

Port Control Protocol  
Internet-Draft  
Intended status: Standards Track  
Expires: April 23, 2012

R. Penno  
Juniper Networks  
D. Wing  
Cisco  
P. Selkirk  
Internet Systems Consortium  
M. Boucadair  
France Telecom  
October 21, 2011

**PCP Support for Nested NAT Environments**  
**draft-penno-pcp-nested-nat-01**

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 currently 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 nested NATs with varying level 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 April 23, 2012.

## Copyright Notice

Copyright (c) 2011 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

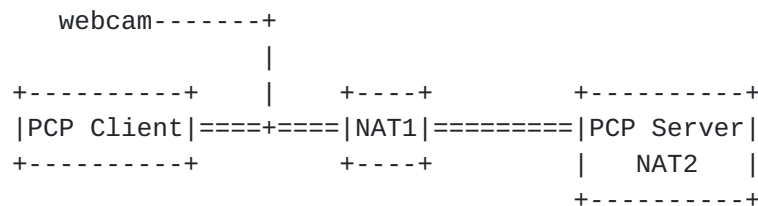
<a href="#">1.</a>	Introduction . . . . .	<a href="#">3</a>
<a href="#">1.1.</a>	Terminology . . . . .	<a href="#">3</a>
<a href="#">1.2.</a>	Problem Statement . . . . .	<a href="#">3</a>
<a href="#">1.3.</a>	Scope . . . . .	<a href="#">4</a>
<a href="#">2.</a>	PCP Nested NAT Methods . . . . .	<a href="#">4</a>
<a href="#">2.1.</a>	NAT and UPnP unaware Intermediate NATs . . . . .	<a href="#">5</a>
<a href="#">2.2.</a>	PCP Server intermediate NAT . . . . .	<a href="#">7</a>
<a href="#">2.3.</a>	UPnP enabled intermediate NAT . . . . .	<a href="#">8</a>
<a href="#">2.4.</a>	PCP Proxy Intermediate NAT . . . . .	<a href="#">8</a>
<a href="#">2.4.1.</a>	PCP Proxy Discovery . . . . .	<a href="#">9</a>
<a href="#">3.</a>	RECEIVED_PORT Option . . . . .	<a href="#">9</a>
<a href="#">4.</a>	SCOPE Option . . . . .	<a href="#">10</a>
<a href="#">5.</a>	IANA Considerations . . . . .	<a href="#">11</a>
<a href="#">6.</a>	Security Considerations . . . . .	<a href="#">11</a>
<a href="#">7.</a>	Acknowledgements . . . . .	<a href="#">11</a>
<a href="#">8.</a>	References . . . . .	<a href="#">11</a>
<a href="#">8.1.</a>	Normative References . . . . .	<a href="#">11</a>
<a href="#">8.2.</a>	Informative References . . . . .	<a href="#">12</a>
	Authors' Addresses . . . . .	<a href="#">12</a>



## 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:

- o NAT1 is detected in the path between the PCP Client and the PCP Server owing to the use of the RCEIVED\_PORT Option and returned perceived IP address in PCP response;
- o 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);
- o 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.

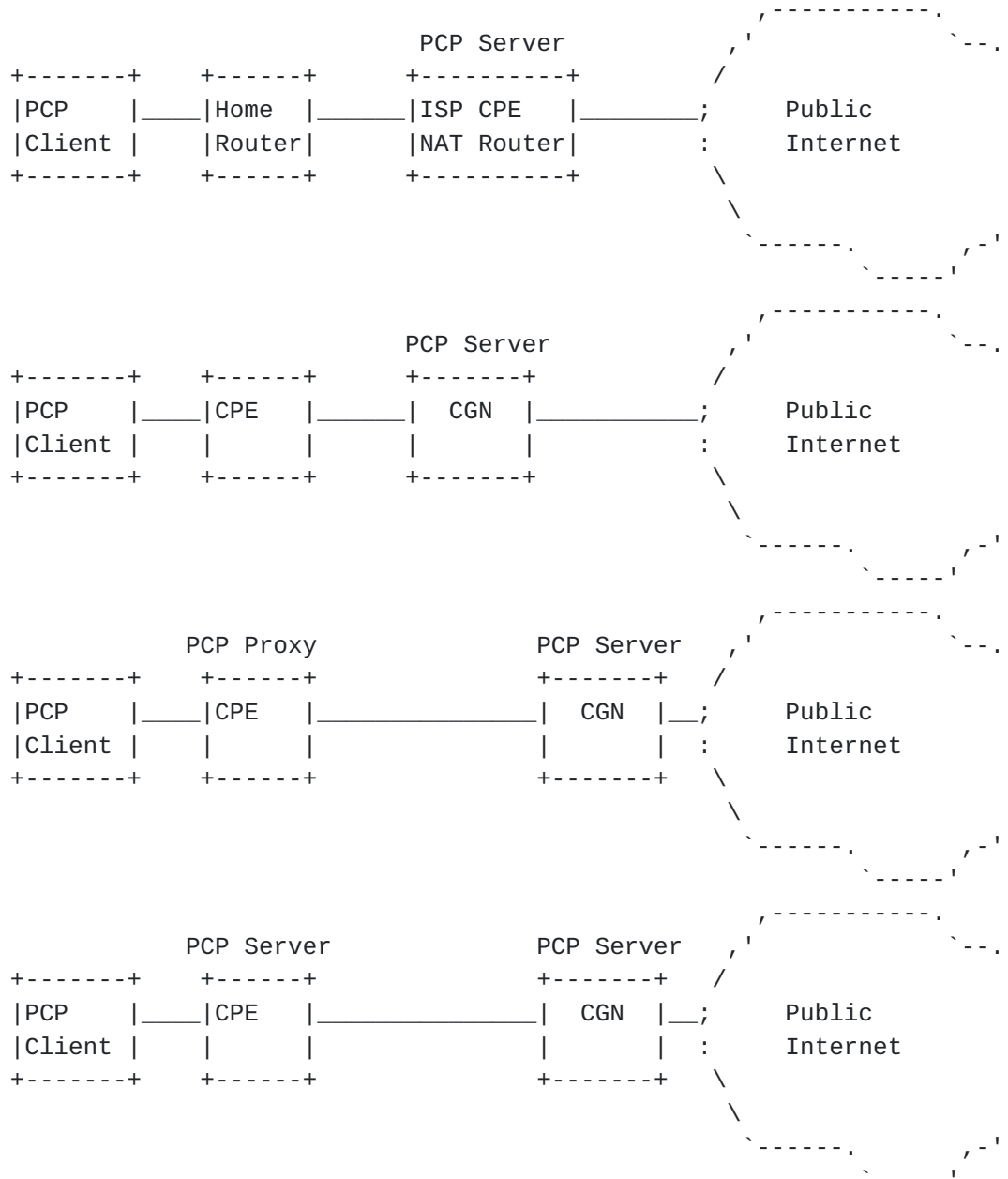
### 1.1. Terminology

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 is a critical piece of PCP deployment in service provider networks.

## 2. PCP 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:

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

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

### **2.1. NAT and UPnP unaware Intermediate NATs**

This method will most likely be used by PCP clients in nested NAT environments while PCP Proxy support is 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

## PCP Client

## PCP-U NAT

## PCP Server

Map request		
Outer sIP:192.68.0.2		
Outer sPort:19268	Map request	
PCP-C Addr:192.168.0.1	Outer sIP:10.0.0.2	
PCP-C port:19268	Outer sPort:10002	
iPort:40000	PCP-C Addr:192.168.0.1	
----->	PCP-C port:19268	
	iPort:40000	
	----->	
	PCP client IP != Outer IP	
	Allocate public IP and port	
	Mapping:	
	(10.0.0.2, 40000) <- (20.0.0.1, 20001)	
	Map response	
	Outer dIP:10.0.0.2	
	Outer dport:10002	
	Assigned E-port:20001	
Map response	Assigned E-IP:20.0.0.1	
Outer dIP:192.168.0.2	PCP-C Addr:10.0.0.2	
Outer dport:19268	PCP-C port:10002	
Assigned E-port:20001	<-----	
Assigned E-IP:20.0.0.1		
PCP-C Addr:10.0.0.2		
PCP-C port:10002		
<-----		

- 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 relies on nested NAT detection the first step is not needed. It is assumed that before 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 need 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.

PCP Client	PCP Server (CPE)	PCP Server (CGN)
PEER request		
Outer sIP:192.168.0.2		
Outer sPort:19216		
PCP-C Addr:192.168.0.2		
PCP-C port:19216		
iPort:19216		
Remote Port:44323		
Remote IP: 10.0.0.1		
----->		
PEER response		
Outer sIP:192.168.0.1		
Outer sPort: 19216		
Assigned E-port: 10002		
Assigned E-IP: 10.0.0.2		
PCP-C Addr:192.168.0.2		
PCP-C port:19216		
iPort:19216		
Remote Port:44323		
Remote IP: 10.0.0.1		
<-----		
(192.68.0.2,19216) -> (10.0.0.2,10002)		
Dest: 10.0.0.1, 44323		
Map request		
Outer sIP:192.168.0.2		
Outer sPort:19216		
PCP-C Addr:10.0.0.2		
PCP-C port:10002		



```

| iPort:40000 | | | |
| -----> | | |
| | | Map request | |
| | | Outer sIP:10.0.0.2 | |
| | | Outer sPort:10002 | |
| | | PCP-C Addr:10.0.0.2 | |
| | | PCP-C port: 10002 | |
| | | iPort:40000 | |
| | | -----> | |
| | | | |
| | | (10.0.0.2, 40000) <- (20.0.0.1, 20001) | |
| | | | |
| | | Map response | |
| | | Outer dIP:10.0.0.2 | |
| | | Outer dport: 10002 | |
| | | Assigned E-port: 20001 | |
| Map response | Assigned E-IP: 20.0.0.1 | |
| Outer dIP:192.168.0.2 | PCP-C Addr: 10.0.0.2 | |
| Outer dport:19216 | PCP-C port: 10002 | |
| Assigned E-port: 20001 | <----- | |
| Assigned E-IP: 20.0.0.1 | | |
| PCP-C Addr: 10.0.0.2 | | |
| PCP-C port: 10002 | | |
| <----- | | |

```

### **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 it 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. RECEIVED\_PORT Option**

This option (Code TBA, Figure 1) is used by a PCP Server to indicate in a PCP response the source 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 port of the received PCP request to set the Received Port.





This Option:

Option Name: PCP Received Port Option (RECEIVED\_PORT)

Number: TBA (IANA)

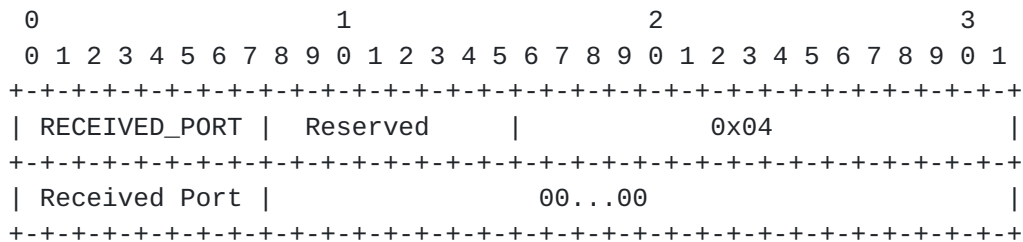
Purpose: Detect the presence of a NAT in the path

Valid for Opcodes: MAP

Length: 0x04

May appear in: both request and response

Maximum occurrences: 1



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

Figure 1: Received IP address/port PCP option

#### 4. 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

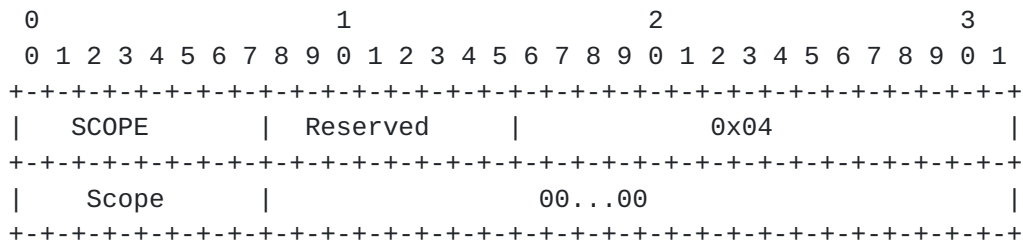


Figure 2: Scope Option

## 5. IANA Considerations

The following PCP Option Codes are to be allocated:

RECEIVED\_PORT

SCOPE

## 6. Security Considerations

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

## 7. Acknowledgements

TBD

## 8. References

### 8.1. 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-16](#) (work in progress), October 2011.

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

## **8.2. 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.

### Authors' Addresses

Reinaldo Penno  
Juniper Networks  
1194 N Mathilda Avenue  
Sunnyvale, California 94089  
USA

Email: [rpenno@juniper.net](mailto:rpenno@juniper.net)

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

Email: [dwing@cisco.com](mailto:dwing@cisco.com)

Paul Selkirk  
Internet Systems Consortium  
950 Charter Street  
Redwood City, California 94063

Phone:

Fax:

Email: [pselkirk@isc.org](mailto:pselkirk@isc.org)

URI:



Mohamed Boucadair  
France Telecom  
Rennes, 35000  
France

Email: mohamed.boucadair@orange.com