Workgroup: Internet Engineering Task Force
Internet-Draft:
draft-scudder-bgp-entropy-label-00
Published: 28 April 2022
Intended Status: Informational
Expires: 30 October 2022
Authors: J. G. Scudder
Juniper Networks
K. Kompella
Juniper Networks

BGP Entropy Label Capability, Version 2

Abstract

RFC 6790 defined the Entropy Label Capability Attribute (ELC); RFC 7447 deprecated that attribute. This specification, dubbed "Entropy Label Capability Attribute version 2" (ELCv2), was intended to be offered for standardization, to replace the ELC as a way to signal that a BGP protocol speaker is capable of processing entropy labels.

Although ultimately a different specification was chosen for that purpose, at least one implementation of ELCv2 was shipped by Juniper Networks and is currently in use in service provider networks. This document is published in order to document what was implemented.

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 30 October 2022.

Copyright Notice

Table of Contents

1. Introduction

[RFC6790] defines the Entropy Label Capability attribute (ELC), an optional, transitive BGP path attribute. For correct operation, it is necessary that any intermediate node modifying the next hop of a route must remove the ELC unless the node so doing is able to process entropy labels. Sadly, these requirements cannot be fulfilled with the ELC as specified, because it is an optional, transitive attribute: by definition, a node that does not support the ELC will propagate the attribute. But such a node might be exactly the one that we desire to remove it.

Ultimately the IDR working group adopted [I-D.ietf-idr-next-hop-capability] as a proposed solution for this and similar problems. However, prior to that, at least one implementation of this specification was shipped, by Juniper Networks. The shipping implementation uses the code point that was assigned by RFC 6790, and deprecated by RFC 7447. This document explains what was implemented and deployed, dubbed "Entropy Label Capability Attribute version 2" (ELCv2).

Although [I-D.ietf-idr-next-hop-capability] uses an optional, non-transitive path attribute, at the time ELCv2 was developed it was decided that an optional, non-transitive solution would over-constrain the deployment options available -- in many cases (for example, route reflectors) it's fine that an intermediate node does propagate an ELC even if it doesn't itself have the ability to process entropy labels.

Instead, in this specification, we take the approach of carrying a copy of the next hop information in the ELCv2. This allows the node processing it to know if it can rely on the information carried therein, while still allowing it to be propagated by all intermediate nodes.

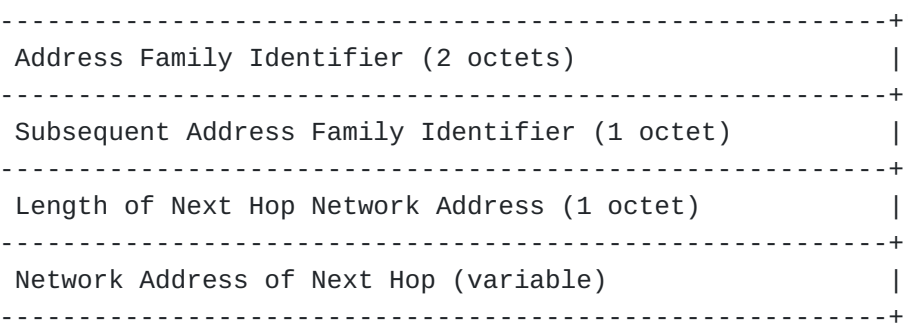
1.1. Requirements Language

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.

2. Entropy Label Capability Path Attribute, Version 2

The Entropy Label Capability Path Attribute, Version 2 (ELCv2) is an optional, transitive BGP attribute (for the attribute type code, see Section 3). The ELCv2 has as its data a network layer address, representing the next hop of the route the ELCv2 accompanies. The ELCv2 signals a useful optimization, so it is desirable to make it transitive; the next hop data is to ensure correctness across BGP speakers that do not understand the ELCv2.

The Attribute Data field of the ELCv2 path attribute is encoded as shown below:



The meanings of the fields are as given in Section 3 of [RFC4760].

When BGP [RFC4271] is used for distributing labeled Network Layer Reachability Information (NLRI) as described in, for example, [RFC8277], the route may include the ELCv2 as part of the Path Attributes. The inclusion of this attribute with a route indicates that the egress of the associated Label Switched Path (LSP) can process entropy labels as an egress Label Switched Router (LSR) for that route -- see Section 4.2 of [RFC6790]. Below, we refer to this for brevity as being "EL-capable."

2.1. Sending the ELCv2

When a BGP speaker S has a route R it wishes to advertise with next hop N to its peer, it MUST NOT include the ELCv2 attribute except if it knows that the egress of the associated LSP L is EL-capable. Specifically, this will be true if S:

- Is itself the egress, and knows itself to be EL-capable, or
- Is re-advertising a BGP route it received with a valid ELCv2 attribute, and is not changing the value of N, or
- Is re-advertising a BGP route it received with a valid ELCv2 attribute, and is changing the value of N, and knows (for example, through configuration) that the router represented by N is either the LSP egress and is EL-capable, or that it will process the outer label(s) without processing the entropy label below, as with a transit LSR, or
- Is redistributing a route learned from another protocol, and that other protocol conveyed the knowledge that the egress of L was EL-capable (for example, this might be known through the LDP ELC TLV, Section 5.1 of [RFC6790]).

In any event, when sending an ELCv2, S MUST set the data portion of the ELCv2 to be equal to N, using the encoding given in Section 2.

The ELCv2 MAY be advertised with routes that are labeled, such as those using SAFI 4 [RFC8277]. It MUST NOT be advertised with unlabeled routes.

We note that due to the nature of BGP optional transitive path attributes, any BGP speaker that does not implement this specification will propagate the ELCv2, the requirements of this section notwithstanding. However, such a speaker will not update the data part of the ELCv2.

2.2. Receiving the ELCv2

When a BGP speaker receives an unlabeled route that includes the ELCv2, it MUST discard the ELCv2.

When a BGP speaker receives a labeled route that includes the ELCv2, it MUST compare the ELCv2's data portion to the next hop of the route. If the two are equal, the egress of the LSP supports entropy labels, which implies that the receiving BGP speaker, if acting as ingress, MAY insert an entropy label below the advertised label, as per Section 4.2 of [RFC6790]. If the two are not equal, either some intermediate router that does not implement this specification modified the next hop, or some router on the path had an incorrect

implementation. In either case, the action taken is the same: the ELCv2 MUST be discarded. The Partial bit MAY be inspected -- if it is equal to zero, then the mismatch must have been caused by an incorrect implementation, and the error MAY be logged.

When a BGP speaker receives a route that includes an ELCv2 whose Attribute Length is less than 4, whose Attribute Length is not equal to 4 plus the value encoded in the Length of Next Hop Network Address carried in the Attribute Data, or whose Attribute Data is otherwise inconsistent with the encoding specified in Section 2, it MUST discard the ELCv2.

3. IANA Considerations

As per [RFC7447], IANA has deprecated BGP attribute 28. That deprecated type code is used by implementations of this specification. IANA is requested to update the references for attribute 28 to include this specification.

4. Security Considerations

Insertion of an ELCv2 by an attacker could cause forwarding to fail. Deletion of an ELCv2 by an attacker could cause one path in the network to be overutilized and another to be underutilized. However, we note that an attacker able to accomplish either of these (below, an "on-path attacker") could equally insert or remove any other BGP path attribute or message. The former attack described above denies service for a given route, which can be accomplished by an on-path attacker in any number of ways even absent ELCv2. The latter attack defeats an optimization but nothing more; it seems dubious that an attacker would go to the trouble of doing so rather than launching some more damaging attack. In sum, the ELCv2 attribute creates no significant issues beyond those analyzed in [RFC4272].

5. Acknowledgements

Thanks to Alia Atlas, Bruno Decraene, Martin Djernaes, John Drake, Adrian Farrell, Keyur Patel, Ravi Singh, and Jim Uttaro for their discussion of this issue. Particular thanks to Kevin Wang for his many valuable contributions.

6. References

6.1. Normative References

[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>.

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI 10.17487/RFC4271, January 2006, <https://www.rfc-editor.org/info/rfc4271>.

[RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, DOI 10.17487/RFC4760, January 2007, <https://www.rfc-editor.org/info/rfc4760>.

[RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and L. Yong, "The Use of Entropy Labels in MPLS Forwarding", RFC 6790, DOI 10.17487/RFC6790, November 2012, <https://www.rfc-editor.org/info/rfc6790>.

[RFC7447] Scudder, J. and K. Kompella, "Deprecation of BGP Entropy Label Capability Attribute", RFC 7447, DOI 10.17487/RFC7447, February 2015, <https://www.rfc-editor.org/info/rfc7447>.

[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>.

6.2. Informative References

[I-D.ietf-idr-next-hop-capability] Decraene, B., Kompella, K., and W. Henderickx, "BGP Next-Hop dependent capabilities", Work in Progress, Internet-Draft, draft-ietf-idr-next-hop-capability-07, 8 December 2021, <https://www.ietf.org/archive/id/draft-ietf-idr-next-hop-capability-07.txt>.

[RFC4272] Murphy, S., "BGP Security Vulnerabilities Analysis", RFC 4272, DOI 10.17487/RFC4272, January 2006, <https://www.rfc-editor.org/info/rfc4272>.

[RFC8277] Rosen, E., "Using BGP to Bind MPLS Labels to Address Prefixes", RFC 8277, DOI 10.17487/RFC8277, October 2017, <https://www.rfc-editor.org/info/rfc8277>.

Authors' Addresses

John G. Scudder
Juniper Networks

Email: jgs@juniper.net

Kireeti Kompella

Juniper Networks

Email: kireeti@juniper.net