intarea
Internet-Draft
Intended status: Standards Track
Expires: December 21, 2018

R. Patterson
Sky UK
M. Abrahamsson
T-Systems
June 19, 2018

IP over Ethernet (IPoE) Client Session 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~~ it is not specified in the IETF 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~~ mechanism 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.

Commentaire [Med1]: to be fair with BBF.

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, for example, ~~—~~a Broadband Network Gateway (BNG) or intermediate DHCP Relay, can leave an IPoE client with a stale DHCP lease for up to the Valid Lifetime.

IPoE session 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.

Section 6.2 of [TR-146] describes this problem, while [TR-124] identifies some requirements to solve the problem.

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

This document makes use of the following terms:

- BNG: Broadband Network Gateway. Often also running a DHCP server or relay.
- CE: Customer Equipment. ~~aka.~~Also known as, ~~—~~Customer Premise Equipment (CPE) or ~~,~~ Residential Gateway (RG).
- IPoE: IP over Ethernet.
- IPoE Client: A network device, often a CE, running a DHCPv4 and/or DHCPv6 client.
- IPoE Health Check: The name of the process described in this document.

2. Alternative Mitigations

- 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.
- Broadband Forum's [TR-146] ~~, Section 6.2.2~~and [TR-124] discuss~~es~~ this problem and ~~suggests~~ recommend the use of BFD echo [RFC5880]. This document acknowledges TR-146 and recommends the use of BFD echo for health

Commentaire [Med2]: You may indicate that some mechanisms are also enabled at the BNG side (add a reference to the BBF TR).

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). As a reminder, DHCP options were proposed in the past to provision BFD-related parameters [I-D.vinokour-bfd-dhcp].

Commentaire [Med3]: Please note that BBF also ACKs the following:

"the BFD protocol stack is not likely to be found on most end-hosts".

Commentaire [Med4]: The use of BFD will require some extra provisioning information (e.g., local discriminator). Those are discussed in that draft.

Adding a statement that some BFD specific information need to be provisioned (by some means) would be helpful.

- 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 (Section xx).

An IPoE client MAY be locally configured for IPoE health checks. Non-default local parameters SHOULD override any dynamically signalled parameters (e.g., 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.

IPoE Health Checks parameters can be configured by other means such as TR-069. It is out of scope of this document to define required extensions to TR-069.

3.1. Parameters

IPoE Health Check ~~specifies~~ uses the following parameters:

- Interval (Integer): The frequency, in seconds, which health checks are sent by the IPoE client.
- Limit (Integer): The number of consecutive checks that can fail before an action is taken.
- Behaviour (Integer): Specifies what additional actions are to be taken when triggered.
- Passive (Boolean): Forces passive health checks instead of active.
- Alternative Target Address (IP address): Overrides the default gateway as the target of health checks.
- Check Method: This parameter indicates the method to be used for checks (0: BFD, 1:ND/ARP).

3.2. BFD Echo

Unless instructed otherwise, ~~An~~ an IPoE client SHOULD use BFD Echo [RFC5880] as the health check mechanism.

Commentaire [Med5]: Do you mean planned maintenance?

Commentaire [Med6]: Is this really needed?

What about side effects on the network (overload, for example) when inappropriate timers are used?

Commentaire [Med7]: What is the rationale here?

Commentaire [Med8]: This text should be positioned right after introducing the parameters.

Commentaire [Med9]: Shouldn't default values be proposed?

Commentaire [Med10]: Shouldn't this be further specified to cover the following cases?

-time interval to follow between two successful checks (e.g., 120s).
-retry time upon check failure (e.g., 10s).

Mis en forme : Avec puces + Niveau : 1 + Alignement : 0,63 cm + Retrait : 1,27 cm

Commentaire [Med11]: This is useful when multiple check methods are available.

Mis en forme : Retrait : Gauche : 1,27 cm

Mis en forme : Avec puces + Niveau : 1 + Alignement : 0,63 cm + Retrait : 1,27 cm

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 or IPoE client is explicitly configured, it MUST send Neighbor Solicits (Section 4.3 of [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 or IPoE client is explicitly configured, 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 target IP address MAY be included ~~within~~ in 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~~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 (Section 3.1). If the neighbor entry
is in state INCOMPLETE for Limit ~~(Section 3.1)~~ checks (Section 3.1), 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~~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~~ these 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.

The 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.

Mis en forme : Surlignage

4.1. Behaviour 0: Renew (Default)

After Limit (Section 3.1) consecutive check failures each transmitted after the expiry of Retry timer, 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 IPoE Health Check DHCP ~~Option~~ 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 single-homed router, would withdraw itself as a default router on the LAN, following requirement G-5 of [RFC7084], Section 4.1.

5. IPoE Health Check DHCP Options

~~IPoE Health Check~~This document 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.

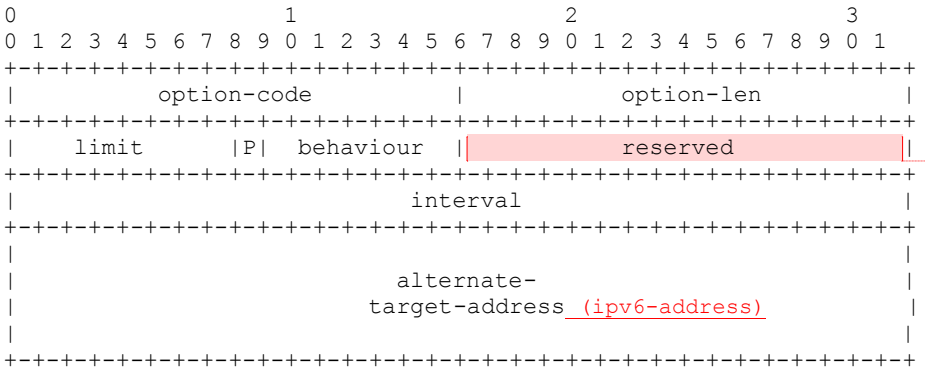

Figure 1: DHCPv6 IPoE Check Option Format

The description of the fields is as follows:

Commentaire [Med12]: I would position the option before the description of the behaviors.

Commentaire [Med13]: Some text may be needed to describe the behavior of the DHCP client.

Commentaire [Med14]: Update with a field to indicate the check method

option-code: OPTION_IPOE_HEALTH (~~TBD~~TBA1).
option-len: 24.
limit: Consecutive failed checks, before an action is taken.
P: Passive Flag. Force passive-only health checks.
behaviour: The following behaviors are defined: ~~Behaviour field.~~
0: Trigger Renew (Section XX).
1: Trigger Rebind (Section XX).
2: Expire lease, start solicit phase (Section XX).
3: Release (Section XX).
4 - 127: ~~Reserved~~Unassigned. New behavior codes can be assigned in the future (see Section XX).
interval: Indicates ~~How~~ how often a health check should be sent.
Expressed in units of seconds.
alternate-target-address: Optional health check target address.
MUST ~~Always~~ always be present~~,~~ ; it MUST be set to zero if no alternate IP address is to be used~~provided~~.

Commentaire [Med15]: see the updated IANA section.

~~Figure 2~~

5.2. DHCPv4

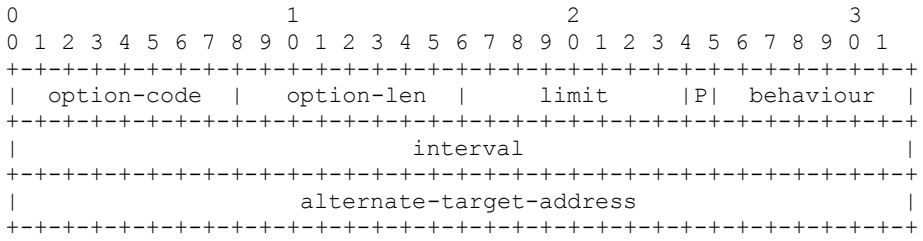

Figure 3: DHCPv4 IPoE Check Option Format

The description of the fields is as follows:

option-code: OPTION_IPOE_HEALTH (TBA2~~D~~).
option-len: 10.
limit: Consecutive failed checks, before an action is taken.
P: Passive Flag. Force passive-only health checks.
behaviour: The following behaviors are defined:~~Behaviour field.~~
0: Trigger Renew (Section XX).
1: Trigger Rebind (Section XX).
2: Expire lease, start discovery phase (Section XX).
3: Release (Section XX).
4 - 127: ~~Reserved~~Unassigned.
interval: Indicates ~~How~~ how often a health check should be sent.
Expressed in units of seconds.
alternate-target-address: Optional health check target address.
MUST ~~Always~~ always be present~~,~~ ; it MUST be set to zero if no alternate IP address is to be used~~ provided~~.

~~Figure 4~~

6. Multihomed Clients

An IPoE client ~~MAY~~ may have multiple leases from the same, or different DHCP servers. These leases ~~MAY~~ 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.

Mis en forme : Surlignage

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~~ network using DHCP ~~Options~~options, but overly zealous, locally configured
IPoE clients, could have an adverse impact. For example, this may induce an overload on the IP access nodes. Means to rate limit such traffic must be enabled at the network side.

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.

Commentaire [Med16]: What is the point of this sentence?

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.~~

IANA is requested to assign a new DHCPv6 Option Code in the registry maintained in <http://www.iana.org/assignments/dhcpv6-parameters>:

Option Name	Value
OPTION_IPOE_HEALTH	TBA1

Also, IANA is requested to assign the following new DHCPv4 Option Code in the registry maintained in <http://www.iana.org/assignments/bootp-dhcp-parameters>:

Option Name	Tag	Data Length	Meaning
OPTION_IPOE_HEALTH	TBA2	10	Provides a set of IPoE check configuration information.

This document requests IANA to create a new registry called "IPoE Check Behaviors" (under https://www.iana.org/assignments/dhcpv6-parameters/dhcpv6-parameters.xhtml). This registry is initially populated with the following values:

Value	Description	Reference
0	Trigger Renew	[ThisRFC]
1	Trigger Rebind	[ThisRFC]
2	Expire lease	[ThisRFC]
3	Release	[ThisRFC]

Values in the range 4-127 are assigned via the "IETF Review" policy defined in [RFC8126].

The same registry is used for both DHCPv4 and DHCPv6.

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.

- Added reference to TR-146.
- 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.

- Emphasised preference for use of BFD echo as the health check mechanism.
- Removed lifetime expiration from Behaviour 2 and clarified usage.
- Updated Behaviour 3 with instructions for whilst mid-renew/rebind.

- Reworded multihoming section.
- 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