Networking Working Group
Internet-Draft
Intended status: Standards Track
Expires: October 20, 2021

R. Chen
S. Peng
ZTE Corporation

April 18, 2021

IGP Extensions to Support Flex-Algorithm Aware Traffic Engineering draft-cp-lsr-fa-aware-te-00

Abstract

[I-D.ietf-lsr-flex-algo]proposes a solution that allows IGPs themselves to compute constraint based paths over the network, and it also specifies a way of using Segment Routing (SR) Prefix-SIDs and SRv6 locators to steer packets along the constraint-based paths.

[RFC8570] describes IS-IS extensions to distribute network-performance information (such as residual bandwidth, and available bandwidth).

This draft describes the IGP extensions to advertise the corresponding network-performance information of the Flex-Algorithm.

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 October 20, 2021.

Copyright Notice

Table of Contents

1. Introduction

[I-D.ietf-lsr-flex-algo]proposes a solution that allows IGPs themselves to compute constraint based paths over the network, and it also specifies a way of using Segment Routing (SR) Prefix-SIDs and SRv6 locators to steer packets along the constraint-based paths.

[RFC8570] describes IS-IS extensions to distribute network-performance information (such as link delay, delay variation, packet loss, residual bandwidth, and available bandwidth).

The current Flex-Algorithm technology may support the implementation of different QoS (Quality of Service QoS) policies of different algorithms (algorithm) at the forwarding plane. It includes different bandwidth, traffic class of service, queue scheduling policies (such as low-delay queue, Priority Queuing priority queue) and discard policies (such as tail discard and random discard). This is actually the local behavior of the forwarding plane inside a node. However, it's not enough. In order to comprehensively optimize the service traffic running in each Flex-Algorithm and avoid unnecessary traffic congestion, the relevant path control technology needs to be implemented in the control plane. That is, to manage and maintain the bandwidth resource reservation and consumption information for each Flex-algo plane, and use the information for path calculation and orchestration of the Flex-Algorithm.

This draft describes the IGP extensions to advertise the corresponding network-performance information of the Flex-Algorithm.

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. IS-IS Extensins for FA-TE

This document registers new IS-IS FA-TE sub-TLVs in the "Sub-TLVs for TLVs 22, 23, 141, 222, and 223" registry. These new sub-TLVs provide ways to distribute network-performance information of the Flex-Algorithm.

This document registers several sub-TLVs:

Type	Description
TBD	Maximum Reservable Link Bandwidth per Algorithm
TBD	Available Bandwidth per Algorithm
TBD	Utilized Bandwidth per Algorithm

3.1. Maximum Reservable Link Bandwidth per Algorithm Sub-TLV

This sub-TLV advertises maximum reservable link bandwidth attributes on a particular link for a Flex-Algorithm.

The Maximum Reservable Link Bandwidth per Algorithm Sub-TLV has the following format:

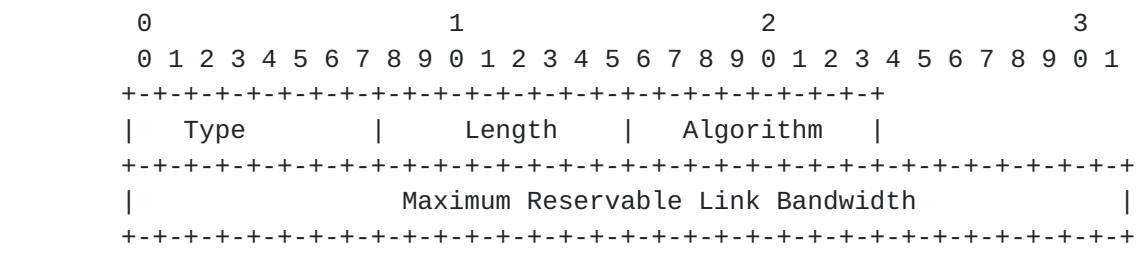

Figure 1

Type: TBD1 (Suggested value to be assigned by IANA)

Length: 1 octets.

Algorithm: 1 octets. Flex-Algorithm number,the value between 128 and 255 inclusive,the same as the definition of Flex-Algorithm in [I-D.ietf-lsr-flex-algo]. Indicates that the value in the Maximum Reserved Link Bandwidth field is for a specific Flex-algo plane.

Maximum Reserved Link Bandwidth: 4 octets.Indicates the maximum link bandwidth that can be reserved for a specific FA-algorithm.

Note that the maximum reserved link bandwidth irrelevant to Flex-algo can be classified as MRLB-FA corresponding to algorithm 0. Therefore, the Sub-TLV (set the Algorithm field to 0) or the Maximum Reservable Link Bandwidth Sub-TLV defined by [RFC5305] can be used.If a node receives two types of notifications at the same time, the bandwidth value carried in the conventional known Sub-TLV is preferred.

3.2. Available Link Bandwidth per Algorithm Sub-TLV

This Sub-TLV advertises the available bandwidth on a particular link for a specific Flex-Algorithm.

The Available Link Bandwidth per Algorithm Sub-TLV has the following format:

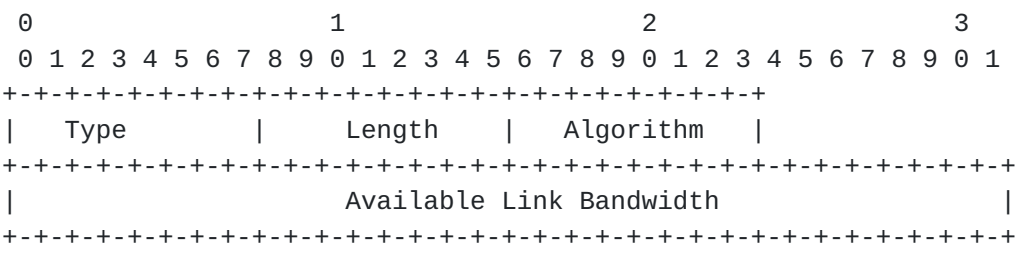

Figure 3

Type: TBD3 (Suggested value to be assigned by IANA)

Length: 1 octets.

Algorithm: 1 octets. Flex-Algorithm number,the value between 128 and 255 inclusive,the same as the definition of Flex-Algorithm in [I-D.ietf-lsr-flex-algo]. Indicates that the value in the Available Link Bandwidth field is for a specific Flex-algo plane.

Available Link Bandwidth: 4 octets. Indicates the available link bandwidth that can be reserved for a specific FA-algorithm.

Note that the available link bandwidth irrelevant to Flex-algo can be classified as ARLB-FA corresponding to algorithm 0. Therefore, the Sub-TLV (set the Algorithm field to 0) or the Unidirectional Available Bandwidth Sub-TLV defined by [RFC8570] can be used.If a node receives two types of notifications at the same time, the bandwidth value carried in the conventional known Sub-TLV is preferred.

3.3. Utilized Link Bandwidth per Algorithm Sub-TLV

This Sub-TLV advertises the utilized bandwidth on a particular link for a specific Flex-Algorithm.

The Utilized Link Bandwidth per Algorithm Sub-TLV has the following format:

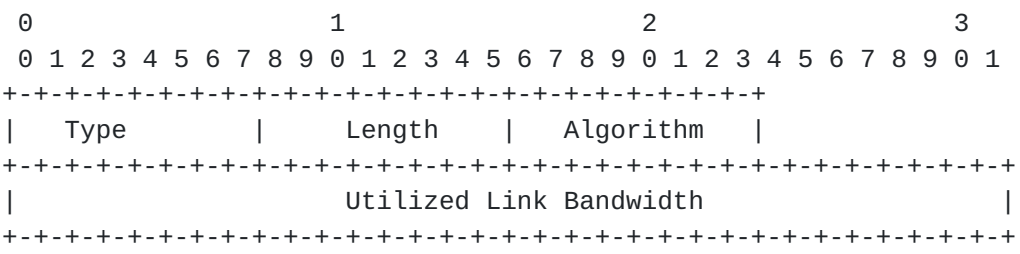

Figure 4

Type: TBD4 (Suggested value to be assigned by IANA)

Length: 1 octets.

Algorithm: 1 octets. Flex-Algorithm number,the value between 128 and 255 inclusive,the same as the definition of Flex-Algorithm in [I-D.ietf-lsr-flex-algo]. Indicates that the value in the Utilized Link Bandwidth field is for a specific Flex-Algorithm plane.

Utilized Link Bandwidth: 4 octets. Indicates the Utilized link bandwidth that can be reserved for a specific Flex-Algorithm.

Note that the Utilized link bandwidth irrelevant to Flex-algo can be classified as ULB-FA corresponding to algorithm 0. Therefore, the Sub-TLV (set the Algorithm field to 0) or the Unidirectional Utilized Bandwidth Sub-TLV defined by [RFC8570] can be used.If a node receives two types of notifications at the same time, the bandwidth value carried in the conventional known Sub-TLV is preferred.

4. OSPF Extensins for FA-TE

OSPF extensions for FA-TE will be defined in next version.

5. Examples

Assume two Flex-Algorithm, FA128 and FA129, with the following configuration parameters:

```
Per TE link resource allocation:
Max-BW(interface): 100G (derived from the physical interface BW)
    Max-resv-BW(interface): 100G (default: equal to Max-BW)

Flex-Algorithm:
Max-resv-link-BW(FA128): 10G
Max-resv-link-BW(FA129): 20G
```

Time0: no LSPs

When the controller is used for path calculation, the controller first collects the network topology information through the BGP-LS, which contains the MRLB-FA/ULB-FA/ALB-FA information of each link. Static bandwidth reservation information maintained by the controller:

TE link	Flex-Algorithm	Bandwidth		
Link 1	FA0	MRLB-FA = 70G	ULB-A = 0	ALB-A = 70G
	FA128	MRLB-FA = 10G	ULB-A = 0	ALB-A = 10G
	FA129	MRLB-FA = 20G	ULB-A = 0	ALB-A = 20G

Time1: Create a new TE path1 in FA 128 with reserved bandwidth of 5G

The controller updates the reserved static bandwidth information:

TE link	Flex-Algorithm	Bandwidth		
Link 1	FA0	MRLB-FA = 70G	ULB-FA = 0	ALB-FA = 70G
	FA128	MRLB-FA = 10G	ULB-FA = 0	ALB-FA = 10G
		(TE path1 statically reserve 5G, remaining 5G}		
	FA129	MRLB-FA = 20G	ULB-FA = 0	ALB-FA = 20G

```
     +-+-+-+-+-+-+-|-+-+-+-+-+-+-+-+-|-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
+-+-+-+-+-+-+-+
```

Time3: After the TE path1 starts to carry traffic, suppose the service traffic on the TE path1 runs at full load.The controller updates the reserved static bandwidth information:

```
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
+-+-+-+-+-+-+-+-+
     |  TE link    |  Flex-Algorithm |
Bandwidth                          |
     |-+-+-+-+-+-+-|-+-+-+-+-+-+-+-+-|-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
+-+-+-+-+-+-+-+-|
     |             |    FA0          |        MRLB-FA = 70G    | ULB-FA = 0  |
ALB-FA = 70G   |
     |             |-+-+-+-+-+-+-+-+-|-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
+-+-+-+-+-+-+-+-|
     |  Link 1     |    FA128        | MRLB-FA = 10G           | ULB-FA = 5G |
ALB-FA = 5G    |
     |             |                 | (TE path1 statically    |
|                 |
     |             |                 |reserve 5G, remaining 5G}|
|                 |
     |             |-+-+-+-+-+-+-+-+-|-+-+-+-+-+-+-+-+-+-+-+-+-|-+-+-+-+-+-+-|-
+-+-+-+-+-+-+-+-|
     |             |    FA129        |     MRLB-FA = 20G       | ULB-FA = 0  |
ALB-FA = 20G   |
     +-+-+-+-+-+-+-|-+-+-+-+-+-+-+-+-|-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
+-+-+-+-+-+-+-+-+
```

Now suppose to continue to create another TE path2 in the Flex-algo 128. This TE path needs to reserve 6G bandwidth. When the controller is used for path computation, path computation will fail, Beacuse there is no link resource that meets the bandwidth requirement.

6. IANA Considerations

IANA maintains the registry for the sub-TLVs. IANA has registered the following sub-TLVs in the "Sub-TLVs for TLVs 22, 23, 141, 222, and 223" registry:

Type	Description
TBD	Maximum Reservable Link Bandwidth per Algorithm
TBD	Residual Bandwidth per Algorithm
TBD	Available Bandwidth per Algorithm
TBD	Utilized Bandwidth per Algorithm

7. Security Considerations

TBD.

8. Acknowledgements

TBD.

9. Normative References

[I-D.ietf-lsr-flex-algo] Psenak, P., Hegde, S., Filsfils, C., Talaulikar, K., and A. Gulko, "IGP Flexible Algorithm", draft-ietf-lsr-flex-algo-13 (work in progress), October 2020.

[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>.

[RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, DOI 10.17487/RFC5305, October 2008, <https://www.rfc-editor.org/info/rfc5305>.

[RFC8570] Ginsberg, L., Ed., Previdi, S., Ed., Giacalone, S., Ward, D., Drake, J., and Q. Wu, "IS-IS Traffic Engineering (TE) Metric Extensions", RFC 8570, DOI 10.17487/RFC8570, March 2019, <https://www.rfc-editor.org/info/rfc8570>.

Authors' Addresses

Ran Chen
ZTE Corporation

Email: chen.ran@zte.com.cn

Shaofu Peng
ZTE Corporation

Email: peng.shaofu@zte.com.cn