                     IPoE Client Health Checking
                  draft-patterson-intarea-ipoe-health-02

Abstract

   PPP over Ethernet clients have the functionality to detect path
   unavailability by using PPP Keepalives. IP over Ethernet does not
   have this functionality, and it's not specified when an IP over
   Ethernet client should consider its WAN connectivity down, unless
   there is a physical layer link down event.

   This document describes a way for IP over Ethernet clients to achieve
   connectivity validation, similar to that of PPP over Ethernet, by
   using BFD Echo, or ARP and Neighbor Discovery functions.

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 https://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 December 21, 2018.

Table of Contents

1. Introduction

   PPP [RFC1661] makes use of regular LCP echos and replies to
   continually test the data link layer, if the peer fails to respond to
   a predetermined number of LCP echos, the PPP session is terminated
   and will return to the Link Dead phase, ready for reestablishing.
   IPoE currently lacks this functionality.

Physical link state change on an IPoE client can trigger the renewing
of a DHCP lease, however any indirect upstream network changes are
not visible to the IPoE client.
An outage or planned maintenance work on a BNG or intermediate DHCP
Relay, can leave an IPoE client with a stale DHCP lease for up to the
Valid Lifetime.

IPoE Health Check allows for an IPoE client to proactively or
passively monitor the state of upstream connectivity, and defines
several actions that may be taken to help the client recover.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
"SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and
"OPTIONAL" in this document are to be interpreted as described in BCP
14 [RFC2119] [RFC8174] when, and only when, they appear in all
capitals, as shown here.

1.2. Terminology

o BNG: Broadband Network Gateway. Often also running a DHCP server
   or relay.

o CE: Customer Equipment. aka. Customer Premise Equipment (CPE),
   Residential Gateway (RG).

o IPoE: IP over Ethernet.

o IPoE Client: A network device, often a CE, running a DHCPv4 and/or
   DHCPv6 client.

o IPoE Health Check: The name of the process described in this
   document.

2. Alternative Mitigations

o Short DHCP lease times reduce the time a client may be left in a
   stale state, but scale poorly, putting extra load on the DHCP
   server.

o Broadband Forum's [TR-146], Section 6.2.2 discusses this problem
   and suggests the use of BFD echo [RFC5880]. This document
   acknowledges TR-146 and recommends the use of BFD echo for health
   checks, but notes that it is not widely available within consumer
   CEs. This document also introduces alternative actions, as the
   renew approach taken in TR-146 is susceptible to the issues
   described in Behaviour 0 (Section 4.1).

   o For planned work, network engineers could include DHCPv4 Force
      Renew [RFC3203] or DHCPv6 Reconfigure [RFC3315]-bis in their
      maintenance plans, however neither of these have been widely
      adopted by CE or BNG vendors due to authentication complexity.

3. IPoE Health Checks

   An IPoE client supporting IPoE Health Check SHOULD begin sending
   health checks at the Interval specified, upon successful binding of a
   lease that contains a valid IPoE Health Check DHCP Option.

   An IPoE client MAY be locally configured for IPoE health checks.
   Non-default local parameters SHOULD override any signalled via DHCP.

   An IPoE client MAY use default parameters in lieu of manually
   configured, or DHCP signalled parameters. Manually configured or
   DHCP signalled parameters SHOULD override any default parameters.

3.1. Parameters

   IPoE Health Check specifies the following parameters:

   o Interval (Integer): The frequency in seconds, which health checks
      are sent by the IPoE client.

   o Limit (Integer): The number of consecutive checks that can fail
      before an action is taken.

   o Behaviour (Integer): Specifies what additional actions are to be
      taken when triggered.

   o Passive (Boolean): Forces passive health checks instead of active.

   o Alternative Target Address (IP): Overrides the default gateway as
      the target of health checks.

3.2. BFD Echo

   An IPoE client SHOULD use BFD Echo [RFC5880] as the health check
   mechanism.

   If BFD echos are used, the destination IP address MUST be locally
   bound on the IPoE client and SHOULD be from the lease triggering the
   IPoE Health Check.

   The use of BFD Echo as the health check mechanism provides the added
   benefit of validating the DHCP lease state, proving layer 3 as well
   as layer 2 connectivity.

3.3. Neighbor Discovery

   If an IPoE client with active DHCPv6 leases is unable to send BFD
   echos, it MUST send Neighbor Solicits [RFC4861], Section 4.3 for the
   target address. If no Alternative Target Address is set, the target
   address SHOULD be the default gateway as obtained from the Operating
   System.
   Neighbor Solicits SHOULD be sent at the frequency set by the Interval
   parameter (Section 3.1).

3.4. ARP

   If an IPoE client with active DHCPv4 leases is unable to send BFD
   echos, it MUST send ARP requests [RFC0826] for the target address.
   If no Alternative Target Address is set, the target address SHOULD be
   the client's default gateway, as received within the DHCPv4 Option 3
   Router option of the lease.
   ARP requests SHOULD be sent at the frequency set by the Interval
   parameter (Section 3.1).

3.5. Alternative Target

   An alternative IP address MAY be included within the IPoE health
   check DHCP option, or locally configured. If an alternative target
   address is specified, it MUST be used as the target for health checks
   instead of the default gateway.

   If an alternative target address provided is outside of a locally
   attached route, health checks SHOULD implicitly fail until a matching
   local route is installed. If a matching locally attached route is
   subsequently installed, health checks SHOULD continue as normal.

3.6. Passive Checks

   If an IPoE client is unable to proactively send health checks itself,
   it SHOULD passively check the operating system's ARP and Neighbor
   cache tables.

   In IPoE Health Check passive mode, alternate target addresses outside
   of locally attached routes MUST NOT be supported.

   Passive IPoE health checks SHOULD use the health check parameters
   signalled by DHCP or configured locally. The IPoE client SHOULD
   passively check the ARP or Neighbor cache tables for the target
   address, every Interval (Section 3.1) seconds. If the neighbor entry
   is in state INCOMPLETE for Limit (Section 3.1) checks, the specified
   IPoE Health Check Action MUST be taken.

Passive-only mode can be forced either by local configuration, or by
a DHCP server setting the Passive flag in the DHCP Option. If
passive-only mode has been set, the IPoE client MUST only use passive
checking for that particular lease health check.

4. Action Behaviours

IPoE Health Check defines four configurable behaviours once the
timeout threshold has been reached. All three behaviours make use of
existing procedures outlined in [RFC2131], Section 4.4.5 for DHCPv4,
and [RFC3315]-bis, Sections 18.2.4, 18.2.5 for DHCPv6.

IPoE Health Check behaviour MAY be signalled per lease by DHCP, or
locally configured. Locally configured, non-default, behaviour
settings SHOULD take precedence over those signalled by DHCP.

4.1. Behaviour 0: Renew (Default)

After Limit (Section 3.1) consecutive failures, the IPoE client MUST
set T1 of the specified lease, to zero. This will trigger a RENEW to
the original DHCP server, as per [RFC3315]-bis and [RFC2131].

If connectivity to the original DHCP server has recovered, and the
server can satisfy the request, the lease may be renewed and timers
updated.
If the original DHCP server cannot satisfy the request, it may reject
the request, to which the DHCP client should begin discovery or
solicit phase anew.

Neither of the above two responses are guaranteed, and as such, an
administrator may elect to use one of the below additional behaviours
to help expedite the IPoE client's recovery process.

Unless specified otherwise, additional actions MUST also be taken if
the DHCP Option Behaviour bits are non-zero. Some behaviours may
offer alternative actions instead of compound ones, they will state
this specifically.

4.2. Behaviour 1: Rebind

If the Behaviour field is set to 1, T2 MUST also be set to zero,
along with T1. This tells the IPoE client to immediately move to the
rebind phase, attempting to renew the lease from any available
server.

This method can be useful in a resilient layer 2 access topology,
with multiple active DHCP servers.

4.3. Behaviour 2: Solicit

   If the Behaviour field is set to 2, T1 and T2 MUST both be set to
   zero as per previous behaviours.

   The IPoE client MUST skip the renew and rebind phases, moving
   straight to the discovery or solicit phase.

   The IPoE client MUST NOT send a DHCP RELEASE.

   The IPoE client MUST keep the address or prefix in the preferred
   state until the preferred lifetime expires, and MUST keep the address
   or prefix until the valid lifetime expires.

   The IPoE client SHOULD include the lease address or prefix in the
   DISCOVER or SOLICIT.

   The DUID and IAID MUST be the same as used in the current lease.

   This method can be useful when using DHCP servers that silently
   discard unknown renew attempts instead of sending back a DHCPv4 NAK
   or DHCPv6 Reply.

4.4. Behaviour 3: Expire & Release

   If the Behaviour field is set to 3, T1, T2, and Valid Lifetime MUST
   all be set to 0, and the IPoE Client MUST send a DHCP RELEASE message
   as per [RFC2131], Section 3.1 for DHCPv4 and [RFC3315]-bis,
   Section 18.2.7

   Once the RELEASE process has completed, the client returns to the
   discovery or solicit phase.

   If the IPoE client is already in the renew or rebind state when
   Behaviour 3 is triggered, the client MUST cease renew or rebind
   attempts and wait for any outstanding messages to time out before
   sending a RELEASE. If an outstanding renew or rebind attempt is
   successful, the IPoE client MUST update T1, T2 and lease lifetimes
   appropriately, and MUST NOT continue with Behaviour 3.

   This method can be useful to clean out state within the network. For
   example, a DHCP relay may be left with stale lease information after
   an outage or maintenance on a DHCP server.

4.5. LAN Considerations

   If all DHCPv6 leases have expired, either naturally or proactively
   with IPoE health checks, it is expected than an IPoE client acting as
   a router, would withdraw itself as a default router on the LAN,
   following requirement G-5 of [RFC7084], Section 4.1.

5. DHCP Option

   IPoE Health Check defines a new option for both DHCPv4 and DHCPv6
   servers to signal suggested health check parameters to clients.
   IPoE clients SHOULD use these values when no locally configured
   parameters have been defined.

   The option data fields are common between DHCPv6 and DHCPv4, with the
   exception of the alternate target address field, which is 32 bits in
   the DHCPv4 option and 128 bits in the DHCPv6 option.

5.1. DHCPv6

   For DHCPv6, this Option MUST be within a specific Identity
   Association as an IPoE client MAY have multiple IAs with different
   health check parameters.

```
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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         option-code           |           option-len          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    limit      |P| behaviour   |            reserved           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             interval                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
|                          alternate-                           |
|                        target-address                         |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

                     Figure 1: DHCPv6 Option Format

```
option-code: OPTION_IPOE_HEALTH (TBD).
option-len: 24.
limit:      Consecutive failed checks, before an action is taken.
P:          Passive Flag. Force passive-only health checks.
behaviour:  Behaviour field.
            0: Trigger Renew.
            1: Trigger Rebind.
            2: Expire lease, start solicit phase.
            3: Release.
            4 - 127: Reserved.
interval:   How often a health check should be sent.
            Expressed in units of seconds.
alternate-target-address: Optional health check target address.
            Always present, set to zero if no address provided.
```

                              Figure 2

5.2. DHCPv4

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  option-code  |   option-len  |     limit     |P|  behaviour  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           interval                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    alternate-target-address                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
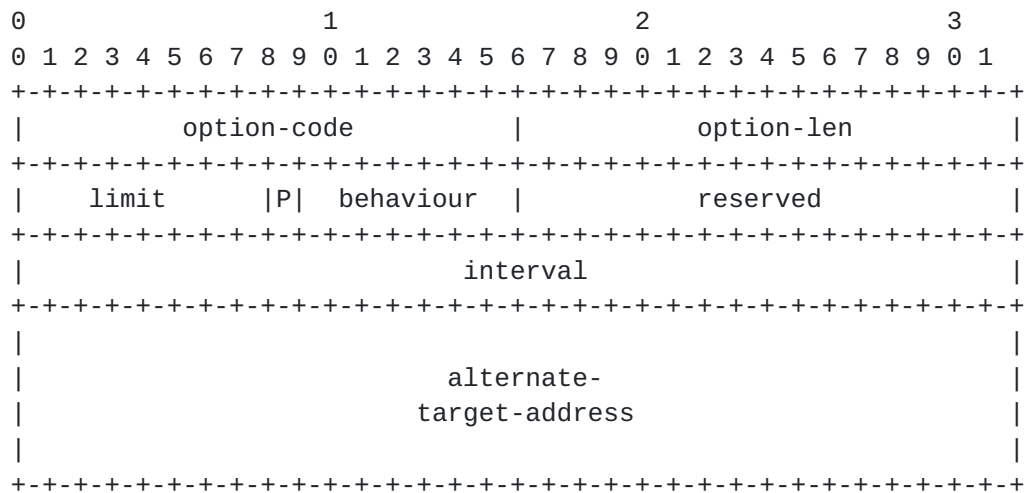
                     Figure 3: DHCPv4 Option Format

```
option-code: OPTION_IPOE_HEALTH (TBD).
option-len: 10.
limit:      Consecutive failed checks, before an action is taken.
P:          Passive Flag. Force passive-only health checks.
behaviour:  Behaviour field.
            0: Trigger Renew.
            1: Trigger Rebind.
            2: Expire lease, start discovery phase.
            3: Release.
            4 - 127: Reserved.
interval:   How often a health check should be sent.
            Expressed in units of seconds.
alternate-target-address: Optional health check target address.
            Always present, set to zero if no address provided.
```

                              Figure 4

[6]. **Multihomed Clients**

   An IPoE client MAY have multiple leases from the same, or different
   DHCP servers. These leases MAY have different IPoE health check
   parameters, and health checks MUST be treated distinctly, tracking
   the particular lease that they belong to.

   Each distinct IPoE health check MUST use an appropriate target
   address as per IPoE Health Check ([Section 3]).

   If an IPoE client is configured with multiple IPoE Health Checks that
   use the same target address, it SHOULD suppress additional checks,
   preferring the parameters with the lowest timeout value.
   I.e. Timeout = Interval * Limit

   Local network administrators may choose to override DHCP-signalled
   parameters in order to facilitate appropriate IPoE Health Check
   operation in a multihomed environment.

[6.1]. **Neighbor Discovery**

   As DHCPv6 does not convey default gateway or other routing
   information, an IPoE client using the ND health check method SHOULD
   obtain the target address by querying the operating system for
   default routes.\

   If multiple default routes exist, ND-based IPoE health checks SHOULD
   attempt to match the target address to the lease by the interface the
   lease is bound to.

   If only a single default route exists, and that default route is not
   routed out the interface the lease was bound to, ND-based health
   checks for that particular lease SHOULD be paused.

[6.2]. **ARP**

   ARP-based IPoE health checks for DHCPv4 make use of the default
   gateway address specified in the lease. As a route for each gateway
   should exist regardless of current route preference, health checks
   SHOULD be run for each lease that is configured for IPoE health
   check.

[7]. **Security Considerations**

   While ARP and Neighbor Discovery are more likely to be handled by
   hardware linecards compared to DHCP messaging, they may be subject to
   protections outlined in [[RFC6192]]. Routers SHOULD ensure that

sufficient quantities of this traffic are permitted to safely ingress
the control plane.

IPoE Health Check frequency would typically be controlled by the
Network using DHCP Options, but overly zealous, locally configured
IPoE clients, could have an adverse impact.

Unlike ARP and ND, BFD echo uses an IP packet destined for the IPoE
client, the peer forwards the packet back to the IPoE client without
any local processing.

Behaviour 2 ([Section 4.3](#)) introduces a privacy risk, possibly leaking
lease information if the IPoE client has been moved to a different
network, e.g., from one fixed line provider to another. The authors
believe this not to be a major concern.

[8](#). IANA Considerations

IPoE Health Check requires the allocation of two new DHCP Options.
One for DHCPv4 and one for DHCPv6. The option for both will be
referred to as OPTION_IPOE_HEALTH.

[9](#). Acknowledgements

The authors would like thank Ian Farrer, Dusan Mudric, Bernie Volz,
Dave Freedman, and Job Snijders for their review and comments on this
and previous versions of this document.

[10](#). [Appendix A](#). Changes from -00

This section should be removed by the RFC Editor.

o Added reference to TR-146.

o Added BFD Echo section, and wording to prefer it as the health
   check mechanism over ARP/ND, if available.

[11](#). [Appendix B](#). Changes from -01

This section should be removed by the RFC Editor.

o Emphasised preference for use of BFD echo as the health check
   mechanism.

o Removed lifetime expiration from Behaviour 2 and clarified usage.

o Updated Behaviour 3 with instructions for whilst mid-renew/rebind.

   o Reworded multihoming section.

   o Added Acknowledgements.

[12](). References

[12.1](). Normative References

   [RFC0826] Plummer, D., "An Ethernet Address Resolution Protocol: Or
              Converting Network Protocol Addresses to 48.bit Ethernet
              Address for Transmission on Ethernet Hardware", STD 37,
              [RFC 826](), DOI 10.17487/RFC0826, November 1982,
              <[https://www.rfc-editor.org/info/rfc826]()>.

   [RFC2131] Droms, R., "Dynamic Host Configuration Protocol",
              [RFC 2131](), DOI 10.17487/RFC2131, March 1997,
              <[https://www.rfc-editor.org/info/rfc2131]()>.

   [RFC3315] Droms, R., Ed., Bound, J., Volz, B., Lemon, T., Perkins,
              C., and M. Carney, "Dynamic Host Configuration Protocol
              for IPv6 (DHCPv6)", [RFC 3315](), DOI 10.17487/RFC3315, July
              2003, <[https://www.rfc-editor.org/info/rfc3315]()>.

   [RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman,
              "Neighbor Discovery for IP version 6 (IPv6)", [RFC 4861](),
              DOI 10.17487/RFC4861, September 2007,
              <[https://www.rfc-editor.org/info/rfc4861]()>.

   [RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection
              (BFD)", [RFC 5880](), DOI 10.17487/RFC5880, June 2010,
              <[https://www.rfc-editor.org/info/rfc5880]()>.

[12.2](). Informative References

   [RFC1661] Simpson, W., Ed., "The Point-to-Point Protocol (PPP)",
              STD 51, [RFC 1661](), DOI 10.17487/RFC1661, July 1994,
              <[https://www.rfc-editor.org/info/rfc1661]()>.

   [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
              Requirement Levels", [BCP 14](), [RFC 2119](),
              DOI 10.17487/RFC2119, March 1997,
              <[https://www.rfc-editor.org/info/rfc2119]()>.

   [RFC3203] T'Joens, Y., Hublet, C., and P. De Schrijver, "DHCP
              reconfigure extension", [RFC 3203](), DOI 10.17487/RFC3203,
              December 2001, <[https://www.rfc-editor.org/info/rfc3203]()>.

   [RFC6192] Dugal, D., Pignataro, C., and R. Dunn, "Protecting the
              Router Control Plane", RFC 6192, DOI 10.17487/RFC6192,
              March 2011, <https://www.rfc-editor.org/info/rfc6192>.

   [RFC7084] Singh, H., Beebee, W., Donley, C., and B. Stark, "Basic
              Requirements for IPv6 Customer Edge Routers", RFC 7084,
              DOI 10.17487/RFC7084, November 2013,
              <https://www.rfc-editor.org/info/rfc7084>.

   [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC
              2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174,
              May 2017, <https://www.rfc-editor.org/info/rfc8174>.

   [TR-146]   "TR-146 Subscriber Sessions", 2013,
              <https://www.broadband-forum.org/technical/download/
              TR-146.pdf>.

Authors' Addresses

   Richard Patterson
   Sky UK

   Email: richard.patterson@sky.uk

   Mikael Abrahamsson
   T-Systems

   Email: mikael.abrahamsson@t-systems.se