

组织：中国互动出版网 (<http://www.china-pub.com/>)

RFC 文档中文翻译计划 (<http://www.china-pub.com/compters/emook/aboutemook.htm>)

E-mail: [ouyang@china-pub.com](mailto:ouyang@china-pub.com)

译者：maggiee (maggiee maggiee@etang.com)

译文发布时间：2001-6-5

版权：本中文翻译文档版权归中国互动出版网所有。可以用于非商业用途自由转载，但必须保留本文档的翻译及版权信息。

Network Working Group  
Request for Comments: 2547  
Category: Informational

E. Rosen  
Y. Rekhter  
Cisco Systems, Inc.  
March 1999

## RFC2547 BGP/MPLS VPNs

Status of this Memo

This memo provides information for the Internet community. It does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

Copyright Notice

Copyright (C) The Internet Society (1999). All Rights Reserved.

摘要

本文档描述了拥有 IP 骨干网的服务提供商 SP 为其客户提供 VPN 服务的一种方法。在骨干网中，使用 MPLS（多协议标签交换）进行包转发，BGP（边界网关协议）进行路由信息分发。这个方法主要目的是为企业网络提供 IP 骨干网服务的外包。这种方法不仅对企业而言很简单，对 SP 而言也有较好的可扩展性和灵活性，还可以提供增值服务。同时，这种方法还可以建立一个为客户提供 IP 服务的 VPN。

## 目录

1. 简介 .....	3
1.1 虚拟专用网 Virtual Private Networks .....	3
1.2 边缘设备 .....	3
1.3 有重叠地址空间的 VPNs .....	4
1.4 由不同路由到达同一系统的 VPN .....	4
1.5 PE 上的转发表 .....	4
1.6 SP 主干网路由器 .....	5
1.7 安全 Security .....	5
2. 站点和 CE .....	5
3. PE 中基于站点的转发表 .....	6
3.1 虚拟站点 .....	6
4. 用 BGP 分发 VPN 路由信息 .....	7
4.1 VPN-IPv4 地址族 .....	7
4.2 控制路由分发 .....	8
4.2.1 目标 VPN 属性 .....	8
4.2.2 用 BGP 在 PE 中分发路由 .....	9
4.2.3 源 VPN 属性 .....	10
4.2.4 用目标和源属性组建 VPN .....	10
5. 在主干网上的转发 .....	11
6. PE 如何从 CE 学习路由 .....	12
7. CE 如何从 PE 学习路由 .....	14
8. CE 支持 MPLS .....	14
8.1 虚站点 .....	14
8.2 用 Stub VPN 表示 ISP VPN .....	14
9. 安全 .....	14
9.1. CE 路由器间的点到点安全隧道 .....	15
9.2. 多方安全关联 .....	15
10. 服务质量 .....	16
12. 版权事宜 .....	16
13. 安全考虑 .....	17
14. 致谢 .....	17
15. 作者地址 .....	17
16. 参考文献 .....	17
17. 版权说明 .....	18

## 1. 简介

### 1.1 虚拟专用网 Virtual Private Networks

与被称为主干网的公共网相连的是一个“站点”集合。我们基于某些原则创建该集合的若干子集，并附加如下规则：只有当两个站点都同在某个子集里时，两者间才可能存在经由该主干网的 IP 互连。

我们创建的这些子集就是“虚拟专用网”(VPNs)。只有同属某个 VPN 时，两个站点间才存在经公共主干网的 IP 连接。不属同一个 VPN 的两个站点间没有经主干网的连接。

如果一个 VPN 中的所有站点都属同一个企业，这个 VPN 就是一个公司“内联网”。如果分属不同企业，该 VPN 就是个“外联网”。一个站点可在多个 VPN 中，如一个内联网和多个外联网中。内联网和外联网我们都视作 VPN。一般而言，我们说的 VPN 并不区分内联网或外联网。

我们主要考虑主干网为一个或多个服务提供商 (SPs) 所拥有的情况。站点为 SP 的用户所有，决定某些站点是否属于一个 VPN 的策略由用户制定。有些用户可能希望完全由 SP 负责这些策略的执行，有些用户可能希望自己独立承担或与 SP 分担这项任务。在本文中，我们主要讨论这些策略的执行机制。这些机制既可以由 SP 单独执行，也可以由 VPN 用户与 SP 共同完成。这里，我们主要讨论前者。

本文中所讨论的机制可用于多种策略的执行。如对给定的一个 VPN，每个站点与其它所有站点都有直接连接（全连接），或者也可以限制某些站点间的直接连接（半连接）。

本文中我们感兴趣的是公共主干网提供 IP 服务的情况。我们关注于在一定契约条件下，一个服务提供商 SP 或多个 SP 为企业提供商主干网的情形，而并非是基于公共 Internet 的 VPN。

下面，我们将详细说明 VPN 应有的特性。本文的其余部分我们描述了一个具备所有这些特性的 VPN 模型。该模型可视为 [4] 中描述的框架结构的实例。

### 1.2 边缘设备

假定每个站点都有一个或多个用户边缘 (CE) 设备，并都通过某种数据连接方式（如 PPP, ATM, ethernet, Frame Relay, GRE tunnel 等）与一个或多个供应商边缘 (PE) 路由器相连。

如果某个站点只有一个主机，这个主机可能就是 CE 设备。如果该站点有一个子网，CE 设备可能是个交换机。一般而言，希望 CE 设备是路由器，我们称之为 CE 路由器。

如果一个 PE 路由器与某个 VPN 的一个 CE 设备相连，我们就说这个路由器与该 VPN 相连。同样，如果一个 PE 路由器与某个站点的一个 CE 设备相连，则称这个路由器与该站点相连。

如果 CE 设备是一个路由器，它是其直接相连的 PE 的路由对等体，而不是其它站点上的 CE

路由器的路由对等体。不同站点上的路由器并不直接交换路由信息，它们甚至无需互相了解（除非因为安全的需要，见第9节）。因为简化了每个站点的路由策略，可以支持大型的VPN（有大量站点的VPN）。

重要的一点是要保持SP和其用户间明晰的管理界限(cf. [4])。PE和P路由器仅由SP管理，SP用户无权介入。CE设备仅由用户管理(除非用户把管理服务委托给了SP)。

### 1.3 有重叠地址空间的 VPNs

任何两个不相交的VPN(如：无公共站点的VPN)可能有重叠的地址空间，而同一地址也可能用在不同VPN的不同系统中。只要端系统的地址在其所属的VPN中是唯一的，该端系统无须对VPN有所了解。

在这种模式下，VPN的拥有者无须管理一个主干网或一个虚拟主干网。SP也无须为每个VPN管理一个单独的主干网或虚拟主干网。主干网中站点间的路由是最优化的(基于组建VPN的限制策略)，并不受限于任何人工的隧道虚拟拓扑。

### 1.4 由不同路由到达同一系统的 VPN

一个站点可以在多个VPN中，但不同VPN到该站点中某一系统的路由无须相同。例如，假设一个内联网中包含站点A、B、C，一个外联网中包含A、B、C和一个外部站点D。若在站点A上有一个服务器，我们希望来自于B、C、D的客户能使用该服务器。同时，在站点B上有一个防火墙，为了对来自外联网的业务进行接纳控制，所有站点D连到服务器的业务都要通过这个防火墙。而来自站点C的业务是内联网的，无须经由该防火墙到达服务器。

也就是说，到服务器有两条路径。站点B和C使用一条路径直接通向站点A，站点D使用第二条路径，先到达站点B的防火墙，如果防火墙允许该业务流通过，该业务就象从站点B发出的业务流一样发往站点A。

### 1.5 PE 上的转发表

每个PE路由器都要维护若干独立的转发表。每个与PE相连的站点必须对应于其中的一个转发表。当从某个站点收取一个包时，需查找与该站点相应的转发表以确定该包的转发路径。只有当路由的目标站点与站点S同在至少一个VPN中时，才产生站点S的转发表，这可以防止两个不在同一VPN的站点间的通信，而且，这样两个没有公共站点的VPN也可以使用重叠的地址空间。

## 1.6 SP 主干网路由器

SP 的主干网包括 PE 路由器和未直接与 CE 设备相连的其它路由器 (P 路由器)。

如果 SP 主干网中的每个路由器都得为所有 VPN 维护路由信息,那么无疑这个模式的扩展能力很差, SP 能支持的站点数取决于一个路由器上可保存的路由信息的数量。因此, 一个重要的要求就是, 一个 VPN 的路由信息只保存在与该 VPN 相连的 PE 路由器上, 而 P 路由器无需保存任何 VPN 的路由信息。

VPN 可以跨越多个服务提供商。如果两个服务提供商 SP 之间是相互信任的, 那么它们的 PE 路由器间的通路可根据一个专用的对等协议穿越 SP 网络间的边界。特别是, 每个提供商信任对方并把正确的路由信息和来源可靠的带有标签的包 (在 MPLS 情况下 [9]) 传递给对方。在此, 我们假定标签交换路径可以穿越 SP 间的边界。

## 1.7 安全 Security

即使不加密, 一个 VPN 模型也应当提供相当于第二层主干网 (如 Frame Relay) 的安全保证。也就是说, 即使在误配置或故意将不同 VPN 间互连的情况下, 一个 VPN 的系统也不能进入另一个 VPN 的系统。

同时, 该模型应该也可以采用标准的安全措施。

## 2. 站点和 CE

从主干网的角度看, 如果一系列 IP 系统相互连接且它们之间的通信无需主干网参与, 则这些 IP 系统组成了一个站点。一般来说, 一个站点由一些地理位置相近的系统组成, 当然并非总是这样。如果地理位置较远的两地间用一根运行 OSPF 的专线相连, 也可以组成一个站点, 因为两地间的通信无须主干网的参与。

一个 CE 设备常被视为在一个单独的站点上 (但一个站点可能由多个“虚拟站点”组成)。而一个站点可能在多个 VPN 中。

一个 PE 路由器可能与多个站点中的 CE 设备相连, 无论它们是否在同一个 VPN 中。出于鲁棒性的考虑, 一个 CE 设备也可能与多个 PE 路由器相连, 这些路由器可能属于同一个或不同的服务提供商。如果这个 CE 设备是路由器, 则它与相连的 PE 路由器互为邻接路由器。

如果互连的基本单元是站点, 这里描述的体系结构可以实现更为精细的互连控制粒度。例如, 一个站点上的某个系统可能是一个内联网的成员, 同时也是一个或多个外联网的成员, 而该站点上的其它系统却只限于是内联网的成员。

### 3. PE 中基于站点的转发表

每个 PE 路由器维护一个或多个“站点转发表”，与 PE 相连的每个站点都对应于其中的一个表。当从某站点接收到一个包时，从与该站点相应的转发表中查找该包的目的地地址。

站点转发表是如何产生的呢？如，PE1, PE2, PE3 是三个 PE 路由器，CE1, CE2, CE3 是三个 CE 路由器。PE1 从 CE1 了解站点 CE1 可达的路由信息。如果 PE2 和 PE3 分别与 CE2、CE3 相连，且 VPN V 包括 CE1、CE2、CE3，PE1 用 BGP 向 PE2、PE3 分发它从 CE1 得到的路由信息。PE2, PE3 使用这些路由信息产生与 CE2、CE3 相应的转发表。来源于 VPN V 以外站点的路由不出现在这些转发表中，也就是说，CE2, CE3 发出的包不会发送到 VPN V 以外的站点。

如果一个站点在多个 VPN 中，其对应的转发表将包含其在所有 VPN 中的路由信息。

一般，一个 PE 只为每个站点维护一个转发表，即使它与该站点间有多个连接。如果几个不同的站点共用相同的路由，它们共享同一个转发表。

假设一个 PE 路由器从一直接相连的站点收到一个包，但在相应的站点转发表中找不到这个包的目的地地址。如果 SP 在该站点处并不提供 Internet 接入服务，那么该包因为无法发送而被丢弃。否则，将会查询 PE 的 Internet 转发表。也就是说，即使提供 Internet 接入，一般每个 PE 也只需要维护一个 Internet 路由转发表。

要保持 VPN 间的隔离，重要的一点就是主干网中的路由器不接收来自于邻接的非主干设备的带标签的包，除非

- 1) 标签栈顶的标签是主干网路由器分配给该设备的；
- 2) 主干网路由器能确定由于该标签的使用，这个包在离开主干网之前，不会检查 IP 包头和标签栈中的其余标签。

这些限制对于防止包进入其它 VPN 相当有必要。

PE 中的站点转发表只用于那些从 PE 直接相连的站点发来的包，而不用于来自于 SP 主干网的包。因此，到同一系统可能有多个不同的路由，包从哪一个站点接入主干网，就由哪一个站点决定选用何种路由。如，你可以为由外联网发往指定系统的包指定一个路由（通向防火墙），为内联网发往同一系统的包（包括那些已经通过防火墙的包）指定另一路由。

#### 3.1 虚拟站点

有时，一个站点可能被用户用 VLAN 分成几个虚拟站点。每个虚拟站点可能是不同 VPN 的成员。PE 要为每个虚拟站点维护一个独立的转发表。例，如果一个 CE 支持 VLAN，并希望每个 VLAN 对应于一个独立的 VPN，PE 与 CE 间发送的包可封装在该站点的 VLAN 中。PE 可以利用这一点，以及接收包的接口，将该包指定到某一虚拟站点。

也可以把接口分成多个“子接口”（尤其是如果接口是 Frame Relay 或 ATM），并根据包到达的子接口把包指定到一个 VPN。或者简单些，每个虚拟站点使用不同的接口。无论何种情况，

即使有多个虚拟站点，每个站点都只需一个 CE 路由器。当然，如果愿意，每个虚拟站点也可以使用不同的 CE 路由器。

注意，无论哪种情况，控制业务流属于何 VPN 的机制和策略都由用户掌握。

如果希望一个主机在多个虚拟站点上，那么主机必须确定，每个包对应于哪一个虚拟站点。例如，它可以在不同的 VLAN 的不同虚拟站点上通过不同的网络接口发送包。

这些并不要求 CE 支持 MPLS。对于支持 MPLS 的 CE 如何支持多个虚拟站点，在第八节中有一个简短的讨论。

## 4. 用 BGP 分发 VPN 路由信息

PE 路由器使用 BGP 相互分发 VPN 路由信息（更确切地说，是引起路由信息的分发）。

一个 BGP 传播者只能安装和分发一个路由到指定的地址前缀。我们允许每个 VPN 有各自的地址空间，也就是说，相同的地址可被若干 VPN 使用，而在每个 VPN 中这个地址代表不同的系统。因此，我们应该允许 BGP 为某个 IP 地址前缀安装和分发多个路由。甚至于，如果 BGP 为某一 IP 地址前缀安装了多个路由，我们必须确保在任何一个站点转发表中只出现其中的一个。

我们使用下面描述的新的地址族来实现上述目标。

### 4.1 VPN-IPv4 地址族

BGP 多协议扩展[3]允许 BGP 携带来自多个“地址族”的路由。我们先介绍 VPN-IPv4 地址族的概念。一个 VPN-IPv4 地址长 12 字节，以 8 字节的“路由标记”RD 开头，后接一个 4 字节的 IPv4 地址。如果两个 VPN 使用相同的地址前缀，PE 可以把它们翻译成不同的 VPN-IPv4 地址前缀。这样，如果在两个 VPN 中使用了相同的地址，也可以分别为每个 VPN 安装与该地址对应的不同的路由。

RD 本身并没什么特别的语义，它并不包含路由来源或向哪些 VPN 分发路由的信息。RD 的目的仅在于可以对一普通的 IPv4 前缀产生一个与众不同的路由，RD 还可以用于决定路由的重新分发。（见 4.2）

RD 还可以用于产生到同一系统的多个不同的路由。在第 3 节中，我们曾给出一个例子：从内联网到某一服务器的路由必须与从外联网的业务流路由不同。这可以通过产生两个 IPv4 地址部分相同，但 RD 不同的 VPN-IPv4 路由来实现。这样，BGP 就可以安装到同一系统的多个路由，还可以使用一定的策略（见第 4.2.3 节）来决定包的路由选择。

RD 的结构使得每个 SP 可以管理各自的“编号空间”（如，可以自主地指定 RD），而不与其它

SP 的 RD 冲突。一个 RD 包括两字节的类别域，一个管理者域，和一个指定号码域。类别域的值决定了其它两个域的长度和管理者域的语义。管理者域表明一个指定的授权号，指定号码域包括由已鉴定的授权方出于某种目的指定的一个数字。如，一个 RD 的管理者域包含一个自治系统编号 (ASN)，IANA 将这个 ASN 分配给一 SP，4 字节的号码域包括的编号就是由该 SP 指定的。RD 采用这种结构是为了确保可以提供 VPN 主干网的 SP 总可以在需要的时候产生一个唯一的 RD。不过，这种结构并没有其它语义。如果 BGP 比较这样的两个地址前缀时，它并不理会这种结构。

如果 VPN-IPv4 地址的管理者域和指定编号域都是全 0，则可以视作是与 IPv4 含义相同。尤其对 BGP 而言，这个 VPN-IPv4 地址与相应的 IPv4 地址被认为是类似的。而其它情况下，BGP 不会认为两者类似。

一个站点转发表中对任何一个 IPv4 地址前缀只有一个 VPN-IPv4 路由。当包的目标地址与一个 VPN-IPv4 路由匹配时，只需 IPv4 部分匹配即可。

一个 PE 需要为通向某一 CE 的路由配置相应的 RD。可以为 PE 中通向同一 CE 的所有路由配置同一个 RD，也可以为通向同一 CE 的不同路由配置不同的 RD。

## 4.2 控制路由分发

在这一部分，我们讨论控制 VPN-IPv4 路由分发的方法。

### 4.2.1 目标 VPN 属性

每个站点转发表都与一个或多个“目标 VPN”属性相关。

当一个 PE 路由器产生一个 VPN-IPv4 路由时，该路由就与一个或多个“目标 VPN”属性相关。这些信息作为路由属性由 BGP 携带。

任何与目标 VPN T 有关的路由都必须分发到每一个存有与目标 VPN T 有关的转发表的 PE 路由器上。当一个 PE 路由器收到这样一个路由时，应当将其安装到每个与目标 VPN T 有关的站点转发表中（实际是否安装取决于 BGP 决定处理的结果）。

一般来说，目标 VPN 属性代表一系列站点。因为路由与某一目标 VPN 属性相关，路由可置于站点转发表中，为来自相应站点的业务流寻路。

PE 路由器用一个目标 VPN 属性集合表明来自站点 S 的路由，用另一个目标 VPN 属性集合来决定是否将一个从其它 PE 路由器接收到的路由信息加入到与站点 S 有关的转发表中。这两个集合是不一样的，也无需相同。

目标 VPN 属性的功能类似于 BGP 群体属性。不过，因为后者只有两字节的编号空间，格式不

够多。扩展 BGP 群体属性以提供一个更大的编号空间相当简单，也是可能的，类似于我们对 RD 的描述（见 4.1 节），因此类别域定义了管理者域的长度，属性的其余部分是一个从指定管理者的编号空间得到的编号。

当一个 BGP 传播者收到对同一 VPN-IPv4 地址前缀的两条路由时，它根据 BGP 关于路由优先级的规则选择其中的一个。

注意一个路由只能有一个 RD，但它可以有多个目标 VPN。在 BGP 中，如果单一路由有多个属性，可扩展性就得到了提高。可以通过产生更多路由（用更多的 RD）的方法去掉目标 VPN 属性，但扩展性就差了。

PE 如何确定哪个目标 VPN 属性与一路由相关呢？有许多可能的方法。PE 可以配置通向某一站点的所有路由都与某一目标 VPN 属性相关，也可以配置通向某一站点的部分路由与一目标 VPN 属性相关，其余的与另一目标 VPN 属性相关。还可以由 CE 路由器在向 PE 分发路由时（见第 6 节），为每一路由指定一个或多个目标 VPN 属性。后一方法把 VPN 策略执行机制的控制权从 SP 转移到了客户方。即使使用这种方法，也希望 PE 能根据自身的配置减少目标 VPN 属性，或者/并且强制性地添加一些目标 VPN 属性。

更确切地说，应该称这种属性为“路由目标”属性而不是“VPN 目标”属性。它只确定一些能使用该路由的站点，而并不关心这些站点是否组成了一个 VPN。

#### 4.2.2 用 BGP 在 PE 中分发路由

如果一个 VPN 的两个站点所连接的 PE 在同一自治系统中，PE 可以通过它们之间的 IBGP 连接分发 VPN-IPv4 路由。或者，它们可以分别与一个路由反射器 RR 有一个 IBGP 连接。

如果 VPN 的两个站点在不同的自治系统中（例如他们连接到不同的 SP），那么一个 PE 路由器要使用 IBGP 把 VPN-IPv4 路由重新分发到一个自治系统边界路由器 ASBR 或是以一 ASBR 为客户的路由反射器 RR 上。ASBR 使用 EBGp 把路由重新分发到另一自治系统的 ASBR 上。这样，就可以连接到不同 SP 的不同 VPN 站点。不过，作为 SP 间互相信任协议的一部分，VPN-IPv4 路由只能被专用对等点间的 EBGp 连接所接受。VPN-IPv4 路由不能在公共 Internet 上分发或被公共 Internet 接受。

如果许多 VPN 的站点连接在不同的自治系统中，不同自治系统间并不需要有一个存有所有 VPN 路由的 ASBR，可以有多个 ASBR，每个 ASBR 只保存 VPN 的部分路由。

当一个 PE 路由器用 BGP 分发一个 VPN-IPv4 路由时，它使用自己的地址作为“BGP 下一跳”地址，并指定和分发一个 MPLS 标签（事实上，PE 路由器分发的并不是 VPN-IPv4 路由，而是带标签的 VPN-IPv4 路由，参见 [8]）。当 PE 接收到一个标签栈项是 MPLS 标签的包时，PE 会弹出该标签，直接把该包发送到路由指定的站点。这意味着它只把包发送到它学习路由的那个 CE 路由器。标签也可以决定数据链路的封装。

一般，接收带标签的包的 PE 并不在转发表中查找包的目的地址，而是利用另一 PE 指定的标

签把包直接发送到 CE。当然 PE 指定的标签也可能隐含地指定了某转发表。这种情况下，PE 接收到这个包后，会根据标签到该转发表中查找包的目的地址。在某些情况下这种方法很有用，但在本文中不做详述。

注意，这种方法分发的 MPLS 标签只在安装该路由的路由器和该路由的 BGP 下一跳之间存在标签交换路径 LSP 时才有用。我们不对标签交换路径 LSP 的建立过程做任何假定，LSP 可能是预先建立，或是在需要时才建立。它可能是个“尽力而为”路由，也可能是个经过流量工程的路由。在某路由的一个 PE 路由器和它的 BGP 下一跳之间可能有一个或多个具有不同 QoS 特性的 LSP。与 VPN 体系结构有关的是路由器及其 BGP 下一跳之间的一些 LSP。

使用路由反射器一般是为了提高可扩展性，如利用路由反射器的层次结构。使用时并不需要某个路由反射器掌握由主干网所支持的所有 VPN 的全部 VPN-IPv4 路由，可以使用若干分离的彼此间无通信的路由反射器，每个都只支持部分 VPN。

如果一个 PE 路由器不连接到一路由的任何一个目标 VPN，它就不必接收那个路由。发送该路由给它的 PE 或路由反射器应该采取出口过滤措施以免继续发送给它无用的路由。当然，如果一个 PE 路由器通过 BGP 接收了一个路由，但它并不连接到该路由的任何一个目标 VPN，PE 也应该对该路由采取入口过滤措施，不安装也不进行重新分发。

一个不连接到任何 VPN 的路由器，如一个 P 路由器，无需安装任何 VPN-IPv4 路由。

这样的分发规则确保了没有一个设备需要掌握主干网所支持的所有 VPN-IPv4 路由。因此，主干网支持的 VPN-IPv4 路由总数不受任何一个设备容量所限，也就可以不受限地增加。

### 4.2.3. 源 VPN 属性

一个 VPN-IPv4 路由可以选择性地通过源属性与一个 VPN 相关。这个属性唯一地代表了一系列站点的集合，并代表了相应的来自于该集合的一个站点的路由。这个属性的典型用途是表明了路由指向的站点的拥有者——某企业，或表明了该站点的内联网。当然，还可能其它用途。这个属性可以象一个扩展的 BGP 群体属性一样来编码。

当需要确定一路由的来源时，应当使用源属性，而不是 RD。如下文所述，这个属性可以用于构建 VPN。

更确切地说，这个属性应称为“源路由”属性而不是“源 VPN”属性。它只确定路由来自于某一站点集合中的一个站点，并不关心这些站点是否组成一个 VPN。

### 4.2.4. 用目标和源属性组建 VPN

如果正确地设置了目标 VPN 和源 VPN 属性，就可以组建各种不同的 VPN。

如果想组建一个包含特定站点集合的封闭用户群 CUG，可以用一个指定的目标 VPN 属性值来代表该 CUG。这个值应该与 CUG 中每个站点的转发表，以及从 CUG 中每个站点中学习的路由相关联。任何有该目标 VPN 属性的路由都应该重新分发，以到达每一个与 CUG 中任一站点相连的 PE 路由器。

如果是想组建一种“hub and spoke”VPN，可以使用两个目标属性值，一个代表“Hub”，一个代表“Spoke”。从 spoke 发出的路由被分发到 hub，但 hub 发出的路由并不被分发到 spoke。

如果一些站点既在内联网上也在外联网上，而另一些站点只在内联网上，那么，有一些内联网和外联网的路由具有代表全体站点的目标 VPN 属性。那些有内联网路由的站点只能过滤有“错误”源 VPN 属性的路由。

利用这两个属性，可以灵活地控制路由信息在不同站点集合间的分发，在 VPN 的组建上也提供了很好的灵活性。

## 5. 在主干网上的转发

如果主干网上的中间路由对到 VPN 的路由一无所知，数据包如何从 VPN 的一个站点转发到另一个站点呢？

这是利用 MPLS 的两层标签栈来实现的。

PE 路由器(和重新分发 VPN-IPv4 地址的 ASBR)要在主干网的 IGP 路由表中插入/32 地址前缀。这样，在主干网的每一个网络节点上，都可以用 MPLS 为到每个 PE 路由器的路由指定一个标签。(在主干网上建立标签交换路径 LSP 的过程中不需要/32 地址前缀)

当 PE 从一个 CE 设备接收到一个包时，它选择一个站点转发表来查找包的目的地地址。假设找到了匹配项。

如果该包的目的地是连接在同一 PE 上的一个 CE 设备，就直接发送到该 CE 设备。

如果该包的目的地不是连接在同一 PE 上，则找到该包的“BGP 下一跳”和 BGP 下一跳为包的目的地地址分配的标签。先把标签压入包的标签栈，成为栈底标签。接着 PE 查找到 BGP 下一跳的 IGP 路由，确定 IGP 下一跳和 IGP 下一跳为 BGP 下一跳地址分配的标签。这个标签也被压入包的标签栈，成为栈顶标签。然后这个包被转发往 IGP 下一跳。(如果 BGP 下一跳就是 IGP 下一跳，第二个标签就用不着入栈了。)

由 MPLS 携带这个包经主干网到达适当的 CE 设备。也就是说，所有的 P 和 PE 路由器做出的转发决定现在都以 MPLS 方式给出，直到包到达该 CE 设备，不需再查看包的 IP 包头。最后的 PE 路由器将在把包发送到 CE 设备前从标签栈中弹出最后的标签，这样，CE 设备看到的仍是一个普通的 IP 包。(第 8 节将讨论 CE 能接收带标签的包的情况)

当一个包通过一个 PE 路由器从某站点进入主干网时，根据 PE 路由器中与该站点相关的转发表内容决定包的路由。与包离开主干网的 PE 路由器中的转发表是无关系的。因此，到同一系统可以有多个路由，为包选择哪一个路由取决于包从哪一个站点进入主干网。

注意，两层标签的运用使保持所有 VPN 路由与 P 路由器的隔离成为可能，这对于保证该模型的扩展性相当重要。主干网甚至只需到 PE 的路由而无需到 CE 的路由。

## 6. PE 如何从 CE 学习路由

与一 VPN 相连的 PE 路由器需要了解在该 VPN 的每个站点上有哪些地址。

如果 CE 设备是一台主机或一个交换机，地址集合一般被配置到连接该设备的 PE 路由器中。如果 CE 设备是一路由器，PE 路由器可以通过多种方法获得该地址集合。

PE 用设定的 RD 把这些地址翻译成 VPN-IPv4 地址，把这些 VPN-IPv4 路由当作 BGP 的输入。这些路由在任何情况下都不会泄露给主干网的 IGP。

实际上，PE/CE 路由分发技术取决于该 CE 是否在一个“传输 VPN”中。一个“传输 VPN”包括一个从第三方(如，不在同一 VPN 中的且不是 PE 的一个路由器)接收路由，并重新分发到一个 PE 路由器的路由器。如果不是一个“传输 VPN”，一个 VPN 则是一个“叶 VPN” stub VPN。在此意义上，大多数 VPN，包括几乎所有企业网络，都希望是后者。

可能的 PE/CE 分发技术有：

1. 静态路由（如，配置）（这只用于 stub VPN）

2. PE 和 CE 路由器可能是 RIP 对等的，而且 CE 可以用 RIP 告诉 PE 路由器在 CE 路由器上的站点的可达地址前缀。当在 CE 中配置 RIP 时，要注意确保从其它站点来的地址前缀(如，CE 路由器从 PE 路由器处学习来的地址前缀)不被广告到 PE。更确切地说，如果一个 PE 路由器，如 PE1，接收到了一个 VPN-IPv4 路由 R1，处理后以 R2 为路由名继续向一个 CE 分发该 IPv4 路由，那么，R2 不能被该 CE 的站点分发至一个 PE 路由器，如 PE2，（这里，PE1 和 PE2 可能是也可能不是同一路由器），除非 PE2 将 R2 映射为一个与 R1 不同的 VPN-IPv4 路由。（如，用一个不同的 RD）

3. PE 和 CE 路由器可能是 OSPF 对等的。这时，站点应该是一个单独的 OSPF 区，CE 则是该区的 ABR，而 PE 是不属于该区的一个 ABR。而且，PE 应该只报告连接到同一站点上的 CE 的路由。（这个技术只能用于 stub VPN）

4. PE 和 CE 路由器可能是 BGP 对等体，CE 路由器可以用 BGP（特别是 EBGP）告诉 PE 路由器该 CE 路由器上的站点的可达地址前缀集合（这个技术既可用于 stub VPN，也可用于传输 VPN）。

从纯技术的角度来说，这是迄今而言最好的技术：

a) 不象 IGP，它不要求 PE 为了与多个 CE 联系而运行多个路由算法实例。

b) BGP 正是为了在不同管理系统之间传递路由信息而设计的

c) 如果站点包括“BGP 后门”，如，路由器除了与 PE 路由器的连接外，还有与其它路由器的 BGP 连接，该过程也能正常工作。其它过程是否能正常工作，要看具体的环境。

d) 使用了 BGP，CE 可以更方便地把路由属性传递给 PE。例如，CE 可以根据 PE 认可的路由目标属性为每个路由提议一个特别的目标属性。

但另一方面，如果用户本身不是一个 ISP，BGP 的使用对 CE 管理员来说是新的工作。注意，如果一个站点并不在一个传输 VPN 中，它并不需要一个自治系统编号 ASN。站点不在一个传输 VPN 中的 CE 都可以用同一个 ASN。而该 ASN 可以从私有的 ASN 空间中选择，PE 将去掉这些 ASN。使用源站点属性可以防止路由环路(见下)。

如果一站点集合组成了一个传输 VPN，可以简单地用一个 BGP 联盟来代表它们，那么对该 VPN 以外的路由器而言，该 VPN 的内部结构就是不可见的。这样，VPN 中的每一个站点需要两个到主干网的 BGP 连接，一个是到联盟内的，一个是联盟外的。考虑到主干网和站点可能采取不同的策略，一般联盟内的处理程序会稍做修改。只在其中的一个连接上，主干网是联盟的成员之一。这种技术允许作为用户的 ISP 得从另一对等的 ISP 处得到 VPN 主干服务，所以该技术对作为 VPN 服务用户的 ISP 而言可能有用。

(如果一个 VPN 用户自身是一个 ISP，而且它的 CE 路由器支持 MPLS，可以用一个更简单的技术，此时把该 ISP 视作一 stub VPN。见第 8 节)

如果我们无需区分向 PE 通知某站点上的地址前缀的各种不同方法，我们只是简单地说 PE 已从该站点学习了路由。

在一个 PE 重新分发它从一个站点处学到的 VPN-IPv4 路由之前，它必须为该路由指定如下三个属性：

- 源站点属性

该属性唯一地标识出 PE 路由器从何站点学习到了此路由。从一个站点学习到的所有路由都应当指定同一个源站点属性，即使该站点与一 PE 有多个连接或连接到多个 PE。不同的源站点属性必须为不同的站点使用。该属性可以编码为一个扩展的 BGP 群体属性(4.2.1 节)。

- 源 VPN 属性(见 4.2.1)

- 目标 VPN 属性(见 4.2.1)

## 7. CE 如何从 PE 学习路由

本节中，我们假设 CE 设备是一个路由器。

一般，一个 PE 会向 CE 分发给在转发表中用于路由来自该 CE 的包的所有路由，但不能把路由重新分发到该路由的源站点属性表明的那个站点上的 CE。

在多数情况下，PE 简单地分发到 CE 的缺省路由就可以了。（在一些情况下，CE 甚至可以只配置一条指向 PE 的缺省路由）。如果一站点无需自己向其它站点分发缺省路由，就可以采用这种方法（例，如果公司 VPN 中的一个站点接入了 Internet，这个站点会向其它站点分发缺省路由，但不会分发回自己）。

从 CE 向 PE 分发路由的任何方法都可以用来从 PE 向 CE 分发路由。

## 8. CE 支持 MPLS

如果 CE 支持 MPLS，也同意从 VPN 中接收所有路由，PE 可以向该 CE 为每个路由分发一个标签。当 PE 从 CE 处收到一个有标签的包时，PE 会 a) 用从 BGP 学来的相应的标签替换这个标签，b) 压入一个对应于该路由 BGP 下一跳的相应标签。

### 8.1 虚站点

如果用 BGP 分发 CE/PE 路由，CE 可以利用 MPLS 支持多个虚站点。CE 会为每个虚站点维护一个独立的转发表，这个转发表是根据它从 PE 得到的路由的源 VPN 和目标 VPN 属性产生的。如果 CE 从 PE 得到了所有路由信息，PE 就不用为来自 CE 的包查找地址。或者，PE 也可以为 CE 上的每一个 VPN 分配一个（有标签）的缺省路由。如果 PE 从 CE 接收到一个带标签的包，它会知道应该查找哪一个转发表，CE 用为包加上的标签表明包来自于哪一个虚站点。

### 8.2 用 Stub VPN 表示 ISP VPN

如果一个 VPN 是一个 ISP，而且 CE 路由器支持 MPLS，可以把它视作一 stub VPN。CE 和 PE 路由器只需交换 VPN 内部的路由。PE 路由器会向 CE 路由器为这些路由分别分发一个标签。在 VPN 不同站点上的路由器是 BGP 对等体。当 CE 路由器查找一个包的目的地时，先将它解析为一个内部地址，通常是包的 BGP 下一跳地址。CE 为包加上标签，再发送到 PE。

## 9. 安全

在下列情况下：

a) 主干网的路由器不接受来源不可信或不可靠的带标签的包，除非它确定这个包在离开主干网之前不检查它的 IP 包头或标签栈里的低层标签。

b) 不接受来源不可信或不可靠的带标签的 VPN-IPv4 路由

这种体系结构提供的安全性与 Frame Relay 或 ATM 主干网提供的 VPN 的安全性相同。

要注意的是，比起用 IP-IP 隧道方法，使用 MPLS 提供安全要简单一些。除非满足以上两个条件之一，拒绝接受一个带标签的包是件简单的事。如果那个包是送错了的经 IP-IP 隧道封装的包，要配置一个路由器拒绝接受这样一个 IP 包相当困难。

MPLS 的使用也使得 VPN 能跨越多个 SP，而无论域内 IPv4 路由信息如何分发。

一个 VPN 用户也可以利用 IPSEC 隧道模式[5]提高自身的安全性。这将在本节的后面讨论。

## 9.1. CE 路由器间的点到点安全隧道

一个注重安全 VPN 用户可能想确保一些或所有穿越主干网的包能够被授权或加密。目前标准的方法是在 VPN 的每一对 CE 路由器之间用 IPSEC 隧道模式建立一个“安全隧道”。

而按我们到现在为止描述的方法，并不能让传输包的 CE 路由器来决定下一个传输包的 CE 路由器。但使用 IPSEC 的隧道模式需要这个信息。所以我们要扩展这些方法以取得这个信息。

[6]中建议了一种实现方法。每一个 VPN-IPv4 路由都有一个标识路由经过的下一 CE 路由器的属性。如果这个信息提供给 VPN 中所有 CE 路由器，就可以用标准 IPSEC 隧道模式。

如果 CE 和 PE 是 BGP 对等体，把这个信息作为一个 BGP 属性是很自然的。

每个使用 IPSEC 的 CE 还需被配置一系列地址前缀，以禁止发送不安全业务流到这些地址。这样可以防止在 CE 因为某些原因无法得到必要的信息时，发送不安全业务流。

当使用 MPLS 在一个 IPSEC 隧道的两个端点传递包时，IPSEC 的外部包头并没有任何作用。如果开发一种 IPSEC 隧道模式，使得在运用 MPLS 时忽略外部包，将是很有益处的。

## 9.2. 多方安全关联

如果建立一个单一的多方的安全关联而不是在每一对 CE 路由器中建立一个安全隧道，是相当有益的。在多方安全关联中，同一 VPN 中的所有 CE 路由器共享相同的安全参数(如，相同的秘密，相同的算法等)。入口 CE 并不知道下一个接收数据的是哪个 CE，它只知道数据将流向哪个 VPN。一个在多个 VPN 中的 CE 为每一个 VPN 配置不同的安全参数，这样可以保护内

联网中的包不会暴露给外联网。

在这种情况下，因为无法向外部包头的 IP 目的地址域填充，标准 IPSEC 隧道模式无法作用。而当使用 MPLS 转发包时，外部包头没有什么用。PE 路由器可以用 MPLS 把包送到隧道另一端而无需知道该端点的 IP 地址，它只需查看内部包头的 IP 目的地址。

这种结构的突出优点就是安全体系对路由的变化(尤其是，一个地址前缀的出口 CE 的变化)是透明的。这对由多个提供者支持的 VPN 来说相当重要，这时，因为路由变化信息的分发只需支持安全体系，增强了系统的可扩展性。

另一个优点就是它用 MPLS 的封装代替了外部 IP 包头。

## 10. 服务质量

尽管不是本文的重点，但 QoS 确实是 VPN 服务中的关键组成部分。利用 MPLS 中 shim 包头中的“实验”位[10]，MPLS/BGP VPN 可提供第三层 QoS 能力。如果主干网用的是 ATM，也可以利用 ATM 的 QoS 特性。[1]中讨论的流量工程方法也可以直接用在 MPLS/BGP VPN 中。还可以用流量工程方法在某些站点对之间建立带有特定 QoS 属性的 LSP。如果 MPLS/BGP VPN 跨越多个 SP，可以用[7]中描述的体系结构。一个 SP 还可以为某一 VPN 提供集成、区分服务能力。

### 11. 可扩展性

我们已在本文中讨论可扩展性的问题。在本节中，我们简单地总结一下本模式在可扩展性方面的主要特性。

服务提供商的主干网络包括 PE 路由器，BGP 路由反射器 RR，P 路由器(既不是 PE 也不是 RR)，如果 VPN 有多个提供者的话还有 ASBR。

P 路由器不维护任何 VPN 路由。为了正确地转发 VPN 流，P 路由器只要维护到 PE 路由器和 ASBR 的路由。两层标签的使用使 VPN 路由得以独立于 P 路由器。

一个 PE 路由器只维护与它直接相连的 VPN 的路由。

路由反射器 RR 和 ASBR 都只需维护服务提供商支持的部分 VPN 的路由。这样，没有一个 RR 或 ASBR 需要维护所有 VPN 的路由。

因此，服务提供商网络中没有一个组件要维护所有 VPN 的所有路由。因此，网络可以支持的 VPN 数量不受网络单一组件容量的限制。

## 12. 版权事宜

Cisco Systems may seek patent or other intellectual property protection for some of all of the technologies disclosed in this document. If any standards arising from this document are or become protected by one or more patents assigned to Cisco Systems, Cisco intends to disclose those patents and license them on reasonable and non-discriminatory terms.

### 13. 安全考虑

安全方面的问题本文已有讨论。

### 14. 致谢

Significant contributions to this work have been made by Ravi Chandra, Dan Tappan and Bob Thomas.

### 15. 作者地址

Eric C. Rosen    Cisco Systems, Inc.    250 Apollo Drive    Chelmsford, MA, 01824  
Email: erosen@cisco.com

Yakov Rekhter    Cisco Systems, Inc.    170 Tasman Drive    San Jose, CA, 95134  
Email: yakov@cisco.com

### 16. 参考文献

- [1] Awduche, Berger, Gan, Li, Swallow, and Srinivasan, "Extensions to RSVP for LSP Tunnels", Work in Progress.
- [2] Bates, T. and R. Chandrasekaran, "BGP Route Reflection: An alternative to full mesh IBGP", RFC 1966, June 1996.
- [3] Bates, T., Chandra, R., Katz, D. and Y. Rekhter, "Multiprotocol Extensions for BGP4", RFC 2283, February 1998.
- [4] Gleeson, Heinanen, and Armitage, "A Framework for IP Based Virtual Private Networks", Work in Progress.
- [5] Kent and Atkinson, "Security Architecture for the Internet Protocol", RFC 2401, November 1998.
- [6] Li, "CPE based VPNs using MPLS", October 1998, Work in Progress.
- [7] Li, T. and Y. Rekhter, "A Provider Architecture for Differentiated Services and Traffic Engineering (PASTE)", RFC 2430, October 1998.
- [8] Rekhter and Rosen, "Carrying Label Information in BGP4", Work in Progress.
- [9] Rosen, Viswanathan, and Callon, "Multiprotocol Label Switching

Architecture”, Work in Progress.

- [10] Rosen, Rekhter, Tappan, Farinacci, Fedorkow, Li, and Conta, “MPLS Label Stack Encoding”, Work in Progress.

## 17. 版权说明

Copyright (C) The Internet Society (1999). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an “AS IS” basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.