Network Working Group
Request for Comments: 4817
Category: Standards Track

M. Townsley
C. Pignataro
S. Wainner
Cisco Systems
T. Seely
Sprint Nextel
J. Young

March 2007

Encapsulation of MPLS over Layer 2 Tunneling Protocol Version 3

Status of This Memo

Copyright Notice

Abstract

The Layer 2 Tunneling Protocol, Version 3 (L2TPv3) defines a protocol for tunneling a variety of payload types over IP networks. This document defines how to carry an MPLS label stack and its payload over the L2TPv3 data encapsulation. This enables an application that traditionally requires an MPLS-enabled core network, to utilize an L2TPv3 encapsulation over an IP network instead.

Table of Contents

1. Introduction

This document defines how to encapsulate an MPLS label stack and its payload inside the L2TPv3 tunnel payload. After defining the MPLS over L2TPv3 encapsulation procedure, other MPLS over IP encapsulation options, including IP, Generic Routing Encapsulation (GRE), and IPsec are discussed in context with MPLS over L2TPv3 in an Applicability section. This document only describes encapsulation and does not concern itself with all possible MPLS-based applications that may be enabled over L2TPv3.

1.1. Specification of Requirements

In this document, several words are used to signify the requirements of the specification. These words are often capitalized. 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].

2. MPLS over L2TPv3 Encoding

MPLS over L2TPv3 allows tunneling of an MPLS stack [RFC3032] and its payload over an IP network, utilizing the L2TPv3 encapsulation defined in [RFC3931]. The MPLS Label Stack and payload are carried in their entirety following IP (either IPv4 or IPv6) and L2TPv3 headers.

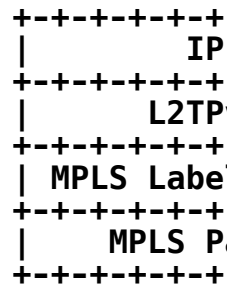

Figure 2.1 MPLS Packet over L2TPv3/IP

The L2TPv3 encapsulation carrying a single MPLS label stack entry is as follows:

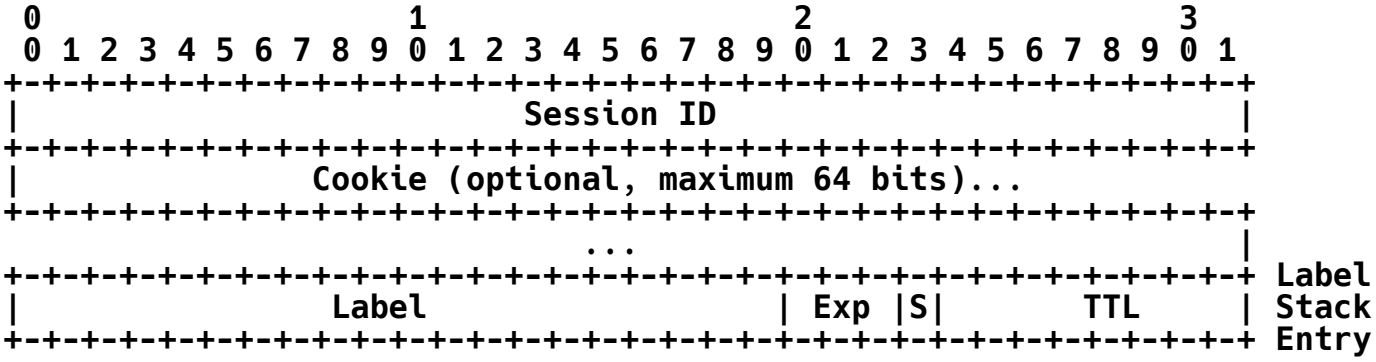

Figure 2.2 MPLS over L2TPv3 encapsulation

When encapsulating MPLS over L2TPv3, the L2TPv3 L2-Specific Sublayer MAY be present. It is generally not present; hence, it is not included in Figure 2.2. The L2TPv3 Session ID MUST be present. The Cookie MAY be present.

Session ID

The L2TPv3 Session ID is a 32-bit identifier field locally selected as a lookup key for the context of an L2TP Session. An L2TP Session contains necessary context for processing a received L2TP packet. At a minimum, such context contains whether the Cookie (see description below) is present, the value it was assigned, the presence and type of an L2TPv3 L2-Specific Sublayer, as well as what type of tunneled encapsulation follows (i.e., Frame Relay, Ethernet, MPLS, etc.)

Cookie

The L2TPv3 Cookie field contains a variable-length (maximum 64 bits), randomly assigned value. It is intended to provide an additional level of guarantee that a data packet has been directed to the proper L2TP session by the Session ID. While the Session ID may be encoded and assigned any value (perhaps optimizing for local lookup capabilities, redirection in a distributed forwarding architecture, etc.), the Cookie MUST be selected as a cryptographically random value [RFC4086], with the added restriction that it not be the same as a recently used value for a given Session ID. A well-chosen Cookie will prevent inadvertent misdirection of a stray packet containing a recently reused Session ID, a Session ID that is subject to packet corruption, and protection against some specific malicious packet insertion attacks, as described in more detail in Section 4 of this document.

Label Stack Entry

An MPLS label stack entry as defined in [RFC3032].

The optional L2-Specific Sublayer (defined in [RFC3931]) is generally not present for MPLS over L2TPv3.

Generic IP encapsulation procedures, such as fragmentation and MTU considerations, handling of Time to Live (TTL), EXP, and Differentiated Services Code Point (DSCP) bits, etc. are the same as the "Common Procedures" for IP encapsulation of MPLS defined in Section 5 of [RFC4023] and are not reiterated here.

3. Assigning the L2TPv3 Session ID and Cookie

Much like an MPLS label, the L2TPv3 Session ID and Cookie must be selected and exchanged between participating nodes before L2TPv3 can operate. These values may be configured manually, or distributed via a signaling protocol. This document concerns itself only with the encapsulation of MPLS over L2TPv3; thus, the particular method of assigning the Session ID and Cookie (if present) is out of scope.

4. Applicability

The methods defined in [RFC4023], [MPLS-IPSEC], and this document all describe methods for carrying MPLS over an IP network. Cases where MPLS over L2TPv3 is comparable to other alternatives are discussed in this section.

It is generally simpler to have one's border routers refuse to accept an MPLS packet than to configure a router to refuse to accept certain MPLS packets carried in IP or GRE to or from certain IP sources or destinations. Thus, the use of IP or GRE to carry MPLS packets increases the likelihood that an MPLS label-spoofing attack will succeed. L2TPv3 provides an additional level of protection against packet spoofing before allowing a packet to enter a Virtual Private Network (VPN) (much like IPsec provides an additional level of protection at a Provider Edge (PE) router rather than relying on Access Control List (ACL) filters). Checking the value of the L2TPv3 Cookie is similar to any sort of ACL that inspects the contents of a packet header, except that we give ourselves the luxury of "seeding" the L2TPv3 header with a value that is very difficult to spoof.

MPLS over L2TPv3 may be advantageous compared to [RFC4023], if:

- Two routers are already "adjacent" over an L2TPv3 tunnel established for some other reason, and wish to exchange MPLS packets over that adjacency.

- Implementation considerations dictate the use of MPLS over L2TPv3. For example, a hardware device may be able to take advantage of the L2TPv3 encapsulation for faster or distributed processing.

- Packet spoofing and insertion, service integrity and resource protection are of concern, especially given the fact that an IP tunnel potentially exposes the router to rogue or inappropriate IP packets from unknown or untrusted sources. IP Access Control Lists (ACLs) and numbering methods may be used to protect the PE routers from rogue IP sources, but may be subject to error and cumbersome to maintain at all edge points at all times. The L2TPv3 Cookie provides a simple means of validating the source of an L2TPv3 packet before allowing processing to continue. This validation offers an additional level of protection over and above IP ACLs, and a validation that the Session ID was not corrupted in transit or suffered an internal lookup error upon receipt and processing. If the Cookie value is assigned and distributed automatically, it is less subject to operator error, and if selected in a cryptographically random nature, less subject to blind guesses than source IP addresses (in the event that a hacker can insert packets within a core network).

(The first two of the above applicability statements were adopted from [RFC4023].)

In summary, L2TPv3 can provide a balance between the limited security against IP spoofing attacks offered by [RFC4023] vs. the greater security and associated operational and processing overhead offered

by [MPLS-IPSEC]. Further, MPLS over L2TPv3 may be faster in some hardware, particularly if that hardware is already optimized to classify incoming L2TPv3 packets carrying IP framed in a variety of ways. For example, IP encapsulated by High-Level Data Link Control (HDLC) or Frame Relay over L2TPv3 may be equivalent in complexity to IP encapsulated by MPLS over L2TPv3.

5. Congestion Considerations

This document specifies an encapsulation method for MPLS inside the L2TPv3 tunnel payload. MPLS can carry a number of different protocols as payloads. When an MPLS/L2TPv3 flow carries IP-based traffic, the aggregate traffic is assumed to be TCP friendly due to the congestion control mechanisms used by the payload traffic. Packet loss will trigger the necessary reduction in offered load, and no additional congestion avoidance action is necessary.

When an MPLS/L2TPv3 flow carries payload traffic that is not known to be TCP friendly and the flow runs across an unprovisioned path that could potentially become congested, the application that uses the encapsulation specified in this document MUST employ additional mechanisms to ensure that the offered load is reduced appropriately during periods of congestion. The MPLS/L2TPv3 flow should not exceed the average bandwidth that a TCP flow across the same network path and experiencing the same network conditions would achieve, measured on a reasonable timescale. This is not necessary in the case of an unprovisioned path through an over-provisioned network, where the potential for congestion is avoided through the over-provisioning of the network.

The comparison to TCP cannot be specified exactly but is intended as an "order-of-magnitude" comparison in timescale and throughput. The timescale on which TCP throughput is measured is the roundtrip time of the connection. In essence, this requirement states that it is not acceptable to deploy an application using the encapsulation specified in this document on the best-effort Internet, which consumes bandwidth arbitrarily and does not compete fairly with TCP within an order of magnitude. One method of determining an acceptable bandwidth is described in [RFC3448]. Other methods exist, but are outside the scope of this document.

6. Security Considerations

There are three main concerns when transporting MPLS-labeled traffic between PEs using IP tunnels. The first is the possibility that a third party may insert packets into a packet stream between two PEs. The second is that a third party might view the packet stream between two PEs. The third is that a third party may alter packets in a

stream between two PEs. The security requirements of the applications whose traffic is being sent through the tunnel characterizes how significant these issues are. Operators may use multiple methods to mitigate the risk, including access lists, authentication, encryption, and context validation. Operators should consider the cost to mitigate the risk.

Security is also discussed as part of the applicability discussion in Section 4 of this document.

6.1. In the Absence of IPsec

If the tunnels are not secured with IPsec, then some other method should be used to ensure that packets are decapsulated and processed by the tunnel tail only if those packets were encapsulated by the tunnel head. If the tunnel lies entirely within a single administrative domain, address filtering at the boundaries can be used to ensure that no packet with the IP source address of a tunnel endpoint or with the IP destination address of a tunnel endpoint can enter the domain from outside.

However, when the tunnel head and the tunnel tail are not in the same administrative domain, this may become difficult, and filtering based on the destination address can even become impossible if the packets must traverse the public Internet.

Sometimes, only source address filtering (but not destination address filtering) is done at the boundaries of an administrative domain. If this is the case, the filtering does not provide effective protection at all unless the decapsulator of MPLS over L2TPv3 validates the IP source address of the packet.

Additionally, the "Data Packet Spoofing" considerations in Section 8.2 of [RFC3931] and the "Context Validation" considerations in Section 6.2 of this document apply. Those two sections highlight the benefits of the L2TPv3 Cookie.

6.2. Context Validation

The L2TPv3 Cookie does not provide protection via encryption. However, when used with a sufficiently random, 64-bit value that is kept secret from an off-path attacker, the L2TPv3 Cookie may be used as a simple yet effective packet source authentication check which is quite resistant to brute force packet-spoofing attacks. It also alleviates the need to rely solely on filter lists based on a list of valid source IP addresses, and thwarts attacks that could benefit by spoofing a permitted source IP address. The L2TPv3 Cookie provides a means of validating the currently assigned Session ID on the packet

flow, providing context protection, and may be deemed complimentary to securing the tunnel utilizing IPsec. In the absence of cryptographic security on the data plane (such as that provided by IPsec), the L2TPv3 Cookie provides a simple method of validating the Session ID lookup performed on each L2TPv3 packet. If the Cookie is sufficiently random and remains unknown to an attacker (that is, the attacker has no way to predict Cookie values or monitor traffic between PEs), then the Cookie provides an additional measure of protection against malicious spoofed packets inserted at the PE over and above that of typical IP address and port ACLs.

6.3. Securing the Tunnel with IPsec

L2TPv3 tunnels may also be secured using IPsec, as specified in Section 4.1.3 of [RFC3931]. IPSec may provide authentication, privacy protection, integrity checking, and replay protection. These functions may be deemed necessary by the operator. When using IPsec, the tunnel head and the tunnel tail should be treated as the endpoints of a Security Association. A single IP address of the tunnel head is used as the source IP address, and a single IP address of the tunnel tail is used as the destination IP address. The means by which each node knows the proper address of the other is outside the scope of this document. However, if a control protocol is used to set up the tunnels, such control protocol MUST have an authentication mechanism, and this MUST be used when the tunnel is set up. If the tunnel is set up automatically as the result of, for example, information distributed by BGP, then the use of BGP's Message Digest 5 (MD5)-based authentication mechanism can serve this purpose.

The MPLS over L2TPv3 encapsulated packets should be considered as originating at the tunnel head and being destined for the tunnel tail; IPsec transport mode SHOULD thus be used.

Note that the tunnel tail and the tunnel head are Label Switched Path (LSP) adjacencies (for label distribution adjacencies, see [RFC3031]), which means that the topmost label of any packet sent through the tunnel must be one that was distributed by the tunnel tail to the tunnel head. The tunnel tail MUST know precisely which labels it has distributed to the tunnel heads of IPsec-secured tunnels. Labels in this set MUST NOT be distributed by the tunnel tail to any LSP adjacencies other than those that are tunnel heads of IPsec-secured tunnels. If an MPLS packet is received without an IPsec encapsulation, and if its topmost label is in this set, then the packet MUST be discarded.

Securing L2TPv3 using IPsec MUST provide authentication and integrity. (Note that the authentication and integrity provided will apply to the entire MPLS packet, including the MPLS label stack.)

Consequently, the implementation MUST support Encapsulating Security Payload (ESP) with null encryption. ESP with encryption MAY be supported if a source requires confidentiality. If ESP is used, the tunnel tail MUST check that the source IP address of any packet received on a given Security Association (SA) is the one expected.

Key distribution may be done either manually or automatically by means of the Internet Key Exchange (IKE) Protocol [RFC2409] or IKEv2 [RFC4306]. Manual keying MUST be supported. If automatic keying is implemented, IKE in main mode with preshared keys MUST be supported. A particular application may escalate this requirement and request implementation of automatic keying. Manual key distribution is much simpler, but also less scalable, than automatic key distribution. If replay protection is regarded as necessary for a particular tunnel, automatic key distribution should be configured.

7. Acknowledgements

Thanks to Robert Raszuk, Clarence Filsfils, and Eric Rosen for their review of this document. Some text was adopted from [RFC4023].

8. References

8.1. Normative References

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

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

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

[RFC4023] Worster, T., Rekhter, Y., and E. Rosen, "Encapsulating MPLS in IP or Generic Routing Encapsulation (GRE)", RFC 4023, March 2005.

[RFC4086] Eastlake, D., 3rd, Schiller, J., and S. Crocker, "Randomness Requirements for Security", BCP 106, RFC 4086, June 2005.

8.2. Informative References

[MPLS-IPSEC] Rosen, E., De Clercq, J., Paridaens, O., T'Joens, Y., and C. Sargor, "Architecture for the Use of PE-PE IPsec Tunnels in BGP/MPLS IP VPNs", Work in Progress, August 2005.

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, January 2001.

[RFC2409] Harkins, D. and D. Carrel, "The Internet Key Exchange (IKE)", RFC 2409, November 1998.

[RFC4306] Kaufman, C., "Internet Key Exchange (IKEv2) Protocol", RFC 4306, December 2005.

[RFC3448] Handley, M., Floyd, S., Padhye, J., and J. Widmer, "TCP Friendly Rate Control (TFRC): Protocol Specification", RFC 3448, January 2003.

Authors' Addresses

W. Mark Townsley
Cisco Systems
USA

Phone: +1-919-392-3741
EMail: mark@townsley.net

Carlos Pignataro
Cisco Systems
7025 Kit Creek Road
PO Box 14987
Research Triangle Park, NC 27709
USA

Phone: +1-919-392-7428
EMail: cpignata@cisco.com

Scott Wainner
Cisco Systems
13600 Dulles Technology Drive
Herndon, VA 20171
USA

EMail: swainner@cisco.com

Ted Seely
Sprint Nextel
12502 Sunrise Valley Drive
Reston, VA 20196
USA

Phone: +1-703-689-6425
EMail: tseely@sprint.net

Jeff Young

EMail: young@jsyoung.net

Full Copyright Statement

Intellectual Property

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.