

# IPv6 Deployment Concepts

**Tony Hain**

**Cisco Systems**

**ahain@cisco.com**

- **Review of basics**
- **Environment descriptions**
- **Tools appropriate for each environment**

# Review

# Do We Really Need a Larger Address Space?

Cisco.com

- **Internet Users or PC**
  - ~530 million users in Q2 CY2002, ~945 million by 2004  
(Source: Computer Industry Almanac)
  - Emerging population/geopolitical and Address space
- **PDA, Pen-Tablet, Notepad,...**
  - ~20 millions in 2004
- **Mobile phones**
  - Already 1 billion mobile phones delivered by the industry
- **Transportation**
  - 1 billion automobiles forecast for 2008
  - Internet access in Planes
- **Consumer devices**
  - Billions of Home and Industrial Appliances

# IP Address Allocation History

Cisco.com

**1981 - IPv4 protocol published**

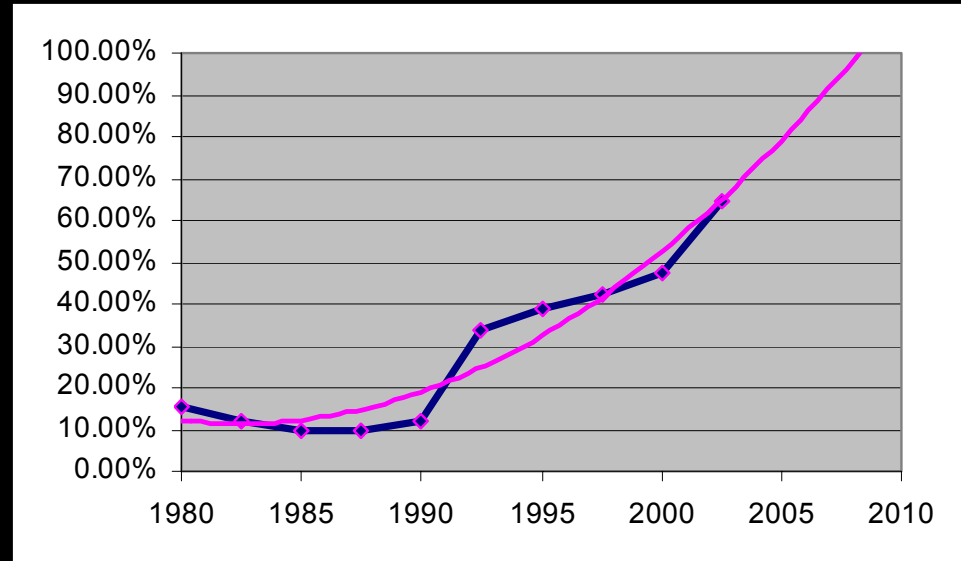
**1985 ~ 1/16 of total space**

**1990 ~ 1/8 of total space**

**1995 ~ 1/3 of total space**

**2000 ~ 1/2 of total space**

**2002.5 ~ 2/3 of total space**

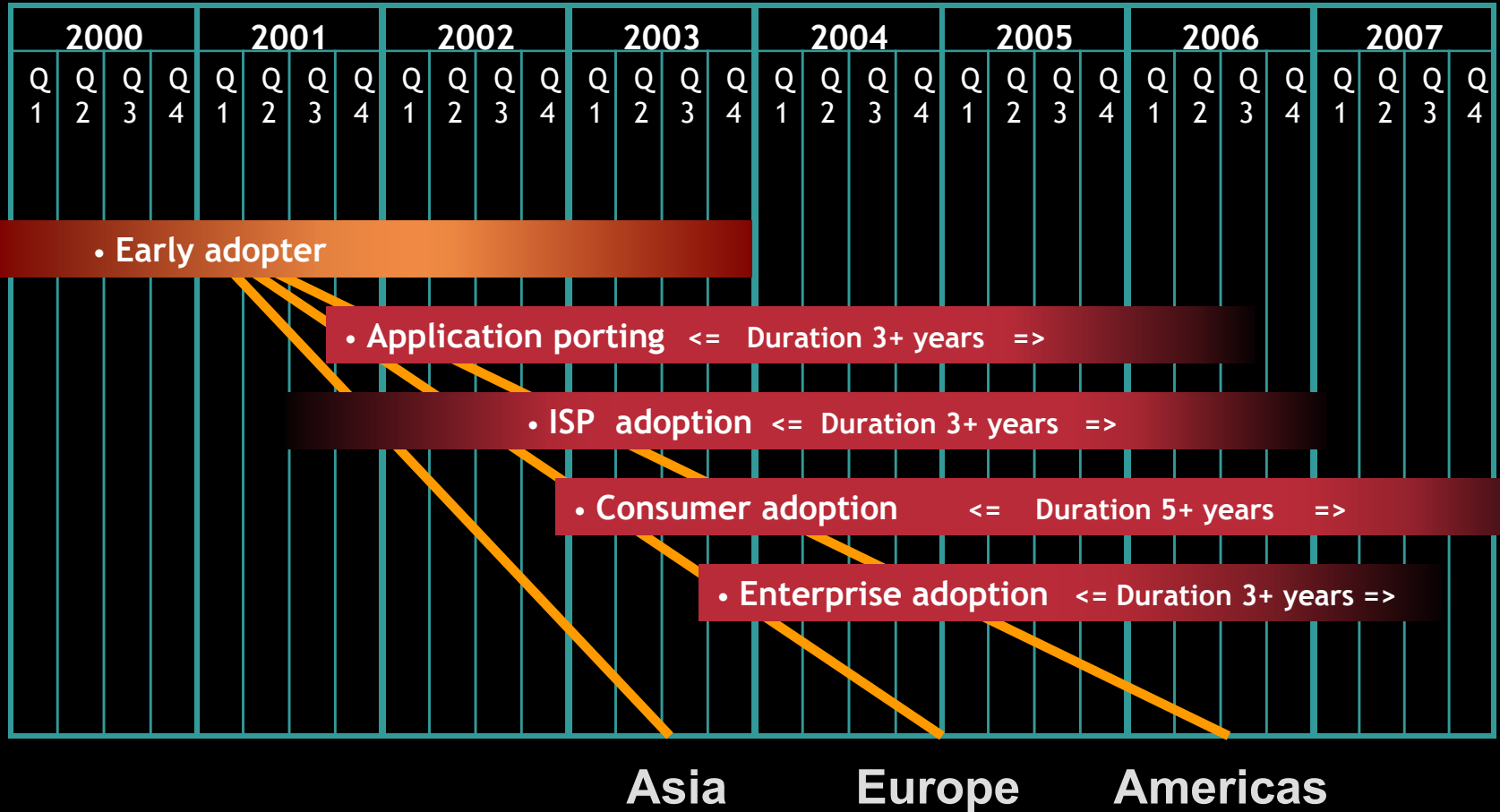


- **This despite increasingly intense conservation efforts**
  - PPP / DHCP address sharing
  - NAT (network address translation)
  - CIDR (classless inter-domain routing)
  - plus some address reclamation
- **Theoretical limit of 32-bit space: ~4 billion devices**  
**Practical limit of 32-bit space: ~250 million devices (RFC 3194)**



# IPv6 Timeline

(A pragmatic projection)



# IPv6 Technology Scope

Cisco.com

<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>



# IPv4 & IPv6 Header Comparison

## IPv4 Header

Version	IHL	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				

- Legend**
- field's name kept from IPv4 to IPv6
  - fields not kept in IPv6
  - Name & position changed in IPv6
  - New field in IPv6

# IPv6 - Addressing Model

Addresses are assigned to interfaces

change from IPv4 model :

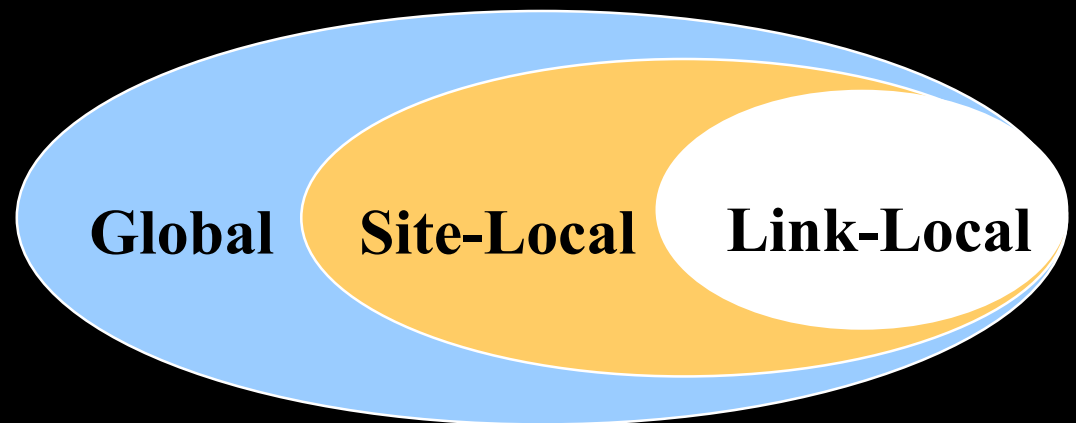
**Interface 'expected' to have multiple addresses**

Addresses have scope

Link Local

Site Local

Global



Addresses have lifetime

Valid and Preferred lifetime

# Interface Address set

Cisco.com

- **Loopback** (only assigned to a single interface per node)
- **Link local** (required on all interfaces)
- **Site local**
- **Auto-configured 6to4** (if IPv4 public is address available)
- **Auto-configured IPv4 compatible** (operationally discouraged)
- **Solicited node Multicast** (required for neighbor discovery)
- **All node multicast**
- **Global anonymous**
- **Global published**

# Unicast Address Formats

## Link Local

FP (10bits)	RESERVED (54bits)	Interface ID (64bits)
1111111010	MUST be 0	MAC derived

## Site Local

FP (10bits)	Subnet (38bits)	Subnet (16bits)	Interface ID (64bits)
1111111011	Locally Administered	Locally Administered	MAC derived or Locally Administered

## Global

FP (3bits)	Registry / provider assigned (45bits)	Subnet (16bits)	Interface ID (64bits)
001	Provider Administered	Locally Administered	MAC derived or Locally Administered or Random

# Tunneling Unicast Address Formats

## Compatible

<b>FP (96bits)</b>	<b>IPv4 ID (32bits)</b>
MUST be 0	Locally administered

## 6to4

<b>FP (16bits)</b>	<b>IPv4 (32bits)</b>	<b>SLA (16bits)</b>	<b>Interface ID (64bits)</b>
00100010	<b>Provider Administered</b>	Locally Administered	MAC derived or Locally Administered or Random

## ISATAP

<b>Any (48bits)</b>	<b>SLA (16bits)</b>	<b>Interface ID (64bits)</b>
Provider Administered	Locally Administered	IPv4 derived

# Multicast Address Format

FP (8bits)	Flags (4bits)	Scope (4bits)	RESERVED (80bits)	Group ID (32bits)
11111111	000T	Lcl/Sit/Gbl	MUST be 0	Locally administered

- **flag field**

low-order bit indicates permanent/transient group  
(three other flags reserved)

- **scope field:**

1 - node local  
2 - link-local  
5 - site-local

8 - organization-local  
B - community-local  
E - global

(all other values reserved)

- **map IPv6 multicast addresses directly into low order 32 bits of the IEEE 802 MAC**

# Multicast Address Format

## Unicast-Prefix based

FP (8bits)	Flags (4bits)	Scope (4bits)	reserved (8bits)	plen (8bits)	Network Prefix (64bits)	Group ID (32bits)
11111 111	00PT	Lcl/Sit/Gbl	MUST be 0	Locally administered	Unicast prefix	Auto configured

- P = 1 indicates a multicast address that is assigned based on the network prefix
- plen indicates the actual length of the network prefix
- Source-specific multicast addresses is accomplished by setting
  - P = 1
  - plen = 0
  - network prefix = 0

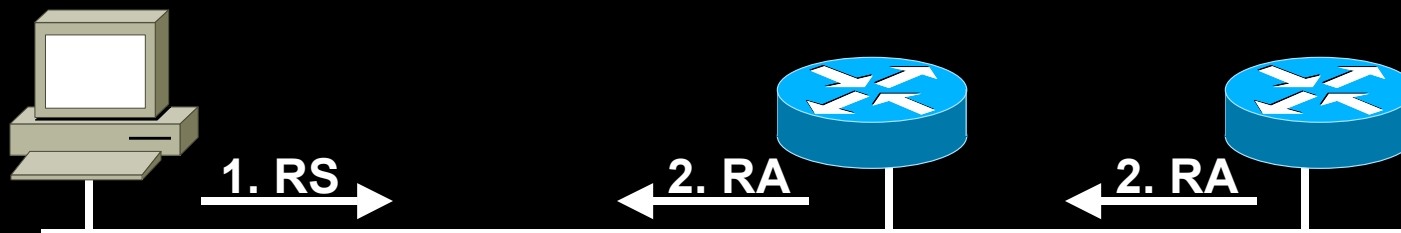
draft-ietf-ipngwg-uni-based-mcast-01.txt

# ICMP / ND Walkthrough



# ND Autoconfiguration, Prefix & Parameter Discovery

Cisco.com



## 1. RS:

ICMP Type = 133

Src = ::

Dst = All-Routers multicast  
Address

query= please send RA

## 2. RA:

ICMP Type = 134

Src = Router Link-local Address

Dst = All-nodes multicast address

Data= options, prefix, lifetime,  
autoconfig flag

- Router solicitation are sent by booting nodes to request RAs for configuring the interfaces.

# ND Address Resolution & Neighbor Unreachability Detection



**ICMP type = 135 (NS)**

**Src = A**

**Dst = Solicited-node multicast of B**

**Data = link-layer address of A**

**Query = what is your link address?**

**ICMP type = 136 (NA)**

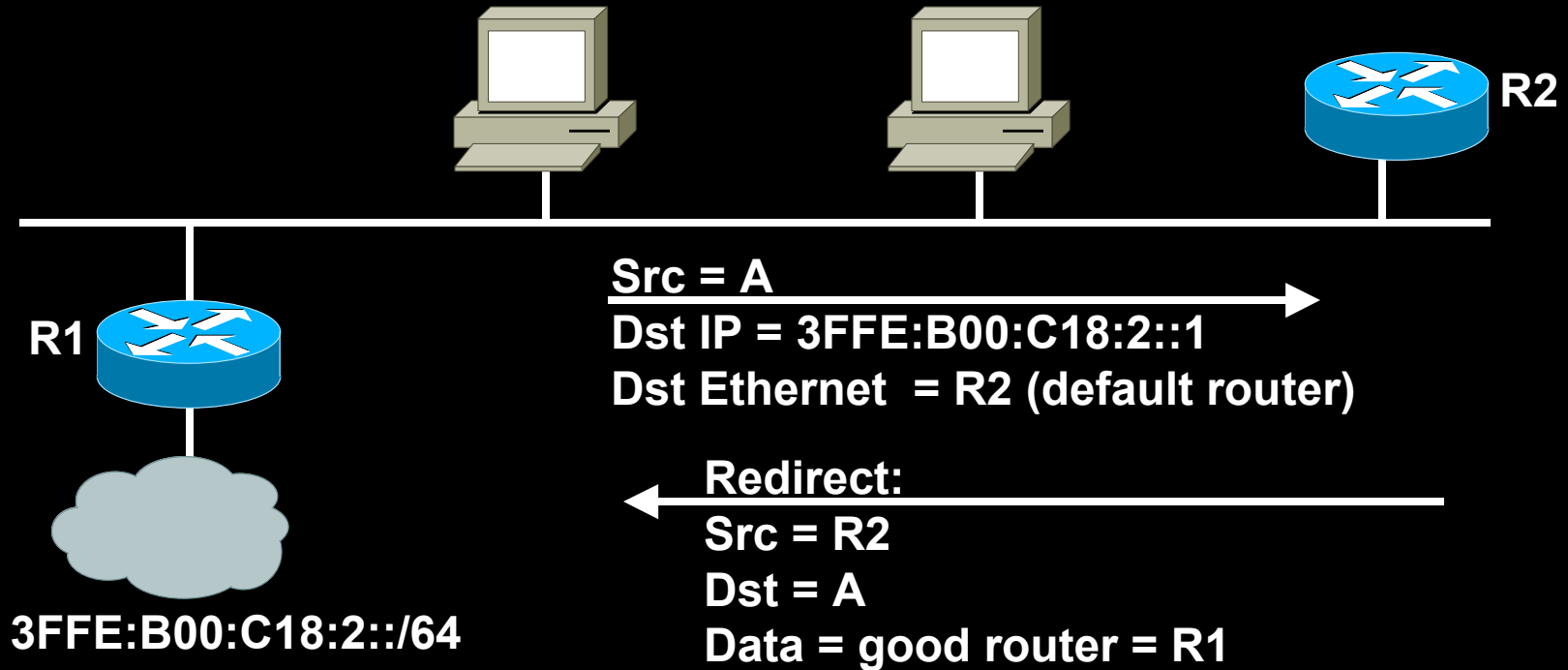
**Src = B**

**Dst = A**

**Data = link-layer address of B**

**A and B can now exchange  
packets on this link**

# ND Redirect

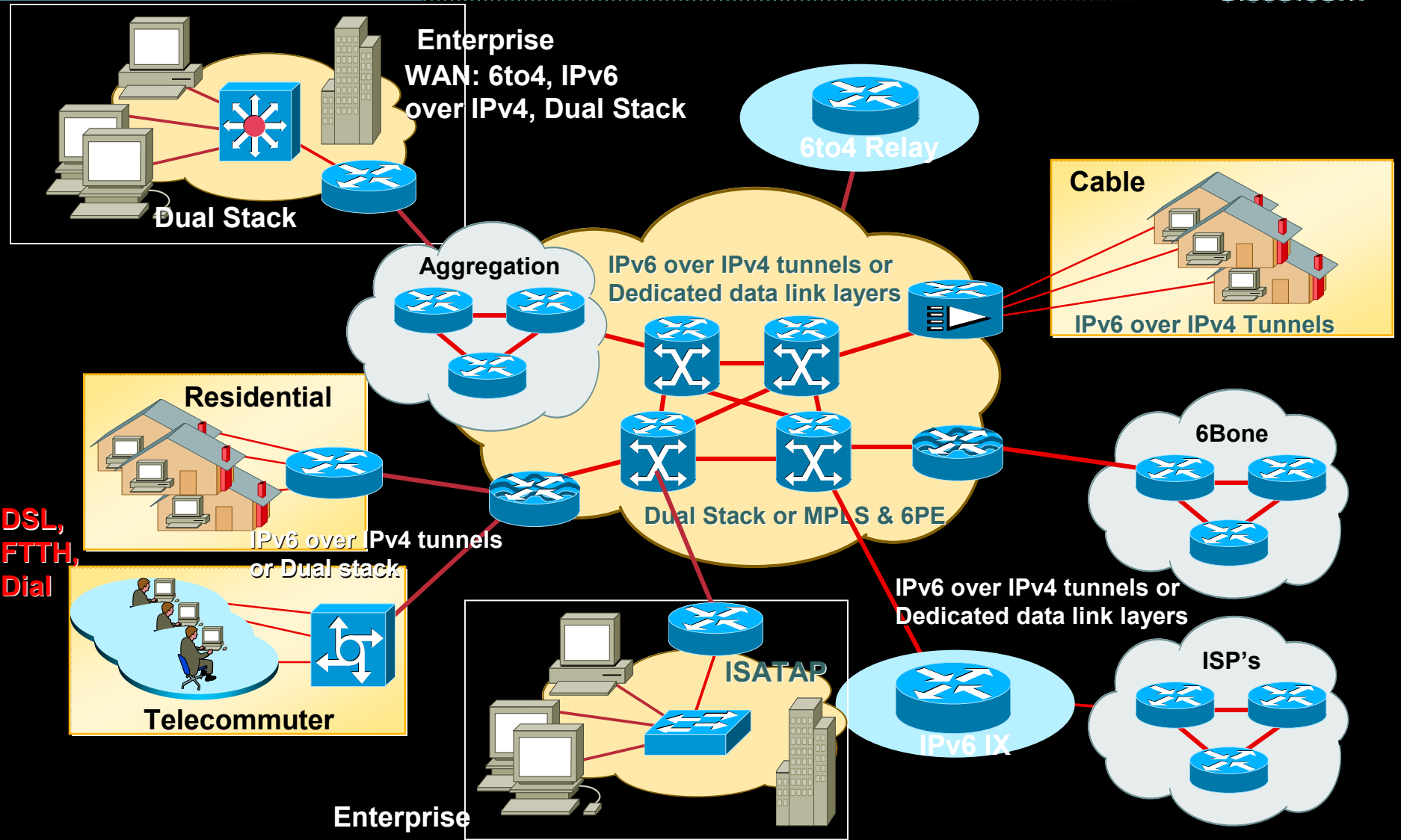


- **Redirect** is used by a router to signal the reroute of a packet to an onlink host to a better router or to another host on the link

- **Review of basics**
- **Environment descriptions**
- **Tools appropriate for each environment**

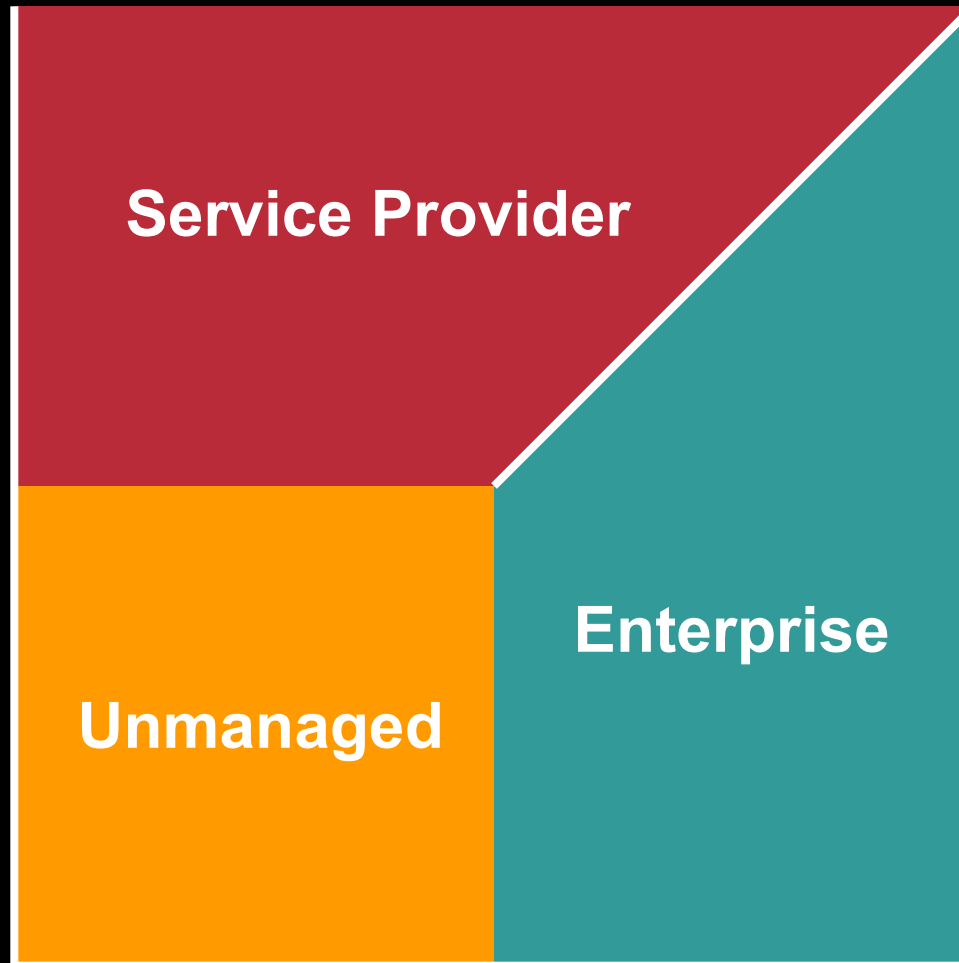
# Environments

# Transition environments

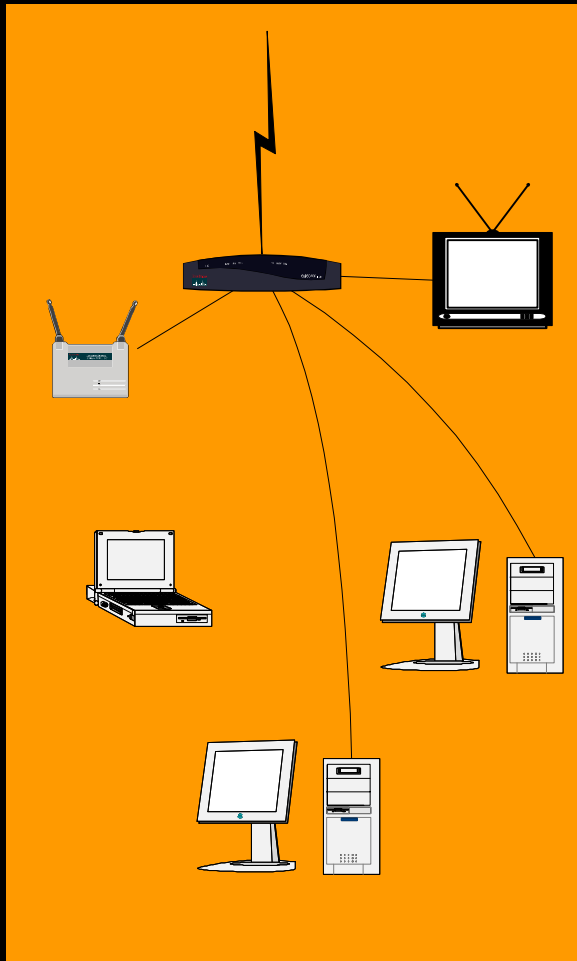


DSL,  
FTTH,  
Dial

# Environments



# Environments – Unmanaged



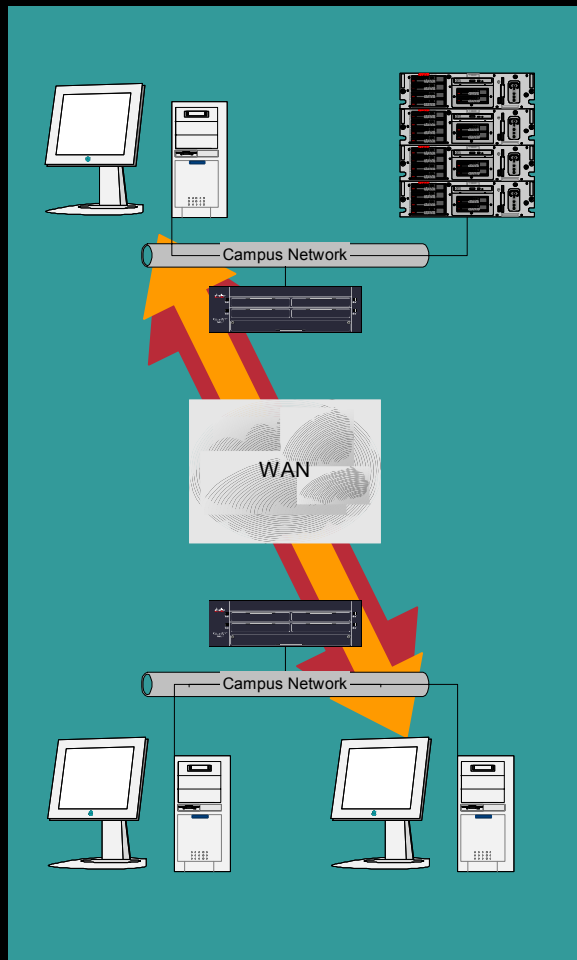
- **No administrative staff to manage configuration or policies**
- **Devices need to be plug-n-play appliances**
- **Network & hosts share administrative policies**
- **Tool automation a primary concern**



- **ISP offers IPv6 service**  
Edge device acquires a prefix to redistribute
- **ISP still IPv4-only service**  
(may be due to device limitations like docsis modems)  
Tunneling required  
Prefix from tunnel broker or automated 6to4/Teredo
- **If no auto-tunnel to native relays, may need both**

# Environments – Managed Enterprise

Cisco.com



- **Dedicated management staff & tools**
- **Network & hosts share administrative policies**
- **Applications will likely require recertification**

# Managed networks differentiation

Cisco.com

**Single geographic region, single administration & policy**

**Multiple geographic regions, single administration & policy**

**Multiple geographic regions, multiple administrations & policy**

**Use of public network for transit service**

Simple routed case looks like multi-multi above

VPN tunneled case would look like multi-single w/circuit setup

**New enterprise, looking to avoid a transition**

Deployment order - All at once by definition

**For each of the 5 categories consider**

Deployment order - Hosts & Apps first vs. Network first

ISP offering - IPv4-only    IPv4 & IPv6    IPv6-only

# Infrastructure concerns

- **Critical Applications**
- **Addressing : Dynamic vs. controlled**
- **DNS : Dynamic vs. controlled**  
**Public visibility of name space**
- **AAA : Internal & external**  
**Mobility of road warrior & telecommuters**  
**Mobility of nodes within the enterprise**
- **ICMP : PMTU & neighbor discovery**
- **Management tools**  
**Trust between host & network management teams**

# Multiple Address Issues

Cisco.com

**Renumbering simplified as old & new can overlap**

**Privacy addresses reduce attack profile**

**Preferred vs. valid lifetimes**

**Improper configuration could lead to 100's per interface**

**Diagnostics require more effort**

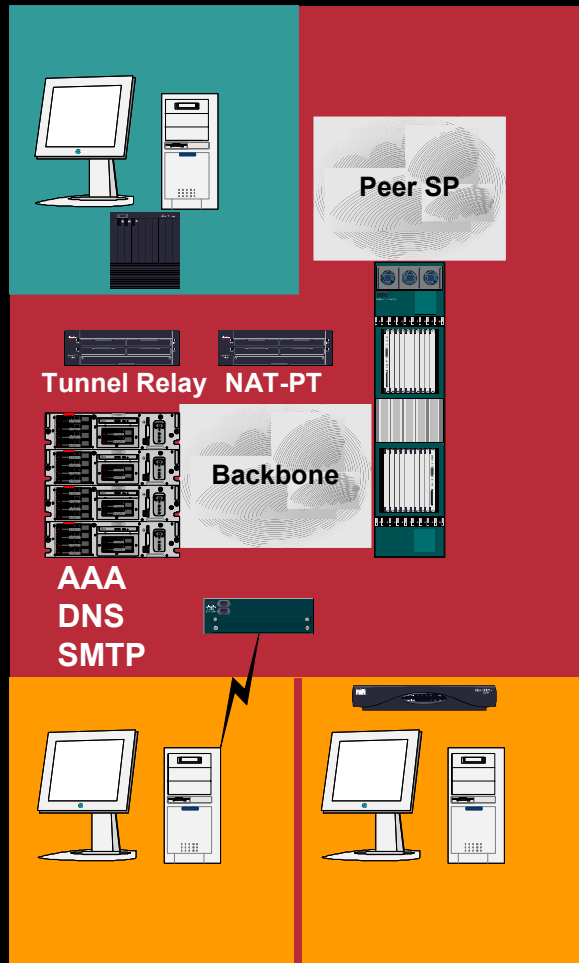
**TE via addressing limits multi-homing flexibility**

**Site-local allows internal stability**

- **Allocations of  $::/48$  should allow self aggregation by organizations with multiple IPv4 prefixes**
- **Tunneling**
  - Decouples network from end system deployment**
  - Multicast less efficient**
- **Native service**
  - May require hardware upgrades**

# Environments – Managed Service Provider

Cisco.com



- **Dedicated management staff & tools**
- **Network has different administrative policies than connected hosts or networks**
- **Interaction with Peer networks may require translation**
- **Services as Dual-stack**
- **Distributed tunnel relay service minimizes overhead**

# Address Allocation Issues

- **From Regional Registries**

  - ::/32 minimum**

    - HD ratio based on .8 utilization of ::/48s

- **To Customers**

  - ::/48 Prefix delegation via DHCPv6**

    - (normal customer allocation)

  - ::/64 Prefix delegation via RA or DHCPv6**

    - (for single subnet sites, ie: 802.11 hotspots)

- **RFC 3041 addresses allow end system anonymity as they move between networks, but the allocated prefix still allows customer identification for LI conformance**



- **Allocations should allow massive aggregation**

**Current allocation policy all PA based, so global BGP table should approach number of origin AS's**

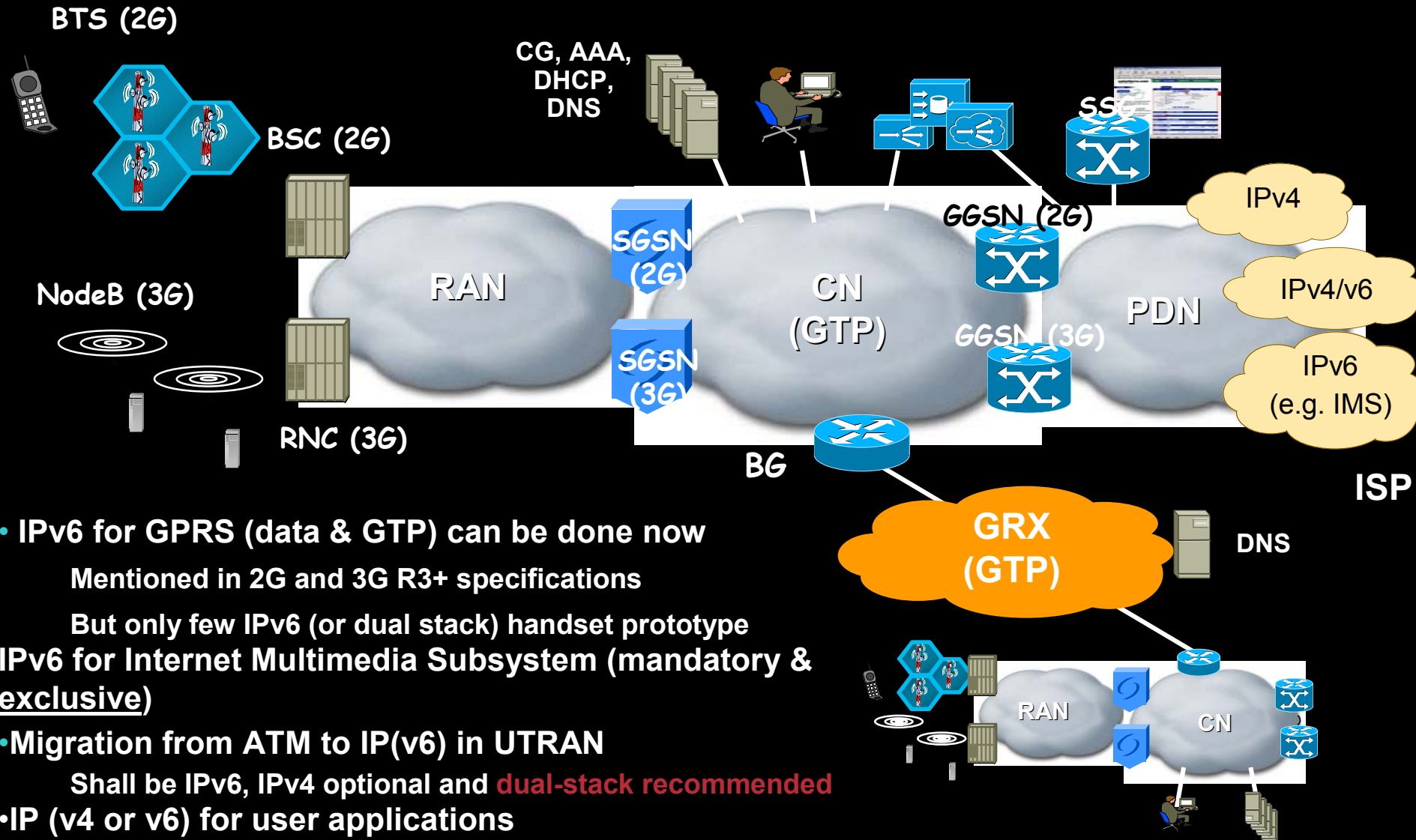
- **Multi-homed sites still an unsolved problem**

- **Dual-stack servers**
- **Consistency of the client and referral chain**
- **IPv6 glue records**
- **Sub-domain delegation to consumer customers?**

- **Dual-stack MTAs**
- **Consistency of clients and MX to A/AAAA mappings**
- **Broken DNS servers return 'nxdomain' for missing AAAA**

# 3GPP Mobile Wireless Network Architecture

Cisco.com



- IPv6 for GPRS (data & GTP) can be done now  
Mentioned in 2G and 3G R3+ specifications  
But only few IPv6 (or dual stack) handset prototype
- IPv6 for Internet Multimedia Subsystem (mandatory & **exclusive**)
- Migration from ATM to IP(v6) in UTRAN  
Shall be IPv6, IPv4 optional and **dual-stack recommended**
- IP (v4 or v6) for user applications

- **Independent PDP contexts for each version**
- **Desire to avoid DAD over expensive air link**

**Only possible when air link end point has full control of a `::/64` or shorter prefix**

- **Review of basics**
- **Environment descriptions**
- **Tools appropriate for each environment**

# Deployment tool set

# Transition Variables

- **Business Requirements**

- Time frame required to meet a set of business requirements**

- Need for applications to communicate between administrative domains**

- New functions that can exist without extensive access to legacy IPv4 nodes**

- Mission critical applications that must interoperate with legacy nodes**

- **Network Security Requirements**

- Firewall support for both IPv4 & IPv6**

- Telecommuters and Mobile Node access methods**

- **Availability of software & hardware upgrades for existing nodes**

- Source code availability for custom applications**

- **Order and rate for IPv6 deployment within a network**

- Current use of IPv4 private addresses and NAT**

- Provider support for IPv6**



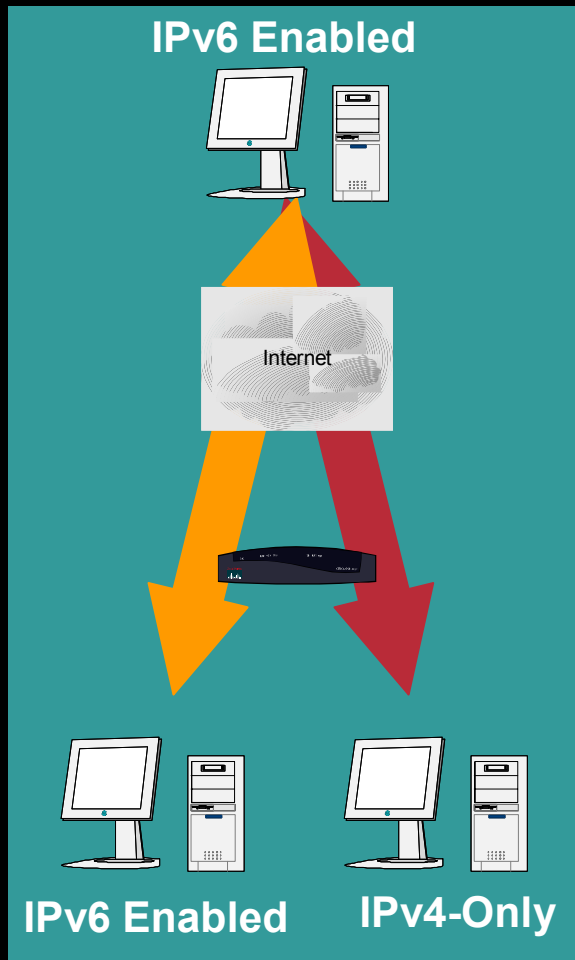
# IPv4-IPv6 Transition / Co-Existence

A wide range of techniques have been identified and implemented, basically falling into three categories:

- (1) **Dual-stack** techniques, to allow IPv4 and IPv6 to co-exist in the same devices and networks
- (2) **Tunneling** techniques, to avoid order dependencies when upgrading hosts, routers, or regions
- (3) **Translation** techniques, to allow IPv6-only devices to communicate with IPv4-only devices

**Expect all of these to be used, in combination**

# Tools – Dual Stack

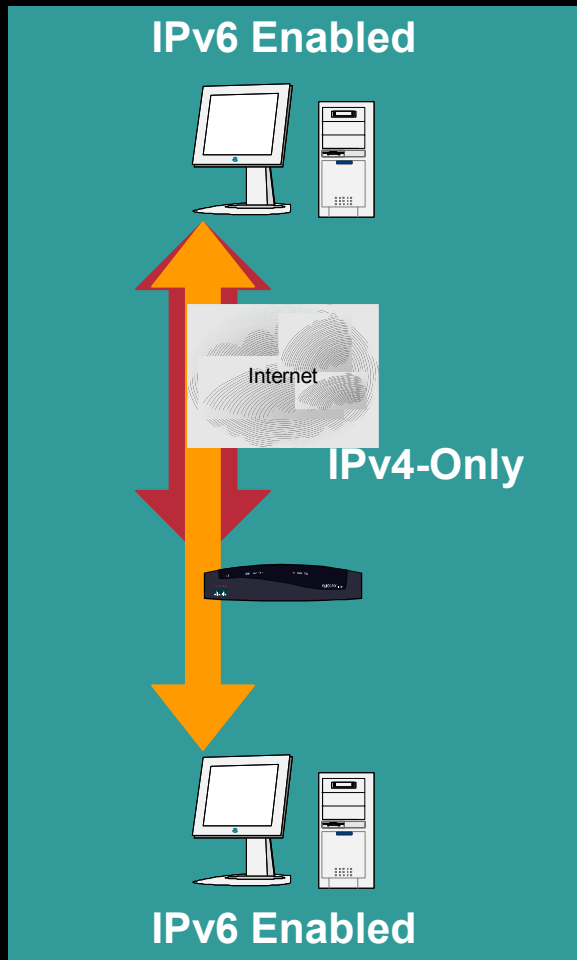


- **Primary tool**
- **Allows continued 'normal' operation with IPv4-only nodes**
- **Address selection rules generally prefer IPv6**
- **DSTM variant allows temporary use of IPv4 pool**

# Dual-Stack Approach

- **When adding IPv6 to a system, do not delete IPv4**  
this multi-protocol approach is familiar and well-understood (e.g., for AppleTalk, IPX, etc.)  
note: in most cases, IPv6 will be bundled with new OS releases, not an extra-cost add-on
- **Applications (or libraries) choose IP version to use when initiating, based on DNS response:**
  - Prefer scope match first, when equal IPv6 over IPv4
  - when responding, based on version of initiating packet
- **This allows indefinite co-existence of IPv4 and IPv6, and gradual app-by-app upgrades to IPv6 usage**

# Tools – Tunneling



- **Nodes view IPv4 network as a logical NBMA link-layer**
- **May be used in conjunction with dual-stack**

# Tunneling issues

- **IPv4 fragmentation needs to be reconstructed at tunnel endpoint.**
- **No translation of Path MTU messages between IPv4 & IPv6.**
- **Translating IPv4 ICMP messages and pass back to IPv6 originator.**
- **May result in an inefficient topology.**

# Tunneling issues II

- **Tunnel interface is always up. Use routing protocol to determine link failures.**
- **Be careful with using the same IPv4 source address for several tunneling mechanisms. Demultiplexing incoming packets is difficult.**

# IPv6 over IPv4 Tunnels

- **Several Tunnelling mechanisms defined by IETF**

**Apply to ISP and Enterprise WAN networks**

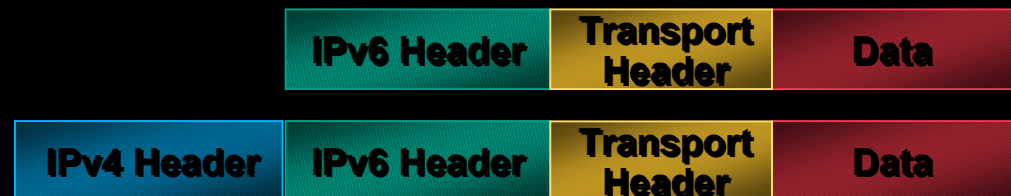
**GRE, Configured Tunnels, Automatic Tunnels using IPv4 compatible IPv6 Address, 6to4**

**Apply to Campus**

**ISATAP, 6over4**

- **No impact on Core infrastructure**

**Either IPv4 or MPLS**



# Tunnels to Get Through IPv6-Ignorant Routers

- **Encapsulate IPv6 packets inside IPv4 packets (or MPLS frames)**
- **Many methods exist for establishing tunnels:**
  - manual configuration**
  - “tunnel brokers” (using web-based service to create a tunnel)**
  - automatic (depricated, using IPv4 as low 32bits of IPv6)**
  - “6-over-4” (intra-domain, using IPv4 multicast as virtual LAN)**
  - “6-to-4” (inter-domain, using IPv4 addr as IPv6 site prefix)**
- **Can view this as:**
  - IPv6 using IPv4 as a virtual NBMA link-layer, or**
  - an IPv6 VPN (virtual public network), over the IPv4 Internet**



# Tunneling Mechanisms (operationally challenging)

Cisco.com

- **Configured**

Prearranged addresses for both IPv4 & IPv6, manually configured

- **Tunnel Broker**

Builds on configured tunnel via IPv4 auth scheme to establish mapping ; typically default route

- **6over4**

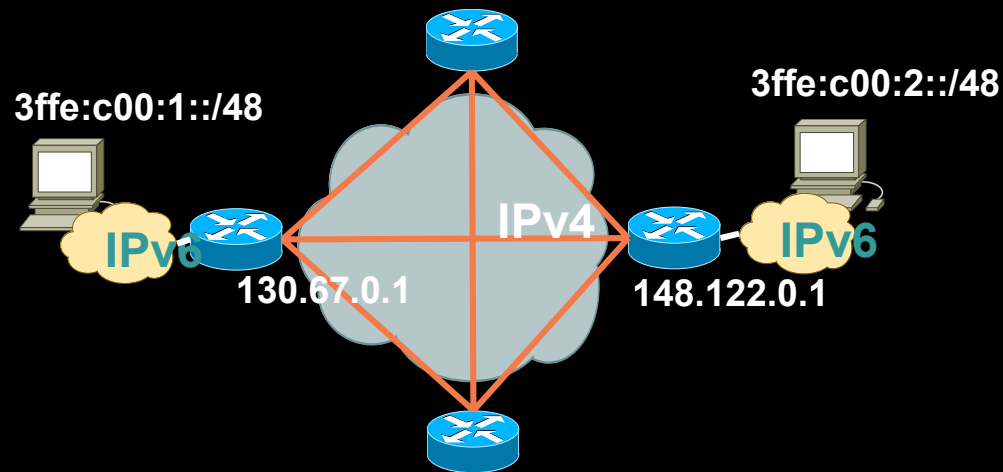
Any address, but requires IPv4 multicast for ND

- **Automatic**

Host-to-host – IPv4 address embedded in low 32 bits with prefix `::/96`

Requires injecting IPv4 BGP table into IPv6 routing

# Configured tunnels



## Pros

As point to point links

Multicast

## Cons

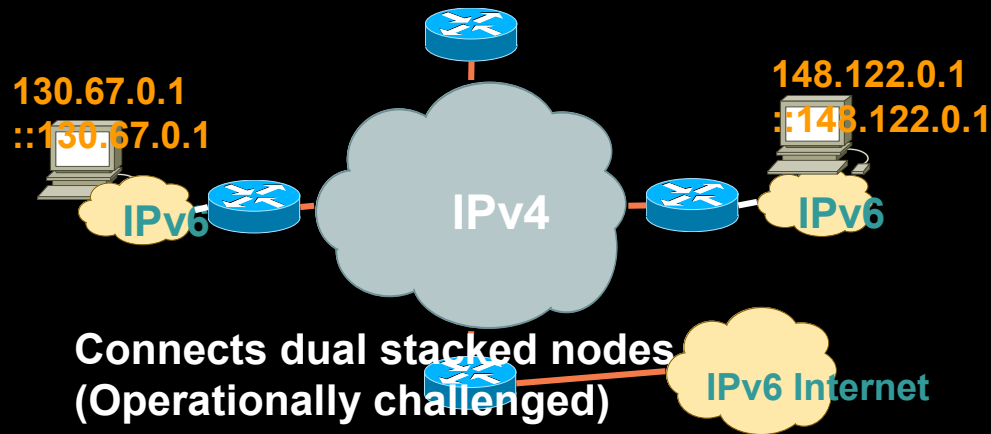
Has to be configured and managed

Inefficient traffic patterns

No keepalive mechanism, interface is always up

# Automatic tunnels

0	IPv4 Address (32bits)
Defined	ISP assigned



## Pros

Useful for some other mechanisms, like BGP tunnels

## Cons

Difficult to reach the native IPv6 Internet, without injecting IPv4 routing information in the IPv6 routing table

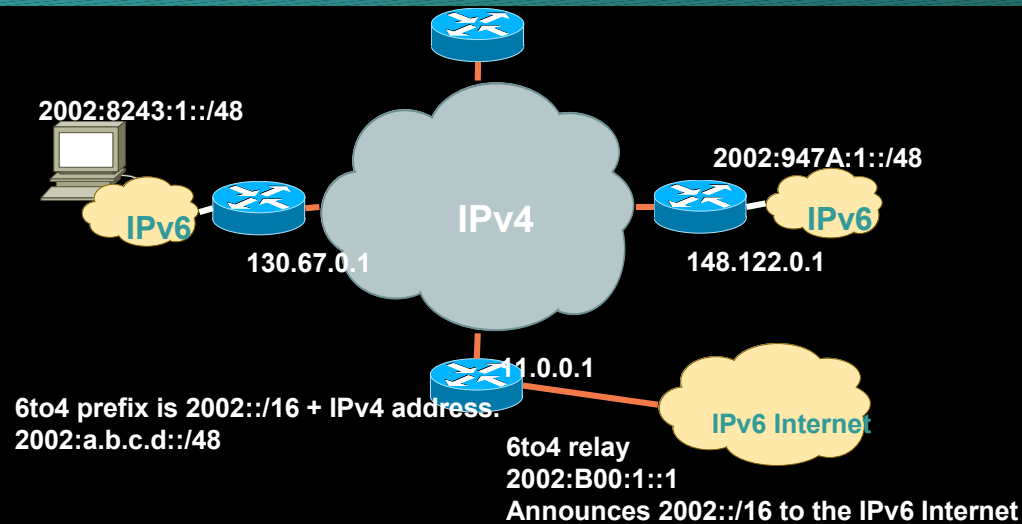
# Tunneling Mechanisms (primary set)

Cisco.com

- **6to4**  
Automatic prefix allocation based on public IPv4
- **ISATAP**  
Intra-site automatic tunneling with any prefix
- **Teredo**  
IPv6 over UDP/IPv4 to traverse NAT

# 6to4 tunnels

FP (3bits)	TLA (13bits)	IPv4 Address (32bits)	SLA ID (16bits)	Interface ID (64bits)
001	0x0002	ISP assigned	Locally administered	Auto configured



## Pros

Works without adjacent native IPv6 routers

Only site border router needs to know about 6to4

## Cons

Requires relay router to reach native IPv6 Internet

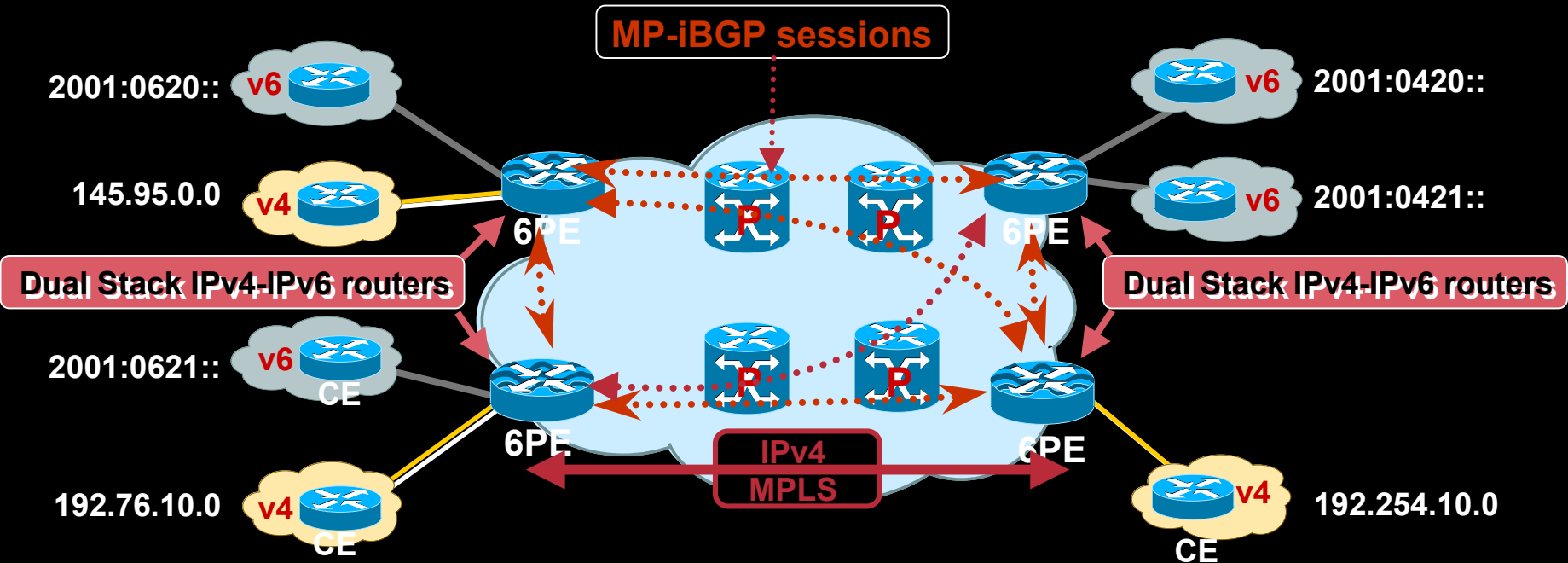
All issues that NMBA networks have.

# IPv6 over MPLS Infrastructure

- **Some Service Providers have already deployed MPLS in their IPv4 backbone for various reasons**
  - MPLS/VPN, MPLS/QoS, MPLS/TE, ATM + IP switching**
- **Several IPv6 over MPLS scenarios**
  - IPv6 Tunnels configured on CE (no impact on MPLS)**
  - IPv6 over Circuit\_over\_MPLS (no impact on IPv6)**
  - IPv6 Provider Edge Router (6PE) over MPLS (no impact on MPLS core)**
  - Native IPv6 MPLS (require full network upgrade)**
- **Upgrading software to IPv6 Provider Edge Router (6PE)**
  - Low cost and risk as only the required Edge routers are upgraded or installed**
  - Allows IPv6 Prefix delegation by ISP**

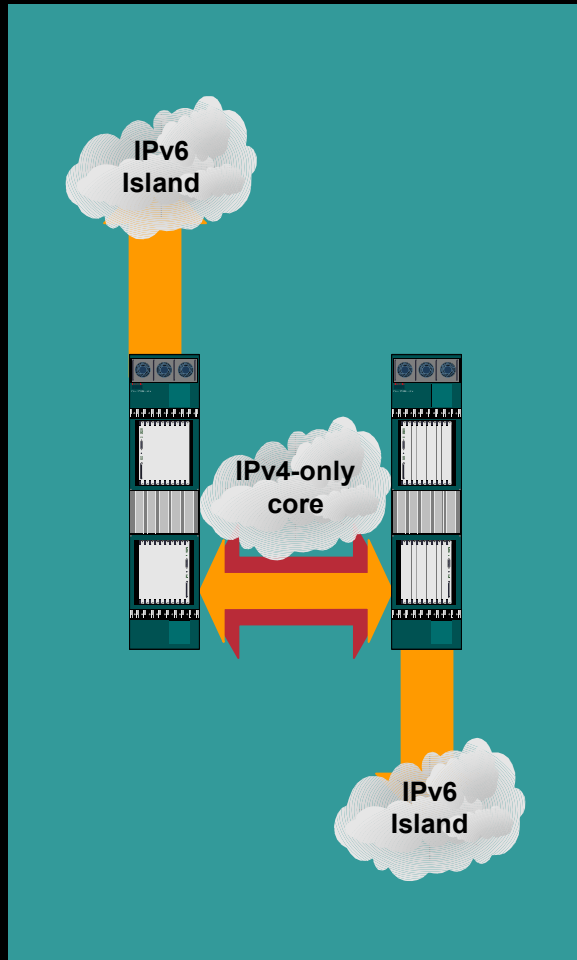
# IPv6 Provider Edge Router (6PE) over MPLS

Cisco.com



- IPv4 or MPLS Core Infrastructure is IPv6-unaware
- PEs are updated to support Dual Stack/6PE
- IPv6 reachability exchanged among 6PEs via iBGP (MP-BGP)
- IPv6 packets transported from 6PE to 6PE inside MPLS

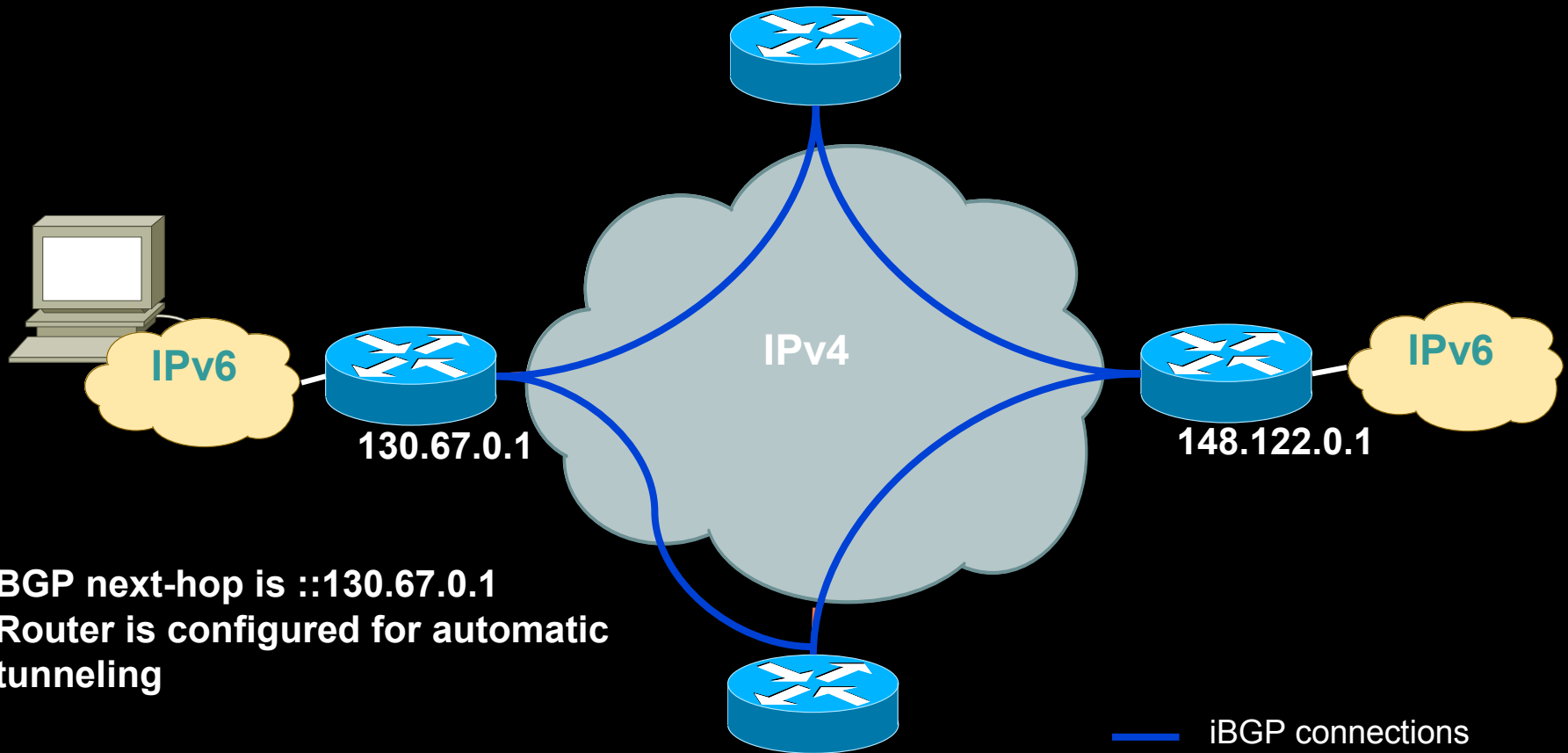
# Tools – BGP tunnel



- Service provider can incrementally upgrade PE routers with active customers
- Sites are connected to Dual Stack MP-BGP-speaking edge router
- Transport across the IPv4 core can be any tunneling mechanism



# BGP tunnels



**Useful for connecting IPv6 PE devices over an IPv4 only core.**

# IPv6 Tunnels & Native Case Study

- **ISP scenario**

Configured Tunnels or Native IPv6 between IPv6 Core Routers

Configured Tunnels or Native IPv6 to IPv6 Enterprise's Customers

Tunnels for specific access technologies eg. Cable

MP-BGP4 Peering with other 6Bone users

Connection to an IPv6 IX

6to4 relay service

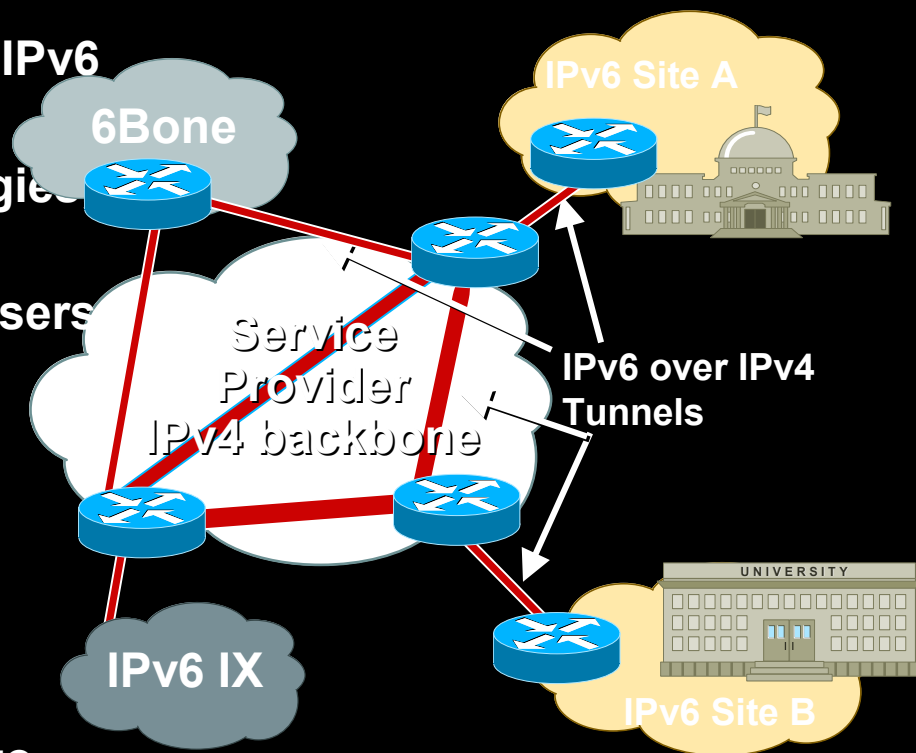
- **Enterprise/Home scenario**

6to4 tunnels between sites, use 6to4 Relay to connect to the IPv6 Internet

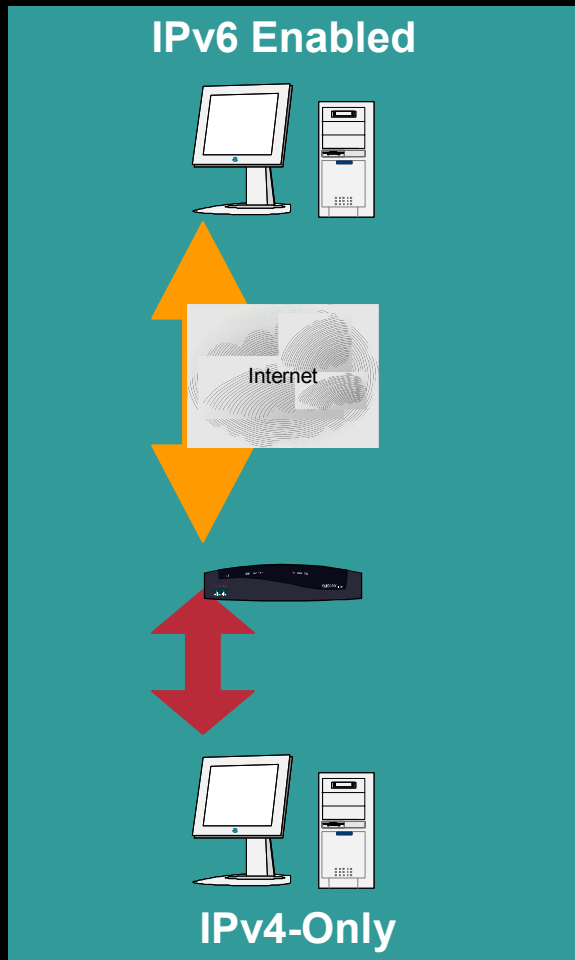
Configured tunnels between sites or to 6Bone users

ISATAP tunnels or Native IPv6 on a Campus

Use the most appropriate



# Tools – Translation

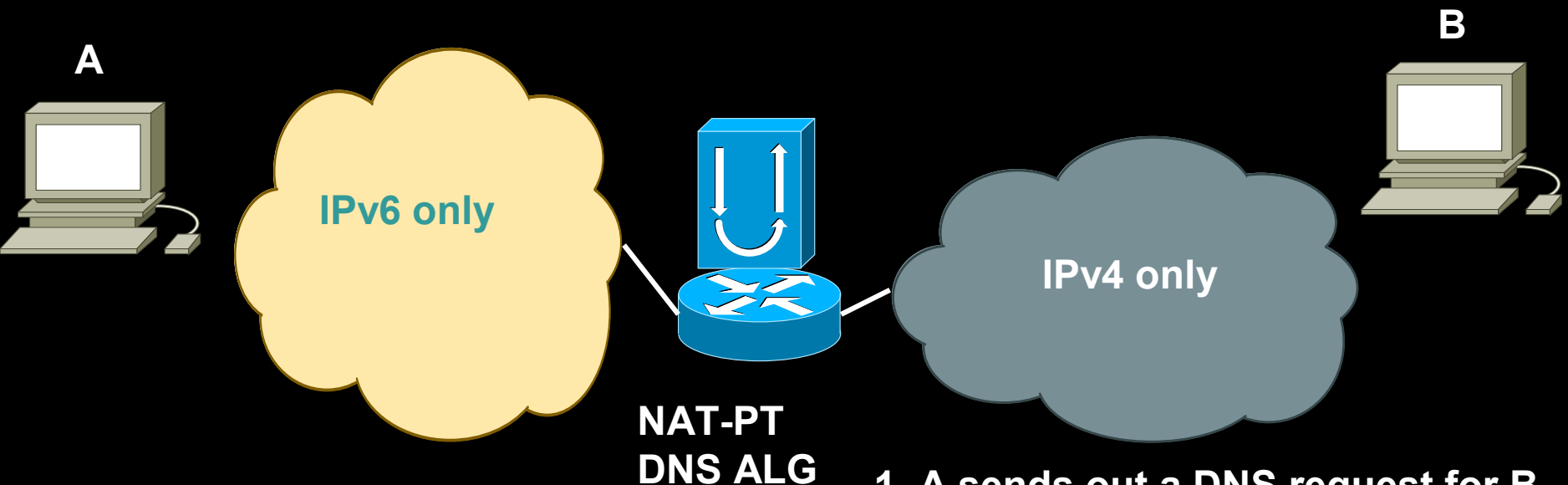


- **Tool of last resort**
- **Allows for the case where some components are IPv6-only while others are IPv4-only**
- **Pay attention to scaling properties**
- **Same application issues as IPv4/IPv4 translation**

# Stateful Translation Mechanisms

- **NAT-PT**  
Address & protocol translation
- **TRT**  
Transport layer relay
- **Socks**  
Application layer gateway
- **IGMP / MLD proxy**  
Joins opposing groups & maps addresses

# NAT-PT



**NAT-Prefix: ipv4-mapped::/96  
announced by NAT-PT**

- 1. A sends out a DNS request for B**
- 2. NAT-PT intercepts the DNS reply translates from A to AAAA, uses prefix:a.b.c.d as the IPv6 address NAT-PT creates translation slot**
- 3. Communication can begin**

# Stateless Translation Mechanisms

- **SIIT**

**Address & protocol translation**

- **BIS**

**Augmentation between IPv4 stack & device driver**

- **BIA**

**Supports IPv4 apps over IPv6 stack**

- **May prefer to use IPv6-IPv4 protocol translation for:**
  - new kinds of Internet devices (e.g., cell phones, cars, appliances)**
  - benefits of shedding IPv4 stack (e.g., serverless autoconfig)**
- **This is a simple extension to NAT techniques, to translate header format as well as addresses**
  - IPv6 nodes behind a translator get full IPv6 functionality when talking to other IPv6 nodes located anywhere**
  - they get the normal (i.e., degraded) NAT functionality when talking to IPv4 devices**
  - drawback : minimal gain over IPv4/IPv4 NAT approach**

# Native IPv6-only Infrastructure?

- **Application focus**

Is the IPv6 traffic important enough?

Worldwide DNS Root not yet reachable via IPv6

- **Requires**

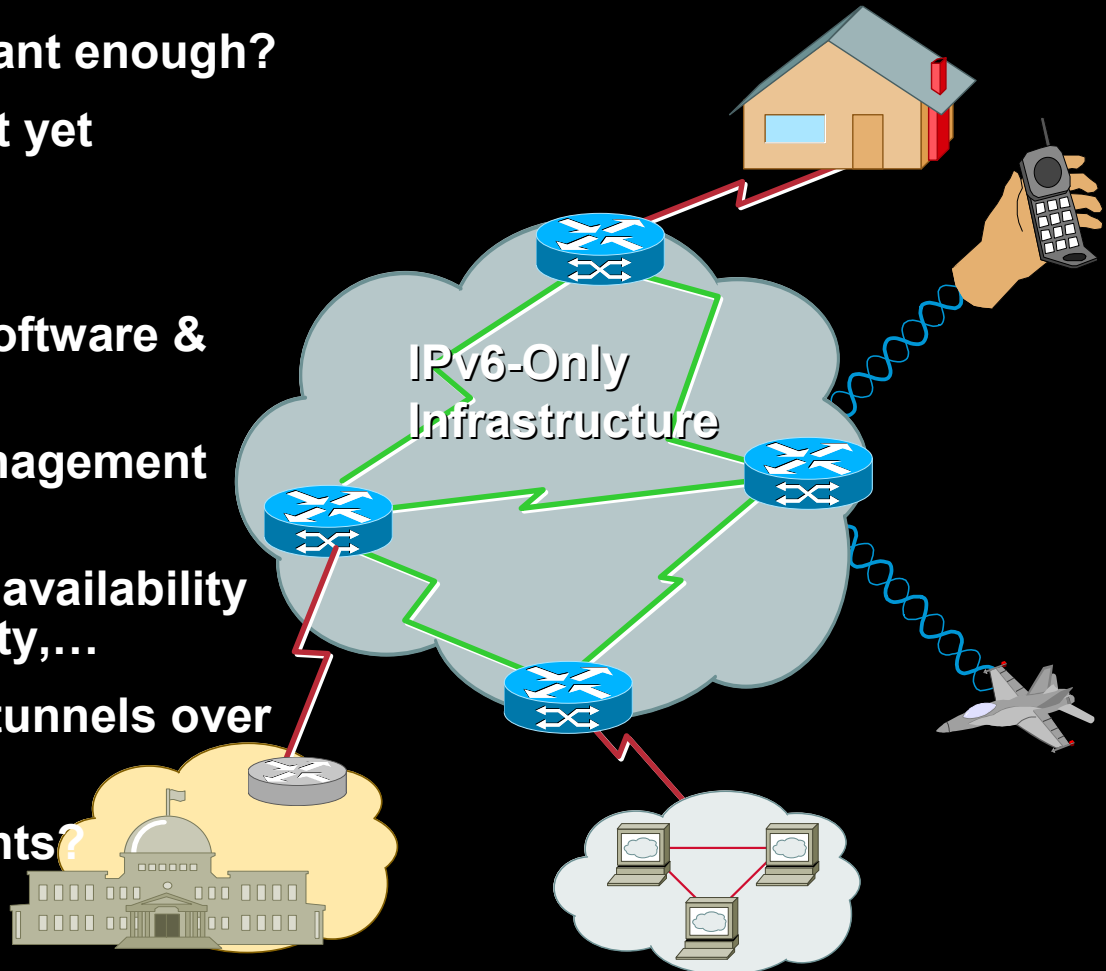
Full Network upgrade (software & potentially hardware)

Native IPv6 Network Management Solutions

Enhanced IPv6 services availability  
Multicast, QoS, security,...

Transport IPv4 through tunnels over IPv6

IPv4 traffic requirements?





# Native IPv6 over Dedicated Data Links

- **Native IPv6 links over dedicated infrastructures**  
ATM PVC, dWDM Lambda, Frame Relay PVC, Serial, Sonet/SDH, Ethernet
- **No impact on existing IPv4 infrastructures**  
Only upgrade the appropriate network paths  
IPv4 traffic (and revenues) can be separated from IPv6
- **Network Management done through IPv4 for now**

- **DNS**

- DNS-ALG in NAT-PT distorts perception**

- Referral chain consistency with resolver**

- Remember glue & reverse records for IPv6**

- **SMTP**

- Create MX records for both IPv4 & IPv6 on dual-stack**

- DNS failure on AAAA may cause mail requeue**

- **DHCP**

- Spec about finished; products will follow**

- Prefix allocation current driver**

# Summary

- **Transition will not be a quick process**
- **Tool set goal : minimize interdependence**
- **Dual-stack & Tunneling before Translation**  
most difficult cases caused by IPv6-only
- **Recognize environment characteristics**
- **Applications will drive deployments**

# Questions?





# For More Information

Cisco.com

- <http://www.ipv6forum.com>
- <http://www.ipv6.org>
- <http://www.cisco.com/ipv6>
- <http://www.microsoft.com/ipv6>
- <http://www.6bone.net>

# For More Information

- <http://www.ietf.org/html.charters/ipngwg-charter.html>
- <http://www.ietf.org/html.charters/v6ops-charter.html>
- <http://playground.sun.com/ipv6/>
- <http://www.6bone.net/ngtrans/>



# For More Information

Cisco.com

- **BGP4+ References**

  - RFC2858 Multiprotocol extension to BGP**

  - RFC2545 BGP MP for IPv6**

  - RFC2842 Capability negotiation**

- **RIPng RFC2080**

# Other Sources of Information

Cisco.com

- **Books**

**IPv6, The New Internet Protocol  
by Christian Huitema (Prentice Hall)**

**Internetworking IPv6 with Cisco Routers  
by Silvano Gai (McGraw-Hill)**

**and many more... (14 hits at Amazon.com)**