AWS 白皮书

构建可扩展且安全的多 VPC AWS 网络基础设施

构建可扩展且安全的多 VPC AWS 网络基础设施: AWS 白皮书

Table of Contents

构建可扩展且安全的多 VPC AWS 网络基础设施

发布日期：2020 年 6 月 10 日（文档历史记录）

摘要

AWS 客户通常依赖数百个账户和 VPC 来细分其工作负载并扩展其业务。这种规模级别通常会带来资源共享、VPC 间连接以及本地到 VPC 连接方面的挑战。

本白皮书介绍了使用 Amazon VPC、AWS Transit Gateway、AWS PrivateLink 和 AWS Direct Connect 网关等 AWS 服务在大型网络中创建可扩展且安全的网络架构的最佳实践。它演示了用于管理不断增长的基础设施的解决方案 – 确保可扩展性、高可用性和安全性，同时保持较低的开销成本。

引言

AWS 客户首先在单个 AWS 账户中构建资源，该账户代表细分权限、成本和服务的管理边界。但是，随着客户组织的发展，必须对服务进行更好的细分，以监控成本、控制访问权限和提供更轻松的环境管理。多账户解决方案通过为组织内的 IT 服务和用户提供特定账户来解决这些问题。AWS 提供了多种工具来管理和配置此基础设施，包括 AWS 登录区和 AWS Control Tower。

图 1 – 登录区账户结构

AWS 登录区和 AWS Control Tower 实现了多个 AWS 服务的设置和集成自动化，从而为基准的、高度受控的多账户环境提供了 Identity and Access Management（IAM）、监管、数据安全、网络设计和日志记录。

图 1 中的 AWS 登录区解决方案包括四个账户 – AWS Organizations 账户（用于管理 AWS 登录区托管账户的配置和访问权限）、共享服务账户（用于创建基础设施共享服务，如目录服务）、日志归档账户（集中记录到 S3 存储桶中）和安全账户（供公司的安全性与合规性团队在分支账户中发生事件时用于审核或执行紧急安全操作）。

本白皮书介绍了管理您的 AWS 基础设施的网络团队所拥有的网络服务账户。该账户的网络服务和网络基础设施由所有账户和 VPC 以集中方式共享（类似于中心辐射型设计）。这种设计使您的登录区具有更好的可管理性，并且无需在每个分支 VPC 和账户中重复网络服务，从而有助于降低成本。

> **Note**
>
> 在本白皮书中，“登录区”是指在其中部署工作负载的可扩展、安全且高性能的多账户/多 VPC 设置的广义术语。可以使用任何工具构建此设置。

大多数客户从几个 VPC 开始部署其基础设施。客户拥有的 VPC 数量通常与其账户、用户和暂存环境（生产、开发、测试等）的数量有关。随着云使用量的增加，与客户交互的用户、业务部门、应用程序和区域的数量将大大增加，从而导致创建新的 VPC。

随着 VPC 数量的增加，跨 VPC 管理对于客户云网络的运营而言变得至关重要。本白皮书介绍了跨 VPC 和混合连接的三个特定方面的最佳实践：

- 网络连接 – 大规模互连 VPC 和本地网络。
- 网络安全 – 构建用于访问互联网和终端节点（如 NAT 网关、VPC 终端节点和 AWS PrivateLink）的集中式出口点。

- DNS 管理 – 解析登录区和混合 DNS 中的 DNS。

VPC 到 VPC 的连接

客户可以使用两种不同的 VPC 流模式来设置多 VPC 环境：多对多或中心辐射型。在多对多方法中，每个 VPC 之间的流量在每个 VPC 之间单独管理。在中心辐射型模型中，所有 VPC 间流量都流经中央资源，该资源根据既定规则路由流量。

主题

VPC 对等

连接两个 VPC 的最简单方法是使用 VPC 对等连接。在此设置中，连接可在 VPC 之间实现完全双向连接。 此对等连接用于在 VPC 之间路由流量。 跨账户和 AWS 区域的 VPC 也可对等连接在一起。VPC 对等连接仅对通过连接传输的流量产生费用（不收取每小时基础设施费用）。

VPC 对等连接是点对点连接，不支持中转路由。例如，如果您在 VPC A 与 VPC B 之间以及 VPC A 与 VPC C 之间有 VPC 对等连接，则 VPC B 中的实例无法通过 VPC A 中转来到达 VPC C。要在 VPC B 与 VPC C 之间路由数据包，您需要创建直接 VPC 对等连接。

在规模上，当您拥有数十个到数百个 VPC 时，通过对等连接来互连它们会产生数百个到数千个对等连接的网格，这很难管理和扩展。每个 VPC 的最大对等连接数限制为 125 个。

图 2 – 使用 VPC 对等连接的网络设置

如果您使用 VPC 对等连接，则必须与每个 VPC 建立本地连接（VPN 和/或 Direct Connect）。一个 VPC 中的资源无法使用对等连接的 VPC 的混合连接在本地访问（图 2）。

当一个 VPC 中的资源必须与另一个 VPC 中的资源通信、这两个 VPC 的环境都受到控制和保护并且要连接的 VPC 数量少于 10（以便可以对每个连接进行单独管理）时，最好使用 VPC 对等连接。与其他 VPC 间连接选项相比，VPC 对等连接的总体成本最低。

Transit VPC 解决方案

Transit VPC 可以通过为 VPC 间连接引入中心辐射型设计来解决 VPC 对等连接的一些缺点。在 Transit VPC 网络中，一个中央 VPC（中心 VPC）通过通常利用 IPsec 上 BGP 的 VPN 连接与所有其他 VPC（分支 VPC）连接。中央 VPC 包含运行软件设备的 EC2 实例，这些设备使用 VPN 覆盖将传入流量路由到其目标（图 3）。Transit VPC 对等连接具有以下优势：

- 使用覆盖 VPN 网络启用了中转路由，从而实现了更简单的中心辐射型设计。
- 在中心 Transit VPC 中的 EC2 实例上使用第三方供应商软件时，可以利用供应商的有关高级安全性的功能（第 7 层防火墙/IPS/IDS）。如果客户在本地使用相同的软件，他们将受益于统一的操作/监控体验。

图 3 – 使用 Cisco CSR 的 Transit VPC

Transit VPC 面临着自身的挑战，例如运行虚拟设备的成本较高、每个 VPC 的吞吐量有限（每个 VPN 隧道的最高吞吐量为 1.25Gbps）以及额外的配置和管理开销（客户必须管理 EC2 实例的可用性和冗余）。

中转网关

AWS Transit Gateway 作为完全托管式服务提供了一种用于连接 VPC 和本地网络的中心辐射型设计，无需您预置 Cisco CSR 之类的虚拟设备。不需要 VPN 覆盖，AWS 管理高可用性和可扩展性。

Transit Gateway 使客户能够连接数千个 VPC。您可以将所有混合连接（VPN 和 Direct Connect 连接）附加到单个 Transit Gateway，从而在一个位置整合和控制组织的整个 AWS 路由配置（图 4）。Transit Gateway 使用路由表控制在所有已连接的分支网络之间路由流量的方式。这种中心辐射型模型简化了管理并降低了运营成本，因为 VPC 只连接到 Transit Gateway 即可访问已连接的网络。

图 4 – 使用 AWS Transit Gateway 的中心辐射型设计

Transit Gateway 是一种区域资源，可以连接同一 AWS 区域内的数千个 VPC。您可以为每个区域创建多个 Transit Gateway，但不能对等连接一个 AWS 区域内的 Transit Gateway，您还可以通过单个 Direct Connect 最多连接到三个 Transit Gateway 以实现混合连接。出于这些原因，您应该将架构限制为仅使用一个 Transit Gateway 来连接给定区域中的所有 VPC，并在需要时使用 Transit Gateway 路由表隔离它们。创建多个 Transit Gateway 的有效案例完全用于限制错误配置的影响范围。

将组织的 Transit Gateway 放入其网络服务账户中。这样，管理网络服务账户的网络工程师就可以进行集中管理。使用 AWS Resource Access Manager（RAM）共享 Transit Gateway，以便在同一区域内跨 AWS Organizations 中的多个账户连接 VPC。 借助 AWS RAM，您可以轻松安全地与任何 AWS 账户共享 AWS 资源，或在 AWS Organizations 内共享 AWS 资源。有关更多信息，请参阅在中央账户中自动将 AWS Transit Gateway 挂载与 Transit Gateway 关联博客文章。

主题

Transit Gateway 与 Transit VPC

与 Transit VPC 相比，Transit Gateway 具有许多优势：

- Transit Gateway 消除了维护与数百个 VPC 进行的 VPN 连接的复杂性。
- Transit Gateway 无需管理和扩展基于 EC2 的软件设备。AWS 负责管理路由流量所需的所有资源。
- Transit Gateway 通过提供高度可用且冗余的多可用区基础设施，无需管理高可用性。
- Transit Gateway 将 VPC 间通信的带宽提高到每个可用区的突发速度 50Gbps。
- Transit Gateway 将用户成本简化为简单的每小时传输的 GB 数模型。
- Transit Gateway 通过移除 EC2 代理和对 VPN 封装的需求来减少延迟。

Transit Gateway 与 VPC 对等连接

Transit Gateway 解决了与大规模创建和管理多个 VPC 对等连接有关的复杂性。尽管这使 TGW 成为大多数网络架构的理想默认设置，但由于与 TGW 相比，VPC 对等连接具有以下优势，因此仍然是一个有效的选择：

- 较低的成本 – 使用 VPC 对等连接，您只需支付数据传输费。除了数据传输费外，Transit Gateway 还对每个挂载收取每小时费用。
- 无带宽限制 – 使用 Transit Gateway，每个 VPC 连接的最大带宽（突发）为 50Gbps。VPC 对等连接没有聚合带宽。单个实例的网络性能限制和流量限制（在置放群组内为 10Gbps，否则为 5Gbps）适用于这两个选项。只有 VPC 对等连接支持置放群组。
- 延迟 – 与 VPC 对等连接不同，Transit Gateway 是 VPC 之间的额外跃点。
- 安全组兼容性 – 安全组引用可与区域内 VPC 对等连接配合使用。它目前不适用于 Transit Gateway。

在登录区设置中，VPC 对等连接可以与 Transit Gateway 启用的中心辐射型模型结合使用。

AWS PrivateLink

客户可能希望以仅使用者 VPC 启动与服务提供商 VPC 的连接的方式，将驻留在一个 VPC（服务提供商）中的服务/应用程序私下公开给 AWS 区域内的其他使用者 VPC。例如，您的私有应用程序能够访问服务提供商 API。

要使用 AWS PrivateLink，请为您的 VPC 中的应用程序创建 Network Load Balancer，然后创建指向该负载均衡器的 VPC 终端节点服务配置。然后，服务使用者会为您的服务创建接口终端节点。此操作将在您的子网中创建一个带有私有 IP 地址的弹性网络接口，用作发送到服务的流量的入口点。使用者和服务不必位于同一 VPC 中。如果 VPC 不同，则使用者和服务提供商 VPC 的 IP 地址范围可能重叠。除了创建接口 VPC 终端节点以访问其他 VPC 中的服务之外，您还可以创建接口 VPC 终端节点，以便通过 AWS PrivateLink 私下访问受支持的 AWS 服务（图 5）。

图 5 – AWS PrivateLink

Transit Gateway、VPC 对等连接和 AWS PrivateLink 之间的选择取决于连接性。

AWS PrivateLink – 如果您设置了客户端/服务器，希望允许一个或多个使用者 VPC 对服务提供商 VPC 中的特定服务或实例集进行单向访问，请使用 AWS PrivateLink。只有使用者 VPC 中的客户端才能启动与服务提供商 VPC 中的服务的连接。当两个 VPC 中的客户端和服务器的 IP 地址重叠时，这也是一个不错的选项，因为 AWS PrivateLink 利用客户端 VPC 中的 ENI，从而不会与服务提供商发生 IP 冲突。您可以通过 VPC 对等连接、VPN 和 AWS Direct Connect 访问 AWS PrivateLink 终端节点。

VPC 对等连接和 Transit Gateway – 如果要在 VPC 之间启用第 3 层 IP 连接，请使用 VPC 对等连接和 Transit Gateway。

您的架构将包含这些技术的组合，以满足不同的使用案例。所有这些服务都可以相互组合和运行。例如，AWS PrivateLink 处理 API 样式的客户端-服务器连接；VPC 对等连接处理区域内可能仍需要置放群组或需要区域间连接的直接连接要求；Transit Gateway 可简化大规模 VPC 的连接以及混合连接的边缘整合。

Amazon VPC 共享

当团队之间的网络隔离不需要由 VPC 拥有者严格管理，但账户级别的用户和权限必须严格管理时，共享 VPC 非常有用。通过共享 VPC，多个 AWS 账户可以在集中管理的共享 Amazon VPC 中创建其应用程序资源（如 Amazon EC2 实例）。在此模型中，拥有 VPC 的账户（拥有者）与其他账户（参与

者）共享一个或多个子网。共享子网之后，参与者可以查看、创建、修改和删除与他们共享的子网中的应用程序资源。参与者无法查看、修改或删除属于其他参与者或 VPC 拥有者的资源。使用安全组和子网网络 ACL 管理共享 VPC 中资源之间的安全性。

VPC 共享的好处：

- 简化了设计 – VPC 间连接没有复杂性
- 减少了托管 VPC
- 网络团队和应用程序拥有者之间的职责划分
- 提高了 IPv4 地址利用率
- 降低了成本 – 属于同一可用区内不同账户的实例之间无数据传输费

注意：当您与多个账户共享子网时，您的参与者应该有一定程度的合作，因为他们共享 IP 空间和网络资源。如有必要，您可以选择为每个参与者账户共享不同的子网。每个参与者一个子网使网络 ACL 能够提供除安全组之外的网络隔离。

大多数客户架构将包含多个 VPC，其中许多 VPC 将与两个或更多账户共享。Transit Gateway 和 VPC 对等连接可用于连接共享 VPC。例如，假设您有 10 个应用程序。每个应用程序都需要自己的 AWS 账户。这些应用程序可以分为两个应用程序组合（同一组合内的应用程序具有相似的联网要求，应用程序 1-5 在“营销”中，应用程序 6-10 在“销售”中）。

每个应用程序组合可以有一个 VPC（总共两个 VPC），并且与该组合内的不同应用程序拥有者账户共享 VPC。应用程序拥有者将应用程序部署到各自的共享 VPC 中（在本例中，在不同的子网中使用 NACL 进行网络路由分段和隔离）。两个共享 VPC 通过 Transit Gateway 进行连接。通过此设置，您可以从必须连接 10 个 VPC 变为仅连接 2 个 VPC（图 6）。

图 6 – 示例设置 – 共享 VPC

> **Note**
>
> VPC 共享参与者无法在共享子网中创建所有 AWS 资源。有关更多信息，请参阅 Amazon VPC 限制。

混合连接

本节重点介绍如何将云资源与本地数据中心安全地连接起来。有两种方法可以实现混合连接：

1. 一对一连接 – 在此设置中，将为每个 VPC 创建 VPN 连接和/或 Direct Connect 私有 VIF。这是通过利用虚拟私有网关（VGW，virtual private gateway）实现的。此选项非常适合少量 VPC，但随着客户扩展 VPC，管理每个 VPC 的混合连接可能会变得困难。
2. 边缘整合 – 在此设置中，客户在单个终端节点上整合多个 VPC 的混合 IT 连接。所有 VPC 共享这些混合连接。这是通过利用 AWS Transit Gateway 和 Direct Connect 网关实现的。

主题

VPN

图 7 – AWS VPN 终止选项

有三种方法可以将 VPN 设置为 AWS：

1. 在 Transit Gateway 上整合 VPN 连接 – 此选项利用 Transit Gateway 上的 Transit Gateway VPN 挂载。Transit Gateway 支持 Site-to-Site VPN 的 IPsec 终止。客户可以创建通往 Transit Gateway 的 VPN 隧道，并可以访问附加到 Transit Gateway 的 VPC。Transit Gateway 支持静态和基于 BGP 的动态 VPN 连接。Transit Gateway 还支持 VPN 挂载上的等价多路径（ECMP，Equal-Cost Multi-Path）。每个 VPN 连接的最大吞吐量为 1.25Gbps，启用 ECMP 可让您跨 VPN 连接聚合吞吐量。在此选项中，您需要支付 Transit Gateway 定价以及 AWS VPN 定价。我们建议对 VPN 连接使用此选项。有关更多信息，请参阅 AWS VPN 概述。
2. 在 EC2 实例上终止 VPN – 如果客户需要特定的供应商软件功能集（如 Cisco DMVPN 或 GRE），或者希望在各种 VPN 部署之间保持操作一致性，则在边缘情况下使用此选项。您可以利用 Transit VPC 设计进行边缘整合，但请务必记住，Transit VPC 的 VPC 到 VPC 连接部分中的所有关键注意事项都适用于混合 VPN 连接。您负责管理高可用性，并需支付 EC2 实例费用以及任何供应商软件许可费用。
3. 在虚拟私有网关（VGW，virtual private gateway）上终止 VPN – 此选项支持一对一连接设计，您可以为每个 VPC 创建一个 VPN 连接（包含一对冗余 VPN 隧道）。这是开始将 VPN 连接到 AWS

的好方法，但随着您扩展 VPC 的数量，利用 Transit Gateway 的边缘整合设计最终会是更好的选项。VPC 的 VPN 吞吐量限制为 1.25Gbps，并且不支持 ECMP 负载均衡。从定价的角度来看，您只需支付 AWS VPN 定价，运行 VGW 无需付费。有关更多信息，请参阅 AWS VPN 定价和虚拟私有网关上的 AWS VPN。

Direct Connect

虽然互联网上的 VPN 是入门的绝佳选择，但互联网连接对于生产流量而言可能并不可靠。由于这种不可靠性，许多客户选择 AWS Direct Connect，它可以在客户的数据中心和 AWS 之间实现一致的、低延迟、高带宽的专用光纤连接。利用 AWS Direct Connect 连接到 VPC 的方法有四种：

图 8 – 将本地数据中心连接到登录区的四种方法

- 创建到附加到 VPC 的 VGW 的私有虚拟接口（VIF，virtual interface）– 您可以为每个 Direct Connect 连接创建 50 个 VIF，从而最多可以连接到 50 个 VPC（一个 VIF 提供与一个 VPC 的连接）。每个 VPC 有一个 BGP 对等连接。此设置中的连接仅限于 Direct Connect 位置所在的 AWS 区域。VIF 到 VPC 的一对一映射（以及缺少全局访问权限）使这成为在登录区中访问 VPC 的最不受欢迎的方法。
- 创建到与多个 VGW 关联的 Direct Connect 网关的私有 VIF（每个 VGW 都附加到一个 VPC）– Direct Connect 网关可以在任何 AWS 账户中连接到全球（中国除外）最多 10 个 VGW。如果登录区由少量 VPC（不超过 10 个 VPC）组成，并且/或者您需要全局访问权限，则这是一个不错的选择。每个 Direct Connect 连接的每个 Direct Connect 网关都有一个 BGP 对等连接。Direct Connect 网关仅适用于北/南流量，不允许 VPC 到 VPC 连接。
- 创建到与 Transit Gateway 关联的 Direct Connect 网关的 Transit VIF – 您可以通过运行速度为 1Gbps 或更高的专用或托管 Direct Connect 连接将 Transit Gateway 关联到 Direct Connect 网关。此选项允许您通过一个 VIF 和 BGP 对等连接将本地数据中心连接到多达三个跨不同 AWS 区域和 AWS 账户的 Transit Gateway（可连接到数千个 VPC）。这是用于大规模连接多个 VPC 的四个选项中最简单的设置，但您应该注意 Transit Gateway 限制。 一个关键限制是，您只能通过 Transit VIF 将 20 个 CIDR 范围从 Transit Gateway 公布到本地路由器。使用选项 1 和 2，您需要支付 Direct Connect 定价。对于选项 3，您还需要支付 Transit Gateway 挂载和数据传输费用。有关更多信息，请参阅 Direct Connect 上的 Transit Gateway 关联文档。
- 通过 Direct Connect 公有 VIF 创建到 Transit Gateway 的 VPN 连接 – 公有虚拟接口允许您使用公有 IP 地址访问所有 AWS 公有服务和终端节点。当您在 Transit Gateway 上创建 VPN 挂载时，将获得两个用于在 AWS 端终止 VPN 的公有 IP 地址。这些公有 IP 可通过公有 VIF 访问。您可以根据需要通过公有 VIF 创建到任意数量的 Transit Gateway 的任意数量的 VPN 连接。当您通过公有 VIF 创建

BGP 对等连接时，AWS 会将整个 AWS 公有 IP 范围公布到您的路由器。为了确保您只允许某些流量（例如，只允许流向 VPN 终止终端节点的流量），建议您在本地使用防火墙。此选项可用于在网络层对 Direct Connect 进行加密。

虽然第三个选项（到 Direct Connect 网关的 Transit VIF）似乎是最好的选项，因为它允许您为每个 Direct Connect 连接使用单个 BGP 会话来在单点（Transit Gateway）上整合给定 AWS 区域的所有本地连接，但考虑到选项 3 的一些限制和注意事项，我们希望客户同时利用选项 2 和选项 3 以满足其登录区的连接要求。图 9 说明了一个示例设置，其中将 Transit VIF 用作连接到 VPC 的默认方法，并将私有 VIF 用于必须将大量数据从本地 DC 传输到媒体 VPC 的边缘使用案例。私有 VIF 用于避免 Transit Gateway 数据传输费。作为最佳实践，您应该在两个不同的 Direct Connect 位置至少有两个连接，以实现最大冗余，总共四个连接。您为每个连接创建一个 VIF，总共四个私有 VIF 和四个 Transit VIF。您还创建一个 VPN 作为到 AWS Direct Connect 连接的备份连接。

图 9 – 混合连接的示例参考架构

使用网络服务账户创建 Direct Connect 资源，从而划分网络管理边界。 Direct Connect 连接、Direct Connect 网关和 Transit Gateway 都可以驻留在网络服务账户中。要与您的登录区共享 AWS Direct Connect 连接，只需通过 RAM 与其他账户共享 Transit Gateway 即可。

集中式互联网出口

在登录区中部署应用程序时，许多应用程序将需要仅出站互联网访问（例如，下载库/补丁/操作系统更新）。使用网络地址转换（NAT，network address translation）网关，或者使用 EC2 实例（配置了源 NAT（SNAT，Source NAT））作为所有出口互联网访问的下一跃点，可以更好地实现这一点。内部应用程序驻留在私有子网中，而 NAT 网关/EC2 NAT 实例驻留在公有子网中。

使用 NAT 网关

在每个分支 VPC 中部署 NAT 网关的成本可能很高，因为您需要为部署的每个 NAT 网关按小时支付费用（请参阅 Amazon VPC 定价），因此集中化 NAT 网关可能是一个可行的选择。为了实现集中化，我们在网络服务账户中创建一个出口 VPC，然后利用 Transit Gateway 通过位于此 VPC 中的 NAT 网关路由来自分支 VPC 的所有出口流量，如图 10 所示。

注意：与在每个 VPC 中运行一个 NAT 网关的分散式方法相比，使用 Transit Gateway 集中化 NAT 网关时，您需要支付额外的 Transit Gateway 数据处理费用。在某些极端情况下，当您通过一个 VPC 中的 NAT 网关发送大量数据时，将 NAT 保留在该 VPC 的本地以避免 Transit Gateway 数据处理费用可能是更具成本效益的选择。

图 10 – 使用 Transit Gateway 的集中式 NAT 网关（概述）

图 11 – 使用 Transit Gateway 的集中式 NAT 网关（路由表设计）

在此设置中，分支 VPC 挂载与路由表 1（RT1）关联并传播到路由表 2（RT2）。我们已显式添加了黑洞路由，以禁止两个 VPC 相互通信。如果要允许 VPC 间通信，可以从 RT1 中删除“10.0.0.0/8 -> Blackhole”路由条目。这使它们能够通过 NAT 网关进行通信。您还可以将分支 VPC 挂载传播到 RT1（或者，您可以使用一个路由表并将所有内容关联/传播到该路由表），从而使用 Transit Gateway 在 VPC 之间实现直接流量。

我们在 RT1 中添加一个将所有流量指向出口 VPC 的静态路由。由于此静态路由，Transit Gateway 会通过其 ENI 在出口 VPC 中发送所有互联网流量。在出口 VPC 中之后，流量将遵循这些 Transit Gateway ENI 所在的子网路由表中定义的规则。我们在此子网路由表中添加一个将所有流量指向 NAT 网关的路由。NAT 网关子网路由表具有互联网网关（IGW，internet gateway）作为下一跃点。要使回程流量回流，您必须在 NAT 网关子网路由表中添加一个将所有分支 VPC 绑定的流量指向作为下一跃点的 Transit Gateway 的静态路由表条目。

高可用性

要获得高可用性，您应该使用两个 NAT 网关（每个可用区中一个）。在可用区内，NAT 网关的可用性 SLA 为 99.9%。AWS 根据 SLA 协议处理针对可用区内组件故障的冗余问题。当 NAT 网关在可用区中不可用时，会丢弃 0.1% 时间内的流量。如果一个可用区完全失败，则该可用区中的 Transit Gateway 终端节点和 NAT 网关将失败，并且所有流量都将通过另一个可用区中的 Transit Gateway 和 NAT 网关终端节点传输。

安全性

您依赖源实例上的安全组、Transit Gateway 路由表中的黑洞路由，以及 NAT 网关所在子网的网络 ACL。

可扩展性

对于每个唯一目的地，NAT 网关最多可以支持 55000 个并发连接。从吞吐量的角度来看，您受到 NAT 网关性能限制的限制。Transit Gateway 不是负载均衡器，不会在多个可用区中的 NAT 网关之间均匀分配流量。如果可能，通过 Transit Gateway 的流量将停留在可用区内。如果发起流量的 EC2 实例位于可用区 1 中，则流量将从出口 VPC 中同一可用区 1 中的 Transit Gateway 弹性网络接口流出，并根据弹性网络接口所在的子网路由表流向下一跃点。有关规则的完整列表，请参阅 NAT 网关规则和限制。

有关更多信息，请参阅使用 AWS Transit Gateway 从多个 VPC 创建单个互联网出口点博客文章。

使用 EC2 实例实现集中式出站

使用来自 AWS Marketplace 的基于软件的防火墙设备（在 EC2 上）作为出口点类似于 NAT 网关设置。如果您想利用各种供应商产品的第 7 层防火墙/入侵防护/检测系统（IPS/IDS，Intrusion Prevention/Detection System）功能，则可以使用此选项。

在图 12 中，我们将 NAT 网关替换为 EC2 实例（在 EC2 实例上启用了 SNAT）。使用此选项需要考虑几个关键因素：

高可用性

在此设置中，您负责监控 EC2 实例、检测故障以及将 EC2 实例替换为备份/备用实例。大多数 AWS 供应商都为其在此设置中部署的软件预先构建了自动化功能。该自动化可以控制以下操作：

- 检测主 EC2-1 实例的故障
- 更改路由表“路由表 Egx 1”，以便在主实例出现故障时将所有流量指向备份 EC2-2 实例。对于可用区 2 中的子网，也必须执行此操作。

图 12 – 使用 EC2 实例和 Transit Gateway 的集中式 NAT

可扩展性

Transit Gateway 不是负载均衡器，不会在两个可用区中的实例之间均匀分配流量。如果可能，通过 Transit Gateway 的流量将停留在可用区内。您受到单个 EC2 实例的带宽容量的限制。随着使用量的增加，您可以纵向扩展此 EC2 实例。

如果您为出口流量检查选择的供应商不支持故障检测自动化，或者您需要横向扩展，则可以使用替代设计。在此设计（图 13）中，我们不在 Transit Gateway 上为出口 VPC 创建 VPC 挂载，而是创建一个 IPsec VPN 挂载，然后创建从 Transit Gateway 到 EC2 实例的 IPsec VPN，利用 BGP 交换路由。

优点

- 由 BGP 处理故障检测和流量的重新路由。无需 VPC 子网路由表自动化。
- BGP ECMP 可用于跨多个 EC2 实例对流量进行负载均衡 – 可以进行横向扩展。

图 13 – 使用 EC2 实例和 Transit Gateway VPN 的集中式 NAT

关键考虑因素

- EC2 实例上的 VPN 管理开销
- Transit Gateway 上的带宽限制为每个 VPN 隧道 1.25Gbps。借助 ECMP，Transit Gateway 可以支持高达 50Gbps 的总 VPN 带宽。供应商设备的 VPN 和数据包处理能力可能是一个限制因素。
- 此设计假设 FW EC2 实例对入站和出站流量使用相同的弹性网络接口。
- 如果您跨多个 EC2 实例启用流量的 ECMP 负载均衡，则必须在 EC2 实例上对流量执行 SNAT 操作才能保证返回流的对称性，这意味着目标不会知道真正的来源。

VPC 到 VPC 和本地到 VPC 流量的集中式网络安全

AWS 提供安全组和子网 NACL，以便在您的登录区内实现网络安全。这些是第 4 层防火墙。在某些情况下，客户可能希望在其登录区内实施第 7 层防火墙/IPS/IDS，以检查 VPC 之间或本地数据中心与 VPC 之间传输的流量。这可以通过使用 Transit Gateway 和运行在 EC2 实例上的第三方软件设备来实现。使用图 14 中的架构，我们可以使 VPC 到 VPC 和本地到 VPC 流量能够通过 EC2 实例传输。该设置与我们在图 12 中讨论过的类似，但我们另外删除了路由表 1 中的黑洞路由以允许暂存 VPC 流量，并将 VPN 挂载和/或 Direct Connect GW 挂载附加到路由表 1 以允许混合流量。这样，来自分支的所有流量都能够先传输到出口 VPC，然后再发送到目标。您需要出口 VPC 子网路由表（防火墙 EC2 设备位于其中）中的静态路由来在流量检查后通过 Transit Gateway 发送发往分支 VPC 和本地 CIDR 的流量。

> **Note**
>
> 路由信息不会从 Transit Gateway 动态传播到子网路由表中，必须静态输入。子网路由表上有 50 个静态路由的软限制。

图 14 – VPC 到 VPC 和 VPC 到本地部署流量控制

将流量发送到 EC2 实例进行嵌入式检查时的关键注意事项：

- 额外的 Transit Gateway 数据处理费用
- 流量必须经过两个附加的跃点（EC2 实例和 Transit Gateway）
- 可能出现带宽和性能瓶颈
- 维护、管理和扩展 EC2 实例的额外复杂性：
  - 检测故障并故障转移到备用实例
  - 跟踪使用情况并进行横向/纵向扩展
  - 防火墙配置、补丁管理
  - 负载均衡时流量的源网络地址转换（SNAT，Source Network Address Translation），用于保证流量对称

您应该选择通过这些 EC2 实例传递哪些流量。一种处理方法是定义安全区域并检查不受信任区域之间的流量。不受信任区域可以是第三方管理的远程站点、您无法控制/信任的供应商 VPC 或者沙盒/开发

VPC，与您的环境的其他部分相比，其安全框架比较宽松。图 15 支持受信任网络之间的直接流量，同时使用嵌入式 EC2 实例检查往返不受信任网络的流量。我们在此示例中创建了三个区域：

- 不受信任区域 – 这适用于来自“VPN 到远程不受信任站点”或第三方供应商 VPC 的任何流量。
- 生产区域 – 这包含来自生产 VPC 和本地客户 DC 的流量。
- 开发区域 – 这包含来自两个开发 VPC 的流量。

以下是我们为跨区域通信定义的示例规则：

1. 不受信任区域与生产区域 – 不允许通信
2. 生产区域与开发区域 – 允许通过出口 VPC 中的 EC2 FW 设备进行通信
3. 不受信任区域与开发区域 – 允许通过出口 VPC 中的 EC2 FW 设备进行通信
4. 生产区域与生产区域以及开发区域与开发区域 – 通过 Transit Gateway 直接通信

这是一个有三个安全区域的设置，但您可以有更多安全区域。您可以使用多个路由表和黑洞路由来实现安全隔离和最佳流量。选择合适的区域取决于您的整体登录区设计策略（账户结构、VPC 设计）。您可以使用区域来实现业务单元、应用程序、环境等之间的隔离。

在此示例中，我们在 Transit Gateway 上终止不受信任的远程 VPN，并将所有流量发送到 EC2 上的软件 FW 设备进行检查。或者，您可以直接在 EC2 实例上终止这些 VPN，而不是在 Transit Gateway 上终止这些 VPN。通过这种方法，不受信任的 VPN 流量永远不会与 Transit Gateway 直接交互。流量中的跃点数减少了 1，这样可以节省 AWS VPN 成本。要启用动态路由交换（让 Transit Gateway 通过 BGP 了解远程 VPN 的 CIDR），防火墙实例应通过 VPN 连接到 Transit Gateway。在本机 TGW 挂载模型中，您必须在 VPN CIDE 的 TGW 路由表中添加静态路由，并将下一跃点作为出口/安全 VPC。在我们的设置（图 15）中，我们为所有流量提供了到出口 VPC 的默认路由，因此不必显式添加任何特定的静态路由。通过这种方法，您从完全托管式 Transit Gateway VPN 终止终端节点迁移到自我管理的 EC2 实例，从而增加 VPN 管理开销以及 EC2 实例在计算和内存方面的额外负载。

图 15 – 通过使用 Transit Gateway 并定义安全区域来进行流量隔离

DNS

当您在非默认 VPC 中启动实例时，AWS 会根据您为 VPC 指定的 DNS 属性以及您的实例是否具有公有 IPv4 地址来为实例提供私有 DNS 主机名（并可能提供公有 DNS 主机名）。 当“enableDnsSupport”属性设置为 true 时，您会从 Route 53 Resolver 获得 VPC 内的 DNS 解析（与 VPC CIDR 的 IP 偏移为 2）。默认情况下，Route 53 Resolver 应答 VPC 域名的 DNS 查询，例如 EC2 实例或 Elastic Load Balancing 负载均衡器的域名。借助 VPC 对等连接，一个 VPC 中的主机可以将对等 VPC 中实例的公有 DNS 主机名解析为私有 IP 地址，前提是启用了这样做的选项。这同样适用于通过 AWS Transit Gateway 连接的 VPC。有关更多信息，请参阅“实现对 VPC 对等连接的 DNS 解析支持”。

如果要将实例映射到自定义域名，可以使用 Amazon Route 53 创建自定义 DNS 到 IP 映射记录。Amazon Route 53 托管区域是一个容器，其中包含有关您希望 Amazon Route 53 如何响应域及其子域的 DNS 查询的信息。公有托管区域包含可通过公有互联网解析的 DNS 信息，而私有托管区域是一种特定实现，仅向已附加到特定私有托管区域的 VPC 提供信息。在您拥有多个 VPC/账户的登录区设置中，您可以将单个私有托管区域与跨 AWS 账户和跨区域的多个 VPC 关联起来。VPC 中的终端主机使用各自的 Route 53 Resolver IP（与 VPC CIDR 的偏移为 2）作为 DNS 查询的名称服务器。VPC 中的 Route 53 Resolver 仅接受来自 VPC 内资源的 DNS 查询。

混合 DNS

在 AWS 登录区设置和本地资源之间协调 DNS 解析是混合网络中最关键的部分之一。实施混合环境的客户通常已经有了 DNS 解析系统，他们想要一个能与当前系统协同工作的 DNS 解决方案。当您将 AWS 区域中的 VPC 的 DNS 与您的网络的 DNS 集成时，您需要一个 Route 53 Resolver 入站终端节点（用于要转发到 VPC 的 DNS 查询）和一个 Route 53 Resolver 出站终端节点（用于要从您的 VPC 转发到您的网络的查询）。 如图 16 所示，您可以配置出站解析程序终端节点，以将其从您 VPC 中的 EC2 实例接收的查询转发到您的网络上的 DNS 服务器。要将选定的查询从 VPC 转发到本地，请创建 Route 53 Resolver 规则，以指定要转发的 DNS 查询的域名（例如 example.com），以及网络上您希望将查询转发到的 DNS 解析程序的 IP 地址。对于从本地到 Route 53 托管区域的入站查询，您网络上的 DNS 服务器可以将查询转发到指定 VPC 中的入站解析程序终端节点。

图 16 – 使用 Route 53 Resolver 的混合 DNS 解析

这样，您的本地 DNS 解析程序就可以轻松解析 AWS 资源（例如 EC2 实例或与该 VPC 关联的 Route 53 私有托管区域中的记录）的域名。

不建议在登录区的每个 VPC 中创建 Route 53 Resolver 终端节点。将它们集中在中央出口 VPC 中（在网络服务账户中）。这种方法可以在保持较低成本的同时实现更好的可管理性（您为创建的每个入站/出站终端节点按小时付费）。您与登录区的其余部分共享集中式入站和出站终端节点。

出站解析 – 使用网络服务账户编写解析程序规则（根据哪些 DNS 查询将转发到本地 DNS 服务器）。使用 Resource Access Manager（RAM），与多个账户共享这些 Route 53 Resolver 规则（并与账户中的 VPC 关联）。分支 VPC 中的 EC2 实例可以将 DNS 查询发送到 Route 53 Resolver，而 Route 53 Resolver 服务会通过出口 VPC 中的出站 Route 53 Resolver 终端节点将这些查询转发到本地 DNS 服务器。 您无需将分支 VPC 与出口 VPC 进行对等连接，也无需通过 Transit Gateway 连接它们。请勿将出站解析程序终端节点的 IP 用作分支 VPC 中的主 DNS。分支 VPC 应在其 VPC 中使用 Route 53 Resolver（以使用 VPC CIDR 偏移）。

图 17 – 将 Route 53 Resolver 终端节点集中在出口 VPC 中

入站 DNS 解析 – 在集中式 VPC 中创建 Route 53 Resolver 入站终端节点，并将登录区中的所有私有托管区域与此集中式 VPC 关联。有关更多信息，请参阅将多个 VPC 与一个私有托管区域关联。与一个 VPC 关联的多个私有托管区域（PHZ，Private Hosted Zone）不能重叠。如图 17 所示，PHZ 与集中式 VPC 的这种关联将使本地服务器能够使用集中式 VPC 中的入站终端节点为任何私有托管区域（与中央 VPC 关联）中的任何条目解析 DNS。 有关混合 DNS 设置的更多信息，请参阅使用 Amazon Route 53 和 AWS Transit Gateway 对混合云进行集中式 DNS 管理和 Amazon VPC 的混合云 DNS 选项。

集中访问 VPC 私有终端节点

VPC 终端节点允许您私下将 VPC 连接到受支持的 AWS 服务，而无需互联网网关或 NAT 设备。使用此接口终端节点，VPC 中的实例无需公有 IP 地址即可与 AWS 服务终端节点进行通信。VPC 与其他服务之间的流量不会离开 AWS 网络骨干。当前可以预置两种类型的终端节点：接口终端节点（由 AWS PrivateLink 提供支持）和网关终端节点。网关终端节点可随意预置，而且没有强大的集中化使用案例。

接口 VPC 终端节点

接口终端节点包含一个或多个具有私有 IP 地址的弹性网络接口，用作发送到受支持 AWS 服务的流量的入口点。当您预置接口终端节点时，用户需要为该终端节点运行的每个小时支付费用。默认情况下，您将在要从中访问 AWS 服务的每个 VPC 中创建一个接口终端节点。在登录区设置中，客户希望跨多个 VPC 与特定 AWS 服务进行交互，这可能非常昂贵且难以管理。为避免这种情况，您可以在一个集中式 VPC 中托管接口终端节点。所有分支 VPC 都将使用这些集中式终端节点。

当您为 AWS 服务创建 VPC 终端节点时，可以启用私有 DNS。启用后，该设置将创建一个 AWS 托管的 Route 53 私有托管区域（PHZ，private hosted zone），从而能够将公有 AWS 服务终端节点解析为接口终端节点的私有 IP。托管式 PHZ 只能在具有接口终端节点的 VPC 内工作。在我们的设置中，当我们希望分支 VPC 能够解析托管在集中式 VPC 中的 VPC 终端节点 DNS 时，托管式 PHZ 将无法正常工作。要解决此问题，请禁用在创建接口终端节点时自动创建私有 DNS 的选项。或者，您可以手动创建 Route 53 PHZ 并添加别名记录，其完整 AWS 服务终端节点名称指向接口终端节点，如图 18 所示。

图 18 – 手动创建的 PHZ

我们将此私有托管区域与登录区内的其他 VPC 关联。此配置允许分支 VPC 将全方位服务终端节点名称解析为集中式 VPC 中的接口终端节点。

> **Note**
>
> 要访问共享的私有托管区域，分支 VPC 中的主机应使用其 VPC 的 Route 53 Resolver IP。接口终端节点还可通过 VPN 和 Direct Connect 从本地网络进行访问。使用条件转发规则将全方位服务终端节点名称的所有 DNS 流量发送到 Route 53 Resolver 入站终端节点，后者将根据私有托管区域解析 DNS 请求。

在图 19 中，Transit Gateway 支持从分支 VPC 到集中式接口终端节点的流量。在网络服务账户中为它创建 VPC 终端节点和私有托管区域，并与分支账户中的分支 VPC 共享它。有关与其他 VPC 共享终端节点信息的更多详细信息，请参阅将 AWS Transit Gateway 与 AWS PrivateLink 和 Amazon Route 53 Resolver 集成博客文章。

注意：分布式 VPC 终端节点方法（即每个 VPC 一个终端节点）允许您对 VPC 终端节点应用最低权限策略。在集中式方法中，您将对单个终端节点上的所有分支 VPC 访问应用和管理策略。随着 VPC 数量的增加，使用单个策略文档维护最低权限的复杂性可能会增加。单个策略文档还会导致更大的影响范围。策略文档的大小也受到限制（20480 个字符）。

图 19 – 集中化接口 VPC 终端节点

总结

随着您扩展 AWS 的使用量并在 AWS 登录区中部署应用程序，VPC 和联网组件的数量会增加。本白皮书介绍了我们如何管理此不断增长的基础设施，从而确保可扩展性、高可用性和安全性，同时保持低成本。在利用 Transit Gateway、共享 VPC、AWS Direct Connect、VPC 终端节点和第三方软件设备等服务时做出正确的设计决策变得至关重要。请务必了解每种方法的关键考虑因素，并根据您的要求进行反向研究，并分析哪个选项或选项组合最适合您。

贡献者

以下是对此文档做出贡献的个人：

- Amazon Web Services 解决方案架构师 Sidhartha Chauhan
- Amazon Web Services 云基础设施架构师 Amir Abu-akeel
- Amazon Web Services 解决方案架构师 Sohaib Tahir

文档历史记录

要获得有关此白皮书的更新通知，请订阅 RSS 源。

更新-历史记录-更改	更新-历史记录-描述	更新-历史记录-日期
次要更新	更新了 Transit Gateway 与 VPC 对等连接部分。	2021 年 4 月 2 日
更新了白皮书	更正了文本，以与图 7 中所示的选项匹配。	2020 年 6 月 10 日
次要更新	更正了文本，以与图 7 中所示的选项匹配。	2020 年 6 月 10 日
初次发布	发布了白皮书。	2019 年 11 月 15 日

声明

客户有责任对本文档中的信息，进行独立评估。本文档：(a) 仅供参考；(b) 代表 AWS 现有的产品服务和实践，如有变更，恕不另行通知；以及 (c) 不构成 AWS 及其联属公司、供应商或授权商的任何承诺或保证。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 对其客户的责任和义务由 AWS 协议决定，本文档与 AWS 和客户之间签订的任何协议无关，亦不影响任何此类协议。