ITUPublications
Recommendations

International Telecommunication Union
Standardization Sector

Recommendation

ITU-T Q.3962 (12/2023)

SERIES Q: Switching and signalling, and associated measurements and tests

Testing specifications – Testing specifications for next generation networks

Requirements and reference model for optimized traceroute of joint Internet protocol/ multiprotocol label switching

ITU-T Q-SERIES RECOMMENDATIONS

Switching and signalling, and associated measurements and tests

SIGNALLING IN THE INTERNATIONAL MANUAL SERVICE	Q.1-Q.3
INTERNATIONAL AUTOMATIC AND SEMI-AUTOMATIC WORKING	Q.4-Q.59
FUNCTIONS AND INFORMATION FLOWS FOR SERVICES IN THE ISDN	Q.60-Q.99
CLAUSES APPLICABLE TO ITU-T STANDARD SYSTEMS	Q.100-Q.119
SPECIFICATIONS OF SIGNALLING SYSTEMS NO. 4, 5, 6, R1 AND R2	Q.120-Q.499
DIGITAL EXCHANGES	Q.500-Q.599
INTERWORKING OF SIGNALLING SYSTEMS	Q.600-Q.699
SPECIFICATIONS OF SIGNALLING SYSTEM NO. 7	Q.700-Q.799
Q3 INTERFACE	Q.800-Q.849
DIGITAL SUBSCRIBER SIGNALLING SYSTEM NO. 1	Q.850-Q.999
PUBLIC LAND MOBILE NETWORK	Q.1000-Q.1099
INTERWORKING WITH SATELLITE MOBILE SYSTEMS	Q.1100-Q.1199
INTELLIGENT NETWORK	Q.1200-Q.1699
SIGNALLING REQUIREMENTS AND PROTOCOLS FOR IMT-2000	Q.1700-Q.1799
SPECIFICATIONS OF SIGNALLING RELATED TO BEARER INDEPENDENT CALL CONTROL (BICC)	Q.1900-Q.1999
BROADBAND ISDN	Q.2000-Q.2999
SIGNALLING REQUIREMENTS AND PROTOCOLS FOR THE NGN	Q.3000-Q.3709
SIGNALLING REQUIREMENTS AND PROTOCOLS FOR SDN	Q.3710-Q.3899
TESTING SPECIFICATIONS	Q.3900-Q.4099
Testing specifications for next generation networks	**Q.3900-Q.3999**
Testing specifications for SIP-IMS	Q.4000-Q.4039
Testing specifications for Cloud computing	Q.4040-Q.4059
Testing specifications for IMT-2020 and IoT	Q.4060-Q.4099
PROTOCOLS AND SIGNALLING FOR PEER-TO-PEER COMMUNICATIONS	Q.4100-Q.4139
PROTOCOLS AND SIGNALLING FOR COMPUTING POWER NETWORKS	Q.4140-Q.4159
PROTOCOLS AND SIGNALLING FOR QUANTUM KEY DISTRIBUTION NETWORKS	Q.4160-Q.4179
SIGNALLING REQUIREMENTS AND PROTOCOLS FOR IMT-2020	Q.5000-Q.5049
COMBATING COUNTERFEITING AND STOLEN ICT DEVICES	Q.5050-Q.5069

For further details, please refer to the list of ITU-T Recommendations.

Recommendation ITU-T Q.3962

Requirements and reference model for optimized traceroute of joint Internet protocol/multi-protocol label switching

Summary

Recommendation ITU-T Q.3962 aims to solve the problems of wrong failure location and performance information brought by the traditional isolated traceroute tools in a joint Internet protocol/ multiprotocol label switching (IP/MPLS) scenario. This Recommendation describes the requirements and reference model for optimized traceroute for joint IP/MPLS.

History *

Edition	Recommendation	Approval	Study Group	Unique ID
1.0	ITU-T Q.3962	2023-12-14	11	11.1002/1000/15720

Keywords

ICMP traceroute, IP/MPLS.

* To access the Recommendation, type the URL https://handle.itu.int/ in the address field of your web browser, followed by the Recommendation's unique ID.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure, e.g., interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents/software copyrights, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the appropriate ITU-T databases available via the ITU-T website at http://www.itu.int/ITU-T/ipr/.

Table of Contents

Recommendation ITU-T Q.3962

Requirements and reference model for optimized traceroute of joint Internet protocol/multiprotocol label switching

1 Scope

The scope of this Recommendation consists of:

1) Requirements of route tracing of joint Internet protocol/multiprotocol label switching (IP/MPLS);
2) Methods of optimized traceroute of joint IP/MPLS;
3) Reference model for optimized traceroute of joint IP/MPLS.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

None.

3 Definitions

3.1 Terms defined elsewhere

None.

3.2 Terms defined in this Recommendation

This Recommendation defines the following term:

3.2.1 s**ubstituted IP address**: An IP address that replaces the original IP address of the device during a test scenario that uses ping or traceroute.

NOTE – The substituted IP address is only used for the purpose of network security and should not be assigned to a customer.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

FIB Forwarding Information Base

ICMP Internet Control Message Protocol

IP Internet Protocol

LFIB Label Forwarding Information Base

MPLS Multiple Protocol Label Switch

PE Provider Edge

TTL Time To Live

VPN Virtual Private Network

VRF Virtual Routing Forwarding

5 Conventions

In this Recommendation:

The keywords "**is required** " indicate a requirement which must be strictly followed and from which no deviation is permitted if conformance to this document is to be claimed.

6 Background

There are several service scenarios using joint IP/MPLS. "Joint" here means that the end-to-end path is joined in several sections using different technologies. Figure 6-1 shows a typical end-to-end enterprise network. In this scenario, customer equipment accesses the carrier network using IP protocol. Within the carrier's network, the carrier uses the MPLS protocol to transfer the enterprise's packets. So, it is a typical service of jointly using different protocols (IP and MPLS).

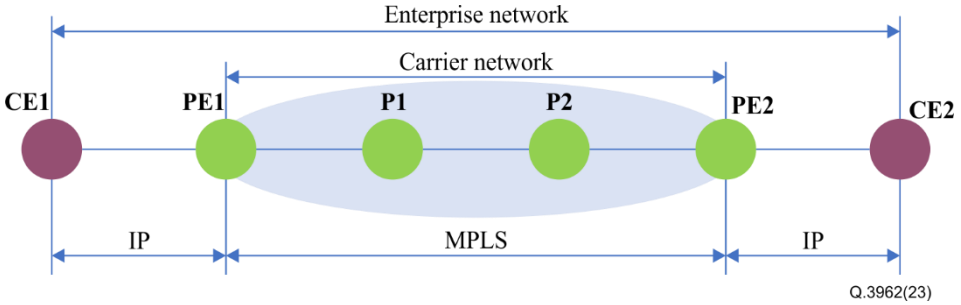

Figure 6-1 – Enterprise IP network carried by service provider's MPLS network

In such a scenario, the current route tracing technologies are IP traceroute technology and MPLS traceroute technology. Those tools are running perfectly in separate IP or MPLS environments. But when they are put together, because they are run in different layers and cannot communicate with each other, it is impossible to find out the breakdown points in an IP and MPLS joint network effectively.

Here is a simple example which explains the behaviour when route trace is triggered from CE1 to remote CE2 through IP and MPLS domain.

1) The first packet is sent with a time to live (TTL) value of 1 in IP header from CE1 to CE2:

   This is a normal IP packet and when it arrives at provider edge 1 (PE1), the TTL value decreases to 0. PE1, generates the Internet Control Message Protocol (ICMP) error message and sends it directly to CE1.

2) The second packet sent with TTL=2 in IP header from CE1 to CE2:

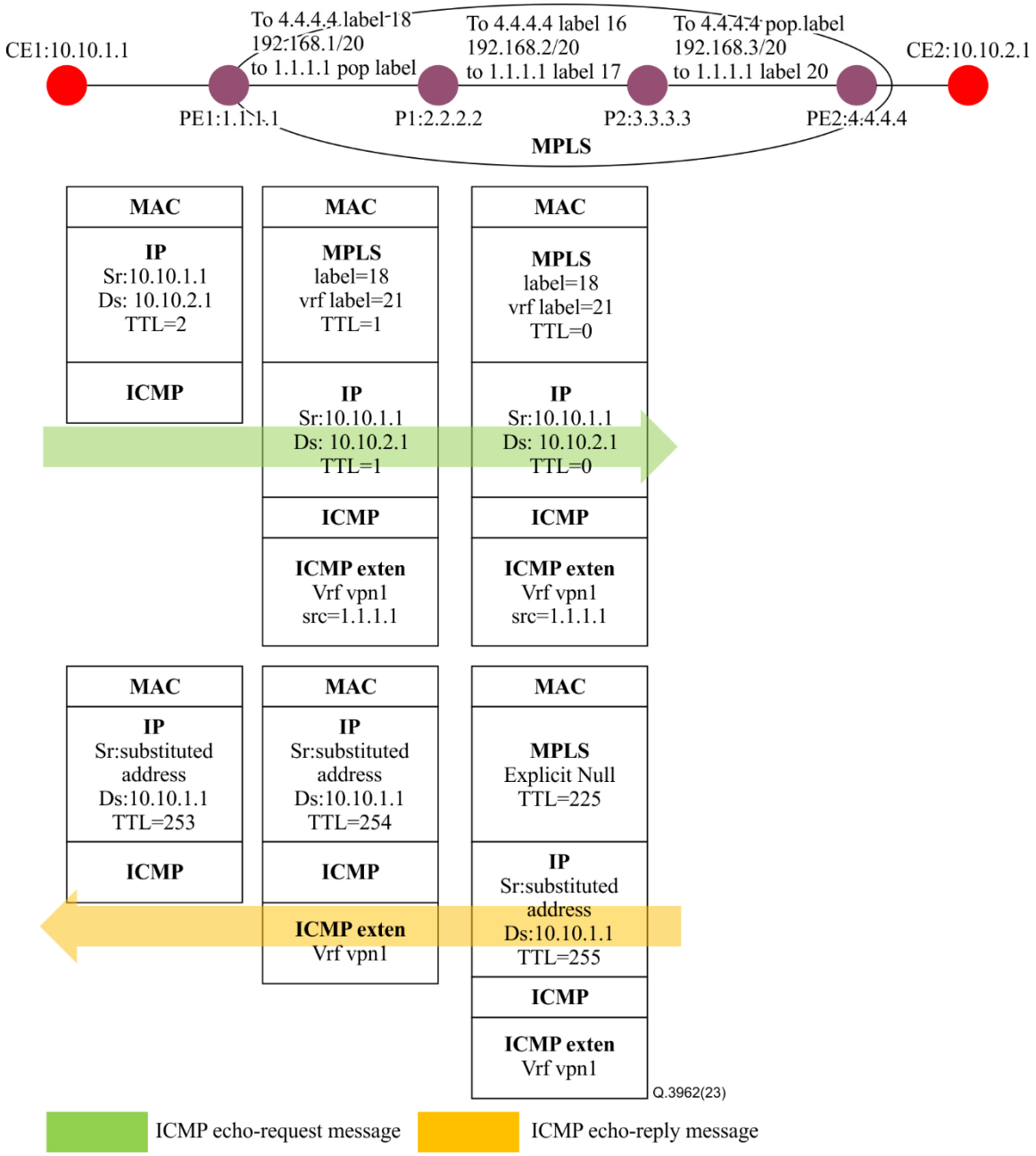

Figure 6-2 – Example of traceroute for an end-to-end enterprise network

The example of TTL=2 traceroute for an end-to-end enterprise network is shown in Figure 6-2.

After the traceroute, the packet arrives at PE1, the TTL of the IP header decreases to 1. The PE1 adds the two MPLS layer tags (outside label = 18; inside virtual routing forwarding (VRF) label = 21) to the packet header and sends the packet to PE2. The traceroute operation within the MPLS domain uses the MPLS traceroute tool. In MPLS scenarios, the traceroute packet is switched based on the MPLS tag values, not on the destination IP addresses of CE2.

There are two different options when PE1 transforms the IP traceroute packet to MPLS traceroute packets:

A. One option is not copying the TTL value to the MPLS header from the IP header in PE1.

Usually, it is forbidden to leak the service providers' network information. For example, it is forbidden to leak the IP address of the routers to the customers. So, the TTL value of the IP header will not be copied to the MPLS header and the service provider's network is transparent to the customers. In this situation, when there are network failures between CE1

and CE2, the customer has no opportunities to know where the network failure has happened, in the service providers' network or in the customer network.

B. The other option is copying the TTL value to the MPLS header from the IP header in PE1.

In this way, PE1 regenerates the traceroute packet with the MPLS header of TTL value = 1 which is copied from the IP header. When the packet arrives at P1, the MPLS TTL value is decreased to 0. P1 buffers the label stack and generates an ICMP error message and includes the incoming label stack from the buffer in the ICMP payload. It further populates the IP header with the source address from the incoming interface (192.168.1.1) of the labelled packet, destination address as the source of the labelled packet (10.10.1.1). The TTL value is set to 255. It now pushes the label stack from the buffer and consults the label forwarding information base (LFIB) table for forwarding action on the top label. In the above topology, for example, the received label stack is {18, 21}. On performing a lookup in LFIB table for top label, 18 will be swapped with label 16 and will be forwarded towards next hop P2. P2 in turn will pop the top label and forward the packet to PE2. PE2 will use the VRF label 21 to identify the VRF and forward the packet back towards CE1(outer label = 20, VRF label = 35). To conclude, the ICMP packet is not sent back to CE1 directly from P1. Instead, this ICMP packet generated from P2 is firstly steered to the end point of the MPLS virtual private network (VPN) tunnel PE2 and then steered back to CE1 from PE2. The drawbacks of this method are obvious:

– The packet takes a long journey with a big circle to reach the destination. This not only costs a large time-delay but also wastes the bandwidth from P2 to PE2.
– Most seriously, the statistic of delay reflected from P2 to CE1 is totally incorrect. It cannot provide the correct time cost between P2 and CE1 and consequently it is meaningless.

In summary, if the TTL value is not copied from the IP header, the customer has no opportunity to find out the location of failure points if they are located in the service provider's network. If the TTL value is copied, the customer has the opportunity to find the failure point, but the delay/loss/jitter information of this failure point is not correct.

7 Requirements for route tracing of an enterprise network

Since the enterprise network is constructed by different technologies and different network operators, to coordinate all of these factors to find out the failures in real time, the traceroute technology has following requirements:

– It is required that the TTL value of the IP header should be copied to the TTL field of the MPLS header. In this way, the customer is capable to identify the faults no matter the faults are located in the MPLS domain of service provider or in the IP domain of the enterprise itself.
– It is required that the service providers' network information won't be leaked to the customer.
– It is required that the delay/packet loss/jitter information of the failure point are correct.

8 Optimized method of route tracing of an enterprise network

It is essential to directly steer the ICMP echo-reply message from tested node to the source node.

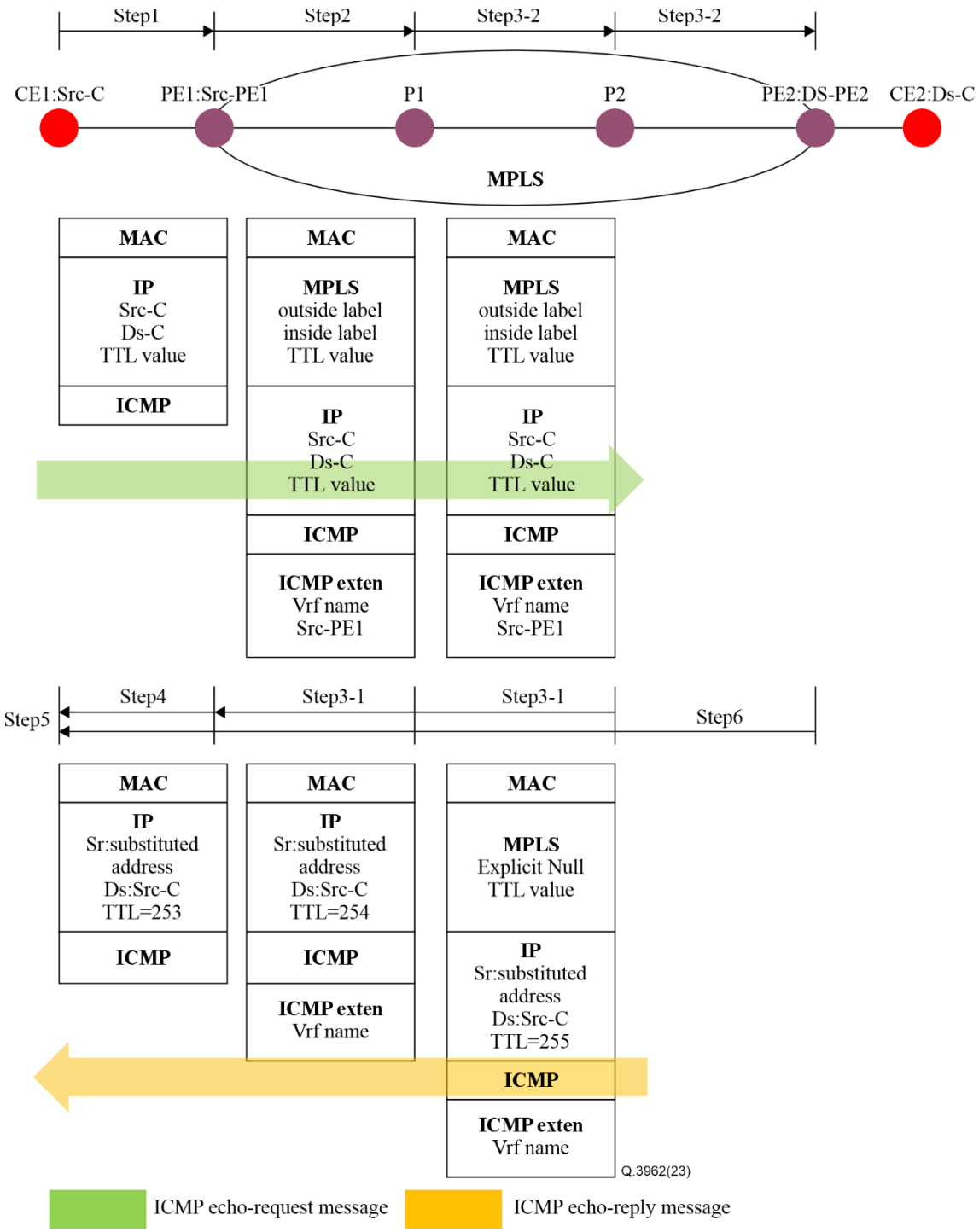

Figure 8-1 – Optimized traceroute for joint IP/MPLS

NOTE – The detailed example of this optimized method of route tracing of an enterprise network is illustrated in Appendix I.

Step 1: In the topology of Figure 8-1, ICMP traceroute packet (ICMP echo-request packets) which has a TTL value = *n* in the IP header is triggered from CE1 to CE2.

Step 2: When ICMP echo-request packets arrives at PE1, PE1 will do the following actions:

Phase 1: To deal with the received IP packet:

a) Minus the TTL value by 1 of the IP header;

b) Decapsulating the source IP address Src-C from the IP header.

Phase 2: To reconstruct the outgoing ICMP echo-request packet:

a) Finding out the mapped VRF name by checking the source IP address from the VRF routing table;

b) Finding out the VRF name related outside outgoing label;

c) Finding out the VRF name related inside outgoing label;

d) Finding out the VRF name related inside incoming label;

e) Adding the outside label to the MPLS header;

f) Adding the inside outgoing VRF label to the MPLS header;

g) Copying the TTL value from IP header to the MPLS header;

h) Remaining the TTL value of the IP header;

i) Adding the VRF name to the ICMP payload;

j) Adding the source end of the MPLS VPN tunnel (Src-PE1) to the ICMP payload;

k) Sending the restructured ICMP echo-request packets to next node.

Step 3: When the restructured packets arrive at the next Pn node (P1 or P2 in Figure 8-1), Pn will take the following actions:

Minusing the TTL of MPLS layer. If TTL value = 0, go to Step 3-1. If the TTL value > 0, go to Step 3-2.

Step 3-1: The TTL value = 0, it means that the packet is time exceeded. Pn will generate ICMP error message (also known as echo-reply message):

a) Checking the source end of the MPLS VPN (Src-PE1) tunnel carried in ICMP payload;

b) Checking the LFIB and find out Src-PE1 mapping to label "Explicit Null";

c) Generating the MPLS header of the ICMP error message:

   i) Adding 'Explicit 'None' to the MPLS outside label field;

   ii) Setting the TTL = 255;

d) Generating the IP header of the ICMP error message:

   i) Checking out the substituted Pn's IP address related to the original Pn's IP address from a dedicated table maintained in Pn which stores the mapping of these two kinds of addresses.

   NOTE 1 – To satisfy the second requirements of clause 7 that "It is required that the service providers' network information won't be leaked to the customer", Pn should acknowledges the substituted IP address instead of the original IP address of itself to the CE1 for the sake of hiding itself.

   NOTE 2 – The substituted IP address and original IP address mapping table maintained in Pn is synchronized by the related server administrated by the network service provider.

   NOTE 3 – The substituted IP address aims to prevent the distributed denial-of-service (DDOS) attack from the hackers. If the traceroute packet is initiated by the operator of the network service provider, the operator will check out the original IP address of Pn based on the substituted address from the mapping table and run the regular operations according to the original IP address to figure out the network failures. The second requirement of clause 7 is meet and the routers within the service provider's domain is able to cancel the ping prohibition.

   ii) Setting the substituted Pn's IP address as the IP source address in the IP header;

   iii) Setting Src-C which is read out from the source IP address field of the IP header of the ICMP request packet;

   iv) Setting the IP TTL=255.

e) Generating the ICMP payload and retaining the VRF name in the ICMP payload.

f) Going to Step 4.

Step 3-2: The TTL value > 0, Pn reconstructes the ICMP echo-request packet by updating the TTL values of MPLS and sending it to the next node. If the next node is a P router, it goes to Step 3. If the next node is the provider edge router, it goes to Step 6.

Step 4: When PE1 receives the ICMP echo-reply packets from the Pn router, it takes the actions as follows:

a) Reading out the VRF name from the ICMP payload;

b) Reading out the IP destination address (Src-C) from the IP header;

c) Checking the forwarding information base (FIB) of the VRF and find the outgoing interface to go to the Src-C and redirect the ICMP the packet to that interface.

d) Going to Step 5.

Step 5: When CE1 receives the packets from PE1, it reads out the Src-C and finds out it is the termination. It also finds out the related time cost from the echo-reply message.

Step 6: When PE2 receives the ICMP echo-request message, it set the Src-C as the destination IP address and its loopback IP address as the source IP address in the echo-reply message. It directly sends the echo-reply message to the CE1.

Hereto, a complete processing of TTL = *n* traceroute initiated by the enterprise customer is finished.

9 Reference model of optimized traceroute for joint IP/MPLS

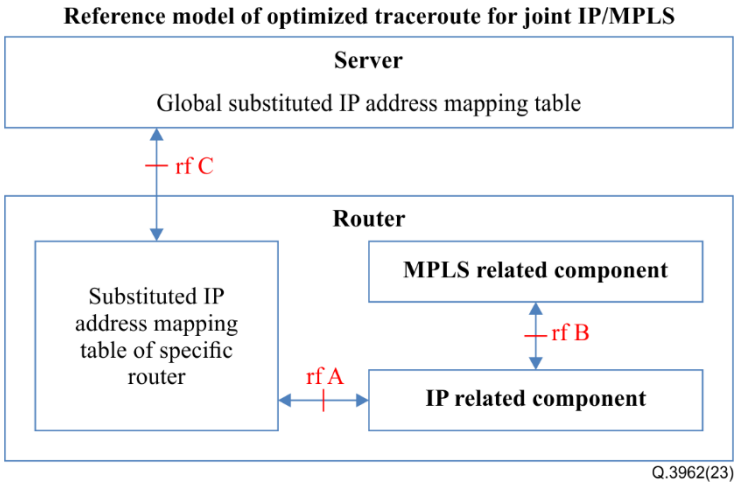

Figure 9-1 – Reference model of optimized traceroute for joint IP/MPLS

As shown in Figure 9-1, the reference model of the optimized traceroute for joint IP/MPLS is composed of four components:

1) MPLS related component. This is responsible for general MPLS operations and additionally copies the IP TTL value to the MPLS TTL value.

2) IP related component. This is responsible for general IP operations and additionally transferring the IP TTL value to the MPLS related component.

3) Substituted IP address mapping table of the specific router: This is responsible for maintaining the substituted IP address mapping table of the specific router.

4) Global substitute IP address mapping table: This is responsible for maintaining the substituted IP addresses mapping table of each router.

The reference model also includes three reference points:

1) Reference point rf A: It is located between substituted IP address mapping table of specific router component and IP related component. It is responsible for transferring substituted IP address for the specific router to the IP related component for the purpose of further IP header encapsulation.

2) Reference point rf B: It is located between MPLS related component and IP related component. It is responsible for transferring IP TTL value to the MPLS related component.

3) Reference point rf C: It is located between global substituted IP address mapping table and substitute IP address mapping table of the specific router. This is responsible for transferring the updating and alignment information of substituted IP address of the specific router and the global substituted IP address mapping table.

Appendix I

An example of optimized method for route tracing of an enterprise network

(This appendix does not form an integral part of this Recommendation.)

This appendix provides an example based on the rules of optimized method of route tracing of enterprise network, which is described in clause 8.

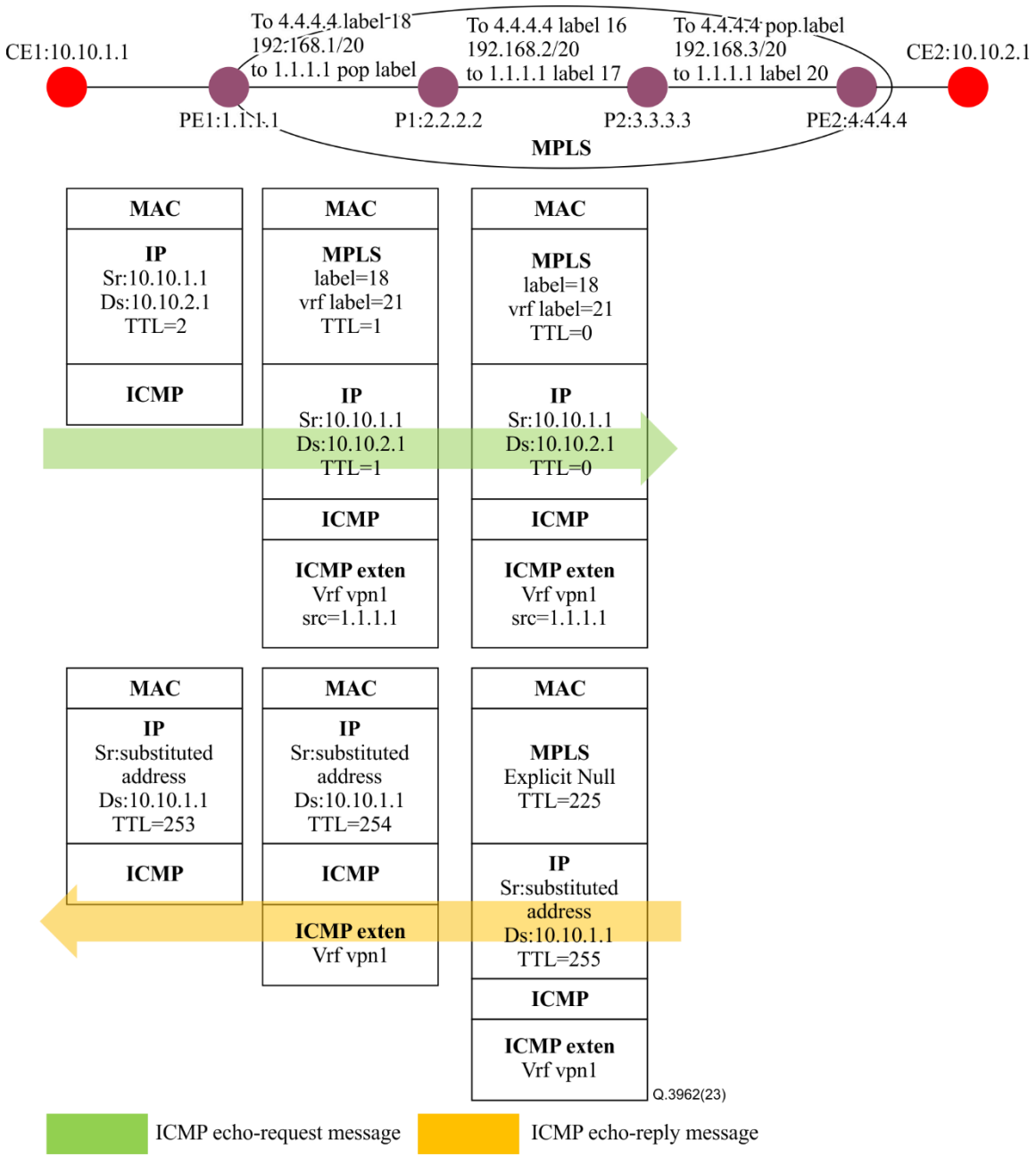

Figure I.1 – The example of optimized ICMP pattern of TTL = 2 for end-to-end enterprise network

In the topology shown in Figure I.1, the ICMP traceroute packet which has TTL = 2 in the IP header is triggered from CE1 to CE2.

When ICMP echo-request packets arrive at PE1, PE1 will check the source IP address 10.10.1.1 from the VRF routing table, and find it is mapped to VRF "vpn1". Then PE1 will find out:

– The VRF VPN1 endpoint is PE2(4.4.4.4) and the related outside outgoing label = 18.
– The VRF vpn1 outgoing label = 21. This label is also known as the inside label.
– The VRF vpn1 incoming label = 35. This label is also known as the inside label.

PE1 encapsulate the ICMP echo-request packets with:

– Adding the outside label = 18 to MPLS header;
– Adding the inside outgoing VRF label = 21 to the MPLS header;
– Copying the TTL = 1 from IP header to the MPLS header;
– Remaining the TTL = 1 of the IP header;
– Adding the VRF name "vpn1" inside to the ICMP payload;
– Adding the source end of MPLS VPN tunnel 1.1.1.1(PE1) to the ICMP payload.

When the restructured packets arrive at P1:

– The TTL of the MPLS layer will minus 1 and decrease to 0, so P1 will generate an ICMP error message (echo-reply message);
– P1 will first check the source end of MPLS VPN tunnel 1.1.1.1 carried in ICMP payload. Then P1 checks the LFIB and finds out 1.1.1.1 is mapping to label "Explicit Null";
– P1 generates the MPLS header of ICMP error message:
  a) It adds 'Explicit None' to the MPLS outside label field
  b) It sets the TTL = 255.
– P2 generates the IP header of ICMP error message:
  a) It checks out the substituted P1's IP address (e.g., 127.0.0.1) related to the original P1's IP address 2.2.2.2 from a dedicated table maintained in P1, which stores the mapping of these two kinds of addresses;
  b) It sets the substituted P1's IP address as the IP source address in the IP header of the ICMP error message;
  c) It sets the destination IP address to 10.10.1.1, which is read out from the IP header of the ICMP request packet;
  d) It sets the IP TTL = 255.
– P2 generates the ICMP payload and it retains the VRF name "vpn1" in the ICMP payload.

When PE1 receives the packets from P1, it will take the actions as follows:

– It reads out the VPN name "vpn1" from the ICMP payload;
– It reads out the IP destination address 10.10.1.1 from the IP header;
– It checks the FIB of "vpn1" and finds the outgoing interface to go to the destination 10.10.1.1 and redirect the ICMP the packet to that interface.

When CE1 receives the packets from PE1, it reads out the "source IP address" = 127.0.0.1 and the related time cost from this node.

Hereto, a complete processing of TTL = 2 traceroute initiated by the enterprise customer is finished.

Through this optimized method, P1 is able to generate the ICMP packets in a better way to make a direct throw between the tested node and source node. Consequently, an accurate traceroute statistic could be collected, and network resources (such as bandwidth) are saved.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	Tariff and accounting principles and international telecommunication/ICT economic and policy issues
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Environment and ICTs, climate change, e-waste, energy efficiency; construction, installation and protection of cables and other elements of outside plant
Series M	Telecommunication management, including TMN and network maintenance
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	**Switching and signalling, and associated measurements and tests**
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks, open system communications and security
Series Y	Global information infrastructure, Internet protocol aspects, next-generation networks, Internet of Things and smart cities
Series Z	Languages and general software aspects for telecommunication systems

Published in Switzerland
Geneva, 2024