DHC Working Group
Internet-Draft
Intended status: Standards Track
Expires: October 12, 2013

Q. Sun
Tsinghua University
Y. Lee
Comcast
Q. Sun
China Telecom
G. Bajko
Nokia
M. Boucadair
France Telecom
April 10, 2013

# Dynamic Host Configuration Protocol (DHCP) Option for Port Set Assignment
draft-sun-dhc-port-set-option-01

## Abstract

Because of the exhaustion of the IPv4 address space, several techniques have been proposed to share the same IPv4 address among several uses.  As an alternative to introducing a level of NAT in the provider's core network, this document provides a mechanism to assign non-overlapping port set to users assigned with the same IPv4 address: Port Set DHCPv4 Option.

## Status of this Memo

This Internet-Draft will expire on October 12, 2013.

## Copyright Notice

Table of Contents

## 1. Introduction

Currently some large ISPs still have a large enough IPv4 address pool to be able to allocate public IPv4 addresses for their subscribers. However, due to the exhaustion of the global IPv4 address space, these ISP expect the situation is unsustainable and they will not be able anymore to assign to every requesting host a public IPv4 address.

Two solutions have been proposed so far: (1) Deploy Network Address Translation (NAT) or (2) Allocate the same public IPv4 address with non-overlapped port sets directly to multiple connected devices (which can be CPEs or end hosts). This document focuses on the second solution.

This document describes a new DHCPv4 option which allows the DHCPv4 server to assign a set of ports to a user device during the IPv4 address provisioning process. By assigning the same IPv4 address with non-overlapped port sets to multiple clients, the clients is enabled to share the IPv4 address and continue to deliver IPv4 services to subscribers.

The Port Set Option described in this document can be used in various deployment scenarios, some of which are described in [RFC6346]

## 2. 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 [RFC2119].

## 3. DHCPv4 Port Set Option

### 3.1. Port Set Option Format

The format of Port Set Option is shown in Figure 1.

```
            0                             1
            0  1  2  3  4  5  6  7  8  9  0  1  2  3  4  5
           +--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
           |   OPTION_PORT_SET     |     option-length     |
           +--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
           |                Port Set Index                 |
           +--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
           |                Port Set Mask                  |
           +--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
```

Figure 1 Port Set Option Format

- option-code: OPTION_PORT_SET (TBD)

- option-length: An 8-bit field indicating the length of the option excluding the 'Option Code' and the 'Option Length' fields. In this option, the option-length is 4 octets.

- Port Set Index: Port Set Index identifies a set of ports assigned to a device. The first k bits on the left of the 2-octet field is the Port Set Index value, with the rest of the field right padding zeros.

- Port Set Mask: Port Set Mask indicates the position of the bits used to build the mask. The first k bits on the left is padding ones while the remained (16-k) bits of the 2-octet field on the right is padding zeros.

In the context of Port Set Option, the port number should consist of port set prefix and port number suffix. The port set prefix can be got from Port Set Index and Port Set Mask, while port number suffix can change continuously. The format of port number is shown in Figure 2.

```
           0                                                15
           +-----------------------+-----------------------------+
           |    port set prefix    |      port number suffix     |
           +-----------------------+-----------------------------+
           |<-------k bits-------->|<--------(16-k) bits-------->|
```

Figure 2 Bit Representation of a port number

In order to exclude the system ports ([I-D.ietf-tsvwg-iana-ports]) or ports saved by SPs, the former port-sets that contains well-known ports SHOULD NOT be assigned.

For example: If k is 10 (the left 10 bits of Port Set Mask is '1'), the first 16 port sets is located in well-known port space, which

should not be allocated.  Or,

For example: If k is 4 (the left 4 bits of Port Set Mask is '1'), the first port set (0 - 4095) contains the well-know port space.  It should be perceived as well.

### 3.2.  Port Set Option Example

The Port Set Option is used to specify one contiguous port set pertaining to the given IP address.

Concretely, this option is used to notify a remote DHCP client about the port set prefix to be applied when selecting a port value as a source port.  The Port Set Option is used to infer a set of allowed contiguous port values.  Two port numbers are said to belong to the same Port Set if and only if, they have the same port set prefix.

The following Port Set Index and Port Set Mask are conveyed using DHCP to assign a contiguous port set with excluding well-know ports (with Port Set Index not zero):

Port Set Index: 0001 0100 0000 0000 (5120)

Port Set Mask: 1111 1100 0000 0000 (64512)

The device will get a contiguous port set: 5120 - 6143

## 4.  Server Behavior

The server will not reply with the option until the client has explicitly listed the option code in the Parameter Request List (Option 55).

Server MUST reply with Port Set Option if the client requested OPTION_PORT_SET in its Parameter Request List.  The server MUST run an address & port-set pool which plays the same role as address pool in regular DHCP server.  The address and port-set pool MUST follow the Port-Mask-format port-set.

If the server receives a DHCPDISCOVER message containing a Port Set Option, this means the client is requesting a specific port set.  The Port Set Mask field in the option indicates the size of port set that the client requests.  The server MAY reply with a Port Set Option whose Port Set Mask is as requested, if the server has such one port set.  Or the server can ignore the request and just assign a port set from the pool.

The port-set assignment SHOULD be coupled with the address assignment process. Therefore server SHOULD assign the address and port set in the same DHCP messages. And the lease information for the address is applicable to the port-set as well.

## 5. Client Behavior

The DHCP client applying for the a port-set MUST include either the OPTION_PORT_SET code in the Parameter Request List (Option 55). The client will retrieve a Port Set Option and use the Port Set Index and Port Set Mask to perform the port mask algorithm to get the contiguous port set. The client renews or releases the DHCP lease with the port set.

The client MAY include a Port Set Option in the DHCPDISCOVER message, in which the Port Set Mask field indicates the requested size of a port set from the client.

## 6. DHCP Unicast Considerations

DHCP messages could be unicasted over UDP port 67. In the context of address sharing, not all the ports are available to the clients. The server cannot use unicast to send the DHCP message to a client which originated the DHCP request. To mitigate this problem, we propose to use the broadcast address (0.0.0.0) when the server replies to the client. Broadcast address is special and won't be assigned to any client.

### 6.1. Server Behavior

DHCP server MUST set broadcast bit of the 'flags' field in DHCP messages (Figure 2 of [RFC2131]) when allocating port sets. And DHCP server MUST NOT unicast responses to DHCP client. In order to identify the DHCP responses are sent to which client, client identifier [I-D.ietf-dhc-client-id] is used. DHCP server MUST return client identifier.

### 6.2. Client Behavior

DHCP client MUST validate client identifier, as specified in [I-D.ietf-dhc-client-id]. DHCP client MUST NOT unicast requests to server: all requests are broadcast. This includes lease renewals. In the case of DHCP relay agent, it will broadcast the server responses to clients.

In some deployment scenarios, DHCP messages containing the proposed

DHCP option can be conveyed by other forwarding carrier than IPv4, saying IPv6 [I-D.ietf-dhc-dhcpv4-over-ipv6], [I-D.scskf-dhc-dhcpv4-over-dhcpv6], etc. The server has to manage to forward DHCP responses to right client.

## 7. Security Consideration

### 7.1. Denial-of-Service

The solution is generally vulnerable to DoS when used in shared medium or when access network authentication is not a prerequisite to IP address assignment. The solution SHOULD only be used on point-to-point links, tunnels, and/or in environments where authentication at link layer is performed before IP address assignment, and not shared medium.

### 7.2. Port Randomization

Preserving port randomization [RFC6056] may be more or less difficult depending on the address sharing ratio (i.e., the size of the port space assigned to a CPE). The host can only randomize the ports inside a fixed port range [RFC6269].

More discussion to improve the robustness of TCP against Blind In-Window Attacks can be found at [RFC5961]. Other means than the (IPv4) source port randomization to provide protection against attacks should be used (e.g., use [I-D.vixie-dnsext-dns0x20] to protect against DNS attacks, [RFC5961] to improve the robustness of TCP against Blind In-Window Attacks, use IPv6).

A proposal to preserve the entropy when selecting port is discussed in [I-D.bajko-pripaddrassign]

## 8. IANA Consideration

IANA is kindly requested to allocate DHCP option code to the OPTION_PORT_SET. The code should be added to the DHCP option code space.

## 9. Contributors List

Many thanks for valuable comments and great efforts from the following contributors:

Peng Wu
Tsinghua University

peng-wu@foxmail.com

Teemu Savolainen
Nokia

teemu.savolainen@nokia.com

Ted Lemon
Nominum, Inc.

mellon@nominum.com

Tina Tsou
Huawei Technologies

tena@huawei.com

Pierre Levis
France Telecom

Email: pierre.levis@orange.com

Cong Liu
Tsinghua University

Email: gnocuil@gmail.com

## 10. References

### 10.1. Normative References

[RFC1918] Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, February 1996.

[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.

[RFC3046] Patrick, M., "DHCP Relay Agent Information Option", RFC 3046, January 2001.

[RFC3527] Kinnear, K., Stapp, M., Johnson, R., and J. Kumarasamy, "Link Selection sub-option for the Relay Agent Information Option for DHCPv4", RFC 3527, April 2003.

[RFC4925] Li, X., Dawkins, S., Ward, D., and A. Durand, "Softwire Problem Statement", RFC 4925, July 2007.

[RFC5961] Ramaiah, A., Stewart, R., and M. Dalal, "Improving TCP's Robustness to Blind In-Window Attacks", RFC 5961, August 2010.

[RFC6056] Larsen, M. and F. Gont, "Recommendations for Transport-Protocol Port Randomization", BCP 156, RFC 6056, January 2011.

[RFC6269] Ford, M., Boucadair, M., Durand, A., Levis, P., and P. Roberts, "Issues with IP Address Sharing", RFC 6269, June 2011.

[RFC6346] Bush, R., "The Address plus Port (A+P) Approach to the IPv4 Address Shortage", RFC 6346, August 2011.

### 10.2. Informative References

[I-D.bajko-pripaddrassign] Bajko, G., Savolainen, T., Boucadair, M., and P. Levis, "Port Restricted IP Address Assignment", draft-bajko-pripaddrassign-04 (work in progress), April 2012.

[I-D.ietf-dhc-client-id] Swamy, N., Halwasia, G., and S. Unit, "Client Identifier Option in DHCP Server Replies",

draft-ietf-dhc-client-id-07 (work in progress), November 2012.

[I-D.ietf-dhc-dhcpv4-over-ipv6]
Cui, Y., Wu, P., Wu, J., and T. Lemon, "DHCPv4 over IPv6 Transport", draft-ietf-dhc-dhcpv4-over-ipv6-06 (work in progress), March 2013.

[I-D.ietf-tsvwg-iana-ports]
Cotton, M., Eggert, L., Touch, J., Westerlund, M., and S. Cheshire, "Internet Assigned Numbers Authority (IANA) Procedures for the Management of the Service Name and Transport Protocol Port Number Registry", draft-ietf-tsvwg-iana-ports-10 (work in progress), February 2011.

[I-D.scskf-dhc-dhcpv4-over-dhcpv6]
Sun, Q., Cui, Y., Siodelski, M., Krishnan, S., and I. Farrer, "DHCPv4 over DHCPv6 Transport", draft-scskf-dhc-dhcpv4-over-dhcpv6-01 (work in progress), April 2013.

[I-D.vixie-dnsext-dns0x20]
Vixie, P. and D. Dagon, "Use of Bit 0x20 in DNS Labels to Improve Transaction Identity", draft-vixie-dnsext-dns0x20-00 (work in progress), March 2008.

## Authors' Addresses

Qi Sun
Tsinghua University
Department of Computer Science, Tsinghua University
Beijing  100084
P.R.China

Phone: +86-10-6278-5822
Email: sunqi@csnet1.cs.tsinghua.edu.cn

Yiu L. Lee
Comcast
One Comcast Center
Philadelphia  PA  19103
USA

Phone:
Email: yiu_lee@cable.comcast.com

Qiong Sun
China Telecom
Room 708, No.118, Xizhimennei Street
Beijing  100035
P.R.China

Phone: +86-10-58552936
Email: sunqiong@ctbri.com.cn

Gabor Bajko
Nokia

Phone:
Email: gabor.Bajko@nokia.com

Mohamed Boucadair
France Telecom
2330 Central Expressway
Rennes  35000
France

Phone:
Email: mohamed.boucadair@orange.com