



- (51) International Patent Classification:  
H04L 12/741 (2013.01) H04L 12/761 (2013.01)
- (21) International Application Number:  
PCT/CN2016/080933
- (22) International Filing Date:  
4 May 2016 (04.05.2016)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
201510220886.5 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) Inventor: ZHANG, Yang; 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, 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))

(54) Title: MULTICAST DATA PACKET FORWARDING

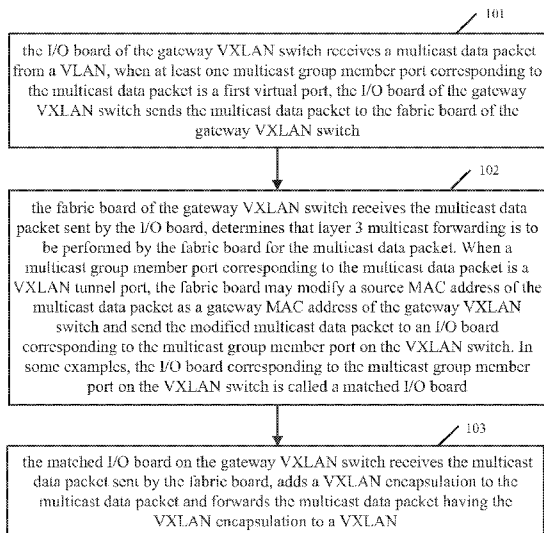
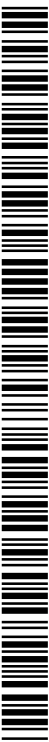


Figure 1

(57) Abstract: An I/O board of a VXLAN switch receives a multicast data packet from a VLAN, and sends the multicast data packet to the fabric board of the VXLAN switch when at least one multicast group member port corresponding to the multicast data packet is a first virtual port. The fabric board determines that layer 3 multicast forwarding is to be performed by the fabric board. When a multicast group member port corresponding to the multicast data packet is a VXLAN tunnel port, the fabric board modifies a source MAC address of the multicast data packet, and sends the modified multicast data packet to an I/O board corresponding to the multicast group member port on the VXLAN switch. The I/O board corresponding to the multicast group member port adds a VXLAN encapsulation to the multicast data packet and forwards the multicast data packet having the VXLAN encapsulation to a VXLAN.



## MULTICAST DATA PACKET FORWARDING

This application claims the benefit of priority from Chinese Patent Application No. 201510220886.5, entitled “Method and apparatus for forwarding a multicast data packet across a VXLAN”, filed on May 4, 2015, the entire content of which is hereby  
5 incorporated by reference.

### Background

Virtual eXtensible Local Area Network (VXLAN) technologies are layer 2 Virtual Private Network (VPN) technologies based on an Internet Protocol (IP) network and a “Media Access Control (MAC) in User Datagram Protocol (UDP)” encapsulation mode is  
10 used. Based on existing service providers or Enterprise IP networks, the VXLAN technologies may provide layer 2 interconnections for spreading physical sites and provide service isolation for different tenants. The VXLAN technologies are mainly used for data center networks.

### Brief Description of the Drawings

15 Figure 1 is a flowchart illustrating a method for forwarding a multicast data packet according to various examples of the present disclosure.

Figure 2 is a flowchart illustrating a method of sending a multicast data packet to a fabric board according to various examples of the present disclosure.

20 Figure 3 is a flowchart illustrating a method of sending a multicast data packet to an I/O board according to various examples of the present disclosure.

Figure 4 is a flowchart illustrating a method of determining a target egress port according to various examples of the present disclosure.

Figure 5 is a diagram illustrating a networking structure for forwarding a multicast data packet according to various examples of the present disclosure.

25 Figure 6 is a diagram illustrating a VXLAN switch according to various examples of the present disclosure.

Figure 7 is a diagram illustrating a hardware structure of a VXLAN switch according

to various examples of the present disclosure.

### **Detailed Description**

For simplicity and illustrative purposes, the present disclosure is described by referring mainly to an example thereof. In the following description, numerous specific  
5 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  
10 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.

Generally, a frame-type switch includes a main board, an I/O board and a fabric board. The main board is a single board for implementing calculations of protocol stacks, distribution and control of forwarding entries, and device management. The I/O board is a  
15 single board for implementing data packet forwarding, e.g. internal forwarding and external forwarding of data packets. The fabric board is a single board for forwarding data packets and control packets between different boards or different chips within the frame-type switch. The chip on the fabric board has a packet forwarding function and is used to forwarding packets between different I/O boards.

20 The structure of a VXLAN switch is similar to the structure of the frame switch, and includes a main board, an I/O board and a fabric board. Generally, the number of the I/O boards is larger than 1.

In VXLAN applications, similarly with the frame-type switch, each chip on the fabric board of the VXLAN switch has the packet forwarding function and is used to  
25 forwarding packets between different I/O boards.

In some examples, one VXLAN switch is selected as a gateway and improvements are made to implement a layer 3 gateway function of the VXLAN. The selected VXLAN switch is called a gateway VXLAN switch.

The improvements of the gateway VXLAN switch are as follows.

The fabric board of the gateway VXLAN switch adopts chips having a forwarding function and various entry functions. The entry functions at least include a function of receiving and storing an entry distributed by the main board, a function of entry searching and etc. In some examples, the fabric board may adopt switch chips adopted by the I/O board of the gateway VXLAN switch.

When the fabric board of the gateway VXLAN switch adopts the chips having the forwarding function and the entry functions, layer 3 forwarding of a multicast data packet entering the gateway VXLAN switch may be implemented via cooperation of the I/O board and the fabric board of the gateway VXLAN switch.

Figure 1 is a flowchart illustrating a method for forwarding a multicast data packet according to various examples of the present disclosure. The method is applied to a gateway VXLAN switch. In some examples, a fabric board of the gateway VXLAN switch adopts chips having a forwarding function and entry functions, e.g. switch chips adopted by an I/O board. As shown in Figure 1, the method includes the following blocks.

At block 101, the I/O board of the gateway VXLAN switch receives a multicast data packet from a Virtual Local Area Network (VLAN), when at least one multicast group member port corresponding to the multicast data packet is a first virtual port, the I/O board of the gateway VXLAN switch sends the multicast data packet to the fabric board of the gateway VXLAN switch.

In some examples, the multicast group member port corresponding to the multicast data packet is included in a layer 3 multicast group entry matching with the multicast data packet. At block 101, the I/O board may search a local layer 3 multicast group table for the layer 3 multicast group entry matching with a destination IP address and a destination MAC address of the multicast data packet.

In some examples, the local layer 3 multicast group table may be directly configured by the main board.

In some examples, the layer 3 multicast group entry may include more than one multicast group member port. When at least one multicast group member port is the first virtual port, the I/O board may send the multicast data packet to the fabric board of the

gateway VXLAN switch. When none of the multicast group member ports is the first virtual port, for example, the multicast group member port may be a VLAN port, the I/O board may perform other processing.

5 The meanings of the first virtual port will be described thereafter and will not be described herein.

In some examples, at block 101, the I/O board may add an internal encapsulation to the multicast data packet, so as to ensure that the multicast data packet is sent from the I/O board to the fabric board successfully. The internal encapsulation is called a first internal encapsulation hereinafter. At block 101, the I/O board may perform the following  
10 processing as shown in Figure 2 when sending the multicast data packet to the fabric board of the gateway VXLAN switch.

At block a1, the I/O board adds the first internal encapsulation to the multicast data packet.

In some examples, the first internal encapsulation includes a first destination chip identifier and a first destination port identifier. The first destination chip identifier may be  
15 an identifier of a first virtual chip, and the first destination port identifier may be an identifier of a first virtual port.

In some examples, the first virtual chip may be a virtual chip configured in advance for indicating that the fabric board may continue to perform entry searching and multicast data packet forwarding. The first virtual port may be a virtual port configured in advance  
20 for indicating that the fabric board may continue to perform entry searching and multicast data packet forwarding.

At block a2, when the I/O board is connected to one fabric board, the multicast data packet having the first internal encapsulation is sent via an inner port connecting the I/O  
25 board with the fabric board. When the I/O board is connected to multiple fabric boards, the multicast data packet having the first internal encapsulation is sent via one of the inner ports connecting the I/O board with the fabric boards respectively.

In some examples, when the I/O board is connected to multiple fabric boards, the inner ports connecting the I/O board with the fabric boards respectively may be bound in

a logic inner port group. At block a2, when sending the multicast data packet having the first internal encapsulation via one of the inner ports connecting the I/O board with the fabric boards respectively, the I/O board may select one inner port from the inner port group and send the multicast data packet having the first internal encapsulation via the selected inner port. Multiple modes may be used to select one inner port from the inner port group, e.g. selecting the inner port randomly, or selecting the inner port according to a certain algorithm, such as Hash algorithm.

After the processing at blocks a1 and a2 is performed, the multicast data packet having the first internal encapsulation may be received by the fabric board of the gateway VXLAN switch.

At block 102, the fabric board of the gateway VXLAN switch receives the multicast data packet sent by the I/O board, determines that layer 3 multicast forwarding is to be performed by the fabric board for the multicast data packet. When a multicast group member port corresponding to the multicast data packet is a VXLAN tunnel port, the fabric board may modify a source MAC address of the multicast data packet as a gateway MAC address of the gateway VXLAN switch and send the modified multicast data packet to an I/O board corresponding to the multicast group member port on the VXLAN switch. In some examples, the I/O board corresponding to the multicast group member port on the VXLAN switch is called a matched I/O board.

In some examples, the multicast group member port corresponding to the multicast data packet is included in a layer 3 multicast group entry matching with the multicast data packet. At block a2, the fabric board may search a local layer 3 multicast group table for the layer 3 multicast group entry matching with a destination IP address and a destination MAC address of the multicast data packet, so as to determine whether the multicast group member port in the searched-out layer 3 multicast group entry is the VXLAN tunnel port.

In some examples, the layer 3 multicast group entry locally stored by the fabric board may be learned by the main board according to a layer 3 multicast group table learning mode and distributed to the fabric board.

In some examples, the VXLAN tunnel port is a virtual tunnel port and may be indicated by using a VXLAN tunnel encapsulation index. For example, the VXLAN

tunnel port may be Tunnel1 which is the VXLAN tunnel encapsulation index.

In some examples, the layer 3 multicast group entry may include more than one multicast group member port. At block 102, when the number of the multicast group member ports is larger than one, for each multicast group member port, the fabric board may determine whether the multicast group member port is the VXLAN tunnel port. When the multicast group member port is the VXLAN tunnel port, the fabric board may modify the source MAC address of the multicast data packet as the gateway MAC address of the gateway VXLAN switch and send the modified multicast data packet to the matched I/O board. When the multicast group member port is not the VXLAN tunnel port, e.g. a VLAN port, the fabric board may perform other operations which will not be described herein.

In some examples, the fabric board may receive one copy of the multicast data packet, and the layer 3 multicast group entry may include more than one multicast group member port. When the number of the multicast group member ports included in the layer 3 multicast group entry is N which is an integer larger than 1, the fabric board may copy the multicast data packet according to the number of the multicast group member ports, and the number of the copied multicast data packets is N-1. The fabric board may distribute the multicast data packet and the copied multicast data packets to the multicast group member ports respectively. Because the contents of the multicast data packet and the copied multicast data packets are the same, the copied multicast data packet is called the multicast data packet in some examples. According to the above processing, each multicast group member port may receive one copy of the multicast data packet.

In some examples, based on the first internal encapsulation at block 101, the fabric board may perform the following processing when determining that the layer 3 multicast forwarding is to be performed by the fabric board for the multicast data packet at block 102.

The fabric board may determine the first destination chip identifier and the first destination port identifier from the header of the first internal encapsulation. If the determined first destination chip identifier is the identifier of the first virtual chip, and the first destination port identifier is the identifier of the first virtual port, the fabric board

may remove the first internal encapsulation from the multicast data packet having the first internal encapsulation. If the destination MAC address of the multicast data packet is the multicast MAC address and the destination IP address of the multicast data packet is the multicast IP address, the fabric board may determine that the layer 3 multicast forwarding is to be performed by the fabric board for the multicast data packet.

In some examples, if the determined first destination chip identifier is not the identifier of the first virtual chip and/or the first destination port identifier is not the identifier of the first virtual port, the fabric board may forward the multicast data packet according to the first destination port identifier without removing the first internal encapsulation from the multicast data packet having the first internal encapsulation.

In some examples, at block 102, the fabric board may perform the following processing as shown in Figure 3 when sending the modified multicast data packet to the matched I/O board.

At block b1, the fabric board may determine a target egress port according to the multicast group member port and select an I/O board including the target egress port as the matched I/O board.

In some examples, the fabric board may perform the following processing as shown in Figure 4 when determining the target egress port according to the multicast group member port.

At block b11, the fabric board may find a locally stored next hop entry corresponding to a VXLAN tunnel associated with the multicast group member port.

At block b12, when the next hop entry includes one next hop, the fabric board may determine the next hop included in the next hop entry as a target next hop; when the next hop entry includes at least two next hops, the fabric board may select one next hop from the at least two next hops and determine the selected next hop as the target next hop. Multiple modes may be used to select the next hop from the at least two next hops, e.g. selecting the next hop randomly, or selecting the next hop according to a certain algorithm, such as Hash algorithm.

At block b13, in the next hop entry, if an egress port corresponding to the target next

hop is a single physical port, the fabric board may determine the single physical port as the target egress port, if the egress port corresponding to the target next hop is a port group formed by binding at least two physical ports, the fabric board may determine one physical port in the port group as the target egress port.

5           The physical port in the port group may be selected from the port group. Multiple modes may be used to select the physical port from the port group, e.g. selecting the physical port randomly, or selecting the physical port according to a certain algorithm, such as Hash algorithm.

10           At block b2, the fabric board may add a second internal encapsulation to the modified multicast data packet, obtain and send a multicast data packet having the second internal encapsulation to the matched I/O board. The second internal encapsulation may include a VXLAN tunnel encapsulation entry index corresponding to the multicast group member port and a Virtual Forwarding Instance (VFI) corresponding to the multicast group member port.

15           In some examples, the VXLAN tunnel encapsulation entry index corresponding to the multicast group member port may be stored in a VXLAN tunnel encapsulation entry corresponding to the multicast group member port. The VFI corresponding to the multicast group member port may be stored in a VFI entry corresponding to the multicast group member port.

20           In some examples, the VXLAN tunnel encapsulation entry and VFI entry may be configured at the fabric board in advance.

          According to the descriptions at blocks b1 and b2, the multicast data packet received by the matched I/O board is the multicast data packet having the second internal encapsulation.

25           At block 103, the matched I/O board on the gateway VXLAN switch receives the multicast data packet sent by the fabric board, adds a VXLAN encapsulation to the multicast data packet and forwards the multicast data packet having the VXLAN encapsulation to a VXLAN.

          In some examples, according to the mode used by the fabric board for sending the

multicast data packet to the matched I/O board at block 102, the multicast data packet received by the matched I/O board is the multicast data packet having the second internal encapsulation. At block 103, after receiving the multicast data packet having the second internal encapsulation, the matched I/O board may determine the VXLAN tunnel encapsulation entry index and the VFI from a header of the second internal encapsulation, remove the second internal encapsulation from the multicast data packet having the second internal encapsulation, find a locally stored VXLAN tunnel encapsulation entry corresponding to the VXLAN tunnel encapsulation entry index, search a preconfigured mapping relation between the VFIs and the VNIDs for a VNID corresponding to the determined VFI, add the VXLAN encapsulation to the multicast data packet according to the VXLAN tunnel encapsulation entry and the searched-out VNID, and send the multicast data packet having the VXLAN encapsulation via an egress port in the found VXLAN tunnel encapsulation entry.

In some examples, when adding the VXLAN encapsulation to the multicast data packet according to the VXLAN tunnel encapsulation entry and the searched-out VNID, the matched I/O board may add a VXLAN outer layer header to the multicast data packet according to VXLAN tunnel encapsulation information in the VXLAN tunnel encapsulation entry, and add a VXLAN field (e.g. a VNID field) to the multicast data packet according to the found VNID. In some examples, the VXLAN outer layer header at least includes an outer layer ETH encapsulation. The outer layer ETH encapsulation at least includes an outer layer source MAC address, an outer layer destination MAC address, an outer layer source IP address, an outer layer destination IP address and etc.

In some examples, when sending the multicast data packet having the VXLAN encapsulation via the egress port in the found VXLAN tunnel encapsulation entry, the following processing may be included. The matched I/O board may identify the egress port from the found VXLAN tunnel encapsulation entry. If the egress port is a single physical port, the matched I/O board may send the multicast data packet having the VXLAN encapsulation via the single physical port. If the egress port is a port group formed by binding at least two physical ports, the matched I/O board may select one physical port from the port group and send the multicast data packet having the VXLAN encapsulation via the selected physical port.

In some examples, the mode used for selecting the physical port from the port group need to ensure that the physical port selected by the matched I/O board is the same as the physical port selected by the fabric board when determining the target egress port. For example, if the Hash algorithm is used by the fabric board to select the physical port, the matched I/O board may select the physical port by using the Hash algorithm.

In some examples, the VXLAN tunnel encapsulation entry stored locally by the matched I/O board may be preconfigured on the matched I/O board.

As can be seen from the processing shown in Figure 1, according to the examples of the present disclosure, the fabric board of the gateway VXLAN switch may adopt chips having the forwarding function and the entry functions, e.g. the switch chips adopted by the I/O board. After receiving the multicast data packet from the VLAN, the I/O board of the gateway VXLAN switch may search the local layer 3 multicast group table for the layer 3 multicast group entry matching with a destination IP address and a destination MAC address of the multicast data packet, when at least one multicast group member port corresponding to the multicast data packet is a first virtual port, the I/O board may send the multicast data packet to the fabric board of the gateway VXLAN switch. The fabric board of the gateway VXLAN switch may search the local layer 3 multicast group table for the layer 3 multicast group entry matching with the destination IP address and the destination MAC address of the multicast data packet. When the multicast group member port corresponding to the multicast data packet in the layer 3 multicast group entry is the VXLAN tunnel port, the fabric board may modify the source MAC address of the multicast data packet as the gateway MAC address of the gateway VXLAN switch and send the modified multicast data packet to the I/O board corresponding to the multicast group member port on the gateway VXLAN switch. The matched I/O board on the gateway VXLAN switch may add the VXLAN encapsulation to the multicast data packet and forward the multicast data packet having the VXLAN encapsulation to the VXLAN. Hence, the layer 3 gateway function of the VXLAN is implemented by using the gateway VXLAN switch, and the multicast data packet is sent from the VLAN to the VXLAN. In addition, the whole processing is implemented within the gateway VXLAN switch, no bandwidth resources are wasted, and wire speed forwarding of the multicast data packet is implemented.

The processing shown in Figure 1 will be described by taking networking shown in Figure 5 as an example. As shown in Figure 5, a Physical Machine (PM) 1, PM2, PM3, PM4 and PM5 are devices in a VLAN, a Virtual Machine (VM)1, VM2, VM3, VM4 and VM5 are devices in a VXLAN. The gateway VXLAN switch is located between the VLAN and the VXLAN, and includes I/O boards, fabric boards and a main board which is not shown in Figure 5. The fabric board adopts the chip having a forwarding function and entry functions, e.g. a switch chip adopted by the I/O board. In the networking shown in Figure 5, an I/O board 1, I/O board 2, and I/O board 3 are taken as examples.

In the example, the PM1 sending a multicast data packet is taken as an example.

10 The I/O board 1 of the gateway VXLAN switch may receive a multicast data packet sent by the PM01 via a local port, e.g. a port10 shown in Figure 5. In the example, the received multicast data packet is called a packet 0.

The I/O board 1 may discover that a destination MAC address of the packet 0 is a multicast MAC address and a destination IP address of the packet 0 is a multicast IP address, and determine the packet 0 as a multicast data packet.

15 The I/O board 1 may search a local layer 3 multicast group table for a layer 3 multicast group entry matching with the destination IP address and the destination MAC address of the multicast data packet.

The I/O board 1 may discover that the layer 3 multicast group entry includes 5 multicast group member ports. Three multicast group member ports are the first virtual port, and the other two multicast group member ports are the VLAN port.

Besides the received packet 0, the fabric board 1 may copy two copies of the packet 0. For the three copies of the packet 0, two copies are sent to the two VLAN ports respectively, and one copy is sent to the first virtual port. The packet sent to the first virtual port is called a packet 2.

25 The I/O board 1 may add a first internal encapsulation to the packet 2. In the first internal encapsulation, the first destination chip identifier is an identifier of a first virtual chip, e.g. chip01 as shown in Figure 5, the first destination port identifier is an identifier of a first virtual port, e.g. Port1 as shown in Figure 5. In the example, a packet obtained

after adding the first internal encapsulation to the packet 2 is called a packet 3.

As shown in Figure 5, the I/O board 1 is connected to the fabric boards via a port group including an inner port 11, an inner port 12 and an inner port 13. The I/O board 1 may select one inner port from the port group, e.g. the port 11, and send the packet 3 to the fabric board 1 via the inner port 11. The mode used by the I/O board 1 for selecting the inner port may be configured in advance. For example, Hash algorithm may be used. In some example, the I/O board 1 may perform Hash calculation of the number of the Port11, and select an inner port having a number corresponding to a result of the Hash calculation.

10 The fabric board 1 may receive the packet 3, and determine that the first destination chip identifier and the first destination port identifier in a header of the first internal encapsulation of the packet 3 are chip01 and Port1 respectively, i.e. the identifier of the first virtual chip and the identifier of the first virtual port. The fabric board 1 may remove the first internal encapsulation, and the packet 2 is obtained.

15 The fabric board 1 may determine the destination MAC address of the packet 2 is the multicast MAC address, and the destination IP address of the packet 2 is the multicast IP address, and determine that the fabric board 1 need to perform layer 3 multicast forwarding for the packet 2.

The fabric board 1 may search a local layer 3 multicast group table for a layer 3 multicast group entry matching with the destination IP address and destination MAC address of the packet 2.

The fabric board 1 may discover three multicast group member ports are included in the layer 3 multicast group entry. For example, the three multicast group member ports are represented as a Port11, Port12, Port13, which are not shown in Figure 5.

25 The fabric board 1 may copy two copies of the packet 2 which are represented as a packet 21 and a packet 22. The fabric board 1 may distribute the packet 2 and two copies of the packet 2 to the three multicast group member ports. For example, the packet 21 is distributed to the Port 11, the packet 22 is distributed to the Port 12, and the packet 2 is distributed to the Port 13.

In the following processing, the Port 11 is taken as an example, and the Port 12 and Port 13 may be processed accordingly.

The fabric board 1 may determine that the Port 11 is a VXLAN tunnel port, and modify the source MAC address of the packet 21 as a gateway MAC address of the gateway VXLAN switch. In the example, the modified packet 21 is called a packet 4.

The fabric board 1 may find a locally stored next hop entry corresponding to a VXLAN tunnel associated with the VXLAN tunnel port. After determining that the found next hop entry includes one next hop and an egress port corresponding to the next hop is a port group formed by binding at least two physical ports, the fabric board 1 may select one physical port of the port group as the target egress port, and an I/O board including the selected physical port is the I/O board associated with the multicast group member port Port11. In the example, the I/O board including the target egress port is the I/O board 2.

The fabric board 1 may find a VXLAN tunnel encapsulation entry corresponding to the multicast group member port Port11 from locally stored VXLAN tunnel encapsulation entries, and find a VFI entry corresponding to the multicast group member port Port11 from locally stored VFI entries.

The fabric board 1 may add a second internal encapsulation to the packet 4, and a packet obtained after adding the second internal encapsulation to the packet 4 is called a packet 5. The second internal encapsulation may include an index of the VXLAN tunnel encapsulation entry found by the fabric board 1 and VFI included in the VFI entry found by the fabric board 1.

The fabric board 1 may send the packet 5 to the I/O board 2 via an inner port connecting the fabric board 1 and the I/O board 2.

The I/O board 2 may receive the packet 5 sent by the fabric board 1, determine the VXLAN tunnel encapsulation entry and the VFI from a header of the second internal encapsulation, and remove the second internal encapsulation from the packet 5, thus the packet 4 is obtained.

The I/O board 2 may find a locally stored VXLAN tunnel encapsulation entry

corresponding to the index of the VXLAN tunnel encapsulation entry, search a preconfigured mapping relation between the VFIs and the VNIDs for a VNID corresponding to the determined VFI, e.g. the VNID200as shown in Figure 5, add a VXLAN encapsulation to the packet 4 according to the found VXLAN tunnel  
5 encapsulation entry and the VNID. In the example, a packet obtained after adding the VXLAN encapsulation to the packet 4 is called a packet 6.

In some examples, the I/O board 2 may add the VXLAN encapsulation to the packet 4 according to the following processing. The I/O board 2 may add a VXLAN outer layer header to the packet 4 according to VXLAN tunnel encapsulation information in the  
10 VXLAN tunnel encapsulation entry, and add a VXLAN field (e.g. a VNID field) to the packet 3 according to the found VNID, e.g. the VNIN200 as shown in Figure 3. As shown in Figure 5, in the VXLAN encapsulation of the packet 6, the VXLAN outer layer header includes an outer ETH encapsulation, and the outer ETH encapsulation includes an outer source MAC address, an outer destination MAC address, an outer source IP address, an  
15 outer destination IP address.

The I/O board 2 may send the packet 6 via an egress port in the found VXLAN tunnel encapsulation entry. In some examples, the I/O board 2 may identify the egress port from the found VXLAN tunnel encapsulation entry. The identified egress port is an egress port of a next hop found by the fabric board. As described above, the egress port is  
20 a port group formed by binding at least two physical ports, and thus the identified egress port is the port group. The I/O board 2 may select one physical port from the port group according to the mode used by the fabric board for selecting the target egress port, thereby ensuring the port via which the packet 5 is sent is the same as the target egress port selected by the fabric board.

In the example, if the multicast group member associated with the Port 11 is the VM11, the VM11 of the VXLAN will receive the multicast data packet sent by the PM1. Similarly, if the multicast group member associated with the Port 12 is the VM12, the VM12 of the VXLAN will receive the multicast data packet sent by the PM1; if the  
25 multicast group member associated with the Port 13 is the VM13, the VM13 of the VXLAN will receive the multicast data packet sent by the PM1. Therefore, the multicast data packet is forwarding from the VLAN to the VXLAN by using the gateway VXLAN  
30

switch.

Figure 6 is a diagram illustrating a VXLAN switch according to various examples of the present disclosure. As shown in Figure 6, the VXLAN switch at least includes an I/O board 61 and a fabric board 62. The fabric board 62 adopts a chip having a forwarding function and entry functions. As shown in Figure 6, the VXLAN switch includes an I/O board processing unit 601 located at the I/O board 61 and a fabric board processing unit 602 located at the fabric board 62.

The I/O board processing unit 601 is to receive a multicast data packet from a VLAN, when at least one multicast group member port corresponding to the multicast data packet is a first virtual port, send the multicast data packet to the fabric board 62 of the VXLAN switch. The I/O board processing unit 601 is further to receives a multicast data packet sent by the fabric board 62, add a VXLAN encapsulation to the multicast data packet and forwards the multicast data packet having the VXLAN encapsulation to a VXLAN.

The fabric board processing unit 602 is to receive the multicast data packet sent by the I/O board 61, determine that layer 3 multicast forwarding is to be performed by the fabric board 62 for the multicast data packet. When a multicast group member port corresponding to the multicast data packet is a VXLAN tunnel port, the fabric board processing unit 602 is to modify a source MAC address of the multicast data packet as a gateway MAC address of the VXLAN switch and send the modified multicast data packet to an I/O board corresponding to the multicast group member port on the VXLAN switch.

In some examples, the multicast group member port corresponding to the multicast data packet is included in a layer 3 multicast group entry matching with the multicast data packet. The I/O board processing unit 601 is further to search a local layer 3 multicast group table for the layer 3 multicast group entry matching with a destination IP address and a destination MAC address of the multicast data packet. The fabric board processing unit 602 is further to search a local layer 3 multicast group table for the layer 3 multicast group entry matching with a destination IP address and a destination MAC address of the multicast data packet.

In some examples, the I/O board processing unit 601 may send the multicast data packet to the fabric board 62 of the VXLAN switch according to the following

processing.

The I/O board processing unit 601 is to add a first internal encapsulation to the multicast data packet. The first internal encapsulation includes a first destination chip identifier and a first destination port identifier. The first destination chip identifier may be an identifier of a first virtual chip, and the first destination port identifier may be an identifier of a first virtual port.

When the I/O board 61 is connected to one fabric board 62, the I/O board processing unit 601 is to send the multicast data packet having the first internal encapsulation via an inner port connecting the I/O board 61 with the fabric board 62. When the I/O board 61 is connected to multiple fabric boards 62, the I/O board processing unit 601 is to send the multicast data packet having the first internal encapsulation via one of the inner ports connecting the I/O board 61 with the fabric boards 62 respectively.

In some examples, the fabric board processing unit 602 may determine that layer 3 multicast forwarding is to be performed by the fabric board 62 for the multicast data packet according to the following processing.

The fabric board processing unit 602 is to determine the first destination chip identifier and the first destination port identifier from the header of the first internal encapsulation. If the determined first destination chip identifier is the identifier of the first virtual chip, and the first destination port identifier is the identifier of the first virtual port, the fabric board processing unit 602 is to remove the first internal encapsulation from the multicast data packet having the first internal encapsulation. If the destination MAC address of the multicast data packet is the multicast MAC address and the destination IP address of the multicast data packet is the multicast IP address, the fabric board processing unit 602 is to determine that layer 3 multicast forwarding is to be performed by the fabric board 62 for the multicast data packet.

In some examples, the fabric board processing unit 602 may send the modified multicast data packet to the I/O board corresponding to the multicast group member port on the VXLAN switch according to the following processing.

The fabric board processing unit 602 is to determine a target egress port according to

the multicast group member port and select an I/O board including the target egress port as the matched I/O board, add a second internal encapsulation to the modified multicast data packet, obtain and send a multicast data packet having the second internal encapsulation to the I/O board including the target egress port. The second internal encapsulation may include a VXLAN tunnel encapsulation entry index corresponding to the multicast group member port and a VFI corresponding to the multicast group member port.

In some examples, the I/O board processing unit 601 may add the VXLAN encapsulation to the multicast data packet and forward the multicast data packet having the VXLAN encapsulation to the VXLAN according to the following processing.

The I/O board processing unit 601 is to determine the VXLAN tunnel encapsulation entry index and the VFI from a header of the second internal encapsulation, and remove the second internal encapsulation from the multicast data packet having the second internal encapsulation, find a locally stored VXLAN tunnel encapsulation entry corresponding to the VXLAN tunnel encapsulation entry index, search a preconfigured mapping relation between the VFIs and the VNIDs for a VNID corresponding to the determined VFI, add the VXLAN encapsulation to the multicast data packet according to the VXLAN tunnel encapsulation entry and the searched-out VNID, and send the multicast data packet having the VXLAN encapsulation via an egress port in the found VXLAN tunnel encapsulation entry.

In some examples, the fabric board processing unit 602 may determine the target egress port according to the multicast group member port according to the following processing.

The fabric board processing unit 602 is to find a locally stored next hop entry corresponding to a VXLAN tunnel associated with the multicast group member port.

When the next hop entry includes one next hop, the fabric board processing unit 602 is to determine the next hop included in the next hop entry as a target next hop; when the next hop entry includes at least two next hops, the fabric board processing unit 602 is to select one next hop from the at least two next hops and determine the selected next hop as the target next hop. In the next hop entry, when an egress port corresponding to the target

next hop is a single physical port, the fabric board processing unit 602 is to determine the single physical port as the target egress port, when the egress port corresponding to the target next hop is a port group formed by binding at least two physical ports, the fabric board processing unit 602 is to determine one physical port in the port group as the target egress port.

Figure 7 is a diagram illustrating a hardware structure of a VXLAN switch according to various examples of the present disclosure. As shown in Figure 7, the hardware structure includes an I/O board 71 and a fabric board 72.

The I/O board 71 includes a first CPU 711 and a first storage 712.

The first storage 712 is to store the I/O board processing unit 601.

The first CPU 711 is to store a running control program of the I/O board processing unit 601, so as to control the I/O board processing unit 601 stored in the first storage 712 to perform the above operations of the I/O board processing unit 601, and the operations will not be described herein.

The I/O board 72 includes a second CPU 721 and a second storage 722.

The second storage 722 is to store the fabric board processing unit 602.

The second CPU 721 is to store a running control program of the fabric board processing unit 602, so as to control the fabric board processing unit 602 stored in the second storage 722 to perform the above operations of the fabric board processing unit 602, and the operations will not be described herein.

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.

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.

5

## Claims

1. A method for forwarding a multicast data packet, comprising:

receiving, by an I/O board of an Virtual eXtensible Local Area Network (VXLAN) switch, a multicast data packet from a Virtual Local Area Network (VLAN); when at least  
5 one multicast group member port corresponding to the multicast data packet is a first virtual port, sending the multicast data packet to the fabric board of the VXLAN switch;

receiving, by the fabric board of the VXLAN switch, the multicast data packet sent by the I/O board; determining that layer 3 multicast forwarding is to be performed by the fabric board for the multicast data packet; when a multicast group member port  
10 corresponding to the multicast data packet is a VXLAN tunnel port, modifying a source Media Access Control (MAC) address of the multicast data packet as a gateway MAC address of the VXLAN switch; sending the modified multicast data packet to an I/O board corresponding to the multicast group member port on the VXLAN switch; and

receiving, by the I/O board corresponding to the multicast group member port on the  
15 VXLAN switch, the multicast data packet sent by the fabric board; adding a VXLAN encapsulation to the multicast data packet and forwarding the multicast data packet having the VXLAN encapsulation to a VXLAN.

2. The method of claim 1, wherein the multicast group member port corresponding  
20 to the multicast data packet is included in a layer 3 multicast group entry matching with the multicast data packet; and the method further comprises:

searching, by the I/O board, a local layer 3 multicast group table for the layer 3 multicast group entry matching with a destination IP address and a destination MAC address of the multicast data packet;

25 searching, by the fabric board, a local layer 3 multicast group table for the layer 3 multicast group entry matching with a destination IP address and a destination MAC address of the multicast data packet.

3. The method of claim 1, wherein sending the multicast data packet to the fabric  
30 board of the VXLAN switch comprises:

adding a first internal encapsulation to the multicast data packet after removing the first VXLAN encapsulation; the first internal encapsulation comprising a first destination

chip identifier and a first destination port identifier; the first destination chip identifier being an identifier of a first virtual chip, and the first destination port identifier being an identifier of a first virtual port;

5 when the I/O board is connected to one fabric board, sending the multicast data packet having the first internal encapsulation via an inner port connecting the I/O board with the fabric board; when the I/O board is connected to multiple fabric boards, sending the multicast data packet having the first internal encapsulation via one of the inner ports connecting the I/O board with the fabric boards respectively.

10 4. The method of claim 3, wherein determining that layer 3 multicast forwarding is to be performed by the fabric board for the multicast data packet comprises:

determining the first destination chip identifier and the first destination port identifier from a header of the first internal encapsulation; if the determined first destination chip identifier is the identifier of the first virtual chip, and the first destination port identifier is the identifier of the first virtual port, removing the first internal encapsulation from the multicast data packet having the first internal encapsulation; if the destination MAC address of the multicast data packet is the multicast MAC address and the destination IP address of the multicast data packet is the multicast IP address, determining that layer 3 multicast forwarding is to be performed by the fabric board for the multicast data packet.

5. The method of claim 1, wherein sending the modified multicast data packet to the I/O board corresponding to the multicast group member port on the VXLAN switch comprises:

25 determining a target egress port according to the multicast group member port and selecting an I/O board including the target egress port as the I/O board corresponding to the multicast group member port;

adding a second internal encapsulation to the modified multicast data packet, obtaining and sending a multicast data packet having the second internal encapsulation to the I/O board including the target egress port; the second internal encapsulation comprising an VXLAN tunnel encapsulation entry index corresponding to the multicast group member port and a VFI corresponding to the multicast group member port;

wherein adding the VXLAN encapsulation to the multicast data packet and

forwarding the multicast data packet having the VXLAN encapsulation to the VXLAN comprises:

determining the VXLAN tunnel encapsulation entry index and the VFI from a header of the second internal encapsulation; removing the second internal encapsulation from the multicast data packet having the second internal encapsulation; finding a locally stored VXLAN tunnel encapsulation entry corresponding to the VXLAN tunnel encapsulation entry index; searching a preconfigured mapping relation between VFIs and VNIDs for a VNID corresponding to the determined VFI; adding the VXLAN encapsulation to the multicast data packet according to the VXLAN tunnel encapsulation entry and the searched-out VNID; and sending the multicast data packet having the VXLAN encapsulation via an egress port in the found VXLAN tunnel encapsulation entry.

6. The method of claim 5, wherein determining the target egress port according to the multicast group member port comprises:

finding a locally stored next hop entry corresponding to a VXLAN tunnel associated with the multicast group member port;

when the next hop entry includes one next hop, determining the next hop included in the next hop entry as a target next hop; when the next hop entry includes at least two next hops, selecting one next hop from the at least two next hops and determining the selected next hop as the target next hop; in the next hop entry, when an egress port corresponding to the target next hop is a single physical port, determining the single physical port as the target egress port, when the egress port corresponding to the target next hop is a port group formed by binding at least two physical ports, determining one physical port in the port group as the target egress port.

7. A VXLAN switch, comprising:

an I/O board processing unit, located at an I/O board of the VXLAN switch, to receive a multicast data packet from a VLAN; when at least one multicast group member port corresponding to the multicast data packet is a first virtual port, send the multicast data packet to a fabric board of the VXLAN switch.; receive a multicast data packet from the fabric board, add a VXLAN encapsulation to the multicast data packet received from the fabric board and forward the multicast data packet having the VXLAN encapsulation

to a VXLAN; and

a fabric board processing unit, located at the fabric board of the VXLAN switch, to receive the multicast data packet sent by the I/O board, determine that layer 3 multicast forwarding is to be performed by the fabric board for the multicast data packet; when a  
5 multicast group member port corresponding to the multicast data packet is a VXLAN tunnel port, modify a source MAC address of the multicast data packet as a gateway MAC address of the VXLAN switch and send the modified multicast data packet to an I/O board corresponding to the multicast group member port on the VXLAN switch.

10 8. The VXLAN switch of claim 7, wherein the multicast group member port corresponding to the multicast data packet is included in a layer 3 multicast group entry matching with the multicast data packet;

the I/O board processing unit is further to search a local layer 3 multicast group table for the layer 3 multicast group entry matching with a destination IP address and a  
15 destination MAC address of the multicast data packet;

the fabric board processing unit is further to search a local layer 3 multicast group table for the layer 3 multicast group entry matching with a destination IP address and a destination MAC address of the multicast data packet.

20 9. The VXLAN switch of claim 7, wherein the I/O board processing unit is to send the multicast data packet to the fabric board of the VXLAN switch through a process of:

adding a first internal encapsulation to the multicast data packet after removing the first VXLAN encapsulation; the first internal encapsulation comprising a first destination chip identifier and a first destination port identifier; the first destination chip identifier  
25 being an identifier of a first virtual chip, and the first destination port identifier being an identifier of a first virtual port;

when the I/O board is connected to one fabric board, sending the multicast data packet having the first internal encapsulation via an inner port connecting the I/O board with the fabric board; when the I/O board is connected to multiple fabric boards, sending  
30 the multicast data packet having the first internal encapsulation via one of the inner ports connecting the I/O board with the fabric boards respectively.

10. The VXLAN switch of claim 7, wherein the fabric board processing unit is to

determine that layer 3 multicast forwarding is to be performed by the fabric board for the multicast data packet through a process of:

determining the first destination chip identifier and the first destination port identifier from a header of the first internal encapsulation; if the determined first destination chip identifier is the identifier of the first virtual chip, and the first destination port identifier is the identifier of the first virtual port, removing the first internal encapsulation from the multicast data packet having the first internal encapsulation; if the destination MAC address of the multicast data packet is the multicast MAC address and the destination IP address of the multicast data packet is the multicast IP address, determining that layer 3 multicast forwarding is to be performed by the fabric board for the multicast data packet.

11. The VXLAN switch of claim 7, wherein the fabric board processing unit is to send the modified multicast data packet to the I/O board corresponding to the multicast group member port on the VXLAN switch through a process of:

determining a target egress port according to the multicast group member port and selecting an I/O board including the target egress port as the I/O board corresponding to the multicast group member port;

adding a second internal encapsulation to the modified multicast data packet, obtaining and sending a multicast data packet having the second internal encapsulation to the I/O board including the target egress port; the second internal encapsulation comprising an VXLAN tunnel encapsulation entry index corresponding to the multicast group member port and a VFI corresponding to the multicast group member port;

wherein adding the VXLAN encapsulation to the multicast data packet and forwarding the multicast data packet having the VXLAN encapsulation to the VXLAN comprises:

determining the VXLAN tunnel encapsulation entry index and the VFI from a header of the second internal encapsulation; removing the second internal encapsulation from the multicast data packet having the second internal encapsulation; finding a locally stored VXLAN tunnel encapsulation entry corresponding to the VXLAN tunnel encapsulation entry index; searching a preconfigured mapping relation between VFIs and VNIDs for a VNID corresponding to the determined VFI; adding the VXLAN encapsulation to the multicast data packet according to the VXLAN tunnel encapsulation

entry and the searched-out VNID; and sending the multicast data packet having the VXLAN encapsulation via an egress port in the found VXLAN tunnel encapsulation entry.

5           12. The VXLAN switch of claim 11, wherein the fabric board processing unit is to determine the target egress port according to the multicast group member port through a process of:

          finding a locally stored next hop entry corresponding to a VXLAN tunnel associated with the multicast group member port;

10           when the next hop entry includes one next hop, determining the next hop included in the next hop entry as a target next hop; when the next hop entry includes at least two next hops, selecting one next hop from the at least two next hops and determining the selected next hop as the target next hop; in the next hop entry, when an egress port corresponding to the target next hop is a single physical port, determining the single physical port as the target egress port, when the egress port corresponding to the target next hop is a port group formed by binding at least two physical ports, determining one physical port in the port group as the target egress port.

15

20

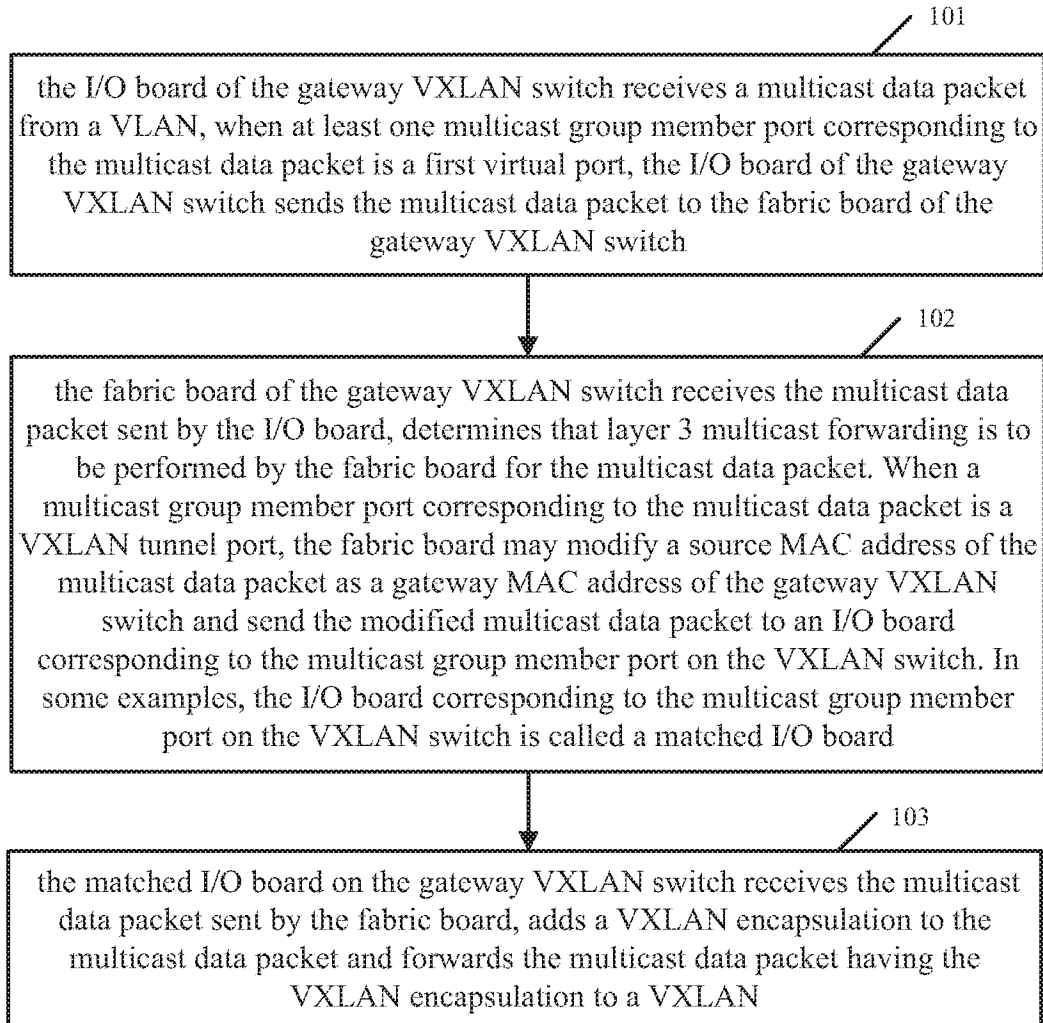


Figure 1

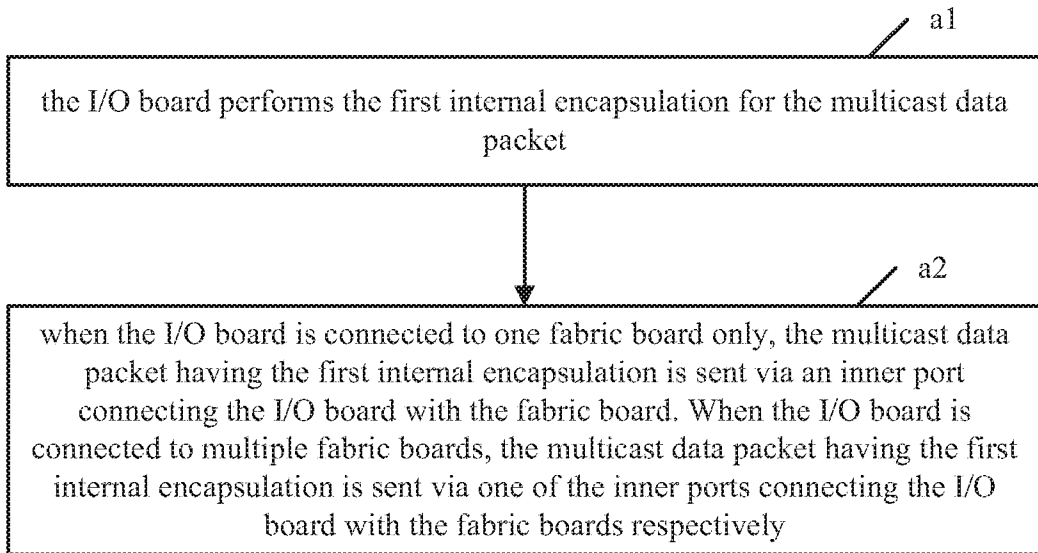


Figure 2

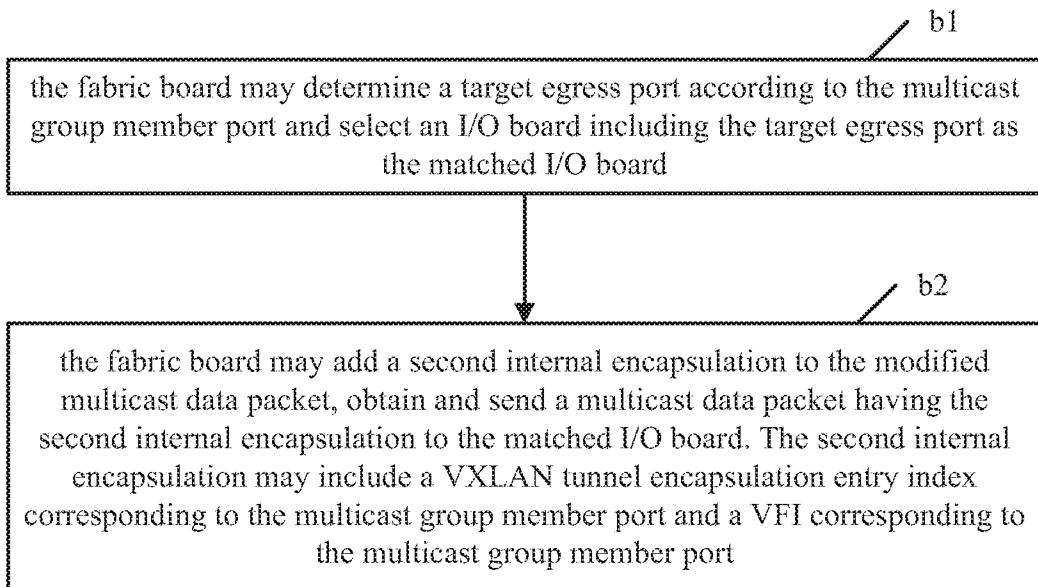


Figure 3

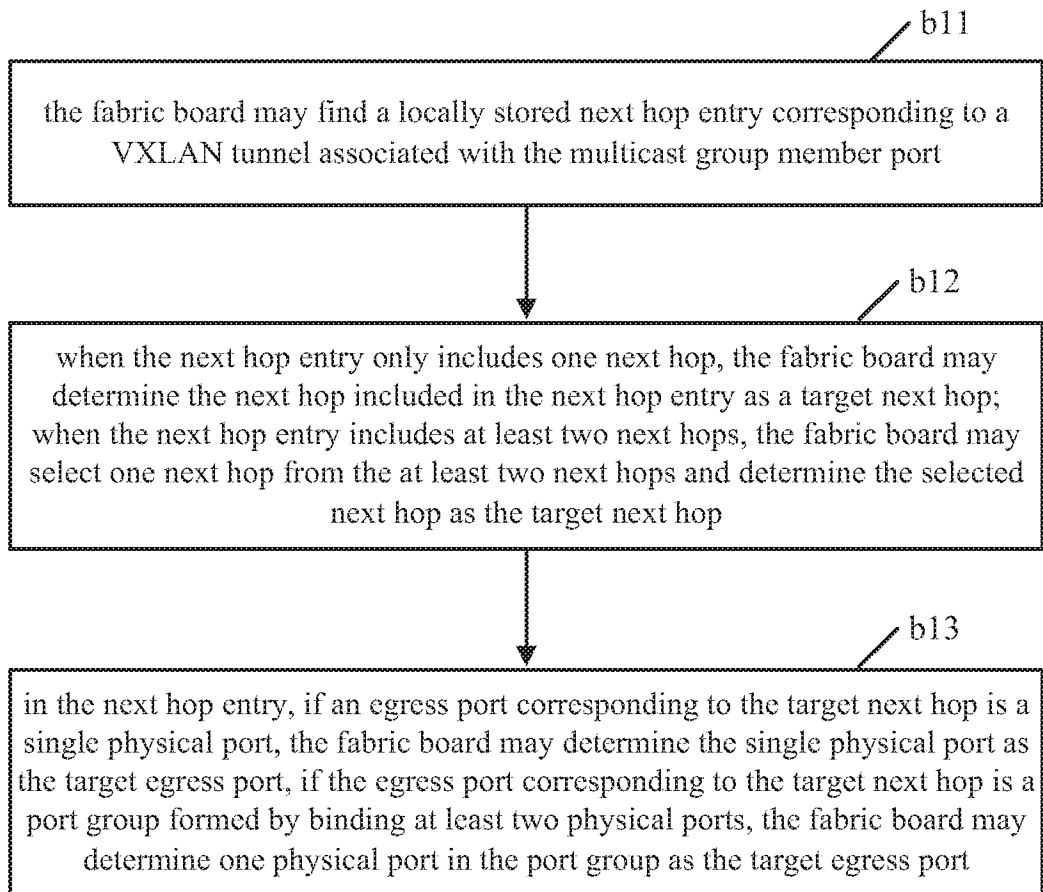


Figure 4

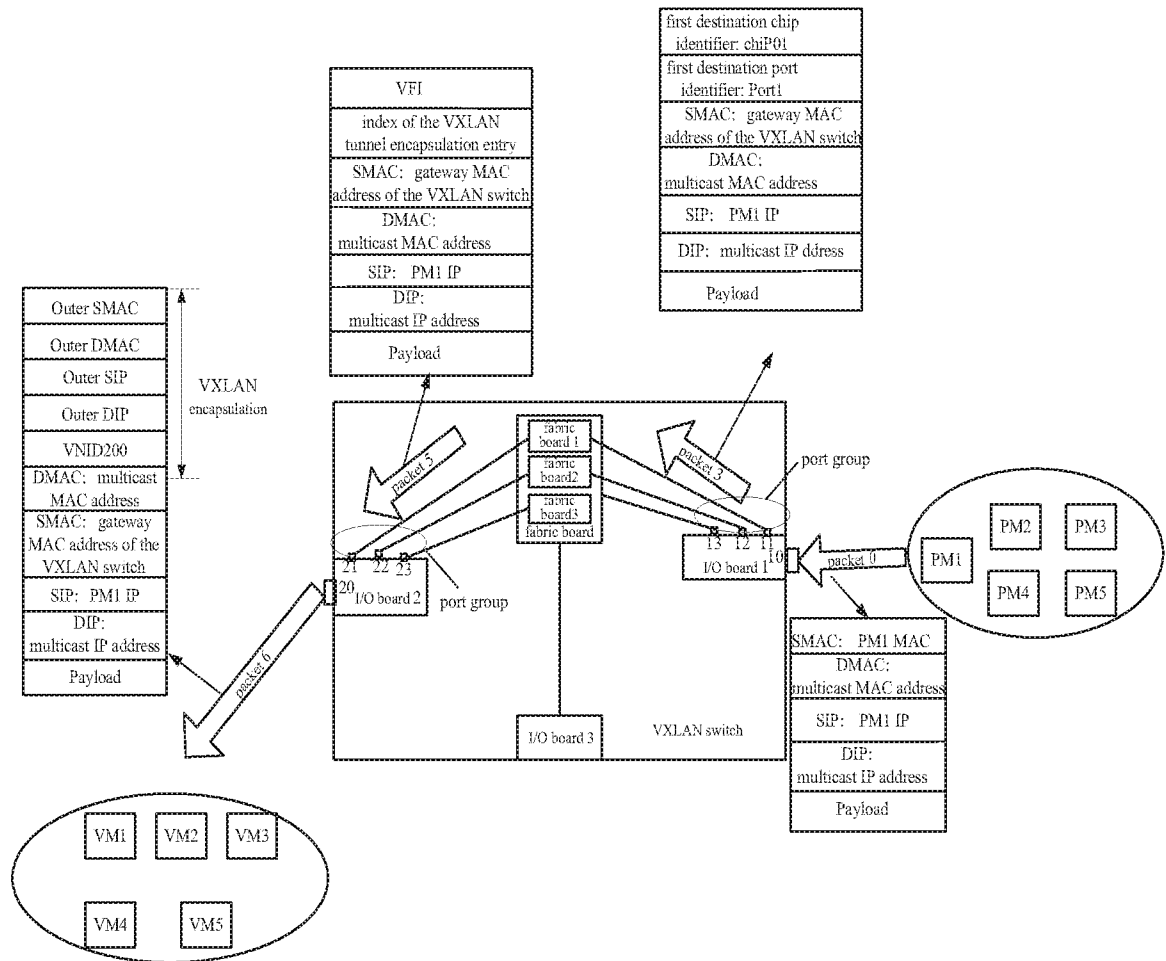


Figure 5

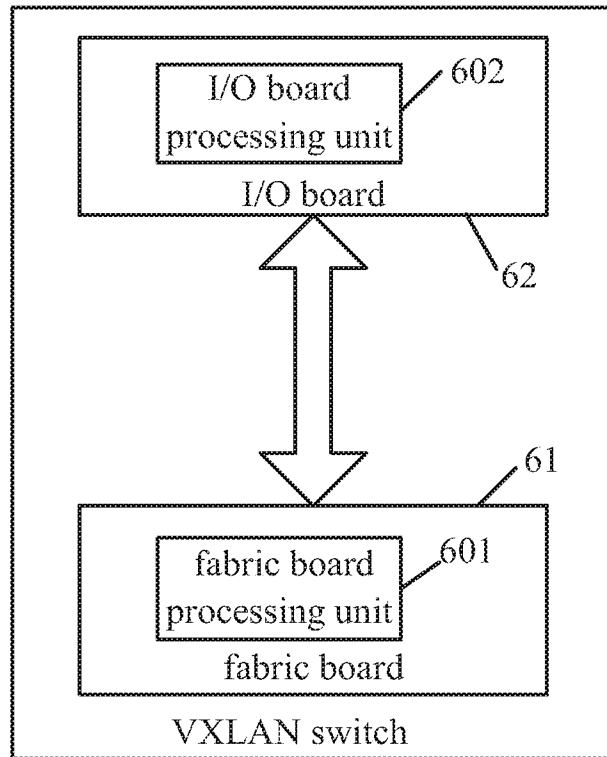


Figure 6

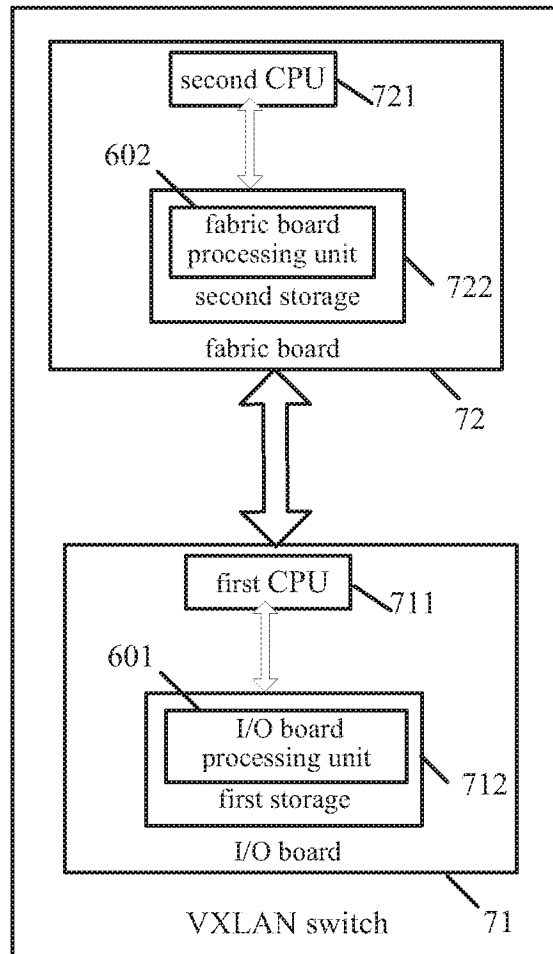


Figure 7

## INTERNATIONAL SEARCH REPORT

International application No.

**PCT/CN2016/080933****A. CLASSIFICATION OF SUBJECT MATTER**

H04L 12/741(2013.01)i; H04L 12/761(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)

USTXT;EPTXT;CNTXT;CNABS;WOTXT;VEN;CNKI:layer, vlan,vxlan, L3, fabric, table, multicast+, extended, extensible, switch,tunnel,mac,I?O,encapsulation,virtual,local area network

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2014067280 A1 (HUAWEI TECH CO LTD) 08 May 2014 (2014-05-08) the whole document	1-12
A	CN 103200069 A (HUAWEI TECH CO LTD) 10 July 2013 (2013-07-10) the whole document	1-12
A	CN 104426773 A (HITACHI METALS LTD) 18 March 2015 (2015-03-18) the whole document	1-12
A	US 2015010001 A1 (DUDA K. J. ET AL.) 08 January 2015 (2015-01-08) the whole document	1-12

 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

“&amp;” document member of the same patent family

Date of the actual completion of the international search

**14 July 2016**

Date of mailing of the international search report

**22 July 2016**

Name and mailing address of the ISA/CN

**STATE INTELLECTUAL PROPERTY OFFICE OF THE  
P.R.CHINA  
6, Xitucheng Rd., Jimen Bridge, Haidian District, Beijing  
100088, China**

Facsimile No. (86-10)62019451

Authorized officer

**FAN,Wenjing**

Telephone No. (86-10)62411255

**INTERNATIONAL SEARCH REPORT**  
**Information on patent family members**

International application No.

**PCT/CN2016/080933**

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
WO	2014067280	A1	08 May 2014	US	2015236871	A1	20 August 2015
				CN	103795636	A	14 May 2014
				EP	2905930	A1	12 August 2015
CN	103200069	A	10 July 2013	CN	103200069	B	27 January 2016
CN	104426773	A	18 March 2015	JP	2015039135	A	26 February 2015
				US	2015049765	A1	19 February 2015
US	2015010001	A1	08 January 2015	WO	2015003029	A1	08 January 2015
				US	2015010002	A1	08 January 2015
				US	9369383	B2	14 June 2016