rtgwg
Internet-Draft
Intended status: Standards Track
Expires: January 20, 2018

D. Lamparter
NetDEF
A. Smirnov
Cisco Systems, Inc.

July 19, 2017

# Destination/Source Routing
# draft-ietf-rtgwg-dst-src-routing-05

## Abstract

This note specifies using packets' source addresses in route lookups as additional qualifier to be used in route lookup. This applies to IPv6 [RFC2460] in general with specific considerations for routing protocol left for separate documents.

## 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 http://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 20, 2018.

## Copyright Notice

Table of Contents

1. Introduction . . . . . 2
1.1. Requirements Language . . . . . 3
2. Use cases . . . . . 3
2.1. Dual-connected home / SOHO network . . . . . 3
2.2. Degree of traffic engineering . . . . . 4
2.3. Distributed filtering based on source address . . . . . 5
2.4. Walled-garden Enterprise services . . . . . 5
2.5. Information Source for Neighbor Management . . . . . 6
3. Principle of operation . . . . . 6
3.1. Lookup ordering and disambiguation . . . . . 6
3.2. Ordering Rationale . . . . . 7
4. Routing protocol considerations . . . . . 7
4.1. Source information . . . . . 8
4.2. Loop-freeness considerations . . . . . 8
4.3. Recursive routing . . . . . 9
5. Applicability To Specific Situations . . . . . 9
5.1. Recursive Route Lookups . . . . . 10
5.1.1. Recursive route expansion . . . . . 11
5.2. Unicast Reverse Path Filtering . . . . . 11
5.3. Multicast Reverse Path Forwarding . . . . . 12
5.4. Testing for Connectivity Availability . . . . . 12
6. Interoperability . . . . . 12
6.1. Interoperability in Distance-Vector Protocols . . . . . 13
6.2. Interoperability in Link-State Protocols . . . . . 14
7. IANA Considerations . . . . . 15
8. Security Considerations . . . . . 15
9. Privacy Considerations . . . . . 15
10. Acknowledgements . . . . . 15
11. Change Log . . . . . 15
12. References . . . . . 16
12.1. Normative References . . . . . 16
12.2. Informative References . . . . . 16
Appendix A. Implementation Options . . . . . 17
A.1. Pre-expanded 2-step lookup without backtracking . . . . . 18
A.2. Translation to Multi-FIB (Policy Routing) perspective . . 18
Authors' Addresses . . . . . 20

## 1. Introduction

Both IPv4 [RFC0791] and IPv6 [RFC2460] architectures specify that determination of the outgoing interface and next-hop gateway for packet forwarding is based solely on the destination address contained in the packet header. There exists class of network design problems which require packet forwarding to consider more than just the destination IP address (see Section 2 for examples). At present these problems are routinely resolved by configuring on routers

special forwarding based on a local policy. The policy enforces packet forwarding decision outcome based not only on the destination address but also on other fields in the packet's IP header, most notably the source address. Such policy-based routing is conceptually similar to static routes in that it is highly static in nature and must be closely governed via the management plane (most frequently - via managing configuration by an operator). Thus policy-based routing configuration and maintenance is costly and error-prone.

Rapid expansion of IPv6 to networks were static configuration is not acceptable due to both its static nature and necessity of frequent intervention by a skilled operator requires change in the paradigm of forwarding IP packets based only on their destination address.

This document describes architecture of source-destination routing. This includes description of making a packet forwarding decision and requirements to dynamic routing protocols which will disseminate source-destination routing information. Specific considerations for particular dynamic routing protocols are outside of the scope of this note and will be covered in separate documents.

General concepts covered by this document are equally applicable to both IPv4 and IPv6. Considering limited backward compatibility of the source-destination routing with the traditional destination-only routing, it appears likely that at this stage of IPv4 deployment change of routing paradigm in existing networks is not feasible (see Section 6 for discussion of backwards compatibility). So examples in this document will be given using IPv6 addresses.

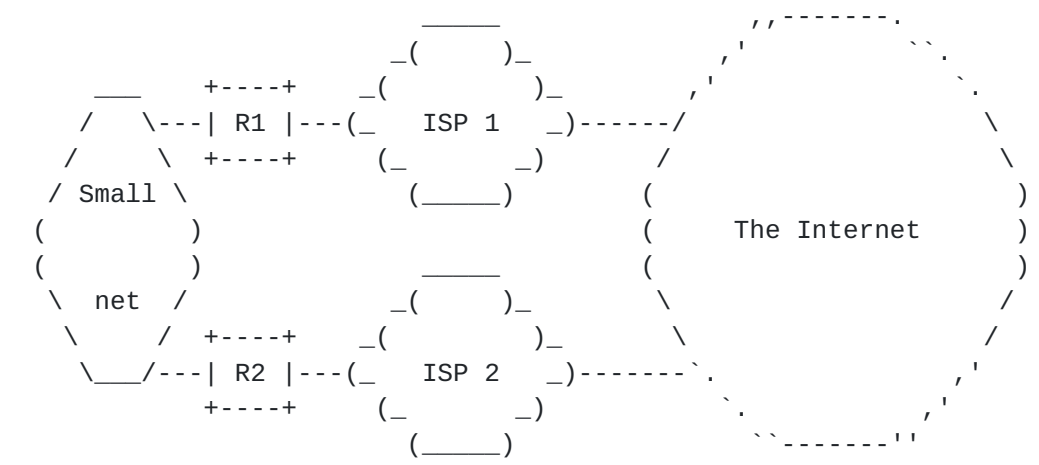
### 1.1. Requirements Language

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. Use cases

### 2.1. Dual-connected home / SOHO network

Small networks - such as SOHO or the home networks (homenet) - may be multihomed (i.e. dual-connected) to two different Internet Service Providers (ISPs). Benefits of doing this may include resiliency or faster access to important resources (for example, video or cloud services) local to ISPs.




Example of multihomed small network

Each ISP will allocate to the network IP address (or small range of IP addresses) to use as source address for Internet communications.

Since connectivity providers generally secure their ingress along the lines of BCP 38 [RFC2827], small multihomed networks have a need to ensure their traffic leaves their network with a correct combination of source address and exit taken. This applies to networks of a particular pattern where the provider's default (dynamic) address provisioning methods are used and no fixed IP space is allocated, e.g. home networks, small business users and mobile ad-hoc setups.

While IPv4 networks would conventionally use NAT or policy routing to produce correct behaviour, this not desirable to carry over to IPv6. Instead, assigning addresses from multiple prefixes in parallel shifts the choice of uplink to the host. However, now for finding the proper exit the source address of packets must be taken into account.

Source-destination routing, when enabled on routers in the multihomed small network (including routers R1 and R2), solves the problem by driving packets originated by internal hosts to the correct Internet exit point considering IP source address assigned to the packet by originating host.

For a general introduction and aspects of interfacing routers to hosts, refer to [RFC8043].

## 2.2. Degree of traffic engineering

Consider enterprise consisting of a headquarter (HQ) and branch offices. A branch office is connected to the enterprise HQ network via 2 links. For performance or security reasons it is desired to

route corporate traffic via one link and Internet traffic via another link. In direction branch -> HQ the problem is easily solvable by having the default route pointing to the Internet link and HQ routes pointing to another link. But destination routing does not provide an easy way to achieve traffic separation in direction HQ -> branch because destination is the same (branch network).

Source-destination routing provides an easy way to sort traffic going to the branch based on its source address.

### 2.3. Distributed filtering based on source address

A network has untrusted zone and secure one (and both zones comprise many links and routers). Computers from the secure zone need to be able to communicate with some selected hosts in the untrusted zone. The secure zone is protected by a firewall. The firewall is configured to check that packets arriving from the untrusted zone have destination address in the range of secure zone and source address of trusted hosts in the untrusted zone. This works but leaves the firewall open to DDOS attack from outside.

If routers in the untrusted zone are configured with source-destination routing (and, possibly, unicast RPF check) and receive via dynamic routing protocol routes <destination: secure zone; source: trusted host in the untrusted zone> then DDOS attack is dropped by routers on the edge of source-destination routing area. DDOS attack does not even reach the firewall whose resources are freed to deal with Deep Packet Inspection. On the other hand, security policy is managed in a single point - on a router injecting relevant source-destination routes into the dynamic routing protocol.

### 2.4. Walled-garden Enterprise services

Apart from transfering from multihomed personal networks to multihomed PA enterprise setups without any changes, source-destination routing can also be used to correctly route services that assign their own prefixes to customers using the particular service. This is distinct from internet connectivity only in that it does not provide a default route. Applying source-destination routing, the entire routing domain is aware of the specific constraints of the routes involved.

Additionally, if the walled-garden's destination prefix is advertised as blackhole route, this ensures that communication with the service will only be routed using the specific D/S route, never leaking onto unintended paths like a default route.

This is very similar to firewall/filtering functionality, except the feature is distributed onto routers.

### 2.5. Information Source for Neighbor Management

Having information on source address restrictions for routes distributed, routers can rely on this additional information to improve their behaviour towards hosts connected to them. This specifically includes IPv6 Router Advertisements, which is described in [I-D.linkova-v6ops-conditional-ras].

## 3. Principle of operation

The mechanism in this document is such that a source prefix is added to all route entries. This document assumes all entries have a source prefix, with ::/0 as default value for entries installed without a specified source prefix. This need not be implemented in this particular way, however the system MUST behave exactly as if it were. In particular, a difference in behaviour between routes with a source prefix of ::/0 and routes without source prefix MUST NOT be visible.

For uniqueness considerations, the source prefix factors MUST be taken into account for comparisons. Two routes with identical information except the source prefix MAY exist and MUST be installed and matched.

### 3.1. Lookup ordering and disambiguation

When a router is making packet forwarding decision, that is consulting its routing table in order to determine outgoing interface and next-hop to forward the packet to, it will use information from packet's header to look up best matching route from the routing table. This section describes lookup into the source-destination routing table.

For longest-match lookups, the source prefix is matched after the destination prefix. This is to say, first the longest matching destination prefix is found, then the table is searched for the route with the longest source prefix match, while only considering routes with exactly the destination prefix previously found. If and only if no such route exists (because none of the source prefixes match), the lookup moves to the next less specific destination prefix.

A router MUST continue to a less specific destination prefix if no route matches on the source prefix. It MUST NOT terminate lookup on such an event.

Using A < B to mean "A is more specific than B", this is represented as:

```
A < B :=    Adst <  Bdst
        || (Adst == Bdst && Asrc < Bsrc)
```

Implementations MAY implement lookup algorithm differently from step-by-step description given above but if they do so then outcome of the algorithm MUST be exactly the same as if above steps were used. A variation providing improved performance, as well as a variation matching existing implementations with reversed order are described in Appendix A.1 and Appendix A.2, respectively.

## 3.2. Ordering Rationale

Ordering of searching for address match is important and reversing it would lead to semantically different behavior. This standard requires most specific match on destination address to be found before looking for match on source address.

Choosing destination to be evaluated first caters to the assumption that local networks should have full, contiguous connectivity to each other. This implies that those specific local routes always match first based on destination, and use a zero ("all sources") source prefix.

If the source prefix were to be matched first, this would result in a less specific (e.g. default) route with a source prefix to match before those local routes. In other terms, this would essentially divide local connectivity into zones based on source prefix, which is not the intention of this document.

Hence, this document describes destination-first match search.

# 4. Routing protocol considerations

As with the destination-only routing, source-destination routes will typically be disseminated throughout the network by dynamic routing protocols. It is expected that multiple dynamic routing protocols will be adapted to the needs of source-destination routing architecture. Specification of dynamic routing protocols is outside of scope of this document. This section lists requirements and considerations for the dynamic source-destination routing protocols.

### 4.1. Source information

Dynamic routing protocols will need to be able to propagate source range information together with destination prefix and other accompanying routing information. Source range information may be propagated with all destination prefixes or only some of them. Destination prefixes advertised without associated source range MUST be treated as having default source range ::/0.

Dynamic routing protocols MUST be able to propagate multiple routes whose destination prefix is the same but associated source ranges are different. Such unique pairs of source and destination MUST be treated as different source-destination routes.

There is no limitation on how source range information is propagated and associated with destination prefixes. Individual protocols may choose to propagate source range together with a destination prefix in the form of prefix, in the form of index to list of known source ranges or in any other form allowing receiver to reconstruct pair of destination prefix and associated source range.

### 4.2. Loop-freeness considerations

It is expected that some existing dynamic routing protocols will be enhanced to propagate source-destination routing information. In this case the protocol may be configured to operate in a network where some, but not all, routers support source-destination routing and others are still using destination-only routing. Even if all routers within a network are capable of source-destination routing, it is very likely that on edges of the network they will have to forward packets to routers doing destination-only routing.

Since a router implementing source-destination routing can have additional, more granular routes than one that doesn't implement it, persistent loops can form between these systems.

Thus specifications of source-destination routing protocols (either newly defined protocols or enhancements to already existing one) MUST take provisions to guarantee loop-free operations.

There are 3 possible approaches to avoid looping condition:

1. Guarantee that next-hop gateway of a source-destination route supports source-destination routing, for example calculate an alternate topology including only routers that support source-destination routing architecture

2. If next-hop gateway is not aware of source-destination routing then a source-destination path can lead to it only if next-hop router is 'closer' to the destination in terms of protocol's routing metric; important particular case of the rule is if destination-only routing is pointing to the same next-hop gateway

3. Discard the packet (i.e. treat source-destination route as unreachable)

In many practical cases routing information on the edges of source-destination routing domain will be provided by an operator via configuration. Dynamic routing protocol will only disseminate this trusted external routing information. For example, returning to the use case of multihomed Home network (Section 2.1), both routers R1 and R2 will have default static routes pointing to ISPs.

Above considerations require a knowledge of the next-hop router's capabilities. For routing protocols based on hop-by-hop flooding (RIP [RFC2080], BGP [RFC4271]), knowing the peer's capabilities is sufficient. Information about if peer supports source-destination routing can either be negotiated explicitly or simply be deduced from the fact that systems would propagate source-destination routing information only if they understand it. Protocols building a link-state database (OSPFv3 [RFC5340], IS-IS [RFC5308]) have the additional opportunity to calculate alternate paths based on knowledge of the entire domain but cannot assume that routers understand source-destination routing information only because they participated in its flooding. Such protocols MUST explicitly advertise support for the source-destination routing.

### 4.3. Recursive routing

Dynamic routing protocols may propagate routing information in a recursive way. Examples of such recursion is forwarding address in OSPFv3 [RFC5340] AS-External-LSAs and NEXT_HOP attribute in BGP [RFC4271] NLRI.

Dynamic routing protocol supporting recursive routes MUST specify how this recursive routing information is interpreted in the context of source-destination routing as part of standardizing source-destination routing extensions for the protocol. Section 5.1 lists several possible strategies protocols can choose from.

## 5. Applicability To Specific Situations

This section discusses how source-destination routing is used together with some common networking techniques dependent on routes in the routing table.

### 5.1. Recursive Route Lookups

Recursive routes provide indirect path information where instead of supplying outgoing interface and next-hop gateway directly they specify that next-hop information must be taken from another route in the same routing table. It is said that one route 'recurses' via another route which is 'resolving' recursion. Recursive routes may either be carried by dynamic routing protocols or provided via configuration as recursive static routes.

Recursive source-destination routes have additional complication in how source address range should be considered while finding source-destination route to resolve recusion.

There are several possible approaches:

1. Ignore source-destination routes, resolve recursion only via destination-only routes (i.e. routes with source range ::/0)

2. Require that both the recursive and resolving routes have the same source range associated with them; this requirement may be too restrictive to be useful in many cases

3. Require that source range associated with recursive route is a subset of source range associated with route resolving recursion (i.e. source range of the resolving route is less specific superset of recursive route's source range)

4. Create multiple instances of the route whose nexthop is being resolved with different source prefixes; this option is further elaborated in Section 5.1.1

When recursive routing information is propagated in a dynamic routing protocol, it is up to the protocol specification to select and standardize appropriate scheme of recusrsive resolution.

Recursive resolution of configured static routes is local to router where recursive static routes were configured, thus behavior is implementation's choice. Implementations SHOULD provide option (3) from the above list as their default method of recursive static route resolution. This is both to guarantee that destination-only recursive static routes do not change their behavior when router's software is upgraded to support source-destination routing and at the same time make source-destination recursive routes useful.

### 5.1.1. Recursive route expansion

When doing recursive nexthop resolution, the route that is being resolved is installed in potentially multiple copies, inheriting all possible more-specific routes that match the nexthop as destination. The algorithm to do this is:

1. form the set of attributes for lookup by using the (unresolved, recursive) nexthop as destination (with full host prefix length, i.e. /128), copy all other attributes from the original route

2. find all routes that overlap with this set of attributes (including both more-specific and less-specific routes)

3. order the result from most to less specific

4. for each route, install a route using the original route's destination and the "logical and" overlap of each extra match attribute with same attribute from the set. Copy nexthop data from the route under iteration. Then, reduce the set of extra attributes by what was covered by the route just installed ("logical AND NOT").

Example recursive route resolution

```
route to be resolved:
2001:db8:1234::/48, source 2001:db8:3456::/48,
                    recursive nexthop via 2001:db8:abcd::1

routes considered for recursive nexthop:
::/0,                                              via fe80::1
2001:db8:abcd::/48,                                via fe80::2
2001:db8:abcd::/48,   source 2001:db8:3456:3::/64, via fe80::3
2001:db8:abcd::1/128, source 2001:db8:3456:4::/64, via fe80::4

recursive resolution result:
2001:db8:1234::/48,   source 2001:db8:3456::/48,   via fe80::2
2001:db8:1234::/48,   source 2001:db8:3456:3::/64, via fe80::3
2001:db8:1234::/48,   source 2001:db8:3456:4::/64, via fe80::4
```

## 5.2. Unicast Reverse Path Filtering

Unicast reverse path filtering MUST use dst-src routes analog to its usage of destination-only routes. However, the system MAY match either only incoming source against routes' destinations, or it MAY match source and destination against routes' destination and source. It MUST NOT ignore dst-src routes on uRPF checks.

### 5.3. Multicast Reverse Path Forwarding

Multicast Reverse Path Lookups are used to find paths towards the (known) sender of multicast packets. Since the destination of these packets is the multicast group, it cannot be matched against the source part of a dst-src route. Therefore, dst-src routes MUST be ignored for Multicast RPF lookups.

### 5.4. Testing for Connectivity Availability

There are situations where systems' behaviour depends on the fact whether "connectivity" is available in a broad sense. These systems may have previously tested for the existence of a default route in the routing table.

Since the default route may now be qualified with a source prefix, this test can fail. If no additional information is available to qualify this test, systems SHOULD test for the existence of any default route instead, e.g. include routes with default destination but non-default source prefix.

However, if the test can be associated with a source address or source prefix, this data SHOULD be used in looking up a default route. Depending on the application, it MAY also be useful to - possibly additionally - consider "connectivity" to be available if any route exists where the route's source prefix covers the prefix or address under consideration, allowing arbitrary destination prefixes.

Note though that this approach to routing SHOULD NOT be used to infer a list of source prefixes in an enumerative manner, or even to guess domain information. Specifically, if an operator uses more specific source prefixes to refine their routing, the inferred information will provide bogus extraneous output. This is distinct from the connectivity tests mentioned above in that those actually inquire the routing system, unlike domain information or enumeration, which is higher-layer application information.

## 6. Interoperability

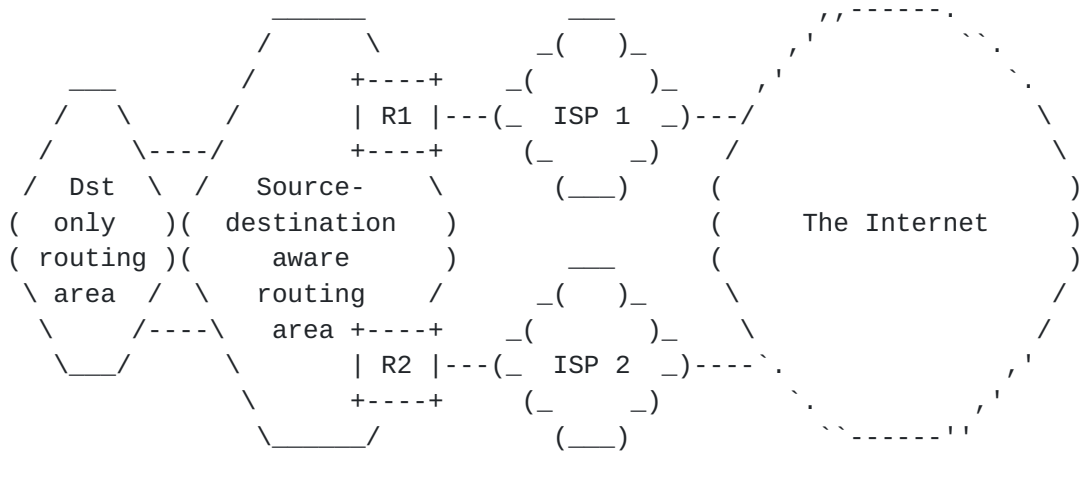
As pointed out in Section 4.2 traffic may permanently loop between routers forwarding packets based only on their destination IP address and routers using both source and destination addresses for forwarding decision.

In networks where the same dynamic routing protocol is being used to propagate routing information between both types of systems the protocol may address some or all traffic looping problems. Recommendations to protocol designers are discussed in Section 4.2.

When routing information is coming from outside of the routing protocol (for example, being provided by operator in the form of static routes or network protocols not aware of source-destination routing paradigm) it may not be possible for the router to ascertain loop-free properties of such routing information. In these cases consistent (and loop-free) packet forwarding is woven into network topology and must be taken into consideration at design time.

It is possible to design network with mixed deployment of routers supporting and not supporting source-destination routing. Thus gradual enablement of source-destination routing in existing networks is also possible but has to be carefully planned and evaluated for each network design individually.

Generally, source-destination routing will not cause traffic loops when disjoint 'islands' of source-destination routing do not exchange source-destination routing information. One particular case of this rule is a network which contains single contiguous 'island' of routers aware of source-destination routing. Example SOHO network from Section 2.1 which demonstrates this design approach:




Example of multihomed small network with partial deployment of source-destination routing

## 6.1. Interoperability in Distance-Vector Protocols

Distance-Vector routing protocols (BGP, RIPng, BABEL), operating on a hop-by-hop basis, can address interoperability and migration concerns on that level. With routing information being flooded in the reverse direction of traffic being forwarded using that information, a hop that floods is the same hop that forwards.

This makes dealing with destination/source-unaware routers easy if destination/source routes are made to be ignored by such unaware routers, and flooding of such routes is inhibited.

If D/S routes are discarded by non-D/S routers, D/S routers will not receive non-working routes and can select from other available working D/S routes.

Note that for this to work, non-D/S routers MUST NOT flood D/S routing information. This can be achieved in 2 ways:

1. Using some preexisting encoding to signal non-D/S routers to not flood these particular routes

2. Ignoring flooded D/S information on D/S routers by having them detect that they received it from a non-D/S router (e.g. using some capability signalling to identify non-D/S routers.) This handling likely needs to be performed on a level of same-link neighborships.

Also note that the considerations in this section only apply if data path and flooding path are congruent.

### 6.2. Interoperability in Link-State Protocols

For Link-State routing protocols (OSPF, IS-IS), there is no relation between route flooding and forwarding. Instead, forwarding decisions are based on shortest-path calculation on top of the received topology information.

For a D/S router to avoid loops, there are again two choices available:

1. Detect that forwarding for a D/S route transits over a non-D/S router and convert the route into a blackhole route to replace looping with blackholing. This obviously impacts connectivity.

2. Perform separate SPF calculations using only the subset of D/S-capable routers; thus D/S routers can forward D/S-routed packets as long as they stay in contiguous islands.

The latter approach is facilitated by Multi-Topology extensions to the respective protocols. These extensions provide a way to both isolate D/S routing information and perform the separate SPF calculation. Note that it is not neccessary to use multiple topologies for distinct source prefixes; only a single additional topology encompassing all D/S-capable routers is sufficient.

## 7. IANA Considerations

This document makes no requests to IANA.

## 8. Security Considerations

Systems operating under the principles of this document can have routes that are more specific than the previously most specific, i.e. host routes. This can be a security concern if an operator was relying on the impossibility of hijacking such a route.

While source/destination routing could be used as part of a security solution, it is not really intended for the purpose. The approach limits routing, in the sense that it routes traffic to an appropriate egress, or gives a way to prevent communication between systems not included in a source/destination route, and in that sense could be considered similar to an access list that is managed by and scales with routing.

## 9. Privacy Considerations

If a host's addresses are known, injecting a dst-src route allows isolation of traffic from that host, which may compromise privacy. However, this requires access to the routing system. As with similar problems with the destination only, defending against it is left to general mechanisms protecting the routing infrastructure.

## 10. Acknowledgements

The base underlying this document was first outlaid by Ole Troan and Lorenzo Colitti in [I-D.troan-homenet-sadr] for application in the homenet area. Significant contributions to source-specific routing as a whole came from Juliusz Chroboczek and Matthieu Boutier.

This document itself is largely the result of discussions with Fred Baker and derives from [I-D.baker-ipv6-isis-dst-src-routing].

Thanks to Chris Bowers, Acee Lindem and Tony Przygienda for their input and review.

The Linux kernel is providing an implementation of the behaviour described here since even before the document was started.

## 11. Change Log

May 2017 [-04]: no changes

November 2016 [-03]:

added DV/LS protocol considerations

note backtracking workaround/caveat

November 2015 [-02]:

added section on source-destination routing use cases

added section on alternative lookup algorithm

added section on requirement for dynamic routing protocols dessiminating source-destination informaton

October 2015 [-00]: renamed to draft-ietf-rtgwg-dst-src-routing-00, no content changes from draft-lamparter-rtgwg-dst-src-routing-01.

April 2015 [-01]: merged routing-extra-qualifiers draft, new ordering rationale section

October 2014 [-00]: Initial Version

## 12. References

### 12.1. 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, <http://www.rfc-editor.org/info/rfc2119>.

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, DOI 10.17487/RFC2460, December 1998, <http://www.rfc-editor.org/info/rfc2460>.

### 12.2. Informative References

[hal-00947234v1] Boutier, M. and J. Chroboczek, "Source-sensitive routing", hal 00947234v1, 2014, <https://hal-univ-diderot.archives-ouvertes.fr/hal-00947234v1>.

[I-D.baker-ipv6-isis-dst-src-routing] Baker, F. and D. Lamparter, "IPv6 Source/Destination Routing using IS-IS", draft-baker-ipv6-isis-dst-src-routing-07 (work in progress), July 2017.

[I-D.linkova-v6ops-conditional-ras]
Linkova, J. and s. stucchi-lists@glevia.com, "Using Conditional Router Advertisements for Enterprise Multihoming", draft-linkova-v6ops-conditional-ras-01 (work in progress), July 2017.

[I-D.troan-homenet-sadr]
Troan, O. and L. Colitti, "IPv6 Multihoming with Source Address Dependent Routing (SADR)", draft-troan-homenet-sadr-01 (work in progress), September 2013.

[RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, DOI 10.17487/RFC0791, September 1981, <http://www.rfc-editor.org/info/rfc791>.

[RFC2080] Malkin, G. and R. Minnear, "RIPng for IPv6", RFC 2080, DOI 10.17487/RFC2080, January 1997, <http://www.rfc-editor.org/info/rfc2080>.

[RFC2827] Ferguson, P. and D. Senie, "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing", BCP 38, RFC 2827, DOI 10.17487/RFC2827, May 2000, <http://www.rfc-editor.org/info/rfc2827>.

[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, <http://www.rfc-editor.org/info/rfc4271>.

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

[RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, DOI 10.17487/RFC5340, July 2008, <http://www.rfc-editor.org/info/rfc5340>.

[RFC8043] Sarikaya, B. and M. Boucadair, "Source-Address-Dependent Routing and Source Address Selection for IPv6 Hosts: Overview of the Problem Space", RFC 8043, DOI 10.17487/RFC8043, January 2017, <http://www.rfc-editor.org/info/rfc8043>.

## Appendix A. Implementation Options

### A.1. Pre-expanded 2-step lookup without backtracking

The backtracking behavior (specified in Section 3.1 as "A router MUST continue to a less specific destination prefix") has been shown to potentially cause a significant loss of forwarding performance since forwarding a single packet may require a large number of table lookups. (The degenerate case is 129 destination lookups in decreasing prefix length, each followed by a failing longest-match on the source prefix.)

To avoid this, implementations can install synthetic routes to achieve the same lookup result. This works as follows, to be evaluated for each unique destination prefix:

1. If there is a route (D, S=::/0), end processing for D.

2. Iterate upwards one level (from D if first iteration, previous D' otherwise) to a less specific destination. Call this D'.

3. For all routes (D', S'), i.e. all source prefixes S' under that destiation prefix, install a copy (D, S') if and only if S' covers some source prefix that isn't covered yet. (In terms of set theory, S' cut by all existing S under D is not empty.)

4. Repeat at step 1.

The effect of this algorithm is that after performing a lookup on the destination prefix, looking up the source prefix directly yields the result that backtracking would give. This eliminates backtracking and provides constant 2 lookup cost (after exactly one destination longest-match, the source longest-match will provide the final, correct result; any no-match is a final no-match).

### A.2. Translation to Multi-FIB (Policy Routing) perspective

The lookup procedure described in this document requires destination-first lookup. This is not a fit with most existing implementations of Policy Routing. While Policy Routing has no formal specification, it generally permits choosing from multiple routing tables / FIBs based on, among other things, source address. Some implementations support using more than one FIB for a single lookup, but not all do.

An implementation that can choose from multiple FIBs based on source address is capable of correct forwarding according to this document, provided that it supports enough FIBs. One FIB will be used for each unique source prefix.

For a complete description of the required translation algorithm, please refer to [hal-00947234v1]. It roughly works as follows:

After source-destination routing information has been collected, one FIB table is created for each source range including the default range ::/0. Source-destination routes then replicated into each destination-only FIB table whose associated source address range is a subset of route's source range. Note that this rule means routes with default source range ::/0 are replicated into each FIB table.

In case when multiple routes with the same destination prefix are replicated into the same FIB table only route with the most specific source address range is installed.

For example, if source-destination routing table contains these routes:

| Destination prefix | Source range | Next Hop |
|---|---|---|
| ::/0, | ::/0, | NH1 |
| 2001:101:1234::/48, | 2001:db8:3456:8000::/56, | NH2 |
| 2001:101:5678::/48, | 2001:db8:3456:8000::/56, | NH3 |
| | ::/0, | NH4 |
| 2001:101:abcd::/48, | 2001:db8:3456::/48, | NH5 |

then 3 FIB tables will be created associated with source ranges ::/0, 2001:db8:3456::/48 and 2001:db8:3456:8000::/56. In this example range 2001:db8:3456:8000::/56 is a subset of less specific range 2001:db8:3456::/48. Such inclusion makes a somewhat artificial example but was intentionally selected to demonstrate hierarchy of route replication.

And content of these FIB tables will be:

FIB 1 (source range ::/0):

| Destination prefix | Next Hop |
|---|---|
| ::/0, | NH1 |
| 2001:101:5678::/48, | NH4 |

FIB 2 (source range 2001:db8:3456::/48):

| Destination prefix | Next Hop |
|---|---|
| ::/0, | NH1 |
| 2001:101:5678::/48, | NH4 |
| 2001:101:abcd::/48, | NH5 |

FIB 3 (source range 2001:db8:3456:8000::/56):

| Destination prefix | Next Hop |
|---|---|
| ::/0, | NH1 |
| 2001:101:1234::/48, | NH2 |
| 2001:101:5678::/48, | NH3 |
| 2001:101:abcd::/48, | NH5 |

During packet forwarding, lookup first matches source address against the list of address ranges associated with FIB tables to select a FIB table with the most specific source address range and then does destination-only lookup in the selected FIB table.

## Authors' Addresses

David Lamparter
NetDEF
Leipzig 04103
Germany

Email: david@opensourcerouting.org

Anton Smirnov
Cisco Systems, Inc.
De Kleetlaan 6a
Diegem 1831
Belgium

Email: as@cisco.com