



AWS ホワイトペーパー

AWS クラウドで Microsoft Power BI を使用する



AWS クラウドで Microsoft Power BI を使用する: AWS ホワイトペーパー

Copyright © Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon の商標およびトレードドレスは、Amazon のものではない製品またはサービスと関連付けてはならず、また、お客様に混乱を招くような形や Amazon の信用を傷つけたり失わせたりする形で使用することはできません。Amazon が所有しない商標はすべてそれぞれの所有者に所属します。所有者は必ずしも Amazon と提携していたり、関連しているわけではありません。また、Amazon 後援を受けているとはかぎりません。

Table of Contents

AWS クラウドで Microsoft Power BI を使用する	1
要約	1
はじめに	2
Microsoft Power BI スイート	4
Power BI Desktop	4
Power BI サービス	4
Power BI Report Server	5
オンプレミスデータゲートウェイ	5
Microsoft Power BI Desktop を AWS データソースに接続する	6
Power BI Desktop をオンプレミスで使用する	6
インターネット経由でデータソースに接続する	6
AWS VPN を介したデータソースへの接続	10
AWS Direct Connect 経由でのデータソースへの接続	14
AWS クラウドで Microsoft Power BI Desktop を使用する	18
オプション 1: Amazon EC2 インスタンスに Microsoft Power BI Desktop をインストールする	19
オプション 2: Amazon WorkSpaces 環境に Microsoft Power BI をインストールする	19
オプション 3: Amazon AppStream 2.0 環境に Microsoft Power BI をインストールする	19
Microsoft Power BI Desktop 接続オプションの概要	22
Microsoft Power BI サービスを AWS データソースに接続する	24
推奨設定	25
追加の考慮事項	26
Amazon QuickSight を使用する	30
まとめ	32
寄稿者	33
その他の資料	34
改訂履歴	35
付録: Microsoft Power BI 対応 AWS データソース	36
Amazon Redshift	36
Amazon RDS	37
Amazon Athena	38
Amazon OpenSearch Service (Amazon Elasticsearch Service の後継サービス)	39
AWS Lake Formation	40
通知	42

AWS クラウドで Microsoft Power BI を使用する

AWS クラウドで Microsoft Power BI を使用する

公開日: 2021 年 11 月 3 日 ([改訂履歴](#))

要約

このホワイトペーパーでは、Microsoft Power BI (デスクトップ、レポートサーバー、サービス、およびオンプレミスデータゲートウェイ) をアマゾン ウェブ サービス (AWS) クラウドと統合して使用する方法について説明します。接続、セキュリティ、パフォーマンス、およびコスト最適化に重点を置いて、Microsoft Power BI 製品を Amazon Redshift、Amazon Athena、Amazon RDS、Amazon OpenSearch、AWS Lake Formation などの AWS のサービスとつなげることを検討しているお客様にオプションを提示します。

このホワイトペーパーは、Microsoft Power BI の概念と、AWS のサービスをデータソースとして使用する場合にこれらのテクノロジーを利用するオプションについて、簡単に理解したい IT の意思決定者やアーキテクトを対象としています。

はじめに

あらゆる規模のビジネスを営むお客様は、AWS の製品とサービスを利用して、信頼性とコスト効率に優れ、安全な方法でデータを保存しています。これは、データストレージや分析に関する成熟した提供物の幅広いエコシステムを利用できるおかげでもあります。提供物には、次のようなサービスが含まれます。

- [Amazon Simple Storage Service](#) (Amazon S3) は、シンプル、スケーラブル、安全で、コスト効率に優れたデータリポジトリを提供します。アプリケーションデータの保存に関する業界標準となっており、お客様がデータレイク用にまず選択するものとなっています。
- [Amazon Athena](#) はインタラクティブなクエリサービスで、Simple Storage Service (Amazon S3) 内のデータを標準 SQL を使用して簡単に分析できるようになります。
- [Amazon Relational Database Service](#) (Amazon RDS) を使用すると、クラウド内でのリレーショナルデータベースの設定、運用、スケーリングを簡単に実行できます。Amazon RDS は、コスト効率の高いサイズ変更が可能な容量を提供しながら、ハードウェアのプロビジョニング、データベースのセットアップ、パッチ適用、バックアップといった時間のかかる管理タスクを自動化します。SQL Server、Oracle Database、MySQL、MariaDB、PostgreSQL の各エンジンを利用できます。
- [Amazon Redshift](#) はフルマネージド型の非常にスケーラブルなデータウェアハウスで、構造化データセットと非構造化データセットの両方を簡単に分析できます。
- [Amazon QuickSight](#) は、クラウド駆動の高速なビジネスインテリジェンスサービスです。組織のすべてのユーザーにインサイトを提供します。
- [Amazon OpenSearch](#) (Amazon Elasticsearch Service の後継サービス) は、Elasticsearch を簡単に、コスト効率よく大規模に、デプロイ、保護、実行するフルマネージドサービスです。
- [AWS Lake Formation](#) は、安全なデータレイクを数日で簡単にセットアップできるサービスです。

サービスの相互関係をよりよく理解するために、データサービスをデータソースまたはデータコンシューマーとしてラベル付けすることがよくあります。データソースでは、お客様やアプリケーションがサービスにデータを保存したり、サービスからデータを取り出したりできます。また、データソースにはコンピューティング機能が組み込まれていることが多く、計算分析やフィルタリングを行うことができます。しかし要するに、データはこれらのデータソースにロードされ、最終的にデータコンシューマーによって取り出されます。Simple Storage Service (Amazon S3)、Amazon Athena、および Amazon Redshift が、データソースの典型例です。

一方、データコンシューマーは、データソースからのデータにアクセスし、通常はそのデータを処理します。オプションでデータを表示することもあります。Amazon QuickSight と Microsoft Power BI スイートが、データコンシューマーの好例です。データソースから読み取り、情報の分析、可視化、公開を支援します。

AWS では、データニーズに合わせて好みのテクノロジーを柔軟に組み合わせることが出来ます。多くのお客様がビジネスインテリジェンス (BI) のニーズに Amazon QuickSight を選択する一方、Microsoft Power BI、Tableau、Qlik などのベンダーを選択するお客様もいます。

このドキュメントでは、Microsoft Power BI の製品とサービスのスイートと、それらを AWS のサービスと組み合わせて使用方法を取り上げます。

Microsoft Power BI スイート

このホワイトペーパーでは、製品名の類似による混乱を避けるため、各 Microsoft Power BI 製品およびサービスについて紹介します。

Power BI Desktop

Power BI Desktop は、ローカルコンピュータにインストールする無料のアプリケーションです。これにより、データへの接続、データの変換、可視化が可能になります。Power BI Desktop を使用すると、複数の異なるデータソースに接続し、それらを組み合わせて (モデリングとも呼ばれる)、データモデルにできます。このデータモデルを使用して、組織内の他のユーザーとレポートとして共有できる可視化情報や、可視化情報のコレクションを構築できます。

Power BI Desktop は、ローカルまたはネットワーク経由で利用可能な、サポートされている任意のデータソースに接続できます。サポートされているデータソースについては、「[付録:Microsoft Power BI 対応 AWS データソース](#)」を参照してください。

ビジネスインテリジェンスプロジェクトに取り組むほとんどのユーザーは、Power BI Desktop を使用してレポートを作成します。次に、コンテンツを Power BI Report Server または Power BI サービスにプッシュして、他のユーザーとレポートを共有します。Power BI Desktop から Power BI Report Server または Power BI サービスにコンテンツをプッシュする操作は、公開と呼ばれます。詳細については、「[Power BI Desktop とは何ですか?](#)」を参照してください。

Note

Power BI Desktop は Windows 専用アプリケーションです。Linux、macOS、その他のオペレーティングシステムでは使用できません。

Power BI サービス

Power BI は、ソフトウェアサービス、アプリケーション、コネクタの集まりです。これらが連携して、お客様とお客様のビジネスに最も効果的なビジネスインサイトの作成、共有、利用を支援します。Power BI サービスは Power BI オンラインと呼ばれることもあり、Power BI のうち Software as a Service (SaaS) の部分です。詳細については、「[Power BI サービスとは何ですか?](#)」を参照してください。

Power BI サービスは、クラウドベースのサービスです。チームや組織の簡単なレポート編集とコラボレーションをサポートします。Power BI サービスのデータソースにも接続できますが、モデリングには制限があります。

ビジネスインテリジェンスプロジェクトに取り組むほとんどのレポート設計者は、Power BI Desktop を使用してレポートを作成し、Power BI サービスを使用してレポートを他のユーザーに配布します。この重要なコンポーネントの詳細については、「[Microsoft Power BI サービスを AWS データソースに接続する](#)」を参照してください。

Power BI Report Server

Power BI Report Server は、レポートと KPI を表示および管理する、Web ポータルを備えたプライベートレポートサーバーです。お客様は、レポートを Power BI サービスに公開したくない場合は、Power BI Report Server を使用します。Power BI Report Server は、もともとはオンプレミス環境を対象としていましたが、AWS でも実行できます。詳細については、「[AWS で Microsoft Power BI Report Server を使用する](#)」を参照してください。

オンプレミスデータゲートウェイ

Microsoft オンプレミスデータゲートウェイは、Power BI デプロイのセキュリティとパフォーマンスを向上できる、一般的にデプロイされているコンポーネントです。これにより、Power BI サービスは、プライベート化されたデータソースにアクセスできるようになります。プライベート化されたデータソースとは、別の施設にあり、データソースとデータゲートウェイ間の内部ネットワーク接続によってアクセスできるものです。通常、パーソナルモードはサーバーコンポーネントとしてインストールされますが、アプリケーションとしてローカルコンピュータにインストールすることもできます。このホワイトペーパーでは、標準 (サーバー) モードのみを取り上げます。詳細については、「[Microsoft Power BI サービスを AWS データソースに接続する](#)」を参照してください。

Microsoft Power BI Desktop を AWS データソースに接続する

Microsoft Power BI Desktop から始めたお客様は、オンプレミスのコンピュータやネットワークから AWS データソースに接続する方法に関心を持つことがほとんどです。デスクトップアプリケーションは通常、ローカルの Windows ラップトップで実行されているため、AWS データソースへの物理的および論理的な接続が、参入の最大の障壁として認識されています。

ただし、別のオプションもあります。Microsoft Power BI Desktop を AWS クラウドで実行するという方法です。このオプションにより、AWS データソースへの接続に関する障壁が大幅に軽減されますが、追加の考慮事項も必要になります。このセクションでは、両方のモデルについて説明します。接続性、セキュリティ、パフォーマンス、コストに関するそれぞれの影響を調査し、お客様が最適なオプションを決定できるようにします。このセクションで示すオプションは、Amazon RDS、Amazon Redshift、および Amazon Athena についてです。すべての AWS データソースの詳細については、「[付録: Microsoft Power BI 対応 AWS データソース](#)」を参照してください。

Power BI Desktop をオンプレミスで使用する

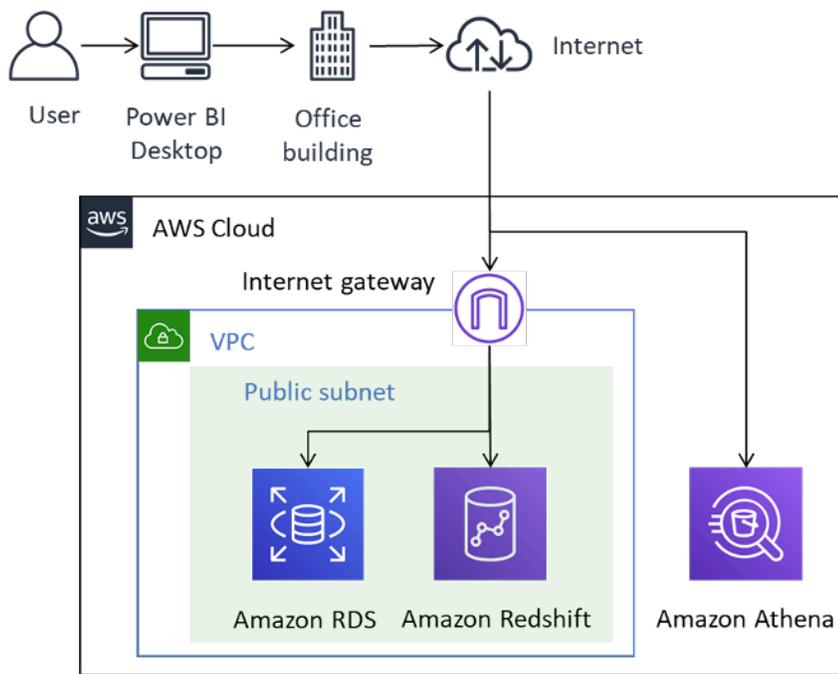
オンプレミスの Power BI Desktop を AWS クラウド内に保存されているデータソースとともに使用する場合、Power BI は次の 3 つの方法のいずれかでこれらのソースにアクセスできます。

- インターネットを使用してデータソースに接続する。
- AWS Virtual Private Network (AWS VPN) を使用してデータソースに接続する。
- AWS Direct Connect を使用してデータソースに接続する。

それぞれの方法については、次のセクションで詳しく説明します。

インターネット経由でデータソースに接続する

このモデルでは、Power BI Desktop アプリケーションは、インターネット経由で、インターネットからアクセス可能な AWS データソースの IP アドレスにルーティングされるアウトバウンド接続を確立します。例えば、お客様の Amazon Virtual Private Cloud (Amazon VPC) 内でインスタンス化されている Amazon RDS と Amazon Redshift は、インターネット経由でインスタンスにアクセスできるパブリックアクセシビリティオプションをサポートしています。お客様固有のリージョンのサービスエンドポイントを使用して、インターネットから直接 Amazon Athena をクエリできます。



インターネット経由での AWS データソースへの Power BI 接続

この接続方法は技術的には可能ですが、少数のユーザー以外にはお勧めしません。次の表に、重要な考慮事項を示します。

表 1 - インターネット経由で AWS データソースにアクセスする場合の考慮事項

条件	インターネット経由で AWS データソースにアクセスする場合の考慮事項
ネットワーク接続	データソースは、VPC 内のプライベート IP アドレスに接続するか、リージョンまたは VPC のサービスエンドポイントを使用することで利用できます。Power BI Desktop は VPN 経由で接続し、DNS 設定に応じて、データソース (Amazon RDS、Amazon Redshift、Amazon EC2 ベースのデータソース) に直接アクセスするか、リージョンエンドポイントを持つサービス (Amazon Athena) のデータソースにプライベート VPC エンドポイントまたはリージョンエンドポイントを使用してアクセスします。
セキュリティ	IP アクセスコントロール

条件	インターネット経由で AWS データソースにアクセスする場合の考慮事項
	<p>セキュリティグループは、インスタンスの仮想ファイアウォールとして機能し、インバウンドトラフィックとアウトバウンドトラフィックをコントロールします。アクセスを信頼できるエンティティに制限するには、既知の Classless Inter-Domain Routing (CIDR) 範囲に関連付けられたインバウンド IP 範囲のみを許可するようにセキュリティグループを設定します。</p> <p>転送時の暗号化</p> <p>AWS では、Amazon RDS、Amazon Redshift、Amazon EC2 ベースのデータソースなど、パブリック IP アドレスを使用するデータソースに対して暗号化を設定することをお勧めします。これにより、転送中にデータや認証情報が漏洩するリスクが軽減されます。暗号化の設定に失敗すると、重大なリスクが生じます。この点を見落とさないでください。</p> <p>Amazon Athena などのリージョンサービスエンドポイントは TLS で暗号化されています。さらに、JDBC または ODBC クライアントにストリーミングされる Amazon Athena のクエリ結果は、Transport Layer Security (TLS) を使用して暗号化されます。</p> <p>認証と認可</p> <p>AWS では、データセットへの読み取り専用アクセスを提供する認証情報を使用し、会社のポリシーに従って認証情報をローテーションするプロセスを設定することを推奨しています。</p>

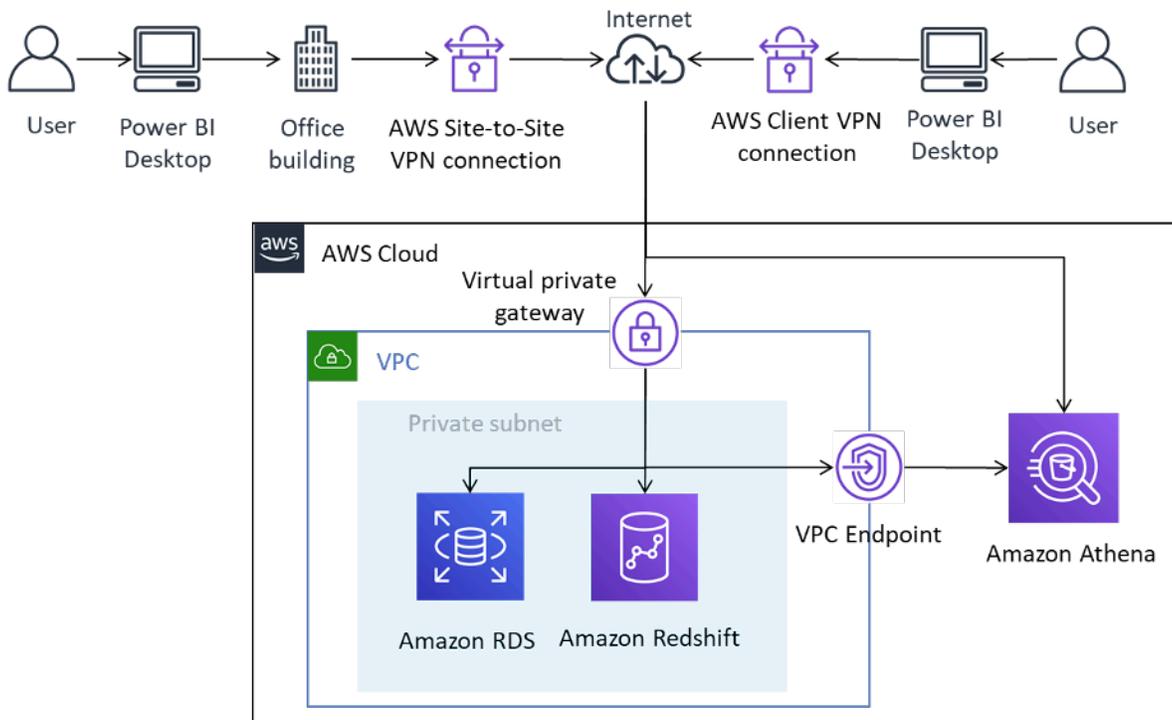
条件	インターネット経由で AWS データソースにアクセスする場合の考慮事項
パフォーマンス	<p>インターネット経由で AWS データソースにアクセスする場合に Power BI Desktop の全体的なパフォーマンスに影響を与える可能性のある要因には、次のようなものがあります。</p> <ul style="list-style-type: none">• アクセスされるデータセットのサイズ。データセットが大きいほど、取得に時間がかかります。インターネット経由で取得されるデータ量を減らすために、クエリを制限し、フィルターを使用することをお勧めします。• 帯域幅、レイテンシー、パケット損失など、インターネット接続の品質。可能な場合は、レイテンシーの影響を軽減するために、地理的に近い AWS リージョンのデータにアクセスしてください。インターネットを共有している場合は、オフピーク時にデータソースをロードし、十分な帯域幅を確保することを検討してください。 <p>一般的に、AWS では、1 日のさまざまな時間帯に、さまざまなデータセットを使用し、ユーザー数を徐々に増やしながらかエクスペリエンスをテストすることを推奨しています。</p>
コスト	<p>VPC 内に存在し、パブリック IP アドレスを使用してインターネット経由でクエリされたデータソースには、標準の Amazon VPC データ送信料金が発生します。コストを削減するために、クエリを制限し、フィルターを使用して、インターネット経由で取得するデータ量を減らすことをお勧めします。</p>

AWS VPN 経由でのデータソースへの接続

このモデルでは、Power BI Desktop インストールは、AWS ネットワーク内のデータソースに、AWS Site-to-Site VPN または AWS Client VPN という 2 つの AWS VPN 方式のいずれかを使用して接続します。各接続タイプは、ネットワークトラフィックを保護する、高可用性かつ伸縮自在なマネージド型クラウド VPN ソリューションを提供します。

Site-to-Site VPN は、お客様のネットワークと AWS VPN または AWS Transit Gateway との間に暗号化されたトンネルを作成します。Client VPN は、無料の VPN ソフトウェアクライアントを使用して、ユーザーを AWS またはオンプレミスのリソースに接続します。

Site-to-Site VPN 接続と Client VPN 接続のどちらも、VPN トラフィックは VPC で停止します。したがって、プライベート IP アドレスにルーティングできるため、インスタンスにパブリック IP アドレスは必要なくなります。Athena など、一般公開されているサービスエンドポイントからアクセス可能なデータパスを持つサービスの場合、これらのサービスリクエストはインターネット経由または VPN 接続経由で VPC エンドポイント経由でルーティングできます。



Power BI Desktop の Site-to-Site VPN および Client VPN を介した AWS データソースへの接続

Site-to-Site VPN は AWS Transit Gateway にも接続でき、複数の VPC にまたがるデータソースへのアクセスを容易にします。

AWS VPN を使用すると、AWS に保存されているデータソースにアクセスするときに、各データソースを明示的に設定しなくても暗号化を採用できるという利点があります。一度設定すると、エンドユーザーにとってほぼシームレスになります。

表 2 - AWS VPN を使用して AWS データソースにアクセスする場合の考慮事項

条件	AWS VPN を使用して AWS データソースにアクセスする場合の考慮事項
ネットワーク接続	<p>データソースは、VPC 内のプライベート IP アドレスに接続するか、リージョンまたは VPC のサービスエンドポイントを使用することで利用できます。Power BI Desktop は VPN 経由で接続し、DNS 設定に応じて、データソース (Amazon RDS、Amazon Redshift、Amazon EC2 ベースのデータソース) に直接アクセスするか、リージョンエンドポイントを持つサービス (Amazon Athena) のデータソースにプライベート VPC エンドポイントまたはリージョンエンドポイントを使用してアクセスします。</p>
セキュリティ	<p>IP アクセスコントロール</p> <p>ルーティングとセキュリティグループを組み合わせ、AWS クラウドに保存されているデータソースへのアクセスをコントロールできます。</p> <p>転送時の暗号化</p> <p>AWS VPN はどちらのタイプも IPsec 暗号化を使用します。つまり、転送されるデータは AWS とオンプレミスの間を移動する際に暗号化されます。これにより、データソースが暗号化された通信を使用するように構成されていなくても、そのデータはインターネットを通過している間も確実に保護されます。</p> <p>認証</p>

条件	AWS VPN を使用して AWS データソースにアクセスする場合の考慮事項
	<p>Site-to-Site VPN は 1 回限りの設定が必要で、一度確立されるとユーザーに対してシームレスになります。エンドユーザーは、Site-to-Site VPN を使用するための認証は不要ですが、データソースに対する認証が必要です。</p> <p>一方、Client VPN では、接続を確立するためにエンドユーザーによる認証が必要です。Client VPN 認証は、アクティブディレクトリ (ユーザーベース)、相互認証 (証明書ベース)、または SAML SSO (ユーザーベース) を介して実行できます。認証されると、エンドユーザーにとって接続はシームレスになります。Power BI Desktop に追加された AWS データソースには認証が必要です。</p> <p>AWS では、必要なデータセットに対してのみ読み取り専用アクセス許可を持つアイデンティティを使用して AWS データソースを認証することを推奨しています。</p>

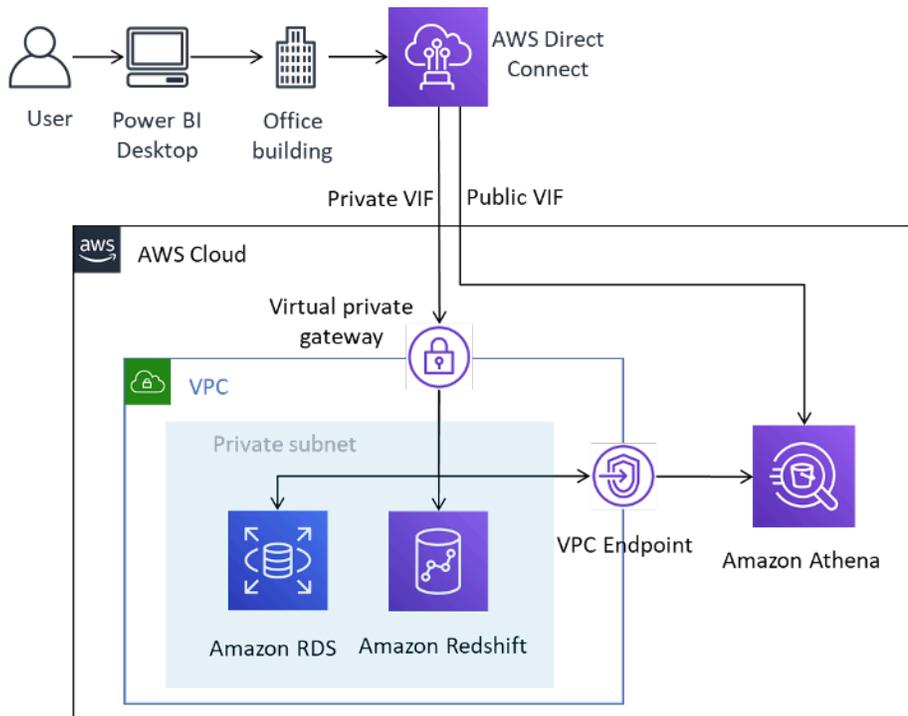
条件	AWS VPN を使用して AWS データソースにアクセスする場合の考慮事項
パフォーマンス	<p>AWS VPN の使用はインターネットを介して行われます。そのため、そのパフォーマンスエンベロープは、最初に示したシナリオと似ています。インターネット経由で AWS データソースにアクセスする場合に、Power BI Desktop の全体的なパフォーマンスに影響を与える可能性のある要因がいくつかあります。例えば以下のようなものです。</p> <ul style="list-style-type: none">• アクセスされるデータセットのサイズ。データセットが大きいほど、取得に時間がかかります。インターネット経由で取得されるデータ量を減らすために、クエリを制限し、フィルターを使用することをお勧めします。• 帯域幅、レイテンシー、パケット損失など、インターネット接続の品質。可能な場合は、地理的に近い AWS リージョンのデータにアクセスします。これにより、レイテンシーの影響が軽減されます。インターネットを共有している場合は、オフピーク時にデータソースをロードし、十分な帯域幅を確保することを検討してください。 <p>一般的に、AWS では、1 日のさまざまな時間帯に、さまざまなデータセットを使用し、ユーザー数を徐々に増やしながらかエクスペリエンスをテストすることを推奨しています。</p>

条件	AWS VPN を使用して AWS データソースにアクセスする場合の考慮事項
コスト	VPC 内に存在し、AWS VPN を使用してクエリされたデータソースには、標準の AWS VPN データ転送料金がかかります。コストを削減するために、クエリを制限し、フィルターを使用して、インターネット経由で取得するデータ量を減らすことをお勧めします。

AWS Direct Connect 経由でのデータソースへの接続

AWS Direct Connect は、お客様の内部ネットワークを AWS Direct Connect ロケーションに、標準のイーサネット光ファイバケーブルを介して接続するサービスです。ケーブルの一端がお客様のルーターに、他方が AWS Direct Connect のルーターに接続されます。この接続を使用すると、パブリックな AWS のサービス (例えば Simple Storage Service (Amazon S3)) または Amazon VPC への仮想インターフェイスを直接作成できるので、お客様のネットワークパスの中でインターネットサービスプロバイダーをバイパスできるようになります。AWS Direct Connect ロケーションとは、対応するリージョンでの AWS へのアクセスの中継地となるものです。パブリックリージョンまたは AWS GovCloud (米国) の単一の接続を使用して、他のすべてのパブリックリージョンのパブリック AWS のサービスにアクセスすることができます。

このモデルでは、お客様のオンプレミスネットワークが AWS Direct Connect 経由で AWS ネットワークに直接接続されます。AWS Direct Connect の設定には複数の方法がありますが、最も単純なモードでは、プライベート仮想インターフェイス (プライベート VIF) と呼ばれる構造を使用して、VPC 内の IP 範囲へのアクセス許可が付与されます。インターネット/パブリック IP 範囲には、パブリック仮想インターフェイス (パブリック VIF) を使用してアクセスします。



Power BI Desktop を AWS Direct Connect 経由で AWS データソースに接続する

Power BI でデータソースを追加する場合、DNS の設定に応じて、VPC にある場合はプライベート IP アドレス、またはサービスのプライベート VPC エンドポイントを選択します。

表 3 - AWS Direct Connect を使用して AWS データソースにアクセスする場合の考慮事項

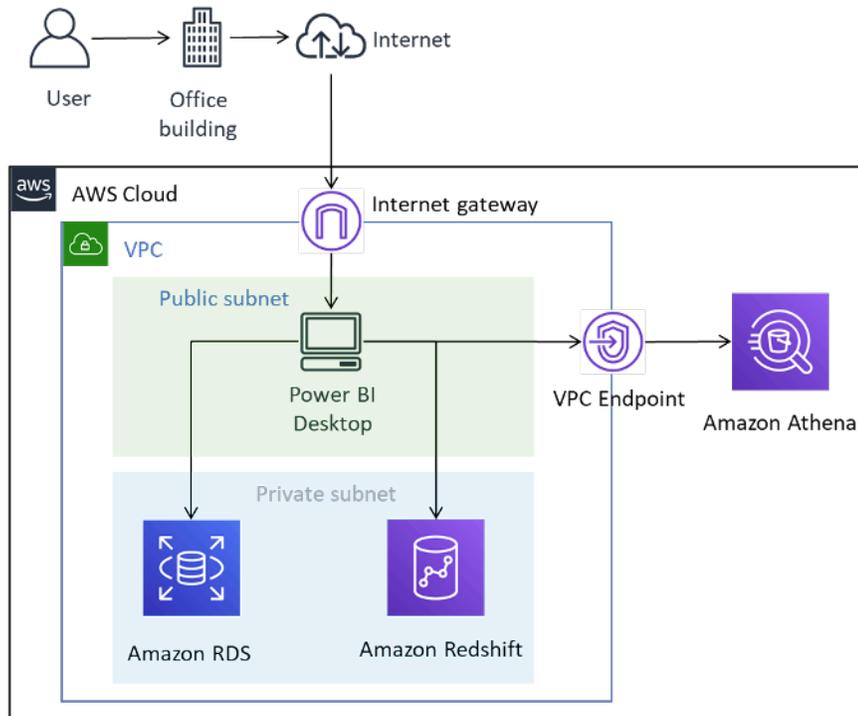
条件	AWS Direct Connect を使用して AWS データソースにアクセスする場合の考慮事項
ネットワーク接続	Direct Connect を設定すると、VPC 内のプライベート IP に接続するか、リージョンのサービスエンドポイントを使用して、データソースにアクセスできます。
セキュリティ	<p>IP アクセスコントロール</p> <p>ルーティングとセキュリティグループを組み合わせ、AWS クラウドに保存されているデータソースへのアクセスをコントロールできます。</p> <p>転送時の暗号化</p>

条件	AWS Direct Connect を使用して AWS データソースにアクセスする場合の考慮事項
	<p>Direct Connect では、行レベルのデータ暗号化は提供されません。データソースレベルで TLS 暗号化を有効にするか、Site-to-Site VPN と Direct Connect を組み合わせて、送信中にデータや認証情報が危険にさらされないようにすることをお勧めします。</p> <p>認証</p> <p>Direct Connect では、構成後に追加の認証を行う必要はありません。</p> <p>AWS では、必要なデータセットに対してのみ読み取り専用アクセス許可を持つアイデンティティを使用して AWS データソースを認証することを推奨しています。</p>

条件	AWS Direct Connect を使用して AWS データソースにアクセスする場合の考慮事項
パフォーマンス	<p>Direct Connect では、1 Gbps ポートと 10 Gbps ポートを使用できます。Direct Connect をサポートする APN パートナーから、50 Mbps、100 Mbps、200 Mbps、300 Mbps、400 Mbps、500 Mbps の速度を注文できます。</p> <p>Direct Connect 接続は、インターネットベースの接続と比較すると、通常、帯域幅が増加し、レイテンシーが減少します。接続の容量によっては、データソースのロード時に Power BI Desktop のパフォーマンスがネットワークによって制約されなくなる場合があります。そのため、Direct Connect が関連付けられている同じリージョンのデータセットにアクセスすると、ユーザーはクエリに対して優れたパフォーマンスを期待できます。</p> <p>大きなデータセットのロードには注意が必要です。Power BI Desktop には 10 GB のデータセット制限があることに注意してください。</p>
コスト	<p>Direct Connect 接続を使用してアクセスされるデータソースには、標準の Direct Connect ポート料金に加えて、AWS Direct Connect の料金で説明されているように、標準の Direct Connect データ転送料金がかかります。Direct Connect 経由で AWS クラウドに送信されるデータには料金は発生しません。</p> <p>コストを削減するために、クエリを制限し、フィルターを使用して、取得するデータ量を減らすことをお勧めします。</p>

AWS クラウドで Microsoft Power BI Desktop を使用する

AWS クラウドでの Microsoft Power BI Desktop の使用は、前のセクションで説明した多くの課題に対する一般的な解決策です。このモデルでは、お客様は AWS クラウドで Microsoft Power BI Desktop をホストし、オンプレミスでリモートアクセスします。次の図表に例を示します。



AWS クラウドにデプロイされた Microsoft Power BI Desktop

この図表は、インターネット経由でのデスクトップへのユーザー接続を示していますが、AWS VPN と Direct Connect も有効な接続タイプです。グラフィカルな管理トラフィックのみが送信されるため、帯域幅の要件は一般的なインターネット接続に適しています。

このモデルでは、Microsoft Power BI Desktop はパブリックサブネット内の Amazon VPC 内でホストされ、Amazon RDS や Amazon Redshift などのプライベート IP アドレスを持つデータソースに直接ネットワーク接続できます。Amazon Athena やリージョンのその他のサービスに接続するには、VPC エンドポイント接続を宛先として使用するか (図表を参照)、リージョンのパブリックサービスエンドポイントを使用します。

AWS クラウドで Microsoft Power BI Desktop をホストするには、いくつかのオプションがあります。

オプション 1: Amazon EC2 インスタンスに Microsoft Power BI Desktop をインストールする

このオプションでは、Microsoft Windows Server がインストールされた Amazon EC2 インスタンスを 1 つ以上作成します。ユーザーはリモートデスクトッププロトコル (RDP、Remote Desktop Protocol) アプリケーションを使用してインスタンスに接続し、Microsoft Power BI を、ラップトップにインストールされているかのようにインストール、構成、使用します。デフォルトでは、Windows Server のインストールは同時に 2 つの RDP クライアントセッションに制限されています。これ以上必要な場合は、Microsoft リセラーからクライアントアクセスライセンスを購入できます。1 つのサーバーで同時ユーザー数を増やすと、リソースの競合が発生することがあります。インスタンス数を増やしてそれぞれのユーザー数を減らすことで、ユーザーエクスペリエンスが向上する場合があります。

オプション 2: Amazon WorkSpaces 環境に Microsoft Power BI をインストールする

Amazon WorkSpaces は、マネージド型でセキュアなサービスとしてのデスクトップ (DaaS) のソリューションです。Amazon WorkSpaces を使用すれば、Windows または Linux のデスクトップをわずか数分でプロビジョンし、すばやくスケールして世界中の作業者に数千のデスクトップを提供できます。

このオプションでは、ユーザー用に Microsoft Power BI Desktop がインストールされた Microsoft Windows デスクトップをプロビジョンします。ユーザーは、Android、iOS、Fire、Mac、PC、Chromebook、Linux デバイスで使用できる軽量のクライアントアプリケーションを使用して Windows デスクトップにアクセスできます。

Amazon WorkSpaces を使用すると、Microsoft Power BI Desktop ユーザーは専用のリソースを持ち、デスクトップを自動的に停止してコストを節約できます。このオプションには、Windows 以外のデバイスから Amazon WorkSpaces にアクセスできるだけでなく、Microsoft Office や他の Windows アプリケーションへのアクセスなど、他のタスクを実行するためのデスクトップ環境をユーザーに提供できるという利点もあります。

オプション 3: Amazon AppStream 2.0 環境に Microsoft Power BI をインストールする

Amazon AppStream 2.0 は、フルマネージド型のアプリケーションストリーミングサービスです。デスクトップアプリケーションを AppStream 2.0 で集中管理し、任意のコンピュータのブラウザへ安

全に配信できます。ユーザー数無制限で世界中に簡単にスケールでき、ハードウェアやインフラストラクチャの調達、プロビジョニング、オペレーションは不要です。

このオプションでは、HTML5 準拠のウェブブラウザにストリーミングされる Microsoft Power BI Desktop アプリケーションのみをオンプレミスユーザーに提供します。プラグインは不要です。Amazon WorkSpaces オプションと同様に、Windows 以外のデバイスも接続でき、ユーザーは専用のリソースにアクセスできます。

このオプションの欠点の 1 つは、AppStream 2.0 アプリケーションはデフォルトで 1 GB のストレージ容量に制限されているため、大規模なモデルやレポートの保存には適さない可能性があることです。Amazon FSx for Windows ファイルサーバーのファイルストレージを組み合わせることで容量を増やすことができますが、お客様が認識すべき複雑さが増します。

表 4 - AWS クラウドで実行される Power BI Desktop に関する考慮事項

条件	AWS クラウドで実行される Power BI Desktop に関する考慮事項
ネットワーク接続	<p>データコンシューマーとデータソースの両方が AWS クラウドに存在するため、データソースへのネットワーク接続は簡単です。Amazon RDS や Amazon Redshift などの Amazon VPC に存在するデータソースには、直接アクセスできません。リージョンエンドポイントを使用するデータソースには、Amazon VPC インターネットゲートウェイまたは Amazon VPC エンドポイントを介してアクセスできます。</p> <p>Microsoft Power BI Desktop への接続は、インターネット、AWS VPN、または AWS Direct Connect を介して行われます。3 つのオプションにはすべて、ほとんどのインターネット接続が満たすことができる適度な要件があります。</p>
セキュリティ	<p>IP アクセスコントロール</p> <p>お客様は、ルーティングとセキュリティグループを組み合わせ、AWS クラウドに保存され</p>

条件	AWS クラウドで実行される Power BI Desktop に関する考慮事項
	<p>ているデータソースへのアクセスをコントロールできます。</p> <p>1 つ目のオプション (Amazon EC2 を使用) では、追加でルーティングとセキュリティグループを組み合わせて使用し、特定のオンプレミスの CIDR 範囲のみを許可することもできます。</p> <p>転送時の暗号化</p> <p>Amazon VPC 内のデータソースは、データの送信に暗号化を使用するよう設定することをお勧めします。リージョン別のサービスは既に TLS 暗号化を利用しています。</p> <p>管理アクセスは 3 つのオプションすべてで暗号化されます。</p> <p>認証</p> <p>AWS では、必要なデータセットに対してのみ読み取り専用アクセス許可を持つアイデンティティを使用して AWS データソースを認証することを推奨しています。</p> <p>3 つのオプションはすべて、リモートユーザーが Microsoft Power BI Desktop をホストする環境にアクセスする前に本人確認を行うために、認証メカニズムを使用する必要があります。通常、このメカニズムはユーザー名とパスワードですが、MFA オプションも使用できます。</p>

条件	AWS クラウドで実行される Power BI Desktop に関する考慮事項
パフォーマンス	AWS クラウドで実行する場合、Microsoft Power BI Desktop のパフォーマンスは通常、オンプレミスで実行する場合よりも優れています。ネットワーキングインスタンスとコンピューティングインスタンスはどちらも最新であり、仕様も高く、アプリケーションとデータソース間のレイテンシーは低いと考えられます。
コスト	<p>Microsoft Power BI Desktop を AWS クラウドに配置することで、データ転送料金を大幅に削減したり、完全に排除したりできます。複数のアベイラビリティーゾーン、異なる VPC、または異なるリージョンのデータソースにアクセスする場合は注意が必要です。データ転送料金が適用される場合があるためです。</p> <p>オプションごとに、考慮する必要がある追加のコスト考慮事項があります。Amazon EC2、Amazon WorkSpaces、および Amazon AppStream 2.0 は、使用量に基づいて料金が設定されます。詳細については、各サービスの料金ページを参照してください。</p>

Microsoft Power BI Desktop 接続オプションの概要

データセットの要件が軽い少数のユーザーには、Microsoft Power BI Desktop をオンプレミスで実行し、インターネット経由で安全に接続する、または AWS VPN を使用して接続するのが適切なソリューションです。このモデルでは、セキュリティが構成され、維持されていることを確認してください。また、この構成をテストして、ユーザーの期待するパフォーマンスを満たしているかどうかを判断することをお勧めします。

ユーザー数が増えるにつれて、AWS Direct Connect を介した接続を検討することを推奨します。Direct Connect を使用すると、より大きなデータセットをロードする際のユーザーエクスペリエ

ンスが向上します。大きなデータセットの転送に伴うコストへの影響をユーザーが認識していることを確認してください。

AWS クラウドでの Microsoft Power BI Desktop の実行を評価することをお勧めします。これにより、エンドユーザーには最高のパフォーマンスエクスペリエンスが、クラウド管理者には最高の管理エクスペリエンスが提供される可能性が高くなります。特に Amazon WorkSpaces は、少数のユーザーから数千人のユーザーまでスケールできます。また、これらのサービスは、セキュリティと管理に多大なメリットをもたらします。

Microsoft Power BI サービスを AWS データソースに接続する

Microsoft Power BI サービス (SaaS) は、インターネットからアクセスできるデータソース、または Amazon VPC 内のプライベートデータソースに、直接接続できます。プライベートデータソースに接続するには、Microsoft オンプレミスデータゲートウェイと呼ばれるアプリケーションコンポーネントが必要です。Microsoft オンプレミスデータゲートウェイは、VPC 内の Amazon EC2 インスタンスにダウンロードおよびインストールされ、Microsoft Power BI 認証情報を使用して設定されます。このゲートウェイは、インターネット経由で Microsoft Azure Service Bus へのアウトバウンド接続を確立し、Microsoft Power BI 内でアクセス可能なデータソースに接続するように構成されます。大規模なデプロイでは、複数のオンプレミスデータゲートウェイを使用して負荷を分散したり、耐障害性を高めたりできます。

Microsoft オンプレミスデータゲートウェイを使用することで、AWS のお客様から報告された多数の大きなメリットがもたらされます。

1. セキュリティ体制の向上

Microsoft オンプレミスデータゲートウェイは Microsoft Azure クラウドからのインバウンド接続を受け付けず、Azure Service Bus へのアウトバウンド接続のみを開始します。この一方向のトラフィックモデルにより、データソースをプライベートに保ち、パブリックインターネットに公開しないようにできます。

2. データ送信量を削減

データソースに接続すると、Microsoft オンプレミスデータゲートウェイは結果セット全体を取得し、スプールと呼ばれるプロセスでローカルに保存します。結果が Power BI サービスに送信される前に、データは圧縮されます。ユーザーからの報告では通常 10:1 の圧縮率です。これにより、インターネット経由でデータを送信する時間が短縮されるだけでなく、下り料金も削減されます。

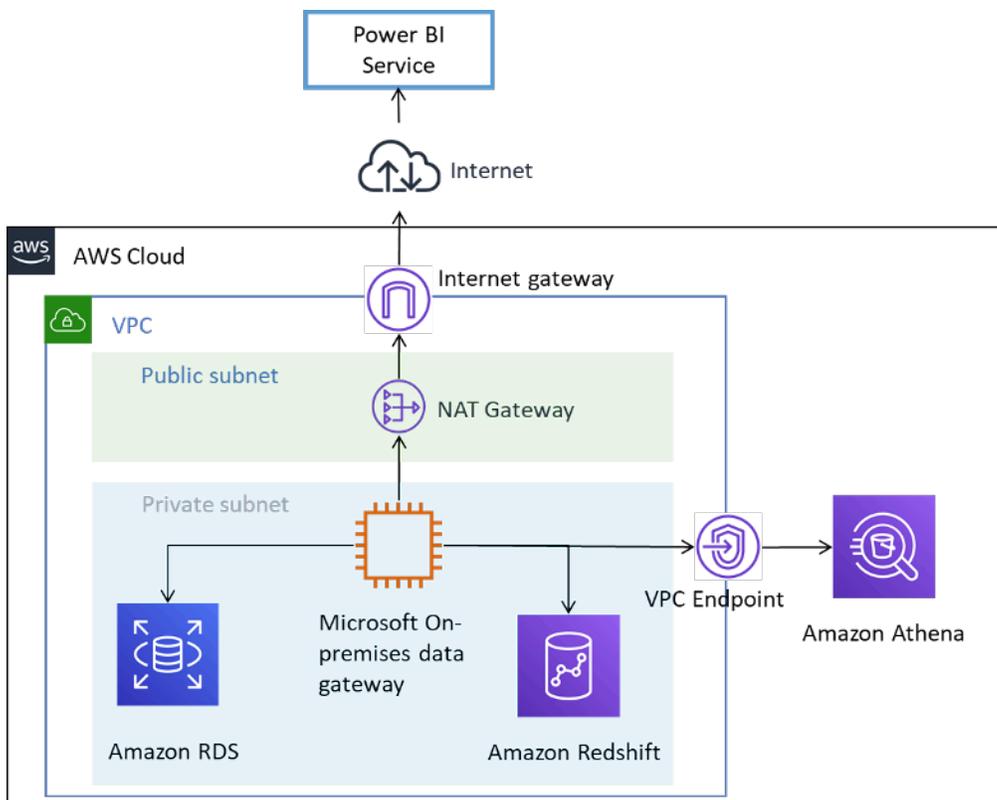
3. ソリューションコストの削減

Microsoft オンプレミスデータゲートウェイを使用すると、サービスに必要なデータ処理の一部が代わりにゲートウェイによって実行されます。Amazon EC2 を Savings Plans やリザーブドインスタンスなどのコスト削減計画と組み合わせて使用すると、BI ソリューション全体のコストを削減できる場合があります。

推奨設定

AWS では、データソースを含むプライベートサブネット内の Amazon EC2 インスタンスに Microsoft オンプレミスデータゲートウェイをインストールすることをお勧めしています。このサブネットは、パブリックサブネットにインストールされた [Amazon VPC NAT ゲートウェイ](#) を経由してリクエストをインターネットにルーティングするように設定されています。ネットワークアドレス変換 (NAT) ゲートウェイを使用して、プライベートサブネットのインスタンスからインターネットや他の AWS のサービスには接続できるが、インターネットからはこれらのインスタンスに接続できないようにできます。可用性の高いデータゲートウェイの実装が必要な場合は、異なる AWS アベイラビリティーゾーンにまたがる複数の EC2 インスタンスにインストールされたオンプレミスデータゲートウェイのクラスターを使用することをお勧めします。詳細については、[もう 1 つのゲートウェイを追加してクラスターを作成する](#) を参照してください。

このセクションで示すオプションは、Amazon RDS、Amazon Redshift、および Amazon Athena についてです。すべての AWS データソースの詳細については、「[付録: Microsoft Power BI 対応 AWS データソース](#)」を参照してください。



AWS データソースを Microsoft Power BI サービスに接続する

追加の考慮事項

表 5 - AWS クラウド内のデータソースを使用する Microsoft Power BI サービスに関する考慮事項

条件	AWS クラウド内のデータソースを使用する Microsoft Power BI サービスに関する考慮事項
ネットワーク接続	<p>データコンシューマーとデータソースの両方が AWS クラウド内に存在するため、データソースへの Microsoft オンプレミスデータゲートウェイの接続は簡単です。Amazon RDS や Amazon Redshift などの Amazon VPC に存在するデータソースには、直接アクセスできません。リージョンエンドポイントを使用するデータソースには、Amazon VPC インターネットゲートウェイ経由、または Amazon VPC エンドポイントを使用してアクセスできます。</p> <p>Microsoft Power BI サービスへの Microsoft オンプレミスデータゲートウェイ接続は、インターネット経由で行われ、アウトバウンド接続のみです。</p>
セキュリティ	<p>IP アクセスコントロール</p> <p>ルーティングとセキュリティグループを組み合わせ、AWS クラウド内に保存されているデータソースへのアクセスをコントロールできます。</p> <p>Microsoft オンプレミスデータゲートウェイは Amazon EC2 インスタンスにインストールされるため、オペレーティングシステムへのインバウンドアクセスを制限するために使用できるセキュリティグループが関連付けられています。ゲートウェイはインバウンドリクエストを受け付けません。インスタンスはパブリック IP ア</p>

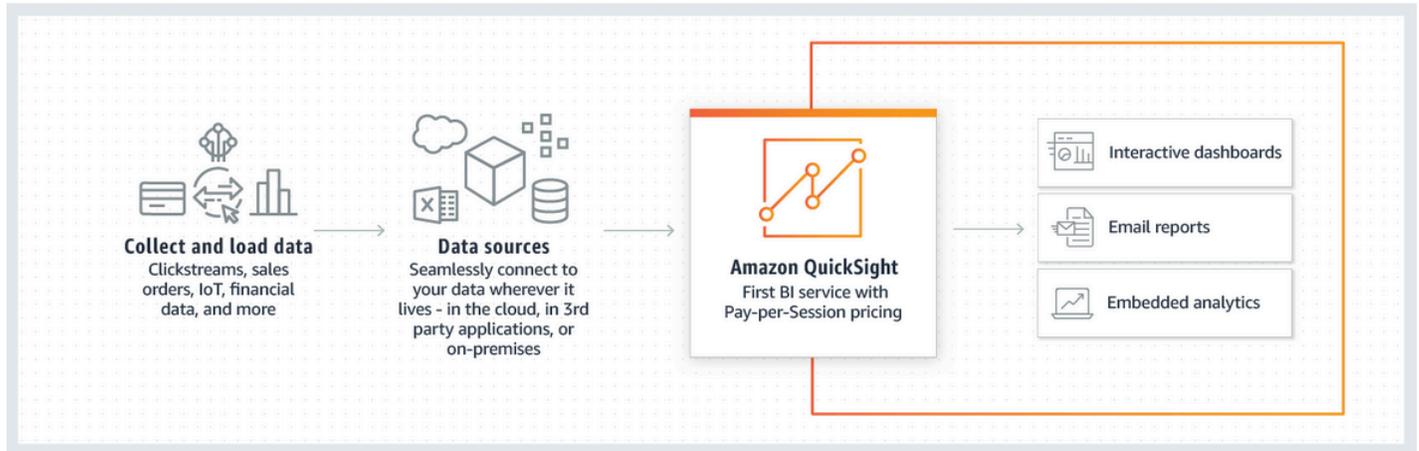
条件	AWS クラウド内のデータソースを使用する Microsoft Power BI サービスに関する考慮事項
	<p>ドレスを必要としないため、パブリック IP アドレスを使用して設定しないでください。</p> <p>転送時の暗号化</p> <p>Amazon VPC 内のデータソースは、データの送信に暗号化を使用するよう設定することをお勧めします。リージョン別のサービスは既に TLS 暗号化を利用しています。</p> <p>Microsoft オンプレミスのデータゲートウェイ接続は、TCP ではなく HTTPS を使用して Microsoft Azure Service Bus に接続するように構成できます。通信には HTTPS モードを使用することをお勧めします。これは、2019 年 6 月のゲートウェイソフトウェアバージョンのリリース以降の新規ゲートウェイインストールのデフォルトでもあります。</p> <p>認証</p> <p>AWS では、必要なデータセットに対してのみ読み取り専用アクセス許可を持つアイデンティティを使用して AWS データソースを認証することを推奨しています。データソースに入力した認証情報は暗号化され、ゲートウェイクラウドサービスに保存されます。認証情報はオンプレミスのゲートウェイで復号化されます。(データソースに入力した認証情報は暗号化され、ゲートウェイクラウドサービスに保存されます。認証情報はオンプレミスのゲートウェイで復号化されます。)</p> <p>Microsoft Power BI の認証情報は必ず安全に管理してください。サービスにアクセスすると、AWS データソースへのアクセスが許可さ</p>

条件	AWS クラウド内のデータソースを使用する Microsoft Power BI サービスに関する考慮事項
	れるほか、そこに含まれる可能性のある機密情報へのアクセスが潜在的に許可されます。
パフォーマンス	AWS クラウドの Microsoft オンプレミスデータゲートウェイは、Amazon EC2 インスタンスのサイジングとスケールアップが可能のため、一般的に高いパフォーマンスを発揮します。また、リージョンネットワークとインターネットへの接続においても高速に動作します。

条件	AWS クラウド内のデータソースを使用する Microsoft Power BI サービスに関する考慮事項
コスト	<p>Amazon EC2 インスタンス料金、データ転送料金、Amazon NAT ゲートウェイ料金の 3 つの要素を考慮する必要があります。</p> <p>Microsoft の要件に従って Amazon EC2 インスタンスをサイジングします。コストを削減するために、Amazon EC2 リザーブドインスタンスまたは AWS Savings Plans を購入できます。</p> <p>Microsoft オンプレミスデータゲートウェイから Microsoft BI サービスに転送されるデータには、VPC 下り料金が発生します。お客様からは、データゲートウェイを使用して 10:1 の圧縮を行うとトラフィック量が減少すると報告されていますが、AWS では、クエリを制限し、関連するデータのみが転送されるようにフィルターを使用することをお勧めします。</p> <p>Microsoft オンプレミスデータゲートウェイが、異なるアベイラビリティーゾーンまたは異なる AWS リージョンのデータソースに接続する場合は、データ転送料金も適用されます。</p> <p>Microsoft オンプレミスデータゲートウェイがプライベートサブネットにあり、AWS NAT ゲートウェイを使用する場合は、時間単位の料金およびデータ処理料金が適用されます。詳細については、「Amazon VPC の料金」を参照してください。</p>

Amazon QuickSight を使用する

AWS で Microsoft Power BI スイートを使用することを検討しているお客様は、代替として [Amazon QuickSight](#) を評価することをお勧めします。このフルマネージド型クラウドサービスは AWS のデータソースにネイティブに接続するため、他の BI ソリューションと比較して複雑さとコストが軽減されます。



Amazon QuickSight のしくみ

他の BI ソリューションと比較すると、Amazon QuickSight には次の利点があります。

- Amazon QuickSight では、クライアントアプリケーションをダウンロードしてインストールする必要はありません。オーサリングやレポート作成を含むすべての機能に、任意のプラットフォーム (Windows、Mac、Linux など) から Web ブラウザでアクセスできます。
- Amazon QuickSight は、フルマネージド型のクラウドネイティブ SaaS アプリケーションとして提供され、ダッシュボードを簡単に構築して本番環境にデプロイできます。このサービスはサーバーレスであるため、ユーザーをサポートするために必要なノードとサーバーの数を計算する必要はありません。また、QuickSight は AWS が提供する高可用性機能を最大限に活用して回復力を高めています。
- QuickSight 内のポイント安堵クリックインターフェイスからユーザーを追加できるため、小規模または大規模な設定でも簡単に開始できます。外部管理者の介入は不要です。
- Amazon QuickSight は、超高速の並列インメモリ計算エンジン (SPICE、Parallel, In-memory Calculation Engine) を搭載しており、応答時間が速く (ミリ秒単位)、インタラクティブな可視化が可能です。現在、データセットは最大 200 GB までスケールアップできます。
- Amazon QuickSight の料金はシンプルで安価であり、レポート作成者とレポートリーダーの 2 つの要素で構成されています。対話型ダッシュボードを作成および公開するレポート作成者は、ユー

ザーごとに料金が設定されます。特定の月にユーザーがログインしなかった場合、そのユーザーに対して料金は発生しません。レポートの読者には 30 分のセッションごとに課金され、1 人あたり 1 か月あたり最大 5.00 USD が課金されます。無料トライアルでは、Amazon QuickSight を無料で評価できます。詳細については、「[Amazon QuickSight の料金](#)」を参照してください。

まとめ

Microsoft Power BI Desktop の使用を検討している場合、お客様はオンプレミスで、インターネット経由でデータソースに接続してソフトウェアの実験を開始するのが一般的です。AWS VPN と Direct Connect を使用するプライベート接続オプションはありますが、Amazon WorkSpaces で Microsoft Power BI Desktop を実行するとパフォーマンスが向上すると多くのお客様が結論付けています。

AWS のデータソースを Microsoft Power BI Service に接続する場合、これが確立されたアーキテクチャパターンであることを理解すれば、安心できるはずですが、Amazon VPC 内に Microsoft オンプレミスデータゲートウェイをインストールし、Amazon RDS、Amazon Redshift、Amazon Athena、Amazon OpenSearch、AWS Lake Formation などのデータソースをサービスにシームレスに接続できます。

AWS で Microsoft Power BI Report Server を使用する予定がある場合は、今後の道筋も確立されています。Microsoft Power BI Report Server は、AWS データソースの近くにある Amazon VPC 内にインストールできます。作成者とレポートユーザーの両方向けに接続オプションが用意されています。

セルフマネージド BI ソリューションのインストール、設定、パッチ適用、スケーリングといった複雑さを増すことなく、同じビジネス成果をもたらすソリューションが必要な場合は、Amazon QuickSight をお勧めします。このフルマネージドサービスは、シンプルなウェブブラウザエクスペリエンスに収められた必要なすべての機能と、ユーザーごとの料金体系を組み合わせたものです。インストールも、追加のコンポーネントも必要ありません。

これが AWS でのビジネスインテリジェンスの旅の始まりに過ぎないことを願っています。使用開始に役立つその他のリソースについては、[付録: Microsoft Power BI 対応 AWS データソース](#) をご覧ください。

寄稿者

本書の作成における寄稿者

- アマゾン ウェブ サービス、シニアソリューションアーキテクト、Ralph Holm
- アマゾン ウェブ サービス、シニアソリューションアーキテクト、Barret Newman
- アマゾン ウェブ サービス、データベースと分析向けスペシャリストソリューションアーキテクト、Fabrizio Napolitano

その他の資料

- [Integrate Power BI with Amazon Redshift for insights and analytics](#)
- [Power BI デスクトップで Amazon Redshift データベースに接続する](#)

改訂履歴

このホワイトペーパーの更新に関する通知を受け取るには、RSS フィードをサブスクライブしてください。

update-history-change	update-history-description	update-history-date
更新	データソース追加: Amazon OpenSearch および AWS Lake Formation Athena、RD S、Redshift データソースの更新。Microsoft オンプレミス データゲートウェイに関する追加のガイダンス。	2021 年 11 月 3 日
レビュー済み	技術的な正確性について確認	2021 年 3 月 1 日
更新	レポートサーバーライセンスの更新	2021 年 1 月 15 日
初版発行	ホワイトペーパーの初回発行	2020 年 11 月 1 日

付録: Microsoft Power BI 対応 AWS データソース

サポートされているデータソースの完全なリストは Microsoft から提供されています ([Power BI データソース](#)を参照)。ただし、各 AWS データソースの以下のセクションに、一部の読者に役立つ使用法と設定のガイダンスが記載されています。

Amazon Redshift

Amazon Redshift は、AWS クラウド内でのフルマネージド型、ペタバイトスケールのデータウェアハウスサービスです。Amazon Redshift データウェアハウスは、ノードと呼ばれるコンピューティングリソースのコレクションであり、これらはクラスターと呼ばれるグループを構成します。各クラスターは Amazon Redshift エンジンを実行し、1 つ以上のデータベースを含みます。

以下のような場合は、Amazon Redshift の使用をご検討ください。

- ネイティブのクラウドデータウェアハウスを構築または移行している。
- 数テラバイトから数百テラバイトまでスケールする必要があるかもしれない。
- Power BI ユーザーが Simple Storage Service (Amazon S3) に格納されているデータレイクからデータに透過的にアクセスし、データウェアハウス内のテーブルと結合できるようにしたい。
- クエリワークロードに次のものが含まれる。
 - 大きな (数ギガバイトや数テラバイト) テーブルで集計を計算するクエリ。
 - 複数の結合とサブクエリを持つ、非常に複雑な SQL。
 - 複雑な分析クエリと、ダッシュボードで使用される単純で高度にフィルターされたクエリが混在している。

Amazon Redshift を Microsoft Power BI とともに使用する場合は、次の点に留意してください。

- Amazon Redshift は、Microsoft Power BI Desktop と Power BI サービスの両方で、Power BI データソースとしてネイティブにサポートされており、それぞれがインポートモードとダイレクトクエリモードをサポートしています。
- Redshift クラスターは、パブリックサブネットで起動しインターネットからのアクセスを許可するように設定できますが、大半のお客様は、セキュリティ強化のためにプライベートサブネットで起動することを好みます。プライベートサブネットを使用する場合は、オンプレミスのデータゲートウェイを使用して Power BI サービスから Amazon Redshift に接続します。

- Redshift コネクタは Power BI Desktop およびサービスで Azure AD 認証をサポートします。
- Spectrum 経由でアクセスされる外部テーブルは、ネイティブの Redshift テーブルとまったく同じように扱われ、Power BI にはそれらを区別する手段がありません。外部テーブルのデータにアクセスするときは、次の点を確認してください。
 - 文字列を含む列が、AWS Glue データカタログで「STRING」ではなく「VARCHAR」としてカタログ化されていること。それ以外の場合、Power BI は次のエラーをスローします。
Exception: OLE DB or ODBC error: [Expression.Error] We couldn't fold the expression to the data source. Please try a simpler expression..
 - ARRAY などの複雑なデータ型を含む列はサポートされていません。複雑なデータ型を含む列を使用すると、Power BI は次のエラーをスローします。Exception: ODBC: ERROR [42703] [Microsoft]Amazon Redshift Error occurred while trying to execute a query

モデルに含める必要がある場合は、JSON シリアル化をユーザーレベルで (Amazon Redshift で) 有効にするか、複雑なデータ型をネイティブテーブルの SUPER 列に格納します。

Amazon RDS

Amazon RDS を利用すると、リレーショナルデータベースをクラウドに簡単に設定、操作、スケーリングできます。Amazon RDS は、メモリ、パフォーマンス、または I/O に最適化されたいくつかのデータベースインスタンスタイプで利用でき、Amazon Aurora、PostgreSQL、MySQL、MariaDB、Oracle データベース、SQL Server など、6 つの使い慣れたデータベースエンジンから選択できます。

以下のような場合は、RDS の使用をご検討ください。

- オペレーショナルデータストアを構築している。
- SQL Server または Oracle Database データウェアハウスをクラウドに移行しているが、リファクタリングは不要である。
- クエリワークロードに次のものが含まれる。
 - 簡単にインデックスを作成できるテーブル上の高度にフィルターされたデータにアクセスするクエリ。
 - 小中規模のテーブル (ギガバイト) に対する分析クエリ。
 - やや複雑な分析クエリと、ダッシュボードで使用される単純で高度にフィルターされたクエリが混在している。

Amazon RDS を Microsoft Power BI とともに使用する場合は、次の点に留意してください。

- Amazon RDS には、SQL Server、MariaDB、MySQL、Oracle Database、PostgreSQL など、複数のデータベースエンジンが用意されています。データベースエンジンは、Amazon RDS サービスではなく、Power BI Desktop および Power BI サービスにリストされていることに注意してください。
- Amazon Aurora の場合は、選択したデータベースエンジンに応じて My SQL または PostgreSQL 接続タイプを使用します。
- Amazon RDS インスタンスは、パブリックサブネットで起動しインターネットからのアクセスを許可するように設定できますが、大半のお客様は、セキュリティ強化のためにプライベートサブネットを起動することを好みます。プライベートサブネットを使用する場合は、オンプレミスのデータゲートウェイを使用して Power BI サービスから RDS に接続します。
- Amazon RDS では、Express、Web、Standard、Enterprise など複数のエディションの SQL Server (2012、2014、2016、2017、および 2019) をデプロイできます。

Amazon Athena

Amazon Athena はインタラクティブなクエリサービスで、Simple Storage Service (Amazon S3) 内のデータを標準 SQL を使用して簡単に分析できるようになります。Athena は初期状態で AWS Glue データカタログと統合されており、さまざまなサービスにわたるメタデータの統合リポジトリを作成できます。データソースのクローリングとスキーマの解析、新規および修正したテーブル定義とパーティション定義のデータカタログへの入力、スキーマのバージョニング保持が可能です。

以下のような場合は、Athena をデータソースとしてご検討ください。

- データレイクを直接クエリしたい。
- クエリワークロードに次のものが含まれる。
 - 大きな (数ギガバイトや数テラバイト) テーブルで集計を計算するクエリ。
 - 予備的な対話型アドホック SQL。

Amazon Athena を Microsoft Power BI とともに使用する場合は、次の点に留意してください。

- 2021 年 7 月の Microsoft Power BI のリリースに伴い、Amazon Athena 用に Microsoft 認定コネクタが導入されました。Amazon Athena 用の Microsoft Power BI コネクタを使用して、Microsoft Power BI Desktop で Amazon Athena からのデータを分析できます。Power BI サービスにコンテ

ンツを公開した後は、Microsoft オンプレミスデータゲートウェイを使用して、オンデマンドまたはスケジュールされた更新によってコンテンツを最新の状態に保つことができます。

- Amazon Athena 用 Microsoft Power BI コネクタは、インポートとダイレクトクエリ両方のデータ接続モードをサポートしています。インポートモードでは、選択したテーブルと列が Power BI Desktop にインポートされ、クエリが実行されます。ダイレクトクエリモードでは、Power BI Desktop にデータがインポートまたはコピーされることはなく、Power BI Desktop は基になるデータソースを直接クエリします。
- Amazon Athena 用の Microsoft Power BI コネクタの詳細については、「[Amazon Athena Power BI コネクタの使用](#)」を参照してください。
- Amazon Athena 用 Microsoft Power BI コネクタを使用するには、Amazon Athena ODBC ドライバーと、システム上の有効な ODBC DSN 設定を使用して Amazon Athena にクエリを実行する必要があります。最新の ODBC ドライバーのダウンロード、および設定情報については、「[ODBC を使用した Amazon Athena への接続](#)」を参照してください。
- Amazon Athena 用 Microsoft Power BI コネクタを使用する場合の設定手順のチュートリアルとベストプラクティスについては、「[Creating dashboards quickly on Microsoft Power BI using Amazon Athena](#)」を参照してください。

Amazon OpenSearch Service (Amazon Elasticsearch Service の後継サービス)

JSON ベースの検索クエリ DSL を使用する代わりに、SQL を使用して Amazon OpenSearch Service をクエリできます。SQL を使用したクエリは、既に言語に慣れている場合や、Microsoft Power BI などの SQL を使用するアプリケーションとドメインを統合する場合に便利です。

以下のような場合は、Amazon OpenSearch Service をデータソースとしてご検討ください。

- ログファイルや JSON 出力などの半構造化データがあり、その情報をすばやく検索、分析、または可視化する必要がある。

Amazon OpenSearch サービスを Microsoft Power BI とともに使用する場合は、次の点に留意してください。

- Amazon OpenSearch Service への接続には、Open Database Connectivity (ODBC) ドライバーが必要です。ODBC ドライバーは Windows および macOS 用の読み取り専用 ODBC ドライバーで、ビジネスインテリジェンス (BI) と [Tableau](#) や [Microsoft Excel](#) などのデータ可視化アプリケーションを接続できます。また、[Power BI](#) をクラスター上の SQL プラグインに接続します。

ドライバーは、[Getting Started & Downloads](#) ウェブサイトで入手できます。構成手順については、[Configuring a DSN on Windows](#) を参照してください。

- 現在は、インポートモードのみがサポートされています。
- Amazon OpenSearch Service への Power BI Desktop 接続には現在、カスタムコネクタを使用する必要があります。開始するには、[Connecting Open Distro For ElasticSearch to Microsoft Power BI Desktop](#) を参照してください。
- Microsoft オンプレミスデータゲートウェイを介した Amazon OpenSearch Service への Power BI サービスの接続の場合は、開始するには [Connecting Open Distro For ElasticSearch to Microsoft Power BI Service](#) を参照してください。

AWS Lake Formation

Lake Formation は、データベースとオブジェクトストレージからデータを収集およびカタログ化し、データを新しい [Simple Storage Service \(Amazon S3\)](#) データレイクに移動し、機械学習アルゴリズムを使用してデータをクリーンアップおよび分類し、機密データへのアクセスを保護します。これらのタスクが完了すると、ユーザーは、一元化されたデータカタログにアクセスできるようになります。この[データカタログ](#)は、利用可能なデータセットおよびその適切な使用方法を示すものです。ユーザーはその後、Apache Spark 向け [Amazon Redshift](#)、[Amazon Athena](#)、および (ベータ版) [Amazon EMR](#) などの分析や機械学習サービスでデータセットを利用します。Lake Formation は、[AWS Glue](#) で利用できる機能を基盤として構築されています。

データレイクに対して、従来の IAM ベースのコントロールではなく、きめ細かい (行と列) レベルのアクセスが必要な場合は、Lake Formation をご検討ください。

Microsoft Power BI で Lake Formation を使用する場合は、次の点に留意してください。

- Power BI Desktop または Power BI サービスを使用して Lake Formation データカタログからデータをクエリするには、Athena のデータのクエリと同じプロセスと設定を使用します。Lake Formation アクセス許可モデルを使用する場合は、Amazon Athena の ODBC DSN 設定で「LakeformationEnabled」プロパティキーの値が「true」に設定されていることを確認してください。これにより、Amazon Athena ODBC ドライバーが認可に AWS Security Token Service を直接使用するのではなく、Lake Formation サービスを使用するよう指示します。詳細については、「[ODBC を使用した Amazon Athena への接続](#)」のドキュメントを参照してください。
- 既存のデータカタログの動作との互換性のために「IAM アクセス制御のみを使用」設定を有効にすると、完全な互換性が得られます。

- AWS Glue データアクセス許可を Lake Formation モデルにアップグレードすると、非互換性が生じる可能性があるため、使用前にテストする必要があります。予備テストでは、列レベルの許可または拒否が受け入れられていることが示されていますが、行およびセルレベルのフィルタリングはまだプレビュー中であり、変更される可能性があるため、作成者によるテストは行われていません。

注意

お客様は、本書の情報について独自の評価を行う責任を負うものとします。本書は、(a) 情報提供のみを目的とし、(b) AWS の現行製品と慣行について説明しており、これらは予告なしに変更されることがあり、(c) AWS およびその関連会社、サプライヤーまたはライセンサーからの契約上の義務や保証をもたらすものではありません。AWS の製品やサービスは、明示または暗示を問わず、一切の保証、表明、条件なしに「現状のまま」提供されます。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 とお客様の間で締結されるいかなる契約の一部でもなく、その内容を修正するものでもありません。

© 2021 Amazon Web Services, Inc. or its affiliates. All rights reserved.