IPv6 (Internet Protocol version 6)

Kazu Yamamoto
IIJ Research Laboratory
kazu@iijlab.net

Contents

- **What's IPv6**
- **Address Architecture**
- **Plug & play**
- **Domain Name System**
- **Transition**
- **The current status of IPv6**

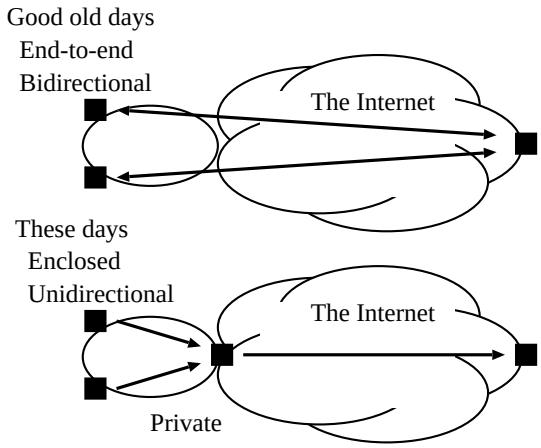
=> What's IPv6

Address Architecture
Plug & play
Domain Name System
Transition
The current status of IPv6

Problems of IPv4

- **Exhaustion of IPv4 addresses**
  - **4 bytes = 4.3 billion**
    - **Much less than the human population (6.1 billion)**
  - **Will be exhausted in around 2008**
  - **Registries are allocating IPv4 addresses by severe policy**
    - **Nobody can obtain enough IPv4 addresses**
- **Increment of routing information**
  - **Routing information cannot be aggregated effectively**
    - **Unaggregatable address assignments**
  - **80,000 entries at present**
  - **Burden for backbone routers**
    - **Unstability, accidents**
- **Proliferation of NAT**
  - **Breaking the Internet architecture**
  - **Enclosure of users**

The Internet architecture(1)



The Internet architecture(2)

- **Flat Internet**
  - **Bidirectional communication**
  - **End-to-end communication**
    - **True communication infrastructure**
    - **Much easier to deploy inventive new applications**

- **Patched Internet by NAT**
  - **Unidirectional communication**
  - **Enclosed communication**
    - **Single point of failure**
    - **Evolution of application are suppressed by NAT**
    - **Accounting from servers is impossible**

IPv6

- **Address extension**
  - **16 bytes = 3.4 x 10^38**
  - **65,536 subnets (minimum) for everybody (/48)**
    - **Class A (in the old term) per site**
- **Starting over**
  - **Some technologies are mandatory**
    - **Plug and play**
    - **End-to-end security (i.e. IPsec)**
  - **Aggregatable global addresses from the start**
    - **Reducing external routing information to 8,192**
- **Paradigm chage for applications**
  - **End-to-end and bidirectional communication**
  - **IPv6 is a NAT-free world**
    - **Cellular phone, automobiles, home network, game machines, ...**

What's IPv6

=> Address Architecture

Plug & play
Domain Name System
Transition
The current status of IPv6

Address notation

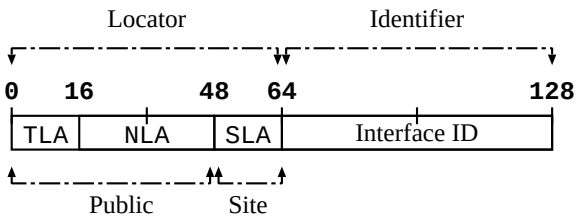
- **Separate 4 figures of hexadecimal by ":"**
  - **ff02:0000:0000:0000:0000:0000:0000:0001**
  - **3ffe:0501:0008:1234:0260:97ff:fe40:efab**
- **Preceding 0 for each piece can be omitted**
  - **ff02:0:0:0:0:0:0:1**
  - **3ffe:501:8:1234:260:97ff:fe40:efab**
- **Continuous 0 piece can be expressed by "::" (at most once)**
  - **ff02::1**
- **Prefix length (0 - 128) is placed after "/"**
  - **3ffe:500::/24**

Address block

- **The address space of IPv6 is divided into 8 blocks (/3)**
  - **3 bits is not friendly to hexadecimal**
  - **Let's consider /4 (i.e. 16 blocks)**
- **The first hexadecimal figure:**
  - **0, 1 Special (e.g loopback)**
  - **2, 3 Global address (aggregatable global address)**
  - **4, 5 Not assigned**
  - **6, 7 Not assigned**
  - **8, 9 Not assigned**
  - **a, b Not assigned**
  - **c, d Not assigned**
  - **e, f Link-local, site-local, multicast**

Global address

- **Aggregatable global address**
  - **3ffe:501:8:1234:260:97ff:fe40:efab**



- **8 byte network part + 8 byte host part**
  - **Prefix length is fixed to /64**
    - **No need to decide prefix length for subnets**
  - **/48 is assigned to a site**
    - **65,536 subnets per a site**

Aggregator

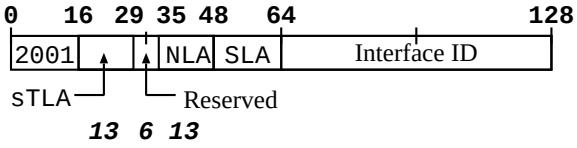
- **TLA (Top Level Aggregator)**
  - **/16**
  - **Big ISPs or IXes**
  - **8,192 (16 - 3 = 13 bit)**
  - **MUST announce /16 routing information to other TLAs**

- **NLA (Next Level Aggregator)**
  - **/17 - /48**
  - **Medium or small ISPs**
    - **NLA1, NLA2,...**
  - **Final NLA is a site**

- **SLA (Site Level Aggregator)**
  - **/49 - /64**
  - **Subnet number in a site**
  - **65,536 subnets per site**

Initial address assignment

- **sTLA (sub TLA)**
  - **Slow start of IPv6 address assignment**
  - **TLA "2001" was divided into 8,192 sub TLA (/35)**
    - **The same number of TLA**
  - **ARIN/RIPE/APNIC have started assigning**
    - **Commercial purpose**
    - **No need to return it to registries in the future**



- **Criteria**
  - **Start IPv6 services within 12 months**
  - **Many IPv4 customers already OR 6 months experience in 6bone**

What's IPv6
Address Architecture

=> Plug & play

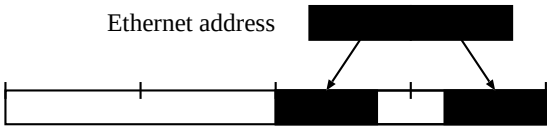
Domain Name System
Transition
The current status of IPv6

Two stories

- **Dilemma of dentists**
  - **Well-educated but a beginner of network**
  - **They believe that computers can start communicating just after they open boxes**
- **Nightmare of network managers**
  - **100 computers arrived on Friday**
  - **They have to get computers ready for the Internet by Monday**

Address auto-configuration

- **Lower 8 bytes from a MAC address**
  - **Link-local address**



- **Upper 8 bytes from a router**
  - **Global address**

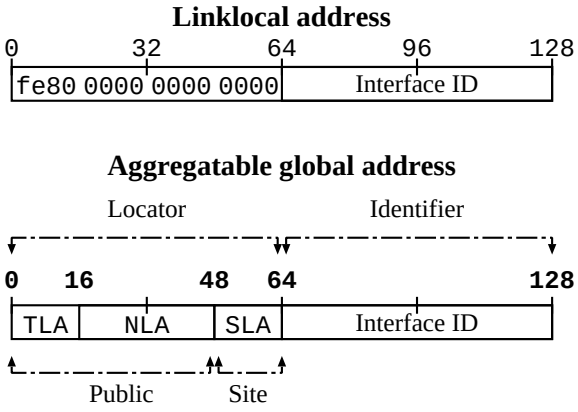
- **DHCP server is not necessary**
  - **Stateless address allocation**

Auto-configuration of link-local

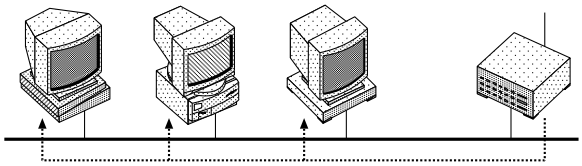
- **MAC address (e.g Ethernet)**
  - **IEEE 802 address (6 bytes)**
    - **00:60:97:40:ef:ab**
- **Converting MAC address to interface ID**
  - **EUI 64 address (8 bytes)**
  - **260:97ff:fe40:efab**
    - **00:60:97 + ff:fe + 40:ef:ab**
    - **Negate one bit**
- **Generating a link-local address**
  - **fe80:: + interface ID**
    - **fe80::260:97ff:fe40:efab**
- **Now IPv6 nodes can communicate within the link!**
  - **Resolving the dilemma of dentists**

Link-local address

- **Unique in a link**
  - **Not necessary unique in a node**
- **fe80:: + interface ID**
  - **fe80::260:97ff:fe40:efab**



Auto-configuration of global



- **Router Advertisement (RA)**
  - **Routers repeatedly announce prefixes (/64)**
- **Generating a global address**
  - **Prefix + interface ID**
- **Default route to one of the router**

- **Now IPv6 nodes can communicate on the Internet!**
  - **Resolving the nightmare of network managers**

Renumbering

- **Precondition**
  - **All IPv6 nodes have address auto-configuration**
  - **Each IPv6 node can obtain multiple IPv6 addresses**
  - **IPv6 addresses have two timers**

- **Switching ISP A to ISP B**
  - **A site is connected to ISP A**
    - **An old address**
  - **Connect to ISP B, then a new prefix is announced**
    - **The old address, a new address**
  - **The first timer is expired**
    - **The old address is not used for further communication**
  - **The second timer is expired**
    - **The new address**
  - **Disconnect the leased to ISP A**

What's IPv6
Address Architecture
Plug & play

=> Domain Name System

Transition
The current status of IPv6

Using IPv6 applications

- Typical users specify "host name" to applications
  - (e.g) ftp ftp.mew.org
- Users are not aware which they are using, IPv4 or IPv6
  - (e.g) "ftp ftp.mew.org" may use IPv4
  - (e.g) "ftp ftp.mew.org" may use IPv6
- Users can specify an IPv6 address to applications, of course
  - (e.g) ftp 3ffe:501:8:1234:260:97ff:fe40:efab
  - (e.g) telnet ::1
- But ":" is unfriendly to some application syntax
  - URL
    - http://[3ffe:501:8:1234:260:97ff:fe40:efab]/

Domain Name System

- **Contents**
  - **AAAA records for forward lookup**

```
$ORIGIN mew.org.
ftp  AAAA  3ffe:501:8:1234:260:97ff:fe40:efab
www  A     133.5.2.1
```

  - **PTR records for reverse lookup**

```
$ORIGIN 4.3.2.1.8.0.0.0.1.0.5.0.e.f.f.3.IP6.INT.
b.a.f.e.0.4.e.f.f.f.7.9.0.6.2.0  PTR  ftp.mew.org.
$ORIGIN 2.5.133.IN-ADDR.ARPA.
1                                PTR  www.mew.org.
```

  - **BIND 4, 8, 9**

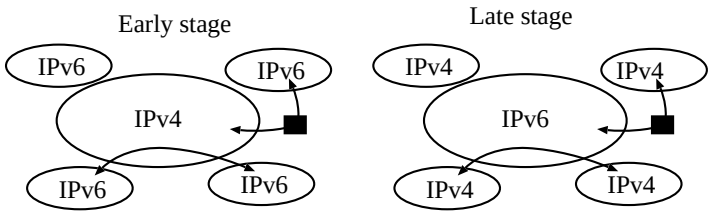
- **Transport**
  - **Both UDP/IPv4 and UDP/IPv6**
  - **BIND 9**

What's IPv6
Address Architecture
Plug & play
Domain Name System

=> Transition

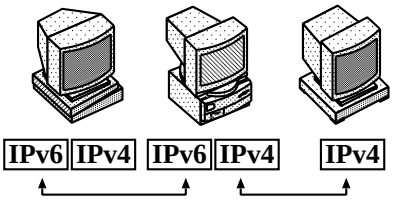
The current status of IPv6

Transition story



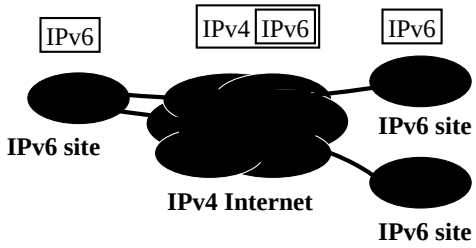
- **Transition technologies**
  - **Dual stack**
  - **Tunnel**
  - **Translator**

Dual stack



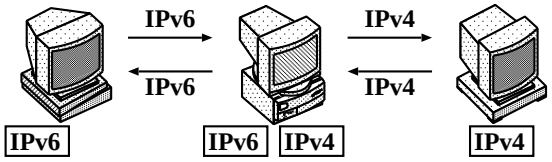
- **Supporting both IPv4 and IPv6**
  - **Dual stack is a MUST in the early stage**
- **BITS (Bump In The Stack)**
  - **Dual stack functionality without OS installation**
    - **Replacement of a driver of IPv4 node**
  - **IPv4 applications are available without modifications**

IPv6 in IPv4 tunnels



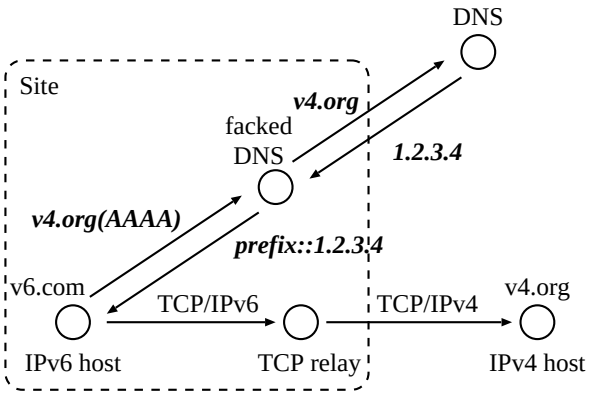
- IPv6 sites are island in the IPv4 ocean
- Connecting IPv6 islands
  - IPv4 as datalink
  - Encapsulating an IPv6 packet into an IPv4 packet
- 6bone is an example
  - About 50 countries

Translator



- The early stage
  - IPv4 nodes exist, of course
  - IPv6 nodes appeared
    - Not enough IPv4 addresses are assigned
- The late stage
  - IPv4 notes will remain
  - IPv6 nodes
    - IPv4 addresses will be unavailable
- Co-existence of IPv4 nodes and IPv6 nodes
  - Translators are necessary

TCP relay(1)



- **Translator in the early stage**
  - **Site is IPv6**
  - **The Internet is IPv4**

TCP relay(2)

- **Precondition**
  - **IPv6 site has also small number of IPv4 addresses**
  - **Both fake DNS and TCP relay are dual stack**
    - **A special prefix is installed on them statically**

- **Connection from v6.com to v4.org**
  - **v6.com asks faked DNS to resolve AAAA record of v4.org**
  - **Faked DNS asks DNS to resolve A record of v4.org**
  - **DNS returns 1.2.3.4**
  - **Faked DNS embeds 1.2.3.4 into special prefix and tells v6.com it**
  - **v6.com tries to make a TCP/IPv6 connection to prefix::1.2.3.4**
  - **TCP relay catches the connection**
  - **TCP relay also make a TCP/IPv4 connection to 1.2.3.4**

What's IPv6
Address Architecture
Plug & play
Domain Name System
Transition

=> The current status of IPv6

Specification and implementation

- Basic specifications have been done
  - Some advanced specifications are being discussed
- Many UNIX systems are IPv6-ready
  - *BSD*, Linux, Solaris 8, DEC(Compaq), IBM...
- Windows 2000 + IE
  - http://msdn.microsoft.com/downloads/sdks/platform/tpipv6.asp
- MacOS X will be ready in this year
  - Based on KAME
- Many open source servers/clients are IPv6 ready
  - Sendmail, BIND, Apache, ...
- Routers
  - Cisco, Hitachi, Fujitsu, NEC, ...

Deployment status

- **Research network**
  - **6bone (spans across 50 countries)**
- **Experimental service of commercial ISP**
  - **IIJ, NTT Com, vBNS, NEC, Fujitsu**
- **IX (Internet Exchange)**
  - **NTT IX, 6TAP, PAIX, NSPIXP6**
- **Commercial service**
  - **IIJ**

KAME Project

- **A single effort**
  - **8 core members from 7 Japanese companies**
  - **Fujitsu, Hitachi, IIJ, NEC, Toshiba, YDC, Yokogawa**
  - **April 1998 - March 2002**
  - **The core members work for IPv6 three days a week**
- **Reference code**
  - **IPv6, IPsec, and advanced networking**
  - **Provided "AS IS" like BSD**
    - **Free and no warranty, commercial use is OK**
- **Adopted**
  - **BSD/OS 4.2, FreeBSD 4.1, NetBSD 1.5 (2000 Autumn),**
  - **OpenBSD 2.7, IIJ SEIL T1, Hitachi GR2000, Fujitsu NetVehicle**
- **TAHI Project**
  - **Conformance test, inter-operability test**

Information

- **URL**
  - `http://www.ipv6.org/`
  - `http://www.6bone.net/`
  - `http://www.v6.wide.ad.jp/`
  - `http://www.kame.net/`
  - `http://www.v6forum.com/`
  - `http://www.freenet6.net/`
  - `http://www.6tap.net/`

- **ML**
  - `6bone@isi.edu`
  - `ngtrans@sunroof.eng.sun.com`
  - `ipng@sunroof.eng.sun.com`
  - `users@ipv6.org`

Global IPv6 Summit

- **Dec 18 - 19, 2000**
- **Osaka, Japan**
- **http://www.jp.ipv6forum.com**

- **Program**
  - **Keynote speech from Jun Murai and Steve Deering**
  - **Business report from Japan**
  - **Business report around the world**
  - **Status report from CN, KR, SG, MY**
  - **Address allocation/assignment session**
  - **Panel on how IPv6 change business**
  - **Panel on how to transit to IPv6**