IPv4 NAT solutions and IPv6 transition technologies

KIS FRI ZU
PrS II

Roman Kaloč

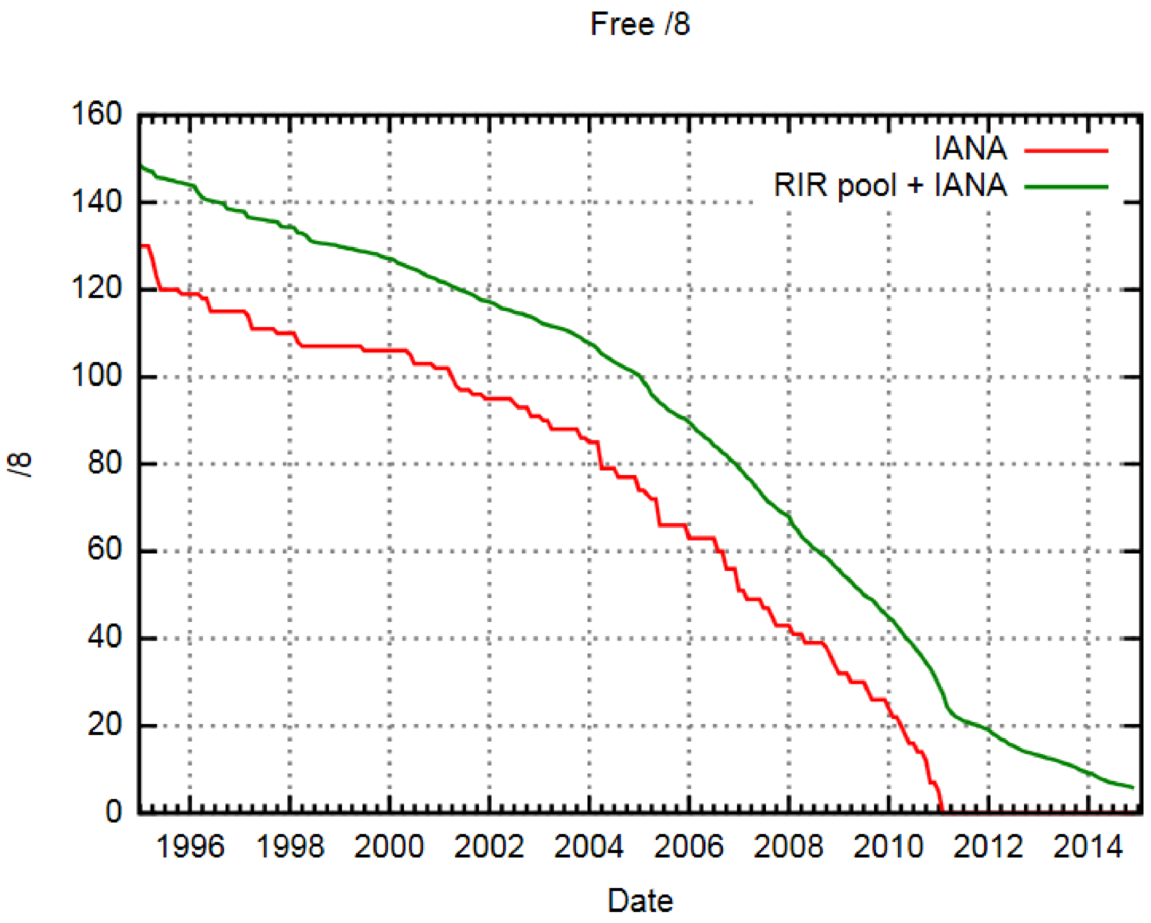
Why use IPv6?

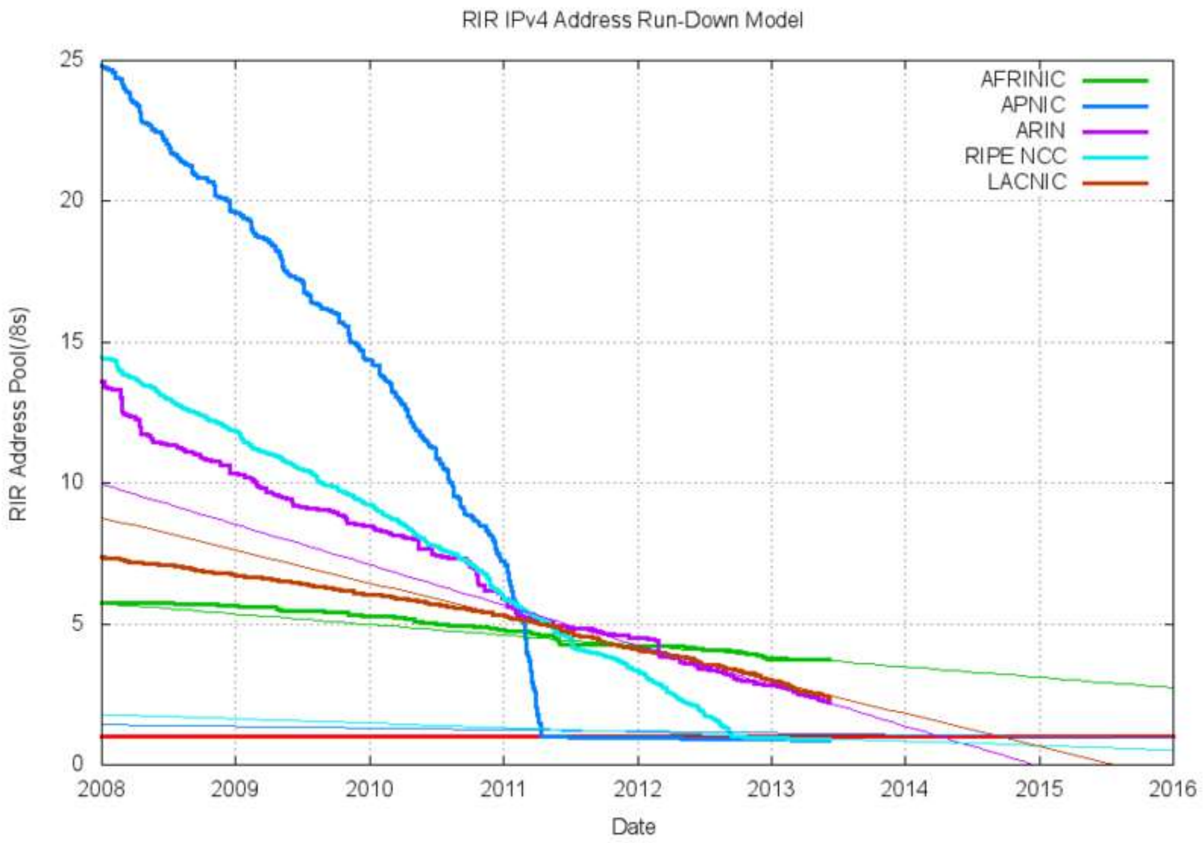
IPv4 exists since 1981, IPv6 since 1994

- IPv6 provides more address space 4,294,967,296 (2^32) for IPv4 versus 340,282,366,920,938,000,000,000,000,000,000,000,000 (2^128) – 340 sextillions
- IPv6 networks provide autoconfiguration capabilities - static or stateful DHCP versus static DHCPv6 and stateless/serverless address autoconfiguration
- E2E transparency Direct addressing is possible due to vast address space - the need for network address translation devices is effectively eliminated
- Security - IPSec is built into the IPv6 protocol
- IPv6 has improved mobility capabilities by the use of Mobile IPv6 (MIPv6)
- Better QoS – new Flow label
- Better multicast and anycast

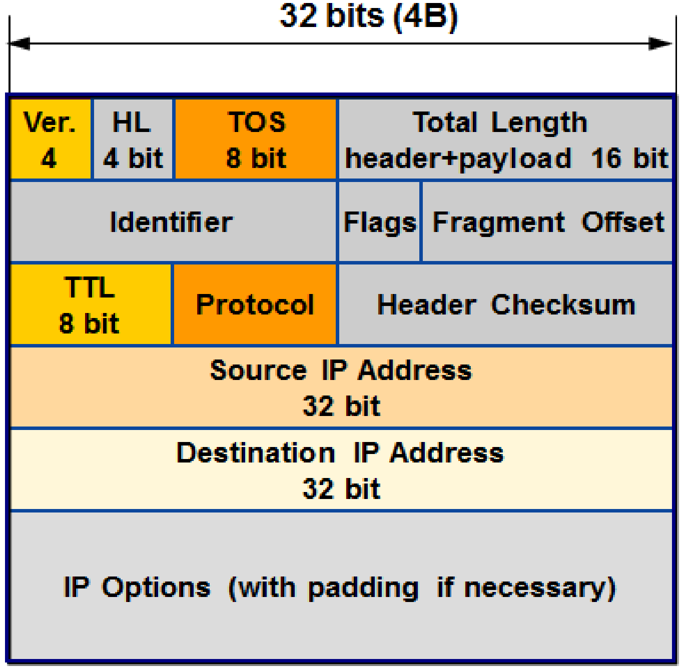
IPv4 Address Depletion

ARIN
RIPE NCC
AFRINIC
APNIC
LACNIC

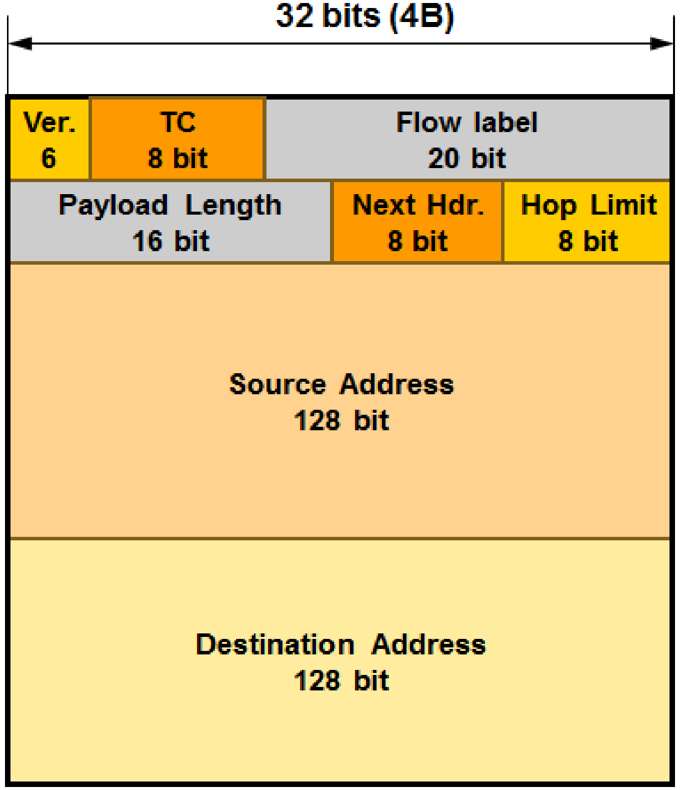

Free /8


RIR IPv4 Address Run-Down Model

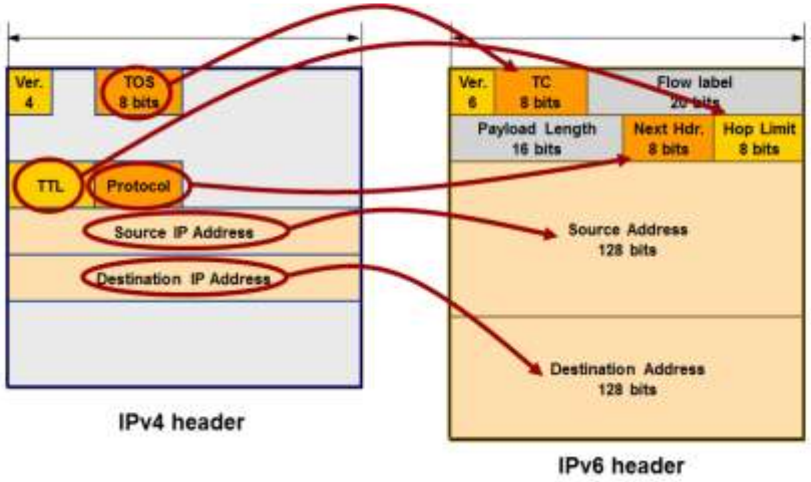
IPv4 versus IPv6 header

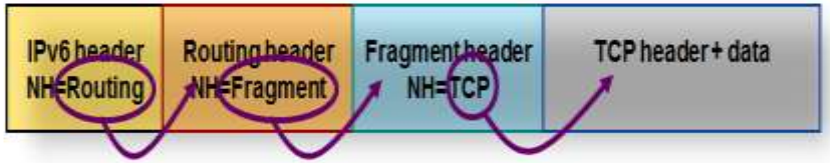


IPv4 header



IPv6 header





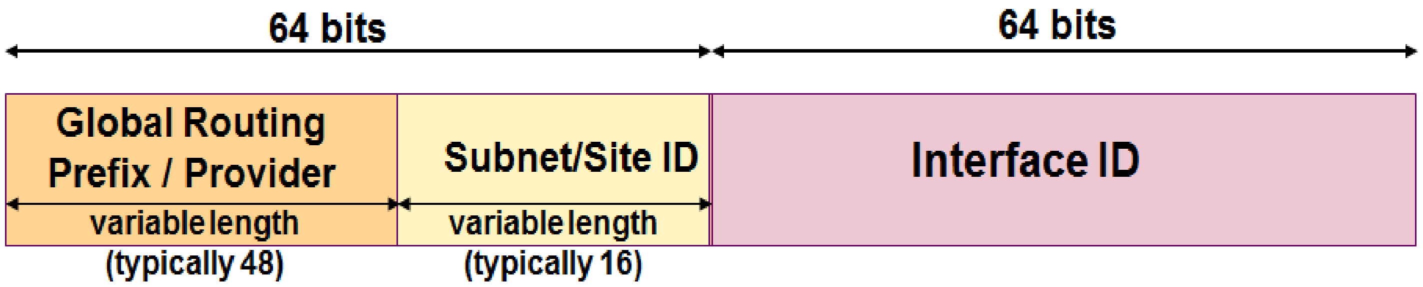
- IPv4 header length 20B, IPv6 header has a length of 40 Byte
- Header checksum is covered by L4 & L2. Fragmentation is covered by an extension header
- Extension headers are external to IPv6 header, routers do not look at these options except for Hop-by-Hop options, no negative impact on routers forwarding performance

IPv6 addressing

- IPv6 is represented as eight groups of four hexadecimal digits with colon separator - long address format is not as readable and memorizeable

2001:0DB8:0000:0000:0000:0000:1428:57ab

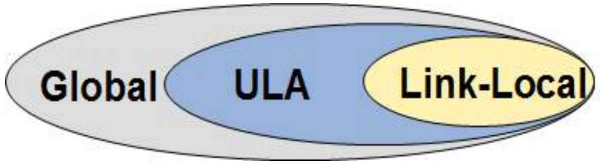
2001:0DB8::1428:57ab


64 bits
64 bits
Global Routing Prefix / Provider
variable length (typically 48)
Subnet/Site ID
variable length (typically 16)
Interface ID

- generally /48 address blocks will be given to organisations, large operators might get shorter prefixes or several /48 prefixes – most common is a /32 block (which is in fact the address space of the IPv4 address range for the subnet IP)

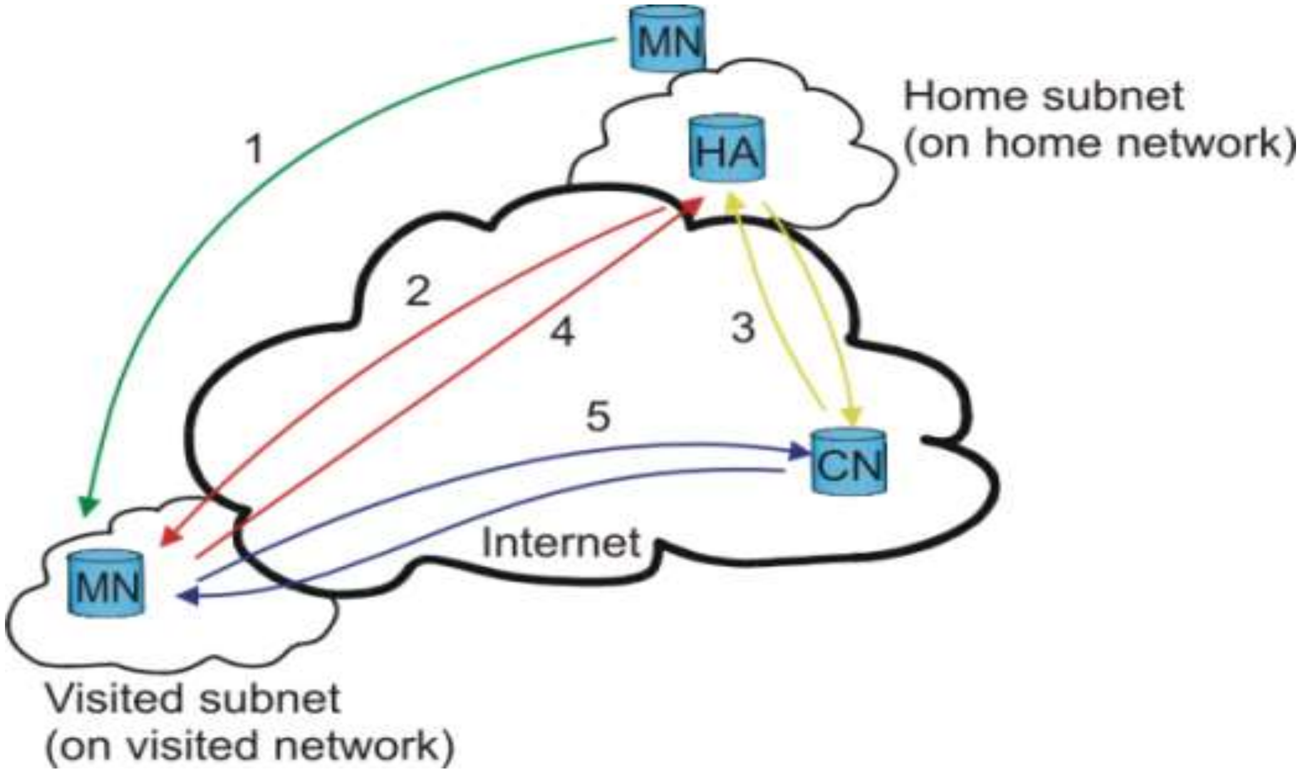
Address block	Usage	RFC
2000::/3	Global unicast address (the first 3 bits are 001)	Currently assigned by IANA
2002::/16	6to4	RFC3056
::1/128	Loopback IP (similar to 127.0.0.1 in IPv4)	
FC00::/7	Unique Local Addresses	RFC4193
FE80::/10	Link-local addresses	
0:0:0:0:FFFF::/96	IPv4 address mapped in IPv6, the last 32 bits are the IPv4 address	
FF00::/8	IPv6 Multicast	

- Link local, Unique local (ULA), Global addresses



Mobile IPv6 principle

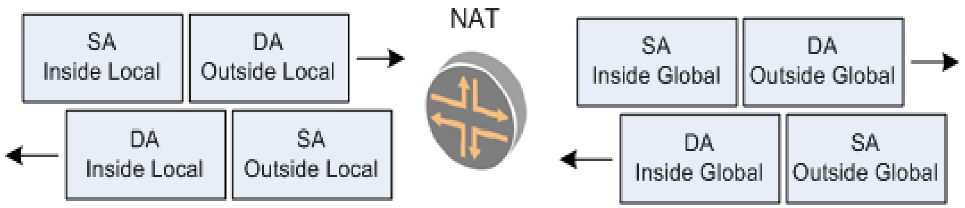
- The Mobile Node (MN) travels to a foreign network and gets a new care-of-address, it does a link-local address configuration to the closest remote router
- The MN performs a binding update to its Home Agent (HA) (the new Care-of-Address gets registered at HA). HA sends a binding acknowledgement to MN.

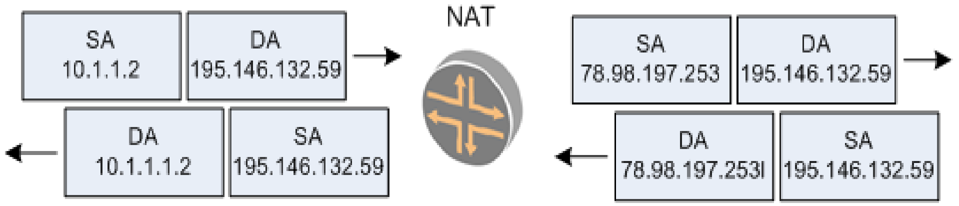


- A Correspondent Node (CN) wants to contact the MN. The HA intercepts packets destined to the MN.
- The HA then tunnels all packets to the MN from the CN using MN's care-of-address.
- When the MN answers the CN, it may use its current Care-of-Address (and perform a binding to the CN) and communicate with the CN directly (optimized routing) or it can tunnel all its packets through the HA.

Basic NAT (Network Address Translation) principles

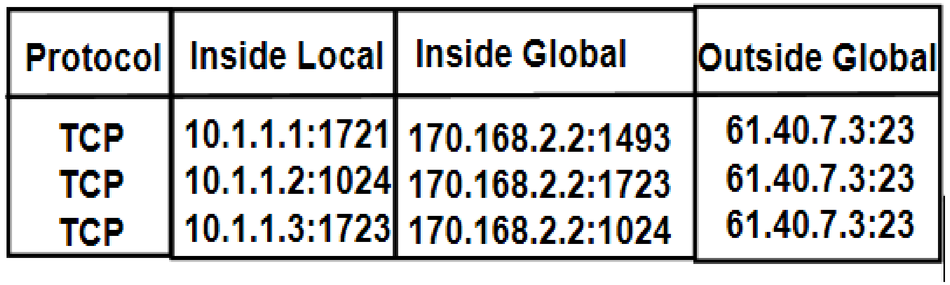
- Static NAT – one-to-one mapping between local and global addresses
- Dynamic NAT –many-to-many mapping between local and global addresses by using a pool of real IP addresses
- Overloading (NAPT, PAT) – many-to-one N:1 local to global mapping using a real IP's transport layer ports but comes with limitations
  - TCP, UDP checksum recalculation
  - ICMP manipulation
  - IPSec Transport mode - Authentication Header (AH) integrity
  - Performance limitation
  - Port number limitation
  - Session initiation from outside





Naming

- Inside Local – inside source address before translation
- Outside Local – destination host before translation
- Inside Global – inside host after translation
- Outside Global –outside destination host after translation



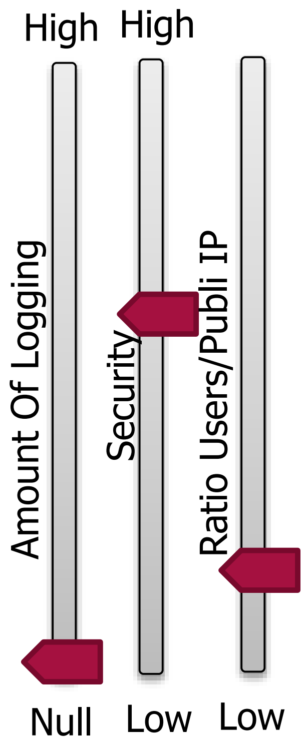
Protocol	Inside Local	Inside Global	Outside Global
TCP	10.1.1.1:1721	170.168.2.2:1493	61.40.7.3:23
TCP	10.1.1.2:1024	170.168.2.2:1723	61.40.7.3:23
TCP	10.1.1.3:1723	170.168.2.2:1024	61.40.7.3:23

NAT – Application Level Gateway (ALG)

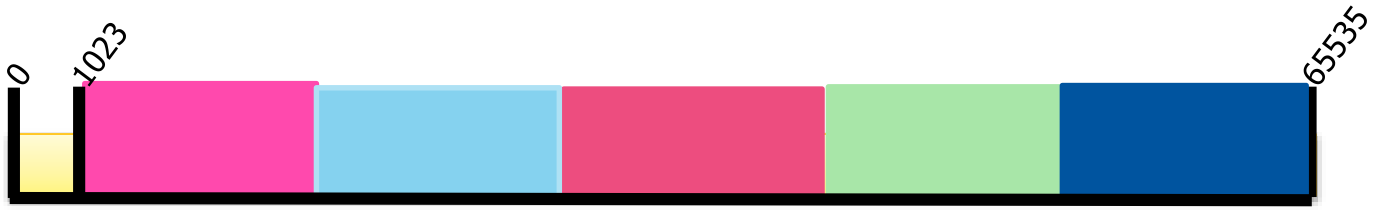
- Several applications include IP addresses in the messages (ASCII or binary formats) and port numbers
- Application Level Gateways (ALG) add some functionalities to NATs for a correct operation with such applications operation with such applications
- Based on the application and messages type, not only IP headers but also message contents are translated, and if needed, TCP segments are modified accordingly
- ALGs are similar to proxy, but they are transparent to hosts
- Rich set of ALG’s for NAT44(4):
  - BOOTP, DNS, FTP, H323, ICMP, IP, PPTP, RTSP, SIP, SNMP, SQLNET, TFTP, Traceroute, Unix Remote Shell Service, WINFrame, etc…

Carrier Grade NAT Specifics & Selected Features

Deterministic NAT



- Algorithmic Allocation of IP addresses and Port bucket per user. The Public IPv4 address and Port range for a given end user are fixed. Once the port range is determined, the allocation of a given port for a new flow is performed dynamically.
- Evaluation:
  - Users keep the same public address all the time (useful for Port forwarding but could be a privacy issue)
  - No log messages needed at all.
  - Lower ratio of Users/Public address (Even Inactive user get a public@)
  - Can’t assign new block if users run out of ports.
  - Load balancing more complex (Need FBF to stick the private subnets to specific MS-DPC).
  - The block of port size is only possible by step of 2^n



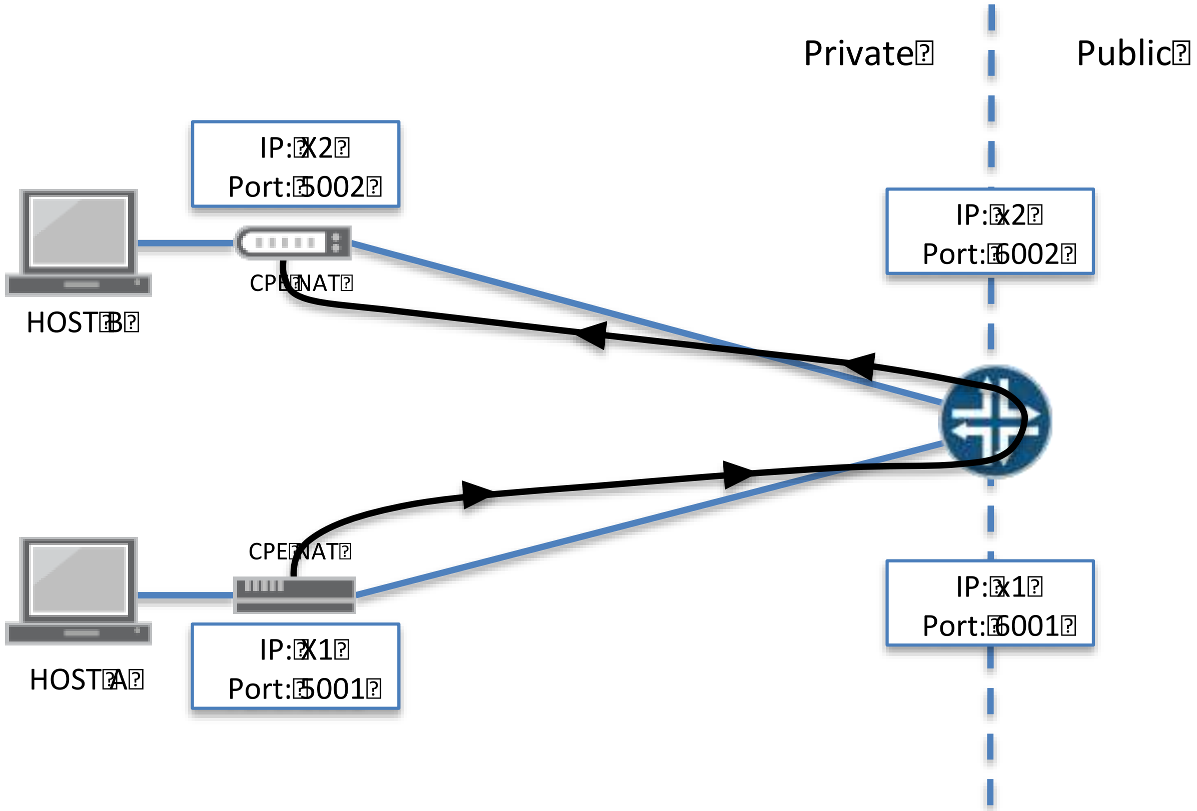
Public address – Ports allocation (one user per color)

Address pooling, EIM, EIF, quotas

- Certain websites such as online banking require that all connections from a given host (SSL or not) come from the same IP address and even the same source port.
  - Address Pooling
    - Address pooling means assigning the same external address for all sessions originating from the same internal host, does not say anything about port, solves the problems of an application opening multiple connections.
  - EIM (Endpoint Independent Mapping)
    - EIM means assigning the same external source address and port for all connections to any dest IP from a given host if they use the same internal source port. This means if they come from a different source port, you are free to assign a different external address. Important for p2p, gaming and the mobile world.
  - Endpoint Independent Filtering
    - Enabling EIM means that you have a stable external IP address + port (for a period of time) that external hosts can use to connect. Note: The determination of who can connect to an internal host is done by End Point Independent Filtering (EIF)
  - Quotas
    - User quotas enable the administrator to restrict the maximum number of simultaneous NAT sessions each subscriber is allowed to use, thereby maintaining a fair distribution of the resources across the entire subscriber

Hairpinning

- Hairpinning is the act of anchoring a session between two hosts inside a CGN environment on the NAT Router



NAT Bindings:

X1:5001 → x1:6001
X2:5002 → x2:6002

Flows, Sessions and CGN Deployment

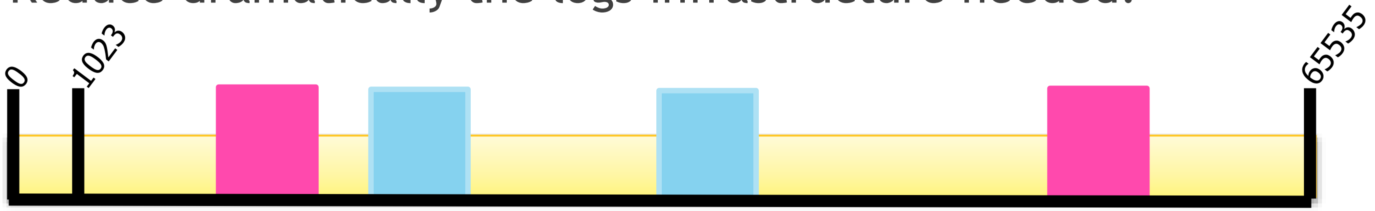
- Flow is a unidirectional 5-tuple (inside source IP, inside source port, (outside source IP, outside source port) dest IP, dest Port, protocol)
- A session is two unidirectional flows
- Flow Analysis is a major part of CGNAT design, the average number of flows per subscriber, duration and break down per protocol and application are key information for storage dimensioning for logging

Session based logging:

```
2011-05-05 23:04:16{in}[FWNAT]: ASP_SFW_CREATE_ACCEPT_FLOW  2.1.1.2:30 [3.3.3.3:1024] ->  192.3.1.2:30  (UDP)
2011-05-05 23:04:22{in}[FWNAT]: ASP_SFW_DELETE_FLOW  2.1.1.2:30 [3.3.3.3:1024] ->  192.3.1.2:30  (UDP)
```

NAT with port bucket allocation (PBA)

- When a session is created, the NAT allocate a contiguous bucket of Ports per user. The port will then be randomly chosen from this bucket.
- New requests for NAT ports will come from this block. Any non-active block (without any ports in use) will get freed from the NAT pool.
- Logs are only generated for each block allocation and release.
- Evaluation:
  - Possible to tune the ratio Logging/Security/Users-per-IP (see next slide)
  - Reduce dramatically the logs infrastructure needed.



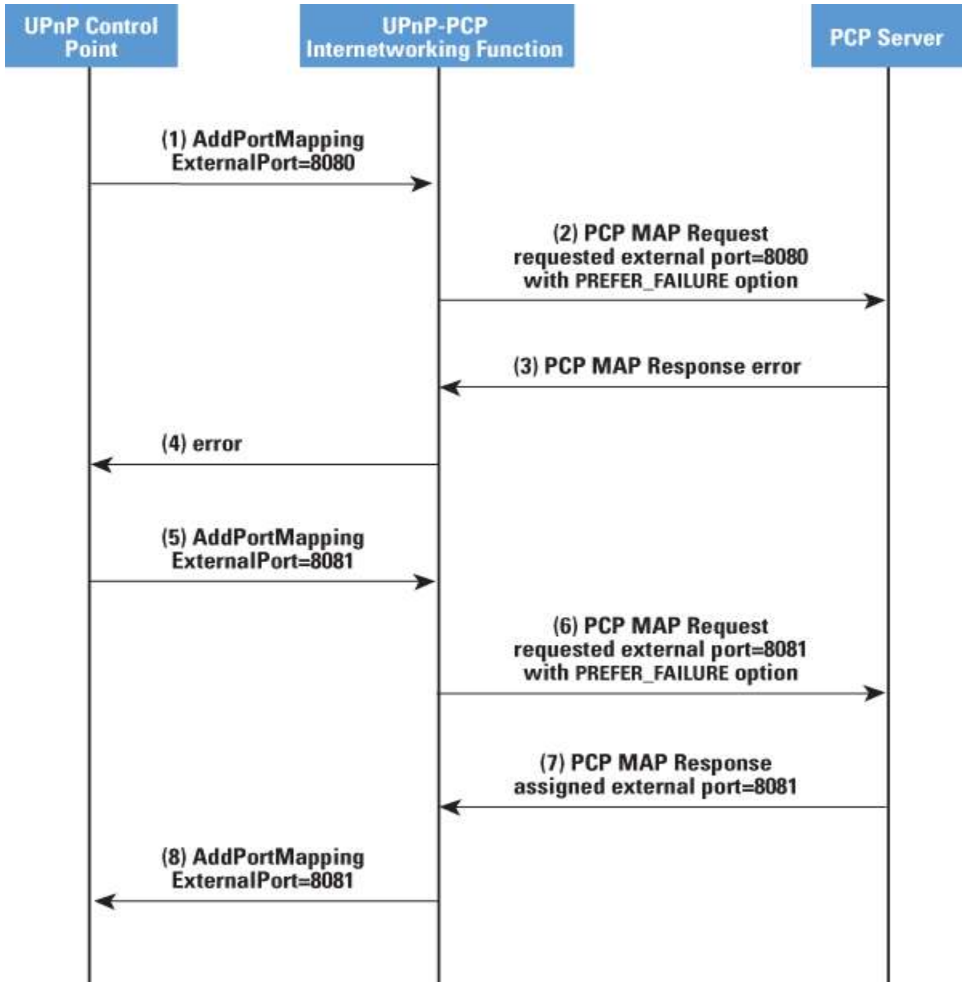
Public address – Ports allocation (one user per color)

PBA based logging:

```
Nov 16 01:23:59  sustain-re0 (FPC Slot 2, PIC Slot 0) 2011-11-16 09:23:59 NAT444[FWNAT]:ASP_NAT_PORT_BLOCK_ALLOC: 2.1.1.2 ->
3.3.3.3:1024-1055
Nov 16 01:23:59  sustain-re0 (FPC Slot 2, PIC Slot 0) 2011-11-16 09:23:59 NAT444[FWNAT]:ASP_NAT_PORT_BLOCK_RELEASE: 2.1.1.2 ->
3.3.3.3:1024-1055
```

Port Forwarding Challenges and PCP

- PCP (Port Control Protocol) objectives are to enable applications to receive incoming connections in the presence of an ISP NAT/Firewall
- Mappings between an external IP address, protocol and port, and an internal IP address, protocol and port
- Additionally, explicit port forwarding rules available through PCP
- IP & port (service) to DNS mapping issue


UPnP Control Point
UPnP-PCP Internetworking Function
PCP Server
(1) AddPortMapping ExternalPort=8080
(2) PCP MAP Request requested external port=8080 with PREFER_FAILURE option
(3) PCP MAP Response error
(4) error
(5) AddPortMapping ExternalPort=8081
(6) PCP MAP Request requested external port=8081 with PREFER_FAILURE option
(7) PCP MAP Response assigned external port=8081
(8) AddPortMapping ExternalPort=8081

IPv4 NAT solutions and IPv6 transition

NAT 44 / NAT 444

- More efficient usage of IPv4 resources
- Application Layer Gateway (ALG) for IP address bound applications
- Short term solution, no IPv6 deployed

Pros

+ No need to change the current CPE spec

+ All consists of existing technologies

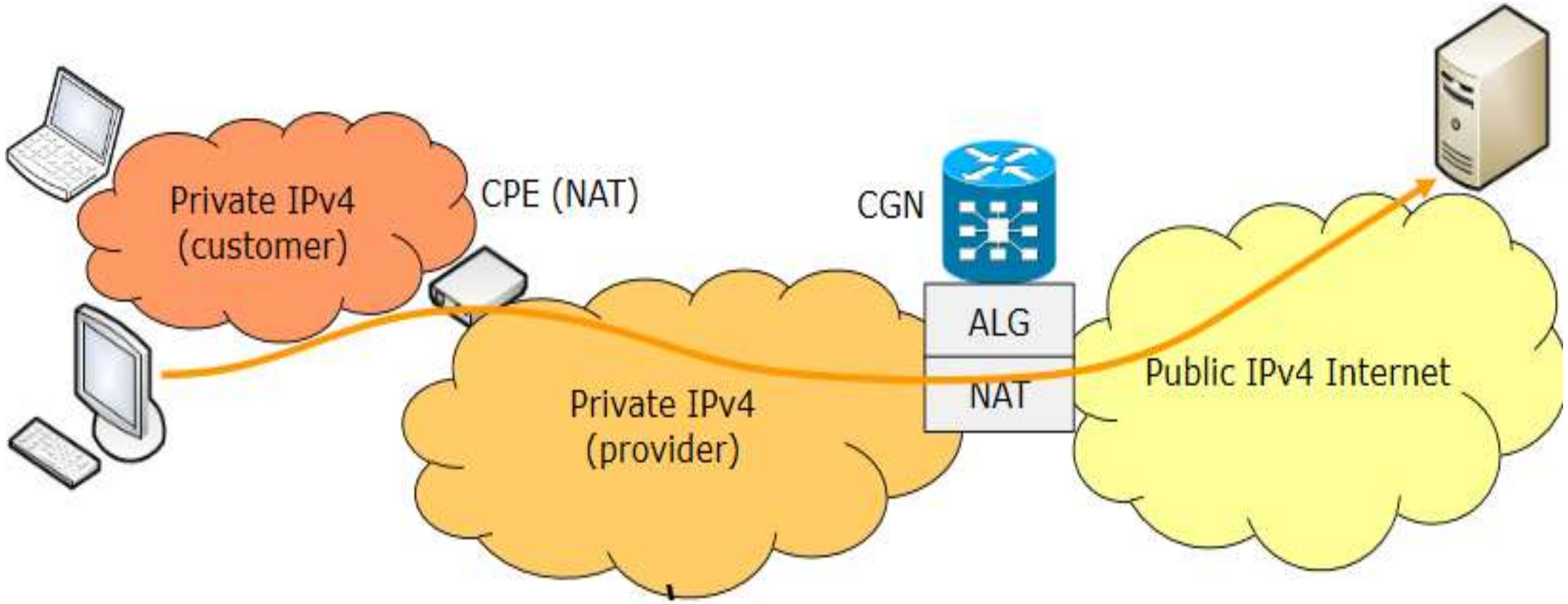
+ Easier to implement

Cons

- Scalability & security concerns

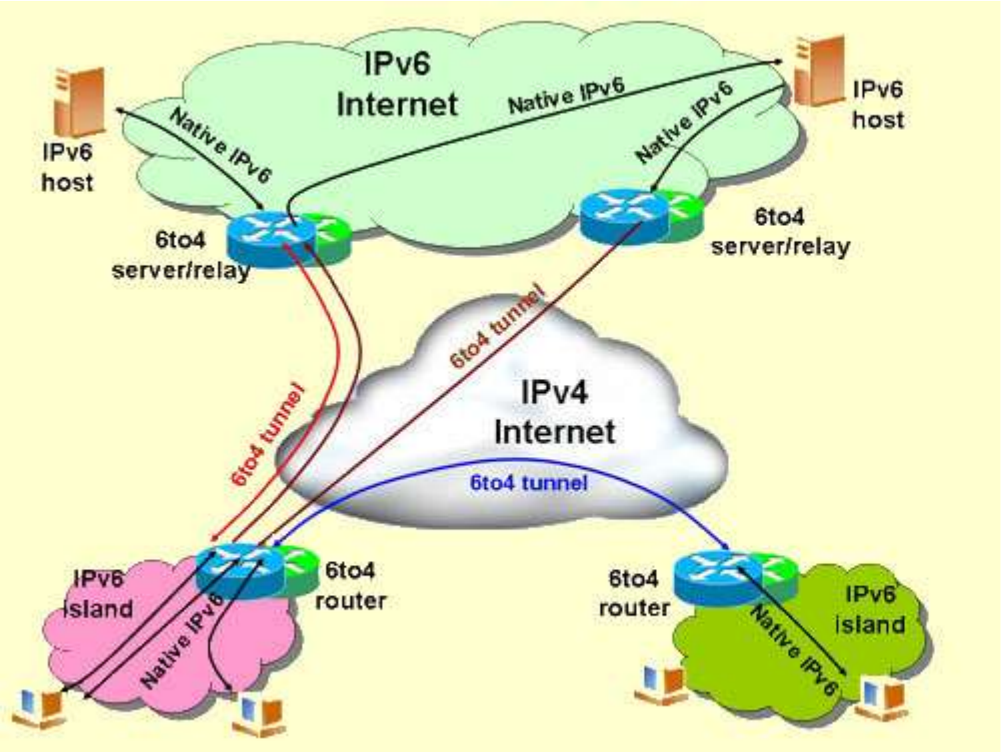
- Two layers of NAT break some applications

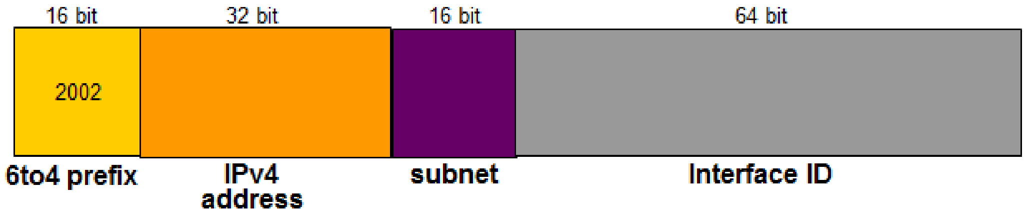
- Difficult to keep records for law-enforcement



6to4 principle

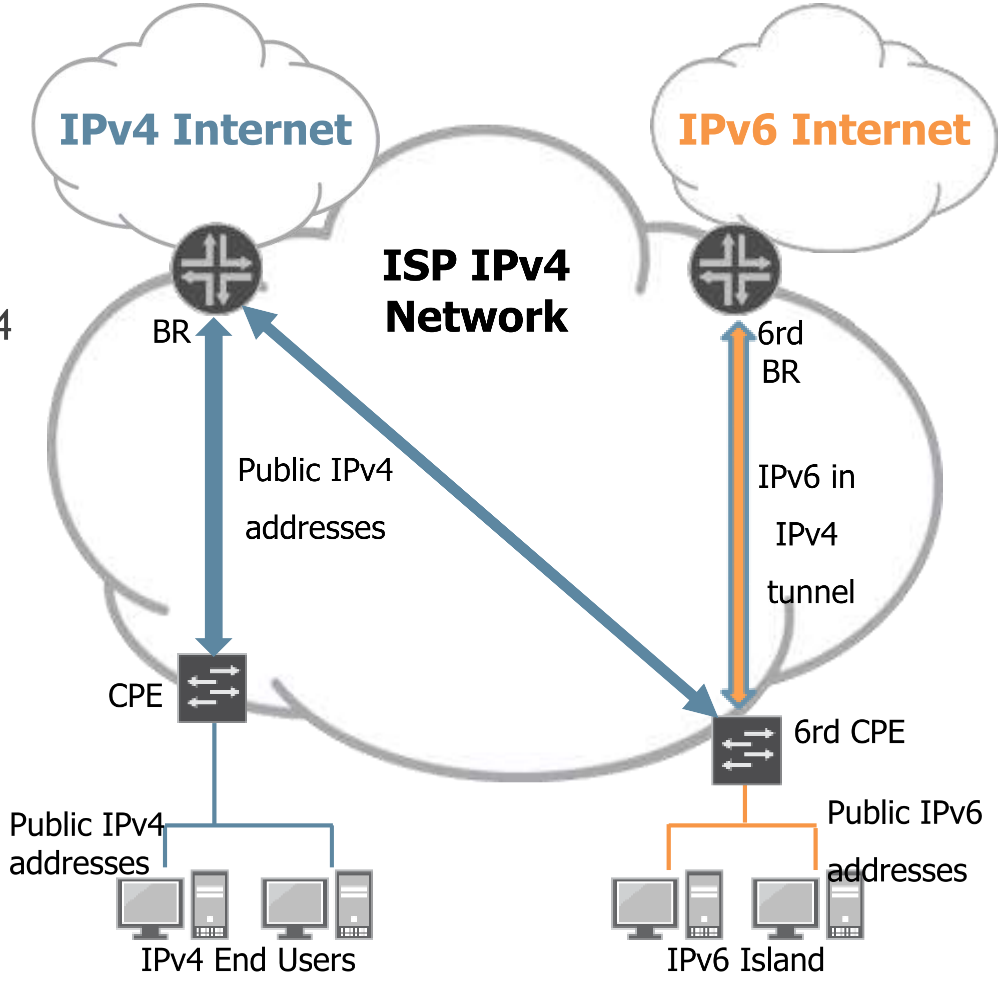
- Allows IPv6 hosts/sites to communicate over IPv4 networks, no need to configure explicit tunnels
- IPv4 unicast addresses are still required
- Uses 6to4 prefix <2002::16>
- 6to4 *does not* facilitate interoperation between IPv4-only hosts and IPv6-only hosts





IPv6 Rapid Deployment (6rd)

- IPv6 tunnelling over IPv4 network from CPE to Border Relay
  - Enhancement of 6to4
  - Standard 6to4 prefix can be replaced by an ISP assigned address space
- Can be considered as an initial transition phase technology: IPv6 islands are connected over IPv4 networks
- As 6rd is stateless, BRs may be reached using anycast for failover and resiliency
- Still requires an IPv4 address, which can be:
  - public → it does not help to solve the IPv4 depletion problem
  - private → helps to solve the IPv4 depletion but requires NAT 444 to connect to the public IPv4 internet



Cons
- Requires CPE change with IPv6 in IPv4 tunnel encapsulation
- Proximity to 6rd relay/AFTR affects geo-location
- Subject to IPv4 address exhaustion for end users

Pros
+ Support IPv6 connectivity
+ Keep existing access, aggregation, and BNG on IPv4

Dual Stack (or Dual Stack with NAT 44/444)

Dual host client must be able to request and read DNS v4 (A record) and DNSv6 (AAAA record)

Dual stack is considered as the easiest and best transition technology as all IPv4 applications will be supported

Dual stack applications must be accessed over IPv6 to save IPv4 addresses

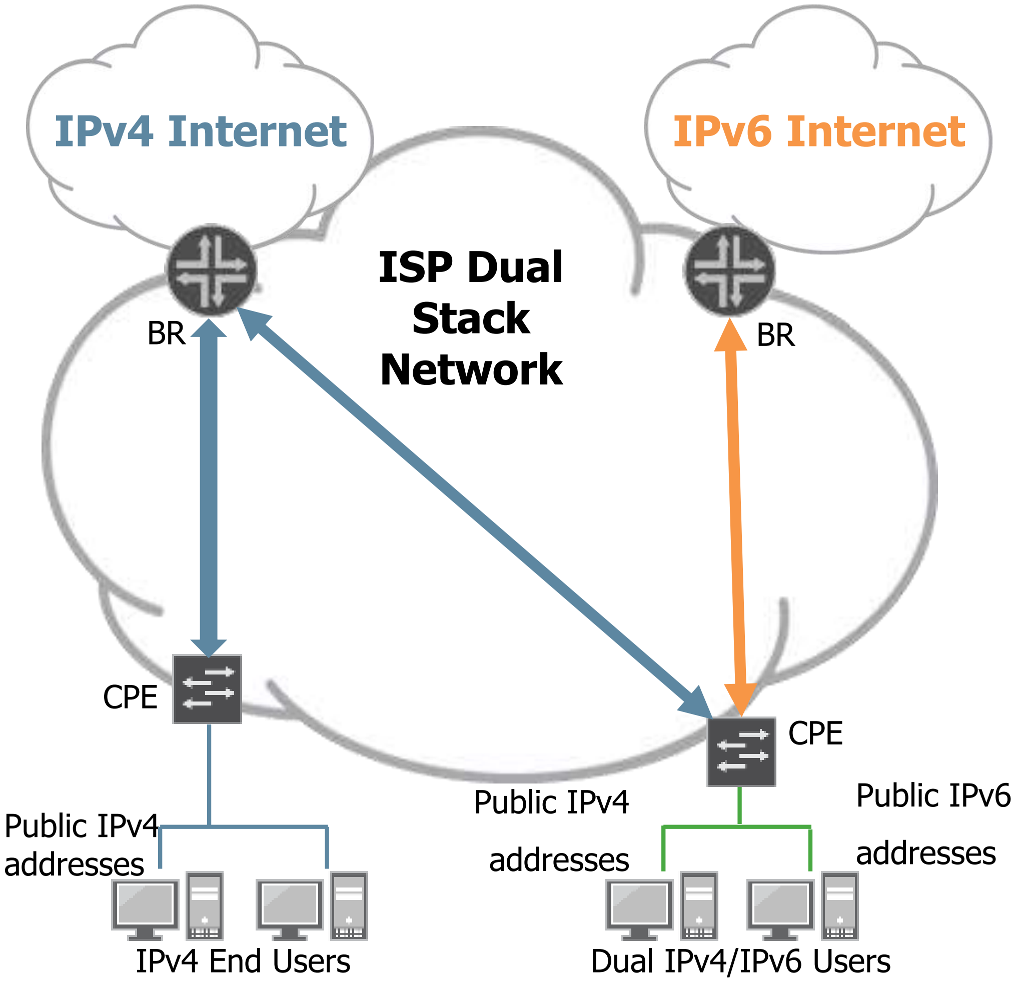
IPv4 and IPv6 protocols must be maintained in the whole chain

Pros

+ No need to tunnel
+ No NAT in network

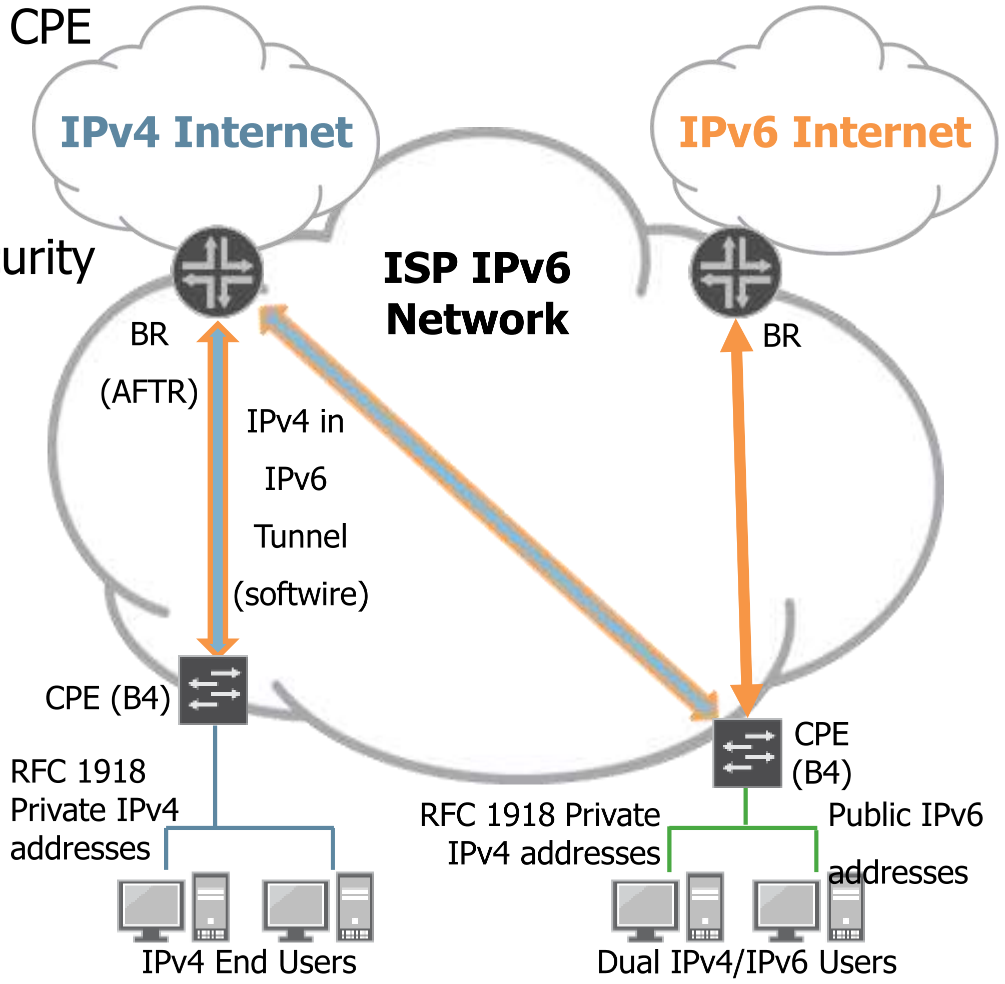
Cons

- Subject to IPv4 address exhaustion for end users
- New CPE investment
- Infrastructure investments for DS



Dual Stack Lite

- P-to-MP tunnelling between tunnel concentrator (AFTR - Address Family Translation Router) and Home router (B4 - Basic Bridging BroadBand)
- Uses IPv4 in IPv6 tunnelling, AFTR discovery mechanism for CPE to AFTR connection establishment
- One IPv4 address for M customers
- port learning rate per user, port limit per user and other security implementation need to be considered



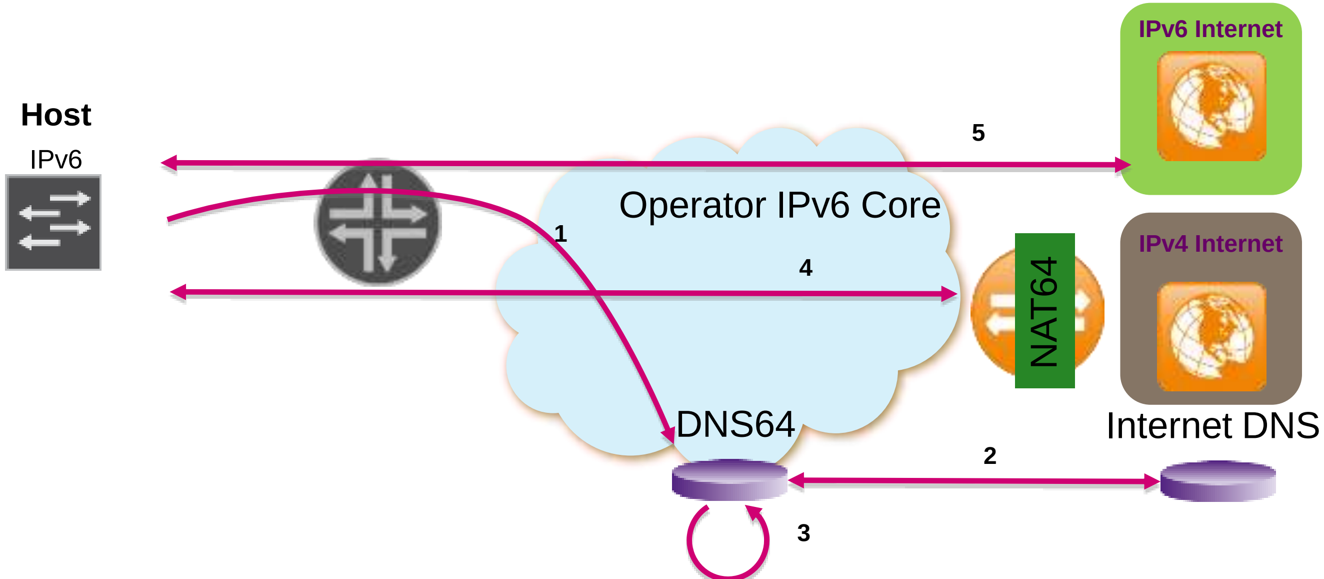
Cons
- Requires CPE change with IPv4 in IPv6 tunnel encapsulation
- Same CGN concerns as with NAT444 because NAT44 in AFTR

Pros
+ Access Network could be IPv6-only
+ Intelligent IPv6/IPv4 bindings

NAT64 / DNS64

- Mechanism to provide connectivity of an IPv6 only user to an IPv4 only application/service
- Requires introduction of NAT64 and DNS 64
- Translation between IPv4 and IPv6 in the NAT64 GW
- Principle:
    1. DNS64 receives a AAAA DNS query from Host
    2. DNS64 attempts resolution
    3. If no AAAA (DNS6) is available DNS64 performs request for A (DNS4) record and synthesizes it into AAAA
    4. The IPv6 representation of the IPv4 is algorithmically generated from the IPv4 address (pref64::/n embedding the IPv4 address)
    5. If AAAA is available Host is directly connected to the IPv6 internet

Well known prefix 64:ff9b::/96



Thank you