



(12)发明专利

(10)授权公告号 CN 104618243 B

(45)授权公告日 2017. 11. 17

(21)申请号 201510092158.0

(22)申请日 2015.02.28

(65)同一申请的已公布的文献号

申请公布号 CN 104618243 A

(43)申请公布日 2015.05.13

(73)专利权人 华为技术有限公司

地址 518129 广东省深圳市龙岗区坂田华为总部办公楼

(72)发明人 李明

(74)专利代理机构 广州三环专利商标代理有限公司 44202

代理人 郝传鑫 熊永强

(51)Int.Cl.

H04L 12/741(2013.01)

H04L 29/12(2006.01)

(56)对比文件

CN 101237378 A,2008.08.06,

CN 101317369 A,2008.12.03,

CN 101729425 A,2010.06.09,

US 2011/0283013 A1,2011.11.17,

审查员 吴晗

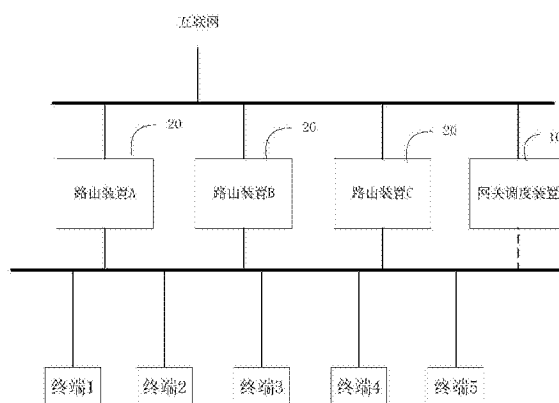
权利要求书5页 说明书24页 附图8页

(54)发明名称

路由方法、装置及系统、网关调度方法及装置

(57)摘要

本发明实施例公开了一种路由系统包括网关调度装置和至少两个路由装置,其中:所述网关调度装置用于接收终端添加指令,所述终端添加指令包含终端的互联网协议IP地址;在所述至少两个路由装置中选择任意一个作为所述终端归属的网关,将所述终端的IP地址添加到所述终端归属的网关的终端IP列表中;所述终端归属的网关用于向所述终端发送地址解析协议ARP响应报文,所述ARP响应报文用于告知所述终端绑定所述终端归属的网关。本发明还公开了一种路由方法及装置、网关调度方法及装置。采用本发明,可规避三角路由问题以及单点性能瓶颈并方便扩容。



1. 一种路由系统,其特征在于,所述路由系统包括网关调度装置和至少两个路由装置,其中:

所述网关调度装置用于接收终端添加指令,所述终端添加指令包含终端的互联网协议IP地址;在所述至少两个路由装置中选择任意一个作为所述终端归属的网关,将所述终端的IP地址添加到所述终端归属的网关的终端IP列表中;

所述终端归属的网关用于向所述终端发送地址解析协议ARP响应报文,所述ARP响应报文用于告知所述终端绑定所述终端归属的网关。

2. 根据权利要求1所述的路由系统,其特征在于,所述终端归属的网关具体用于:在向所述终端发送ARP响应报文之前,接收所述终端发送的询问网关的ARP请求报文,所述ARP请求报文携带所述终端的IP地址;判断所述终端归属的网关的终端IP列表中是否包含所述终端的IP地址,在所述终端归属的网关的终端IP列表包含所述终端的IP地址时,获取所述终端归属的网关的MAC地址,通过ARP响应报文将所述MAC地址返回所述终端。

3. 根据权利要求1所述的路由系统,其特征在于,所述终端归属的网关还用于接收所述终端发送的上行数据报文,所述上行数据报文中携带网关MAC地址;在所述携带的网关MAC地址与所述终端归属的网关的MAC地址相同时,转发所述上行数据报文。

4. 根据权利要求2所述的路由系统,其特征在于,所述终端归属的网关还用于接收所述终端发送的上行数据报文,所述上行数据报文中携带网关MAC地址;在所述携带的网关MAC地址与所述终端归属的网关的MAC地址相同时,转发所述上行数据报文。

5. 根据权利要求3所述的路由系统,其特征在于,所述终端归属的网关具有预设的外网IP地址;

所述终端归属的网关还用于根据所述终端归属的网关的外网IP地址对所述上行数据报文进行源地址转换SNAT。

6. 根据权利要求4所述的路由系统,其特征在于,所述终端归属的网关具有预设的外网IP地址;

所述终端归属的网关还用于根据所述终端归属的网关的外网IP地址对所述上行数据报文进行源地址转换SNAT。

7. 根据权利要求1-6任意一项所述的路由系统,其特征在于,所述网关调度装置还用于接收广播的ARP请求报文,获取所述广播的ARP请求报文的源IP地址;在所述源IP地址不属于所述至少两个路由装置各自的终端IP列表时,在至少两个路由装置中选择任意一个作为所述源IP地址对应的终端归属的网关。

8. 根据权利要求1-6任意一项所述的路由系统,其特征在于,所述网关调度装置还用于检测所述至少两个路由装置的负载信息;根据检测到的所述负载信息在所述至少两个路由装置中选择任意一个作为所述终端归属的网关。

9. 根据权利要求7所述的路由系统,其特征在于,所述网关调度装置还用于检测所述至少两个路由装置的负载信息;根据检测到的所述负载信息在所述至少两个路由装置中选择任意一个作为所述终端归属的网关。

10. 根据权利要求8所述的路由系统,其特征在于,所述网关调度装置还用于根据检测到所述负载信息在所述至少两个路由装置中选择作为迁移源的路由装置和作为迁移目的的路由装置;选取所述作为迁移源的路由装置的终端IP列表中的终端的IP地址作为迁移IP

地址,并将所述迁移IP地址通知给所述作为迁移目的的路由装置;

所述作为迁移目的的路由装置还用于向所述迁移IP地址对应的终端发送免费ARP报文,通知所述迁移IP地址对应的终端更新所述迁移IP地址对应的终端的网关MAC地址。

11.根据权利要求9所述的路由系统,其特征在于,所述网关调度装置还用于根据检测到所述负载信息在所述至少两个路由装置中选择作为迁移源的路由装置和作为迁移目的的路由装置;选取所述作为迁移源的路由装置的终端IP列表中的终端的IP地址作为迁移IP地址,并将所述迁移IP地址通知给所述作为迁移目的的路由装置;

所述作为迁移目的的路由装置还用于向所述迁移IP地址对应的终端发送免费ARP报文,通知所述迁移IP地址对应的终端更新所述迁移IP地址对应的终端的网关MAC地址。

12.根据权利要求10所述的路由系统,其特征在于,所述网关调度装置还用于将所述迁移IP地址通知所述作为迁移源的路由装置;

所述作为迁移源的路由装置还用于将所述迁移IP地址由所述作为迁移源的路由装置的终端IP列表中移除。

13.根据权利要求11所述的路由系统,其特征在于,所述网关调度装置还用于将所述迁移IP地址通知所述作为迁移源的路由装置;

所述作为迁移源的路由装置还用于将所述迁移IP地址由所述作为迁移源的路由装置的终端IP列表中移除。

14.根据权利要求12所述的路由系统,其特征在于,所述作为迁移源的路由装置还用于查找与所述迁移IP地址对应的弹性互联网协议EIP地址,将查找到的所述EIP地址移除。

15.根据权利要求13所述的路由系统,其特征在于,所述作为迁移源的路由装置还用于查找与所述迁移IP地址对应的弹性互联网协议EIP地址,将查找到的所述EIP地址移除。

16.一种路由方法,其特征在于,所述方法包括:

接收询问网关的地址解析协议ARP请求报文,获取所述ARP请求报文携带的终端的IP地址;

获取预设的终端IP列表,在所述终端IP列表包含所述获取到终端的IP地址时,通过ARP响应报文将自身的MAC地址返回所述终端;

接收上行数据报文,提取所述上行数据报文中携带的网关MAC地址;

在所述上行数据报文中携带的网关MAC地址与自身的MAC地址相同时,转发所述上行数据报文。

17.根据权利要求16所述的路由方法,其特征在于,所述转发所述上行数据报文之前,所述方法还包括:

根据预设的外网IP地址对所述上行数据报文进行源地址转换SNAT,生成相应的地址转换记录。

18.根据权利要求17所述的路由方法,其特征在于,所述转发所述上行数据报文之后,所述方法还包括:

接收其目的IP地址为所述外网IP地址的下行数据报文,获取所述下行数据报文的端口;

查找与所述目的端口对应的地址转换记录,根据所述查找到的地址转换记录对所述下行数据报文进行目的地址转换DNAT,转发所述DNAT后的下行数据报文。

19. 根据权利要求16-18任意一项所述的路由方法,其特征在于,所述方法还包括:

接收弹性互联网协议EIP地址绑定指令,获取EIP地址绑定指令中包含的EIP地址以及与所述EIP地址为映射关系的终端的IP地址;

接收目的地址为所述EIP地址的下行数据报文,将所述下行数据报文的目IP地址设置为所述EIP地址对应的终端的IP地址,并转发所述下行数据报文至与所述EIP地址对应的终端的IP地址。

20. 根据权利要求19所述的路由方法,其特征在于,所述方法还包括:

接收路由服务移除指令,获取路由服务移除指令对应的终端的IP地址;

查找所述终端的IP地址对应的EIP地址,将查找到的EIP地址移除。

21. 根据权利要求16-18任意一项所述的路由方法,其特征在于,所述方法还包括:

接收路由服务添加指令,获取路由服务添加指令对应的终端的IP地址;

向所述终端发送免费地址解析协议报文,所述终端根据所述免费地址解析协议报文更新网关MAC地址。

22. 根据权利要求19所述的路由方法,其特征在于,所述方法还包括:

接收路由服务添加指令,获取路由服务添加指令对应的终端的IP地址;

向所述终端发送免费地址解析协议报文,所述终端根据所述免费地址解析协议报文更新网关MAC地址。

23. 根据权利要求20所述的路由方法,其特征在于,所述方法还包括:

接收路由服务添加指令,获取路由服务添加指令对应的终端的IP地址;

向所述终端发送免费地址解析协议报文,所述终端根据所述免费地址解析协议报文更新网关MAC地址。

24. 一种路由装置,其特征在于,包括ARP报文接收模块、与ARP报文接收模块连接的MAC地址返回模块、数据报文接收模块、与所述数据报文接收模块连接的数据报文转发模块:

ARP报文接收模块,用于接收询问网关的地址解析协议ARP请求报文,获取所述ARP请求报文携带的终端的IP地址;

MAC地址返回模块,用于获取预设的终端IP列表,在所述终端IP列表包含所述获取到终端的IP地址时,通过ARP响应报文将自身的MAC地址返回与所述终端;

数据报文接收模块,用于接收上行数据报文,提取所述上行数据报文中携带的网关MAC地址;

数据报文转发模块,用于在所述上行数据报文中携带的网关MAC地址与自身的MAC地址相同时,转发所述上行数据报文。

25. 根据权利要求24所述的路由装置,其特征在于,所述数据报文转发模块还用于根据预设的外网IP地址对所述上行数据报文进行源地址转换SNAT,生成相应的地址转换记录。

26. 根据权利要求25所述的路由装置,其特征在于,所述数据报文转发模块还用于接收其目的IP地址为所述外网IP地址的下行数据报文,获取所述下行数据报文的端口;查找与所述目的端口对应的地址转换记录,根据所述查找到的地址转换记录对所述下行数据报文进行目的地址转换DNAT,转发所述DNAT后的下行数据报文。

27. 根据权利要求24-26任意一项所述的路由装置,其特征在于,所述装置还包括EIP地址绑定模块,用于接收弹性互联网协议EIP地址绑定指令,获取EIP地址绑定指令中包含的

EIP地址以及与所述EIP地址为映射关系的终端的IP地址；

所述数据报文转发模块还用于接收目的地址为所述EIP地址的下行数据报文，将所述下行数据报文的IP地址设置为所述EIP地址对应的终端的IP地址，并转发所述下行数据报文至所述终端的IP地址对应的终端。

28. 根据权利要求27所述的路由装置，其特征在于，所述装置还包括路由服务移除模块，用于接收路由服务移除指令，获取路由服务移除指令对应的终端的IP地址；查找所述终端的IP地址对应的EIP地址，将查找到的EIP地址移除。

29. 根据权利要求24至26任意一项所述的路由装置，其特征在于，所述装置还包括路由服务添加模块，用于接收路由服务添加指令，获取路由服务添加指令对应的终端的IP地址；向所述终端发送免费地址解析协议报文，所述终端根据所述免费地址解析协议报文更新网关MAC地址。

30. 根据权利要求27所述的路由装置，其特征在于，所述装置还包括路由服务添加模块，用于接收路由服务添加指令，获取路由服务添加指令对应的终端的IP地址；向所述终端发送免费地址解析协议报文，所述终端根据所述免费地址解析协议报文更新网关MAC地址。

31. 根据权利要求28所述的路由装置，其特征在于，所述装置还包括路由服务添加模块，用于接收路由服务添加指令，获取路由服务添加指令对应的终端的IP地址；向所述终端发送免费地址解析协议报文，所述终端根据所述免费地址解析协议报文更新网关MAC地址。

32. 一种网关调度方法，其特征在于，所述方法包括：

接收终端添加指令，所述终端添加指令包含终端的IP地址；

在至少两个路由装置中选择任意一个作为所述终端归属的网关；

将所述终端的IP地址添加到所述终端归属的网关的终端IP列表中，所述终端归属的网关通过ARP响应报文告知所述终端该路由装置的MAC地址。

33. 根据权利要求32所述的网关调度方法，其特征在于，所述在至少两个路由装置中选择任意一个作为所述终端归属的网关之前，所述方法还包括：

接收广播的ARP请求报文，获取所述广播的ARP请求报文的源IP地址；在所述源IP地址不属于所述至少两个路由装置各自的终端IP列表时，在至少两个路由装置中选择任意一个作为所述源IP地址对应的终端归属的网关。

34. 根据权利要求32或33所述的网关调度方法，其特征在于，所述方法还包括：

检测所述至少两个路由装置的负载信息；

则所述在至少两个路由装置中选择路由装置具体为：

根据检测到的所述负载信息在所述至少两个路由装置中选择路由装置终端的IP地址。

35. 根据权利要求34所述的网关调度方法，其特征在于，所述检测所述至少两个路由装置的负载信息之后，所述方法还包括：

根据检测到的路由装置的负载信息在至少两个路由装置选择作为迁移源的路由装置和作为迁移目的的路由装置；

选取所述作为迁移源的路由装置的终端IP列表中的终端的IP地址作为迁移IP地址，并将所述迁移IP地址通知给所述作为迁移目的的路由装置。

36. 一种网关调度装置，其特征在于，所述装置包括终端IP获取模块、与终端IP获取模块连接的路由装置选择模块以及与所述路由装置选择模块连接的网关分配模块，其中：

终端IP获取模块,用于接收终端添加指令,所述终端添加指令包含终端的IP地址;

路由装置选择模块,用于在至少两个路由装置中选择任意一个作为所述终端归属的网关;

网关分配模块,用于将所述终端的IP地址添加到所述终端归属的网关的终端IP列表中,所述终端归属的网关通过ARP响应报文告知所述终端该路由装置的MAC地址。

37.根据权利要求36所述的网关调度装置,其特征在于,所述终端IP获取模块还用于接收广播的ARP请求报文,获取所述广播的ARP请求报文的源IP地址;在所述源IP地址不属于所述至少两个路由装置各自的终端IP列表时,在至少两个路由装置中选择任意一个作为所述源IP地址对应的终端归属的网关。

38.根据权利要求36或37所述的网关调度装置,其特征在于,所述装置还包括负载检测模块,用于检测所述至少两个路由装置的负载信息;

所述路由装置选择模块还用于根据检测到的所述负载信息在所述至少两个路由装置中选择路由装置终端的IP地址。

39.根据权利要求38所述的网关调度装置,其特征在于,所述装置还包括负载均衡模块,用于根据检测到的路由装置的负载信息在至少两个路由装置选择作为迁移源的路由装置和作为迁移目的的路由装置;选取所述作为迁移源的路由装置的终端IP列表中的终端的IP地址作为迁移IP地址,并将所述迁移IP地址通知给所述作为迁移目的的路由装置。

路由方法、装置及系统、网关调度方法及装置

技术领域

[0001] 本发明涉及网络技术领域,特别是涉及一种路由方法、装置及系统、网关调度方法及装置。

背景技术

[0002] 传统技术中的路由器通常采用集中式的路由方式,即所有终端接入该路由器,该路由器负责转发所有接入该路由器的终端的数据流量。如图1所示,传统技术中的集中式路由方式通常采用主备模式,主路由器与多个终端组成以太网局域网,终端与外部的互联网之间的数据流量由该主路由器负责转发,该主路由器即为以太网局域网中的终端访问外部的互联网的网关设备。

[0003] 为避免上述集中式路由产生的单点性能瓶颈问题,传统技术中采用了等价路由(英文:Equal Cost Multi Path,简称:ECMP)的方式。若路由表中存在多条目的地址相同、权重相同,但是下一跳不同的路由,这些路由都称为等价路由。若路由器A、B、C为等价路由,终端发出的上行数据报文被随机分配到某个路由器进行转发,外部网络返回的下行数据报文也被随机分配到某个路由器进行转发,从而使得转发业务的负载得到均衡。但等价路由的方式则存在三角路由的情况。例如,若在路由器A、B、C上实现SNAT/NAT操作,则会出现三角路由的情况导致负责转发同一终端的上行数据报文和下行数据报文不一致的情况,造成数据包丢失。因此,传统技术中的路由方法无法解决三角路由问题。

发明内容

[0004] 基于此,为了上述传统技术中的路由方式的存在的三角路由问题,提供了一种路由系统。

[0005] 本发明实施例第一方面公开了一种路由系统,所述路由系统包括网关调度装置和至少两个路由装置,其中:

[0006] 所述网关调度装置用于接收终端添加指令,所述终端添加指令包含终端的互联网协议IP地址;在所述至少两个路由装置中选择任意一个作为所述终端归属的网关,将所述终端的IP地址添加到所述终端归属的网关的终端IP列表中;

[0007] 所述终端归属的网关用于向所述终端发送地址解析协议ARP响应报文,所述ARP响应报文用于告知所述终端绑定所述终端归属的网关。

[0008] 在本发明实施例第一方面的第一种可能的实现方式中,所述终端归属的网关具体用于:在向所述终端发送ARP响应报文之前,接收所述终端发送的询问网关的ARP请求报文,所述ARP请求报文携带所述终端的IP地址;判断所述终端归属的网关的终端IP列表中是否包含所述终端的IP地址,在所述终端归属的网关的终端IP列表包含所述终端的IP地址时,获取所述终端归属的网关的MAC地址,通过ARP响应报文将所述MAC地址返回所述终端。

[0009] 结合本发明实施例第一方面和第一方面的第一种可能的实现方式,在本发明实施例第一方面的第二种可能的实现方式中,所述终端归属的网关还用于接收所述终端发送的

上行数据报文,所述上行数据报文中携带网关MAC地址;在所述携带的网关MAC地址与所述终端归属的网关的MAC地址相同时,转发所述上行数据报文。

[0010] 结合本发明实施例第一方面的第二种可能的实现方式,在本发明实施例第一方面的第三种可能的实现方式中,所述终端归属的网关具有预设的外网IP地址;所述终端归属的网关还用于根据所述终端归属的网关的外网IP地址对所述上行数据报文进行源地址转换SNAT。

[0011] 结合本发明实施例第一方面或第一方面的第一种至第三种中任一种可能的实现方式,在本发明实施例第一方面的第四种可能的实现方式中,在本发明实施例第一方面的第三种可能的实现方式中,所述网关调度装置还用于接收广播的ARP请求报文,获取所述广播的ARP请求报文的源IP地址;在所述源IP地址不属于所述至少两个路由装置各自的终端IP列表时,在至少两个路由装置中选择任意一个作为所述源IP地址对应的终端归属的网关。

[0012] 结合本发明实施例第一方面或第一方面的第一种至第四种中任一种可能的实现方式,在本发明实施例第一方面的第五种可能的实现方式中,所述网关调度装置还用于检测所述至少两个路由装置的负载信息;根据检测到的所述负载信息在所述至少两个路由装置中选择任意一个作为所述终端归属的网关。

[0013] 结合本发明实施例第一方面的第五种可能的实现方式,在本发明实施例第一方面的第六种可能的实现方式中,所述网关调度装置还用于根据检测到所述负载信息在所述至少两个路由装置中选择作为迁移源的路由装置和作为迁移目的的路由装置;选取所述作为迁移源的路由装置的终端IP列表中的终端的IP地址作为迁移IP地址,并将所述迁移IP地址通知给所述作为迁移目的的路由装置;

[0014] 所述作为迁移目的的路由装置还用于向所述迁移IP地址对应的终端发送免费ARP报文,通知所述迁移IP地址对应的终端更新所述迁移IP地址对应的终端的网关MAC地址。

[0015] 结合本发明实施例第一方面的第六种可能的实现方式,在本发明实施例第一方面的第七种可能的实现方式中,所述网关调度装置还用于将所述迁移IP地址通知所述作为迁移源的路由装置;

[0016] 所述作为迁移源的路由装置还用于将所述迁移IP地址由所述作为迁移源的路由装置的终端IP列表中移除。

[0017] 结合本发明实施例第一方面的第七种可能的实现方式,在本发明实施例第一方面的第八种可能的实现方式中,所述作为迁移源的路由装置还用于查找与所述迁移IP地址对应的弹性互联网协议EIP地址,将查找到的所述EIP地址移除。

[0018] 此外,为了上述传统技术中的路由方式的存在的三角路由问题,提供了一种路由方法。

[0019] 本发明实施例第二方面公开了一种路由方法,所述方法包括:

[0020] 接收询问网关的地址解析协议ARP请求报文,获取所述ARP请求报文携带的终端的IP地址;

[0021] 获取预设的终端IP列表,在所述终端IP列表包含所述获取到终端的IP地址时,通过ARP响应报文将自身的MAC地址返回所述终端;

[0022] 接收上行数据报文,提取所述上行数据报文中携带的网关MAC地址;

[0023] 在所述上行数据报文中携带的网关MAC地址与自身的MAC地址相同时,转发所述上行数据报文。

[0024] 在本发明实施例第二方面的第一种可能的实现方式中,所述转发所述上行数据报文之前,所述方法还包括:

[0025] 根据预设的外网IP地址对所述上行数据报文进行源地址转换SNAT,生成相应的地址转换记录。

[0026] 结合本发明实施例第二方面的第一种可能的实现方式,在本发明实施例第二方面的第二种可能的实现方式中,所述转发所述上行数据报文之后,所述方法还包括:

[0027] 接收其目的IP地址为所述外网IP地址的下行数据报文,获取所述下行数据报文的端口;

[0028] 查找与所述端口对应的地址转换记录,根据所述查找到的地址转换记录对所述下行数据报文进行目的地址转换DNAT,转发所述DNAT后的下行数据报文。

[0029] 结合本发明实施例第二方面或第二方面的第一种或第二种可能的实现方式,在本发明实施例第二方面的第三种可能的实现方式中,所述方法还包括:

[0030] 接收弹性互联网协议EIP地址绑定指令,获取EIP地址绑定指令中包含的EIP地址以及与所述EIP地址为映射关系的终端的IP地址;

[0031] 接收目的地址为所述EIP地址的下行数据报文,将所述下行数据报文的端口IP地址设置为所述EIP地址对应的终端的IP地址,并转发所述下行数据报文至与所述EIP地址对应的终端的IP地址。

[0032] 结合本发明实施例第二方面的第三种可能的实现方式,在本发明实施例第二方面的第四种可能的实现方式中,所述方法还包括:

[0033] 接收路由服务移除指令,获取路由服务移除指令对应的终端的IP地址;

[0034] 查找所述终端的IP地址对应的EIP地址,将查找到的EIP地址移除。

[0035] 结合本发明实施例第二方面或第二方面的第一种至第四种中任一种可能的实现方式,在本发明实施例第二方面的第五种可能的实现方式中,所述方法还包括:

[0036] 接收路由服务添加指令,获取路由服务添加指令对应的终端的IP地址;

[0037] 向所述终端发送免费地址解析协议报文,所述终端根据所述免费地址解析协议报文更新网关MAC地址。

[0038] 此外,为了上述传统技术中的路由方式的存在的三角路由问题,提供了一种路由装置。

[0039] 本发明实施例第三方面公开了一种路由装置,包括ARP报文接收模块、与ARP报文接收模块连接的MAC地址返回模块、数据报文接收模块、与所述数据报文接收模块连接的数据报文转发模块:

[0040] ARP报文接收模块,用于接收询问网关的地址解析协议ARP请求报文,获取所述ARP请求报文携带的终端的IP地址;

[0041] MAC地址返回模块,用于获取预设的终端IP列表,在所述终端IP列表包含所述获取到终端的IP地址时,通过ARP响应报文将自身的MAC地址返回与所述终端;

[0042] 数据报文接收模块,用于接收上行数据报文,提取所述上行数据报文中携带的网关MAC地址;

[0043] 数据报文转发模块,用于在所述上行数据报文中携带的网关MAC地址与自身的MAC地址相同时,转发所述上行数据报文。

[0044] 在本发明实施例第三方面的第一种可能的实现方式中,所述数据报文转发模块还用于根据预设的外网IP地址对所述上行数据报文进行源地址转换SNAT,生成相应的地址转换记录。

[0045] 结合本发明实施例第三方面的第一种可能的实现方式,在本发明实施例第三方面的第二种可能的实现方式中,所述数据报文转发模块还用于接收其目的IP地址为所述外网IP地址的下行数据报文,获取所述下行数据报文的端口;查找与所述端口对应的地址转换记录,根据所述查找到的地址转换记录对所述下行数据报文进行目的地址转换DNAT,转发所述DNAT后的下行数据报文。

[0046] 结合本发明实施例第三方面或第二方面的第一种或第二种可能的实现方式,在本发明实施例第三方面的第三种可能的实现方式中,所述装置还包括EIP地址绑定模块,用于接收弹性互联网协议EIP地址绑定指令,获取EIP地址绑定指令中包含的EIP地址以及与所述EIP地址为映射关系的终端的IP地址;

[0047] 所述数据报文转发模块还用于接收目的地址为所述EIP地址的下行数据报文,将所述下行数据报文的端口IP地址设置为所述EIP地址对应的终端的IP地址,并转发所述下行数据报文至所述终端的IP地址对应的终端。

[0048] 结合本发明实施例第三方面的第三种可能的实现方式,在本发明实施例第三方面的第四种可能的实现方式中,所述装置还包括路由服务移除模块,用于接收路由服务移除指令,获取路由服务移除指令对应的终端的IP地址;查找所述终端的IP地址对应的EIP地址,将查找到的EIP地址移除。

[0049] 结合本发明实施例第三方面或第二方面的第一种至第四种中任一种可能的实现方式,在本发明实施例第三方面的第五种可能的实现方式中,所述装置还包括路由服务添加模块,用于接收路由服务添加指令,获取路由服务添加指令对应的终端的IP地址;向所述终端发送免费地址解析协议报文,所述终端根据所述免费地址解析协议报文更新网关MAC地址。

[0050] 此外,为了上述传统技术中的路由方式的存在的三角路由问题,提供了一种网关调度方法。

[0051] 本发明实施例第四方面公开了一种网关调度方法,所述方法包括:

[0052] 接收终端添加指令,所述终端添加指令包含终端的IP地址;

[0053] 在至少两个路由装置中选择任意一个作为所述终端归属的网关;

[0054] 将所述终端的IP地址添加到所述终端归属的网关的终端IP列表中,所述终端归属的网关通过ARP响应报文告知所述终端该路由装置的MAC地址。

[0055] 在本发明实施例第四方面的第一种可能的实现方式中,所述在至少两个路由装置中选择任意一个作为所述终端归属的网关之前,所述方法还包括:

[0056] 接收广播的ARP请求报文,获取所述广播的ARP请求报文的源IP地址;在所述源IP地址不属于所述至少两个路由装置各自的终端IP列表时,在至少两个路由装置中选择任意一个作为所述源IP地址对应的终端归属的网关。

[0057] 结合本发明实施例第四方面和第四方面的第一种可能的实现方式,在本发明实施

例第四方面的第二种可能的实现方式中,所述方法还包括:

[0058] 检测所述至少两个路由装置的负载信息;

[0059] 则所述在至少两个路由装置中选择路由装置具体为:

[0060] 根据检测到的所述负载信息在所述至少两个路由装置中选择路由装置终端的IP地址。

[0061] 结合本发明实施例第四方面的第二种可能的实现方式,在本发明实施例第四方面的第三种可能的实现方式中,所述检测所述至少两个路由装置的负载信息之后,所述方法还包括:

[0062] 根据检测到的路由装置的负载信息在至少两个路由装置选择作为迁移源的路由装置和作为迁移目的的路由装置;

[0063] 选取所述作为迁移源的路由装置的终端IP列表中的终端的IP地址作为迁移IP地址,并将所述迁移IP地址通知给所述作为迁移目的的路由装置。

[0064] 此外,为了上述传统技术中的路由方式的存在的三角路由问题,提供了一种网关调度装置。

[0065] 本发明实施例第四方面公开了一种网关调度装置,所述装置包括终端IP获取模块、与终端IP获取模块连接的路由装置选择模块以及与所述路由装置选择模块连接的网关分配模块,其中:

[0066] 终端IP获取模块,用于接收终端添加指令,所述终端添加指令包含终端的IP地址;

[0067] 路由装置选择模块,用于在至少两个路由装置中选择任意一个作为所述终端归属的网关;

[0068] 网关分配模块,用于将所述终端的IP地址添加到所述终端归属的网关的终端IP列表中,所述终端归属的网关通过ARP响应报文告知所述终端该路由装置的MAC地址。

[0069] 在本发明实施例第四方面的第一种可能的实现方式中,所述终端IP获取模块还用于接收广播的ARP请求报文,获取所述广播的ARP请求报文的源IP地址;在所述源IP地址不属于所述至少两个路由装置各自的终端IP列表时,在至少两个路由装置中选择任意一个作为所述源IP地址对应的终端归属的网关。

[0070] 结合本发明实施例第五方面和第四方面的第一种可能的实现方式,在本发明实施例第五方面的第二种可能的实现方式中,所述装置还包括负载检测模块,用于检测所述至少两个路由装置的负载信息;

[0071] 所述路由装置选择模块还用于根据检测到的所述负载信息在所述至少两个路由装置中选择路由装置终端的IP地址。

[0072] 结合本发明实施例第五方面的第二种可能的实现方式,在本发明实施例第五方面的第三种可能的实现方式中,所述装置还包括负载均衡模块,用于根据检测到的路由装置的负载信息在至少两个路由装置选择作为迁移源的路由装置和作为迁移目的的路由装置;选取所述作为迁移源的路由装置的终端IP列表中的终端的IP地址作为迁移IP地址,并将所述迁移IP地址通知给所述作为迁移目的的路由装置。

[0073] 实施本发明实施例,将具有如下有益效果:

[0074] 也就是说,上述路由方法、装置及系统、网关调度方法及装置可根据各个路由装置上的负载情况为终端选择合适的路由装置作为该终端的网关、创建或接入新的路由装置、

在现有的路由装置之间迁移数据转发业务,使得数据转发业务不需要集中由某一个路由装置集中处理,而可分散在多个路由装置上,由该多个路由装置进行负载均衡,从而避免了传统技术中出现的单点性能瓶颈的问题。

[0075] 同时,由于路由装置相对独立,若现有的路由装置负载均较大需要进行扩容时,也只需要添加新的路由装置,再由网关调度装置为其分配新接入的终端的IP地址或者向其迁入其他路由装置上的数据转发业务,即可实现扩容。因此,上述网关调度方法使得在对路由系统进行扩容的操作较方便。

[0076] 相应的,采用上述网关调度方法之后,终端被指定特定的路由转发装置作为该终端的网关,该终端发起的上行数据报文以及外部网络发送至终端的下行数据报文均由该固定的路由装置进行转发,该路由装置固定成为该终端与外部网络之间通信的网关设备。因此在进行SNAT/NAT操作时,SNAT的地址转换记录或NAT操作对应EIP则可在同一个路由装置上顺利找到,因此有效地避免了三角路由的情况。

附图说明

[0077] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0078] 其中:

[0079] 图1为传统技术中集中式路由方式的逻辑组网图;

[0080] 图2为本发明一个实施例中路由系统的逻辑组网图;

[0081] 图3为本发明一个实施例中基于多个实体路由装置的实际组网图;

[0082] 图4为本发明一个实施例中基于多网卡服务器或服务器集群的内部组网图;

[0083] 图5为本发明一个实施例中路由系统进行网关调度过程的流程图;

[0084] 图6为本发明一个实施例中体现多种存储终端IP列表方式的示意图;

[0085] 图7为地址解析协议报文的结构示意图;

[0086] 图8为SNAT转换过程的示意图;

[0087] 图9为基于EIP的NAT转换过程的示意图;

[0088] 图10为体现传统技术中的三角路由问题的示意图;

[0089] 图11为传统技术中存在三角路由问题的等价路由组网方式的示意图;

[0090] 图12为本发明一个实施例中一种网关调度方法的流程图;

[0091] 图13为本发明一个实施例中一种网关调度装置的结构示意图;

[0092] 图14为本发明一个实施例中一种路由方法的流程图;

[0093] 图15为本发明一个实施例中一种路由装置的结构示意图。

具体实施方式

[0094] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他

实施例,都属于本发明保护的范围。

[0095] 为解决上述传统技术中集中式的路由方式的不利于扩容的技术问题,在一个实施例中,特提出了一种路由系统,如图2所示,图2展示了一个实施例中实现上述路由系统的逻辑结构图。

[0096] 该路由转发系统由两部分构成:负责为以太网中的终端选择路由装置作为网关的一种网关调度装置,以及作为网关负责数据转发的至少两个路由装置。该至少两个路由装置可以是路由器集群、设置有多个网卡具有网关功能的计算机设备或者提供云服务的服务器集群等。

[0097] 需要说明的是,图2中的终端、路由装置以及网关调度装置可以是实体的路由器设备,也可以是具有多个网卡的计算机设备通过虚拟化技术生成的虚拟机对象。

[0098] 例如,在一个应用场景中,如图3所示,路由装置A、B和C、网关调度装置和终端1至5构成了以太局域网。图2中的路由装置A、B和C则可以是图3中的实体的路由器A、B和C;终端1至5则可以是实体的个人电脑、web服务器等硬件装置;网关调度装置则可以是图4中独立设置的网关调度装置设备,也可以基于实体的路由器(即该实体的路由器既作为路由装置进行数据转发,也作为该以太网的网关调度装置)。网关调度装置

[0099] 再例如,在另一个应用场景中,如图4所示,上述终端、网关调度装置以及至少两个路由装置为多网卡服务器或者服务器集群上的虚拟机。路由装置A、B和C可以是多网卡服务器通过虚拟化技术虚拟的分别基于网卡A、B和C的虚拟路由器(在服务器集群中则为网卡分别为A、B和C服务器上通过虚拟化技术虚拟的虚拟路由器,在该应用场景中,路由装置可以以多网卡服务器或服务器集群上的计算机程序执行时生成的对象或结构体的数据形式存在,但本质上数据转发任务仍由该多网卡服务器或服务器集群执行),终端1至5则可以是多网卡服务器或服务器集群上虚拟化的虚拟机1至5。网关调度装置即为多网卡服务器或服务器集群上虚拟化的网关调度装置。

[0100] 具体的,在本实施例中,该路由转发系统包括包括网关调度装置10和至少两个路由装置20,其中:

[0101] 网关调度装置10用于接收终端添加指令,所述终端添加指令包含终端的互联网协议IP地址;在至少两个路由装置中选择任意一个作为所述终端归属的网关,将该终端的IP地址添加到该终端归属的网关的终端IP列表中。

[0102] 该终端归属的网关用于向所述终端发送地址解析协议ARP响应报文,该ARP响应报文用于告知所述终端绑定该终端归属的网关。

[0103] 在本实施例中,如图3所示,以太局域网的管理人员在将图3中的多个路由器和终端组网成以太局域网,或者在向该以太局域网中添加新的终端时,可向网关调度装置输入终端添加指令,并输入终端的IP地址作为该终端添加指令的参数。而在图4对应的虚拟机的应用场景中,承载虚拟机的服务器可在创建了作为终端的虚拟机之后,生成终端添加指令,并将在创建虚拟机时为其分配的终端的IP地址作为终端添加指令的参数,发送给实体或者虚拟化的网关调度装置。

[0104] 如图2所示,以太局域网中的路由装置的数量可以是两个或两个以上,且均与网关调度装置连接,由网关调度装置管理。网关调度装置在至少两个路由装置中选择路由装置,将该终端的IP地址添加到选择的路由装置的终端IP列表中之后,选择的路由装置则仅对该

路由装置的终端IP列表中的终端的IP地址对应的终端返回ARP响应请求,并作为该终端的网关。例如,该至少两个路由装置可具有相同的以太网IP地址(也可不同)。在前述的组建以太网的过程中,每个终端上均可设置网关的IP地址为该以太网IP地址,但终端上缓存的与网关的IP地址对应的MAC地址(英文:Media Access Control,简称:MAC)可能不同,终端中存储的网关IP地址对应的路由装置的MAC地址由网关调度装置分配。也就是说,每个终端上可设置相同的网关IP地址,但该网关IP地址对应的MAC地址可因终端而异,且通过网关调度装置指定。

[0105] 在一个实施例中,网关调度装置可随机选择路由装置与通过终端添加指令输入的终端的IP地址建立映射关系,建立映射关系的过程即为将该终端的IP地址添加到与该分配的路由装置对应的终端IP列表中。

[0106] 终端IP列表用于存储路由装置与终端的IP地址的映射关系,可以存在各种数据结构上的表现形式,例如,列表、数组、键值表等。终端IP列表中存储的终端的IP地址即为相应的路由装置需要对其进行地址解析协议(英文:Address Resolution Protocol,简称:ARP)报文响应以及数据报文转发的终端或虚拟机的IP地址。终端IP列表可存储于网关调度装置中,也可存储于各个路由装置中,也可同时存储于网关调度装置和路由装置中。

[0107] 例如,如图6所示,若路由装置A、B和C的MAC地址分别为MAC_A、MAC_B和MAC_C,组网时以太网中共有6台终端设备,他们的IP地址分别为IP_1至IP_6,且IP_1和IP_2的网关被选为路由装置A,IP_3和IP_4的网关被选为路由装置B,IP_5和IP_6的网关被选为路由装置C,则在一个实施例中,在路由装置A、B和C中可分别存储终端IP列表,如图6中所示,路由装置A中存储的终端IP列表则包含IP_1和IP_2,路由装置B中存储的终端IP列表则包含IP_3和IP_4,路由装置C中存储的终端IP列表则包含IP_5和IP_6。

[0108] 而在另一个实施例中,终端IP列表可存储在网关调度装置中。如图6所示,网关调度装置中存储的终端IP列表可以是映射表,记录有各个路由装置的MAC地址以及与其对应的终端的IP地址。路由装置可通过查询网关调度装置得到与其自身对应的终端IP列表。

[0109] 也就是说,对于IP_1和IP_2的终端或虚拟机而言,MAC_B和MAC_C对应的路由装置由于没有返回ARP响应报文,因此MAC_B和MAC_C的路由装置为不可见状态。虽然路由装置B和C与IP_1和IP_2的终端或虚拟机存在物理上的链路连接,但由于IP_1和IP_2的终端或虚拟机无法获知MAC_B和MAC_C,因此,路由装置B和C与IP_1和IP_2的终端或虚拟机并不在一个逻辑上的以太网中,从而使得IP_1和IP_2的终端或虚拟机的数据转发业务与路由装置B和C无关。同时,由于路由装置A、B和C仅对各自的终端IP列表中的终端的IP地址对应的终端返回ARP响应请求,使得路由装置A、B和C互相无法发现对方,因此不会存在以太网中IP地址冲突的问题。

[0110] 在另一个实施例中,可不需要接收终端添加指令,网关调度装置还可用于在至少两个路由装置中选择路由装置之前,接收广播的ARP请求报文,获取所述广播的ARP请求报文的源IP地址;在所述源IP地址不属于所述至少两个路由装置各自的终端IP列表时,在至少两个路由装置中选择任意一个作为该源IP地址对应的终端归属的网关,将广播的ARP请求报文的源IP地址添加到该路由装置的终端IP列表中。

[0111] 如图3所示,参考图3中的网关调度装置与以太网虚线连接。若有新的终端加入该以太网或者以太网中现有的某个终端更新了IP地址,则该终端即会广播

ARP请求报文,且在该ARP请求报文的源IP地址(即发送该ARP请求的终端的IP地址)数据段即为该终端在以太局域网中的终端的IP地址。网关调度装置可通过前述的虚线连接接收到以太网中终端广播的ARP报文。

[0112] 如图6所示,网关调度装置中可存储各个路由装置的MAC地址及其对应的终端IP列表,网关调度装置可持续接收以太局域网中广播的ARP报文,并对ARP报文进行解析,得到其源IP地址,然后在存储的各个路由装置的MAC地址对应的终端IP列表中查找提取的源IP地址。若查找到,则表示已经为该源IP地址分配了路由装置,该源IP地址已与某个路由装置建立了映射关系,也就是说,以太局域网中已存在该源IP地址对应的终端的网关设备;若未查找到,则表示没有给该源IP地址对应的终端设置该终端在该以太局域网中的网关,从而可继续为终端的IP地址选择作为网关的路由装置。

[0113] 具体的,所述终端归属的网关具体用于在向所述终端发送ARP响应报文之前,接收所述终端发送的询问网关的ARP请求报文,所述ARP请求报文携带所述终端的IP地址;判断该终端归属的网关的终端IP列表中是否包含所述终端的IP地址,在所述终端归属的路由装置的终端IP列表包含所述终端的IP地址时,获取该终端归属的网关的MAC地址,通过ARP响应报文将该MAC地址返回该终端。终端即可将该MAC地址对应的路由装置设置为网关。

[0114] 如前所述,如图3所示,以太网中的终端配置了网关IP地址之后,则会向该网关IP地址发起ARP请求,该ARP请求报文用于询问网关MAC地址。对于路由装置A,可接收以太局域网中终端1至5发起的ARP请求报文,ARP请求报文的数据格式可参考图7所示,其中发送者IP地址即为ARP请求报文的源IP地址。路由装置A接收到某个ARP请求报文之后,可根据ARP协议定义的数据格式提取该源IP地址,然后再与路由装置A对应的终端IP列表中查找该源IP地址,若查找到,则该终端IP列表包含获取到的ARP请求报文的源IP地址,可返回相应的ARP响应报文。且在返回ARP响应报文时,需要在先获取路由装置A自身的MAC地址(如在图3对应的应用场景中,返回路由器自身的MAC地址,而在图4对应的虚拟机的应用场景中,则返回虚拟路由器绑定的网卡的MAC地址),然后将该MAC地址添加到ARP响应报文的头部,即图7中的“以太网源地址”数据段处。

[0115] 而该ARP请求报文的源IP地址对应的终端或虚拟机在接收到该ARP响应报文后,则可获取到路由装置A的MAC地址并将其设置为网关MAC地址,该终端后续向外部网络发送的数据报文则将向该网关MAC地址对应的路由装置A进行转发。

[0116] 若未查找到,则路由装置A的终端IP列表不包含获取到的ARP请求报文的源IP地址,路由装置A不响应该ARP请求报文而将其丢弃,该ARP请求报文的源IP地址对应的终端或虚拟机不会将该路由装置A作为网关。而同样的ARP请求报文也同时会被路由装置B和路由装置C接收到,若路由装置B在其对应的终端IP列表中查找到了该ARP请求报文的源IP地址,则可返回以路由装置B的MAC地址作为以太网源地址的ARP响应报文,从而使得路由装置B成为该ARP请求报文的源IP地址对应的终端或虚拟机的网关。

[0117] 需要说明的是,路由装置在判断的接收到的ARP请求报文携带的终端的IP地址是否存在于与该路由装置对应的终端IP列表中时,若路由装置上存储有终端IP列表,则可在其自身存储的终端IP列表中查找(例如图4对应的多个实体装置构成的分布式路由系统的应用场景中,每个路由器上均存储有与其对应的终端IP列表);若终端IP列表存储在如图6所示的网关调度装置中,则路由装置可通过向网关调度装置查询来判断终端IP列表中是否

包含ARP请求报文携带的终端的IP地址(例如图5对应的提供云服务的多网卡服务器或服务器集群的应用场景,路由装置可与网关调度装置处于同一硬件装置上,可通过共享内存方式进行终端IP列表的共享,路由装置通过访问共享的内存即可进行终端IP列表的查询)。

[0118] 也就是说,该网关调度系统在进行网关调度时的执行过程可如图5所示:

[0119] 步骤S102:网关调度装置接收终端添加指令,所述终端添加指令包含终端的互联网协议IP地址;在该至少两个路由装置中选择任意一个作为所述终端归属的网关,将所述终端的IP地址添加到该终端归属的网关的终端IP列表中。

[0120] 步骤S104:该终端归属的网关用于向所述终端发送地址解析协议ARP响应报文,所述ARP响应报文用于告知所述终端绑定该终端归属的网关。该终端归属的网关还用于接收所述终端发送的上行数据报文,提取所述上行数据报文中携带的网关MAC地址;在所述携带的网关MAC地址与该终端归属的网关的MAC地址相同时,转发所述上行数据报文。

[0121] 上行数据报文即为以太网局域网中的终端向外部网络发送的数据报文。相应的,下行数据报文即为外部网络向以太网局域网中的终端发送的数据报文。例如,用户在以太网局域网中的终端打开浏览器应用访问架设在互联网上的web应用时,用户点击链接产生的网页访问请求即会产生上行数据报文。而用户点击链接后互联网上的web应用向该终端下发网页响应则会产生下行数据报文。

[0122] 如上例中,若终端1接收到了路由装置A返回的ARP响应报文,则将路由装置A的MAC地址作为网关MAC地址,并将上行数据报文发送至该网关MAC地址处请求转发。路由装置A接收到该上行数据报文之后,获取上行数据报文的网关MAC地址是否为其自身路由装置A,若是,则对该上行数据报文进行转发操作;否则,将该上行数据报文丢弃。

[0123] 也就是说,路由装置仅作为该路由装置的终端IP列表中存在的终端的IP地址对应的终端的网关,转发该终端的上行数据报文和下行数据报文。不在该路由装置的终端IP列表中的终端的IP地址对应的终端由于无法获知该路由装置的MAC地址,因此无法通过该路由装置转发数据。

[0124] 进一步的,在一个实施例中,网关调度装置还可为路由装置设置外网IP地址。路由装置还可以用于在转发上行数据报文之前,根据该路由装置的预设的外网IP地址对该上行数据报文进行源地址转换SNAT。

[0125] 源地址转换(英文:Source Network Address Translation,简称:SNAT)即为以太网局域网中的终端向外部网络访问时,发起访问的以太网局域网中的IP地址转换为指定的IP地址(可指定具体的服务以及相应的端口或端口范围),这可以使内网中使用保留IP地址的主机访问外部网络,即以太网局域网中的多个终端可以通过一个有效的外网IP地址访问外部网络。

[0126] 也就是说,路由装置在转发上行数据报文之前还可获取预设的外网IP地址,根据该外网IP地址对所述上行数据报文进行源地址转换,生成相应的地址转换记录。

[0127] 路由装置还可用于在转发上行数据报文之后,还可接收目的IP地址为该外网IP地址的下行数据报文,获取所述下行数据报文的端口;查找与获取到的端口对应的地址转换记录,根据所述查找到的地址转换记录对所述下行数据报文进行目的地址转换DNAT(英文:Destination Network Address Translation,简称:DNAT),转发所述DNAT后的下行数据报文。

[0128] 例如,同时参考图3和图8,网关调度装置可为路由装置A、B和C分别配置不同的外网IP地址。若为路由装置A分配的外网IP地址为202.1.1.1,以路由装置A为网关的终端的IP地址为10.1.1.1,则当用户在该终端上打开浏览器网页中的链接时,将发起网络访问请求产生上行数据报文,且该上行数据报文的源IP即为10.1.1.1,源端口即为80(浏览器通常以80端口发起网页访问请求)。该上行数据报文由路由装置A接收到后,路由装置先判断该上行数据报文的网关MAC地址是否为路由装置A自身的MAC地址,若否,则丢弃该上行数据报文;若是,则先为10.1.1.1:80的组合分配555的映射端口,然后将该上行数据报文的源IP地址替换为外网IP地址202.1.1.1,将源端口替换为555,然后向外部网络转发该数据报文。SNAT过程结束后,还会在地址转换记录中记录10.1.1.1:80与202.1.1.1:555的映射关系。互联网上的web服务器接收到该网页访问请求的上行数据报文后,提取到的源IP地址即为202.1.1.1,源端口即为555。Web服务器返回网页访问请求的响应数据(例如网页、图片等),该响应数据的下行数据报文的目IP即为202.1.1.1,目的端口即为555。由于路由装置A绑定了外网IP地址202.1.1.1,因此可接收到该下行数据报文,通过查询前述地址转换记录中的10.1.1.1:80与202.1.1.1:555的映射关系,则可将接收到的下行数据报文的目IP地址替换为10.1.1.1,目的端口替换为80。通过查询ARP地址表即可得到10.1.1.1对应的终端或虚拟机的MAC地址,然后通过以太网局域网络将该下行数据报文转发至该MAC地址。

[0129] 在一个实施例中,进一步的,网关调度装置还可接收EIP地址(英文:ElasticIP,简称:EIP)申请指令,获取EIP地址申请指令对应的终端的IP地址和路由装置,为该路由装置分配EIP地址;然后通过EIP地址绑定指令将该终端的IP地址和分配的EIP发送给该路由装置。

[0130] 弹性IP地址(英文:Elastic IP,简称:EIP),指云服务系统中,与用户账户关联的而不是与云服务器关联的,用户完全独占的互联网公共IP地址。如在上述图5所示的应用场景中,云服务器可为用户提供建站服务,用户可在云服务器上创建虚拟机来创建web应用,但云服务器为虚拟机分配的私有IP地址为随机分配。若用户希望将其购买的域名与其创建的虚拟机进行绑定,则需要其虚拟机具有独立的互联网公共IP地址。此时,用户可向云服务器申请为其虚拟机分配EIP地址。

[0131] 网关调度装置可根据申请EIP的虚拟机的终端的IP地址查找已为该终端的IP地址分配的作为其网关的路由装置。然后将该终端的IP地址和EIP通过EIP绑定指令发送给该路由装置并由其缓存。路由装置即可根据该EIP对该虚拟机的数据流量进行地址转换(英文:Network Address Translation,简称:NAT)转换。

[0132] 也就是说,路由装置接收EIP地址绑定指令,获取EIP地址绑定指令中包含的EIP地址以及与所述EIP地址为映射关系的终端的IP地址;接收目的地址为所述EIP地址的下行数据报文,将该下行数据报文的目IP地址设置为所述EIP地址对应的终端的IP地址,并转发所述下行数据报文至所述终端的IP地址对应的终端。

[0133] 例如,同时参考图4和图8,用户操作虚拟机1向网关调度装置申请EIP之后,网关调度装置则可为IP地址为10.1.1.1的虚拟机1的分配EIP为202.1.1.1,若虚拟机1的网关被分配为路由转发虚拟机A,则网关调度装置通知路由转发虚拟机A存储10.1.1.1与202.1.1.1的映射关系。

[0134] 当用户在该终端上打开浏览器网页中的链接时,将发起网络访问请求产生上行数

据报文,且该上行数据报文的源IP即为10.1.1.1。该上行数据报文由路由装置A接收到后,路由装置先判断该上行数据报文的网关MAC地址是否为路由装置A自身的MAC地址,若否,则丢弃该上行数据报文;若是,则将该上行数据报文的源IP地址替换为EIP地址202.1.1.1,然后向外部网络转发该上行数据报文。

[0135] 而在本应用场景中,若该虚拟机1作为web服务器使用,则用户可将其为web应用绑定的域名(英文:Domain Name System,简称:DNS)先在域名服务提供商处映射到EIP地址202.1.1.1处,其他用户访问该web服务器发出的数据请求对于该路由转发虚拟机A产生的下行数据报文即先由路由转发虚拟机A接收,该下行数据报文的目的IP地址即为202.1.1.1,路由转发虚拟机A通过查询前述映射即可得到EIP地址202.1.1.1映射的虚拟机的终端的IP地址为10.1.1.1。将下行数据报文的目的IP地址替换为10.1.1.1后,通过查询ARP地址表即可得到10.1.1.1对应的终端或虚拟机的MAC地址,然后通过以太局域网将该下行数据报文转发至该MAC地址。

[0136] 因此,采用上述网关调度方法,可为每个路由装置设置不同的外网IP地址,每个路由装置可根据分配给自己的外网IP地址独立的进行SNAT转换的操作。以太局域网中的终端或虚拟机也可以申请EIP,从而可使终端或虚拟机的终端的IP地址能够被NAT为独立的互联网公共IP地址。也就是说,上述网关调度方法在进行SNAT/NAT操作时,不会出现三角路由的情况。

[0137] 一个业务流通常同时包含了上行数据报文和下行数据报文的流量。当上行数据报文和下行数据报文的转发路径不一致时,则出现三角路由。如图10所示,若用户在终端上点击链接生成的网页访问请求的上行数据报文通过路由装置A转发至互联网并在转发之前进行了SNAT/NAT操作,则该SNAT的地址转换记录或NAT操作对应EIP存储在路由装置A中。而若返回的网页响应的下行数据报文若不经由路由装置A转发而经由路由装置B转发至该终端,则路由装置B在接收到该下行数据报文之后,无法找到相应的SNAT的地址转换记录或NAT操作对应EIP,使得该下行数据报文无法抵达该终端。

[0138] 传统技术中,为解决上述单点性能瓶颈问题引入的等价路由(英文:Equal Cost Multi Path,简称:ECMP)的方式则存在上述三角路由的情况。路由表中,存在多条目的地址相同、权重相同,但是下一跳不同的路由,这些路由都称为等价路由。如图11所示,路由器A、B、C为等价路由,终端发出的上行数据报文被随机分配到某个路由器进行转发,外部网络返回的下行数据报文也被随机分配到某个路由器进行转发,从而使得转发业务的负载得到均衡。若需要在路由器A、B、C上实现SNAT/NAT操作,则会出现三角路由的情况导致数据报文丢失。

[0139] 而采用上述网关调度方法之后,终端1发起的上行数据报文以及外部网络发送至终端1的下行数据报文均由同一个路由装置进行转发,该路由装置固定成为该终端与外部网络之间通信的网关设备。因此在进行SNAT/NAT操作时,SNAT的地址转换记录或NAT操作对应EIP则可在同一个路由装置上顺利找到,因此有效地避免了三角路由的情况。

[0140] 可选的,网关调度装置还可用于对各个路由装置的负载信息进行检测。路由装置的负载信息可包括:流量信息、带宽使用比例、处理器使用率和/或连接数等。

[0141] 在本实施例中,网关调度装置还用于根据检测到的负载信息在至少两个路由装置中选择任意一个作为所述终端归属的网关。

[0142] 例如,在图5对应的应用场景中,若有用户新创建了虚拟机,则需要为该虚拟机分配路由装置作为其网关。若路由装置A、B、C三者的负载信息中路由装置B的负载最小,则可将路由装置B分配给该虚拟机作为该虚拟机的网关。或者,若路由装置A的负载较大,而路由装置B和C的负载较小均小于负载阈值,则可在路由装置B和C中随机选择作为虚拟机的网关。

[0143] 可选的,网关调度装置还可用于在检测路由装置的负载信息之后根据检测到的负载信息创建路由装置。例如,若路由装置A、B和C的负载均大于或等于负载阈值,则可创建新的路由装置D(可以是创建新的虚拟机也可以是新添加一台路由器设备)。

[0144] 在本实施例中,网关调度装置还可根据路由装置的负载信息调整路由装置的负载,也就是说为已有的终端重新指定路由装置作为网关。具体为:

[0145] 根据检测到的路由装置的负载信息在至少两个路由装置选择作为迁移源的路由装置和作为迁移目的的路由装置;选取作为迁移源的路由装置的终端IP列表中的终端的IP地址作为迁移IP地址,并将该迁移IP地址通知给作为迁移目的的路由装置。

[0146] 可选的,根据检测到的路由装置的负载信息在至少两个路由装置选择作为迁移源的路由装置和作为迁移目的的路由装置可具体为:

[0147] 查找其负载信息大于或等于负载阈值的路由装置作为迁移源的路由装置,查找负载信息最小的路由装置作为迁移目的的路由装置。

[0148] 而作为迁移目的的路由装置则可向迁移IP地址对应的终端发送免费ARP报文,通知所述迁移IP地址对应的终端更新网关MAC地址。

[0149] 也就是说,参考图2所示,若路由装置A的负载较高,其终端列表中包含终端的IP地址:IP_1和IP_2,路由装置B的负载较低。则可将IP_2由路由装置A的终端IP列表中移除,然后添加到路由装置B的终端IP列表中,然后由路由装置B向IP_2发送免费ARP报文。IP_2接收到该免费ARP报文后,则可将网关MAC地址更新为路由装置B的MAC地址。

[0150] 免费ARP报文是一种特殊的ARP报文,该报文中携带的发送端IP地址和目标IP地址都是本机IP地址,报文源MAC地址是本机MAC地址,报文的目的MAC地址是广播地址。通常为设备在改变了MAC地址时,通过发送免费ARP报文通知其它设备更新ARP表项。终端接收到该免费ARP报文后,则可更新网关MAC地址为该路由装置B的MAC地址。

[0151] 网关调度装置选取作为迁移源的路由装置的终端IP列表中的终端的IP地址作为迁移IP地址之后还可将所述迁移IP地址通知作为迁移源的路由装置;作为迁移源的路由装置则将迁移IP地址由该路由装置的终端IP列表中移除。也就是说,迁移IP地址对应的终端在接收到免费ARP报文后更新了网关MAC地址,不会再向作为迁移源的路由装置发送数据报文,且该迁移IP地址对应的终端广播的ARP请求报文不会得到作为迁移源的路由装置的ARP响应。

[0152] 进一步的,网关调度装置将迁移IP地址通知所述作为迁移源的路由装置之后,所述方法还包括:作为迁移源的路由装置查找与迁移IP地址对应的EIP地址,将查找到的EIP地址移除。

[0153] 如前例中,若终端的IP地址为10.1.1.1,其在路由装置A上绑定的EIP为202.1.1.1,则若10.1.1.1被迁移到路由装置B上时,需要路由装置A移除该绑定的EIP:202.1.1.1,使得路由装置A不在接收外部网络发送至202.1.1.1的下行数据报文。用户可申

请在路由装置B上重新绑定该EIP:202.1.1.1,也可由网关调度装置自动在作为目的路由装置的路由装置B上重新绑定该EIP:202.1.1.1与IP地址:10.1.1.1的对应关系。

[0154] 也就是说,上述网关调度方法可根据各个路由装置上的负载情况分配终端的IP地址、创建或接入新的路由装置、在现有的路由装置之间迁移数据转发业务,使得数据转发业务不需要集中由某一个路由装置集中处理,而可分散在多个路由装置上,由该多个路由装置进行负载均衡,从而避免了传统技术中出现的单点性能瓶颈的问题。

[0155] 同时,由于路由装置相对独立,若现有的路由装置负载均较大需要进行扩容时,也只需要添加新的路由装置,再由网关调度装置为其分配新接入的终端的IP地址或者向其迁入其他路由装置上的数据转发业务,即可实现扩容。因此,上述网关调度方法使得在对路由系统进行扩容的操作较方便。

[0156] 为解决上述传统技术中集中式的路由方式的不利于扩容的技术问题,在一个实施例中,为解决上述传统技术中集中式的路由方式的不利于扩容的技术问题,在一个实施例中,对应于前述的路由转发方法和系统,还提供了一种网关调度方法,该网关调度方法可由前述图3所示的网关调度装置执行。

[0157] 具体的,如图12所示,该网关调度方法可具体包括:

[0158] 步骤S202:接收终端添加指令,所述终端添加指令包含终端的IP地址。

[0159] 在本实施例中,可在组网时或在以太网局域网中加入了新的终端时执行上述步骤S102。例如,在图4中,以太网局域网的管理人员在将图4中的多个路由器和终端组网成以太网局域网,或者在向该以太网局域网中添加新的终端时,可向网关调度装置输入终端添加指令,并输入终端的IP地址作为该终端添加指令的参数。而在图5对应的虚拟机的应用场景中,承载虚拟机的服务器可在创建了作为终端的虚拟机之后,生成终端添加指令,并将在创建虚拟机时为其分配的终端的IP地址作为终端添加指令的参数,发送给实体或者虚拟化的网关调度装置。

[0160] 步骤S204:在至少两个路由装置中选择任意一个作为所述终端归属的网关。

[0161] 步骤S206:将该终端的IP地址添加到该终端归属的网关的终端IP列表中,所述该终端归属的网关通过ARP响应报文告知所述终端该路由装置的MAC地址。

[0162] 如图2所示,以太网局域网中的路由装置的数量可以是两个或两个以上,且均与网关调度装置连接,由网关调度装置管理。网关调度装置在至少两个路由装置中选择路由装置,将该终端的IP地址添加到选择的路由装置的终端IP列表中之后,选择的路由装置则仅对该路由装置的终端IP列表中的终端的IP地址对应的终端返回ARP响应请求,并作为该终端的网关。例如,该至少两个路由装置可具有相同的以太网IP地址(也可不同)。在前述的组建以太网的过程中,每个终端上均可设置网关的IP地址为该以太网IP地址,但终端上缓存的与网关的IP地址对应的MAC地址(英文:Media Access Control,简称:MAC)可能不同,终端中存储的网关IP地址对应的路由装置的MAC地址由网关调度装置分配。也就是说,每个终端上可设置相同的网关IP地址,但该网关IP地址对应的MAC地址可因终端而异,且通过网关调度装置指定。

[0163] 在一个实施例中,网关调度装置可随机选择路由装置与通过终端添加指令输入的终端的IP地址建立映射关系,建立映射关系的过程即为将该终端的IP地址添加到与该分配的路由装置对应的终端IP列表中。

[0164] 终端IP列表用于存储路由装置与终端的IP地址的映射关系,可以存在各种数据结构上的表现形式,例如,列表、数组、键值表等。终端IP列表中存储的终端的IP地址即为相应的路由装置需要对其进行地址解析协议(英文:Address Resolution Protocol,简称:ARP)报文响应以及数据报文转发的终端或虚拟机的IP地址。终端IP列表可存储于网关调度装置中,也可存储于各个路由装置中,也可同时存储于网关调度装置和路由装置中。

[0165] 例如,如图6所示,若路由装置A、B和C的MAC地址分别为MAC_A、MAC_B和MAC_C,组网时以太网局域网中共有6台终端设备,他们的IP地址分别为IP_1至IP_6,且IP_1和IP_2的网关被选为路由装置A,IP_3和IP_4的网关被选为路由装置B,IP_5和IP_6的网关被选为路由装置C,则在一个实施例中,在路由装置A、B和C中可分别存储终端IP列表,如图6中所示,路由装置A中存储的终端IP列表则包含IP_1和IP_2,路由装置B中存储的终端IP列表则包含IP_3和IP_4,路由装置C中存储的终端IP列表则包含IP_5和IP_6。

[0166] 而在另一个实施例中,终端IP列表可存储在网关调度装置中。如图6所示,网关调度装置中存储的终端IP列表可以是映射表,记录有各个路由装置的MAC地址以及与其对应的终端的IP地址。路由装置可通过查询网关调度装置得到与其自身对应的终端IP列表。

[0167] 也就是说,对于IP_1和IP_2的终端或虚拟机而言,MAC_B和MAC_C对应的路由装置由于没有返回ARP响应报文,因此MAC_B和MAC_C的路由装置为不可见状态。虽然路由装置B和C与IP_1和IP_2的终端或虚拟机存在物理上的链路连接,但由于IP_1和IP_2的终端或虚拟机无法获知MAC_B和MAC_C,因此,路由装置B和C与IP_1和IP_2的终端或虚拟机并不在一个逻辑上的以太网局域网中,从而使得IP_1和IP_2的终端或虚拟机的数据转发业务与路由装置B和C无关。同时,由于路由装置A、B和C仅对各自的终端IP列表中的终端的IP地址对应的终端返回ARP响应请求,使得路由装置A、B和C互相无法发现对方,因此不会存在以太网中IP地址冲突的问题。

[0168] 在另一个实施例中,可不需要接收终端添加指令,网关调度装置在至少两个路由装置中选择路由装置之前,还可接收广播的ARP请求报文,获取广播的ARP请求报文的源IP地址;在获取的源IP地址不属于至少两个路由装置各自的终端IP列表时,在至少两个路由装置中选择任意一个作为该源IP地址对应的终端归属的网关,将该终端的IP地址添加到选择的路由装置的终端IP列表中。

[0169] 如图3所示,参考图3中的网关调度装置与以太网局域网的虚线连接。若有新的终端加入该以太网局域网或者以太网局域网中现有的某个终端更新了IP地址,则该终端即会广播ARP请求报文,且在该ARP请求报文的源IP地址(即发起该ARP请求报文的终端的IP地址)数据段即为该终端在以太网局域网中的终端的IP地址。网关调度装置可通过前述的虚线连接接收到以太网中终端广播的ARP报文。

[0170] 如图6所示,网关调度装置中可存储各个路由装置的MAC地址及其对应的终端IP列表,网关调度装置可持续接收以太网局域网中广播的ARP报文,并对ARP报文进行解析,得到其源IP地址,然后在存储的各个路由装置的MAC地址对应的终端IP列表中查找提取的源IP地址。若查找到,则表示已经为该源IP地址分配了路由装置,该源IP地址已与某个路由装置建立了映射关系,也就是说,以太网局域网中已存在该源IP地址对应的终端的网关设备;若未查找到,则表示没有给该源IP地址对应的终端设置该终端在该以太网局域网中的网关,从而可继续执行为终端的IP地址分配路由装置的步骤,为该终端的IP地址设置网关。

[0171] 路由装置接收所述终端发送的询问网关的ARP请求报文,所述ARP请求报文携带所述终端的IP地址;判断该路由装置的终端IP列表中是否包含所述终端的IP地址,在该终端IP列表包含所述终端的IP地址时,获取该路由装置的MAC地址,通过ARP响应报文将所述MAC地址返回所述终端。

[0172] 如前所述,如图3所示,以太网中的终端配置了网关IP地址之后,则会向该网关IP地址发起ARP请求,该ARP请求报文用于询问网关MAC地址。对于路由装置A,可接收以太局域网中终端1至5发起的ARP请求报文,ARP请求报文的数据格式可参考图7所示,其中发送者IP地址即为ARP请求报文的源IP地址。路由装置A接收到某个ARP请求报文之后,可根据ARP协议定义的数据格式提取该源IP地址,然后再与路由装置A对应的终端IP列表中查找该源IP地址,若查找到,则该终端IP列表包含获取到的ARP请求报文的源IP地址,可返回相应的ARP响应报文。且在返回ARP响应报文时,需要在先获取路由装置A自身的MAC地址(如在图3对应的应用场景中,返回路由器自身的MAC地址,而在图4对应的虚拟机的应用场景中,则返回虚拟路由器绑定的网卡的MAC地址),然后将该MAC地址添加到ARP响应报文的头部,即图7中的“以太网源地址”数据段处。

[0173] 而该ARP请求报文的源IP地址对应的终端或虚拟机在接收到该ARP响应报文后,则可获取到路由装置A的MAC地址并将其设置为网关MAC地址,该终端后续向外部网络发送的数据报文则将向该网关MAC地址对应的路由装置A进行转发。

[0174] 若未查找到,则路由装置A的终端IP列表不包含获取到的ARP请求报文的源IP地址(即发送该ARP请求报文的终端的IP地址),路由装置A不响应该ARP请求报文而将其丢弃,该ARP请求报文的源IP地址对应的终端或虚拟机不会将该路由装置A作为网关。而同样的ARP请求报文也同时会被路由装置B和路由装置C接收到,若路由装置B在其对应的终端IP列表中查找到了该ARP请求报文的源IP地址,则可返回以路由装置B的MAC地址作为以太网源地址的ARP响应报文,从而使得路由装置B成为该ARP请求报文的源IP地址对应的终端或虚拟机的网关。

[0175] 需要说明的是,路由装置在判断的接收到的ARP请求报文携带的终端的IP地址是否存在于与该路由装置对应的终端IP列表中时,若路由装置上存储有终端IP列表,则可在其自身存储的终端IP列表中查找(例如如图4对应的多个实体装置构成的分布式路由系统的应用场景中,每个路由器上均存储有与其对应的终端IP列表);若终端IP列表存储在如图6所示的网关调度装置中,则路由装置可通过向网关调度装置查询来判断终端IP列表中是否包含ARP请求报文的源IP地址(例如如图5对应的提供云服务的多网卡服务器或服务器集群的应用场景,路由装置可与网关调度装置处于同一硬件装置上,可通过共享内存方式进行终端IP列表的共享,路由装置通过访问共享的内存即可进行终端IP列表的查询)。

[0176] 可选的,还可根据路由装置的负载信息调整路由装置的负载,也就是说为已有的终端重新指定路由装置作为网关。也就是说,可在路由装置A上的终端IP列表中选择终端的IP地址生成路由服务添加指令,发送给路由装置B,路由装置B则执行:

[0177] 接收路由服务添加指令,获取路由服务添加指令对应的终端的IP地址;终端的IP地址发送免费地址解析协议报文,该终端的IP地址对应的终端根据所述免费地址解析协议报文更新网关MAC地址。

[0178] 而对于路由装置A,则可根据在路由装置A上的终端IP列表中选择终端的IP地址向

路由装置A其发送路由服务移除指令,路由装置A则执行:

[0179] 接收路由服务移除指令,获取路由服务移除指令对应的终端的IP地址;判断所述终端的IP地址是否对应所述缓存的EIP地址,若是,则移除所述EIP地址。

[0180] 例如,参考图2所示,若路由装置A的负载较高,其终端列表中包含终端的IP地址:IP_1和IP_2,路由装置B的负载较低。则可将IP_2由路由装置A的终端IP列表中移除,然后添加到路由装置B的终端IP列表中,然后由路由装置B向IP_2发送免费ARP报文。IP_2接收到该免费ARP报文后,则可将网关MAC地址更新为路由装置B的MAC地址。

[0181] 免费ARP报文是一种特殊的ARP报文,该报文中携带的发送端IP地址和目标IP地址都是本机IP地址,报文源MAC地址是本机MAC地址,报文的目的地MAC地址是广播地址。通常为设备在改变了MAC地址时,通过发送免费ARP报文通知其它设备更新ARP表项。终端接收到该免费ARP报文后,则可更新网关MAC地址为该路由装置B的MAC地址。

[0182] 而前例中,若终端的IP地址为10.1.1.1,其在路由装置A上绑定的EIP为202.1.1.1,则若10.1.1.1被迁移到路由装置B上时,需要路由装置A移除该绑定的EIP:202.1.1.1,使得路由装置A不在接收外部网络发送至202.1.1.1的下行数据报文。用户可申请在路由装置B上重新绑定该EIP:202.1.1.1,也可由网关调度装置自动在作为目的的路由装置的路由装置B上重新绑定该EIP:202.1.1.1与IP地址:10.1.1.1的对应关系。

[0183] 在一个实施例中,为解决上述传统技术中集中式的路由方式的不利于扩容的技术问题,特提出了一种与前述网关调度方法对应的网关调度装置,如图13所示,包括终端IP获取模块102、与终端IP获取模块102连接的路由装置选择模块104和与路由装置选择模块连接的网关分配模块106,其中:

[0184] 终端IP获取模块102,用于接收终端添加指令,所述终端添加指令包含终端的IP地址。

[0185] 路由装置选择模块104,用于在至少两个路由装置中选择任意一个作为所述终端归属的网关。

[0186] 网关分配模块106,用于将所述终端的IP地址添加到所述该终端归属的网关的终端IP列表中,该终端归属的网关通过ARP响应报文告知所述终端该路由装置的MAC地址。

[0187] 在本实施例中,终端IP获取模块102还用于接收广播的ARP请求报文,获取所述广播的ARP请求报文的源IP地址;在所述源IP地址不属于所述至少两个路由装置各自的终端IP列表时,在至少两个路由装置中选择任意一个作为该源IP地址对应的终端归属的网关。

[0188] 在本实施例中,如图13所示,网关调度装置还包括负载检测模块108,用于检测所述至少两个路由装置的负载信息。

[0189] 路由装置选择模块104还用于根据检测到的所述负载信息在所述至少两个路由装置中选择路由装置终端的IP地址。

[0190] 在本实施例中,如图13所示,网关调度装置还包括负载均衡模块110,用于根据检测到的路由装置的负载信息在至少两个路由装置选择作为迁移源的路由装置和作为迁移目的的路由装置;选取所述作为迁移源的路由装置的终端IP列表中的终端的IP地址作为迁移IP地址,并将所述迁移IP地址通知给所述作为迁移目的的路由装置。

[0191] 为解决上述传统技术中集中式的路由方式的不利于扩容的技术问题,在一个实施例中,为解决上述传统技术中集中式的路由方式的不利于扩容的技术问题,在一个实施例

中,对应于前述的路由转发方法和系统,还提供了一种路由方法,该路由方法可由前述图3所示的路由装置执行。

[0192] 具体的,如图14所示,该路由方法可具体包括:

[0193] 步骤S302,接收询问网关的地址解析协议ARP请求报文,获取所述ARP请求报文携带的终端的IP地址。

[0194] 步骤S304,获取预设的终端IP列表,在所述终端IP列表包含所述获取到终端的IP地址时,通过ARP响应报文将自身的MAC地址返回所述终端。

[0195] 在本实施例中,如图4中,以太局域网的管理人员在将图4中的多个路由器和终端组网成以太局域网,或者在向该以太局域网中添加新的终端时,可向网关调度装置输入终端添加指令,并输入终端的IP地址作为该终端添加指令的参数。而在图5对应的虚拟机的应用场景中,承载虚拟机的服务器可在创建了作为终端的虚拟机之后,生成终端添加指令,并将在创建虚拟机时为其分配的终端的IP地址作为终端添加指令的参数,发送给实体或者虚拟化的网关调度装置。

[0196] 如图2所示,以太局域网中的路由装置的数量可以是两个或两个以上,且均与网关调度装置连接,由网关调度装置管理。网关调度装置在至少两个路由装置中选择路由装置,将该终端的IP地址添加到选择的路由装置的终端IP列表中之后,选择的路由装置则仅对该路由装置的终端IP列表中的终端的IP地址对应的终端返回ARP响应请求,并作为该终端的网关。例如,该至少两个路由装置可具有相同的以太网IP地址(也可不同)。在前述的组建以太网的过程中,每个终端上均可设置网关的IP地址为该以太网IP地址,但终端上缓存的与网关的IP地址对应的MAC地址(英文:Media Access Control,简称:MAC)可能不同,终端中存储的网关IP地址对应的路由装置的MAC地址由网关调度装置分配。也就是说,每个终端上可设置相同的网关IP地址,但该网关IP地址对应的MAC地址可因终端而异,且通过网关调度装置指定。

[0197] 终端IP列表用于存储路由装置与终端的IP地址的映射关系,可以存在各种数据结构上的表现形式,例如,列表、数组、键值表等。终端IP列表中存储的终端的IP地址即为相应的路由装置需要对其进行数据报文转发的终端或虚拟机的IP地址。终端IP列表可存储于网关调度装置中,也可存储于各个路由装置中,也可同时存储于网关调度装置和路由装置中。

[0198] 例如,如图6所示,若路由装置A、B和C的MAC地址(英文:Media Access Control,简称:MAC)分别为MAC_A、MAC_B和MAC_C,组网时以太局域网中共有6台终端设备,他们的IP地址分别为IP_1至IP_6,且IP_1和IP_2的网关被选为路由装置A,IP_3和IP_4的网关被选为路由装置B,IP_5和IP_6的网关被选为路由装置C,则在一个实施例中,在路由装置A、B和C中可分别存储终端IP列表,如图6中所示,路由装置A中存储的终端IP列表则包含IP_1和IP_2,路由装置B中存储的终端IP列表则包含IP_3和IP_4,路由装置C中存储的终端IP列表则包含IP_5和IP_6。

[0199] 而在另一个实施例中,终端IP列表可存储在网关调度装置中。如图6所示,网关调度装置中存储的终端IP列表可以是映射表,记录有各个路由装置的MAC地址以及与其对应的终端的IP地址。路由装置可通过查询网关调度装置得到与其自身对应的终端IP列表。

[0200] 也就是说,对于IP_1和IP_2的终端或虚拟机而言,MAC_B和MAC_C对应的路由装置由于没有返回ARP响应报文,因此MAC_B和MAC_C的路由装置为不可见状态。虽然路由装置B

和C与IP_1和IP_2的终端或虚拟机存在物理上的链路连接,但由于IP_1和IP_2的终端或虚拟机无法获知MAC_B和MAC_C,因此,路由装置B和C与IP_1和IP_2的终端或虚拟机并不在一个逻辑上的以太局域网中,从而使得IP_1和IP_2的终端或虚拟机的数据转发业务与路由装置B和C无关。

[0201] 如图2中,对于路由装置A,可接收以太网中终端1至5发起的ARP请求报文,ARP请求报文的数据格式可参考图7所示,其中发送者IP地址即为ARP请求报文的源IP地址。路由装置A接收到某个ARP请求报文之后,可根据ARP协议定义的数据格式提取该源IP地址,然后再与路由装置A对应的终端IP列表中查找该源IP地址,若查找到,则该终端IP列表包含获取到的ARP请求报文的源IP地址,可返回相应的ARP响应报文。且在返回ARP响应报文时,需要在先获取路由装置A自身的MAC地址(如在图3对应的应用场景中,返回路由器自身的MAC地址,而在图4对应的虚拟机的应用场景中,则返回虚拟路由器绑定的网卡的MAC地址),然后将该MAC地址添加到ARP响应报文的头部,即图7中的“以太网源地址”数据段处。

[0202] 而该ARP请求报文的源IP地址对应的终端或虚拟机在接收到该ARP响应报文后,则可获取到路由装置A的MAC地址并将其设置为网关MAC地址,该终端后续向外部网络发送的数据报文则将向该网关MAC地址对应的路由装置A发送。

[0203] 若未查找到,则该终端IP列表包含获取到的ARP请求报文的源IP地址,路由装置A可不响应该ARP请求报文而将其丢弃,该ARP请求报文的源IP地址对应的终端或虚拟机不会将该路由装置A作为网关。而同样的ARP请求报文也同时会被路由装置B和路由装置C接收到,若路由装置B在其对应的终端IP列表中查找到了该ARP请求报文的源IP地址,则可返回以路由装置B的MAC地址作为以太网源地址的ARP响应报文,从而使得路由装置B称为该ARP请求报文的源IP地址对应的终端或虚拟机的网关。

[0204] 需要说明的是,路由装置在判断的接收到的ARP请求报文的源IP地址是否存在于与该路由装置对应的终端IP列表中时,若路由装置上存储有终端IP列表,则可在其自身存储的终端IP列表中查找(例如图3对应的多个实体装置构成的分布式路由系统的应用场景中,每个路由器上均存储有与其对应的终端IP列表);若终端IP列表存储在如图6所示的网关调度装置中,则路由装置可通过向网关调度装置查询来判断终端IP列表中是否包含ARP请求报文的源IP地址(例如图4对应的提供云服务的多网卡服务器或服务器集群的应用场景,路由装置可与网关调度装置处于同一硬件装置上,可通过共享内存方式进行终端IP列表的共享,路由装置通过访问共享的内存即可进行终端IP列表的查询)。

[0205] 步骤S306,接收上行数据报文,提取所述上行数据报文中携带的网关MAC地址。

[0206] 步骤S308,在所述上行数据报文中携带的网关MAC地址与自身的MAC地址相同时,转发所述上行数据报文。

[0207] 上行数据报文即为以太网中的终端向外部网络发送的数据报文。相应的,下行数据报文即为外部网络向以太网中的终端发送的数据报文。例如,用户在以太网中的终端打开浏览器应用访问架设在互联网上的web应用时,用户点击链接产生的网页访问请求即会产生上行数据报文。而用户点击链接后互联网上的web应用向该终端下发网页响应则会产生下行数据报文。

[0208] 如上例中,若终端1接收到了路由装置A返回的ARP响应报文,则将路由装置A的MAC地址作为网关MAC地址,并将上行数据报文发送至该网关MAC地址处请求转发。路由装置A接

收到该上行数据报文之后,获取上行数据报文的网关MAC地址是否为其自身路由装置A,若是,则对该上行数据报文进行转发操作;否则,将该上行数据报文丢弃。

[0209] 也就是说,路由装置仅作为该路由装置的终端IP列表中存在的终端的IP地址对应的终端的网关,转发该终端的上行数据报文和下行数据报文。不在该路由装置的终端IP列表中的终端的IP地址对应的终端由于无法获知该路由装置的MAC地址,因此无法通过该路由装置转发数据。

[0210] 进一步的,在一个实施例中,网关调度装置还可为路由装置设置外网IP地址。路由装置转发上行数据报文的步骤之前还可根据该路由装置的外网IP地址对该上行数据报文进行源地址转换SNAT。

[0211] 源地址转换(英文:Source Network Address Translation,简称:SNAT)即为以太网局域网中的终端向外部网络访问时,发起访问的以太网局域网中的IP地址转换为指定的IP地址(可指定具体的服务以及相应的端口或端口范围),这可以使内网中使用保留IP地址的主机访问外部网络,即以太网局域网中的多个终端可以通过一个有效的外网IP地址访问外部网络。

[0212] 也就是说,路由装置在转发上行数据报文之前还可获取预设的外网IP地址,根据该外网IP地址对所述上行数据报文进行源地址转换,生成相应的地址转换记录。

[0213] 路由装置在转发上行数据报文之后,还可接收目的IP地址为该外网IP地址的下行数据报文,获取所述下行数据报文的端口;查找与获取到的端口对应的地址转换记录,根据所述查找到的地址转换记录对所述下行数据报文进行目的地址转换DNAT(英文:Destination Network Address Translation,简称:DNAT),转发所述DNAT后的下行数据报文。

[0214] 例如,同时参考图3和图8,网关调度装置可为路由装置A、B和C分别配置不同的外网IP地址。若为路由装置A分配的外网IP地址为202.1.1.1,以路由装置A为网关的终端的IP地址为10.1.1.1,则当用户在该终端上打开浏览器网页中的链接时,将发起网络访问请求产生上行数据报文,且该上行数据报文的源IP即为10.1.1.1,源端口即为80(浏览器通常以80端口发起网页访问请求)。该上行数据报文由路由装置A接收到后,路由装置先判断该上行数据报文的网关MAC地址是否为路由装置A自身的MAC地址,若否,则丢弃该上行数据报文;若是,则先为10.1.1.1:80的组合分配555的映射端口,然后将该上行数据报文的源IP地址替换为外网IP地址202.1.1.1,将源端口替换为555,然后向外部网络转发该数据报文。SNAT过程结束后,还会在地址转换记录中记录10.1.1.1:80与202.1.1.1:555的映射关系。互联网上的web服务器接收到该网页访问请求的上行数据报文后,提取到的源IP地址即为202.1.1.1,源端口即为555。Web服务器返回网页访问请求的响应数据(例如网页、图片等),该响应数据的下行数据报文的端口IP即为202.1.1.1,目的端口即为555。由于路由装置A绑定了外网IP地址202.1.1.1,因此可接收到该下行数据报文,通过查询前述地址转换记录中的10.1.1.1:80与202.1.1.1:555的映射关系,则可将接收到的下行数据报文的端口IP地址替换为10.1.1.1,目的端口替换为80。通过查询ARP地址表即可得到10.1.1.1对应的终端或虚拟机的MAC地址,然后通过以太网局域网将该下行数据报文转发至该MAC地址。

[0215] 在一个实施例中,进一步的,网关调度装置还可接收EIP地址(英文:Elastic IP,简称:EIP)申请指令,获取EIP地址申请指令对应的终端的IP地址和路由装置,为该路由装

置分配EIP地址;然后通过EIP地址绑定指令将该终端的IP地址和分配的EIP发送给该路由装置。

[0216] 弹性IP地址(英文:Elastic IP,简称:EIP),指云服务系统中,与用户账户关联的而不是与云服务器关联的,用户完全独占的互联网公共IP地址。如在上述图5所示的应用场景中,云服务器可为用户提供建站服务,用户可在云服务器上创建虚拟机来创建web应用,但云服务器为虚拟机分配的私有IP地址为随机分配。若用户希望将其购买的域名与其创建的虚拟机进行绑定,则需要其虚拟机具有独立的互联网公共IP地址。此时,用户可向云服务器申请为其虚拟机分配EIP地址。

[0217] 网关调度装置可根据申请EIP的虚拟机的终端的IP地址查找已为该终端的IP地址分配的作为其网关的路由装置。然后将该终端的IP地址和EIP通过EIP绑定指令发送给该路由装置并由其缓存。路由装置即可根据该EIP对该虚拟机的数据流量进行地址转换(英文:Network Address Translation,简称:NAT)转换。

[0218] 也就是说,路由装置接收EIP地址绑定指令,获取EIP地址绑定指令中包含的EIP地址以及与所述EIP地址为映射关系的终端的IP地址;接收目的地址为所述EIP地址的下行数据报文,将该下行数据报文的源IP地址设置为所述EIP地址对应的终端的IP地址,并转发所述下行数据报文至所述终端的IP地址对应的终端。

[0219] 例如,同时参考图4和图8,用户操作虚拟机1向网关调度装置申请EIP之后,网关调度装置则可为IP地址为10.1.1.1的虚拟机1的分配EIP为202.1.1.1,若虚拟机1的网关被分配为路由转发虚拟机A,则网关调度装置通知路由转发虚拟机A存储10.1.1.1与202.1.1.1的映射关系。

[0220] 当用户在该终端上打开浏览器网页中的链接时,将发起网络访问请求产生上行数据报文,且该上行数据报文的源IP即为10.1.1.1。该上行数据报文由路由装置A接收到后,路由装置先判断该上行数据报文的网关MAC地址是否为路由装置A自身的MAC地址,若否,则丢弃该上行数据报文;若是,则将该上行数据报文的源IP地址替换为EIP地址202.1.1.1,然后向外部网络转发该上行数据报文。

[0221] 而在本应用场景中,若该虚拟机1作为web服务器使用,则用户可将其为web应用绑定的域名(英文:Domain Name System,简称:DNS)先在域名服务提供商处映射到EIP地址202.1.1.1处,其他用户访问该web服务器发出的数据请求对于该路由转发虚拟机A产生的下行数据报文即先由路由转发虚拟机A接收,该下行数据报文的源IP地址即为202.1.1.1,路由转发虚拟机A通过查询前述映射即可得到EIP地址202.1.1.1映射的虚拟机的终端的IP地址为10.1.1.1。将下行数据报文的源IP地址替换为10.1.1.1后,通过查询ARP地址表即可得到10.1.1.1对应的终端或虚拟机的MAC地址,然后通过以太网局域网将该下行数据报文转发至该MAC地址。

[0222] 因此,采用上述网关调度方法,可为每个路由装置设置不同的外网IP地址,每个路由装置可根据分配给自己的外网IP地址独立的进行SNAT转换的操作。以太网局域中的终端或虚拟机也可以申请EIP,从而可使终端或虚拟机的终端的IP地址能够被NAT为独立的互联网公共IP地址。也就是说,上述网关调度方法在进行SNAT/NAT操作时,不会出现三角路由的情况。

[0223] 一个业务流通常同时包含了上行数据报文和下行数据报文的流量。当上行数据报

文和下行数据报文的转发路径不一致时,则出现三角路由。如图10所示,若用户在终端上点击链接生成的网页访问请求的上行数据报文通过路由装置A转发至互联网并在转发之前进行了SNAT/NAT操作,则该SNAT的地址转换记录或NAT操作对应EIP存储在路由装置A中。而若返回的网页响应的下行数据报文若不经由路由装置A转发而经由路由装置B转发至该终端,则路由装置B在接收到该下行数据报文之后,无法找到相应的SNAT的地址转换记录或NAT操作对应EIP,使得该下行数据报文无法抵达该终端。

[0224] 传统技术中,为解决上述单点性能瓶颈问题引入的等价路由(英文:Equal Cost Multi Path,简称:ECMP)的方式则存在上述三角路由的情况。路由表中,存在多条目的地址相同、权重相同,但是下一跳不同的路由,这些路由都称为等价路由。如图11所示,路由器A、B、C为等价路由,终端发出的上行数据报文被随机分配到某个路由器进行转发,外部网络返回的下行数据报文也被随机分配到某个路由器进行转发,从而使得转发业务的负载得到均衡。若需要在路由器A、B、C上实现SNAT/NAT操作,则会出现三角路由的情况导致数据报文丢失。

[0225] 而采用上述网关调度方法之后,终端1发起的上行数据报文以及外部网络发送至终端1的下行数据报文均由同一个路由装置进行转发,该路由装置固定成为该终端与外部网络之间通信的网关设备。因此在进行SNAT/NAT操作时,SNAT的地址转换记录或NAT操作对应EIP则可在同一个路由装置上顺利找到,因此有效地避免了三角路由的情况。

[0226] 可选的,网关调度装置还可对各个路由装置的负载信息进行检测。路由装置的负载信息可包括:流量信息、带宽使用比例、处理器使用率和/或连接数等。

[0227] 在本实施例中,网关调度装置为终端的IP地址分配路由装置的步骤还包括:根据检测到的负载信息在至少两个路由装置中选择路由装置终端的IP地址。

[0228] 例如,在图5对应的应用场景中,若有用户新创建了虚拟机,则需要为该虚拟机分配路由装置作为其网关。若路由装置A、B、C三者的负载信息中路由装置B的负载最小,则可将路由装置B分配给该虚拟机作为该虚拟机的网关。或者,若路由装置A的负载较大,而路由装置B和C的负载较小均小于负载阈值,则可在路由装置B和C中随机选择作为虚拟机的网关。

[0229] 可选的,在检测路由装置的负载信息的步骤之后还可根据检测到的负载信息创建路由装置。例如,若路由装置A、B和C的负载均大于或等于负载阈值,则可创建新的路由装置D(可以是创建新的虚拟机也可以是新添加一台路由器设备)。

[0230] 在本实施例中,网关调度装置还可根据路由装置的负载信息调整路由装置的负载,也就是说为已有的终端重新指定路由装置作为网关。具体为:

[0231] 根据检测到的路由装置的负载信息在至少两个路由装置选择作为迁移源的路由装置和作为迁移目的的路由装置;选取作为迁移源的路由装置的终端IP列表中的终端的IP地址作为迁移IP地址,并将该迁移IP地址通知给作为迁移目的的路由装置。

[0232] 可选的,还可根据路由装置的负载信息调整路由装置的负载,也就是说为已有的终端重新指定路由装置作为网关。也就是说,可在路由装置A上的终端IP列表中选择终端的IP地址生成路由服务添加指令,发送给路由装置B,路由装置B则执行:

[0233] 接收路由服务添加指令,获取路由服务添加指令对应的终端的IP地址;向该终端发送免费ARP报文,所述终端根据所述免费ARP报文更新网关MAC地址。

[0234] 而对于路由装置A,则可根据在路由装置A上的终端IP列表中选择终端的IP地址向路由装置A其发送路由服务移除指令,路由装置A则执行:

[0235] 接收路由服务移除指令,获取路由服务移除指令对应的IP地址;查找该IP地址对应的EIP地址,移除查找到的EIP地址。

[0236] 例如,参考图2所示,若路由装置A的负载较高,其终端列表中包含终端的IP地址:IP_1和IP_2,路由装置B的负载较低。则可将IP_2由路由装置A的终端IP列表中移除,然后添加到路由装置B的终端IP列表中,然后由路由装置B向IP_2发送免费ARP报文。IP_2接收到该免费ARP报文后,则可将网关MAC地址更新为路由装置B的MAC地址。

[0237] 免费ARP报文是一种特殊的ARP报文,该报文中携带的发送端IP地址和目标IP地址都是本机IP地址,报文源MAC地址是本机MAC地址,报文的目的地MAC地址是广播地址。通常为设备在改变了MAC地址时,通过发送免费ARP报文通知其它设备更新ARP表项。终端接收到该免费ARP报文后,则可更新网关MAC地址为该路由装置B的MAC地址。

[0238] 而前例中,若终端的IP地址为10.1.1.1,其在路由装置A上绑定的EIP为202.1.1.1,则若10.1.1.1被迁移到路由装置B上时,需要路由装置A移除该绑定的EIP:202.1.1.1,使得路由装置A不在接收外部网络发送至202.1.1.1的下行数据报文。用户可申请在路由装置B上重新绑定该EIP:202.1.1.1,也可由网关调度装置自动在作为目的路由装置的路由装置B上重新绑定该EIP:202.1.1.1与IP地址:10.1.1.1的对应关系。

[0239] 也就是说,上述网关调度方法可根据各个路由装置上的负载情况分配终端的IP地址、创建或接入新的路由装置、在现有的路由装置之间迁移数据转发业务,使得数据转发业务不需要集中由某一个路由装置集中处理,而可分散在多个路由装置上,由该多个路由装置进行负载均衡,从而避免了传统技术中出现的单点性能瓶颈的问题。

[0240] 同时,由于路由装置相对独立,若现有的路由装置负载均较大需要进行扩容时,也只需要添加新的路由装置,再由网关调度装置为其分配新接入的终端的IP地址或者向其迁入其他路由装置上的数据转发业务,即可实现扩容。因此,上述网关调度方法使得在对路由系统进行扩容的操作较方便。

[0241] 在一个实施例中,为解决上述传统技术中集中式的路由方式的不利于扩容的技术问题,特提出了一种与前述路由方法对应的路由装置,如图15所示,包括:ARP报文接收模块202、与ARP报文接收模块202连接的MAC地址返回模块204、数据报文接收模块206以及与数据报文接收模块206连接的数据报文转发模块208,其中:

[0242] ARP报文接收模块202,用于接收询问网关的地址解析协议ARP请求报文,获取所述ARP请求报文携带的终端的IP地址。

[0243] MAC地址返回模块204,用于获取预设的终端IP列表,在所述终端IP列表包含所述获取到终端的IP地址时,通过ARP响应报文将自身的MAC地址返回所述终端。

[0244] 数据报文接收模块206,用于接收上行数据报文,提取所述上行数据报文中携带的网关MAC地址。

[0245] 数据报文转发模块208,用于在所述上行数据报文中携带的网关MAC地址与自身的MAC地址相同时,转发所述上行数据报文。

[0246] 在本实施例中,数据报文转发模块208还用于根据预设的外网IP地址对所述上行数据报文进行源地址转换SNAT,生成相应的地址转换记录。

[0247] 在本实施例中,数据报文转发模块208还用于接收其目的IP地址为所述外网IP地址的下行数据报文,获取所述下行数据报文的目的端口;查找与所述目的端口对应的地址转换记录,根据所述查找到的地址转换记录对所述下行数据报文进行目的地址转换DNAT,转发所述DNAT后的下行数据报文。

[0248] 在本实施例中,如图15所示,路由装置还包括EIP地址绑定模块210,用于接收弹性互联网协议EIP地址绑定指令,获取EIP地址绑定指令中包含的EIP地址以及与所述EIP地址为映射关系的终端的IP地址;

[0249] 数据报文转发模块208还用于接收目的地址为所述EIP地址的下行数据报文,将所述下行数据报文的目的IP地址设置为所述EIP地址对应的终端的IP地址,并转发所述下行数据报文至与所述EIP地址对应的终端的IP地址。

[0250] 在本实施例中,如图15所示,路由装置还包括路由服务移除模块212,用于接收路由服务移除指令,获取路由服务移除指令对应的终端的IP地址;查找所述终端的IP地址对应的EIP地址,将查找到的EIP地址移除EIP。

[0251] 在本实施例中,如图15所示,路由装置还包括路由服务添加模块214,用于接收路由服务添加指令,获取路由服务添加指令对应的终端的IP地址;向所述终端发送免费地址解析协议报文,所述终端根据所述免费地址解析协议报文更新网关MAC地址。

[0252] 综上所述,实施本发明实施例,将具有如下有益效果:

[0253] 也就是说,上述路由方法、装置及系统、网关调度方法及装置可根据各个路由装置上的负载情况为终端选择合适的路由装置作为该终端的网关、创建或接入新的路由装置、在现有的路由装置之间迁移数据转发业务,使得数据转发业务不需要集中由某一个路由装置集中处理,而可分散在多个路由装置上,由该多个路由装置进行负载均衡,从而避免了传统技术中出现的单点性能瓶颈的问题。

[0254] 同时,由于路由装置相对独立,若现有的路由装置负载均较大需要进行扩容时,也只需要添加新的路由装置,再由网关调度装置为其分配新接入的终端的IP地址或者向其迁入其他路由装置上的数据转发业务,即可实现扩容。因此,上述网关调度方法使得在对路由系统进行扩容的操作较方便。

[0255] 相应的,采用上述网关调度方法之后,终端被指定特定的路由转发装置作为该终端的网关,该终端发起的上行数据报文以及外部网络发送至终端的下行数据报文均由该固定的路由装置进行转发,该路由装置固定成为该终端与外部网络之间通信的网关设备。因此在进行SNAT/NAT操作时,SNAT的地址转换记录或NAT操作对应EIP则可在同一个路由装置上顺利找到,因此有效地避免了三角路由的情况。

[0256] 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程,是可以通过计算机程序来指令相关的硬件来完成,所述的程序可存储于一计算机可读取存储介质中,该程序在执行时,可包括如上述各方法的实施例的流程。其中,所述的存储介质可为磁碟、光盘、只读存储记忆体(Read-Only Memory,ROM)或随机存储记忆体(Random Access Memory,RAM)等。

[0257] 以上所揭露的仅为本发明较佳实施例而已,当然不能以此来限定本发明之权利范围,因此依本发明权利要求所作的等同变化,仍属本发明所涵盖的范围。

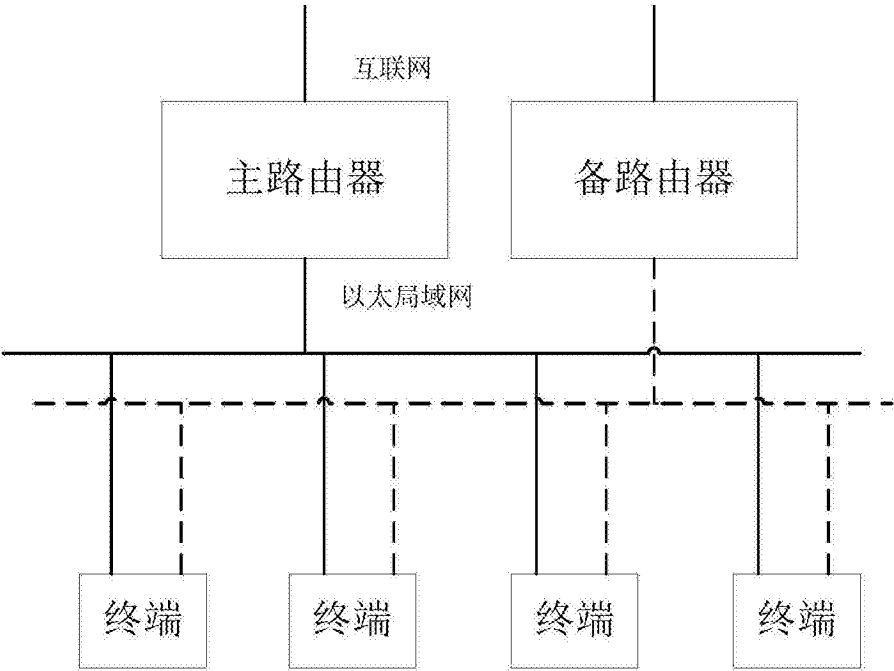


图1

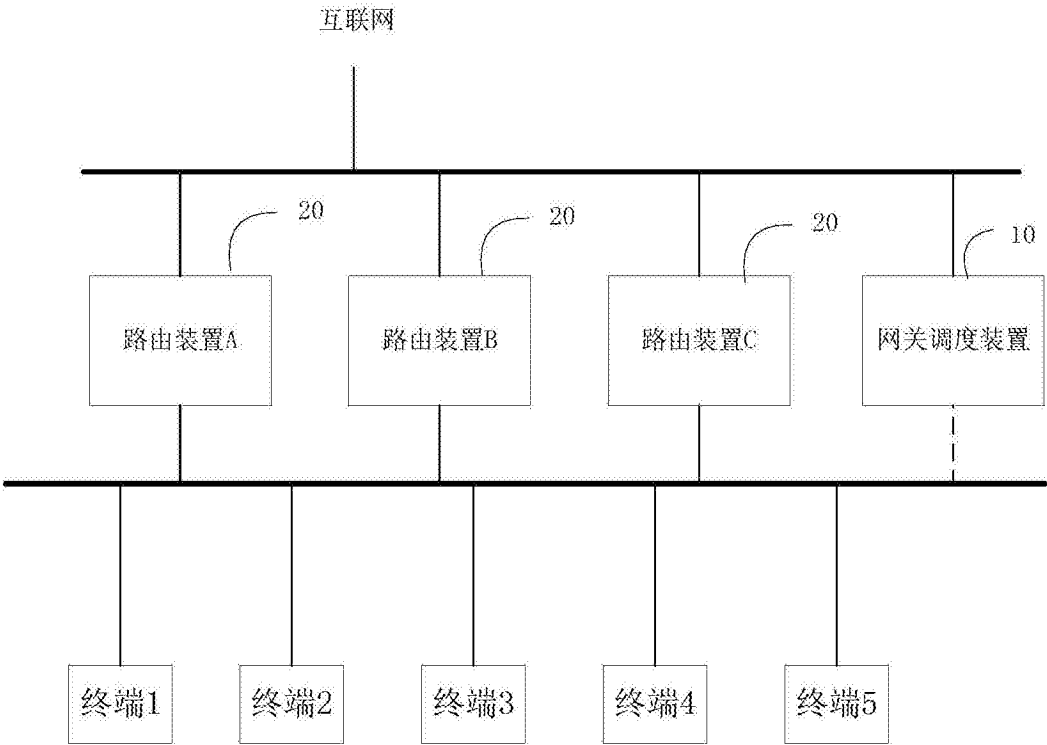


图2

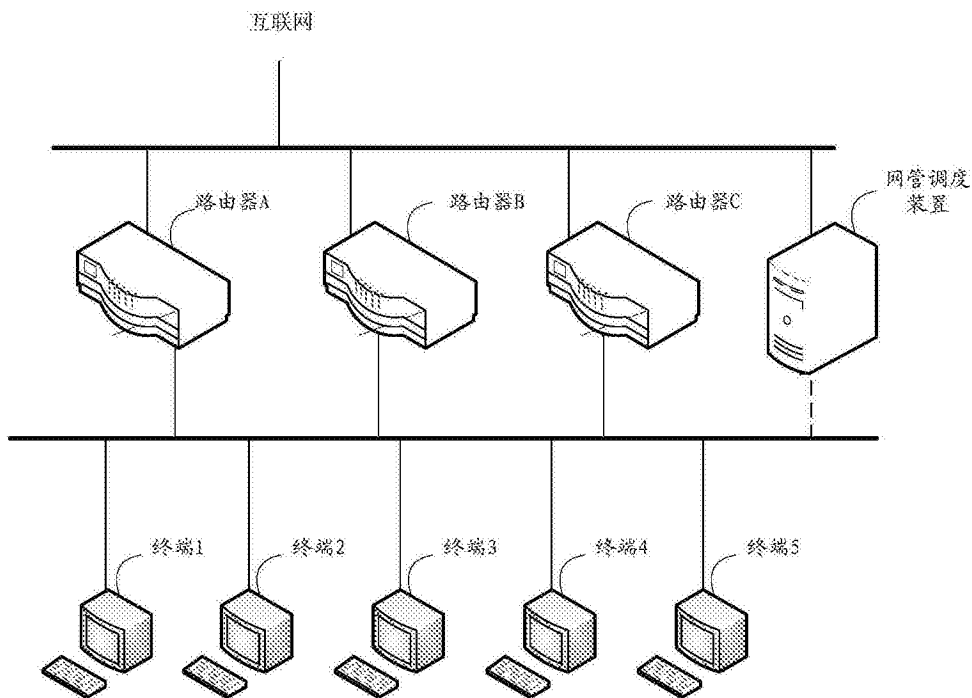


图3

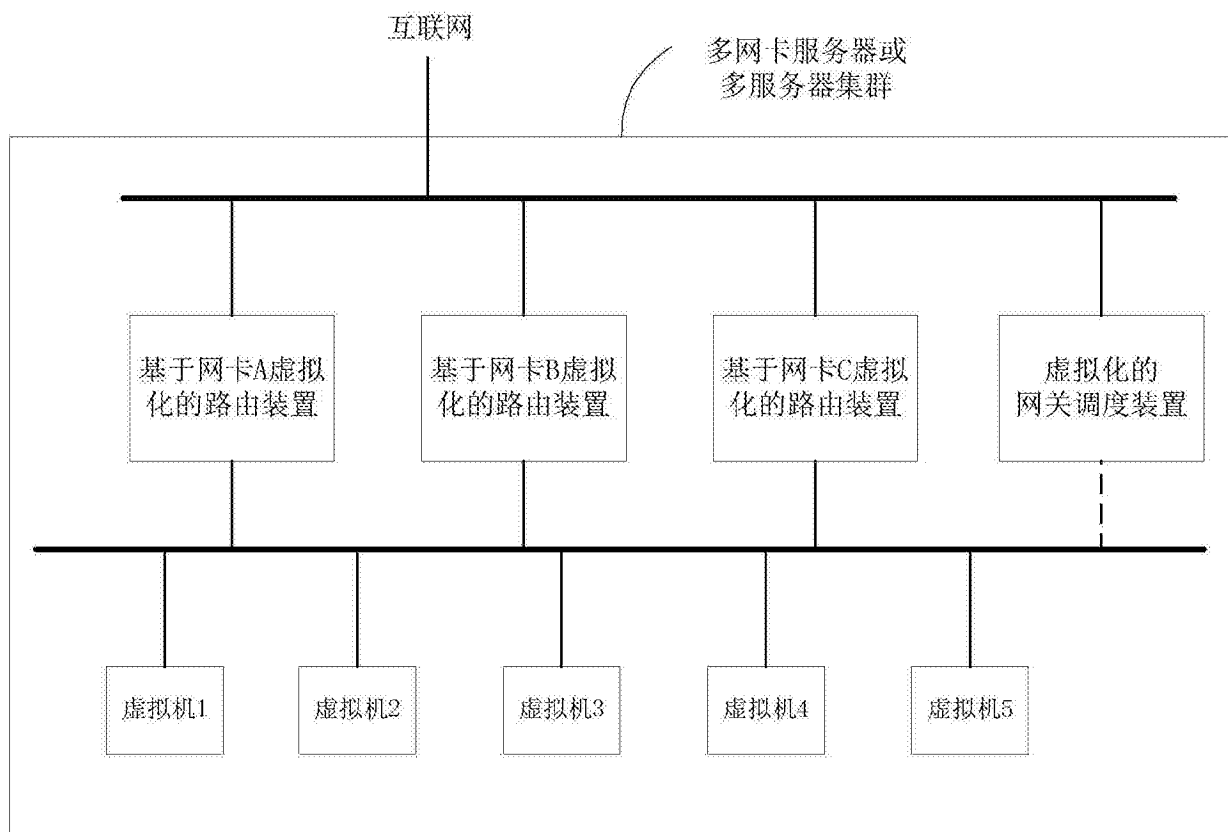


图4

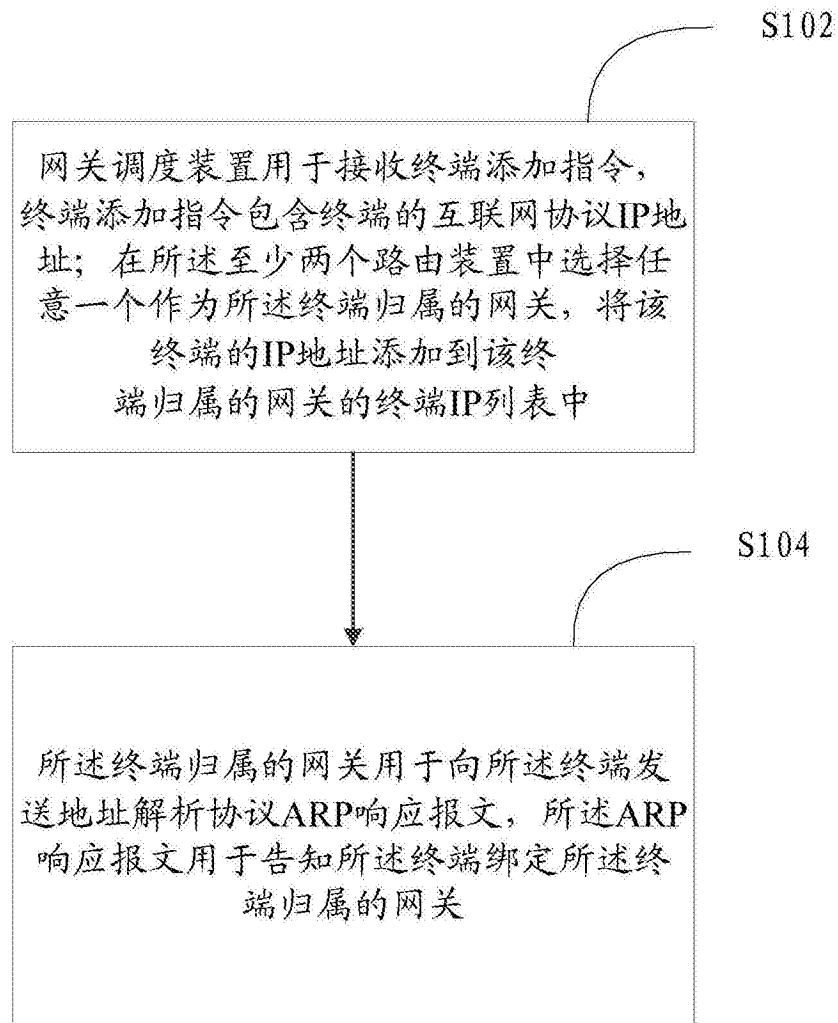


图5

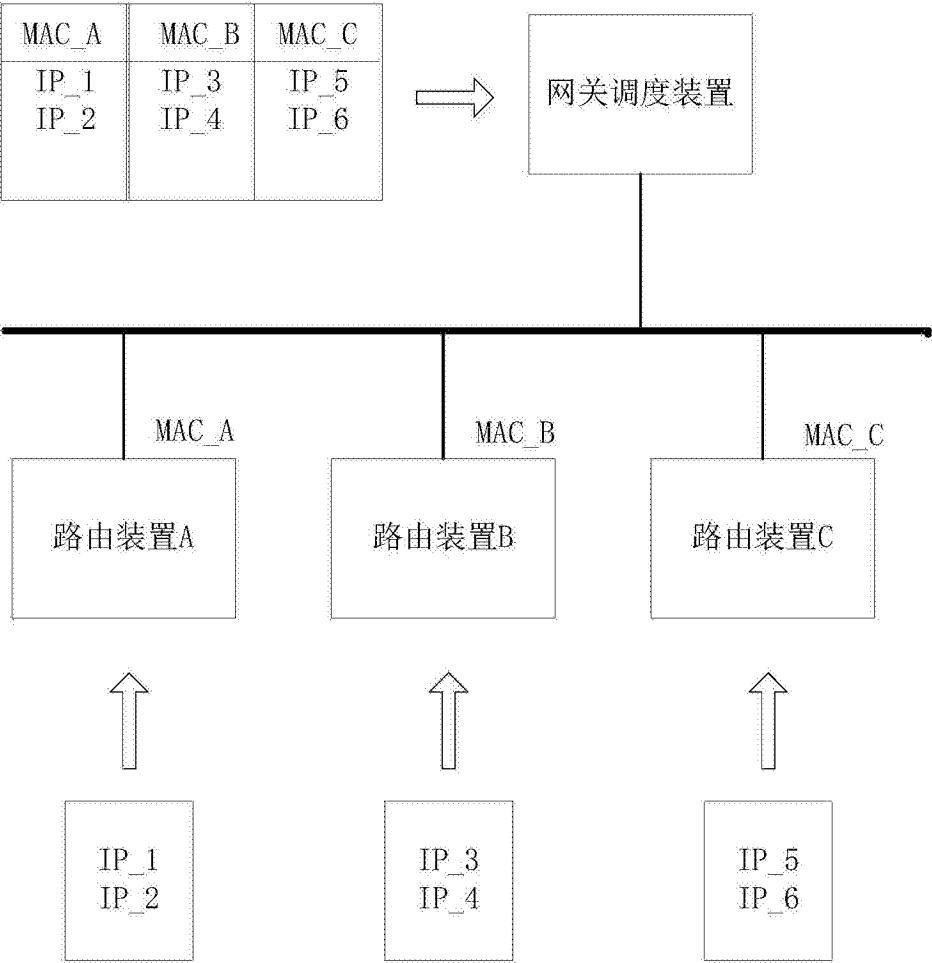


图6

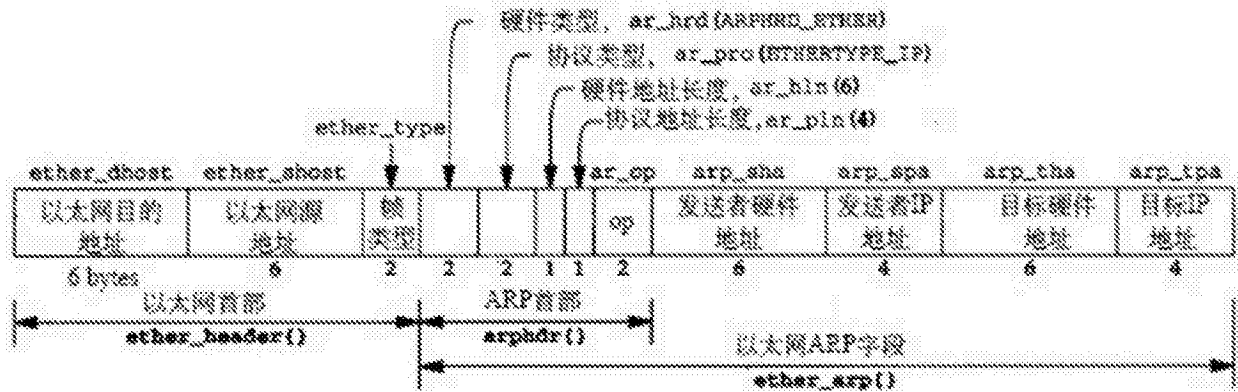


图7

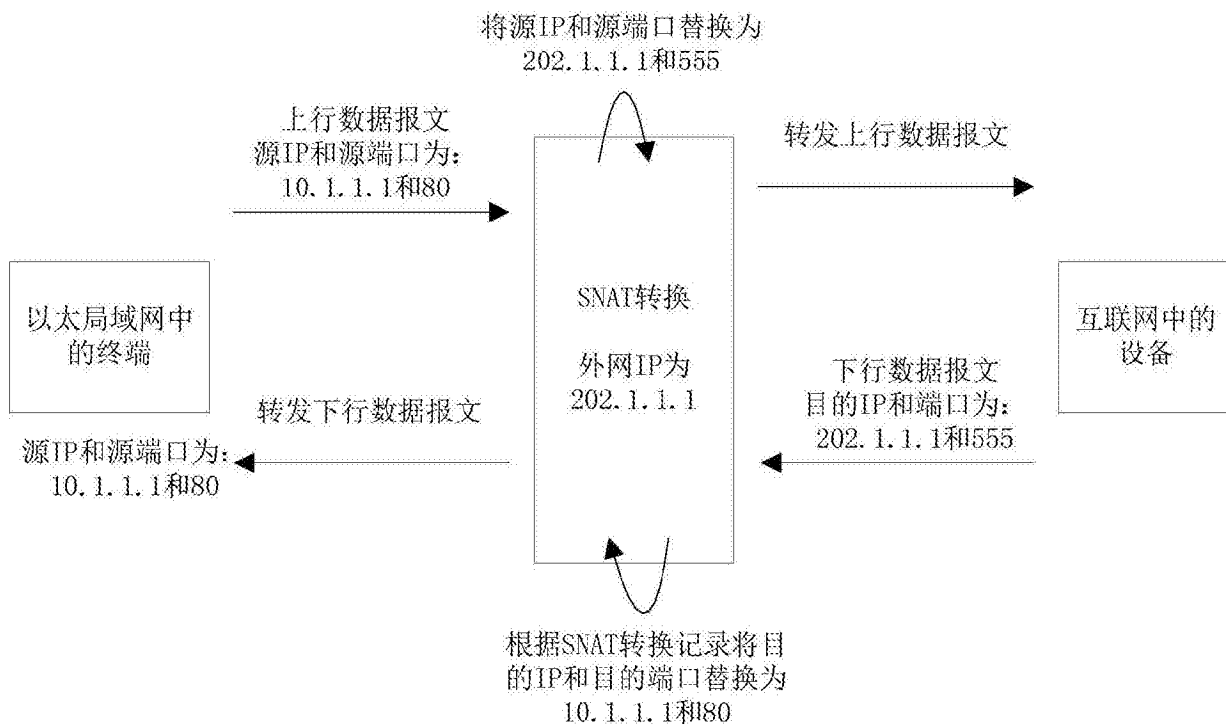


图8

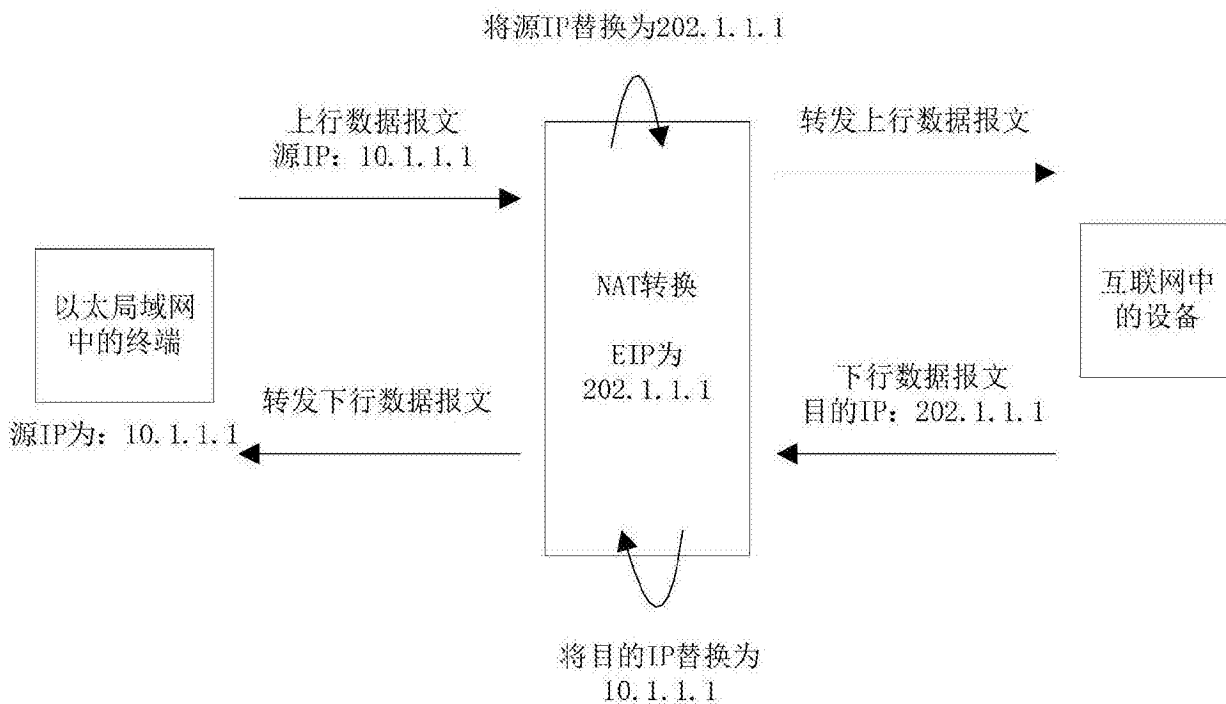


图9

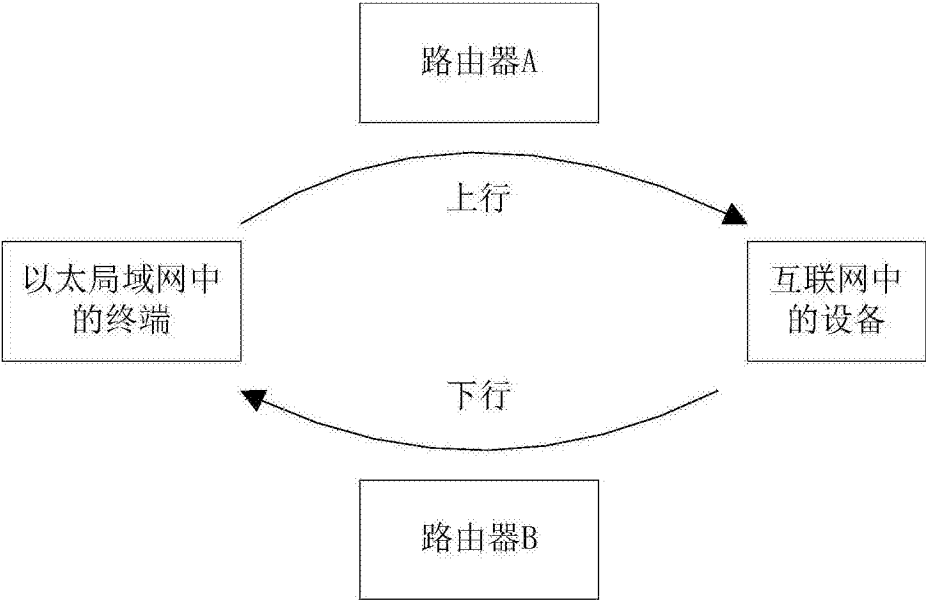


图10

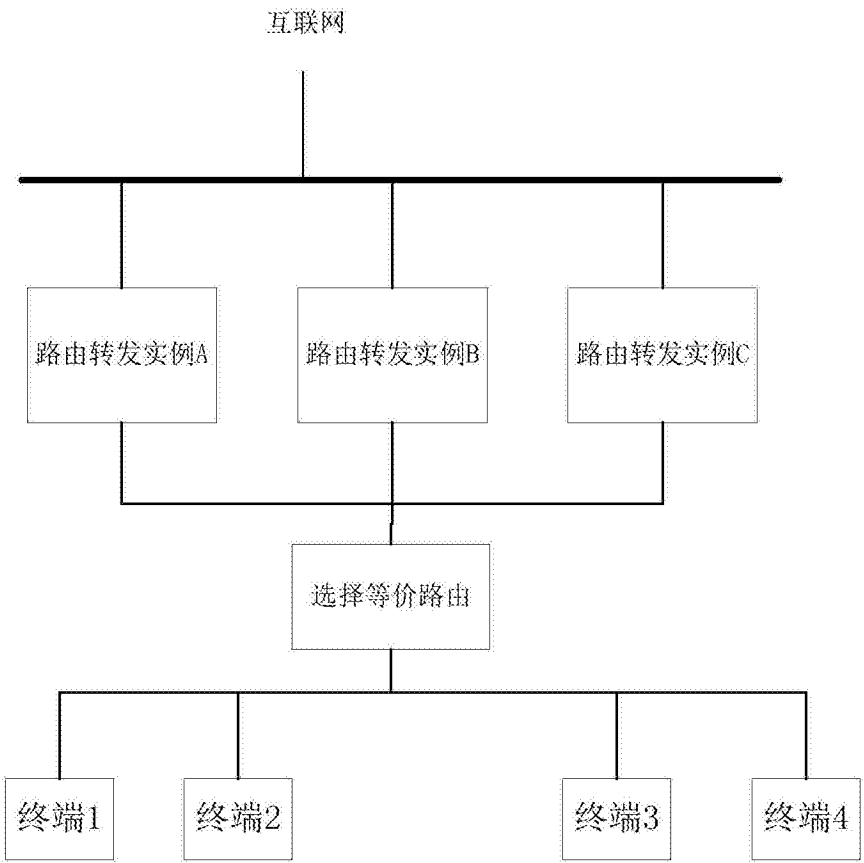


图11

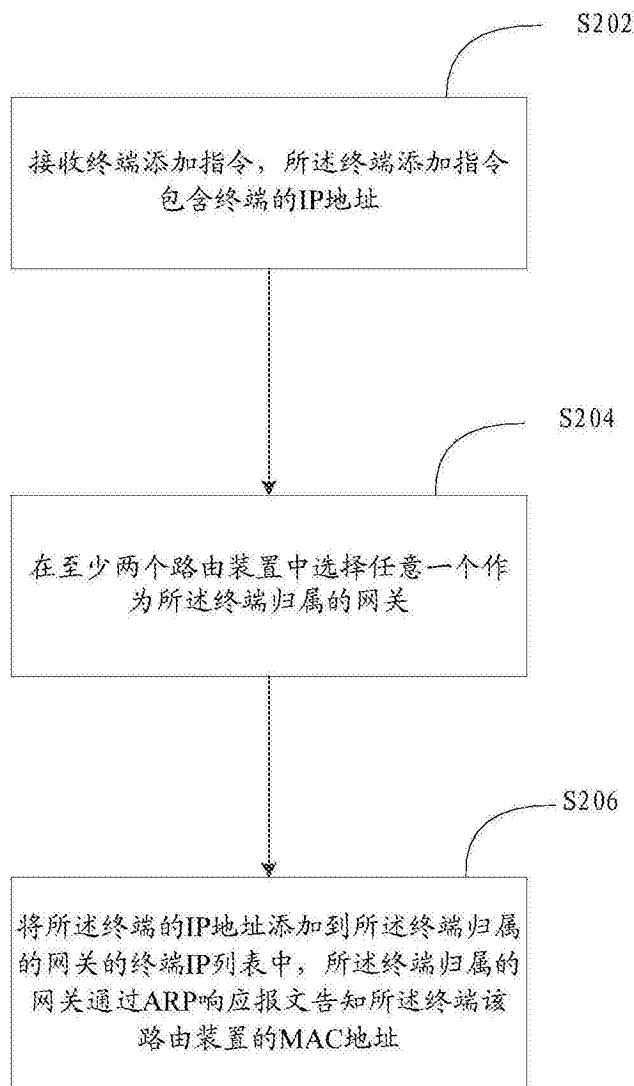


图12

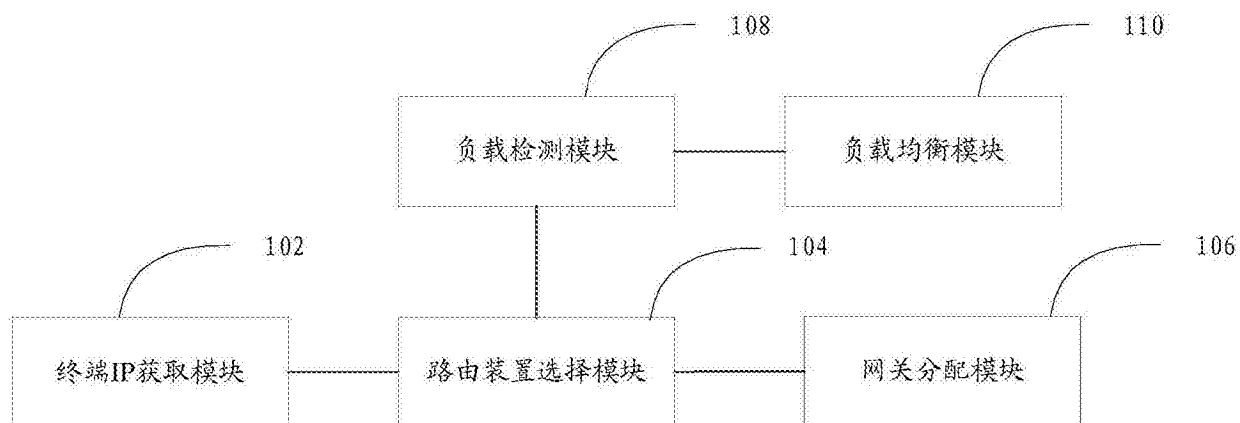


图13

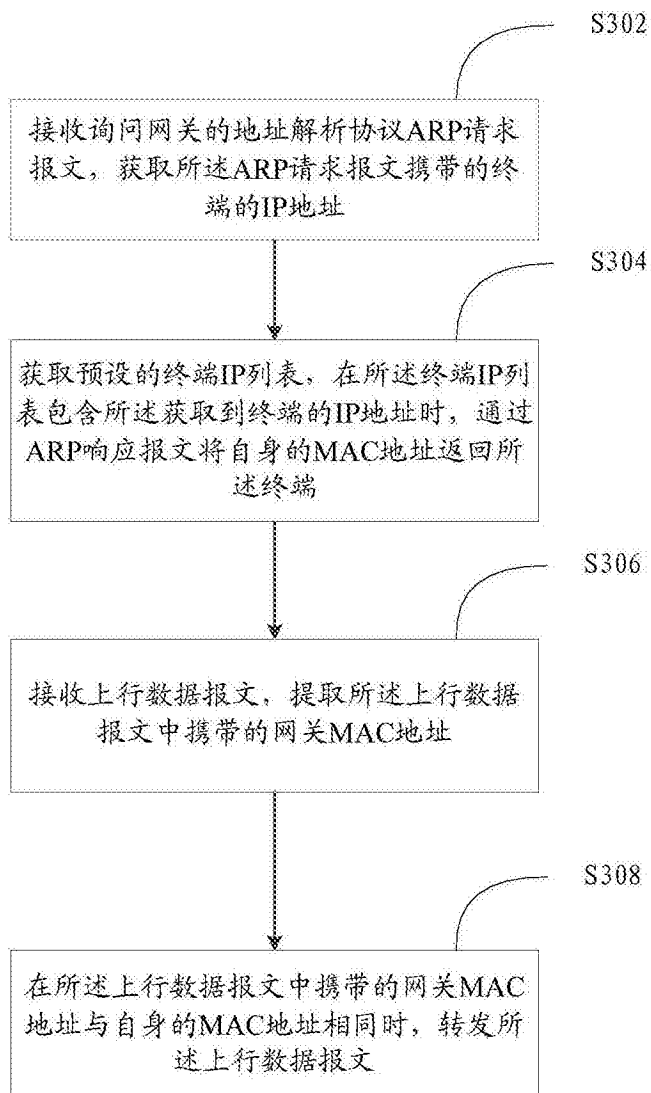


图14

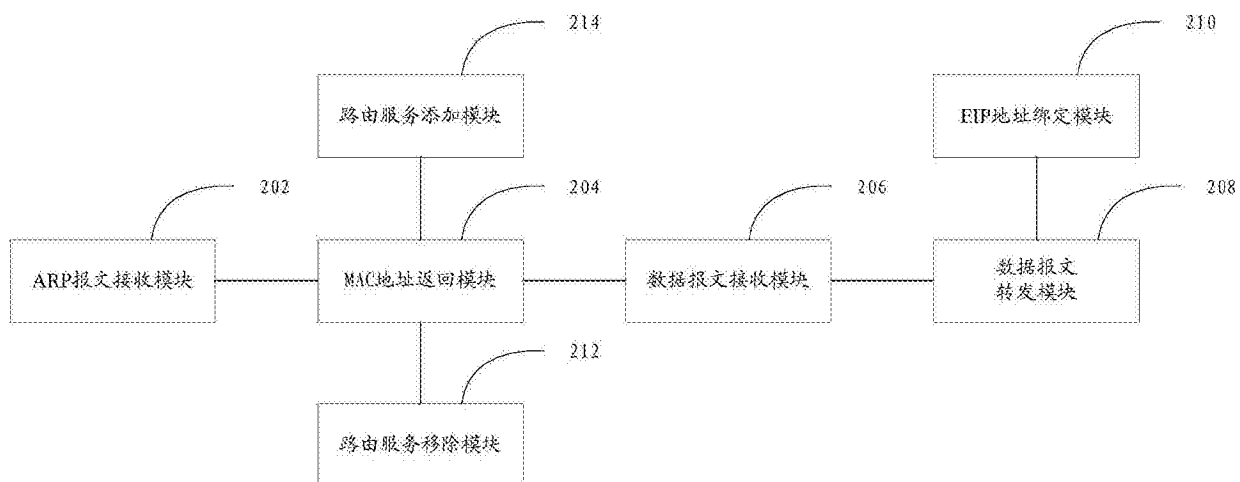


图15