



(12) 发明专利

(10) 授权公告号 CN 111786882 B

(45) 授权公告日 2022.04.22

(21) 申请号 202010613741.2

(22) 申请日 2020.06.30

(65) 同一申请的已公布的文献号  
申请公布号 CN 111786882 A

(43) 申请公布日 2020.10.16

(73) 专利权人 中国联合网络通信集团有限公司  
地址 100033 北京市西城区金融大街21号

(72) 发明人 张余

(74) 专利代理机构 北京天昊联合知识产权代理有限公司 11112  
代理人 彭瑞欣 冯建基

(51) Int. Cl.  
H04L 45/24 (2022.01)  
H04L 45/745 (2022.01)  
H04L 12/46 (2006.01)

(56) 对比文件

CN 106878047 A, 2017.06.20  
CN 106878047 A, 2017.06.20  
US 2016036728 A1, 2016.02.04  
CN 105743816 A, 2016.07.06

审查员 孟姗

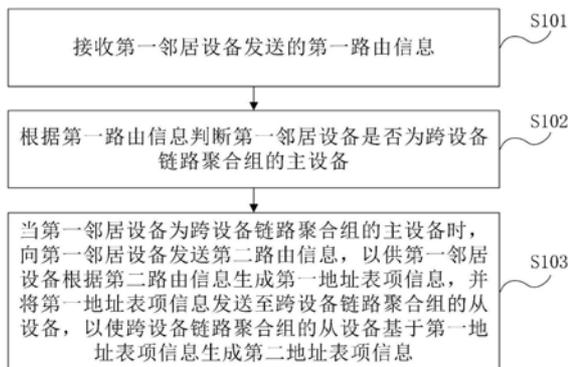
权利要求书2页 说明书12页 附图3页

(54) 发明名称

一种路由处理方法及装置

(57) 摘要

本发明公开了一种路由处理方法及装置,属于通信技术领域。该路由处理方法包括:接收第一邻居设备发送的第一路由信息;根据第一路由信息判断第一邻居设备是否为跨设备链路聚合组的主设备;并在第一邻居设备为跨设备链路聚合组的主设备时,向第一邻居设备发送第二路由信息,以供第一邻居设备根据第二路由信息生成第一地址表项信息,并将第一地址表项信息发送至跨设备链路聚合组的从设备,以减少当前设备向处于跨设备链路聚合组中的邻居设备发送重复路由信息,从而降低跨设备链路聚合组中邻居设备处理路由信息的复杂度。



1. 一种路由处理方法,其特征在於,应用于不处于跨设备链路聚合组的设备,所述方法包括:

接收第一邻居设备发送的第一路由信息;其中,所述第一邻居设备为处于跨设备链路聚合组的邻居设备;

根据所述第一路由信息判断所述第一邻居设备是否为所述跨设备链路聚合组的主设备;

当所述第一邻居设备为所述跨设备链路聚合组的主设备时,向所述第一邻居设备发送第二路由信息,以供所述第一邻居设备根据所述第二路由信息生成第一地址表项信息,并将所述第一地址表项信息发送至所述跨设备链路聚合组的从设备,以使所述跨设备链路聚合组的从设备基于所述第一地址表项信息生成第二地址表项信息。

2. 根据权利要求1所述的路由处理方法,其特征在於,所述接收第一邻居设备发送的第一路由信息之前,还包括:

与所述第一邻居设备建立邻居关系;其中,所述第一邻居设备包括所述跨设备链路聚合组的主设备和所述跨设备链路聚合组的从设备。

3. 根据权利要求1所述的路由处理方法,其特征在於,当判断所述第一邻居设备为所述跨设备链路聚合组的主设备之后,所述向所述第一邻居设备发送第二路由信息之前,还包括:

创建与所述第一邻居设备进行信息交互的虚拟可扩展局域网络隧道。

4. 根据权利要求1所述的路由处理方法,其特征在於,所述第一路由信息包括设备主备标识;

所述根据所述第一路由信息判断所述第一邻居设备是否为所述跨设备链路聚合组的主设备,包括:

根据所述设备主备标识判断所述第一邻居设备是否为所述跨设备链路聚合组的主设备。

5. 一种路由处理方法,其特征在於,应用于处于跨设备链路聚合组的主设备,所述方法包括:

向第二邻居设备发送第一路由信息;其中,所述第二邻居设备为不处于跨设备链路聚合组的邻居设备;

接收所述第二邻居设备返回的第二路由信息;

根据所述第二路由信息生成第一地址表项信息;

将所述第一地址表项信息发送至所述跨设备链路聚合组的从设备,以供所述跨设备链路聚合组的从设备基于所述第一地址表项信息生成对应的第二地址表项信息。

6. 根据权利要求5所述的路由处理方法,其特征在於,所述将所述第一地址表项信息发送至所述跨设备链路聚合组的从设备,包括:

通过直连聚合链路将所述第一地址表项信息发送至所述跨设备链路聚合组的从设备。

7. 根据权利要求5所述的路由处理方法,其特征在於,所述向第二邻居设备发送第一路由信息之前,还包括:

与所述第二邻居设备和所述跨设备链路聚合组的从设备建立邻居关系。

8. 根据权利要求7所述的路由处理方法,其特征在於,所述与所述第二邻居设备和所述

跨设备链路聚合组的从设备建立邻居关系之前,还包括:

通过直连聚合链路向所述跨设备链路聚合组的从设备发送同步报文消息。

9. 一种路由处理装置,其特征在于,所述装置设置于不处于跨设备链路聚合组的设备中,所述装置包括:

第一接收模块,用于接收第一邻居设备发送的第一路由信息;其中,所述第一邻居设备为处于跨设备链路聚合组的邻居设备;

判断模块,用于根据所述第一路由信息判断所述第一邻居设备是否为所述跨设备链路聚合组的主设备;

第一发送模块,用于当所述第一邻居设备为所述跨设备链路聚合组的主设备时,向所述第一邻居设备发送第二路由信息,以供所述第一邻居设备根据所述第二路由信息生成第一地址表项信息,并将所述第一地址表项信息发送至所述跨设备链路聚合组的从设备,以使所述跨设备链路聚合组的从设备基于所述第一地址表项信息生成第二地址表项信息。

10. 一种路由处理装置,其特征在于,所述装置设置于处于跨设备链路聚合组的主设备中,所述装置包括:

第二发送模块,用于向第二邻居设备发送第一路由信息;其中,所述第二邻居设备为不处于跨设备链路聚合组的邻居设备;

第二接收模块,用于接收所述第二邻居设备返回的第二路由信息;

生成模块,用于根据所述第二路由信息生成第一地址表项信息;

第三发送模块,用于将所述第一地址表项信息发送至所述跨设备链路聚合组的从设备,以供所述跨设备链路聚合组的从设备基于所述第一地址表项信息生成对应的第二地址表项信息。

## 一种路由处理方法及装置

### 技术领域

[0001] 本发明涉及通信技术领域,具体涉及一种路由处理方法及装置。

### 背景技术

[0002] 跨设备链路聚合组(Multi-chassis Link Aggregate Group,MLAG)技术可以将多台设备进行跨设备链路聚合,形成跨设备链路聚合组,从而提升链路的可靠性。而且,MLAG中的设备可以与不处于MLAG的设备建立邻居连接。如,使用可扩展虚拟局域网隧道端点(VXLAN Tunnel End Point,VTEP)设备构建MLAG,并与不处于该MLAG的VTEP设备建立邻居连接,且假设MLAG中的设备分别是VTEP1和VTEP2,不处于该MLAG的邻居设备为VTEP3。在处理路由信息时,首先由VTEP3分别向VTEP1和VTEP2发送Type2路由信息;VTEP1接收到VTEP3发送的Type2路由信息后,生成与VTEP1对应的本地MAC(Media Access Control Address,媒体访问控制地址)地址表项,并把MAC地址表项信息通过peer-link链路(即直连聚合链路)同步给VTEP2;同样的,VTEP2接收到VTEP3发送的Type2路由信息后,生成与VTEP2对应的本地MAC地址表项,并把MAC地址表项信息通过peer-link链路同步给VTEP1。这些在peer-link链路上同步的MAC地址表项信息存在重复内容,且占用了VTEP1和VTEP2之间的链路资源,同时也消耗VTEP1和VTEP2的处理能力。

[0003] 因此,如何减少当前设备向处于跨设备链路聚合组中的邻居设备发送重复路由信息,以降低跨设备链路聚合组中邻居设备处理路由信息的复杂度成为本领域亟待解决的问题。

### 发明内容

[0004] 为此,本发明提供一种路由处理方法及装置,以解决当前设备向跨设备链路聚合组中的邻居设备发送重复路由信息,导致跨设备链路聚合组中的邻居设备处理路由信息复杂度较高的问题。

[0005] 为了实现上述目的,本发明第一方面提供一种路由处理方法,包括:

[0006] 接收第一邻居设备发送的第一路由信息;其中,所述第一邻居设备为处于跨设备链路聚合组的邻居设备;

[0007] 根据所述第一路由信息判断所述第一邻居设备是否为所述跨设备链路聚合组的主设备;

[0008] 当所述第一邻居设备为所述跨设备链路聚合组的主设备时,向所述第一邻居设备发送第二路由信息,以供所述第一邻居设备根据所述第二路由信息生成第一地址表项信息,并将所述第一地址表项信息发送至所述跨设备链路聚合组的从设备,以使所述跨设备链路聚合组的从设备基于所述第一地址表项信息生成第二地址表项信息。

[0009] 进一步地,所述接收第一邻居设备发送的第一路由信息之前,还包括:

[0010] 与所述第一邻居设备建立邻居关系;其中,所述第一邻居设备包括所述跨设备链路聚合组的主设备和所述跨设备链路聚合组的从设备。

[0011] 进一步地,当判断所述第一邻居设备为所述跨设备链路聚合组的主设备之后,所述向所述第一邻居设备发送第二路由信息之前,还包括:

[0012] 创建与所述第一邻居设备进行信息交互的虚拟可扩展局域网络隧道。

[0013] 进一步地,所述第一路由信息包括设备主备标识;

[0014] 所述根据所述第一路由信息判断所述第一邻居设备是否为所述跨设备链路聚合组的主设备,包括:

[0015] 根据所述设备主备标识判断所述第一邻居设备是否为所述跨设备链路聚合组的主设备。

[0016] 为了实现上述目的,本发明第二方面提供一种路由处理方法,包括:

[0017] 向第二邻居设备发送第一路由信息;其中,所述第二邻居设备为不处于跨设备链路聚合组的邻居设备;

[0018] 接收所述第二邻居设备返回的第二路由信息;

[0019] 根据所述第二路由信息生成第一地址表项信息;

[0020] 将所述第一地址表项信息发送至所述跨设备链路聚合组的从设备,以供所述跨设备链路聚合组的从设备基于所述第一地址表项信息生成对应的第二地址表项信息。

[0021] 进一步地,所述将所述第一地址表项信息发送至所述跨设备链路聚合组的从设备,包括:

[0022] 通过直连聚合链路将所述第一地址表项信息发送至所述跨设备链路聚合组的从设备。

[0023] 进一步地,所述向第二邻居设备发送第一路由信息之前,还包括:

[0024] 与所述第二邻居设备和所述跨设备链路聚合组的从设备建立邻居关系。

[0025] 进一步地,所述与所述第二邻居设备和所述跨设备链路聚合组的从设备建立邻居关系之前,还包括:

[0026] 通过直连聚合链路向所述跨设备链路聚合组的从设备发送同步报文消息。

[0027] 为了实现上述目的,本发明第三方面提供一种路由处理装置,包括:

[0028] 第一接收模块,用于接收第一邻居设备发送的第一路由信息;其中,所述第一邻居设备为处于跨设备链路聚合组的邻居设备;

[0029] 判断模块,用于根据所述第一路由信息判断所述第一邻居设备是否为所述跨设备链路聚合组的主设备;

[0030] 第一发送模块,用于当所述第一邻居设备为所述跨设备链路聚合组的主设备时,向所述第一邻居设备发送第二路由信息,以供所述第一邻居设备根据所述第二路由信息生成第一地址表项信息,并将所述第一地址表项信息发送至所述跨设备链路聚合组的从设备,以使所述跨设备链路聚合组的从设备基于所述第一地址表项信息生成第二地址表项信息。

[0031] 为了实现上述目的,本发明第四方面提供一种路由处理装置,包括:

[0032] 第二发送模块,用于向第二邻居设备发送第一路由信息;其中,所述第二邻居设备为不处于跨设备链路聚合组的邻居设备;

[0033] 第二接收模块,用于接收所述第二邻居设备返回的第二路由信息;

[0034] 生成模块,用于根据所述第二路由信息生成第一地址表项信息;

[0035] 第三发送模块,用于将所述第一地址表项信息发送至所述跨设备链路聚合组的从设备,以供所述跨设备链路聚合组的从设备基于所述第一地址表项信息生成对应的第二地址表项信息。

[0036] 本发明具有如下优点:

[0037] 本发明提供的路由处理方法,接收第一邻居设备发送的第一路由信息;根据第一路由信息判断第一邻居设备是否为跨设备链路聚合组的主设备;并在第一邻居设备为跨设备链路聚合组的主设备时,向第一邻居设备发送第二路由信息,以供第一邻居设备根据第二路由信息生成第一地址表项信息,并将第一地址表项信息发送至跨设备链路聚合组的从设备,以减少当前设备向处于跨设备链路聚合组中的邻居设备发送重复路由信息,从而降低跨设备链路聚合组中邻居设备处理路由信息的复杂度。

## 附图说明

[0038] 附图是用来提供对本发明的进一步理解,并且构成说明书的一部分,与下面的具体实施方式一起用于解释本发明,但并不构成对本发明的限制。

[0039] 图1为本发明第一实施例提供的一种路由处理方法的流程图;

[0040] 图2为本发明第二实施例提供的一种路由处理方法的流程图;

[0041] 图3为本发明第三实施例提供的一种路由处理方法的流程图;

[0042] 图4为本发明第四实施例提供的一种路由处理方法的流程图;

[0043] 图5为本发明第五实施例提供的一种路由处理装置的原理框图;

[0044] 图6为本发明第六实施例提供的一种路由处理装置的原理框图。

[0045] 在附图中:

[0046] 501:第一接收模块                      502:判断模块

[0047] 503:第一发送模块

[0048] 601:第二发送模块                      602:第二接收模块

[0049] 603:生成模块                              604:第三发送模块

## 具体实施方式

[0050] 以下结合附图对本发明的具体实施方式进行详细说明。应当理解的是,此处所描述的具体实施方式仅用于说明和解释本发明,并不用于限制本发明。

[0051] 本发明提供的路由处理方法及装置,考虑到MLAG中的VTEP1和VTEP2与不处于MLAG的VTEP3建立两两邻居关系以后,在处理路由信息时,首先由VTEP3分别向VTEP1和VTEP2发送Type2路由信息。VTEP1接收到VTEP3发送的Type2路由信息后,生成与VTEP1对应的本地MAC地址表项,并把MAC地址表项信息通过peer-link链路同步给VTEP2;同样的,VTEP2接收到VTEP3发送的Type2路由信息后,生成与VTEP2对应的本地MAC地址表项,并把MAC地址表项信息通过peer-link链路同步给VTEP1。这些在peer-link链路上同步的MAC地址表项信息存在重复内容,且占用了VTEP1和VTEP2之间的链路资源,同时也消耗VTEP1和VTEP2的处理能力。有鉴于此,提出一种新的路由信息处理方法,可以避免重复路由信息占用peer-link链路资源,同时还可降低VTEP1和VTEP2处理路由信息的复杂度。

[0052] 图1是本发明第一实施例提供的一种路由处理方法的流程图,应用于不处于跨设

备链路聚合组的设备。如图1所示,该路由处理方法可包括如下步骤:

[0053] 步骤S101,接收第一邻居设备发送的第一路由信息。

[0054] 在本实施例中,当前设备(即执行主体)为不处于跨设备链路聚合组的设备,在当前设备与处于跨设备链路聚合组中的设备建立两两邻居关系之后,当前设备与跨链路聚合组中的设备即互为邻居设备。对于当前设备来说,它的邻居设备就是跨设备链路聚合组中的设备。进一步地,将处于跨设备链路聚合组中的邻居设备定义为第一邻居设备。而且,跨设备链路聚合组中的设备包括主设备和从设备,因此,当前设备的第一邻居设备就包括跨设备链路聚合组的主设备和跨设备链路聚合组的从设备。如,假设有三台设备,分别为equ1、equ2和equ3;其中,equ1和equ2处于跨设备链路聚合组,且equ1为跨设备链路聚合组的主设备,equ2为跨设备链路聚合组的从设备,equ3为不处于跨设备链路聚合组的设备。equ1、equ2和equ3之间分别建立邻居关系。equ3即为当前设备(执行主体),equ3的邻居设备即为equ1和equ2,而且,由于equ1和equ2均为处于跨设备链路聚合组中的设备,因此,equ1和equ2是equ3的第一邻居设备。

[0055] 在当前设备与跨设备链路聚合组中的设备建立邻居关系,跨设备链路聚合组中的设备成为当前设备的第一邻居设备之后,第一邻居设备为了进行设备发现以及传输通道的建立,会向当前设备发送第一路由信息,以供当前设备基于第一路由信息实现设备发现与传输通道的建立。

[0056] 上述内容中提到的跨设备链路聚合组为基于跨设备链路聚合组技术构建的聚合组。跨设备链路聚合组技术是业界为提高网络可用性而提出的一种实现跨设备链路聚合的机制。MLAG技术将一台设备与另外一台或多台设备进行跨设备链路聚合,从而将不同设备的物理链路组成逻辑上的链路聚合组,进而将链路可靠性从单板级提高到设备级,并组成对应的双活系统(或称双归系统)。一般情况下,跨设备链路聚合组中包括一台主设备和一台或多台从设备。在正常情况下,跨设备链路聚合组主设备和跨设备链路聚合组备设备同时进行业务流量的转发;当跨设备链路聚合组主设备故障时,会从跨设备链路聚合组的从设备中选取一台作为新的主设备继续负责转发业务流量。

[0057] 在一些实施方式中,构建跨设备链路聚合组的设备为可扩展虚拟局域网隧道端点(VXLAN Tunnel End Point,VTEP)设备。VTEP设备为可扩展虚拟局域网(Virtual Extensible Local Area Network,VXLAN)中的端点设备。VXLAN技术是基于IP(Internet Protocol,国际互连协议)网络的、将以太网报文封装在UDP(User Datagram Protocol,用户报文协议)传输层上进行转发的二层VPN(Virtual Private Network,虚拟专用网络)技术。VXLAN主要应用于数据中心网络,可以基于已有的服务提供商或企业IP网络,为分散的物理站点提供二层互联,并能够为不同的租户提供业务隔离。具体地,VXLAN技术将已有的三层物理网络作为Underlay网络(即下层网络),在其上构建出虚拟的二层网络,即Overlay网络(即上层网络)。Overlay网络通过封装技术,并利用Underlay网络提供的三层转发路径,实现租户二层报文跨越三层网络在不同站点间传递。对于租户来说,Underlay网络是透明的,同一租户的不同站点就像工作在一个局域网中。

[0058] VXLAN技术具有支持大量租户和易于维护等特点,具体地:

[0059] 1、支持大量租户:VXLAN技术使用的标识符具有24位,最多可支持 $2^{24}$ 次方(16777216)个VXLAN,使得支持的租户数量得以大规模增加,解决了传统二层网络VLAN资源

不足的问题。

[0060] 2、易于维护:VXLAN技术基于IP网络组建大二层网络,使得网络部署和维护更加容易,并且可以充分地利用现有的IP网络技术,例如可以利用等价路由进行负载分担等。而且,只有IP核心网络的边缘设备需要进行VXLAN处理,网络中间设备只需根据IP头转发报文,从而可以降低网络部署的难度和费用。

[0061] VXLAN的典型网络模型包括虚拟机、VTEP设备、VXLAN隧道、核心设备、虚拟交换实例和接入电路等,具体地:

[0062] 1、虚拟机(Virtual Machine,VM):在一台服务器上可以创建多台虚拟机,不同的虚拟机可以属于不同的VXLAN。属于相同VXLAN的虚拟机处于同一个逻辑二层网络,彼此之间二层互通;属于不同VXLAN的虚拟机之间二层隔离。VXLAN通过VXLANID来标识,VXLANID又称VNI(VXLAN Network Identifier,VXLAN网络标识符),其长度为24比特。

[0063] 2、VTEP设备:VXLAN的边缘设备。VXLAN的相关处理都在VTEP设备上进行,例如识别以太网数据帧所属的VXLAN、基于VXLAN对数据帧进行二层转发、封装/解封装报文等。VTEP可以是一台独立的物理设备,也可以是虚拟机所在的服务器。

[0064] 3、VXLAN隧道:两个VTEP之间点到点的逻辑隧道。VTEP设备为数据帧封装VXLAN头、UDP头和IP头,通过VXLAN隧道将封装后的报文转发给远端VTEP,远端VTEP对其进行解封装。

[0065] 4、核心设备:IP核心网络中的设备。核心设备不参与VXLAN处理,仅需要根据封装后报文的目的IP地址对报文进行三层转发。

[0066] 5、虚拟交换实例(Virtual Switch Instance,VSI):VTEP设备为一个VXLAN提供二层交换服务的虚拟交换实例。VSI可以看做是VTEP设备中的一台基于VXLAN进行二层转发的虚拟交换机,它具有传统以太网交换机的所有功能,包括源MAC地址学习、MAC地址老化、泛洪等。VSI与VXLAN一一对应。

[0067] 6、接入电路(Attachment Circuit,AC):VTEP连接本地站点的物理电路或虚拟电路。在VTEP中,与VSI关联的三层接口或以太网服务实例(service instance,SI)称为AC。其中,以太网服务实例在二层以太网接口上创建,它定义了一系列匹配规则,用来匹配从该二层以太网接口上接收到的数据帧。

[0068] 在一个实施方式中,VTEP1、VTEP2和VTEP3均为可扩展虚拟局域网隧道端点,其中,VTEP1与VTEP2基于跨设备链路聚合组机制组成跨设备链路聚合组,且VTEP1为该跨设备链路聚合组的主设备,VTEP2为该跨设备链路聚合组的从设备,VTEP1与VTEP2构成双活系统。VTEP1与VTEP2之间使用不同的地址作为BGP对等体地址,相互建立边界网关协议(Border Gateway Protocol,BGP)邻居关系,并分别与不处于该跨设备链路聚合组的VTEP3也建立BGP邻居关系。

[0069] 在本实施方式中,当前设备为VTEP3,VTEP3对应的第一邻居设备包括VTEP1和VTEP2。在VTEP1、VTEP2和VTEP3建立两两BGP邻居关系之后,VTEP1和VTEP2分别向VTEP3发送第一路由信息,VTEP3接收其第一邻居设备VTEP1和VTEP2分别发送的第一路由信息(即VTEP3接收的第一路由信息的数量为两条,一条来自于VTEP1,另一条来自于VTEP2)。其中,第一路由信息为Type3路由。Type3路由即inclusive multicast路由(即包含模式组播路由),该类型路由主要用于完成VTEP地址的自动发现和VXLAN隧道的动态建立,以及通告二层VXLAN网络标识(VXLAN Network Identifier,VNI)信息和VTEP IP地址信息。

[0070] 步骤S102,根据第一路由信息判断第一邻居设备是否为跨设备链路聚合组的主设备。

[0071] 第一路由信息中可以携带特定标识,通过该特定标识来判断第一邻居设备是否为跨设备链路聚合组的主设备。如,在第一路由信息中设置设备主备标识,通过设备主备标识来判断第一邻居设备是否为跨设备链路聚合组的主设备。

[0072] 在一个实施方式中,第一路由信息为Type3路由,在Type3路由的运营商组播服务虚拟出接口(Provider Multicast Service Interface,PMSI)属性集中新增两个属性,分别是跨设备链路聚合组编号和设备主备标识。当前设备接收Type3路由后,根据跨设备链路聚合组编号和设备主备标识对应的属性值即可判断发送Type3路由的第一邻居设备是否为跨设备链路聚合组的主设备。

[0073] 步骤S103,当第一邻居设备为跨设备链路聚合组的主设备时,向第一邻居设备发送第二路由信息,以供第一邻居设备根据第二路由信息生成第一地址表项信息,并将第一地址表项信息发送至跨设备链路聚合组的从设备,以使跨设备链路聚合组的从设备基于第一地址表项信息生成第二地址表项信息。

[0074] 在传统的邻居设备路由处理方法中,当前设备为不处于跨设备链路聚合组的设备,如果当前设备与处于跨设备链路聚合组的设备建立两两邻居关系之后,当前设备与跨设备链路聚合组中的设备即互为邻居设备,而且当前设备的邻居设备均为第一邻居设备。当前设备接收多个第一邻居设备发送的第一路由信息之后,需要分别向这多个第一邻居设备返回第二路由信息。第一邻居设备接收第二路由信息之后,均会生成自己的本地地址表项信息,而且,第一邻居设备还会向自己的邻居设备发送自己生成的地址表项信息,即跨设备链路聚合组中的主设备与从设备之间会互相发送本地生成的地址表项信息,然而,这些地址表项信息本身存在重复的内容。而且,发送这些地址表项信息需要占用peer-link链路资源,还会增加跨设备链路聚合组中的主设备和从设备处理路由信息的复杂度。因此,提出本申请的路由处理方法,当前设备只向第一邻居设备中的跨设备链路的主设备发送第二路由信息,而不再向第一邻居设备中的跨设备链路的从设备发送第二路由信息,当跨设备链路的主设备生成其本地的第一地址表项信息之后,将第一地址表项信息发送至跨设备链路的从设备,跨设备链路的从设备根据第一地址表项信息生成其本地的第二地址表项信息,从而可以避免邻居设备向处于跨设备链路聚合组中的邻居设备发送重复路由信息,进而可以降低跨设备链路聚合组中邻居设备处理路由信息的复杂度。

[0075] 在一个实施方式中,VTEP1、VTEP2和VTEP3建立两两BGP邻居关系,其中,VTEP1与VTEP2分别为跨设备链路聚合组的主设备和跨设备链路聚合组的从设备,VTEP3为不处于该跨设备链路聚合组的邻居设备。在本实施方式中,当前设备为VTEP3,VTEP3对应的第一邻居设备包括VTEP1和VTEP2。

[0076] 当VTEP3判断VTEP1为跨设备链路聚合组的主设备时(相当于VTEP3也获知VTEP2为该跨设备链路聚合组的从设备),VTEP3向VTEP1发送第二路由信息,而不向VTEP2发送第二路由信息。其中,第二路由信息为type2路由,也即MAC/IP路由,该类型路由主要用于通告mac地址和主机IP地址。

[0077] VTEP1收到type2路由后,基于type2路由生成VTEP1的本地第一地址表项信息,并通过VTEP1与VTEP2之间的peer-link链路将第一地址表项信息发送至VTEP2。VTEP2收到第

一地址表项信息之后,根据第一地址表项信息生成VTEP2本地的第二地址表项信息。其中,peer-link链路为VTEP1与VTEP2之间的直连聚合链路,一般情况下,peer-link链路是一条二层链路,用于VTEP1与VTEP2之间协商报文的交互以及部分流量的传输。

[0078] 图2是本发明第二实施例提供的一种路由处理方法的流程图,应用于不处于跨设备链路聚合组的设备,与本发明第一实施例基本相同,区别之处在于:当判断第一邻居设备为跨设备链路聚合组的主设备之后,创建与第一邻居设备进行信息交互的虚拟可扩展局域网络隧道。如图2所示,该路由处理方法可包括如下步骤:

[0079] 步骤S201,接收第一邻居设备发送的第一路由信息。

[0080] 本实施例中的步骤S201与本发明第一实施例中步骤S101的内容相同,在此不再赘述。

[0081] 步骤S202,根据第一路由信息判断第一邻居设备是否为跨设备链路聚合组的主设备。

[0082] 本实施例中的步骤S202与本发明第一实施例中步骤S102的内容相同,在此不再赘述。

[0083] 步骤S203,当第一邻居设备为跨设备链路聚合组的主设备时,创建与第一邻居设备进行信息交互的虚拟可扩展局域网络隧道。

[0084] 虚拟可扩展局域网络隧道是VTEP设备之间进行数据传输的虚拟网络通道。具体地,当前VTEP设备将VXLAN报文进行封装后,通过虚拟可扩展局域网络隧道将其发送至对端VTEP设备,对端VTEP设备接收VXLAN报文后,根据被封装的MAC地址进行VXLAN报文转发或者对VXLAN报文进行解封装。VXLAN隧道包括静态单播隧道、动态单播隧道和组播隧道等类型,在实际使用中可根据需要择选某一类型VXLAN隧道进行创建。

[0085] 在本实施例中,当前设备为不处于跨设备链路聚合组的设备,其第一邻居设备为处于跨设备链路聚合组的邻居设备,具体地,第一邻居设备包括跨设备链路聚合组的主设备和跨设备链路聚合组的从设备。当前设备判断其第一邻居设备为跨设备链路聚合组的主设备时,创建与第一邻居设备进行信息交互的虚拟可扩展局域网络隧道。

[0086] 在一个实施方式中,VTEP1、VTEP2和VTEP3建立两两BGP邻居关系,其中,VTEP1与VTEP2分别为跨设备链路聚合组的主设备和跨设备链路聚合组的从设备,VTEP3为不处于该跨设备链路聚合组的邻居设备。

[0087] 在本实施方式中,当前设备为VTEP3,VTEP3对应的第一邻居设备包括VTEP1和VTEP2。当VTEP3判断出VTEP1为跨设备链路聚合组的主设备时,VTEP3和VTEP1分别创建对应的VXLAN隧道端口,并基于VXLAN隧道端口创建VXLAN隧道,以通过VXLAN隧道进行信息交互。

[0088] 步骤S204,向第一邻居设备发送第二路由信息,以供第一邻居设备根据第二路由信息生成第一地址表项信息,并将第一地址表项信息发送至跨设备链路聚合组的从设备,以使跨设备链路聚合组的从设备基于第一地址表项信息生成第二地址表项信息。

[0089] 本实施例中的步骤S204与本发明第一实施例中步骤S103的内容相同,在此不再赘述。

[0090] 图3是本发明第三实施例提供的一种路由处理方法的流程图,应用于处于跨设备链路聚合组的主设备。如图3所示,该路由处理方法可包括如下步骤:

[0091] 步骤S301,向第二邻居设备发送第一路由信息。

[0092] 在本实施例中,当前设备(即执行主体)为跨设备链路聚合组中的主设备,当前设备与跨设备链路聚合组中的从设备以及不处于跨设备链路聚合组的设备建立邻居连接之后,当前设备与跨设备链路聚合组中的从设备以及不处于跨设备链路聚合组的设备互为邻居设备。对于当前设备而言,其邻居设备包括两类,一类是处于该跨设备链路聚合组的从设备,另一类是不处于跨设备链路聚合组的邻居设备。进一步地,将不处于跨设备链路聚合组的邻居设备定义为第二邻居设备。

[0093] 在一个实施方式中,VTEP1、VTEP2和VTEP3均为可扩展虚拟局域网隧道端点,其中,VTEP1与VTEP2跨设备链路聚合组中的设备,且VTEP1为该跨设备链路聚合组的主设备,VTEP2为该跨设备链路聚合组的从设备。VTEP1、VTEP2和VTEP3之间建立两两BGP邻居关系。

[0094] 在本实施方式中,当前设备为VTEP1,VTEP1对应的第二邻居设备为VTEP3。VTEP1、VTEP2和VTEP3建立两两邻居关系之后,VTEP1向VTEP3发送第一路由信息,其中,第一路由信息为Type3路由,而且在Type3路由中新增两个属性,分别是跨设备链路聚合组编号和设备主备标识,以供VTEP3根据Type3路由中的跨设备链路聚合组编号和设备主备标识确定其第一邻居设备中的跨设备链路聚合组的主设备。

[0095] 需要说明的是,VTEP2也会向VTEP3发送第一路由信息。

[0096] 步骤S302,接收第二邻居设备返回的第二路由信息。

[0097] 第二邻居设备接收其第一邻居设备(包括当前设备以及当前设备所处跨设备链路聚合组的从设备)发送的第一路由信息之后,根据第一路由信息确定第一邻居设备中的跨设备链路聚合组的主设备,并且只向该跨设备链路聚合组的主设备返回第二路由信息。

[0098] 在一个实施方式中,VTEP1与VTEP2分别为跨设备链路聚合组的主设备和跨设备链路聚合组的从设备,VTEP1、VTEP2与不处于该跨设备链路聚合组的VTEP3建立两两邻居关系。

[0099] 在本实施方式中,当前设备为VTEP1,VTEP1对应的第二邻居设备为VTEP3(可以理解的是,VTEP1的邻居设备还包括VTEP2,只不过VTEP2不属于本实施例中定义的第二邻居设备类型)。

[0100] VTEP3分别收到VTEP1和VTEP2发送的第一路由信息之后,根据第一路由判断VTEP1和VTEP2是否为跨设备链路聚合组的主设备。当VTEP3判断出VTEP1为跨设备链路聚合组的主设备时,VTEP3向VTEP1发送第二路由信息。其中,第二路由信息为type2路由,也即MAC/IP路由,该类型路由主要用于通告mac地址和主机IP地址。

[0101] 需要说明的是,虽然VTEP1和VTEP2均会向VTEP3发送第一路由信息。但是,VTEP3只向VTEP1返回第二路由信息,从而可以避免后续VTEP1与VTEP2之间发送重复的地址表项信息,同时还可降低VTEP1与VTEP2的处理路由信息的复杂程度。

[0102] 步骤S303,根据第二路由信息生成第一地址表项信息。

[0103] 其中,第一地址表项信息为当前设备(即跨设备链路聚合组的主设备)对应的地址表项信息。地址表项信息为设备在本地MAC地址表中生成的地址表项,它是识别网络节点的标识。

[0104] 在一个实施方式中,当前设备接收第二邻居设备发送的第二路由信息,其中,第二路由信息为type2路由。当前设备根据type2路由在本地MAC地址表中生成对应的地址表项。

[0105] 步骤S304,将第一地址表项信息发送至跨设备链路聚合组的从设备,以供跨设备

链路聚合组的从设备基于第一地址表项信息生成对应的第二地址表项信息。

[0106] 在一个实施方式中,当前设备通过peer-link链路将第一地址表项信息发送至跨设备链路聚合组的从设备。跨设备链路聚合组的从设备接收第一地址表项信息后,根据第一地址表项信息在本地MAC地址表中生成对应的第二地址表项信息。

[0107] 图4是本发明第四实施例提供的一种路由处理方法的流程图,应用于处于跨设备链路聚合组的主设备,与本发明第三实施例基本相同,区别之处在于:在向第二邻居设备发送第一路由信息之前,通过直连聚合链路向跨设备链路聚合组的从设备发送同步报文消息。如图4所示,该路由处理方法可包括如下步骤:

[0108] 步骤S401,通过直连聚合链路向跨设备链路聚合组的从设备发送同步报文消息。

[0109] 跨设备链路聚合组的主设备和跨设备链路聚合组的从设备在与不处于该跨设备链路聚合组的设备建立邻居关系之前,跨设备链路聚合组的主设备和跨设备链路聚合组的从设备基于跨设备链路聚合方式构建跨设备链路聚合组,形成双活系统,并通过peer-link链路发送同步报文消息。

[0110] 在一个实施方式中,当跨设备链路聚合方式下的两台设备VTEP1和VTEP2完成配置后,VTEP1和VTEP2通过peer-link链路向对端设备(VTEP1的对端设备为VTEP2,VTEP2的对端设备为VTEP1)发送Hello报文。相应的,VTEP1和VTEP2收到对端设备发送的Hello报文,并判断Hello报文中携带的跨设备链路聚合组编号是否和本端的跨设备链路聚合组编号相同。如果对端与本端的跨设备链路聚合组编号相同,则VTEP1和VTEP2跨设备链路聚合组配对成功。

[0111] 配对成功之后,VTEP1和VTEP2会向对端设备发送跨设备链路聚合组的设备信息报文。同样的,VTEP1和VTEP2接收对端设备发送的设备信息报文,并根据设备信息报文中携带的跨设备链路聚合组优先级以及系统MAC地址确定出跨设备链路聚合组的主备状态。

[0112] 在跨设备链路聚合组协商出主备状态后,VTEP1和VTEP2会通过peer-link链路发送接口信息报文,接口信息报文中携带了成员接口的配置信息。VTEP1和VTEP2根据接口信息报文进行接口信息同步,并在接口信息同步之后,进一步协商成员接口的主备状态。

[0113] 在协商出接口主备状态后,VTEP1和VTEP2通过双主检测链路按照预设第一周期(如,15s)发送双主检测报文。一旦VTEP1或VTEP2感知peer-link链路存在故障,就会按照预设第二周期(如,100ms)发送三个双主检测报文,从而加速检测。

[0114] 当VTEP1和VTEP2均能够收到对端发送的双主检测报文时,双活系统即开始正常工作。正常工作之后,VTEP1和VTEP2通过peer-link链路发送同步报文,并根据同步报文实时同步对端的信息。其中,同步报文中包括MAC表项、ARP表项、发送成员端口状态以及同步STP、VRRP协议报文等内容。

[0115] 步骤S402,向第二邻居设备发送第一路由信息。

[0116] 本实施例中的步骤S402与本发明第三实施例中步骤S301的内容相同,在此不再赘述。

[0117] 步骤S403,接收第二邻居设备返回的第二路由信息。

[0118] 本实施例中的步骤S403与本发明第三实施例中步骤S302的内容相同,在此不再赘述。

[0119] 步骤S404,根据第二路由信息生成第一地址表项信息。

[0120] 本实施例中的步骤S404与本发明第三实施例中步骤S303的内容相同,在此不再赘述。

[0121] 步骤S405,将第一地址表项信息发送至跨设备链路聚合组的从设备,以供跨设备链路聚合组的从设备基于第一地址表项信息生成对应的第二地址表项信息。

[0122] 本实施例中的步骤S405与本发明第三实施例中步骤S304的内容相同,在此不再赘述。

[0123] 上面各种方法的步骤划分,只是为了描述清楚,实现时可以合并为一个步骤或者对某些步骤进行拆分,分解为多个步骤,只要包括相同的逻辑关系,都在本专利的保护范围内;对算法中或者流程中添加无关紧要的修改或者引入无关紧要的设计,但不改变其算法和流程的核心设计都在该专利的保护范围内。

[0124] 图5是本发明第五实施例提供的一种路由处理装置的原理框图,应用于不处于跨设备链路聚合组的邻居设备。如图5所示,该路由处理装置包括:第一接收模块501、判断模块502和第一发送模块503。

[0125] 第一接收模块501,用于接收第一邻居设备发送的第一路由信息。

[0126] 在本实施例中,当前设备(即执行主体)为不处于跨设备链路聚合组的设备,在当前设备与处于跨设备链路聚合组中的设备建立两两邻居关系之后,当前设备与跨链路聚合组中的设备即互为邻居设备。对于当前设备来说,它的邻居设备就是跨设备链路聚合组中的设备。进一步地,将处于跨设备链路聚合组中的邻居设备定义为第一邻居设备。而且,跨设备链路聚合组中的设备包括主设备和从设备,因此,当前设备的第一邻居设备就包括跨设备链路聚合组的主设备和跨设备链路聚合组的从设备。

[0127] 在一个实施方式中,VTEP1、VTEP2和VTEP3均为可扩展虚拟局域网隧道端点,其中,VTEP1与VTEP2基于跨设备链路聚合组机制组成跨设备链路聚合组,且VTEP1为该跨设备链路聚合组的主设备,VTEP2为该跨设备链路聚合组的从设备,VTEP1与VTEP2构成双活系统。VTEP1与VTEP2之间使用不同的地址作为BGP对等体地址,相互建立边界网关协议(Border Gateway Protocol,BGP)邻居关系,并分别与不处于该跨设备链路聚合组的VTEP3也建立BGP邻居关系。

[0128] 在本实施方式中,当前设备为VTEP3,VTEP3对应的第一邻居设备包括VTEP1和VTEP2。在VTEP1、VTEP2和VTEP3建立两两BGP邻居关系之后,VTEP1和VTEP2分别向VTEP3发送第一路由信息,VTEP3通过第一接收模块501接收其第一邻居设备VTEP1和VTEP2分别发送的第一路由信息(即VTEP3通过第一接收模块501接收的第一路由信息的数量为两条,一条来自于VTEP1,另一条来自于VTEP2)。其中,第一路由信息为Type3路由。Type3路由即inclusive multicast路由,该类型路由主要用于完成VTEP地址的自动发现和VXLAN隧道的动态建立,以及通告二层VXLAN网络标识(VXLAN Network Identifier,VNI)信息和VTEPIP地址信息。

[0129] 判断模块502,用于根据第一路由信息判断第一邻居设备是否为跨设备链路聚合组的主设备。

[0130] 在一个实施方式中,第一路由信息为Type3路由,在Type3路由的运营商组播服务虚拟出接口(Provider Multicast Service Interface,PMSI)属性集中新增两个属性,分别是跨设备链路聚合组编号和设备主备标识。当前设备接收Type3路由后,根据跨设备链路

聚合组编号和设备主备标识对应的属性值即可判断发送Type3路由的第一邻居设备是否为跨设备链路聚合组的主设备。

[0131] 第一发送模块503,用于当第一邻居设备为跨设备链路聚合组的主设备时,向第一邻居设备发送第二路由信息,以供第一邻居设备根据第二路由信息生成第一地址表项信息,并将第一地址表项信息发送至跨设备链路聚合组的从设备,以使跨设备链路聚合组的从设备基于第一地址表项信息生成第二地址表项信息。

[0132] 在一个实施方式中,VTEP1、VTEP2和VTEP3建立两两邻居关系,其中,VTEP1与VTEP2分别为跨设备链路聚合组的主设备和跨设备链路聚合组的从设备,VTEP3为不处于该跨设备链路聚合组的邻居设备。

[0133] 在本实施方式中,当前设备为VTEP3,VTEP3对应的第一邻居设备包括VTEP1和VTEP2。当VTEP3判断VTEP1为跨设备链路聚合组的主设备时(相当于VTEP3也获知VTEP2为该跨设备链路聚合组的从设备),VTEP3通过第一发送模块503向VTEP1发送第二路由信息(VTEP3不向VTEP2发送第二路由信息)。其中,第二路由信息为type2路由,也即MAC/IP路由,该类型路由主要用于通告mac地址和主机IP地址。

[0134] VTEP1收到type2路由后,基于type2路由生成VTEP1的本地第一地址表项信息,并通过VTEP1与VTEP2之间的peer-link链路将第一地址表项信息发送至VTEP2。VTEP2收到第一地址表项信息之后,根据第一地址表项信息生成VTEP2本地的第二地址表项信息。其中,peer-link链路为VTEP1与VTEP2之间的直连聚合链路,一般情况下,peer-link链路是一条二层链路,用于VTEP1与VTEP2之间协商报文的交互以及部分流量的传输。

[0135] 图6是本发明第六实施例提供的一种路由处理装置的原理框图,应用于处于跨设备链路聚合组的主设备。如图6所示,该路由处理装置包括:第二发送模块601、第二接收模块602、生成模块603和第三发送模块604。

[0136] 第二发送模块601,用于向第二邻居设备发送第一路由信息。

[0137] 在本实施例中,当前设备(即执行主体)为跨设备链路聚合组中的主设备,当前设备与跨设备链路聚合组中的从设备以及不处于跨设备链路聚合组的设备建立邻居连接之后,当前设备与跨设备链路聚合组中的从设备以及不处于跨设备链路聚合组的设备互为邻居设备。对于当前设备而言,其邻居设备包括两类,一类是处于该跨设备链路聚合组的从设备,另一类是不处于跨设备链路聚合组的邻居设备。进一步地,将不处于跨设备链路聚合组的邻居设备定义为第二邻居设备。

[0138] 在一个实施方式中,VTEP1、VTEP2和VTEP3均为可扩展虚拟局域网隧道端点,其中,VTEP1与VTEP2跨设备链路聚合组中的设备,且VTEP1为该跨设备链路聚合组的主设备,VTEP2为该跨设备链路聚合组的从设备。VTEP1、VTEP2和VTEP3之间建立两两BGP邻居关系。

[0139] 在本实施方式中,当前设备为VTEP1,VTEP1对应的第二邻居设备为VTEP3。VTEP1、VTEP2和VTEP3建立两两邻居关系之后,VTEP1通过第二发送模块601向VTEP3发送第一路由信息,其中,第一路由信息为Type3路由,而且在Type3路由中新增两个属性,分别是跨设备链路聚合组编号和设备主备标识,以供VTEP3根据Type3路由中的跨设备链路聚合组编号和设备主备标识确定其第一邻居设备中的跨设备链路聚合组的主设备。

[0140] 第二接收模块602,用于接收第二邻居设备返回的第二路由信息。

[0141] 在一个实施方式中,VTEP1与VTEP2分别为跨设备链路聚合组的主设备和跨设备链

路聚合组的从设备, VTEP1、VTEP2与不处于该跨设备链路聚合组的VTEP3建立两两邻居关系。

[0142] 在本实施方式中, 当前设备为VTEP1, VTEP1对应的第二邻居设备为VTEP3(可以理解的是, VTEP1的邻居设备还包括VTEP2, 只不过VTEP2不属于本实施例中定义的第二邻居设备类型)。

[0143] VTEP3分别收到VTEP1和VTEP2发送的第一路由信息之后, 根据第一路由判断VTEP1和VTEP2是否为跨设备链路聚合组的主设备。当VTEP3判断出VTEP1为跨设备链路聚合组的主设备时, VTEP3向VTEP1发送第二路由信息。VTEP1通过第二接收模块602接收VTEP3发送的第二路由信息。其中, 第二路由信息为type2路由, 也即MAC/IP路由, 该类型路由主要用于通告mac地址和主机IP地址。

[0144] 需要说明的是, 虽然VTEP1和VTEP2均会向VTEP3发送第一路由信息。但是, VTEP3只向VTEP1返回第二路由信息, 从而可以避免后续VTEP1与VTEP2之间发送重复的地址表项信息, 同时还可降低VTEP1与VTEP2的处理路由信息的复杂程度。

[0145] 生成模块603, 用于根据第二路由信息生成第一地址表项信息。

[0146] 在一个实施方式中, 当前设备接收第二邻居设备发送的第二路由信息, 其中, 第二路由信息为type2路由。当前设备根据type2路由通过生成模块603在本地MAC地址表中生成对应的地址表项。

[0147] 第三发送模块604, 用于将第一地址表项信息发送至跨设备链路聚合组的从设备, 以供跨设备链路聚合组的从设备基于第一地址表项信息生成对应的第二地址表项信息。

[0148] 在一个实施方式中, 当前设备通过第三发送模块604由peer-link链路将第一地址表项信息发送至跨设备链路聚合组的从设备。跨设备链路聚合组的从设备接收第一地址表项信息后, 根据第一地址表项信息在本地MAC地址表中生成对应的第二地址表项信息。

[0149] 值得一提的是, 本实施方式中所涉及到的各模块均为逻辑模块, 在实际应用中, 一个逻辑单元可以是一个物理单元, 也可以是一个物理单元的一部分, 还可以以多个物理单元的组合实现。此外, 为了突出本发明的创新部分, 本实施方式中并没有将与解决本发明所提出的技术问题关系不太密切的单元引入, 但这并不表明本实施方式中不存在其它的单元。

[0150] 可以理解的是, 以上实施方式仅仅是为了说明本发明的原理而采用的示例性实施方式, 然而本发明并不局限于此。对于本领域内的普通技术人员而言, 在不脱离本发明的精神和实质的情况下, 可以做出各种变型和改进, 这些变型和改进也视为本发明的保护范围。

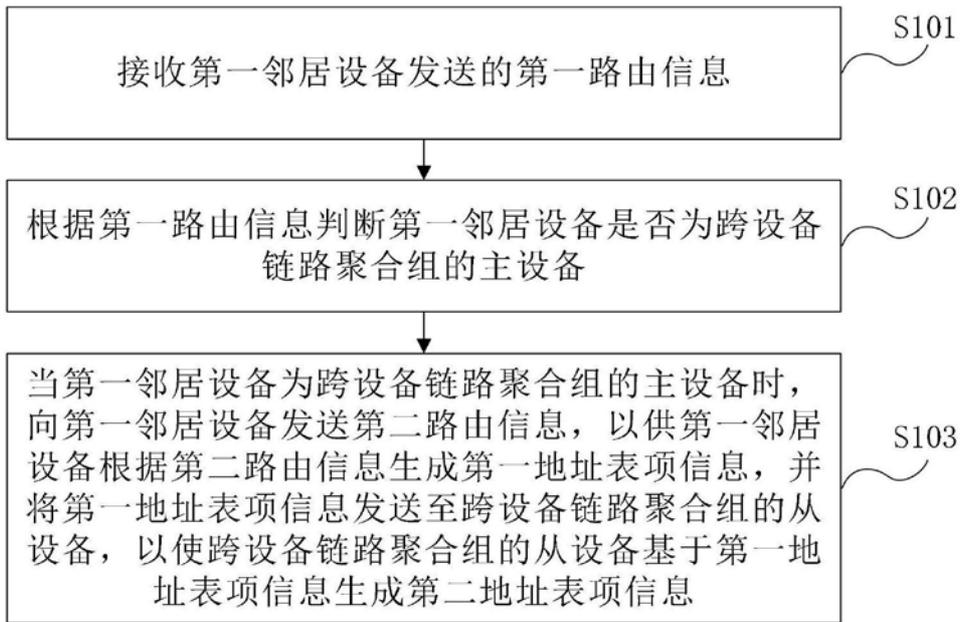


图1

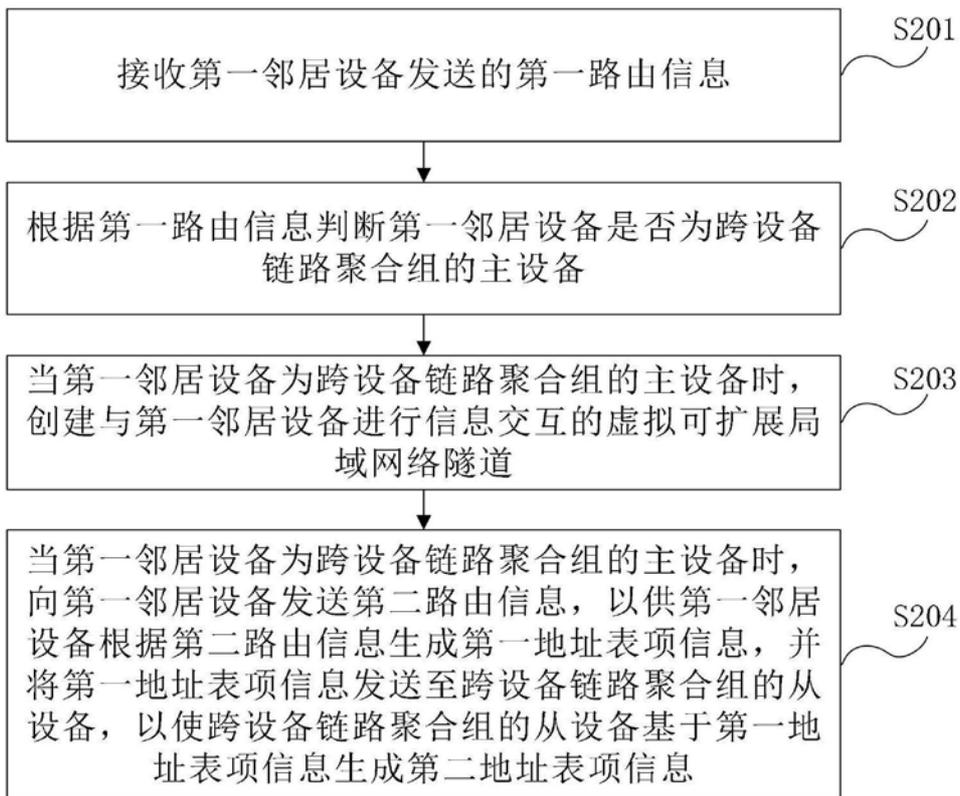


图2

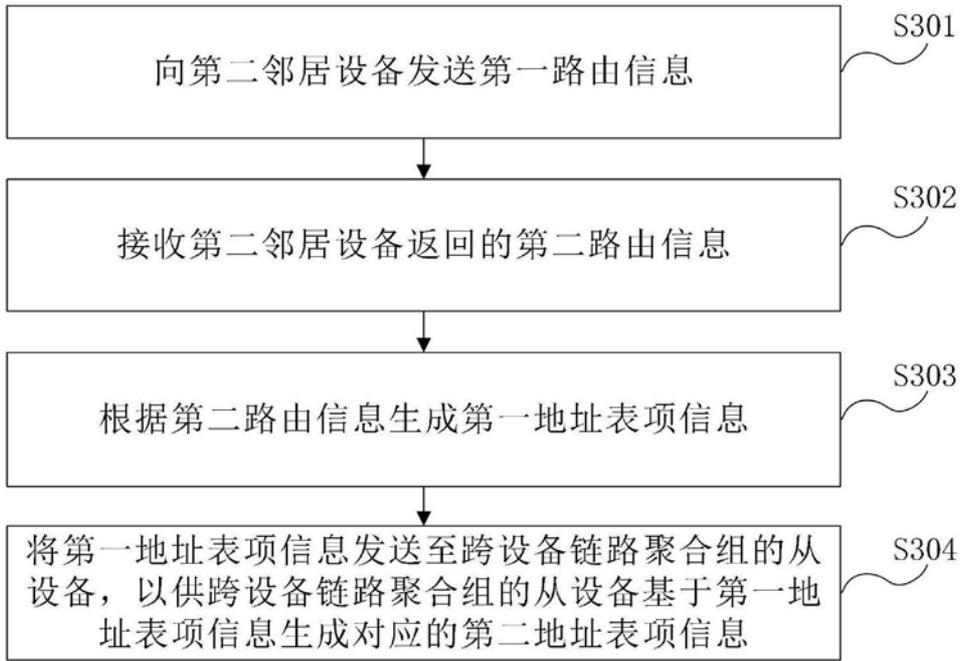


图3

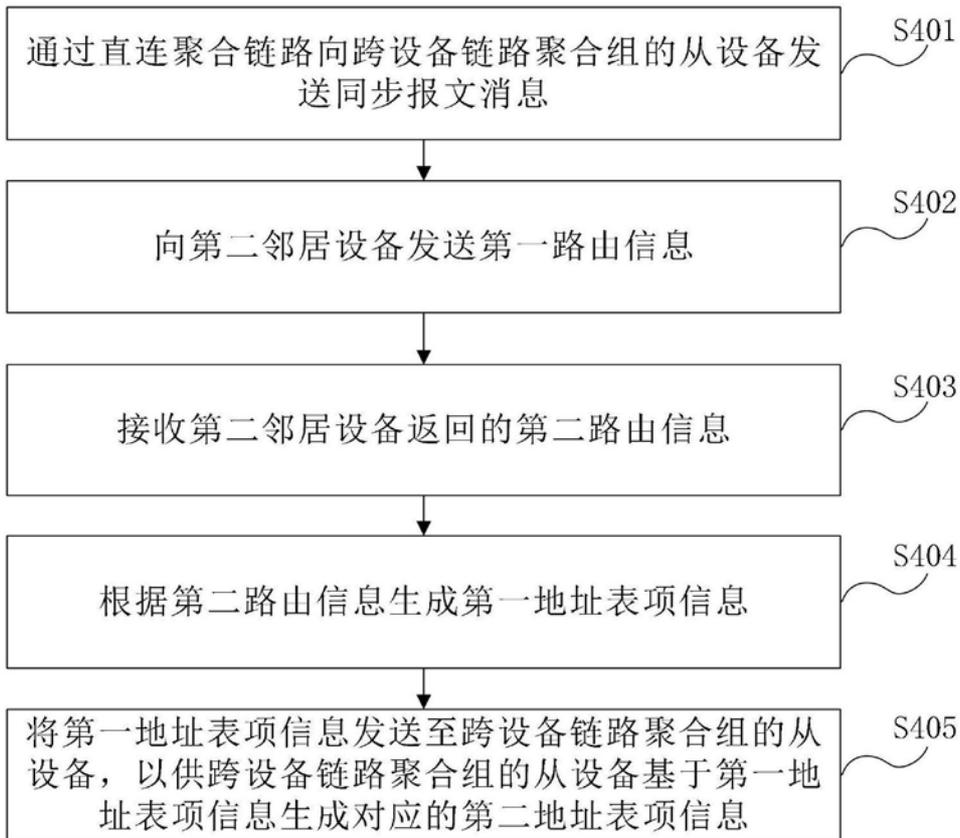


图4

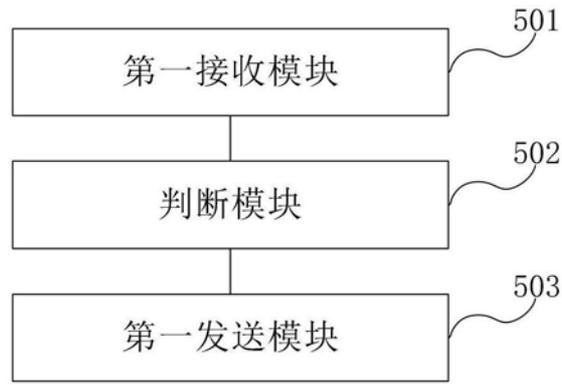


图5

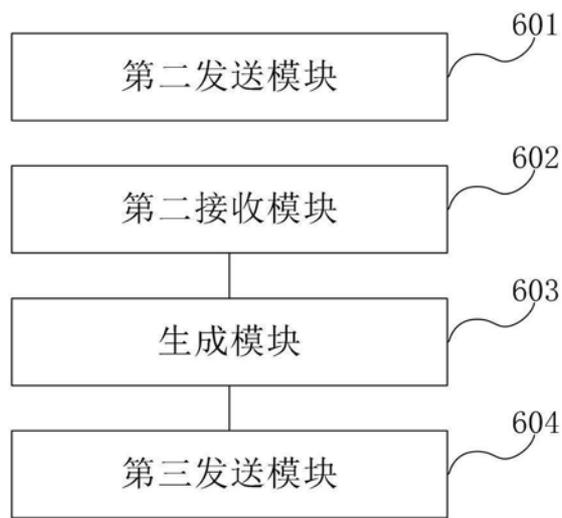


图6