rtgwg
Internet-Draft
Intended status: Standards Track
Expires: December 29, 2015

D. Lamparter
NetDEF

June 27, 2015

Destination/Source Routing
draft-lamparter-rtgwg-dst-src-routing-01

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 December 29, 2015.

Table of Contents

1. Introduction

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.

For a general introduction and aspects of interfacing routers to hosts, refer to [I-D.sarikaya-6man-sadr-overview].

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. 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.

2.1. Lookup ordering and disambiguation

Adding further criteria to be looked up when forwarding packets on a hop-by-hop basis has the very fundamental requirement that all routers behave the same way in choosing the most specific route when there are multiple eligible routes.

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)
```

2.2. Ordering Rationale

The ordering described by this document (destination before source) could as well be reversed, which would lead to semantically different behavior.

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 lookup.

3. Applicability To Specific Situations

3.1. Recursive Route Lookups

TBD, multiple possible approaches:

variant 1: ignore dst-src routes, only use routes with src ::/0

variant 2: exact-match src prefixes from resolvee to resolvent (will not work for a lot of cases)

variant 3: longer-match src prefixes from resolvee to resolvent (nexthop src may be superset of looked-up route)

variant 4: create multiple instances of the route whose nexthop is resolved, with different source prefixes

(Variant 4:)

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
```

3.2. 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.

3.3. 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.

4. Interoperability

Since a router implementing source/destination routing can have additional, more specific routes than one that doesn't implement source/destination routing, persistent loops can form between these systems. To prevent this from happening, a simple rule must be followed:

The set of qualifiers used to route a particular packet MUST be a subset of the qualifiers supported by the next hop.

This means in particular that a router using the source address as extra qualifier MUST NOT route packets based on a source/destination route to a system that doesn't support source/destination routes (and hence doesn't understand the route).

There are 3 possible approaches to avoid such a condition:

1. discard the packet (treat as destination unreachable)

2. calculate an alternate topology including only routers that support qualifier A

3. if the lookup returns the same nexthop without using qualifier A, use that result (i.e., the nexthop is known to correctly route the packet)

Above considerations require under all circumstances a knowledge of the next router's capabilities. For routing protocols based on hop-by-hop flooding (RIP [RFC2080], BGP [RFC4271]), knowing the peer's capabilities - or simply relying on systems to only flood what they understand - is sufficient. Protocols building a link-state database (OSPF [RFC5340], IS-IS [RFC5308]) have the additional opportunity to calculate alternate paths based on knowledge of the entire domain, but cannot rely on routers flooding only link state they support themselves.

5. IANA Considerations

This document makes no requests to IANA.

6. 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.

7. 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.

8. 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.

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

9. Change Log

Initial Version: April 2015: merged routing-extra-qualifiers draft, new ordering rationale section

Initial Version: October 2014

10. References

10.1. Normative References

[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.

10.2. Informative References

[I-D.baker-ipv6-isis-dst-src-routing]
Baker, F., "IPv6 Source/Destination Routing using IS-IS", draft-baker-ipv6-isis-dst-src-routing-01 (work in progress), August 2013.

[I-D.sarikaya-6man-sadr-overview]

Sarikaya, B., "Overview of Source Address Dependent Routing", draft-sarikaya-6man-sadr-overview-01 (work in progress), September 2014.

[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.

[RFC2080] Malkin, G. and R. Minnear, "RIPng for IPv6", RFC 2080, January 1997.

[RFC2827] Ferguson, P. and D. Senie, "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing", BCP 38, RFC 2827, May 2000.

[RFC4271] Rekhter, Y., Li, T., and S. Hares, "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, January 2006.

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

[RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, July 2008.

Author's Address

David Lamparter
NetDEF
Leipzig 04103
Germany

Email: david@opensourcerouting.org