



(12)发明专利

(10)授权公告号 CN 106878166 B

(45)授权公告日 2020.04.03

(21)申请号 201710047654.3

H04L 12/703(2013.01)

(22)申请日 2017.01.22

H04L 12/741(2013.01)

(65)同一申请的已公布的文献号

申请公布号 CN 106878166 A

H04L 12/46(2006.01)

(43)申请公布日 2017.06.20

(73)专利权人 新华三技术有限公司

地址 310052 浙江省杭州市滨江区长河路
466号

(56)对比文件

CN 106130819 A,2016.11.16,

CN 105577417 A,2016.05.11,

CN 104486214 A,2015.04.01,

US 2016149751 A1,2016.05.26,

审查员 岳晋

(72)发明人 陈岩 王伟

(74)专利代理机构 北京博思佳知识产权代理有

限公司 11415

代理人 陈蕾

(51)Int.Cl.

H04L 12/707(2013.01)

H04L 12/709(2013.01)

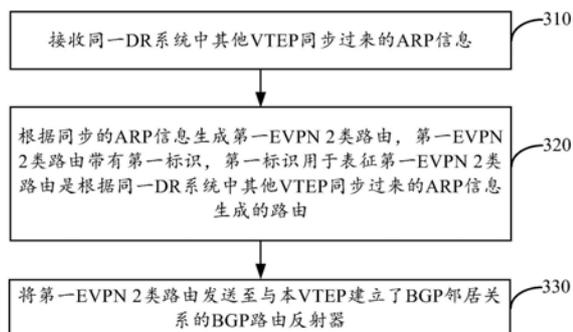
权利要求书3页 说明书10页 附图7页

(54)发明名称

路由通告方法及装置

(57)摘要

本申请提供一种路由通告方法和装置,该方法应用于DR系统中的VTEP,包括:接收同一DR系统中其他VTEP同步过来的ARP信息;根据同步的ARP信息生成第一以太网虚拟专用网络EVPN 2类路由,所述第一EVPN 2类路由带有第一标识,所述第一标识用于表征所述第一EVPN 2类路由是根据同一DR系统中其他VTEP同步过来的ARP信息生成的路由;将所述第一EVPN 2类路由发送至与本VTEP建立了边界网关协议BGP邻居关系的BGP路由反射器。因此,本申请可以实现接收到同步过来的ARP信息的VTEP也能触发上报路由,避免了单一路由所在链路出现故障而造成的流量中断,从而提高了EVPN网络的可靠性。



1. 一种路由通告方法,其特征在于,所述方法应用于分布式聚合DR系统中的可扩展虚拟局域网络隧道端点VTEP,所述方法包括:

所述DR系统中接收到地址解析协议ARP请求报文的第一VTEP通过学习所述ARP请求报文,得到ARP信息;所述第一VTEP基于所述ARP信息生成不带第一标识的第二以太网虚拟专用网络EVPN 2类路由,并发送至路由反射器,以使所述路由反射器将所述第二EVPN 2类路由发送至不属于所述DR系统的远端VTEP;所述第一VTEP向属于所述DR系统的第二VTEP同步所述ARP信息;

所述第二VTEP根据同步的ARP信息生成带有所述第一标识的第一EVPN 2类路由,并发送至所述路由反射器,以使所述路由反射器将所述第一EVPN 2类路由发送至所述远端VTEP;所述第一标识用于表征根据同步的ARP信息生成的EVPN 2类路由;

以使所述远端VTEP根据不带所述第一标识的第二EVPN 2类路由生成主可扩展虚拟局域网络VXLAN隧道,并通过所述主VXLAN隧道进行报文转发;根据带有所述第一标识的第一EVPN 2类路由生成备份VXLAN隧道,并在所述主VXLAN隧道故障时,激活所述备份VXLAN隧道,通过所述备份VXLAN隧道进行报文转发。

2. 根据权利要求1所述的方法,其特征在于,所述第二VTEP根据同步的ARP信息生成带有所述第一标识的第一EVPN 2类路由之前,所述方法还包括:

所述第二VTEP接收所述ARP信息;

若所述ARP信息包括第二标识,所述第二标识用于表征所述ARP信息是同步信息,则确定接收到的ARP信息是所述DR系统中所述第一VTEP同步过来的信息。

3. 根据权利要求2所述的方法,其特征在于,所述第一VTEP向属于所述DR系统的第二VTEP同步所述ARP信息,包括:

所述第一VTEP对学习到的所述ARP信息添加所述第二标识,将添加所述第二标识后的ARP信息同步给所述第二VTEP。

4. 根据权利要求3所述的方法,其特征在于,所述第一EVPN 2类路由和所述第二EVPN 2类路由均带有本地桥MAC地址。

5. 一种路由通告方法,其特征在于,所述方法应用于BGP路由反射器,所述方法包括:

接收分布式聚合DR系统中与本BGP路由反射器建立了BGP邻居关系的第一VTEP发送的EVPN 2类路由;

当所述EVPN 2类路由带有第一标识,所述第一标识用于表征所述EVPN 2类路由是所述第一VTEP根据所述DR系统中其他VTEP同步过来的ARP信息生成的路由,则将所述EVPN 2类路由设置为次优路由;

当所述EVPN 2类路由不带有所述第一标识,则将所述EVPN 2类路由设置为主优路由;

将所述次优路由和/或所述主优路由发送至与本BGP路由反射器建立了BGP邻居关系且不属于所述DR系统的远端VTEP,以使所述远端VTEP根据不带有所述第一标识的EVPN 2类路由生成主可扩展虚拟局域网络VXLAN隧道,并通过所述主VXLAN隧道进行报文转发;根据带有所述第一标识的EVPN 2类路由生成备份VXLAN隧道,并在所述主VXLAN隧道故障时,激活所述备份VXLAN隧道,通过所述备份VXLAN隧道进行报文转发。

6. 根据权利要求5所述的方法,其特征在于,所述将所述次优路由和/或所述主优路由发送至与本BGP路由反射器建立了BGP邻居关系且不属于所述DR系统的远端VTEP,包括:

根据同一ARP信息对应的主优路由和次优路由生成待反射的EVPN 2类路由,所述待反射的EVPN 2类路由包括所述对应的主优路由的下一跳和所述对应的次优路由的下一跳;
将所述待反射的EVPN 2类路由发送至所述远端VTEP。

7. 一种路由通告方法,其特征在于,所述方法应用于远端VTEP,所述方法包括:

接收与本VTEP建立了BGP邻居关系的BGP路由反射器发送的EVPN 2类路由,其中,所述EVPN 2类路由为分布式聚合DR系统中与所述BGP路由反射器建立了BGP邻居关系的第一VTEP发送的;

当所述EVPN 2类路由不带有第一标识,所述第一标识用于表征所述EVPN 2类路由是所述第一VTEP根据DR系统中其他VTEP同步过来的ARP信息生成的路由,则确定所述EVPN 2类路由为主优路由,并根据所述主优路由生成主可扩展虚拟局域网VXLAN隧道,以及通过所述主VXLAN隧道进行报文转发;

当所述EVPN 2类路由带有所述第一标识,则确定所述EVPN 2类路由为次优路由,并根据所述次优路由生成所述主VXLAN隧道备份的备份VXLAN隧道;

当检测到所述主优路由所在链路出现故障时,激活所述备份VXLAN隧道,以及通过所述备份VXLAN隧道进行报文转发。

8. 一种路由通告装置,其特征在于,所述装置应用于分布式聚合DR系统中的可扩展虚拟局域网隧道端点VTEP,所述装置包括:

第二路由生成单元,用于所述DR系统中接收到地址解析协议ARP请求报文的第一VTEP通过学习所述ARP请求报文,得到ARP信息;所述第一VTEP基于所述ARP信息生成不带第一标识的第二以太网虚拟专用网络EVPN 2类路由;

第二发送单元,用于所述第一VTEP将所述第二EVPN 2类路由发送至路由反射器,以使所述路由反射器将所述第二EVPN 2类路由发送至不属于所述DR系统的远端VTEP;所述第一VTEP向属于所述DR系统的第二VTEP同步所述ARP信息;

第一路由生成单元,用于所述第二VTEP根据同步的ARP信息生成带有所述第一标识的第一EVPN 2类路由,所述第一标识用于表征根据同步的ARP信息生成的EVPN 2类路由;

第一发送单元,用于所述第二VTEP将所述第一EVPN 2类路由发送至所述路由反射器,以使所述路由反射器将所述第一EVPN 2类路由发送至所述远端VTEP,以使所述远端VTEP根据不带所述第一标识的第二EVPN 2类路由生成主可扩展虚拟局域网VXLAN隧道,并通过所述主VXLAN隧道进行报文转发;根据带有所述第一标识的第一EVPN 2类路由生成备份VXLAN隧道,并在所述主VXLAN隧道故障时,激活所述备份VXLAN隧道,通过所述备份VXLAN隧道进行报文转发。

9. 根据权利要求8所述的装置,其特征在于,所述装置还包括:

接收子单元,用于所述第二VTEP接收所述ARP信息;

确定子单元,用于若所述ARP信息包括第二标识,所述第二标识用于表征所述ARP信息是同步信息,则确定接收到的ARP信息是所述DR系统中所述第一VTEP同步过来的信息。

10. 根据权利要求9所述的装置,其特征在于,所述装置还包括:

添加单元,用于所述第一VTEP对学习到的所述ARP信息添加所述第二标识,将添加所述第二标识后的ARP信息同步给所述第二VTEP。

11. 根据权利要求10所述的装置,其特征在于,所述第一EVPN 2类路由和所述第二EVPN

2类路由均带有本地桥MAC地址。

12. 一种路由通告装置,其特征在於,所述装置应用于BGP路由反射器,所述装置包括:

接收单元,用于接收分布式聚合DR系统中与本BGP路由反射器建立了BGP邻居关系的第一VTEP发送的EVPN 2类路由;

次优路由设置单元,用于当所述EVPN 2类路由带有第一标识,所述第一标识用于表征所述EVPN 2类路由是所述第一VTEP根据所述DR系统中其他VTEP同步过来的ARP信息生成的路由,则将所述EVPN 2类路由设置为次优路由;

主优路由设置单元,用于当所述EVPN 2类路由不带有所述第一标识,则将所述EVPN 2类路由设置为主优路由;

发送单元,用于将所述次优路由和/或所述主优路由发送至与本BGP路由反射器建立了BGP邻居关系且不属于所述DR系统的远端VTEP,以使所述远端VTEP根据不带有所述第一标识的EVPN 2类路由生成主可扩展虚拟局域网络VXLAN隧道,并通过所述主VXLAN隧道进行报文转发;根据带有所述第一标识的EVPN 2类路由生成备份VXLAN隧道,并在所述主VXLAN隧道故障时,激活所述备份VXLAN隧道,通过所述备份VXLAN隧道进行报文转发。

13. 根据权利要求12所述的装置,其特征在於,所述发送单元包括:

生成子单元,用于根据同一ARP信息对应的主优路由和次优路由生成待反射的EVPN 2类路由,所述待反射的EVPN 2类路由包括所述对应的主优路由的下一跳和所述对应的次优路由的下一跳;

发送子单元,用于将所述待反射的EVPN 2类路由发送至所述远端VTEP。

14. 一种路由通告装置,其特征在於,所述装置应用于远端VTEP,所述装置包括:

接收单元,用于接收与本VTEP建立了BGP邻居关系的BGP路由反射器发送的EVPN 2类路由,其中,所述EVPN 2类路由为分布式聚合DR系统中与所述BGP路由反射器建立了BGP邻居关系的第一VTEP发送的;

主隧道生成单元,用于当所述EVPN 2类路由不带有第一标识,所述第一标识用于表征所述EVPN 2类路由是所述第一VTEP根据DR系统中其他VTEP同步过来的ARP信息生成的路由,则确定所述EVPN 2类路由为主优路由,并根据所述主优路由生成主可扩展虚拟局域网络VXLAN隧道,以及通过所述主VXLAN隧道进行报文转发;

备份隧道生成单元,用于当所述EVPN 2类路由带有所述第一标识,则确定所述EVPN 2类路由为次优路由,并根据所述次优路由生成为所述主VXLAN隧道备份的备份VXLAN隧道;

激活单元,用于当检测到所述主优路由所在链路出现故障时,激活所述备份VXLAN隧道,以及通过所述备份VXLAN隧道进行报文转发。

路由通告方法及装置

技术领域

[0001] 本申请涉及网络通信技术领域,特别涉及一种路由通告方法及装置。

背景技术

[0002] EVPN(Ethernet Virtual Private Network,以太网虚拟专用网络)是一种二层VPN技术,在控制平面上采用BGP(Border Gateway Protocol,边界网关协议)通告路由信息,在数据平面上采用VXLAN(Virtual eXtensible LAN,可扩展虚拟局域网)封装方式转发用户报文。

[0003] 在基本的EVPN组网架构中主要包括:VM(Virtual Machine,虚拟机)、CE(Customer Edge,用户边缘)设备和VTEP(VXLAN Tunnel End Point,VXLAN隧道端点)。

[0004] 目前,为了提高EVPN网络的可靠性,通常会在EVPN网络中应用分布式聚合技术,如图1所示。在如图1所示的组网中,VTEP1和VTEP2组成一个DR(Distributed Relay,分布式聚合)系统。并且,VTEP1、VTEP2、VTEP3和RR(Route Reflector,路由反射器)建立了IBGP(Internal Border Gateway Protocol,内部边界网关协议)EVPN邻居关系。

[0005] 但是,即使在EVPN网络中应用了分布式聚合技术,该EVPN网络的可靠性还需要进一步提高。

发明内容

[0006] 本申请提供一种路由通告方法及装置,以解决现有技术中EVPN网络的可靠性还需要进一步提高的问题。

[0007] 根据本申请实施例的第一方面,提供了一种路由通告方法,所述方法应用于DR系统中的VTEP,所述方法包括:

[0008] 接收同一DR系统中其他VTEP同步过来的ARP信息;

[0009] 根据同步的ARP信息生成第一EVPN 2类路由,所述第一EVPN 2类路由带有第一标识,所述第一标识用于表征所述第一EVPN 2类路由是根据同一DR系统中其他VTEP同步过来的ARP信息生成的路由;

[0010] 将所述第一EVPN 2类路由发送至与本VTEP建立了边界网关协议BGP邻居关系的BGP路由反射器。

[0011] 根据本申请实施例的第二方面,提供了一种路由通告方法,所述方法应用于BGP路由反射器,所述方法包括:

[0012] 接收分布式聚合DR系统中与本BGP路由反射器建立了BGP邻居关系的第一VTEP发送的EVPN 2类路由;

[0013] 当所述EVPN 2类路由带有第一标识,所述第一标识用于表征所述EVPN 2类路由是所述第一VTEP根据所述DR系统中其他VTEP同步过来的ARP信息生成的路由,则将所述EVPN 2类路由设置为次优路由;

[0014] 当所述EVPN 2类路由不带有所述第一标识,则将所述EVPN 2类路由设置为主优路

由；

[0015] 将所述次优路由和/或所述主优路由发送至与本BGP路由反射器建立了BGP邻居关系且不属于所述DR系统的远端VTEP。

[0016] 根据本申请实施例的第三方面,提供了一种路由通告方法,所述方法应用于远端VTEP,所述方法包括:

[0017] 接收与本VTEP建立了BGP邻居关系的BGP路由反射器发送的EVPN 2类路由,其中,所述EVPN 2类路由为分布式聚合DR系统中与所述BGP路由反射器建立了BGP邻居关系的第一VTEP发送的;

[0018] 当所述EVPN 2类路由不带有第一标识,所述第一标识用于表征所述EVPN 2类路由是所述第一VTEP根据DR系统中其他VTEP同步过来的ARP信息生成的路由,则确定所述EVPN 2类路由为主优路由,并根据所述主优路由生成主VXLAN隧道,以及通过所述主VXLAN隧道进行报文转发;

[0019] 当所述EVPN 2类路由带有所述第一标识,则确定所述EVPN 2类路由为次优路由,并根据所述次优路由生成为所述主VXLAN隧道备份的备份VXLAN隧道;

[0020] 当检测到所述主优路由所在链路出现故障时,激活所述备份VXLAN隧道,以及通过所述备份VXLAN隧道进行报文转发。

[0021] 根据本申请实施例的第四方面,提供了一种路由通告装置,所述装置应用于DR系统中的VTEP,所述装置包括:

[0022] 接收单元,用于接收同一DR系统中其他VTEP同步过来的地址解析协议ARP信息;

[0023] 第一路由生成单元,用于根据同步的ARP信息生成第一以太网虚拟专用网络EVPN 2类路由,所述第一EVPN 2类路由带有第一标识,所述第一标识用于表征所述第一EVPN 2类路由是根据同一DR系统中其他VTEP同步过来的ARP信息生成的路由;

[0024] 第一发送单元,用于将所述第一EVPN 2类路由发送至与本VTEP建立了边界网关协议BGP邻居关系的BGP路由反射器。

[0025] 根据本申请实施例的第五方面,提供了一种路由通告装置,所述装置应用于BGP路由反射器,所述装置包括:

[0026] 接收单元,用于接收分布式聚合DR系统中与本BGP路由反射器建立了BGP邻居关系的第一VTEP发送的EVPN 2类路由;

[0027] 次优路由设置单元,用于当所述EVPN 2类路由带有第一标识,所述第一标识用于表征所述EVPN 2类路由是所述第一VTEP根据所述DR系统中其他VTEP同步过来的ARP信息生成的路由,则将所述EVPN 2类路由设置为次优路由;

[0028] 主优路由设置单元,用于当所述EVPN 2类路由不带有所述第一标识,则将所述EVPN 2类路由设置为主优路由;

[0029] 发送单元,用于将所述次优路由和/或所述主优路由发送至与本BGP路由反射器建立了BGP邻居关系且不属于所述DR系统的远端VTEP。

[0030] 根据本申请实施例的第六方面,提供了一种路由通告装置,所述装置应用于远端VTEP,所述装置包括:

[0031] 接收单元,用于接收与本VTEP建立了BGP邻居关系的BGP路由反射器发送的EVPN 2类路由,其中,所述EVPN 2类路由为分布式聚合DR系统中与所述BGP路由反射器建立了BGP

邻居关系的第一VTEP发送的；

[0032] 主隧道生成单元,用于当所述EVPN 2类路由不带有第一标识,所述第一标识用于表征所述EVPN 2类路由是所述第一VTEP根据DR系统中其他VTEP同步过来的ARP信息生成的路由,则确定所述EVPN 2类路由为主优路由,并根据所述主优路由生成主VXLAN隧道,以及通过所述主VXLAN隧道进行报文转发；

[0033] 备份隧道生成单元,用于当所述EVPN 2类路由带有所述第一标识,则确定所述EVPN 2类路由为次优路由,并根据所述次优路由生成为所述主VXLAN隧道备份的备份VXLAN隧道；

[0034] 激活单元,用于当检测到所述主优路由所在链路出现故障时,激活所述备份VXLAN隧道,以及通过所述备份VXLAN隧道进行报文转发。

[0035] 应用本申请实施例,通过接收同一DR系统中其他VTEP同步过来的ARP信息,根据同步的ARP信息生成第一EVPN 2类路由,第一EVPN 2类路由带有第一标识,将第一EVPN 2类路由发送至与本VTEP建立了BGP邻居关系的BGP路由反射器,使得接收到同步过来的ARP信息的VTEP也能触发上报路由,避免了单一路由所在链路出现故障而造成的流量中断,从而提高了EVPN网络的可靠性。

附图说明

[0036] 图1是EVPN网络的一个架构示意图；

[0037] 图2是EVPN网络的另一个架构示意图；

[0038] 图3是本申请路由通告方法的一个实施例流程图；

[0039] 图4是本申请路由通告方法的另一个实施例流程图；

[0040] 图5是本申请路由通告方法的一个实施例流程图；

[0041] 图6是本申请路由通告方法的一个实施例流程图；

[0042] 图7是本申请路由通告装置所在设备的一种硬件结构示意图；

[0043] 图8是本申请路由通告装置的一个实施例框图；

[0044] 图9是本申请路由通告装置的一个实施例框图；

[0045] 图10是本申请路由通告装置的一个实施例框图。

具体实施方式

[0046] 这里将详细地对示例性实施例进行说明,其示例表示在附图中。下面的描述涉及附图时,除非另有表示,不同附图中的相同数字表示相同或相似的要素。以下示例性实施例中所描述的实施方式并不代表与本申请相一致的所有实施方式。相反,它们仅是与如所附权利要求书中所详述的、本申请的一些方面相一致的装置和方法的例子。

[0047] 在本申请使用的术语是仅仅出于描述特定实施例的目的,而非旨在限制本申请。在本申请和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式,除非上下文清楚地表示其他含义。还应当理解,本文中使用的术语“和/或”是指并包含一个或多个相关联的列出项目的任何或所有可能组合。

[0048] 应当理解,尽管在本申请可能采用术语第一、第二、第三等来描述各种信息,但这些信息不应限于这些术语。这些术语仅用来将同一类型的信息彼此区分开。例如,在不脱离

本申请范围的情况下,第一信息也可以被称为第二信息,类似地,第二信息也可以被称为第一信息。取决于语境,如在此所使用的词语“如果”可以被解释成为“在……时”或“当……时”或“响应于确定”。

[0049] 参见图1,图1是EVPN网络的一个架构示意图。该EVPN网络中应用了分布式聚合技术。如图1所示。在如图1所示的EVPN组网中,DR系统包括两个VTEP(即VTEP1、VTEP2),并且,VTEP1和VTEP2之间的链路为IPL链路(Intra-Portal Link,分布式聚合内部链路)。另外,该DR系统还可以包括两个以上的VTEP,比如:3个VTEP等。

[0050] VTEP1和VTEP2通过分布式聚合口DR1与CE1设备连接,CE1设备通过分布式聚合口LAGG1与VTEP1和VTEP2连接,且CE1下挂VM(Virtual Machine,虚拟机)1。还有,VTEP1和VTEP2通过分布式聚合口DR2与CE2设备连接,CE2设备通过分布式聚合口LAGG2与VTEP1和VTEP2连接,且CE2下挂VM2。并且,VTEP1、VTEP2均与RR建立了IBGP EVPN邻居关系。

[0051] 在如图1所示的EVPN组网中,包括一个远端VTEP(即VTEP3),且该VTEP3与RR也建立了IBGP EVPN邻居关系。另外,该EVPN网络中还可以包括多个远端VTEP,比如:2个远端VTEP等。

[0052] 当CE1向DR系统发送ARP请求报文,LAGG1通过哈希算法选取其中一条链路发送该ARP(Address Resolution Protocol,地址解析协议)请求报文,假如选取的链路是VTEP1和CE1这条链路,则:

[0053] 当VTEP1接收到CE1发送的ARP请求报文后,学习该ARP请求报文,得到ARP信息,会根据该ARP信息生成EVPN 2类路由,并将该EVPN 2类路由发送至RR,以及由RR反射给VTEP3。同时,由于VTEP1和VTEP2位于同一个DR系统,VTEP1还会将该ARP信息同步给VTEP2。

[0054] 假如,VTEP2接收到ARP信息后,也会根据该ARP信息生成EVPN 2类路由,并将该EVPN 2类路由发送至RR,以及由RR反射给VTEP3,这样可能会存在多种时序问题,并对EVPN网络的可靠性带来新的风险。比如:VTEP3先接收到VTEP2上报的ARP信息,可能会认为VTEP1下的CE1发生迁移。

[0055] 假如,VTEP2接收到ARP信息后,不会根据该ARP信息生成EVPN 2类路由并发送至RR。当VTEP1和RR之间的链路出现故障、或VTEP1出现故障时,VTEP1发送给RR的EVPN 2类路由会被撤销,导致发射给VTEP3的EVPN 2类路由也会被撤销,这样导致VTEP3上没有到CE1的转发表项,不能指导转发报文,从而导致单一链路出现故障而造成流量中断。

[0056] 为了解决上述问题,本申请实施例中提供了一种路由通告方法,以及一种可以应用该方法的路由通告装置。

[0057] 参见图2,图2是EVPN网络的另一个架构示意图:

[0058] 当VTEP1接收到CE1发送的ARP请求报文后,学习该ARP请求报文,得到ARP信息,会将学习到的ARP信息同步给VTEP2。

[0059] VTEP2接收到VTEP1同步过来的ARP信息后,会根据同步的ARP信息生成第一EVPN 2类路由,第一EVPN 2类路由带有第一标识,第一标识用于表征第一EVPN 2类路由是根据同一DR系统中其他VTEP同步过来的ARP信息生成的路由,以及将第一EVPN 2类路由发送至RR。另外,VTEP1也会根据学习到的ARP信息生成第二EVPN 2类路由,将第二EVPN 2类路由发送至RR。

[0060] RR接收到EVPN 2类路由后,会识别EVPN 2类路由是否带有第一标识,若带有第一

标识,表明接收到的EVPN 2类路由是VTEP2根据同步的ARP信息生成的路由,则将该EVPN 2类路由设置为次优路由;若不带有第一标识,表明接收到的EVPN 2类路由是VTEP1根据学习到的ARP信息生成的路由,则将该EVPN 2类路由设置为主优路由,以及将次优路由和/或主优路由发送至与VTEP3。

[0061] VTEP3接收到EVPN 2类路由后,也会识别EVPN 2类路由是否带有第一标识,若不带有第一标识,则确定该EVPN 2类路由为主优路由,并根据主优路由生成主VXLAN隧道,以及通过主VXLAN隧道进行报文转发;若带有第一标识,确定该EVPN 2类路由为次优路由,并根据次优路由生成为主VXLAN隧道备份的备份VXLAN隧道;以及当检测到主优路由所在链路出现故障时,激活备份VXLAN隧道,以及通过备份VXLAN隧道进行报文转发。

[0062] 因此,通过VTEP2向RR发送第一EVPN 2类路由,这样可以避免出现单一链路出现故障而造成流量中断的情形;并且,第一EVPN 2类路由带有第一标识,以使RR和VTEP3接收到EVPN 2类路由后,可以根据第一标识来区分接收到的EVPN 2类路由是VTEP2发送的还是VTEP1发送的,这样可以避免由于VTEP2也发送EVPN 2类路由而导致的时序问题和新的风险,从而提高了EVPN网络的可靠性。

[0063] 下面结合附图对本申请路由通告实施例进行详细描述。

[0064] 参见图3,为本申请路由通告方法的一个实施例流程图,该方法可以应用于DR系统中的VTEP,比如,图2中的VTEP1或VTEP2,可以包括以下步骤:

[0065] 步骤310,接收同一DR系统中其他VTEP同步过来的ARP信息。

[0066] 本申请实施例中,当本VTEP接收到本地CE设备发送的ARP请求报文时,会对ARP请求报文进行学习,得到ARP信息,并将该ARP信息同步至同一DR系统中其他VTEP。同理,其他VTEP学习到ARP信息后,也会同步至本VTEP。

[0067] 在一实施例中,为了区分接收到的ARP信息是否为同一DR系统中其他VTEP同步过来的,可以利用第二标识进行识别,具体过程包括:

[0068] 接收ARP信息;

[0069] 若该ARP信息包括第二标识,第二标识用于表征ARP信息是同步信息,则确定接收到的ARP信息是同一DR系统中其他VTEP同步过来的信息。

[0070] 其中,第二标识可以是预先设置的标识,比如:MLAG (Multi-Chassis Link Aggregation,分布式聚合链路)标识。

[0071] 步骤320,根据同步的ARP信息生成第一EVPN 2类路由,第一EVPN 2类路由带有第一标识,第一标识用于表征第一EVPN 2类路由是根据同一DR系统中其他VTEP同步过来的ARP信息生成的路由。

[0072] 本申请实施例中,第一标识可以是预先设置的标识,比如:MLAG标识。

[0073] 在一实施例中,第一EVPN 2类路由还可以带有本地桥MAC地址。其中,同一DR系统中各个VTEP的本地桥MAC地址是相同的。

[0074] 比如:第一EVPN 2类路由带有本地桥MAC地址和MLAG标识。其中,本地桥MAC地址可以表征本VTEP所在的DR系统,MLAG标识可以表征本VTEP是根据同一DR系统中其他VTEP同步过来的ARP信息生成的第一EVPN 2类路由。

[0075] 步骤330,将第一EVPN 2类路由发送至与本VTEP建立了BGP邻居关系的BGP路由反射器。

[0076] 参图2所示,VTEP1接收到CE1发送的ARP请求报文后,通过对ARP请求报文的学习,得到ARP信息,该ARP信息可以包括VM1的MAC地址和IP地址;VTEP1会将该ARP信息同步至VTEP2,VTEP2接收到VTEP1同步过来的ARP信息后,会根据同步的ARP信息生成EVPN 2类路由,该EVPN 2类路由带有VTEP2的桥MAC地址和MLAG标记,将EVPN 2类路由发送至RR,并由RR根据路由中带有桥MAC地址和MLAG标记来确定该EVPN 2类路由为次优路由;VTEP1还会根据ARP信息生成EVPN 2类路由,该EVPN 2类路由带有VTEP1的桥MAC地址,然后将该EVPN 2类路由发送至RR,并由RR根据路由中仅带有桥MAC地址来确定该EVPN 2类路由为主优路由。

[0077] 同理,VTEP2接收到CE2发送的ARP请求报文后,通过对ARP请求报文的学习,得到ARP信息,该ARP信息可以包括CE2下挂的VM2的MAC地址和IP地址;VTEP2会将ARP信息同步至VTEP1.VTEP1接收到VTEP2同步过来的ARP信息后,会根据同步的ARP信息生成EVPN 2类路由,该EVPN 2类路由带有VTEP1的桥MAC地址和MLAG标记,将该EVPN 2类路由发送至与RR,并由RR根据路由中带有桥MAC地址和MLAG标记来确定该EVPN 2类路由为次优路由;VTEP2还会根据该ARP信息生成EVPN 2类路由,该EVPN 2类路由带有VTEP2的桥MAC地址,然后将该EVPN 2类路由发送至RR,并由RR根据路由中仅带有桥MAC地址来确定该EVPN 2类路由为主优路由。

[0078] 由上述实施例可见,通过接收同一DR系统中其他VTEP同步过来的ARP信息,根据同步的ARP信息生成第一EVPN 2类路由,第一EVPN 2类路由带有第一标识,将第一EVPN 2类路由发送至与本VTEP建立了BGP邻居关系的BGP路由反射器,使得接收到同步过来的ARP信息的VTEP也能触发上报路由,避免了单一路由所在链路出现故障而造成的流量中断,从而提高了EVPN网络的可靠性。

[0079] 参见图4,为本申请路由通告方法的另一个实施例流程图,该方法应用于DR系统中的VTEP,比如,图2中的VTEP1或VTEP2,可以包括以下步骤:

[0080] 步骤410,对本地CE设备发送的ARP请求报文进行学习得到ARP信息。

[0081] 步骤420,对学习到的ARP信息添加第二标识,以及将添加第二标识后的ARP信息同步给同一DR系统中其他VTEP,该第二标识用于表征ARP信息是同步信息。

[0082] 其中,第二标识可以是预先设置的标识,比如:MLAG标识。

[0083] 步骤430,根据学习到的ARP信息生成第二EVPN 2类路由。

[0084] 在一实施例中,第二EVPN 2类路由可以带有本地桥MAC地址。由于同一DR系统中各个VTEP的本地桥MAC地址是相同的,故此可以利用路由中自带的本地桥MAC地址,来判断路由发送方是否位于同一DR系统。

[0085] 另外,上述步骤420和步骤430没有执行顺序的限定,可以先执行步骤420,再执行步骤430;也可以先执行步骤430,再执行步骤420;还可以同时执行步骤420和步骤430。步骤440,将第二EVPN 2类路由发送至BGP路由反射器。

[0086] 由上述实施例可见,可以将学习到的ARP信息添加第二标识后,同步给同一DR系统中其他VTEP,并由其他VTEP触发上报路由,还可以根据学习到的ARP信息触发上报路由,增加了单个VTEP的路由通告功能,从而提高了EVPN网络的可靠性。

[0087] 参见图5,为本申请路由通告方法的一个实施例流程图,该方法应用于BGP路由反射器,比如,图2中的RR,可以包括以下步骤:

[0088] 步骤510,接收DR系统中与本BGP路由反射器建立了BGP邻居关系的第一VTEP发送

的EVPN 2类路由。

[0089] 步骤520,当EVPN 2类路由带有第一标识,第一标识用于表征EVPN 2类路由是第一VTEP根据指定DR系统中其他VTEP同步过来的ARP信息生成的路由时,则将EVPN 2类路由设置为次优路由。

[0090] 步骤530,当EVPN 2类路由不带有第一标识,则将EVPN 2类路由设置为主优路由。

[0091] 本申请实施例中,第一标识可以是预先设置的标识,比如:MLAG标识。BGP路由反射器接收到EVPN 2类路由后,会根据该EVPN 2类路由是否带有第一标识来设置路由等级,若带有第一标识,则设置为次优路由,若不带有第一标识,则设置为主优路由。

[0092] 比如:如图2所示,VTEP1根据学习到的ARP信息生成带有桥MAC地址的EVPN 2类路由,以及将EVPN 2类路由发送至RR;VTEP2根据同步的ARP信息生成携带有桥MAC地址和MLAG标识的EVPN 2类路由,然后将该EVPN 2类路由发送至RR;此时,RR会将带有MLAG标识的EVPN 2类路由设置为次优路由,而将不带有MLAG标识的EVPN 2类路由设置为主优路由。

[0093] 步骤540,将次优路由和/或主优路由发送至与本BGP路由反射器建立了BGP邻居关系且不属于指定DR系统的远端VTEP。

[0094] 在一个实施例中,RR可以利用扩展BGP的能力,使BGP反射时可以携带多个路由,即携带主优路由和次优路由,具体包括:

[0095] 根据同一ARP信息对应的主优路由和次优路由生成待反射的EVPN 2类路由,该待反射的EVPN 2类路由包括主优路由的下一跳和次优路由的下一跳;

[0096] 将待反射的EVPN 2类路由发送至与本BGP路由反射器建立了BGP邻居关系且不属于所述指定DR系统的远端VTEP。

[0097] 由上述实施例可见,通过第一标识确定主优路由和次优路由,并将主优路由和次优路由发送至远端VTEP,使得远端VTEP根据主优路由和次优路由生成主VXLAN隧道和备份VXLAN隧道,从而出现了针对同一ARP信息出现主优路由和次优路由并存的情形,避免了主优路由所在链路出现故障而造成的流量中断,从而提高了EVPN网络的可靠性。

[0098] 参见图6,为本申请路由通告方法的一个实施例流程图,该方法应用于远端VTEP,比如,图2中的远端VTEP3,可以包括以下步骤:

[0099] 步骤610,接收与本VTEP建立了BGP邻居关系的BGP路由反射器发送的EVPN 2类路由,其中,EVPN 2类路由为DR系统中与BGP路由反射器建立了BGP邻居关系的第一VTEP发送的。

[0100] 步骤620,当EVPN 2类路由不带有第一标识,第一标识用于表征EVPN 2类路由是第一VTEP根据指定DR系统中其他VTEP同步过来的ARP信息生成的路由时,则确定EVPN 2类路由为主优路由,并根据主优路由生成主VXLAN隧道,以及通过主VXLAN隧道进行报文转发。

[0101] 步骤630,当EVPN 2类路由带有第一标识,则确定EVPN 2类路由为次优路由,并根据次优路由生成为主VXLAN隧道备份的备份VXLAN隧道。

[0102] 本申请实施例中,第一标识可以是预先设置的标识,比如:MLAG标识。远端VTEP接收到EVPN 2类路由后,会根据该EVPN 2类路由是否带有第一标识来建立主备VXLAN隧道,若带有第一标识,则确定为次优路由,并建立备份VXLAN隧道,若不带有第一标识,则确定为主优路由,并建立主VXLAN隧道。

[0103] 比如:如图2所示,VTEP1根据学习到的ARP信息生成带有桥MAC地址的EVPN 2类路

由,以及将EVPN 2类路由发送至RR,RR会将不带有MLAG标识的EVPN 2类路由设置为主优路由,并发送至远端VTEP,远端VTEP也会根据不带有MLAG标识的EVPN 2类路由建立主VXLAN隧道;

[0104] VTEP2根据同步的ARP信息生成携带有桥MAC地址和MLAG标识的EVPN 2类路由,然后将该EVPN 2类路由发送至RR,RR会将带有MLAG标识的EVPN 2类路由设置为次优路由,并发送至远端VTEP,远端VTEP也会根据带有MLAG标识的EVPN 2类路由建立备份VXLAN隧道。

[0105] 步骤640,当检测到主优路由所在链路出现故障时,激活备份VXLAN隧道,以及通过备份VXLAN隧道进行报文转发。

[0106] 由上述实施例可见,通过第一标识确定主优路由和次优路由,并根据主优路由和次优路由,可以生成对应的主VXLAN隧道和备份VXLAN隧道,即使主优路由所在链路出现故障时,还可以激活备份VXLAN隧道,避免了由于出现故障而造成的流量中断,从而提高了EVPN网络的可靠性。

[0107] 与前述路由通告方法实施例相对应,本申请还提供了路由通告装置的实施例。

[0108] 本申请路由通告装置的实施例可以应用在路由通告设备上,该路由通告设备分别为DR系统中的VTEP、BGP路由反射器和远端VTEP。装置实施例可以通过软件实现,也可以通过硬件或者软硬件结合的方式实现。以软件实现为例,作为一个逻辑意义上的装置,是通过其所在设备的处理器将非易失性存储器中对应的计算机程序指令读取到内存中运行形成的。从硬件层面而言,如图7所示,为本申请路由通告装置所在设备的一种硬件结构示意图,除了图7所示的处理器、网络接口、内存以及非易失性存储器之外,实施例中装置所在的设备通常还可以包括其他硬件,如负责处理报文的转发芯片等等;从硬件结构上来讲该设备还可能是分布式的设备,可能包括多个接口卡,以便在硬件层面进行网络地址转换的扩展。

[0109] 参见图8,为本申请路由通告装置的一个实施例框图,该装置应用于DR系统中的VTEP,比如,图2中的VTEP1或VTEP2,该装置可以包括:接收单元81、第一路由生成单元82和第一发送单元83。

[0110] 接收单元81,用于接收同一DR系统中其他VTEP同步过来的地址解析协议ARP信息;

[0111] 第一路由生成单元82,用于根据同步的ARP信息生成第一以太网虚拟专用网络EVPN 2类路由,所述第一EVPN 2类路由带有第一标识,所述第一标识用于表征所述第一EVPN 2类路由是根据同一DR系统中其他VTEP同步过来的ARP信息生成的路由;

[0112] 第一发送单元83,用于将所述第一EVPN 2类路由发送至与本VTEP建立了边界网关协议BGP邻居关系的BGP路由反射器。

[0113] 由上述实施例可见,通过接收同一DR系统中其他VTEP同步过来的ARP信息,根据同步的ARP信息生成第一EVPN 2类路由,第一EVPN 2类路由带有第一标识,将第一EVPN 2类路由发送至与本VTEP建立了BGP邻居关系的BGP路由反射器,使得接收到同步过来的ARP信息的VTEP也能触发上报路由,避免了单一路由所在链路出现故障而造成的流量中断,从而提高了EVPN网络的可靠性。

[0114] 在一个可选的实现方式中,所述接收单元81可以包括:接收子单元和确定子单元(图8中未示出)。

[0115] 接收子单元,用于接收ARP信息;

[0116] 确定子单元,用于若所述ARP信息包括第二标识,所述第二标识用于表征所述ARP

信息是同步信息,则确定接收到的ARP信息是同一DR系统中其他VTEP同步过来的信息。

[0117] 在一个可选的实现方式中,所述装置还可以包括:学习单元、添加单元、第二路由生成单元和第二发送单元(图8中未示出)。

[0118] 学习单元,用于对本地CE设备发送的ARP请求报文进行学习得到ARP信息;

[0119] 添加单元,用于对学习到的ARP信息添加所述第二标识,以及将添加所述第二标识后的ARP信息同步给同一DR系统中其他VTEP;

[0120] 第二路由生成单元,用于根据学习到的ARP信息生成第二EVPN 2类路由;

[0121] 第二发送单元,用于将所述第二EVPN 2类路由发送至所述BGP路由反射器。

[0122] 上述第一EVPN 2类路由和第二EVPN 2类路由还可以均带有本地桥MAC地址。

[0123] 由上述实施例可见,可以将学习到的ARP信息添加第二标识后,同步给同一DR系统中其他VTEP,并由其他VTEP触发上报路由,还可以根据学习到的ARP信息触发上报路由,增加了单个VTEP的路由通告功能,从而提高了EVPN网络的可靠性。

[0124] 参见图9,为本申请路由通告装置的一个实施例框图,该装置应用于BGP路由反射器,比如,图2中的RR,该装置可以包括:接收单元91、次优路由设置单元92、主优路由设置单元93和发送单元94。

[0125] 接收单元91,用于接收DR系统中与本BGP路由反射器建立了BGP邻居关系的第一VTEP发送的EVPN 2类路由;

[0126] 次优路由设置单元92,用于当所述EVPN 2类路由带有第一标识,所述第一标识用于表征所述EVPN 2类路由是所述第一VTEP根据所述DR系统中其他VTEP同步过来的ARP信息生成的路由,则将所述EVPN 2类路由设置为次优路由;

[0127] 主优路由设置单元93,用于当所述EVPN 2类路由不带有所述第一标识,则将所述EVPN 2类路由设置为主优路由;

[0128] 发送单元94,用于将所述次优路由和/或所述主优路由发送至与本BGP路由反射器建立了BGP邻居关系且不属于所述指定DR系统的远端VTEP。

[0129] 在一个可选的实现方式中,发送单元94可以包括:生成子单元和发送子单元(图9中未示出)。

[0130] 生成子单元,用于根据同一ARP信息对应的主优路由和次优路由生成待反射的EVPN 2类路由,所述待反射的EVPN 2类路由包括所述对应的主优路由的下一跳和所述对应的次优路由的下一跳;

[0131] 发送子单元,用于将所述待反射的EVPN 2类路由发送至所述远端VTEP。

[0132] 由上述实施例可见,通过第一标识确定主优路由和次优路由,并将主优路由和次优路由发送至远端VTEP,使得远端VTEP根据主优路由和次优路由生成主VXLAN隧道和备份VXLAN隧道,从而出现了针对同一ARP信息出现主优路由和次优路由并存的情形,避免了主优路由所在链路出现故障而造成的流量中断,从而提高了EVPN网络的可靠性。

[0133] 参见图10,为本申请路由通告装置的一个实施例框图,该装置应用于装置应用于远端VTEP,比如,图2中的VTEP3,该装置可以包括:接收单元101、主隧道生成单元102、备份隧道生成单元103和激活单元104。

[0134] 接收单元101,用于接收与本VTEP建立了BGP邻居关系的BGP路由反射器发送的EVPN 2类路由,其中,所述EVPN 2类路由为DR系统中与所述BGP路由反射器建立了BGP邻居

关系的第一VTEP发送的；

[0135] 主隧道生成单元102,用于当所述EVPN 2类路由不带有第一标识,所述第一标识用于表征所述EVPN 2类路由是所述第一VTEP根据DR系统中其他VTEP同步过来的ARP信息生成的路由,则确定所述EVPN 2类路由为主优路由,并根据所述主优路由生成主VXLAN隧道,以及通过所述主VXLAN隧道进行报文转发；

[0136] 备份隧道生成单元103,用于当所述EVPN 2类路由带有所述第一标识,则确定所述EVPN 2类路由为次优路由,并根据所述次优路由生成为所述主VXLAN隧道备份的备份VXLAN隧道。

[0137] 激活单元104,用于当检测到所述主优路由所在链路出现故障时,激活所述备份VXLAN隧道,以及通过所述备份VXLAN隧道进行报文转发。

[0138] 由上述实施例可见,通过第一标识确定主优路由和次优路由,并根据主优路由和次优路由,可以生成对应的主VXLAN隧道和备份VXLAN隧道,即使主优路由所在链路出现故障时,还可以激活备份VXLAN隧道,避免了由于出现故障而造成的流量中断,从而提高了EVPN网络的可靠性。

[0139] 上述装置中各个单元的功能和作用的实现过程具体详见上述方法中对应步骤的实现过程,在此不再赘述。

[0140] 对于装置实施例而言,由于其基本对应于方法实施例,所以相关之处参见方法实施例的部分说明即可。以上所描述的装置实施例仅仅是示意性的,其中所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本申请方案的目的。本领域普通技术人员在不付出创造性劳动的情况下,即可以理解并实施。

[0141] 以上所述仅为本申请的较佳实施例而已,并不用以限制本申请,凡在本申请的精神和原则之内,所做的任何修改、等同替换、改进等,均应包含在本申请保护的范围之内。

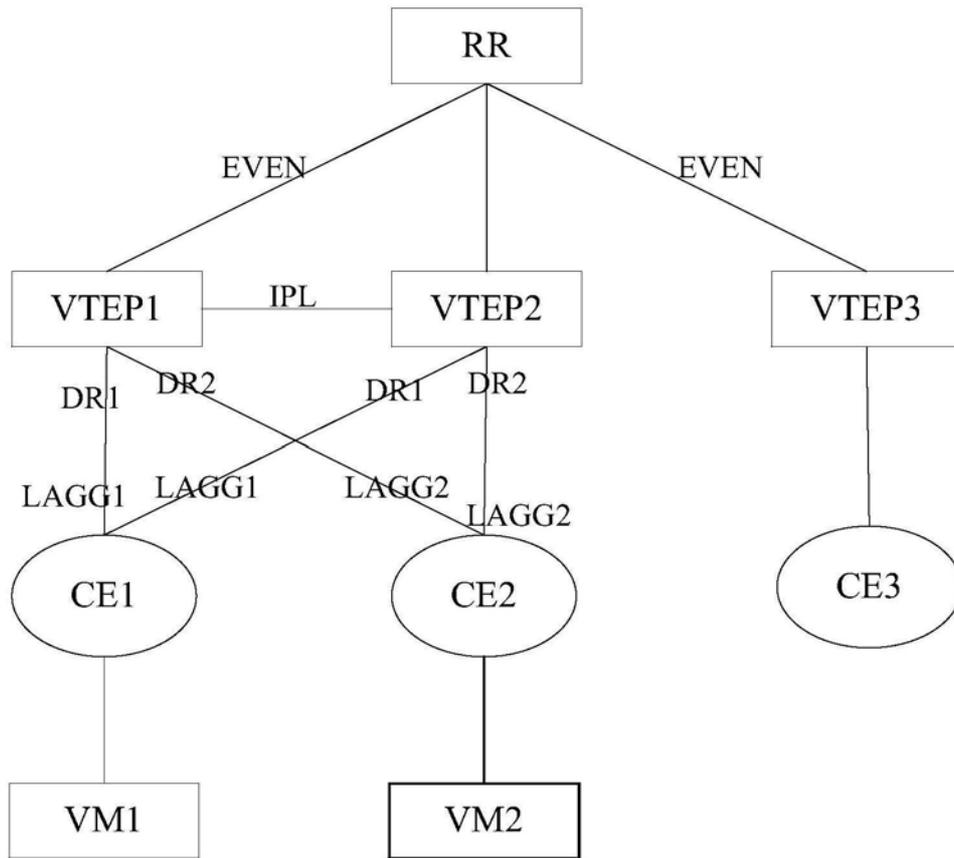


图1

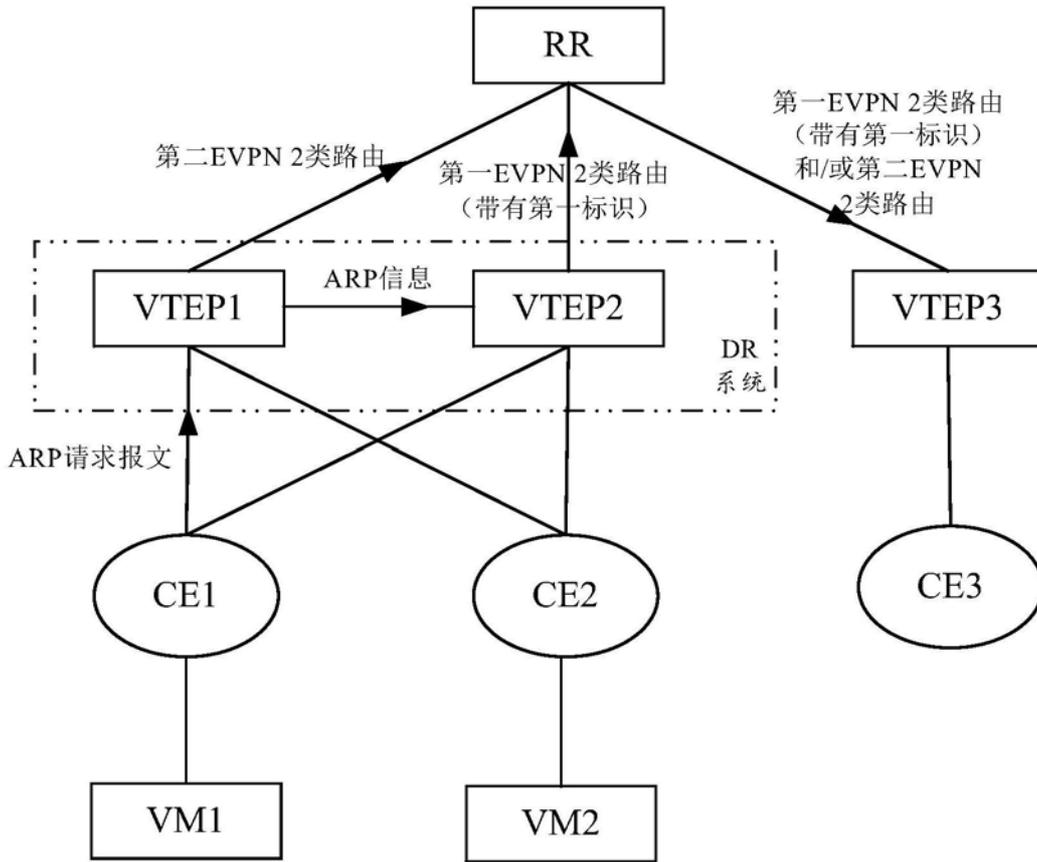


图2

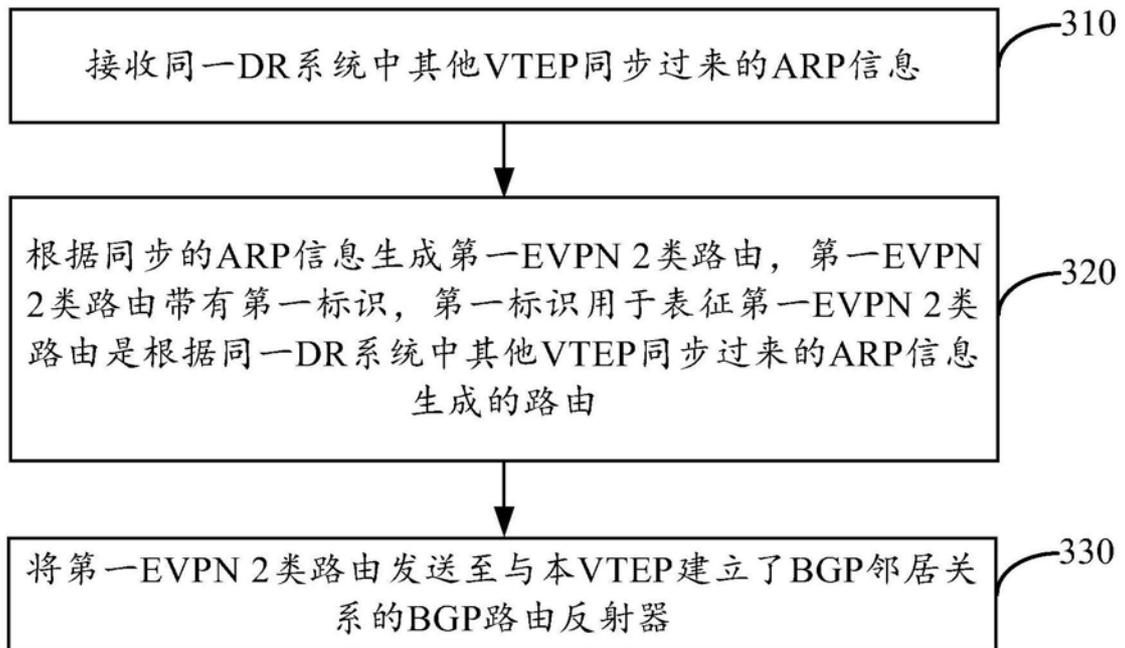


图3

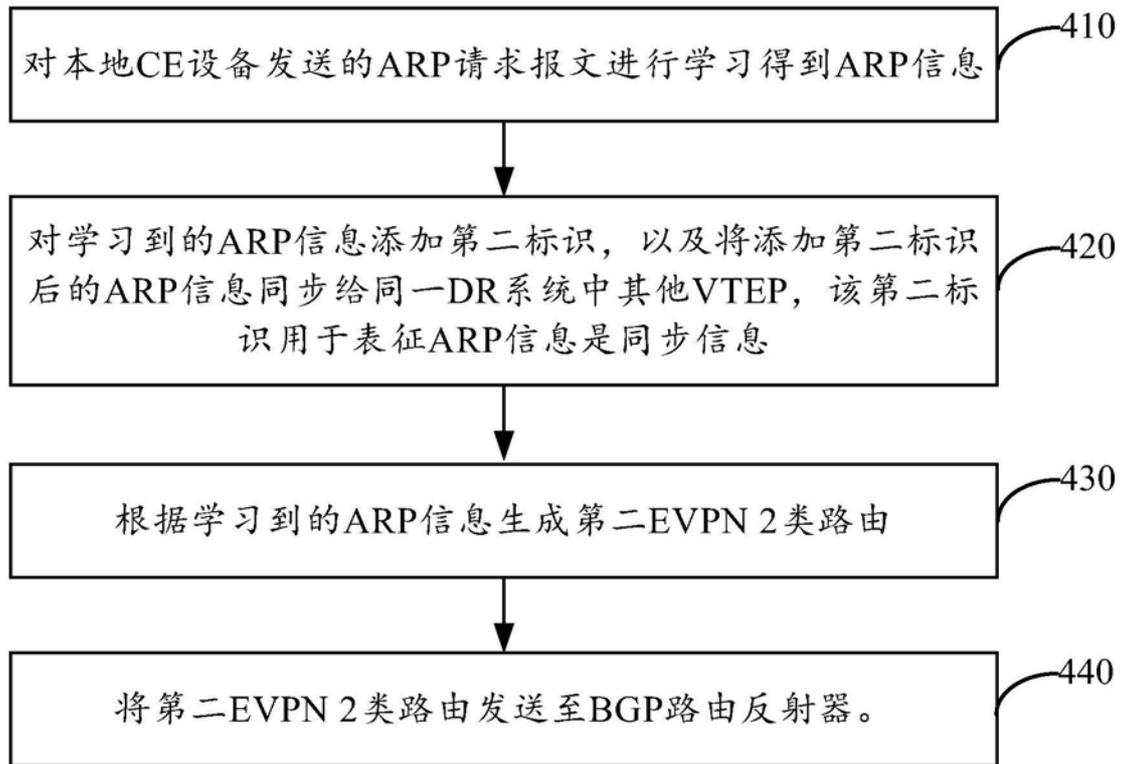


图4

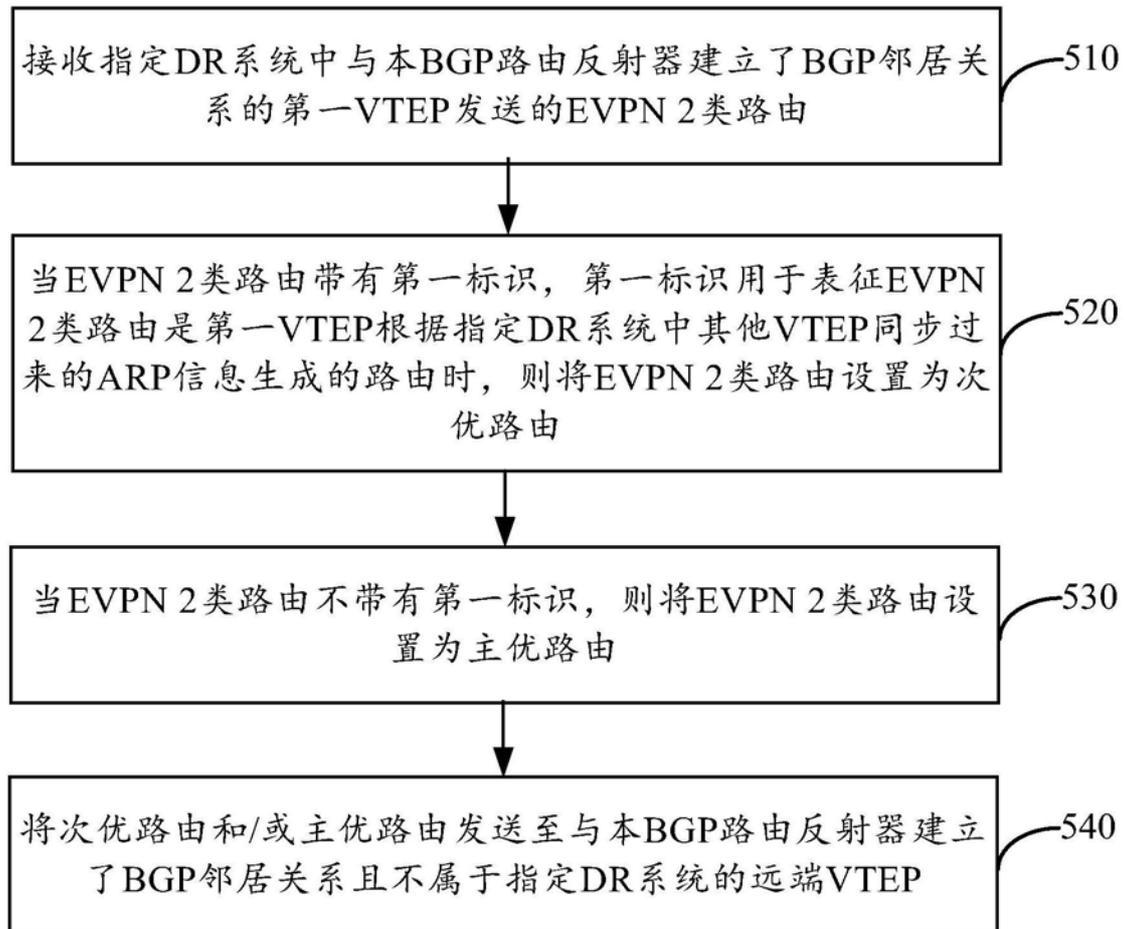


图5

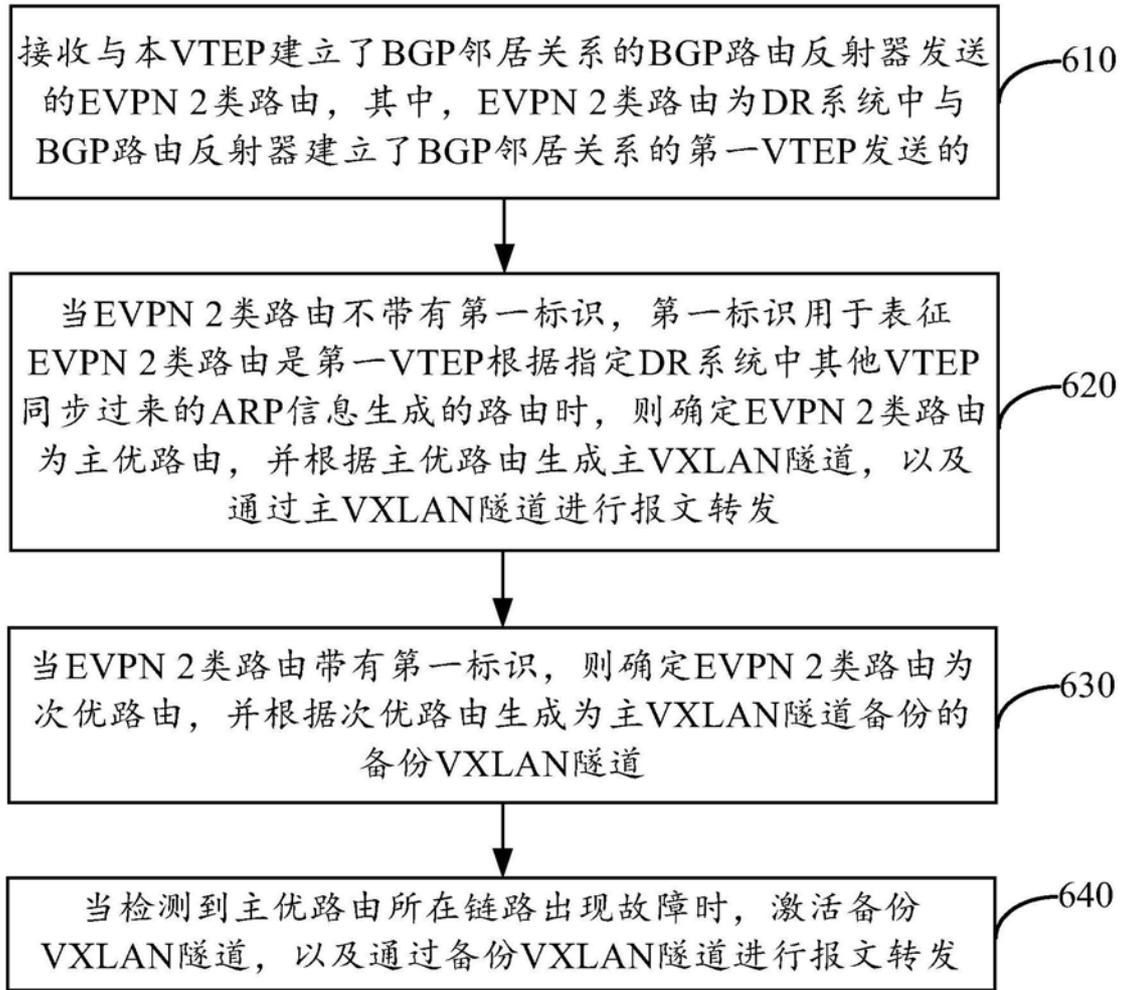


图6

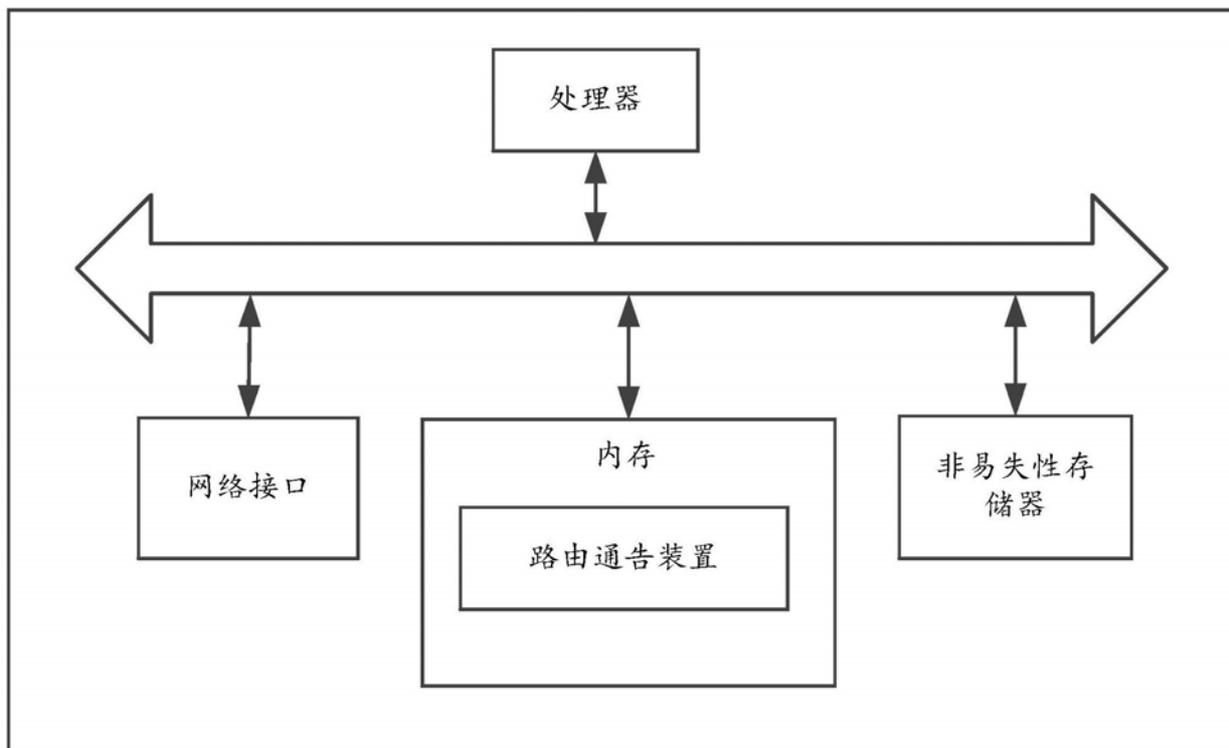


图7

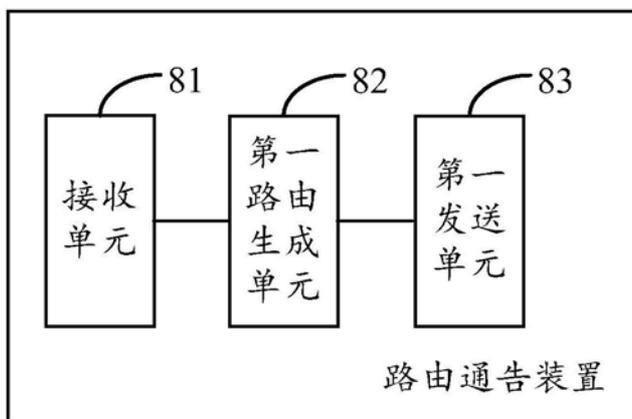


图8

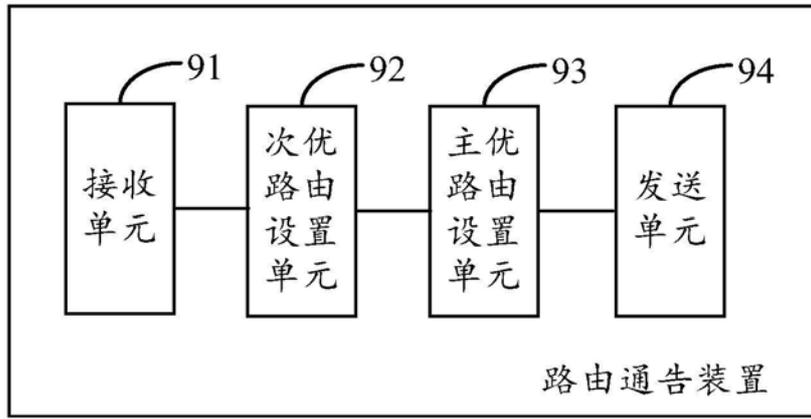


图9

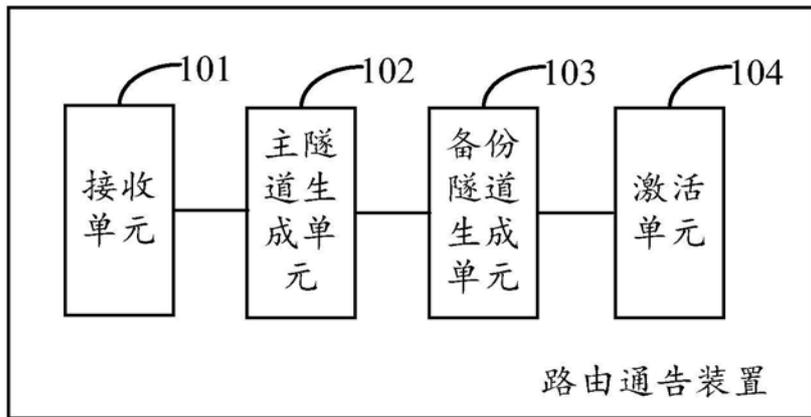


图10