NETWORKWORLD PRESENTS The Critical Path to IPv6

IPv6 Address Design

A Few Practical Principles

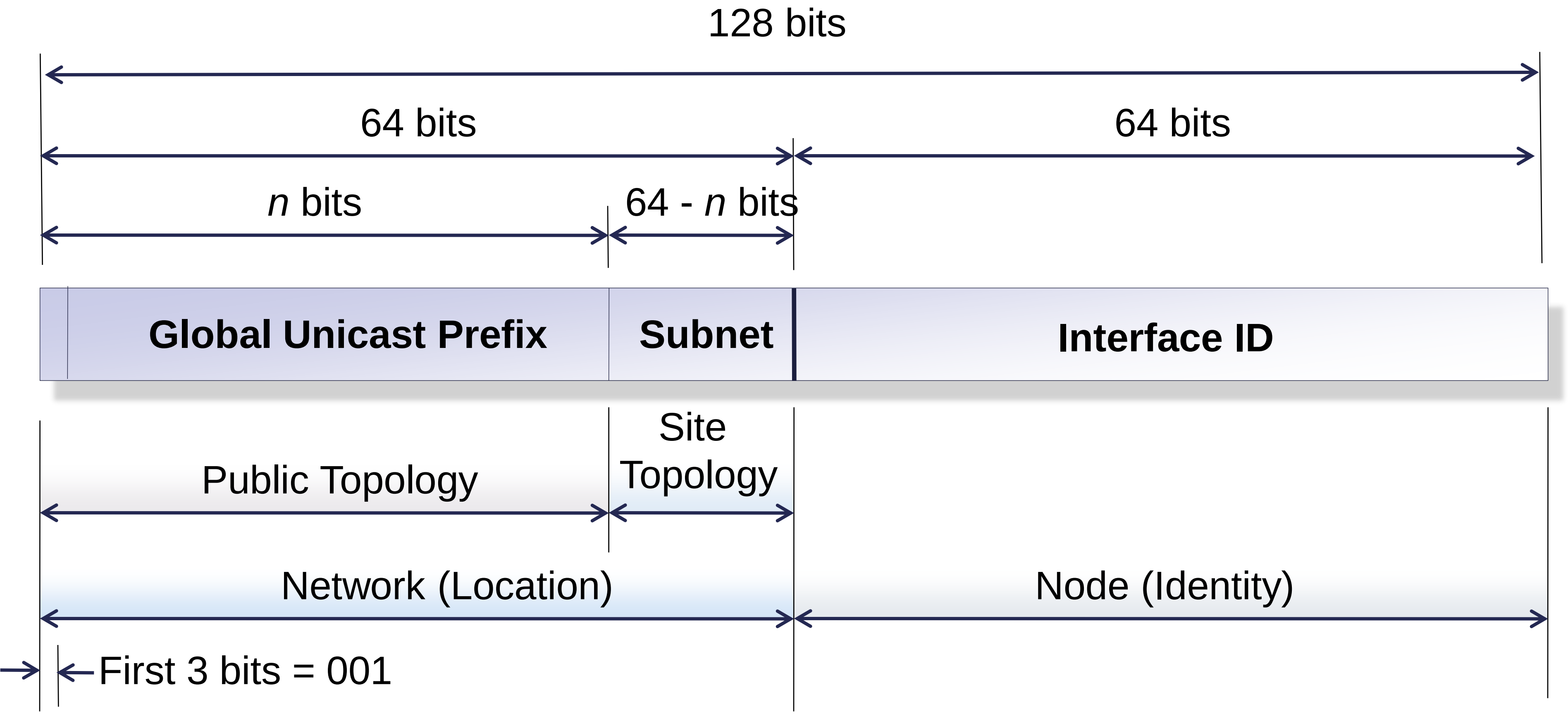
Jeff Doyle
Jeff Doyle and Associates, Inc.

Abandon IPv4 Thinking!

- Foremost IPv4 address design consideration: Address Conservation
- Balancing act between:
  - Number of subnets
  - Number of hosts on *each* subnet
- Result: VLSM
  - Complex
  - Hard to manage
- Legacy “class” categories still occasionally used in IPv4
  - Outdated and misleading
- No such thing as subnet masks in IPv6
  - CIDR-style prefix length notation always used

2001:db8:1234:abcd:5401:3c:15:85/48

IPv6 Global Unicast Address Structure



128 bits

64 bits

64 bits

n bits

64 - *n* bits

Global Unicast Prefix

Subnet

Interface ID

Public Topology

Site Topology

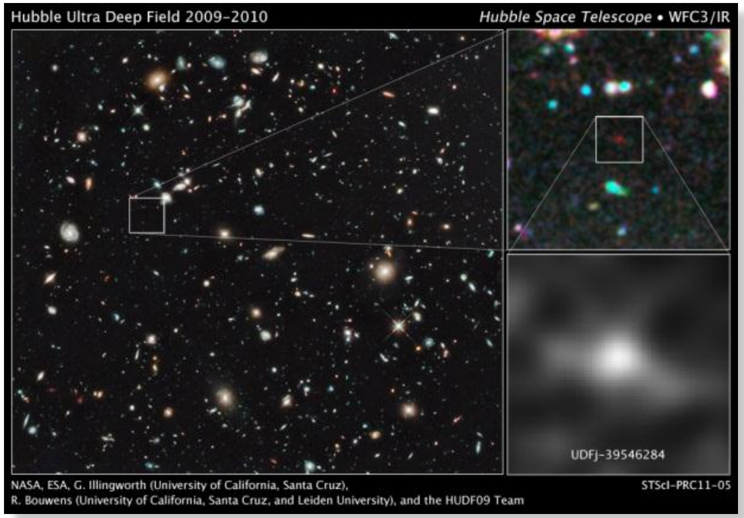
Network (Location)

Node (Identity)

First 3 bits = 001

How Big is the IPv6 Address Space?

- IPv4 developed 1973 – 1977
  - 2^{32} = 4.3 billion addresses
  - More than anyone could possibly use!
- IPv6 developed mid-1990s
  - $2^{128} = 3.4 \times 10^{38}$ addresses
  - More than anyone could possibly use?


Hubble Ultra Deep Field 2009–2010
Hubble Space Telescope • WFC3/IR
UDFj-39546284
STScI-PRC11-05

NASA, ESA, G. Illingworth (University of California, Santa Cruz),
R. Bouwens (University of California, Santa Cruz, and Leiden University), and the HUDF09 Team

Galaxy UDFj-39546284

How Big is the IPv6 Address Space?

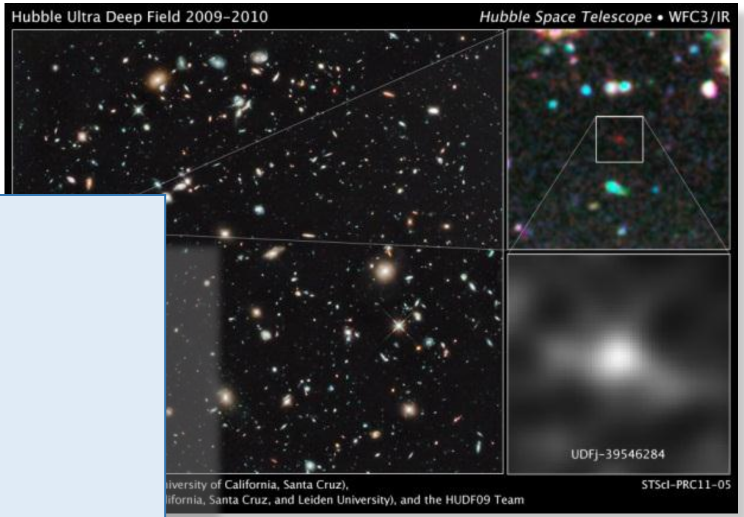
- IPv4 developed 1973 – 1977
  - 2^{32} = 4.3 billion addresses
  - More than anyone could possibly use!
-

Some Perspective:

1 picometer = 10^{-12} (one trillionth) meter

2^{32} picometers = 4.29 millimeters
- length of a small ant

2^{128} picometers = 3.4 x 10^{23} kilometers
- 34 billion light years
- Furthest visible object in universe: 13.2B LYs


Hubble Ultra Deep Field 2009-2010
Hubble Space Telescope • WFC3/IR
UDFj-39546284
STScI-PRC11-05

iversity of California, Santa Cruz), ifornia, Santa Cruz, and Leiden University), and the HUDF09 Team

Galaxy UDFj-39546284

In Practical Terms…

- Typical IPv6 prefix assignments:
  - Service provider (LIR): /32 ➔ 2^{32} /64 subnets
  - Large end user: /48 ➔ 65,536 /64 subnets
  - Small end user: /56 ➔ 256 /64 subnets
  - SOHO: /64 or /60 ➔ 1 or 16 /64 subnets
- Address conservation is *not* a major consideration
  - Is this wasteful?
  - Yes! (But that’s okay)
- If you don’t have enough subnets, you don’t have the right prefix allocation

What Do I Get in Exchange for Waste?

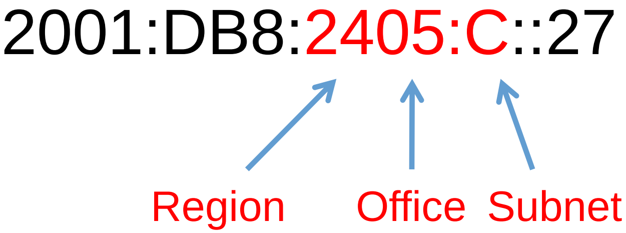
- Simplicity
  - One-size-fits-all subnets
- Manageability
  - Hex is much easier to interpret at binary level than decimal
- Scalability
  - Room to grow
- Flexibility
  - Room to change

Designing for Simplicity

- Start by mapping “working” bits
  - Generally the bits between assigned prefix and Interface-ID
- Group by hex digit
  - 4 bits per hex digit
- Define “meanings” you need to operate
  - Geographic area? Logical topology? Type designation? User ID?
- Try to keep “meanings” on hex boundaries
  - Defined meanings will then be some multiple of 2^{4n}
  - Ex: 16, 256, 4096, 65536…
- Don’t get carried away with meanings
  - No need for 10 layers of address hierarchy if 4 will do

Designing for Simplicity (continued)

- Use zero space as much as possible
  - Which address is easier to read?
    - 2001:DB8:2405:83FC:72A6:3452:19ED:4727
    - 2001:DB8:2405:C::27
- Benefit: Operations quickly learns to focus on meaningful bits
  - Ignore public prefix (usually)
  - Ignore Interface-ID (usually)
  - A few hex digits tell operations most of what they need to know


2001:DB8:2405:C::27

Designing for Scale

- Leave “zero” space whenever possible
  - Designate as Reserved
- Insert between “meaningful” digits or bits
  - Allows future expansion in two directions

Designing for the Future

- Trying to anticipate the unanticipated
  - A challenge for any kind of design
- Another reason for well-placed Reserved (zero) space
  - Horizontal Reserved space
  - Vertical Reserved space
- Do not integrate IPv4 into an IPv6 design!
  - Reading IPv4 in hex is (almost) meaningless
  - IPv4 will (eventually) go away

What About Point-to-Point Links?

- 18 million trillion addresses in a /64 link
  - And I will only *ever* use 2 of them?
  - Are you kidding???
- People have a very hard time accepting this
  - Again: This is not IPv4!
  - What else are you going to do with those addresses?
- It’s a matter of comprehending the scale
  - 500 out of 2^{64} is not really any bigger than 2 out of 2^{64}

Point-to-Point Subnets

- Reasons for using /64:
  - – RFC 3627
  - – RFC 5375 => /64 usage endorsed and encouraged
    - IANA and RIRS also encourage /64 everywhere
  - – Design consistency
  - – *Required* for SLAAC
  - – Anycast problems are not significant on PtP links
    - Subnet-Router Anycast
    - MIPv6 Home Agent Anycast

Point-to-Point Subnets

- Reasons for using /127:
  - – RFC 6164
  - – Ping-pong vulnerability
    - This is an issue with older version of ICMPv6 (RFC 2463)
    - Issue is corrected in newer version of ICMPv6 (RFC 4443)
    - Vendors: Upgrade your code!
  - – Neighbor cache exhaustion vulnerability

Point-to-Point Subnets

- Don’t use /126
  - This is IPv4 thinking
  - “Subnet number” is meaningless in IPv6
  - IPv6 does not use broadcast addresses
- Potential compromise:
  - Assign /64 per PtP subnet
  - Address /127 out of the /64

What About Provider Independence?

- There is (currently) no NAT66
- PI address assignment rules (varies by RIR):
  - Must not be an LIR
  - Must be an end site
  - Must have previously justified a PI IPv4 assignment; or
  - Must currently be multihomed with IPv4; or
    - And have an assigned ASN
    - Proposals to end this requirement
  - Will make active use of 2000 IPv6 addresses within 12 months; or
  - Will make active use of 200 /64s within 12 months; or
  - Technical justification why cannot use assignment from LIR
- PI assignment: One or more /48s
  - Larger based on number of sites
- Micro-allocations available for *critical Internet infrastructure*

Link Local vs Global Unicast

- Some conflict of interpretation
  - Static route next hops
  - BGP peering
- IPv6 says use link local for direct connections
- Accepted practice is to use global unicast
- Recommendation: Stick with accepted practice
  - Link-local harder to manage
  - Interface changes can change link-local address

Other Issues

- DNS design and management is critical
  - DNS issues are well documented
- IP Address Management is critical
  - IPv6 design is not easy to manage via spreadsheets
  - Good luck finding integrated DNS and DHCPv6 management
- Stateful vs Stateless Address Configuration
- Abandon IPv4 thinking!

Questions?

jdoyle@doyleassociates.net

www.doyleassociates.net

+1-303-428-4680