Network Working Group
INTERNET-DRAFT
Intended status: Standard Track
Expires: Sep 6, 2023
Mar 7, 2023

Louis Chan
Juniper Networks

Enhanced Port Forwarding functions with CGNAT
draft-chan-tsvwg-eipf-cgnat-02.txt

Abstract

There is a need for peer-to-peer (P2P) communication under the use of CGNAT in service providers. With the combination of home gateway, this becomes NAT444.

In RFC5128, methods of using UDP hole punching solves the problem partially when EIM (Endpoint-Independent Mapping) is supported in NAT device in the path, and there exists a common rendezvous server.

The success rate of UDP hole punching is high, but not TCP hole punching in practical world. Also, the P2P solution requires a common server in the public internet to exchange the IP and port information.

In this draft, a method is described to achieve incoming TCP or UDP session without a common rendezvous server in NAT444 situation.

Status of this Memo

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 Sep 6, 2023.

Copyright Notice

Table of Contents

1. Introduction

The purpose of this document is to describe to a way to allow incoming
TCP or UDP
sessions under NAT444 situation.

The success rate of TCP and UDP session would be guaranteed under this proposal.

There would be two sections in the draft.

- The first section describes a procedure for an application in end device to detect and allocate TCP or UDP port for its use for incoming session. The required tools are STUN [RFC5389] and UPNP [RFC6970].

- The second section describes a method for residential gateway RG to discover the usable port range under a CGNAT deployment with port-block-allocation. In turn, the home gateway could allocate TCP or UDP to the end devices via UPNP, NAT-PMP [RFC6886] or PCP [RFC6887].

2. Conventions used in this document

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].

In this document, these words will appear with that interpretation only when in ALL CAPS. Lower case uses of these words are not to be interpreted as carrying significance described in RFC 2119.

3. Port acquiring procedure in Application

```
PC1-----RG-------CGNAT------Internet------PC2
                               |
                               +-----STUN server
```

- Private network: PC1: 192.168.1.10, RG: 192.168.1.1
- WAN: RG: 10.1.1.20, CGNAT: 10.1.1.1
- CGNAT: public IP 100.1.1.1, PBA (port block allocation for RG) 1024-1055
- PC2: public IP 201.1.1.10

Here is an example of step to acquire a TCP or UDP port

- Application in PC1 sends a STUN request to STUN servers in public internet. The STUN server would reply the XOR-mapped-address. E.g.

  100.1.1.1:1024 ;public ip is 100.1.1.1 with port 1024

  This detects both public IP address and the UDP port available. This assumes the same TCP port is also available since most CGNAT implementations allocate the same port number for both TCP and UDP with EIM enabled.

The application will then send UPNP request to residential gateway RG, 192.168.1.1, for port forward TCP port 1024 to the local device IP, 192.168.1.10.

- CGNAT, due to PBA allocation and allow incoming session enabled, TCP traffic sent to 100.1.1.1:1024 as destination would be forwarded to RG 10.1.1.20:1024 without changing port value, when EIPF function is enabled. Then, RG would pass the TCP traffic to PC1 with 192.168.10.1:1024 as destination due to the registration of UPNP. In this case, PC2 could initiate a direct TCP session to PC1 via 100.1.1.1:1024.

Please note that in PBA allocation, 100.1.1.1:1024-1055 port range is always associated with this RG 10.1.1.20 only. This port range is not shared with other RGs or private IP.

- UDP would work in the same way. Any host in the internet could create TCP or UDP session directly with the application in PC1

The above procedure assumes both RG and CGNAT have EIM capability enabled.

The application in PC1, optionally, could release the UPNP mapping after finishing the session.

4. Endpoint Independent Port Forwarding (EIPF) Enhancement

4.1. When this EIPF feature enabled in CGNAT together with EIM

- It is possible that the public IP:PORT is already used in established outgoing connections. This is possible when port resource is re-usable.
- If there is a packet with destination 100.1.1.1:1024 in inbound direction, CGNAT first would check if there is an existing session established. If yes, it should follow the session table for translation. This session might be created by other outgoing session which could potentially share the same 100.1.1.1:1024 port.
- if there is no matching session in the CGNAT, it is a new incoming session. Then the associated TCP or UDP port is UNCHANGED, and just change the destination IP to 10.1.1.20.
- It is working like port forward function in a NAT44
- In the example, any IP source address, 202.1.1.1 or 222.1.1.1, sending traffic to 100.1.1.1:1024. CGNAT would translate the traffic as 10.1.1.20:1024 as destination.
- UDP hole punching would be compatible if the UDP session is still in RG and CGNAT session table. Port 1024 would follow the translation.

4.2. When this feature is enabled in CGNAT with both EIM and EIF

- EIF (Endpoint-Independent Filtering), described in RFC5128, will happen only if the external host already has a session through EIM.
- The TCP or UDP port is kept UNCHANGED for any other external hosts sending inbound traffic.
- For example, there is a session originated from PC1 to PC3, 201.1.1.20

```
            PC1-----------RG----------CGNAT-----------Internet---------PC3
                                                          |
+---------------------PC4

        Src: 192.168.1.10:3333    10.1.1.20:4444          100.1.1.1:1033

        Dst: 201.1.1.20:5555      201.1.1.20:5555         201.1.1.20:5555
```

When PC3 sends traffic with different source port, 201.1.1.20:6666 and destination 100.1.1.1:1033, CGNAT should honor the EIF behavior. It would be translated back to 10.1.1.20:4444.

When other host without any session established through EIM, and it sends traffic with destination port 1033, the port 1033 should not be changed at CGNAT.

When PC4 send traffic to 100.1.1.1:1033, the port 1033 is kept UNCHANGED. PC4 has no previous established sessions with PC1. This follows the EIPF behavior.

This implementation is an optional with EIF enabled.

Another option is to make EIPF and EIF exclusive. EIPF could be just implemented with or without EIM enabled.

5. Co-existence with established session in CGNAT

It is allowed that a TCP or UDP port could be shared with outgoing sessions from CGNAT perspective. Here is an illustration based on the example in section 3.

Behind RG, there are additional PC1a, 192.168.1.11 and PC1b, 192.168.1.12, and all of these can share port 1024 for outgoing at CGNAT. For example,

	dest ip:port	src ip:port@PC	src ip:port@RG	src ip:port@CGNAT
PC1a	2.2.2.2:888	192.168.1.11:1234	10.1.1.20:4444	100.1.1.1:1024
PC1b	3.3.3.3:999	192.168.1.12:5678	10.1.1.20:5555	100.1.1.1:1024

there are two sessions appeared in internet (src 100.1.1.1:1024, dst 2.2.2.2:888) and (src 100.1.1.1:1024, dst 3.3.3.3:999).

In fact, it is possible to have more sessions from PC1a or PC1b sharing port 1024 as long as destination IP:port is different. It is up to CGNAT implementation.

Since sessions are created in NAT tables inside RG and CGNAT, these sessions will co-exist with any new incoming sessions, providing that no clash of IP and port pairs.

When PC2, initiates a session (src 201.1.1.10:6666, dst 100.1.1.1:1024) toward PC1, CGNAT will look up the local NAT session table first. If there is no match, it is a new session, and accepted with EIPF behavior.

RG will do the same. Traffic with destination port 1024 will be forwarded to device who make a reservation via UPNP or NAT-PMP.

6. Requirement on CGNAT and RG

6.1. CGNAT requirement

For CGNAT, the mandatory requirement is that one public IP:port must only associate with one private IP. This allows unique translation in the incoming direction. EIM or EIF are optional features, and it is discussed in Section 4.

It should be noted that same UDP and TCP port should be assigned to the same private IP. Otherwise, the device behind RG is required to perform STUN based on TCP, which is less commonly available today.

Port block allocation, in the example, is not mandatory but recommended in fixed line use case.

6.2. RG requirement

For RG, it is mandatory to support port forwarding with UPNP. NAT-PMP support is optional but recommended.

EIM support is optional. If EIM is support, it would be a fallback means with UDP hole punching.

7. Other considerations

In this draft, it assumes the port number that STUN procedure detects externally via UDP, is also available for TCP. In practical world, this is likely the case.

When the port is successfully allocated from RG, the application should make a verification of the incoming connection via other means. If CGNAT supports hair-pinning session, it could be verified without external help.

And how the IP and port information is conveyed to third parties is not discussed here. It is out the scope of this document.

There is a chance that RG would receive new private IP due to reboot or IP refresh.
And there is a chance of change in CGNAT translation due to failure recovery. In
this case, it is the responsibility of application to detect such change. It is
advised that the application should periodically detect any IP change.

8. Retrieval of IP and port information via HTTP

The internet service provider host a HTTP web server for the enquiry of IP and port
information. Two URI's are suggested

8.1. IP and port - URI /ipport/

With the URI /ipport/, the HTTP response is clear text with IP:PORT, where IP is the external public IP address and the PORT is external port as seen.

For example, the response is

100.1.1.1:1040

The HTTP response should be human readable with a web browser.

Although TCP port 1040 is seen here, it is assumed that UDP port 1040 is also available from CGNAT for incoming mapping.

8.2. IP and port range - URI /ipportrange/

With the URI /ipportrange/, the HTTP response is clear text with

IP:PORT_START:PORT_END<LF>

IP:PORT_START:PORT_END<LF>

IP:PORT_START .. ..

Where <LF> is ASCII character for line feed.

The response is a human readable format in a normal web browser.

For examples, here are valid responses

a) Single line

100.1.1.1:1024:1031

Port range 1024 to 1031 assigned for both TCP and UDP.

b) Two lines

100.1.1.1:1024:1031

100.1.1.1:1064:1071

Port ranges 1024 to 1031 and range 1064 and 1071 are assigned for both TCP and UDP.

It is possible to have multiple port block allocated to the same private

IP address
from CGNAT perspective.

If the RG device or application could not support multiple entries of IP and port range, it should take one of the lines, preferably the first line.

Human user or RG could use this information to plan for incoming services. For example, when PC1 requests a TCP 8888 port forward from RG via UPNP [RFC6970], NAT-PMP [RFC6886] or PCP [RFC6887], RG would counter offer another TCP port 1031.

9. Compatibility

There would be no obvious compatibility problem with existing implementation methods.

There is a possibility when more than 2 level of NAT is used. This is not scoped in this document.

10. Security Considerations

When EIPF is enabled in CGNAT, more incoming traffic would be allowed sending to RG.

It would be the RG as the gatekeeper for blocking unwanted sessions. This would be the same scenario as public IP is assigned.

For the CGNAT, there would be more session attempt to handle. Incoming session limit or attempted session per seconds kind of parameters could be considered as security measure. Optionally, it is allowed to open the port only when STUN procedure for the port is seen.

11. References

11.1. Normative References

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

11.2. Informative References

[RFC5128] Srisuresh, P., Ford, B., and D. Kegel, "State of Peer-to-Peer (P2P) Communication across Network Address Translators (NATs)", March 2008.

[RFC5389] Rosenberg, J., Mahy, R., Matthews, P., and D. Wing, "Session Traversal Utilities for NAT (STUN)", October 2008.

[RFC6886] S. Cheshire and M. Krochmal. "NAT Port Mapping Protocol (NAT-PMP)", April 2013.

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

[RFC6970] Boucadair, M., Penno, R., and D. Wing, "Universal Plug and Play (UPnP) Internet Gateway Device - Port Control Protocol Interworking Function (IGD-PCP IWF)", July 2013

12. Acknowledgments

The following people have contributed to this document:

Author's Address

Louis Chan (editor)
Juniper Networks
Suite 3001-7, 30/F, Tower 2
Time Square, Causeway Bay
Hong Kong

Phone: +852-38562100
Email: louisc@juniper.net