NAT Behavioral Requirements, as Defined by the IETF (RFC 4787) - Part 1. Mapping Behavior

September 12, 2013 | By Netmanias (tech@netmanias.com)

Many applications like Skype, online games, etc. are all Peer-to-Peer (P2P) applications based on UDP. When running these applications, two devices communicate with each other without a server in between. We learned in our previous post that Korean telecom operators have already employed NATs in all of their access networks (Wi-Fi, 3G and LTE) except for wired networks.

These P2P applications and NATs do not really get along with each other (usually P2P apps are victimized by NATs).
It is impossible for two user devices with different private IP addresses in different locations (one behind a NAT, the other on the other side of the NAT) to directly communicate with each other through a NAT. This is because an external host cannot connect (i.e. send a packet) to an internal host behind a NAT first. For example, if device 1 (e.g. host A in the figure below) sends a packet to device 2 (e.g. host B), a NAT ahead of device 2 will drop the packet. If device 2 sends one to device 1, this time the NAT ahead of device 1 will drop it.

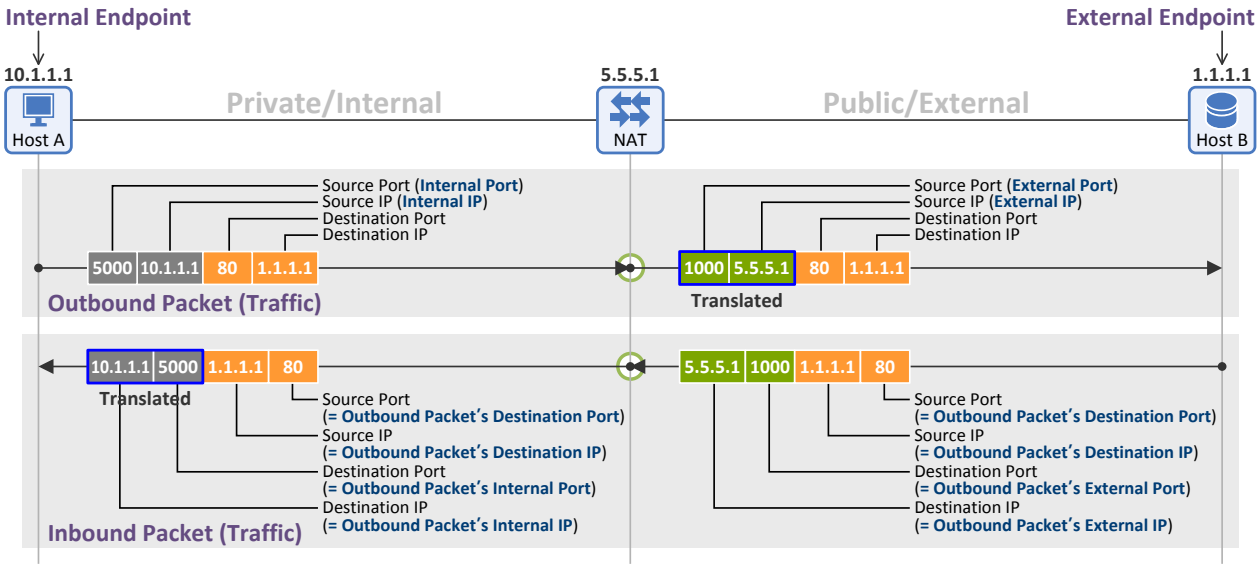
To address this issue, NAT Traversal techniques, such as Session Traversal Utilities for NAT (STUN), RFC 5389/RFC 5780, Traversal Using Relays around NAT (TURN), RFC 5766, Interactive Connectivity Establishment (ICE), RFC 5245, etc., have been standardized. These techniques can be summarized as follows:

- **STUN**: This technique describes how a host (a STUN client) communicates with a STUN server (a server with public IP addresses) to find out i) if the host is within a private network (i.e. has a NAT), ii) if any, how the NAT behaves, and iii) what public IP addresses and source port values are translated by the NAT, etc.
- **TURN**: This describes how a host behind a NAT communicates with another host located on the opposing side of the NAT through a "Turn Server", a relay server with public IP addresses. One drawback of this method is that the communication has to be through the relay server. So, this will inevitably slow response.
- **ICE**: It describes how to find an optimal way of establishing sessions between hosts using STUN or TURN.

As such, decisions on which NAT Traversal technique to use are made based on the operational characteristics of a NAT. So, in 2007, "NAT Behavioral Requirements for Effective Nat Traversal" were standardized in RFC 4787.

The three subsequent posts will present the ideal NAT behaviors for P2P applications described in the RFC 4787.

Before we continue, please see the following definitions of some important terms first.



- **Internal Endpoint**: a user device that has a private IP address and is located behind a NAT, e.g. Host A in the figure above (e.g. a user device that is located in the same operator network as a NAT)
- **External Endpoint**: a user device that has a public IP address and is located on the opposing side of the NAT, e.g. Host B in the figure above (e.g. a user device that is not located in the same operator network as a NAT)
- **Outbound Packet (Traffic)**: a packet (traffic) sent from an Internal Endpoint to an External Endpoint via a NAT
- **Inbound Packet (Traffic)**: a packet (traffic) sent from an External Endpoint to an Internal Endpoint via a NAT
- **Internal Address and Internal Port**: the source IP (10.1.1.1) and source Port (5000) of a packet sent by an Internal Endpoint (Host A)
- **External Address and External Port**: the source IP (5.5.5.1) and source Port (1000) of a packet translated by a NAT and then sent to an External Endpoint (Host B)
- In general, the destination information (i.e. destination IP (1.1.1.1) and destination Port (80)) of a packet sent by an Internal Endpoint (Host A) is forwarded to an External Endpoint (Host B) transparently without being translated by a NAT.
- When an External Endpoint (Host B) receives the packet, it returns a packet comprising of following information to the Internal Endpoint as a response:
  - Destination IP = the source IP address of the received packet, i.e. the External Address (5.5.5.1)

- Destination Port = the source Port of the received packet, i.e. the External Port (1000)
- Source IP = the destination IP address of the received packet (1.1.1.1), i.e. the IP address of the External Endpoint (Host B)
- Source Port = the destination port of the received packet (80)

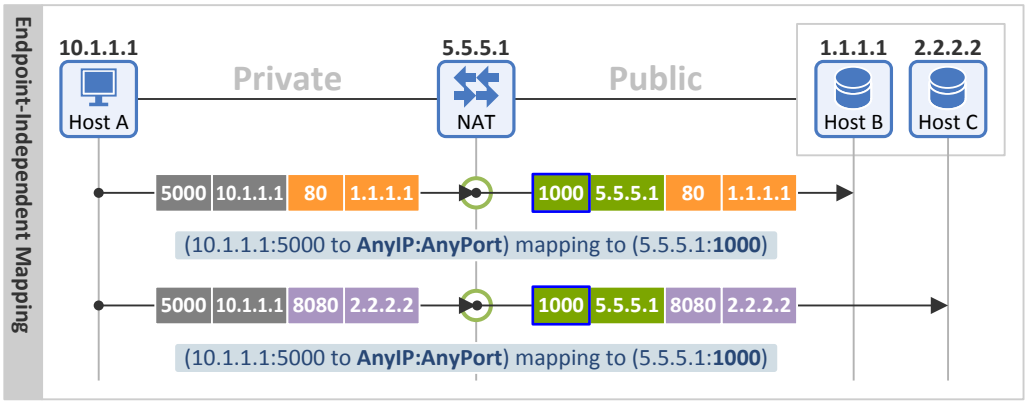
1. Network Address and Port Translation Behavior

1.1 Address and Port Mapping

■ Endpoint-Independent Mapping

In "Endpoint-Independent Mapping", the Endpoint refers to an External Endpoint.

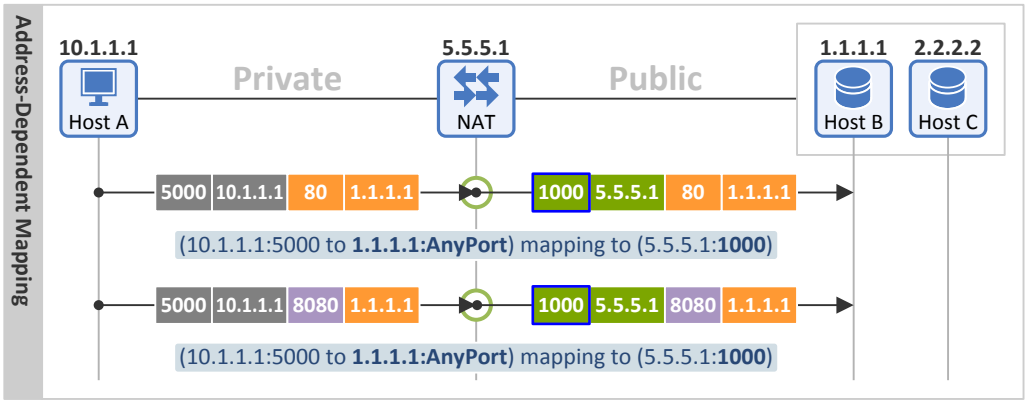
"Endpoint-Independent Mapping" assigns the same External Port Mapping value (translated port = 1000) to packets sent by an Internal Endpoint (Host A), as long as the packets have i) the same **source IP address** (10.1.1.1) and ii) the same **source port** (5000), regardless of their destination IP address (1.1.1.1 or 2.2.2.2) or destination port (80 or 8080).



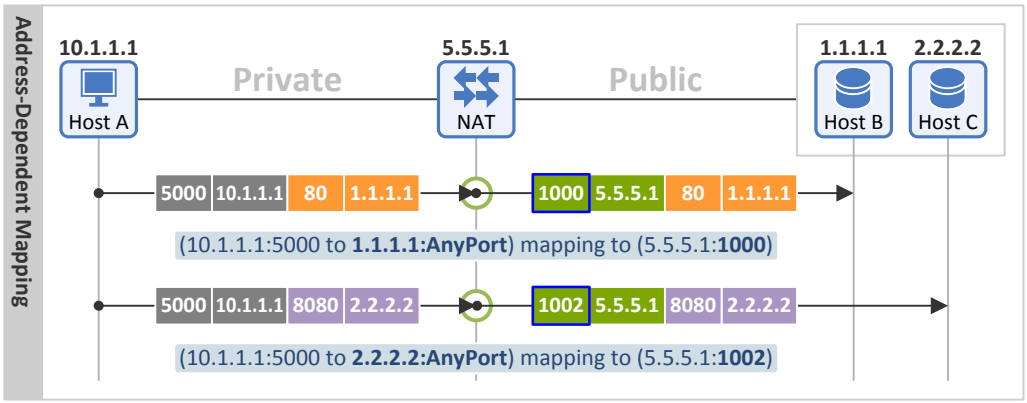
■ Address-Dependent Mapping

In "Address-Dependent Mapping", the Address refers to the destination IP address of a packet sent by an Internal Endpoint.

"Address-Dependent Mapping" assigns the same External Port Mapping value (translated port = 1000) to packets sent by an Internal Endpoint (Host A) if the packets have i) the same **source IP address** (10.1.1.1), ii) the same **source port** (5000) AND iii) the same **destination IP address** (1.1.1.1), regardless of their destination port (80 or 8080).



In the figure below, for the two packets with the same source IP address (10.1.1.1) and source port (5000), but with different destination IP addresses (1.1.1.1 and 2.2.2.2), different External Port Mapping values (translated port = 1000 and 1002) are used.

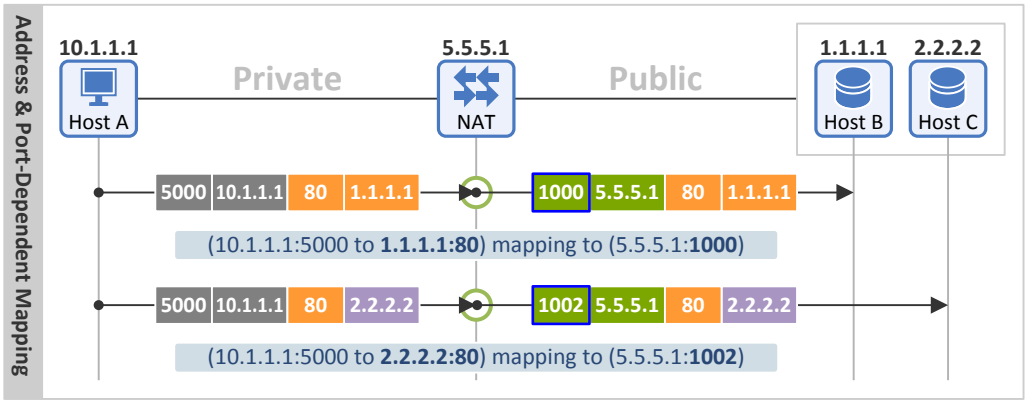


■ **Address and Port-Dependent Mapping**

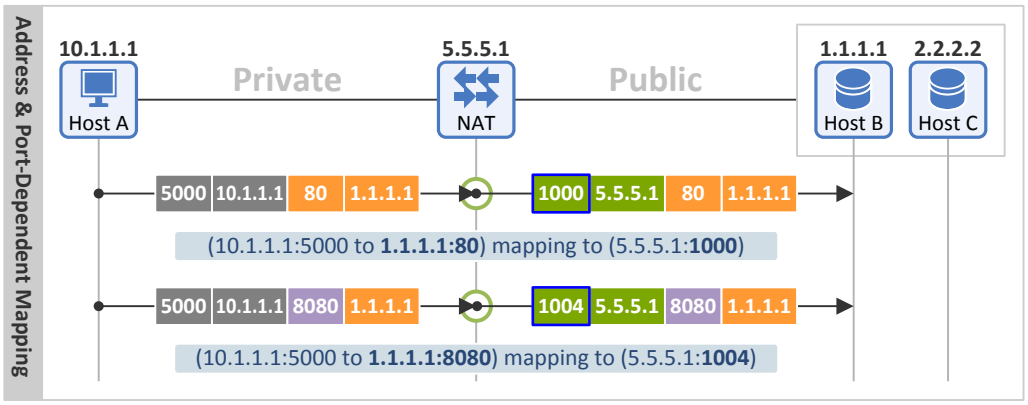
In "Address and Port-Dependent Mapping", the Address and Port respectively refer to the destination IP address and destination port of the packet sent by an Internal Endpoint.

"Address and Port-Dependent Mapping" assigns the same External Port Mapping value to packets sent by an Internal Endpoint (Host A), only if the packets have i) the same **source IP address**, ii) the same **source port**, iii) the same **destination IP address**, AND iv) the same **destination port**.

In the figure below, different External Port Mapping values (translated Port = 1000 and 1002) are used for the two packets because they have different destination IP addresses (1.1.1.1 and 2.2.2.2).



In the figure below, different External Port Mapping values (translated port = 1000 and 1004) are used because the two packets have different destination ports (80 and 8080).



RFC 4787 Recommendation (REQ-1): A NAT MUST have an "Endpoint-Independent Mapping" behavior

RFC 4787 states that "Failure to meet REQ-1 will force the use of a UDP relay, which is very often impractical".

1.2 IP Address Pooling

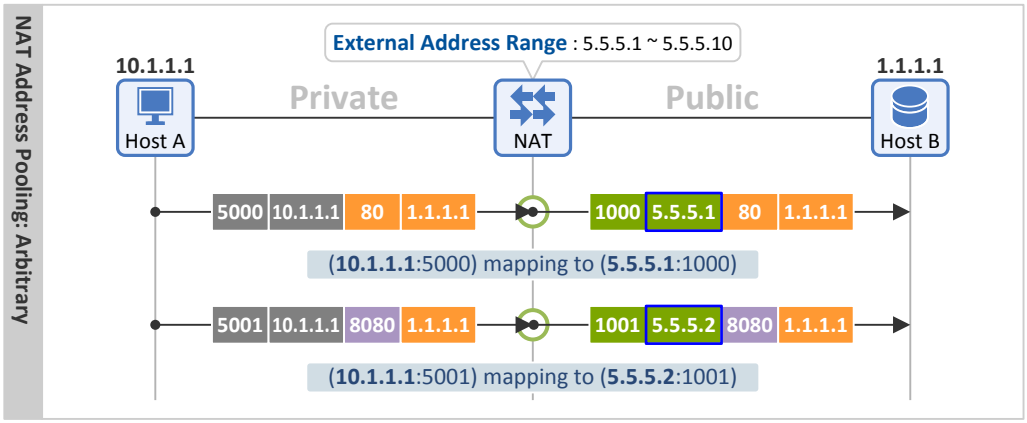
Home APs and Wi-Fi hotspot APs use one public IP address to perform NAPT (Network Address Port Translation). However, Large Scale NATs (LSNs, also called Carrier Glade NATs (CGNs)) employed in 3G/LTE networks use multiple public IP addresses (a pool of IP addresses on the external side of the NAT).

■ **Arbitrary**

NATs use different External IP addresses even for packets sent by one Internal Endpoint (i.e. those with the same source IP addresses), if their sessions (tuple of {source IP, source port, destination IP, destination port}) are different.

As seen in the figure below, the NAT allocates two different External IP addresses (5.5.5.1 and 5.5.5.2) for two different sessions that the Internal Endpoint 10.1.1.1 (Host A) has established to the External Endpoint 1.1.1.1 (Host B).

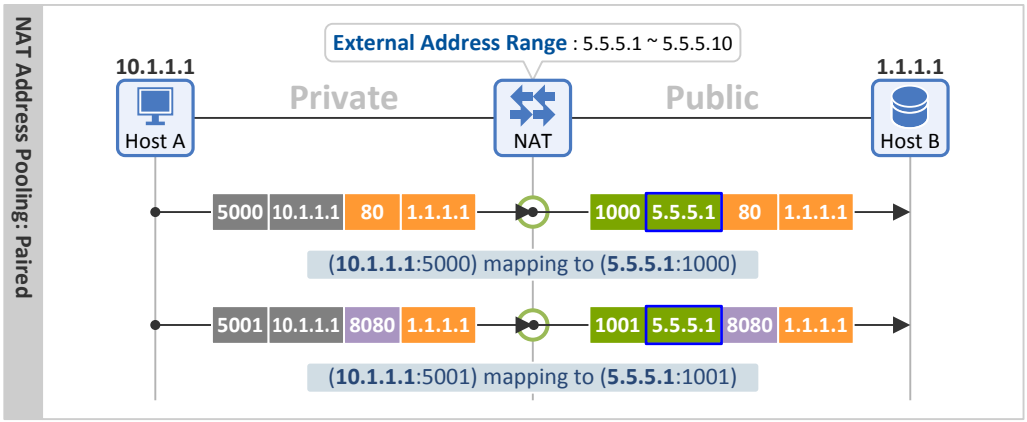
- Session 1: {10.1.1.1:5000 to 1.1.1.1:80} -> {5.5.5.1:1000 to 1.1.1.1:80}
- Session 2: {10.1.1.1:5001 to 1.1.1.1:8080} -> {5.5.5.2:1001 to 1.1.1.1:8080}



■ **Paired**

NATs use the same External IP address for packets sent by one Internal Endpoint (i.e. those with the same source IP addresses) even when their sessions (tuple of {source IP, source port, destination IP, destination port}) are different.

As seen in the figure below, the NAT allocates the same External IP addresses (5.5.5.1) for two different sessions that the Internal Endpoint 10.1.1.1 (Host A) has established to the External Endpoint 1.1.1.1 (Host B).

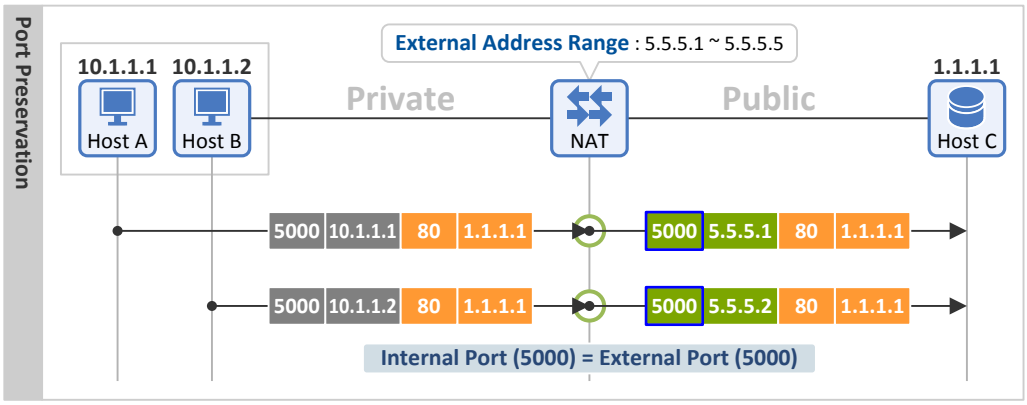


RFC 4787 Recommendation (REQ-2): It is RECOMMENDED that a NAT have an "IP address pooling" behavior of "Paired"

1.3 Port Assignment

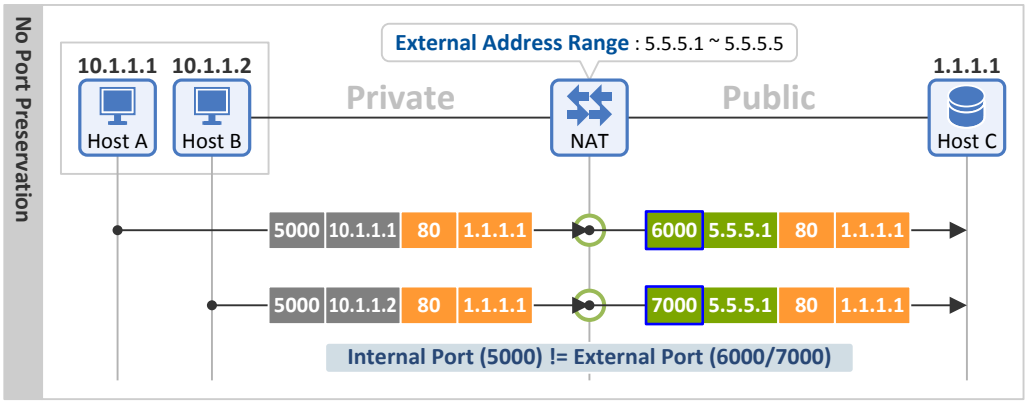
■ **Port Preservation**

NATs preserve the source port (Internal/Local Port) value used by the Internal Endpoint which sent a packet, even after implementing NAT (External Port = Internal Port).



■ **No Port Preservation**

NATs do not preserve the source port (Internal Port) value used by the Internal Endpoint, and allocate a source port (External Port) value randomly (External Port != Internal Port).



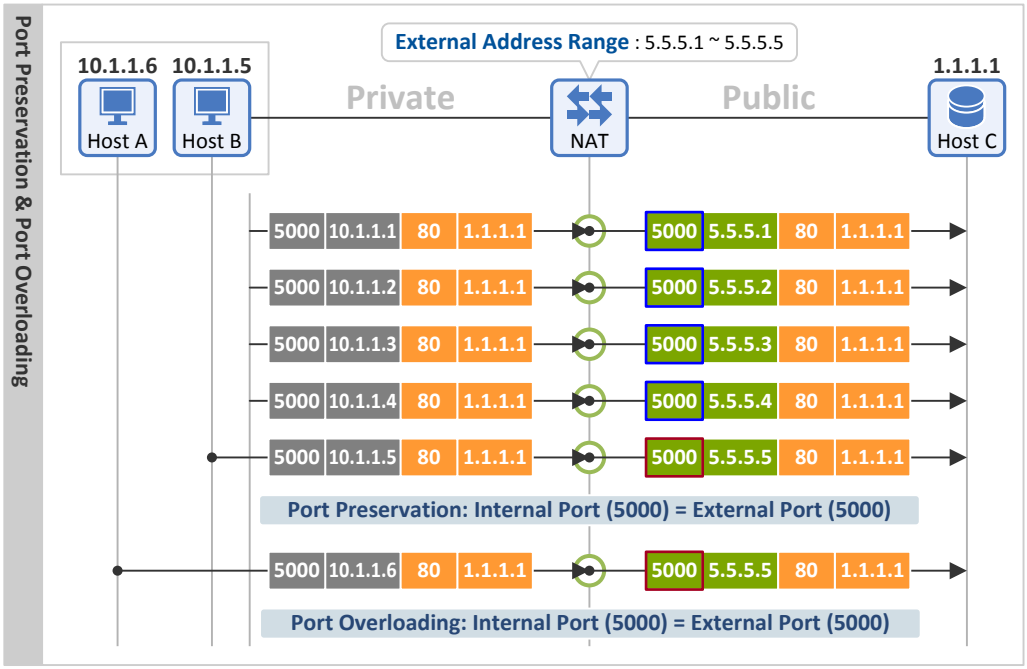
■ **Port Overloading (in Port Preservation)**

Let us suppose a NAT supporting port preservation runs out of External IP addresses (public IP addresses). What should the NAT do if an outbound packet with the same source port value arrives?

Port overloading is a simple but reckless method. NATs simply overwrite an existing binding entry in case of port collisions. That is, they always use port preservation. Then, as seen in the figure below, the binding entry generated for Host B would be overwritten by Host A.

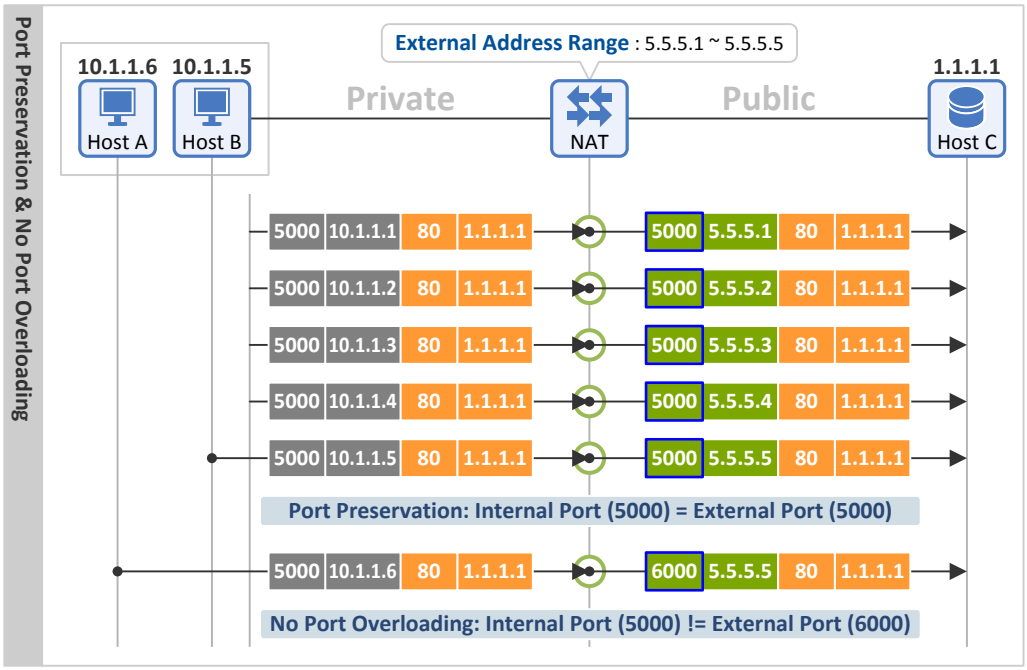
As a result, each packet that Host A and Host B sent to Host C through the NAT would have the same External Address (5.5.5.5) and External Port (5000). This will lead the NAT to forward all the inbound

packets from Host C to Host A. Eventually, communication between Host B and Host C will be impossible. Of course, no vendor would actually manufacture products that support this method.



■ **No Port Overloading (in Port Preservation)**

In the event of port collisions, NATs do not use port preservation, and instead they allocate the External Port a value different from that of the Internal Port.



RFC 4787 Recommendation (REQ-3): A NAT MUST NOT have a "Port assignment" behavior of "Port overloading"

1.4 Port Assignment Rule

The RFC also includes "Export Port Assignment Rules" for NATs supporting "No Port Preservation". The Internet Assigned Numbers Authority (IANA) defines port ranges as follows:

- Well-Known: 0 ~ 1023 (as standardized by IANA. e.g. HTTP = 80)
- Registered: 1024 ~ 49151 (not standardized by IANA, but widely used)
- Dynamic/Private: 49192 ~ 65535

Depending on NAT implementation, the following External Port assignment rules can be applied:

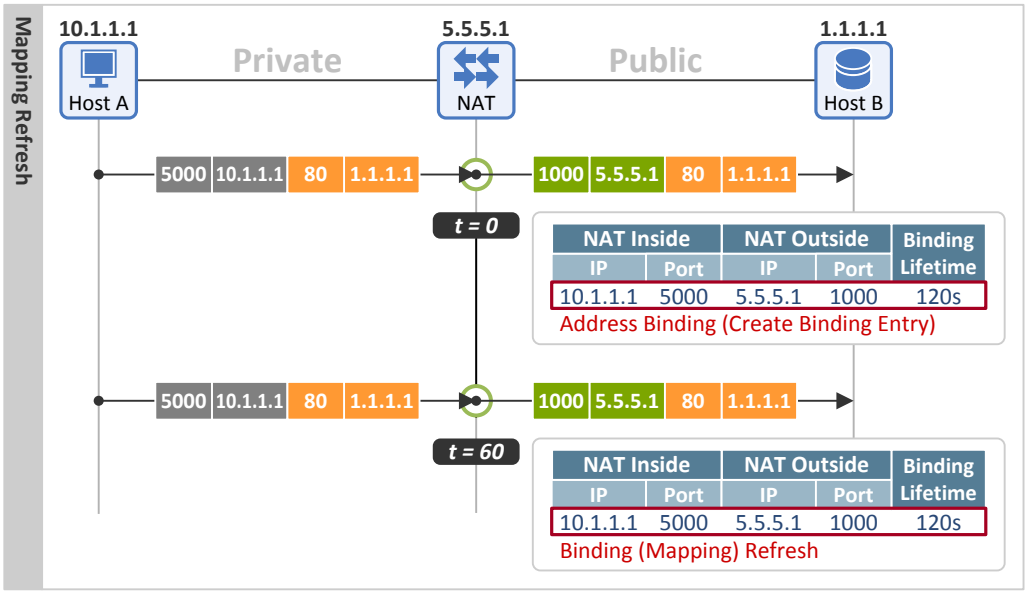
- If a NAT uses values from dynamic/private port range (49192 ~ 65535) for External Ports, the maximum number of NAT sessions supportable by one public IP address is limited to 16,000.
- If a NAT uses values from the ranges excluding well-known range (1024 ~ 65535) for External Ports, the NAT uses values from dynamic/private port range first, and then ones from registered range if needed.

1.5 Mapping Timer (Binding Refresh Timer)

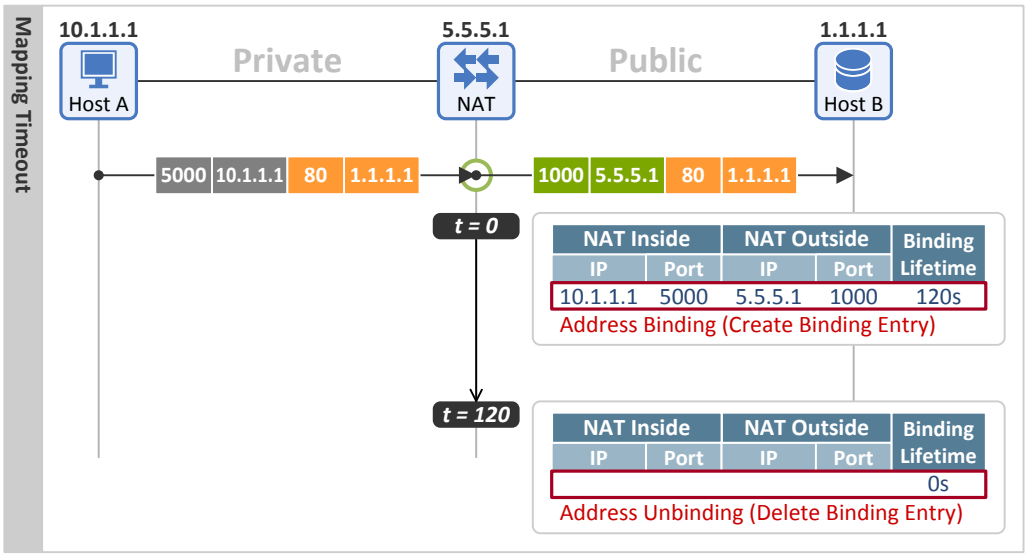
A binding entry generated in a NAT table by outbound traffic remains valid if there is traffic that is mapped using the same binding entry. However, if there is no traffic, the entry is deleted from the table when a mapping timer (also called binding refresh timer, or binding lifetime) expires.

If a NAT has a short mapping timer, a device (NAT-friendly application) will have to send keep-alive packets quite often to keep NAT sessions active. This will not be problematic for subscribers using wired or Wi-Fi networks. But, for 3G/LTE network subscribers who pay for what they use, this can be problematic.

In the figure below, the NAT mapping timer is set two minutes. At t=0, Host A sends its first packet and generates a binding entry for the packet. Then one minute later, when Host A sends another packet, the binding entry is refreshed (reset) to two minutes again.



As opposed to that, the NAT binding entry in the figure below is deleted after two minutes without traffic.



RFC 4787 Recommendation (REQ-5): A NAT UDP mapping timer MUST NOT expire in less than two minutes, unless REQ-5a applies

a) For specific destination ports in the well-known port range (ports 0-1023), a NAT MAY have shorter UDP mapping timers that are specific to the IANA-registered application running over that specific destination port

b) The value of the NAT UDP mapping timer MAY be configurable

c) A default value of five minutes or more for the NAT UDP mapping timer is RECOMMENDED

1.6 Mapping Refresh Behavior

■ **NAT Outbound refresh behavior of "True"**

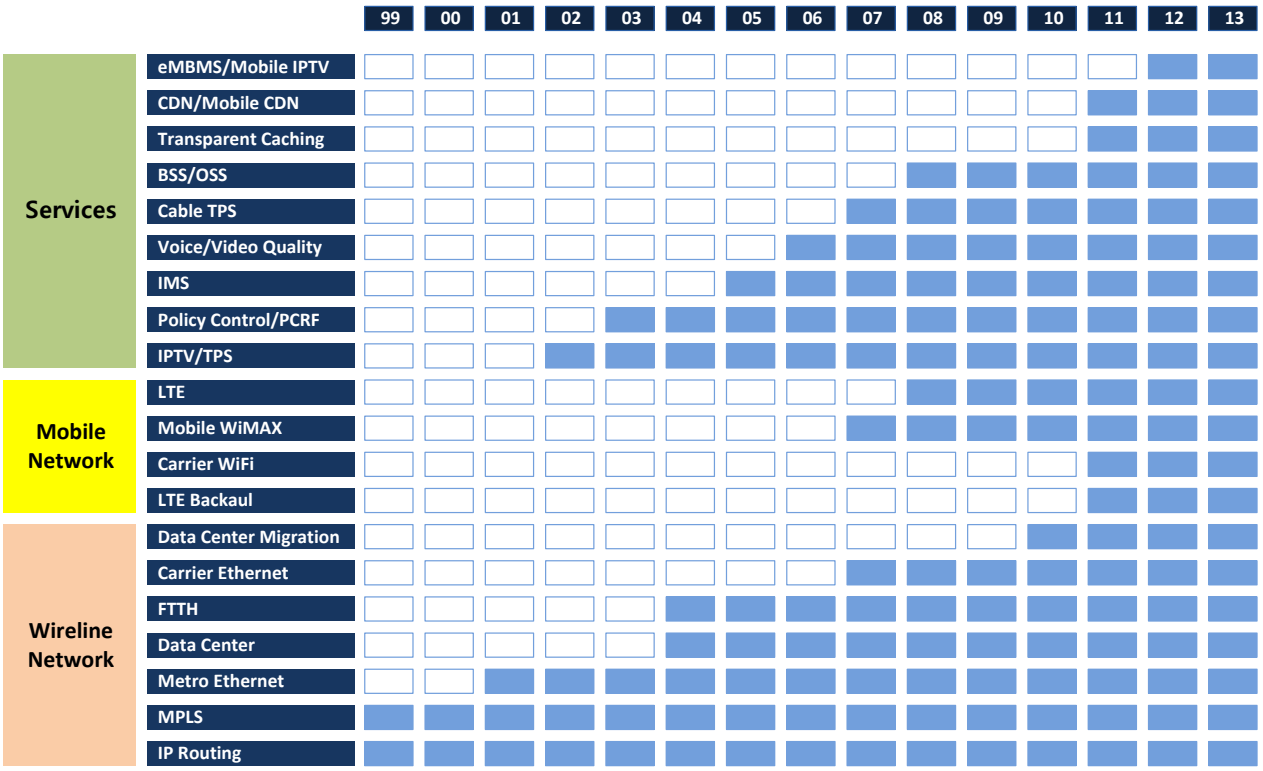
When a mapping timer is refreshed by outbound traffic (a packet sent by an Internal Endpoint to an External Endpoint)

■ **NAT Inbound refresh behavior of "True"**

When a mapping timer is refreshed by inbound traffic (a packet sent by an External Endpoint to an Internal Endpoint)

RFC 4787 Recommendation (REQ-6): The NAT mapping Refresh Direction MUST have a "NAT Outbound refresh behavior" of "True"

Netmanias Research and Consulting Scope

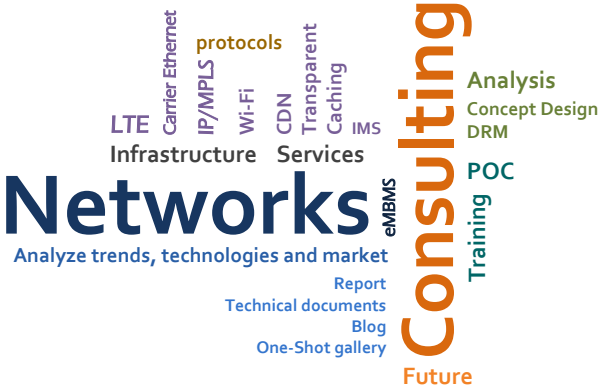


Visit http://www.netmanias.com to view and download more technical documents.

We design the future

We design the future

We design the future



About NMC Consulting Group (www.netmanias.com)

NMC Consulting Group is an advanced and professional network consulting company, specializing in IP network areas (e.g., FTTH, Metro Ethernet and IP/MPLS), service areas (e.g., IPTV, IMS and CDN), and wireless network areas (e.g., Mobile WiMAX, LTE and Wi-Fi) since 2002.