



AWS 백서

# AWS 클라우드에서 Microsoft Power BI 사용



# AWS 클라우드에서 Microsoft Power BI 사용: AWS 백서

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

Amazon의 상표 및 브랜드 디자인은 Amazon 외 제품 또는 서비스와 함께, Amazon 브랜드 이미지를 떨어뜨리거나 고객에게 혼동을 일으킬 수 있는 방식으로 사용할 수 없습니다. Amazon이 소유하지 않은 기타 모든 상표는 Amazon과 제휴 관계이거나 관련이 있거나 후원 관계 여부에 관계없이 해당 소유자의 자산입니다.

# Table of Contents

|  |    |
|--|----|
| AWS 클라우드에서 Microsoft Power BI 사용 .....                           | 1  |
| 요약 .....   | 1  |
| 소개 .....   | 2  |
| Microsoft Power BI 제품군 .....                                     | 4  |
| Power BI Desktop .....   | 4  |
| Power BI 서비스 .....   | 4  |
| Power BI Report Server .....                                     | 5  |
| 온프레미스 데이터 게이트웨이 .....  | 5  |
| Microsoft Power BI Desktop을 AWS 데이터 원본에 연결 .....                 | 6  |
| 온프레미스에서 Power BI Desktop 사용 .....                                | 6  |
| 인터넷을 통해 데이터 원본에 연결 .....   | 6  |
| AWS VPN을 통해 데이터 원본에 연결 .....                                     | 9  |
| AWS Direct Connect를 통해 데이터 원본에 연결 .....                          | 13 |
| AWS 클라우드에서 Microsoft Power BI Desktop 사용 .....                   | 16 |
| 옵션 1: Amazon EC2 인스턴스에 Microsoft Power BI Desktop 설치 .....       | 17 |
| 옵션 2: Amazon WorkSpaces 환경에 Microsoft Power BI 설치 .....          | 17 |
| 옵션 3: Amazon AppStream 2.0 환경에 Microsoft Power BI 설치 .....       | 18 |
| Microsoft Power BI Desktop 연결 옵션 요약 .....                        | 20 |
| Microsoft Power BI 서비스를 AWS 데이터 원본에 연결 .....                     | 22 |
| 권장 구성 .....  | 22 |
| 추가 고려 사항 .....   | 23 |
| Amazon QuickSight 사용 .....                                       | 28 |
| 결론 .....   | 30 |
| 기여자 .....  | 31 |
| 추가 자료 .....  | 32 |
| 문서 개정 .....  | 33 |
| 부록: Microsoft Power BI에서 지원하는 AWS 데이터 원본 .....                   | 34 |
| Amazon Redshift .....  | 34 |
| Amazon RDS .....   | 35 |
| Amazon Athena .....  | 36 |
| Amazon OpenSearch Service(Amazon Elasticsearch Service 후속) ..... | 37 |
| AWS Lake Formation .....   | 38 |
| 고지 사항 .....  | 39 |

# AWS 클라우드에서 Microsoft Power BI 사용

AWS 클라우드에서 Microsoft Power BI 사용

게시 날짜: 2021년 11월 3일([문서 개정](#))

## 요약

이 백서에서는 Microsoft Power BI(Desktop, Report Server, 서비스, 온프레미스 데이터 게이트웨이)를 Amazon Web Services(AWS) 클라우드와 통합하고 사용하는 방법에 대해 설명합니다. 연결, 보안, 성능, 비용 최적화에 중점을 두고 Microsoft Power BI 제품을 Amazon Redshift, Amazon Athena, Amazon RDS, Amazon OpenSearch, AWS Lake Formation 같은 AWS 서비스에 연결하려는 고객을 위한 옵션을 제공합니다.

이 백서는 Microsoft Power BI 개념과 AWS 서비스를 데이터 원본으로 사용할 때 사용할 수 있는 기술 옵션을 빠르게 이해하기를 원하는 IT 의사 결정권자 및 설계자를 대상으로 합니다.

# 소개

모든 규모의 비즈니스를 운영하는 고객은 AWS 제품 및 서비스를 사용하여 데이터를 안정적이고 비용 효율적이며 안전하게 저장하고 있습니다. 부분적으로 이는 성숙한 데이터 스토리지 및 분석 상품을 사용할 수 있는 광범위한 에코시스템 때문입니다. 이러한 상품 중 일부에는 다음과 같은 서비스가 포함됩니다.

- [Amazon Simple Storage Service](#)(Amazon S3)는 간단하고 확장 가능하며 안전하고 비용 효율적인 데이터 리포지토리를 제공합니다. 이는 애플리케이션 데이터 저장을 위한 업계 표준이 되었으며 고객 데이터 레이크를 위한 첫 번째 선택이 되었습니다.
- [Amazon Athena](#)는 표준 SQL을 사용해 Amazon S3에 저장된 데이터를 간편하게 분석할 수 있는 대화식 쿼리 서비스입니다.
- [Amazon Relational Database Service](#)(Amazon RDS)를 사용하면 클라우드에서 관계형 데이터베이스를 더욱 간편하게 설정, 관리 및 확장할 수 있습니다. 하드웨어 프로비저닝, 데이터베이스 설정, 패치 적용, 백업 등 시간이 많이 소요되는 관리 태스크를 자동화하는 동시에 비용 효율적이고 크기 조정이 가능한 용량을 제공합니다. SQL Server, Oracle Database, MySQL, MariaDB 및 PostgreSQL 엔진을 사용할 수 있습니다.
- [Amazon Redshift](#)는 대규모로 확장이 가능한 완전관리형 데이터 웨어하우스로, 정형 및 비정형 데이터 집합 모두 간편하게 분석할 수 있습니다.
- [Amazon QuickSight](#)는 조직 내 모든 구성원에게 통찰력을 손쉽게 제공할 수 있게 지원하는 빠른 클라우드 기반 비즈니스 인텔리전스 서비스입니다.
- [Amazon OpenSearch](#)(Amazon Elasticsearch Service 후속)는 보다 간편하게 대규모 Elasticsearch를 비용 효과적으로 배포, 보호, 실행할 수 있는 완전관리형 서비스입니다.
- [AWS Lake Formation](#)은 안전한 데이터 레이크를 며칠 만에 손쉽게 설정할 수 있도록 지원하는 서비스입니다.

서비스가 어떻게 서로 관련되는지 더 잘 이해하기 위해 데이터 서비스를 데이터 원본 또는 데이터 소비자로 분류하는 경우가 많습니다. 데이터 원본을 사용하여 고객 및 애플리케이션이 서비스에서 데이터를 저장하고 검색할 수 있습니다. 데이터 원본은 컴퓨팅 기능이 내장되어 계산 분석 및 필터링을 제공할 수 있는 경우가 많습니다. 그러나 궁극적으로 데이터는 이러한 데이터 원본에 로드되고 결국 데이터 소비자가 해당 데이터 원본에서 데이터를 검색합니다. Amazon S3, Amazon Athena, Amazon Redshift는 데이터 원본의 좋은 예입니다.

반면 데이터 소비자는 데이터 원본의 데이터에 액세스하고 일반적으로 데이터를 처리합니다. 선택적으로 데이터를 표시할 수도 있습니다. Amazon QuickSight와 Microsoft Power BI 제품군은 데이터 소비자의 좋은 예입니다. 데이터 원본에서 데이터를 읽은 다음 정보의 분석, 시각화 및 게시를 지원합니다.

AWS에서는 고객이 데이터 요구 사항에 따라 선호하는 기술을 혼합할 수 있는 완벽한 유연성을 제공합니다. 많은 고객이 비즈니스 인텔리전스(BI) 요구 사항에 따라 Amazon QuickSight를 선택하지만 다른 고객은 Microsoft Power BI, Tableau, Qlik과 같은 공급 업체를 선택합니다.

이 문서에서는 Microsoft Power BI 제품군 및 서비스와 이를 AWS 서비스와 함께 사용하는 방법을 중점적으로 다룹니다.

# Microsoft Power BI 제품군

제품명 유사성으로 인한 혼동을 줄이기 위해 이 백서에서는 각 Microsoft Power BI 제품 및 서비스에 대해 설명합니다.

## Power BI Desktop

Power BI Desktop은 로컬 컴퓨터에 설치하는 무료 애플리케이션입니다. 이를 통해 데이터를 연결, 변환 및 시각화할 수 있습니다. Power BI Desktop을 사용하면 여러 데이터 원본에 연결하고 이를 데이터 모델로 결합(흔히 모델링이라고도 함)할 수 있습니다. 이 데이터 모델을 사용하면 조직 내 다른 사용자와 보고서로 공유할 수 있는 시각적 개체 및 시각적 개체 모음을 만들 수 있습니다.

Power BI Desktop은 로컬로 또는 네트워크를 통해 사용할 수 있는 모든 지원되는 데이터 원본에 연결할 수 있습니다. 지원되는 데이터 원본은 [부록: Microsoft Power BI에서 지원하는 AWS 데이터 원본](#)을 참조하십시오.

비즈니스 인텔리전스 프로젝트에서 작업하는 대부분의 사용자는 Power BI Desktop을 사용하여 보고서를 작성합니다. 그런 다음 다른 사용자와 보고서를 공유하기 위해 콘텐츠를 Power BI Report Server 또는 Power BI 서비스로 푸시합니다. Power BI Desktop에서 Power BI Report Server 또는 Power BI 서비스로 콘텐츠를 푸시하는 행위를 게시라고 합니다. 자세한 내용은 [Power BI Desktop이란 무엇입니까?](#)를 참조하십시오.

### Note

Power BI Desktop은 Windows 전용 애플리케이션입니다. Linux, macOS 등 다른 운영 체제에서는 사용할 수 없습니다.

## Power BI 서비스

Power BI는 함께 작동하는 소프트웨어 서비스, 앱 및 커넥터의 모음으로, 사용자와 사용자의 비즈니스를 가장 효과적으로 지원하는 방식으로 비즈니스 인사이트를 생성, 공유, 사용할 수 있도록 지원합니다. Power BI 온라인이라고도 하는 Power BI 서비스는 Power BI의 서비스형 소프트웨어(SaaS) 부분입니다. 자세한 내용은 [Power BI 서비스란 무엇인가요?](#)를 참조하십시오.

Power BI 서비스는 클라우드 기반 서비스입니다. 팀과 조직을 위한 간단한 보고서 편집 및 공동 작업을 지원합니다. Power BI 서비스에서도 데이터 원본에 연결할 수 있지만 모델링은 제한적입니다.

비즈니스 인텔리전스 프로젝트에서 작업하는 대부분의 보고서 디자이너는 Power BI Desktop을 사용하여 보고서를 작성한 다음 Power BI 서비스를 사용하여 다른 사용자에게 보고서를 배포합니다. 이 중요한 구성 요소에 대한 자세한 내용은 [Microsoft Power BI 서비스를 AWS 데이터 원본에 연결](#)을 참조하십시오.

## Power BI Report Server

Power BI Report Server는 보고서 및 KPI를 표시하고 관리하는 웹 포털이 있는 비공개 보고서 서버입니다. 고객은 보고서를 Power BI 서비스에 게시하지 않으려는 경우 Power BI Report Server를 사용합니다. Power BI Report Server는 원래 온프레미스 환경을 위한 것이었지만 AWS에서도 실행할 수 있습니다. 자세한 내용은 [AWS에서 Microsoft Power BI Report Server 사용](#)을 참조하십시오.

## 온프레미스 데이터 게이트웨이

Microsoft 온프레미스 데이터 게이트웨이는 Power BI 배포의 보안 및 성능을 향상시킬 수 있는 일반적으로 배포되는 구성 요소입니다. 이를 통해 Power BI 서비스는 다른 시설에 상주하고 데이터 원본과 데이터 게이트웨이 간의 내부 네트워크 연결을 통해 액세스할 수 있는 비공개 데이터 원본에 액세스할 수 있습니다. 일반적으로 서버 구성 요소로 설치되지만 로컬 컴퓨터에 애플리케이션으로 개인 모드를 설치할 수도 있습니다. 이 백서는 표준(서버) 모드에만 초점을 맞추고 있습니다. 자세한 내용은 [Microsoft Power BI 서비스를 AWS 데이터 원본에 연결](#)을 참조하십시오.



# Microsoft Power BI Desktop을 AWS 데이터 원본에 연결

Microsoft Power BI Desktop을 처음 사용하는 고객은 대부분 온프레미스 컴퓨터 및 네트워크에서 AWS 데이터 원본에 연결하는 방법에 관심이 있습니다. 데스크톱 애플리케이션은 일반적으로 로컬 Windows 랩톱에서 실행되며 AWS 데이터 원본에 대한 물리적 및 논리적 연결이 가장 큰 진입 장벽으로 인식됩니다.

하지만 AWS 클라우드에서 Microsoft Power BI Desktop을 실행하는 또 다른 옵션이 있습니다. 이 옵션은 AWS 데이터 원본에 대한 연결 장벽을 크게 줄이지만 추가로 몇 가지 사항을 고려해야 합니다. 이 섹션에서는 두 모델 모두에 대해 설명합니다. 연결성, 보안, 성능 및 비용과 관련하여 각각 미치는 영향을 조사하여 가장 적합한 옵션을 결정할 수 있습니다. 이 섹션에 제시된 옵션은 Amazon RDS, Amazon Redshift 및 Amazon Athena를 보여줍니다. 모든 AWS 데이터 원본에 대한 자세한 내용은 [부록: Microsoft Power BI에서 지원하는 AWS 데이터 원본](#)을 참조하십시오.

## 온프레미스에서 Power BI Desktop 사용

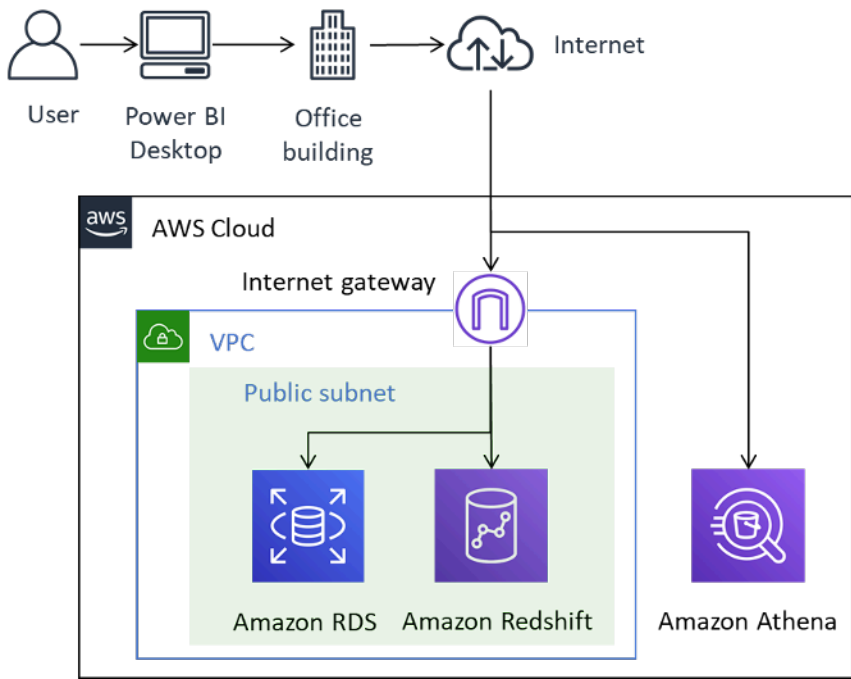
온프레미스에서 Power BI Desktop을 AWS 클라우드에 저장된 데이터 원본과 함께 사용하려는 경우 Power BI는 다음 세 가지 방법 중 하나로 이러한 원본에 액세스할 수 있습니다.

- 인터넷을 사용하여 데이터 원본에 연결
- AWS 가상 사설 네트워크(AWS AWS VPN)를 사용하여 데이터 원본에 연결
- AWS Direct Connect를 사용하여 데이터 원본에 연결

각 방법은 다음 섹션에 자세히 나와 있습니다.

## 인터넷을 통해 데이터 원본에 연결

이 모델에서 Power BI Desktop 애플리케이션은 인터넷을 통해 인터넷 액세스가 가능한 AWS 데이터 원본의 IP 주소로 라우팅되는 아웃바운드 연결을 배치합니다. 예를 들어, 고객의 Amazon Virtual Private Cloud(Amazon Amazon VPC) 내에서 인스턴스화되는 Amazon RDS 및 Amazon Redshift는 퍼블릭 액세스 가능성 옵션을 지원하여 인터넷을 통해 인스턴스에 액세스할 수 있도록 합니다. Amazon Athena는 특정 리전의 서비스 엔드포인트를 사용하여 인터넷에서 직접 쿼리할 수 있습니다.



Power BI에서 인터넷을 통해 AWS 데이터 원본에 연결

이 연결 방법은 기술적으로 가능하지만 소수의 사용자 외에는 권장하지 않습니다. 다음 표에는 중요한 고려 사항이 나열되어 있습니다.

표 1 - 인터넷을 통해 AWS 데이터 원본에 액세스할 때의 고려 사항

| 기준      | 인터넷을 통해 AWS 데이터 원본에 액세스할 때의 고려 사항   |
|---------|---|
| 네트워크 연결 | VPC의 프라이빗 IP 주소에 연결하거나 리전 또는 VPC 서비스 엔드포인트를 사용하여 데이터 원본을 사용할 수 있습니다. Power BI Desktop은 VPN을 통해 연결하고 데이터 원본(Amazon RDS, Amazon Redshift, Amazon EC2 기반 데이터 원본)에 직접 액세스하거나, 리전 엔드포인트(Amazon Athena)가 있는 서비스의 경우 DNS 구성에 따라 프라이빗 VPC 엔드포인트 또는 리전 엔드포인트를 사용하여 액세스합니다. |
| 보안      | IP 액세스 제어   |

|           |  |
|-----------|--|
| <p>기준</p> | <p>인터넷을 통해 AWS 데이터 원본에 액세스할 때의 고려 사항</p>   |
|           | <p>보안 그룹은 인스턴스에 대한 인바운드 및 아웃바운드 트래픽을 제어하는 가상 방화벽 역할을 합니다. 신뢰할 수 있는 엔터티로 액세스를 제한하려면 알려진 CIDR(Classless Inter-Domain Routing) 범위와 연결된 인바운드 IP 범위만 허용하도록 보안 그룹을 구성합니다.</p> <p>전송 중 데이터 암호화</p> <p>퍼블릭 IP 주소를 사용하는 모든 데이터 원본(예: Amazon RDS, Amazon Redshift 또는 모든 Amazon EC2 기반 데이터 원본)에 대해 암호화를 구성하는 것이 좋습니다. 이렇게 하면 전송 중에 데이터 또는 자격 증명이 손상될 위험이 줄어듭니다. 암호화를 구성하지 않으면 심각한 위험이 따릅니다. 이 측면을 간과하지 마십시오.</p> <p>Amazon Athena와 같은 리전 서비스 엔드포인트는 TLS로 암호화됩니다. 또한 JDBC 또는 ODBC 클라이언트로 스트리밍되는 Amazon Athena 쿼리 결과는 전송 계층 보안(TLS)을 사용하여 암호화됩니다.</p> <p>인증 및 권한 부여</p> <p>데이터 집합에 대한 읽기 전용 액세스를 제공하는 자격 증명을 사용하고 회사 정책에 따라 자격 증명을 회전하는 프로세스를 설정하는 것이 좋습니다.</p> |

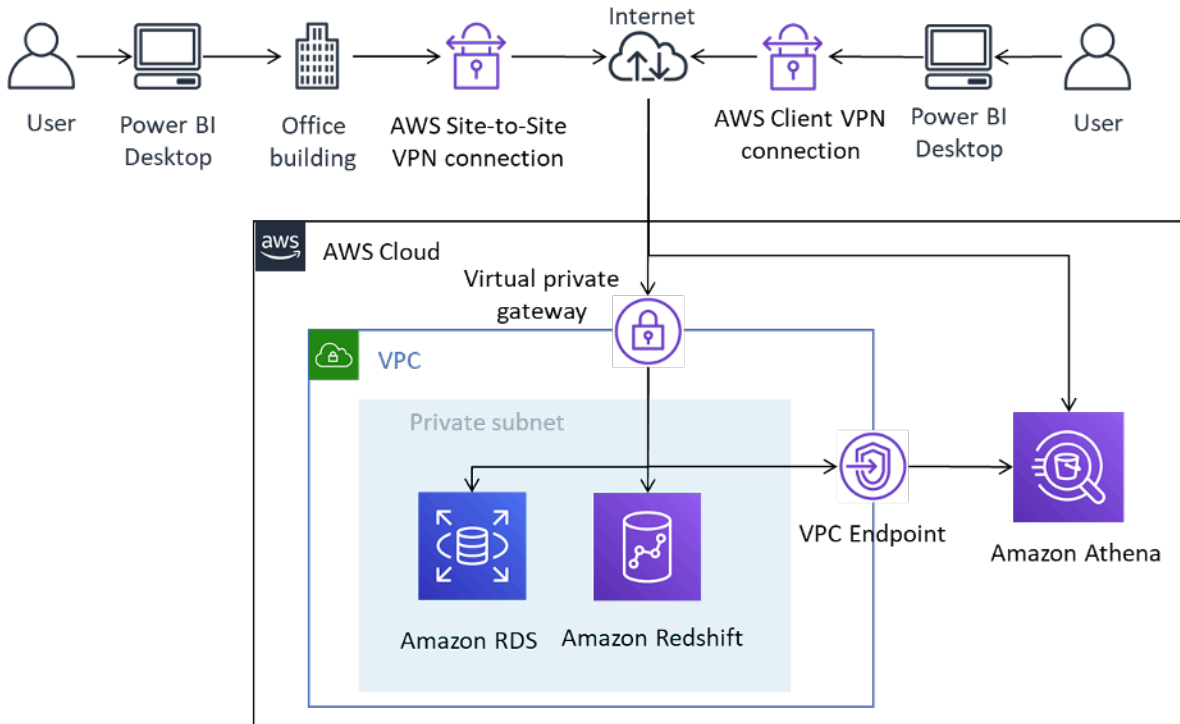
|    |  |
|----|--|
| 기준 | 인터넷을 통해 AWS 데이터 원본에 액세스할 때의 고려 사항  |
| 성능 | <p>인터넷을 통해 AWS 데이터 원본에 액세스할 때 전체 Power BI Desktop 성능에 영향을 줄 수 있는 몇 가지 요소는 다음과 같습니다.</p> <ul style="list-style-type: none"> <li>• 액세스되는 데이터 집합의 크기. 데이터 집합이 클수록 검색하는 데 시간이 더 오래 걸립니다. 쿼리를 제한하고 필터를 사용하여 인터넷을 통해 검색되는 데이터의 양을 줄이는 것이 좋습니다.</li> <li>• 대역폭, 대기 시간, 패킷 손실 등의 인터넷 연결 품질. 가능한 경우 대기 시간의 영향을 줄이기 위해 지리적으로 가까운 AWS 리전의 데이터에 액세스합니다. 인터넷을 공유하는 경우 사용량이 적은 시간에 데이터 원본을 로드하고 사용 가능한 대역폭을 충분히 확보하는 것이 좋습니다.</li> </ul> <p>일반적으로 서로 다른 일중 시간대에 다양한 데이터 집합을 사용하고 점차 사용자를 늘려가며 환경을 테스트하는 것이 좋습니다.</p> |
| 비용 | <p>VPC에 상주하고 퍼블릭 IP 주소를 사용하여 인터넷을 통해 쿼리되는 데이터 원본에는 표준 Amazon VPC 데이터 송신 요금이 부과됩니다. 비용을 절감하려면 쿼리를 제한하고 필터를 사용하여 인터넷을 통해 검색되는 데이터의 양을 줄이는 것이 좋습니다.</p>   |

## AWS VPN을 통해 데이터 원본에 연결

이 모델에서 Power BI Desktop 설치하는 두 가지 AWS VPN 방법인 AWS Site-to-Site VPN 또는 AWS Client VPN 중 하나를 사용하여 AWS 네트워크의 데이터 원본에 연결합니다. 각 연결 유형은 네트워크 트래픽을 보호하는 탄력적인 고가용성 관리형 클라우드 VPN 솔루션을 제공합니다.

Site-to-Site VPN은 네트워크와 AWS VPN 또는 AWS Transit Gateway 사이에 암호화된 터널을 생성합니다. Client VPN은 무료 VPN 소프트웨어 클라이언트를 사용하여 사용자를 AWS 또는 온프레미스 리소스에 연결합니다.

Site-to-Site VPN 및 Client VPN 연결을 통한 VPN 트래픽은 VPC에서 멈춥니다. 따라서 프라이빗 IP 주소로 라우팅할 수 있으므로 인스턴스에 더 이상 퍼블릭 IP 주소가 필요하지 않습니다. Athena와 같이 공개 서비스 엔드포인트에서 액세스할 수 있는 데이터 경로가 있는 서비스의 경우 이러한 서비스 요청은 인터넷을 통해 또는 VPN 연결 및 VPC 엔드포인트를 통해 라우팅될 수 있습니다.



Power BI Desktop을 Site-to-Site VPN 및 Client VPN을 통해 AWS 데이터 원본에 연결

Site-to-Site VPN은 또한 AWS Transit Gateway에 연결할 수 있습니다. 그러면 여러 VPC에 분산되어 있는 데이터 원본에 쉽게 액세스할 수 있습니다.

AWS VPN을 사용하면 각 데이터 원본을 명시적으로 구성할 필요 없이 AWS에 저장된 데이터 원본에 액세스할 때 암호화를 사용할 수 있다는 이점이 있습니다. 일단 구성되면 최종 사용자가 VPN 기술을 대부분 원활하게 사용할 수 있습니다.

표 2 - AWS VPN을 사용하여 AWS 데이터 원본에 액세스할 때의 고려 사항

|                |  |
|----------------|--|
| <p>기준</p>      | <p>AWS VPN을 사용하여 AWS 데이터 원본에 액세스할 때의 고려 사항</p>   |
| <p>네트워크 연결</p> | <p>VPC의 프라이빗 IP 주소에 연결하거나 리전 또는 VPC 서비스 엔드포인트를 사용하여 데이터 원본을 사용할 수 있습니다. Power BI Desktop은 VPN을 통해 연결하고 데이터 원본(Amazon RDS, Amazon Redshift, Amazon EC2 기반 데이터 원본)에 직접 액세스하거나, 리전 엔드포인트(Amazon Athena)가 있는 서비스의 경우 DNS 구성에 따라 프라이빗 VPC 엔드포인트 또는 리전 엔드포인트를 사용하여 액세스합니다.</p>   |
| <p>보안</p>      | <p>IP 액세스 제어</p> <p>라우팅과 보안 그룹을 조합하여 AWS 클라우드에 저장된 데이터 원본에 대한 액세스를 제어할 수 있습니다.</p> <p>전송 중 데이터 암호화</p> <p>두 유형의 AWS VPN 모두 <a href="#">IPsec</a> 암호화를 사용합니다. 즉, 전송되는 데이터가 AWS와 온프레미스 간에 이동할 때 암호화됩니다. 이렇게 하면 데이터 원본이 암호화된 통신을 사용하도록 구성되지 않은 경우에도 해당 데이터가 인터넷을 통과하는 동안 보호됩니다.</p> <p>인증</p> <p>Site-to-Site VPN은 일회성 구성이 필요하며 일단 설정되면 사용자가 원활하게 사용할 수 있습니다. 최종 사용자는 Site-to-Site VPN을 사용하기 위해 인증할 필요가 없지만 데이터 원본에 대한 인증이 필요합니다.</p> <p>반면 Client VPN은 최종 사용자가 연결을 설정하려면 인증이 필요합니다. Client VPN 인증은</p> |

|           |  |
|-----------|--|
| <p>기준</p> | <p>AWS VPN을 사용하여 AWS 데이터 원본에 액세스할 때의 고려 사항</p>   |
|           | <p>Active Directory(사용자 기반), 상호 인증(인증서 기반) 또는 SAML SSO(사용자 기반)를 통해 수행할 수 있습니다. 인증되면 최종 사용자가 원활하게 연결됩니다. Power BI Desktop에 추가된 AWS 데이터 원본을 사용하려면 인증이 필요합니다.</p> <p>필요한 데이터 집합에 대해서만 읽기 전용 액세스 권한이 있는 자격 증명을 사용하여 AWS 데이터 원본에 인증하는 것이 좋습니다.</p>  |
| <p>성능</p> | <p>AWS VPN의 사용은 인터넷을 통해 이루어집니다. 따라서 성능 범위는 첫 번째로 제시된 시나리오와 유사합니다. 인터넷을 통해 AWS 데이터 원본에 액세스할 때 몇 가지 요소가 전체 Power BI Desktop 성능에 영향을 줄 수 있습니다. 여기에는 다음이 포함됩니다.</p> <ul style="list-style-type: none"> <li>• 액세스되는 데이터 집합의 크기. 데이터 집합이 클수록 검색하는 데 시간이 더 오래 걸립니다. 쿼리를 제한하고 필터를 사용하여 인터넷을 통해 검색되는 데이터의 양을 줄이는 것이 좋습니다.</li> <li>• 대역폭, 대기 시간, 패킷 손실 등의 인터넷 연결 품질. 가능한 경우 지리적으로 가까운 AWS 리전의 데이터에 액세스합니다. 이렇게 하면 대기 시간의 영향이 줄어듭니다. 인터넷을 공유하는 경우 사용량이 적은 시간에 데이터 원본을 로드하거나 사용 가능한 대역폭을 충분히 확보하는 것이 좋습니다.</li> </ul> <p>일반적으로 서로 다른 일중 시간대에 다양한 데이터 집합을 사용하고 점차 사용자를 늘려가며 환경을 테스트하는 것이 좋습니다.</p> |

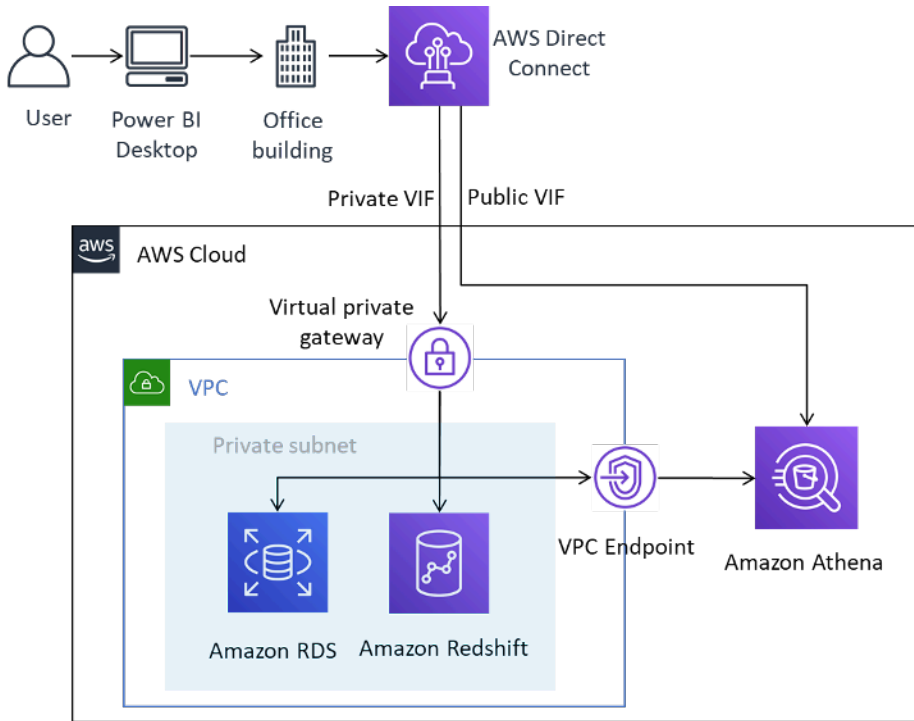
|    |  |
|----|--|
| 기준 | AWS VPN을 사용하여 AWS 데이터 원본에 액세스할 때의 고려 사항  |
| 비용 | VPC에 상주하고 AWS VPN을 사용하여 쿼리되는 데이터 원본에는 AWS VPN 표준 데이터 전송 요금이 부과됩니다. 비용을 절감하려면 쿼리를 제한하고 필터를 사용하여 인터넷을 통해 검색되는 데이터의 양을 줄이는 것이 좋습니다. |

## AWS Direct Connect를 통해 데이터 원본에 연결

AWS Direct Connect를 사용하면 표준 이더넷 광섬유 케이블을 통해 내부 네트워크를 AWS Direct Connect 위치에 연결할 수 있습니다. 케이블의 한쪽 끝을 사용자의 라우터에 연결하고 다른 쪽 끝을 AWS Direct Connect 라우터에 연결합니다. 이 연결을 사용하면 네트워크 경로에서 인터넷 서비스 제공업체를 우회하여 퍼블릭 AWS 서비스(예: Amazon S3) 또는 Amazon VPC에 직접 가상 인터페이스를 만들 수 있습니다. AWS Direct Connect 위치에서 해당 위치와 연결된 리전의 AWS에 액세스할 수 있습니다. 퍼블릭 리전 또는 AWS GovCloud(미국)의 단일 연결을 사용하여 다른 모든 퍼블릭 리전의 퍼블릭 AWS 서비스에 액세스할 수 있습니다.

이 모델에서는 고객의 온프레미스 네트워크가 AWS Direct Connect를 통해 AWS 네트워크에 직접 연결됩니다. AWS Direct Connect를 구성하는 여러 방법이 있지만, 가장 단순한 모드에서는 프라이빗 가상 인터페이스(프라이빗 VIF)라는 구성을 사용하여 VPC 내의 IP 범위에 액세스할 수 있습니다. 퍼블릭 가상 인터페이스(퍼블릭 VIF)를 사용하여 인터넷/퍼블릭 IP 범위에 액세스합니다.





**AWS Direct Connect를 통해 Power BI Desktop을 AWS 데이터 원본에 연결**

Power BI에서 데이터 원본을 추가할 때 프라이빗 IP 주소가 VPC에 있는 경우 DNS 구성 방법에 따라 해당 서비스의 프라이빗 IP 주소 또는 프라이빗 VPC 엔드포인트를 선택합니다.

표 3 - AWS Direct Connect를 사용하여 AWS 데이터 원본에 액세스할 때의 고려 사항

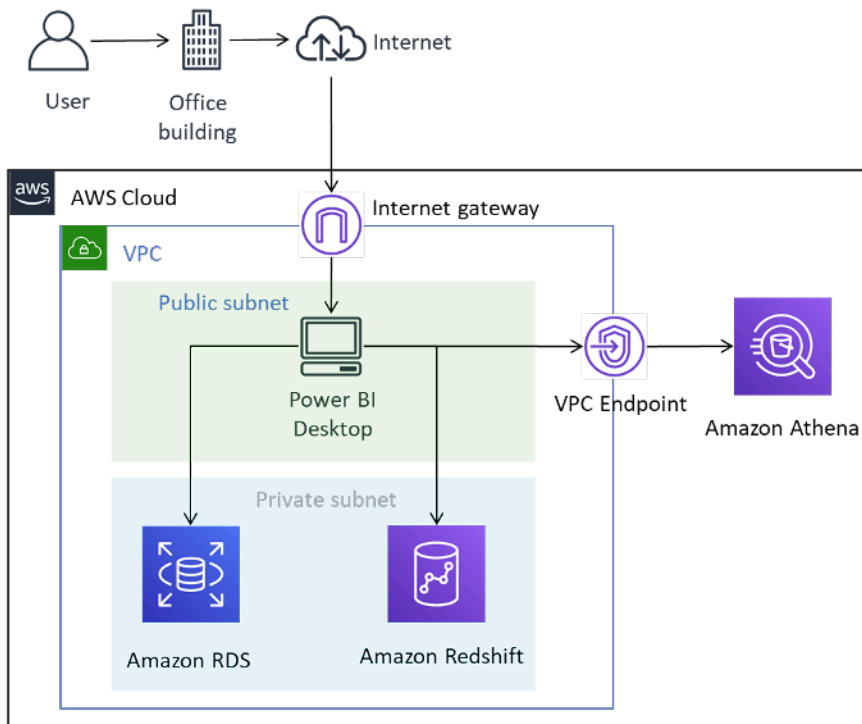
|                |  |
|----------------|--|
| <p>기준</p>      | <p>AWS Direct Connect를 사용하여 AWS 데이터 원본에 액세스할 때의 고려 사항</p>  |
| <p>네트워크 연결</p> | <p>구성된 Direct Connect는 VPC의 프라이빗 IP에 연결하거나 리전 서비스 엔드포인트를 사용하여 데이터 원본에 액세스할 수 있습니다.</p>                 |
| <p>보안</p>      | <p>IP 액세스 제어</p> <p>라우팅과 보안 그룹을 조합하여 AWS 클라우드에 저장된 데이터 원본에 대한 액세스를 제어할 수 있습니다.</p> <p>전송 중 데이터 암호화</p> |

|           |   |
|-----------|---|
| <p>기준</p> | <p>AWS Direct Connect를 사용하여 AWS 데이터 원본에 액세스할 때의 고려 사항</p>   |
|           | <p>Direct Connect는 라인 수준 데이터 암호화를 제공하지 않습니다. 데이터 원본 수준에서 TLS 암호화를 사용하거나 Site-to-Site VPN을 Direct Connect와 결합하여 전송 중에 데이터 및 자격 증명이 손상되지 않도록 하는 것이 좋습니다.</p> <p>인증</p> <p>Direct Connect를 구성한 후에는 추가 인증이 필요하지 않습니다.</p> <p>필요한 데이터 집합에 대해서만 읽기 전용 액세스 권한이 있는 자격 증명을 사용하여 AWS 데이터 원본에 인증하는 것이 좋습니다.</p>  |
| <p>성능</p> | <p>Direct Connect를 사용하면 1Gbps 및 10Gbps 포트를 사용할 수 있습니다. Direct Connect를 지원하는 모든 APN 파트너에게 50Mbps, 100Mbps, 200Mbps, 300Mbps, 400Mbps 및 500Mbps의 속도를 주문할 수 있습니다.</p> <p>인터넷 기반 연결과 비교할 때 Direct Connect 연결은 일반적으로 대역폭이 늘어나고 대기 시간이 줄어듭니다. 연결 용량에 따라 데이터 원본을 로드할 때 더 이상 네트워크 때문에 Power BI Desktop 성능이 제한되지 않을 수 있습니다. 따라서 Direct Connect가 연결된 동일한 리전의 데이터 집합에 액세스할 때 사용자는 쿼리에서 우수한 성능을 기대할 수 있습니다.</p> <p>대용량 데이터 집합을 로드할 때는 Power BI Desktop에 10GB의 데이터 집합 제한이 있다는 점을 유의해야 합니다.</p> |

|           |   |
|-----------|---|
| <p>기준</p> | <p>AWS Direct Connect를 사용하여 AWS 데이터 원본에 액세스할 때의 고려 사항</p>   |
| <p>비용</p> | <p><a href="#">AWS Direct Connect 요금</a>에 설명된 대로 Direct Connect 연결을 사용하여 액세스하는 데이터 원본에는 표준 Direct Connect 포트 요금 외에 표준 Direct Connect 데이터 전송 요금이 부과됩니다. Direct Connect를 통해 AWS 클라우드로 전송된 데이터에는 요금이 부과되지 않습니다.</p> <p>비용을 절감하려면 쿼리를 제한하고 필터를 사용하여 검색되는 데이터의 양을 줄이는 것이 좋습니다.</p> |

## AWS 클라우드에서 Microsoft Power BI Desktop 사용

AWS 클라우드에서 Microsoft Power BI Desktop을 사용하는 것은 이전 섹션에서 설명한 많은 과제에 널리 사용되는 솔루션입니다. 이 모델에서 고객은 AWS 클라우드에서 Microsoft Power BI Desktop을 호스팅한 다음 온프레미스에서 원격으로 액세스합니다. 다음 다이어그램은 한 가지 예를 보여줍니다.



### AWS 클라우드에 배포된 Microsoft Power BI Desktop

다이어그램에서는 사용자가 인터넷을 통해 Desktop에 연결하는 방법을 보여주지만 AWS VPN 및 Direct Connect도 모두 유효한 연결 유형입니다. 그래픽 관리 트래픽만 전송되므로 대역폭 요구 사항은 일반적인 인터넷 연결에 적합합니다.

이 모델에서 Microsoft Power BI Desktop은 퍼블릭 서브넷의 Amazon VPC 내에서 호스트되며 프라이빗 IP 주소가 있는 데이터 원본(예: Amazon RDS 및 Amazon Redshift)에 직접 네트워크로 연결됩니다. VPC 엔드포인트 연결을 대상으로 사용하거나(다이어그램 참조) 리전 퍼블릭 서비스 엔드포인트를 사용하여 Amazon Athena 및 기타 리전별 서비스에 연결할 수 있습니다.

AWS 클라우드에서 Microsoft Power BI Desktop을 호스트하는 방법은 여러 가지가 있습니다.

## 옵션 1: Amazon EC2 인스턴스에 Microsoft Power BI Desktop 설치

이 옵션에서는 Microsoft Windows Server가 설치된 Amazon EC2 인스턴스를 하나 이상 생성합니다. 사용자는 RDP(Remote Desktop Protocol) 애플리케이션을 사용하여 인스턴스에 연결하고 랩톱에 설치된 것처럼 Microsoft Power BI를 설치, 구성, 사용합니다. 기본적으로 Windows Server 설치하는 두 개의 동시 RDP 클라이언트 세션으로 제한됩니다. 더 필요한 경우 Microsoft 리셀러로부터 클라이언트 액세스 라이선스를 구입할 수 있습니다. 단일 서버에서 동시 사용자 수를 늘리면 리소스 경합이 발생할 수 있습니다. 인스턴스 수를 늘려 각 인스턴스의 사용자 수를 줄이면 더 나은 사용자 경험을 얻을 수 있습니다.

## 옵션 2: Amazon WorkSpaces 환경에 Microsoft Power BI 설치

Amazon WorkSpaces는 안전하게 관리형으로 운영되는 서비스형 데스크톱(DaaS) 솔루션입니다. Amazon WorkSpaces를 사용하면 단 몇 분 만에 Windows 또는 Linux 데스크톱을 프로비저닝하고 전 세계 작업자에게 수천 대의 데스크톱을 제공하도록 신속하게 확장할 수 있습니다.

이 옵션에서는 사용자에게 Microsoft Power BI Desktop이 설치된 Microsoft Windows 데스크톱을 프로비저닝합니다. 사용자는 Android, iOS, Fire, Mac, PC, Chromebook 및 Linux 디바이스에서 사용할 수 있는 경량 클라이언트 애플리케이션을 사용하여 Windows 데스크톱에 액세스할 수 있습니다.

Amazon WorkSpaces를 사용하면 Microsoft Power BI Desktop 사용자가 전용 리소스를 사용할 수 있으며 데스크톱을 자동으로 중지하여 비용을 절감할 수 있습니다. 또한 이 옵션은 비 Windows 디바이스에서 Amazon WorkSpaces에 액세스할 수 있을 뿐만 아니라 사용자에게 Microsoft Office 또는 기타 Windows 애플리케이션에 액세스하는 등 다른 작업을 수행할 수 있는 데스크톱 환경을 제공하는 이점이 있습니다.

## 옵션 3: Amazon AppStream 2.0 환경에 Microsoft Power BI 설치

Amazon AppStream 2.0은 완전관리형 애플리케이션 스트리밍 서비스입니다. AppStream 2.0을 통해 중앙에서 데스크톱 애플리케이션을 관리하고 모든 컴퓨터로 안전하게 제공할 수 있습니다. 하드웨어 또는 인프라를 구매, 프로비저닝 및 운영할 필요 없이 전 세계 수많은 사용자로 손쉽게 확장할 수 있습니다.

이 옵션에서는 온프레미스 사용자에게 플러그인을 사용할 필요 없이 HTML5 호환 웹 브라우저로 스트리밍되는 Microsoft Power BI Desktop 애플리케이션만 제공합니다. Amazon WorkSpaces 옵션과 마찬가지로 비 Windows 디바이스도 연결할 수 있으며 사용자는 전용 리소스에 액세스할 수 있습니다.

이 옵션의 한 가지 단점은 기본적으로 AppStream 2.0 애플리케이션의 스토리지 용량이 1GB로 제한되어 더 큰 모델 및 보고서를 저장하는 데 적합하지 않을 수 있다는 것입니다. Amazon FSx for Windows File Server의 파일 스토리지를 결합하여 용량을 늘릴 수 있지만, 이로 인해 고객이 인지해야 할 복잡성이 추가로 발생합니다.

표 4 - AWS 클라우드에서 실행되는 Power BI Desktop에 대한 고려 사항

| 기준      | AWS 클라우드에서 실행되는 Power BI Desktop에 대한 고려 사항  |
|---------|---|
| 네트워크 연결 | <p>데이터 소비자와 데이터 원본 모두 AWS 클라우드에 상주하므로 데이터 원본에 대한 네트워크 연결은 간단합니다. Amazon RDS 및 Amazon Redshift 같이 Amazon VPC에 상주하는 데이터 원본은 직접 액세스할 수 있습니다. 리전 엔드포인트를 사용하는 데이터 원본은 Amazon VPC 인터넷 게이트웨이 또는 Amazon VPC 엔드포인트를 통해 액세스할 수 있습니다.</p> <p>Microsoft Power BI Desktop에 대한 연결은 인터넷, AWS VPN 또는 AWS Direct Connect를 통해 이루어집니다. 세 가지 옵션 모두 대부분의 인터넷 연결이 충족할 수 있는 적당한 요구 사항이 있습니다.</p> |
| 보안      | IP 액세스 제어   |

| 기준 | AWS 클라우드에서 실행되는 Power BI Desktop에 대한 고려 사항  |
|----|---|
|    | <p>고객은 라우팅과 보안 그룹을 조합하여 AWS 클라우드에 저장된 데이터 원본에 대한 액세스를 제어할 수 있습니다.</p> <p>첫 번째 옵션(Amazon EC2 사용)의 경우에는 라우팅 및 보안 그룹의 조합을 사용하여 특정 온프레미스 CIDR 범위만 허용할 수도 있습니다.</p> <p>전송 중 데이터 암호화</p> <p>Amazon VPC 내의 데이터 원본은 데이터 전송에 암호화를 사용하도록 구성하는 것이 좋습니다. 리전별 서비스는 이미 TLS 암호화를 사용하고 있습니다.</p> <p>관리 액세스는 세 가지 옵션 모두에서 암호화됩니다.</p> <p>인증</p> <p>필요한 데이터 집합에 대해서만 읽기 전용 액세스 권한이 있는 자격 증명을 사용하여 AWS 데이터 원본에 인증하는 것이 좋습니다.</p> <p>세 가지 옵션 모두 Microsoft Power BI Desktop을 호스트하는 환경에 액세스하기 전에 원격 사용자가 자신의 자격 증명을 증명할 수 있도록 인증 메커니즘을 사용해야 합니다. 이 메커니즘은 일반적으로 사용자 이름 및 암호이지만 MFA 옵션도 사용할 수 있습니다.</p> |

|    |   |
|----|---|
| 기준 | AWS 클라우드에서 실행되는 Power BI Desktop에 대한 고려 사항  |
| 성능 | AWS 클라우드에서 실행할 때 Microsoft Power BI Desktop 성능은 일반적으로 온프레미스에서 실행할 때보다 우수합니다. 네트워킹 및 컴퓨팅 인스턴스 모두 최신형이고 더 고사양이며 애플리케이션과 데이터 원본 간의 대기 시간이 짧을 가능성이 높기 때문입니다.  |
| 비용 | <p>Microsoft Power BI Desktop을 AWS 클라우드에 배치하면 데이터 전송 요금을 크게 줄이거나 완전히 없앨 수 있습니다. 다른 가용 영역, VPC 또는 리전의 데이터 원본에 액세스할 때는 데이터 전송 요금이 부과될 수 있으므로 주의해야 합니다.</p> <p>각 옵션에서 감안해야 할 추가 비용 고려 사항이 있습니다. Amazon EC2, Amazon WorkSpaces, Amazon AppStream 2.0은 사용량을 기준으로 요금이 부과됩니다. 자세한 내용은 각 서비스의 요금 페이지를 참조하십시오.</p> |

## Microsoft Power BI Desktop 연결 옵션 요약

데이터 집합 요구 사항이 단순한 소수의 사용자에게는 온프레미스에서 Microsoft Power BI Desktop을 실행하고 인터넷을 통해 안전하게 연결하거나 AWS VPN을 사용하는 것이 적절한 솔루션일 수 있습니다. 이 모델에서는 보안을 구성하고 유지 관리해야 합니다. 또한 이 구성을 테스트하여 사용자의 성능 기대치를 충족하는지 확인하는 것이 좋습니다.

사용자 수가 증가하면 AWS Direct Connect를 통한 연결을 고려하는 것이 좋습니다. Direct Connect는 더 큰 데이터 집합을 로드할 때 더 나은 사용자 경험을 제공합니다. 사용자가 대규모 데이터 집합을 전송할 때 비용에 미치는 영향을 인지하도록 해야 합니다.

AWS 클라우드에서의 Microsoft Power BI Desktop 실행을 평가할 것을 권장합니다. 최종 사용자에게 최상의 성능 경험과 클라우드 관리자에게 최상의 관리 환경을 제공할 가능성이 높기 때문입니다. 특히

Amazon WorkSpaces는 소수의 사용자에서 수천 명의 사용자로 확장할 수 있습니다. 또한 이러한 서비스는 상당한 보안 및 관리 이점을 제공합니다.



# Microsoft Power BI 서비스를 AWS 데이터 원본에 연결

Microsoft Power BI 서비스(SaaS)는 인터넷에서 액세스할 수 있는 데이터 원본 또는 Amazon VPC의 비공개 데이터 원본에 직접 연결할 수 있습니다. 비공개 데이터 원본에 연결하려면 Microsoft 온프레미스 데이터 게이트웨이라는 애플리케이션 구성 요소가 필요합니다. Microsoft 온프레미스 데이터 게이트웨이는 VPC의 Amazon EC2 인스턴스에 다운로드 및 설치되고 Microsoft Power BI 자격 증명을 사용하여 구성됩니다. 이 게이트웨이는 인터넷을 통해 Microsoft Azure Service Bus에 대한 아웃바운드 연결을 설정하며 Microsoft Power BI에서 액세스할 수 있는 데이터 원본에 연결하도록 구성됩니다. 대규모 배포에서는 여러 온프레미스 데이터 게이트웨이를 사용하여 로드 밸런싱하거나 내결함성을 높일 수 있습니다.

Microsoft 온프레미스 데이터 게이트웨이를 사용하면 AWS 고객들이 보고하고 있듯이 여러 가지 실질적 이점을 얻을 수 있습니다.

## 1. 보안 태세 개선

Microsoft 온프레미스 데이터 게이트웨이는 Microsoft Azure 클라우드로부터의 인바운드 연결은 허용하지 않으며 Azure Service Bus로의 아웃바운드 연결만 시작합니다. 이 단방향 트래픽 모델을 사용하면 데이터 원본을 비공개로 유지하고 퍼블릭 인터넷에 노출하지 않을 수 있습니다.

## 2. 데이터 전송 비용 절감

데이터 원본에 연결할 때 Microsoft 온프레미스 데이터 게이트웨이는 스포링이라는 프로세스에서 전체 결과 집합을 검색하여 로컬로 저장합니다. 결과가 Power BI 서비스로 전송되기 전에 데이터가 압축됩니다. 사용자는 일반적으로 10:1 압축률을 보고하는데, 이는 인터넷을 통해 데이터를 전송하는 시간을 단축할 뿐만 아니라 송신 요금을 줄입니다.

## 3. 솔루션 비용 절감

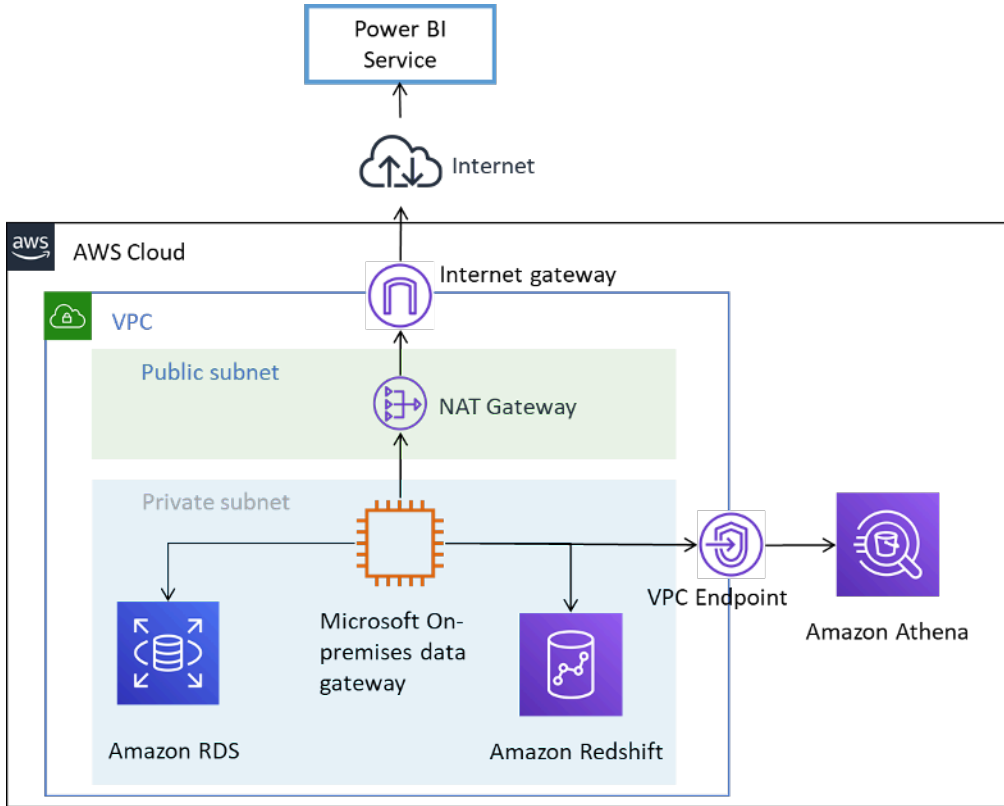
Microsoft 온프레미스 데이터 게이트웨이를 사용하는 경우 서비스에 필요한 일부 데이터 처리가 게이트웨이에서 대신 수행됩니다. Amazon EC2를 Savings Plans 또는 예약 인스턴스와 같은 비용 절감 플랜과 함께 사용하면 전체 BI 솔루션 비용을 줄이는 데 도움이 될 수 있습니다.

## 권장 구성

데이터 원본이 포함된 프라이빗 서브넷의 Amazon EC2 인스턴스에 Microsoft 온프레미스 데이터 게이트웨이를 설치하는 것이 좋습니다. 이 서브넷은 퍼블릭 서브넷에 설치된 [Amazon VPC NAT 게이트웨이](#)를 통해 인터넷으로 요청을 라우팅하도록 구성됩니다. 네트워크 주소 변환(NAT) 게이트웨이를 사용하면 프라이빗 서브넷의 인스턴스를 인터넷 또는 기타 AWS 서비스에 연결하는 한편, 인터넷에서 해

당 인스턴스에 연결하지 못하게 할 수 있습니다.고가용성 데이터 게이트웨이 구현이 필요한 경우 여러 AWS 가용 영역에 걸쳐 있는 여러 EC2 인스턴스에 설치된 온프레미스 데이터 게이트웨이 클러스터를 사용하는 것이 좋습니다. 자세한 내용은 [다른 게이트웨이를 추가하여 클러스터 생성](#)을 참조하십시오.

이 섹션에 제시된 옵션은 Amazon RDS, Amazon Redshift 및 Amazon Athena를 보여줍니다. 모든 AWS 데이터 원본에 대한 자세한 내용은 [부록: Microsoft Power BI에서 지원하는 AWS 데이터 원본](#)을 참조하십시오.



AWS 데이터 원본을 Microsoft Power BI 서비스에 연결

## 추가 고려 사항

표 5 - AWS 클라우드에서 데이터 원본을 사용하는 Microsoft Power BI 서비스에 대한 고려 사항

|         |   |
|---------|---|
| 기준      | AWS 클라우드에서 데이터 원본을 사용하는 Microsoft Power BI 서비스에 대한 고려 사항                        |
| 네트워크 연결 | 데이터 소비자와 데이터 원본이 모두 AWS 클라우드에 상주하기 때문에 데이터 원본에 대한 Microsoft 온프레미스 데이터 게이트웨이 연결은 |

|    |  |
|----|--|
| 기준 | <p>AWS 클라우드에서 데이터 원본을 사용하는 Microsoft Power BI 서비스에 대한 고려 사항</p> <p>간단합니다. Amazon RDS 및 Amazon Redshift 같이 Amazon VPC에 상주하는 데이터 원본은 직접 액세스할 수 있습니다. 리전 엔드포인트를 사용하는 데이터 원본은 Amazon VPC 인터넷 게이트웨이 또는 Amazon VPC 엔드포인트를 통해 액세스할 수 있습니다.</p> <p>Microsoft Power BI 서비스에 대한 Microsoft 온프레미스 데이터 게이트웨이 연결은 인터넷을 통해 이루어지며 아웃바운드 연결만 가능합니다.</p> |
|----|--|

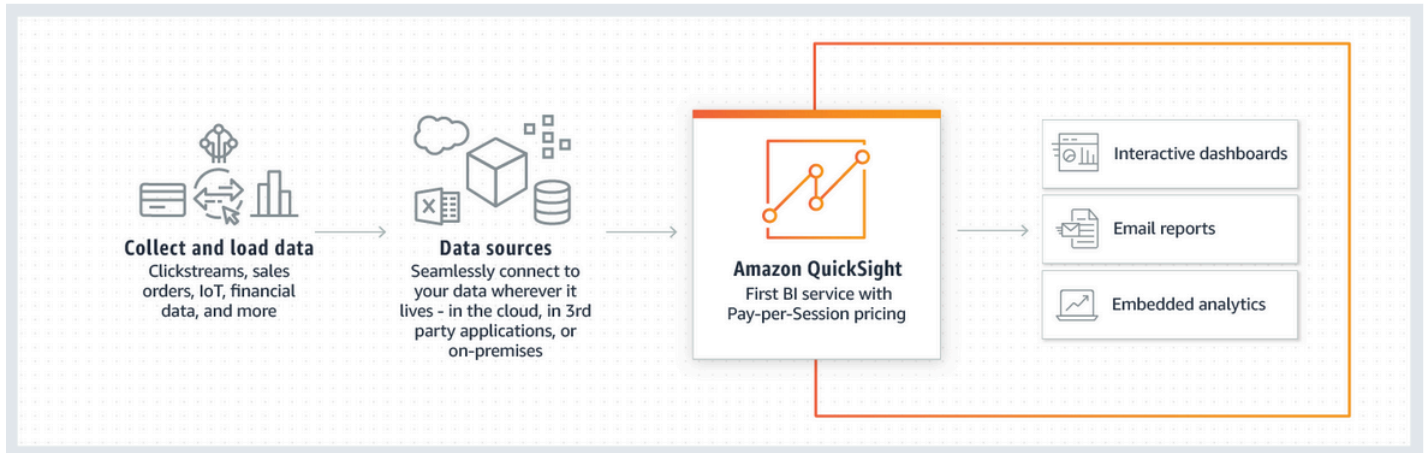
|           |   |
|-----------|---|
| <p>기준</p> | <p>AWS 클라우드에서 데이터 원본을 사용하는 Microsoft Power BI 서비스에 대한 고려 사항</p>   |
| <p>보안</p> | <p>IP 액세스 제어</p> <p>라우팅과 보안 그룹을 조합하여 AWS 클라우드 내에 저장된 데이터 원본에 대한 액세스를 제어할 수 있습니다.</p> <p>Microsoft 온프레미스 데이터 게이트웨이는 Amazon EC2 인스턴스에 설치되므로 연결된 보안 그룹을 사용하여 운영 체제에 대한 인바운드 액세스를 제한할 수 있습니다. 이 게이트웨이는 인바운드 요청을 수락하지 않습니다. 인스턴스에는 퍼블릭 IP 주소가 필요하지 않으며 퍼블릭 IP 주소로 구성해서는 안 됩니다.</p> <p>전송 중 데이터 암호화</p> <p>Amazon VPC 내의 데이터 원본은 데이터 전송에 암호화를 사용하도록 구성하는 것이 좋습니다. 리전별 서비스는 이미 TLS 암호화를 사용하고 있습니다.</p> <p>TCP 대신 HTTPS를 사용하여 Microsoft Azure Service Bus에 연결하도록 Microsoft 온프레미스 데이터 게이트웨이 연결을 구성할 수 있습니다. 통신에는 HTTPS 모드를 사용하는 것이 좋습니다. 이는 2019년 6월 게이트웨이 소프트웨어 버전 릴리스 이후 새 게이트웨이 설치의 기본값이기도 합니다.</p> <p>인증</p> <p>필요한 데이터 집합에 대해서만 읽기 전용 액세스 권한이 있는 자격 증명을 사용하여 AWS 데이터 원본에 인증하는 것이 좋습니다. 데이터 원본에 입력하는 자격 증명은 암호화되어 게이트웨이 클라우드 서비스에 저장됩니다. 자격 증명</p> |

|    |   |
|----|---|
| 기준 | AWS 클라우드에서 데이터 원본을 사용하는 Microsoft Power BI 서비스에 대한 고려 사항은 온프레미스의 게이트웨이에서 해독됩니다.   |
|    | <p>은 온프레미스의 게이트웨이에서 해독됩니다. (데이터 원본에 입력하는 자격 증명은 암호화된 게이트웨이 클라우드 서비스에 저장됩니다. 자격 증명은 온프레미스의 게이트웨이에서 해독됩니다.)</p> <p>Microsoft Power BI 자격 증명에 안전하게 제어되도록 해야 합니다. 서비스에 액세스하면 AWS 데이터 원본 및 여기에 포함될 수 있는 중요한 정보에 액세스할 수 있기 때문입니다.</p> |
| 성능 | AWS 클라우드의 Microsoft 온프레미스 데이터 게이트웨이는 Amazon EC2 인스턴스의 크기를 지정하고 확장할 수 있기 때문에 일반적으로 성능이 우수합니다. 또한 리전 네트워킹 및 인터넷 연결에서도 속도가 빠릅니다.   |

|    |   |
|----|---|
| 기준 | AWS 클라우드에서 데이터 원본을 사용하는 Microsoft Power BI 서비스에 대한 고려 사항  |
| 비용 | <p>Amazon EC2 인스턴스 요금, 데이터 전송 요금 및 Amazon NAT 게이트웨이 요금의 세 가지 요소를 고려해야 합니다.</p> <p>Microsoft의 <a href="#">요구 사항</a>에 따라 Amazon EC2 인스턴스의 크기를 지정합니다. 비용을 절감하기 위해 Amazon EC2 예약 인스턴스 또는 <a href="#">AWS Savings Plans</a>를 구매할 수 있습니다.</p> <p>Microsoft 온프레미스 데이터 게이트웨이에서 Microsoft BI 서비스로 전송된 데이터에는 VPC 송신 요금이 부과됩니다. 고객들은 데이터 게이트웨이를 사용하여 트래픽 양이 줄어드는 것(10:1 압축)을 보고하고 있지만, 관련 데이터만 전송되도록 쿼리를 제한하고 필터를 사용하는 것이 좋습니다.</p> <p>Microsoft 온프레미스 데이터 게이트웨이가 다른 가용 영역 또는 다른 AWS 리전의 데이터 원본에 연결하는 경우 데이터 전송 요금도 적용됩니다.</p> <p>Microsoft 온프레미스 데이터 게이트웨이가 프라이빗 서브넷에 있고 AWS NAT 게이트웨이를 사용하는 경우 시간당 사용 및 데이터 처리 요금이 적용됩니다. 자세한 내용은 <a href="#">Amazon VPC 요금</a>을 참조하십시오.</p> |

# Amazon QuickSight 사용

Microsoft Power BI 제품군을 AWS와 함께 사용하려는 고객은 [Amazon QuickSight](#)를 대안으로 평가하는 것이 좋습니다. 이 완전관리형 클라우드 서비스는 기본적으로 AWS의 데이터 원본에 연결되므로 다른 BI 솔루션에 비해 복잡성과 비용이 줄어듭니다.



## Amazon QuickSight 작동 방식

다른 BI 솔루션과 비교할 때 Amazon QuickSight는 다음과 같은 이점이 있습니다.

- Amazon QuickSight를 사용하면 클라이언트 애플리케이션을 다운로드하여 설치할 필요가 없습니다. 작성 및 보고를 포함한 모든 기능은 모든 플랫폼(Windows, Mac, Linux 등)에서 웹 브라우저를 통해 액세스할 수 있습니다.
- Amazon QuickSight는 완전관리형 클라우드 네이티브 SaaS 애플리케이션으로 제공되며 간편하게 대시보드를 구축하고 프로덕션에 배포할 수 있습니다. 이 서비스는 서버리스이므로 사용자를 지원하는 데 필요한 노드 및 서버 수를 계산할 필요가 없습니다. 또한 QuickSight는 복원력을 위해 AWS에서 제공하는 고가용성 기능을 최대한 활용합니다.
- QuickSight 내에서 포인트 앤 클릭 인터페이스를 통해 사용자를 추가할 수 있으므로 소규모 또는 대규모 설정에서 손쉽게 시작할 수 있습니다. 외부 관리자의 개입이 필요하지 않습니다.
- Amazon QuickSight는 빠른 응답 시간(밀리초)과 대화형 시각화를 위해 SPICE(초고속 병렬 인 메모리 계산 엔진)로 구동됩니다. 데이터 집합은 현재 최대 200GB까지 확장할 수 있습니다.
- Amazon QuickSight 요금은 단순하고 저렴하며 보고서 작성자와 보고서 리더라는 두 가지 구성 요소가 있습니다. 대화형 대시보드를 생성하고 게시하는 보고서 작성자는 사용자당 요금이 부과됩니다. 특정 월에 로그인하지 않은 사용자가 있는 경우 해당 사용자에 대한 요금이 부과되지 않습니다. 보고서 리더에는 30분 세션당 요금이 부과되며, 리더당 월 최대 5.00 USD가 부과됩니다. 무료 평가판을

사용하면 비용 없이 Amazon QuickSight를 평가할 수 있습니다. 자세한 내용은 [Amazon QuickSight 요금](#)을 참조하십시오.



## 결론

Microsoft Power BI Desktop을 사용하려는 경우, AWS에서 파악하기로는 고객은 일반적으로 인터넷을 통해 데이터 원본에 연결하여 온프레미스에서 소프트웨어를 실험하기 시작합니다. AWS VPN 및 Direct Connect를 사용하기 위한 프라이빗 연결 옵션이 있지만, 많은 고객은 Amazon WorkSpaces에서 Microsoft Power BI Desktop을 실행하면 성능이 더 우수한 환경을 제공한다는 결론을 내렸습니다.

AWS의 데이터 원본을 Microsoft Power BI 서비스에 연결하려는 경우 이것이 확립된 아키텍처 패턴이라는 것을 알기 때문에 안심할 수 있습니다. Microsoft 온프레미스 데이터 게이트웨이를 Amazon VPC 내에 설치하고 Amazon RDS, Amazon Redshift, Amazon Athena, Amazon OpenSearch, AWS Lake Formation과 같은 데이터 원본을 서비스에 원활하게 연결할 수 있습니다.

AWS에서 Microsoft Power BI Report Server를 사용하려는 경우에도 확립된 패턴이 있습니다. Microsoft Power BI Report Server는 AWS 데이터 원본과 가까이 위치한 Amazon VPC 내에 설치할 수 있으며 작성자와 보고서 사용자 모두를 위한 연결 옵션이 있습니다.

자체 관리형 BI 솔루션을 설치, 구성, 패치 및 확장하는 복잡성을 가중하지 않으면서 동일한 비즈니스 성과를 제공하는 솔루션을 원하는 경우 Amazon QuickSight를 사용하는 것이 좋습니다. 이 완전관리형 서비스는 간단한 웹 브라우저 환경에서 필요한 모든 기능을 사용자당 요금제로 제공합니다. 설치할 필요가 없으며 추가 구성 요소가 필요하지 않습니다.

이것이 AWS와 함께 하는 비즈니스 인텔리전스 여정의 시작일 뿐이길 바랍니다. 시작하는 데 도움이 되는 추가 리소스는 [부록: Microsoft Power BI에서 지원하는 AWS 데이터 원본](#)을 참조하십시오.

## 기여자

이 문서를 작성하는 데 도움을 주신 분들입니다.

- Ralph Holm, 선임 솔루션스 아키텍트, Amazon Web Services
- Barret Newman, 선임 솔루션스 아키텍트, Amazon Web Services
- Fabrizio Napolitano, 데이터베이스 및 분석 스페셜리스트 솔루션스 아키텍트, Amazon Web Services

## 추가 자료

- [인사이트 및 분석을 위해 Power BI를 Amazon Redshift와 통합](#)
- [Power BI Desktop에서 Amazon Redshift 데이터베이스에 연결](#)

# 문서 개정

이 백서의 업데이트에 대한 알림을 받으려면 RSS 피드를 구독하세요.

| 업데이트 기록-변경            | update-history-description   | update-history-date |
|-----------------------|--|---------------------|
| <a href="#">업데이트</a>  | 데이터 원본 추가: Amazon OpenSearch 및 AWS Lake Formation. 업데이트: Athena, RDS 및 Redshift 데이터 원본. Microsoft 온프레미스 데이터 게이트웨이에 대한 추가 지침. | 2021년 11월 3일        |
| <a href="#">검토</a>    | 기술 정확성이 검토되었습니다.   | 2021년 3월 1일         |
| <a href="#">업데이트</a>  | Report Server 라이선싱이 업데이트되었습니다.   | 2021년 1월 15일        |
| <a href="#">최초 게시</a> | 처음 게시된 백서  | 2020년 11월 1일        |

## 부록: Microsoft Power BI에서 지원하는 AWS 데이터 원본

지원되는 데이터 원본의 전체 목록은 Microsoft에서 제공합니다([Power BI 데이터 원본 참조](#)). 하지만 다음 섹션에서 일부 독자에게 도움이 될 수 있도록 각 AWS 데이터 원본에 대한 사용 및 구성 지침을 제공합니다.

### Amazon Redshift

Amazon Redshift는 AWS 클라우드에서 제공되는 페타바이트 규모의 완전관리형 데이터 웨어하우스 서비스입니다. Amazon Redshift 데이터 웨어하우스는 노드라고 하는 컴퓨팅 리소스의 모음으로, 노드는 클러스터라고 하는 그룹을 구성합니다. 각 클러스터는 Amazon Redshift 엔진을 실행하며, 하나 이상의 데이터베이스를 포함합니다.

다음과 같은 경우 Amazon Redshift를 사용하는 것이 좋습니다.

- 네이티브 클라우드 데이터 웨어하우스를 구축 또는 마이그레이션하는 경우
- 몇 테라바이트에서 수백 테라바이트로 확장해야 할 수 있는 경우
- Power BI 사용자가 Amazon S3에 저장된 데이터 레이크의 데이터에 투명하게 액세스하고 이를 데이터 웨어하우스의 테이블과 조인하도록 허용하려는 경우
- 쿼리 워크로드에 다음이 포함되는 경우
  - 대규모(기가바이트 및 테라바이트 규모) 테이블에 대한 집계를 계산하는 쿼리
  - 여러 조인 및 하위 쿼리가 있는 매우 복잡한 SQL
  - 대시보드에서 사용되는 복잡한 분석 쿼리와 단순하고 고도로 필터링된 쿼리가 혼합됨

Amazon Redshift를 Microsoft Power BI와 함께 사용할 때는 다음 사항에 유의하십시오.

- Amazon Redshift는 기본적으로 Microsoft Power BI Desktop 및 Power BI 서비스 모두에서 Power BI 데이터 원본으로 지원되며, 각각 가져오기 및 직접 쿼리 모드를 지원합니다.
- Redshift 클러스터는 퍼블릭 서브넷에서 시작할 수 있고 인터넷으로부터 액세스를 허용하도록 구성할 수 있지만, 대부분의 고객은 보안을 강화하기 위해 프라이빗 서브넷에서 클러스터를 시작하는 것을 선호합니다. 프라이빗 서브넷을 사용하는 경우 온프레미스 데이터 게이트웨이를 사용하여 Power BI 서비스에서 Amazon Redshift로 연결합니다.
- Redshift 커넥터는 Power BI Desktop 및 서비스에서 Azure AD 인증을 지원합니다.

- Spectrum을 통해 액세스되는 외부 테이블은 기본 Redshift 테이블과 다르게 취급되지 않으며 Power BI는 이러한 테이블을 구별할 수 있는 수단이 없습니다. 외부 테이블의 데이터에 액세스할 때는 다음 사항을 확인해야 합니다.
- 문자열이 포함된 열은 AWS Glue 데이터 카탈로그에서 'STRING'이 아닌 'VARCHAR'로 분류해야 합니다. 그렇지 않으면 Power BI에서 다음 오류가 발생합니다. Exception: OLE DB or ODBC error: [Expression.Error] We couldn't fold the expression to the data source. Please try a simpler expression..
- ARRAY와 같은 복잡한 데이터 형식을 포함하는 열은 지원되지 않습니다. 복잡한 데이터 형식이 포함된 열을 사용하면 Power BI에서 다음 오류가 발생합니다. Exception: ODBC: ERROR [42703] [Microsoft]Amazon Redshift Error occurred while trying to execute a query

모델에 포함해야 하는 경우 사용자 수준에서 JSON 직렬화를 활성화하거나(Amazon Redshift에서) 복잡한 데이터 형식을 네이티브 테이블의 SUPER 열에 저장할 수 있습니다.

## Amazon RDS

Amazon RDS를 사용하면 클라우드에서 관계형 데이터베이스를 손쉽게 설치, 운영 및 조정할 수 있습니다. Amazon RDS는 여러 데이터베이스 인스턴스 유형(메모리, 성능 또는 I/O 최적화)으로 제공되며 Amazon Aurora, PostgreSQL, MySQL, MariaDB, Oracle Database, SQL Server 등 6개의 익숙한 데이터베이스 엔진 중에서 선택할 수 있습니다.

다음과 같은 경우 RDS를 사용하는 것이 좋습니다.

- 운영 데이터 스토어를 구축하는 경우
- SQL Server 또는 Oracle Database 데이터 웨어하우스를 클라우드로 마이그레이션하지만 리팩토링에는 관심이 없는 경우
- 쿼리 워크로드에 다음이 포함되는 경우
  - 쉽게 인덱싱할 수 있는 테이블에서 고도로 필터링된 데이터에 액세스하는 쿼리
  - 중소 규모 테이블(기가바이트 규모)에 대한 분석 쿼리
  - 대시보드에서 사용되는 중간 복잡도의 분석 쿼리와 단순하고 고도로 필터링된 쿼리가 혼합됨

Amazon RDS를 Microsoft Power BI와 함께 사용할 때는 다음 사항에 유의하십시오.

- Amazon RDS는 SQL Server, MariaDB, MySQL, Oracle Database, PostgreSQL 등 여러 데이터베이스 엔진을 제공합니다. 데이터베이스 엔진은 Amazon RDS 서비스가 아니라 Power BI Desktop 및 Power BI 서비스에 나열됩니다.
- Amazon Aurora의 경우 선택한 데이터베이스 엔진에 따라 내 SQL 또는 PostgreSQL 연결 유형을 사용합니다.
- Amazon RDS 인스턴스는 퍼블릭 서브넷에서 시작할 수 있고 인터넷으로부터 액세스를 허용하도록 구성할 수 있지만, 대부분의 고객은 보안을 강화하기 위해 프라이빗 서브넷에서 클러스터를 시작하는 것을 선호합니다. 프라이빗 서브넷을 사용하는 경우 온프레미스 데이터 게이트웨이를 사용하여 Power BI 서비스에서 RDS로 연결합니다.
- Amazon RDS를 사용하면 Express, Web, Standard 및 Enterprise 버전을 비롯하여 여러 SQL Server(2012, 2014, 2016, 2017 및 2019) 버전을 배포할 수 있습니다.

## Amazon Athena

Amazon Athena는 표준 SQL을 사용해 Amazon S3에 저장된 데이터를 간편하게 분석할 수 있는 대화식 쿼리 서비스입니다. Athena는 AWS Glue 데이터 카탈로그와 즉시 통합되므로, 다양한 서비스에 걸쳐 통합된 메타데이터 리포지토리를 생성하고, 데이터 원본을 크롤링하여 스키마를 검색하고 데이터 카탈로그를 신규 및 수정된 테이블 정의와 파티션 정의로 채우며, 스키마 버전을 관리할 수 있습니다.

다음과 같은 경우 데이터 원본으로 Athena를 사용하는 것이 좋습니다.

- 데이터 레이크를 직접 쿼리하려는 경우
- 쿼리 워크로드에 다음이 포함되는 경우
  - 대규모(기가바이트 및 테라바이트 규모) 테이블에 대한 집계를 계산하는 쿼리
  - 탐색을 위한 대화형 임시 SQL

Amazon Athena를 Microsoft Power BI와 함께 사용할 때는 다음 사항에 유의하십시오.

- Microsoft Power BI의 2021년 7월 릴리스에서 Amazon Athena용 Microsoft 인증 커넥터가 도입되었습니다. Amazon Athena용 Microsoft Power BI 커넥터를 사용하여 Microsoft Power BI Desktop에서 Amazon Athena의 데이터를 분석할 수 있습니다. Power BI 서비스에 콘텐츠를 게시한 후 Microsoft 온프레미스 데이터 게이트웨이를 사용하여 온디맨드 또는 예약된 새로 고침을 통해 콘텐츠를 최신 상태로 유지할 수 있습니다.
- Amazon Athena용 Microsoft Power BI 커넥터는 가져오기 및 직접 쿼리 데이터 연결 모드를 모두 지원합니다. 가져오기 모드에서는 선택한 테이블 및 열을 쿼리할 수 있도록 Power BI Desktop으로 가

저입니다. 직접 쿼리 모드에서는 데이터를 Power BI Desktop으로 가져오거나 복사하지 않는 대신 Power BI Desktop이 기본 데이터 원본을 직접 쿼리합니다.

- Amazon Athena용 Microsoft Power BI 커넥터에 대한 자세한 내용은 [Amazon Athena Power BI 커넥터 사용](#)을 참조하십시오.
- Amazon Athena용 Microsoft Power BI 커넥터를 사용하려면 Amazon Athena를 쿼리하기 위해 시스템에서 Amazon Athena ODBC 드라이버와 유효한 ODBC DSN 구성을 사용해야 합니다. 최신 ODBC 드라이버를 다운로드하고 구성 정보는 확인하려면 [ODBC를 사용하여 Amazon Athena에 연결](#)을 참조하십시오.
- Amazon Athena용 Microsoft Power BI 커넥터를 사용할 때의 구성 단계 및 모범 사례에 대한 자습서는 [Amazon Athena를 사용하여 Microsoft Power BI에서 신속하게 대시보드 만들기](#)를 참조하십시오.

## Amazon OpenSearch Service(Amazon Elasticsearch Service 후속)

JSON 기반 검색 쿼리 DSL을 사용하는 대신 SQL을 사용하여 Amazon OpenSearch Service를 쿼리할 수 있습니다. SQL을 사용한 쿼리는 이 언어에 이미 익숙하거나 도메인을 사용하는 애플리케이션(예: Microsoft Power BI)과 통합하려는 경우에 유용합니다.

다음과 같은 경우 데이터 원본으로 Amazon OpenSearch Service를 사용하는 것이 좋습니다.

- 로그 파일 또는 JSON 출력과 같은 반정형 데이터가 있고 정보를 빠르게 검색, 분석 또는 시각화해야 하는 경우

Amazon OpenSearch Service를 Microsoft Power BI와 함께 사용할 때는 다음 사항에 유의하십시오.

- Amazon OpenSearch Service에 연결하려면 ODBC(Open Database Connectivity) 드라이버가 필요합니다. 이 Windows 및 macOS용 읽기 전용 ODBC 드라이버를 사용하여 비즈니스 인텔리전스(BI)와 데이터 시각화 애플리케이션(예: [Tableau](#), [Microsoft Excel](#), [Power BI](#))을 클러스터의 SQL 플러그인에 연결할 수 있습니다. 이 드라이버는 Open Distro for Elasticsearch [시작 및 다운로드](#) 웹 사이트에서 사용할 수 있습니다. 구성 지침은 [Windows에서 DSN 구성](#)을 참조하십시오.
- 지금은 가져오기 모드만 지원됩니다.
- 현재, Power BI Desktop을 Amazon OpenSearch Service에 연결하려면 사용자 지정 커넥터를 사용해야 합니다. 시작하려면 [Open Distro For ElasticSearch를 Microsoft Power BI Desktop에 연결](#)을 참조하십시오.
- Microsoft 온프레미스 데이터 게이트웨이를 통해 Power BI 서비스를 Amazon OpenSearch Service에 연결하는 자세한 내용은 [Open Distro For ElasticSearch를 Microsoft Power BI 서비스에 연결](#)을 참조하십시오.



## AWS Lake Formation

Lake Formation이 데이터베이스 및 객체 스토리지의 데이터를 수집하고 카탈로그화한 후, 새로운 [Amazon S3](#) 데이터 레이크로 옮긴 다음, 기계 학습 알고리즘을 사용해 정리 및 분류하고, 민감한 데이터에 대한 액세스를 보호하도록 지원합니다. 이 과정이 마무리되면 사용자는 사용 가능한 데이터 세트 및 적절한 사용 방법이 설명된 중앙 집중식 [데이터 카탈로그](#)에 액세스할 수 있습니다. 그런 다음 [Amazon Redshift](#), [Amazon Athena](#) 및 [Amazon EMR](#) for Apache Spark(베타) 등 원하는 분석 및 기계 학습 서비스를 통해 이러한 데이터 세트를 활용할 수 있습니다. Lake Formation은 [AWS Glue](#)에서 제공되는 기능을 기반으로 합니다.

기존의 IAM 기반 컨트롤 대신 데이터 레이크에 대한 세분화된 (행 및 열) 수준의 액세스가 필요한 경우 Lake Formation을 사용하는 것이 좋습니다.

Lake Formation을 Microsoft Power BI와 함께 사용할 때는 다음 사항에 유의하십시오.

- Power BI Desktop 또는 Power BI 서비스를 사용하여 Lake Formation 데이터 카탈로그의 데이터를 쿼리하려면 Athena에서 데이터를 쿼리하는 것과 동일한 프로세스 및 구성을 사용합니다. Lake Formation 권한 모델을 사용하는 경우 Amazon Athena의 ODBC DSN 구성에 "LakeformationEnabled" 속성 키가 "true"로 설정되었는지 확인합니다. 이 설정은 Amazon Athena ODBC 드라이버에게 권한 부여를 위해 AWS Security Token Service를 직접 사용하는 대신 Lake Formation 서비스를 사용하도록 지시합니다. 자세한 내용은 [ODBC를 사용하여 Amazon Athena에 연결](#) 설명서를 참조하십시오.
- 기존 데이터 카탈로그 동작과의 호환성을 위해 활성화된 "IAM 액세스 제어만 사용" 설정은 완벽한 호환성을 제공합니다.
- AWS Glue 데이터 권한을 Lake Formation 모델로 업그레이드하면 호환성 문제가 발생할 수 있으므로 사용하기 전에 테스트해야 합니다. 예비 테스트에 따르면 열 수준 부여 또는 거부 호환되지만 행 및 셀 수준 필터링은 아직 평가판 상태이고 변경될 수 있으므로 테스트하지 않았습니다.

## 고지 사항

고객은 본 문서에 포함된 정보를 독자적으로 평가할 책임이 있습니다. 본 문서는 (a) 정보 제공만을 위한 것이며, (b) 사전 고지 없이 변경될 수 있는 현재의 AWS 제품 제공 서비스 및 사례를 보여 주며, (c) AWS 및 자회사, 공급업체 또는 라이선스 제공자로부터 어떠한 약정 또는 보증도 하지 않습니다. AWS 제품 또는 서비스는 명시적이든 묵시적이든 어떠한 종류의 보증, 진술 또는 조건 없이 '있는 그대로' 제공됩니다. 고객에 대한 AWS의 책임과 법적 책임은 AWS 계약서에 준하며 본 문서는 AWS와 고객 간의 계약에 포함되지 않고 계약을 변경하지도 않습니다.

© 2021 Amazon Web Services, Inc. 또는 자회사. All rights reserved.