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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
5 February 2015 (05.02.2015)

WIPO | PCT

(10) International Publication Number
WO 2015/015505 A1

(51) International Patent Classification:
H04L 12/00 (2006.01)

(21) International Application Number:
PCT/IN2013/000477

(22) International Filing Date:
1 August 2013 (01.08.2013)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant: **HEWLETT-PACKARD DEVELOPMENT COMPANY, L.P.** [US/US]; 11445 Compaq Center Drive West, Houston, Texas 77070 (US).

(72) Inventor; and
(71) Applicant *(for US only)*: **DEVARAJAN, Venkatavaradhan** [IN/IN]; Sy. No.192, Whitefield Road, Mahadevapura Post, Bangalore, Karnataka 560048 (IN).

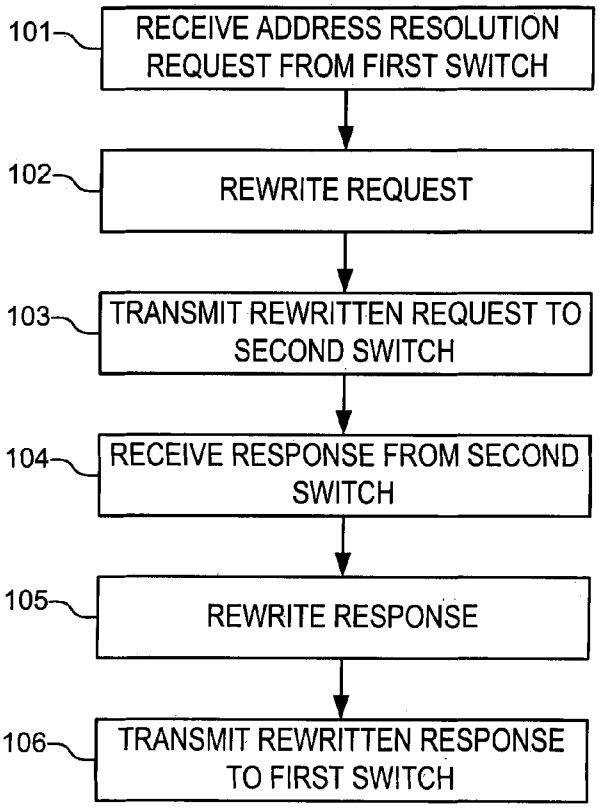
(74) Agent: **ALANKI, N. V. Pradeep Kumar**; Global IP Services, 198F, 27th Cross, 3rd Block, Jayanagar, Bangalore, Karnataka 560011 (IN).

(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, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, 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, 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, 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,

[Continued on next page]

(54) Title: ADDRESS RESOLUTION REWRITING



(57) Abstract: A network controller may rewrite an address resolution request and an address resolution response. The address resolution request may be rewritten by changing a request source link address to a link address of a first switch. The address resolution response may be rewritten by changing a response source link address to a link address of a second switch.

WO 2015/015505 A1

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).

Declarations under Rule 4.17:

— *as to the identity of the inventor (Rule 4.17(i))*

— *of inventorship (Rule 4.17(iv))*

Published:

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

ADDRESS RESOLUTION REWRITING

BACKGROUND

[1] In networks using software defined networking (SDN), one or more network controllers may manage the control plane of network switches, such as bridges and routers. The network controller may also manage the data plane of the switches by providing flow handling rules to the switches.

[2] In some SDN networks, such as networks implementing an OPENFLOW standard (an "OPENFLOW network"), when a packet is received by a controlled switch, the switch determines if it has a rule applicable to the packet. Such a rule may be defined to apply if a matching criterion is satisfied. For example, the matching criterion may be based on source or destination IP or MAC address, port numbers, transport protocol type, frame type, class of service indicators, or frame control information. If the matching criterion is met, the rule may define various actions for the switch to take. For example, the rule may direct the switch to forward the packet out of a particular port. If the switch does not have a rule applicable to the packet, the switch may transmit all or a portion of the packet to a network controller to allow the network controller to provide a rule applicable to the packet.

BRIEF DESCRIPTION OF THE DRAWINGS

[3] Certain examples are described in the following detailed description and in reference to the drawings, in which:

[4] Figure 1 illustrates an example process of network control including rewriting address resolution packets;

[5] Figure 2 illustrates an example process of network control including managing an endpoint database;

[6] Figure 3 illustrates an example process of network control including establishing a network path between edge switches;

[7] Figure 4 illustrates an example network controller including an address resolution rewriter and rule manager; and

[8] Figure 5 illustrates an example network switch including a non-transitory computer readable medium storing instructions to cause the switch to forward address resolution packets and rewrite packets received from or destined for a host.

DETAILED DESCRIPTION OF SPECIFIC EXAMPLES

[9] In an SDN network, a network controller may define flow rules to implement network paths on a per endpoint basis. For example, to implement a flow on a single subnet of a virtual local area network (VLAN) between a first host, *H1*, and a second host, *H2*, a network controller may provide each switch in the flow path with a rule of the type illustrated in Table 1:

TABLE 1

Match Condition	Action
Destination Address == MAC_H2	Output port = {i}
Other Fields == wildcards	

In Table 1, MAC_H2 refers to the Media Access Control (MAC) address of the second host, *H2*, and *i* is specific to each switch to implement the path. Accordingly, the rule causes a switch to output a packet to output port *i* if a packet is received with *H2*'s MAC address as the destination address, where the value of *i* depends on the switch.

[10] Programing flow rules on a per endpoint basis may consume significant forwarding table resources and render it difficult to scale up the number of active endpoints in a network. Additionally, the workload of programming flows may be directly proportional to the number of switches and endpoints in a network. Workload scalability may become a concern as network sizes increase. For example, to cope with the number of endpoint-to-endpoint flows, a network administrator may need to utilize more expensive and power intensive switches having larger ternary content addressable memory (TCAM) tables.

[11] Additionally, a network controller may add and remove flow rules on a per flow basis. This may require significant overhead, particularly if there are many transient or ephemeral flows. For example, a network controller may send out a set of flow rules for each new Transport Control Protocol (TCP) connection, and may remove the flow rules when the TCP connection ends.

[12] Some implementations of the disclosed technology rewrite packets such that packets between endpoint hosts are rewritten to appear to be addressed to and from the edge switches connected to the hosts. Accordingly, the edge switches may act as proxies for their connected endpoint host devices. This may simplify the number of different paths through a network because the number of switches in the

network may be less than the number of endpoints in the network. In turn, this may reduce the forwarding table resources needed to implement the network.

[13] Figure 1 illustrates an example process of network control including rewriting address resolution packets. For example, the illustrated process may be performed by a network controller in an SDN network, such as an OPENFLOW network. The network controller may be connected to a number of switches, such as bridges and routers, through a control interface. For example, the control interface may be an out-of-band network coupling the switches to the network controller. In various implementations, the network controller and connected network devices may be physical or virtual network devices. By performing the example process, the network controller may cause edge switches to act as proxies for connected devices, reducing the number of flow rules needed to provide paths between devices.

[14] In some examples, edge switches may be at the edge of the SDN network controlled by the network controller. In these examples, the edge switch may be directly connected to the host or may be connected to the host via one or more devices that are not controlled by the network controller.

[15] The example process includes a step 101 of receiving an address resolution request from a first switch. For example, step 101 may be performed by the network controller. The address resolution request may be forwarded by the first switch from a source host. For example, the first switch may receive and forward the address resolution request from a directly or indirectly connected source host. For example, the address resolution request may be an Address Resolution Protocol (ARP) request or a Neighbor Discovery Protocol (NDP) solicitation.

[16] In some implementations, a network controller may provide address resolution request forwarding rules to its controlled switches. The address resolution request forwarding rules may cause the switches to forward packets to the network controller if the packets have a packet type matching the address resolution request packet type. For example, the forwarding action may occur if the packet has an ARP EtherType or NDP Internet Control Message Protocol version 6 (ICMPv6) type.

[17] The example process includes a step 102 of rewriting the address resolution request by changing a request source link address to a first switch link address. For example, the request source link address may be a MAC address of a source host that generated an ARP request. The first switch link address may be a MAC address of an edge switch that is connected to the source host. After rewriting

the request source link address, the source host's internetwork address may be associated with the first switch's link address. For example, the first switch's MAC address may be associated with the source host's IP address in address tables maintained by other network devices.

[18] In some cases, the request source link address may be included at multiple locations in the request. For example, the request source link address may occur in an Ethernet frame header and in an ARP packet transported in the Ethernet frame. Step 102 may include changing the request source link address at any location it occurs. For example, the network controller may rewrite the request as indicated in Table 2.

TABLE 2

	Ethernet Frame Header Fields			ARP Packet Fields				
	DA	SA	Type	Header	Sender MAC	Sender IP	Target MAC	Target IP
Received Request	FFFF...	MAC _H1	0x0806	ARP Request Header	MAC_H1	IP_H1	000...	IP_H2
Rewritten Request	FFFF...	MAC_S1	0x0806	ARP Request Header	MAC_S1	IP_H1	000...	IP_H2

In Table 2, DA denotes a destination MAC address, which is a placeholder value in an ARP request frame; SA denotes a source MAC address; MAC_H1 denotes the source host's MAC address; MAC_S1 denotes MAC address of the edge switch connected to the source host;0x0806 is the ARP EtherType code; IP_H1 denotes the source host's Internet Protocol (IP) address; IP_H2 denotes the destination host's IP address; and the target MAC address in the ARP packet is another placeholder value.

[19] The example process includes a step 103 of transmitting the rewritten address resolution request to a second switch. For example, the second switch may be an edge switch connected to a destination host specified in the address resolution request. In some implementations, step 103 includes broadcasting the rewritten address resolution request to the controlled switches of the network. For example, the controller may send a separate outgoing message to each of the controlled switches in a subnet of the network using an out-of-band SDN control network. As another example, the controller may broadcast rewritten messages in-band in frames

addressed to the switches. As a further example, the second switch may be a router or gateway connected to another subnet.

[20] The example process also includes a step 104 of receiving an address resolution response from the second switch. For example, step 104 may be performed by the network controller. In some cases, the address resolution response may be forwarded to the controller by the second switch from a target host specified in the address resolution request. In other cases, the address resolution response may be sent to the controller by a router or gateway. For example, the address resolution response may be an ARP response to an ARP request or an NDP advertisement in response to an NDP solicitation.

[21] In some implementations, the network controller may provide address resolution response forwarding rules to its controlled switches. The address request resolution response forwarding rules may cause the switches to forward packets to the network controller if the packets have a packet type matching the address resolution request packet type. The response received in step 104 may be an address resolution response forwarded by the second switch according to such a rule. In some implementations, the address request resolution response forwarding rules may be the same as the address request resolution request forwarding rules. For example, a single rule may be provided for both request and response forwarding if the rule specifies forwarding all packets having an ARP EtherType or NDP ICMPv6.

[22] The example process further includes a step 105 of rewriting the address resolution response by changing a response source link address to a second switch link address. For example, the response source link address may be a MAC address of a destination host that generated the ARP response. The second switch link address may be the MAC address of an edge switch that is connected to the destination host. After rewriting the response source link address, the destination host's internetwork address may be associated with the second switch's link address. For example, the second switch's MAC address may be associated with the destination host's IP address in address tables, such as ARP caches, maintained by network devices.

[23] In some cases, the response source link address may be included at multiple locations in the request. For example, the response source link address may occur in an Ethernet frame header and in an ARP packet transported in the

Ethernet frame. Step 104 may include changing the response source link address at any location it occurs. For example, if the response is to the rewritten request described in relation to Table 2, the network controller may rewrite the response as indicated in Table 3.

TABLE 3

	Ethernet Frame Header Fields			ARP Packet Fields				
	DA	SA	Type	Header	Sender MAC	Sender IP	Target MAC	Target IP
Received Response	MAC_S1	MAC _H2	0x0806	ARP Response Header	MAC_H2	IP_H2	MAC_S1	IP_H1
Rewritten Response	MAC_S1	MAC _S2	0x0806	ARP Response Header	MAC_S2	IP_H2	MAC_S1	IP_H1

In Table 3, MAC_H2 denotes the destination host's MAC address; MAC_S2 denotes the second switch's MAC address; and the remaining entries have the same meaning as in Table 2.

[24] In further implementations, step 105 may further include changing a response destination link address to the source host link address. In some implementations, the network controller may obtain the host link address from the address resolution request received in step 101. Changing the response destination link address may prevent the source host from dropping the address resolution response. For example, if the received response is a response to the rewritten request described in relation to Table 2, the network controller may rewrite the response as indicated in Table 4.

TABLE 4

	Ethernet Frame Header Fields			ARP Packet Fields				
	DA	SA	Type	Header	Sender MAC	Sender IP	Target MAC	Target IP
Received Response	MAC_S1	MAC _H2	0x0806	ARP Response Header	MAC_H2	IP_H2	MAC_S1	IP_H1
Rewritten Response	MAC_H1	MAC _S2	0x0806	ARP Response Header	MAC_S2	IP_H2	MAC_H1	IP_H1

In Table 4, entries have the same meaning as in Tables 2 and 3. In other implementations, the network controller may provide a rule to the first switch to rewrite the destination link of the response packet to the source host's link address.

[25] The example process further includes a step 106 of transmitting the rewritten address resolution response to the first switch. For example, step 106 may include transmitting the rewritten address resolution response out-of-band using an SDN control network. As another example, step 106 may include transmitting the rewritten response in-band encapsulated in a frame addressed to the first switch. The first switch may then transmit the rewritten response to the source host. Accordingly, the source host may update its address table to associate the destination host's internetworking address with the second switch's link address. For example, the source host may update its ARP cache to associate the destination host's IP address with the second switch's MAC address.

[26] Figure 2 illustrates an example process of network control including managing an endpoint database. For example, the illustrated process may be performed by a network controller managing an SDN network, such as an OPENFLOW network.

[27] The example process includes a step 201 of receiving an address resolution request. For example, like the example process illustrated in Figure 1, the address resolution request may be received by a network controller after being forwarded by a first switch. In some implementations, the request may be received out-of-band via a control interface.

[28] The example process further includes a step 202 of obtaining a host link address from the address resolution request. For example, a network controller may obtain a source host's MAC address from an ARP request or an NDP solicitation obtained in step 201.

[29] In the illustrated example, the network controller may maintain an endpoint database 204. The endpoint database 204 may have entries associating host devices' link and internet addresses with their connected edge switches' identifications. Entries in the endpoint database 204 may also include a port or interface number connecting the edge switch and its associated host device. For example, Table 5 shows an example entry of an example endpoint database 204.

TABLE 5

IP Address	MAC address	Edge Switch ID	Edge Switch Interface
IP_H1	MAC_H1	S1	*K*

In Table 5, S1 is a device identification number used by the controller to identify the first switch, such as a link address of the first switch; *K* is the interface to which the source host is connected, such as a port number; and the remaining entries have the same meaning as in Tables 2-4.

[30] The process includes a step 203 of determining if the host link address obtained in step 202 is in the endpoint database 204. If not, then a step 205 of storing the first host link address may be performed. In step 205, the stored first host link address may be associated with a first host internetwork address and the first switch link address in the endpoint database 204.

[31] The illustrated process further includes a step 206 of rewriting the address resolution request. Step 206 may include rewriting the address resolution request by change the request source link address to the first switch's link address. If the request source link address corresponds to a host link address in the database 204, then step 206 may be performed after step 203 using information obtained from the database 204. For example, during a subsequent operation, the controller may perform step 201 by receiving a second address resolution request. The second address resolution request may have a request destination internetwork address that is a stored internetwork address stored in the endpoint database 204. In this example, the controller may perform step 206 after step 203.

[32] The process further includes a step 207 of transmitting the rewritten request and a step 208 of receiving a response to the rewritten request. For example, steps 207 and 208 may be performed in a manner similar to steps 103 and 104 of Figure 1.

[33] The example process includes a step 209 of obtaining a second host link address from the address resolution response. For example, the second host link address may be the requested MAC address of a target host indicated in the address resolution request.

[34] The example process further includes a step 210 of storing the second host link address associated with a second host internetwork address and the second switch link address in the endpoint database 204. For example, Table 6 illustrates an example entry for the endpoint database 204.

TABLE 6

IP Address	MAC address	Edge Switch ID	Edge Switch Interface

IP_H2	MAC_H2	S2	*L*

In Table 6, S2 is a device identification number used by the controller to identify the second switch, such as a link address of the second switch; *L* is the interface to which the destination host is connected, such as a port number; and the remaining entries have the same meaning as in Tables 2-4.

[35] After step 210, the example process may proceed as described with respect to steps 105 and 106 of Figure 1.

[36] In some implementations, the network controller may use information in the endpoint database 204 to bypass a remainder of the address resolution process. For example, in the illustrated process, during a subsequent operation the network controller may perform step 201 by receiving a second address resolution request. During step 203, the network controller may determine that the request has a request destination internetwork address that is a stored internetwork address stored in the endpoint database 204. For example, the request may have a target IP address that corresponds to a host IP address stored in the database 204.

[37] In these cases, the process may include a step 211 of generating an address resolution response. For example, step 208 may include generating an address resolution response having a response destination link that is switch link address associated with the stored internetwork address. As an example, in step 201, the controller may receive an ARP request from host *H1* requesting the MAC address of a second destination host, *H3*. The network controller may determine in step 203 that *H3*'s IP address is stored in the endpoint database 204 and associated with a third edge switch *S3*. If so, the network controller may perform step 211 to generate an ARP response including S3's MAC address as the destination MAC address

[38] The network controller may then perform step 212 by transmitting the generated address resolution response. In some implementations, step 212 may comprise broadcasting the generated address resolution response to the controlled switches. For example, the controller may broadcast the response to allow the network to update address tables. As another example, the network controller may perform step 212 by transmitting the generated address resolution response only to the edge switch connected to the host performing the address resolution request.

[39] In some implementations, a network controller may provide path management rules to its controlled switches. For example, rules may be formatted and transmitted in compliance with an SDN networking protocol, such as an OPENFLOW standard. Figure 3 illustrates an example process of establishing a network path between edge switches. For example, the illustrated process may be performed to cause edge switches to translate packet addresses so that packets are delivered to the correct host devices. As a further example, the illustrated process may be performed to manage paths between edge switches through transit switches.

[40] The example process may also include a step 303 of establishing network paths. For example, with respect to Figure 1, the step 301 may be performed to establish a path between the first switch connected to the first host transmitting the address resolution request and the second switch connected to the second host transmitting the address resolution response. The path established in step 301 may be used for subsequent flows between the first host and the second host. The path may be implemented as forwarding rules that define how transit switches between the edge switches handle packets from the first edge switch to the second edge switch. For example, the forwarding rules may be based only on destination link addresses as match criteria. As another example, the forwarding rules may be based on a destination link and source link address pair. In the second example, multiple link level paths may be used for traffic to a single edge switch.

[41] In further implementations, step 303 may include establishing multiple paths between a pair of edge switches. For example, multiple paths may be established to implement Quality of Service (QoS). In this example, QoS values may be included in forwarding rule criteria to indicate different paths for the packet. Additionally, edge switches may be provided with rewriting rules to establish the QoS values. For example, QoS value rewriting rules may be provided if hosts may not be trustworthy to set their own QoS values.

[42] The example process may include a step 302 of transmitting a first rule to a first switch. For example, the first switch may be an edge switch connected to a host that transmitted an address resolution request or address resolution response that was rewritten during the process illustrated and described with respect to Figure 1. The first rule may instruct the first switch to rewrite a first packet received from a host to have a source link address that is the link address of the switch. The first rule may specify the host's internetwork address as a match condition. For example,

the rule may specify that any packet having the host's IP address as the source IP address be rewritten so that the source MAC address is replaced with the first switch's MAC address. In some implementations, this rule may ensure that the source link address of a packet matches the entry in the ARP cache of the destination host. This may prevent destination host security or firewall software from interfering with packet delivery.

[43] In further implementations, the first rule may include further actions. For example, the first rule may instruct the first switch to output the packet on a specific output port. This may allow the network manager to select the path to be used to transmit packets from the host to other network devices. For example, the first rule may include an output instruction to implement the path established in step 301. In some implementations, the first rule may also include any QoS rewriting rules established as part of step 301.

[44] The example process may also include a step 302 of transmitting a second rule to the first switch. The second rule may instruct the first switch to rewrite a second packet destined for the host to have a destination link address that is the host's link address. The second rule may specify the host's internetwork address as a match condition. For example, the rule may specify that any packet having the host's IP address as the destination IP address be rewritten so that the destination MAC address is replaced with the host's MAC address. The second rule may include further actions. For example, the second rule may instruct the switch to output the packet on the interface connected to the host. In some implementations, after an address resolution rewriting process, other network devices may associate the host's internetworking address with the switch's link address. Accordingly, packets destined for the host's internetworking address may be addressed to the switch's link address. After implementing the second rule, packets may be written to be correctly delivered to the host.

[45] The example process may also include a step 304 of transmitting a forwarding rule to a third switch in the path to implement the path established in step 301. For example, the forwarding rule may define a match criterion based on source and destination link addresses and may specify an outgoing forwarding interface. In some implementations, because the source and destination link addresses are rewritten to be edge switch addresses, the number of forwarding rules needed to handle packets is reduced. This may reduce the flow rule programming overhead on

the controller. This may also decouple the scalability of the number of supportable endpoints from the size of the forwarding tables on the network switches.

[46] In some implementations, step 304 is repeated to provide a second forwarding rule to a fourth switch in a second path between the first and second switches. For example, to establish different levels of QoS, the network controller may establish multiple paths between pairs of switches. In these implementations, the rules transmitted in step 304 may further include QoS values as match criteria.

[47] Additionally, in some implementations, the forwarding rule transmitted in step 304 may be maintained while the hosts are present on the network and connected to the same edge switches. In these implementations, clearing non-active endpoint flows does not require removing rules from transit switches. Rather, clearing non-active endpoint flows may be implemented by transmitting clearing instructions only to the corresponding edge switches. This may reduce the burden on the network controller.

[48] Figure 4 illustrates an example network controller 401 including an address resolution rewriter 406 and rule manager 407. For example, the example network controller 401 may comprise an SDN controller managing an SDN network, such as an OPENFLOW network. In some examples, the illustrated modules may be implemented as processor-executable instructions on non-transitory computer readable media, as hardware, or as a combination of the two. In some implementations, during operation, the example network controller 401 may perform the processes described and illustrated with respect to Figures 1, 2, or 3.

[49] The example network controller 401 may include an interface 408. For example, the interface 408 may be an out-of-band control interface to connect to controlled switches. As another example, the interface 408 may be a network interface to connect in-band with controlled switches.

[50] In some implementations, the interface 408 may allow the controller 401 to receive an address resolution request from a first switch and to receive an address resolution response from a second switch. The interface 408 may also be used to transmit rewritten address resolution packets. For example, the interface 408 may be used to transmit a rewritten address resolution response to the second switch and a rewritten address resolution request to the first switch.

[51] The network controller 401 may also include an address resolution rewriter 406. The address resolution rewriter 406 may rewrite the address resolution

request received using the interface 408. For example, the rewriter 406 may rewrite the address resolution request by changing a request source link address to the first switch's link address.

[52] The address resolution rewriter 406 may also rewrite the address resolution response by changing a response source address to the second switch's link address. In further implementations, the rewriter 406 may rewrite the address resolution response by changing the destination link address of the response to a host's link address. For example, the host's link address may be the link of the host that generated the address resolution request.

[53] The example network controller 401 may also include a database manager 404. For example, the database manager 404 may manage an endpoint database 402. The database manager 404 may store host link addresses associated with host internetwork addresses and switch link addressed. For example, the database manager 404 may obtain and store a first host link address, first host internetwork address, and first switch link address from an address resolution request received through the interface 408. As another example, the database manager 404 may obtain and store a second host link address, second host internetwork address, and second switch link address from an address resolution response received through the interface 408. The database manager 404 may obtain this information directly from the interface 408, or from the address resolution rewriter 406. Additionally, the database manager 404 may provide the entries from the endpoint database 402 to the rewriter 406.

[54] The example network controller 401 may also include a path manager 405. The path manager 405 may be used to establish paths to destination edge switches or between pairs of switches. For example, the path manager 405 may establish a path to the switch that forwarded the address resolution response. As another example, the path manager 405 may establish a path between the switch that forwarded the address resolution request received and the switch that forwarded the address resolution response.

[55] In some implementations, the path manager 405 may establish multiple paths between pairs of switches. To implement the paths, the path manager 405 may transmit a first forwarding rule to a switch in a first path and may transmit a second forwarding rule to a switch in a second path. For example, the path manager 405 may establish multiple paths between pairs of switches to implement QoS. In

this example, the path manager 405 may transmit a first set of rules to transit switches in a first QoS path and a second set of rules to transit switches in a second QoS path.

[56] The example network controller 401 further includes a rule manager 407. In some implementations, the rule manager 407 may provide a first rule to a switch to rewrite a source link address of a received packet. For example, the rule manager 407 may provide a rule to the first switch that forwarded the address resolution request. The first rule may instruct the first switch to rewrite packets received from the host that generated the address resolution request. The first rule may cause the first switch to rewrite the packets' source link address to the switch's link address. For example, the rule manager 407 may specify the host's internetwork address as a match criterion to cause the second switch to perform the rewrite.

[57] In further implementations, the rule manager 407 may provide a second rule to the second switch that forwarded the address resolution response. The second rule may instruct the second switch to rewrite the source link address of packets received from the host that generated the address resolution response. The second rule may cause the second switch to rewrite the packets' source link address to second switch's link address. For example, the rule manager 407 may specify the host's internetwork address as a match criterion to cause the second switch to perform the rewrite.

[58] In further implementations, the rule manager 407 may provide a rule to a switch to rewrite the destination link address of a received packet. For example, the rule manager 407 may provide a second rule to the first switch that forwarded the address resolution request. The second rule may instruct the first switch to rewrite the destination link address to the link address of the host that transmitted the address resolution request. For example, the rule manager 407 may specify the host's internetwork address as a match criterion to cause the first switch to perform the rewrite.

[59] In further implementations, the rule manager 407 may provide a rule to provide a second rule to the second switch to rewrite a source link address of a second packet to the second switch link address if the second packet is received from a second host.

[60] Figure 5 illustrates an example network switch 501 including a non-transitory computer readable medium 506 storing instructions 508 to cause the switch 501 to rewrite packets. In some implementations, the switch 501 may be a switch an SDN network controlled by a network controller. For example, the switch 501 may be an edge switch interacting with a network controller performing a process, such as the example processes illustrated in Figures 1-3.

[61] The example switch 501 may include a processor 503 coupled to a control interface 502, a non-transitory computer readable medium 505, and switch hardware 506. For example, the processor 503 may be a network processor implementing an SDN protocol stack 507 to configure switch hardware 506, such as TCAM tables, according to rules 508 provided by a controller over the control interface 502. In some implementations, the non-transitory computer readable medium 505 may include read-only memory (ROM), random-access memory (RAM), TCAM, flash memory, other storage, or a combination thereof.

[62] In the illustrated example, the rules 508 may include instructions 509 to cause the switch 501 to forward an address resolution packet to a network controller. In various implementations, the address resolution packet may be an address resolution request or an address resolution response. For example, the instructions 509 may cause the switch 501 to forward an address resolution request when acting as an edge switch for a source host. As another example, the instructions 509 may cause the switch 501 to forward an address resolution response when acting as an edge switch for a target host. In some implementations, when executing the instructions 509, the processor 503 may forward an address resolution packet received from a host via the network interface 504 to a controller via control interface 502. In some implementations, the instructions 509 may be received from a network controller via the controller interface 502. For example, the instructions 509 may be received during configuration when the switch 501 joins a network.

[63] In the illustrated example, the rules 508 further include instructions 510 to rewrite a source link address of a first packet if a first packet is received from a host. For example, the host may be a host that generated an address resolution packet forwarded in accordance with instructions 509. In some implementations, the instructions 510 may be executable by the processor 503 to rewrite the source link address of a first packet received from the host via the network interface 504. In some implementations, the instructions 510 instruct the processor to rewrite the

source link address of the first packet to the switch's link address 512. For example, the link address 512 may be the switch's 501 MAC address.

[64] The instructions 510 may cause the processor 503 to determine if the first packet is from the host by determining if the first packet has a source internetwork address that is the host's internetwork address. For example, the instructions 510 may be provided via the control interface 502 as a match criterion and an action. The match criterion may match the packet's source internetwork address and the host's internetwork address. The action may be to rewrite the packet's source link address to the switch's link address 512.

[65] In the illustrated example, the rules 508 further include instructions 511 to rewrite a destination link address of a second packet if a second packet destined for the host is received. In some implementations, the instructions 511 may be executable by the processor 503 to rewrite the destination link address of the second packet to the host's link address.

[66] The instructions 511 may cause the processor 503 to determine if the second packet is destined for the host by determining if the second packet has a destination internetwork address that is the host's internetwork address. For example, the instructions 511 may be provided via the control interface 502 as a match criterion and an action. The match criterion may match the packet's destination internetwork address and the host's internetwork address. The action may be to rewrite the second packet's destination link address to the host's link address.

[67] In the foregoing description, numerous details are set forth to provide an understanding of the subject disclosed herein. However, implementations may be practiced without some or all of these details. Other implementations may include modifications and variations from the details discussed above. It is intended that the appended claims cover such modifications and variations.

CLAIMS

1. A method, comprising:

receiving an address resolution request from a first switch;

rewriting the address resolution request by changing a request source link address to a first switch link address;

transmitting the rewritten address resolution request to a second switch;

receiving an address resolution response from the second switch;

rewriting the address resolution response by changing a response source link address to a second switch link address; and

transmitting the rewritten address resolution response to the first switch.

2. The method of claim 1, further comprising:

obtaining a host link address from the address resolution request; and

wherein the step of rewriting the address resolution response further comprises changing a response destination link address to the host link address.

3. The method of claim 2, further comprising:

storing the first host link address associated with a first host internetwork address and the first switch link address in an endpoint database;

obtaining a second host link address from the address resolution response; and

storing the second host link address associated with a second host internetwork address and the second switch link address in the endpoint database.

4. The method of claim 3, further comprising:

receiving a second address resolution request having a request destination internetwork address that is a stored internetwork address stored in the endpoint database; and

transmitting a second address resolution response having a second response destination link address that is an associated switch link address associated with the stored internetwork address.

5. The method of claim 1, further comprising:

transmitting, to the first switch, a first rule to rewrite a first packet received from a host to have a source link address that is the first switch link address; and

transmitting, to the first switch, a second rule to rewrite a second packet destined for the host to have a destination link address that is a host link address.

6. The method of claim 1, further comprising:

establishing a path between the first switch and the second switch; and

transmitting a forwarding rule to a third switch in the path to implement the path.

7. A network controller, comprising:

an interface to:

receive an address resolution request from a first switch, and

receive an address resolution response from a second switch; and

an address resolution rewriter to:

rewrite the address resolution request by changing a request source link address to a first switch link address, and

rewrite the address resolution response by changing a response source address to a second switch link address;

wherein the interface is to:

transmit the rewritten address resolution request to the second switch, and

transmit the rewritten address resolution response to the first switch.

8. The network controller of claim 7, wherein:

the address resolution rewriter is to rewrite the address resolution response by changing a destination link address to a host link address.

9. The network controller of claim 7, further comprising:

a database manager to store in an endpoint database:

a first host link address associated with a first host internetwork address and the first switch link address; and

a second host link address associated with a second host internetwork address and the second switch link address.

10. The network controller of claim 7, further comprising:

a path manager to establish a first path between the first switch and the second switch by transmitting a first forwarding rule to a third switch in the first path,

and to establish a second path between the first switch and the second switch by transmitting a second forwarding rule to a fourth switch in the second path.

11. The network controller of claim 7, further comprising:

a rule manager to provide a rule to the first switch to rewrite a source link address of a packet to the first switch link address if the packet is received from a host.

12. The network controller of claim 11, wherein:

the rule manager is to provide a second rule to the second switch to rewrite a source link address of a second packet to the second switch link address if the second packet is received from a second host.

13. The network controller of claim 11, wherein:

the rule manager is to provide a second rule to the first switch to rewrite a destination link address of a second packet to a host link address if the second packet is destined for the host.

14. A non-transitory computer readable medium, storing instructions executable by a processor of a switch to cause the switch to:

forward an address resolution packet to a network controller;

if a first packet is received from a host, rewrite a source link address of the first packet to a switch link address; and

if a second packet is received that is destined for the host, rewrite a destination link address of the second packet to a host link address.

15. The non-transitory computer readable medium of claim 14, wherein, when executed, the instructions cause the switch to:

determine if the first packet is received from the host by determining if the first packet has a source internetwork address that is a host internetwork address; and

determine if the second packet is destined for the host by determining if the second packet has a destination internetwork address that is the host internetwork address.

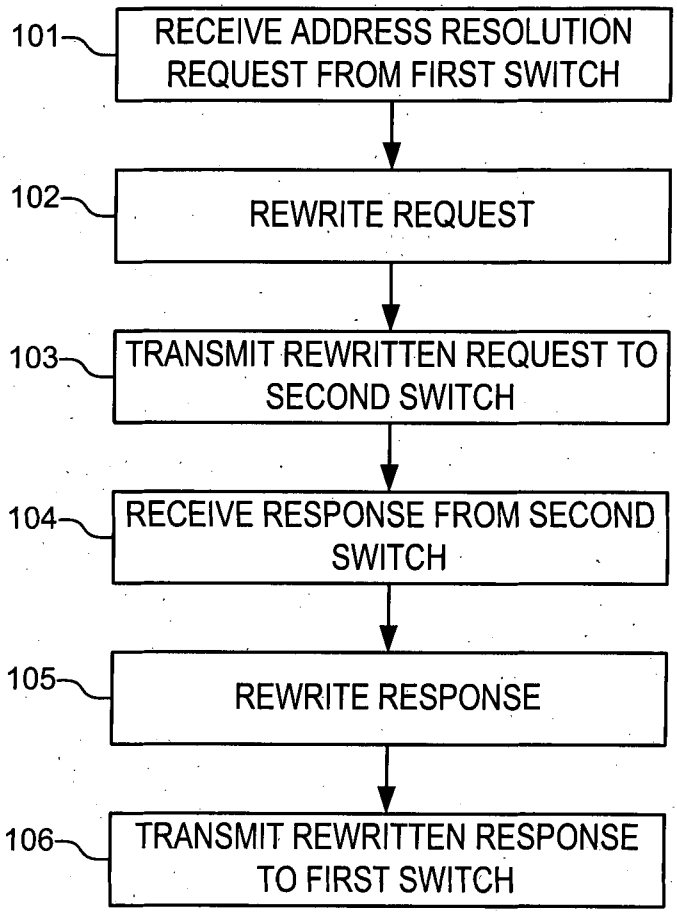

101
RECEIVE ADDRESS RESOLUTION REQUEST FROM FIRST SWITCH
102
REWRITE REQUEST
103
TRANSMIT REWRITTEN REQUEST TO SECOND SWITCH
104
RECEIVE RESPONSE FROM SECOND SWITCH
105
REWRITE RESPONSE
106
TRANSMIT REWRITTEN RESPONSE TO FIRST SWITCH

FIG. 1

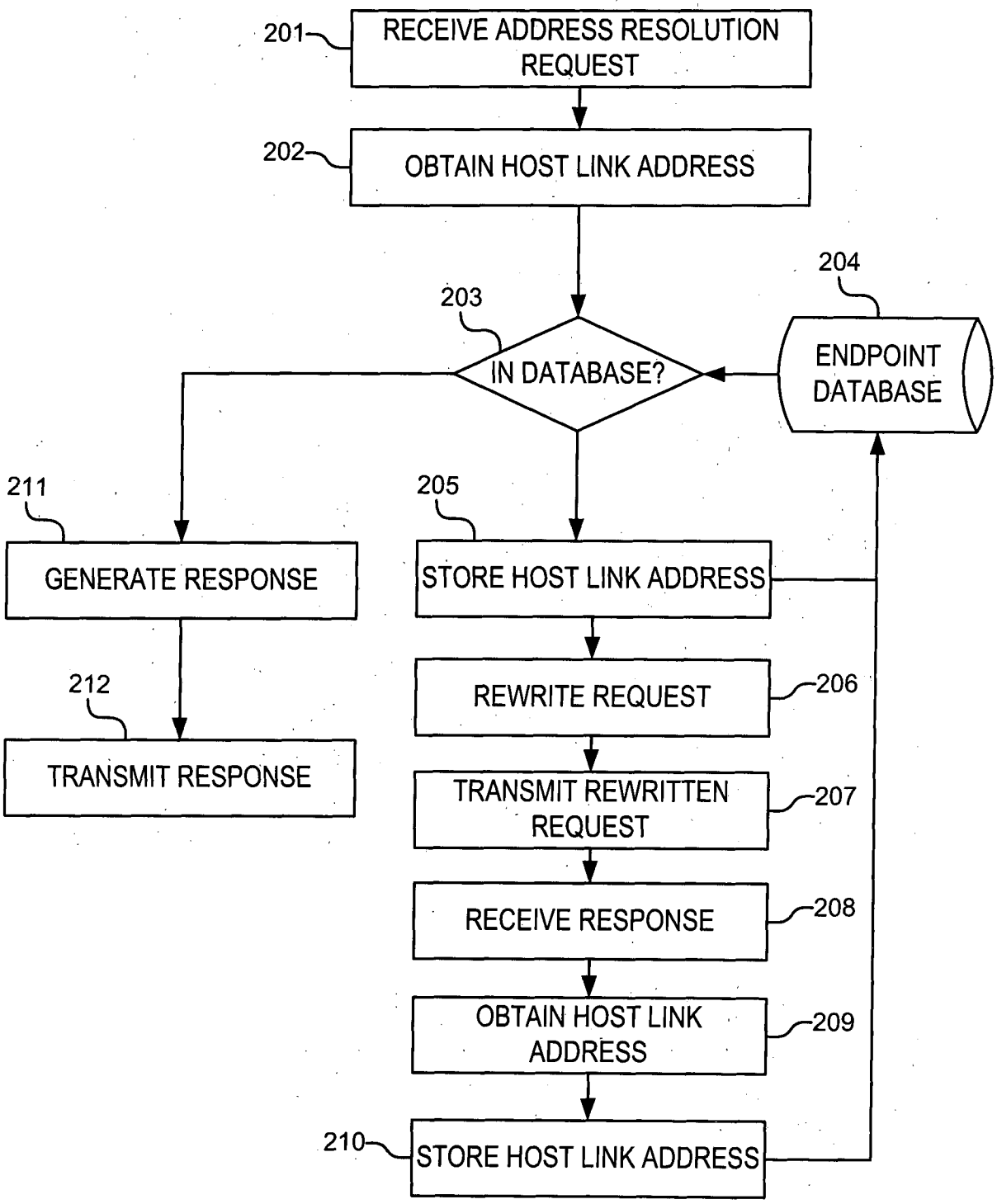

201
RECEIVE ADDRESS RESOLUTION REQUEST
202
OBTAIN HOST LINK ADDRESS
203
IN DATABASE?
204
ENDPOINT DATABASE
211
GENERATE RESPONSE
212
TRANSMIT RESPONSE
205
STORE HOST LINK ADDRESS
REWRITE REQUEST
206
TRANSMIT REWRITTEN REQUEST
207
RECEIVE RESPONSE
208
OBTAIN HOST LINK ADDRESS
209
210
STORE HOST LINK ADDRESS

FIG. 2

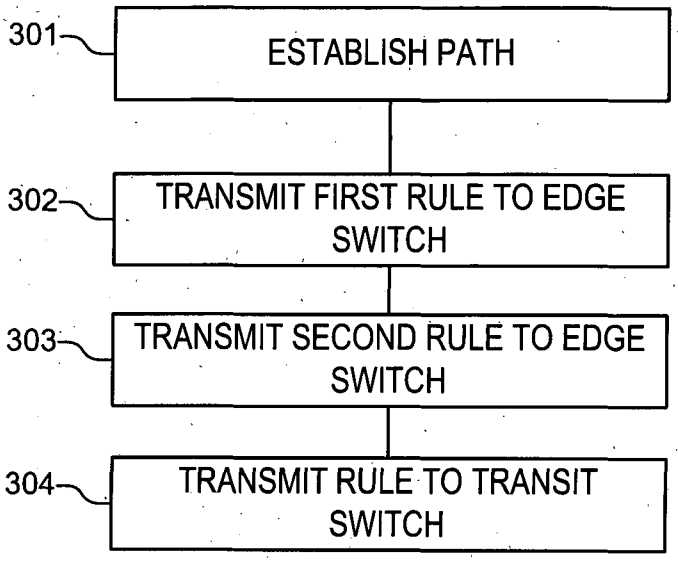

301
ESTABLISH PATH
302
TRANSMIT FIRST RULE TO EDGE SWITCH
303
TRANSMIT SECOND RULE TO EDGE SWITCH
304
TRANSMIT RULE TO TRANSIT SWITCH

FIG. 3

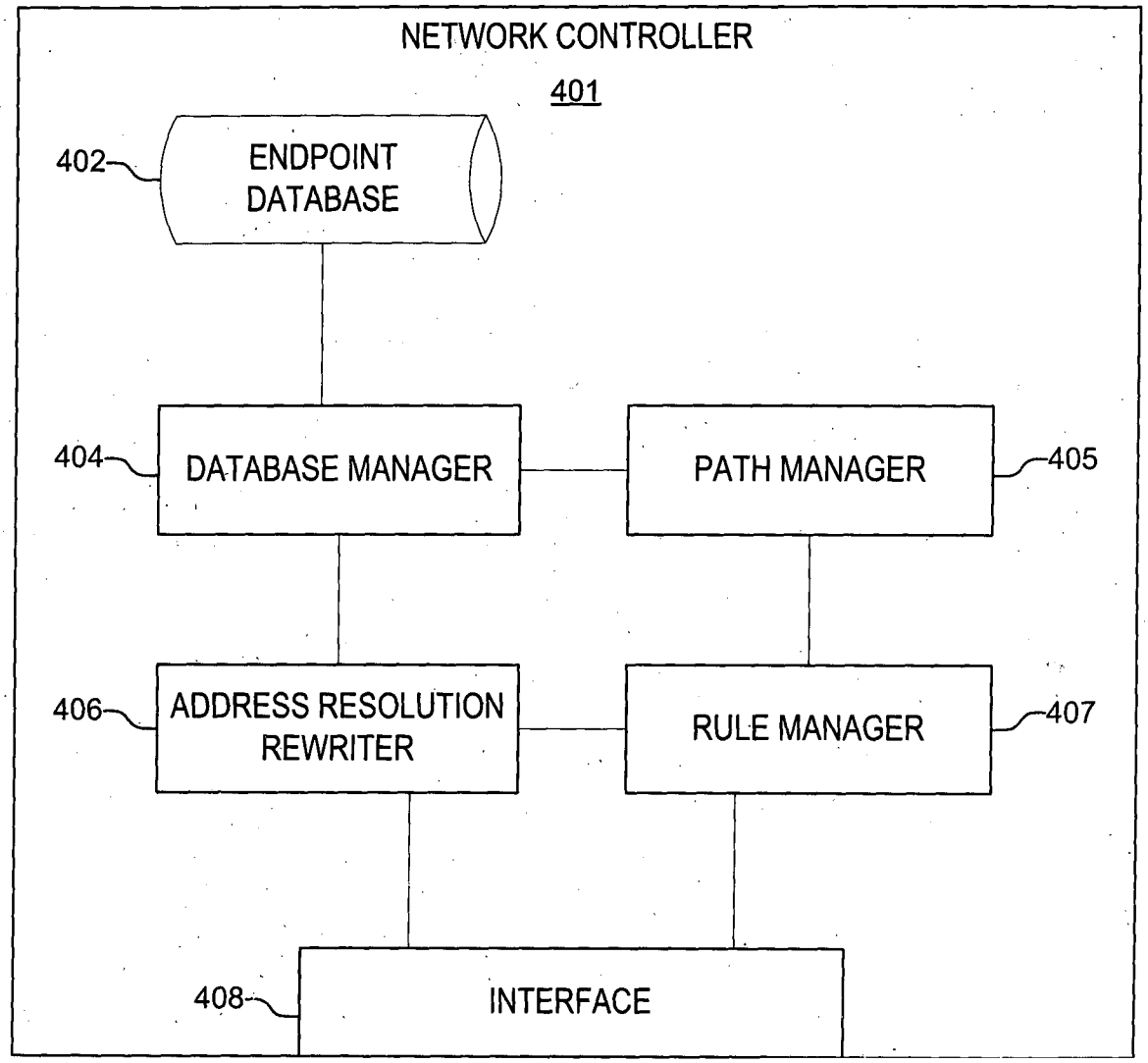

NETWORK CONTROLLER
401
402
ENDPOINT DATABASE
404
DATABASE MANAGER
PATH MANAGER
405
406
ADDRESS RESOLUTION REWRITER
RULE MANAGER
407
408
INTERFACE

FIG. 4

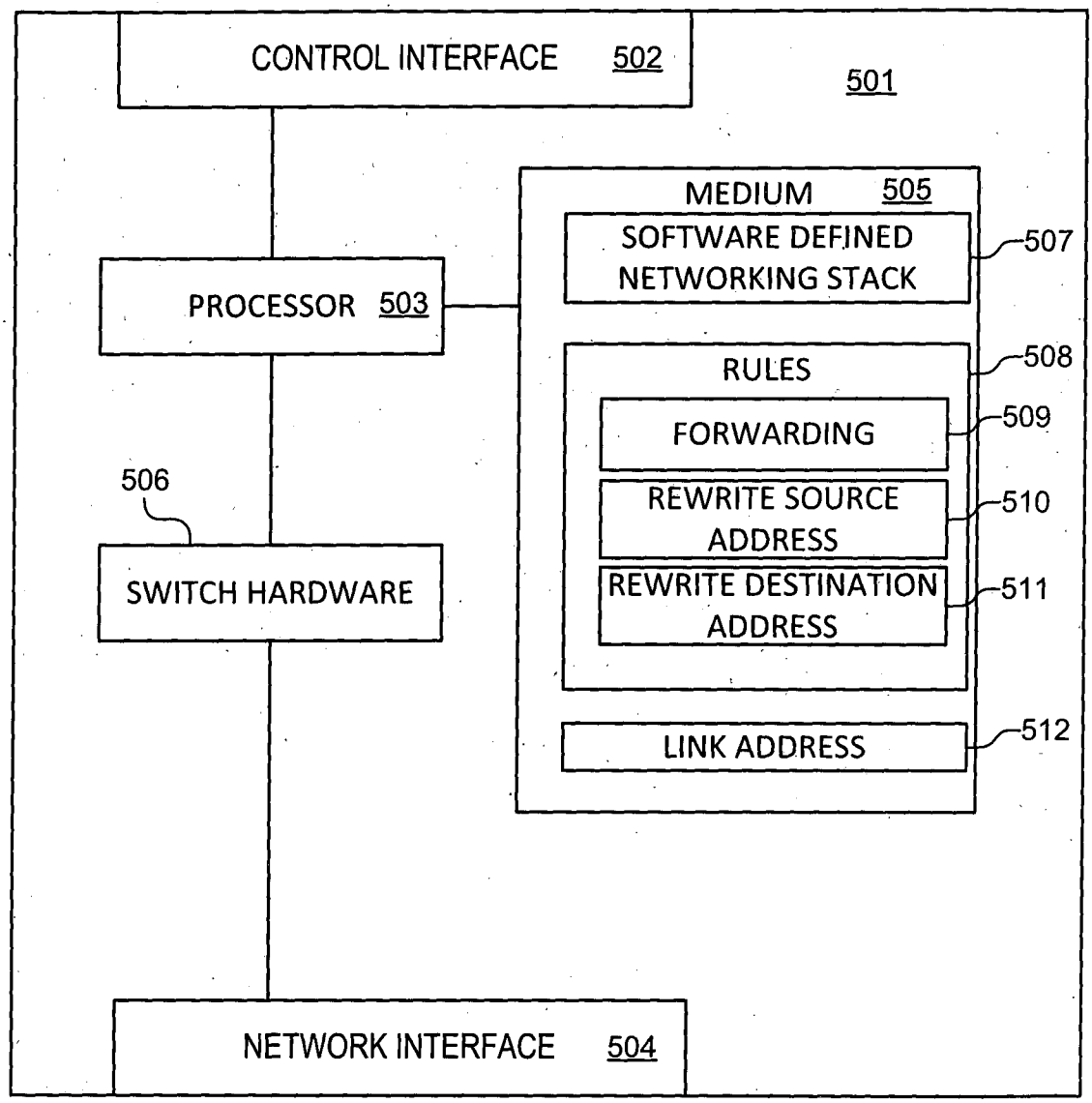

CONTROL INTERFACE 502
501
MEDIUM 505
SOFTWARE DEFINED NETWORKING STACK
507
PROCESSOR 503
RULES
508
FORWARDING
509
506
REWRITE SOURCE ADDRESS
510
SWITCH HARDWARE
REWRITE DESTINATION ADDRESS
511
LINK ADDRESS
512
NETWORK INTERFACE 504

FIG. 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IN2013/000477

A. CLASSIFICATION OF SUBJECT MATTER

H04L 12/00(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)

H04Q; H04W; H04L; H04B; H04J; H04M

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;CNPAT;CNKI: address resolution protocol request, address resolution request, ARP request, address resolution protocol response, address resolution respnse, ARP response, address, MAC, rewrite, change, alter, modify, amend, host, switch, packet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 103023818A ((HANGZHOU H3C TECHNOLOGIES CO., LTD.)) 03 April 2013 (2013-04-03) description, paragraphs [0056]-[0160]	1-15
A	CN 102355417A ((HANGZHOU H3C TECHNOLOGIES CO., LTD.)) 15 February 2012 (2012-02-15) the whole document	1-15
A	JP 2008109357A ((MATSUSHITA ELECTRIC WORKS LTD.)) 08 May 2008 (2008-05-08) 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
14 April 2014	**08 May 2014**

Name and mailing address of the ISA/	Authorized officer
STATE INTELLECTUAL PROPERTY OFFICE OF THE P.R.CHINA(ISA/CN) **6,Xitucheng Rd., Jimen Bridge, Haidian District, Beijing, China** **100088 China**	**ZHANG,Fan**
Facsimile No. **(86-10)62019451**	Telephone No. **(86-10)62413355**

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/IN2013/000477

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 103023818A	03 April 2013	None	
CN 102355417A	15 February 2012	WO 2013049989A1	11 April 2013
JP 2008109357A	08 May 2008	None	