



(12) 发明专利申请

(10) 申请公布号 CN 102244589 A

(43) 申请公布日 2011. 11. 16

(21) 申请号 201110202642. 6

(22) 申请日 2011. 07. 19

(71) 申请人 北京星网锐捷网络技术有限公司
地址 100036 北京市海淀区复兴路 33 号翠
微大厦东 1106

(72) 发明人 石伟玉 陈佑建

(74) 专利代理机构 北京同立钧成知识产权代理
有限公司 11205

代理人 刘芳

(51) Int. Cl.

H04L 12/24 (2006. 01)

H04L 12/56 (2006. 01)

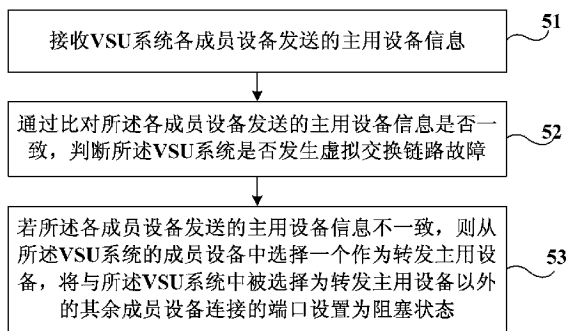
权利要求书 2 页 说明书 10 页 附图 4 页

(54) 发明名称

处理虚拟交换单元系统中链路故障的方法及
对端设备

(57) 摘要

本发明涉及一种处理虚拟交换单元系统中链
路故障的方法及对端设备,方法包括:接收虚拟
交换单元系统各成员设备发送的主用设备信息;
通过比对所述各成员设备发送的主用设备信息
是否一致,判断所述虚拟交换单元系统是否发生
虚拟交换链路故障;若所述各成员设备发送的主
用设备信息不一致,则从所述虚拟交换单元系统
的成员设备中选择一个作为转发主用设备,将与
所述虚拟交换单元系统中被选择为转发主用设备
以外的其余成员设备连接的端口设置为阻塞状
态。解决了在 VSL 链路故障后当主用设备出现
故障,进入恢复模式的成员设备无法接替成为
新主用设备的问题,增强了 VSU 组网应用的
可靠性,有效提高了 VSU 网络的自行修复抗
故障能力。



1. 一种处理虚拟交换单元系统中链路故障的方法,其特征在于,包括:
接收虚拟交换单元系统各成员设备发送的主用设备信息;
通过比对所述各成员设备发送的主用设备信息是否一致,判断所述虚拟交换单元系统是否发生虚拟交换链路故障;
若所述各成员设备发送的主用设备信息不一致,则从所述虚拟交换单元系统的成员设备中选择一个作为转发主用设备,将与所述虚拟交换单元系统中被选择为转发主用设备以外的其余成员设备连接的端口设置为阻塞状态。
2. 根据权利要求1所述的处理虚拟交换单元系统中链路故障的方法,其特征在于,还包括:
再次接收所述虚拟交换单元系统各成员设备发送的主用设备信息;
通过比对所述各成员设备发送的主用设备信息是否一致,判断所述虚拟交换单元系统是否发生虚拟交换链路故障;
若所述各成员设备发送的主用设备信息一致,则将设置为阻塞状态的端口恢复为转发状态。
3. 根据权利要求1或2所述的处理虚拟交换单元系统中链路故障的方法,其特征在于,接收虚拟交换单元系统各成员设备发送的主用设备信息的同时,还包括:接收所述各成员设备发送的链路故障信息,所述链路故障信息为与发送所述链路故障信息的成员设备之间的链路发生故障的上下联对端设备的数量。
4. 根据权利要求3所述的处理虚拟交换单元系统中链路故障的方法,其特征在于,从所述虚拟交换单元系统的成员设备中选择一个作为转发主用设备的过程,包括:从所述虚拟交换单元系统的成员设备中,选择发送的链路故障信息中链路发生故障的上下联对端设备的数量最少的成员设备作为转发主用设备。
5. 根据权利要求3所述的处理虚拟交换单元系统中链路故障的方法,其特征在于,在所述各成员设备发送的链路故障信息一致的情况下,从所述虚拟交换单元系统的成员设备中选择一个作为转发主用设备之前,还包括:
判断不一致的所述主用设备信息是否是在收到所述各成员设备发送的主用设备信息一致后第一次接收到的主用设备信息,所述主用设备信息为作为主用设备的成员设备编号;
从所述虚拟交换单元系统的成员设备中选择一个作为转发主用设备的过程,包括:
若是第一次,则从所述虚拟交换单元系统中选择发送的主用设备信息的值最小的成员设备作为转发主用设备,否则,保持之前选择的转发主用设备不变。
6. 根据权利要求1或2所述的处理虚拟交换单元系统中链路故障的方法,其特征在于,将自身与所述虚拟交换单元系统成员设备相连的、物理状态为开启且未收到所述虚拟交换单元系统成员设备的主用设备信息的端口设置为阻塞状态。
7. 根据权利要求1或2所述的处理虚拟交换单元系统中链路故障的方法,其特征在于,还包括:
在与选择的转发主用设备相连的端口的物理状态转变为关闭状态的情况下,将自身收到其余成员设备发送的主用设备信息的端口设置为转发状态。
8. 一种虚拟交换单元系统的上下联对端设备,其特征在于,包括:

信息接收单元,用于接收虚拟交换单元系统各成员设备发送的主用设备信息;

设备信息判断单元,用于通过比对所述各成员设备发送的主用设备信息是否一致,判断所述虚拟交换单元系统是否发生虚拟交换链路故障;

处理单元,用于若所述各成员设备发送的主用设备信息不一致,则从所述虚拟交换单元系统的成员设备中选择一个作为转发主用设备,将与所述虚拟交换单元系统中被选择为转发主用设备以外的其余成员设备连接的端口设置为阻塞状态。

9. 根据权利要求 8 所述的虚拟交换单元系统的上下联对端设备,其特征在于,

所述信息接收单元还用于再次接收所述虚拟交换单元系统各成员设备发送的主用设备信息;

所述设备信息判断单元还用于通过比对所述各成员设备发送的主用设备信息是否一致,判断所述虚拟交换单元系统是否发生虚拟交换链路故障;

所述处理单元还用于若所述各成员设备发送的主用设备信息一致,则将设置为阻塞状态的端口恢复为转发状态。

10. 根据权利要求 8 或 9 所述的虚拟交换单元系统的上下联对端设备,其特征在于,所述信息接收单元还用于在接收虚拟交换单元系统各成员设备发送的主用设备信息的同时,接收所述各成员设备发送的链路故障信息,所述链路故障信息为与发送所述链路故障信息的成员设备之间的链路发生故障的上下联对端设备的数量。

11. 根据权利要求 10 所述的虚拟交换单元系统的上下联对端设备,其特征在于,所述处理单元具体用于从所述虚拟交换单元系统的成员设备中,选择发送的链路故障信息中链路发生故障的上下联对端设备的数量最少的成员设备作为转发主用设备。

12. 根据权利要求 10 所述的虚拟交换单元系统的上下联对端设备,其特征在于,还包括:

接收次数判断单元,用于在所述各成员设备发送的链路故障信息一致的情况下,在从所述虚拟交换单元系统的成员设备中选择一个作为转发主用设备之前,判断不一致的所述主用设备信息是否是在收到所述各成员设备发送的主用设备信息一致后第一次接收到的主用设备信息,所述主用设备信息为作为主用设备的成员设备编号;

所述处理单元,具体用于若是第一次,则从所述虚拟交换单元系统中选择发送的主用设备信息的值最小的成员设备作为转发主用设备,否则,保持之前选择的转发主用设备不变。

13. 根据权利要求 8 或 9 所述的虚拟交换单元系统的上下联对端设备,其特征在于,所述处理单元还用于将物理状态为开启且未收到所述虚拟交换单元系统成员设备的主用设备信息的端口设置为阻塞状态。

14. 根据权利要求 8 或 9 所述的虚拟交换单元系统的上下联对端设备,其特征在于,所述处理单元还用于在与选择的转发主用设备相连的端口的物理状态转变为关闭状态的情况下,将自身收到其余成员设备发送的主用设备信息的端口设置为转发状态。

处理虚拟交换单元系统中链路故障的方法及对端设备

技术领域

[0001] 本发明涉及虚拟交换单元技术,尤其涉及一种处理虚拟交换单元系统中链路故障的方法及虚拟交换单元系统的上下联对端设备。

背景技术

[0002] 虚拟交换单元 (Virtual Switch Unit, VSU) 系统是将传统网络结构中的两台或者多台设备通过虚拟交换链路 (Virtual Switching Link, VSL) 连接在一起,组成单一的逻辑实体,易于增加网络设备的带宽、提高报文处理能力、以及提供更加稳定的多链路备份。

[0003] 如图 1 所示,汇聚层的 VSU 系统可以看作一台设备,与核心层、接入层进行交互。接入层的设备通过二层聚合链路接入到 VSU 汇聚层,VSU 汇聚层通过三层聚合链路接入到核心层的设备,即接入层和核心层设备通常都是通过聚合链路连接到 VSU 系统中。VSU 系统的成员设备之间通过内部 VSL 组成逻辑实体,并通过一定的选举协议确定主备身份:其中一台 VSU 成员设备选举为主用设备,其余的作为备用设备。

[0004] 以两台成员设备组成的 VSU 系统为例,如果主用设备和备用设备之间的 VSL 链路出现故障,VSU 系统将分成两个工作在 VSU 模式下的独立设备,备用设备也会切换成为主用设备,但由于原来的主用设备仍然在运行,此时 VSU 系统中就会存在两台全局配置完全相同的主用设备,这些主用设备同时接入到网络中,引起 IP 地址冲突、协议计算错误等一系列问题。因此,对于 VSU 系统需要提供一种检测出现多主用设备冲突的方法。

[0005] 现有技术中,针对上述 VSU 系统可能出现多个主用设备的问题,通常采用多 Active 检测 (Multi-Active Detection, MAD) 机制来解决。MAD 机制主要有:利用双向转发检测 (Bidirectional Forwarding Detection, BFD) 技术、利用链路聚合控制协议 (Link Aggregation Control Protocol, LACP) 检测技术及利用免费地址解析协议 (Address Resolution Protocol, ARP) 检测技术等。

[0006] BFD 技术如图 2 所示,在 VSU 系统的两台成员设备上各选取一个路由口,用来配置不同网段的 IP 地址,以建立连接,组成用于传输 BFD 检测消息的 BFD 心跳线。当 VSL 正常工作时,BFD 会话处于关闭 (down) 状态;当 VSL 出现故障时,BFD 检测功能被激活,两台成员设备通过 BFD 心跳线建立会话交换信息,检测出 VSU 系统当前存在双主用设备的情况。当检测出双主用设备时,非主用设备关闭自身的除了 VSL 端口和例外端口之外的所有端口,进入恢复模式。其中,例外端口是在配置 VSU 系统时,事先把一些端口定义为例外端口,以在进入恢复模式时不关闭该端口。

[0007] LACP 检测技术如图 3 所示,VSU 系统与上下联对端设备进行跨成员设备聚合,VSU 系统的各成员设备向上下联对端设备发送 LACP 报文。该 LACP 报文预留字段中携带 ACTIVE_ID 参数,ACTIVE_ID 表示 VSU 主用设备号,其值由作为主用设备的成员设备编号决定。当上下联对端设备接收到该 LACP 报文,发现其中的预留字段不为 0,确定该 LACP 报文的发送端是 VSU 系统的成员设备,则将该 LACP 报文从其他汇聚成员端口转发出去。当 VSU 系统的成员设备收到 LACP 报文后,把接收到的 LACP 报文中的 ACTIVE_ID 参数的值与自身 ACTIVE_

ID 进行比较,如果相同,则表示 VSL 正常,VSU 系统中只有一个主用设备;如果不同,则表示 VSL 故障,VSU 系统中有多个主用设备,从而检测出冲突。当检测出多主用设备冲突,选取 ACTIVE_ID 较小的成员设备继续工作,其余成员设备关闭本地除了 VSL 端口和例外端口之外的所有业务接口进入恢复模式。

[0008] 免费 ARP 检测技术如图 4 所示,在 VSU 系统的成员设备之间增加一条 ARP 检测链路,并对免费 ARP 报文进行扩展,在扩展字段中携带当前 VSU 系统主用设备的 ACTIVE_ID 参数,方法与 LACP 检测技术类似。当检测出多主用设备冲突,选取 ACTIVE_ID 较小的设备继续工作,其余成员设备关闭本地除了 VSL 端口和例外端口之外的所有业务接口进入恢复模式。

[0009] 现有技术存在的缺陷在于:当检测出多主用设备时,被选择为非主用的设备关闭除了 VSL 端口和例外端口之外的所有业务接口进入恢复模式,以保证 VSU 系统中只有一个主用设备正常工作,但是在 VSL 故障恢复前,当被选择为主用设备的成员设备运行也出现故障时,进入恢复模式的非主用设备将无法接替成为新的主用设备。

发明内容

[0010] 本发明提出一种处理虚拟交换单元系统中链路故障的方法及虚拟交换单元系统的上下联对端设备,以解决在 VSL 故障后当主用设备出现故障,进入恢复模式的成员设备无法接替成为新主用设备的问题。

[0011] 本发明提供了一种处理虚拟交换单元系统中链路故障的方法,包括:

[0012] 接收虚拟交换单元系统各成员设备发送的主用设备信息;

[0013] 通过比对所述各成员设备发送的主用设备信息是否一致,判断所述虚拟交换单元系统是否发生虚拟交换链路故障;

[0014] 若所述各成员设备发送的主用设备信息不一致,则从所述虚拟交换单元系统的成员设备中选择一个作为转发主用设备,将与所述虚拟交换单元系统中被选择为转发主用设备以外的其余成员设备连接的端口设置为阻塞状态。

[0015] 本发明还提供了一种虚拟交换单元系统的上下联对端设备,包括:

[0016] 信息接收单元,用于接收虚拟交换单元系统各成员设备发送的主用设备信息;

[0017] 判断单元,用于通过比对所述各成员设备发送的主用设备信息是否一致,判断所述虚拟交换单元系统是否发生虚拟交换链路故障;

[0018] 处理单元,用于若所述各成员设备发送的主用设备信息不一致,则从所述虚拟交换单元系统的成员设备中选择一个作为转发主用设备,将与所述虚拟交换单元系统中被选择为转发主用设备以外的其余成员设备连接的端口设置为阻塞状态。

[0019] 本发明提供的处理虚拟交换单元系统中链路故障的方法及虚拟交换单元系统的上下联对端设备,通过将未被选择为转发主用设备的其余成员设备相连的端口设置为阻塞状态,使得未被选择为转发主用设备的其余成员设备仍处于正常工作状态且不会转发报文影响转发主用设备的报文转发,解决了在 VSL 链路故障后当主用设备出现故障,进入恢复模式的成员设备无法接替成为新主用设备的问题,增强了 VSU 组网应用的可靠性,有效提高了 VSU 网络的自行修复抗故障能力。

附图说明

[0020] 为了更清楚地说明本发明实施例中的技术方案,下面将对实施例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性的前提下,还可以根据这些附图获得其他的附图。

[0021] 图 1 为汇聚层的 VSU 系统的结构示意图;

[0022] 图 2 为 BFD 技术的应用示意图;

[0023] 图 3 为 LACP 检测技术的应用示意图;

[0024] 图 4 为免费 ARP 检测技术的应用示意图;

[0025] 图 5 为本发明实施例提供的一种处理虚拟交换单元系统中链路故障的方法的流程图;

[0026] 图 6 为本发明实施例提供的另一种处理虚拟交换单元系统中链路故障的方法的流程图;

[0027] 图 7 为本发明实施例提供的处理虚拟交换单元系统中链路故障的方法中免费 ARP 报文格式示意图;

[0028] 图 8 为本发明实施例提供的处理虚拟交换单元系统中链路故障的方法应用的 VSU 系统示意图;

[0029] 图 9 为本发明实施例提供的处理虚拟交换单元系统中链路故障的方法中逻辑端口组的示意图;

[0030] 图 10a 为本发明实施例提供的处理虚拟交换单元系统中链路故障的方法的一种应用环境示意图;

[0031] 图 10b 为本发明实施例提供的处理虚拟交换单元系统中链路故障的方法的另一种应用环境示意图;

[0032] 图 11 为本发明实施例提供的虚拟交换单元系统的上下联对端设备的结构示意图。

具体实施方式

[0033] 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。

[0034] 图 5 为本发明实施例提供的一种处理虚拟交换单元系统中链路故障的方法的流程图。如图 5 所示,处理方法包括:

[0035] 步骤 51、接收 VSU 系统各成员设备发送的主用设备信息;

[0036] 步骤 52、通过比对所述各成员设备发送的主用设备信息是否一致,判断所述 VSU 系统是否发生虚拟交换链路故障;

[0037] 步骤 53、若所述各成员设备发送的主用设备信息不一致,则从所述 VSU 系统的成员设备中选择一个作为转发主用设备,将与所述 VSU 系统中被选择为转发主用设备以外的其余成员设备连接的端口设置为阻塞状态。

[0038] 本实施例中,当VSL故障导致VSU系统中出现多主用设备时,上下联对端设备根据收到的主用设备信息,检测出VSU系统中存在多主用设备的情况,从而选举出一台转发主用设备为网络提供转发任务,并把自身与其余VSU系统成员设备相连的汇聚成员端口设置成阻塞(block)状态。这样,处于阻塞状态的汇聚成员端口将不能转发报文,只能接收报文,所有的数据流报文就只能经由转发主用设备进行转发。

[0039] 本发明实施例提供的处理VSU系统中链路故障的方法还可包括:

[0040] 再次接收所述虚拟交换单元系统各成员设备发送的主用设备信息;

[0041] 通过比对所述各成员设备发送的主用设备信息是否一致,判断所述VSU系统是否发生虚拟交换链路故障;

[0042] 若所述各成员设备发送的主用设备信息一致,则将设置为阻塞状态的端口恢复为转发状态。

[0043] 接收VSU系统各成员设备发送的主用设备信息的同时,还可包括:接收所述各成员设备发送的链路故障信息,所述链路故障信息为与发送所述链路故障信息的成员设备之间的链路发生故障的上下联对端设备的数量。

[0044] 从所述VSU系统的成员设备中选择一个作为转发主用设备的过程,包括:从所述虚拟交换单元系统的成员设备中,选择发送的链路故障信息中链路发生故障的上下联对端设备的数量最少的成员设备作为转发主用设备。

[0045] 从所述VSU系统的成员设备中选择一个作为转发主用设备之前,还可包括:

[0046] 判断不一致的所述主用设备信息是否是在收到所述各成员设备发送的主用设备信息一致后,第一次接收到的主用设备信息,所述主用设备信息为作为主用设备的成员设备编号;

[0047] 从所述VSU系统的成员设备中选择一个作为转发主用设备的过程,包括:

[0048] 若是第一次,则从所述VSU系统中选择发送的主用设备信息的值最小的成员设备作为转发主用设备,否则,保持之前选择的主用设备不变。

[0049] 上述步骤51~步骤53的执行主体可为上述步骤51~步骤53中VSU系统的上下联对端设备。

[0050] 上述主用设备信息及链路故障信息可通过链路聚合控制协议(Link Aggregation Control Protocol, LACP)、双向转发检测(Bidirectional Forwarding Detection, BFD)、链路层发现协议(Link Layer Discovery Protocol, LLDP)、标签分发协议(Label Distribution Protocol, LDP)、免费ARP等协议,从VSU系统的各成员设备发送到上下联对端设备。

[0051] 在VSL故障期间,当进行转发的转发主用设备也出现故障时,上下联对端设备会根据收到的主用设备信息选举出新的转发主用设备。具体地,将与出现故障的转发主用设备互连的汇聚成员端口的状态设置成阻塞(block)状态,并将与重新选举出的转发主用设备互连的汇聚成员端口从阻塞(block)状态转化为转发(forwarding)状态。这样,后续所有的数据流报文仅经由重新选举出的转发主用设备进行转发。

[0052] 当VSL故障恢复,VSU系统中各成员设备通过VSL恢复成正常的VSU主备模式,则各成员设备后续发送的主用设备信息一致,上下联对端设备收到后,把所有汇聚成员端口中处于阻塞(block)状态的成员端口都转为转发(forwarding)状态。

[0053] 需要注意的是：转发主用设备是由 VSU 系统的上下联对端设备选择出来的，而主用设备是 VSU 系统内部各成员设备通过 VSL 选举出来或者在 VSL 故障时备用设备自动切换得到的。

[0054] 图 6 为本发明实施例提供的另一种处理虚拟交换单元系统中链路故障的方法的流程图。本实施例中，VSU 系统中各成员设备定期向上下联设备发送免费 ARP 报文，也就是说，VSU 系统中各成员设备采用免费 ARP 协议向上下联对端设备发送主用设备信息及链路故障信息。如图 6 所示，处理方法包括：

[0055] 步骤 61、VSU 系统的上下联对端设备接收 VSU 系统中各成员设备发送的免费 ARP 报文。

[0056] 免费 ARP 报文格式如图 7 所示，扩展了字段，用来携带主用设备信息和链路故障信息。用来携带主用设备信息的字段定义为 ACTIVE_ID，当主设备信息为 VSU 系统中作为主用设备的成员设备编号时，ACTIVE_ID 的值为成员设备编号。例如编号为 1 的成员设备作为主用设备时，ACTIVE_ID 的值为 1，编号为 2 的成员设备作为主用设备时，ACTIVE_ID 的值为 2。本领域技术人员应理解为，主用设备信息还可为 VSU 系统中作为主用设备的成员设备的 MAC 地址等，只要通过 ACTIVE_ID 的值能够确定 VSU 系统中哪一个成员设备为主用设备即可。

[0057] 当上下联对端设备从 VSU 系统各成员设备接收到的免费 ARP 报文的 ACTIVE_ID 值相同时，表明此时 VSU 系统只有一个主用设备，VSL 正常；当上下联对端设备从 VSU 系统各成员设备接收到的免费 ARP 报文的 ACTIVE_ID 值不同时，表明此时 VSU 系统存在多个主用设备，VSL 故障。

[0058] 用来携带链路故障信息的字段为 FAULT。如 FAULT 的值为 0，表示发送该免费 ARP 报文的成员设备与上下联对端设备之间的互联链路均正常，没有出现故障；该 FAULT 的值为非 0，表示发送该免费 ARP 报文的成员设备与上下联对端设备之间的互联链路发生故障，且该 FAULT 的值等于与发送该免费 ARP 报文的成员设备之间链路发生故障的上下联对端设备的个数。

[0059] 这里，VSU 系统的成员设备与上下联对端设备之间链路发生的故障，指一个成员设备与一个上下联对端设备之间的所有链路都断开，对应的所有成员端口的状态都为关闭 (down) 状态。例如，当 VSU 系统中，某一成员设备与一个上下联对端设备间的链路发生故障，则该成员设备能够发送出去的免费 ARP 报文中 $FAULT = 1$ 。相应地，该成员设备与 N 个上下联对端设备之间链路发生故障，则该成员设备能够发送出去的免费 ARP 报文中 $FAULT = N$ 。

[0060] 免费 ARP 报文的原始部分中，发送端 IP 和目的 IP 都用发送报文端自己的 IP，源 MAC 是发送免费 ARP 报文的成员设备的 MAC，目的 MAC 是广播。

[0061] 步骤 62、上下联对端设备比对各成员设备发送的免费 ARP 报文中 ACTIVE_ID 的值是否一致，若一致，说明 VSL 正常，保持自身与 VSU 系统各成员设备连接的端口状态不变；若不一致，说明发送免费 ARP 报文的 VSU 系统中 VSL 故障，则执行步骤 63。

[0062] 步骤 63、上下联对端设备从发送免费 ARP 报文的 VSU 系统中选择一个成员设备作为转发报文的转发主用设备，将与所述 VSU 系统中被选择为转发主用设备以外的其余成员设备连接的端口设置为阻塞 (block) 状态。

[0063] 具体地,上下联对端设备先比较各成员设备发送的免费 ARP 报文中 FAULT 的值,当 FAULT 值不同时,从 VSU 系统中选择发送的免费 ARP 报文中 FAULT 值最小的成员设备作为转发报文的转发主用设备,将 VSU 系统中的其余成员设备作为非转发主用设备,将与转发主用设备相连的端口设置为转发 (forwarding) 状态,并将自身与非转发主用设备互联的汇聚成员端口设置成阻塞 (block) 状态。

[0064] 当 VSU 系统中各成员设备发出的免费 ARP 报文中 FAULT 的值相同时,可以从 VSU 系统中选择任意一台成员设备作为转发报文的转发主用设备。为了避免频繁的二次切换,在选择转发主用设备前,可先判断接收到的 VSU 系统各成员设备发送的前一个免费 ARP 报文的 ACTIVE_ID 值是否一致,也即判断当前接收的免费 ARP 报文是否是在收到所述各成员设备发送的主用设备信息一致后第一次接收到的免费 ARP 报文,若 VSU 系统各成员设备发送的前一个免费 ARP 报文的 ACTIVE_ID 值一致,也即上述步骤 61 中上下联对端设备接收的免费 ARP 报文是在收到所述各成员设备发送的主用设备信息一致后第一次接收到的免费 ARP 报文,则从 VSU 系统中选取发送免费 ARP 报文中 ACTIVE_ID 最小的成员设备作为转发主用设备;否则,说明 VSU 系统发生 VSU 故障上下联对端设备选择转发主用设备时,非转发主用设备与上下联对端设备之间的链路发生故障,之后非转发主用设备与上下联对端设备之间的链路故障恢复,而 VSU 系统中 VSL 故障仍未恢复,导致未被选择为转发主用设备的成员设备恢复发送免费 ARP 报文,使得上下联对端设备接收到转发主用设备及 VSU 系统中其余成员设备发送的 ACTIVE_ID 值不同的多个免费 ARP 报文,这种情况下,保持之前选择的转发主用设备不变,以避免重复切换。

[0065] VSU 系统中,每个成员设备与上下联对端设备互联的汇聚接口可能存在多个成员端口,为了便于描述,将 VSU 系统的每个成员设备分别与上下联对端设备互联的汇聚接口可能包含的所有成员端口,视为一个逻辑端口组。本领域技术人员应理解为:上联或下联对端设备上某一个逻辑端口组在一个周期内未收到免费 ARP 报文,如可能由于当 VSU 系统中主用设备重启、主用设备与上下联对端设备相连的链路断开或是主用设备 CPU 处理负荷太高等原因而导致,则对该未收到免费 ARP 报文的逻辑端口组作以下处理:若该逻辑端口组物理状态为关闭 (down) 状态,则不作处理;若逻辑端口组物理状态为开启 (up) 状态,则设置为阻塞 (block) 状态。上联或下联对端设备对于其余有收到免费 ARP 报文的逻辑端口组作以下处理:如果有多个逻辑端口组收到免费 ARP 报文,则按以上比较免费 ARP 报文参数方式处理;如果只剩一个逻辑端口组收到免费 ARP 报文,则不管该逻辑端口组之前是什么状态,直接设置为转发 (forwarding) 状态,与该端口相连的主用设备即成为转发主用设备。

[0066] 步骤 64、上下联对端设备再次接收 VSU 系统各设备成员发送的免费 ARP 报文。

[0067] 具体地,当 VSU 系统中的 VSL 恢复时,各主用设备即 VSU 系统中的所有成员设备通过 VSL 重新选举,其中一个成员设备被选举为主用设备,其余未被选举为主用设备的成员设备自动重启,重启完成后,成为 VSU 系统的备用设备。VSU 系统中,重新恢复成主备模式的各成员设备均向上行联对端设备发送免费 ARP 报文。

[0068] 步骤 65、上下联对端设备根据再次接收到的免费 ARP 报文修改相应的端口状态。

[0069] 由于 VSL 恢复后,只有一个成员设备选举为主用设备,其余成员设备都为备用设备,因此,各成员设备发出的免费 ARP 报文中 ACTIVE_ID 的值都一样。如选举出编号为 1 的成员设备为主用设备,则各成员设备发送的免费 ARP 报文中的 ACTIVE_ID 的值都为 1;如选

举出编号为 2 的成员设备为主用设备,则各成员设备发送的免费 ARP 报文中的 ACTIVE_ID 的值都为 2。

[0070] 上下联对端设备在各成员设备发送的免费 ARP 报文中 ACTIVE_ID 的值相同的情况下,把所有与 VSU 系统相连且处于阻塞状态的端口都置为转发状态。

[0071] 本实施例中,免费 ARP 报文由 VSU 系统各成员设备周期性发送,发送周期可以根据实际应用如网络中断时间的要求进行调节。对上下联对端设备而言,当 VSL 故障时,VSU 成员设备发送的免费 ARP 报文中 ACTIVE_ID 的值不同,可以将免费 ARP 报文称为 VSL 链路故障通告报文;当 VSL 链路恢复,VSU 成员设备发送的免费 ARP 报文中 ACTIVE_ID 的值相同,可以把免费 ARP 报文称为 VSL 恢复通告报文。

[0072] 在 VSL 故障期间,当正在运行的转发主用设备出现整机故障时,则与该转发主用设备相连的上下联对端设备的端口状态变为关闭(down)状态,收不到任何通告报文。如果上下联对端设备上与另一台主用设备相连的端口可以收到通告报文,则该端口自动由阻塞(block)状态转成转发(forwarding)状态,实现转发主用设备的切换。当出现整机故障的主用设备重新启动时,因检测到 VSL 故障,则启动后仍然为 VSU 系统中的主用设备,与转发主用设备一起,继续发送链路故障通告报文,再由上下联对端设备选举转发主用设备。

[0073] 当 VSL 故障恢复时,各主用设备则通过 VSL 实现主备选举,恢复到 VSU 系统的主备模式,并继续发送 VSL 恢复通告报文给上下联对端设备。上下联对端设备收到 VSL 恢复通告报文后,把自身所有与 VSU 系统相连的端口都转为转发(forwarding)状态。因此,不管 VSL 是否有故障,本实施例都能实现主备设备间的转换,即:VSL 无故障时,通过 VSL 选举主用设备和备用设备;VSL 故障期间,通过上下联对端设备修改自身端口状态,选择转发主用设备。

[0074] 下面以如图 8 所示的 VSU 系统为例,对处理 VSU 系统中链路故障的方法进行更为详细说明。VSU 系统包括两个成员设备:设备编号 1 的汇聚层设备 A 和设备编号 2 的设备 B,汇聚层设备 A 与汇聚层设备 B 之间通过 VSL 连接。下联对端设备 D 通过二层聚合链路接入到 VSU 汇聚层,VSU 汇聚层通过三层聚合链路接入到上联对端设备 C。C1、D1 分别为上下联对端设备 C、D 与汇聚层设备 A 互联的逻辑端口组;C2、D2 分别为上下联对端设备 C、D 与汇聚层设备 B 互联的逻辑端口组。如图 9 所示,每个 VSU 成员端口与上下联对端设备互联都通过一个逻辑端口组,故障通告报文和故障恢复通告报文均在逻辑端口组上进行收发:当上下联对端设备从逻辑端口组中的任何 1 个成员端口收到故障通告报文或故障恢复通告报文时,即认为收到从该 VSU 系统成员设备发送的通告报文。

[0075] 假设 VSU 系统开始运行时,汇聚层设备 A 为主用设备,汇聚层设备 B 为备用设备。

[0076] 当汇聚层设备 A、汇聚层设备 B 与上联对端设备 C、下联对端设备 D 相连的链路正常时,执行如下操作:

[0077] 当 VSL 正常时,汇聚层设备 A 和汇聚层设备 B 周期性的发送通告报文给上联对端设备 C、下联对端设备 D。二者发送的通告报文中扩展字段的值相同,均为:FAULT = 0 ACTIVE_ID = 1。上下联对端设备确定收到的是 ACTIVE 值相同的 VSL 恢复通告报文,自身又没有处于阻塞(block)状态的汇聚成员端口,因此,对接收到的 VSL 恢复通告报文不作处理。

[0078] 当 VSL 故障时,作为备用设备的汇聚层设备 B 由于无法通过 VSL 接收到作为主用

设备的汇聚层设备 A 发来的信息,就认为作为主用设备的汇聚层设备 A 出现故障,而自己切换成为新主用设备。作为主用设备的汇聚层设备 A、汇聚层设备 B 向上联对端设备 C、下联对端设备 D 发送故障通告报文。汇聚层设备 A 发送的通告报文中 :FAULT = 0 ACTIVE_ID = 1 ;汇聚层设备 B 发送的通告报文中 :FAULT = 0 ACTIVE_ID = 2。

[0079] 上联对端设备 C、下联对端设备 D 对比汇聚层设备 A 与汇聚层设备 B 发送的通告报文中的 ACTIVE_ID 的值,由于二者发送的通告报文中的 ACTIVE_ID 的值不同,因而可判定接收到的是故障通告报文。由于两个作为主用设备的汇聚层设备 A、汇聚层设备 B 发出的通告报文中 FAULT 参数值一样,则选取 ACTIVE_ID 值较小的汇聚层设备 A 为转发主用设备。因此上联对端设备 C、下联对端设备分别把与汇聚层设备 A 相连的逻辑端口组 C1、逻辑端口组 D1 设置为转发 (forwarding) 状态,把与汇聚层设备 B 相连的逻辑端口组 C2、逻辑端口组 D2 置为阻塞 (block) 状态。

[0080] VSU 系统继续运行一段时间后,当作为转发主用设备的汇聚层设备 A 出现如掉电等整机故障,而导致上联对端设备 C、下联对端设备 D 与汇聚层设备 A 相连的逻辑端口组 C1、逻辑端口组 D1 的状态变为关闭 (down)。此时,只有与汇聚层设备 B 相连的逻辑端口组 C2、逻辑端口组 D2 能收到通告报文。则把与汇聚层设备 B 相连的逻辑端口组 C2、逻辑端口组 D2 由阻塞 (block) 状态转为转发 (forwarding) 状态,切换汇聚层设备 B 成为转发主用设备。

[0081] 备注 :这里整机故障,简单指设备掉电等故障。

[0082] 当作为非转发主用设备的汇聚层设备 A 整机重启,故障恢复后,会继续和作为转发主用设备的汇聚层设备 B 一起发送故障通告报文给上下联对端设备。汇聚层设备 A 发送的通告报文中 :FAULT = 0 ACTIVE_ID = 1,汇聚层设备 B 发送的通告报文中 :FAULT = 0 ACTIVE_ID = 2。由于上联对端设备 C、下联对端设备 D 不是第一次在收到 VSL 恢复通告报文后比较故障通告报文,此时即使故障通告报文中的 FAULT 值一样,也不比较第二个参数 ACTIVE_ID,还是由汇聚层设备 B 作为转发主用设备。

[0083] 备注 :汇聚层设备 A 整机重启时,因检测到 VSL 故障,则启动后还是为 VSU 系统中的主用设备。

[0084] 当 VSL 故障恢复,汇聚层设备 A、汇聚层设备 B 通过 VSL 恢复成 VSU 主备模式,并同时发送 ACTIVE_ID 值一样的 VSL 恢复通告报文给上联对端设备 C、下联对端设备 D。上联对端设备 C、下联对端设备 D 收到 VSL 恢复通告报文后,把与汇聚层设备 A、汇聚层设备 B 相连且处于阻塞 (block) 状态的逻辑端口组 C1、逻辑端口组 D1 都变为转发 (forwarding) 状态,恢复正常的 VSU 主备模式。

[0085] 备注 :如果 VSL 恢复,则汇聚层设备 A 和汇聚层设备 B 会重新进行 VSU 主备选举,被选为备用的汇聚层设备将重启。备用设备重启过程中,只有主用设备会发送通告报文,上下联对端设备对主用设备发送的通告报文处理方式类似上述作为转发主用设备时的汇聚层设备 A 出现如掉电等整机故障的处理方式。

[0086] 利用上述处理故障的方法,当网络中可能出现如下故障时,处理方式如下 :

[0087] 第一种故障情况 :当 VSL 出现故障的同时,汇聚层设备 A、汇聚层设备 B 也出现整机故障,那么,上下联对端设备逻辑端口组接口中无法收到报文,不作处理。

[0088] 第二种故障情况 :当 VSL 出现故障的同时,汇聚层设备 A 与上联对端设备 C 之间的

链路也有故障。此时,汇聚层设备 A 发送的通告报文中 :FAULT = 1 ACTIVE_ID = 1 ;汇聚层设备 B 发送的通告报文中 :FAULT = 0 ACTIVE_ID = 2。由于上联对端设备 C 的逻辑端口组 C1 接口状态为关闭 (down) 状态,无法收到故障通告报文,则不作处理,逻辑端口组 C2 收到通告报文中 FAULT = 0 ACTIVE_ID = 2,则设置为转发 (forwarding) 状态 ;下联对端设备 D 的逻辑端口组 D1 收到通告报文中 FAULT = 1 ACTIVE_ID = 1,逻辑端口组 D2 收到通告报文中 FAULT = 0 ACTIVE_ID = 2,通过比较通告报文中 FAULT 的值,把收到的通告报文中 FAULT 值较大的逻辑端口组 D1 置为阻塞 (block) 状态,把收到的通告报文中 FAULT 值较小的逻辑端口组 D2 置为转发 (forwarding) 状态。此时上联对端设备 C、下联对端设备 D 都选择汇聚层设备 B 作为转发主用设备。

[0089] 第三种故障情况 :当 VSL 正常,而汇聚层设备 A 与上联对端设备 C 之间的链路出现故障。此时,A 发送的通告报文 :FAULT = 1 ACTIVE_ID = 1 ;B 发送的通告报文 :FAULT = 0 ACTIVE_ID = 1。上联对端设备 C 的逻辑端口组 C1 收不到故障通告报文,则不作处理,逻辑端口组 C2 收到的通告报文中 FAULT = 0 ACTIVE_ID = 1,由于逻辑端口组 C2 本身就处于转发 (forwarding) 状态,因此,也不作处理 ;下联对端设备 D 逻辑端口组 D1 收到的通告报文中 FAULT = 1 ACTIVE_ID = 1,逻辑端口组 D2 收到的通告报文中 FAULT = 0 ACTIVE_ID = 1,由于两个逻辑端口组收到的通告报文中 ACTIVE_ID 值相同,确定通告报文为 VSL 恢复通告报文,且由于逻辑端口组 D1、逻辑端口组 D2 本身为转发 (forwarding) 状态,因此,不作处理。

[0090] 第四种故障情况 :VSL 故障期间,汇聚层设备 A 由于 CPU 利用率太高等原因而导致无法发送通告报文,则上下联对端设备按上述步骤 63 中的说明进行处理,选取汇聚层设备 B 作为转发主用设备。

[0091] 第五种故障情况 :VSL 故障期间,汇聚层设备 A 或汇聚层设备 B 与上联对端设备 C、下联对端设备 D 之间的链接出现出故障,处理过程类似上述第二种故障情况下的处理。

[0092] 第六种故障情况 :VSL 链路正常时,汇聚层设备 A 或汇聚层设备 B 与上联对端设备 C、下联对端设备 D 之间的链接出现出故障,处理过程类似上述第三种故障情况下的处理。

[0093] 上述处理故障的方法也可应用于如图 10a 所示的环境,即上下联对端设备可以是多台,但单个 VSU 成员设备与单个上下联对端设备连接的 AP 口内只有一个成员端口。当单个 VSU 成员设备与单个上下联设备连接的 AP 口内有多个成员端口时,如图 10b 所示,此时,与单个上下联设备连接的多个成员端口视为一个逻辑端口组,上述故障通告报文及 VSL 恢复通告报文在这个逻辑端口组上进行收发。

[0094] 上述实施例利用现有的网络结构,在 VSL 出现故障的情况下,实现多主用设备冲突的检测,并利用上下联对端设备对转发主用设备进行选取和切换,使得 VSU 系统中各成员设备间无需增加额外链路,并有效地解决了 VSU 系统出现多主用设备时所带来的 IP 冲突、网络不通等问题。

[0095] 本领域普通技术人员可以理解 :实现上述方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成,前述的程序可以存储于一计算机可读取存储介质中,该程序在执行时,执行包括上述方法实施例的步骤 ;而前述的存储介质包括 :ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

[0096] 图 11 为本发明实施例提供的虚拟交换单元系统的上下联对端设备的结构示意图

图。如图 11 所示,VSU 系统的上下联对端设备包括:信息接收单元 111、设备信息判断单元 112 及处理单元 113。

[0097] 信息接收单元 111 用于接收 VSU 系统各成员设备发送的主用设备信息;设备信息判断单元 112 用于通过比对所述各成员设备发送的主用设备信息是否一致,判断所述 VSU 系统是否发生 VSL 故障;处理单元 113 用于若所述各成员设备发送的主用设备信息不一致,则从所述 VSU 系统的成员设备中选择一个作为转发主用设备,将与所述 VSU 系统中被选择为转发主用设备以外的其余成员设备连接的端口设置为阻塞状态。

[0098] 所述信息接收单元 111 还可用于再次接收所述虚拟交换单系统各成员设备发送的主用设备信息;所述设备信息判断单元 112 还可用于通过比对所述各成员设备发送的主用设备信息是否一致,判断所述 VSU 系统是否发生 VSL 故障;所述处理单元 113 还可用于若所述各成员设备发送的主用设备信息一致,则将设置为阻塞状态的端口恢复为转发状态。

[0099] 所述信息接收单元 111 还可用于在接收 VSU 系统各成员设备发送的主用设备信息的同时,接收所述各成员设备发送的链路故障信息,所述链路故障信息为与发送所述链路故障信息的成员设备之间的链路发生故障的上下联对端设备的数量。

[0100] 所述处理单元 113 还可具体用于从所述 VSU 系统的成员设备中,选择发送的链路故障信息中链路发生故障的上下联对端设备的数量最少的成员设备作为转发主用设备。

[0101] 本发明实施例提供的 VSU 系统的上下联对端设备还可包括:接收次数判断单元,用于从所述 VSU 系统的成员设备中选择一个作为转发主用设备之前,判断不一致的所述主用设备信息是否是在收到所述各成员设备发送的主用设备信息一致后第一次接收到的主用设备信息,所述主用设备信息为作为主用设备的成员设备编号,详见上述方法实施例中的说明。

[0102] 所述处理单元 113 还可具体用于若是第一次,则从所述 VSU 系统中选择发送的主用设备信息的值最小的成员设备作为转发主用设备,否则,保持之前选择的转发主用设备不变。

[0103] 所述处理单元 113 还可用于将物理状态为开启且未收到所述 VSU 系统成员设备的主用设备信息的端口设置为阻塞状态。

[0104] 所述处理单元 113 还可用于若仅收到所述 VSU 系统中一个成员设备发送的主用设备信息,则将与所述 VSU 系统中其余成员设备连接的端口设置为阻塞状态。

[0105] 所述处理单元 113 还可用于在与选择的转发主用设备相连的端口的物理状态转变为关闭状态的情况下,将自身收到其余成员设备发送的主用设备信息的端口设置为转发状态。

[0106] 最后应说明的是:以上实施例仅用以说明本发明的技术方案,而非对其限制;尽管参照前述实施例对本发明进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本发明各实施例技术方案的精神和范围。

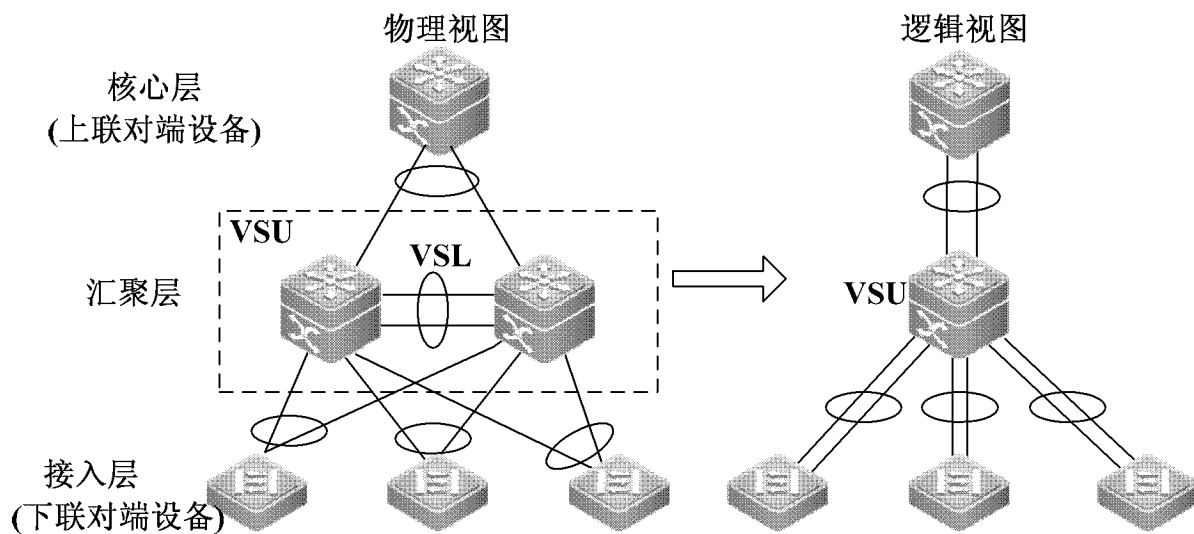


图 1

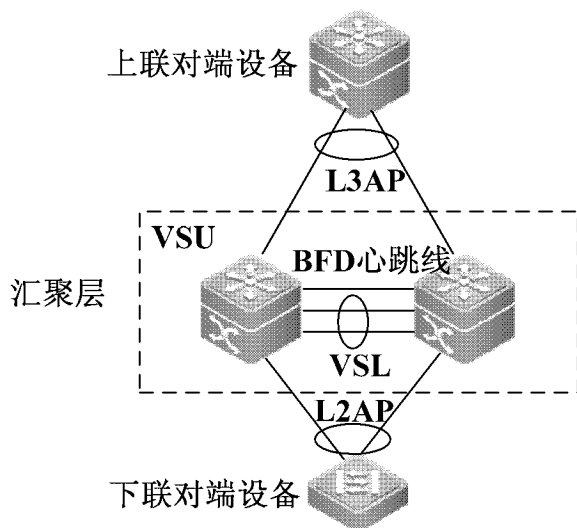


图 2

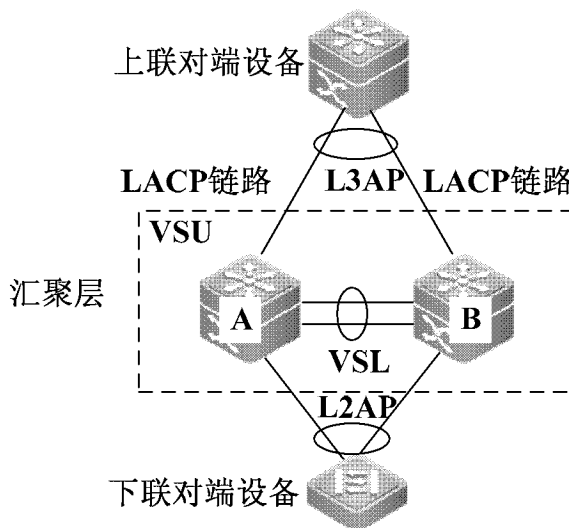


图 3

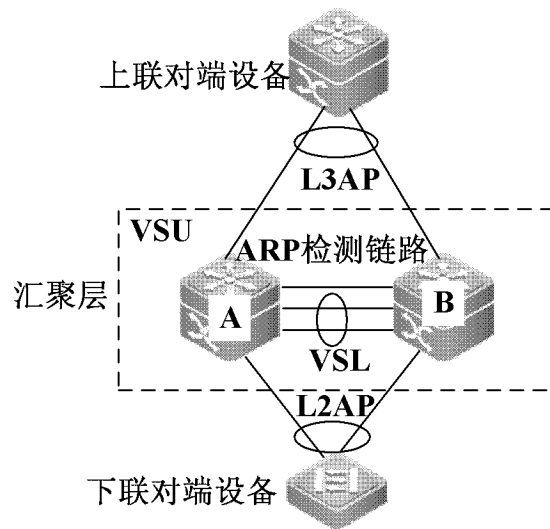


图 4

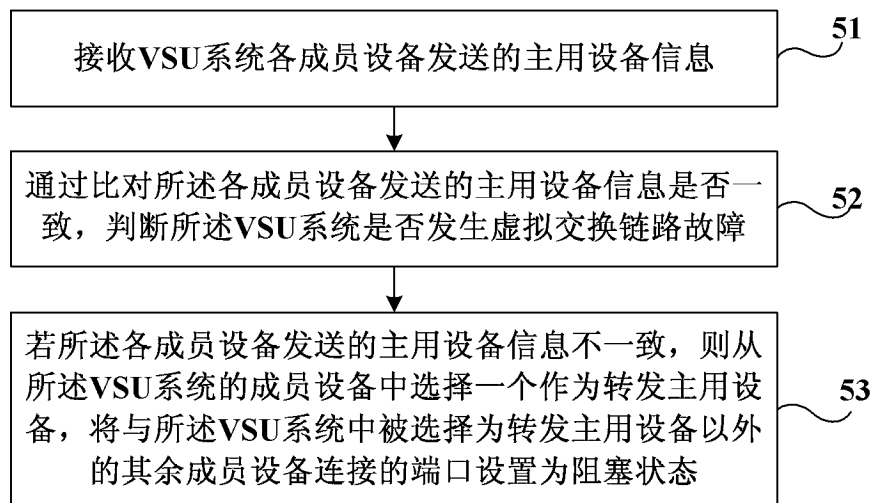


图 5

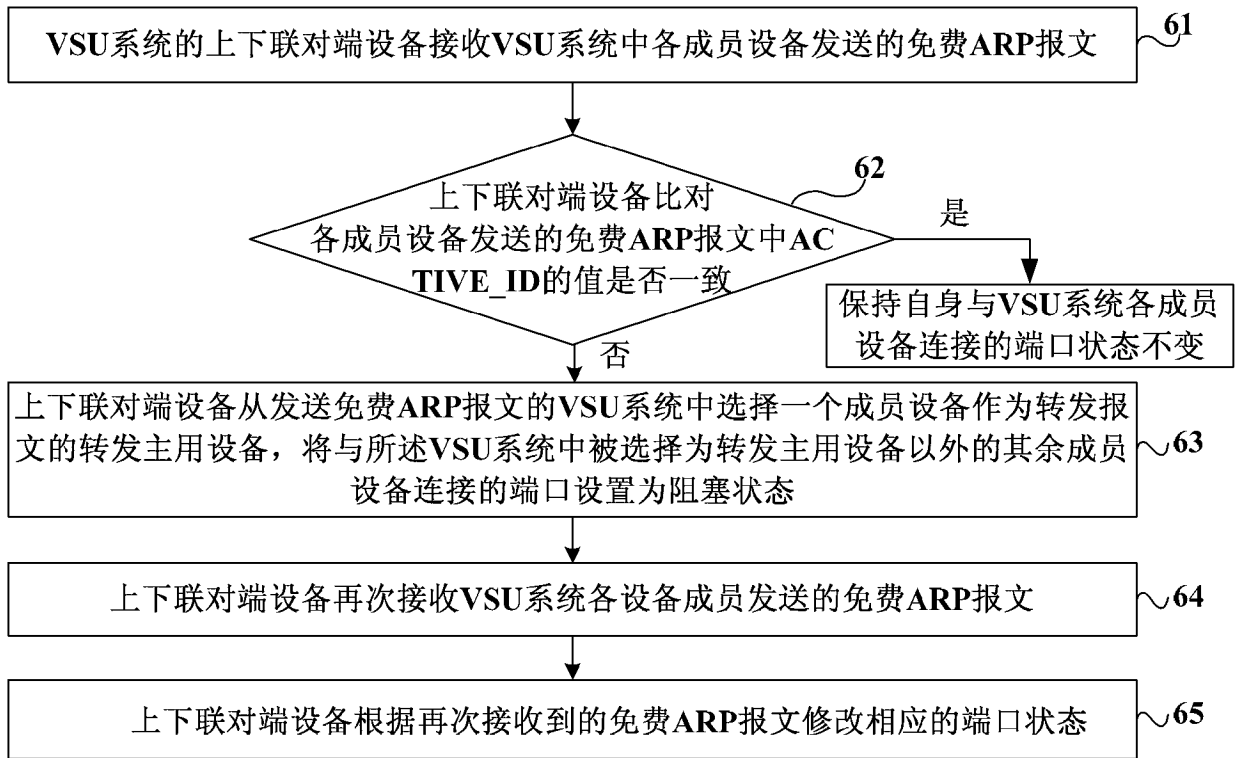


图 6

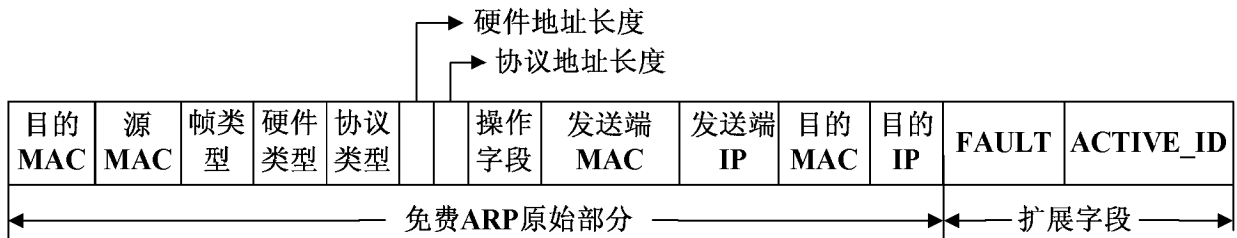


图 7

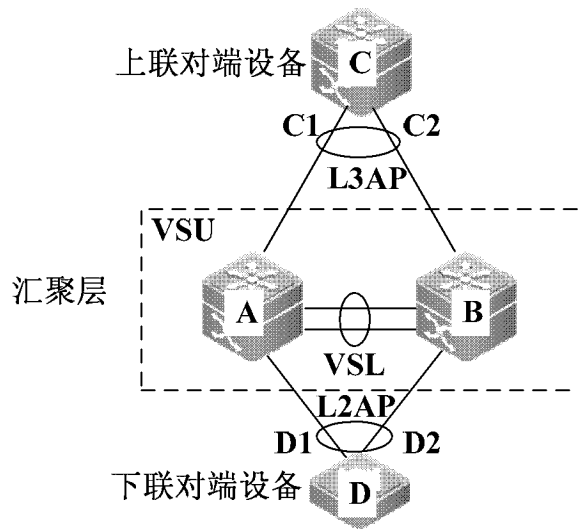


图 8

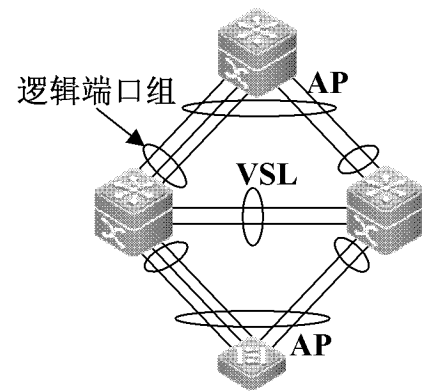


图 9

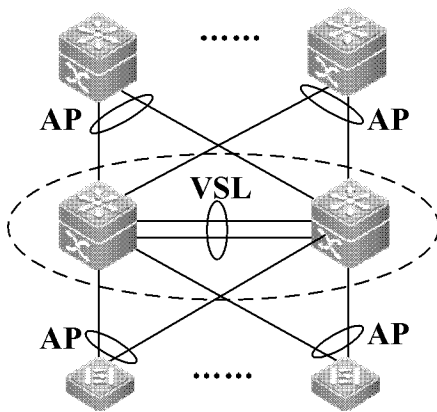


图 10a

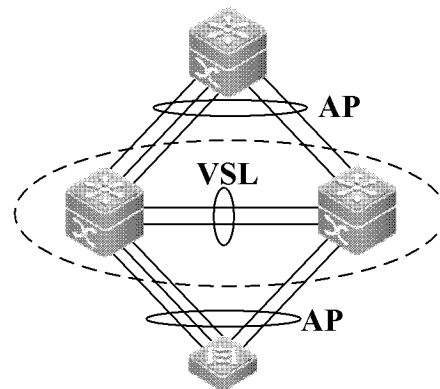


图 10b

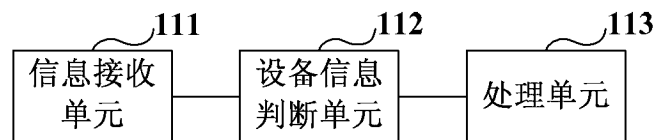


图 11