

(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
2017 年 1 月 19 日 (19.01.2017)



(10) 国际公布号
WO 2017/008712 A1

- (51) 国际专利分类号:
H04L 12/26 (2006.01)
- (21) 国际申请号: PCT/CN2016/089597
- (22) 国际申请日: 2016 年 7 月 11 日 (11.07.2016)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201510408264.5 2015 年 7 月 10 日 (10.07.2015) CN
- (71) 申请人: 杭州华三通信技术有限公司 (HANG-ZHOU H3C TECHNOLOGIES CO., LTD.) [CN/CN]; 中国浙江省杭州市滨江区长河路 466 号, Zhejiang 310052 (CN)。
- (72) 发明人: 高瑞昌 (GAO, Ruichang); 中国北京市海淀区上地信息产业基地创业路 2 号东方电子大厦 730 室, Beijing 100085 (CN)。 乔剌 (QIAO, Yan); 中国北京市海淀区上地信息产业基地创业路 2 号东方电子大厦 730 室, Beijing 100085 (CN)。 鄢能 (YAN, Neng); 中国北京市海淀区上地信息产业基地创业路 2 号东方电子大厦 730 室, Beijing 100085 (CN)。 缪琛 (MIAO, Chen); 中国北京市海淀区上地信息产业基地创业路 2 号东方电子大厦 730 室, Beijing 100085 (CN)。

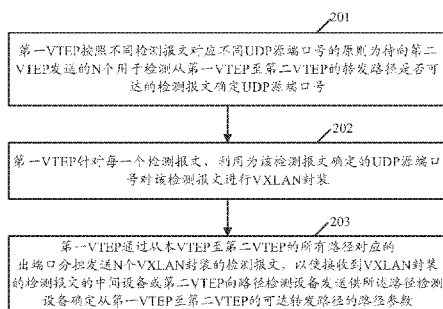
- (74) 代理人: 北京德琦知识产权代理有限公司 (DEQI INTELLECTUAL PROPERTY LAW CORPORATION); 中国北京市海淀区知春路 1 号学院国际大厦 7 层, Beijing 100083 (CN)。
- (81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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))。

(54) Title: REACHABILITY DETECTION FOR PATH BETWEEN VXLAN TUNNEL END POINTS (VTEP)

(54) 发明名称: VXLAN 隧道端点 VTEP 之间的路径可达检测



201 A FIRST VTEP DETERMINES UDP SOURCE PORT NUMBERS FOR A NUMBER N OF DETECTION PACKETS TO BE TRANSMITTED TO A SECOND VTEP AND FOR USE IN DETECTING WHETHER OR NOT FORWARD PATHS FROM THE FIRST VTEP TO A SECOND VTEP ARE REACHABLE ACCORDING TO THE PRINCIPLE OF DIFFERENT DETECTION PACKETS CORRESPOND TO DIFFERENT UDP SOURCE PORT NUMBERS
202 WITH RESPECT TO EACH DETECTION PACKET, THE FIRST VTEP VXLAN ENCAPSULATES THE DETECTION PACKET BY UTILIZING THE UDP SOURCE PORT NUMBER DETERMINED FOR THE DETECTION PACKET
203 THE FIRST VTEP DISTRIBUTIVELY TRANSMITS THE NUMBER N OF VXLAN ENCAPSULATED DETECTION PACKETS VIA OUT PORTS CORRESPONDING TO ALL OF THE PATHS FROM THE PRESENT VTEP TO THE SECOND VTEP, THUS ALLOWING AN INTERMEDIATE DEVICE RECEIVING THE VXLAN ENCAPSULATED DETECTION PACKETS OR THE SECOND VTEP TO TRANSMIT TO A PATH DETECTION DEVICE A PATH PARAMETER FOR THE PATH DETECTION DEVICE TO DETERMINE A REACHABLE FORWARD PATH FROM THE FIRST VTEP TO THE SECOND VTEP

图 1

(57) Abstract: In the present application, a first VTEP determines different UDP source port numbers for use in VXLAN encapsulation for a number N of detection packets to be transmitted to a second VTEP. This allows an intermediate device receiving VXLAN encapsulated detection packets to utilize the UDP source port numbers in a VXLAN encapsulation in selecting a forward path when two or more forward paths to the second VTEP are determined, where the number N of the detection packets is greater than the number M of all forward paths from the first VTEP to the second VTEP, and the first VTEP distributively transmits the number N of VXLAN encapsulated detection packets via out ports corresponding to all of the paths from the first VTEP to the second VTEP.

(57) 摘要: 本申请中, 第一 VTEP 为待向第二 VTEP 发送的 N 份检测报文确定不同的用于 VXLAN 封装的 UDP 源端口号, 这能够使接收到 VXLAN 封装的检测报文的中间设备在确定出至第二 VTEP 的转发路径有两条以上时利用 VXLAN 封装中的 UDP 源端口号选择转发路径, 检测报文的数量 N 大于从第一 VTEP 至第二 VTEP 的所有转发路径数量 M, 并且, 第一 VTEP 从第一 VTEP 至第二 VTEP 的所有路径对应的出端口分担发送 N 个 VXLAN 封装的检测报文。



WO 2017/008712 A1

VXLAN 隧道端点 VTEP 之间的路径可达检测

发明背景

虚拟扩展局域网（VXLAN：Virtual eXtensible Local Area
5 Network），是一种实现 Overlay 网络架构的技术。VXLAN 技术将二
层报文用三层协议进行封装，主要原理为：引入一个用户数据包协议
（UDP：User Datagram Protocol）格式的外层隧道，作为数据路径层，
而原有的二层报文数据作为净荷来传输。

由于净荷的外层采用了 UDP 作为传输手段，这可以保证净荷顺
10 利在二、三层网络中传输，实现了二层网络覆盖（overlay）三层网络
的能力。这里，为方便描述，将覆盖的虚拟网络称为 overlay 网络，
而将被覆盖的物理承载网络称为 underlay 网络。

附图简要说明

本申请将通过非限制性示例结合以下附图来描述：

- 15 图 1 为本申请示例提供的方法流程示意图；
图 2 为本申请示例提供的 VXLAN 封装结构示意图；
图 3 为本申请示例提供的应用组网示意图；
图 4 为本申请示例提供的应用组网示意图；
图 5 为本申请示例提供的装置结构示意图；
20 图 6 为本申请示例提供的装置的硬件结构示意图。

实施本申请的方式

在提供服务的数据中心，例如提供云服务的数据中心，使用

VXLAN 技术架构 overlay 网络和 underlay 网络。

Overlay 网络中包括隧道端点(VTEP: VXLAN Tunnel End Point)。VXLAN 的相关处理都在 VTEP 上进行, 例如报文的 VXLAN 封装和 VXLAN 解封装。VTEP 可以是虚拟交换机或者物理交换机。

5 Underlay 网络可以是一个物理承载网络, 包括多个主机和其他设备。

Overlay 网络中的 VTEP 与 underlay 网络连接, 分配有 underlay 网络的物理承载网络的 IP 地址, 该 IP 地址与 overlay 网络无关。

在 overlay 网络中, 不同 VTEP 之间的路径(因为 VTEP 与 underlay
10 网络连接, 这里不同 VTEP 之间的路径也称 underlay 路径) 可能有多条, 甚至是多条等价多路径 (ECMP: Equal-Cost Multi-Path), 然而目前还没有有一种方法能够检测 VTEP 之间的 underlay 路径可达。

参见图 1, 图 1 为本申请示例提供的方法流程图。如图 1 所示, 该方法可包括以下步骤:

15 步骤 201, 第一 VTEP 按照不同的检测报文对应不同的 UDP 源端口号的原则, 为待向第二 VTEP 发送的用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达的 N 个检测报文确定 UDP 源端口号。

假设第一 VTEP 到第二 VTEP 之间的路径一共有 M 条, 那么会发送例如 N 个检测报文, 其中检测报文的 UDP 源端口号是连续变化的,
20 的, 确保发出去的检测报文后 HASH (哈希) 均匀, 可以覆盖所有的 M 条路径。

本申请示例中, 第一 VTEP、第二 VTEP 只是为便于描述进行的命名而已, 并非限定本申请。

本申请示例中, N 大于或等于从第一 VTEP 至第二 VTEP 的路径
25 数量 M。

本申请示例中，N 个检测报文可以具有不同的报文关键信息，这里，报文关键信息为报文五元组或者报文七元组包含的至少一个参数。其中，报文五元组包括以下参数：目的 IP 地址、源 IP 地址、目的端口号、源端口号、协议类型；报文七元组包括以下参数：目的 IP 地址、源 IP 地址、目的 MAC 地址、源 MAC 地址、目的端口号、源端口号、协议类型。举例描述，以报文关键信息为目的 MAC 地址为例，为保证 N 个检测报文的目的 MAC 地址不同，这里可预先定义一个目的 MAC 地址范围，比如 00-00-5E-90-00-01 ~ 00-00-5E-90-00-80，N 份检测报文的目的 MAC 地址可为上述目的 MAC 地址范围中不同的目的 MAC 地址。在报文关键信息为目的 MAC 地址的前提下，N 个检测报文的其它参数比如源 MAC 地址、源 IP 地址、目的 IP 地址、目的端口号、源端口号、协议类型是否相同本申请并不具体限定。以其他参数相同为例，则 N 份检测报文的源 MAC 地址可为源端 VTEP 的桥 MAC 地址，源 IP 地址可为源端 VTEP 的 IP 地址，目的 IP 地址可固定设置为一个用于表示目的端 VTEP 收到后丢弃报文的数值，比如 127.0.0.1。至于目的端口号、源端口号、协议类型可预先根据实际需求设置，这里不再一一赘述。

本申请示例中，N 个检测报文也可以具有相同的报文关键信息。在报文关键信息相同的情况下，N 个检测报文的其它参数可以相同也可以不同，这里不具体限制。

如上所述，假如 N 个检测报文具有不同的报文关键信息，则作为本申请的一个示例，本步骤 201 中，按照不同的检测报文对应不同的 UDP 源端口号的原则为待向第二 VTEP 发送的用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达的 N 个检测报文确定外层 UDP 源端口号可通过以下方式 1 或方式 2 实现：

方式 1: 针对每一个检测报文, 按照设定的 UDP 源端口号确定算法对该检测报文具有的报文关键信息进行计算, 将得到的结果确定为该检测报文的 UDP 源端口号。例如, 按照哈希算法, 依据该检测报文的报文五元组计算出该检测报文的 UDP 源端口号。

5 方式 2: 将 N 个检测报文排成一个队列, 为队列中的第一个检测报文确定 UDP 源端口号为预设端口号 K, 为队列中的第二个检测报文确定与第一份检测报文的外层 UDP 源端口号不同的 UDP 源端口号比如预设端口号 K+1, 依次类推, 直至最终为队列中最后一个检测报文确定出与队列中任一个检测报文的外层 UDP 源端口号不同的 UDP
10 源端口号比如预设端口号 K+N。

如上所述, N 个检测报文可具有相同报文关键信息, 基于此, 作为本申请的另一个示例, 本步骤 201 中, 按照不同检测报文对应不同 UDP 源端口号的原则, 为待向第二 VTEP 发送的用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达的 N 个检测报文确定 UDP
15 源端口号可通过上述的方式 2 实现。

通过步骤 201, 最终能够为用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达的 N 个检测报文确定出不同的 UDP 源端口号。

步骤 202, 第一 VTEP 针对每一个检测报文, 利用为该检测报文确定的 UDP 源端口号对该检测报文进行 VXLAN 封装。

20 作为本申请的一个示例, 步骤 202 中第一 VTEP 针对每一个检测报文, 利用为该检测报文确定的 UDP 源端口号对该检测报文进行 VXLAN 封装可包括以下步骤 a1 至步骤 a2:

步骤 a1, 确定用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数。

25 如上所述, 上述 N 个检测报文用于检测从第一 VTEP 至第二

VTEP 的转发路径是否可达，基于此，在本申请示例中，上述 N 个检测报文执行 VXLAN 封装时的 VXLAN 封装参数相同，均为步骤 a1 确定的 VXLAN 封装参数。图 2 示出了 VXLAN 封装的结构。这里的 VXLAN 封装参数包含 VXLAN 封装中除了外层 UDP 头内 UDP 源端口号（Source Port）之外的参数。

步骤 a2，针对每一个检测报文，利用为该检测报文确定的外层 UDP 源端口号和所述 VXLAN 封装参数对该检测报文进行 VXLAN 封装。

也即，在本申请示例中，每一个检测报文的 VXLAN 封装中外层 UDP 头内的 UDP 源端口号为该检测报文确定的 UDP 源端口号，其他参数为上述的 VXLAN 封装参数。

步骤 203，第一 VTEP 通过从本 VTEP 至第二 VTEP 的所有路径对应的出端口分担发送 N 个 VXLAN 封装的检测报文，以使接收到 VXLAN 封装的检测报文的中间设备或第二 VTEP 向路径检测设备发送供所述路径检测设备确定从第一 VTEP 至第二 VTEP 的可达转发路径的路径参数。第一 VTEP 和第二 VTEP 之间是 underlay 三层转发链路，中间设备可以为三层交换机或者路由器。

在一个示例中，第一 VTEP 通过从本 VTEP 至第二 VTEP 的所有路径对应的出端口分担发送 N 个 VXLAN 封装的检测报文可包括以下步骤：

步骤 b1，确定从第一 VTEP 至第二 VTEP 的路径对应的出端口。

这里，从第一 VTEP 至第二 VTEP 的路径对应的出端口可由第一 VTEP 依据本地路由表查找目的地为第二 VTEP 的路径对应的出端口，不再具体描述。

如上所述，M 为从第一 VTEP 至第二 VTEP 的转发路径数量，基于

此，本步骤 b1 中确定的路径对应的出端口的数量（以 L 表示）小于或等于 M。

步骤 b2，在确定的路径对应的出端口的数量 L 为 1 时，通过该确定出的 1 个路径对应的出端口发送 N 个 VXLAN 封装的检测报文；在确定的路径对应的出端口的数量 L 大于 1 时，按照均衡分担原则从 L 个路径对应的出端口中为每一个 VXLAN 封装的检测报文分配一个目标出端口，通过为 VXLAN 封装的检测报文分配的目标出端口发送 VXLAN 封装的检测报文。

本步骤 b2 中，按照均衡分担原则从 L 个路径对应的出端口中为每一个检测报文分配一个目标出端口具体实现时可有多种实现形式，这里仅举一个实现形式，比如：将 N 份检测报文分组，分成的任一组中检测报文的数量小于等于 L，从 L 个出端口中为每一组中的检测报文分配不同出端口作为目标出端口。

通过步骤 b1 至步骤 b2 能够实现第一 VTEP 将 N 个 VXLAN 封装的检测报文分担至从本 VTEP 至第二 VTEP 的所有路径对应的出端口转发。

在 VXLAN 中，一般会有控制器控制管理 VTEP，基于此，当 VXLAN 中存在控制器时，作为本申请的一个示例，上述的路径检测设备可为控制器。

其中，当路径检测设备为控制器时，上述从第一 VTEP 至第二 VTEP 的转发路径是否可达的 N 个检测报文是控制器在确定检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时生成并向第一 VTEP 发送的。

当路径检测设备为控制器时，上述步骤 a1 确定用于检测从第一

VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数包括: 接收所述控制器发送的用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数。

作为本申请的另一个示例, 不管 VXLAN 中是否存在上述的控制器, 上述的路径检测设备也可为第一 VTEP。

当路径检测设备为第一 VTEP 时, 用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达的 N 个检测报文是第一 VTEP 生成的。

当路径检测设备为第一 VTEP 时, 上述步骤 a1 中确定用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数可包括: 从本地预先配置的所有 VXLAN 封装参数中确定用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数。

如图 2 所示的 VXLAN 封装的结构, 在本申请示例中, 上述的 VXLAN 封装参数包含外层 UDP 头中的外层 UDP 目的端口 (Dest Port)、外层以太头 (Outer Ethernet Header)、外层 IP 头 (Outer IP Header)、VXLAN 头 (VXLAN Header)。

在本申请示例中, 外层 UDP 头内外层 UDP 目的端口号 (Dest Port) 为 VXLAN 协议规定的设定端口号。在本申请示例中, 外层以太头的外层目的 MAC 地址为第二 VTEP 的桥 MAC 地址, 外层以太头的外层源 MAC 地址为第一 VTEP 的桥 MAC 地址, 外层 IP 头的外层目的 IP 地址为第二 VTEP 的 IP 地址、外层 IP 头的外层源 IP 地址为第一 VTEP 的 IP 地址。在本申请示例中, VXLAN 头内的指定保留字段 (比如图 2 所示的第一个保留字段) 内的数值为设定值比如 100, VXLAN 头内的 VXLAN 标识 (VNID) 为第一 VTEP 所处的 VXLAN 的标识比如 VNID200。

在本申请示例中,当第一 VTEP 发送 VXLAN 封装的检测报文后,处于从第一 VTEP 至第二 VTEP 之间的中间设备,以及第二 VTEP 会接收到 VXLAN 封装的检测报文,当中间设备或第二 VTEP 接收到 VXLAN 封装的检测报文后,可利用检测报文的 VXLAN 封装匹配本地保存的匹配规则,如果匹配成功,则向路径检测设备发送供所述路径检测设备确定从第一 VTEP 至第二 VTEP 的可达转发路径的路径参数。这里的匹配规则可为路径检测设备发送的,也可为预先配置的,本申请示例并不具体限定。下述两个示例均以路径检测设备发送匹配规则为例。具体地,结合上面描述的 VXLAN 封装,这里的匹配规则具体可为:外层 UDP 头内外层 UDP 目的端口号为 VXLAN 协议规定的设定端口号,且 VXLAN 头内的指定保留字段比如图 2 所示的第一个保留字段内的数值为设定值比如 100。

本申请示例中,中间设备发送的路径参数具体实现时至少可包括中间设备的设备标识、中间设备收到 VXLAN 封装的检测报文的端口的标识。并且,中间设备除了向路径检测设备发送路径参数外,还会继续按照 VXLAN 封装中外层 IP 头的目的 IP 地址选择用于转发检测报文的出端口转发接收的 VXLAN 封装的检测报文。其中,中间设备选择用于转发检测报文的出端口转发接收的 VXLAN 封装的检测报文具体为:按照 VXLAN 封装中外层 IP 头的目的 IP 地址确定用于转发检测报文的出端口,如果确定的出端口数量为 1,则直接选择该确定的出端口转发接收的 VXLAN 封装的检测报文;如果确定的出端口的数量大于 1,则对 VXLAN 封装中外层 UDP 头内 UDP 源端口号进行哈希运算,依据哈希运算结果从确定出的出端口中选择一个出端口转发接收的 VXLAN 封装的检测报文。

因为第一 VTEP 发送的检测报文数量 N 大于从第一 VTEP 至第二

VTEP 的路径数量 M ，这样，中间设备会接收到多个目的端为第二 VTEP 的检测报文，由于这多个检测报文的 VXLAN 封装中外层 UDP 头内的 UDP 源端口号不同，则多个检测报文将从不同的出端口转发，使得检测报文可以遍历不同的转发路径。

5 本申请示例中，第二 VTEP 发送的路径参数具体实现时至少可包括第二 VTEP 的设备标识、第二 VTEP 收到 VXLAN 封装的检测报文的端口的标识。并且，为防止影响正常的业务，第二 VTEP 会丢弃接收的 VXLAN 封装的检测报文。

需要说明的是，在实际应用中，可能存在同时检测多对 VTEP 之间的转发路径是否可达时的情况，比如同时存在如上所述的检测第一 VTEP 至第二 VTEP 之间的转发路径是否可达，以及检测第三 VTEP 至第四 VTEP 之间的转发路径是否可达，基于此种情况，为便于区分检测的转发路径，上述中间设备、第二 VTEP 发送的路径参数还可进一步包括第一 VTEP 的标识、第二 VTEP 的标识。也即，本申请示例
10 中，并不局限路径参数的具体内容，只要能够最终保证检测出从第一 VTEP 至第二 VTEP 之间的转发路径是否可达即可。

在本申请示例中，第一 VTEP 也可以接收到其他 VTEP 发送的 VXLAN 封装的检测报文，当第一 VTEP 接收到 VXLAN 封装的检测报文，则第一 VTEP 的处理过程类似上述的第二 VTEP，这里不再赘述。
20 述。

可以看出，本申请示例中，通过图 1 所示的流程能够遍历从第一 VTEP 至第二 VTEP 之间所有的转发路径是否可达。

下面通过示例对图所示流程进行描述：

25 图 3 为本申请提供的示例的应用组网图。在图 3 中，Overlay 网

络仅示出两个 VTEP: VTEP1 和 VTEP2, 控制器 30 控制 VTEP1 和 VTEP2, underlay 网络仅示出 5 个设备, 分别为设备 31 至设备 35。在图 3 中, 路径检测设备为控制器 30。

在图 3 中, 当控制器 30 确定需要检测两个 VTEP (以 VTEP1 为源端 VTEP、VTEP2 为目的端 VTEP 为例) 之间的转发路径可达情况, 则控制器 30 依据已收集的网络拓扑确定从 VTEP1 至 VTEP2 的转发路径数量, 假如转发路径数量为 5, 分别为: 转发路径 1, VTEP1->设备 31->设备 32->设备 34->VTEP2, 转发路径 2, VTEP1->设备 31->设备 32->设备 35->VTEP2, 转发路径 3, VTEP1->设备 31->设备 33->设备 34->VTEP2, 转发路径 4, VTEP1->设备 32->设备 35->VTEP2, 转发路径 5, VTEP1->设备 32->设备 34->VTEP2; 并且, 控制器 30 还依据确定的从 VTEP1 至 VTEP2 的转发路径数量 5 生成 5 份以上检测报文, 这里以生成 5 份检测报文为例。

控制器 30 将生成的 5 份检测报文发送给 VTEP1。本示例中以 5 份检测报文具有不同报文关键信息、且报文关键信息为目的 MAC 地址为例。假如 5 份检测报文分别记为报文 01, 报文 02, 报文 03、报文 04、报文 05, 其中, 报文 01 的目的 MAC 地址为 00-00-5E-90-00-01, 报文 02 的目的 MAC 地址为 00-00-5E-90-00-02, 报文 03 的目的 MAC 地址为 00-00-5E-90-00-03, 报文 04 的目的 MAC 地址为 00-00-5E-90-00-04, 报文 05 的目的 MAC 地址为 00-00-5E-90-00-05。

控制器 30 还将用于检测从 VTEP1 至 VTEP2 的转发路径是否可达时使用的 VXLAN 封装参数(VXLAN 封装中除了外层 UDP 头中外层 UDP 源端口之外的参数) 通知给 VTEP1。这里, VXLAN 封装参数中外层 UDP 头中 UDP 目的端口号为 VXLAN 协议规定的设定端口号(以端口号 200 为例), 外层以太头的外部目的 MAC 地址为 VTEP2

的桥 MAC 地址、外部源 MAC 地址为 VTEP1 的桥 MAC 地址，外层 IP 头的外层目的 IP 地址为 VTEP2 的 IP 地址、外层 IP 头的外层源 IP 地址为 VTEP1 的 IP 地址，VXLAN 头中 VXLAN 标识（VNID）为 VTEP1 所处的第一 VXLAN 的 ID（这里以 VNID100 为例），指定保留
5 留字段比如 VXLAN 头中第一个保留字段内的数值为设定值比如 100。

控制器 30 下发匹配规则给 VTEP2。

控制器 30 下发匹配规则给处于从 VTEP1 至 VTEP2 的任一转发路径上的中间设备（这里以中间设备为图 3 中的设备 31 至设备 35 为例，图 3 因为线路比较多，没有示出控制器 30 下发匹配规则给设备
10 31 至设备 35 的线路）。

这里，匹配规则可为：VXLAN 封装中 VXLAN 头内的指定保留字段内的数值为设定值，且 VXLAN 封装中外层 UDP 头内的 UDP 目的端口号为 VXLAN 协议规定的设定端口号。

VTEP1 接收控制器 30 发送的报文 01 至报文 05 共 5 份检测报文，
15 以及接收控制器 30 发送的 VXLAN 封装参数。

VTEP1 发现报文 01 至报文 05 具有不同报文关键信息、且报文关键信息为目的 MAC 地址为例，则 VTEP1 按照设定的 UDP 源端口号计算方法并利用报文 01 的目的 MAC 地址 00-00-5E-90-00-01 计算报文 01 的 UDP 源端口号，VTEP1 按照设定的 UDP 源端口号计算方法
20 法并利用报文 02 的目的 MAC 地址 00-00-5E-90-00-02 计算报文 02 的 UDP 源端口号；VTEP1 按照设定的 UDP 源端口号计算方法并利用报文 03 的目的 MAC 地址 00-00-5E-90-00-03 计算报文 03 的 UDP 源端口号；VTEP1 按照设定的 UDP 源端口号计算方法并利用报文 04 的目的 MAC 地址 00-00-5E-90-00-04 计算报文 04 的 UDP 源端口号；
25 VTEP1 按照设定的 UDP 源端口号计算方法并利用报文 05 的目的

MAC 地址 00-00-5E-90-00-05 计算报文 05 的 UDP 源端口号。

VTEP1 利用控制器 30 通知的 VXLAN 封装参数和报文 01 的 UDP 源端口号对报文 01 进行 VXLAN 封装，即 VXLAN 封装中外层 UDP 头内的 UDP 源端口号为报文 01 被确定的 UDP 源端口号，VXLAN 封装中其他参数如上述的 VXLAN 封装参数。同样，VTEP1 利用控制器 30 通知的 VXLAN 封装参数和报文 02 的 UDP 源端口号对报文 02 进行 VXLAN 封装；VTEP1 利用控制器 30 通知的 VXLAN 封装参数和报文 03 的 UDP 源端口号对报文 03 进行 VXLAN 封装；VTEP1 利用控制器 30 通知的 VXLAN 封装参数和报文 04 的 UDP 源端口号对报文 04 进行 VXLAN 封装；VTEP1 利用控制器 30 通知的 VXLAN 封装参数和报文 05 的 UDP 源端口号对报文 05 进行 VXLAN 封装。为便于描述，这里将进行了 VXLAN 封装的报文 01 至报文 05 依次记为报文 11、报文 12、报文 13、报文 14、报文 15。

VTEP1 确定从本 VTEP1 至 VTEP2 的路径对应的出端口。假如 VTEP1 确定出的路径对应的出端口的数量 L 为 2，分别为端口 11、端口 12，则 VTEP1 通过端口 11 发送报文 11，通过端口 12 发送报文 12，通过端口 11 发送报文 13，通过端口 12 发送报文 14、通过端口 11 发送报文 15。

处于从 VTEP1 至 VTEP2 任一路径的任一中间设备（这里仅以设备 31 为例）收到报文 11，发现报文 11 的 VXLAN 封装中 VXLAN 头内的指定保留字段内的数值为设定值 100，且 VXLAN 封装中外层 UDP 头内的 UDP 目的端口号为 VXLAN 协议规定的设定端口号 200，满足控制器下发的匹配规则，则上送路径参数至控制器，这里的路径参数至少包括：设备 31 的设备标识、设备 31 收到报文 11 的端口的端口号（以端口 311 为例）。

同时，设备 31 会依据报文 11 的 VXLAN 封装中外层 IP 头的外层目的 IP 地址确定到达该外层目的 IP 地址的路径，结果发现该路径不止一条，为三条等价路径，依次为：设备 31->设备 32->设备 34->VTEP2；设备 31->设备 32->设备 35->VTEP2，设备 31->设备 33->设备 34->VTEP2，基于此，设备 31 对报文 11 的 VXLAN 封装中外层 UDP 头的 UDP 源端口号进行哈希运算，依据哈希结果从三条等价路径中找到与哈希结果对应的一条路径比如设备 31->设备 33->设备 34->VTEP2 并通过找到的一条路径转发收到的报文 11。

设备 31 收到报文 13、收到报文 15 的处理类似收到报文 11 的处理，这里不再赘述。

控制器 30 接收并保存设备 31 发送的路径参数。

设备 32 至设备 35 收到报文后的处理类似设备 31 的处理，这里不再赘述。

最终，VTEP2 收到报文 11 至报文 15 中任一报文时，发现报文的 VXLAN 封装中 VXLAN 头内的第一个保留字段内的数值为设定值 100，且 VXLAN 封装中外层 UDP 头内的目的 UDP 端口号为 VXLAN 协议规定的设定端口号 200，满足控制器下发的匹配规则，则上送路径参数至控制器，这里的路径参数至少包括：VTEP2 的设备标识、VTEP2 收到报文的端口的标识。

VTEP2 发现 VXLAN 封装中外层 IP 头的外层目的 IP 地址为本 VTEP2 的 IP 地址，则丢弃接收的报文。

控制器 30 收到 VTEP2 上送的路径参数后，依据收到的所有路径参数对从 VTEP1 至 VTEP2 的可达路径进行汇总和统计。比如控制器 30 收到的路径参数为：

(设备 31、端口 311)、(设备 32、端口 321)、(设备 35、端口 352)、

(设备 32、端口 322)、(设备 34、端口 342)、(VTEP2、端口 25)、(VTEP2、端口 24)。

控制器 30 基于已收集的网络拓扑发现设备 31 上端口 311 对应的端口连接 VTEP1，如此，通过（设备 31、端口 311）确定从 VTEP1 至设备 31 的路径可达；
5 控制器 30 基于已收集的网络拓扑发现设备 32 上端口 321 对应的端口连接设备 31，如此，通过（设备 32、端口 321）确定从设备 31 至设备 32 的路径可达，结合之前确定的从 VTEP1 至设备 31 的路径可达，则此时进一步确定从 VTEP1 至设备 32 的路径可达；
控制器 30 基于已收集的网络拓扑发现设备 35 上端口 352 对应的端口连接设备 32，如此，通过（设备 35、端口 352）确定从设备 32 至设备 35 的路径可达，结合之前确定的从 VTEP1 至设备 32 的路径可达，则此时进一步确定从 VTEP1 至设备 35 的路径可达；
10 控制器 30 基于已收集的网络拓扑发现 VTEP2 上端口 25 对应的端口连接设备 35，如此，通过（VTEP2、端口 25）确定从设备 35 至 VTEP2 的路径可达，结合之前确定的从 VTEP1 至设备 35 的路径可达，则此时进一步确定从 VTEP1 至 VTEP2 的路径可达，即上述的转发路径 2 可达。

控制器 30 基于已收集的网络拓扑发现设备 32 上端口 322 对应的端口连接 VTEP1，如此，通过（设备 32、端口 322）确定从 VTEP1 至设备 32 的路径可达；
20 控制器 30 基于已收集的网络拓扑发现设备 34 上端口 342 对应的端口连接设备 32，如此，通过（设备 34、端口 342）确定从设备 32 至设备 34 的路径可达，结合之前确定的从 VTEP1 至设备 32 的路径可达，则此时进一步确定从 VTEP1 至设备 34 的路径可达；
控制器 30 基于已收集的网络拓扑发现 VTEP2 上端口 24 对应的端口连接设备 34，如此，通过（VTEP2、端口 24）确定从设备 34 至 VTEP2 的路径可达，结合之前确定的从 VTEP1 至设备 34 的路径可达，则此时进一
25

步确定从 VTEP1 至 VTEP2 的路径可达，即上述的转发路径 5 可达。

也即，控制器 30 通过对收到的路径参数进行汇总和统计确定出只有上述的转发路径 2 和转发路径 5 可达，而上述的转发路径 1、3、4 不可达。

5 图 4 为本申请提供的示例的应用组网示意图。在图 4 中，Overlay 网络仅示出两个 VTEP：VTEP3 和 VTEP4，underlay 网络仅示出 5 个设备，分别为设备 61 至设备 65。在图 4 中，以检测从 VTEP3 至 VTEP4 的转发路径是否可达为例。在图 4 中，路径检测设备为 VTEP3。

VTEP3 依据之前已收集的网络拓扑确定从本 VTEP3 至 VTEP4
10 的转发路径数量，假如转发路径数量为 5，分别为：转发路径 1，VTEP3->设备 61->设备 62->设备 64->VTEP4，转发路径 2，VTEP3->设备 61->设备 62->设备 65->VTEP4，转发路径 3，VTEP3->设备 61->设备 63->设备 64->VTEP4，转发路径 4，VTEP3->设备 62->设备 65->VTEP4，转发路径 5，VTEP3->设备 62->设备 64->VTEP4。

15 VTEP3 依据确定的从本 VTEP3 至 VTEP4 的转发路径数量 5 生成 5 份以上检测报文，这里以生成 5 份检测报文为例。本示例中以 5 份检测报文具有不同报文关键信息、且报文关键信息为目的 MAC 地址为例。假如 5 份检测报文分别记为报文 61，报文 62，报文 63、报文 64、报文 65，其中，报文 61 的目的 MAC 地址为 00-00-5E-90-00-06，
20 报文 62 的目的 MAC 地址为 00-00-5E-90-00-07，报文 63 的目的 MAC 地址为 00-00-5E-90-00-08，报文 64 的目的 MAC 地址为 00-00-5E-90-00-09，报文 65 的目的 MAC 地址为 00-00-5E-90-00-10。

VTEP3 发送匹配规则给 VTEP4。

VTEP3 依据之前已收集的网络拓扑确定从本 VTEP3 至 VTEP4
25 的路径上的中间设备（即图 4 中的设备 61 至设备 65），VTEP3 发送

匹配规则给图 4 中的设备 61 至设备 65 (图 4 因为线路比较多, 没有示出发送匹配规则的线路)。

这里, 匹配规则可为: VXLAN 封装中 VXLAN 头内的第一个保留字段内的数值为设定值, 且 VXLAN 封装中外层 UDP 头内的 UDP
5 目的端口号为 VXLAN 协议规定的设定端口号。

VTEP3 按照设定的 UDP 源端口号计算方法分别确定报文 61 至报文 65 的 UDP 源端口号。

VTEP3 对报文 61 至报文 65 分别进行 VXLAN 封装; 报文 61 至报文 65 的 VXLAN 封装中外层 UDP 头的 UDP 源端口号分别为报文
10 61 至报文 65 被确定的 UDP 源端口号。为便于描述, 这里将进行了 VXLAN 封装的报文 61 至报文 65 依次记为报文 71 至报文 75。

VTEP3 确定从本 VTEP3 至 VTEP4 的路径对应的出端口。假如 VTEP3 确定出的出端口的数量 L 为 2, 分别为端口 31、端口 32 (图 4 未示出), 则 VTEP3 通过端口 31 发送报文 71, 通过端口 32 发送报文
15 72, 通过端口 31 发送报文 73, 通过端口 32 发送报文 74、通过端口 31 发送报文 75。

处于从 VTEP3 至 VTEP4 任一路径的任一中间设备 (这里仅以设备 61 为例) 收到报文 71, 发现报文 71 的 VXLAN 封装中 VXLAN 头内的第一个保留字段内的数值为设定值, 且 VXLAN 封装中外层 UDP
20 头内的 UDP 目的端口号为 VXLAN 协议规定的设定端口号, 满足 VTEP3 发送的匹配规则, 则发送路径参数至 VTEP3, 路径参数包括: 设备 61 的设备标识、设备 61 收到报文 71 的端口的端口号。

同时, 设备 61 会依据报文 71 的 VXLAN 封装中外层 IP 头的外层目的 IP 地址确定到达该外层目的 IP 地址的路径, 结果发现该路径
25 不止一条, 为三条等价路径, 基于此, 设备 61 对报文 71 的 VXLAN

封装中外层 UDP 头的 UDP 源端口号进行哈希运算,依据哈希结果从三条等价路径中找到与哈希结果对应的一条路径比如设备 61->设备 63->设备 64->VTEP4 并通过找到的一条路径转发收到的报文 71。

设备 61 收到报文 73、收到报文 75 的处理类似收到报文 71 的处理,
5 这里不再赘述。

VTEP3 接收并保存设备 61 发送的路径参数。

设备 62 至设备 65 收到报文后的处理类似设备 61 的处理,这里不再赘述。

最终, VTEP4 收到报文 71 至报文 75 中任一报文时,发现报文的
10 VXLAN 封装中 VXLAN 头内的第一个保留字段内的数值为设定值,且 VXLAN 封装中外层 UDP 头内的 UDP 目的端口号为 VXLAN 协议规定的设定端口号,满足 VTEP3 发送的匹配规则,则发送路径参数至 VTEP3,并发现 VXLAN 封装中外层 IP 头的外层目的 IP 地址为本 VTEP4 的 IP 地址,则丢弃接收的报文。

15 VTEP3 收到 VTEP4 发送的路径参数后,依据收到的所有路径参数对从 VTEP3 至 VTEP4 的可达转发路径进行汇总和统计,这里的汇总和统计的方式类似前述示例的方法,不再赘述,图 4 不具体再示出各个设备收到 VXLAN 封装的检测报文的端口。最终 VTEP3 会全局统计出从 VTEP3 至 VTEP4 的所有路径可达情况包括有多少可达路径等,以供
20 上层查询。这里 VTEP3 最终全局统计出从 VTEP3 至 VTEP4 的所有路径可达情况可能没有控制器统计出的所有路径可达情况直观。

以上对本申请示例提供的方法进行了描述。下面对本申请示例提供的装置进行描述:

参见图 5,图 5 为本申请示例提供的装置结构示意图。如图 5 所示,
25 该装置 50 包括:

确定单元 51, 用于按照不同检测报文对应不同 UDP 源端口号的原则, 为第一 VTEP 待向第二 VTEP 发送的用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达的 N 个检测报文确定 UDP 源端口号; 其中, N 大于或等于从第一 VTEP 至第二 VTEP 的路径数量 M。

- 5 VXLAN 封装单元 52, 用于针对每一个检测报文, 利用为该检测报文确定的 UDP 源端口号对该检测报文进行 VXLAN 封装;

检测报文发送单元 53, 用于通过从第一 VTEP 至第二 VTEP 的所有路径对应的出端口分担发送 N 个 VXLAN 封装的检测报文, 以使接收到 VXLAN 封装的检测报文的中间设备或第二 VTEP 向路径检测设备发送供所述路径检测设备确定从第一 VTEP 至第二 VTEP 的可达转发路径的路径参数。

10

根据一个示例, 如果所述路径检测设备为 VXLAN 中用于管理 VTEP 的控制器, 所述 N 个用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达的检测报文是所述控制器向第一 VTEP 发送的;

- 15 如果所述路径检测设备为第一 VTEP, 所述 N 个用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达的检测报文是所述第一 VTEP 生成的。

根据一个示例, 所述 VXLAN 封装单元 52 针对每一个检测报文, 利用为该检测报文确定的 UDP 源端口号对该检测报文进行 VXLAN 封装包括:

20

确定用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数;

针对每一个检测报文, 利用为该检测报文确定的 UDP 源端口号和所述 VXLAN 封装参数对该检测报文进行 VXLAN 封装。

- 25 根据一个示例, 如果所述路径检测设备为 VXLAN 中用于管理

VTEP 的控制器，所述确定用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数包括：接收所述控制器发送的用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数；

5 如果所述路径检测设备为第一 VTEP，所述确定用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数包括：从本地预先配置的所有 VXLAN 封装参数中确定用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数。

10 根据一个示例，所述检测报文发送单元 53 通过从第一 VTEP 至第二 VTEP 的所有路径对应的出端口分担发送 N 个 VXLAN 封装的检测报文包括：

确定从本 VTEP 至第二 VTEP 的所有路径对应的出端口；

在确定的路径对应的出端口的数量 L 为 1 时，通过该确定出的 1 个
15 路径对应的出端口发送 N 个 VXLAN 封装的检测报文；

在确定的路径对应的出端口的数量 L 大于 1 时，按照均衡分担原则从 L 个路径对应的出端口中为每一个 VXLAN 封装的检测报文分配一个目标出端口，通过为 VXLAN 封装的检测报文分配的目标出端口发送 VXLAN 封装的检测报文。

20 根据一个示例，如图 5 所示，该装置 50 进一步包括：

接收单元 54，用于接收 VXLAN 封装的检测报文；

路径参数发送单元 55，用于在所述接收单元 54 接收 VXLAN 封装的检测报文时，向路径检测设备发送接收到的 VXLAN 封装的检测报文对应的路径参数，所述接收到的 VXLAN 封装的检测报文对应的路
25 径参数至少包括：第一 VTEP 的设备标识、第一 VTEP 收到 VXLAN

封装的检测报文的端口的标识;

处理单元 56, 用于确定 VXLAN 封装中外层 IP 头的目的 IP 地址为本 VTEP 的 IP 地址, 丢弃所述接收单元接收的 VXLAN 封装的检测报文。

5

参见图 6, 图 6 为本申请示例提供的装置的硬件结构示意图。该装置 60 为第一 VTEP, 如图 6 所示, 装置 60 包括: 处理器 61 和存储器 62。处理器 61 例如为中央处理单元 (CPU)。存储器 62 可以为非易失性存储介质, 存储有计算机可执行指令, 该计算机可执行指令可以具体为路径可达检测指令。

10

处理器 61 读取该路径可达检测指令, 以执行:

按照不同检测报文对应不同 UDP 源端口号的原则, 为第一 VTEP 待向第二 VTEP 发送的用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达的 N 个检测报文确定 UDP 源端口号;

15

针对每一个检测报文, 利用为该检测报文确定的 UDP 源端口号对该检测报文进行 VXLAN 封装;

20

通过从第一 VTEP 至第二 VTEP 的所有路径对应的出端口分担发送 N 个 VXLAN 封装的检测报文, 以使接收到 VXLAN 封装的检测报文的中间设备或第二 VTEP 向路径检测设备发送供所述路径检测设备确定从第一 VTEP 至第二 VTEP 的可达转发路径的路径参数;

其中, N 大于或等于从第一 VTEP 至第二 VTEP 的路径数量 M。

如果所述路径检测设备为 VXLAN 中用于管理 VTEP 的控制器, 所述用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达的 N 个检测报文是所述控制器向第一 VTEP 发送的;

25

如果所述路径检测设备为第一 VTEP, 所述 N 个用于检测从第一

VTEP 至第二 VTEP 的转发路径是否可达的检测报文是所述第一 VTEP 生成的。

所述针对每一个检测报文，利用为该检测报文确定的 UDP 源端口号对该检测报文进行 VXLAN 封装包括：

- 5 确定用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数；

针对每一个检测报文，利用为该检测报文确定的 UDP 源端口号和所述 VXLAN 封装参数对该检测报文进行 VXLAN 封装。

- 10 如果所述路径检测设备为 VXLAN 中用于管理 VTEP 的控制器，所述确定用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数包括：接收所述控制器发送的用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数；

- 15 如果所述路径检测设备为第一 VTEP，所述确定用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数包括：从本地预先配置的所有 VXLAN 封装参数中确定用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数。

所述通过从第一 VTEP 至第二 VTEP 的所有路径对应的出端口分担发送 N 个 VXLAN 封装的检测报文包括：

- 20 确定从第一 VTEP 至第二 VTEP 的所有路径对应的出端口；

在确定的路径对应的出端口的数量 L 为 1 时，通过该确定出的 1 个路径对应的出端口发送 N 个 VXLAN 封装的检测报文；

- 25 在确定的路径对应的出端口的数量 L 大于 1 时，按照均衡分担原则从 L 个路径对应的出端口中为每一个 VXLAN 封装的检测报文分配一个目标出端口，通过为 VXLAN 封装的检测报文分配的目标出端口发送

VXLAN 封装的检测报文。

处理器 61 读取路径可达检测指令，还执行：接收 VXLAN 封装的检测报文；

5 在接收 VXLAN 封装的检测报文时，向路径检测设备发送接收到的 VXLAN 封装的检测报文对应的路径参数，所述接收到的 VXLAN 封装的检测报文对应的路径参数至少包括：第一 VTEP 的设备标识、第一 VTEP 收到 VXLAN 封装的检测报文的端口的标识；

确定 VXLAN 封装中外层 IP 头的目的 IP 地址为第一 VTEP 的 IP 地址，丢弃所述接收的 VXLAN 封装的检测报文。

10 其中，通过以下方式确定 UDP 源端口号：

针对每一个检测报文，按照设定的 UDP 源端口号确定算法对该检测报文具有的报文关键信息进行计算，将得到的结果确定为该检测报文的 UDP 源端口号；或者

15 将 N 个检测报文排成一个队列，为队列中的每一个检测报文确定不同的预设端口号。

以上所述仅为本申请的示例而已，并不用以限制本申请，凡在本申请的精神和原则之内，所做的任何修改、等同替换、改进等，均应包含在本申请保护的范围之内。

权利要求书

1、一种虚拟扩展局域网 VXLAN 隧道端点 VTEP 之间的路径可达检测方法，其特征在于，该方法包括：

第一 VTEP 按照不同检测报文对应不同 UDP 源端口号的原则，为待向第二 VTEP 发送的用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达的 N 个检测报文确定 UDP 源端口号；

第一 VTEP 针对每一个检测报文，利用为该检测报文确定的 UDP 源端口号对该检测报文进行 VXLAN 封装；

第一 VTEP 通过从本 VTEP 至第二 VTEP 的所有路径对应的出端口分担发送 N 个 VXLAN 封装的检测报文，以使接收到 VXLAN 封装的检测报文的中间设备或第二 VTEP 向路径检测设备发送供所述路径检测设备确定从第一 VTEP 至第二 VTEP 的可达转发路径的路径参数；

其中，N 大于或等于从第一 VTEP 至第二 VTEP 的路径数量 M。

2、根据权利要求 1 所述的方法，其特征在于，

所述路径检测设备为 VXLAN 中用于管理 VTEP 的控制器，所述用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达的 N 个检测报文是所述控制器向第一 VTEP 发送的；或者

所述路径检测设备为第一 VTEP，所述用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达的 N 个检测报文是所述第一 VTEP 生成的。

3、根据权利要求 1 所述的方法，其特征在于，所述第一 VTEP 针对每一个检测报文，利用为该检测报文确定的 UDP 源端口号对该检测报文进行 VXLAN 封装包括：

确定用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时

使用的 VXLAN 封装参数;

针对每一个检测报文,利用为该检测报文确定的 UDP 源端口号和所述 VXLAN 封装参数对该检测报文进行 VXLAN 封装。

- 4、根据权利要求 3 所述的方法,其特征在于,所述路径检测设备为 VXLAN 中用于管理 VTEP 的控制器,所述确定用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数包括:接收所述控制器发送的用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数; 或者

所述路径检测设备为第一 VTEP,所述确定用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数包括:从本地预先配置的所有 VXLAN 封装参数中确定用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数。

- 5、根据权利要求 1 所述的方法,其特征在于,所述第一 VTEP 通过从本 VTEP 至第二 VTEP 的所有路径对应的出端口分担发送 N 个 VXLAN 封装的检测报文包括:

第一 VTEP 确定从本 VTEP 至第二 VTEP 的所有路径对应的出端口;

在确定的路径对应的出端口的数量 L 为 1 时,通过该确定出的 1 个路径对应的出端口发送 N 个 VXLAN 封装的检测报文;

- 20 在确定的路径对应的出端口的数量 L 大于 1 时,按照均衡分担原则从 L 个路径对应的出端口中为每一个 VXLAN 封装的检测报文分配一个目标出端口,通过为 VXLAN 封装的检测报文分配的目标出端口发送 VXLAN 封装的检测报文。

- 6、根据权利要求 1 所述的方法,其特征在于,该方法进一步包括:

第一 VTEP 接收 VXLAN 封装的检测报文;

第一 VTEP 向路径检测设备发送接收到的 VXLAN 封装的检测报文对应的路径参数, 所述接收到的 VXLAN 封装的检测报文对应的路径参数至少包括: 第一 VTEP 的设备标识、第一 VTEP 收到 VXLAN
5 封装的检测报文的端口的标识;

第一 VTEP 确定 VXLAN 封装中外层 IP 头的目的 IP 地址为本 VTEP 的 IP 地址, 丢弃所述接收到的 VXLAN 封装的检测报文。

7、根据权利要求 1 所述的方法, 其中, 通过以下方式确定 UDP 源端口号:

10 针对每一个检测报文, 按照设定的 UDP 源端口号确定算法对该检测报文具有的报文关键信息进行计算, 将得到的结果确定为该检测报文的 UDP 源端口号; 或者

将 N 个检测报文排成一个队列, 为队列中的每一个检测报文确定不同的预设端口号。

15 8、一种虚拟扩展局域网 VXLAN 隧道端点 VTEP 之间的路径可达检测装置, 其特征在于, 所述装置为第一 VTEP, 该装置包括:

处理器;

非易失性存储介质, 存储有计算机可执行指令, 所述处理器读取所述计算机可执行指令, 以执行:

20 按照不同检测报文对应不同 UDP 源端口号的原则, 为第一 VTEP 待向第二 VTEP 发送的用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达的 N 个检测报文确定 UDP 源端口号;

针对每一个检测报文, 利用为该检测报文确定的 UDP 源端口号对该检测报文进行 VXLAN 封装;

25 通过从第一 VTEP 至第二 VTEP 的所有路径对应的出端口分担发

送 N 个 VXLAN 封装的检测报文,以使接收到 VXLAN 封装的检测报文的中间设备或第二 VTEP 向路径检测设备发送供所述路径检测设备确定从第一 VTEP 至第二 VTEP 的可达转发路径的路径参数;

其中, N 大于或等于从第一 VTEP 至第二 VTEP 的路径数量 M。

5 9、根据权利要求 8 所述的装置,其特征在于,

所述路径检测设备为 VXLAN 中用于管理 VTEP 的控制器,所述用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达的 N 个检测报文是所述控制器向第一 VTEP 发送的;或者

10 所述路径检测设备为第一 VTEP,所述 N 个用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达的检测报文是所述第一 VTEP 生成的。

10、根据权利要求 8 所述的装置,其特征在于,所述针对每一个检测报文,利用为该检测报文确定的 UDP 源端口号对该检测报文进行 VXLAN 封装包括:

15 确定用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数;

针对每一个检测报文,利用为该检测报文确定的 UDP 源端口号和所述 VXLAN 封装参数对该检测报文进行 VXLAN 封装。

20 11、根据权利要求 10 所述的装置,其特征在于,所述路径检测设备为 VXLAN 中用于管理 VTEP 的控制器,所述确定用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数包括:接收所述控制器发送的用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数;或者

25 所述路径检测设备为第一 VTEP,所述确定用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数包括:

从本地预先配置的所有 VXLAN 封装参数中确定用于检测从第一 VTEP 至第二 VTEP 的转发路径是否可达时使用的 VXLAN 封装参数。

12、根据权利要求 8 所述的装置，其特征在于，所述通过从第一 VTEP 至第二 VTEP 的所有路径对应的出端口分担发送 N 个 VXLAN

5 封装的检测报文包括：

确定从第一 VTEP 至第二 VTEP 的所有路径对应的出端口；

在确定的路径对应的出端口的数量 L 为 1 时，通过该确定出的 1 个路径对应的出端口发送 N 个 VXLAN 封装的检测报文；

10 在确定的路径对应的出端口的数量 L 大于 1 时，按照均衡分担原则从 L 个路径对应的出端口中为每一个 VXLAN 封装的检测报文分配一个目标出端口，通过为 VXLAN 封装的检测报文分配的目标出端口发送 VXLAN 封装的检测报文。

13、根据权利要求 8 所述的装置，其特征在于，所述处理器读取所述计算机可执行指令，还执行：

15 接收 VXLAN 封装的检测报文；

向路径检测设备发送接收到的 VXLAN 封装的检测报文对应的路径参数，所述接收到的 VXLAN 封装的检测报文对应的路径参数至少包括：第一 VTEP 的设备标识、第一 VTEP 收到 VXLAN 封装的检测报文的端口的标识；

20 确定 VXLAN 封装中外层 IP 头的目的 IP 地址为第一 VTEP 的 IP 地址，丢弃所述接收的 VXLAN 封装的检测报文。

14、根据权利要求 8 所述的装置，其特征在于，通过以下方式确定 UDP 源端口号：

25 针对每一个检测报文，按照设定的 UDP 源端口号确定算法对该检测报文具有的报文关键信息进行计算，将得到的结果确定为该检测

报文的 UDP 源端口号；或者

将 N 个检测报文排成一个队列，为队列中的每一个检测报文确定不同的预设端口号。

1/6

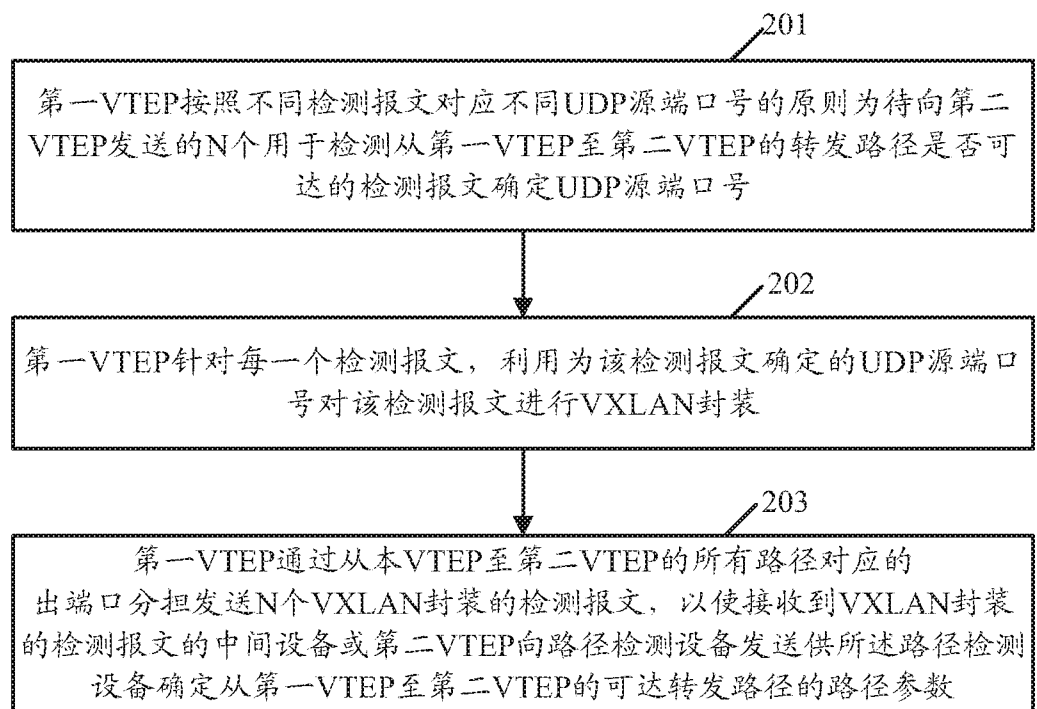


图 1

2/6

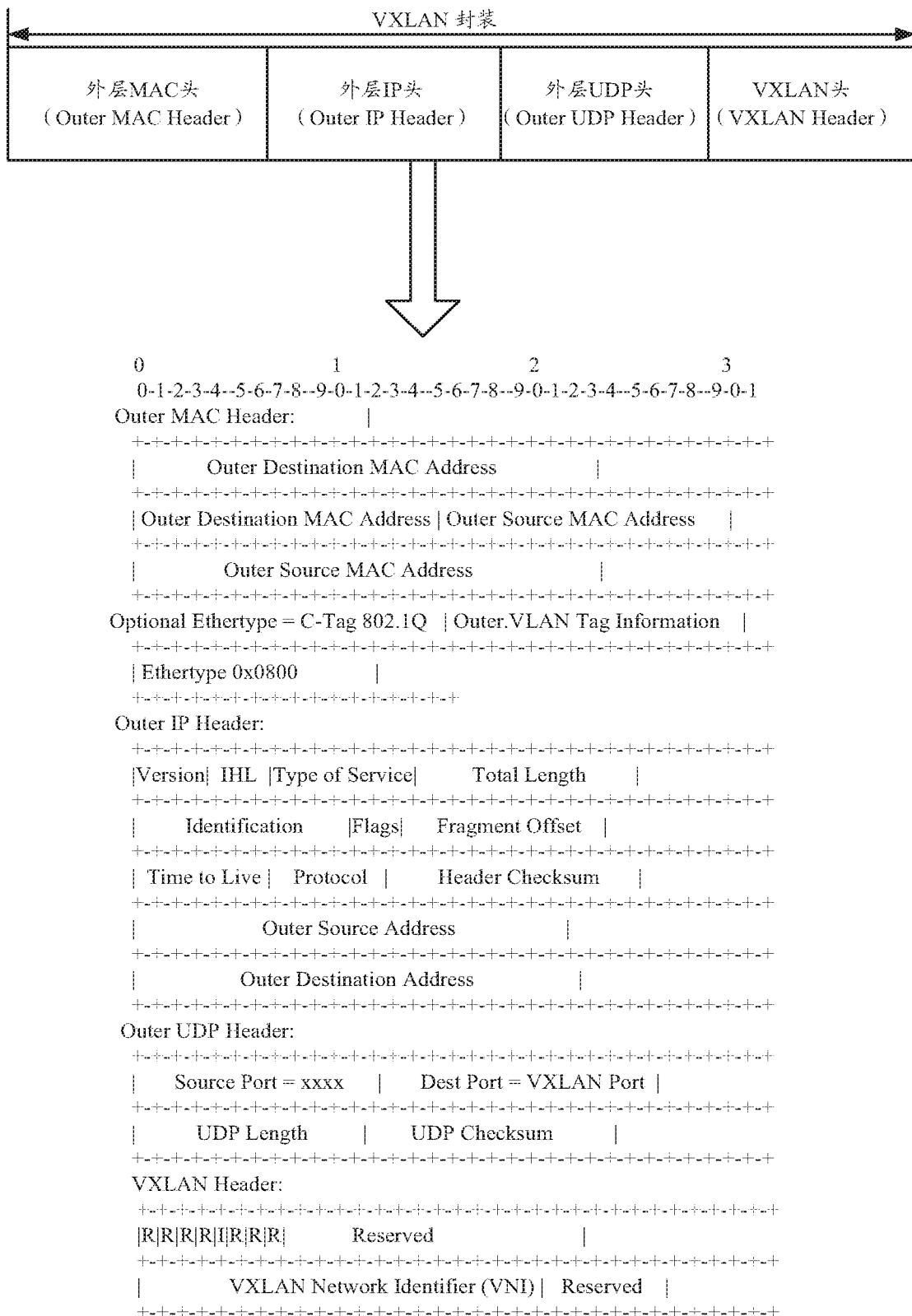


图 2

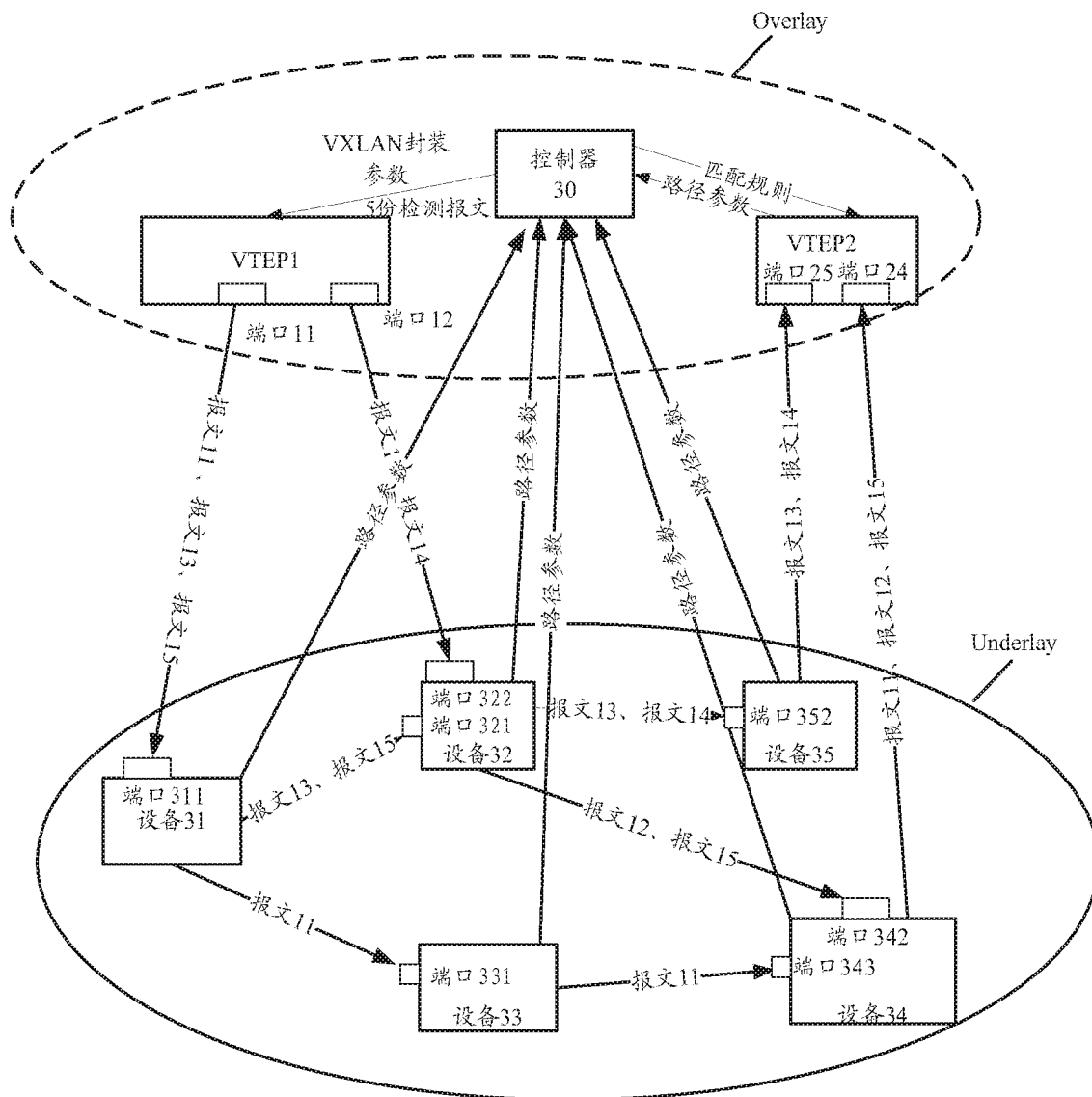


图 3

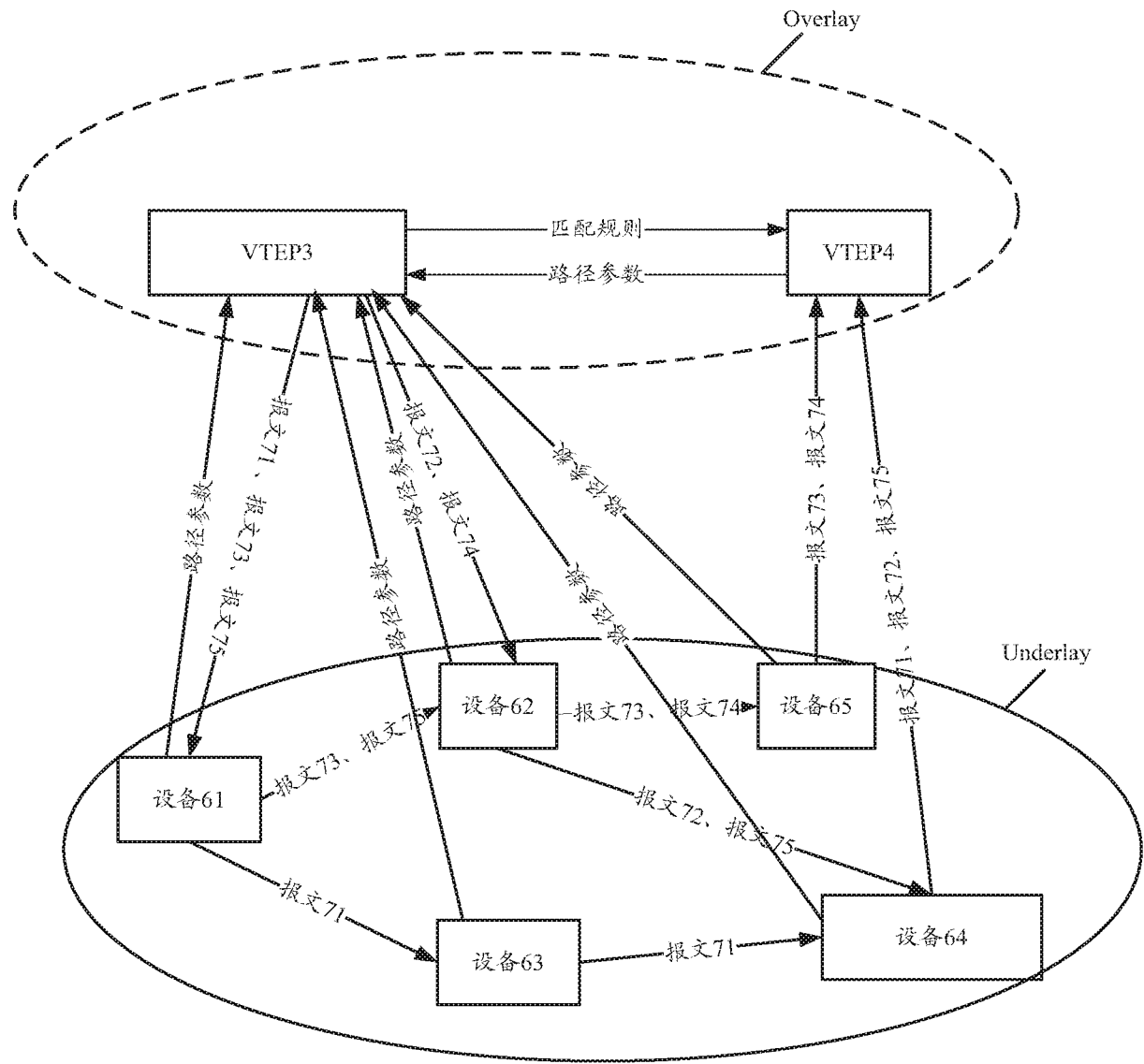


图 4

5/6

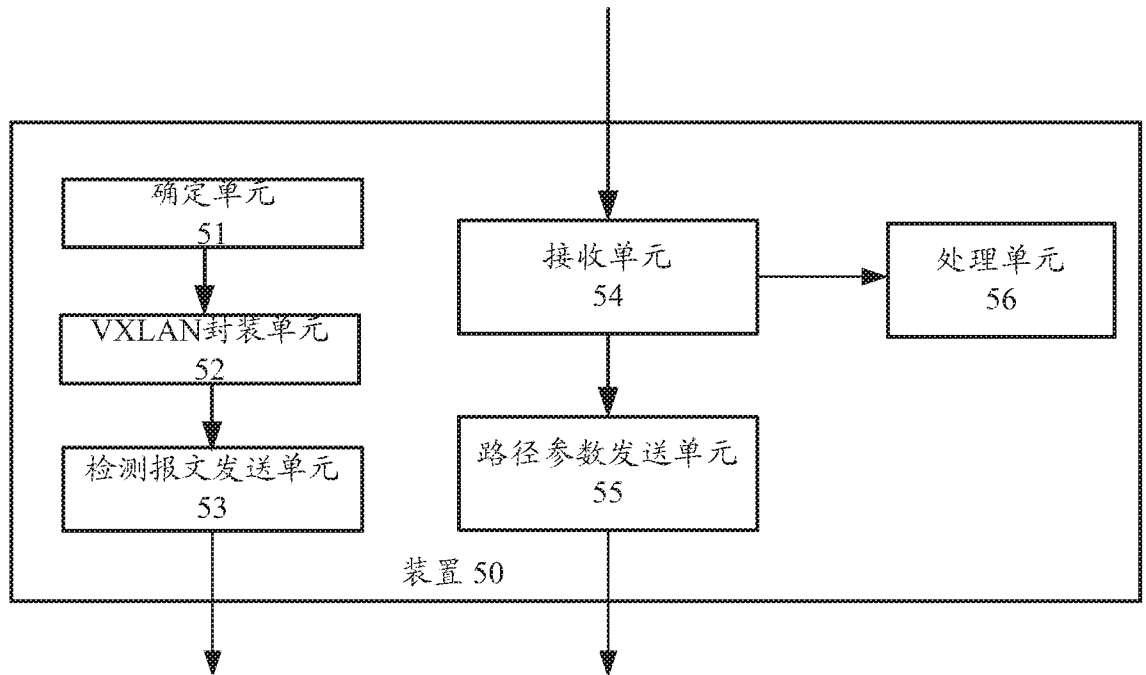


图 5

6/6

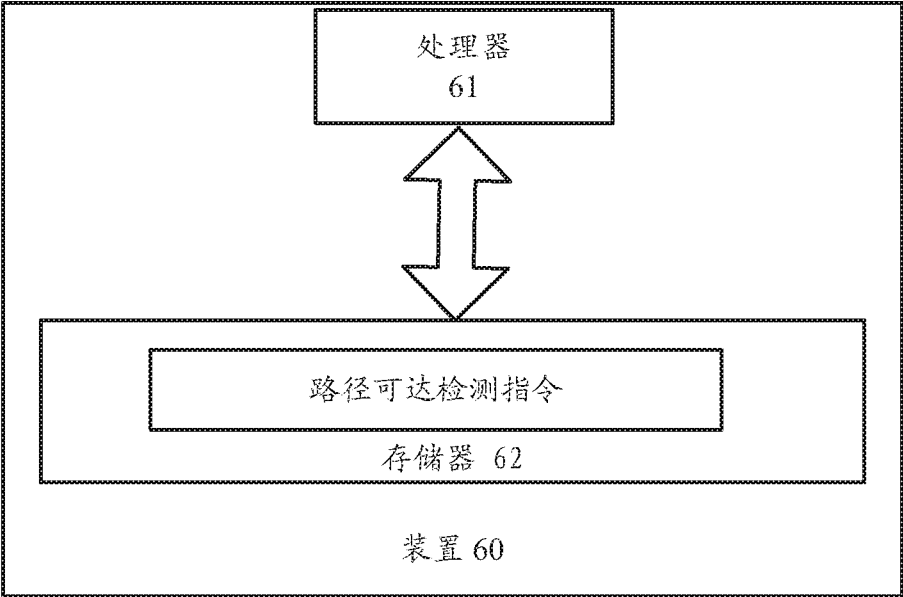


图 6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2016/089597

A. CLASSIFICATION OF SUBJECT MATTER

H04L 12/26 (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)

CNPAT, CNKI, WPI, EPODOC, Google: VXLAN, VTEP, UDP, path, link, failure, fault, disconnect, connectivity, detect, sense, measure, source, port number

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	PANG, Junying et al., Path Detection in VXLAN Overlay Network, draft-pang-nvo3-vxlan-path-detection-00, Network Working Group INTERNET-DRAFT, 22 May 2015 (22.05.2015), section 4	1-14
Y	JAIN P. et al., "Detecting VXLAN Segment Failure, draft-jain-nvo3-vxlan-ping-00 " Network Working Group INTERNET-DRAFT, 06 June 2013 (06.06.2013), section 6.1	1-14
A	CN 102177681 A (HUAWEI TECHNOLOGIES) 07 September 2011 (07.09.2011) the whole document	1-14
A	CN 104283738 A (HANGZHOU H3C TECHNOLOGIES) 14 January 2015 (14.01.2015) the whole document	1-14

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	"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
"A" document defining the general state of the art which is not considered to be of particular relevance	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"&" document member of the same patent family
"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	

Date of the actual completion of the international search 28 September 2016	Date of mailing of the international search report 13 October 2016
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer SUN, Shurong Telephone No. (86-10) 62413335

INTERNATIONAL SEARCH REPORTInternational application No.
PCT/CN2016/089597

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2014348006 A1 (ALCATEL-LUCENT USA INC.) 27 November 2014 (27.11.2014) the whole document	1-14

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2016/089597

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 102177681 A	07 September 2011	EP 2611075 A2	03 July 2013
		WO 2011110118 A2	15 September 2011
		US 2013286859 A1	31 October 2013
CN 104283738 A	14 January 2015	None	
US 2014348006 A1	27 November 2014	EP 3005627 A1	13 April 2016
		WO 2014189709 A1	27 November 2014
		JP 2016522645 A	28 July 2016

A. 主题的分类 H04L 12/26 (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)) CNPAT, CNKI, WPI, EPDOC, Google: VXLAN, VTEP, 路径, 链路, 故障, 断路, 连通性, 检测, 探测, UDP, 源, 端口号, path, link, failure, fault, disconnect, connectivity, detect, sense, measure, source, port number		
C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
Y	PANG, Junying等. "Path Detection in VXLAN Overlay Network, draft-pang-nvo3-vxlan-path-detection-00" Network Working Group INTERNET-DRAFT, 2015年 5月 22日 (2015 - 05 - 22), 第4节	1-14
Y	JAIN P. 等. "Detecting VXLAN Segment Failure, draft-jain-nvo3-vxlan-ping-00" Network Working Group INTERNET-DRAFT, 2013年 6月 6日 (2013 - 06 - 06), 第6.1节	1-14
A	CN 102177681 A (华为技术有限公司) 2011年 9月 7日 (2011 - 09 - 07) 全文	1-14
A	CN 104283738 A (杭州华三通信技术有限公司) 2015年 1月 14日 (2015 - 01 - 14) 全文	1-14
A	US 2014348006 A1 (ALCATEL-LUCENT USA INC.) 2014年 11月 27日 (2014 - 11 - 27) 全文	1-14
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。		
* 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 2016年 9月 28日		国际检索报告邮寄日期 2016年 10月 13日
ISA/CN的名称和邮寄地址 中华人民共和国国家知识产权局 (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		受权官员 孙淑蓉 电话号码 (86-10) 62413335

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2016/089597

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	102177681	A	2011年 9月 7日	EP	2611075	A2	2013年 7月 3日
				WO	2011110118	A2	2011年 9月 15日
				US	2013286859	A1	2013年 10月 31日
CN	104283738	A	2015年 1月 14日	无			
US	2014348006	A1	2014年 11月 27日	EP	3005627	A1	2016年 4月 13日
				WO	2014189709	A1	2014年 11月 27日
				JP	2016522645	A	2016年 7月 28日