

Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: January 17, 2013

Q. Sun
China Telecom
C. Zhou
Huawei Technologies
July 16, 2012

Multicast transition path optimization in IPv4 and IPv6 networks
draft-zhou-mboned-multtrans-path-optimization-02

Abstract

This document describes a mechanism to optimize the path between the multicast router and multicast source in both IPv4 and IPv6 networks. The basic idea is that when a multicast translation router has an IPv4 path and an IPv6 path to the same multicast data source, and both IPv4 and IPv6 joins are received, only one path is used. One path is pruned, instead of the same traffic flowing over both v4 and v6 paths. By adding a metric to the IPv4 path, the multicast translation router can determine which path to receive multicast data: IPv4 path, IPv6 path or both. Therefore, an optimization path will typically be chosen when an identical v4/v6 traffic flow exists.

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 .

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 17, 2013.

Copyright Notice

Copyright (c) 2012 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.

Table of Contents

1. Introduction	3
2. Terminology	3
3. Scenarios	4
4. Solution Overview	4
4.1. A general topology for IPv4 and IPv6 multicast networks	5
4.2. Parsing MTR to two virtual Routers	6
4.3. Selecting interfaces to Source or RP	7
4.4. Selecting a multicast data flow from upstream interface	7
4.5. Modifications to multicast Router	7
5. Security Considerations	8
6. Acknowledgments	8
7. IANA Considerations	8
8. Informative References	8
Authors' Addresses	9

1. Introduction

It is common to use multi-access LANs such as Ethernet for transmitting multicast data in networks. Section 3.6 of[RFC4601] describes Multi-Access Transit LANs.

The PIM Assert message could be used when there are two identical multicast data flows (IPv4 and IPv6). When duplicate data packets appear on the LAN from different routers, the routers notice this and then select a single forwarder. This selection is performed using PIM Assert messages, which solve the problem in favor of the upstream router that has (S,G) state; Or, if neither or both router has (S,G) state, then the problem is solved in favor of the router with the best metric to the RP for RP trees, or the best metric to the source via source-specific trees.

During IPv6 transition, it is common that there are many IPv4 networks and IPv6 networks that connected to each other, which means that multiple multicast translation routers(MTR) exist at the edge of a network. For robustness, reliability and load balance purpose, MTR function could be implemented in several nodes in the network. MTR can be the mAFTR (Multicast AFTR) mentioned in [draft-ietf-softwire-dslite-multicast]. mAFTR can encapsulate IPv4 multicast data in IPv6 tunnel. MTR can also be the mXlate (Multicast Translator) as mentioned in [draft-lee-behave-v4v6-mcast-fwk]. mXlate can translate IPv4 multicast data to IPv6 multicast data.

As a result, MTR (mXlate or mAFTR) will have more than one path to reach the RP or source S in IPv4 networks and IPv6 networks. Or in other words, they will have two upstream routers: one is IPv6 router, and the other is IPv4 router. MTR can reach the RP or source S by both paths. Since MTR can receive both IPv4 and IPv6 (*,G) (or (S,G)) Join request, it needs to select a best path to RP or S in both IPv4 and IPv6 networks. When it receives the two identical multicast data flows via IPv4 and IPv6 interfaces, MTR needs to send Prune Message to the worse path interface. Figure 1 shows the scenario that MTR can reach source S through both IPv4 path and IPv6 path.

2. Terminology

This document makes use of the following terms:

mXlate: A multicast translator mentioned in [draft-lee-behave-v4v6-mcast-fwk].

mAFTR: A multicast Address Family Transition Router mentioned in [draft-ietf-softwire-dslite-multicast].

MTR: A multicast translation router, it can be mAFTR or mXlate.

PIM-SM: Protocol Independent Multicast-Sparse Mode

RP:Rendezvous Point

3. Scenarios

During the multicast transition from IPv4 to IPv6, there may be a router which receives IPv4 join (PIM or IGMP) on one interface, and an IPv6 join (PIM or MLD) on another interface (or it could even be the same interface). The router should support IPv4 PIM and IPv6 PIM that is translation capable. Assume these joins are for both IPv4 (S4,G4) and IPv6 (S6,G6), and that there are active sources for both, sending basically the same content. Either because there is a real source for both, or some upstream router is translating. The router could then simply send upstream joins for both of these, and forward the traffic as needed without translation.

However, if the router is aware that these two sources are really the same content, it could select to join just one of the streams, and translate as needed for the one downstream that wants a different protocol. In this case, there will be a tradeoff between bandwidth on the upstream links, and the cost of translation (both on this device, and perhaps the quality of the stream). When PIM Assert message is used to achieve this, the metrics for IPv4 and IPv6 should be comparable and all the PIM devices on the link should support PIM assert.

4. Solution Overview

This section gives a sloution for the issues mentioned above.

4.1. A general topology for IPv4 and IPv6 multicast networks

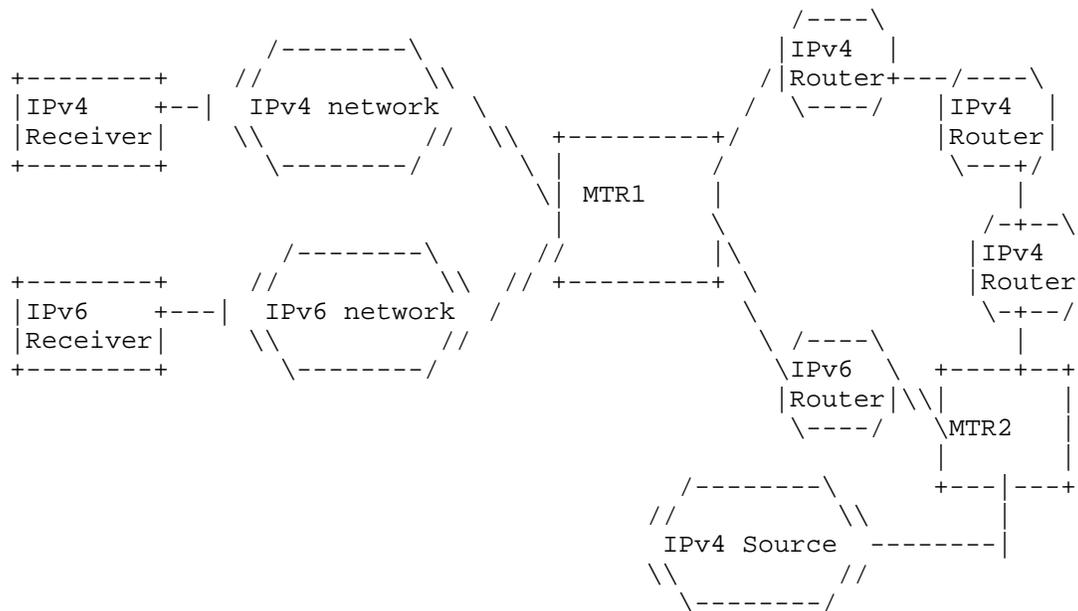
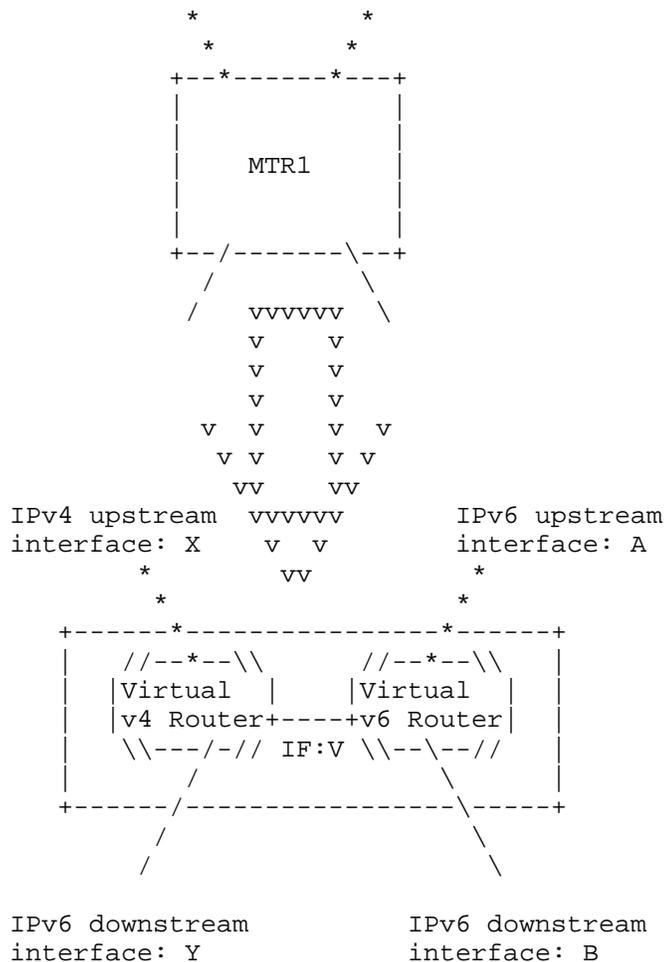


Figure 1: MTR can reach IPv4 Source through IPv4 path and IPv6 path

Figure 1 shows that MTR1 can access IPv4 Source through IPv4 path or IPv6 path. MTR1 has two upstream routers, one is IPv4 Router and the other is IPv6 Router. MTR1 receives IPv4 (*,G) or (S,G) Join request from IPv4 network and IPv6 (*,G) or (S,G) Join request from IPv6 network. MTR1 can send Join request to RP or source S from interface connected to IPv4 Router or from interface connected to IPv6 Router. MTR1 may also send Join request from both upstream interfaces. In this case, MTR1 need to select a best path to RP or S in both IPv4 and IPv6 networks. MTR1 sends Prune Message to the worse path, when it receives two identical multicast data flows in IPv4 and IPv6 upstream interface. MTR1 may receive two identical multicast data flows at the same time and stop interworking multicast data flow between IPv4 network and IPv6 network.

4.2. Parsing MTR to two virtual Routers



For simplification, we use two virtual Routers to replace MTR1 Router. Figure 2 shows that MTR1 can be taken as two virtual Routers. The one on the left is a Virtual IPv4 Router, the one on the right is a Virtual IPv6 Router. Virtual IPv4 Router has an IPv4 upstream interface X and an IPv4 downstream interface Y. Virtual IPv6 Router has an IPv6 upstream interface A and an IPv6 downstream interface B. The interface between two virtual Routers is V.

When MTR receives two multicast data flows (one from IPv4 interface and the other from IPv6 interface), it compares two flows according to [draft-ietf-mboned-64-multicast-address-format] to confirm whether they are identical data flows. If they are the same, select one or

two. When MTR Receives a IPv6 (S, G) or (*, G)Join, virtual IPv6 Router selects an interface to send Join message. The interface can be IPv6 upstream interface A or IPv4 upstream interface X (via interface V).

4.3. Selecting interfaces to Source or RP

The steps to select an interface to S or RP.

- 1.Set the Metric value $m1$ for translation or encapsulation from IPv4 multicast to IPv6 multicast data.
- 2.From interface A connecting IPv6 Router, MTR can get the metric $m2$ to reach S or RP by PIM assert message sent from IPv6 Router.
- 3.From interface X connecting IPv4 Router, MTR can get the metric $m3$ to reach S or RP by PIM assert message from IPv4 Router.
- 4.When MTR receives a IPv6 PIM Join message, virtual IPv6 Router compares $m2$ and $m3+m1$. If $m2 > m3+m1$, sending PIM Join message from IPv4 interface; If $m2 < m3+m1$, sending PIM Join message from IPv6 interface; If $m2 = m3+m1$, MTR can choose interface X or A to send PIM Join message.

4.4. Selecting a multicast data flow from upstream interface

The steps for selecting a multicast data flow from upstream interface

- 1.MTR receives two identical mulitcast flows from IPv6 and IPv4 Router. The address formats of the two flows follows [draft-ietf-mboned-64-multicast-address-format].
- 2.If virtual IPv6 Router receives multicast data from interface V, it will compare $m2$ and $m3+m1$ (the value is from last section).
- 3.If $m2 > m3+m1$, MTR will send PIM Prune Messages to IPv6 interface A; If $m2 < m3+m1$ MTR will send PIM Prune Messages to interfaceX via virtual interface V. MTR will not translate multicast data from IPv4 to IPv6 or encapsulate IPv4 multicast data in IPv6 packets. If $m2 = m3+m1$, MTR selects any interface to receive multicast data and sends PIM Prune Messages to the other interface.

4.5. Modifications to mulitcast Router

The main modifications to the edge PIM-SM Router include:

Edge PIM-SM Router needs to check multicast data flow from IPv4 and IPv6 interfaces based on

[draft-ietf-mboned-64-multicast-address-format] to determine whether they are the same multicast data flow.

Edge PIM-SM Router sends PIM Assert messages via IPv4 and IPv6 interfaces with different Metric value.

Edge PIM-SM Router may stop translating/encapsulating IPv4 multicast flow to IPv6 multicast flow or send Prune Messages to stop receiving IPv6/IPv4 multicast flow.

5. Security Considerations

6. Acknowledgments

Thanks Ronald Bonica, Stig Venaas and Yiu Lee for their valuable comments.

7. IANA Considerations

8. Informative References

[RFC4601] IETF, "Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol Specification (Revised)", Aug 2006, <<http://datatracker.ietf.org/doc/rfc4601/>>.

[draft-ietf-mboned-64-multicast-address-format]
IETF, "IPv4-Embedded IPv6 Multicast Address Format", Feb 2012, <<http://datatracker.ietf.org/doc/draft-ietf-mboned-64-multicast-address-format/>>.

[draft-ietf-software-dslite-multicast]
IETF, "Multicast Extensions to DS-Lite Technique in Broadband Deployments", Oct 2011, <<http://datatracker.ietf.org/doc/draft-ietf-software-dslite-multicast/>>.

[draft-lee-behave-v4v6-mcast-fwk]
IETF, "IPv4/IPv6 Multicast Translation Framework", Feb 2011, <<http://tools.ietf.org/id/draft-lee-behave-v4v6-mcast-fwk-00.txt>>.

Authors' Addresses

Qiong Sun
China Telecom
Xizhimenneidajie Xicheng District
Beijing, 100035
China

Phone:
Fax:
Email: sunqiong@ctbri.com.cn
URI:

Cathy Zhou
Huawei Technologies
Section F, R&D Building, Huawei Longgang Production Base
Shenzhen, 518129
China

Phone:
Fax:
Email: cathy.zhou@huawei.com
URI:

