(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 November 2016 (10.11.2016)

WIPO | PCT

(10) International Publication Number
WO 2016/177314 A1

(51) International Patent Classification:
H04L 12/741 (2013.01)

(21) International Application Number:
PCT/CN2016/080932

(22) International Filing Date:
4 May 2016 (04.05.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
201510220901.6 4 May 2015 (04.05.2015) CN

(71) Applicant: **HANGZHOU H3C TECHNOLOGIES CO., LTD.** [CN/CN]; 466 Changhe Road, Binjiang District, Hangzhou, Zhejiang 310052 (CN).

(72) Inventors: **ZHANG, Yang**; Room 730, Oriental Electronic BLD., NO.2, Chuangye Road, Shangdi Information Industry Base, Haidian District, Beijing 100085 (CN). **KUANG, Ling**; Room 730, Oriental Electronic BLD., NO.2, Chuangye Road, Shangdi Information Industry Base, Haidian District, Beijing 100085 (CN). **WANG, Minghui**; Room 730, Oriental Electronic BLD., NO.2, Chuangye Road, Shangdi Information Industry Base, Haidian District, Beijing 100085 (CN). **ZHANG, Hongyuan**; Room 730, Oriental Electronic BLD., NO.2, Chuangye Road, Shangdi Information Industry Base, Haidian District, Beijing 100085 (CN). **HU, Hai**; Room 730, Oriental Electronic BLD., NO.2, Chuangye Road, Shangdi Information Industry Base, Haidian District, Beijing 100085 (CN). **WEN, Guangliang**; Room 730, Oriental Electronic BLD., NO.2, Chuangye Road, Shangdi Information Industry Base, Haidian District, Beijing 100085 (CN).

(74) Agent: **DEQI INTELLECTUAL PROPERTY LAW CORPORATION**; 7/F, Xueyuan International Tower, NO.1 Zhichun Road, Haidian District, Beijing 100083 (CN).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: 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) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU,

[Continued on next page]

(54) Title: PACKET FORWARDING

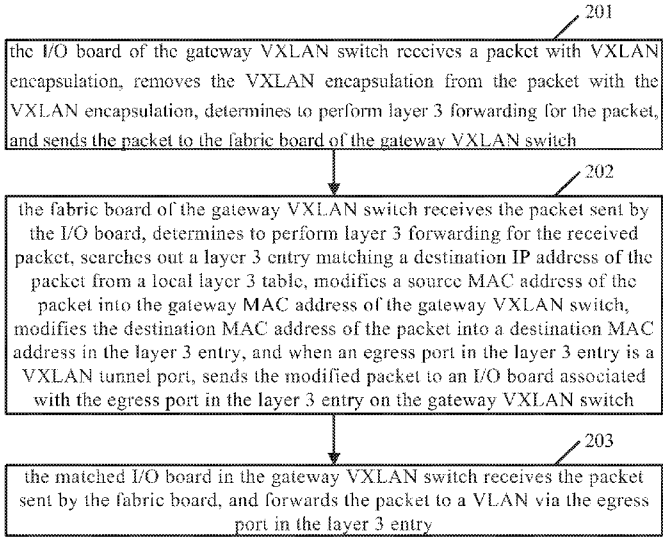

FIG. 1

(57) Abstract: I/O board of a VXLAN switch removes VXLAN encapsulation from a packet, determines to perform layer 3 forwarding for the packet and sends the packet to a fabric board of the VXLAN switch. The fabric board determines to perform layer 3 forwarding for the packet, searches out a layer 3 entry matching a destination IP address of the packet from a local layer 3 table, modifies a source MAC address of the packet into a gateway MAC address of the VXLAN switch, modifies a destination MAC address of the packet into a destination MAC address in the layer 3 entry, and when an egress port in the layer 3 entry is a VLAN port, sends the packet to an I/O board associated with the egress port. The I/O board associated with the egress port forwards the packet to a VLAN via the egress port in the layer 3 entry.

WO 2016/177314 A1

TJ, TM), European (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).

Published:

— *with international search report (Art. 21(3))*

PACKET FORWARDING

Background

[0001] Virtual eXtensible Local Area Network (VXLAN) is a layer 2 Virtual Private Network (VPN) technology based on an Internet Protocol (IP) network and adopting "Media Access Control (MAC) in User Data Protocol (UDP)" encapsulation. The VXLAN may implement layer 2 interconnection among distributed physical sites based on service providers or enterprise IP networks, and may provide service isolation for different tenants. The VXLAN may be applied to a data center network.

Brief Description of the Drawings

[0002] FIG. 1 is a flowchart illustrating a method for forwarding a packet from a VXLAN to a VLAN according to some examples of the present disclosure.

[0003] FIG. 2 is a flowchart illustrating a method for sending a packet to a fabric board of by an I/O board according to some examples of the present disclosure.

[0004] FIG. 3 is a flowchart illustrating a method for sending to a matched I/O board a packet modified by a fabric board according to some examples of the present disclosure.

[0005] FIG. 4 is a diagram illustrating a networking structure for forwarding a packet from a VXLAN to a VLAN according to some examples of the present disclosure.

[0006] FIG. 5 is a diagram illustrating a structure of an apparatus for forwarding a packet from a VXLAN to a VLAN according to some examples of the present disclosure.

[0007] FIG. 6 is a diagram illustrating a hardware structure of an apparatus for forwarding a packet from a VXLAN to a VLAN according to some examples of the present disclosure.

Detailed Description

[0008] For simplicity and illustrative purposes, the present disclosure is described by referring mainly to an example thereof. In the following description, numerous specific details are set forth in order to provide a thorough understanding of the present disclosure. It will be readily apparent however, that the present disclosure may be practiced without limitation to these specific details. In other instances, some methods and structures have not been described in detail so as not to unnecessarily obscure the present disclosure. Throughout the present disclosure, the terms "a" and "an" are intended to denote at least one of a particular element. As used herein, the term "includes" means includes but not limited to, the term "including" means including but not limited to. The term "based on" means based at least in part on.

[0009] A frame-type switch includes a main board, an Input/Output (I/O) board and a fabric board. The main board is a single board for implementing calculation of a protocol stack, distribution and control of a forwarding entry and device management. The I/O board is a single board for implementing forwarding of data packets, for example, internal forwarding and external forwarding of data packets. The fabric board is a single board for forwarding a data packet and a control packet between boards and between chips inside a device. A chip on the fabric board has a packet forwarding function and may forward packets between different I/O boards.

[0010] Similar to the structure of the frame-type switch, a VXLAN switch includes a main board, an I/O board and a fabric board. In an example, the number of I/O boards is larger than 1.

[0011] In a VXLAN application, the fabric board of the VXLAN switch is set as the fabric board of the frame-type switch. A chip on the fabric board of the VXLAN switch has a packet forwarding function and may forward packets between different I/O boards.

[0012] In some examples of the present disclosure, a VXLAN switch may be selected as a gateway. The VXLAN switch selected as the gateway is called a gateway VXLAN switch and may be improved.

[0013] The gateway VXLAN switch may be improved as follows.

[0014] The fabric board of the gateway VXLAN switch uses a chip with a forwarding function and various entry functions. The entry functions include a function of

receiving and saving an entry distributed by the main board, an entry searching function and so on. For example, the chip used by the fabric board may be a switch chip used by the I/O board of the gateway VXLAN switch.

[0015] When the fabric board of the gateway VXLAN switch uses the chip with the forwarding function and the entry functions, layer 3 forwarding of a packet entering the gateway VXLAN switch may be implemented via the cooperation of the I/O board and the fabric board of the gateway VXLAN switch. A method for forwarding a packet from a VXLAN to a Virtual Local Area Network (VLAN) will be described hereinafter with reference to some examples and FIG. 1.

[0016] FIG. 1 is a flowchart illustrating a method for forwarding a packet from a VXLAN to a VLAN according to some examples of the present disclosure. The method may be applied to a gateway VXLAN switch. The fabric board of the gateway VXLAN switch uses a chip with a forwarding function and various entry functions. For example, the chip used by the fabric board may be the same as that used by the I/O board of the gateway VXLAN switch. Accordingly, as shown in FIG. 1, the method includes following blocks 201 to 203.

[0017] At block 201, the I/O board of the gateway VXLAN switch receives a packet with VXLAN encapsulation, removes the VXLAN encapsulation from the packet with the VXLAN encapsulation, determines to perform layer 3 forwarding for the packet, and sends the packet to the fabric board of the gateway VXLAN switch.

[0018] In an example of block 201, before the I/O board of the gateway VXLAN switch removes the VXLAN encapsulation from the packet with the VXLAN encapsulation, the method further includes: identifying a Virtual Network ID (VNID) from a VXLAN encapsulation header of the packet.

[0019] In an example of block 201, the I/O board of the gateway VXLAN switch may determine to perform layer 3 forwarding for the packet according to the identified VNID. A method for determining to perform layer 3 forwarding for the packet according to the identified VNID by the I/O board of the gateway VXLAN switch is implemented as follows.

[0020] The I/O board of the gateway VXLAN switch searches a local MAC table for a MAC entry matching the VNID and a destination MAC address of the packet. If the MAC entry is searched out and the destination MAC address of the packet is a gateway MAC address of the gateway VXLAN switch, the I/O board of

the gateway VXLAN switch determines to perform layer 3 forwarding for the packet.

[0021] If the MAC entry is searched out, but the destination MAC address of the packet is not the gateway MAC address of the gateway VXLAN switch, the packet may be forwarded according to a layer 2 packet forwarding method.

[0022] In an example of the present disclosure, the local MAC table of the I/O board may be learned by the I/O board according to a MAC entry learning method, or configured by the main board, or learned by the main board according to the MAC entry learning method and distributed to the I/O board.

[0023] In an example of block 201, in order to ensure that the packet can be sent from the I/O board to the fabric board after the VXLAN encapsulation is removed, the I/O board may add internal encapsulation to the packet. For example, this internal encapsulation may be called first internal encapsulation. Accordingly, in an example of block 201, a method for sending the packet to the fabric board by the I/O board includes following blocks a1 and a2, referring to FIG. 2.

[0024] At block a1, the first internal encapsulation is added to the packet after the VXLAN encapsulation is removed.

[0025] The first internal encapsulation includes a first destination chip identity and a first destination port identity. The first destination chip identity is a first virtual chip identity, and the first destination port identity is a first virtual port identity.

[0026] In an example of the present disclosure, the first virtual chip identity may be a pre-configured virtual chip identity for indicating the fabric board to continue entry searching and packet forwarding. The first virtual port identity may be a pre-configured virtual port identity for indicating the fabric board to continue entry searching and packet forwarding. When the gateway VXLAN switch includes multiple fabric boards, all fabric boards are configured with the first virtual chip identity and the first virtual port identity.

[0027] At block a2, when the I/O board is connected with one fabric board, the packet with the first internal encapsulation is sent via an inner port through which the I/O board is connected with the fabric board; when the I/O board is connected with multiple fabric boards, the packet with the first internal encapsulation is sent via one of inner ports through which the I/O board is connected with the multiple

fabric boards respectively.

[0028] When the I/O board is connected with multiple fabric boards, the inner ports through which the I/O board is connected with the multiple fabric boards respectively may be bundled in an inner port group. Accordingly, in an example of block a2, a process of sending the packet with the first internal encapsulation via one of inner ports through which the I/O board is connected with the multiple fabric boards respectively includes: selecting an inner port from the inner port group and sending the packet with the first internal encapsulation via the selected inner port. There are multiple methods for selecting an inner port from the inner port group, for example, selecting an inner port randomly, or selecting an inner port according to an algorithm such as a Hash algorithm.

[0029] According to blocks a1 and a2, the I/O board of the gateway VXLAN switch can send the packet to the fabric board of the gateway VXLAN switch after the VXLAN encapsulation is removed.

[0030] At block 202, the fabric board of the gateway VXLAN switch receives the packet sent by the I/O board, determines to perform layer 3 forwarding for the received packet, searches out a layer 3 entry matching a destination IP address of the packet from a local layer 3 table, modifies a source MAC address of the packet into the gateway MAC address of the gateway VXLAN switch, modifies the destination MAC address of the packet into a destination MAC address in the layer 3 entry, and when an egress port in the layer 3 entry is a VXLAN tunnel port, sends the modified packet to an I/O board associated with the egress port in the layer 3 entry on the gateway VXLAN switch.

[0031] In an example of the present disclosure, the I/O board associated with the egress port in the layer 3 entry is called a matched I/O board.

[0032] Based on the first internal encapsulation described at block 201, a process of determining to perform layer 3 forwarding for the received packet by the fabric board is implemented as follows.

[0033] The fabric board determines the first destination chip identity and the first destination port identity in a first internal encapsulation header of the received packet. If the first destination chip identity is the first virtual chip identity and the first destination port identity is the first virtual port identity, the fabric board removes the first internal encapsulation from the packet. If the destination MAC

address of the packet is the gateway MAC address of the gateway VXLAN switch after the first internal encapsulation is removed, the fabric board determines to perform layer 3 forwarding for the received packet.

[0034] In an example of the present disclosure, when determining that the first destination chip identity is not the first virtual chip identity and/or the first destination port identity is not the first virtual port identity, the fabric board may not remove the first internal encapsulation from the packet, but forward the packet according to the first destination port identity in the first internal encapsulation.

[0035] In an example of block 202, the local layer 3 table of the fabric board may be configured by the main board, or learned by the main board according to a layer 3 entry learning method and distributed to the fabric board. In an example, the layer 3 entry may be a routing entry, and may include a destination IP address, a Virtual Local Area Network (VLAN) Identity (ID), a destination MAC address, a source MAC address, an egress port and so on. According to the contents included in the layer 3 entry, the layer 3 entry matching the destination IP address of the packet at block 202 is a layer 3 entry including the destination IP address of the packet.

[0036] In an example of the present disclosure, a VLAN port may be a single physical port, or a physical port group in which multiple physical ports are bundled. The single physical port and the physical port group in which multiple physical ports are bundled are called VLAN ports.

[0037] In an example of the present disclosure, the fabric board adds second internal encapsulation to the modified packet and sends the packet with the second internal encapsulation to the matched I/O board. Accordingly, the matched I/O board may send the packet with the second internal encapsulation to a VXLAN.

[0038] Accordingly, in an example of block 202, a method for sending the modified packet to the matched I/O board includes following blocks, referring to FIG. 3.

[0039] At block b1, a target egress port is determined according to the egress port in the layer 3 entry.

[0040] In an example of the present disclosure, the target egress port may be determined according to the type of the egress port in the layer 3 entry. For example, the egress port in the layer 3 entry is a VLAN port. When the VLAN port is a single

physical port, the target egress port is determined as the physical port. When the VLAN port is a physical port group in which multiple physical ports are bundled, the target egress port is determined as one physical port in the physical port group. The physical port may be selected from the physical port group. There are multiple methods for selecting one physical port from the physical port group, for example, selecting one physical port randomly, or selecting one physical port according to a selection algorithm such as a Hash algorithm.

[0041] At block b2, an I/O board where the target egress port is located is determined as the matched I/O board.

[0042] In an example of the present disclosure, the matched I/O board may be determined according to the target egress port described at block b1.

[0043] At block b3, the second internal encapsulation is added to the modified packet.

[0044] The second internal encapsulation may include a second destination chip identity and a second destination port identity. The second destination chip identity is an identity of a chip for forwarding packets on the matched I/O board. If the matched I/O board has one chip, the second destination chip identity may be an identity of the chip. If the matched I/O board has multiple chips, one of the multiple chips may be assigned to forward packets in advance, and the second destination chip identity may be an identity of the assigned chip. The second destination port identity may be an identity of the target egress port described at block b1.

[0045] At block b4, the packet with the second internal encapsulation is sent to the matched I/O board.

[0046] According to blocks b1 to b4, the packet received from the fabric board by the matched I/O board is the packet with the second internal encapsulation.

[0047] At block 203, the matched I/O board in the gateway VXLAN switch receives the packet sent by the fabric board, and forwards the packet to a VLAN via the egress port in the layer 3 entry.

[0048] In an example of block 203, a process of forwarding the packet to the VLAN via the egress port in the layer 3 entry is implemented as follows.

[0049] The matched I/O board determines the second destination chip identity and the second destination port identity in a second internal encapsulation header of the packet. If the second destination chip identity is an identity of a chip for

forwarding packets on the I/O board and the second destination port identity is an identity of a physical port connected with the I/O board, it is indicated that the packet is to be forwarded by the I/O board via the physical port, the matched I/O board removes the second internal encapsulation from the packet, and sends the packet to a VLAN via the physical port corresponding to the second destination port identity.

[0050] In an example of the present disclosure, the local encapsulation entries of the I/O board may be configured on the I/O board in advance.

[0051] In an example of the present disclosure, the fabric board of the gateway VXLAN switch uses a chip with a forwarding function and various entry functions, for example, a chip that is the same as that used by the I/O board. Accordingly, when the gateway VXLAN switch receives the packet with the VXLAN encapsulation via the I/O board, the VXLAN encapsulation is removed from the packet with the VXLAN encapsulation; when determining to perform layer 3 forwarding for the packet after removing the VXLAN encapsulation, the packet is sent to the fabric board of the gateway VXLAN switch. Afterwards, the layer 3 entry matching the destination IP address of the packet is searched out from the local layer 3 table of the fabric board. When the egress port in the layer 3 entry is a VXLAN tunnel port, the I/O board associated with the egress port in the layer 3 entry in the gateway VXLAN switch sends the modified packet to a VLAN. Accordingly, a layer 3 gateway function of VXLAN may be implemented via the gateway VXLAN switch, the packet may be forwarded from the VXLAN to the VLAN, and the whole process is performed inside the gateway VXLAN switch. Accordingly, any bandwidth resources are not wasted and the wire speed forwarding of the packet can be implemented.

[0052] The flowchart shown in FIG. 1 will be described according to an example.

[0053] FIG. 4 is a diagram illustrating a networking structure for forwarding a packet from a VXLAN to a VLAN according to some examples of the present disclosure. As shown in FIG. 4, Physical Machine (PM) 1 to PM5 are devices in a VLAN, and Virtual Machine (VM) 1 to VM5 are devices in a VXLAN. A gateway VXLAN switch is connected between the VLAN and the VXLAN, and includes a main board, an I/O board and a fabric board. The main board is not shown in FIG. 4.

The fabric board uses a chip with a forwarding function and various entry functions. The chip used by the fabric board is the same as that used by the I/O board. FIG. 4 shows three I/O boards of the VXLAN switch.

[0054] A process of forwarding a packet from a VXLAN to a VLAN will be described, in which the VM1 visits the PM1.

[0055] The I/O board 1 of the gateway VXLAN switch receives a packet from the VM1 in the VXLAN via a local port (for example, port 10 in FIG. 4). Herein, the packet received by the I/O board is called packet 0. The packet 0 has VXLAN encapsulation corresponding to the VXLAN. The VXLAN encapsulation includes the VNID100 of the VXLAN and a VXLAN outer layer header corresponding to the VXLAN. The VXLAN outer layer header of the packet 0 shown in FIG. 4 includes outer ETH encapsulation. The outer ETH encapsulation includes an outer source MAC address, an outer destination MAC address, an outer source IP address and an outer destination IP address. FIG. 4 shows the structure of the packet with the VXLAN encapsulation.

[0056] The I/O board 1 obtains the VNID100 from the VXLAN encapsulation, and removes the VXLAN encapsulation from the packet 0. Herein, the packet 0 from which the VXLAN encapsulation has been removed is called packet 1.

[0057] The I/O board 1 searches a local MAC table for a MAC entry matching the destination MAC address of the packet 1 and the obtained VNID100. If the MAC entry is searched out and the destination MAC address of the packet 1 is the gateway MAC address of the gateway VXLAN switch, the I/O board 1 determines to perform layer 3 forwarding for the packet 1.

[0058] The I/O board 1 adds first internal encapsulation to the packet 1. A first destination chip identity in the first internal encapsulation is a first virtual chip identity (for example, chip01 shown in FIG. 4), and a first destination port identity in the first internal encapsulation is a first virtual port identity (for example, port1 shown in FIG. 4). Herein, the packet 1 with the first internal encapsulation is called packet 2.

[0059] As shown in FIG. 4, the I/O board 1 is connected with each fabric board via an inner port group in which inner port 11, inner port 12 and inner port 13 are bundled. The I/O board 1 selects one inner port from the inner port group, for example, selects the inner port 11 shown in FIG. 4 to send the packet 2 to fabric

board 1. A method for selecting an inner port by the I/O board 1 may be configured in advance. For example, the I/O board 1 may perform a Hash operation for the destination IP address of the packet 2 according to a Hash algorithm, and selects an inner port whose number corresponds to a Hash operation result.

[0060] The fabric board 1 receives the packet 2, determines that the first destination chip identity and the first destination port identity in the first internal encapsulation header of the packet 2 are chip01 and port1 respectively, which are the first virtual chip identity and the first virtual port identity respectively, and removes the first internal encapsulation from the packet 2. In this case, the packet 1 is restored.

[0061] The fabric board 1 identifies the destination MAC address of the packet 1, finds that the destination MAC address of the packet 1 is the gateway MAC address of the gateway VXLAN switch, and determines to perform layer 3 forwarding for the packet 1.

[0062] The fabric board 1 searches out a layer 3 entry matching the destination IP address of the packet 1 from a local layer 3 table, and finds that an egress port in the searched-out layer 3 entry is a VLAN port. The fabric board 1 modifies the source MAC address of the packet 1 into the gateway MAC address of the gateway VXLAN switch, and modifies the destination MAC address of the packet 1 into a destination MAC address in the searched-out layer 3 entry. In FIG. 4, the destination MAC address in the searched-out layer 3 entry may be the MAC address of the PM1. Herein, the packet 1 whose source MAC address and destination MAC address have been modified is called packet 3.

[0063] The fabric board 1 finds that the egress port in the searched-out layer 3 entry is a physical port group in which multiple physical ports are bundled, and selects an I/O board where a physical port in the physical port group is located as a matched I/O board. When the egress port in the searched-out layer 3 entry is a single physical port, the fabric board 1 selects an I/O board where the single physical port is located as the matched I/O board. Herein, it is supposed that the I/O board where the single physical port is located is I/O board 2.

[0064] The fabric board 1 adds second internal encapsulation to the packet 3. Herein, the packet 3 with the second internal encapsulation is called packet 4. The second internal encapsulation includes a second destination chip identity and a

second destination port identity. The second destination chip identity is an identity of a chip for forwarding packets on the I/O board 2. In FIG. 4, the identity of the chip is chip02. The second destination port identity is a target egress port. In FIG. 4, the target egress port is port20.

[0065] The fabric board 1 sends the packet 4 to the I/O board 2 via an inner port through which the fabric board 1 is connected with the I/O board 2.

[0066] The I/O board 2 receives the packet 4 sent by the fabric board 1, and determines the second destination chip identity and the second destination port identity in a second internal encapsulation header of the packet 4. If the second destination chip identity is the identity of the chip (for example, chip02) for forwarding packets on the I/O board 2 and the second destination port identity is an identity of a physical port (for example, port20) connected with the I/O board 2, the I/O board 2 removes the second internal encapsulation from the packet 4. In this case, the packet 3 is restored.

[0067] The I/O board 2 sends the packet 3 via an egress port corresponding to the second destination port identity (for example, port20).

[0068] The PM1 can receive the packet sent by the VM1 in the VXLAN. Accordingly, the packet can be forwarded from the VXLAN to the VLAN via the gateway VXLAN switch.

[0069] An apparatus for forwarding a packet from a VXLAN to a VLAN will be described hereinafter with reference to some examples.

[0070] FIG. 5 is a diagram illustrating a structure of an apparatus for forwarding a packet from a VXLAN to a VLAN according to some examples of the present disclosure. The apparatus may be a VXLAN switch used as a gateway. The VXLAN switch includes an I/O board and a fabric board. Both the number of I/O boards and the number of fabric boards are larger than or equal to 1. The fabric board uses a chip with a forwarding function and various entry functions. As shown in FIG. 5, the VXLAN switch further includes an I/O board processing unit 501 and a fabric board processing unit 502.

[0071] The I/O board processing unit 501 is located on an I/O board of the VXLAN switch, and may receive a packet with first VXLAN encapsulation, remove the first VXLAN encapsulation from the packet with the first VXLAN encapsulation, determine to perform layer 3 forwarding for the packet, and send the packet to a

fabric board of the VXLAN switch; receive a packet sent by the fabric board, and forward the packet to a VLAN via an egress port in a layer 3 entry searched out by the fabric board.

[0072] The fabric processing unit 502 is located on the fabric board of the VXLAN switch, and may receive the packet sent by the I/O board, determine to perform layer 3 forwarding for the received packet, search out the layer 3 entry matching a destination IP address of the packet from a local layer 3 table, modify a source MAC address of the packet into a gateway MAC address of the VXLAN switch, modify a destination MAC address of the packet into a destination MAC address in the layer 3 entry, and when the egress port in the layer 3 entry is a VLAN port, send the modified packet to an I/O board associated with the egress port in the layer 3 entry in the VXLAN switch.

[0073] In an example, before removing the first VXLAN encapsulation from the packet with the first VXLAN encapsulation, the I/O board processing unit 501 may identify a VNID from a VXLAN encapsulation header of the packet.

[0074] The I/O board processing unit 501 may determine to perform layer 3 forwarding for the packet through following processes. The I/O board processing unit 501 searches a local MAC table for a MAC entry matching the VNID and the destination MAC address of the packet. If the MAC entry is searched out and the destination MAC address of the packet is the gateway MAC address of the VXLAN switch, the I/O board processing unit 501 determines to perform layer 3 forwarding for the packet.

[0075] In an example, the I/O board processing unit 501 may send the packet to the fabric board of the VXLAN switch through following processes.

[0076] The I/O board processing unit 501 adds first internal encapsulation to the packet after the first VXLAN encapsulation is removed. The first internal encapsulation includes a first destination chip identity and a first destination port identity. The first destination chip identity is a first virtual chip identity, and the first destination port identity is a first virtual port identity.

[0077] In an example, when the I/O board is connected with one fabric board, the I/O board processing unit 501 sends the packet with the first internal encapsulation via an inner port through which the I/O board is connected with the fabric board. When the I/O board is connected with multiple fabric boards, the I/O

board processing unit 501 sends the packet with the first internal encapsulation via one of inner ports through which the I/O board is connected with the multiple fabric boards respectively.

[0078] In an example, the fabric board processing unit 502 may determine to perform layer 3 forwarding for the received packet with the first internal encapsulation through following processes.

[0079] The fabric board processing unit 502 determines the first destination chip identity and the first destination port identity in a first internal encapsulation header of the received packet with the first internal encapsulation. If the first destination chip identity is the first virtual chip identity and the first destination port identity is the first virtual port identity, the fabric board processing unit 502 removes the first internal encapsulation from the packet with the first internal encapsulation. When determining that the destination MAC address of the packet is the gateway MAC address of the VXLAN switch after the first internal encapsulation is removed, the fabric board processing unit 502 determines to perform layer 3 forwarding for the packet.

[0080] In an example, the fabric board processing unit 502 may send the packet modified by the fabric board to the I/O board associated with the egress port in the layer 3 entry in the VXLAN switch through following processes. The fabric board processing unit 502 determines a target egress port according to the egress port in the layer 3 entry, determines an I/O board where the target egress port is located as the I/O board associated with the egress port in the layer 3 entry, adds second internal encapsulation to the modified packet, and send the packet with the second internal encapsulation to the I/O board associated with the egress port in the layer 3 entry. The second internal encapsulation may include a second destination chip identity and a second destination port identity. The second destination chip identity is an identity of a chip for forwarding packets on the I/O board associated with the egress port in the layer 3 entry. The second destination port identity is an identity of the target egress port.

[0081] In an example, the I/O board processing unit 501 may forward the packet to the VLAN via the egress port in the searched-out layer 3 entry through following processes. The I/O board processing unit 501 determines the second destination chip identity and the second destination port identity in a second internal

encapsulation header of the packet. If the second destination chip identity is an identity of a chip for forwarding packets on the I/O board associated with the egress port in the layer 3 entry and the second destination port identity is an identity of a physical port connected with the I/O board associated with the egress port in the layer 3 entry, the I/O board processing unit 501 removes the second internal encapsulation from the packet, and sends the packet to the VLAN via the physical port corresponding to the second destination port identity.

[0082] The fabric board processing unit 502 may determine the target egress port according to the egress port in the layer 3 entry through following processes. When the egress port in the layer 3 entry is a single physical port, the fabric board processing unit 502 determines the physical port as the target egress port. When the egress port in the layer 3 entry is a physical port group in which multiple physical ports are bundled, the fabric board processing unit 502 determines one physical port in the physical port group as the target egress port.

[0083] A hardware structure of the apparatus for forwarding a packet from a VXLAN to a VLAN is also provided according to some examples of the present disclosure. The apparatus is a VXLAN switch used as a gateway. As shown in FIG. 6, the hardware structure of the VXLAN switch may include an I/O board 60 and a fabric board 61.

[0084] The I/O board 60 may include a first processor 601 (for example, a CPU) and a first non-transitory storage602.

[0085] The first non-transitory storage 602 may store machine-readable instructions, which include I/O board processing instructions that can be executed by the first processor 601.

[0086] The first processor 601 may read and execute the I/O board processing instructions stored in the first non-transitory storage 602 to implement the functions of the I/O board processing unit shown in FIG. 5.

[0087] The fabric board 61 may include a second processor 611 (for example, a CPU) and a second non-transitory storage 612.

[0088] The second non-transitory storage 612 may store machine-readable instructions, which include fabric board processing instructions that can be executed by the second processor 611.

[0089] The second processor 611 may read and execute the fabric board

processing instructions stored in the second non-transitory storage 612 to implement the functions of the fabric board processing unit shown in FIG. 5.

[0090] Although described specifically throughout the entirety of the instant disclosure, representative examples of the present disclosure have utility over a wide range of applications, and the above discussion is not intended and should not be construed to be limiting, but is offered as an illustrative discussion of aspects of the disclosure.

[0091] What has been described and illustrated herein is an example along with some of its variations. The terms, descriptions and figures used herein are set forth by way of illustration only and are not meant as limitations. Many variations are possible within the spirit and scope of the subject matter, which is intended to be defined by the following claims -- and their equivalents -- in which all terms are meant in their broadest reasonable sense unless otherwise indicated.

What is claimed is:

1. A packet forwarding method, comprising:

receiving, by an Input/Output (I/O) board of a Virtual eXtensible Local Area Network (VXLAN) switch, a packet with VXLAN encapsulation, removing the VXLAN encapsulation from the packet with the VXLAN encapsulation, determining to perform layer 3 forwarding for the packet, and sending the packet to a fabric board of the VXLAN switch;

receiving, by the fabric board of the VXLAN switch, the packet sent by the I/O board, determining to perform layer 3 forwarding for the received packet, searching out a layer 3 entry matching a destination Internet Protocol (IP) address of the packet from a local layer 3 table, modifying a source Media Access Control (MAC) address of the packet into a gateway MAC address of the VXLAN switch, modifying a destination MAC address of the packet into a destination MAC address in the layer 3 entry, and when an egress port in the layer 3 entry is a VLAN port, sending the modified packet to an I/O board associated with the egress port in the layer 3 entry in the VXLAN switch; and

receiving, by the I/O board associated with the egress port in the layer 3 entry in the VXLAN switch, the packet sent by the fabric board, and forwarding the packet to a VLAN via the egress port in the layer 3 entry.

2. The method of claim 1, before the I/O board of the VXLAN switch removes the VXLAN encapsulation from the packet with the VXLAN encapsulation, further comprising: identifying a Virtual Network ID (VNID) from a VXLAN encapsulation header of the packet, wherein

the determining, by the I/O board of the VXLAN switch, to perform layer 3 forwarding for the packet after removing the VXLAN encapsulation comprises: searching a local MAC table for a MAC entry matching the VNID and the destination MAC address of the packet after the VXLAN encapsulation is removed; when the MAC entry is searched out and the destination MAC address of the packet is a the gateway MAC address of the VXLAN switch, determining to perform layer 3 forwarding for the packet after the VXLAN encapsulation is removed.

3. The method of claim 1, wherein the sending the packet to the fabric board of

the VXLAN switch comprises:

adding first internal encapsulation to the packet, wherein the first internal encapsulation comprises a first destination chip identity and a first destination port identity, the first destination chip identity is a first virtual chip identity, and the first destination port identity is a first virtual port identity;

when the I/O board is connected with one fabric board, sending the packet with the first internal encapsulation via an inner port through which the I/O board is connected with the fabric board; when the I/O board is connected with multiple fabric boards, sending the packet with the first internal encapsulation via one of inner ports through which the I/O board is connected with the multiple fabric boards respectively;

the determining, by the fabric board, to perform layer 3 forwarding for the received packet comprises:

determining the first destination chip identity and the first destination port identity in a first internal encapsulation header of the received packet; when the first destination chip identity is the first virtual chip identity and the first destination port identity is the first virtual port identity, removing the first internal encapsulation from the packet with the first internal encapsulation; when determining that the destination MAC address of the packet is the gateway MAC address of the VXLAN switch, determining to perform layer 3 forwarding for the packet.

4. The method of claim 1, wherein the sending the modified packet to an I/O board associated with the egress port in the layer 3 entry in the VXLAN switch comprises:

determining a target egress port according to the egress port in the layer 3 entry, determining an I/O board where the target egress port is located as the I/O board associated with the egress port in the layer 3 entry, adding second internal encapsulation to the modified packet, and sending the packet with the second internal encapsulation to the I/O board associated with the egress port in the layer 3 entry, wherein the second internal encapsulation includes a second destination chip identity and a second destination port identity, the second destination chip identity is an identity of a chip for forwarding packets on the I/O board associated with the egress port in the layer 3 entry, and the second destination port identity is an identity of the target egress port; and

the forwarding the packet to the VLAN via the egress port in the layer 3 entry comprises:

determining the second destination chip identity and the second destination port identity in a second internal encapsulation header of the packet; when the second destination chip identity is an identity of a chip for forwarding packets on the I/O board associated with the egress port in the layer 3 entry and the second destination port identity is an identity of a physical port connected with the I/O board associated with the egress port in the layer 3 entry, removing the second internal encapsulation from the packet, and sending the packet to the VLAN via the physical port corresponding to the second destination port identity.

5. The method of claim 4, wherein the determining the target egress port according to the egress port in the layer 3 entry comprises:

when the egress port in the layer 3 entry is a single physical port, determining the physical port as the target egress port; when the egress port in the layer 3 entry is a physical port group in which multiple physical ports are bundled, determining one physical port in the physical port group as the target egress port.

6. A Virtual eXtensible Local Area Network (VXLAN) switch, comprising:

an Input/Ouput (I/O) board processing unit, located on an I/O board of the VXLAN switch, and to receive a packet with first VXLAN encapsulation, remove the first VXLAN encapsulation from the packet with the first VXLAN encapsulation, determine to perform layer 3 forwarding for the packet, and send the packet to a fabric board of the VXLAN switch; receive a packet sent by the fabric board, and forward the packet to a Virtual Local Area Network (VLAN) via an egress port in a layer 3 entry searched out by the fabric board; and

a fabric processing unit, located on the fabric board of the VXLAN switch, and to receive the packet sent by the I/O board, determine to perform layer 3 forwarding for the received packet, search out the layer 3 entry matching a destination Internet Protocol (IP) address of the packet from a local layer 3 table, modify a source Media Access Control (MAC) address of the packet into a gateway MAC address of the VXLAN switch, modify a destination MAC address of the packet into a destination MAC address in the layer 3 entry, and when the egress port in the layer 3 entry is a VLAN port, send the modified packet to an I/O board associated with the egress port in the layer 3 entry in the VXLAN switch.

7. The VXLAN switch of claim 6, wherein, before removing the first VXLAN encapsulation from the packet with the first VXLAN encapsulation, the I/O board processing unit 501 is further to identify a Virtual Network ID (VNID) from a VXLAN encapsulation header of the packet;

the I/O board processing unit is to determine to perform layer 3 forwarding for the packet through a process of: searching a local MAC table for a MAC entry matching the VNID and the destination MAC address of the packet, and when the MAC entry is searched out and the destination MAC address of the packet is the gateway MAC address of the VXLAN switch, determining to perform layer 3 forwarding for the packet.

8. The VXLAN switch of claim 6, wherein the I/O board processing unit is to send the packet to the fabric board of the VXLAN switch through a process of:

adding first internal encapsulation to the packet after the first VXLAN encapsulation is removed, wherein the first internal encapsulation includes a first destination chip identity and a first destination port identity, the first destination chip identity is a first virtual chip identity, and the first destination port identity is a first virtual port identity;

when the I/O board is connected with one fabric board, sending the packet with the first internal encapsulation via an inner port through which the I/O board is connected with the fabric board; when the I/O board is connected with multiple fabric boards, sending the packet with the first internal encapsulation via one of inner ports through which the I/O board is connected with the multiple fabric boards respectively; and

the fabric board processing unit is to determine to perform layer 3 forwarding for the received packet with the first internal encapsulation through a process of:

determining the first destination chip identity and the first destination port identity in a first internal encapsulation header of the received packet with the first internal encapsulation; when the first destination chip identity is the first virtual chip identity and the first destination port identity is the first virtual port identity, removing the first internal encapsulation from the packet with the first internal encapsulation; when determining that the destination MAC address of the packet is the gateway MAC address of the VXLAN switch after the first internal encapsulation is removed, determining to perform layer 3 forwarding for the packet.

9. The VXLAN switch of claim 6, wherein the fabric board processing unit is to send the packet modified by the fabric board to the I/O board associated with the egress port in the layer 3 entry in the VXLAN switch through a process of:

determining a target egress port according to the egress port in the layer 3 entry, determining an I/O board where the target egress port is located as the I/O board associated with the egress port in the layer 3 entry, adding second internal encapsulation to the modified packet, and sending the packet with the second internal encapsulation to the I/O board associated with the egress port in the layer 3 entry, wherein the second internal encapsulation includes a second destination chip identity and a second destination port identity, the second destination chip identity is an identity of a chip for forwarding packets on the I/O board associated with the egress port in the layer 3 entry, and the second destination port identity is an identity of the target egress port; and

the I/O board processing unit is to forward the packet to the VLAN via the egress port in the searched-out layer 3 entry through a process of:

determining the second destination chip identity and the second destination port identity in a second internal encapsulation header of the packet; when the second destination chip identity is an identity of a chip for forwarding packets on the I/O board associated with the egress port in the layer 3 entry and the second destination port identity is an identity of a physical port connected with the I/O board associated with the egress port in the layer 3 entry, removing the second internal encapsulation from the packet, and sending the packet to the VLAN via the physical port corresponding to the second destination port identity.

10. The VXLAN switch of claim 6, wherein the fabric board processing unit 502 may determine the target egress port according to the egress port in the layer 3 entry through a process of:

when the egress port in the layer 3 entry is a single physical port, determining the physical port as the target egress port; when the egress port in the layer 3 entry is a physical port group in which multiple physical ports are bundled, determining one physical port in the physical port group as the target egress port.

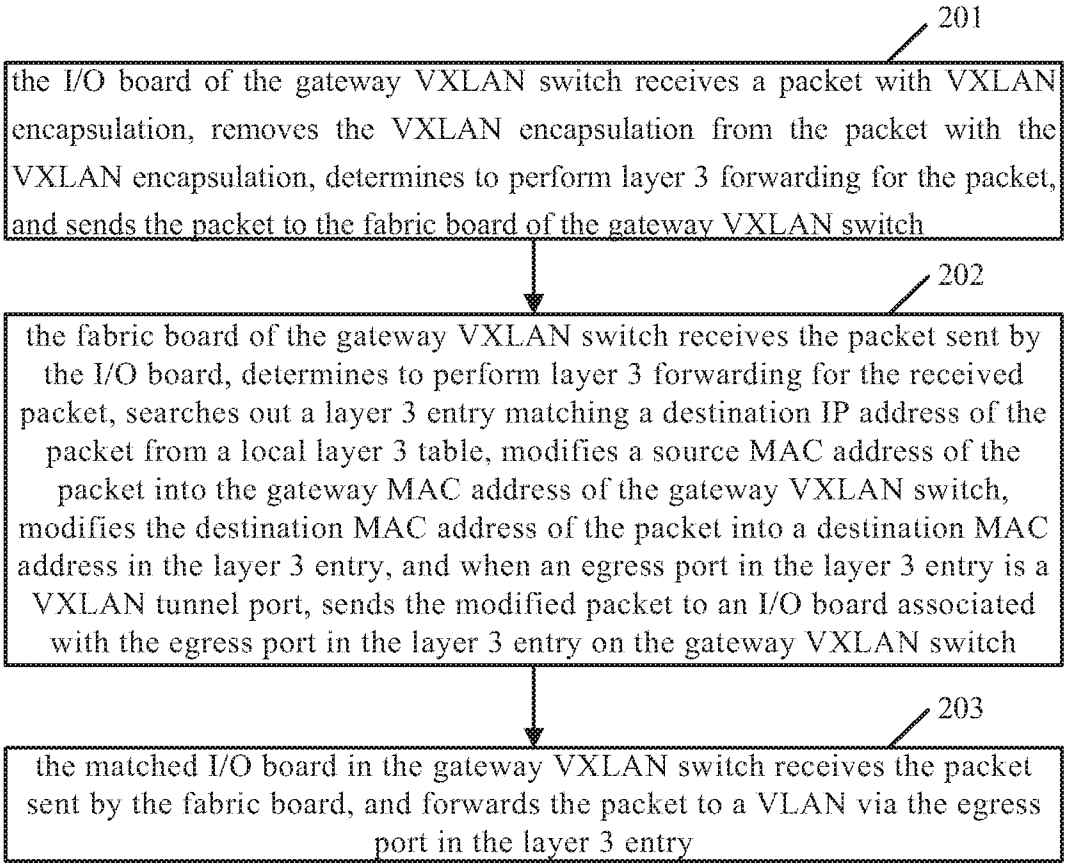
201
the I/O board of the gateway VXLAN switch receives a packet with VXLAN encapsulation, removes the VXLAN encapsulation from the packet with the VXLAN encapsulation, determines to perform layer 3 forwarding for the packet, and sends the packet to the fabric board of the gateway VXLAN switch
202
the fabric board of the gateway VXLAN switch receives the packet sent by the I/O board, determines to perform layer 3 forwarding for the received packet, searches out a layer 3 entry matching a destination IP address of the packet from a local layer 3 table, modifies a source MAC address of the packet into the gateway MAC address of the gateway VXLAN switch, modifies the destination MAC address of the packet into a destination MAC address in the layer 3 entry, and when an egress port in the layer 3 entry is a VXLAN tunnel port, sends the modified packet to an I/O board associated with the egress port in the layer 3 entry on the gateway VXLAN switch
203
the matched I/O board in the gateway VXLAN switch receives the packet sent by the fabric board, and forwards the packet to a VLAN via the egress port in the layer 3 entry

FIG. 1

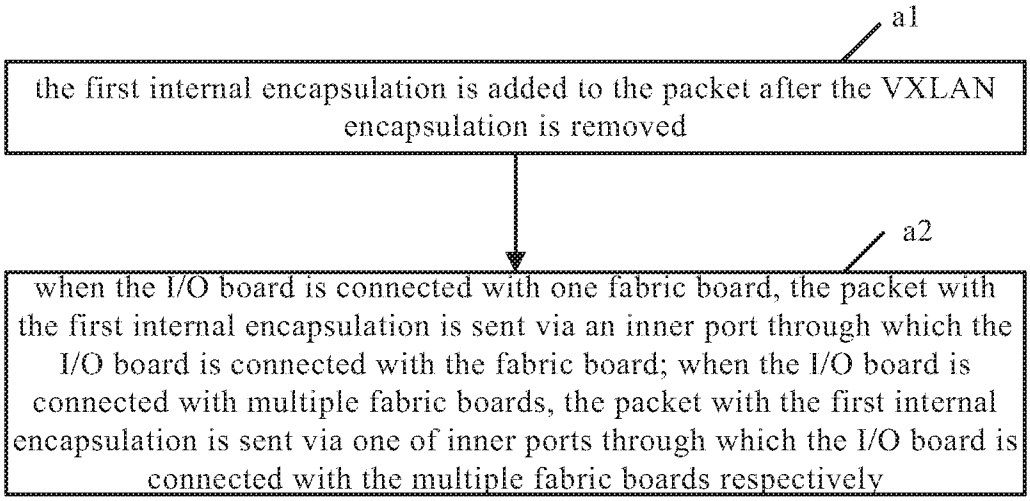
a1
the first internal encapsulation is added to the packet after the VXLAN encapsulation is removed
a2
when the I/O board is connected with one fabric board, the packet with the first internal encapsulation is sent via an inner port through which the I/O board is connected with the fabric board; when the I/O board is connected with multiple fabric boards, the packet with the first internal encapsulation is sent via one of inner ports through which the I/O board is connected with the multiple fabric boards respectively

FIG. 2

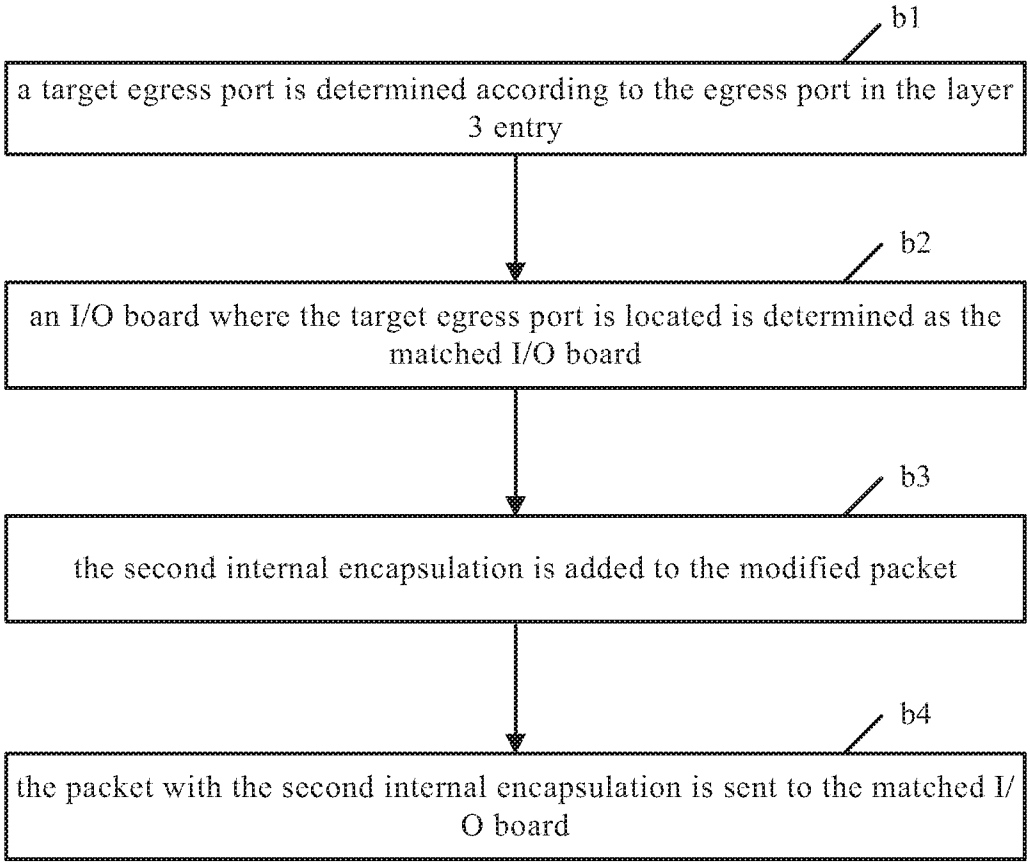
b1
a target egress port is determined according to the egress port in the layer 3 entry
b2
an I/O board where the target egress port is located is determined as the matched I/O board
b3
the second internal encapsulation is added to the modified packet
b4
the packet with the second internal encapsulation is sent to the matched I/O board

FIG. 3

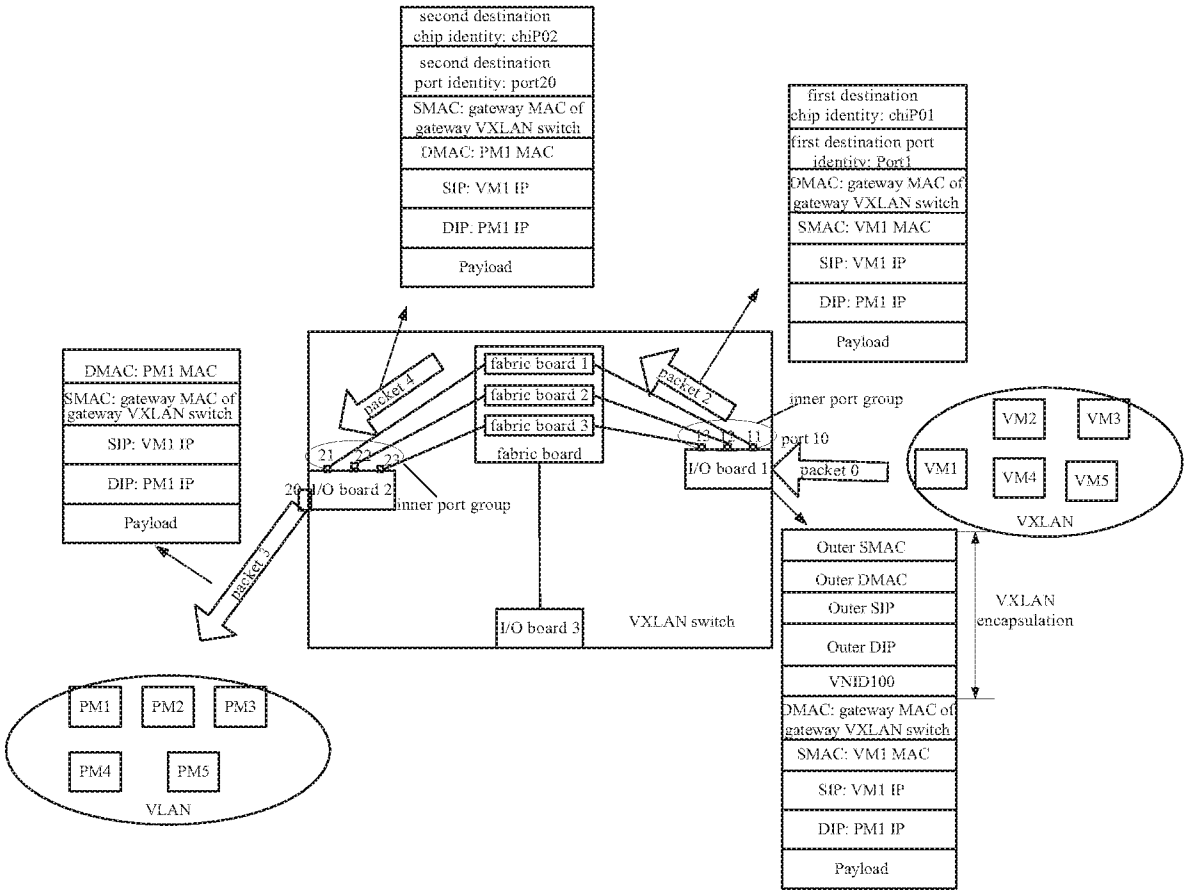
second destination chip identity: chiP02
second destination port identity: port20
SMAC: gateway MAC of gateway VXLAN switch
DMAC: PM1 MAC
SIP: VM1 IP
DIP: PM1 IP
Payload
first destination chip identity: chiP01
first destination port identity: Port1
DMAC: gateway MAC of gateway VXLAN switch
SMAC: VM1 MAC
SIP: VM1 IP
DIP: PM1 IP
Payload
DMAC: PM1 MAC
SMAC: gateway MAC of gateway VXLAN switch
SIP: VM1 IP
DIP: PM1 IP
Payload
packet 4
packet 2
packet 3
fabric board 1
fabric board 2
fabric board 3
fabric board
inner port group
port 10
packet 0
I/O board 1
I/O board 2
I/O board 3
VXLAN switch
VM1
VM2
VM3
VM4
VM5
VXLAN
PM1
PM2
PM3
PM4
PM5
VLAN
Outer SMAC
Outer DMAC
Outer SIP
Outer DIP
VNID100
DMAC: gateway MAC of gateway VXLAN switch
SMAC: VM1 MAC
SIP: VM1 IP
DIP: PM1 IP
Payload
VXLAN encapsulation

FIG. 4

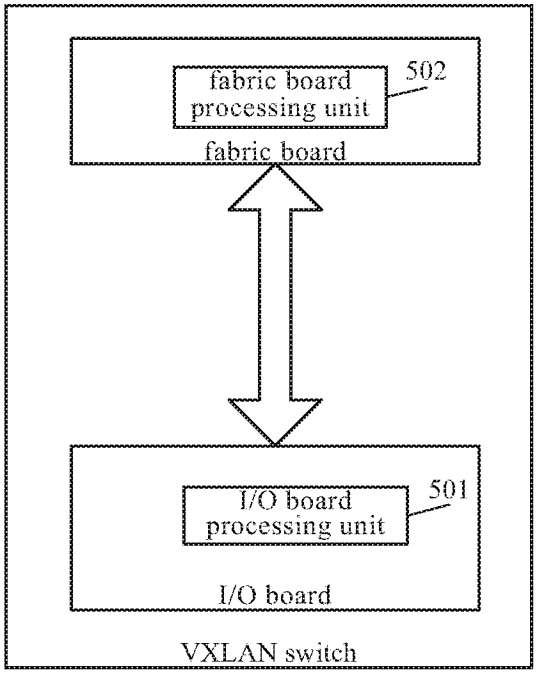
fabric board processing unit
502
fabric board
I/O board processing unit
501
I/O board
VXLAN switch

FIG. 5

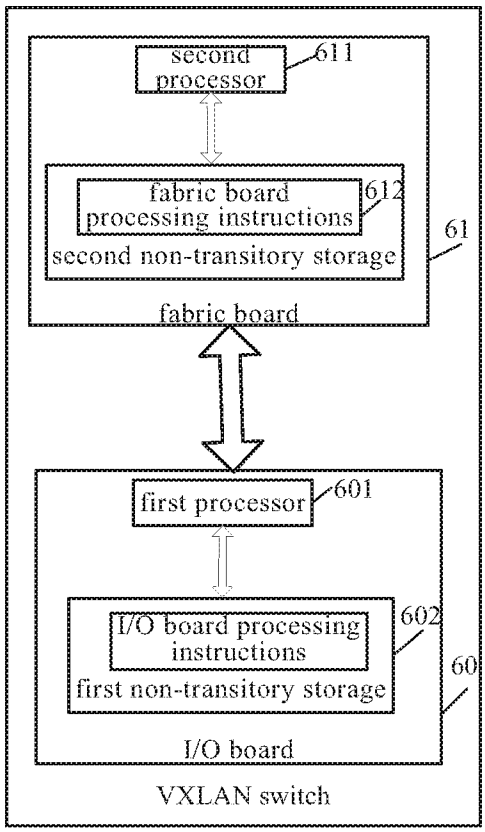
second processor
611
fabric board processing instructions
612
second non-transitory storage
61
fabric board
first processor
601
I/O board processing instructions
602
first non-transitory storage
60
I/O board
VXLAN switch

FIG. 6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2016/080932

A. CLASSIFICATION OF SUBJECT MATTER

H04L 12/741(2013.01)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)

WPI; EPODOC; CNKI; CNPAT; 3GPP: virtual extensible LAN, VXLAN, Virtual Local Area Network, VLAN, gateway, switch, Layer 3, L3, I/O, input/output, board, fabric, encapsulation, media access control, MAC, table, port

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2015063353 A1 (CISCO TECHNOLOGY, INC.) 05 March 2015 (2015-03-05) description, paragraph [0020], figure 1	1-10
A	CN 104158718 A (HANGZHOU H3C TECHNOLOGIES CO., LTD.) 19 November 2014 (2014-11-19) the whole document	1-10
A	CN 103095546 A (HUAWEI TECHNOLOGIES CO., LTD.) 08 May 2013 (2013-05-08) the whole document	1-10

☐ 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	Date of mailing of the international search report
28 June 2016	**22 July 2016**

Name and mailing address of the ISA/CN	Authorized officer
STATE INTELLECTUAL PROPERTY OFFICE OF THE P.R.CHINA **6, Xitucheng Rd., Jimen Bridge, Haidian District, Beijing 100088, China**	**LIU,Qingfeng**
Facsimile No. **(86-10)62019451**	Telephone No. **(86-10)62413279**

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2016/080932

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
US 2015063353 A1	05 March 2015	None	
CN 104158718 A	19 November 2014	None	
CN 103095546 A	08 May 2013	None	