

Partitioning Tag Space among Multicast Routers on a Common Subnet
<[draft-farinacci-multicast-tag-part-00.txt](#)>

Status of this Memo

This document is an Internet-Draft. Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

To learn the current status of any Internet-Draft, please check the "l1d-abstracts.txt" listing contained in the Internet-Drafts Shadow Directories on ftp.is.co.za (Africa), nic.nordu.net (Europe), munnari.oz.au (Pacific Rim), ds.internic.net (US East Coast), or ftp.isi.edu (US West Coast).

Abstract

There are 3 major functions that must be performed to achieve multicast tagswitching. 1) Tag Allocation, which requires each multicast Tag Switching Router (TSR) to have a tag value range that it uses. 2) Tag Binding, using the tags allocated, a TSR must assign them to multicast routes. 3) Tag Binding Distribution, after binding tag values to routes, they must be distributed to other TSRs so they all forward on a common and consistent distribution tree.

In this document we present how tags are allocated uniquely across multicast capable TSRs on a LAN and point-to-point IP subnets.

1. Introduction

This memo specifies a simple algorithm to partition a tag allocation space among multicast TSRs on a common media. So that each range allocated to a router is unique and non overlapping with any range allocated to any other TSR on that network media.

Since an upstream TSR will use a single tag for a multicast packet, all downstream TSRs, on the same subnet as the upstream router, must be ready to use it to do multicast tag forwarding. Therefore, no other TSR, on the subnet, can use the same tag or it would be ambiguous which multicast distribution tree to forward on. Therefore, each TSR might be potentially an upstream TSR for different multicast distribution trees and needs its own tag range.

This procedure can be used for any tag binding and distribution scheme, be it upstream or downstream multicast tag distribution.

2. LAN Procedure

Each multicast TSR on a LAN is configured with the tag table size it will use for the LAN interface. An approximate router-count will also be configured. This allows a router to allocate a range equal to $1/\text{router-count}$ of the tag table size. This tag table is used for multicast tag forwarding only.

When a multicast TSR boots up or enables the LAN interface to do multicast routing, it will advertise in PIM Hello messages the tag table size, router count and the tag range it randomly selects. The lower range tag value and the higher range tag value accompany the advertisement. A TSR verifies that no other TSR on the LAN is using the tag range before it advertises the one it selected.

If there is another TSR that has selected the same range and it has a higher IP address than the newly booted TSR, the other TSR will trigger a PIM Hello message unicasted to that newly booted TSR indicating that it owns that range. The newly booted TSR will then select another range randomly. If the newly booted TSR has a higher IP address than the other TSR, the other TSR will withdraw his tag bindings and randomly allocate another tag range. Then it will reallocate its bindings.

A TSR can be configured to use more than one tag range if one believes it will be an upstream TSR for many flows. It just inserts additional advertisements in the same PIM Hello message. The tag table size and router-count should be the same in all advertisements contained in a message.

3. Point-to-point Procedure

On point-to-point links since there can only be one forwarder from a TSR's point of view (the other end of the link), each TSR can use the entire tag table size as their tag range. There is no conflict

because there can be one and only one downstream neighbor for a given distribution tree.

Also, the tag table size advertised in PIM Hello messages over point-to-point links don't have to be the same size. The router-count is ignored for point-to-point advertisements.

4. Configuration Errors

If the tag table size is not configured consistently on all TSRs connected to a LAN, the smallest table size advertised by any TSR will be used.

If the router-count is not configured consistently on all TSRs connected to a LAN, the smallest router-count value advertised by any TSR will be used. This means the largest division of the tag table space will occur.

5. Subnet Partitioning

When a subnet partitions and new multicast TSRs come up, they will allocate tag ranges that are unique to their partition. When the partition heals, there may be conflicts. Once the PIM Hellos messages are received by TSRs on the other side of the partition, they will determine there is a tag range allocation conflict and immediately perform the tie breaking rules described above.

6. Exceeding the Tag Table Size

When a TSR cannot allocate a tag range because all ranges within the tag table size have been allocated, it will not participate in binding tags to multicast routes. Packets for these routes will not be tagswitched. However, the TSR is still capable of tagswitching a packet as either an upstream or downstream TSR on that LAN. This is the case when another router is binding tags for the multicast route and has an allocated a tag range successfully.

7. PIM Hello Packet Format

The PIM Hello message will carry a new OptionType (called "Tag Parameters") as specified in [2]. A router that sends a PIM Hello with this option is regarded as being tag-capable. This Option can appear multiple times in a Hello packet if a TSR wants to allocate multiple ranges. When this option appears multiple times in the Hello message, the Tag Table Size and Router Count must be the same for each Tag Parameters Option supplied in the message.

When sent on point-to-point links, this option should have Router Count, Lower Tag Range, and Upper Tag Range set to 0. These fields are ignored on receipt.

```

0                               1                               2                               3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      OptionType = 17      |      OptionLength = 16      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Tag Table Size                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Router Count                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Lower Tag Range                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Upper Tag Range                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

OptionType "Tag Parameters"
Set to value 17 decimal.

OptionLength
The option is 16 bytes in length.

Tag Table Size
The size of the TIB, in number of entries, the sending router supports on the interface the Hello is sent on.

Router Count
The approximate maximum number of routers that may be connected to the subnet the Hello is sent on.

Lower Tag Range
The lower tag value in the tag range that was been randomly selected by the sending router. This value must be less than the Upper Tag Range value.

Upper Tag Range

The upper tag value in the tag range that was been randomly selected by the sending router. This value must be greater than the Lower Tag Range value.

8. Security Considerations

Security considerations are not discussed in this memo.

9. Acknowledgments

Contributions to this work has been made by Yakov Rekhter.

10. Author's Address:

Dino Farinacci
Cisco Systems, Inc.
170 Tasman Drive
San Jose, CA, 95134

Email: dino@cisco.com

11. References

- [1] Tag Switching Architecture Overview, [draft-rfcd-tag-switching-overview-00.txt](#), Rekhter, Davie, Katz, Rosen, Swallow
- [2] Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol Specification, <[draft-ietf-idmr-pim-sm-spec-09.txt](#)>, Estrin, Farinacci, Helmy, Thaler, Deering, Handley, Jacobson, Liu, Sharma, Wei, October, 1996

