PCP Working Group
Internet-Draft
Intended status: Standards Track
Expires: February 7, 2013

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

August 6, 2012

DHCP Options for the Port Control Protocol (PCP)
draft-ietf-pcp-dhcp-04

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

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 http://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 February 7, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of

publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

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.
- DHCP server (or server) refers to a node that responds to requests from DHCP clients.
- Name is a domain name 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 has some drawbacks such as inducing errors (See Section 7 of [I-D.ietf-dhc-option-guidelines]). 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.

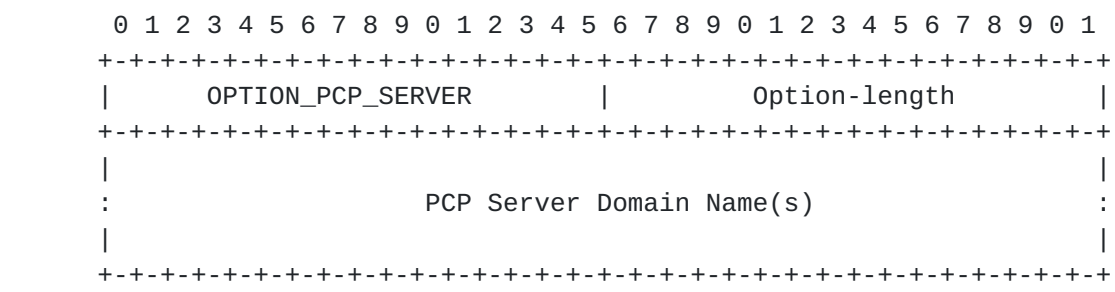

Figure 1: PCP Server Name DHCPv6 Option

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(s): The domain name s) of the PCP Server to be used by the PCP Client. The OPTION_PCP_SERVER option can include multiple PCP Server Domain Names; each Name is treated as a separate PCP Server. The domain name(s) is encoded as string. When several names are included, a space character is used as separator.

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 validate it.

Once each Name conveyed in the OPTION_PCP_SERVER option is validated, the DHCPv6 client MUST follow the procedure specified in Section 6.

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.

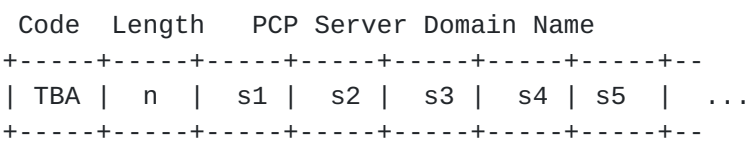

Figure 2: PCP Server Name DHCPv4 Option

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(s): The domain name(s) of the PCP Server to be used by the PCP Client when issuing PCP messages. The OPTION_PCP_SERVER option can include multiple PCP Server Domain Names; each Name is treated as a separate PCP Server. The domain name(s) is encoded as strings. When several names are included, a space character is used as separator.

The OPTION_PCP_SERVER DHCPv4 option is a concatenation-requiring option. As such, the mechanism specified in [RFC3396] MUST be used if the PCP Server Name option exceeds the maximum DHCPv4 option size of 255 octets.

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.

Once each Name conveyed in the OPTION_PCP_SERVER option is validated, the DHCPv4 client MUST follow the procedure specified in Section 6.

6. Use of PCP Server Names

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. This section is not specific to DHCP; it is applicable to any mechanism that configures server names.

Multiple Names may be configured to a PCP Client in some deployment contexts such as multi-homing. It is out of scope of this document to enumerate all deployment scenarios which require multiple Names to be configured.

6.1. Name Resolution

Each configured 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 6.2 to contact its PCP Server(s).

It is RECOMMENDED to associate a validity lifetime (e.g., TTL of DNS record if the Name is resolved using DNS) with any address resulting from resolving the PCP Server Name 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.

6.2. 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 configured: if a list of IP addresses is returned as a result of resolving the PCP Server Name, the PCP Client follows the procedure specified in Section 6.2.1.
2. If several PCP Names are configured: 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 6.2.1 for the list of IP addresses returned for each Name. Section 6.2.2 provides some examples to illustrate this procedure.

The discovery procedure may result in a PCP Client instantiating multiple mappings maintained by distinct PCP Servers. The decision to use all these mappings or delete some of them is deployment-specific. Only the client can decide whether all the mappings are needed or only a subset of them.

6.2.1. 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 Server addresses. 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 address on that interface, it sends subsequent PCP requests to that same server address 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 address on its list.

6.2.2. Examples

The following sub-sections provide three examples to illustrate the procedure.

For all these examples, let's suppose pcpserver-x, pcpserver-y and pcpserver-z are configured as PCP Names.

6.2.2.1. Example 1

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

6.2.2.2. Example 2

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

6.2.2.3. Example 3

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

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-26 (work in progress), June 2012.

[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.

[RFC3396] Lemon, T. and S. Cheshire, "Encoding Long Options in the Dynamic Host Configuration Protocol (DHCPv4)", RFC 3396, November 2002.

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-08 (work in progress), July 2012.

[I-D.ietf-dhc-option-guidelines]
Hankins, D., Mrugalski, T., Siodelski, M., Jiang, S., and S. Krishnan, "Guidelines for Creating New DHCPv6 Options", draft-ietf-dhc-option-guidelines-08 (work in progress), June 2012.

[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.

[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.

[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