文章编号:1001 - 9081(2006)04 - 0980 - 03

BGP/MPLS VPN 在 NS-2 中的实现

韩 波[1],沈富可[2],刘 莉[2]
(1. 华东师范大学 计算机科学技术系,上海 200062; 2. 华东师范大学 网络中心,上海 200062)
(bhan@admin.ecnu.edu.cn)

摘 要:研究了 BGP/MPLS VPN 的原理,分析了边界网关协议(BGP)和多协议标签交换(MPLS)在 NS 中的实现,进而针对 NS-2 环境提出了一套符合 RFC 标准的 BGP/MPLS VPN 的具体实现方案。最后给出了该实现对 BGP/MPLS VPN 的模拟方法和实验结果,证明了该方案的可行性。

关键词:边界网关协议;多协议标签交换;虚拟专用网;NS-2
中图分类号: TP393.01 **文献标识码**:A

Implementation of BGP/MPLS VPN in NS-2

HAN Bo[1], SHEN Fu-Ke[2], LIU Li[2]
(1. *Department of Computer Science, East China Normal University, Shanghai* 200062, *China*;
2. *Network Information Center, East China Normal University, Shanghai* 200062, *China*)

Abstract: The principle of BGP/MPLS VPN was researched, and the ways of implementing BGP and MPLS in NS was analyzed respectively. A new solution to implement BGP/MPLS VPN for the environment of NS-2 was presented which is consistent with the RFC standard. At last, it described how to use this solution to simulate this kind of VPN and some testing results, which proved the feasibility of this solution.

Key words: Border Gateway Protocol(BGP); Multiprotocol Label Switching(MPLS); Virtual Private Network(VPN); NS-2

0 引言

BGP/MPLS VPN 是一种采用边界网关协议(Border Gateway Protocol,BGP)在边界路由器间分发路由、使用多协议标签交换(Multiprotocol Label Switching,MPLS)技术在 VPN(Virtual Private Network,虚拟专用网)站点之间传送数据的三层 VPN。其中有三种类型的路由器:CE(Customer Edge)路由器、PE(Provider Edge)路由器和 P(Provider)路由器。CE 路由器是客户端路由器,为用户提供到 PE 路由器的连接;PE 路由器是运营商边缘路由器,也就是 MPLS 网络中的标签边缘路由器(Label Edge Router,LER),它根据存放的路由信息将来自 CE 路由器或标签交换路径(Label Switch Path,LSP)的 VPN 数据处理后进行转发,同时负责和其他 PE 路由器交换路由信息;P 路由器是运营商网络主干路由器,也就是 MPLS 网络中的标签交换路由器(Label Switch Router,LSR),它根据分组的外层标签对 VPN 数据进行透明转发。P 路由器只维护到 PE 路由器的路由信息而不维护 VPN 相关的路由信息。其基本结构可以见图 1 所示。

NS(Network Simulator)网络仿真器具有开放性好、扩展性强等特点,是一个出色的研究网络拓扑结构、分析网络传输的仿真工具。随着众多科研学者对它的不断扩充和完善,NS 在近年来的科学研究中越来越受到重视。

BGP/MPLS VPN 是一个大规模的 VPN,采用真实的环境对它进行研究比较困难。然而在 NS 模拟器中还没有实现这种 VPN,因此研究在 NS 中如何实现 BGP/MPLS VPN 以进行各方面的研究具有很重要的意义。本文就给出了一种 BGP/MPLS VPN 在 NS 中的实现方案。

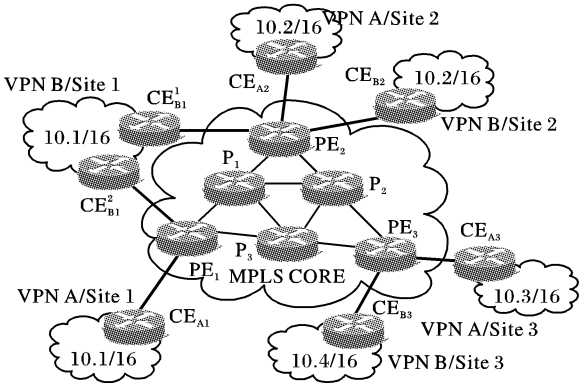

图 1 BGP/MPLS VPN 基本结构

1 基于 NS 的 BGP 和 MPLS 分析

1.1 NS 中的 BGP(ns-BGP)

本文采用的 BGP 是加拿大 Simon Fraser 大学 ns-BGP 项目[1]给出的适用于 NS-2.27 版本的 BGP 扩展。该扩展移植自 SSFNET 模拟器的 BGP-4 模型,实现了 BGP-4 的大部分功能,是目前比较完善的 BGP 实现。ns-BGP 主要对 NS 作了以下四个方面的扩展:

1)类 TcpSocket:实现了一套 Sockt API,给 BGP 提供 TCP 的支持。

2)类 IPv4Classifer:重载 NS 的类 Classifier,实现了基于 IPv4 地址的包分类和转发。

3)路由模型 rtModule/BGP:用 TCL 实现的一个新的路由模型,对 ns-BGP 节点提供注册等支持。

4)类 rtProtoBGP.:建立 BGP peer 间的会话、从 IBGP 和

收稿日期:2005 - 10 - 14;**修订日期**:2005 - 12 - 19

作者简介:韩波(1980 -),男,山东莱芜人,硕士研究生,主要研究方向:计算机网络安全; 沈富可(1967 -),男,山东莱阳人,副教授,主要研究方向:计算机网络通信; 刘莉(1980 -),女,山东滨州人,助理工程师,硕士,主要研究方向:计算机网络安全.

EBGP 学习路径信息、最佳路径选择和对 IP 转发表的更改以及管理 BGP 状态机。

1.2　NS 中的 MPLS

MNS(MPLS NS)是对 NS 作的 MPLS 扩展,以便在 NS 中模拟 MPLS。该扩展主要包括 MPLSClassifier 和 Agent/LDP。其中 MPLSClassfier 负责对数据包进行分类、压入/弹出标签、交换标签等操作;LDP(Label Distribution Protocol,标签分发协议)则负责标签的分发以建立 LSP。

MPLS 节点使用三个表结构来管理有关 LSP 的信息和标签的分配信息:PFT(Partial Forwarding Table)、LIB(Lable Information Base)和 ERB(Explicit Routing Information Base)。

2　BGP/MPLS VPN 在 NS 中的实现

MPLS VPN 使用多个 VPN 路由转发表(VRF)解决地址重叠的问题。在运营商 PE 路由器上使用基于每 VPN 的路由转发表隔离不同 VPN 的路由。通过路由信息的隔离,实现支持 VPN 地址的重叠。对于重叠 VPN 的情况,重叠发生的站点需要使用独立的 VRF 表存储来自其所属 VPN 的路由信息。为了区别来自不同 VPN 的路由信息,PE 使用 8 octet 的路由标识(RD)对来自不同 VPN 的路由信息进行标识。这个 8 octet 的路由标识作为 4 octet 的 IP 地址前缀的扩展构成了一个新的地址类(VPN-IPv4 地址)。为了防止 PE 路由器接收到不属于该 PE 上 VPN 成员的路由信息而浪费 PE 的资源,MPLS VPN 使用 BGP 的扩展属性 RT 来控制运营商网络中路由信息的发布。

根据文献[2]中作出的规范, BGP/MPLS VPN 的实现主要包括在 PE 路由器上实现多个 VPN 路由转发表(VRF)及相应处理机制,对现有的 BGP-4 进行多协议扩展[3]以分发 VPN 路由信息(其中包括 VPN 使用的内部标签),利用 MPLS 标签嵌套实现 VPN 隧道以转发数据。

2.1　实现 VRF

ns-BGP 的 PE 节点中只有一张用来处理普通 BGP 报文的转发表 RouteTable。为了实现 PE 节点上的多个路由转发表 VRF,本文所用的方法实现了一个专门用来处理 VRF 的类 VRF,并将这个类的一个实例向量嵌入到 PE 节点的分类器 IPv4Classifier 中。类 VRF 实现了一个以数据结构 VRFEntry 为条目的 VRF 表和相应的 VRF 表操作,同时还带有成员路由标识 RD,输入路由目标向量 IMPORT_RT 和输入路由目标向量 EXPORT_RT。其中数据结构 VRFEntry 如下所示:

```
struct VRFEntry
 {
   int network_;                    //网络地址
   int prefixlen_;                  //地址前缀长度
   int nexthop_;                    //下一跳地址
   int label_;                      //VPN 内部标签
   int olface_;                     //输出接口
 };
```

VRF 表所存储和处理的路由信息都是带有 RD 的 VPN-IPv4 路由。每个 VRF 有一个唯一的 RD 值。收到 VPNIPv4 路由时根据 RD 选择 VRF,根据比较 RT 来确定是否安装该路由。如可以安装则将其中的网络地址、前缀长度、下一跳和标签构造结构 VRFEntry 安装到 VRF 表中。此时的输出接口为 －1。

本方法通过对 IPv4Classifier 扩展一个配置 VRF 的接口,实现对客户 VPN 的配置。该配置包括设置 VRF 的 RD、RT,添加到客户 CE 的接口路由,具体可见第 4 节的测试脚本。

2.2　实现 MP-BGP 分发 VPN 路由

ns-BGP 没有实现 MP-BGP,为了实现 MP-BGP 需要对 ns-BGP扩展两个路由属性:MP_REACH_NLRI 和 MP_UNREACH_NLRI。为了传输 VRF 的路由目标,还需增加一个扩展的集合属性 RouteTarget。相应地,要扩展 Agent/rtProto/BGP 为能够处理 MP-BGP 和 RoutTarget 的 Agent/rtProto/MBGP。

VPN 所使用的内部标签是由 MP_REACH_NLRI 来携带的。本文采用的方法中 CE 节点和 PE 节点之间是采用 EBGP 来交互路由信息的。当 PE 节点收到来自 CE 节点的一条路由更新报文时,它首先根据网络接口来查找 VRF 来确定该接口是否是一个 VPN 的接口。如果该接口属于一个 VPN,并在相应 VRF 中没有找到该路由信息,则它会首先分配一个 VPN 标签,然后使用它和从 VRF 查找到的 RD 与 EXPORT_RT 把这条路由信息变成一条 VPN-IPv4 路由更新信息发布给 PE 路由器所属 AS 内的所有 BGP peer。

当 PE 节点收到一条 VPN-IPv4 路由更新时,它首先比较 VRF 的 IMPORT_RT 向量和该路由的 RouteTarget 属性来确定是否进行安装。如果可以安装,它将从 MP_REACH_NLRI 中取出路由添加到相应的 VRF 中,从相应的 VRF 中把与 MP_UNREACH_NLRI 中的路由相同路由删除。当 BGP 要发布 VPN-IPv4 路由时,它要查找 VRF 的 EXPORT_RT 向量,并将其作为路由的 RouteTarget 属性发布出去。所有的 VRF 更新都来源于更新路由的 MP_REACH_NLRI 和 MP_UNREACH_NLRI 属性,而不是路由本身。

图 2 所示为 BGP 分发路由信息的过程。

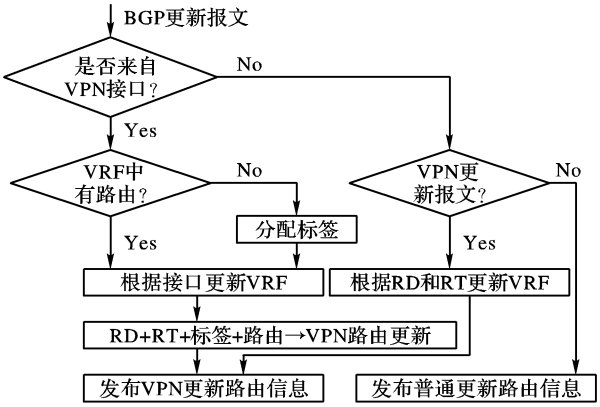

图 2　BGP 更新报文处理流程

2.3　实现数据转发:RtModule/VPNIPv4 路由模型

为了使 PE 节点同时具有处理 BGP 和 MPLS 的功能,本方法新建了一个路由模型 RtModule/VPNIPv4,并将 PE 的节点结构扩展为如图 3 所示。

(1)控制平面:多协议 BGP 的控制操作由 Agent/roProto/MBGP 和 Agent/PeerEntry 实现;LDP 的控制操作仍旧由 Agent/LDP 实现。

(2)数据平面:由 MPLS Classfier、IPv4 Classifier 和 Port Classifier 构成数据包的三级转发。其中 MPLS Classifer 与 LIB 表交互,IPv4 Classifier 和 VRF 表交互。VRF 表由 Agent/rtProto/BGP 通过 RtModule/VPNIPv4 进行相应的添加和删除。

具有这种节点模型的 PE 路由器的转发流程如图 4 所示。当 PE 节点的 MPLS Classifier 收到数据包之后将首先判

断该数据是否是 VPN 数据。如果是无标签的 VPN 数据则进行标签封装并转发至 VRF 表中的下一跳,如果是只有 VPN 标签的数据将弹出 VPN 标签,并根据 VRF 转发给相应输出接口的 CE 节点。

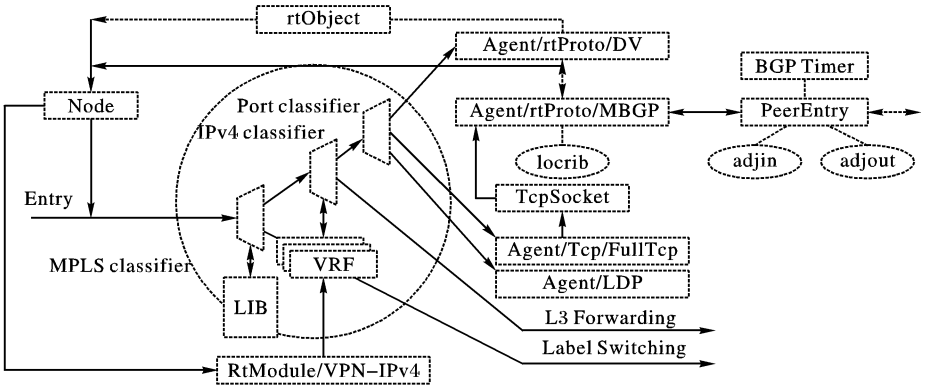

图 3 BGP/MPLS VPN PE 节点结构

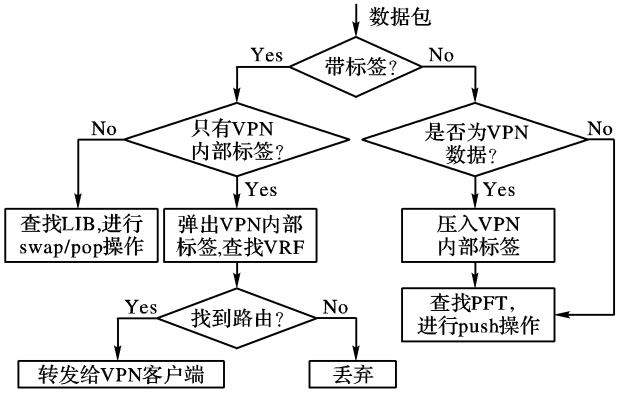

图 4 PE 节点上的数据转发流程

3 结果与测试

对按照前述方法实现的 VPN,我们采用了如图 5 的拓扑来进行测试。其中节点 2 和节点 6 为 PE 节点,节点 0、1、7、8 都为 CE 节点,节点 3、4、5 为 P 节点。

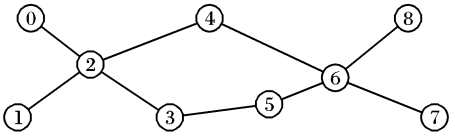

图 5 测试拓扑图

图示节点 0 ~ 8 的地址为 10.0.0.1 ~ 10.0.8.1,其中 0、1、7、8 的 AS 号为 2,其余节点 AS 号为 1。

在节点 2 上配置如下:

```
$classifier_ vrf vpn-test
$classifier_ vrf vpn-test rd 100:1
$classifier_ vrf vpn-test route-target import 100:1
$classifier_ vrf vpn-test route-target export 100:1
$classifier_ vrf vpn-test interface 0 10.0.0.0/24
```

在节点 6 上配置如下:

```
$classifier_ vrf vpn-test
$classifier_ vrf vpn-test rd 100:1
$classifier_ vrf vpn-test route-target import 100:1
$classifier_ vrf vpn-test route-target export 100:1
$classifier_ vrf vpn-test interface 8 10.0.8.0/24
```

此时节点 0 和节点 8 便加入了 VPN vpn-test。当 CE 节点作如下配置后,节点 2 和节点 6 上的 VRF 表将会产生 VPN-IPv4 路由。

在节点 0 上配置如下:

```
$bgp_agent neighbor 10.0.2.1 remote-as 1
$bgp_agent network 10.0.100.0/24
```

在节点 8 上配置如下:

```
$bgp_agent neighbor 10.0.6.1 remote-as 1
$bgp_agent network 10.0.100.0/24
```

此时节点 2 和节点 6 的 VRF 表如下:

节点 2 上用新增调试命令 $clssifier dumpvrf 显示如下:

```
VRF: vpn-test RD: 100:1
Route-Target: import: 100:1 export: 100:1
network nexthop label oIface
10.0.2.1/32 -1 -1 2
10.0.10.0/24 10.0.0.1 -1 0
10.0.100.0/24 10.0.0.1 0 0
```

节点 6 上用新增调试命令 $clssifier dumpvrf 显示如下:

```
VRF: vpn-test RD: 100:1
Route-Target: import: 100:1 export: 100:1
network nexthop label oIface
10.0.6.1/32 10.0.6.1 -1 6
10.0.8.0/24 10.0.8.1 -1 8
10.0.100.0/24 10.0.8.1 8 8
```

上述结果显示 VRF 已经为 VPN 站点 0 和 8 建立了 VPN-IPv4 路由,数据将会安照图 4 所示流程顺利通过。事实上当在 CE 节点 0 上配置一个数据源并采用 VPN 地址传递给 CE 节点 8 时,节点 8 的确收到了数据。

4 结语

本文在 NS-2 中现有的 BGP 扩展和 MPLS 扩展的基础上进一步对 NS-2 进行扩展,给出了文献[2]规定的 BGP/MPLS VPN 的具体实现方案,为对 BGP/MPLS VPN 上的进一步研究打下了基础。本方案也有一定的局限性,比如核心网是在同一个 AS 内的、只支持 IPv4 地址、LSP 的建立也没有利用流量工程的方法等。这些都是以后需要继续改进的。

参考文献:

[1] TONY DONGLIANG FENG. Implementation of BGP in a Network Simulator[EB/OL]. http://www.ensc.sfu.ca/~ljilja/cnl/projects/BGP/Tony_thesis.pdf, 2005-09.

[2] ROSEN EC. BGP/MPLS IP VPNs[Z]. IETF Ineternet Draft, 2004.

[3] Multiprotocol Extensions for BGP-4, RFC2858[S]. 2000.

[4] Sangli SR. BGP Extended Communities Attribute[Z]. IETF Internet Draft, 2004.

[5] A Border Gateway Protocol 4(BGP-4), RFC 1771[S]. 1995.

[6] Carrying Label Information in BGP-4, RFC3107[S]. 2001.

[7] MPLS Label Stack Encoding, RFC3032[S]. 2001.

[8] 徐需鸣. NS 与网络模拟[M]. 北京: 人民邮电出版社, 2003.