



AWS白皮书

将微软 Power BI 与AWS Cloud



将微软 Power BI 与AWS Cloud: AWS白皮书

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon 的商标和商业外观不得用于任何非 Amazon 的商品或服务，也不得以任何可能引起客户混淆、贬低或诋毁 Amazon 的方式使用。所有非 Amazon 拥有的其他商标均为各自所有者的财产，这些所有者可能附属于 Amazon、与 Amazon 有关联或由 Amazon 赞助，也可能不是如此。

Table of Contents

摘要	i
简介	2
微软 Power BI 套件	3
Power BI 桌面	3
Power BI 服务	3
Power BI 报表服务器	4
On-premises 数据网关	4
将微软 Power BI 桌面连接到 AWS 数据源	5
在本地使用 Power BI 桌面	5
通过互联网连接到数据源	5
通过连接数据源Site-to-Site VPN	8
通过以下方式连接到数据源AWS Direct Connect	11
在 AWS 云中使用微软 Power BI 桌面	14
选项 1：在亚马逊 EC2 实例上安装微软 Power BI 桌面	15
选项 2：在亚马逊 WorkSpaces 环境中安装微软 Power BI	15
选项 3：在亚马逊 AppStream 2.0 环境中安装微软 Power BI	15
微软 Power BI 桌面连接选项摘要	17
将微软 Power BI 服务连接到 AWS 数据源	19
建议的配置	19
其他注意事项	20
使用 Amazon Quick	23
结论	24
贡献者	25
延伸阅读	26
文档修订	27
附录：微软 Power BI 支持的 AWS 数据源	28
Amazon Redshift	28
Amazon RDS	29
Amazon Athena	30
亚马逊 OpenSearch 服务	30
AWS Lake Formation	31
版权声明	33
.....	xxxiv

将微软 Power BI 与 AWS Cloud

发布日期：2021 年 11 月 3 日 ([文档修订](#))

本白皮书讨论了如何将 Microsoft Power BI (桌面、服务和本地数据网关) 与亚马逊网络服务 (AWS) 云集成和使用。它为希望将 Microsoft Power BI 产品与 Amazon Redshift、Amazon Athena、Amazon RDS、Amazon Service 等 AWS 服务连接起来的客户提供了选项，重点是连接、安全、性能 AWS Lake Formation 和成本优化。OpenSearch

本白皮书适用于希望快速理解 Microsoft Power BI 概念的 IT 决策者和架构师，以及在使用 AWS 服务作为数据源时有哪些选项可以利用这些技术。

简介

各种规模企业的客户都在使用AWS产品和服务来可靠、经济实惠且安全地存储数据。这在一定程度上要归功于可用的成熟数据存储和分析产品组成的广泛生态系统。其中一些产品包括以下服务：

- [亚马逊简单存储服务](#) (Amazon S3) 提供了一个简单、可扩展、安全且经济实惠的数据存储库。它已成为存储应用程序数据的行业标准，也是客户数据湖的首选。
- [Amazon Athena](#) 是一项交互式查询服务，可使用标准 SQL 轻松分析 Amazon S3 中的数据。
- [Amazon Relational Database Service](#) (Amazon RDS) 让您可以轻松地在云中设置、操作和扩展关系数据库。它提供了经济实惠且可调整大小的容量，同时自动执行耗时的管理任务，例如硬件配置、数据库设置、修补和备份。SQL Server、Oracle 数据库、MySQL、MariaDB 和 PostgreSQL 引擎可用。
- [Amazon Redshift](#) 是一个完全托管、可大规模扩展的数据仓库，可以轻松分析结构化和非结构化数据集。
- [Amazon Quick](#) 是一项快速、基于云的商业智能服务，可让您轻松地组织中的每个人提供见解。
- [Amazon S OpenSearch er](#) vice 可让您轻松执行交互式日志分析、近乎实时的应用程序监控、网站搜索等。
- [AWS Lake Formation](#) 是一项服务，便于您在短短几天内轻松地设置安全数据湖。

为了更好地了解服务之间的关系，我们经常将数据服务标记为数据源或数据使用者。数据源允许客户和应用程序存储和检索服务中的数据。通常，数据源还具有内置计算功能，可以提供计算分析和筛选。但最终，数据会加载到这些数据源中，最终由数据使用者从中检索数据。亚马逊 S3、亚马逊 Athena 和 Amazon Redshift 就是很好的数据源示例。

另一方面，数据使用者访问数据源中的数据，通常会对其进行处理。他们也可以选择显示它。Amazon Quick 和 Microsoft Power BI 套件就是数据消费者的好例子。他们读取数据源，然后协助分析、可视化和发布信息。

AWS为客户提供了充分的灵活性，让他们可以根据自己的数据需求混合使用他们喜欢的技术。虽然许多客户选择 Amazon Quick 来满足他们的商业智能 (BI) 需求，但其他客户则选择了微软 Power BI、Tableau 和 Qlik 等供应商。

本文档重点介绍 Microsoft Power BI 产品和服务套件，以及如何将它们与AWS服务结合使用。

微软 Power BI 套件

为了减少由于产品命名相似而造成的混乱，本白皮书介绍了每个 Microsoft Power BI 产品和服务的含义。

Power BI 桌面

Power BI Desktop 是一款安装在本地计算机上的免费应用程序。它允许您连接、转换和可视化数据。使用 Power BI Desktop，您可以连接到多个不同的数据源，并将它们合并（通常称为建模）到一个数据模型中。此数据模型允许您构建视觉对象和视觉对象集合，您可以将其作为报告与组织内的其他人共享。

Power BI Desktop 可以连接到本地或通过网络提供的任何支持的数据源。有关支持的数据源，请参阅[附录：Microsoft Power BI 支持的 AWS 数据源](#)。

大多数从事商业智能项目的用户都使用 Power BI Desktop 来创建报告。然后，他们将内容推送到 Power BI 报表服务器或 Power BI 服务，以便与其他人共享报表。将内容从 Power BI Desktop 推送到 Power BI 报表服务器或 Power BI 服务的行为称为发布。有关更多信息，请参阅[什么是 Power BI 桌面？](#)

Note

Power BI 桌面是一款 Windows-only 应用程序。它不适用于 Linux、macOS 或其他操作系统。

Power BI 服务

Power BI 是一系列软件服务、应用程序和连接器，它们协同工作，可帮助您以最有效地为您和您的业务服务的方式创建、共享和使用业务见解。Power BI 服务，有时也称为 Power BI 在线，是 Power BI 中的软件即服务 (SaaS) 部分。有关更多信息，请参阅[什么是 Power BI 服务？](#)

Power BI 服务是一项基于云的服务。它支持团队和组织的简易报告编辑和协作。你也可以连接到 Power BI 服务中的数据源，但建模有限。

大多数从事商业智能项目的报表设计者都使用 Power BI Desktop 来创建报表，然后使用 Power BI 服务与其他人分发报表。有关此关键组件的更多信息，请参阅[将 Microsoft Power BI 服务连接到 AWS 数据源](#)。

Power BI 报表服务器

Power BI 报表服务器 (PBIRS) 是一款私有报表服务器，带有助于显示和管理报表和 KPI 的 Web 门户。当你不想将报表发布到 Power BI 服务时，你可以使用 Power BI 报表服务器。

随着 SQL Server 2025 的发布，PBIRS 已成为 SQL Server 的默认报告解决方案，并且不会发布任何新版本的 SQL Server 报告服务 (SSRS)。这意味着，要部署 SQL Server 2025，你可以在 AWS 上部署 Power BI 报表服务器。这适用于 SQL Server 企业版和标准版，包括许可证部署和 BYOL 部署。对于 2025 年之前的 SQL Server 版本，AWS 不允许使用 PBIRS。相反，请使用 SSRS 或 Power BI 服务 (SaaS)。

在包含许可证的实例上，通过运行 SQL Server 安装程序 `C:\SQLServerSetup\x64\DefaultSetup.ini` 或通过运行 SQL Server 安装程序来检索产品密钥。有关 BYOL 许可的详细信息，包括核心计数要求和故障转移实例规则，请参阅 [Microsoft SQL Server 2025 产品条款](#)。

Note

发布共享报表需要 Power BI Pro 用户订阅许可证。

有关硬件要求，请参阅 [Power BI 报表服务器的系统要求](#)。

有关安装说明，请参阅 [安装 Power BI 报表服务器](#)。

On-premises 数据网关

Microsoft 本地数据网关是一个常用部署的组件，可以提高 Power BI 部署的安全性和性能。它允许 Power BI 服务访问私有化的数据源，这些数据源位于另一个设施中，可通过数据源和数据网关之间的内部网络连接进行访问。尽管它通常作为服务器组件安装，但您也可以将个人模式作为应用程序安装在本地计算机上。本白皮书仅关注标准（服务器）模式。有关更多信息，请参阅 [将微软 Power BI 服务连接到 AWS 数据源](#)。

将微软 Power BI 桌面连接到 AWS 数据源

大多数情况下，开始使用 Microsoft Power BI Desktop 的客户对如何从本地计算机和网络连接到 AWS 数据源感兴趣。桌面应用程序通常在其本地 Windows 笔记本电脑上运行，与 AWS 数据源的物理和逻辑连接是最大的进入障碍。

但是，还有另一种选择，那就是在 AWS 云中运行 Microsoft Power BI 桌面。此选项可显著减少与 AWS 数据源的连接障碍，但仍需要一些额外的注意事项。本节将讨论这两种模型。我们会研究每种方案对连接性、安全性、性能和成本的影响，以便您可以决定哪个选项最适合您。本节中介绍的选项说明了亚马逊 RDS、Amazon Redshift 和亚马逊 Athena。有关所有 AWS 数据源的完整讨论，请参阅[附录：Microsoft Power BI 支持的 AWS 数据源](#)。

在本地使用 Power BI 桌面

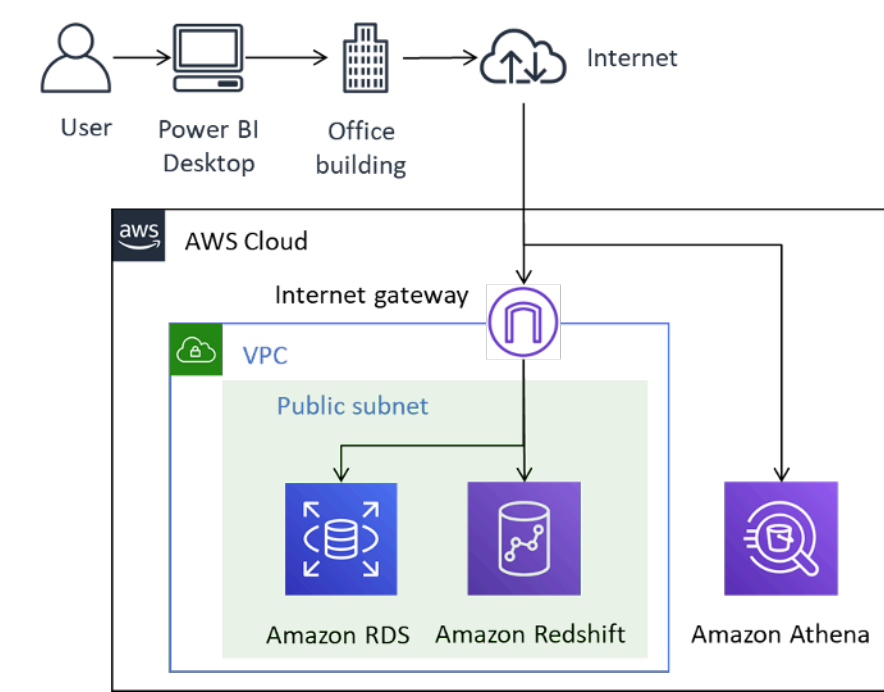
如果您计划在本地使用 Power BI 桌面，数据源存储在 AWS 云中，Power BI 可以通过以下三种方式之一访问这些源：

- 使用互联网连接到数据源。
- 使用AWS Virtual Private Network(Site-to-Site VPN) 连接到数据源。
- 使用连接到数据源Direct Connect。

以下各节将详细介绍每种方法。

通过互联网连接到数据源

在此模型中，Power BI Desktop 应用程序放置了一个出站连接，该连接通过互联网路由到可通过互联网访问的 AWS 数据源的 IP 地址。例如，在客户的亚马逊虚拟私有云 (Amazon Amazon VPC) 中实例化的亚马逊 RDS 和 Amazon Redshift 支持公共可访问性选项，使实例可以通过互联网访问。使用您特定地区的服务终端节点，可以直接从互联网上查询 Amazon Athena。



支持通过互联网与 AWS 数据源的 BI 连接

尽管这种连接方式在技术上是可行的，但除了少数用户之外，我们不建议其他任何方式使用它。下表列出了重要的注意事项。

表 1 — 通过互联网访问 AWS 数据源的注意事项

标准	通过互联网访问 AWS 数据源的注意事项
网络连接	通过连接到 VPC 中的私有 IP 地址，或者使用区域或 VPC 服务终端节点可以访问数据源。Power BI Desktop 通过 VPN 连接，要么直接访问数据源（亚马逊 RDS、Amazon Redshift、EC2-based 亚马逊数据源），要么使用私有 VPC 终端节点或区域终端节点访问具有区域终端节点（Amazon Athena）的服务，具体取决于 DNS 配置。
安全	IP 访问控制 安全组 充当实例的虚拟防火墙以控制入站和出站流量。为了限制对可信实体的访问，请将安全

标准	通过互联网访问 AWS 数据源的注意事项
	组配置为仅允许与已知无类 Inter-Domain 路由 (CIDR) 范围关联的入站 IP 范围。 传输中加密 AWS 建议您为任何使用公有 IP 地址的数据源（例如 Amazon RDS、Amazon Redshift 或任何亚马逊 EC2-based 数据源）配置加密。这样可以确保降低数据或凭证在传输过程中被泄露的风险。未能配置加密会带来重大风险。不要忽视这方面。 区域服务终端节点（例如 Amazon Athena）采用 TLS 加密。此外，流向 JDBC 或 ODBC 客户端的 Amazon Athena 查询结果使用传输层安全 (TLS) 进行加密。 身份验证和授权 AWS 建议您使用提供对数据集的只读访问权限的证书，并根据您的公司政策设置轮换证书的流程。

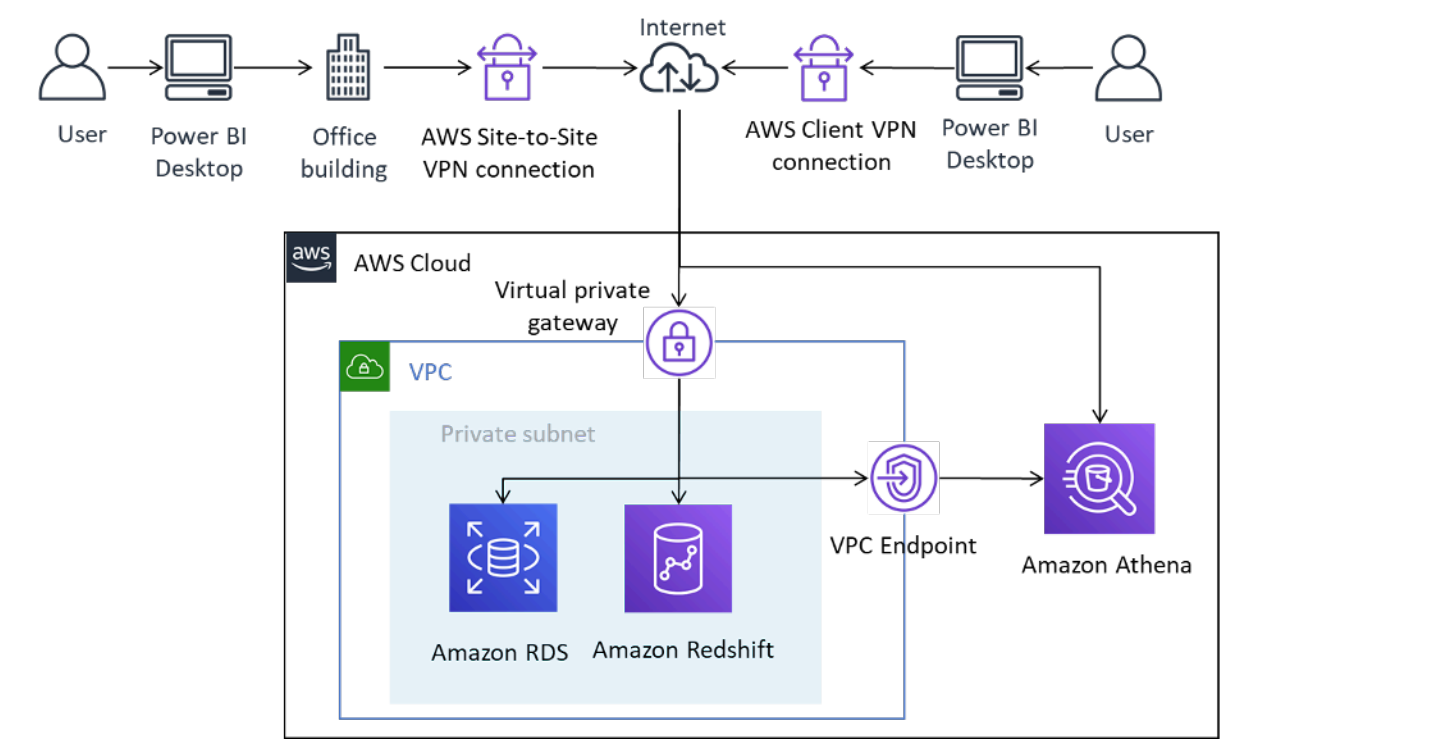
标准	通过互联网访问 AWS 数据源的注意事项
性能	通过互联网访问 AWS 数据源时，可能会影响 Power BI Desktop 整体性能的一些因素包括： - 正在访问的数据集的大小。较大的数据集需要更长的时间才能检索。我们建议限制查询并使用过滤器来减少通过互联网检索的数据量。 - 互联网连接的质量，包括带宽、延迟和丢包。如有可能，请访问地理位置靠近的 AWS 区域的数据，以减少延迟的影响。如果您的互联网是共享的，请考虑在非高峰时段加载数据源，并确保有足够的带宽可用。 通常，AWS 建议在一天中的不同时间、不同的数据集以及用户数量逐渐增加的情况下测试体验。
成本	驻留在 VPC 中并使用公有 IP 地址通过互联网进行查询的数据源会产生标准的 Amazon VPC 数据出站费用。为了降低成本，我们建议限制查询并使用过滤器来减少通过互联网检索的数据量。

通过以下方式连接到数据源Site-to-Site VPN

在此模型中，Power BI Desktop 安装使用以下两种Site-to-Site VPN方法之一连接到 AWS 网络中的数据源：AWS Site-to-Site VPN 或 AWS Client VPN。每种连接类型都提供高度可用、托管且具有弹性的云 VPN 解决方案，以保护您的网络流量。

Site-to-Site VPN 在您的网络和您的 AWS VPN 或 AWS Transit Gateway 之间创建加密隧道。Client VPN 使用免费的 VPN 软件客户端将您的用户连接到 AWS 或本地资源。

来自 VPN 和 Client Site-to-Site VPN 连接的 VPN 流量在您的 VPC 中停止。因此，它可以路由到私有 IP 地址，因此您的实例不再需要面向公众的 IP 地址。对于具有可从面向公众的服务终端节点访问的数据路径的服务（例如 Athena），这些服务请求可以通过 Internet 路由，也可以通过 VPN 连接和 VPC 终端节点路由。



通过 VPN 和 Client Site-to-Site VPN 将 Power BI Desktop 连接到 AWS 数据源

Site-to-Site VPN 还可以连接到 AWS Transit Gateway，从而便于访问分布在多个 VPC 中的数据源。

使用Site-to-Site VPN提供了在访问存储在 AWS 中的数据源时采用加密的好处，而无需明确配置每个数据源。配置完成后，VPN 技术在很大程度上可以与最终用户无缝连接。

表 2 — 使用访问 AWS 数据源的注意事项Site-to-Site VPN

标准	使用访问 AWS 数据源的注意事项Site-to-Site VPN
网络连接	通过连接到 VPC 中的私有 IP 地址，或者使用区域或 VPC 服务终端节点可以访问数据源。Power BI Desktop 通过 VPN 连接，要么直接访问数据源（亚马逊 RDS、Amazon Redshift、亚马逊 EC2-based 数据源），要么使用私有 VPC 终端节点或区域终端节点访问具有区域终端节点（Amazon Athena）的服务，具体取决于 DNS 配置。
安全	IP 访问控制

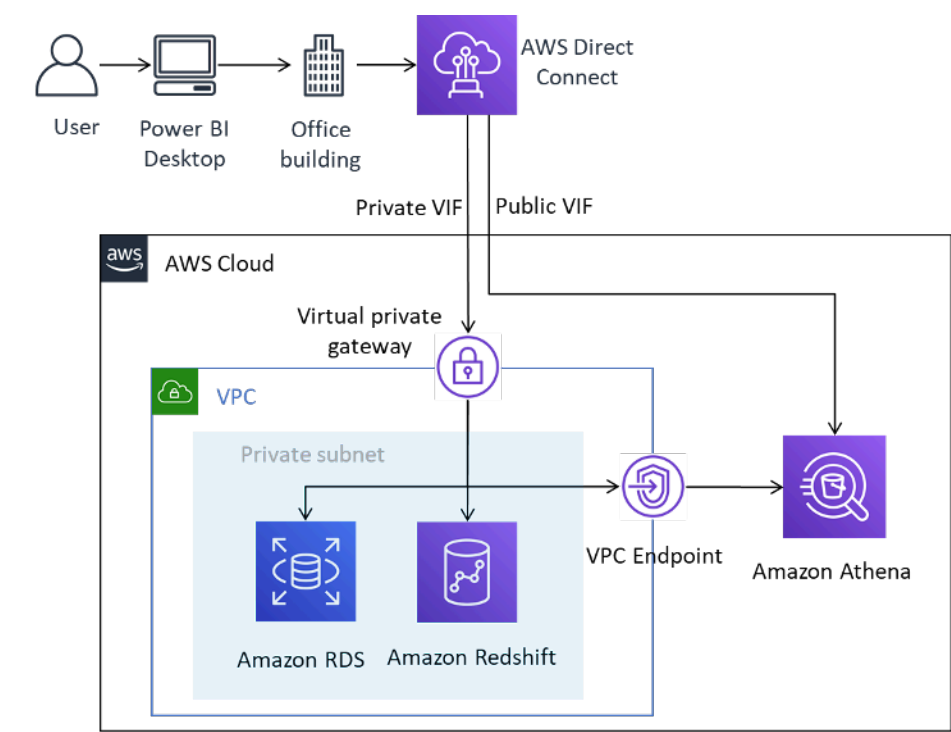
标准	使用访问 AWS 数据源的注意事项Site-to-Site VPN
	您可以组合使用路由组和安全组来控制对存储在 AWS 云中的数据源的访问。 传输中加密 这两种类型都Site-to-Site VPN使用 IPsec 加密，这意味着传输的数据在 AWS 和本地之间传输时会被加密。这样可以确保即使数据源未配置为使用加密通信，也能在穿越互联网时保护数据。 身份验证 Site-to-Site VPN 需要一次性配置，一旦建立，即可与用户无缝连接。最终用户无需进行身份验证即可使用 Site-to-Site VPN，但他们需要对数据源进行身份验证。 另一方面，Client VPN 确实需要最终用户进行身份验证才能建立连接。Client VPN 身份验证可通过 Active Directory (基于用户)、相互身份验证 (基于证书) 或 SAML SSO (基于用户) 进行。通过身份验证后，终端用户即可实现无缝连接。添加到 Power BI 桌面的 AWS 数据源需要身份验证。 AWS 建议您使用仅对所需数据集具有只读访问权限的身份对 AWS 数据源进行身份验证。

标准	使用访问 AWS 数据源的注意事项Site-to-Site VPN
性能	的使用是Site-to-Site VPN通过互联网进行的。因此，其性能极限与提出的第一个场景类似。通过互联网访问 AWS 数据源时，某些因素可能会影响 Power BI Desktop 的整体性能。其中包括： - 正在访问的数据集的大小。较大的数据集需要更长的时间才能检索。我们建议限制查询并使用过滤器来减少通过互联网检索的数据量。 - 互联网连接的质量，包括带宽、延迟和丢包。如有可能，请访问地理位置靠近的 AWS 区域的数据。这可以减少延迟的影响。如果您的互联网是共享的，请考虑在非高峰时段加载数据源， and/or确保有足够的带宽可用。 通常，AWS 建议在一天中的不同时间、不同的数据集以及用户数量逐渐增加的情况下测试体验。
成本	驻留在 VPC 中并使用查询的数据源Site-to-Site VPN会产生标准Site-to-Site VPN的数据传输费用。为了降低成本，我们建议限制查询并使用过滤器来减少通过互联网检索的数据量。

通过以下方式连接到数据源AWS Direct Connect

AWS Direct Connect通过标准的以太网光纤电缆将您的内部网络链接到某个AWS Direct Connect位置。电缆的一端连接到您的路由器，另一端连接到AWS Direct Connect路由器。通过此连接，您可以绕过网络路径中的互联网服务提供商，直接创建通往公有 AWS 服务（例如 Amazon S3）或 Amazon VPC 的虚拟接口。某个AWS Direct Connect位置允许访问与其关联的区域中的 AWS。您可以使用公共区域或 AWS GovCloud（美国）中的单个连接访问所有其他公共区域中的公有 AWS 服务。

在此模型中，客户的本地网络AWS Direct Connect直接连接到 AWS 网络。虽然有多种配置方法AWS Direct Connect，但在最简单的模式下，您可以使用称为私有虚拟接口（私有 VIF）的结构访问 VPC 内的 IP 范围。您可以使用公共虚拟接口（公共 VIF）访问 internet/public IP 范围。



通过 Power BI Desktop 连接到 AWS 数据源AWS Direct Connect

在 Power BI 中添加数据源时，如果私有 IP 地址位于 VPC 中，则可以选择该地址或服务的私有 VPC 终端节点，具体取决于您的 DNS 配置方式。

表 3 — 使用访问 AWS 数据源的注意事项AWS Direct Connect

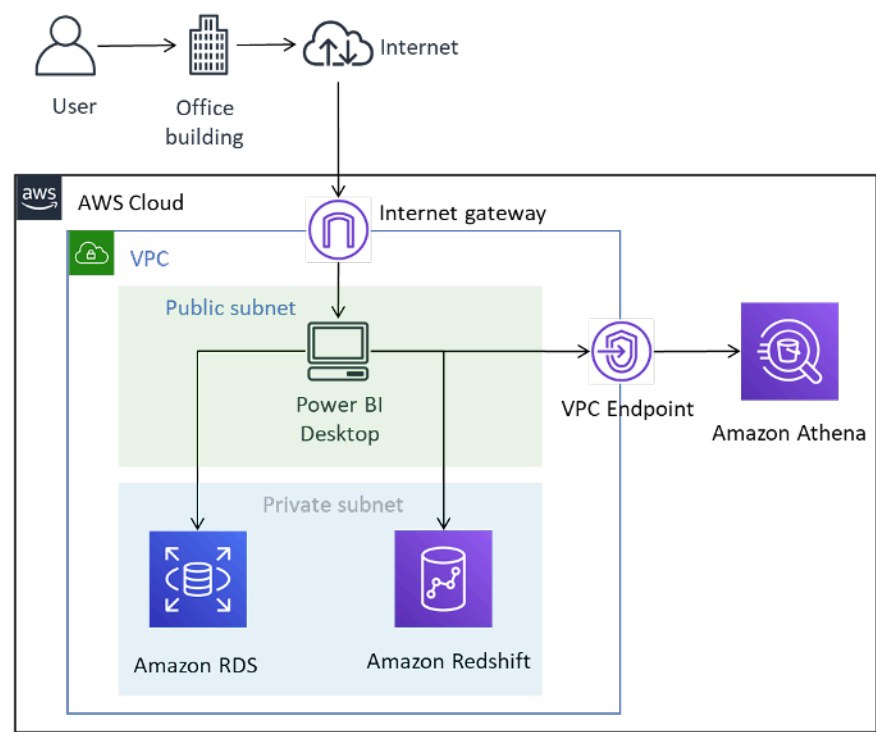
标准	使用 AWS Direct Connect 访问 AWS 数据源的注意事项
网络连接	在您配置 Direct Connect 后，它可以通过连接到 VPC 中的私有 IP 或使用区域服务终端节点来访问数据源。
安全	IP 访问控制 您可以组合使用路由组和安全组来控制对存储在 AWS 云中的数据源的访问。 传输中加密

标准	使用 AWS Direct Connect 访问 AWS 数据源的注意事项 Direct Connect 不提供数据线路级加密。我们建议您在数据源级别启用 TLS 加密，或者将 Site-to-Site VPN 与 Direct Connect 结合使用，以确保数据和凭据在传输过程中不会被泄露。 身份验证 配置完成后，Direct Connect 不需要任何额外的身份验证。 AWS 建议您使用仅对所需数据集具有只读访问权限的身份对 AWS 数据源进行身份验证。
性能	使用 Direct Connect，可以使用 1 Gbps 和 10 Gbps 端口。你可以从任何支持 Direct Connect 的 APN 合作伙伴那里订购 50 Mbps、100 Mbps、200 Mbps、300 Mbps、400 Mbps 和 500 Mbps 的速度。 与基于互联网的连接相比，Direct Connect 连接通常提供更高的带宽和更短的延迟。根据连接的容量，加载数据源时，Power BI Desktop 的性能可能不再受网络的限制。因此，当访问与 Direct Connect 关联的同一个区域中的数据集时，您的用户可以期望查询性能良好。 请注意加载大型数据集，并注意 Power BI Desktop 的数据集限制为 10 GB。

标准	使用 AWS Direct Connect 访问 AWS 数据源的注意事项
成本	除了标准的 Direct Connect 端口费用外，使用 Direct Connect 连接访问的数据源还会产生标准的 Direct Connect 数据传输费用，如 AWS Direct Connect 定价 中所述。通过 Direct Connect 发送到 AWS 云的数据不会产生任何费用。 为了降低成本，AWS 建议限制查询并使用筛选器来减少检索的数据量。

在 AWS 云中使用微软 Power BI 桌面

在 AWS 云中使用 Microsoft Power BI 桌面是应对上一节中描述的许多挑战的热门解决方案。在此模型中，客户在 AWS 云中托管 Microsoft Power BI 桌面，然后在本地对其进行远程访问。下图显示了一个示例。



部署在 AWS 云中的微软 Power BI Desktop

尽管该图描绘了用户通过互联网连接到桌面，Site-to-Site VPN 但 Direct Connect 也是有效的连接类型。由于仅传输图形管理流量，因此带宽要求非常适合典型的互联网连接。

在此模型中，Microsoft Power BI Desktop 托管在亚马逊 VPC 内的公有子网中，并且可以直接通过网络连接到具有私有 IP 地址的数据源，例如亚马逊 RDS 和 Amazon Redshift。您可以使用 VPC 终端节点连接作为目的地（如图所示）或使用区域公共服务终端节点连接到 Amazon Athena 和其他区域服务。

在 AWS 云中托管 Microsoft Power BI Desktop 有多种选择：

选项 1：在亚马逊 EC2 实例上安装微软 Power BI 桌面

在此选项中，您可以创建一个或多个安装了微软 Windows 服务器的 Amazon EC2 实例。您的用户使用远程桌面协议 (RDP) 应用程序连接到实例，并像安装在笔记本电脑上一样安装、配置和使用 Microsoft Power BI。默认情况下，Windows 服务器安装仅限于两个并发的 RDP 客户端会话。如果您需要更多，可以从 Microsoft 经销商处购买客户端访问许可证。您可能会发现，增加单台服务器上的并发用户数会导致资源争用。您可以通过增加实例数量来获得更好的用户体验，每个实例都有更少的用户。

选项 2：在亚马逊 WorkSpaces 环境中安装微软 Power BI

Amazon WorkSpaces 是一种托管、安全 Desktop-as-a-Service (DaaS) 解决方案。您可以使用 Amazon WorkSpaces 在短短几分钟内配置 Windows 或 Linux 桌面，并快速扩展，为全球员工提供数千个桌面。

在此选项中，您可以为用户配置安装微软 Power BI 桌面的微软 Windows 桌面。用户可以使用适用于安卓、iOS、Fire、Mac、PC、Chromebook 和 Linux 设备的轻量级客户端应用程序访问 Windows 桌面。

使用亚马逊 WorkSpaces，Microsoft Power BI Desktop 用户拥有专用资源，并且可以自动停止桌面以节省成本。此选项还具有允许非 Windows 设备访问亚马逊 WorkSpaces，以及为用户提供用于执行其他任务（例如访问微软 Office 或其他 Windows 应用程序）的桌面环境。

选项 3：在亚马逊 AppStream 2.0 环境中安装微软 Power BI

Amazon AppStream 2.0 是一项完全托管的应用程序流媒体服务。您可以在 AppStream 2.0 上集中管理桌面应用程序，并将其安全地交付到任何计算机。无需购买、预置和操作硬件或基础设施，即可轻松扩展到全球任意数量的用户。

在此选项中，您只为本地用户提供流式传输到 HTML5-compliant 网络浏览器的 Microsoft Power BI 桌面应用程序，无需任何插件。与 Amazon WorkSpaces 选项一样，非 Windows 设备可以连接，用户可以访问专用资源。

此选项的一个缺点是，默认情况下，AppStream 2.0 应用程序的存储容量限制为 1 GB，这可能不适合保存更大的模型和报告。通过组合适用于 Windows File Server 的 Amazon FsX 的文件存储空间可以增加容量，但这会带来额外的复杂性，客户应注意。

表 4 — 在 AWS 云中运行 Power BI Desktop 的注意事项

标准	在 AWS 云中运行 Power BI Desktop 的注意事项
网络连接	与数据源的网络连接非常简单，因为数据使用者和数据源都位于 AWS 云中。可以直接访问位于亚马逊 VPC 中的数据源，例如 Amazon RDS 和 Amazon Redshift。使用区域终端节点的数据源可通过 Amazon VPC 互联网网关或 Amazon VPC 终端节点进行访问。 与 Microsoft Power BI Desktop 的连接是通过互联网进行的 Site-to-Site VPN，或 AWS Direct Connect。所有三个选项都有适度的要求，大多数互联网连接都能满足。
安全	IP 访问控制 客户可以使用路由组和安全组的组合来控制对存储在 AWS 云中的数据源的访问。 对于第一个选项（使用 Amazon EC2），您还可以使用路由和安全组的组合，仅允许特定的本地 CIDR 范围。 传输中加密 我们建议将 Amazon VPC 内的数据源配置为使用加密来传输数据。区域服务已经使用 TLS 加密。

标准	在 AWS 云中运行 Power BI Desktop 的注意事项
	所有三个选项的管理访问权限均已加密。 身份验证 AWS 建议您使用仅对所需数据集具有只读访问权限的身份对 AWS 数据源进行身份验证。 所有三个选项都需要使用身份验证机制，以便远程用户在访问托管 Microsoft Power BI Desktop 的环境之前证明自己的身份。这种机制通常是登录凭证，但也提供了 MFA 选项。
性能	在 AWS 云中运行时，Microsoft Power BI Desktop 的性能通常比在本地运行时要好。网络和计算实例都可能是现代化的，并且具有更高的规格，并且应用程序和数据源之间的延迟可能很低。
成本	通过将 Microsoft Power BI Desktop 放入 AWS 云中，可以显著降低或完全取消数据传输费用。在跨可用区、不同 VPC 或不同区域访问数据源时要小心，因为可能会收取数据传输费用。 对于每种选择，都需要考虑其他成本因素。亚马逊 EC2 WorkSpaces、Amazon 和 Amazon AppStream 2.0 是根据使用量定价的。有关更多信息，请参阅每项服务的定价页面。

微软 Power BI 桌面连接选项摘要

对于少数对数据集要求较少的用户来说，在本地运行 Microsoft Power BI Desktop 并通过互联网安全连接或使用Site-to-Site VPN可能是一个合适的解决方案。确保在此模型中配置并维护了安全性。我们还建议测试此配置，以确定它是否符合用户的性能预期

随着用户数量的增加，我们建议您考虑通过连接AWS Direct Connect。在加载更大的数据集时，Direct Connect 可提供更好的用户体验。确保用户了解传输大型数据集的成本影响。

我们建议你评估在 AWS 云中运行 Microsoft Power BI Desktop。这可能会为最终用户提供最佳的性能体验，也为云管理员提供最佳的管理体验。特别是 Amazon WorkSpaces 可以从少量用户扩展到成千上万的用户。这些服务还具有显著的安全和管理优势。

将微软 Power BI 服务连接到 AWS 数据源

微软 Power BI 服务 (SaaS) 可以直接连接到可访问互联网的数据源或亚马逊 VPC 中的私有数据源。连接私有数据源需要一个名为 Microsoft 本地数据网关的应用程序组件。微软本地数据网关下载并安装在 VPC 中的亚马逊 EC2 实例上，并使用微软 Power BI 凭证进行配置。该网关通过互联网与微软 Azure 服务总线建立出站连接，并在 Microsoft Power BI 中配置为连接到其可以访问的数据源。较大规模的部署可以使用多个本地数据网关来平衡负载或提高容错能力。

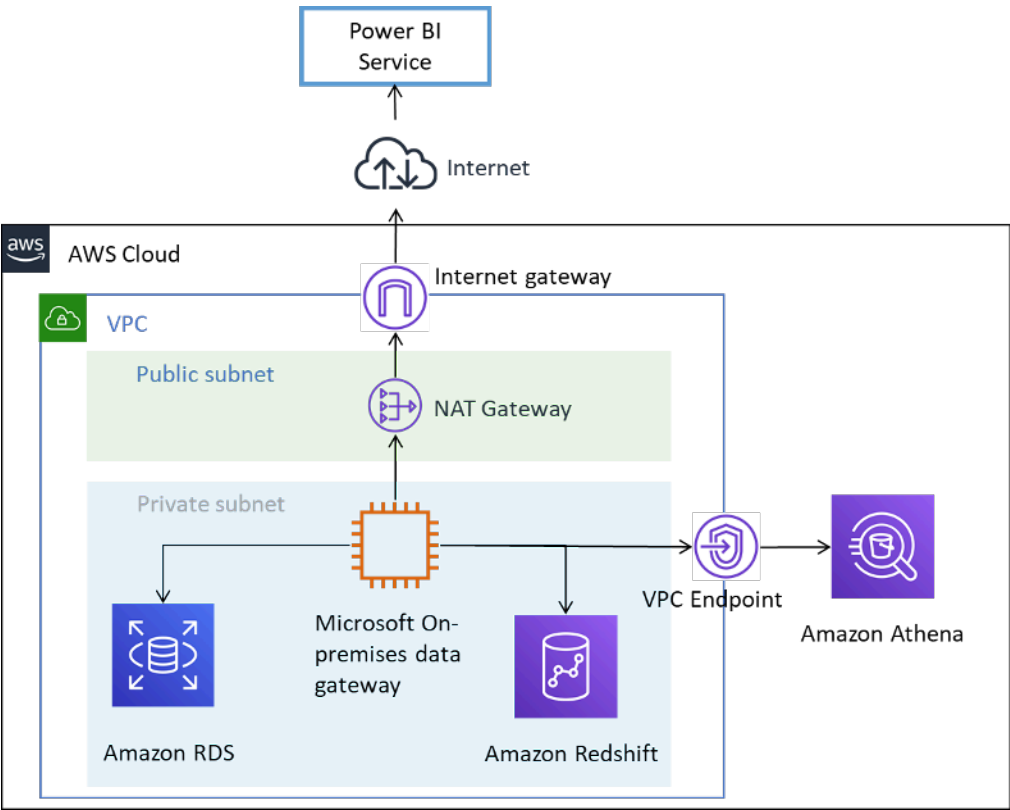
使用 Microsoft 本地数据网关可以带来许多实质性的好处，AWS 客户已经报告了这些好处：

- **改善安全状况：**[微软本地数据网关](#)不接受来自微软 Azure 云的入站连接，只启动与 Azure 服务总线的出站连接。这种单向流量模型允许您将数据源保密，而不会将其暴露在互联网上。
- **减少数据传出：**连接到数据源时，Microsoft 本地数据网关会检索整个结果集，并通过名为后台处理的过程将其存储在本地。在将结果传输到 Power BI 服务之前，会对数据进行压缩。用户通常报告压缩率为 10:1，这不仅缩短了通过互联网传输数据的时间，还减少了出站费用。
- **降低解决方案成本：**使用 Microsoft 本地数据网关时，服务所需的部分数据处理由网关执行。使用 Amazon EC2，再加上 Savings Plans 或预留实例等成本削减计划，可能有助于降低商业智能解决方案的总体成本。

建议的配置

AWS 建议您在包含您的数据源的私有子网中的 Amazon EC2 实例上安装 Microsoft 本地数据网关。该子网配置为通过安装在公有子网中的 [Amazon VPC NAT 网关](#) 将请求路由到互联网。您可以使用网络地址转换 (NAT) 网关使私有子网中的实例能够连接到 Internet 或其他 AWS 服务，但阻止 Internet 连接到这些实例。如果您需要高可用性的数据网关实施，我们建议您使用跨越不同 AWS 可用区的多个 EC2 实例上安装的本地数据网关集群。有关信息，请参阅[添加其他网关以创建集群](#)。

本节中介绍的选项说明了亚马逊 RDS、Amazon Redshift 和亚马逊 Athena。有关所有 AWS 数据源的完整讨论，请参阅[附录：微软 Power BI 支持的 AWS 数据源](#)。



将AWS数据源连接到微软 Power BI 服务

其他注意事项

表 5 — 包含数据源的微软 Power BI 服务的注意事项AWS Cloud

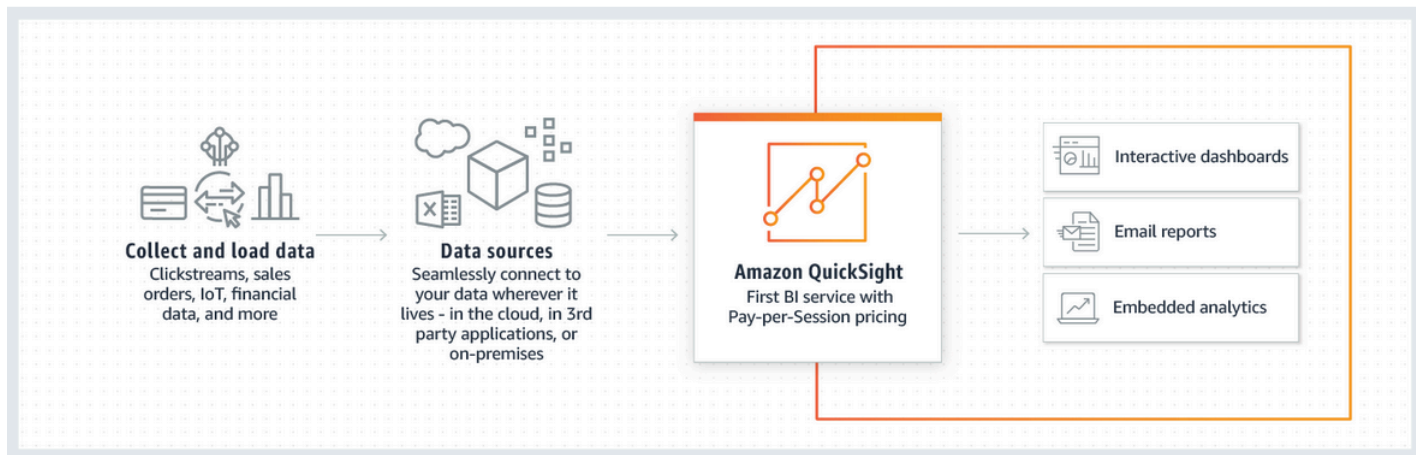
标准	包含数据源的微软 Power BI 服务的注意事项AWS Cloud
网络连接	Microsoft 本地数据网关与数据源的连接非常简单，因为数据使用者和数据源都位于 AWS 云中。可以直接访问位于亚马逊 VPC 中的数据源，例如 Amazon RDS 和 Amazon Redshift。使用区域终端节点的数据源可通过 Amazon VPC 互联网网关或亚马逊 VPC 终端节点进行访问。 微软本地数据网关与 Microsoft Power BI 服务的连接是通过互联网进行的，并且仅是出站连接。

标准	包含数据源的微软 Power BI 服务的注意事项 AWS Cloud
安全	IP 访问控制 您可以结合使用路由组和安全组来控制对存储在 AWS 云中的数据源的访问。 由于 Microsoft 本地数据网关安装在 Amazon EC2 实例上，因此它将有一个关联的安全组，可用于限制对操作系统的入站访问。网关不接受入站请求。该实例不需要公有 IP 地址，也不应配置公有 IP 地址。 传输中加密 我们建议将 Amazon VPC 内的数据源配置为使用加密来传输数据。区域服务已经使用 TLS 加密。 可以将微软本地数据网关连接配置为使用 HTTPS 而不是 TCP 连接到微软 Azure 服务总线。我们建议使用 HTTPS 模式进行通信。这也是自 2019 年 6 月网关软件版本发布以来安装的新网关的默认设置。 身份验证 AWS 建议您使用仅对所需数据集具有只读访问权限的身份对 AWS 数据源进行身份验证。您为数据源输入的凭据经过加密并存储在网关云服务中。凭证在本地网关进行解密。（您为数据源输入的凭据经过加密并存储在网关云服务中。） 确保微软 Power BI 凭据得到安全控制。访问该服务允许访问 AWS 数据源及其可能包含的潜在敏感信息。

标准	包含数据源的微软 Power BI 服务的注意事项 AWS Cloud
性能	由于能够调整和扩展 Amazon EC2 实例，AWS 云中的 Microsoft 本地数据网关通常表现良好。它在区域联网和互联网连接方面也表现得很快。
成本	需要考虑三个因素：Amazon EC2 实例费用、数据传输费用和 Amazon NAT 网关费用。 根据微软的要求调整您的 Amazon EC2 实例的大小。为了降低成本，您可以购买 Amazon EC2 预留实例或 AWS Savings Plans。 从微软本地数据网关传输到微软 BI 服务的数据会产生 VPC 出站费。客户报告使用数据网关进行了 10:1 的压缩，这将减少流量，但我们建议您限制查询并使用过滤器来确保仅传输相关数据。 如果 Microsoft 本地数据网关连接到不同可用区或不同 AWS 区域的数据源，则还会收取数据传输费用。 如果 Microsoft 本地数据网关位于私有子网中并使用 AWS NAT 网关，则按小时收取费用和数据处理费。有关更多信息，请参阅 Amazon VPC pricing。

使用 Amazon Quick

鼓励考虑在 AWS 中使用 Microsoft Power BI 套件的客户评估 [Amazon QuickSight](#) 作为替代方案。这种完全托管的云服务以原生方式连接到 AWS 中的数据源，与其他商业智能解决方案相比，它降低了复杂性和成本。



Amazon Quick 的工作原理

与其他商业智能解决方案相比，Quick具有以下优点：

- 使用 Quick，无需下载和安装客户端应用程序。所有功能，包括创作和报告，都可以通过 Web 浏览器从任何平台（Windows、Mac、Linux 等）进行访问。
- Quick 作为完全托管的云原生 SaaS 应用程序提供，易于构建仪表板并将其部署到生产环境中。该服务是无服务器的，这意味着您无需计算需要多少节点和服务器来支持您的用户。QuickSight还充分利用 AWS 提供的高可用性功能来实现弹性。
- 无论是小型还是大型设置，都可以轻松入门，并且能够通过点击界面添加用户。QuickSight无需外部管理员干预。
- Quick 由 Par Super-fast all-in-memory 计算引擎 (SPICE) 提供支持，可实现快速的响应时间（以毫秒为单位）和交互式可视化。数据集目前可以扩展到 200 GB。
- 快速定价既简单又便宜，有两个组成部分：报告作者和报告阅读器。创建和发布交互式仪表板的报告作者按用户定价。如果用户在给定月份内没有登录，则这些用户无需支付任何费用。报告读者按每 30 分钟的会话收费，每位读者每月最高收取 5.00 美元。免费试用版允许您免费评估 Quick。有关更多信息，请参阅 [Amazon 快速定价](#)。

结论

如果你想使用 Microsoft Power BI Desktop，我们通常会发现客户开始在本地上用该软件，通过互联网连接到数据源。虽然有私有连接选项可供使用Site-to-Site VPN和Direct Connect，但许多客户得出的结论是，在亚马逊上运行 Microsoft Power BI Desktop WorkSpaces 可以提供更好的性能体验。

如果你想将数据源连接AWS到 Microsoft Power BI 服务，那么知道这是一种既定的架构模式，你应该感到放心。你可以在亚马逊 VPC 中安装微软本地数据网关，将亚马逊 RDS、Amazon Redshift、Amazon Athena、OpenSearch 亚马逊服务等数据源无缝连接到该服务。AWS Lake Formation

如果您想要一款既能提供相同业务成果，又不增加安装、配置、修补和扩展自我管理商业智能解决方案的复杂性的解决方案，我们建议使用 Amazon Quick。这项完全托管的服务将所有必需的功能与按用户付费定价相结合，提供简单的 Web 浏览器体验。无需安装，也不需要其他组件。

希望这仅仅是您的商业智能之旅的开始AWS。有关帮助您入门的其他资源，请参阅[附录：微软 Power BI 支持的 AWS 数据源](#)。

贡献者

本文档的贡献者包括：

- Ralph Holm , Amazon Web Services 高级解决方案架构师
- Barret Newman , Amazon Web Services 高级解决方案架构师
- Fabrizio Napolitano , Amazon Web Services 数据库和分析专业解决方案架构师

延伸阅读

- [将 Power BI 与 Amazon Redshift 集成，获取见解和分析](#)
- [在 Power BI Desktop 中连接亚马逊 Redshift 数据库](#)

文档修订

如需获取有关该白皮书更新的通知，请订阅 RSS 信息源。

变更	说明	日期
已更新白皮书	由于 Microsoft 许可限制，已删除报表服务器的内容。	2023 年 8 月 29 日
已更新	增加了数据源：亚马逊 OpenSearch 和 AWS Lake Formation。Athena、Amazon RDS 和 Redshift 数据源的更新。有关 Microsoft 本地数据网关的更多指南。	2021 年 11 月 3 日
已审核	已审查技术准确性	2021 年 3 月 1 日
已更新	更新了报表服务器许可	2021 年 1 月 15 日
初次发布	白皮书首次发布	2020 年 11 月 1 日

附录：微软 Power BI 支持的 AWS 数据源

支持的数据源的完整列表由 Microsoft 提供（请参阅 [Power BI 数据源](#)）；但是，以下各节针对每个 AWS 数据源提供了可能对某些读者有所帮助的使用和配置指南。

Amazon Redshift

Amazon Redshift 是 AWS 云中一项完全托管的 PB 级数据仓库服务。Amazon Redshift 数据仓库是一个由称作节点的各种计算资源构成的集合，这些节点已整理到名为集群的组中。每个集群运行一个 Amazon Redshift 引擎并包含一个或多个数据库。

在以下情况下，你应该考虑使用亚马逊 Redshift：

- 您正在构建或迁移到原生云数据仓库。
- 您可能需要从几太字节扩展到数百 TB。
- 您希望允许 Power BI 用户透明地访问存储在 Amazon S3 中的数据湖中的数据，并将其与数据仓库中的表连接起来。
- 您的查询工作负载包括：
 - 在大型（多 GB 和多 TB）表上计算聚合的查询。
 - 具有多个联接和子查询的极其复杂的 SQL。
 - 仪表板中混合了复杂的分析查询和经过高度筛选的简单查询。

将亚马逊 Redshift 与微软 Power BI 配合使用时，请记住以下几点：

- 微软 Power BI Desktop 和 Power BI 服务原生支持 Amazon Redshift 作为 Power BI 数据源，每种服务都支持导入和直接查询模式。
- 虽然 Redshift 集群可以在公有子网中启动并配置为允许从 Internet 进行访问，但大多数客户更愿意在私有子网中启动它以提高安全性。使用私有子网时，请使用本地数据网关从 Power BI 服务连接到 Amazon Redshift。
- Redshift 连接器支持 Power BI 桌面和服务中的 Azure AD 身份验证。
- 通过 Spectrum 访问的外部表与原生 Redshift 表的处理方式没有任何区别，而且 Power BI 无法区分它们。访问外部表中的数据时，请确保：
 - 包含字符串的列在 AWS Glue 数据目录中被归类为“VARCHAR”，而不是“字符串”，否则 Power BI 将抛出以下错误：Exception: OLE DB or ODBC error: [Expression.Error] We

couldn't fold the expression to the data source. Please try a simpler expression..

- 不支持包含复杂数据类型 (例如 ARRAY) 的列。当使用包含复杂数据类型的列时, Power BI 会抛出以下错误: Exception: ODBC: ERROR [42703] [Microsoft]Amazon Redshift Error occurred while trying to execute a query

如果您需要将它们包含在模型中, 则可以在用户级别启用 (在 Amazon Redshift 中) JSON 序列化, 也可以将复杂数据类型存储在本机表的 SUPER 列中。

Amazon RDS

Amazon RDS 可以轻松地在云中设置、操作和扩展关系数据库。Amazon RDS 适用于多种数据库实例类型 (针对内存、性能或 I/O), 并为您提供六种熟悉的数据库引擎可供选择, 包括亚马逊 Aurora、PostgreSQL、MySQL、MariaDB、Oracle 数据库和 SQL Server。

在以下情况下, 您应该考虑使用 RDS :

- 您正在构建运营数据存储。
- 您正在将 SQL Server 或 Oracle 数据库数据仓库迁移到云端, 但对重构不感兴趣。
- 您的查询工作负载包括:
 - 访问可轻松编制索引的表上经过高度筛选的数据的查询。
 - 对中小型表 (千兆字节) 进行分析查询。
 - 混合了中等复杂度的分析查询和仪表板中使用的简单、经过高度过滤的查询。

将 Amazon RDS 与微软 Power BI 配合使用时, 请记住以下几点 :

- Amazon RDS 提供多个数据库引擎, 包括 SQL Server、MariaDB、MySQL、Oracle 数据库和 PostgreSQL。请注意, 数据库引擎列在 Power BI Desktop 和 Power BI 服务中, 而不是 Amazon RDS 服务中。
- 对于 Amazon Aurora, 使用我的 SQL 或 PostgreSQL 连接类型, 具体取决于您选择的数据库引擎。
- 虽然 Amazon RDS 实例可以在公有子网中启动并配置为允许从 Internet 进行访问, 但大多数客户更愿意在私有子网中启动它以提高安全性。使用私有子网时, 请使用本地数据网关从 Power BI 服务连接到 RDS。
- 使用 Amazon RDS, 您可以部署多个版本的 SQL Server (2012、2014、2016、2017 和 2019 年), 包括 Express、Web、标准版和企业版。

Amazon Athena

Amazon Athena 是一种交互式查询服务，方便使用标准 SQL 分析 Amazon S3 的数据。Athena 与 AWS Glue 数据目录开箱即用，允许您跨各种服务创建统一的元数据存储库，爬取数据源以发现架构，使用新的和修改过的表和分区定义填充数据目录，以及维护架构版本控制。

在以下情况下，您应该将 Athena 视为数据源：

- 你想直接查询你的数据湖。
- 您的查询工作负载包括：
 - 在大型（多 GB 和多 TB）表上计算聚合的查询
 - 交互式临时 SQL，用于探索目的。

将亚马逊 Athena 与 Microsoft Power BI 配合使用时，请记住以下几点：

- 随着 2021 年 7 月微软 Power BI 的发布，亚马逊 Athena 推出了 Microsoft-certified 连接器。您可以使用适用于亚马逊 Athena 的微软 Power BI 连接器在微软 Power BI Desktop 中分析来自亚马逊 Athena 的数据。将内容发布到 Power BI 服务后，您可以使用 Microsoft 本地数据网关，通过按需或计划刷新来保持内容最新。
- 适用于亚马逊 Athena 的 Microsoft Power BI 连接器支持导入和直接查询数据连接模式。在导入模式下，选定的表和列将导入 Power BI Desktop 中进行查询。在直接查询模式下，不会将任何数据导入或复制到 Power BI Desktop 中，而是 Power BI Desktop 直接查询底层数据源。
- 有关适用于亚马逊 Athena 的 Microsoft Power BI 连接器的更多信息，请参阅使用亚马逊 Athena [Power BI 连接器](#)。
- 请注意，适用于亚马逊 Athena 的 Microsoft Power BI 连接器需要在系统上使用亚马逊 Athena ODBC 驱动程序和有效的 ODBC DSN 配置才能查询亚马逊 Athena。要下载最新的 ODBC 驱动程序和配置信息，请参阅使用 ODBC [连接到 Amazon Athena](#)。
- 有关使用适用于亚马逊 Athena 的 Microsoft Power BI 连接器时的配置步骤和最佳实践的教程，请参阅使用亚马逊 Athena [在 Microsoft Power BI 上快速创建](#) 仪表板。

亚马逊 OpenSearch 服务

您可以使用 SQL 来查询您的亚马逊 OpenSearch 服务，而不必使用 JSON-based 搜索查询 DSL。如果您已经熟悉该语言，或者想将您的域与使用该语言的应用程序（例如 Microsoft Power BI）集成，那么使用 SQL 进行查询非常有用。

在以下情况下，您应该考虑将亚马逊 OpenSearch 服务作为数据源：

- 您有半结构化数据，例如日志文件或 JSON 输出，需要快速搜索、分析或可视化这些信息。

将亚马逊 OpenSearch 服务与 Microsoft Power BI 配合使用时，请记住以下几点：

- 连接亚马逊 OpenSearch 服务需要开放数据库连接 (ODBC) 驱动程序，这是一款适用于 Windows 和 macOS 的只读 ODBC 驱动程序，允许您将商业智能 (BI) 和数据可视化应用程序（如 Tableau、Microsoft Excel 和 Power BI）连接到集群上的 SQL 插件。该驱动程序可在“OpenSearch [下载并入门](#)”网站上找到。[有关配置说明，请参阅 ODBC 驱动程序网站中的“自定义 ODBC 驱动程序”部分。OpenSearch](#)
- 目前仅支持导入模式。
- 与亚马逊 OpenSearch 服务的 Power BI 连接目前需要使用测试版连接器。要开始使用，请参阅[微软 Power Query 文档——连接器参考：亚马逊 Opensearch 服务（测试版）](#)。

AWS Lake Formation

Lake Formation 可帮助您收集数据库和对象存储中的数据并对其进行分类，将数据移动到新的 [Amazon S3](#) 数据湖中，使用机器学习算法对数据进行清理和分类，以及安全访问敏感数据。您的用户可以访问集中式[数据目录](#)，[该目录](#)描述了可用数据集及其适当用法。然后，您的用户将这些数据集与他们选择的分析和机器学习服务结合起来，例如[亚马逊 Redshift](#)、[Amazon Athena](#) 和 [（测试版）适用于 Apache Spark 的 Amazon EMR](#)。Lake Formation 建立在 [AWS Glue 中提供的功能之上](#)。

如果您需要对数据湖进行细粒度（行和列）级别的访问权限，而不是传统的基于 IAM 的控件，则应考虑 Lake Formation。

在微软 Power BI 中使用 Lake Formation 时，请记住以下几点：

- 要使用 Power BI Desktop 或 Power BI 服务查询 Lake Formation 数据目录中的数据，请使用与在 Athena 中查询数据相同的流程和配置。如果您使用的是 Lake Formation 权限模型，请确保亚马逊 Athena 的 ODBC DSN 配置的属性密钥 LakeformationEnabled 值设置为 true。这告诉亚马逊 Athena ODBC 驱动程序使用 Lake Formation 服务进行授权，而不是直接使用 AWS Security Token Service 进行授权。有关更多信息，请参阅使用 ODBC [连接至亚马逊 Athena](#) 的文档。
- 为与现有数据目录行为兼容而启用的“仅使用 IAM 访问控制”设置将提供完全兼容性。

- 将 AWS Glue 数据权限升级到 Lake Formation 模型可能会带来不兼容性，应在使用前进行测试。初步测试表明，列级允许或拒绝，但作者尚未测试行级和单元级筛选，因为这仍处于预览阶段，可能会发生变化。

版权声明

客户有责任对本文档中的信息进行单独评测。本文档：(a) 仅供参考；(b) 代表当前提供的 AWS 产品和实操，如有更改，恕不另行通知；并且 (c) AWS 及其附属机构、供应商或许可方不做任何承诺或保证。AWS 产品或服务“按原样”提供，不提供任何形式的保证、陈述或条件，无论是明示还是暗示。AWS 对其客户承担的责任和义务受 AWS 协议制约，本文档不是 AWS 与客户直接协议的一部分，也不构成对该协议的修改。

© 2021 , Amazon Web Services, Inc. 或其附属公司。保留所有权利。

本文属于机器翻译版本。若本译文内容与英语原文存在差异，则一律以英文原文为准。