

一种通信方法、系统、装置、网络设备、目标设备

技术领域

本申请涉及通信技术领域，特别是涉及一种通信方法、系统、装置、网络设备、目标设备。

背景技术

5 MLAG (Multichassis Link Aggregation Group, 跨设备链路聚合组) 是一种实现跨设备链路聚合的机制, 可以将多个网络设备在逻辑上虚拟成一台网络设备, 形成一个统一的二层逻辑节点。相当于将多台网络设备作为一台网络设备进行流量转发, 从而把链路可靠性从单板级提高到了设备级。

 参见图 1, 为相关技术中提供的一种 MLAG 场景示意图。如下图所示, 网络设备 A 与网络设备 B 互为邻居, 组成 MLAG 系统, MLAG 系统中的网络设备也可以被称为 MLAG 设备。网络设备 A 为主设备, 网络设备 B 为从设备。网络设备 A 与网络设备 B 的 MAC (Media Access Control Address, 媒体存取控制位址) 地址被配置为同一 MAC 地址, LACP (Link Aggregation Control Protocol, 链路汇聚控制协议) 优先级被配置为同一 LACP 系统优先级, 从而被虚拟成一台网络设备, 形成负载分担, 共同进行流量转发。图中示出的 MLAG 系统中包含两台网络设备仅为一种示例, 本申请对 MLAG 系统中网络设备的数量不进行限定。

15 网络设备 C 分别与网络设备 A 和网络设备 B 相连。且网络设备 C 与用户设备相连, 网络设备 A 与网络设备 B 分别与 IP (Internet Protocol, 网际互连协议) 网络相连。

 MLAG 为每个网络设备 (即网络设备 A 与网络设备 B) 定义的接口角色包括: MLAG 接口、peer-link (Intra-Portal, 内部控制链路) 接口、keepalive (保活) 接口。其中, MLAG 接口是网络设备与外部设备连接的二层聚合接口。图 1 中网络设备 A、网络设备 B 上与网络设备 C 和 IP 网络相连接接口均为 MLAG 接口。与外部设备上相同聚合组相连的 MLAG 接口属于同一 MLAG 组。图 1 中网络设备 A 上的 MLAG 接口 1 与网络设备 B 上的 MLAG 接口 2 属于同一 MLAG 组。

20 另外, peer-link 接口用于连接对端网络设备实现内部控制。MLAG 系统中的每台网络设备仅有一个 peer-link 接口。网络设备的 peer-link 接口之间的链路是 peer-link 链路, 网络设备通过 peer-link 链路交互协议报文以及传输数据流量。每个 MLAG 系统中仅有一条 peer-link 链路。

25 网络设备之间通过 keepalive 链路交互 keepalive 报文检测邻居的状态, 即通过交互 keepalive 报文来进行 peer-link 链路故障时的双主检测。

 在 MLAG 系统正常运行时, 网络设备 C 向网络设备 A 或网络设备 B 转发流量均可以实现网络通信。当 MLAG 系统中的一台网络设备发生故障时, 流量可以快速切换到另一台网络设备, 保证业务的正常运行, 也就是链路可靠性被提高到设备级。

30 相关技术中, 如果网络设备之间的 peer-link 链路与 keepalive 链路发生故障, 则 MLAG 系统分裂。为了避免 MLAG 系统分类后网络设备均作为主设备进行流量转发, 需要配置各个网络设备进入独立模式, 开始独立工作。配置独立模式时各个网络设备的 MAC 地址均会发生变化, 在配置 MAC 地址变化期间, 网络设备重新学习表项的过程中无法转发流量, 导致 MLAG 接口发生震荡。

35 尤其在一台网络设备由于故障或重启进入关闭状态的情况下, 由于网络设备整体关闭导致网络设备之间的 peer-link 链路与 keepalive 链路断开, 其他的网络设备会误判进入独立模式。在此过程中所有网络设备均无法进行流量转发, 会导致 MLAG 接口发生不必要的震荡, 影响网络流量转发。

发明内容

 本申请实施例的目的在于提供一种通信方法、系统、装置、网络设备、目标设备, 保证 MLAG 系统中的部分网络设备进入关闭状态, 不会影响网络流量的转发。具体技术方案如下:

40 第一方面, 本申请实施例提供了一种通信方法, 应用于组成跨设备聚合组 MLAG 系统的多个网络设备中的任一网络设备, 所述方法包括:

接收目标设备发送的携带路径开销值的生成树协议STP报文，其中，所述目标设备包括多个接口，每一接口用于接入所述MLAG系统中的一个网络设备，所述路径开销值用于指示：所述多个接口中处于正常状态的接口数量；

5 若所述多个网络设备之间已建立的保活keepalive链路与内部控制peer-link链路均断开，且确定第一路径开销值与第二路径开销值不同，则维持所述MLAG系统，其中，所述第一路径开销值为：keepalive链路
与peer-link链路均断开之前接收到的STP报文中携带的路径开销值，所述第二路径开销值为：
keepalive链路与peer-link链路均断开之后接收到的STP报文中携带的路径开销值。

10 第二方面，本申请实施例提供了一种通信方法，应用于接入跨设备聚合组MLAG系统的目标设备，所述MLAG系统由多个网络设备组成，所述目标设备包括多个接口，每一接口用于接入所述MLAG系统
中的一个网络设备，所述方法包括：

基于哈希hash原则，向所述MLAG系统中的网络设备发送第一生成树协议STP报文，所述第一STP
报文包括第一路径开销值；

15 基于hash原则，向所述MLAG系统中的网络设备发送第二STP报文，所述第二STP报文包括第二路
径开销值，以使得所述MLAG系统中的任一网络设备在确定所述多个网络设备之间已建立的保活
keepalive链路与内部控制peer-link链路均断开，且第一路径开销值与第二路径开销值不同的情况下，维
持所述MLAG系统；

其中，所述第一路径开销值为：keepalive链路与peer-link链路均断开之前所述网络设备接收到的STP
报文中携带的路径开销值，所述第二路径开销值为：keepalive链路与peer-link链路均断开之后所述网
络设备接收到的STP报文中携带的路径开销值。

20 第三方面，本申请实施例提供了一种通信系统，所述系统包括组成跨设备聚合组MLAG系统的多个
网络设备、接入所述MLAG系统的目标设备，所述目标设备包括多个接口，每一接口用于接入所述MLAG
系统中的一个网络设备；

所述多个网络设备中的任一网络设备，用于实现前述第一方面中任一项的方法步骤；

所述目标设备，用于实现前述第二方面中任一项的方法步骤。

25 第四方面，本申请实施例提供了一种网络设备，所述网络设备为：组成跨设备聚合组MLAG系统的
多个网络设备中的任一网络设备，所述网络设备包括：

处理器；

收发器；

30 机器可读存储介质，所述机器可读存储介质存储有能够被所述处理器执行的机器可执行指令；所述
机器可执行指令促使所述处理器执行以下步骤：

接收目标设备发送的携带路径开销值的生成树协议STP报文，其中，所述目标设备包括多个接口，
每一接口用于接入所述MLAG系统中的一个网络设备，所述路径开销值用于指示：所述多个接口中处于
正常状态的接口数量；

35 若所述多个网络设备之间已建立的保活keepalive链路与内部控制peer-link链路均断开，且确定第一
路径开销值与第二路径开销值不同，则维持所述MLAG系统，其中，所述第一路径开销值为：keepalive
链路
与peer-link链路均断开之前接收到的STP报文中携带的路径开销值，所述第二路径开销值为：
keepalive链路与peer-link链路均断开之后接收到的STP报文中携带的路径开销值。

40 第五方面，本申请实施例提供了一种目标设备，所述目标设备接入跨设备聚合组MLAG系统，所述
MLAG系统由多个网络设备组成，所述目标设备包括多个接口，每一接口用于接入所述MLAG系统中的
一个网络设备，所述目标设备包括：

处理器；

收发器；

机器可读存储介质，所述机器可读存储介质存储有能够被所述处理器执行的机器可执行指令；所述机器可执行指令促使所述处理器执行以下步骤：

基于哈希hash原则，向所述MLAG系统中的网络设备发送第一生成树协议STP报文，所述第一STP报文包括第一路径开销值；

5 基于hash原则，向所述MLAG系统中的网络设备发送第二STP报文，所述第二STP报文包括第二路径开销值以使得所述MLAG系统中的任一网络设备在确定所述多个网络设备之间已建立的保活keepalive链路与内部控制peer-link链路均断开，且第一路径开销值与第二路径开销值不同的情况下，维持所述MLAG系统；

10 其中，所述第一路径开销值为：keepalive链路与peer-link链路均断开之前所述网络设备接收到的STP报文中携带的路径开销值，所述第二路径开销值为：keepalive链路与peer-link链路均断开之后所述网络设备接收到的STP报文中携带的路径开销值。

第六方面，本申请实施例提供了一种通信装置，应用于组成跨设备聚合组MLAG系统的多个网络设备中的任一网络设备，所述装置包括：

15 STP报文接收模块，用于接收目标设备发送的携带路径开销值的生成树协议STP报文，其中，所述目标设备包括多个接口，每一接口用于接入所述MLAG系统中的一个网络设备，所述路径开销值用于指示：所述多个接口中处于正常状态的接口数量；

20 系统维持模块，用于若所述多个网络设备之间已建立的保活keepalive链路与内部控制peer-link链路均断开，且确定第一路径开销值与第二路径开销值不同，则维持所述MLAG系统，其中，所述第一路径开销值为：keepalive链路与peer-link链路均断开之前接收到的STP报文中携带的路径开销值，所述第二路径开销值为：keepalive链路与peer-link链路均断开之后接收到的STP报文中携带的路径开销值。

第七方面，本申请实施例提供了一种通信装置，应用于接入跨设备聚合组MLAG系统的目标设备，所述MLAG系统由多个网络设备组成，所述目标设备包括多个接口，每一接口用于接入所述MLAG系统中的一个网络设备，所述装置包括：

25 第一STP报文发送模块，用于基于哈希hash原则，向所述MLAG系统中的网络设备发送第一生成树协议STP报文，所述第一STP报文包括第一路径开销值；

 第二STP报文发送模块，用于基于hash原则，向所述MLAG系统中的网络设备发送携带路径开销值的第二STP报文，所述第二STP报文包括第二路径开销值，以使得所述MLAG系统中的任一网络设备在确定所述多个网络设备之间已建立的保活keepalive链路与内部控制peer-link链路均断开，且第一路径开销值与第二路径开销值不同的情况下，维持所述MLAG系统；

30 其中，所述第一路径开销值为：keepalive链路与peer-link链路均断开之前所述网络设备接收到的STP报文中携带的路径开销值，所述第二路径开销值为：keepalive链路与peer-link链路均断开之后所述网络设备接收到的STP报文中携带的路径开销值。

第八方面，本申请实施例提供了一种机器可读存储介质，存储有机器可执行指令，在被处理器调用和执行时，所述机器可执行指令促使所述处理器：实现第一方面或第二方面任一所述的方法。

35 第九方面，一种计算机程序产品，所述计算机程序产品促使处理器：实现第一方面或第二方面任一所述的方法。

40 本申请实施例有益效果：本申请实施例提供的通信方法中，在MLAG系统中的多个网络设备中的任一网络设备确定多个网络设备之间已建立的keepalive链路与peer-link链路均断开的情况下，不会直接进入MLAG系统的分裂状态。而是确定接入MLAG系统的目标设备发送的STP报文中携带的路径开销值是否发生变化。若该网络设备之外的其他网络设备重启或故障，导致上述其他网络设备未正常工作。则上述其他网络设备的所有接口均会处于关闭状态，导致该其他网络设备与目标设备之间的通道断开，目标设备上与其他网络设备相连的接口也会变为关闭状态。在此情况下，目标设备上处于正常状态

的接口的数量会发生变化，进而目标设备发送的 STP 报文中携带的路径开销值会发生变化。若网络设备确定第一路径开销值和第二路径开销值不同，则可以确定上述其他网络设备的所有接口均关闭，上述其他网络设备当前未正常工作。

在此情况下，该网络设备不会由于多个网络设备之间的链路断开，直接进入 MLAG 系统的分裂状态。而是会结合 STP 报文携带的第一路径开销值和第二路径开销值是否相同，进行进一步判断。在确定第一路径开销值与第二路径开销值不同，即上述其他网络设备当前未正常工作的情况下，维持 MLAG 系统，继续保持自身当前的 MAC 地址和优先级进行流量转发。从而避免 MAC 地址与优先级变化器件 MLAG 接口发生震荡，避免影响网络流量的转发。

附图说明

为了更清楚地说明本申请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申请的一些实施例，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实施例。

图 1 为相关技术中提供的一种 MLAG 场景示意图；

图 2 为本申请实施例提供的一种通信系统的结构示意图；

图 3 为本申请实施例提供的第一种应用场景示意图；

图 4 为本申请实施例提供的第二种应用场景示意图；

图 5 为本申请实施例提供的第一种通信方法的流程示意图；

图 6 为本申请实施例提供的第三种应用场景示意图；

图 7 为本申请实施例提供的第二种通信方法的流程示意图；

图 8 为本申请实施例提供的一种网络设备的结构示意图；

图 9 为本申请实施例提供的一种目标设备的结构示意图；

图 10 为本申请实施例提供的第一种通信装置的结构示意图；

图 11 为本申请实施例提供的第二种通信装置的结构示意图。

具体实施方式

下面将结合本申请实施例中的附图，对本申请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申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

相关技术中，对于 Layer2（层二）而言，MLAG 技术可以被理解作为一种横向虚拟化技术，将多台网络设备在逻辑上虚拟成一台设备，形成一个统一的层二逻辑节点。MLAG 技术提供了一个没有环路的二层拓扑，且同时实现冗余备份，不再需要繁琐的生成树协议配置，极大地简化了组网和配置。

在网络设备之间的 peer-link 链路与 keepalive 链路均处于 DOWN（关闭）状态时，MLAG 系统进入分裂状态，当前主要有以下方式（一）-（二）管理 MLAG 系统的分裂。

方式（一）MAD（Multi-Active Detection，多活检测）DOWN 保持功能。

在 peer-link 链路故障，keepalive 链路正常工作的情况下，主设备正常工作，从设备会自动关闭从设备上除 IRF（Intelligent Resilient Framework，智能弹性架构）保留接口、MLAG 保留接口外的其他接口。将关闭的接口配置为 MLAG MAD DOWN 状态。

在此情况下，若 keepalive 链路也发生故障，则从设备上被配置为 MLAG MAD DOWN 状态的接口会升级为主设备。在此情况下，网络设备都可以作为主设备转发流量，引起网络故障。为了避免以上情况，可以开启 MLAG MAD DOWN 保持功能，是从设备上的接口持续处于 MLAG MAD DOWN 状态，不参与流量转发。

但方式（一）仅适用于 peer-link 链路先故障，keepalive 链路后故障的场景。如果 keepalive 链路先故障，peer-link 链路后故障，该机制就无法起效。

方式（二）网络设备独立模式。

当 MLAG 系统分裂时，为了避免 MLAG 系统中的网络设备都作为主设备转发流量，需要配置网络设备独立工作。在 peer-link 链路和 keepalive 链路均处于 DOWN 状态时，从设备会立即或经过一段时间后开始切换到独立模式。

5 从设备切换到独立模式后，聚合接口发送的 LACP 报文中携带的 MLAG 系统参数还原为接口原本的 MAC 地址和 LACP 系统优先级，也就是将接口还原为被配置为 MLAG 系统中的接口之前的状态。从而使同一 MLAG 系统中的聚合接口的 MAC 地址和 LACP 系统优先级不一致。这样两个网络设备开始独立工作，在进行流量转发时只有一边聚合接口的成员端口可以被选中。通过被选中的设备转发业务流量，可以避免流量转发异常。

10 但虽然方式（二）适用于方式（一）不适用的场景，但采用方式（二）进行 MLAG 系统分裂管理时，若一个网络设备由于故障或重启进入关闭状态，其他的网络设备会误判进入独立模式，使得 MLAG 系统分裂。在 MLAG 系统进入分裂配置期间，所有网络设备均无法进行流量转发，会导致 MLAG 接口发生不必要的震荡，影响网络流量转发。

为了解决上述问题，本申请实施例提供了一种通信方法、系统、装置、网络设备、目标设备。

15 参见图2，为本申请实施提供的一种通信系统的结构示意图。上述系统包括组成MLAG系统20的多个网络设备201、接入MLAG系统20的目标设备202，上述目标设备202包括多个接口，每一接口用于接入上述MLAG系统20中的一个网络设备201。

参见图3，为本申请实施例提供的第一种应用场景示意图。

20 此场景中包含网络设备A1、网络设备A2、网络设备A3、网络设备A4以及网络设备A5。网络设备A1与网络设备A2组成MLAG系统，两者对应图2中的网络设备201。网络设备A1与网络设备A2之间的虚线连线表示keepalive链路。网络设备A1与网络设备A2之间的实线连线表示peer-link链路。网络设备A1和网络设备A2分别和网络设备A3以及网络设备A4通过MLAG接口连接。网络设备A3和网络设备A4对应图2中的目标设备202。例如，上述MLAG接口的数据传输速率为千兆级别。

网络设备A1-网络设备A4开启STP（Spanning Tree Protocol，生成树协议）。

25 网络设备A1与网络设备A5之间通过三层口ECMP（Equal-Cost Multipath Routing，等价路由）的方式连接。

在此场景中，各个网络设备均处于正常状态，因此各个接口也均处于正常状态。

在图3所示的应用场景的基础上，若网络设备A1重启，则应用场景发生变化。在此情况下，参见图4，为本申请实施例提供的第二种应用场景示意图。

30 在网络设备A1重启的过程中，网络设备A1关闭，因此与网络设备A1相连的所有链路均处于关闭状态。在图中以“×号”表示。除此之外，图4所示的应用场景中的其他部分与图3相同，在此不再赘述。

需要说明的是，图3和图4示出的MLAG系统中包含的网络设备的数量以及目标设备的数量仅为一种示例，本申请实施例两者的数量对此不进行具体限定。

35 基于图3与图4所示的应用场景，参见图5，为本申请实施例提供的第一种通信方法的流程示意图，应用于MLAG系统的多个网络设备中的任一网络设备，上述方法包括以下步骤S501-步骤S502。

步骤S501：接收目标设备发送的携带路径开销值的STP报文。

其中，上述目标设备包括多个接口，每一接口用于接入上述MLAG系统中的一个网络设备，上述路径开销值用于指示：上述多个接口中处于正常状态的接口数量。

40 以前述图3所示的场景为例，图中的所有接口都处于正常状态。在此情况下，网络设备A3与网络设备A4上处于正常状态的接口的数量均为2。因此，在此情况下，网络设备A3与网络设备A4向网络设备A1和网络设备A2上报的STP报文中携带的路径开销值表示的正常状态的接口的数量为2。STP报文的发送遵循hash（哈希）原则，STP报文可能被发送至网络设备A1，或者被发送至网络设备A2。网络设备

A1与网络设备A2中的一个会接收到STP报文。

但若网络设备A1重启期间，网络设备A1的接口关闭。参见图4，网络设备A3与网络设备A4上处于正常状态的接口的数量均下降为1。因此，在此情况下，网络设备A3与网络设备A4向网络设备201上报的STP报文中携带的路径开销值与图3所示的应用场景相比发生变化，变更为表示接口的数量为1的路径开销值。由于网络设备A1的重启，接口关闭，因此STP报文不可能被发送至网络设备A1，只有网络设备A2会接收到上述STP报文。

参见表1，为本申请实施例提供的一种链路速率、端口类型与路径开销值之间的对应关系表。

表1

链路速率	接口类型	路径开销值		
		IEEE 802.1D-1998	IEEE 802.1t	其他标准
1000Mbps	单个接口	4	20000	20
	聚合接口（含两个接口）		10000	18
	聚合接口（含三个接口）		6666	16
	聚合接口（含四个接口）		5000	14

其中，表1中链路速率为1000Mbps仅为一种示例，仅是为了对路径开销值进行举例设置的一种链路速率的可能取值，本申请对链路速率不进行限定。不同的接口类型分别对应设备连接的处于正常状态的接口的不同数量。由此表可见，若采用IEEE 802.1D-1998标准，则无论设备所连接的处于正常状态的接口的数量为何，路径开销值均相同。因此IEEE 802.1D-1998标准不适用于本方案。但IEEE 802.1t标准适用于本方案。上述其他标准为除IEEE规定的标准之外的自定义标准，其他标准中路径开销值的取值仅为一种示例，只要针对不同的接口类型路径开销值不同，则该标准即可适用于本申请实施例。

另外，上述STP报文包括MST BPDU（Multiple Spanning Tree Bridge Protocol Data Unit，多生成树桥协议数据单元）字段，上述MST BPDU字段承载上述路径开销值。

步骤S502：若上述多个网络设备之间已建立的keepalive链路与peer-link链路均断开，且确定第一路径开销值与第二路径开销值不同，则维持上述MLAG系统。

其中，上述第一路径开销值为：keepalive链路与peer-link链路均断开之前接收到的STP报文中携带的路径开销值，上述第二路径开销值为：keepalive链路与peer-link链路均断开之后接收到的STP报文中携带的路径开销值。

具体的，上述第一路径开销值可以为图3所示的应用场景中的路径开销值，上述第二路径开销值可以为图4所示的应用场景中的路径开销值。

具体的，若网络设备201确定多个网络设备之间已建立的keepalive链路与peer-link链路均断开，但路径开销值没有发生变化，则可以确定不仅多个网络设备之间的链路断开，目标设备上接口也关闭。在此情况下，MLAG系统中的其他网络设备应当是整个设备整体故障或关闭。在此情况下，未故障的网络设备需要维持MLAG系统，保持继续工作，保证网络流量正常转发，因此需要保持当前的MAC地址以及LACP系统优先级。

在图3与图4所示的应用场景中，网络设备A1重启器件。网络设备A2能够感知到网络设备A1上与网络设备A2之间的keepalive链路与peer-link链路均断开。在此情况下，若按照相关技术，网络设备A2会进入独立模式，导致流量转发受到影响。

但本申请提供的方案中，在STP基于IEEE 802.1t标准的情况下，在网络设备A1重启之前，网络设备A3与网络设备A4连接的聚合接口包含两个处于正常状态的接口。在此情况下参见表1，STP报文携带的

第一路径开销值为10000。在网络设备A1重启期间，网络设备A3与网络设备A4仅有单个正常状态的接口。在此情况下参见表1，STP报文携带的第二路径开销值为20000。网络设备A2确定第一路径开销值与第二路径开销值不相同。则网络设备A2可以确定，不仅网络设备A1上与网络设备A2连接的接口关闭，网络设备A1上与目标设备之间连接的接口也关闭。也就是不仅网络设备A1与网络设备A2之间的
5 keepalive链路和peer-link链路断开，网络设备A1与目标设备之间的链路也断开。并不是keepalive链路和peer-link链路发生故障，而是网络设备A1整体关闭或故障。在此情况下，网络设备A2需要维持MLAG系统，继续进行流量转发，不能够进入独立模式。也就是需要保持当前的MAC地址和LACP系统优先级。即网络设备A2的MAC地址保持为MLAG系统的MAC地址，网络设备A1的LACP系统优先级保持为MLAG系统的LACP系统优先级。由于网络设备A2的MAC地址和LACP系统优先级都没有发生变化，因此
10 此MLAG接口不会发生不必要的震荡，流量转发不会受到影响。

另外，上述STP报文中携带Flags（标识）字段，Flags字段的数据量为1字节。对于STP，第一个Bit，即左侧的高位Bit表示TCA（Topology Change Acknowledge，拓扑改变响应）。最后一个Bit，即右边、低位Bit表示TC（Topology Change，拓扑改变），其他字段预留。

对于RSTP（Rapid Spanning Tree Protocol，快速生成树协议）和MSTP（Multi-Service Transport Platform，多业务传送平台），Bit7为TCA，Bit6为Agreement（协定）标识，Bit5为Forwarding（发送）标识，Bit4为Learning（学习）标识，Bit3和Bit2表示接口角色，Bit1为Proposal（请求标识），Bit0为TC。

其中，Bit3和Bit2取值为00，表示接口角色未知；取值为01(低位为1高位为0)，表示接口角色为Alternate（候补）接口或Backup（备份）接口；取值为10(低位为0高位为1)，表示接口角色为根接口；取值为11，表示接口角色为指定端口。

由以上可见，本申请实施例提供的通信方法中，在MLAG系统中的多个网络设备中的任一网络设备确定多个网络设备之间已建立的keepalive链路与peer-link链路均断开的情况下，不会直接进入MLAG系统的分裂状态。而是确定接入MLAG系统的目标设备发送的STP报文中携带的路径开销值是否发生变化。若该网络设备之外的其他网络设备重启或故障，导致上述其他网络设备未正常工作。则上述其他网络设备的所有接口均会处于关闭状态，导致该其他网络设备与目标设备之间的通道断开，目标设备上
25 与其他网络设备相连的接口也会变为关闭状态。在此情况下，目标设备上处于正常状态的接口的数量会发生变化，进而目标设备发送的STP报文中携带的路径开销值会发生变化。若网络设备确定第一路径开销值和第二路径开销值不同，则可以确定上述其他网络设备的所有接口均关闭，上述其他网络设备当前未正常工作。

在此情况下，该网络设备不会由于多个网络设备之间的链路断开，直接进入MLAG系统的分裂状态。
30 而是会结合STP报文携带的第一路径开销值和第二路径开销值是否相同，进行进一步判断。在确定第一路径开销值与第二路径开销值不同，即上述其他网络设备当前未正常工作的情况下，维持MLAG系统，继续保持自身当前的MAC地址和优先级进行流量转发。从而避免MAC地址与优先级变化器件MLAG接口发生震荡，避免影响网络流量的转发。

本申请的另一个实施例中，参见图6，为本申请实施例提供的第三种应用场景示意图。在图3所示的网络设备A1上与网络设备A2相连的接口故障，因此图中的keepalive链路与peer-link链路断开。其中，keepalive链路和peer-link链路处于断开状态可以是两者同时断开，也可以是keepalive链路先断开，peer-link链路后断开。

在图中以“×号”表示链路断开。除此之外，图6所示的应用场景中的其他部分与图3相同，在此不再赘述。

需要说明的是，图3和图6示出的MLAG系统中包含的网络设备的数量以及目标设备的数量仅为一种示例，本申请实施例两者的数量对此不进行具体限定。

参见图7，为本申请实施例提供的第二种通信方法的流程示意图，与前述图5所示的实施例相比，还

包括以下步骤S503。

步骤S503：若上述多个网络设备之间已建立的keepalive链路和peer-link链路均断开，且确定第一路径开销值与第二路径开销值相同，则上述MLAG系统处于分裂状态。

具体的，网络设备之间的keepalive链路和peer-link链路处于断开状态，但目标设备第一路径开销值和第二路径开销值相同。则可以确定对于目标设备而言，目标设备上正常状态的接口的数量未发生变化。说明网络设备与目标设备之间的接口未断开。在此情况下，网络设备可以确定仅多个网络设备之间的keepalive链路和peer-link链路处于断开状态，但网络设备本身没有发生整体故障或重启。在此情况下，MLAG系统进入分裂状态，网络设备需要进入独立模式。在此情况下，需要调整网络设备的MAC地址，并降低该网络设备的LACP系统优先级。

本申请的一个实施例中，网络设备可以将MAC地址和LACP系统优先级调整回自身配置MLAG系统之前原本的MAC地址和元隔板LACP系统优先级。

在图3与图6所示的应用场景中，网络设备A1能够确定keepalive链路和peer-link链路均断开状态。

本申请提供的方案中，在STP基于IEEE 802.1t标准的情况下，若网络设备A1处于正常状态，网络设备A3与网络设备A4连接的聚合接口包含两个处于正常状态的接口。在此情况下参见表1，STP报文携带的第一路径开销值为10000。之后，keepalive链路和peer-link链路均断开状态。但网络设备A3与网络设备A4连接的聚合接口依旧包含两个处于正常状态的接口。在此情况下参见表1，STP报文携带的第二路径开销值依旧为10000。网络设备A2接收到的STP报文携带的第一路径开销值和第二路径开销值相同。则网络设备A2可以确定，仅有网络设备A1与网络设备A2之间的keepalive链路和peer-link链路发生故障。在此情况下，MLAG系统处于分裂状态，网络设备A2需要进入独立模式。否则网络设备A1和网络设备A2之间会互相影响报文转发。因此网络设备需要更新当前的MAC地址和LACP系统优先级。

由以上可见，本申请实施例提供的通信方法中，在MLAG系统中的网络设备确定keepalive链路和peer-link链路均断开的情况下，不会直接确定MLAG系统处于分裂状态。而是确定第一路径开销值与第二路径开销值是否相同。若两者相同，则说明网络设备与目标设备之间接口的状态未发生变化。因此，可以确定仅有keepalive链路和peer-link链路故障，MLAG系统中的网络设备之间的链路断开。MLAG系统需要进入分裂状态，网络设备可以正常进入独立模式，以防止多个网络设备均作为主设备影响流量转发。

另外，不同设备上接口的STP接口角色不同。在设备的接口的STP接口角色为能够主动发送STP报文的第一类型（即STP接口角色为指定接口）的情况下，该设备能够通过接口主动发送STP报文。在设备的接口的STP接口角色为不能够主动发送STP报文的第二类型（即STP接口角色为根接口）的情况下，该设备不能通过接口主动发送STP报文。因此，目标设备向网络设备发送STP报文存在以下两种情形，分别为情形一与情形二。

情形一：上述目标设备的接口的STP接口角色为第一类型。

则一种情况中，上述STP报文为：上述目标设备在确定上述多个接口中的第一接口处于关闭状态时通过第二接口发送的报文。上述第一接口与第二接口不同。

在目标设备的接口的STP接口角色为第一类型的情况下，目标设备能够主动发送STP报文。因此在目标设备确定第一接口处于关闭状态的情况下，目标设备可以通过第二接口直接发送STP报文。在此情况下，STP报文仅能够被发送至第二接口连接的网络设备。

另一种情况中，上述STP报文为：上述目标设备按照预设周期，基于hash原则，向上述MLAG系统中的网络设备发送的报文。

在目标设备的接口的STP接口角色为第一类型的情况下，目标设备可以主动发送STP报文。目标设备可以按照预设周期基于hash原则，向MLAG系统中的网络设备发送STP报文。

基于hash原则发送的一个报文可以被发送至MLAG系统中的任一网络设备中。

此方式可以适用于目标设备上与MLAG系统中的网络设备相连的接口中不存在处于关闭状态的接口的情况，包括MLAG系统中的网络设备上的接口未发生任何故障。或MLAG系统中的网络设备发生整体故障或重启之前，或MLAG系统中的多个网络设备之间的keepalive链路(peer-link)链路均断开，但网络设备与目标设备之间的链路未发生故障。

5 情形二：在上述网络设备的接口的STP接口角色为第一类型的情况下，上述MLAG系统中的任一网络设备还执行以下步骤A。

 步骤A：若上述多个网络设备之间已建立的keepalive链路(peer-link)链路均断开，向上述目标设备发送STP请求报文，以使得上述目标设备接收到上述STP请求报文后，向上述MLAG系统中的网络设备反馈STP报文。

10 由于上述网络设备的接口的STP接口角色为第一类型，目标设备的接口的STP接口角色为第二类型。因此网络设备能够通过接口主动发送STP报文，目标设备不能够主动发送STP报文。在此情况下，为了让目标设备能够向网络设备发送STP报文，网络设备需要向目标设备发送STP请求报文。

 其中，上述STP请求报文携带的Flags中的位于Bit1处的proposal（请求）标识置位，也就是取值为第一预设取值，例如可以将proposal标识置位为1或0，表示该报文为STP请求报文。

15 目标设备在接收到STP请求报文之后，确定网络设备发起了针对STP报文的请求，因此目标设备需要回应相对应的STP报文，该STP报文可以被称为STP协定报文。STP响应报文携带的Flags中位于Bit6处的Agreement标识置位，也就是取值为第二预设取值，例如第二预设取值可以Agreement标识置位为1或0，表示该报文为STP协定报文。

 与前述应用于网络设备的通信方法相对应，本申请实施例还提供了一种应用于目标设备的通信方法。

20 本申请的一个实施例中，应用于接入MLAG系统的目标设备，包括以下步骤B-步骤C。

 步骤B：基于hash原则，向上述MLAG系统中的网络设备发送第一STP报文。

 上述第一STP报文包括第一路径开销值。

 步骤C：基于hash原则，向上述MLAG系统中的网络设备发送第二STP报文，上述第二STP报文包括第二路径开销值，以使得上述MLAG系统中的任一网络设备在确定上述多个网络设备之间已建立的keepalive链路(peer-link)链路均断开，且第一路径开销值与第二路径开销值不同的情况下，维持上述MLAG系统。

 其中，上述第一路径开销值为：keepalive链路(peer-link)链路均断开之前上述网络设备接收到的STP报文中携带的路径开销值，上述第二路径开销值为：keepalive链路(peer-link)链路均断开之后上述网络设备接收到的STP报文中携带的路径开销值。

30 由以上可见，本申请实施例提供的通信方法中，在MLAG系统中的多个网络设备中的任一网络设备确定多个网络设备之间已建立的keepalive链路(peer-link)链路均断开的情况下，不会直接进入MLAG系统的分裂状态。而是确定接入MLAG系统的目标设备发送的STP报文中携带的路径开销值是否发生变化。若该网络设备之外的其他网络设备重启或故障，导致上述其他网络设备未正常工作。则上述其他网络设备的所有接口均会处于关闭状态，导致该其他网络设备与目标设备之间的通道断开，目标设备上与其他网络设备相连的接口也会变为关闭状态。在此情况下，目标设备上处于正常状态的接口的数量会发生变化，进而目标设备发送的STP报文中携带的路径开销值会发生变化。若网络设备确定第一路径开销值和第二路径开销值不同，则可以确定上述其他网络设备的所有接口均关闭，上述其他网络设备当前未正常工作。

40 在此情况下，该网络设备不会由于多个网络设备之间的链路断开，直接进入MLAG系统的分裂状态。而是会结合STP报文携带的第一路径开销值和第二路径开销值是否相同，进行进一步判断。在确定第一路径开销值与第二路径开销值不同，即上述其他网络设备当前未正常工作的情况下，维持MLAG系统，继续保持自身当前的MAC地址和优先级进行流量转发。从而避免MAC地址与优先级变化器件MLAG接

口发生震荡，避免影响网络流量的转发。

本申请的一个实施例中，在上述目标设备的上述多个接口的STP接口角色为第一类型的情况下，通过以下步骤D向网络设备发送第一STP报文和第二STP报文。

步骤D：基于hash原则，在确定上述多个接口中的第一接口处于关闭状态时，通过第二接口向上述
5 MLAG系统中的网络设备发送第一STP报文和第二STP报文。

上述第一接口与第二接口不同。

本申请的一个实施例中，在上述目标设备的上述多个接口的STP接口角色为第一类型的情况下，通过以下步骤E向网络设备发送第一STP报文和第二STP报文。

步骤E：基于hash原则，按照预设周期，向上述MLAG系统中的网络设备发送第一STP报文和第二
10 STP报文。

本申请的一个实施例中，在上述网络设备的接口的STP接口角色为第一类型的情况下，通过以下步骤F向网络设备发送第二STP报文。

步骤F：在接收到上述MLAG系统中的网络设备发送的STP请求报文的情况下，向该网络设备反馈第二STP报文。

其中，上述STP请求报文是该网络设备确定上述MLAG系统中的多个网络设备之间已建立的
15 keepalive链路与peer-link链路均断开的情况下发送的。

本申请的一个实施例中，上述STP报文包括MST BPDU字段，上述MST BPDU字段承载上述路径开销值。

与前述通信方法相对应，本申请实施例还提供了一种通信系统。

本申请实施例提供了一种通信系统，所述系统包括组成跨设备聚合组MLAG系统的多个网络设备、
20 接入所述MLAG系统的目标设备，所述目标设备包括多个接口，每一接口用于接入所述MLAG系统中的一个网络设备；

所述多个网络设备中的任一网络设备，用于实现前述应用于网络设备的通信方法中任一项的方法步骤；

所述目标设备，用于实现前述应用于目标设备的通信方法中任一项的方法步骤。

上述系统的结构可以为前述图2所示的结构，在此不再赘述。

与前述应用于网络设备的通信方法相对应，本申请实施例还提供了一种网络设备。

参见图8，为本申请实施例提供了一种网络设备的结构示意图，所述网络设备为：组成跨设备聚合组MLAG系统的多个网络设备中的任一网络设备，所述网络设备包括：

30 处理器801；

收发器804；

机器可读存储介质802，所述机器可读存储介质802存储有能够被所述处理器801执行的机器可执行指令；所述机器可执行指令促使所述处理器801执行以下步骤：

接收目标设备发送的携带路径开销值的生成树协议STP报文，其中，所述目标设备包括多个接口，
35 每一接口用于接入所述MLAG系统中的一个网络设备，所述路径开销值用于指示：所述多个接口中处于正常状态的接口数量；

若所述多个网络设备之间已建立的保活 keepalive 链路与内部控制 peer-link 链路均断开，且确定第一路径开销值与第二路径开销值不同，则维持所述 MLAG 系统，其中，所述第一路径开销值为：keepalive 链路与 peer-link 链路均断开之前接收到的 STP 报文中携带的路径开销值，所述第二路径开销值为：

40 keepalive 链路与 peer-link 链路均断开之后接收到的 STP 报文中携带的路径开销值。

如图 8 所示，网络设备还可以包括通信总线 803。处理器 801、机器可读存储介质 802 及收发器 804 之间通过通信总线 803 完成相互间的通信，通信总线 803 可以是外设部件互连标准(Peripheral Component

Interconnect, PCI) 总线或扩展工业标准结构 (Extended Industry Standard Architecture, EISA) 总线等。该通信总线 803 可以分为地址总线、数据总线、控制总线等。

收发器 804 可以为无线通信模块, 收发器 804 在处理器 801 的控制下, 与其他设备进行数据交互。

5 机器可读存储介质 802 可以包括随机存取存储器 (Random Access Memory, RAM), 也可以包括非易失性存储器 (Non-Volatile Memory, NVM), 例如至少一个磁盘存储器。另外, 机器可读存储介质 802 还可以是至少一个位于远离前述处理器的存储装置。

10 处理器 801 可以是通用处理器, 包括中央处理器 (Central Processing Unit, CPU)、网络处理器 (Network Processor, NP) 等; 还可以是数字信号处理器 (Digital Signal Processing, DSP)、专用集成电路 (Application Specific Integrated Circuit, ASIC)、现场可编程门阵列 (Field-Programmable Gate Array, FPGA) 或其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。

15 由以上可见, 本申请实施例提供的通信方法中, 在 MLAG 系统中的多个网络设备中的任一网络设备确定多个网络设备之间已建立的 keepalive 链路 peer-link 链路均断开的情况下, 不会直接进入 MLAG 系统的分裂状态。而是确定接入 MLAG 系统的目标设备发送的 STP 报文中携带的路径开销值是否发生变化。若该网络设备之外的其他网络设备重启或故障, 导致上述其他网络设备未正常工作。则上述其他网络设备的

15 网络设备的接口均会处于关闭状态, 导致该其他网络设备与目标设备之间的通道断开, 目标设备上与其他网络设备相连的接口也会变为关闭状态。在此情况下, 目标设备上处于正常状态的接口的数量会发生变化, 进而目标设备发送的 STP 报文中携带的路径开销值会发生变化。若网络设备确定第一路径开销值和第二路径开销值不同, 则可以确定上述其他网络设备的接口均关闭, 上述其他网络设备当前未正常工作。

20 在此情况下, 该网络设备不会由于多个网络设备之间的链路断开, 直接进入 MLAG 系统的分裂状态。而是会结合 STP 报文携带的第一路径开销值和第二路径开销值是否相同, 进行进一步判断。在确定第一路径开销值与第二路径开销值不同, 即上述其他网络设备当前未正常工作的情况下, 维持 MLAG 系统, 继续保持自身当前的 MAC 地址和优先级进行流量转发。从而避免 MAC 地址与优先级变化器件 MLAG 接口发生震荡, 避免影响网络流量的转发。

25 本申请的一个实施例中, 所述机器可执行指令还促使所述处理器执行以下步骤:

若所述多个网络设备之间已建立的 keepalive 链路 peer-link 链路均断开, 且确定第一路径开销值与第二路径开销值相同, 则所述 MLAG 系统处于分裂状态。

30 由以上可见, 本申请实施例提供的通信方法中, 在 MLAG 系统中的网络设备确定 keepalive 链路 peer-link 链路均断开的情况下, 不会直接确定 MLAG 系统处于分裂状态。而是确定第一路径开销值与第二路径开销值是否相同。若两者相同, 则说明网络设备与目标设备之间接口的状态未发生变化。因此, 可以确定仅有 keepalive 链路 peer-link 链路故障, MLAG 系统中的网络设备之间的链路断开。MLAG 系统需要进入分裂状态, 网络设备可以正常进入独立模式, 以防止多个网络设备均作为主设备影响流量转发。

35 本申请的一个实施例中, 在所述目标设备的所述多个接口的 STP 接口角色为第一类型的情况下, STP 报文为: 所述目标设备在确定所述多个接口中的第一接口处于关闭状态时通过第二接口发送的报文, 所述第一接口与第二接口不同。

本申请的一个实施例中, 在所述目标设备的所述多个接口的 STP 接口角色为第一类型的情况下, STP 报文为: 所述目标设备按照预设周期, 基于哈希 hash 原则, 向所述 MLAG 系统中的网络设备发送的报文。

40 本申请的一个实施例中, 在所述网络设备的接口的 STP 接口角色为第一类型的情况下, 所述机器可执行指令还促使所述处理器执行以下步骤:

若所述多个网络设备之间已建立的 keepalive 链路 peer-link 链路均断开, 向所述目标设备发送 STP 请求报文, 以使得所述目标设备接收到所述 STP 请求报文后, 向所述 MLAG 系统中的网络设备反馈 STP

报文。

本申请的一个实施例中，所述STP报文包括多生成树桥协议数据单元MST BPDU字段，所述MST BPDU字段承载所述路径开销值。

与前述应用于目标设备的通信方法相对应，本申请实施例还提供了一种目标设备。

5 参见图9，为本申请实施例提供的一种目标设备的结构示意图，所述目标设备接入跨设备聚合组MLAG系统，所述MLAG系统由多个网络设备组成，所述目标设备包括多个接口，每一接口用于接入所述MLAG系统中的一个网络设备，所述目标设备包括：

处理器901；

收发器904；

10 机器可读存储介质902，所述机器可读存储介质902存储有能够被所述处理器执行的机器可执行指令；所述机器可执行指令促使所述处理器执行以下步骤：

基于哈希hash原则，向所述MLAG系统中的网络设备发送第一生成树协议STP报文，所述第一STP报文包括第一路径开销值；

15 基于hash原则，向所述MLAG系统中的网络设备发送第二STP报文，所述第二STP报文包括第二路径开销值，以使得所述MLAG系统中的任一网络设备在确定所述多个网络设备之间已建立的保活keepalive链路与内部控制peer-link链路均断开，且第一路径开销值与第二路径开销值不同的情况下，维持所述MLAG系统；

20 其中，所述第一路径开销值为：keepalive 链路与 peer-link 链路均断开之前所述网络设备接收到的STP 报文中携带的路径开销值，所述第二路径开销值为：keepalive 链路与 peer-link 链路均断开之后所述网络设备接收到的 STP 报文中携带的路径开销值。

如图9所示，网络设备还可以包括通信总线903。处理器901、机器可读存储介质902及收发器904之间通过通信总线903完成相互间的通信，通信总线903可以是外设部件互连标准(Peripheral Component Interconnect, PCI)总线或扩展工业标准结构(Extended Industry Standard Architecture, EISA)总线等。该通信总线903可以分为地址总线、数据总线、控制总线等。

25 收发器904可以为无线通信模块，收发器904在处理器901的控制下，与其他设备进行数据交互。

机器可读存储介质902可以包括随机存取存储器(Random Access Memory, RAM)，也可以包括非易失性存储器(Non-Volatile Memory, NVM)，例如至少一个磁盘存储器。另外，机器可读存储介质902还可以是至少一个位于远离前述处理器的存储装置。

30 处理器901可以是通用处理器，包括中央处理器(Central Processing Unit, CPU)、网络处理器(Network Processor, NP)等；还可以是数字信号处理器(Digital Signal Processing, DSP)、专用集成电路(Application Specific Integrated Circuit, ASIC)、现场可编程门阵列(Field-Programmable Gate Array, FPGA)或其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。

35 由以上可见，本申请实施例提供的通信方法中，在MLAG系统中的多个网络设备中的任一网络设备确定多个网络设备之间已建立的keepalive链路与peer-link链路均断开的情况下，不会直接进入MLAG系统的分裂状态。而是确定接入MLAG系统的目标设备发送的STP报文中携带的路径开销值是否发生变化。若该网络设备之外的其他网络设备重启或故障，导致上述其他网络设备未正常工作。则上述其他网络设备的所有接口均会处于关闭状态，导致该其他网络设备与目标设备之间的通道断开，目标设备上与其他网络设备相连的接口也会变为关闭状态。在此情况下，目标设备上处于正常状态的接口的数量会发生变化，进而目标设备发送的STP报文中携带的路径开销值会发生变化。若网络设备确定第一路径
40 开销值和第二路径开销值不同，则可以确定上述其他网络设备的所有接口均关闭，上述其他网络设备当前未正常工作。

在此情况下，该网络设备不会由于多个网络设备之间的链路断开，直接进入MLAG系统的分裂状

态。而是会结合 STP 报文携带的第一路径开销值和第二路径开销值是否相同，进行进一步判断。在确定第一路径开销值与第二路径开销值不同，即上述其他网络设备当前未正常工作的情况下，维持 MLAG 系统，继续保持自身当前的 MAC 地址和优先级进行流量转发。从而避免 MAC 地址与优先级变化器件 MLAG 接口发生震荡，避免影响网络流量的转发。

5 本申请的一个实施例中，在所述目标设备的所述多个接口的STP接口角色为第一类型的情况下，所述基于哈希hash原则，向所述MLAG系统中的网络设备发送第一生成树协议STP报文，以及所述基于hash原则，向所述MLAG系统中的网络设备发送第二STP报文，包括：

 基于hash原则，在确定所述多个接口中的第一接口处于关闭状态时，通过第二接口向所述MLAG系统中的网络设备发送第一STP报文和第二STP报文，所述第一接口与第二接口不同。

10 本申请的一个实施例中，在所述目标设备的所述多个接口的STP接口角色为第一类型的情况下，所述基于哈希hash原则，向所述MLAG系统中的网络设备发送第一生成树协议STP报文，以及所述基于hash原则，向所述MLAG系统中的网络设备发送第二STP报文，包括：

 基于hash原则，按照预设周期，向所述MLAG系统中的网络设备发送第一STP报文和第二STP报文。

15 本申请的一个实施例中，在所述网络设备的接口的STP接口角色为第一类型的情况下，所述所述基于hash原则，向所述MLAG系统中的网络设备发送第二STP报文，包括：

 在接收到所述MLAG系统中的网络设备发送的STP请求报文的情况下，向该网络设备反馈第二STP报文，其中，所述STP请求报文是该网络设备确定所述MLAG系统中的多个网络设备之间已建立的keepalive链路和peer-link链路均断开的情况下发送的。

20 本申请的一个实施例中，所述 STP 报文包括多生成树桥协议数据单元 MST BPDU 字段，所述 MST BPDU 字段承载所述路径开销值。

 与前述应用于网络设备的通信方法相对应，本申请实施例还提供了一种应用于网络设备的通信装置。

 参见图 10，为本申请实施例提供的第一种通信装置的结构示意图，应用于组成 MLAG 系统的多个网络设备中的任一网络设备，所述装置包括：

25 STP报文接收模块1001，用于接收目标设备发送的携带路径开销值的生成树协议STP报文，其中，所述目标设备包括多个接口，每一接口用于接入所述MLAG系统中的一个网络设备，所述路径开销值用于指示：所述多个接口中处于正常状态的接口数量；

30 系统维持模块1002，用于若所述多个网络设备之间已建立的保活keepalive链路和内部控制peer-link链路均断开，且确定第一路径开销值与第二路径开销值不同，则维持所述MLAG系统，其中，所述第一路径开销值为：keepalive链路和peer-link链路均断开之前接收到的STP报文中携带的路径开销值，所述第二路径开销值为：keepalive链路和peer-link链路均断开之后接收到的STP报文中携带的路径开销值。

35 由以上可见，本申请实施例提供的通信方法中，在 MLAG 系统中的多个网络设备中的任一网络设备确定多个网络设备之间已建立的 keepalive 链路和 peer-link 链路均断开的情况下，不会直接进入 MLAG 系统的分裂状态。而是确定接入 MLAG 系统的目标设备发送的 STP 报文中携带的路径开销值是否发生变化。若该网络设备之外的其他网络设备重启或故障，导致上述其他网络设备未正常工作。则上述其他网络设备的所有接口均会处于关闭状态，导致该其他网络设备与目标设备之间的通道断开，目标设备上与其他网络设备相连的接口也会变为关闭状态。在此情况下，目标设备上处于正常状态的接口的数量会发生变化，进而目标设备发送的 STP 报文中携带的路径开销值会发生变化。若网络设备确定第一路径开销值和第二路径开销值不同，则可以确定上述其他网络设备的所有接口均关闭，上述其他网络设备当前未正常工作。

40 在此情况下，该网络设备不会由于多个网络设备之间的链路断开，直接进入 MLAG 系统的分裂状态。而是会结合 STP 报文携带的第一路径开销值和第二路径开销值是否相同，进行进一步判断。在确定第一路径开销值与第二路径开销值不同，即上述其他网络设备当前未正常工作的情况下，维持 MLAG

系统, 继续保持自身当前的 MAC 地址和优先级进行流量转发。从而避免 MAC 地址与优先级变化器件 MLAG 接口发生震荡, 避免影响网络流量的转发。

本申请的一个实施例中, 所述装置还包括:

5 系统分裂模块, 用于若所述多个网络设备之间已建立的keepalive链路和peer-link链路均断开, 且确定第一路径开销值与第二路径开销值相同, 则所述MLAG系统处于分裂状态。

10 由以上可见, 本申请实施例提供的通信方法中, 在MLAG系统中的网络设备确定keepalive链路和peer-link链路均断开的情况下, 不会直接确定MLAG系统处于分裂状态。而是确定第一路径开销值与第二路径开销值是否相同。若两者相同, 则说明网络设备与目标设备之间接口的状态未发生变化。因此, 可以确定仅有keepalive链路和peer-link链路故障, MLAG系统中的网络设备之间的链路断开。MLAG系统需要进入分裂状态, 网络设备可以正常进入独立模式, 以防止多个网络设备均作为主设备影响流量转发。

本申请的一个实施例中, 在所述目标设备的所述多个接口的STP接口角色为第一类型的情况下, STP报文为: 所述目标设备在确定所述多个接口中的第一接口处于关闭状态时通过第二接口发送的报文, 所述第一接口与第二接口不同。

15 本申请的一个实施例中, 在所述目标设备的所述多个接口的STP接口角色为第一类型的情况下, STP报文为: 所述目标设备按照预设周期, 基于哈希hash原则, 向所述MLAG系统中的网络设备发送的报文。

本申请的一个实施例中, 在所述网络设备的接口的STP接口角色为第一类型的情况下, 所述装置还包括:

20 请求报文发送模块, 用于若所述多个网络设备之间已建立的keepalive链路和peer-link链路均断开, 向所述目标设备发送STP请求报文, 以使得所述目标设备接收到所述STP请求报文后, 向所述MLAG系统中的网络设备反馈STP报文。

本申请的一个实施例中, 所述STP报文包括多生成树桥协议数据单元MST BPDU字段, 所述MST BPDU字段承载所述路径开销值。

25 与前述应用于目标设备的通信方法相对应, 本申请实施例提供了一种应用于接入MLAG系统的目标设备的通信装置。参见图11, 为本申请实施例提供的第二种通信装置的结构示意图, 上述MLAG系统由多个网络设备组成, 上述目标设备包括多个接口, 每一接口用于接入上述MLAG系统中的一个网络设备, 所述装置包括:

第一STP报文发送模块1101, 用于基于哈希hash原则, 向所述MLAG系统中的网络设备发送第一生成树协议STP报文, 所述第一STP报文包括第一路径开销值;

30 第二STP报文发送模块1102, 用于基于hash原则, 向所述MLAG系统中的网络设备发送第二STP报文, 所述第二STP报文包括第二路径开销值, 以使得所述MLAG系统中的任一网络设备在确定所述多个网络设备之间已建立的保活keepalive链路和内部控制peer-link链路均断开, 且第一路径开销值与第二路径开销值不同的情况下, 维持所述MLAG系统;

35 其中, 所述第一路径开销值为: keepalive链路和peer-link链路均断开之前所述网络设备接收到的STP报文中携带的路径开销值, 所述第二路径开销值为: keepalive链路和peer-link链路均断开之后所述网络设备接收到的STP报文中携带的路径开销值。

40 由以上可见, 本申请实施例提供的通信方法中, 在 MLAG 系统中的多个网络设备中的任一网络设备确定多个网络设备之间已建立的 keepalive 链路和 peer-link 链路均断开的情况下, 不会直接进入 MLAG 系统的分裂状态。而是确定接入 MLAG 系统的目标设备发送的 STP 报文中携带的路径开销值是否发生变化。若该网络设备之外的其他网络设备重启或故障, 导致上述其他网络设备未正常工作。则上述其他网络设备的所有接口均会处于关闭状态, 导致该其他网络设备与目标设备之间的通道断开, 目标设备上与其他网络设备相连的接口也会变为关闭状态。在此情况下, 目标设备上处于正常状态的接口的数量会

发生变化，进而目标设备发送的 STP 报文中携带的路径开销值会发生变化。若网络设备确定第一路径开销值和第二路径开销值不同，则可以确定上述其他网络设备的所有接口均关闭，上述其他网络设备当前未正常工作。

在此情况下，该网络设备不会由于多个网络设备之间的链路断开，直接进入MLAG系统的分裂状态。而是会结合STP报文携带的第一路径开销值和第二路径开销值是否相同，进行进一步判断。在确定第一路径开销值与第二路径开销值不同，即上述其他网络设备当前未正常工作的情况下，维持MLAG系统，继续保持自身当前的MAC地址和优先级进行流量转发。从而避免MAC地址与优先级变化器件MLAG接口发生震荡，避免影响网络流量的转发。

本申请的一个实施例中，在所述目标设备的所述多个接口的STP接口角色为第一类型的情况下，所述基于哈希hash原则，向所述MLAG系统中的网络设备发送第一生成树协议STP报文，以及所述基于hash原则，向所述MLAG系统中的网络设备发送第二STP报文，包括：

基于 hash 原则，在确定所述多个接口中的第一接口处于关闭状态时，通过第二接口向所述 MLAG 系统中的网络设备发送第一 STP 报文和第二 STP 报文，所述第一接口与第二接口不同。

本申请的一个实施例中，在所述目标设备的所述多个接口的STP接口角色为第一类型的情况下，所述基于哈希hash原则，向所述MLAG系统中的网络设备发送第一生成树协议STP报文，以及所述基于hash原则，向所述MLAG系统中的网络设备发送第二STP报文，包括：

基于 hash 原则，按照预设周期，向所述 MLAG 系统中的网络设备发送第一 STP 报文和第二 STP 报文。

本申请的一个实施例中，在所述网络设备的接口的STP接口角色为第一类型的情况下，所述基于hash原则，向所述MLAG系统中的网络设备发送第二STP报文，包括：

在接收到所述 MLAG 系统中的网络设备发送的 STP 请求报文的情况下，向该网络设备反馈第二 STP 报文，其中，所述 STP 请求报文是该网络设备确定所述 MLAG 系统中的多个网络设备之间已建立的 keepalive 链路和 peer-link 链路均断开的情况下发送的。

本申请的一个实施例中，所述 STP 报文包括多生成树桥协议数据单元 MST BPDU 字段，所述 MST BPDU 字段承载所述路径开销值。

基于同一种发明构思，根据上述本申请实施例提供的通信方法，一种机器可读存储介质，存储有机器可执行指令，在被处理器调用和执行时，所述机器可执行指令促使所述处理器：实现任一应用于网络设备或目标设备的通信方法的步骤。

在本申请提供的又一实施例中，还提供了一种包含指令的计算机程序产品，当其在计算机上运行时，使得计算机执行上述实施例中任一应用于网络设备或目标设备的通信方法的步骤。

在上述实施例中，可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时，可以全部或部分地以计算机程序产品的形式实现。所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机程序指令时，全部或部分地产生按照本申请实施例所述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。所述计算机指令可以存储在计算机可读存储介质中，或者从一个计算机可读存储介质向另一个计算机可读存储介质传输，例如，所述计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线（例如同轴电缆、光纤、数字用户线（DSL））或无线（例如红外、无线、微波等）方式向另一个网站站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机能够存取的任何可用介质或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。所述可用介质可以是磁性介质，（例如，软盘、硬盘、磁带）、光介质（例如，DVD）、或者半导体介质（例如固态硬盘 Solid State Disk（SSD））等。

需要说明的是，在本文中，诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另

- 一个实体或操作区分开来,而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。
- 5 本说明书中的各个实施例均采用相关的方式描述,各个实施例之间相同相似的部分互相参见即可,每个实施例重点说明的都是与其他实施例的不同之处。尤其,对于应用于目标设备的通信方法、系统、网络设备、目标设备、装置、计算机可读存储介质和计算机程序产品实施例而言,由于其基本相似于方法实施例,所以描述的比较简单,相关之处参见方法实施例的部分说明即可。
- 10 以上所述仅为本申请的较佳实施例,并非用于限定本申请的保护范围。凡在本申请的精神和原则之内所作的任何修改、等同替换、改进等,均包含在本申请的保护范围内。

权 利 要 求

1、一种通信方法，其特征在于，应用于组成跨设备聚合组MLAG系统的多个网络设备中的任一网络设备，所述方法包括：

5 接收目标设备发送的携带路径开销值的生成树协议STP报文，其中，所述目标设备包括多个接口，每一接口用于接入所述MLAG系统中的一个网络设备，所述路径开销值用于指示：所述多个接口中处于正常状态的接口数量；

若所述多个网络设备之间已建立的保活keepalive链路与内部控制peer-link链路均断开，且确定第一路径开销值与第二路径开销值不同，则维持所述MLAG系统，其中，所述第一路径开销值为：keepalive链路
10 与peer-link链路均断开之前接收到的STP报文中携带的路径开销值，所述第二路径开销值为：keepalive链路与peer-link链路均断开之后接收到的STP报文中携带的路径开销值。

2、根据权利要求1所述的方法，其特征在于，所述方法还包括：

若所述多个网络设备之间已建立的keepalive链路与peer-link链路均断开，且确定第一路径开销值与第二路径开销值相同，则所述MLAG系统处于分裂状态。

3、根据权利要求1或2所述的方法，其特征在于，在所述目标设备的所述多个接口的STP接口角色
15 为第一类型的情况下，STP报文为：所述目标设备在确定所述多个接口中的第一接口处于关闭状态时通过第二接口发送的报文，所述第一接口与第二接口不同。

4、根据权利要求1或2所述的方法，其特征在于，在所述目标设备的所述多个接口的STP接口角色为第一类型的情况下，STP报文为：所述目标设备按照预设周期，基于哈希hash原则，向所述MLAG系统中的网络设备发送的报文。

5、根据权利要求1或2所述的方法，其特征在于，在所述网络设备的接口的STP接口角色为第一类型的情况下，所述方法还包括：

若所述多个网络设备之间已建立的keepalive链路与peer-link链路均断开，向所述目标设备发送STP请求报文，以使得所述目标设备接收到所述STP请求报文后，向所述MLAG系统中的网络设备反馈STP
20 报文。

6、根据权利要求1或2所述的方法，其特征在于，所述STP报文包括多生成树桥协议数据单元MST BPD
25 U字段，所述MST BPD U字段承载所述路径开销值。

7、一种通信方法，其特征在于，应用于接入跨设备聚合组MLAG系统的目标设备，所述MLAG系统由多个网络设备组成，所述目标设备包括多个接口，每一接口用于接入所述MLAG系统中的一个网络设备，所述方法包括：

30 基于哈希hash原则，向所述MLAG系统中的网络设备发送第一生成树协议STP报文，所述第一STP报文包括第一路径开销值；

基于hash原则，向所述MLAG系统中的网络设备发送第二STP报文，所述第二STP报文包括第二路径开销值，以使得所述MLAG系统中的任一网络设备在确定所述多个网络设备之间已建立的保活keepalive链路
35 与内部控制peer-link链路均断开，且第一路径开销值与第二路径开销值不同的情况下，维持所述MLAG系统；

其中，所述第一路径开销值为：keepalive链路与peer-link链路均断开之前所述网络设备接收到的STP报文中携带的路径开销值，所述第二路径开销值为：keepalive链路与peer-link链路均断开之后所述网络设备接收到的STP报文中携带的路径开销值。

8、根据权利要求7所述的方法，其特征在于，在所述目标设备的所述多个接口的STP接口角色为第一类型的情况下，所述基于哈希hash原则，向所述MLAG系统中的网络设备发送第一生成树协议STP
40 报文，以及所述基于hash原则，向所述MLAG系统中的网络设备发送第二STP报文，包括：

基于hash原则，在确定所述多个接口中的第一接口处于关闭状态时，通过第二接口向所述MLAG系

统中的网络设备发送第一STP报文和第二STP报文，所述第一接口与第二接口不同。

9、根据权利要求7所述的方法，其特征在于，在所述目标设备的所述多个接口的STP接口角色为第一类型的情况下，所述基于哈希hash原则，向所述MLAG系统中的网络设备发送第一生成树协议STP报文，以及所述基于hash原则，向所述MLAG系统中的网络设备发送第二STP报文，包括：

5 基于hash原则，按照预设周期，向所述MLAG系统中的网络设备发送第一STP报文和第二STP报文。

10、根据权利要求7所述的方法，其特征在于，在所述网络设备的接口的STP接口角色为第一类型的情况下，所述基于hash原则，向所述MLAG系统中的网络设备发送第二STP报文，包括：

在接收到所述MLAG系统中的网络设备发送的STP请求报文的情况下，向该网络设备反馈第二STP报文，其中，所述STP请求报文是该网络设备确定所述MLAG系统中的多个网络设备之间已建立的
10 keepalive链路与peer-link链路均断开的情况下发送的。

11、根据权利要求7-10中任一项所述的方法，其特征在于，所述STP报文包括多生成树桥协议数据单元MST BPDU字段，所述MST BPDU字段承载所述路径开销值。

12、一种通信系统，其特征在于，所述系统包括组成跨设备聚合组MLAG系统的多个网络设备、接入所述MLAG系统的目标设备，所述目标设备包括多个接口，每一接口用于接入所述MLAG系统中的一个
15 网络设备；

所述多个网络设备中的任一网络设备，用于实现前述权利要求1-6中任一项所述的方法步骤；

所述目标设备，用于实现前述权利要求7-11中任一项所述的方法步骤。

13、一种网络设备，其特征在于，所述网络设备为：组成跨设备聚合组MLAG系统的多个网络设备中的任一网络设备，所述网络设备包括：

20 处理器；

收发器；

机器可读存储介质，所述机器可读存储介质存储有能够被所述处理器执行的机器可执行指令；所述机器可执行指令促使所述处理器执行以下步骤：

25 接收目标设备发送的携带路径开销值的生成树协议STP报文，其中，所述目标设备包括多个接口，每一接口用于接入所述MLAG系统中的一个网络设备，所述路径开销值用于指示：所述多个接口中处于正常状态的接口数量；

若所述多个网络设备之间已建立的保活 keepalive 链路与控制 peer-link 链路均断开，且确定第一路径开销值与第二路径开销值不同，则维持所述 MLAG 系统，其中，所述第一路径开销值为：keepalive 链路与控制 peer-link 链路均断开之前接收到的 STP 报文中携带的路径开销值，所述第二路径开销值为：
30 keepalive 链路与控制 peer-link 链路均断开之后接收到的 STP 报文中携带的路径开销值。

14、根据权利要求13所述的网络设备，其特征在于，所述机器可执行指令还促使所述处理器执行以下步骤：

若所述多个网络设备之间已建立的keepalive链路与控制peer-link链路均断开，且确定第一路径开销值与第二路径开销值相同，则所述MLAG系统处于分裂状态。

35 15、根据权利要求13或14所述的网络设备，其特征在于，在所述目标设备的所述多个接口的STP接口角色为第一类型的情况下，STP报文为：所述目标设备在确定所述多个接口中的第一接口处于关闭状态时通过第二接口发送的报文，所述第一接口与第二接口不同。

16、根据权利要求13或14所述的网络设备，其特征在于，在所述目标设备的所述多个接口的STP接口角色为第一类型的情况下，STP报文为：所述目标设备按照预设周期，基于哈希hash原则，向所述MLAG
40 系统中的网络设备发送的报文。

17、根据权利要求13或14所述的网络设备，其特征在于，在所述网络设备的接口的STP接口角色为第一类型的情况下，所述机器可执行指令还促使所述处理器执行以下步骤：

若所述多个网络设备之间已建立的keepalive链路与peer-link链路均断开，向所述目标设备发送STP请求报文，以使得所述目标设备接收到所述STP请求报文后，向所述MLAG系统中的网络设备反馈STP报文。

18、根据权利要求13或14所述的网络设备，其特征在于，所述STP报文包括多生成树桥协议数据单元MST BPDU字段，所述MST BPDU字段承载所述路径开销值。

19、一种目标设备，其特征在于，所述目标设备接入跨设备聚合组MLAG系统，所述MLAG系统由多个网络设备组成，所述目标设备包括多个接口，每一接口用于接入所述MLAG系统中的一个网络设备，所述目标设备包括：

处理器；

收发器；

机器可读存储介质，所述机器可读存储介质存储有能够被所述处理器执行的机器可执行指令；所述机器可执行指令促使所述处理器执行以下步骤：

基于哈希hash原则，向所述MLAG系统中的网络设备发送第一生成树协议STP报文，所述第一STP报文包括第一路径开销值；

基于hash原则，向所述MLAG系统中的网络设备发送第二STP报文，所述第二STP报文包括第二路径开销值，以使得所述MLAG系统中的任一网络设备在确定所述多个网络设备之间已建立的保活keepalive链路与内部控制peer-link链路均断开，且第一路径开销值与第二路径开销值不同的情况下，维持所述MLAG系统；

其中，所述第一路径开销值为：keepalive 链路与 peer-link 链路均断开之前所述网络设备接收到的STP 报文中携带的路径开销值，所述第二路径开销值为：keepalive 链路与 peer-link 链路均断开之后所述网络设备接收到的 STP 报文中携带的路径开销值。

20、根据权利要求19所述的目标设备，其特征在于，在所述目标设备的所述多个接口的STP接口角色为第一类型的情况下，所述基于哈希hash原则，向所述MLAG系统中的网络设备发送第一生成树协议STP报文，以及所述基于hash原则，向所述MLAG系统中的网络设备发送第二STP报文，包括：

基于hash原则，在确定所述多个接口中的第一接口处于关闭状态时，通过第二接口向所述MLAG系统中的网络设备发送第一STP报文和第二STP报文，所述第一接口与第二接口不同。

21、根据权利要求19所述的目标设备，其特征在于，在所述目标设备的所述多个接口的STP接口角色为第一类型的情况下，所述基于哈希hash原则，向所述MLAG系统中的网络设备发送第一生成树协议STP报文，以及所述基于hash原则，向所述MLAG系统中的网络设备发送第二STP报文，包括：

基于hash原则，按照预设周期，向所述MLAG系统中的网络设备发送第一STP报文和第二STP报文。

22、根据权利要求19所述的目标设备，其特征在于，在所述网络设备的接口的STP接口角色为第一类型的情况下，所述所述基于hash原则，向所述MLAG系统中的网络设备发送第二STP报文，包括：

在接收到所述MLAG系统中的网络设备发送的STP请求报文的情况下，向该网络设备反馈第二STP报文，其中，所述STP请求报文是该网络设备确定所述MLAG系统中的多个网络设备之间已建立的keepalive链路与peer-link链路均断开的情况下发送的。

23、根据权利要求19-22中任一项所述的目标设备，其特征在于，所述STP报文包括多生成树桥协议数据单元MST BPDU字段，所述MST BPDU字段承载所述路径开销值。

24、一种通信装置，其特征在于，应用于组成跨设备聚合组MLAG系统的多个网络设备中的任一网络设备，所述装置包括：

STP报文接收模块，用于接收目标设备发送的携带路径开销值的生成树协议STP报文，其中，所述目标设备包括多个接口，每一接口用于接入所述MLAG系统中的一个网络设备，所述路径开销值用于指示：所述多个接口中处于正常状态的接口数量；

系统维持模块，用于若所述多个网络设备之间已建立的保活keepalive链路与内部控制peer-link链路均断开，且确定第一路径开销值与第二路径开销值不同，则维持所述MLAG系统，其中，所述第一路径开销值为：keepalive链路与peer-link链路均断开之前接收到的STP报文中携带的路径开销值，所述第二路径开销值为：keepalive链路与peer-link链路均断开之后接收到的STP报文中携带的路径开销值。

5 25、一种通信装置，其特征在于，应用于接入跨设备聚合组MLAG系统的目标设备，所述MLAG系统由多个网络设备组成，所述目标设备包括多个接口，每一接口用于接入所述MLAG系统中的一个网络设备，所述装置包括：

 第一STP报文发送模块，用于基于哈希hash原则，向所述MLAG系统中的网络设备发送第一生成树协议STP报文，所述第一STP报文包括第一路径开销值；

10 第二STP报文发送模块，用于基于hash原则，向所述MLAG系统中的网络设备发送携带路径开销值的第二STP报文，所述第二STP报文包括第二路径开销值，以使得所述MLAG系统中的任一网络设备在确定所述多个网络设备之间已建立的保活keepalive链路与内部控制peer-link链路均断开，且第一路径开销值与第二路径开销值不同的情况下，维持所述MLAG系统；

15 其中，所述第一路径开销值为：keepalive链路与peer-link链路均断开之前所述网络设备接收到的STP报文中携带的路径开销值，所述第二路径开销值为：keepalive链路与peer-link链路均断开之后所述网络设备接收到的STP报文中携带的路径开销值。

 26、一种机器可读存储介质，其特征在于，存储有机器可执行指令，在被处理器调用和执行时，所述机器可执行指令促使所述处理器：实现权利要求 1-6 或 7-11 任一所述的方法。

20 27、一种计算机程序产品，其特征在于，所述计算机程序产品促使处理器：实现权利要求 1-6 或 7-11 任一所述的方法。