Network Working Group
Internet-Draft
Intended status: Informational
Expires: October 5, 2014

M. Konstantynowicz, Ed.
G. Heron, Ed.
Cisco Systems
R. Schatzmayr
Deutsche Telekom AG
W. Henderickx
Alcatel-Lucent, Inc.
April 3, 2014

Keyed IPv6 Tunnel
draft-ietf-l2tpext-keyed-ipv6-tunnel-00

Abstract

This document describes a simple L2 Ethernet over IPv6 tunnel encapsulation with mandatory 64-bit key for connecting L2 Ethernet attachment circuits identified by IPv6 addresses. The encapsulation is based on L2TPv3 over IP.

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 [RFC2119].

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 5, 2014.

Copyright Notice

Table of Contents

1. Introduction

L2TPv3, as defined in RFC3931 [RFC3931], provides a dynamic mechanism for tunneling Layer 2 (L2) "circuits" across a packet-oriented data network (e.g., over IP), with multiple attachment circuits multiplexed over a single pair of IP address endpoints (i.e. a tunnel) using the L2TPv3 session ID as a circuit discriminator.

Implementing L2TPv3 over IPv6 provides the opportunity to utilize unique IPv6 addresses to identify Ethernet attachment circuits directly, leveraging the key property that IPv6 offers, a vast number of unique IP addresses. In this case, processing of the L2TPv3 Session ID may be bypassed upon receipt as each tunnel has one and only one associated session. This local optimization does not hinder the ability to continue supporting the multiplexing of circuits via the Session ID on the same router for other L2TPv3 tunnels.

2. Static 1:1 Mapping Without a Control Plane

Static local configuration creates a one-to-one mapping between the access-side L2 attachment circuit and the IP address used in the network-side IPv6 encapsulation. The L2TPv3 Control Plane defined in RFC3931 [RFC3931] is not used.

The IPv6 L2TPv3 tunnel encapsulating device uniquely identifies each Ethernet L2 attachment connection by a port ID or a combination of port ID and VLAN ID(s) on the access side, and by an IPv6 address on the network side.

Any VLAN identifiers, S-VID, C-VID or tuple (S-VID, C-VID) are treated with local significance within the Ethernet L2 port and are not forwarded over the IPv6 L2TPv3 tunnel. IPv6 address is treated as the IPv6 L2TPv3 tunnel endpoint.

Certain deployment scenarios may require using a single IPv6 address to identify a tunnel endpoint for many IPv6 L2TPv3 tunnels. For such cases the tunnel encapsulating device identifies each tunnel by a unique combination of tunnel source and destination IPv6 addresses.

As mentioned above Session ID processing is not required as each keyed IPv6 tunnel has one and only one associated session. However for compatibility with existing RFC3931 [RFC3931] implementations, the packets need to be sent with Session ID. The router implementing L2TPv3 according to RFC3931 [RFC3931] can be configured with multiple L2TPv3 tunnels, with one session per tunnel, to interoperate with the router implementing the keyed IPv6 tunnel as specified by this document.

Note that a previous IETF draft [I.D.ietf-pppext-l2tphc] introduces the concept of an L2TP tunnel carrying a single session and hence not requiring session ID processing.

3. 64-bit Cookie

In line with RFC3931 [RFC3931], the key in the cookie field is used for additional tunnel endpoint context check. All packets MUST carry a 64-bit key in the L2TPv3 cookie field. The cookie MUST be 64-bits long in order to provide sufficient protection against spoofing and brute force blind insertion attacks.

In the absence of the L2TPv3 Control Plane, the L2TPv3 encapsulating router must be provided with local configuration of the 64-bit cookie for each local and remote IPv6 endpoint - note that cookies are asymmetric, so local and remote endpoints may send different cookie values. The value of the cookie must be able to be changed at any

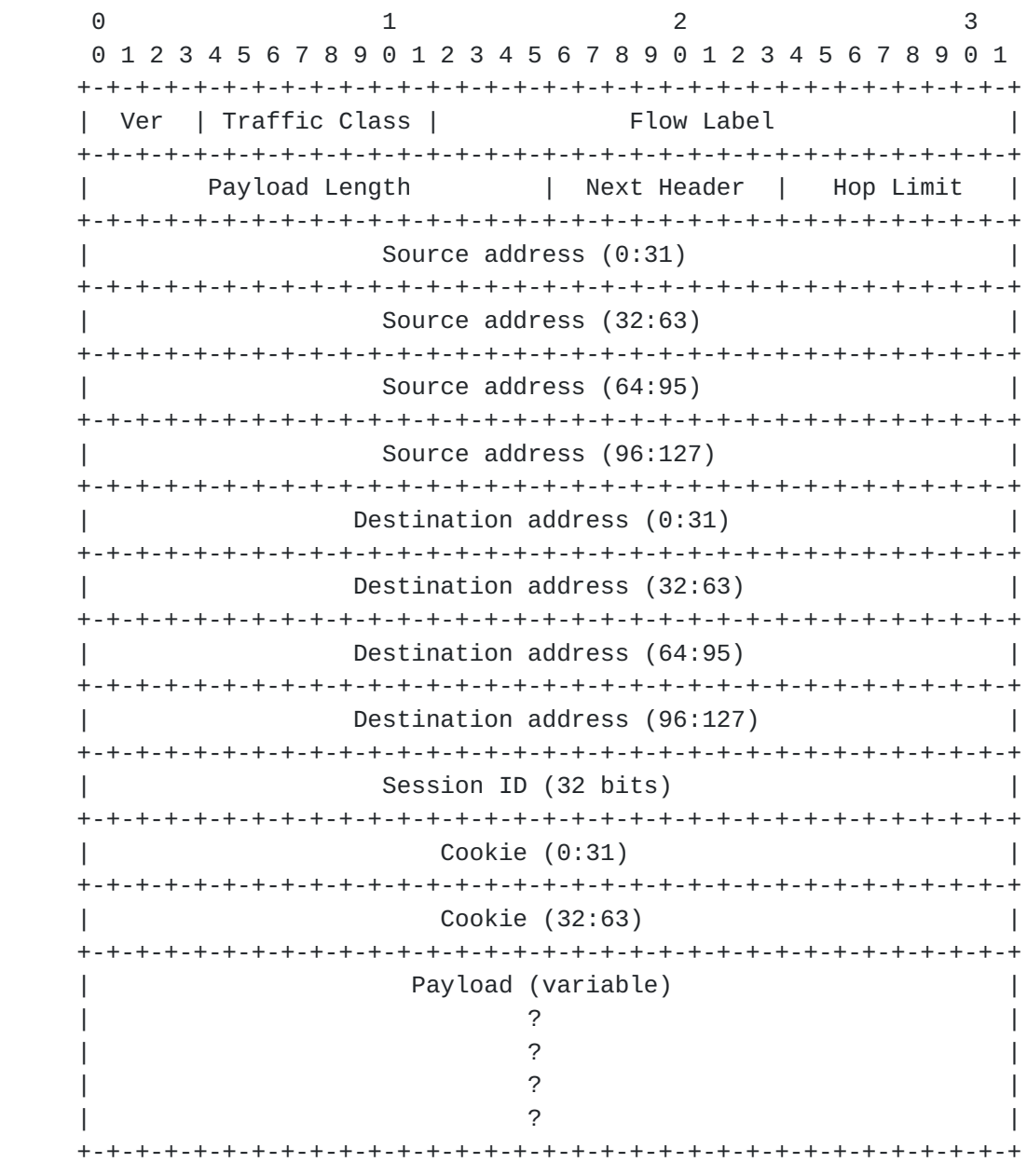
time in a manner that does not drop any legitimate tunneled packets - i.e. the receiver must be willing to accept both "old" and "new" cookie values during a change of cookie value.

4. Encapsulation

RFC4719 [RFC4719] describes encapsulation of Ethernet over L2TPv3. Paraphrasing from this document, the Ethernet frame, without the preamble or frame check sequence (FCS), is encapsulated in L2TPv3 and is sent as a single packet by the ingress router.

The s-tag (or in the multi-stack access case the s-tag and c-tag) SHOULD be removed before the packet is encapsulated.

The full encapsulation is as follows:



The combined IPv6 and L2TPv3 header contains the following fields:

- Ver. Set to 0x6 to indicate IPv6.

- Traffic Class. May be set by the ingress router to ensure correct PHB treatment by transit routers between the ingress and egress, and correct QoS disposition at the egress router.

- Flow Label. May be set by the ingress router to indicate a flow of packets from the client which may not be reordered by the

network (if there is a requirement for finer grained ECMP load balancing than per-circuit load balancing).

- Payload Length. Set to the length of the packet, excluding the IPv6 header (i.e. the length from the Session ID to the end of the packet).

- Next Header. Set to 0x73 to indicate that the next header is L2TPv3.

- Hop Limit. Set to 0xFF, and decremented by one by each router in the path to the egress router.

- Source Address. IPv6 source address for the tunnel. In the "Static 1:1" case the IPv6 source address may correspond to a port or VLAN being transported as an L2 circuit, or may be a loopback address terminating inside the router (e.g. if L2 circuits are being used within a multipoint VPN) or may be an anycast address terminating on a data center virtual machine.

- Destination Address. IPv6 destination address for the tunnel. As with the source address this may correspond to a port or VLAN being transported as an L2 circuit or may be a loopback or anycast address.

- Session ID. In the "Static 1:1 mapping" case described in Section 2, the IPv6 address resolves to an L2TPv3 session immediately, thus the Session ID may be ignored upon receipt. For compatibility with other tunnel termination platforms supporting only 2-stage resolution (IPv6 Address + Session ID), this specification recommends supporting explicit configuration of Session ID to any value other than zero. For cases where both tunnel endpoints support one-stage resolution (IPv6 Address only), this specification recommends setting the Session ID to all ones for easy identification in case of troubleshooting. The Session ID of zero MUST NOT be used, as it is reserved for use by L2TP control messages RFC3931 [RFC3931].

- Cookie. 64 bits, configured and described as in Section 3. All packets for a destined L2 Circuit (or L2TPv3 Session) must match the configured Cookie value or be discarded (see RFC3931 [RFC3931] for more details).

- Payload. The customer data, with s-tag or s-tag/c-tag removed. As noted above preamble and FCS are stripped before encapsulation. A new FCS will be added at each hop when the IP packet is transmitted.

5. Fragmentation and Reassembly

Using tunnel encapsulation, Ethernet L2 datagrams in IPv6 in this case, will reduce the effective MTU of the Ethernet L2 datagram.

The recommended solution to deal with this problem is for the network operator to increase the MTU size of all the links between the devices acting as IPv6 L2TPv3 tunnel endpoints to accommodate both the IPv6 L2TPv3 encapsulation header and the Ethernet L2 datagram without fragmenting the IPv6 packet.

If it is impossible to increase the link MTU across the network, the IPv6 L2TPv3 encapsulating device MUST perform fragmentation and reassembly if the outgoing link MTU cannot accommodate the extra IPv6 L2TPv3 header for specific Ethernet L2 payload. Fragmentation MUST happen after the encapsulation of the IPv6 L2TPv3 packet. Reassembly MUST happen before the decapsulation of the IPv6 L2TPv3 packet.

The proposed approach is in line with the DS-Lite specification RFC6333 [RFC6333].

6. OAM Considerations

OAM is an important consideration when providing circuit-oriented services such as those described in this document, and all the more so in the absence of a dedicated tunnel control plane, as OAM becomes the only way to detect failures in the tunnel overlay.

Note that in the context of keyed IP tunnels, failures in the IPv6 underlay network can be detected using the usual methods such as through the routing protocol.

Since keyed IP tunnels always carry an Ethernet payload, and since OAM at the tunnel layer is unable to detect failures in the Ethernet service processing at the ingress or egress router, or on the Ethernet attachment circuit between the router and the Ethernet client, this document recommends that Ethernet OAM as defined in IEEE 802.1ag [IEEE802.1ag] and/or ITU Y.1731 [Y.1731] is enabled for keyed IP tunnels. More specifically the following Connecitivity Fault Management (CFM) and/or Ethernet continuity check (ETH-CC) configurations are to be used in conjunction with keyed IPv6 tunnels:

- Connectivity verification between the tunnel endpoints across the tunnel - use an Up MEP located at the tunnel endpoint for transmitting the CFM PDUs towards, and receiving them from the direction of the tunnel.

- Connectivity verification from the tunnel endpoint across the local attachment circuit - use a Down MEP located at the tunnel endpoint for transmitting the CFM PDUs towards, and receiving them from the direction of the local attachment circuit.

- Intermediate connectivity verifcation - use a MIP located at the tunnel endpoint to generate CFM PDUs in response to received CFM PDUs.

In addition the Pseudowire Virtual Circuit Connectivity Verfiication (VCCV) RFC5085 [RFC5085] MAY be used.

7. IANA Considerations

None.

8. Security Considerations

Packet spoofing for any type of Virtual Private Network (VPN) tunneling protocol is of particular concern as insertion of carefully constructed rogue packets into the VPN transit network could result in a violation of VPN traffic separation, leaking data into a customer VPN. This is complicated by the fact that it may be particularly difficult for the operator of the VPN to even be aware that it has become a point of transit into or between customer VPNs.

Keyed IPv6 encapsulation provides traffic separation for its VPNs via use of separate 128-bit IPv6 addresses to identify the endpoints. The mandatory authentication key carried in the L2TPv3 cookie field, provides an additional check to ensure that an arriving packet is intended for the identified tunnel.

In the presence of a blind packet spoofing attack, the authentication key provides security against inadvertent leaking of frames into a customer VPN, like in case of L2TPv3 RFC3931 [RFC3931]. To illustrate the type of security that it is provided in this case, consider comparing the validation of a 64-bit Cookie in the L2TPv3 header to the admission of packets that match a given source and destination IP address pair. Both the source and destination IP address pair validation and Cookie validation consist of a fast check on cleartext header information on all arriving packets. However, since L2TPv3 uses its own value, it removes the requirement for one to maintain a list of (potentially several) permitted or denied IP addresses, and moreover, to guard knowledge of the permitted IP addresses from hackers who may obtain and spoof them. Further, it is far easier to change a compromised L2TPv3 Cookie than a compromised IP address," and a cryptographically random RFC4086 [RFC4086] value

is far less likely to be discovered by brute-force attacks compared to an IP address.

For protection against brute-force, blind, insertion attacks, a 64-bit Cookie MUST be used with all tunnels.

Note that the Cookie provides no protection against a sophisticated man-in-the-middle attacker who can sniff and correlate captured data between nodes for use in a coordinated attack.

The L2TPv3 64-bit cookie must not be regarded as a substitute for security such as that provided by IPsec when operating over an open or untrusted network where packets may be sniffed, decoded, and correlated for use in a coordinated attack.

9. Contributing Authors

Peter Weinberger
Cisco Systems

Email: peweinbe@cisco.com

Michael Lipman
Cisco Systems

Email: mlipman@cisco.com

Mark Townsley
Cisco Systems

Email: townsley@cisco.com

10. Acknowledgements

The authors would like to thank Carlos Pignataro for his suggestions and review.

11. References

11.1. Normative References

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

[RFC3931] Lau, J., Townsley, M., and I. Goyret, "Layer Two Tunneling Protocol - Version 3 (L2TPv3)", RFC 3931, March 2005.

[RFC4086] Eastlake, D., Schiller, J., and S. Crocker, "Randomness Requirements for Security", BCP 106, RFC 4086, June 2005.

[RFC4719] Aggarwal, R., Townsley, M., and M. Dos Santos, "Transport of Ethernet Frames over Layer 2 Tunneling Protocol Version 3 (L2TPv3)", RFC 4719, November 2006.

[RFC5085] Nadeau, T. and C. Pignataro, "Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires", RFC 5085, December 2007.

11.2. Informative References

[I.D.ietf-pppext-l2tphc]
Valencia, A., "L2TP Header Compression", December 1997.

[IEEE802.1ag]
IEEE, "IEEE Standard for Local and metropolitan area networks - Virtual Bridged Local Area Networks, Amendment 5: Connectivity Fault Managements", 2007.

[RFC6333] Durand, A., Droms, R., Woodyatt, J., and Y. Lee, "Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion", RFC 6333, August 2011.

[Y.1731] ITU, "ITU-T Recommendation G.8013/Y.1731 - OAM functions and mechanisms for Ethernet based networks", 2011.

Authors' Addresses

Maciek Konstantynowicz (editor)
Cisco Systems

Email: maciek@cisco.com

Giles Heron (editor)
Cisco Systems

Email: giheron@cisco.com

Rainer Schatzmayr
Deutsche Telekom AG

Email: rainer.schatzmayr@telekom.de

Wim Henderickx
Alcatel-Lucent, Inc.

Email: wim.henderickx@alcatel-lucent.com