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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
11 August 2016 (11.08.2016)

(10) International Publication Number
WO 2016/124155 A1

(51) International Patent Classification:
H04L 12/883 (2013.01) *H04L 29/12* (2006.01) *H04L 12/46* (2006.01)

(21) International Application Number: PCT/CN2016/073608

(22) International Filing Date: 5 February 2016 (05.02.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
201510064873.3 6 February 2015 (06.02.2015) CN

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

(72) Inventors: **YANG, Donghong**; Room 730, Oriental Electronic BLD., No.2 Chuangye Road, Shangdi Information Industry Base, Haidian District, Beijing 100085 (CN). **LIANG, Yin**; Room 730, Oriental Electronic BLD., No.2 Chuangye Road, Shangdi Information Industry Base, Haidian District, Beijing 100085 (CN). **PAN, Xiyuan**; Room 730, Oriental Electronic BLD., No.2 Chuangye Road, Shangdi Information Industry Base, Haidian District, Beijing 100085 (CN). **WANG, Dian**; 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: PACKET TRANSMISSION

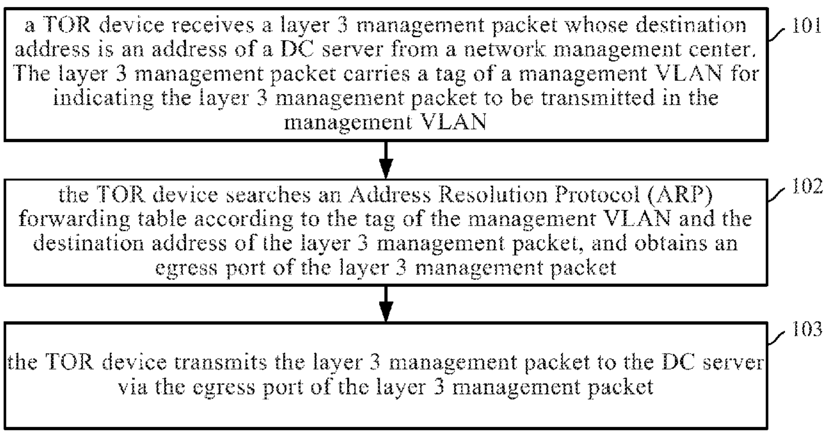

Figure 1

(57) Abstract: In an example, a TOR device receives a layer 3 management packet whose destination address is an address of a DC server from a network management center, the layer 3 management packet carries a tag of a management VLAN for indicating the layer 3 management packet to be transmitted in the management VLAN. The TOR device searches an ARP forwarding table according to the tag of the management VLAN and the destination address, and obtains an egress port of the layer 3 management packet. The egress port is a physical port in an aggregation port, the physical port is bound with the management VLAN and connected with a physical port bound with the management VLAN on the DC server, and is bound with a service VLAN. The TOR device transmits the layer 3 management packet to the DC server via the egress port of the layer 3 management packet.

WO 2016/124155 A1

PACKET TRANSMISSION

This application claims the benefit of priority from Chinese Patent Application No. 201510064873.3, entitled “Method and apparatus for managing data center servers”, filed on February 6, 2015, the entire content of which is hereby incorporated by reference.

Background

[0001] A Data Center (DC) server may be managed via an Integrated Light-Out (ILO) port, also called a server out-of-band management port. ILO is also referred to as “in case of line out”. The ILO port is a separate dedicated physical port, and the DC server sets a physical port aside as the ILO port. Management packets are transmitted between a network management device and the DC server via the ILO port.

[0002] A dedicated management network is necessary so that the network management device can manage the DC server via the ILO port. The management network is separate from a service network of the DC server.

Brief Description of the Drawings

[0003] Features of the present disclosure are illustrated by way of example and not limited in the following figure(s), in which like numerals indicate like elements, in which:

[0004] Figure 1 is a schematic flowchart illustrating a packet transmission method according to some examples of the present disclosure.

[0005] Figure 2 is a schematic flowchart illustrating a packet transmission method according to some examples of the present disclosure.

[0006] Figure 3 is a schematic diagram illustrating the structure of a DC network according to some examples of the present disclosure.

[0007] Figure 4 is a schematic diagram illustrating a TOR device according to some examples of the present disclosure.

[0008] Figure 5 is a schematic diagram illustrating a hardware structure of a TOR device according to some examples of the present disclosure.

Detailed Description

[0009] 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.

[0010] In a management packet transmission method, a DC server sets a physical port aside as an ILO port. A network management device transmits management packets via the ILO port, so as to manage the DC server. Accordingly, physical port resources of the DC server are occupied. In addition, a dedicated management network is necessary so that the network management device can manage the DC server, which causes high network construction costs.

[0011] Figure 1 is a schematic flowchart illustrating a packet transmission method according to some examples of the present disclosure. The method may be applicable to an Intelligent Resilient Framework (IRF) system. In the IRF system, multiple network devices may be connected through physical IRF ports and may be virtualized into a device after necessary configurations are performed.

[0012] As shown in Figure 1, the method includes the following blocks.

[0013] At block 101, a top-of-rack (TOR) device receives a layer 3 management packet whose destination address is an address of a DC server from a network management center. The layer 3 management packet carries a tag of a management Virtual Local Area Network (VLAN) for indicating the layer 3 management packet to be transmitted in the management VLAN.

[0014] At block 102, the TOR device searches an Address Resolution Protocol (ARP) forwarding table according to the tag of the management VLAN and the destination address of the layer 3 management packet, and obtains an egress port of the layer 3 management packet. The egress port is a physical port in an aggregation port, and the egress port is bound with the management VLAN and is connected with a physical port bound with the management VLAN on the DC server. The physical ports bound with the management VLAN may be also bound with a service VLAN.

[0015] At block 103, the TOR device transmits the layer 3 management packet to the DC server via the egress port of the layer 3 management packet.

[0016] In some examples, a physical port in the aggregation port for transmitting service packets of the DC server is bound with the management VLAN. The management packet is transmitted via the physical port bound with the management VLAN after the management packet is received. Thus, the management packet may be transmitted via a service network and it is unnecessary to set a physical port for transmitting management packets aside, thereby reducing network construction costs and utilizing physical port resources of the IRF system effectively.

[0017] In some examples, forwarding entries used for transmitting management packets in the ARP forwarding table may be preconfigured by a user or obtained by the TOR device via learning. In some examples, the TOR device broadcasts a first ARP request packet, and the first ARP request packet carries the address of the DC server and the tag of the management VLAN. The TOR device receives a first ARP response packet from the DC server. In the ARP forwarding table, a port receiving the first ARP response packet is configured as the egress port corresponding to the address of the DC server and the management VLAN. The port receiving the first ARP response packet is a physical port included in the aggregation port.

[0018] In some examples, the TOR device may learn and obtain the ARP forwarding table via the following processing.

[0019] a. after a TOR switch receives a layer 3 management VLAN packet from a network management device, if it is determined that an ARP forwarding table does not include a forwarding entry corresponding to a management VLAN tag and a destination Internet Protocol (IP) address which are carried by the layer 3 management VLAN packet, the TOR switch broadcasts an ARP request packet. The ARP request packet carries the management VLAN tag, a destination IP address of the ARP request packet is the destination address of the layer 3 management VLAN packet, and a destination Media Access Control (MAC) address is a broadcast MAC address.

[0020] b. after any one DC server connected to the TOR switch receives the ARP request packet if the DC server determines that the destination IP address of the ARP request packet is an IP address of the DC server, the DC server returns a ARP response packet. The ARP response packet carries the management VLAN tag, and a source IP address and a source MAC address of the ARP response packet are the IP address and a MAC address of the DC server respectively.

[0021] c. the TOR switch receives the ARP response packet, learns and obtains a forwarding entry as follows.

[0022] IP address: the source IP address of the ARP response packet;

[0023] MAC address: the source MAC address of the ARP response packet;

[0024] VLAN tag: the management VLAN tag;

[0025] egress port: an ingress physical port of the ARP response packet.

[0026] In some examples, if an ingress port of the ARP response packet is an aggregation port, the egress port in the forwarding entry is configured as an ingress physical port, and the ingress physical port is a physical port connected with a physical port bound with the management VLAN on the DC server. While in processing of learning a forwarding entries of a service VLAN, if the ingress port of the ARP response packet is the aggregation port, the egress port of the forwarding entry is configured as the aggregation port.

[0027] In some examples, the TOR device is located in an Intelligent Resilient Framework (IRF) system. When the TOR device determines that a new forwarding entry in the ARP forwarding table is configured or learned, the new forwarding entry is synchronized to other TOR devices in the IRF system.

[0028] In some examples, the TOR device is located in the IRF system. The TOR device receives a layer 3 service packet from a core network, and the layer 3 service packet carries an address of a DC server and a tag of a service VLAN. The TOR device searches a ARP forwarding table according to the address of the DC server and the tag of the service VLAN, and obtains an egress port of the layer 3 service packet. The egress port of the layer 3 service packet is an aggregation port. According to preconfigured load sharing algorithm, the TOR device distributes the layer 3 service packet to a physical port included in the aggregation port for transmission.

[0029] In the ARP forwarding table, the forwarding entries used for transmitting service packets may be preconfigured by a user or obtained by the TOR device via learning. In some examples, the TOR device broadcasts a second ARP request packet, and the second ARP request packet carries the address of the DC server and the tag of the service VLAN. The TOR device receives a second ARP response packet from the DC server. In the ARP forwarding table, an aggregation port including a physical port receiving the second ARP response packet is recorded as the egress port corresponding to the address of the DC server and the service VLAN.

[0030] In some examples, the physical port bound with the management VLAN may be used for transmit management packets, and the physical port bound with the service VLAN may be used for transmit service packets. The physical port may bound with both the management VLAN and the service VLAN.

[0031] If all of or parts of the physical ports in the aggregation port are bound with the service VLAN, the aggregation port may be regarded as being bound with the service VLAN. After the aggregation port is found according to the destination address of the service packet and the tag of the service VLAN, one or more physical ports may be selected from the physical ports bounding with the service VLAN in the aggregation port according to load sharing algorithm and may be used for transmitting service packets.

[0032] In some examples, the TOR device synchronizes the ARP forwarding table to other TOR devices in the same IRF system.

[0033] In some examples, the TOR device is a TOR switch.

[0034] Figure 2 is a schematic flowchart illustrating a packet transmission method according to some examples of the present disclosure. As shown in Figure 2, the method includes the following block.

[0035] At block 200, a management VLAN is configured to manage DC servers in advance. Generally, a tag of the management VLAN is VLAN 401.

[0036] AT block 201, for any one of the DC servers, one physical port of the DC server is bound with the management VLAN.

[0037] In various examples of the present disclosure, the physical port bound with the management VLAN on the DC server is called a first physical port.

[0038] The first physical port may be also bound with a service VLAN, so as to transmit service packets.

[0039] At block 202, following configurations are performed for any one TOR switch connected with the DC server in advance. In an ARP forwarding table, an egress port in a forwarding entry corresponding to the management VLAN is configured as each physical port which is on the TOR switch and is connected with a first physical port of each DC server connected with the TOR switch.

[0040] At block 203, when the TOR switch generates a forwarding entry in which a VLAN is the management VLAN and a destination address is an address of a DC server connected with the TOR switch, according to the configurations at 202, an egress port in the forwarding entry is configured as a physical port which is on the TOR switch and is connected with a first physical port of the DC server.

[0041] At block 204, the TOR switch receives a layer 3 management packet from a network management device, finds a forwarding entry according to the tag of the management VLAN and a destination address of the layer 3 management packet, and sends the layer 3 management packet to a DC server according to an egress port in the found forwarding entry.

[0042] According to the above examples, the management VLAN for managing DC servers is configured. When the TOR device learns the forwarding entry of the managing VLAN in the ARP forwarding table, the egress port in the forwarding entry is configured as the physical port on the TOR device which is connected with the physical port bound with the managing WLAN on the DC server.

[0043] Because one physical port may support multiple VLANs, the physical port on the DC server may be bound with the service VLAN after bounded with the management VLAN. Accordingly, both management packets and service packets may be transmitted via the physical port. Therefore, it is unnecessary for the DC server to set a physical port for transmitting management packets aside, thereby utilizing physical port resources of the DC server effectively.

[0044] In addition, layer 3 management packets may be successfully transmitted in an service network. Accordingly, it is unnecessary to establish a dedicated management network separate from the service network, thereby saving network establishment costs.

[0045] An application example is provided.

[0046] Figure 3 is a schematic diagram illustrating the structure of a DC network according to some examples of the present disclosure. As shown in Figure 3, a IRF system includes a TOR switch 1 and a TOR switch 2, and a DC server 1 and a DC server 2 are connected with the IRF system.

[0047] The physical port P11 of the DC server 1 is connected with the physical port G1/0/1 of the TOR switch 1.

[0048] The physical port P12 of the DC server 1 is connected with the physical port G2/0/1 of the TOR switch 2.

[0049] The physical port P21 of the DC server 2 is connected with the physical port G1/0/2 of the TOR switch 1.

[0050] The physical port P22 of the DC server 2 is connected with the physical port G2/0/2 of the TOR switch 2.

[0051] In the IRF system, the physical port G1/0/1 of the TOR switch 1 and the physical port G2/0/1 of the TOR switch 2 are bound as an aggregation port G1 according to Link Aggregation Control Protocol (LACP), and the physical port G1/0/2 of the TOR switch 1 and the physical port G2/0/2 of the TOR switch 2 are bound as an aggregation port G2 according to the LACP.

[0052] The following configurations are performed in advance.

[0053] The physical port P11 of the DC server 1 is bound with a management VLAN 401, and the physical port P21 of the DC server 2 is bound with the management VLAN 401.

[0054] On the TOR switch 1, the physical port G1/0/1 connected with the physical port P11 is bound with the management VLAN 401, and the physical port G1/0/2 connected with the physical port P21 is bound with the management VLAN 401.

[0055] Configurations are performed for the TOR switch 1 and the TOR switch 2: in an ARP forwarding table, an egress port in a forwarding entry corresponding to the management VLAN 401 may be configured as the physical port G1/0/1 or the physical port G1/0/2.

[0056] The following procedure is performed when the TOR switch learns the ARP forwarding table.

[0057] The TOR switch 1 receives an ARP response packet sent by the DC server 1 via the physical port G1/0/1 and learns a forwarding entry. The TOR switch 1 discovers that a VLAN tag in the ARP response packet is VLAN 401, and configures an egress port of the forwarding entry as the ingress physical port G1/0/1 (rather than the ingress aggregation port G1) according to the above configurations.

[0058] The TOR switch 1 receives an ARP response packet sent by the DC server 2 via the physical port G1/0/2 and learns a forwarding entry. The TOR switch 1 discovers that a VLAN tag in the ARP response packet is VLAN 401, and configures an egress port of the forwarding entry as the ingress physical port G1/0/2 (rather than the ingress aggregation port G1) according to the above configurations in B.

[0059] The TOR switch 1 receives an ARP response packet sent by the DC server 1 via the physical port G1/0/1 and learns a forwarding entry. The TOR switch 1 discovers that a VLAN tag in the ARP response packet is a service VLAN, and configures an egress port of the forwarding entry as the aggregation port G1 corresponding to the ingress physical port G1/0/1.

[0060] The TOR switch 1 receives an ARP response packet sent by the DC server 2 via the physical port G1/0/2 and learns a forwarding entry. The TOR switch 1 discovers that a VLAN tag in the ARP response packet is the service VLAN, and configures an egress port of the forwarding entry as the aggregation port G2 corresponding to the ingress physical port G1/0/2.

[0061] Similarly with the processing performed by the TOR switch 1, when the TOR switch 2 learns a forwarding entry of the service VLAN, an egress port is configured as the aggregation port G1 or G2.

[0062] According to characteristics of the IRF system, once each device in the IRF system, such as the TOR switches 1 and 2 in this example, learns a new forwarding entry, the device may synchronize the new forwarding entry to other devices in the IRF system.

[0063] The procedure of forwarding the layer 3 management packet by the TOR switch includes the following processing.

[0064] The TOR switch 1 receives a layer 3 management packet sent from a network management device to the DC server 1. The VLAN tag in the packet is the VLAN 401, and the destination IP address of the packet is the address of the DC server 1. The TOR switch 1 locally searches out a forwarding entry corresponding to the packet, and sends the packet according to the egress port in the forwarding entry, i.e. the physical port G1/0/1. The packet may arrive at the physical port P11 of the DC server 1.

[0065] The TOR switch 1 receives a layer 3 management packet sent from the network management device to the DC server 2. The VLAN tag in the packet is the VLAN 401, and the destination IP address of the packet is the address of the DC server 2. The TOR switch 1 locally searches out a forwarding entry corresponding to the packet, and sends the packet according to the egress port in the forwarding entry, i.e. the physical port G1/0/2. The packet may arrive at the physical port P21 of the DC server 2.

[0066] Because the TOR switch 1 has synchronized the learned forwarding entry of the management VLAN 401 to the TOR switch 2, if the TOR switch 2 receives a layer 3 management packet sent from the network management device to the DC server 1 or the DC server 2, the TOR switch 2 may find an forwarding entry corresponding to the received packet in an ARP forwarding table, and determine the egress port (G1/0/1 or G1/0/2) of the forwarding entry is located on the TOR switch 1. The TOR switch 2 sends the packet and information of the egress port to the TOR switch 1. After receiving the packet and information of the egress port, the TOR switch 1 directly sends the packet according to the information of the egress port.

[0067] The procedure of forwarding the layer 3 service packet by the TOR switch includes the following processing.

[0068] The TOR switch 1 receives a layer 3 service packet sent from a core network, searches the ARP forwarding table stored locally according to the VLAN tag and the destination address in the service packet, and finds a forwarding entry corresponding to the packet. The TOR switch 1 determines that the egress port in the forwarding entry is the aggregation port, and according to load sharing algorithm, distributes the packet to a physical port included in the aggregation port to send the packet.

[0069] In some examples, if the egress port in the forwarding entry is the aggregation port G1, a calculation process is performed by the TOR switch according to a predefined hash algorithm, e.g. a hash calculation may be performed for five tuple information of the received layer 3 service packet, and maps a calculation result to a physical port included in the aggregation port, e.g. G1/0/1 or G2/0/1. The TOR switch 1 then sends the packet via the mapped physical port. If the calculation result is mapped to the G2/0/1, the TOR switch 1 sends the service packet to the TOR switch 2, and the TOR switch 2 sends the service packet via the port G2/0/1.

[0070] The procedure of forwarding the layer 3 service packet by the TOR switch 2 is similar with that performed by the TOR switch 1.

[0071] According to the above examples, the layer 3 management packets may be forwarded via a fixed physical port in the IRF system, and the layer 3 service packets may be forwarded via the physical ports in the IRF system according to the load balance algorithm.

[0072] Figure 4 is a schematic diagram illustrating a TOR device according to some examples of the present disclosure. As shown in Figure 4, the TOR device includes a non-transitory storage medium 410 and a processor 420. The non-transitory storage medium 410 may store machine readable instructions that can be executed by the processor 420. The processor 420 may read the machine readable instructions stored in the non-transitory storage 410 to perform a process of:

[0073] receiving a layer 3 management packet whose destination address is an address of a DC server from a network management center, the layer 3 management packet carrying a tag of a management VLAN for indicating the layer 3 management packet to be transmitted in the management VLAN;

[0074] searching an ARP forwarding table according to the tag of the management VLAN and the destination address of the layer 3 management packet, and obtaining an egress port of the layer 3 management packet; the egress port being a physical port in an aggregation port, and the physical port being bound with the management VLAN and connected with a physical port bound with the management VLAN on the DC server; the physical port being bound with a service VLAN; and

[0075] transmitting the layer 3 management packet to the DC server via the egress port of the layer 3 management packet.

[0076] In some examples, the processor 420 may read the machine readable instructions stored in the non-transitory storage 410 to further perform a process of:

[0077] broadcasting a first ARP request packet; the first ARP request packet carrying the address of the DC server and the tag of the management VLAN;

[0078] receiving a first ARP response packet sent by the DC server;

[0079] configuring a port receiving the first ARP response packet as an egress corresponding to the address of the DC server and the management VLAN in the ARP forwarding table; the port receiving the first ARP response packet being a physical port included in the aggregation port.

[0080] In some examples, the TOR device is located in a IRF system, and the processor 420 may read the machine readable instructions stored in the non-transitory storage 410 to further perform a process of:

[0081] receiving a layer 3 service packet from a core network, and the layer 3 service packet carrying the address of the DC server and a tag of the service VLAN;

[0082] searching the ARP forwarding table according to the tag of the service VLAN and the address of the DC server, and obtaining an egress port of the layer 3 service packet; the egress port of the layer 3 service packet being the aggregation port;

[0083] transmitting the service packet via a physical port included in the aggregation port according to preconfigured load sharing algorithm.

[0084] In some examples, the processor 420 may read the machine readable instructions stored in the non-transitory storage 410 to further perform a process of:

[0085] broadcasting a second ARP request packet; the second ARP request packet carrying the address of the DC server and the tag of the service VLAN;

[0086] receiving a second ARP response packet sent by the DC server;

[0087] configuring an aggregation port including a physical port receiving the second ARP response packet as an egress corresponding to the address of the DC server and the service VLAN in the ARP forwarding table.

[0088] In some examples, the processor 420 may read the machine readable instructions stored in the non-transitory storage 410 to further perform a process of: synchronizing the ARP forwarding table to each of other TOR devices in the IRF system.

[0089] A non-transitory storage medium is provided according to some examples of the present disclosure. The non-transitory storage medium may store machine readable instructions, and a processor may read the machine readable instructions stored in the non-transitory storage medium to perform a process of:

[0090] receiving a layer 3 management packet whose destination address is an address of a DC server from a network management center; the layer 3 management packet carrying a tag of a management VLAN for indicating the layer 3 management packet to be transmitted in the management VLAN;

[0091] searching an ARP forwarding table according to the tag of the management VLAN and the destination address of the layer 3 management packet, and obtaining an egress port of the layer 3 management packet; the egress port being a physical port in an aggregation port, and the physical port being bound with the management VLAN and connected with a physical port bound with the management VLAN on the DC server; the physical port being bound with a service VLAN; and

[0092] transmitting the layer 3 management packet to the DC server via the egress port of the layer 3 management packet.

[0093] In some examples, the processor may read the machine readable instructions stored in the non-transitory storage to further perform a process of:

[0094] broadcasting a first ARP request packet; the first ARP request packet carrying the address of the DC server and the tag of the management VLAN;

[0095] receiving a first ARP response packet sent by the DC server;

[0096] configuring a port receiving the first ARP response packet as an egress corresponding to the address of the DC server and the management VLAN in the ARP forwarding table; the port receiving the first ARP response packet being a physical port included in the aggregation port.

[0097] In some examples, the TOR device is located in a IRF system, and the processor may read the machine readable instructions stored in the non-transitory storage to further perform a process of:

[0098] receiving a layer 3 service packet from a core network, and the layer 3 service packet carrying the address of the DC server and a tag of the service VLAN;

[0099] searching the ARP forwarding table according to the tag of the service VLAN and the address of the DC server, and obtaining an egress port of the layer 3 service packet; the egress port of the layer 3 service packet being the aggregation port; and

[0100] transmitting the service packet via a physical port included in the aggregation port according to preconfigured load sharing algorithm.

[0101] In some examples, the processor may read the machine readable instructions stored in the non-transitory storage to further perform a process of:

[0102] broadcasting a second ARP request packet; the second ARP request packet carrying the address of the DC server and the tag of the service VLAN;

[0103] receiving a second ARP response packet sent by the DC server;

[0104] configuring an aggregation port including a physical port receiving the second ARP response packet as an egress corresponding to the address of the DC server and the service VLAN in the ARP forwarding table.

[0105] In some examples, the processor may read the machine readable instructions stored in the non-transitory storage to further perform a process of: synchronizing the ARP forwarding table to each of other TOR devices in the IRF system.

[0106] The non-transitory storage medium for storing the machine readable instructions may be any electric, magnetic or optical storage medium or other physical storage medium. The machine readable storage medium may include or store information, e.g. executable instructions and data etc. For example, the non-transitory storage medium may be a Radom Access Memory (RAM), a Flash card, a storage driver (e.g. a disk driver), a solid-state disk, a storage disk of any type (e.g. an optical disk, a DVD), or a similar storage medium, or any combination of the above.

[0107] Figure 5 is a schematic diagram illustrating a hardware structure of a TOR device according to some examples of the present disclosure. As shown in Figure 5, the TOR device includes a receiving unit 510, a searching unit 520 and a transmitting unit 530.

[0108] The receiving unit 510 is to receive a layer 3 management packet whose destination address is an address of a DC server from a network management center. The layer 3 management packet carries a tag of a management VLAN for indicating that the layer 3 management packet to be transmitted in the management VLAN.

[0109] The searching unit 520 is to search an ARP forwarding table according to the tag of the management VLAN and the destination address of the layer 3 management packet, and obtain an egress port of the layer 3 management packet. The egress port is a physical port in an aggregation port, and the physical port is bound with the management VLAN and connected with a physical port bound with the management VLAN on the DC server; the physical port is bound with a service VLAN.

[0110] The transmitting unit 530 is to transmit the layer 3 management packet to the DC server via the egress port of the layer management packet.

[0111] In some examples, the transmitting unit 530 is further to broadcast a first ARP request packet; the first ARP request packet carries the address of the DC server and the tag of the management VLAN.

[0112] The receiving unit 510 is further to receive a first ARP response packet sent by the DC server.

[0113] As shown in Figure 6, the TOR device further includes a configuring unit 540. The configuring unit 540 is to configure a port receiving the first ARP response packet as an egress corresponding to the address of the DC server and the management VLAN in the ARP forwarding table; the port receiving the first ARP response packet is a physical port included in the aggregation port.

[0114] In some examples, the TOR device is located in a IRF system.

[0115] The receiving unit 510 is further to receive a layer 3 service packet from a core network, and the layer 3 service packet carries the address of the DC server and a tag of the service VLAN.

[0116] The searching unit 520 is further to search the ARP forwarding table according to the tag of the service VLAN and the address of the DC server, and obtain an egress port of the layer 3 service packet. The egress port of the layer 3 service packet is the aggregation port.

[0117] The transmitting unit 530 is further to transmit the service packet via a physical port included in the aggregation port according to preconfigured load sharing algorithm.

[0118] In some examples, the transmitting unit 530 is further to broadcast a second ARP request packet; the second ARP request packet carries the address of the DC server and the tag of the management VLAN.

[0119] The receiving unit 510 is further to receive a second ARP response packet sent by the DC server.

[0120] The configuration unit 540 is further to configure an aggregation port including a physical port receiving the second ARP response packet as an egress corresponding to the address of the DC server and the service VLAN in the ARP forwarding table.

[0121] In some examples, the transmitting unit 530 is further to synchronize the ARP forwarding table to each of other TOR devices in the IRF system.

[0122] 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.

[0123] 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.

Claims

1. A packet transmission method, comprising:

receiving, by a top-of-rack (TOR) device, a layer 3 management packet whose destination address is an address of a Data Center (DC) server from a network management center; the layer 3 management packet carrying a tag of a management Virtual Local Area Network (VLAN) for indicating the layer 3 management packet to be transmitted in the management VLAN;

searching an Address Resolution Protocol (ARP) forwarding table according to the tag of the management VLAN and the destination address of the layer 3 management packet, and obtaining an egress port of the layer 3 management packet; the egress port being a physical port in an aggregation port, and the physical port being bound with the management VLAN and connected with a physical port bound with the management VLAN on the DC server; the physical port being bound with a service VLAN; and

transmitting the layer 3 management packet to the DC server via the egress port of the layer 3 management packet.

2. The method of claim 1, further comprising:

before searching the ARP forwarding table,

broadcasting a first ARP request packet; the first ARP request packet carrying the address of the DC server and the tag of the management VLAN;

receiving a first ARP response packet sent by the DC server;

configuring a port receiving the first ARP response packet as an egress corresponding to the address of the DC server and the management VLAN in the ARP forwarding table; the port receiving the first ARP response packet being a physical port included in the aggregation port.

3. The method of claim 1, wherein the TOR device is located in an Intelligent Resilient Framework (IRF) system, and the method further comprises:

receiving a layer 3 service packet from a core network, the layer 3 service packet carrying the address of the DC server and a tag of the service VLAN;

searching the ARP forwarding table according to the tag of the service VLAN and the address of the DC server, and obtaining an egress port of the layer 3 service packet;

the egress port of the layer 3 service packet being the aggregation port;

transmitting the service packet via a physical port included in the aggregation port according to a preconfigured load sharing algorithm.

4. The method of claim 3, further comprising:

before searching the ARP forwarding table,

broadcasting a second ARP request packet; the second ARP request packet carrying the address of the DC server and the tag of the service VLAN;

receiving a second ARP response packet sent by the DC server;

configuring an aggregation port including a physical port receiving the second ARP response packet as an egress corresponding to the address of the DC server and the service VLAN in the ARP forwarding table.

5. The method of claim 3, further comprising:

synchronizing the ARP forwarding table to each of other TOR devices in the IRF system.

6. A top-of-rack (TOR) device, comprising a non-transitory storage medium and a processor; wherein

the non-transitory storage medium storing machine readable instructions executable by the processor; and

the processor is to read the machine readable instructions stored in the non-transitory storage medium to perform a process of:

receiving a layer 3 management packet whose destination address is an address of a Data Center (DC) server from a network management center; the layer 3 management packet carrying a tag of a management Virtual Local Area Network (VLAN) for indicating the layer 3 management packet to be transmitted in the management VLAN;

searching an Address Resolution Protocol (ARP) forwarding table according to the tag of the management VLAN and the destination address of the layer 3 management packet, and obtaining an egress port of the layer 3 management packet; the egress port being a physical port in an aggregation port, and the physical port being bound with the management VLAN and connected with a physical port bound with the management VLAN on the DC server; the physical port being bound with a service VLAN; and transmitting the layer 3 management packet to the DC server via the egress port of the layer 3 management packet.

7. The TOR device of claim 6, wherein the processor is to read the machine readable instructions stored in the non-transitory storage medium to further perform a process of:

broadcasting a first ARP request packet; the first ARP request packet carrying an address of the DC server and the tag of the management VLAN;

receiving a first ARP response packet sent by the DC server;

configuring a port receiving the first ARP response packet as an egress corresponding to the address of the DC server and the management VLAN in the ARP forwarding table; the port receiving the first ARP response packet being a physical port included in the aggregation port.

8. The TOR device of claim 6, wherein the TOR device is located in an Intelligent Resilient Framework (IRF) system, and the processor is to read the machine readable instructions stored in the non-transitory storage medium to further perform a process of:

receiving a layer 3 service packet from a core network, and the layer 3 service packet carrying the address of the DC server and a tag of the service VLAN;

searching the ARP forwarding table according to the tag of the service VLAN and the address of the DC server, and obtaining an egress port of the layer 3 service packet; the egress port of the layer 3 service packet being the aggregation port;

transmitting the service packet via a physical port included in the aggregation port according to preconfigured load sharing algorithm.

9. The TOR device of claim 8, wherein the processor is to read the machine readable instructions stored in the non-transitory storage medium to further perform a process of:

broadcasting a second ARP request packet; the second ARP request packet carrying the address of the DC server and the tag of the service VLAN;

receiving a second ARP response packet sent by the DC server;

configuring an aggregation port including a physical port receiving the second ARP response packet as an egress corresponding to the address of the DC server and the service VLAN in the ARP forwarding table.

10. The TOR device of claim 8, wherein the processor is to read the machine readable instructions stored in the non-transitory storage medium to further perform a process of:

synchronizing the ARP forwarding table to each of other TOR devices in the IRF system.

11. A non-transitory storage medium, located in a top-of-rack (TOR) device, storing machine readable instructions executable by a processor, wherein the machine readable instructions are executed by the processor to perform a process of:

receiving a layer 3 management packet whose destination address is an address of a Data Center (DC) server from a network management center; the layer 3 management packet carrying a tag of a management Virtual Local Area Network (VLAN) for indicating the layer 3 management packet to be transmitted in the management VLAN;

searching an Address Resolution Protocol (ARP) forwarding table according to the tag of the management VLAN and the destination address of the layer 3 management packet, and obtaining an egress port of the layer 3 management packet; the egress port being a physical port in an aggregation port, and the physical port being bound with the management VLAN and connected with a physical port bound with the management VLAN on the DC server; the physical port being bound with a service VLAN;

transmitting the layer 3 management packet to the DC server via the egress port of the layer 3 management packet.

12. The non-transitory storage medium of claim 11, wherein the machine readable instructions are executed by the processor to further perform a process of:

broadcasting a first ARP request packet; the first ARP request packet carrying the address of the DC server and the tag of the management VLAN;

receiving a first ARP response packet sent by the DC server;

configuring a port receiving the first ARP response packet as an egress corresponding to the address of the DC server and the management VLAN in the ARP forwarding table; the port receiving the first ARP response packet being a physical port included in the aggregation port.

13. The non-transitory storage medium of claim 11, wherein the TOR device is

located in an Intelligent Resilient Framework (IRF) system, and the machine readable instructions are executed by the processor to further perform a process of:

receiving a layer 3 service packet from a core network, and the layer 3 service packet carrying the address of the DC server and a tag of the service VLAN;

searching the ARP forwarding table according to the tag of the service VLAN and the address of the DC server, and obtaining an egress port of the layer 3 service packet; the egress port of the layer 3 service packet being the aggregation port;

transmitting the service packet via a physical port included in the aggregation port according to preconfigured load sharing algorithm.

14. The non-transitory storage medium of claim 13, wherein the machine readable instructions are executed by the processor to further perform a process of:

broadcasting a second ARP request packet; the second ARP request packet carrying the address of the DC server and the tag of the service VLAN;

receiving a second ARP response packet sent by the DC server;

configuring an aggregation port including a physical port receiving the second ARP response packet as an egress corresponding to the address of the DC server and the service VLAN in the ARP forwarding table.

15. The non-transitory storage medium of claim 13, wherein the machine readable instructions are executed by the processor to further perform a process of:

synchronizing the ARP forwarding table to each of other TOR devices in the IRF system.

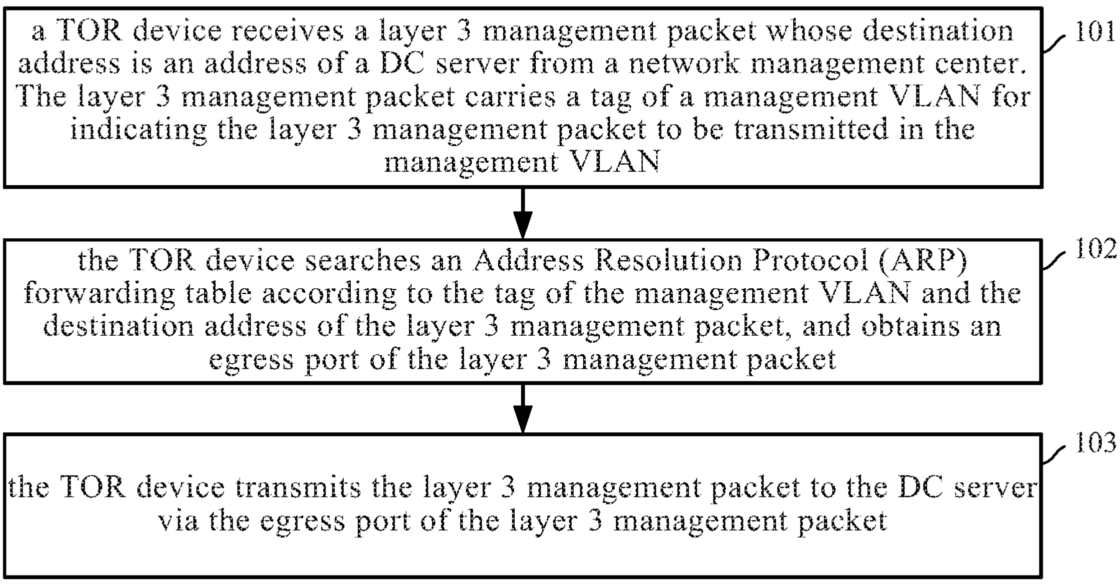
a TOR device receives a layer 3 management packet whose destination address is an address of a DC server from a network management center. The layer 3 management packet carries a tag of a management VLAN for indicating the layer 3 management packet to be transmitted in the management VLAN
101
the TOR device searches an Address Resolution Protocol (ARP) forwarding table according to the tag of the management VLAN and the destination address of the layer 3 management packet, and obtains an egress port of the layer 3 management packet
102
the TOR device transmits the layer 3 management packet to the DC server via the egress port of the layer 3 management packet
103

Figure 1

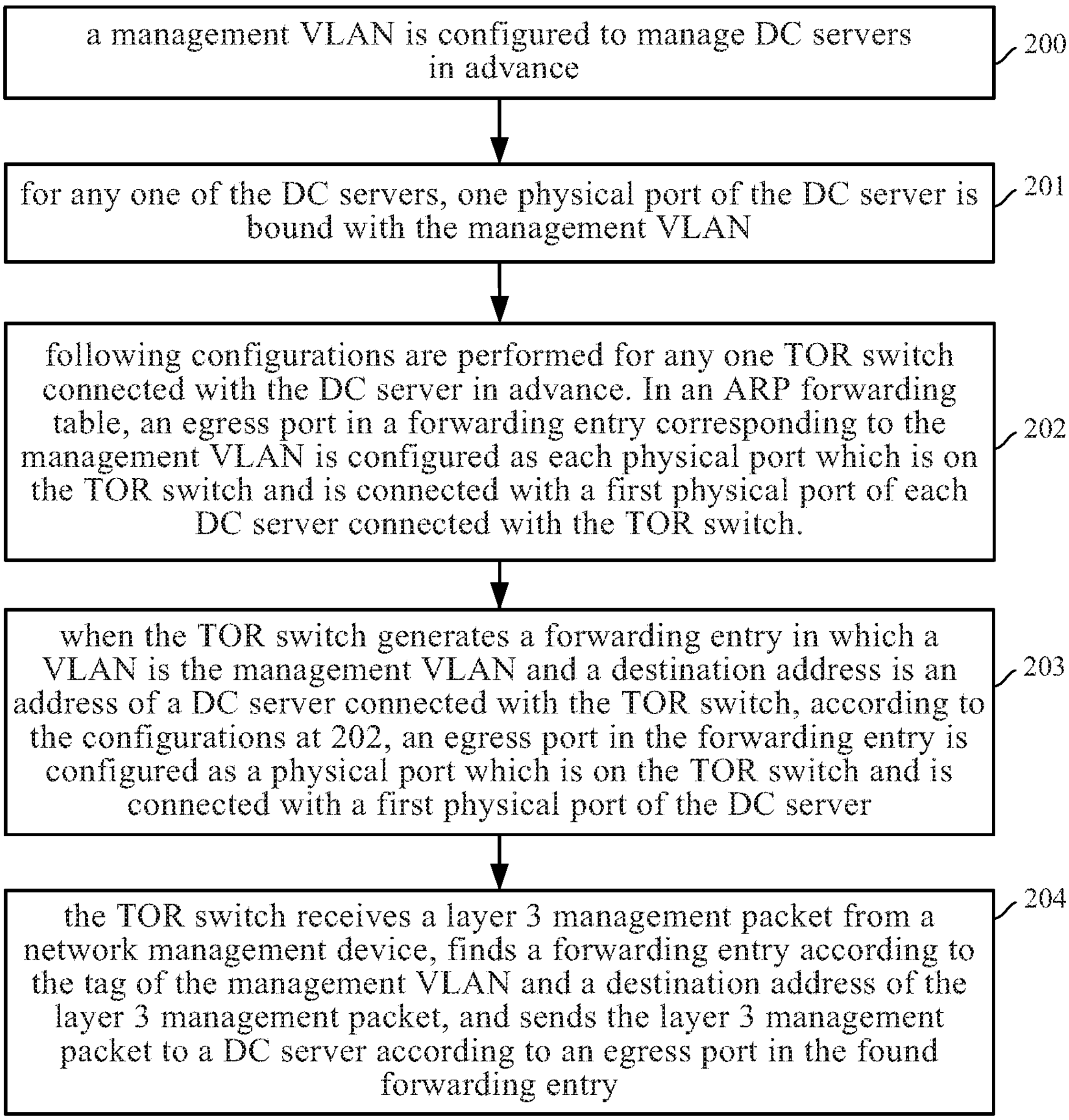
a management VLAN is configured to manage DC servers in advance
200
for any one of the DC servers, one physical port of the DC server is bound with the management VLAN
201
following configurations are performed for any one TOR switch connected with the DC server in advance. In an ARP forwarding table, an egress port in a forwarding entry corresponding to the management VLAN is configured as each physical port which is on the TOR switch and is connected with a first physical port of each DC server connected with the TOR switch.
202
when the TOR switch generates a forwarding entry in which a VLAN is the management VLAN and a destination address is an address of a DC server connected with the TOR switch, according to the configurations at 202, an egress port in the forwarding entry is configured as a physical port which is on the TOR switch and is connected with a first physical port of the DC server
203
the TOR switch receives a layer 3 management packet from a network management device, finds a forwarding entry according to the tag of the management VLAN and a destination address of the layer 3 management packet, and sends the layer 3 management packet to a DC server according to an egress port in the found forwarding entry
204

Figure 2

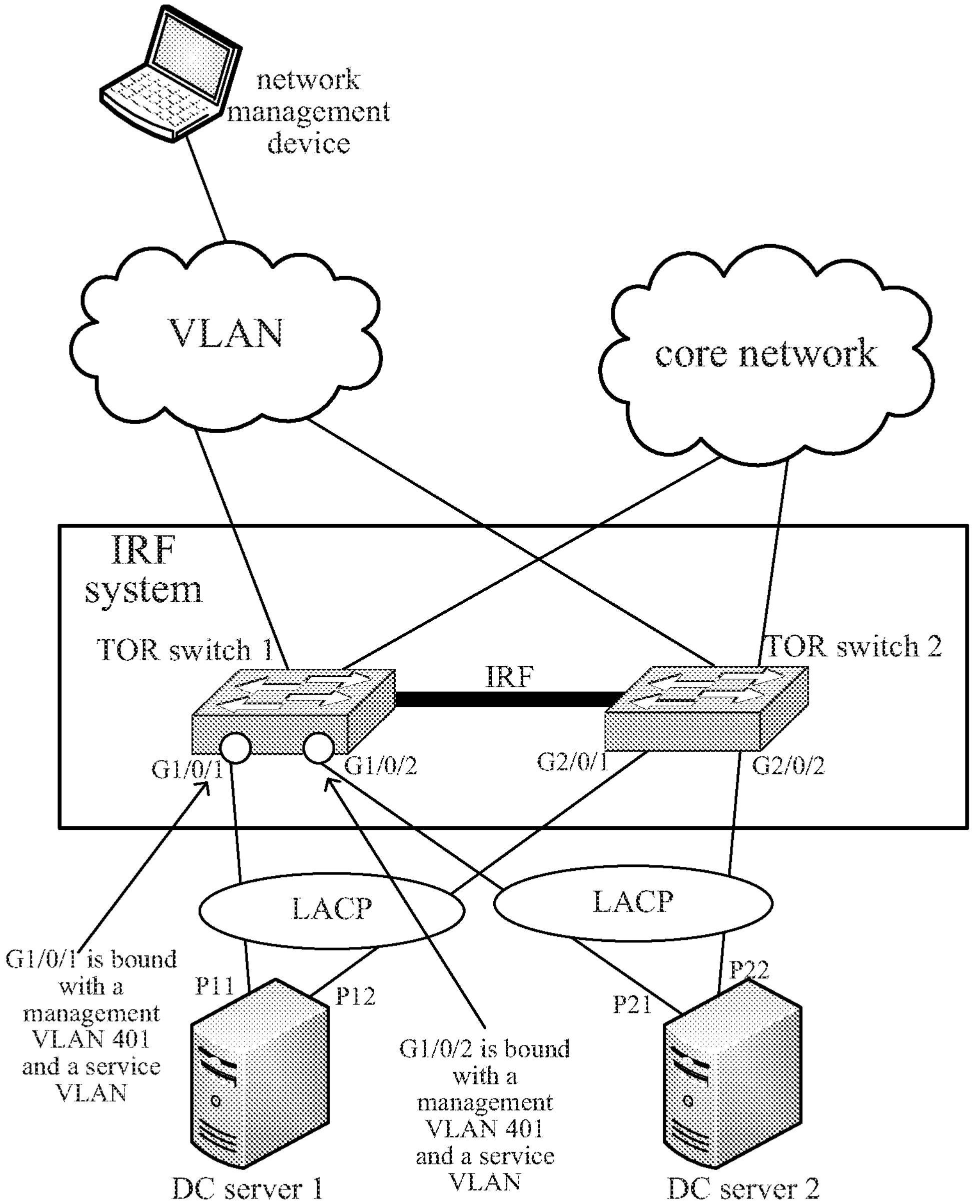
network management device
VLAN
core network
IRF system
TOR switch 1
IRF
TOR switch 2
G1/0/1
G1/0/2
G2/0/1
G2/0/2
LACP
LACP
G1/0/1 is bound with a management VLAN 401 and a service VLAN
P11
P12
P21
P22
G1/0/2 is bound with a management VLAN 401 and a service VLAN
DC server 1
DC server 2

Figure 3

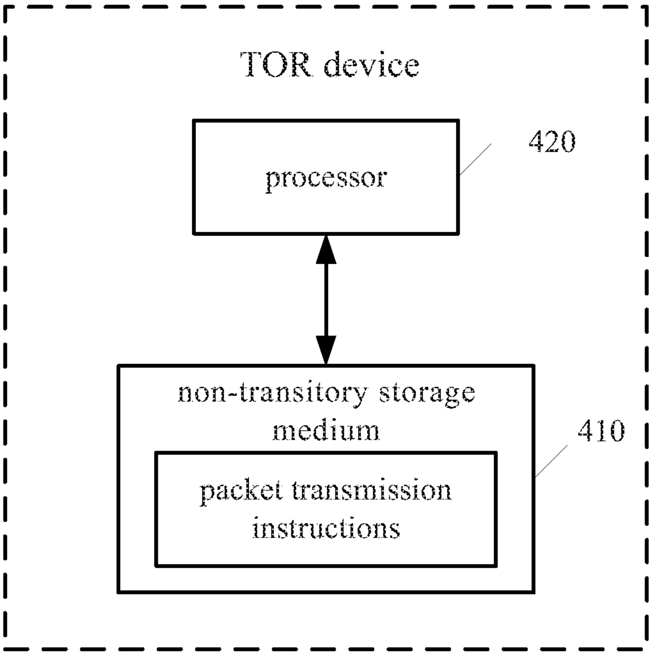

Figure 4

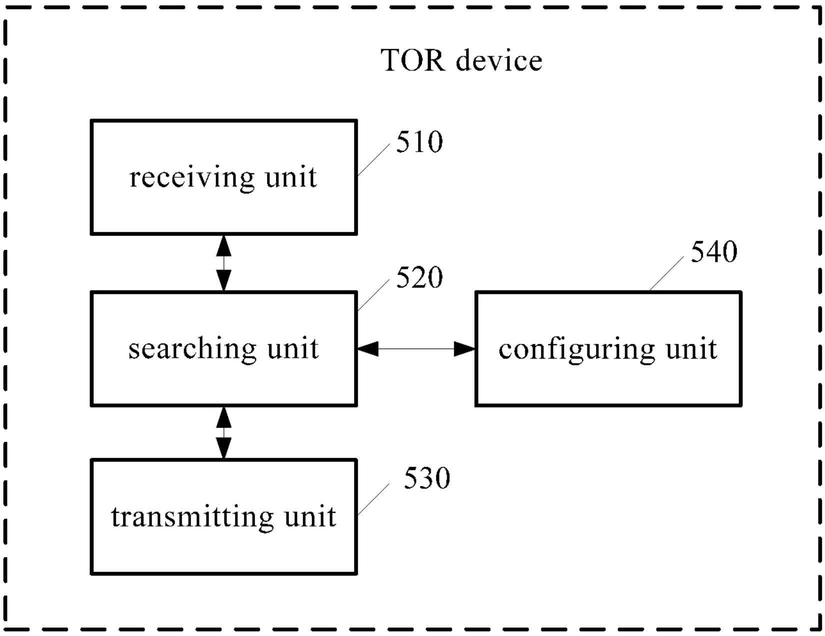

Figure 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2016/073608

A. CLASSIFICATION OF SUBJECT MATTER

H04L 12/883(2013.01)i; H04L 12/46(2006.01)i; H04L 29/12(2006.01)i

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04L

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

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

WPI;EPODOC;IEEE;CNPAT;CNKI:packet, forward+, data center, DC, layer, three, destination address, management, VLAN, virtual local area network, egress, port, TOR, ARP

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 101764752 A (HANGZHOU H3C TECHNOLOGIES CO., LTD.) 30 June 2010 (2010-06-30) paragraphs [0007]-[0051] in description	1-15
A	CN 103227757 A (HANGZHOU H3C TECHNOLOGIES CO., LTD.) 31 July 2013 (2013-07-31) the whole document	1-15
A	CN 103166874 A (HANGZHOU H3C TECHNOLOGIES CO., LTD.) 19 June 2013 (2013-06-19) the whole document	1-15
A	US 2012287785 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 15 November 2012 (2012-11-15) the whole document	1-15

☐ 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
15 April 2016	**13 May 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**	**YAN,Yue**
Facsimile No. **(86-10)62019451**	Telephone No. **(86-10)010-62413472**

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2016/073608

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 101764752 A	30 June 2010	None	
CN 103227757 A	31 July 2013	CN 103227843 A	31 July 2013
		WO 2014032620 A1	06 March 2014
		US 2015172190 A1	18 June 2015
		US 2015169351 A1	18 June 2015
		EP 2891284 A1	08 July 2015
		WO 2014032623 A1	06 March 2014
		WO 2014032624 A1	06 March 2014
		CN 103259727 A	21 August 2013
		EP 2891287 A1	08 July 2015
		EP 2891304 A1	08 July 2015
		US 2015222543 A1	06 August 2015
CN 103166874 A	19 June 2013	US 2016065463 A1	03 March 2016
		WO 2014154123 A1	02 October 2014
		EP 2979402 A1	03 February 2016
US 2012287785 A1	15 November 2012	US 2012320749 A1	20 December 2012
		CN 102821038 A	12 December 2012