Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: October 19, 2016

F.J. Baker
Cisco Systems
D. Lamparter
NetDEF

April 17, 2016

# IPv6 Source/Destination Routing using IS-IS
## draft-baker-ipv6-isis-dst-src-routing-05

## 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



## Copyright Notice



## 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] routing environments. To this extent, a new sub-TLV for an IPv6 [RFC2460] 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 |

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.

As this compatibility mechanism is not considered optional, M-ISIS MUST therefore be implemented for supporting the protocol outlined in this document. Even installations that previously used only MTID 0 (i.e. no M-ISIS) would need to start using MTID TBD-MT0.

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

## 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:

```
 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     |     Length    |Prefix Length  |    Prefix
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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

## 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 and Acee Lindem for valuable feedback on this document.  (TODO: incomplete.)

## 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-01 (work in progress), March 2016.

[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, March 1997.

[RFC2460] Deering, S.E. and R.M. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", RFC 5120, February 2008.

[RFC5308] Hopps, C., "Routing IPv6 with IS-IS", RFC 5308, October 2008.

### 8.2. Informative References

[I-D.baker-ipv6-isis-dst-flowlabel-routing]
Baker, F., "Using IS-IS with Token-based Access Control", draft-baker-ipv6-isis-dst-flowlabel-routing-01 (work in progress), August 2013.

[I-D.baker-rtgwg-src-dst-routing-use-cases]
Baker, F., "Requirements and Use Cases for Source/ Destination Routing", draft-baker-rtgwg-src-dst-routing-use-cases-01 (work in progress), October 2014.

## 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. 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.

If, for case (b1), the system knows of the existence of a more specific D/S route, but cannot calculate a valid path, it may either apply non-D/S routing (i.e. not install any route) or discard the packet (i.e. install a discard route). The next hop will either be a non-D/S system, or a D/S system with the same link-state information (and thus again unable to calculate a valid path -- or, more specifically, won't calculate a path that includes the previous router).

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

## 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
Cisco Systems
Santa Barbara, California  93117
USA

Email: fred@cisco.com

David Lamparter
NetDEF
Leipzig  04229
Germany

Email: david@opensourcerouting.org