

BGP MPLS VPN 中 CE 双归属 场景下路由优化研究

Research on Routing Optimization in CE Dual Attribution Scenario in BGP MPLS VPN

钟兴宇, 韩梅 (天津理工大学, 天津 300384)

Zhong Xingyu, Han Mei (Tianjin University of Technology, Tianjin 300384, China)

摘要:

BGP MPLS VPN 网络架构是一种由 MP-BGP 技术和 MPLS 技术相结合的新型 VPN 网络架构。为了提高 BGP MPLS VPN 网络架构中接入层 PE 的可靠性, 通常部署 CE 双归属架构, 然而该方案在特殊组网下, 可能会存在次优路径。详细分析了一种基于 CE 双归属的特殊组网场景, 针对存在的问题部署了 OSPF 虚连接优化方案, 并使用 ENSP 软件进行模拟仿真, 验证该方案的可行性。

关键词:

边界网关协议; 多协议标签交换; CE 双归属; 路由优化

doi: 10.12045/j.issn.1007-3043.2024.05.011

文章编号: 1007-3043(2024)05-0062-06

中图分类号: TN919

文献标识码: A

开放科学(资源服务)标识码(OSID):



Abstract:

BGP MPLS VPN network architecture is a new VPN network architecture combined by MP-BGP technology and MPLS technology. In order to improve the reliability of access layer PE in BGP MPLS VPN network architecture, the common solution is to deploy CE dual attribution architecture. However, this scheme may have suboptimal paths under special networking. It analyzes a special networking scenario based on a CE dual attribution, deploys OSPF virtual connection optimization scheme for the existing problems, and uses ENSP software to verify the feasibility of the optimization scheme.

Keywords:

Border gateway protocol; Multi-protocol label switching; CE dual attribution; Route optimization

引用格式: 钟兴宇, 韩梅. BGP MPLS VPN 中 CE 双归属场景下路由优化研究[J]. 邮电设计技术, 2024(5): 62-67.

0 引言

近年来, 为了提高企业网络的可扩展性及可维护性, 众多企业纷纷部署 BGP MPLS VPN 架构, OSPF 协议作为企业中运用最广泛的 IGP 协议, 通常被用作底层协议来承载 MPLS VPN 网络^[1]。同时, 为了提高 MPLS VPN 网络中接入层 PE 的可靠性, 除了提升网络设备自身的可靠性外, 还可以将用户边缘路由器 (Custom Edge Router, CE) 双上联至 2 台服务商边缘路由器 (Provider Edge Router, PE) 上, 与 2 个 PE 共同组成 CE

双归属架构^[2]。这种组网方案虽然提高了网络的可靠性, 但却存在路由环路及次优路径的问题。为了避免这些问题, RFC4364 文档在 OSPF 协议的 Option 字段中新增了 DN 特性^[3], 然而在某些特殊组网场景下, 次优路径仍然存在, 还需要设计方案进一步优化。

1 CE 双归属场景相关原理概述

1.1 CE 双归属场景介绍

BGP MPLS VPN 架构是一种由 MP-BGP 技术和 MPLS 技术相结合的 VPN 新型网络架构, 解决了传统 Site-to-Site VPN 网络架构可扩展性及可维护性不强的问题。该网络架构由 CE、PE 和服务商路由器 (Pro-

收稿日期: 2024-03-06

vider Router, P)设备组成。其中CE是与公网相连的网络设备,可以是路由器或一台主机,P是运营商核心设备,负责MPLS报文的快速交换和转发,而VPN用户的接入、VPNv4路由信息的管理维护都由PE完成^[4-6]。在该架构中,若CE仅上联至一台PE,则存在单点故障的风险。为了提高网络的可靠性,通常会将CE1通过2条链路分别连接至双上层PE,实现流量传输的负载分担和冗余备份(见图1)。

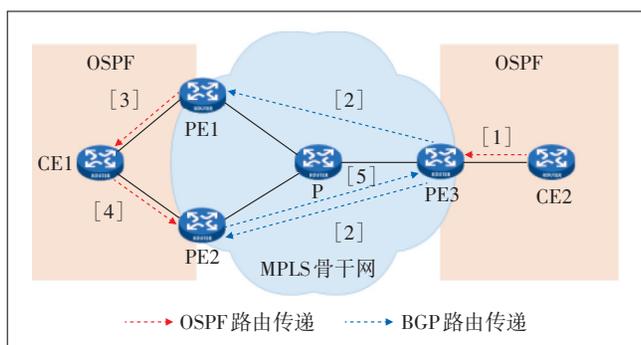


图1 CE双归属场景网络拓扑

1.2 CE双归属场景问题分析

在企业BGP MPLS VPN网络基本实现中,MPLS骨干网通常采用OSPF协议来承载,并在PE之间通过MP-BGP协议跨MPLS骨干网交互VPN路由信息,这种可扩展性好、灵活性强、转发效率高的组网方案受到了各大公司的青睐^[7-8]。但CE双归属场景存在路由环路和次优路径的问题。以图1的拓扑为例,对该问题的产生原理及解决方案进行分析。

当CE1的私网路由需要和CE2交互时,执行的报文转发流程如下。

步骤1:CE2通过OSPF将私网路由汇总成1类LSA或2类LSA的形式发送给PE3。

步骤2:PE3收到从CE2传来的OSPF路由后,将该路由引入到MP-BGP中,并通过MP-BGP邻居通告给PE1和PE2,路由优先级为255。

步骤3:PE1收到PE3通告的BGP VPNv4路由后,将该路由引入到多实例OSPF中,然后以3类LSA的形式在OSPF区域中泛洪(假设两端OSPF域的Domain-ID相同)。

步骤4:PE2收到PE1泛洪的3类LSA后保存至自己的LSDB中,并计算出优先级为10的OSPF内部路由。

步骤5:PE2将该OSPF内部路由重新引入到MP-BGP,再通过MP-BGP邻居通告给PE3。由于这条

VPN私网路由是PE3始发的,路由形成环路。

当PE2上有用户需要访问CE2时,根据路由转发原则,在目的地址相同的情况下,会进行路由优先级的比较。因此PE2会优先选择从PE1传来的优先级为10的OSPF路由,导致PE2上的VPN用户访问PE3的路径为PE2-CE1-PE1-P-PE3,产生了次优路径。

1.3 CE双归属场景解决方案

为了解决CE双归属场景下产生的路由环路及次优路径问题,RFC4364文档在OSPF协议的Option字段中新增了DN特性。

如图2所示,当PE1收到由PE3传来的MP-BGP路由,再以3类LSA的形式引入到多实例OSPF时,会将该LSA的DN位设置为1。当PE2运行的多实例OSPF收到这条DN位设置为1的LSA后,仅将其保存在LSDB中,不再参与OSPF路由计算^[9],因此该路由不会被PE2再次引入到MP-BGP中,消除了路由环路的可能性。PE2上连接的VPN用户若要访问CE2,也会选择PE3传递过来的BGP路由,消除了次优路径。

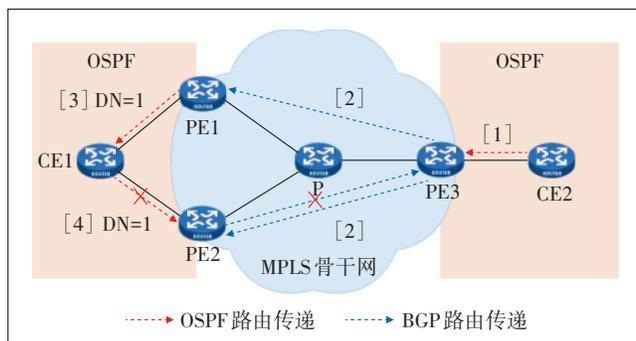


图2 携带DN位的CE双归属场景示意

2 特殊组网下CE双归属场景研究

随着网络规模的日益扩大,企业在部署OSPF时,通常会采用多区域的部署方式,即把大型网络划分成若干个较小的区域,用于控制区域内泛洪的LSA数量和缩小路由震荡产生的范围,减轻OSPF路由器的工作负担,使网络更便于维护管理^[10-11]。然而在某些特殊组网场景下,即便引入了DN位,次优路径的问题仍然存在,还需要结合其他技术来规避这些问题,下文对该问题的产生原因及解决方案进行分析。

2.1 特殊组网下CE双归属场景分析

特殊场景下的CE双归属拓扑如图3所示。通过OSPF将自治系统划分成不同的区域(见图3左侧),其中PE1和CE1之间的链路被划分在骨干区域,CE1和

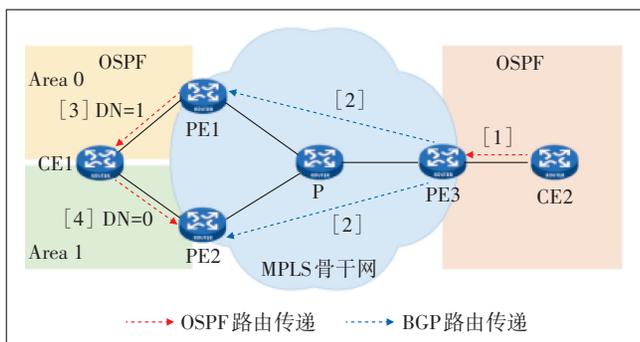


图3 特殊场景下的CE双归属拓扑

PE2之间的链路被划分在非骨干区域,然后在PE1和PE2上引入BGP和多实例OSPF部署的双点双向路由。

在该场景下,PE1会把PE3传来的BGP路由以3类LSA的形式引入到多实例OSPF(假设两端OSPF域的Domain-ID相同),并将DN位设置为1。但CE1作为连接骨干区域和非骨干区域的区域边界路由器(Area Border Router, ABR),在收到3类LSA并泛洪至Area 1时,会清除原先LSA携带的DN位,即将DN位再次置0。

同时,PE2也会把PE3传来的BGP路由以3类LSA的形式引入到多实例OSPF中,由于两端OSPF域中的Domain相同,中间的MPLS骨干网可以视为一个超级骨干区域(Super Backbone)。在OSPF的基本实现中,区域的优先级大小为:Super Backbone>非骨干区域(Area 1),导致PE2上的VPN用户访问CE1时,会优先选择PE3通过MP-BGP传来的路由(属于Super Backbone),而非直接通过CE1访问(属于Area 1),产生次优路径。

2.2 OSPF 虚连接技术

为了解决特殊场景下的CE双归属产生的次优路径,在该场景中部署OSPF虚连接(Virtual link)。由于初期网络规划不合理等原因,经常会出现OSPF骨干区域被分割的情况,但在OSPF的基本实现中,所有非骨干区域必须和骨干区域相连,一旦骨干区域被分割,就会导致OSPF路由学习不正常。OSPF虚连接技术主要用于解决骨干区域不连续导致的OSPF路由学习错误问题。

当骨干区域被分割后,RT2和RT4都各自连接着骨干区域,因此它们都成为了ABR(见图4)。根据OSPF的基本实现原理,ABR从骨干区域学习到的LSA不能再向骨干区域泛洪,因此RT2从RT5学习到的路由不能再向RT1传播。同理,RT4也不会向RT5传递

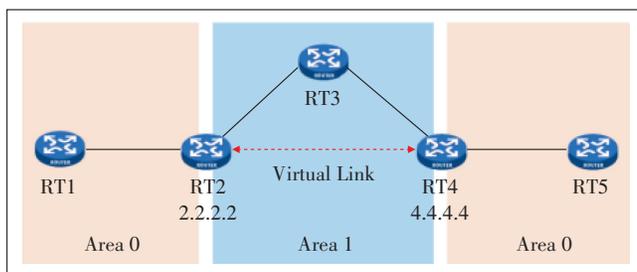


图4 OSPF骨干区域被分割示意

从RT1学习的路由。这样一来,RT1与RT5之间就无法交互OSPF路由信息,导致路由学习出现错误。

此时若要使RT1和RT5之间仍能交互OSPF路由信息,就需要引入虚连接解决方案。该方案的原理是通过相关的配置,使骨干区域两端的ABR通过双方的Router-ID建立点对点的连接,从而建立一条穿越非骨干区域的逻辑通道,并使用单播方式交互OSPF报文,最终实现被分割的骨干区域的连接和扩展。基于上述分析,OSPF虚连接参考配置如表1所示。

表1 OSPF虚连接参考配置

RT2的配置	RT4的配置	命令作用
OSPF 1 router-ID 2.2.2.2	OSPF 1 router-ID 4.4.4.4	进入OSPF进程
area 1	area 1	进入OSPF区域
Vlink-peer 4.4.4.4	Vlink-peer 2.2.2.2	创建并配置OSPF虚连接,虚连接的目的IP必须在区域内可达

3 特殊组网下CE双归属场景仿真

为了解决特殊组网下CE双归属场景产生的次优路径,本文采用了OSPF虚连接技术,并在ENSP上设计实验进行仿真验证,设计的网络拓扑如图5所示。

3.1 IP网络具体规划

由图5可知,MPLS骨干网的AS号为100,使用OSPF协议来保证连通性。为了确保接入层PE的可靠性,将CE1双归属到2台PE上(见图5左侧组网),2台PE之间采用OSPF进行VPN私网路由的交互。为了控制区域内LSA泛洪的数量,对OSPF路由器进行了区域划分。其中CE1和PE1之间的链路被划分为骨干区域,PC1通过CE1接入到Area 0中;CE1和PE2之间的链路被划分为非骨干区域,PC2通过PE2接入到Area 1中。此外,还需要在PE上互相引入VPN私网路由和MP-BGP路由。最后,跟踪PC2访问PC1的流量路径,测试该路径是否为最优路径,具体的IP地址规划如表2所示。

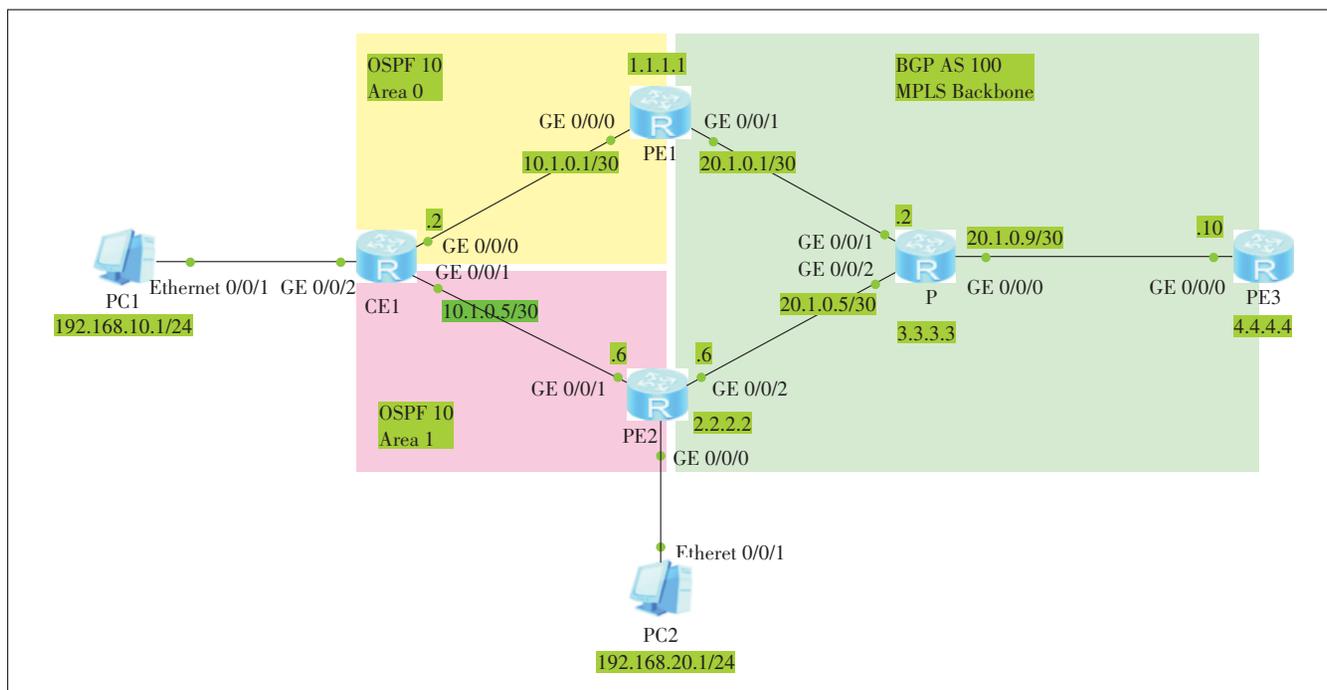


图5 特殊组网下 CE 双归属仿真拓扑

表2 IP地址规划表

网络设备	接口	IP地址/掩码	VPN实例
PE1	G0/0/0	10.1.0.1/30	VPN1
	G0/0/1	20.1.0.1/30	-
	Loopback0	1.1.1.1/32	-
PE2	G0/0/0	192.168.20.254/24	VPN1
	G0/0/1	10.1.0.6/30	VPN1
	G0/0/2	20.1.0.6/30	-
PE3	Loopback0	2.2.2.2/32	-
	G0/0/0	20.1.0.10/30	-
P	Loopback0	4.4.4.4/32	-
	G0/0/0	20.1.0.9/30	-
	G0/0/1	20.1.0.2/30	-
	G0/0/2	20.1.0.5/30	-
CE1	Loopback0	3.3.3.3/32	-
	G0/0/0	10.1.0.2/30	-
	G0/0/1	10.1.0.5/30	-
PC1	-	192.168.10.1/24	-
PC2	-	192.168.20.1/24	-

3.2 网络设备配置

3.2.1 骨干网 OSPF 配置

在 PE1、PE2、PE3 和 P 上使能 OSPF 协议,使得骨干网之间 IP 互通。PE1 的 OSPF 配置如图 6 所示,PE2、PE3 和 P 的 OSPF 配置和 PE1 类似,不再赘述。

```
PE1:
[PE1]ospf 1
[PE1-ospf-1]area 0
[PE1-ospf-1-area-0.0.0.0]network 1.1.1.1 0.0.0.0
[PE1-ospf-1-area-0.0.0.0]network 20.1.0.0 0.0.0.3
```

图6 PE1 OSPF 配置

3.2.2 骨干网 MPLS VPN 配置

在骨干网实现 IP 互通的基础上,在 PE1、PE2、PE3 和 P 上的互连接口使能 MPLS 转发及 MPLS 标签分发功能,并进行相应的 VPN 规划,通过设计合理的 RT 属性值来实现 VPN 路由信息的交互。PE1 的 MPLS VPN 配置如图 7 所示,PE2、PE3 和 P 的 MPLS VPN 配置和

```
PE1:
[PE1]mpls lsr-id 1.1.1.1
[PE1]mpls
[PE1-mpls]mpls ldp
[PE1-mpls-ldp]int g0/0/0
[PE1-GigabitEthernet0/0/0]mpls
[PE1-GigabitEthernet0/0/0]mpls ldp
[PE1]ip vpn-instance vpn1
[PE1-vpn-instance-vpn1]ipv4-family
[PE1-vpn-instance-vpn1-af-ipv4]route-distinguisher 100:1
[PE1-vpn-instance-vpn1-af-ipv4]vpn-target 100:1 export-extcommunity
[PE1-vpn-instance-vpn1-af-ipv4]vpn-target 100:1 import-extcommunity
[PE1]int g0/0/0
[PE1-GigabitEthernet0/0/0]ip binding vpn-instance vpn1
[PE1-GigabitEthernet0/0/0]ip address 10.1.0.1 255.255.255.252
```

图7 PE1 MPLS VPN 配置

PE1 类似, 不再赘述。

3.2.3 PE 和 CE 之间的路由配置

在 PE 和 CE 之间配置 OSPF 协议进行私网路由交互。同时, 为了控制区域内 LSA 泛洪的数量, 减轻路由器的负担, 对 OSPF 进行区域划分, 其配置如图 8 所示。

3.2.4 PE 之间的 MP-BGP 配置

在 PE1、PE2 和 PE3 之间建立 MP-BGP 邻居, 通过 MP-BGP 交互 VPN 私网路由, 并在 PE1 和 PE2 上引入 MP-BGP 和多实例 OSPF 部署双点双向路由。PE1 的 MP-BGP 配置如图 9 所示, PE2 和 PE3 的 MP-BGP 配置和 PE1 类似, 不再赘述。

3.3 实验效果验证

完成上述配置后, 在 PC2 (属于非骨干区域) 上通过 ping 命令测试其能否与 PC1 (属于骨干区域) 正常通信, 并使用 tracert 命令跟踪数据到达 PC1 的传输路径, 对前文的分析结果进行验证。PC2 访问 PC1 测试结果如图 10 所示。

由图 10 可知, PC2 访问 PC1 的路径为 PE2-P-PE1-CE1, 数据流量没有直接通过 OSPF 10 内部的

```
PE1:
[PE1]ospf 10 vpn-instance vpn1 router-id 1.1.1.1
[PE1-ospf-10]area 0
[PE1-ospf-10-area-0.0.0.0]network 10.1.0.0 0.0.0.3
PE2:
[PE2]ospf 10 vpn-instance vpn1 router-id 2.2.2.2
[PE2-ospf-10]area 1
[PE2-ospf-10-area-0.0.0.1]network 10.1.0.4 0.0.0.3
[PE2-ospf-10-area-0.0.0.1]network 192.168.20.0 0.0.0.255
CE1:
[CE1]ospf 10 router-id 5.5.5.5
[CE1-ospf-10]area 0
[CE1-ospf-10-area-0.0.0.0]network 10.1.0.0 0.0.0.3
[CE1-ospf-10-area-0.0.0.0]network 192.168.10.0 0.0.0.255
[CE1-ospf-10]area 1
[CE1-ospf-10-area-0.0.0.1]network 10.1.0.4 0.0.0.3
```

图 8 PE、CE 路由配置

```
PE1:
[PE1]bgp 100
[PE1-bgp]peer 2.2.2.2 as-number 100
[PE1-bgp]peer 2.2.2.2 connect-interface LoopBack0
[PE1-bgp]peer 4.4.4.4 as-number 100
[PE1-bgp]peer 4.4.4.4 connect-interface LoopBack0
[PE1-bgp]ipv4-family vpnv4
[PE1-bgp-af-vpnv4]peer 2.2.2.2 enable
[PE1-bgp-af-vpnv4]peer 4.4.4.4 enable
[PE1-bgp]ipv4-family vpn-instance vpn1
[PE1-bgp-vpn1]import-route ospf 10
[PE1]ospf 10 vpn-instance vpn1
[PE1-ospf-10]import-route bgp
```

图 9 PE1 MP-BGP 配置

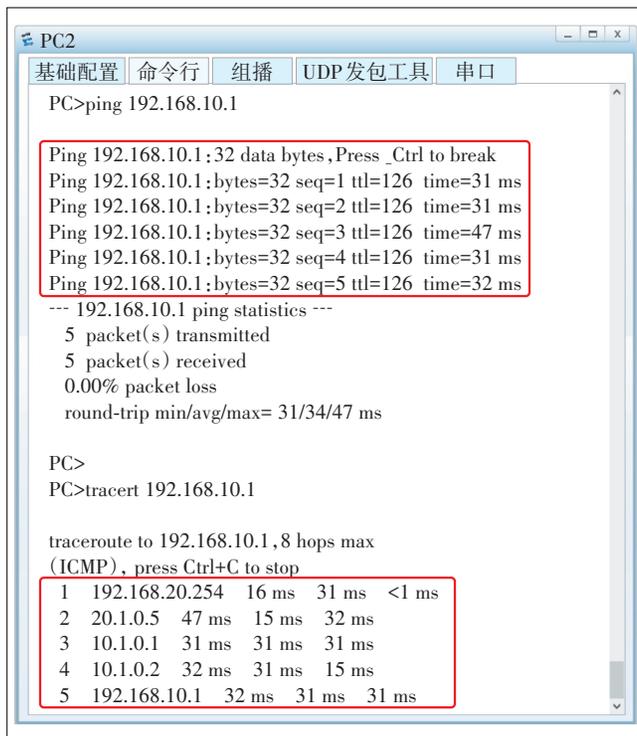


图 10 PC2 访问 PC1 测试结果

CE1 进行访问, 而是跨越了 MPLS 骨干网。这是由于 PE1 和 PE2 上多实例 OSPF 的 Domain-ID 相同, 此时 MPLS 骨干网可视为一个 Super Backbone, 路由优先级大于 OSPF 10 中的 Area 1, 因此产生了次优路径。

3.4 路由优化分析

为了解决 Super Backbone 产生的次优路径问题, 在 CE1 和 PE2 上部署 OSPF 虚连接, 其配置如图 11 所示。配置完成后在 PE2 上通过 dis ospf vlink 命令查看 OSPF 虚连接信息 (见图 12)。

由图 12 可知, PE2 与 CE1 的 OSPF 虚连接状态为 Full, 表明它们之间 OSPF 虚连接的邻居已成功建立, 传输区域为 Area 1。此时该传输区域已经被虚拟成骨干区域, 因此 PE2 能够直接学习到 OSPF 10 中骨干区域的 1 类 LSA 和其他区域的 3 类 LSA。再次在 PC2 上使用 ping 命令和 tracert 命令测试路由优化效果, 具体

```
CE1:
[CE1]ospf 10 router-id 5.5.5.5
[CE1-ospf-10]area 1
[CE1-ospf-10-area-0.0.0.1]vlink-peer 2.2.2.2
PE2:
[PE2]ospf 10 router-id 2.2.2.2
[PE2-ospf-10]area 1
[PE2-ospf-10-area-0.0.0.1]vlink-peer 5.5.5.5
```

图 11 CE1 和 PE2 部署 OSPF 虚连接

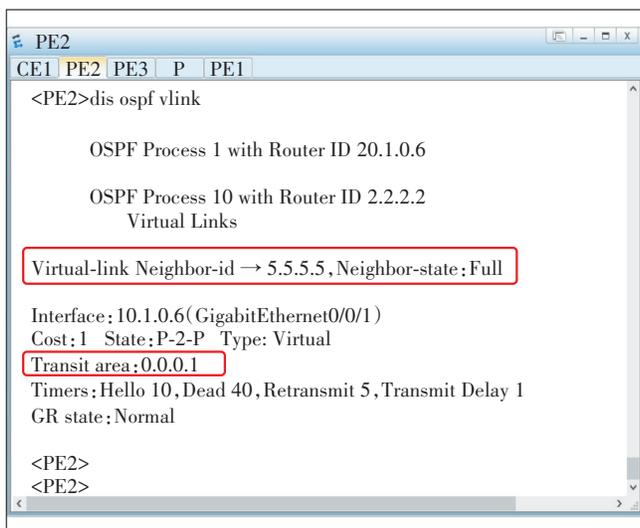


图 12 OSPF 虚连接信息

测试结果如图 13 所示。

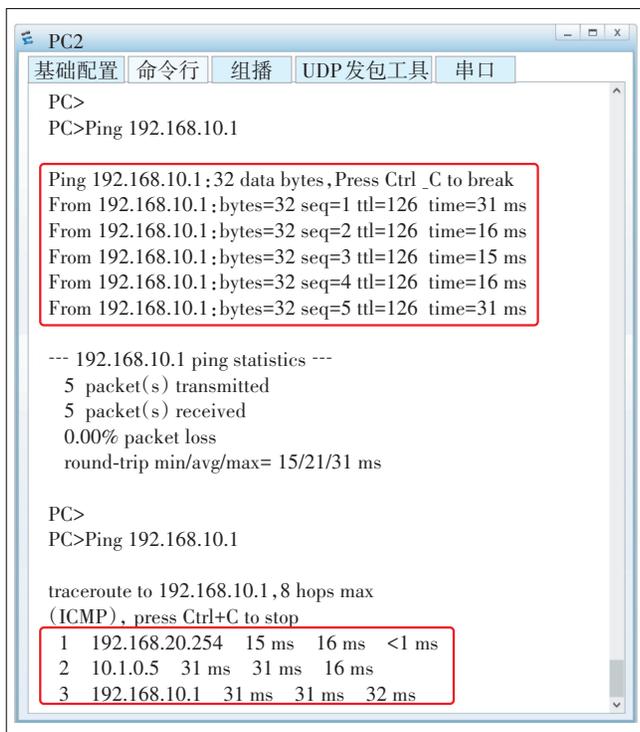


图 13 优化后 PC2 访问 PC1 测试结果

由图 13 可知, PC2 访问 PC1 的路径为 PE2-CE1, 数据流量不再经过 MPLS 骨干网, 次优路径完全消除。

最后, 将 PE2 和 CE1 的互联口断开, 模拟虚链路出现故障的情况。此时继续在 PC2 上访问 PC1, 测试发现若 OSPF 虚连接因链路故障不能保持连通性时, PC2 依然可以通过 Super Backbone 访问 PC1, 提高了 BGP MPLS VPN 网络的可靠性, 实现了特殊组网下 CE 双归

属场景的路由优化目标。

4 结语

传统的 CE 双归属场景通过 OSPF 的 DN 特性进行路由防环和次优路径消除。然而在一些特殊组网下, 次优路径的问题依然存在。本文针对一种特殊组网的 CE 双归属场景, 提出了 OSPF 虚连接的优化方案, 并使用 ENSP 软件进行模拟仿真。仿真结果表明, 该方案不但消除了特殊组网下 CE 双归属场景存在的次优路径, 还能为网络提供一定的冗余性, 满足了 BGP MPLS VPN 网络建设中高效、可靠的组网需求。

参考文献:

- [1] 侯剑锋, 马明凯. MPLS VPN 中 PE-CE 互连仿真研究[J]. 计算机工程, 2010, 36(12): 123-125.
- [2] 陈佑建. 防止双归属组网路由环路方法的研究[J]. 科技资讯, 2015, 13(4): 9-10.
- [3] ROSEN E, REKHTER Y. BGP/MPLS IP Virtual Private Networks (VPNs) [EB/OL]. [2024-01-26]. <https://www.rfc-editor.org/info/rfc4364>.
- [4] 李永芳. 基于 BGP MPLS VPN 企业跨域组网仿真设计[J]. 实验室研究与探索, 2021, 40(3): 121-128.
- [5] 李永芳. 一种跨域铁路数据网综合组网设计与仿真[J]. 实验室研究与探索, 2021, 40(2): 102-108, 126.
- [6] 陈丰琴, 窦军, 张丹. 基于 BGP/MPLS VPN 的 VRF 配置设计与仿真[J]. 成都信息工程大学学报, 2020, 35(4): 378-381.
- [7] 陶志勇, 王如龙, 张锦. 基于负载均衡的跨域虚拟专用网解决方案[J]. 计算机工程, 2014, 40(9): 106-110.
- [8] 王丽娜, 刘炎. BGP/MPLS VPN 中 CE 接入 Internet 仿真[J]. 实验室研究与探索, 2014, 33(3): 76-80, 97.
- [9] 华为技术有限公司. OSPF VPN [EB/OL]. [2024-01-26]. https://support.huawei.com/hedex/pages/EDOC1100196216AZK03238/03/EDOC1100196216AZK03238/03/resources/dc/dc_feature_ospf_0008.html?ft=99&id=&text=null&docid=EDOC1100196216.
- [10] 郭丽, 刘海燕. 基于 Packet Tracer 的 OSPF 多区域中 LSA 的类型研究[J]. 电脑知识与技术, 2020, 16(36): 1-2.
- [11] 郭佳. 多区域 OSPF 路由协议实验的设计与实现[J]. 科技创新与应用, 2017(29): 114, 116.

作者简介:

钟兴宇, 毕业于天津理工大学, 高级工程师, 硕士, 主要从事计算机网络的研究工作; 韩梅, 毕业于天津师范大学, 教授, 硕士生导师, 主要从事计算机应用研究工作。

