Intro to IPv6

Hari Kurup

AfriNIC

AfNOG-8 May 26th 2008

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:-

4. imminent exhaustion of address space
5. 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.

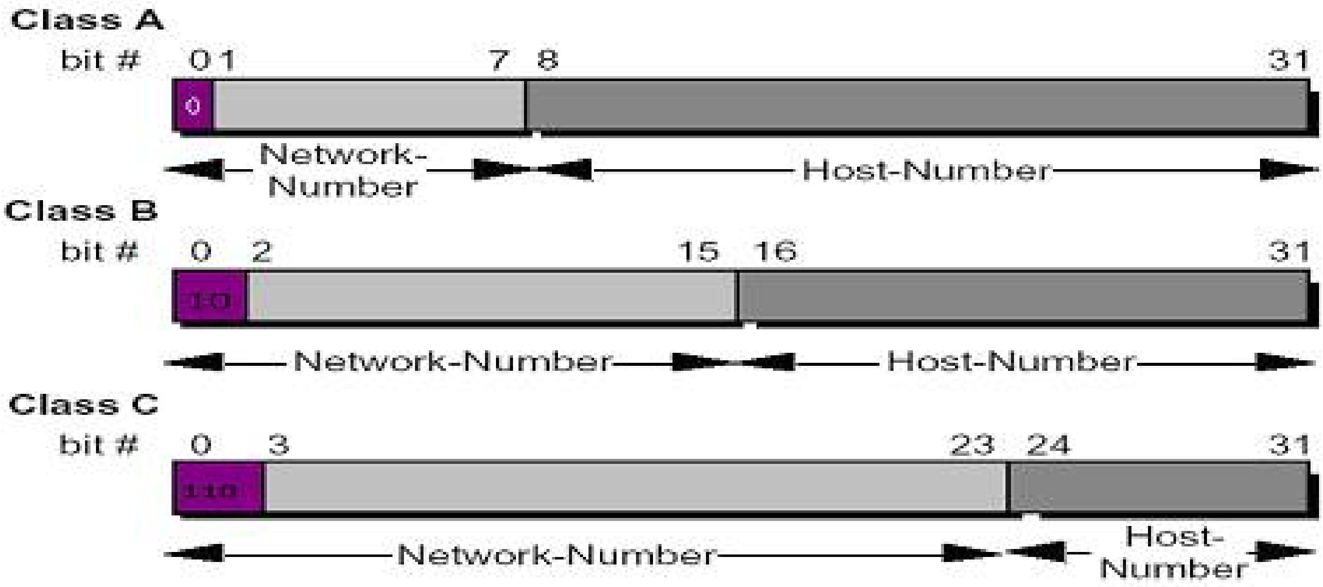
IPv4 space utilization:

As of May 2008, only 41 (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)

IPV6

What happened to IPv5?

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

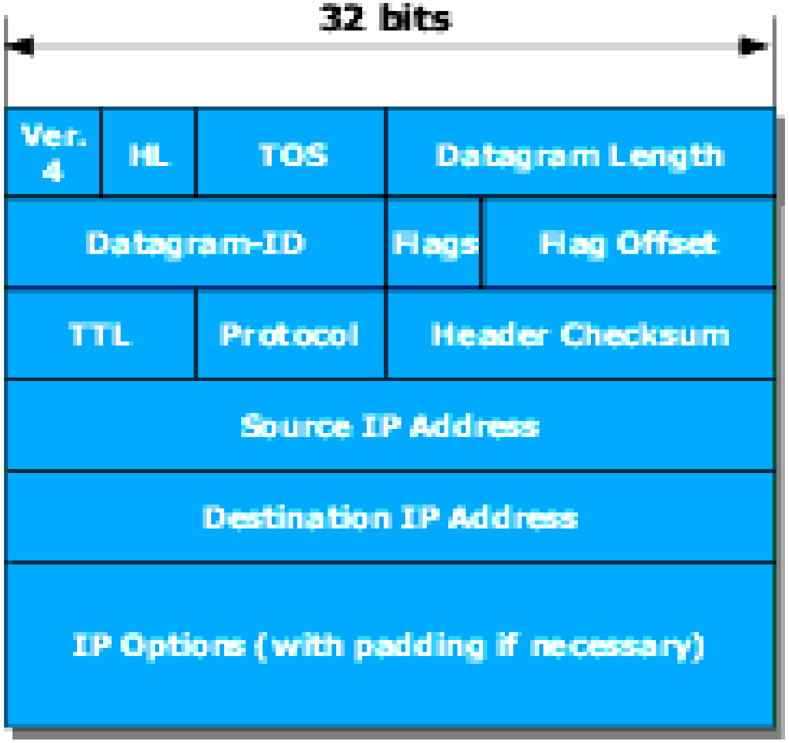
- Experimental non-IP real-time streaming protocol
- Never widely used

- RFC 1819

IPv6 Features

- Large Address Space
- Scoped Addresses
- Auto-configuration
- IPv6 has a new (fixed) 40-bytes header
- Multicast (no broadcast)
- Improved Mobility Support
- Support for IPsec

IPv6 Datagram



32 bits

Ver. 4 | HL | TOS | Datagram Length

Datagram-ID | Flags | Flag Offset

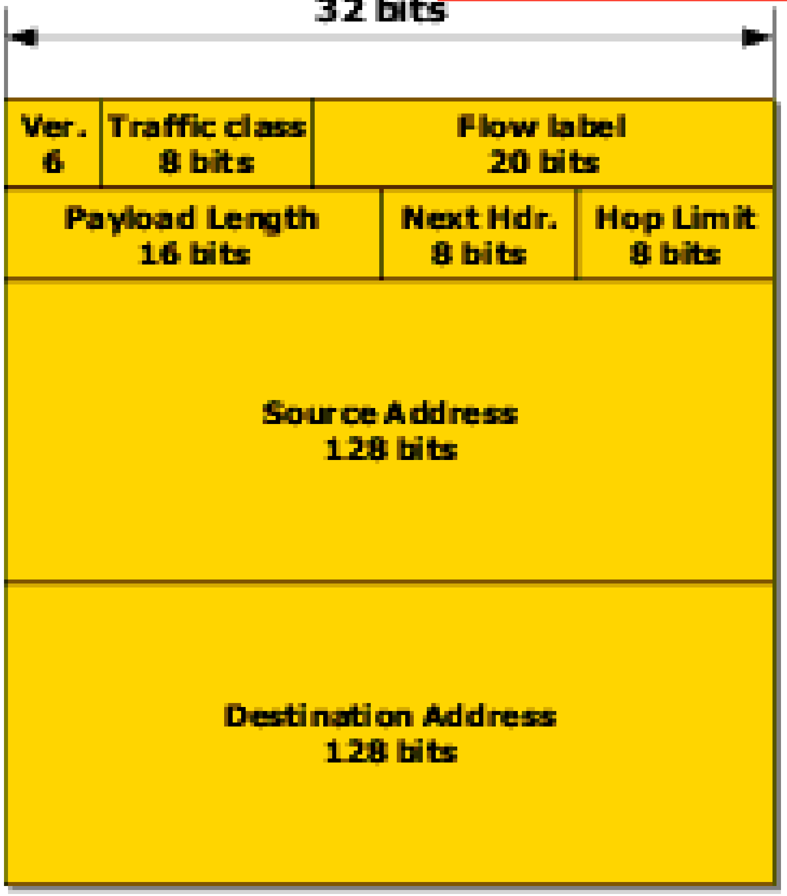
TTL | Protocol | Header Checksum

Source IP Address

Destination IP Address

IP Options (with padding if necessary)

IPv4 header



32 bits

Ver. 6 | Traffic class 8 bits | Flow label 20 bits

Payload Length 16 bits | Next Hdr. 8 bits | Hop Limit 8 bits

Source Address 128 bits

Destination Address 128 bits

IPv6 header

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.

ICMPv6

- 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

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]

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 ::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.
- A packet   received on an interface with a destination address of loopback must be dropped.

Global Unicast

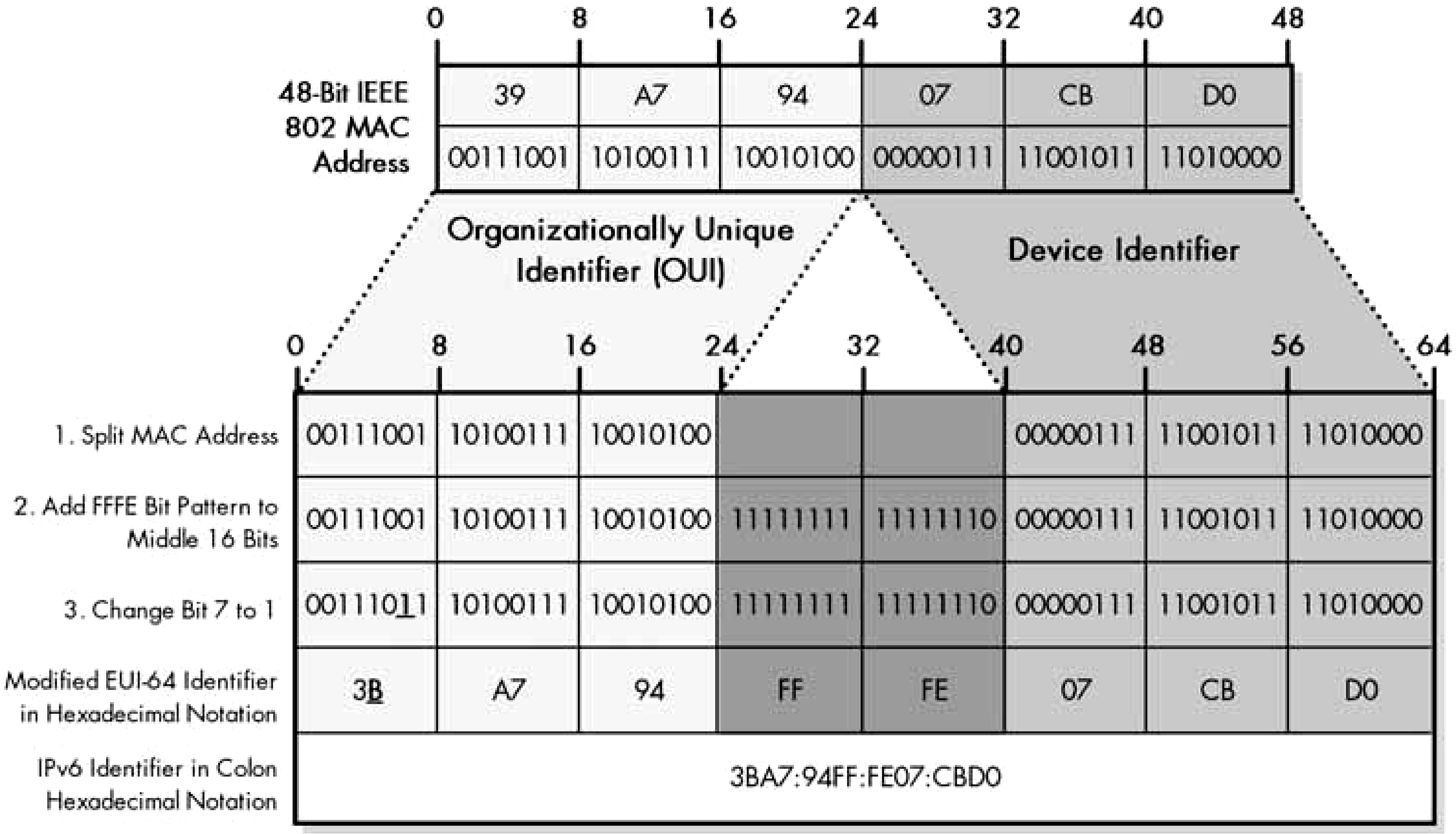
n	m	64
Prefix	Subnet ID	Interface ID

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

IPv6 Modified EUI-64 Format 1/2

- A 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



48-Bit IEEE 802 MAC Address	39	A7	94	07	CB	D0
	00111001	10100111	10010100	00000111	11001011	11010000

1. Split MAC Address	00111001	10100111	10010100			00000111	11001011	11010000
2. Add FFFE Bit Pattern to Middle 16 Bits	00111001	10100111	10010100	11111111	11111110	00000111	11001011	11010000
3. Change Bit 7 to 1	00111011	10100111	10010100	11111111	11111110	00000111	11001011	11010000
Modified EUI-64 Identifier in Hexadecimal Notation	3B	A7	94	FF	FE	07	CB	D0
IPv6 Identifier in Colon Hexadecimal Notation	3BA7:94FF:FE07:CBD0							

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.

Unique Local Addesses (ULA)

ULA addresses 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.

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

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. Its use is defined in RFC 3053.

Free tunnel providers:-

Hurricane Electric http://www.tunnelbroker.net/

SixXS http://www.sixxs.net/

Hexago/go6 http://www.go6.net/

6to4 (RFC 3056)

- Automatic tunnels using special 2002::/16 addresses.
- 2002:<ipv4 external address in hex>::/48
- 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.

Teredo (RFC 4380)

- Teredo provides IPv4 NAT traversal capabilities by tunneling IPv6 over the top of IPv4 using UDP
- Teredo provides IPv6 connectivity when behind an Internet IPv4 NAT device
- Is designed to be a universal method for NAT traversal for most types of NAT used

Applications

- Apache Web Server (v2)

In /etc/httpd/conf/httpd.conf

```
Listen [2001:42d0::2:20:1]:80
NameVirtualHost [2001:42d0::2:20:1]
<VirtualHost [2001:42d0::2:20:1]>
   ServerAdmin webmaster@mydomain.net
   DocumentRoot /usr/local/var/www/data/mydomain
   ServerName www.mydomain.net
   ErrorLog logs/www.mydomain-error_log
   CustomLog logs/www.mydomain-access_log common
</VirtualHost>
```

Applications

- Sendmail (v8.10 and above)

In sendmail.mc

```
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 (v2.2 and above)

In main.cf:-

Inet_protocols=all

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

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

Applications

Exim:-

o Support since version 4.30

o HAVE_IPV6=yes in Local/Makefile at build-time

o Change config literal separator from ‘:’ to, e.g. ‘;’

+ Otherwise parsing literal addresses is tricky.

```
local_interfaces = <; 127.0.0.1 ; \
192.0.2.3 ; \
::1 ; \
2001:db8:1::25
```

Applications

- SSH

Listen address 2001:42d0::1:20:1

Listen address [2001:42d0::1:20:1]:2345

Applications

- BIND9

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

or

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

Useful websites

Some RFCs for your bedtime reading ;-)

- http://www.ietf.org/html.charters/OLD/ipv6-charter.html