

How PCP applies in the Multi-Egress NAT scene?

Author:
Zhangzhan(Channy.zhang@huawei.com)

Feb 14, 2014

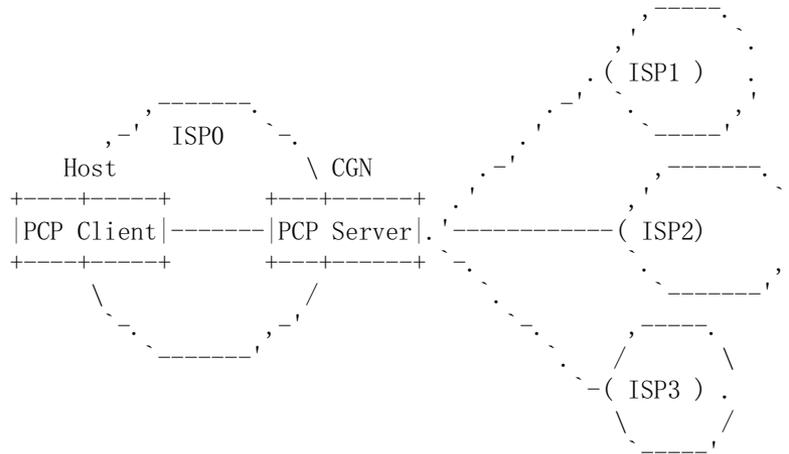


Table of contents

- ▶ Question
- ▶ Solution 1
- ▶ Solution 2
- ▶ Solution 3
- ▶ Conclusion

Question

In such a network topology:



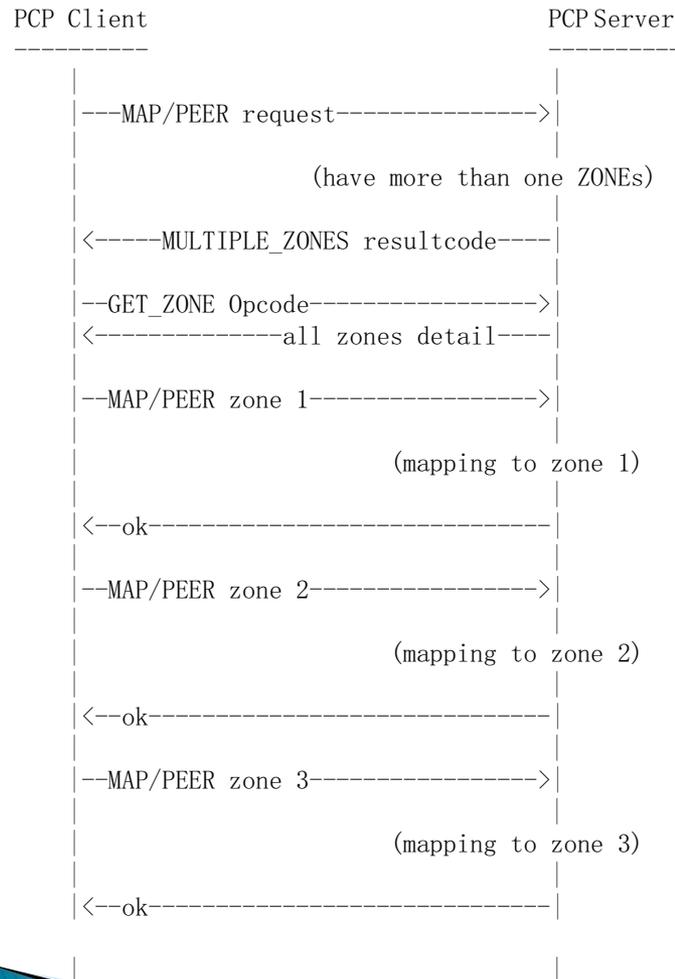
If the NAT gateway device (such as CGN) is PCP Server and has two or more out interfaces to internet were doing different NAT conversion for different ISP. Maybe using NAT address pool 1, 2, 3... And the routing entries to internet on NAT gateway is equal default route. In this case, the out interface is uncertain, fully in accordance with the equal route selection. Service traffic can be selected out interface according to the destination address and do different NAT conversion.

But PCP MAP mode message does not include the destination address, the NAT gateway device how to choose the out interfaces and the different NAT address pools?

As the upon topology, if users in ISP1 want to visit the host in ISPO, suppose the global NAT address of host is IP pool 1. Normally, if no PCP, if users in ISP2 want to visit the host, the global NAT address of host is IP pool 2. IP pool 1 and IP pool 2 are certainly different address segments, and belonging to different ISP.

Now PCP is used to mapping the global NAT address, suppose CGN select IP pool 1 as the optimal one. If users in ISP2 want to visit the host, must visit the IP pool 1 through ISP1's network. But usually this is not acceptable to IPS2.

Solution 1



We define a new result code, such as `MULTIPLE_ZONES`, and define a new Opcode, such as `GET_ZONE`.

If the PCP server have more than one ZONE, reply a `MULTIPLE_ZONES` result code message when receiving any PCP requests.

The client receives the result code, must send a `GET_ZONE` Opcode request message.

The PCP server must replay a message which contains the list of all ZONE ID information.

Then the client get the ZONE ID information, sends request messages (PEER or MAP) which include ZONEID option. If there are N zones, client must send N request messages.

Finally The PCP server can be selected the different address pools according to the ZONEID information and create multiple different mapping tables.

If the zone information is changed in PCP server, mapping tables must be deleted and create new mapping table when receiving renewing messages from clients.

Solution 2

```
PCP Client                                PCP Server
-----                                -----
|--MAP w/LEARN_ZONES option----->|
|                                     |
|         (one mapping is created)   |
|<--mapped to zone=1, total zones=3-|
|--GET_ZONE 1 details----->|
|<--zone 1 details----->|
|--GET_ZONE 2 details----->|
|<--zone 2 details----->|
|--GET_ZONE 3 details----->|
|<--zone 3 details----->|
|--MAP, ZONE=2----->|
|<--ok----->|
|--MAP, ZONE=3----->|
|<--ok----->|
```

LEARN_ZONE would be in the optional-to-process range, so the PCP client can safely send that option with any MAP request.

If the PCP server understands LEARN_ZONE, it includes LEARN_ZONE information in the response.

It show that the LEARN_ZONE information would merely indicate the number of zones supported.

If the PCP server does not understand LEARN_ZONE, the PCP server omits the LEARN_ZONE option from the response, but otherwise creates the requested MAP normally (that is, the PCP server behaves as if the optional-to-process LEARN_ZONE was not in the request at all).

Conclusion

The problem and the scene is very common in the enterprise network, especially the ISP side, it's an urgent problem to be solved. If this problem is not solved, maybe affect to generalize the PCP protocol.

These solutions are the initial results discussed on those days, hope to confirm a formal extending protocol instead of draft-penno-pcp-zones-01 as soon as possible.

Major players in this discussion:
Zhangzhan(channy.zhang@huawei.com)
Dan Wing(dwing@cisco.com)
Ted Lemon(ted.lemon@nominum.com)
Dave Thaler(dthaler@microsoft.com)