

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2020 年 12 月 3 日 (03.12.2020)



(10) 国际公布号
WO 2020/238875 A1

- (51) 国际专利分类号:
H04L 12/24 (2006.01)
- (21) 国际申请号: PCT/CN2020/092209
- (22) 国际申请日: 2020 年 5 月 26 日 (26.05.2020)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201910470325.9 2019年5月31日 (31.05.2019) CN
- (71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (72) 发明人: 尹超(YIN, Chao); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 陈永献(CHEN, Yongxian); 中国广东省深

圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 刘和洋(LIU, Heyang); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 于翔(YU, Xiang); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 郑合文(ZHENG, Hewen); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 孙黎阳(SUN, Liyang); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(74) 代理人: 北京龙双利达知识产权代理有限公司 (LONGSUN LEAD IP LTD.); 中国北京市海淀区北清路68号院3号楼101, Beijing 100094 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS,

(54) Title: METHOD AND APPARATUS FOR DETERMINING PORT ATTRIBUTE

(54) 发明名称: 确定端口属性的方法和装置

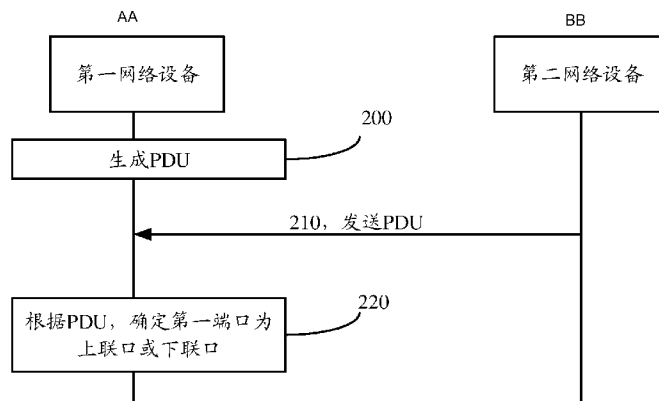


图 2

- 200 Generate a PDU
210 Send the PDU
220 Determine that a first port is an uplink port or a downlink port according to the PDU
AA First network device
BB Second network device

(57) Abstract: The present application provides a method and apparatus for determining a port attribute, which can be applied to a data center network, such as a fat-tree network. The method comprises: a first network device receives, by means of a first port, a Protocol Data Unit (PDU) sent by a second network device; and the first network device determines that the first port is an uplink port or a downlink port according to the PDU. In embodiments of the present application, the first network device can determine that the first port is the uplink port or the downlink port on the basis of the PDU sent by the second network device, so that a basis is provided for subsequently helping an SDN controller to reduce a load. The problem in the prior art that it is impossible to help the SDN controller to reduce the load because the network device cannot know the aforementioned configuration situation is avoided.



JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

(57) 摘要: 本申请提供一种确定端口属性的方法和装置, 可以应用于数据中心网络, 例如胖树(fat-tree)网络。该方法包括: 第一网络设备通过第一端口接收第二网络设备发送的协议数据单元PDU; 所述第一网络设备根据所述PDU, 确定所述第一端口为上联口或下联口。在本申请实施例中, 第一网络设备可以基于第二网络设备发送的PDU, 可以确定第一端口为上联口还是下联口, 为后续帮助SDN控制器减轻负载提供基础。避免了现有技术中, 网络设备无法获知上述配置情况, 导致无法帮助SDN控制器减轻负载。

说明书

确定端口属性的方法和装置

5 本申请要求于 2019 年 05 月 31 日提交中国专利局、申请号为 201910470325.9、申请名称为“确定端口属性的方法和装置”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

10 本申请涉及本申请涉及信息技术领域，并且更具体地，涉及确定端口属性的方法和装置。

背景技术

15 软件定义网络（Software Defined Network, SDN）是网络虚拟化的一种实现方式，旨在将网络设备的控制面与数据面分离开来，从而实现了数据流的灵活控制，使网络作为管道变得更加智能，为核心网络及应用的创新提供了良好的平台。

20 目前，SDN 技术被广泛应用于数据中心网络，由 SDN 控制器对数据中心网络中的交换机进行管理。例如，获取数据中心网络的网络拓扑，并基于网络拓扑对数据中心中传输的数据流情况生成流量控制策略，发送至网络设备，由网络设备基于流量控制策略对数据流进行控制。也就是说，SDN 控制器需要获取数据中心网络的配置情况，例如，数据中心网络网络拓扑，网络设备的端口属性等。而对于网络设备而言，网络设备无需知道数据中心的配置情况，仅基于 SDN 控制器的指示，传输数据流即可。

25 随着数据中心网络的规模不断的扩张，SDN 控制器的负载越来越大，会对数据中心的性能产生影响。为了减少 SDN 控制器的负载，希望将原先由 SDN 控制器实现的功能可以转由网络设备自主实现。然而，目前由于网络设备无法获知上述配置情况，导致无法帮助 SDN 控制器减轻负载。

发明内容

30 本申请提供一种确定端口属性的方法和装置，以使网络设备可以获知自身端口的端口属性，为后续帮助 SDN 控制器减轻负载提供基础。

 第一方面，提供了一种确定端口属性的方法，包括：第一网络设备通过第一端口接收第二网络设备发送的协议数据单元 PDU；所述第一网络设备根据所述 PDU，确定所述第一端口为上联口或下联口。

35 在本申请实施例中，第一网络设备可以基于第二网络设备发送的 PDU，可以确定第一端口为上联口还是下联口，为后续帮助 SDN 控制器减轻负载提供基础。避免了现有技术中，网络设备无法获知上述配置情况，导致无法帮助 SDN 控制器减轻负载。

 在一种可能的实现方式中，所述 PDU 携带网络中所述第二网络设备的级数，和/或所述 PDU 携带所述第二网络设备发送所述 PDU 使用的第二端口的端口属性，所述第二端口

的端口属性包括所述第二端口为上联口或所述第二端口为下联口。

在本申请实施例中，通过在 PDU 中携带第二网络设备的级数，以及第二端口的端口属性中的至少一种信息，有利于更加直观的辅助第一网络设备确定第一端口为上联口还是下联口。

5 在一种可能的实现方式中，所述第一网络设备根据所述 PDU，确定所述第一端口为上联口或下联口，包括：所述第一网络设备根据所述 PDU 确定在所述网络中所述第一网络设备的级数；所述第一网络设备根据所述第一网络设备的级数，确定所述第一端口为上联口或下联口。

10 在本申请实施例中，第一网络设备根据第一网络设备的级数确定第一端口为上联口还是下联口，有利于提高确定第一端口的端口属性的准确性。

在一种可能的实现方式中，所述网络包括多个网络设备，所述多个网络设备包括所述第二网络设备，所述第一网络设备通过第一端口接收第二网络设备发送的协议数据单元 PDU，包括：所述第一网络设备通过多个端口接收所述多个网络设备发送的多个 PDU，所述多个端口包括所述第一端口，所述多个 PDU 包括所述第二网络设备发送的所述 PDU；
15 所述第一网络设备根据所述 PDU 确定在所述网络中所述第一网络设备的级数，包括：所述第一网络设备根据所述多个 PDU，确定在所述网络中所述多个网络设备的级数；所述第一网络设备从所述多个网络设备的级数中选择目标级数；所述第一网络设备根据所述目标级数以及预设规则确定所述第一网络设备的级数。

20 在本申请实施例中，第一网络设备基于多个网络设备的级数确定目标级数，再根据目标级数以及预设的规则确定第一网络设备的级数，使得第一网络设备可以基于多个网络设备的级数确定第一网络设备的级数，有利于提高确定第一网络设备的级数的可靠性。

可选地，上述多个网络设备是与第一网络设备直接相连的网络设备，也就是说，网络中多个网络设备所在的层级与第一网络设备所在的层级相邻。

需要说明的是，上述多个网络设备所在的层级可以是一个层级也可以是多个层级。当
25 多个网络设备所在的层级可以是一个层级时，即多个网络设备所在的层级相同，且该层级与第一网络设备的层级相邻。当多个网络设备所在的层级可以是多个层级时，多个网络设备中至少两个网络设备所在的层级不同，且该多个层级与第一网络设备的层级相邻。例如，多个层级可以包括网络中位于第一网络设备所在层级以北的一层，以及第一网络设备所在层级以南的一层。

30 可选地，上述多个网络网络设备还可以是与第一网络设备不相邻的网络设备。例如，多个网络设备为根节点，第一网络设备的层次与多个网络设备的层级间隔一层或多层。此时，多个网络设备可以将根节点的级数通过多个 PDU 发送给第一网络设备，多个 PDU 中可以分别记录将对应的 PDU 传输至第一网络设备时进过的网络设备的数量，即传输 PDU 至第一网络设备时所需的跳数，第一网络设备可以基于 PDU 以及跳数确定第一网络设备的级数。
35

可选地，上述预设规则可以是以目标级数为基础，增加预设值，例如，1。上述预设规则还可以是以目标级数为基础，减少预设值，例如，1。上述预设规则还可以是通过非线性的函数表示。

在一种可能的实现方式中，所述第一网络设备根据所述第一网络设备的级数，确定所

述第一端口为所述上联口或所述下联口，包括：所述第一网络设备根据所述第一网络设备的级数，以及所述第二网络设备的级数确定所述第一端口为上联口或下联口。

在本申请实施例中，第一网络设备基于第一网络设备的级数，以及第二网络设备的级数，确定第一端口为上联口还是下联口，有利于提高确定第一端口的端口属性的可靠性。

- 5 在一种可能的实现方式中，所述第一网络设备根据所述 PDU，确定所述第一端口为上联口或下联口，包括：所述第一网络设备根据所述 PDU 确定所述第二网络设备的级数；所述第一网络设备根据所述第二网络设备的级数确定所述第一端口为上联口或下联口。

在本申请实施例中，第一网络设备可以直接基于第二网络设备的级数，确定第一端口为上联口还是下联口，有利于简化确定第一端口为上联口还是下联口的过程。

- 10 在一种可能的实现方式中，所述第二网络设备为所述网络中的根节点，所述第一网络设备通过第一端口接收第二网络设备发送的 PDU，包括：所述第一网络设备通过所述第一端口接收所述第二网络设备发送的所述 PDU，所述 PDU 中携带的所述第二网络设备的级数为预设值；所述第一网络设备根据所述第二网络设备的级数确定所述第一端口为上联口或下联口，包括：所述第一网络设备根据所述预设值确定所述第一端口为上联口。

- 15 在本申请实施例中，若第二网络设备为根节点时，第二网络设备的级数为预设值，这样第一网络设备通过 PDU 确定第二网络设备的级数为预设值时，可以直接确定第二网络设备为根节点，相应地，与第二网络设备通信的第一端口为上联口，有利于简化确定第一端口为上联口还是下联口的过程。

- 20 在一种可能的实现方式中，所述方法还包括：所述第一网络设备根据所述预设值，确定所述第一网络设备的级数，其中，在网络中，所述第一网络设备的级数表示的层级与所述根节点所在的层级相邻。

- 25 在本申请实施例中，若第二网络设备为根节点时，第二网络设备的级数为预设值，这样，第一网络设备通过 PDU 确定第二网络设备的级数为预设值时，可以直接确定第一网络设备所在的层级与根节点（即第二网络设备）的层级相邻，有利于简化确定第一网络设备在网络中的层级的过程。

- 30 在一种可能的实现方式中，所述第一网络设备根据所述 PDU，确定所述第一端口为上联口或下联口，包括：所述第一网络设备根据所述 PDU，确定所述第二网络设备为服务器，所述 PDU 携带第一 TLV，所述第一 TLV 指示所述第二网络设备为服务器；所述第一网络设备确定所述第一端口为下联口。

- 35 在本申请实施例中，通过在 PDU 中携带第一 TLV，以指示第二网络设备为服务器，这样，第一网络设备通过 PDU 确定第二网络设备为服务器时，可以直接确定与第二网络设备通信的第一端口为下联口，有利于简化确定第一端口的端口属性的过程。

可选地，上述第一 TLV 可以占用 PDU 中的可选 TLV，或者占用 PDU 中的保留 TLV。

可选地，所述 PDU 为链路层发现协议 LLDP PDU。

- 35 在一种可能的实现方式中，所述第一网络设备根据所述 PDU 确定所述第一端口为上联口或下联口，包括：所述第一网络设备根据所述 LLDP PDU 中端口 ID 的子类型，确定所述第二网络设备为服务器；所述第一网络设备确定所述第一端口为下联口。

在本申请实施例中，基于 LLDP 中协议规定交换机使用的端口 ID 的子类型与服务器使用的端口 ID 的子类型不同这一规律，可以通过 LLDP PDU 中携带的端口 ID 的子类型，

确定第二网络设备为服务器，这样，与第二网络设备通信的第一端口即为下联口，有利于简化确定第一端口的端口属性的过程。

在一种可能的实现方式中，所述方法还包括：所述第一网络设备根据所述 LLDP PDU 中端口 ID 的子类型，确定所述第一网络设备的级数，其中，在网络中所述第一网络设备的级数表示的层级与所述服务器所在的层级相邻。

在本申请实施例中，基于 LLDP 中协议规定交换机使用的端口 ID 的子类型与服务器使用的端口 ID 的子类型不同这一规律，可以通过 LLDP PDU 中携带的端口 ID 的子类型，确定第二网络设备为服务器，这样，第一网络设备在网络中的层级与服务器相邻，有利于简化确定第一网络设备在网络中的层级的过程。

在一种可能的实现方式中，在所述第一网络设备接收第二网络设备发送的协议数据单元 PDU 之前，所述方法还包括：所述第一网络设备向所述第二网络设备发送 LLDP 请求，所述 LLDP 请求用于请求所述第二网络设备的第二端口的端口属性，和/或所述第二网络设备在网络中的级数。

在本申请实施例中，通过第一网络设备向第二网络设备发送 LLDP 请求，以触发第二网络设备向第一网络设备反馈 LLDP PDU，以通过主动请求的方式，请求第二端口的端口属性和/或所述第二网络设备在网络中的级数，有利于提高第一网络设备获取上述信息的灵活性。

在一种可能的实现方式中，所述第一网络设备向所述第二网络设备发送 LLDP 请求，包括：若所述第一端口的端口属性未知，所述第一网络设备向所述第二网络设备发送所述 LLDP 请求。

在本申请实施例中，当第一端口的端口属性未知时，第一网络设备向第二网络设备发送 LLDP 请求，以触发第二网络设备向第一网络设备反馈 LLDP PDU，即第一网络设备有端口的端口属性需要确定时，才向第二网络设备发送 LLDP 请求，有利于减少第一网络设备与第二网络设备之间的通信占用的传输资源。

在一种可能的实现方式中，所述 LLDP PDU 包括第二 TLV，所述第二 TLV 用于指示所述第二网络设备的级数，和/或所述 LLDP PDU 包括第三 TLV，所述第三 TLV 用于指示所述第二端口的端口属性。

在本申请实施例中，上述第二网络设备的级数和/或第二端口的端口属性可以以 TLV 的格式携带在 LLDP PDU 中，即复用传统 LLDP PDU 的封装格式携带上述信息，以避免各厂商提供的网络设备需要兼容新的机制，以支持本申请提供的确定端口的端口属性的方法，有利于减少各厂商提供的网络设备为兼容本申请提供的方法所做的改动。

可选地，上述第二 TLV 或第三 TLV 可以占用 LLDP PDU 中的可选 TLV，例如，系统能力 TLV，以减少对传统 LLDP PDU 的封装格式的改动。

可选地，上述第二 TLV 或第三 TLV 可以占用 LLDP PDU 中的保留 TLV，以减少对传统 LLDP PDU 的中可用 TLV 的改动。

在一种可能的实现方式中，所述第一网络设备包括多个发送队列，所述多个发送队列包括第一发送队列和第二发送队列，所述第一发送队列为数据流对应的原始发送队列，所述方法还包括：所述第一网络设备确定所述第一端口以及所述第一网络设备的第三端口为上联口；所述第一网络设备生成防优先级流量控制 PFC 死锁规则，所述防 PFC 死锁规则

用于指示将所述数据流缓存至所述第一发送队列，所述第一数据流的入端口为所述第一端口，所述第一数据流的出端口为所述第三端口。

在本申请实施例中，第一网络设备可以获知第一端口以及第三端口为上联口，并为第一端口和第三端口生成防 PFC 死锁规则，以防止从第一端口转入并从第三端口转储的第一数据流在网络中与其他数据流形成 PFC 死锁。避免了传统的防 PFC 死锁规则部署方式中，需要由 SDN 控制器为第一端口以及第三端口生成防 PFC 死锁规则，并最终部署到第一网络设备中。有利于减轻 SDN 控制器在部署防 PFC 死锁规则方面的负载。

在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在预设时间段内，所述第一端口与所述第三端口为 PFC 死锁关联的端口的次数；所述第一网络设备生成防 PFC 死锁规则，包括：当所述次数高于预设阈值，所述第一网络设备确定生成所述防 PFC 死锁规则。

在本申请实施例中，在预设时间段内，当次数高于预设阈值时，第一网络设备才为第一端口和第三端口生成防 PFC 死锁规则，有利于减少第一网络设备存储防 PFC 死锁规则占用的存储空间。避免了无论 PFC 死锁是否产生，第一网络设备都为第一端口和第三端口生成防 PFC 死锁规则，导致第一网络设备存储防 PFC 死锁规则占用较大的存储空间。

在一种可能的实现方式中，所述方法还包括：所述第一网络设备通过所述第一端口接收所述第二网络设备发送的指示信息，所述指示信息用于指示所述第二网络设备中的第三发送队列对应的缓存的容量超过阈值，且所述第三发送队列的入流量为 0，所述第三发送队列缓存的数据流的优先级与所述第一数据流的优先级相同；所述第一网络设备根据指示信息确定所述次数。

在本申请实施例中，当第二网络设备中第三发送队列对应的缓存的容量超过阈值，且所述第三发送队列的入流量为 0 时，第二网络设备可以向第一网络设备发送指示信息，相应地，第一网络设备可以根据指示信息确定为与第三发送队列中缓存的流量的优先级的流量（即第一数据流）部署防 PFC 死锁规则。有利于降低网络中产生 PFC 死锁的可能性。

可选地，上述指示信息可以为第三发送队列缓存的数据流的优先级。

可选地，上述指示信息中可以承载于 PDU 中。具体地，可以封装为 PDU 中的一个 TLV。需要说明的是，上述承载指示信息的 PDU 与上文中确定端口的端口属性的 PDU 可以是相同的 PDU 也可以是不同的 PDU，本申请对此不作限定。

在一种可能的实现方式中，在所述网络的网络拓扑中，以北向到南向的方向网络设备的级数递增。

在一种可能的实现方式中，在所述网络的网络拓扑中，以北向到南向的方向网络设备的级数递减。

上述北向到南向的方向还可以理解为从根节点到架顶交换机的方向。通常，架顶交换机为网络中的一类网络设备，通常，上述架顶交换机位于网络中与服务器相邻的层级。

上述北向到南向的方向与从北向接口（Northbound Interface）到南向接口的方向相同，其中，北向接口用于为厂商或运营商进行接入和管理网络提供的接口，即向上提供的接口。南向接口用于为厂商提供管理其他厂商的网络设备提供的接口。即向下提供的接口。

第二方面，提供了一种确定端口属性的方法，包括：第二网络设备生成协议数据单

元 PDU, 所述 PDU 用于确定第一网络设备的第一端口为上联口或下联口; 所述第二网络设备向所述第一网络设备发送所述 PDU。

在本申请实施例中, 第一网络设备可以基于第二网络设备发送的 PDU, 可以确定第一端口为上联口还是下联口, 为后续帮助 SDN 控制器减轻负载提供基础。避免了现有技术中, 网络设备无法获知上述配置情况, 导致无法帮助 SDN 控制器减轻负载。

在一种可能的实现方式中, 所述 PDU 携带网络中所述第二网络设备的级数, 和/或所述 PDU 携带所述第二网络设备发送所述 PDU 使用的第二端口的端口属性, 所述第二端口的端口属性包括所述第二端口为上联口或所述第二端口为下联口。

在本申请实施例中, 通过在 PDU 中携带第二网络设备的级数, 以及第二端口的端口属性中的至少一种信息, 有利于更加直观的辅助第一网络设备确定第一端口为上联口还是下联口。

在一种可能的实现方式中, 所述第二网络设备为所述网络中的根节点, 所述 PDU 中携带的所述第二网络设备的级数为预设值。

在本申请实施例中, 若第二网络设备为根节点时, 第二网络设备的级数为预设值, 这样第一网络设备通过 PDU 确定第二网络设备的级数为预设值时, 可以直接确定第二网络设备为根节点, 相应地, 与第二网络设备通信的第一端口为上联口, 有利于简化确定第一端口为上联口还是下联口的过程。

在一种可能的实现方式中, 所述第二网络设备为服务器, 所述 PDU 携带第一 TLV, 所述第一 TLV 指示所述第二网络设备为服务器。

在本申请实施例中, 通过在 PDU 中携带第一 TLV, 以指示第二网络设备为服务器, 这样, 第一网络设备通过 PDU 确定第二网络设备为服务器时, 可以直接确定与第二网络设备通信的第一端口为下联口, 有利于简化确定第一端口的端口属性的过程。

可选地, 上述第一 TLV 可以占用 PDU 中的可选 TLV, 或者占用 PDU 中的保留 TLV。

在一种可能的实现方式中, 所述 PDU 为链路层发现协议 LLDP PDU。

在一种可能的实现方式中, 在所述第二网络设备向所述第一网络设备发送所述 PDU 之前, 所述方法还包括: 所述第二网络设备接收所述第一网络设备发送的 LLDP 请求, 所述 LLDP 请求用于请求所述第二端口的端口属性, 和/或所述第二网络设备的级数。

在本申请实施例中, 通过第一网络设备向第二网络设备发送 LLDP 请求, 以触发第二网络设备向第一网络设备反馈 LLDP PDU, 以通过主动请求的方式, 请求第二端口的端口属性和/或所述第二网络设备在网络中的级数, 有利于提高第一网络设备获取上述信息的灵活性。

在一种可能的实现方式中, 所述 LLDP PDU 包括第二 TLV, 所述第二 TLV 用于指示所述第二网络设备的级数, 和/或所述 LLDP PDU 包括第三 TLV, 所述第三 TLV 用于指示所述第二端口的端口属性。

在本申请实施例中, 上述第二网络设备的级数和/或第二端口的端口属性可以以 TLV 的格式携带在 LLDP PDU 中, 即复用传统 LLDP PDU 的封装格式携带上述信息, 以避免各厂商提供的网络设备需要兼容新的机制, 以支持本申请提供的确定端口的端口属性的方法, 有利于减少各厂商提供的网络设备为兼容本申请提供的方法所做的改动。

可选地, 上述第二 TLV 或第三 TLV 可以占用 LLDP PDU 中的可选 TLV, 例如, 系

统能力 TLV，以减少对传统 LLDP PDU 的封装格式的改动。

可选地，上述第二 TLV 或第三 TLV 可以占用 LLDP PDU 中的保留 TLV（或称保留位），以减少对传统 LLDP PDU 的中可用 TLV 的改动。

5 在一种可能的实现方式中，所述方法还包括：所述第二网络设备向所述第一网络设备发送指示信息，所述指示信息用于指示所述第二网络设备中的第三发送队列对应的缓存的容量超过阈值，且所述第三发送队列的入流量为 0，所述第三发送队列缓存的数据流的优先级与所述第一数据流的优先级相同。

10 在本申请实施例中，当第二网络设备中第三发送队列对应的缓存的容量超过阈值，且所述第三发送队列的入流量为 0 时，第二网络设备可以向第一网络设备发送指示信息，相应地，第一网络设备可以根据指示信息确定为与第三发送队列中缓存的流量的优先级相同的流量（即第一数据流）部署防 PFC 死锁规则。有利于降低网络中产生 PFC 死锁的可能性。

可选地，上述指示信息可以为第三发送队列缓存的数据流的优先级。

15 可选地，上述指示信息中可以承载于 PDU 中。具体地，可以封装为 PDU 中的一个 TLV。需要说明的是，上述承载指示信息的 PDU 与上文中确定端口的端口属性的 PDU 可以是相同的 PDU 也可以是不同的 PDU，本申请对此不作限定。

第三方面，提供了一种第一网络设备，包括实现上述方法中第一网络设备需要执行的功能的各个模块。

20 第四方面，提供了一种第二网络设备，包括实现上述方法中第二网络设备需要执行的功能的各个模块。

第五方面，提供了一种第一网络设备，包括存储器以及处理器，该存储器用于存储计算机程序，该处理器用于从存储器中调用并运行该计算机程序，使得该第一网络设备执行上述第一方面中的方法。

25 第六方面，提供了一种第二网络设备，包括存储器以及处理器，该存储器用于存储计算机程序，该处理器用于从存储器中调用并运行该计算机程序，使得该第二网络设备执行上述第二方面中的方法。

第七方面，提供了一种计算机程序产品，所述计算机程序产品包括：计算机程序代码，当所述计算机程序代码在计算机上运行时，使得计算机执行上述各方面中的方法。

30 需要说明的是，上述计算机程序代码可以全部或者部分存储在所述第一存储介质上，其中第一存储介质可以与处理器封装在一起的，也可以与处理器单独封装，本申请实施例对此不作具体限定。

第八方面，提供了一种计算机可读介质，所述计算机可读介质存储有程序代码，当所述计算机程序代码在计算机上运行时，使得计算机执行上述各方面中的方法。

35 附图说明

图 1 是本申请实施例适用的胖树网络的网络拓扑的示意图。

图 2 是本申请实施例的确定端口属性的方法的示意性流程图。

图 3 示出了本申请实施例的 PDU 的封装格式的示意图。

图 4 示出了本申请实施例 LLDP 报文的封装格式的示意图。

图 5 示出了本申请实施例的 PDU 的封装格式的示意图。

图 6 示出了本申请实施例的 PDU 的封装格式的示意图。

图 7 示出了本申请实施例的 LLDP 请求的封装格式。

图 8 示出了本申请实施例提供的基于 LLDP 协议确定端口的端口属性的方法

5 图 9 示出了本申请另一实施例提供的基于 LLDP 协议确定端口的端口属性的方法。

图 10 示出了本申请实施例适用的胖树网络的网络拓扑的示意图。

图 11 示出了本申请实施例的防 PFC 死锁规则的部署方法的流程图。

图 12 示出了本申请实施例的确定端口的端口属性的装置的示意性框图。

图 13 示出了本申请另一实施例的确定端口属性的装置的示意性框图。

10 图 14 是本申请另一实施例的网络设备的示意性框图。

具体实施方式

为了便于理解，先介绍本申请实施例涉及的术语。

一、网络设备，即连接到网络中的物理实体。例如，交换机、路由器、以及服务器等。

15 网络设备可以包括多个端口，用于与其他网络设备进行通信。

应理解，上述端口可以是网络设备和网络设备之间级联时使用的端口，例如，可以是交换机和交换机之间级联时使用的端口，还可以是交换机和服务器之间级联使用的端口。

二、端口属性，又称“端口角色”，可以用于描述网络设备的端口在网络拓扑中的角色。具体可以包括上联口（uplink）和下联口（downlink）。

20 上联口，用于指网络设备发送上行数据使用的端口。其中，上行数据可以理解为与从服务器到根节点方向同向传输的数据。

下联口，用于指网络设备发送下行数据使用的端口。其中，下行数据可以理解为与根节点到叶子节点方向同向传输的数据。

三、网络设备的级数，用于指示网络设备在网络（或者说在网络拓扑）中位于哪一层。例如，在数据中心网络中，可以将网络设备到根节点的最短跳数作为该网络设备在网络中的级数。或者，可以将网络设备到服务器的最短跳数作为该网络设备在网络中的级数。

需要说明的是，上述根节点在脊-叶子（spine-leaf）网络架构中，又称为脊节点。

四、协议数据单元（Protocol Data Unit，PDU），是指对等层次之间传递的数据单位。物理层的 PDU 可以指数据位（bit），数据链路层的 PDU 可以指数据帧（frame），网络层的 PDU 可以指数据包（packet），传输层的 PDU 可以指数据段（segment），其他更高层次的 PDU 可以指报文（message）。

下面将结合附图，对本申请中的技术方案进行描述。

为了便于理解，下文先结合图 1 介绍本申请实施例适用的胖树网络。应理解，图 1 所示的网络仅作为本申请实施例适用的一种可能网络架构，本申请实施例还可以适用于数据中心网络的其他网络形式，例如，传统的数据中心三层网络架构等。

图 1 是本申请实施例适用的胖树（fat-tree）网络的网络拓扑的示意图。图 1 所示的胖树网络包括多个根节点 110，多个叶子节点 120，多个服务器 130 以及 SDN 控制器 140。需要说明的是，图 1 中仅示出了 CLOS 中的部分节点，例如，多个叶子节点中的部分叶子节点，即叶子节点 1、叶子节点 2 以及叶子节点 3。多个根节点中的部分根节点，即根节

点 1、根节点 2。本申请实施例对胖树网络中网络设备的具体数量不作限定。

在该胖树网络架构中属于同层的网络设备之间不能直接通信，且每个根节点以全连接的方式连接全部叶子节点。

服务器 130，用于对数据进行数据处理，提供存储服务等。

5 叶子节点 120，用于从服务器中读取数据，或者向服务器中写数据。

根节点 110，用于将帮助同层的叶子节点转发数据流。例如，根节点 1 可以将叶子节点 1 的数据流转发至叶子节点 2。

通常，上述叶子节点或根节点为交换器，或者具有数据转发功能的其他网络设备。

SDN 控制器 140，用于对数据中心网络中的网络设备（例如，交换机）进行管理。

10 随着数据中心网络的规模不断的扩张，SDN 控制器需要管理的网络设备原来越多，SDN 控制器的负载越来越大，SDN 控制器的处理速度会对数据中心的性能产生影响。为了减少 SDN 控制器的负载，希望将原先由 SDN 控制器实现的功能可以转由网络设备自主实现。例如，将 SDN 控制器从生成防优先级流量控制（Priority Flow Control, PFC）死锁规则的功能中释放出来，让网络设备自主生成防 PFC 死锁规则。然而，网络设备自身实
15 现某些功能（例如，生成防 PFC 死锁规则）时，需要获知各个端口的端口为上联口或下联口。

因此，本申请实施例提供了一种确定端口属性的方法，以使网络设备可以获知自身端口的端口属性，为后续帮助 SDN 控制器减轻负载提供基础。下文结合图 2 介绍本申请实施例的确定端口属性的方法。

20 图 2 是本申请实施例的确定端口属性的方法的示意性流程图。图 2 所示的方法包括步骤 200 至步骤 220。需要说明的是，下文中的第一网络设备和第二网络设备可以是图 2 所示的网络架构中的任意两个网络设备。

200，第二网络设备生成 PDU。

上述 PDU 可以用于第一网络设备确定第一网络设备的第一端口的端口属性。

25 210，第二网络设备通过第二端口向第一网络设备发送 PDU，相应地，第一网络设备通过第一端口接收该 PDU。

上述第一端口可以是第一网络设备的多个端口中的任意一个端口。

上述第二端口可以是第二网络设备的多个端口中的任意一个端口。

220，第一网络设备根据 PDU，确定第一端口为上联口或下联口。

30 可选地，上述 PDU 可以携带网络中第二网络设备的级数，或者，PDU 携带第二端口的端口属性，其中，第二端口为第二网络设备发送 PDU 使用的端口。又或者，上述 PDU 可以携带第二网络设备的级数以及第二端口的端口属性。上述 PDU 还可以携带第一 TLV，第一 TLV 用于指示第二网络设备为服务器还是交换机。

第一网络设备基于 PDU 中携带的上述信息，可以确定第一端口的端口属性，即确定
35 第一端口为上联口还是下联口。下文将基于 PDU 中携带的信息，分情况介绍确定第一端口的端口属性的方法。

为了便于理解，在介绍确定端口属性的方式之前，先介绍本申请实施例提供的确定第一网络设备的级数的方式。需要说明的是，下文介绍的确定第一网络设备的级数的方式可以单独使用，以确定第一网络设备的级数；也可以与确定端口属性的方式结合，以辅助第

一网络设备确定第一端口的端口属性，本申请实施例对此不限定。

第一网络设备根据 PDU 确定第一网络设备的级数的方式很多，基于 PDU 中携带的信息不同，主要可以分为以下五种方式。当然，下面的五种级数的确定方式除了与确定端口数据的方式结合，还可以单独使用，本申请对此不限定。

- 5 确定级数的方式一，若 PDU 中携带第二网络设备的级数，则第一网络设备可以根据第二网络设备的级数确定第一网络设备的级数。

例如，以图 1 所示的网络架构为例，从根节点到服务器方向上，不同层级的网络设备的级数递增，且第一网络设备为叶子节点 1，第二网络设备为根节点。预配置网络中的网络设备按照从根节点到服务器的方向，逐层发送 PDU。即根节点先通过所有端口发送
10 PDU，并且在 PDU 中携带根节点的级数 0，叶子节点 1 在接收到 PDU 后，可以以级数 0 为基准加 1，作为叶子节点 1 的级数 1。

又例如，以图 1 所示的网络架构为例，从根节点到服务器方向上，不同层级的网络设备的级数递减，且第一网络设备为叶子节点 1，第二网络设备为根节点。预配置网络中的网络设备按照从根节点到服务器的方向，逐层发送 PDU。即根节点先通过所有端口（即
15 下联口）发送 PDU，并且在 PDU 中携带根节点的级数 max，叶子节点 1 在接收到 PDU 后，可以以级数 max 为基准减 1，作为叶子节点 1 的级数 max-1。

需要说明的是，上述 max 可以是网络中网络架构的最大级数，还可以是预设的大于该最大级数的任意一个值。

- 20 确定级数的方式二，若 PDU 中携带发送该 PDU 的网络设备的级数，则第一网络设备可以基于多个网络设备发送的 PDU，确定第一网络设备的级数，多个网络设备可以包括上文中的第二网络设备。

即，第一网络设备通过第一端口接收多个网络设备发送的多个 PDU，第一网络设备根据多个 PDU，确定在网络中多个网络设备的级数；第一网络设备从多个网络设备的级数中选择目标级数；第一网络设备根据目标级数以及预设规则确定第一网络设备的级数。

- 25 上述预设规则可以包括若目标级数为多个网络设备的级数中的最大级数，则减小目标级数以得到第一网络设备的级数，或者，上述预设规则还可以包括若目标级数为多个网络设备的级数中的最小级数，则增大目标级数以得到第一网络设备的级数。当然，还可以配置其他规则，以确定第一网络设备的级数，本申请实施例对此不限定。

- 30 上述增大或减小目标级数，可以指基于预设的变化量，增大或减少目标级数。例如预设的变化量可以为 1，或者其他值。

- 需要说明的是，上述方法中，选择目标级数的步骤与确定第一网络设备的级数的步骤，还可以替换为，第一网络设备先基于预设规则对多个网络设备的级数进行调整，得到多个调整后的级数，在从多个调整后的级数中选择目标级数。例如，第一网络设备先对多个网络设备的级数都加 1，得到多个调整后的级数，再从多个调整后的级数中选择最小的级数
35 作为目标级数。

可选地，上述多个网络设备是与第一网络设备直接相连的网络设备，也就是说，网络中多个网络设备所在的层级与第一网络设备所在的层级相邻。

需要说明的是，上述多个网络设备所在的层级可以是一个层级也可以是多个层级。当多个网络设备所在的层级可以是一个层级时，即多个网络设备所在的层级相同，且该层级

与第一网络设备的层级相邻。当多个网络设备所在的层级可以是多个层级时，多个网络设备中至少两个网络设备所在的层级不同，且该多个层级与第一网络设备的层级相邻。多个层级可以包括网络中位于第一网络设备所在层级以北的一层，以及第一网络设备所在层级以南的一层。例如，第一网络设备为图 1 中的叶子节点 1 时，多个网络设备可以包括根节点 1，即根节点 1 在网络中位于第一网络设备所在层级以北的一层，多个网络设备可以包括服务器 a，即服务器 a 在网络中位于第一网络设备所在层级以南的一层。

下文以图 1 所示的网络架构为例，介绍本申请实施例确定网络设备的级数的方法。假设从根节点到服务器方向上，不同层级的网络设备的级数递增。通过预配置的方式配置根节点的级数为 0，叶子节点的级数的初始值为 max。且上述预设变化量为 1。每个网络设备以 T 为周期，周期性地通过所有端口发送 PDU，该 PDU 中携带发送该 PDU 的网络设备的级数。

由于通常不知道网络的总层数，上述初始值 max 可以是 PDU 中表示网络设备的级数的 bit 位能表示的最大。例如，PDU 中通过 2bit 表示网络设备的级数时，可以配置初始值 max 为 4。即 2bit 最多可以表示 4 层网络。如果需要表示更多的网络层级，可以扩大表示网络设备的级数的 bit 位的数量。

对于网络中的每个根节点而言，都以周期 T 主动从所有端口发出 PDU，PDU 携带根节点的级数 0。

需要说明的是，由于每个网络设备需要通过所有的端口发送 PDU，因此，对于根节点而言，也会收到其他节点发送的 PDU，此时，可以配置根节点丢弃 PDU。

对于网络中的任意一个叶子节点（例如叶子节点 3）而言，在周期 1 中，会从多个端口接收分别接收多个叶子节点发送的 PDU，每个 PDU 会携带发送该 PDU 的叶子节点的级数。由于网络中叶子节点的级数的初始值为 max，因此，叶子节点 3 获取到的多个叶子节点的级数中可能包括初始值 max。

这样，叶子节点 3 可以从接收到的多个叶子节点的级数中选择最小的级数作为目标级数，并以目标级数为基准增加 1，作为周期 1 内叶子节点 3 的级数。在下一个周期 2 中，叶子节点 3 再将周期 1 确定的级数封装在 PDU 中，发送到其他的叶子节点。经过预设数量的周期后，网络中各叶子节点都可以通过上述方式获知自己在网络中的级数。

可选地，上述多个网络网络设备还可以是与第一网络设备不相邻的网络设备。例如，多个网络设备为根节点，第一网络设备的层次与多个网络设备的层级间隔一层或多层。此时，多个网络设备可以将根节点的级数通过多个 PDU 发送给第一网络设备，多个 PDU 中可以分别记录将对应的 PDU 传输至第一网络设备时进过的网络设备的数量，即传输 PDU 至第一网络设备时所需的跳数，第一网络设备可以基于 PDU 以及跳数确定第一网络设备的级数。

确定级数的方式三，若 PDU 中携带发送该 PDU 的网络设备的级数为预设值，则第一网络设备直接根据预设值确定自身的级数。

对于根节点来说，通常根节点的级数会被配置为预设值，例如上文中介绍的 0。此时，若 PDU 中携带发送该 PDU 的网络设备的级数为预设值，则第一网络设备可以确定自身所在的层级与根节点所在的层级相邻，且第一网络设备的级数为以预设值为基准，调整预设变化量。需要说明的是，上述预设变化量的介绍可以参见上文确定级数的方式二中的介绍。

当然，也可以设置网络中服务器的级数为预设值，则第一网络设备可以确定自身所在的层级与服务器所在的层级相邻，且第一网络设备的级数为以预设值为基准，调整预设变化量。

确定级数的方式四，若上述 PDU 为链路层发现协议（Link Layer Discovery Protocol, LLDP）PDU，则第一网络设备根据所述 LLDP PDU 中端口 ID 的子类型，确定所述第一网络设备的级数。

通常，链路层发现协议中规定，交换机遵循 IETF RFC 2863，使用的端口 ID 的子类型（subtype）为接口名称（Interface name），此时，PDU 中携带的端口 ID 的子类型为 5。服务器遵循 IEEE Std 802，使用的端口 ID 的子类型为媒体访问控制地址（media access control address），此时，PDU 中携带的端口 ID 的子类型为 3。

这样，当第一网络设备收到 LLDP PDU 时，如果 PDU 中携带的端口 ID 的子类型为 3，则第一网络设备可以确定自身所在的层级与服务器所在的层级相邻，则可以将自身的级数记为预设值，例如，0。

确定级数的方式五，若 PDU 携带第一 TLV，且第一 TLV 指示发送 PDU 的网络设备（即，第二网络设备）为服务器，则第一网络设备可以确定自身所在的层级与服务器所在的层级相邻，并可以将自身的级数记为预设值，例如，0。

需要说明的是，上述确定级数的方式三至确定级数的方式五，通常用于网络中特定节点的级数的确定，例如，根节点，或者位于与服务器所在层相邻的层的节点。网络中其他层的网络设备的级数的确定可以通过其他方式确定，例如，确定级数的方式二中介绍的方式。

下文介绍本申请实施例提供的确定端口属性的五种方式，下文确定端口属性的方式一，确定端口属性的方式二以及确定端口属性的方式三中，涉及的确定第一网络设备的级数，和/或确定第二网络设备的级数的方式可以参见上文介绍。当然，第一网络设备的级数，和/或第二网络设备的级数还可以由 SDN 控制器指示给第一网络设备和第二网络设备，本申请实施例对此不作限定。

确定端口属性的方式一，第一网络设备确定第一网络设备的级数，并基于第一网络设备的级数确定第一端口的端口属性。

第一网络设备可以仅基于第一网络设备的级数确定第一端口的端口属性，例如，第一网络设备在确定自身为根节点后，可以直接确定全部端口的端口属性为下联口。又例如，第一网络设备在确定自身所在的层级为与服务器所在的层级相邻后，可以直接确定接收服务器发送的 PDU 的第一端口的下联口。当然，第一网络设备还可以结合第一网络设备的级数以及其他信息确定第一端口的端口属性。具体的确定方式在确定端口的端口属性的方式二中具体介绍。

确定端口的端口属性的方式二、第一网络设备根据第一网络设备的级数，所述第二网络设备的级数确定第一端口的端口属性。其中，第一网络设备的级数的确定可以参见上述确定级数的方式一至确定级数的方式六，为了简洁，在此不再赘述。第二网络设备的级数可以承载于 PDU 中。

当从根节点到服务器方向上，不同层级的网络设备的级数递增时，若第一网络设备的级数小于第二网络设备的级数，则第一网络设备可以确定第一端口为下联口。若第一网络

设备的级数大于第二网络设备的级数，则第一网络设备可以确定第一端口为上联口。

当从根节点到服务器方向上，不同层级的网络设备的级数递减时，若第一网络设备的级数大于第二网络设备的级数，则第一网络设备可以确定第一端口为上联口。若第一网络设备的级数小于第二网络设备的级数，则第一网络设备可以确定第一端口为下联口。

- 5 确定端口的端口属性的方式三、第一网络设备根据第二网络设备的级数确定第一端口的端口属性。

应理解，上述第一网络设备根据第二网络设备的级数确定第一端口的端口属性，可以包括第一网络设备仅根据第二网络设备的级数确定第一端口的端口属性，或者，第一网络设备根据第二网络设备的级数以及其他信息确定第一端口的端口属性。例如，可以根据第
10 二网络设备的级数以及第一网络设备的级数确定第一端口的端口属性，具体的确定方式可以参见确定端口的端口属性的方式二中的介绍，本申请实施例对此不作限定。

例如，基于上文的介绍可知，如果第二网络设备为根节点或服务器，第二网络设备的级数可以被配置为预设值，那么，当第一网络设备确定第二网络设备为预设值后，第二网络设备可以直接将第一端口确定为上联口。

- 15 确定端口的端口属性的方式四、第一网络设备根据第一 TLV 确定第一端口的端口属性。

基于上文的介绍，如果第二网络设备为服务器时，可以在 PDU 中携带上述第一 TLV，第一 TLV 指示所述第二网络设备为服务器。如此，第一网络设备在通过第一端口接收该 PDU 后，可以直接将第一端口确定为下联口。

- 20 确定端口的端口属性的方式五、第一网络设备可以根据 PDU 中端口 ID (port ID) 的子类型，确定第一端口的端口属性。

如上文所述，链路层发现协议中规定，交换机遵循 IETF RFC 2863，使用的端口 ID 的子类型 (subtype) 为接口名称 (Interface name)，此时，PDU 中携带的端口 ID 的子类型为 5。服务器遵循 IEEE Std 802，使用的端口 ID 的子类型为媒体访问控制地址 (media access
25 control address)，此时，PDU 中携带的端口 ID 的子类型为 3。

这样，当第一网络设备收到的 PDU 为 LLDP PDU 时，如果 PDU 中携带的端口 ID 的子类型为 3，则第一网络设备可以确定第一端口为下联口。

- 30 在上文介绍的确定端口的端口属性以及确定网络设备的级数的方法中，网络中的多个网络设备（例如，第二网络设备）可以周期性的自主发送 PDU，以便于网络中的其他网络设备（例如，第一网络设备）可以确定端口属性或级数。当然，第一网络设备还可以借用请求应答的机制，从第二网络设备获取 PDU。这种请求应答的机制可以应用于多种协议，下文仅以应用于 LLDP 协议为例进行介绍。

可选地，上述 PDU 为 LLDP PDU，在步骤 210 之前，所述方法还包括：所述第一网络设备向所述第二网络设备发送 LLDP 请求，所述 LLDP 请求用于请求所述第二网络设备的第二端口的端口属性，和/或所述第二网络设备在网络中的级数。相应地，第二网络设备可以向第一网络设备回复上述 LLDP PDU，即响应 LLDP 请求的过程。
35

当然，第一网络设备可以以周期向第二网络设备发送上述 LLDP 请求，但是为了减少第一网络设备与第二网络设备之间的通信次数，减少发送 LLDP 请求占用的传输资源，第一网络设备还可以按需向第二网络设备发送 LLDP 请求。即，若所述第一端口的端口属性

未知，所述第一网络设备向所述第二网络设备发送所述 LLDP 请求。

如上文所述，端口属性和/或网络设备的级数可以承载于 PDU 中。需要说明的是，关于 PDU 的标准化协议有很多种，使得在不同的标准中封装格式存在一定差异。下文结合图 3 至图 6，主要介绍在慢协议和 LLDP 协议中的几种可能的封装格式。需要说明的是，图 3 至图 6 所示的封装格式中还可以包括通信协议中已经规定好的其他字段，或者未来通信协议中规定的其他字段，本申请实施例对此不作限定。

下文中结合图 3 和图 4 介绍携带网络设备的级数的 PDU 格式或报文格式。结合图 5 和图 6 介绍携带网络设备的级数以及端口的端口属性后的 PDU 格式或报文格式。

图 3 示出了本申请实施例的 PDU 的封装格式的示意图。图 3 所示的 PDU 包括：目的地址（Destination Address，DA）310，源地址（Source Address，SA）320，类型（Type）330，子类型（Subtype）340，标志（Flag）域 350，代码（Code）360，数据（Data）域 370，帧检验序列（Frame Check Sequence，FCS）380。

DA 310，表示 PDU 的目的 MAC 地址。通常，为组播 MAC 地址 0180c2000002。

SA 320，表示 PDU 的源 MAC 地址。通常，为一个单播 MAC 地址。

类型 330，表示 PDU 的协议类型，通常为 0x8809。

子类型 340，表示 PDU 的协议子类型，通常为 0x03。

标志域 350：包含了网络中网络设备的状态信息。

代码 360，指示 PDU 的类型。例如，该字段的取值为 0x00，表示 PDU 为消息（Information）PDU；该字段的取值为 0x01，表示 PDU 为事件通知（Event Notification）PDU；该字段的取值为 0x04，表示 PDU 为环路控制（Loopback Control）PDU。

数据域 370，用于承载 OAM PDU 的有效载荷。

FCS 380，为 PDU 的尾部字段，俗称为“帧尾”，用于验证 PDU 在传输过程中的完整性。通常为一段 4 字节的循环冗余校验码。

本申请实施例中，表示网络设备的级数的 TLV（又称“第二 TLV”）可以承载于图 3 所示的 PDU 中的 FCS 380 之前的任意字段。例如，如图 3 所示的，第二 TLV 390 可以占用 FCS 380 之前的字段，且与 FCS 380 相邻。

需要说明的是，上述图 3 所示的 PDU 的封装格式还可以仅携带端口的端口属性（即第三 TLV），即将上文中的“第二 TLV”替换为“第三 TLV”即可，为了简洁，在此不再赘述。

图 4 示出了本申请实施例 LLDP 报文的封装格式的示意图。图 4 所示的 LLDP 报文的封装格式包括：DA 410，SA 420，类型（Type）430，LLDP PDU 440，FCS 440。

其中，DA 410，用于表示 LLDP 报文的接收端的媒体访问控制（Media Access Control，MAC）地址，通常可以设定为组播 MAC 地址，例如，0x0180-C200-000E。

SA 420，用于表示发送 LLDP 报文的发送端的 MAC 地址。

类型 430，用于指示 LLDP 报文的类型。通常 LLDP 报文中，该字段的取值为 0x88CC。

LLDP PDU 440，用于承载信息或数据等，可以理解为 LLDP 报文中进行信息交换的主体。通常，LLDP PDU 采用类型长度值（type-length-value，TLV）的格式封装数据或信息。具体可以包括：底盘 ID（Chassis ID TLV）441、端口 ID TLV（Port ID TLV）442、生存时间（Time to Live，TTL）TLV443 以及 LLDP PDU 的结束（End of LLDP PDU）TLV

444。其中，Chassis ID TLV 441，用于表示发送端的底盘 MAC 地址。端口 ID TLV 442，用于表示 PDU 的发送端的端口。TTL TLV 443，用于限制 PDU 在网络中存在的时间，防止 PDU 环路。LLDP PDU 的结束 444，用于标识 LLDP PDU 的结束。

为提高协议的可扩展性，LLDP PDU 440 中还预留了多个保留 TLV (reserved TLV)
5 449，例如，保留 TLV #1 至保留 TLV #n，其中 n 为正整数。通常，保留 TLV 对应的 TLV 类型为 9 至 126。在本申请实施例中，表示网络设备的级数的 TLV (即“第二 TLV”) 可以承载于上述多个保留 TLV 中的任意一个保留 TLV 中，例如，可以占用保留 TLV #1 445。具体地，第二 TLV 可以包括第二 TLV 的类型 446，第二 TLV 的长度 (TLV length) 447 以及第二 TLV 的值 (value) 448。其中，第二 TLV 的类型 446 可以是保留 TLV #1 对应
10 的 TLV 类型，即为 9。第二 TLV 的长度 447，用于指示第二 TLV 的值中承载的数据的长度。第二 TLV 的值 448，表示实际的数值，即网络设备的级数。

FCS 440，LLDP 报文的尾部字段，俗称为“帧尾”，用于验证 LLDP 报文在传输过程中的完整性。通常为一段 4 字节的循环冗余校验码。

可选地，上文仅列出了 LLDP PDU 440 中包含的部分 TLV。目前，LLDP 中规定 LLDP
15 PDU 中的 TLV 可以分为两种类型，即强制 TLV 和可选 TLV。其中，强制 TLV 主要包括底盘 ID (Chassis ID TLV) 441、端口 ID TLV (Port ID TLV) 442、生存时间 (Time to Live, TTL) TLV 443 等。可选 TLV 主要包括端口描述 (port description) TLV，系统名称 (system name) TLV，系统描述 (system description) TLV，系统能力 (system capabilities) TLV 等。为了减少对 LLDP PDU 中保留 TLV 的改动，本申请实施例还提供了一种携带网络设备的级数的 TLV (即，第二 TLV) 的方式，即配置第二 TLV 占用上述任一种可选 TLV
20 对应的字段。

需要说明的是，上述图 4 所示的 LLDP 报文的封装格式还可以仅携带端口的端口属性 (即第三 TLV)，即将上文中的“第二 TLV”替换为“第三 TLV”即可，为了简洁，在此不再赘述。

25 图 5 示出了本申请实施例的 PDU 的封装格式的示意图。需要说明的是，图 5 所示的 PDU 的封装格式中与图 3 所示的 PDU 的封装格式中含义相同的字段采用相同的编号，具体的含义可以参见上文的介绍，为了简洁，在此不再赘述。

图 5 所示的 PDU 包括：目的地址 (Destination Address, DA) 310，源地址 (Source Address, SA) 320，类型 (Type) 330，子类型 (Subtype) 330，标志 (Flag) 域 350，代码 (Code) 360，数据 (Data) 370，帧检验序列 (Frame Check Sequence, FCS) 380。本
30 申请实施例中，表示网络设备的级数的 TLV (又称“第二 TLV”)，以及表示端口的端口属性的 TLV (又称“第三 TLV”) 可以承载于图 5 所示的 PDU 中的 FCS 380 之前的任意字段。例如，如图 5 所示的，第二 TLV 510 以及第三 TLV 520 可以占用 FCS 380 之前的字段，且与 FCS 380 相邻。

35 需要说明的是，上述第二 TLV 510 与第三 TLV 520 之间的位置可以调换，即第三 TLV 可以位于第二 TLV 之前的字段。上述第二 TLV 和第三 TLV 之间可以不相邻也可以相邻。本申请实施例对此不作限定。

图 6 示出了本申请实施例的 PDU 的封装格式的示意图。需要说明的是，图 6 所示的 PDU 的封装格式中与图 4 所示的 PDU 的封装格式中含义相同的字段采用相同的编号，具

体的含义可以参见上文的介绍，为了简洁，在此不再赘述。

图 6 所示的 LLDP 报文的封装格式包括：DA 410，SA 420，类型（Type）430，LLDP PDU 440，FCS 440。其中，LLDP PDU 440 中预留有多个保留 TLV（reserved TLV）449，例如，保留 TLV #1 至保留 TLV #n，其中 n 为正整数。通常，保留 TLV 对应的 TLV 类型为 9 至 126。

在本申请实施例中，表示网络设备的级数的 TLV（即“第二 TLV”）可以承载于上述多个保留 TLV 中的任意一个保留 TLV 中，例如，可以占用保留 TLV #1 445。具体地，第二 TLV 可以包括第二 TLV 的类型 446，第二 TLV 的长度（TLV length）447 以及第二 TLV 的值（value）448。其中，第二 TLV 的类型 446 可以是保留 TLV #1 对应的 TLV 类型，即为 9。第二 TLV 的长度 447，用于指示第二 TLV 的值中承载的数据的长度。第二 TLV 的值 448，表示实际的数值，即网络设备的级数。

表示端口的端口属性的 TLV（即“第三 TLV”）可以承载于上述多个保留 TLV 中的任意一个保留 TLV 中，例如，可以占用保留 TLV #2 610。具体地，第三 TLV 可以包括 TLV 类型 611，TLV 长度（TLV length）612 以及 TLV 的值（value）613。其中，第三 TLV 的类型 612 可以是保留 TLV #2 对应的 TLV 类型，即 10。第三 TLV 的长度 612，用于指示第三 TLV 的值中承载的数据的长度。第三 TLV 的值 613，表示实际的数值，即端口的端口属性。

可选地，上文仅列出了 LLDP PDU 440 中包含的部分 TLV。目前，LLDP 中规定 LLDP PDU 中的 TLV 可以分为两种类型，即强制 TLV 和可选 TLV。其中，强制 TLV 主要包括底盘 ID（Chassis ID）TLV 441、端口 ID TLV（Port ID TLV）442、生存时间（Time to Live，TTL）TLV 443 等。可选 TLV 主要包括端口描述（port description）TLV，系统名称（system name）TLV，系统描述（system description）TLV，系统能力（system capabilities）TLV 等。为了减少对 LLDP PDU 中保留 TLV 的改动，本申请实施例还提供了一种携带第二 TLV 以及第三 TLV 的方式，即配置第二 TLV 以及第三 TLV 占用上述任两种可选 TLV 对应的字段。

下文结合图 7 介绍本申请实施例的 LLDP 请求的封装格式。图 7 所示的 LLDP 请求包括：DA 710，SA 720，类型（Type）730，LLDP PDU 740，FCS 770。

其中，DA 710，用于表示 LLDP 请求的接收端的媒体访问控制（Media Access Control，MAC）地址，通常可以设定为组播 MAC 地址，例如，0x0180-C200-000E。

SA 720，用于表示发送 LLDP 请求的发送端的 MAC 地址。

类型 730，用于指示 LLDP 请求的类型。

LLDP PDU 740，用于承载信息或数据等，可以理解为 LLDP 信息交换的主体。通常，LLDP PDU 采用类型长度值（type-length-value，TLV）的格式封装数据或信息。具体可以包括：底盘 ID（Chassis ID）TLV 741、端口 ID TLV（Port ID TLV）742、生存时间（Time to Live，TTL）TLV 743 以及 LLDP PDU 的结束（End of LLDP PDU）TLV 744。其中，底盘 ID TLV 741，用于表示发送端的底盘 MAC 地址。端口 ID TLV 742，用于表示 LLDP 请求的发送端的端口。TTL TLV 743，用于限制 LLDP 请求在网络中存在的时间，防止 LLDP 请求环路。End of LLDP PDU 744，用于标识 LLDP PDU 的结束。

为提高协议的可扩展性，LLDP PDU 740 中还预留了多个保留 TLV（reserved TLV）

749, 例如, 保留 TLV #1 至保留 TLV #n, 其中 n 为正整数。通常, 保留 TLV 对应的 TLV 类型为 9 至 126。在本申请实施例中, 用于请求网络设备的级数的 TLV 和/或用于请求端口的端口属性的 TLV (统称为“请求 TLV”) 可以承载于上述多个保留 TLV 中的任意一个保留 TLV 中。

5 例如, 请求 TLV 仅可以仅占用保留 TLV #1 745。具体地, 请求 TLV 可以包括请求 TLV 的类型 746, 请求 TLV 的长度 (TLV length) 747 以及请求 TLV 的值 (value) 748。其中, 请求 TLV 的类型 746 可以是保留 TLV #1 对应的 TLV 类型, 即为 9。请求 TLV 的长度 747, 用于指示请求 TLV 的值中承载的数据的长度。请求 TLV 的值 748, 表示实际的数值, 即网络设备的级数和/或端口的端口属性。

10 FCS 740, LLDP 报文的尾部字段, 俗称为“帧尾”, 用于验证 LLDP 报文在传输过程中的完整性。通常为一段 4 字节的循环冗余校验码。

可选地, 上文仅列出了 LLDP PDU 740 中包含的部分 TLV。目前, LLDP 中规定 LLDP PDU 中的 TLV 可以分为两种类型, 即强制 TLV 和可选 TLV。其中, 强制 TLV 主要包括底盘 ID (Chassis ID TLV) 741、端口 ID TLV (Port ID TLV) 742、生存时间 (Time to Live, 15 TTL) TLV 743 等。可选 TLV 主要包括端口描述 (port description) TLV, 系统名称 (system name) TLV, 系统描述 (system description) TLV, 系统能力 (system capabilities) TLV 等。为了减少对 LLDP PDU 中保留 TLV 的改动, 本申请实施例还提供了一种携带请求 TLV 的方式, 即配置请求 TLV 占用上述任一种可选 TLV 对应的字段。

需要说明的是, 上文中的 LLDP 请求的封装格式与 LLDP 报文的封装格式对应。例如, 20 当上述 LLDP 请求用于请求网络设备的级数时, LLDP 报文可以采用图 4 所示的封装格式。当上述 LLDP 请求用于请求网络设备的级数以及端口的端口类型时, LLDP 报文可以采用图 5 所示的封装格式。

下文以图 1 所示的网络架构为例, 结合具体的例子介绍本申请实施例的确定端口的端口属性的方法。需要说明的是, 下文中的具体例子仅仅为了便于理解本申请, 并不会限定 25 本申请的范围。

图 7 是本申请实施例的确定端口的端口属性的方法的示意性流程图。图 7 所示的方法包括步骤 710 至步骤 750。假设从根节点到服务器方向上, 不同层级的网络设备的级数以步长 1 递增。通过预配置的方式配置根节点的级数为 0, 叶子节点的级数的初始值为 max, 以及预配置网络中每个节点 (包括根节点以及叶子节点) 的端口的初始端口属性为下联口。 30 每个网络设备以 T 为周期, 周期性地通过所有下联口发送 PDU, 该 PDU 中携带发送该 PDU 的网络设备的级数。

由于通常不知道网络的总层数, 上述初始值 max 可以是 PDU 中表示网络设备的级数的 bit 位能表示的最大。例如, PDU 中通过 2bit 表示网络设备的级数时, 可以配置初始值 max 为 4。即 2bit 最多可以表示 4 层网络。如果需要表示更多的网络层级, 可以扩大表示 35 网络设备的级数的 bit 位的数量。

对于网络中的每个根节点而言, 都以周期 T 主动从所有端口发出 PDU, PDU 携带根节点的级数 0。需要说明的是, 对于根节点而言, 所有的端口都是下联口。

还应理解, 由于每个网络设备 (包括根节点和叶子节点) 都需要通过所有的端口发送 PDU, 因此, 对于根节点而言, 也会收到叶子节点发送的 PDU, 此时, 可以配置根节点

丢弃其他节点发送的 PDU。

对于网络中的任意一个叶子节点而言，叶子节点中配置有变量 switch_level，switch_level 用于存放叶子节点的级数，其初始值为 max。叶子节点中还为每个端口配置有变量 port_level，port_level 用于存储根据从端口接收的 PDU 中携带的网络设备的级数，
5 确定的本节点的级数，其初始值也为 max。

下文以叶子节点 3 为例介绍确定端口的端口属性的方法。网络中的其他叶子节点确定端口的端口属性的方法与该方法类似，为了简洁，不再赘述。假设叶子节点 3 有三个端口分别用 P_1 、 P_2 和 P_3 表示，则 P_1 对应的变量为 port_level₁， P_2 对应的变量为 port_level₂， P_1 对应的变量为 port_level₃。

10 710，在周期 1 中，叶子节点 3 通过 P_1 、 P_2 和 P_3 分别接收 PDU₁，PDU₂ 和 PDU₃，其中，PDU₁ 中携带的网络设备的级数为 X_1 ，PDU₂ 中携带的网络设备的级数为 X_2 ，PDU₃ 中携带的网络设备的级数为 X_3 。

720，叶子节点 3 将 X_1 保存在 port_level₁ 中， X_2 保存在 port_level₂ 中， X_3 保存在 port_level₃ 中。

15 需要说明的是，若上述周期 1 为确定端口的端口属性的初始周期，且叶子节点的级数的初始值为 max，那么上述 X_1 ， X_2 ， X_3 中表示叶子节点的级数的取值为 max，表示叶子节点的级数的取值为 0。

730，叶子节点 3 从 X_1 ， X_2 ， X_3 中，选择最小值 $X_{\min}$ ，并将 $X_{\min}+1$ 存储至变量 switch_level。

20 需要说明的是，若 X_1 ， X_2 ， X_3 相等，则 $X_{\min}$ 等于 X_1 ， X_2 ， X_3 中任一值。

740，叶子节点 3 根据 X_i 以及 $X_{\min}+1$ ，确定 P_1 、 P_2 和 P_3 的端口属性，其中 $i=1,2,3$ 。

具体地，若 $X_i < X_{\min}+1$ ，则可以确定 X_i 对应的叶子节点 3 的端口为上联口；若 $X_i > X_{\min}+1$ ，则可以确定 X_i 对应的叶子节点 3 的端口为下联口。

25 750，在周期 2 中，将 $X_{\min}+1$ 封装在 PDU 中作为叶子节点 3 的级数，并通过全部的下联口发送携带 $X_{\min}+1$ 的 PDU。

需要说明的是，上述周期 2 是周期 1 的下一个周期。叶子节点 3 在周期 2 中依然会收到其他节点通过下联口发送的 PDU。叶子节点 3 可以重新执行上述步骤 710 至步骤 750，确定各端口的端口属性。

30 经过预设的多个周期，网络中每个叶子节点的端口的端口属性趋于稳定，则可以停止发送 PDU，此时，每个叶子节点确定的端口的端口属性较为准确。

下文以图 1 所示的网络架构为例，结合具体的例子，介绍本申请实施例提供的基于 LLDP 协议确定端口的端口属性的方法。图 8 所示的方法包括步骤 810 至步骤 830。假设从根节点到服务器方向上，不同层级的节点的级数以步长 1 递减，且预先预定若网络设备与服务器相邻则该网络设备的级数为 1。下文以叶子节点 3 发送 LLDP 请求为例，进行说明。
35 假设叶子节点 3 的端口 P_1 的端口属性未知。

810，叶子节点 3 通过端口 P_1 向相邻的节点发送 LLDP 请求，以请求相邻的节点的级数以及端口属性。

当相邻节点为叶子节点时，叶子节点 2 通过端口 Q_1 接收该第一 LLDP 请求，第一 LLDP 请求用于请求叶子节点 2 的级数以及第二端口的端口属性。

若叶子节点 2 的级数以及端口 Q_1 的端口属性已知, 则叶子节点 2 通过端口 Q_1 向叶子节点 3 发送第一 LLDP 报文, 第一 LLDP 报文携带叶子节点 2 的级数, 以及端口 Q_1 的端口属性。若叶子节点 2 的级数以及端口 Q_1 的端口属性未知, 则叶子节点 2 被配置忽略上述第一 LLDP 请求。

- 5 当相邻节点为服务器时, 叶子节点 3 通过 P_2 向服务器 1 发送第二 LLDP 请求, 用于请求服务器 1 的级数以及服务器 1 的端口的端口属性。相应地, 服务器向叶子节点 2 反馈第二 LLDP 报文。

820, 叶子节点 3 通过端口 P_1 接收第一 LLDP 报文, 或第二 LLDP 报文。

- 10 830, 叶子节点 3 根据接收的 LLDP 报文, 确定叶子节点 3 的级数以及发送 LLDP 报文的端口的端口属性。

上述 LLDP 报文为第一 LLDP 报文, 若叶子节点 3 从第一 LLDP 报文中获取叶子节点 2 的级数为 M , 且端口 Q_1 的端口属性为下联口, 则叶子节点 3 可以确定叶子节点 3 的级数为 $M+1$, 且端口 P_1 为上联口, 其中, M 为正整数。

- 15 若叶子节点 3 从第一 LLDP 报文中获取叶子节点 2 的级数为 Z , 且端口 Q_1 的端口属性为上联口, 则叶子节点 3 可以确定叶子节点 3 的级数为 $Z-1$, 且端口 P_1 为下联口, 其中, Z 为大于或等于 1 的正整数。

上述 LLDP 报文为第二 LLDP 报文时, 叶子节点 3 可以根据第二 LLDP 报文中端口 ID 的子类型确定相邻的节点为服务器, 此时, 叶子节点 3 可以确定端口 P_1 为下联口, 且叶子节点 3 的级数为 1。

- 20 需要说明的是, 如上文所述, LLDP 协议中规定, 服务器适用的端口 ID 的子类型通常为 MAC 地址, 即第二 LLDP 报文中携带的端口 ID 的子类型的取值为 5。与交换机使用的端口 ID 的子类型不同。因此, 叶子节点 3 可以基于第二 LLDP 报文中的端口 ID 的子类型, 直接确定第二 LLDP 报文的发送端为服务器。

- 25 还应理解, 相邻的节点为服务器时, 由于叶子节点 3 可以直接基于端口 ID 的子类型, 确定叶子节点 3 的级数以及端口 P_1 的端口属性, 因此, 上述第二 LLDP 报文中可以无需携带服务器的级别 0 以及服务器发送第二 LLDP 报文使用的端口的端口属性。当然, 为了提高叶子节点 3 确定级数以及端口的端口属性的准确率, 也可以在第二 LLDP 报文中携带服务器的级数, 以及发送第二 LLDP 报文使用的端口的端口属性。

- 30 需要说明的是, 在上述步骤 810 至步骤 830 描述的过程之前, 叶子节点 3 可能会获得多个叶子节点 3 的级数。例如, 叶子节点 3 在确定端口 P_1 的端口属性之前, 已经获知了叶子节点 3 的级数为 m , m 为正整数, 且与叶子节点 3 在上述步骤 830 中确定的级数不同。又或者, 叶子节点 3 有多个端口的端口属性未知, 则叶子节点 3 可能会从多个相邻节点处获取多个不同的相邻节点级数, 导致叶子节点 3 确定了不同的多个叶子节点 3 的级数。在这种情况下, 叶子节点 3 可以从多个叶子节点 3 的级数选择最大的级数, 作为叶子节点 3 的级数。

下文以图 1 所示的网络架构为例, 结合具体的例子, 介绍本申请另一实施例提供的基于 LLDP 协议确定端口的端口属性的方法。图 9 所示的方法包括步骤 910 至步骤 920。假设从根节点到服务器方向上, 不同层级的节点的级数以步长 1 递减, 且预先预定若网络设备与服务器相邻则该网络设备的级数为 1。在预设时间段内, 图 9 所述的网络中网络设备

通过全部的端口向相邻的网络设备发送 LLDP 报文。

下文以叶子节点 3 为例介绍确定端口的端口属性的方法。网络中的其他叶子节点确定端口的端口属性的方法与该方法类似，为了简洁，不再赘述。假设叶子节点 3 有三个端口分别用 P_1 、 P_2 和 P_3 表示。

5 910，叶子节点 3 通过端口 P_1 接收相邻节点发送的 LLDP 报文。

若 LLDP 报文中携带的端口 ID 的子类型为服务器使用的端口 ID 的子类型，叶子节点 3 可以直接确定端口 P_1 为下联口，且叶子节点 3 的级数为 1。

需要说明的是，关于上述端口 ID 的子类型的介绍可以参见上文描述，为了简洁，在此不再赘述。

10 若 LLDP 报文中携带相邻节点的级数 A' 以及端口的端口属性为下联口，且叶子节点 3 记录的叶子节点 3 的级数为 B' ，其中， A' 为大于 1 的正整数， B' 为正整数。当 $A'-1 < B'$ 时，叶子节点 3 将 $A'-1$ 更新为叶子节点 3 的级数，并将端口 P_1 标记为上联口。当 $A'-1 > B'$ 时，叶子节点 3 不更新叶子节点 3 的级数，并将端口 P_1 标记为上联口。

15 当然，在上述情况中，若叶子节点 3 未记录的叶子节点 3 的级数，则叶子节点 3 可以直接将 $A'-1$ 作为叶子节点 3 的级数。

若 LLDP 报文中携带相邻节点的级数 A 以及端口的端口属性为上联口，且叶子节点 3 记录的叶子节点 3 的级数为 B ，其中， A 为大于 1 的正整数， B 为正整数。当 $A+1 < B$ 时，叶子节点 3 将 $A+1$ 更新为叶子节点 3 的级数，并将端口 P_1 标记为下联口。当 $A+1 > B$ 时，叶子节点 3 不更新叶子节点 3 的级数，并将端口 P_1 标记为下联口。

20 当然，在上述情况中，若叶子节点 3 未记录的叶子节点 3 的级数，则叶子节点 3 可以直接将 $A+1$ 作为叶子节点 3 的级数。

920，叶子节点 3 通过各个端口向相邻节点发送 LLDP 报文，该报文中携带叶子节点 3 的级数，以及发送 LLDP 报文使用的端口的端口属性。

25 上文中介绍的确定网络设备的级数，以及确定端口的端口属性的方法可以应用在各种场景中。下文以将确定端口的端口属性的方法应用于部署防 PFC 死锁规则的场景为例，介绍本申请实施例的方法。为了便于理解，下文先基于图 1 所示的网络拓扑，先介绍优先级流量控制（Priority-based flow control, PFC）策略以及 PFC 死锁的形成。

30 在图 1 所示的网络中，可以采用上述 PFC 策略传输数据流。PFC 一种基于优先级的流量控制策略。PFC 允许在一条以太网链路上创建 8 个虚拟通道，并为每条虚拟通道指定一个优先级，允许单独暂停和重启其中任意一条虚拟通道，同时允许其它虚拟通道的流量无中断通过。这一方法使网络能够为单个虚拟通道创建无丢包类别的服务，使其能够与同一接口上的其它流量共存。

35 下文结合图 1 所示数据流 1、数据流 2 以及数据流 3 的传输路径为例，介绍 PFC 死锁的产生。其中，数据流 1、数据流 2 以及数据流 3 的优先级相同，且在网络中通过队列 5 组成的虚拟通道进行传输。

在图 1 所示的网络中，每个根节点以全连接的方式连接至所有的叶子节点，且同层的网络设备之间不能直接通信，因此，网络中无链路故障时，数据流的流向通常是仅上行的方向，或者仅下行的方向。又或者是由上行（up）到下行（down）的方向，例如，图 1 中数据流 2 的传输路径以及数据流 3 的传输路径。其中，数据流 2 为需要从服务器 c 传输

至服务器 d 的数据流, 数据流 2 的传输路径为服务器 c-叶子节点 2-根节点 1-叶子节点 3-服务器 d。数据流 3 为需要从服务器 e 传输至服务器 b 的数据流, 数据流 3 的传输路径为服务器 e-叶子节点 3-根节点 2-叶子节点 2-服务器 c。

然而, 当不同层的两个网络设备之间的链路故障后, 网络中会产生由下行到上行流向的数据流。例如, 根节点 2 与叶子节点 4 之间的链路故障后, 需要从服务器 a 传输至服务器 f 的数据流 1 只能沿着图 1 所示的数据流 1 的传输路径进行传输, 即服务器 a-叶子节点 1-根节点 1-叶子节点 3-根节点 2-叶子节点 4-服务器 f。其中, 数据流 1 在根节点 1 到叶子节点 3 之间为下行方向的数据流, 数据流 1 在叶子节点 3 到根节点 2 之间为上行方向的数据流。

一旦网络中产生上述下行-上行流向的数据流后, 该流向的数据流就可能与网络中相同优先级的其他通过正常流向传输数据流形成一个环路, 该环路的形状类型数字 8。例如, 图 1 中数据流 1、数据流 2 以及数据流 3 在根节点 1、根节点 2、叶子节点 2 以及叶子节点 3 之间形成的环路。

为了便于说明 PFC 死锁, 图 10 仅示出了图 1 中的环路涉及的 4 个节点以及组成环路的 3 个数据流。在图 10 所示的环路中, 当叶子节点 3 中队列 5 对应的缓存的容量的消耗超过阈值, 即出现瞬时拥塞, 则叶子节点 3 即向数据进入的方向发送反压信息 (又称“PFC 暂停帧”), 即叶子节点 3 沿着数据流 1 和数据流 2 流进的方向, 将反压信息发送至根节点 1。相应地, 根节点 1 接收到反压信息后, 会根据反压信息指示停止发送或延迟发送数据流 1 以及数据流 2 剩下的数据, 并将剩下的数据缓存在本端队列 5 对应的缓存中。如果根节点 1 中队列 5 对应的缓存的消耗超过阈值, 则沿着数据流 2 流进的方向, 再将反压信息发送至叶子节点 2。相应地, 叶子节点 2 接收到反压信息后, 会根据反压信息指示停止发送或延迟发送数据流 2 剩下的数据, 并将剩下的数据缓存在本端队列 5 对应的缓存中。如果叶子节点 2 中队列 5 对应的缓存的消耗超过阈值, 则继续沿着数据流 3 流进的方向, 再将反压信息发送至根节点 2。相应地, 根节点 2 接收到反压信息后, 会根据反压信息指示停止发送或延迟发送数据流 3 剩下的数据, 并将剩下的数据缓存在本端队列 5 对应的缓存中。如果根节点 2 的队列 5 对应的缓存的消耗超过阈值, 则继续沿着数据流 3 流进的方向, 再将反压信息发送至叶子节点 3。如此一级级反压, 最终导致根节点 1、根节点 2、叶子节点 2 以及叶子节点 3 的队列 5 中都没有数据流进行传输, 也就是说, 上述 4 个节点之间队列 5 中缓存的流量不能发送, 使得 4 个节点的缓存容量的消耗长久处于超过阈值的状态, 即产生 PFC 死锁。

综上, 可以看出, PFC 死锁产生的根本原因是由于网络中产生了下行-上行流向的数据流。并且该下行-上行流向的数据流在一个网络设备被传输时, 使用的是该网络设备中的两个上联口。例如, 数据流 1 在流经叶子节点 3 时, 入端口 (inport) P_r , 以及出端口 (outport) P_{r+1} 的端口属性都为上联口。因此, 为了避免网络中产生下行-上行流向的数据流, 从而预防 PFC 死锁产生, 可以在网络中的每个叶子节点部署防 PFC 死锁规则。例如, 可以在叶子节点 3 中部署防 PFC 死锁规则, 该防 PFC 死锁规则可以包括: 若 {inport = P_r , output = P_{r+1} , queue = 5}, 动作: 切换到队列 6, 且将数据流的优先级字段修改为队列 6 对应的优先级。该防 PFC 死锁规则表示, 当数据流的入端口为 P_r , 出端口为 P_{r+1} , 且对应的队列为队列 5 时, 叶子节点 3 需要将该数据流直接缓存在队列 6 中, 并将该数据流的优先级从队列 5 对

应的优先级修改为队列 6 对应的优先级。即对流经一个网络设备的两个上联口的流量，执行切换优先级的操作，以防止产生 PFC 死锁。

如上文所述，下行-上行流向的数据流产生的原因通常是由于某一链路故障导致的，也就是说，下行-上行流向的数据流的产生具有一定的随机性。而这种随机性是不可预测的，因此，为了避免下行-上行流向的数据流产生，通常以预先部署的方式，在网络中有数据流被传输之前，由 SDN 控制器（例如，SDN 控制器 140）为网络中的每个网络设备的每两个上联口都部署防 PFC 死锁规则。如此一来，可能导致 SDN 控制器的负载过大。

为了减少 SDN 控制器部署防 PFC 死锁规则产生的压力，本申请还提供了一种新的防 PFC 死锁规则的部署方式，即由网络设备自己基于本端的上联口，自主生成防 PFC 死锁规则。下文结合图 11 介绍本申请实施例的防 PFC 死锁规则的部署方法的流程图。图 11 所示的方法可以由第一网络设备执行，第一网络设备可以是图 1 所示的网络设备中的任意一个叶子节点，第一网络设备包括多个发送队列，多个发送队列包括第一发送队列和第二发送队列，所述第一发送队列为第一数据流对应的原始发送队列。图 11 包括步骤 1110 至步骤 1120。

1110，第一网络设备确定第一端口以及第一网络设备的第三端口为上联口。

需要说明的是，第一网络设备确定第一端口以及第三端口的方法可以参见上文中关于确定端口的端口属性的方法，为了简洁，在此不再赘述。当然，SDN 控制器也可以向第一网络设备直接指示上述第一端口以及第三端口为上联口。

1120，第一网络设备生成防优先级流量控制 PFC 死锁规则，防 PFC 死锁规则用于指示将所述第一数据流缓存至第二发送队列，第一数据流的入端口为第一端口，第一数据流的出端口为所述第三端口。

可选地，上述第一发送队列对应的优先级与第二发送队列对应的优先级不同，但是在最终进行数据调度的过程中，可以将第一发送队列与第二发送队列视为对应相同优先级的队列，并按照预设的比例进行数据调度，以避免由切换队列导致的数据流乱序传输。

可选地，对于相同的一组上联口可以部署一条或多条防 PFC 死锁规则。

例如，可以针对两个上联口的不同数据流的流向部署多条防 PFC 死锁规则。对于上述叶子节点 2 中的上联口 P_r 以及上联口 P_{r+1} 可以部署的防 PFC 死锁规则包括：若 $\{\text{inport} = P_r, \text{output} = P_{r+1}, \text{queue} = 5\}$ ，动作：切换到队列 6，且将数据流的优先级字段修改为队列 6 对应的优先级。以及若 $\{\text{inport} = P_{r+1}, \text{output} = P_r, \text{queue} = 5\}$ ，动作：切换到队列 6，且将数据流的优先级字段修改为队列 6 对应的优先级。

又例如，可以针对不同的传输方式为两个上联口部署多条防 PFC 死锁规则。对于上述叶子节点 2 中的上联口 P_r 以及上联口 P_{r+1} 可以部署的防 PFC 死锁规则包括：若 $\{\text{inport} = P_r, \text{output} = P_{r+1}, \text{queue} = 5\}$ ，动作：切换到队列 6，且将数据流的优先级字段修改为队列 6 对应的优先级。以及若 $\{\text{inport} = P_{r+1}, \text{output} = P_r, \text{queue} = 5\}$ ，动作：切换到队列 3，且将数据流的优先级字段修改为队列 3 对应的优先级，其中队列 5、队列 6 为无损队列，即用于传输无损流量。队列 3 为有损队列，即用于传输有损流量。

需要说明的是，上述流量可以基于预设的规则，确定切换无损流量传输模式还是有损流量传输模式，从而确定基于上述防 PFC 死锁规则确定切换至队列 6 还是队列 3。

可选地，上述步骤 1110 以及步骤 1120 可以是由于网络中网络拓扑结构发生变化而触

发的，例如，新增部分网络设备，或者减少部分网络设备等。还可以是网络拓扑建立完成后，在传输数据流之前触发的。本申请实施例对此不作限定。

可选地，所述方法还包括：所述第一网络设备确定在预设时间段内，所述第一端口与
5 所述第三端口为 PFC 死锁关联的端口的次数；上述步骤 1120，包括：当所述次数高于预设
阈值，所述第一网络设备确定生成所述防 PFC 死锁规则。

在本申请实施例中，根据次数确定生成防 PFC 死锁规则的上联口，无需为每组上联
口都生成防 PFC 死锁规则，有利于减少网络设备生成防 PFC 死锁规则的数量，减少防 PFC
死锁规则占用的存储空间。

可选地，所述方法还包括：所述第一网络设备通过所述第一端口接收所述第二网络设
10 备发送的指示信息，所述指示信息用于指示所述第二网络设备中的第三发送队列对应的缓
存的容量超过阈值，且所述第三发送队列的入流量为 0，所述第三发送队列缓存的数据流
的优先级与所述第一数据流的优先级相同；所述第一网络设备根据指示信息确定所述次
数。

需要说明的是，上述根据指示信息确定所述次数的方式有很多种，例如指示信息可以
15 指示在预设时间段内第三发送队列容量超过阈值且入流量为 0 产生的总次数。还可以当每
次第三发送队列容量超过阈值且入流量为 0 时，第二网络设备都向第一网络设备发送指示
信息，由第一网络设备统计预设时间段内，第三发送队列容量超过阈值且入流量为 0 产生
的总次数。

例如，第一网络设备的控制面可以维护一张链表，其中一个表项为<第一端口，第三
20 端口，优先级 1，Z>，表示预设时间段内，第一端口作为入端口，第三端口作为出端口，
且优先级 1 对应的次数为 Z。当次数 Z 大于预设阈值时，第一网络设备为第一端口，第三
端口生成防 PFC 死锁规则。

又例如，第一网络设备的控制面可以维护一张链表，其中一个表项为<第一端口，第
三端口，优先级 1，Z'>，表示预设时间段内，第一端口作为入端口，第三端口作为出端
25 口，且优先级 1 对应的次数为 Z。其中 Z' 为预先配置的初始次数，且大于 0。每收到指示
信息后，Z' 都减少预设的次数（例如，1），当次数 Z' 为 0 时，第一网络设备为第一端
口，第三端口生成防 PFC 死锁规则。

可选地，上述指示信息可以通过携带第三发送队列缓存的数据流的优先级，指示所述
第二网络设备中的第三发送队列对应的缓存的容量超过阈值，且所述第三发送队列的入流
30 量为 0。

可选地，上述指示信息可以携带在 PDU 中，上述数据流的优先级可以通过 PDU 中的
可选 TLV 承载，或通过新增 TLV 承载，具体的封装方式可以参见上文关于 PDU 的封装
方式的介绍，为了简洁，在此不再赘述。

上文结合图 1 至图 11 详细地描述了本申请实施例的方法，下文结合图 12 至图 14 详
35 细地描述本申请实施例的确定端口的端口属性的装置。需要说明的是，图 12 至图 14 所示
的装置可以实现上述方法中各个步骤，为了简洁，在此不再赘述。

图 12 是本申请实施例的确定端口的端口属性的装置的示意性框图。图 12 所示的装置
1200 例如可以是上文中的第一网络设备，具体包括接收模块 1210 以及处理模块 1220。可
选地，接收模块 1210 用于执行上述步骤 210，处理模块 1220 用于执行上述步骤 220。可

选地，处理模块 1220 还用于执行上述步骤 1110 以及步骤 1120。

即，接收模块 1210，用于通过第一端口接收第二网络设备发送的协议数据单元 PDU；

处理模块 1220，用于根据所述接收模块接收的所述 PDU，确定所述第一端口为上联口或下联口。

- 5 可选地，作为一个实施例，所述 PDU 携带网络中所述第二网络设备的级数，和/或所述 PDU 携带所述第二网络设备发送所述 PDU 使用的第二端口的端口属性，所述第二端口的端口属性包括所述第二端口为上联口或所述第二端口为下联口。

10 可选地，作为一个实施例，所述处理模块，具体用于：根据所述 PDU 确定在所述网络中所述第一网络设备的级数；根据所述第一网络设备的级数，确定所述第一端口为上联口或下联口。

可选地，作为一个实施例，所述网络包括多个网络设备，所述多个网络设备包括所述第二网络设备，所述接收模块，还用于：通过多个端口接收所述多个网络设备发送的多个 PDU，所述多个端口包括所述第一端口，所述多个 PDU 包括所述第二网络设备发送的所述 PDU；所述处理模块还用于：根据所述多个 PDU，确定在所述网络中所述多个网络设备的级数；从所述多个网络设备的级数中选择目标级数；根据所述目标级数以及预设规则确定所述第一网络设备的级数。

可选地，作为一个实施例，所述处理模块，具体用于：根据所述第一网络设备的级数，以及所述第二网络设备的级数确定所述第一端口为上联口或下联口。

20 可选地，作为一个实施例，所述处理模块，具体用于：根据所述 PDU 确定所述第二网络设备的级数；根据所述第二网络设备的级数确定所述第一端口为上联口或下联口。

可选地，作为一个实施例，所述第二网络设备为所述网络中的根节点，所述接收模块还用于：通过所述第一端口接收所述第二网络设备发送的所述 PDU，所述 PDU 中携带的所述第二网络设备的级数为预设值；所述处理模块还用于：根据所述预设值确定所述第一端口为上联口。

25 可选地，作为一个实施例，所述处理模块，还用于：根据所述预设值，确定所述第一网络设备的级数，其中，在网络中，所述第一网络设备的级数表示的层级与所述根节点所在的层级相邻。

30 可选地，作为一个实施例，所述处理模块，还用于：根据所述 PDU，确定所述第二网络设备为服务器，所述 PDU 携带第一 TLV，所述第一 TLV 指示所述第二网络设备为服务器；确定所述第一端口为下联口。

可选地，作为一个实施例，所述 PDU 为 LLDP PDU。

可选地，作为一个实施例，所述处理模块，还用于：根据所述 LLDP PDU 中端口 ID 的子类型，确定所述第二网络设备为服务器；确定所述第一端口为下联口。

35 可选地，作为一个实施例，所述处理模块，还用于：根据所述 LLDP PDU 中端口 ID 的子类型，确定所述第一网络设备的级数，其中，在网络中所述第一网络设备的级数表示的层级与所述服务器所在的层级相邻。

可选地，作为一个实施例，所述装置还包括发送模块，所述发送模块用于：向所述第二网络设备发送 LLDP 请求，所述 LLDP 请求用于请求所述第二网络设备的第二端口的端口属性，和/或所述第二网络设备在网络中的级数。

可选地，作为一个实施例，若所述第一端口的端口属性未知，所述发送模块还用于向所述第二网络设备发送所述 LLDP 请求。

可选地，作为一个实施例，所述 LLDP PDU 包括第二 TLV，所述第二 TLV 用于指示所述第二网络设备的级数，和/或所述 LLDP PDU 包括第三 TLV，所述第三 TLV 用于指示所述第二端口的端口属性。

可选地，作为一个实施例，所述第二 TLV 或所述第三 TLV 占用所述 LLDP PDU 中的保留位。

可选地，作为一个实施例，所述装置还包括多个发送队列，所述多个发送队列包括第一发送队列和第二发送队列，所述第一发送队列为第一数据流对应的原始发送队列，所述处理模块，还用于确定所述第一端口以及所述第一网络设备的第三端口为上联口；以及生成防优先级流量控制 PFC 死锁规则，所述防 PFC 死锁规则用于指示将所述第一数据流缓存至所述第二发送队列，所述第一数据流的入端口为所述第一端口，所述第一数据流的出端口为所述第三端口。

可选地，作为一个实施例，所述处理模块，还用于：确定在预设时间段内，所述第一端口与所述第三端口为 PFC 死锁关联的端口的次数；当所述次数高于预设阈值，确定生成所述防 PFC 死锁规则。

可选地，作为一个实施例，所述发送模块，还用于通过所述第一端口接收所述第二网络设备发送的指示信息，所述指示信息用于指示所述第二网络设备中的第三发送队列对应的缓存的容量超过阈值，且所述第三发送队列的入流量为 0，所述第三发送队列缓存的数据流的优先级与所述第一数据流的优先级相同；所述处理模块，用于根据指示信息确定所述次数。

图 13 是本申请另一实施例的确定端口属性的装置的示意性框图，图 13 所示的装置 1300 例如可以是上文中的第二网络设备，具体可以包括：处理模块 1310 以及发送模块 1320。可选地，处理模块 1310 用于执行上述步骤 200，发送模块 1320 用于执行上述步骤 210 中第二网络设备实现的功能。

即，处理模块 1310，用于生成协议数据单元 PDU，所述 PDU 用于确定第一网络设备的第一端口为上联口或下联口；

发送模块 1320，用于通过第二端口向所述第一网络设备发送所述处理模块生成的所述 PDU。

可选地，作为一个实施例，所述 PDU 携带网络中所述第二网络设备的级数，和/或所述 PDU 携带所述第二网络设备发送所述 PDU 使用的第二端口的端口属性，所述第二端口的端口属性包括所述第二端口为上联口或所述第二端口为下联口。

可选地，作为一个实施例，所述第二网络设备为所述网络中的根节点，所述 PDU 中携带的所述第二网络设备的级数为预设值。

可选地，作为一个实施例，所述第二网络设备为服务器，所述 PDU 携带第一 TLV，所述第一 TLV 指示所述第二网络设备为服务器。

可选地，作为一个实施例，所述 PDU 为 LLDP PDU。

可选地，作为一个实施例，所述装置还包括接收模块，所述接收模块用于接收所述第一网络设备发送的 LLDP 请求，所述 LLDP 请求用于请求所述第二端口的端口属性，和/

或所述第二网络设备的级数。

可选地，作为一个实施例，所述 LLDP PDU 包括第二 TLV，所述第二 TLV 用于指示所述第二网络设备的级数，和/或所述 LLDP PDU 包括第三 TLV，所述第三 TLV 用于指示所述第二端口的端口属性。

5 可选地，作为一个实施例，所述第二 TLV 或所述第三 TLV 占用所述 LLDP PDU 中的保留位。

可选地，作为一个实施例，所述发送模块，还用于向所述第一网络设备发送指示信息，所述指示信息用于指示所述装置中的第三发送队列对应的缓存的容量超过阈值，且所述第三发送队列的入流量为 0，所述第三发送队列缓存的数据流的优先级与所述第一数据流的
10 优先级相同。

在可选的实施例中，所述接收模块 1210 可以为入端口 1430，所述处理模块 620 可以为处理器 1420，所述网络设备还可以包括存储器 1410 以及出端口 1440，具体如图 14 所示。

在可选的实施例中，所述发送模块 1310 可以为出端口 1440，所述处理模块 1320 可以
15 以为处理器 1420，所述网络设备还可以包括存储器 1410 以及入端口 1430，具体如图 14 所示。

图 14 是本申请另一实施例的网络设备的示意性框图。图 14 所示的网络设备 1400 可以包括：存储器 1410、处理器 1420、入端口 1430，以及出端口 1440。其中，存储器 1410、处理器 1420，入端口 1430，以及出端口 1440 通过内部连接通路相连，该存储器 1410 用于
20 存储指令，该处理器 1420 用于执行该存储器 1420 存储的指令，以控制入端口 1430 接收输入的 PDU 或数据流，控制出端口 1440 转发待发送的 PDU 或者数据流。

需要说明的是，上述入端口 1430 可以是网络设备包括的多个端口中的一个。上述出端口 1440 可以网络设备包括的多个端口中的一个。上述入端口 1430 和出端口 1440 可以理解
25 为两个逻辑端口，也就是说可以分别对应两个物理端口，也可以对应一个物理端口，本申请对此不作限定。

在实现过程中，上述方法的各步骤可以通过处理器 1420 中的硬件的集成逻辑电路或者软件形式的指令完成。结合本申请实施例所公开的方法可以直接体现为硬件处理器执行完成，或者用处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器，闪存、只读存储器，可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的
30 存储介质中。该存储介质位于存储器 1410，处理器 1420 读取存储器 1410 中的信息，结合其硬件完成上述方法的步骤。为避免重复，这里不再详细描述。

应理解，本申请实施例中，该处理器可以为中央处理单元（central processing unit，CPU），该处理器还可以是其他通用处理器、数字信号处理器（digital signal processor，DSP）、专用集成电路（application specific integrated circuit，ASIC）、现成可编程门阵列
35 （field programmable gate array，FPGA）或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

还应理解，本申请实施例中，该存储器可以包括只读存储器和随机存取存储器，并向处理器提供指令和数据。处理器的一部分还可以包括非易失性随机存取存储器。例如，处

理器还可以存储设备类型的信息。

应理解，在本申请实施例中，“与 A 对应的 B”表示 B 与 A 相关联，根据 A 可以确定 B。但还应理解，根据 A 确定 B 并不意味着仅仅根据 A 确定 B，还可以根据 A 和/或其它信息确定 B。

5 应理解，本文中术语“和/或”，仅仅是一种描述关联对象的关联关系，表示可以存在三种关系，例如，A 和/或 B，可以表示：单独存在 A，同时存在 A 和 B，单独存在 B 这三种情况。另外，本文中字符“/”，一般表示前后关联对象是一种“或”的关系。

10 应理解，在本申请的各种实施例中，上述各过程的序号的大小并不意味着执行顺序的先后，各过程的执行顺序应以其功能和内在逻辑确定，而不对本申请实施例的实施过程构成任何限定。

15 本领域普通技术人员可以意识到，结合本文中所公开的实施例描述的各示例的单元及算法步骤，能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行，取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本申请的范围。

所属领域的技术人员可以清楚地了解到，为描述的方便和简洁，上述描述的系统、装置和单元的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

20 在本申请所提供的几个实施例中，应该理解到，所揭露的系统、装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

25 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

另外，在本申请各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。

30 所述功能如果以软件功能单元的形式实现并作为独立的产品销售或使用，可以存储在一个计算机可读取存储介质中。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来，该计算机软件产品存储在一个存储介质中，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等）执行本申请各个实施例所述方法的全部或部分步骤。而前述的存储介质包括：U 盘、移动硬盘、只读存储器（Read-Only Memory, ROM）、随机存取存储器（Random Access Memory, RAM）、磁碟或者光盘等各种可以存储程序代码的介质。

以上所述，仅为本申请的具体实施方式，但本申请的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本申请揭露的技术范围内，可轻易想到变化或替换，都应涵盖在本申请的保护范围之内。因此，本申请的保护范围应以所述权利要求的保护范围为准。

权 利 要 求 书

1、一种确定端口属性的方法，其特征在于，包括：

第一网络设备通过第一端口接收第二网络设备发送的协议数据单元 PDU；

5 所述第一网络设备根据所述 PDU，确定所述第一端口为上联口或下联口。

2、如权利要求 1 所述的方法，其特征在于，所述 PDU 携带网络中所述第二网络设备的级数，和/或

所述 PDU 携带所述第二网络设备发送所述 PDU 使用的第二端口的端口属性，所述第二端口的端口属性包括所述第二端口为上联口或所述第二端口为下联口。

10 3、如权利要求 2 所述的方法，其特征在于，所述第一网络设备根据所述 PDU，确定所述第一端口为上联口或下联口，包括：

所述第一网络设备根据所述 PDU 确定在所述网络中所述第一网络设备的级数；

所述第一网络设备根据所述第一网络设备的级数，确定所述第一端口为上联口或下联口。

15 4、如权利要求 3 所述的方法，其特征在于，所述网络包括多个网络设备，所述多个网络设备包括所述第二网络设备，

所述第一网络设备通过第一端口接收第二网络设备发送的协议数据单元 PDU，包括：

所述第一网络设备通过多个端口接收所述多个网络设备发送的多个 PDU，所述多个端口包括所述第一端口，所述多个 PDU 包括所述第二网络设备发送的所述 PDU；

20 所述第一网络设备根据所述 PDU 确定在所述网络中所述第一网络设备的级数，包括：

所述第一网络设备根据所述多个 PDU，确定在所述网络中所述多个网络设备的级数；

所述第一网络设备从所述多个网络设备的级数中选择目标级数；

所述第一网络设备根据所述目标级数以及预设规则确定所述第一网络设备的级数。

25 5、如权利要求 3 或 4 所述的方法，其特征在于，所述第一网络设备根据所述第一网络设备的级数，确定所述第一端口为所述上联口或所述下联口，包括：

所述第一网络设备根据所述第一网络设备的级数，以及所述第二网络设备的级数确定所述第一端口为上联口或下联口。

6、如权利要求 2 所述的方法，其特征在于，所述第一网络设备根据所述 PDU，确定所述第一端口为上联口或下联口，包括：

30 所述第一网络设备根据所述 PDU 确定所述第二网络设备的级数；

所述第一网络设备根据所述第二网络设备的级数确定所述第一端口为上联口或下联口。

7、如权利要求 6 所述的方法，其特征在于，所述第二网络设备为所述网络中的根节点，所述第一网络设备通过第一端口接收第二网络设备发送的 PDU，包括：

35 所述第一网络设备通过所述第一端口接收所述第二网络设备发送的所述 PDU，所述 PDU 中携带的所述第二网络设备的级数为预设值；

所述第一网络设备根据所述第二网络设备的级数确定所述第一端口为上联口或下联口，包括：

所述第一网络设备根据所述预设值确定所述第一端口为上联口。

8、如权利要求1所述的方法，其特征在于，所述第一网络设备根据所述PDU，确定所述第一端口为上联口或下联口，包括：

所述第一网络设备根据所述PDU，确定所述第二网络设备为服务器，所述PDU携带
5 第一TLV，所述第一TLV指示所述第二网络设备为服务器；

所述第一网络设备确定所述第一端口为下联口。

9、权利要求1-8中任一项所述的方法，其特征在于，所述PDU为链路层发现协议LLDP
PDU。

10、如权利要求9所述的方法，其特征在于，在所述第一网络设备接收第二网络设备
10 发送的协议数据单元PDU之前，所述方法还包括：

所述第一网络设备向所述第二网络设备发送LLDP请求，所述LLDP请求用于请求所
述第二网络设备的第二端口的端口属性，和/或所述第二网络设备在网络中的级数。

11、如权利要求10所述的方法，其特征在于，所述第一网络设备向所述第二网络设
备发送LLDP请求，包括：

15 若所述第一端口的端口属性未知，所述第一网络设备向所述第二网络设备发送所述
LLDP请求。

12、如权利要求1-11中任一项所述的方法，其特征在于，所述第一网络设备包括多
个发送队列，所述多个发送队列包括第一发送队列和第二发送队列，所述第一发送队列为
第一数据流对应的原始发送队列，

20 所述方法还包括：

所述第一网络设备确定所述第一端口以及所述第一网络设备的第三端口为上联口；

所述第一网络设备生成防优先级流量控制PFC死锁规则，所述防PFC死锁规则用于
指示将所述第一数据流缓存至所述第二发送队列，所述第一数据流的入端口为所述第一端
口，所述第一数据流的出端口为所述第三端口。

25 13、一种确定端口属性的方法，其特征在于，包括：

第二网络设备生成协议数据单元PDU，所述PDU用于确定第一网络设备的所述第一端口
为上联口或下联口；

所述第二网络设备通过第二端口向所述第一网络设备发送所述PDU。

30 14、如权利要求13所述的方法，其特征在于，所述PDU携带网络中所述第二网络设
备的级数，和/或

所述PDU携带所述第二网络设备的所述第二端口的端口属性，所述第二端口的端口
属性包括所述第二端口为上联口或所述第二端口为下联口。

15、权利要求14所述的方法，其特征在于，所述第二网络设备为所述网络中的根节
点，所述PDU中携带的所述第二网络设备的级数为预设值。

35 16、权利要求13所述的方法，其特征在于，所述第二网络设备为服务器，所述PDU
携带第一TLV，所述第一TLV指示所述第二网络设备为服务器。

17、权利要求13-16中任一项所述的方法，其特征在于，所述PDU为链路层发现协
议LLDP PDU。

18、如权利要求17所述的方法，其特征在于，在所述第二网络设备向所述第一网络

设备发送所述 PDU 之前, 所述方法还包括:

所述第二网络设备接收所述第一网络设备发送的 LLDP 请求, 所述 LLDP 请求用于请求所述第二端口的端口属性, 和/或所述第二网络设备的级数。

19、一种第一网络设备, 其特征在于, 包括:

5 接收模块, 用于通过第一端口接收第二网络设备发送的协议数据单元 PDU;

处理模块, 用于根据所述接收模块接收的所述 PDU, 确定所述第一端口为上联口或下联口。

20、如权利要求 19 所述的第一网络设备, 其特征在于, 所述 PDU 携带网络中所述第二网络设备的级数, 和/或

10 所述 PDU 携带所述第二网络设备发送所述 PDU 使用的第二端口的端口属性, 所述第二端口的端口属性包括所述第二端口为上联口或所述第二端口为下联口。

21、如权利要求 20 所述的第一网络设备, 其特征在于, 所述处理模块还用于:

根据所述 PDU 确定在所述网络中所述第一网络设备的级数;

根据所述第一网络设备的级数, 确定所述第一端口为上联口或下联口。

15 22、如权利要求 21 所述的第一网络设备, 其特征在于, 所述网络包括多个网络设备, 所述多个网络设备包括所述第二网络设备,

所述接收模块, 还用于通过多个端口接收所述多个网络设备发送的多个 PDU, 所述多个端口包括所述第一端口, 所述多个 PDU 包括所述第二网络设备发送的所述 PDU;

20 所述处理模块, 还用于根据所述多个 PDU, 确定在所述网络中所述多个网络设备的级数;

所述处理模块, 还用于从所述多个网络设备的级数中选择目标级数;

所述处理模块, 还用于根据所述目标级数以及预设规则确定所述第一网络设备的级数。

25 23、如权利要求 21 或 22 所述的第一网络设备, 其特征在于, 所述处理模块, 还用于: 根据所述第一网络设备的级数, 以及所述第二网络设备的级数确定所述第一端口为上联口或下联口。

24、如权利要求 20 所述的第一网络设备, 其特征在于, 所述处理模块, 还用于:

根据所述 PDU 确定所述第二网络设备的级数;

根据所述第二网络设备的级数确定所述第一端口为上联口或下联口。

30 25、如权利要求 24 所述的第一网络设备, 其特征在于, 所述第二网络设备为所述网络中的根节点,

所述接收模块, 还用于通过所述第一端口接收所述第二网络设备发送的所述 PDU, 所述 PDU 中携带的所述第二网络设备的级数为预设值;

所述处理模块, 还用于根据所述预设值确定所述第一端口为上联口。

35 26、如权利要求 19 所述的第一网络设备, 其特征在于, 所述处理模块, 还用于:

根据所述 PDU, 确定所述第二网络设备为服务器, 所述 PDU 携带第一 TLV, 所述第一 TLV 指示所述第二网络设备为服务器;

确定所述第一端口为下联口。

27、权利要求 19-26 中任一项所述的第一网络设备, 其特征在于, 所述 PDU 为链路

层发现协议 LLDP PDU。

28、如权利要求 27 所述的第一网络设备，其特征在于，所述第一网络设备还包括：

发送模块，用于向所述第二网络设备发送 LLDP 请求，所述 LLDP 请求用于请求所述第二网络设备的第二端口的端口属性，和/或所述第二网络设备在网络中的级数。

5 29、如权利要求 28 所述的第一网络设备，其特征在于，所述发送模块，还用于若所述第一端口的端口属性未知，向所述第二网络设备发送所述 LLDP 请求。

30、如权利要求 19-29 中任一项所述的第一网络设备，其特征在于，所述第一网络设备包括多个发送队列，所述多个发送队列包括第一发送队列和第二发送队列，所述第一发送队列为第一数据流对应的原始发送队列，

10 所述处理模块还用于：

确定所述第一端口以及所述第一网络设备的第三端口为上联口；

生成防优先级流量控制 PFC 死锁规则，所述防 PFC 死锁规则用于指示将所述第一数据流缓存至所述第二发送队列，所述第一数据流的入端口为所述第一端口，所述第一数据流的出端口为所述第三端口。

15 31、一种第二网络设备，其特征在于，包括：

处理模块，用于生成协议数据单元 PDU，所述 PDU 用于确定第一网络设备的第一端口为上联口或下联口；

发送模块，用于通过第二端口向所述第一网络设备发送所述处理模块生成的所述 PDU。

20 32、如权利要求 31 所述的第二网络设备，其特征在于，所述 PDU 携带网络中所述第二网络设备的级数，和/或

所述 PDU 携带所述第二网络设备的所述第二端口的端口属性，所述第二端口的端口属性包括所述第二端口为上联口或所述第二端口为下联口。

25 33、权利要求 32 所述的第二网络设备，其特征在于，所述第二网络设备为所述网络中的根节点，所述 PDU 中携带的所述第二网络设备的级数为预设值。

34、权利要求 31 所述的第二网络设备，其特征在于，所述第二网络设备为服务器，所述 PDU 携带第一 TLV，所述第一 TLV 指示所述第二网络设备为服务器。

35、权利要求 31-34 中任一项所述的第二网络设备，其特征在于，所述 PDU 为链路层发现协议 LLDP PDU。

30 36、如权利要求 35 所述的第二网络设备，其特征在于，所述第二网络设备还包括：

接收模块，用于接收所述第一网络设备发送的 LLDP 请求，所述 LLDP 请求用于请求所述第二端口的端口属性，和/或所述第二网络设备的级数。

37、一种网络设备，其特征在于，包括处理器，所述处理器用于与存储器耦合，读取并执行所述存储器中的指令，以执行如权利要求 1-18 中任一项所述的方法。

35 38、一种计算机可读介质，所述计算机可读介质存储有程序代码，当所述计算机程序代码在计算机上运行时，使得计算机执行如权利要求 1-18 中任一项所述的方法。

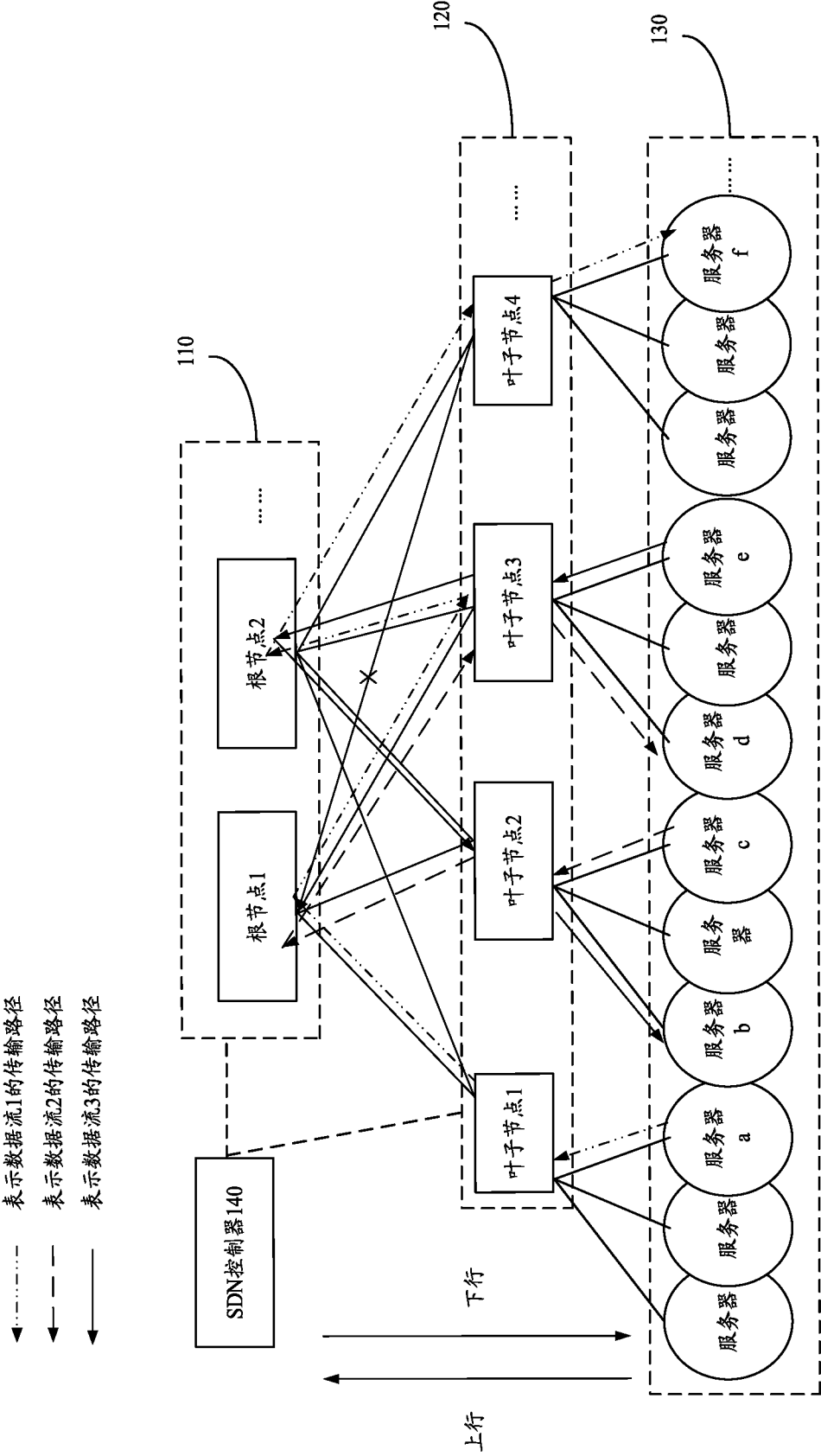


图1

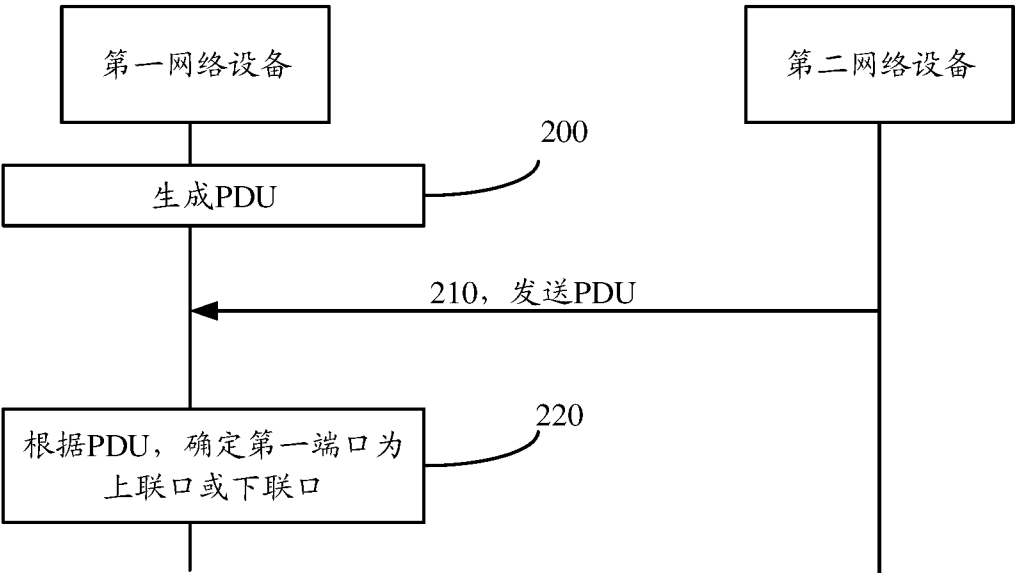


图 2



图 3

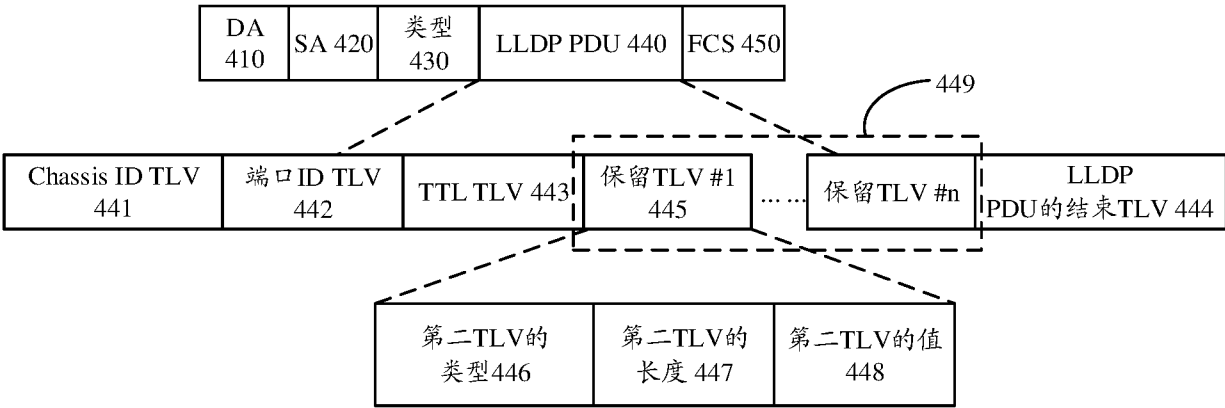


图 4



图 5

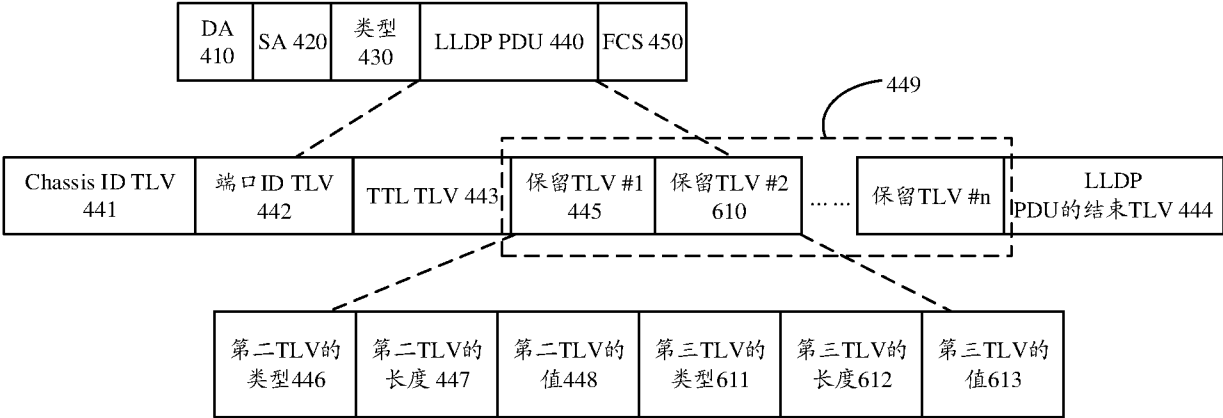


图 6

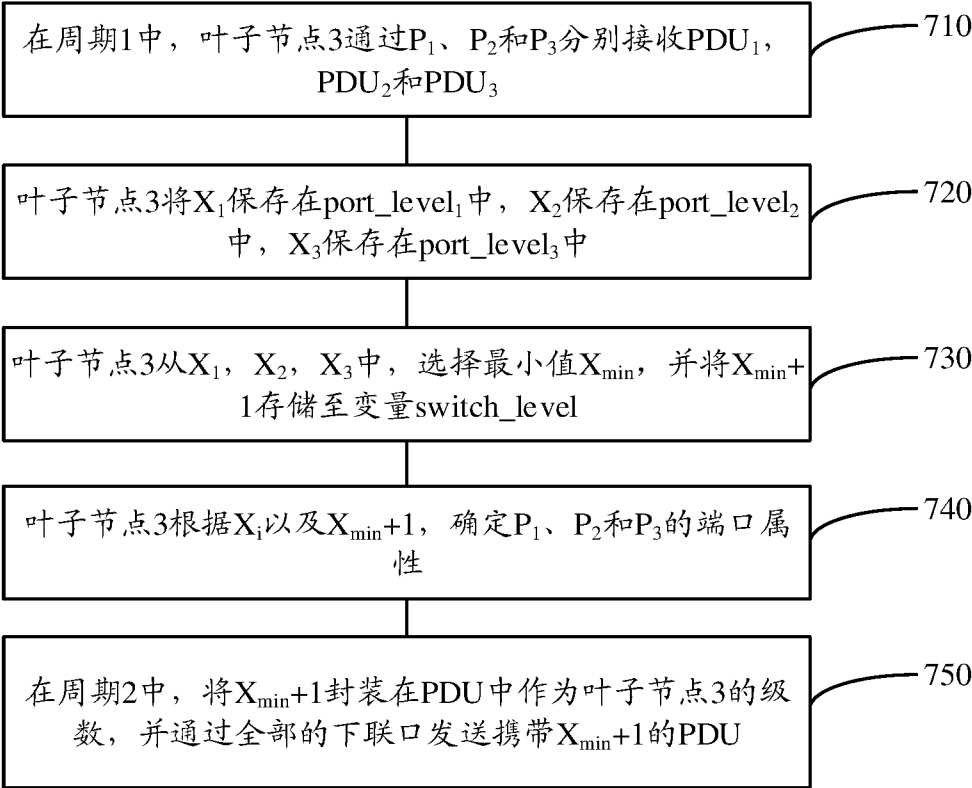


图 7

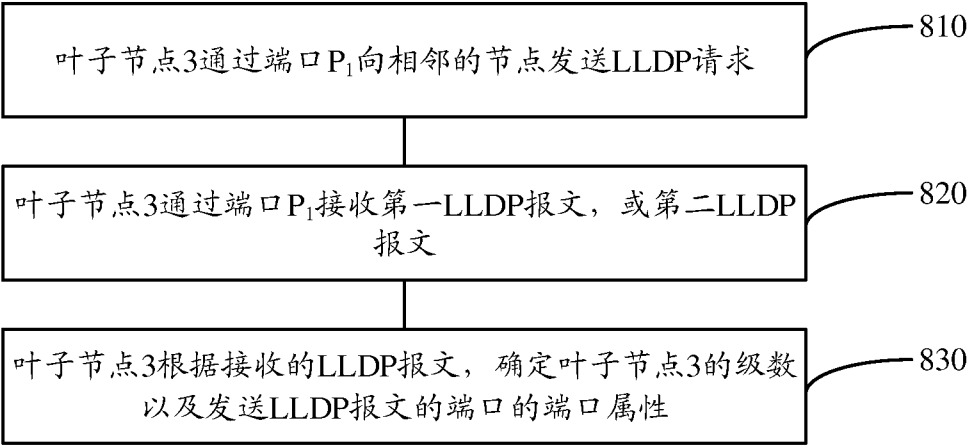


图 8

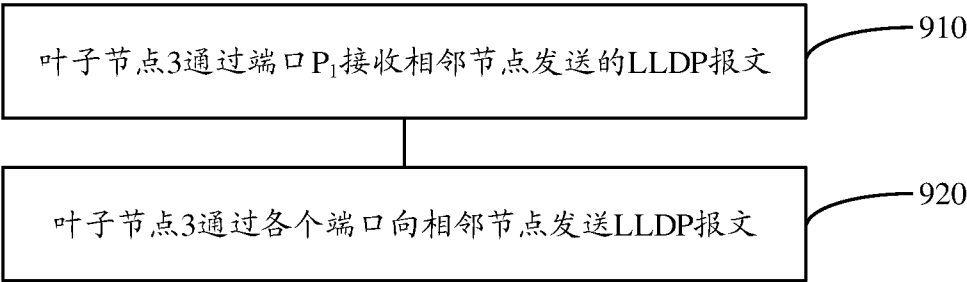


图 9

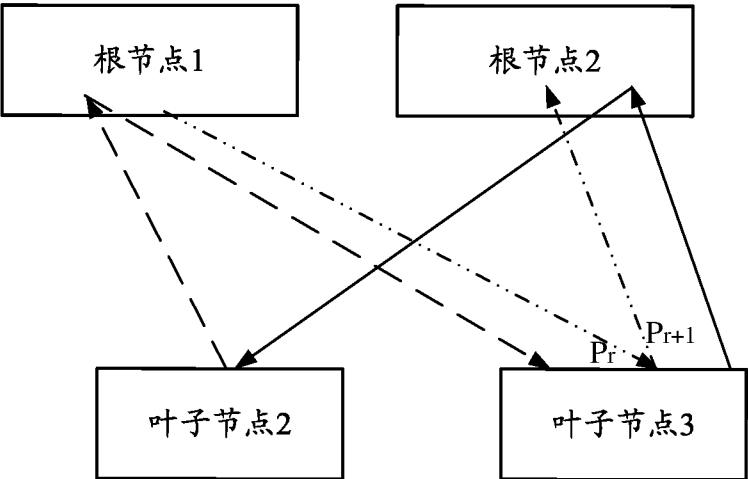
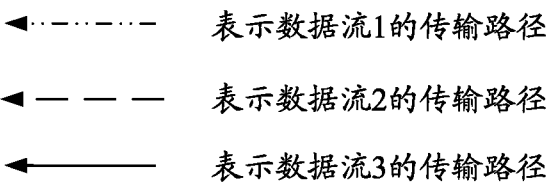


图 10

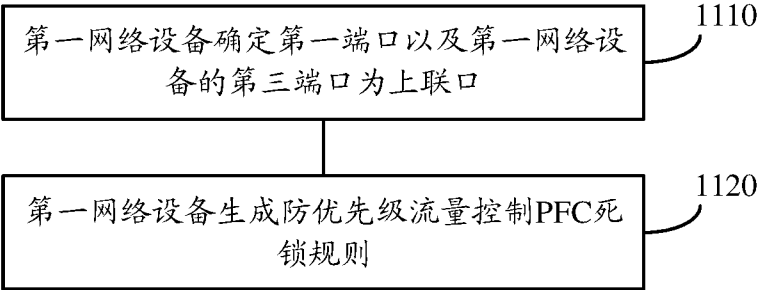


图 11

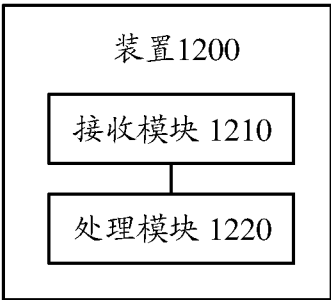


图 12

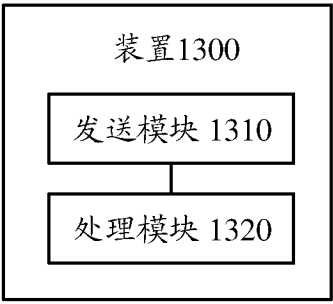


图 13

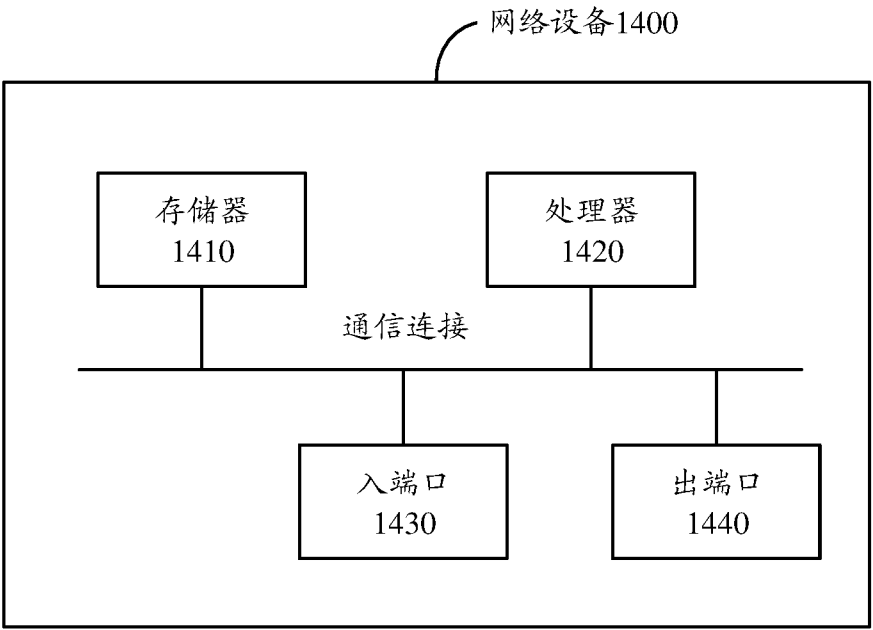


图 14

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/092209

A. CLASSIFICATION OF SUBJECT MATTER

H04L 12/24(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, CNTXT, CNKI: 网络设备, 端口, 接口, 属性, 协议数据单元, PDU, 级数, 上联, 下联, TLV, 链路层发现协议, LLDP, 防优先级控制, PFC, 死锁 VEN, USTXT, WOTXT, EPTXT, GOOGLE: network equipment, port, interface, attribute, protocol data unit, PDU, level, uplink, upper link, downlink, TLV, link layer discovery protocol, LLDP, anti-priority control, PFC, deadlock

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 102377679 A (FIBERHOME TELECOMMUNICATION TECHNOLOGIES CO., LTD.) 14 March 2012 (2012-03-14) see description, paragraphs [0007]-[0014]	1, 13, 19, 31, 37-38
A	CN 101262412 A (ALCATEL SHANGHAI BELL CO., LTD.) 10 September 2008 (2008-09-10) entire document	1-38
A	CN 101043422 A (ALCATEL SHANGHAI BELL CO., LTD.) 26 September 2007 (2007-09-26) entire document	1-38
A	WO 2012171378 A1 (ZTE CORPORATION et al.) 20 December 2012 (2012-12-20) entire document	1-38

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

11 August 2020

Date of mailing of the international search report

17 August 2020

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Authorized officer

Facsimile No. (86-10)62019451

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2020/092209

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	102377679	A	14 March 2012	CN	102377679	B	31 December 2014
CN	101262412	A	10 September 2008	CN	101262412	B	29 September 2010
CN	101043422	A	26 September 2007	CN	100571205	C	16 December 2009
WO	2012171378	A1	20 December 2012	CN	102244609	B	21 December 2016
				CN	102244609	A	16 November 2011

国际检索报告

国际申请号

PCT/CN2020/092209

A. 主题的分类 H04L 12/24 (2006.01) i 按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类																	
B. 检索领域 检索的最低限度文献 (标明分类系统和分类号) H04L 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用)) CNABS, CNTXT, CNKI: 网络设备、端口、接口、属性、协议数据单元、PDU、级数、上联、下联、TLV、链路层发现协议、LLDP、防优先级控制、PFC、死锁 VEN, USTXT, WOTXT, EPTXT, GOOGLE: network equipment, port, interface, attribute, protocol data unit, PDU, level, uplink, upper link, downlink, TLV, link layer discovery protocol, LLDP, anti-priority control, PFC, deadlock																	
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>CN 102377679 A (烽火通信科技股份有限公司) 2012年 3月 14日 (2012 - 03 - 14) 参见说明书第[0007]段至第[0014]段</td> <td>1、13、19、31、37-38</td> </tr> <tr> <td>A</td> <td>CN 101262412 A (上海贝尔阿尔卡特股份有限公司) 2008年 9月 10日 (2008 - 09 - 10) 全文</td> <td>1-38</td> </tr> <tr> <td>A</td> <td>CN 101043422 A (上海贝尔阿尔卡特股份有限公司) 2007年 9月 26日 (2007 - 09 - 26) 全文</td> <td>1-38</td> </tr> <tr> <td>A</td> <td>WO 2012171378 A1 (中兴通讯股份有限公司等) 2012年 12月 20日 (2012 - 12 - 20) 全文</td> <td>1-38</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	CN 102377679 A (烽火通信科技股份有限公司) 2012年 3月 14日 (2012 - 03 - 14) 参见说明书第[0007]段至第[0014]段	1、13、19、31、37-38	A	CN 101262412 A (上海贝尔阿尔卡特股份有限公司) 2008年 9月 10日 (2008 - 09 - 10) 全文	1-38	A	CN 101043422 A (上海贝尔阿尔卡特股份有限公司) 2007年 9月 26日 (2007 - 09 - 26) 全文	1-38	A	WO 2012171378 A1 (中兴通讯股份有限公司等) 2012年 12月 20日 (2012 - 12 - 20) 全文	1-38
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求															
X	CN 102377679 A (烽火通信科技股份有限公司) 2012年 3月 14日 (2012 - 03 - 14) 参见说明书第[0007]段至第[0014]段	1、13、19、31、37-38															
A	CN 101262412 A (上海贝尔阿尔卡特股份有限公司) 2008年 9月 10日 (2008 - 09 - 10) 全文	1-38															
A	CN 101043422 A (上海贝尔阿尔卡特股份有限公司) 2007年 9月 26日 (2007 - 09 - 26) 全文	1-38															
A	WO 2012171378 A1 (中兴通讯股份有限公司等) 2012年 12月 20日 (2012 - 12 - 20) 全文	1-38															
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																	
<table border="0"> <tr> <td> * 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 </td> <td> “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件 </td> </tr> </table>			* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件	“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件													
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件	“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件																
国际检索实际完成的日期 2020年 8月 11日		国际检索报告邮寄日期 2020年 8月 17日															
ISA/CN的名称和邮寄地址 中国国家知识产权局 (ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		授权官员 唐文森 电话号码 62411230															

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/092209

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	102377679	A	2012年 3月 14日	CN	102377679 B 2014年 12月 31日
CN	101262412	A	2008年 9月 10日	CN	101262412 B 2010年 9月 29日
CN	101043422	A	2007年 9月 26日	CN	100571205 C 2009年 12月 16日
WO	2012171378	A1	2012年 12月 20日	CN	102244609 B 2016年 12月 21日
				CN	102244609 A 2011年 11月 16日