

TEAS
Internet-Draft
Intended status: Standards Track
Expires: January 14, 2021

Z. Zhang
S. Hegde
Juniper Networks
A. Gulko
Refinitiv
July 13, 2020

Network Slicing with Flexible Traffic Engineering
draft-zzhang-teas-network-slicing-with-flex-te-00

Abstract

This document specifies procedures and signaling enhancements to Flexible Algorithm to ease provisioning and to scale it better via Flexible Traffic Engineering, which is an integration of Flexible Algorithm and Segment Routing [RFC8402] Traffic Engineering (SR-TE).

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [BCP 14](#) [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

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 January 14, 2021.

Copyright Notice

Copyright (c) 2020 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](https://trustee.ietf.org/license-info) and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	2
1.1.	FlexAlgo Background	3
1.2.	Central Provisioning and Signaling of FlexAlgo	4
1.3.	Flexible Traffic Engineering (FlexTE) and Targeted Signaling	4
1.4.	Traffic Isolation and FlexAlgo/FlexTE	5
2.	Specification	6
2.1.	Southbound BGP-LS Encoding of FAD	6
2.2.	Southbound BGP-LS Encoding of Link Administrative Group Information	7
2.3.	OSPF/ISIS Encoding of Link Administrative Group Information for Cetralized Advertisement	7
2.4.	FlexAlgo and Link AG Signaling from Controllers	8
3.	Security Considerations	8
4.	IANA Considerations	8
5.	Acknowledgements	8
6.	References	9
6.1.	Normative References	9
6.2.	Informative References	9
	Authors' Addresses	10

[1.](#) Introduction

[dong-network-slicing-problem-statement] defines Network Slicing Problem Statement for IP/MPLS networks. While Virtual Private Networks (VPNs) have been widely deployed to provide many different virtual networks on the same physical operator network, and can be reused to provide network slicing service to applications, currently those VPNs share the same underlay operator network without any separation and isolation.

Multi-Topology Routing (MTR) [[RFC4915](#)] [[RFC5120](#)] provides a mechanism to have a set of independent IP topologies referred to as Multi-Topologies (MTs) over the same underlay network. It can be used to provide separation and isolation required by network slicing, though MTR has not been widely deployed over the years, except limited usage

of maintaining separate IGP routing domains for isolated multicast islands within the backbone.

Some reasons for MTR's lack of success are listed below:

1. Lack of strong demand for mapping traffic to different MTs
2. Lack of good mechanism for mapping traffic to different MTs
3. Lack of operating tools to ease provisioning and monitoring

In 5G era, 1) is no longer the case and 3) needs to be addressed given the network slicing requirements. 2) is addressed by SR and Flexible Algorithm (FlexAlgo) [[ietf-lsr-flex-algo](#)]. This document specifies signaling enhancements to FlexAlgo to ease provisioning and to scale it better via Flexible Traffic Engineering (FlexTE), which is an integration of FlexAlgo and Segment Routing [[RFC8402](#)] Traffic Engineering (SR-TE).

1.1. FlexAlgo Background

FlexAlgo can be viewed as a more flexible and light-weighted mechanism of MTR. A Flexible Algorithm is defined as a <Calculation Type, Metric Type, Included/excluded Administrative Group> tuple. The "Included/excluded Administrative Group" defines the (sub-)topology used for the algorithm. While there are different metric types to be used, there are no per-algorithm metric values advertised for links.

Routers are configured to use certain algorithms for its SPF calculations. Definitions for the algorithms are locally configured and/or learnt through signaling.

The Administrative Groups (AG) of a link, which are often referred to as link colors, are advertised in a 32-bit AG BitMask as specified in [[RFC3630](#)] [[RFC5305](#)] or an arbitrary length EAG BitMask as specified

in [\[RFC7308\]](#). The advertisements are originated from the router owning the link based on local provisioning.

While not a mandatory part of FlexAlgo, Segment Routing can be integrated with FlexAlgo seamlessly to map traffic to different algorithms: prefix SIDs can optionally be associated with algorithms, so that a prefix can be reached via different SIDs or SID lists, following different paths.

[1.2.](#) Central Provisioning and Signaling of FlexAlgo

More and more operators use controllers for centralized orchestration, provisioning and signaling. In fact, even before the controllers were used, the network planning was centralized, albeit done manually, and then configuration information resulting from centralized planning was entered into individual routers via out of band means.

The centralized model can be applied to FlexAlgo very well - instead of provisioning FAD and link AGs on individual routers after centralized planning, the provisioning is be done centrally on the controllers and then constraints and link AG information are signaled to routers.

Given that it is very common for controllers to learn network information via northbound BGP-LS [\[RFC7752\]](#), this document uses southbound BGP-LS to distribute FAD and link AG information from controllers to BGP-LS speakers. This can also take advantages of inherent BGP mechanisms for optimized large scale state distribution. If not all routers but only IGP border routers run BGP-LS, the border routers will then flood received information via IGP.

[1.3.](#) Flexible Traffic Engineering (FlexTE) and Targeted Signaling

With current FlexAlgo mechanisms, Flexible Algorithm Definitions (FADs) and link AG information are flooded throughout an IGP area and every router will do an SPF calculation for each algorithm. This may work for a few algorithms but it will not scale for larger number of

algorithms that are necessary for large number of slices in some 5G scenarios.

To address the scaling problem, the above flooded information and SPF caculation may be restricted to network edge only. The idea is first introduced in [[I-D.drake-bess-enhanced-vpn](#)] but adapted to FlexAlgo in this document.

With this scheme, the internal routers don't have per-algorithm information and do not do per-algorithm based SPF or per-algorithm prefix-SID based forwarding. The edge routers use SR-TE Adjacency SID-lists to explicitly steer traffic through the network. This is referred to as Flexible Traffic Engineering (FlexTE), an integraion of Flexible Algorithm and SR Traffic Engineering.

Specifically, with BGP Route Target [[RFC4364](#)] and Route Target Constrains [[RFC4684](#)] mechanisms, the FADs and link AG information are only propagated to and imported by edge routers that need that information. For example, if a network slice is presented to

application as a VPN and instantiated in the underlay with a Flexible Algorithm that utilizes only "red" links, then that specific FAD and "red" link AG information are advertised to and imported by only PEs for that VPN (if the same algorithm is used by many VPNs then all PEs of those VPNs will import the relevant information).

For better scalability, the link AG information is encoded in a new type of Route Target (RT) used for the control of route propagation and importation, as detailed in [Section 2.2](#) and [[I-D.zzhang-idr-bitmask-route-target](#)].

Because a router may not be able to push too deep a label stack, per-algorithm Binding SIDs may have to be used. For example, if there are 10 hops from PE1 to PE2 while the maximum labels that PE1 can push is 5, then PE1 has to use a label stack that specifies the explicit hop-by-hop path (calculated by an algorithm) to an intermediate router P1 and a binding SID advertised by P1 for PE2. For P1 to calculate the per-algorithm explicit path to PE2, it also needs to know the information for that algorithm, and it can do so following the same way as how PE1 learns the information.

[1.4.](#) Traffic Isolation and FlexAlgo/FlexTE

FlexAlgo as described in [[I-D.ietf-lsr-flex-algo](#)] separates the routing domain into different planes. The primary and backup paths are computed based on the topology that corresponds to the plane. FlexAlgo provides strict isolation of the data traffic between the different planes. Notice that FlexAlgo is suitable for slices that need complete isolation. Packet transport networks are expected to have a limited number of such isolated routing planes.

With FlexTE, traffic isolation is achieved via SR-TE Adjacency SID lists, but during local Fast ReRoute (FRR) traffic may flow through paths that don't satisfy constraints. If the SR-TE SID list is too long, Node SIDs may be used but the traffic isolation is not possible on the path between node SIDs, unless some internal routers also get targeted signaling, behave as edge routers, and advertise per-algorithm Node/Binding SIDs (targeted at the edge routers and those selected internal routers). Therefore, FlexTE is more suitable for soft slicing where traffic isolation is not critical in certain situations.

With 5G, network slicing requires high number of slices though they may not necessarily require routing plane isolation but they may need to satisfy certain constraints and have guaranteed Quality Of Service, and FlexTE as a flexible soft slicing solution allows for slice creation inside specific isolated planes or in a generic plane.

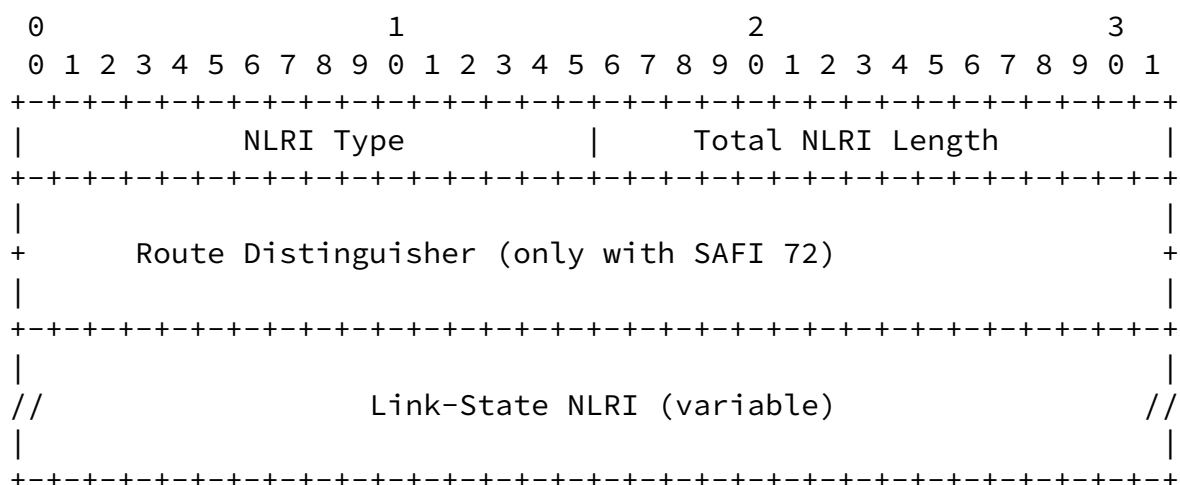
The QOS guarantees for the slices are outside the scope of this document.

[2.](#) Specification

BGP-LS [[RFC7752](#)] is for "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP". This document extends it for south-bound distribution of FlexAlgo/TE constraint related information, and specifies relevant procedures for FlexAlgo based on centralized, targeted signaling.

[2.1.](#) Southbound BGP-LS Encoding of FAD

Currently BGP-LS uses the following NLRI format with AFI 16388 and SAFI 71/72:



A new NLRI type TBD1 is added to advertise FAD. For simplicity, the variable Link-State NLRI field has the exactly same TLV format of ISIS FAD Sub-TLV as specified in [[I-D.ietf-lsr-flex-algo](#)], including the Type number and Sub-TLVs.

Type	NLRI Type
1	Node NLRI
2	Link NLRI
3	IPv4 Topology Prefix NLRI
4	IPv6 Topology Prefix NLRI
TBD1	Flexible Algorithm Definition

2.2. Southbound BGP-LS Encoding of Link Administrative Group Information

Currently, BGP-LS encodes link Administrative Group information as a Type 1088 Administrative Group TLV in BGP-LS attribute attached to Link NLRI that are propagated northbound from routers to controllers. For the southbound signaling of Administrative Group information from controllers, for the purpose of targeted propagation

and importation, the Administrative Group information are encoded in a new Bitmask Route Target as specified in [I-D.zhang-idr-bitmask-route-target]. The Administrative Group TLV is omitted from the BGP-LS Attribute for the link because the information is already encoded in the BitMask RT.

Specifically, the EAG BitMask is encoded into the Bitmask field of a Bitmask RT that is attached to the Link NLRI for the link. The Global Administrator (GA) Type, GA Length, and Local Administrator fields are set according to the operator's need to provide a context.

To distinguish from Link NLRIs signaled northbound by routers, the Protocol-ID of the Link NLRI is set to BGP (to be assigned by IANA).

[2.3.](#) OSPF/ISIS Encoding of Link Administrative Group Information for Centralized Advertisement

When centralized provisioning and signaling is not used, an OSPF router advertises its local links' attributes in OSPFv2 Extended Link Opaque LSAs. The LSA includes OSPFv2 Extended Link TLVs, one for each link, which in turn includes sub-TLVs for specific link attributes.

The same OSPFv2 Extended Link TLVs can be used for ABRs to flood link attributes that are centrally provisioned on and signaled from controllers, but they MUST additionally carry a new sub-TLV to indicate the routers that host the links, because these Extended Link TLVs are in the Extended Link Opaque LSAs originated by the ABRs not those originated by the routers hosting the links. The sub-TLV is called Hosting Router sub-TLV, with a new TBD2 type and a 4-octet value for the Router ID of the router hosting the link.

For OSPFv3, a router advertises its local links' TE attributes in Intra-Area-TE LSAs, which contains Link TLVs with link attribute sub-TLVs. Similarly to OSPFv2, when ABRs flood the link attributes that are centrally provisioned on and signaled from controllers, the Link TLVs MUST carry the Hosting Router sub-TLV.

For ISIS, the Link Administrative Group information is signaled as sub-TLVs in Extended IS Reachability TLV [[RFC5305](#)]. Similarly, when

signaled from controllers, the Extended IS Reachability TLV MUST carry a new Hosting System sub-TLV. The sub-TLV has a new type TBD3 and a 7-octet value for system ID and pseudonode number.

When a router receives a OSPFv2/OSPFv3 Link TLV with Hosting Router sub-TLV or an ISIS Extended IS Reachability TLV with Hosting System sub-TLV, it MUST associate the link with the advertised hosting router/system, not with the originator of the OSPF LSA or ISIS LSP.

[2.4.](#) FlexAlgo and Link AG Signaling from Controllers

With centralized provisioning and signaling, a controller signals Link AG information using BGP-LS Link NLRI with a BitMask RT attached, as specified in [Section 2.2](#).

The controller signals FADs used in the domain using the BGP-LS NLRI type TBD1. The updates carry a Bitmask RT with the bits set for the AGs that the FADs care about.

Routers that need to learn the information MUST have a BitMask RT locally configured, with the bits set for the AGs that they care about, so that they will import corresponding NLRIs. In case of FlexTE, only edge routers and some internal routers will have the BitMask RT locally configured. Otherwise, all BGP-LS routers are configured with a BitMask RT to import all FAD and Link NLRIs.

To optimize the propagation of south-bound BGP-LS NLRIs, Route Target Constrain [[RFC4684](#)] mechanisms SHOULD be used for Bitmask RT as well, as specified in [I-D.zhang-idr-bgp-rt-constrain-extension].

[3.](#) Security Considerations

To be added.

[4.](#) IANA Considerations

To be added.

[5.](#) Acknowledgements

The authors thank Jeff Haas, Srihari Sangli and Colby Barth for their comments and suggestions.

6. References

6.1. 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-07](#) (work in progress), April 2020.
- [I-D.zhang-idr-bgp-rt-constrains-extension]
Zhang, Z. and J. Haas, "Route Target Constrain Extension", [draft-zhang-idr-bgp-rt-constrains-extension-00](#) (work in progress), July 2020.
- [I-D.zhang-idr-bitmask-route-target]
Zhang, Z., Ramachandra, S., and J. Haas, "Bitmask Route Target", [draft-zhang-idr-bitmask-route-target-00](#) (work in progress), July 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>>.
- [RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", [RFC 7752](#), DOI 10.17487/RFC7752, March 2016, <<https://www.rfc-editor.org/info/rfc7752>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in [RFC 2119](#) Key Words", [BCP 14](#), [RFC 8174](#), DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.
- [RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", [RFC 8402](#), DOI 10.17487/RFC8402, July 2018, <<https://www.rfc-editor.org/info/rfc8402>>.

6.2. Informative References

- [I-D.dong-network-slicing-problem-statement]
Dong, J. and S. Bryant, "Problem Statement of Network Slicing in IP/MPLS Networks", [draft-dong-network-slicing-problem-statement-00](#) (work in progress), October 2016.

Internet-Draft

slicing-with-flex-te

July 2020

[I-D.drake-bess-enhanced-vpn]

Drake, J., Farrel, A., Jalil, L., and A. Lingala, "BGP-LS Filters : A Framework for Network Slicing and Enhanced VPNs", [draft-drake-bess-enhanced-vpn-03](#) (work in progress), May 2020.

[RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", [RFC 3630](#), DOI 10.17487/RFC3630, September 2003, <<https://www.rfc-editor.org/info/rfc3630>>.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", [RFC 4364](#), DOI 10.17487/RFC4364, February 2006, <<https://www.rfc-editor.org/info/rfc4364>>.

[RFC4684] Marques, P., Bonica, R., Fang, L., Martini, L., Raszuk, R., Patel, K., and J. Guichard, "Constrained Route Distribution for Border Gateway Protocol/MultiProtocol Label Switching (BGP/MPLS) Internet Protocol (IP) Virtual Private Networks (VPNs)", [RFC 4684](#), DOI 10.17487/RFC4684, November 2006, <<https://www.rfc-editor.org/info/rfc4684>>.

[RFC4915] Psenak, P., Mirtorabi, S., Roy, A., Nguyen, L., and P. Pillay-Esnault, "Multi-Topology (MT) Routing in OSPF", [RFC 4915](#), DOI 10.17487/RFC4915, June 2007, <<https://www.rfc-editor.org/info/rfc4915>>.

[RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", [RFC 5120](#), DOI 10.17487/RFC5120, February 2008, <<https://www.rfc-editor.org/info/rfc5120>>.

[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>>.

[RFC7308] Osborne, E., "Extended Administrative Groups in MPLS Traffic Engineering (MPLS-TE)", [RFC 7308](#),

Authors' Addresses

Zhaohui Zhang
Juniper Networks

EMail: zzhang@juniper.net

Zhang, et al.

Expires January 14, 2021

[Page 10]

Internet-Draft

slicing-with-flex-te

July 2020

Shraddha Hegde
Juniper Networks

EMail: shraddha@juniper.net

Arkadiy Gulko
Refinitiv

EMail: arkadiy.gulko@refinitiv.com

