



(12) 发明专利

(10) 授权公告号 CN 108199947 B

(45) 授权公告日 2020.10.09

(21) 申请号 201810096635.4

H04L 12/725 (2013.01)

(22) 申请日 2018.01.31

(56) 对比文件

(65) 同一申请的已公布的文献号
申请公布号 CN 108199947 A

CN 106656843 A, 2017.05.10
CN 106936682 A, 2017.07.07
CN 106713005 A, 2017.05.24
CN 106254203 A, 2016.12.21

(43) 申请公布日 2018.06.22

US 9735983 B1, 2017.08.15
US 2017141963 A1, 2017.05.18
US 2017126587 A1, 2017.05.04

(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

审查员 王璐

(72) 发明人 黄李伟 王伟

(74) 专利代理机构 北京林达刘知识产权代理事
务所(普通合伙) 11277
代理人 刘新宇

(51) Int. Cl.

H04L 12/46 (2006.01)
H04L 12/721 (2013.01)

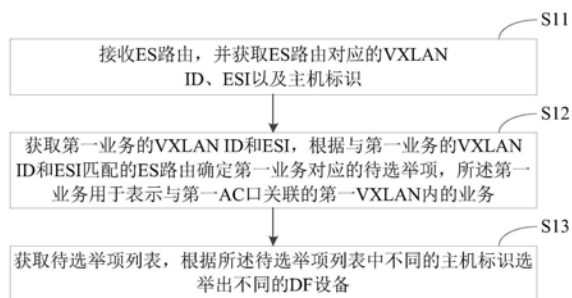
权利要求书2页 说明书8页 附图4页

(54) 发明名称

指定转发者DF选举方法及装置

(57) 摘要

本公开涉及一种指定转发者DF选举方法及装置,应用于VTEP设备,包括接收ES路由,并获取ES路由对应的可扩展虚拟局域网络标识VXLAN ID、以太网段标识ESI以及主机标识;获取第一业务的VXLAN ID和ESI,将与第一业务的VXLAN ID和ESI匹配的ES路由确定为第一业务对应的待选举项,所述第一业务用于表示与第一AC口关联的第一VXLAN内的业务;获取待选举项列表,根据所述待选举项列表中不同的主机标识选举出不同的DF设备。通过针对不同主机标识选举出不同的DF设备,根据本公开实施例的DF选举方法及装置能够实现同一VXLAN内的业务流量的负载分担。



1. 一种指定转发者DF选举方法,其特征在于,所述方法应用于以太虚拟专用网络EVPN多归属组网中的可扩展虚拟局域网络隧道端点VTEP设备,所述方法包括:

接收以太网段ES路由,并获取ES路由对应的可扩展虚拟局域网络标识VXLAN ID、以太网段标识ESI以及主机标识;

获取第一业务的VXLAN ID和ESI,根据与第一业务的VXLAN ID和ESI匹配的ES路由确定第一业务对应的待选举项,所述第一业务用于表示与第一AC口关联的第一VXLAN内的业务,其中,所述第一业务对应的待选举项包括与第一业务的VXLAN ID和ESI匹配的ES路由对应的主机标识;

获取待选举项列表,根据所述待选举项列表中不同的主机标识选举出不同的DF设备,其中,待选举项列表包括与第一业务对应的多个待选举项;

其中,根据所述待选举项列表中不同的主机标识选举出不同的DF设备,包括:

针对不同的主机标识,从所述待选举项列表中选举出与某主机标识对应的DF设备,并且,将所述待选举项列表中已被选取的待选举项设置为不可用状态,处于不可用状态的待选举项不能用于DF设备的选举,已被选举的待选举项用于表示被选出的DF设备对应的待选举项。

2. 根据权利要求1所述的方法,其特征在于,针对不同的主机标识,从所述待选举项列表中选取与某主机标识对应的DF设备,并且,将所述待选举项列表中已被选取的待选举项设置为不可用状态,包括:

针对第一主机标识,从所述待选举项列表中选举出与第一主机标识对应的第一DF设备,并且,将所述待选举项列表中包括所述第一DF设备的IP地址的待选举项设置为不可用状态;

针对第二主机标识,从所述待选举项列表中选举出与第二主机标识对应的第二DF设备,并且,将所述待选举项中包括第二DF设备的IP地址的待选举项设置为不可用状态,所述第二主机标识与所述第一主机标识不同。

3. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

获取第一AC口配置的ESI,与第一AC口关联的第一VXLAN的VXLAN ID,以及接入第一VXLAN的主机的主机标识;

针对每个主机标识,发送ES路由,所述ES路由中携带有所述第一VXLAN的VXLAN ID和所述第一AC口配置的ESI,以及该主机标识。

4. 根据权利要求1至3任一项所述的方法,其特征在于,所述主机标识包括MAC地址和IP地址中的一者或两者。

5. 一种指定转发者DF选举装置,其特征在于,所述装置应用于EVPN多归属组网中的VTEP设备,所述装置包括:

接收模块,用于接收ES路由,并获取ES路由对应的VXLAN ID、ESI以及主机标识;

第一获取模块,用于获取第一业务的VXLAN ID和ESI,根据与第一业务的VXLAN ID和ESI匹配的ES路由确定第一业务对应的待选举项,所述第一业务用于表示与第一AC口关联的第一VXLAN内的业务,其中,所述第一业务对应的待选举项包括与第一业务的VXLAN ID和ESI匹配的ES路由对应的主机标识;

选举模块,用于获取待选举项列表,根据所述待选举项列表中不同的主机标识选举出

不同的DF设备,其中,待选举项列表包括与第一业务对应的多个待选举项;

其中,所述选举模块包括:

选举子模块,用于针对不同的主机标识,从所述待选举项列表中选举出与某主机标识对应的DF设备,并且,将所述待选举项列表中已被选取的待选举项设置为不可用状态,处于不可用状态的待选举项不能用于DF设备的选举,已被选举的待选举项用于表示被选出的DF设备对应的待选举项。

6. 根据权利要求5所述的装置,其特征在于,所述选举子模块还用于:

针对第一主机标识,从所述待选举项列表中选举出与第一主机标识对应的第一DF设备,并且,将所述待选举项列表中包括所述第一DF设备的IP地址的待选举项设置为不可用状态;

针对第二主机标识,从所述待选举项列表中选举出与第二主机标识对应的第二DF设备,并且,将所述待选举项中包括第二DF设备的IP地址的待选举项设置为不可用状态,所述第二主机标识与所述第一主机标识不同。

7. 根据权利要求5所述的装置,其特征在于,所述装置还包括:

第二获取模块,用于获取第一AC口配置的ESI,与第一AC口关联的第一VXLAN的VXLAN ID,以及接入第一VXLAN的主机的主机标识;

发送模块,用于针对每个主机标识,发送ES路由,所述ES路由中携带有所述第一VXLAN的VXLAN ID和所述第一AC口配置的ESI,以及该主机标识。

8. 根据权利要求5至7任一项所述的装置,其特征在于,所述主机标识包括MAC地址和IP地址中的一者或两者。

指定转发者DF选举方法及装置

技术领域

[0001] 本公开涉及通信技术领域,尤其涉及一种指定转发者DF (Designated Forwarder, 指定转发者) 选举方法及装置。

背景技术

[0002] EVPN (Ethernet Virtual Private Network, 以太虚拟专用网) 是一种二层VPN技术。为了提高接入侧的可靠性,协议规定了EVPN多归属接入网,从而提高业务数据报文的可靠性。在EVPN多归属接入网中,一个VM (Virtual Machine, 虚拟机) 可以通过多条链路同时接入不同的VTEP (VXLAN Tunnel End Point, 可扩展虚拟局域网络隧道端点) 设备。VM接入不同VTEP设备的多条链路形成一个ES (Ethernet Segment, 以太网段)。ESI (Ethernet Segment Identifier, 以太网段标识) 可以用来标识一个ES。

[0003] 对应一个VM接入不同的VTEP设备,为了让BUM (Broadcast, Unknown Unicast or Multicast, 广播、未知单播、多播) 流量不存在环路和多份的问题,多个VTEP设备可以从满足相同ESI值的VTEP设备中选举出一个DF,由该DF负责转发来自VM以及发往VM的报文。

发明内容

[0004] 有鉴于此,本公开提出了一种DF选举方法及装置能够在在一个VXLAN内进行业务的负载分担。

[0005] 根据本公开的一方面,提供了一种DF选举方法,所述方法应用于EVPN多归属组网中的VTEP设备,所述方法包括:接收以太网段ES路由,并获取ES路由对应的可扩展虚拟局域网络标识VXLAN ID、以太网段标识ESI以及主机标识;获取第一业务的VXLAN ID和ESI,根据与第一业务的VXLAN ID和ESI匹配的ES路由确定第一业务对应的待选举项,所述第一业务用于表示与第一AC口关联的第一VXLAN内的业务;获取待选举项列表,根据所述待选举项列表中不同的主机标识选举出不同的DF设备。

[0006] 根据本公开的另一方面,提供了一种DF选举装置,所述装置应用于EVPN多归属组网中的VTEP设备,所述装置包括:接收模块,用于接收ES路由,并获取ES路由对应的VXLAN ID、ESI以及主机标识;第一获取模块,用于获取第一业务的VXLAN ID和ESI,根据与第一业务的VXLAN ID和ESI匹配的ES路由确定第一业务对应的待选举项,所述第一业务用于表示与第一AC口关联的第一VXLAN内的业务;选举模块,用于获取待选举项列表,根据所述待选举项列表中不同的主机标识选举出不同的DF设备。

[0007] 通过接收ES路由,并获取ES路由对应的VXLAN ID、ESI以及主机标识;获取第一业务的VXLAN ID和ESI,根据与第一业务的VXLAN ID和ESI匹配的ES路由确定与第一AC口关联的第一VXLAN内的第一业务对应的待选举项,获取待选举项列表,根据所述待选举项列表中不同的主机标识选举出不同的DF设备,根据本公开的各方面实施例的DF选举方法及装置能够将一个VXLAN内的业务按照主机标识进行划分,针对不同主机标识对应业务分别进行DF选举,从而实现同一VXLAN内的业务流量的负载分担。

[0008] 根据下面参考附图对示例性实施例的详细说明,本公开的其它特征及方面将变得清楚。

附图说明

[0009] 包含在说明书中并且构成说明书的一部分的附图与说明书一起示出了本公开的示例性实施例、特征和方面,并且用于解释本公开的原理。

[0010] 图1示出根据本公开一实施例的DF选举方法的流程图。

[0011] 图2示出根据本公开一实施例的EVPN多归属组网的示意图。

[0012] 图3示出根据本公开一实施例的DF选举方法的流程图。

[0013] 图4示出根据本公开一实施例的DF选举方法的流程图。

[0014] 图5示出根据本公开一实施例的DF选举装置的框图。

[0015] 图6示出根据本公开一实施例的DF选举装置的框图。

[0016] 图7示出根据本公开一实施例的DF选举装置的框图。

具体实施方式

[0017] 以下将参考附图详细说明本公开的各种示例性实施例、特征和方面。附图中相同的附图标记表示功能相同或相似的元件。尽管在附图中示出了实施例的各种方面,但是除非特别指出,不必按比例绘制附图。

[0018] 在这里专用的词“示例性”意为“用作例子、实施例或说明性”。这里作为“示例性”所说明的任何实施例不必解释为优于或好于其它实施例。

[0019] 另外,为了更好的说明本公开,在下文的具体实施方式中给出了众多的具体细节。本领域技术人员应当理解,没有某些具体细节,本公开同样可以实施。在一些实例中,对于本领域技术人员熟知的方法、手段、元件和电路未作详细描述,以便于凸显本公开的主旨。

[0020] 图1示出根据本公开一实施例的DF选举方法的流程图。图2示出根据本公开一实施例的EVPN多归属组网的示意图。如图2所示,VTEP1、VTEP2、VTEP3和VTEP4表示VTEP设备;VM1和VM2表示VM(Virtual Machine,虚拟机);LSW1和LSW2表示交换机。VM1通过一个交换机LSW1同时接入3个VTEP设备,分别为VTEP1、VTEP2和VTEP3。VM2通过一个交换机LSW2同时接入3个VTEP设备,分别为VTEP1、VTEP2和VTEP3。其中,VM1和VM2接入同一VTEP设备的同一个物理端口。VTEP1、VTEP2和VTEP3上分别创建了AC(Attachment Circuit,接入电路)口1、AC口2和AC口3,AC口1、AC口2和AC口3配置了相同的ESI的值。

[0021] 在EVPN中,各VTEP设备之间建立邻居关系后互相发送ES路由(EVPN路由中的4类路由,该ES路由可以用于进行DF选举);根据ES路由中携带的ESI值,各VTEP设备上会生成DF选举列表,DF选举列表中包含配置相同ESI的所有VTEP设备的信息。

[0022] 相关技术中,各VTEP设备可以通过从其他VTEP设备收到的ES路由的“Originating Router's IP address”字段确定其他VTEP设备的Router ID;各VTEP设备可以Router ID大小的顺序对DF选举列表内的VTEP设备进行排序,按照一定算法选出配置该ESI的所有VTEP设备中的DF。例如,各VTEP设备可以按照Router ID由小到大的顺序对DF选举列表内的VTEP设备进行排序,并按照顺序分配由0开始的序号。根据 $i = \text{VLAN ID} \bmod n$ 公式获取可以称为DF的VTEP设备序号*i*,其中VLAN ID为VTEP设备上绑定的VLAN的ID,*n*为DF选举列表内包含的

VTEP设备数量, i 表示VLAN ID除以 n 后的余数。

[0023] 由此可见,对于某个特定的VLAN,DF选举列表中的VTEP设备选举出的DF设备为固定的VTEP设备。这样对于某个特定的VLAN,只有某个VTEP设备能从VM接收流量或转发流量到VM。由于在VTEP设备上,来自VM的某个VLAN的流量接入的VXLAN是固定的,而该VLAN对应选举出的DF设备是固定的。这样,同一VXLAN内的业务只能通过同一个VTEP设备,这导致负载分担无法在一个VXLAN业务内进行。需要说明的是,上述的VLAN ID可以替换为VXLAN ID。

[0024] 图1所示的DF选举方法可以应用于EVPN多归属组网中的VTEP设备,例如图2所示的VTEP1、VTEP2和VTEP3。如图1所示,该DF选举方法包括:

[0025] 步骤S11,接收ES路由,并获取ES路由对应的VXLAN ID、ESI以及主机标识。

[0026] 其中,其中,VXLAN ID是VXLAN的标识,可以用于识别唯一的VXLAN,其长度为24比特。ESI是以太网段标识,可以用来识别一个ES,ESI是一个10字节的具有全局意义的ID。主机标识可以用于识别唯一的主机,这里的主机可以为一台服务器,也可以为一台服务器上的一个VM,本公开对此不做限制。在一种可能的实现方式中,主机标识可以为MAC地址和IP地址中的一者或两者。

[0027] 在一种可能的实现方式中,ES路由中携带有发送该ES路由的VTEP设备的某个AC口上配置的ESI,以及与该AC口关联的VXLAN的VXLAN ID。VTEP设备可以在通告ES路由的同时,通告接入该VXLAN的主机的主机标识。VTEP设备接收到一个ES路由后,可以将该ES路由中获取的ESI和VXLAN ID作为ES路由对应的ESI和VXLAN ID,并将在接收到该ES路由时接收到的主机标识,作为该ES路由对应的主机标识。

[0028] 在一种可能的实现方式中,ES路由中可以携带有发送该ES路由的VTEP设备的某个AC口上配置的ESI、与该AC口关联的VXLAN的VXLAN ID,以及接入该VXLAN的主机的主机标识。如此,VTEP设备可以将该ES路由中获取的VXLAN ID、ESI以及主机标识,作为ES路由对应的VXLAN ID、ESI以及主机标识。

[0029] 步骤S12,获取第一业务的VXLAN ID和ESI,根据与第一业务的VXLAN ID和ESI匹配的ES路由确定第一业务对应的待选举项,所述第一业务用于表示与第一AC口关联的第一VXLAN内的业务。

[0030] 其中,第一业务可以用于表示与第一AC口关联的第一VXLAN内的业务。对于第一业务的其他方面,例如类型、内容等,本公开不做限制。第一AC口可以用于表示VTEP设备上任意一个AC口,第一AC口上配置有ESI。第一VXLAN可以用于表示与第一AC口关联的VXLAN。第一业务的VXLAN ID为第一VXLAN的VXLAN ID,第一业务的ESI为第一AC口配置的ESI。

[0031] 若ES路由对应的VXLAN ID和ESI与第一业务的VXLAN ID和ESI一致,则表示该ES路由是针对第一业务进行DF选举时采用的ES路由,VTEP设备可以确定该ES路由与第一业务匹配。

[0032] 第一业务对应的待选举项包括与第一业务的VXLAN ID和ESI匹配的ES路由对应的主机标识和发送与第一业务的VXLAN ID和ESI匹配的ES路由的VTEP设备的IP地址。当然,由于不同的主机中可以包括同一业务,例如,不同的VTEP中,均有相同VXLAN ID的业务,所以,本公开中,根据第一业务可以确定出多个待选举项。

[0033] 步骤S13,获取待选举项列表,根据所述待选举项列表中不同的主机标识选举出不同的DF设备。

[0034] 由于根据第一业务可以确定出多个待选举项,所以该待选举项列表可以包括与第一业务对应的多个待选举项。

[0035] 由于不同的待选举项会对应不同的主机标识,所以待选举项列表包括了不同的主机标识。针对每个主机标识,VTEP设备可以在待选举项列表中,查找包括该主机标识的待选举项,并根据包括该主机标识的待选举项选出该主机标识对应的DF设备。其中,针对不同的主机标识,VTEP设备选举的DF设备不同。

[0036] 通过接收ES路由,并获取ES路由对应的VXLAN ID、ESI以及主机标识;获取第一业务的VXLAN ID和ESI,根据与第一业务的VXLAN ID和ESI匹配的ES路由确定与第一AC口关联的第一VXLAN内的第一业务对应的待选举项,获取待选举项列表,根据所述待选举项列表中不同的主机标识选举出不同的DF设备,根据本公开的实施例的DF选举方法能够将一个VXLAN内的业务按照主机标识进行划分,针对不同主机标识对应业务分别进行DF选举,从而实现同一VXLAN内的业务流量的负载分担。

[0037] 在一种可能的实现方式中,步骤S13可以包括:针对不同的主机标识,VTEP设备可以采用不同的DF选举方法选举出不同的DF设备。在一个示例中,针对主机标识1,VTEP设备可以按照Router ID由小到大的顺序对DF选举列表内的VTEP设备进行排序,并按照顺序分配由0开始的序号。而针对主机标识2,VTEP设备可以按照Router ID由大到小的顺序对DF选举列表内的VTEP设备进行排序,并按照顺序分配0开始的序号。这样,主机标识1和主机标识2,根据 $i = \text{VLAN ID} \bmod n$ 公式获取的作为DF设备的VTEP设备序号*i*虽然是相同的,但该序号*i*对应的VTEP设备是不同的。

[0038] 在一种可能的实现方式中,步骤S13可以包括:针对不同的主机标识,VTEP设备可以从所述待选举项列表中选举出与某主机标识对应的DF设备,并且,将所述待选举项列表中已被选取的待选举项设置为不可用状态,其中,处于不可用状态的待选举项不能用于DF设备的选举,已被选举的待选举项可以用于表示被选出的DF设备对应的待选举项。在一个示例中,已被选举的待选举项可以为包括被选举出的DF设备的IP地址的待选举项。

[0039] 在一种可能的实现方式中,图3示出根据本公开一实施例的DF选举方法的流程图。如图3所示,针对不同的主机标识,VTEP设备可以从所述待选举项列表中选举出与某主机标识对应的DF设备,并且,将所述待选举项列表中已被选取的待选举项设置为不可用状态,包括:

[0040] 步骤S131,针对第一主机标识,从所述待选举项列表中选举出与第一主机标识对应的第一DF设备,并且,将所述待选举项列表中包括所述第一DF设备的IP地址的待选举项设置为不可用状态。

[0041] 第一主机标识可以用于表示通过第一AC口接入第一VXLAN的任意一个主机的标识。第一DF设备可以用于表示与第一主机标识对应的DF设备。

[0042] 待选举项列表中的待选举项可以包括不同的主机标识。VTEP设备在确定第一主机标识对应的DF设备时,可以先从待选举项列表中,查找出包括该第一主机标识的待选举项,再根据DF选举规则,从包括第一主机标识的待选举项对应的VTEP设备中,选择出与该第一主机标识对应的DF设备,即第一DF设备。

[0043] 由于处于不可用状态待选举项不能用于DF设备的选举,因此,VTEP设备在确定第一主机标识对应的DF设备的过程中,不会使用处于不可用状态的待选举项。在一个示例中,

VTEP设备从待选举项列表中查找出来的包括第一主机标识的待选举项均为处于可用状态的待选举项。

[0044] VTEP设备选择出第一DF设备后,可以将待选举项列表中包括第一DF设备的IP得知的待选举项设置为不可用状态。

[0045] 在一种可能的实现方式中,VTEP设备可以为待选举项添加用于表示待选举项是否可用的状态标识。例如,VTEP设备可以为待选举项添加0或1作为状态标识,当待选举项的状态标识为0时表示该待选举项处于不可用状态,不可用于后续DF设备的选举;当待选举项的状态标识为1时表示该待选举项处于可用状态,可以用于后续DF设备的选举。

[0046] 步骤S132,针对第二主机标识,从所述待选举项列表中选举出与第二主机标识对应的第二DF设备,并且,将所述待选举项中包括第二DF设备的IP地址的待选举项设置为不可用状态,所述第二主机标识与所述第一主机标识不同。

[0047] 第二主机标识可以用于表示通过第一AC口接入第一VXLAN的主机的主机标识,且第二主机标识对应的主机与第一主机标识对应的主机不同。第二DF设备可以用于表示与第二主机标识对应的DF设备。步骤S132中选举第二DF设备的过程可以参照步骤S131中选举第一DF设备的过程,这里不再赘述。

[0048] 由于处于不可用状态的待选举项不能用于DF设备的选举。VTEP设备将包括第一DF设备的IP地址的待选举项设置为不可用状态的情况下,在进行第二DF设备的选举过程中,VTEP设备使用的待选举项都不包括第一DF设备的IP地址,VTEP设备选举出来的第二DF设备也就不可能是第一DF设备。这样,第二主机标识和第一主机标识对应的DF设备不同,从而达到针对不同主机标识对应业务分别进行DF选举的目的,实现同一VXLAN内的业务流量的负载分担。

[0049] 选取出第二主机标识对应的第二DF设备后,同样将所述待选举项中包括第二DF设备的IP地址的待选举项设置为不可用状态,这样,在后续针对其他主机标识对应的DF设备的选举过程中,不会选到第二DF设备,达到针对不同主机标识对应业务分别进行DF选举的目的,实现同一VXLAN内的业务流量的负载分担。

[0050] 图4示出根据本公开一实施例的DF选举方法的流程图。如图4所示,该方法还包括:

[0051] 步骤S21,获取第一AC口配置的ESI,与第一AC口关联的第一VXLAN的VXLAN ID,以及接入第一VXLAN的主机的主机标识。

[0052] 其中,第一AC口,ESI、第一VXLAN,以及主机标识的描述可以参见步骤S11和步骤S12,这里不再赘述。

[0053] 步骤S22,针对每个主机标识,发送ES路由,所述ES路由中携带有所述第一VXLAN的VXLAN ID和所述第一AC口配置的ESI,以及该主机标识。

[0054] 举例来说,针对主机标识1,VTEP设备发送一个ES路由,该ES路由中携带有所述第一VXLAN的VXLAN ID和所述第一AC口配置的ESI,以及主机标识1。针对主机标识2,VTEP设备发送一个ES路由,该ES路由中携带有所述第一VXLAN的VXLAN ID和所述第一AC口配置的ESI,以及主机标识2。直至遍历了通过第一AC口接入第一VXLAN的所有主机的主机标识。

[0055] 这样,针对通过第一AC口接入第一VXLAN的每一个主机,发送一个携带该主机的主机标识的ES路由,使得接收到ES路由的VTEP设备可以针对不同的主机标识分别进行DF选举,选举出不同的DF,实现同一VXLAN内的业务流量的负载分担。

[0056] 应用示例

[0057] 如图2所示,VM1、VM2多归属接入VTEP设备,分别为VTEP1、VTEP2、VTEP3。其中,VM1的MAC地址为A-A-A,IP地址为10.1.1.2,VM2的MAC地址为B-B-B,IP地址为10.1.1.3。VTEP1的IP地址为1.1.1.1,VTEP2的IP地址为2.2.2.2,VTEP3的IP地址为3.3.3.3。以VM1的MAC地址为第一主机标识(或者以VM1的IP地址为第一主机标识,或者以VM1的MAC地址和IP地址为第一主机标识),VM2的MAC地址为第二主机标识(或者以VM2的IP地址为第二主机标识,或者以VM2的MAC地址和IP地址为第二主机标识),VTEP1、VTEP2、VTEP3上分别创建了第一AC口,且各第一AC口的ESI配置为1,与第一AC口关联的第一VXLAN的VXLAN ID为100,VM1和VM2分别通过第一AC口接入第一VXLAN为例,对本公开实施例的DF选举方法进行说明。其中,VTEP1、VTEP2、VTEP3、VTEP4两两之间建立BGP EVPN邻居。

[0058] 针对VM1,VTEP1分别向VTEP2、VTEP3和VTEP4发送ES路由,该ES路由中携带有值为100的VXLAN ID、值为1的ESI、值为A-A-A的主机标识,以及VTEP1的IP地址1.1.1.1;VTEP2分别向VTEP1、VTEP3和VTEP4发送ES路由,该ES路由中携带有值为100的VXLAN ID、值为1的ESI、值为A-A-A的主机标识,以及VTEP2的IP地址2.2.2.2;VTEP3分别向VTEP1、VTEP2和VTEP4发送ES路由,该ES路由中携带有值为100的VXLAN ID、值为1的ESI、值为A-A-A的主机标识,以及VTEP3的IP地址3.3.3.3。

[0059] 针对VM2,VTEP1分别向VTEP2、VTEP3和VTEP4发送ES路由,该ES路由中携带有值为100的VXLAN ID、值为1的ESI、值为B-B-B的主机标识,以及VTEP1的IP地址1.1.1.1;VTEP2分别向VTEP1、VTEP3和VTEP4发送ES路由,该ES路由中携带有值为100的VXLAN ID、值为1的ESI、值为B-B-B的主机标识,以及VTEP2的IP地址2.2.2.2;VTEP3分别向VTEP1、VTEP2和VTEP4发送ES路由,该ES路由中携带有值为100的VXLAN ID、值为1的ESI、值为B-B-B的主机标识,以及VTEP3的IP地址3.3.3.3。

[0060] 如此,VTEP1、VTEP2和VTEP3上分别形成了第一业务(VXLAN ID为100,ESI为1)对应的待选举项列表,该待选举项列表包括(与第一业务的VXLAN ID和ESI匹配的ES路由对应的)主机标识和(发送与第一业务的VXLAN ID和ESI匹配的ES路由的)VTEP设备的IP地址,例如表1所示待选举项列表。

[0061] 表1

[0062]

	主机标识	VXLAN ID	ESI	VTEP设备的IP地址
1	A-A-A	100	1	1.1.1.1
2	A-A-A	100	1	2.2.2.2
3	A-A-A	100	1	3.3.3.3
4	B-B-B	100	1	1.1.1.1
5	B-B-B	100	1	2.2.2.2
6	B-B-B	100	1	3.3.3.3

[0063] 根据表1所示的待选举项列表中的主机标识A-A-A和主机标识B-B-B,选举出不同的DF设备。在一个示例中,针对主机标识A-A-A从表1所示的待选举项列表中选举出IP地址为1.1.1.1的VTEP设备(VTEP1)为与A-A-A对应的第一DF设备,并且将表1所示的待选举项列表中包括1.1.1.1的待选举项设置为不可用状态,即将第1行和第4行的待选举项设置为不可用状态。此时,第1行和第4行的待选举项不能参与后续DF设备的选举。这样,在针对主机

标识B-B-B从表1所示的待选举项列表中选择出的第二DF设备在IP地址为2.2.2的VTEP设备(VTEP2)与IP地址为3.3.3的VTEP设备(VTEP)之间产生,不可能是VTEP1。这样,VM1和VM2接入同一VXLAN网络的业务,可以采用不同的VTEP设备进行转发,实现了同一VXLAN内的业务流量的负载分担。

[0064] 假设A-A-A对应的DF设备为VTEP1,B-B-B对应的DF设备为VTEP2。

[0065] VM1向VTEP1、VTEP2和VTEP3上报免费ARP报文,VTEP2和VTEP3会直接丢弃该报文,VTEP1会让该报文进入,并在本地产生VM1的IP/MAC表项。VTEP1将VM1的IP/MAC地址表项通告给VTEP4。VTEP4会将VM1的转发表项下发。

[0066] VM2向VTEP1、VTEP2和VTEP3上报免费ARP报文,VTEP1和VTEP3会直接丢弃该报文,VTEP2会让该报文进入,并在本地产生VM2的IP/MAC表项。VTEP2将VM2的IP/MAC地址表项通告给VTEP4。VTEP4会将VM2的转发表项下发。

[0067] 这样,当VM3要访问VM1时,通过在VTEP4上进行表项查询,将流量向VTEP1转发,再转发给VM1。当VM3要访问VM2时,通过在VTEP4上进行表项查询,将流量向VTEP2转发,再发送给VM2。可见,VM1的流量通过VTEP1转发,VM2的流量通过VTEP2转发,达到了同一VXLAN内的业务流量的负载分担。

[0068] 图5示出根据本公开一实施例的DF选举装置的框图。所述装置应用于EVPN多归属组网中的VTEP设备。如图5所示,该装置50包括:

[0069] 接收模块51,用于接收ES路由,并获取ES路由对应的VXLAN ID、ESI以及主机标识。

[0070] 第一获取模块52,用于获取第一业务的VXLAN ID和ESI,根据与第一业务的VXLAN ID和ESI匹配的ES路由确定第一业务对应的待选举项,所述第一业务用于表示与第一AC口关联的第一VXLAN内的业务。

[0071] 选举模块53,用于获取待选举项列表,根据所述待选举项列表中不同的主机标识选举出不同的DF设备。

[0072] 图6示出根据本公开一实施例的DF选举装置的框图。如图6所示,在一种可能的实现方式中,所述选举模块53包括:

[0073] 选举子模块531,用于针对不同的主机标识,从所述待选举项列表中选择出与某主机标识对应的DF设备,并且,将所述待选举项列表中已被选取的待选举项设置为不可用状态,处于不可用状态的待选举项不能用于DF设备的选举,已被选举的待选举项用于表示被选出的DF设备对应的待选举项。

[0074] 在一种可能的实现方式中,所述选举子模块531还用于:

[0075] 针对第一主机标识,从所述待选举项列表中选择出与第一主机标识对应的第一DF设备,并且,将所述待选举项列表中包括所述第一DF设备的IP地址的待选举项设置为不可用状态;

[0076] 针对第二主机标识,从所述待选举项列表中选择出与第二主机标识对应的第二DF设备,并且,将所述待选举项中包括第二DF设备的IP地址的待选举项设置为不可用状态,所述第二主机标识与所述第一主机标识不同。

[0077] 在一种可能的实现方式中,所述装置50还包括:

[0078] 第二获取模块54,用于获取第一AC口配置的ESI,与第一AC口关联的第一VXLAN的VXLAN ID,以及接入第一VXLAN的主机的主机标识。

[0079] 发送模块55,用于针对每个主机标识,发送ES路由,所述ES路由中携带有所述第一VXLAN的VXLAN ID和所述第一AC口配置的ESI,以及该主机标识。

[0080] 在一种可能的实现方式中,所述主机标识包括MAC地址和IP地址中的一者或两者。

[0081] 通过接收ES路由,并获取ES路由对应的VXLAN ID、ESI以及主机标识;获取第一业务的VXLAN ID和ESI,根据与第一业务的VXLAN ID和ESI匹配的ES路由确定与第一AC口关联的第一VXLAN内的第一业务对应的待选举项,获取待选举项列表,根据所述待选举项列表中不同的主机标识选举出不同的DF设备,根据本公开的实施例的DF选举装置能够将一个VXLAN内的业务按照主机标识进行划分,针对不同主机标识对应业务分别进行DF选举,从而实现同一VXLAN内的业务流量的负载分担。

[0082] 图7是根据一示例性实施例示出的一种用于DF选举装置900的框图。参照图7,该装置900可包括处理器901、存储有机器可执行指令的机器可读存储介质902。处理器901与机器可读存储介质902可经由系统总线903通信。并且,处理器901通过读取机器可读存储介质902中与DF选举逻辑对应的机器可执行指令以执行上文所述的DF选举方法。

[0083] 本文中提到的机器可读存储介质902可以是任何电子、磁性、光学或其它物理存储装置,可以包含或存储信息,如可执行指令、数据,等等。例如,机器可读存储介质可以是:RAM (Random Access Memory,随机存取存储器)、易失存储器、非易失性存储器、闪存、存储驱动器(如硬盘驱动器)、固态硬盘、任何类型的存储盘(如光盘、dvd等),或者类似的存储介质,或者它们的组合。

[0084] 以上已经描述了本公开的各实施例,上述说明是示例性的,并非穷尽性的,并且也不限于所披露的各实施例。在不偏离所说明的各实施例的范围和精神的情况下,对于本技术领域的普通技术人员来说许多修改和变更都是显而易见的。本文中所用术语的选择,旨在最好地解释各实施例的原理、实际应用或对市场中的技术的技术改进,或者使本技术领域的其它普通技术人员能理解本文披露的各实施例。

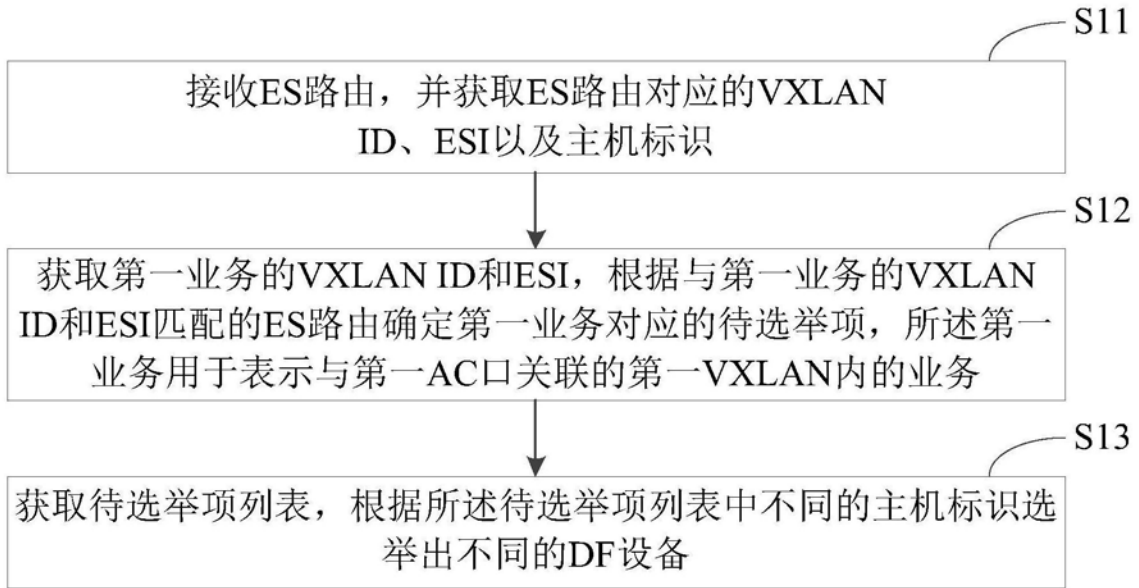


图1

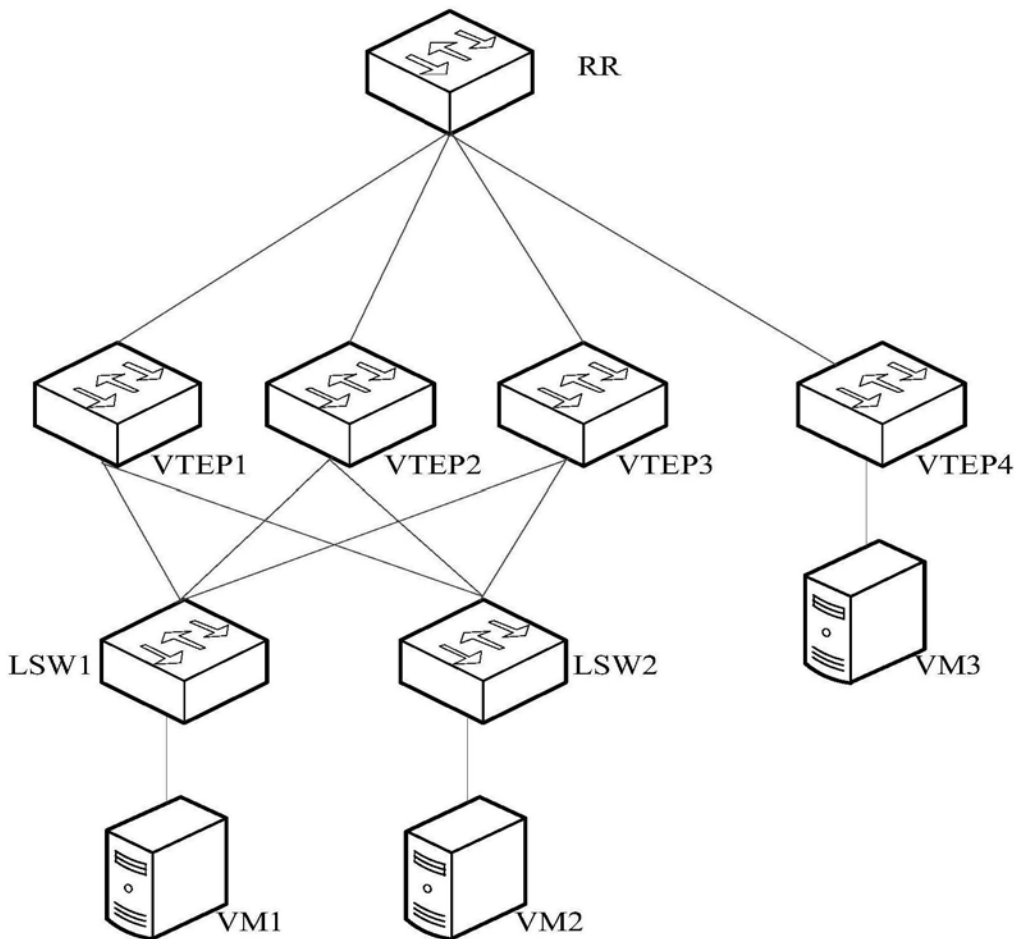


图2

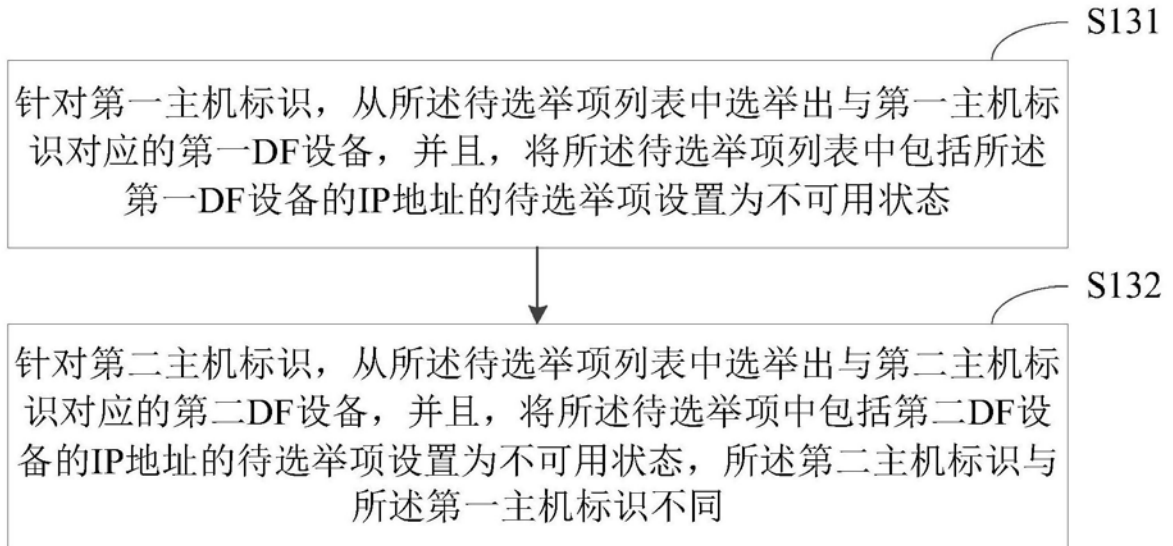


图3



图4

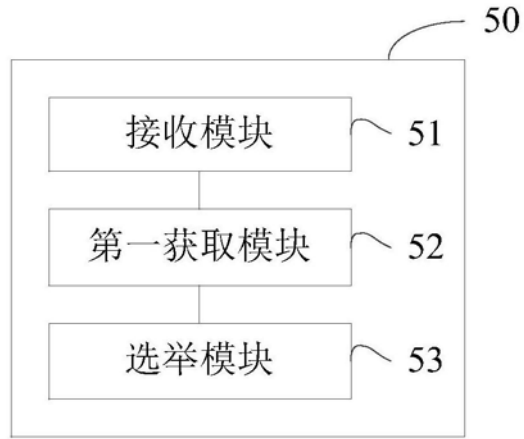


图5

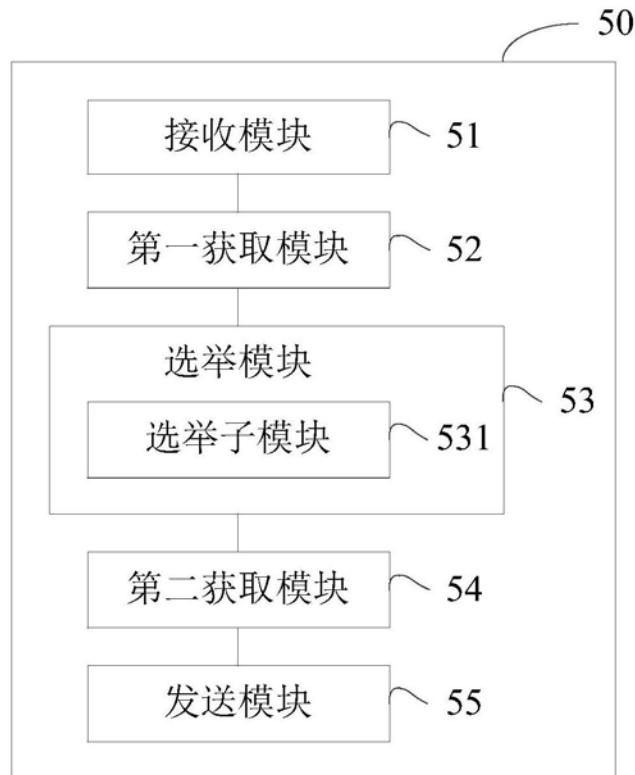


图6

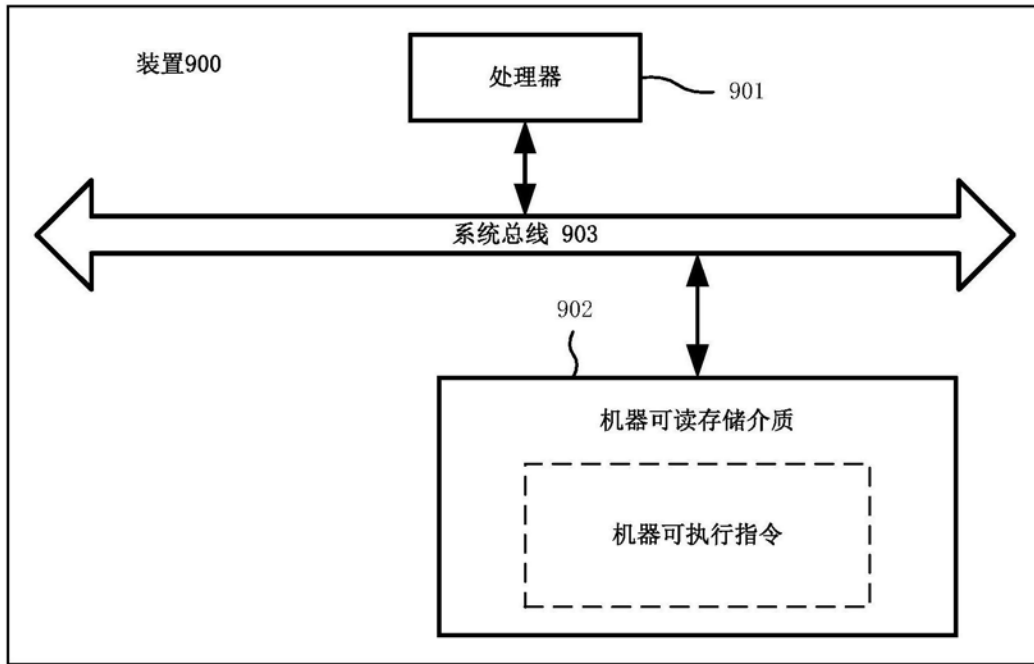


图7