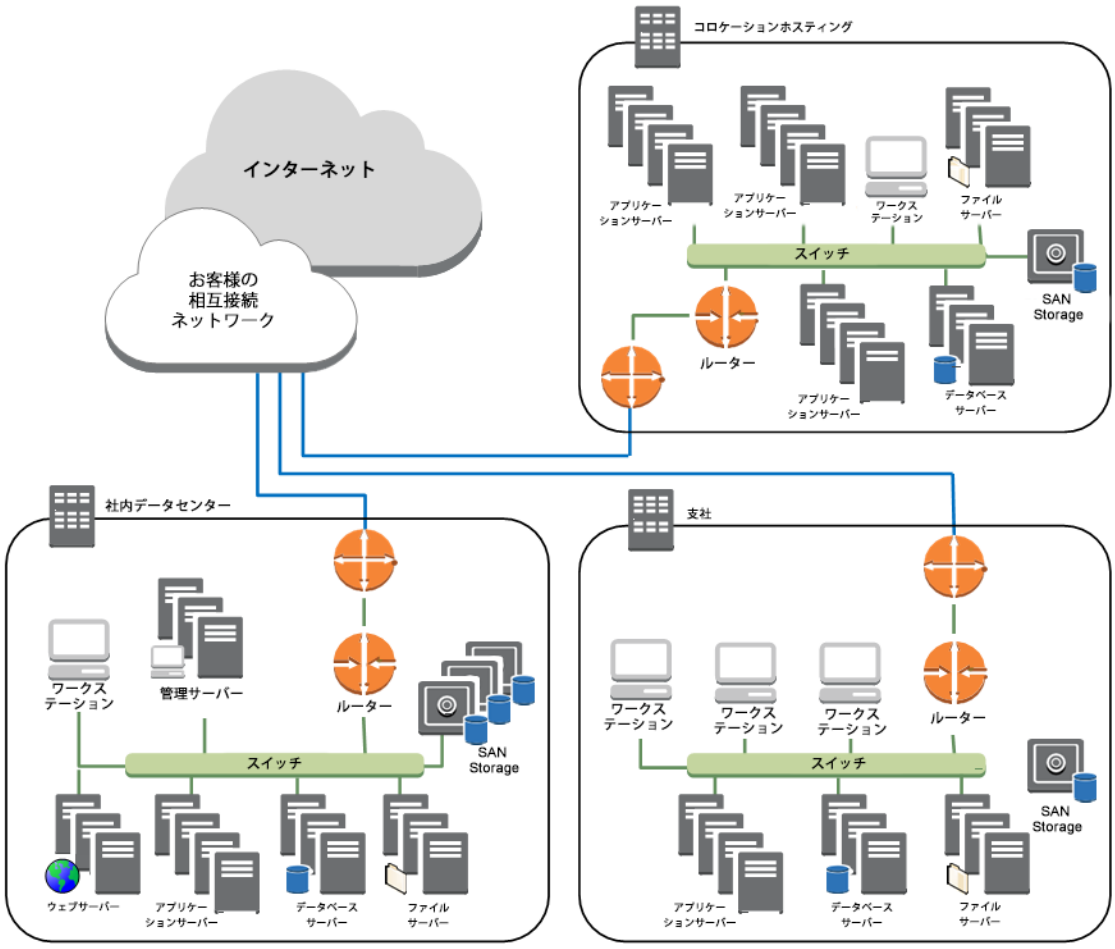

図 5: オンプレミス環境

このシナリオで AWS ストレージサービスを使用することで、バックアップおよびアーカイブタスクに集中できます。バックアップタスクを完了するために、ストレージのスケーリングまたはインフラストラクチャの容量について心配する必要はありません。

Amazon S3 および Amazon Glacier はネイティブな API ベースで、インターネットを通じて利用できます。これにより、ソフトウェアベンダーは次の図に示すようにアプリケーションを直接 AWS ストレージソリューションに統合できます。

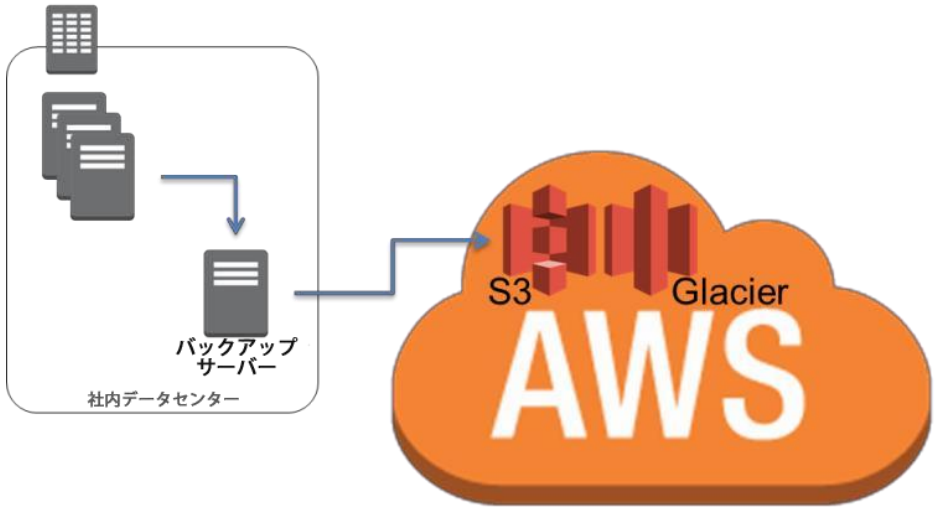

図 6: Amazon S3 または Amazon Glacier にコネクタをバックアップする

このシナリオでは、バックアップおよびアーカイブソフトウェアは API を通じて直接 AWS と連結されます。バックアップソフトウェアは AWS に対応しているため、オンプレミスサーバーから直接 Amazon S3 または Amazon Glacier にデータをバックアップします。

既存のバックアップソフトウェアがネイティブに AWS クラウドをサポートしていない場合は、AWS Storage Gateway 製品を使用できます。AWS Storage Gateway[13] は、お客様のデータセンターと AWS のストレージインフラストラクチャ間でシームレスかつセキュアな統合を実現する仮想アプライアンスです。このサービスを使用すると、AWS クラウドにデータを安全に保存し、スケーラブルで費用効率が高いストレージを利用できます。Storage Gateway はお客様の既存のアプリケーションと連携し、すべてのデータを暗号化して Amazon S3 または Amazon Glacier に安全に保存する、業界標準のストレージプロトコルをサポートします。

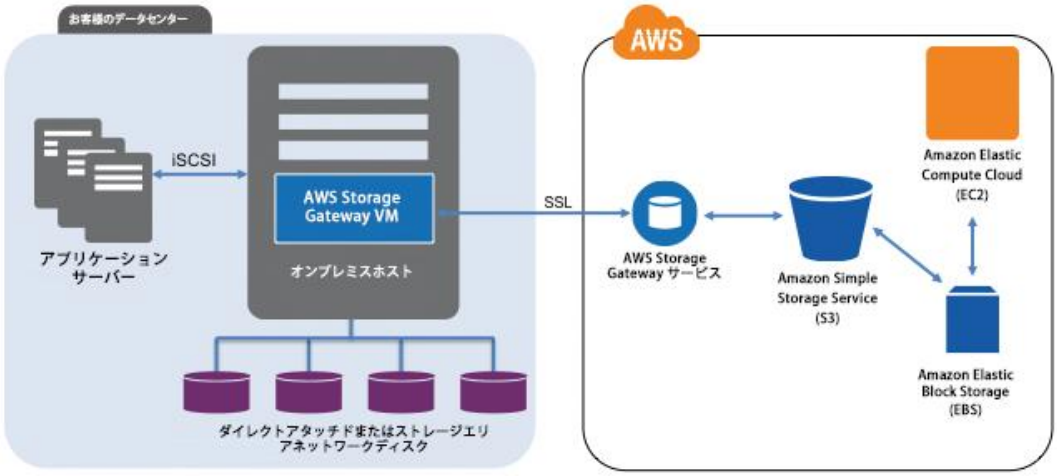

図 7: オンプレミスから AWS ストレージに接続する

AWS Storage Gateway は次の設定をサポートします。

[13] http://aws.amazon.com/storagegateway/

- **ボリュームゲートウェイ:** ボリュームゲートウェイは、オンプレミスのアプリケーションサーバーから iSCSI (Internet Small Computer System Interface) デバイスとしてマウントできる、クラウドベースのストレージボリュームを提供します。サポートするボリューム構成は以下のとおりです。
  - **ゲートウェイキャッシュ型ボリューム:** プライマリデータを Amazon S3 に保存し、頻繁にアクセスするデータをローカルに保持します。ゲートウェイキャッシュ型ボリュームはプライマリストレージのコストを大幅に削減し、ストレージをオンプレミスで拡張する必要を最小限に抑えます。また、頻繁にアクセスするデータへのアクセスを低レイテンシーに保つことができます。
  - **ゲートウェイ保管型ボリューム:** データセット全体への低レイテンシーアクセスが必要な場合に可能なオンプレミスデータゲートウェイ設定で、プライマリデータをローカルに保存し、そのデータのポイントインタイムスナップショットを Amazon S3 に非同期バックアップします。ゲートウェイ保管型ボリュームは、ローカルに復元するか Amazon EC2 から復元できる、耐久性が高く低コストのオフサイトバックアップを提供します。
- **ゲートウェイ仮想テープライブラリ (ゲートウェイ VTL):** ゲートウェイ VTL では、仮想テープの無制限のコレクションを持つことができます。各仮想テープシェルフは Amazon S3 によってバックアップされる仮想テープライブラリ、または Amazon Glacier によってバックアップされる仮想テープシェルフに保存できます。仮想テープライブラリは業界標準の iSCSI インターフェイスを公開します。これにより、お客様のバックアップアプリケーションは仮想テープにオンラインでアクセスすることができます。仮想テープに含まれているデータにすぐにアクセスする必要はない場合や、頻繁にアクセスする必要がない場合は、バックアップアプリケーションを使って仮想テー

プライブラリから仮想テープシェルフに移動して、さらにストレージコストを減らすことができます。

これらのゲートウェイは標準の iSCSI デバイスを提供するプラグアンドプレイデバイスとして機能し、バックアップまたはアーカイブフレームワークと統合できます。iSCSI ディスクデバイスはバックアップソフトウェアまたはゲートウェイ VTL 用のストレージプールとして使用し、テープベースのバックアップの負荷を軽減するか、Amazon S3 または Amazon Glacier に直接アーカイブすることができます。

この方法を使用すると、バックアップとアーカイブは自動的にオフサイトになり(コンプライアンスのため)、堅牢なメディアに保存されます。これにより、オフサイトのテープ管理の複雑さとセキュリティのリスクがなくなります。

ハイブリッド環境

これまでに説明した 2 つのインフラストラクチャデプロイであるクラウドネイティブとオンプレミスは、ワークロード環境にオンプレミスと AWS インフラストラクチャのコンポーネントがある、ハイブリッドシナリオに統合できます。ウェブサーバー、アプリケーションサーバー、モニタリングサーバー、データベース、Active Directory などを含むリソースが、お客様のデータセンターまたは AWS でホストされます。AWS クラウドで実行中のアプリケーションは、オンプレミスで実行中のアプリケーションに接続されます。

これは企業ワークロードで一般的になりつつあるシナリオです。多くの企業は独自のデータセンターを持っており、AWS を使用して容量を補強しています。多くの場合、それらのデータセンターは、高容量のネットワークリンクによって

AWS ネットワークに接続されています。たとえば、AWS Direct Connect[14] では、施設から AWS へのプライベートな専用接続を確立することができます。これにより、データ保護およびハイブリッドワークロードに対する一貫したパフォーマンスとレイテンシーのために、クラウドにデータをアップロードする帯域幅と一貫したレイテンシーが得られます。

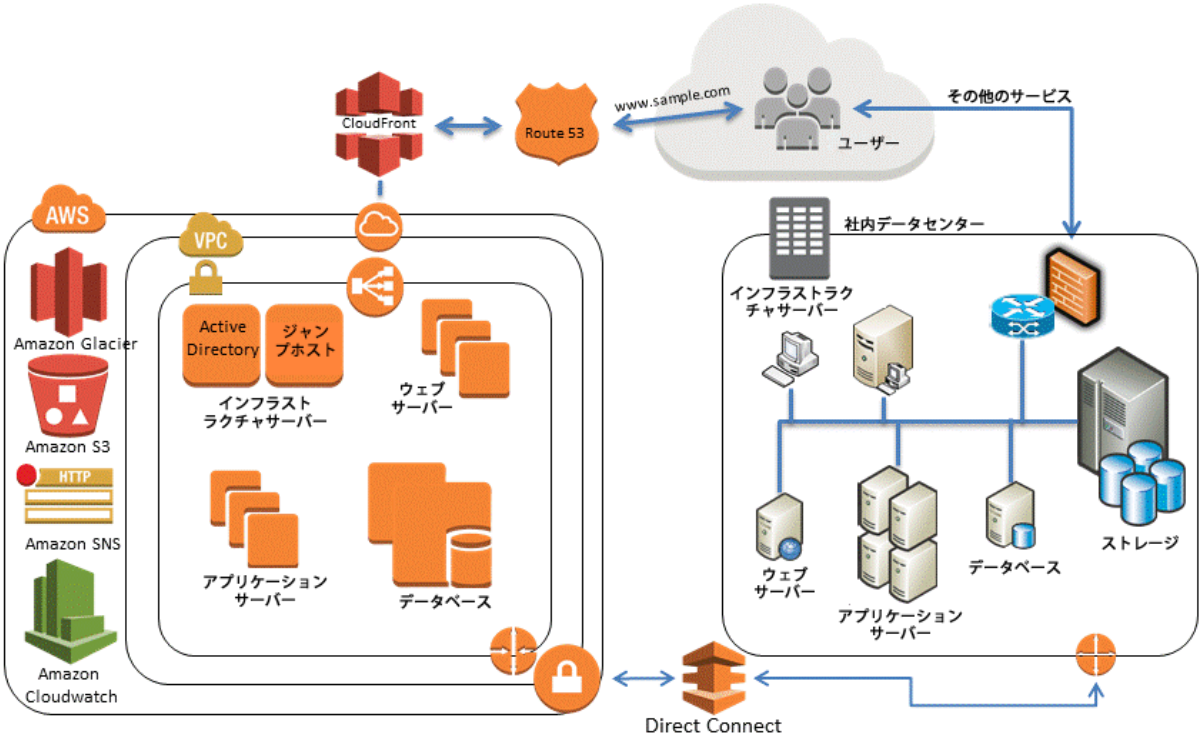

図 8: ハイブリッドインフラストラクチャのシナリオ

通常、優れた設計のデータ保護ソリューションでは、クラウドネイティブおよびオンプレミスソリューションで説明した方法を組み合わせて使用します。

[14] http://aws.amazon.com/directconnect/

AWS ベースのアプリケーションのデータセンターへのバックアップ

オンプレミスサーバーのデータをバックアップするフレームワークがすでにある場合は、それを VPN 接続または AWS Direct Connect 経由で簡単に AWS リソースに拡大できます。Amazon EC2 インスタンスでバックアップエージェントをインストールし、データ保護ポリシーに従ってバックアップできます。

可用性のためのバックアップ管理のクラウドへの移行

バックアップアーキテクチャによっては、マスターバックアップサーバーと 1 つ以上のメディアまたはストレージサーバーを、保護対象のサービスとともにオンプレミスに配置する場合があります。この場合、マスターバックアップサーバーを Amazon EC2 インスタンスに移動してオンプレミスの障害から保護し、可用性の高いバックアップインフラストラクチャを持つことができます。

バックアップデータの流れを管理するためには、1 つ以上のメディアサーバーを Amazon EC2 インスタンスで作成することもできます。Amazon EC2 インスタンスの近くにメディアサーバーがあると、インターネット転送の費用を節約でき、S3 または Amazon Glacier にバックアップすると、バックアップおよび復旧の全体的なパフォーマンスが向上します。

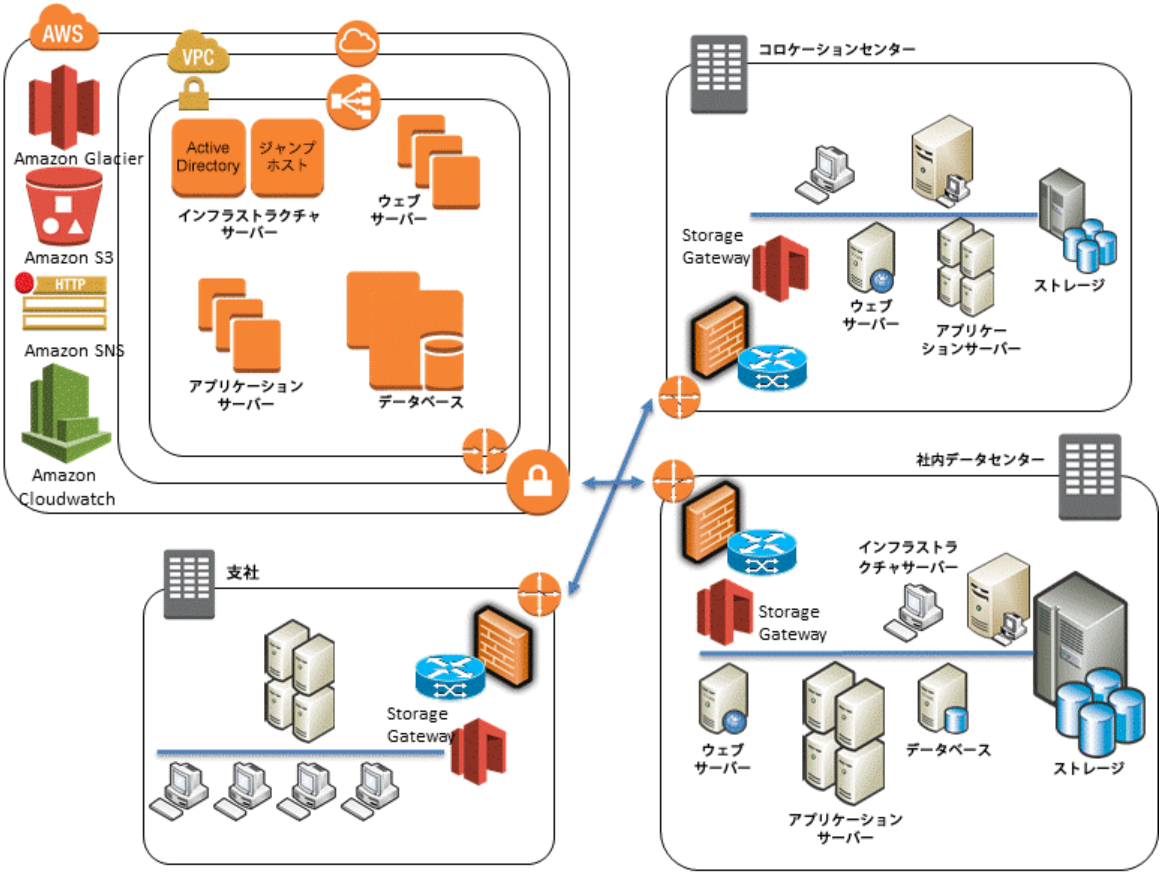

図 9: ハイブリッドシナリオでゲートウェイを使用する

ハイブリッドシナリオの例

Amazon EC2 インスタンス、スタンドアロンサーバー、仮想マシン、およびデータベースをバックアップする環境を管理しているとします。この環境には 1,000 台のサーバーがあり、オペレーティングシステム、ファイルデータ、仮想マシンイメージ、およびデータベースをバックアップします。バックアップするデータベースが 20 (MySQL、Microsoft SQL Server、Oracle の組み合わせ) あります。

バックアップソフトウェアには、オペレーティングシステム、仮想マシンイメージ、データボリューム、SQL Server データベース、および Oracle データベース (RMAN を使用) をバックアップするエージェントがあります。バックアップソフトウェアにエージェントがない MySQL などのアプリケーションでは、mysqldump クライアントユーティリティを使用して、標準のバックアップエージェントがデータをバックアップできるディスクにデータベースダンプファイルを作成します。

この環境を保護するため、おそらくサードパーティーのバックアップソフトウェアには、バックアップ、アーカイブ、および復元アクティビティや、ディスクベースのストレージ、Linear Tape-Open (LTO) テープドライブ、および AWS ストレージサービスに接続された複数のメディアサーバーを管理するグローバルカタログサーバーまたはマスターサーバーがあります。

AWS ストレージサービスを使ってバックアップソリューションを補強するための最も簡単な方法は、Amazon S3 または Amazon Glacier に対するバックアップベンダーのサポートを活用することです。ベンダーと連携して、ベンダーの統合およびコネクタオプションについて理解することをお勧めします。AWS と連携しているバックアップソフトウェアベンダーの完全なリストについては、「パートナーディレクトリ[15]」を参照してください。

既存のバックアップソフトウェアが、バックアップまたはアーカイブ用にクラウドストレージをネイティブにサポートしていない場合は、バックアップソフトウェアと Amazon S3 または Amazon Glacier の間で、ブリッジなどのストレージゲートウェイデバイスを使用できます。

---

[15] http://www.aws-partner-directory.com/PartnerDirectory/PartnerSearch?type=ISV

サードパーティーによる数多くのゲートウェイソリューションがあります。AWS Storage Gateway 仮想アプライアンスを使用して、このギャップを埋めることもできます。これは、iSCSI ベースのボリュームや仮想テープライブラリ (VTL) などの汎用的なテクニックが使用されるためです。この設定では、アプライアンスをホストするために、サポートされるハイパーバイザー (VMware または Microsoft Hyper-V) とローカルストレージが必要です。

AWS を使用したデータのアーカイブ

コンプライアンスまたは企業の理由でデータを保持する必要がある場合は、アーカイブすることができます。通常はデータ破損またはデータ損失から復旧するために短い期間のみ運用データのコピーを維持するために実行されるバックアップとは異なり、アーカイブでは保持ポリシーの有効期限が切れるまで、データのすべてのコピーが維持されます。

優れたアーカイブは以下の基準を満たします。

- 長期的な完全性のためのデータの耐久性
- データのセキュリティ
- 復元のしやすさ
- 低コスト

イミュータブルのデータストアが、規制またはコンプライアンスの要件となる場合があります。

Amazon Glacier は低コストでアーカイブを提供し、保管時のデータのネイティブな暗号化、ほぼ完全な耐久性、および無制限のキャパシティーを提供します。

データの迅速な取得を必要とするユースケースでは、Amazon S3 低頻度アクセスが最適です。Amazon Glacier は、データのアクセス頻度が低く、取得時間が数時間であるユースケースに最適です。

オブジェクトは、S3 のライフサイクルルールまたは Amazon Glacier API を通じて Amazon Glacier の階層にできます。Amazon Glacier のボールトロック機能では、ボールトロックポリシーを使用して、Amazon Glacier の各ボールトに対するコンプライアンス管理を簡単にデプロイして適用することができます。ボールトロックポリシーで「write once, read many」(WORM) などのコントロールを指定して、ポリシーをロックし、今後編集できないようにします。詳細については、「Amazon Glacier」を参照してください。

AWS でのバックアップデータの保護

データのセキュリティは共通の懸念事項です。AWS ではセキュリティを非常に真剣にとらえています。これは AWS が提供するすべてのサービスの基礎になります。Amazon S3 などのストレージサービスでは、保管時と転送時の両方で、アクセス制御と暗号化に関する強力な機能を提供しています。すべての Amazon S3 および Amazon Glacier API エンドポイントは、転送時のデータの SSL 暗号化をサポートします。デフォルトでは、Amazon Glacier は保管時のすべてのデータを暗号化します。Amazon S3 では、AWS に暗号化キーの管理を任せ、オブジェクトをアップロードするときに独自のキーを提供するか、暗号化キーに AWS Key Management Service (AWS KMS) [16]の統合を使用して、保管時のオブジェクトのサーバー側の暗号化を選択できます。または、AWS にアップロードする前に、いつでもデータを暗号化することができます。詳細については、「アマゾン ウェブ サービス: AWS セキュリティプロセスの概要」を参照してください。

---

[16] http://docs.aws.amazon.com/AmazonS3/latest/dev/UsingKMSEncryption.html

まとめ

ガートナーは AWS をパブリッククラウドストレージサービスにおけるリーダーに認定しました[17]。AWS は組織が次世代のバックアップであるクラウドベースのプラットフォームにワークロードを移行する支援を行うために最適な立場にあります。AWS は費用効率が高くスケーラブルなソリューションを提供し、組織がバックアップとアーカイブの要件をバランスできるようにします。これらのサービスは、現在お客様が使用しているテクノロジーとうまく統合します。

寄稿者

本書の執筆に当たり、次の者が寄稿しました。

- Pawan Agnihotri、ソリューションアーキテクト、アマゾン ウェブ サービス
- Lee Kear、ソリューションアーキテクト、アマゾン ウェブ サービス
- Peter Levett、ソリューションアーキテクト、アマゾン ウェブ サービス

ドキュメントの改訂

2016 年 5 月に更新

---

[17] http://www.gartner.com/technology/reprints.do?id=1-1WWKTQ3&ct=140709&st=sb