IETF 110 Path Computation Element (PCE) WG

Wednesday, March 10, 2021

Chairs
Julien Meuric (julien.meuric@orange.com)
Dhruv Dhody (dd@dhruvdhody.com)
Secretary
Hariharan Ananthakrishnan (hari@netflix.com)

This session is being recorded

IETF 110
hosted by

Google

Internet Engineering Task Force
© 2021 IETF Trust
Production by Meetecho

IETF

Note Well

This is a reminder of IETF policies in effect on various topics such as patents or code of conduct. It is only meant to point you in the right direction. Exceptions may apply. The IETF's patent policy and the definition of an IETF "contribution" and "participation" are set forth in BCP 79; please read it carefully.

As a reminder:

- By participating in the IETF, you agree to follow IETF processes and policies.
- If you are aware that any IETF contribution is covered by patents or patent applications that are owned or controlled by you or your sponsor, you must disclose that fact, or not participate in the discussion.
- As a participant in or attendee to any IETF activity you acknowledge that written, audio, video, and photographic records of meetings may be made public.
- Personal information that you provide to IETF will be handled in accordance with the IETF Privacy Statement.
- As a participant or attendee, you agree to work respectfully with other participants; please contact the ombudsteam (https://www.ietf.org/contact/ombudsteam/) if you have questions or concerns about this.

Definitive information is in the documents listed below and other IETF BCPs. For advice, please talk to WG chairs or ADs:

- BCP 9 (Internet Standards Process)
- BCP 25 (Working Group processes)
- BCP 25 (Anti-Harassment Procedures)
- BCP 54 (Code of Conduct)
- BCP 78 (Copyright)
- BCP 79 (Patents, Participation)
- https://www.ietf.org/privacy-policy/(Privacy Policy)

I E T F

This session is being recorded

IETF 110 Online Meeting Tips

- Make sure your video is off unless you are presenting during a session
- Mute your microphone unless you are speaking
- Use of a headset is strongly recommended
- Session bluesheet is automatically generated based on IETF Datatracker logins
- Chatrooms in Meetecho are connected to the Jabber chatrooms on IETF Datatracker agenda
- More information and assistance:
  - Participant guide https://www.ietf.org/how/meetings/110/session-participant-guide
  - Request assistance and report issues via: http://www.ietf.org/how/meetings/issues/

Administrivia

- Minute taker(s), jabber scribe(s)
- Meetecho Etiquette
  - Join the queue if you would like to speak/present
    - Do not send audio directly
  - Please state your name before speaking
  - Be mindful of the agenda time
    - Longer discussion on mailing list (or jabber)
- Collaborative minutes
  - https://codimd.ietf.org/notes-ietf-110-pce?both

Thank you Deborah & Welcome John Scudder!

Using the Mailing List

- Please use the mailing list actively to discuss all working group business
- Open issues with drafts should be discussed on the list, and conclusions reported to the list
- New drafts should be introduced to the working group first on the mailing list, to gauge interest
- Working group consensus is determined from the mailing list
- Priority in meetings is given to drafts that have been discussed on the list

Please be Vocal

- During WG Adoption and WG LC calls, response number is low.
- Please be vocal on the list to help us gauge the consensus better.
- The WG mailing lists are looked at by the IESG, IAB, and others (internal and external to IETF) to determine interest/participation level in our standards process.
- Please review ideas from your peers, these are community outputs of the working group as a whole.
- Also help flushing our queues faster
  - we had to extend the calls when response was lacking!

Using the Wiki

- A way to give you visibility as the document progress through the WG
  - adoption queue
  - WG LC queue
  - balancing work between chairs
  - shepherding responsibilities and opportunities
  - pending actions
  - IPR polls
- Use this wiki
  - make sure this is up to date!
- https://trac.ietf.org/trac/pce/wiki/WikiStart

Early Codepoint Allocation

- If you have an implementation of a WG I-D
  - that requires inter-operation with other implementations
    - Please request for early IANA codepoint allocation
  - Make sure to include an Implementation Status section in your I-D
  - Make sure the IANA section is correct and complete
    - And meets the condition set out in RFC 7120
- Maintained at
  - https://trac.ietf.org/trac/pce/wiki/WikiStart#CandidateforearlyIANAAllocations

Agenda Bashing

- **Introduction**
  1.1 Administrivia, Agenda Bashing (chairs, 5 min)
  1.2 WG Status (chairs, 10 min) [15/120]
  1.3 State of WG I-Ds and next steps (chairs, 15 min) [30/120]
- **Segment Routing**
  2.1 SR Policy (Mike Koldychev, 10 min) [40/120]
  draft-ietf-pce-segment-routing-policy-cp-03
  2.2 Algorithm in SID (Samuel, 10 min) [50/120]
  draft-tokar-pce-sid-algo-03
  2.3 SID Verification (Ran, 5 min) [55/120]
  draft-chen-pce-sr-mpls-sid-verification-01
  2.4 Multipath ERO (Mike Koldychev, 10 min) [65/120]
  draft-koldychev-pce-multipath-05
- **Stateful PCE & PCECC**
  3.1 Native IP (Boris/Aijun, 10 min) [75/120]
  draft-ietf-pce-pcep-extension-native-ip-11
  3.2 Inter-Domain (Olivier, 10 min) [85/120]
  draft-ietf-pce-stateful-interdomain-01
  3.3 IFIT (Giuseppe, 10 min) [95/120]
  draft-chen-pce-pcep-ifit-02
  3.4 Operational Clarification (Mike Koldychev, 5 min) [100/120]
  draft-koldychev-pce-operational-03
- **Ohers**
  4.1 BIER-TE (Huaimo, 10 min) [110/120]
  draft-chen-pce-bier-te-path-00
  4.2 PCEP-LS (Gyan, 10 min) [120/120]
  draft-dhodylee-pce-pcep-ls-20
- **If time permits**
  5.1 RSVP Color (Balaji Rajagopalan, 5 min) [125/120]
  draft-rajagopalan-pcep-rsvp-color-00

WG Status

Beyond the WG

- No new RFCs since IETF 109
- RFC Editor Queue
  - draft-ietf-pce-association-bidir
  - draft-ietf-pce-pcep-extension-for-pce-controller
  - draft-ietf-pce-association-policy (AUTH48)
  - draft-ietf-pce-pcep-flowspec (MISREF on IDR L2VPN flowspec)

In the WG's Hands

- 1 new Errata
  - RFC 5088 - Editorial (Reported)
- Early IANA codepoint allocation
  - draft-ietf-pce-local-protection-enforcement
    - Expires 2022-01-28
  - draft-ietf-pce-segment-routing-policy-cp
    - In process
  - draft-ietf-pce-binding-label-sid
    - Requested

Status of WG I-Ds & Next Steps

WG documents “nearing” WG LC

draft-ietf-pce-binding-label-sid

- -07 posted on 2021-02-19
- PCE Allocation of Binding SID moved from PCECC to this I-D
  - was discussed on the list
- Early Allocation Requested
- Ready for WG LC

draft-ietf-pce-enhanced-errors

- -09 posted on 2021-02-17
  - A refresh!
- Not used by other I-Ds yet
  - Is there still interest in this work?

WG documents “nearing” WG LC

draft-ietf-pce-pcep-yang

- -16 posted on 2021-02-22
- Some Typedef are changed to Identity
- Aligned to PCEP’s LSP-IDENTIFIERS TLV, instead of referring to ietf-te model
- Dependencies are making progress
- Very early YANG Doctor review was done
  - ready for another one
- WGLC next!

WG documents “nearing” WG LC

draft-ietf-pce-pcep-stateful-pce-gmpls

- -14 posted on 2020-12-29
- Reorganization done by authors
  - After the merge of 2 I-Ds
- Almost ready for WG-LC?

WG I-Ds

draft-ietf-pce-pcep-extension-native-ip

- -11 posted on 2021-02-07
- On Agenda
- Coordination with IDR on BGP considerations is needed

draft-ietf-pce-flexible-grid

- -05 posted on 2021-02-22
- A refresh!

WG I-Ds

draft-ietf-pce-segment-routing-ipv6

- -08 posted on 2020-11-24
- no major changes
- Comment from 109 - Align the SID Structure

draft-ietf-pce-vn-association

- No update

WG I-Ds

draft-ietf-pce-sr-path-segment

- -03 posted on 2021-02-08
- Support for SRv6
- P flag moved to BSID

draft-ietf-pce-sr-bidir-path

- -05 posted on 2021-01-26
- Sync with the post-WGLC changes of Bidir I-D

WG I-Ds

draft-ietf-pce-segment-routing-policy-cp

- -03 posted on 2021-02-22
- -04 posted on 2021-03-08
- On agenda
  - Issue with Association source resolved
  - Early code point allocation request in progress

draft-ietf-pce-local-protection-enforcement

- -02 posted on 2021-02-03
- Early code point allocated

Recent WG I-Ds

draft-ietf-pce-pcep-extension-pce-controller-sr

- -01 posted on 2021-02-21
- Comments received during adoption are handled

draft-ietf-pce-stateful-interdomain

- -01 posted on 2021-02-22
- On the agenda

draft-ietf-pce-lsp-extended-flags-00

- Recently adopted

WG Adoption Poll Queue

- draft-koldychev-pce-multipath
- draft-litkowski-pce-state-sync
- …
- Refer to wiki

Thanks!

IETF 110 – Online
PCE Working Group

PCEP SR Policy Association

M. Koldychev – Cisco Systems (mkoldych@cisco.com) – Presenter
M. Sivabalan – Ciena Corporation (ssivabal@ciena.com)
C. Barth – Juniper Networks (cbarth@juniper.net)
S. Peng – Huawei Technologies (pengshuping@Huawei.com)
H. Bidgoli – Nokia (Hooman.Bidgoli@Nokia.com)

Recent updates

Version-02

- Encode Color + Endpoint in Extended Association ID TLV
- SR Policy Identifiers TLV is not necessary because SR Policy ID is encoded into the Association ID

Version-03 and Version-04

- Preparation for IANA allocation
- Cleaned up the wording and re-organize to be more clear

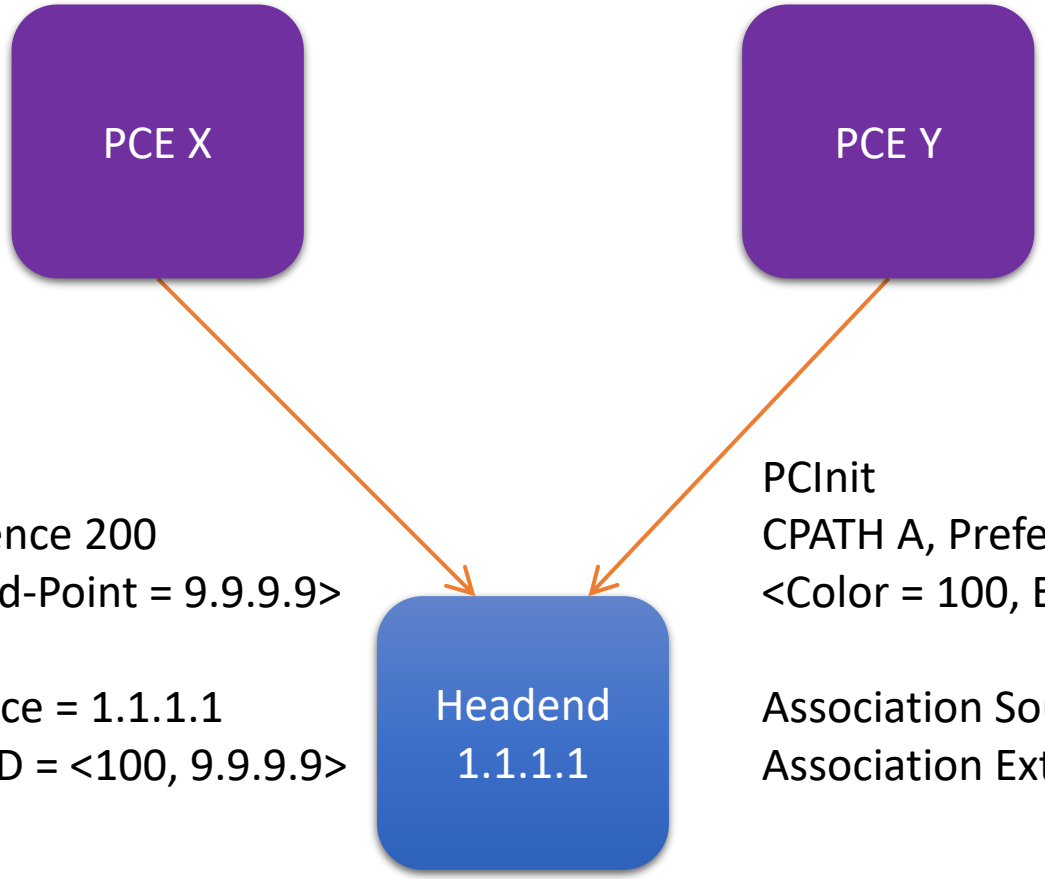
Extended Association ID TLV

- We encode the <headend, color, endpoint> Policy ID tuple directly into Association Parameters
- Association Key is the same as the Policy ID
- Association Source encodes the <headend> part of Policy ID
- Extended Association ID TLV encodes the <color, endpoint> part

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Type = 31           |       Length = 8 or 20        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             Color                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                           Endpoint                            ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: Extended Association ID TLV format

Example: two CPATHs of same SR Policy



PCInit
CPATH B, Preference 200
<Color = 100, End-Point = 9.9.9.9>

Association Source = 1.1.1.1
Association Ext ID = <100, 9.9.9.9>

PCInit
CPATH A, Preference 100
<Color = 100, End-Point = 9.9.9.9>

Association Source = 1.1.1.1
Association Ext ID = <100, 9.9.9.9>

Next steps

- Requested IANA code points
- Interop testing once IANA code points are allocated

IETF 110 – Online
PCE Working Group

Carrying SID Algorithm information in PCE-based Networks

A. Tokar – Cisco Systems (atokar@cisco.com)
S. Sidor – Cisco Systems (ssidor@cisco.com) – Presenter
M. Sivabalan – Ciena Corporation (ssivabal@ciena.com)
S. Peng – Huawei Technologies (pengshuping@huawei.com)
M. Negi – RtBrick Inc (mahend.ietf@gmail.com)

Motivation

- A PCE can compute SR-TE paths using prefix SIDs with different Algorithms depending on the use-case, constraints, etc. While this information is available on the PCE, there is no method of conveying this information to the headend router

- The headend can also compute SR-TE paths using different Algorithms, and this information also needs to be conveyed to the PCE for collection or troubleshooting purposes

- An operator may also want to constrain the path computed by the PCE to a specific SID Algorithm. For example, in order to only use SID Algorithms for a low-latency path

Prefix-SID Algorithm

- Described in Segment Routing Architecture RFC
  - RFC 8402 - Section 3.1.1
  - Two algorithms defined
    - Shortest Path First (SPF) = 0
    - Strict Shortest Path First (Strict-SPF) = 1
- IGP Algorithm Types in IANA registry
  - Two existing algorithms and Flexible Algorithms range

TLV for the LSPA Object - changes

- Removed F flag

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Type=TBD             |            Length=4           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Reserved             |   Flags   |L|F|   Algorithm   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- ~~F (Fallback): If set to 1 and the PCE is unable to compute a path using only prefix SIDs with the specified Algorithm, the PCE MAY compute an alternate fallback path without constraining to the specified Algorithm.~~
- L (Loose): If set to 1, the PCE MAY insert prefix SIDs with a different Algorithm, but it MUST prefer the specified Algorithm whenever possible.

TLV for the LSPA Object - clarification

- SID Algorithm is not replacing Objective Function
  - Specified algorithm acts as filter restricting, which SIDs can be used
- Algorithm constraint in LSPA is specifying end-to-end algorithm
  - Applicable only if L flag is not set
- Different algorithms can be used in one path if L flag is set
  - Loop avoidance is implementation specific

Next steps

- Comments and discussion are welcome

PCEP Extensions for SID Verification for SR-MPLS

draft-chen-pce-sr-mpls-sid-verification-01

Ran Chen（ZTE）----Presenter
Samuel SidorChun（Cisco systems）
Chun Zhu（ZTE）
Alex. Tokar（Cisco systems）
Mike. Koldychev（Cisco systems）

Virtual PCE WG IETF-110 Meeting, March 2021

Motivation

- draft-ietf-spring-segment-routing-policy describes the "SID verification" bit usage. SID verification is performed when the headend is explicitly requested to verify SID(s) by the controller via the signaling protocol used. Implementations MAY provide a local configuration option to enable verification on a global or per policy or per candidate path basis.

- RFC8664 specifies extensions to the PCEP that allow a stateful PCE to compute and initiate TE paths, as well as a Path Computation Client (PCC) to request a path subject to certain constraints and optimization criteria in SR networks.

- This document updates RFC8664 to clarify usage of "SID verification" bit signaled in Path Computation Element Protocol (PCEP), and this document proposes to define a new flag for indicating the headend is explicitly requested to verify SID(s) by the PCE.

SID verification flag(V-Flag)

- Extended V-Flag in SR-ERO Subobject for indicating the PCC is explicitly requested to verify SID(s) by the PCE.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|L|   Type=36   |     Length    |  NT   |     Flags   |V|F|S|C|M|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        SID (optional)                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                  NAI (variable, optional)                   //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- Extended V-Flag in SR-RRO Subobject
  - The format of the SR-RRO subobject is the same as that of the SR-ERO subobject, but without the L-Flag, per RFC8664.
  - The V flag has no meaning in the SR-RRO and is ignored on receipt at the PCE.

Next Step

- Request review and feedback from WG

Thanks!

IETF 110 – Online
PCE Working Group

PCEP Extensions for Signaling Multipath Information

M. Koldychev – Cisco Systems (mkoldych@cisco.com) – Presenter
M. Sivabalan – Ciena Corporation (ssivabal@ciena.com)
T. Saad – Juniper Networks (tsaad@juniper.net)
V. Beeram – Juniper Networks (vbeeram@juniper.net)
H. Bidgoli – Nokia (Hooman.Bidgoli@Nokia.com)
S. Peng – Huawei Technologies (pengshuping@Huawei.com)
B. Yadav – Ciena (byadav@ciena.com)

Recent updates

Version-03 and Version-04

- Editorial changes

Version-05

- Composite candidate path signaling

Composite Candidate Path

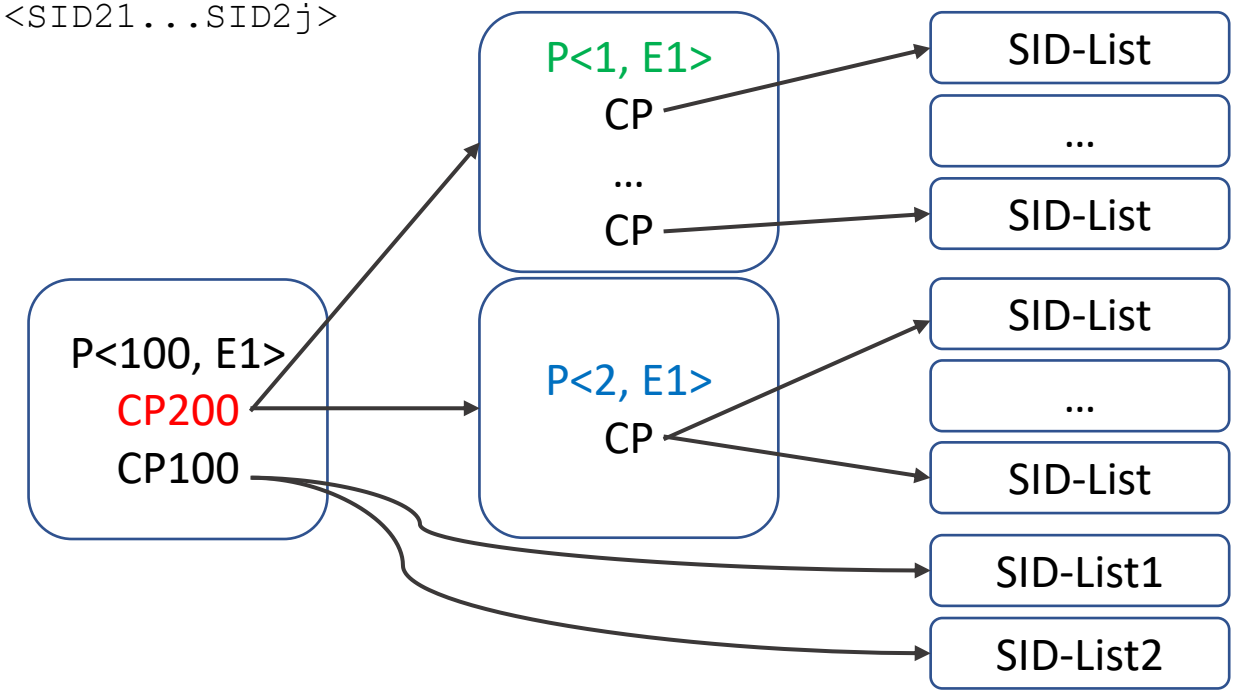
Composite Candidate Path, introduced in [draft-ietf-spring-segment-routing-policy-09], allows an SR Policy Candidate Path to ECMP/UCMP traffic to other SR Policies on the same headend.

Example of Composite Candidate Path and Explicit Candidate Path:

```
SR policy POL100 <headend = H1, color = 100, endpoint = E1>
    Candidate-path CP200 <protocol-origin = 20, originator = 100:1.1.1.1, discriminator = 1>
        Preference 200
        Weight W1, SR policy <color = 1>
        Weight W2, SR policy <color = 2>
    Candidate-path CP100 <protocol-origin = 20, originator = 100:1.1.1.1, discriminator = 2>
         Preference 100
         Weight W1, SID-List1 <SID11...SID1i>
         Weight W2, SID-List2 <SID21...SID2j>
```

Example

```
SR policy POL100 <headend = H1, color = 100, endpoint = E1>
    Candidate-path CP200 <protocol-origin = 20, originator = 100:1.1.1.1, discriminator = 1>
        Preference 200
        Weight W1, SR policy <color = 1>
        Weight W2, SR policy <color = 2>
    Candidate-path CP100 <protocol-origin = 20, originator = 100:1.1.1.1, discriminator = 2>
         Preference 100
         Weight W1, SID-List1 <SID11...SID1i>
         Weight W2, SID-List2 <SID21...SID2j>
```



PCEP Extensions

Composite Candidate Path can be signaled by putting COLOR TLV [I-D.peng-pce-te-constraints] into the PATH-ATTRIB object and sending an empty SR-ERO object.

```
<LSP>
<ASSOCIATION Type=SRPOLICY>
<PATH-ATTRIB <WEIGHT-TLV Weight=W1> <COLOR-TLV Color=1>>
<SR-ERO (empty)>
<PATH-ATTRIB <WEIGHT-TLV Weight=W2> <COLOR-TLV Color=2>>
<SR-ERO (empty)>
```

Next steps

- Wait for WG adoption
- Discuss

PCEP Extension for Native IP Network

draft-ietf-pce-pcep-extension-native-ip

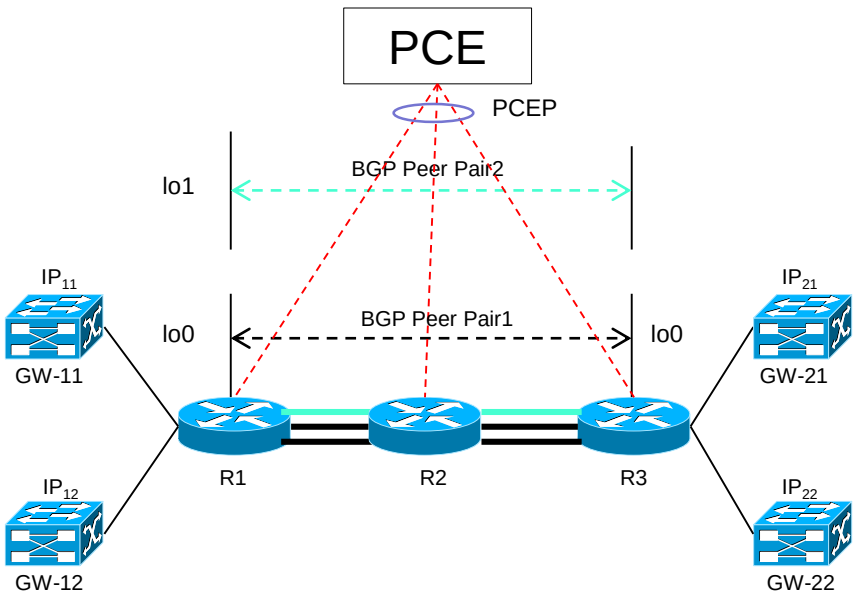
A. Wang (China Telecom)
B. Khasanov (Yandex)
Sheng Fang (Huawei Technologies)
Ren Tan(Huawei Technologies)
Chun Zhun(ZTE Corporation)

IETF-110, March 2021

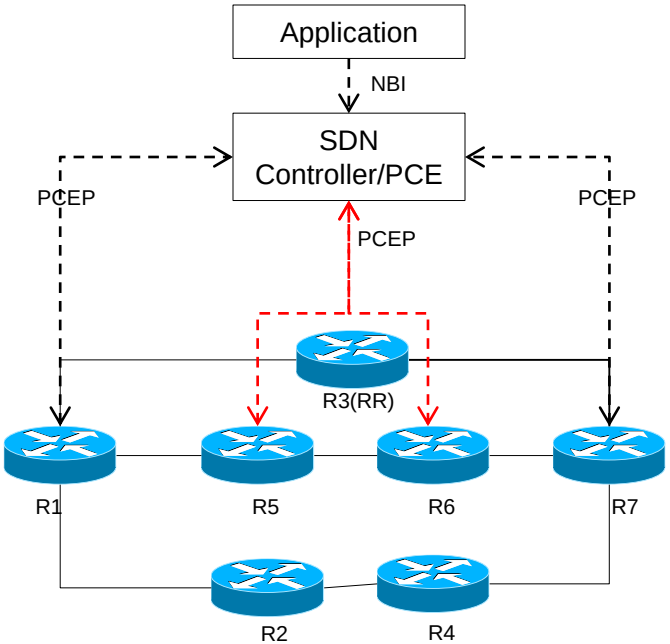
Motivation

- RFC8735 describes the scenarios and simulation results for TE in Native IP network
- Draft-ietf-teas-pce-native-ip describes an architecture for providing traffic engineering in a native IP network by using multiple BGP sessions and a PCE-based central control mechanism.
- This document describes the PCEP extensions and procedures to practically build such architecture.

Overview of the Solution



Dual/Multi-BGP Solution



Simplified CCDR* Architecture in a Large Network

- Building Dual/Multi BGP sessions between edge routers upon request via PCEP
- Advertises different prefixes via different BGP sessions, w/PCEP-based setup
- Steer traffic towards particular routes via BGP next-hop w/PCEP-based setup

CCDR* - Centralized Control Dynamic Routing, RFC 8735

PCEP Extensions

From Section 4-10 of the document:

- Capability Advertisement
- Related PCEP messages
- New PCEP Objects
- New Error-Type and Error-Value
- CCDR Native IP Procedures
- Operations Consideration

PCEP Capability Advertisement

- RFC8408 defines the Path Setup Type Capability TLV to indicate the path type supported by the PCE and PCC

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Type (34)            |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Reserved              |  Num of PSTs  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     PST#1     |      ...      |     PST#N     |    Padding    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//              Optional sub-TLVs (variable)                   //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: PATH-SETUP-TYPE-CAPABILITY TLV

  - New PST (TBD) is defined for Native IP path

- Draft-ietf-pce-pcep-extension-for-pce-controller defines the PCECC capability sub-tlv to exchange information about the PCECC capability
  - N(NATIVE-IP-TE-CAPABILITY-1 bit=TBD) is defined for PCEP speaker

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Type=TBD12       |          Length=4             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             Flags                           N|L|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 7: PCECC Capability sub-TLV

Updated PCEP Messages

```
<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>)|
                                           ((<BPI>|<EPR>|<PPA>)
                                           <CCI>)

   <cci-list> ::=  <CCI>
                   [<cci-list>]

Where:
    <cci-list> is as per
    [I-D.ietf-pce-pcep-extension-for-pce-controller].
    <PCE-initiated-lsp-instantiation> and
    <PCE-initiated-lsp-deletion> are as per
    [RFC8281].

    The LSP and SRP object is defined in [RFC8231].
```

When PCInitiate message is used create Native IP instructions, the SRP and CCI objects MUST be present. The error handling for missing SRP or CCI object is as per [I-D.ietf-pce-pcep-extension-for-pce-controller]. Further either one of BPI, EPR, or PPA object MUST be present. If none of them are present, the receiving PCC MUST send a PCErr message with Error-type=6 (Mandatory Object missing) and Error-value=TBD (Native IP object missing).

To cleanup the SRP object must set the R (remove) bit.

```
<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>)|
                                ((<BPI>|<EPR>|<PPA>)
                                <CCI>)

Where:
   <path> is as per [RFC8231] and the LSP and SRP object are
   also defined in [RFC8231].
```

The error handling for missing CCI object is as per [I-D.ietf-pce-pcep-extension-for-pce-controller]. Further either one of BPI, EPR, or PPA object 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=TBD (Native IP object missing).

SRP – Stateful Request parameters Object RFC8231, CCI – Central Controller Instructions Object, BPI - BGP Peer Info Object, ERP – Explicit Peer Route Object, PPA - Peer Prefix Association Object.

New PCEP Objects(1/4)

CCI: Central Controller Instructions

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            CC-ID                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Reserved             |             Flags             |
+---------------------------------------------------------------+
|                                                               |
//                        Optional TLV                         //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5: CCI Object for Native IP

CC-ID is described in draft-ietf-pce-pcep-extension-for-pce-controller

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

Symbolic Path Name TLV(RFC 8231) MUST be included in the above CCI-Object

New PCEP Objects(2/4)

BPI (BGP Peer Info) Object-Class is TBD

BPI Object-Type is for IPv4 and 2 for IPv6

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Peer AS Number                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   ETTL        |               Reserved                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Local IP Address                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Peer IP Address                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Additional TLVs                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 6: BGP Peer Info Object Body Format for IPv4

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Peer AS Number                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   ETTL        |               Reserved                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
+                                                               +
|                Local IP Address (16 bytes)                    |
+                                                               +
|                                                               |
+                                                               +
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
+                                                               +
|                Peer IP Address (16 bytes)                     |
+                                                               +
|                                                               |
+                                                               +
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Additional TLVs                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 7: BGP Peer Info Object Body Format for IPv6

New PCEP Objects(3/4)

EPR (Explicit Peer Route) Object-Class is TBD

EPR Object-Type is 1 for IPv4 and 2 for IPv6

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       Route Priority        |           Reserved              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                IPv4 Peer Address                                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Next Hop Address to the IPv4 Peer Address            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Additional TLVs                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 8: Explicit Peer Route Object Body Format for IPv4

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       Route Priority        |           Reserved              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
+                                                               +
|                      IPv6 Peer Address                        |
+                                                               +
|                                                               |
+                                                               +
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
+                                                               +
|          Next Hop Address to the IPv6 Peer Address            |
+                                                               +
|                                                               |
+                                                               +
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Additional TLVs                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 9: Explicit Peer Route Object Body Format for IPv6

New PCEP Objects(4/4)

PPA (Peer Prefix Association) Object-Class is TBD

PPA Object-Type is 1 for IPv4 and 2 for IPv6

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Peer IPv4 Address                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//               IPv4 Prefix subobjects                        //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Additional TLVs                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 10: Peer Prefix Association Object Body Format for IPv4

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Peer IPv6 Address                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//               IPv6 Prefix subobjects                        //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Additional TLVs                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 11: Peer Prefix Association Object Body Format for IPv6

IPv4 Prefix sub-object/IPv6 Prefix sub-object is defined in RFC3209

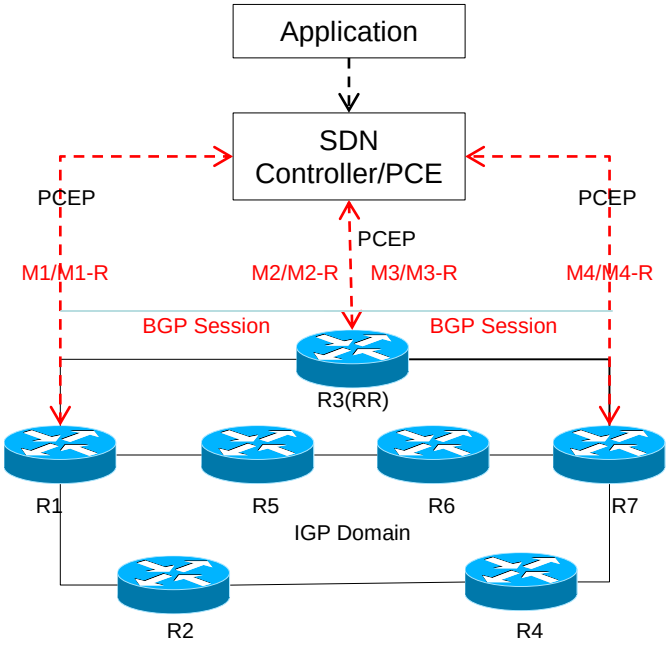
New Error-Type and Error-Values

Error-Type	Meaning	Error-value
TBD6	Native IP TE failure	
		0: Unassigned
		TBD7: Peer AS not match
		TBD8:Peer IP can't be reached
		TBD9:Local IP is in use
		TBD10:Remote IP is in use
		TBD11:Exist BGP session broken
		TBD12:Explicit Peer Route Error
		TBD17:EPR/BPI Peer Info mismatch
		TBD18:BPI/PPA Address Family mismatch
		TBD19:PPA/BPI Peer Info mismatch

Figure 12: Newly defined Error-Type and Error-Value

CCDR Native IP Procedures(1/3)

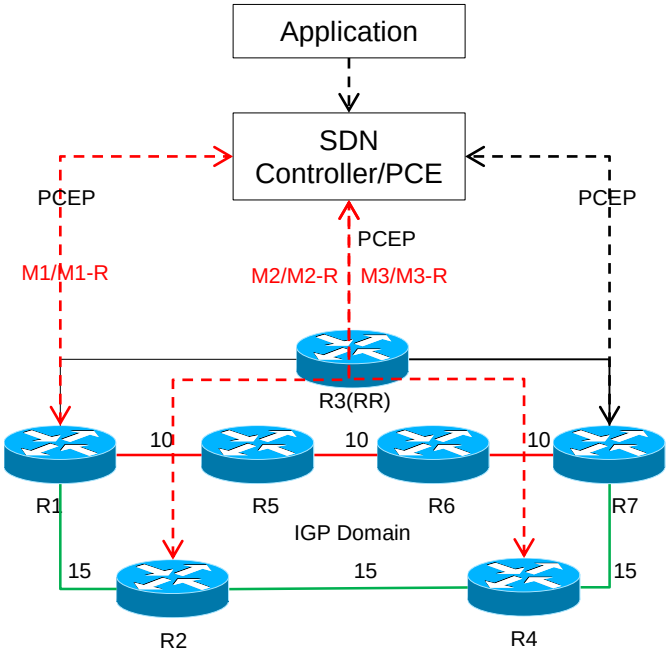
- PCEP messages (PCInitiate/PCRpt) are sent to R1、R3(RR)、R7 respectively, to build/remove the BGP sessions between R1/R3(RR) and R3(RR)/R7
- BPI Object is included
- “CC-ID” is unique, but “Symbolic Path Name” is constant for the E2E path
- “Local/Peer IP address” and “Peer AS” number is assigned within the BPI Object



BGP Session Establishment Procedures

CCDR Native IP Procedures(2/3)

- PCEP messages (PCInitiate/PCRpt) are sent to on-path routers R1、R2、R4 respectively, to install/remove the explicit route to the BGP nexthop
- EPR Object is included
- “CC-ID” is unique, but “Symbolic Path Name” is constant for the E2E path
- “Peer Address” and “Next Hop” information is assigned within the BPI Object. The route priority etc. for such path could also be assigned
- Reverse Path is built similarly from R7/R4/R2

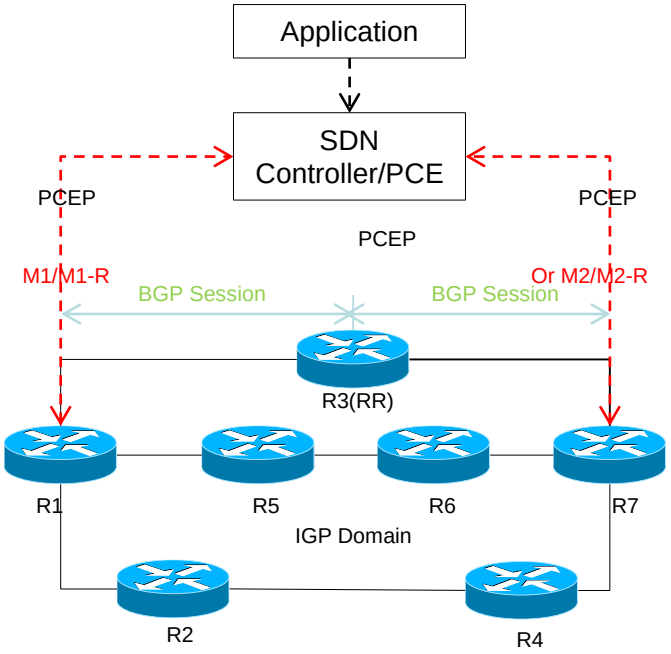


Explicit Route Establish Procedures
(From R1 to R7))

Red Link represent congested path (IGP shortest path)
Green Link represent idle path (explicit route from PCE)

CCDR Native IP Procedures(3/3)

- PCEP messages (PCInitiate/PCRpt) are sent only to edge routers R1 or R7, to advertise/revoke the prefixes that associated with different BGP sessions
- PPA Object is included
- “CC-ID” is unique, but “Symbolic Path Name” is constant for the E2E path
- “Peer Address” and “Advertised Prefixes” information is assigned within the PPA Object
- Same AF prefixes should be advertised via the same AF sessions



BGP Prefix Advertisement Procedures

Operations Consideration

- The information transferred in this draft is mainly used for the light weight BGP session setup, explicit route deployment and the prefix distribution
- The planning, allocation and distribution of the peer addresses within IGP domain should be done in advance
- The state synchronization between PCE and PCC should follow the procedure that described in RFC8232
- When PCE detects one or some of PCCs are out of control, it should recompute and redeploy a traffic engineering path for native IP on active PCCs
- When PCC detects that connection with a PCE is lost, it should stale the information that initiated by PCE. A PCE should assure the avoidance of possible transient loop in such node failure case when it deploys the explicit peer route on the PCCs.

Next Step

1. Comments/Q&A
2. WG Last Call?

Aijun Wang@ChinaTelecom
Khasanov.Boris@Yandex
Sheng Fang@Huawei
Ren Tan@Huawei
C.Zhu@ZTE

IETF110@Online

PCEP Extension for Stateful Inter-Domain Tunnels

Olivier Dugeon & Julien Meuric (Orange Labs)

Y. Lee (Samsung)

D. Ceccarelli (Ericsson)

draft-ietf-pce-stateful-interdomain-01

Update since previous versions

- Version 00:
  - Working Group adoption based on draft-dugeon-pce-stateful-interdomain-04.txt

- Version 01:
  - Take into account comments received during the WG call adoption
    - Except the comment about the implementation option

Implementation requirements

- The stitching label principle requires at least a certain number of modifications in the current PCEP version
  - A new PCE Capability to announce the inter-domain behavior
  - A new PCE Association Group to associate the local paths identifier to the inter-domain identifier
  - A new PCEP Errors message to manage the Stitching Label exchange
  - A mechanism to convey the Stitching Label
    - The WG should choose between several options

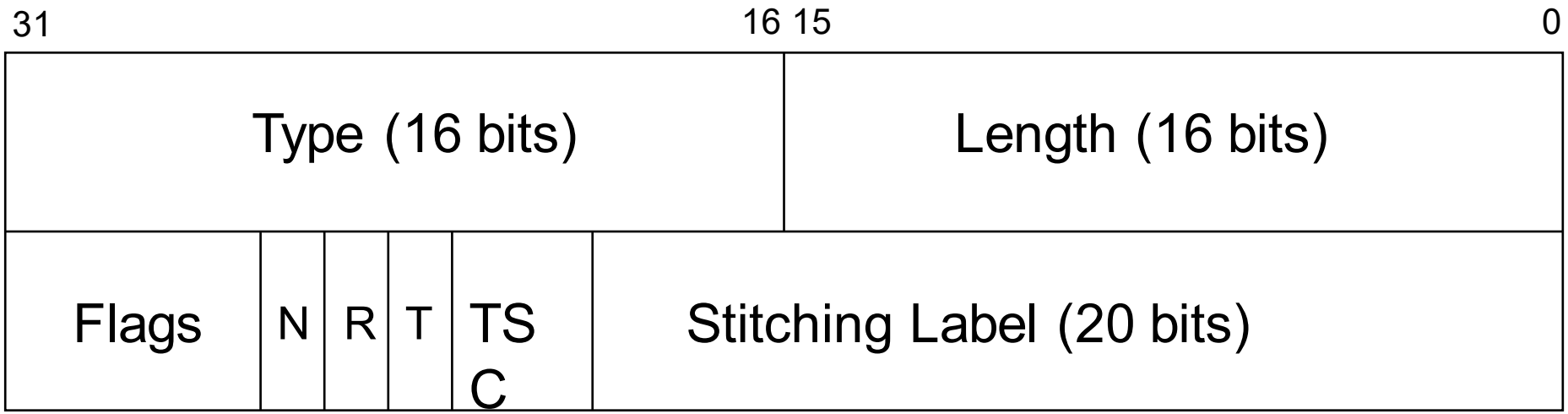
Technical solution to convey Stitching Label

- Use ERO and RRO in conjunction to new Path Setup code points
  - Solution proposed in the current version of the draft
  - Pro: Simplest implementation
  - Cons: As mention by Dhruv, each time a new path enforcement appear, a new PST code point must be allocated (e.g. new Segment Routing v6)

- Use ERO and RRO in conjunction to a new flag in LSP
  - Pro: Simple as PST code points
  - Cons: Need to use the new LSP Extended Flag sub-Object
  - Alternate solution: find another place for the flag e.g. SRP or LSPA Object

- Define a new PCEP sub-Objet TLV within the LSP to convey the stitching label
  - Pro: More independent and explicitly convey the Stitching Label
  - Cons: Need extra parsing in the PCEP Grammar from an implementation point of view

New Stitching Label sub-Object format



- Stitching Label (20 bits): Must equal to 0 for request
- TSC (3 bits): Traffic Class for the Stitching Label if T flag is set
- Flags (8 bits):
  - R: Request Stitching Label
  - T: Traffic Class must be used
  - N: Nested path

Path Computation Element Communication Protocol (PCEP) Extensions to Enable IFIT

draft-chen-pce-pcep-ifit-02

Online, Mar 2021, IETF 110

Huanan Chen (China Telecom)
Hang Yuan (UnionPay)
Tianran Zhou (Huawei)
Weidong Li (Huawei)
Giuseppe Fioccola (Huawei)
Yali Wang (Huawei)

Background and Motivation

- In-situ Flow Information Telemetry (**IFIT**) refers to dataplane on-path telemetry techniques, including In-situ OAM (**IOAM**) (draft-ietf-ippm-ioam-data) and **Alternate Marking** (RFC8321, RFC8889)

- The **PCEP extension** defined in this document allows to signal the IFIT capabilities. In this way IFIT methods are automatically activated and running.

The IFIT attributes can be generalized and included as TLVs carried inside the LSPA (**LSP Attributes**) object in order to be applied for all path types, as long as they support the relevant data plane telemetry method

Changes from -01 to -02

Comments during IETF 109

- Mike Koldychev: consider the case of different IFIT methods for each Segment Lists (draft-koldychev-pce-multipath).
  - If it is needed to apply different IFIT methods for each Segment List, the IFIT attributes can be added into the PATH-ATTRIB object, instead of the LSPA object, according to draft-koldychev-pce-multipath that defines PCEP Extensions for Signaling Multipath Information.

- It has been added the AltMark Mode (HbH, DOH).

- The companion document (draft-ietf-idr-sr-policy-ifit) has been adopted in IDR WG

IFIT capability advertisement TLV

A new **IFIT-CAPABILITY TLV**, that is an optional TLV for use in the OPEN Object for IFIT attributes via PCEP capability advertisement

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Type                |             Length=4          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Flags                       P|I|D|E|M|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

P: IOAM Pre-allocated Trace Option Type-enabled flag (draft-ietf-ippm-ioam-data)

I: IOAM Incremental Trace Option Type-enabled flag (draft-ietf-ippm-ioam-data)

D: IOAM DEX Option Type-enabled flag (draft-ietf-ippm-ioam-data)

E: IOAM E2E Option Type-enabled flag (draft-ietf-ippm-ioam-data)

M: Alternate Marking enabled flag (RFC8321)

- If set to 1 by a PCC, the flag indicates that the PCC allows instantiation of the feature by a PCE
- If set to 1 by a PCE, the flag indicates that the PCE supports the feature instantiation
- The flag MUST be set by both PCC and PCE in order to support the instantiation

IFIT Attributes TLV

The **IFIT-ATTRIBUTES TLV** provides the configurable knobs of the IFIT feature, and it can be included as an optional TLV in the **LSPA object**

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Type              |           Length              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//                          sub-TLVs                           //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

IFIT attribute TLVs, carried inside the LSPA object and applicable to all path types

- IFIT TLVs are optional and can be taken into account by the PCE during path computation and by the PCC during path setup.
- In general, the LSPA object can be carried within a PCInitiate message, a PCUpd message, or a PCRpt message in the stateful PCE model.
- IFIT for SR Policies: IFIT attibutes also complement draft-ietf-pce-segment-routing-policy-cp

IOAM Sub-TLVs

- IOAM Pre-allocated Trace Option Sub-TLV

```
+-------------------------------+-------------------------------+
|          Type=1               |           Length=8            |
+---------------------------------------------------------------+
|       Namespace ID            |            Rsvd1              |
+-------------------------------+-----------------------+-------+
|         IOAM Trace Type                               | Flags | Rsvd2 |
+-----------------------------------------------------+-------+-------+
```

- IOAM Incremental Trace Option Sub-TLV

```
+-------------------------------+-------------------------------+
|          Type=2               |           Length=8            |
+---------------------------------------------------------------+
|       Namespace ID            |            Rsvd1              |
+-------------------------------+-----------------------+-------+
|         IOAM Trace Type                               | Flags | Rsvd2 |
+-----------------------------------------------------+-------+-------+
```

- IOAM Directly Export Option Sub-TLV

```
+-------------------------------+-------------------------------+
|           Type=3              |          Length=12            |
+---------------------------------------------------------------+
|        Namespace ID           |            Flags              |
+-------------------------------+---------------+---------------+
|               IOAM Trace Type                 |      Rsvd     |
+-----------------------------------------------+---------------+
|                        Flow ID                                |
+---------------------------------------------------------------+
```

- IOAM Edge-to-Edge Option Sub-TLV

```
+-------------------------------+-------------------------------+
|         Type=4                |           Length=4            |
+---------------------------------------------------------------+
|        Namespace ID           |          IOAM E2E Type        |
+-------------------------------+-------------------------------+
```

Enhanced Alternate Marking Sub-TLV

- Enhanced Alternate Marking Sub-TLV

```
+-------------------------------+-------------------------------+
|             Type=5            |            Length=4           |
+-------------------------------+-------+---------------+-------+
|            FlowMonID                  |     Period    |H|E| R |
+---------------------------------------+---------------+-------+
```

New fields added:

H: A flag indicating that the measurement is Hop-By-Hop.

E: A flag indicating that the measurement is end to end.

PCEP Messages and Example of application to SR Policy

The Examples of PCC Initiated SR Policy and PCE Initiated SR Policy are reported in draft-ietf-pce-segment-routing-policy-cp and this draft describes the addition of IFIT TLVs through LSPA object:

- For the PCE-initiated LSP with the IFIT feature enabled, IFIT-ATTRIBUTES TLV MUST be included in the LSPA object with the **PCInitiate message**
- The PCC creates the LSP using the attributes communicated by the PCE and the local values for the unspecified parameters
- After the successful instantiation of the LSP, the PCC automatically delegates the LSP to the PCE and generates a **PCRpt message** to provide the status report for the LSP
- When the LSP is instantiated the IFIT methods are applied as specified for the corresponding data plane, e.g. draft-ietf-ippm-ioam-ipv6-options and draft-ietf-6man-ipv6-alt-mark
- To enabling/disabling some features, the IFIT-ATTRIBUTES TLV MUST be included in the LSPA object with the **PCUpd message**

Discussion & Next Steps

- Evaluate WG adoption
- Welcome questions, comments

Thank you

IETF 110 – Online
PCE Working Group

PCEP Operational Clarification

M. Koldychev – Cisco Systems (mkoldych@cisco.com) – Presenter
M. Sivabalan – Ciena Corporation (ssivabal@ciena.com)
S. Peng – Huawei Technologies (pengshuping@Huawei.com)
Diego Achaval – Nokia (diego.achaval@nokia.com)
Hari Kotni – Juniper Networks (hkotni@juniper.net)

Recent updates

Version-03

- RRO object usage in Segment Routing

RRO in Segment Routing

- The original PCEP RRO object is taken directly from RSVP-TE
- In RSVP-TE the RRO is OPTIONAL, meaning that an RSVP-TE LSP can be setup using ONLY the ERO object, without RRO
- In RSVP-TE the RRO is used for functions like label recording along the route which is used for FRR Node Protection, for example
- It is not clear how/if this applies to SR-TE
- SR-RRO and SRv6-RRO objects were “cloned” from the RRO object, but there is no definition of their purpose in SR-TE
- Propose to use SR-ERO/SRv6-ERO and to NOT use SR-RRO/SRv6-RRO
- This seems to match other vendors implementations

Next steps

- We would like to request WG Adoption of this draft

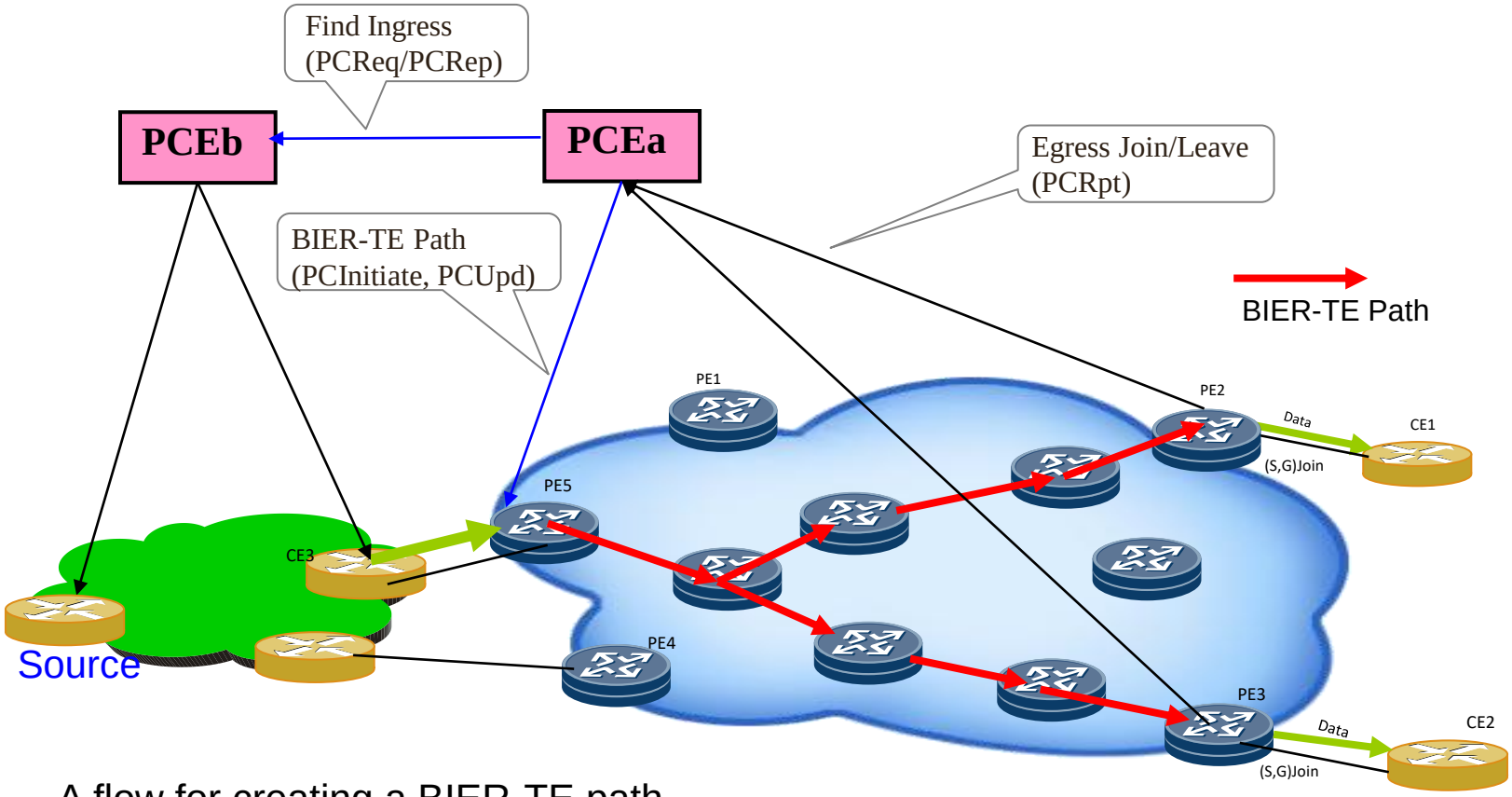
PCE for BIER-TE Path

draft-chen-pce-bier-te-path-00

Huaimo Chen, Mike McBride(Futurewei)
Aijun Wang (China Telecom)
Gyan S. Mishra (Verizon Inc.)
Yisong Liu (China Mobile)
Yanhe Fan (Casa Systems)
Lei Liu (Fujitsu)
Xufeng Liu (Volta Networks)

IETF 110

PCE for BIER-TE Path Overview



A flow for creating a BIER-TE path

1. PCEa gets reports about egresses in PCRpt
2. PCEa requests PCEb to find Ingress by PCReq
3. PCEb finds ingress, and sends it to PCEa in PCRep
4. PCEa computes path, sends it to Ingress using PCInitiate
5. Ingress creates path, sends PCEa report in PCRpt
6. PCEa gets report about path creation, records it

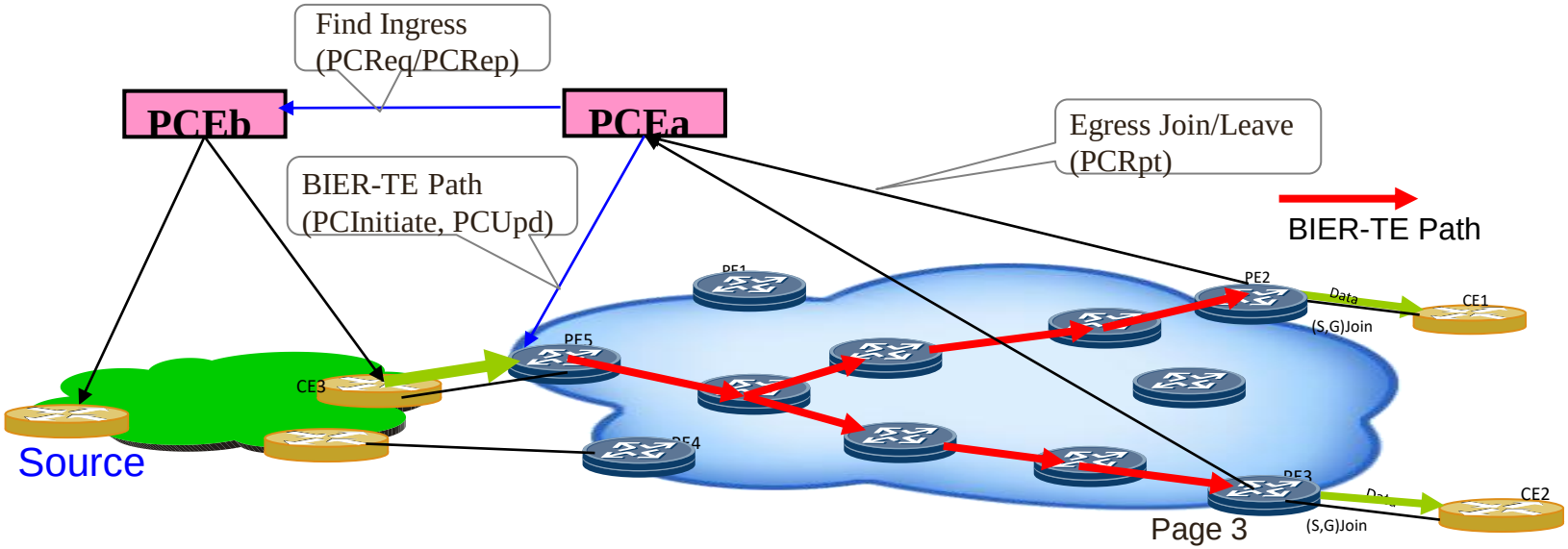
More on BIER-TE Path

A flow for updating a BIER-TE path

1. PCEa gets reports about existing egress leave in PCRpt
2. PCEa computes a new path, sends it to Ingress using PCUpd
3. Ingress replaces old path with new one, sends PCEa report in PCRpt
4. PCEa gets report about path update in PCRpt, records it

Another flow for creating a BIER-TE path

1. PCEa gets a request from a user for a path from ingress to egresses
2. PCEa computes a path, sends it to Ingress using PCInitiate
3. Ingress creates path, sends PCEa report in PCRpt
4. PCEa gets report about path creation in PCRpt, records it



Extensions to PCE

Extensions to PCEP Objects

1. BIER-TE Path Capability
2. Extensions to SRP
3. Ingress Node Object
4. Objective Functions
5. BIER-TE Path Subobject

Extensions to PCEP Messages

1. PCRpt Message
2. PCUpd Message
3. PCInitiate Message
4. PCReq Message
5. PCRep Message

Procedures

1. BIER-TE Path Creation
2. BIER-TE Path Update
3. BIER-TE Path Deletion

Capability for BIER-TE Path

PCC and PCE exchange capabilities of BIER-TE path

PATH_SETUP_TYPE_CAPABILITY TLV with PST = TBD1 + sub-TLV in Open

* PST = TBD1: Path is setup using BIER-TE.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Type = TBD2          |            Length=4           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Reserved                    |  SILen  |   BitStrLen   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

PCE-BIER-TE-Path-Capability sub-TLV

- SILen (SI Length) - 5 bits: The length in bits of the SI field

- BitStrLen (Bit String Length) - 8 bits: The length in bits of the BitString field according to RFC8296. If k is the length of the BitString, the value of BitStrLen is log2(k)-5. For example, BitStrLen = 1 indicates k = 64, BitStrLen = 7 indicates k = 4096.

Extensions to SRP (1/4)

For BIER-TE path, SRP MUST include:

- PATH-SETUP-TYPE TLV with PST = TBD1 for path setup using BIER-TE

Three contiguous bits (bit 27-29) in SRP Flag Field are defined as AOP (Assistant Operation).

Multicast Traffic TLV

```
AOP Value     Meaning (Assistant Operation)
0x001 (J):  Join with Multicast Group and Source
0x010 (L):  Leave from Multicast Group and Source
0x011 (I):  Ingress node computation
```

- When the PCC running on an edge node of a BIER-TE domain sends the PCE for the domain a PCEP message such as PCRpt to report that the edge node receives a multicast join, the message MUST include a SRP object with AOP == 0x001 (J).

- When the PCC running on an edge node of a BIER-TE domain sends the PCE for the domain a PCEP message such as PCRpt to report that the edge node receives a multicast leave, the message MUST include a SRP object with AOP == 0x010 (L).

- When the PCE for the domain sends a PCEP message such as PCReq to another PCE for requesting to find an ingress node for a BIER-TE path, the message MUST include a SRP object with AOP == 0x011 (I).

Extensions to SRP (2/4)

Multicast Traffic TLV

- For a PCE-Initiated BIER-TE path, when a PCE sends a PCC a PCInitiate message to create a BIER-TE path in a BIER-TE domain, the message MUST contain the Multicast Traffic Information TLV in SRP.
- When PCC on an edge node of a BIER-TE domain sends PCE for the domain a message to report that the edge node receives a multicast join or leave with a multicast group/address and source, the message MUST contain the Multicast Traffic Information TLV in SRP.
- When the PCE for a BIER-TE domain sends another PCE a message to request for finding an ingress node of a BIER-TE path, the message MUST contain the Multicast Traffic Information TLV in SRP.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Type = TBD3         |            Length=4           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                       sub-TLVs (optional)                     ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                 Multicast Traffic Information TLV
```

```
Two groups of sub-TLVs for IPv4/IPv6:
 o  IPv4/IPv6 multicast group address prefix sub-TLV and
 o  IPv4/IPv6 multicast source address prefix sub-TLV.
```

Extensions to SRP (3/4)

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Type = TBD4          |        Length (Variable)      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Prefix-Len   |     IPv4 Multicast Group Address Prefix       ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

IPv4 Multicast Group Address Prefix sub-TLV

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Type = TBD5          |        Length (Variable)      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Prefix-Len   |     IPv6 Multicast Group Address Prefix       |
+-+-+-+-+-+-+-+-+                                               |
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

IPv6 Multicast Group Address Prefix sub-TLV

Extensions to SRP (4/4)

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Type = TBD6          |        Length (Variable)      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Prefix-Len   |     IPv4 Multicast Source Address Prefix      ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

IPv4 Multicast Source Address Prefix sub-TLV

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Type = TBD7          |        Length (Variable)      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Prefix-Len   |     IPv6 Multicast Source Address Prefix      |
+-+-+-+-+-+-+-+-+                                               |
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

IPv6 Multicast Source Address Prefix sub-TLV

Ingress Node Object

Ingress node object to represent ingress node

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|ObjectClass=TBD|  OT=1 |Res|P|I|      Object Length (bytes)    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Ingress Node IPv4 address                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Cost to Ingress Node                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                          Optional TLVs                        ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Ingress Node Object for IPv4

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|ObjectClass=TBD|  OT=2 |Res|P|I|      Object Length (bytes)    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Ingress Node IPv6 address(16 bytes)            |
~                                                                ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Cost to Ingress Node                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                          Optional TLVs                        ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Ingress Node Object for IPv6

Objective Functions

New OF defined for BIER-TE path

```
OF Code:     TBD8
Name:        Minimum Bit Sets (MBS)
Description: Find a path represented by BitPositions that has
             the minimum number of bit sets.
```

```
OF Code:     TBD9
Name:        Minimum Bits (MB)
Description: Find a path represented by BitPositions that has
             the minimum bit distance. The bit distance of
             BitPositions is the distance from the lowest bit
             to the highest bit in BitPositions.
```

BIER-TE Path Subobject

New subobject, called BIER-TE Path subobject (or BIER-TE-ERO subobject), is defined to contain the information about one or more BitPositions.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|L| Type = TBDa |     Length    | sub-domain-id |     MT-ID     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:                          BitPositions                         :
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

BIER-TE Path Subobject in ERO

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0| Type = TBDa |  Length = 10  |       0       |       0       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       8       |0 0 0 0 0 0 1 0|       7       |1 0 0 0 0 0 0 0|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       6       |0 0 1 0 1 0 1 0|       0       |0 0 0 0 1 0 1 0|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

BIER-TE Path Subobject for Path {18', 16', 6', 4', 2', 4, 2}

{18'(8:00000010), 16'(7:10000000), 6'(6:00100000), 4'(6:00001000),
2'(6:00000010), 4 (0:00001000), 2 (0:00000010)}.

BIER-TE Path Subobject in RRO

BIER-TE Path Subobject in RRO (Record Route Object) has the same format as BIER-TE Path subobject in ERO except for L flag.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type = TBDa |     Length    | sub-domain-id |     MT-ID     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:                          BitPositions                         :
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

BIER-TE Path Subobject in RRO

Extensions to PCEP Messages

Every PCEP message for BIER-TE path MUST include

- SRP containing PATH-SETUP-TYPE TLV with PST = TBD1 for path setup using BIER-TE

PCRpt includes

- BIER-TE path represented by a BIER-TE path subobject or
- Multicast Traffic TLV in SRP
- Flags in SRP set for egress join or leave

PCUpd includes

- BIER-TE path represented by a BIER-TE path subobject

PCInitiate includes

- BIER-TE path represented by a BIER-TE path subobject
- Multicast Traffic TLV in SRP

PCReq includes

- Multicast Traffic TLV in SRP
- OF for BIER-TE path or
- Flags in SRP set for computing ingress

PCRep includes

- BIER-TE path represented by a BIER-TE path subobject or
- Ingress nodes represented by ingress node objects

Next Step

Comments

PCEP-LS: PCEP extensions for Distribution of Link-State and TE Information

draft-dhodylee-pce-pcep-ls-20

Gyan Mishra, Verizon

A very quick recap...

- Use of PCEP to also learn the network topology and state
- Applicable to Device to controller as well as controller to controller (H-PCE)
- Complementary extension (or another tool in the tool-box)
  - Not a replacement for running IGP in your network!
  - Or BGP-LS, Or Netconf!
  - Enable use of a single control plane protocol as an SBI in some scenarios
- A new PCEP Message and Object and reuse the TLVs already defined
  - Default is local-only (remote learned information can be enabled)

A rough summary of where we left off...

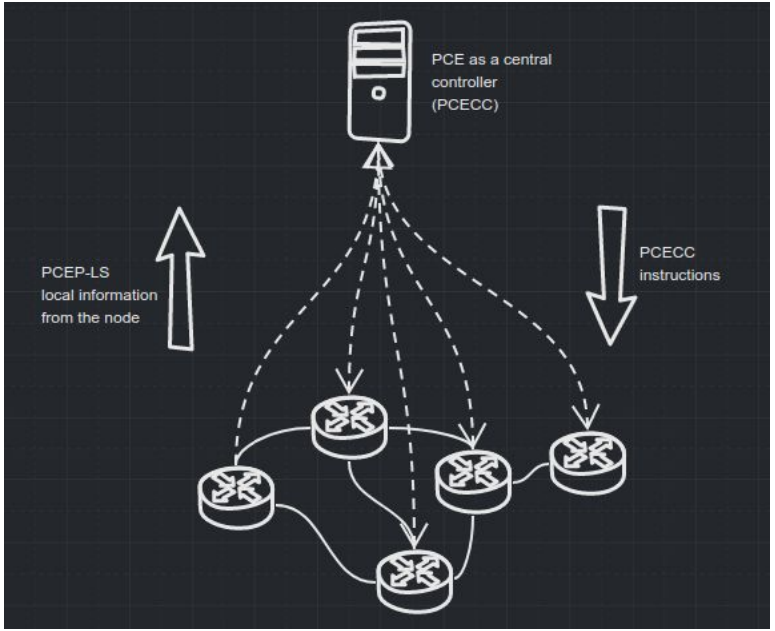
- Presence of other ways to do this
  - and some consider them to be better!
- PCEP scalability worries!
- Operational Complexity!
- Does this require multi-vendor inter-operable RFC?

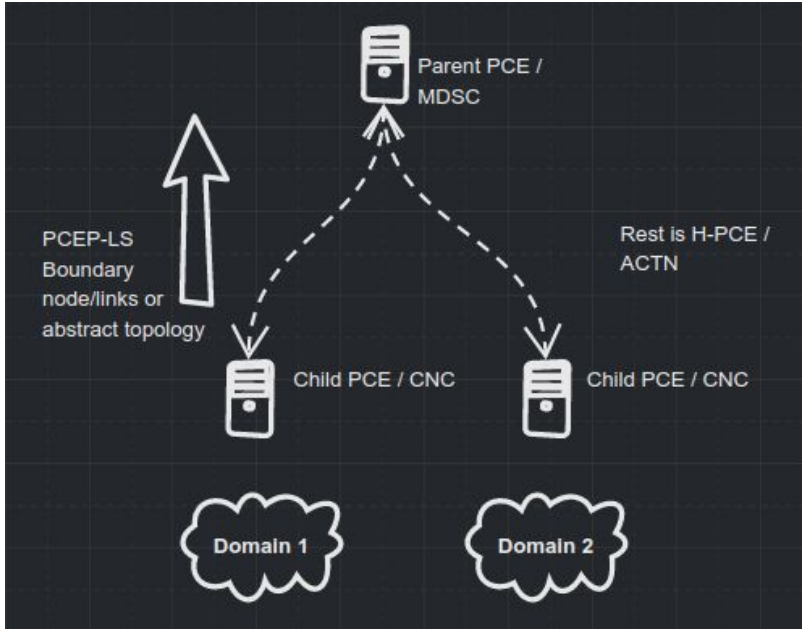
- Consensus on use of PCEP as SBI
- In some PCECC scenarios, there is a direct PCEP session with the nodes
  - Leveraging the direct PCEP session to also learn topology (and changes) is an attractive option!
- Usefulness in H-PCE, Inter-layer, Optical etc
- Another tool in the tool-box (and not replacing any other mechanism)
  - For instance we recognize that some may want to use YANG Path computation RPC instead of PCEP in some scenarios and we support both approaches!

Some Use Cases & Scenarios where PCEP-LS is an attractive choice!

- PCECC
  - Some use cases require direct PCEP session to all nodes
  - Reusing the same session to also learn local network state is attractive
  - Enable the possibility of a single SBI protocol for some use cases
- H-PCE (and ACTN)
  - Between controllers for boundary nodes/links as part of the abstract domain topology
- Partial
  - Some information such as Optical extension learned via PCEP-LS for faster learning
  - Reusing PCEP synchronization optimization techniques and incremental updates
  - Other mechanism can co-exist

Flow of information/control





Question to the WG

- Is there enough interest by some in the WG to work on this?
- Are there targeted experiments, demo, implementations
  - Some were showcased in the past in Hackathon and Bits-n-Bytes
  - Some open source implementation exist and documented
  - Some researchers have shown interest and experimented
  - Some operators have shown interest
- Is there a possibility of a somewhat rough consensus/support for this as an Experimental I-D?
  - Scope of the experiment and results to collect would be the next step!

Useful References

- Chairs Slide from IETF 101: https://datatracker.ietf.org/meeting/101/materials/slides-101-pce-update-on-pcep-sdn-discussion-00.pdf
- Mailing List Thread: https://mailarchive.ietf.org/arch/msg/pce/TXS2v8tXWCxXmp8Vxx59K2dOwCg/
- Implementation: https://mailarchive.ietf.org/arch/msg/pce/0zEEJv-u7mQ1drkkWkAJXLQnDpo/ and https://mailarchive.ietf.org/arch/msg/pce/HF_X3oUS7rIrjyymaw7miUQurpI/
- Researcher: https://mailarchive.ietf.org/arch/msg/pce/p1vKMyCWVxAd-Dpb5IcKX42BcVA/

Thank You!

Backup

- Scalability Concern
  - Some PCECC scenarios already have session to all nodes
  - Reusing the same session to also carry local node information is okay
  - Bulk of the work during PCEP session establishment and before any other PCEP interactions!
- Some benefits of PCEP-LS procedures
  - Incremental changes only
    - Use of stateful techniques: LS-ID to uniquely identify node/link and only the attributes that have changed need to be encoded
  - Synchronization Optimization techniques for PCEP
    - Can be leveraged for PCEP-LS as well during session up/down

RSVP Color in PCEP

9th Mar 2021

Balaji Rajagopalan (balajir@juniper.net)

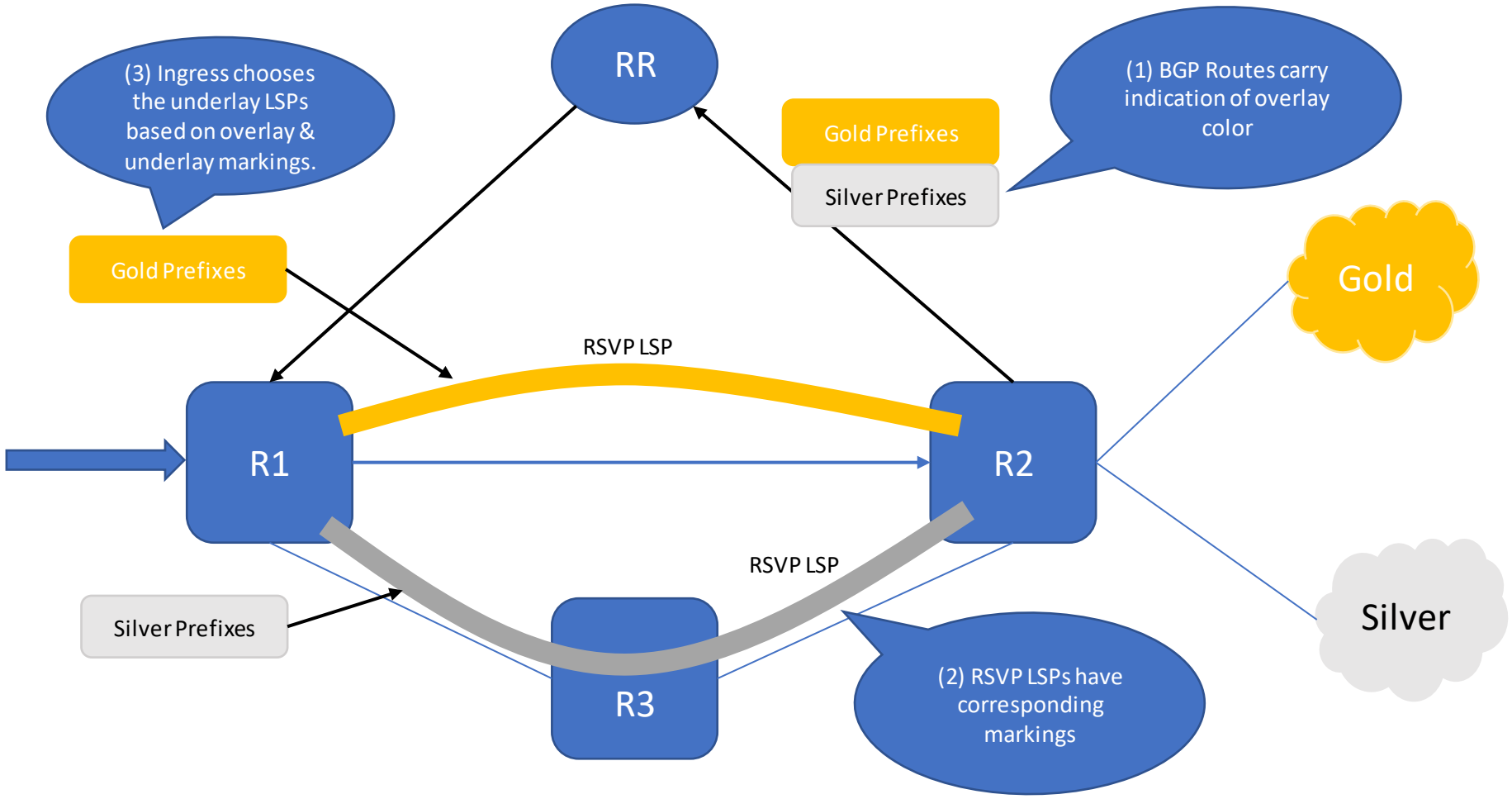
Vishnu Pavan Beeram (vbeeram@juniper.net)

Gyan Mishra (hayabusagsm@gmail.com)

PROBLEM STATEMENT

- Provide a convenient way to associate service prefixes with RSVP underlay tunnels
- Need has existed for some time. Existing solutions using ‘color’ marking for SR-TE:
  - https://tools.ietf.org/html/draft-ietf-idr-segment-routing-te-policy-11 carries SR-TE color in BGP SR-TE NLRI
  - https://tools.ietf.org/html/draft-barth-pce-segment-routing-policy-cp-06 carries SR-TE color in PCEP
- Carrying color markings across domain boundaries in BGP:
  - https://tools.ietf.org/html/draft-kaliraj-idr-bgp-classful-transport-planes-06
- Existing RSVP deployments have a similar need.

ENVISIONED USAGE



SOLUTION OVERVIEW: PCEP PROTOCOL

- LSP Object carries color marking in a new TLV
  - Can be used by PCE to create an LSP with the appropriate color
  - Can be used by PCC to report color
- ‘Color’ is a property of a tunnel, rather than individual LSP’s of the tunnel.
  - Attach with “primary” LSP

Type (2 bytes)	Length(2 bytes)
Color (4 bytes)	

Note: Need to reconcile with https://tools.ietf.org/html/draft-peng-pce-te-constraints-04

MISMATCH HANDLING

PCE	PCC	Implication
Y	N	LSP can't honor color. Server needs to know.
N	Y	The values reported by the PCC are ignored by server, which is benign. But, client can as well not insert color.

STATEFUL-PCE-CAPABILITY TLV

Type (2 bytes)	Length (2 bytes)
Flags (4 bytes) A new C-bit to represent color	

Usage with BGP-CT

- While BGP-CT is not a pre-requisite for using the color specified in this draft, BGP-CT & RSVP PCEP color can inter-operate.
- In BGP-CT, overlay marking (mapping community) selects "resolution scheme"
  - Resolution scheme is associated with an ordered set of transport classes
  - The PCEP RSVP 'color' field can be used to associate LSP's with transport classes

THANK YOU!