IPv6 Tutorial

Ing. Gunter Van de Velde
Dr. Athanassios Liakopoulos
Ir. Wim Verrydt
Dr. Ciprian Popoviciu

www.6diss.org

6DISS - Intro

- 6DISS: IPv6 Dissemination and Exploitation
- Key Data:
  - Partners:
    - Martel
    - Cisco, Alcatel
    - GRnet, Renater, FCCN, NIIF/Hungarnet, Terena
    - University College London, University of Southampton
  - Duration:
    - 1st April 2005, for 30 months
  - EC Funding

Objectives (1)

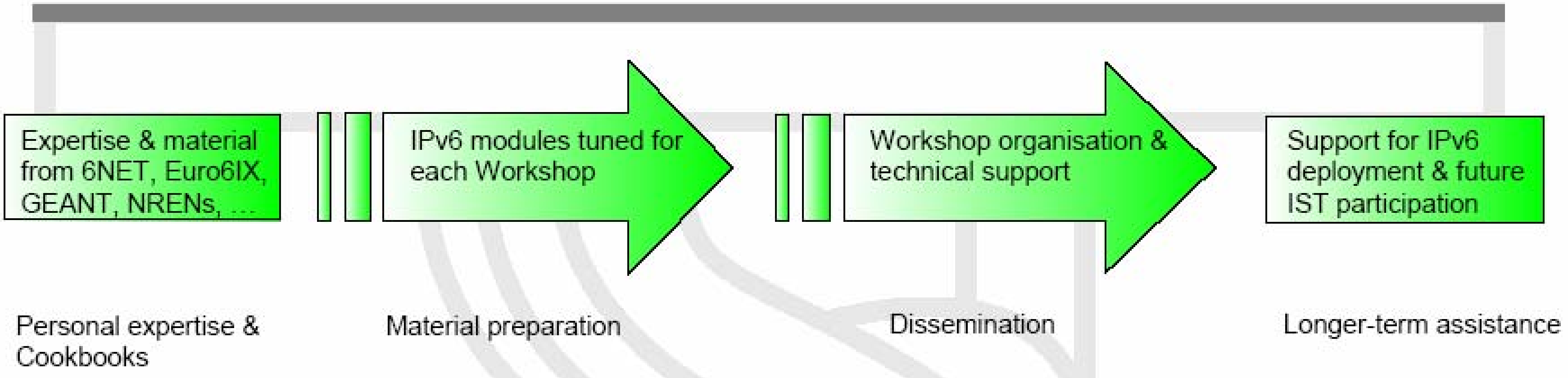
To transfer knowledge and deployment experiences from European IPv6 projects (6NET, Euro6IX, GEANT), TERENA, European NRENs, etc. to research network operators, universities, commercial organisations, governments and regulators in:

- Balkan countries (inc. Bulgaria, Romania, Moldova & Turkey)
- Mediterranean countries
- Newly-Independent States (Central Asia, including Afghanistan)
- Sub-Saharan Africa
- Southern Africa
- The Caribbean
- Asia-Pacific region
- South and Central America
- … and will exchange information with such organisations in India & China

Objectives (2)

To build constituencies, raise awareness, exchange best practices, and engage the organisations in future IST projects.

Technical Approach: Workshops



Expertise & material from 6NET, Euro6IX, GEANT, NRENs, …

IPv6 modules tuned for each Workshop

Workshop organisation & technical support

Support for IPv6 deployment & future IST participation

Personal expertise & Cookbooks

Material preparation

Dissemination

Longer-term assistance

8 Workshops (1 per targeted region, “owned” by a specific partner)

Balkan countries	GRNET
Mediterranean, Sub-Saharan Africa & Caribbean	RENATER
Newly-Independent States	UCL
South & Central America	TERENA/FCCN
Southern Africa	Cisco
Southern Asia	Uni SOTON

plus exchange of best practices and deployment examples at events in India and China — Uni SOTON, Martel

Other Technical Approaches

1. “Tiger Teams” (1 expert per topic for back-up technical support, maintaining FAQ lists, etc.)
2. IPv6 Training at Cisco in Brussels (hands-on)
3. “Training the Trainers” (people can be trained in all topics and go back to their regions to teach others)
4. e-learning (on-line guide to where reference information can be found – eg. 6NET Cookbooks)

Agenda

- Technology Introduction
- IPv6 Protocol Basics
- IPv6 Protocol Specifics
- IPv6 Transition and Coexistence with IPv4

What is IPv6? Basic Perspectives

The Network Manager Perspective

Infrastructure focus

- **Stability of a given technology, implementations and benefits**
- **Cost of deployment and operation**

Care but…has to get confident

The End-User Perspective

Applications focus

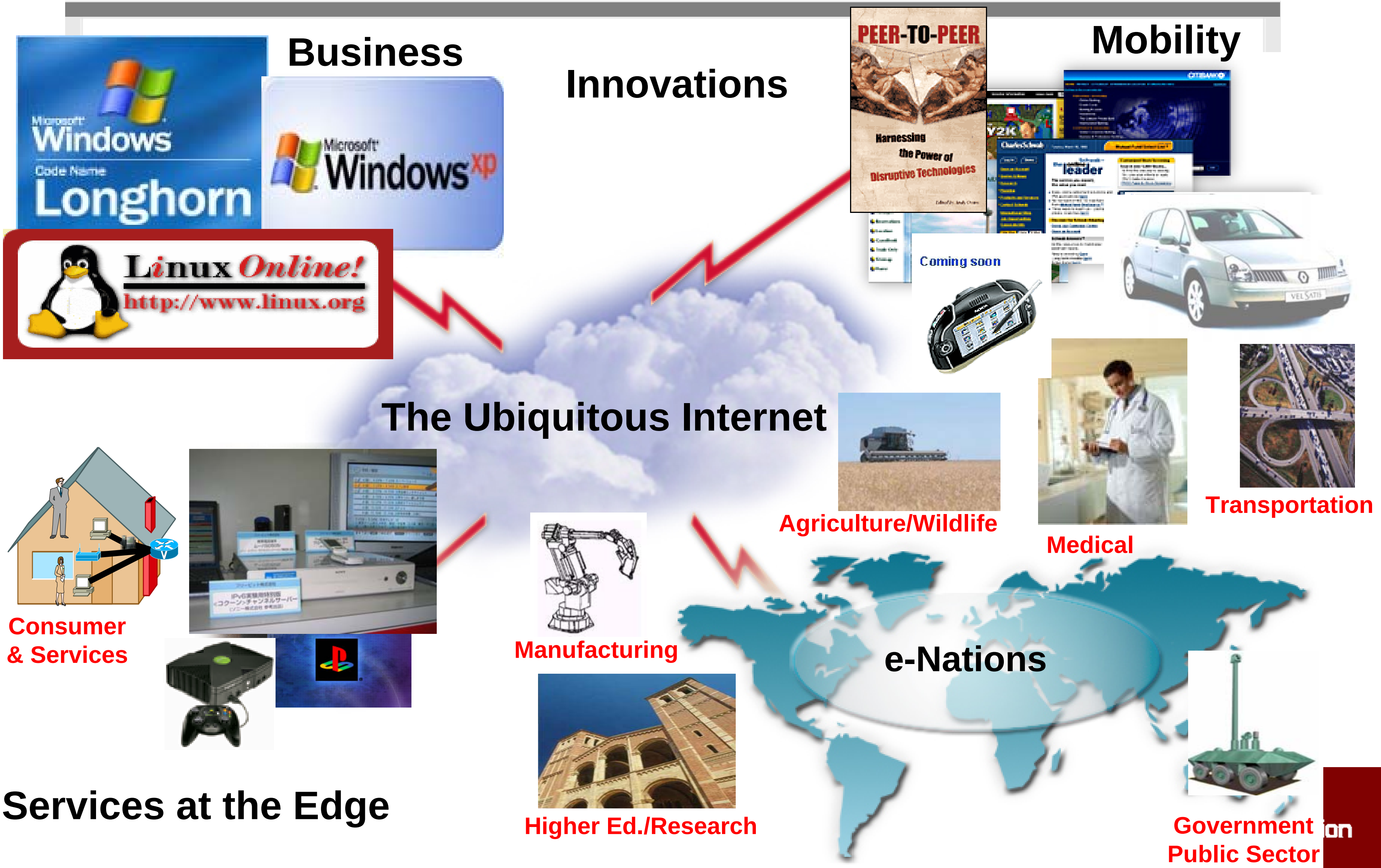
- **The network capability to provide the desired services**
- **It's all about the applications, and their services**

Don't care about IPv6!!!

Reminder

The Future?!
Nobody really knows what the End-users may ask or accept as services by the time a given technology reaches the market.

IPv6 - Key driver for Next Generation Ubiquitous Networking



Business

Innovations

Mobility

The Ubiquitous Internet

Agriculture/Wildlife

Medical

Transportation

Consumer & Services

Manufacturing

e-Nations

Higher Ed./Research

Government Public Sector

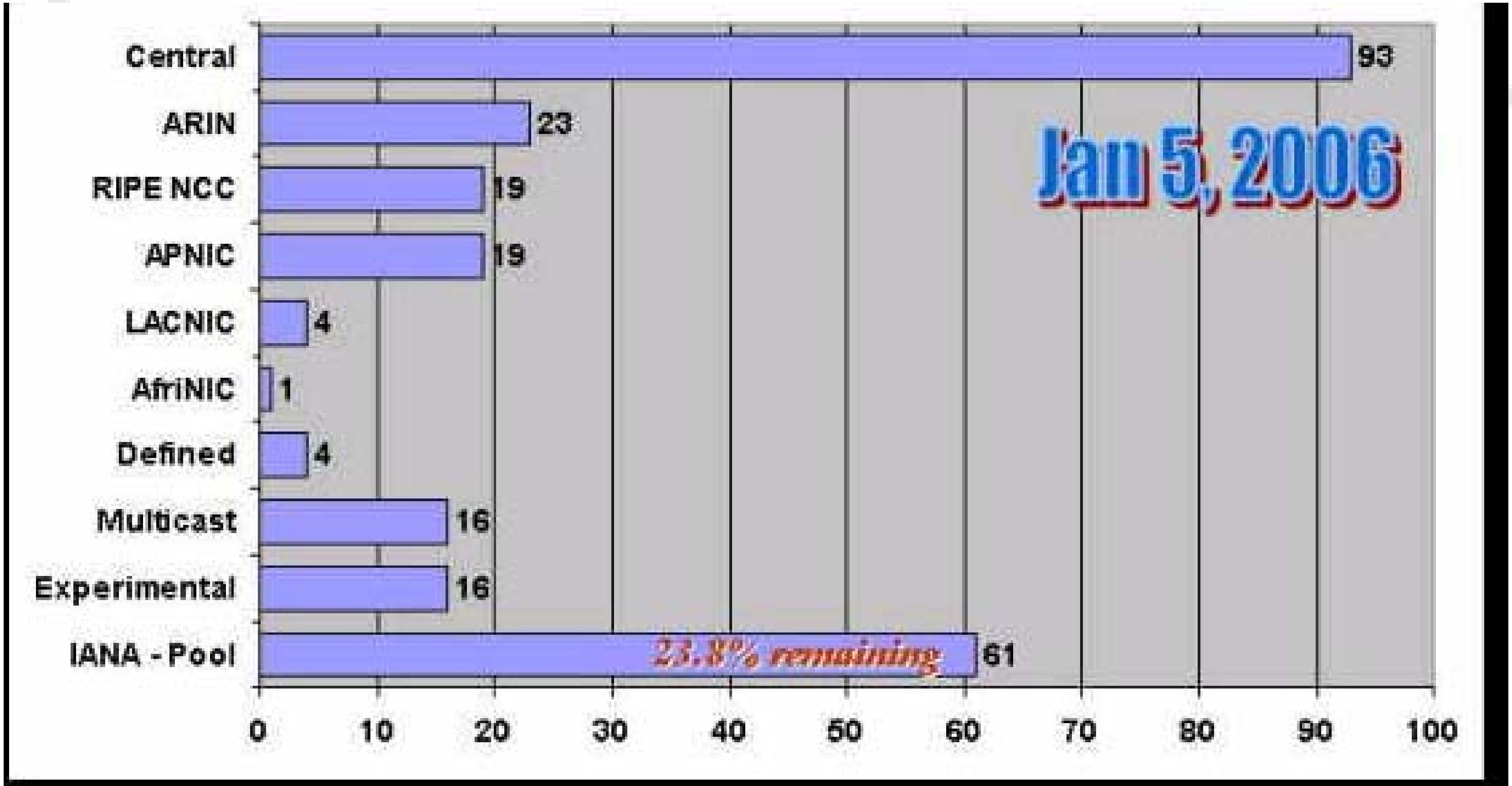
Services at the Edge

ion

Why Not ‘NAT’?

- Exhaustion of address space
- NAT breaks the end-to-end model
- Growth of NAT has slowed down growth of transparent applications
- No easy way to maintain states of NAT in case of node failures
- NAT break security
- NAT complicates mergers, double NATing is needed for devices to communicate with each other
- Multicast through NAT is troublesome

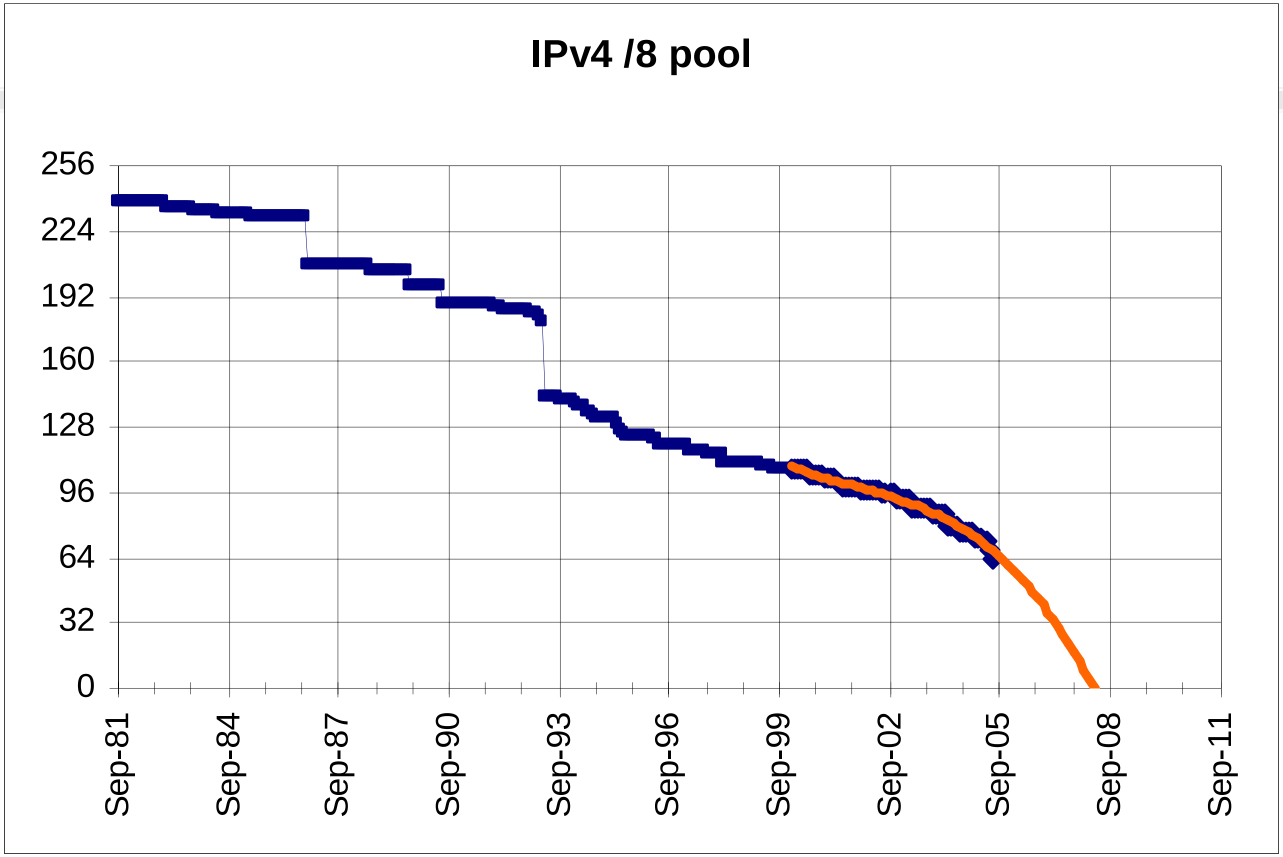
Distribution of IPv4 addresses by /8



IANA allocated 25 /8's between Jan. 1, 2004 and Jan. 5, 2006

typical RIR re-allocation period 9-12 months

Pool exhaustion



Full discussion in the Sept. 2005 issue of the Internet Protocol Journal www.cisco.com/ipj

Do We Really Need a Larger Address Space?

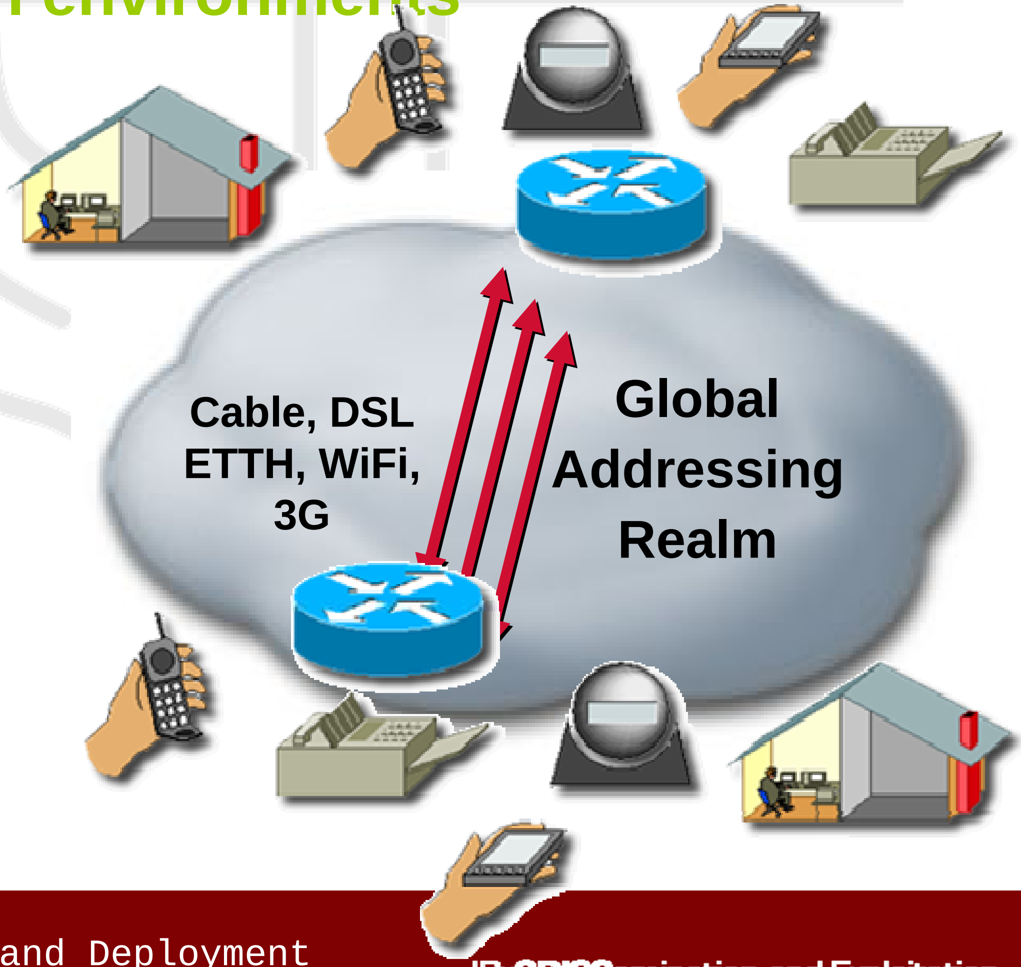
During the life cycle of a technology, a new product is often considered to have reached the early majority – or the mass market – after achieving 22 percent penetration.

- **Internet Population**
  - –~945M by end CY 2004 (source Computer Industry Almanac) – only 10-15%
  - –How to address the future Worldwide population? (~9B in CY 2050)
  - –Emerging Internet countries need address space, eg: China uses nearly 2 class A (11/2002), ~20 class A needed if every student (320M) has to get an IP address
- **Mobile Internet introduces new generation of Internet devices**
  - –PDA (~20M in 2004), Mobile Phones (~1.5B in 2003), Tablet PC
  - –Enable through several technologies, eg: 3G, 802.11,…
- **Transportation – Mobile Networks**
  - –1B automobiles forecast for 2008 – Begin now on vertical markets
  - –Internet access on planes, eg. Lufthansa – trains, eg. Narita express
- **Consumer, Home and Industrial Appliances**

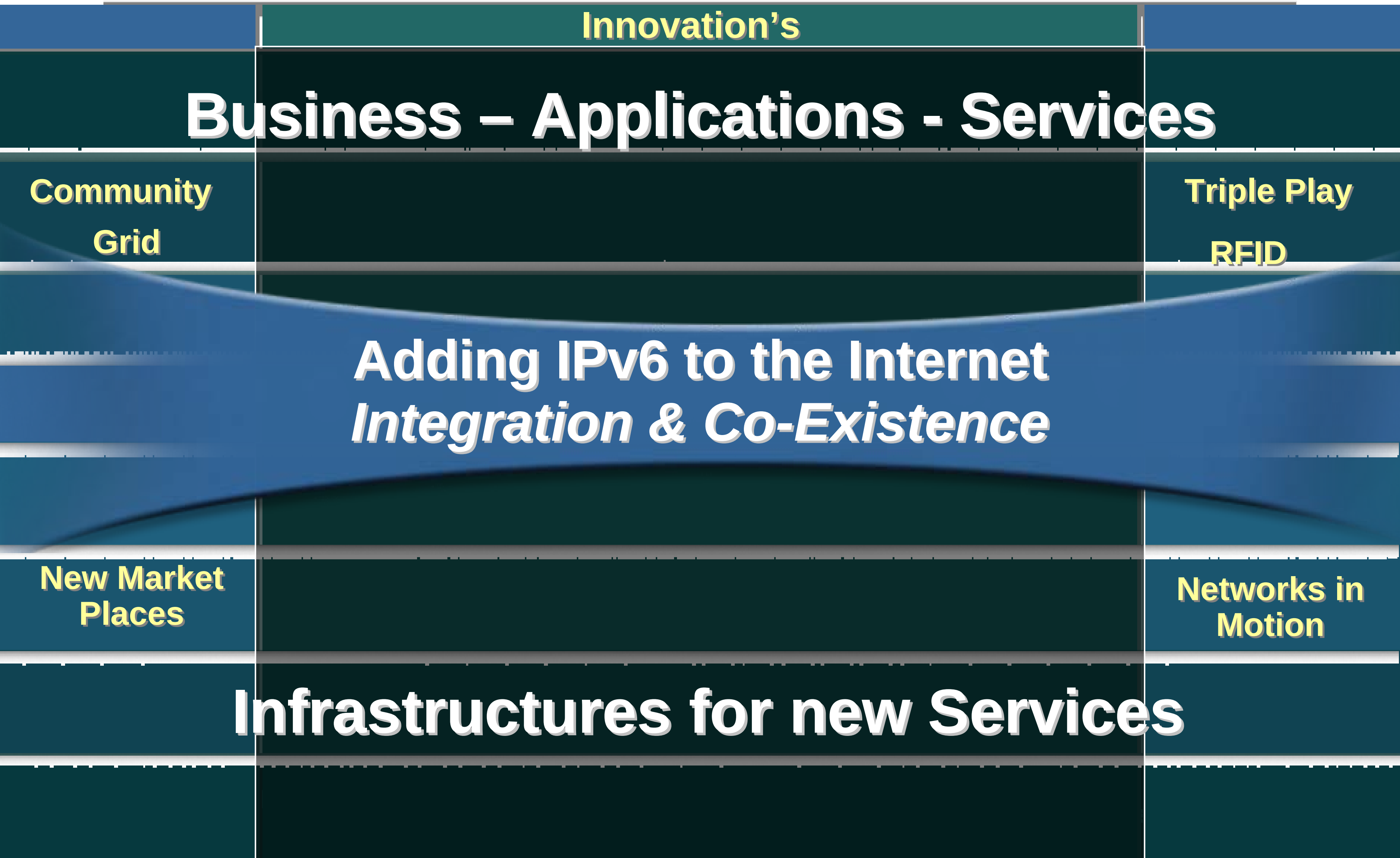
IPv6 Drivers—Network's Architecture

"Always-on" technologies enable new application environments

- Today, Network Address Translation (NAT) and application-layer gateways connect disparate networks
  - •Internet started with end-to-end connectivity for any application
- Peer-to-peer or server-to-client applications mean global addresses
  - •IP telephony, fax, video
  - •Mobility, GRID,
  - •Distributed gaming
  - -Remote monitoring
  - -Instant messaging



Expanding the Internet with IPv6



Innovation’s

Business – Applications - Services

Community

Grid

Triple Play

RFID

Adding IPv6 to the Internet

Integration & Co-Existence

New Market Places

Networks in Motion

Infrastructures for new Services

Broadband Home and IPv6 – a Must!

Home Networking

- IPv6 enables bi-directional reachability for multiple devices, is not intended to a single PC
- Bandwidth increase and symetric access to generate contents
- Easy plug and play

IP Video

Printer

PDA

IP Phone & Fax

Wireless Laptop

- Distance learning
- Video calls
- MP3/MP4 downloads

Wired Devices

- Streaming Video/Audio
- Print/file sharing

Broadband Internet Access

Triple Play Services

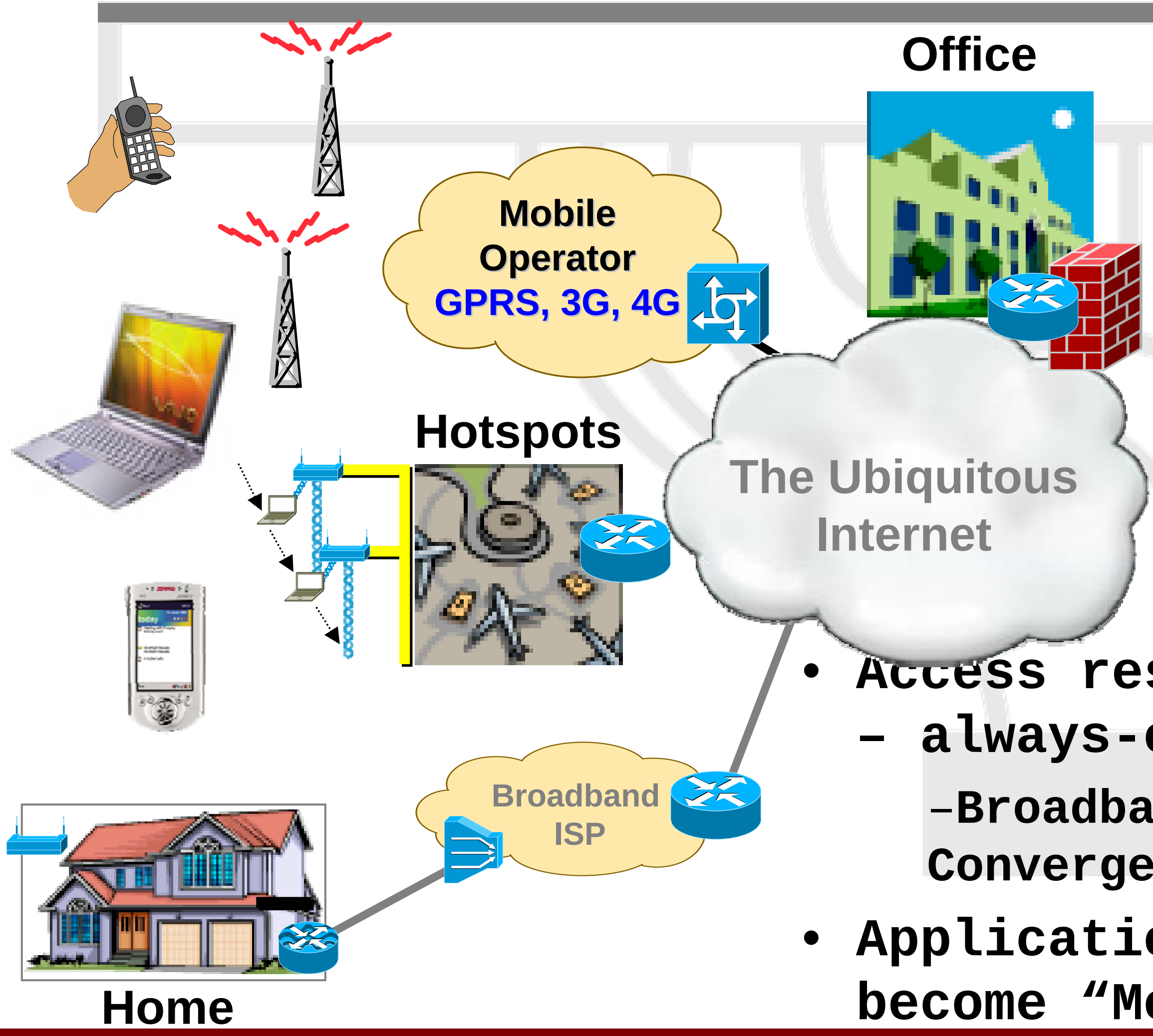
- Multiple devices served in a Home
- Commercial download
- TV guide

Broadband Access Point

- Multiplayer gaming
- Video on demand
- Home security
- Digital audio
- Domestic appliances

Wireless Gaming

IPv6 Mobility Vision



Independent from the Access Technologies

- Unlicensed Band (WiFi,…)
  - Personal mobility
  - high data rate
  - incremental infrastructure
- Licensed Band (GPRS, 3G/4G, WiMax, DVB,…)
  - Full mobility
  - New infrastructure

- Access resources from anywhere – always-on
  - –Broadband/Wireless services Convergence
- Applications and Services to become “Mobile”

IPv6 for the Military

Soldiers
Weapons
Sensors
Command/Control
Logistics

- Massive Address Space (Billions)
  (IP addressed 3d battlespace)
- Mobile IP
- Security/Encryption
- Simplified Management
- Inter-service Interoperability

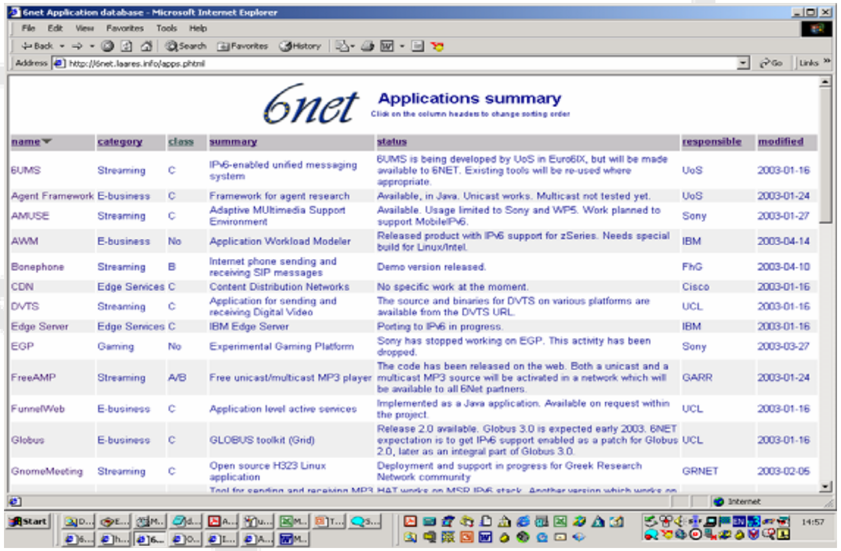
FCS (Future Combat Systems)

WIN-T (Warfighter Information Network – Tactical)

IPv6 Integration – Per Application Model

Today, all O.S. are Dual-Stack

- As soon as the infrastructure is IPv6 capable…IPv6 integration can follow a non-disruptive "per application" model


6net Application database - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites History
Address http://6net.laares.info/apps.phtml
Go Links
Internet
Start
14:57

6net **Applications summary**

Click on the column headers to change sorting order

name	category	class	summary	status	responsible	modified
6UMS	Streaming	C	IPv6-enabled unified messaging system	6UMS is being developed by UoS in Euro6IX, but will be made available to 6NET. Existing tools will be re-used where appropriate.	UoS	2003-01-16
Agent Framework	E-business	C	Framework for agent research	Available, in Java. Unicast works. Multicast not tested yet.	UoS	2003-01-24
AMUSE	Streaming	C	Adaptive MUltimedia Support Environment	Available. Usage limited to Sony and WP5. Work planned to support MobileIPv6.	Sony	2003-01-27
AWM	E-business	No	Application Workload Modeler	Released product with IPv6 support for zSeries. Needs special build for Linux/Intel.	IBM	2003-04-14
Bonephone	Streaming	B	Internet phone sending and receiving SIP messages	Demo version released.	FhG	2003-04-10
CDN	Edge Services	C	Content Distribution Networks	No specific work at the moment.	Cisco	2003-01-16
DVTS	Streaming	C	Application for sending and receiving Digital Video	The source and binaries for DVTS on various platforms are available from the DVTS URL.	UCL	2003-01-16
Edge Server	Edge Services	C	IBM Edge Server	Porting to IPv6 in progress.	IBM	2003-01-16
EGP	Gaming	No	Experimental Gaming Platform	Sony has stopped working on EGP. This activity has been dropped.	Sony	2003-03-27
FreeAMP	Streaming	A/B	Free unicast/multicast MP3 player	The code has been released on the web. Both a unicast and a multicast MP3 source will be activated in a network which will be available to all 6Net partners.	GARR	2003-01-24
FunnelWeb	E-business	C	Application level active services	Implemented as a Java application. Available on request within the project.	UCL	2003-01-16
Globus	E-business	C	GLOBUS toolkit (Grid)	Release 2.0 available. Globus 3.0 is expected early 2003. 6NET expectation is to get IPv6 support enabled as a patch for Globus 2.0, later as an integral part of Globus 3.0.	UCL	2003-01-16
GnomeMeeting	Streaming	C	Open source H323 Linux application	Deployment and support in progress for Greek Research Network community	GRNET	2003-02-05
			Tool for sending and receiving MP3	HAT works on MSR IPv6 stack. Another version which works on		

A Case Study – IP in Schools – Today

- School’s business is Education
  - Reading, Writing, Mathematics and Foreign Languages as foundations to Knowledge
  - The above are minimum end-users requirements to access the Internet
  - Analytic mind is key to value the data retrieved from the Internet
- Schools are part of the Information Society
  - Today, more and more schools get an Internet connection – a Must
  - Lease lines, Broadband Access,…
  - Linked to NRN or local government

- Today, Applications and Services
  - Client-Server: e-mails, web browsing
  - Servers generally hosted externally
  - Most of the time using PAT (a single global IPv4 address)

A Case Study – IPv6 in Schools - Tomorrow

- **Developing new Class of Applications and Services**
  - –Class to Class collaboration – internal to the school, between schools (national & international)
  - – Sharing Database, creating server's,…
  - –Teachers-Students collaboration
  - – "After-time" support, digital pupil desk, foreign languages class,…
  - –Content delivery between schools or Information Providers – Multimedia streaming
  - –IP Telephony between schools
  - –Tele-surveillance – Physical security
  - –Secure Information – Transfer between schools-academy, teachers-school

- **Integrating those services over IPv6**
  - –IPv6 could easily be configured on routers connecting the schools
  - –NRN or Local Government can delegate production IPv6 prefixes to the schools.
- **It can be done Today**
  - –IPv4 applications must not get disturbed, keep IPv4 as it is.

Agenda

- Technology Introduction
- IPv6 Protocol Basics
  - IP Address Space
  - IPv6 header – Extension Headers
  - Addressing
  - MTU
  - IPv6 & DNS
  - Enable IPv6 in operating systems
- IPv6 Protocol Specifics
- IPv6 Transition and Coexistence with IPv4

IPv6 Address Space

IPv4 32-bits

IPv6 128-bits

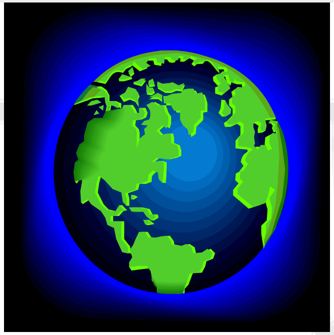
2^{32} = 4,294,967,296

2^{128} = 340,282,366,920,938,463,463,374,607,431,768,211,456

$2^{128} = 2^{32} * 2^{96}$

2^{96} = 79,228,162,514,264,337,593,543,950,336 times the number of possible IPv4 Addresses (79 trillion trillion)

IPv6 Address Space

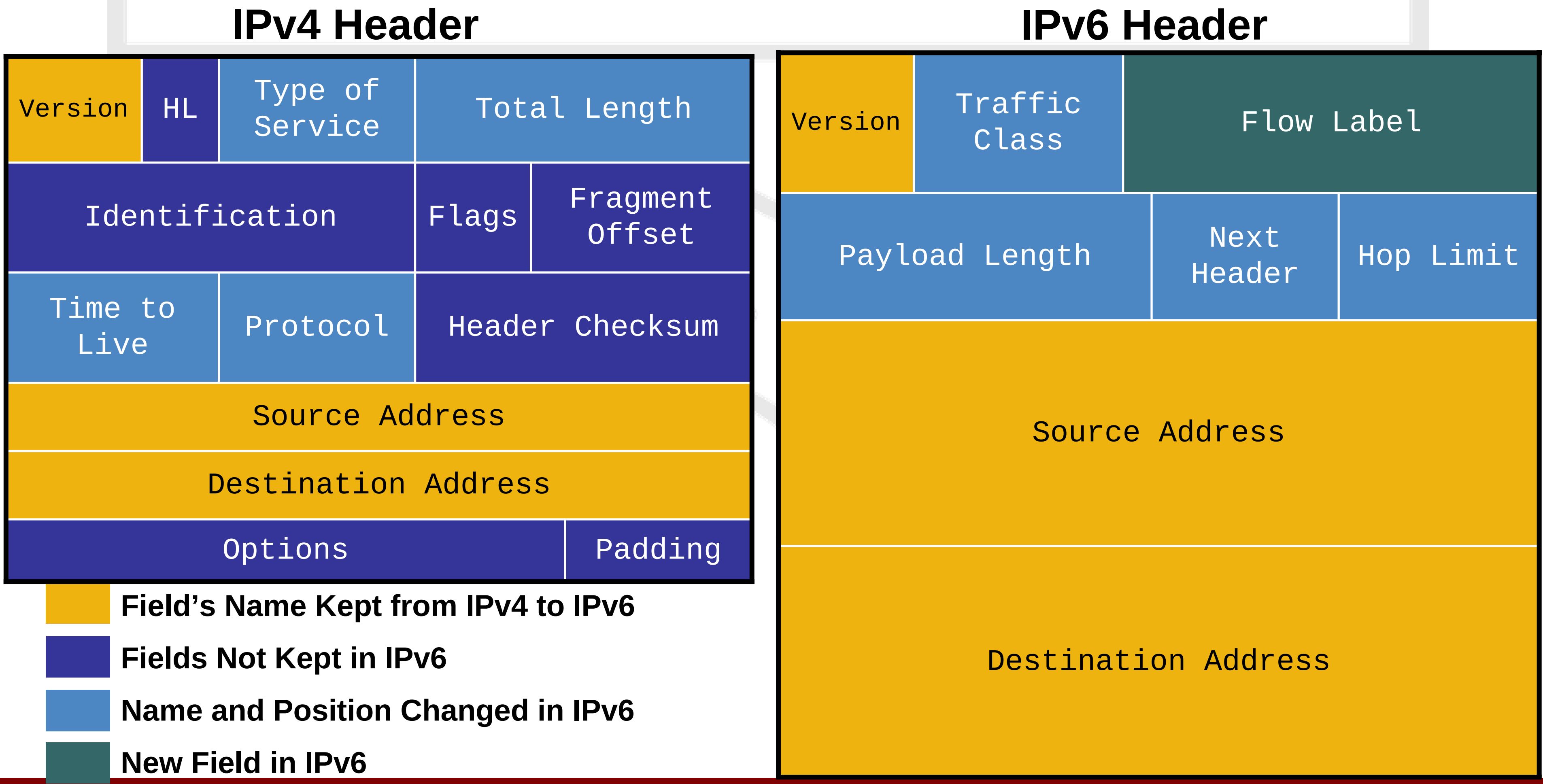


IPv6 Header

- The IPv6 header is redesigned.
- Minimize header overhead and reduce the header process for the majority of the packets.
- Less essential and optional fields are moved to extension headers

IPv6 and IPv4 headers are not *interoperable!*

IPv4 and IPv6 Header Comparison



IPv4 Header

Version | HL | Type of Service | Total Length

Identification | Flags | Fragment Offset

Time to Live | Protocol | Header Checksum

Source Address

Destination Address

Options | Padding

IPv6 Header

Version | Traffic Class | Flow Label

Payload Length | Next Header | Hop Limit

Source Address

Destination Address

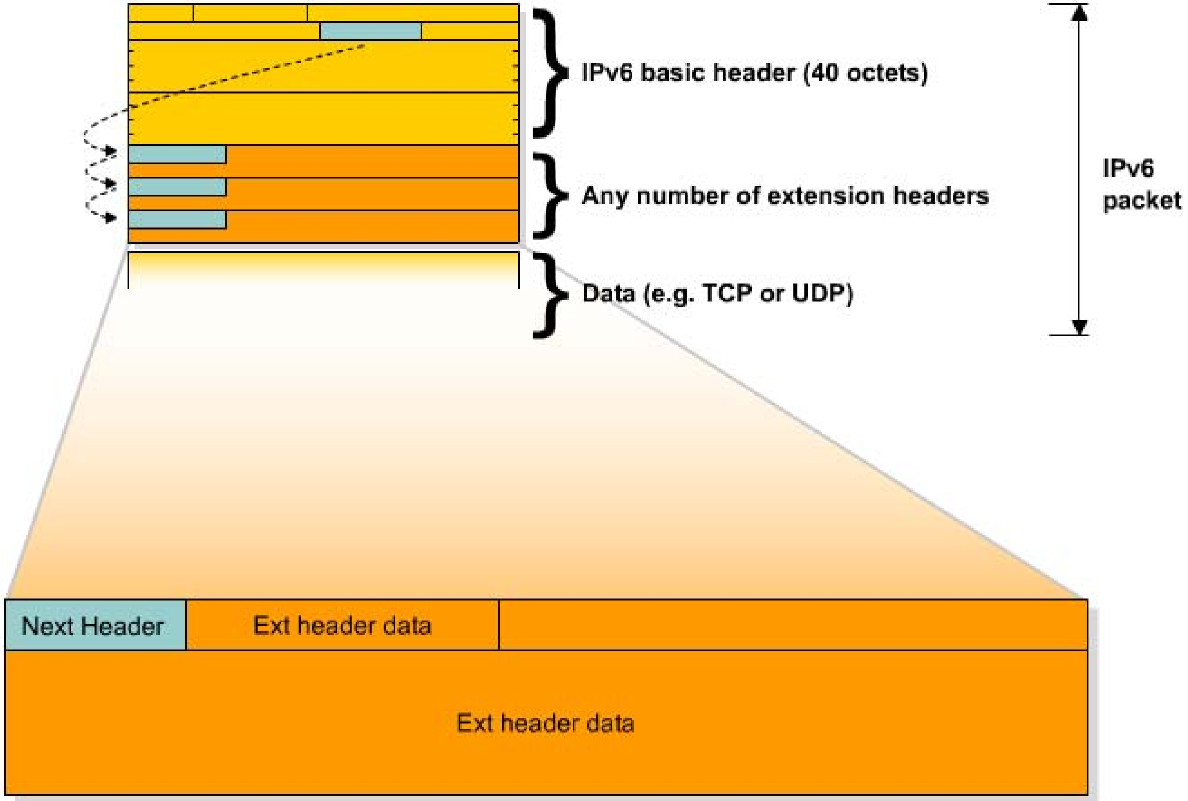
🟨 Field’s Name Kept from IPv4 to IPv6

🟪 Fields Not Kept in IPv6

🟦 Name and Position Changed in IPv6

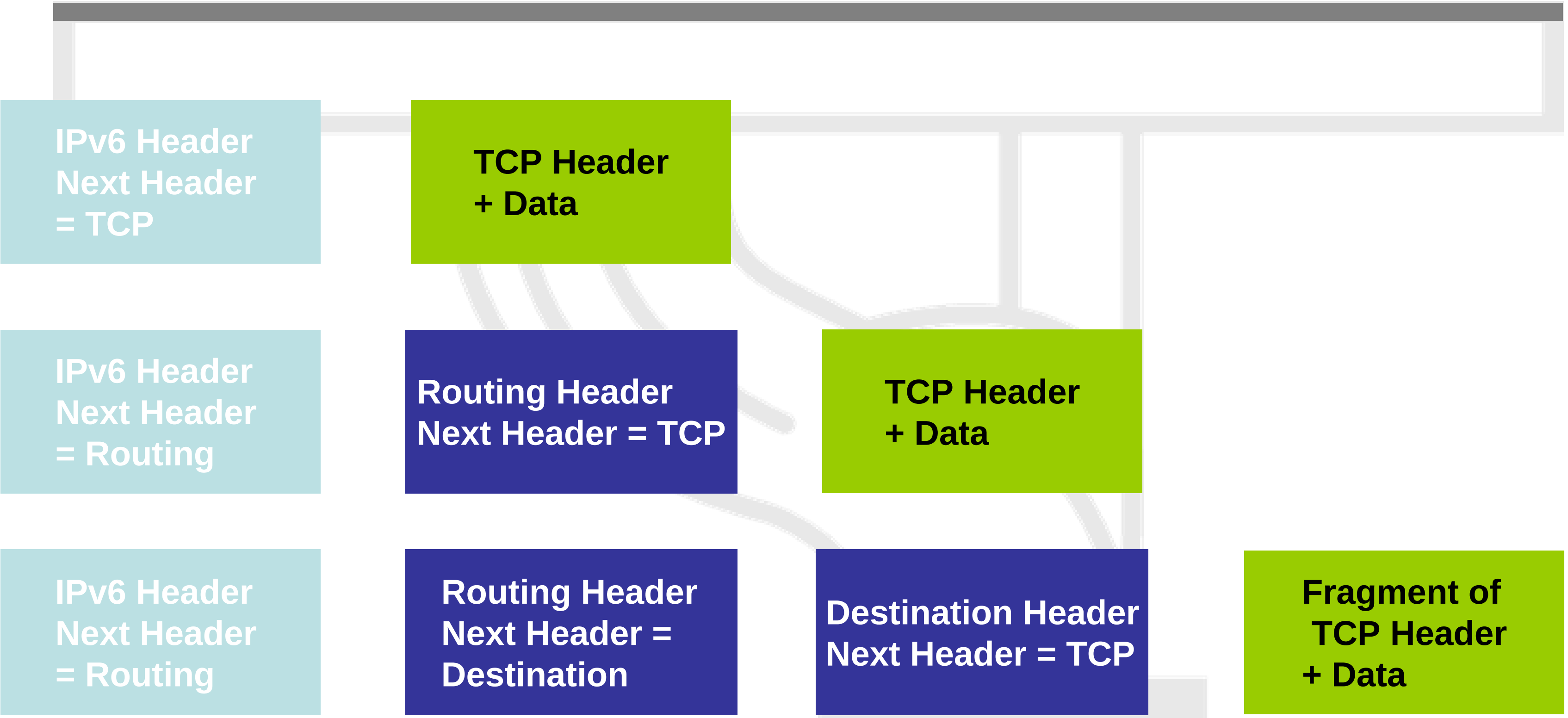
🟩 New Field in IPv6

Extension Headers (RFC2460)



- Processed only by node identified in IPv6 Destination Address field => much lower overhead than IPv4 options
  - exception: Hop-by-Hop Options header
- Eliminated IPv4's 40-octet limit on options
  - In IPv6, limit is total packet size, or Path MTU in some cases

Extension Headers



IPv6 Header
Next Header
= TCP

TCP Header
+ Data

IPv6 Header
Next Header
= Routing

Routing Header
Next Header = TCP

TCP Header
+ Data

IPv6 Header
Next Header
= Routing

Routing Header
Next Header =
Destination

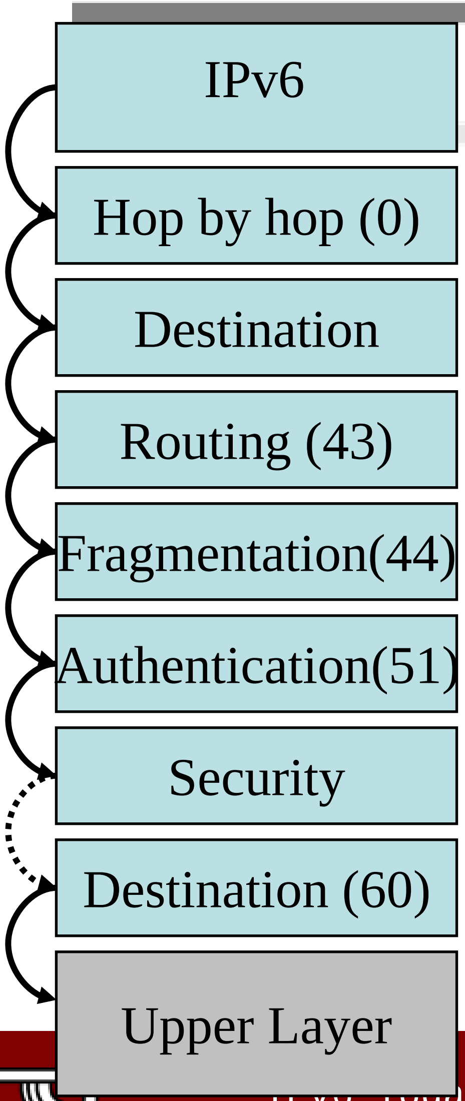
Destination Header
Next Header = TCP

Fragment of
TCP Header
+ Data

Extension Headers Are Daisy Chained

IPv6 extension headers: order is important

RFC 2460



IPv6

Hop by hop (0)

Destination

Routing (43)

Fragmentation(44)

Authentication(51)

Security

Destination (60)

Upper Layer

} Processed by every router

} Processed by routers listed in Routing extension

} List of routers to cross

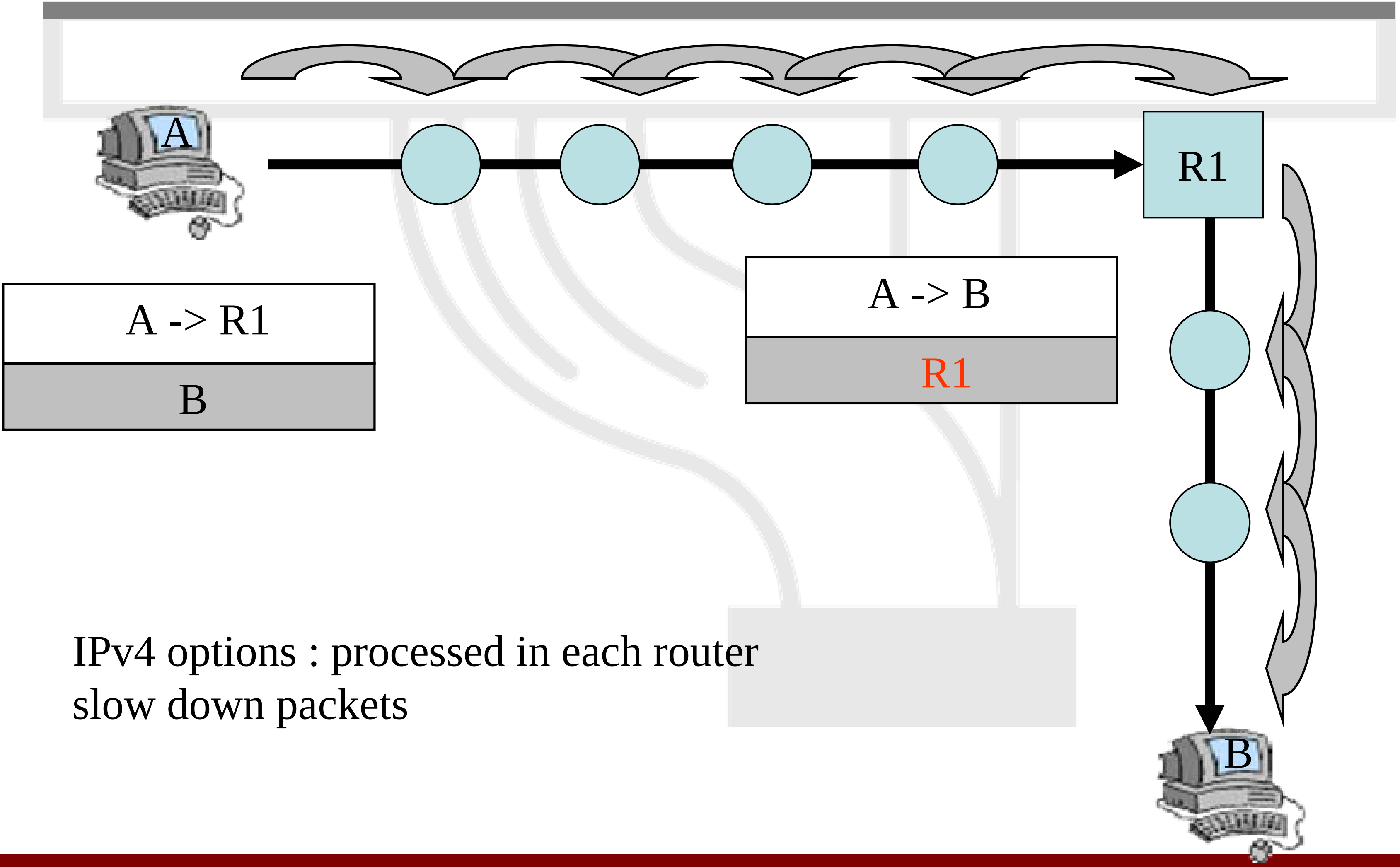
} Processed by the destination

} After reassembling the packet

} Cipher the content of the remaining information

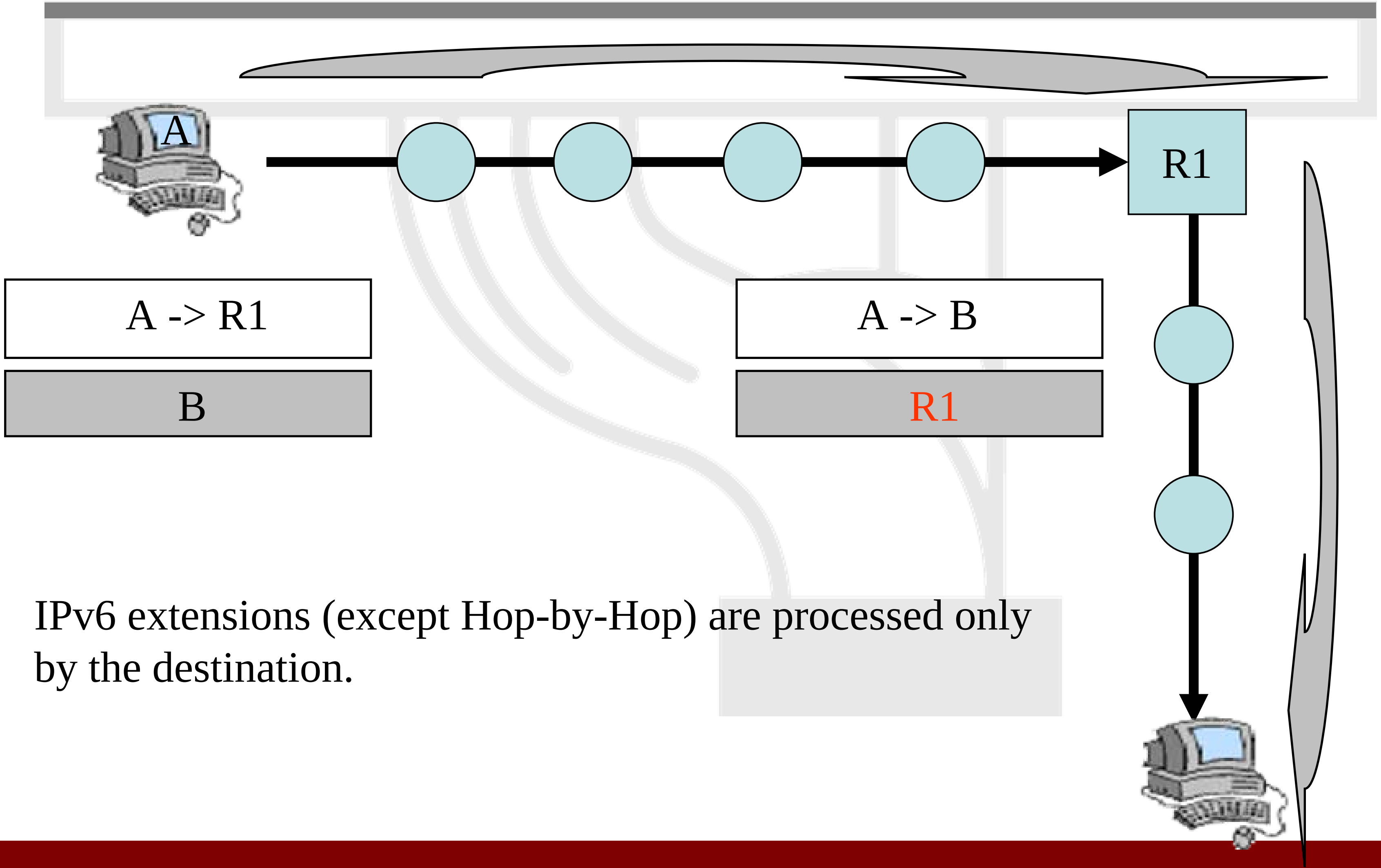
} Processed **only** by the destination

v4 options vs. v6 extensions



IPv4 options : processed in each router
slow down packets

v4 options vs. v6 extensions



IPv6 extensions (except Hop-by-Hop) are processed only by the destination.

IPv6 Address Representation (Example)

- Base format (16-byte)

2001:0660:3003:0001:0000:0000:6543:210F

- Compact Format:

2001:660:3003:1::6543:210F

- Litteral representation
  - [2001:660:3003:2:a00:20ff:fe18:964c]

IPv6 Addressing

Prefix Representation

- Representation of prefix is just like CIDR [address prefix / prefix length]
- In this representation you attach the prefix length
  - IPv4 address: 198.10.0.0/16
  - IPv6 address: 3ef8:ca62:12FE::/48

IPv6 Address Representation

- **Loopback address representation**
  - 0:0:0:0:0:0:0:1=> ::1
  - Same as 127.0.0.1 in IPv4
  - Identifies self

- **Unspecified address representation**
  - 0:0:0:0:0:0:0:0=> ::
  - Used as a placeholder when no address available
  - (Initial DHCP request, Duplicate Address Detection DAD)

IPv6 Address Representation

- **IPv4 mapped**
  - 0:0:0:0:0::FFFF:IPv4 = ::FFFF:IPv4
  - 0:0:0:0:0:FFFF:192.168.30.1 = ::FFFF:C0A8:1E01
- **IPv4 compatible**
  - 0:0:0:0:0:0:IPv4 = ::IPv4
  - 0:0:0:0:0:0:192.168.30.1 = ::192.168.30.1 = ::C0A8:1E01

IPv6 Addressing Architecture

- IPv6 Addressing rules are covered by multiples RFC’s
  - –Architecture initially defined by RFC 2373
  - –Now RFC rfc4291.txt (obsoletes 3513 which obsoletes RFC 2373)
- Address Types are :
  - –Unicast : One to One (Global, Link local, Site local, Compatible)
  - –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)

IPv6 - Addressing Model

Addresses are assigned to interfaces

change from IPv4 model :

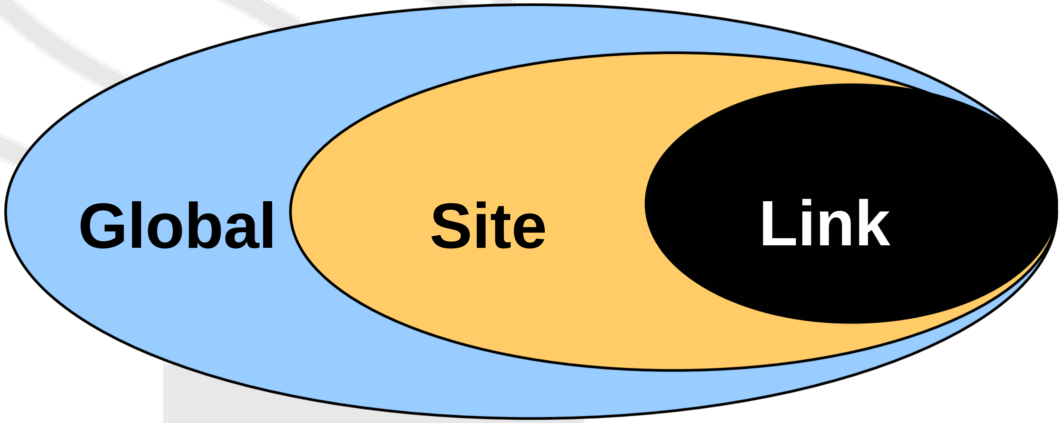
Interface 'expected' to have multiple addresses

Addresses have scope

Link Local

Site Local

Global

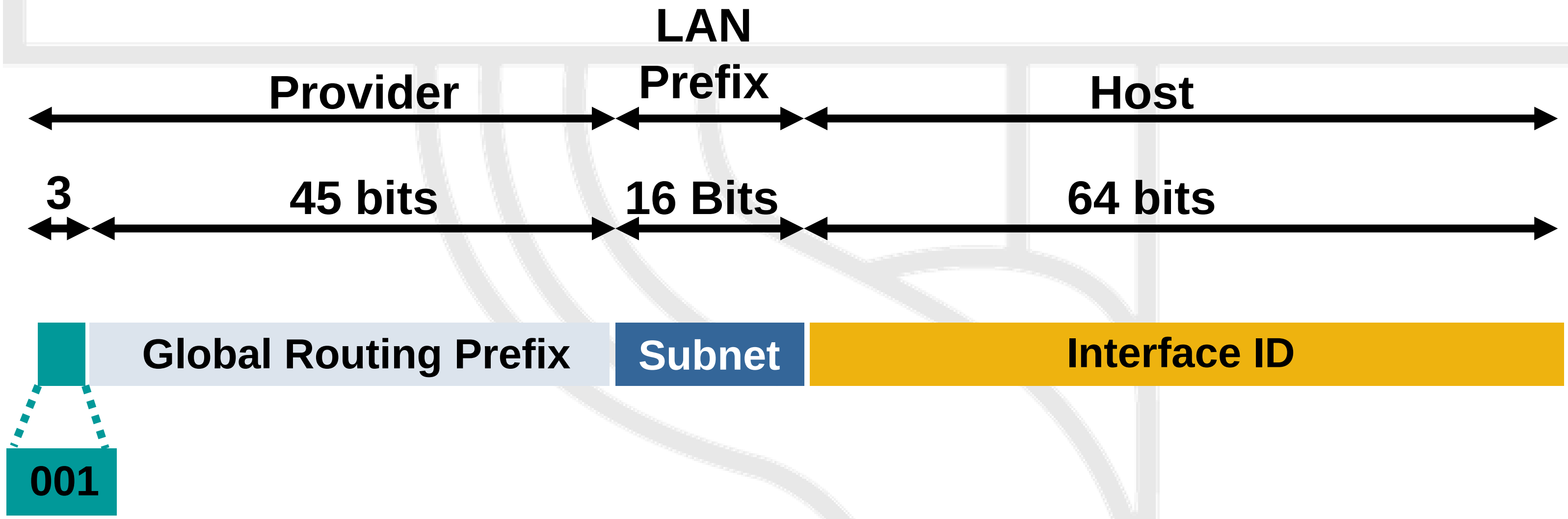


Site-Local Address Deprecated in RFC 3879 now it is Unique Local Address (ULA) RFC 4193

Addresses have lifetime

Valid and Preferred lifetime

Aggregatable Global Unicast Addresses



LAN Prefix

Provider

Host

3

45 bits

16 Bits

64 bits

Global Routing Prefix

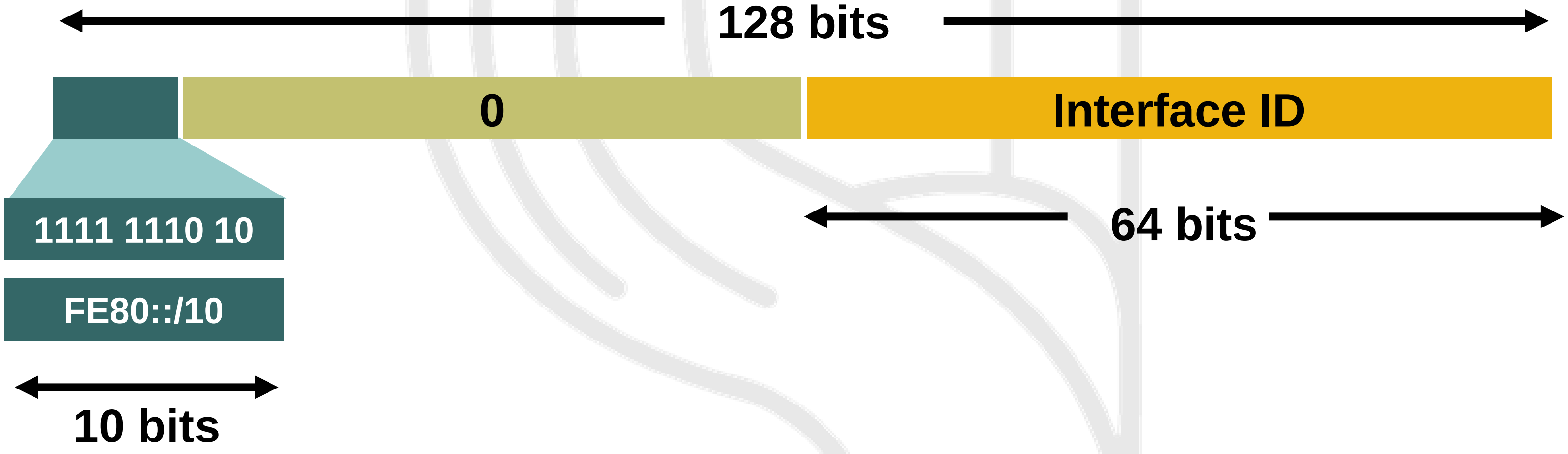
Subnet

Interface ID

001

- Aggregatable global unicast addresses are:
  - Addresses for generic use of IPv6
  - Structured as a hierarchy to keep the aggregation
- See RFC 4291

Link-Local



- Link-local addresses:
  - Have a limited scope of the link
  - Are automatically configured with the interface ID

Link-Local

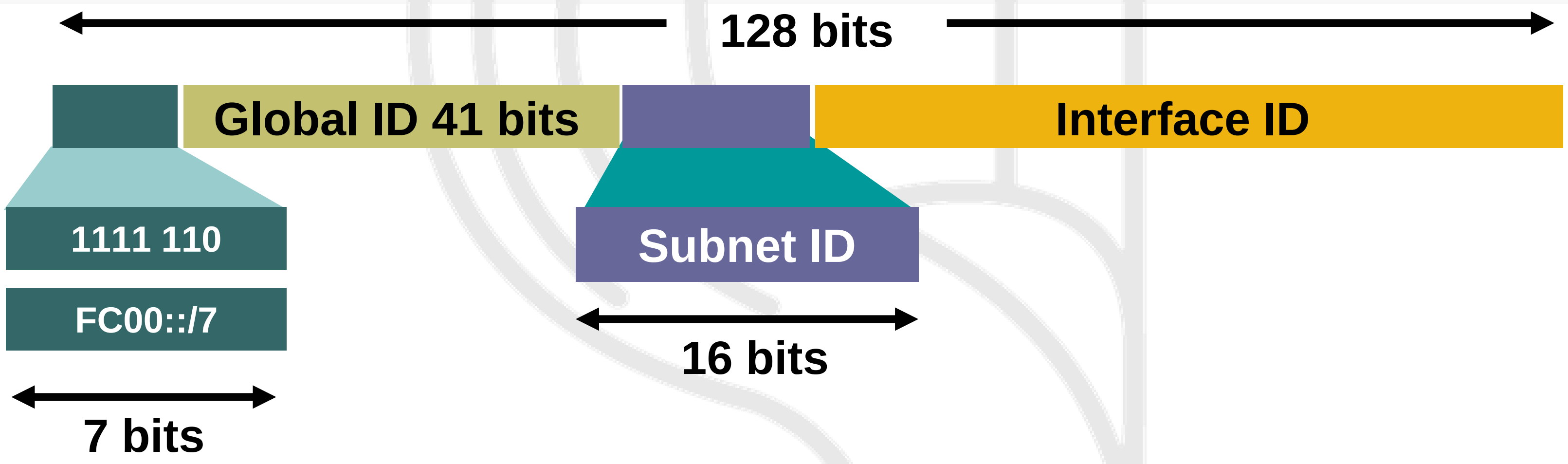
Aggregatable Address

2001::4: 204:9AFF:FEAC:7D80

Link-Local Address

FE80:0:0:0 204:9AFF:FEAC:7D80

Unique-Local

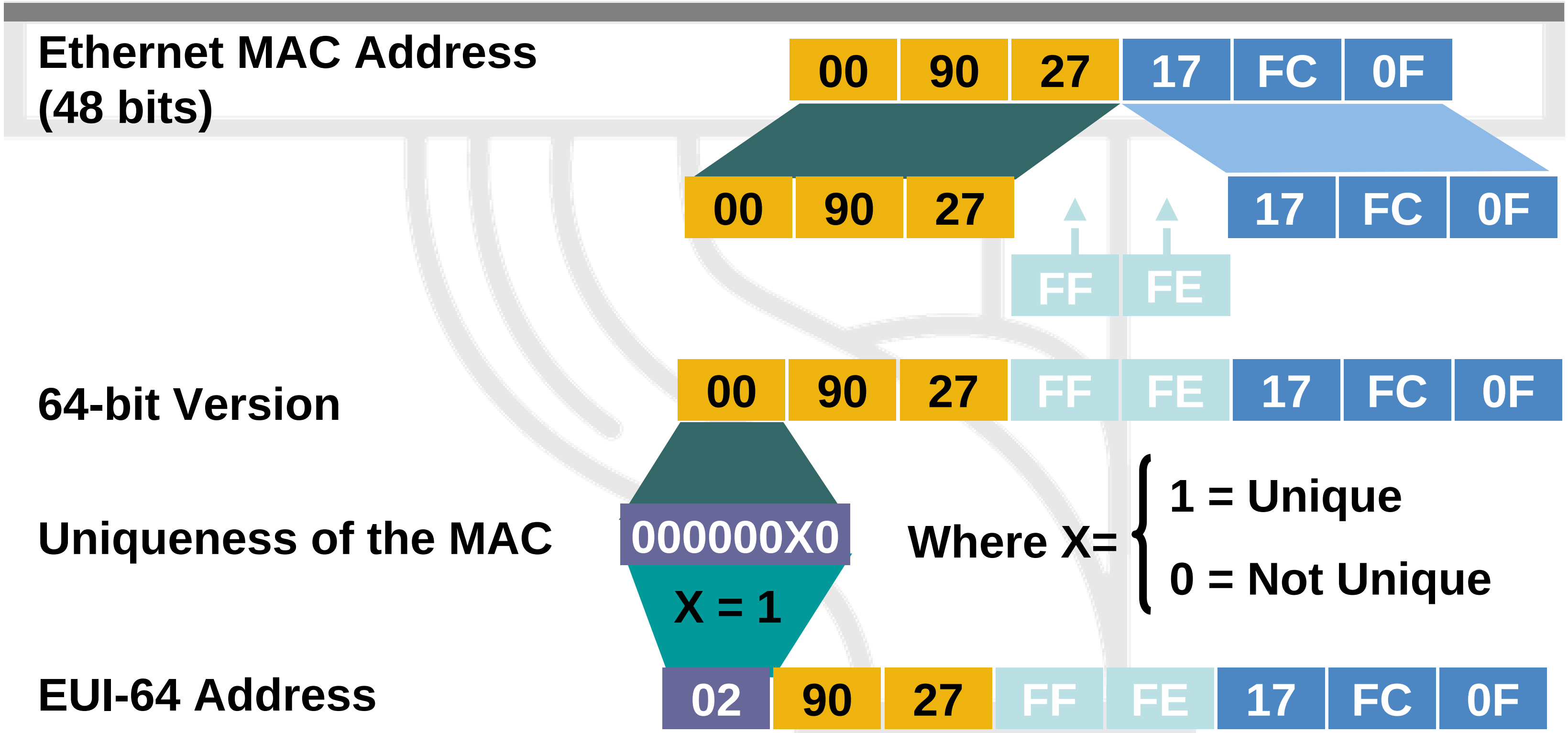


- Unique-local addresses used for:
  - Local communications
  - Inter-site VPNs
  - Not routable on the Internet

Aggregatable Global Unicast Addresses

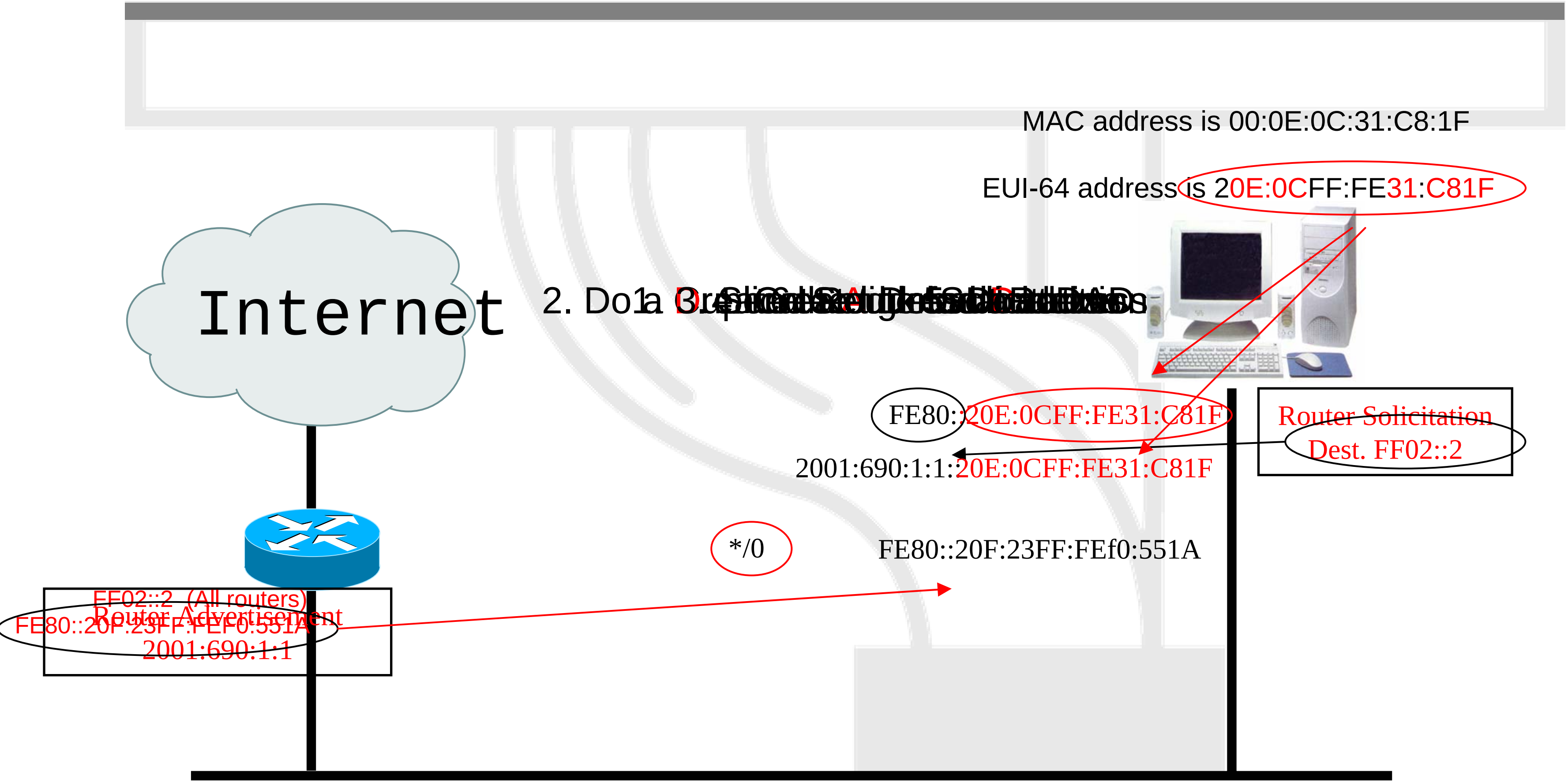
- Lowest-order 64-bit field of unicast addresses may be assigned in several different ways:
  - – Manually configured
  - – Auto-configured from a 64-bit EUI-64, or expanded from a 48-bit MAC address (e.g. Ethernet address)
  - – Auto-generated pseudo-random number (to address privacy concerns)
  - –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

EUI-64

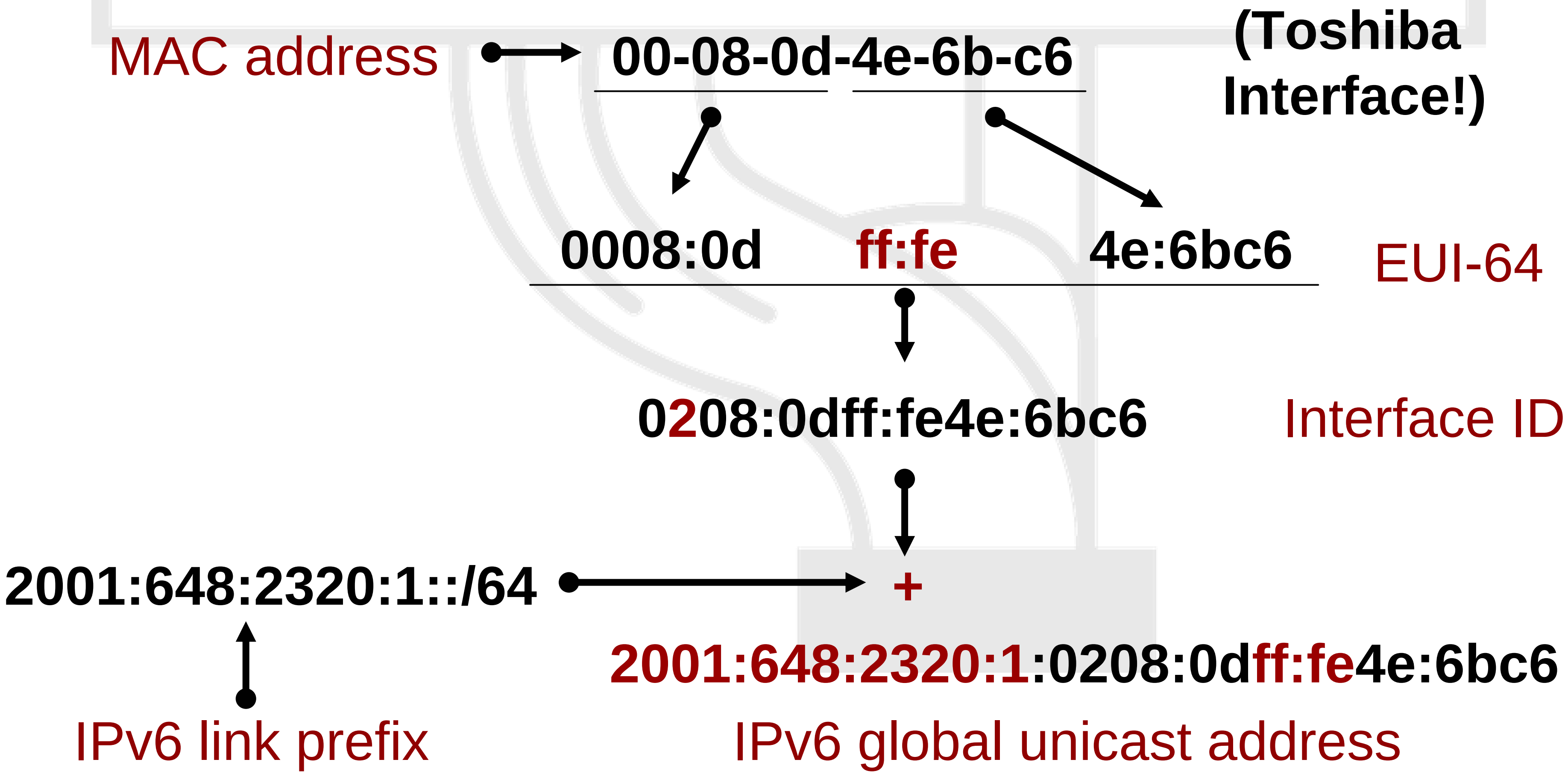


- EUI-64 address is formed by inserting "FFFE" and ORing a bit identifying the uniqueness of the MAC address

Stateless Autoconfiguration example



Interface Identifier: Example



MAC address → 00-08-0d-4e-6b-c6

(Toshiba Interface!)

0008:0d ff:fe 4e:6bc6

EUI-64

0208:0dff:fe4e:6bc6

Interface ID

2001:648:2320:1::/64 → +

IPv6 link prefix

2001:648:2320:1:0208:0dff:fe4e:6bc6

IPv6 global unicast address

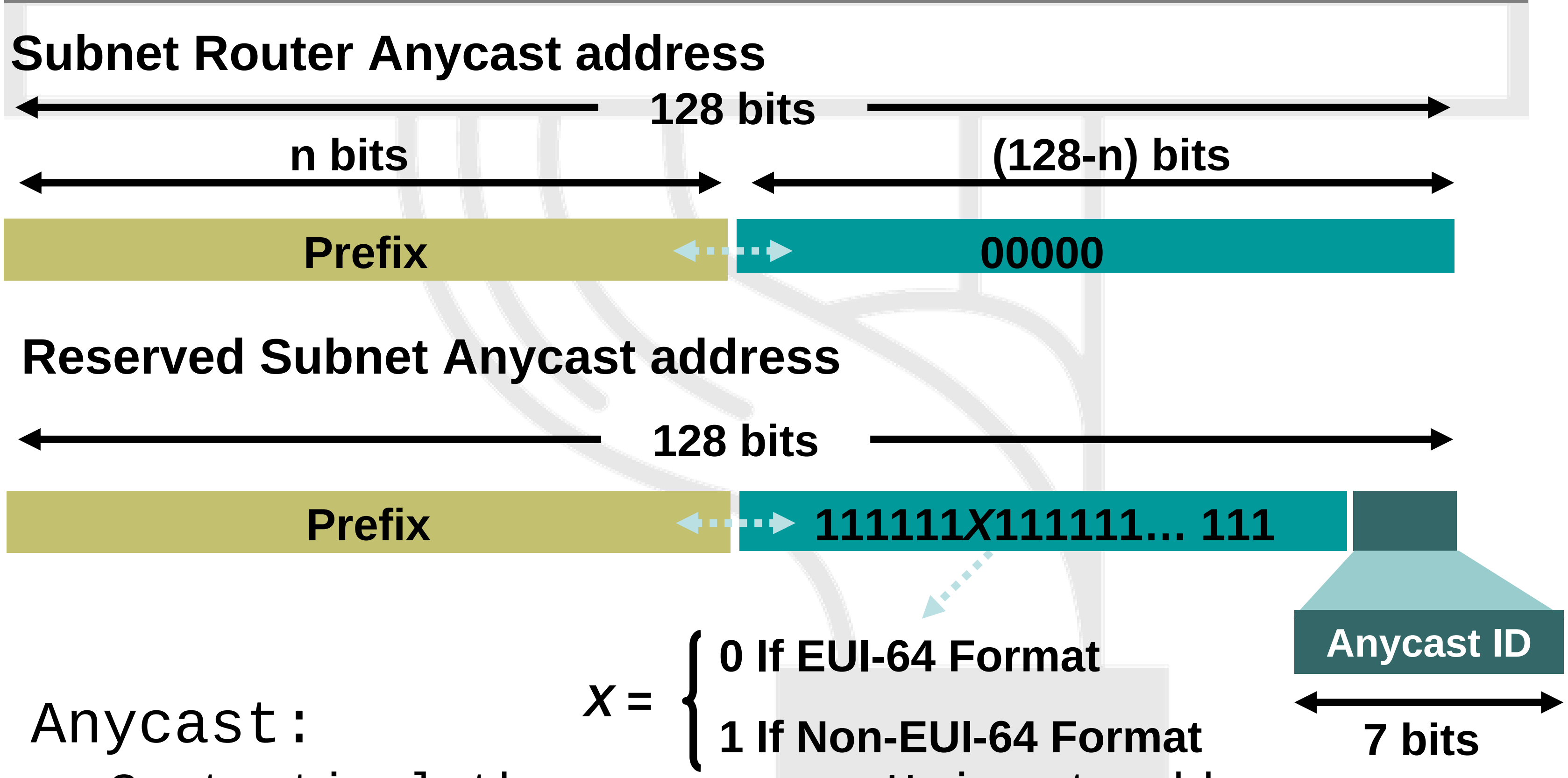
Anycast Addresses (RFC 3513)

- «Anycast addresses allow a packet to be **routed to one of a number** of different nodes all responding to the same address »
- «Anycast addresses are taken from the unicast address spaces (of any scope) and are not syntactically distinguishable from unicast addresses … it may be assigned to an IPv6 router only »

Anycast Addresses (RFC 3513)

- Anycast address …
  - … can not be a used as a source address of an IPv6 packet
  - … must be assigned only to routers
- Reserved anycast addresses are defined in RFC 2526

Anycast Address



Subnet Router Anycast address

128 bits

n bits

(128-n) bits

Prefix

00000

Reserved Subnet Anycast address

128 bits

Prefix

111111*X*111111… 111

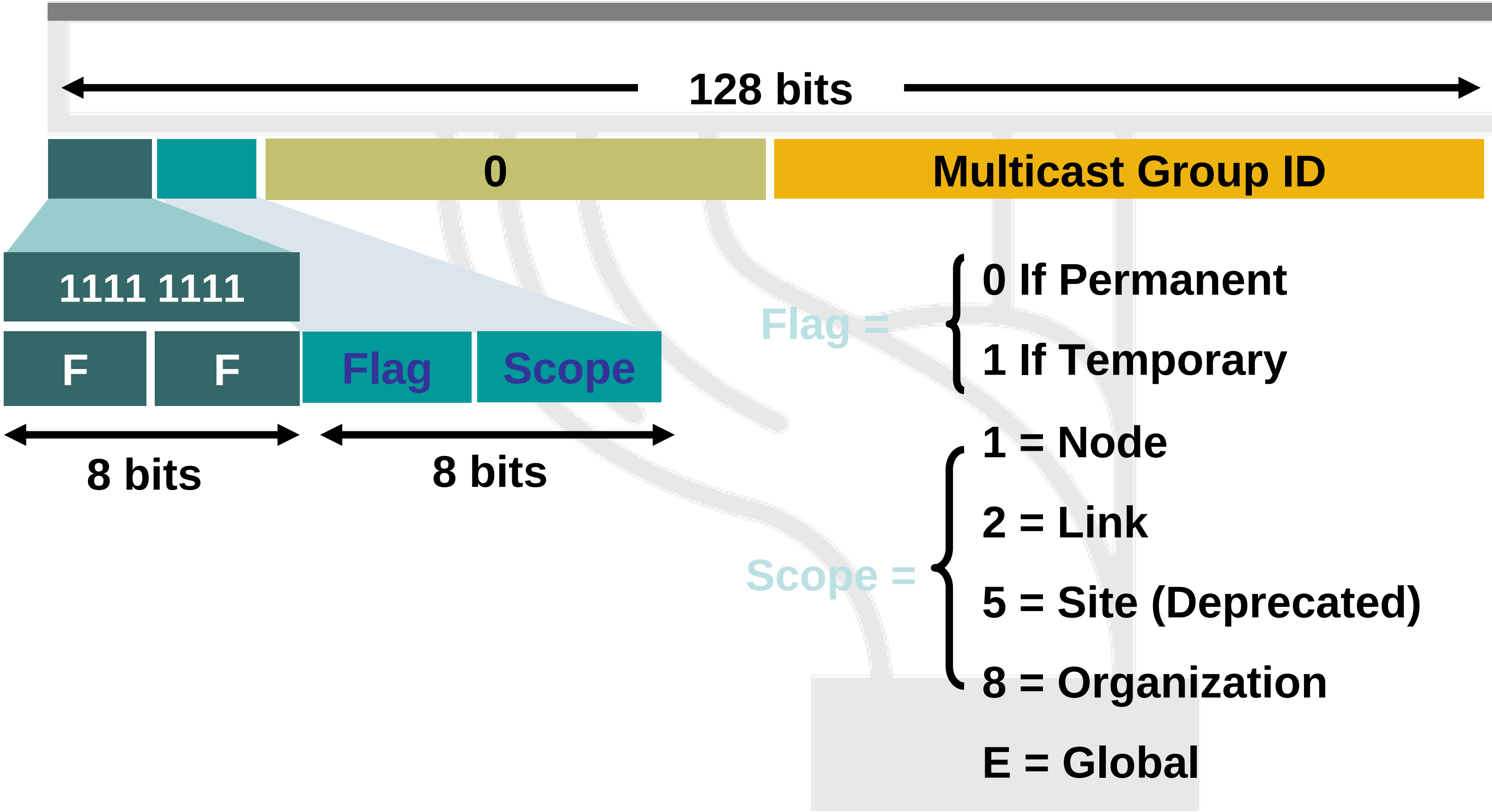
Anycast ID

7 bits

$$X = \begin{cases} 0 \text{ If EUI-64 Format} \\ 1 \text{ If Non-EUI-64 Format} \end{cases}$$

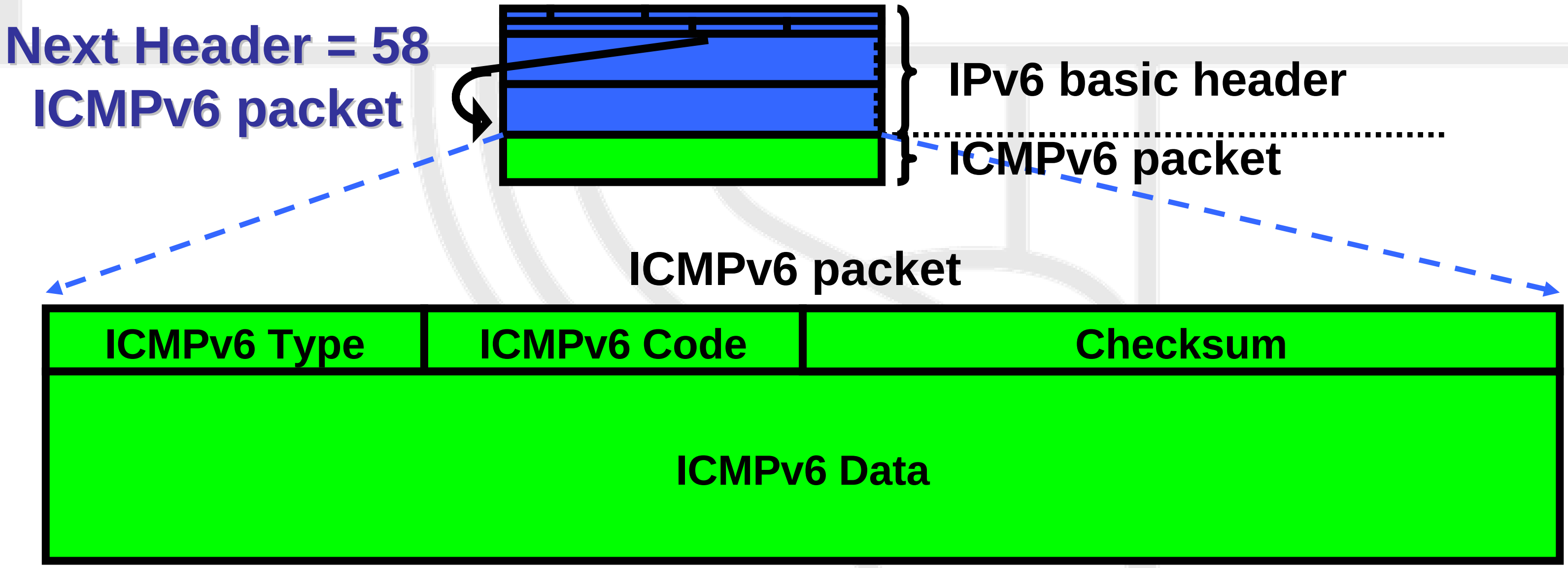
- Anycast:
  - Syntactical the same as a Unicast address
  - Is one-to-nearest type of address
  - Has a current limited use

Multicast



- Multicast is used in the context of one-to-many
- A multicast scope is new in IPv6

ICMPv6



- ICMPv6 (RFC 2463 DS) "Super" Protocol that :
  - Covers ICMP (v4) features (Error control, Administration, …)
  - Transports ND messages
  - Transports MLD messages (Queries, Reports, …)

DNS Extensions for IPv6

RFC 1886→ RFC 3596 (upon successful interoperability tests)

AAAA : forward lookup ('Name IPv6 → Address'):

Equivalent to '**A**' record

Example:

```
ns3.nic.fr.             IN      A       192.134.0.49
                        IN      AAAA    2001:660:3006:1::1:1
```

PTR : reverse lookup ('IPv6 Address → Name'):

Reverse tree equivalent to in-addr.arpa

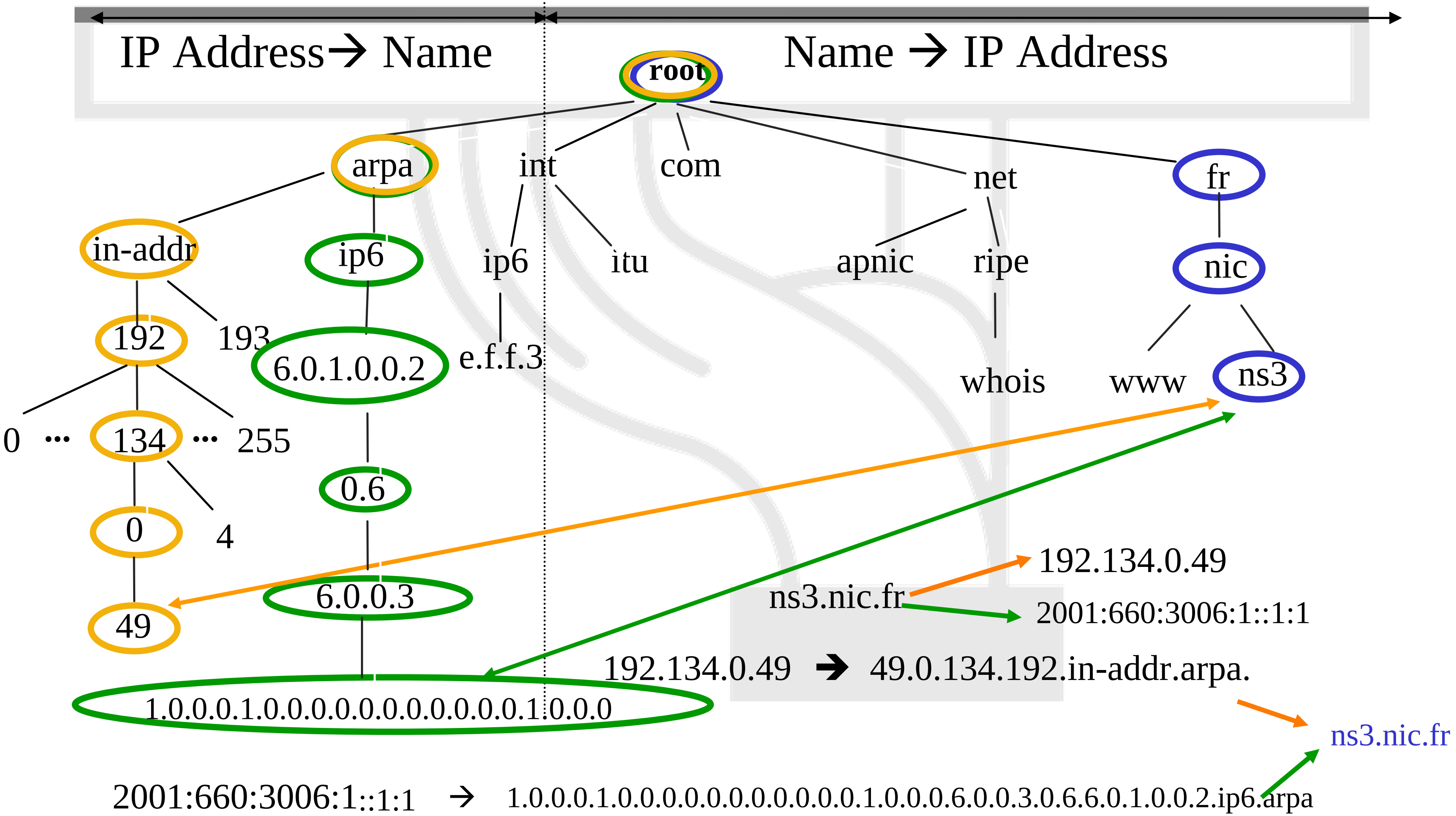
New tree: **ip6.arpa** (under deployment)

Former tree: **ip6.int** (deprecated)

Example:

```
$ORIGIN 1.0.0.0.6.0.0.3.0.6.6.0.1.0.0.2.ip6.arpa.
  1.0.0.0.1.0.0.0.0.0.0.0.0.0.0.0       PTR
  ns3.nic.fr.
```

Lookups in an IPv6-aware DNS Tree



IP Address→ Name

Name → IP Address

ns3.nic.fr → 192.134.0.49

ns3.nic.fr → 2001:660:3006:1::1:1

192.134.0.49 → 49.0.134.192.in-addr.arpa. → ns3.nic.fr

2001:660:3006:1::1:1 → 1.0.0.0.1.0.0.0.0.0.0.0.0.0.0.0.0.1.0.0.0.6.0.0.3.0.6.6.0.1.0.0.2.ip6.arpa → ns3.nic.fr

About Required IPv6 Glue in DNS Zones

When the DNS zone is delegated to a DNS server (among others) contained in the zone itself

Example: In zone file rennes.enst-bretagne.fr

```
@         IN        SOA        rsm.rennes.enst-bretagne.fr. fradin.rennes.enst-
    bretagne.fr.
                    (2005040201 ;serial
                    86400     ;refresh
                    3600      ;retry
                    3600000   ;expire}

                    IN        NS        rsm
                    IN        NS        univers.enst-bretagne.fr.
[…]
ipv6                IN        NS        rhadamanthe.ipv6
          IN        NS        ns3.nic.fr.
          IN        NS        rsm
;
rhadamanthe.ipv6              IN        A         192.108.119.134
                              IN        AAAA      2001:660:7301:1::1
[…]
```

IPv4 glue (A 192.108.119.134) is required to reach rhadamanthe over IPv4 transport

IPv6 glue (AAAA 2001:660:7301:1::1) is required to reach rhadamanthe over IPv6 transport

IPv6 DNS and root servers

- DNS root servers are critical resources!
- 13 roots « around » the world (#10 in the US)
- Not all the 13 servers already have IPv6 enabled and globally reachable via IPv6.
- Need for (mirror) root servers to be installed in other locations (EU, Asia, Africa, …)
- New technique : anycast DNS server
  - To build a clone from the master/primary server
  - Containing the same information (files)
  - Using the same IP address
- Such anycast servers have already begun to be installed :
  - F root server: Ottawa, Paris(Renater), Hongkong, Lisbon (FCCN)…
  - Look at http://www.root-servers.org for the complete and updated list.

Path MTU discovery (RFC 1981)

Derived from RFC 1191, (IPv4 version of the protocol)

- **Path** : set of links followed by an IPv6 packet between source and destination
- **Link MTU** : maximum packet length (bytes) that can be transmitted on a given link without fragmentation
- **Path MTU (or pMTU)** : min { link MTUs } for a given path
- **Path MTU Discovery** : automatic pMTU discovery for a given path

Path MTU discovery (2)

- Protocol operation
  - makes assumption that pMTU = link MTU to reach a neighbor (first hop)
  - if there is an intermediate router such that link MTU < pMTU ➔ it sends an ICMPv6 message: "Packet size Too Large"
  - source reduces pMTU by using information found in the ICMPv6 message

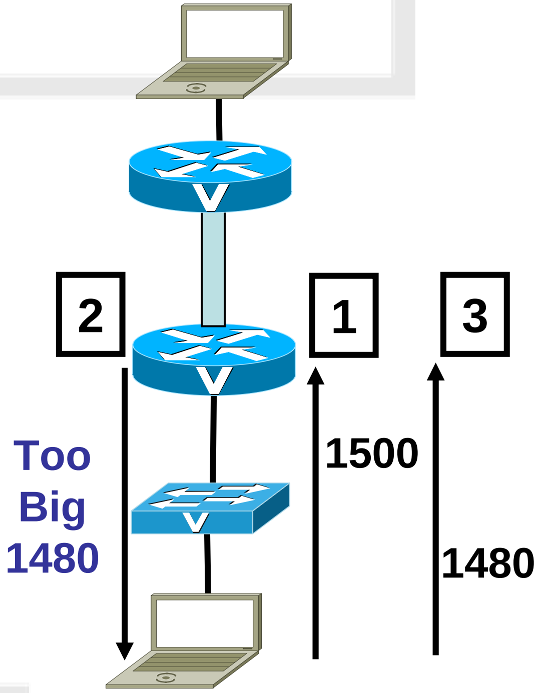
=> Intermediate equipments aren’t allowed to perform packet fragmentation

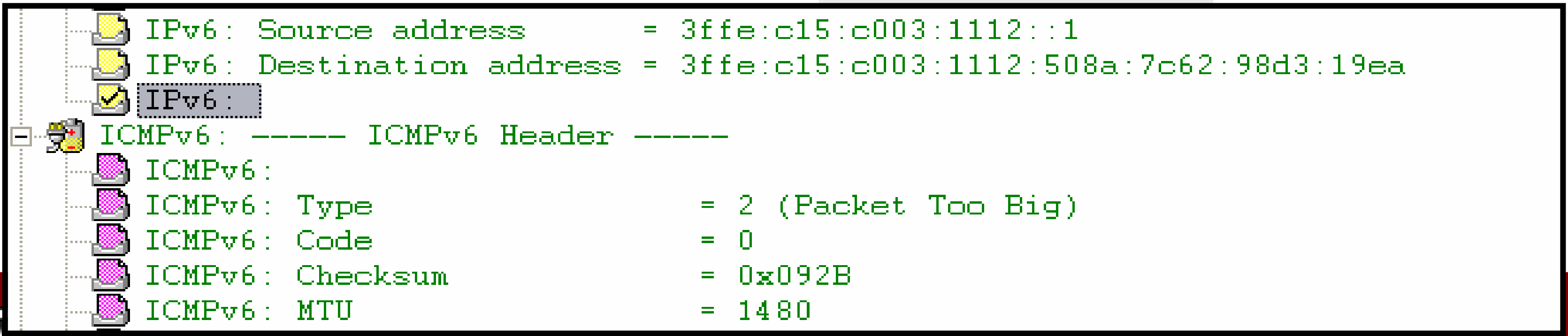
Path MTU Discovery

```
D:\>ping -l 1500 toshiba-redhat                                   1

Pinging toshiba-redhat [3ffe:c15:c003:1114:210:a4ff:fec7:5fcf]
                                                                  2
Request timed out.
Reply from 3ffe:c15:c003:1114:210:a4ff:fec7:5fcf : time=3ms       3
Reply from 3ffe:c15:c003:1114:210:a4ff:fec7:5fcf : time=3ms
Reply from 3ffe:c15:c003:1114:210:a4ff:fec7:5fcf : time=3ms

netsh interface ipv6>show destinationcache
Interface 6: LAN
PMTU Destination Address                    Next Hop Address
---- -------------------------------------- -------------------------
1480 3ffe:c15:c003:1112::1                  3ffe:c15:c003:1112::1
```





```
IPv6: Source address       = 3ffe:c15:c003:1112::1
IPv6: Destination address  = 3ffe:c15:c003:1112:508a:7c62:98d3:19ea
IPv6:
ICMPv6: ------ ICMPv6 Header ------
  ICMPv6:
  ICMPv6: Type                 = 2 (Packet Too Big)
  ICMPv6: Code                 = 0
  ICMPv6: Checksum             = 0x092B
  ICMPv6: MTU                  = 1480
```

IPv6 Support: Windows

- WinXP
  - SP0: Autoconfiguration, tunnels, ISATAP, etc. IPv6 has explicitly to be activated!
  - SP1: GUI installation, **netsh** command line interface
  - SP2: Teredo, firewall, and other additions
- Win2000
  - Only developer edition available
- Windows 95/98/ME
  - No official support

Enable IPv6: Windows

- WinXP
  - Execute “ipv6 install” at a command prompt (SP0)
  - Add ‘*Microsoft IPv6 Developer Edition*’ component as a new protocol in the Network Connections Control Panel pane (SP1)
  - Add ‘Microsoft TCP/IP version 6’ as a new protocol in the Network Connections Control Panel pane (SP2)

IPv6 commands: WinXP

- Command line interface (netsh):

```
c:\>netsh interface ipv6
```

- Well known (IPv4/6) commands

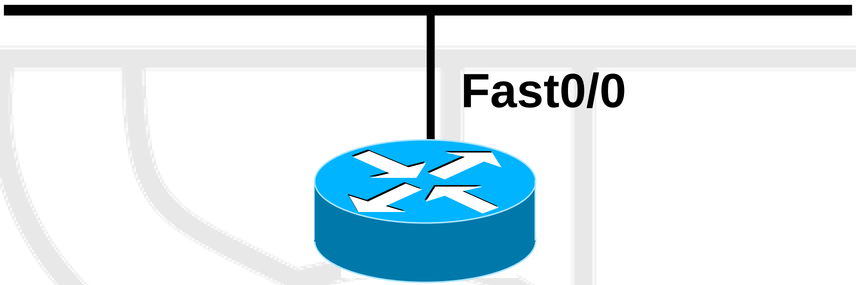
```
ipconfig, netstat, ping6,
tracert6, pathping
```

IPv6 Support: Linux distributions

- Redhat (6.2+), Fedora 1&2, SuSE (7.3+), Debian (2.2+), Mandrake (8.0+), Scientific Linux (3.0+), *BSD, etc
  - Look for IPv6 support at kernel!
- USAGI
  - Collaboration between WIDE, KAME and TAHI in order to improve kernel

IPv6 Support: Cisco IOS Example (1)

Manual Interface Identifier



```
!
interface FastEthernet0/0
 ip address 10.151.1.1 255.255.255.0
 ip pim sparse-mode
 duplex auto
 speed auto
 ipv6 address 2006:1::1/64
 ipv6 enable
 ipv6 nd ra-interval 30
 ipv6 nd prefix 2006:1::/64 300 300
!
```

IPv6 Support: Cisco IOS Example (1)

Manual Interface Identifier

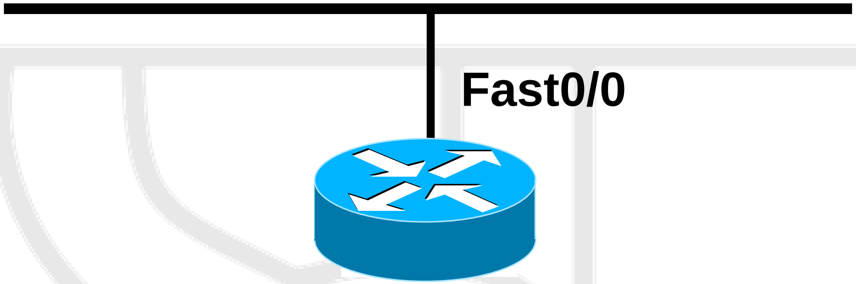
```
r1#sh ipv6 int fast0/0
FastEthernet0/0 is up, line protocol is up
  IPv6 is enabled, link-local address is FE80::207:50FF:FE5E:9460
  Global unicast address(es):
    2006:1::1, subnet is 2006:1::/64
  Joined group address(es):
    FF02::1
    FF02::2
    FF02::1:FF00:1
    FF02::1:FF5E:9460
  MTU is 1500 bytes
  ICMP error messages limited to one every 100 milliseconds
  ICMP redirects are enabled
  ND
  ND
  ND
  ND advertised retransmit interval is 0 milliseconds
  ND router advertisements are sent every 30 seconds
  ND router advertisements live for 1800 seconds
  Hosts use stateless autoconfig for addresses.
r1#
```

```
r1#sh int fast0/0
FastEthernet0/0 is up, line protocol is up
  Hardware is AmdFE, address is 0007.505e.9460 (bia 0007.505e.9460)
```

MAC Address : 0007.505e.9460

IPv6 Support: Cisco IOS Examples (2)

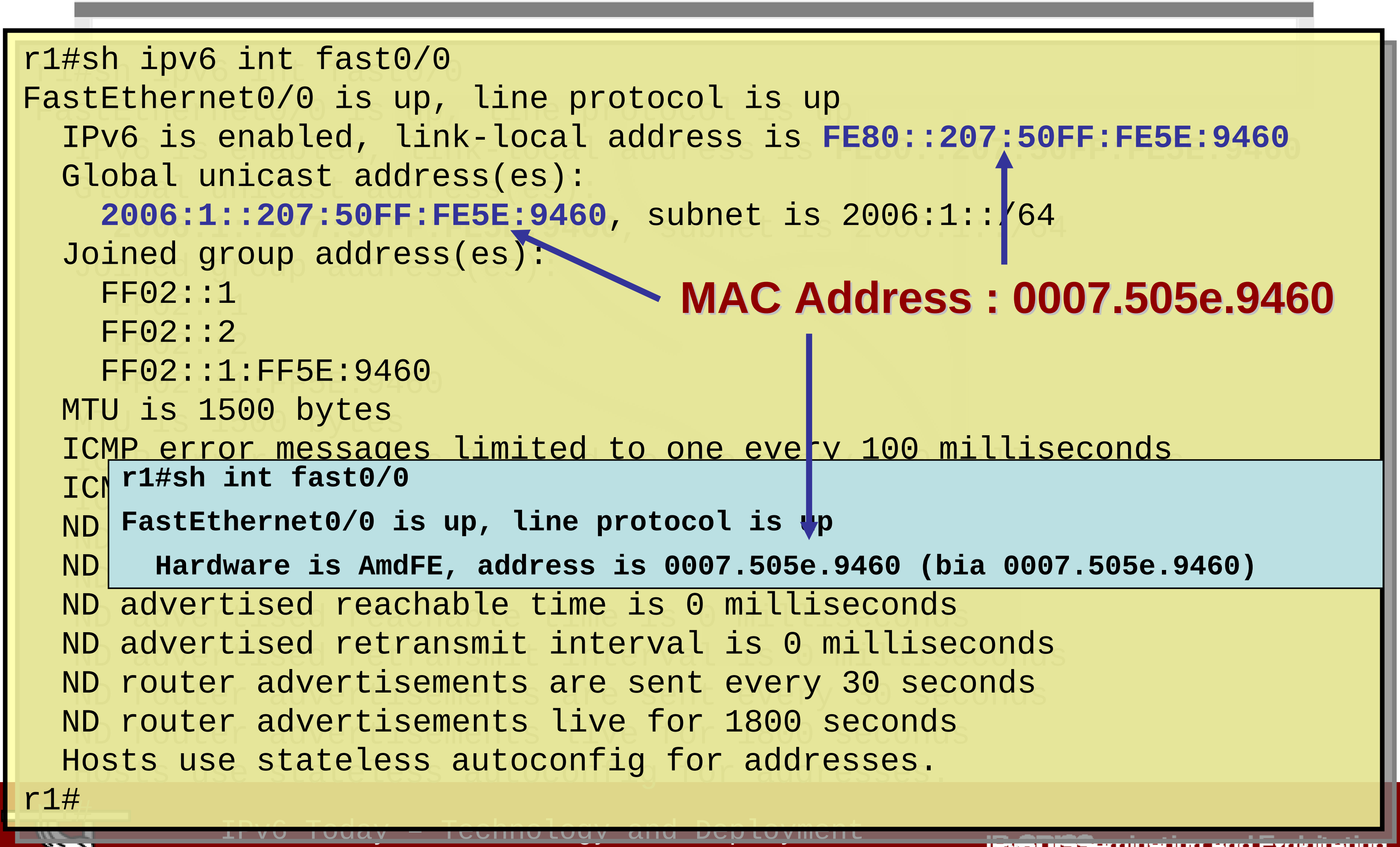
EUI-64 Interface Identifier



```
!
interface FastEthernet0/0
 ip address 10.151.1.1 255.255.255.0
 ip pim sparse-mode
 duplex auto
 speed auto
 ipv6 address 2006:1::/64 eui-64
 ipv6 enable
 ipv6 nd ra-interval 30
 ipv6 nd prefix 2006:1::/64 300 300
!
```

IOS IPv6 Addressing Examples (2)

EUI-64 Interface Identifier



```
r1#sh ipv6 int fast0/0
FastEthernet0/0 is up, line protocol is up
  IPv6 is enabled, link-local address is FE80::207:50FF:FE5E:9460
  Global unicast address(es):
    2006:1::207:50FF:FE5E:9460, subnet is 2006:1::/64
  Joined group address(es):
    FF02::1
    FF02::2
    FF02::1:FF5E:9460
  MTU is 1500 bytes
  ICMP error messages limited to one every 100 milliseconds
  ICM
  ND
  ND
  ND advertised reachable time is 0 milliseconds
  ND advertised retransmit interval is 0 milliseconds
  ND router advertisements are sent every 30 seconds
  ND router advertisements live for 1800 seconds
  Hosts use stateless autoconfig for addresses.
r1#
```

```
r1#sh int fast0/0
FastEthernet0/0 is up, line protocol is up
  Hardware is AmdFE, address is 0007.505e.9460 (bia 0007.505e.9460)
```

MAC Address : 0007.505e.9460

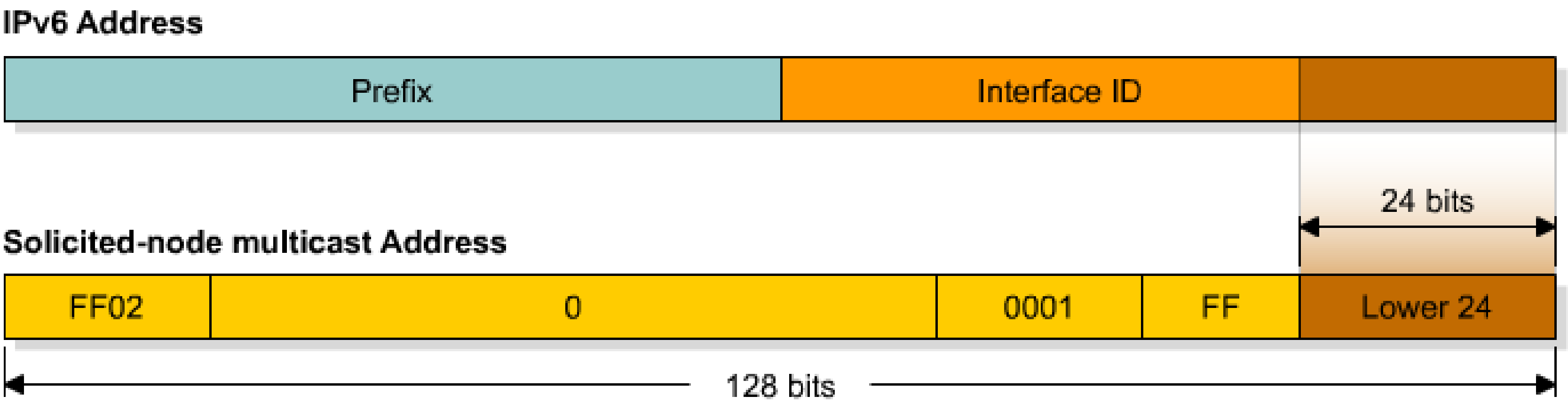
Agenda

- Technology Introduction
- IPv6 Protocol Basics
- IPv6 Protocol Specifics
- IPv6 Transition and Coexistence with IPv4

Neighbor Discovery (RFC 2461)

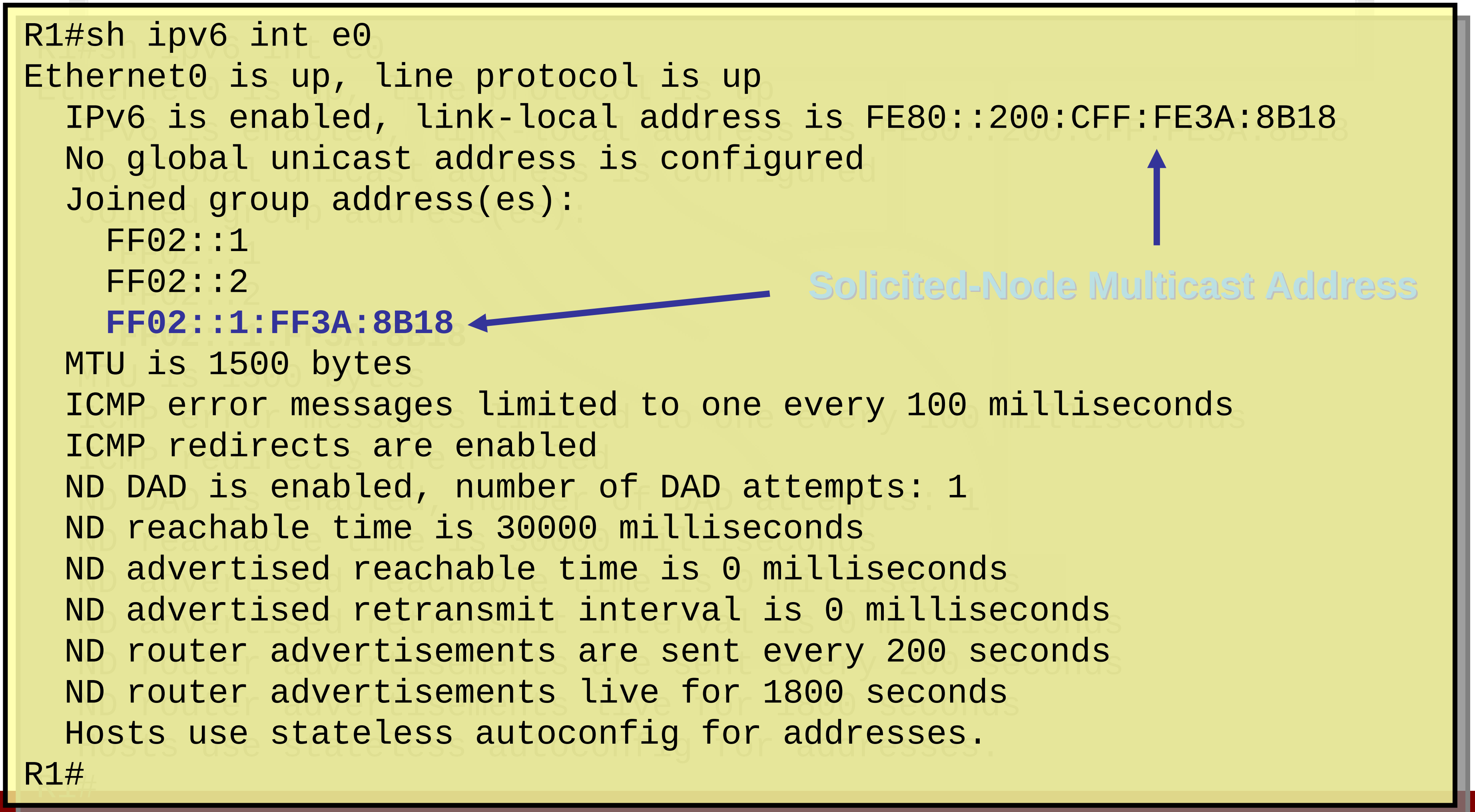
- IPv6 nodes which share the same physical medium (link) use Neighbor Discovery (ND) to:
  - discover their mutual presence
  - determine link-layer adresses of their neighbors
  - find routers
  - maintain neighbors’ reachability information (NUD)
- Defines 5 ICMPv6 packet types
  - Router Solicitation / Router Advertisements
  - Neighbor Solicitation / Neighbor Advertisements
  - Redirect

Solicited-Node Multicast Address



- Used in neighbor solicitation messages
- Multicast address with a link-local scope
- Solicited-node multicast consists of prefix + lower 24 bits from unicast, FF02::1:FF:

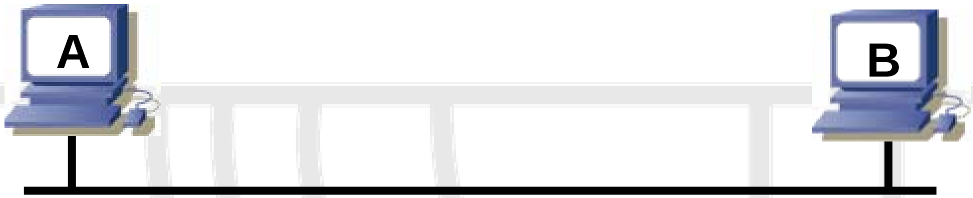
Router Interface



```
R1#sh ipv6 int e0
Ethernet0 is up, line protocol is up
  IPv6 is enabled, link-local address is FE80::200:CFF:FE3A:8B18
  No global unicast address is configured
  Joined group address(es):
    FF02::1
    FF02::2
    FF02::1:FF3A:8B18
  MTU is 1500 bytes
  ICMP error messages limited to one every 100 milliseconds
  ICMP redirects are enabled
  ND DAD is enabled, number of DAD attempts: 1
  ND reachable time is 30000 milliseconds
  ND advertised reachable time is 0 milliseconds
  ND advertised retransmit interval is 0 milliseconds
  ND router advertisements are sent every 200 seconds
  ND router advertisements live for 1800 seconds
  Hosts use stateless autoconfig for addresses.
R1#
```

Solicited-Node Multicast Address

Neighbor Solicitation



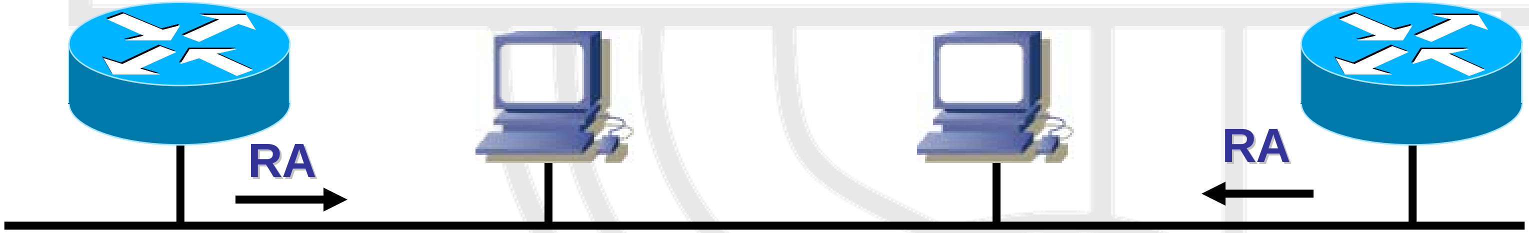
Neighbor Solicitation
ICMP type = 135

Src = A
Dst = Solicited-node multicast of B Data = link-layer address of A
Query = what is your link address?

Neighbor Advertisement
ICMP type = 136
Src = B
Dst = A
Data = link-layer address of B

A and B can now exchange packets on this link

Router Advertisements (RA)

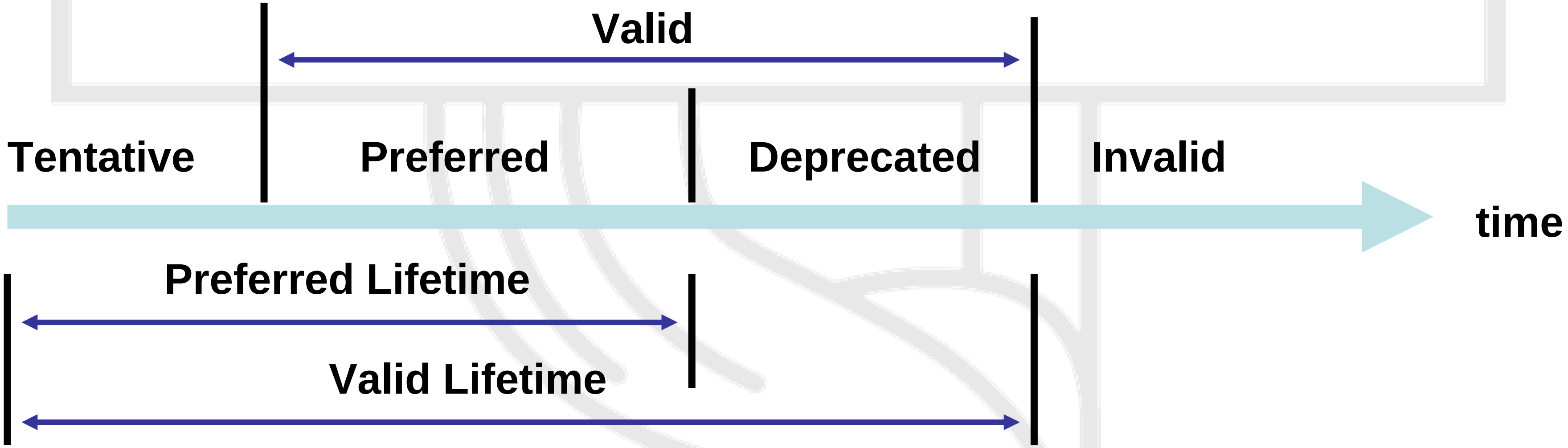


RA packet definitions:
ICMP Type = 134

Src = Router Link-local Address
Dst = All-nodes multicast address
Data= options, prefix, lifetime,
autoconfig flag

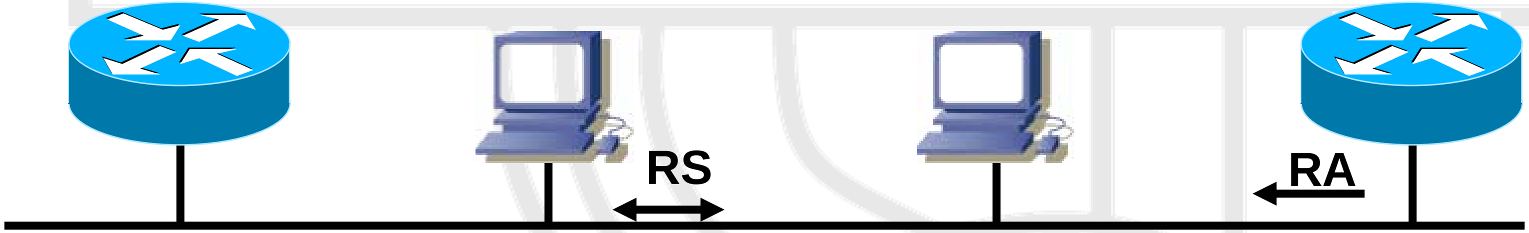
- Routers send periodic Router Advertisements (RA) to the all-nodes multicast address.

Address Lifetime



- **Tentative** : the address is in the process of being verified as unique
- **Preferred** : a node can send and receive unicast traffic to and from a preferred address
- **Deprecated** : the address is still valid, but using it for new communication is discouraged
- **Invalid** : the address can no longer send unicast traffic to or receive it from a node. An address enters this state after the valid lifetime expires.

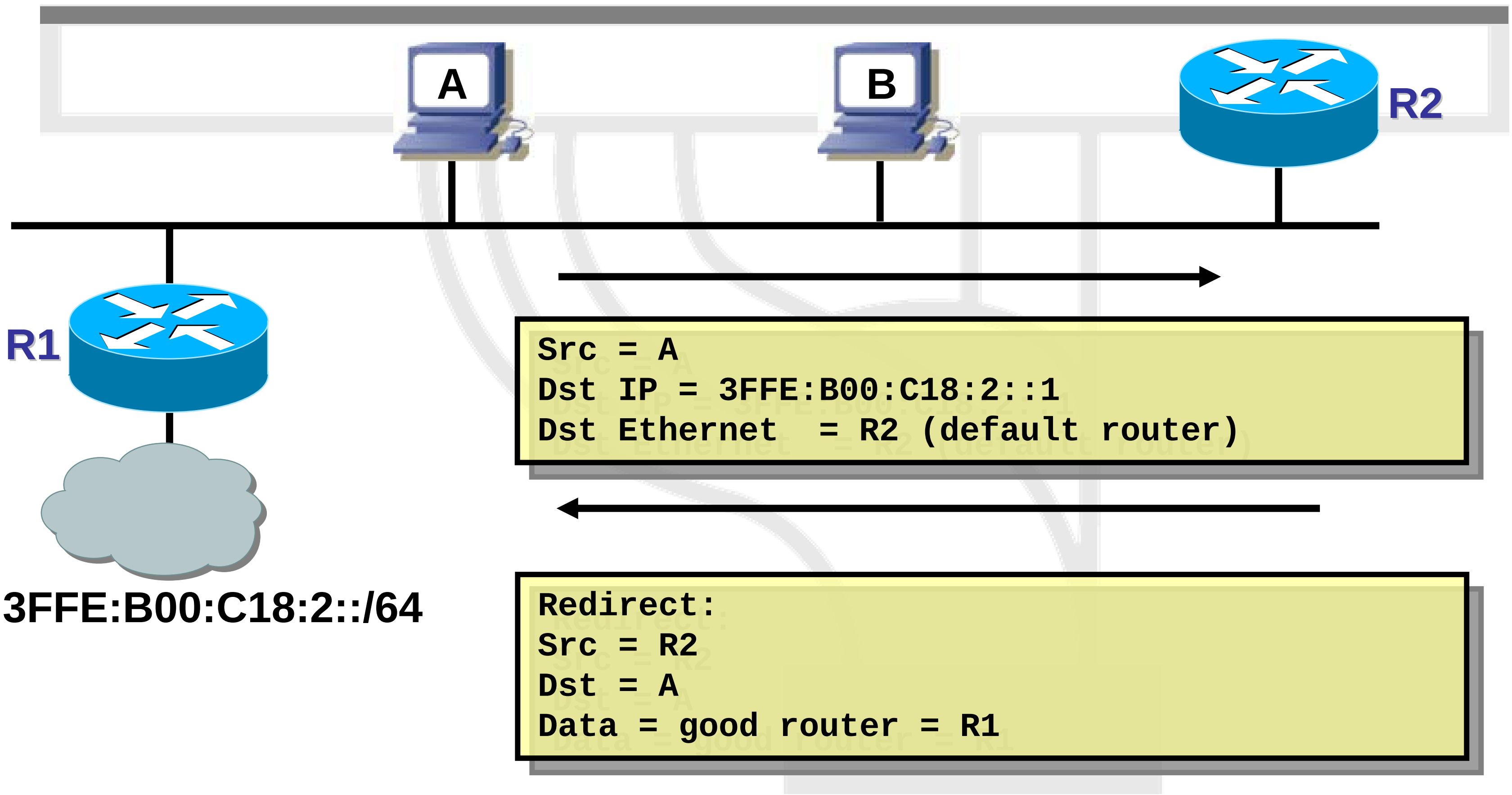
Router Solicitations



RS packet definitions:
ICMP Type = 133
Src = Unspecified Address
Dst = All-routers multicast address

- At boot time, nodes sends Router Solicitations to receive promptly Router Advertisements.

Redirect

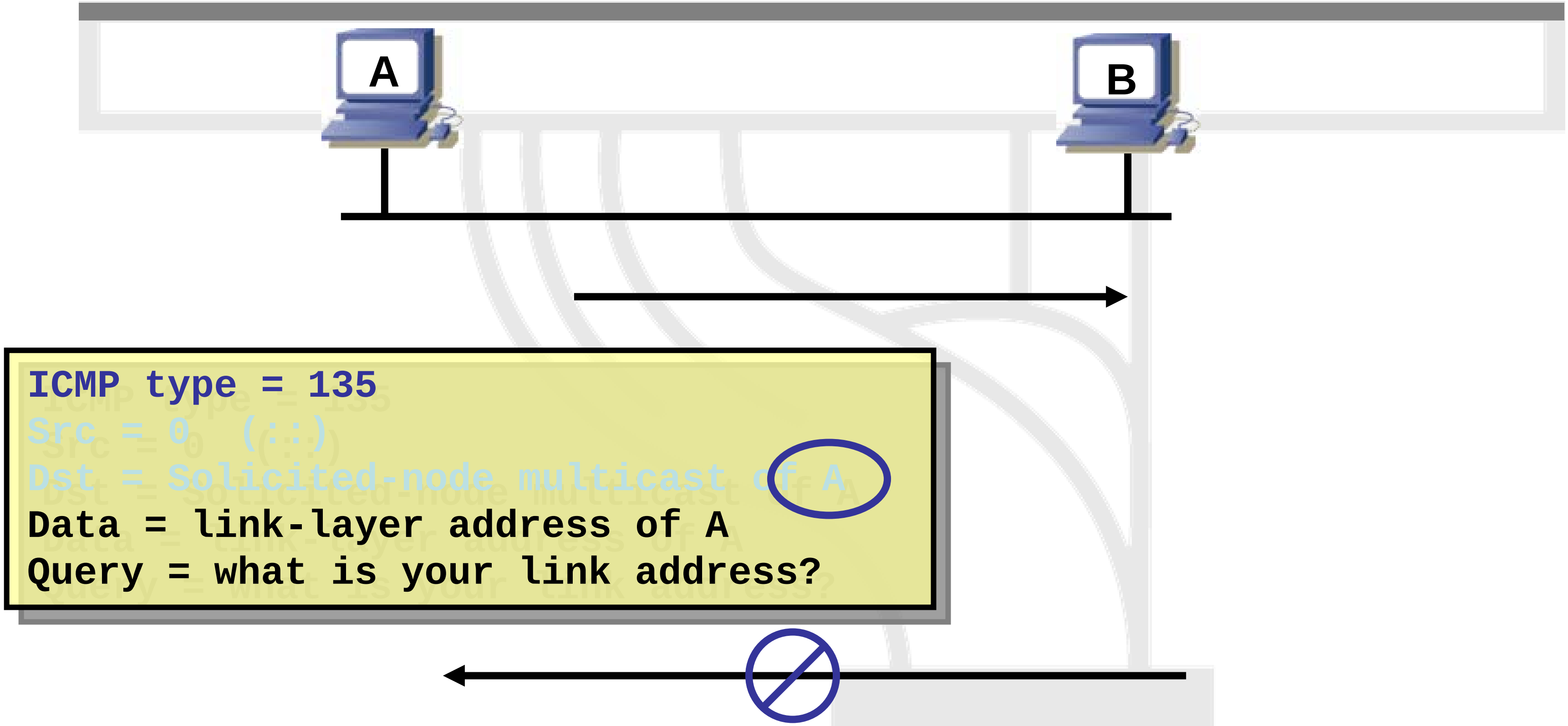


```
Src = A
Dst IP = 3FFE:B00:C18:2::1
Dst Ethernet  = R2 (default router)
```

```
Redirect:
Src = R2
Dst = A
Data = good router = R1
```

- Redirect is used by a router to signal the reroute of a packet to a better router

Duplicate Address Detection

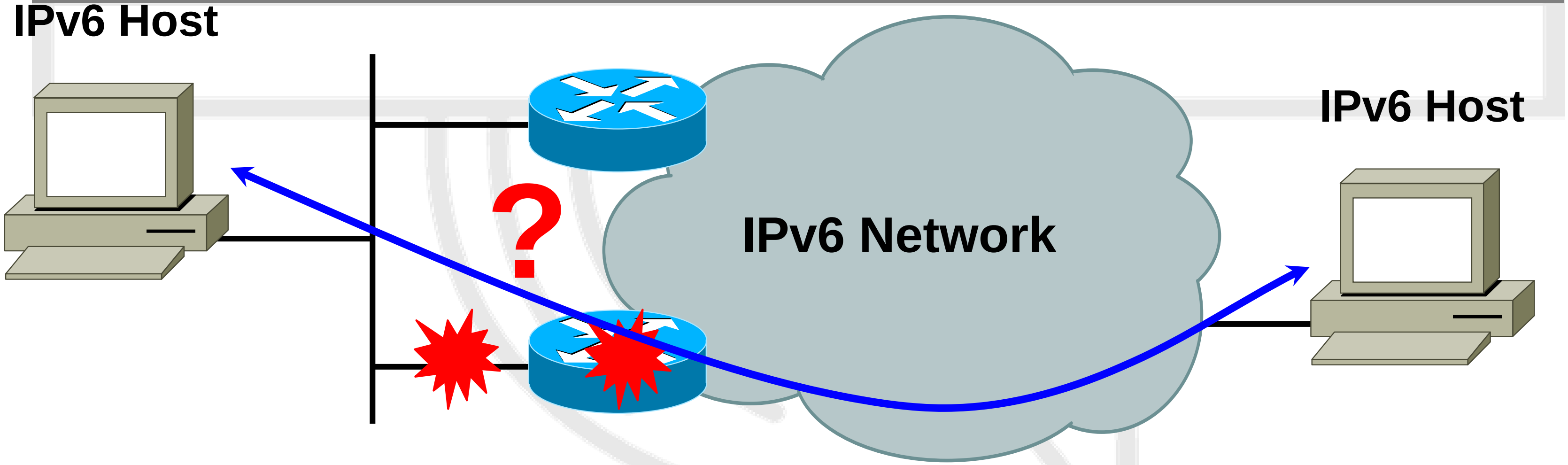


ICMP type = 135
Src = 0 (::)
Dst = Solicited-node multicast of A
Data = link-layer address of A
Query = what is your link address?

- Duplicate Address Detection (DAD) uses neighbor solicitation to verify the existence of an address to be configured.

First Hop Router Redundancy

Problem Definition



- IPv6 routing protocols ensure router-to-router resilience and failover
- But what if the path between a host and the first-hop router or the first-hop router itself fail?
- First Hop Redundancy Protocols (FHRP) ensure host-to-router resilience and failover

Agenda

- Introduction
- Tuning Neighbor Discovery Protocol
- Hot Standby Routing Protocol (HSRP) for IPv6
- Gateway Load Balancing Protocol (GLBP) for IPv6
- Default Router Selection
- Conclusions

Introduction

- First Hop Redundancy for IPv6 can be achieved in different ways
  - Tuning ICMPv6 / Neighbor Discovery (ND) protocol
  - Enabling one of the First Hop Redundancy Protocols (Cisco HSRP, Cisco GLBP, or VRRP)
  - Enabling Default Router Selection
- Reference
  - www.cisco.com/en/US/products/sw/iosswrel/ps5187/products_configuration_guide_chapter09186a00801d65ed.html

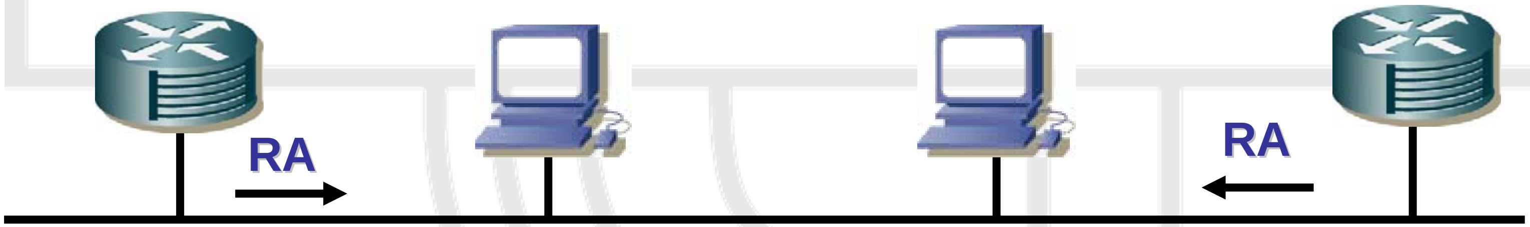
Agenda

- Introduction
- Tuning Neighbor Discovery Protocol
- HSRP for IPv6
- GLBP for IPv6
- Default Router Selection
- Conclusion

Neighbor Discovery RFC 2461

- Neighbor Discovery defines five ICMPv6 packet types
  - A pair of Router Solicitation (RS) and Router Advertisement (RA) messages
  - A pair of Neighbor Solicitation (NS) and Neighbor Advertisements (NA) messages
  - A redirect message

RA and Neighbor Unreachability Detection



- Routers announce their availability by sending out RA messages
- Hosts use these RAs to
  - Discover routers
  - Determine on/off-link destination addresses
  - Perform stateless address autoconfiguration
  - Determine routers’ reachable time, used for Neighbor Unreachability Detection (NUD)

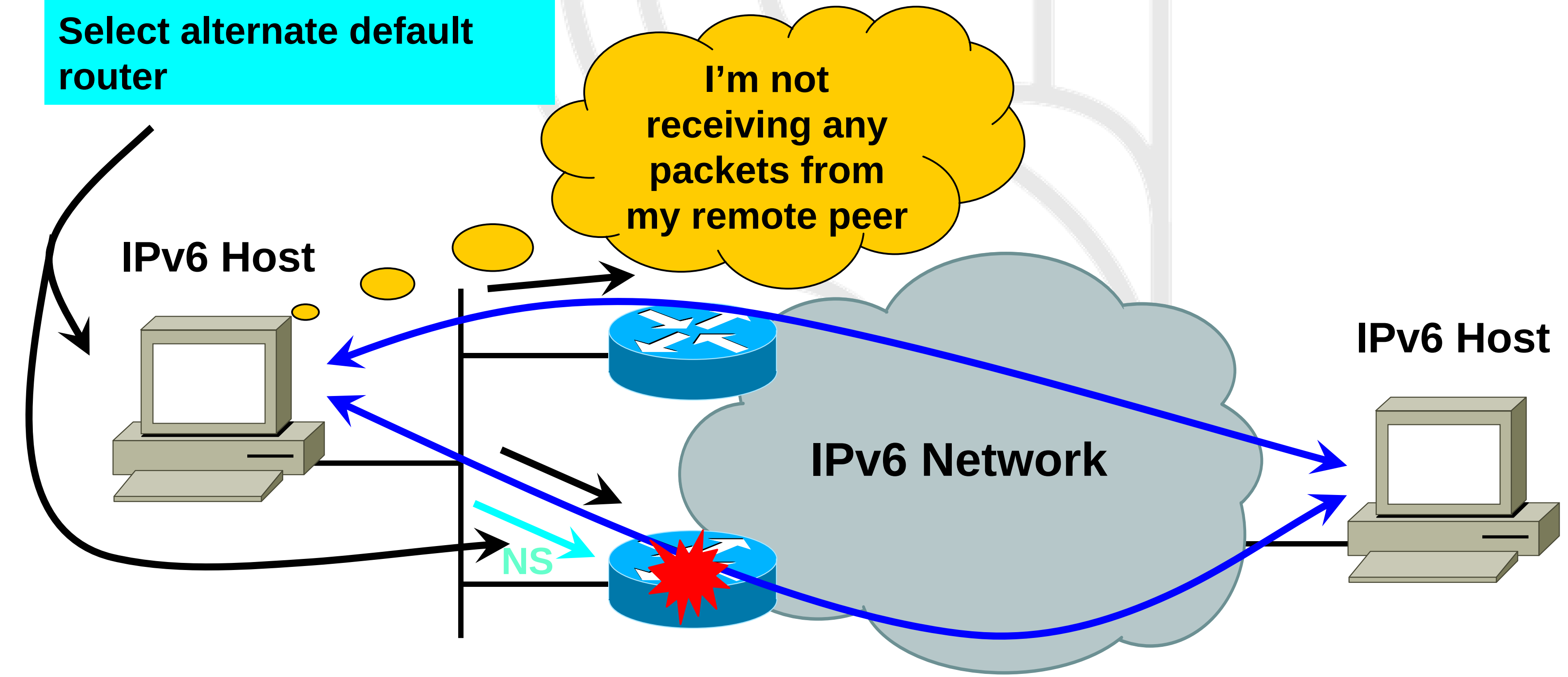
Neighbor Unreachability Detection

- RFC 2461: “How nodes determine that a neighbor is no longer reachable. For neighbors used as routers, alternate default routers can be tried. For both routers and hosts, address resolution can be performed again.”
- NUD uses confirmation from two sources
  - When possible, upper-layer protocols provide a positive confirmation that a connection is making "forward progress", that is, previously sent data is known to have been delivered correctly
  - When positive confirmation is not forthcoming through such "hints", a node sends unicast Neighbor Solicitation messages that solicit Neighbor Advertisements as reachability confirmation from the next hop

Neighbour Unreachability Detection (NUD)

Reachable Time timer for my neighbor router has Neighbor state: PROBE → DELETE

Select alternate default router



IPv6 ND Timers
Which to Tune?

- Parameters that do not reduce default router failover
  - ipv6 nd prefix
    - Valid lifetime - for on/off-link determination / preferred lifetime - for stateless address autoconfiguration
  - ipv6 nd ra-interval
    - Interval between RA transmissions, jittered
  - ipv6 nd ra-lifetime
    - Router lifetime - validity of the router as a default router

IPv6 ND Timers Which to Tune? (Cont.)

- Parameters that do reduce default router failover
  - ipv6 nd reachable-time
    - Reachable time - time a node will consider a neighbor (router) to be reachable after receiving a reachability confirmation

IPv6 ND Tuning Default Router

```
!
interface ethernet x/y
 ipv6 nd prefix 2001:XXXX::/64
   2592000 604800
 ipv6 nd ra-interval 200
 ipv6 nd ra-lifetime 1800
 ipv6 nd reachable-time 15000
!
```

- Default Reachable Time is 30 seconds
- Tuning the Reachable Time to a lower value will ensure faster failover between default routers
  - Test performed with Reachable Time 5 → 15 seconds
- Trade-off is increase in NS/NA messages, processing impact on IPv6 nodes (function of # nodes on subnet)

IPv6 ND Tuning
Host

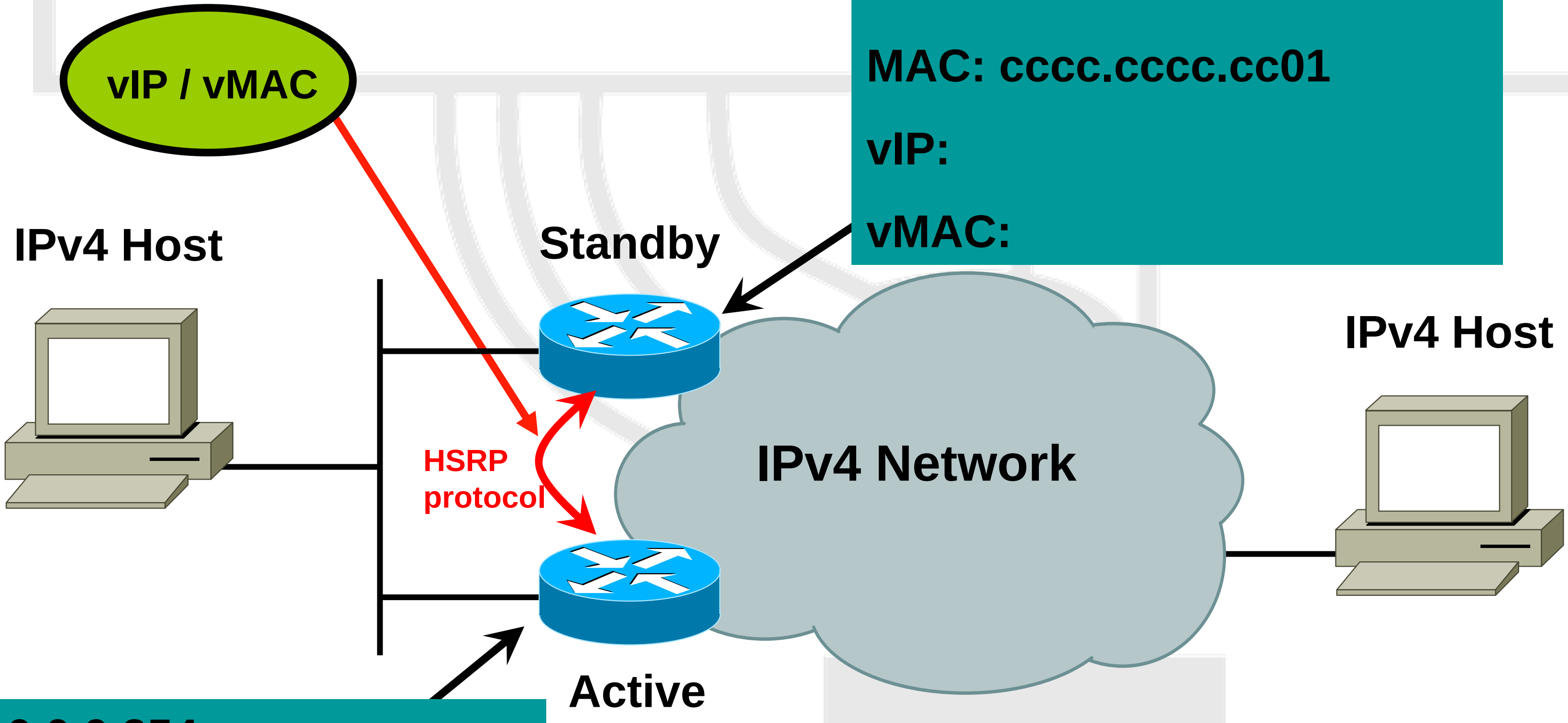
```
R200#sh ipv6 routers
Router FE80::A8BB:CCFF:FE00:C900 on Ethernet0/0, last update 1 min
  Hops 64, Lifetime 1800 sec, AddrFlag=0, OtherFlag=0, MTU=1500
  HomeAgentFlag=0, Preference=Medium
  Reachable time 15000 msec, Retransmit time 0 msec
  Prefix 2001:1::/64 onlink autoconfig
    Valid lifetime 2592000, preferred lifetime 604800
Router FE80::A8BB:CCFF:FE00:CA00 on Ethernet0/0, last update 1 min
  Hops 64, Lifetime 1800 sec, AddrFlag=0, OtherFlag=0, MTU=1500
  HomeAgentFlag=0, Preference=Medium
  Reachable time 15000 msec, Retransmit time 0 msec
  Prefix 2001:1::/64 onlink autoconfig
    Valid lifetime 2592000, preferred lifetime 604800
```

Note: host is router configured as a host in this example

Agenda

- Introduction
- Tuning Neighbor Discovery Protocol
- Cisco HSRP for IPv6
- GLBP for IPv6
- Default Router Selection
- Conclusions

Hot Standby Routing Protocol



IP: 10.0.0.253

MAC: cccc.cccc.cc01

vIP:

vMAC:

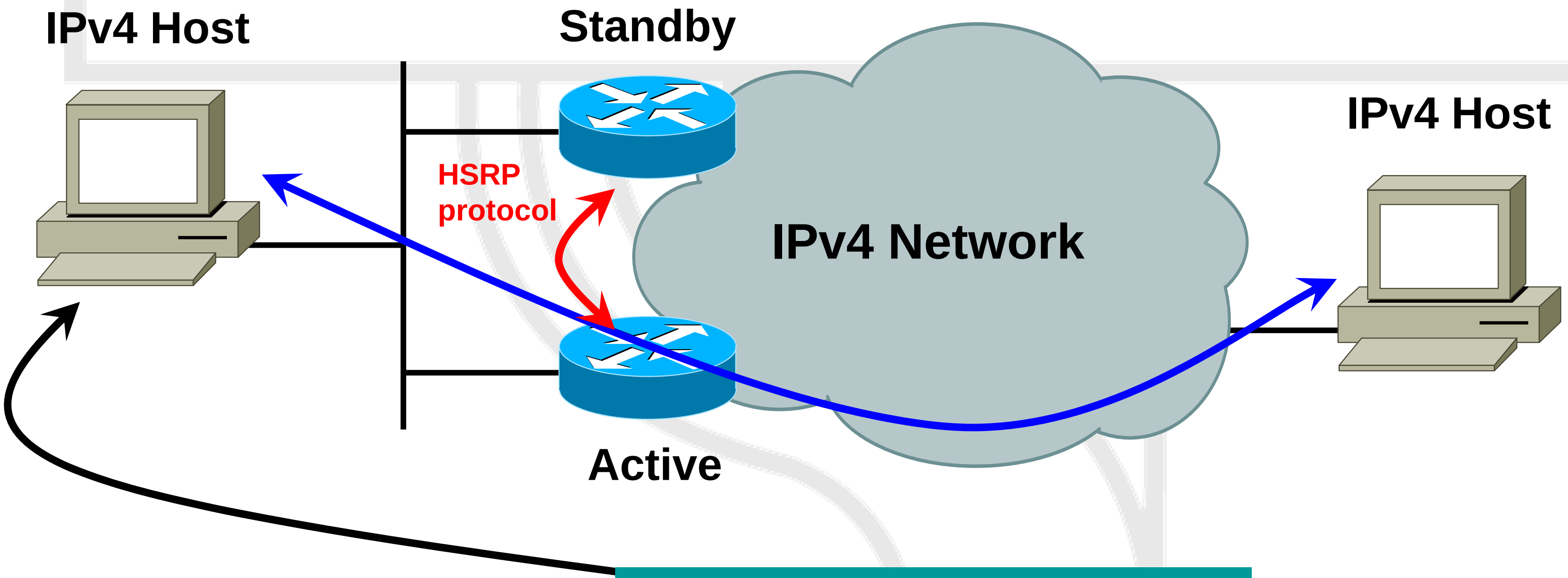
IP: 10.0.0.254

MAC: bbbb.bbbb.bb01

vIP: 10.0.0.10

vMAC: 0000.5e00.0101

Hot Standby Routing Protocol (Cont.)



IP: 10.0.0.1

MAC: aaaa.aaaa.aa01

GW: 10.0.0.10

GW MAC: 0000.5e00.0101

Hot Standby Routing Protocol (Cont.)

- HSRP for IPv4 and IPv6 have similar state-machine
- HSRP IPv4 differences
  - Host will learn the default gateway through router RA messages (no need to configure default gateway)
  - Active HSRP router will by default send RA every 200 seconds
  - Standby HSRP router will suppress its RA messages
- HSRP for IPv6 vs. IPv6 ND
  - Provides predictable IPv6 Host-to-Router redundancy and faster failover – default 10 seconds vs. default 30 seconds
  - Reduces ND traffic overhead (NS/NA messages) associated with reducing ND Reachable Time timer

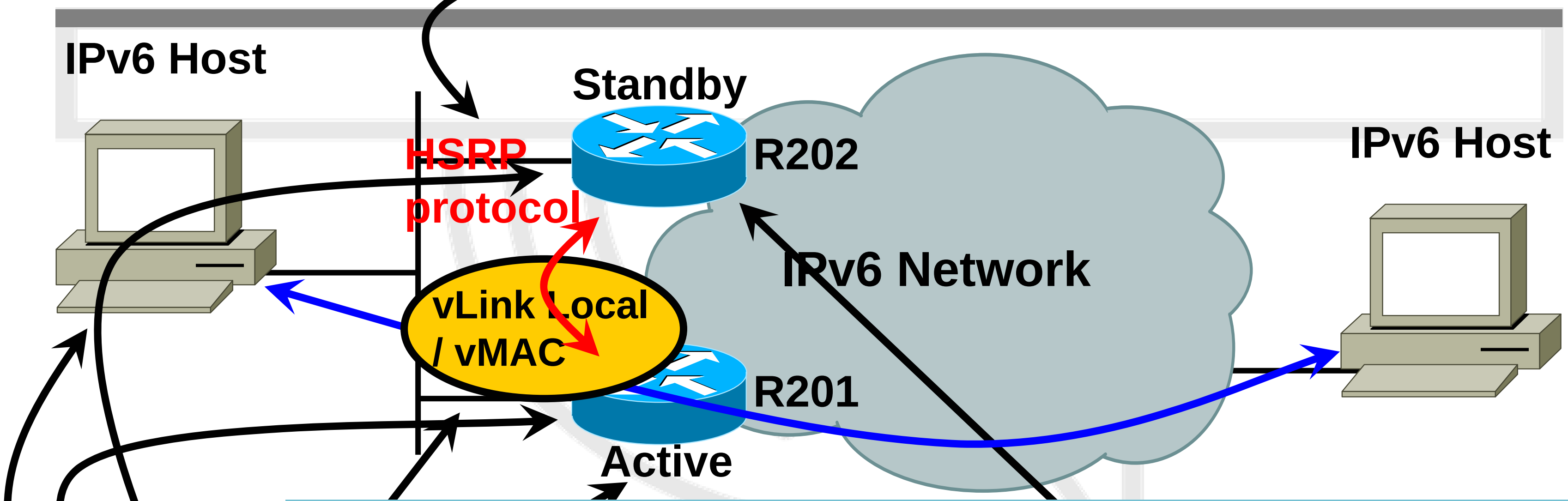
Hot Standby Routing Protocol (Cont.)

- Virtual MAC addresses associated with HSRP for IPv6
  - 0005.73A0.0000 … 0005.73A0.0FFF (= 4096 available addresses)
  - HSRP group number (4096) → virtual MAC address → virtual Link Local address (modified EUI-64 derived)
- UDP port 2029 for HSRP packets
  - standby version 2
  - standby <group> ipv6 {autoconfig | <IPv6 address>}
  - Autoconfig creates a Link Local IPv6 address derived from the virtual MAC address through modified EUI-64
  - If an IPv6 address is entered then it MUST be Link Local

HSRP for

UDP Port 2029

LL Scope Mcast: FF02::66



```
R202#sh standby
Ethernet0/0 - Group 0 (version 2)
  State is Standby
    1 state change, last state change 00:00:25
  R200#sh ipv6 routers
  Router FE80::5:73FF:FEA0:0 on Ethernet0/0, last update 2 min
    Hops 64, Lifetime 1800 sec, AddrFlag=0, OtherFlag=0,
  MTU=1500
    HomeAgentFlag=0, Preference=Medium
    Reachable time 0 msec, Retransmit time 0 msec
    Prefix 2001:1::/64 onlink autoconfig
      Valid lifetime 2592000, preferred lifetime 604800
  Priority 100 (default 100)
  IP redundancy name is "hsrp-Et0/0-0" (default)
```

in 8.068 sec)

LL: FE80::A8BI

MAC : aabb.cc(

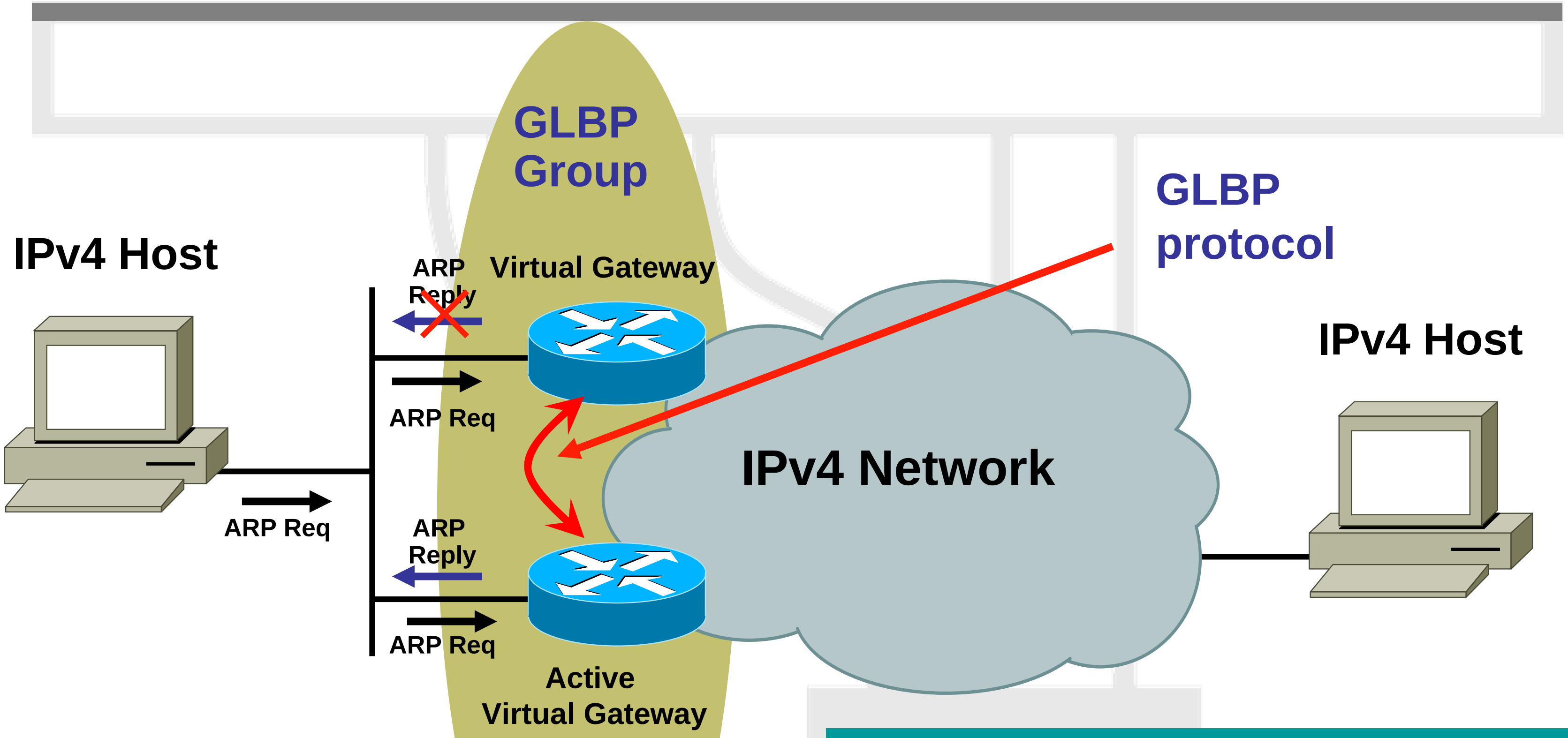
vLL: FE80::5:7:

vMAC: 0005.73

Agenda

- Introduction
- Tuning Neighbor Discovery Protocol
- HSRP for IPv6
- GLBP for IPv6
- Default Router Selection
- Conclusions

Gateway Load Balancing Protocol



Active Virtual Gateway (AVG)

is responsible for answering ARP requests

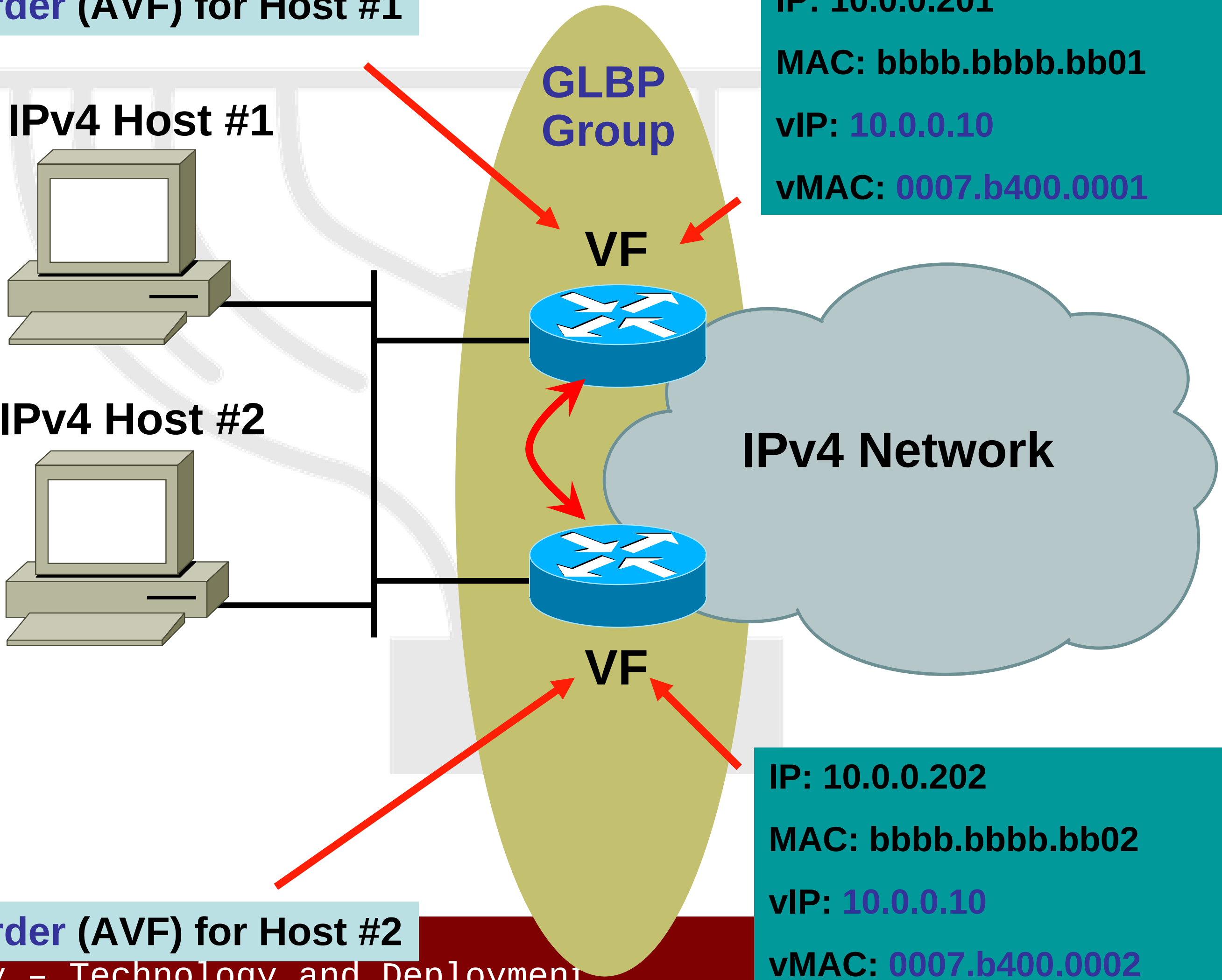
Active Virtual Gateway (AVG) selects:

Virtual IP address
- **One per GLBP group**

Virtual MAC addresses
- **One per Active Virtual Forwarder (AVF)**

Gateway Load Balancing Protocol

Active Virtual Forwarder (AVF) for Host #1



IP: 10.0.0.1
MAC: bbbb.bbbb.bb01
Gateway IP: 10.0.0.10
vMAC: 0007.b400.0001

IP: 10.0.0.2
MAC: bbbb.bbbb.bb02
Gateway IP: 10.0.0.10
vMAC: 0007.b400.0002

IP: 10.0.0.201
MAC: bbbb.bbbb.bb01
vIP: 10.0.0.10
vMAC: 0007.b400.0001

IP: 10.0.0.202
MAC: bbbb.bbbb.bb02
vIP: 10.0.0.10
vMAC: 0007.b400.0002

Active Virtual Forwarder (AVF) for Host #2

Gateway Load Balancing Protocol(Cont.)

- Original GLBP specification already catered for IPv6 addresses
- Virtual Gateway (VG) redundancy
  - Same state-machine as with HSRP
  - One gateway is elected as Active VG (AVG)
  - One gateway is elected as Standby VG
  - Remaining gateways are in listening state

Gateway Load Balancing Protocol
Load-Balancing Schemes

- Weighted
  - Ability to place a weight on each device when calculating the amount of load sharing that will occur through MAC assignment
- Host dependent
  - The MAC address of a host is used to determine which VF MAC address the host is directed towards. This ensures that a host will be guaranteed to use the same virtual MAC address.
- Round robin
  - Each VF MAC address is used sequentially in ARP replies for the virtual IP address. Round robin load balancing is suitable for any number of end hosts.

Gateway Load Balancing Protocol for IPv6

- Same state-machine as IPv4 is used
- GLBP-v6 will make use of technology designed in IPv6 to force hosts to use NS instead of the default Router Advertise (RA) mechanism
  - Quote from RFC2461:
  - Load balancing is handled by allowing routers to omit the source link-layer address from Router Hello packets, thereby forcing neighbors to use Neighbor Solicitation messages to learn link-layer addresses of routers.

Gateway Load Balancing Protocol for IPv6 (Cont.)

- The optional source MAC address will not be included in multicast Router Hello messages
- IPv6 virtual address (link-local IPv6 address) will be used instead of IPv4 virtual address
- The GLBP IPv6 Multicast address
  - FF02::224.0.0.102 or FF02::0100.5E00.0066

Agenda

- Introduction
- Tuning Neighbor Discovery Protocol
- HSRP for IPv6
- GLBP for IPv6
- Default Router Selection
- Conclusions

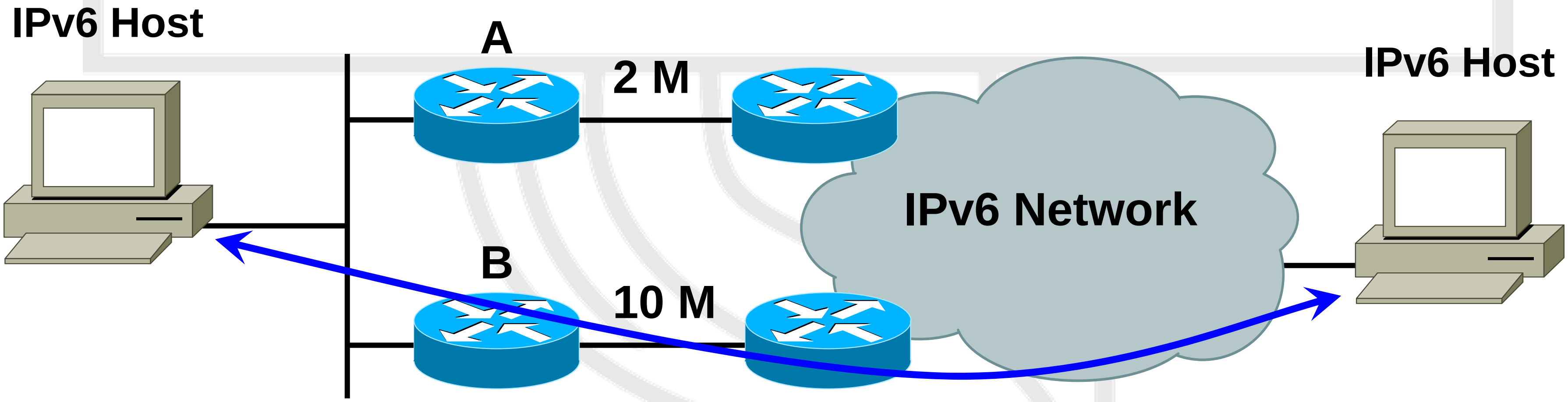
Default Router Selection

- Hosts maintain a default router list from which one is selected for traffic to off-link destinations and is then cached
  - "round-robin", or "always the same" selection is implementation dependent
- RFC 4191 – two optional extensions to RA messages
  - Default Router Preferences (DRP): A very coarse preference metric for default routers
  - More-Specific Routes (MSR): More specific routes than the default route, together with a very coarse preference metric for each such route
  - DRP can be implemented without implementing MSR

Default Router Selection (Cont.)

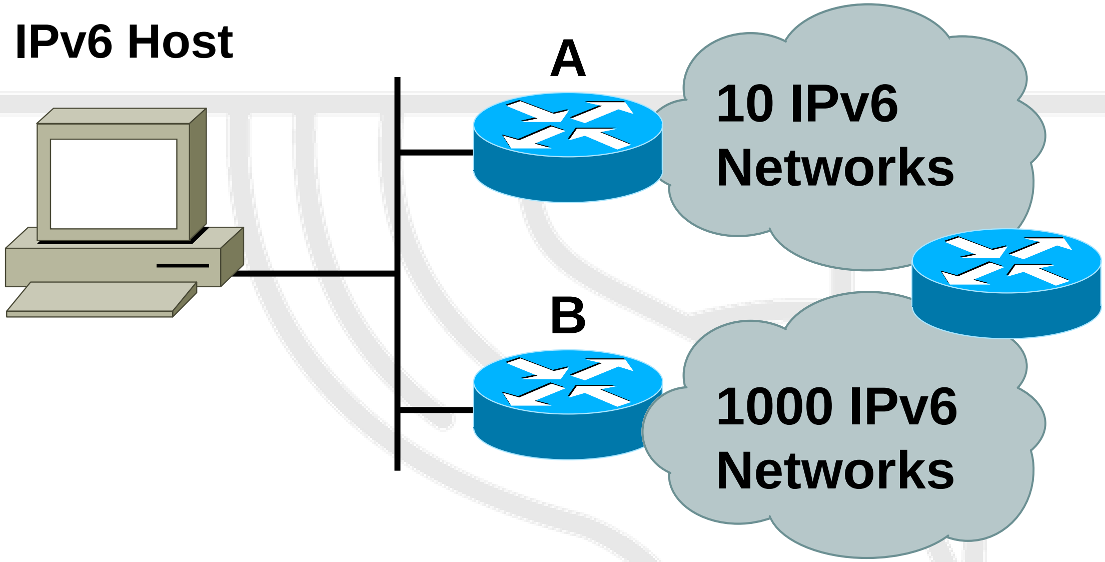
- Default Router Selection
  - Enhances hosts’ selection mechanism from a set of default routers
  - Complementary to mechanisms to improve First Hop Redundancy (ND tuning, HSRP)

Default Router Selection Example One



- One default router may provide much better performance than another for a destination
- It makes sense to adopt “B” as the default router

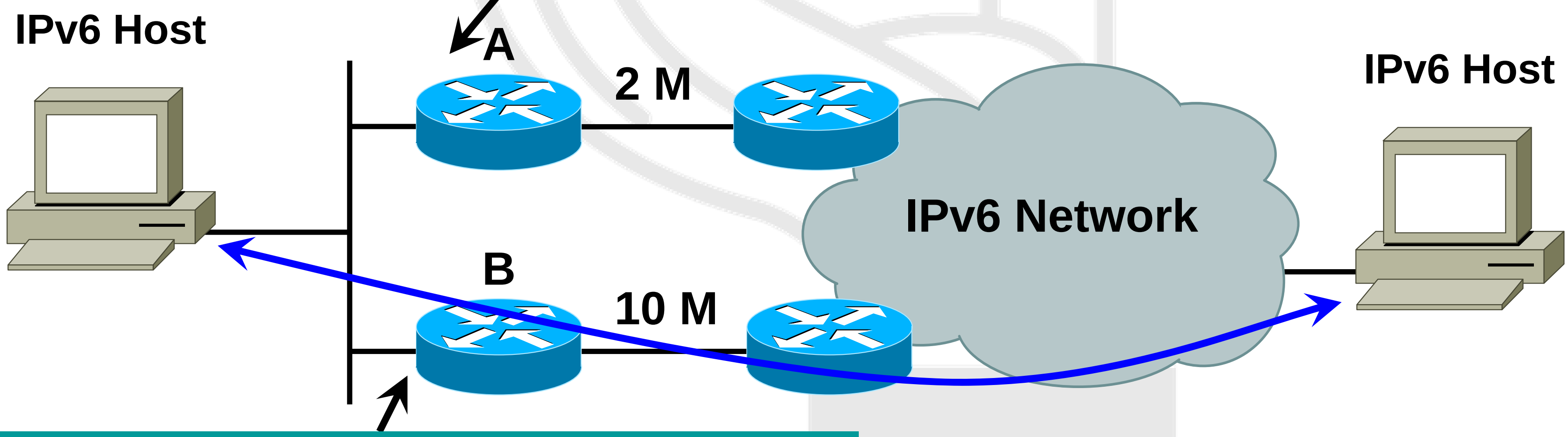
Default Router Selection Example Two



- If most traffic is routed through “B”, than “B” is least likely to redirect traffic
- In order to minimize redirects, it makes sense to adopt “B” as the default router

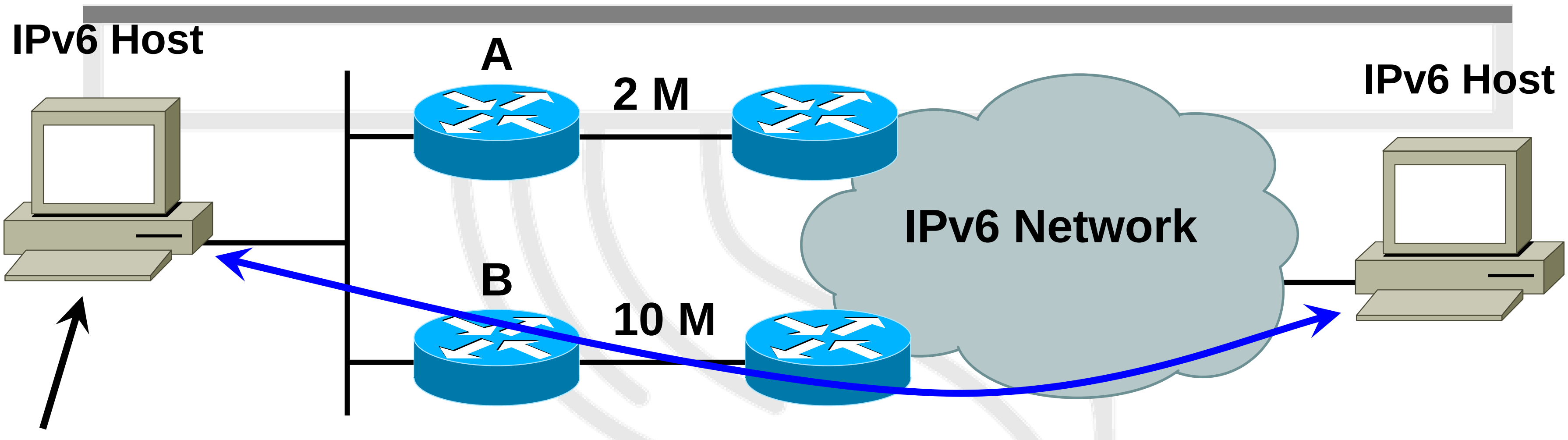
Default Router Selection

```
!
interface Ethernet0/0
 ipv6 nd reachable-time 15000
 ipv6 nd router-preference Low
!
```



```
!
interface Ethernet0/0
 ipv6 nd reachable-time 15000
 ipv6 nd router-preference High
!
```

Default Router Selection



```
R200#sh ipv6 router
Router FE80::A8BB:CCFF:FE00:CA00 on Ethernet0/0, last update 0 min
  Hops 64, Lifetime 1800 sec, AddrFlag=0, OtherFlag=0, MTU=1500
  HomeAgentFlag=0, Preference=High
  Reachable time 15000 msec, Retransmit time 0 msec
  Prefix 2001:1::/64 onlink autoconfig
    Valid lifetime 2592000, preferred lifetime 604800
Router FE80::A8BB:CCFF:FE00:C900 on Ethernet0/0, last update 2 min
  Hops 64, Lifetime 1800 sec, AddrFlag=0, OtherFlag=0, MTU=1500
  HomeAgentFlag=0, Preference=Low
  Reachable time 15000 msec, Retransmit time 0 msec
  Prefix 2001:1::/64 onlink autoconfig
    Valid lifetime 2592000, preferred lifetime 604800
```

Agenda

- Introduction
- Tuning Neighbor Discovery Protocol
- HSRP for IPv6
- GLBP for IPv6
- Default Router Selection
- Conclusion

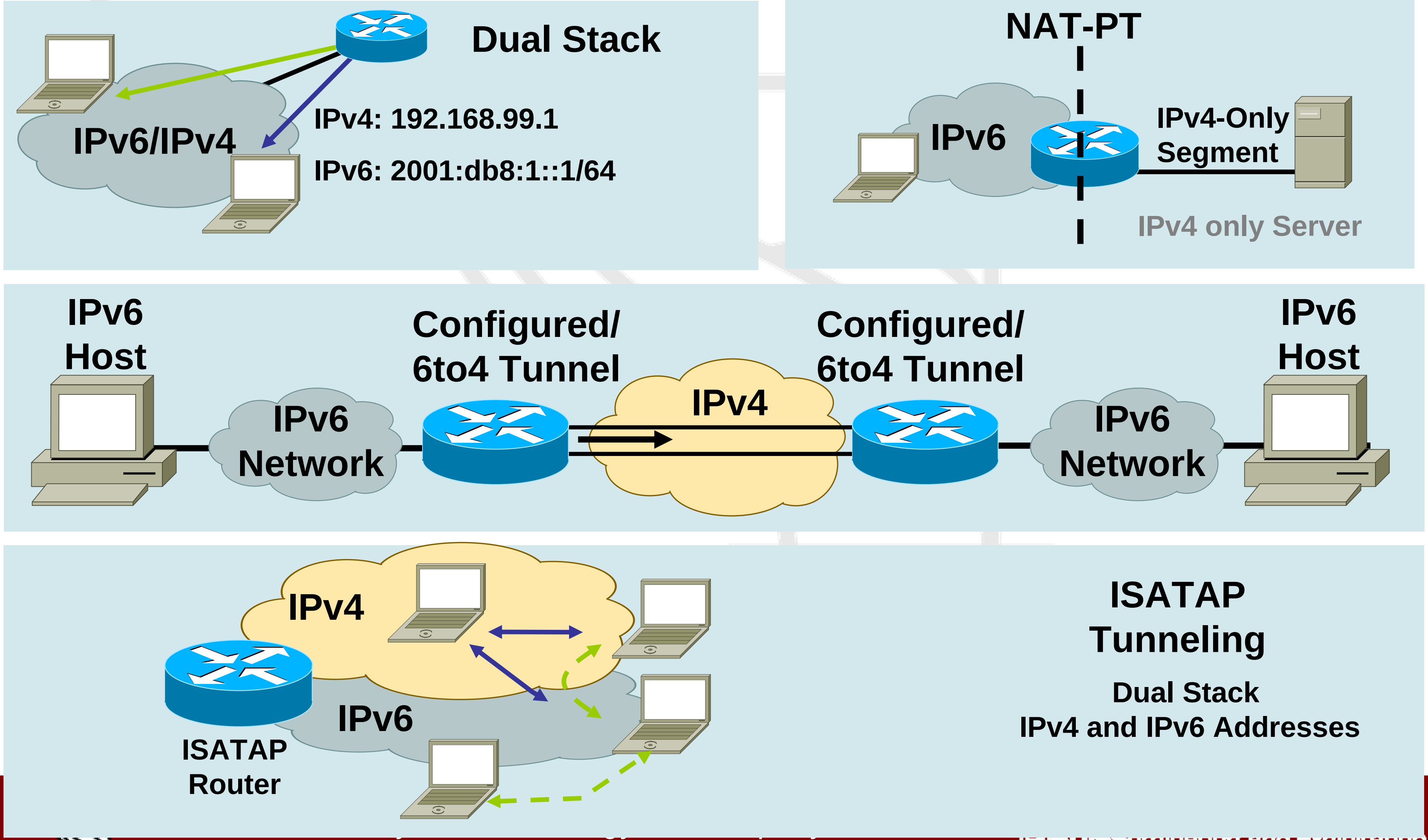
Conclusion

- Tuning ICMPv6 and Neighbor Discovery Reachable Time can be achieved on any Cisco IOS Software release supporting IPv6
- Default Router Selection on Release 12.4(2)T
- First Hop Redundancy Protocol for IPv6
  - Cisco HSRP for IPv6 on Release 12.4(4)T
  - Cisco GLBP for IPv6 on Release 12.4(6)T
  - Later support for VRRP for IPv6

Agenda

- Technology Introduction
- IPv6 Protocol Basics
- IPv6 Protocol Specifics
- IPv6 Transition and Coexistence with IPv4

IPv6 Coexistence in the Enterprise



Dual Stack

IPv4: 192.168.99.1

IPv6: 2001:db8:1::1/64

NAT-PT

Configured/ 6to4 Tunnel

ISATAP Tunneling

Dual Stack
IPv4 and IPv6 Addresses

IPv6 Transition and Coexistence with IPv4

Agenda

- Approaches to deploying IPv6
  - Standalone (IPv6-only) or alongside IPv4
- Considerations for IPv4 and IPv6 coexistence
- Approaches to coexistence
  - 1: Tunnelling
  - 2: Translation
  - 3: Dual-stack
- Specific examples
  - 6to4
  - Tunnel broker
  - ISATAP

Deploy IPv6 standalone

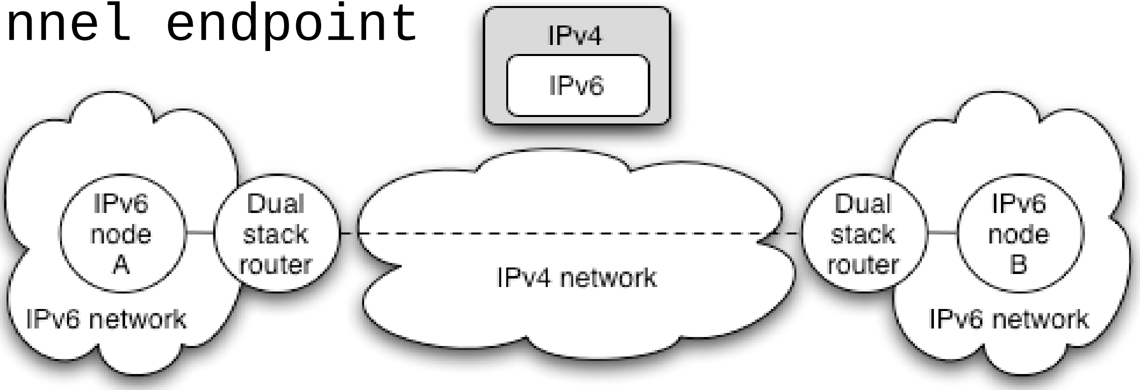
- One option is to deploy an IPv6-only network
- Introduces specific requirements:
  - All components must be IPv6-capable
  - Likely to need to talk to IPv4-only systems
    - So need some way to ‘translate’ between the protocols at some layer
  - Likely to want to communicate with remote IPv6 network ‘islands’ that may only be connected through existing IPv4 networks
    - Need a way to send IPv6 packets over/through an intermediate IPv4-only network

Deploy IPv6 alongside IPv4

- Existing network runs IPv4
- Introduce IPv6 to the same network
- Deploy IPv6 in parallel to IPv4
  - Known as ‘dual-stack’ operation
  - Hosts and routers are able to talk using either protocol
- Choice of protocol is application-specific
  - DNS returns IPv4 and IPv6 addresses for a given hostname
  - As an example, MS Internet Explorer by default prefers IPv6 connectivity, but can fall back to IPv4 (after a timeout)
  - Thus need to be confident IPv6 connectivity is good, else the application may perform worse than in an IPv4-only network

1: Tunnelling

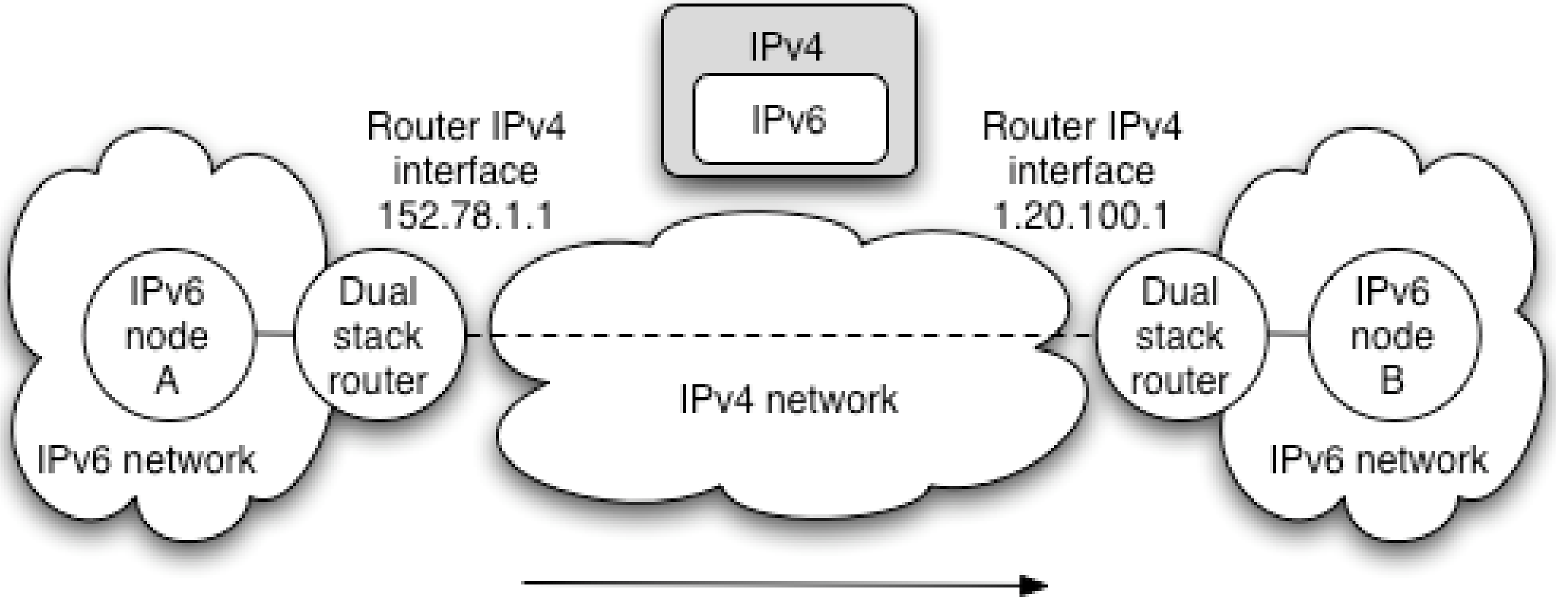
- IPv6 packets encapsulated in IPv4 packets
  - IPv6 packet is payload of IPv4 packet
- Usually used between edge routers to connect IPv6 ‘islands’
  - Edge router talks IPv6 to internal systems
  - Encapsulates IPv6 in IPv4 towards remote tunnel endpoint



Packet delivery over the tunnel

- IPv6 node A sends packet to IPv6 node B
  - Routed internally to edge router A
- Edge router A sees destination network B is reachable over tunnel interface
  - Encapsulates IPv6 packet in IPv4 packet(s)
  - Sends resulting IPv4 packet(s) to edge router B
  - Delivered over existing IPv4 Internet infrastructure
- Edge router B decapsulates IPv6 packet from payload of received IPv4 packet
  - Packet routed internally in network B to node B
  - Node B receives the IPv6 packet

Tunnel addressing view



IPv6 node A
2001:db8:1::1

IPv6 Network
2001:db8:1::/48

IPv4 src: 152.78.1.1
IPv4 dst: 1.20.100.1
IPv6 src: 2001:db8:1::1
IPv6 dst: 2001:db8:2:1

IPv6 node B
2001:db8:2::1

IPv6 Network
2001:db8:2::/48

Fragmentation

- IPv6 requires that packet fragmentation only occurs at end systems, not on intermediate routers
  - Use Path Maximum Transmission Unit (PMTU) Discovery to choose the MTU
  - Achieved using special ICMP messages
  - Minimum MTU is 1280 bytes in IPv6
- When tunnelling IPv6 in IPv4, the IPv4 packets may be fragmented
  - Depends on the IPv4 packet size
  - Additional IPv6 headers (e.g. Authentication Header) will affect this

Tunnel solution considerations

- These include:
  - Security
  - Manual or automatic setup
  - Ease of management
  - Handling dynamic IPv4 addresses
  - Support for hosts or sites to be connected
  - Scalability: 10, 100, or 10,000 served tunnels?
  - Support for NAT traversal
  - Tunnel service discovery
  - Support for special services (e.g. multicast)
  - Tunnel concentration/bandwidth usage issues
- We’ll come back to these later…

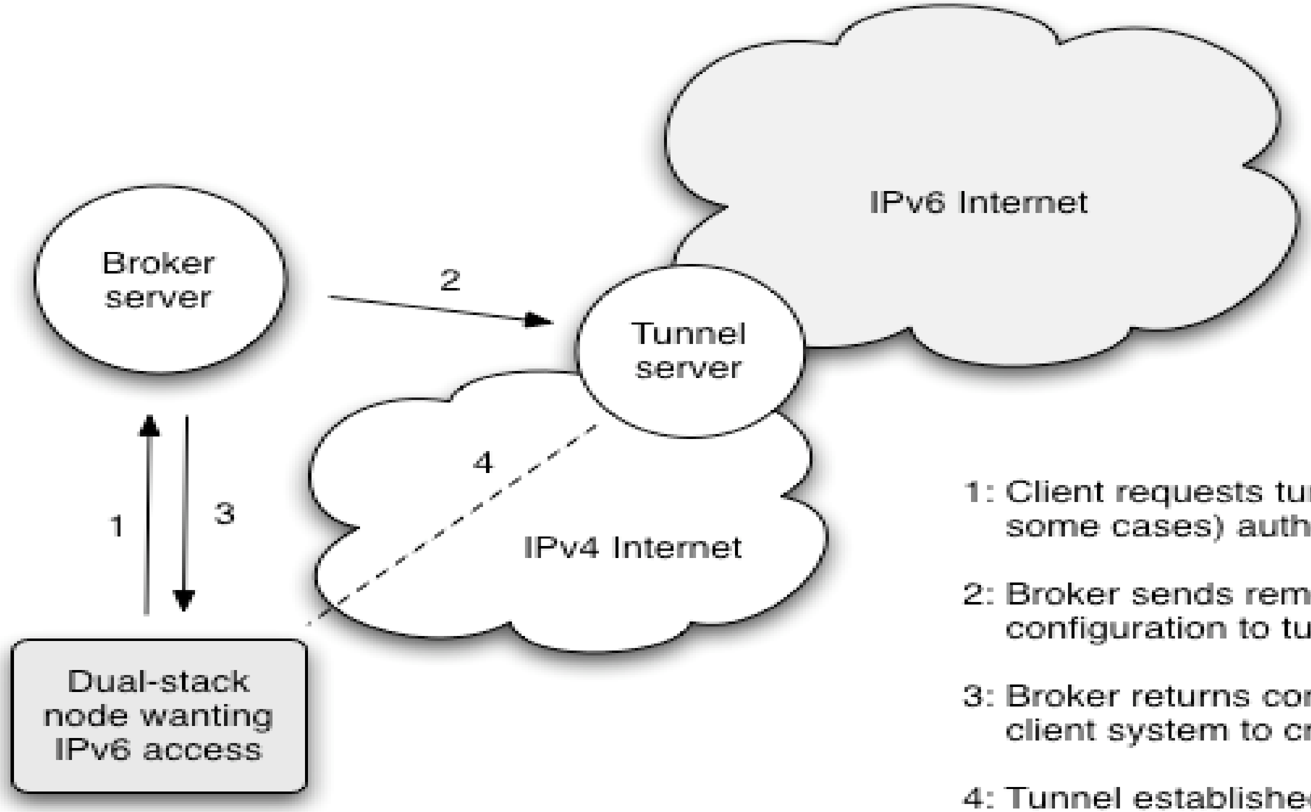
Manually configured tunnels

- Very easy to setup and configure
- Good management potential
  - ISP configures all tunnels, so is in control of its deployment
  - This is the current approach used by many NRENs (including UKERNA and Renater) to connect academic sites/users over IPv6 where native IPv6 connectivity is not available
- Usually used router-to-router or host-to-router
  - Desirable to allow end user to register (and subsequently authenticate) to request a tunnel
  - The IPv6 Tunnel Broker (RFC3053) offers such a system, usually for host-to-router connectivity, but sometimes for router-to-router.

Tunnel broker

- Very popular in IPv6 user community
- Most well-known broker is www.freenet6.net
  - Hosted in Canada by Hexago
- General mode of operation is:
  - User/client registers with the broker system
  - A tunnel is requested from a certain IPv4 address
  - The broker sets up its end of the requested tunnel on its tunnel server
  - The broker communicates the tunnel settings to the user, for client-side configuration
- Can traverse a NAT, e.g. if UDP tunnelling used

Broker architecture



1: Client requests tunnel and (in some cases) authenticates

2: Broker sends remote endpoint configuration to tunnel server

3: Broker returns configuration to client system to create tunnel

4: Tunnel established

Broker issues

- Broker’s key advantage is its manageability
  - ISP can track usage levels
- A few downsides:
  - If broker is topologically remote, round trip times for data may suffer
    - e.g. using freenet6 in Canada to reach UK sites
  - Not well-suited if IPv4 address is dynamic
    - Common problem in home DSL networks
  - If using a remote tunnel broker, your own ISP may not perceive a demand for IPv6

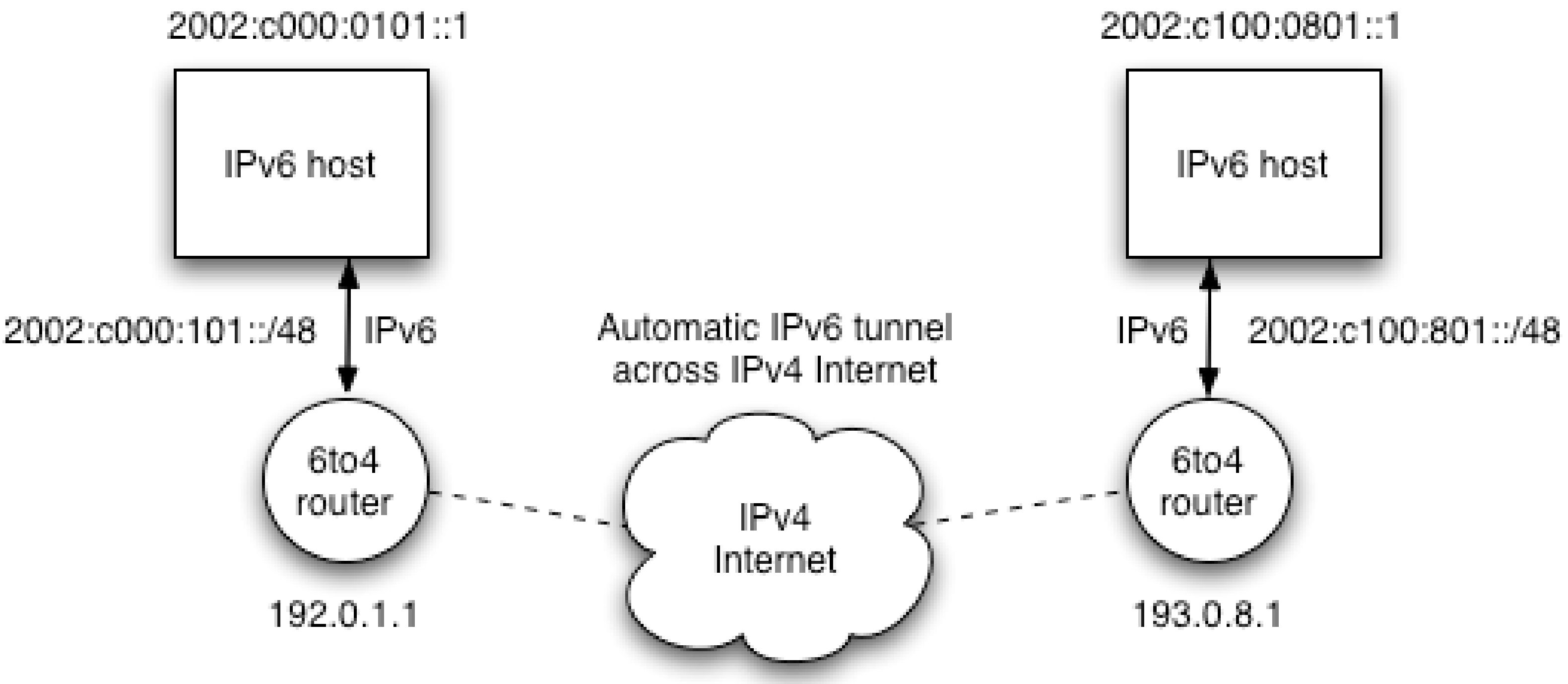
Automatic tunnelling

- Goal is to avoid requiring support staff effort to setup and maintain tunnels
- Set up required tunnels on demand
- Make deployment and usage simple(r) for the end user
- Most common automatic method is 6to4 (RFC3056)
  - Generally used router-to-router
  - Well supported in commercial routing platforms
- Other methods include ISATAP (RFC4214) and Teredo
  - We don’t cover Teredo (RFC4380) here; it is a NAT-traversing IPv6 connectivity method used by Microsoft in XP/Vista.

6to4

- In its basic configuration, 6to4 is used to connect two IPv6 islands across an IPv4 network
- Uses special ‘trick’ for the 2002::/16 IPv6 prefix that is reserved for 6to4 use
  - Next 32 bits of the prefix are the 32 bits of the IPv4 address of the 6to4 router
  - For example, a 6to4 router on 192.0.1.1 would use an IPv6 prefix of 2002:c000:0101::/48 for its site network
- When a 6to4 router sees a packet with destination prefix 2002::/16, it knows to tunnel the packet in IPv4 towards the IPv4 address indicated in the next 32 bits

6to4 basic overview



6to4 features

- On the plus side:
  - Simple to deploy and use
  - Fully automatic; no administrator effort per tunnel
  - Tunnelled packets automatically route efficiently to the destination network (following the best IPv4 path over the IPv4 Internet)
- But there’s an important capability missing:
  - How does a node on a 6to4 site communicate with an IPv6 node on a regular, ‘real’ IPv6 site?
    - Without requiring all IPv6 sites to support 6to4

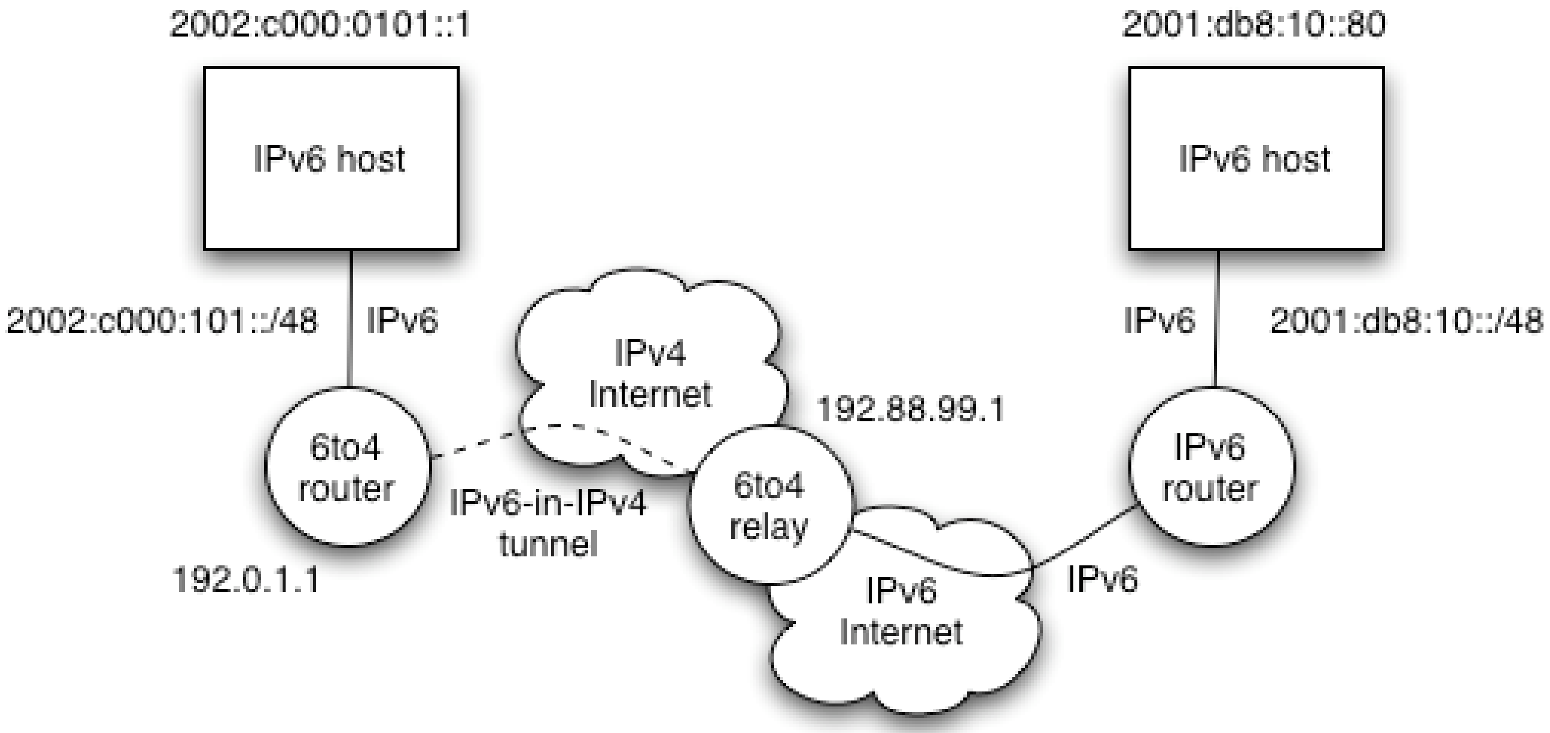
6to4 relay

- A 6to4 relay has a 6to4 interface and a ‘real’ IPv6 interface
- Two cases to consider:
  - IPv6 packets sent from a 6to4 site to a destination address outside 2002::/16 are tunnelled using 6to4 to the relay, are decapsulated, and then forwarded on the relay’s ‘real’ IPv6 interface to the destination site
  - IPv6 packets sent from a ‘real’ IPv6 site towards an address using the 2002::/16 prefix (a 6to4 site) are routed to the 6to4 relay and then tunnelled using 6to4 to the destination 6to4 site

Routing to/from the relay

- The 6to4 relay needs to be ‘discovered’ by routers in the 6to4 world and in the ‘real’ IPv6 Internet
  - All 6to4 routers are configured to tunnel to an anycast address for the relay, for which 192.88.99.1 has been assigned. The 6to4 relay effectively advertises a host route for this address, allowing 6to4 routers to use the topologically nearest 6to4 relay
  - The 6to4 relay advertises 2002::/16 to the ‘real’ IPv6 Internet using BGP or a similar routing protocol

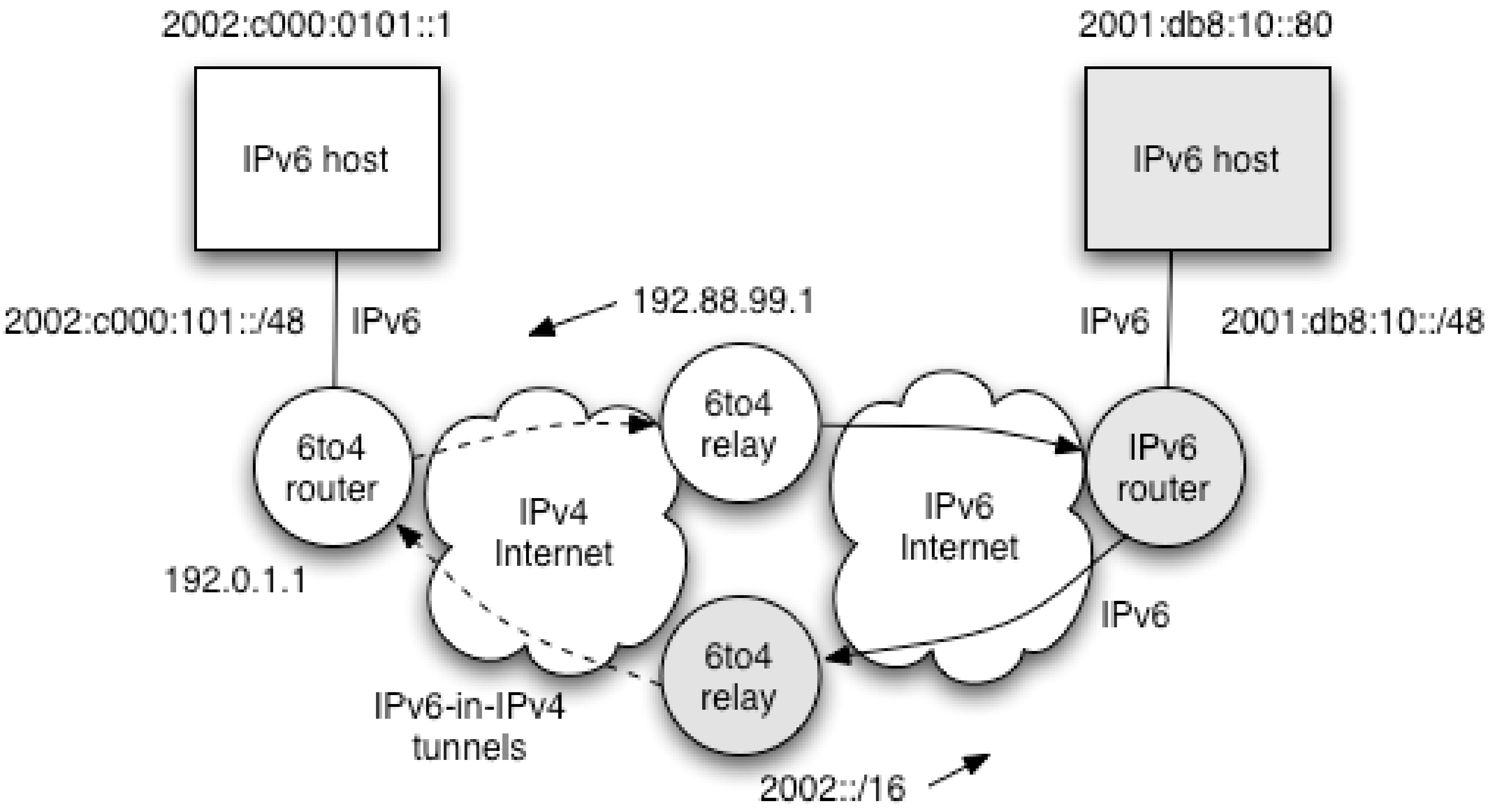
6to4 with relay



6to4 issues

- In principle 6to4 is attractive
  - But there are operational concerns
- Problem 1: possible relay abuse
  - Relay could be used for a DoS attack
  - Tunnelled IPv6 traffic addresses may be spoofed
- Problem 2: asymmetric model/reliability
  - The 6to4 site may use a different 6to4 relay to the ‘real’ IPv6 site
  - One of the sites may not see a 6to4 relay at all, if ISPs choose to only deploy relays for their own customers, and thus filter routing information
- But for 6to4 relay to 6to4 relay operation, it’s good

Asymmetric 6to4



Looking back at considerations

- How do 6to4 and the tunnel broker fare for:
  - Security
  - Manual or automatic setup
  - Ease of management
  - Handling dynamic IPv4 addresses
  - Support for hosts or sites to be connected
  - Scalability: 10, 100, or 10,000 served tunnels?
  - Support for NAT traversal
  - Tunnel service discovery
  - Support for special services (e.g. multicast)
  - Tunnel concentration/bandwidth usage issues
- Have a think and we’ll discuss next time

6to4 and broker features

Feature	6to4	Tunnel broker
Security	Potential for abuse	Supports authentication
Setup	Automatic	Manual
Ease of management	Poor (automatic)	Good
Dynamic IPv4 addresses	Poor	Poor
Host or site tunnels	Primarily site	Primarily host
Scalability	Very good	Good
NAT traversal	Tricky	Yes, with TSP
Tunnel service discovery	Automatic	Manual configuration
Special service support	Variable	Variable
Bandwidth concentration	Only at 6to4 relay	At tunnel server

ISATAP

- Intra-Site Automatic Tunnel Addressing Protocol (RFC4214)
  - Automatic tunneling
  - Designed for use *within* a site
  - Used where dual-stack nodes are sparsely deployed in the site (very early deployment phase)
- Host-to-host or host-to-router automatic tunnels
  - Works by using a specific EUI-64 host address format
  - Format can be recognized and acted upon by ISATAP-aware nodes and routers

ISATAP addresses

- The EUI-64 is formed by
  - A reserved IANA prefix (00-00-5e)
  - A fixed 8-bit hex value (fe)
  - The 32-bit IPv4 address of the node
  - Toggling the globally unique (u) bit
- For example, 152.78.64.1 would have an EUI-64 host address for IPv6 of:
  - 0200:5efe:984e:4001

ISATAP tunneling

- Relies on the OS supporting ISATAP
- Use one ISATAP router per site, usually advertised under FQDN ‘isatap.domain’
  - Virtual IPv6 link over the IPv4 network
  - Know the IPv4 tunnel end-point address from last 32-bits of the IPv6 ISATAP address
  - Get network prefix via ND from router
- Not widely deployed
- Better to deploy proper dual-stack
  - Allows much better managed control of deployment

2: Translation

- When an IPv4-only system needs to communicate with an IPv6-only system some form of translation is required
- Can be done at various layers
- Network layer
  - Rewrite IP headers
- Transport layer
  - Use a TCP relay
- Application layer
  - Use an application layer gateway (ALG)
- Ideally avoid translation
  - Use IPv4 to speak to IPv4 systems and IPv6 for IPv6 systems

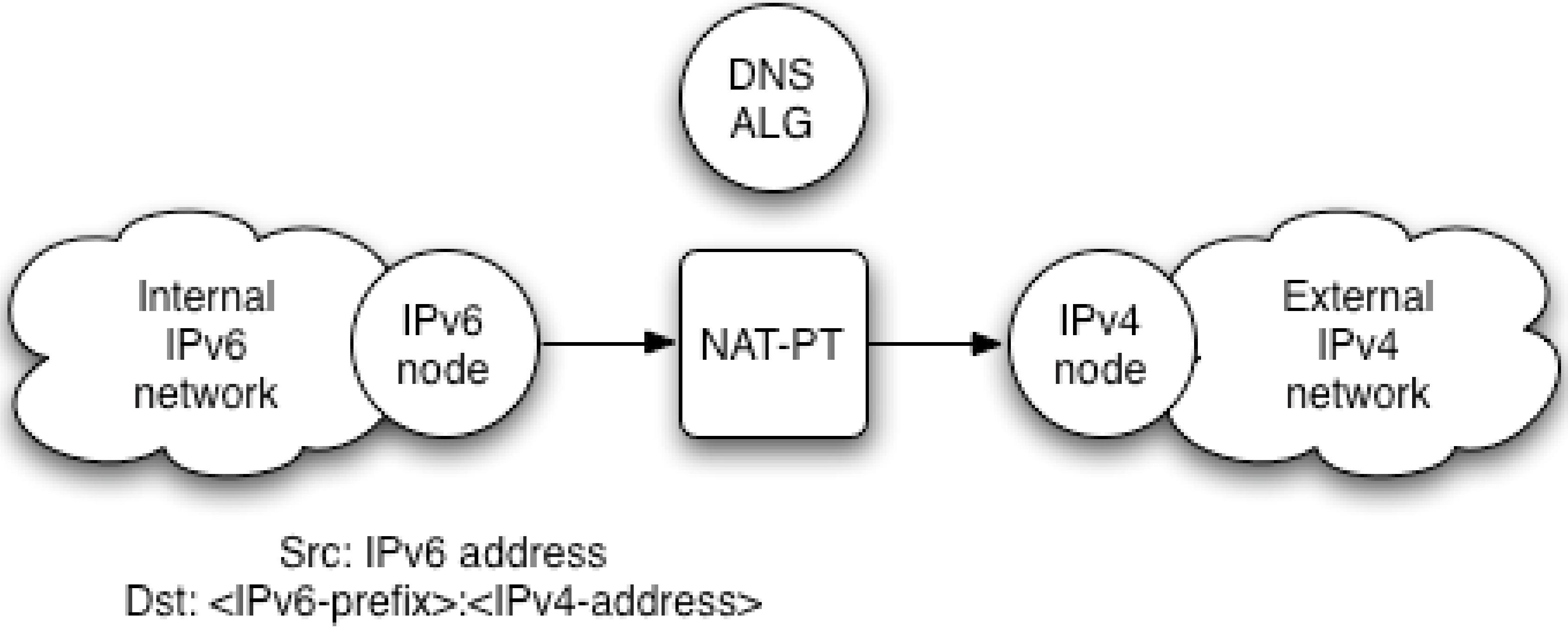
Translation scenarios

- Generally when deploying IPv6-only network elements and you need them to communicate with IPv4-only systems
  - Legacy applications that cannot be ported to support IPv6
    - Or perhaps source code not available
  - Legacy IPv4-only operating systems
    - For example Windows 98
  - Legacy IPv4-only hardware
    - Printers

Network layer: NAT-PT

- Network Address Translation - Protocol Translation
  - Defined in RFC2766
  - Like IPv4 NAT, but with protocol translation
- Uses Stateless IP/ICMP Translation (SIIT)
  - Defined in RFC2765
  - SIIT defines algorithms to translate between the IPv4 and IPv6 header fields, where possible to do so
- NAT-PT extends SIIT by using IPv4 address pools
  - IPv4-to-IPv6 and IPv6-to-IPv4 supported

NAT-PT topology



Src: IPv6 address
Dst: <IPv6-prefix>:<IPv4-address>

NAT-PT and DNS

- Internal network IPv6 only
- DNS ALG watches for IPv6 (AAAA) DNS queries outbound, and translates to IPv4 (A) queries
- When IPv4 DNS response comes back, DNS ALG maps the result to an IPv6 address
  - <IPv6-prefix>:<IPv4 address>
  - A special NAT-PT IPv6 prefix is taken from the IPv6 network’s address space
- Querying host can now use an IPv6 destination that NAT-PT can map to the real IPv4 destination

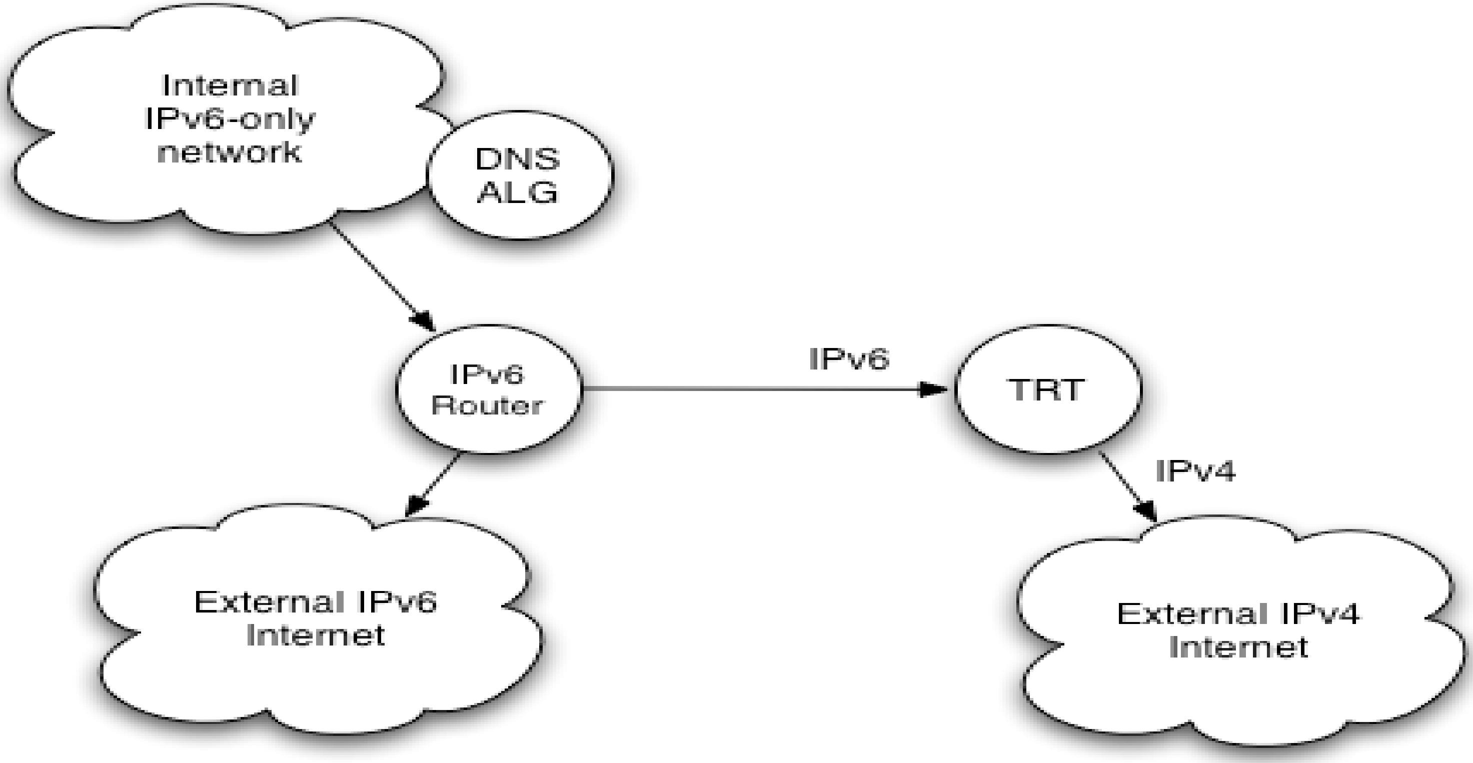
NAT-PT downsides

- Has all the shortcomings of IPv4 NAT, and more
  - Needs state to be held in the NAT-PT device
  - Needs to handle IP addresses information embedded in packet payload (e.g. FTP)
  - DNS considerations are complex
- Can use from IPv4 network into IPv6 network
  - If have enough IPv4 global addresses available to advertise special NAT-PT prefix addresses externally
- It’s considered a ‘last resort’ mechanism
  - NAT-PT is deprecated within the IETF

Transport layer: TRT

- Transport Relay Translator (TRT)
  - Designed for use in IPv6-only networks wishing to connect to external IPv4-only systems
  - TRT has internal IPv6 and external IPv4 interfaces
- External IPv6 connections work as usual
- Trick is handling connections to IPv4 networks
  - Relies on use of a DNS proxy
  - Internal IPv6 host looks up IP address of destination
  - If an IPv6 address, traffic is sent out to IPv6 Internet
  - If an IPv4 address, traffic needs to route to the TRT

TRT topology



DNS proxy address mapping

- If internal IPv6 host is trying to reach an IPv4-only system, the DNS proxy (ALG) returns a special IPv6 destination
  - First 64 bits assigned to be unique locally
  - Next 32 bits all zero
  - Last 32 bits are the real IPv4 destination
    - <IPv6-prefix>:0:0:<IPv4 address>
- The <ipv6-prefix> is routed internally to the TRT
  - Which terminates the TCP/IPv6 connection
  - And opens a connection to the real IPv4 destination

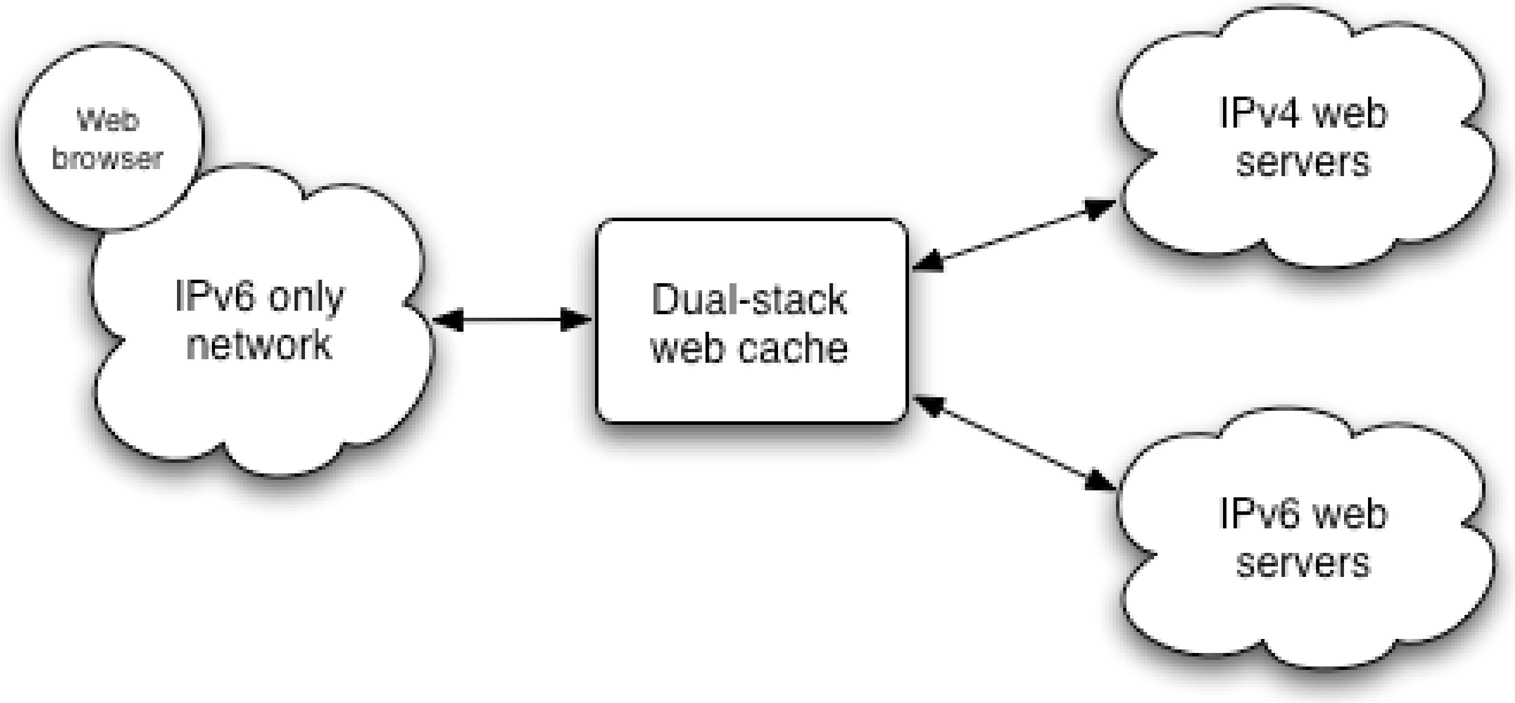
TRT pros and cons

- Pros
  - Transparent to hosts/applications
  - Scalable - can use multiple TRTs, with one internal /64 prefix used per TRT device
  - TRT can work with one global IPv4 address
- Cons
  - Like NAT, problems with embedded IP addresses in payload (e.g. FTP)
  - No simple way to allow connections initiated inbound from external IPv4 to internal IPv6 hosts

Application: ALGs

- NAT-PT and TRT are somewhat complex
- Luckily, application layer gateways (ALGs) offer a simpler alternative
- Many applications support ALGs already
  - Web cache
  - SMTP gateway
  - DNS resolver
  - SIP proxy
  - etc
- We can leverage this in a simple way

ALG topology



ALG pros and cons

- Pros
  - Simple to deploy
  - ALGs already commonly in use, e.g.
    - Web cache to reduce bandwidth usage
    - SMTP relay to channel mail through one server
  - Avoids complexity of NAT-PT or TRT
- Cons
  - Requires client configuration to use ALG
  - Only works for specific ALG-supported applications

3: Dual-stack

- Support both protocols on nodes
- Requires support in:
  - Host platforms
  - Router platforms
  - Applications and services
    - e.g. web, DNS, SMTP
- Adds considerations for
  - Security in all components
  - New policies dependent on IPv6-specific features

Dual-stack issues

- Application must choose which IP protocol to use
  - Given DNS returns IPv4 and IPv6 addresses
  - e.g. MSIE prefers IPv6
  - Don’t advertise AAAA record for a host unless you have good IPv6 connectivity (for all services on host)
- Enabling IPv6 should not adversely impact IPv4 performance
- Security should be no worse
  - Hosts listen on both protocols; secure both

Conclusions

- There is a large set of IPv6 transition tools available
  - No single ‘best’ solution
  - Transition plan is likely to be site-specific
- Current ‘best practice’ is dual-stack deployment
  - Natural path via procurement cycles
  - Allows experience in IPv6 operation to be gained early
- IPv6-only networks can be deployed
  - But very limited in number to date, and missing some apps
- Ultimate driver is IPv4 address space availability
  - But also need IPv4 addresses for a smooth transition