OSPF 作为 PE/CE 协议

BGP/MPLS IP VPNs

RFC 4577

陶志豪

zhihao.tao@outlook.com

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

目录

Network Working Group
Request for Comments: 4577
Updates: 4364
Category: Standards Track

E. Rosen
P. Psenak
P. Pillay-Esnault
Cisco Systems, Inc.
June 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

OSPF 作为 PE/CE 协议
BGP/MPLS IP VPNs

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

许多服务提供商提供虚拟专用网络（VPN）服务给其客户，使用的技术是客户边缘路由器（CE 路由器）作为提供商边缘路由器（PE 路由器）的路由 peer。边界网关协议（BGP）通过供应商的 IP 骨干网用于分发客户的路由，以及多协议标签交换（MPLS）通过提供商的骨干网用隧道传输客户数据包。这被称为“BGP/MPLS IP VPN“，BGP/MPLS IP VPN 的基本规范假设 PE 路由器和 CE 路由器之间接口上的路由协议是 BGP。本文档扩展了该规范，通过允许 PE/CE 接口的路由协议为开放最短路径优先（OSPF）协议。

本文档更新 RFC 4364。

1. 介绍

[VPN]描述了一种服务提供商（SP）可以使用它的 IP 骨干网提供 VPN 服务给顾客的方法。在这种方法中，客户的边缘设备（CE 设备）连接到提供商的边缘路由器（PE 路由器）。如果 CE 设备是路由器，那么 PE 路由器可能成为 CE 路由器的路由 peer（在某些路由协议中），作为结果，并可能学习通往 CE 站点的路由，且需要分发连接到相同 VPN 的其他 PE 路由器的路由。

连接到公共 VPN 的 PE 路由器使用 BGP 将 VPN 路由分发给彼此。然后一个 CE 路由器可以通过其毗邻的 PE 路由器用路由协议来学习到 VPN 中其他站点的路由。不过，不同站点的 CE 路由器并不彼此毗邻。

可以预料，许多 VPNs 将使用 OSPF 作为他们的 IGP（内部网关协议），即在那个网络里用于分发一个网络内部路由的路由协议。这并不一定意味着 PE 路由器需要使用 OSPF 与 CE 路由器成为 peer。在一个 VPN 里的每个站点可以使用 OSPF 作为其站内路由协议，同时使用，例如 BGP [BGP]或 RIP（路由信息协议）[RIP]分发路由到 PE 路由器。但是，当 OSPF 在站内使用时，最方便的肯定是在 PE-CE 链接上使用它，[VPN]明确允许这样做。

像其他任何协议一样，在 PE-CE 链路上使用 OSPF 具有优势和劣势。在 PE-CE 链路上使用 OSPF 的缺点是 SP 的 PE 路由器，而在外围，参与了一个 VPN 站点的 IGP。虽然优点是：

- CE 路由器的管理员不需要任何除 OSPF 之外的任何路由协议中的专业知识。
- CE 路由器不需要支持 OSPF 以外的任何路由。
- 如果一个客户正在转变其网络，从一个传统的使用 OSPF 骨干网络到[VPN]描述的 VPN 服务时，使用 PE-CE 链路上的 OSPF 简化了过渡问题。

一些 SPs 和他们的客户在赞同 PE-CE 链路上使用 OSPF 似乎可能会做出权衡。因此，为了使这成为可能，我们需要指定 PE 路由器必须执行的规程。（不过在 CE 路由器中不需要特别的规程；CE 路由器只运行他们可能有的任何 OSPF 的实现）。

2. 要求规范

关键词“必须”，“不得”，“所需”，“已”，“不”，“应该”，“不应该”，“推荐”，“可能”和“可选”在文档[RFC2119]中所述。

3. 要求

考虑一组被认为是相同“OSPF 域”的 VPN 站点。两个站点被认为是在同一个 OSPF 域中，如果它打算从一个站点到另一个站点的路由被考虑为网内路由。在相同域中的一组 OSPF 站点，几乎肯定的是一个站点集合以前构成一个“内部网”，每个内部网都运行 OSPF 作为它的站内路由协议。

根据[VPN]，VPN 路由通过 PE 路由器使用 BGP 分发。如果 PE 使用 OSPF 分发路由给 CE 路由器，管理 BGP/OSPF 交互的标准规程[OSPFv2]，使路由在 5 类 LSA（链接状态公告）中从一个站点被交付到另一个站点，作为“AS 外部”路由。这是不合需要的；在 3 类 LSA（作为区域间路由）中交付这样的路由会好得多，因此它们可以区分，任何可能在 VPN 中

流通的“真实的”AS 外部路由（也就是说，所以它们可以通过 OSPF 区分真正不是来自 VPN 内的路由）。因此，PE 路由器必须实现 BGP/OSPF 的交互规程的改良版本。

实际上，我们希望有一套非常通用的规程允许客户轻松替换传统的私有 OSPF 骨干网为 VPN 服务。我们希望这个规程能够满足以下要求：

- 规程不应该对 OSPF 拓扑结构做出假设。特别是，不应该假定客户站点是 OSPF 存根站点或 NSSA（Not So Stubby Area）站点。也不应该假设客户站点只包含一个 OSPF 区域，或者它没有区域 0 路由器。
- 如果 VPN 站点 A 和 B 在同一个 OSPF 域中，则路由从一个站点应该被传递给另一个站点是作为 OSPF 网内路由。一般来说，这可以通过作为 3 类 LSA 的区域间路由传递这样的路由来完成。

  请注意，这允许两个 VPN 站点通过一个“OSPF 后门链路”连接。也就是说，这两个站点之间可以有一个 OSPF 链接，只有当 VPN 骨干网不可用时才能使用。（这对普通 BGP/OSPF 交互规程来说是不可能的。普通规程会使用 AS 外部路由通过 VPN 主干网传递路由，这些永远不会成为网内路由的首选）。在从遗留 OSPF 骨干网转换到 VPN 骨干网的期间内可能是非常有用。
- 很可能两个站点间使用“OSPF 后门链路”，即使两个站点在同一个 OSPF 区域中，且连接站点之间后门链路的路由器不是一个区域 0 路由器。这在过渡期间是非常有用的，并且消除了任何需要将站点的路由器重新配置为 ABRs（区域边界路由器）。

  假设希望有通过 VPN 的主干网的路由优于后门路由，VPN 骨干必须呈现给每个站点的 CE 路由器，VPN 骨干作为 CE 路由器分别连接的两个 PE 路由器之间一条链路。
- 连接到 VPN 服务 PE 路由器的 CE 路由器可以它们本身作为 OSPF 主干（区域 0）路由器。一个 OSPF 主干甚至可能由几个只能通过 VPN 服务相互连接的“片段”组成。在这样一个场景，连接到 OSPF 骨干的不同部分的站点之间的完全互通仍然应该是可能。
- 从传统私有 OSPF 骨干网到 VPN 服务的转换必须简单直接。过渡是可能会被分阶段，这样的客户站点一个接一个的由传统的专用 OSPF 骨干网迁移向 VPN 服务。在过渡期间，任何给定的站点可能是连接到 VPN 服务，也可能是连接到传统的 OSPF 主干网，或者二者皆是。所有这些站点之间的完整连接必须被保持。

  由于 VPN 服务是要取代传统的骨干网，必须能够通过适当调整 OSPF 度量，使 OSPF 更喜欢穿越 SP 的 VPN 骨干网的路由，而不是没有替代的路由。
- 分配给给定路由的 OSPF 度量值应透传过 VPN 骨干网。

来自不同 OSPF 域的站点的路由将显示为 AS 外部路由。

我们假设您熟悉[OSPFv2]的内容，包括 OSPF LSA 类型，并且将不进一步解释而引用的 1，2，3 类等 LSA。熟悉[VPN]也是先决条件。

4. PE 路由器的 BGP/OSPF 交互过程

4.1 概述

4.1.1 VRFs 和 OSPF 实例

连接到多个 OSPF 域的 PE 路由器必须为每个域独立的运行一个 OSPF 实例。如果 PE 正在运行 OSPF 作为其 IGP（内部网关协议），即作为 IGP 运行 OSPF 的实例必须是独立的，独立于任何其他 PE 正在运行的 OSPF 实例。（不论这些实例是作为单独的进程来实现，或者仅仅作为在一个公共的进程的分离的上下文，就是一个实现的事情）。连接到一个 VPN 站点的每个接口不能属于多个 OSPF 实例。

[VPN]定义了“每个站点路由和转发表”的概念，或 VRF。每个 VRF 都与一组接口相关联。如果一个 VRF 与特定的接口相关联，并且该接口属于到一个特定的 OSPF 实例，那么 OSPF 实例就是这个与 VRF 相关联。如果两个接口属于同一个 OSPF 实例，那么这两个接口必须与相同的 VRF 关联。

如果一个接口将一个 PE 连接到一个 CE 上，并且该接口与 VRF 相关联，我们会说 CE 关联这个 VRF。

4.1.2 VRFs 和路由

OSPF 用于从 CE 发布路由到 PE。标准 OSPF 的决策过程用来安装最优的 OSPF 分发的路由在 VRF 中。

根据[VPN]，BGP 用于在 PE 路由器间分发 VPN-IPv4 路由。安装在 VRF 中的 OSPF 路由可能被“导出”，作为 VPN-IPv4 路由重新分发到 BGP。那可能是由 BGP 分发给其他 PEs。在其他 PEs 上，VPN-IPv4 路由可以由 VRF“导入”，然后可以被重新分发进一个或多个与该 VRF 关联的 OSPF 实例。

导入和导出特定的 VRF 由路由目标拓展团体属性的使用来控制（或者更简单地说，路由目标或 RT），如[VPN]中所述。

如果 VPN-IPv4 路由的 RT 等于 VRF 的导入 RT 中的一个，则其“符合导入”到特定的 VRF 的条件。标准的 BGP 决策过程用于从在有资格导入的路由中选择一组 VPN-IPv4 路由“安装”在 VRF 中。

如果 VRF 中同时包含对相同的 IPv4 前缀的 OSPF 分发的路由和 VPN-IPv4 路由，则 OSPF 分发的路由是首选。一般来说，这意味着转发是根据到 OSPF 路由完成的。这个规则的一个例外是“sham link”。如果安装路由的下一跳接口（OSPF-分发的）是 sham link，转发是按照一条对应的 BGP 路由完成的。这在第 4.2.7.4 节中详细描述。

为了满足第 3 节的要求，PE 安装了一个特定路由到特定 VRF，需要知道这条路由最初是否是一个 OSPF 路由，如果是的话，引入进 BGP 的 OSPF 实例是否与该路由重新分发到的 OSPF 实例是相同的域。因此，域标识符被编码为 BGP 拓展团体属性[EXTCOMM]，通过 BGP 伴随 VPN-IPv4 路由进行分发。路由的 OSPF 度量和 OSPF 路由类型也作为路由的 BGP 属性被携带。

4.1.3 区域间，区域内和外部路由

如果一个 PE 安装一个特定的 VPN-IPv4 路由（通过 BGP 学习）到 VRF 中，如果这是相应的 IPv4 前缀首选的 BGP 路由，那么相应的 IPv4 路由就是“符合条件重新分发”到每个与该 VRF 相关联的 OSPF 实例中。因此，可以在 LSA 中通告给每个 CE。

一条路由是否有资格重新分发到 OSPF（实际上重新分发到一个特定的 OSPF 实例）可能依赖于配置。例如，PE 可以配置为只将缺省路由分布到给定的 OSPF 实例中。在这种情况下，有资格进行再分发的路由实际上不会被重新分发。

在下面，我们讨论重新分发一个 BGP 分发的 VPN-IPv4 路由到 OSPF 规程；无论何时只要这样的路由有资格重新分发进 OSPF 或者配置不阻止这种重新分发，就应该遵守这些规程。

如果路由来自与 OSPF 将要被重新分发的实例是不同的 OSPF 域，或者路由不是来自一个 OSPF 域，那么这个路由被认为是一个外部路由。

如果路由来自与 OSPF 将要被重新分发的实例是相同的 OSPF 域，如果它最初通告作为 OSPF 外部路由或 OSPF NSSA 路由到 PE，其被视为外部路由。按照正常的 OSPF 规程，外部路由会通过 5 类的 LSA 或 7 类 LSA 发布给 CE，或根本不发送，其取决于 PE/CE 链路属于区域的类型。

如果路由来自与 OSPF 将要被重新分发的实例是相同的 OSPF 域，如果它最初通告作为区域间和区域内的路由到 PE，路由通常将被作为区域间路由（3 类 LSA）通告给 CE。

作为特殊情况，假设 PE1 连接到 CE1，而 PE2 连接到 CE2，其中：

- 包含 PE1-CE1 链路的 OSPF 实例和包含 PE2-CE2 链路的 OSPF 实例在同一个 OSPF 域中，
- PE1-CE1 和 PE2-CE2 链路在同一个 OSPF 区域 A（由 OSPF 区域号的配置决定），

那么 PE1 可能向 CE1 泛洪一条 1 类 LSA，其发布到 PE2 的链路；PE2 可能向 CE2 泛洪一条 1 类 LSA，其发布到 PE1 的链路。该在这些 LSAs 中通告的链路被称为“sham link/假链路”，并且是通告作为区域 A 中的链路。这使得它看起来区域 A 内的路由器，就好像 CE1 到 PE1，然后穿过服务提供商网络到 PE2，再到 CE2 的路径一样，是区域内路径。Sham links 是本规范的一个可选功能，仅在需要把服务提供商的网络视为一个区域内的链路时才使用。Sham link 的进一步细节参见 4.2.7 节。

在 4.2.8 节中规定了，PE 决定用于通告通向 CE 的特定路由的 LSA 类型的确切细节。请注意，如果 VRF 与多个 OSPF 实例相关联，用于通告路由的 LSA 的类型在不同的实例中可能会有所不同。

请注意，如果 VRF 与多个 OSPF 实例相关联，则给定路由可以被重新分发到这些 OSPF 实例中的某些或全部，取决于每个实例的特征。如果重新分发进两个或两个以上的 OSPF 实例，每个实例都可以使用不同类型的 LSA 来通告，同样取决于每个实例的特征。

4.1.4 PE 和 OSPF 区域 0

在给定的 OSPF 域内，PE 可以连接到多个 CE。每个 PE/CE 链路被分配（通过配置）到一个 OSPF 区域。任何链路可以分配到任何区域，包括区域 0。

如果一个 PE 通过一个非零区域的链路连接到一个 CE，那么 PE 作为该区域的 ABR。

因此 PE 可以被认为是 OSPF 的“区域 0 路由器”，即它们可以是被认为是“OSPF 骨干”的一部分。因此，他们被允许通过 3 类 LSA 向 CE 发布区域间路由。

如果 OSPF 域有除了 PE 路由器以外的区域 0 路由器，那么他们中至少有一个必须是一个 CE 路由器，并且必须有一条区域 0 的链路到至少一个 PE 路由器。这个邻接可以通过 OSPF 虚拟链接。（以这种方式，使用 OSPF 虚拟链路的能力是一个可选功能）。必须确保区域间路由和 AS 外部路由可能在 PE 路由器和非 PE 的骨干网之间泄漏。

两个不在同一个 OSPF 区域的站点，将会理解 VPN 骨干网作为 OSPF 骨干的一个组成部分。但是，如果有区域 0 路由器不是 PE 路由器，那么 VPN 骨干实际上是一种更高层次的骨干，区域 0 以上提供三级层次。传统的 OSPF 骨干网将在转换期间断开连接，只要各个部分都连接到 VPN 骨干。

4.1.5 预防环路

如果从 PE 路由器发送到 CE 路由器的路由，然后可以由另一个 PE 路由器从它自己的一个 CE 路由器接收，将有可能发生路由环路。为了防止环路，PE 在发往 CE 的任意 LSA 上配置 DN[OSPF-DN]，且 PE 忽略从 CE 接收到的已经设置 DN 位的任何 LSA。较早的实现可能使用 OSPF 路由标记，而不是 DN 位，在某些情况下，见 4.2.5.1 和 4.2.5.2 节。

4.2 细节

4.2.1 PEs 中独立的 OSPF 实例

PE 必须为其连接的每个 OSPF 域支持一个 OSPF 实例。这些 OSPF 实例功能是独立的，不会互相泄漏彼此的路由。每个 OSPF 实例必须关联一个单独的 VRF。如果 n 个 CEs 与 VRF 相关联，此 VRF 在他们各自的 PE/CE 链路上正在运行 OSPF，那么这 n 个 CEs 与对应 OSPF 实例中 PE 是 OSPF 邻接关系。

一般来说，虽然不一定，如果 PE 连接到在同一个 OSPF 域中几个 CEs，它将把这些接口关联到那些具有单个 VRF 的 PEs。

4.2.2 路由器 ID

如果 PE 和 CE 通过 OSPF 进行通信，则 PE 具有的 OSPF 路由器 ID 在 OSPF 域内有效（即唯一）。更确切地说，每个 OSPF 实例都有一个路由器 ID。不同的 OSPF 实例可能有不同的路由器 ID。

4.2.3 OSPF 区域

PE-CE 链路可以在任何区域，包括区域 0；这是一个 OSPF 配置的问题。

如果 PE 有链路属于非零区域，则 PE 作用是作为该区域的区域边界路由器（ABR）。

PEs 不会将链路状态拓扑从一个站点传递到另一个（除非使用 sham link，请参见章节 4.2.7）。

根据[OSPFv2，Section 3.1]，“OSPF 骨干总是包含全部区域边界路由器”，因此 PE 路由器被认为是区域 0 路由器。[OSPFv2]的第 3.1 节也要求区域 0 是连续的。因此，如果 OSPF 域有任何除了 PE 路由器之外的其他区域 0 路由器，那么他们中至少有一个必须是一个 CE 路由器，并且必须有一条区域 0 的链路（可能是虚连接）到至少一个 PE 路由器。

4.2.4 OSPF 域标识符

每个 OSPF 实例必须与一个或多个域标识符相关联。这必须是可配置的，并且默认值（如果没有被配置）应该是 NULL。

如果一个 OSPF 实例有多个域标识符，其中之一被认为是其“主”域标识符；这必须可以通过配置确定。如果一个 OSPF 实例只有一个域标识符，理所当然的是它的主域标识符。如果一个 OSPF 实例有多个域标识符，即 NULL 域标识符不能是其中之一。

如果一个路由被特定的 OSPF 实例安装在一个 VRF 中，该 OSPF 实例的主域标识符被认为是路由的域标识符。

考虑由 OSPF 实例 I1 安装路由 R 在 VRF 中，然后作为 VPN-IPv4 路由重新分发到 BGP，然后通过 BGP 安装在另一个 VRF 中。如果 R 需要重新分发到与后面的 VRF 相关联的 OSPF 实例 I2，其中 R 在 I2 中的通告的方式将取决于 R 的域名标识符是否是 I2 的域名标识符之一。如果 R 的域名标识符不是一个 I2 的域标识符，那么，如果 R 被重新分发到 I2，R 将被发布为 AS 外部路由，不管它的 OSPF 路由类型是什么。另一方面，如果 R 的域名标识符是一个 I2 的域标识符之一，R 如何被公告将取决于 R 的 OSPF 路由类型。

如果两个 OSPF 实例在同一个 OSPF 域中，则：

1. 他们都有 NULL 域标识符，或者
2. 每个 OSPF 实例都有主域标识符作为其他实例的域标识符之一。

如果两个 OSPF 实例在不同的 OSPF 域中，则：

3. 他们都有 NULL 域标识符，或者
4. 没有 OSPF 实例主域标识符作为其他实例的域标识符之一。

（请注意，如果两个 OSPF 实例都有 NULL 域标识符，我们无法从域标识符中知道他们是否在同一个 OSPF 域中。如果他们在不同的域，如果从一个路由分配到另一个，路由将显示为网内路由，这可能不是预期的）。

域标识符是八字节，是一个有效的 BGP 扩展团体属性，如第 4.2.4 节所述。如果一个特定的 OSPF 实例具有非空的域标识符时，来自该 OSPF 实例的路由被 BGP 分发为 VPN-IPv4 路由，这些路由必须携带域标识符与 OSPF 实例的主域标识符对应的扩展团体属性。如果 OSPF 实例的域标识符为 NULL，当来自该 OSPF 实例的由 BGP 分发时路由，域标识符扩展社区属性可以省略；或者，也可以由携带在路由中域标识符扩展团体属性的值代表 NULL（见第 4.2.4 节）。

如果一个 OSPF 域的 OSPF 实例被赋予一个或多个非 NULL 域标识符，这个规程允许我们确定一个特定的 OSPF 发起的 VPN-IPv4 路由是否属于，与给定的 OSPF 实例相同的域。然后我们可以确定是否应将该路由作为区域间路由或作为 OSPF AS 外部路由，重新分给该 OSPF 实例。详情可以在 4.2.4 节和 4.2.8.1 节找到。

4.2.5 环路预防

4.2.5.1 DN 位

当从 PE 路由器发送 3 类 LSA 到 CE 路由器时，必须设置 LSA 选项字段中的 DN 位[OSPF-DN]。这是用来确保 CE 路由器将 3 类 LSA 发送给 PE 路由器后，PE 路由器不会将其进一步重新分发回 CE。

当 PE 路由器需要向 CE 路由器分发一条，来自后者的 OSPF 域之外的一个站点路由时，PE 路由器呈现为 ASBR（自治系统边界路由器），并且在 5 类 LSA 中分发路由。在这些 LSA 中，DN 位[OSPF-DN]必须被设置，以确保它们将被其他任何接收它们的 PE 路由器忽略。

有部署的实现不设置 DN 位，而是使用 OSPF 路由标记来确保由 PE 路由器生成的 5 类 LSA，将被任何其他可能收到的它 PE 路由器忽略。一个用于此目的特殊的 OSPF 路由标记，我们称之为 VPN 路线标记（见第 4.2.5.2 节）。为了确保向后兼容性，所有的实现都坚持这个规范，必须默认支持 VPN 路由标记过程在 4.2.5.2,4.2.8.1 和 4.2.8.2 节中的规定。当在特定的部署中不再需要使用 VPN 路由标记，它的使用（发送和接收）可能被配置禁用。

4.2.5.2 使用 OSPF 路由标记

如果 PE 中的特定 VRF 与一个 OSPF 实例相关联，那么默认情况下，它必须配置一个特殊的 OSPF 路由标记值，我们称之为 VPN 路由标记。默认情况下，这个路由标记务必包含在 PE 发起的 5 类 LSAs 中（如接收 BGP 分发的 VPN-IPv4 路由的结果，请参见 4.2.8 节），并发送给任何连接的 CEs。

配置和包含 VPN 路由标记需要向后兼容未设置 5 类 LSA 中的 DN 位的部署实现。包含的 VPN 路由标记可能通过配置被失效，如果向后兼容不再需要。

VPN 路由标记的值是任意的，但必须不同于 OSPF 域内使用任何 OSPF 路由标记。因此，它的值必须是可配置的。如果 VPN 骨干的自治系统编号长度是两个字节，默认值应该是一个基于该自治系统编号自动计算的标签：

```
Tag = <Automatic = 1, Complete = 1, PathLength = 01>
 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 3 3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|1|1|0|1|     ArbitraryTag       |       AutonomousSystem        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

 1 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 _AS number of the VPN Backbone_
```

如果自治系统号码是四个字节长，那么一个路由标记的值必须被配置，并且必须与在 VPN 本身使用的任何路由标签不同。

如果一台 PE 路由器需要使用 OSPF 向 CE 路由器，分发一条来自 CE 路由器的 OSPF 域之外站点的路由，即 PE 路由器应该把自己作为一个 ASBR 呈现给 CE 路由器，应该报告这些路由为 AS 外部的路由。也就是说，这些 PE 路由器发起 5 类 LSA 报告域外路由当作 AS 外部路由。每个这样的 5 型 LSA 必须包含一个 OSPF 路由标记，其值是 VPN 路由的标记。该标识将路由认为来自 PE 路由器。VPN 路由标记务必用于确保，由 PE 路由器发起的一个 5 类 LSA 不通过 OSPF 区域重新分发给另一个 PE 路由器。

4.2.5.3 其他可能的循环

本文档中描述的规程确保了，从 BGP 分发的 VPN-IPv4 路由派生的路由信息是被分布到 OSPF 中，不能作为 VPN-IPv4 路由重新分发回 BGP，只要是在 OSPF 域内维护 DN 位和/或 VPN 路由标记。这并没有消除所有可能的环路源头。例如，如果一条 BGP VPN-IPv4 路由是分发到 OSPF，然后分发到 RIP（所有的防止循环所需的信息丢失），然后分发回到 OSPF，那么它可能会被分发回去作为 VPN-IPv4 路由进入 BGP，造成一个环路。

因此，如果在 OSPF 域和第三方路由域（即，不是 VPN 骨干）之间有任何路由的相互重新分发，必须非常小心。如果第三方路由域是一个 BGP 域（如公共互联网），普通的 BGP 环路预防措施将防止路由重新进入 OSPF 域。

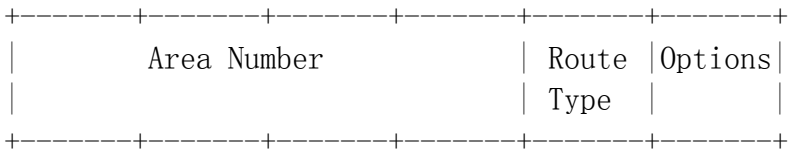
4.2.6 处理来自 CE 的 LSA

本节规定了 PE 路由器处理从 CE 路由器收到的 OSPF LSAs 的方式。

当 PE 路由器从 CE 路由器接收到设置了 DN 位[OSPF-DN]的任何 LSAs 时，来自该 LSA 的信息不能被用于路由计算。如果从 CE 接收到 5 类 LSA，并且如果它具有与 VPN 路由标记相同的 OSPF 路由标记值（请参阅第 4.2.5.2 节），则不得使用来自该 LSA 的信息进行路由计算。

否则，PE 必须检查相应的 VRF。对于每一个地址前缀，由其关联的 OSPF 实例安装在一个 VRF 中的，PE 必须在 BGP 中创建 VPN-IPv4 路由。每个这样的路由将具有以下一些扩展团体属性：

- OSPF 域标识符扩展社区属性。如果 OSPF 实例安装了具有非 NULL 值的主域标识符的路由，这必须存在；如果那个 OSPF 实例只有一个 NULL 域标识符，可以省略。这个属性是用一个两字节的类型字段编码的，类型是 0005,0105 或 0205。为了向后兼容，8005 型也可以使用，并被视为 0005。如果 OSPF 实例有一个 NULL 域标识符，且 OSPF 域标识符扩展团体属性是存在的，属性的值域必须全为零，其类型字段可能是 0005,0105,0205 或 8005 中的任何一个。
- OSPF 路由类型扩展路由属性。这个属性必须存在。它是用一个双字节的类型字段编码的，它的类型是 0306。为了确保向后兼容，类型 8000 也应该被接受，并被视为类型 0306。属性的其余六个字节被编码为如下：



* 区域号码：4 个字节，编码一个 32 位区域号。对于 AS 外部路由，值为 0。非 0 值将路由标识为 OSPF 域内的路由，并在所标识的区域内。区域号码相对于一个特定的 OSPF 域。

* OSPF 路由类型：1 字节，编码如下：

  ** 1 或 2 的区域内路由（取决于路由是否来自 1 类或 2 类 LSA）。
  ** 3 为区域之间的路由。
  ** 5 为外部路由（区号必须为 0）。
  ** 7 为 NSSA 路由。

  请注意，4.2.8 节的规程，路由类型 1,2 和 3 的处理不做任何区分。如果 BGP 安装 VRF 中的这些类型之一的路由，且如果该路由被选择重新分发到 OSPF，它将被 OSPF 公告在 3 类或 5 类 LSA 中，具体取决于域标识符。

* 选项：1 个字节。目前 ，只用于路由类型是 5 或 7。设置最低有效位字段表示该路由携带 2 类度量值。

- OSPF 路由器 ID 扩展团体属性。这个可选属性指定系统的 OSPF 路由器 ID，其在 BGP 下一跳属性中标识。更确切地说，它指定 OSPF 实例中 PE 的 OSPF 路由器 ID，其将路由安装到该路由要输出的所在的 VRF 中。该属性使用两字节类型字段进行编码，类型为 0107，在值字段的前 4 个字节中携带路由器 ID。类型 8001 应该被接受，以确保向后兼容性，应该被看作是 0107。

- MED/Multi_EXIT_DISC 属性。默认情况下，这应该设置为与该路由关联的 OSPF 距离的值，加 1。

这一切的意图如下。从一个站点收到的 OSPF 路由被转换为 BGP，在 VPN 骨干网上分发，在分发到另一个站点之前，可能转换回 OSPF 路由。通过这些属性，BGP 可以传送关于该路由的足够信息，使路由能够“透明”转换回 OSPF 的路由，就好像没有涉及到 BGP 一样。

PE 在 4 类 LSA 中接收到的路由不能重新分发给 BGP。

上面指定的属性之外的任何其它的属性,其必须按照[VPN]被路由携带。

[VPN]通常需要的“始发站点”属性，对于 PE 通过 OSPF 从 CE 学习的路由是可选的。

在多个宿主站点（即，连接到多个 PE 路由器的站点）的情况下，使用站点起源属性阻止站内路由从 VPN 骨干重新注入到站点。这样的回注不会损害路由，因为经由 VPN 骨干的路由将在 3 类 LSA 中被通告，并且因此似乎是一个区域间的路由；真正的区域内路由将是首选。但是引进不必要的开销。另一方面，如果站点起源属性不被使用，分区的站点会发现自动修复，因为从一个分区到另一个分区的流量将自动通过 VPN 骨干网传输。因此，使用站点起源属性是可选的，这样就可以在增加成本开销和分区自动修理的价值之间进行折衷了。

4.2.7 Sham Links

本节介绍了支持本文定义的“Sham Links”必要的协议和规程。支持 Sham Links 是本规范的可选功能。

4.2.7.1 区域内路由

假设在同一个 OSPF 区域有两个站点。每个站点连接到不同的 PE 路由器，并且存在连接两个站点的 OSPF 区域内链路。

有可能把这两个站点当作一个 VPN 站点恰好是多宿主的骨干。事实上最简单的事情是，只要这两个站点之间的首选路由是通过区域内的 OSPF 链路（“后门链接”），而不是通过 VPN 骨干，这就完全足够了。存在通过 PE 路由器的站点间的路由，但这些路由似乎是区域间的路线，而 OSPF 将会考虑它们比通过后门链路的区域内路由具有较少偏好度。

如果希望让 OSPF 更喜欢通过骨干网的路由而不是通过后门链路的路由，然后通过主干的路由必须看起来是区域内路线。为了使一个通过骨干的路由似乎是一个区域内的路线，有必要使它看起来好像有一个区域内的链接存在于连接的两台 PE 路由器之间。这就是我们所说的“sham link“（如果这两个站点连接到同一个 PE 路由器，这是当然没有必要）。

sham link 可以被认为是两个 VRFs 之间的关系。如果两个 VRFs 要通过 sham link 进行连接，每个 VRF 必须关联一个“Sham Link 末端地址”，一个 32 位的 IPv4 地址，其被视为包含该 VRF 的 PE 路由器的地址。Sham Link 末端地址是 VPN 地址空间中的地址，而不是 SP 的地址空间。与 VRF 关联的 Sham link 端点地址必须是可配置的。如果 VRF 仅与单个 OSPF 实例关联，以及该 OSPF 实例中的 PE 的路由器 ID 是一个 IP 地址，然后是 Sham Link 末端地址默认为该路由器 ID。如果 VRF 与几个 OSPF 关联，每个 Sham Link 都属于一个单独的 OSPF 实例。

对于给定的 OSPF 实例，VRF 只需要一个 Sham Link 末端地址，不管它有多少 Sham Links。Sham Link 末端地址必须由 BGP 作为 VPN-IPv4 分发，其 IPv4 地址前缀部分长度为 32 位的地址。Sham Link 末端地址不能被 OSPF 发布；如果没有到 Sham Link 末端地址的 BGP 路由，该地址似乎无法访问，所以 Sham Links 似乎是关闭的。

4.2.7.2 创建 Sham Links

Sham Links 是手动配置的。

为了在两个 VRFs 之间存在 Sham Links，每个 VRF 必须配置创建到其它 VRF 的 Sham Links，通过 Sham Links 末端地址来标识。同一对 sham link 末端地址仅仅有一个 Sham Link 将会被创建。本规范不包括单端 sham link 手动配置的规程。

请注意，可以为任何区域（包括区域 0）创建 sham link。

当且仅当 VRF 中已经安装了到 32 位远程末端地址的路由，连接两个 VRFs 的 sham link 被认为是 UP。

sham link 末端地址不能用作 OSPF 虚拟链路末端地址。

4.2.7.3 Sham Links 上的 OSPF 协议

从一个 PE 到另一个 PE 的 Sham Link 上发送的 OSPF 协议报文，必须有作为其 IP 源地址的发送者的 Sham Link 末端地址，以及作为其 IP 目的地址的接收者 Sham Link 末端地址。数据包将从一个 PE 路由器通过 VPN 骨干传送到另一个 PE，这意味着它可以遍历多跳。因此，其 TTL（生存时间）字段必须被适当设置。

一个 OSPF 协议报文被认为是在一个特别 sham link 上收到的，唯有如果满足以下三个条件：

- 数据包作为一个 MPLS 数据包到达，以及它的 MPLS 标签栈使其“交付/delivered”到本地的一个 sham link 末端地址。
- 数据包的 IP 目的地址是本地一个特别 sham link 末端地址。
- 数据包的 IP 源地址是远程 sham link 末端地址。

sham link 应该被 OSPF 视为 OSPF 按需电路。这意味着 LSAs 将在其上面泛洪，但定期刷新流量被避免的。请注意，只要后门链路启动/UP，在 sham link 上泛洪 LSAs 是没有用的。然而，如果后门链路断开/DOWN，OSPF 没有机制，来使能一个站点中的路由器能够快速清除从另一个站点中收到的 LSAs。因此，仍然需要在这两个站点的 LSA 数据库保持同步，因此会在 sham link 上泛洪。

sham link 是一个无编号的点到点的区域内连接，在 1 类 LSA 中作为 1 类链路通告。

与 sham link 相关的 OSPF 度量必须是可配置的（和必须有一个可配置的默认值）。站点之间是否有流量通过后门链路或通过 VPN 骨干网（即通过 sham link）取决于 OSPF 链路度量的设置。该度量值可以被设置，以便后门链接不被使用，除非例如，通过 VPN 主干连接失败。

sham link 默认的 Hello Interval 是 10 秒，而 sham link 默认路由器死时间间隔是 40 秒。

4.2.7.4 在 Sham 链路上路由和转发

如果 PE 确定某个特定路由的下一跳接口是一个 sham link，那么 PE 路由器不应该把这个路由重新分发到 BGP 作为 VPN-IPv4 路由。

在 LSA 中通告的任何其他路由通过 sham link 传输时也必须重新分发到 BGP（通过 PE 在 sham link 上泛洪 LSA）。这意味着如果对于给定地址前缀的首选（OSPF）路由具有 sham link 作为其下一跳接口，则也会存在一个相同地址前缀的“相应的 BGP 路由”安装在 VRF 中。根据第 4.1.2 节，OSPF 路由是首选。但是，在转发数据包时，如果该数据包的首选路由具有 sham link 作为其下一跳接口，那么这个包必须按照对应的 BGP 路由转发。也就是说，它会像首选路由是相应的 BGP 路由一样被转发。该“相应的 BGP 路由”始终是 VPN-IPv4 路由；在[VPN]中描述了通过 VPN-IPv4 路由转发数据包的规程。

同样的规则适用于任何 IP 目的地址为一个 sham link 远程末端地址的数据包。这样的数据包必须是按照相应的 BGP 路由进行转发。

4.2.8 通过 BGP 接收的 VPN-IPv4 路由

本节介绍 PE 路由器如何处理通过 BGP 接收 VPN-IPv4 路由。

如果收到的 BGP VPN-IPv4 路由没有安装在 VRF 中，没有什么会向 CE 汇报。收到的路由将不会被安装到 VRF，如果 BGP 决策过程把其他路由看作是优选的。当安装在 VRF 中时，路由似乎是一个 IPv4 路由。

安装在 VRF 中的 BGP 路由不一定用于转发。如果同一个 IPv4 地址前缀的 OSPF 路由已经安装在 VRF 中，将使用 OSPF 路由转发，除了 OSPF 路由的下一跳接口是一个 sham 的情况。

如果使用安装在 VRF 中的 BGP 路由进行转发，BGP 路由被重新分配到 OSPF，并可能在 OSPF LSA 中报告给 CE。LSA 的种类，如果有的话，取决于 BGP 路由的各种特征，详见本文件的后续部分。

通过 VPN-IPv4 路由转发报文的规程如[VPN]中所述。

在下文中，我们描述 PE 到 CE 使用 OSPF LSA 报告的内容，假设在向 CE 报告路由之前，PE 没有配置做任何进一步汇总或过滤路由信息。

当向 CE 发送 LSA 时，可能需要设置 DN 位。 有关 DN 位的规则，请参见第 4.2.5.1 节。

当向 CE 发送 LSA 时，可能需要设置 OSPF 路由标记。有关设置 OSPF 路由标记的规则，请参见第 4.2.5.2 节。

当发送 5 类 LSA 时，转发地址被设置为 0。

4.2.8.1 外部路由

对于与 VRF 相关的特定 OSPF 实例，VPN-IPv4 路由被安装在 VRF 中，然后选择为首选路由，如果满足以下条件成立其中一个，则被视为外部路由：

- OSPF 路由类型扩展团体的路由类型字段有“external”OSPF 路由类型。
- 路由来自与 OSPF 实例不同的域。

确定来自一个域的路由和一个特定的 OSPF 实例是否具有不同的规则，如下所示。由路由携带的 OSPF 域标识符扩展团体属性与 OSPF 实例中已配置的 OSPF 域标识符扩展团体属性（若有的话）进行比较。一般来说，当比较两个这样的属性时，全部八个字节必须进行比较。因此，两个 OSPF 域标识符扩展团体属性，当且仅当以下三个条件其中之一成立时，被认为是相等的：

1. 它们在所有八个字节中是相同的。
2. 它们在低六字节（值字段）是相同的，但是一个属性的两个高位字节（类型字段）是 0005，另一个的两个高位字节（类型字段）是 8005。（这个条件是为了向后兼容）
3. 两个属性的低六位字节（值字段）完全由零组成。在这种情况下，这两个属性被认为是相同的，不论其类型字段，它们被视为表示空域标识符。

如果 VPN-IPv4 路由具有 OSPF 域标识符扩展团体属性，我们说这条路由是在确定的域中。如果扩展团体属性的值字段由全零组成，那么标识的域就是 NULL 域，并且路由被认为属于 NULL 域。如果路由没有 OSPF 域标识扩展团体属性，那么路由属于 NULL 域。

每个 OSPF 实例都与一个或多个域标识符相关联，尽管可能只有 NULL 域标识符。如果 OSPF 实例与特定的域标识符相关联，我们会说它属于标识的域。

如果一个 VPN-IPv4 路由将被重新分发到一个特定的实例，必须确定该路由和那个 OSPF 实例是否属于同一个域。一个路由和一个 OSPF 实例当且仅当以下条件之一成立时，属于同一个域：

1. 路由和 OSPF 实例都属于 NULL 域。
2. 该路由所属的域是 OSPF 实例所属的域。（也就是路由的域标识符等于 OSPF 实例的域标识符，正如本节前面给出的定义所确定的那样）。

如果路由和 VRF 不属于同一个域，则路由被视为外部路由。

如果外部路由被重新分发到一个 OSPF 实例中，路由可能会或也可能不会被通告给特定的 CE，具体取决于配置以及 PE/CE 链路所属的区域类型。如果路由发布了，PE/CE 链路属于 NSSA 区域，其会在一个 7 类 LSA 中公告。否则，如果路由会在 5 类 LSA 被公告。LSA 由 PE 生成。

DN 位（4.2.5.1 节）必须在 LSA 中设置。VPN 路由标记（见第 4.2.5.2 节）必须放置在 LSA 中，除非 VPN 路由标记的使用已被配置关闭。

默认情况下，类型 2 度量值包含在 LSA 中，除非 OSPF 路由类型扩展团体属性的选项字段的 VPN-IPv4 路由指定度量标准应该是类型 1。

默认情况下，度量值取自 VPN-IPv4 路由的 MED 属性。如果 MED 不存在，则默认度量值被使用。（默认类型 1 度量标准和默认类型 2 度量可能不同。）

请注意，这种处理外部路由的方式会使每个 PE 成为连接所有外部路由的 ASBR。在一个多宿主站点，这可能导致包含许多相同的 5 类 LSA 信息。

4.2.8.2 路由汇总

如果路由和它所导入的 VRF 属于同一域，则路由应该被视为好似在一个 OSPF 的 3 类 LSA 中被收到。这意味着 PE 将在 3 类 LSA 中报告给 CE。（请注意，即使这种情况是可能的，即使 VPN-IPv4 路由携带的区域号码与 CE 路由器相同。这意味着如果一个区域被“分割”了，两件只能通过 VPN 骨干连接，看来是两个区域，他们之间有区域间的路由。）

4.2.8.3 NSSA 路由

NSSA 路由的处理方式与外部路由相同，如第 4.2.8.1 节所述。

5. IANA 考虑

[EXTCOMM]第 11 部分呼吁 IANA 为 BGP 扩展团体类型字段和扩展类型字段值创建一个注册表。本文的第 4.2.6 节为 BGP 扩展团体的扩展类型字段分配了新的值。这些值在

[EXTCOMM]指定的值范围内都下降了，“由于 IANA 使用 RFC 2434 中定义的“先到先得”政策”。

在本文件第 4.2.6 节中分配的 BGP 扩展团体扩展类型字段值，如下：

- OSPF 域标识符： 扩展类型 0005,0105 和 0205。
- OSPF 路由类型： 扩展类型 0306
- OSPF 路由器 ID： 扩展类型 0107

6. 安全考虑

安全性考虑因素与在[VPN]和[VPN-AS]中讨论的 BGP/MPLS IP VPNS 相关。我们只讨论这些安全考虑是特定于 OSPF 作为 PE/CE 协议的使用。

单个 PE 可以运行 OSPF 作为 SP 骨干网的 IGP，以及运行 OSPF 作为一个或多个 VPNs 的 IGP。这需要使用多个独立的 OSPF 实例，因此路由不会无意中在主干和任何 VPN 之间泄漏。不同 VPN 的 OSPF 实例也必须是独立的 OSPF 实例，以防止 VPNs 间无意的路由泄漏。

OSPF 提供了许多允许 PE 和 CE 之间的 OSPF 控制消息进行认证的过程。PE 和 CE 之间应该使用 OSPF“密码认证”。它必须在每个 PE 上实现。

在没有这种认证的情况下，有可能的 CE 不属于 PE 分配给它的 VPN。它可能被攻击者在其 PE/CE 链接上插入伪造消息，在任何方向。欺骗消息发送到 CE 可能会损害 CE 网站的路由。欺骗消息发送到 PE 可能会导致不正确的 VPN 路由，或者在 VPN 上拒绝服务攻击。