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

Abstract

   The Port Control Protocol (PCP) Anycast Address enables PCP clients
   to transmit signaling 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. This document
   establishes one well-known IPv4 address and one well-known IPv6
   address to be used as PCP Anycast Address.

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

1. Introduction

The Port Control Protocol (PCP) [RFC6887] provides a mechanism to
control how incoming packets are forwarded by upstream devices such
as Network Address Translator IPv6/IPv4 (NAT64), Network Address
Translator IPv4/IPv4 (NAT44), IPv6 and IPv4 firewall devices, and a
mechanism to reduce application keep alive traffic.

The PCP document [RFC6887] 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. The properties and drawbacks of
various mechanisms are discussed in Appendix A.

This document follows a different approach: it establishes a well-
known anycast address for the PCP Server. PCP clients are expected
to send requests to this address during the PCP Server discovery
process. A PCP Server configured with the anycast address could
optionally redirect or return a list of unicast PCP Servers to the
client. For a more extensive discussion on anycasting see
Appendix B.

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.

[2](#). **PCP Server Discovery based on well-known IP Address**

[2.1](#). **PCP Discovery Client behavior**

   PCP Clients that need to discover PCP servers SHOULD first send a PCP
   request to its default router. This is important because in the case
   of cascaded PCP Servers, all of them need to be discovered in order
   of hop distance from the client. The PCP client then SHOULD send a
   PCP request to the anycast address. PCP Clients must be prepared to
   receive an error and try other discovery methods.

[2.2](#). **PCP Discovery Server behavior**

   PCP Server can be configured to listen on the anycast address for
   incoming PCP requests.

   PCP responses are sent from that same IANA-assigned address (see Page
   5 of [[RFC1546](#)]).

3. Deployment Considerations

   Network operators should install one or more PCP Servers as specified
   above. Depending on the network deployment scenario they may use IP
   routing tables, or other suitable mechanisms to direct PCP requests
   to one of these servers.

3.1. Multiple PCP Servers, Symmetric Routing

   In the case of symmetric routing all inbound and outbound packets
   from a PCP client traverse the same PCP Server or controlled device.
   Multiple PCP Servers sharing an anycast address in a symmetric
   routing scenario are used for two purposes: ease of network
   configuration and redundancy. In the case of redundancy, if there is
   a network or routing change a PCP client might start interacting with
   a different PCP Server sharing the same anycast address. From a PCP
   Client point of view this would be the same as a PCP Server reboot
   and a PCP Client could find out about it by examining the Epoch field
   during the next PCP response or ANNOUNCE message.

3.2. Multiple PCP Servers, Asymetric Routing

   In the case of asymmetric routing inbound packets from a PCP client
   traverse a different PCP Server or controlled device than outbound
   packets. If these PCP Servers are firewalls, the PCP client would
   need to create mappings on both or rely on state-synch between them
   in order to properly communicate with other hosts. But if these PCP
   Servers share an anycast address the PCP Client will create mappings
   in only one, when in fact should create mappings on both of them.

   Therefore in order to support this scenario the firewalls need to
   synch state, alleviating the development on the client, or
   alternatively a new option could be used that informs the client of
   the unicast address of each PCP Servers, alleviating the development
   on the PCP Server or Firewall.

   At the time of this writing there is no clear consensus on the PCP WG
   if a new option is desirable. Current deployments rely on state
   synch between Firewall devices to solve assymetric routing issues.
   If those devices are upgraded to be controlled by PCP Servers it
   might be desirable to maintain the client simple.

4. IANA Considerations

4.1. Registration of IPv4 Special Purpose Address

   IANA is requested to register a single IPv4 address in the IANA IPv4
   Special Purpose Address Registry [RFC5736].

   [RFC5736] itemizes some information to be recorded for all
   designations:

      1. The designated address prefix.

      Prefix: TBD by IANA. Prefix length: /32

      2. The RFC that called for the IANA address designation.

      This document.

      3. The date the designation was made.

      TBD.

      4. The date the use designation is to be terminated (if specified
      as a limited-use designation).

      Unlimited. No termination date.

      5. The nature of the purpose of the designated address (e.g.,
      unicast experiment or protocol service anycast).

      protocol service anycast.

      6. For experimental unicast applications and otherwise as
      appropriate, the registry will also identify the entity and
      related contact details to whom the address designation has been
      made.

      N/A.

      7. The registry will also note, for each designation, the
      intended routing scope of the address, indicating whether the
      address is intended to be routable only in scoped, local, or
      private contexts, or whether the address prefix is intended to be
      routed globally.

Typically used within a network operator's network domain, but in principle globally routable.

8. The date in the IANA registry is the date of the IANA action, i.e., the day IANA records the allocation.

TBD.

4.2. Registration of IPv6 Special Purpose Address

IANA is requested to register a single IPv6 address in the IANA IPv6 Special Purpose Address Block [RFC4773].

[RFC4773] itemizes some information to be recorded for all designations:

1. The designated address prefix.

Prefix: TBD by IANA. Prefix length: /128

2. The RFC that called for the IANA address designation.

This document.

3. The date the designation was made.

TBD.

4. The date the use designation is to be terminated (if specified as a limited-use designation).

Unlimited. No termination date.

5. The nature of the purpose of the designated address (e.g., unicast experiment or protocol service anycast).

protocol service anycast.

6. For experimental unicast applications and otherwise as appropriate, the registry will also identify the entity and related contact details to whom the address designation has been made.

N/A.

7. The registry will also note, for each designation, the
intended routing scope of the address, indicating whether the
address is intended to be routable only in scoped, local, or
private contexts, or whether the address prefix is intended to be
routed globally.

Typically used within a network operator's network domain, but in
principle globally routable.

8. The date in the IANA registry is the date of the IANA action,
i.e., the day IANA records the allocation.

TBD.

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

[RFC1546] Partridge, C., Mendez, T., and W. Milliken, "Host
           Anycasting Service", RFC 1546, November 1993.

[RFC3958] Daigle, L. and A. Newton, "Domain-Based Application
           Service Location Using SRV RRs and the Dynamic Delegation
           Discovery Service (DDDS)", RFC 3958, January 2005.

[RFC4773] Huston, G., "Administration of the IANA Special Purpose
           IPv6 Address Block", RFC 4773, December 2006.

[RFC5736] Huston, G., Cotton, M., and L. Vegoda, "IANA IPv4 Special
           Purpose Address Registry", RFC 5736, January 2010.

[RFC6887] Wing, D., Cheshire, S., Boucadair, M., Penno, R., and P.
           Selkirk, "Port Control Protocol (PCP)", RFC 6887,
           April 2013.

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.

[DhcpRequestParams]
           OpenFlow, "OpenFlow Switch Specification", February 2011,
           <http://msdn.microsoft.com/en-us/library/windows/desktop/
           aa363298%28v=vs.85%29.aspx>.

[I-D.chen-pcp-mobile-deployment]
           Chen, G., Cao, Z., Boucadair, M., Ales, V., and L.
           Thiebaut, "Analysis of Port Control Protocol in Mobile
           Network", draft-chen-pcp-mobile-deployment-04 (work in
           progress), July 2013.

[I-D.ietf-dhc-container-opt]
           Droms, R. and R. Penno, "Container Option for Server
           Configuration", draft-ietf-dhc-container-opt-07 (work in
           progress), April 2013.

[I-D.ietf-pcp-dhcp]
           Boucadair, M., Penno, R., and D. Wing, "DHCP Options for
           the Port Control Protocol (PCP)", draft-ietf-pcp-dhcp-08
           (work in progress), August 2013.

[Appendix A](#).   Discussion of other PCP Discovery methods

   Several algorithms have been specified that allows PCP Client to
   discover the PCP Servers on a network .   However, each of this
   approaches has technical or operational issues that will hinder the
   fast deployment of PCP.

[A.1](#).   Default Router

   The PCP specification allows one mode of operation in which the PCP
   client sends its requests to the default router.   This approach is
   appropriate in a typical single-link residential network but has
   limitations in more complex network topologies.

   If PCP server does not reside in first hop router, whether because
   subscriber has a existing home router or in the case of Wireless
   Networks (3G, LTE) [[I-D.chen-pcp-mobile-deployment](#)], trying to send a
   request to default router will not work.

[A.2](#).   DHCP PCP Options

   One general drawback of relying on external configuration mechanisms,
   such as DHCP [[I-D.ietf-pcp-dhcp](#)], 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.

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

   There are two problems with DHCP Options: DHCP Server on Home
   Gateways (HGW) and Operating Systems DHCP clients

   Currently what the HGW does with the options it receives from the ISP
   is not standardized in any general way.   As a matter of practice, the
   HGW is most likely to use its own customer-LAN-facing IP address for
   the DNS server address.   As for other options, it's free to offer the
   same values to the client, offer no value at all, or offer its own IP
   address if that makes sense, as it does (sort of) for DNS.

   In scenarios where PCP Server resides on ISP network and is intended
   to work with arbitrary home gateways that don't know they are being
   used in a PCP context, that won't work, because there's no reason to
   think that the HGW will even request the option from the DHCP server,

much less offer the value it gets from the server on the customer-
facing LAN. There is work on the DHC WG to overcome some of these
limitations [I-D.ietf-dhc-container-opt] but in terms of deployment
it also needs HGW to be upgraded.

The problems with Operating Systems is that even if DHCP PCP Option
were made available to customer-facing LAN, host stack DHCP
enhancements are required to process or request new DHCP PCP option.
One exception is Windows [DhcpRequestParams]

Finally, in the case of IPv6 there are networks where there is DHCPv6
infrastructure at all or some hosts do not have a DHCPv6 client.

A.3. User Input

A regular subscriber can not be expected to input IP address of PCP
Server or network domain name. Moreover, user can be at a Wi-Fi
hotspot, Hotel or related. Therefore relying on user input is not
reliable.

A.4. Domain Name System Based

There are three separate category of problems with NAPTR [RFC3958]

1. End Points: It relies on PCP client determining the domain name
    and supporting certain DNS queries

2. DNS Servers: DNS server need to be provisioned with the necessary
    records

3. CPEs: CPEs might interfere with DNS queries and the DHCP domain
    name option conveyed by ISP that could be used to bootstrap NAPTR
    might not be relayed to home network.

A.5. Addressing only based on Destination Port

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.)

Appendix B. Discussion of IP Anycast Address usage for PCP

B.1. Motivation

   The two issues identified in Appendix A.5 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.

B.2. Scenarios

   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 [RFC6887].

   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.

B.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:

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

o 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.

o 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 appropriate 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:

o 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.

o 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).

Authors' Addresses

   Sebastian Kiesel
   University of Stuttgart Computing Center
   Allmandring 30
   Stuttgart 70550
   Germany

   Email: ietf-pcp@skiesel.de
   URI: http://www.rus.uni-stuttgart.de/nks/

   Reinaldo Penno
   Cisco Systems, Inc.
   San Jose, CA
   US

   Phone:
   Fax:
   Email: repenno@cisco.com
   URI:

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

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