

Network Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: April 11, 2015

D. Fan  
L. Xia  
Huawei  
Z. Cao  
China Mobile  
N. Kim  
KT

October 8, 2014

L2TP-VP: Layer Two Tunneling Protocol - Virtualization Profile  
draft-fan-l2tp-vp-02

## Abstract

This document describes Layer Two Tunneling Protocol (L2TP)'s virtualization profile (L2TP-VP), which reuses session header of L2TP data message to securely support overlay networks for multiple tenants, and simplifies tunnel setup by disabling all kinds of L2TP control messages.

## 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 April 11, 2015.

## Copyright Notice

Copyright (c) 2014 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

Internet-Draft

L2TP-VP

October 2014

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

<a href="#">1.</a>	Introduction . . . . .	<a href="#">2</a>
<a href="#">2.</a>	Terminology . . . . .	<a href="#">3</a>
<a href="#">3.</a>	L2TP-VP Frame Format . . . . .	<a href="#">3</a>
<a href="#">4.</a>	Control Plane Consideration . . . . .	<a href="#">6</a>
<a href="#">5.</a>	Data Plane Consideration . . . . .	<a href="#">6</a>
<a href="#">5.1.</a>	Address Learning . . . . .	<a href="#">6</a>
<a href="#">5.2.</a>	Forwarding . . . . .	<a href="#">6</a>
<a href="#">5.2.1.</a>	Unicast Traffic . . . . .	<a href="#">6</a>
<a href="#">5.2.2.</a>	Broadcast/Unknown/Multicast(BUM) Traffic . . . . .	<a href="#">6</a>
<a href="#">5.3.</a>	MTU Configuration . . . . .	<a href="#">7</a>
<a href="#">5.4.</a>	Qos Consideration . . . . .	<a href="#">7</a>
<a href="#">5.5.</a>	ECMP . . . . .	<a href="#">7</a>
<a href="#">6.</a>	Management Plane Consideration . . . . .	<a href="#">7</a>
<a href="#">7.</a>	Deployment Consideration . . . . .	<a href="#">8</a>
<a href="#">8.</a>	Security Considerations . . . . .	<a href="#">8</a>
<a href="#">9.</a>	IANA Considerations . . . . .	<a href="#">8</a>
<a href="#">10.</a>	References . . . . .	<a href="#">8</a>
<a href="#">10.1.</a>	Normaative References . . . . .	<a href="#">8</a>
<a href="#">10.2.</a>	Informative References . . . . .	<a href="#">8</a>
	Authors' Addresses . . . . .	<a href="#">8</a>

## [1.](#) Introduction

Traditional data centre network uses global VLAN ID to distinguish different tenants. Usually, a tenant consumes several VLAN IDs, for example, one for web server, one for application server and one for database server. When the number of tenants increases, the number of available VLAN IDs becomes insufficient.

When services provided by cloud via Internet becomes popular, a tenant's local area network needs to securely and smoothly reach anywhere via Internet if it wants. For example, a tenant can access its office IT services hosted in cloud data centers consisting of many geographically dispersed physical data centers. So, VPN access to cloud data centers becomes very important.

Layer Two Tunneling Protocol - Version 3 (L2TPv3) [[RFC3931](#)] is a mature and practical protocol that provides secure remote access service and layer 2 over IP service, but L2TPv3 also uses complicated control messages to setup tunnel. At the same time, L2TPv3 uses dynamical session id that is controlled by signaling mechanism and

Internet-Draft

L2TP-VP

October 2014

only has local significance. Currently, L2TPv3 is complex and does not support multiple tenants though it provides basic overlay functions.

This document will describe Layer Two Tunneling Protocol (L2TP)'s virtualization profile (L2TP-VP), which reuses session header of L2TP data message to securely support overlay networks for multiple tenants, and simplifies tunnel setup by disabling all kinds of L2TP control messages. Essentially, L2TP-VP defines a subset of L2TPv3 via fine and back-compatible reuse, and then extends L2TP's usage to network virtualization. L2TP is widely deployed and used whatever for operators' network or enterprises' network, L2TP-VP brings L2TP to the entire cloud network by further covering data center network.

The motivation of this draft is to propose an alternative L3-based overlay technology, besides the existed VxLAN [[RFC7348](#)], NVGRE [[NVGRE](#)], based on the following consideration:

- o L2TPv3 is a mature IP-based tunnelling technology that is widely supported and implemented on current operators' deployed networks. Directly reusing it can help operators to save their costs;
- o L2TPv3 inherent Cookie mechanism provides security protection against network attacks for tenant service;
- o L2TP-VP solution mainly focuses on the changing on network side, i.e. router or switch, to be transparent to client/server for alleviating their complexity and burden.

## [2.](#) Terminology

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](#)].

In this document, these words will appear with that interpretation

only when in ALL CAPS. Lower case uses of these words are not to be interpreted as carrying [RFC-2119](#) significance.

### 3. L2TP-VP Frame Format

L2TPv3 message format is specified in [[RFC3931](#)]. In order to support virtualization and reduce complexity from the control messages, two key fields are added into L2TP header to carry the original payload type and TNI (Tenant Network Identifier). The example of packet format for Ethernet encapsulation in L2TP-VP is shown in Figure 1.

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
Outer Ethernet Header:
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Outer Destination MAC Address                               |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Outer Destination MAC Address | Outer Source MAC Address |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Outer Source MAC Address                               |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Optional Ethertype=C-Tag 802.1Q | Outer VLAN Tag Information |
+-----+-----+-----+-----+-----+-----+-----+-----+
|           Ethertype 0x0800           |
+-----+-----+-----+-----+-----+-----+-----+-----+
Outer IPv4 Header:
+-----+-----+-----+-----+-----+-----+-----+-----+
| Version | IHL | Type of Service | Total Length |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Identification | Flags | Fragment Offset |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Time to Live | Protocol 115 | Header Checksum |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Outer Source Address                               |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Outer Destination Address                               |
+-----+-----+-----+-----+-----+-----+-----+-----+
L2TP-VP Header:
+-----+-----+-----+-----+-----+-----+-----+-----+
| M | Reserved#0 | Type |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

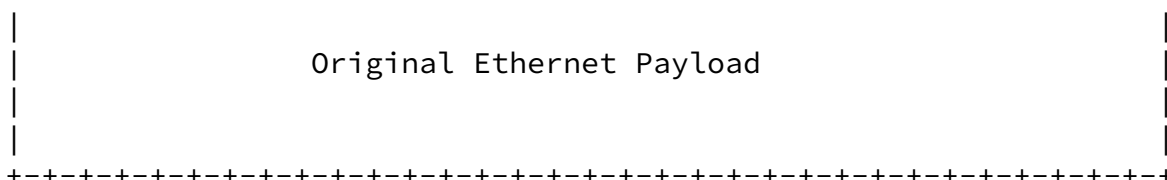
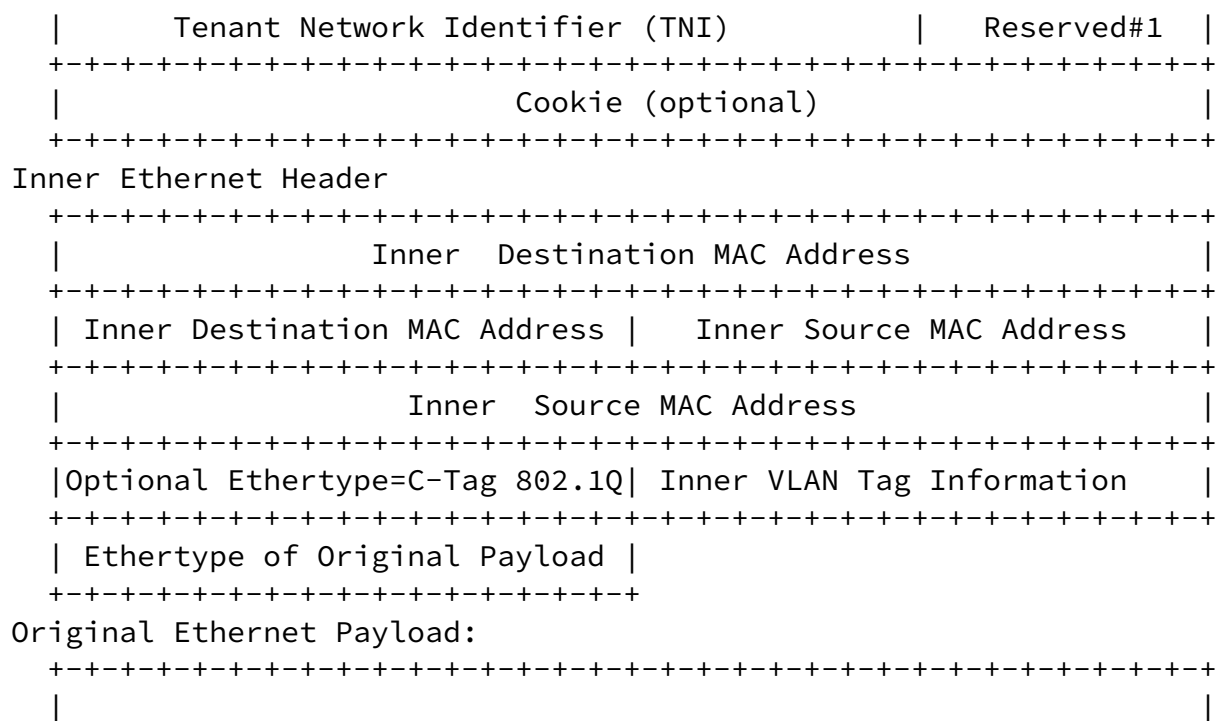


Figure 1 L2TP-VP Encapsulation Frame Format

The original Ethernet frame is encapsulated with L2TP-VP Header, Outer IP Header and Outer Ethernet Header.

#### L2TP-VP Header:

- o M bit: Modified Identifier. The M bit MUST be set to 1 to indicate the header is modified to L2TP-VP header format. If M=0, it indicates the header format following the L2TPv3 Session Header Over IP format and should to refer to [\[RFC3931\]](#) (Session ID is changed to a 31-bit field);
- o Type field : A 16-bit field is used to carry original payload type (e.g., frame type). Payload type can be Layer2 type such as ATM, FR, Ethernet, etc. It also can be Layer3 type such as IPv4 , IPv6

,etc. In Figure 1 the type of original packet is Ethernet;

- o TNI field : A 24-bit field allows up to 16 million tenants in the same management domain. The packets with different TNI will be isolated logically;
- o Cookie field : The optional Cookie field inherits all the functions from cookie field in [\[RFC3931\]](#) . It is used to check the association of a received data message with TNI. Only need to change its length to be 32-bit.

Outer IP Header: Both IPv4 and IPv6 can be used as encapsulation IP header. Figure 1 shows an example of IPv4. The source IP address is filled with IP address of L2TP-VP endpoint which encapsulates the original packet with L2TP-VP frame format. The destination IP address is unicast address obtained by lookup of address table. Also it may be a multicast address representing this packet may be used for address learning.

Outer Ethernet Header: The destination MAC address in Figure 1 may be the address of next hop device. The Optional Vlan Tag may be used to limit the area of the broadcast.

## [4.](#) Control Plane Consideration

In order to reduce complexity coming from control messages, there is no separate control plane in L2TP-VP. All kinds of control messages defined in [\[RFC3931\]](#) are disabled. All tunnel endpoints are expected to be configured by management plane(e.g., OSS).

## [5.](#) Data Plane Consideration

### [5.1.](#) Address Learning

For the E2E link and tunnel setup of L2TP-VP overlay network, the forwarding information including tenant systems' address, and its associated L2TP-VP endpoint address and TNI should be populated in the network. There are several options to support address learning:

- o Through the management plane, L2TP-VP endpoints will be configured part or all of the address table;
- o L2TP-VP endpoints directly acquire the forwarding information through data plane by flooding mechanism;
- o L2TP-VP endpoints join the multicast group and populate the forwarding information to the other endpoints in the same virtual network by the multicast tree.

## [5.2.](#) Forwarding

### [5.2.1.](#) Unicast Traffic

Ingress L2TP-VP endpoint firstly gets the destination address from the unicast traffic, then obtains IP address of the egress endpoint and the TNI by lookup of address table, at last encapsulates the original packet in L2TP-VP frame format. The source IP address in outer IP header is filled with its own IP address and the destination IP address is filled with egress endpoint's IP address.

### [5.2.2.](#) Broadcast/Unknown/Multicast(BUM) Traffic

There are several proven methods to process BUM traffic.

One method needs the multicast support of underlay network. All BUM traffic originating from within a TNI is terminated by the L2TP-VP endpoint, then encapsulated and sent to the assigned multicast address. The binding relation of the TNI and the multicast address of underlay network can be configured by the management plane.

Another method is ingress replication. One BUM frame in a TNI can be replicated to multiple unicast frames which will be sent to all the egress L2TP-VP endpoints in the same TNI.

## [5.3.](#) MTU Configuration

L2TP-VP overlay header can cause the MTU of the path to the egress tunnel endpoint to be exceeded. Here lists some solutions:

- o Modifying the MTU support configuration in the network devices, including L2TP-VP endpoints and other network devices which will transmit the encapsulation packets;
- o Classical ICMP-based MTU Path Discovery [[RFC1191](#)] [[RFC1981](#)] or Extended MTU Path Discovery techniques such as defined in [[RFC4821](#)].

#### [5.4.](#) Qos Consideration

QoS of underlay network can be provided without problem due to the fact that it's an IP network.

QoS of the overlay network may need to support the mapping of CoS marking between different network layers (e.g., Tenant Systems, Overlays, and/or Underlay) in L2TP-VP endpoints, for enabling each networking layer to independently enforce its own CoS policies.

TS's QoS fields (e.g. IP DSCP and/or Ethernet 802.1p) and policies can be defined to indicate application level CoS requirements. L2TP-VP endpoint can use the new service CoS fields in the overlay header to indicate the proper service CoS to be applied across the overlay network. This field can be mapped from the TS's QoS fields or other mechanism (e.g. DPI).

#### [5.5.](#) ECMP

Because the outer header is standard IP header, the L2TP-VP endpoint SHOULD provide ECMP. Basically the L2TP-VP endpoint uses a hash of various fields of the outer Ethernet header and outer IP header, furthermore it can use the fields of L2TP-VP header or even inner original packet. And the endpoint can select different fields for hash according to the requirement.

### [6.](#) Management Plane Consideration

Management plane is needed to configure access type, TNI, QoS, Cookie, etc. In some scenarios, management plane should support to configure the forwarding information or policies for data plane and

control plane , such as routing table, address table, etc.



Management plane can be OSS or SDN controller.

## 7. Deployment Consideration

TBD.

## 8. Security Considerations

Like L2TPv3, L2TP-VP continues to adopt Cookie Field as an additional check to the received packet. A 32-bit random field is difficult to be cracked so that it can afford protection against brute-force, blind and insertion attacks.

When the network is open network and someone can sniff the whole traffic through the network, it will need other security measures. Traditional security mechanisms based on IP technique will provide authentication/encryption function, such as IPSec.

## 9. IANA Considerations

TBD.

## 10. References

### 10.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", March 1997.
- [RFC3931] Lau, J., "Layer Two Tunneling Protocol - Version 3 (L2TPv3)", March 2005.
- [RFC7348] Mahalingam, M., "Virtual eXtensible Local Area Network (VXLAN): A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks", August 2014.

### 10.2. Informative References

- [NVGRE] Sridharan, M., "NVGRE: Network Virtualization using Generic Routing Encapsulation", ID [draft-sridharan-virtualization-nvgre-06](#), October 2014.

Authors' Addresses

Duoliang Fan  
Huawei  
101 Software Avenue, Yuhua District  
Nanjing, Jiangsu 210012  
China

Email: fanduoliang@huawei.com

Liang Xia  
Huawei  
101 Software Avenue, Yuhua District  
Nanjing, Jiangsu 210012  
China

Email: frank.xialiang@huawei.com

Zhen Cao  
China Mobile  
Xuanwumenxi Ave. No.32 , Xicheng District  
Beijing 100053  
China

Email: zehn.cao@gmail.com, caozhen@chinamobile.com

Namgon Kim  
KT  
463-1 Jeonmin-Dong, Yuseoung-Gu Daejeon, 305-811  
Korea

Email: ng.kim@kt.com

