



(51) International Patent Classification:
H04L 12/46 (2006.01)

(21) International Application Number:
PCT/CN2014/089959

(22) International Filing Date:
31 October 2014 (31.10.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
201310534091.2 31 October 2013 (31.10.2013) CN

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

(72) Inventor: **CHANG, Huifeng**; 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 FORWARDING IN DATA CENTER NETWORK

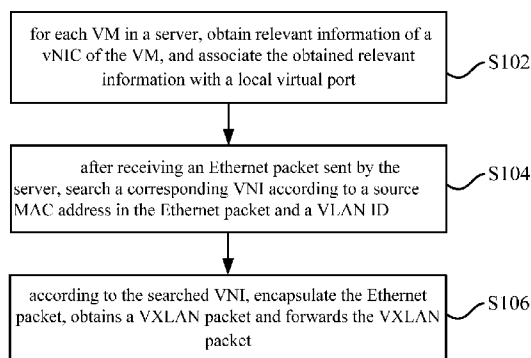


FIG. 2

(57) Abstract: The present disclosure provides a method and apparatus for forwarding a packet, wherein the method comprising: for each Virtual Machine (VM) in the server, obtaining relevant information of a virtual Network Interface Controller (vNIC) of the VM, and associating the obtained relevant information with a local virtual port, wherein the relevant information includes a Media Access Control (MAC) address of the vNIC, a Virtual eXtensible Local Area Network (VXLAN) Network Identifier (VNI) of the VXLAN where the vNIC is located, and a Virtual Local Area Network (VLAN) Identifier (ID) of a VLAN associated with the VXLAN where the vNIC is located; after receiving an Ethernet packet sent by the server, searching a VNI according to a source MAC address of the Ethernet packet and a VLAN ID; and according to the searched VNI, encapsulating the Ethernet packet, obtaining a VXLAN packet, and forwarding the VXLAN packet.

PACKET FORWARDING IN DATA CENTER NETWORK

BACKGROUND

[0001] Virtual eXtensible Local Area Network (VXLAN) is a simple mechanism for encapsulating an Ethernet packet in User Data Protocol (UDP). VXLAN can establish a virtual two-layer sub-network across multiple physical Internet Protocol (IP) sub-networks. VXLAN can deploy millions of virtual networks within a data center through a tenant ID (i.e., VXLAN ID, or VXLAN Network Identifier (VNI), or VXLAN Segment ID) with 24 bits. In addition, it is much easy to implement a format of a VXLAN packet in hardware, hence it may be possible to integrate VXLAN with physical network devices more closely in the future.

BRIEF DESCRIPTION OF THE DRAWINGS

[0002] FIG. 1 is a schematic diagram illustrating a network including a data center network, a Virtual Machine (VM) management center and a network management server according to a first example of the present disclosure.

[0003] FIG. 2 is a flow diagram illustrating a method for forwarding a packet in a data center network according to a first example of the present disclosure.

[0004] FIG. 3 is a flow diagram illustrating a method for obtaining relevant information of a virtual Network Interface Controller (vNIC) according to a first example of the present disclosure.

[0005] FIG. 4 is a schematic diagram illustrating a format of Type Length Value (TLV) in a Virtual Station Interface (VSI) Discovery and configuration Protocol (VDP) packet according to a first example of the present disclosure.

[0006] FIG. 5 is a schematic diagram illustrating a format of Filter Info domain when the value of Filter Format domain in FIG. 4 is 0x03 according to a first example of the present disclosure.

[0007] FIG. 6 is a schematic diagram illustrating a format of Filter Info domain when the value of Filter Format domain in FIG. 4 is 0x04 according to a first example of the present disclosure.

[0008] FIG. 7 is a flow diagram illustrating a method for obtaining relevant information of a vNIC according to an example of the present disclosure.

[0009] FIG. 8 is a flow diagram illustrating a method for obtaining relevant information of a vNIC according to another example of the present disclosure.

[0010] FIG. 9 is a flow diagram illustrating a method for obtaining relevant information of a vNIC according to yet another example of the present disclosure.

[0011] FIG. 10 is a flow diagram illustrating a method for forwarding a packet in a data center network according to a second example of the present disclosure.

[0012] FIG. 11 is a schematic diagram illustrating a structure of an apparatus for forwarding a packet in a data center network according to a second example of the present disclosure.

[0013] FIG. 12 is a schematic diagram illustrating a hardware structure of a device for forwarding a packet in a data center network according to a third example of the present disclosure.

DETAILED DESCRIPTION

[0014] In order to solve the problem of low forwarding efficiency caused by using a virtual Switch (vSwitch) as an VXLAN Tunnel End Point (VTEP) implemented by a software in the prior art, the following examples provide a method for forwarding a packet in a data center network and an apparatus for applying the method.

[0015] In examples of the present disclosure, as shown in FIG. 1, a data center network includes at least one server, a vSwitch and a VM configured in the server, a physical switch outside of the server and connected with a VM via a vSwitch. The server may be a direct connect server which is directly connected with the physical switch. An overlay function (i.e., a two-layer packet carried over a three-layer network) is enabled in the physical switch. For

example, the overlay may be VXLAN, or Network Virtualization using Generic Routing Encapsulation (NVGRE), etc.

[0016] In the following examples, a VXLAN function is enabled in the physical switch, and the physical switch works as a VTEP. A vNIC in VM connected with a vSwitch may also be called as a VSI.

[0017] Example 1

[0018] The method for forwarding a pack described in examples of the present disclosure is performed by the physical switch in the data center network as shown in FIG. 1. FIG. 2 is a flow diagram illustrating a method for forwarding a packet in a data center network according to a first example of the present disclosure. As shown in FIG. 2, the method includes the following procedures.

[0019] At block S102, for each VM in a server, obtain relevant information of a vNIC of the VM, and associate the obtained relevant information with a local virtual port.

[0020] The relevant information includes a MAC address of the vNIC, a VNI of the VXLAN where the vNIC is located, and a Virtual Local Area Network (VLAN) ID of a VLAN associated with the VXLAN where the vNIC is located.

[0021] At block S104, after receiving an Ethernet packet sent by the server, search a corresponding VNI according to a source MAC address in the Ethernet packet and a VLAN ID.

[0022] In particular, after a VM in the server sends a data packet, a directly connected vSwitch receives the data packet, and encapsulates the data packet as an Ethernet packet, then sends it to the directly connected physical switch.

[0023] At block S106, according to the searched VNI, encapsulate the Ethernet packet, obtain a VXLAN packet and forward the VXLAN packet. The specific way of encapsulating may follow the prior art, which is not described in detail herein.

[0024] In implementation, there are some ways to obtain relevant information of a vNIC. The following are four examples for illustration.

[0025] FIG. 3 is a flow diagram illustrating a method for obtaining relevant information of a vNIC according to a first example of the present disclosure. The method includes the following procedures.

[0026] In block S202, a physical switch establishes a VXLAN in advance, and specifies a VNI corresponding to the VXLAN.

[0027] In block S204, a network management server divides service types of the VM into groups, and each group of service types corresponds to a port strategy; allocates a vNIC type for each port strategy, and each vNIC type has a vNIC Type Identifier (VTID) and a version number; and stores the above-described relationship in a vNIC Type DataBase (VTDB) of the network management server, i.e., stores the corresponding relationship among the service type group, the port strategy and the vNIC type to the VTDB.

[0028] Service types of the VM may include webserver, database, e-mail, etc.

[0029] The port strategy may include some port behavior, e.g., restricting the speed, allowing a data flow to pass the VLAN, etc.

[0030] In block S206, by communication with the network management server, a VM manager (also called as a VM management apparatus) searches an applicable vNIC type in the VTDB, associates the vNIC of the VM with the searched vNIC type, and specifies a tenant where the vNIC is located at the same time, i.e., specifies the VXLAN where the vNIC is located.

[0031] In particular, associating the vNIC of the VM with the searched vNIC type includes, associating the MAC address of the vNIC with the VTID and the version number of the vNIC type; or, associating other information for uniquely identifying the vNIC with the VTID and the version number of the vNIC type.

[0032] In block S208, the VM manager pushes the association information in block S206 to the server, i.e., pushes the association relationship among the vNIC of the VM, the vNIC type (in particular, the VTID and the version number), and the VXLAN where the vNIC is located (in particular, the VNI) to the server.

[0033] In block S210, before the VM is on line, the server sends information corresponding to the vNIC of the VM to the directly connected physical switch. The information includes the MAC address of the vNIC and the VNI of the VXLAN where the vNIC is located. Furthermore, it may also include the VTID and the version number of the vNIC type associated with the vNIC.

[0034] Specifically, the server may carry the above-described information through a VDP packet. As shown in FIG. 4, a TLV of the VDP packet may include a Type domain, a Length domain, a Reason domain, a VSI Type ID domain, a VSI ID Format domain, a VSI ID domain, a Filter Format domain, and a Filter Info domain.

[0035] Besides the Filter Format domain and the Filter Info domain, values and meaning of other domains may be same with the prior art, which is not described in detail herein. The following will describe the Filter Format domain and the Filter Info domain.

[0036] (1) Filter Format domain

[0037] Values of the Filter Format domain are shown in Table 1.

Value	Meaning
0x01	VLAN (C-VLAN)
0x02	MAC+VLAN
0x03	Group ID+VLAN
0x04	Group ID+ MAC+ VLAN
Others	undefined

Table 1 Values and meaning of Filter Format domain

[0038] (2) Filter Info domain

[0039] Content carried in the Filter Info domain is related with values of the Filter Format domain. For example, when the value of Filter Format domain is 0x03 or 0x04, a Group ID may be carried in the Filter Info domain. When the value of Filter Format domain is 0x03, the format of the Filter Info domain is shown in FIG. 5. When the value of Filter Format domain is 0x04, the format of the Filter Info domain is shown in FIG. 6.

[0040] In the example, a VNI may be carried in the Group ID field of FIG. 5 or FIG. 6. Because the Group ID field uses 4 bytes, in implementation, it may set a first high byte of those 4 bytes to be 0, and three low bytes to be a VNI (24 bits).

[0041] Hence, in block S210, it may use the Filter Info domain to send the MAC address of the vNIC of the VM and the VNI of the VXLAN where the vNIC of the VM is located to the physical switch. In addition, because there is a VLAN domain in the Filter Info domain, it may fill an invalid VLAN ID when sending, e.g., VLAN 0 or VLAN 4095.

[0042] In block S212, after a VDP packet is received, the physical switch extracts the MAC address, VNI, VTID and the version number, and then searches whether there is a corresponding VXLAN available in itself according to the extracted VNI (i.e., whether the corresponding VXLAN is already established).

[0043] If no, perform block S214; if yes, perform block S216.

[0044] In block S214, the physical switch returns a VDP failure response to the server, and the procedure ends.

[0045] In block S216, the physical switch determines whether the searched VXLAN is associated with a VLAN. If yes, perform block S218; if no, perform block S220.

[0046] In block S218, when the VXLAN is already associated with a VLAN, the physical switch searches the VLAN associated with the VXLAN, and then perform block S222.

[0047] In block S220, when the VXLAN is not associated with a VLAN, the physical switch allocates a VLAN which is not associated with any other VXLANs for the VXLAN, and associates the VXLAN with the VLAN.

[0048] In block S222, the physical switch establishes a VSI virtual port (also called as a local virtual port) to correspond to the VM or the vNIC of the VM, associates the VSI virtual port with the MAC address, VLAN and VNI, and binds the VSI virtual port with the physical port which receives the VDP packet.

[0049] In block S224, the physical switch obtains the corresponding port strategy from the VTDB of the network management server according to the extracted VTID and version number, and applies the obtained port strategy to the VSI virtual port.

[0050] Applying the obtained port strategy to the VSI virtual port includes, for example, associating the port strategy with the VSI virtual port.

[0051] In addition, on the contrary, before the VM is off line, the server may also perform block S210 to send a VDP packet to the physical switch. After receiving the VDP packet, the physical switch extracts the MAC address and VNI from the VDP packet, searches a corresponding VSI virtual port locally according to the extracted MAC address and VNI, deletes the searched VSI virtual port, and associates the VSI virtual port with its corresponding information.

[0052] FIG. 7 is a flow diagram illustrating a method for obtaining relevant information of a vNIC according to an example of the present disclosure. The method includes the following procedures.

[0053] In block S302, the network management server divides service types of the VM into groups, and each group of service types corresponds to a port strategy; allocates a vNIC type for each port strategy, and each vNIC type has a VTID and a version number; and stores the above-described relationship in a VTDB of the network management server, i.e., stores the corresponding relationship among the service type group, the port strategy and the vNIC type to the VTDB.

[0054] Service types of the VM may include webserver, database, e-mail, etc.

[0055] The port strategy may include some port behavior, e.g., restricting the speed, allowing a data flow to pass the VLAN, etc.

[0056] In block S304, by communication with the network management server, a VM manager (also called as a VM management apparatus) searches an applicable vNIC type in the VTDB, associates the vNIC of the VM with the searched vNIC type, and specifies a tenant where the vNIC is located at the same time, i.e., specifies the VXLAN where the vNIC is located.

[0057] In particular, associating the vNIC of the VM with the searched vNIC type includes, associating the MAC address of the vNIC with the VTID and the version number of the vNIC type; or, associating other information for uniquely identifying the vNIC with the VTID and the version number of the vNIC type.

[0058] In block S306, the VM manager pushes the association information in block S304 to the server, i.e., pushes the association relationship among the vNIC of the VM, the vNIC type (in particular, the VTID and the version number), and the VXLAN where the vNIC is located (in particular, the VNI) to the server.

[0059] In block S308, before the VM is on line, the server sends information corresponding to the vNIC of the VM to the directly connected physical switch. The information includes the MAC address of the vNIC and the VNI of the VXLAN where the vNIC is located. Furthermore, it may also include the VTID and the version number of the vNIC type associated with the vNIC.

[0060] Specifically, the server may carry the above-described information through a VDP packet.

[0061] In block S310, after a VDP packet is received, the physical switch extracts the MAC address, VNI, VTID and the version number, and then searches whether there is a corresponding VXLAN available in the physical switch according to the extracted VNI (i.e., whether the corresponding VXLAN is already established).

[0062] If no, perform block S312; if yes, perform block S314.

[0063] In block S312, the physical switch establishes a corresponding VXLAN in itself, allocates a VLAN which is not associated with any other VXLANs for the VXLAN, and associates the VXLAN with the VLAN.

[0064] In block S314, the physical switch determines whether the searched VXLAN is associated with a VLAN. If yes, perform block S316; if no, perform block S318.

[0065] In block S316, when the VXLAN is already associated with a VLAN, search the VLAN associated with the VXLAN, and then perform block S320.

[0066] In block S318, when the VXLAN is not associated with a VLAN, allocate a VLAN which is not associated with any other VXLANs for the VXLAN, and associate the VXLAN with the VLAN.

[0067] In block S320, the physical switch establishes a VSI virtual port (also called as a local virtual port) to correspond to the VM or the vNIC of the VM, associates the VSI virtual port with the MAC address, VLAN and VNI, and binds the VSI virtual port with the physical port which receives the VDP packet.

[0068] In block S322, the physical switch obtains the corresponding port strategy from the VTDB of the network management server according to the extracted VTID and version number, and applies the obtained port strategy to the VSI virtual port.

[0069] Applying the obtained port strategy to the VSI virtual port includes, for example, associating the port strategy with the VSI virtual port.

[0070] In addition, on the contrary, before the VM is off line, the operations are same with those in the above-described example, which are not described in detail herein.

[0071] Compared with the above-described example, in this example, the physical switch may not establish a VXLAN in advance, and automatically establish based on needs in block S312.

[0072] FIG. 8 is a flow diagram illustrating a method for obtaining relevant information of a vNIC according to another example of the present disclosure. As shown in FIG. 8, the method includes the following procedures.

[0073] In block S402, a physical switch establishes a VXLAN in advance, and specifies a VNI corresponding to the VXLAN.

[0074] In block S404, a network management server divides service types of the VM into groups, and each group of service types corresponds to a port strategy; allocates a vNIC type for each port strategy, and each vNIC type has a VTID and a version number; and stores the above-described relationship in a VTDB of the network management server, i.e., stores the corresponding relationship among the service type group, the port strategy and the vNIC type to the VTDB.

[0075] Service types of the VM may include webserver, database, e-mail, etc.

[0076] The port strategy may include some port behavior, e.g., restricting the speed, allowing a data flow to pass the VLAN, etc.

[0077] In block S406, by communication with the network management server, a VM manager (also called as a VM management apparatus) searches an applicable vNIC type in the VTDB, associates the vNIC of the VM with the searched vNIC type, and specifies a tenant where the vNIC is located at the same time, i.e., specifies the VXLAN where the vNIC is located.

[0078] In particular, associating the vNIC of the VM with the searched vNIC type includes, associating the MAC address of the vNIC with the VTID and the version number of the vNIC type; or, associating other information for uniquely identifying the vNIC with the VTID and the version number of the vNIC type.

[0079] In block S408, the VM manager pushes the association information in block S406 to the server, i.e., pushes the association relationship among the vNIC of the VM, the vNIC type (in particular, the VTID and the version number), and the VXLAN where the vNIC is located (in particular, the VNI) to the server.

[0080] In block S410, before the VM is on line, the server sends information corresponding to the vNIC of the VM to the directly connected physical switch. The information includes the MAC address of the vNIC and the VNI of the VXLAN where the vNIC is located. Furthermore, it may also include the VTID and the version number of the vNIC type associated with the vNIC.

[0081] Specifically, the server may carry the above-described information through a VDP packet.

[0082] In block S412, after a VDP packet is received, the physical switch extracts the MAC address, VNI, VTID and the version number, and then searches whether there is a corresponding VXLAN available in the physical switch according to the extracted VNI (i.e., whether the corresponding VXLAN is already established).

[0083] If no, perform block S414; if yes, perform block S416.

[0084] In block S414, the physical switch returns a VDP failure response to the server, and the procedure ends.

[0085] In block S416, the physical switch determines whether the searched VXLAN is associated with a VLAN. If yes, perform block S418; if no, perform block S422.

[0086] In block S418, the physical switch establishes a VSI virtual port to correspond to the VM or the vNIC of the VM, associates the VSI virtual port with the MAC address and VNI (i.e., the VNI of the VXLAN), and binds the VSI virtual port with the physical port which receives the VDP packet; then performs block S420.

[0087] In block S420, the physical switch obtains the corresponding port strategy from the VTDB of the network management server according to the extracted VTID and version number, applies the obtained port strategy to the VSI virtual port, obtains the corresponding VLAN from the port strategy, associates the obtained VLAN with the virtual port, and associates the obtained VLAN with the VNI.

[0088] For example, the port strategy may include allowing the data flow to pass a VLAN, and such VLAN is the corresponding VLAN in the port strategy.

[0089] In block S422, the physical switch searches the VLAN associated with the VXLAN, establishes a VSI virtual port to correspond to the VM or the vNIC of the VM, associates the VSI virtual port with the MAC address, VLAN and VNI, and binds the VSI virtual port with the physical port which receives the VDP packet; then performs block S424.

[0090] In block S424, the physical switch obtains the corresponding port strategy from the VTDB of the network management server according to the extracted VTID and version number, applies the obtained port strategy to the VSI virtual port.

[0091] Or, in this example, block S402 may not be performed, such that in block S414, the physical switch may establish a corresponding VXLAN (the newly established VXLAN is not associated with any VLANs); then perform block S418.

[0092] In addition, on the contrary, before the VM is off line, the operations are same with those in the above-described examples, which are not described in detail herein.

[0093] Compared with the above-described two examples, in this example, when it is determined that the searched VXLAN is not associated with a VLAN in block S416, it may not allocate a VLAN for the VXLAN and associate them in block S418. But rather, in block S420, obtain the corresponding port strategy and the corresponding VLAN from the port strategy, and associates the VLAN with the virtual port and the VXLAN.

[0094] FIG. 9 is a flow diagram illustrating a method for obtaining relevant information of a vNIC according to yet another example of the present disclosure. As shown in FIG. 9, the method includes the following procedures.

[0095] In block S502, a physical switch establishes a VXLAN in advance, and specifies a VNI corresponding to the VXLAN.

[0096] In block S504, a network management server divides service types of the VM into groups, and each group of service types corresponds to a port strategy; allocates a vNIC type for each port strategy, and each vNIC type has a VTID and a version number; and stores the above-described relationship in a VTDB of the network management server, i.e., stores the corresponding relationship among the service type group, the port strategy and the vNIC type to the VTDB.

[0097] The port strategy may include some port behavior, e.g., restricting the speed, allowing a data flow to pass the VLAN, etc.

[0098] In block S506, by communication with the network management server, a VM manager (also called as a VM management apparatus) searches an applicable vNIC type in the VTDB, and associates the vNIC of the VM with the searched vNIC type.

[0099] In block S508, the VM manager pushes the association information in block S506 to the server, i.e., pushes the association relationship between the vNIC of the VM and the vNIC type to the server.

[00100] In block S510, before the VM is on line, the server sends information corresponding to the VM to the directly connected physical switch. The information includes the MAC address of the vNIC. Furthermore, it may also include the VTID and the version number of the vNIC type associated with the VM.

[00101] In block S512, after a VDP packet is received, the physical switch extracts the MAC address, VNI, VTID and the version number, obtains the corresponding port strategy from the VTDB of the network management server according to the extracted VTID and version number, and obtains the corresponding VNI from the port strategy.

[00102] In block S514, according to the extracted VNI obtained in block S512, the physical switch searches whether there is a corresponding VXLAN available in the physical switch (i.e., whether the corresponding VXLAN is already established).

[00103] If no, perform block S516; if yes, perform block S518.

[00104] In block S516, the physical switch returns a VDP failure response to the server, and the procedure ends.

[00105] In block S518, the physical switch determines whether the searched VXLAN is associated with a VLAN. If yes, perform block S522; if no, perform block S520.

[00106] In block S520, the physical switch allocates a VLAN which is not associated with any other VXLANs for the VXLAN, and associate the VXLAN with the VLAN; then perform block S524.

[00107] In block S522, the physical switch searches the VLAN associated with the VXLAN, and then perform block S524.

[00108] In block S524, the physical switch establishes a VSI virtual port to correspond to the VM or the vNIC of the VM, associates the VSI virtual port with the MAC address, VLAN and VNI, applies the obtained port strategy in block S512 to the VSI virtual port, and binds the VSI virtual port with the physical port which receives the VDP packet.

[00109] Or, in this example, block S502 may not be performed, such that in block S516, the physical switch may establish a corresponding VXLAN, allocate a VLAN which is not associated with any other VXLANs for the VXLAN, and associate the VXLAN with the VLAN; then perform block S524.

[00110] In addition, on the contrary, before the VM is off line, the server may also perform block S510 to send a VDP packet to the physical switch, and the physical switch may

perform block S512 to obtain the corresponding MAC address and VNI, search a corresponding VSI virtual port locally, deletes the searched VSI virtual port, and associates the VSI virtual port with its corresponding information.

[00111] Compared with the above-described examples, in this example, the VM manager may not specify a VNI for the vNIC of the VM. Then, in block S512, after a VDP packet is received, the physical switch obtains the corresponding port strategy from the VTDB of the network management server according to the extracted VTID and version number, and obtains the corresponding VNI from the port strategy.

[00112] In implementation, any two examples of the above-described four examples can be combined for obtaining relevant information of a vNIC.

[00113] Example 2

[00114] FIG. 10 is a flow diagram illustrating a method for forwarding a packet in a data center network according to a second example of the present disclosure. Besides the block S102, S104 and S106 as described in the first example, the method also includes S108.

[00115] Specifically, at block S108, for the VXLAN packet received from the IP network, the physical switch may also perform the following operations. After the VXLAN packet is received, the physical switch de-encapsulates the VXLAN packet, obtains a VXLAN header and an Ethernet packet, searches the corresponding local virtual port according to a destination MAC address in the Ethernet packet and a VNI in the VXLAN header obtained by de-encapsulation, and forwards the Ethernet packet obtained by de-encapsulation to the server through a local physical port binded with the searched local virtual port.

[00116] After the Ethernet packet is received, the vSwitch in the server may search matched flow table entries locally, and forward the data packet obtained by de-encapsulating the Ethernet packet to a destination VM according to the table entries.

[00117] According to the above-described examples in the present disclosure, the physical switch obtains relevant information of a vNIC of each VM in the server, and associates the obtained relevant information with a local virtual port. Then, after receiving an Ethernet packet sent by the server, according to a source MAC address in the Ethernet packet and

VLAN, the physical switch identifies the corresponding local virtual port and searches a corresponding VNI. At last, according to the searched VNI, the physical switch encapsulates the Ethernet packet as a VXLAN packet and forwards the VXLAN packet by mapping the VLAN data flow into the VXLAN. In the present disclosure, an overlay technology, e.g., VXLAN, is initiated in the physical switch, and the physical switch works as a VTEP. Because the physical switch uses hardware for forwarding, the forwarding efficiency is higher than the vSwitch implemented by software. Further, the VXLAN data flow is not only visible in the physical switch, but also operated between the physical switch and the server. In the service plane, only the VXLAN is visible for the user. By dividing a tenant according to the VXLAN, the VLAN is invisible for the user. VLAN works as a local service dividing tool, and a VLAN tag is automatically allocated by the system, so the VLAN is invisible for the final user.

[00118] According to the method described in the first example, the present disclosure provides an apparatus for applying the method for forwarding a packet. The apparatus for forwarding a packet is applied in the physical switch. As shown in FIG. 11, the apparatus for forwarding a packet includes a processing module 10, a receiving module 20, a searching module 30, an encapsulation and de- encapsulation module 40 and a forwarding module 50.

[00119] The processing module 10 is configured to, for each VM in the server of the physical switch, obtain relevant information of a vNIC of the VM, and associate the obtained relevant information with a local virtual port. The relevant information includes a MAC address of the vNIC, a VNI of the VXLAN where the vNIC is located, and a VLAN ID of a VLAN associated with the VXLAN where the vNIC is located.

[00120] The receiving module 20 is configured to receive an Ethernet packet sent by the server.

[00121] The searching module 30 is configured to, after an Ethernet packet is received from the server in the receiving module 20, search a corresponding VNI according to a source MAC address in the Ethernet packet and a VLAN ID.

[00122] The encapsulation and de-encapsulation module 40 is configured to, according to the VNI searched in the searching module 30, encapsulate the Ethernet packet received in the receiving module 20 and obtain a VXLAN packet.

[00123] The forwarding module 50 is configured to, forward the VXLAN packet encapsulated in the encapsulation and de- encapsulation module 40.

[00124] In addition, the apparatus may also include a storage module, configured to store the local virtual port and relevant information of the associated vNIC.

[00125] The receiving module is further configured to receive a VXLAN packet.

[00126] The encapsulation and de-encapsulation module is further configured to, after the VXLAN packet is received in the receiving module, de-encapsulate the VXLAN packet, and obtain a VXLAN header and an Ethernet packet.

[00127] The searching module is further configured to search the corresponding local virtual port according to a destination MAC address in the Ethernet packet and a VNI in the VXLAN header de-encapsulated in the encapsulation and de-encapsulation module.

[00128] The forwarding module is further configured to forward the Ethernet packet de-encapsulated in the encapsulation and de-encapsulation module to the server through a local physical port binded with the local virtual port searched in the searching module.

[00129] The processing module includes a MAC receiving unit, a first obtaining unit and a second obtaining unit.

[00130] The MAC receiving unit is configured to receive a MAC address of the vNIC sent by the server.

[00131] The first obtaining unit is configured to obtain a VNI of the VXLAN where the vNIC is located.

[00132] The second obtaining unit is configured to obtain a VLAN ID of a VLAN associated with the VXLAN where the vNIC is located.

Specifically, the first obtaining unit is configured to receive the VNI of the VXLAN where the vNIC is located sent by the server, where the VNI is sent to the server after the VM manager specifies the VXLAN for the vNIC; or

after receiving the vNIC type associated with the vNIC sent by the server, obtain the corresponding port strategy from the network management server according to the vNIC type, and obtain a VNI from the port strategy as the VNI of the VXLAN where the vNIC is located, where the vNIC type associated with the vNIC is sent to the server after the VM manager specifies the associated vNIC type for the vNIC.

Specifically, the second obtaining unit is configured to, after the VNI of the VXLAN where the vNIC is located is obtained in the first obtaining unit, search whether there is a corresponding VXLAN available in the physical switch according to the obtained VNI; if yes, when the searched VXLAN is not associated with a VLAN, allocate a VLAN which is not associated with any other VXLANs for the searched VXLAN, and associate the searched VXLAN with the allocated VLAN; when the searched VXLAN is already associated with a VLAN, search the VLAN associated with the VXLAN; or

after receiving the vNIC type associated with the vNIC sent by the server, obtain the corresponding port strategy from the network management server according to the vNIC type, and obtain a VNI from the port strategy as the VNI of the VXLAN where the vNIC is located.

[00133] The apparatus may further includes an establishing module, configured to establish a VXLAN in the physical switch in advance, and allocate a VNI for the VXLAN. Thus, the second obtaining unit is further configured to, after searching whether there is a corresponding VXLAN available in the physical switch according to the obtained VNI, if it is no, end the procedure; or

after searching whether there is a corresponding VXLAN available in the physical switch according to the obtained VNI, if it is no, establish a VXLAN identified by the obtained VNI in the physical switch, allocate a VLAN which is not associated with any other VXLANs for the VXLAN, and associate the VXLAN and the VLAN.

[00134] Example 3

[00135] The third example of the present disclosure provides a hardware structure of a device for forwarding a packet.

[00136] The device for forwarding a packet may be a programmable device with a combined software and hardware. FIG. 12 is a schematic diagram illustrating a hardware structure of a device for forwarding a packet in a data center network according to a third example of the present disclosure. As shown in FIG. 12, the device for forwarding a packet includes a processor, e.g., a Central Process Unit (CPU) 1201, a forwarding chip 1202, a storage, including a non-transitory storage 1203 and a memory 1204, and other hardware 1205.

[00137] The memory 1204 is configured to store instructions. The operations performed when the instructions are executed include the functions performed by the processing module, the receiving module, the searching module, the encapsulation and de-encapsulation module, the forwarding module and the establishing module in the apparatus for forwarding a packet, which are not described in detail herein.

[00138] The CPU 1201 is configured to communicate with the forwarding chip 1202, forward and receive packets, including: receiving an Ethernet packet sent by the server of the physical switch, forwarding the VXLAN packet obtained by encapsulation, receiving a VXLAN packet, and forwarding the Ethernet packet obtained by de-encapsulation to the server of the physical switch; configured to communicate with the memory 1204, read and execute the instructions stored in the memory 1204, to perform functions of the processing module, the receiving module, the searching module, the encapsulation and de-encapsulation module, the forwarding module and the establishing module in the above-described apparatus for forwarding a packet, and process the packet forwarded by the forwarding chip 1202; and configured to communicate with the non-transitory storage 1203, read/write the data in the non-transitory storage 1203, including the local virtual port and relevant information of the associated vNIC.

[00139] The processing module is configured to, for each VM in the server of the physical switch, obtain relevant information of a vNIC of the VM, and associate the obtained relevant information with a local virtual port, and store them in the non-transitory storage 1203. The relevant information includes a MAC address of the vNIC, a VNI of the VXLAN where the

vNIC is located, and a VLAN ID of a VLAN associated with the VXLAN where the vNIC is located. The receiving module is configured to receive an Ethernet packet sent by the server. The searching module is configured to, after an Ethernet packet is received from the server in the receiving module, search a corresponding VNI from the non-transitory storage 1203 according to a source MAC address in the Ethernet packet and a VLAN ID. The encapsulation and de-encapsulation module is configured to, according to the VNI searched in the searching module, encapsulate the Ethernet packet received in the receiving module and obtain a VXLAN packet. The forwarding module is configured to forward the encapsulated VXLAN packet.

[00140] The receiving module is further configured to receive a VXLAN packet. The encapsulation and de-encapsulation module is further configured to, after the VXLAN packet is received in the receiving module, de-encapsulate the VXLAN packet, and obtain a VXLAN header and an Ethernet packet. The searching module is further configured to search the corresponding local virtual port from the non-transitory storage 1203 according to a destination MAC address in the Ethernet packet and a VNI in the VXLAN header de-encapsulated in the encapsulation and de-encapsulation module. The forwarding module is further configured to forward the Ethernet packet de-encapsulated in the encapsulation and de-encapsulation module to the server through a local physical port binded with the local virtual port searched in the searching module.

[00141] The forwarding chip 1202 is configured to connect with the physical switch, the server and the IP network through ports in the chip, and be responsible for forwarding and receiving the above-described packets. In particular, the forwarding chip forwards the Ethernet packet from the server to the CPU 1201, forwards the VXLAN packet from the CPU 1201 to the IP network, forwards the VXLAN packet from the IP network to the CPU 1201, and forwards the Ethernet packet from the CPU 1201 to the server.

[00142] The non-transitory storage 1203 is configured to store data, including the local virtual port and relevant information of the associated vNIC, and perform the functions of the storage module in the device for forwarding the packet. The storage module is configured to store the local virtual port and relevant information of the associated vNIC.

[00143] It should be noted that, the device for forwarding a packet as shown in FIG. 12 is just an example for illustration. It is possible to implement using other structures different from the example of the present disclosure. For example, operations performed when executing the above-described instructions may also implemented by an Application Specific Integrated Circuit (ASIC). In addition, the above-described CPU may be one or more CPUs. If there are more CPUs, the above-described instructions are read and executed by the more CPUs jointly. Hence, the present disclosure does not limit the specific structure of the apparatus for forwarding a packet.

[00144] In a word, the examples of the present disclosure can achieve the following technical effects.

[00145] The physical switch obtains relevant information of a vNIC of each VM in the server, and associates the obtained relevant information with a local virtual port. Then, after receiving an Ethernet packet sent by the server, according to a source MAC address in the Ethernet packet and VLAN, the physical switch identifies the corresponding local virtual port and searches a corresponding VNI. At last, according to the searched VNI, the physical switch encapsulates the Ethernet packet as a VXLAN packet and forwards the VXLAN packet by mapping the VLAN data flow into the VXLAN. In the present disclosure, an overlay technology, e.g., VXLAN, is initiated in the physical switch, and the physical switch works as a VTEP. Because the physical switch uses hardware for forwarding, the forwarding efficiency is higher than the vSwitch implemented by software. Further, the VXLAN data flow is not only visible in the physical switch, but also operated between the physical switch and the server. In the service plane, only the VXLAN is visible for the user. By dividing a tenant according to the VXLAN, the VLAN is invisible for the user. VLAN works as a local service dividing tool, and a VLAN tag is automatically allocated by the system, so the VLAN is invisible for the final user.

[00146] The foregoing is preferred examples of the present disclosure, which is not intended to limit the present disclosure. Any modifications, equivalents, and improvements made within the spirit and principle of the present disclosure should be covered by the scope of the present disclosure.

CLAIMS

WHAT IS CLAIMED IS:

1. A method for forwarding a packet in a data center network, the data center network comprises: at least one server, a virtual Switch (vSwitch) and a Virtual Machine (VM) configured in the server, a physical switch outside of the server and connected with the VM via the vSwitch, the method applied in the physical switch, comprising:

for each VM in the server, obtaining relevant information of a virtual Network Interface Controller (vNIC) of the VM, and associating the obtained relevant information with a local virtual port;

after receiving an Ethernet packet sent by the server, searching a Virtual eXtensible Local Area Network (VXLAN) Network Identifier (VNI) according to a source Media Access Control (MAC) address of the Ethernet packet and a Virtual Local Area Network (VLAN) Identifier (ID); and

according to the searched VNI, encapsulating the Ethernet packet, obtaining a VXLAN packet, and forwarding the VXLAN packet.

2. The method according to claim 1, further comprising:

After receiving an VXLAN packet, de-encapsulating the VXLAN packet, obtaining a VXLAN header and an Ethernet packet, and searching a local virtual port according to a destination MAC address in the Ethernet packet and a VNI in the VXLAN header obtained by de-encapsulation; and

forwarding the Ethernet packet obtained by de-encapsulation to the server through a local physical port binded with the searched local virtual port.

3. The method according to claim 1, wherein obtaining relevant information of the vNIC of the VM comprising:

receiving a MAC address of the vNIC sent by the server; and

obtaining a VNI of a VXLAN where the vNIC is located, and a VLAN ID of a VLAN associated with a VXLAN where the vNIC is located.

4. The method according to claim 3, wherein obtaining a VNI of a VXLAN where the vNIC is located comprising:

receiving a VNI of a VXLAN where the vNIC is located sent by the server, wherein the VNI is sent to the server after a VM manager specifies the VXLAN where the vNIC is located for the vNIC;

or, after receiving a vNIC type associated with the vNIC sent by the server, obtaining a port strategy from a network management server according to the vNIC type, and obtaining a VNI from the port strategy as the VNI of the VXLAN where the vNIC is located, wherein the vNIC type associated with the vNIC is sent to the server after the VM manager specifies the associated vNIC type for the vNIC.

5. The method according to claim 4, wherein obtaining a VLAN ID of a VLAN associated with a VXLAN where the vNIC is located comprising:

after a VNI of a VXLAN where the vNIC is located is obtained, searching whether there is a VXLAN available in the physical switch according to the obtained VNI; if yes, when the searched VXLAN is not associated with a VLAN, allocating a VLAN which is not associated with any other VXLANs for the searched VXLAN, and associating the searched VXLAN with the allocated VLAN; when the searched VXLAN is already associated with a VLAN, searching the VLAN associated with the searched VXLAN;

or, after receiving a vNIC type associated with the vNIC sent by the server, obtaining a port strategy from the network management server according to the vNIC type, and obtaining a VLAN ID from the port strategy as a VLAN ID of a VLAN associated with a VXLAN where the vNIC is located.

6. The method according to claim 5, further comprising:

establishing a VXLAN in advance, and allocating a VNI for the VXLAN;

after searching whether there is a VXLAN available according to the obtained VNI, if it is no, ending the procedure.

7. The method according to claim 5, further comprising:

after searching whether there is a VXLAN available according to the obtained VNI, if it is no, establishing a VXLAN identified by the obtained VNI, allocating a VLAN which is not associated with any other VXLANs for the VXLAN, and associating the VXLAN and the VLAN.

8. The method according to claim 4, wherein a VNI of a VXLAN where the vNIC is located sent by the server is carried in a Group ID field of a Filter Info domain in a Virtual Station Interface (VSI) Discovery and configuration Protocol (VDP) packet.

9. An apparatus for forwarding a packet in a data center network, the data center network comprises: at least one server, a virtual Switch (vSwitch) and a Virtual Machine (VM) configured in the server, a physical switch outside of the server and connected with the VM via the vSwitch, the apparatus applied in the physical switch, comprising:

a processing module, to, for each VM in the server of the physical switch, obtain relevant information of a virtual Network Interface Controller (vNIC) of the VM, and associate the obtained relevant information with a local virtual port;

a receiving module, to receive an Ethernet packet sent by the server;

a searching module, to, after receiving an Ethernet packet from the server in the receiving module, search a Virtual eXtensible Local Area Network (VXLAN) Network Identifier (VNI) according to a source Media Access Control (MAC) address in the Ethernet packet and a Virtual Local Area Network (VLAN) Identifier (ID);

an encapsulation and de-encapsulation module, to, according to the VNI searched in the searching module, encapsulate the Ethernet packet received in the receiving module and obtain a VXLAN packet; and

a forwarding module, to forward the VXLAN packet encapsulated in the encapsulation and de- encapsulation module.

10. The apparatus according to claim 9, wherein

the receiving module, is further to receive a VXLAN packet;

the encapsulation and de-encapsulation module, is further to, after receiving a VXLAN packet in the receiving module, de-encapsulate the VXLAN packet, and obtain a VXLAN header and an Ethernet packet;

the searching module, is further to search a local virtual port according to a destination MAC address in the Ethernet packet and a VNI in the VXLAN header de-encapsulated in the encapsulation and de-encapsulation module; and

the forwarding module, is further to forward the Ethernet packet de-encapsulated in the encapsulation and de-encapsulation module to the server through a local physical port binded with the local virtual port searched in the searching module.

11. The apparatus according to claim 9, wherein the processing module comprises:

a MAC receiving unit, to receive a MAC address of the vNIC sent by the server;

a first obtaining unit, to obtain a VNI of the VXLAN where the vNIC is located; and

a second obtaining unit, to obtain a VLAN ID of a VLAN associated with the VXLAN where the vNIC is located.

12. The method according to claim 11, wherein the first obtaining unit is further to:

receive a VNI of a VXLAN where the vNIC is located sent by the server, wherein the VNI is sent to the server after a VM manager specifies the VXLAN where the vNIC is located for the vNIC;

or, after receiving a vNIC type associated with the vNIC sent by the server, obtain a port strategy from a network management server according to the vNIC type, and obtain a VNI from the port strategy as the VNI of the VXLAN where the vNIC is located, wherein the vNIC type associated with the vNIC is sent to the server after the VM manager specifies the associated vNIC type for the vNIC.

13. The apparatus according to claim 12, wherein the second obtaining unit is further to:

after a VNI of a VXLAN where the vNIC is located is obtained in the first obtaining unit, search whether there is a VXLAN available in the physical switch according to the obtained VNI; if yes, when the searched VXLAN is not associated with a VLAN, allocate a VLAN which is not associated with any other VXLANs for the searched VXLAN, and associate the searched VXLAN with the allocated VLAN; when the searched VXLAN is already associated with a VLAN, search the VLAN associated with the searched VXLAN;

or, after receiving a vNIC type associated with the vNIC sent by the server, obtain a port strategy from the network management server according to the vNIC type, and obtain a VLAN ID from the port strategy as a VLAN ID of a VLAN associated with a VXLAN where the vNIC is located.

14. The apparatus according to claim 13, the apparatus further comprises:

an establishing module, to establish a VXLAN in advance, and allocate a VNI for the VXLAN;

wherein the second obtaining unit is further to, after searching whether there is a VXLAN available according to the obtained VNI, if it is no, end the procedure.

15. The apparatus according to claim 13, wherein the second obtaining unit is further to:

after searching whether there is a VXLAN available according to the obtained VNI, if it is no, establish a VXLAN identified by the obtained VNI, allocate a VLAN which is not

associated with any other VXLANs for the VXLAN, and associate the VXLAN and the VLAN.

1/10

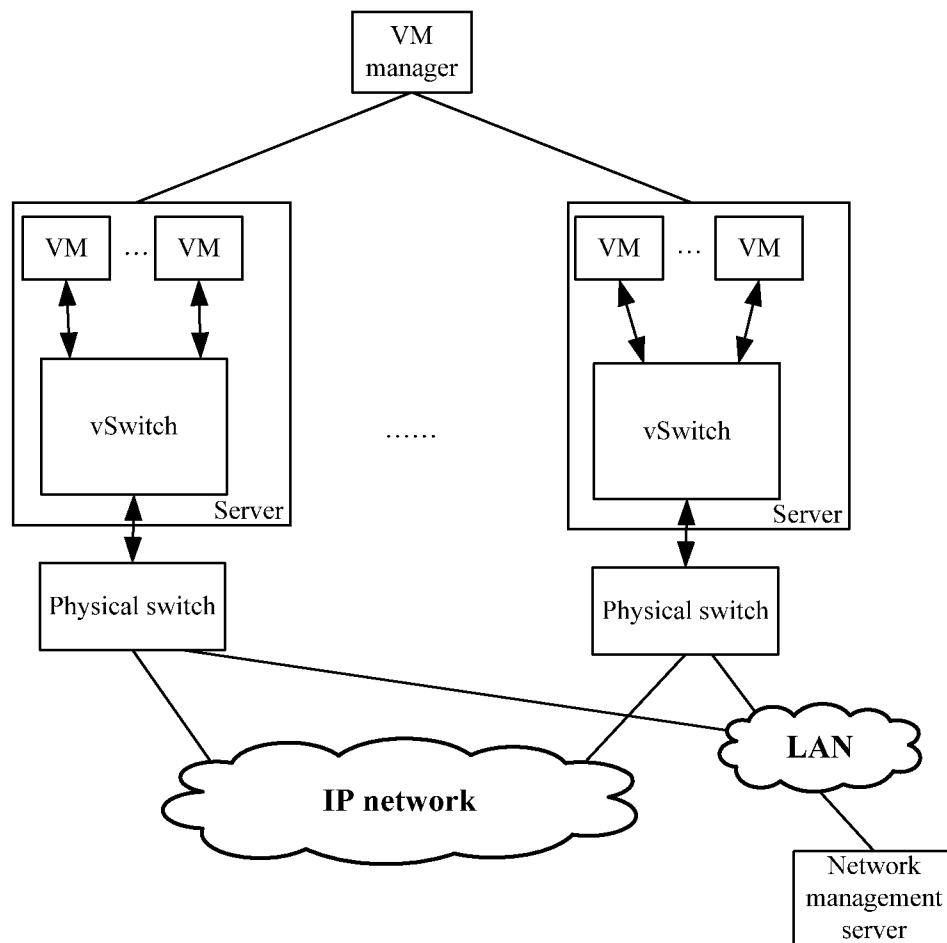


FIG. 1

2/10

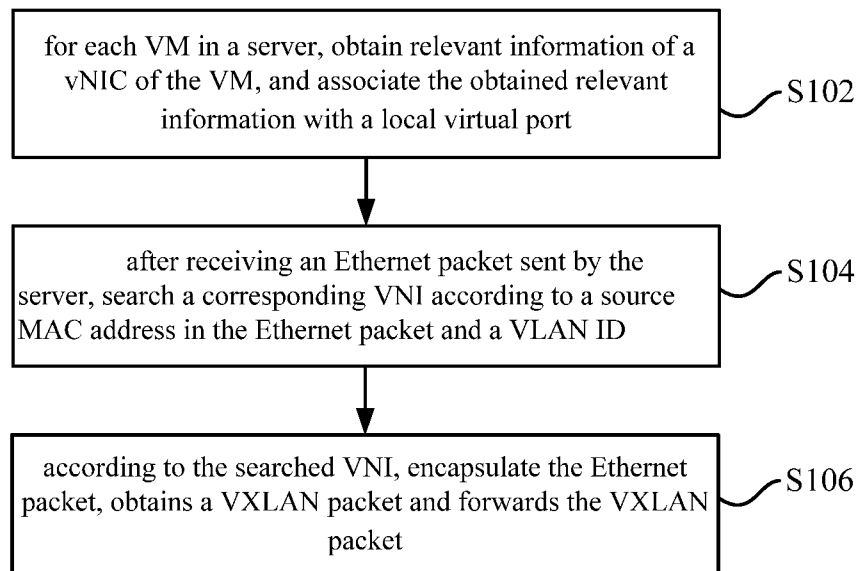


FIG. 2

3/10

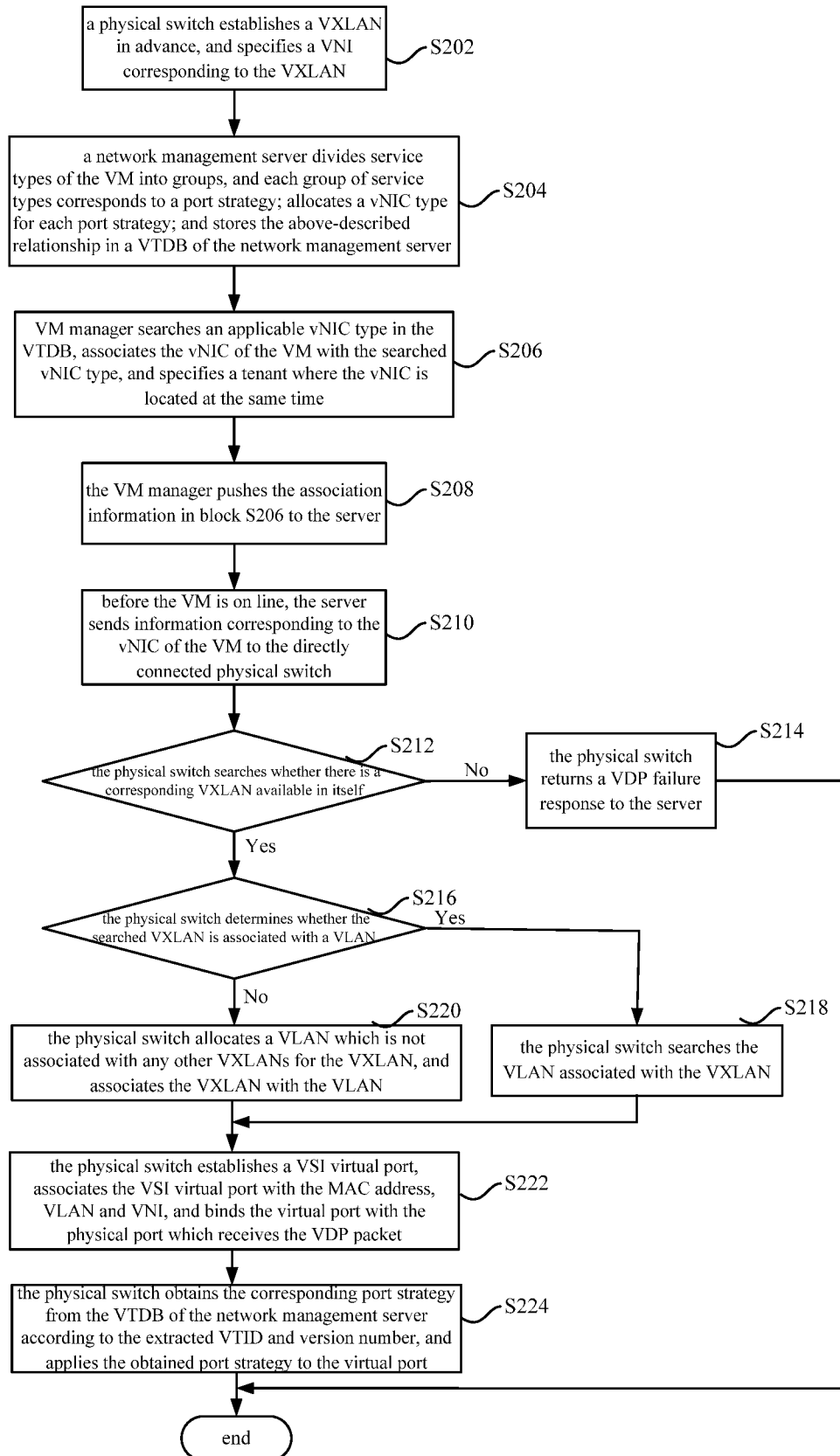


FIG. 3

Type (7 bits)	Length (9 bits)	Reason (1 byte)	VSI Type ID (3 bytes)	VSI Type Version (1 byte)	VSI ID Format (1 byte)	VSI ID (16 bytes)	Filter Format (1 byte)	Filter Info (n bytes)
-----------------	-------------------	-------------------	-------------------------	-----------------------------	--------------------------	---------------------	--------------------------	-------------------------

FIG. 4

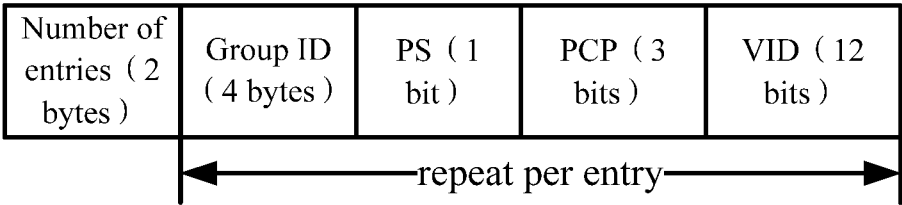


FIG. 5

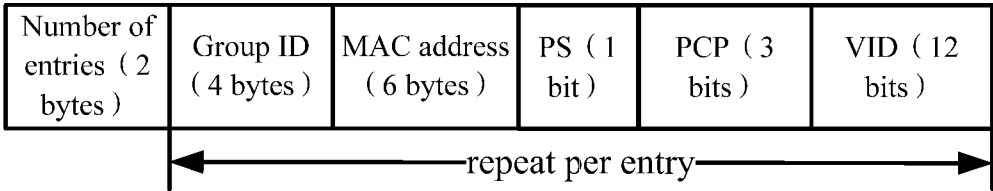


FIG. 6

5/10

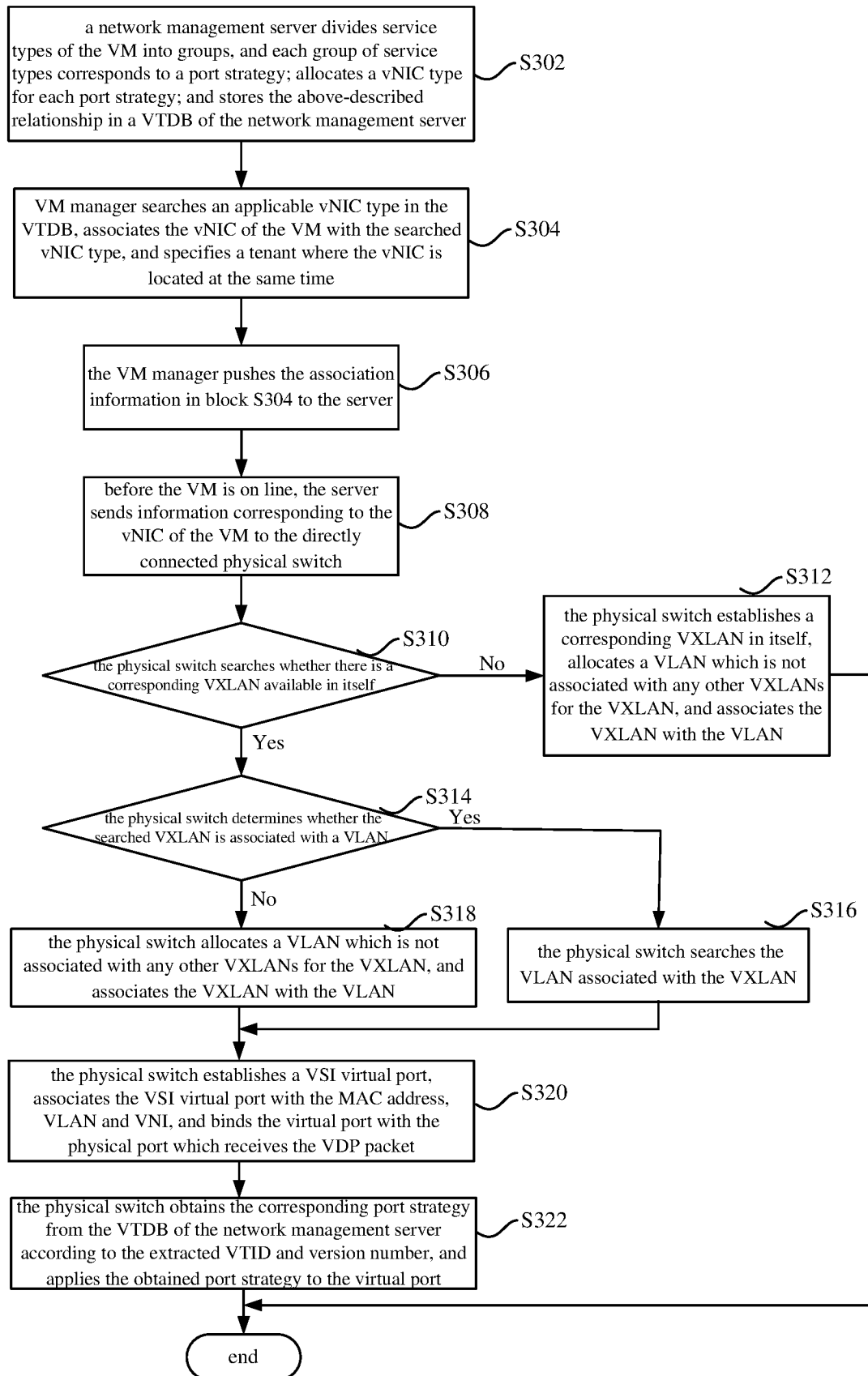


FIG. 7

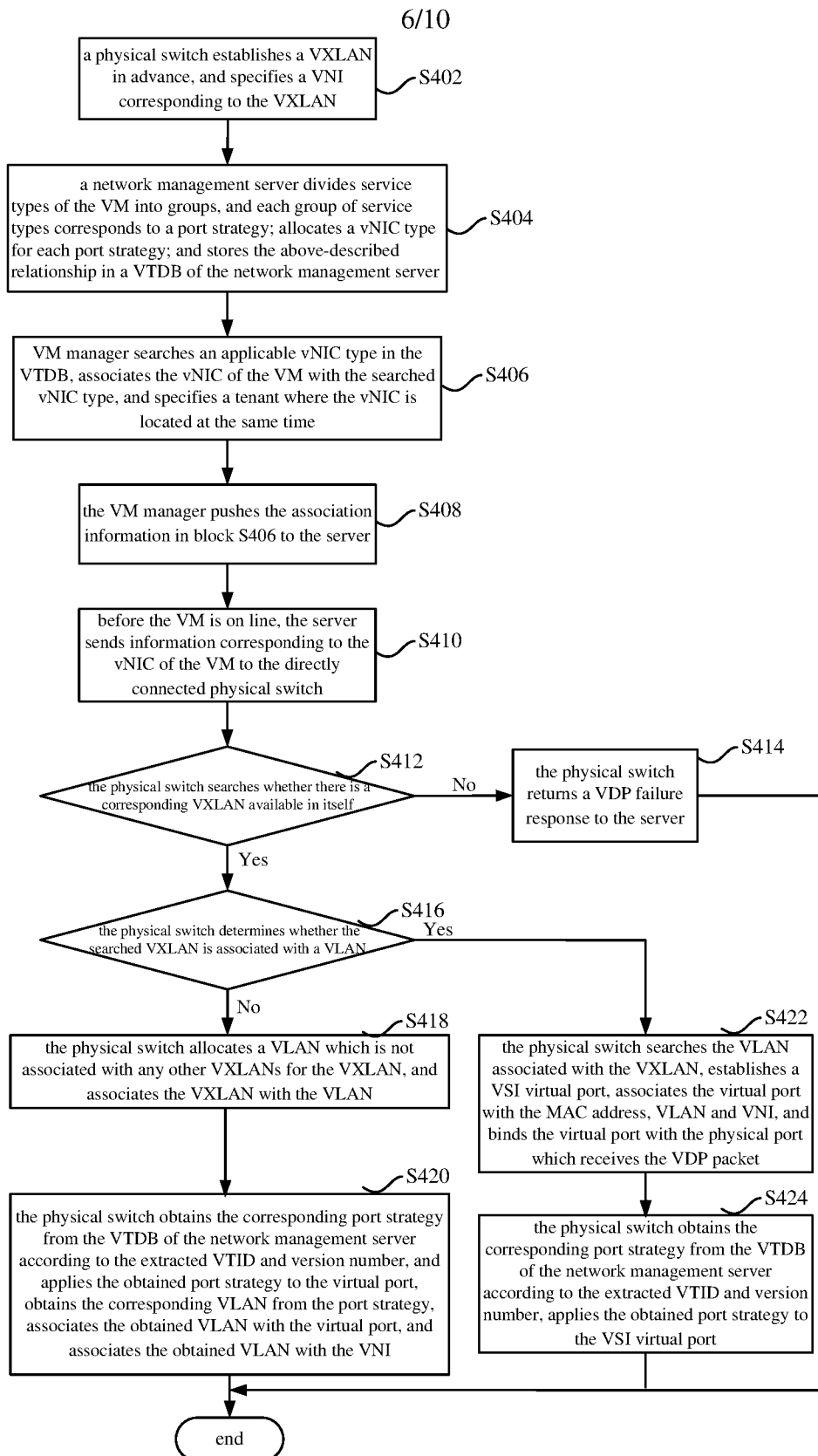


FIG. 8

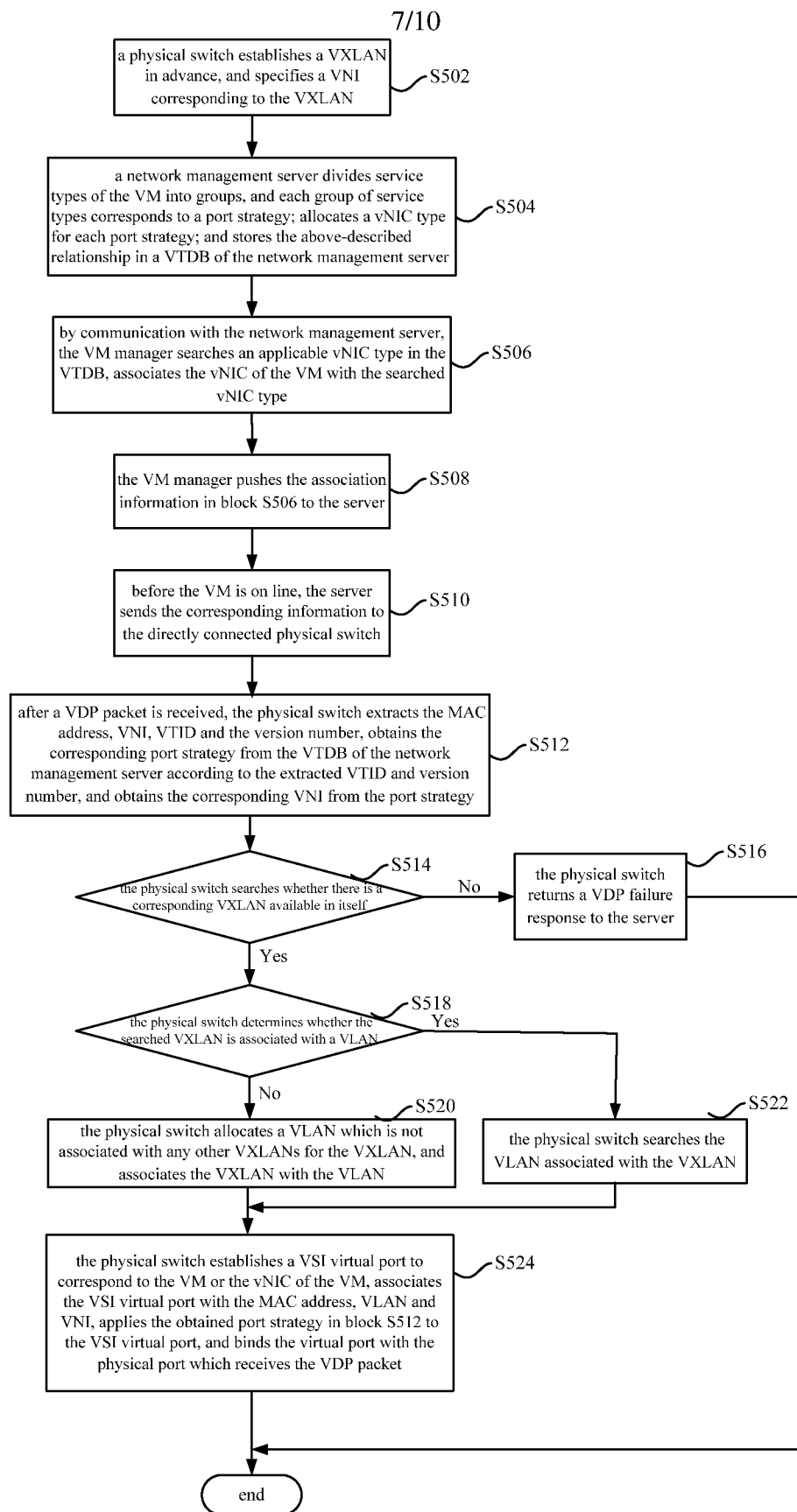


FIG. 9

8/10

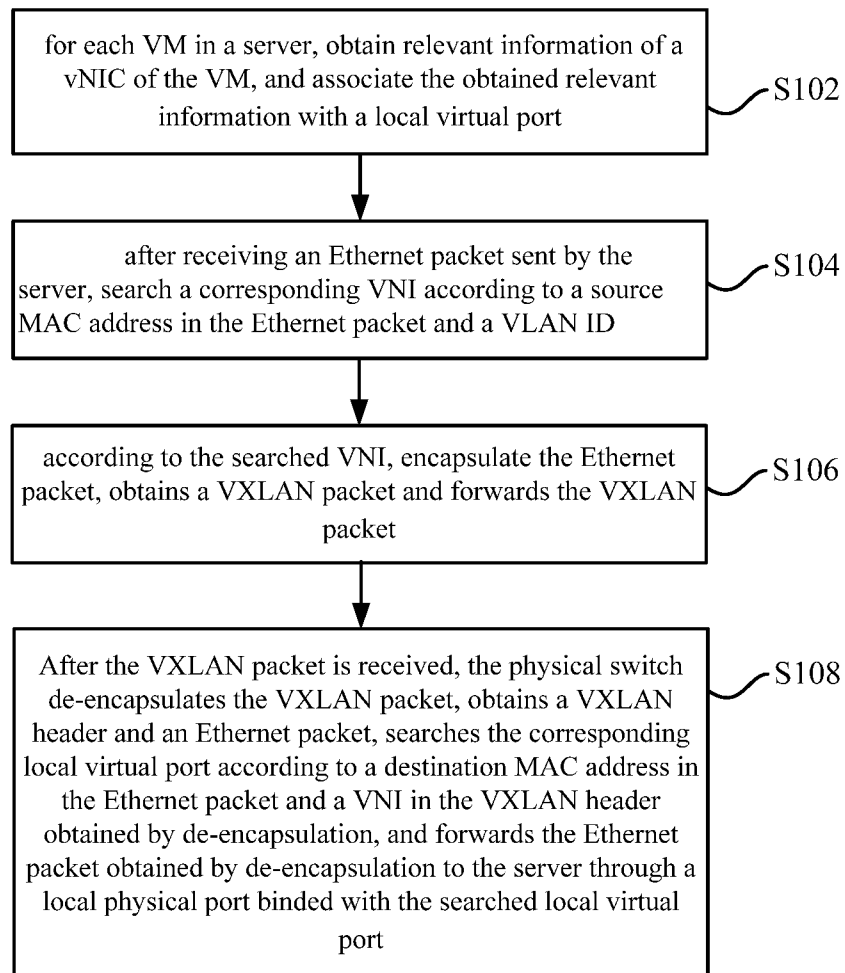


FIG. 10

9/10

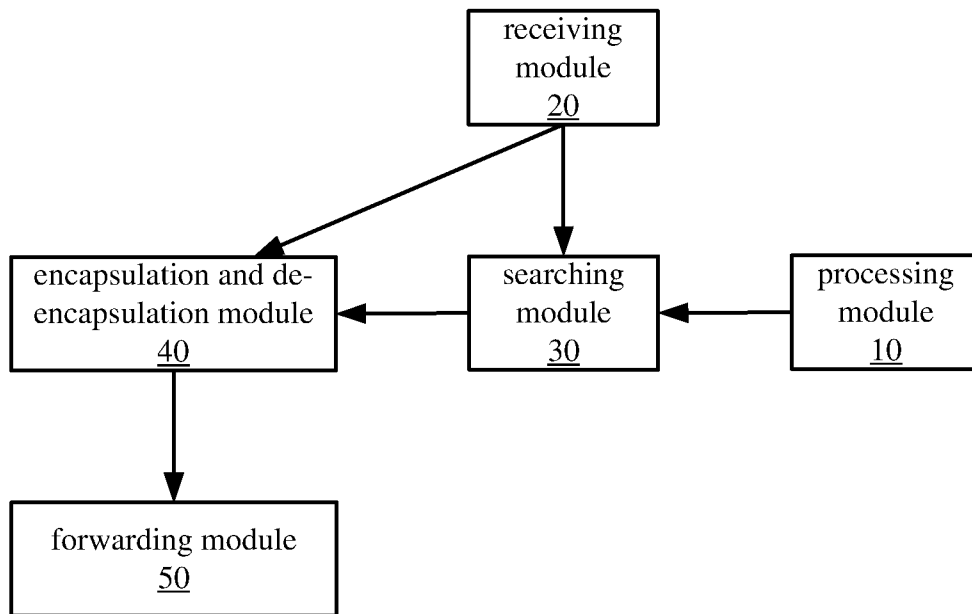


FIG. 11

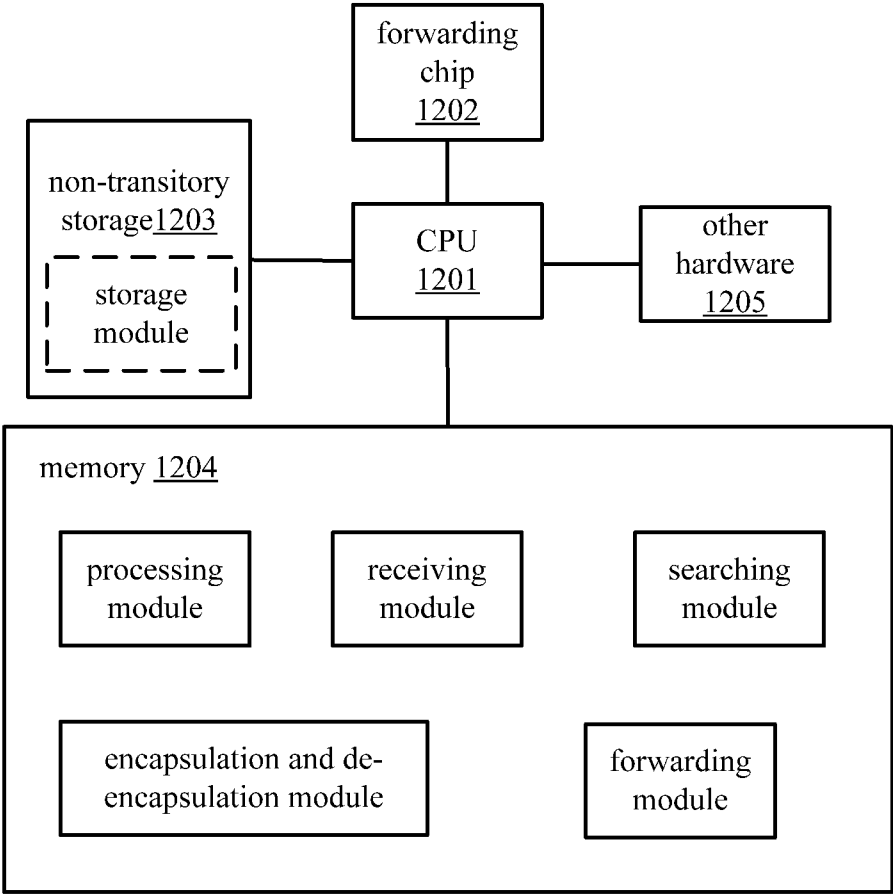


FIG. 12

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2014/089959

A. CLASSIFICATION OF SUBJECT MATTER

H04L 12/46(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)

CNTXT, CNABS, VEN, IEEE: VLAN, VXLAN, virtual, extensible, local area network, physical, switch, machine, vswitch, pswitch, VM, MAC, media access control, VNI, virtual network identifier, vNIC, network interface controller, port, message, encapsulate

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	H3C. "CLOUD COMPUTING OVERLAY NETWORK BASED ON MULTIPLE TENANTS" http://cloud.chinabyte.com/news/303/12745303.shtml , 15 October 2013 (2013-10-15), page 6, line 1 to page 9, line 15	1-15
Y	"VXLAN GATEWAY COMBINING VIRTUAL AND PHYSICAL NETWORK FUNCTIONS" http://www.sdxidx.com/Newsdetail.asp?id=760VXLAN , 27 November 2012 (2012-11-27), page 1, line 16 to page 2, line 2	1-15
A	CN 103200069 A (HUAWEI TECHNOLOGIES CO LTD) 10 July 2013 (2013-07-10) 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

10 January 2015

Date of mailing of the international search report

06 February 2015

Name and mailing address of the ISA/CN

STATE INTELLECTUAL PROPERTY OFFICE OF THE
P.R.CHINA(ISA/CN)
6,Xitucheng Rd., Jimen Bridge, Haidian District, Beijing
100088 China

Authorized officer

WANG,Wei

Facsimile No. (86-10)62019451

Telephone No. (86-10)62089398

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2014/089959

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 103200069 A	10 July 2013	None	