ETSI TR 103 829 V1.1.1 (2022-08)

Lawful Interception (LI);
IP address retention and traceability

Reference
DTR/LI-00204

Keywords
CGNAT, IP address resolution, IP retention, NAT, PAT, traceability

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - APE 7112B
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° w061004871

Contents

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™** and **LTE™** are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM**® and the GSM logo are trademarks registered and owned by the GSM Association.

Foreword

This Technical Report (TR) has been produced by ETSI Technical Committee Lawful Interception (LI).

Modal verbs terminology

In the present document "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Executive summary

The present document acts as a guide for policy makers, communication service providers and law enforcement agencies, regarding the retention of IP addresses for law enforcement purposes.

Introduction

The present document provides information regarding typical IP usage and Network Address Translation within a Communication or Internet Service Providers network. The present document does not attempt to describe all possible implementation variations of NAT but instead focuses on the key underlying principles of IP Communication and typical NAT implementation patterns.

Through understanding these concepts, the reader can gain an appreciation of the impact these techniques have on traffic attribution as observed from outside the Communication Service Providers' private network.

1 Scope

The present document considers the following aspects of IP address retention and traceability:

1) Basic Internet Protocol principles, highlighting specifically how IP addresses and ports are used to access the internet.

2) Key differences between IPv4, IPv6, Dual Stack and other relevant layer 3 protocols.

3) How IP addresses are allocated within networks, including EPC and 5GC, documenting any differences in approach.

4) The role of Network and Port address translation within a CSP network.

5) The impact address translation technologies on IP address attribution as observed from outside the CSP network. This includes a discussion on the different translation technologies commonly used by CSPs (e.g. NAT, PAT, CGNAT, NAT64).

6) Description of the key elements which define user and IP address association and therefore make up the minimal set of stored attributes for a viable IP retention solution.

7) Methods for accessing records of IP and port allocation from within a CSP network.

8) Methods for retaining and querying stored IP association records; consideration is given to the storage volumes, durations and accuracy.

The present document does not consider TOR®, VPN services or over top identity protection services, which may impact the ability to attribute observed IP addresses to a specific User Equipment.

2 References

2.1 Normative references

Normative references are not applicable in the present document.

2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication ETSI cannot guarantee their long term validity.

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

[i.1] Wikipedia: "IPv4".

NOTE: Available at https://en.wikipedia.org/wiki/IPv4.

[i.2] Wikipedia: "IPv6".

NOTE: Available at https://en.wikipedia.org/wiki/IPv6.

[i.3] IETF RFC 8799: "Limited Domains and Internet Protocols".

NOTE: Available at https://datatracker.ietf.org/doc/html/rfc8799.

[i.4] ETSI TS 103 280: "Lawful Interception (LI); Dictionary for common parameters".

[i.5] Wikipedia: "Regional Internet Registry".

NOTE: Available at https://en.wikipedia.org/wiki/Regional_Internet_registry.

[i.6] Wikipedia: "IPv6-deployment".

NOTE: Available at https://en.wikipedia.org/wiki/IPv6-deployment.

[i.7] ETSI TS 123 501: "5G; System architecture for the 5G System (5GS) (3GPP TS 23.501)".

[i.8] Recommendation ITU-T E.164: "The international public telecommunication numbering plan".

[i.9] IETF RFC 7422: "Deterministic Address Mapping to Reduce Logging in Carrier-Grade NAT Deployments".

NOTE: Available at https://datatracker.ietf.org/doc/html/rfc7422.

3 Definition of terms, symbols and abbreviations

3.1 Terms

Void.

3.2 Symbols

Void.

3.3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

4G	4th Generation Mobile Technology
5G	5th Generation Mobile Technology
5GC	5G Core Network
AAA	Authentication, Authorization, and Accounting
APN	Access Point Name
CGNAT	Carrier Grade Network Address Translation
CPE	Consumer Premises Equipment
CSP	Communication Service Provider
DNS	Domain Name Server
DHCP	Dynamic Host Configuration Protocol
EPC	Evolved Packet Core
FTP	File Transport Protocol
GB	Gigabyte
IANA	Internet Assigned Numbers Authority
ICANN	Internet Corporation for Assigned Names and Numbers
ID	Identifier
IMEI	International Mobile Equipment Identifier
IMEISV	International Mobile Subscriber Identity and Software Version
IMSI	International Mobile Subscriber Identity
IP	Internet Protocol
IPv4	Internet Protocol version 4
IPv6	Internet Protocol version 6
ISP	Internet Service Provider
LEA	Law Enforcement Agency
LI	Lawful Intercept
LSN	Large Scale Network Address Translation
MAC	Media Access Control
MSISDN	Mobile Station International Subscriber Directory Number

NAT	Network Address Translation
NIC	Network Interface Controller
OTT	Over The Top
PAT	Port Address Translation
PCF	Policy and Charging Function
PCRF	Policy and Charging Rules Function
PDU	Protocol Data Unit
PGw	Packet Gateway
RADIUS	Remote Authentication Dial-In User Service
RIR	Regional Internet Registry
SBI	Service Based Interface
SMF	Session Management Function
SPAN	Switched Port Analyser
SQL	Structured Query Language
SSL	Secure Sockets Layer
TB	Terabyte
TOR	The Onion Router
TR	Technical Report
UDM	Unified Data Management
UE	User Equipment
UPF	User Plane Function
URL	Uniform Resource Locator
UTC	Coordinated Universal Time
VPN	Virtual Private Network
Wi-Fi	Wireless Fidelity

4 IP address retention and traceability

4.1 Basic Internet Protocol principles, highlighting specifically how IP addresses and ports are used to access the internet

What is an IP?

The Internet Protocol (IP) is a fundamental building block for communication between entities over a computer network, such as the internet. The Internet Protocol provides a mechanism for passing information between connected end points through the use of IP addresses. Within the context of the internet an IP address can be considered analogous to a postal address, it is owned by a specific entity that can use this address to send and receive information to and from other similar end points that are connected to the same network.

IPv4 addresses are expressed as a set of four numbers separated by the '.' character, an example address might be 192.168.10.1. Each of the four numbers within the IP Address string can range from 0 to 255 allowing the full IP addressing range to go from 0.0.0.0 to 255.255.255.255.

IPv6 addresses consist of eight blocks of 16 bits each. Each group is written as four hexadecimal digits (sometimes called hextets or more formally hexadectets and informally a quibble or quad-nibble) and the groups are separated by colons (:). An example of this representation is 2001:0db8:0000:0000:0000:ff00:0042:8329.

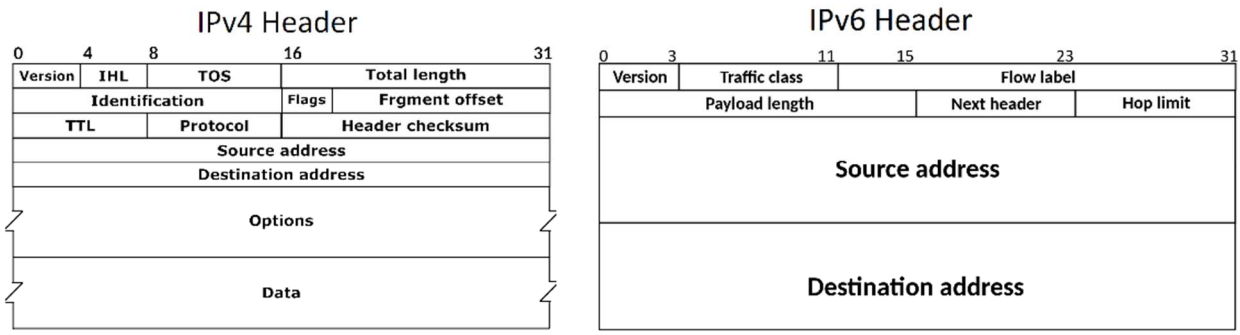

Figure 1: IPv4 and IPv6 Headers (see [i.1] and [i.2])

IP addresses are produced and allocated by the Internet Assigned Numbers Authority (IANA), which is part of the not for profit Internet Corporation for Assigned Names and Numbers (ICANN). Each time anyone registers a domain (website address) on the internet, they go through a domain name registrar, who pays a small fee to ICANN to register the domain.

Are there different types of subscriber IP address?

All IPv4 or IPv6 addresses are expressed in the same way but there are different categories of IP addresses, and within each category, different types.

Private IP addresses

Devices that connect to the internet usually belong to a private network and as such are allocated a private IP address, this includes laptops and mobile devices. Private IP addresses are generated and allocated for each device within the private network allowing them to communicate.

For further details regarding private IP addresses and their used in limited domains see IETF RFC 8799 [i.3].

Public IP addresses

A public IP address is an IP address that can be accessed directly over the internet, in the case of a broadband connection this could be the Public IP which is allocated to your router by the ISP. In the case of a mobile network, this is likely to the IP address that will be used to route IP traffic onto the internet after Network Address Translation (NAT).

Public IP addresses come in two forms - dynamic and static.

Dynamic IP addresses

These are IP addresses that change automatically and regularly. CSPs buy large contiguous IP address ranges and assign individual IP addresses automatically to their customers or to their customers individual internet sessions. Unallocated addresses are held in a pool to be used for other customers. The rationale for this approach is to generate cost savings for the CSP through better utilization of their available IP addresses space. There are security benefits, too, as a dynamic IP address makes it harder for targeted cyber-attacks.

Static IP addresses

Unlike dynamic IP addresses, static addresses are allocated to a specific customer and remain consistent for the lifetime of the agreement. Most individuals and businesses do not need a static IP address, but for businesses that plan to host their own web server, it is typically a requirement to have one. This is because a static IP address ensures hosted services (e.g. websites and email servers) can be reached with a consistent IP address, this is necessary if they want other devices to be able to find them on the internet.

Subscriber IP address allocation

Service Provider Subscribers will be allocated either a private IP address which is translated to a public address at the perimeter of Service Providers network or are allocated a Public IP address which is routable and addressable from the public internet, in either cases the IP address can be statically or dynamically allocated.

Website IP addresses

If a subscriber is allocated a static public IP address by its Service Provider, then it is possible that this IP address can be used to host (run) a web server which would be addressable through this IP address.

For website owners who do not host their own server and instead rely on a web-hosting provider, which is the case for most websites, there are two types of website IP addresses, shared and dedicated.

Shared IP addresses

Websites that rely on shared hosting from a web-hosting provider will typically be one of a number of websites hosted on the same server. This tends to be the case for personal and small business websites, where traffic volumes are low. Websites hosted in this way will share the same IP addresses.

Dedicated IP addresses

Some web-hosting providers offer the option to use one or more dedicated IP addresses. This can give more control over things such as SSL certificates and make the support of some services such as FTP possible. A dedicated IP address also allows the website to be reached using the IP address alone rather than the domain name.

4.2 Introduction to Network Address Translation

When observing internet activity from outside of a Communication Service Provider's private network, whether this be in real time or through transaction logs generated by web servers, the Source IP and Source Port are key values that identify the origin of the communication session.

While the combination of these identifiers will be unique for a specific data session as observed by the destination host server, due to the implementation of Network Address Translation (NAT) and Port Address Translation (PAT) at the premier edge of the CSP's network, these identifiers may not be unique to a specific connected device within the CSP's private network:

- **Private Source IP** - the private (internal) IP address assigned by the CSP to the user device for communication within their private network.
- **Public Source IP** - the public (external) IP address assigned by the CSP to the user's communications for outside of their network (i.e. the internet).
- **Destination IP** - the external (public) IP address that the user is trying to reach (often resolved by DNS) on the internet.
- **Terminology** - Source and Destination IP addresses are defined as viewed from the User Equipment (UE) point of view.

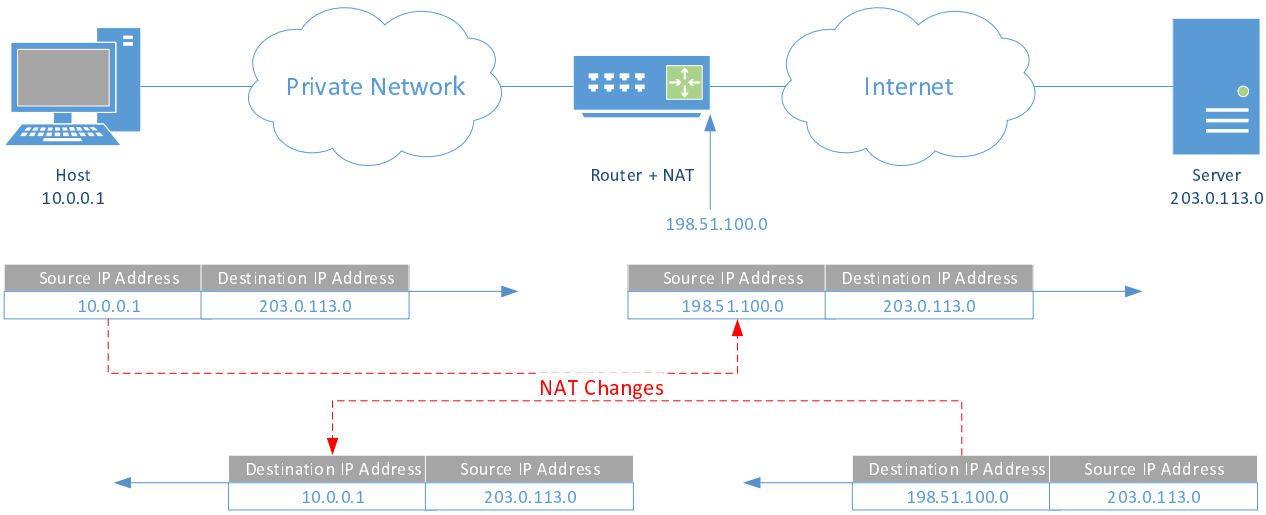

Figure 2: IP communication with Network Address Translation

For Law Enforcement Agencies (LEAs) the attribution of a permanent subscriber identifier such as a Mobile Station International Subscriber Directory Number (MSISDN) to observed internet transactions can be a critical part of an investigation and therefore the ability to trace data sessions across a CSP's address NAT or PAT capability is a necessary requirement if attribution is to be achieved.

The IP address to user attribution problem is typically largest in mobile telephony networks, where NAT is universal, although it is also a common problem for public Wi-Fi® networks. However, it is generally not a problem for fixed line or broadband services where each subscriber is typically allocated a static or long held IP address.

The present document focuses on the attribution of observed identifiers, as seen by the external destination, to permanent subscriber identifiers within the CSP's network, such as MSISDN or MAC address, for the use case where IP communication traverses a Carrier Grade NAT (CGNAT).

Why do CSPs use NAT?

The implementation of NAT within CSP network is most often through the use of CGNAT technology. CGNAT differs from standard NAT primarily through the introduction of predefined deterministic NAT and the support of large volumes of NAT transactions per second. The primary reason for CSPs to implement a CGNAT is to maximize the use of their available IP public address space as CGNAT allows the same public address to be used by multiple, potentially thousands, of subscribers simultaneously.

A secondary benefit of CGNAT is that it provides a layer of security for CSP subscribers by obscuring the source address, preventing the observed Public IP address from being tracked or attributed from outside the CSP domain.

How is CGNAT implemented?

When a device, allocated with a private address, needs to communicate with a publicly addressable entity (e.g. a website), in order for the response to be routed back to the initiating source, the private address needs to be changed for a publicly routable address. This is termed Network Address Translations (NATs) and is normally transparent to both the internal and external hosts.

CSP deployed address translation capabilities are typically implemented at the gateway between the CSP's private network and the public internet. This translation allows the CSP to define their own private address ranges, allocate IP addresses within these ranges to devices connected to the CSP's private network, and support internal traffic routing.

Similarly, to support bidirectional communication, sessions initiated from an internet host towards the UE, the sessions will pass through the CSP NAT device with the public source address being translated into a private source address for the purposes of routing.

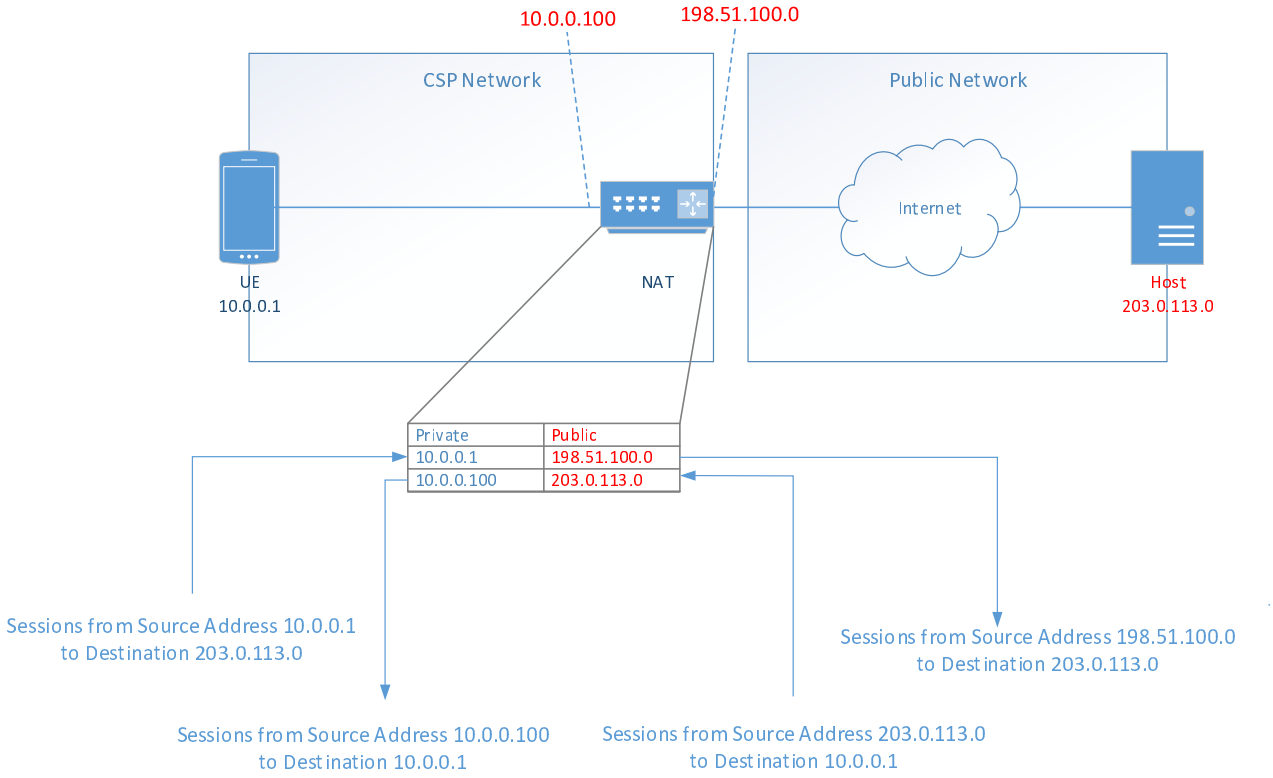

Figure 3: Bidirectional NAT

What NAT variants are there?

NAT can be implemented as:

- **Static** - Internal and external IP addresses have a 1:1 mapping; this is typically used at a CSP where a subscriber has requested or has paid for a static public IP address.

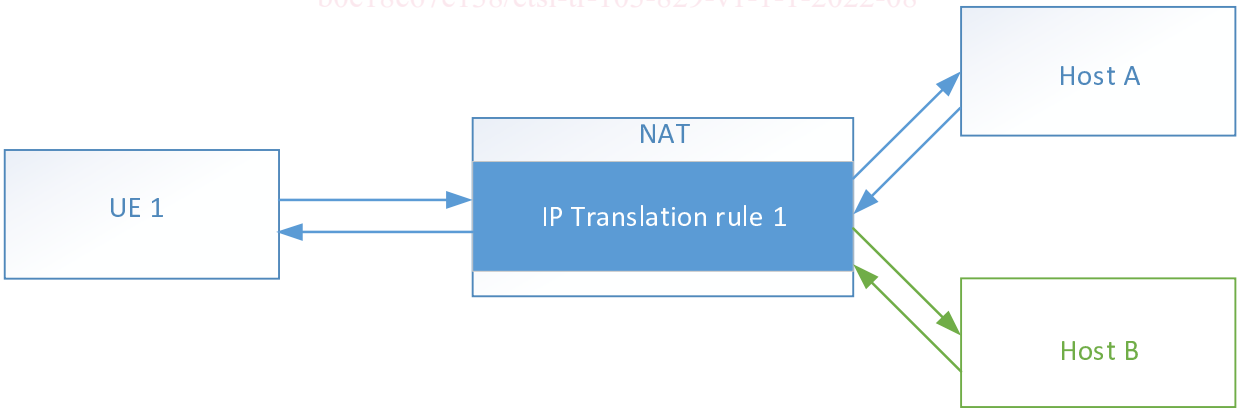

Figure 4: Static NAT

- **Dynamic** - Internal IP addresses are dynamically allocated a public IP address from an available public address pool; this is commonly used to maximize the CSP's available public address space.

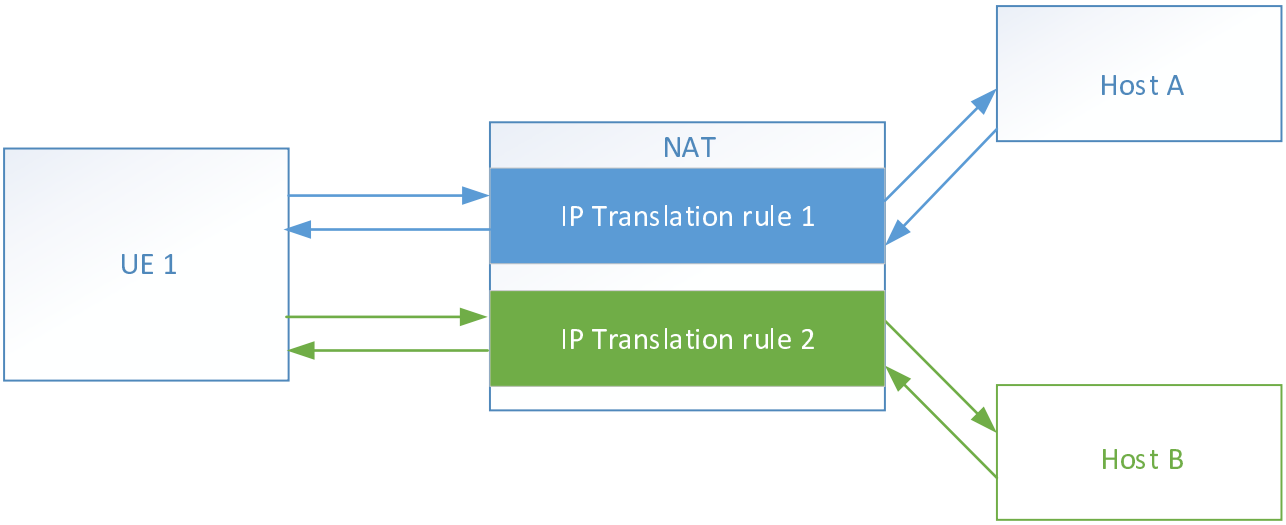

Figure 5: Dynamic NAT

- **PAT** - This is the most popular of the three types and is a variant of dynamic NAT, but maps multiple private IP addresses to a single public IP address by making use of source ports to distinguish between individual sessions.

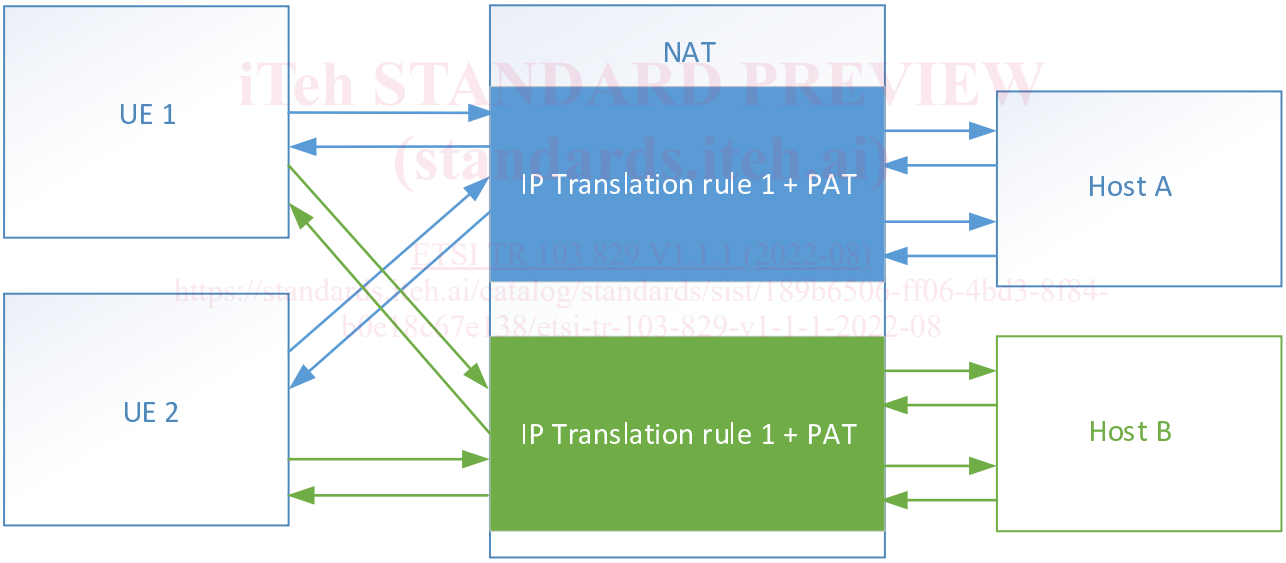

Figure 6: PAT

- **Predefined/Deterministic NAT** - In this mode of operation, private IP addresses and source ports are translated using a predefined logic which selects a public IP address and a source port from a subscriber specific range. This avoids the need for full session-based NAT log retention since the public IP address, and source port can be mapped back to a source IP address using the CGNATs deterministic logic, which may be logged or provided as part of a static configuration file.

Note that IP addresses are not assigned to users but to user devices. In some specific cases IP addresses are assigned to subscriptions (e.g. where a fixed public IP address can be provided as part of the service offering of the ISP).

Why is IP address retention and traceability difficult?

- IP addresses are often dynamically assigned and only for short periods of time.
- The private IP address to customer identifier mapping is held separately to the private/public IP address mapping, and both sessions are managed independently.
- Correlation of the records requires high accuracy and precise timestamps.