Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: September 15, 2011

T. Tsou
C. Zhou
Huawei Technologies
Q. Sun
China Telecom
M. Boucadair
France Telecom
G. Bajko
Nokia

March 14, 2011

# Using PCP To Coordinate Between the CGN and Home Gateway Via Port Allocation
draft-tsou-pcp-natcoord-01

## Abstract

Consider a situation where a subscriber's packets are subject to two levels of NAT, with both NATs operating under the control of the ISP. An example of this would be a NATing Home Gateway forwarding packets to a Large Scale NAT. This memo proposes that advantage be taken of the presence of the second NAT, to offload the burden on the Large Scale NAT by delegation to the Home Gateway. Enhancements to the Port Control Protocol are specified to achieve this. The proposed solution applies also for DS-Lite where the AFTR offloads it NAT to the B4 element.

## 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 September 15, 2011.

Table of Contents

## 1. Application Scenario

A Large Scale NAT (LSN) is responsible for translating source addresses and ports for packets passing into and out of the provider network. Especially for large scale service providers, one LSN may need to support at least tens of thousands of customers, resulting in heavy processing requirements for the LSN.

In some broadband scenarios an additional NAT is present at the edge of the customer network. For convenience we will call this the Home Gateway. The load on the LSN could be reduced if address and port translation were actually done at the Home Gateway. Achieving such an outcome would require coordination between the two devices. This memo makes a detailed proposal for the required coordination mechanism.

## 2. Proposed Solution

### 2.1. Delegation of Port Ranges

The basic proposal made in this memo is to provide the means for the Home Gateway to request that the LSN delegate to it a set of ports and optionally an external address that will be associated with those ports. It is proposed to use the Port Control Protocol (PCP) [ID.port-control-protocol] to achieve this. The procedure is illustrated in Figure 1.

> The LSN allocation of port sets MAY take into account the advice given in [ID.behave-natx4-log-reduction].
>
> [Open Issue: if we want to make the port sets discontinuous, we must either allow negotiation of the algorithm or parameters of that algorithm for determining the complete set from a given starting point, or specify it here. Specifying it all here is probably counter-productive, given that this is a security measure to make port guessing harder.]

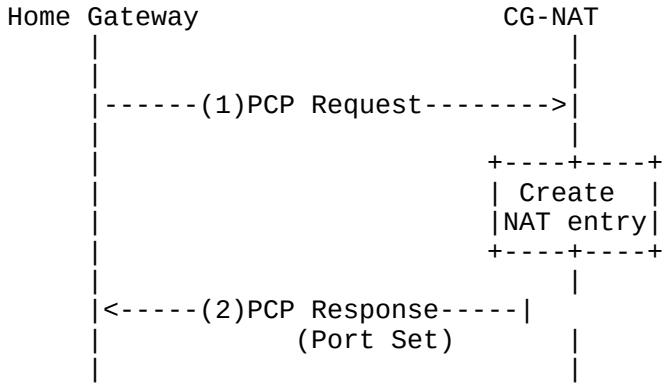


Figure 1: Acquiring a Delegated Port Set

If the Home Gateway allocates all of the ports that have been delegated to it for a given protocol, it MAY send a request to the LSN for another delegated set of ports. If the LSN satisfies that request, the Home Gateway MUST release the additional set as soon as possible. To achieve this, the Home Gateway May follow a policy for allocation of additional ports to flows, that has the same effect as searching for "free" ports in the port sets in the order in which they were delegated to the Home Gateway. A port SHOULD be considered "free" if no traffic has been observed through it for the timeout interval specified for the protocol concerned, as discussed in [ID.behave-natx4-log-reduction], or if the Home Gateway knows through other means (e.g., host reboot) that it is no longer in use.

## 2.2. Packet Processing At the Home Gateway and LSN

The Home Gateway maps outgoing flows to the delegated ports. If an external address was received it uses that for the source address; otherwise it retains the private address of the Home Gateway as the source address.

> The procedures are more complicated, of course, if the IP version running externally to the LSN is different from the IP version running between the Home Gateway and the LSN, since the destination address also has to be translated. The details depend on the particular transition mechanism in use, and are left as an exercise for the reader.

If the private address is retained, the LSN recognizes it from the original delegation request and changes the source address but not the port before forwarding the packet. If the external public address was used, the LSN is not useful and another device may be needed to allocate the port range.

In the reverse direction, the LSN recognizes the public destination address and port of an incoming packet as belonging to a delegated set for the Home Gateway. It translates the destination address, if necessary, leaving the destination port unchanged. The Home Gateway translates the destination port and address to the corresponding values in the customer network and forwards the packet in turn.

## 2.3. Proposed Enhancements To and Usage Of the Port Control Protocol

This document proposes the following new option for MAP opcodes: PORT_SET_REQUESTED.

option number: to be allocated

is valid for OpCodes: MAP44, MAP64, MAP46, or MAP66

is included in responses: MUST

has length: 0 in requests, 4 in successful responses. [As mentioned above, if non-consecutive sets of ports are allocated, we may want to add parameters of the algorithm for deriving the complete set from the initial value provided in the "assigned external port" field of the response.]

may appear more than once: no

When constructing a PCP request with the PORT_SET_REQUESTED option, the client MUST set the "internal port" field of the request to zero. If requesting a new set of delegated ports, the client MAY set the "requested external port" field to a non-zero value. If releasing a set of delegated ports (i.e., by setting the "Requested lifetime" field to zero), the client MUST set the "requested external port" field to the value of the "assigned external port" field of the earlier response from the server. The remaining fields of the PCP request MUST be set as directed by [ID.port-control-protocol]

[Open issue: for a release, should the PORT_SET_REQUESTED option have the same contents as it had in the earlier response?]

Upon receiving a PCP request with the PORT_SET_REQUESTED option, the server MAY reject it using return codes 151 - NOT_AUTHORIZED, or 152 - USER_EX_QUOTA. In this case, the PORT_SET_REQUESTED option in the response MUST have zero length (no data). If the server chooses to honour the request, it MUST place the value of the first port in the assigned set in the "assigned external port" field of the response. It MUST set the length of the PORT_SET_REQUESTED option in the response to 4, and MUST provide the number of ports in the delegated set as the value of the option.

## 3. Port Range Options

The Port Range option is used to specify one range of ports (contiguous or not contiguous) pertaining to a given IP address. The starting point of the ports and the number of delegated ports are used to infer a set of allowed port values. This section provides only one method to request the port range values. Other ways and Optcode can be proposed in later versions.

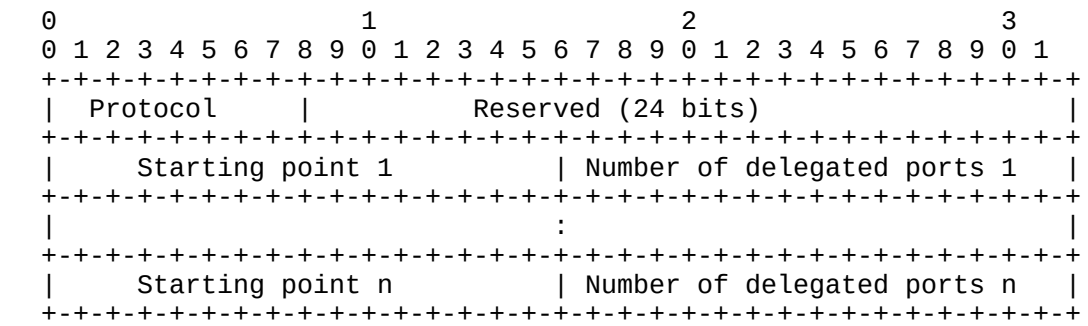


Figure 2: Port_Range_Option

These fields are described as below:

- Starting Port: A 16 bit value used as an input to the specified function.
- Number of delegated ports: A 16 bit value specifying the number of ports delegated to the client for use as source port values.
- The value "n" indicates that the port range is not contiguous.

## 4. Security Considerations

Will do later. Trust issues between the client and server, plus the port randomization issues discussed in [ID.behave-natx4-log-reduction].

## 5. IANA Considerations

Will register the new option if this draft goes through as a standalone document rather than being incorporated into the base protocol.

# 6. References

## 6.1. Normative References

[ID.port-control-protocol]
Wing, D., "Port Control Protocol (PCP)", January 2011.

## 6.2. informative References

[ID.behave-natx4-log-reduction]
Tsou, T., Li, W., and T. Taylor, "Port Management To Reduce Logging In Large-Scale NATs", September 2010.

# Appendix A. NAT By-pass PCP

## A.1. Introduction

This section defines a new PCP Informational Element denoted NAT by-pass IE. The purpose of this IE is to instruct a PCP- controlled device to not enforce NAT operation on a set of flows destined to a given device located behind the PCP-controlled device.

### A.1.1. Use Cases

PCP can be used to control an upstream device to achieve the following goals:

1. A plain (i.e., a non-shared) IP address can be assigned to a given subscriber because it subscribed to a service which uses a protocol don’t embedding a transport number or because the NAT is the only deployed platform to manage IP addresses.

2. An application (e.g., sensor) does not need to listen to a whole range of ports available on a given IP address. Only a limited set of ports are used to bind its running services. For such devices, the external port(s) and IP address can be delegated to that application and therefore avoid enforcing NAT for its associated flows. The NAT in the PCP- controlled device should be bypassed.

3. A device able to restrict its source ports can be delegated an external port restricted IP address. The PCP- controlled device should be instructed to by-pass the NAT when handling flows destined/issued to that device.

### A.1.2. Scope

As currently defined in PCP Base document, PCP is unable to instruct a PCP-controlled device to de-activate the NAT for a given customer, given flows, etc.

This document defines new PCP Informational Elements (IE) which are meant to instruct a PCP-controlled device to by-pass the NAT function whenever required.

## A.2. NAT Bypass PCP Informational Element

This IE (Figure 3) is used by a PCP Client to indicate to the PCP Server to not apply any NAT operation to a corresponding binding.

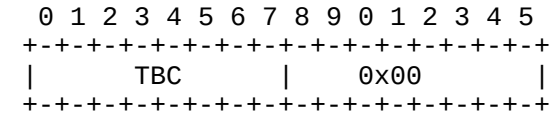


Figure 3: NAT Bypass IE

The code of this IE is to be assigned by IANA.

The length MUST be set to 0.

A PCP Client inserts this IE in a PCP request to indicate to the PCP Server to not apply the NAT function. The NAT is then by-passed in the PCP-controlled device.

A PCP Server which supports the NAT by-pass feature MUST include this IE in its response to the requesting PCP Client. In particular, when the PCP Server does not include this IE in its response, the PCP Client should deduce that the NAT will be enforced in the PCP-controlled device; a NAT will be then enforced in the PCP-controlled device.

The NAT bypass feature can be associated with a plain IP address. In such case, a full external IP address is returned to the requesting PCP Client. The client is then able to use all ports associated with that IP address (i.e., without any restriction). Furthermore, this "full" address can be used to access services which do not rely on protocols embedding a port number (e.g., some IPsec modes).

In some cases, the PCP Client can request the by-pass of the NAT but without requiring a full IP address (e.g., for the use cases described in bullet 2 and 3 of Appendix A.1.1). In such scenario, in

addition to the NAT by-pass IE, the PCP Client includes in its PCP request a Port Set Option IE (Appendix A.3). More information about this IE is provided hereafter.

## A.3. Port Set Option IE

This IE (Figure 4) is used to indicate a request for a contiguous port set. This IE conveys the length of the requested ports set. It is up to the PCP Server to decide whether the request will be satisfied or not. In particular, the PCP Server may discard the request or accept to assign a port range with a length distinct than the one requested by the PCP Client. The PCP Server can assign bigger or shorter ports set compared to is actually requested by a PCP Client.

If the PCP Server supports the ability to delegate a set of ports to a requesting PCP Client, it should include in its PCP response the external port set IE described in Figure 5.

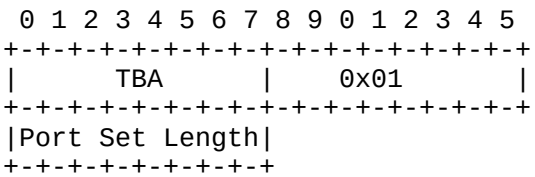


Figure 4: Port Set Option IE

The code of this IE is to be assigned by IANA.

Port Set Length indicates the length of the requested port range.

If the PCP Server is configured to assign port ranges, it should use the External Port Set IE (Appendix A.4) in its response to convey a range of port to a requesting PCP Client.

## A.4. External Port Set IE

This IE is used to enclose contiguous ports set in a PCP message sent by the PCP Server to a requesting Client. This IE may be included in a PCP response to delegate a set of ports associated with the same external IP address.

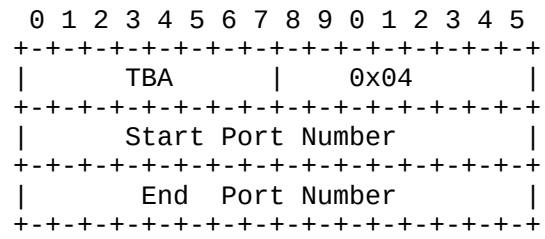


Figure 5: External Ports Set IE

The code of this IE is to be assigned by IANA.

The length filed MUST be equal to 4 bytes.

The data part of this IE indicate the bounds of the assigned ports range.

A PCP Client which receives this IE from a PCP Server is delegated all the port numbers within that range.

## Authors' Addresses

Tina Tsou
Huawei Technologies
Bantian, Longgang District
Shenzhen 518129
P.R. China

Phone:
Email: tena@huawei.com

Cathy Zhou
Huawei Technologies
Bantian, Longgang District
Shenzhen 518129
P.R. China

Phone:
Email: cathyzhou@huawei.com

Qiong Sun
China Telecom
Room 708 No.118, Xizhimenneidajie
Beijing, xicheng District 100035
China

Phone: +86 10 58552923
Email: sunqiong@ctbri.com.cn

Mohamed Boucadair
France Telecom
Rennes, 35000
France

Email: mohamed.boucadair@orange-ftgroup.com

Gabor Bajko
Nokia

Email: gabor.bajko@nokia.com