Network working group
Internet Draft
Category: Standard Track

X. Xu
Huawei
N. Sheth
Contrail Systems
L. Yong
Huawei
C Pignataro
Cisco
Y. Fan
China Telecom

Expires: July 2013 January 11, 2013

Encapsulating MPLS in UDP

draft-ietf-mpls-in-udp-00

Abstract

Existing technologies to encapsulate Multi-Protocol Label Switching (MPLS) over IP are not adequate for efficient load balancing of MPLS application traffic, such as MPLS-based Layer2 Virtual Private Network (L2VPN) or Layer3 Virtual Private Network (L3VPN) traffic across IP networks. This document specifies additional IP-based encapsulation technology, referred to as MPLS-in-User Datagram Protocol (UDP), which can facilitate the load balancing of MPLS application traffic across IP networks.

Status of this Memo

The list of Internet-Draft Shadow Directories can be accessed at

http://www.ietf.org/shadow.html.

This Internet-Draft will expire on July 11, 2013.

Copyright Notice

Conventions used in this document

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 RFC-2119 [RFC2119].

Table of Contents

1. Introduction

To fully utilize the bandwidth available in IP networks and/or facilitate recovery from a link or node failure, load balancing of traffic over Equal Cost Multi-Path (ECMP) and/or Link Aggregation Group (LAG) across IP networks is widely used. In effect, most existing core routers in IP networks are already capable of distributing IP traffic flows over ECMP paths and/or LAG based on the hash of the five-tuple of User Datagram Protocol (UDP)[RFC768] and Transmission Control Protocol (TCP) packets (i.e., source IP address, destination IP address, source port, destination port, and protocol).

In practice, there are some scenarios for Multi-Protocol Label Switching (MPLS) applications (e.g., MPLS-based Layer2 Virtual Private Network (L2VPN) or Layer3 Virtual Private Network (L3VPN)) where the MPLS application traffic needs to be transported through IP-based tunnels, rather than MPLS tunnels. For example, MPLS-based L2VPN or L3VPN technologies may be used for interconnecting geographically dispersed enterprise data centers or branch offices across IP Wide Area Networks (WAN) where enterprise own router devices are deployed as L2VPN or L3VPN Provider Edge (PE) routers. In this case, efficient load balancing of the MPLS application traffic across IP networks is much desirable.

1.1. Existing Technologies

With existing IP-based encapsulation methods for MPLS applications, such as MPLS-in-IP and MPLS-in-Generic Routing Encapsulation (GRE) [RFC4023] or even MPLS-in-Layer Two Tunneling Protocol - Version 3 (L2TPv3)[RFC4817], distinct customer traffic flows between a given PE router pair would be encapsulated with the same IP-based tunnel headers prior to traversing the core of the IP WAN. Since the encapsulated traffic is neither TCP nor UDP traffic, for many existing core routers which could only perform hash calculation on fields in the IP headers of those tunnels (i.e., source IP address, destination IP address), it would be hard to achieve a fine-grained load balancing of these traffic flows across the network core due to the lack of adequate entropy information.

[RFC5640] describes a method for improving the load balancing efficiency in a network carrying Softwire Mesh service over L2TPv3 and GRE encapsulation. However, this method requires core routers to be capable of performing hash calculation on the "load-balancing" field contained in the tunnel encapsulation headers (i.e., the Session ID field in the L2TPv3 header or the Key field in the GRE header), which means a non-trivial change to the date plane of many existing core routers.

1.2. Motivations for MPLS-in-UDP Encapsulation

On basis of the fact that most existing core routers (i.e., P routers in the context of MPLS-based L2VPN or L3VPN) are already capable of balancing IP traffic flows over the IP networks based on the hash of the five-tuple of UDP packets, it would be advantageous to use MPLS-in-UDP encapsulation instead of MPLS-in-GRE or MPLS-in-L2TPv3 in the environments where the load balancing of MPLS application traffic across IP networks is much desired but the load balancing mechanisms defined in [RFC5640] have not yet been widely supported by most existing core routers. In this way, the default load balancing capability of most existing core routers as mentioned above can be utilized directly without requiring any change to them.

2. Terminology

This memo makes use of the terms defined in [RFC4364] and [RFC4664].

3. Encapsulation in UDP

MPLS-in-UDP encapsulation format is shown as follows:

```
         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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

```
|    Source Port = entropy      |       Dest Port = MPLS        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           UDP Length          |        UDP Checksum           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
~                       MPLS Label Stack                        ~
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
~                         Message Body                          ~
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Source Port of UDP

This field contains an entropy value that is generated by the ingress PE router. For example, the entropy value can be generated by performing hash calculation on certain fields in the customer packets (e.g., the five tuple of UDP/TCP packets).

Destination Port of UDP

This field is set to a value (TBD) indicating the MPLS packet encapsulated in the UDP header is a MPLS one or a MPLS one with upstream-assigned label.

UDP Length

The usage of this field is in accordance with the current UDP specification.

UDP Checksum

The usage of this field is in accordance with the current UDP specification. To simplify the operation on egress PE routers, this field is recommended to be set to zero.

MPLS Label Stack

This field contains an MPLS Label Stack as defined in [RFC3032].

Message Body

This field contains one MPLS message body.

4. Processing Procedures

This MPLS-in-UDP encapsulation causes MPLS packets to be forwarded through "UDP tunnels". When performing MPLS-in-UDP encapsulation by an ingress PE router, the entropy value would be generated by the ingress PE router and then be filled in the Source Port field of the UDP header.

P routers, upon receiving these UDP encapsulated packets, could balance these packets based on the hash of the five-tuple of UDP packets.

Upon receiving these UDP encapsulated packets, egress PE routers would decapsulate them by removing the UDP headers and then process them accordingly.

As for other common processing procedures associated with tunneling encapsulation technologies including but not limited to Maximum Transmission Unit (MTU) and preventing fragmentation and reassembly, Time to Live (TTL) and differentiated services, the corresponding procedures defined in [RFC4023] which are applicable for MPLS-in-IP and MPLS-in-GRE encapsulation formats SHOULD be followed.

5. Applicability

Besides the MPLS-based L3VPN [RFC4364] and L2VPN [RFC4761, RFC4762] [E-VPN] applications, MPLS-in-UDP encapsulation could apply to other MPLS applications including but not limited to 6PE [RFC4798] and PWE3 services.

6. Security Considerations

Just like MPLS-in-GRE and MPLS-in-IP encapsulation formats, the MPLS-in-UDP encapsulation format defined in this document by itself cannot ensure the integrity and privacy of data packets being transported through the MPLS-in-UDP tunnels and cannot enable the tunnel decapsulators to authenticate the tunnel encapsulator. In the case where any of the above security issues is concerned, the MPLS-in-UDP tunnels SHOULD be secured with IPsec in transport mode. In this way, the UDP header would not be seeable to P routers anymore. As a result, the meaning of adopting MPLS-in-UDP encapsulation format as an alternative to MPLS-in-GRE and MPLS-in-IP encapsulation formats is lost. Hence, MPLS-in-UDP encapsulation format SHOULD be used only in the scenarios where all the security issues as mentioned above are not significant concerns. For example, in a data center environment, the whole network including P routers and PE routers are under the control of a single administrative entity and therefore there is no need to worry about the above security issues.

7. IANA Considerations

Two distinct UDP destination port numbers indicating MPLS and MPLS with upstream-assigned label respectively need to be assigned by IANA.

8. Acknowledgements

Thanks to Shane Amante, Dino Farinacci, Keshava A K, Ivan Pepelnjak, Eric Rosen, Andrew G. Malis, Kireeti Kompella, Marshall Eubanks, Vivek Kumar, Weiguo Hao, Zhenxiao Liu and Xing Tong for their valuable comments on the idea of MPLS-in-UDP encapsulation. Thanks to Daniel King, Gregory Mirsky and Eric Osborne for their valuable reviews on this draft.

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

9.2. Informative References

[RFC4364] Rosen, E and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, February 2006.

[RFC4664] Andersson, L. and Rosen, E. (Editors),"Framework for Layer 2 Virtual Private Networks (L2VPNs)", RFC 4664, Sept 2006.

[RFC4023] Worster, T., Rekhter, Y., and E. Rosen, "Encapsulating MPLS in IP or GRE", RFC4023, March 2005.

[RFC5332] Eckert, T., Rosen, E., Aggarwal, R., and Y. Rekhter, "MPLS Multicast Encapsulations", RFC 5332, August 2008.

[RFC4817] M. Townsley, C. Pignataro, S. Wainner, T. Seely and J. Young, "Encapsulation of MPLS over Layer 2 Tunneling Protocol Version 3, March 2007.

[RFC5640] Filsfils, C., Mohapatra, P., and C. Pignataro, "Load-Balancing for Mesh Softwires", RFC 5640, August 2009.

[RFC6391] Bryant, S., Filsfils, C., Drafz, U., Kompella, V., Regan, J., and S. Amante, "Flow Aware Transport of Pseudowires over an MPLS Packet Switched Network", RFC6391, November 2011

[RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and L. Yong, "The Use of Entropy Labels in MPLS Forwarding", draft-ietf-mpls-entropy-label-01, work in progress, October, 2011.

[RFC5512] Mohapatra, P. and E. Rosen, "The BGP Encapsulation Subsequent Address Family Identifier (SAFI) and the BGP Tunnel Encapsulation Attribute", RFC 5512, April 2009.

[RFC4798] J Declerq et al., "Connecting IPv6 Islands over IPv4 MPLS using IPv6 Provider Edge Routers (6PE)", RFC4798, February 2007.

[RFC4761] Kompella, K. and Y. Rekhter, "Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling", RFC 4761, January 2007.

[RFC4762] Lasserre, M. and V. Kompella, "Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling", RFC 4762, January 2007.

[E-VPN] Aggarwal et al., "BGP MPLS Based Ethernet VPN", draft-ietf-l2vpn-evpn-00.txt, work in progress, February, 2012.

[RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", RFC 3032, January 2001.

[RFC768] Postel, J., "User Datagram Protocol", STD 6, RFC 768, August 1980.

[I-D.ietf-6man-udpchecksums] Eubanks, M., Chimento, P., and M. Westerlund, "UDP Checksums for Tunneled Packets", draft-ietf-6man-udpchecksums-04 (work in progress), September 2012.

[I-D.ietf-6man-udpzero] Fairhurst, G. and M. Westerlund, "Applicability Statement for the use of IPv6 UDP Datagrams with Zero Checksums", draft-ietf-6man-udpzero-07 (work in progress), October 2012.

Authors' Addresses

Xiaohu Xu

Huawei Technologies,

Beijing, China

Phone: +86-10-60610041

Email: xuxiaohu@huawei.com

Nischal Sheth

Contrail Systems

Email: nsheth@contrailsystems.com

Lucy Yong

Huawei USA

5340 Legacy Dr.

Plano TX75025

Phone: 469-277-5837

Email: Lucy.yong@huawei.com

Carlos Pignataro
Cisco Systems
7200-12 Kit Creek Road
Research Triangle Park, NC 27709
USA
EMail: cpignata@cisco.com

Yongbing Fan
China Telecom
Guangzhou, China.
Phone: +86 20 38639121
Email: fanyb@gsta.com

Zhenbin Li
Huawei Technologies,
Beijing, China
Phone: +86-10-60613676
Email: lizhenbin@huawei.com