Port Control Protocol
Internet-Draft
Intended status: Standards Track
Expires: July 25, 2013

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

January 21, 2013

PCP Support for Nested NAT Environments
draft-penno-pcp-nested-nat-03

Abstract

Nested NATs or multi-layer NATs are already widely deployed. They are characterized by two or more NAT devices in the path of packets from the subscriber to the Internet. Moreover, NAT devices current deployed are PCP unaware and It is assumed that NAT aware PCP devices will take a long time to be rolled out. Therefore in order to lower the adoption barrier of PCP and make it work for current deployed networks, this document proposes a few mechanisms for PCP-enabled applications to work through NATs with varying levels of PCP protocol support.

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 July 25, 2013.

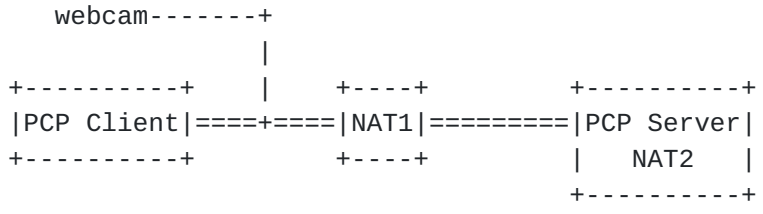
Copyright Notice

Table of Contents

1. Introduction

Nested NATs are widely deployed and come in different topology flavors. It could be a home subscriber which has an ISP provided NAT CPE chained with another personal NAT router. It could be an ISP provided CPE chained with a CGN.

An example of the use of the proposed options is illustrated in the following figure where there is a NAT in the path between the PCP Client and the PCP Server.



An example of instructing mappings in the PCP Server is as follows:

- NAT1 is detected in the path between the PCP Client and the PCP Server owing to the use of the RECEIVED_SOURCE_IP_PORT Option and the returned IP address (IP Header) of PCP request in PCP response;

- After learning about that NAT, the PCP Client uses UPnP IGD, NAT-PMP or manual configuration to interact with NAT1 and open the necessary port on NAT1 (e.g., IP address= IPx, port=X);

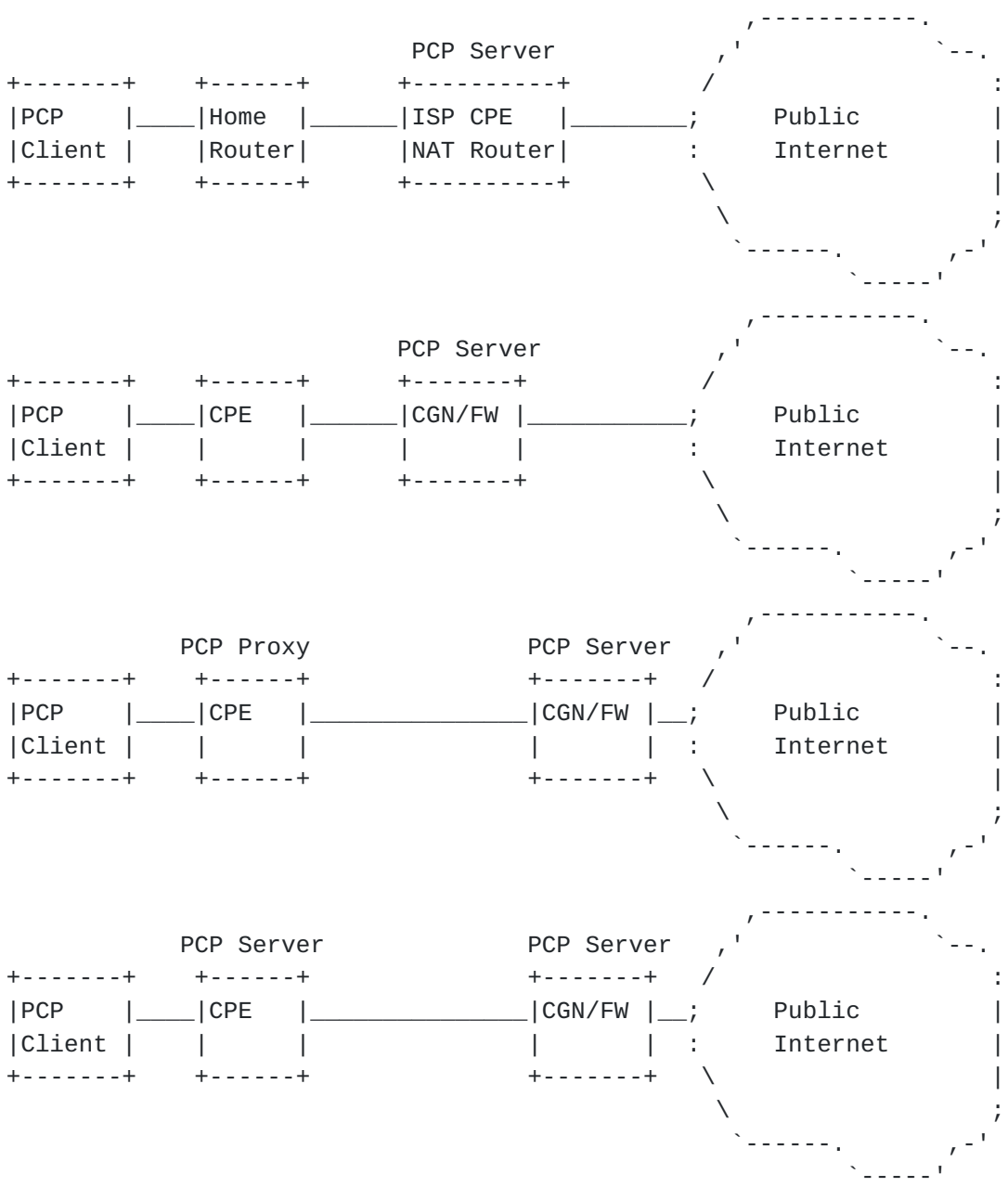
- The PCP Client then sends PCP message to the PCP Server, indicating IPx and X as the internal IP address and port. The PCP Server opens pinhole towards IPx and X.

1.1. Terminology

This document uses PCP terminology defined in [I-D.ietf-pcp-base]].

1.2. Problem Statement

The current NAT deployed devices will take years to be replaced or upgraded to become PCP aware. Moreover, nested NATs are common and come in a variety of flavors (examples below). Therefore, as applications become PCP enabled, it is important that they can work through nested NAT networks as is, without requiring infrastructure changes. From the point of view of a PCP-enabled application running on an end host, the core problem is common across different nested NAT topologies: how to install PCP mappings in a nested NAT scenario where the different NATs in the path have varying level of PCP protocol support.



1.3. Scope

This proposal considers the discovery of the PCP Server out of scope. Nonetheless, it s a critical piece of PCP deployment in service provider networks.

2. PCP MAP Nested NAT Methods

There are a few methods to make PCP work through nested NATs. They differ mainly based on the level of support that can be expected from

intermediate NATs, which can be:

- PCP and UPnP unaware or disabled
- PCP Server
- UPnP Server
- PCP Proxy

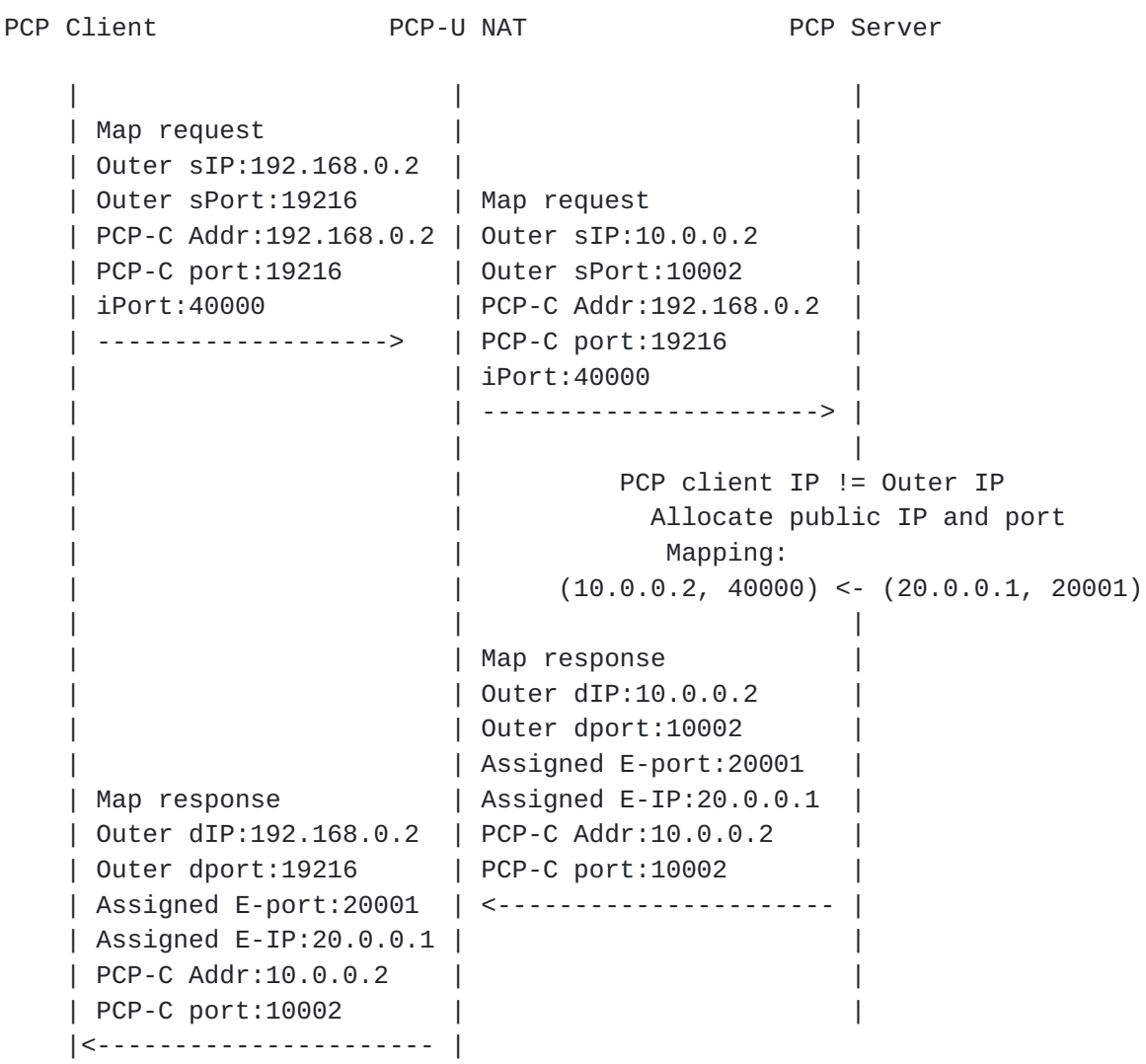
The next sections discuss each scenario on the basis of protocol support on intermediate NATs.

2.1. PCP and UPnP unaware Intermediate NATs

This method will most likely be used by PCP clients in nested NAT environments while PCP Proxy support in not ubiquitous. It assumes no UPnP or PCP Proxy support on intermediate NATs. This proposal leverages the current behavior of PCP [I-D.ietf-pcp-base] which allows a PCP Client and Server to detect intervening nested NATs. The PCP Server uses the information on the outer IP and PCP headers to detect and install a proper NAT mapping and return the source IP: port from the IP header on the PCP response. It does not assume any change to current deployed NATs.

1. The PCP Client sends the MAP request as it normally would without any changes.
2. As the message goes through one (or more) PCP-unaware NAT, the source IP:port of the IP header will change accordingly
3. The PCP Server compares the PCP Client IP:port in the PCP header with the source IP:port of the IP header
4. If these are different, the server knows that the PCP message went through a PCP-unaware NAT. Therefore it installs a mapping directed to the source IP address found on the IP header and internal port of the PCP header.

s/dport: source/destination port
s/dIP : source/destination IP
PCP-C : PCP client
iport : Internal port
PCP-U : PCP Unaware NAT
E-port : External port
E-IP : External IP



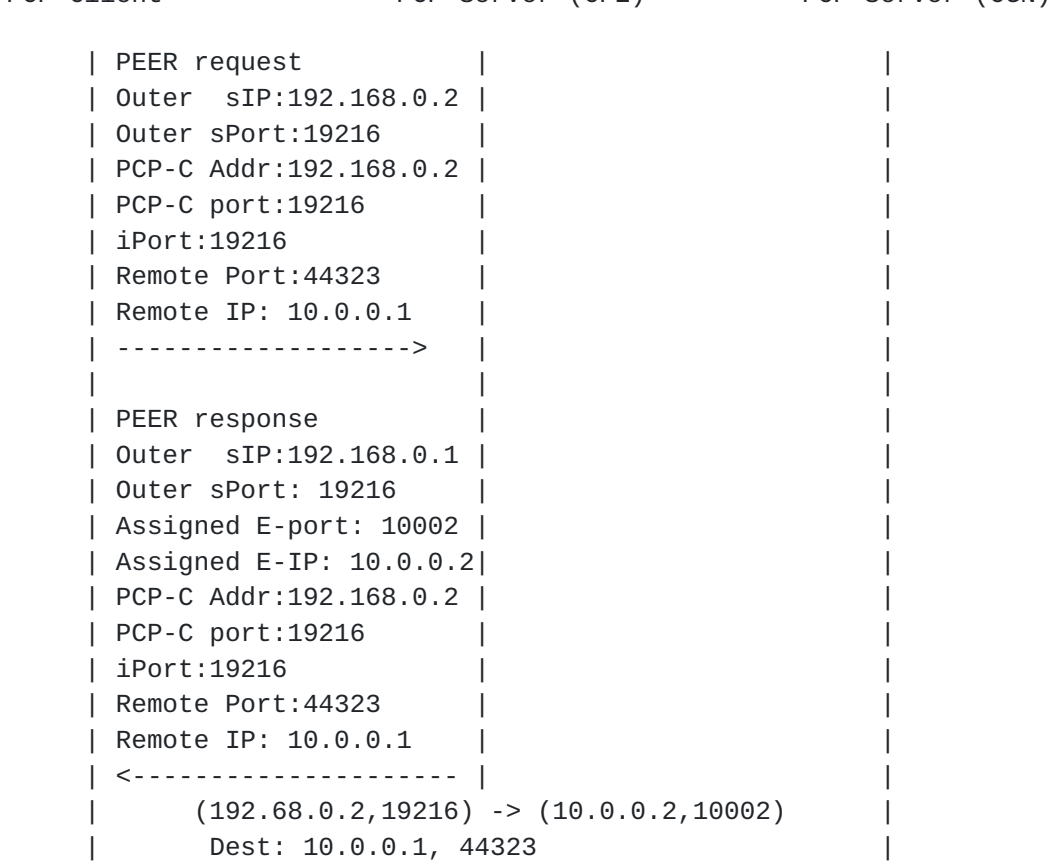
- Subscriber installs a port forwarding or DMZ entry on its home CPE (PCP U-NAT) through manual configuration. The entry would be (*, 40000) -> (10.0.0.1, 40000). Alternatively the application could use UPnP for the same purpose.

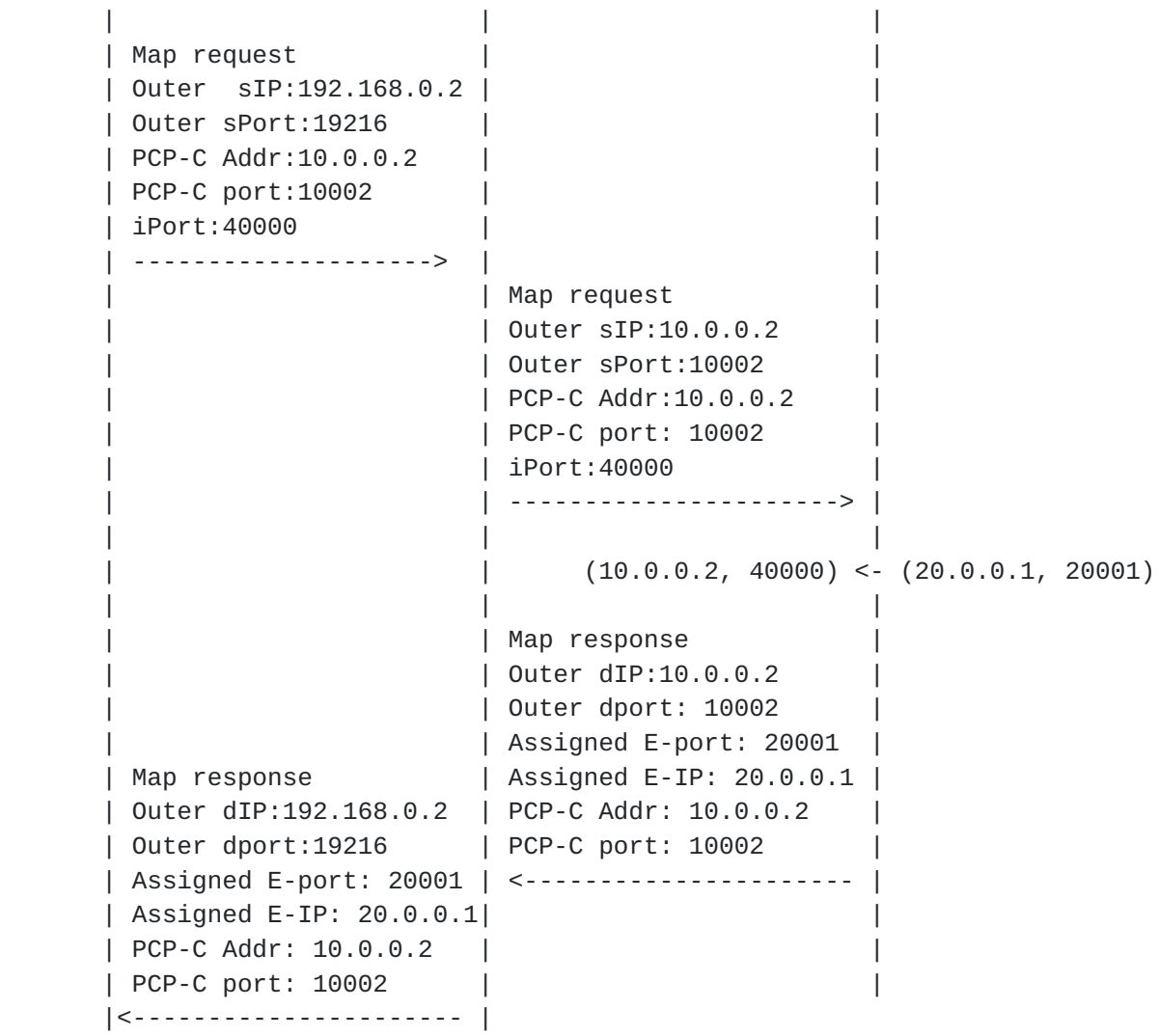
2.2. PCP Server intermediate NAT

If the intermediate NAT implements a PCP Server (but not a Proxy), a two-step iterative process is needed in order to install PCP PEER mappings for the PCP control message itself followed by another PCP mapping for the data path. If the PCP Client Address does not match the IP address of IP header, PCP Server (CGN) will reject request with ADDRESS_MISMATCH error. Therefore PCP Client first needs to know the IP address and port the CPE NAT will use for the actual PCP request to CGN.

If the PCP client relies on nested NAT detection the first step is not needed. It is assumed that before sending the PCP MAP request to the CGN the client would install the following map on the NAT Home Gateway: (192.168.0.2, 40000) <- (10.0.0.2, 40000). The internal port that the server listens on does not necessarily needs to be 40000, it could be different than the internal port used between the CGN and CPE.

The drawback of this technique is that there is no obvious way for the PCP Client to know the PCP Servers downstream. One possibility is for each PCP Server in the path to return the address of the upstream PCP Server to the PCP Client.





2.3. UPnP enabled intermediate NAT

This scenario is very similar to the PCP Server intermediate NAT, but the CPE implements a UPnP Server instead of PCP Server. The mechanics are the same with the difference that first PEER message to setup the PCP Control messages mapping is substituted by its UPnP equivalent.

2.4. PCP Proxy Intermediate NAT

This method assumed that the intermediate NATs implement a PCP Proxy function. There are two non-exclusive types of proxy functions: interception (ALG) and server-client based. In the interception case the PCP Proxy intercepts PCP messages destined to a PCP Server downstream, modifies IP, UDP and PCP headers, allocates a mapping and send them to the downstream PCP Server. Ideally if the interception PCP Proxy also implements a PCP server it would let the PCP Client

know of its existence in a PCP response through an option (TBD) and henceforth the PCP Client would start directing messages to it.

In the server-client scenario the PCP Client sends PCP messages to the proxy which acts as both PCP Server and Client. This proxy in turn will terminate the PCP request and generate a new one acting as a PCP Client to its own PCP Server. Therefore mappings are installed in all NAT devices in a recursive manner. This is the recommended method since its does not need a special discovery procedure and works with any number of NATs. More information about this method can be found in [I-D.bpw-pcp-proxy].

2.4.1. PCP Proxy Discovery

TBD

3. PCP PEER Nested NAT Methods

All techniques discussed for PCP MAP methods do not work for PCP PEER messages. PCP PEER is a different beast and another set of techniques need to be used to overcome intervening NATs. The critical issue related to PEER is that the client needs to know the external source port NAT1 will use to translate packets for the actual data session. There are two scenarios to consider: send-then-connect and connect-then-send.

3.1. Send-then-connect

In this scenario the client sends a PEER message to install a mapping which later will be used by a regular UDP or TCP data session. In order for this to work reliably, the following procedure needs to the followed:

1. PCP Client needs to allocate a binding on the intervening NAT thorugh STUN, UPnP or other method. Let's suppose this binding is (192.168.0.2, 19216 <-> 10.0.0.2, 10002).

2. PCP Client constructs a PCP PEER request like the following

   * Internal port: 10002

   * Remote Peer Port: 20002 (upcoming data connection destination port)

   * Remote Peer address: 20.0.0.2 (upcoming data connection destination IP address)

* PCP Client Address: 10.0.0.2

* Protocol: TCP

3. Application will connect to remote peer using the same source IP address and port of the existing mapping on the intervening NAT. If the intervening NAT supports Protocol Independent Endpoint Independent Mapping (PI-EIM) [I-D.penno-behave-rfc4787-5382-5508-bis] , it will allocate the same external IP:port of step 1 to the new connection. Therefore 5-tuple of the new connection will match those of the previously installed PEER map.

3.2. Connect-then-send

If the data connection to the remote peer is established before the PEER message, the challenge for the PCP client is to find out which source IP:port the intervening NAT is using to translate the data packets. If the PCP Client has the necessary permissions to reuse the socket used by the data connection and the intervening NAT support EIM, two solutions are possible:

1. PCP Client sends a request from the same source IP:port as the data connection. Since the intervening NAT supports PI-EIM, it should allocate the same external IP:port of the data connection, which would be returned in the RECEIVED_PORT_OPTION. The PCP Client then can send an appropriate PEER message to take over the data connection. The advantage of this solution is that is built around a single protocol, PCP, and the disadvatange is that it requires a PCP extension.

2. If the PCP Client is also a STUN client it can send a binding request from the same source IP:port as the data connection and since the intervening NAT supports EIM the client will find out the external IP:port that is used to translate data packets. The PCP Client then can send an appropriate PEER message to take over the data connection. The advantage of this solution is that no extensions to PCP are needed. The disadvantage is that STUN client and server are needed, specially the fact that in case of nested NATs the STUN server needs to be located between NAT1 and NAT2.

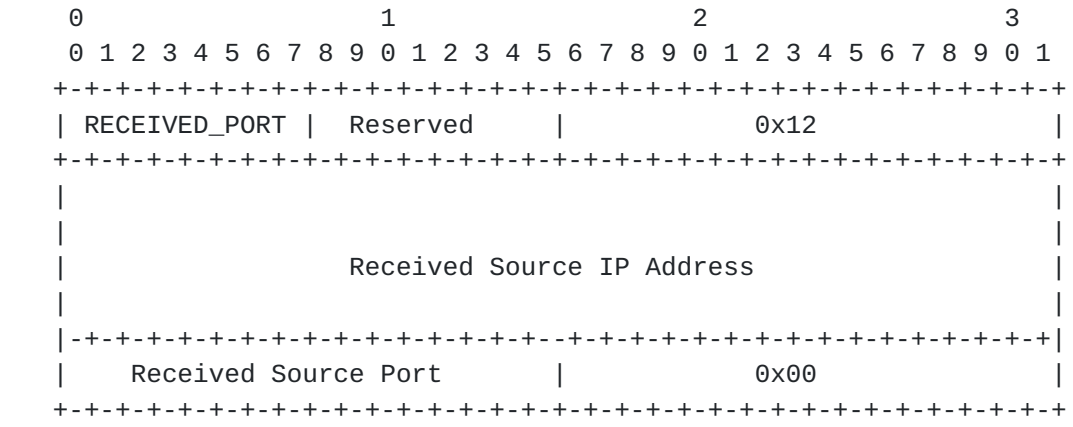
4. RECEIVED_SOURCE_IP_PORT Option

This option (Code TBA, Figure 1) is used by a PCP Server to indicate in a PCP response the source IP and port of PCP messages received from a PCP Client. Together with the IP Address of the PCP Client

conveyed in the common PCP header, a PCP Client uses this information to detect whether a NAT is present in the path to reach its PCP Server.

A PCP Client MAY include this option to learn the port number as perceived by the PCP Server. When this option is received by the PCP Server, it uses the source IP:port of the received PCP request to set the Received Port.

This Option:
Option Name: PCP Received Port Option (RECEIVED_SOURCE_IP_PORT)
Number: TBA (IANA)
Purpose: Detect the presence of a NAT in the path and discover externally allocate IP:port
Valid for Opcodes: MAP and PEER
Length: 0x12
May appear in: both request and response
Maximum occurrences: 1



Received Source Port and IP: The source IP:port number of the received PCP request as seen by the PCP Server.

Figure 1: Received IP address/port PCP option

5. SCOPE Option

The Scope Option (Code TBA, Figure 2) is used by a PCP Client to indicate to the PCP Server the scope of the flows that will use a given mapping. This object is meant to be used in the context of cascaded PCP Servers/NAT levels. Two values are defined:

Value	Meaning
0x00	Internet
0x01	Internal

When 0x00 value is used, the PCP Proxy MUST propagate the mapping request to its upstream PCP Server. When 0x01 value is used, the mapping is to be instantiated only in the first PCP-controlled device; no mapping is instantiated in the upstream PCP-controlled device.

When no Scope Option is included in a PCP message, this is equivalent to including a Scope Option with a scope value of "Internet".

This Option:
Option Name: PCP Scope Policy Option (SCOPE)
Number: TBA (IANA)
Purpose: Restrict the scope of PCP requests
Valid for Opcodes: MAP
Length: 0x04
May appear in: both request and response
Maximum occurrences: 1

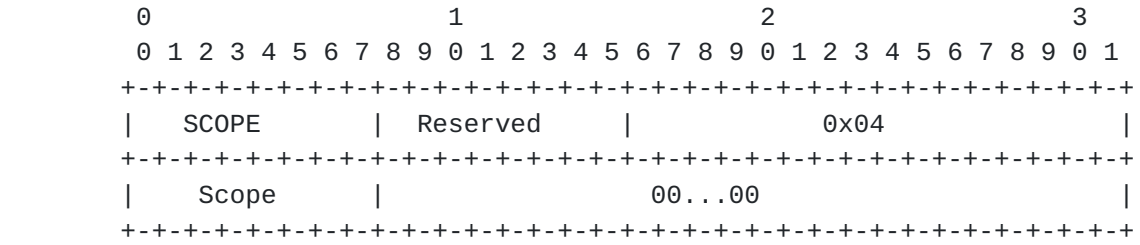


Figure 2: Scope Option

6. IANA Considerations

The following PCP Option Codes are to be allocated:

RECEIVED_PORT

SCOPE

7. Security Considerations

Security considerations discussed in [I-D.ietf-pcp-base] must be considered.

8. Acknowledgements

Thanks to Linda (wang.cui1@zte.com.cn) for her review.

9. References

9.1. 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-29 (work in progress), November 2012.

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

9.2. Informative References

[I-D.bpw-pcp-proxy]
Boucadair, M., Penno, R., Wing, D., and F. Dupont, "Port Control Protocol (PCP) Proxy Function", draft-bpw-pcp-proxy-02 (work in progress), September 2011.

[I-D.penno-behave-rfc4787-5382-5508-bis]
Penno, R., Perreault, S., Kamiset, S., Boucadair, M., and K. Naito, "Network Address Translation (NAT) Behavioral Requirements Updates", draft-penno-behave-rfc4787-5382-5508-bis-04 (work in progress), January 2013.

Authors' Addresses

Reinaldo Penno
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, California 95134
USA

Email: repenno@cisco.com

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

Email: dwing@cisco.com

Mohamed Boucadair
France Telecom
Rennes, 35000
France

Email: mohamed.boucadair@orange.com