Internet2
IPv6 Workshop
Addressing

IPv6 Addressing
IPv6 Addressing

- RFC 2173 - 98
- RFC 3513 - 02
- RFC 4291 - 06
Overview of Addressing

• Historical aspects
• Types of IPv6 addresses
• Work-in-progress
• Abilene IPv6 addressing, a v6 network
Historical Aspects of IPv6

• IPv4 address space not big enough
  – Can’t get needed addresses
    (particularly outside the Americas)
  – Routing table issues
  – Resort to private (RFC1918) addresses - NATS

• Competing plans to address problem
  – Some 64-bit, some 128-bit

• Current scheme unveiled at Toronto IETF (July 1994)
Private Address Space

• Led to the development of NAT.
• Increased use of NAT has had an effect on the uses the Internet may be put to.
  – Due to the loss of transparency
• Increasingly could lead to a bifurcation of the Internet.
  – Application rich
  – Application poor
• Affects our ability to manage and diagnose the network – attendees are operators
Types of IPv6 Addresses

• Like IPv4...
  – 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, RFC 1546
    • 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).

• Specified in the v6 address architecture RFC 4291.
IPv6 Address Categories

• Traditional Unicast (one-to-one)

• Multicast (one-to-many)
  – Destination is a group of computers that may reside on many networks
  – Broadcast is just a special case of multicast

• Anycast/Cluster (one-to-nearest)
  – Destination is one from a group of nodes (probably routers or servers)
  – Probably local scope only (requires cooperation of routers)
Classless Addressing

• IPv6 classifies addresses based on a variable-length *Address Type Prefix*
  – Prefix ranges from 3 to 10 bits, depending on Address Type

• Different than IPv4 address classes (A/B/C), similar to CIDR (classless)

• The longest-match algorithm is used to optimize routing table lookup
What is not in IPv6

• Broadcast
  – There is no broadcast in IPv6.
  – This functionality is taken over by multicast.
    • multicast is ‘restricted’ broadcast, removing the ‘all or none’ approach to broadcast, in favor of groups
    • But now must manage groups with Multicast Addresses...
Interface Identifiers

- 64-bit field – New mac address
- Guaranteed unique on subnet
- Essentially same as EUI-64
  - See Appendix A, RFC 4291
- Formula for mapping IEEE 802 MAC address into interface identifier
- Used in many forms of unicast addressing
- Basically: modified Ethernet MAC address expanded to 64 bits which is where Ethernet is going
Interface Identifiers

• IPv6 addresses of all types are assigned to interfaces, not nodes.
  – An IPv6 unicast address refers to a single interface. Since each interface belongs to a single node, any of that node's interfaces' unicast addresses may be used as an identifier for the node – multi-homed machines
Interface Identifiers

• EUI-64 from Mac addresses:
  – 00-02-2D-02-82-34 – 48 bits but moving to 64
  – 0202:2dff:fe02:8234 - 64 bit

• The rules are:
  – Insert $\text{ffe}$ after the first 3 octets – ‘company’
    was encoded in upper octets
  – Last 3 octets remain the same
  – Invert the 2\textsuperscript{nd} to the last low order bit of the first
    octet.
    • Universal/local bit
  – why? – assume this follows move to 64 bit Mac
Interface Identifiers

• Privacy addresses:
  – Some concern was expressed about having one’s MAC address be public - h/w identifier, persistent
  – The response was to standardize privacy addresses (RFC 3041).
  – These use random 64-bit numbers instead of EUI-64.
    • May change for different connections
    • On by default in Windows, off by default in Linux (net.ipv6.conf.all.use_tempaddr), OSX and BSD (net.inet6.ip6.use_tempaddr)
Interface Identifiers

- A host is required to recognize the following addresses as identifying itself:
  - Its link-local address for each interface
  - Assigned unicast and anycast addresses
  - Loopback address
  - All-nodes multicast addresses
  - Solicited-node multicast address for each of its unicast and anycast addresses
  - Multicast addresses of all other groups to which the node belongs
Interface Identifiers

• A router is required to recognize:
  – All addresses it must recognize as a host, plus
  – The subnet-router anycast addresses for the interfaces it is configured to act as a router on
  – All other anycast addresses with which the router has been configured
  – All-routers multicast addresses
Representation of Addresses

- All addresses are 128 bits
- Written as sequence of 8 groups of 4 hex digits (16 bits each) separated by colons
  - Leading zeros in group may be omitted
  - A contiguous all-zero group may be replaced by “::”
    - Only one such group can be replaced
Examples of Writing Addresses

• Consider
  – 3ffe:3700:0200:00ff:0000:0000:0000:0001

• This can be written as
  – 3ffe:3700:200:ff:0:0:0:1 or
  – 3ffe:3700:200:ff::1

• Both reduction methods are used here.
Types of Unicast Addresses

• Unspecified address
  – All zeros (::)
  – Used as source address during initialization
  – Also used in representing default route

• Loopback address
  – Low-order one bit (::1)
  – Same as 127.0.0.1 in IPv4
Types of Unicast Addresses

- Link-local address (private addressing)
  - Unique on a subnet – never routed
  - Auto configured
  - High-order: FE80::/10
  - Low-order: interface identifier
  - Routers must not forward any packets with link-local source or destination addresses.
  - Why? DHCP, creating temporary networks, 2 guys in a bar
Types of Unicast Addresses

• Unique local addresses
  – RFC 4193
  – replacing site-local addresses, which were deprecated in RFC 3879
  – globally unique but used locally within a site
  – fd(8), uniqueid(40), site subnet(16), interfaceid (48)
  – e.g., printer or other devices unknown outside site
  – independent of ISP,
  – unique global prefix, allow filtering at network boundaries
Types of Unicast Addresses

- Mapped IPv4 addresses
  - Of form ::FFFF:a.b.c.d
  - Used by dual-stack machines to communicate over IPv4 using IPv6 addressing in system calls

- Compatible IPv4 addresses
  - Of form ::a.b.c.d
  - Used by IPv6 hosts to communicate over automatic tunnels
Address Deployment

• There have been many discussions of how to make use of the immense IPv6 address space.

• Suggestions included:
  – Provider-Independent (PI)
  – Provider-Assigned (PA)
  – Geographical

• At least for now, PA addressing was selected.
  – It is important to understand the difference between allocation and assignment.
Types of Unicast Addresses

- Aggregatable global unicast address
  - Used in production IPv6 networks
  - Goal: minimize global routing table size
  - From range 2000::/3
  - First 64 bits divided into two parts:
    - Global routing prefix
    - Subnet identifier
Aggregatable Global Unicast Addresses

Global Routing Prefix includes a value assigned to the ISP (used for backbone routing) and a value assigned to the subscriber.

Subnet ID is assigned by that subscriber to a part of its intranet.

Interface ID would probably be a hardware address (e.g. Ethernet) to facilitate easy autoconfiguration.
Internet Registry Hierarchy

- Regional IR - designated by IANA (ARIN, RIPE, APNIC, AfriNIC, LACNIC)
- Local IR - ISP, or other network provider
- RIR -> LIR, LIR -> customer (or smaller provider)

<table>
<thead>
<tr>
<th>ARIN</th>
<th>2001:0400::/23</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abilene</td>
<td>2001:0468::/32</td>
</tr>
<tr>
<td>NYSERNet</td>
<td>2001:0468:0900::/40</td>
</tr>
<tr>
<td>Columbia</td>
<td>2001:0468:0904::/48</td>
</tr>
</tbody>
</table>
Abilene Address Space

• ARIN gave 2001:468::/32 to Abilene
• The bit-level representation of this is:

\[0010 \ 0000 \ 0000 \ 0001 : 0000 \ 0100 \ 0110 \ 1000 ::\]

• This leaves 32 bits of network space available.
• We will see later how this is to be used – serious design question
Current Practice and Aggregation

• The overarching goal of the PA addressing scheme is aggregation.
  – As you move up the provider chain all addresses are aggregated into larger blocks.
  – If implemented completely the result would be a default-free zone with a very small number of prefixes — only those allocated by the RIRs.
Anycast Address

- Interfaces (I > 1) can have the same address. The low-order bits (typically 64 or more) are zero.
- A packet sent to that address will be delivered to the topologically-closest instance of the set of hosts having that address.
- Examples:
  - subnet-router anycast address (RFC 2373)
  - reserved subnet anycast address (RFC 2526)
  - 6to4 relay anycast address (RFC 3068)
Possible Anycast Applications

• Mobile IP - packets sent by a mobile computer could use same default router address, even though it is moving from one router to another

• Service Abstraction - a way to locate an HTTP proxy server or DNS server
  – Do not care what the name or IP address of the server is, just need access to a machine that provides that service
Multicast Address

• From FF00::/8
  – 1111 1111 | flgs (4) | scope (4) | group id (112) |

• Flags
  – 000t
    • t=0 means this is a well-known address, i.e., from set of multicast groups
    • t=1 means this is a transitory address

• Low-order 112 bits are group identifier, not interface identifier

• Scope and Flags are independent of each other
  – Well-known and local is different from well-known and global
Multicast address scope

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
Abilene Allocation Procedures

• **GigaPoPs** allocated /40s
  – Expected to delegate to participants
    • The minimum allocation is a /48
  – No BCP *(yet)* for gigaPoP allocation procedures

• **Direct connectors** allocated /48s
  – Will *(for now)* provide addresses to participants behind gigaPoPs which haven’t received IPv6 addresses
Abilene Registration Procedures

• Providers allocated address space must register suballocations
  – ARIN allows rwhois or SWIP
  – For now, Abilene will use SWIP
  – Will eventually adopt rwhois
  – GigaPoPs must also maintain registries
    • Will probably have central Abilene registry
Obtaining Addresses

• If you are a gigaPoP or a direct connect send a note to the Abilene NOC (noc@abilene.iu.edu) with a request.
  – Will set the wheels in motion

• If you connect to a gigaPoP you should obtain your address block from that gigaPoP—talk to them first
  – Remember the minimum you should receive is a /48.
  – More is OK if you can negotiate for a larger block.

• Note ongoing transition from Abilene to the new Internet2 Network
  – See www.internet2.edu/network/
Allocation Schemes

CIDR representation and IPv6 allocations
IPv4 Subnet Masking

- Originally the network size was based on the first few bits (classful addressing)
- Getting rid of address classes was *painful*!
  - routing protocols, stacks, applications
- Modern IPv4 allows subnet boundaries anywhere within the address (classless addressing)
- But decimal addresses still make figuring out subnets unnecessarily difficult. . .
CIDR

In IPv4 you would see representations like:
- 129.93.0.0/16
- 129.93.0.0 255.255.0.0

At the bit level this is:
- 10000001.01011101.00000000.00000000

*network* . *subnet* or *host*
Reasons for CIDR

• To try to preserve the address space.

• To control the growth of the routing table.
IPv6 Notation

• In IPv6 every address is written:
  – IPv6 address / prefix length
• For example (which is the larger network?)
  – 2001:0468::/35
  – 2001:0468::/32
• At the bit level:
  – 0010 0000 0000 0001: 0000 0100 0110 1000: 000 0::/35
  – 0010 0000 0000 0001: 0000 0100 0110 1000 0::/32
Allocation Strategies Example

• We wish to allocate /48s out of the /35.
• Which are available:
  – 2001:0468:0000 through
  – 2001:0468:1fff
• Recall that the bit structure is:
  – 0010 0000 0000 0001: 0000 0100 0110 1000: 000 | 0:0000:0000:0000
  – 0010 0000 0000 0001: 0000 0100 0110 1000: 000 | 1:1111:1111:1111
• So there are 8192 /48s in a /35

IPv6 Addressing
Why Allocation?

• To try to control the growth of the routing table in the default-free zone, i.e., (DFZ) refers to the collection of all Internet autonomous systems that do not require a default route.

• It is a necessary consequence of using a provider-based aggregatable address scheme.

• It makes the address space more manageable.
How would allocations work?

• Suppose you wish to give out /40s in the /35.
  – 2001:0468:000 | 0 0000 | or 2001:0468::/40
  – 2001:0468:000 | 1 1111 | or 2001:0468:1f00::/40

• Thus there are 32 /40s in the /35, each of which has 256 /48s.
  – 5 bits
  – 8 bits
How would allocations work?

- The same idea holds for /41s or /42s.
  - 2001:0468:000 | 0:0000:0 | or 2001:0468::/41
  - 2001:0468:000 | 1:1111:1 | or 2001:0468:1f80::/41
  - 2001:0468:000 | 0:0000:00 - :000 | 1:1111:11
  - 2001:0468::/42 – 2001:0468:1fc0::/42
Mixed Allocations

• The interesting case is how to handle mixed allocations.
• Some sites need a /40, others a /42. How can you handle this case?
• See
  – RFC 3531 (Marc Blanchet)
  – A flexible method for managing the assignment of bits of an IPv6 address block
  – A perl script is included.
Mixed Allocations

• Take 2001:468::/32. Out of that allocate:
  – 2 subnets of /34
  – 3 subnets of /37
  – 5 subnets of /38

• Review address allocations (separate slide)

• Assign addresses:
  – Assign /34s for the two top-tier routers.
  – Assign /35s for their downstream routers.
  – Assign /37s for the third-tier routers.
  – Remember at each level to retain some /64s for "local" use, and allocate them for point-to-point links in the network diagram.
  – When you're done, your network diagram will have loopbacks, point-to-points, and appropriately-sized network blocks allocated at each level.