BSc in Telecommunications Engineering

TEL3214

Computer Communication Networks

Lecture 06
The Internet Protocol – Next Generation

Eng Diarmuid O'Briain, CEng, CISSP

Department of Electrical and Computer Engineering,
College of Engineering, Design, Art and Technology,
Makerere University

Table of Contents

1. Introduction to Internet Protocol version 6 (IPv6, IPng)....5
1.1 Features of IPv6....6
1.2 IPv6 Address Scope....7
2. IPv6 Address Architecture....8
2.1 The Zero Suppression rule....8
2.2 The Zero Compression rule....8
2.3 IPv6 Address Notation....9
2.4 Special IPv6 addresses....10
2.5 IPv6 Packet Structure....10
2.6 IPv6 Option headers....12
3. IPv6 Prefix Terminology....13
3.1 Link-Local Address (LLA)....14
4. IPv6 Multicast address....16
4.1 Flags....16
4.2 Scope....16
4.3 GroupID....17
4.4 Solicited-Node Multicast Group Address....18
4.5 Special Prefix's....18
5. Applications for IPv6....20
5.1 DHCP for IPv6 (DHCPv6)....20
5.2 DNS Extensions to Support IP Version 6 (DNSv6)....20
5.3 ICMPv6 for IPv6....20
6. IPv6 Stateless Address Auto-configuration (SLAAC)....22
6.1 MLD joins multicast group....23
6.2 Neighbour Solicitation (135)....24
6.3 MLD joins multicast group for the second time....24
6.4 Router Solicitation (133)....25
6.5 Router Advertisement (134)....26
6.6 Neighbour Solicitation (135) for the second time....28
6.7 MLD joins multicast group for the third time....29
7. IPv6 Address Resolution and redirection....30
7.1 Neighbour Unreachability Detection (NUD)....30
7.2 ICMPv6 Redirect....31
8. IPv6 Configuration best practice – Inter-router links....32
8.1 Using LLA on Inter-router links....32
9. IPv6 Address planning....33
10. IPv6 Interior Gateway Routing....34
10.2 IPv6 Exterior Gateway Routing....34
11. IPv6 transition mechanisms....35
11.1 Dual Stack....35
11.2 Proxying and translation....35
11.3 Tunnelling....38

Illustration Index

Illustration 1: IPv6 network notation....9
Illustration 2: IPv6 packet structure....10
Illustration 3: Forming an EUI-64 MAC from an EUI-48 MAC....14
Illustration 4: IPv6 Multicast identifier....16
Illustration 5: Solicited-Node Multicast Group Address formation....18
Illustration 6: MLD joins multicast group....23
Illustration 7: Neighbour Solicitation (135)....24
Illustration 8: MLD joins multicast group (2)....24
Illustration 9: Router Solicitation (133)....25
Illustration 10: Router Advertisement (134)....26
Illustration 11: RA flags....27
Illustration 12: RA - Prefix flags....27
Illustration 13: Neighbour Solicitation (135) (2)....28
Illustration 14: MLD joins multicast group (3)....29
Illustration 15: Address Resolution....30
Illustration 16: Neighbour Unreachability Detection (NUD)....30
Illustration 17: ICMPv6 Redirect....31
Illustration 18: Inter-router link....32
Illustration 19: Inter-router link with LLA....32
Illustration 20: NAT-PT....36
Illustration 21: NAT64 / DNS64....37
Illustration 22: Overlay tunnels for IPv6....38
Illustration 23: Tunnel Broker (TB)....39

1. Introduction to Internet Protocol version 6 (IPv6, IPng)

IPv6 also called IPng is the replacement for IPv4. It has 3.4 x 10^{38} addresses (2^{128}) more than 7.9 x 10^{28} times as many as IPv4. This updated version of IP was invented by Steve Deering and Craig Mudge at Xerox PARC, it was then adopted by the Internet Engineering Task Force in 1994 as IPng.

The adoption of IPv6 has been slowed by the introduction of Network Address Translation (NAT), which partially alleviates IPv4 address exhaustion. Japan and Korea had started with their implementation of IPv6 in the late 1990's. The European Union (EU) formed an IPv6 Task Force as a steering committee in 2001 and member states all had their own IPv6 Task Forces by 2004. The United States of America (US) has specified that the network backbones of all federal agencies must have deployed IPv6 by 2008.

In October 2007 Vint Cerf the founder of the Internet issued a warning to ISP urgently need to roll out IPv6 because the IPv4 pool is finite and has all but run out in 2012 (The RIPE NCC body started allocating its last */8* in September 2012). Each Local Internet Registry (LIR) can receive one final /22 allocation (1,024 IPv4 addresses) upon application for IPv4 resources. No new IPv4 Provider Independent (PI) space will be assigned.

By 2017 all the RIRs has exhausted their allocations with AFRINIC, the African RIR being the last to declare on the 3 April 2017 [1].

It is expected that IPv4 will be supported alongside IPv6 for the foreseeable future with hosts running dual-stack software.

1 http://www.afrinic.net/en/library/news/2053-afrinic-enters-ipv4-exhaustion-phase-1

1.1 Features of IPv6

IPv6 supports many new features over IPv4, these features were developed considering the problems that were showing in IPv4.

- **Much larger address space** - StateLess Address Auto-configuration (SLAAC)
  - IPv6 hosts can be configured automatically when connected to a routed IPv6 network. When first connected to a network, a host sends a link-local multicast request for its configuration parameters; if configured suitably, routers respond to such a request with a router advertisement packet that contains network-layer configuration parameters.
- **Multicast**
  - Multicast (both on the local link and across routers) is part of the base protocol suite in IPv6. This is different to IPv4, where multicast is optional.
  - IPv6 does not have a link-local broadcast facility; the same effect can be achieved by multicasting to the all-hosts group with a hop count of one.
- **Jumbograms**
  - In IPv4, packets are limited to 64KB of payload. When used between capable communication partners, IPv6 has support for packets over this limit, referred to as jumbograms. Use of jumbograms improves performance over high throughput networks.
- **Faster routing**
  - By using a simpler and more systematic header structure, IPv6 improves the performance of routing. Recent advances in router technology, however, may have made this improvement irrevalent.
- **Network-layer security**
  - IPsec, the protocol for IP network-layer encryption and authentication, is an integral part of the base protocol suite in IPv6. It is, however, not yet deployed widely except for securing BGP traffic between IPv6 routers.
- **Mobility**
  - IPv6 was designed to support mobility. IPv6 Neighbour Discovery (ND) and SLAAC allow hosts to operate in any locations without any special support. This makes it more scalable and the performance is better because less traffic passes through the home link and less redirection and less rerouting. It also means no single point of failure.

1.2 IPv6 Address Scope

IPv6 address have a *scope* to specify where the address is valid. Within unicast addressing, Link-local Addresses (LLA) and the loopback address have ***link-local*** scope, which means they are to be used in the directly attached network (link) only. All other addresses, including Unique Local Addresses (ULA) and Global Unicast Addresses (GUA), have global (or universal) scope, which means they are globally routable, and can be used to connect to addresses with global scope anywhere, or addresses with *link-local* scope on the directly attached network. The scope of an *anycast* address is defined identically to that of a *unicast* address.

For multicasting, the four least-significant bits of the second address octet of a multicast address (ff0X::) define the address scope, the span over which the multicast address is propagated.

2. IPv6 Address Architecture

IPv6 addresses are normally written as 32 nibbles within 8 groups of 4 hexadecimal digits. For example, *2a02:2158:435a:0000:83:314:ea21:b33f* is a valid IPv6 address.

128 bits

⇣

32 nibbles

⇣

8 x 4 nibble groups

2.1 The Zero Suppression rule

Leading zeros in a group can be omitted. Thus

2a02:0201:0000:0000:0000:0000:00a1:b33f is shortened to *2a02:201:0:0:0:0:a1:b33f*.

2.2 The Zero Compression rule

Contiguous groups of '0' can be replaced with '::' as long as there is only one double colon used in an address.

2a02:201:0:0:0:0:0a1:b33f maybe shortened to 2a02:201::a1:b33f .

Having more than one double-colon abbreviation in an address is invalid as it would make the notation ambiguous. Leading zeros in a group can be omitted. Thus

2a02:0000:0000:0000:0022:0000:0000:b33f may be shortened as follows:

2a02:0:0:0:22:0:0:b33f (Zero suppression rule)

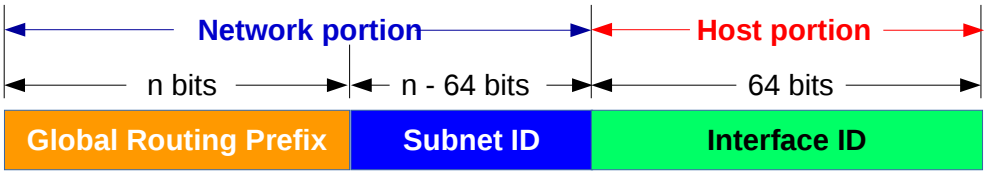
2a02::22:0:0:b33f (Zero compression rule)

Following the rules above, confirm if the addresses below are all valid and equivalent:

2a02:2158:0000:0000:0000:0000:00a1:b33f
2a02:2158:0000:0000:0000::00a1:b33f
2a02:2158:0:0:0:0:0a1:b33f
2a02:2158:0::0:0a1:b33f
2a02:2158::0a1:b33f
2a02:2158::a1:b33f

A sequence of 4 bytes at the end of an IPv6 address can also be written in decimal, using dots as separators. This notation is often used with compatibility addresses. Thus, *::ffff:1.2.3.4* is the same address as *::ffff:102:304*.

2.3 IPv6 Address Notation



2a02:2158:435a:330::9bc2:45/64

Illustration 1: IPv6 network notation

IPv6 networks are written using Classless Inter-Domain Routing (CIDR) notation. An IPv6 network is a contiguous group of IPv6 addresses the size of which must be a power of two; the initial bits of addresses which are identical for all hosts in the network are called the network's prefix.

A network is denoted by the first address in the network and the size in bits of the prefix, separated with a slash. For example, *2a02:2158:435a:330::/64* stands for the network with

- First address: *2a02:2158:435a:330::*
- Last address: *2a02:2158:435a:330:ffff:ffff:ffff:ffff*

Because a single host can be seen as a network with a 128-bit prefix, a host address may be shown with */128* mask.

Like IPv4 the IPv6 Address is constructed of two parts the Prefix + host Identifier (ID) (Sometimes the Interface ID). The idea is to separate *who u are* from *where u are connected to*. The Prefix is dependant on the routing topology and the Interface ID identifies a node. IPv6 removes the Broadcast address and instead uses special Multicast addresses *all hosts ff0X::1* or *all routers ff0X::2* where X is replaced by the scope number. IPv6 also introduces a new *anycast* address. An *anycast* address is an IPv6 address that is assigned to one or more network interfaces, with the property that a packet sent to an *anycast* address is routed to the *nearest* interface having that address, according to the routing protocols measure of distance.

- *Unicast*: from one host to another.
- *Multicast*: from one to all belonging to a group.
- *Anycast*: from one to the nearest belonging to a group.

2.4 Special IPv6 addresses

2.4.1 Unspecified Address

Such an address is similar to IPv4, IPv6 has a special address reserved for loopback.

- ::
- ::/128

2.4.2 Default route Address

This special address format is for the default route. Similiar to 0.0.0.0/0 in IPv4.

- ::/0

2.4.3 Loopback Address

Similar to the loopback function in IPv4, IPv6 has a special address reserved for loopback.

- 0:0:0:0:0:0:0:1
- ::1
- ::1/128

2.5 IPv6 Packet Structure

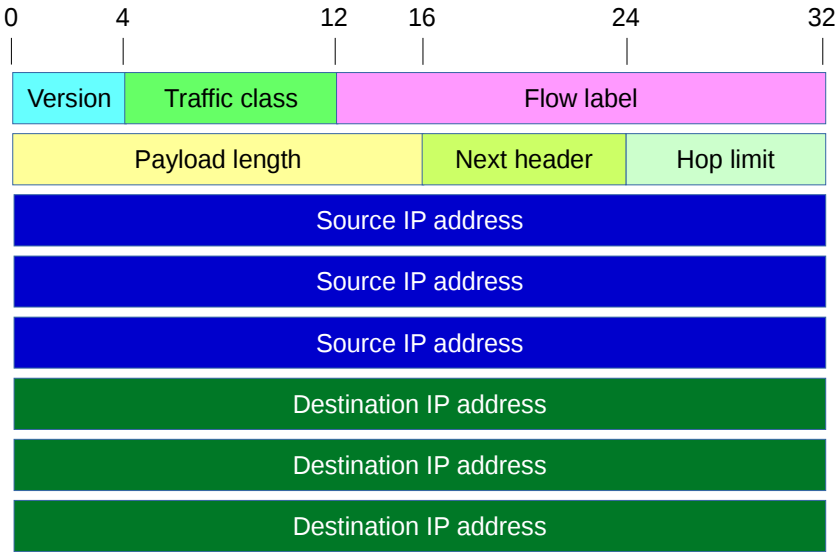


Illustration 2: IPv6 packet structure

The IPv6 packet header has many changes compared to the IPv4 header while maintaining necessary elements. Unlike the IPv4 packet header the IPv6 header is a fixed 40 bytes in size and adds extension headers to add additional information. Fragmentation is IPv6 only occurs at the source, intermeditary routers will not fragment like IPv4 and if a router cannot pass the packet it drops it. This reduces processing by intermediatary routers.

The IPv6 header contains:

Header	Bytes	Description
Version	4	Describes the version as 6
Traffic Class	8	One byte field
Flow Label	20	20 bit flow label for label tagging
Payload Length	16	Two byte integer giving the length of the packet less the base header but inclusing the extension headers
Next Header	8	Specifies IPv6 extension headers or a upper layer protocol
Hop Limit	8	Single byte decremented at each router, packet discarded if zero
Source Address	128	Address of originator
DestinationAddress	128	Address of the destination

Here is an example IPv6 packet which has an IPv6 Hop by Hop extension header followed by Internet Control Message Protocol v6 (ICMPv6) as the upper layer protocol. This will become clearer in the next section.

```
Frame: 90 bytes on wire (720 bits)
    Encapsulation type: Ethernet (1)
Ethernet II
    Destination: IPv6mcast_16 (33:33:00:00:00:16)
        .... ..1. .... .... .... .... = LG bit: Locally administered address
        .... ...1 .... .... .... .... = IG bit: Group address (multicast/broadcast)
    Source: 00:00:00_aa:00:02
        .... ..0. .... .... .... .... = LG bit: Globally unique address (factory default)
        .... ...0 .... .... .... .... = IG bit: Individual address (unicast)
    Type: IPv6 (0x86dd)
Internet Protocol Version 6, Src: ::, Dst: ff02::16
    0110 .... = Version: 6
        .... 0000 0000 .... .... .... .... .... = Traffic class: 0x00 (DSCP: CS0, ECN: Not-ECT)
        .... 0000 00.. .... .... .... .... .... = Differentiated Services Codepoint: Default (0)
        .... .... ..00 .... .... .... .... .... = Explicit Congestion Notification:
                                                   Not ECN-Capable Transport (0)
        .... .... .... 0000 0000 0000 0000 0000 = Flowlabel: 0x00000000
    Payload length: 36
    Next header: IPv6 Hop-by-Hop Option (0)
    Hop limit: 1
    Source: ::
    Destination: ff02::16
    Hop-by-Hop Options
        Next Header: ICMPv6 (58)
        Length: 0 (8 bytes)
        IPv6 Option (Router Alert)
            Type: Router Alert (5)
            Length: 2
            Router Alert: MLD (0)
        IPv6 Option (PadN)
            Type: PadN (1)
            Length: 0
            PadN: <MISSING>
Internet Control Message Protocol v6
    Type: Multicast Listener Report Message v2 (143)
    Code: 0
    Checksum: 0x6ede [correct]
    Reserved: 0000
    Number of Multicast Address Records: 1
    Multicast Address Record Changed to exclude: ff02::1:ffaa:2
        Record Type: Changed to exclude (4)
        Aux Data Len: 0
        Number of Sources: 0
        Multicast Address: ff02::1:ffaa:2
```

2.6 IPv6 Option headers

Unlike IPv4 the IPv6 options are handled outside the IPv6 header. This is achieved by the addition of extensions headers which are only processed as necessary. For example only routers process the *Hop by Hop options header*. With this method it is easier to define new extensions and options as the protocol evolves. Here is a list of some optional headers that are used with IPv6 today. They always appear in this order within packets if they are being added.

Header	Code	Description
Hop by Hop options	0	
Destination options	60	Examined only by destination node
Routing	43	Specify the route for a datagram
Fragment	44	Fragmentation parameters
Authentication header (AH)	51	Verify packet authenticity
Encapsulation security payload (ESP)	50	Encrypted data
Destination options	60	Examined only by destination node
Mobility (Mobile IPv6)	135	Parameters for use with mobile IPv6

3. IPv6 Prefix Terminology

IPv6 does not have a *classful* concept like IPv4 but within GUA assignments have a number of prefixes, with different prefix lengths. Here is a table outlining four of the key terms.

Prefix Term	Assigned by	Example prefix
Registry Prefix	Assigned to Regional Registry (RR)	2a02::/12
ISP Prefix	Assigned to Internet Service Provider (ISP)	2a02:2158::/32
Site Prefix	Assigned to Large Organisation	2a02:2158:1111::/48
Site Prefix	Assigned to Smaller Organisation	2a02:2158:1111:100::/56
Subnet Prefix	Internal subnet within Organisation	2a02:2158:1111:110::/64
A host address	Organisation/Residential home user	2a02:2158:1111:110::10/128

The following table give an indication of IPv6 Relative Network Sizes.

Mask	Size	Description
128	1 IPv6 Address	A network interface
64	1 IPv6 subnet	18,446,744,073,709,551,616 IPv6 addresses
56	256 LAN segments	Popular prefix size for smaller subscriber site
48	65,536 LAN segments	Popular prefix size for larger subscriber site
32	65,536 /48 subscriber sites	Minimum IPv6 allocation by RR
24	16,777,216 subscriber sites	256 times larger than the min IPv6 allocation

3.1 Link-Local Address (LLA)

The address block fe80::/10 has been reserved for link-local unicast addressing. To conform to standard /64 addressing on subnets, the actual LLA are assigned with the prefix fe80::/64. The 54 bits after the most significant ten bits must be zero.

IPv6 requires an LLA on every network interface on which the IPv6 protocol is enabled, even when routable addresses are also assigned. Therefore IPv6 hosts usually have more than one IPv6 address assigned to each of their IPv6-enabled network interfaces. The LLA is required for the Neighbour Discovery Protocol (NDP) and DHCPv6.

LLA addresses can be assigned automatically by a process called StateLess Address AutoConfiguration (SLAAC) using NDP or manually.

3.1.1 Forming an LLA

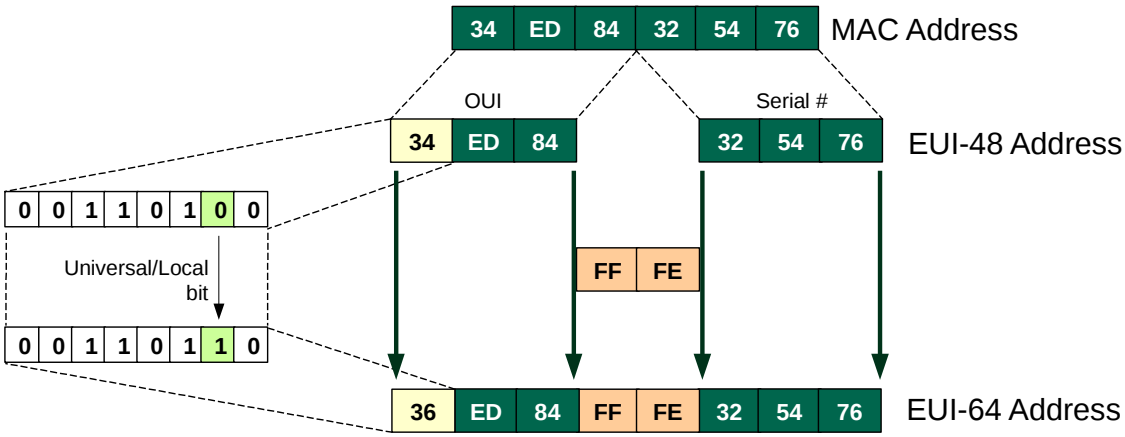


Illustration 3: Forming an EUI-64 MAC from an EUI-48 MAC

Nodes, both hosts and routers generate an LLA for each interface on boot. Such an LLA is formed by appending the interface's identifier (Interface ID) to the well-known Link-local prefix *fe80::*.

The first step is to generate an EUI-64 MAC address. IPv6 uses 64 bits for the network and subnets while it reserves the last 64 bits to identify the host. Traditionally MAC addresses are of the EUI-48 type with 48 bits. However the range of unique EUI-48 MAC addresses are running out and it was decided to migrate to EUI-64 format in the future. IPv6 was built for EUI-64 addresses.

To convert process of an EUI-48 to EUI-64 refer to Illustration 3 where the original address is split and FF:FE inserted. The 6th bit, called the Universal/local bit in the first octet is changed to a '1' to indicate the new MAC is not unique.

Universal/local bit 0 = Unique MAC
1 = non-Unique MAC

3.1.2 Resolving LLA ambiguity with zone IDs.

As LLA fe80::/10 address will exist on all interfaces it and therefore each interface is part of the same network it is necessary to let the device know which interface to use when communicating with a neighbouring node. Basically the node has no way to determine the interface to send packet out on.

Here is an example where a ping to an IPv6 LLA address fails on a router.

```
n1# ping ipv6 fe80::200:ff:feaa:0
connect: Invalid argument
```

Appending the **%<interfaceID>** to the LLA resolves this ambiguity on the router.

```
n1# ping ipv6 fe80::200:ff:feaa:0%eth1
PING fe80::200:ff:feaa:0%eth1(fe80::200:ff:feaa:0) 56 data bytes
64 bytes from fe80::200:ff:feaa:0: icmp_seq=1 ttl=64 time=0.025 ms
64 bytes from fe80::200:ff:feaa:0: icmp_seq=2 ttl=64 time=0.036 ms
```

Another example, this time from a host.

```
$ ssh fe80::287b:8236:8feb:df65%wlp4s0
alovelace@fe80::287b:8236:8feb:df65%wlp4s0's password: babbage
Welcome to Linux Mint 18.1 Serena (GNU/Linux 4.4.0-53-generic x86_64)

 * Documentation:  https://www.linuxmint.com
Last login: Tue Jan  3 23:04:33 2017
alovelace@remote ~ $
```

3.1.3 Reserved Interface IDs (RFC 5453)

Interface Identifier Range	Description
0000:0000:0000:0000	Subnet router anycast
FDFF:FFFF:FFFF:FF80 - FDFF:FFFF:FFFF:FFFF	Reserved Subnet anycast

4. IPv6 Multicast address

An IPv6 multicast address is an identifier for a group of interfaces that are typically on different nodes. An interface may belong to any number of multicast groups. Multicast addresses have the following format:

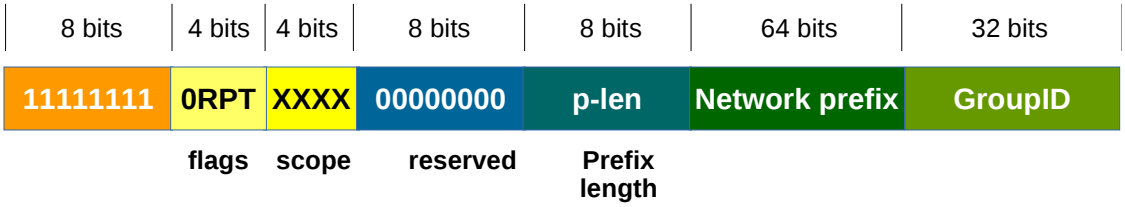


Illustration 4: IPv6 Multicast identifier

4.1 Flags

The multicast identifier containf a 4 bit flag field.

Nibble value	flag	Meaning when 0	Meaning when 1
8	reserved	reserved	reserved
9	R (Rendezvous)	Rendezvous point not embedded	Rendezvous point embedded
10	P (Prefix)	Without prefix information	Address based on network prefix
11	T (Transient)	Well-known multicast address	Dynamically assigned multicast address

4.2 Scope

Multicast scope is a 4-bit value used to limit the scope of the multicast group.

ff0X::/8	Meaning
0x1	Interface local
0x2	Link local
0x4	Admin local
0x5	Site local
0x8	Organisation local
0xE	Global
0x0	Reserved
0xF	Reserved

4.3 GroupID

The group ID identifies the multicast group, either permanent or transient, within the given scope.

ID	Meaning
1	All nodes on the local network segment
2	All routers on the local network segment
5	OSPFv3 All SPF routers
6	OSPFv3 All DR routers
8	IS-IS for IPv6 routers
9	Routing Internet Protocol (RIP) routers
a	Enhanced Interior Gateway Routing Protocol (EIGRP) routers
d	Protocol Independent Multicast (PIM) routers
16	MLDv2 reports (defined in RFC 3810)
1:2	All Dynamic Host Configuration Protocol (DHCP) servers and relay agents on the local network segment (defined in RFC 3315)
1:ff	Solicited-Node Multicast Address (SNMA)
fb	Multicast DNS
101	Network Time Protocol (NTP)
108	Network Information Service (NIS)
181	Precision Time Protocol (PTP)
114	Used for experiments

Examples:

- ff05::1 All nodes on the local site
- ff02::2 All routers on the link local
- ff02::5 All OSPF routers on the link local
- ff02::a All EIGRP routers on the link local
- ff05::101 All Network Time Protocol (NTP) Servers on the local site
- ff02::1:3 All DHCPv6 servers on the link local

4.3.1 Multicast MAC

A corresponding multicast MAC is asociated with each ,ulticast address. For example:

- ff02::1 → 33:33:00:00:00:01
- ff02::2 → 33:33:00:00:00:02
- ff02::5 → 33:33:00:00:00:05
- ff02::6 → 33:33:00:00:00:06

4.4 Solicited-Node Multicast Group Address

Every device that uses an IPv6 address will also compute and join a Solicited Node Multicast Group Address (SNMA). This address is required for the IPv6 NDP.

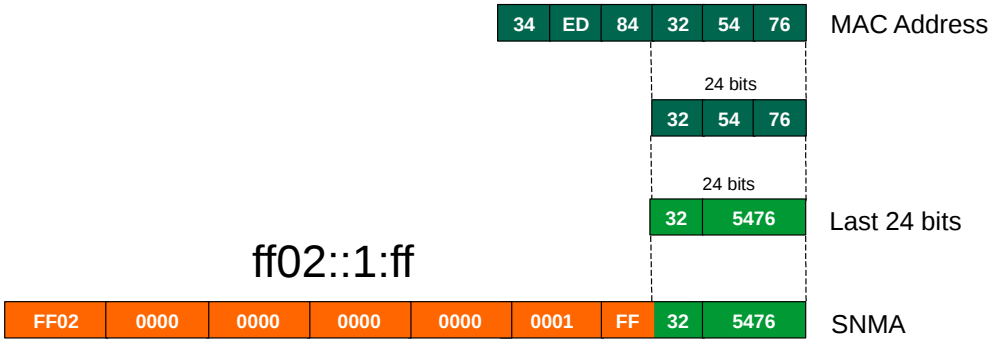


Illustration 5: Solicited-Node Multicast Group Address formation

The process of generating such an address is simple. The last 24-bits of the interface MAC address is aquired. These 24 bits are prepended with the special SNMA network ff02::1:ff/104. The example demonstrates the EUI-48: *34ed:8432:5476* MAC address being used to generate an SNMA: *ff02::1::ff32:5476*. A corresponding MAC address for the multicast IPv6 address is also created. In this example that would be *33:33:ff:32:54:76.*

4.5 Special Prefix's

There are a number of specific addresses within IPv6 with special meaning:

Prefix	Meaning
::/0	The default unicast route address (similar to 0.0.0.0/0 in IPv4)
::/128	The address with all zeroes is an unspecified address, and is only to be used in software
::1/128	The loopback address is a localhost address. (like 127.0.0.1 in IPv4)
::ffff:0:0/96	This prefix is used for IPv4 mapped addresses. Transparent use of Transport Layer protocols over IPv4 through IPv6 API
64:ff9b::/96	Well known prefix for 6to4 address translation.
0400::/7	Internetwork Packet Exchange (IPX) from the IPX/SPX protocol stack routed via IPv6
2000::/3	**Global Unicast Address (GUA):** 2000:: - 3fff::
fc00::/7	**Unique Local Address (ULA)**: are only routable within a set of cooperating sites.
fe80::/10	**Link-local Address (LLA)**: Prefix specifies that the address only is valid in the local physical link. (like the Auto-configuration address 169.254.x.x in IPv4)
ff00::/8	The multicast prefix for multicast addresses
ff01::0/12	Pre-defined Multicast addresses
ff01::1/12	All host addresses (interface-local)
ff01::2/12	All routers (interface-local)
ff02::1/12	All host addresses (link-local)
ff02::2/12	All routers (link-local)
ff05::2/12	All routers (site-local)

4.5.1 Depreciated Prefix's

The following prefixes were originally defined as part of IPv6 but have since been depreciated or obsoleted. I have added them here for information in case you come across such addresses.

Prefix	Meaning
::/96	The zero prefix was used for IPv4-compatible addresses. Depreciated in February 2006.
fec0::/10	Site-local prefix specifies that the address is only valid inside the local organisation. Its use has been deprecated in September 2004 by IPv6 Deprecating Site Local Addresses RFC and future systems must not implement any support for this special type of address any more.
0200::/7	Network Service Access Point (NSAP) addresses from ISO/IEC 8348 routed via IPv6. Deprecated in December 2004.

5. Applications for IPv6

5.1 DHCP for IPv6 (DHCPv6)

DHCP for IPv6 (DHCPv6). Although IPv6's StateLess Address Auto Configuration (SLAAC) removes the primary motivation for DHCP in IPv4, DHCPv6 can still be used to statefully assign addresses if the network administrator desires more control over addressing. It can also be used to distribute information which is not otherwise discoverable; the most important case of this is the DNS server.

A major difference with DHCPv4 Servers is that hosts send broadcasts to find DHCP Servers whereas with DHCPv6 Servers IPv6 hosts send IPv6 multicast. The reserved address for hosts to send packets to an unknown DHCPv6 Server is *FF02::1:2*.

5.2 DNS Extensions to Support IP Version 6 (DNSv6)

DNS is similar for IPv4 and IPv6 (DNSv6). The main difference is that the *A* record is replaced by the *AAAA* record which maps a hostname to a 128-bit IPv6 address for forward lookups. Reverse lookups take place under ip6.arpa, where address space is delegated on nibble boundaries. This scheme is a straightforward adaptation of the familiar A record and in-addr.arpa schemes for IPv4.

5.3 ICMPv6 for IPv6

ICMP version 6 (ICMPv6) is a new version of ICMP and is an integral part of the IPv6 architecture that must be completely supported by all IPv6 implementations and nodes. ICMPv6 combines functions previously subdivided among different protocols, such as ICMP, IGMP (Internet Group Membership Protocol version 3), and ARP (Address Resolution Protocol) and it introduces some simplifications by eliminating obsolete types of messages no longer in use.

ICMPv6 is a multi-purpose protocol and it is used for reporting errors encountered in processing packets, performing diagnostics, performing ND and reporting IPv6 multicast memberships. For this reason, ICMPv6 messages are subdivided into two classes:

5.3.1 Error messages

The first type of ICMPv6 message is the error message. ICMPv6 is used by IPv6 nodes to report errors encountered.

Type	Message
1	Destination Unreachable
2	Packet Too Big
3	Time Exceeded
4	Parameter Problem

5.3.2 Information messages

The second type of ICMPv6 message is the informational message type which is subdivided into three groups: diagnostic, management of multicast groups, and ND messages.

Type	Message
128	Echo Request
129	Echo Reply
130	Group Membership Query
131	Group Membership Report
132	Group Membership Reduction
133	Router Solicitation
134	Router Advertisement
135	Neighbour Solicitation
136	Neighbour Advertisement
137	Redirect
138	Router Renumbering

6. IPv6 Stateless Address Auto-configuration (SLAAC)

SLAAC is an IPv6 process that removes the requirement for the manual configuration of hosts, minimal configuration of routers, and no additional servers. The stateless mechanism enables a host to generate its own global address. It is based on ICMPv6. The stateless mechanism uses local information as well as non-local information that is advertised by routers to generate the addresses.

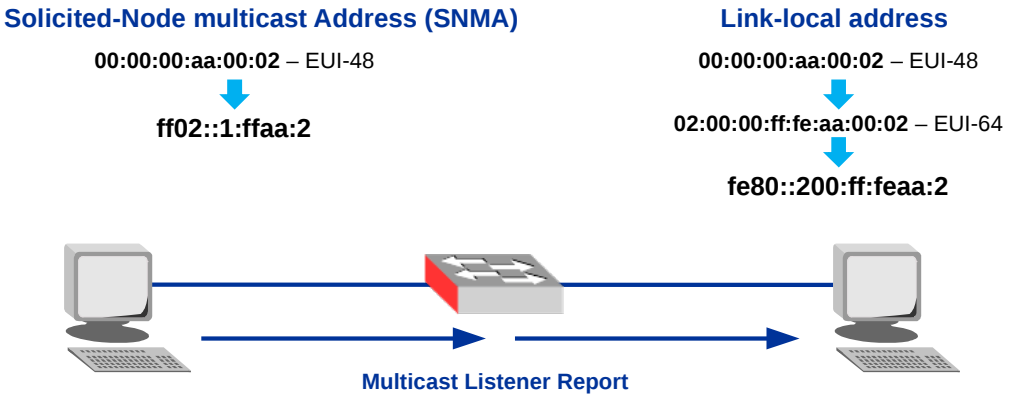
Routers advertise prefixes that identify the subnet or subnets that are associated with a link. Hosts generate an interface identifier that uniquely identifies an interface on a subnet. An address is formed by combining the prefix and the interface identifier. In the absence of routers, a host can generate only *link-local* addresses. However, *link-local* addresses are only sufficient for allowing communication among nodes that are attached to the same link.

The steps to SLAAC are:

- Host creates a SNMA
  - Host registers a Multicast Listener Report for SNMA to join group
  - from (::) to ff02::16 Multicast Listener Discovery (MLD).
- Host creates a LLA.
- Sends Neighbour Solicitation (NS) (135) from (::) to SNMA with LLA as target
  - If Neighbour Advertisement (NA) (136) received auto-configuration stops.
- Host registers a Multicast Listener Report for SNMA address to join group
  - from LLA to ff02::16 MLD.
- Host sends Router Solicitation (133) to ff02::2 ‘All routers’ from LLA.
- Router sends Router Advertisement (134) to ff02::1 ‘All nodes’
  - from its LLA with prefix.
- Host creates GUA from prefix and EUI-64 MAC
  - Sends NS (135) from (::) to SNMA with GUA target
  - If NA (136) received auto-configuration stops.
- Finish SLAAC.

There are security concers of mapping the MAC and the IPv6. RFC 4941 presents Privacy Extensions for Stateless Address Autoconfiguration in IPv6 to combat the issue. In reality within enterprises the linking of EUI-64 to IPv6 addresses may be useful but on public networks it is recommended to apply privacy extensions.

6.1 MLD joins multicast group

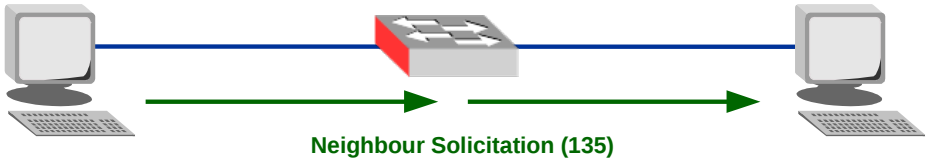



```
Ethernet II, Dst: 33:33:ff:aa:00:02, Src: 00:00:00:aa:00:02; Type: IPv6 (0x86dd)
Internet Protocol Version 6, Src: ::, Dst: ff02::16
    Payload length: 36
    Next header: IPv6 Hop-by-Hop Option (0)
    Hop limit: 255
    Hop-by-Hop Options
        Next Header: ICMPv6 (58)
Internet Control Message Protocol v6
    Type: Multicast Listener Report Message v2 (143)
    Multicast Address Record Changed to exclude: fe80::200:ff:feaa:2
```

Illustration 6: MLD joins multicast group

The host sends a Multicast Listener Report to the MLD multicast address *ff02::16* for the LLA address it generated *fe80::200:ff:feaa:2*. The purpose of MLD is to enable multicast routers to learn which multicast addresses and which sources have interested listeners on that link. The information gathered by MLD is provided to whichever multicast routing protocol is used by the router, in order to ensure that multicast packets are delivered to all links where there are listeners interested in such packets.

6.2 Neighbour Solicitation (135)




```
Ethernet II, Dst: 33:33:ff:aa:00:02, Src: 00:00:00:aa:00:02; Type: IPv6 (0x86dd)
Internet Protocol Version 6, Src: ::, Dst: ff02::1:ffaa:2
   Payload length: 24
   Next header: ICMPv6 (58)
   Hop limit: 255
Internet Control Message Protocol v6
   Type: Neighbor Solicitation (135)
   Target Address: fe80::200:ff:feaa:2
```

Illustration 7: Neighbour Solicitation (135)

The host then sends a NS (135) ICMPv6 message to the SNMA address from the unassigned address (::) with the LLA as the target address. If a NA (136) is received then a duplicate address has been detected and the process stops. If no NA (136) is detected then the process continues.

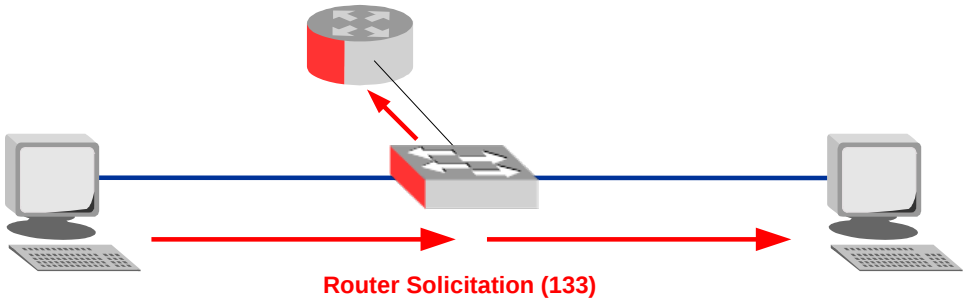
6.3 MLD joins multicast group for the second time


```
Ethernet II, Dst: 33:33:00:00:00:16, Src: 00:00:00:aa:00:02; Type: IPv6 (0x86dd)
Internet Protocol Version 6, Src: fe80::200:ff:feaa:2, Dst: ff02::16
   Payload length: 36
   Next header: IPv6 Hop-by-Hop Option (0)
   Hop limit: 255
   Hop-by-Hop Options
      Next Header: ICMPv6 (58)
Internet Control Message Protocol v6
   Type: Multicast Listener Report Message v2 (143)
   Multicast Address Record Changed to exclude: ff02::1:ffaa:2
```

Illustration 8: MLD joins multicast group (2)

The host sends a second Multicast Listener Report to the MLD multicast address *ff02::16* but this time with the LLA as the source and the SNMA as the Multicast Address Record Changed to exclude.

6.4 Router Solicitation (133)

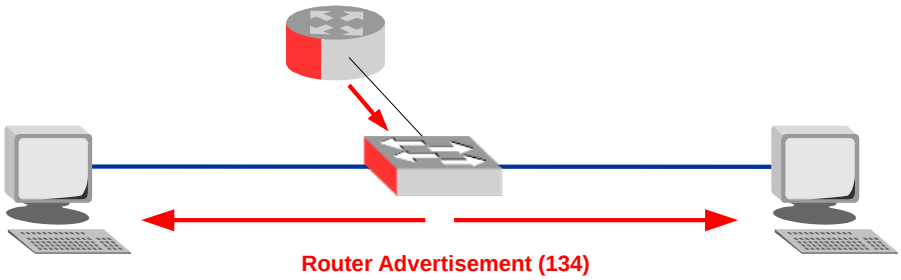


Ethernet II, Dst: **33:33**:00:00:00:02, Src: 00:00:00:aa:00:02;
Type: IPv6 (0x86dd)
Internet Protocol Version 6, Src: **fe80::200:ff:feaa:2**, Dst: **ff02:2**
Payload length: 16
Next header: **ICMPv6 (58)**
Hop limit: 255
Internet Control Message Protocol v6
Type: **Router Solicitation (133)**
Link-layer address: **00:00:00:aa:00:02**

Illustration 9: Router Solicitation (133)

The host now sends an RS (133) ICMPv6 message to the multicast group *ff02::2*, the group of all routers on the link-local ising its LLA as the source IPv6 address.

6.5 Router Advertisement (134)




```
Ethernet II, Dst: 33:33:00:00:00:01, Src: 00:00:00:aa:00:03; Type: IPv6 (0x86dd)
Internet Protocol Version 6, Src: fe80::200:ff:feaa:3, Dst: ff02:1
   Payload length: 16
   Next header: ICMPv6 (58)
   Hop limit: 255
Internet Control Message Protocol v6
   Type: Router Advertisement (134)
   Flags: 0xc0          .1.. .... = Other configuration: Set
   ICMPv6 Option, Prefix information
      Prefix Length: 64
      Flag: 0xc0        1... .... = On-link flag(L): Set
                        .1.. .... = Autonomous address-configuration flag(A): Set
      Valid Lifetime: 86400, Preferred Lifetime: 86400
      Prefix: 2001:a::
   ICMPv6 Option
      Link-layer address: 00:00:00:aa:00:03
```

Illustration 10: Router Advertisement (134)

The router responds with an RA (134) to the multicast group *ff02:1*, the group of all hosts on the link-local. As part of this message it sends the network prefix *2001:1::/64* with the prefix length to the hosts plus some flags.

6.5.1 Router Advertisement flags

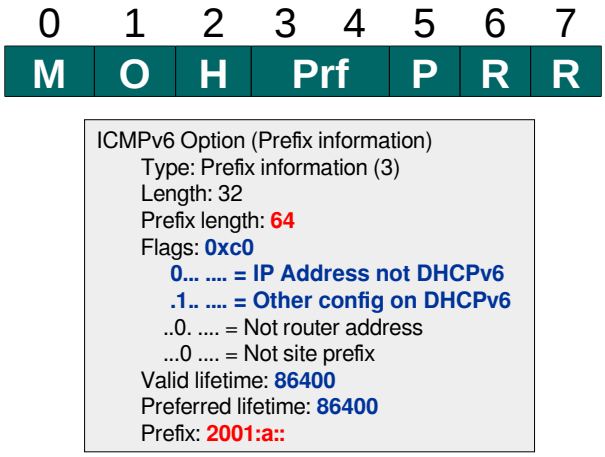


Illustration 11: RA flags

RA (134) informational messages contain two flags that indicate what type of stateful Auto-configuration should be performed. A Managed address configuration flag (M-Flag) indicates whether hosts should use stateful auto-configuration to obtain global scope IPv6 addresses. The other stateful configuration flag (O-Flag) if set (1) indicates that hosts should use stateful auto-configuration to obtain additional information, excluding addresses, from a stateless DHCPv6 Server. Additionally the RA (134) – Prefix flags have the options On-link flag (L-flag) and Autonomous address configuration flag (A-flag).

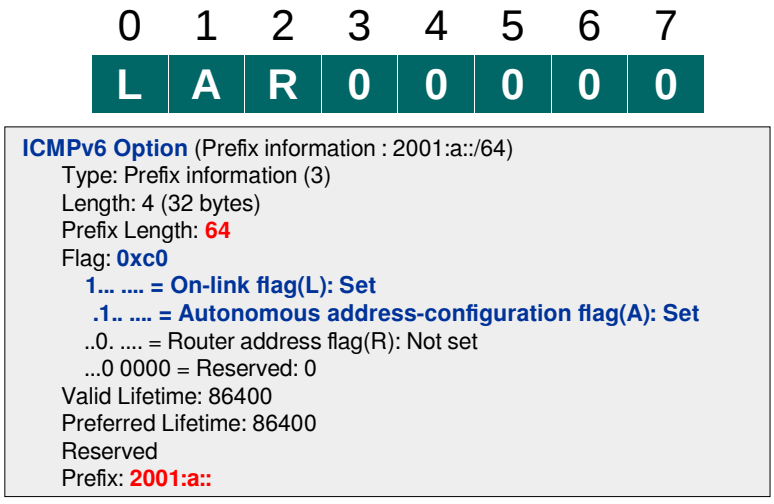
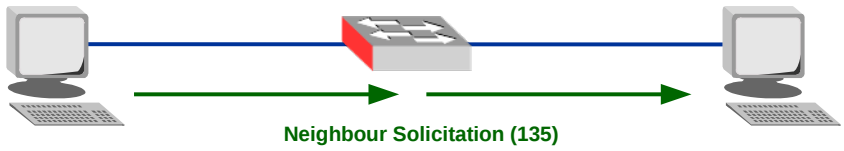


Illustration 12: RA - Prefix flags

The following table outlines the infleuence of the '**M**' & '**A**' flags on auto-configuration. Remember that hosts must be set to obtain IP address *automatically* and all hosts continue to generate and use a LLAs.

M	A	Resulting non-Link Local addresses on client
0	0	No addresses will be auto-configured
0	1	Address generated from prefix in RAs
1	1	Address generated from prefix) in RAs; Full address from DHCP server
1	0	Full address(es) from DHCP server

6.6 Neighbour Solicitation (135) for the second time



Ethernet II, Dst: 33:33:ff:aa:00:02, Src: 00:00:00:aa:00:02; Type: IPv6 (0x86dd)
Internet Protocol Version 6, Src: ::, Dst: ff02::1:ffaa:2
Payload length: 24
Next header: ICMPv6 (58)
Hop limit: 255
Internet Control Message Protocol v6
Type: Neighbor Solicitation (135)
Target Address: 2001:a::200:ff:feaa:2

Illustration 13: Neighbour Solicitation (135) (2)

Having received a prefix from the router the host generates a GUA by adding its EUI-64 MAC that it used to generate the LLA to the prefix. It then sends a NS (135) to the SNMA address from the unassigned address (::) with the new GUA as the target address. It doesn't expect a responding NA (136) and should it receive one the SLAAC process will stop.

6.7 MLD joins multicast group for the third time


```
Ethernet II, Dst: 33:33:00:00:00:16, Src: 00:00:00:aa:00:02; Type: IPv6 (0x86dd)
Internet Protocol Version 6, Src: fe80::200:ff:feaa:2, Dst: ff02::16
   Payload length: 36
   Next header: IPv6 Hop-by-Hop Option (0)
   Hop limit: 255
   Hop-by-Hop Options
      Next Header: ICMPv6 (58)
Internet Control Message Protocol v6
   Type: Multicast Listener Report Message v2 (143)
   Multicast Address Record Changed to exclude: ff02::1:ffaa:2
```

Illustration 14: MLD joins multicast group (3)

The host then sends a third Multicast Listener Report to the MLD multicast address *ff02::16* from the LLA as the source and the SNMA as the Multicast Address Record Changed to exclude value.

7. IPv6 Address Resolution and redirection

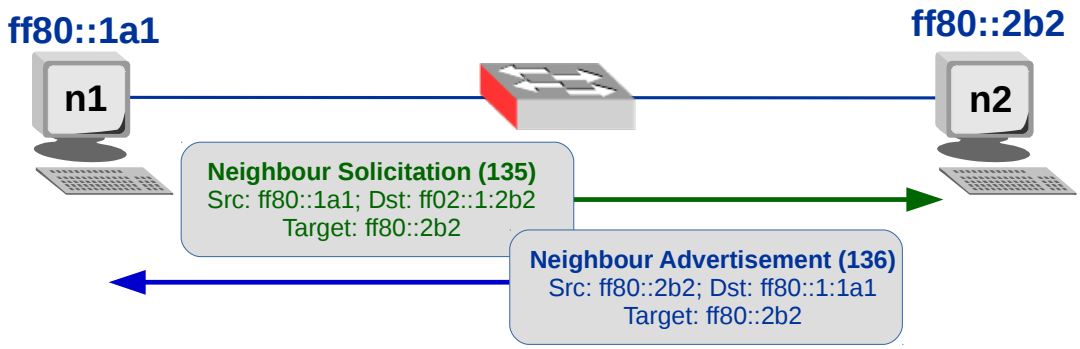


Illustration 15: Address Resolution

The IPv6 Address Resolution process uses Neighbour Discovery (ND) to verify connectivity and establish two-way connectivity. A NS (135) packed is sent from the hosts LLA to the SNMA address of the target node with the target LLA within the ICMPv6 header. If the targetted node is available it responds with a NA (136) to the LLA of the requestor from its own LLA and with its own LLA as the target address within the ICMPv6 header. Node n1 can now record an established two-way connectivity with node n2.

7.1 Neighbour Unreachability Detection (NUD)

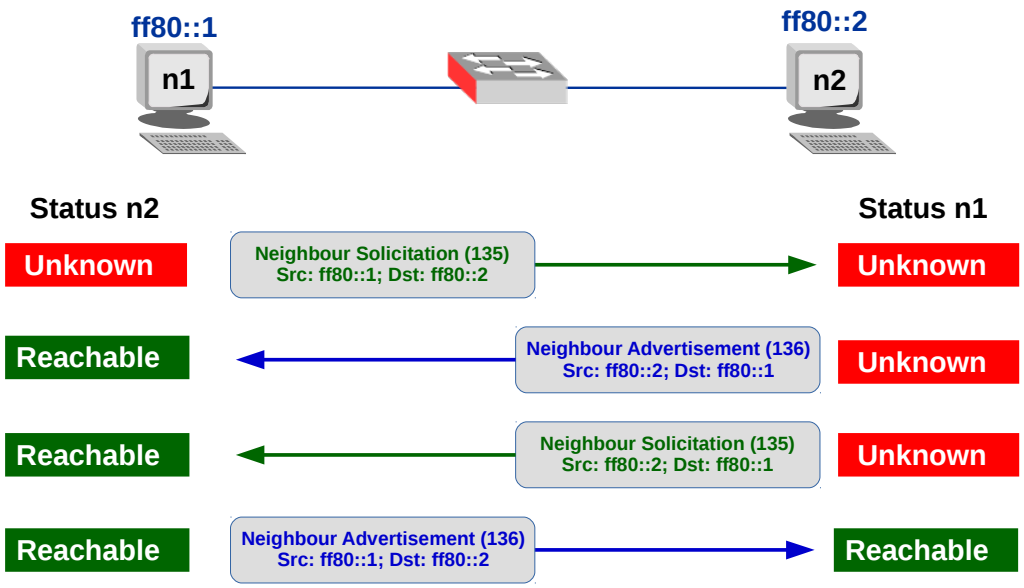


Illustration 16: Neighbour Unreachability Detection (NUD)

IPv6 Neighbour Unreachability Detection (NUD) detection improves packet delivery in the presence of failing routers. This capability improves packet delivery over partially failing or partitioned links. This capability also improves packet delivery over nodes that change their link-local addresses. For example, mobile nodes can move off the local network without losing any connectivity because of stale ARP caches.

IPv4 has no corresponding method for NUD. Unlike ARP, ND detects half-link failures by using NUD and it therefore avoids sending traffic to neighbours when two-way connectivity is absent. NUD improves the robustness of packet delivery in the presence of failing routers or links, or mobile nodes.

Illustration 16 Demonstrates the process, node n1 sends a unicast NS (135) message to n2 who responds with a unicast NA (136). Node n2 then sends a unicast NS (135) message to node n1 who responds with a unicast NA (136) of its own. In this was the two nodes confirm their established two-way connectivity.

7.2 ICMPv6 Redirect

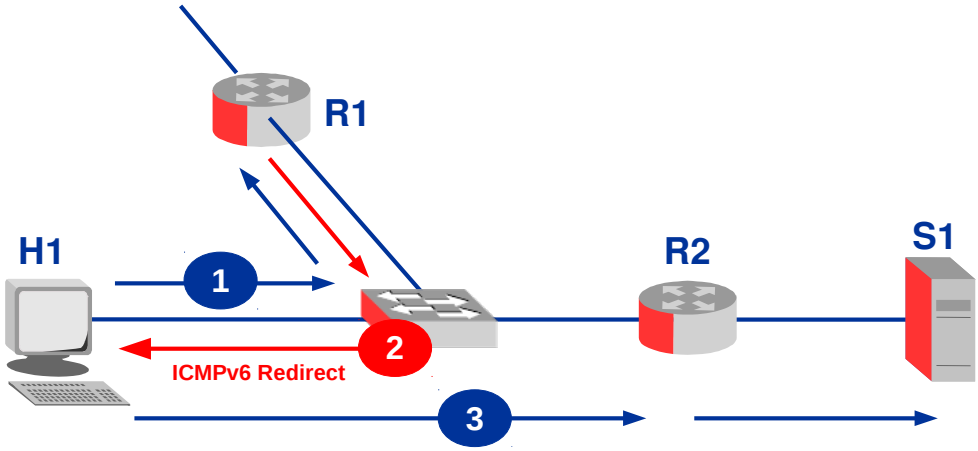


Illustration 17: ICMPv6 Redirect

On an IPv6 network, hosts don't know much about routes. They send their packets to a default gateway which handles the routing for them.

In the case of the local network having a single router, then hosts will send all non-local traffic to the router. However if there is more than one local router, the host then must decide which router to use for which traffic. Generally the host does not know the most efficient choice of gateway router for frame it needs to send.

As demonstrated in Illustration 17 when a router receives datagrams destined for certain networks, it may realise that it would be more efficient if such traffic was sent by the host to a different router on the local network. In this case it can invoke the Redirect function by sending an ICMPv6 Redirect message to the device that sent the original packet. Redirect messages are always sent unicast to the address of the device that originally sent the packet that led to the Redirect being created.

8. IPv6 Configuration best practice – Inter-router links

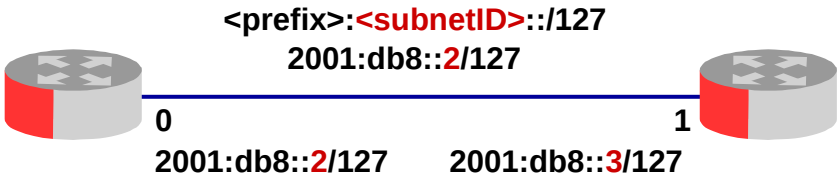


Illustration 18: Inter-router link

Best practice: use /127 for inter-router links. Addresses with the following 64 bits must NOT be used:

- 0000:0000:0000:0000
- ffff:ffff:ffff:ff7f ➠ :ffff

8.1 Using LLA on Inter-router links

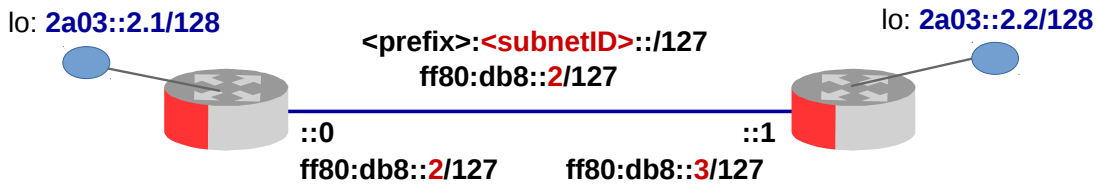


Illustration 19: Inter-router link with LLA

It is possible to use LLA on infrastructure links and it has advantages. However it is important to configure a GUA on a loopback address on each device for management plane traffic (ssh, telnet, SNMP, etc). The source ICMPv6 error messages destined off-subnet.

The advantages of this approach are:

- Smaller routing tables which leads to
  - less memory consumption
  - faster routing convergence
  - Accelerated forwarding due to smaller RIBs & FIBs.
- Simpler address management
- Lower configuration complexity
- Simpler DNS as LLAs are not put into zone files)
- Reduced attack surface

Caveats

- Router-interfaces not ping-able from off-link (fix: ping the loopback)
- Traceroutes to these interfaces break
- Hardware dependency – LLAs change if line cards change
- NMS functions that are interface-address specific will break

9. IPv6 Address planning

An ISP performing address planning considering how many addresses to request from the RIR needs to plan their address space. Take for example an ISP in a country with 10 regions, each with 50 Points of Presence (POP) and each of these supporting 3,500 clients. Calculate the number of bits in the mask for each tier in multiples of 4 (i.e. nibbles). Take the value that gives a result higher that the requirement. For example 12 bits = 2^{12} = 4,096 where 2^8 = 256 so 12 bits are required for the blocks of 3,500 clients. A similar process is carried out for each item as outlined below. Assuming each client is assigned a /48 the mask for POPs can be determined by subtracting 12 from 48 giving a /36 for POPs and subtract 8 from 36 to give a /28 for Regions and finally subtracting 4 from 28 gives a /24 for the ISP. Therefore in this example the IPS requires a /24 from the RIR.

Item	#	Bits (multiple of 4)	Possible #	Mask
ISP	1	1	2	/24
Regions	10	4	16	/28
POPs	50	8	256	/36
Clients	3,500	12	4,096	/**48**

24

10. IPv6 Interior Gateway Routing

10.1.1 RIPng

Like its IPv4 variant *RIPng* is a Distance vector algorithm. It has a number of implementations: GateD, MRTd, Kame, route6d, Quagga as well as vendor equipment solutions from companies like Cisco, Juniper, HP, Huawei, MikroTik etc....

10.1.2 OSPFv3

OSPFv3 is a Link State algorithm like the IPv4 version. It is the recommended IGP of IETF. The main differences from OSPFv2 are the removal of security as IPv6 has its own implementation embedded and the format of addresses are for IPv6. Implementations: GateD, MRTd, Kame, route6d, Quagga and vendor hardware solutions from companies like Cisco, Juniper, HP, Huawei, MikroTik etc....

10.2 IPv6 Exterior Gateway Routing

10.2.1 BGP4+

BGP4+ is the standard Inter domain routing protocol for IPv6. It is used between ISPs and carriers and its extensions to BGP4 are defined in RFC 2858. RFC 2545 defines how to use IPv6 extensions. It is used in 6BONE and the following are implementations today: GateD, MTRd, Kame, BGPd, Quagga and vendor hardware solutions from companies like Cisco, Juniper, HP, Huawei, MikroTik etc....

11. IPv6 transition mechanisms

Until IPv6 completely replaces IPv4, a number of transition mechanisms are needed to enable IPv6-only hosts to reach IPv4 services and to allow isolated IPv6 hosts and networks to reach the IPv6 Internet over the IPv4 infrastructure. As the IPv6 Internet grows larger, the need also arises for carrying IPv4 traffic over the IPv6 infrastructure.

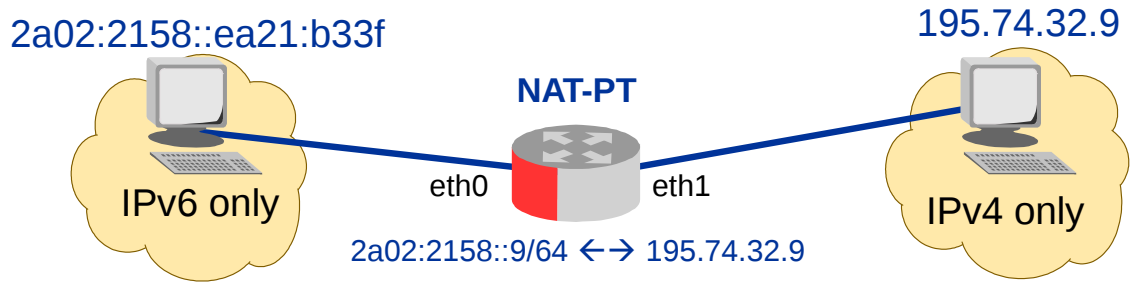
11.1 Dual Stack

IPv6 is a form of extension of IPv4 and therefore it is relatively easy to write a network stack that supports both IPv4 and IPv6 while sharing most of the code. Dual Stack is implemented by the various OS today. Some early experimental implementations used independent IPv4 and IPv6 stacks. There are no known implementations that implement IPv6 only. Actually when used in IPv4 communications, hybrid stacks tend to use an IPv6 API and represent IPv4 addresses in a special address format, the IPv4-mapped IPv6 address.

11.2 Proxying and translation

When an IPv6 only host needs to access an IPv4 only host, translation is necessary. The one form of translation that actually works is the use of a dual stack application-layer proxy. Techniques for application agnostic translation at the lower layers have also been proposed, but they have been found to be too unreliable in practice due to the wide range of functionality required by common application-layer protocols, and are commonly considered to be obsolete.

11.2.1 Network Address Translation - Protocol Translation (NAT-PT)




```
NAT-PT Router(config)# interface ethernet 1/0
NAT-PT Router(config-if)# ipv6 address 2a02:2158::9/64
NAT-PT Router(config-if)# ipv6 nat
NAT-PT Router(config-if)# exit
NAT-PT Router(config)# interface ethernet 1/1
NAT-PT Router(config-if)# ip address 195.74.32.9 255.255.255.0
NAT-PT Router(config-if)# ipv6 nat
```

Illustration 20: NAT-PT

NAT-PT is a protocol translator between IPv6 and IPv4 that allows direct communication between hosts speaking different network protocols. However RFC 4966 outline reasons to Move the Network Address Translator - Protocol Translator (NAT-PT) to historic status.

11.2.2 NAT64 and DNS64

NAT64 is an IPv6 transition mechanism that facilitates communication between IPv6 and IPv4 hosts by using a form of NAT. The NAT64 gateway is a translator between IPv4 and IPv6 protocols. NAT64 requires at least one IPv4 address and an IPv6 network segment comprising a 32-bit address space to cater for the IPv4 network.

An IPv6 client embeds the IPv4 address it wishes to communicate with using the host part of the IPv6 network segment, resulting in an IPv4-embedded IPv6 addresses, and sends packets to the resulting address. The NAT64 gateway creates a mapping between the IPv6 and the IPv4 addresses, which may be manually configured or determined automatically.

Typically, NAT64 is designed to be used when the communication is initiated by IPv6 hosts. Some mechanisms, including static address mapping, exist to allow the inverse scenario.

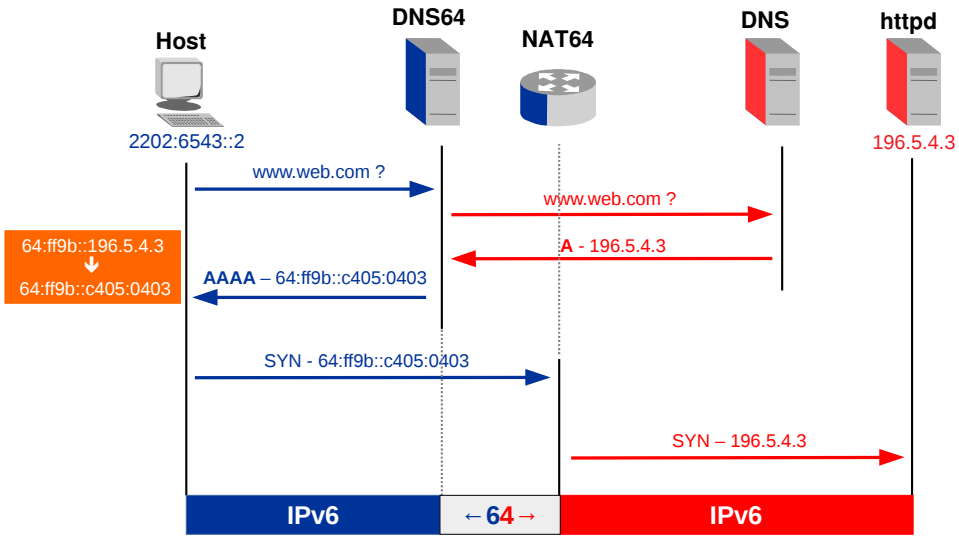


Illustration 21: NAT64 / DNS64

Illustration 21 demonstrates how an IPv6 host acccesses an IPv4 service using NAT64/DNS64. The IPv4 host makes a IPv4 DNS request to the DNS64 server which forwards the request on the IPv4 network to the DNS server. The DNS server responds with an A record, 196.5.4.3 for the webserver. The DNS64 server translates the IPv4 address by combining it with the NAT64 prefix 64:ff9b::/96 to form an IPv6 address 64:ff9b::c405:0403 and forwards this as an AAAA record to the requesting host.

The host forms a packet and sends it to the NAT64 router, it being its gateway which translates the IPv6 packet header to an IPv4 packet header changing the IP address to the IPv4 format. It continues to translate the packet headers in each direction for the duration of the connection.

11.3 Tunnelling

In order to reach the IPv6 Internet, an isolated host or network must be able to use the existing IPv4 infrastructure to carry IPv6 packets. This is achieved using a technique known as tunnelling which consists of the encapsulation of IPv6 packets within IPv4, in effect using IPv4 as a link layer for IPv6.

IPv6 packets can be directly encapsulated within IPv4 packets using protocol number 41. They can also be encapsulated within UDP packets e.g. in order to cross a router or NAT device that blocks protocol 41 traffic. Another options is to use generic encapsulation schemes like Generic Routing Encapsulation (GRE).

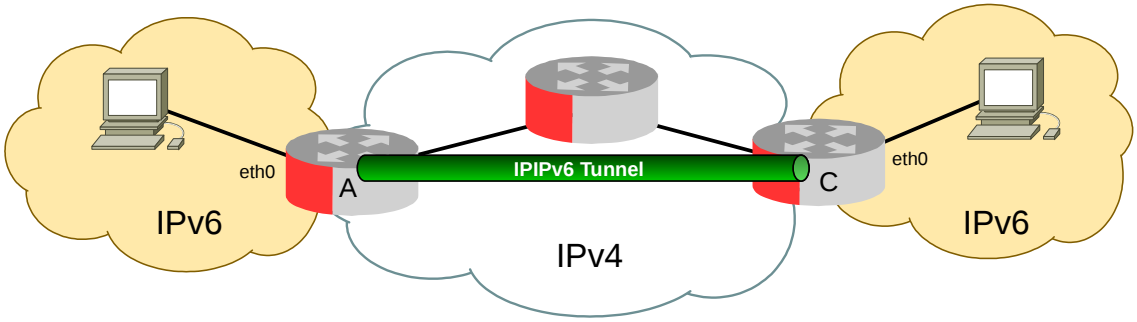


Illustration 22: Overlay tunnels for IPv6

11.3.1 Manual tunnelling

Manual tunnelling is done by manually configuring the end points of the tunnel. This tunnelling method can be used for sites with few nodes or for a limited number of remote connections. As is the case with static routing, scalability and management overhead are major issues limiting the use of manual tunnelling.

11.3.2 Automatic tunnelling

Automatic tunnelling refers to a technique where the tunnel endpoints are automatically determined by the routing infrastructure.

11.3.3 Connection of IPv6 Domains via IPv4 Clouds (6to4)

The recommended technique for automatic tunnelling is 6to4 tunnelling, which uses protocol 41 encapsulation. Tunnel endpoints are determined by using a well-known IPv4 anycast address on the remote side, and embedding IPv4 address information within IPv6 addresses on the local side. 6to4 is widely deployed today.

6to4 performs three functions:

- Assigns a block of IPv6 address space to any host or network that has a global IPv4 address.
- Encapsulates IPv6 packets inside IPv4 packets for transmission over an IPv4 network using 6in4.
- Routes traffic between 6to4 and "native" IPv6 networks.

11.3.4 Teredo: Tunnelling IPv6 over UDP through NATs

Teredo is an automatic tunnelling technique that uses UDP encapsulation and is capable of crossing multiple NAT devices. Teredo gives IPv6 connectivity to IPv6 capable hosts which are on the IPv4 Internet but have have no direct native connection to an IPv6 network. Teredo is not widely deployed today. *Miredo* is the GNU/Linux and BSD UNIX open-source implementation of Teredo.

11.3.5 Intra-Site Automatic Tunnel Addressing Protocol (ISATAP)

ISATAP is an IPv6 transition mechanism designed to transmit IPv6 packets between nodes with dual-stack (IPv6/IPv4) over IPv4 networks. ISATAP views the IPv4 network as a link layer for IPv6 and supports an automatic tunnelling abstraction similar to a Non-Broadcast Multiple Access (NBMA) model. ISATAP is designed for transporting IPv6 packets *within* a site, not *between* sites.

11.3.6 RFC 3053 Tunnel Broker (TB)

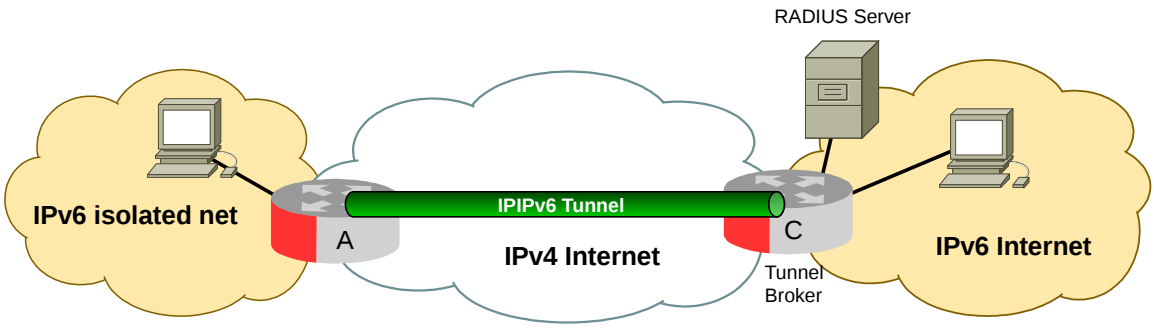


Illustration 23: Tunnel Broker (TB)

A Tunnel Broker (TB) allows isolated users/routers to connect to the IPv6 network. The router or host establishes an IPv6 encapsulation over IPv4 to the TB who then authenticates the connection using RADIUS. If authorised the router or host is assigned an IPv6 address and can now route IPv6.

This page is intentionally blank