

Network Working Group
Internet-Draft
Expires: October 23, 2014

M. Xu
S. Yang
J. Wu
Tsinghua University
F. Baker
Cisco Systems
April 21, 2014

Traffic Class Routing Protocol in Home Networks
draft-xu-homenet-traffic-class-02

Abstract

Home IT staff is generally unfamiliar with network operations, making it desirable to provide a configuration-free mode of operation. Policy-based routing (in the sense of configuring one router to redirect traffic to another based on access control) and multi-topology routing both require configuration, making them undesirable.

In this document, we propose a configuration-free mechanism, in which packets will be routed towards the corresponding upstream ISPs based on both destination and source addresses.

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 23, 2014.

Copyright Notice

Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents

(<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Table of Contents

1.	Introduction	2
2.	Terminology	3
3.	Overview	3
4.	Router Behavior	5
4.1.	Egress Router Behavior	5
4.2.	Interior Router Behavior	5
5.	TC-LSA Format	6
6.	Routing Table Structure	8
7.	Calculation of the Routing Table	8
8.	Matching Rule	9
9.	Forwarding Table Structure	9
10.	Implementation	10
11.	Compatibility	10
12.	IANA Considerations	10
13.	Acknowledgments	10
14.	References	11
14.1.	Normative References	11
14.2.	Informative References	11
	Authors' Addresses	11

[1.](#) Introduction

Home networks are growing both in device count and in complexity. Today they generally contain both wired and wireless components, and may require routing to place audio/visual entertainment traffic one path, office services on another, and wireless LANs (both IEEE-

style and 4G/LTE-style) on a third. Traditionally, we have simplified networks using a single exit router and a default route. Today, we might have multiple routers to wired upstream networks, and by separate paths LTE services, "Smart Grid" services, or health network services. Increasingly, such networks are multihomed, and multihomed using diverse access network technologies.

Traditionally, routing protocols make routing decisions solely based on destination IP addresses, packets towards the same destination will be delivered to the same next hop no matter where they come from. These protocols work well with simple home networks that have only one egress router. However, in the multi-homing scenario, packets may be dropped if forwarded only based on destination addresses [[RFC3704](#)].

Although many patch-like solutions, like policy-based routing (PBR), multi-topology routing (MTR) and layer-3 VPN can solve the problem, they complex the configurations in home networks, and are not suitable for home IT staffs. We need a configuration-free solution to help operators set up their home networks in the multi-homing scenario.

In this document, we propose a configuration-free mechanism - traffic class routing, based on OSPFv3, such that home networks can route packets towards the corresponding upstream ISPs, according to both destination and source addresses.

2. Terminology

Terminology used in this document:

- o Traffic Class (TC): Identified by (destination prefix, source prefix), all packets falling in the domain belong to the traffic class.
- o TC-Route: Identified by (destination prefix, source prefix, value), where value is the administrative value applied to the traffic class (destination prefix, source prefix).
- o TC-LSA: Link state advertisement that communicates the reachability for a traffic classes.

3. Overview

In a home network, traditionally, egress routers obtain delegated prefixes from upstream ISPs using DHCPv6 with prefix options [[RFC3633](#)]. The egress routers will then assign longer sub-prefixes to the other links in the home network. Each router inside the home

network will act as standard OSPFv3 router, and forward packets based on their destination addresses.

With traffic class routing, after obtaining delegated prefixes and assigning sub-prefixes, egress routers will populate traffic classes (with extended LSAs), rather than destination address only, into the home network. Each router inside the home network will flood these traffic classes information. When calculating the path towards a destination address, routers will take the traffic classes into considerations. Intrinsically, in traditional routing model, the object being routed to is a destination prefix; in our routing model, the object being routed might be a destination prefix given that the packet sports a certain source prefix.

Each traffic class is associated with a cost, which is a single dimensionless metric.

For example, a site is connected to the Internet through two ISPs, ISP1 and ISP2. ISP1 delegates prefix P1 to the site, and ISP2 delegates prefix P2 to the site. After being delegated with P1, the egress router E1 of the site will advertise a traffic class - $\{::/0, P1\}$, into the site. After being delegated with P2, the egress router E2 of the site will advertise a traffic class - $\{::/0, P2\}$, into the site. Receiving these advertisements, interior router I1 will compute two paths towards $::/0$, one through router E1 for traffic from P1, the other through E2 for traffic from P2.

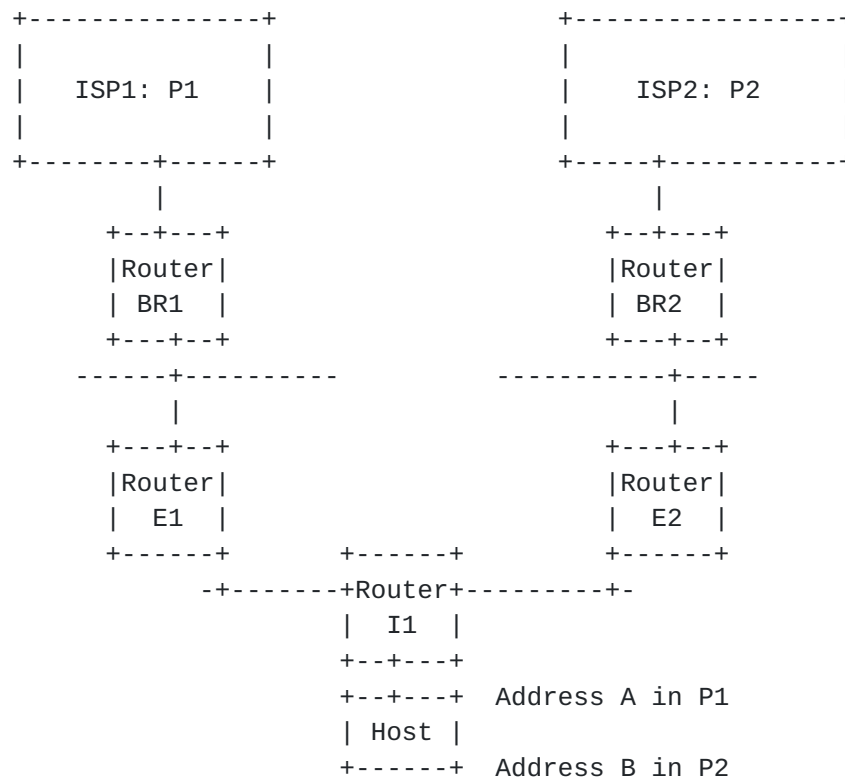


Figure 1: Multi-homing Scenario in Home Networks

4. Router Behavior

All routers behave like traditional OSPFv3 routers, however, the following behaviors are different with traditional OSPFv3 routers.

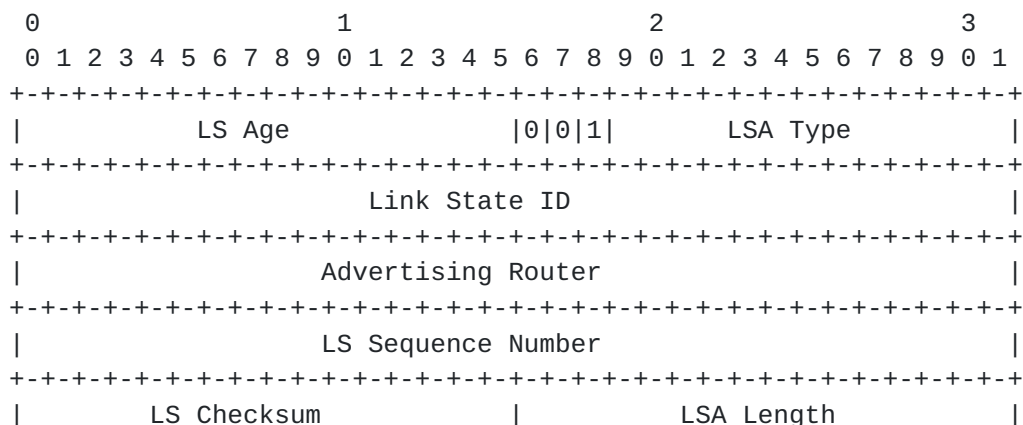
4.1. Egress Router Behavior

After obtaining delegated prefixes using DHCPv6 with prefix options, an egress router should originate TC-LSAs, i.e., extended LSAs with source prefixes appended. Egress routers then will advertise these TC-LSAs into the home network.

Note that an egress router behaves like an interior router if it receives a TC-LSA from other egress routers.

4.2. Interior Router Behavior

Receiving TC-LSAs from egress routers, an interior router should store the TC-LSAs into its LSDB, and flood it to other routers. After calculating a path to an egress router advertising reachability, i.e., a destination prefix, the interior router should decide which traffic class can follow this path towards the egress




```

+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      # Prefixes      |      Referenced LS Type      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      Referenced Link State ID      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      Referenced Advertising Router      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| PrefixLength | PrefixOptions |      Metric      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      Address Prefix      |
|      ...      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      ...      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| PrefixLength | PrefixOptions |      Metric      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      Address Prefix      |
|      ...      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      TLV Type      |      TLV Length      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| SPrefixLength | SPrefixOptions|      0      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      Source Address Prefix      |
|      ...      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      ...      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      TLV Type      |      TLV Length      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| SPrefixLength | SPrefixOptions|      0      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      Source Address Prefix      |
|      ...      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Figure 1: Multi-homing Scenario in Home Networks

All LSA header fields are the same as defined in [[RFC5340](#)], except the following:

- o LSA type: The LSA type value is 0xA029, according to [[I-D.acee-ospfv3-lsa-extend](#)];
- o LSA length: The length of the whole LSA header, including the TLVs;

- o TLV type: The type of IPv6 source prefix TLV, assigned by IANA;
- o TLV length: The value is 20 as defined in [\[I-D.baker-ipv6-ospf-dst-src-routing\]](#);
- o SPrefixLength, SPrefixOptions, Source Address Prefix: Representation of the IPv6 address prefix, which is delegated from the upstream ISP providers;

For simplicity, each extended LSA should only carry one source prefix, suppose there are n destination prefix d_1, d_2, \dots, d_n , and the source prefix is s , then the LSA carries n TC-route announcement, $(d_1, s, v_1), (d_2, s, v_2), \dots, (d_n, s, v_n)$, where v_i is the metric associated with destination prefix d_i .

6. Routing Table Structure

For traditional routing, the routing table structure contains all needed information to forward IP packets to the right destination. For example, destination prefixes are commonly structured into a prefix trie, where each trie nodes contain the necessary information. Routers can lookup and update the prefix trie.

With traffic classes, the routing table structure must contain all needed information to forward IP packets following the right traffic class, i.e., towards the related destination and from the related source. For each routing table entry, there are two additional fields other than the fields mentioned in [\[RFC5340\]](#):

- o Source IP Address: The IP address of the source in traffic class.
- o Source Address Mask: If the source is a subnet, then it is referred to as the subnet mask.

The routing table must provide interface for update and lookup in it. For example, traffic classes can be structured into a two dimensional (or two level) trie, where each trie node in the first dimension points to a sub-trie in the second dimension. The trie nodes in the second dimension contain the necessary information to forward IP packets following the right traffic class.

7. Calculation of the Routing Table

The fundamental algorithm in OSPFv3 doesn't change. The algorithm uses the SPF approach to calculate a path to the router advertising reachability, and then uses the reachability advertisement to decide what traffic should follow that route. What we are changing is the reachability advertisement, in traditional OSPFv3, the

advertisements, which is one or several kinds of LSAs, represent destination prefixes; in this document, the advertisements, which is one or several kinds of TC-LSAs, represent traffic classes.

Note that we do not have to change router-LSA and network-LSA in [\[RFC5340\]](#). Thus, the first stage of [Section 4.8.1 in \[RFC5340\]](#) remains the same in this document. However, the second stage of [Section 4.8.1 in \[RFC5340\]](#) should change by a little bit. Instead of examining the list of the intra-area-prefix-LSAs, the list of extended intra-area-prefix-LSAs is examined. The cost of any advertised traffic class is the sum of the class' advertised metric plus the cost of the transit vertex (either router or transit network) indentified by extended intra-area-prefix-LSAs' referenced LS type, referenced link state ID, and referenced advertising router field.

8. Matching Rule

We also adopt the LMF (longest match first) rule when a packet matches multiple routing entries. However, traffic class has two dimensions, there might exist ambiguity. For example, if there exists two routing entries, (d1, s1, nexthop1), (d2, s2, nexthop2), where d1 is longer than d2 and s2 is longer than s1, then none entry is longer than the other in both dimensions. In this situation, we must insert an additional entry into the routing table, e.g., (d1, s2, nexthop1) in the above example. The entry directs to nexthop1 rather than nexthop2, because we must guarantee reachability according to the destination prefix.

9. Forwarding Table Structure

As the format of routing table entries changes from (destination prefix, nexthop) to (destination prefix, source prefix, nexthop), The Forwarding Information Base (FIB), based on which routers forward IP packets, should be redesigned. There can be different possible FIB structures, e.g., TCAM-based structure, linear table structure, trie-based structure, etc. we use a trie-based solution in this document.

In traditional routing, the destination prefixes are constructed using a prefix trie. However, in traffic class routing, the forwarding decision is based on both the destination and source prefixes. To store them, we design a new structure, based on patricia-trie to meet the needs.

In traffic class routing, the router forwards IP packets following the right traffic class based on destination and source prefixes. Routers construct a two dimensional (or two level) patricia-trie, where each trie node in the first dimension points to a sub-trie (the

sub-trie can be empty). When a packet arrives, the packet should match a trie node (destination prefix) in the first dimension using the destination address. Following the sub-trie under the node, the packet should match another trie node (source prefix) in the second dimension, i.e., the sub-trie.

10. Implementation

We implemented the routing protocol in Quagga, an open-source routing software. In Quagga, we modified the LSA format to support E-Intra-Area-Prefix-LSA, the routing table format, and the routing table calculation algorithm.

To implement the FIB structure, we modified Click, which is a modular software router. The FIB structure could support lookup and update. When the routing table changes, routers can update the FIB through pre-defined interfaces.

11. Compatibility

Routers can also announce the traditional destination-based LSAs, e.g., Intra-Area-Prefix-LSA, at the same time. When a router receives traditional destination-based LSAs, it has two choices. In the first choice, the destination-based LSAs are treated as TC-LSAs where the source prefix equal the wildcard, and routers keep all routes in a common routing table. When a packet arrives, routers need to lookup in the routing table according to the matching rule. In the second choice, routers have to keep two routing tables, one for destination prefix only, and the other for traffic classes. When a packet arrives, routers first lookup in the routing table storing traffic classes; If none entry matches, then routers lookup in the routing table storing destination prefixes.

12. IANA Considerations

The newly LSA types and TLVs should be assigned by IANA, please refer to [[I-D.baker-ipv6-ospf-dst-src-routing](#)] and [[I-D.acee-ospfv3-lsa-extend](#)].

13. Acknowledgments

Zheng Liu and Gautier Bayzelon provided useful input into this document.

14. References

14.1. Normative References

- [RFC3704] Baker, F. and P. Savola, "Ingress Filtering for Multihomed Networks", [BCP 84](#), [RFC 3704](#), March 2004.
- [RFC3633] Troan, O. and R. Droms, "IPv6 Prefix Options for Dynamic Host Configuration Protocol (DHCP) version 6", [RFC 3633](#), December 2003.
- [RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", [RFC 5340](#), July 2008.

14.2. Informative References

- [I-D.ietf-homenet-arch]
Chown, T., Arkko, J., Brandt, A., Troan, O., and J. Weil,
"Home Networking Architecture for IPv6", [draft-ietf-homenet-arch-07](#) (work in progress), February 2013.
- [I-D.baker-ipv6-ospf-dst-src-routing]
Baker, F., "IPv6 Source/Destination Routing using OSPFv3",
[draft-baker-ipv6-ospf-dst-src-routing-02](#) (work in progress), May 2013.
- [I-D.acee-ospfv3-lsa-extend]
Lindem, A., Mirtorabi, S., Roy, A., and F. Baker, "OSPFv3 LSA Extendibility", [draft-acee-ospfv3-lsa-extend-00](#) (work in progress), May 2013.
- [I-D.baker-fun-routing-class]
Baker, F., "Routing a Traffic Class", [draft-baker-fun-routing-class-00](#) (work in progress), July 2011.

Authors' Addresses

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

Phone: +86-10-6278-5822
Email: xmw@csnet1.cs.tsinghua.edu.cn

Shu Yang
Tsinghua University
Department of Computer Science, Tsinghua University
Beijing 100084
P.R. China

Phone: +86-10-6278-5822
Email: yangshu@csnet1.cs.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

Fred Baker
Cisco Systems
Santa Barbara, California 93117
USA

Email: fred@cisco.com

