

(51) International Patent Classification:
H04L 12/26 (2006.01)(21) International Application Number:
PCT/EP2016/076154(22) International Filing Date:
28 October 2016 (28.10.2016)

(25) Filing Language: English

(26) Publication Language: English

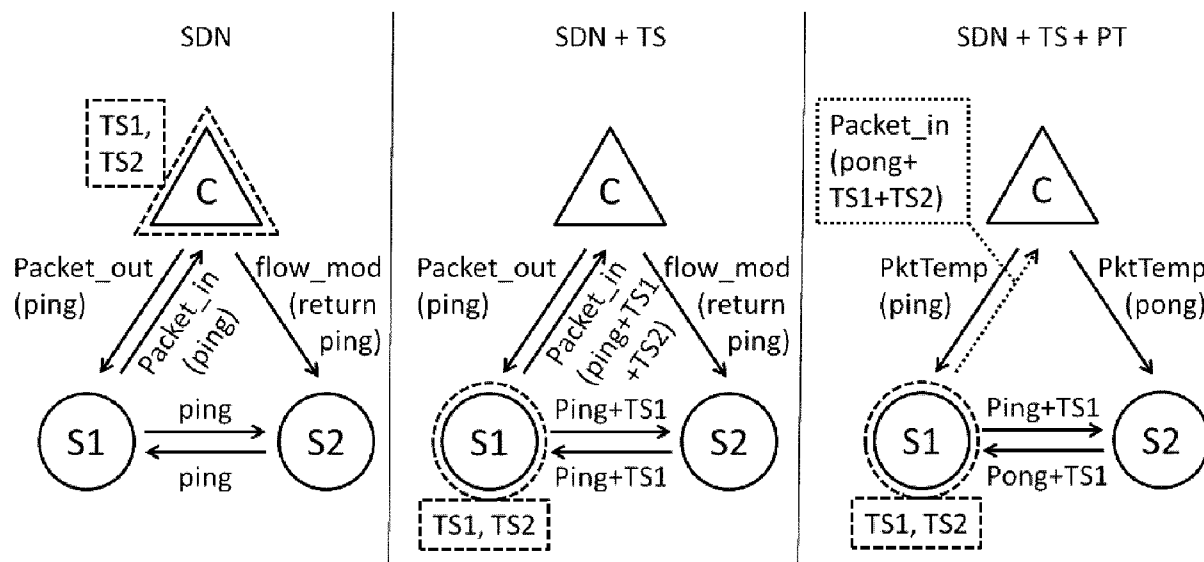
(71) Applicant: NEC LABORATORIES EUROPE GMBH
[DE/DE]; Kurfuersten-Anlage 36, 69115 Heidelberg (DE).(72) Inventors: SCHNEIDER, Fabian; Langer Anger 71,
69115 Heidelberg (DE). MATSIUK, Anton; Bautzen-
strasse 1, Whg-Nr. 12.01, 69115 Heidelberg (DE). GK-
OUNIS, Dimitrios; Franz-Knauff-Strasse 17, 69115 Hei-
delberg (DE).(74) Agent: ULLRICH & NAUMANN; Schneidmuehlstrasse
21, 69115 Heidelberg (DE).(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, JP, KE, KG, 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: METHOD FOR SUPPORTING NETWORK MONITORING IN A SOFTWARE DEFINED NETWORK AND COR-
RESPONDING SOFTWARE DEFINED NETWORK

Fig. 2

(57) Abstract: A method for supporting network monitoring in a software defined network, wherein said software defined network comprises forwarding elements and a software defined network controller for controlling the forwarding elements, wherein a source forwarding element and a target forwarding element are employed to perform a measurement task for network monitoring such that said source forwarding element and said target forwarding element function as monitoring probes, and wherein said software defined network controller configures said monitoring probes such that one or more probe packets for performing the measurement task are generated by at least one of said monitoring probes. Furthermore, a corresponding software defined network is disclosed.



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))*

METHOD FOR SUPPORTING NETWORK MONITORING IN A SOFTWARE DEFINED NETWORK AND CORRESPONDING SOFTWARE DEFINED NETWORK

5 The present invention relates to a method for supporting network monitoring in a software defined network.

Furthermore, the present invention relates to a software defined network for supporting network monitoring.

10

In recent years network monitoring is a crucial task for Internet service providers (ISPs) to ensure their network is working as expected and that the service level agreements (SLAs) with their customers are fulfilled.

15

According to known methods and systems for network monitoring, the network monitoring is deployed by using dedicated monitoring probes that generate and receive test traffic. The term test traffic may also be designated as probe traffic, wherein probe packets are generated and transmitted in the network for testing and determining network parameters. Fig. 1 illustrates an application scenario of a typical network monitoring probe deployment. The goal is to perform full end-to-end path monitoring to ensure that the whole network is working as intended. The monitoring probes are located close to the customer edge inside the points of presence of the ISPs. The monitoring probes are attached to the network like regular customers would be connected, i.e. by using the access technologies such as DSL, Cable, PON, FttH, etc. that are present at the point of presence (PoP).

20

25

30

While the network is managed by the network management system (NMS) the monitoring probes are orchestrated by a monitoring manager, which may receive some information from the Operations Support System and Business Support System (OSS/BSS system). Typically such information may include guarantees given to the customers according to the service level agreements (SLAs). Furthermore, this information may include information about the topology of the network, especially with respect to the location of the monitoring probes.

The reasons for placing monitoring probes at the customer edge of the network are

- (i) that this is the place where the customer would also connect, and
- (ii) that this does not interfere with the management of the network.

These reasons coincide with the goal of end-to-end probing.

In the field of network monitoring, there are diverse network probing mechanisms which may include the following:

- **Round-trip time (RTT) measurement:** Usually utilizing a mechanism similar to the ping tool. A monitoring probe A sends a packet to a monitoring probe B, which sends back a reply to the monitoring probe A, and then the monitoring probe A can calculate the RTT by subtracting the time-stamp of the reply from the time-stamp of the initial probe packet.
- **One-way delay (OWD) measurement:** Being similar to round-trip time measurements, however this requires synchronized and accurate clocks in both monitoring probes.
- **Jitter:** For jitter or packet delay variation measurements typically equally spaced sequences of packets are generated at the source monitoring probe and their inter-arrival time variations are recorded at the target monitoring probe.
- **Throughput measurement:** A source monitoring probe generates as much traffic as possible and tries to saturate the available capacity of the network toward the target monitoring probe, the target monitoring probe will then determine how much of the generated traffic was received.
- **Available Bandwidth (AvBW) estimation:** This is similar to a throughput measurement, but the difference is that here the probe traffic is not filling up the link capacity. This is achieved by sending only probe packet pairs or

- 3 -

probe packet trains to the target monitoring probe and then inferring the bandwidth from the differences of the inter-packet timings when sending and receiving the probe traffic.

5 Conventional methods and systems for network monitoring have the problem that relevant probing techniques require careful crafting of the probe traffic. However, this is not possible in network elements, neither in legacy networks nor in typical software defined networks (SDNs) such as networks controlled by OpenFlow. With regard to OpenFlow it is exemplarily referred to non-patent literature of Open
10 Networking Foundation: "OpenFlow Switch Specification", Version 1.3.4 (March 27, 2014 – TS - 019) retrievable at "<https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-switch-v1.3.4.pdf>". Hence, the known network monitoring systems are implemented with dedicated monitoring probe devices and a probing orchestrator, which is
15 cumbersome and inflexible.

In view of the above it is therefore an objective of the present invention to improve and further develop a method of the initially described type for supporting network monitoring in a software defined network in such a way that an improved and a
20 more flexible network monitoring is achieved.

In accordance with the invention, the aforementioned objective is accomplished by a method for supporting network monitoring in a software defined network, wherein said software defined network comprises forwarding elements and a
25 software defined network controller for controlling the forwarding elements, wherein a source forwarding element and a target forwarding element are employed to perform a measurement task for network monitoring such that said source forwarding element and said target forwarding element function as monitoring probes, and wherein said software defined network controller
30 configures said monitoring probes such that one or more probe packets for performing the measurement task are generated by at least one of said monitoring probes.

Furthermore, the aforementioned objective is accomplished by a software defined network for supporting network monitoring, the network comprising forwarding elements and a software defined network controller for controlling the forwarding elements, wherein said forwarding elements include a source forwarding element and a target forwarding element that are programmable to perform a measurement task for network monitoring such that said source forwarding element and said target forwarding element function as monitoring probes, and wherein said software defined network controller is operable to configure said monitoring probes such that one or more probe packets for performing the measurement task are generated by at least one of said monitoring probes.

According to the invention it has first been recognized that in the context of network monitoring in software defined networks (SDNs) an enormous improvement, in particular with regard to the flexibility of the monitoring, can be provided by employing forwarding elements, e.g. switches, of the software defined network (SDN) as monitoring probes. Specifically, a source forwarding element and a target forwarding element are employed to perform a measurement task for network monitoring, such that the source forwarding element and the target forwarding element function as monitoring probes. Further, it has been recognized that more flexible and more accurate measurements can be achieved when the measurement task is performed in a more autonomous way by the forwarding elements functioning as monitoring probes. To this extent, according to the invention, the software defined network controller configures the monitoring probes such that one or more probe packets for performing the measurement task are generated directly on at least one of the monitoring probes. By doing this, the one or more probe packets are sent out by at least one of the monitoring probes. Thus, the measurement task can be initiated by the forwarding element, in particular by the source forwarding element. Hence, there is no need for dedicated monitoring probe devices, as the network elements themselves can generate the probe packets. For example, it might be conceivable that any forwarding element of a software defined network or at least the most of the forwarding elements of a software defined network is able to be turned into a monitoring probe, which in turn enables a finer granular network probing and does not require at first to deploy physically a dedicated monitoring probe device in the network. Hence, a method

and a software defined network according to the present invention provide an improved and more flexible network monitoring.

5 The term “monitoring probe” may refer in particular in the claims, preferably in the description, to an entity such as a program or other device, preferably located at a key juncture, in a network for the purpose of monitoring or collecting data about network activity. Furthermore, for example, a monitoring probe may perform an action and/or send out an object, such as a probe packet, used for the purpose of learning something about the state of the network. For example, an empty
10 message can be sent simply to see whether the destination actually exists. Thus, according to the present invention, a source forwarding element and/or a target forwarding element are configured in such a way that they may function as monitoring probes for performing a measurement task.

15 The term “probe packet” may refer in particular in the claims, preferably in the description, to a packet used in an active measurement experiment/task to collect knowledge on a given network parameter of interest.

20 According to embodiments of the invention, it may be provided that generation and/or handling of probe packets, which are used for performing the measuring task, is/are temporally decoupled from the software defined network controller. Thus, one or more measurement tasks – and thus a measurement campaign – can be programmed in a flexible and efficient way on network elements, e.g. on switches in the context of OpenFlow.

25 According to embodiments of the invention the one or more probe packets may include a ping probe packet for performing the measuring task, wherein the ping probe packet is generated by the source forwarding element. Thus, there is no need to deploy a dedicated monitoring probe device for generating a ping probe packet that is to be used for performing a measurement task. Rather, it is possible
30 that ping probe packets can be generated directly on a forwarding element of the network. Furthermore, timing inaccuracies can be reduced, in particular compared to an approach where the software defined network controller would perform probe packet generation and/or time stamping. Further, interactions on the control

channel, i.e. interactions between the software defined network controller and the source forwarding element can be kept to a minimum.

5 The term “ping probe packet” may refer in particular in the claims, preferably in the description, to an initial probe packet that is sent from the source forwarding element functioning as source monitoring probe to the target forwarding element functioning as target monitoring probe.

10 According to embodiments of the invention the one or more probe packets may include a pong probe packet for performing the measurement task, wherein the pong probe packet is generated by the target forwarding element. Thus, at the target forwarding element, the pong probe packet may be generated as new probe packet that is used for the return path of the measurement task. For example, in the context of a round-trip measurement procedure, the target forwarding element
15 needs to have the capability to modify the header fields of a received ping probe packet in which e.g. the time stamp can be stored. Furthermore, returning the ping probe packet back to the source forwarding element will require changes to the packet header of the ping probe packet, especially when there is no direct connection between the source forwarding element and the target forwarding
20 element and the ping probe packet that is to be returned will need to traverse other network elements. Hence, in this case, the packets need to be routed appropriately and might involve also changes on the IP addresses, etc. Thus, by generating the pong probe packet, which is a new probe packet and which is used as reply packet, this requirement can be enabled with minimal changes to the
25 packet header and facilitates routing.

30 The term “pong probe packet” may refer in particular in the claims, preferably in the description, to a reply probe packet that is sent out from the target forwarding element functioning as target monitoring probe. For example, the pong probe packet may be sent as reply to the source forwarding element, e.g. in the context of a round-trip time measurement as measuring task. Furthermore, it may be provided that the pong probe packet is transmitted directly to the software defined network controller, e.g. in the context of a one-way delay measurement as measurement task.

According to embodiments of the invention the software defined network controller may configure the monitoring probes to generate the one or more probe packets by creating probe packet template information that is provided to the monitoring probes, wherein the probe packet template information includes one or more probe packet templates and/or probe packet template handling instructions. Thus, the software defined network controller can configure the monitoring probes for probe packet generation. For example, the concept of Packet Template (PT) as exemplarily specified in the document WO 2015/000517 A1 can be used in an advantageous way for probe packet generation on forwarding elements. Document WO 2015/000517 A1 describes a mechanism for generating packets directly on forwarding elements of a software defined network. This mechanism for programming in-switch generation may be used according to an embodiment of the present invention in order to improve, enrich and simplify network monitoring, namely by generating probe packets using the Packet Template (PT) mechanism. Advantageously, the most network elements from the customer edge all the way up to the backbone gateway are SDN-enabled and in-switch packet generation enabled. Hence, by using one or more probe packet templates the load on the control channel between the forwarding elements and the software defined network controller is significantly reduced.

According to embodiments of the invention the monitoring probes may be triggered to generate the one or more probe packets based on the provided packet template information. This enables the generation of probe packets in a way that is temporally decoupled from the supervision of the software defined network controller.

According to embodiments of the invention the triggering instructions may be provided to the monitoring probes by including them into the packet template information. Thus, it is enabled in an easy and simple way to provide triggering instructions for the generation of probe packets, e.g. ping probe packets and/or pong probe packets, by the software defined network controller.

According to embodiments of the invention the triggering instructions may include configuring a timer as a trigger for generating the one more probe packets.

Advantageously, the triggering instructions may include configuring a timer at the source forwarding element, wherein the timer triggers the generation of one or more ping probe packets that are sent to the target forwarding element. For example, the timer may be configured such that a predetermined and/or periodic event is set as trigger. Hence, it may be provided that a timer is employed to trigger the generation of probe packets. Advantageously, hierarchical timers may be employed. E.g. for packet trains large scale is used to trigger the train, and small scale is used to trigger the individual probe packets in the train.

According to embodiments of the invention the triggering instructions may include configuring a reception of a predetermined probe packet as trigger for generating a new probe packet, in particular a pong probe packet. For example, the triggering instructions may include that, at the target forwarding element, a reception of a predetermined ping probe packet is configured as trigger for generating a new pong probe packet. This enables in an easy way a triggering of a generation of a pong probe packet based on a received ping probe packet. Thus, the measurement task can be suitably programmed, e.g. for performing a round-trip time measurement.

According to embodiments of the invention, it may be provided that local information of the monitoring probes is employed to directly trigger the one or more probe packets for performing the measurement task. This enables to perform the measurement task in a flexible and efficient way. Thus, a predetermined measurement task and/or probing may be directly triggered dependent on a predetermined condition and/or status of the monitoring probe, e.g. of the source forwarding element.

According to embodiments of the invention time stamping may be performed directly on at least one of the monitoring probes. Thus, time stamps may be recorded and/or added to probe packets directly on and by the forwarding elements of the software defined network. Hence, the measurement task is more accurate. For example, time delays on the control channel can be avoided. Advantageously, an internal clock of the forwarding element may be employed for copying a time stamp into the probe packet.

According to embodiments of the invention the measurement task may include a round-trip time measurement, a one-way delay measurement, a jitter measurement, a throughput measurement and/or an available bandwidth estimation.

5

According to embodiments of the invention, it may be provided that a measurement result is computed after collecting measurement feedback, wherein the measurement feedback is based on returned probe packets, in particular returned pong probe packets. The measurement result may be computed by the software defined network controller and/or by a dedicated measurement analysis component. The measurement feedback may include probe packets that are sent back to the software defined network controller, wherein the probe packets may comprise collected information such as time stamps, statistic counters, derived, etc. Thus, effective and flexible measurement tasks and campaigns may be provided.

10

15

Generally, network monitoring and/or network measurement may be performed on multi-hop paths through the network, due to the way that monitoring probes are deployed on the edge of the network, or at selected locations within the network. Yet with an software defined network controlled measurement system all the above measurement techniques – such as round-trip time measurement, one-way delay measurement, jitter measurement, throughput measurement and/or an available bandwidth estimation – which support multi-hop measurement can also be run on shorter paths and even individual links. Thus, with the option to perform accurate timed measurements while being able to generate the probe packets at each forwarding element in the network, once a SLA violation (i.e. a fault or degradation of performance) is detected on a multi-hop path, the software defined network controller can issue additional measurements on a sub-path until the source of the problem is identified. Thus, pinpointing sources of SLA violations may be provided.

20

25

30

At least one embodiment of the present invention may have one or more of the following advantages:

- 5 - No need for deploying dedicated monitoring probes for network monitoring, as the forwarding elements themselves can generate the probe packets, both the ping probe packets as well as the pong probe packets as reply packet. Embodiments of the present invention also may enable to turn every forwarding element into a monitoring probe, which in turn provides a finer granular probing and does not require first to deploy physically a probe
10 device in the network.
- End-to-end probing is no longer the only option. E.g. if the end-to-end probing reveals a problem, finer grained per-hop probing can enable pinpointing the exact location of the problem.
15
- The monitoring manager function is now part of the software defined network controller, and can therefore much better coordinate the measurement task and campaigns with the current network state:
 - 20 • E.g. a measurement task can be executed deliberately in busy hours to understand worse case situations or in times of low network load in order to establish a performance baseline.
 - 25 • Given that the software defined network controller knows the rules and the configuration of the network, it is now easily possible to craft packet headers for the probe packets that fit with the network, and e.g. take the correct path to be probed. This enables to break out of the limitation of probing with IP (L3) packet only.
- 30 - Switch-local information may now be used to directly trigger certain probing.
- Timing inaccuracies are reduced in comparison to an approach where the SDN controller performs probe packet generation and time stamping.

- Control channel interactions are reduced, in particular they are kept to a minimum.
- Ability to zoom in and pinpoint the source of a SLA violation.
- Possibility to take into account controller knowledge for measurement control (avoid busy hours) and probe packet forming (tailor towards installed packet forwarding rules).

5

10 Maybe, some of the above advantages might also be achieved by extensively using an software defined network controller that generates the probe traffic itself and thereby acts as both the source forwarding element (sending monitoring probe) as well as the target forwarding element (receiving monitoring probe). But this comes at several disadvantages:

15

- For delay based measurements involving the software defined network controller adds at least the following rather unpredictable components of inaccuracy:

20

- The latency between the software defined network controller and the forwarding elements, e.g. switches. This latency might vary depending on the load of the control network and the individual switch.

25

- The latency inside the software defined network switch, namely between the SDN agent and the forwarding pipeline.

30

- In general the approach using only software defined network functionality, wherein the controller is used as both sending monitoring probe and receiving monitoring probe, adds load to the control network (called control channel in OpenFlow) and the software defined network controller. Overloading the control channel or the controller with measurement traffic will prevent the network from functioning as desired.

One or more embodiments of the present invention provide a mechanism that leverages the capabilities of in-switch packet generation for the purpose of network monitoring. Thus, one or more embodiments of the invention utilize in-switch packet generation capabilities to:

5

- Programming measurement tasks and campaigns to be executed on forwarding elements of the network.

10

- Enabling in-network measurements as compared to edge-to-edge only network measurement.

- Coordinating network measurements with the current network state.

15

- Enabling automated failure search, e.g. when SLA violations are detected by zooming in on a particular link on a longer path.

20

There are several ways how to design and further develop the teaching of the present invention in an advantageous way. To this end it is to be referred to the patent claims subordinate to patent claim 1 on the one hand and to the following explanation of further embodiments of the invention by way of example, illustrated by the figure on the other hand. In connection with the explanation of the further embodiments of the invention by the aid of the figure, generally further embodiments and further developments of the teaching will be explained.

25

In the drawings

Fig. 1 is a schematic view illustrating an application scenario of a typical network monitoring probe deployment,

30

Fig. 2 is a schematic view illustrating an application scenario of a method and a software defined network according to an embodiment of the present invention,

Fig. 3 is a schematic view illustrating an application scenario of a method and a software defined network according to an embodiment of the present invention,

5 Fig. 4 is a schematic view illustrating an application scenario of a method and a software defined network according to an embodiment of the present invention.

10 Fig. 1 shows an application scenario of a typical network monitoring probe deployment. Fig. 1 shows an Internet backbone 1 of an Internet service provider (ISP). The Internet backbone 1 is accessible by a Point of Presence (PoP) 2, wherein network connection may be provided from one or more backbone gateways 3 via an aggregation layer 4 through to the customer edge 5. The
15 network is managed by a network management system (NMS) 6 that is located within a system/business support system (OSS/BSS) 7. The OSS/BSS 7, operated together by the ISP, are used to support a range of telecommunication services. According to Fig. 1, monitoring probes 8 are placed at the customer edge 5 of the network. The monitoring probes 8 are orchestrated by a monitoring manager 9 that
20 may receive information from the OSS/BSS 7. This information may include the guarantees given to the customers according to service level agreements (SLAs) 10. Furthermore, this information may include information about the topology of the network, in particular with respect to the location of the monitoring probes.

25 Fig. 2 shows three different exemplary approaches "SDN", "SDN+TS" and "SDN+TS+PT" that are employed for an application scenario in the context of round-trip time (RRT) measurements, wherein the approach SDN+TS+PT represents a method and a software defined network (SDN) according to an embodiment of the present invention. According to the approach SDN only the
30 functionality of a software defined network is used. According to the approach SDN+TS the functionality of a software defined network (SDN) is used together with time stamping (TS), wherein the time stamping is performed directly on a switch. According to the approach SDN+TS+PT the functionality of a software defined network (SDN) is used together with time stamping (TS) that is performed

directly on a switch as forwarding element and together with in-switch packet generation, wherein the in-switch packet generation is performed by using, e.g., the Packet Template (PT) mechanism.

5 When using software defined network functionality alone according to the approach SDN illustrated in Fig. 2, the software defined network controller C generates a ping probe packet – denoted as ping in Fig. 2 –, records a time stamp TS1 and uses the Packet_out message to send out the ping probe packet from a source switch S1 as forwarding element. The Packet_out message is an
10 OpenFlow protocol message for instructing a switch to output an encapsulated packet – here the ping probe packet – on a given port. Before doing so – i.e. prior to instructing the switch by using the Packet_out message – the software defined network controller C also installs a rule on another forwarding element, namely on a target switch S2, to return the received ping probe packet to the source switch
15 S1, for example by using the IN_PORT reserved port in OpenFlow. The rule is installed on the target switch S2 by using a Flow_mod message. The Flow_mod message is an OpenFlow command message for instructing the switch to install a rule on its flow table. The Flow_mod message allows the controller to modify the state of an OpenFlow switch.

20 With regard to OpenFlow, it is exemplarily referred to the non-patent literature of Open Networking Foundation: “OpenFlow Switch Specification”, Version 1.3.4 (March 27, 2014 – TS-019) retrievable at “<https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-switch-v1.3.4.pdf>”.

30 In the data plane the source switch S1 will send the ping probe packet to the target switch S2, which will return the ping probe packet to the source switch S1. Assuming a corresponding rule is present in the source switch S1, the source switch S1 will send the return ping probe packet via the Packet_in message mechanism to the software defined network controller C. The Packet_in message is an OpenFlow protocol message for sending packets to the software defined network controller. The packet_in message is for example used in cases when there is no rule installed in the forwarding element that prescribes how to handle a

specific packet. The software defined network controller C can then perform time stamping again. Hence, the software defined network controller C records a time stamp TS2 and calculates the difference between the time stamps TS1 and TS2. With some approximation of the delay between the source switch S1 and software defined network C the resulting round-trip time (RTT) will be $RTT = TS2 - TS1 - 2 * \text{delay}(S1, C)$.

In the approach SDN+TS of Fig. 2, wherein software defined network functionality is used together with performing time stamping, the step of putting time stamps is shifted from the software defined network controller C to the source switch S1. So before sending the ping probe packet to the target switch S2, the source switch S1 adds a time-stamp to the ping probe packet (denoted as ping+TS1 in Fig. 2). The target switch S2 may slightly modify (within the limits of the OpenFlow protocol capabilities) the headers of the ping probe packet such that the ping probe packet is able to be returned to the source switch S1. The returned ping probe packet will still have the time stamp TS1. Upon reception of the returned ping probe packet, the source switch S1 will add a target time stamp TS2 and send this adapted ping probe packet (denoted as ping+TS1+TS2) to the software defined network controller C via a Packet_in message, namely Packet_in(ping+TS1+TS2). This approach requires a dedicated space in the headers of the ping probe packet to store the time stamps.

The approaches according to the approach SDN and SDN+TS of Fig. 2 have the following limitations, which are resolved by the approach SDN+TS+PT illustrated in Fig. 2:

(i) Accuracy of the RTT measurement: Approach SDN illustrated in Fig. 2 with “only SDN functionality” will suffer from the following sources of inaccuracy for the RTT calculation:

- a. The control channel delay between the software defined network controller C and the source switch S1 cannot be determined with high accuracy and depends on the current load on the control channel.

- 5 b. The software defined network controller C performs the time stamping in software. Thus, depending on when the processing of the time stamping is scheduled relatively to the Packet_out message event and the Packet_in message event the time-stamp may be inaccurate.
- 10 c. In the switch, the OpenFlow Agent will need to translate the instructions of the software defined network controller to changes in the forwarding pipeline, which may be delayed by other tasks. Moreover, it is difficult to establish a baseline delay between the OpenFlow Agent and the forwarding pipeline of the source switch S1.
- 15 (ii) Repeated measurements: Usually continuous measurements may be necessary when networks are monitored. According to the approach SDN with “only SDN functionality” and according to the approach SDN+TS with “SDN functionality plus time stamping”, the software defined network controller C has to generate a probe packet for every ping probe packet. Moreover, for every ping probe packet, the software defined network controller C has to use Packet_out message to have a switch send it out. Hence, a lot of data traffic is transmitted on the control channel between the software control channel C and the switches such as the source switch S1.
- 20 (iii) The SDN switches need to have the capability to modify the header fields in which the time stamp is stored. For example, this is not possible for a typical OpenFlow switch and assuming time stamps in the TCP time stamp option.
- 25 (iv) Returning the ping probe packet back to the source switch S1 will require changes to the probe packet header, especially when there is no direct connection between the switches S1 and S2 and the returned ping probe packet will need to traverse other network forwarding elements. Hence, in this case, the ping probe packet needs to be routed appropriately (and
- 30

might involve also changes on the IP addresses, etc.). However, the concept of in-switch packet generation, e.g. by using Packet Templates (PT), as described below may enable this requirement with minimal changes to the probe packet header and facilitates routing.

5

- (v) Some forwarding pipelines of switches are not designed to forward to the ingress port.

10

Approach SDN+TS+PT illustrated in Fig. 2 is different from approach SDN+TS in that a Packet Template (PT) is installed only once in both switches S1 and S2 for configuring a measurement task. Another difference is that instead of just sending the ping probe packet back to the origin, a new pong probe packet is generated in the target switch S2 dependent on the received ping probe packet.

15

With regard to in-switch packet generation, e.g. by using Packet Templates, it is exemplarily referred to document WO 2015/000517 A1 that discloses a method for operating a software defined network and a corresponding software defined network.

20

Furthermore, it is exemplarily referred to the non-patent literature of

25

Roberto Bifulco, Julien Boite, Mathieu Bouet and Fabian Schneider: "Improving SDN with InSPired Switches", 2016, In Proceedings of the Symposium on SDN Research (SOSR '16), ACM, New York, NY, USA retrievable at "http://conferences.sigcomm.org/sosr/2016/papers/sosr_paper_42.pdf"

and to the non-patent literature of

30

F. Schneider, R. Bifulco and A. Matsiuk: "Better ARP handling with InSPired SDN switches", In 2016 IEEE International Symposium on Local and Metropolitan Area Networks (LANMAN), Rome, Italy, pp. 1-6.

Probe packet template information provided to a forwarding element such as a switch may comprise triggering instructions, a probe packet template and probe

packet handling instructions. Thus, the triggering instructions may represent a trigger for the probe packet generation. The probe packet template may be represented by a byte array, including the packet headers and content, for generating the probe packet on the forwarding element. The probe packet handling instructions represent instructions what to do with the byte array before sending it out as probe packet.

According to the embodiment illustrated by approach SDN+TS+PT in Fig. 2, probe packet templates are installed at the source forwarding element and at the target forwarding element by the following steps such that both forwarding elements function as monitoring probes for network monitoring:

- In switch S1 as source forwarding element: For generating the ping probe packet, a timer (e.g. a periodic event) is configured as a trigger. The byte array should resemble a valid ping probe packet that would reach the target switch S2 as target forwarding element based on the configuration of forwarding rules installed in the network. The ping probe packet also needs to provide space for inserting two time stamps TS1 and TS2. As probe packet template handling instructions, it is necessary to copy the locally generated time stamp into the ping probe packet and optionally copy an incrementing sequence number into the ping probe packet. This step is denoted as PktTemp(ping) in Fig. 2.
- In switch S2 as target forwarding element: For generating the pong probe packet, the reception of a ping probe packet is configured as a trigger. The byte array will be similar to the ping probe packet, however with a header that allows reaching the origin of the ping probe packet. As probe packet template handling instructions, it is necessary to copy the time stamp from the received ping probe packet, and optionally also the sequence number. This step is denoted as PktTemp(pong) in Fig. 2.

Once the pong probe packet is received at switch S1, the switch S1 will need to time stamp the received pong probe packet again (by adding TS2) and send the pong probe packet via a Packet_in message to the software defined network

controller C. This final step might again be implemented by using a probe packet template that is triggered by the reception of the pong probe packet. With some additional intelligence in switch S1, the double time-stamped pong probe packets might only be sent to the software defined network controller C if a predetermined threshold for the RTT is reached to prevent flooding the controller C.

According to approach SDN+TS+PT illustrated in Fig. 2, it is much easier to set the appropriate header for the return path leveraging the concept of Packet Templates.

Employing and installing probe packet templates at forwarding elements for the purpose of network monitoring is based on the goal to be able to respond to incoming probe packets by generating a new dedicated response probe packet. And as stated before the option for periodic generation of probe packets, in particular ping probe packets, will significantly reduce the load in the control path and the software defined network controller.

Fig. 3 shows a further application scenario of a method and a software defined network according to an embodiment of the present invention. By analogy to Fig. 2, Fig. 3 illustrates three different exemplary approaches SDN, SDN+TS and SDN+TS+PT that are employed for the application scenario in the context of one-way delay (OWD) measurements, wherein approach SDN+TS+PT represents a method and a software defined network (SDN) according to an embodiment of the present invention. One-way delay measurements can be considered as a particular case of RTT measurements, e.g. as described and illustrated by Fig. 2.

According to the approach SDN, the functionality of a software defined network (SDN) is used for the measuring task. According to the approach SDN+TS, the functionality of a software defined network (SDN) is used together with time stamping (TS), wherein the time stamping is performed directly on a switch. According to the approach SDN+TS+PT the functionality of a software defined network (SDN) is used together with time stamping (TS) that is performed directly on a switch as forwarding element and together with in-switch packet generation. The in-switch packet generation is implemented by using Packet Template (PT).

Since the delay for one-way delay measurements is measured in a single direction, the ping probe packet is not returned to the source switch S1. The target switch S2 may encapsulate the received ping probe packet into the Packet_in message and sends the encapsulated ping probe packet directly to the software defined network controller C. This condition relieves previously described limitations related to the need to return the ping probe packet back to the source switch S1.

The approach SDN with only software defined network functionality suffers from the same sources of inaccuracy as described above with regard to the application scenario of Fig. 2. Though, the latency measurement framework proposed in the non-patent literature of C. Yu, C. Lumezanu, A. Sharma, Q. Xu, G. Jiang and H. Madhyastha: "Software-defined Latency Monitoring in Data Center Networks", 2015 Passive and Active Measurement Conference (PAM), New York, March, 2015 retrievable at "<http://www.nec-labs.com/~gfj/yu-pam-15.pdf>" aims to decompose and to estimate the control channel delay (between the controller and the OpenFlow agent) and forwarding pipeline delay (between OpenFlow agent and the ASIC), various implementations of OpenFlow agents and their hidden processing logic do not allow to estimate and correlate processing delays of one type of messages based on the measurements for the others. Moreover, these delays and their cross-correlations may change significantly under dynamically changing load of a switch. Furthermore, such approaches do not consider time stamping inaccuracy introduced by the controller's processing logic.

In the approach SDN+TS illustrated by Fig. 3, the OWD value is measured by the software defined network controller C as the difference between the time stamp TS2 and TS1 values attached by the switches S2 and S1 to the ping probe packet respectively. OWD measurements rely on the fact that internal clocks of both switches are synchronous. This requirement can be supported by a specific clock synchronization protocol with sufficient precision (e.g. NTP, IEEE 1588 PTP etc.). Apart from the clock synchronization need, the approach SDN+TS has also the limitations (ii) and (iii) described above with regard to the embodiment of Fig. 2.

The approach SDN+TS+PT of Fig. 3 overcomes the limitations of the previous both approaches SDN and SDN+TS. Furthermore, the approach SDN+TS+PT of Fig. 3 merely requires slight modification to the RTT measurements application scenario of Fig. 2. Here, the Packet Template (PT) installed into the target switch S2 has a probe packet template and probe packet template handling instructions that allow to generate the Packet_in message from target switch S2 directly to the software defined network controller instead of returning it back to the source S1 through the fast path.

Since the target switch S2 does not need to send the pong probe packet back immediately and generally can defer the generation of the Packet_in message, the target switch S2 can perform additional computations on the time stamps, e.g. encoding TS2-TS1 difference directly into the Packet_in message or reporting only OWD threshold crossing to the controller. Also, since the approach SDN+TS+PT implies synchronized switches, the switch S1 may avoid the “full” time stamp writing into the pong probe packet, but can encode into the probe packet header an increasing counter relatively to some reference point which is known to all switches, e.g. start of the second. Such simplification may easily complement the constant-rate periodic probe packet generation, relieve the demand for a dedicated packet header for the time stamp TS1 and simplify the Packet Template generation logic at the source switch S1. Since the reference point is known to the switch S2, the switch S2 can simply derive the TS1 from the encoded counter and can preform related calculations if needed.

Jitter measurements are another application scenario for an embodiment of the present invention. Jitter measurements may be considered as a straightforward modification of one-way delay measurements with the difference that the target switch S2 needs to evaluate the variation of the OWD delay of subsequent received ping probe packets instead of its absolute value. The ping probe packets are transmitted at every fixed time interval and the variation of the packet inter-arrival time is considered as jitter. Thus, jitter measurements relax the need for synchronized clocks of both switches. As described above, the approach SDN will suffer from instability of control channels between the controller and both switches masking the actual jitter of data plane links.

The approach SDN+TS+PT effectively eliminates this instability and relies on the generation of equally spaced packets on the source switch S1 which in turn depends only on its internal clock stability. The target switch S2 needs to measure the delay between any two subsequent ping probe packets and to provide this value to the software defined network controller C. This step can be accomplished with a dedicated timer at the target switch S2 and probe packet template handling instructions to fix the timer's value on an arrival of ping probe packet, i.e. this would be the trigger for the pong probe packet generation and to reset the timer afterwards. The fixed value, i.e. the delay between two ping probe packets, can be directly written into the Packet_in message towards the software defined network controller C or, more effectively, several measurements can be accumulated into the same Packet_in message. Later, the software defined network controller C may derive the actual jitter value from these inter-arrival ping probe packet delays.

Fig. 4 shows a further application scenario of a method and a software defined network according to an embodiment of the present invention. Specifically, Fig. 4 illustrates a throughput measurement task with an approach SDN+TS+PT. Generally, throughput measurements imply link saturation with maximum-sized probe packets and counting the number of probe packets at the end of the link over a time unit. An implementation of such measurements with only software defined network functionality or according to an SDN+TS approach will face severe limitations making it impractical without an extended support at the switches:

- Link saturation requires high-rate generation of maximum-sized probe packets and encapsulation of them into Packet_out messages by the software defined network controller as well as parsing of these messages by the OpenFlow agent and injecting the probe packets into the data plane at the source switch S1. Similar operations but in the reverse order should be applied at the target switch S2 to send all probe packets back to the software defined network controller C. These operations may exhaust CPU resources both at the software defined network controller C and the switches S1, S2.

- Typically, control channels have smaller bandwidth than the data plane creating a bottleneck in a measurements setup.

Similarly to the approaches SDN and SDN+TS, the throughput measurement with the approach SDN+TS+PT according to an embodiment of the present invention may require the following steps:

- Installing a rule matching the ping probe packets in the target switch S2. This is denoted as step I in Fig. 4.
- In the source switch S1 a probe packet template is installed for generating the maximum-sized ping probe packets. An internal timer serves as a trigger for generation of ping probe packets and may be pre-calculated given the maximum link throughput. This is denoted as step II in Fig. 4.
- The actual ping probe packets that are automatically triggered at switch S1 (denoted as step III in Fig. 4) may be simply dropped as for throughput calculations the software defined network controller C requires only an access to the statistics counters (bytes, packets) that can be provided by standard OpenFlow semantic (denoted as step IV in Fig. 4).
- The setup requires only two time stamps TS1 and TS2 indicating the beginning and the end of the throughput measurement task. Long-term measurements do not require the time stamping at the switches, since the control channel inaccuracies become negligible relatively to the measurement period. Here, both time stamps TS1 and TS2 may be safely performed by the software defined network controller C.

With regard to available bandwidth estimation it may be exemplarily referred to the following non-patent literature of

- Jacob Strauss, Dina Katabi, and Frans Kaashoek: „A measurement study of available bandwidth estimation tools”, 2003, In Proceedings of the 3rd ACM SIGCOMM conference on Internet measurement (IMC '03). ACM, New

York, NY, USA, 39-44 retrievable at "<https://pdos.csail.mit.edu/papers/spruce/imc03.pdf>";

- 5 - Oana Goga and Renata Teixeira: "Speed measurements of residential internet access", 2012, In Proceedings of the 13th international conference on Passive and Active Measurement (PAM'12), Nina Taft and Fabio Ricciato (Eds.), Springer-Verlag, Berlin, Heidelberg, 168-178 retrievable at "www.mpi-sws.org/~ogoga/papers/PAM12-speed.pdf"; and
- 10 - Alok Shriram, Margaret Murray, Young Hyun, Nevil Brownlee, Andre Broido, Marina Fomenkov, and kc claffy.: "Comparison of public end-to-end bandwidth estimation tools on high-speed links", 2005, In Proceedings of the 6th international conference on Passive and Active Network Measurement (PAM'05), Constantinos Dovrolis (Ed.). Springer-Verlag,
15 Berlin, Heidelberg, 306-320 retrievable at „<https://www.caida.org/publications/papers/2005/pam-bwest/pam-bwest.pdf>“

for an overview of available bandwidth measurement approaches. Many of these approaches rely on well-defined spacing between probe packets and the recording
20 of the inter-probe spacing at the destination. Such a well-defined spacing either requires accurate timed generation of probe packets or an ensured back-to-back generation.

By the use of timers as triggers for in-switch generation of packets and by
25 generating probe packets with increasing sequence numbers using Packet Templates, available bandwidth estimation measurements can be particularly improved. Specifically, the advantage of using Packet Templates for available bandwidth estimations is that possible inaccuracies due to delay and jitter on the control channel can be eliminated as well as the load of packet generation at the
30 controller and transmission through the control channel can be reduced.

The limitations that are overcome by using Packet Templates are similar to the improvements around RTT and OWD measurements. The inter-probe times can

- 25 -

be recorded either at the receiving switch (again assuming time-stamping capabilities) or at the software defined network controller (less accurate).

5 Many modifications and other embodiments of the invention set forth herein will come to mind to the one skilled in the art to which the invention pertains having the benefit of the teachings presented in the foregoing description and the associated drawings. Therefore, it is to be understood that the invention is not to be limited to the specific embodiments disclosed and that modifications and other embodiments are intended to be included within the scope of the appended claims. Although
10 specific terms are employed herein, they are used in a generic and descriptive sense only and not for purposes of limitation.

List of reference signs

1	Internet Backbone
2	Point of Presence
3	Backbone Gateways
4	Aggregation Layer
5	Customer Edge
6	Network Management System
7	OSS/BSS
8	Monitoring Probes
9	Monitoring Manager
10	SLA
C	SDN Controller
S1	Source Switch
S2	Target Switch
TS1	Time stamp
TS2	Time stamp

Claims

1. Method for supporting network monitoring in a software defined network, wherein said software defined network comprises forwarding elements and a software defined network controller for controlling the forwarding elements,
5 wherein a source forwarding element and a target forwarding element are employed to perform a measurement task for network monitoring such that said source forwarding element and said target forwarding element function as monitoring probes, and
10 wherein said software defined network controller configures said monitoring probes such that one or more probe packets for performing the measurement task are generated by at least one of said monitoring probes.
2. Method according to claim 1, wherein generation and/or handling of probe packets for performing the measuring task is/are temporally decoupled from said software defined network controller.
3. Method according to claim 1 or 2, wherein said one or more probe packets include a ping probe packet for performing the measuring task, wherein said ping probe packet is generated by said source forwarding element.
20
4. Method according to any of claims 1 to 3, wherein said one or more probe packets include a pong probe packet for performing said measurement task, wherein said pong probe packet is generated by said target forwarding element.
25
5. Method according to any of claims 1 to 4, wherein said software defined network controller configures said monitoring probes by creating probe packet template information that is provided to said monitoring probes, wherein said probe packet template information includes one or more probe packet templates and/or probe packet template handling instructions.
30
6. Method according to any of claims 5, wherein said monitoring probes are triggered to generate said one or more probe packets based on the provided probe packet template information.

7. Method according to claim 5 or 6, wherein triggering instructions are provided to said monitoring probes by including them into the packet template information.

5 8. Method according to claim 7, wherein said triggering instructions include configuring, in particular at said source forwarding element, a timer as a trigger for generating said one more probe packets, in particular for generating a ping probe packet on said source forwarding element.

10 9. Method according to claim 7 or 8, wherein said triggering instructions include configuring, in particular at said target forwarding element, a reception of a predetermined probe packet, in particular a ping probe packet, as trigger for generating a new probe packet, in particular a pong probe packet.

15 10. Method according to any of claims 1 to 9, wherein local information of said monitoring probes is employed to trigger said one or more probe packets for performing the measurement task.

20 11. Method according to any of claims 1 to 10, wherein time stamping is performed by at least one of said monitoring probes.

25 12. Method according to any of claims 1 to 11, wherein an internal clock of a forwarding element functioning as monitoring probe is employed for adding time stamps into said one or more probe packets.

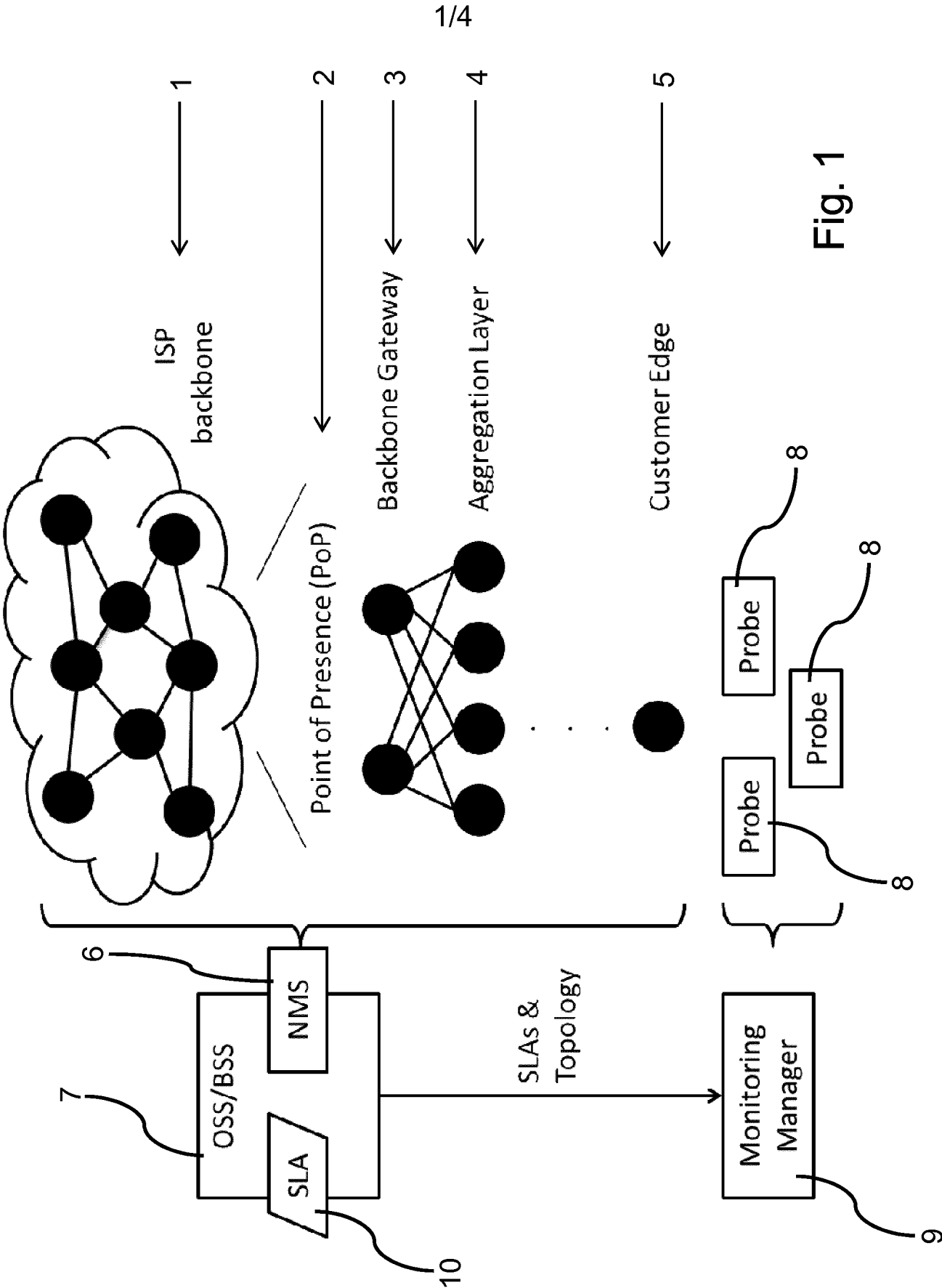
30 13. Method according to any of claims 1 to 12, wherein said measurement task includes a round-trip time measurement, a one-way delay measurement, a jitter measurement, a throughput measurement and/or an available bandwidth estimation.

14. Method according to any of claims 1 to 13, wherein a measurement result is computed after collecting measurement feedback, wherein said measurement feedback is based on returned probe packets.

15. Software defined network for supporting network monitoring, in particular for executing a method according to any of claims 1 to 14, the network comprising forwarding elements and a software defined network controller for controlling the forwarding elements,

5 wherein said forwarding elements include a source forwarding element and a target forwarding element that are programmable to perform a measurement task for network monitoring such that said source forwarding element and said target forwarding element function as monitoring probes, and

10 wherein said software defined network controller is operable to configure said monitoring probes such that one or more probe packets for performing the measurement task are generated by at least one of said monitoring probes.



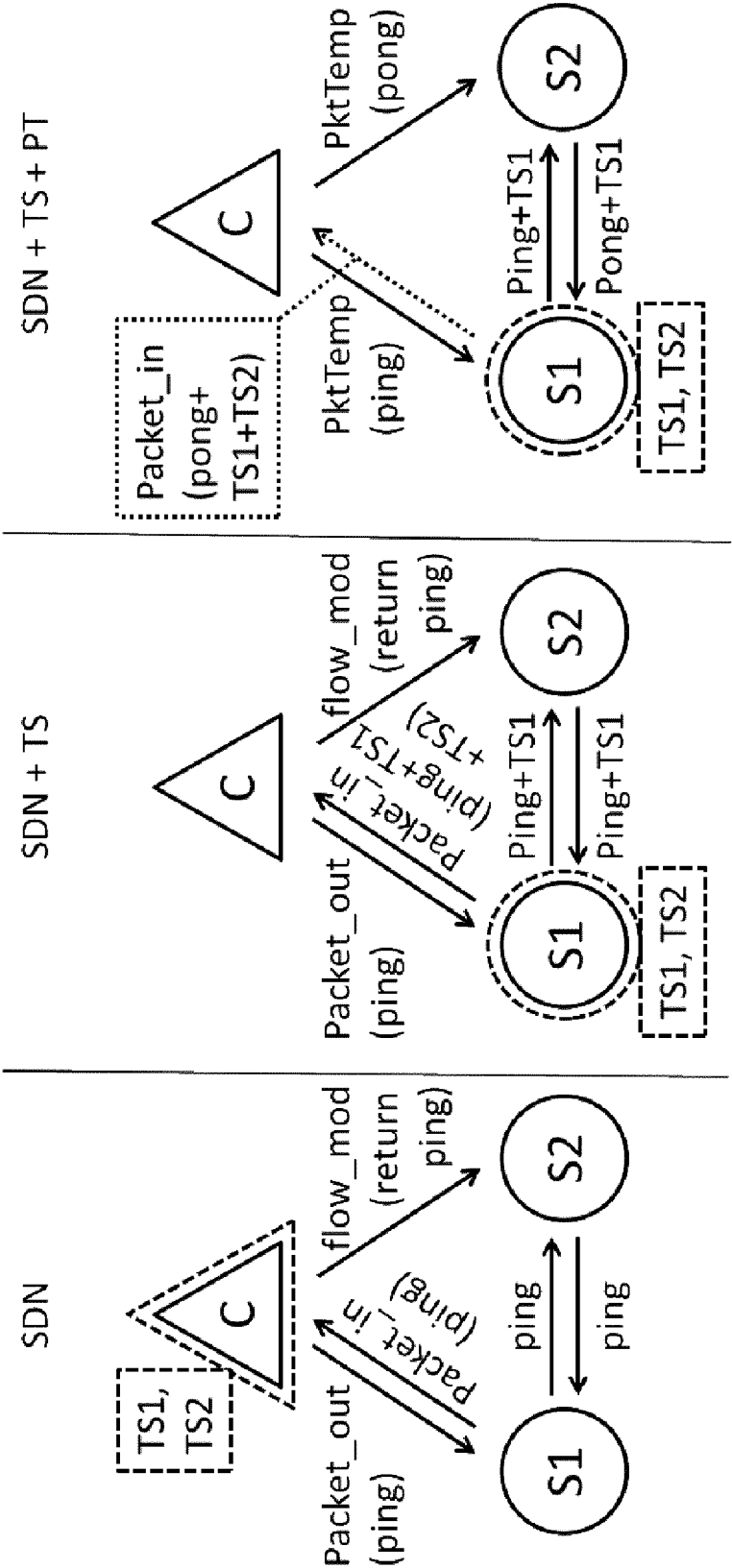


Fig. 2

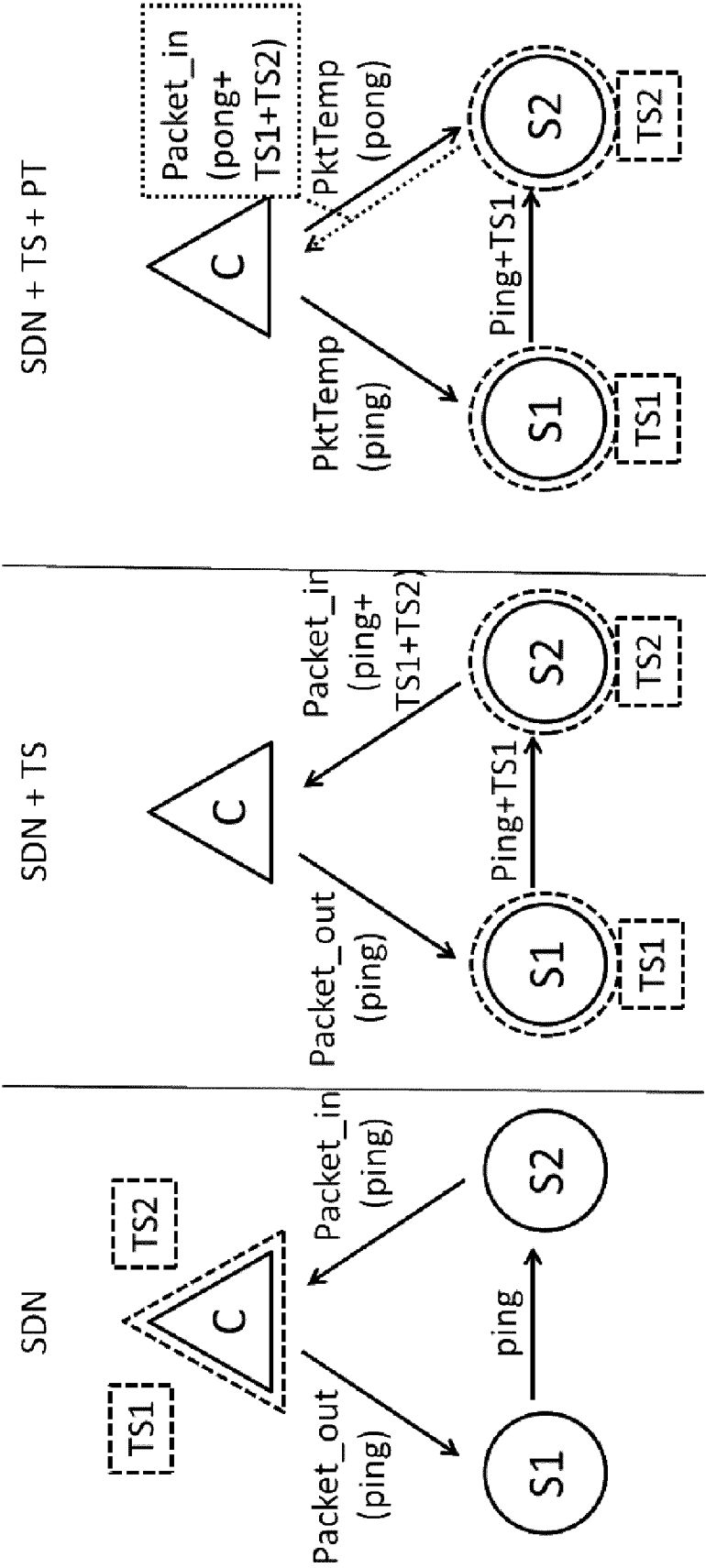


Fig. 3

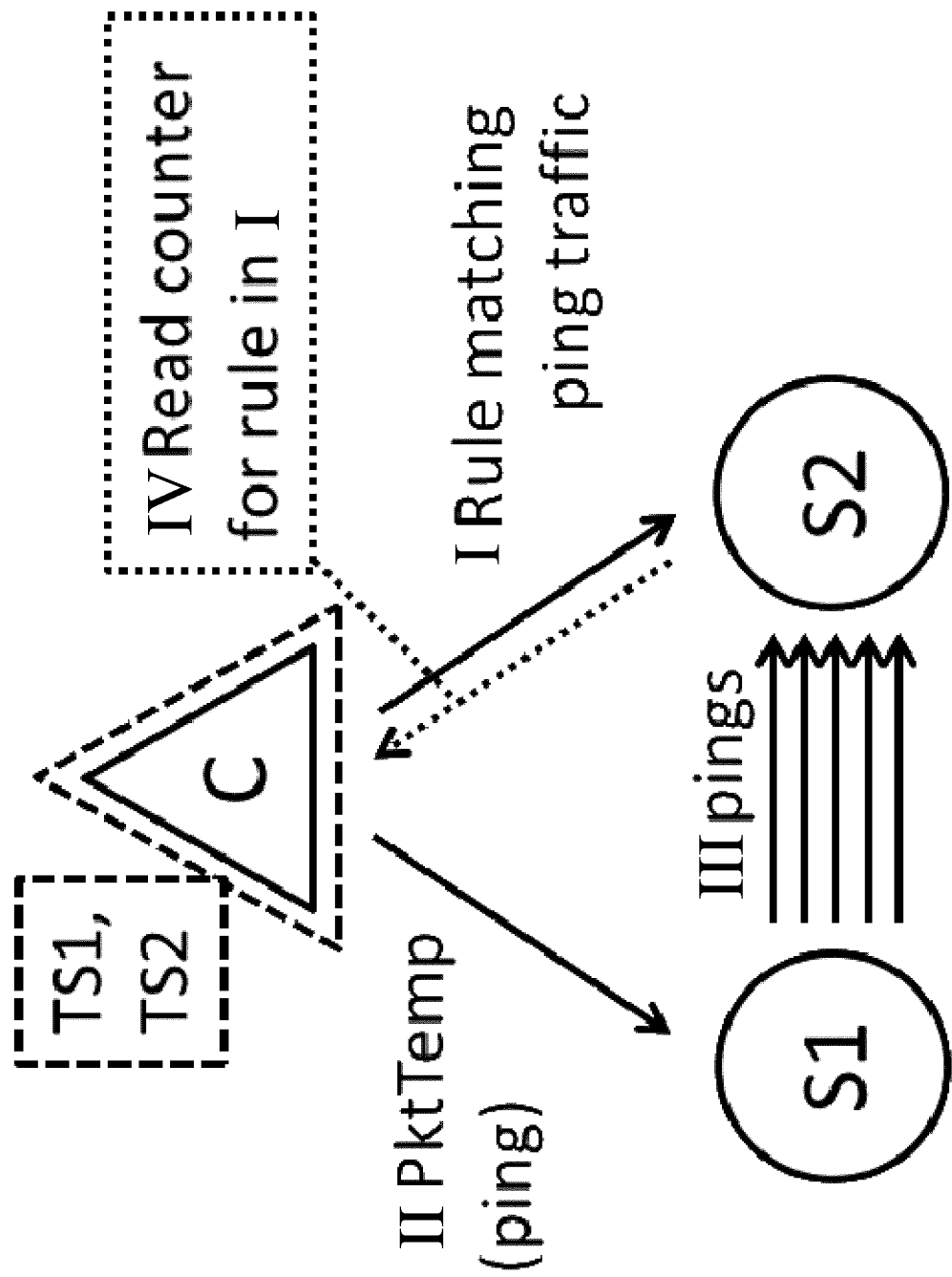


Fig. 4

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2016/076154

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04L12/26
ADD.

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)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2015/071108 A1 (LUMEZANU CRISTIAN [US] ET AL) 12 March 2015 (2015-03-12) abstract paragraphs [0006], [0027] - [0036]; figures 2, 3	1-15
A	----- JAMES KEMPF ET AL: "Scalable fault management for OpenFlow", COMMUNICATIONS (ICC), 2012 IEEE INTERNATIONAL CONFERENCE ON, IEEE, 10 June 2012 (2012-06-10), pages 6606-6610, XP032274514, DOI: 10.1109/ICC.2012.6364688 ISBN: 978-1-4577-2052-9 abstract page 6607, paragraph III ----- -/-	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

18 April 2017

Date of mailing of the international search report

25/04/2017

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Fantacone, Vincenzo

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2016/076154

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2016/285729 A1 (CHINTHALAPATI ESWARA [US] ET AL) 29 September 2016 (2016-09-29) abstract; figures 4, 5 -----	1-15

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2016/076154

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2015071108 A1	12-03-2015	EP 3042477 A1	13-07-2016
		JP 2016524401 A	12-08-2016
		US 2015071108 A1	12-03-2015
		WO 2015034564 A1	12-03-2015

US 2016285729 A1	29-09-2016	US 2016285729 A1	29-09-2016
		US 2016285750 A1	29-09-2016
		WO 2016153713 A1	29-09-2016
