PCE Working Group
Internet-Draft
Intended status: Experimental
Expires: 8 January 2026

Y. Wang
A. Wang
China Telecom
B. Khasanov
MTS Web Services (MWS)
F. Qin
China Mobile
H. Chen
Futurewei
C. Zhu
ZTE Corporation

25 July 2025

PCEP Procedures and Extension for VLAN-based Traffic Forwarding
draft-wang-pce-vlan-based-traffic-forwarding-13

Abstract

This document defines the Path Computation Element Communication Protocol (PCEP) extension for VLAN-based traffic forwarding in native IP network and describes the essential elements and key procedures of the data packet forwarding system based on VLAN info to accomplish the End to End (E2E) traffic assurance for VLAN-based traffic forwarding in native IP network.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 8 March 2025.

Table of Contents

1. Introduction

[RFC8283] introduces the extension to the architecture described in [RFC4655] where PCE acts as a central controller and gets more responsibility for LSP provisioning on each hop. Based on such mechanism, the PCE can calculate the optimal path for various applications and send the instructions to the network equipment via PCEP protocol, thus control the packet forwarding and achieve the QoS assurance for prioritized traffic. .

[RFC8735] describes the scenarios of QoS assurance for hybrid cloud-based application within one domain and traffic engineering in multi-domains. It proposes also the following requirements for the potential solution:

1. Should be applied both in native IPv4 and IPv6 environment.

2. Should be same procedures for the intra-domain and inter-domain scenario.

3. Should utilize the existing forwarding capabilities of the deployed network devices.

Due to large scale of Ethernet interfaces used in operators network and need to establish E2E QoS assurance for them, an operator should currently use either VPWS or EVPN with MPLS signaling. This is not suitable for Native IP scenarios. Thus PCECC architecture can solve that problem for Native IP networks by building the end-to-end dedicated path based on a VLAN header to control the forwarding behavior of a packet. Similar with the PCECC for LSP [RFC9050], this document defines a Path Computation Element Communication Protocol (PCEP) Extension for VLAN-based traffic forwarding by using the VLAN info contained in the Ethernet frame in native IP network and the mechanism is actually the PCECC for VSP(VLAN Switching Path). It is an end to end traffic guarantee mechanism based on the PCEP protocol in the native IP environment, which can ensure the connection-oriented network communication. The overall QoS assurance effect is achieved via the central controller by calculating and deploying the optimal VSP to bypass the congested nodes and links, thus avoids the resource reservation on each nodes in advance.

Compared with other traffic assurance technologies such as MPLS or SRv6 which is supported only in IPv6 environment and has the obvious packet overhead problems, the VLAN-based traffic forwarding (VTF) mechanism uses a completely new address space which will not conflict with other existing protocols and can easily avoid these problems and be deployed in IPv4 and IPv6 environment simultaneously. It is suitable for IPv4 and IPv6 networks and can leverage the existing PCE technologies as much as possible.

2. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119] .

3. Terminology

The following terminology is used in this document:

* PCC: Path Computation Client, defined in [RFC5440]

* PCE: Path Computation Element, defined in [RFC5440]

* PCEP: PCE Communication Protocol, defined in [RFC5440]

* PCECC: PCE-based Central Controller, defined in[RFC5440]

* SRP: Stateful PCE Request Parameters, defined in [RFC8231]

* LSP: Lable Switching Path, defined in [RFC5440]

* PST: Path Setup Type, defined in[RFC9050]

* CCI: Central Controller Instructions, defined in [RFC9050]

The following terms are defined in this draft:

* VFCCI: VLAN Forwarding Central Controller Instructions

* VCCCI: VLAN Crossing Central Controller Instructions

4. Capability Advertisement

During the PCEP Initialization Phase, PCEP Speakers (PCE or PCC) advertise their support of VLAN-based traffic forwarding extensions. This document defines a new Path Setup Type (PST)[RFC8408] for PCECC, as follows:

* PST=TBD1: Path is a VLAN-based traffic forwarding type.

A PCEP speaker MUST indicate its support of the function described in this document by sending a PATH-SETUP-TYPE-CAPABILITY TLV in the OPEN object with this new PST included in the PST list.

Because the path is set up through PCE, a PCEP speaker MUST advertise the PCECC capability by using PCECC-CAPABILITY sub-TLV which is used to exchange information about their PCECC capability as per PCEP extensions defined in [RFC9050]

A new flag is defined in PCECC-CAPABILITY sub-TLV for VLAN-based traffic forwarding.

V (VLAN-based-forwarding-CAPABILITY - 1 bit - TBD2): If set to 1 by a PCEP speaker, it indicates that the PCEP speaker supports the capability of VLAN based traffic forwarding as specified in this document. The flag MUST be set by both the PCC and PCE in order to support this extension.

If a PCEP speaker receives the PATH-SETUP-TYPE-CAPABILITY TLV with the newly defined path setup type, but without the V bit set in PCECC-CAPABILITY sub-TLV, it MUST:

* Send a PCErr message with Error-Type=10(Reception of an invalid object) and Error-Value TBD3(PCECC VLAN-based-forwarding-CAPABILITY bit is not set).

* Terminate the PCEP session

5. PCEP message

As per [RFC8281] ,the PCInitiate message sent by a PCE was defined to trigger LSP instantiation or deletion with the SRP and LSP object included during the PCEP initialization phase. The Path Computation LSP State Report message (PCRpt message) was defined in [RFC8231], which is used to report the current state of a LSP. A PCC can send a LSP State Report message in response to a LSP instantiation. Besides, the message can either in response to a LSP Update Request from a PCE or asynchronously when the state of a LSP changes .

[RFC9050] defines an object called Central Controller Instructions (CCI) to specify the forwarding instructions to the PCC. During the coding procedure used for central controller instructions, the CCI object contains the label information and is carried within PCInitiate or PCRpt message.

This document specify two new CCI object-types for VLAN-based traffic forwarding in the Native IP network and are said to be mandatory in a PCEP message when the object MUST be included and are considered to be valid. In addition, this document extends the PCEP message to handle the VLAN-based traffic forwarding path in the native IP network with the new CCI object.

5.1. The PCInitiate message

The PCInitiate message[RFC8281] extended in[RFC9050] can be used to download or remove labels by using the CCI Object.

Based on the extended PCInitiate message and PCRpt described in [I-D.ietf-pce-pcep-extension-native-ip], the BGP Peer Info (BPI) Object and the Peer Prefix Association (PPA) Object are used to establish multi BGP sessions and advertise route prefixes among different BGP sessions before setting up a VLAN-based traffic forwarding path.

This document extends the PCInitiate message as shown below:

```
     <PCInitiate Message> ::= <Common Header>
                                <PCE-initiated-lsp-list>
        Where:
           <Common Header> is defined in [RFC5440]

           <PCE-initiated-lsp-list> ::= <PCE-initiated-lsp-request>
                                        [<PCE-initiated-lsp-list>]

           <PCE-initiated-lsp-request> ::=
                                (<PCE-initiated-lsp-instantiation>|
                                 <PCE-initiated-lsp-deletion>|
                                 <PCE-initiated-lsp-central-control>)

           <PCE-initiated-lsp-central-control> ::= <SRP>
                                                   <LSP>
                                                   <cci-list>

           <cci-list> ::=  <new-CCI>
                           [<BPI>|<PPA>]
                           [<cci-list>]
```

Where:

<cci-list> is as per [RFC9050].
<PCE-initiated-lsp-instantiation> and <PCE-initiated-lsp-deletion> are as per [RFC8281].
<BPI> and <PPA> are as per [draft-ietf-pce-pcep-extension-native-ip]

When PCInitiate message is used to create VLAN-based forwarding instructions, the SRP, LSP and CCI objects MUST be present. The error handling for missing SRP, LSP or CCI object is as per [RFC9050]. Further only one of BPI, PPA or one type of CCI objects MUST be present. If none of them are present, the receiving PCE MUST send a PCErr message with Error-type=6 (Mandatory Object missing) and Error-value=TBD4 (VLAN-based forwarding object missing). If there are more than one of BPI, PPA or more than one type of CCI objects, the receiving PCC MUST send a PCErr message with Error-type=19(Invalid Operation) and Error-value=TBD5(Only one of BPI, PPA or one type of the CCI objects for VLAN can be included in this message).

5.2. The PCRpt message

The PCRpt message is used to report the state and confirm the VLAN info that was allocated by the PCE, to be used during the state synchronization phase or as acknowledgement to PCInitiate message.

The format of the PCRpt message is as follows:

```
 <PCRpt Message> ::= <Common Header>
                          <state-report-list>
      Where:

         <state-report-list> ::= <state-report>[<state-report-list>]

         <state-report> ::= (<lsp-state-report>|
                          <central-control-report>)

         <lsp-state-report> ::= [<SRP>]
                             <LSP>
                             <path>

         <central-control-report> ::= [<SRP>]
                                   <LSP>
                                   <cci-list>

        <cci-list> ::=  <new-CCI>
                        [<BPI>|<PPA>]
                        [<cci-list>]
       Where:
         <path> is as per [RFC8231] and the LSP and SRP object are
         also defined in [RFC8231].
         <BPI> and <PPA> are as per
         [draft-ietf-pce-pcep-extension-native-ip]
```

The error handling for missing LSP or CCI object is as per [RFC9050]. Further only one of BPI, PPA or one type of CCI objects MUST be present. If none of them are present, the receiving PCE MUST send a PCErr message with Error-type=6 (Mandatory Object missing) and Error-value=TBD4 (VLAN-based forwarding object missing). If there are more than one of BPI, PPA or more than one type of CCI objects, the receiving PCC MUST send a PCErr message with Error-type=19(Invalid Operation) and Error-value=TBD5(Only one of BPI, PPA or one type of the CCI objects for VLAN can be included in this message).

6. VLAN-based traffic forwarding Procedures

The target deployment environment of VLAN based traffic forwarding mechanism is for both Native IPv4 and IPv6. In such scenarios, the BGP is used for the prefix distribution among underlying devices(PCCs), no MPLS is involved.

In order to set up the VLAN-based traffic forwarding paths for different applications in native IP network, multiple BGP sessions should be deployed between the ingress PCC and egress PCC at the edge of the network respectively.

Based on the business requirements, the PCE calculates the explicit route and sends the route information to the PCCs through PCInitiate messages. When the PCInitiate message is received, the packet to be guaranteed will be labeled with corresponding VLAN tag, that is done by the ingress PCC. The labeled packet will be further sent to the PCC's specific subinterface identified by the VLAN tag and then be forwarded. Similarly, after receive of the PCInitiate message, the packet will be re-labeled with new VLAN tag and then be forwarded by the transit PCC and the egress PCC. The mechanism of allocating and managing VLAN ID by PCC can refer to the appendix.

The whole procedures mainly focused on the end-to-end traffic assurance for key applications so that it can ensure the adequacy of VLAN quantity.

6.1. Multiple BGP Session Establishment Procedures

As described in section 4, multiple BGP sessions should be deployed between the ingress device and egress device at the edge of the network respectively in order to carry information of different applications. As per [I-D.ietf-pce-pcep-extension-native-ip], the PCE should send the BPI (BGP Peer Info) Object to the ingress and egress device with the indicated Peer AS and Local/Peer IP address. The Ingress and egress devices will receive multiple BPI objects to establish sessions with different next hop. The specific procedure is as follows:

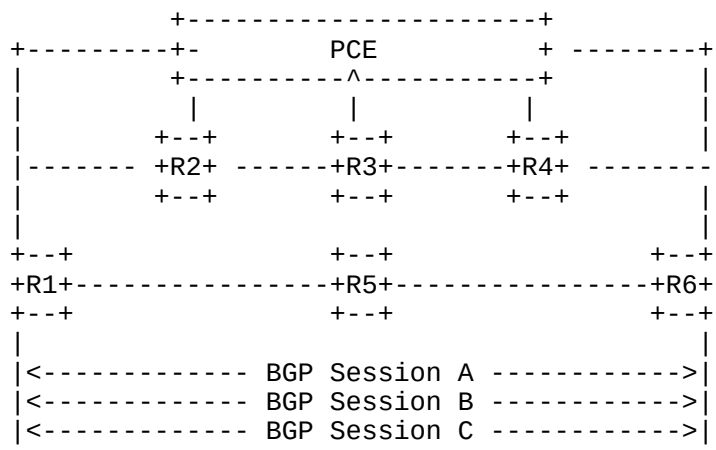

Figure 1: BGP Session Establishment Procedures

6.2. BGP Prefix Advertisement Procedures

The detail procedures for BGP prefix advertisement procedures is introduced in [I-D.ietf-pce-pcep-extension-native-ip], using PCInitiate and PCRpt message pair.

The BGP prefix for different BGP sessions should be sent to the ingress and egress device respectively. The end-to-end traffic for key application can be identified based on these BGP prefix informations and be further assured. As per [I-D.ietf-pce-pcep-extension-native-ip], the PPA(Peer Prefix Association) object with list of prefix sub-objects and the peer address will be sent through the PCInitiate and PCRpt message pair. Through BGP protocol, the ingress device can learn different BGP prefix of the egress device based on the different sessions.

6.3. VLAN mapping info Advertisement Procedures

After the BGP prefix for different BGP session are successfully advertised, information of different applications should be forwarded to different VLAN-based traffic forwarding paths. Based on [RFC8281] and [RFC9050], in order to set up a a PCE-initiated VSP based on the PCECC mechanism, the PCE MUST send the VLAN forwarding CCI (VFCCI) Object with the VLAN-ID included to the ingress PCC and the VLAN crossing CCI (VCCCI) Object to the transit PCC and egress PCC through a PCInitiate message with the PST set to TBD1 in SRP.

6.3.1. VLAN-Based forwarding info Advertisement Procedures

The detail procedures for VLAN-Based forwarding info advertisement contained in the VLAN forwarding CCI Object are shown below, using PCInitiate and PCRpt message pair.

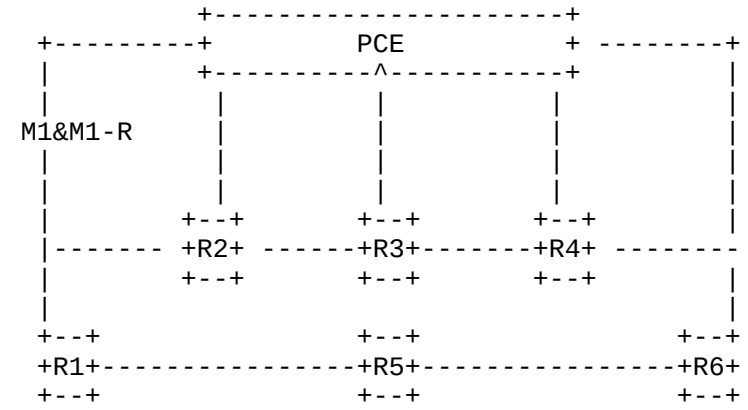

Figure 2: VLAN-Based Forwarding Info Advertisement Procedures for Ingress PCC

The VLAN-forwarding instructions contained in the VFCCI object from the PCE needs to be sent after the initial PCInitiate and PCRpt message exchange with the ingress PCC. On receipt of a PCInitiate message for the PCECC VSP, the PCC responds with a PCRpt message with the status set to 'Going- up', carrying the assigned PLSP-ID and set the D(Delegate) flag and C(Create) flag(see Figure 1).

When the PCE receives this PCRpt message with the PLSP-ID, it assigns VLANs along the path and sets up the path by sending a PCInitiate message to each node along the path of the VSP, as per the PCECC technique. The ingress PCC would receive one VFCCI Object which contains VLAN on the logical subinterface and the Peer IP address. Once the VLAN operations are completed, the PCE MUST send a PCUpd message to the ingress PCC.

The new CCI for the VLAN operations in PCEP are sent via the PCInitiate message by defining a new PCEP object for CCI operations. The fields in the LSP-IDENTIFIERS TLV are described for the RSVP-signaled LSPs but are applicable to the PCECC VSP as well. So the LSP object is included in the PCInitiate message can still be used to identify the PCECC VSP for this instruction and the process is the same.

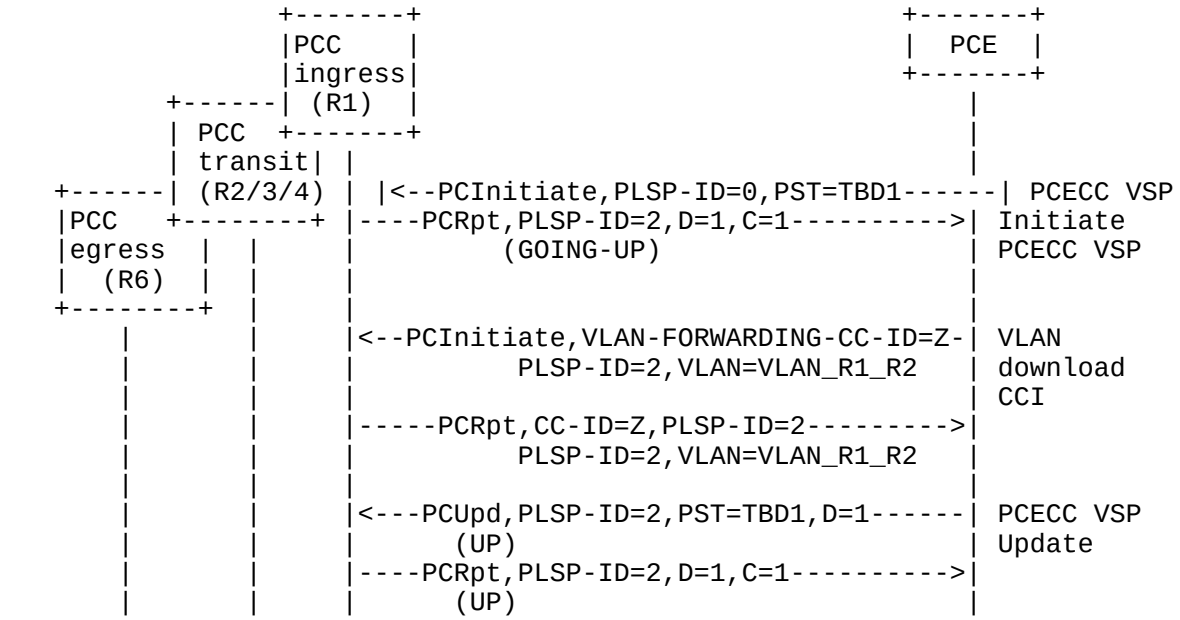

Figure 3: PCE-Initiated VSP On VLAN Forwarding

After the PCC receives the CCI object (with the R bit set to 0 in SRP object) in PCInitiate message, the PCC's subinterface will set up the specific VLAN based on the VFCCI object, source and destination BGP prefix learnt before. When the ingress PCC receives a packet, based on the source and destination IP, the packet that needs to be guaranteed will be matched and then be labeled with corresponding VLAN tag. After that, The labeled packet will be further forwarded to the specific subinterface according to the appendix.

When PCC receives the VFCCI Object with the R bit set to 1 in SRP object in PCInitiate message, the PCC MUST withdraw the VLAN-Based forwarding info advertisement to the peer that indicated by this object.

On receipt of a PCInitiate message for the PCECC VSP, the PCC MUST report the result via the PCRpt messages, with the corresponding SRP and CCI object included.

The message number, message peers, message types and message key parameters in the above figures are shown in the table below:

Table 1: Message Information

No.	Peers	Type	Message Key Parameters
M1 M1-R	PCE/R1	PCInitiate PCRpt	CC-ID=X1(Symbolic Path Name=Class A) VLAN Forwarding CCI Object (Peer_IP=R6_A,Interface_Address=INF1, VLAN_ID=VLAN_R1_R2)

6.3.2. VLAN-Based crossing info Advertisement Procedures

After the VLAN-Based forwarding info advertisement procedures, the PCE needs to send a PCInitiate message to each node along the path to download the VLAN crossing instructions.The detail procedures for VLAN-Based crossing info advertisement contained in the VCCCI Object are shown below, using PCInitiate and PCRpt message pair.

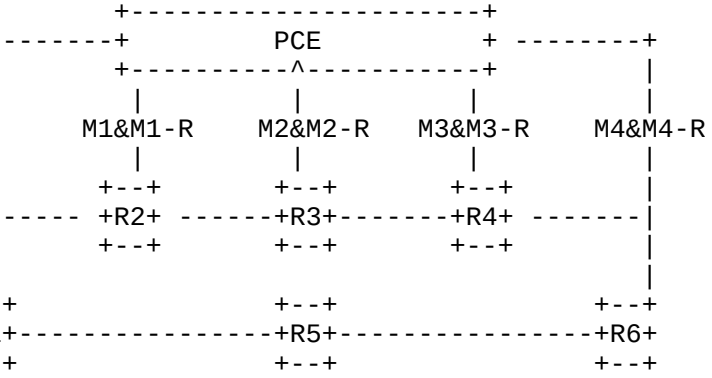

Figure 4: VLAN-Based Crossing Info Advertisement Procedures for Transit PCC and Egress PCC

Similarly, as per the PCECC, The transit PCC would receive two VCCCI Objects with the O bit set for the out-VLAN on the egress subinterface and the O bit unset for the in-VLAN on the ingress subinterface. Similar with the transit PCC, the egress PCC would receive two VCCCI Objects but the out-VLAN on the egress subinterface is set to 0. The in-VLAN tag and an out-VLAN tag in the CCI Objects specifies a new VLAN forwarding path. After the procedure of VLAN-Based forwarding info advertisement mentioned above, the PCC's subinterface will set up the specific VLAN based on the VCCCI Object(with the R bit set to 0 in SRP object) contained in the PCInitiate message. When the transit PCC receives a data packet that has been labeled with VLAN by ingress PCC before, based on matching procedure of the VLAN tag, the in-VLAN tag of this data packet will be replaced by a new out-VLAN tag of the current transit PCC according to the appendix. The packet with the new VLAN tag will be further forwarded to the next hop.

For the egress PCC, the out-VLAN tag MUST be 0 which indicates it is the last hop of the transmission. So the egress PCC will directly remove the in-VLAN tag of the packet and the packet will be forwarded.

When PCC receives the VCCCI Object with the R bit set to 1 in SRP object in PCInitiate message, the PCC MUST withdraw the VLAN-Based crossing info advertisement to the peer that indicated by this object.

On receipt of a PCInitiate message for the PCECC VSP, the PCC MUST report the result via the PCRpt messages, with the corresponding SRP and CCI object included.

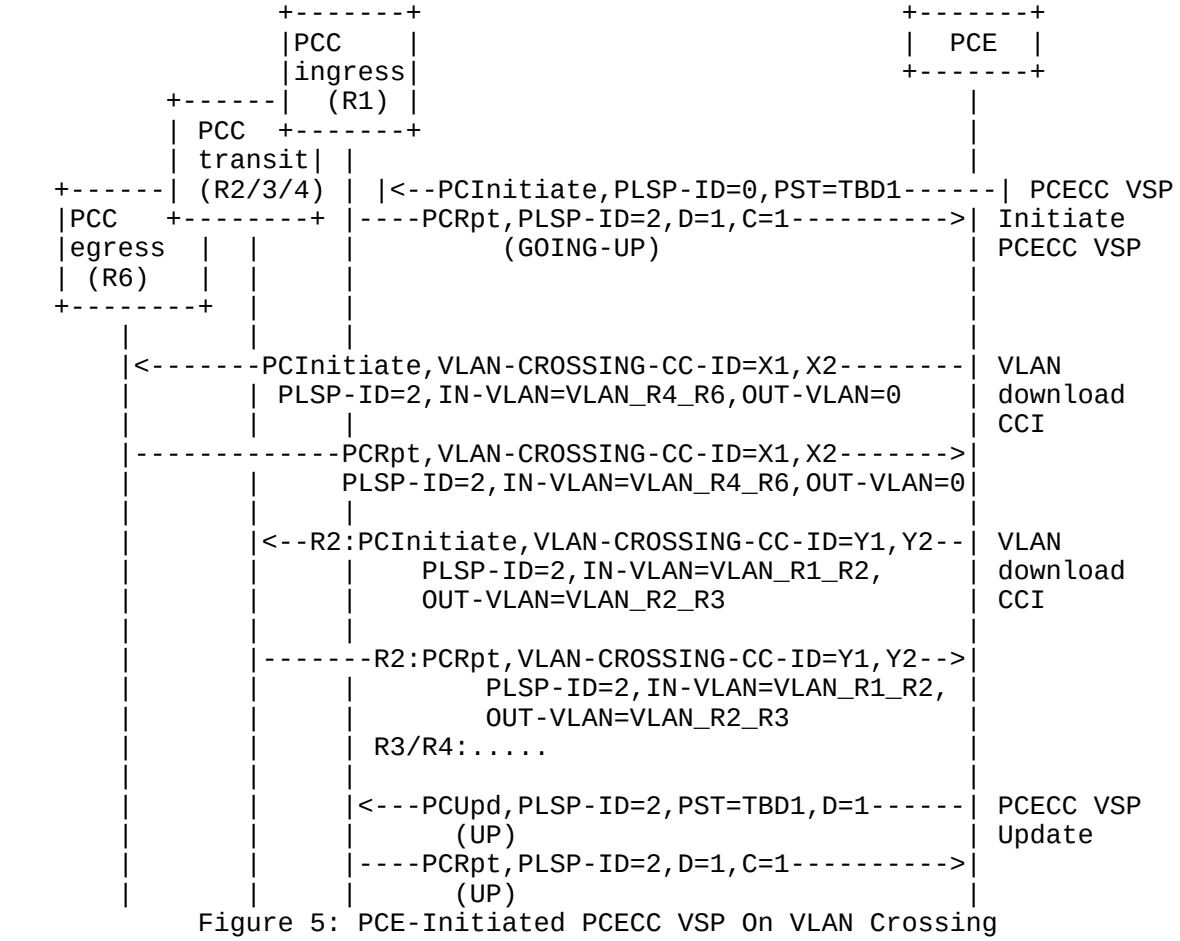

Figure 5: PCE-Initiated PCECC VSP On VLAN Crossing

When the out-VLAN tag conflicts with a pre-defined VLAN tag or the PCC can not set up a VLAN forwarding path with the out-VLAN tag, an error (Error-type=TBD6, VLAN-based forwarding failure, Error-value=TBD7, VCCCI Object peer info mismatch) MUST be reported via the PCRpt message.

The message number, message peers, message type and message key parameters in the above figures are shown in below table:

Table 2: Message Information

No.	Peers	Type	Message Key Parameters
M1 M1-R	PCE/R2	PCInitiate PCRpt	CC-ID=X1(Symbolic Path Name=Class A) VLAN crossing CCI Object(IN) (O=0,Interface_Address=INF1,IN_VLAN_ID=VLAN_R1_R2) VLAN crossing CCI Object(OUT) (O=1,Interface_Address=INF2,OUT_VLAN_ID=VLAN_R2_R3)
M2 M2-R	PCE/R3	PCInitiate PCRpt	CC-ID=X1(Symbolic Path Name=Class A) VLAN crossing CCI Object(IN) (O=0,Interface_Address=INF1,IN_VLAN_ID=VLAN_R2_R3) VLAN crossing CCI Object(OUT) (O=1,Interface_Address=INF2,OUT_VLAN_ID=VLAN_R3_R4)
M3 M3-R	PCE/R4	PCInitiate PCRpt	CC-ID=X1(Symbolic Path Name=Class A) VLAN crossing CCI Object(IN) (O=0,Interface_Address=INF1,IN_VLAN_ID=VLAN_R3_R4) VLAN crossing CCI Object(OUT) (O=1,Interface_Address=INF2,OUT_VLAN_ID=VLAN_R4_R6)
M4 M4-R	PCE/R6	PCInitiate PCRpt	CC-ID=X1(Symbolic Path Name=Class A) VLAN crossing CCI Object(IN) (O=0,Interface_Address=INF1,IN_VLAN_ID=VLAN_R4_R6) VLAN crossing CCI Object(OUT) (O=1,Interface_Address=INF2,OUT_VLAN_ID=0)

7. New PCEP Objects

The Central Control Instructions (CCI) Object is used by the PCE to specify the forwarding instructions is defined in [RFC9050]. This document defines two other CCI object-types for VLAN-based traffic forwarding. All new PCEP objects are compliant with the PCEP object format defined in [RFC5440].

7.1. VLAN forwarding CCI Object

The VLAN forwarding CCI Object is used to set up the specific VLAN forwarding path including the logical subinterface that will be used for traffic forwarding to the specific hop. Combined with this type of CCI Object and the Peer Prefix Association object(PPA) defined in [I-D.ietf-pce-pcep-extension-native-ip], the ingress PCC will identify the traffic that needs to be protected. This object MUST only be included and sent to the ingress PCC of the end2end path.

CCI Object-Class is 44.

CCI Object-Type is TBD8 for VLAN forwarding info in the native IP network.

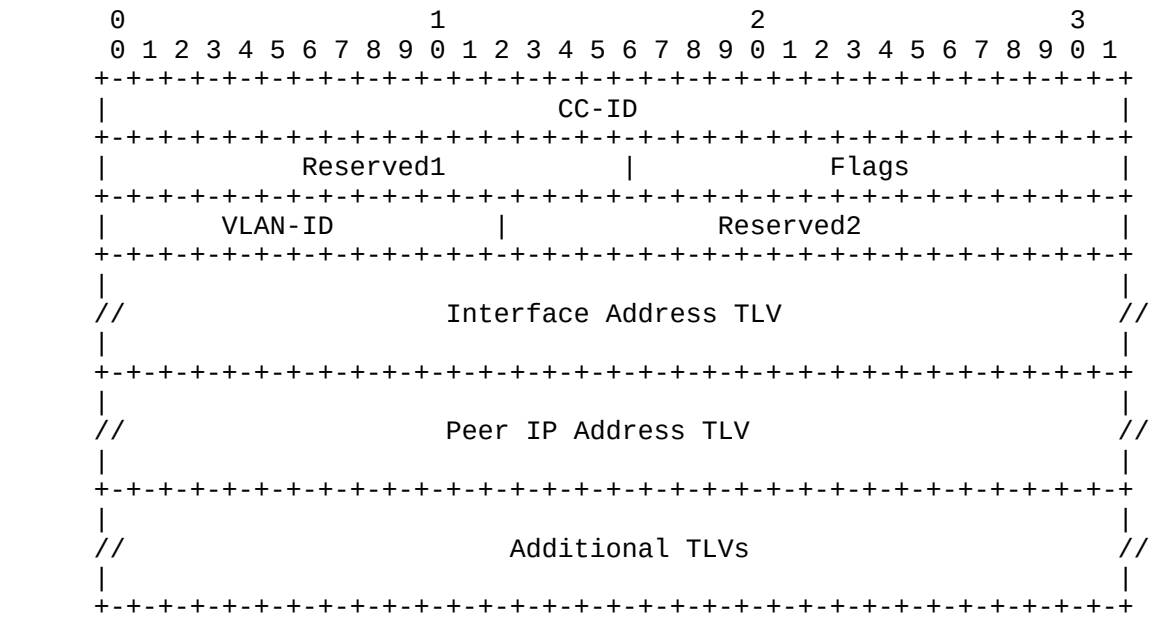

Figure 6: VLAN Forwarding CCI Object

The fields in the CCI object are as follows:

CC-ID: is as described in [RFC9050]. Following fields are defined for CCI Object-Type TBD8.

Reserved1(16 bits): is set to zero while sending, ignored on receipt.

Flags(16 bits): is used to carry any additional information pertaining to the CCI. Currently no flag bits are defined.

VLAN ID(12 bits):the ID of the VLAN forwarding path that the PCC will set up on its logical subinterface in order to transfer the packet to the specific hop.

Reserved2(20 bits): is set to zero while sending, ignored on receipt.

Interface Address TLV [RFC8779] MUST be included in this CCI Object-Type TBD8 to specify the interface which will set up the VLAN defined in the VFCCI Object.

The Peer IP Address TLV [RFC8779]MUST be included in this CCI Object-Type TBD8 to identify the end to end TE path in VLAN-based traffic forwarding network and MUST be unique.

7.2. Address TLVs

[RFC8779] defines IPV4-ADDRESS, IPV6-ADDRESS, and UNNUMBERED-ENDPOINT TLVs for the use of Generalized Endpoint. The same TLVs can also be used in the CCI object to find the Peer address that matches egress PCC and further identify the packet to be guaranteed. If the PCC is not able to resolve the peer information or can not find the corresponding ingress device, it MUST reject the CCI and respond with a PCErr message with Error-Type = TBD6 ("VLAN-based forwarding failure") and Error Value = TBD9 ("Invalid egress PCC information").

7.3. VLAN crossing CCI Object

The VLAN crossing CCI object is defined to control the transmission-path of the packet by VLAN-ID. This new type of CCI Object can be carried within a PCInitiate message sent by the PCE to the transit PCC and the egress PCC in the VLAN-based traffic forwarding scenarios.

CCI Object-Class is 44.

CCI Object-Type is TBD10 for VLAN crossing info in the native IP network.

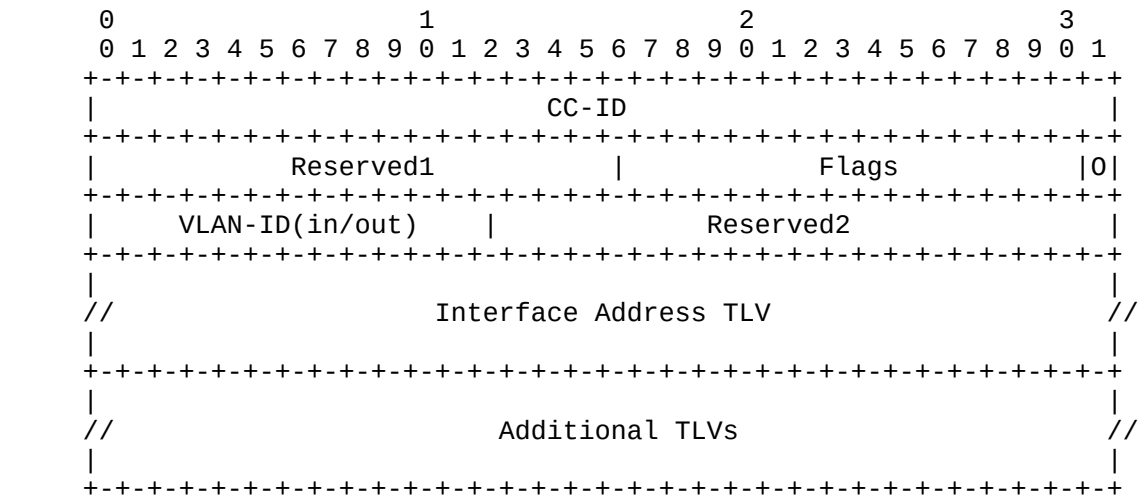

Figure 7: VLAN Crossing CCI Object

CC-ID: is as described in [RFC9050]. Following fields are defined for CCI Object-Type TBD10.

Reserved1(16 bits): is set to zero while sending, ignored on receipt.

Flags(16 bits): is used to carry any additional information pertaining to the CCI. Currently, the following flag bit are defined:

* O bit (out-label) : If the bit is set to '1', it specifies the VLAN is the out-VLAN, and it is mandatory to encode the egress interface information(via Interface Address TLVs in the CCI object). If the bit is not set or set to '0', it specifies the VLAN is the in-VLAN, and it is mandatory to encode the ingress interface information.

VLAN ID(12 bits): The ID of the VLAN switching path. When the O bit is set to 0, the VLAN is the in-VLAN and the ID indicates a VLAN forwarding path which is used to identify the traffic that needs to be protected. When the O bit is set to 1, the VLAN is the out-VLAN and it indicates the ID of the VLAN forwarding path that the PCC will set up on its logical subinterface in order to transfer the packet labled with this VLAN ID to the specific hop. To the transit PCC, the value MUST not be 0 to indicate it is not the last hop of the VLAN-based traffic forwarding path. To the egress PCC, the value MUST be 0 to indicate it is the last hop of the VLAN-based traffic forwarding path.

Reserved2(8 bits): is set to zero while sending, ignored on receipt.

Interface Address TLV [RFC8779] MUST be included in this CCI Object-Type TBD8 to specify the interface which will set up the VLAN defined in the VCCCI Object.

8. IANA Considerations

8.1. Path Setup Type Registry

[RFC8408] created a sub-registry within the "Path Computation Element Protocol (PCEP) Numbers" registry called "PCEP Path Setup Types". IANA is requested to allocate a new code point within this registry, as follows:

Value	Description	Reference
TBD1	VLAN-Based Traffic Forwarding Path	This document

8.2. PCECC-CAPABILITY sub-TLV's Flag field

[RFC9050] created a sub- registry within the "Path Computation Element Protocol (PCEP) Numbers" registry to manage the value of the PCECC-CAPABILITY sub- TLV's 32-bits Flag field. IANA is requested to allocate a new bit position within this registry, as follows:

Value	Description	Reference
TBD2(V)	VLAN-Based Forwarding CAPABILITY	This document

8.3. PCEP Object Types

IANA is requested to allocate new registry for the PCEP Object Type:

Object-Class Value	Name	Reference
44	CCI Object-Type TBD8: VLAN forwarding CCI TBD10: VLAN crossing CCI	This document

8.4. PCEP-Error Object

IANA is requested to allocate new error types and error values within the "PCEP-ERROR Object Error Types and Values" sub-registry of the PCEP Numbers registry for the following errors:

Error-Type	Meaning	Error-value	Reference
6	Mandatory Object missing	TBD4:VLAN-based forwarding object missing	This document
10	Reception of an invalid object	TBD3:PCECC VLAN-based-forwarding -CAPABILITY bit is not set	This document
19	Invalid Operation	TBD5: Only one of BPI, PPA or one type of the CCI objects for VLAN can be included in this message	This document
TBD6	VLAN-based forwarding failure	TBD7: VLAN crossing CCI Object peer info mismatch	This document
		TBD9: Invalid egress PCC information	This document

9. Normative References

[I-D.ietf-pce-pcep-extension-for-pce-controller]
Li, Z., Peng, S., Negi, M. S., Zhao, Q., and C. Zhou, "Path Computation Element Communication Protocol (PCEP)

Procedures and Extensions for Using the PCE as a Central Controller (PCECC) of LSPs", Work in Progress, Internet-Draft, draft-ietf-pce-pcep-extension-for-pce-controller-14, 5 March 2021, <https://datatracker.ietf.org/doc/html/draft-ietf-pce-pcep-extension-for-pce-controller-14>.

[I-D.ietf-pce-pcep-extension-native-ip]
Wang, A., Khasanov, B., Fang, S., Tan, R., and C. Zhu, "Path Computation Element Communication Protocol (PCEP) Extensions for Native IP Networks", Work in Progress, Internet-Draft, draft-ietf-pce-pcep-extension-native-ip-39, 1 September 2024, <https://datatracker.ietf.org/doc/html/draft-ietf-pce-pcep-extension-native-ip-39>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC4655] Farrel, A., Vasseur, J.-P., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, DOI 10.17487/RFC4655, August 2006, <https://www.rfc-editor.org/info/rfc4655>.

[RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, DOI 10.17487/RFC5440, March 2009, <https://www.rfc-editor.org/info/rfc5440>.

[RFC8231] Crabbe, E., Minei, I., Medved, J., and R. Varga, "Path Computation Element Communication Protocol (PCEP) Extensions for Stateful PCE", RFC 8231, DOI 10.17487/RFC8231, September 2017, <https://www.rfc-editor.org/info/rfc8231>.

[RFC8281] Crabbe, E., Minei, I., Sivabalan, S., and R. Varga, "Path Computation Element Communication Protocol (PCEP) Extensions for PCE-Initiated LSP Setup in a Stateful PCE Model", RFC 8281, DOI 10.17487/RFC8281, December 2017, <https://www.rfc-editor.org/info/rfc8281>.

[RFC8283] Farrel, A., Ed., Zhao, Q., Ed., Li, Z., and C. Zhou, "An Architecture for Use of PCE and the PCE Communication Protocol (PCEP) in a Network with Central Control", RFC 8283, DOI 10.17487/RFC8283, December 2017, <https://www.rfc-editor.org/info/rfc8283>.

[RFC8408] Sivabalan, S., Tantsura, J., Minei, I., Varga, R., and J. Hardwick, "Conveying Path Setup Type in PCE Communication Protocol (PCEP) Messages", RFC 8408, DOI 10.17487/RFC8408, July 2018, <https://www.rfc-editor.org/info/rfc8408>.

[RFC8735] Wang, A., Huang, X., Kou, C., Li, Z., and P. Mi, "Scenarios and Simulation Results of PCE in a Native IP Network", RFC 8735, DOI 10.17487/RFC8735, February 2020, <https://www.rfc-editor.org/info/rfc8735>.

[RFC8779] Margaria, C., Ed., Gonzalez de Dios, O., Ed., and F. Zhang, Ed., "Path Computation Element Communication Protocol (PCEP) Extensions for GMPLS", RFC 8779, DOI 10.17487/RFC8779, July 2020, <https://www.rfc-editor.org/info/rfc8779>.

[RFC9050] Li, Z., Peng, S., Negi, M., Zhao, Q., and C. Zhou, "Path Computation Element Communication Protocol (PCEP) Procedures and Extensions for Using the PCE as a Central Controller (PCECC) of LSPs", RFC 9050, DOI 10.17487/RFC9050, July 2021, <https://www.rfc-editor.org/info/rfc9050>.

Appendix A. Dataplane Operations for VLAN Switching

A.1. VLAN Functional Categorization

According to IEEE 802.1Q, VLAN-enabled ports are generally categorized in one of two ways: tagged or untagged. During the VLAN switching procedure, the usage of VLAN can be further refined into three categories:

VLAN of ingress interface: A VLAN originally tagged by the devices which is used for VLAN-based flow labeling behavior. The Ingress VLAN refers to the initial VLAN tagging performed by devices when they send traffic into the network. It helps identify and manage the flow of traffic as it enters the network.

VLAN of transit interface: A VLAN transporting transit traffic, in other words, the traffic tagged with transit VLAN does not have the final source or destination. The Transit VLAN facilitates the movement of traffic between different segments of the network, ensuring efficient routing to reach its ultimate destination.

VLAN of egress interface: A VLAN through which traffic exits a network, and the devices within the network may remove the VLAN tag before sending the traffic to its final destination. The Egress VLAN

handles traffic leaving the network from end-user devices. Devices within this VLAN perform necessary operations, such as VLAN tag removal or other processing, to ensure that outgoing traffic is appropriately formatted for the next segment of its journey in the network.

A.2. VLAN switching process

A.2.1. VLAN-Forwarding routing table

Based on the three categories of VLANs above, the ingress devices needs to maintain a VLAN-Forwarding routing table(VFR) table defined below which is used to match the packet based on the source and destination IP.

Table 1: VLAN-Forwarding routing table

Src IP Prefix	Dst IP Prefix	Interface	VLAN
Source IP_A	Destination IP_A	INF X	VLAN_X
Source IP_B	Destination IP_B	INF Y	VLAN_Y
	...		

The source and destination IP address is used to identify the end to end TE path in VLAN-based traffic forwarding network. The VLAN indicates a VLAN forwarding path which is used to mark the traffic that needs to be protected. Through the VLAN in the VFR table, a VLAN forwarding path will be set up on its logical subinterface in order to transfer the packet to the specific hop.

A.2.2. VLAN-Crossing routing table

Accordingly, the egress devices and transit devices need to maintain a VLAN-Crossing routing table(VCR) table. The VLAN IDs of inbound and outbound form a key-value pair which indicates a new VSP. The interface addresses indicate the inbound and out bound sub-interface addresses that carries the specific service traffic which needs to be guaranteed. The in-VLAN is used to identify the traffic that needs to be protected while the out-VLAN indicates the ID of the VLAN forwarding path that the device will set up on its logical subinterface in order to transfer the packet labled with this VLAN ID to the specific hop. To the transit device, the value MUST not be 0 to indicate it is not the last hop of the VLAN-based traffic forwarding path. To the egress device, the value MUST be 0 to indicate it is the last hop of the VLAN-based traffic forwarding path. Through the mapping of the in-VLAN and the out-VLAN in the

table, the data packet will be transferred to the specific interface and be switched on the out VLAN for the transit devices or 0 for the egress devices.

Table 2: VLAN-Crossing routing table

In-Interface	In-VLAN	Out-Interface	Out-VLAN
INF1	VLAN_X1_X2	INF2	VLAN_X2_X3
INF3	VLAN_Y1_Y2	INF4	VLAN_Y2_Y3
INF5	VLAN_Z1_Z2	INF6	0
	...		

A.2.3. VLAN-based traffic switching Procedures

In order to implement VLAN switching, the routers and switches MUST support VLANs. Based on the business requirements, the packet to be guaranteed will be labeled with corresponding VLAN tag. The tag may be added or removed by a host, a router, or a switch which is out of the scope of this document. The labeled packet will be further sent to the router's specific subinterface identified by the VLAN tag and then be forwarded.

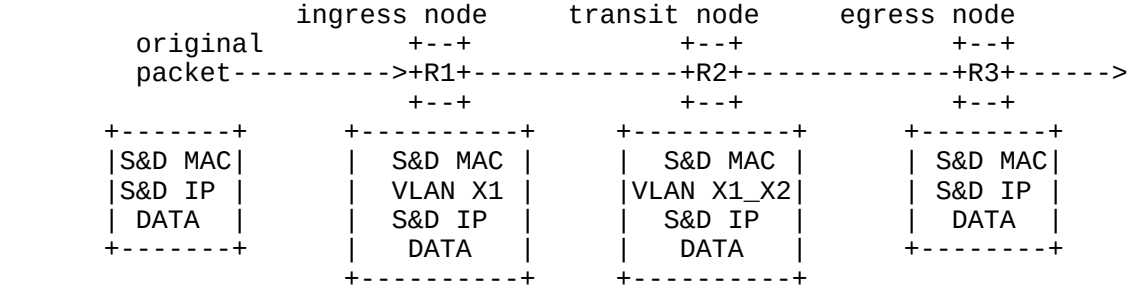

Figure 1: Data Packet Encapsulation Procedure

Figure8 shows the data packet encapsulation procedure of VLAN switching operations. As a ingress node, R1 maintains a VFR table shown in the table 3 below. Similarly, as a transit node and an egress node, R2 and R3 maintain a VCR routing table shown in the table 4&5 below separately. Based on these tables, the specific VLAN will be set up on their sub-interfaces. When the ingress node R1 receives a packet, based on the source and destination IP, the packet that needs to be guaranteed will hit the first entry in the routing table and then be labeled with corresponding VLAN tag VLAN_X1.

Table 3: VLAN-Forwarding routing table to R1

Src IP Address	Dst IP Address	Interface	VLAN
Source IP_A	Destination IP_A	INF X	VLAN_X1
Source IP_B	Destination IP_B	INF Y	VLAN_Y1
	...		

After that, The labeled packet will be further forwarded to the specific subinterface specified by VLAN. When the data packet tagged with VLAN_X1 which is done by R1 is delivered to R2, it will look up the VCR table via tagged VLAN. If the VLAN is consistent, the in-VLAN as VLAN_X1 will be replaced with a out-VLAN as VLAN_X1X2 by the current transit node R2. The packet labeled with new VLAN will be further delivered to the next hop.

Table 4: VLAN-Crossing routing table to R2

In-Interface	In-VLAN	Out-Interface	Out-VLAN
INF1	VLAN_X1	INF2	VLAN_X1_X2
INF3	VLAN_Y1	INF4	VLAN_Y1_Y2
INF5	VLAN_Z1	INF6	0
	...		

R3, as the egress node, its out-VLAN in the VCR table MUST be 0 which indicates it' s the final hop in the whole transit procedure. Therefore, the egress node will strip the in-VLAN and the packet will be transited directly.

Table 5: VLAN-Crossing routing table to R3

In-Interface	In-VLAN	Out-Interface	Out-VLAN
INF1	VLAN_X1_X2	INF2	0
INF3	VLAN_Y1_Y2	INF4	VLAN_Y2_Y3
INF5	VLAN_Z1_Z2	INF6	VLAN_Z2_Z3
	...		

Based on the VFR table and VCR table, the original packet will be transmitted along the path of the VSP through the exchange of VLAN labels. Via calculating and deploying the optimal VSP by the central controller, the overall QoS assurance effect is achieved, and there is no more need to reserve resources for physical links in advance.

Authors' Addresses

Yue Wang
China Telecom
Beiqijia Town, Changping District
Beijing
Beijing, 102209
China
Email: wangy73@chinatelecom.cn

Aijun Wang
China Telecom
Beiqijia Town, Changping District
Beijing
Beijing, 102209
China
Email: wangaj3@chinatelecom.cn

Boris Khasanov
MTS Web Services (MWS)
Vorontsovskaya street, Bldg. 4
Moscow
Email: bhassanov@yahoo.com

Fengwei Qin
China Mobile
32 Xuanwumenxi Ave.
Beijing
100032
China
Email: qinfengwei@chinamobile.com

Huaimo Chen
Futurewei
Boston,
United States of America
Email: Huaimo.chen@futurewei.com

Chun Zhu
ZTE Corporation
50 Software Avenue, Yuhua District
Nanjing
Jiangsu, 210012
China
Email: zhu.chun1@zte.com.cn