

# Introduction to IPv6



Scalable Infrastructure  
Workshop  
AfNOG 2008

# Agenda

---

- Background
- Protocols & Standards
- Addressing
- Co-existence & Transition

# Early Internet History

---

- Late 1980s
  - Exponential growth of the Internet
- Late 1990: CLNS proposed as IP replacement
- 1991-1992
  - Running out of “class-B” network numbers
  - Explosive growth of the “default-free” routing table
  - Eventual exhaustion of 32-bit address space
- Two efforts – short-term vs. long-term
  - More at “The Long and Windy ROAD”  
<http://rms46.vlsm.org/1/42.html>

# Early Internet History

---

- ❑ CIDR and Supernetting proposed in 1992-3
  - Deployment started in 1994
- ❑ IETF "ipng" solicitation – RFC1550, Dec 1993
- ❑ Direction and technical criteria for ipng choice – RFC1719 and RFC1726, Dec 1994
- ❑ Proliferation of proposals:
  - TUBA – RFC1347, June 1992
  - PIP – RFC1621, RFC1622, May 1994
  - CATNIP – RFC1707, October 1994
  - SIP – RFC1710, October 1994
  - NIMROD – RFC1753, December 1994
  - ENCAPS – RFC1955, June 1996

# Early Internet History

→ 1996

---

- Other activities included:
  - Development of NAT, PPP, DHCP,...
  - Some IPv4 address reclamation
  - The RIR system was introduced
- → Brakes were put on IPv4 address consumption
- IPv4 32 bit address = 4 billion hosts
  - HD Ratio (RFC3194) realistically limits IPv4 to 250 million hosts

# Recent Internet History

## The “boom” years → 2001

---

- IPv6 Development in full swing
  - Rapid IPv4 consumption
  - IPv6 specifications sorted out
  - (Many) Transition mechanisms developed
- 6bone
  - Experimental IPv6 backbone sitting on top of Internet
  - Participants from over 100 countries
- Early adopters
  - Japan, Germany, France, UK,...

# Recent Internet History

## The “bust” years: 2001 → 2004

---

- The DotCom “crash”
  - i.e. Internet became mainstream
- IPv4:
  - Consumption slowed
  - Address space pressure “reduced”
- Indifference
  - Early adopters surging onwards
  - Sceptics more sceptical
  - Yet more transition mechanisms developed

# 2004 → Today

---

- Resurgence in demand for IPv4 address space
  - 18.6% address space still unallocated (05/2008)
  - Exhaustion predictions range from wild to conservative
  - ...but late 2010 seems realistic at current rates
  - ...but what about the market for address space?
- Market for IPv4 addresses:
  - Creates barrier to entry
  - Condemns the less affluent to use of NATs
- IPv6 offers vast address space
  - **The only compelling reason for IPv6**



# Current Situation

---

- General perception is that “IPv6 has not yet taken hold”
  - IPv4 Address run-out is not “headline news” yet
    - More discussions and run-out plans proposed
  - Private sector requires a business case to “migrate”
    - No easy Return on Investment (RoI) computation
- But reality is very different from perception!
  - Something needs to be done to sustain the Internet growth
  - IPv6 or NAT or both or something else?

# Do we really need a larger address space?

---

- Internet population
  - ~630 million users end of 2002 – 10% of world pop.
  - ~1320 million users end of 2007 – 20% of world pop.
  - Future? (World pop. ~9B in 2050)
- US uses 81 /8s – this is 3.9 IPv4 addresses per person
  - Repeat this the world over...
  - 6 billion population could require 23.4 billion IPv4 addresses
  - (6 times larger than the IPv4 address pool)
- Emerging Internet economies need address space:
  - China uses more than 94 million IPv4 addresses today (5.5 /8s)

# Do we really need a larger address space?

---

- RFC 1918 is not sufficient for large environments
  - Cable Operators (e.g. Comcast – NANOG37 presentation)
  - Mobile providers (fixed/mobile convergence)
  - Large enterprises
- The Policy Development process of the RIRs turned down a request to increase private address space
  - RIR membership guideline is to use global addresses instead
  - This leads to an accelerated depletion of the global address space
- 240/4 being proposed as new private address space

# IPv6 OS and Application Support

---

- All software vendors officially support IPv6 in their latest Operating System releases
- Application Support
  - Applications must be IPv4 and IPv6 agnostic
  - User should not have to “pick a protocol”
  - Successful deployment is driven by Applications

# ISP Deployment Activities

---

- Several Market segments
  - IX, Carriers, Regional ISP, Wireless
- ISP have to get an IPv6 prefix from their Regional Registry
- Large carriers planning driven by customer demand:
  - Some running trial networks (e.g. Sprint)
  - Others running commercial services (e.g. NTT, FT,...)
- Regional ISP focus on their specific markets
- Much discussion by operators about transition
  - [www.civil-tongue.net/6and4/](http://www.civil-tongue.net/6and4/)
  - [www.nanog.org/mtg-0710/presentations/Bush-v6-op-reality.pdf](http://www.nanog.org/mtg-0710/presentations/Bush-v6-op-reality.pdf)

# Why not use Network Address Translation?

---

- ❑ Private address space and Network address translation (NAT) could be used instead of IPv6
- ❑ But NAT has many serious issues:
  - Breaks the end-to-end model of IP
  - Layered NAT devices
  - Mandates that the network keeps the state of the connections
  - How to scale NAT performance for large networks?
  - Makes fast rerouting difficult
  - Service provision inhibited

# NAT has many implications

---

- ❑ Inhibits end-to-end network security
- ❑ When a new application is not NAT-friendly, NAT device requires an upgrade
- ❑ Some applications cannot work through NATs
- ❑ Application-level gateways (ALG) are not as fast as IP routing
- ❑ Complicates mergers
  - Double NATing is needed for devices to communicate with each other
- ❑ Breaks security
- ❑ Makes multihoming hard
- ❑ Simply does not scale
- ❑ RFC2993 – architectural implications of NAT

# Conclusion

---

- There is a need for a larger address space
  - IPv6 offers this – will eventually replace NAT
  - But NAT will be around for a while too
  - Market for IPv4 addresses looming also
- Many challenges ahead



# Protocols & Standards



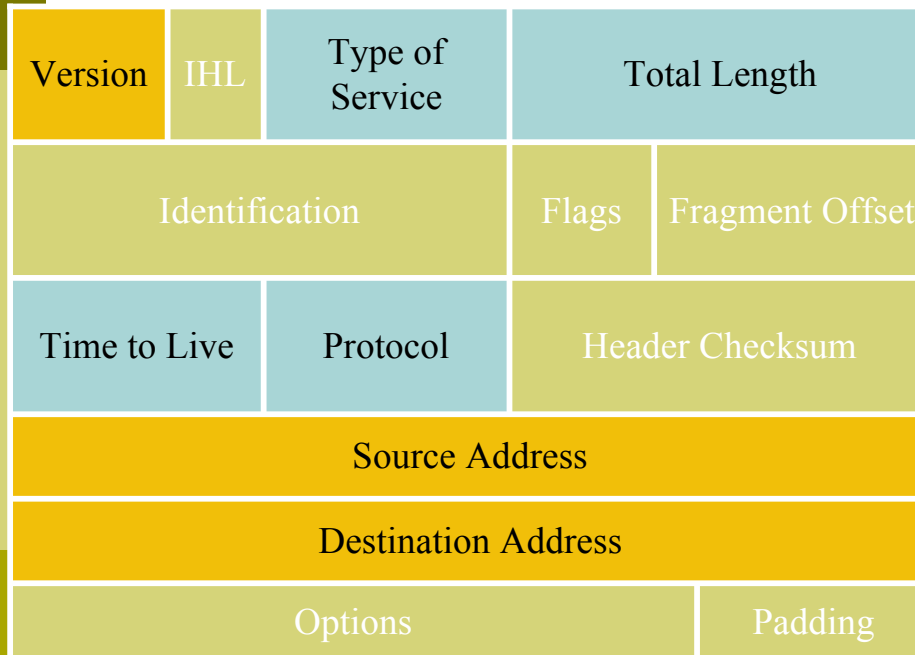
# So what has really changed?

---

- ❑ Expanded address space
  - Address length quadrupled to 16 bytes
- ❑ Header Format Simplification
  - Fixed length, optional headers are daisy-chained
  - IPv6 header is twice as long (40 bytes) as IPv4 header without options (20 bytes)
- ❑ No checksum at the IP network layer
- ❑ No hop-by-hop segmentation
  - Path MTU discovery
- ❑ 64 bits aligned
- ❑ Authentication and Privacy Capabilities
  - IPsec is mandated
- ❑ No more broadcast

# IPv4 and IPv6 Header Comparison

## IPv4 Header



## IPv6 Header

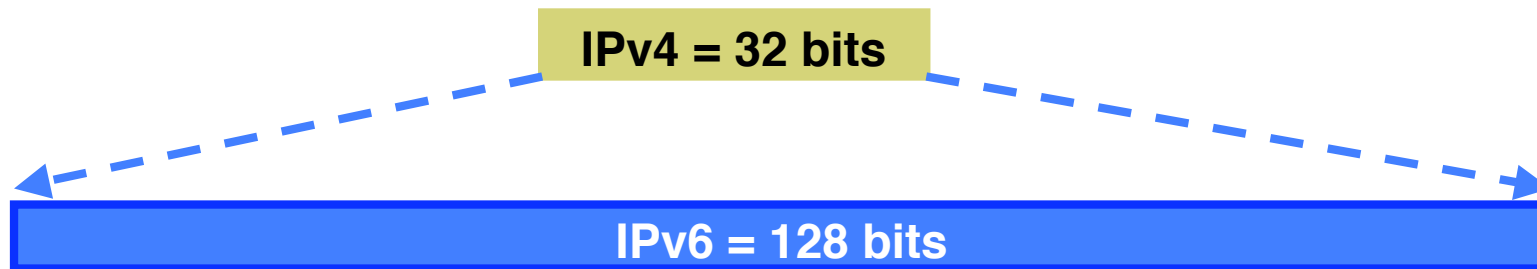


### Legend

- Field's name kept from IPv4 to IPv6
- Fields not kept in IPv6
- Name and position changed in IPv6
- New field in IPv6

# Larger Address Space

---



- IPv4
  - 32 bits
  - = 4,294,967,296 possible addressable devices
- IPv6
  - 128 bits: 4 times the size in bits
  - =  $3.4 \times 10^{38}$  possible addressable devices
  - = 340,282,366,920,938,463,463,374,607,431,768,211,456
  - ~  $5 \times 10^{28}$  addresses per person on the planet

# How was the IPv6 Address Size Chosen?

---

- ❑ Some wanted fixed-length, 64-bit addresses
  - Easily good for  $10^{12}$  sites,  $10^{15}$  nodes, at .0001 allocation efficiency (3 orders of magnitude more than IPv6 requirement)
  - Minimizes growth of per-packet header overhead
  - Efficient for software processing
- ❑ Some wanted variable-length, up to 160 bits
  - Compatible with OSI NSAP addressing plans
  - Big enough for auto-configuration using IEEE 802 addresses
  - Could start with addresses shorter than 64 bits & grow later
- ❑ Settled on fixed-length, 128-bit addresses

# IPv6 Address Representation

---

- 16 bit fields in case insensitive colon hexadecimal representation
  - 2031:0000:130F:0000:0000:09C0:876A:130B
- Leading zeros in a field are optional:
  - 2031:0:130F:0:0:9C0:876A:130B
- Successive fields of 0 represented as ::, but only once in an address:

- 2031:0:130F::9C0:876A:130B is ok
- 2031::**130F**::9C0:876A:130B is **NOT** ok



- 0:0:0:0:0:0:0:1 → ::1 (loopback address)
- 0:0:0:0:0:0:0:0 → :: (unspecified address)

# IPv6 Address Representation

---

- ❑ In a URL, it is enclosed in brackets (RFC3986)
  - `http://[2001:db8:4f3a::206:ae14]:8080/index.html`
  - Cumbersome for users
  - Mostly for diagnostic purposes
  - Use fully qualified domain names (FQDN)
- ❑ Prefix Representation
  - Representation of prefix is same as for IPv4 CIDR
    - ❑ Address and then prefix length
  - IPv4 address:
    - ❑ `198.10.0.0/16`
  - IPv6 address:
    - ❑ `2001:db8:12::/40`

# IPv6 Addressing

---

- ❑ IPv6 Addressing rules are covered by multiples RFCs
  - Architecture defined by RFC 4291
- ❑ Address Types are :
  - Unicast : One to One (Global, Unique Local, Link local)
  - Anycast : One to Nearest (Allocated from Unicast)
  - Multicast : One to Many
- ❑ A single interface may be assigned multiple IPv6 addresses of any type (unicast, anycast, multicast)
  - No Broadcast Address → Use Multicast

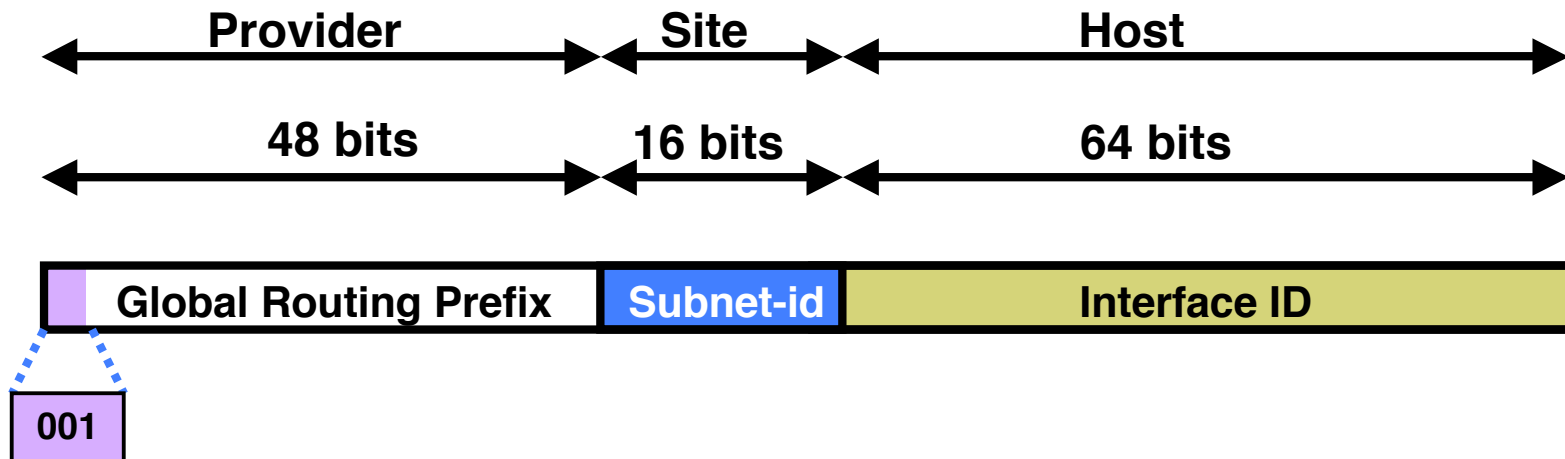


# IPv6 Addressing

Type	Binary	Hex
Unspecified	000...0	::/128
Loopback	000...1	::1/128
Global Unicast Address	0010	2000::/3
Link Local Unicast Address	1111 1110 10	FE80::/10
Unique Local Unicast Address	1111 1100 1111 1101	FC00::/7
Multicast Address	1111 1111	FF00::/8

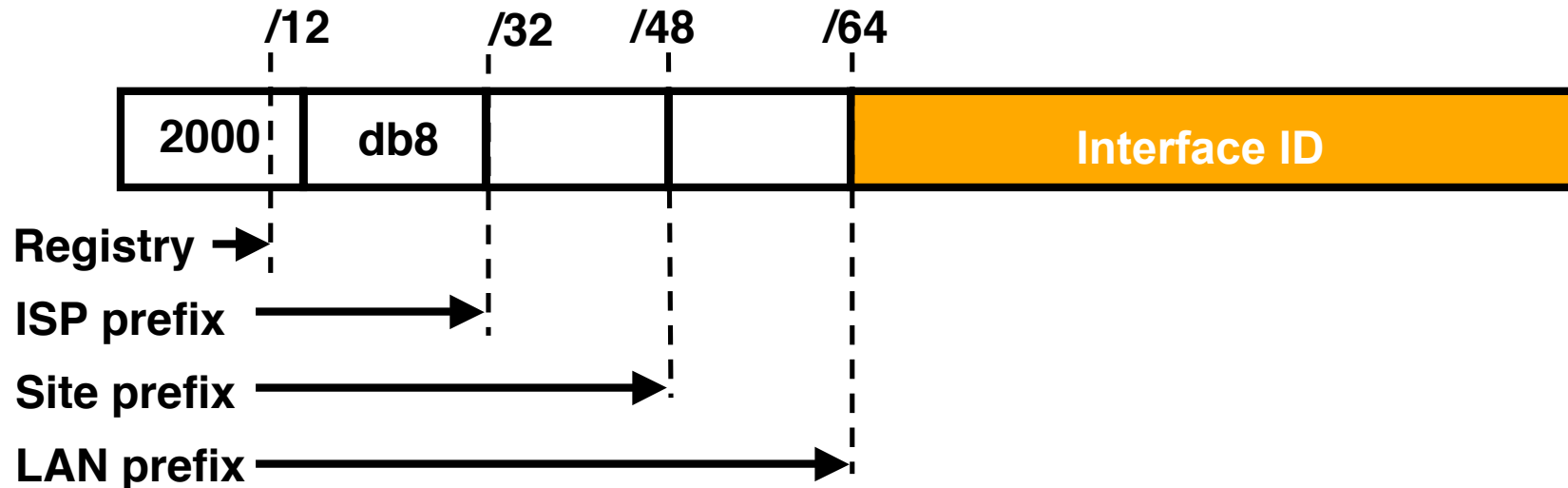
# IPv6 Global Unicast Addresses

---



- IPv6 Global Unicast addresses are:
  - Addresses for generic use of IPv6
  - Hierarchical structure intended to simplify aggregation

# IPv6 Address Allocation



- The allocation process is:
  - The IANA is allocating out of 2000::/3 for initial IPv6 unicast use
  - Each registry gets a /12 prefix from the IANA
  - Registry allocates a /32 prefix (or larger) to an ISP
  - Policy is that an ISP allocates a /48 prefix to each end customer

# IPv6 Addressing Scope

---

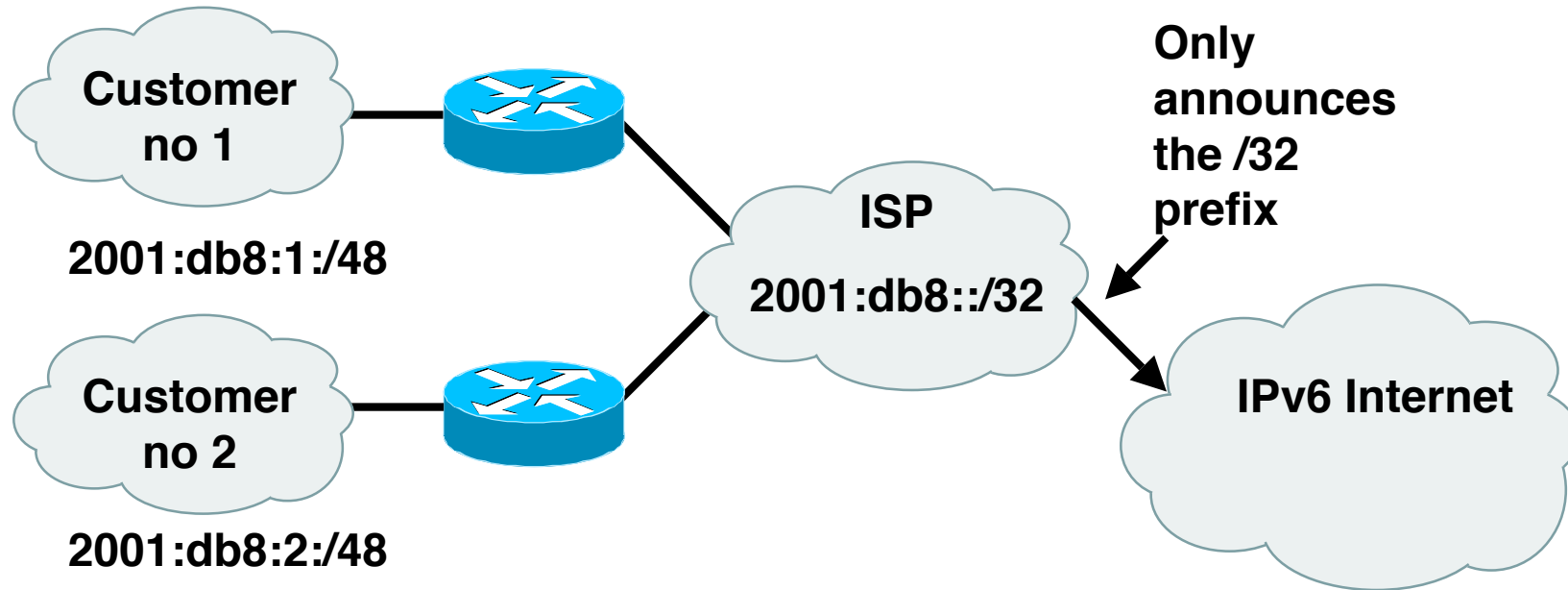
- 64 bits reserved for the interface ID
  - Possibility of  $2^{64}$  hosts on one network LAN
  - Arrangement to accommodate MAC addresses within the IPv6 address
- 16 bits reserved for the end site
  - Possibility of  $2^{16}$  networks at each end-site
  - 65536 subnets equivalent to a /12 in IPv4 (assuming 16 hosts per IPv4 subnet)

# IPv6 Addressing Scope

---

- 16 bits reserved for the service provider
  - Possibility of  $2^{16}$  end-sites per service provider
  - 65536 possible customers: equivalent to each service provider receiving a /8 in IPv4 (assuming a /24 address block per customer)
- 32 bits reserved for service providers
  - Possibility of  $2^{32}$  service providers
  - i.e. 4 billion discrete service provider networks
    - Although some service providers already are justifying more than a /32
  - Equivalent to the size of the entire IPv4 address space

# Aggregation hopes



- ❑ Larger address space enables aggregation of prefixes announced in the global routing table
- ❑ Idea was to allow efficient and scalable routing
- ❑ **But current Internet multihoming solution breaks this model**

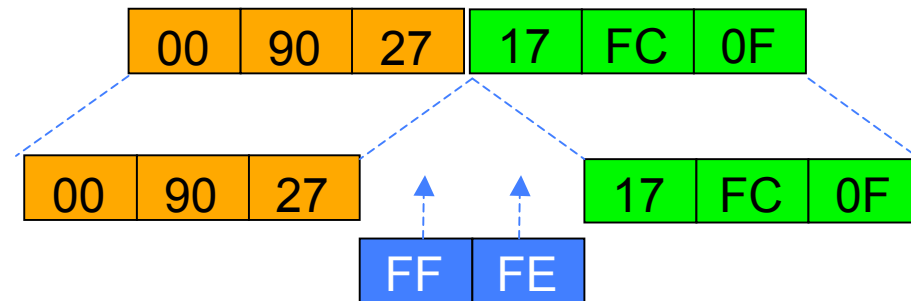
# Interface IDs

---

- Lowest order 64-bit field of unicast address may be assigned in several different ways:
  - Auto-configured from a 64-bit EUI-64, or expanded from a 48-bit MAC address (e.g., Ethernet address)
  - Auto-generated pseudo-random number (to address privacy concerns)
  - Assigned via DHCP
  - Manually configured

# EUI-64

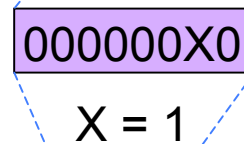
**Ethernet MAC address  
(48 bits)**



**64 bits version**

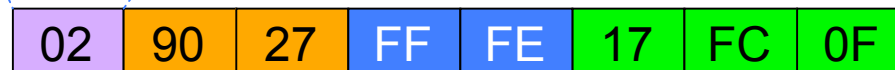


**Uniqueness of the MAC**



where  $X = \begin{cases} 1 = \text{unique} \\ 0 = \text{not unique} \end{cases}$

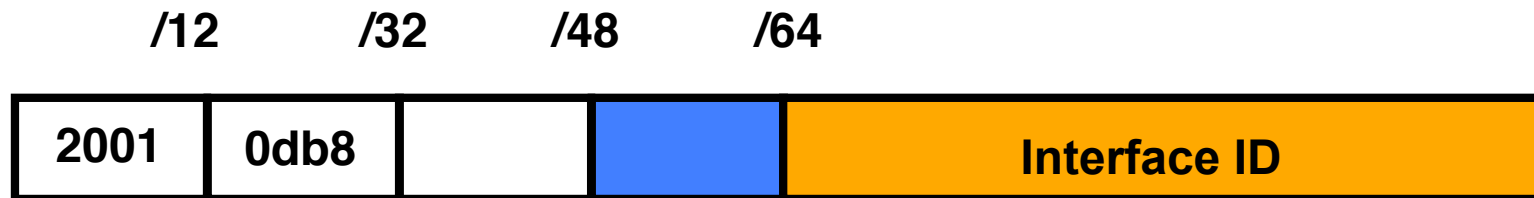
**Eui-64 address**



- EUI-64 address is formed by inserting FFFE and OR'ing a bit identifying the uniqueness of the MAC address



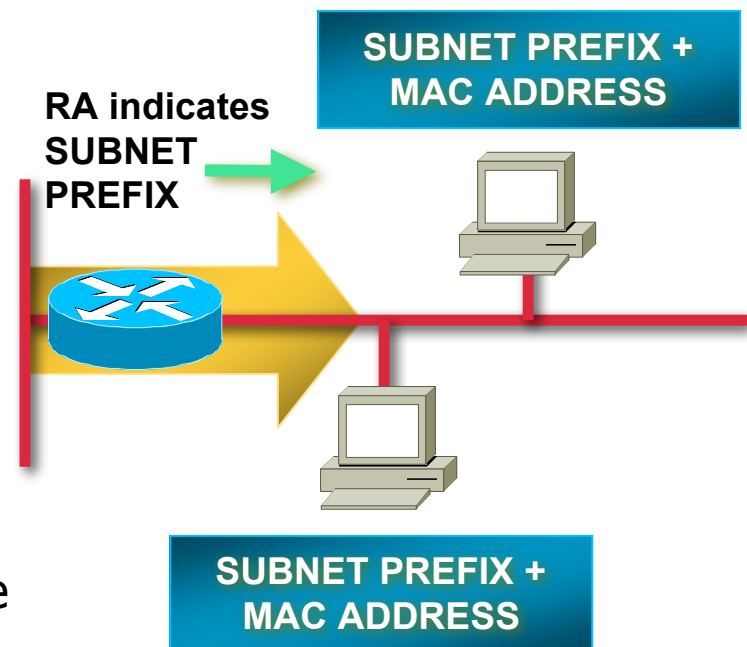
# IPv6 Address Privacy (RFC 3041)



- Temporary addresses for IPv6 host client application, e.g. Web browser
- Intended to inhibit device/user tracking but is also a potential issue
  - More difficult to scan all IP addresses on a subnet
  - But port scan is identical when an address is known
- Random 64 bit interface ID, run DAD before using it
- Rate of change based on local policy
- **Implemented on Microsoft Windows XP & Vista only**

# IPv6 Auto-Configuration

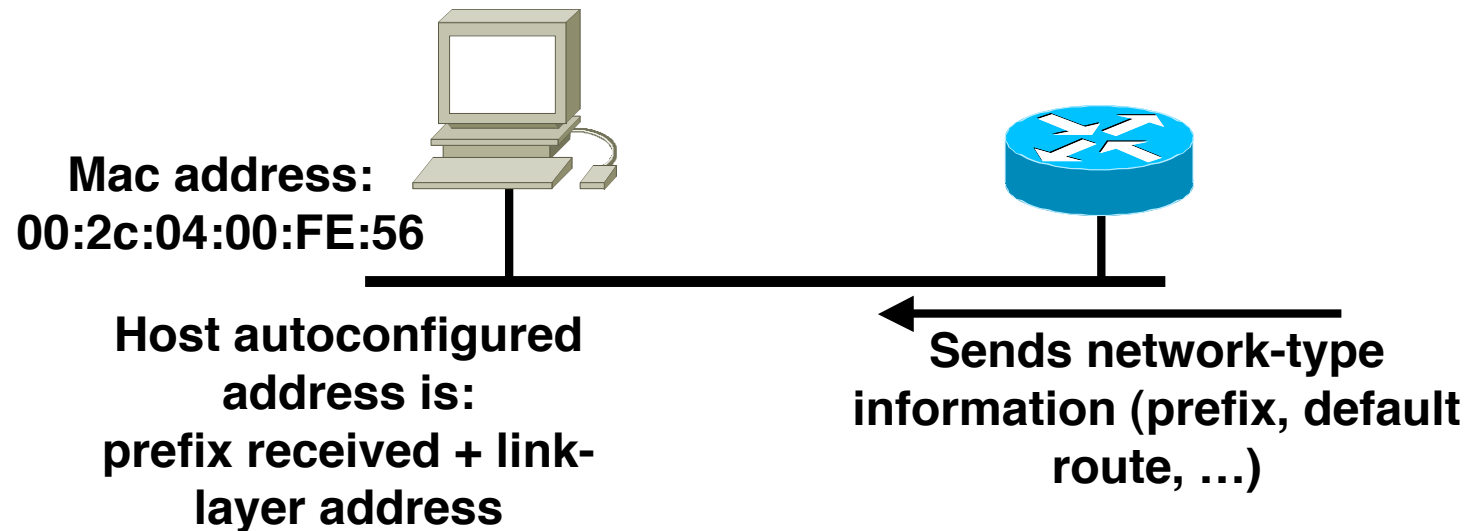
- Stateless (RFC2462)
  - Host autonomously configures its own Link-Local address
  - Router solicitation are sent by booting nodes to request RAs for configuring the interfaces.
- Stateful
  - DHCPv6 – required by most enterprises
- Renumbering
  - Hosts renumbering is done by modifying the RA to announce the old prefix with a short lifetime and the new prefix
  - Router renumbering protocol (RFC 2894), to allow domain-interior routers to learn of prefix introduction / withdrawal



**At boot time, an IPv6 host build a Link-Local address, then its global IPv6 address(es) from RA**

# Auto-configuration

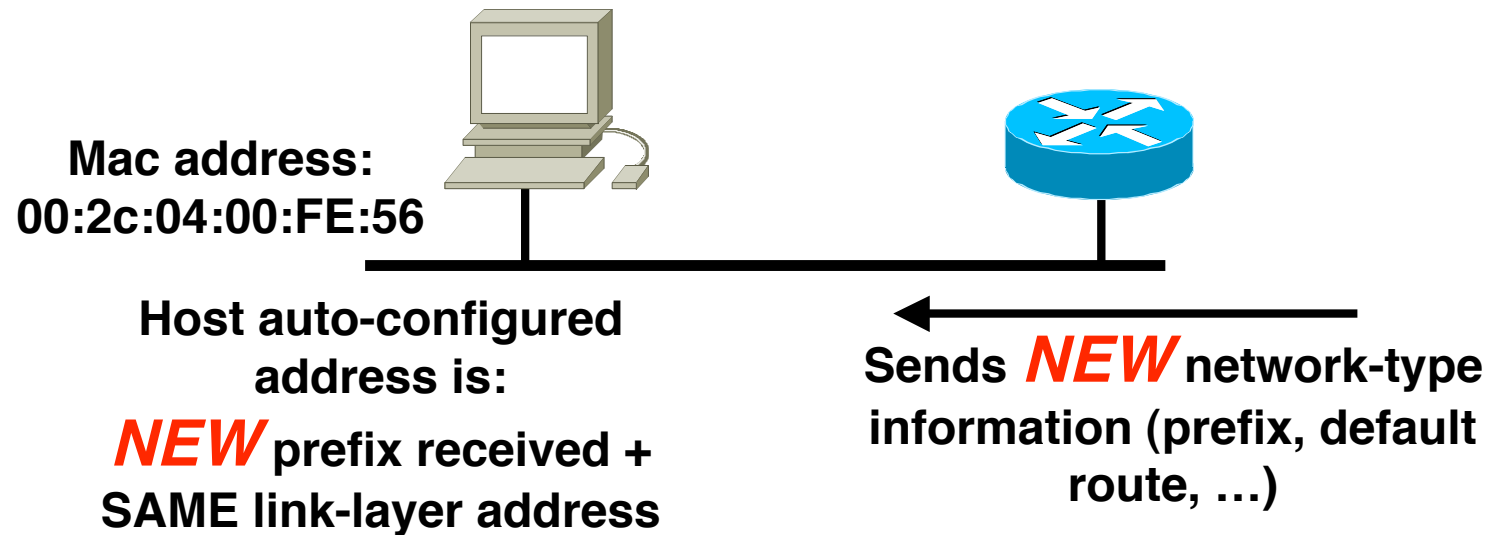
---



- ❑ Client sends router solicitation (RS) messages
- ❑ Router responds with router advertisement (RA)
  - This includes prefix and default route
- ❑ Client configures its IPv6 address by concatenating prefix received with its EUI-64 address

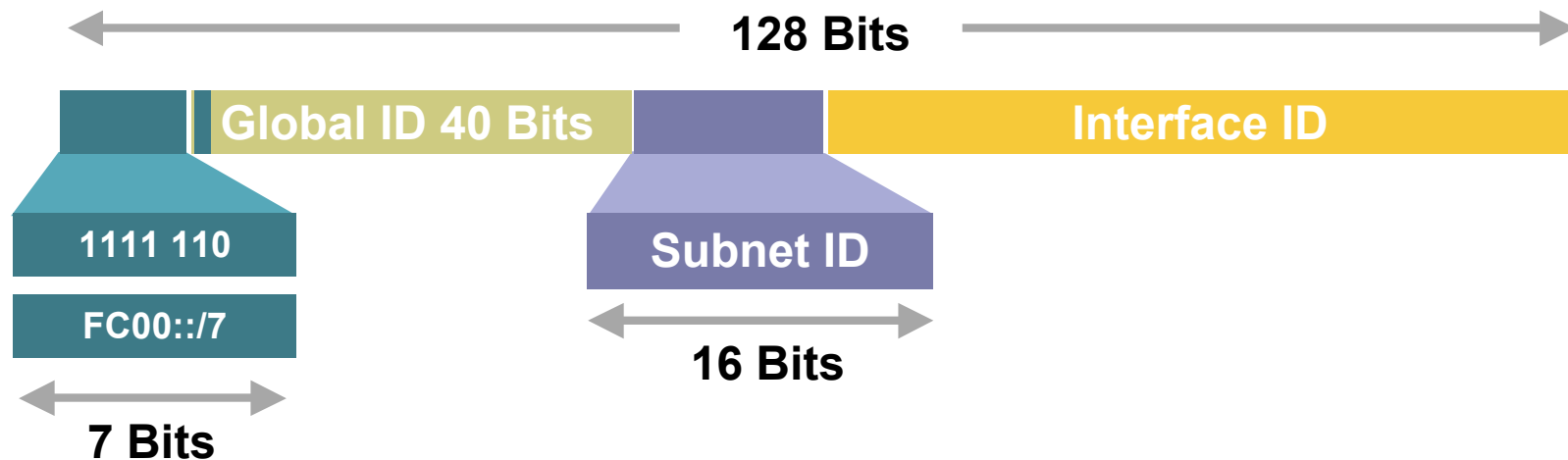
# Renumbering

---



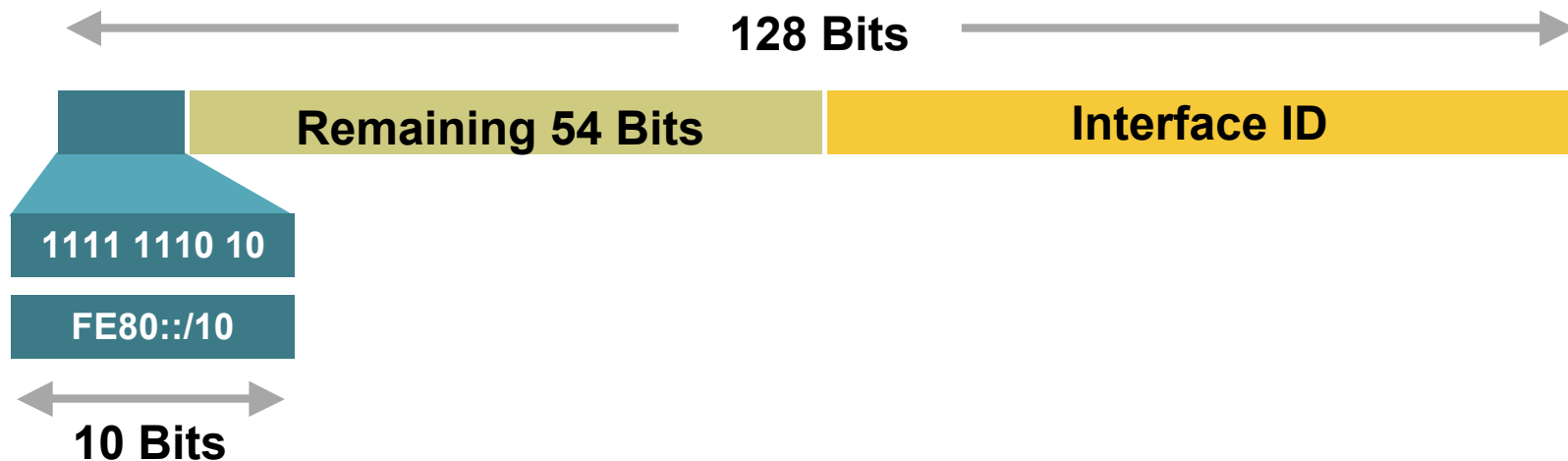
- Router sends router advertisement (RA)
  - This includes the new prefix and default route (and remaining lifetime of the old address)
- Client configures a new IPv6 address by concatenating prefix received with its EUI-64 address
  - Attaches lifetime to old address

# Unique-Local



- ❑ Unique-Local Addresses Used For:
  - Local communications
  - Inter-site VPNs
- ❑ **Not** routable on the Internet
- ❑ Reinvention of the deprecated site-local? It's future is unclear.

# Link-Local



- ❑ Link-Local Addresses Used For:
  - Communication between two IPv6 device (like ARP but at Layer 3)
  - Next-Hop calculation in Routing Protocols
- ❑ Automatically assigned by Router as soon as IPv6 is enabled
  - Mandatory Address
- ❑ Only Link Specific scope
- ❑ Remaining 54 bits could be Zero or any manual configured value

# Multicast use

---

- Broadcasts in IPv4
  - Interrupts all devices on the LAN even if the intent of the request was for a subset
  - Can completely swamp the network (“broadcast storm”)
- Broadcasts in IPv6
  - Are not used and replaced by multicast
- Multicast
  - Enables the efficient use of the network
  - Multicast address range is much larger

# IPv6 Multicast Address

- IP multicast address has a prefix FF00::/8
- The second octet defines the lifetime and scope of the multicast address.

8-bit	4-bit	4-bit	112-bit
1111 1111	Lifetime	Scope	Group-ID

Lifetime	
0	If Permanent
1	If Temporary

Scope	
1	Node
2	Link
5	Site
8	Organization
E	Global



# IPv6 Multicast Address Examples

---

## □ RIPng

- The multicast address **AllRIPRouters** is **FF02::9**

- Note that 02 means that this is a permanent address and has link scope

## □ OSPFv3

- The multicast address **AllSPFRouters** is **FF02::5**

- The multicast address **AllDRouters** is **FF02::6**

## □ EIGRP

- The multicast address **AllEIGRPRouters** is **FF02::A**

# IPv6 Anycast

---

- ❑ An IPv6 anycast address is 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 protocol’s measure of distance).
  - RFC4291 describes IPv6 Anycast in more detail
- ❑ In reality there is no known implementation of IPv6 Anycast as per the RFC
  - Most operators have chosen to use IPv4 style anycast instead

# Anycast on the Internet

---

- A global unicast address is assigned to all nodes which need to respond to a service being offered
  - This address is routed as part of its parent address block
- The responding node is the one which is closest to the requesting node according to the routing protocol
  - Each anycast node looks identical to the other
- Applicable within an ASN, or globally across the Internet
- Typical (IPv4) examples today include:
  - Root DNS and ccTLD/gTLD nameservers
  - SMTP relays within ISP autonomous systems

# MTU Issues

---

- ❑ Minimum link MTU for IPv6 is 1280 octets (versus 68 octets for IPv4)
  - $\Rightarrow$  on links with  $MTU < 1280$ , link-specific fragmentation and reassembly must be used
- ❑ Implementations are expected to perform path MTU discovery to send packets bigger than 1280
- ❑ Minimal implementation can omit PMTU discovery as long as all packets kept  $\geq 1280$  octets
- ❑ A Hop-by-Hop Option supports transmission of “jumbograms” with up to  $2^{32}$  octets of payload

# Neighbour Discovery (RFCs 2461 & 4311)

---

- Protocol built on top of ICMPv6 (RFC 4443)
  - combination of IPv4 protocols (ARP, ICMP, IGMP,...)
- Fully dynamic, interactive between Hosts & Routers
  - defines 5 ICMPv6 packet types:
    - Router Solicitation / Router Advertisements
    - Neighbour Solicitation / Neighbour Advertisements
    - Redirect

# IPv6 and DNS

---

	IPv4	IPv6
<b>Hostname to IP address</b>	<b>A record:</b> www.abc.test. A 192.168.30.1	<b>AAAA record:</b> www.abc.test AAAA 2001:db8:c18:1::2
<b>IP address to hostname</b>	<b>PTR record:</b> 1.30.168.192.in-addr.arpa. PTR www.abc.test.	<b>PTR record:</b> 2.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.1.0.0.0.8.1.c.0. 8.b.d.0.1.0.0.2.ip6.arpa PTR www.abc.test.

# IPv6 Technology Scope

<i>IP Service</i>	<i>IPv4 Solution</i>	<i>IPv6 Solution</i>
Addressing Range	32-bit, Network Address Translation	<b>128-bit, Multiple Scopes</b>
Autoconfiguration	DHCP	<b>Serverless, Reconfiguration, DHCP</b>
Security	IPSec	<b>IPSec Mandated, works End-to-End</b>
Mobility	Mobile IP	<b>Mobile IP with Direct Routing</b>
Quality-of-Service	Differentiated Service, Integrated Service	Differentiated Service, Integrated Service
IP Multicast	IGMP/PIM/Multicast BGP	<b>MLD/PIM/Multicast BGP, Scope Identifier</b>

# What does IPv6 do for:

---

## □ Security

- Nothing IPv4 doesn't do – IPsec runs in both
- But IPv6 architecture mandates IPsec

## □ QoS

- Nothing IPv4 doesn't do –
  - Differentiated and Integrated Services run in both
  - So far, Flow label has no real use



# IPv6 Status – Standardisation

---

## □ Several key components on standards track...

Specification (RFC2460)	Neighbour Discovery (RFC4861 & 4311)
ICMPv6 (RFC4443)	IPv6 Addresses (RFC4291 & 3587)
RIP (RFC2080)	BGP (RFC2545)
IGMPv6 (RFC2710)	OSPF (RFC2740)
Router Alert (RFC2711)	Jumbograms (RFC2675)
Autoconfiguration (RFC4862)	Radius (RFC3162)
DHCPv6 (RFC3315 & 4361)	Flow Label (RFC3697)
IPv6 Mobility (RFC3775)	Mobile IPv6 MIB (RFC4295)
GRE Tunnelling (RFC2473)	Unique Local IPv6 Addresses (RFC4193)
DAD for IPv6 (RFC4429)	Teredo (RFC4380)

## □ IPv6 available over:

PPP (RFC5072)	Ethernet (RFC2464)
FDDI (RFC2467)	Token Ring (RFC2470)
NBMA (RFC2491)	ATM (RFC2492)
Frame Relay (RFC2590)	ARCnet (RFC2497)
IEEE1394 (RFC3146)	FibreChannel (RFC4338)

# Addressing



# Getting IPv6 address space

---

- Become a member of your Regional Internet Registry and get your own allocation
  - Require a plan for a year ahead
  - General allocation policies and specific details for IPv6 are on the individual RIR website
- or
- Take part of upstream ISP's PA space
- or
- Use 6to4
- There is **plenty** of IPv6 address space
  - The RIRs require high quality documentation

# Getting IPv6 address space

---

- From the RIR
  - Receive a /32 (or larger if you have more than 65k /48 assignments)
- From your upstream ISP
  - Get one /48 from your upstream ISP
  - More than one /48 if you have more than 65k subnets
- Use 6to4
  - Take a single public IPv4 /32 address
  - 2002:<ipv4 /32 address>::/48 becomes your IPv6 address block, giving 65k subnets
  - Requires a 6to4 gateway

# Addressing Plans – ISP Infrastructure

---

- ❑ ISPs should receive /32 from their RIR
- ❑ Address block for router loop-back interfaces
  - Generally number all loopbacks out of **one** /64
- ❑ Address block for infrastructure
  - /48 allows 65k subnets
  - /48 per PoP or region (for large networks)
  - /48 for whole backbone (for small to medium networks)
  - Summarise between sites if it makes sense

# Addressing Plans – ISP Infrastructure

---

- What about LANs?
  - /64 per LAN
- What about Point-to-Point links?
  - Expectation is that /64 is used
  - People have used /126s
    - Mobile IPv6 Home Agent discovery won't work
  - People have used /112s
    - Leaves final 16 bits free for node IDs
  - See RFC3627 for more discussion

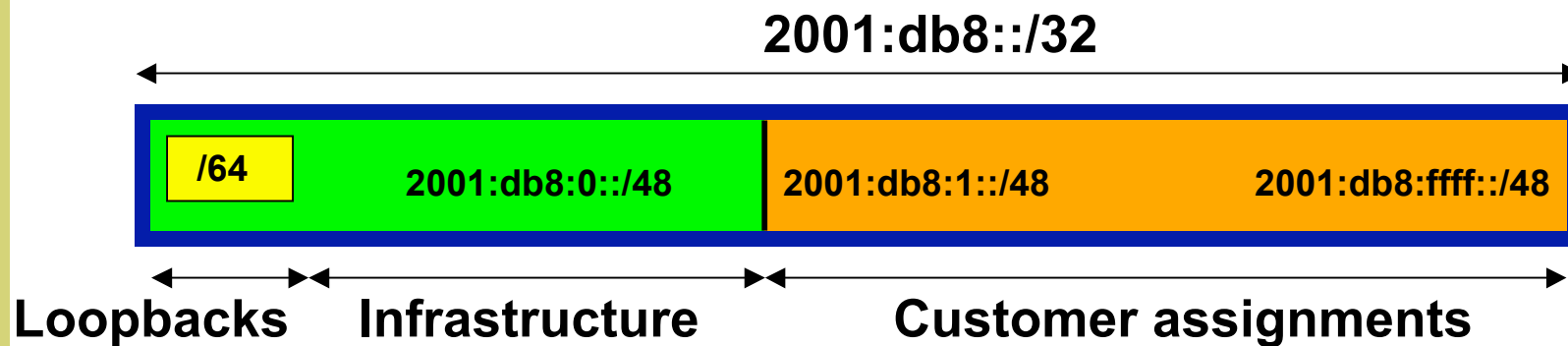
# Addressing Plans – Customer

---

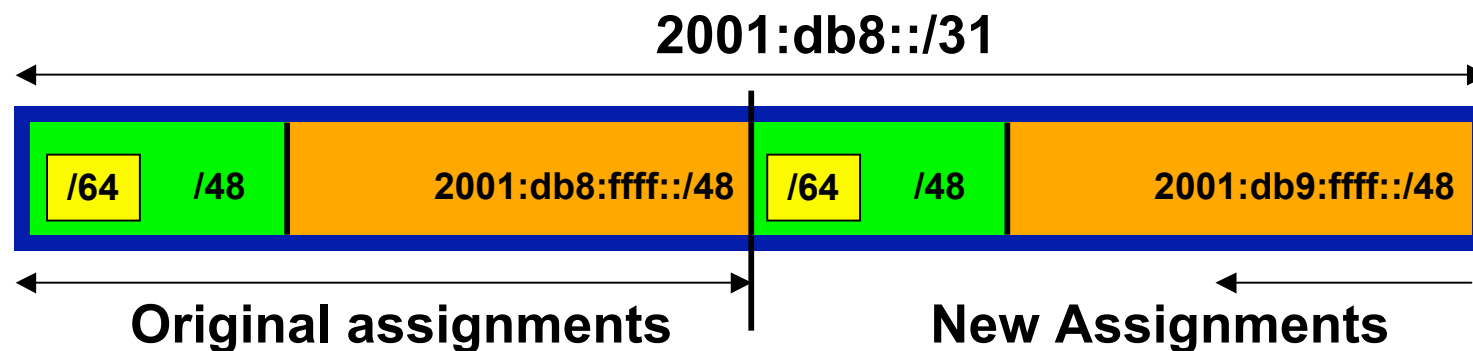
- Customers get **one** /48
  - Unless they have more than 65k subnets in which case they get a second /48 (and so on)
  - (Still on going RIR policy discussion about giving “small” customers a /56 and single LAN end-sites a /64)
- Should not be reserved or assigned on a per PoP basis
  - ISP iBGP carries customer nets
  - Aggregation within the iBGP not required and usually not desirable
  - Aggregation in eBGP is very necessary

# Addressing Plans – ISP Infrastructure

## Phase One



## Phase Two – second /32





# Addressing Plans

## Planning

---

- Registries will usually allocate the next block to be contiguous with the first allocation
  - Minimum allocation is /32
  - Very likely that subsequent allocation will make this up to a /31
  - So plan accordingly

# Transition & Coexistence

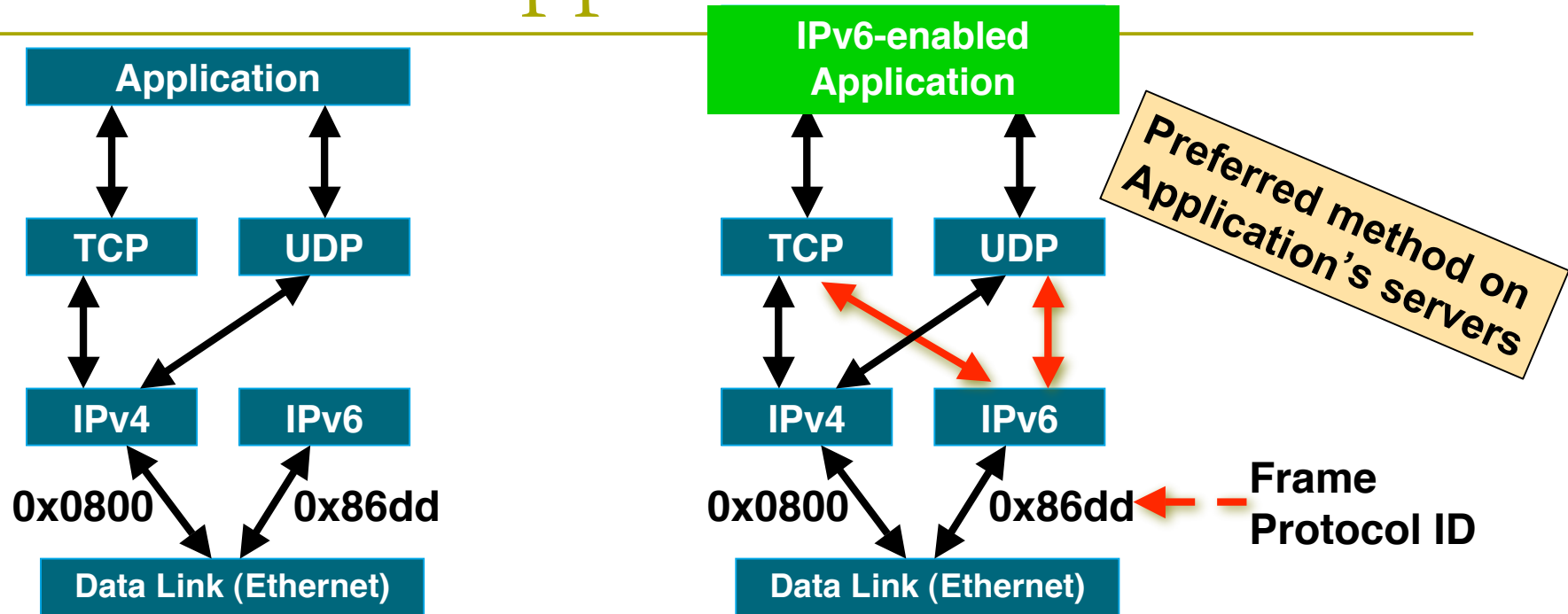


# IPv4-IPv6 Co-existence/Transition

---

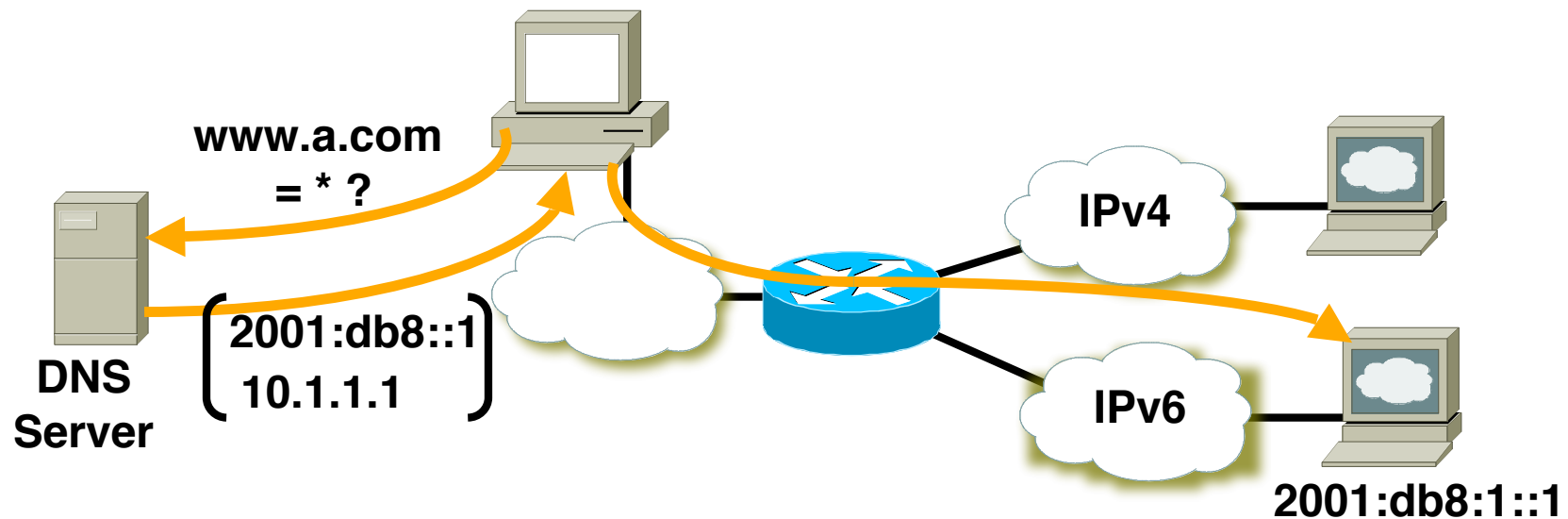
- A wide range of techniques have been identified and implemented, basically falling into three categories:
  - **Dual-stack techniques**, to allow IPv4 and IPv6 to co-exist in the same devices and networks
  - **Tunneling techniques**, to avoid dependencies when upgrading hosts, routers, or regions
  - **Translation techniques**, to allow IPv6-only devices to communicate with IPv4-only devices
- Expect all of these to be used, in combination

# Dual Stack Approach



- Dual stack node means:
  - Both IPv4 and IPv6 stacks enabled
  - Applications can talk to both
  - Choice of the IP version is based on name lookup and application preference

# Dual Stack & DNS



- On a system running dual stack, an application that is both IPv4 and IPv6 enabled will:
  - Ask the DNS for an IPv6 address (AAAA record)
  - If that exists, IPv6 transport will be used
  - If it does not exist, it will then ask the DNS for an IPv4 address (A record) and use IPv4 transport instead

# Using Tunnels for IPv6 Deployment

---

- ❑ Many techniques are available to establish a tunnel:
  - Manually configured
    - ❑ Manual Tunnel (RFC 4213)
    - ❑ GRE (RFC 2473)
  - Semi-automated
    - ❑ Tunnel broker
  - Automatic
    - ❑ 6to4 (RFC 3056)
    - ❑ ISATAP (RFC 4214)
    - ❑ TEREDO (RFC 4380)

ISATAP & TEREDO are more useful for Enterprises than for Service Providers

# Summary

---

- IPv6 offers vast address space
- Distinct addressing hierarchy between ISPs, end-sites, and LANs
  - Planning is not so “confined” as for IPv4
- Coexistence with, **NOT** replacement of IPv4
- Clients prefer IPv6 before IPv4
  - If IPv6 is configured & available