

路由信息维护方法、装置、网络设备和存储介质

技术领域

本申请涉及通信技术领域，尤其涉及一种路由信息维护方法、装置、网络设备和存储介质。

5 背景技术

自适应路由技术是根据网络拓扑和流量负载的变化，进行动态路由决策的技术。基于自适应路由技术，网络设备可以主动感知链路的拥塞状态，进而选择路径最短且不拥塞的路径转发报文。

10 网络设备可以基于路由信息和链路的拥塞状态，创建并维护最优路径（best path）表，每个 best path 表用于存放一个目的互联网协议（Internet Protocol，IP）地址对应的最优路径。该最优路径可以是与目的 IP 地址之间的最短路径，若该最短路径发生拥塞，则该最优路径将被修改为与目的 IP 地址之间的非最短路径。

在 best path 表中的最优路径被修改为非最短路径后，网络设备使用该非最短路径转发包括该目的 IP 地址的报文，最短路径将被停用，导致网络的带宽利用率降低。

发明内容

15 本申请实施例提供一种路由信息维护方法、装置、网络设备和存储介质，用于解决 best path 表中的最短路径被修改后，网络带宽利用率降低的问题，具体包括以下方案：

第一方面，本申请实施例提供一种路由信息维护方法，应用于第一网络设备，所述第一网络设备包括多路径组转发表，所述多路径组转发表对应的第一多路径组包括最短路径和非最短路径，每条路径均用于指示第一网络设备与第二网络设备之间的路径；所述最短路径的转发表项包括所述第一网络设备的第一出口；所述方法包括：

20 当所述最短路径承载的流量使得所述第一出口拥塞时，获取所述多路径组转发表对应的预设路由策略；

基于所述预设路由策略，调整所述第一多路径组中每条路径的转发权重值，每条路径的转发权重值用于表示从所述第一多路径组中选择该路径进行报文转发时，该路径承载的流量比例。

25 在一种可能的实现方式中，所述多路径组转发表包括软件转发表和硬件转发表；所述方法还包括：计算与所述第二网络设备之间的最短路径和非最短路径；

分别针对每条最短路径和每条非最短路径生成一个软件转发表项，得到所述软件转发表，其中，每个软件转发表项包括一条路径的转发权重值，非最短路径的初始转发权重值为 0；

30 针对转发权重值非 0 的每个软件转发表项，将该软件转发表项中除转发权重值之外的信息复制 N 份，得到 N 个硬件转发表项，将所述 N 个硬件转发表项下发至硬件资源，得到所述硬件转发表；

其中， $N = \frac{\text{该软件转发表项的转发权重值}}{\text{所述硬件转发表的预设表项数}} \times \text{所述软件转发表项包括的各转发权重值的和值}$ 。

在一种可能的实现方式中，所述多路径组转发表包括路由类型标记；所述获取所述多路径组转发表对应的预设路由策略，包括：

35 获取所述路由类型标记对应的预设路由策略；
所述预设路由策略至少包括以下策略中的一种：将拥塞出口的转发负载调整为 0、减少拥塞出口的转发负载且不增加非最短路径的转发负载、减少拥塞出口的转发负载并增加非最短路径的转发负

载。

在一种可能的实现方式中，所述第一网络设备中存储有第一接口状态表和所述多路径组转发表对应的第二接口状态表，所述第一接口状态表包括所述第一网络设备的每个出接口的接口状态，所述第二接口状态表包括所述第一多路径组中的每条路径承载的流量对出接口的带宽占用量；

5 在所述当所述最短路径承载的流量导致所述第一出接口拥塞时，获取所述多路径组转发表对应的预设路由策略之前，所述方法还包括：

若所述第一接口状态表中的第一接口的接口状态变更为拥塞状态，则从所述第二接口状态表中获取所述最短路径承载的流量对所述第一出接口的第一带宽占用量；

10 若所述第一带宽占用量满足预设拥塞条件，则确定所述最短路径承载的流量使得所述第一出接口拥塞。

在一种可能的实现方式中，所述第一接口状态表还包括每个出接口的带宽占用量和缓冲区占用量；所述方法还包括：

每隔预设时长，针对所述第一接口状态表中的每个出接口，更新该出接口的带宽占用量和缓冲区占用量；

15 若该出接口更新后的带宽占用量达到预设带宽阈值，或该出接口更新后的缓冲区占用量达到预设缓冲区阈值，则将该出接口的接口状态设置为拥塞状态。

在一种可能的实现方式中，所述基于所述预设路由策略，调整所述第一多路径组中每条路径的转发权重值，包括：

20 基于所述预设路由策略以及所述第二接口状态表包括的各出接口的带宽占用量，调整所述第一多路径组中每条路径的转发权重值。

在一种可能的实现方式中，所述第一网络设备与第三网络设备之间的多条路径组成第二多路径组，在所述第一多路径组与所述第二多路径组相同的情况下，所述第一多路径组和所述第二多路径组共用所述多路径组转发表。

25 第二方面，本申请实施例提供一种路由信息维护装置，应用于第一网络设备，所述第一网络设备包括多路径组转发表，所述多路径组转发表对应的第一多路径组包括最短路径和非最短路径，每条路径均用于指示第一网络设备与第二网络设备之间的路径；所述最短路径的转发表项包括所述第一网络设备的所述第一出接口；所述装置包括：

获取模块，用于当所述最短路径承载的流量使得所述第一出接口拥塞时，获取所述多路径组转发表对应的预设路由策略；

30 调整模块，用于基于所述预设路由策略，调整所述第一多路径组中每条路径的转发权重值，每条路径的转发权重值用于表示从所述第一多路径组中选择该路径进行报文转发时，该路径承载的流量比例。

在一种可能的实现方式中，所述多路径组转发表包括软件转发表和硬件转发表；所述生成模块，用于：

计算与所述第二网络设备之间的最短路径和非最短路径；

35 分别针对每条最短路径和每条非最短路径生成一个软件转发表项，得到所述软件转发表，其中，每个软件转发表项包括一条路径的转发权重值，非最短路径的初始转发权重值为 0；

针对转发权重值非 0 的每个软件转发表项，将该软件转发表项中除转发权重值之外的信息复制 N 份，得到 N 个硬件转发表项，将所述 N 个硬件转发表项下发至硬件资源，得到所述硬件转发表；

其中， $N = \text{该软件转发表项的转发权重值} \times \text{所述硬件转发表的预设表项数} / \text{所述软件转发表项包括的各转发权重值的和值}$ 。

在一种可能的实现方式中，所述多路径组转发表包括路由类型标记；所述获取模块，具体用于：

5 获取所述路由类型标记对应的预设路由策略；所述预设路由策略至少包括以下策略中的一种：将拥塞出接口的转发负载调整为 0、减少拥塞出接口的转发负载且不增加非最短路径的转发负载、减少拥塞出接口的转发负载并增加非最短路径的转发负载。

10 在一种可能的实现方式中，所述第一网络设备中存储有第一接口状态表和所述多路径组转发表对应的第二接口状态表，所述第一接口状态表包括所述第一网络设备的每个出接口的接口状态，所述第二接口状态表包括所述第一多路径组中的每条路径承载的流量对出接口的带宽占用量；所述装置还包括确定模块：

所述获取模块，还用于若所述第一接口状态表中的第一接口的接口状态变更为拥塞状态，则从所述第二接口状态表中获取所述最短路径承载的流量对所述第一出接口的第一带宽占用量；

所述确定模块，用于若所述第一带宽占用量满足预设拥塞条件，则确定所述最短路径承载的流量使得所述第一出接口拥塞。

15 在一种可能的实现方式中，所述第一接口状态表还包括每个出接口的带宽占用量和缓冲区占用量；所述装置还包括更新模块和设置模块：

所述更新模块，用于每隔预设时长，针对所述第一接口状态表中的每个出接口，更新该出接口的带宽占用量和缓冲区占用量；

20 所述设置模块，用于若该出接口更新后的带宽占用量达到预设带宽阈值，或该出接口更新后的缓冲区占用量达到预设缓冲区阈值，则将该出接口的接口状态设置为拥塞状态。

在一种可能的实现方式中，所述调整模块，具体用于：

基于所述预设路由策略以及所述第二接口状态表包括的各出接口的带宽占用量，调整所述第一多路径组中每条路径的转发权重值。

25 在一种可能的实现方式中，所述第一网络设备与第三网络设备之间的多条路径组成第二多路径组，在所述第一多路径组与所述第二多路径组相同的情况下，所述第一多路径组和所述第二多路径组共用所述多路径组转发表。

30 第三方面，本申请实施例提供一种第一网络设备，所述第一网络设备包括多路径组转发表，所述多路径组转发表对应的第一多路径组包括最短路径和非最短路径，每条路径均用于指示第一网络设备与第二网络设备之间的路径；所述最短路径的转发表项包括所述第一网络设备的第一出接口；所述第一网络设备包括：

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：

35 当所述最短路径承载的流量使得所述第一出接口拥塞时，获取所述多路径组转发表对应的预设路由策略；

基于所述预设路由策略，调整所述第一多路径组中每条路径的转发权重值，每条路径的转发权重值用于表示从所述第一多路径组中选择该路径进行报文转发时，该路径承载的流量比例。

在一种可能的实现方式中，所述多路径组转发表包括软件转发表和硬件转发表；所述机器可执行指令还促使所述处理器执行以下步骤：

计算与所述第二网络设备之间的最短路径和非最短路径；

5 分别针对每条最短路径和每条非最短路径生成一个软件转发表项，得到所述软件转发表，其中，每个软件转发表项包括一条路径的转发权重值，非最短路径的初始转发权重值为 0；

针对转发权重值非 0 的每个软件转发表项，将该软件转发表项中除转发权重值之外的信息复制 N 份，得到 N 个硬件转发表项，将所述 N 个硬件转发表项下发至硬件资源，得到所述硬件转发表；

其中， $N = \text{该软件转发表项的转发权重值} \times \text{所述硬件转发表的预设表项数} / \text{所述软件转发表项包括的各转发权重值的和值}$ 。

10 在一种可能的实现方式中，所述多路径组转发表包括路由类型标记；所述机器可执行指令还促使所述处理器执行以下步骤：

获取所述路由类型标记对应的预设路由策略；所述预设路由策略至少包括以下策略中的一种：将拥塞出接口的转发负载调整为 0、减少拥塞出接口的转发负载且不增加非最短路径的转发负载、减少拥塞出接口的转发负载并增加非最短路径的转发负载。

15 在一种可能的实现方式中，所述第一网络设备中存储有第一接口状态表和所述多路径组转发表对应的第二接口状态表，所述第一接口状态表包括所述第一网络设备的每个出接口的接口状态，所述第二接口状态表包括所述第一多路径组中的每条路径承载的流量对出接口的带宽占用量；

机器可执行指令还促使所述处理器执行以下步骤：

20 若所述第一接口状态表中的第一接口的接口状态变更为拥塞状态，则从所述第二接口状态表中获取所述最短路径承载的流量对所述第一出接口的第一带宽占用量；

若所述第一带宽占用量满足预设拥塞条件，则确定所述最短路径承载的流量使得所述第一出接口拥塞。

在一种可能的实现方式中，所述第一接口状态表还包括每个出接口的带宽占用量和缓冲区占用量；所述机器可执行指令还促使所述处理器执行以下步骤：

25 每隔预设时长，针对所述第一接口状态表中的每个出接口，更新该出接口的带宽占用量和缓冲区占用量；

若该出接口更新后的带宽占用量达到预设带宽阈值，或该出接口更新后的缓冲区占用量达到预设缓冲区阈值，则将该出接口的接口状态设置为拥塞状态。

在一种可能的实现方式中，所述机器可执行指令还促使所述处理器执行以下步骤：

30 基于所述预设路由策略以及所述第二接口状态表包括的各出接口的带宽占用量，调整所述第一多路径组中每条路径的转发权重值。

在一种可能的实现方式中，所述第一网络设备与第三网络设备之间的多条路径组成第二多路径组，在所述第一多路径组与所述第二多路径组相同的情况下，所述第一多路径组和所述第二多路径组共用所述多路径组转发表。

35 第四方面，本申请实施例提供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：实现第一方面所述的方法。

第五方面，本申请实施例提供一种计算机程序产品，所述计算机程序产品促使所述处理器：实现第一方面所述的方法。

采用上述技术方案，当最短路径承载的流量使得第一出接口拥塞时，获取多路径组转发表对应的预设路由策略，然后第一网络设备基于多路径组转发表对应的预设路由策略，调整第一多路径组中每条路径的转发权重值。可见，本申请实施例中，当最短路径承载的流量使得第一出接口拥塞时，并未直接停用第一出接口所在的最短路径，而是对第一多路径组中的每条路径的转发权重值进行调整。由于每条路径的转发权重值用于表示第一多路径组中选择该路径进行报文时，该路径承载的流量比例，所以在调整每条路径的转发权重值之后，每条路径的转发负载会发生变化，可以减少最短路径的负载，从而可以缓解第一出接口的拥塞，充分利用各路径的资源，提高网络的整体带宽利用率。

附图说明

此处所说明的附图用来提供对本申请的进一步理解，构成本申请的一部分，本申请的示意性实施例及其说明用于解释本申请，并不构成对本申请的不当限定。

- 图 1 为本申请实施例提供的一种组网系统的结构示意图；
- 图 2 为本申请实施例提供的一种路由信息维护方法流程图；
- 图 3a 为本申请实施例提供的第一种软件转发表的示例性示意图；
- 图 3b 为本申请实施例提供的第一种硬件转发表的示例性示意图；
- 图 4a 为本申请实施例提供的第二种软件转发表的示例性示意图；
- 图 4b 为本申请实施例提供的第二种硬件转发表的示例性示意图；
- 图 5a 为本申请实施例提供的第三种软件转发表的示例性示意图；
- 图 5b 为本申请实施例提供的第三种硬件转发表的示例性示意图；
- 图 6 为本申请实施例提供的另一种组网系统的结构示意图；
- 图 7a 为本申请实施例提供的第四种软件转发表的示例性示意图；
- 图 7b 为本申请实施例提供的第四种硬件转发表的示例性示意图；
- 图 8 为本申请实施例提供的又一种组网结构的结构示意图；
- 图 9a 为本申请实施例提供的第五种软件转发表的示例性示意图；
- 图 9b 为本申请实施例提供的第五种硬件转发表的示例性示意图；
- 图 10 为本申请实施例提供的一种路由信息维护装置的结构示意图；
- 图 11 为本申请实施例提供的一种网络设备的结构示意图。

具体实施方式

为使本申请的目的、技术方案、及优点更加清楚明白，以下参照附图并举实施例，对本申请进一步详细说明。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。基于本申请中的实施例，本领域普通技术人员所获得的所有其他实施例，都属于本申请保护的范围。

如图 1 所示，图 1 示例性示出了本申请实施例所应用的组网系统，该组网系统包括：源端设备、目的端设备、入口（Ingress）设备、中转（Transit）设备 1、中转设备 2 和中转设备 3。其中，图 1 中 S 表示源端设备，D1 和 D2 表示目的端设备，入口设备接收到源端设备发送的报文后，可通过中转设备将报文转发至目的端设备。源端设备、目的端设备可以是服务器、虚拟机等装置。

入口设备和中转设备可以基于最短路径优先算法，计算与目的 IP 地址或目的 IP 网段之间的最短路径和无环路非最短路径。以入口设备为例，入口设备可以计算和 D1 之间的最短路径和无环路非最短路径。其中，出接口 1 到 D1 的路径和出接口 2 到 D1 的路径为最短路径，出接口 3 到 D1 的路径和出接口 4 到 D1 的路径为无环路非最短路径。

无环路非最短路径是指该非最短路径中未存在环路, 本申请后续实施例中介绍的非最短路径均未存在环路。

为了提高网络的带宽利用率, 本申请实施例提供一种路由信息维护方法, 该方法可以应用于互联网协议第四版 (Internet Protocol version 4, IPv4)、互联网协议第六版 (Internet Protocol version 6, IPv6) 或互联网协议第六版段路由 (Segment Routing Internet Protocol Version 6, SRv6) 网络中, 具体可以在园区网络、数据中心网络以及广域网中使用。该方法可以由第一网络设备执行, 作为示例, 该第一网络设备可以为图 1 中的入口设备或中转设备。

第一网络设备包括多路径组转发表, 多路径组转发表包括第一多路径组中每条路径的转发表项, 多路径组转发表对应的第一多路径组包括最短路径和非最短路径, 最短路径的转发表项包括第一网络设备的第一出接口。

每条路径均用于指示第一网络设备与第二网络设备之间的路径。可以理解为, 每条路径均用于指示到达同一目的 IP 地址或同一目的 IP 网段的路径, 该目的 IP 地址为第二网络设备的 IP 地址, 该目的 IP 网段为第二网络设备的 IP 地址所在的 IP 网段。

其中, 多路径组转发表用于指导报文转发, 多路径组转发表的每个转发表项包括下一跳、出接口和该转发表项对应路径的转发权重值。

需要说明的是, 本申请实施例中的多路径组转发表位于一个整体转发表中, 该整体转发表中可以包括多个多路径组转发表, 换言之, 每个多路径组转发表可以作为整体转发表的一个表项。

如图 2 所示, 本申请实施例提供的路由信息维护方法包括以下步骤:

S201、当最短路径承载的流量使得第一出接口拥塞时, 获取多路径组转发表对应的预设路由策略。

可以理解的是, 第一出接口为上述多路径组转发表中最短路径的转发表项包括的出接口, 第一出接口还可能位于其他路径中, 即多个路径可以共用第一出接口转发流量。

当第一出接口拥塞时, 且最短路径承载的流量使得第一出接口拥塞, 说明最该短路径的转发负载较大, 可以获取该多路径组转发表对应的预设路由策略, 以对多条路径的转发负载进行调整。

S202、基于预设路由策略, 调整第一多路径组中每条路径的转发权重值。

其中, 每条路径的转发权重值用于表示从第一多路径组中选择该路径进行报文转发时, 该路径承载的流量比例。可以理解的是, 一条路径转发权重值越大, 代表利用该路径进行报文转发时, 该路径承载的流量比例越大, 即该路径将被用于转发更多的报文。

采用上述方法, 当最短路径承载的流量使得第一出接口拥塞时, 获取多路径组转发表对应的预设路由策略, 然后第一网络设备基于多路径组转发表对应的预设路由策略, 调整第一多路径组中每条路径的转发权重值。可见, 本申请实施例中, 当最短路径承载的流量使得第一出接口拥塞时, 并未直接停用第一出接口所在的最短路径, 而是对第一多路径组中的每条路径的转发权重值进行调整。由于每条路径的转发权重值用于表示第一多路径组中选择该路径进行报文时, 该路径承载的流量比例, 所以在调整每条路径的转发权重值之后, 每条路径的转发负载会发生变化, 可以减少最短路径的负载, 从而可以缓解第一出接口的拥塞, 充分利用各路径的资源, 提高网络的整体带宽利用率。

需要说明的是, 第一网络设备与第三网络设备之间的多条路径组成第二多路径组, 在第一多路径组与第二多路径组相同的情况下, 第一多路径组和第二多路径组共用上述多路径组转发表。

可以理解的是, 若第一网络设备的多个多路径组均相同, 则多个多路径组均可共用同一多路径组转发表。如此无需分别为每个多路径组创建和存储一个多路径组转发表, 可以节约存储资源。且更重要的

是能够提高路径更新速度，若需要对多个多路径组对应的多路径组转发表进行调整，则可以直接对共用的一个多路径组转发表进行调整，无需分别对多个多路径组转发表进行调整，可以提高多路径组转发表的更新速度和降低处理开销。

- 例如，第一网络设备到目的端设备 1 的路径包括：最短路径 A、非最短路径 B 和非最短路径 C；
- 5 第一网络设备到目的端设备 2 的路径包括：最短路径 A、非最短路径 B 和非最短路径 C。即第一网络设备到目的端设备 1 的路径和第一网络设备到目的端设备 2 的路径完全相同。这种情况下，目的端设备 1 的 IP 地址和目的端设备 2 的 IP 地址可以共用一个多路径组转发表。

本申请实施例中，上述多路径组转发表包括软件转发表和硬件转发表；在执行图 2 的流程之前，第一网络

10 设备中已生成多路径转发表和硬件转发表。

其中，软件转发表和硬件转发表通过以下步骤生成：

步骤 1、计算与第二网络设备之间的最短路径和非最短路径。

其中，第一网络设备可以基于最短路径优先算法计算得到第一网络设备与第二网络设备之间的最短路径和非最短路径。

- 具体的，可以获取第一网络设备的目的 IP 地址或目的 IP 网段，进而计算第一网络设备与该目的 IP
- 15 地址之间的最短路径和非最短路径，或者计算第一网络设备与该目的 IP 网段之间的最短路径和非最短路径。

第一网络设备与第二网络设备之间的最短路径数量为至少一条，非最短路径的数量可以为 0 条至多条。在不具有非最短路径，且具有多条最短路径的情况下，第一多路径组包括的多条最短路径为等价多

20 路径路由（Equal-cost multi-path routing, ECMP）。

- 步骤 2、分别针对每条最短路径和每条非最短路径生成一个软件转发表项，得到软件转发表，其中，每个软件转发表项包括一条路径的转发权重值，非最短路径的初始转发权重值为 0。

可以理解的是，软件转发表包括步骤 1 计算得到的最短路径和非最短路径。第一网络设备可以优先利用最短路径进行报文转发，在最短路径未发生拥塞的情况下，可以不使用非最短路径进行报文转发，因此最短路径的初始权重值可以为 1，非最短路径的初始权重值可以为 0。

- 25 如图 3a 所示，图 3a 示例性示出了 n 个目的 IP 地址共用的一个软件转发表，图 3a 左侧为 n 个目的 IP 地址和对应的路径组标识，分别为目的端设备 D1 至 Dn 的目的 IP 地址和对应的路径组标识。

例如，D1 目的 IP 地址是 119.1.1.2/32，Dn 目的 IP 地址是 121.1.1.2/32。假设第一网络设备与图 3a 中的各目的端设备之间的路径组，均包括两条最短路径和两条非最短路径。

- 图 3a 右侧为目的端设备 D1 至 Dn 的目的 IP 地址共用的软件转发表，作为示例，该软件转发表包
- 30 括 4 个软件转发表项。

第一个软件转发表项包括一条最短路径对应的下一跳（nexthop）1，以及该最短路径的转发权重值 1。

第二个软件转发表项包括另一条最短路径对应的下一跳 2，以及该最短路径的转发权重值 1。

第三个软件转发表项包括一条非最短路径对应的下一跳 1，以及该非最短路径的转发权重值 0。

- 35 第四个软件转发表项包括另一条非最短路径对应的下一跳 2，以及该非最短路径的转发权重值 0。

需要说明的是，图 3a 中每条软件转发表项还包括出接口，图 3a 中未示出。

步骤 3、针对转发权重值非 0 的每个软件转发表项，将该软件转发表项中除转发权重值之外的信息复制 N 份，得到 N 个硬件转发表项，将 N 个硬件转发表项下发至硬件资源，得到硬件转发表。

其中， $N = \text{该软件转发表项的转发权重值} \times \text{硬件转发表的预设表项数} / \text{软件转发表项包括的各转发权重值的和值}$ 。

预设表项数可以根据经验设置，例如预设表项数可以为 128。

5 可以理解的，在第一网络设备将硬件转发表项下发到硬件资源，得到硬件转发表之后，该第一网络设备的转发芯片可以利用硬件资源中的硬件转发表对报文进行转发。

其中，转发权重值为 0 的软件转发表项当前不被用于报文转发，处于备用状态，因此无需将转发权重值为 0 的软件转发表项下发到硬件资源中。

作为示例，基于图 3a 所示的软件转发表得到的硬件转发表如图 3b 所示，图 3b 左侧的 IP 地址与图 3a 左侧的 IP 地址相同，此处不再赘述。

10 图 3b 以硬件转发表的预设表项数是 128 为例，图 3a 中第三条表项和第四条表项的转发权重值均为 0，所以无需被下发到硬件转发表。图 3a 中的第一条软件转发表项和第二条软件转发表项的转发权重值均为 1，可以各被复制为 64 份，进而下发到硬件转发表。

所以，图 3b 的硬件转发表包括的前 64 条硬件转发表项相同，每条转发表项均包括一条最短路径的下一跳 1。

15 并且，该硬件转发表包括的后 64 条硬件转发表项相同，每条转发表项均包括另一条最短路径的下一跳 2。

需要说明的是，图 3b 中每条硬件转发表项还包括出接口，图 3b 中未示出。

20 可以理解的是，若硬件转发表中的表项数量由多变少，可能会存在报文匹配到一条硬件转发表项后，且在报文被转发之前，该硬件转发表项被删除，将造成丢包。本申请实施例中，硬件转发表包括的表项数量是固定值，即路径的权重发生变化后，不会导致硬件转发表中的表项数量发生变化，可以避免丢包。

在本申请的另一实施例中，第一网络设备中存储有第一接口状态表和多路径组转发表对应的第二接口状态表。其中，第一接口状态表包括第一网络设备的每个出接口的带宽占用量、缓冲区占用量和接口状态，第二接口状态表包括第一多路径组中的每条路径承载的流量对出接口的带宽占用量。

25 第一接口状态表是第一网络设备中已开启自适应路由功能的每个出接口上所有流量的综合状态。作为示例，以图 1 中入口设备的出接口分别为出接口 1、出接口 2、出接口 3 和出接口 4 为例，该入口设备的第一接口状态表如表 1 所示。

表 1

出接口	带宽占用量，单位 kbps	缓冲区占用量，单位字节	接口状态
出接口 1	BW1	BF1	正常状态
出接口 2	BW2	BF2	正常状态
出接口 3	BW3	BF3	正常状态
出接口 4	BW4	BF4	正常状态

接口状态可以为正常状态、拥塞状态或恢复状态，各出接口的初始状态设置为正常状态。

30 第一网络设备可以每隔预设时长，针对第一接口状态表中的每个出接口，更新该出接口的带宽占用量和缓冲区占用量，若该出接口更新后的带宽占用量达到预设带宽阈值，或该出接口更新后的缓冲区占用量达到预设缓冲区阈值，则将该出接口的接口状态设置为拥塞状态。

例如，预设时长可以为 5 毫秒。

其中，预设带宽阈值为带宽高阈值，预设缓冲区阈值为缓冲区高阈值。

第一网络设备确定接口状态的方法具体包括：

5 若某一接口的宽占用量达到带宽高阈值，或缓冲区占用量达到缓冲区高阈值，则将该接口的接口状态设置为拥塞状态，若该接口从正常状态或恢复状态变更为拥塞状态，则第一网络设备确定该接口拥塞；。

在某一接口处于拥塞状态时，若该接口更新后的宽占用量低于带宽高阈值，并且更新后的缓冲区占用量低于缓冲区高阈值，则将该接口的接口状态设置为恢复状态。

在某一接口处于恢复状态时，若该接口更新后的宽占用量低于带宽低阈值，并且更新后的缓冲区占用量低于缓冲区低阈值，则将该接口的接口状态设置为正常状态。

10 上述带宽高阈值、带宽低阈值、缓冲区高阈值和缓冲区低阈值是根据经验预先静态设置的。

或者也可以通过人工智能的方式，根据网络状态智能动态设置带宽高阈值、带宽低阈值、缓冲区高阈值和缓冲区低阈值。

例如，可以预先训练好一个深度学习模型，并且预先设置阈值更新周期，每隔一个阈值更新周期，将各出接口的带宽占用量和缓冲区占用量输入该深度学习模型，从而获取该深度模型输出的更新后的带宽高阈值、带宽低阈值、缓冲区高阈值和缓冲区低阈值。

15

并且第一网络设备中存储有第二接口状态表，第二接口状态表包括多路径组转发表对应的每个出接口的带宽占用量。作为示例，多路径组转发表 1 包括的出接口分别为出接口 1、出接口 2、出接口 3 和出接口 4，多路径组转发表 2 包括的出接口分别为出接口 1、和出接口 3，多路径组转发表 3 包括的出接口分别为出接口 2 和出接口 4，第一网络设备中存储的第二接口状态表如表 2 所示。

20 表 2

出接口	多路径组转发表 1 (路径类型：带宽占用量)	多路径组转发表 2（路 径类型：带宽占用量）	多路径组转发表 3（路 径类型：带宽占用量）
出接口 1	M:a	N:c	
出接口 2	M:b		N:h
出接口 3	N:c	M:f	
出接口 4	M:d		M:g

表 2 中，M 表示最短路径，N 表示非最短路径；“:”后的数字表示路径承载的流量对出接口的带宽占用量。

由表 2 可知，多路径组转发表 1 包括三条最短路径和一条非最短路径。其中，三条最短路径分别为最短路径 1、最短路径 2 和最短路径 3。最短路径 1 承载的流量对出接口 1 的带宽占用量为 a；最短路径 2 承载的流量对出接口 2 的带宽占用量为 b；最短路径 3 承载的流量对出接口 4 的带宽占用量为 d；非最短路径承载的流量对出接口 3 的带宽占用量为 c。

25

多路径转发表 2 包括一条最短路径和一条非最短路径。最短路径在出接口 3 上的带宽占用量为 f；

非最短路径在出接口 1 上的带宽占用量为 e。

多路径转发表 3 包括一条最短路径和一条非最短路径。最短路径在出接口 4 上的带宽占用量为 g；非最短路径在出接口 2 上的带宽占用量为 h。

在上述 S201、当最短路径承载的流量使得第一出接口拥塞时，获取多路径组转发表对应的预设路由策略之前，该方法还包括：

若第一接口状态表中第一接口的接口状态变更为拥塞状态，则从第二接口状态表中获取最短路径承载的流量对第一出接口的第一带宽占用量，并判断第一带宽占用量是否满足预设拥塞条件。

其中，预设拥塞条件可以为：使用第一出接口的各路径的带宽占用量中，第一带宽占用量最大。

也就是说若第一出接口所在的最短路径对第一出接口的带宽占用量最大，则说明是该最短路径承载的流量使得第一出接口拥塞。

或者，预设拥塞条件还可以为第二带宽占用量大于指定带宽阈值，该指定带宽阈值小于上文中的带宽高阈值。

以下结合具体例子进行说明，假设各出接口的最大带宽为 1000kbps、缓冲区占用量最大值为 1000 字节；带宽高阈值为 800kbps、带宽低阈值为 200 kbps，缓冲区高阈值为 800 字节、缓冲区低阈值为 200 字节。

以图 1 中入口设备的出接口分别为出接口 1、出接口 2、出接口 3 和出接口 4 为例，入口设备的第一接口状态表如表 3 所示。

表 3

出接口	带宽占用量，单位 kbps	缓冲区占用量，单位字节	接口状态
出接口 1	600	800	拥塞状态
出接口 2	600	500	正常状态
出接口 3	400	200	正常状态
出接口 4	300	400	正常状态

根据表 3，当表 3 中的出接口 1 被设置为拥塞状态时，入口设备确定出接口 1 进入拥塞状态。

假设入口设备包括多路径组转发表 1、多路径组转发表 2 和多路径组转发表 3，多路径组转发表 1 包括的出接口分别为出接口 1、出接口 2、出接口 3 和出接口 4，多路径组转发表 2 包括的出接口分别为出接口 1、和出接口 3，多路径组转发表 3 包括的出接口分别为出接口 2、和出接口 4，第一网络设备存储的多路径组转发表对应的第二接口状态表如表 4 所示。

表 4

出接口	多路径组转发表 1	多路径组转发表 2	多路径组转发表 3
出接口 1	M:600	N:0	
出接口 2	M:600		N:0
出接口 3	N:0	M:400	

出接口 4	N:0		M:300
-------	-----	--	-------

从表 4 中可以看出，对于出接口 1，多路径组转发表 1 对应的最短路径承载的流量对出接口 1 的带宽占用量最大，因此可确定是该最短路径承载的流量使得出接口 1 拥塞。

采用该方法，基于第一接口状态表可以及时确定第一出接口是否拥塞，在确定第一出接口拥塞的情况下，还可以确定使得第一出接口拥塞的路径，从而对该路径对应的多路径转发表对应的各路径的权重值进行调整，可以避免拥塞恶化，提高带宽利用率。

在本申请的另一实施例中，多路径组转发表包括路由类型标记，作为示例，路由类型标记可以占用软件转发表的一个字段。

上述 S201 中，获取多路径组转发表对应的预设路由策略，具体可以实现为：

获取路由类型标记对应的预设路由策略；预设路由策略至少包括以下策略中的一种：将拥塞出接口的转发负载调整为 0、减少拥塞出接口的转发负载且不增加非最短路径的转发负载、减少拥塞出接口的转发负载并增加非最短路径的转发负载。

作为示例，路由类型标记 1 对应的预设路由策略为将拥塞出接口的转发负载调整为 0，即不继续利用拥塞出接口转发报文；路由类型标记 2 对应的预设路由策略为减少拥塞出接口的转发负载且不增加非最短路径的转发负载；路由类型标记 3 对应的预设路由策略为减少拥塞出接口的转发负载并增加非最短路径的转发负载。本申请实施例不限制路由类型标记的形式。

需要说明的是，本申请实施例中的预设路由策略不限于以上三种，可以根据实际场景灵活设置。

采用上述方法，可以通过在多路径组转发表中设置路由类型标记的方式，匹配路由类型标记对应的预设路由策略。如此，可以灵活的通过路由类型标记调整多路径组转发表项使用的预设路由策略，以提高整体网络性能。

在确定预设路由策略后，S202、基于该预设路由策略，调整第一多路径组中每条路径的转发权重值，具体包括：

基于预设路由策略以及第二接口状态表包括的各接口的带宽占用量，调整第一多路径组中每条路径的转发权重值。

其中，第一网络设备可基于预设路由策略确定需要被调整转发权重值的路径。

若预设路由策略为将拥塞出接口的转发负载调整为 0，则可将拥塞出接口的转发权重值调整为 0，利用第一多路径组中未发生拥塞的出接口所在的路径负载分担第一出接口所在的最短路径的流量。

若预设路由策略为减少拥塞出接口的转发负载且不增加非最短路径的转发负载，则可按多个预设比例分别计算：若将第一出接口所在的最短路径的流量平均分配给其他出接口未拥塞的最短路径后，其他最短路径的出接口是否会发生拥塞。从而确定不会使得其他最短路径的出接口拥塞的最大预设比例，按照该预设比例调整每条最短路径的转发权重值。其中，多个预设比例为依次降低的多个比例，比如多个预设比例可以包括 1/2、1/4、1/8。

假设经计算确定，将第一出接口所在的最短路径承载的 1/2 的流量调整给其他最短路径后，其他最短路径的出接口不会发生拥塞，则按照将 1/2 的比例调整各最短路径的转发权重值。

若预设路由策略为减少拥塞出接口的转发负载并增加非最短路径的转发负载，则可按多个预设比例分别计算：若将第一出接口所在的最短路径的流量平均分配给其他出接口未拥塞的非最短路径后，其他非最短路径的出接口是否会发生拥塞。从而确定出不会使得其他非最短路径的出接口拥塞的最大预设

比例，按照该预设比例调整每条路径的转发权重值。

结合上述表 4 中的举例，假设预设路由策略为减少拥塞出接口的转发负载并增加非最短路径的转发负载，则可调整多路径组转发表 1 对应的多路径组中每条路径的转发权重值。

5 为方便描述，将表 4 中多路径组转发表项 1 对应的 4 条路径分别称为路径 1、路径 2、路径 3 和路径 4。根据表 3 可确定非最短路径的出接口 3 和出接口 4 均不拥塞，则可计算若将路径 1 的出接口 1 上 1/2 的流量平均分配到路径 3 的出接口 3 和路径 4 的出接口 4 后，出接口 3 和出接口 4 是否会发生拥塞，若不会，则可按照 1/2 的比例重新确定路径 1 至 4 对应的转发权重值。

10 计算得到的各出接口的带宽占用量如表 5 所示，即将出接口 1 中 300kbps 的流量均分给出口 3 和出口 4，可以看出调整后出接口 3 的带宽占用量为 550kbps，出接口 4 的带宽占用量为 450kbps，均未超出带宽高阈值，均不会发生拥塞。

表 5

出接口	带宽占用量，单位 kbps
出接口 1	$600-300=300$
出接口 2	600
出接口 3	$400+150=550$
出接口 4	$300+150=450$

15 假设计算得到出接口 2、出接口 3 和出接口 4 中任意一个出接口会发生拥塞，则计算若将路径 1 的出接口 1 上 1/4 的流量平均分配到路径 2 的出接口 3、路径 3 的出接口 3 和路径 4 的出接口 4 后，出接口 2、出接口 3 和出接口 4 是否会发生拥塞，若不会，则可按照 1/4 的比例重新确定路径 1 至 4 对应的转发权重值。若计算得到任意一个出接口会发生拥塞，则继续按照 1/8 进行计算，依次类推。

按照表 5 的例子，确定可按照 1/2 的比例重新确定路径 1 至 4 对应的转发权重值，则多路径组转发表 1 的 4 个软件转发表项包括的转发权重值由 1、1、0、0 调整为 1/2、1、1/4、1/4，也就是 2、4、1、1。

以下介绍对软件转发表和硬件转发表进行更新的方法：

20 在一个例子中，假设预设路由策略为减少拥塞出接口的转发负载并增加非最短路径的转发负载。在图 3a 的基础上，可以将 4 个软件转发表项包括的转发权重值由 1、1、0、0 调整为 2、4、1、1，调整后如图 4a 所示，如此可以将第一个软件转发表项对应的路径上一半的流量均分到两条非最短路径。

25 作为示例，以第一网络设备转发的 1000 个报文为例，在调整转发权重值之前，出接口 1 转发的报文数量为 $1000*1/2=500$ 个，出接口 2 转发的报文数量为 $1000*1/2=500$ 个，出接口 3 转发的报文数量和出接口 4 转发的报文数量均为 0。在调整转发权重值之后，出接口 1 转发的报文数量为 $1000*2/(2+4+1+1)=250$ 个，出接口 2 转发的报文数量为 $1000*4/(2+4+1+1)=500$ 个，出接口 3 转发的报文数量为 $1000*1/(2+4+1+1)=125$ 个，出接口 4 转发的报文数量为 $1000*1/(2+4+1+1)=125$ 个。即出接口 1 上的一半的流量被均分到出接口 3 和出接口 4。

30 在得到图 4a 中的软件转发表项后，需适应对硬件转发表进行更新，更新后的硬件转发表如图 4b 所示。

其中，每个软件转发表项需被复制的数量分别为：

第一个软件转发表项被复制的数量为： $2 \times 128 / (2+4+1+1) = 32$;

第二个软件转发表项被复制的数量为： $4 \times 128 / (2+4+1+1) = 64$;

第三个软件转发表项被复制的数量为： $1 \times 128 / (2+4+1+1) = 16$;

第四个软件转发表项被复制的数量为： $1 \times 128 / (2+4+1+1) = 16$ 。

- 5 在完成复制后，可以将各软件转发表中除转发权重值之外的字段下发至硬件转发表，从而得到如图 4b 所示的硬件转发表。

在另一个例子中，假设预设路由策略为减少拥塞出接口的转发负载且不增加非最短路径的转发负载，在图 3a 的基础上，可以将 4 个软件转发表项包括的转发权重值由 1、1、0、0 调整为 1、3、0 和 0，调整后如图 5a 所示。如此可以将一条最短路径上一半的流均分到另一条最短路径。

- 10 以第一网络设备转发的 1000 个报文为例，在调整转发权重值之前，出接口 1 转发的报文数量为 $1000 \times 1/2 = 500$ 个，出接口 2 转发的报文数量为 $1000 \times 1/2 = 500$ 个。在调整转发权重值之后，出接口 1 转发的报文数量为 $1000 \times 1 / (1+3) = 250$ 个，出接口 2 转发的报文数量为 $1000 \times 3 / (1+3) = 750$ 个，出接口 3 转发的报文数量为 0 个，出接口 4 转发的报文数量为 0 个。即将出接口 1 上的一半的报文分到出接口 2。

- 15 在得到图 5a 所示的软件转发表项后，需适应对硬件转发表进行更新，更新后的硬件转发表如图 5b 所示。

其中，每个软件转发表项需被复制的数量分别为：

第一个软件转发表项被复制的数量为： $1 \times 128 / (1+3+0+0) = 32$;

第二个软件转发表项被复制的数量为： $3 \times 128 / (1+3+0+0) = 96$;

- 20 第三个软件转发表项被复制的数量为： $0 \times 128 / (1+3+0+0) = 0$;

第四个软件转发表项被复制的数量为： $0 \times 128 / (1+3+0+0) = 0$ 。

在完成复制后，可以将各软件转发表中除转发权重值之外的字段下发至硬件转发表，从而得到如图 5b 所示的硬件转发表。

- 25 采用上述方法，当第一出接口拥塞时，第一网络设备可以根据多路径组转发表对应的预设路由策略，调整每个转发表项对应的路径的转发权重值。即第一网络设备可以利用其他转发表项对应的路径分担原本需要通过第一出接口进行转发的报文，减轻了第一出接口的负载压力。充分利用网络中的其他路径转发报文，提升了带宽利用率，在支持大规模组网的同时，可以满足高吞吐、低时延、低成本的要求。

- 30 本申请实施例提供的方法还可以应用于叶 (Leaf) -脊 (Spine) 拓扑组网系统，如图 6 所示，图 6 示例性示出了两个 Spine 设备和两个 Leaf 设备。两个 Spine 设备分别为 Spine 设备 1 和 Spine 设备 2，两个 Leaf 设备分别为 Leaf 设备 1 和 Leaf 设备 2。

其中，Leaf 设备 1 为源端设备，Leaf 设备 2 为目的端设备。Leaf 设备 1 可以向 Leaf 设备 2 发送报文。以 Leaf 设备 1 为例，Leaf 设备 1 计算得到两条路径，出接口 1 到 Leaf 设备 2 的路径和出接口 2 到 Leaf 设备 2 的路径均为最短路径，即 Leaf 设备 1 的软件转发表中，包括上述两条最短路径，上述两条最短路径形成 ECMP 路径。该软件转发表具体如图 7a 所示。

- 35 图 7a 示例性示出了在图 6 所示的组网系统中，Leaf 设备 1 中的软件转发表。图 7a 左侧是 D1 目的 IP 地址和对应的路径组标识，D1 目的 IP 地址为 Leaf 设备 2 的 IP 地址。例如 D1 目的 IP 地址是 119.1.1.2/32。

Leaf 设备 1 基于该路径组生成的软件转发表包括两个软件转发表项，第一个软件转发表项包括最短路径的下一跳 (nexthop) 1，以及转发权重值 1，第二个软件转发表项包括另一条最短路径的下一跳

nexthop 2, 以及转发权重值 1。

图 7b 为图 7a 对应的硬件转发表, 图 7b 的硬件转发表包括的前 64 条硬件转发表项相同, 每条转发表项均包括一条最短路径的下一跳 1。

并且, 该硬件转发表包括的后 64 条硬件转发表项相同, 每条转发表项均包括另一条最短路径的下一跳 2。

本申请实施例提供的方法还可以应用于蜻蜓 (Dragonfly) 拓扑组网系统中的每个组 (Group)。如图 8 所示, 图 8 中矩形虚线框代表组 1, 组 1 包括 Ingress 设备、Transit 设备和边缘 (Edge) 设备。图 8 中 S 表示源端设备, D1 表示目的端设备, 源端设备通过入口设备接入组 1, 目的端设备通过边缘设备接入组 1。

以入口设备为例, 入口设备可以计算和目的端设备之间的最短路径和非最短路径, 出接口 1 到目的端设备的路径和出接口 2 到目的端设备的路径为最短路径, 出接口 3 到目的端设备的路径和出接口 4 到目的端设备的路径为非最短路径。

图 9a 示例性示出了在图 8 所示的组网系统中, 入口设备中的软件转发表。其中, 图 9a 左侧是 D1 目的 IP 地址和对应的路径组标识, D1 目的 IP 地址为目的端设备的 IP 地址。例如 D1 目的 IP 地址是 119.1.1.2/32。

该软件转发表中, 第一个软件转发表项包括一条最短路径对应的下一跳 (nexthop) 1, 以及该最短路径的转发权重值 1。

第二个软件转发表项包括另一条最短路径对应的下一跳 2, 以及该最短路径的转发权重值 1。

第三个软件转发表项包括一条非最短路径对应的下一跳 1, 以及该非最短路径的转发权重值 0。

第四个软件转发表项包括另一条非最短路径对应的下一跳 2, 以及该非最短路径的转发权重值 0。

图 9b 为图 9a 对应的硬件转发表, 图 7b 的硬件转发表包括的前 64 条硬件转发表项相同, 每条转发表项均包括一条最短路径的下一跳 1。

并且, 该硬件转发表包括的后 64 条硬件转发表项相同, 每条转发表项均包括另一条最短路径的下一跳 2。

基于相同的构思, 本申请实施例提供一种路由信息维护装置, 该装置应用于第一网络设备, 第一网络设备包括多路径组转发表, 多路径组转发表对应的第一多路径组包括最短路径和非最短路径, 每条路径均用于指示第一网络设备与第二网络设备之间的路径; 最短路径的转发表项包括所述第一网络设备的第一出接口; 如图 10 所示, 该装置包括:

获取模块 1001, 用于当最短路径承载的流量使得第一出接口拥塞时, 获取多路径组转发表对应的预设路由策略;

调整模块 1002, 用于基于预设路由策略, 调整第一多路径组中每条路径的转发权重值, 每条路径的转发权重值用于表示从第一多路径组中选择该路径进行报文转发时, 该路径承载的流量比例。

可选的, 多路径组转发表包括软件转发表和硬件转发表; 该装置还包括生成模块:

生成模块, 用于:

计算与第二网络设备之间的最短路径和非最短路径;

分别针对每条最短路径和每条非最短路径生成一个软件转发表项, 得到软件转发表, 其中, 每个软件转发表项包括一条路径的转发权重值, 非最短路径的初始转发权重值为 0;

针对转发权重值非 0 的每个软件转发表项, 将该软件转发表项中除转发权重值之外的信息复制 N

份，得到 N 个硬件转发表项，将 N 个硬件转发表项下发至硬件资源，得到硬件转发表；

其中， $N = \text{该软件转发表项的转发权重值} \times \text{硬件转发表的预设表项数} / \text{软件转发表项包括的各转发权重值的和值}$ 。

可选的，多路径组转发表包括路由类型标记；获取模块 1001，具体用于：

5 获取路由类型标记对应的预设路由策略；预设路由策略至少包括以下策略中的一种：将拥塞出接口的转发负载调整为 0、减少拥塞出接口的转发负载且不增加非最短路径的转发负载、减少拥塞出接口的转发负载并增加非最短路径的转发负载。

10 可选的，第一网络设备中存储有第一接口状态表和多路径组转发表对应的第二接口状态表，第一接口状态表包括第一网络设备的每个出接口的接口状态，第二接口状态表包括第一多路径组中的每条路径承载的流量对出接口的带宽占用量；该装置还包括确定模块：

获取模块 1001，还用于若第一接口状态表中的第一接口的接口状态变更为拥塞状态，则从第二接口状态表中获取最短路径承载的流量对第一出接口的第一带宽占用量；

确定模块，用于若第一带宽占用量满足预设拥塞条件，则确定最短路径承载的流量使得第一出接口拥塞。

15 可选的，第一接口状态表还包括每个出接口的带宽占用量和缓冲区占用量；该装置还包括更新模块和设置模块：

更新模块，用于每隔预设时长，针对第一接口状态表中的每个出接口，更新该出接口的带宽占用量和缓冲区占用量；

20 设置模块，用于若该出接口更新后的带宽占用量达到预设带宽阈值，或该出接口更新后的缓冲区占用量达到预设缓冲区阈值，则将该出接口的接口状态设置为拥塞状态。

可选的，调整模块 1002，具体用于：

基于预设路由策略以及第二接口状态表包括的各出接口的带宽占用量，调整第一多路径组中每条路径的转发权重值。

25 可选的，第一网络设备与第三网络设备之间的多条路径组成第二多路径组，在第一多路径组与第二多路径组相同的情况下，第一多路径组和第二多路径组共用多路径组转发表。

基于相同的构思，本申请实施例还提供一种第一网络设备，第一网络设备包括多路径组转发表，多路径组转发表对应的第一多路径组包括最短路径和非最短路径，每条路径均用于指示第一网络设备与第二网络设备之间的路径；最短路径的转发表项包括所述第一网络设备的第一出接口；如图 11 所示，该第一网络设备包括：

30 处理器 1101；

收发器 1104；

机器可读存储介质 1102，机器可读存储介质 1102 存储有能够被处理器 1101 执行的机器可执行指令；机器可执行指令促使处理器 1101 执行以下步骤：

当最短路径承载的流量使得第一出接口拥塞时，获取多路径组转发表对应的预设路由策略；

35 基于预设路由策略，调整第一多路径组中每条路径的转发权重值，每条路径的转发权重值用于表示从第一多路径组中选择该路径进行报文转发时，该路径承载的流量比例。

可选的，多路径组转发表包括软件转发表和硬件转发表；机器可执行指令还促使处理器 1101 执行以下步骤：

计算与第二网络设备之间的最短路径和非最短路径；

分别针对每条最短路径和每条非最短路径生成一个软件转发表项，得到软件转发表，其中，每个软件转发表项包括一条路径的转发权重值，非最短路径的初始转发权重值为 0；

5 针对转发权重值非 0 的每个软件转发表项，将该软件转发表项中除转发权重值之外的信息复制 N 份，得到 N 个硬件转发表项，将 N 个硬件转发表项下发至硬件资源，得到硬件转发表；

其中， $N = \text{该软件转发表项的转发权重值} \times \text{硬件转发表的预设表项数} / \text{软件转发表项包括的各转发权重值的和值}$ 。

可选的，多路径组转发表包括路由类型标记；机器可执行指令还促使处理器 1101 执行以下步骤：

10 获取路由类型标记对应的预设路由策略；预设路由策略至少包括以下策略中的一种：将拥塞出接口的转发负载调整为 0、减少拥塞出接口的转发负载且不增加非最短路径的转发负载、减少拥塞出接口的转发负载并增加非最短路径的转发负载。

可选的，第一网络设备中存储有第一接口状态表和多路径组转发表对应的第二接口状态表，第一接口状态表包括第一网络设备的每个出接口的接口状态，第二接口状态表包括第一多路径组中的每条路径承载的流量对出接口的带宽占用量；

15 机器可执行指令还促使处理器 1101 执行以下步骤：

若第一接口状态表中的第一接口的接口状态变更为拥塞状态，则从第二接口状态表中获取最短路径承载的流量对第一出接口的第一带宽占用量；

若第一带宽占用量满足预设拥塞条件，则确定最短路径承载的流量使得第一出接口拥塞。

20 可选的，第一接口状态表还包括每个出接口的带宽占用量和缓冲区占用量；机器可执行指令还促使处理器 1101 执行以下步骤：

每隔预设时长，针对第一接口状态表中的每个出接口，更新该出接口的带宽占用量和缓冲区占用量；

若该出接口更新后的带宽占用量达到预设带宽阈值，或该出接口更新后的缓冲区占用量达到预设缓冲区阈值，则将该出接口的接口状态设置为拥塞状态。

可选的，机器可执行指令还促使处理器 1101 执行以下步骤：

25 基于预设路由策略以及第二接口状态表包括的各出接口的带宽占用量，调整第一多路径组中每条路径的转发权重值。

可选的，第一网络设备与第三网络设备之间的多条路径组成第二多路径组，在第一多路径组与第二多路径组相同的情况下，第一多路径组和第二多路径组共用多路径组转发表。

30 在图 11 中，还可以包括通信总线 1103。处理器 1101、机器可读存储介质 1102 及收发器 1104 之间通过通信总线 1103 完成相互间的通信，通信总线 1103 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。该通信总线可以分为地址总线、数据总线、控制总线等。

收发器 1104 可以为无线通信模块，收发器 1104 在处理器 1101 的控制下，与其他设备进行数据交互。

35 机器可读存储介质 1102 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，机器可读存储介质还可以是至少一个位于远离前述处理器的存储装置。

处理器 1101 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。

5 基于同一种发明构思,根据上述本申请实施例提供的路由信息维护方法,本申请实施例还提供了一种机器可读存储介质,机器可读存储介质存储有能够被处理器执行的机器可执行指令。处理器被机器可执行指令促使实现上述路由信息维护方法的步骤。

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路由信息维护方法的步骤。

10 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。

15 以上所述仅为本发明的较佳实施例,并不用以限制本发明,凡在本发明的精神和原则之内,所做的任何修改、等同替换、改进等,均应包含在本发明保护的范围之内。

权 利 要 求

- 1、一种路由信息维护方法，其特征在于，应用于第一网络设备，所述第一网络设备包括多路径组转发表，所述多路径组转发表对应的第一多路径组包括最短路径和非最短路径，每条路径均用于指示第一网络设备与第二网络设备之间的路径；所述最短路径的转发表项包括所述第一网络设备的第一出接口；
- 5 所述方法包括：
- 当所述最短路径承载的流量使得所述第一出接口拥塞时，获取所述多路径组转发表对应的预设路由策略；
- 基于所述预设路由策略，调整所述第一多路径组中每条路径的转发权重值，每条路径的转发权重值用于表示从所述第一多路径组中选择该路径进行报文转发时，该路径承载的流量比例。
- 10 2、根据权利要求1所述的方法，其特征在于，所述多路径组转发表包括软件转发表和硬件转发表；所述方法还包括：
- 计算与所述第二网络设备之间的最短路径和非最短路径；
- 分别针对每条最短路径和每条非最短路径生成一个软件转发表项，得到所述软件转发表，其中，每个软件转发表项包括一条路径的转发权重值，非最短路径的初始转发权重值为0；
- 15 针对转发权重值非0的每个软件转发表项，将该软件转发表项中除转发权重值之外的信息复制N份，得到N个硬件转发表项，将所述N个硬件转发表项下发至硬件资源，得到所述硬件转发表；
- 其中， $N = \frac{\text{该软件转发表项的转发权重值}}{\text{所述硬件转发表的预设表项数}} \times \text{所述软件转发表项包括的各转发权重值的和值}$ 。
- 3、根据权利要求2所述的方法，其特征在于，所述多路径组转发表包括路由类型标记；所述获取所述多路径组转发表对应的预设路由策略，包括：
- 20 获取所述路由类型标记对应的预设路由策略；
- 所述预设路由策略至少包括以下策略中的一种：将拥塞出接口的转发负载调整为0、减少拥塞出接口的转发负载且不增加非最短路径的转发负载、减少拥塞出接口的转发负载并增加非最短路径的转发负载。
- 25 4、根据权利要求1所述的方法，其特征在于，所述第一网络设备中存储有第一接口状态表和所述多路径组转发表对应的第二接口状态表，所述第一接口状态表包括所述第一网络设备的每个出接口的接口状态，所述第二接口状态表包括所述第一多路径组中的每条路径承载的流量对出接口的带宽占用量；
- 在所述当所述最短路径承载的流量导致所述第一出接口拥塞时，获取所述多路径组转发表对应的预设路由策略之前，所述方法还包括：
- 30 若所述第一接口状态表中的第一接口的接口状态变更为拥塞状态，则从所述第二接口状态表中获取所述最短路径承载的流量对所述第一出接口的第一带宽占用量；
- 若所述第一带宽占用量满足预设拥塞条件，则确定所述最短路径承载的流量使得所述第一出接口拥塞。
- 5、根据权利要求4所述的方法，其特征在于，所述第一接口状态表还包括每个出接口的带宽占用量和缓冲区占用量；所述方法还包括：
- 35 每隔预设时长，针对所述第一接口状态表中的每个出接口，更新该出接口的带宽占用量和缓冲区占用量；
- 若该出接口更新后的带宽占用量达到预设带宽阈值，或该出接口更新后的缓冲区占用量达到预设缓

缓冲区阈值，则将该出接口的接口状态设置为拥塞状态。

6、根据权利要求4所述的方法，其特征在于，所述基于所述预设路由策略，调整所述第一多路径组中每条路径的转发权重值，包括：

5 基于所述预设路由策略以及所述第二接口状态表包括的各出接口的带宽占用量，调整所述第一多路径组中每条路径的转发权重值。

7、根据权利要求1所述的方法，其特征在于，所述第一网络设备与第三网络设备之间的多条路径组成第二多路径组，在所述第一多路径组与所述第二多路径组相同的情况下，所述第一多路径组和所述第二多路径组共用所述多路径组转发表。

10 8、一种路由信息维护装置，其特征在于，应用于第一网络设备，所述第一网络设备包括多路径组转发表，所述多路径组转发表对应的第一多路径组包括最短路径和非最短路径，每条路径均用于指示第一网络设备与第二网络设备之间的路径；所述最短路径的转发表项包括所述第一网络设备的第一出接口；所述装置包括：

获取模块，用于当所述最短路径承载的流量使得所述第一出接口拥塞时，获取所述多路径组转发表对应的预设路由策略；

15 调整模块，用于基于所述预设路由策略，调整所述第一多路径组中每条路径的转发权重值，每条路径的转发权重值用于表示从所述第一多路径组中选择该路径进行报文转发时，该路径承载的流量比例。

9、根据权利要求8所述的装置，其特征在于，所述多路径组转发表包括软件转发表和硬件转发表；所述装置还包括生成模块：

所述生成模块，用于：

20 计算与所述第二网络设备之间的最短路径和非最短路径；

分别针对每条最短路径和每条非最短路径生成一个软件转发表项，得到所述软件转发表，其中，每个软件转发表项包括一条路径的转发权重值，非最短路径的初始转发权重值为0；

针对转发权重值非0的每个软件转发表项，将该软件转发表项中除转发权重值之外的信息复制N份，得到N个硬件转发表项，将所述N个硬件转发表项下发至硬件资源，得到所述硬件转发表；

25 其中， $N = \text{该软件转发表项的转发权重值} \times \text{所述硬件转发表的预设表项数} / \text{所述软件转发表项包括的各转发权重值的和值}$ 。

10、根据权利要求9所述的装置，其特征在于，所述多路径组转发表包括路由类型标记；所述获取模块，具体用于：

获取所述路由类型标记对应的预设路由策略；

30 所述预设路由策略至少包括以下策略中的一种：将拥塞出接口的转发负载调整为0、减少拥塞出接口的转发负载且不增加非最短路径的转发负载、减少拥塞出接口的转发负载并增加非最短路径的转发负载。

35 11、根据权利要求8所述的装置，其特征在于，所述第一网络设备中存储有第一接口状态表和所述多路径组转发表对应的第二接口状态表，所述第一接口状态表包括所述第一网络设备的每个出接口的接口状态，所述第二接口状态表包括所述第一多路径组中的每条路径承载的流量对出接口的带宽占用量；所述装置还包括确定模块：

所述获取模块，还用于若所述第一接口状态表中的第一接口的接口状态变更为拥塞状态，则从所述第二接口状态表中获取所述最短路径承载的流量对所述第一出接口的第一带宽占用量；

所述确定模块，用于若所述第一带宽占用量满足预设拥塞条件，则确定所述最短路径承载的流量使得所述第一出接口拥塞。

12、根据权利要求 11 所述的装置，其特征在于，所述第一接口状态表还包括每个出接口的带宽占用量和缓冲区占用量；所述装置还包括更新模块和设置模块：

5 所述更新模块，用于每隔预设时长，针对所述第一接口状态表中的每个出接口，更新该出接口的带宽占用量和缓冲区占用量；

所述设置模块，用于若该出接口更新后的带宽占用量达到预设带宽阈值，或该出接口更新后的缓冲区占用量达到预设缓冲区阈值，则将该出接口的接口状态设置为拥塞状态。

13、根据权利要求 11 所述的装置，其特征在于，所述调整模块，具体用于：

10 基于所述预设路由策略以及所述第二接口状态表包括的各出接口的带宽占用量，调整所述第一多路径组中每条路径的转发权重值。

14、根据权利要求 8 所述的装置，其特征在于，所述第一网络设备与第三网络设备之间的多条路径组成第二多路径组，在所述第一多路径组与所述第二多路径组相同的情况下，所述第一多路径组和所述第二多路径组共用所述多路径组转发表。

15 15、一种第一网络设备，其特征在于，所述第一网络设备包括多路径组转发表，所述多路径组转发表对应的第一多路径组包括最短路径和非最短路径，每条路径均用于指示第一网络设备与第二网络设备之间的路径；所述最短路径的转发表项包括所述第一网络设备的第一出接口；所述第一网络设备包括：

处理器；

收发器；

20 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：

当所述最短路径承载的流量使得所述第一出接口拥塞时，获取所述多路径组转发表对应的预设路由策略；

基于所述预设路由策略，调整所述第一多路径组中每条路径的转发权重值，每条路径的转发权重值用于表示从所述第一多路径组中选择该路径进行报文转发时，该路径承载的流量比例。

25 16、根据权利要求 15 所述的第一网络设备，其特征在于，所述多路径组转发表包括软件转发表和硬件转发表；所述机器可执行指令还促使所述处理器执行以下步骤：

计算与所述第二网络设备之间的最短路径和非最短路径；

30 分别针对每条最短路径和每条非最短路径生成一个软件转发表项，得到所述软件转发表，其中，每个软件转发表项包括一条路径的转发权重值，非最短路径的初始转发权重值为 0；

针对转发权重值非 0 的每个软件转发表项，将该软件转发表项中除转发权重值之外的信息复制 N 份，得到 N 个硬件转发表项，将所述 N 个硬件转发表项下发至硬件资源，得到所述硬件转发表；

其中， $N = \text{该软件转发表项的转发权重值} \times \text{所述硬件转发表的预设表项数} / \text{所述软件转发表项包括的各转发权重值的和值}$ 。

35 17、根据权利要求 16 所述的第一网络设备，其特征在于，所述多路径组转发表包括路由类型标记；所述机器可执行指令还促使所述处理器执行以下步骤：

获取所述路由类型标记对应的预设路由策略；所述预设路由策略至少包括以下策略中的一种：将拥塞出接口的转发负载调整为 0、减少拥塞出接口的转发负载且不增加非最短路径的转发负载、减少拥塞

出接口的转发负载并增加非最短路径的转发负载。

18、根据权利要求 15 所述的第一网络设备，其特征在于，所述第一网络设备中存储有第一接口状态表和所述多路径组转发表对应的第二接口状态表，所述第一接口状态表包括所述第一网络设备的每个出接口的接口状态，所述第二接口状态表包括所述第一多路径组中的每条路径承载的流量对出接口的带宽占用量；

所述机器可执行指令还促使所述处理器执行以下步骤：

若所述第一接口状态表中的第一接口的接口状态变更为拥塞状态，则从所述第二接口状态表中获取所述最短路径承载的流量对所述第一出接口的第一带宽占用量；

若所述第一带宽占用量满足预设拥塞条件，则确定所述最短路径承载的流量使得所述第一出接口拥塞。

19、根据权利要求 18 所述的第一网络设备，其特征在于，所述第一接口状态表还包括每个出接口的带宽占用量和缓冲区占用量；所述机器可执行指令还促使所述处理器执行以下步骤：

每隔预设时长，针对所述第一接口状态表中的每个出接口，更新该出接口的带宽占用量和缓冲区占用量；

若该出接口更新后的带宽占用量达到预设带宽阈值，或该出接口更新后的缓冲区占用量达到预设缓冲区阈值，则将该出接口的接口状态设置为拥塞状态。

20、根据权利要求 18 所述的第一网络设备，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：

基于所述预设路由策略以及所述第二接口状态表包括的各出接口的带宽占用量，调整所述第一多路径组中每条路径的转发权重值。

21、根据权利要求 15 所述的第一网络设备，其特征在于，所述第一网络设备与第三网络设备之间的多条路径组成第二多路径组，在所述第一多路径组与所述第二多路径组相同的情况下，所述第一多路径组和所述第二多路径组共用所述多路径组转发表。

22、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 1-7 任一所述的方法。

23、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所述处理器：实现权利要求 1-7 任一所述的方法。