



(12)发明专利申请

(10)申请公布号 CN 109150709 A
(43)申请公布日 2019.01.04

(21)申请号 201710508989.0

(22)申请日 2017.06.28

(71)申请人 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 宋波 钱锋 赵晖 王冲

(51)Int.Cl.

H04L 12/709(2013.01)

H04L 12/823(2013.01)

H04L 12/26(2006.01)

权利要求书4页 说明书23页 附图5页

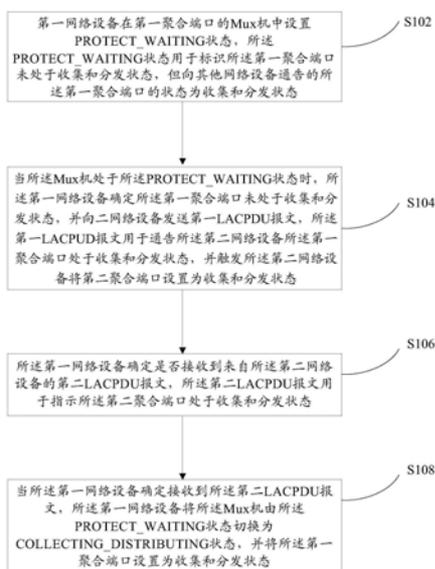
(54)发明名称

一种实现Mux机的方法、设备及系统

(57)摘要

一种实现Mux机的方法、设备及系统。所述方法包括,第一网络在第一聚合端口的Mux机中设置PROTECT_WAITING状态。当Mux机处于PROTECT_WAITING状态时,第一网络设备确定第一聚合端口未处于收集和分发状态,并向第二网络设备发送通告第一聚合端口处于收集和分发状态的第一LACPDU报文。当第一网络设备确定接收到来自第二网络设备的、用于指示第二聚合端口处于收集和分发状态的LACPDU报文,将Mux机由PROTECT_WAITING状态切换为COLLECTING_DISTRIBUTING状态,并将第一聚合端口设置为收集和分发状态。本申请实施方式实现Mux机中的聚合链路单端口UP的检测和处理,从而有助于减少业务流量在传输过程中产生的丢包。

CN 109150709 A



1. 一种实现复用Mux机的方法,其特征在于,所述方法应用于运行链路聚合控制协议LACP的网络中,所述网络包括第一网络设备和第二网络设备,所述第一网络设备的第一聚合端口经由聚合链路与所述第二网络设备的第二聚合端口连接,所述方法包括:

所述第一网络设备在所述第一聚合端口的Mux机中设置保护等待PROTECT_WAITING状态,所述PROTECT_WAITING状态用于标识所述第一聚合端口未处于收集和分发状态,但向其他网络设备通告的所述第一聚合端口的状态为收集和分发状态;

当所述Mux机处于所述PROTECT_WAITING状态时,所述第一网络设备确定所述第一聚合端口未处于收集和分发状态,并向所述第二网络设备发送第一链路聚合控制协议数据单元LACPDU报文,所述第一LACPDU报文用于通告所述第二网络设备所述第一聚合端口处于收集和分发状态,并触发所述第二网络设备将所述第二聚合端口设置为收集和分发状态;

所述第一网络设备确定是否接收到来自所述第二网络设备的第二LACPDU报文,所述第二LACPDU报文用于指示所述第二聚合端口处于收集和分发状态;

当所述第一网络设备确定接收到所述第二LACPDU报文,所述第一网络设备将所述Mux机由所述PROTECT_WAITING状态切换为收集分发COLLECTING_DISTRIBUTING状态,并将所述第一聚合端口设置为收集和分发状态。

2. 根据权利要求1所述的方法,其特征在于,在所述第一聚合端口的Mux机处于PROTECT_WAITING状态之前,所述Mux机处于附接ATTACHED状态,所述方法还包括:

所述第一网络设备接收来自所述第二网络设备的第三LACPDU报文;

当所述第一网络设备确定所述第一聚合端口处于选中状态,并且根据所述第三LACPDU报文确定所述第二聚合端口处于同步状态,以及根据所述第三LACPDU报文确定所述第二聚合端口未处于收集和分发状态时,所述第一网络设备将所述Mux机由所述ATTACHED状态切换为所述PROTECT_WAITING状态。

3. 根据权利要求2所述的方法,其特征在于,所述第一网络设备根据所述第三LACPDU报文确定所述第二聚合端口处于同步状态,具体包括:

所述第一网络设备确定所述第三LACPDU报文包括的第一信息与所述第一网络设备存储所述第一聚合端口的第二信息匹配,所述第一信息包括对端端口Partner_Port、对端端口优先级Partner_Port_Priority、对端系统Partner_System、对端系统优先级Partner_System_Priority、对端关键字Partner_Key和对端状态聚合Partner_State.Aggregation,所述第二信息包括本端端口号Actor_Port_Number、本端端口优先级Actor_Port_Priority、本端系统Actor_System、本端系统优先级Actor_System_Priority、本端操作端口关键字Actor_Oper_Port_Key和本端操作端口聚合Actor_Oper_Port_State.Aggregation;

所述第一网络设备确定所述第三LACPDU报文包括的本端状态同步Actor_State.Synchronization处于同步状态。

4. 根据权利要求2或3所述的方法,其特征在于,所述方法还包括:

当所述第一网络设备根据所述第三LACPDU报文确定所述第二聚合端口处于同步状态,并且根据所述第三LACPDU报文确定所述第二聚合端口处于收集和分发状态时,所述第一网络设备将所述Mux机由所述ATTACHED状态切换为所述COLLECTING_DISTRIBUTING状态,并将所述第一聚合端口设置为收集和分发状态。

5. 根据权利要求1-4中任一项所述的方法,其特征在于,所述方法还包括:

所述第一网络设备接收来自所述第二网络设备的第四LACPDU报文,所述第四LACPDU报文用于指示所述第二聚合端口未处于收集和分发状态;

当所述第一网络设备根据所述第四LACPDU报文确定所述第二聚合端口未处于收集和分发状态时,将所述Mux机由所述COLLECTING_DISTRIBUTING状态切换为所述PROTECT_WAITING状态。

6. 根据权利要求2-5中任一项所述的方法,其特征在于,所述方法还包括:

所述第一网络设备确定以下条件中至少一个条件成立时,将所述Mux机由所述PROTECT_WAITING状态切换为所述ATTACHED状态:

所述第一网络设备确定所述第一聚合端口处于不选中状态;

所述第一网络设备确定所述第一聚合端口处于备用状态;和,

所述第一网络设备接收来自所述第二网络设备的第五LACPDU报文,并根据所述第五LACPDU报文确定所述第二聚合端口未处于同步状态。

7. 根据权利要求1所述的方法,其特征在于,所述第一网络设备包括定时器,在所述第一聚合端口的Mux机处于PROTECT_WAITING状态中,所述方法还包括:

所述第一网络设备启动所述定时器;

直到所述定时器超时,所述第一网络设备确定未接收到来自所述第二网络设备的所述第二LACPDU报文,所述第一网络设备将所述Mux机由所述PROTECT_WAITING状态切换为ATTACHED状态。

8. 根据权利要求7所述的方法,其特征在于,所述方法还包括:

当所述Mux机处于PROTECT_WAITING状态,并且所述定时器超时时,所述第一网络设备将所述聚合链路设置为单端口UP,所述单端口UP用于指示所述聚合链路的一端聚合端口处于收集和分发状态,另一端聚合端口未处于收集和分发状态。

9. 根据权利要求7或8所述的方法,其特征在于,所述定时器的时长大于等于3秒,并且所述定时器的时长小于等于90秒。

10. 一种第一网络设备,其特征在于,所述第一网络设备运行有链路聚合控制协议LACP,所述第一网络设备的第一聚合端口经由聚合链路与所述第二网络设备的第二聚合端口连接,所述第一网络设备包括处理单元和发送单元,

所述处理单元,用于在所述第一聚合端口的Mux机中设置保护等待PROTECT_WAITING状态,所述PROTECT_WAITING状态用于标识所述第一聚合端口未处于收集和分发状态,但向其他网络设备通告的所述第一聚合端口的状态为收集和分发状态;

所述处理单元还用于当所述Mux机处于所述PROTECT_WAITING状态时,确定所述第一聚合端口未处于收集和分发状态;

所述发送单元,用于向所述第二网络设备发送第一链路聚合控制协议数据单元LACPDU报文,所述第一LACPDU报文用于通告所述第二网络设备所述第一聚合端口处于收集和分发状态,并触发所述第二网络设备将所述第二聚合端口设置为收集和分发状态;

所述处理单元还用于确定是否接收到来自所述第二网络设备的第二LACPDU报文,所述第二LACPDU报文用于指示所述第二聚合端口处于收集和分发状态;

当所述处理单元确定接收到所述第二LACPDU报文,所述处理单元还用于将所述Mux机

由所述PROTECT_WAITING状态切换为收集分发COLLECTING_DISTRIBUTING状态,并将所述第一聚合端口设置为收集和分发状态。

11. 根据权利要求10所述的第一网络设备,其特征在于,在所述第一聚合端口的Mux机处于PROTECT_WAITING状态之前,所述Mux机处于附接ATTACHED状态,所述第一网络设备还包括接收单元,

所述接收单元用于接收来自所述第二网络设备的第三LACPDU报文;

当所述处理单元确定所述第一聚合端口处于选中状态,并且根据所述第三LACPDU报文确定所述第二聚合端口处于同步状态,以及根据所述第三LACPDU报文确定所述第二聚合端口未处于收集和分发状态时,所述处理单元还用于将所述Mux机由所述ATTACHED状态切换为所述PROTECT_WAITING状态。

12. 根据权利要求11所述的第一网络设备,其特征在于,所述处理单元根据所述第三LACPDU报文确定所述第二聚合端口处于同步状态,具体包括:

所述处理单元还用于确定所述第三LACPDU报文包括的第一信息与所述第一网络设备存储所述第一聚合端口的第二信息匹配,所述第一信息包括对端端口Partner_Port、对端端口优先级Partner_Port_Priority、对端系统Partner_System、对端系统优先级Partner_System_Priority、对端关键字Partner_Key和对端状态聚合Partner_State.Aggregation,所述第二信息包括本端端口号Actor_Port_Number、本端端口优先级Actor_Port_Priority、本端系统Actor_System、本端系统优先级Actor_System_Priority、本端操作端口关键字Actor_Oper_Port_Key和本端操作端口聚合Actor_Oper_Port_State.Aggregation;

所述处理单元还用于确定所述第三LACPDU报文包括的本端状态同步Actor_State.Synchronization处于同步状态。

13. 根据权利要求11或12所述的第一网络设备,其特征在于,

当所述处理单元根据所述第三LACPDU报文确定所述第二聚合端口处于同步状态,并且根据所述第三LACPDU报文确定所述第二聚合端口处于收集和分发状态时,所述处理单元还用于将所述Mux机由所述ATTACHED状态切换为所述COLLECTING_DISTRIBUTING状态,并将所述第一聚合端口设置为收集和分发状态。

14. 根据权利要求10-13中任一项所述的第一网络设备,其特征在于,

所述接收单元还用于接收来自所述第二网络设备的第四LACPDU报文,所述第四LACPDU报文用于指示所述第二聚合端口未处于收集和分发状态;

当所述处理单元根据所述第四LACPDU报文确定所述第二聚合端口未处于收集和分发状态时,所述处理单元还用于将所述Mux机由所述COLLECTING_DISTRIBUTING状态切换为所述PROTECT_WAITING状态。

15. 根据权利要求11-14中任一项所述的第一网络设备,其特征在于,

所述处理单元确定以下条件中至少一个条件成立时,所述处理单元还用于将所述Mux机由所述PROTECT_WAITING状态切换为所述ATTACHED状态:

所述处理单元还用于确定所述第一聚合端口处于不选中状态;

所述处理单元还用于确定所述第一聚合端口处于备用状态;和,

所述处理单元还用于根据接收到的来自所述第二网络设备的第五LACPDU报文确定所

述第二聚合端口未处于同步状态。

16. 根据权利要求10所述的第一网络设备,其特征在于,在所述第一聚合端口的Mux机处于PROTECT_WAITING状态中,

所述处理单元还用于启动定时器;

直到所述定时器超时,所述处理单元确定未接收到来自所述第二网络设备的所述第二LACPDU报文,所述处理单元还用于将所述Mux机由所述PROTECT_WAITING状态切换为ATTACHED状态。

17. 根据权利要求16所述的第一网络设备,其特征在于,

当所述Mux机处于PROTECT_WAITING状态,并且所述定时器超时时,所述处理单元还用于将所述聚合链路设置为单端口UP,所述单端口UP用于指示所述聚合链路的一端聚合端口处于收集和分发状态,另一端聚合端口未处于收集和分发状态。

18. 根据权利要求16或17所述的第一网络设备,其特征在于,所述定时器的时长大于等于3秒,并且所述定时器的时长小于等于90秒。

19. 一种实现复用Mux机的网络系统,所述网络系统包括第一网络设备和第二网络设备,所述第一网络设备和/或所述第二网络设备为权利要求10至18中任一所述的第一网络设备。

一种实现Mux机的方法、设备及系统

技术领域

[0001] 本申请涉及通信技术领域,尤其涉及一种实现复用(英文:Mux)机的方法、设备及系统。更具体的,本申请涉及在运行链路聚合控制协议(英文:Link Aggregation Control Protocol,缩写:LACP)的网络中实现Mux机的技术。

背景技术

[0002] LACP为实现数据交换的网络设备提供一种标准的协商方式,采用LACP聚合的两端网络设备可以根据自身配置自动形成聚合链路,并利用聚合链路收发数据。聚合链路形成后,聚合链路两端的网络设备负责维护链路状态,在聚合条件发生变化时,自动调整或解散链路聚合。

[0003] LACP状态机基于端口维护各个聚合端口及其对端聚合端口的信息,计算每个聚合端口的聚合状态,并与其它网络设备交换LACP报文,调用聚合端口进行聚合组以及聚合成员管理。LACP状态机包括接收状态机(英文:Receive machine,缩写:RXmachine)、周期发送机(英文:Periodic Transmission machine,缩写:PTX machine)、选择逻辑(英文:Selection Logic)、Mux机(英文:Mux machine)和发送机(英文:Transmission machine,缩写:TX machine)。其中,Mux机用于将聚合端口附接到选定的聚合器(英文:aggregator),将聚合端口从不选定的聚合器分离,和基于当前的协议信息打开或关闭聚合端口的收集和分发状态。

[0004] 在实际的应用场景中,网络设备运行与所述网络设备的聚合端口相关联的Mux机。在满足触发条件的情况下,所述网络设备将所述网络设备的所述聚合端口的Mux机由附接(英文:ATTACHED)状态切换为收集分发(英文:COLLECTING_DISTRIBUTING)状态。但是,所述网络设备无法确保所述网络设备的对端网络设备的聚合端口相关联的Mux机处于COLLECTING_DISTRIBUTING状态。因此,在实际的应用场景中,所述网络设备和所述对端网络设备之间的聚合链路可能出现一端聚合端口处于收集和分发状态,另一端聚合端口未处于收集和分发状态的情况,从而导致业务流量出现丢包。

发明内容

[0005] 有鉴于此,本申请实施例提供了一种实现Mux机的方法、设备及系统,应用在运行LACP的网络中,实现Mux机中的聚合链路的两端聚合端口状态的检测和处理,有助于减少业务流量在传输过程中产生的丢包。

[0006] 本申请实施例提供的技术方案如下。

[0007] 第一方面,提供了一种实现Mux机的方法。所述方法应用于运行LACP的网络中,所述网络包括第一网络设备和第二网络设备,所述第一网络设备的第一聚合端口经由聚合链路与所述第二网络设备的第二聚合端口连接。所述第一网络设备在所述第一聚合端口的Mux机中设置PROTECT_WAITING状态,所述PROTECT_WAITING状态用于标识所述第一聚合端口未处于收集和分发状态,但向其他网络设备通告的所述第一聚合端口的状态为收集和分

发状态。当所述Mux机处于所述PROTECT_WAITING状态时,所述第一网络设备确定所述第一聚合端口未处于收集和分发状态,并向所述第二网络设备发送第一LACPDU报文,所述第一LACPDU报文用于通告所述第二网络设备所述第一聚合端口处于收集和分发状态,并触发所述第二网络设备将所述第二聚合端口设置为收集和分发状态。然后,所述第一网络设备确定是否接收到来自所述第二网络设备的第二LACPDU报文,所述第二LACPDU报文用于指示所述第二聚合端口处于收集和分发状态。并且,当所述第一网络设备确定接收到所述第二LACPDU报文,所述第一网络设备将所述Mux机由所述PROTECT_WAITING状态切换为COLLECTING_DISTRIBUTING状态,并将所述第一聚合端口设置为收集和分发状态。

[0008] 基于实施例提供的方案,所述第一网络在所述第一聚合端口的Mux机中设置PROTECT_WAITING状态。当所述Mux机处于所述PROTECT_WAITING状态时,所述第一网络设备确定所述第一聚合端口处于未处于收集和分发状态。然后,所述第一网络设备向所述第二网络设备发送通告所述第一聚合端口处于收集和分发状态的第一LACPDU报文。并且,当所述第一网络设备确定接收到来自所述第二网络设备的、用于指示所述第二聚合端口处于收集和分发状态的第二LACPDU报文,所述第一网络设备将所述Mux机由所述PROTECT_WAITING状态切换为COLLECTING_DISTRIBUTING状态,并将所述第一聚合端口设置为收集和分发状态。本申请实施方式实现Mux机中的聚合链路单端口UP的检测和处理,从而有助于减少业务流量在传输过程中产生的丢包。

[0009] 在第一方面的一种可能的实现方式中,在所述第一聚合端口的Mux机处于PROTECT_WAITING状态之前,所述Mux机处于ATTACHED状态,所述方法还包括,所述第一网络设备接收来自所述第二网络设备的第三LACPDU报文。当所述第一网络设备确定所述第一聚合端口处于选中状态,并且根据所述第三LACPDU报文确定所述第二聚合端口处于同步状态,以及根据所述第三LACPDU报文确定所述第二聚合端口未处于收集和分发状态时,所述第一网络设备将所述Mux机由所述ATTACHED状态切换为所述PROTECT_WAITING状态。

[0010] 基于上述实现方式,所述第一网络设备不仅通过所述第二LACPDU报文检测所述第二聚合端口的同步状态,还通过所述第二LACPDU报文检测所述第二聚合端口的收集和分发状态,从而确保在Mux机进入PROTECT_WAITING状态后,启动对所述第二聚合端口从未处于收集和分发状态到处于收集和分发状态的检测。

[0011] 在第一方面的又一种可能的实现方式中,所述第一网络设备根据所述第三LACPDU报文确定所述第二聚合端口处于同步状态,具体包括,所述第一网络设备确定所述第三LACPDU报文包括的第一信息与所述第一网络设备存储所述第一聚合端口的第二信息匹配,所述第一信息包括对端端口(英文:Partner_Port)、对端端口优先级(英文:Partner_Port_Priority)、对端系统(英文:Partner_System)、对端系统优先级(英文:Partner_System_Priority)、对端关键字(英文:Partner_Key)和对端状态聚合(英文:Partner_State.Aggregation),所述第二信息包括本端端口号(英文:Actor_Port_Number)、本端端口优先级(英文:Actor_Port_Priority)、本端系统(英文:Actor_System)、本端系统优先级(英文:Actor_System_Priority)、本端操作端口关键字(英文:Actor_Oper_Port_Key)和本端操作端口聚合(英文:Actor_Oper_Port_State.Aggregation)。所述第一网络设备确定所述第三LACPDU报文包括的本端状态同步Actor_State.Synchronization处于同步状态。

[0012] 在第一方面的再一种可能的实现方式中,所述方法还包括,当所述第一网络设备

根据所述第三LACPDU报文确定所述第二聚合端口处于同步状态,并且根据所述第三LACPDU报文确定所述第二聚合端口处于收集和分发状态时,所述第一网络设备将所述Mux机由所述ATTACHED状态切换为所述COLLECTING_DISTRIBUTING状态,并将所述第一聚合端口设置为收集和分发状态。

[0013] 基于上述实现方式,所述第一网络设备在确定所述第二聚合端口处于收集和分发状态时,快速的进入双端口收集和分发的状态,有助于减少业务流量在传输过程中产生的丢包。

[0014] 在第一方面的再一种可能的实现方式中,所述方法还包括,所述第一网络设备接收来自所述第二网络设备的第四LACPDU报文,所述第四LACPDU报文用于指示所述第二聚合端口未处于收集和分发状态。当所述第一网络设备根据所述第四LACPDU报文确定所述第二聚合端口未处于收集和分发状态时,将所述Mux机由所述COLLECTING_DISTRIBUTING状态切换为所述PROTECT_WAITING状态。

[0015] 基于上述实现方式,当所述第一网络设备收到指示所述第二聚合端口状态为未处于收集和分发状态的第三LACPDU报文后,及时将Mux机切换为PROTECT_WAITING状态,有助于减少业务流量在传输过程中产生的丢包。

[0016] 在第一方面的再一种可能的实现方式中,所述方法还包括,所述第一网络设备确定以下条件中至少一个条件成立时,将所述Mux机由所述PROTECT_WAITING状态切换为所述ATTACHED状态:所述第一网络设备确定所述第一聚合端口处于不选中状态;所述第一网络设备确定所述第一聚合端口处于备用状态;和,所述第一网络设备接收来自所述第二网络设备的第五LACPDU报文,并根据所述第五LACPDU报文确定所述第二聚合端口未处于同步状态。

[0017] 在第一方面的再一种可能的实现方式中,所述第一网络设备包括定时器,在所述第一聚合端口的Mux机处于PROTECT_WAITING状态中,所述方法还包括,所述第一网络设备启动所述定时器;直到所述定时器超时,所述第一网络设备确定未接收到来自所述第二网络设备的所述第二LACPDU报文,所述第一网络设备将所述Mux机由所述PROTECT_WAITING状态切换为ATTACHED状态。

[0018] 在第一方面的再一种可能的实现方式中,所述方法还包括,当所述Mux机处于PROTECT_WAITING状态,并且所述定时器超时时,所述第一网络设备将所述聚合链路设置为单端口UP,所述单端口UP用于指示所述聚合链路的一端聚合端口处于收集和分发状态,另一端聚合端口未处于收集和分发状态。

[0019] 基于上述实现方式,可以避免所述Mux的ATTACHED状态和PROTECT_WAITING状态之间的反复震荡。

[0020] 在第一方面中,可选的,在分离(英文:DETACHED)状态中,所述第一网络设备将所述单端口UP标志位的值设置为FALSE。

[0021] 在第一方面中,可选的,所述定时器的时长大于等于3秒,并且所述定时器的时长小于等于90秒。

[0022] 第二方面,提供了第一网络设备,所述第一网络设备具有实现上述方法中第一网络设备行为的功能。所述功能可以基于硬件实现,也可以基于硬件执行相应的软件实现。所述硬件或软件包括一个或多个与上述功能相对应的模块。

[0023] 在一个可能的设计中,第一网络设备的结构中包括处理器和接口,所述处理器被配置为支持第一网络设备执行上述方法中相应的功能。所述接口用于支持第一网络设备与第二网络设备之间的通信,向第二网络设备发送上述方法中所涉及的信息或者指令,或者从第二网络设备接收上述方法中所涉及的信息或者指令。所述第一网络设备还可以包括存储器,所述存储器用于与处理器耦合,其保存第一网络设备必要的程序指令和数据。

[0024] 在另一个可能的设计中,所述第一网络设备包括:处理器、发送器、随机存取存储器、只读存储器以及总线。其中,处理器通过总线分别耦接发送器、随机存取存储器以及只读存储器。其中,当需要运行第一网络设备时,通过固化在只读存储器中的基本输入/输出系统或者嵌入式系统中的bootloader引导系统进行启动,引导第一网络设备进入正常运行状态。在第一网络设备进入正常运行状态后,在随机存取存储器中运行应用程序和操作系统,使得该处理器执行第一方面或第一方面的任意可能的实现方式中的方法。

[0025] 第三方面,提供一种第一网络设备,所述第一网络设备包括:主控板和接口板,进一步,还可以包括交换网板。所述第一网络设备用于执行第一方面或第一方面的任意可能的实现方式中的方法。具体地,所述第一网络设备包括用于执行第一方面或第一方面的任意可能的实现方式中的方法的模块。

[0026] 第四方面,提供一种第一网络系统,所述第一网络系统包括控制器和第一转发设备。所述第一转发设备包括:接口板,进一步,还可以包括交换网板。所述第一转发设备用于执行第三方面中的接口板的功能,进一步,还可以执行第三方面中交换网板的功能。所述控制器包括接收器、处理器、发送器、随机存取存储器、只读存储器以及总线。其中,处理器通过总线分别耦接接收器、发送器、随机存取存储器以及只读存储器。其中,当需要运行控制器时,通过固化在只读存储器中的基本输入/输出系统或者嵌入式系统中的bootloader引导系统进行启动,引导控制器进入正常运行状态。在控制器进入正常运行状态后,在随机存取存储器中运行应用程序和操作系统,使得该处理器执行第三方面中主控板的功能。

[0027] 第五方面,提供一种实现Mux机的网络系统,所述网络系统包括第一网络设备和第二网络设备,所述第一网络设备和/或第二网络设备为前述第三方面或第四方面或第五方面中的第一网络设备。

[0028] 第六方面,提供了一种计算机存储介质,用于储存为上述第一网络设备所用的程序、代码或指令,当处理器或硬件设备执行这些程序、代码或指令时可以完成上述方面中第一网络设备的功能或步骤。

[0029] 通过上述方案,本申请实施例提供的用于实现Mux机的方法、设备及系统。所述第一网络在所述第一聚合端口的Mux机中设置PROTECT_WAITING状态。当所述Mux机处于所述PROTECT_WAITING状态时,所述第一网络设备确定所述第一聚合端口处于未处于收集和分发状态。然后,所述第一网络设备向所述第二网络设备发送通告所述第一聚合端口处于收集和分发状态的第一LACPDU报文。并且,当所述第一网络设备确定接收到来自所述第二网络设备的、用于指示所述第二聚合端口处于收集和分发状态的第二LACPDU报文,所述第一网络设备将所述Mux机由所述PROTECT_WAITING状态切换为COLLECTING_DISTRIBUTING状态,并将所述第一聚合端口设置为收集和分发状态。本申请实施方式实现Mux机中的聚合链路单端口UP的检测和处理,从而有助于减少业务流量在传输过程中产生的丢包。

附图说明

- [0030] 图1为本申请实施例的一种运行LACP的网络结构示意图；
- [0031] 图2为本申请实施例的一种Mux机的状态图；
- [0032] 图3为本申请实施例的一种实现Mux机的方法流程图；
- [0033] 图4为本申请实施例的另一种Mux机的状态图；
- [0034] 图5为本申请实施例的第一网络设备的结构示意图；
- [0035] 图6为本申请实施例的第一网络设备的硬件结构示意图；
- [0036] 图7为本申请实施例的另第一网络设备的硬件结构示意图；
- [0037] 图8为本申请实施例的第一网络系统的硬件结构示意图。

具体实施方式

[0038] 下面通过具体实施例,分别进行详细的说明。

[0039] 本申请所涉及的LACP可以参见电气和电子工程师协会(英文:Institute of Electrical and Electronics Engineers,缩写:IEEE)802.1AX和IEEE 802.3AD的说明,IEEE 802.1AX和IEEE 802.3A以全文引用的方式并入本申请中。在本申请实施例中,TRUE表示“真”或“条件成立”,在计算机编码中,TRUE的取值可以是“1”;FALSE表示“假”或“条件不成立”,在计算机编码中,FALSE的取值可以是“0”。在本申请实施例中,UP表示端口打开或可用,即:允许端口发送和接收数据报文;DOWN表示端口关闭或不可用,即:不允许端口发送和接收数据报文。在本申请实施例中,“确定”可以包括通过“设置”手段实现的方式,或者通过“保持”手段实现的方式。

[0040] 图1为本申请实施例的一种运行LACP的网络结构示意。如图1所示,所述网络包括第一网络设备和第二网络设备。所述第一网络设备经由链路聚合组(英文:Link Aggregation Group,缩写:LAG)与所述第二网络设备通信。所述LAG包括至少一条聚合链路(英文:aggregation link)。举例说明,所述LAG包括聚合链路1,所述第一网络设备经由所述聚合链路1与所述第二网络设备通信。所述聚合链路1的一端为所述第一网络设备上的聚合端口(英文:aggregation port)11,所述聚合链路1的另一端为所述第二网络设备上的聚合端口21。所述第一网络设备上的聚合器(英文:aggregator)1和所述第二网络设备上的聚合器2与所述LAG关联。所述聚合器1和聚合器2分别用于管理所述LAG包括的聚合链路的聚合端口,例如aggregator1管理聚合端口11,aggregator2管理聚合端口21。

[0041] 图1中示出了2台网络设备,应当理解,所述网络中包括的网络设备的数量不受限制,可以为2台或者3台或者更多数量。例如,所述网络还包括第三网络设备,所述第一网络设备经由LAG1与所述第二网络设备通信,所述第一网络设备经由LAG2与所述第三设备通信。经由LAG连接的两个网络设备之间可以包括其他网络设备。例如,所述第一网络设备与所述第二网络设备之间的LAG包括一台交换机。所述第一网络设备和所述第二网络设备可以是物理设备,所述物理设备包括路由器或交换机。所述第一网络设备和所述第二网络设备均可以连接其他网络设备。例如,所述第一网络设备连接客户机,所述第二网络设备连接服务器。所述客户端经由图1所示的网络与所述服务器通信。

[0042] 本申请实施例中的LAG、聚合链路、聚合端口和聚合器的定义可以参见

IEEE802.1AX中的说明,此处不进行赘述。其中,网络设备可以包括一个或多个聚合器。所述网络设备包括的聚合器的数量可以多于、少于或等于所述网络设备包括的聚合端口的数量。所述网络设备上的同一个聚合端口可以同时归属不同的聚合器。

[0043] 在图1中,所述第一网络设备和所述第二网络设备分别运行LACP。具体的,所述第一网络设备和所述第二网络设备的每个聚合端口关联一个LACP状态机,例如,聚合端口11关联LACP状态机11,聚合端口21关联LACP状态机21。由于聚合端口与LACP状态机是一一对应的,因此LACP状态机可以理解为:所述第一网络设备和所述第二网络设备在每个聚合接口上实现的LACP状态机。所述LACP状态机包括接收状态机(英文:Receive machine,缩写:RX machine)、周期发送机(英文:Periodic Transmission machine,缩写:PTX machine)、选择逻辑(英文:Selection Logic)、Mux机(英文:Mux machine)和发送机(英文:Transmission machine,缩写:TX machine)。上述LACP状态机包括的5个子状态机的定义和实现原理可以参见IEEE802.1AX中的说明,此处不进行赘述。所述第一网络设备和所述第二网络设备的每个聚合接口通过调用上述状态机,实现链路的聚合和解聚合。

[0044] 以图1中的聚合链路1实现Mux机为例进行说明。所述聚合链路1的一端为所述第一网络设备上的聚合端口11,所述聚合链路1的另一端为所述第二网络设备上的聚合端口21。在所述聚合端口11和所述聚合端口21上分别实现Mux机。具体的,所述第一网络设备在所述聚合端口11上实现Mux机,所述第二网络设备在所述聚合端口21上实现Mux机。图2示出了一种耦合控制(英文:coupled control)模式的Mux机的状态图。如图2所示,所述Mux机包括分离(英文:DETACHED)、等待(英文:WAITING)、附接(英文:ATTACHED)和收集分发(英文:COLLECTING_DISTRIBUTING)4个状态。在图2中,在所述聚合端口11上实现的Mux机中,所述第一网络设备被视为本端(英文:actor),所述第二网络设备被视为对端(英文:partner);在所述聚合端口21上实现的Mux机中,所述第二网络设备被视为actor,所述第一网络设备被视为partner。图2中,“+”表示“或”,“*”表示“和”。

[0045] 所述第一网络设备在所述聚合端口11上实现Mux机,初始化时默认进入DETACHED状态。当Mux机处于DETACHED状态时,所述第一网络设备进行以下处理:

[0046] 将本端操作端口同步状态(英文:Actor_Oper_Port_State.Synchronization,缩写:Actor.Sync)确定为FALSE,表示所述第一网络设备的聚合端口11未处于同步状态,或者可以认为是不同步状态。其中,所述Actor.Sync用于指示所述第一网络设备的聚合端口11的同步状态。

[0047] 将本端操作端口分发状态(英文:Actor_Oper_Port_State.Distributing,缩写:Actor.Distributing)确定为FALSE,表示所述聚合端口11未处于分发状态,或者可以认为是不分发状态。其中,所述Actor.Distributing用于指示所述聚合端口11的分发状态。

[0048] 将本端操作端口收集状态(英文:Actor_Oper_Port_State.Collecting,缩写:Actor.Collecting)确定为FALSE,表示所述聚合端口11未处于收集状态,或者可以认为是不收集状态。其中,所述Actor.Collecting用于指示所述聚合端口11的收集状态。

[0049] 将需要传送(英文:Need To Transmit,缩写:NTT)确定为TRUE,所述NTT用于指示是否允许所述第一网络设备与所述第二网络设备之间传送链路聚合控制协议数据单元(英文:Link Aggregation Control Protocol data unit,缩写:LACPDU)报文。

[0050] 在本申请的实施例中,可以使用两个字段分别指示本端操作端口的分发状态和收

集状态,例如上述中的Actor.Collecting和Actor.Distributing;也可以使用一个字段指示本端操作端口的分发状态和收集状态,例如Actor.Collecting_Distributing。同样的原则,可以使用两个字段分别指示对端操作端口的分发状态和收集状态,例如Partner.Collecting和Partner.Distributing;也可以使用一个字段指示本端操作端口的分发状态和收集状态,例如Partner.Collecting_Distributing。在本申请实施例中,使用Collecting和Distributing字段分别指示的方式和使用Collecting_Distributing指示的方式均可以表述为“收集和分发状态”。即,使用Collecting和Distributing字段分别指示的方式和使用Collecting_Distributing指示的方式均是“收集和分发状态”的实现方式。

[0051] 当所述第一网络设备确定所述聚合端口11处于选中(英文:SELECTED)状态(如图2中所示,Selected=SELECTED)或备用(英文:STANDBY)状态(如图2中所示,Selected=STANDBY),所述第一网络设备将Mux机由DETACHED状态切换为WAITING状态。具体的,所述第一网络设备可以调用Selection Logic选择与所述聚合端口11相关联的聚合器,根据图1,该聚合器为聚合器1。所述SELECTED状态用于指示合适的聚合器被选中。所述STANDBY状态用于指示合适的聚合器被选中,并且,与聚合器相关联的端口不使能聚合功能。当Mux机处于WAITING状态时,所述第一网络设备进行以下处理:启动等待时间计时器(英文:wait_while_timer)。在WAITING状态中,如果所述第一网络设备接收到触发事件,所述第一网络设备将所述聚合端口11确定为不选中(英文:UNSELECTED)状态(如图2中所示,Selected=UNSELECTED),所述第一网络设备将Mux机由WAITING状态切换为DETACHED状态。具体的,所述UNSELECTED状态可以由Selection Logic进行触发,也可以由RX machine进行触发。

[0052] 当所述聚合端口11被确定为SELECTED状态,并且Ready状态为TRUE时,所述第一网络设备将Mux机由WAITING状态切换为ATTACHED状态。其中,如果在WAITING状态中,所述聚合端口11处于SELECTED状态,则所述wait_while_timer超时时,所述第一网络设备将Ready状态确定为TRUE。其中,Ready状态被确定TRUE表示所述wait_while_timer超时,并且等待聚合端口被附接到聚合器。如果在WAITING状态中,所述聚合端口11的处于STANDBY状态,则即使所述wait_while_timer超时,Ready状态不会被确定为TRUE,WAITING状态也不会改变。在ATTACHED状态下,所述第一网络设备将所述聚合端口11附接到所述聚合器1,并且在完成附接过程后进行以下处理:将所述Actor.Sync确定为TRUE,表示所述第一网络设备的聚合端口11处于同步状态。将所述Actor.Collecting确定为FALSE,表示所述聚合端口11未处于收集状态。将所述Actor.Distributing确定为FALSE,表示所述聚合端口11未处于分发状态。将所述NTT确定为TRUE。

[0053] 如果所述第一网络设备接收到触发事件,导致所述第一网络设备将所述聚合端口11的确定为UNSELECTED状态或STANDBY状态,所述第一网络设备将Mux机由ATTACHED状态切换为DETACHED状态。Mux机进入DETACHED状态后,所述第一网络设备将所述聚合端口11从所述聚合器1中分离,并且,将聚合端口11的相应的状态信息确定为图2中DETACHED状态中所指示的状态。

[0054] 当所述聚合端口11的处于SELECTED状态,并且,对端操作端口同步状态(英文:Partner_Oper_Port_State.Synchronization,缩写:Partner.Sync)被确定为TRUE,所述第一网络设备将Mux机由ATTACHED状态切换为COLLECTING_DISTRIBUTING状态。其中,所述Partner.Sync被确定为TRUE表示所述第一网络设备确定所述第二网络设备的聚合端口21

处于同步状态,所述Partner.Sync用于指示所述第一网络设备确定所述聚合端口21的同步状态。具体的,所述Partner.Sync被确定为TRUE可以被解释为:所述第一网络设备获知所述聚合端口21处于同步状态,并且,所述第一网络设备认可所述聚合端口21的同步状态。其中,所述第一网络设备根据从所述第二网络设备接收到的LACPDU报文,确定所述Partner.Sync的状态。当Mux机处于COLLECTING_DISTRIBUTING状态时,所述第一网络设备进行以下处理:将所述Actor.Collecting确定为TRUE,表示所述聚合端口11处于收集状态。将所述Actor.Distributing确定为TRUE,表示所述聚合端口11处于分发状态。将所述NTT确定为TRUE。如此这样,Mux机进入COLLECTING_DISTRIBUTING状态后,所述第一网络设备根据所述聚合端口11处于收集和分发状态,将所述聚合端口11设置为UP,从而,所述第一网络设备允许所述聚合端口11接收和发送数据报文。

[0055] 如果所述第一网络设备接收到触发事件,导致所述聚合端口11的被确定为UNSELECTED状态或STANDBY状态,或者导致所述Partner.Sync被确定为FALSE,所述第一网络设备将Mux机由COLLECTING_DISTRIBUTING状态切换为ATTACHED状态。其中,Partner.Sync被确定为FALSE表示所述第一网络设备确定所述第二网络设备的聚合端口21未处于同步状态。

[0056] 图2示出的实现Mux机的方法流程图中的各个状态的解释也可以参见IEEE802.1AX中的说明。图2所示的Mux机可以应用在所述第一网络设备和第二网络设备的每个聚合端口上。所述第一网络设备在所述聚合端口11上实现图2所示的Mux机,并且所述Mux机进入COLLECTING_DISTRIBUTING状态后,所述第一网络设备无法确保所述第二网络设备的聚合端口21上实现的Mux机也处于COLLECTING_DISTRIBUTING状态。例如,如果所述第二网络设备存在异常,导致所述聚合端口21上的Mux机不能进入COLLECTING_DISTRIBUTING状态。这将导致聚合链路1的聚合端口11为UP状态,可以接收和发送数据报文;聚合链路1的聚合端口21为DOWN状态,不能接收和发送数据报文。从而,所述聚合链路1出现一端聚合端口处于收集和分发状态,另一端聚合端口未处于收集和分发状态的情况,或者可以认为是,聚合链路的一端聚合端口为UP,另一端聚合端口为DOWN。聚合链路上处于UP的聚合端口发出的数据报文,无法被聚合链路上处于DOWN的聚合端口接收,从而导致业务流量丢包。

[0057] 在本申请实施例中,聚合链路的一端聚合端口处于收集和分发状态,另一端聚合端口未处于收集和分发状态的情况可以被称为单端口UP(英文:Single-port_UP),也就是说,聚合链路的一端聚合端口为UP,另一端聚合端口为DOWN的情况可以被称为单端口UP。其中,所述第二网络设备存在异常,可能是由于不同厂商网络设备的实现差异或者网络设备故障或者链路故障等导致的。

[0058] 以图1为例说明,假设第一网络设备向第二网络设备发送的第一LACPDU报文是正确的,所述第二网络设备向所述第一网络设备发送的第二LACPDU报文存在错误。所述错误例如是所述第二LACPDU报文携带的本端系统(英文:Actor_System)或本端接口(英文:Actor_Port)等出现错误。其中,所述第二LACPDU报文中的Actor_System用于指示所述第二网络设备的系统标识,Actor_Port用于指示第二网络设备为聚合端口分配的端口号。所述第二LACPDU报文存在错误的原因可能包括:所述第一网络设备和所述第二网络设备之间链路上的其他网络设备因为故障造成所述第二LACPDU报文错误、所述第二网络设备在将所述第二LACPDU报文从控制平面下发到转发平面过程中因为故障造成所述第二LACPDU报文错

误,或所述第一网络设备在将所述第二LACPDU报文从转发平面上报到控制平面过程中因为故障造成所述第二LACPDU报文错误。这样,所述第一网络设备根据第二LACPDU中的actor状态信息,保存错误的第二网络设备的状态信息,所述第一网络设备根据所述第二LACPDU报文识别出的partner不是所述第二网络设备,而是错误的第二网络设备;所述第二网络设备根据第一LACPDU中的actor状态信息,保存正确的第一网络设备的状态信息,所述第二网络设备根据所述第一LACPDU报文识别出的partner是正确的所述第一网络设备。根据LACP的实现,所述第一网络设备再次接收到来自所述第二网络设备的LACPDU报文(该LACPDU报文可能是正确的LACPDU报文,也可能是不同于第二LACPDU报文的另一错误的LACPDU报文),根据来自所述第二网络设备的LACPDU报文中的partner状态信息,将Mux机的状态切换为COLLECTING_DISTRIBUTING状态。所述第二网络设备再次接收到来自所述第一网络设备的LACPDU报文,根据来自所述第一网络设备的LACPDU报文中的partner状态信息(该信息为之前存储的错误的所述第二网络设备的状态信息),确定Mux机的不进入COLLECTING_DISTRIBUTING状态。从而,所述第一网络设备的聚合端口进入稳态的UP状态,而聚合链路另一端的所述第二网络设备的聚合端口处于DOWN状态,聚合链路出现了单端口UP,导致业务流量丢包。

[0059] 为了解决上述问题,本申请实施例对图2所示的Mux机进行重新设计,并且,该Mux机可以应用在图1所示的网络结构中。从而,所述第一网络在所述第一聚合端口的Mux机中设置保护等待(英文:PROTECT_WAITING)状态。当所述Mux机处于所述PROTECT_WAITING状态时,所述第一网络设备确定所述第一聚合端口处于未处于收集和分发状态。然后,所述第一网络设备向所述第二网络设备发送第一LACPDU报文,所述第一LACPDU报文用于通告所述第二网络设备所述第一聚合端口处于收集和分发状态,并触发所述第二网络设备将所述第二聚合端口设置为收集和分发状态。并且,当所述第一网络设备确定接收到来自所述第二网络设备的、用于指示所述第二聚合端口处于收集和分发状态的第二LACPDU报文,所述第一网络设备将所述Mux机由所述PROTECT_WAITING状态切换为COLLECTING_DISTRIBUTING状态,并将所述第一聚合端口设置为收集和分发状态。

[0060] 本申请实施方式实现Mux机中的聚合链路单端口UP的检测和处理,有助于减少业务流量在传输过程中产生的丢包。具体Mux机的实现方式请参见本申请后续实施例的说明。

[0061] 图3为本申请实施例的一种实现Mux机的方法流程图。图3所示的方法可以应用于图1所示的运行LACP的网络中。所述网络包括第一网络设备和第二网络设备,所述第一网络设备的第一聚合端口经由聚合链路与所述第二网络设备的第二聚合端口连接,所述第一网络设备在所述第一聚合端口上实现所述Mux机。

[0062] 在本实施例中,以图1中的LAG中的聚合链路1为例进行说明,其中,所述第一网络设备的聚合端口11经由聚合链路1与所述第二网络设备的聚合端口21连接。应当理解,所述第一网络设备包括的所有聚合接口和所述第二网络设备包括的所有聚合接口均可以运行图3所示的方法。图4示出了申请实施例的一种Mux机的状态图。图3所示的方法可以用于实现图4所示的Mux机。图4中未示出DETACHED状态、WAITING状态和ATTACHED状态下的操作,以及,图4中未示出DETACHED状态、WAITING状态和ATTACHED状态之间的切换条件。相应的,图4中的DETACHED状态、WAITING状态和ATTACHED状态下的操作,以及上述3个状态之间的切换条件可以参见图2的相应部分,图4中不再进行赘述;并且,上述3个状态之间相互转换的解

释可以参见前述实施例的描述,此处不进行赘述。另外,关于actor和partner,以及“+”和“*”的解释,图4与前述图2的描述相同,此处不进行赘述。在本申请实施例中,所述第一聚合端口对应举例说明中的聚合端口11,所述第二聚合端口对应举例说明中的聚合端口21。本实施例结合图1和图4对图3所示的方法进行说明,所述方法包括:

[0063] S102、所述第一网络设备在所述第一聚合端口的Mux机中设置PROTECT_WAITING状态,所述PROTECT_WAITING状态用于标识所述第一聚合端口未处于收集和分发状态,但向其他网络设备通告的所述第一聚合端口的状态为收集和分发状态。

[0064] 举例说明,所述第一网络设备在所述第一网络设备的聚合端口11上实现Mux机。所述第一网络设备在所述Mux机中设置PROTECT_WAITING状态,如图4所示。相比较图2所示的Mux机,图4所示的Mux机在ATTACHED状态和COLLECTING_DISTRIBUTING状态之间增加了PROTECT_WAITING状态。当满足切换条件,所述第一网络设备可以实现所述聚合端口11的Mux机的ATTACHED状态和PROTECT_WAITING状态之间的切换。并且,当满足切换条件,所述第一网络设备还可以实现所述聚合端口11的Mux机的COLLECTING_DISTRIBUTING状态和PROTECT_WAITING状态之间的切换。

[0065] 所述PROTECT_WAITING状态用于标识所述聚合端口11未处于收集和分发状态,但向其他网络设备通告的所述聚合端口11的状态为收集和分发状态。也就是说,在PROTECT_WAITING状态中,所述第一网络设备会将所述聚合端口11确定为未处于收集和分发状态。所述第一网络设备将所述聚合端口11确定为未处于收集和分发状态包括:所述第一网络设备将所述聚合端口11保持为未处于收集和分发状态,或者,所述第一网络设备将所述聚合端口11设置为未处于收集和分发状态。具体的,在一种可能的实现方式中,所述第一网络设备检测到所述聚合端口11未处于收集和分发状态,则维持所述聚合端口11未处于收集和分发状态。例如,当所述聚合端口11的Mux机由ATTACHED状态切换为PROTECT_WAITING状态时,所述聚合端口11未处于收集和分发状态。Mux机进入PROTECT_WAITING状态后,所述第一网络设备保持所述聚合端口11的未处于收集和分发状态。在另一种可能的实现方式中,所述第一网络设备检测到所述聚合端口11处于收集和分发状态,则将所述聚合端口11设置为未处于收集和分发状态。例如,当所述聚合端口11的Mux机由COLLECTING_DISTRIBUTING状态切换为PROTECT_WAITING状态时,所述聚合端口11处于收集和分发状态。Mux机进入PROTECT_WAITING状态后,所述第一网络设备将所述聚合端口11设置为未处于收集和分发状态。在又一种可能的实现方式中,所述第一网络设备不检测所述聚合端口11的状态,而是直接将所述聚合端口11设置为未处于收集和分发状态。

[0066] 在PROTECT_WAITING状态中,所述第一网络设备将所述聚合端口11确定为未处于收集和分发状态后,所述第一网络设备向其他网络设备通告的所述聚合端口11的状态为收集和分发状态。所述其他网络设备包括与所述第一网络设备通过聚合链路通信的网络设备。所述第一网络设备向其他网络设备宣告:所述聚合端口11的状态为收集和分发状态。但实际上,所述聚合端口11的真实的状态为未处于收集和分发状态。

[0067] S104、当所述Mux机处于所述PROTECT_WAITING状态时,所述第一网络设备确定所述第一聚合端口未处于收集和分发状态,并向所述第二网络设备发送第一LACPDU报文,所述第一LACPDU报文用于通告所述第二网络设备所述第一聚合端口处于收集和分发状态,并触发所述第二网络设备将所述第二聚合端口设置为收集和分发状态。

[0068] 举例说明,当所述聚合端口11的Mux机处于所述PROTECT_WAITING状态时,所述第一网络设备进行以下处理:将Actor.Collecting确定为FALSE,表示所述聚合端口11未处于收集状态。将Actor.Distributing确定为FALSE,表示所述聚合端口11未处于分发状态。向其他网络设备通告Actor.Collecting为TRUE和Actor.Distributing为TRUE,表示所述聚合端口11处于收集和分发状态。将所述NTT确定为TRUE。

[0069] 其中,上述Actor.Collecting和Actor.Distributing也可以由Actor.Collecting_Distributing表示。例如,Actor.Collecting_Distributing被确定为TRUE,表示所述聚合端口11处于收集状态和分发状态;Actor.Collecting_Distributing被确定为FALSE,表示所述聚合端口11未处于收集状态和分发状态。在本申请实施例中,上述两种表示方式均被涵盖在“收集和分发状态”的范围内。所述第一网络设备将Actor.Collecting和Actor.Distributing确定为FALSE的实现方式,可以参见前述S102的说明,此处不进行赘述。

[0070] 所述第一网络设备还向所述第二网络设备发送第一LACPDU报文,所述第一LACPDU报文用于通告所述第二网络设备所述聚合端口11处于收集和分发状态。另外,所述第一LACPDU报文还用于触发所述第二网络设备将所述聚合端口21设置为收集和分发状态。具体的,所述第二网络设备接收到所述第一LACPDU报文后,根据所述第一LACPDU报文确定所述聚合端口11的状态为收集和分发状态,触发将所述聚合端口21设置为收集和分发状态的过程。

[0071] 其中,所述第一LACPDU报文是由所述第一网络设备向所述第二网络设备发送的LACPDU报文。所述第一LACPDU报文携带有本端状态(英文:Actor_State)字段,所述Actor_State字段长度为1字节。所述Actor_State字段包括Collecting标志位和Distributing标志位。所述Collecting标志位和Distributing标志位用于指示所述聚合端口11是否处于收集和分发状态。例如,当Collecting标志位和Distributing标志位均被置位为1时,表示聚合端口11处于收集和分发状态。其中,LACPDU报文和Actor_State的格式也可以参见IEEE 802.1AX中图5-8和图5-9的说明。另外,所述第二网络设备可以调用RX machine完成对所述第一LACPDU报文的接收和处理。

[0072] S106、所述第一网络设备确定是否接收到来自所述第二网络设备的第二LACPDU报文,所述第二LACPDU报文用于指示所述第二聚合端口处于收集和分发状态。

[0073] 举例说明,如果所述第二网络设备被所述第一LACPDU报文触发成功,所述第二网络设备将生成所述第二LACPDU报文,并且将所述第二LACPDU报文向所述第一网络设备发送。或者,如果所述第二网络设备存在故障或异常,可能导致所述聚合端口21未处于收集和分发状态,从而,所述第二网络设备在接收到所述第一LACPDU报文的情况下,无法向所述第一网络设备发送所述第二LACPDU报文。

[0074] 所述第一网络设备检测是否接收到来自所述第二网络设备的第二LACPDU报文,所述第二LACPDU报文用于指示所述聚合端口21处于收集和分发状态。从而,所述第一网络设备可以确定所述聚合链路1是否存在一端的聚合端口处于收集和分发状态,另一端的聚合端口未处于收集和分发状态(不收集和转发状态)的情况。也就是说,所述第一网络设备可以检测所述聚合链路1是否存在单端口UP的情况。具体的,所述第一网络设备检测聚合链路1是否为单端口UP的实现方式是:如果所述第二网络设备的聚合端口21处于收集和分发

状态,则所述聚合链路1不属于单端口UP。如果所述第二网络设备的聚合端口21始终未处于收集和分发状态,则所述聚合链路1为单端口UP。

[0075] 其中,所述第二LACPDU报文是由所述第二网络设备向所述第一网络设备发送的LACPDU报文。所述第二LACPDU报文携带有Actor_State字段,所述Actor_State字段长度为1字节。所述Actor_State字段包括Collecting标志位和Distributing标志位。所述Collecting标志位和Distributing标志位用于指示所述聚合端口21是否处于收集和分发状态。例如,当Collecting标志位和Distributing标志位均被置位为1时,表示聚合端口21处于收集和分发状态。其中,LACPDU报文和Actor_State的格式也可以参见IEEE 802.1AX中图5-8和图5-9的说明。另外,所述第一网络设备可以调用RX machine完成对所述第二LACPDU报文的接收和处理。

[0076] S108、当所述第一网络设备确定接收到所述第二LACPDU报文,所述第一网络设备将所述Mux机由所述PROTECT_WAITING状态切换为COLLECTING_DISTRIBUTING状态,并将所述第一聚合端口设置为收集和分发状态。

[0077] 根据LACP,actor或partner的聚合端口发生状态信息的变化,可以通过LACPDU报文通知对方。例如,所述第二网络设备的聚合端口21的Mux机由ATTACHED状态切换为COLLECTING_DISTRIBUTING状态,所述第二网络设备将所述聚合端口21的收集状态切换为收集和将所述聚合端口21的分发状态切换为分发。如此这样,所述聚合端口21发生了状态信息的变化。所述第二网络设备将向所述第一网络设备发送第二LACPDU报文,所述第二LACPDU报文用于指示所述聚合端口21处于收集和分发状态。在图4所示的Mux机中,所述第一网络设备确定是否接收到所述第二LACPDU报文。从而,第一网络设备可以根据所述第二LACPDU报文,确定所述聚合端口21是否处于收集和分发状态。

[0078] 如果所述第一网络设备接收到所述第二LACPDU报文,则说明所述第二设备将聚合端口21的收集状态确定为收集和将聚合端口21的分发状态确定为分发。所述第一网络设备根据指示所述第二设备的聚合端口21处于收集和分发状态的所述第二LACPDU报文,将所述Mux机由所述PROTECT_WAITING状态切换为COLLECTING_DISTRIBUTING状态。具体的,所述第一网络设备根据所述第二LACPDU报文携带的Actor_State字段中的Collecting标志位为TRUE,所述第一网络设备确定对端操作端口收集状态(英文:Partner_Oper_Port_State.Collecting,缩写:Partner.Collecting)的值为TRUE。所述第一网络设备根据所述第二LACPDU报文携带的所述Actor_State字段中的Distributing标志位为TRUE,所述第一网络设备确定对端操作端口分发状态(英文:Partner_Oper_Port_State.Distributing,缩写:Partner.Distributing)的值为TRUE。所述第一网络设备根据Partner.Collecting=TRUE和Partner.Distributing=TRUE,将所述Mux机由所述PROTECT_WAITING状态切换为COLLECTING_DISTRIBUTING状态。

[0079] 当所述聚合端口11的Mux机处于COLLECTING_DISTRIBUTING状态,所述第一网络设备将所述聚合端口11设置为收集和分发状态。从而,所述聚合端口11和聚合端口21均处于收集和分发状态。也就是说,所述聚合链路1的聚合端口11和聚合端口21均为UP。即,所述第一网络设备允许所述聚合端口11接收和发送数据报文,并且,所述第二网络设备允许所述聚合端口21接收和发送数据报文。

[0080] 图4所示的Mux机通过增加PROTECT_WAITING状态,确保了聚合链路两端的聚合端

口均UP的情况下,将聚合端口的Mux机切换到稳定的收集和分发状态。

[0081] 通过上述实施方式,在运行LACP的网络场景中,所述第一网络在所述第一聚合端口的Mux机中设置PROTECT_WAITING状态。当所述Mux机处于所述PROTECT_WAITING状态时,所述第一网络设备确定所述第一聚合端口处于未处于收集和分发状态。然后,所述第一网络设备向所述第二网络设备发送第一LACPDU报文,所述第一LACPDU报文用于通告所述第二网络设备所述第一聚合端口处于收集和分发状态,并触发所述第二网络设备将所述第二聚合端口设置为收集和分发状态。并且,当所述第一网络设备确定接收到来自所述第二网络设备的、用于指示所述第二聚合端口处于收集和分发状态的第二LACPDU报文,所述第一网络设备将所述Mux机由所述PROTECT_WAITING状态切换为COLLECTING_DISTRIBUTING状态,并将所述第一聚合端口设置为收集和分发状态。本申请实施方式实现Mux机中的聚合链路单端口UP的检测和处理,从而有助于减少业务流量在传输过程中产生的丢包。

[0082] 应当理解,本申请实施例的图3所示的方法是以所述第一网络设备的聚合端口实现Mux机为例说明,图3所示的方法同样可以应用在所述第二网络设备的聚合端口中。而且,对于同一条聚合链路而言,所述聚合链路的两个聚合端口均实现图4所示的Mux机。举例说明,图1所示的聚合链路1的聚合端口11在实现图4所示的Mux机期间,聚合链路1的聚合端口21也在实现图4所示的Mux机。并且,在所述聚合端口11的Mux机中,所述第一网络设备被视为actor,所述第二网络设备被视为partner;在所述聚合端口21的Mux机中,所述第二网络设备被视为actor,所述第一网络设备被视为partner。

[0083] 可选的,所述第一网络设备包括定时器,在所述第一聚合端口的Mux机处于PROTECT_WAITING状态中,所述方法还包括:所述第一网络设备启动定时器;直到所述定时器超时,所述第一网络设备确定未接收到来自所述第二网络设备的所述第二LACPDU报文,所述第一网络设备将所述Mux机由所述PROTECT_WAITING状态切换为ATTACHED状态。

[0084] 举例说明,所述第一网络设备中还设置有定时器,所述定时器与所述聚合端口11相关联。所述第一网络设备在计时器计期间检测是否接收到来自所述第二网络设备的第二LACPDU报文。所述定时器可以由硬件计时器或者软件计时器实现。通过设置所述定时器,可以防止所述聚合端口11的Mux机长时间处于所述PROTECT_WAITING状态。在一种可能的实现方式中,所述定时器的时长可以根据组网的实际场景和网络设备的性能要求进行设定,例如, $3s \leq T \leq 90s$,其中T表示定时器的时长,s表示秒。

[0085] 如果从所述定时器开始计时起,直到所述定时器超时,所述第一网络设备始终没有接收到所述第二LACPDU报文,所述第一网络设备确定所述聚合链路1为单端口UP。也就是说,直到所述定时器超时,所述第一网络设备始终没有接收到所述第二LACPDU报文,所述第一网络设备可以确定所述聚合端口21未处于收集和分发状态。基于上述情况,所述第一网络设备会将所述聚合端口11的所述Mux机由所述PROTECT_WAITING状态切换为ATTACHED状态(如图4所示)。其中,所述直到所述定时器超时,所述第一网络设备始终没有接收到所述第二LACPDU报文包括:从所述定时器开始计时起,直到所述定时器超时,所述第一网络设备没有接收到任何来自所述第二网络设备的LACPDU报文;或者,从所述定时器开始计时起,直到所述定时器超时,所述第一网络设备接收到来自所述第二网络设备的LACPDU报文,但该LACPDU报文并没有指示所述聚合端口21处于收集和分发状态。

[0086] 可选的,所述方法还包括:当所述Mux机处于COLLECTING_DISTRIBUTING状态,并且

所述定时器超时,所述第一网络设备将所述聚合链路设置为单端口UP,所述单端口UP用于指示所述聚合链路的一端聚合端口处于收集和分发状态,另一端聚合端口未处于收集和分发状态

[0087] 举例说明,所述第一网络设备还可以设置Single-port_UP。所述Single-port_UP用于指示聚合链路是否为单端口UP。其中,所述Single-port_UP被部署在Mux机中,根据所述定时器进行触发。具体的,在PROTECT_WAITING状态中,直到所述定时器超时,所述第一网络设备未接收到来自所述第二网络设备的所述第二LACPDU报文,所述第一网络设备将所述Single-port_UP标志位的值被设置为TRUE,即Single-port_UP=TRUE。其中,Single-port_UP标志位的值被设置为TRUE表示所述聚合链路1存在单端口UP的情况。因此,直到所述定时器超时,所述第一网络设备未接收到所述第二LACPDU报文,所述第一网络设备会将所述聚合端口11的Mux机由PROTECT_WAITING状态切换为ATTACHED状态,并且,所述第一网络设备将所述Single-port_UP标志位的值被设置为TRUE。

[0088] 所述第一网络设备的聚合端口11的所述Mux机进入ATTACHED状态后,由于所述Single-port_UP标志位的值为TRUE,所述第一网络设备不会将所述Mux由ATTACHED状态切换为PROTECT_WAITING状态。因此,进一步的有益效果是:所述Single-port_UP可以避免所述Mux的ATTACHED状态和PROTECT_WAITING状态之间的反复震荡。

[0089] 可选的,Mux机处于DETACHED状态时,所述第一网络设备将所述单端口UP标志位的值设置为FALSE。

[0090] 所述DETACHED状态可以被看作Mux机的初始化阶段,在DETACHED状态中,所述第一网络设备将Single-port_UP标志位的值设置为FALSE (Single-port_UP=FALSE),表示聚合链路1未处于单端口UP。在本申请实施例中,所述未处于单端口UP可能存在两种情况:所述聚合链路1的聚合端口11和聚合端口21均为DOWN,或者,所述聚合链路1的聚合端口11和聚合端口21均为UP(聚合端口11为UP包括对外通告为UP的情况)。在所述DETACHED状态中,将Single-port_UP标志位的值设置为FALSE,有助于避免旧的Single-port_UP标志位的值影响Mux机的运行。

[0091] 可选的,在S102或S104之前,所述第一聚合端口的Mux机处于所述ATTACHED状态,所述方法还包括S1001和S1002,以便完成所述第一网络设备将所述Mux机由ATTACHED状态切换为PROTECT_WAITING状态的过程,如图4所示。

[0092] S1001、所述第一网络设备接收来自所述第二网络设备的第三LACPDU报文;

[0093] S1002、当所述第一网络设备确定所述第一聚合端口处于选中状态,并且根据所述第三LACPDU报文确定所述第二聚合端口处于同步状态,以及根据所述第三LACPDU报文确定所述第二聚合端口未处于收集和分发状态时,所述第一网络设备将所述Mux机由所述ATTACHED状态切换为所述PROTECT_WAITING状态。

[0094] 举例说明,所述第一网络设备可以调用Selection Logic选择与所述聚合端口11相关联的聚合器,根据图1,该聚合器为聚合器1。所述SELECTED用于指示合适的聚合器被选中。当所述Selection Logic完成所述聚合端口11相关联的聚合器的选择后,所述第一网络设备可以确定所述第一网络设备的聚合端口11处于选中状态 (Selected=SELECTED)。

[0095] 所述第一网络设备可以调用RX machine,使用RX machine完成对所述第三LACPDU报文的接收和处理。所述第三LACPDU报文是由所述第二网络设备发送的。所述第三LACPDU

报文被用于确定所述聚合端口21是否处于同步状态,并且所述第三LACPDU报文还被用于确定所述聚合端口21是否处于收集和分发状态。当所述聚合端口11处于选中状态,并且所述第一网络设备根据所述第三LACPDU报文确定所述聚合端口21处于同步状态,以及根据所述第三LACPDU报文确定所述聚合端口21未处于收集和分发状态时,所述第一网络设备将所述聚合端口11的Mux机由ATTACHED状态切换为PROTECT_WAITING状态。

[0096] 其中,所述第三LACPDU报文中的用于指示所述聚合端口21的状态信息包括Actor_State,所述Actor_State包括Collecting标志位和Distributing标志位。所述第一网络设备确定所述第三LACPDU报文中所述Actor_State包括的Collecting标志位和Distributing标志位是否被设置为TRUE。当Collecting标志位和Distributing标志位被设置为FALSE时,所述第一网络设备确定所述聚合端口21未处于收集和分发状态。

[0097] 通过上述S1001和S1002的实现方式,所述第一网络设备的聚合端口11的Mux机处于ATTACHED状态中,当所述第一网络设备确定所述聚合端口21处于同步状态和所述聚合端口21未处于收集和分发状态的情况下,所述第一网络设备将聚合端口11的Mux机切换为PROTECT_WAITING状态。从而,确保在Mux机进入PROTECT_WAITING状态后,启动对所述第二聚合端口从未处于收集和分发状态到处于收集和分发状态的检测。

[0098] 在S1002中,进一步可选的,所述方法还包括S10021和S10022。S10021和S10022说明了所述第一网络设备根据所述第三LACPDU报文确定所述第二聚合端口处于同步状态的实现方式。

[0099] S10021、所述第一网络设备确定所述第三LACPDU报文包括的第一信息与所述第一网络设备存储所述第一聚合端口的第二信息匹配,所述第一信息包括Partner_Port、Partner_Port_Priority、Partner_System、Partner_System_Priority、Partner_Key和Partner_State.Aggregation,所述第二信息包括Actor_Port_Number、Actor_Port_Priority、Actor_System、Actor_System_Priority、Actor_Oper_Port_Key和Actor_Oper_Port_State.Aggregation。

[0100] S10022、所述第一网络设备确定所述第三LACPDU报文包括的本端状态同步Actor_State.Synchronization处于同步状态。

[0101] 所述第一网络设备根据所述第三LACPDU报文确定所述聚合端口21是否处于同步状态。具体的,所述第三LACPDU报文中的所述第一信息包括:对端端口(英文:Partner_Port)、对端端口优先级(英文:Partner_Port_Priority)、对端系统(英文:Partner_System)、对端系统优先级(英文:Partner_System_Priority)、对端关键字(英文:Partner_Key)和对端状态聚合(英文:Partner_State.Aggregation)。所述第一网络设备确定上述信息与所述第一网络设备存储的所述第二信息,即本端操作参数值(英文:corresponding operational parameter values for the Actor)是否匹配。所述第一网络设备存储的本端操作参数值包括:本端端口号(英文:Actor_Port_Number)、本端端口优先级(英文:Actor_Port_Priority)、本端系统(英文:Actor_System)、本端系统优先级(英文:Actor_System_Priority)、本端操作端口关键字(英文:Actor_Oper_Port_Key)和本端操作端口聚合(英文:Actor_Oper_Port_State.Aggregation)。而且,所述第三LACPDU报文中的用于指示所述聚合端口21的状态信息包括本端状态同步(英文:Actor_State.Synchronization),所述Actor_State.Synchronization是所述第三LACPDU报文中的Actor_State字段的

Synchronization标志位。所述第一网络设备确定所述第三LACPDU报文中Actor_State.Synchronization是否被设置为TRUE。其中,上述各个信息的定义可以参见IEEE 802.1AX中的说明,此处不进行赘述。

[0102] 其中,可选的,用于确定所述第二网络设备的聚合端口21的同步状态的信息和用于确定所述聚合端口21的收集和分发状态的信息可以被携带在不同的第三LACPDU报文中。例如,用于确定所述聚合端口21的同步状态的信息被携带在一个第三LACPDU报文中;用于确定所述聚合端口21的收集和分发状态的信息被携带在另一个第三LACPDU报文中。两个第三LACPDU报文被所述第二网络设备分别发送到所述第一网络设备。

[0103] 所述第一网络设备确定所述第三LACPDU报文中包括的Partner_Port、Partner_Port_Priority、Partner_System、Partner_System_Priority、Partner_Key和Partner_State.Aggregation与所述第一网络设备存储的Actor_Port_Number、Actor_Port_Priority、Actor_System、Actor_System_Priority、Actor_Oper_Port_Key和Actor_Oper_Port_State.Aggregation相匹配,并且,所述第一网络设备确定所述第三LACPDU报文中包括的Actor_State.Synchronization为TRUE,所述第一网络设备确定所述聚合端口21处于同步状态,即所述第一网络设备将Partner.Sync确定为TRUE。具体的,所述Partner.Sync被确定为TRUE可以被解释为:所述第一网络设备获知所述聚合端口21处于同步状态(根据Actor_State.Synchronization=TRUE),并且,所述第一网络设备认可所述聚合端口21处于同步状态(根据上述匹配的结果)。

[0104] 可选的,所述方法还包括S1003,基于前述S1001中的第三LACPDU报文,以便完成所述第一网络设备将所述第一聚合端口的Mux机由ATTACHED状态切换为COLLECTING_DISTIBUTING状态的过程,如图4所示。

[0105] S1003、当所述第一网络设备根据所述第三LACPDU报文确定所述第二聚合端口处于同步状态,并且根据所述第三LACPDU报文确定所述第二聚合端口处于收集和分发状态时,所述第一网络设备将所述Mux机由所述ATTACHED状态切换为所述COLLECTING_DISTIBUTING状态,并将所述第一聚合端口设置为收集和分发状态。

[0106] 举例说明,根据S1001,所述第一网络设备接收到所述第三LACPDU报文后,确定所述第三LACPDU报文中包括的Partner_Port、Partner_Port_Priority、Partner_System、Partner_System_Priority、Partner_Key和Partner_State.Aggregation与所述第一网络设备存储的Actor_Port_Number、Actor_Port_Priority、Actor_System、Actor_System_Priority、Actor_Oper_Port_Key和Actor_Oper_Port_State.Aggregation相匹配,并且,所述第一网络设备确定所述第三LACPDU报文中包括的Actor_State.Synchronization为TRUE,所述第一网络设备确定所述聚合端口21处于同步状态。所述第一网络设备还确定所述第三LACPDU报文中所述Actor_State包括的Collecting标志位和Distributing标志位为TRUE,从而,所述第一网络设备确定所述聚合端口21处于收集和分发状态。因此,所述第一网络设备可以根据上述确定结果,将所述Mux机由所述ATTACHED状态切换为所述COLLECTING_DISTIBUTING状态。

[0107] S1003说明了在ATTACHED状态中,所述第一网络设备已经确认所述聚合端口21处于收集和分发状态,或者可以认为是,所述聚合端口21为UP。因此,所述聚合端口11运行的Mux机不再进入PROTECT_WAITING状态,而是由ATTACHED状态直接切换为COLLECTING_

DISTRIBUTING状态。进入COLLECTING_DISTRIBUTING状态后,所述第一网络设备将所述聚合端口11设置为收集和分发状态。也就是说,所述Actor.Collecting和Actor.Distributing被确定为TRUE,表示所述聚合端口11处于收集和分发状态。所述聚合端口11被设置为UP,聚合链路1进入稳定的转发状态。

[0108] 在S1003中,可选的,所述第一网络设备将所述Single-port_UP标志位的值设置为FALSE。这样设置的目的是为了防止Mux机中可能存在Single-port_UP=TRUE的旧信息造成Mux机的误操作。

[0109] 通过上述S1003的实现方式,所述聚合端口11的Mux机处于ATTACHED状态中,当所述第一网络设备确定所述聚合端口21处于同步状态和所述聚合端口21已经处于收集和分发状态的情况下,所述第一网络设备将所述聚合端口11上运行的Mux机切换为COLLECTING_DISTRIBUTING状态。基于上述实现方式,所述第一网络设备在确定所述第二聚合端口处于收集和分发状态时,快速的进入双端口收集和分发的状态,有助于减少业务流量在传输过程中产生的丢包。

[0110] 可选的,所述方法还包括S1101和S1102,以便完成所述第一网络设备将所述第一聚合端口的Mux机由COLLECTING_DISTRIBUTING状态切换为PROTECT_WAITING状态的过程,如图4所示。

[0111] S1101、所述第一网络设备接收来自所述第二网络设备的第四LACPDU报文,所述第四LACPDU报文用于指示所述第二聚合端口未处于收集和分发状态;

[0112] S1102、当所述第一网络设备根据所述第四LACPDU报文确定所述第二聚合端口未处于收集和分发状态时,将所述Mux机由所述COLLECTING_DISTRIBUTING状态切换为所述PROTECT_WAITING状态。

[0113] 根据前述实施例的描述,当所述聚合端口11的Mux机处于COLLECTING_DISTRIBUTING状态中,所述聚合端口11和所述聚合端口21均为收集和分发状态,聚合链路1的两端均为UP,所述第一网络设备和所述第二网络设备经由所述聚合链路1互相转发数据报文。

[0114] 根据LACP,actor或partner的聚合端口发生状态信息的变化,均会通过LACPDU报文通知对方。假设所述第二网络设备发生异常,导致所述聚合端口21收集状态由收集切换为不收集和所述聚合端口21分发状态由分发切换为不分发。所述第二网络设备向所述第一网络设备发送所述第四LACPDU报文,所述第四LACPDU报文中的Actor_State包括Collecting标志位和Distributing标志位。其中,所述Collecting被确定为FALSE,表示所述聚合端口21未处于收集状态。所述Distributing被确定为FALSE,表示所述聚合端口21未处于分发状态。所述第一网络设备根据所述第四LACPDU报文携带的所述Actor_State字段中的Collecting标志位为FALSE,所述第一网络设备确定Partner.Collecting的值为FALSE。所述第一网络设备根据所述第四LACPDU报文携带的所述Actor_State字段中的Distributing标志位为FALSE,所述第一网络设备确定Partner.Distributing的值为FALSE。所述第一网络设备根据Partner.Collecting=FALSE和Partner.Distributing=FALSE,将所述Mux机由COLLECTING_DISTRIBUTING状态切换为PROTECT_WAITING状态。

[0115] 通过上述S1101和S1102的实现方式,所述聚合端口11的Mux机处于COLLECTING_DISTRIBUTING状态中,当所述第一网络设备确定所述聚合端口21未处于收集和分发状态

时,所述第一网络设备将聚合端口11的所述Mux机由COLLECTING_DISTRIBUTING状态切换为PROTECT_WAITING状态。基于上述实现方式,当所述第一网络设备收到指示所述第二聚合端口状态为未处于收集和分发状态的第四LACPDU报文后,及时将Mux机切换为PROTECT_WAITING状态,有助于减少业务流量在传输过程中产生的丢包。

[0116] 可选的,所述方法还包括S1201,以便完成所述第一网络设备将所述第一聚合端口的Mux机由PROTECT_WAITING状态切换为ATTACHED状态的过程,如图4所示。

[0117] S1201、所述第一网络设备确定以下条件中至少一个条件成立时,将所述Mux机由所述PROTECT_WAITING状态切换为所述ATTACHED状态:所述第一网络设备确定所述第一聚合端口处于不选中状态;所述第一网络设备确定所述第一聚合端口处于备用状态;和,所述第一网络设备接收来自所述第二网络设备的第五LACPDU报文,并根据所述第五LACPDU报文确定所述第二聚合端口未处于同步状态。

[0118] 举例说明,当所述聚合端口11的Mux机处于PROTECT_WAITING状态中,所述聚合端口11可能发生状态变化。例如,所述聚合端口11变更为不选择聚合器1,所述第一网络设备将聚合端口11的选中状态切换为UNSELECTED。又例如,所述第一网络设备的聚合端口11变更为选择聚合器1,但聚合端口11不使能聚合功能,所述第一网络设备将聚合端口11的选中状态切换为STANDBY。基于所述聚合端口11处于UNSELECTED状态或STANDBY状态,所述第一网络设备将所述聚合端口11的Mux机由PROTECT_WAITING状态切换为ATTACHED状态。其中,所述UNSELECTED状态可以由Selection Logic进行触发,也可以由RX machine进行触发;所述STANDBY状态可以由Selection Logic进行触发。

[0119] 当所述聚合端口11运行的Mux机处于PROTECT_WAITING状态中,所述第一网络设备还可能接收来自所述第二网络设备的第五LACPDU报文,所述第五LACPDU报文被用于确定所述聚合端口21是否处于同步状态。所述第一网络设备根据所述第五LACPDU报文确定所述第二网络设备的聚合端口21未处于同步状态。其中,所述第一网络设备根据所述第五LACPDU报文确定所述第二网络设备的聚合端口21未处于同步状态的实现方式可以参见前述S10021和S10022的相应描述,此处不进行赘述。基于聚合端口21未处于同步状态,所述第一网络设备将所述聚合端口11的Mux机由PROTECT_WAITING状态切换为ATTACHED状态。

[0120] 可选的,所述方法还包括S1301,以便完成所述第一网络设备将Mux机由COLLECTING_DISTRIBUTING状态切换为ATTACHED状态的过程,如图4所示。

[0121] S1301、所述第一网络设备确定以下条件中至少一个条件成立时,将所述第一聚合接口的所述Mux机由所述COLLECTING_DISTRIBUTING状态切换为所述ATTACHED状态:所述第一网络设备确定所述第一聚合端口处于不选中状态;所述第一网络设备确定所述第一聚合端口处于备用状态;和,所述第一网络设备接收来自所述第二网络设备的第五LACPDU报文,并根据所述第五LACPDU报文确定所述第二聚合端口未处于同步状态。

[0122] 所述第一网络设备根据触发条件,将所述聚合端口11的Mux机由COLLECTING_DISTRIBUTING状态切换为ATTACHED状态的实现方式可以参见前述S1201的相应描述,此处不进行赘述。

[0123] 可选的,在S1301后,所述方法还包括:

[0124] S1302、所述Mux机在ATTACHED状态中,所述第一网络设备停止所述定时器。

[0125] 在PROTECT_WAITING状态中,所述第一网络设备启动了所述定时器。所述第一网络

设备根据上述触发条件将所述Mux机由所述PROTECT_WAITING状态切换为所述ATTACHED状态后,所述定时器仍然在计时。在ATTACHED状态中,为了防止所述定时器超时导致Single-port_UP=TRUE,所述第一网络设备停止所述定时器的计时。

[0126] 图5为本申请实施例的第一网络设备1000的结构示意图。图5所示的第一网络设备1000可以执行上述实施例的方法中第一网络设备执行的相应步骤。所述第一网络设备1000运行有LACP,所述第一网络设备1000的第一聚合端口经由聚合链路与所述第二网络设备的第二聚合端口连接。如图5所示,所述第一网络设备1000包括处理单元1002和发送单元1004。

[0127] 所述处理单元1002,用于在所述第一聚合端口的Mux机中设置PROTECT_WAITING状态,所述PROTECT_WAITING状态用于标识所述第一聚合端口未处于收集和分发状态,但向其他网络设备通告的所述第一聚合端口的状态为收集和分发状态;

[0128] 所述处理单元1002还用于当所述Mux机处于所述PROTECT_WAITING状态时,确定所述第一聚合端口未处于收集和分发状态;

[0129] 所述发送单元1004,用于向所述第二网络设备发送第一LACPDU报文,所述第一LACPDU报文用于通告所述第二网络设备所述第一聚合端口处于收集和分发状态,并触发所述第二网络设备将所述第二聚合端口设置为收集和分发状态;

[0130] 所述处理单元1002还用于确定是否接收到来自所述第二网络设备的第二LACPDU报文,所述第二LACPDU报文用于指示所述第二聚合端口处于收集和分发状态;

[0131] 当所述处理单元1002确定接收到所述第二LACPDU报文,所述处理单元1002还用于将所述Mux机由所述PROTECT_WAITING状态切换为COLLECTING_DISTRIBUTING状态,并将所述第一聚合端口设置为收集和分发状态。

[0132] 可选的,在所述第一聚合端口的Mux机处于PROTECT_WAITING状态之前,所述Mux机处于附接ATTACHED状态,所述第一网络设备还包括接收单元。所述接收单元用于接收来自所述第二网络设备的第三LACPDU报文。当所述处理单元1002确定所述第一聚合端口处于选中状态,并且根据所述第三LACPDU报文确定所述第二聚合端口处于同步状态,以及根据所述第三LACPDU报文确定所述第二聚合端口未处于收集和分发状态时,所述处理单元1002还用于将所述Mux机由所述ATTACHED状态切换为所述PROTECT_WAITING状态。

[0133] 可选的,所述处理单元1002根据所述第三LACPDU报文确定所述第二聚合端口处于同步状态,具体包括:所述处理单元1002还用于确定所述第三LACPDU报文包括的第一信息与所述第一网络设备存储所述第一聚合端口的第二信息匹配,所述第一信息包括Partner_Port、Partner_Port_Priority、Partner_System、Partner_System_Priority、Partner_Key和Partner_State.Aggregation,所述第二信息包括Actor_Port_Number、Actor_Port_Priority、Actor_System、Actor_System_Priority、Actor_Oper_Port_Key和Actor_Oper_Port_State.Aggregation。所述处理单元1002还用于确定所述第三LACPDU报文包括的Actor_State.Synchronization处于同步状态。

[0134] 可选的,当所述处理单元1002根据所述第三LACPDU报文确定所述第二聚合端口处于同步状态,并且根据所述第三LACPDU报文确定所述第二聚合端口处于收集和分发状态时,所述处理单元1002还用于将所述Mux机由所述ATTACHED状态切换为所述COLLECTING_DISTRIBUTING状态,并将所述第一聚合端口设置为收集和分发状态。

[0135] 可选的,所述接收单元还用于接收来自所述第二网络设备的第四LACPDU报文,所述第四LACPDU报文用于指示所述第二聚合端口未处于收集和分发状态。当所述处理单元1002根据所述第四LACPDU报文确定所述第二聚合端口未处于收集和分发状态时,所述处理单元1002还用于将所述Mux机由所述COLLECTING_DISTRIBUTING状态切换为所述PROTECT_WAITING状态。

[0136] 可选的,所述处理单元1002确定以下条件中至少一个条件成立时,所述处理单元1002还用于将所述Mux机由所述PROTECT_WAITING状态切换为所述ATTACHED状态:所述处理单元1002还用于确定所述第一聚合端口处于不选中状态;所述处理单元1002还用于确定所述第一聚合端口处于备用状态;和,所述处理单元1002还用于根据接收到的来自所述第二网络设备的第五LACPDU报文确定所述第二聚合端口未处于同步状态。

[0137] 可选的,在所述第一聚合端口的Mux机处于PROTECT_WAITING状态中,所述处理单元1002还用于启动定时器。直到所述定时器超时,所述处理单元1002确定未接收到来自所述第二网络设备的所述第二LACPDU报文,所述处理单元1002还用于将所述Mux机由所述PROTECT_WAITING状态切换为ATTACHED状态。

[0138] 可选的,当所述Mux机处于PROTECT_WAITING状态,并且所述定时器超时时,所述处理单元1002还用于将所述聚合链路设置为单端口UP,所述单端口UP用于指示所述聚合链路的一端聚合端口处于收集和分发状态,另一端聚合端口未处于收集和分发状态。

[0139] 可选的,所述定时器的时长大于等于3秒,并且所述定时器的时长小于等于90秒。

[0140] 图5所示的第一网络设备可以执行上述实施例的方法中第一网络设备执行的相应步骤。应用在运行LACP的网络中,实现Mux机中的聚合链路的两端聚合端口状态的检测和处理,有助于减少业务流量在传输过程中产生的丢包。应当理解,所述图5中的结构同样适用于所述图1中的第二网络设备。

[0141] 图6为本申请实施例的第一网络设备1100的硬件结构示意图。图6所示的第一网络设备1100可以执行上述实施例的方法中第一网络设备执行的相应步骤。

[0142] 如图6所示,所述第一网络设备1100包括处理器1101、存储器1102、接口1103和总线1104。其中接口1103可以通过无线或有线的方式实现,具体来讲可以是网卡。上述处理器1101、存储器1102和接口1103通过总线1104连接。

[0143] 所述接口1103具体可以包括发送器和接收器,用于第一网络设备与上述实施例中的第二网络设备之间收发信息。作为举例,所述接口1103用于支持与所述第二网络设备收发LACPDU报文。所述处理器1101用于执行上述实施例中由第一网络设备进行的处理。例如,所述处理器1101用于设置所述第一聚合端口的收集和分发状态;还用于确定所述第一聚合端口是否处于收集和分发状态;还用于启动或停止定时器;还用于确定是否接收到第二网络设备发送的LACPDU报文;还用于实现Mux机各个状态之间的切换;和/或用于本文所描述的技术的其他过程。作为举例,所述处理器1101用于支持图3中的过程S102、S104、S106和S108。存储器1102包括操作系统11021和应用程序11022,用于存储程序、代码或指令,当处理器或硬件设备执行这些程序、代码或指令时可以完成方法实施例中涉及第一网络设备的处理过程。可选的,所述存储器1102可以包括只读存储器(英文:Read-only Memory,缩写:ROM)和随机存取存储器(英文:Random Access Memory,缩写:RAM)。其中,所述ROM包括基本输入/输出系统(英文:Basic Input/Output System,缩写:BIOS)或嵌入式系统;所述RAM包

括应用程序和操作系统。当需要运行第一网络设备1100时,通过固化在ROM中的BIOS或者嵌入式系统中的bootloader引导系统进行启动,引导第一网络设备1100进入正常运行状态。在第一网络设备1100进入正常运行状态后,运行在RAM中的应用程序和操作系统,从而,完成方法实施例中涉及第一网络设备的处理过程。

[0144] 可以理解的是,图6仅仅示出了第一网络设备1100的简化设计。在实际应用中,第一网络设备可以包含任意数量的接口,处理器或者存储器。另外,本实施例仅以第一网络设备为例进行说明,应当理解,第二网络设备,或者更多的网络设备具有与所述第一网络设备相同的功能,此处不再一一赘述。

[0145] 图7为本申请实施例的另第一网络设备1200的硬件结构示意图。图7所示的第一网络设备1200可以执行上述实施例的方法中第一网络设备执行的相应步骤。

[0146] 如图7所述,第一网络设备1200包括:主控板1210、接口板1230、交换网板1220和接口板1240。主控板1210用于完成系统管理、设备维护、协议处理等功能。交换网板1220用于完成各接口板(接口板也称为线卡或业务板)之间的数据交换。接口板1230和1240用于提供各种业务接口(例如,POS接口、GE接口、ATM接口等),并实现数据包的转发。主控板1210、接口板1230和1240,以及交换网板1220之间通过系统总线与系统背板相连实现互通。接口板1230上的中央处理器1231用于对接口板进行控制管理并与主控板上的中央处理器进行通信。

[0147] 接口板1230上的物理接口卡1233接收来自所述第二网络设备的LACPDU报文,并且,经由接口板1230上的中央处理器1231向主控板1210上的中央处理器1211上送所述LACPDU报文。主控板1210上的中央处理器1211用于获取所述LACPDU报文。

[0148] 中央处理器1211还用于在所述第一聚合端口的Mux机中设置PROTECT_WAITING状态,所述PROTECT_WAITING状态用于标识所述第一聚合端口未处于收集和分发状态,但向其他网络设备通告的所述第一聚合端口的状态为收集和分发状态。所述中央处理器1211还用于当所述Mux机处于所述PROTECT_WAITING状态时,确定所述第一聚合端口未处于收集和分发状态。所述中央处理器1211还用于控制所述接口板1230向所述第二网络设备发送第一LACPDU报文,所述第一LACPDU报文用于通告所述第二网络设备所述第一聚合端口处于收集和分发状态,并触发所述第二网络设备将所述第二聚合端口设置为收集和分发状态。所述中央处理器1211还用于确定是否接收到来自所述第二网络设备的第二LACPDU报文,所述第二LACPDU报文用于指示所述第二聚合端口处于收集和分发状态。当所述中央处理器1211确定接收到所述第二LACPDU报文,所述中央处理器1211还用于将所述Mux机由所述PROTECT_WAITING状态切换为收集分发COLLECTING_DISTRIBUTING状态,并将所述第一聚合端口设置为收集和分发状态。

[0149] 中央处理器1211还用于生成LACPDU报文。中央处理器1211将生成的LACPDU报文由经由接口板1230上的中央处理器1231向物理接口卡1233发送。接口板1230上的物理接口卡1233向所述第二网络设备发送所述LACPDU报文。

[0150] 接口板1230上的转发表项存储器1234用于保存转发表项。接口板1230上的中央处理器1231用于控制网络存储器1232获取转发表项存储器1234中的转发表项,并且,中央处理器1231用于控制网络存储器1232经由物理接口卡1233完成流量的接收和发送。

[0151] 应理解,本发明实施例中接口板1240上的操作与所述接口板1230的操作一致,为

了简洁,不再赘述。应理解,本实施例的第一网络设备1200可对应于上述方法实施例所具有的功能和/或所实施的各种步骤,在此不再赘述。另外,本实施例仅以第一网络设备为例进行说明,应当理解,第二网络设备,或者更多的网络设备具有与所述第一网络设备相同的功能,此处不再一一赘述。

[0152] 此外,需要说明的是,主控板可能有一块或多块,有多块的时候可以包括主用主控板和备用主控板。接口板可能有一块或多块,第一网络设备的数据处理能力越强,提供的接口板越多。接口板上的物理接口卡也可以有一块或多块。交换网板可能没有,也可能有一块或多块,有多块的时候可以共同实现负荷分担冗余备份。在集中式转发架构下,第一网络设备可以不需要交换网板,接口板承担整个系统的业务数据的处理功能。在分布式转发架构下,第一网络设备可以有至少一块交换网板,通过交换网板实现多块接口板之间的数据交换,提供大容量的数据交换和处理能力。所以,分布式架构的第一网络设备的数据接入和处理能力要大于集中式架构的设备。具体采用哪种架构,取决于具体的组网部署场景,此处不做任何限定。

[0153] 图8为本申请实施例的第一网络系统1300的硬件结构示意图。图8所示的第一网络系统1300可以执行上述实施例的方法中第一网络设备执行的相应步骤。

[0154] 第一网络系统1300的这种产品形态适用于基于控制与转发分离的网络架构(例如,软件定义网络(英文:Software Defined Network,缩写:SDN))。在SDN中,如图7所示的第一网络设备1200的主控板1210从设备中分离出来,形成新的独立的物理设备(即如图8所示的控制器1210A),剩下的形成另一独立的物理设备(即如图8所示的第一转发设备1200A)。控制器1210A与第一转发设备1200A通过控制通道协议实现交互。控制通道协议可以是开放流(英文:OpenFlow)协议、路径计算通信协议(英文:Path Computation Element Communication Protocol,缩写:PCEP)、BGP、路由系统接口(英文:Interface to the Routing System,缩写:I2RS)等。也就是说,与上述图7所对应的实施例相比,本实施中的第一网络系统1300包括分离出去的控制器1210A和第一转发设备1200A。

[0155] 控制器1210A可以是基于通用的物理服务器实现或者是专用的硬件结构实现,在一个设计示例中,所述控制器包括接收器、处理器、发送器、RAM、ROM以及总线(图中未示出)。其中,处理器通过总线分别耦接接收器、发送器、RAM以及ROM。其中,当需要运行控制器时,通过固化在ROM中的BIOS或者嵌入式系统中的bootloader引导系统进行启动,引导控制器进入正常运行状态。在控制器进入正常运行状态后,在RAM中运行应用程序和操作系统,使得该处理器执行上述图7中主控板1210的所有功能和步骤。

[0156] 第一转发设备1200A可以是基于专用的硬件结构实现,其功能和结构与上述图7中的接口板1230、接口板1240和交换网板1220的功能和结构保持一致,执行相应的功能和步骤。也可以是基于通用的物理服务器和网络功能虚拟化(英文:Network Function Virtualization,缩写:NFV)技术实现的虚拟第一转发设备,所述虚拟第一转发设备为虚拟路由器。在虚拟第一转发设备的场景下,上述实体第一转发设备实施例中提到的该第一转发设备包括接口板、交换网板以及处理器在虚拟环境下可以认为是其所基于通用的物理服务器分配给该虚拟第一转发设备所使用的接口资源、网络资源以及处理资源。采用通用物理服务器实施该第一转发设备的功能或步骤,或者采用通用物理服务器并利用NFV技术实施该第一转发设备的功能或步骤具体可以参考图6的实施例。

[0157] 应理解,本实施例中第一网络系统1300中的控制器1210A和该第一转发设备1200A可以实现方法实施例中的第一网络设备所实施的各种功能、步骤,为了简洁,在此不再赘述。另外,本实施例仅以第一网络设备为例进行说明,应当理解,第二网络设备,或者更多的网络设备具有与所述第一网络设备相同的功能,此处不再一一赘述。

[0158] 另外,本申请实施例提供了一种计算机存储介质,用于储存为上述第一网络设备所用的计算机软件指令,其包含用于执行上述方法实施例所设计的程序。

[0159] 如图1所示,本申请实施例还包括一种实现Mux机的网络系统,所述网络系统包括第一网络设备和第二网络设备,所述第一网络设备和/或第二网络设备为前述图5或图6或图7;或者所述第一网络设备和/或第二网络设备为前述图8中的第一网络系统。

[0160] 结合本申请公开内容所描述的方法或者算法的步骤可以硬件的方式来实现,也可以是由处理器执行软件指令的方式来实现。软件指令可以由相应的软件模块组成,软件模块可以被存放于RAM存储器、闪存、ROM存储器、EPROM存储器、EEPROM存储器、寄存器、硬盘、移动硬盘、CD-ROM或者本领域熟知的任何其它形式的存储介质中。一种示例性的存储介质耦合至处理器,从而使处理器能够从该存储介质读取信息,且可向该存储介质写入信息。当然,存储介质也可以是处理器的组成部分。处理器和存储介质可以位于ASIC中。另外,该ASIC可以位于用户设备中。当然,处理器和存储介质也可以作为分立组件存在于用户设备中。

[0161] 本领域技术人员应该可以意识到,在上述一个或多个示例中,本申请所描述的功能可以用硬件、软件、固件或它们的任意组合来实现。当使用软件实现时,可以将这些功能存储在计算机可读介质中或者作为计算机可读介质上的一个或多个指令或代码进行传输。计算机可读介质包括计算机存储介质和通信介质,其中通信介质包括便于从一个地方向另一个地方传送计算机程序的任何介质。存储介质可以是通用或专用计算机能够存取的任何可用介质。

[0162] 以上所述的具体实施方式,对本申请的目的、技术方案和有益效果进行了进一步详细说明。所应理解的是,以上所述仅为本申请的具体实施方式而已。

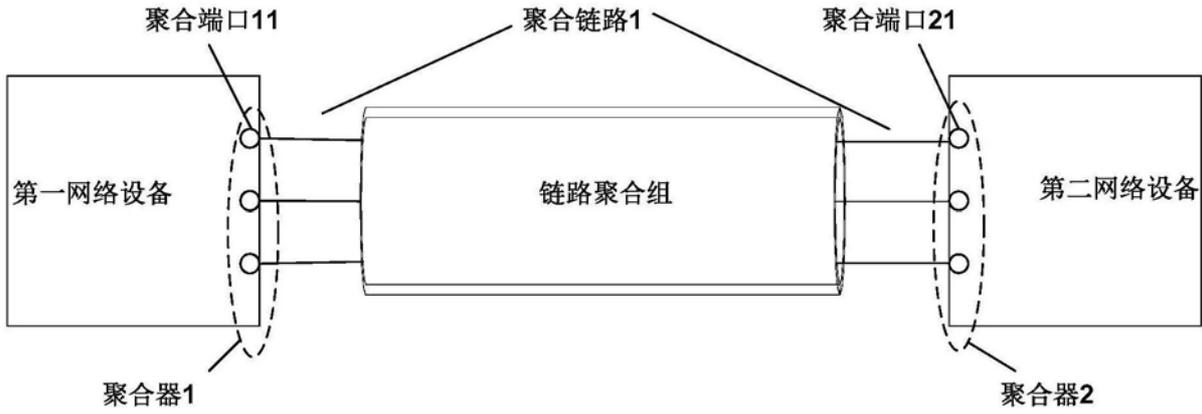


图1

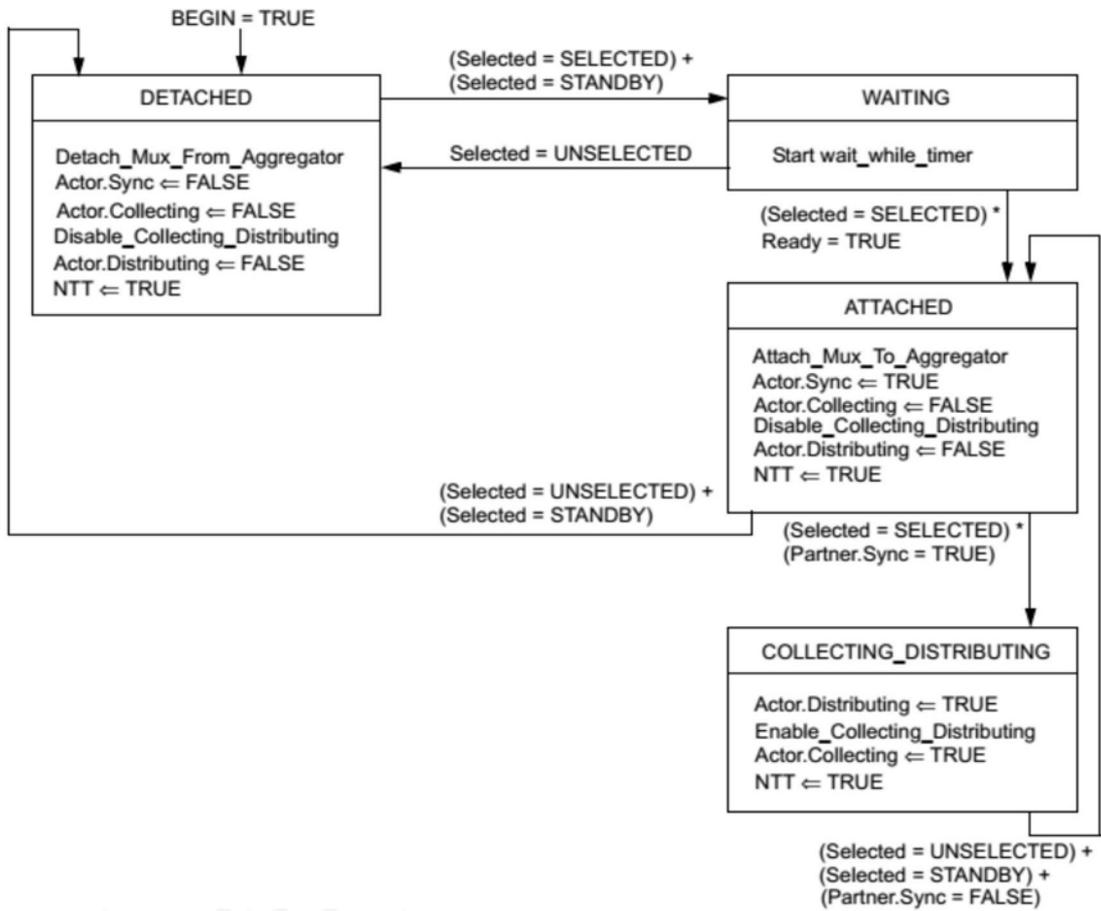


图2

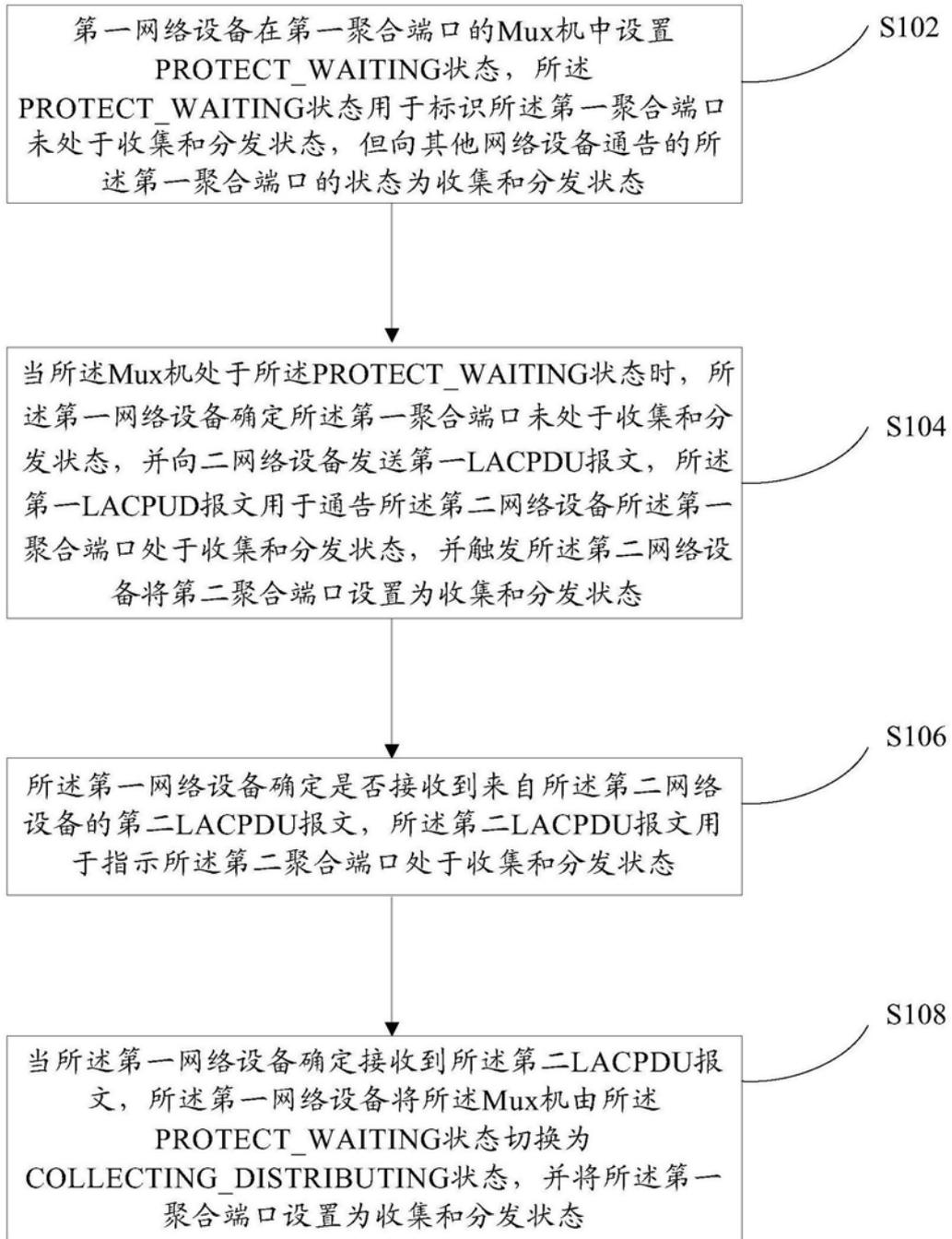


图3

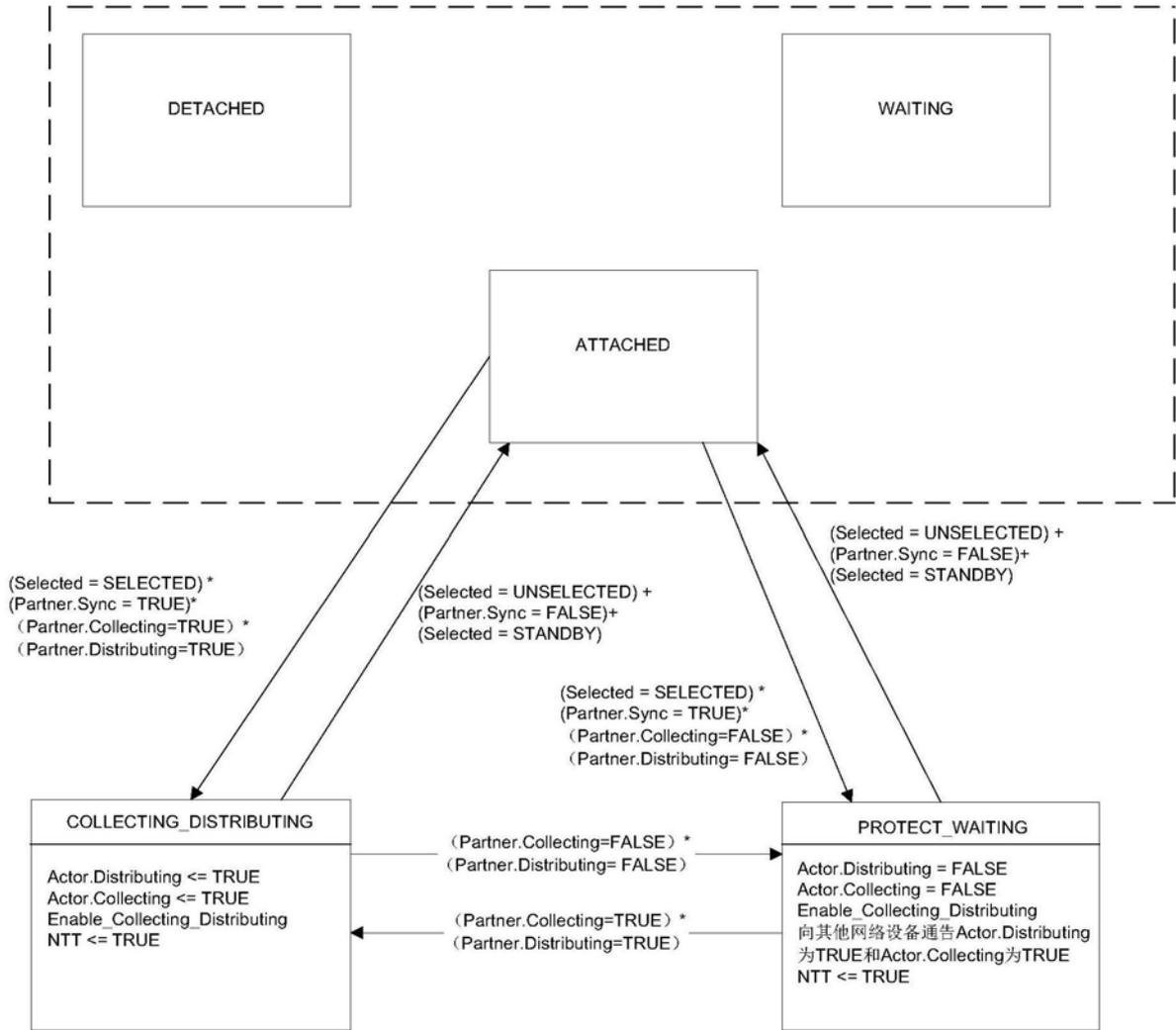


图4

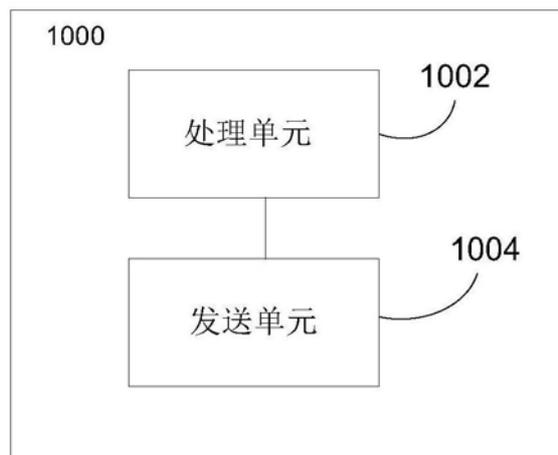


图5

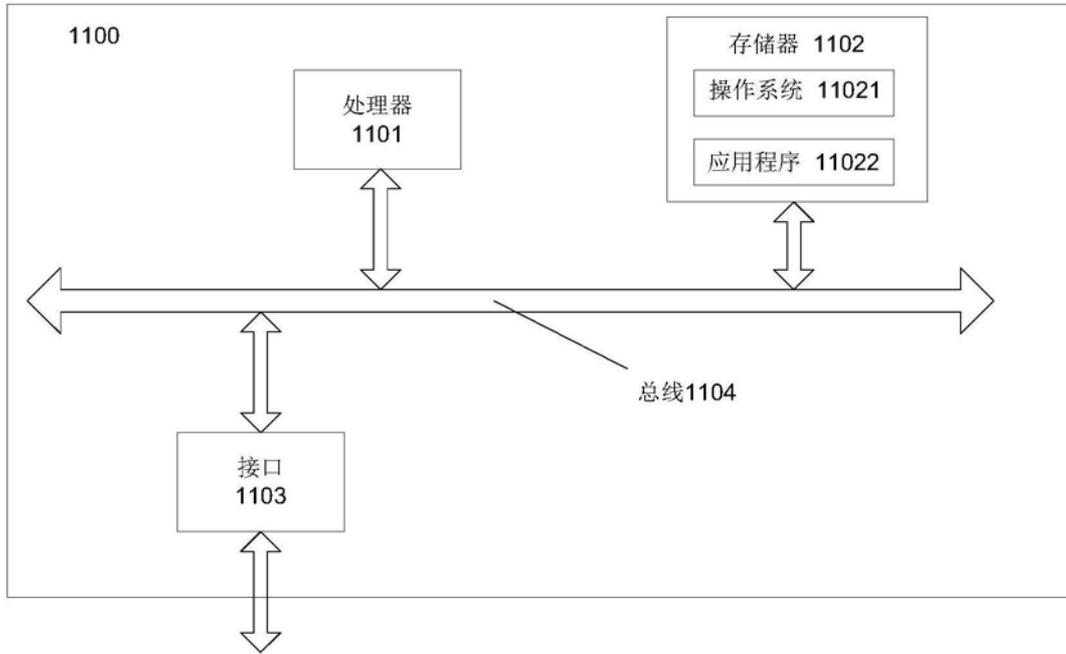


图6

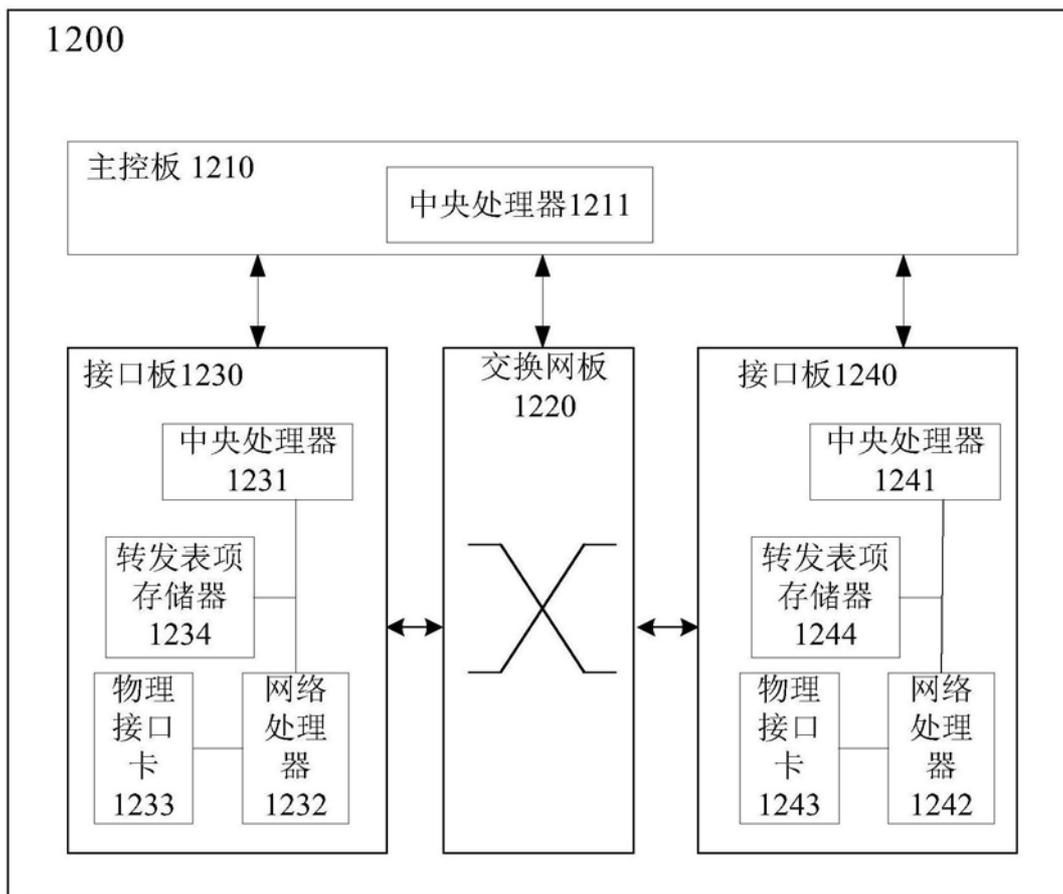


图7

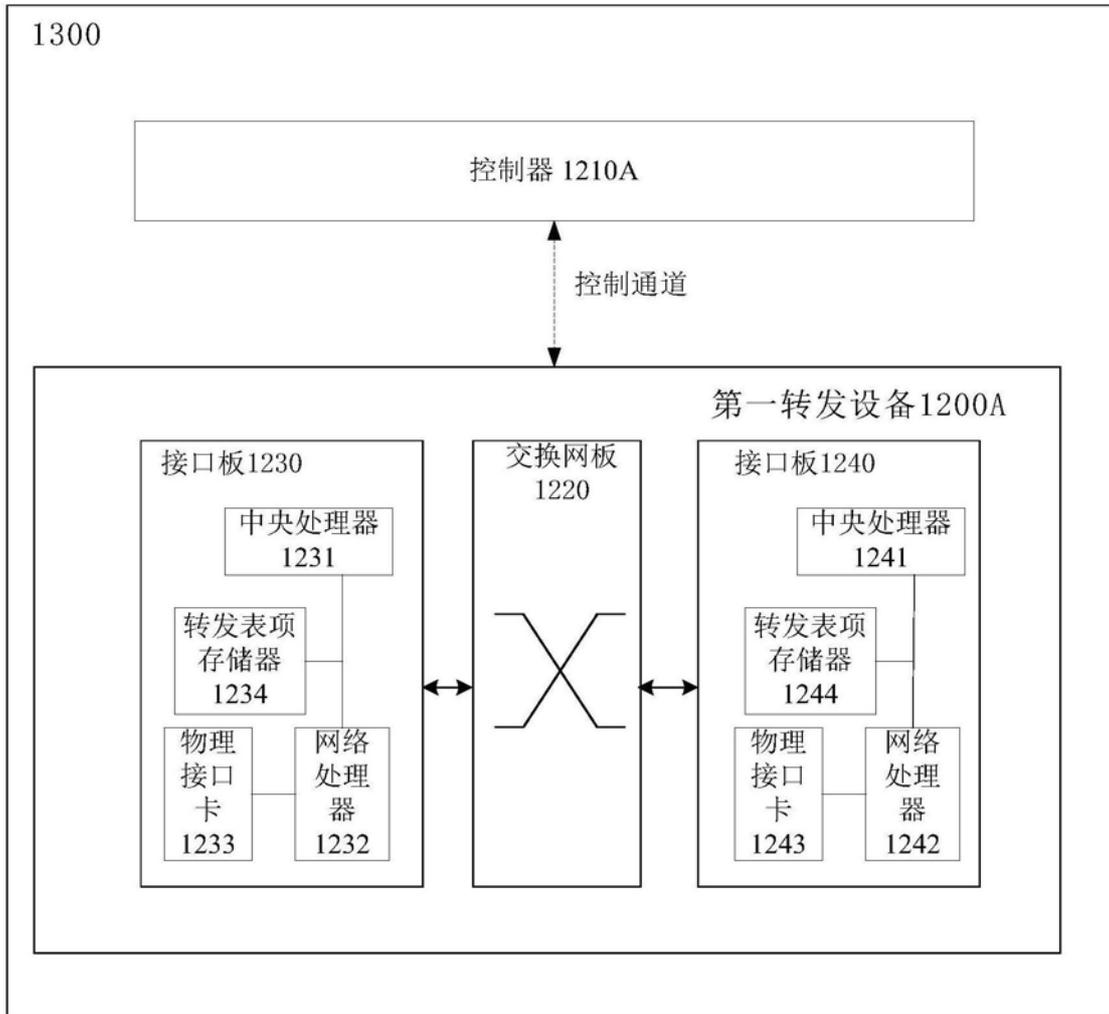


图8