BGP / MPLS IP 虚拟专用网（VPN）

RFC 4364

陶志豪

zhihao.tao@outlook.com

https://github.com/netwiki/share-doc

目录

备忘录状态 3
版权声明 3
概述 3
1. 介绍 4
1.1 虚拟专用网络 4
1.2 客户边缘/CE 和提供商边缘/PE 5
1.3 具有重叠地址空间的 VPNs 6
1.4 到同一系统具有不同路由的 VPNs 6
1.5 SP 骨干路由器 6
1.6 安全 7
2. 站点和 CEs 7
3. VRFs：PE 中的多个转发表 7
3.1 VRF 和接入链路 7
3.2 将 IP 数据包与 VRF 关联 8
3.3 填充 VRF 9
4. 通过 BGP 进行 VPN 路由分发 9
4.1 VPN-IPv4 地址族 10
4.2 路由识别器的编码 10
4.3 控制路由分发 11
4.3.1 路由目标属性 12
4.3.2 通过 BGP 分发 PEs 间的路由 13
4.3.3 使用路由反射器 14
4.3.4 如何在 BGP 中携带 VPN-IPv4 NLRI 16
4.3.5 使用路由目标构建 VPN 16
4.3.6 单个 PE 中 VRFs 之间的路由分发 16
5. 转发 16
6. 维护 VPN 的正确隔离 18
7. PE 如何学习 CE 的路由 19
8. CEs 如何学习 PEs 的路由 20
9. 运营商的运营商 21

10. 多 AS 骨干网.....22
11. 从 VPN 接入互联网.....23
12. 管理 VPN.....24
13. 安全注意事项.....25
13.1 数据平面.....25
13.2 控制面.....26
13.3 P 和 PE 设备的安全性.....26
14. 服务质量.....26
15. 可扩展性.....26

Network Working Group
Request for Comments: 4364
Obsoletes: 2547
Category: Standards Track

E. Rosen
Cisco Systems, Inc.
Y. Rekhter
Juniper Networks, Inc.

February 2006

Thank Zhihao Tao for your hard work in Translation. The translator spent countless nights and weekends, using his hard work to make it convenient for everyone.
If you have any questions, please send a email to zhihao.tao@outlook.com

BGP / MPLS IP 虚拟专用网（VPN）

备忘录状态

本文档为互联网社区规定的互联网标准化协议，并请讨论和建议来改进。

请参考当前版本的“互联网官方协议标准”（STD 1）和该协议的状态。此备忘录的传播不受限制。

版权声明

概述

本文档描述了服务提供商可以使用 IP 骨干网为其客户提供 IP 虚拟专用网（VPN）的方法。该方法使用“peer 模型”，其中，客户边缘路由器（CE 路由器）将路由发送到服务提供商边缘路由器（PE 路由器）；不同站点的 CE 路由器相互不对等，所以 VPN 的路由算法看不到“重叠”。数据包通过骨干区隧道，使核心路由器不需要知道 VPN 路由。

本文档作废 RFC 2547。

1. 介绍

本文档描述了服务提供商可以使用 IP 骨干网为其客户提供 IP 虚拟专用网（VPN）的方法。该方法使用“peer 模型”，其中，客户边缘路由器（CE 路由器）将路由发送到服务提供商边缘路由器（PE 路由器）；边界网关协议（BGP）[BGP，BGP-MP]被服务提供商用来交换 PE 路由器所连接的那个 VPN 间特定的 VPN 路由。这样做是为了确保来自不同 VPN 的路由仍然是分开的，即使两个 VPN 都有重叠地址空间。PE 路由器分发从特定 VPN 中其它 CE 路由器收到的路由，给此 VPN 中的 CE 路由器。不同站点的 CE 路由器相互不对等，所以 VPN 的路由算法看不到“重叠”。术语“IP VPN”中的“IP”用于指示 PE 从 CE 接收 IP 数据报文，检查他们的 IP 头，并相应地为其寻路。

VPN 内的每个路由都被分配了多协议标签交换（MPLS）[MPLS-ARCH，MPLS-BGP，MPLS-ENCAPS]标签；当 BGP 分发时 VPN 路由，也为该路由分配 MPLS 标签。在客户数据包通过服务提供商骨干区之前，它被封装在与到达数据包目的地址最匹配路由，所对应的客户 VPN 的 MPLS 标签中。此 MPLS 报文进一步封装（例如，使用另一个 MPLS 标签或 IP 或 Generic 路由封装（GRE）隧道报头[MPLS-in-IP-GRE]）以便它通过骨干网隧道到适当的 PE 路由器。从而，骨干核心路由器不需要知道 VPN 路由。

这种方法的主要目标是为了支持客户端从服务提供商或与、保持合同关系的服务提供商们那里获取 IP 骨干服务。客户可能是一个企业、一组需要外部网络的企业、互联网服务提供商、应用服务提供商、另一个使用相同方提供 VPNs 给它自己客户的 VPN 服务提供商等。该方法使其客户能够非常简单使用骨干服务。

对服务提供商是可扩展的和灵活的，并允许服务提供商提升价值。

1.1 虚拟专用网络

考虑一组连接到公共网络称之为“骨干”的“站点”集。现在应用一些策略来创建该集合的一些子集，并施加以下规则：只要两个站点的子集中至少各有一个包含在骨干区域内，则这两个站点可能有骨干网的 IP 连通性。

这些子集是虚拟专用网（VPN）。两个站点仅在各有一些 VPN 包含在公共骨干网的情况下，两个站点才有公用的骨干网 IP 连通性。两个站点在公共骨干网上没有 VPN，两个站点没有骨干网连通性。

如果 VPN 中的所有站点都由同一企业拥有，则 VPN 可能被认为是企业的“内部网”。如果在 VPN 中的各个站点由不同的企业所有，VPN 可能被认为是作为“外联网”。一个站点可以在多个 VPN 中；例如，在内部网和几个外联网。一般来说，当我们使用“VPN”这个术语时，我们不会区分内部网和外联网。

我们将站点的所有者称为“客户”。我们把骨干的业主/运营商称作为“服务提供商”（SPs）。客户从 SP 获取“VPN 服务”。

一个客户可能是一个企业，一组企业，一个互联网服务提供商，应用服务提供商，另一个为自己的客户提供同样的 VPN 服务的 SP，等等。

客户策略决定特定的站点集合是否是一个 VPN。有些客户会想这些策略的实施完全是 SP 的责任。其他客户可能想与 SP 分担负责实现这些策略。这个文件描述了可用于实施这

些策略的机制。我们描述了普通的机制，足以允许由 SP 单独或由 VPN 客户与 SP 一起实现这些策略。然而，大部分讨论集中在前一种情况。

本文档中讨论的机制允许广泛策略的实现。例如，在给定的 VPN 中，可以允许每个站点都有直连路由到其他每个站点（“全连接”）。或者，可以强制某对网站之间的流量通过第三个网站进行寻路。这可能是有用的，例如，如果希望一对站点之间的流量通过防火墙传输，防火墙位于第三个站点。

在这个文档中，我们将限制我们的讨论在这种情况下，客户明确地从 SP 或一组已经同意合作提供 VPN 服务的 SP 购买 VPN 服务。也就是说，客户不仅从一个 SP 购买互联网接入服务，并且该 VPN 流量不通过一个随机互联的 SP 网络集合。

我们也将讨论限于这种情况，为客户提供 IP 骨干网服务，而不是帧中继，异步传输模式（ATM），以太网，高级数据链路控制（HDLC）或点到点协议（PPP）等这些二层服务。客户可以通过他们其中一个（或其他）二层服务连接到骨干网，但第二层服务是终止于骨干网的“边缘”，客户的 IP 数据报移除任何 2 层封装。

在本介绍的其余部分，我们描述了 VPN 应该有的一些属性。本文档的其余部分描述了一组可以部署来提供具有全部功能属性的 VPN 模型的机制。本节还介绍了一些在文件的其余部分使用技术术语。

1.2 客户边缘/CE 和提供商边缘/PE

路由器可以彼此相互连接，或者连接到末端系统，通过各种不同的方式：PPP 连接，ATM 虚拟电路（VCs），帧中继 VCs，以太网接口，以太网接口上的虚拟局域网（VLAN），GRE 隧道，2 层隧道协议（L2TP）隧道，IPsec 隧道等。我们将使用术语“接入链路”指一些通常连接到路由器手段。接入链路可能通常被认为是“数据链接”这样的连接，或者可能是某种隧道；重要的是通过接入链路两个设备有可能成为网络层 peer。

每个 VPN 站点必须包含一个或多个客户端（CE）设备。每个 CE 设备通过某种接入链路连接到一个或多个提供商边缘（PE）路由器。

SP 的网络中没有连接到 CE 设备的路由器是被称为“P 路由器”。

CE 设备可以是主机或路由器。在一个典型的情况下，一个网站包含一个或多个路由器，其中一些连接到 PE 路由器。然后连接到 PE 路由器的站点路由器作为 CE 设备或“CE 路由器”。但是，没有什么方式可以防止非路由主机直接连接到 PE 路由器，这种情况下主机将会成为 CE 设备。

有时，物理上连接到 PE 路由器的是 2 层交换机。在这种情况下，我们不能说二层交换机是 CE 设备。相反，通过 2 层交换机与 PE 路由器通信的主机和路由器是 CE 设备；2 层设施是透明的。如果是二层设施提供多点服务，然后可以有多个 CE 设备通过相同的接入链路连接到 PE 路由器。

在逻辑上 CE 设备是客户 VPN 的一部分。在逻辑上 PE 和 P 路由器是 SP 的网络的一部分。

当 CE 到 PE 的数据包在接入链路上传输时，被称为该报文的“入接入链路”，并且 PE 作为数据包的“入 PE”。当 PE 到 CE 的数据包在接入链路上传输时，被称为该报文的“出接入链路”，并且 PE 作为数据包的“出 PE”。

如果 VPN 的站点中的 CE 设备连接到 PE 路由器，我们会说 PE 路由器是连接到一个特定的 VPN。同样，我们会说一个 PE 路由器连接到一个特定的站点，如果该站点中的 CE 设备连接到 PE 路由器。

当 CE 设备是路由器时，它是其连接的 PE 的路由 peer，但它不是其它站点的 CE 路由器路由 peer。不同站点的路由器不直接交换相互的路由信息；其实他们甚至不需要相互了解。因此，客户没有骨干网或“虚拟骨干网”的管理，并且不必处理与任何站点间路由问题。换句话说，在本文档中描述的这个方案中，VPN 不“覆盖”SP 的网络。

关于边缘设备的管理，需要明确 SP 与其客户之间保持的管理界限。管理目的是客户不需要访问 PE 或 P 路由器，SP 设备不需要访问 CE 设备。

1.3 具有重叠地址空间的 VPNs

如果两个 VPN 没有共同的站点，那么它们可能有重叠地址空间。也就是说，VPN V1 中可能会使用给定的地址是系统 S1 的地址，但在 VPN V2 中使用完全不同的系统 S2 作为地址。VPN 都使用 RFC 1918 私有地址空间时这是一个常见的情况。当然在里面每个 VPN，每个地址必须明确。

即使两个具有共同站点的 VPN 也可有重叠地址空间，只要具有这样地址的系统与公共站点中的系统之间不需要存在任何通信即可。

1.4 到同一系统具有不同路由的 VPNs

虽然站点可能在多个 VPN 中，但不需要该站点中到给定系统的路由在所有 VPN 中应该是相同的。假设我们有一个由 A，B 和 C 组成内部网，和由 A，B，C 和“外部”站点 D 组成的外部网.假设在 A 站有一个服务器，我们希望 B，C 或 D 的客户能够使用这个服务器。也假设在站点 B 有一个防火墙。我们想要所有从站点 D 到服务器的流量都通过防火墙，使来自外联网的流量可以被访问控制。但是，我们不希望使用 C 的流量通过防火墙到服务器的方式，因为这是内部流量。

可以设置两条到服务器的路由。一条路由被站点 B 和 C 使用，将流量直接传送到站点 A.第二条路由，由站点 D 使用，采取流量改为到站点 B 防火墙.如果防火墙允许流量通过，则显示来自站点 B 的流量，并且遵循到站点 A 的路由。

1.5 SP 骨干路由器

SP 的骨干网由 PE 路由器以及其他不连接到 CE 设备的路由器（“P 路由器”）组成。

如果 SP 的骨干网中的每个路由器都必须维护 SP 支持的所有 VPN 的路由信息，将会有严重的可扩展性问题；支持站点的数量可能被单个路由器中可以保存的路由信息量限制。因此，重要的是关于特定的 VPN 的路由信息仅需要存在于连接到该 VPN 的 PE 路由器中。特别是 P 路由器不需要任何 VPN 的任何路由信息。（在组播路由被考虑这种情况可能需要稍微放松一些。这篇文章不被进一步考虑，但是在[VPN-MCAST]中进行了检查。）

所以就像 VPN 所有者没有骨干网或“虚拟骨干”管理权，SPs 本身没有为一个单独的骨干网或“虚拟骨干”管理每个 VPN。骨干中的站到站路由是最优的（在的限制之内用于形成 VPN 的策略），并不受任何通过人造“虚拟拓扑”隧道的限制。

第 10 节讨论了当骨干跨越了几个服务提供商时出现的一些特殊问题。

1.6 安全

这里讨论的 VPN，即使没有使用加密安全措施，旨在提供一个相当于 2 层骨干网安全级别（例如，帧中继）。也就是说，在没有配置错误或故意互连不同的 VPN 的情况下，一个 VPN 中的系统是不可能访问另一个 VPN 中的系统的。当然，这里描述的方法本身不加密隐私的数据，也不提供确定数据在途中是否被篡改的方式。如果需要，必须另外采取加密措施。（例如，[MPLS / BGP-IPsec]）。安全性在第 13 节中有更详细的讨论。

2. 站点和 CEs

从特定的骨干网络的角度来看，如果一套 IP 系统不需要使用骨干网就具有相互 IP 互连性，则该系统可以被视为“站点”。一般来说，一个站点将包含一组地理接近的系统。但是，这并不绝对。如果两个地理位置通过专线连接，在该线路上开放最短路径优先（OSPF）协议[OSPFv2]正在运行，如果是那样两个地点之间通信的首选方式是线路（line），而且这两个地点可以被看作是一个单一的站点，即使每一个地点有自己的 CE 路由器。（这个“站点”的概念是指拓扑而不是地理。如果专线线路 DOWN，或其它导致不再是首选路线，但两个地点仍可以通过使用 VPN 骨干网继续进行通信，这时一个站点已经成为两个。）

CE 设备总是被认为是在一个单一的站点（尽管我们将在第 3.2 节中看到，一个站点可能由多个“虚拟站点”组成）。然而，站点可能属于多个 VPN。

PE 路由器可以连接任何数量的不同站点的 CE 设备，无论这些 CE 设备是在相同还是不同的 VPNs。出于健壮性的考虑，CE 设备可以连接到多个相同或不同的服务提供商的 PE 路由器。如果 CE 设备是路由器，PE 路由器和 CE 路由器将显示为路由器彼此相邻。

虽然我们主要以“站点”为基本互连单元，在互连控制上没有任何措施可以防止细粒化。例如，在一个站点的某些系统可能是一个内部网的成员以及一个或更多的外部网的成员，而同一站点的其他系统可能是仅是内部网的成员。不过可能要求这个站点到骨干网有两个接入链路，一个用于内部网，一个用于外部网；它可能进一步要求在外部接入链路上应用防火墙功能。

3. VRFs：PE 中的多个转发表

每个 PE 路由器维护多个单独的转发表。一个转发表是“默认转发表”。其他的是“VPN 路由和转发表”或“VRF”。

3.1 VRF 和接入链路

通过配置，每个 PE/CE 的接入链路与一个或多个 VRF 关联。与 VRF 相关的接入线路被称为“VRF 接入链路”。

在最简单的情况和最典型的情况是，PE/CE 的接入链路与正好一个 VRF 相关联。当通过特定的接入链路接收 IP 包，在其相关联的 VRF 对目的地 IP 进行查找。其查找结果确定如何对数据包寻路。数据包入 PE 上用于特定数据包寻路的 VRF 被称为数据包的“入 VRF”。（还有一个数据包“出 VRF”的概念，位于数据包的出 PE；在第 5 节讨论。）

如果从接入链路收到的 IP 数据包不与任何 VRF 相关联，在默认转发表中查找数据包的目的地址，并且数据包被相应的寻路。根据默认转发表转发包括来自相邻 P 或 PE 路由器的数据包，以及来自面向客户的尚未与 VRF 相关联的接入链路的数据包。

直观地，一方面认为将默认转发表视为包含“公共路由”，VRF 视为包含“私有路由”。另一方面也可以将 VRF 接入链路视为“私有”，而非 VRF 接入链路视为“公开的”。

如果特定的 VRF 接入链路将站点 S 连接到 PE 路由器，然后与 S 的连通性（通过该接入链路）可以通过控制从相应的 VRF 输入的路由集合来限制。该 VRF 中的路由集合应该限制为通往站点（与 S 至少有一个相同 VPN 的站点）的路由集合。然后从 S 通过 VRF 接入链路发送的数据包，只能由 PE 寻路到另一个站点 S’，如果 S’与 S 存在一个相同的 VPN。也就是说，在任何一对没有共同的 VPN 的 VPN 站点之间的通信（通过 PE 路由器）被阻止。通过阻止到 VPN 站点的路由进入默认转发表，来防止 VPN 站点和非 VPN 站点之间的通信。

如果有多个接入链路从 S 到一个或更多的 PE 路由器，则可能会有多个 VRFs 可以用于为来自 S 的流量寻路。为了适当地限制 S 的连通性，所有 VRFs 中都必须存在同一组路由。或者，另一种方式可以在与 S 不同的接入链路上施加不同的连接限制。在这种情况下，一些与 S 接入链路相关联的 VRF 与其它的 VRFs 相比将包含不同的路由集合。

我们允许单个接入链路与一组 VRF 而不是一个 VRF 相关联的情况。这是有用的，如果希望将单个 VPN 划分为几个“子 VPN”，每个具有不同的连接限制，从子 VPN 中筛选一些客户数据包的特性。为了简单起见，我们通常会说一个接入链路与单个 VRF 相关联。

3.2 将 IP 数据包与 VRF 关联

当 PE 路由器接收到来自 CE 设备的数据包时，它必须确定收到数据包的接入链路，依次确定可用于转发该数据包的 VRF（或一组 VRF）。一般来说，要确定收到数据包的接入链路，PE 路由器需要关注收到数据包的物理接口，也可能关注数据包的第 2 层头的某些方面。例如，如果数据包的入接入链路是帧中继 VC，接入链路的身份可以从收到数据包的物理帧中继接口，以及数据包的帧中继头中的数据链路连接标识符（DLCI）字段确定。

虽然 PE 由数据包的第 2 层头来断定特定数据包从特定接入链路收到，这是不全面的，客户绝不可能通过改写头字段欺骗 SP 认为一个数据包是通过一个接入链路收到的，而其实际上从另一个不同的接入链路收到的。在上面的例子中，虽然接入链路部分通过检查帧中继头中的 DLCI 字段来确定，此字段无法由客户自由设置。

相反，它必须被设置为 SP 指定的值，否则数据包不能到达 PE 路由器。

在某些情况下，一个特定站点可能会被客户分成几个“虚拟站点”。SP 可以指定一组特定的 VRFs 用于寻路来自该站点的数据包，并允许客户设置一些包的特性，然后其用于从集合中选择特定的 VRF。

例如，每个虚拟站点可能被实现为一个 VLAN。SP 和客户可以同意来自特定的 CE 的数据包，某些 VLAN 值将用于标识某些 VRFs。当然，如果从 CE 收到的数据报携带的 VLAN 标签值没有达成一致，报文将被 PE 丢弃。另一种方法是使用 IP 源地址。在这种情况下，PE 使用从 CE 接收到的数据包中的 IP 源地址以及接收到数据包的接口，将数据包分配给特定的 VRF。此外，客户只能从允许客户使用的特定 VRFs 的集合进行选择。

如果希望将特定主机置于多个虚拟站点中，那么主机必须确定每个数据包关联哪个虚拟站点。它可以做到这一点，例如，通过发送来自不同 VLAN 的不同虚拟站点或出自不同的网络接口的数据包。

3.3 填充 VRF

VRF 用什么路由集合填充？

例如，PE1，PE2 和 PE3 作为三台 PE 路由器，CE1，CE2 和 CE3 作为三个 CE 路由器。假设 PE1 从 CE1 中学习到达 CE1 的站点的路由。如果 PE2 和 PE3 分别连接 CE2 和 CE3，还有某个 VPN V 包含 CE1，CE2 和 CE3，则 PE1 使用 BGP 分发从 CE1 学到的路由给 PE2 和 PE3。PE2 和 PE3 使用这些路由分别填充他们与 CE2 和 CE3 的站点关联的 VRFs。不在 VPN V 的站点学习的路由不能出现在这些 VRFs 中，这意味着来自 CE2 或 CE3 的数据包不能发送到不在 VPN V 的站点。

当我们谈论一个 PE 从一个 CE“学习”路由时，我们不能预料任何特定的学习技术。PE 可以通过动态路由算法学习路由，但也可通过配置那些路由（例如，静态路由）来“学习”路由。（在这种情况下，就是说 PE“学习”来自 CE 路由，也许是运用一点诗意破格。）

PEs 还需要从其他 PEs 学习属于某个给定 VPN 的路由。使用适当的路由组合来填充 VRFs 的规程在第 4 节介绍。

如果从一个特定的 PE 路由器到一个特定的站点有多个接入链路，它们可能都映射到同一个转发表。但是，如果策略要求，他们可以被映射到不同的转发表。例如，策略可能是这样，来自一个站点的一个特定接入链路仅用于内部网流量络，而来自该站点的另一个接入链路仅用于外部网络流量。（例如，也许 CE 链接的外部接入链路是防火墙，而 CE 连接的内部网接入链路不是。）在这种情况下，两个接入链路将与不同的 VRFs 相关联。

请注意，如果两个接入链路关联同一个 VRF，那么 PE 通过他们其中一个收到的数据包可以与 PE 从另一个接收到的数据包达到完全相同的目的地集合。所以两个接入链路不能关联相同的 VRF，除非每个 CE 与另一个都在完全相同的 VPNs 集合中。

如果一条接入链路通向一个处于多个 VPN 中的站点，其仍然可以与单个 VRF 相关联，在这种情况下，VRF 将包含该站点所属的全部 VPNs 集合的路由。

4. 通过 BGP 进行 VPN 路由分发

PE 路由器使用 BGP 分发彼此的 VPN 路由（更多准确地说，将 VPN 路由分发给彼此）。

我们允许每个 VPN 都有自己的地址空间，这意味着在不同 VPN 中给定地址可以表示不同系统。如果两个到相同 IP 地址前缀的路由实际上是不同的路由系统，极为重要的是，

确保 BGP 不将它们视为类似的。否则，BGP 可能选择只安装其中一个，使其他系统无法访问。此外，我们必须确保 POLICY 用于确定在哪些路由上发送哪些数据包；由于 BGP 安装了几条路由，在任何特定的 VRFs 中必须只有一条路由出现。

我们通过使用指定的新地址族来实现这些目标下面。

4.1 VPN-IPv4 地址族

BGP 多协议扩展[BGP-MP]允许 BGP 携带来自多个“地址族”路由。我们介绍一下“VPN-IPv4 地址族”这个概念。VPN-IPv4 地址是 12 字节数量，以 8 字节的路由识别器（RD）开始，以 4 字节 IPv4 地址结尾。如果几个 VPN 使用相同的 IPv4 地址前缀，PE 将它们转换为唯一的 VPN-IPv4 地址前缀。这确保了如果在几个不同的 VPNs 中使用相同的地址，BGP 可以携带几个到该地址的完全不同路由，每个 VPN 对应一个。

由于 VPN-IPv4 地址和 IPv4 地址是不同的地址族，BGP 从不将它们视为可类比的地址。

RD 只是一个数字，它不包含任何固有的信息；它不会识别路由的起点或路由将被分发的 VPNs 集合。RD 的目的是只允许创建到普通 IPv4 地址前缀的独特路由。其他手段用于确定重新分发路由到哪里（见第 4.3 节）。

RD 也可以用来创建去往同一系统多条不同的路由。我们已经讨论过一个这样的情况，去往特定服务器的路由对于内部网流量和外部网络流量应该是不同的。这可以通过创建两个具有相同 IPv4 部分，但具有不同 RD 的不同的 VPN-IPv4 路由来实现。这允许 BGP 到相同的系统安装多个不同的路由，并允许使用策略（见第 4.3.5 节）决定哪个包使用哪个路由。

RDs 是构造化的，以便每个服务提供商都可以管理它自己的“编号空间”（即可以自己分配的 RD），而不会与任何其他服务提供商产生 RD 分配冲突。RD 由三个字段组成：一个 2 字节类型字段，管理员字段和分配的号码字段。该类型字段的值确定其他两个字段的长度，以及管理员字段的语义。该管理员字段标识分配遍号权限，并且分配遍号字段包含由指定的机构已分配的号码，用于特定目的。例如，一个可以有一个 RD，其管理员字段包含一个自治系统号（ASN），其（4 字节）号码字段包含一个由 ASN 所属的 SP（已经由专门的机构分配给该 SP）分配的号码。

给出 RDs 这个结构，为了确保一个提供 VPN 骨干服务 SP 可以随时创建一个独特的 RD。但是，这个结构对 BGP 是没有意义的；当 BGP 比较两个这样的地址前缀时，它完全忽略该结构。

需要配置 PE 使得通向一个特定 CE 的路由与特定的 RD 相关联。该配置可能使通往相同 CE 的所有路由与相同的 RD 相关联，或可能引起不同的路由与不同的 RDs 相关联，即使它们通往相同的 CE。

4.2 路由识别器的编码

如上所述，VPN-IPv4 地址由 8 字节的路由识别器后跟一个 4 字节的 IPv4 地址组成。RD 编码如下：

- 类型字段： 2 字节

- 值字段： 6 字节

值字段的解释取决于类型字段的值。目前，类型字段的三个值是定义为：0,1 和 2。

- 类型 0 ： 值字段由两个子字段组成：
  * 管理员子字段： 2 字节
  * 分配号码子字段： 4 字节

  管理员子字段必须包含一个自治系统号。如果这个 ASN 来自公共 ASN 空间，它必须由专门的机构分配（使用来自私有的 ASN 空间的 ASN 值是非常不鼓励的）。该分配号码子字段包含一个编号，其来自由专门机构已经指派 ASN 的企业（SP）管理的编码空间。
- 类型 1 ： 值字段由两个子字段组成：
  * 管理员子字段： 4 字节
  * 分配号码子字段： 2 字节

  管理员子字段必须包含一个 IP 地址。如果这 IP 地址是从公共 IP 地址空间，由专门的机构分配（使用来自私人 IP 地址空间的地址是非常不鼓励的）。该分配号码子字段包含一个编号，其来自由已经指派 IP 地址的企业管理的编码空间。
- 类型 2 ： 值字段由两个子字段组成：
  * 管理员子字段： 4 字节
  * 分配号码子字段： 2 字节

  管理员子字段必须包含一个 4 字节的自治系统号[BGP-AS4]。如果这个 ASN 来自公共 ASN 空间，它必须由专门的机构分配（使用私有 ASN 空间的 ASN 值是非常不鼓励的）。分配号码子字段包含一个编号，其来自由专门机构已经指派 ASN 的企业（SP）管理的编码空间。

4.3 控制路由分发

在本节中，我们将讨论控制 VPN-IPv4 路由分发的方式。

如果 PE 路由器连接到一个特定的 VPN（通过连接到该 VPN 中的特定 CE），它会从连接的 CE 路由器学习一些 VPN 的 IP 路由。通过特定的接入链路从 CE 路由 peer 学到的路由可以安装在与该接入链路相关联 VRF 中。正是由 PE 从 CE 学习路由的方式决定以这种方式安装路由。特别是 PE 和 CE 是路由协议 peer 时，这是由路由协议决策过程决定的；这在第 7 节中讨论。

然后将这些路由转换为 VPN-IP4 路由，并将其导出 BGP。如果有多条路由到特定的 VPN-IP4 地址前缀，BGP 利用决策过程选择“最好”的一个。然后由 BGP 分发该路由到其他需要知道的 PE 的集合。在这些其他 PE 上，BGP 将再次为特定 VPN-IP4 地址前缀选择的最佳路由。然后选择的 VPN-IP4 路由被转换回 IP 路由，并导入一个或多个 VRF。它们是否实际安装在 VRF 中，取决于 PE 和这些与正在讨论的 VRF 相关的 CE 间使用的路由决策过程方法。最后，可以将安装在 VRF 中的任何路由分发给相关的 CE 路由器。

4.3.1 路由目标属性

每个 VRF 与一个或多个路由目标（RT）属性相关联。

当一个 PE 路由器创建一条 VPN-IPv4 路由（PE 路由器从 CE 学习 IPv4 路由）时，它与一个或多个路由目标属性相关联。携带在 BGP 路由属性中。

任何与 RT T 相关联的路由都必须分发给每个关联 RT T 的 VRF 的 PE 路由器。PE 路由器接收如此路由时，它有资格安装在与 RT T 相关联的 PEs 的 VRF 中。（实际上是否安装取决于 BGP 决策过程结果，以及运行于 PE/CE 接口上的 IGP（即域内路由协议）决策过程的结果）。

一组 RT 属性可以被认为是标识一组站点。（尽管将它更为精确标识一组 VRFs）。关联特定 RT 属性的路由允许将路由放置在 VRFs 中，用于为从相应的站点中接收的流量进行寻路。

存在一组 RT，PE 路由器附加其到一个从站点 S 收到路由的；这些可能被称为“出口目标”。而存在一组 RT,PE 路由器用其来确定是否可以放置从另一个 PE 路由器接收的路由于与站点 S 相关联的 VRF 中；这些可能被称为“入口目标”。这两组是有区别的，不一定相同。特定的 VPN-IPv4 路由有资格安装在一个特定的 VRF 中，如果存在某个 RT 即是路由 RT 中的一个又是 VRF 的入口目标中的一个。

RT 属性执行的功能类似于 BGP 团体属性执行。然而，后者的格式不适用于目前的目的，由于其只有一个 2 字节的编号空间。希望构造类似于我们描述的 RDs（参见第 4.2 节）的格式，于是类型字段定义了管理员字段的长度，属性的其余部分是详细解释管理员的编号空间的数字。这可以使用 BGP 扩展团体来完成。这里讨论的 RT 被编码为 BGP 扩展团体 RT[BGP-EXTCOMM]。它们的结构类似于 RDs。

当一个 BGP speaker 接收到多条到同一个 VPN-IPv4 前缀的路由，对于 BGP 路由优先级规则用于选择 BGP 安装哪个 VPN-IPv4 路由。

请注意，一条路由只能有一个 RD，但它可以有多个 RT。在 BGP 中，如果有一个具有多个属性的路由，而不是多个路由，则可以提高可扩展性。一方面可以通过创建更多路由（即，使用更多的 RDs）来消除 RT 属性，但是无利于拓展性。

PE 如何确定哪个 RT 属性关联一个给定路由？有许多不同的可能方法。PE 可能被配置到达一个特定站点的所有路由与特定的 RT 相管理。或 PE 可能配置为将通往指定站点的某些路由与一个 RT 相关联，某些与另一个相关联。

如果 PE 和 CE 是他们自己的 BGP peer（见第 7 节），那么 SP 可能允许客户在限制范围内指定如何分发路由。SP 和客户将需要提前同意允许附加 RTs 集合到客户的 VPN 路由。然后，CE 可以附加一个或多个这些 RTs 到（要分发给 PE）每个 IP 路由。在实际情况下，客户可以在商定的限制下，自由地指定其路由分发策略。如果允许 CE 将 RT 附加到其路由，PE 必须过滤出那些包含客户不允许使用 RTs 的所有路由。如果不允许 CE 将 RTs 附加到其路由，但无论如何，PE 必须在将客户的路由转换为 VPN-IPv4 路由前移除 RT。

4.3.2 通过 BGP 分发 PEs 间的路由

如果一个 VPN 的两个站点连接的 PEs 在同一个自治系统中，PEs 可以通过他们之间的 IBGP 连接方式彼此分发 VPN-IPv4 路由。（术语“IBGP”是指当在同一个自治系统中的两个 BGP speakers 之间有 BGP 连接时使用的一组协议和规程。这与“EBGP”不同，在不同的自治系统中两个 BGP speakers 之间使用的一组规程）。或者，每个 speaker 存在到路由反射器的 IBGP 连接[BGP-RR]。

当 PE 路由器通过 BGP 分发 VPN-IPv4 路由时，它将使用它自己的地址作为“BGP 下一跳”。该地址编码为一个 RD 为 0 的 VPN-IPv4 地址.（[BGP-MP]要求下一跳地址与网络层可达性信息（NLRI）属于同一个地址族）。它还分配和分发 MPLS 标签。（本质上，PE 路由器不分发没有标签的 VPN-IPv4 路由，但分发带标签的 VPN-IPv4 路由。参看[MPLS-BGP]）。PE 处理接收到的堆栈顶部有标签的报文时，即 PE 将弹出堆栈，并适当地处理数据包。

PE 可以精确的对出现在 VRF 中的路由分发，或者可以执行聚合和分发那些聚合路由，或者它可以这么做，也可以那么做。

假设 PE 已经给路由 R 分配了标签 L，并且通过 BGP 分发此标签映射。如果 R 是 VRF 中路由集合的聚合路由，PE 将知道来自骨干网的携带这个标签的数据包，必须在 VRF 中查出有它们的目的地地址。当 PE 在其标签信息库中查找标签时，它将知道必须使用哪个 VRF。另一方面，如果 R 不是聚合路由，那么当 PE 查找封装在数据包头中的标签时，它会得知出口接入链路。在这种情况下，不需要查找 VRF。

我们期望最常见是未聚合的路由的情况。聚合的情况可能非常有用，如果 VRF 包含大量的主机路由（例如，如拨号接入），或者如果 VRF 关联一个局域网（LAN）接口（其中在 LAN 上的每个系统存在不同的出 2 层头，但一条路由不能够为每个这样的系统分发信息）。

每个路由是否具有不同的标签是一个实现事情。存在一些合适的算法，可以用来确定两条路由是否分配相同的标签：

- 可以选择为整个 VRF 设置单个标签，因此单个标签由该 VRF 的所有路由共享。则当出口 PE 收到带有该标签的数据包时，必须在该 VRF（的分组的“出口 VRF”）查找该数据包的 IP 地址，以便确定数据包的出口接入链路和相应的数据链路封装。

- 每个接入链路可以选择单个标签，使得在相同的“输出接入链路”上所有路由共享单个标签。这使得避免在出口 VRF 中进行查找，尽管如此可能需要进行某些查找才能确定数据链接封装，例如地址解析协议（ARP）查找。

- 可以选择为每个路由选择不同的标签。那么如果一条路由可能有一个以上的接入链路，PE/CE 路由可以切换一个路由的首选路径从一个接入链路到另一个接入链路，而不需要分配任何新标签给这个路由。

也可能存在其他合适的算法。算法的选择完全由出口 PE 自行决定，否则透明。

用 BGP 分发 MPLS 标签的方式，我们假设携带这样的标签的 MPLS 报文可以被隧道从安装相应的 BGP 分发路由的路由器传送给该路由的 BGP 下一跳路由器。这也要求这两个路

由器之间存在标签交换路径，否则一些其他的隧道技术（例如，[MPLS-in-IP-GRE]）在他们之间使用。

这条隧道可能遵循“尽力而为”的路由，或者它可能遵循 TE 路由。在给定的一对路由器之间，可能存在一个或数个这样的隧道，也许具有不同服务质量（QoS）特征。VPN 架构所有重要的是存在一些这样的隧道。为确保实现此 VPN 架构的系统之间的互操作性，使用 MPLS 标签交换路径作为隧道技术，所有这些系统必须支持标签分发协议（LDP）[MPLS-LDP]。特别是下游主动模式必须被既不是标签控制 ATM（LC-ATM）[MPLS-ATM]接口也不是标签控制帧中继（LC-FR）[MPLS-FR]的接口支持。下行按需模式必须被 LC-ATM 接口和 LC-FR 接口支持。

如果隧道遵循“尽力而为”的路由，则 PE 通过查找其默认转发表中 IP 地址来发现到远端的路由。

一个 PE 路由器，除非是路由反射器（见第 4.3.3 节）或者用于互联网供应商 VPN 的自治系统边界路由器（ASBR）（见第 10 节），不应该安装 VPN-IPv4 路由，除非它有至少一个 VRF 的导入目标与路由 RT 属性中的一个相同。应使用入站过滤来丢弃这样的路由。如果稍后添加新的导入目标到一个 PE 的 VRF（“VPN 加入”操作），那么它必须获取以前可能丢弃的路由。这可以使用[BGP-RFSH]中描述的刷新机制完成。出站路由过滤机制[BGP-ORF]也可以有利于使过滤更加动态。

同样，如果一个特定的导入目标不再存在于任何一个 PE 的 VRF 中（一个或多个“VPN 剪枝”操作的结果），PE 可以丢弃其所有路由，导致的结果是 PE 的任何 VRFs 中不再有 RT 属性如导入目标的路由。

没有连接到任何 VPN 并且不是路由反射器（即，P 路由器）的路由器不会在其中安装任何 VPN-IPv4 路由。

请注意，VPN Join/加入和 Prune/剪枝操作是无破坏性的且不要求关闭任何 BGP 连接，只要使用[BGP-RFSH]的刷新机制。

由于这些分配规则，没有一个 PE 需要维护所有 VPNs 的所有路由；这是一个重要的可扩展考虑。

4.3.3 使用路由反射器

与其在 PE 之间建立完整的 IBGP 全连接，不如利用 BGP 路由反射器[BGP-RR]改进扩展性。所有使用路由反射器的常规技术都有利于提高扩展性（例如，路由反射器层次结构）。

路由反射器是唯一需要知道不直接连接到其的 VPN 路由信息的系统。然而，任何一个路由反射器没有必要知道主干网所支持的所有 VPN 的所有 VPN-IPv4 路由。

下面我们将以两种不同的方式概述，如何区分一组路由反射器之间的 VPN-IPv4 路由的集合。

1. 每个路由反射器都预先配置了一系列 RTs。以备不时之需，相同的列表上可能预先配置了多于一个的路由反射器。路由反射器使用预配置 RTs 列表构建其入站路由过滤。路由反射器可以使用[BGP-ORF]技术在每个对等体上（不管 peer

是否是另一个路由反射器或 PE）安装出站路由过滤器集（ORFs），ORFs 包含其预配置的 RT 列表。请注意，路由反射器应该接受来自其他路由反射器的 ORF，这意味着路由反射器应向其他路由反射器宣告 ORF 能力。

服务提供商可以修改路由反射器上预先配置的 RT 列表。当这样做，路由反射器修改其安装在所有 IBGP peer 上的 ORFs。为了减少路由反射器上配置变化的频率，每个路由反射器可以被预先配置一大批 RTs。这样，当一条新的 VPN 需要新的 RT 时，存在一个或多个路由反射器已经（预）配置此 RT。

除非给定的 PE 是所有路由反射器的客户端，当一个新的 VPN 被添加到 PE（“VPN Join”），它需要成为维护那个 VPN 路由的路由反射器的客户端。同样，从 PE 中删除（“VPN 修剪”）现有的 VPN 可能导致 PE 不再需要成为一些路由反射器的客户端的情况。无论如何，加入或修剪操作是无中断的（如果应用[BGP-RFSH]，不需要断开 BGP 连接，只需要做好备份。

（通过“添加新的 VPN 到 PE”，我们的意思是增加一个新的导入 RT 到其中一个 VRFs，或添加新的 VRF，其导入 RT 没有在任何其他的 PE 的 VRFs 中）。

2. 另一种方式是让每个 PE 成为一个路由反射器子集的客户端。路由反射器未预先配置 RT 列表，并且从客户端（PES）收到的路由不执行入站路由过滤；而是接受从所有客户（PEs）收到的所有路由。路由反射器跟踪所有收到的路由携带的 RTs 集。当路由反射器从其客户端接收一条路由，该路由的一个 RT 不是此集合中时，RT 立即添加到集合中。另一方面，当路由反射器集合中特定 RT 不再有任何路由，路由反射器应该延迟（几个小时）从集合中删除此 RT。

   路由反射器使用此组合形成入站路由过滤器，适用于从其他路由反射器接收的路由。路由反射器也可以使用 ORFs 在其他路由反射器安装适当的出站路由过滤器。就像第一种方法，一个路由反射器应接受来自其他路由反射器的 ORFs。为了做到这一点，路由反射器向其它路由反射器通告 ORF 能力。

   当路由反射器更改集合时，应立即进行更改其入站路由过滤器。另外，如果路由反射器使用了 ORFs，那么 ORFs 必须立即改变来反映集合中的变化。如果路由反射器不使用 ORFs，并且添加了一个新的 RT 到该集合，在更改其入站路由过滤器后，路由反射器必须向其他路由反射器发布 BGP 刷新。

   上述“几个小时”的延迟允许路由反射器，即使在它感兴趣的路由的最后的客户失去之后，也可以保持给定 RT 的路由。如果客户的“消失”只是暂时的，这样可以防止重新获得这些路由。

   通过此过程，VPN Join 和 Prune 操作也是非破坏性的。

   请注意，如果某些客户端 PE 的 VRF 存在导入 RT 不是其出口 RT 中的一个，此技术将无法正常工作。

在这些规程中，连接到特定 VPN 的 PE 路由器“自动发现”连接到同一 VPN 的其他 PEs。当一个添加新的 PE 路由器，或者现有 PE 路由器连接到新的 VPN 时，不需要重新配置其他 PE 路由器。

就像没有一台 PE 路由器需要知道所有骨干网上支持的 VPN-IPv4 一样，这些分发规则确保没有一个路由反射器（RR）需要知道所有骨干网上支持的 VPN-IPv4 一样。结果，在骨干网上可以支持的路由总数不受任何单一设备的容量限制，因此几乎可以没有限制地增加。

4.3.4 如何在 BGP 中携带 VPN-IPv4 NLRI

BGP 多协议扩展[BGP-MP]用于对 NLRI 编码。如果地址族标识符（AFI）字段设置为 1，并且子地址族标识符（SAFI）字段设置为 128，NLRI 是一个 MPLS 标签的 VPN-IPv4 地址。AFI 1 用于网络层协议，其相关联的 NLRI 仍然是 IP。请注意，该 VPN 架构不需要具备分发未携带标签的 VPN-IPv4 地址能力。

为了使两个 BGP speaker 交换带有标签的 VPN-IPv4 NLRI，必须使用 BGP 能力公告来确保它们都能够正确处理这种 NLRI。这是按照[BGP-MP]中规定完成的，通过使用能力代码 1（多协议 BGP）AFI 为 1，SAFI 为 128。

带标签的 VPN-IPv4 NLRI 本身按照[MPLS-BGP]中的规定进行编码，其中前缀由 8 字节的 RD 组成后跟一个 IPv4 前缀。

4.3.5 使用路由目标构建 VPN

通过正确设置导入目标和导出目标，可以构建不同种类的 VPN。

假设想要创建一个全连接的封闭用户组，即一组站点其中每个站点都可以直接发送流量到组内其他站点，但不能发送流量到组外其他站点或从组外其他站点接收流量。然后每个站点都与一个 VRF 相关联，唯一的 RT 属性被选择，其分配给每个 VRF 的 RT 作为导入目标和导出目标，以及没有分配给任何其他 VRFs 的 RT 作为导入目标或出口目标。

或者，假设一个想要的，无论什么原因，创造一个“中心和辐射”的类型的 VPN。这可以通过使用两个 RT 的值来完成，一个意思是“Hub”，一个意思是“Spoke”。在连接到枢纽/Hub 站点的 VRFs，“Hub”是“出口目标”，而“Spoke”是导入目标。在连接到辐射/Spoke 站点的 VRFs，“Hub”是导入目标，“Spoke”是导出目标。

因此，控制各种网站之间的路由信息分发的方法是非常灵活，其循环为构建 VPN 提供了很大的灵活性。

4.3.6 单个 PE 中 VRFs 之间的路由分发

可以将路由从一个 VRF 分发到另一个 VRF，即使两个 VRFs 都在同一个 PE 中，虽然在这种情况下不能说路由已由 BGP 分发。不过，在一个单一 PE 内将特定路由从一个 VRF 分发到另一个 VRF 的决策，与在不同的 PEs 上 VRFs 上做出的决策相同。也就是说，其取决于被分发路由（或如果路由由 BGP 分发）的 RT 属性和第二个 VRF 的导入目标。

5. 转发

如果骨干中的中间的路由器没有任何关于到 VPNs 的路由的信息，如何从一个 VPN 站点到另一个 VPN 站点转发数据包？

当 PE 从 CE 设备收到 IP 包时，会选择一个特定的 VRF，用于查找数据包的目的地址。这个选择是基于数据包的入口接入链路。假设找到了匹配项。结果是我们学习了数据包“下一跳”。

如果数据包的下一跳通过该 PE 的 VRF 接入链路直接达到（即，数据包的出口接入链路与其入口接入链路在相同的 PE 上），则数据包在出口接入链路上被发送，没有 MPLS 标签被 PUSH 到数据包的标签堆栈上。如果入口和出口接入链路在同一个 PE 上，但是与不同的 VRFs 相关联，如果在入口接入链路的 VRF 中最匹配目的地址的路由是由出口接入链路中的数条路由聚合而成，其可能也需要在出口 VRF 中查找数据包的目的地址。

如果数据包的下一跳通过 VRF 接入链路不能到达，则数据包必须经过至少一跳骨干网。因此该数据包具有“BGP 下一跳”，且 BGP 下一跳将为数据包目的地址最匹配的路由分配一个 MPLS 标签。将此标签称为“VPN 路由标签”。IP 报文变成具有 VPN 路由标签（作为标签堆栈上的唯一标签）的 MPLS 报文。

该数据包必须被隧道传送到 BGP 下一跳。

如果骨干网支持 MPLS，则完成如下：

- PE 路由器（和任何自治系统边界路由器）重新分配 VPN-IPv4 地址，其需要为其自己插入/32 地址前缀到骨干网的 IGP 路由表。这使得在骨干网络中的每个节点处启用了 MPLS，为每个 PE 路由器分配与该路由相对应的标签。为确保不同实现的互操作性时，需要在这些横跨骨干网的标签交换路径上支持 LDP 的设置。但是，其他设置这种标签交换路径的方法也是可能的（其他一些方法可能不需要在 IGP 中存在/32 地址前缀）。

- 如果存在任何到达 BGP 下一跳的 TE 隧道，且如果其中存在一个或多个可供该问题中的数据包使用的 TE 隧道，选择这些隧道其中的一个。这条隧道将是与 MPLS 标签相关联，即“隧道标签”。隧道标签被 PUSH 进 MPLS 标签栈，数据包被转发到隧道的下一跳。

- 除此以外，

  * 该数据包将有一个“IGP 下一跳”，即沿着 IGP 路由到 BGP 下一跳的“下一跳”。

  * 如果 BGP 下一跳和 IGP 下一跳相同，且如果使用倒数第二跳弹出/ PHP，然后将数据包发送到 IGP 下一跳，仅携带 VPN 路由标签。

  * 否则，IGP 将为与 BGP 下一跳的地址最匹配的路由分配一个标签。称之为“隧道标签”。隧道标签被 PUSH 为数据包的栈顶标签。然后转发数据包到 IGP 下一跳。

- 然后，MPLS 将通过骨干网将数据包传递到 BGP 下一跳，其 VPN 标签将被检查。

如果骨干网不支持 MPLS，则携带 VPN 路由标签的 MPLS 报文可能使用[MPLS-in-IP-GRE]的技术，被隧道传输到 BGP 下一跳。当数据包脱离隧道时，它将在 BGP 下一跳，其 VPN 路由标签将被检查。

在 BGP 下一跳，数据包的处理取决于 VPN 路由标签（见第 4.3.2 节）。在许多情况下，PE 将能够从该标签确定应将数据包传输（到 CE 设备）的接入链路，以及这个接口的正确数据链路层头。在其他情况下，在转发到 CE 设备之前，PE 可能只决定需要某个特定 VRF 来查找数据包的目的地址。还有中间情景，VPN 路由标签可以确定数据包的出口接入链路，但是查找操作（例如，ARP）仍然需要，以确定在这个接入链路上的数据包的数据链路头。

MPLS 标头本身的信息和/或相关的标签信息，也可以用在到 CE 的接口上提供 QoS。

无论如何，如果到达入口 PE 的数据包是未携带标签的 IP 数据包，离开它的出口 PE 时它将再次成为一个未携带标签的数据包。

事实上，具有 VPN 路由标签的报文通过骨干网用隧道传输，这使得 P 路由器可以不需要保留所有的 VPN 路由。这对于方案可扩展性非常重要。主干网甚至不需要有到 CE 路由，只需要有到 PE 路由。

关于隧道，这个规范值得注意的是：

- 不要求隧道是点对点的；点对多点可以使用；
- 不要求有明确的隧道设置（通过信令或通过手动配置）；
- 不要求有任何隧道的特定信令；
- 不要求 P 或 PE 路由器中存在任何隧道的特定状态，超出所必需的路由信息和（如果使用）MPLS 标签信息的维护。

当然，这个规范与点对点隧道的使用相兼容，必须明确配置和/或发出信号，在某些情况下，可能有使用这种隧道的原因。

与选择特定隧道技术的相关考虑因素不在本规范的范围之内。

6. 维护 VPN 的正确隔离

为了保持一个 VPN 与另一个 VPN 的正确隔离，骨干网中没有路由器从骨干网外接收一个隧道报文，这很重要，除非确定该隧道的两个端点都在骨干之外。

如果使用 MPLS 作为隧道技术，这意味着骨干路由器不得从任何相邻的非骨干设备接受带标签的数据包，除非满足以下两个条件：

1. 标签堆栈顶部的标签实际上是由该骨干路由器分发到该非骨干网设备
2. 骨干路由器可以确定，在任何栈底标签和 IP 头被检查之前，该标签的使用将会使数据包离开骨干网。

第一个条件确保从非骨干路由器接收到的任何带标签的数据包，在标签堆栈的顶部合法且正确的分配的标签。第二个条件确保了骨干路由器永远不会看到栈顶标签下面。当然，满足这两个条件的最简单的方法就是骨干设备拒绝接受来自非骨干网设备的标签数据包。

如果不使用 MPLS 作为隧道技术，则过滤必须做到这一点来确保 MPLS-in-IP 或 MPLS-in-GRE 的数据包被接受进骨干网(只有在数据包的 IP 目的地址会使其发送到骨干网以外)。

7. PE 如何学习 CE 的路由

连接到特定 VPN 的 PE 路由器需要知道对于每个都通向该 VPN 的接入链路，哪个 VPN 的地址应该通过该接入链路是可达的。

PE 将这些地址转换为 VPN-IPv4 地址，使用配置的 RD。然后，PE 将这些 VPN-IPv4 路由视为 BGP 的输入。VPN 站点的路由不会泄漏到骨干网的 IGP 中。

准确来说，PE/CE 的路由分发技术可能取决于特定 CE 是否在“中转 VPN”中。一个“中转 VPN”是指包含一个从“第三方”接收路由的路由器（即来自 VPN 外的路由器，但是不是 PE 路由器），并将这些路由重新分发到 PE 路由器。不是中转 VPN 的 VPN 是“stub VPN”。绝大多数的 VPN，包括所有的企业网，在这个意义上，这将被认为是“存根”。

可能的 PE/CE 分发技术有：

1. 可以使用静态路由（即配置）。（这是可能仅在存根 VPN 中有用。）

2. PE 和 CE 路由器可能是路由信息协议（RIP）[RIP]对等体，CE 可以使用 RIP 告诉 PE 路由器，CE 路由器站点可达的地址前缀集合。在 CE 中配置 RIP 时，必须注意确保来自其他站点的地址前缀（即 CE 路由器从 PE 路由器学到的地址前缀）从不向 PE 发布。更确切地说，如果一个 PE 路由器 PE1 接收 VPN-IPv4 路由 R1，作为结果分发 IPv4 路由 R2 到 CE，则 R2 不能被从 CE 站点经 PE 路由器 PE2 分发回来，（其中 PE1 和 PE2 可能是相同的路由器或不同的路由器），除非 PE2 将 R2 映射为不同于 R1 的 VPN-IPv4 路由（即，包含不同的 RD）。

3. PE 和 CE 路由器可能是 OSPF 对等体。一个 PE 路由器就是 CE 路由器的 OSPF 对等体，出现在 CE 路由器上作为区域 0 路由器。如果 PE 路由器是不同的 VPN 中 CE 路由器的 OSPF 对等体，PE 必须运行多个 OSPF 实例。

PE 通过 OSPF 学习的 IPv4 路由是作为 VPN-IPv4 路由被引入 BGP。扩展团体属性用于携带，路由信息，以及在正确类型的 OSPF 链路状态公告（LSA）中分发路由信息到 VPN 中的其他 CE 路由器所需的所有信息。OSPF 路由标记用于确保从 MPLS / BGP 骨干网收到的路由不能被送回骨干网。

PE 和 CE 之间的 OSPF 使用的完整处理的规范可以在[VPN-OSPF]和[OSPF-2547-DNBIT]发现。

4. PE 和 CE 路由器可以是 BGP 对等体，CE 路由器可以使用 BGP（特别是 EBGP 告诉 PE 路由器，CE 路由器的站点的地址前缀集合。这个技术可用于存根 VPN 或中转 VPN）。

这种技术比其他技术有很多优点：

a） 与 IGP 替代方案不同，这不需要 PE 为了与多个 CE 进行通话而运行多个路由算法实例。

b） 明确设计 BGP 仅是为此功能：在不同的管理机构运行的系统之间传递路由信息。

c） 如果站点包含“BGP 后门”，即路由器通过 BGP 连接到非 PE 的路由器，这个程序在任何情况下都能正常工作。其他程序可能或可能不工作，具体取决于具体的情况。

d） 使用 BGP 使 CE 易于传递路由属性到 PE。属性集合及其使用的完整的规范这个文件的描述范围。但是，这方面的一些例子使用如下：

- CE 可以从 PE 授权用来附加到路由的 RTs，为每个路由提议一个特定的 RT。而 PE 只附加建议的 RT，而不是全部集合。这允许 CE 管理员动态控制来自 CE 的路由。

- 其他类型的扩展团体属性可能被定义，意图是要那些属性从 CE 到 CE 透明地传递（即，不被 PE 路由器改变）。这允许 CE 管理员实现额外的路由过滤，远超 PEs 完成的过滤。这额外的过滤不需要与 SP 协调。

另一方面，使用 BGP 可能对 CE 管理员是新东西。

如果站点不在中转 VPN 中，请注意，它不需要独一无二的自治系统编号（ASN）。不在同一中转 VPN 中每个 CE 的站点是可以使用相同的 ASN。可以从私人 ASN 空间选择，它将会从被 PE 中剥离。通过使用站点的源属性来防止路由环路（见下文）。

如果一组站点构成中转 VPN 怎么办？一般情况下，只要 VPN 本身就是一个互联网服务提供商（ISP‘s）网络，ISP 本身从另一个 SP 购买骨干网服务。后者 SP 可能被称为“承运人承运人”。在这种情况下，提供 VPN 最好的方法是使 CE 路由器支持 MPLS，且使用第 9 节所述的技术。

当我们不需要区分 PE 学习给定站点的地址前缀的不同的方式，我们只会说 PE 已经从那个站点中学到了路由。这包括 PE 手动配置路由的情况。

在 PE 重新分发从一个站点学到的 VPN-IPv4 路由之前，必须为这个路由分配一个 Route Target 属性（见第 4.3.1 节），并且可以向该路由分配站点源属性。

站点源属性（如果使用的话）被编码为路由源拓展团体[BGP-EXTCOMM]。这个属性的目的是唯一标识从特定的站点学到的路由集合。在某些情况下，需要此属性来确保，通过特定的 PE/CE 连接从一个特定站点学习到路由，不能通过不同的 PE/CE 连接分发回该站点。如果使用 BGP 作为 PE/CE 协议，但不同的站点尚未分配 ASN 的，则这个特别有用。

8. CEs 如何学习 PEs 的路由

在本节中，我们假设 CE 设备是路由器。

如果 PE 在 VRF 中放置一个路由，其用于对来自特定 CE 的数据包进行寻路，则一般来说，PE 可以将该路由分发给 CE。当然，只有在 PE/CE 协议规则允许的情况下 PE 可能会分发该路由给 CE。（例如，如果特定的 PE/CE 协议具有“水平分割”，VRF 中的某些路由不能重新分发回到 CE。）我们对从 PE 分发到 CE 的路由再增加一个限制：如果路由的站点源属性标识一个特定的站点，该路由不能重新分发到该站点中的任何 CE。

然而，在大多数情况下，对 PE 来说只要简单地将默认路由分发给 CE 就足够了。（在某些情况下，甚至只要 CE 配置默认路由指向 PE 就足够了）。这通常下工作在任何站点（站点本身不需要将默认路由分发到其他站点）。（例如，如果企业 VPN 中的一个站点具有该公司的访问互联网权限，该站点可能需要将默认路由分发到其他站点，但无法将默认分发到该网站本身）。用于 CE 到 PE 路由分发的过程怎也可用于从 PE 到 CE 的分发路由。

9. 运营商的运营商

有时，VPN 实际上可能是 ISP 的网络，拥有自己的对等体和路由策略。有时 VPN 可能是一个轮流为自己的客户提供 VPN 服务的 SP 网络。像这样的 VPN 也可以从另一个 SP 获得骨干网服务，“运营商的运营商”，基本上使用这个文件描述的相同方法。但是，在这些情况下，CE 路由器是必须支持 MPLS。尤其是：

- CE 路由器只能分发 VPN 内部的路由给 PE 路由器。这允许 VPN 作为存根 VPN 处理。
- CE 路由器应该支持 MPLS，以便它们应该能够从 PE 路由器接收标签，并发送带标签的数据包到 PE 路由器。他们不需要分发标签自己。
- PE 路由器应该向 CE 路由器，为其分配到 CE 路由器的路由分发标签。

  PE 不得将相同的标签分发给两个不同的 CE 除非满足以下条件之一：

  * 两个 CE 与完全相同的 VRFs 集合相关联;
  * PE 为每个 CE 保留不同的传入标签映射（[MPLS-ARCH]）。

  此外，当 PE 从 CE 接收到带标签的数据包时，其必须验证顶级标签是分发给该 CE 的标签。
- 不同站点的路由器彼此应建立 BGP 连接为了交换外部路由（即，VPN 外部的路由）。
- CE 路由器必须知道所有外部路由。

然后当 CE 路由器查找数据包的目的地址时，路由查找将解析为内部地址，通常是数据包 BGP 下一跳的地址。CE 为数据包打上标签并将数据包发送给 PE。PE 使用数据包的顶层 MPLS 标签选择 BGP 下一跳，而不是在 VRF 中查找数据包的 IP 目标地址。结果，如果 BGP 下一跳不止一跳，顶级标签将是由两个标签，隧道标签和 VPN 路由标签替代。如果 BGP 下一跳是一跳，顶级标签可能只是被替换为 VPN 路由标签。如果入口 PE 也是出口 PE，顶级标签将被弹出。当数据包从其出口 PE 发送到 CE 时，该数据包将比其从入口 PE 收到时少一个 MPLS 标签。

在上述步骤中，CE 路由器是 VPN 中唯一需要支持 MPLS 的路由器。如果，另一方面，所有的特定 VPN 站点的路由器支持 MPLS，则不再要求 CE 路由器知道所有外部路由。需要的是，对于任何知道外部路由的路由器，负责将标签堆栈放在迄今为止未带标签的数据包上，并且存在引导那些路由器到达其他站点的 BGP 对等体的标签交换路径。在这种情况下，对于 CE 路由器分发到 PE 路由器的每个内部路由，还必须分发标签。

10. 多 AS 骨干网

如果 VPN 的两个站点连接到不同的 AS，该怎么办（例如，因为站点连接到不同的 SPs）？连接到该 VPN 的 PE 路由器将无法维护彼此 IBGP 连接，或公共路由反射器。相反，需要一些使用 EBGP 方法来分发 VPN-IPv4 地址。

处理这种情况有很多不同的方法，我们这样做以增加可扩展性的顺序呈现。

a）ASBR 中的 VRF-to-VRF 连接。

在此过程中，一个 AS 中的 PE 路由器直接连接到另一个 AS 的 PE 路由器。两个 PE 路由器通过多个子接口直接连接，每个 VPN 中至少一个路由需要从 AS 传递到 AS。每个 PE 都会把对方视为 CE 路由器。那就是 PE 将每个这样的子接口与 VRF 相关联，并使用 EBGP 将彼此分发未携带标签的 IPv4 地址。

这是一个“正常工作”的过程，而在 ASes 边界之间不需要 MPLS。但是，它不能像下面讨论的其他规程一样扩展。

b）EBGP 从 AS 重新分发携带标签的 VPN-IPv4 路由到相邻 AS。

在此过程中，PE 路由器使用 IBGP 重分发携带标签的 VPN-IPv4 路由，到自治系统边界路由器（ASBR）或到 ASBR 为其客户的路由反射器。ASBR 然后使用 EBGP 重新分发那些携带标签的 VPN-IPv4 路由到另一个 AS 中的 ASBR，依次将它们分发到该 AS 中的 PE 路由器，或者也可以到另一个 ASBR 又分发它们，等等。

使用此过程时，只能在私有对等直连点的 EBGP 连接上接受 VPN-IPv4 路由，作为 SPs 之间信任安排的一部分。VPN-IPv4 路由应该既不分发给开放互联网或任何不信任的 BGP 对等体。ASBR 不应该接受来自 EBGP 对等体的标签数据包，除非它实际上已经分发栈顶标签给该对等体了。

如果许多 VPNs 中不同的站点连接到不同的 ASes，在那两个 AS 之间单个 ASBR（仅存在一个 ASBR）不需要拥有所有 VPN 的所有路由；可以有多个 ASBRs，每个 ASBR 只保存一个 VPN 特定子集的的路由。

该过程要求存在一个标签交换路径，从数据包的入口 PE 引导到其出口 PE。因此沿着路径的一组 ASes 之间必须存在适当的信任关系。此外，在一组 SPs 中，必须存在协议，关于边界路由器需要接收哪些 RTs 的路由。

c）多跳 EBGP 在源和目的 AS 之间重分发标签 VPN-IPv4 路由，EBGP 从 AS 到相邻 AS 重新分发的携带标签的 IPv4 路由。

在此规程中，VPN-IPv4 路由既不由 ASBR 维护也不由 ASBR 分发。ASBR 必须维护到其 AS 内 PE 路由器的携带标签的 IPv4/32 路由。它使用 EBGP 将这些路由分发给其他 ASes。任何传输 ASes 中的 ASBR 也必须使用 EBGP 传递携带标签的/32 路由。由此导致从入口 PE 路由器到出口 PE 路由器标签交换路径的被创建。现在不同 AS 中的 PE 路由器可以彼此建立多跳 EBGP 连接，可以通过这些连接交换 VPN-IPv4 路由。

如果每个 AS 的 P 路由器知晓 PE 路由器的/32 路由，一切工作正常。如果 PE 路由器的/32 路由不被 P 路由器所知（ASBR 除外），则此过程需要数据包的入口 PE 将三层标签堆叠放在其上。

栈底标签由出口 PE 分配，对应于报文特定 VRF 中的目标地址。中间标签是由 ASBR 分配，对应到出口 PE 的/32 路由。栈顶标签由入口 PE 的 IGP 下一跳分配，对应到 ASBR 的/32 路由。

为了提高可扩展性，多跳 EBGP 连接仅存在于一个 AS 中的路由反射器和另一个 AS 中的路由反射器之间。（但是，当路由反射器通过这种连接分发路由，但是它们没有修改路由中 BGP 下一跳属性）。实际上 PE 路由器与其 AS 内的路由反射器只有 IBGP 连接。

这个过程与在第 9 节中描述“运营商的运营商”过程非常相似。像上面过程一样，它要求存在引导从数据包的入口 PE 到其出口 PE 的标签交换路径。

11. 从 VPN 接入互联网

许多 VPN 站点将需要能够访问公共互联网，以及访问其他 VPN 站点。下面介绍一些这样做的替代方法。

1. 在某些 VPN 中，一个或多个站点通过连接到 ISP 的非 VRF 接口的“互联网网关”（可能是防火墙）访问互联网。ISP 与提供 VPN 服务的 SP 可能是也可能不是相同的组织。那么到/来（to/from）自互联网网关的流量就可以了根据 PE 路由器的默认转发路由表转发数据。

   在这种情况下，有互联网访问权限的站点可能将默认路由分发给他们的 PEs，其 PEs 依次将重新分发到其他 PEs，从而进入其他站点的 VPN。这为所有 VPN 的站点提供互联网接入。

   为了正确处理来自互联网上的流量，ISP 必须向互联网发送到达 VPN 内中地址的路由。这完全不受任何本文件描述的任何路线分发过程约束。一般来说，不能从互联网上看到 VPN 的内部结构；这样的路由只会引导到连接到 VPN 的互联网网关的非 VRF 接口。

   在这种模式下，PE 路由器的默认转发表及其任何 VRF 之间没有路由交换。VPN 路由分发程序和互联网路由分发程序是完全独立的。

   请注意，虽然 VPN 的某些站点使用 VRF 接口与互联网通信，最终所有到/从（to/from）互联网的数据包离开/进入 VPN 之前遍历非 VRF 接口，所以我们把它称为“非 VRF 互联网”。

   请注意，非 VRF 接口连接的 PE 路由器不一定需要在其默认转发表中维护所有的互联网路由。

   默认转发表可能只有一条路由，“默认”，通向有互联网路由的另一个路由器（可能是相邻的路由器）。该方案的一个变体是通过 PE 路由器到另一个路由器非 VRF 接口接收到隧道报文，其它路由器保持全部的互联网路由集合。

2. 一些 VPN 可以通过 VRF 接口获得互联网访问（“VRF 互联网访问“）。如果 PE 通过 VRF 接口接收到数据包，如果数据包的目标地址没有匹配 VRF 中的任何路由，而可以匹配 PE 的默认转发表。如果匹配完成，那么数据包可以通过本地转发从骨干网到互联网，而不是被 MPLS 转发。

   为了使交通向相反的方向流动（从互联网到 VRF 接口），来自 VRF 的一些路由必须导出到互联网的转发表。不用说，任何这样的路由必须对应全球唯一地址。

   在该方案中，默认转发表可能具有全套互联网路由，或者可能只有一个单个默认路由通往另一个路由器，在其默认转发表中全套互联网路由。

3. 假设 PE 有能力在一个 VRF 存储“非 VPN 路由”。如果数据包的目的地址与“非 VPN 路由“匹配，那么数据包是本地传输的，而不是通过 MPLS 传输。如果 VRF 包含非 VPN 默认路由，公共互联网的所有数据包将匹配它，并被本地转发到默认路由的下一跳。在下一跳，数据包的目标地址将查找默认转发表，并可能匹配更多具体路由。

   这种技术只有在没有 CE 路由器正在分发默认路由的情况下才可用。

4. 还可以通过 VRF 接口（VRF 包含互联网路由）获得互联网访问。与模型 2 相比，这消除了第二次查找，但是它的缺点是需要在每个这样的 VRF 中复制互联网路由。

   如果使用这种技术，SP 可能想接入互联网的接口是一个 VRF 接口，并使用第 4 节技术分发互联网路由，作为 VPN-IPv4 路由，到其他 VRFs。

   应该清楚的是，默认情况下，在 VRF 和默认转发表之间没有路由的交换。只有在客户和 SP 之间达成协议后才能完成，只有在此情况下符合客户的策略。

12. 管理 VPN

本规范不要求子接口连接的 PE 路由器和 CE 路由器是“编号”接口。如果这是编号接口，这个规范允许分配到接口的地址来自 VPN 的地址空间或 SP 的地址空间。

如果 CE 路由器由服务提供商管理，那么服务提供商可能会有一个网络管理系统，其需要能够与 CE 路由器通信。在这种情况下，分配给连接 CE 和 PE 路由器的子接口的地址应该来自 SP 的地址空间，应该在这个空间内是唯一的。网络管理系统本身通过 VRF 接口连接到 PE 路由器（更准确地说，在连接到 PE 路由器的站点上）。网络管理系统的地址将被输出到所有 VRFs，这些 VRFs 与到 SP 管理的 CE 路由器的接口相关联。CE 路由器的地址将被导出到网络管理系统相关的 VRF 中，但不能到任何其他 VRFs。

这允许 CE 和网络管理系统之间的通信，但不允许任何不期望的通信到或在 CE 路由器之间。

确保路由导入/导出正确的完成的一种方法是使用两个路由目标；称它们为 T1 和 T2。如果一个特定的 VRF 接口连接到由 SP 管理的 CE 路由器，则该 VRF 配置为：

- 附加 T1 的 import/导入路由，

- 将 T2 附加到给其 VRF 接口每一端分配的地址。

如果特定的 VRF 接口连接到 SP 的网络管理系统，那么 VRF 被配置为将 T1 附加到该系统地址，并导入附有 T2 的路由。

13. 安全注意事项

13.1 数据平面

在“数据平面”中的安全性，我们的意思是保护以下可能性：

- 从 VPN 内传输到 VPN 之外的站点的数据包，除非与 VPN 的策略一致。
- 从 VPN 外部进入 VPN 的一个站点的数据包，除非与 VPN 的策略一致。

在以下条件下：

1. 一个骨干路由器不接受一个特定数据链接上携带标签的数据包，除非已知数据链接仅接入到受信任的系统，或者除非已知这样的数据包将在 IP 头或堆栈中的任何标签被检查之前都将离开骨干网，并且
2. 不接受来自不可信或不可靠的路由对等体的 VPN-IPv4 路由，
3. 在控制平面上没有成功的攻击，

实际上，这种架构提供的数据平面安全性与通过帧中继或 ATM 骨干网提供的 VPNs 是相同。如果 SP 控制下的设备配置正确，除非获得授权，否则数据不会进入或离开 VPN。

上述条件 1 可以更准确地说明。应该丢弃从特定邻居接收到的携带标签数据包，除非满足以下两个条件中的一个：

- 数据包的栈顶标签具有接收系统分配给该邻居的标签值，或
- 数据包的栈顶标签具有接收系统已经分配给越过该邻居的系统的标签值（即，从系统到标签被分发到的接收系统的路径，可能是经过该邻居）。

上述条件 2 对于跨 VPN 供应商（参见第 10 节）来说是最有意义的。对于跨 VPN 供应商的构建根据第 10 节的方案 b），条件 2 容易检查。（当使用第 10 节方案（c）时的安全性问题进一步研究。）

值得注意的是，使用 MPLS 使其提供数据平面安全性，比尝试使用某种形式的 IP 隧道代替 MPLS 外部标签简单得多。让边界路由器拒绝接受一个标签包是一个简单的事情，除非上述第一项条件适用它。

如果该数据包是 IP 隧道数据包，其目的地址是 PE 路由器的地址。当然这不是不可能做的，但它有管理和表现影响。

[MPLS-in-IP-GRE]规定了 MPLS-in-IP 和 MPLS-in-GRE 隧道。如果希望使用这样的隧道运载 VPN 数据包，然后必须充分理解该文件第 8 节中描述的安全考虑。BGP/MPLS IP VPNs 的任何实现，允许 VPN 报文按照该文件上述方式进行隧道传送，必须在其中包含一个 IPsec 的实现。如果隧道未被 IPsec 保护，那么该文件第 8.2 节所述边界路由器的 IP

地址过滤技术，是确保数据包在特定出口 PE（实际上是放置在隧道中，正确的隧道头节点）处离开隧道的唯一的手段（即，该数据包没有伪造源地址）。由于边界路由器经常只过滤源地址，数据包过滤可能无效，除非出口 PE 可以检查它接收的任何隧道数据包的 IP 源地址，并将其与一系列有效的隧道头地址的 IP 地址列表进行比较。

任何允许没有 IPsec 的情况下使用 MPLS-in-IP 和/或 MPLS-in-GRE 隧道的实现，必须允许出口 PE 以此方式对收到的任何隧道数据包的 IP 源地址进行验证。

在多个 CE 路由器通过 LAN 接口连接到 PE 路由器的情况下，以确保适当的安全性，必须满足以下条件之一：

1. LAN 上的所有 CE 路由器都属于同一个 VPN，或者

2. 可靠且安全的 LAN 交换机将 LAN 划分为多个 VLANs，每个 VLAN 只包含单个 VPN 的系统；在这种情况下，交换机转发任何数据包到 PE 路由器之前，将附加相应的 VLAN 标记到数据包。

此架构不提供密码隐私，也不提供帧中继或 ATM VPN。如果需要的话，这些架构在 CE-CE 的基础上都兼容使用密码。

在 PE-PE 基础上使用密码会进一步研究的。

13.2 控制面

上一节的数据平面安全性取决于控制平面的安全。为了保证安全性，BGP 和 LDP 连接都不应该在不信任的对等体间进行。TCP/IP MD5 身份验证选项[TCP-MD5]都应该应用于两个协议。SP 的网络中的路由协议也应该以类似的方式进行保护。

13.3 P 和 PE 设备的安全性

如果这些设备的物理安全受到威胁，则数据平面安全性也可能受到影响。

应采取通常的步骤，确保来自公共互联网 IP 的流量不能用来修改这些设备配置，或对其进行拒绝服务攻击。

14. 服务质量

虽然不是本文的重点，但服务质量是任何 VPN 服务的关键组件。在 MPLS/BGP VPN 中，现有的 L3 QoS 功能可以通过使用垫片/shim 头[MPLS-ENCAPS]中的“实验/experimental”位应用于标签的数据包，或在 ATM 用作骨干网中使用 ATM QoS 能力。[MPLS-RSVP]中讨论的流量工程工作也直接适用于 MPLS/BGP VPN。交通工程甚至可以在特定站点对之间，用于建立具有特定 QoS 特征的标签交换路径，如果想要的话。在 MPLS/BGP VPN 跨越多个 SPs 的地方，[PASTE]中描述的架构可能很有用。SP 可能应用 intserv（集成服务）或 diffserv（区分服务）功能，适用于特定的 VPN。

15. 可扩展性

我们在本文中讨论了可扩展性问题。在这个部分，我们简要总结了我们的模型在可扩展性方面的主要特征。

服务提供商骨干网由（a）PE 路由器，（b）BGP 路由反射器，（c）P 路由器（既不是 PE 路由器也不是 PE 路由器路由反射器），并且在多供应商 VPN 的情况下，（d）ASBRs。

P 路由器不维护任何 VPN 路由。为了正确转发 VPN 流量，P 路由器只需维护到 PE 路由器和 ASBR 的路由。使用两层标签使得可以将 VPN 路由保持在 P 路由器之外。

PE 路由器维护 VPN 路由，但仅适用于那些直连的 VPN。

路由反射器可以在 VPNs 之间进行分区，以便每个分区仅为该服务供应商支持的 VPNs 的一部分子集 承载路由。因此，不需要单个路由反射器维护所有 VPNs 的路由。

对于提供商间 VPNs，如果 ASBRs 维护和分发 VPN-IPv4 路由，则 ASBR 可以以一种类似的方式在 VPN 中进行分区，结果是不需要单个 ASBR 维护所有提供商间 VPNs 的路由。如果是多跳 EBGP，则 ASBR 不需要维护和分发 VPN-IPv4 所有路由。

因此，服务提供商网络中没有单个组件必须维护所有 VPNs 的所有路由。所以网络支持 VPNs 增长的总能力不受任何单个组件的容量限制。