PCP Working Group
Internet-Draft
Intended status: Standards Track
Expires: November 4, 2012

M. Boucadair
France Telecom
R. Penno
D. Wing
Cisco

May 3, 2012

DHCP Options for the Port Control Protocol (PCP)
draft-ietf-pcp-dhcp-03

Abstract

This document specifies DHCP (IPv4 and IPv6) options to configure hosts with Port Control Protocol (PCP) Server names. The use of DHCPv4 or DHCPv6 depends on the PCP deployment scenario.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

Status of this Memo

Copyright Notice

Table of Contents

1. Introduction

This document defines DHCPv4 [RFC2131] and DHCPv6 [RFC3315] options which can be used to provision PCP Server [I-D.ietf-pcp-base] names.

In order to make use of these options, this document assumes appropriate name resolution means (e.g., Section 6.1.1 of [RFC1123]) are available on the host client.

The use of DHCPv4 or DHCPv6 depends on the PCP deployment scenarios.

2. Terminology

This document makes use of the following terms:

- PCP Server denotes a functional element which receives and processes PCP requests from a PCP Client. A PCP Server can be co-located with or be separated from the function (e.g., NAT, Firewall) it controls. Refer to [I-D.ietf-pcp-base].
- PCP Client denotes a PCP software instance responsible for issuing PCP requests to a PCP Server. Refer to [I-D.ietf-pcp-base].
- DHCPv4 refers to the Dynamic Host Configuration Protocol [RFC2131] for IPv4.
- DHCP refers to both DHCPv4 [RFC2131] and DHCPv6 [RFC3315].
- DHCP client (or client) denotes a node that initiates requests to obtain configuration parameters from one or more DHCP servers [RFC3315].
- DHCP server (or server) refers to a node that responds to requests from DHCP clients [RFC3315].
- Name is a domain name (as per Section 3.1 of [RFC1035]) that contains one or more labels. In particular, a PCP name may be structured as DNS qualified name or be composed of strings such as can be passed to getaddrinfo (Section 6.1 of [RFC3493]), including address literals, etc.

3. Rationale

Both IP Address and Name DHCP options have been considered in early stages of this specification. This flexibility aims to let service providers to make their own engineering choices and use the convenient option according to their deployment context. Nevertheless, DHC WG's position is this flexibility have some drawbacks such as inducing errors. Therefore, only the Name option is maintained within this document.

This document defines an option to carry a name rather than an IP

address. This choice is motivated by operational considerations: In particular, some Service Providers are considering two levels of redirection:

(1) The first level is national-wise and undertaken by DHCP: a regional-specific Name will be returned;
(2) The second level is done during the resolution of the regional-specific Name to redirect the customer to a regional PCP server among a pool deployed regionally.

Distinct operational teams are responsible for each of the above mentioned levels. A clear separation between the functional perimeter of each team is a sensitive task for the maintenance of the offered services. Regional teams will require to introduce new resources (e.g., new PCP-controlled devices such as Carrier Grade NATs (CGNs, [I-D.ietf-behave-lsn-requirements])) to meet an increase of customer base. Operations related to the introduction of these new devices (e.g., addressing, redirection, etc.) are implemented locally. Having this regional separation provides flexibility to manage portions of network operated by dedicated teams. This two-level redirection can not be met by the IP Address option.

In addition to the operational considerations:

- The use of the Name for NAT64 [RFC6146] might be suitable for load-balancing purposes;
- For the DS-Lite case [RFC6333], if the encapsulation mode is used to send PCP messages, an IP address may be used since the AFTR selection is already done via the AFTR_NAME DHCPv6 option [RFC6334]. Of course, this assumes that the PCP Server is co-located with the AFTR function. If these functions are not co-located, conveying the Name would be more convenient.

4. DHCPv6 PCP Server Option

This DHCPv6 option conveys a domain name to be used to retrieve the IP addresses of PCP Server(s). Appropriate name resolution queries should be issued to resolve the conveyed name.

4.1. Format

The format of the DHCPv6 PCP Server option is shown in Figure 1.

```
  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_PCP_SERVER        |         Option-length         |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                                                               |
 :                     PCP Server Domain Name                    :
 |                                                               |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: PCP Server Name DHCPv6 Option

The fields of the option shown in Figure 1 are as follows:

- Option-code: OPTION_PCP_SERVER (TBA, see Section 9.1)
- Option-length: Length of the 'PCP Server Domain Name' field in octets.
- PCP Server Domain Name: The domain name of the PCP Server to be used by the PCP Client. The domain name is encoded as specified in Section 8 of [RFC3315].

The OPTION_PCP_SERVER option can include multiple PCP Server Domain Names; each Name is treated as a separate PCP Server.

4.2. Client Behaviour

To discover a PCP Server [I-D.ietf-pcp-base], the DHCPv6 client MUST include an Option Request Option (ORO) requesting the DHCPv6 PCP Server Name option as described in Section 22.7 of [RFC3315] (i.e., include OPTION_PCP_SERVER on its OPTION_ORO).

If the DHCPv6 client receives an OPTION_PCP_SERVER option from the DHCPv6 server, it extracts the Name(s) conveyed in the OPTION_PCP_SERVER option and proceeds to validating it. The DHCPv6 client MUST verify that the option length does not exceed 255 octets ([RFC1035]). The DHCPv6 client MUST verify the name(s) is properly encoded as detailed in Section 8 of [RFC3315].

Once each Name conveyed in the OPTION_PCP_SERVER option is validated, each included Name is passed to the name resolution library (e.g., Section 6.1.1 of [RFC1123] or [RFC6055]) to retrieve the corresponding IP address(es) (IPv4 or IPv6). Then, the PCP Client MUST follow the procedure specified in Section 6 to contact its PCP Server(s).

It is RECOMMENDED to associate a validity lifetime with any address resulting from resolving the Name conveyed in a OPTION_PCP_SERVER DHCPv6 option when stored in a local name resolution cache. Considerations on how to flush out a local cache are out of the scope

of this document.

A host may have multiple network interfaces (e.g, 3G, WiFi, etc.); each configured differently. Each PCP Server learned MUST be associated with the interface via which it was learned. When an application issues a PCP request to a PCP Server, the source address of the request MUST be among those assigned on the interface to which the destination PCP Server is bound.

5. DHCPv4 PCP Option

5.1. Format

The PCP Server Name DHCPv4 option can be used to configure a name to be used by the PCP Client to contact a PCP Server. The format of this option is illustrated in Figure 2.

```
       Code  Length   PCP Server Domain Name
      +-----+-----+-----+-----+-----+-----+-----+--
      | TBA |  n  |  s1 |  s2 |  s3 |  s4 | s5  |  ...
      +-----+-----+-----+-----+-----+-----+-----+--
```

The values s1, s2, s3, etc. represent the domain name labels in the domain name encoding.

Figure 2: PCP Server Name DHCPv4 Option

The description of the fields is as follows:

- Code: OPTION_PCP_SERVER (TBA, see Section 9.2);
- Length: Includes the length of the "PCP Server Domain Name" field in octets; The maximum length is 255 octets.
- PCP Server Domain Name: The domain name of the PCP Server to be used by the PCP Client when issuing PCP messages. The encoding of the domain name is described in Section 3.1 of [RFC1035].

The OPTION_PCP_SERVER option can include multiple PCP Server Domain Names; each Name is treated as a separate PCP Server.

5.2. Client Behaviour

DHCPv4 client expresses the intent to get OPTION_PCP_SERVER by specifying it in Parameter Request List Option [RFC2132].

If the DHCPv4 client receives an OPTION_PCP_SERVER option from the DHCPv4 server, it extracts the Name(s) conveyed in the option and

proceeds to validating it. The DHCPv4 client MUST verify that the option length does not exceed 255 octets [RFC1035]). If more than one Name is included in a OPTION_PCP_SERVER option, and once each name conveyed in the OPTION_PCP_SERVER option is validated, each included Name is passed to the name resolution library (e.g., Section 6.1.1 of [RFC1123] or [RFC6055]) to retrieve the corresponding IPv4 address(es).

The PCP Client MUST follow the procedure specified in Section 6 to contact its PCP Server(s).

It is RECOMMENDED to associate a validity lifetime with any address resulting from resolving the Name conveyed in a OPTION_PCP_SERVER DHCPv4 option when stored in a local name resolution cache. Considerations on how to flush out a local cache are out of the scope of this document.

A host may have multiple network interfaces (e.g, 3G, WiFi, etc.); each configured differently. Each PCP Server learned MUST be associated with the interface via which it was learned. When an application issues a PCP request to a PCP Server, the source address of the request MUST be among those assigned on the interface to which the destination PCP Server is bound.

6. IP Address Selection

This section specifies the behavior to be followed by the PCP Client to contact its PCP Server(s) when receiving one or several PCP Names:

1. If only one PCP Name is received: if a list of IP addresses is returned as a result of resolving the name conveyed in the PCP Name DHCP option, the PCP Client follows the procedure specified in Section 6.1.
2. If several PCP Names are received: each Name is treated as a separate PCP Server. Moreover, each Name may be resolved into one IP address or a list of IP addresses. The PCP Client contacts in parallel the first IP address of each Name and follows the procedure specified in Section 6.1 for the list of IP addresses returned for each Name. Section 6.2 provides some examples to illustrate this procedure.

6.1. Serial Queries

The PCP Client initializes its retransmission timer, RETRY_TIMER, to 2 seconds. The PCP Client sends its PCP message to the PCP Server and waits 2 seconds for a response. If no response is received, it doubles the value of RETRY_TIMER, sends another (identical) PCP

message and waits 2*RETRY_TIMER. This procedure is repeated three (3) times, doubling the value of RETRY_TIMER each time. If no response is received after four (4) attempts, the PCP Client tries with the next IP address in its list of PCP Servers. If it has exhausted its list, the procedure is repeated every fifteen minutes until the PCP request is successfully answered. If, when sending PCP requests the PCP Client receives an ICMP error (e.g., port unreachable, network unreachable) it SHOULD immediately try the next IP address in the list. Once the PCP Client has successfully received a response from a PCP Server on that interface, it sends subsequent PCP requests to that same server until that PCP Server becomes non-responsive, which causes the PCP client to attempt to re-iterate the procedure starting with the first PCP Server on its list.

6.2. Examples

Let's suppose pcpserver-x, pcpserver-y and pcpserver-z are returned as PCP Names in a OPTION_PCP_SERVER option. Let's also suppose:

* IPx1 and IPx2 are returned for pcpserver-x; IPx1 is not reachable.
* IPy1 and IPy2 are returned for pcpserver-y; IPy1 is reachable
* IPz1 and IPz2 are returned for pcpserver-z; IPz1 is reachable

The procedure to contact the PCP Servers is as follows:

* Send PCP requests to all servers: IPx1, IPy1 and IPz1
* Responses are received from IPy1 and IPz1 but not from IPx1
  - The request is re-sent to IPx1
  - If no response is received after four attempts, the request is sent to IPx2

Now, if the following conditions are made:

* IPx1 and IPx2 are returned for pcpserver-x; IPx1 is not reachable.
* IPy1 and IPy2 are returned for pcpserver-y; IPy1 is reachable
* IPz1 and IPz2 are returned for pcpserver-z; IPz1 is not reachable

The procedure to contact the PCP Servers lead to the following:

* Send PCP requests to all servers: IPx1, IPy1 and IPz1
* A response is received from IPy1 but not from IPx1 and IPz1
  - the requests are re-sent to IPx1 and IPz1
  - If no response is received after four attempts, the request is then sent to IPx2 and IPz2

Let's suppose now that:

* IPx1 and IPx2 are returned for pcpserver-x; IPx1 is not reachable.
* IPy1 and IPy2 are returned for pcpserver-y; IPy1 is not reachable
* IPz1 and IPz2 are returned for pcpserver-z; IPz1 is not reachable

The procedure to contact the PCP Servers is as follows:

* Send PCP requests to all servers: IPx1, IPy1 and IPz1
* No answer is received for all requests
  - the requests are re-sent to IPx1, IPy1 and IPz1
  - If no response is received after four attempts, the request is then sent to IPx2, IPy2 and IPz2

7. Dual-Stack Hosts

A PCP Server configured using OPTION_PCP_SERVER over DHCPv4 is likely to be resolved to IPv4 address(es).

A PCP Server configured using OPTION_PCP_SERVER over DHCPv6 may be resolved to IPv4-mapped IPv6 address(es) or IPv6 address(es) (e.g., NAT64 [RFC6146], IPv6 firewall [RFC6092], NPTv6 [RFC6296]).

In some deployment contexts, the PCP Server may be reachable with an IPv4 address but DHCPv6 is used to provision the PCP Client. In such scenarios, a plain IPv4 address or an IPv4-mapped IPv6 address can be configured to reach the PCP Server.

A Dual-Stack host may receive OPTION_PCP_SERVER via both DHCPv4 and DHCPv6. The content of these OPTION_PCP_SERVER options may refer to the same or distinct PCP Servers. This is deployment-specific and as such it is out of scope of this document.

8. Security Considerations

The security considerations in [RFC2131], [RFC3315] and [I-D.ietf-pcp-base] are to be considered.

9. IANA Considerations

9.1. DHCPv6 Option

Authors of this document request the following DHCPv6 option code:

Option Name	Value
OPTION_PCP_SERVER	TBA

9.2. DHCPv4 Option

Authors of this document request the following DHCPv4 option code:

Option Name	Value
OPTION_PCP_SERVER	TBA

10. Acknowledgements

Many thanks to B. Volz, C. Jacquenet, R. Maglione, D. Thaler, T. Mrugalski, T. Lemon and M. Wasserman for their review and comments.

11. References

11.1. Normative References

[I-D.ietf-pcp-base]
Wing, D., Cheshire, S., Boucadair, M., Penno, R., and P. Selkirk, "Port Control Protocol (PCP)", draft-ietf-pcp-base-24 (work in progress), March 2012.

[RFC1035] Mockapetris, P., "Domain names - implementation and specification", STD 13, RFC 1035, November 1987.

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

[RFC2131] Droms, R., "Dynamic Host Configuration Protocol", RFC 2131, March 1997.

[RFC2132] Alexander, S. and R. Droms, "DHCP Options and BOOTP Vendor Extensions", RFC 2132, March 1997.

[RFC3315] Droms, R., Bound, J., Volz, B., Lemon, T., Perkins, C., and M. Carney, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3315, July 2003.

11.2. Informative References

[I-D.ietf-behave-lsn-requirements]
Perreault, S., Yamagata, I., Miyakawa, S., Nakagawa, A.,

and H. Ashida, "Common requirements for Carrier Grade NATs (CGNs)", draft-ietf-behave-lsn-requirements-05 (work in progress), November 2011.

[RFC1123] Braden, R., "Requirements for Internet Hosts - Application and Support", STD 3, RFC 1123, October 1989.

[RFC3493] Gilligan, R., Thomson, S., Bound, J., McCann, J., and W. Stevens, "Basic Socket Interface Extensions for IPv6", RFC 3493, February 2003.

[RFC6055] Thaler, D., Klensin, J., and S. Cheshire, "IAB Thoughts on Encodings for Internationalized Domain Names", RFC 6055, February 2011.

[RFC6092] Woodyatt, J., "Recommended Simple Security Capabilities in Customer Premises Equipment (CPE) for Providing Residential IPv6 Internet Service", RFC 6092, January 2011.

[RFC6146] Bagnulo, M., Matthews, P., and I. van Beijnum, "Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers", RFC 6146, April 2011.

[RFC6296] Wasserman, M. and F. Baker, "IPv6-to-IPv6 Network Prefix Translation", RFC 6296, June 2011.

[RFC6333] Durand, A., Droms, R., Woodyatt, J., and Y. Lee, "Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion", RFC 6333, August 2011.

[RFC6334] Hankins, D. and T. Mrugalski, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6) Option for Dual-Stack Lite", RFC 6334, August 2011.

Authors' Addresses

Mohamed Boucadair
France Telecom
Rennes,   35000
France

Email: mohamed.boucadair@orange.com

Reinaldo Penno
Cisco
USA

Email: repenno@cisco.com

Dan Wing
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, California 95134
USA

Email: dwing@cisco.com