Network Working Group
Request for Comments: 4938
Category: Informational

B. Berry
H. Holgate
Cisco Systems,Inc.

June 2007

PPP Over Ethernet (PPPoE) Extensions for Credit Flow and Link Metrics

Status of This Memo

This memo provides information for the Internet community. It does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

Copyright Notice

Copyright (C) The IETF Trust (2007).

IESG Note

The PPP Extensions Working Group (PPPEXT) has reservations about the desirability of the feature described in this document. In particular, it solves a general problem at an inappropriate layer and it may have unpredictable interactions with higher and lower level protocols. The techniques described in this document are intended for use with a particular deployment technique that uses a PPP termination separated from a radio termination by an Ethernet, and that has radio-side flow control for a slower PPP-only link to remote nodes. Implementors are better advised to avoid split termination with inter-media protocol translation, and use standard Internet Protocol routing instead.

Abstract

This document extends the Point-to-Point over Ethernet (PPPoE) Protocol with a credit-based flow control mechanism and Link Quality Metric report. This optional extension should improve the performance of PPPoE over media with variable bandwidth and limited buffering, such as mobile radio links.

Table of Contents

1. Introduction

PPP over Ethernet (PPPoE) [2] is a protocol for establishing and encapsulating sessions between hosts and traffic aggregators (Access Concentrators) for PPP [1] transport over real or emulated Ethernet. PPPoE works well when both session endpoints have similar bandwidth, forwarding, and buffering capabilities that do not vary over time. However, it is insufficient for applications with variable bandwidth and limited buffering (for example, mobile radio links). This document addresses this problem by suggesting an extension to PPPoE to support credit-based session flow control and session-based link metric exchanges.

The diagram below illustrates the problem that this extension is intended to solve, for the case of a radio link. Here PPPoE sessions are used between access concentrators (routers) and radio transmission systems that are shown as radio neighbors. Each radio transmission system establishes point-to-point Radio Link Protocol (RLP) sessions with its neighbors and establishes a corresponding PPPoE session for each neighbor with the transmission system's associated access concentrator (router). The radio logically associates the PPPoE session with the corresponding RLP session.

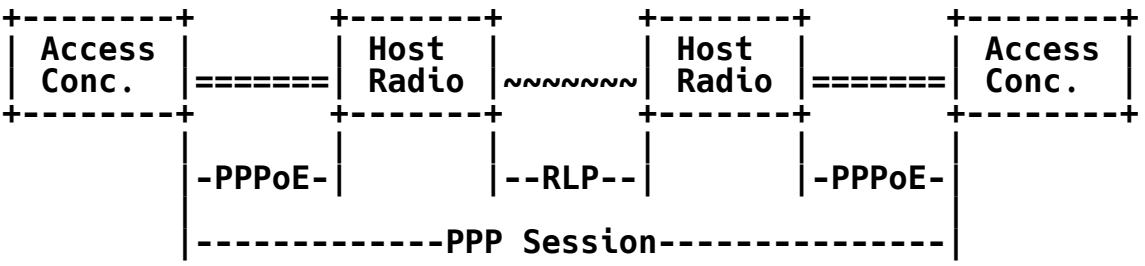


Figure 1. PPPoE Network

The capabilities of the RF links between RLP neighbors may vary over time due to mobility and environmental conditions. In many instances, the Host Radio has limited buffering capability to handle capacity changes in the RLP sessions. To limit buffering in the Host Radio, the PPPoE credit flow control mechanism provides dynamic buffering feedback to the access concentrator.

In the diagram above, from the access concentrator's perspective, each PPPoE session between it and the Host Radio represent a connection to a remote routable peer. For efficient routing, the local Host Radio uses the link metric mechanism to dynamically update the access concentrator route cost of the associated link.

While the example shows an RF-based application, the extensions are applicable to other media.

2. Payload

The Ethernet payload version field retains its value of 0x01. The extensions for credit flow control and link quality metrics are optional and backward compatible.

3. Overview of Protocol Extensions

PPPoE has two distinct stages. There is a Discovery Stage and a PPP Session Stage. During the Discovery Stage, the Host can optionally request a flow controlled PPP Session Stage. Once the Access Concentrator acknowledges the Host flow control request, all PPP Session Stage traffic must be flow controlled.

4. Discovery Stage

The packet exchange of the Discovery Stage is unchanged by this specification. The specifications of the Session Request (PADR) and the Session Confirmation (PADS) packets are extended to include the optional Credit Tag Type-Length-Value (TLV).

In addition, the optional Credit Grant (PADG) packet, the Credit Response (PADC) packet, and the Link Quality Metric (PADQ) packets are introduced.

4.1. PPPoE Active Discovery Request (PADR)

The PADR packet is extended to optionally contain a single Credit Tag TLV, indicating that the Host requests credit flow control for this session. The Credit Tag contains the Forward Credit Notification (FCN) and the Backward Credit Notification (BCN) to be applied to the PPP Session Stage. The FCN provides the initial credits granted to the Access Concentrator by the Host. The BCN value is set to 0.

An example packet is shown in Appendix B.

4.2. PPPoE Active Discovery Session-confirmation (PADS)

The PADS packet is extended to optionally contain a single Credit Tag TLV, indicating the Forward Credit Notification (FCN) and the Backward Credit Notification (BCN) of the PPP Session Stage.

If the PADR contained a Credit Tag, then the Access Concentrator PADS packet indicates support for credit flow control by including a Credit Tag. The PADS Credit Tag FCN represents the number of credits being initially granted to the Host. The Credit Tag BCN is an echo of the number of credits that the Host had granted to the Access Concentrator in the previous PADR packet.

Exchange of the Credit Tag TLV in the PADR and PADS indicates that credit flow control is supported by both the Access Concentrator and the Host for the designated PPP Session Stage. This is binding and must be followed for the entire duration of the PPP Session Stage. A session's credit binding must be established prior to any other credit indications can be exchanged.

The Access Concentrator PADS should only carry the Credit Tag in response to a Host PADR with Credits. If the Access Concentrator does not support credit flow, it should not include the Credit Tag in its PADS response. The Host must terminate a credit-based session that cannot be supported by the Access Concentrator. Credit Tags transmitted outside an established credit based session must be ignored.

An example packet is shown in Appendix B.

4.3. PPPoE Active Discovery Session-Grant (PADG)

The PPPoE Active Discovery Session-Grant (PADG) is a new packet defined in this specification. An Access Concentrator or Host MAY send a PADG at any time after the PADR/PADS exchange to grant incremental flow control credits. The CODE field is set to 0x0A and the SESSION_ID must be set to the unique value generated for this PPPoE Session.

The peer may then transmit data until the credits are exhausted.

When the peer receives a PADG packet, it adds the incremental credits to its working credit count and responds with a PPPoE Active Discovery Session-Credit (PADC) packet indicating the accumulated credits.

The PADG packet must contain a single Credit Tag TLV, indicating the Forward Credit Notification (FCN) and the Backward Credit Notification (BCN) of the PPP Session.

The Credit Tag FCN indicates the number of incremental credits being granted to the peer by the node. A value between 1 and 0xffff represents an incremental credit grant. The peer must add these credits to its accumulated transmit credit count. A value of 0x0000 represents a NULL grant, meaning that there are no additional credits being granted.

The Credit Tag BCN indicates the remaining absolute credits that have been granted by the peer to the node.

Once a credit has been granted, it must be honored. The largest number of outstanding credits at any time is 0xffff.

The PADG packet must contain a single Sequence Number Tag TLV. This tag is used to carry a unique 16-bit sequence number to uniquely identify each request. The sequence number should be initialized to zero and incremented by one for each new PADG. For retransmitted PADGs, the same sequence number that was used in the previous packet transmission is repeated.

An example packet is shown in Appendix B.

4.4. PPPoE Active Discovery Session-Credit Response (PADC)

The PPPoE Active Discovery Session-Credit Response (PADC) is a new packet defined in this specification. An Access Concentrator or Host must send a PADC in response to a PADG. The CODE field is set to

0x0B, and the SESSION_ID must be set to the unique value generated for this PPPoE session.

The PADC packet must contain a single Credit Tag TLV, indicating the Forward Credit Notification (FCN) and the Backward Credit Notification (BCN) of the PPPoE session, and any number of other Tag types.

The Credit Tag FCN represents the absolute credits remaining that have been granted to the peer by the node. The Credit Tag BCN represents the remaining absolute credits that have been granted to the node from the peer.

The PADC packet must contain a single Sequence Number Tag. The sequence number must be the sequence number associated with the PADG.

An example packet is shown in Appendix B.

4.5. PPPoE Active Discovery Quality (PADQ)

The PPPoE Active Discovery Quality (PADQ) is a new packet defined in this specification. An Access Concentrator or Host may send an optional PADQ at any time to query or report link quality metrics.

When transmitting PPP [1] streams over wireless links through radio modems, the quality of the RF link directly affects the throughput. The PPPoE Active Discovery Quality (PADQ) packet can be used by the radio modem to report RF link metrics. The CODE field is set to 0x0C, and the SESSION_ID must be set to the unique value generated for this PPPoE session.

The PADQ must carry a single Metric Tag TYPE, which contains the following fields:

Receive only - a bit that indicates whether the link is bi-directional or receive only. A value of -1- indicates that the link is receive-only.

Maximum data rate - the maximum theoretical data rate, in kilobits per second (kbps), that the Host link is capable of providing. When metrics are reported, the maximum data rate must be reported.

Current data rate - the current data rate, in kilobits per second (kbps), achieved on the Host link. If there is no distinction between maximum data rate and current data rate, current data rate should equal to the maximum data rate.

Latency - the transmission delay that a packet encounters as it is transmitted over the Host link. This is reported in absolute delay, milliseconds. If latency cannot be calculated, a value of 0 should be reported.

Resources - a percentage, 0-100, representing the amount of remaining or available resources, such as battery power. If resources cannot be calculated, a value of 100 should be reported.

Relative Link Quality (RLQ) - a non-dimensional number, 0-100, representing the relative link quality. A value of 100 represents a link of the highest quality. If the RLQ cannot be calculated, a value of 100 should be reported.

The PPPoE Active Discovery Quality (PADQ) packet can be used to query link metrics by setting the PADQ Metric Tag Length to zero.

An example packet is shown in Appendix B.

5. PPP Session Stage

This specification defines the optional use of TLV Tags in the PPP Session Stage. The first field following the PPP Session Stage LENGTH must be checked. If the value is equal to the PPP Protocol identifier (0xc021), then normal packet (payload) processing occurs. When the field following the PPP Session Stage LENGTH is not the PPP Protocol identifier (0xc021), a TLV is assumed. In this case, the Tag length is subtracted from the overall payload length.

The Credit Tag is the only optional TLV permitted in the PPP Session Stage. The Credit Tag TLV is used to support in-band flow control.

A PPP Session Stage packet with Credits is shown in Appendix B.

6. Credit Flow Considerations

For a given session, credit grants exchanged in the Discovery Stage, PADG-PADC, are referred to as out-of-band. Credit grants exchanged in the PPP Session Stage are referred to as in-band. Credit processing is only applied to the packets transmitted in the PPP Session Stage.

Out-of-band credit management is handled by periodic exchange of the PPPoE Active Discovery Grant (PADG) and PPPoE Active Discovery Credit (PADC) packets.

In-band credit management allows credits to be incrementally granted with each PPP Session Stage packet. These in-band incremental credit grants are not explicitly unacknowledged. However, they are reflected in the in-band credit flow from the peer node. This offers the greatest credit granting efficiency when traffic rates are high.

Once agreed upon during the Discovery Stage, credit grants are required to transmit packets in the PPP Session Stage. A node must grant credits to its peer, before the peer can transmit packets to the granting node.

Credits are granted incrementally in the forward direction. Locally, a node manages the credits that it has granted to a peer, as well as the credits that a peer has granted to it.

Grants received from a peer are added to a local running credit counter. The accumulated credits are decremented with each packet the node transmits to the peer. When the running counter reaches zero, the node stops transmitting packets to the peer. The values of the PADC are not simply an echo of the PADG. They represent the current internal FCN/BCN values of that node.

To manage the credits that a node has granted, the node maintains a running counter. With each PPP Session Stage packet received from the peer, the running counter is decremented. When the running counter reaches zero, no additional packets are expected. The node incrementally grants more credits to the peer to maintain packet flow. Packets received when granted credits that have been exhausted are discarded.

The largest possible credit limit is 0x0ffff. If an incremental credit grant causes the accumulated count to exceed this value, the max value is used.

One unit of credit represents 64-bytes, so a grant of 4 credits translates to 256 bytes.

7. PADG and PADC Retransmission

When a node does not receive a PADC packet in response to a PADG within a specified amount of time, it should transmit a new PADG packet with zero credits, using the same sequence number and double the waiting period. A PADC response with the associated sequence number will indicate whether or not the previously granted credits were accumulated. If they were not, a PADG with credits, with an incremented sequence number, should be transmitted. This process should be repeated until granted credits are properly acknowledged or as many times as desired.

When a node does not receive a PADQ metric packet within a specified amount of time, it should resend the PADQ query packet and double the waiting period. This can be repeated as many times as desired.

8. Other Considerations

A node may autonomously generate PADQ metric packets. The rate of autonomously generated PADQ metric packets may need to be throttled so as not to overrun the peer.

The sending and receiving of PPPoE control packets are independent of credit counts. For example, a node must always be able to receive a PADG and send a PADC.

During normal operation, nodes may disagree about the number of credits. Operational credit mismatches would occur due to packets in transit on the wire. Much larger credit mismatches can occur if there are transmission errors. To correct these larger errors, the BCN fields of the PADG and PADC packets and in-band credit grants from a peer should be used by the receiving node to set the credit values of its peer.

9. IANA Considerations

IANA has assigned the following PPPoE TAG Values as noted in [3]:

TAG Value	TAG Name	Tag Description	Reference
262 0x0106	Credits	See the reference	[RFC4938]
263 0x0107	Metrics	See the reference	[RFC4938]
264 0x0108	Sequence Number	See the reference	[RFC4938]

IANA has assigned the following PPPoE Code fields as noted in [3]:

Code	PPPoE Packet Name	Description	Reference
10 0x0a	PADG, Session-Grant	See the reference	[RFC4938]
11 0x0b	PADC, Session-Credit Response	See the reference	[RFC4938]
12 0x0c	PADQ, Quality	See the reference	[RFC4938]

10. Security Considerations

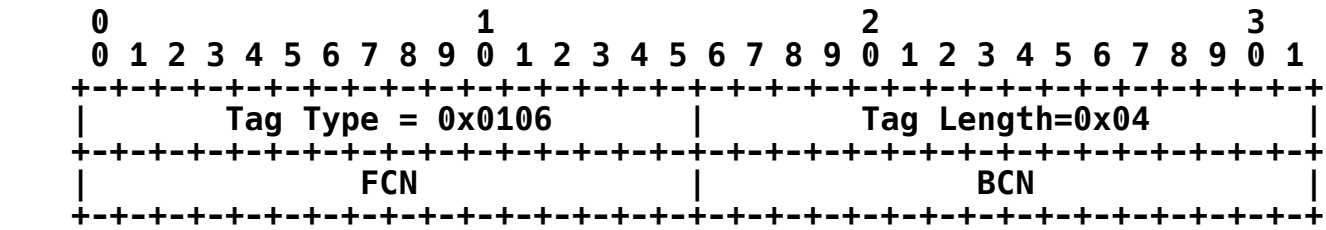
This memo defines a mechanism for adding flow control to the existing PPP Over Ethernet (PPPoE) sessions. These extensions are subsequent to the existing PPPoE security mechanisms as described in RFC 2516 [2]. It is required that the Service Tag and Session ID always be validated prior to processing credits.

Appendix A: Tag Values

Feature Tag_Types and Tag_Values

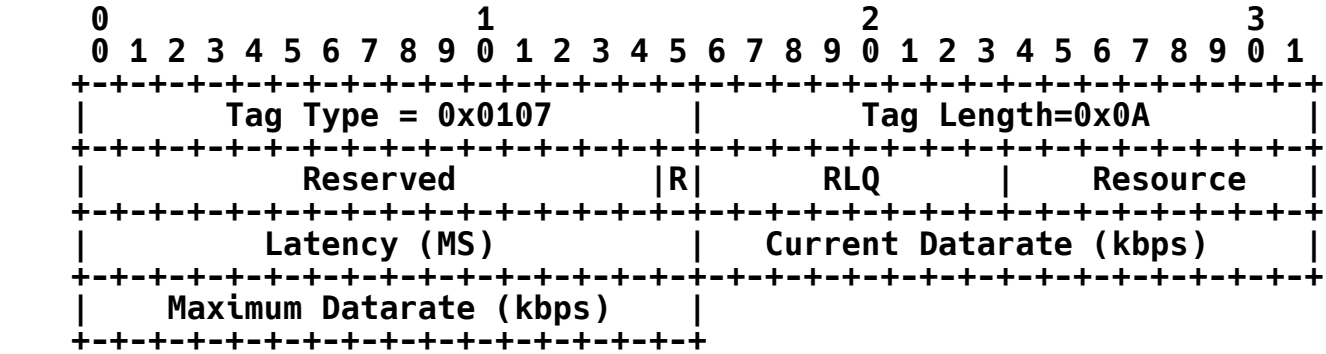
0x0106 Credits

This tag contains the Forward Credit Notification (FCN) and the Backward Credit Notification (BCN). The Credit Tag TLV is OPTIONAL with the PADR, PADS, and the PPPoE data payload packet (ETHER_TYPE=8864).



0x0107 Metrics

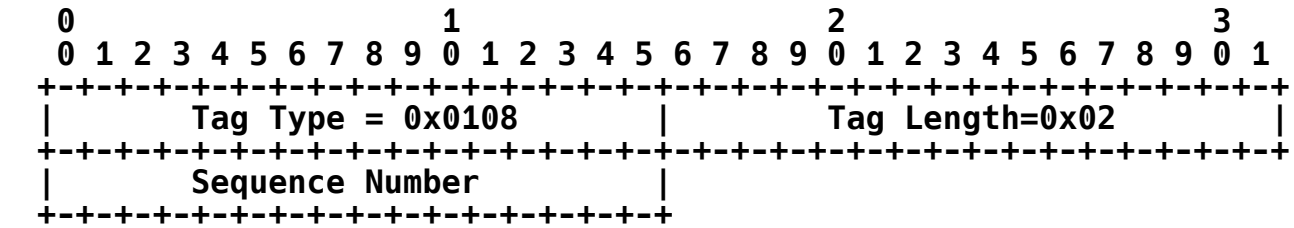
This tag is used to report the link quality and performance. The Metrics Tag TLV contains the Receive Only indicator, Resource status, Latency, Relative Link Quality (RLQ), Current data rate, and Maximum data rate. The Metrics TLV is required by the PADQ packet.



0x0108 Sequence Number

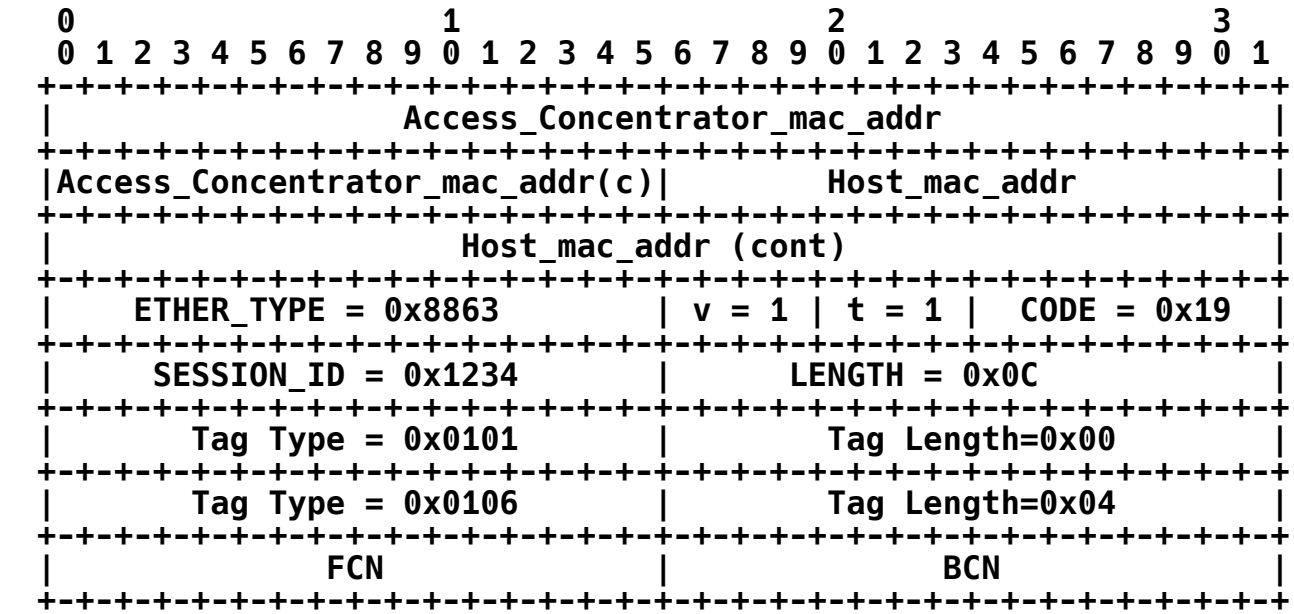
This tag is used to carry a unique 16-bit sequence number in order to identify a specific request and the associated response. The sequence number should be initialized to zero and incremented by one for each new request. For retransmitted packets, the same sequence number that was used in the previous packet transmission is repeated. The PADG and PADC packets require the Sequence Number Tag.

For example, the sequence number sent in the PADG request is echoed in the PADC response. This ties a specific PADC response to a specific PADG request.

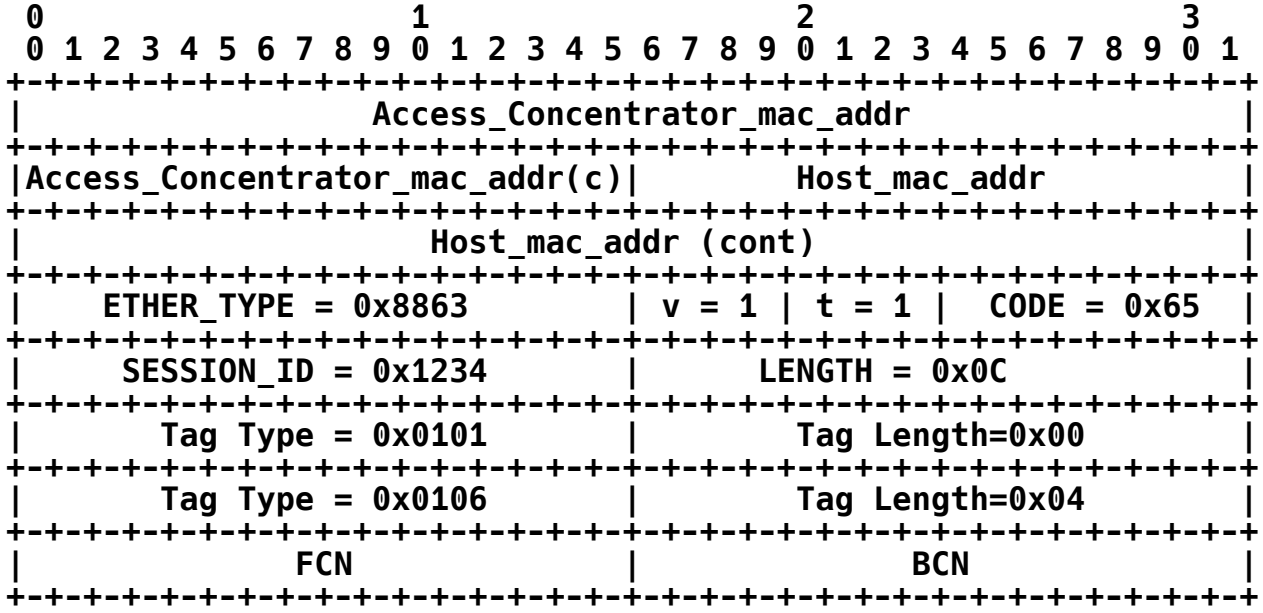


Appendix B: Example Message Formats

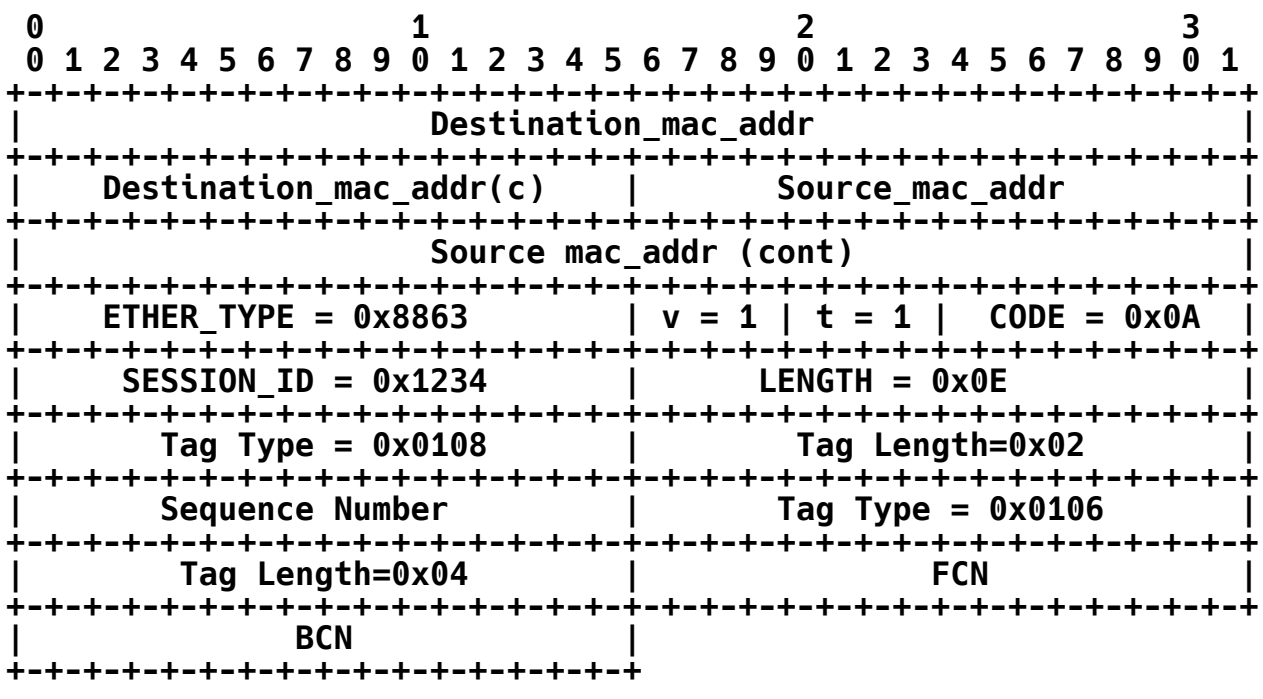
A PADR packet with OPTIONAL Credit Tag Type 0x0106:



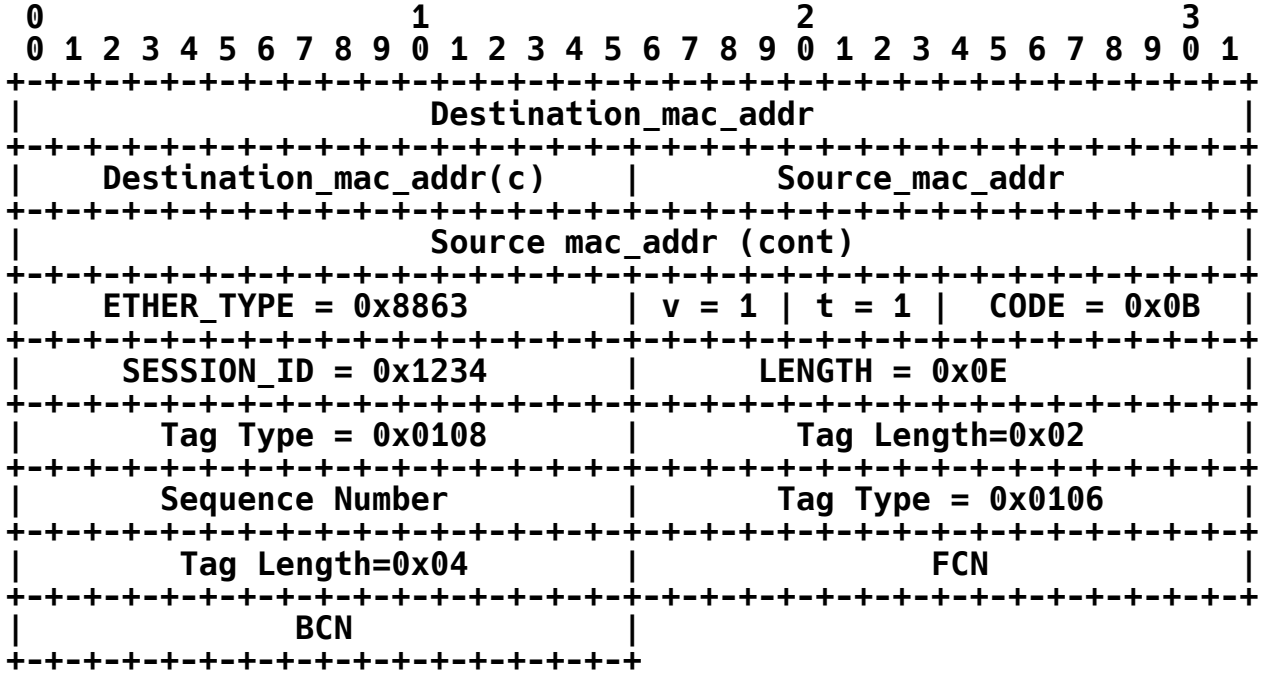
A PADS packet with OPTIONAL Credit Tag Type 0x0106:



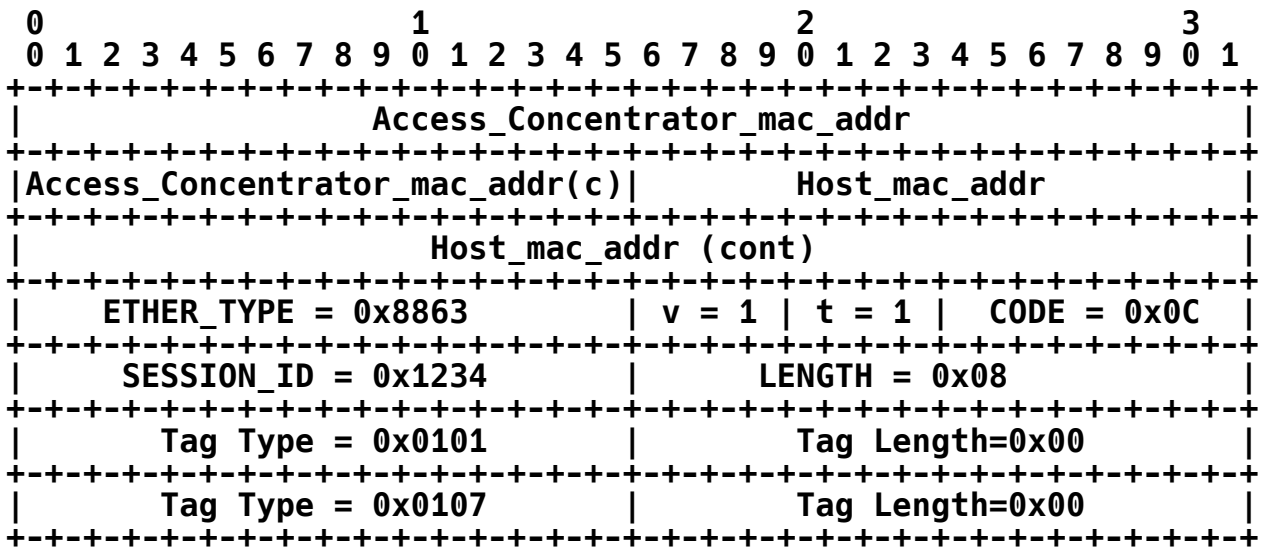
A PADG packet with Credit Tag Type 0x0106:



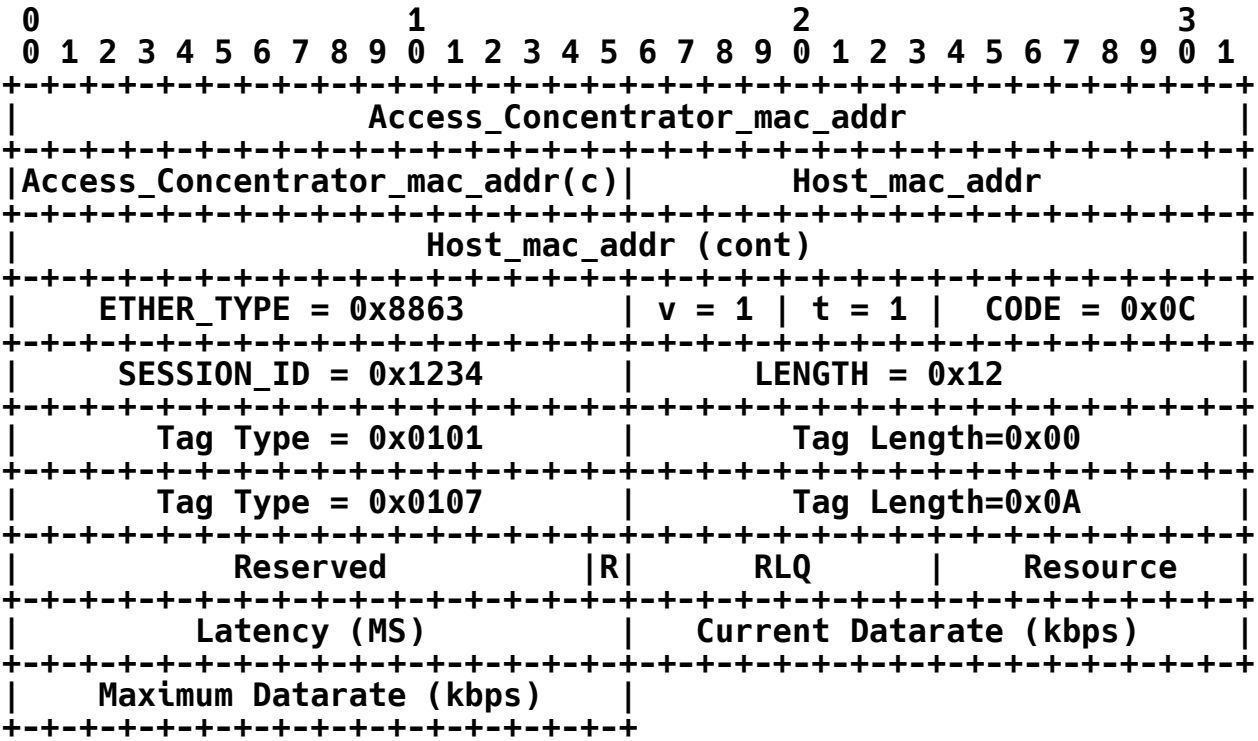
A PADC packet with Credit Tag Type 0x0106:



A PADQ packet to query for the link metrics: This is indicated by the Metric Tag Length=0.

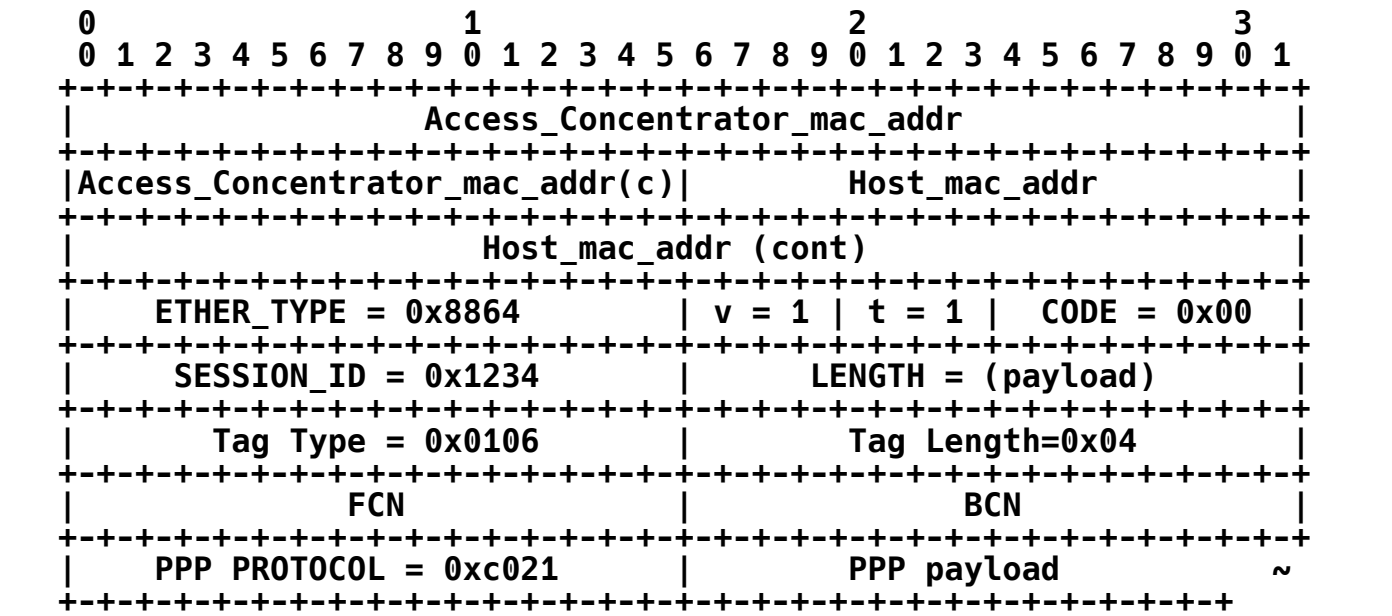


A PADQ packet with Metric Tag Type 0x0107:



A PPP LCP packet with optional Credit Tag Type 0x0106:

While the PPP protocol value is shown (0xc021), the PPP payload is left to the reader. This is a packet from the Host to the Access Concentrator.



Acknowledgements

The authors would like to acknowledge the influence and contributions from Billy Moon, Fred Baker, Stan Ratliff, and Ed Paradise.

Normative References

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

[2] Mamakos, L., Lidl, K., Evarts, J., Carrel, D., Simone, D., and R. Wheeler, "A Method for Transmitting PPP Over Ethernet (PPPoE)", RFC 2516, February 1999.

[3] Arberg, P. and V. Mammoliti, "IANA Considerations for PPP over Ethernet (PPPoE)", RFC 4937, June 2007.

Authors' Addresses

Bo Berry
Cisco
170 West Tasman Drive
San Jose, CA 95134
USA
EMail: bberry@cisco.com

Howard Holgate
Cisco
170 West Tasman Drive
San Jose, CA 95134
USA
EMail: hholgate@cisco.com

Full Copyright Statement

Copyright (C) The IETF Trust (2007).

This document is subject to the rights, licenses and restrictions contained in BCP 78 and at www.rfc-editor.org/copyright.html, and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Acknowledgement

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