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

(19) **World Intellectual Property Organization**
International Bureau

(43) **International Publication Date**
17 December 2020 (17.12.2020)

WIPO | PCT

(10) **International Publication Number**
WO 2020/248117 A1

(51) **International Patent Classification:**
H04L 12/723 (2013.01)

(21) **International Application Number:**
PCT/CN2019/090673

(22) **International Filing Date:**
11 June 2019 (11.06.2019)

(25) **Filing Language:** English

(26) **Publication Language:** English

(71) **Applicant: TELEFONAKTIEBOLAGET LM ERICSSON (PUBL)** [SE/SE]; S-164 83 Stockholm (SE).

(72) **Inventor; and**
(71) **Applicant** *(for SC only)*: **HE, Jiang** [CN/CN]; ET2, No. 5 Lize East Street, Chaoyang District, Beijing 100102 (CN).

(72) **Inventors: NIE, Bolin**; ET2, No. 5 Lize East Street, Chaoyang District, Beijing 100102 (CN). **ZHAO, Zhenning**; ET2, No. 5 Lize East Street, Chaoyang District, Beijing 100102 (CN).

(74) **Agent: CHINA PATENT AGENT (H. K.) LTD.**; 22/F, Great Eagle Centre, 23 Harbour Road, Wanchai, Hong Kong (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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, 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,

(54) **Title:** METHODS AND DEVICES FOR PACKET ROUTING IN COMMUNICATION NETWORKS

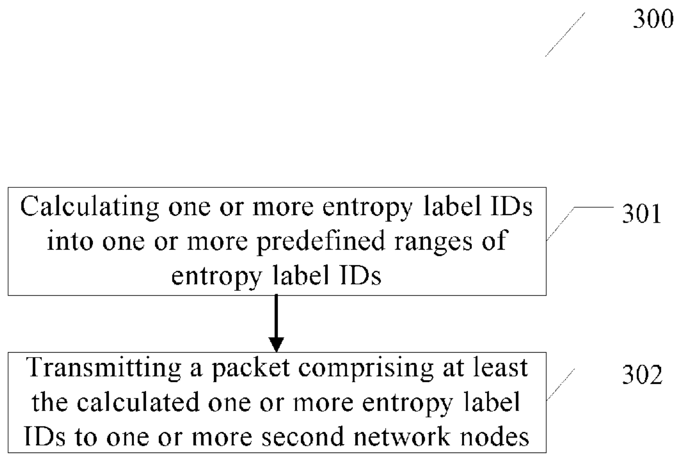

Fig. 3

(57) **Abstract:** A method implemented by a first network node in a communication network is provided. The method comprises: calculating one or more entropy label identifiers into one or more predefined ranges of entropy label identifiers; and transmitting a packet comprising at least the calculated one or more entropy label identifiers to one or more second network nodes. In the present disclosure, an MPLS label stack depth may be saved to mitigate stack depth challenges on the ingress node, especially for TE tunnel applications. Moreover, a light weighted solution may be achieved so that there is no impact on the transit nodes and the egress node, and on the ingress node, the switch chipset may support entropy with different ranges.

WO 2020/248117 A1

[Continued on next page]

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))*

METHODS AND DEVICES FOR PACKET ROUTING IN COMMUNICATION NETWORKS

TECHNICAL FIELD

The present disclosure generally relates to Multi-Protocol Label Switching (MPLS) communication networks, and more specifically to methods and devices for entropy label handling in the MPLS communication networks.

BACKGROUND

This section introduces aspects that may facilitate better understanding of the present disclosure. Accordingly, the statements of this section are to be read in this light and are not to be understood as admissions about what is in the prior art or what is not in the prior art.

Traffic load-balancing over Equal Cost Multi Path (ECMP) or Link Aggregation Group (LAG) is widely used in IP(Internet Protocol)/MPLS networks. Entropy Label and Flow-Aware Transport are mechanisms for MPLS transit nodes to perform load balance without deep packet inspection.

Traffic Engineering (TE) tunnels are popular in 5^{th} Generation (5G) networks to accommodate application specific performance such as low latency. Usually, multiple labels shall be encapsulated into packets for traffic steering.

VXLAN(Virtual eXtensible Local Area Network)/NVGRE(Network Virtualization using Generic Routing Encapsulation) are widely supported by switch chipsets in the market. Inside entropy information is calculated and set into packets with different sizes, e.g., about 14bit for VXLAN and about 8bit for NVGRE.

However, in the Entropy Label mechanism, a label pair (Entropy Label Indicator (ELI) + Entropy Label (EL)) may consume excessive parts of an MPLS label stack, and ingress nodes and transit nodes may be heavily impacted. When a plurality of label pairs are required for the transit nodes with limited entropy reachability, an ingress PE has to push to many labels. This may impose heavy impacts on customer networks and generally may not be acceptable for the TE tunnels.

Furthermore, in the Flow-Aware Transport mechanism, only Layer 2 Virtual Private Network (L2VPN) scenarios may be supported. Also, TE tunnel deployment may be limited. For instance, a flow label may always be located behind a pseudo wire label. For TE tunnels with deep tunnel labels, the flow label may not be reachable for the transit nodes.

SUMMARY

In the present disclosure, an MPLS entropy block may define a label ID range to carry entropy information. The entropy label is identified by the fact that the value of entropy label ID falls into the predefined label ID range. An entropy label having the entropy information may be pushed behind a tunnel label. On transit nodes, the entropy information carried in the entropy label may assist in traffic load balance. On egress tunnel endpoints, the entropy label may be popped.

According to a first aspect of the present disclosure, a method implemented by a first network node in a communication network is provided. The method comprises: calculating one or more entropy label identifiers into one or more predefined ranges of entropy label identifiers; and transmitting a packet comprising at least the calculated one or more entropy label identifiers to one or more second network nodes.

In an alternative embodiment of the first aspect, the one or more predefined ranges may be determined by the first network node.

In a further alternative embodiment of the first aspect, the one or more predefined ranges may be advertised by the first network node through Multi-Protocol Label Switching protocols.

In an alternative embodiment of the first aspect, the one or more predefined ranges may be received from the one or more second network nodes respectively.

In a further alternative embodiment of the first aspect, the one or more predefined ranges may be advertised by the one or more second network nodes respectively through Multi-Protocol Label Switching protocols.

In another alternative embodiment of the first aspect, each of the predefined ranges may indicated by a minimum entropy label identifier and a maximum entropy label identifier, or by a base entropy label identifier and a length of this range.

In still another alternative embodiment of the first aspect, the one or more predefined ranges may be configured.

In yet another alternative embodiment of the first aspect, the one or more entropy label identifiers may be pushed behind respective tunnel labels.

In yet another alternative embodiment of the first aspect, the one or more entropy label identifiers may be calculated from a 5-tuple for the packet.

According to a second aspect of the present disclosure, a method implemented by a second network node in a communication network is provided. The method comprises: receiving a packet comprising at least one or more entropy label identifiers from a first network node; and popping an entropy label having an entropy label identifier of the entropy label identifiers based on a predefined range of entropy label identifiers.

According to a third aspect of the present disclosure, a method implemented by a third network node in a communication network is provided. The method comprises: determining a range of entropy label identifiers; and advertising the range to a plurality of network nodes.

According to a fourth aspect of the present disclosure, a first network node in a communication network is provided. The first network node comprises a processor and a memory communicatively coupled to the processor. The memory is adapted to store instructions which, when executed by the processor, may cause the first network node to perform operations of the method according to the above first aspect.

According to a fifth aspect of the present disclosure, a second network node in a communication network is provided. The second network node comprises a processor and a memory communicatively coupled to the processor. The memory is adapted to store instructions which, when executed by the processor, may cause the second network node to perform operations of the method according to the above second aspect.

According to a sixth aspect of the present disclosure, a third network node in a communication network is provided. The third network node comprises a processor and a memory communicatively coupled to the processor. The memory is adapted to store instructions which, when executed by the processor, may cause the third network node to perform operations of the method according to the above third aspect.

According to a seventh aspect of the present disclosure, a non-transitory computer readable medium having a computer program stored thereon is provided. When the computer program is executed by a set of one or more processors of a first network node, the computer program may cause the first network node to perform operations of the method according to the above first aspect.

According to an eighth aspect of the present disclosure, a non-transitory computer readable medium having a computer program stored thereon is provided. When the computer program is executed by a set of one or more processors of a second network node, the computer program may cause the second network node to perform operations of the method according to the above second aspect.

According to a ninth aspect of the present disclosure, a non-transitory computer readable medium having a computer program stored thereon is provided. When the computer program is executed by a set of one or more processors of a third network node, the computer program may cause the third network node to perform operations of the method according to the above third aspect.

In this way, the present disclosure may be applicable to various MPLS based scenarios, such as L3VPN, L2VPN, etc. An MPLS label stack depth may be saved to mitigate stack depth challenges on the ingress node, especially for TE tunnel applications. Moreover, a light weighted solution may be achieved so that there is no impact on the transit nodes and the egress node, and on the ingress node, the switch chipset may support entropy with different ranges.

BRIEF DESCRIPTION OF THE DRAWINGS

The present disclosure may be best understood by way of example with reference to the following description and accompanying drawings that are used to illustrate embodiments of the present disclosure. In the drawings:

Fig. 1 is an exemplary schematic diagram illustrating a fundamental instance for entropy label handling according to some embodiments of the present disclosure;

Fig. 2 is an exemplary schematic diagram illustrating a comprehensive instance for entropy label handling according to some embodiments of the present disclosure;

Fig. 3 is a flow chart illustrating a method implemented on a first network node according to some embodiments of the present disclosure;

Fig. 4 is a flow chart illustrating a method implemented on a second network node according to some embodiments of the present disclosure;

Fig. 5 is a flow chart illustrating a method implemented on a third network node according to some embodiments of the present disclosure;

Fig. 6 is a block diagram illustrating a first network node according to some embodiments of the present disclosure;

Fig. 7 is another block diagram illustrating a first network node according to some embodiments of the present disclosure;

Fig. 8 is a block diagram illustrating a second network node according to some embodiments of the present disclosure;

Fig. 9 is another block diagram illustrating a second network node according to some embodiments of the present disclosure;

Fig. 10 is a block diagram illustrating a third network node according to some embodiments of the present disclosure; and

Fig. 11 is another block diagram illustrating a third network node according to some embodiments of the present disclosure.

DETAILED DESCRIPTION

The following detailed description describes methods and devices for entropy label handling in the MPLS communication networks. In the following detailed description, numerous specific details such as logic implementations, types and interrelationships of system components, etc. are set forth in order to provide a more thorough understanding of the present disclosure. It should be appreciated, however, by one skilled in the art that the present disclosure may be practiced without such specific details. In other instances, control structures, circuits and instruction sequences have not been shown in detail in order not to obscure the present disclosure. Those of ordinary skill in the art, with the included descriptions, will be able to implement appropriate functionality without undue experimentation.

References in the specification to “one embodiment”, “an embodiment”, “an example embodiment” etc. indicate that the embodiment described may include a particular feature, structure, or characteristic, but every embodiment may not necessarily include the particular feature, structure, or characteristic. Moreover, such phrases are not necessarily referring to the same embodiment. Further, when a particular feature, structure, or characteristic is described in connection with an embodiment, it is submitted that it is within the knowledge of one skilled in the art to affect such feature, structure, or characteristic in connection with other embodiments whether or not explicitly described.

Bracketed text and blocks with dashed borders (e.g., large dashes, small dashes, dot-dash, and dots) may be used herein to illustrate optional operations that add additional features to embodiments of the present disclosure. However, such notation should not be taken to mean that these are the only options or optional operations, and/or that blocks with solid borders are not optional in certain embodiments of the present disclosure.

In the following detailed description and claims, the terms “coupled” and “connected,” along with their derivatives, may be used. It should be understood that these terms are not intended as synonyms for each other.

"Coupled" is used to indicate that two or more elements, which may or may not be in direct physical or electrical contact with each other, cooperate or interact with each other. "Connected" is used to indicate the establishment of communication between two or more elements that are coupled with each other.

As used herein, the terms "first", "second" and so forth refer to different elements. The singular forms "a" and "an" are intended to include the plural forms as well, unless the context clearly indicates otherwise. The terms "comprises", "comprising", "has", "having", "includes" and/or "including" as used herein, specify the presence of stated features, elements, and/or components and the like, but do not preclude the presence or addition of one or more other features, elements, components and/or combinations thereof.

The term "terminal device" refers to any end device/client that can access a communication network and receive services therefrom. By way of example and not limitation, the terminal device may refer to a mobile terminal, a user equipment (UE), or other suitable devices. The UE may be, for example, a subscriber station, a portable subscriber station, a mobile station (MS) or an access terminal (AT). The terminal device may include, but not limited to, portable computers, image capture terminal devices such as digital cameras, gaming terminal devices, music storage and playback appliances, a mobile phone, a cellular phone, a smart phone, a tablet, a wearable device, a personal digital assistant (PDA), a vehicle, and the like. In the following description, the terms "terminal device", "client" and "UE" may be used interchangeably.

Fig. 1 is an exemplary schematic diagram illustrating a fundamental instance for entropy label handling according to some embodiments of the present disclosure.

As shown in Fig.1, PE2, which acts as a tunnel endpoint, may define an entropy block, including <min label ID, max label ID> or an equivalent one, such as <base label ID, range length>. Typically, 8-10bit width may be sufficient for the entropy block. As an example, a predefined entropy label ID range included in the entropy block may be configured by a user, e.g., as 5000-5999 shown in Fig. 1.

The entropy block may be advertised by PE2 through extension of existing MPLS protocols, such as ISIS(Intermediate System to Intermediate System)/OSPF(Open Shortest Path First) for Segment Routing, RSVP (Resource reSerVation Protocol) for RSVP-TE tunnel, BGP (Border Gateway Protocol) for BGP-LU (Labeled Unicast), etc.

Preferably, all nodes in the network may employ the same entropy block to simplify the deployment.

When the tunnel endpoint defines the entropy block as described above, it is called downstream assignment. In contrast to the downstream assignment case, such an entropy block may be upstream assigned by an ingress node PE1, which may further advertise this entropy block through the extension of the MPLS protocols. In the upstream assignment case (not shown), a tunnel label may also be upstream assigned by PE1.

In both cases, an entropy label ID may be calculated by PE1 into the entropy label ID range. As an example, the entropy label ID may be calculated from a 5-tuple for a packet, which may include e.g. a source IP, a destination IP, a protocol (e.g., TCP (Transmission Control Protocol) or UDP (User Datagram Protocol)), a source port and a destination port. For example, the entropy label ID may be calculated (hashed) from the 5-tuple to be 432. In order to fit in the range of 5000-5999, the entropy label ID is shifted to be 5432. As an example, the entropy label may be pushed by PE1 behind the tunnel label, e.g., 5432 behind 50 as shown in Fig. 1.

As shown in Fig. 1, the packet comprising at least the entropy label ID may traverse the transit nodes and arrive at the egress node PE2. In e.g. node P1, this entropy label ID is used for calculation of path selection for load balance.

At PE2, once the entropy block is defined (downstream assignment) or obtained (upstream assignment), corresponding Label Forwarding Information Base (LFIB) entries may be configured. For instance, if the entropy block is <5000, 5999>, then 1000 entries may be added to the

LFIB as “in label = 5xxx, action = pop”. When the packet is received by PE2, after tunnel termination, the entropy label may be popped by PE2.

Fig. 2 is an exemplary schematic diagram illustrating a comprehensive instance for entropy label handling according to some embodiments of the present disclosure.

The difference from the scenario shown in Fig. 1 is that the entropy label comprises more than one entropy label, e.g., a first entropy label with ID 3666 and a second entropy label with ID 5666, and more than one corresponding tunnel label, e.g., 30 and 50.

When a plurality of entropy labels are pushed into one packet due to an entropy reachability limitation issue, if entropy blocks are the same (not shown) for all tunnels encapsulated in the packet, one entropy label may be copied to all places in the MPLS label stack. If the entropy blocks are different, the entropy label should use the range according to the tunnel label ahead, as shown in Fig. 2.

In the case of different entropy blocks, not only PE2 but also P2 or many other nodes may act as tunnel endpoints. As shown in Fig. 2, since a node-label of P2 is 30, P2 may pop an entropy label within the range 3000-3999 and leave the remaining parts of the packet to subsequent tunnel endpoints. Then PE2, which has a node-label of 50, may pop an entropy label within the range 5000-5999.

Fig. 3 is a flow chart illustrating a method 300 implemented on a first network node in an MPLS communication network according to some embodiments of the present disclosure. As an example, operations of this flow chart may be performed by PE1 as shown in Fig. 1 or Fig. 2.

In one embodiment, the first network node may calculate one or more entropy label IDs into one or more predefined ranges of entropy label IDs (block 301). Then, the first network node may transmit a packet comprising at least the calculated one or more entropy label IDs to one or more second network nodes (block 302). As an example, the second network nodes may be tunnel endpoints, such as PE2 as shown in Fig. 1 or Fig. 2, or P2 as shown in Fig. 2. The predefined range may be associated with the first network node, e.g., associated with a node label of the first network node which acts as a tunnel endpoint. Alternatively, the predefined range may be associated with the tunnel of which the first network node acts as a tunnel endpoint.

As an optional example, the one or more predefined ranges may be determined by the first network node itself. As a further example, the one or more predefined ranges may be advertised by the first network node through MPLS protocols.

As an optional example, the one or more predefined ranges may be received from the one or more second network nodes respectively. As a further example, the one or more predefined ranges may be advertised by the one or more second network nodes respectively through MPLS protocols.

As another optional example, each of the predefined ranges may be indicated by a minimum entropy label ID and a maximum entropy label ID, or by a base entropy label ID and a length of this range.

As an additional optional example, the one or more predefined ranges may be configured, e.g., by users.

As an additional optional example, the one or more entropy label IDs may be pushed behind respective tunnel labels.

As an additional optional example, the one or more entropy label IDs may be calculated from a 5-tuple for the packet.

Fig. 4 is a flow chart illustrating a method 400 implemented on a second network node in an MPLS communication network according to some embodiments of the present disclosure. As an example, operations of this flow chart may be performed by tunnel endpoints, such as PE2 as shown in Fig. 1 or Fig. 2, or P2 as shown in Fig. 2.

In one embodiment, the second network node may receive a packet comprising at least one or more entropy label IDs from a first network node (block 401). As an example, the first network node may be PE1 as shown in Fig. 1 or Fig. 2. Then, the second network node may pop an entropy label having an entropy label ID of the entropy label IDs based on a predefined range of entropy label IDs (block 402). For instance, the entropy label ID of the entropy label may be within the predefined range. The predefined range may be associated with the second network node, e.g., associated with a node label of the second network node which acts as a tunnel endpoint. Alternatively, the predefined range may be associated with the tunnel of which the second network node acts as a tunnel endpoint.

As an optional example, the predefined range may be determined by the second network node. As a further example, the predefined range may be advertised by the second network node through an MPLS protocol prior to receipt of the packet.

As an optional example, the predefined range may be received from the first network node. As a further example, the predefined range may be advertised by the first network node through an MPLS protocol.

As another optional example, the predefined range may be indicated by a minimum entropy label ID and a maximum entropy label ID, or by a base entropy label ID and a length of this range.

As an additional optional example, the predefined range may be configured, e.g., by users.

Fig. 5 is a flow chart illustrating a method 500 implemented on a third network node in an MPLS communication network according to some embodiments of the present disclosure. As an example, operations of this flow chart may be performed by a node which specifies an entropy label ID range, whether an ingress node or an egress node.

In one embodiment, the third network node may determine a range of entropy label IDs (block 501). Then, the third network node may advertise the range to a plurality of network nodes (block 502).

As an example, if the third network node is the ingress node, the third network node may calculate an entropy label ID into the range (block 503), and transmit a packet comprising at least the calculated entropy label ID to the plurality of network nodes (block 504). As a further example, the entropy label ID may be pushed behind a tunnel label. As a still further example, the entropy label ID may be calculated from a 5-tuple for the packet.

As another example, if the third network node is the egress node, the third network node may receive a packet comprising at least one or more entropy label IDs from a fourth network node (block 505), and pop an entropy label having an entropy label ID of the entropy label IDs based on the range (block 506). For instance, the fourth network node may be an ingress node, and the entropy label ID of the entropy label may be within the predefined range. The predefined range may be associated with the third network node, e.g., associated with a node label of the third network node. Alternatively, the predefined range may be associated with the tunnel of which the third network node acts as the tunnel endpoint.

As an additional example, the range may be indicated by a minimum entropy label ID and a maximum entropy label ID, or by a base entropy label ID and a length of this range.

Fig. 6 is a block diagram illustrating a first network node 600 in an MPLS communication network according to some embodiments of the present disclosure. As an example, the first network node 600 may act as PE1 as shown in Fig. 1 or Fig. 2, but it is not limited thereto. It should be appreciated that the first network node 600 may be implemented using components other than those illustrated in Fig. 6.

With reference to Fig. 6, the first network node 600 may comprise at least a processor 601, a memory 602, a network interface 603 and a communication medium 604. The processor 601, the memory 602 and the network interface 603 may be communicatively coupled to each other via the communication medium 604.

The processor 601 may include one or more processing units. A processing unit may be a physical device or article of manufacture comprising one or more integrated circuits that read data and instructions from computer readable media, such as the memory 602, and selectively execute the instructions. In various embodiments, the processor 601 may be implemented in various ways. As an example, the processor 601 may be implemented as one or more processing cores. As another example, the processor 601 may comprise one or more separate microprocessors. In yet another example, the processor 601 may comprise an application-specific integrated circuit (ASIC) that provides specific functionality. In still another example, the processor 601 may provide specific functionality by using an ASIC and/or by executing computer-executable instructions.

The memory 602 may include one or more computer-usable or computer-readable storage medium capable of storing data and/or computer-executable instructions. It should be appreciated that the storage medium is preferably a non-transitory storage medium.

The network interface 603 may be a device or article of manufacture that enables the first network node 600 to send data to or receive data from other network nodes. In different embodiments, the network interface 603 may be implemented in different ways. As an example, the network interface 603 may be implemented as an Ethernet interface, a token-ring network interface, or another type of network interface.

The communication medium 604 may facilitate communication among the processor 601, the memory 602 and the network interface 603. The communication medium 604 may be implemented in various ways. For example, the communication medium 604 may comprise a Peripheral Component Interconnect (PCI) bus, a PCI Express bus, an accelerated graphics port (AGP) bus, a serial Advanced Technology Attachment (ATA) interconnect, a parallel ATA interconnect, a Fiber Channel interconnect, a USB bus, a Small Computing System Interface (SCSI) interface, or another type of communications medium.

In the example of Fig. 6, the instructions stored in the memory 602 may include those that, when executed by the processor 601, cause the first network node 600 to implement the method described with respect to Fig. 3.

Fig. 7 is another block diagram illustrating a first network node 700 in an MPLS communication network according to some embodiments of the present disclosure. As an example, the first network node 700 may act as

PE1 as shown in Fig. 1 or Fig. 2, but it is not limited thereto. It should be appreciated that the first network node 700 may be implemented using components other than those illustrated in Fig. 7.

With reference to Fig. 7, the first network node 700 may comprise at least a calculation unit 701 and a transmission unit 702.The calculation unit 701 may be adapted to perform at least the operation described in the block 301 of Fig. 3. The transmission unit 702 may be adapted to perform at least the operation described in the block 302 of Fig. 3.

Fig. 8 is a block diagram illustrating a second network node 800 in an MPLS communication network according to some embodiments of the present disclosure. As an example, the second network node 800 may act as PE2 as shown in Fig. 1 or Fig. 2, or P2 as shown in Fig. 2, but it is not limited thereto. It should be appreciated that the second network node 800 may be implemented using components other than those illustrated in Fig. 8.

With reference to Fig. 8, the second network node 800 may comprise at least a processor 801, a memory 802, a network interface 803 and a communication medium 804. The processor 801, the memory 802 and the network interface 803 are communicatively coupled to each other via the communication medium 804.

The processor 801, the memory 802, the network interface 803 and the communication medium 804 are structurally similar to the processor 601, the memory 602, the network interface 603 and the communication medium 604 respectively, and will not be described herein in detail.

In the example of Fig. 8, the instructions stored in the memory 802 may include those that, when executed by the processor 801, cause the second network node 800 to implement the method described with respect to Fig. 4.

Fig. 9 is another block diagram illustrating a second network node 900 in an MPLS communication network according to some embodiments of the present disclosure. As an example, the second network node 900 may act as a PE2 as shown in Fig. 1 or Fig. 2, or P2 as shown in Fig. 2, but it is not limited thereto. It should be appreciated that the second network node 900 may be implemented using components other than those illustrated in Fig. 9.

With reference to Fig. 9, the second network node 900 may comprise at least a receiving unit 901 and a popping unit 902. The receiving unit 901 may be adapted to perform at least the operation described in the block 401 of Fig. 4. The popping unit 902 may be adapted to perform at least the operation described in the block 402 of Fig. 4.

Fig. 10 is a block diagram illustrating a third network node 1000 in an MPLS communication network according to some embodiments of the present disclosure. As an example, the third network node 1000 may act as a node which specifies an entropy label ID range, whether an ingress node or an egress node, but it is not limited thereto. It should be appreciated that the third network node 1000 may be implemented using components other than those illustrated in Fig. 10.

With reference to Fig. 10, the third network node 1000 may comprise at least a processor 1001, a memory 1002, a network interface 1003 and a communication medium 1004. The processor 1001, the memory 1002 and the network interface 1003 are communicatively coupled to each other via the communication medium 1004.

The processor 1001, the memory 1002, the network interface 1003 and the communication medium 1004 are structurally similar to the processor 601 or 801, the memory 602 or 802, the network interface 603 or 803 and the communication medium 604 or 804 respectively, and will not be described herein in detail.

In the example of Fig. 10, the instructions stored in the memory 1002 may include those that, when executed by the processor 1001, cause the third network node 1000 to implement the method described with respect to Fig. 5.

Fig. 11 is another block diagram illustrating a third network node 1100 in an MPLS communication network according to some embodiments of the present disclosure. As an example, the third network node 1100 may act as a node which specifies an entropy label ID range, whether an ingress node or an egress node, but it is not limited thereto. It should be appreciated that the third network node 1100 may be implemented using components other than those illustrated in Fig. 11.

With reference to Fig. 11, the third network node 1100 may comprise at least a determination unit 1101 and an advertising unit 1102. The determination unit 1101 may be adapted to perform at least the operation described in the block 501 of Fig. 5. The advertising unit 1102 may be adapted to perform at least the operation described in the block 502 of Fig. 5.

In an optional example, the third network node 1100 may further comprise at least a calculation unit 1103, a transmission unit 1104, a receiving unit 1105 and a popping unit 1106. The calculation unit 1103 may be adapted to perform at least the operation described in the block 503 of Fig. 5. The transmission unit 1104 may be adapted to perform at least the operation described in the block 504 of Fig. 5. The receiving unit 1105 may be adapted to perform at least the operation described in the block 505 of Fig. 5. The popping unit 1106 may be adapted to perform at least the operation described in the block 506 of Fig. 5.

The units 701-702, 901-902 and 1101-1106 are illustrated as separate units in Figs. 7, 9 and 11. However, this is merely to indicate that the functionality is separated. The units may be provided as separate elements. However, other arrangements are possible, e.g., some of them may be combined as one unit in each figure. Any combination of the units may be implemented in any combination of software, hardware, and/or firmware in any suitable location.

The units shown in Figs. 7, 9 and 11 may constitute machine-executable instructions embodied within a machine, e.g., readable medium, which when executed by a machine will cause the machine to perform the operations described. Besides, any of these units may be implemented as hardware, such as an application specific integrated circuit (ASIC), Digital Signal Processor (DSP), Field Programmable Gate Array

(FPGA) or the like.

Moreover, it should be appreciated that the arrangements described herein are set forth only as examples. Other arrangements may be used in addition to or instead of those shown, and some units may be omitted altogether. Functionality and cooperation of these units are correspondingly described in more detail with reference to Figs. 3-5.

Some portions of the foregoing detailed description have been presented in terms of algorithms and symbolic representations of transactions on data bits within a computer memory. These algorithmic descriptions and representations are ways used by those skilled in the data processing arts to most effectively convey the substance of their work to others skilled in the art. An algorithm is here, and generally, conceived to be a self-consistent sequence of transactions leading to a desired result. The transactions are those requiring physical manipulations of physical quantities. Usually, though not necessarily, these quantities take the form of electrical or magnetic signals capable of being stored, transferred, combined, compared, and otherwise manipulated. It has proven convenient at times, principally for reasons of common usage, to refer to these signals as bits, values, elements, symbols, characters, terms, numbers, or the like.

It should be appreciated, however, that all of these and similar terms are to be associated with the appropriate physical quantities and are merely convenient labels applied to these quantities. Unless specifically stated otherwise as apparent from the above discussion, it is appreciated that throughout the description, discussions utilizing terms such as "processing" or "computing" or "calculating" or "determining" or "displaying" or the like, refer to actions and processes of a computer system, or a similar electronic computing device, that manipulates and transforms data represented as physical (electronic) quantities within the computer system's registers and memories into other data similarly represented as physical quantities within the computer system memories or registers or other such information storage, transmission or display devices.

The algorithms and displays presented herein are not inherently related to any particular computer or other apparatus. Various general-purpose systems may be used with programs in accordance with the teachings herein, or it may prove convenient to construct more specialized apparatus to perform the required method transactions. The required structure for a variety of these systems will appear from the description above. In addition, embodiments of the present disclosure are not described with reference to any particular programming language. It should be appreciated that a variety of programming languages may be used to implement the teachings of embodiments of the present disclosure as described herein.

An embodiment of the present disclosure may be an article of manufacture in which a non-transitory machine-readable medium (such as microelectronic memory) has stored thereon instructions (e.g., computer code) which program one or more data processing components (generically referred to here as a "processor") to perform the operations described above. In other embodiments, some of these operations might be performed by specific hardware components that contain hardwired logic (e.g., dedicated digital filter blocks and state machines). Those operations might alternatively be performed by any combination of programmed data processing components and fixed hardwired circuit components.

In the foregoing detailed description, embodiments of the present disclosure have been described with reference to specific exemplary embodiments thereof. It will be evident that various modifications may be made thereto without departing from the spirit and scope of the present disclosure as set forth in the following claims. The specification and drawings are, accordingly, to be regarded in an illustrative sense rather than a restrictive sense.

Throughout the description, some embodiments of the present disclosure have been presented through flow diagrams. It should be appreciated that the order of transactions and transactions described in these flow diagrams are only intended for illustrative purposes and not intended as a limitation of the present disclosure. One having ordinary skill in the art would recognize that variations can be made to the flow diagrams without departing from the spirit and scope of the present disclosure as set forth in the following claims.

CLAIMS

1. A method (300) implemented by a first network node in a communication network, comprising:

calculating (301) one or more entropy label identifiers into one or more predefined ranges of entropy label identifiers; and

transmitting (302) a packet comprising at least the calculated one or more entropy label identifiers to one or more second network nodes.

2. The method of Claim 1, wherein the one or more predefined ranges are determined by the first network node.

3. The method of Claim 2, wherein the one or more predefined ranges are advertised by the first network node through Multi-Protocol Label Switching protocols.

4. The method of Claim 1, wherein the one or more predefined ranges are received from the one or more second network nodes respectively.

5. The method of Claim 4, wherein the one or more predefined ranges are advertised by the one or more second network nodes respectively through Multi-Protocol Label Switching protocols.

6. The method of any of Claims 1-5, wherein each of the predefined ranges is indicated by a minimum entropy label identifier and a maximum entropy label identifier, or by a base entropy label identifier and a length of this range.

7. The method of Claim 1, wherein the one or more predefined ranges are configured.

8. The method of any of Claims 1-7, wherein the one or more entropy label identifiers are pushed behind respective tunnel labels.

9. The method of any of Claims 1-8, wherein the one or more entropy label identifiers are calculated from a 5-tuple for the packet.

10. A method (400) implemented by a second network node in a communication network, comprising:

receiving (401) a packet comprising at least one or more entropy label identifiers from a first network node; and popping (402) an entropy label having an entropy label identifier of the entropy label identifiers based on a predefined range of entropy label identifiers.

11. The method of Claim 10, wherein the predefined range is determined by the second network node.

12. The method of Claim 11, wherein the predefined range is advertised by the second network node through a Multi-Protocol Label Switching protocol prior to receipt of the packet.

13. The method of Claim 10, wherein the predefined range is received from the first network node.

14. The method of Claim 13, wherein the predefined range is advertised by the first network node through a Multi-Protocol Label Switching protocol.

15. The method of any of Claims 10-14, wherein the predefined range is indicated by a minimum entropy label identifier and a maximum entropy label identifier, or by a base entropy label identifier and a length of this range.

16. The method of claim 10, wherein the predefined range is configured.

17. A method (500) implemented by a third network node in a communication network, comprising:

determining (501) a range of entropy label identifiers; and

advertising (502) the range to a plurality of network nodes.

18. The method of Claim 17, further comprising:

calculating (503) an entropy label identifier into the range; and

transmitting (504) a packet comprising at least the calculated entropy label identifier to the plurality of network nodes.

19. The method of Claim 18, wherein the entropy label identifier is pushed behind a tunnel label.

20. The method of Claim 18 or 19, wherein the entropy label identifier is calculated from a 5-tuple for the packet.

21. The method of Claim 17, further comprising:

receiving (505) a packet comprising at least one or more entropy label identifiers from a fourth network node of the plurality of network nodes; and

popping (506) an entropy label having an entropy label identifier of the entropy label identifiers based on the range.

22. The method of any of Claims 17-21, wherein the range is indicated by a minimum entropy label identifier and a maximum entropy label identifier, or by a base entropy label identifier and a length of this range.

23. A first network node (600) in a communication network, comprising:

a processor (601); and

a memory (602) communicatively coupled to the processor and adapted to store instructions which, when executed by the processor, cause the first network node to perform operations of the method of any of Claims 1 to 9.

24. A second network node (800) in a communication network, comprising:

a processor (801); and

a memory (802) communicatively coupled to the processor and adapted to store instructions which, when executed by the processor, cause the second network node to perform operations of the method of any of Claims 10 to 16.

25. A third network node (1000) in a communication network, comprising:

a processor (1001); and

a memory (1002) communicatively coupled to the processor and adapted to store instructions which, when executed by the processor, cause the third network node to perform operations of the method of any of Claims 17 to 22.

26. A non-transitory computer readable medium having a computer program stored thereon which, when executed by a set of one or more

processors of a first network node in a communication network, causes the first network node to perform operations of the method of any of Claims 1 to 9.

27. A non-transitory computer readable medium having a computer program stored thereon which, when executed by a set of one or more processors of a second network node in a communication network, causes the second network node to perform operations of the method of any of Claims 10 to 16.

28. A non-transitory computer readable medium having a computer program stored thereon which, when executed by a set of one or more processors of a third network node in a communication network, causes the third network node to perform operations of the method of any of Claims 17 to 22.

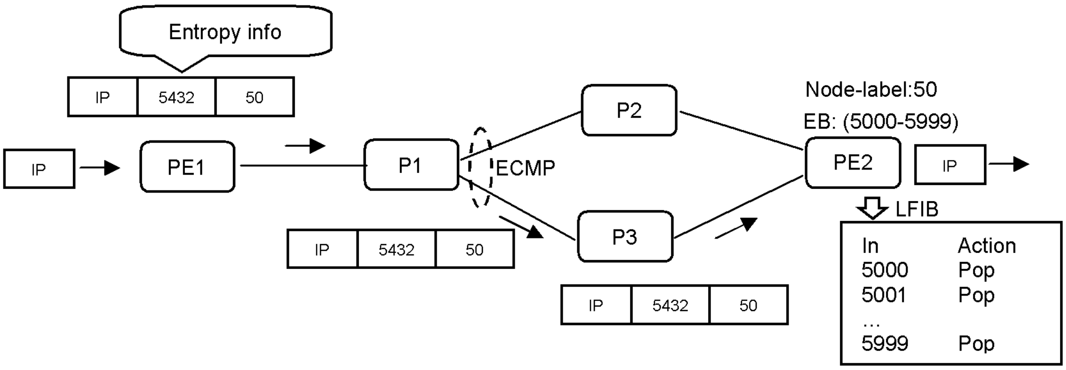

Fig.1

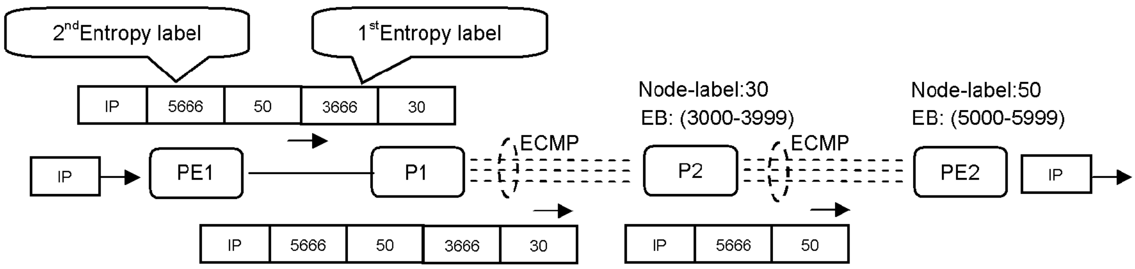

Fig.2

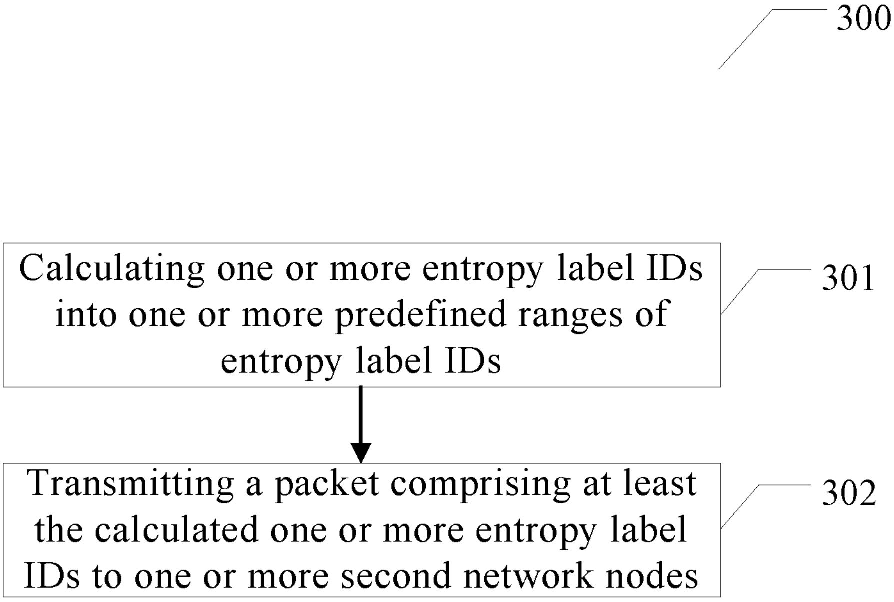

Fig. 3

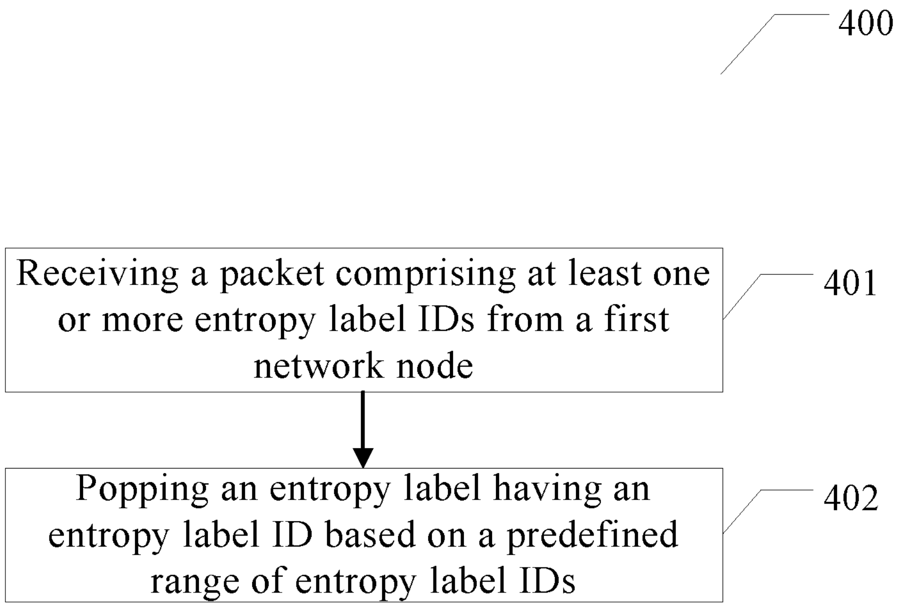
400
Receiving a packet comprising at least one or more entropy label IDs from a first network node
401
Popping an entropy label having an entropy label ID based on a predefined range of entropy label IDs
402

Fig. 4

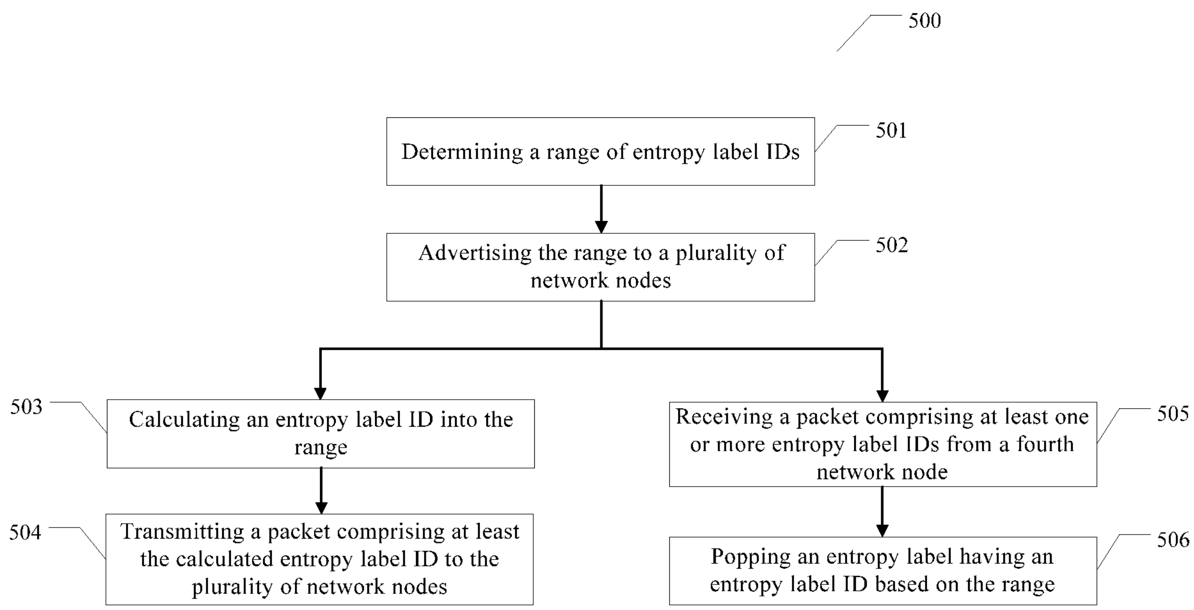
500
Determining a range of entropy label IDs
501
Advertising the range to a plurality of network nodes
502
503
Calculating an entropy label ID into the range
504
Transmitting a packet comprising at least the calculated entropy label ID to the plurality of network nodes
Receiving a packet comprising at least one or more entropy label IDs from a fourth network node
505
Popping an entropy label having an entropy label ID based on the range
506

Fig. 5

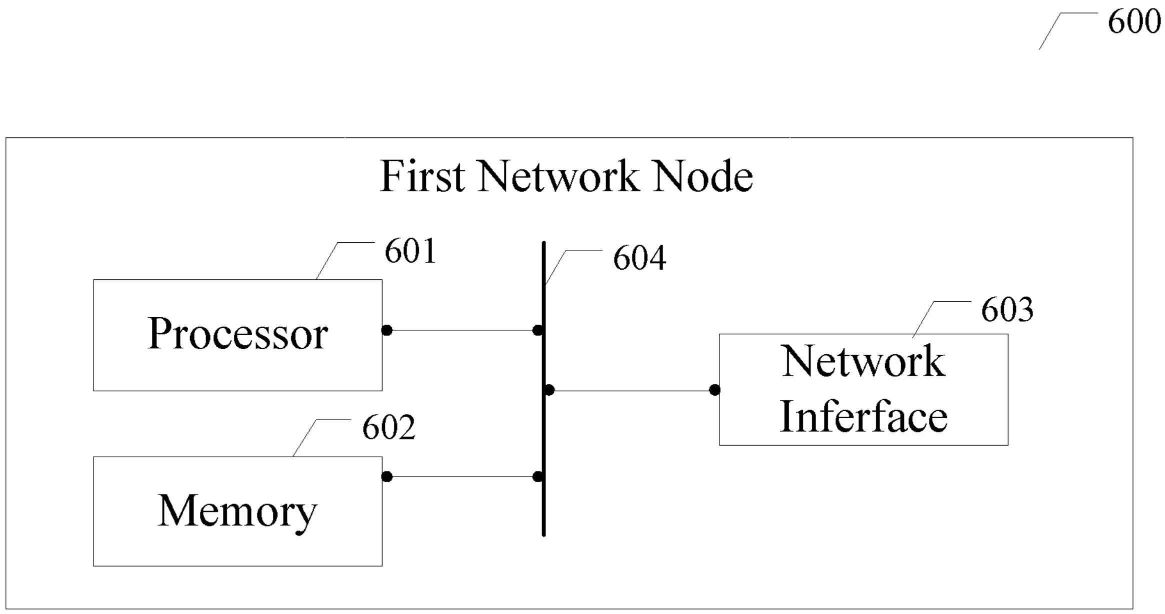
600
First Network Node
601
Processor
604
603
Network
Inferface
602
Memory

Fig. 6

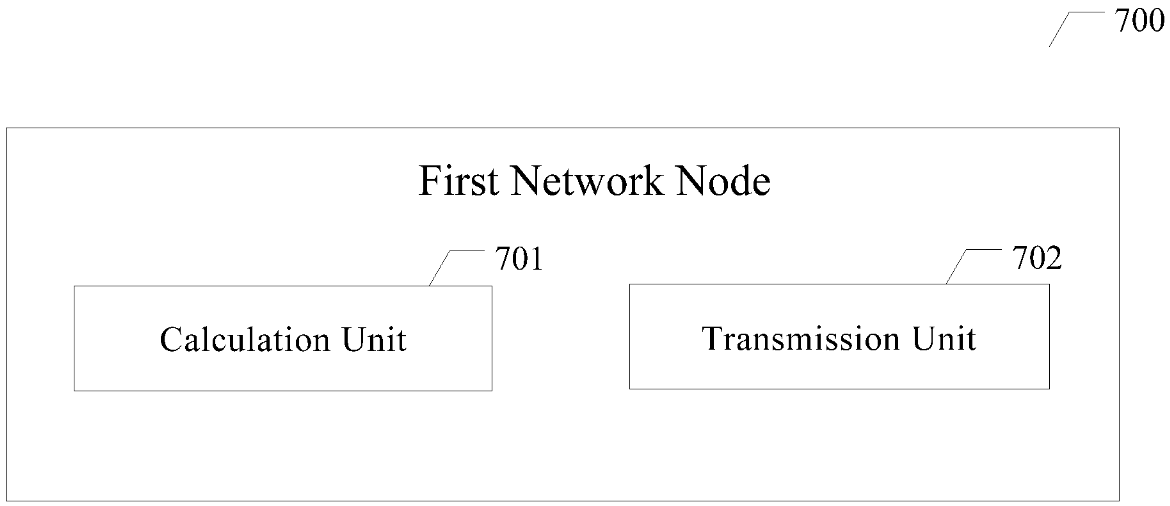
700
First Network Node
701
702
Calculation Unit
Transmission Unit

Fig. 7

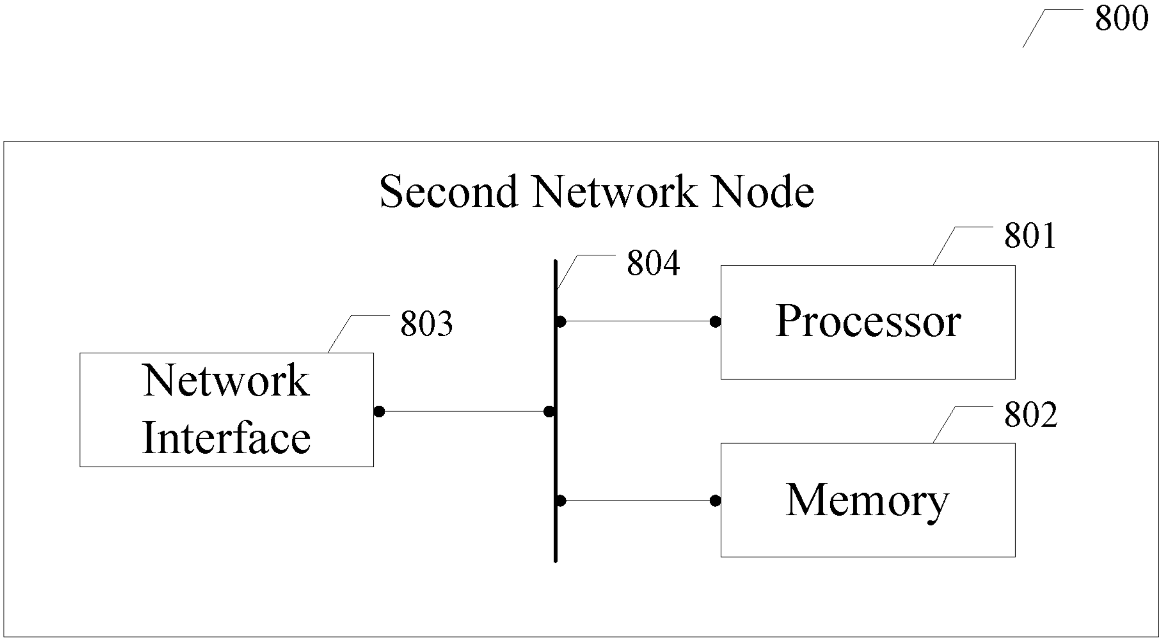
800
Second Network Node
801
804
Processor
803
Network Interface
802
Memory

Fig. 8

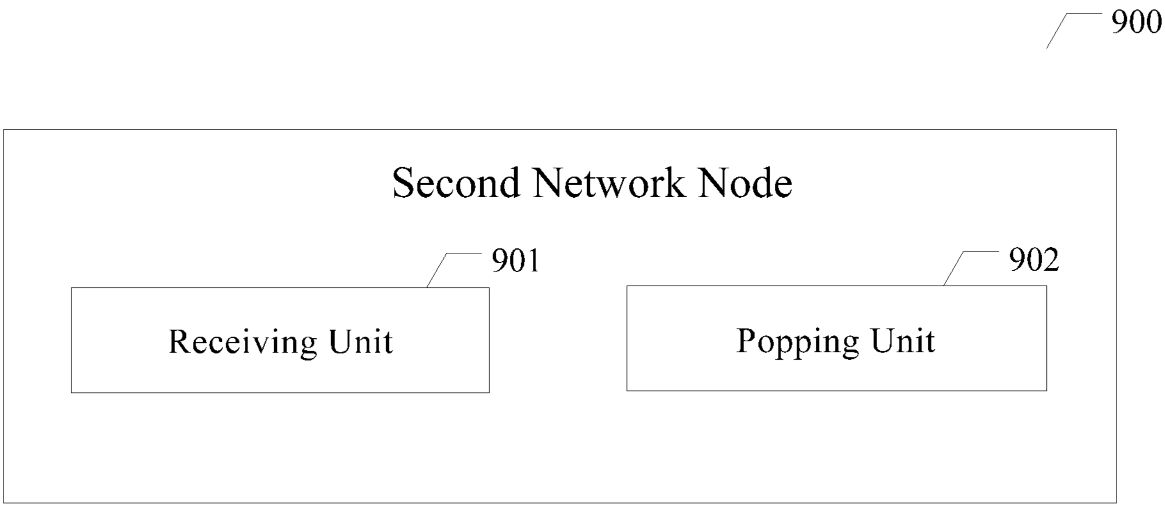
900
Second Network Node
901
902
Receiving Unit
Popping Unit

Fig. 9

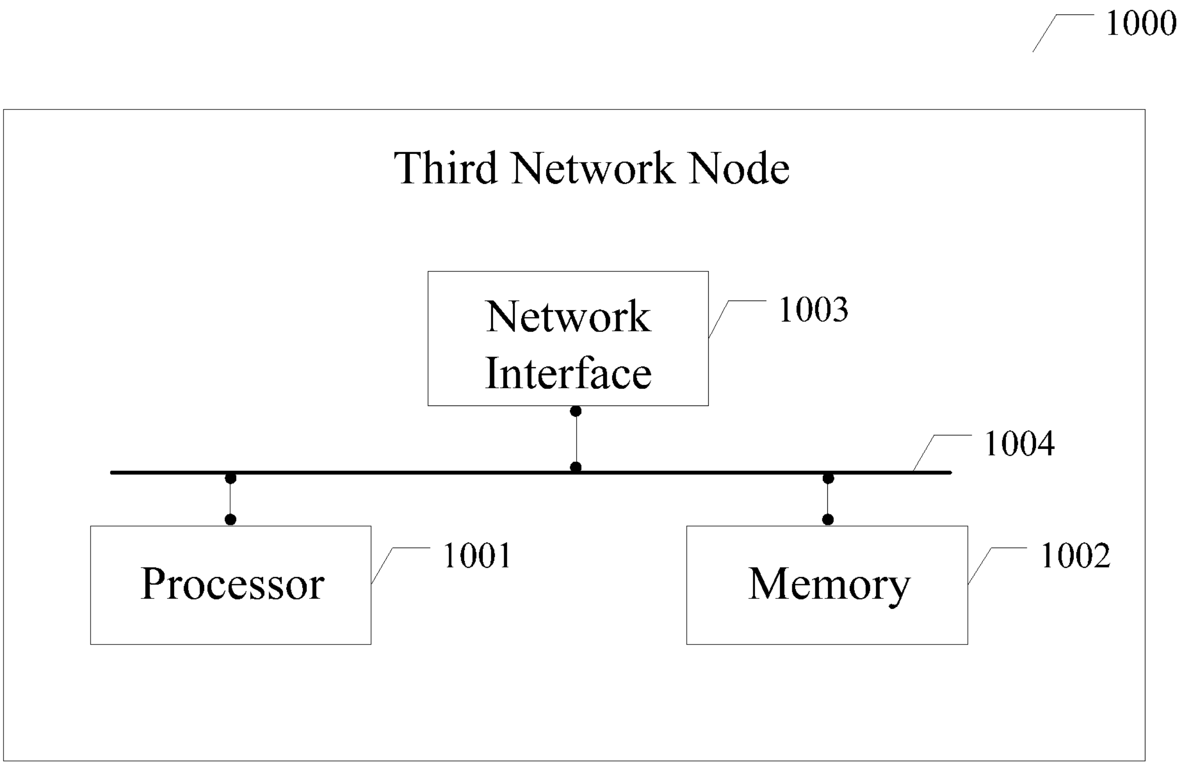
1000
Third Network Node
Network Interface
1003
1004
Processor
1001
Memory
1002

Fig. 10

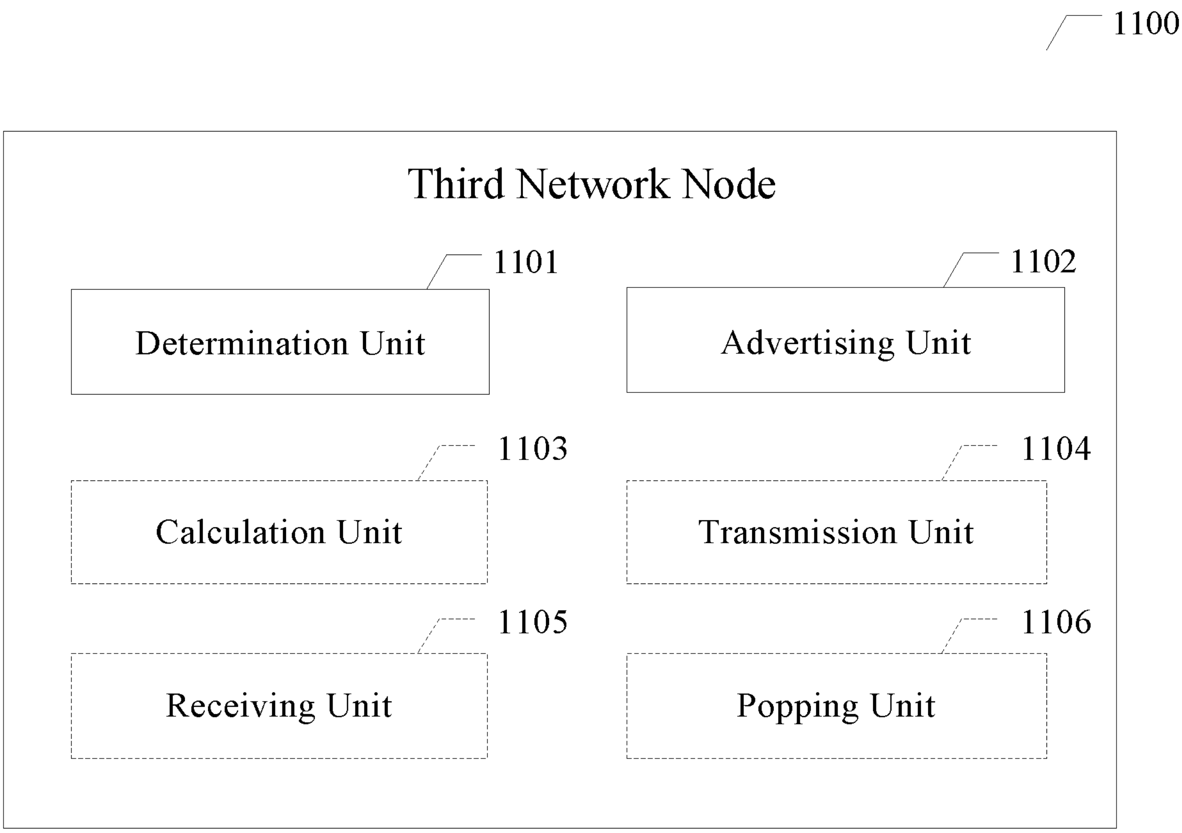
1100
Third Network Node
1101
Determination Unit
1102
Advertising Unit
1103
Calculation Unit
1104
Transmission Unit
1105
Receiving Unit
1106
Popping Unit

Fig. 11

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2019/090673

A. CLASSIFICATION OF SUBJECT MATTER

H04L 12/723(2013.01)i

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04L

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

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

CNPAT,CNKI,WPI,EPODOC,3GPP： multi protocol label switching MPLS entropy label identifiers predefined range EL ELI ID

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2015030020 A1 (TELEFONAKTIEBOLAGET L M ERICSSON PUBL) 29 January 2015 (2015-01-29) description paragraphs 8-22,claim 1	1-28
A	US 2015029849 A1 (CISCO TECHNOLOGY INC.) 29 January 2015 (2015-01-29) the whole document	1-28
A	EP 2978176 A1 (HUAWEI TECHNOLOGIES CO., LTD.) 27 January 2016 (2016-01-27) the whole document	1-28
A	CN 104639470 A (ZTE CORPORATION) 20 May 2015 (2015-05-20) the whole document	1-28
A	CN 104468391 A (CENTEC NETWORK SUZHOU CO., LTD.) 25 March 2015 (2015-03-25) the whole document	1-28

☐ 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
24 February 2020	**10 March 2020**

Name and mailing address of the ISA/CN

National Intellectual Property Administration, PRC
6, Xitucheng Rd., Jimen Bridge, Haidian District, Beijing 100088
China

Facsimile No. **(86-10)62019451**

Authorized officer

FANG,Ting

Telephone No. **86-010-53961654**

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2019/090673

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
US	2015030020	A1	29 January 2015	WO	2015015359	A1	05 February 2015
				CN	105745880	A	06 July 2016
				EP	3028412	A1	08 June 2016
US	2015029849	A1	29 January 2015		None		
EP	2978176	A1	27 January 2016	CN	108632146	A	09 October 2018
				US	2018183711	A1	28 June 2018
				CN	104243311	A	24 December 2014
				WO	2014194751	A1	11 December 2014
				US	2016087884	A1	24 March 2016
				US	2019297009	A1	26 September 2019
CN	104639470	A	20 May 2015	WO	2015070601	A1	21 May 2015
				EP	3070895	A1	21 September 2016
				US	2016261492	A1	08 September 2016
				EP	3070895	A4	16 November 2016
CN	104468391	A	25 March 2015		None		