

S8500交换机实现BGP/MPLS VPN技术介绍

一、简单原理介绍

BGP/MPLS VPN的主要原理是：利用BGP在运营商骨干网上传播VPN的私网路由信息，用MPLS来转发VPN业务流。

下面从VPN路由信息的发布和VPN报文转发两个方面介绍BGP/MPLS VPN的实现。

VPN路由信息发布：

(1) CE到PE间的路由信息交换

PE可以通过静态路由、RIP（应支持多实例）、OSPF（应支持多实例）或EBGP学习到与它相连的CE的路由信息，并将此路由安装到VPN-instance中。

(2) 入口PE到出口PE的路由信息交换

入口PE路由器利用MBGP穿越公网，把它从CE学习到的路由信息发布给出口PE（带着MPLS标签），同时，获得出口PE学习到的CE路由信息。

PE之间通过IGP（如RIP、OSPF）或者配置静态路由来保证VPN内部节点之间的连通性。

(3) PE之间的LSP建立

为了使用MPLS LSP转发VPN的数据流量，一定要在PE之间建立LSP。从CE接收报文并建立标签栈的PE路由器是Ingress LSR，BGP的下一跳（即出口PE路由器）是Egress LSR。使用LDP建立LSP将在PE之间形成全连接的LSP。

(4) PE到CE间的路由信息交换

CE可以通过静态路由、RIP、OSPF、或EBGP，从相连的PE上学习远端的VPN路由。

经过以上的步骤，CE之间将建立可达的路由，完成VPN私网路由信息在公网上的传播。

VPN报文的转发：

VPN报文在入口PE路由器上形成两层标签栈：

内层标签，也称MPLS标签，是由出口PE向入口PE发布路由时由M-BGP分配的（安装在VPN转发表中），在标签栈中处于栈底位置。当从公网上发来的VPN报文从PE到达CE时，根据标签查找MPLS转发表就可以从指定的接口将报文发送到指定的CE或者Site。

外层标签，也称LSP的初始化标签，由MPLS LDP分配，指示了从入口PE到出口PE的一条LSP，在标签栈中处于栈顶位置。VPN报文利用这层标签的交换，就可以沿着LSP到达对端PE。

二、S8500典型配置实例

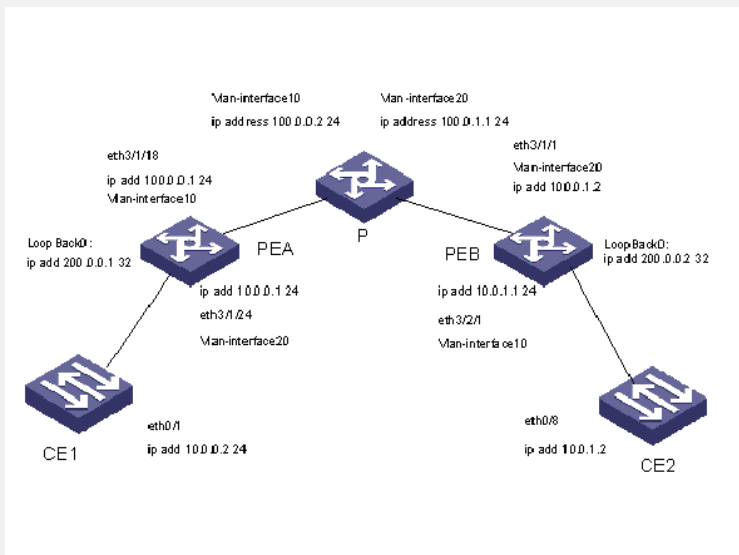
2.1 组网需求

CE1、CE2构成VPNA；

VPNA使用的VPN-target属性为100:1；

PE、P为支持MPLS的Quidway交换机，CE为一层的三层交换机。

2.2 组网图



2.3 配置命令

Switch PE A的配置：

```

[PEA]vlan 10
[PEA-vlan10]port eth3/1/18
[PEA -vlan10]int vlan 10
[PEA -Vlan-interface10]ip add 100.0.0.1 24
[PEA]mpls lsr-id 100.0.0.1
[PEA]mpls
[PEA]mpls ldp
#在接口上启动mpls。
[PEA -Vlan-interface10]mpls
[PEA -Vlan-interface10]mpls ldp en
[PEA -Vlan-interface10]vlan 20
[PEA -vlan20]port eth3/1/24
#创建vpn-instance实例，并设置RD、RT。
[PEA]ip vpn-instance 1
[PEA-vpn-1]route-distinguisher 100:1
[PEA-vpn-1]vpn-target 100:1 both
#PEA和CE1之间利用BGP来交换路由。
[PEA]bgp 100
[PEA-bgp]ipv4-family vpn-instance 1
[PEA-bgp-af-vpn-instance]import-route direct
[PEA-bgp-af-vpn-instance]group 10 ex
[PEA-bgp-af-vpn-instance]peer 10.0.0.2 group 10 as-number 65000
[PEA-bgp-af-vpn-instance]int vlan20
#vpn实例和int vlan20绑定。
[PEA-Vlan-interface20]ip binding vpn-instance 1
[PEA-Vlan-interface20]ip add 10.0.0.1 24
#设置loopback接口，并使掩码位32位，此接口地址用来触发mpls标签的生成和传递
。
[PEA-LoopBack0]ip add 200.0.0.1 32
#使能ospf来传递PEA和P之间的路由。
[PEA]ospf
[PEA-ospf-1]area 0
[PEA-ospf-1-area-0.0.0.0]network 100.0.0.1 0.0.0.255
[PEA-ospf-1-area-0.0.0.0]network 200.0.0.1 0.0.0.0
[PEA-ospf-1]import-route direct
#设置PEA和PEB之间为IBGP邻居。同时传递IPV4族路由。
[PEA]bgp 100
[PEA-bgp]group 200 in
[PEA-bgp]peer 200.0.0.2 connect-interface loopback 0
[PEA-bgp]ipv4-family vpnv4
[PEA-bgp-af-vpn]peer 200 enable
[PEA-bgp-af-vpn]peer 200.0.0.2 group 200
P的基本配置：
[P]mpls lsr-id 100.0.0.2
[P]mpls
[P]mpls ldp
[P]vlan 10
[P-vlan10]port eth3/1/18
[P-vlan10]int vlan 10
[P-Vlan-interface10]ip address 100.0.0.2 255.255.255.0
#在接口上使能mpls。
[P-Vlan-interface10]mpls
[P-Vlan-interface10]mpls ldp enable
[P]vlan 20
[P-vlan20]port eth3/1/24
[P-vlan20]int vlan 20
[P-Vlan-interface20]ip address 100.0.1.1
255.255.255.0
[P-Vlan-interface20]mpls
[P-Vlan-interface20]mpls ldp enable
[P]ospf
[P-ospf-1]area 0
[P-ospf-1-area-0.0.0.0]network 100.0.0.0 0.0.0.255
[P-ospf-1-area-0.0.0.0]network 100.0.1.0 0.0.0.255

```

Switch PEB的配置：

```
[PEB]mpls lsr-id 100.0.0.1
[PEB]mpls
[PEB]mpls ldp
[PEB]vlan 20
[PEB-vlan20]port eth3/1/1
[PEB-vlan20]int vlan 10
[PEB-Vlan-interface20]ip add 100.0.1.2 24
#在接口上启动mpls。
[PEB-Vlan-interface20]mpls
[PEB-Vlan-interface20]mpls ldp en
[PEB-Vlan-interface20]vlan 10
[PEB-vlan10]port eth3/2/1
#创建vpn-instance实例，并设置RD、RT。
[PEB]ip vpn-instance 1
[PEB-vpn-1]route-distinguisher 100:1
[PEB-vpn-1]vpn-target 100:1 both
#PEB和CE1之间利用BGP来交换路由。
[PEB]bgp 100
[PEB-bgp]ipv4-family vpn-instance 1
[PEB-bgp-af-vpn-instance]import-route direct
[PEB-bgp-af-vpn-instance]group 10 ex
[PEB-bgp-af-vpn-instance]peer 10.0.1.2 group 10 as-number 65001
[PEB-bgp-af-vpn-instance]int vlan10
#vpn实例和int vlan10绑定。
[PEB-Vlan-interface10]ip binding vpn-instance 1
[PEB-Vlan-interface10]ip add 10.0.1.1 24
#设置loopback接口，并使掩码位32位，此接口地址用来触发mpls标签的生成和传递
```

```
。
[PEB]int loop 0
[PEB-LoopBack0]ip add 200.0.0.2 32
#使能ospf来传递PEA和P之间的路由。
[PEB]ospf
[PEB-ospf-1]area 0
[PEB-ospf-1-area-0.0.0.0]network 100.0.1.2 0.0.0.255
[PEB-ospf-1-area-0.0.0.0]network 200.0.0.2 0.0.0.0
[PEB-ospf-1]import-route direct
#设置PEB和PEA之间为IBGP邻居。同时传递IPv4族路由。
[PEB]bgp 100
[PEB-bgp]group 200 in
[PEB-bgp]peer 200.0.0.1 group 200
[PEB-bgp]peer 200.0.0.1 connect-interface LoopBack 0
[PEB-bgp]ipv4-family vpnv4
[PEB-bgp-af-vpn]peer 200 enable
[PEB-bgp-af-vpn]peer 200.0.0.1 group 200
```

三、正常状态信息查看

#显示mpls lsp的信息，可以看出有两条lsp信息，分别是200.0.0.1和200.0.0.2。其中200.0.0.2的下一跳是100.0.0.2，即为P。

```
[PEA]dis mpls lsp
          LSP Information: Ldp Lsp
NO  FEC          NEXTHOP    I/O-LABEL  OUT-INTERFACE
1   200.0.0.1/32  127.0.0.1  3/-----  -----
2   200.0.0.2/32  100.0.0.2  -----/1024  Vlan10
TOTAL: 2 Record(s) Found.
```

#显示vpn实例1的bgp vpnv4路由，可以看到路由表中存在到10.0.0.0的路由，下一跳是200.0.0.1，即PEA。

```
[PEB]dis bgp vpnv4 vpn 1 rou
Flags: # - valid    ^ - active    I - internal
       D - damped   H - history   S - aggregate suppressed
       B - balance
       Dest/Mask    Next-Hop     Med      Local-pref Origin Path
Route Distinguisher:100:1 (VPN instance:1)
#^I 10.0.0.0/24    200.0.0.1   0        100     INC
#^ 10.0.1.0/24    0.0.0.0     0         0       INC
```

Routes total: 2

#显示PEB的路由表, 发现并没有vpn实例1中的路由条目, 可见, 其路由是相互独立的。

[PEB]dis ip rou

Routing Table: public net

Destination/Mask	Protocol	Pre	Cost	NextHop	Interface
100.0.0.0/24	OSPF	10	20	100.0.1.1	Vlan-interface20
100.0.1.0/24	DIRECT	0	0	100.0.1.2	Vlan-interface20
100.0.1.2/32	DIRECT	0	0	127.0.0.1	InLoopBack0
127.0.0.0/8	DIRECT	0	0	127.0.0.1	InLoopBack0
127.0.0.1/32	DIRECT	0	0	127.0.0.1	InLoopBack0
200.0.0.1/32	OSPF	10	21	100.0.1.1	Vlan-interface20
200.0.0.2/32	DIRECT	0	0	127.0.0.1	InLoopBack0

#显示mpls lsp, 可见有两条lsp, 其中200.0.0.1的下一跳是100.0.1.1。

[PEB]dis mpls lsp

LSP Information: Ldp Lsp

NO	FEC	NEXTHOP	I/O-LABEL	OUT-INTERFACE
1	200.0.0.1/32	100.0.1.1	----/1024	Vlan20
2	200.0.0.2/32	127.0.0.1	3/----	-----

TOTAL: 2 Record(s) Found.

CE1的基本配置:

[CE1]bgp 65000

[CE1-bgp]group 10 ex

[CE1-bgp]peer 10.0.0.1 group 10 as 100

[CE1-bgp]import-route direct

[CE1]int eth0/1

[CE1-Ethernet0/1]ip add 10.0.0.2 24

#显示CE1的路由表, 可见有一条bgp路由, 可以到达10.0.1.0, 即可到CE2。

[CE1]dis ip rou

Routing Table: public net

Destination/Mask	Protocol	Pre	Cost	NextHop	Interface
10.0.0.0/24	DIRECT	0	0	10.0.0.2	Ethernet0/1
10.0.0.2/32	DIRECT	0	0	127.0.0.1	InLoopBack0
10.0.1.0/24	BGP	256	0	10.0.0.1	Ethernet0/1
127.0.0.0/8	DIRECT	0	0	127.0.0.1	InLoopBack0
127.0.0.1/32	DIRECT	0	0	127.0.0.1	InLoopBack0

CE2的基本配置:

[CE2]vlan 20

[CE2-vlan20]port eth0/8

[CE2-vlan20]int vlan20

[CE2-Vlan-interface20]ip add 10.0.1.2 255.255.255.0

[CE2-Vlan-interface20]bgp 65001

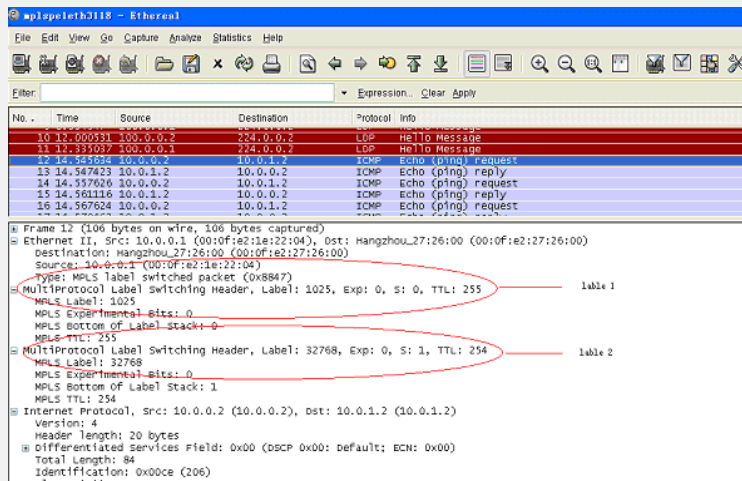
[CE2-bgp]peer 10.0.1.1 group 10

#显示bgp的peer对象, 看出是peer的地址是10.0.1.1。

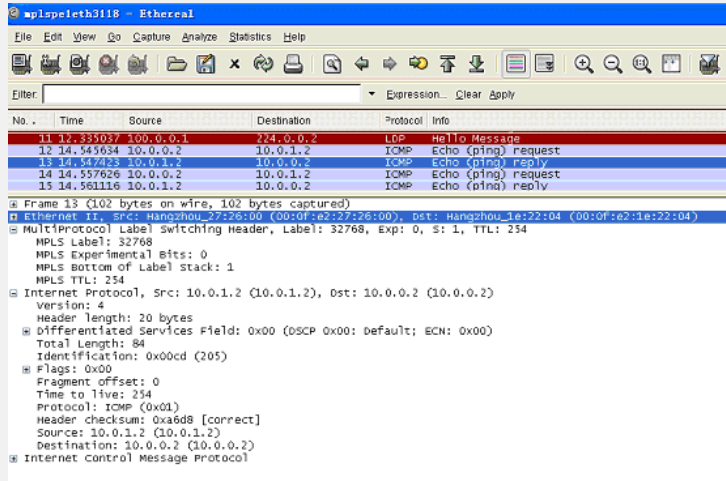
[CE2-bgp]dis bgp peer

Peer	AS-number	Ver	Queued-Tx	Msg-Rx	Msg-Tx	Up/Down	State
10.0.1.1	100	4	0	0	0	00:00:55	Connect

#下图是10.0.0.2 ping 10.0.1.2时, 在PEA上抓到的包, 可以看到, 传递的数据打上了两层的mpls标签。



#当10.0.1.2 回应 10.0.0.2的ping包时，在PEA上抓到的包，可以看出，当数据包到了PEA后，只能下一层的私网mpls的标签了。



#拓扑发生变化，BGP路由更新，在PEA上抓到的包，可以看出，在路由属性的withdr awn条目中，有RD的存在，同时传递路由的过程中也被打上了一层公网标签。

