



**IPv6 Workshop  
Gaborone, Botswana  
22<sup>nd</sup> February 2008**

# Contents

---

**Introduction**

The IPv6 datagram

Addressing

Transition Mechanisms

Applications

---

# Introduction

---

In the 70s, the Internet Protocol (IP) was designed.

- IP is based on publicly available standards
  - – Published by Internet Engineering Task Force

<http://www.ietf.org>

- – RFCs

<http://www.ietf.org/rfc.html>

- – IETF Working Groups

<http://www.ietf.org/html.charters/wg-dir.html>

---

# Why another protocol?

---

Despite the success of IPv4 the Internet Protocol needs an important revision.

- exponential internet growth has led to the imminent exhaustion of address space and global routing table growth.
-

# IPv4 Addressing

---

- In theory, 32 bits of IPv4 enables 4 billion hosts. Realistically, the HD ratio limits IPv4 to 250million hosts.

## **IPv4 space utilization:**

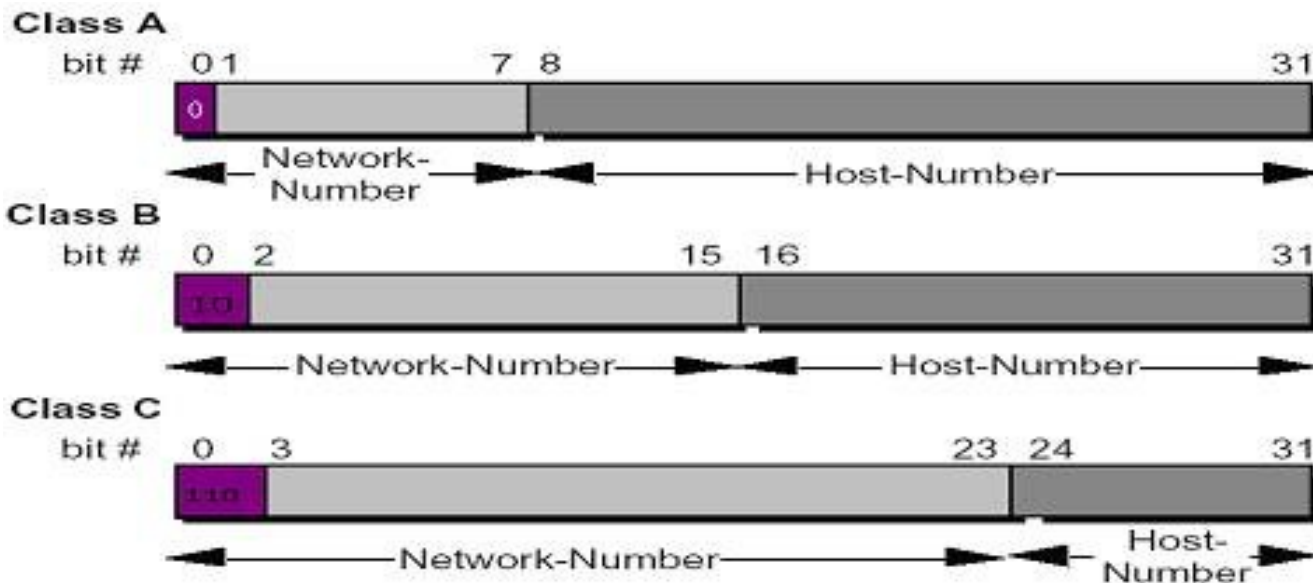
As of Feb 2007, only 41/256 (16%) blocks are left for allocation by the IANA to the RIRs.

Predictions indicate exhaustion around 2010-2011.

---

# IPv4 Primary Address Classes

Historically, IPv4 was divided into three address classes-Class A, Class B, and Class C to provide flexibility for networks of varying sizes



## Problems of classful addressing

---

The classful A, B, and C boundaries were easy to understand and implement, but they did not foster the efficient allocation of the finite address space.

- Exhaustion of IP Class B Address Space.
  - Exhaustion of IP Address Space in General.
  - Non-hierarchical nature of address allocation leading to flat routing space.
-

# Solutions to addressing problems

---

Short term solution:-

- CIDR (RFC 4632)
- Other
  - Use Private space (RFC 1918)
  - Use NAT (RFC 3022)

Long term:

- New Protocol with larger address space (RFC 1752)
-

## CIDR

- The change from Class A/B/C network numbers to “classless” prefixes.
- Make explicit which bits in a 32-bit IPv4 address are interpreted as the network number (or prefix) associated with a site and which are the bits used to number individual end systems within the site.
- In CIDR notation, a prefix is shown as a 4-octet quantity, just like a traditional IPv4 address or network number, followed by the "/" (slash) character, followed by a decimal value between 0 and 32 that describes the number of significant bits.

# Private Network addresses

---

- The IETF defined a series of IP ranges which may be freely used by any user/organization on their PRIVATE network. These addresses may not be used on the PUBLIC internet and all PUBLIC INTERNET gateways (routers) will silently discard the traffic (or unusually reject with an ICMP message).
  - The use of PRIVATE address means that the same IP addresses may be reused by many organisations and only when the traffic is sent to the PUBLIC network must it be translated using mediating gateway (ALG or NAT) to a PUBLIC address.
-

# Network Address Translation (NAT)

---

- Network Address Translation is a method by which IP addresses are mapped from one address realm to another, providing transparent routing to end hosts.
  - Types of NAT
    - Basic NAT
    - NATPT
    - Bidirectional (Two way) NAT
    - Twice NAT
-

# NAT has many implications

---

- Inhibits end-to-end network security
  - When a new application is not NAT-friendly, NAT device requires an upgrade
  - Some applications cannot work through NATs
  - Application-level gateways (ALG) are not as fast as IP routing
  - Complicates mergers
  - Double NATing is needed for devices to communicate with each other
  - Breaks security
  - Makes multihoming hard
  - Simply does not scale
-

## IPng (IP Next Generation)

---

- In December 1993, the IETF chartered a new working group named Internet Protocol – Next Generation, or IPng which solicited proposals from interested parties on requirements and concerns for the new protocol (RFC 1550)
  - These responses were considered, listing 17 criteria developing RFC 1726 in December 1994
  - Several proposed protocols were evaluated vis-à-vis these criteria, with those evaluations published in January 1995 with RFC 1752. It gave the IPng protocol a new name – IPv6.
-

# IPv6

---

What happened to IPv5?

Version 5 in IP header was assigned to ST protocol (a.k.a, Internet Streaming Protocol)

- – Experimental non-IP real-time streaming protocol
  - – Never widely used
  
  - RFC 1819
-

# IPv6 Features

---

- Large Address Space
  - New Types of Addresses
  - Auto-configuration
  - IPv6 has a new 40-bytes header
  - Better Network Management
  - Improved Mobility Support
  - Support for IPsec
  - QoS
-

# Contents

---

Introduction

**The IPv6 datagram**

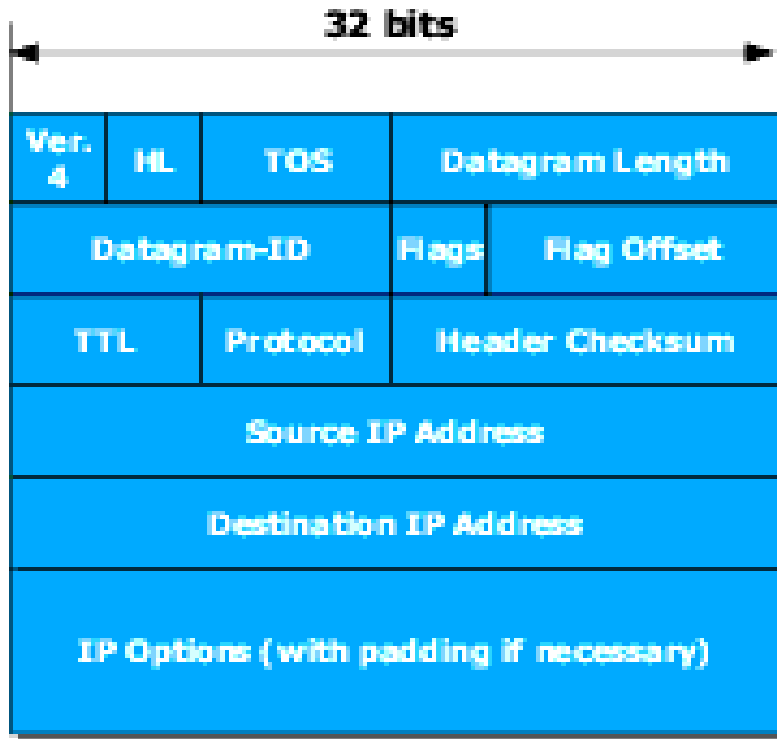
Addressing

Transition Mechanisms

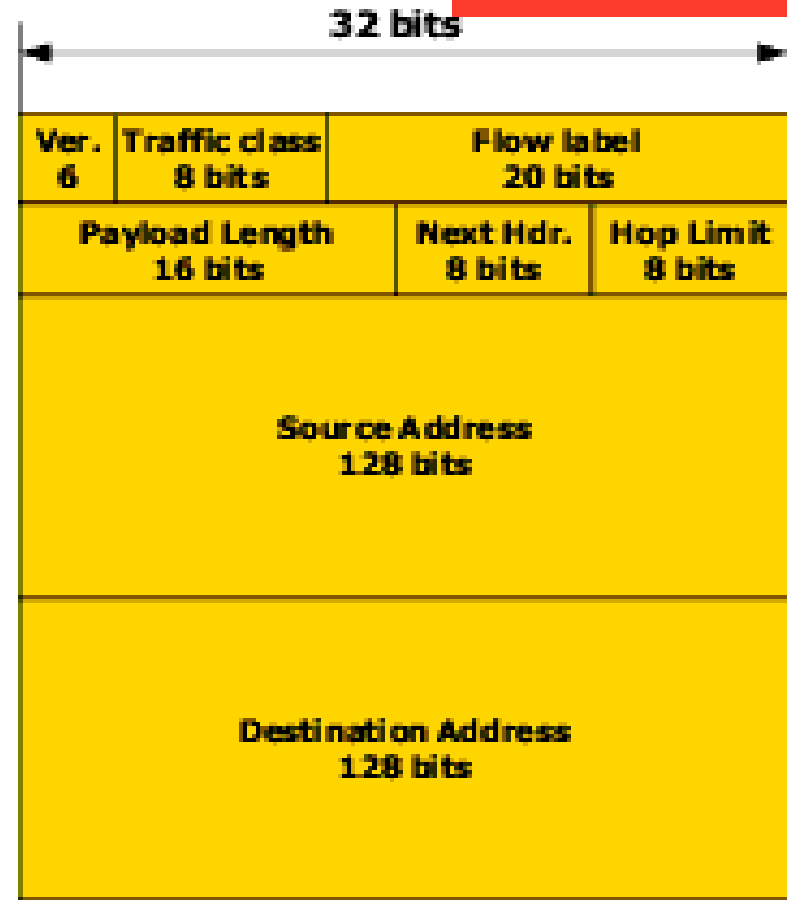
Applications

---

# IPv6 Datagram



**IPv4 header**



**IPv6 header**

---

The basic IPv6 header is simpler than the IPv4 header: It has 8 fields instead of 12. The IP header length and checksum fields are removed, the flow label field is added and the fragmentation fields are moved in the extension headers. The size of the basic header is 40 bytes, which is double the IPv4 header size (but the size of the addresses is 4 times).

---

# IPv6 header fields

The 40-byte IPv6 header consists of the following eight fields:

- **Version** - Indicates the version of the Internet Protocol.
- **Traffic class** - Previously the type-of-service (ToS) field in IPv4, the traffic class field defines the class-of-service (CoS) priority of the packet.
- **Flow label** - The flow label identifies all packets belonging to a specific flow (that is, packet flows requiring a specific class of service [CoS]); routers can identify these packets and handle them in a similar fashion.
- **Payload length** - Previously the total length field in IPv4, the payload length field specifies the length of the IPv6 payload.
- **Next header** - Previously the protocol field in IPv4, the Next Header field indicates the next extension header to examine.
- **Hop limit** - Previously the time-to-live (TTL) field in IPv4, the hop limit indicates the maximum number of hops allowed.
- **Source address** - Identifies the address of the source node sending the packet.
- **Destination address** - Identifies the final destination node address for the packet.

# Extension headers

---

- IPv6 extension headers are similar to IPv4 options
  - Extension headers were chosen for the purpose of compromising generality and efficiency. IPv6 needs to include mechanisms to support functions such as fragmentation, source routing and authentication. However choosing to allocate fixed fields in the datagram header for all mechanisms is inefficient because most datagrams do not use all fields. E.g. a header that contains an empty field can occupy a substantial fraction of each frame.
-

## Next Header

Each type of extension header is identified by a next header specific value. Ext headers are daisy chained, the next header field of the successive extension headers points to the next ext header till the last ext. header where the next header points to the transport header.

- Hop-by-hop Option
- Routing Header
- Fragment Header
- Encapsulation Security Payload
- Authentication Header
- Destination Options
- No next header

- new version of ICMP integral to IPv6 that must be completely supported by all IPv6 implementations and nodes.
- a multipurpose protocol used for reporting errors encountered in processing packets, performing diagnostics, performing Neighbor Discovery and reporting IPv6 multicast membership.
- ICMPv6 messages are grouped into two classes: error messages and informational messages.

Error messages: Destination Unreachable, Packet Too Big, Time Exceeded, Parameter Problem.

Info messages: Echo Request, Echo Reply.

# Neighbor Discovery 1/3

---

- ND protocol manages interactions between nodes via message exchanges
  - Nodes use ND to determinate the link-layer addresses for neighbors known to reside on attached links and purge cached invalid values
  - Hosts use ND to find neighboring routers that are willing to forward packets on their behalf
  - Nodes use the protocol to actively keep track of which neighbors are reachable and which are not, and to detect changed link-layer addresses
  - Replaces ARP, ICMP Router Discovery, and ICMP redirect used in IPv4
-

## Neighbor Discovery 2/3

---

Discovery functions performed by routers

- Router Advertisements (RA): Routers periodically transmit messages about the router and the network
  - Parameter maintenance: Routers maintain information about key parameters of the local network
  - Process Router Solicitations (RS): Routers listen for solicitation messages and when one is received, immediately send the RA to requestor.
-

## Neighbor Discovery 3/3

---

Discovery functions performed by hosts

- Process advertisements: Hosts listen for RAs and set parameters based on received info.
  - Generate Solicitations: Sometimes they generate SAs without having to wait for RAs say when a host has just been turned on.
  - Auto-configuration: When required and if the network supports it, the host will use the info from the router to auto-configure itself with an IP and other parameters.
-

## IPv6 (stateless) autoconfiguration

---

- The device generates a link-local address
  - The node tests to ensure that the address is not used on the network
  - If it passes, the device assigns the link-local address to its IP interface
  - The node establishes contact with the local router either by listening to RAs or by sending RS
  - Router either informs node of DHCP server to use or how to determine its address using autoconfiguration
  - Host will configure itself with a globally unique address
-

# Contents

---

Introduction

The IPv6 datagram

**Addressing**

Transition Mechanisms

Applications

---

# Addressing

---

- IPv6 addresses are 128-bit wide

xxxx:xxxx:xxxx:xxxx:xxxx:xxxx:xxxx:xxxx

- Hexadecimal representation (Too Long For Dotted Decimal Notation)
  - Interfaces have several IPv6 addresses
  - CIDR principles [address prefix / prefix length]
    - 2001:42d0::/32
-

## Rules applied for address representation:-

- Letters are case insensitive
- Leading zeroes in a field are optional
- Successive fields of 0 are represented as :: but only once

Eg

2001:0000:2ABC:0000:0000:0000:0000:0001

2001:0000:2ABC::0001

2001:0:2ABC::1

# IPv6 Address Space

(last updated 2007-07-19)

IPv6 Prefix	Allocation	Reference
-----	-----	-----
0000::/8	Reserved by IETF	[RFC4291]
0100::/8	Reserved by IETF	[RFC4291]
0200::/7	Reserved by IETF	[RFC4048]
0400::/6	Reserved by IETF	[RFC4291]
0800::/5	Reserved by IETF	[RFC4291]
1000::/4	Reserved by IETF	[RFC4291]
2000::/3	Global Unicast	[RFC4291]
4000::/3	Reserved by IETF	[RFC4291]
6000::/3	Reserved by IETF	[RFC4291]
8000::/3	Reserved by IETF	[RFC4291]
A000::/3	Reserved by IETF	[RFC4291]
C000::/3	Reserved by IETF	[RFC4291]
E000::/4	Reserved by IETF	[RFC4291]
F000::/5	Reserved by IETF	[RFC4291]
F800::/6	Reserved by IETF	[RFC4291]
FC00::/7	Unique Local Unicast	[RFC4193]
FE00::/9	Reserved by IETF	[RFC4291]
FE80::/10	Link Local Unicast	[RFC4291]
FEC0::/10	Reserved by IETF	[RFC3879]
FF00::/8	Multicast	[RFC4291]

# Types of Addresses

---

- Unicast

An identifier for a single interface. A packet sent to a unicast address is delivered to the interface identified by that address.

- Multicast

An identifier for a set of interfaces (typically belonging to different nodes). A packet sent to a multicast address is delivered to all interfaces identified by that address.

- Anycast:

An identifier for a set of interfaces (typically belonging to different nodes). A packet sent to an anycast address is delivered to one of the interfaces identified by that address (the "nearest" one, according to the routing protocols' measure of distance)

---

# The unspecified address

---

- The address 0:0:0:0:0:0:0:0 is called the unspecified address. It must never be assigned to any node. It indicates the absence of an address.  
An IPv6 packet with a source address of unspecified must never be forwarded by an IPv6 router.
-

# The loopback address

---

- The unicast address 0:0:0:0:0:0:0:1 is called the loopback address. It may be used by a node to send an IPv6 packet to itself.
  - It must not be assigned to any physical interface. An IPv6 packet with a destination address of loopback must never be sent outside of a single node and must never be forwarded by an IPv6 router.
  - A packet received on an interface with a destination address of loopback must be dropped.
-

# Global Unicast

n	m	128-m-n
Prefix	Subnet ID	Interface ID

- The implementation chosen for IPv6 assigns 48 bits to the routing prefix and 16 bits to the subnet identifier. This means 64 bits are available for interface identifiers, which are constructed based on the IEEE “EUI-64” format.

## IPv6 addresses with embedded v4 addresses (1/2)

- IPv4-compatible IPv6 addresses (deprecated)

80 bits	16 bits	32 bits
zeros	zeros	IPv4 address

For example `::201.212.11.13`

It was defined to assist in the IPv6 transition but was deprecated because current transition mechanisms no longer use them.

## IPv6 addresses with embedded v4 addresses (2/2)

- IPv4-Mapped IPv6 addresses

80 bits	16 bits	32 bits
zeros	FFFF	IPv4 address

For example

::FFFF:201.212.11.13

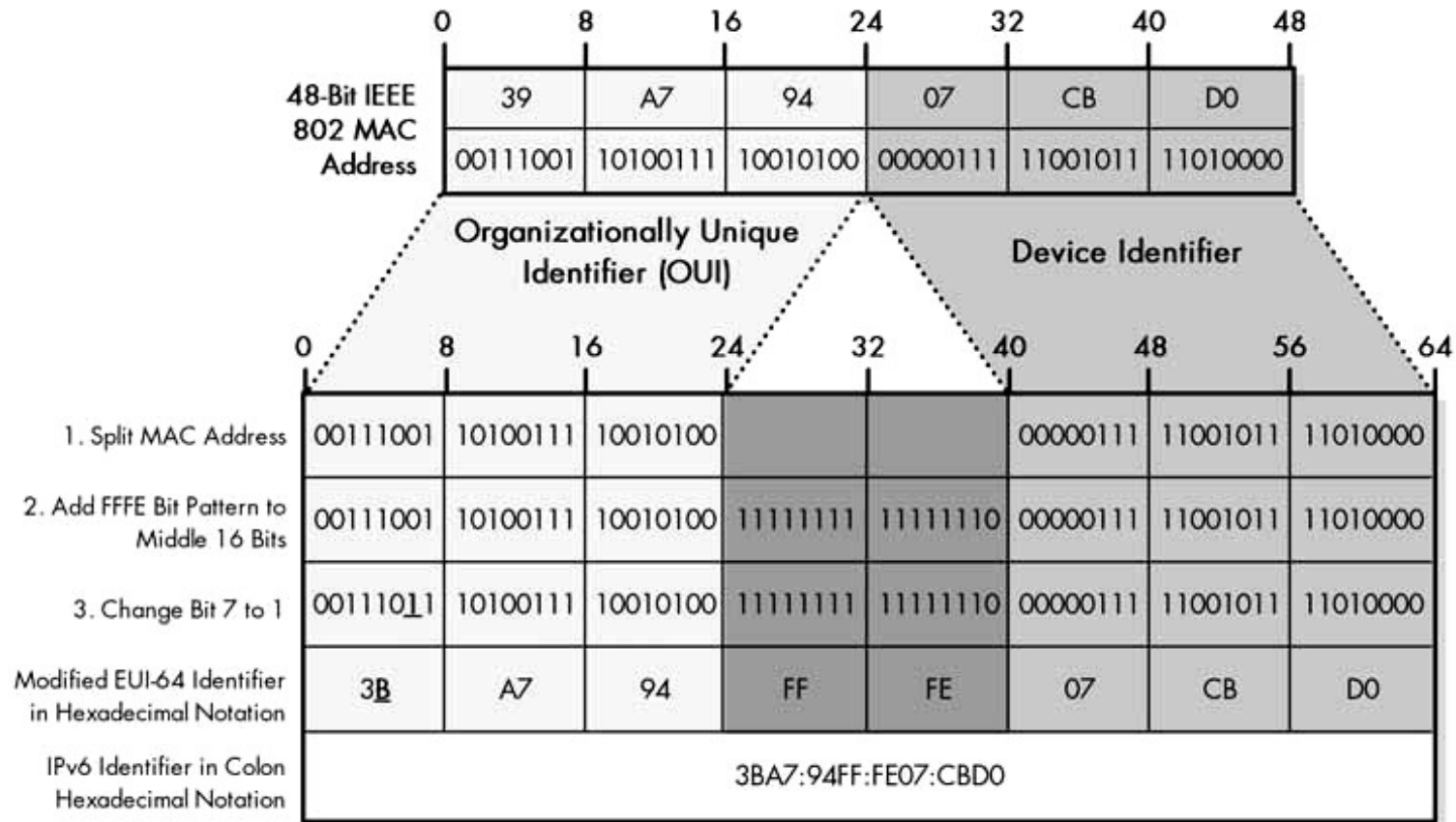
They are used to represent the addresses of IPv4 nodes as IPv6 addresses.

## IPv6 Modified EUI-64 Format 1/2

---

- better way of mapping IP unicast addresses and physical network addresses
  - IEEE has also defined a format called the 64-bit extended unique identifier, abbreviated EUI-64.
  - To get the modified EUI-64 interface ID for a device, you simply take the EUI-64 address and change the 7th bit from the left (the “universal/local” or “U/L” bit) from a zero to a one
-

# IPv6 Modified EUI-64 Format 2/2



64-Bit IPv6 Modified EUI-64 Interface Identifier

# Link-Local IPv6 unicast addresses

10 bits	54 bits	64 bits
1111111010	0	Interface ID

- Link-Local addresses are for use on a single link for purposes such as automatic address configuration, neighbor discovery, or when no routers are present.
- Routers must not forward any packets with Link-Local source or destination addresses to other links.

## Site-Local IPv6 unicast addresses

10 bits	38 bits	16 bits	64 bits
1111111011	0	Subnet ID	Interface ID

- Site-Local addresses were originally designed to be used for addressing inside of a site without the need for a global prefix.
- Site-local addresses are now deprecated but existing implementations and deployments may continue to use this prefix.

# Local IPv6 Unicast Addresses (ULA)

Local IPv6 unicast addresses, as defined in [ULA], have the following characteristics:

- Globally unique prefix.
- Well known prefix to allow for easy filtering at site boundaries.
- Allows sites to be combined or privately interconnected without creating any address conflicts or requiring renumbering of interfaces using these prefixes.
- Internet Service Provider independent and can be used for communications inside of a site without having any permanent or intermittent Internet connectivity.
- If accidentally leaked outside of a site via routing or DNS, there is no conflict with any other addresses
- In practice, applications may treat these addresses like global scoped addresses

# ULA format

7 bits	1	40 bits	16 bits	64 bits
Prefix	L	Global ID	Subnet ID	Interface ID

Where:-

Prefix                      FC00::/7 prefix to identify Local IPv6 unicast addresses.

L                              Set to 1 if the prefix is locally assigned.  
Set to 0 may be defined in the future.

Global ID                    40-bit global identifier used to create a globally unique  
prefix.

Subnet ID                    16-bit Subnet ID is an identifier of a subnet within the  
site.

Interface ID                64-bit Interface ID.

# Advantages of ULA

- Provides Local IPv6 prefixes that can be used independently of any provider-based IPv6 unicast address allocations.
- Applications can treat these addresses in an identical manner as any other type of global IPv6 unicast addresses.
- Sites can be merged without any renumbering of the Local IPv6 addresses.
- Sites can change their provider-based IPv6 unicast address without disrupting any communication that uses Local IPv6 addresses.
- Well-known prefix that allows for easy filtering at site boundary.
- Can be used for inter-site VPNs.
- If accidentally leaked outside of a site via routing or DNS, there is no conflict with any other addresses

## Disadvantages of ULA

---

- Not possible to route Local IPv6 prefixes on the global Internet . Consequentially, it is necessary to filter these addresses.
  - While there is a very low probability of non-unique locally assigned Global IDs, this risk can be ignored for all practical purposes, but it leads to a theoretical risk of clashing address prefixes.
-

# Multicast addresses

8 bits	4 bits	4 bits	112 bits
11111111	flags	scop	Group ID

- Binary 11111111 at the start identifies it as multicast
- Flag is a set of 4 flags. The first bit is reserved and must be 0. The last bit can be 0 indicating a permanently assigned multicast address by IANA or 1 indicating a non-permanently-assigned or transient multicast address.

## Pre-defined multicast addresses

---

- All Nodes addresses - group of all IPv6 nodes, within scope 1 (interface-local) or 2 (link-local)
    - FF01::1
    - FF02::1
  - All routers addresses - group of all IPv6 routers, within scope 1 (interface-local), 2 (link-local), or 5 (site-local)
    - FF01::2
    - FF02::2
    - FF05::2
-

- scop is a 4-bit multicast scope value used to limit the scope of the multicast group.
  - 0 reserved
  - 1 Interface-Local scope
  - 2 Link-Local scope
  - 3 reserved
  - 4 Admin-Local scope
  - 5 Site-Local scope
  - 6 (unassigned)
  - 7 (unassigned)
  - 8 Organization-Local scope
  - 9 (unassigned)
  - A (unassigned)
  - B (unassigned)
  - C (unassigned)
  - D (unassigned)
  - E Global scope
  - F reserved

# Contents

---

Introduction

The IPv6 datagram

Addressing

**Transition Mechanisms**

Applications

---

# Transition

---

IPv6 was designed, at the beginning, with transition in mind: No D-day!

Basic mechanisms:

- Dual stack host: IPv6 and IPv4 co-exist in the network devices, e.g. router or end-system
  - Tunneling: IPv6(/4) traffic is encapsulated into IPv4(/6)
  - Translation (Rewrite packet headers)
-

## Dual Stack

---

- End-systems and routers support both protocols
  - Routers have to support two routing tables
  - Security has to be applied into both protocols
  - Applications chooses which protocol to use
    - IPv6 is usually selected first.
    - If there is an “IPv6 connectivity problem”, IPv4 protocols is used after a while!
-

# Tunnelling

---

- Tunnelling allows IPv6 connectivity through IPv4-only networks
- IPv6 traffic is encapsulated into IPv4 packets

Data is carried through that tunnel using a process called encapsulation, in which the IPv6 packet is carried inside an IPv4 packet, which makes IPv4 appear as a Data Link Layer with respect to IPv6 packet transport. The encapsulating IPv4 header is created at the entry point of the tunnel, and then removed at the exit point of the tunnel

---

## Tunnel Broker (RFC 3053)

---

- Semi manual installation of tunnels using a server.

A tunnel broker is a service which provides a network tunnel. These tunnels can provide encapsulated connectivity over existing infrastructure to a new infrastructure. Its use is defined in RFC 3053.

---

## 6to4 (RFC 3056)

---

- Automatic tunnels using special 2002::/16 addresses.
  - This process allows IPv6 sites to communicate with each other via an IPv4 network without using explicit tunnels, and for these sites to communicate with native IPv6 domains via relay routers.
  - An IPv4 anycast address 192.88.99.1 is used to reach the nearest 6 to 4 relay router.
-

## 6 to 4 (2/2)

### The main advantages are:

- No need to register anything, if you have an IPv4 address then you also have IPv6 6to4 addresses;
- Traffic between separate 6to4 sites takes the most direct route possible. (lower latency)

### The main disadvantages are:

- If you only have a dynamic IPv4 address then your IPv6 6to4 addresses will also be dynamic.
- There is currently no support for setting reverse DNS entries when using 6to4 addresses.
- The tunnelled IPv6 packets may arrive from any IPv4 addresses and therefore filtering becomes both more difficult and more important.

## Teredo (RFC 4380)

---

A proposed enhancement to the 6to4 method is another automatic tunnelling technique called Teredo, which is supported by Microsoft, and defined. Teredo enables nodes that are located behind an IPv4 Network Address Translation (NAT) device to tunnel packets using the User Datagram Protocol (UDP), and thus obtain IPv6 connectivity. Teredo requires the use of server and relay elements to assist with path connectivity.

---

## 6over4 (RFC 2529)

---

- “Virtual Ethernet”. multicast is required.
  - 6over4 defines a method for generating a link-local IPv6 address from an IPv4 address, and a mechanism to perform Neighbor Discovery on top of IPv4 using multicast.
-

## ISATAP (RFC 4214)

- considered an experimental protocol.  
ISATAP is used to connect IPv6 hosts and routers over an IPv4 network using a process that views the IPv4 network as a link layer for IPv6, and other nodes on the network as potential IPv6 hosts or routers. This process acts as a host-to-host, host-to-router, or router-to host automatic tunnel.
- In practice it solves the same problem as 6over4, but doesn't require IPv4 multicast.

## Translation (1/2)

---

The "translators" are equipment able to ensure the translation of traffic IPv4 towards IPv6 and vice versa.

- Supposed to eliminate the need for double stack
  - Solution of last resort, because the translation interferes with the end to end security
-

## Translation (2/2)

---

- NAT-PT (RFC 2766) was designed by the NGtrans working group and has been moved to historic by RFC 4966 for various reasons.
  - But because it is now clear that complete transition to IPv6 will not be completed before the exhaustion of the IPv4, many scenarios of IPv6 only nodes communicating with IPv4 only nodes are being envisaged, creating need for translators. There are many proposals of modified versions of NAT-PT that are on the table now.
-

# Contents

---

Introduction

The IPv6 datagram

Addressing

Transition Mechanisms

**Applications**

---

# Applications

- Apache Web Server

In /etc/httpd/conf/httpd.conf

```
Listen [2001:42d0::200:80:1]:80
```

```
NameVirtualHost [2001:42d0::200:80:1]
```

```
<VirtualHost [2001:42d0::200:80:1]>
```

```
    ServerAdmin webmaster@afrinic.net
```

```
    DocumentRoot /var/www/html/afrinic.net.new
```

```
    ServerName www.afrinic.net
```

```
    ErrorLog logs/www.afrinic.net-error_log
```

```
    CustomLog logs/www.afrinic.net-access_log common
```

```
</VirtualHost>
```

# Applications

## ■ Sendmail

```
DAEMON_OPTIONS('Name=IPv4, Family=inet')dnl
```

```
DAEMON_OPTIONS('Name=IPv6, Family=inet6')dnl
```

In the config files for mailertable, access and relay-domains, v6 addresses are added by prefixing them with the keyword IPv6:

E.g.

```
IPV6:2001:42d0: :200:80:1
```

# Applications

---

- Postfix

In main.cf:-

```
Inet_protocols=all
```

```
smtp_bind_address6= 2001:42d0::200:80:1
```

```
Mynetworks=[2001:42d0::200:80:0]/64
```

---

# Applications

---

- SSH

Listen address 2001:42d0::200:80:1

Listen address [2001:42d0::200:80:1]:2345

---

# Applications

---

- BIND9

```
listen-on-v6 { ::1 ; 2001:4f8:feec::1; };
```

or

```
listen-on-v6 { any; };
```

---

# Configuring radvd

```
interface eth0
{
AdvSendAdvert on;
MinRtrAdvInterval 3;
MaxRtrAdvInterval 10;
AdvHomeAgentFlag off;
# example of a standard prefix
prefix 2001:4f8:feec::/64{
    AdvOnLink on;
    AdvAutonomous on;
    AdvRouterAddr on;
    AdvPreferredLifetime 604800; # 7days
    AdvValidLifetime 259200; #30 days
};
};
```

# IPv6 Transition Problems

---

- Network built on NAT is fine and there is no real need to look at IPv6.
  - Cannot see the immediate ‘cost versus benefit’ of implementing IPv6.
  - Limited number of applications that are IPv6 enabled.
  - Lack of educational material from service providers, equipment vendors, and R&E institutions.
  - Poor government/regulatory support and awareness.
  - IPv6 is about 90% complete work still needed on: PPP, DNS standards, multicast/anycast, etc.
  - Network renumbering, service disruption and other operational issues.
-

## References

---

- RFC 1752
  - RFC 2460
  - RFC 4291
  - RFC 4862
  - RFC 4193
  - RFC 4443
  - [tools.ietf.org/wg/ipv6](https://tools.ietf.org/wg/ipv6)
  - [tools.ietf.org/wg/v6ops](https://tools.ietf.org/wg/v6ops)
-