Network Working Group
Internet Draft
Document: draft-ietf-l2tpext-mcast-05.txt
Category: Experimental

G. Bourdon
France Telecom

October 2004

Extensions to support efficient carrying of multicast traffic in Layer-2 Tunneling Protocol (L2TP) <draft-ietf-l2tpext-mcast-05.txt>

Status of this Memo

Abstract

The Layer Two Tunneling Protocol (L2TP) provides a standard method for tunneling PPP packets. This document describes an extension to L2TP, in order to have an efficient use of L2TP tunnels within the context of deploying multicast services whose data will have to be conveyed by such tunnels.

Table of Contents

1. Introduction

The deployment of IP multicast-based services may have to deal with L2TP tunnel engineering. From this perspective, the forwarding of multicast data within L2TP sessions may impact the throughput of L2TP tunnels because the same traffic may be sent multiple times within the same L2TP tunnel, but in different sessions. This proposal aims to reduce this impact by applying replication mechanism of multicast traffic only when necessary.
The solution described herein provides a mechanism for transmitting multicast data only once for all the L2TP sessions that have been established in a tunnel, each multicast flow having a dedicated L2TP session.

Within the context of deploying IP multicast-based services, it is assumed that the routers of the IP network that embed a L2TP Network Server (LNS) capability may be involved in the forwarding of

multicast data, towards users who access the network through an L2TP tunnel. Then the LNS is in charge of replicating the multicast data for each L2TP session that is used by a receiver who has requested a multicast flow. The solution described here gives the ability for a LNS to send multicast data only once and let the L2TP Access Concentrator (LAC) perform the traffic replication. By doing so, it is expected to spare transmission resources in the core network that supports L2TP tunnels. This multicast extension to L2TP is designed so that it does not affect the behavior of L2TP equipment under normal conditions. A solution to carry multicast data only once in a L2TP tunnel is interesting for service providers since edge devices are aggregating more and more users. This is particularly true for operators who are deploying xDSL (Digital Subscriber Line) services and cable infrastructures. Therefore, L2TP tunnels that may be supported by the network will have to carry multiple redundant multicast data more often. The solution described in this document applies to downstream traffic exclusively, i.e. data coming from the LNS towards end-users connected to the LAC. This downstream multicast traffic is not framed by the LNS but by the LAC, thus ensuring compatibility for all users in a common tunnel whatever the framing scheme.

1.1. 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 [RFC2119].

1.2. Terminology

Unicast session

This term refers to the definition of "Session", as it is described in the terminology section of [RFC2661]. (Also: L2TP unicast session)

Multicast session

This term refers to a connection between the LAC and the LNS. Additional Control Messages and Attribute-Value-Pairs (AVPs) are defined in this document to open and maintain this connection for the particular purpose of multicast traffic transportation. This connection between the LAC and the LNS is intended to convey multicast traffic only. (Also: L2TP multicast session)

Session

This term is used when there is no need to dissociate multicast from unicast sessions, and thus designates both. (Also: L2TP session)

Designates a Multicast Interior Gateway Protocol.

Multicast flow

Designates datagrams sent to a group from a set of sources for which multicast reception is desired.

GMP

Group Management Protocol, such as:
- IGMPv1 ([RFC1112])
- IGMPv2 ([RFC2236])
- MLD ([RFC2710], [RFC3590])

SFGMP

Source Filtering Group Management Protocol such as:
- IGMPv3 ([RFC3376])
- MLDv2 ([RFC3810])

2. Motivation for a session-based solution

Multicast data have to be seen as a singular flow that may be conveyed into all the L2TP sessions that have been established in a tunnel. It means that a given L2TP session can be dedicated for the forwarding of a multicast flow that will be forwarded to multiple receivers, including those that can be reached by one or several of these L2TP sessions. A session carrying IP multicast data is independent from the underlying framing scheme and is therefore compatible with any new framing scheme that may be supported by the L2TP protocol.

Using a single L2TP session per multicast flow is motivated by the following arguments:

- The administrator of the LNS is presumably in charge of the IP multicast-based services and the related engineering aspects. As such, he must be capable of filtering multicast traffic on a multicast source basis, on a multicast group basis, and/or on a user basis (users who access the network using a L2TP session that terminates in this LNS).
- Having a L2TP session dedicated for a multicast flow gives the ability to enforce specific policies for multicast traffic. For instance, it is possible to change the priority treatment for multicast packets against unicast packets.
- It is not always acceptable nor possible to have multicast forwarding performed within the network between the LAC and the LNS. Having the multicast traffic conveyed within a L2TP tunnel ensures a

multicast service between the LNS and end-users, alleviating the need for activating multicast capabilities in the underlying network.

3. Control Connection establishment

3.1. Negotiation phase

The multicast extension capability is negotiated between the LAC and the LNS during the control connection establishment phase. However, establishment procedures defined in [RFC2661] remain unchanged. A LAC indicates its multicast extension capability by using a new AVP, the "Multicast Capability" AVP. There is no explicit acknowledgement sent by the LNS during the control connection establishment phase. Instead, the LNS is allowed to use multicast extension messages to open and maintain multicast session(s) (Section 5).

3.2. Multicast Capability AVP (SCCRQ, SCCRP)

In order to inform the LNS that a LAC has the ability to handle multicast sessions, the LAC sends a Multicast Capability AVP during the control connection establishment phase.
This AVP is sent either in a SCCRQ or a SCCRP control message by the LAC towards the LNS.

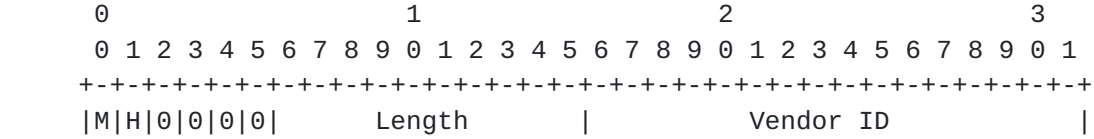
Upon receipt of the Multicast Capability AVP, a LNS may adopt two distinct behaviors:

1) The LNS does not implement the L2TP multicast extension: any multicast-related information (including the Multicast Capability AVP) will be silently ignored by the LNS.
2) The LNS implements L2TP multicast extensions, and therefore supports the Multicast Capability AVP: the LNS is allowed to send L2TP specific commands for conveying multicast traffic towards the LAC.

The multicast capability exclusively refers to the tunnel for which the AVP has been received during control connection establishment phase. It SHOULD be possible for a LNS administrator to shut down L2TP multicast extension features towards one or a set of LAC(s). In this case, the LNS behavior is similar to 1).

The AVP has the following format:

Vendor ID = 0
Attribute = TBA1 (16 bits) (Note: to be assigned by IANA)



```
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |              TBA1                |
```

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

The M-bit MUST be set to 0, the AVP MAY be hidden (H-bit set to 0 or 1).

The length of this AVP is 6 octets.

4. L2TP multicast session establishment decision

4.1. Multicast states in LNS

The router that embeds the LNS feature MUST support at least one Group Management Protocol (GMP) such as:

- IGMPv1
- IGMPv2
- MLD

or a Source Filtering Group Management Protocol (SFGMP) such as:

- IGMPv3
- MLDv2

The LAC does not have any group management activity: GMP or SFGMP processing is performed by the LNS. The LAC is a layer-2 equipment, and is not supposed to track GMP or SFGMP messages between the receivers and the LNS in this context.
The LNS MUST always be at the origin of the creation of a multicast L2TP session dedicated for the forwarding of IP multicast datagrams destined to a multicast group.
The LNS acts as a GMP or SFGMP Querier for every logical interface associated to a L2TP session.

As a multicast router, the equipment that embeds the LNS function will keep state per group per attached network (i.e. per L2TP session). The LNS-capable equipment activating multicast extensions for L2TP will have to classify and analyze GMP and SFGMP states in order to create L2TP multicast session(s) within the appropriate L2TP tunnel(s). This is performed in three steps:

1- The LNS has to compute group states for each L2TP tunnel, using group states recorded for each L2TP session of the tunnel. Group state determination for L2TP tunnels is discussed in section 4.2. For each L2TP tunnel, the result of this computation will issue a list of states of the form (group, filter-mode, source-list):

- group: denotes the multicast group
- filter-mode: either INCLUDE or EXCLUDE, as defined in [RFC3376]
- source-list: list of IP unicast addresses from which multicast reception is desired or not, depending on the filter-mode.

2- According to each group state, the LNS will create one or multiple replication contexts, depending on the filter-mode for the considered group and the local policy configured in the LNS.
For groups in INCLUDE mode, the LNS SHOULD implement two different policies:

- One session per (source, group) pair: the LNS creates one replication context per (source, group) pair.

or

- One session per group: the LNS creates one replication context per (source-list, group) pair.

For groups in EXCLUDE mode, the LNS will create one replication context per (list of sources excluded by *all* the receivers, group). The list of sources represents the intersection of the sets, not the union.

3- For each replication context, the LNS will create one L2TP multicast session (if threshold conditions are met, see Section 4.3) and its associated Outgoing Session List (OSL). The OSL lists L2TP sessions that requested the multicast flow corresponding to the group and the associated source-filtering properties. There is one OSL per replication context, i.e. per L2TP multicast session.

For a group member running a SFGMP, it is therefore possible to receive multicast traffic from sources that have been explicitly excluded in its SFGMP membership report if other group members in the same L2TP tunnel wish to receive packets from these sources. This behavior is comparable to the case where group members are connected to the same multi-access network. When a group is in EXCLUDE mode or in INCLUDE mode with a policy allowing one session per (group, source-list), sharing the same L2TP tunnel is equivalent to be connected to the same multi-access network in term of multicast traffic received. For groups in INCLUDE mode with a policy allowing one L2TP multicast session per (source, group), the behavior is slightly improved because it prevents group members to receive traffic from non-requested sources. On the other hand, this policy potentially increases the number of L2TP multicast sessions to establish and maintain. Examples are provided in Appendix A.

In order for the LAC to forward the multicast traffic received through the L2TP multicast session to group members, the LNS sends the OSL to the LAC for the related multicast session (see Section 6).

4.2. Group state determination

Source Filtering Group Management Protocols require querier routers to keep a filter-mode per group per attached network, to condense the total desired reception state of a group to a minimum set such that

all systems' memberships are satisfied.

Within the context of L2TP, each L2TP session has to be considered as an attached network by GMP and SFGMP protocols. When activating L2TP multicast extension, each L2TP Control Connection has to be considered as a pseudo attached network as well in order to condense group membership reports for every L2TP session in the tunnel.

Therefore, a list of group states is maintained for each L2TP Control Connection into which the membership information of each of its L2TP sessions is merged. This list of group states is a set of membership records of the form (group, filter-mode, source-list).
Each group state represents the result of a merging process applied to subscriptions on L2TP sessions of a Control Connection for a considered group. This merging process is performed in three steps:

1- Conversion of any GMP subscription into SFGMP subscription (IGMPv1/v2 to IGMPv3, MLDv1 to MLDv2);
2- Removal of subscription timers and, if filter-mode is EXCLUDE, sources with source timer > 0;
3- Then, resulting subscription are merged using merging rules described in SFGMP specifications ([RFC3376] Section 3.2, [RFC3810] Section 4.2).

This process is also described in [PROXY]. Examples of group state determination are provided in Appendix A.

4.3. Triggering

The rules to be enforced by the LNS so as to decide when to open a dedicated L2TP multicast session for a multicast group SHOULD be configurable by the LNS administrator. This would typically happen whenever a threshold of MULTICAST_SESSION_THRESHOLD receivers/sessions referenced in a replication context is reached. This threshold value SHOULD be valued at 2 by default, if we consider that it is worth opening a dedicated L2TP multicast session for two group members sharing the same desired reception state (which means that two L2TP unicast sessions are concerned). In this case, the OSL will reference two distinct L2TP sessions.

The actual reception by the LNS of multicast traffic requested by end-users can also be taken into account to decide whether the associated L2TP multicast session has to be opened or not.

Whenever an OSL gets empty, the LNS MUST stop sending multicast traffic over the corresponding L2TP multicast session. Then the L2TP multicast session MUST be torn down as described in Section 7 of this document.

Filter-mode changes for a group can also trigger the opening or the termination of L2TP multicast session(s):

a- From INCLUDE mode to EXCLUDE mode

When a group state filter-mode switches from INCLUDE to EXCLUDE, only one replication context (and its associated L2TP multicast session) issued from this group state can exist (see Section 4.1). The LNS SHOULD keep one replication context previously created for this group state and has to update it with:

-a new source-list that has to be excluded from forwarding

-a new OSL

The LNS MUST send an OSL update to the LAC to reflect L2TP sessions list changes (section 6.2), whenever appropriate. The unused L2TP multicast sessions that correspond to previously created replication contexts for the group SHOULD be terminated, either actively or passively by emptying their corresponding OSLs.
The remaining L2TP multicast session MAY also be terminated if the number of receivers is below a predefined threshold (see Section 7).
To limit the duration of temporary packet loss or duplicates to receivers, the LNS has to minimize delay between OSL updates messages sent to the LAC. Therefore, one can assume that terminating a multicast session passively gives the smoothest transition.

b- From EXCLUDE mode to INCLUDE mode

When a group state filter-mode switches from EXCLUDE to INCLUDE, multiple replication contexts issued by this group state may be created (see Section 4.1). The LNS SHOULD keep the replication context previously created for this group state and has to update it accordingly with the following information:

-a new list of sources that has to be forwarded, this list having only one record if there is one replication context per (group, source)

-a new OSL

The LNS MUST send an OSL update to the LAC to reflect L2TP sessions list changes, whenever appropriate. If the LNS is configured to create one replication context per (group, source), L2TP multicast sessions will be opened in addition to the existing one, depending on the number of sources for the group.
If new L2TP multicast sessions need to be opened, the LNS SHOULD wait until these multicast sessions are established before updating the OSL of the original multicast session. To limit the duration of temporary packet loss or duplicates to receivers, the LNS has to minimize delay between OSL updates messages sent to the LAC.

4.4. Multicast traffic sent from group members

The present document proposes a solution to enhance the forwarding of downstream multicast traffic exclusively, i.e. data coming from the LNS towards end-users connected to the LAC.

In the case where a group member that uses a L2TP session is also a multicast source for traffic conveyed in a multicast session,

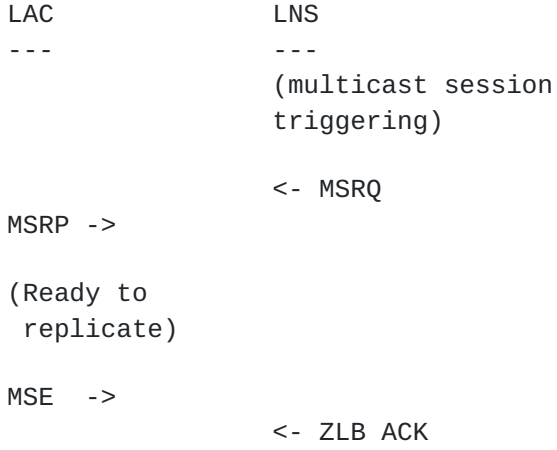
datagrams may be sent back to the source. To prevent this behavior, two options can be used in the LNS:

1- Disable the multicast packets forwarding capability, for those multicast datagrams sent by users connected to the network by the means of a L2TP tunnel. Protocols using well-known multicast addresses MUST NOT be impacted.

2- Exclude from the OSL the L2TP session used by a group member that sends packets matching the replication context of this OSL. Therefore, the corresponding multicast flow is sent by the LNS over the user L2TP unicast session, using standard multicast forwarding rules.

5. L2TP multicast session opening process

The opening of an L2TP multicast session is initiated by the LNS. A three-message exchange is utilized to set up the session. Following is a typical sequence of events:



ZLB ACK is sent if there are no further messages waiting in queue for that peer.

5.1. Multicast-Session-Request (MSRQ)

Multicast-Session-Request (MSRQ) is a control message sent by the LNS to the LAC to indicate that a multicast session can be created. The LNS initiates this message according to the rules mentioned in section 4.2. It is the first in a three-message exchange used for establishing a multicast session within a L2TP tunnel.

A LNS MUST NOT send a MSRQ control message if the remote LAC did not open the L2TP tunnel with the Multicast Capability AVP. The LAC MUST ignore MSRQ control messages sent in a L2TP tunnel, if the L2TP

tunnel was not opened with control messages including a Multicast Capability AVP.

The following AVPs MUST be present in MSRQ:

- Message Type
- Assigned Session ID

The following AVP MAY be present in MSRQ:

- Random Vector
- Maximum BPS

The Maximum BPS value is set by the LNS administrator. However, this value should be chosen in accordance with the line capabilities of the end-users. The Maximum BPS value SHOULD NOT be higher than the highest speed connection for all end-users within the L2TP tunnel.

The associated Message Type AVP is encoded with the values:

- Vendor ID = 0
- Attribute Type = 0
- Attribute Value = TBA2 (16 bits) (Note: to be assigned by IANA)

The M-bit MUST be set to 0, the H-bit MUST be set to 0.

5.2. Multicast-Session-Response (MSRP)

Multicast-Session-Response (MSRP) is a control message sent by the LAC to the LNS in response to a received MSRQ message. It is the second in a three-message exchange used for establishing a multicast session within a L2TP tunnel.

MSRP is used to indicate that the MSRQ was successful and the LAC will attempt to reserve appropriate resources to perform multicast replication for unicast sessions managed in the pertaining control connection.

The following AVPs MUST be present in MSRP:

- Message Type
- Assigned Session ID

The following AVP MAY be present in MSRP:

- Random Vector

The associated Message Type AVP is encoded with the values:

- Vendor ID = 0

Attribute Type = 0
Attribute Value = TBA3 (16 bits) (Note: to be assigned by IANA)

The M-bit MUST be set to 0, the H-bit MUST be set to 0.

5.3. Multicast-Session-Established (MSE)

Multicast-Session-Established (MSE) is a control message sent by the LAC to the LNS to indicate that the LAC is ready to receive necessary multicast information (Section 6) for the group using the newly created multicast session. It is the third message in the three-message sequence used for establishing a multicast session within a L2TP tunnel.

The following AVP MUST be present in MSE:

Message Type

The following AVP MAY be present in MSE:

Sequencing Required

Sequencing will occur only from the LNS to the LAC since a multicast session is only used to forward multicast traffic downstream.

The associated Message Type AVP is encoded with the values:

Vendor ID = 0
Attribute Type = 0
Attribute Value = TBA4 (16 bits) (Note: to be assigned by IANA)

The M-bit MUST be set to 0, the H-bit MUST be set to 0.

6. Session maintenance and management

Once the multicast session is established, the LAC has to be informed of the L2TP unicast sessions interested in receiving the traffic from the newly-created multicast session, as well as a related optional priority parameter defined in Section 6.3. To achieve this, a new control message type is defined: Multicast-Session-Information (MSI).

6.1. Multicast-Session-Information (MSI)

Multicast-Session-Information (MSI) control messages carry AVPs to keep the OSL synchronized between the LNS and the LAC, and to set the optional priority parameter for multicast traffic versus unicast traffic. MSI may be extended to update the multicast session with additional parameters, as needed.

Each MSI message is specific to a particular multicast session.

Therefore, the control message MUST use the assigned session ID associated to the multicast session (assigned by the LAC), except for

the case mentioned in 6.3.2.

The associated Message Type AVP is encoded with the values:

Vendor ID = 0
Attribute Type = 0
Attribute Value = TBA5 (16 bits) (Note: to be assigned by IANA)

The M-bit MUST be set to 0, the H-bit MUST be set to 0.

The following AVPs MUST be present in MSI:

Message Type

The following AVP MAY be present in MSI:

Random Vector
New Outgoing Sessions
New Outgoing Sessions Acknowledgement
Withdraw Outgoing Sessions
Multicast Packets Priority

New Outgoing Sessions, New Outgoing Sessions Acknowledgement, Withdraw Outgoing Sessions and Multicast Packets Priority are new AVPs defined in sections 6.2 and 6.3.

6.2. Outgoing Sessions List updates

Whenever a change occurs in the Outgoing Sessions List, the LNS MUST inform the LAC of that change. The OSL is built upon subscription reports recorded by GMP or SFGMP processes running in the LNS (Section 4.1).
The LAC maintains an OSL as a local table transmitted by the LNS. As for the LNS, the LAC has to maintain an OSL for each L2TP multicast session within a L2TP tunnel. To update the LAC OSL, the LNS sends a New Outgoing Sessions AVP for additional(s) session(s), or sends a Withdraw Outgoing Sessions AVP to remove session(s). All sessions mentioned in these AVPs MUST be added or removed by the LAC from the relevant OSL. The Outgoing Sessions List is identified by the tunnel ID and the multicast session ID the updating AVP is referring to.
To update the OSL, the following AVPs are used:

Additional session(s): New Outgoing Sessions AVP
Session(s) removal: Withdraw Outgoing Sessions AVP

These new AVPs MUST be sent in a MSI message.

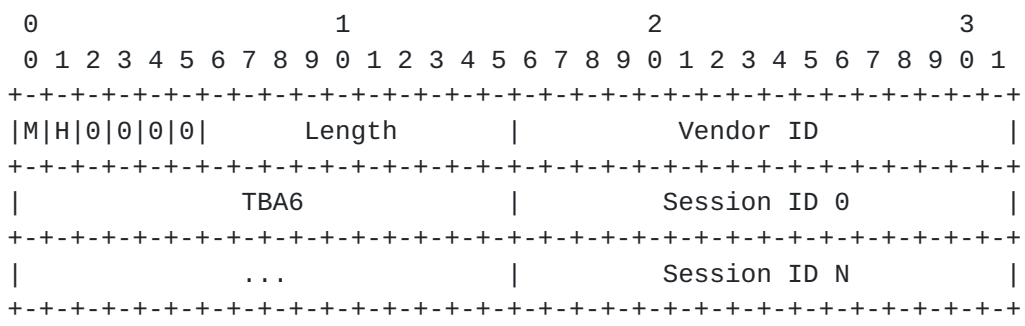
6.2.1. New Outgoing Sessions AVP (MSI)

The New Outgoing Sessions AVP can only be carried within a MSI

message type. This AVP piggybacks every Session ID to which the multicast traffic has to be forwarded.

The AVP has the following format:

Vendor ID = 0
Attribute = TBA6 (16 bits) (Note: to be assigned by IANA)



There can be from 1 to N Session IDs present in the New Outgoing Sessions AVP (considering the maximum value of the Length field). This AVP must be placed in a MSI message and sent after the establishment of the multicast session to indicate the LAC what are the initial outgoing sessions, and at any time when one or more outgoing sessions appear during the multicast session lifetime. Upon reception of this AVP, the LAC sends a New Outgoing Sessions Acknowledgment AVP to the LNS to notify that the LAC is ready to replicate the multicast traffic towards the indicated sessions.

Usage of this AVP is incremental: only new outgoing sessions have to be listed in the AVP.

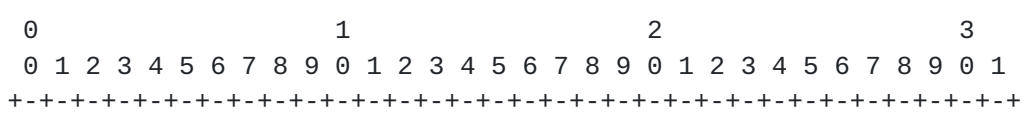
The M-bit MUST be set to 1, the AVP MAY be hidden (H-bit set to 0 or 1).

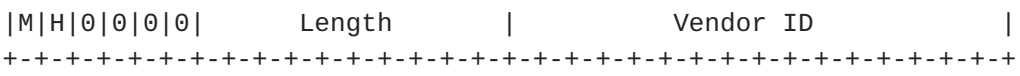
6.2.2. New Outgoing Sessions Acknowledgement AVP (MSI)

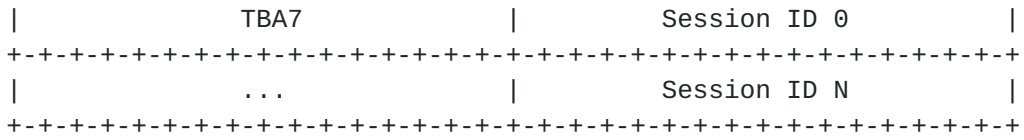
The New Outgoing Sessions Acknowledgement AVP can only be carried within a MSI message type. This AVP informs the LNS that the LAC is ready to replicate traffic for every Session ID listed in the AVP.

The AVP has the following format:

Vendor ID = 0
Attribute = TBA7 (16 bits) (Note: to be assigned by IANA)




|M|H|0|0|0|0|
Length
Vendor ID



This AVP must be placed in a MSI message and sent by the LAC towards the LNS to acknowledge the reception of a New Outgoing Sessions list received in a New Outgoing Sessions AVP from the LNS.

A LNS is allowed to send multicast traffic within the L2TP multicast session as soon as a New Outgoing Sessions Acknowledgement AVP is received for the corresponding L2TP multicast session.

A LNS is allowed to stop sending packets of the corresponding multicast flow within L2TP unicast sessions only if it receives a MSI message with the New Outgoing Session Acknowledgement AVP, and only for the unicast Session IDs mentioned in this AVP. The multicast traffic can then be conveyed in L2TP unicast sessions again when the L2TP multicast session goes down. From this standpoint, packets related to this multicast flow SHOULD NOT be conveyed within the L2TP unicast sessions mentioned in the AVP to avoid the duplication of multicast packets.

There can be from 1 to N Session IDs present in the New Outgoing Sessions Acknowledgement AVP (considering the maximum value of the Length field). Session IDs mentioned in this AVP that have not been listed in a previous New Outgoing Sessions AVP should be ignored. Non-acknowledged Session IDs MAY be listed in forthcoming New Outgoing Sessions AVPs, but multicast traffic MUST be sent to logical interfaces associated to these Session IDs as long as these Session IDs are not acknowledged for replication by the LAC.

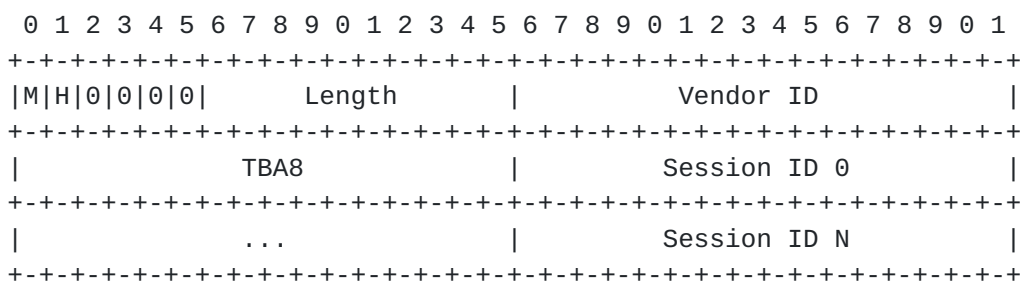
The M-bit MUST be set to 1, the AVP MAY be hidden (H-bit set to 0 or 1).

6.2.3. Withdraw Outgoing Sessions AVP (MSI)

The Withdraw Outgoing Sessions AVP is sent whenever there is one or more withdrawn subscriptions for the corresponding multicast flow (designated by the session ID on which the MSI is sent).
The LAC can stop forwarding packets to Session IDs mentioned in the AVP for the corresponding multicast flow as soon as it receives the MSI message embedding this Withdraw Target Session AVP.

The AVP has the following format:

Vendor ID = 0
Attribute = TBA8 (16 bits) (Note: to be assigned by the IANA)



There can be from 1 to N Session IDs present in the Withdraw Outgoing Sessions AVP (considering the value of the Length field). The M-bit MUST be set to 1, the AVP MAY be hidden (H-bit set to 0 or 1).

6.3. Multicast Packets Priority AVP (MSI)

The Multicast Packets Priority AVP is an optional AVP intended to provide the LAC with an indication on how to process multicast traffic against unicast traffic. Even though the LAC behavior is partially described here, the nature of the traffic (layer-2 frames for unicast traffic and pure IP packets for multicast traffic) is not a criteria for enforcing a traffic prioritization policy. Traffic processing for the provisioning of a uniformly-framed traffic for the final user is described is section 8.

Three different behaviors can be adopted:

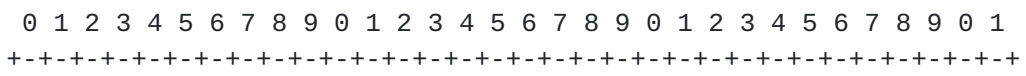
1) Best effort: the traffic is forwarded from the LAC to the end-user in the order it comes from the LNS, whatever the type of traffic.
2) Unicast traffic priority: traffic coming down the L2TP unicast session has priority over traffic coming down the L2TP multicast session.
3) Multicast traffic priority: traffic coming down the L2TP multicast session has priority over traffic coming down the L2TP unicast session.

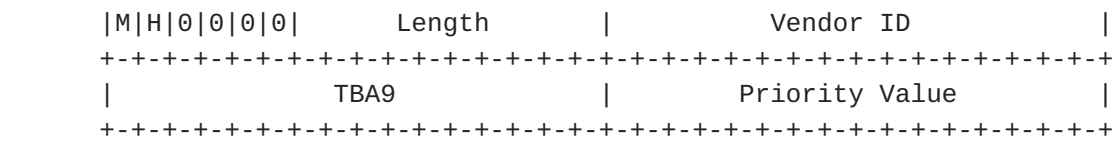
The priority is encoded as a 16-bit quantity, which can take the following values:

0: Best effort (default)
1: Unicast traffic priority
2: Multicast traffic priority

The AVP has the following format:

Vendor ID = 0
Attribute = TBA9 (16 bits) (Note: to be assigned by the IANA)


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



It is important to note that the multicast traffic rate can reach up to Maximum BPS (as indicated in MSRQ). This rate can exceed the maximum rate allowed for a particular end-user. This means that even with a priority value of 0, the end-user may receive multicast traffic only: unicast packets might be dropped because the multicast flow overwhelms the LAC forwarding buffer(s).

The default Priority Value is 0. The M-bit MUST be set to 0, the AVP MAY be hidden (H-bit set to 0 or 1).

There are two ways of using this AVP: global configuration and individual configuration.

6.3.1. Global configuration

The Multicast Priority Packet AVP is sent for all L2TP unicast sessions concerned with a specific multicast flow represented by a L2TP multicast session.
In this case, the AVP is sent in a L2TP MSI control message for the corresponding multicast session ID (Session ID = L2TP session for the corresponding multicast group). The priority value applies to all L2TP unicast sessions to which the multicast group designated by the L2TP multicast session is intended, as soon as this AVP is received.

6.3.2. Individual configuration

The Multicast Priority Packet AVP is sent for a specific L2TP unicast session that SHALL adopt a specific behavior for both unicast and multicast traffics. In this case, the AVP is sent in a L2TP MSI control message for the L2TP unicast session (Session ID = L2TP session for the concerned user). The priority value applies to the targeted session only, and does not affect the other sessions. It is important to note that in this case, all multicast packets carried in L2TP multicast sessions are treated the same way by the LAC for the concerned user.
This is the only case where a MSI control message can be sent for a L2TP unicast session.

6.3.3. Priority

It is the responsibility of the network administrator to decide which behavior to adopt between global or individual configurations, if the AVP is sent twice (one for a multicast group and one for a specific end-user). By default, only the individual configurations SHOULD be

taken into consideration in that case.

Support of the Multicast Packets Priority AVP is optional and SHOULD be configurable by the LAC administrator, if relevant.

7. Multicast session teardown

A L2TP multicast session should be torn down whenever there are no longer users interested in receiving the corresponding multicast traffic. More specifically, we consider that a multicast session becomes useless as soon as the related OSL has less than a predefined number of entries, this number being defined by a threshold.
Multicast session flapping may occur when the number of OSL entries is oscillating around the threshold, if the same value is used to trigger the creation or the deletion of a L2TP multicast session.
To avoid this behavior, two methods can be used:

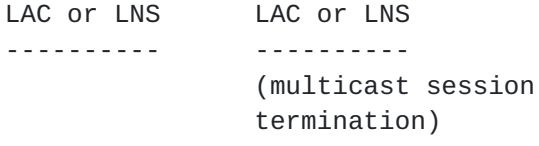
- The threshold value used to determine if the L2TP multicast session has to be torn down is lower than the MULTICAST_SESSION_THRESHOLD value;
- The MULTICAST_SESSION_THRESHOLD value is used to determine if the L2TP multicast session has to be torn down. A multicast session SHOULD be killed after a period of MULTICAST_SESSION_HOLDTIME seconds if the corresponding OSL maintains less than MULTICAST_SESSION_THRESHOLD entries. The MULTICAST_SESSION_HOLDTIME value is 10 seconds by default, and SHOULD be configurable either by the LAC or the LNS administrator.

The multicast session can be torn down for multiple reasons, including specific criteria not described here (can be vendor-specific).
A multicast session teardown can be initiated either by the LAC or the LNS. However, multicast session teardown MUST be initiated by the LNS if the termination decision is motivated by the number of users interested in receiving the traffic corresponding to a multicast flow.

7.1. Operations

The actual termination of a multicast session is initiated with a new Multicast-Session-End-Notify (MSEN) control message, sent either by the LAC or by the LNS.

The following is an example of a control messages exchange that terminates a multicast session:



<- MSEN

```
                    (Clean up)
        ZLB ACK ->
        (Clean up)
```

7.2. Multicast-Session-End-Notify (MSEN)

The Multicast-Session-End-Notify (MSEN) is a L2TP control message sent by either the LAC or the LNS to request the termination of a specific multicast session within the tunnel. Its purpose is to inform the peer with the relevant termination information, including the reason why the termination occurred. The peer MUST clean up any associated resources, and does not acknowledge the MSEN message.

As defined in [RFC2661], termination of a control connection will terminate all sessions managed within, including multicast sessions if any.

The MSEN message carries a Result Code AVP with an optional Error Code.

The following AVPs MUST be present in a MSEN message:

- Message Type
- Result Code
- Assigned Session ID

The associated Message Type AVP is encoded with the following values:

- Vendor ID = 0
- Attribute Type = 0
- Attribute Value = TBA10 (16 bits) (Note: to be assigned by IANA)

The M-bit MUST be set to 0, the H-bit MUST be set to 0.

7.3. Result Codes

The following values are the defined result codes for MSEN control messages:

- TBA11 (16 bits) - Session terminated for the reason indicated in the error code
- TBA12 (16 bits) - No multicast traffic to forward
- TBA13 (16 bits) - No more receivers
- TBA14 (16 bits) - No more receivers (filter-mode change)

(Note: TBA11, TBA12, TBA13 and TBA14 to be defined by the IANA)

o The code TBA11 refers to General Error Codes maintained by the IANA for L2TP.

o The code TBA12 MAY be used when the LAC detects that no traffic is coming down the multicast session, or when the LNS doesn't receive

multicast traffic to be conveyed over the L2TP multicast session during a certain period of time.

o The code TBA13 MAY be used by the LAC or the LNS when the OSL is empty.

o The code TBA14 MAY be used by the LNS when a multicast session is torn down because of a filter-mode change. This result code SHOULD also be used when the OSL becomes empty after a filter-mode change (passive termination when filter-mode change from INCLUDE to EXCLUDE, Section 4.3).

8. Traffic merging

Both unicast and multicast traffics have to be merged by the LAC in order to forward properly framed data to the end-user. Multicast packets are framed by the LAC and transmitted towards the proper end-user. Methods to achieve this function are not described here, since it is an implementation-specific issue.

All frames conveyed from the LAC to the end-users have to follow the framing scheme applied for the considered peer to which the traffic is destined (e.g. the LAC is always aware of the PPP link parameters, as described in [RFC2661], Section 6.14). It has to be noted that using L2TP Multicast Extension features is not appropriate for end-users who have negotiated a sequenced layer-2 connection with the LNS: while inserting PPP-encapsulated multicast packets in a session, the LAC cannot modify PPP sequencing performed by the LNS for each PPP session.

9. IANA Considerations

This document defines:

- 5 new Message Type (Attribute Type 0) Values:
  - o Multicast-Session-Request (MSRQ) : TBA2
  - o Multicast-Session-Response (MSRP) : TBA3
  - o Multicast-Session-Establishment (MSE) : TBA4
  - o Multicast-Session-Information (MSI) : TBA5
  - o Multicast-Session-End-Notify (MSEN) : TBA10
- 5 new Control Message Attribute Value Pairs:
  - o Multicast Capability : TBA1
  - o New Outgoing Sessions : TBA6
  - o New Outgoing Sessions Acknowledgement : TBA7
  - o Withdraw Outgoing Sessions : TBA8
  - o Multicast Packets Priority : TBA9
- 4 Result Codes for the MSEN message:
  - o Session terminated for the reason indicated in the error code : TBA11
  - o No multicast traffic for the group : TBA12

- No more receivers : TBA13
- No more receivers (filter-mode change): TBA14

IANA will assign, register and maintain values for these new attributes ([RFC3438]).

10. Security Considerations

It is possible for one receiver to be able to make additional multicast traffic that has not been requested go down the link of another receiver. This can happen if a single replication context per group is used in INCLUDE mode with receivers having divergent source lists, as well as in EXCLUDE mode if a receiver has a source list not shared by another. This behavior can be encountered every time receivers are connected to a common multi-access network.

The extension described in this document does not introduce any additional security issues as far as the activation of the L2TP protocol is concerned.

Injecting appropriate control packets in the tunnel towards a LAC to modify Outgoing Session List and flood end-users with unwanted multicast traffic is only possible if the control connection is hacked. As for any reception of illegitimate L2TP control messages:

- If the spoofed control message embeds consistent sequence numbers, next messages will appear out of synch yielding the control connection to terminate.
- If sequence numbers are inconsistent with current control connection states, the spoofed control message will be queued or discarded, as described in [RFC2661] section 5.8.

The activation of the L2TP multicast capability on a LAC could make the equipment more sensitive to Denial of Service attacks if the control connection or the related LNS is hacked. The LAC might also be sensitive to the burden generated by the additional replication work.

As mentioned in [RFC2661] section 9.2, securing L2TP requires that the underlying transport makes encryption, integrity and authentication services available for all L2TP traffic, including L2TP multicast traffic (control and data).

11. References

11.1. Normative References

[RFC1112] S. Deering, "Host Extensions for IP Multicasting", RFC 1112, August 1989.

[RFC1661] W. Simpson, "The Point-to-Point Protocol (PPP)", STD 51, RFC 1661, July 1994.

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

[RFC2236] W. Fenner, "Internet Group Management Protocol, Version 2", RFC 2236, November 1997.

[RFC2661] W. Townsley, A. Valencia, A. Rubens, G. Pall, G. Zorn, B. Palter, "Layer 2 Tunneling Protocol "L2TP" ", RFC2661, August 1999.

[RFC2710] S. Deering, W. Fenner, B. Haberman, "Multicast Listener Discovery (MLD) for IPv6", RFC 2710, October 1999.

[RFC3376] B. Cain et. al., "Internet Group Management Protocol, Version 3", RFC 3376, October 2002.

[RFC3438] W. Townsley, "Layer Two Tunneling Protocol (L2TP) Internet Assigned Numbers Authority (IANA) Considerations Update", RFC 3438, December 2002.

[RFC3810] Vida, R., L. Costa, S.Fdida, S. Deering, B. Fenner, I. Kouvelas, and B. Haberman, "Multicast Listener Discovery Version 2 (MLDv2) for IPv6", RFC 3810, June 2004.

11.2. Informative References

[PROXY] Fenner, B., He, H., Haberman, B., Sandick, H., "IGMP/MLD-based Multicast Forwarding ("IGMP/MLD Proxying")", Work in Progress.

12. Acknowledgments

Thanks to Christian Jacquenet for all the corrections done on this document and his precious advice, Pierre Levis for his contribution about IGMP, Francis Houllier for PPP considerations and Xavier Vinet for his input about thresholds. Many thanks to W. Mark Townsley, Isidor Kouvelas and Brian Haberman for their highly valuable input on protocol definition.

13. Author's Addresses

Gilles Bourdon
France Telecom
38-40, rue du General Leclerc
92794 Issy les Moulineaux Cedex 9 - FRANCE
Phone: +33 1 4529-4645
Email: gilles.bourdon@francetelecom.com

Appendix A. Examples of group states determination

*Example 1:

All users are managed in the same control connection.

Users {1, 2, 3} subscribe to (Group G1, EXCLUDE {})
Users {3, 4, 5} subscribe to (Group G2, EXCLUDE {})

Group states for this L2TP tunnel will be:
(G1, EXCLUDE, {})
(G2, EXCLUDE, {})

Therefore, two replication contexts will be created:
-RC1:
(*, G1) packets, Multicast Session MS1, OSL = 1, 2, 3
-RC2:
(*, G2) packets, Multicast Session MS2, OSL = 3, 4, 5

*Example 2:

All users are managed in the same control connection.

Users {1, 2, 3} subscribe to (Group G1, INCLUDE {S1})
Users {4, 5, 6} subscribe to (Group G1, INCLUDE {S1,S2})
Users {7, 8, 9} subscribe to (Group G1, INCLUDE {S2})

The group state for this L2TP tunnel will be:
(G1, INCLUDE, {S1, S2)})

If the LNS policy allows one replication context per (group, source), two replication contexts will be created:
-RC1:
(S1, G1) packets, Multicast Session MS1, OSL = 1, 2, 3, 4, 5, 6
-RC2:
(S2, G1) packets, Multicast Session MS2, OSL = 4, 5, 6, 7, 8, 9

If the LNS policy allows one replication context per (group, source-list), one replication context will be created:
-RC1:
({S1, S2}, G1) packets, Multicast Session MS1, OSL = [1..9]

*Example 3:

All users are managed in the same control connection.

Users {1, 2} subscribe to (Group G1, EXCLUDE {S1})

User {3} subscribes to (Group G1, EXCLUDE {S1, S2})

The group state for this L2TP tunnel will be:
(G1, EXCLUDE, {S1})

Therefore, one replication context will be created:
-RC1:
(*-{S1}, G1) packets, Multicast Session MS1, OSL = 1, 2, 3

Next, user {4} subscribes to (Group G1, INCLUDE {S1}). The group state for the L2TP tunnel is changed to:
(G1, EXCLUDE, {})

The replication context RC1 is changed to:
-RC1: (*, G1) packets, Multicast Session MS1, OSL = 1, 2, 3, 4

*Example 4:

All users are managed in the same control connection. The LNS policy allows one replication context per (group, source).

Users {1, 2, 3} subscribe to (Group G1, INCLUDE {S1, S2})

The group state for this L2TP tunnel will be:
(G1, INCLUDE, {S1, S2)})

Therefore, two replication contexts will be created:
-RC1:
(S1, G1) packets, Multicast Session MS1, OSL = 1, 2, 3
-RC2:
(S2, G1) packets, Multicast Session MS2, OSL = 1, 2, 3

Next, user {4} subscribes to (Group G1, EXCLUDE {}), equivalent to an IGMPv2 membership report. The group state for the L2TP tunnel is changed to:
(G1, EXCLUDE, {})

The replication context RC1 is changed to:
-RC1: (*, G1) packets, Multicast Session MS1, OSL = 1, 2, 3, 4

The replication context RC2 is changed to:
-RC2: no packets to forward, Multicast Session MS2, OSL = {}
(Multicast Session MS2 will be deleted)

When user {4} leaves G1, the group state for the L2TP tunnel goes back to:
(G1, INCLUDE, {S1, S2})

Replication contexts become:
-RC1:

(S1, G1) packets, Multicast Session MS1, OSL = 1, 2, 3
-RC2:

(S2, G1) packets, Multicast Session MS2, OSL = 1, 2, 3
(Multicast Session MS2 is re-established)

Full Copyright Statement