文章编号：2096-1618(2020)04-0378-04

基于 BGP/MPLS VPN 的 VRF 配置设计与仿真

陈丰琴， 窦 军， 张 丹
(西南交通大学信息科学与技术学院，四川 成都 611756)

摘要：BGP/MPLS VPN 是 ISP 骨干网的重要模型，是通信领域备受关注的研究对象。针对该模型中 PE 节点内 VRF 转发技术及路由部署进行研究，首先介绍 BGP/MPLS VPN 模型及 VRF 技术原理；然后，基于 Linux 4.14 内核交换机上 L3mdev 机制设计 VRF 配置；最后，通过仿真测试 VRF 上部署 OSPF 协议，验证其有效的隔离性。

关 键 词：BGP/MPLS VPN；Linux 4.14 内核；L3mdev；VRF；OSPF
中图分类号：TP393.1 **文献标志码**：A
doi：10.16836/j.cnki.jcuit.2020.04.003

0 引言

虚拟专用网 VPN 技术的快速发展，多协议标签交换 MPLS 为弥补 VPN 静态扩展等需求，成为企业应用的重要技术，特别是具有 3 层路由功能的 BGP/MPLS VPN 目前已成为 ISP 骨干网的核心技术[1]。大量学者对 BGP/MPLS VPN 进行了深入的研究，在 VPN 隔离方面，大多采用传统的 VPN 隔离或基于 Linux 4.0～4.8内核的策略路由加虚拟转发路由技术 VRF 进行隔离，后者需要显式配置策略路由，操作不够灵活[2]。

为此，在基于 Linux 4.14 内核版本支持 L3mdev 新机制的交换机上，加入 3 层 VRF 功能，搭建小型 BGP/MPLS VPN 自治系统环境，并在 VRF 上部署 OSPF 协议，测试新机制 VRF 功能实现，以及 OSPF 在 VRF 中的有效隔离性。

1 VPN 及 VRF 技术原理

1.1 BGP/MPLS VPN 技术原理

BGP/MPLS VPN 是一种具有路由方式的三层 MPLS VPN，网络模型由运营商的骨干网与用户的各个 Site 组成。其架构主要包括 3 大组件：PE(provider edge)路由器、CE(customer edge)路由器和 P(provider)路由器。PE 是 ISP 骨干网的主要实现者，负责不同 VPN 用户的接入与处理，并支持多个私有路由表的存储与转发；CE 为用户提供到 PE 的路由连接；P 是各个 PE 间连接的桥梁，只负责 MPLS 协议包的高速转发[3]。BGP/MPLS VPN 架构如图 1 所示。

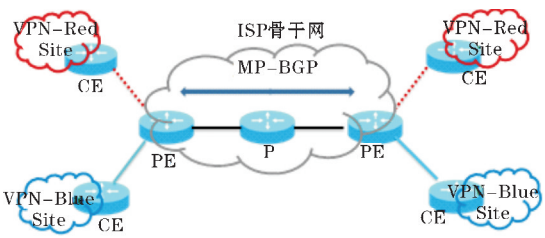

图 1 BGP/MPLS VPN 架构图

BGP/MPLS VPN 构建的 ISP 骨干网中，PE-PE 之间采用支持 IPv6 的扩展 BGP 协议 MP-BGP，跨自治域系统，创建 MP-BGP 隧道使得 VPN 报文在骨干网上透明转发[4]。CE-PE 之间通过静态路由或动态路由将 VPN 本地 Site 路由信息通告给 PE，同时从 PE 学习该 VPN 其他 Site 路由信息[5]。不同 VPN 路由信息都将作为 VRF 实例存储在 PE 中，以实现不同 VPN 用户之间独立的通信。

1.2 VRF 技术转发原理

VRF 作为 VPN 隔离的关键技术，主要应用于 PE 上，每个 VRF 实例包括一张路由表和转发表以及一组 VRF 接口集合和一组相对应的策略规则[6]。

早期 VRF 在 Linux 中的实现是通过 Net Namespace 技术，该技术通过虚拟化整个七层协议栈，进而实现 3 层隔离，这种虚拟方式太过笨重。Linux 4.3 内核出现后，此时的 VRF 隔离需要依靠外部策略路由的配合，但无法构建完整的 VRF 体系[7]。到了 Linux 4.8内核的出现，则采用 L3mdev 新技术来支持一种 3 层的虚拟网卡，通过隐藏网卡之间的可见性，从而实现 VRF 的隔离。

基于 L3mdev 技术的 VRF 隔离不需要显式配置策略路由，直接在 VRF 网卡中关联路由表查询，操作更

收稿日期：2019-11-28

加简洁。因此,选用 Linux 4.14 内核的交换机,搭建模拟环境,实现 L3mdev 机制的 VRF 功能。

2　VRF 配置设计及仿真实现

2.1　基于 L3mdev 的 VRF 配置设计

L3mdev 机制在创建一个 VRF 虚拟网卡时,系统就将其与一个特定的策略路由表自动关联,完成定向操作[8]。其中,VRF 的配置分两部分:控制路径部分和数据路径部分。

控制路径部分主要完成 VRF 虚拟网卡的创建和策略路由表的关联,相关配置操作如下:

(1)创建与 FIB 表关联的 VRF 设备

ip link add [vrf-name] type vrf [table]

(2)设置默认路由

ip route add [table] unreachable default

(3)三层端口加入 VRF 设备

ip link set dev [ethx] master [vrf-name]

(4)添加路由到路由表

ip route add [table]…

通过以上的 Linux 命令配置后,本地路由和连接路由分别自动从本地表和主表移动到 VRF 表中,至此完成 VRF 设备的创建。

数据路径部分包括网卡收包和本地始发包的 VRF 配置。网卡收包的 VRF 配置部分,需要依次经过网卡驱动程序,netif_receive_skb,ip_rcv 的调用,然后在 L3mdev_ip_rcv 中定位到与收包网卡关联的 VRF Master 虚拟网卡,最后在 L3mdev_fib_rule_match 中取出与 VRF Master 关联的策略路由表,进行路由查找。本地始发包的 VRF 配置时,由于数据包来自 Socket 而不是网卡,因此要为 Socket 绑定一个网卡,在 ip_queue_xmit 中查找路由时,就会在 fib_rule_match 中定位到与 VRF 设备关联的策略路由表,进行路由查找[9-10]。

2.2　仿真实现环境的搭建

为展示 BGP/MPLS VPN 模型中 VRF 隔离技术的实现过程及 OSPF 协议的部署,将模拟 BGP/MPLS VPN 模型搭建小型自治网络系统环境,如图 2 所示。因只需验证 CE 与 PE 之间 VRF 转发功能实现及其 OSPF 协议部署,所以将 PE-P-PE 进行透明化为一台 PE。其中,PC1-(CE-Red)-PC2 与 PC3-(CE-Blue)-PC4 分别代表两个不同分支的用户位于 PE 的左、右两边。

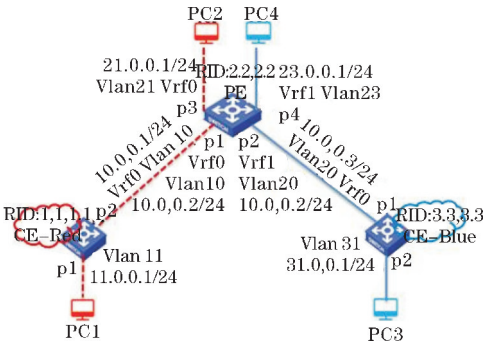

图 2　VRF 功能测试仿真环境

PE 与 CE 均采用 Pica8-5648 型交换机,内核 4.14 版本,支持 L3mdev 机制,并实现了 3 层 VRF 功能。PC1 ~4 分别采用 Ixia 的 4 个接口模拟主机组包与打流,包属性包括源 IP、目的 IP、源 Mac、目的 Mac 及包数目等。

2.3　VRF 中 OSPF 协议部署

搭建好环境后,为三台交换机配置相应的属性,包括:Router ID、Vlan ID、IP、Mac 和 VRF 等,具体属性值参见图 2。其中,CE-Red-p2 与 PE-p1 端口同属于 Vlan10,但分别位于不同子网;CE-Blue-p1、与 PE-p2 端口同属于 Vlan20,但分别位于不同子网;PE-p1 与 PE-p2 端口同属于同一子网,但位于不同 Vlan 中。然后,分别在 CE-Red-p2、PE-p1、PE-p2、CE-Blue-p1 端口 VRF 实例中部署对应的 OSPF 协议[11-12]。

Ixia 上 4 个端口模拟的 4 台 PC 开始组包和打流,其目的是 3 台设备之间进行 OSPF 邻居状态检测与学习。此时,OSPF 邻居状态如图 3 所示。由图 3,CE-Red 中只包含一条通往10.0.0.2、Vlan10 的邻居条目,且存在于默认 VRF0 中。CE-Blue 中只包含一条通往 10.0.0.2、Vlan20 的邻居条目,且存在于默认 VRF0 中。PE 中有一条通往 10.0.0.1、Vlan10 的邻居条目,且存在于默认 VRF0 中;另一条通往 10.0.0.3、Vlan20 的邻居条目,且存在于 VRF1 中。

```
admin@CE-Red# run  show  ospf4  neighbor
Address          Interface              State    Router ID        Pri    Dead
---------------  ---------------------  -------  ---------------  -----  ----
10.0.0.2         vlan10/vlan10          Full     2.2.2.2          128    34

admin@PE# run  show  ospf4  neighbor
Address          Interface              State    Router ID        Pri    Dead
---------------  ---------------------  -------  ---------------  -----  ----
10.0.0.1         vlan10/vlan10          Full     1.1.1.1          128    37
admin@PE# run  show  ospf4  vrf  vrf1 neighbor
Address          Interface              State    Router ID        Pri    Dead
---------------  ---------------------  -------  ---------------  -----  ----
10.0.0.3         vlan20/vlan20          Full     3.3.3.3          128    36

admin@CE-Blue# run  show  ospf4  neighbor
Address          Interface              State    Router ID        Pri    Dead
---------------  ---------------------  -------  ---------------  -----  ----
10.0.0.2         vlan20/vlan20          Full     2.2.2.2          128    30
```

图 3　OSPF 邻居状态图

2.4 VRF 功能及 OSPF 协议测试

2.4.1 不同 VRF 中 OSPF 邻居更新相互隔离

关闭 PE-p2 端口,此时 VRF1 设备关闭将不能学习邻居路由,但 VRF0 不受影响。其更新后的 OSPF 邻居状态如图 4 所示。

```
admin@CE-Red# run  show  ospf4  neighbor
Address          Interface              State     Router ID        Pri    Dead
---------------  ---------------------  --------  ---------------  -----  ----
10.0.0.2         vlan10/vlan10          Full      2.2.2.2          128    34

admin@PE# run  show  ospf4  neighbor
Address          Interface              State     Router ID        Pri    Dead
---------------  ---------------------  --------  ---------------  -----  ----
10.0.0.1         vlan10/vlan10          Full      1.1.1.1          128    36

admin@PE# run  show  ospf4  vrf  vrf1 neighbor
Address          Interface              State     Router ID        Pri    Dead
---------------  ---------------------  --------  ---------------  -----  ----


admin@CE-Blue# run  show  ospf4  neighbor
Address          Interface              State     Router ID        Pri    Dead
---------------  ---------------------  --------  ---------------  -----  ----
```

图 4 OSPF 邻居更新状态图

2.4.2 不同 VRF 中数据转发相互隔离

启用 CE-Red-Vlan11、CE-Blue-Vlan31、PE-Vlan21、PE-Vlan23 的 OSPF Network 功能,命令如下:

CE-Red# set protocols ospf4 area 0.0.0.0 interface vlan11 vif vlan11 address 11.0.0.1

PE# set protocols ospf4 area 0.0.0.0 interface vlan21 vif vlan21 address 21.0.0.1

PE# set protocols ospf4 area 0.0.0.0 interface vlan23 vif vlan23 address 23.0.0.1

CE-Blue# set protocols ospf4 area 0.0.0.0 interface vlan31 vif vlan31 address 31.0.0.1

查看 3 台交换机当前路由表,CE-Red 只包含去往 21.0.0.0/24 的 OSPF 路由,存在于 VRF0 中。PE 包含去往 11.0.0.0/24 的 OSPF 路由,存在于 VRF0 中;去往 31.0.0.0/24 的 OSPF 路由,存在于 VRF1 中。CE-Blue 只包含去往 23.0.0.0/24 的 OSPF 路由,存在于 VRF0 中,如图 5 所示。

```
admin@CE-Red# run  show  route table ipv4 unicast ospf
IPv4 Routing table: 1 routes
21.0.0.0/24          [ospf(110)/2]
                     > to 10.0.0.2 via vlan10/vlan10

admin@PE# run  show  route table ipv4 unicast ospf
IPv4 Routing table: 1 routes
11.0.0.0/24          [ospf(110)/2]
                     > to 10.0.0.1 via vlan10/vlan10

admin@PE# run  show  route vrf vrf1 table ipv4 unicast ospf
IPv4 Routing table: 1 routes
31.0.0.0/24          [ospf(110)/2]
                     > to 10.0.0.3 via vlan20/vlan20

admin@CE-Blue# run  show  route table ipv4 unicast ospf
IPv4 Routing table: 1 routes
23.0.0.0/24          [ospf(110)/2]
                     > to 10.0.0.2 via vlan20/vlan20
```

图 5 设备路由表

在 CE-Red 上分别发送数据到 PE 的左分支 Vlan21 和右分支 Vlan23 端口,如图 6 所示。CE-Red 与 Vlan21 端口同属于 VRF0 相通,Vlan23 端口属于 VRF1 相隔离。

```
admin@CE-Red# run ping 23.0.0.1 -c 5
PING 23.0.0.1 (23.0.0.1) 56(84) bytes of data.

--- 23.0.0.1 ping statistics ---
5 packets transmitted, 0 received, 100% packet loss, time 403

admin@CE-Red# run ping 21.0.0.1 -c 5
PING 21.0.0.1 (21.0.0.1) 56(84) bytes of data.
64 bytes from 21.0.0.1: icmp_req=1 ttl=63 time=2.67 ms
64 bytes from 21.0.0.1: icmp_req=2 ttl=63 time=2.73 ms
64 bytes from 21.0.0.1: icmp_req=3 ttl=63 time=2.50 ms
64 bytes from 21.0.0.1: icmp_req=4 ttl=63 time=2.67 ms
64 bytes from 21.0.0.1: icmp_req=5 ttl=63 time=2.83 ms

--- 21.0.0.1 ping statistics ---
5 packets transmitted, 5 received, 0% packet loss, time 4007m
rtt min/avg/max/mdev = 2.500/2.685/2.839/0.123 ms
```

图 6 VRF 数据传送

通过在以上小型 BGP/MPLS VPN 仿真环境中对 VRF 上 OSPF 协议部署更新及其数据传送的测试,验证基于 Linux 4.14 内核交换机上 VRF 及 OSPF 协议部署的实现。

3 结束语

在基于 Linux 4.14 内核的交换机上实现了 L3mdev 机制的 VRF 功能,并基于 BGP/MPLS VPN 模型搭建仿真测试环境。针对 VRF 功能及其上部署 OSPF 协议进行测试,验证基于 L3mdev 机制的 VRF 上 OSPF 路由隔离的有效性。

只对 ISP 骨干网中 PE 设备上 VRF 功能及 OSPF 协议部署进行测试,针对 VRF 中同时部署 OSPF 和 BGP 混合协议的情况,将是下一步的研究工作。

参考文献:

[1] 王蔚旻,于子恒. 基于 MPLS-VPN 的虚拟网技术实现[J]. 指挥信息系统与技术,2017,8(1):84-90.

[2] Mehraban S, Vora K B, Upadhyay D. Deploy Multi Protocol Label Switching (MPLS) Using Virtual Routing and Forwarding (VRF) [C]. The 2nd International Conference on Trends in Electronics and Informatics(ICOEI-2018),2018:543-548.

[3] 曹弘坚. 浅析 MPLS VPN 技术框架[J]. 科学技术创新,2018(10):72-73.

[4] Tamanna T, T Fatema. MPLS VPN over mGRE design and implementation for a service provider´s

network using GNS3 simulator[C]. International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET-2017) (ICOEI 2017),2017:2339-2341.

[5] 董玲,黄杨,徐塞虹. BGP/MPLS VPN 实现细节探讨[J]. 计算机工程与应用,2005(29):117-119.

[6] 邓丽云. 浅析 MPLS BGP VPN 技术及应用网络[J]. 中国新通信,2019,21(8):112-114.

[7] Pico J A, Fajardo J O, Munoz A, et al. MPLS-VRF integration: forwarding capabilities of BGP/MPLS IP VPN in GNU/Linux[C]. International Conference on Optical Network Design and Modeling on the IEEE 2008.

[8] David Ahern. Using the Linux VRF Solution[EB/OL]. https://cumulusnetworks.com/,2017-9.

[9] Fathima, K M M. A Survey on Multiprotocol Label Switching in Virtual Private Networks [C]. The 2nd International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC 2018),2018(10):737-741.

[10] Yadav S, A Jeyakumar. MPLS multi-VRF design and implementation using GNS simulator[J]. The 2nd IEEE International Conference on Engineering and Technology(ICETECH-2016),2016.

[11] Doyle J. TCP/IP 路由技术[M]. 北京:人民邮电出版社,2009:234-384.

[12] 赵新胜,陈美娟. 路由与交换技术[M]. 北京:人民邮电出版社,2018:89-159.

VRF Configuration Design and Simulation based on BGP/MPLS VPN

CHEN Fengqin, DOU Jun, ZHANG Dan

(College of Information Science & Technology, Southwest Jiaotong University, Chengdu 611756, China)

Abstract: BGP/MPLS VPN is an important model of ISP backbone network, and it is a research object that attracts much attention in the communication field. This paper aims to study the VRF forwarding technology and routing deployment in PE nodes in this model. First, this paper introduces BGP / MPLS VPN model and VRF technology principles; then the VRF function based on the L3mdev mechanism on the Linux 4.14 kernel switch is configured; finally, simulation tests of the VRF and the deployment of OSPF protocol are done to verify its effective isolation.

Keywords: BGP/MPLS VPN; Linux 4.14 kernel; L3mdev; VRF; OSPF