Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: July 28, 2018

F. Baker

D. Lamparter
NetDEF

January 24, 2018

IPv6 Source/Destination Routing using IS-IS
draft-ietf-isis-ipv6-dst-src-routing-00

Abstract

This note describes the changes necessary for IS-IS to route IPv6 traffic from a specified prefix to a specified prefix.

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 July 28, 2018.

Table of Contents

1. Introduction

This specification defines how to exchange destination/source routing [I-D.ietf-rtgwg-dst-src-routing] information in IS-IS for IPv6 [RFC5308][IS-IS] routing environments. To this extent, a new sub-TLV for an IPv6 [RFC8200] Source Prefix is added, and Multi Topology Routing [RFC5120] is employed to address compatibility and isolation concerns.

The router MUST implement the Destination/Source Routing mechanism described in [I-D.ietf-rtgwg-dst-src-routing]. This implies not simply routing "to a destination", but routing "to that destination AND from a specified source". The obvious application is egress routing, as required for a multihomed entity with a provider-allocated prefix from each of several upstream networks. Traffic within the network could be source/destination routed as well, or could be implicitly or explicitly routed from "any prefix", ::/0. Other use cases are described in [I-D.baker-rtgwg-src-dst-routing-use-cases]. If a FIB contains a route to a given destination from one or more prefixes not including ::/0, and a given packet destined there that has a source address that is in none of them, the packet in effect has no route, just as if the destination itself were not in the route table.

1.1. Requirements Language

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].

2. Theory of Routing

Both IS-IS and OSPF perform their calculations by building a lattice of routers and links from the router performing the calculation to each router, and then use routes (sequences in the lattice) to get to destinations that those routes advertise connectivity to. Following the SPF algorithm, calculation starts by selecting a starting point (typically the router doing the calculation), and successively adding {link, router} pairs until one has calculated a route to every router in the network. As each router is added, including the original router, destinations that it is directly connected to are turned into routes in the route table: "to get to 2001:db8::/32, route traffic to {interface, list of next hop routers}". For immediate neighbors to the originating router, of course, there is no next hop router; traffic is handled locally.

In this context, the route is qualified by a source prefix; It is installed into the FIB with the destination prefix, and the FIB applies the route if and only if the IPv6 source address also matches the advertised prefix. Of course, there may be multiple LSPs in the RIB with the same destination and differing source prefixes; these may also have the same or differing next hop lists. The intended forwarding action is to forward matching traffic to one of the next hop routers associated with this destination and source prefix, or to discard non-matching traffic as "destination unreachable".

TLVs that lack a source prefix sub-TLV match any source address (i.e., the source prefix TLV defaults to ::/0), by definition.

To ensure that routers without support for Destination/Source routing are excluded from path calculation for routes with a non-default source prefix, a separate MTID is used to carry Destination/Source routes. A router MUST NOT participate in a topology with such an MTID unless it implements Destination/Source routing.

There is a distinct Destination/Source Routing MTID for each of the underlying base MT topologies the information applies to. The set of routes propagated towards the forwarding plane is the union of the information in the base topology and the D/S Routing MTID. Incoming connectivity information with a default or non-present source prefix is advertised in the base topology, routes with non-default source prefix are advertised in the D/S Routing MTID.

2.1. Notation

For the purposes of this document, a route from the prefix A to the prefix B (in other words, whose source prefix is A and whose destination prefix is B) is expressed as A->B. A packet with the source address A and the destination address B is similarly described as A->B.

2.2. Dealing with ambiguity

In any routing protocol, there is the possibility of ambiguity. For example, one router might advertise a fairly general prefix - a default route, a discard prefix (which consumes all traffic that is not directed to an instantiated subnet), or simply an aggregated prefix while another router advertises a more specific one. In source/destination routing, potentially ambiguous cases include cases in which the link state database contains two routes A->B' and A'->B, in which A' is a more specific prefix within the prefix A and B' is a more specific prefix within the prefix B. Traditionally, we have dealt with ambiguous destination routes using a "longest match first" rule. If the same datagram matches more than one destination prefix advertised within an area, we follow the route with the longest matching prefix.

With source/destination routes, as noted in [I-D.baker-rtgwg-src-dst-routing-use-cases], we follow a similar but slightly different rule; the FIB lookup MUST yield the route with the longest matching destination prefix that also matches the source prefix constraint. In the event of a tie on the destination prefix, it MUST also match the longest matching source prefix among those options.

An example of the issue is this. Suppose we have two routes:

1. 2001:db8:1::/48 -> 2001:db8:3:3::/64

2. 2001:db8:2::/48 -> 2001:db8:3::/48

and a packet

   2001:db8:2::1 -> 2001:db8:3:3::1

If we require the algorithm to follow the longest destination match without regard to the source, the destination address matches 2001:db8:3:3::/64 (the first route), and the source address doesn't match the constraint of the first route; we therefore have no route. The FIB algorithm, in this example, must therefore match the second

route, even though it is not the longest destination match, because it also matches the source address.

2.3. Multi-topology Routing

As outlined in Section 2, this document specifies the use of separate topologies for Multi Topology Routing [RFC5120] to carry Destination/Source routing information. These topologies form pairs with a base topology each as follows:

base designated usage	base MTID	D/S MTID
default topology	0	TBD-MT0
IPv4 management	1	n/a
IPv6 default	2	TBD-MT2
IPv4 multicast	3	n/a
IPv6 multicast	4	n/a
IPv6 management	5	TBD-MT5

Figure 1: Destination/Source Routing MTIDs

The rationale for in-/excluding base MTIDs to provide a D/S MTID for is as follows:

MTID 0: The base (non-MTR) topology in some installations carries all routing information, including IPv6 reachabilities. In such a setup, the topology with MTID TBD-MT0 is used to carry associated D/S reachabilities.

MTIDs 1 and 3: Topologies with MTID 1 and 3 carry exclusively IPv4 reachabilities. Thus, no IPv6 D/S topology is created to associate with them.

MTID 2: The topology with MTID 2 carries IPv6 reachabilities in common M-ISIS setups. (MTID 0 in such cases carries exclusively IPv4 reachability information.) Associated IPv6 D/S reachabilities MUST be carried in MTID TBD-MT2.

MTID 4: MTID 4, while carrying IPv6 connectivity information, is used for multicast RPF lookups. Since Destination/Source routing is not compatible with multicast RPF lookups, no associated D/S MTID is defined for IS-IS.

MTID 5: An alternate management/administration topology may carry its routing information in MTID 5. Destination/Source routing is applicable to this and MUST use MTID TBD-MT5 to carry associated reachability TLVs.

Note that the different topology ID is the sole and only mechanism of both capability detection and backwards compatibility. D/S routing will operate correctly if D/S routing information is put in the same topology as non-D/S information, but adding an IS that does not support D/S routing will then -undetectably- lead to incorrect routing decisions, possibly including loops.

Therefore, all routers participating in D/S routing MUST implement M-ISIS and participate in the appropriate D/S topology per the table above. Conversely, routers not supporting D/S routing (or not configured to participate) MUST NOT participate in these topologies. Even installations that previously used only MTID 0 (i.e. no M-ISIS) would need to start using M-ISIS on all D/S routers. This results in correct operation in the face of partial deployment of D/S routing.

Note it is implied by the separate topology that there is a separate SPF calculation for that topology - using only the participants of that topology - and D/S routes use paths according to the result from that calculation. This is an aspect of Multi-topology operation itself, not this document.

Routers MUST NOT advertise non-D/S routing information using a D/S-Routing MTID. This includes both reachability information with a source prefix TLV with value ::/0, as well as without a source prefix sub-TLV. On receipt, routers MUST ignore any reachability information in a D/S-Routing MTID that does not have non-default source prefix information.

To limit complexity, each IPv6 Reachability TLV in a D/S-Routing MTID MUST have exactly one Source Prefix sub-TLV. Routers MUST NOT advertise TLVs with more than one Source Prefix sub-TLV, and MUST ignore any received TLV with more than one Source Prefix sub-TLV.

Systems that use topology IDs different than the values reserved by IANA should apply the considerations from this section analogously.

2.4. Migration and partial deployments

The Multi-topology mechanism described in the previous section introduces a distinct, independently operating topology that covers D/S routers. This easily allows partial and incremental deployments.

Such deployments then contain one or more D/S "subdomains" of neighboring routers that have D/S routing capability. Since shortest paths for D/S routes are calculated using a separate topology, traffic routed on D/S routes will be forwarded inside such a subdomain until it reaches the router originating the reachability.

Routers unaware or not participating in D/S routing will in such a case forward traffic according to only non-D/S routes. This can produce 2 distinct outcomes:

1. Traffic traverses a D/S router, where a more specific D/S route matches (and SPF in the D/S topology has found a valid path). It is then kept inside the D/S subdomain, reaching an originator of the D/S route.

2. Traffic reaches a system originating a non-D/S route or is considered unroutable even without regard to D/S routes.

Since the latter case provides no guarantee that there is no D/S route in the routing domain that could have matched, operators must pay careful attention to where non-D/S reachabilities are originated when more specific D/S routes are covered by them.

A very simple configuration that guarantees correct operation is to ensure covering destination-only reachabilities for D/S routes are originated by D/S routers themselves, and only by them. This results in traffic entering the D/S subdomain and D/S routes applying.

Lastly, in partial deployments, disconnected D/S subdomains may exist. Routers in such a subdomain cannot calculate a path for reachabilities in a subdomain they're not in. In this case a router MAY discard packets matching a D/S reachability for which it was unable to calculate a valid path. Alternatively, it MAY behave as if the D/S reachability didn't exist to begin with, i.e. routing the packet using the next less specific route (which could be D/S or non-D/S). It MUST NOT keep stale SPF calculation results that have become invalid as result of the topology partition.

This can be remediated by the operator adding connectivity between the subdomains, for example using some tunneling interface. The new link is then used to form an IS-IS adjacency fusing the previously split subdomains. The link will then be used to forward D/S traffic, possibly incurring some tunnel encapsulation overhead. To the IS-IS implementation, this link is no different from other links.

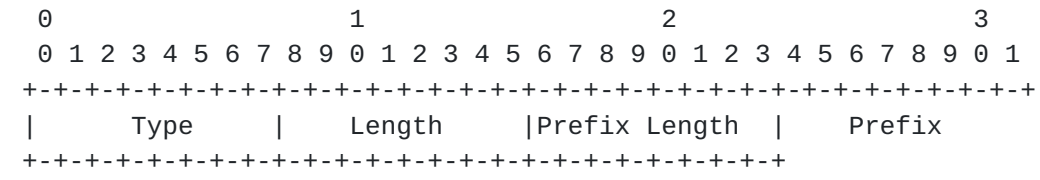
3. Protocol encoding for IPv6 Source Prefix information

Destination/Source reachabilities are originated using TLV 237, using an additional sub-TLV to carry the source prefix as follows.

As noted in Section 2, any IPv6 Reachability TLV that does not specify a source prefix is functionally identical to specifying ::/0 as the source prefix. Such routes SHOULD NOT be originated into the D/S MTID, but rather into the base MTID.

3.1. Source Prefix sub-TLV

The following Sub-TLV is defined for TLV 237:



Source Prefix Sub-TLV

Source Prefix Type: TBD-TLV (assigned by IANA)

TLV Length: Length of the sub-TLV in octets

Prefix Length: Length of the prefix in bits

Prefix: (source prefix length+7)/8 octets of prefix

This Sub-TLV MUST occur exactly once in all reachability originated in any of the D/S topologies listed in Figure 1. A reachability in these topologies with the Sub-TLV either missing or present more than once MUST be ignored in its entirety.

4. IANA Considerations

IANA is requested to allocate Values from the "IS-IS Multi-Topology ID Values" registry as follows:

TBD-MT0: IPv6 Dest/Source routing corresponding to topology 0

TBD-MT2: Reserved for IPv6 Dest/Source routing corresponding to topology 2

TBD-MT5: Reserved for IPv6 Dest/Source routing corresponding to topology 5

Additionally, IANA is requested to allocate an IS-IS codepoint from the "Sub-TLVs for TLVs 135, 235, 236, and 237" registry:

Type: TBD-TLV

Description: IPv6 SADR Source Prefix

Applicable to TLV 237: Yes

Applicable to TLVs 135, 235, 236: No

5. Security Considerations

The same injection and resource exhaustion attack scenarios as with all routing protocols apply.

Security considerations from [I-D.ietf-rtgwg-dst-src-routing] are particularly relevant to this document, in particular the possibility to inject (more) specific routes to hijack traffic.

6. Privacy Considerations

No privacy considerations apply to this document, as it only specifies routing control plane information.

7. Acknowledgements

Thanks to Les Ginsberg, Chris Hopps, Acee Lindem, Chris Bowers and Tony Przygienda for valuable feedback on this document. (TODO: incomplete, and sort by name.)

8. References

8.1. Normative References

[I-D.ietf-rtgwg-dst-src-routing]
Lamparter, D. and A. Smirnov, "Destination/Source Routing", draft-ietf-rtgwg-dst-src-routing-06 (work in progress), October 2017.

[IS-IS] ISO/IEC, "Intermediate System to Intermediate System Intra-Domain Routing Exchange Protocol for use in Conjunction with the Protocol for Providing the Connectionless-mode Network Service (ISO 8473)", ISO/IEC 10589:2002, Second Edition, 2002.

[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>.

[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>.

[RFC5308] Hopps, C., "Routing IPv6 with IS-IS", RFC 5308, DOI 10.17487/RFC5308, October 2008, <https://www.rfc-editor.org/info/rfc5308>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

8.2. Informative References

[I-D.baker-rtgwg-src-dst-routing-use-cases] Baker, F., Xu, M., Yang, S., and J. Wu, "Requirements and Use Cases for Source/Destination Routing", draft-baker-rtgwg-src-dst-routing-use-cases-02 (work in progress), April 2016.

Appendix A. Correctness considerations

While Multi-Topology routing in general can be assumed to work correctly when used on its own, this may not apply to a scenario mixing route calculation results as suggested in this document. However, this specific application is easily understandable as correct:

Systems that do not implement D/S routing will not participate in the D/S topology. They will calculate SPF in the base topology. Packets routed by such system will either (a) cross only non-D/S routers and reach the last hop as intended, or (b) cross a D/S router at some point.

For case (b), the D/S router may (b1) or may not (b2) have a more specific D/S route with a valid path. In case (b2), packets will be routed based on the same decisions that a non-D/S system would apply, so they will reach their last hop without any differences.

For case (b1), a break in forwarding behaviour happens for packets as they hit the first D/S-capable router, possibly after traversing some non-D/S systems. That router will apply D/S routing - which, since the path calculation is performed in the D/S topology, means that the packet is from there on routed on a path that only contains D/S capable systems. It will thus reach the D/S last hop as intended.

Packets starting out on a D/S-capable router fall into cases (b1) or (b2) as if a non-D/S router routed them first.

For both cases (b1) and (b2), a situation where a D/S router is aware (by flooding) of a more specific D/S route, but can't calculate a valid path (because the MT topology is not contiguous), this is for correctness concerns identical to the D/S route not existing to begin with. Note below on the correctness of this.

The compatibility mechanics thus rest on 2 pillars:

D/S routes will match as more specific if applicable

Packets will transit into D/S routing but not out of it

Note that the latter assumption holds true even if D/S routers fall back to non-D/S paths if they cannot calculate a shortest path towards the advertising system (either because SPF reaches the maximum path metric, or because there are multiple discontiguous D/S subdomains). This is because if a router A receives a packet routed on a D/S path, this implies the previous router B was able to successfully calculate SPF, via A, and that A has a path towards the originating system with a lower path metric than B. Conversely, if router A is unable to find a valid path, it is safe to assume router B was unable to do so either, and B forwarded the packet on a path calculated on non-D/S information.

Lastly, in terms of application use cases, it is also worth pointing out that loops will always result if (for example on a boundary to an upstream) the prefix routed incoming to the IS-IS domain is not fully covered by routes. Just as in non-D/S routing, this may cause a less specific (default) route to apply and loop packets back onto the same upstream. With D/S routing, this can now also occur if the incoming prefix is not covered for all sources. The solution remains the same: making sure the entire prefix is covered (for all sources), usually with a discard route. This is not an IS-IS consideration.

Appendix B. Change Log

(to be removed)

Initial Version: February 2013

updated Version: August 2013

Added MTR: August 2014

Split into 4 drafts: October 2014

Dropped 'Critical Sub-TLV' drafts June 2015

MT clarifications October 2015

Authors' Addresses

Fred Baker
Santa Barbara, California 93117
USA

Email: FredBaker.IETF@gmail.com

David Lamparter
NetDEF
Leipzig 04229
Germany

Email: david@opensourcerouting.org