IP Next Generation (IPv6)

Shivkumar Kalyanaraman
Rensselaer Polytechnic Institute
shivkuma@ecse.rpi.edu
http://www.ecse.rpi.edu/Homepages/shivkuma

- **Limitations of current Internet Protocol (IP)**
- **How many addresses do we need?**
- **IPv6 Addressing**
- **IPv6 header format**
- **IPv6 features: routing flexibility, plug-n-play, multicast support, flows**

IP Addresses

- Example**: 164.107.134.5**
  = 1010 0100 : 0110 1011 : 1000 0110 : 0000 0101
  = A4:6B:86:05 (32 bits)
- **Maximum number of address = 2^{32} = 4 Billion**
- **Class A Networks: 15 Million nodes**
- **Class B Networks: 64,000 nodes or less**
- **Class C Networks: 250 nodes or less**

IP Address Format

- Three all-zero network numbers are reserved
- 127 Class A + 16,381 Class B + 2,097,151 Class C networks = 2,113,659 networks total
- Class B is most popular.
- 20% of Class B were assigned by 7/90 and doubling every 14 months ⇒ Will exhaust by 3/94
- Question: Estimate how big will you become?
  Answer: More than 256!
  Class C is too small. Class B is just right.

How Many Addresses?

- 10 Billion people by 2020
- Each person has more than one computer
- Assuming 100 computers per person ⇒ 10^{12} computers
- More addresses may be required since
  - Multiple interfaces per node
  - Multiple addresses per interface
- Some believe 2^6 to 2^8 addresses per host
- Safety margin ⇒ 10^{15} addresses
- IPng Requirements ⇒ 10^{12} end systems and 10^9 networks. Desirable 10^{12} to 10^{15} networks

Address Size

- H Ratio = $\log_{10}$(# of objects)/available bits
- 2^n objects with n bits: H-Ratio = $\log_{10}2$ = 0.30103
- French telephone moved from 8 to 9 digits at 10^7 households ⇒ H = 0.26 (~3.3 bits/digit)
- US telephone expanded area codes with 10^8 subscribers ⇒ H = 0.24
- Physics/space science net stopped at 15000 nodes using 16-bit addresses ⇒ H = 0.26
- 3 Million Internet hosts currently using 32-bit addresses ⇒ H = 0.20 ⇒ A few more years to go

IPv6 Addresses

- 128-bit long. Fixed size
- $2^{128} = 3.4 \times 10^{38}$ addresses
  $\Rightarrow 665 \times 10^{21}$ addresses per sq. m of earth surface
- If assigned at the rate of $10^6/\mu s$, it would take 20 years
- Expected to support 8×10^{17} to 2×10^{33} addresses
  $8 \times 10^{17} \Rightarrow$ 1,564 address per sq. m
- Allows multiple interfaces per host.
- Allows multiple addresses per interface
- Allows unicast, multicast, anycast
- Allows provider based, site-local, link-local
- 85% of the space is unassigned

Colon-Hex Notation

- Dot-Decimal: **127.23.45.88**
- Colon-Hex:
  FEDC:0000:0000:0000:3243:0000:0000:ABCD
  - Can skip leading zeros of each word
  - Can skip one sequence of zero words, e.g., FEDC::3243:0000:0000:ABCD *or* ::3243:0000:0000:ABCD
  - Can leave the last 32 bits in dot-decimal, e.g., ::127.23.45.88
  - Can specify a prefix by /length, e.g., 2345:BA23:7::/40

Header

- **IPv6:**

Version	Class	Flow Label
Payload Length	Next Header	Hop Limit
Source Address		
Destination Address		

- IPv4:

Version	IHL	Type of Service	Total Length
Identification		Flags	Fragment Offset
Time to Live	Protocol	Header Checksum	
Source Address			
Destination Address			
Options			Padding

IPv6 vs IPv4

- IPv6 only twice the size of IPv4 header
- Only version number has the same position and meaning as in IPv4
- *Removed*:
  - Header length, fragmentation fields (identification, flags, fragment offset), header checksum
- *Replaced*:
  - Datagram length by payload length
  - Protocol type by next header
  - Time to live by hop limit
  - Type of service by "class" octet
- *Added*: flow label
- All fixed size fields.

IPv6 vs IPv4

- No optional fields. Replaced by extension headers.
  - Idea: avoid unnecessary processing by intermediate routers while not sacrificing the flexibility possible due to options
  - Next Header = 6 (TCP), 17 (UDP), etc

Extension Headers

Base Header	Extension Header 1		Extension Header *n*	Data

- Most extension headers are examined only at destination
- Routing: Loose or tight source routing
- Fragmentation: All IPv6 routers can carry 536 Byte payload
- Authentication
- Hop-by-Hop Options
- Destination Options:

Extension Header (Cont)

- Only Base Header:

Base Header Next = TCP	TCP Segment

- Only Base Header and One Extension Header:

Base Header Next = TCP	Route Header Next = TCP	TCP Segment

- Only Base Header and Two Extension Headers:

Base Header Next = TCP	Route Header Next = Auth	Auth Header Next = TCP	TCP Segment

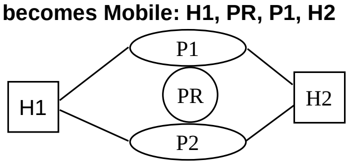
Routing Header

Next Header	Hdr Ext Len	Routing Type	Sgmts left
Reserved			
Address 1			
Address 2			
Address n			

- **Type = 0 ⇒ Current source routing**
- **Router will look at RH if its address is in the destination field**
- **New Functionality: Provider selection, Host mobility, Auto-readdressing (route to new address)**

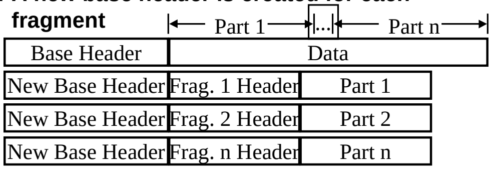
Provider Selection

- **Possible using routing extension header**
- **Source specified intermediate systems**
- **No preference: H1, H2**
- **P1 Preferred: H1, P1, H2**
- **H1 becomes Mobile: H1, PR, P1, H2**



Fragmentation

- Routers cannot fragment. Only source hosts can.
  ⇒ Need path MTU discovery or tunneling
- Fragmentation requires an extension header
- Payload is divided into pieces
- A new base header is created for each fragment



IPv6 addressing and routing

- Aggregatable Global Unicast Addresses
- Link-local and Site-local addresses
- Multicast and Anycast support
- Provider-based inter-domain routing & IDRP

Initial IPv6 Prefix Allocation

Allocation	Prefix	Allocation	Prefix
Reserved	0000 0000	Unassigned	101
Unassigned	0000 0001	Unassigned	110
NSAP	0000 001	Unassigned	1110
IPX	0000 010	Unassigned	1111 0
Unassigned	0000 011	Unassigned	1111 10
Unassigned	0000 1	Unassigned	1111 110
Unassigned	0001	Unassigned	1111 1110
Unassigned	001	Unassigned	1111 1110 0
Provider-based*	010	Link-Local	1111 1110 10
Unassigned	011	Site-Local	1111 1110 11
Geographic	100	Multicast	1111 1111

*Has been renamed as “Aggregatable global unicast”

Aggregatable Global Unicast Addresses

- Address allocation:“provider-based” plan
- Format: TLA + NLA + SLA + 64-bit interface ID
- TLA = “Top level aggregator.” (13 bits)
  - Ranges of TLA values allocated to various registries
  - For “backbone” providers or “exchange points”
- NLA = “Next Level Aggregator” (32 bits)
  - Second tier provider and a subscriber
  - More levels of hierarchy possible within NLA
- SLA = “Site level aggregator” = 16 bits for link

Aggr. Global Unicast Addrs

  - Renumbering after change of provider => change the TLA and NLA. But have same SLA & I/f ID
- Interface ID = 64 bits
  - Will be based on IEEE EUI-64 format
  - An extension of the IEEE 802 (48 bit) format.
  - Possible to derive the IEEE EUI-64 equivalent of current IEEE 802 addresses
  - Along with neighbor discovery procedures, obviates need for ARP.

Local-Use Addresses

- Link Local: Not forwarded outside the link, FE:80::xxx

10 bits	n bits	118-n
1111 1110 10	0	Interface ID

  - Site Local: Not forwarded outside the site, FE:C0::xxx

10 bits	n bits	m bits	118-n-m bits
1111 1110 11	0	Subnet ID	Interface ID

  - Provides plug and play

Multicast Addresses

8 bits	4 bits	4 bits	112 bits
1111 1111	Flags	Scope	Group ID
	0 0 0 T		

- All routers recognize this format => all routers can route multicast packets. Also IGMP part of ICMPv6 => required.
- Flags: T = 0 ⇒ Permanent (well-known) multicast address, 1 ⇒ Transient
- Scope: 1 Node-local, 2 Link-local, 5 Site-local, 8 Organization-local, E Global => routers reqd to honor this.
- Predefined: 1 ⇒ All nodes, 2 ⇒ Routers, 1:0 ⇒ DHCP servers

Multicast & Anycast

- Example: 43 ⇒ NTP Servers
  - FF01::43 ⇒ All NTP servers on this node
  - FF02::43 ⇒ All NTP servers on this link
  - FF05::43 ⇒ All NTP servers in this site
  - FF08::43 ⇒ All NTP servers in this org.
  - FF0F::43 ⇒ All NTP servers in the Internet
- Structure of Group ID:
  - First 80 bits = zero (to avoid risk of group collision, because IP multicast mapping uses only 32 bits)

Inter-domain routing

- CIDR supports aggregation using address allocation based on geographical constraints
- Required: aggregation to be correlated to topology structure to allow routing
- Future: providers will define routing topology => aggregation should be based on providers
- IPv6: allows both provider-based (aggregatable global unicast addrs) and geographic address allocation
- Multiple addresses/interface => free customers from providers. Cost of switching is small.

Inter-domain routing (contd)

- BGP-4 too much optimized for 32-bits
- Inter-domain routing protocol (IDRP) from the OSI world is the current choice for IPv6.
  - IDRP has a superset of BGP functionalities
  - It does not use TCP => can send new (and modified) routing packets if the old ones do not make it {instead of retransmitting stale information}
  - It uses address-prefixes instead of AS numbers {builds on TLA/NLA and avoids AS assignment by IANA}

Address Autoconfiguration

- Allows plug and play
- BOOTP and DHCP are used in IPv4
- DHCPng will be used with IPv6
- Two Methods: Stateless and Stateful
- Stateless:
  - A system uses link-local address as source and multicasts to "All routers on this link"
  - Router replies and provides all the needed prefix info

Address Autoconfiguration

  - All prefixes have a associated lifetime
  - System can use link-local address permanently if no router
- Stateful:
  - Problem w stateless: Anyone can connect
  - Routers ask the new system to go DHCP server (by setting managed configuration bit)
  - System multicasts to "All DHCP servers"
  - DHCP server assigns an address

Neighbor Discovery

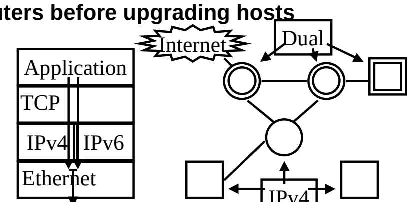
- Media addresses allowed to be upto 128 bits long
- Part of ICMPv6 functionality
- Subsumes ARP, Router discovery.
- Source maintains several caches: destination cache, neighbor cache, prefix cache, router cache
- Multicast solicitation for neighbor media address if unavailable in neighbor cache
- Neighbor advertisement message sent to soliciting station.
- Redirects also part of ICMPv6

Real-time support

- Flow label and the "class" octet field
- Flow = sequence of packets from a single source to a particular (unicast/multicast) destinations requiring special handling by intermediate routers
- Applications becoming adaptive
  - Even adaptive voice available for IP telephony
- Hierarchical transmissions:
  - Can cause congestion {Steve McCanne, SIGCOMM'96} =>"priority" renamed as "class"
  - "Class" field currently being worked upon by differentiated services group

Transition Mechanisms

- Dual-IP Hosts, Routers, Name servers
- Tunneling IPv6 over IPv4
- Nodes can be partially upgraded to IPv6
- It is better (though not required) to upgrade routers before upgrading hosts



Application Issues

- Most application protocols will have to be upgraded: FTP, SMTP, Telnet, Rlogin
- 27 of 51 Full Internet standards, 6 of 20 draft standards, 25 of 130 proposed standards will be revised for IPv6
- No checksum ⇒ checksum at upper layer is mandatory, even in UDP
- non-IETF standards: X-Open, Kerberos, ... will be updated
- Should be able to request and receive new DNS records

Implementations

- 4.4-lite BSD by US Naval Research Laboratory
- UNIX, OPEN-VMS by DEC
- DOS/WINDOWS by FTP Software
- HP-UX SICS (Swedish Institute of Comp. Science)
- Linux
- NetBSD by INRIA Rocquencourt
- Solaris 2 by Sun
- Streams by Mentat

Summary

- IPv6 uses 128-bit addresses
- Allows provider-based, site-local, link-local, multicast, anycast addresses
- Fixed header size. Extension headers instead of options for provider selection, security etc
- Allows auto-configuration
- Dual-IP router and host implementations for transition

IPng: Key References

- C. Huitema, "IPv6: The New Internet Protocol," Prentice-Hall, 1998, 247 pp.
- IP Next Generation, http://playground.sun.com/pub/ipng/html/ipng-implementations.html
- IP: Next Generation, http://www.cnri.reston.va.us/ipng/ipng.html
- 6-bone:
  http://www-6bone.lbl.gov/6bone/