Network Working Group
Internet-Draft
Intended status: Informational
Expires: October 29, 2016

F. Baker
Cisco Systems
M. Xu
S. Yang
J. Wu
Tsinghua University
April 27, 2016

Requirements and Use Cases for Source/Destination Routing
draft-baker-rtgwg-src-dst-routing-use-cases-02

Abstract

This note attempts to capture important use cases for source/ destination routing.

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 October 29, 2016.

Table of Contents

1. Introduction

Source/Destination routing has been proposed in the IPv6 community and specifically in homenet as a means of dealing with multihomed networks whose upstream networks give them provider-allocated addresses. An initial approach was suggested in [RFC3704], which assumed that a packet following a default route to an egress CPE Router might arrive at the wrong one, and need to be redirected to the right CPE Router. Subsequent approaches, including those listed in the bibliography, have focused on using routing protocols or routing procedures with extensions that make decisions based on both the source and the destination address.

"Source/Destination Routing" is defined as routing in which both the source and the destination address must be considered in selecting the next hop. It might be thought of as routing "to a destination with a constraint" - a router might have multiple routes to a given destination, and follow the one that also obeys the constraint, or it might have only one route to a destination but correctly fail to forward a packet that doesn't meet the constraint. From that perspective, the logic here extends to other cases in which a constraint might be placed on the route. As with all routing, a primary requirement is to follow the longest-match-first rule to the destination; following a less specific route may well take traffic to the wrong place.

As a side note, source address spoofing in this case will be limited to addresses from the indicated source prefixes, obviating the need for upstream ingress filtering. Ingress filtering within the domain in LAN switches can prevent spoofing of addresses within those prefixes.

This note attempts to capture common use cases. These will be in terms of a general statement of intent coupled with a specific example of the intent for clarity. The use cases are obviously not limited to these, but these should be a reasonably complete set.

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

The use cases proposed here are not an exhaustive set, but are representative of a set of possibilities. At least three are presently-deployed use cases; the fourth is a possible use case within an edge network.

2.1. Simple Egress Routing

One use case is as shown in Figure 1. A customer network has two or more upstream networks, and a single CPE Router. Each upstream network allocates a prefix for use in the customer network, and the customer network configures a subnet from each of those ISP prefixes on each of its LANs. The CPE Router advertises default routes into the network that are "from" each PA prefix. Apart from prefix itself, the services of the upstream ISPs are indistinguishable; they each get the customer to the Internet.

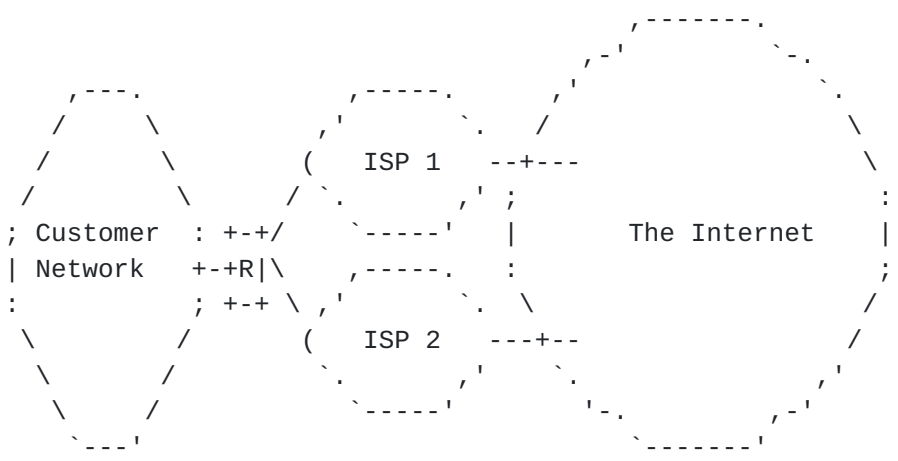

Figure 1: Egress Routing in a Multihomed Environment with One CPE Router

The big issue in this network is, of course, ingress filtering [RFC2827] by the upstream ISP. If packets intended for a remote destination pass through the wrong ISP, they will be blocked. In the ideal case, traffic following default route gets to the upstream network indicated by its source address.

The CPE Router could, at least in concept, advertise a single default route into the network, as all traffic to an upstream ISP must pass through that CPE Router. However, should another CPE Router be added later, it would have to change its behavior to accommodate that CPE Router (as in Section 2.2). Hence, the single CPE Router must advertise two default routes into the network, one "from" each PA prefix.

In this case, the destination prefix in routing is a default route, ::/0. The source prefix is the prefix allocated by the ISP. In this case, routing within the network is largely unchanged, as all traffic to another network goes to the CPE Router, but the CPE Router must send it to the correct ISP.

Note that in this use case, if there are other routers or internal routes in the network, there is no need for them to specify source prefixes on their routes, and if they do, the prefix specified is likely to be :;/0. The reason is that traffic arriving from the ISPs must be delivered to destinations within the network, so routing cannot preclude them.

2.2. General Egress Routing

A more general use case is as shown in Figure 2. A customer network has two or more upstream networks, with a separate CPE Router for each one. Each upstream network allocates a prefix for use in the customer network, and the customer network configures a subnet from each of those ISP prefixes on each of its LANs. Each CPE Router advertises a default route into the customer network. Apart from prefix itself, the services of the upstream ISPs are indistinguishable; they each get the customer to the Internet.

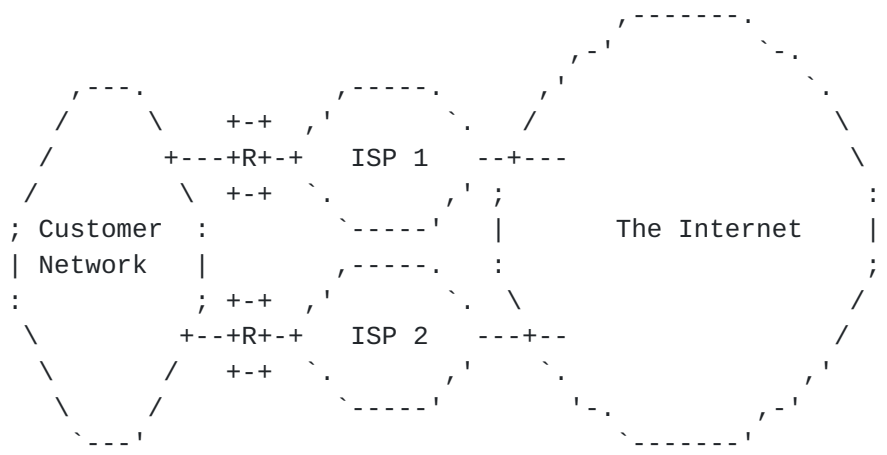

Figure 2: Egress Routing in a Multihomed Environment

The big issue in this network is again ingress filtering [RFC2827] by the upstream ISP. If packets intended for a remote destination pass through the wrong ISP, they will be blocked. Traffic following default route gets to the upstream network indicated by its source address.

In this case, the destination prefix in routing is a default route, ::/0. The source prefix is the prefix allocated by the ISP. We want a routing algorithm that sends packets matching such a specification to the CPE Router advertising that default route.

Note that in this use case, if there are other routers or internal routes in the network, there is no need for them to specify source prefixes on their routes, and if they do, the prefix specified is likely to be :;/0. The reason is that traffic arriving from the ISPs must be delivered to destinations within the network, so routing cannot preclude them.

2.3. Specialized Egress Routing

A more specialized use case is as shown in Figure 3. A customer network has two or more upstream networks, with one or more CPE Routers; the example shows a separate CPE Router for each one. Each upstream network allocates a prefix for use in the customer network, and the customer network configures a subnet from each of those ISP prefixes on each of its LANs. Some CPE Routers might advertise a default route into the customer network; one or more of the other CPE Routers, perhaps all of them, advertise a more-specific route. The services offered by the upstream networks differ in some important way.

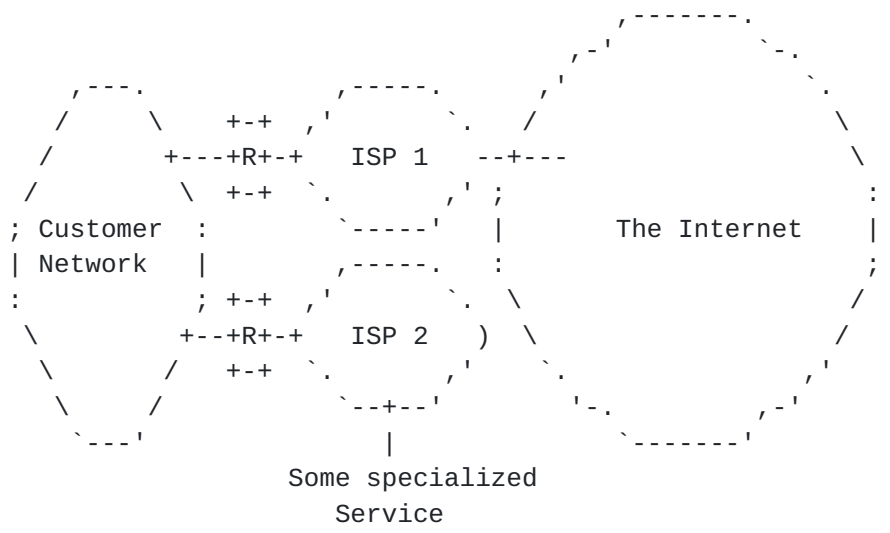

Figure 3: Egress Routing with a specialized upstream network

A specific example of such a service is the NTT B-FLETS video service in Japan; however, the use case describes any use with one or more walled gardens. In the B-FLETS case, a customer may purchase services from a number of ISPs, providing general Internet access. However, the video service requires customers accessing it to use its allocated prefix, and other ISPs (following [RFC2827]) will not accept that prefix as a source address. This is similar to the previous use cases, but

- the only application at that "ISP" is the video service,
- packets using the video service MUST use the video service's source and destination addresses, and
- no other service will accept a video service address as a source address.

The big issue in this network is, once again, ingress filtering [RFC2827] by the upstream ISP, with the additional caveat that the upstream services are far from identical. If packets intended for a remote destination pass through the wrong ISP, they will be blocked. Additionally, while other ISPs advertise access to the general Internet, they may not provide service to the specialized service in question. Hence, egress routing in this case also ensures delivery to the intended destination using the bandwidth it provides. In the ideal case, traffic following default route gets to the upstream network indicated by its source address.

In this case, one or more ISPs might offer a default route as a destination prefix in routing, ::/0. The source prefix is the prefix allocated by the ISP. In addition, the ISP offering the specialized service advertises one or more specific prefixes for those services, with appropriate source prefixes for their use. We want a routing algorithm that sends packets matching such a specification to the CPE Router advertising that indicated route, and dropping, perhaps with an ICMPv6 response, packets for which it effectively has no route.

Note that in this use case, if there are other routers or internal routes in the network, there is no need for them to specify source prefixes on their routes, and if they do, the prefix specified is likely to be :;/0. The reason is that traffic arriving from the ISPs must be delivered to destinations within the network, so routing cannot preclude them.

2.4. Intra-domain access control

A use case within the confines of a single network is as shown in Figure 4. A network has one or more internal networks with differing access permission sets; the financial servers might only be accessible from a set of other prefixes that financial people are located in, or university grade records is only reachable from the offices of professors. This could be implemented using firewalls between the domains, or using application layer filters; in this case, the routing architecture replaces an exclusive firewall rule.

In this case, each domain advertises reachability to its prefix, listing acceptable source prefixes. Domains that are willing to be generally reached might advertise ::/0 as a source prefix, or the prefix in use in the general domain.

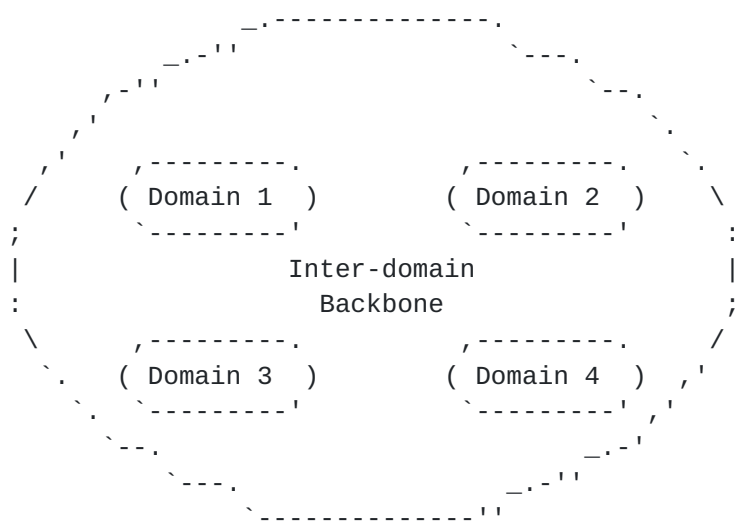

Figure 4: Intradomain Access Control

The big issue in this network is a difference in policy.

2.5. Traffic Engineering

This use case derives from real requirements of CERNET2, an IPv6 network with 59 PoPs and sites from 22 cities. The network shown in Figure 5 has multiple internal networks with different priorities when accessing the target network. For example, domain 1 and domain 2 need higher speed. At the same time, the egress router R1 is much more congested than R2, because traffic from almost all domains (including 1, 2, 3, 4) travel through R1. It is anticipated that network can divert traffic (from some domain to target network) to another egress router for reducing the total latency.

For a mid-size network, CERNET2 wants to make the operations more dynamic and does not want to use static routing or PBR. Also, CERNET2 does not want to use MPLS and MTR, because it does not have MPLS/MTR operators and the learning curve is quite high. So, CERNET2 desires to deploy src/dst routing.

In this case, the egress router advertises reachability from specific source prefixes to the target network, with different metric representing the priority. For example, by adjusting the advertised metrics, the path from domain 1 and 2 towards the target network will have much smaller metrics when going through R2 than through R1. Thus, the routers across the intra-domain will divert the traffic from domain 1 and 2 to R2 when forwarding to the target network.

This implementation uses Source/Destination Routing Using BGP-4 [I-D.xu-src-dst-bgp].

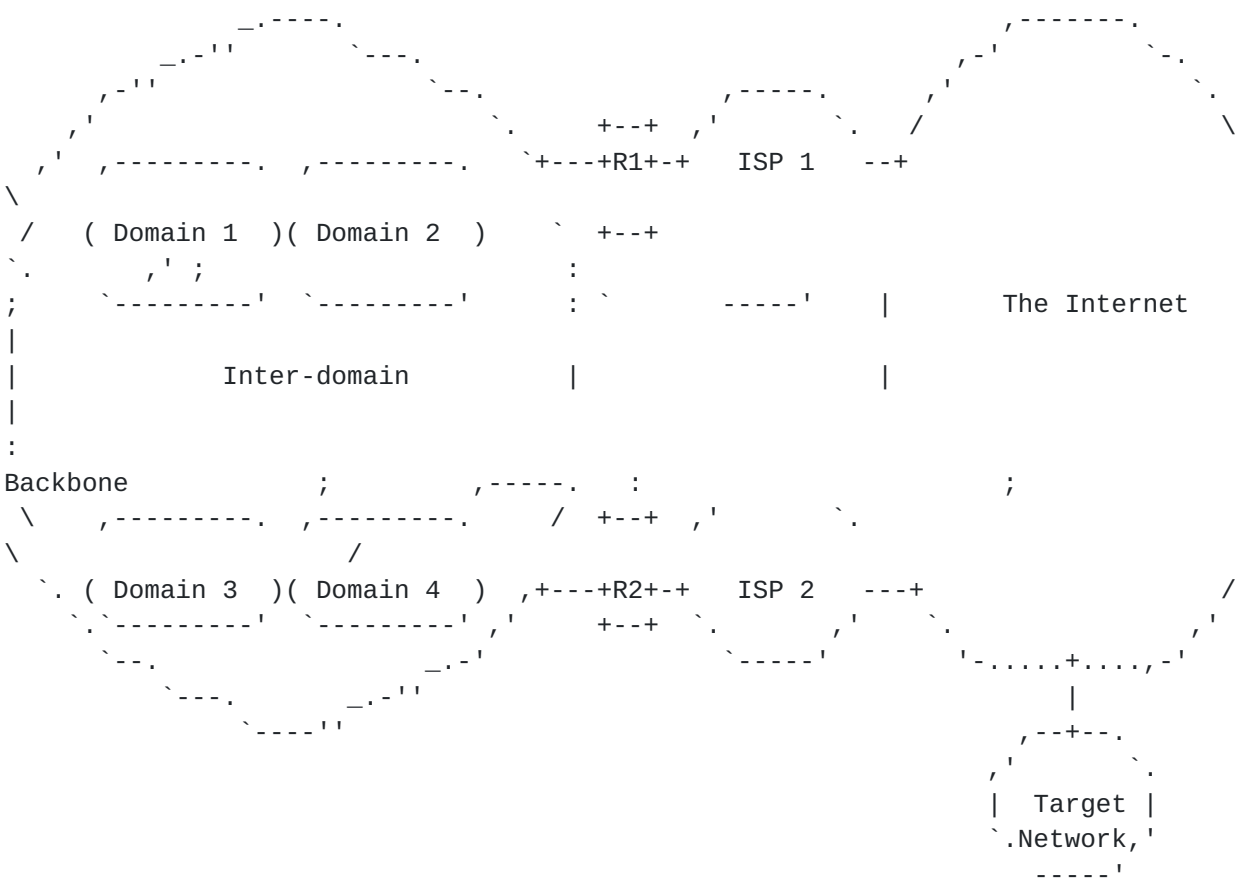

Figure 5: Traffic Engineering

3. Derived Requirements

The use cases in can each be met if:

- The routing protocol or mechanism includes a source prefix. It is acceptable that a default source prefix of ::/0 (all addresses) applies to routes that don't specify a prefix.

- The routing protocol or mechanism includes a destination prefix, which may be a default route (::/0) or any more specific prefix up to and including a host route (/128).

- The FIB lookup yields the route with the most specific (e.g. longest-match) destination prefix that also matches the source prefix constraint, or no match.

4. IANA Considerations

This memo asks the IANA for no new parameters.

5. Security Considerations

As a descriptive document, this note adds no new security risks to the network.

6. Privacy Considerations

As a descriptive document, this note adds no new privacy risks to the network.

7. Acknowledgements

This note was discussed with Acee Lindem, Jianping Wu, Juliusz Chroboczek, Les Ginsberg Lorenzo Colitti, Mark Townsley, Markus Stenberg, Matthieu Boutier, Ole Troan, Ray Bellis, Shu Yang, and Xia Yin.

8. References

8.1. 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>.

8.2. Informative References

[I-D.baker-fun-routing-class]
Baker, F., "Routing a Traffic Class", draft-baker-fun-routing-class-00 (work in progress), July 2011.

[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-05 (work in progress), April 2016.

[I-D.baker-ipv6-ospf-dst-src-routing]
Baker, F., "IPv6 Source/Destination Routing using OSPFv3", draft-baker-ipv6-ospf-dst-src-routing-03 (work in progress), August 2013.

[I-D.boutier-homenet-source-specific-routing]
Boutier, M. and J. Chroboczek, "Source-specific Routing", draft-boutier-homenet-source-specific-routing-00 (work in progress), July 2013.

[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.

[I-D.xu-homenet-traffic-class]
Xu, M., Yang, S., Wu, J., and F. Baker, "Traffic Class Routing Protocol in Home Networks", draft-xu-homenet-traffic-class-02 (work in progress), April 2014.

[I-D.xu-src-dst-bgp]
Xu, M., Yang, S., and J. Wu, "Source/Destination Routing Using BGP-4", draft-xu-src-dst-bgp-00 (work in progress), March 2016.

[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>.

[RFC3704] Baker, F. and P. Savola, "Ingress Filtering for Multihomed Networks", BCP 84, RFC 3704, DOI 10.17487/RFC3704, March 2004, <http://www.rfc-editor.org/info/rfc3704>.

Appendix A. Change Log

Initial Version: August 2013

Repost: October 2014, initial draft reposted on request.

CERNET2: April 2016, CERNET2 use cases added.

Authors' Addresses

Fred Baker
Cisco Systems
Santa Barbara, California 93117
USA

Email: fred@cisco.com

Mingwei Xu
Tsinghua University
Department of Computer Science, Tsinghua University
Beijing 100084
P.R. China

Phone: +86-10-6278-1572
Email: xumw@tsinghua.edu.cn

Shu Yang
Graduate School at Shenzhen, Tsinghua University
Division of Information Science and Technology
Shenzhen 518055
P.R. China

Phone: +86-755-2603-6059
Email: yang.shu@sz.tsinghua.edu.cn

Jianping Wu
Tsinghua University
Department of Computer Science, Tsinghua University
Beijing 100084
P.R. China

Phone: +86-10-6278-5983
Email: jianping@cernet.edu.cn