PCP working group
Internet-Draft
Intended status: Standards Track
Expires: August 11, 2013

S. Cheshire
Apple

Feb 7, 2013

PCP Anycast Address
draft-cheshire-pcp-anycast-00

Abstract

The Port Control Protocol Anycast Address enables PCP clients to transmit messages to their closest on-path NAT, Firewall, or other middlebox, without having to learn the IP address of that middlebox via some external channel.

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 August 11, 2013.

1. Introduction

The Port Control Protocol document [PCP] specifies the message formats used, but the address to which a client sends its request is either assumed to be the default router (which is appropriate in a typical single-link residential network) or has to be configured otherwise via some external mechanism, such as DHCP.

One drawback of relying on external configuration is that it creates an external dependency on another piece of network infrastructure which must be configured with the right address for PCP to work. In some environments the staff managing the DHCP servers may not be the same staff managing the NAT gateways, and in any case, constantly keeping the DHCP server address information up to date as NAT gateways are added, removed, or reconfigured, is burdensome. In addition DHCP clients would have to be updated to request the new "PCP Address" DHCP option, which for practical purposes limits this capability to operating system vendors, precluding creation of a pure user-level PCP library that can be linked with (and shipped) with an application that requires PCP.

Another drawback of relying on DHCP for configuration is that one of the target deployment environments for PCP -- 3GPP for mobile telephones -- does not use DHCP.

One design option that was considered for Apple's NAT gateways was to have the NAT gateway simply handle and respond to all packets addressed to UDP port 5351, regardless of the destination address in the packet. Since the device is a NAT gateway, it already examines every packet in order to rewrite port numbers, so also detecting packets addressed to UDP port 5351 is not a significant additional burden. Also, since this device is a NAT gateway which rewrites port numbers, any attempt by a client to talk *though* this first NAT gateway to create mappings in some second upstream NAT gateway is futile and pointless. Any mappings created in the second NAT gateway are useful to the client only if there are also corresponding mappings created in the first NAT gateway. Consequently, there is no case where it is useful for PCP requests to pass transparently through the first PCP-aware NAT gateway on their way to the second PCP-aware NAT gateway. In all cases, for useful connectivity to be established, the PCP request must be handled by the first NAT gateway, and then the first NAT gateway generates a corresponding new upstream request to establish a mapping in the second NAT gateway. (This process can be repeated recursively for as many times as necessary for the depth of nesting of NAT gateways; this is transparent to the client device.)

These two issues result in the following related observations: the PCP client may not *know* what destination address to use in its PCP request packets; the PCP server doesn't *care* what destination address is in the PCP request packets.

Given that the devices neither need to know nor care what destination address goes in the packet, all we need to do is pick one and use it. It's little more than a placeholder in the IP header. Any globally routable unicast address will do. Since this address is one that automatically routes its packet to the closest on-path device that implements the desired functionality, it is an anycast address.

In the simple case where the first-hop router is also the NAT gateway (as is common in a typical single-link residential network), sending to the PCP anycast address is equivalent to sending to the client's default router, as specified in the PCP base document [PCP].

In the case of a larger corporate network, where there may be several internal routed subnets and one or more border NAT gateway(s) connecting to the rest of the Internet, sending to the PCP anycast address has the interesting property that it magically finds the right border NAT gateway for that client. Since we posit that other network infrastructure does not need (and should not have) any special knowledge of PCP (or its anycast address) this means that to other non-NAT routers, the PCP anycast address will look like any other unicast destination address on the public Internet, and consequently the packet will be forwarded as for any other packet destined to the public Internet, until it reaches a NAT or firewall device that is aware of the PCP anycast address. This will result in the packet naturally arriving the NAT gateway that handles this client's outbound traffic destined to the public Internet, which is exactly the NAT gateway that the client wishes to communicate with when managing its port mappings.

2. Benefit of using a PCP Anycast Address

The benefit of using an anycast address is simplicity and reliability. In an example deployment scenario:

1. A network administrator installs a PCP-capable NAT.

2. An end user (who may be the same person) runs a PCP-enabled application. This application can implement PCP with purely user-level code -- no operating system support is required.

3. This PCP-enabled application sends its PCP request to the PCP anycast address. This packet travels through the network like any other, without any special support from DNS, DHCP, other routers, or anything else, until it reaches the PCP-capable NAT, which receives it, handles it, and sends back a reply.

Using the PCP anycast address, the only two things that need to be deployed in the network are the two things that actually use PCP: The PCP-capable NAT, and the PCP-enabled application. Nothing else in the network needs to be changed or upgraded, and nothing needs to be configured, including the PCP client.

3. Historical Objections to Anycast

In March 2001 a draft document entitled "Analysis of DNS Server Discovery Mechanisms for IPv6" [DNSDisc] proposed using anycast to discover DNS servers, a proposal that was subsequently abandoned in later revisions of that draft document.

There are legitimate reasons why using anycast to discover DNS servers is not compelling, mainly because it requires explicit configuration of routing tables to direct those anycast packets to the desired DNS server. However, DNS server discovery is very different to NAT gateway discovery. A DNS server is something a client explicitly talks to, via IP address. The DNS server may be literally anywhere on the Internet. Various reasons make anycast an uncompelling technique for DNS server discovery:

- DNS is a pure application-layer protocol, running over UDP.

- On an operating system without appropriate support for configuring anycast addresses, a DNS server would have to use something like Berkeley Packet Filter (BPF) to snoop on received packets to intercept DNS requests, which is inelegant and inefficient.

- Without appropriate routing changes elsewhere in the network, there's no reason to assume that packets sent to that anycast address would even make it to the desired DNS server machine. This places an addition configuration burden on the network administrators, to install approprate routing table entries to direct packets to the desired DNS server machine.

In contrast, a NAT gateway is something a client's packets stumble across as they try to leave the local network and head out onto the public Internet. The NAT gateway has to be on the path those packets naturally take or it can't perform its NAT functions. As a result, the objections to using anycast for DNS server discovery do not apply to PCP:

- No routing changes are needed (or desired) elsewhere in the local network, because the whole *point* of using anycast is that we want the client's PCP request packet to take the same forwarding path through the network as a TCP SYN to any other remote destination address, because we want the *same* NAT gateway that would have made a mapping in response to receiving an outbound TCP SYN packet from the client to be the the one that makes a mapping in response to receiving a PCP request packet from the client.

- A NAT engine is already snooping on (and rewriting) every packet it forwards. As part of that snooping it could trivially look for packets addressed to the PCP UDP port and process them locally (just like the local processing it already does when it sees an outbound TCP SYN packet).

4. IANA Considerations

IANA should allocate an IPv4 and an IPv6 well-known PCP anycast address.

192.0.0.0/24 and 2001:0000::/23 are reserved for IETF Protocol Assignments, as listed at <http://www.iana.org/assignments/iana-ipv4-special-registry/> and <http://www.iana.org/assignments/iana-ipv6-special-registry/>

Suitable addresses in these ranges, such as 192.0.0.8, and a corresponding suitable IPv6 address, should be allocated.

5. Security Considerations

In a network without any border gateway, NAT or firewall that is aware of the PCP anycast address, outgoing PCP requests could leak out onto the external Internet, possibly revealing information about internal devices.

Using an IANA-assigned well-known PCP anycast address enables border gateways to block such outgoing packets. In the default-free zone, routers should be configured to drop such packets. Such configuration can occur naturally via BGP messages advertising that no route exists to said address.

Sensitive clients that do not wish to leak information about their presesence can set an IP TTL on their PCP requests that limits how far they can travel into the public Internet.

6. References

6.1. Normative References

[PCP] 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.

6.2. Informative References

[DNSDisc] Hagino, J. and D. Thaler, "Analysis of DNS Server Discovery Mechanisms for IPv6", draft-ietf-ipngwg-dns-discovery-01 (work in progress), November 2001.

Author's Address

Stuart Cheshire
Apple Inc.
1 Infinite Loop
Cupertino, California 95014
USA

Phone: +1 408 974 3207
Email: cheshire@apple.com