IPv6 (Internet Protocol version 6)

APNIC meeting, 3 September 2002

Internet Initiative Japan, Inc. / KAME Project
Keiichi SHIMA <keiichi@iij.ad.jp>

Contents

Why do we use IPv6?

IPv6 Addresses
Link-layer address resolution
Auto-configuration mechanism
DNS
Transision mechanisms
Deployment status
Recent event report

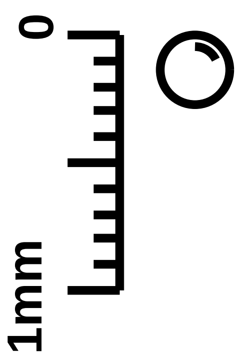
Why do we use IPv6?

- Because IPv6 is better than IPv4
  - Almost infinite address space
    - Everything can have its own address
    - No restriction to allocate addresses any more
  - Easy to use
    - Address auto-configuration
    - Default route discovery
  - Restore the end-to-end communication
  - Enhanced security

IPv6 address space

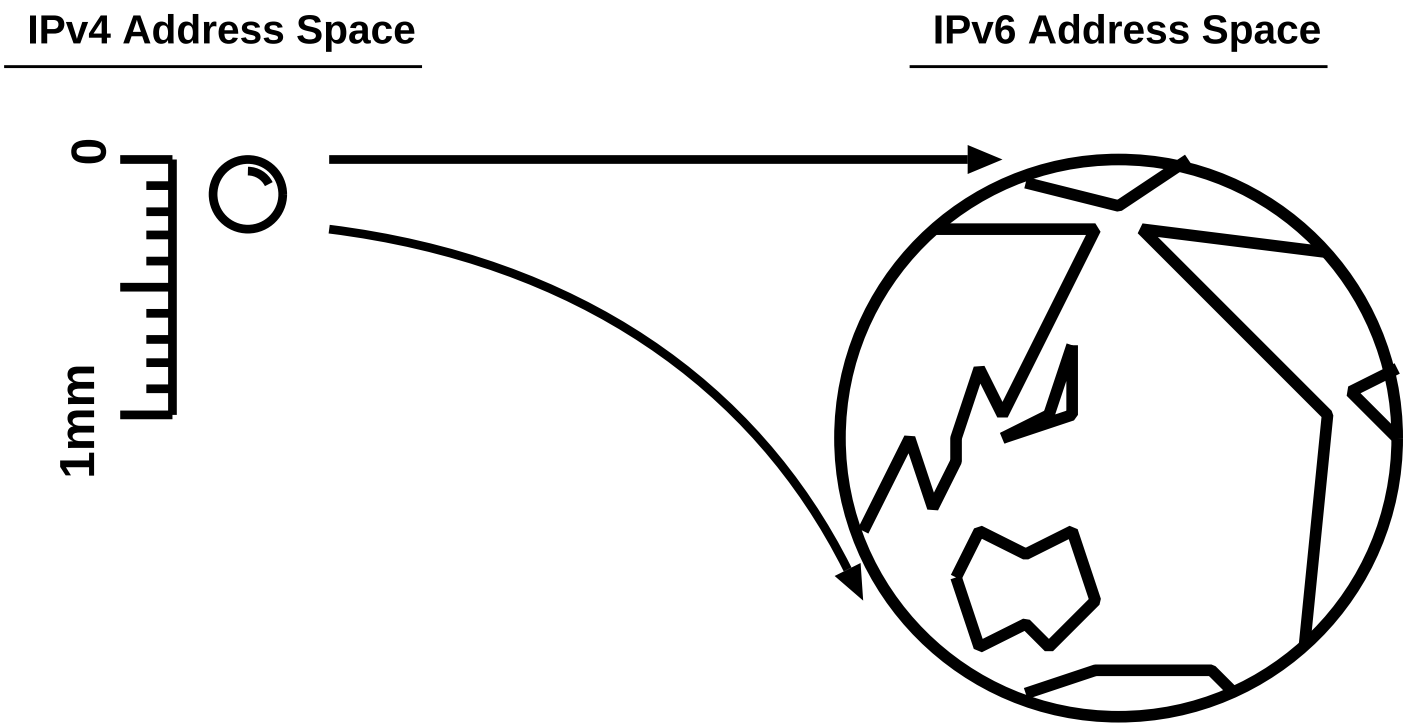
- IPv6 address is 128-bit (= 3.4 x 10^38)
  - IPv4 is 32-bit (= only 4 billions)
- We can assign address to whatever we want
  - Small devices, Electrical appliances, even Thermometers

IPv4 Address Space



IPv6 address space

- IPv6 address is 128-bit (= 3.4 x 10^38)
  - IPv4 is 32-bit (= only 4 billions)
- We can assign address to whatever we want
  - Small devices, Electrical appliances, even Thermometers



Plug-and-Play

- Auto-configuration is mandated
- Just plug a node and we will get addresses
- Defualt routers are automatically installed

End-to-end communication

- Global address for everything makes it possible
- No need for NAT any more
  - NAT does not enhance security
    - Think about HTTP attack, Mail virus, etc..
  - NAT breaks end-to-end communication
  - NAT breaks end-to-end security
- Encourage development of new applications
  - Remember the old Internet where we have had various protocols and various applications on the net

Enhanced security

- IPsec is optional in IPv4

- IPsec is mandatory for all IPv6 nodes
- Security features of IPv6
  - Protect from data forgery
  - Protect from wiretapping
  - Easy to make VPN connections

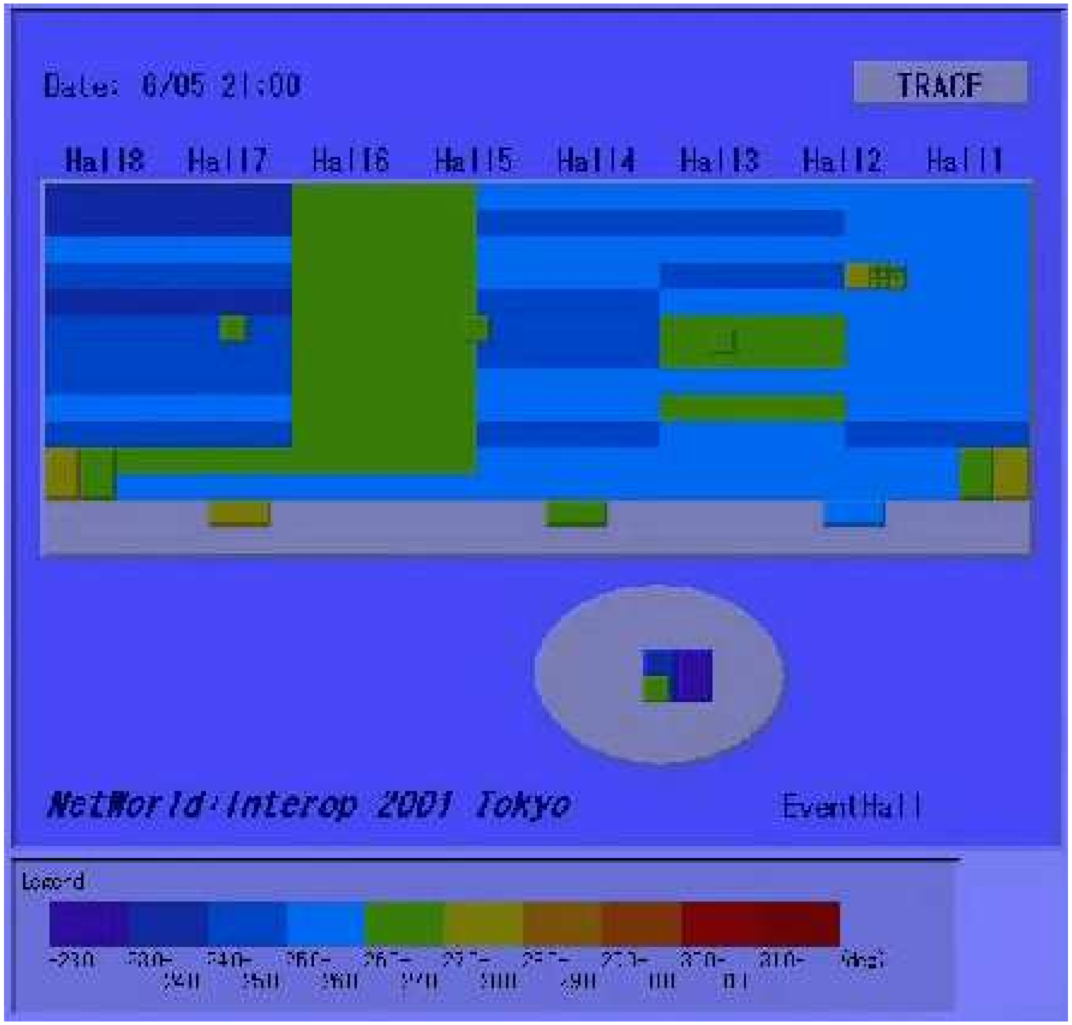
What can we do with IPv6? (1)

- Put addresses to everything!
- At N+I 2001 Tokyo, we put an address to a thermometer
- Hotnode

- The information that one hotnode creates is little, but...

What can we do with IPv6? (1)

- 100 hotnodes made a temperature map

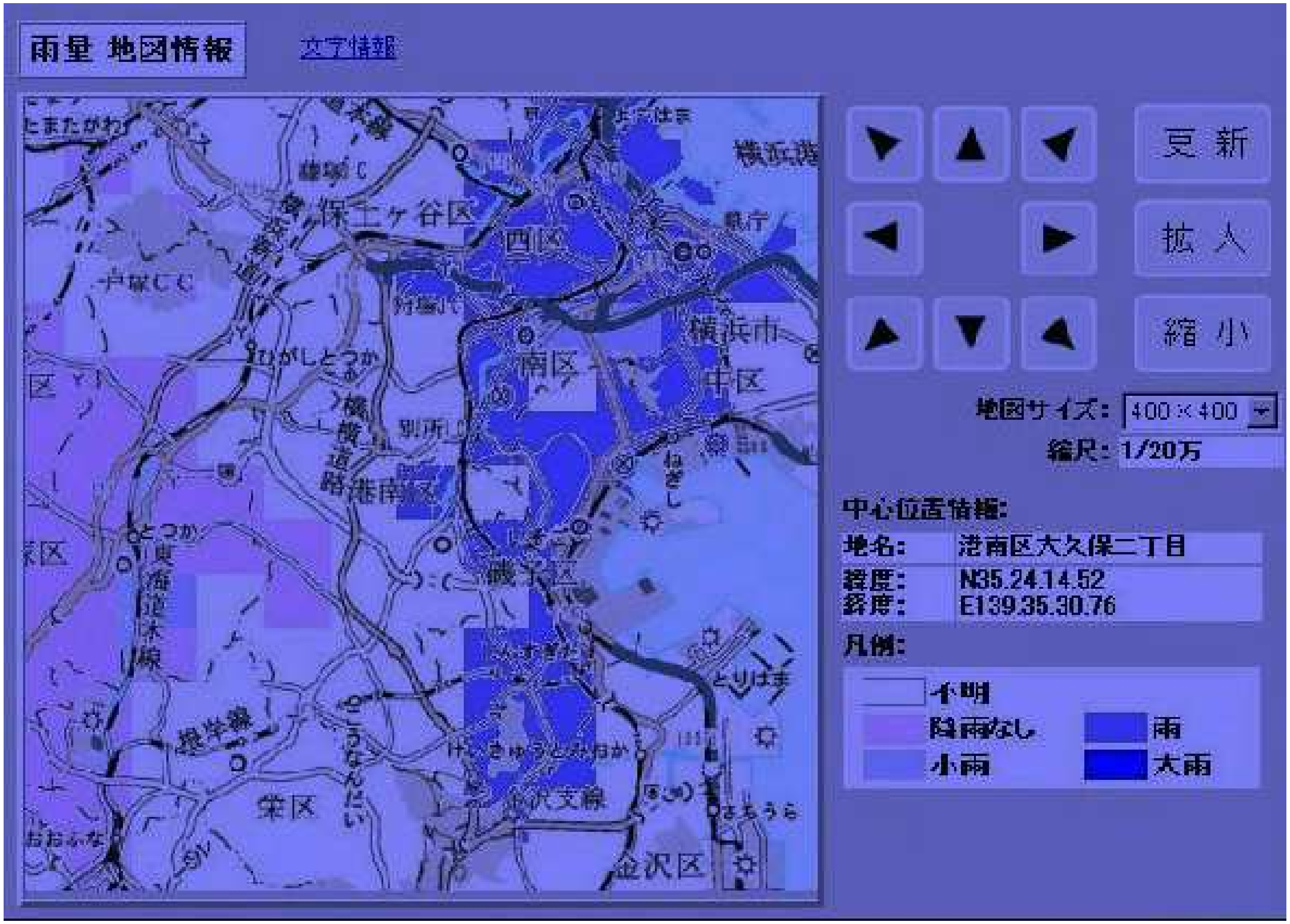


What can we do with IPv6? (2)

- Put addresses to everything!
- Internet ITS Project (2001.2 - 2002.5)
  - http://www.internetits.org/
- We put addresses to hundreds of cars
  - In Nagoya city, 15 hundreds of taxies are addressed
  - In Yokohama city, 70 cars are addressed
- Each sensors has an address
  - Wipers
  - Speed meters

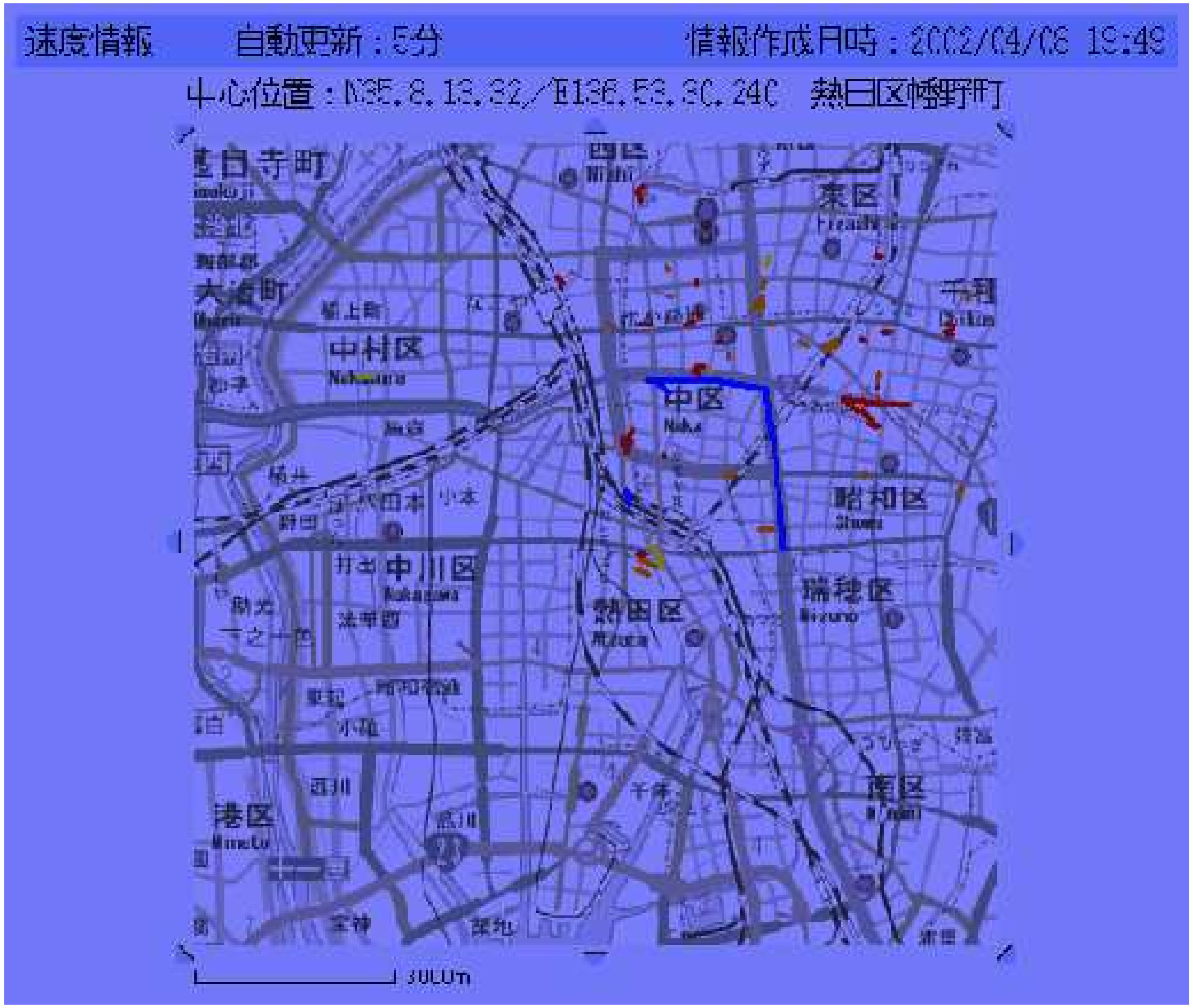
What can we do with IPv6? (2)

- Rain map



What can we do with IPv6? (2)

- Traffic map



Why do we use IPv6?

IPv6 Addresses

Link-layer address resolution

Auto-configuration mechanism

DNS

Transision mechanisms

Deployment status

Recent event report

IPv6 address types

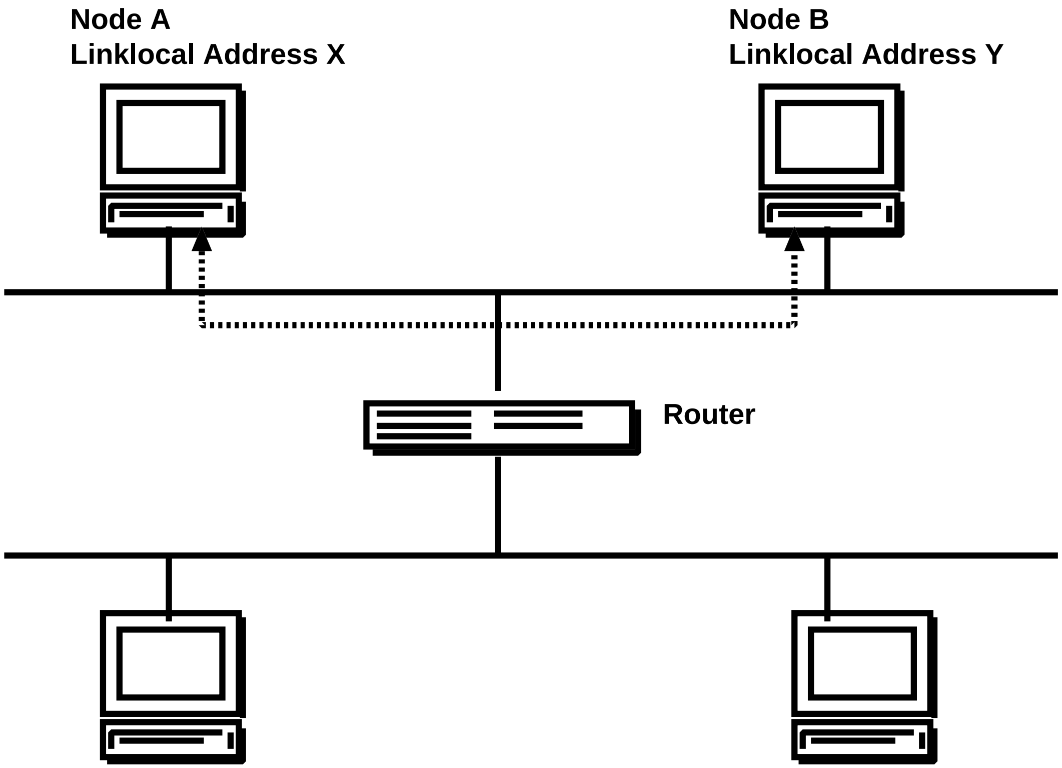
- Unicast address
  - Represents one interface
- Multicast address
  - Represents a set of interfaces those have joined to this multicast address
- Anycast address
  - Represents a nearest interfaces which has this address
  - Anycast address format is same as unicast address

Unicast address

- Basically same as IPv4 unicast address
- IPv6 addresses have "SCOPE"
  - Each scope has a special address block
  - Easily distinguishable from its address form
- Link-local address
  - Unique only in a single link
  - Used by link-layer address resolution, default router discovery
- Site-local address
  - Unique only in a single site
  - Not well researched
- Global address
  - Globally unique

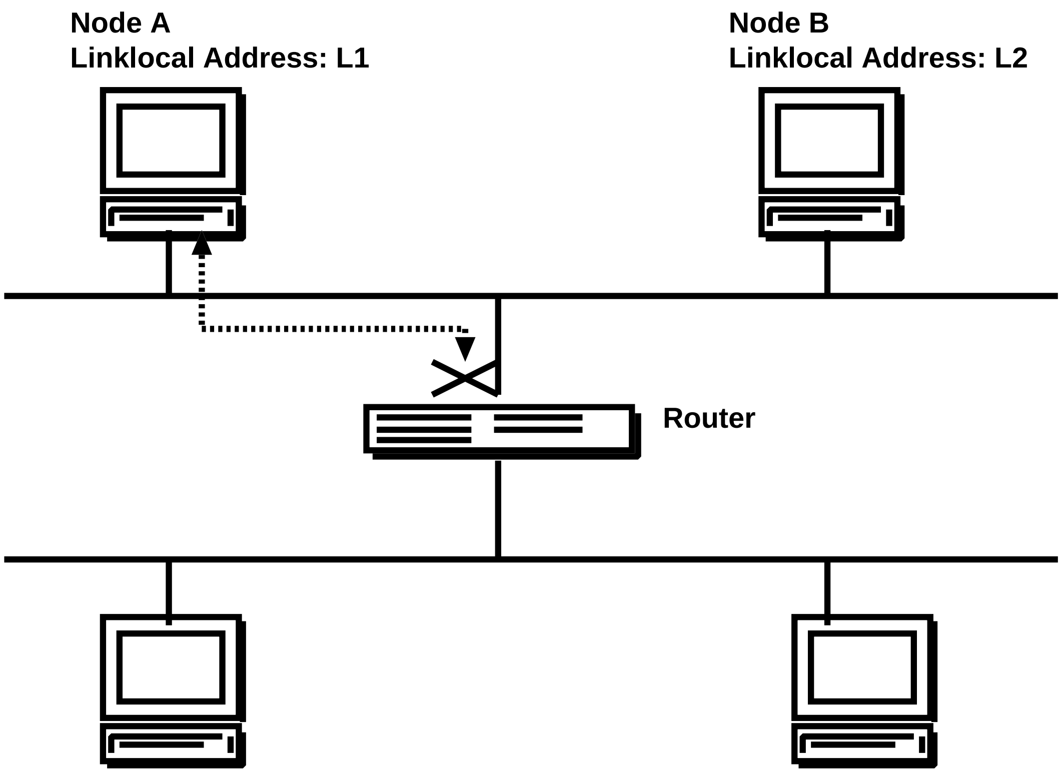
Link-local address

- Unique only in a single link



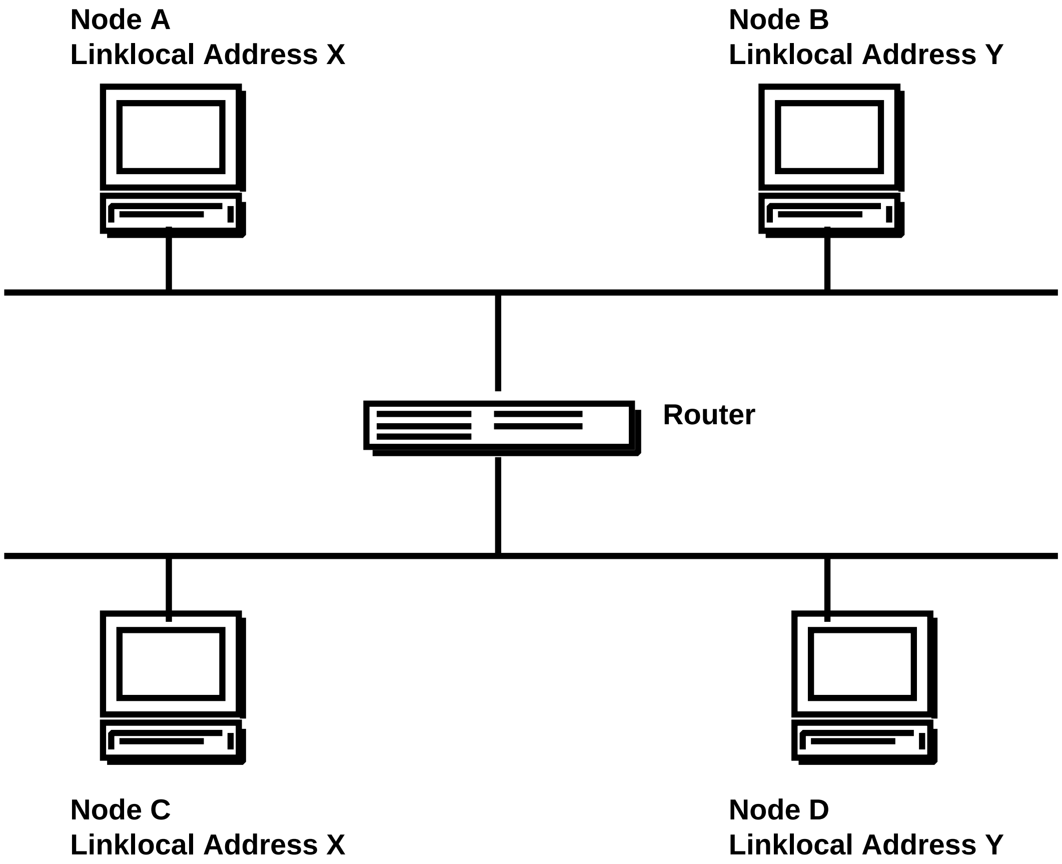
Link-local address

- Unique only in a single link
- Can’t be forwarded to another link



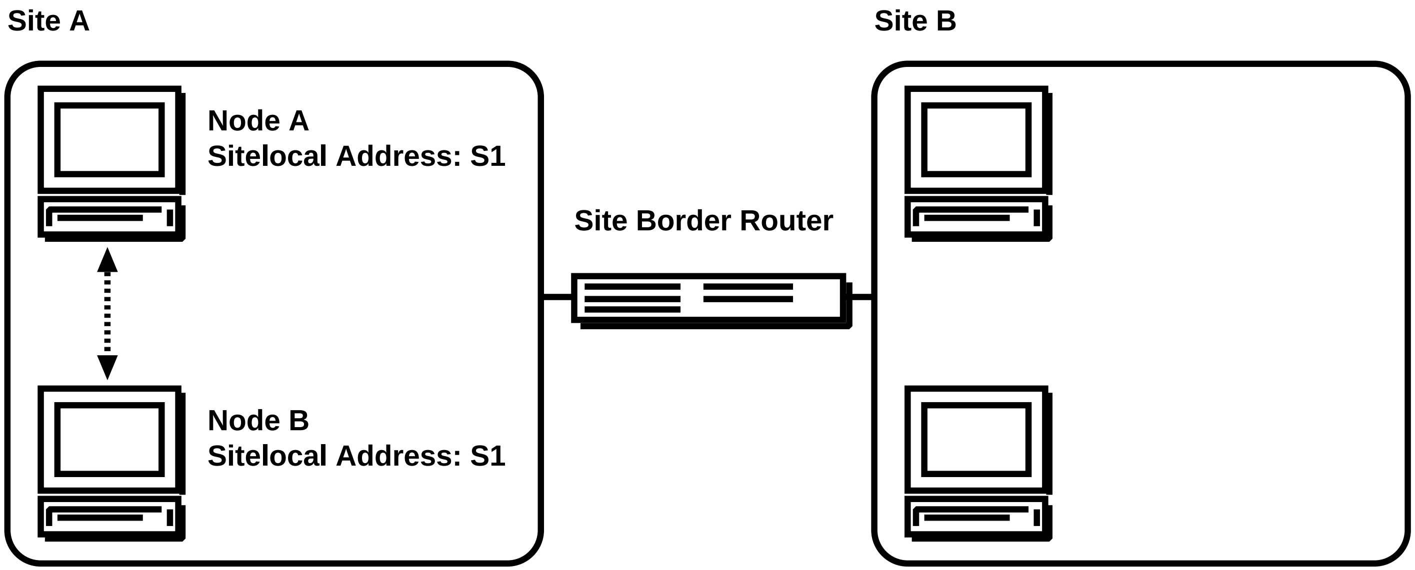
Link-local address

- Unique only in a single link
- Can’t be forwarded to another link
- Same addresses may exist on other links



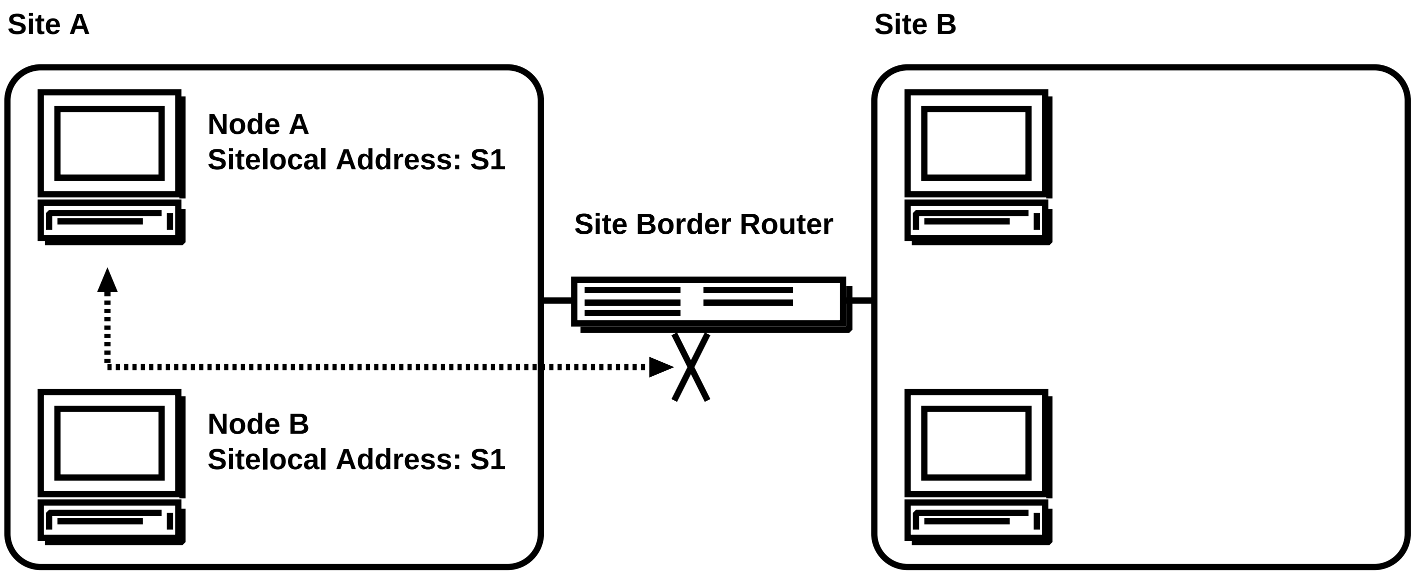
Site-local address

- Unique on a single site



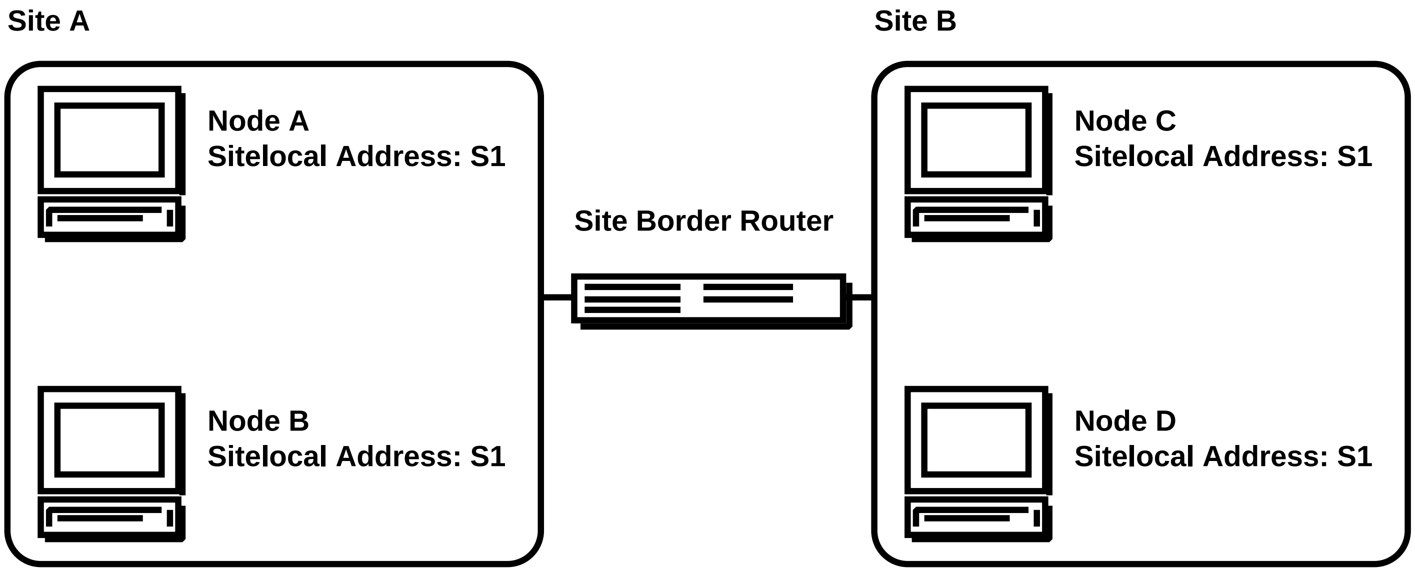
Site-local address

- Unique on a single site
- Can't be forwarded to another site



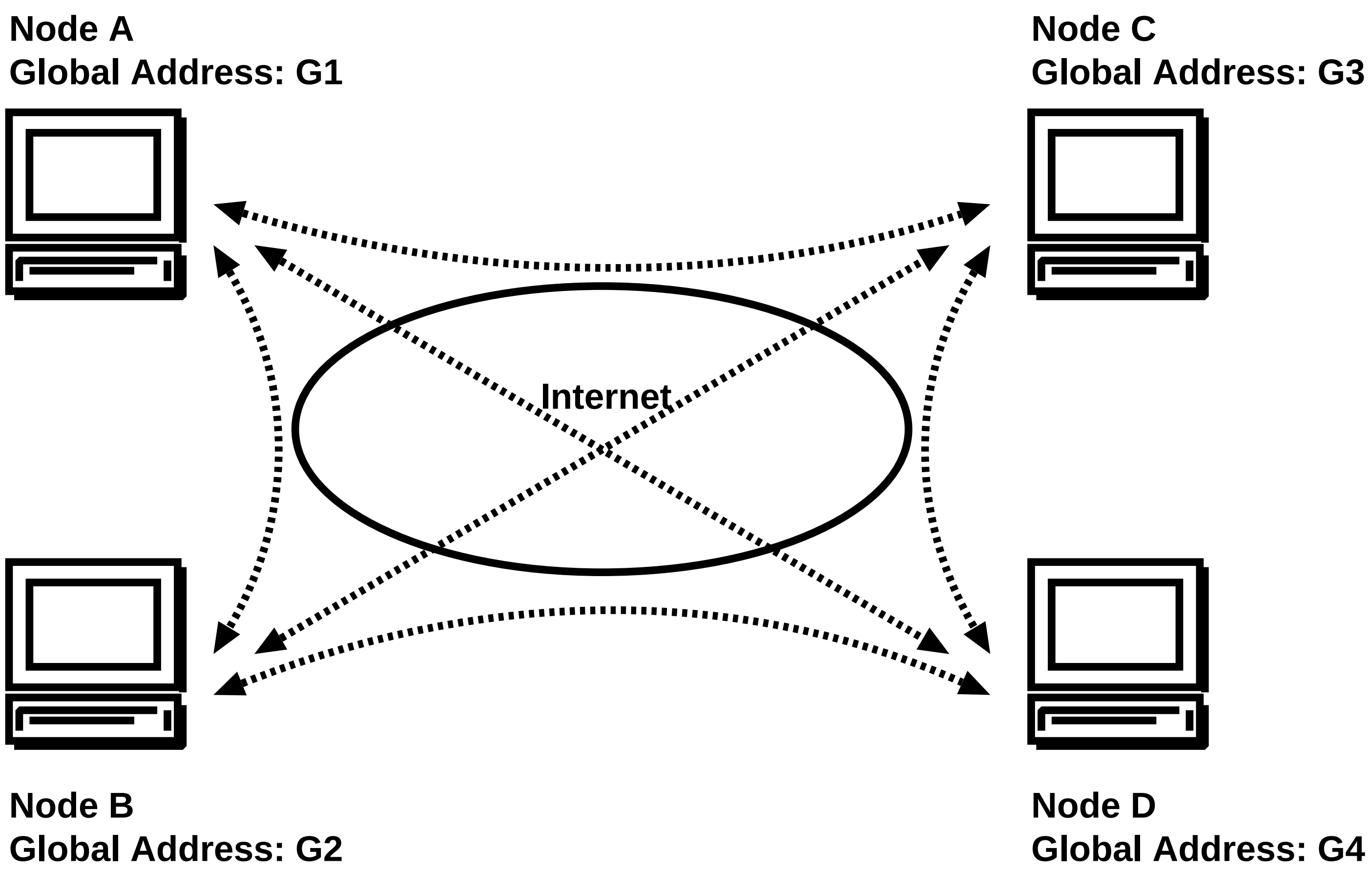
Site-local address

- Unique on a single site
- Can’t be forwarded to another site
- Same addresses may exist on other sites



Global address

- Unique entirely



Multicast address

- Basically same as IPv4 multicast address
- Multicast addresses also have "SCOPE"
  - Interface-local
  - Link-local
  - Subnet-local
  - Admin-local
  - Site-local
  - Organization-local
  - Global
- Scope values are embedded to the address format
- Typical usage of multicast addresses
  - Link-local scope for link-layer address resolution, default router discovery
  - Global scope for video conferences-like applications

Broadcast address ?

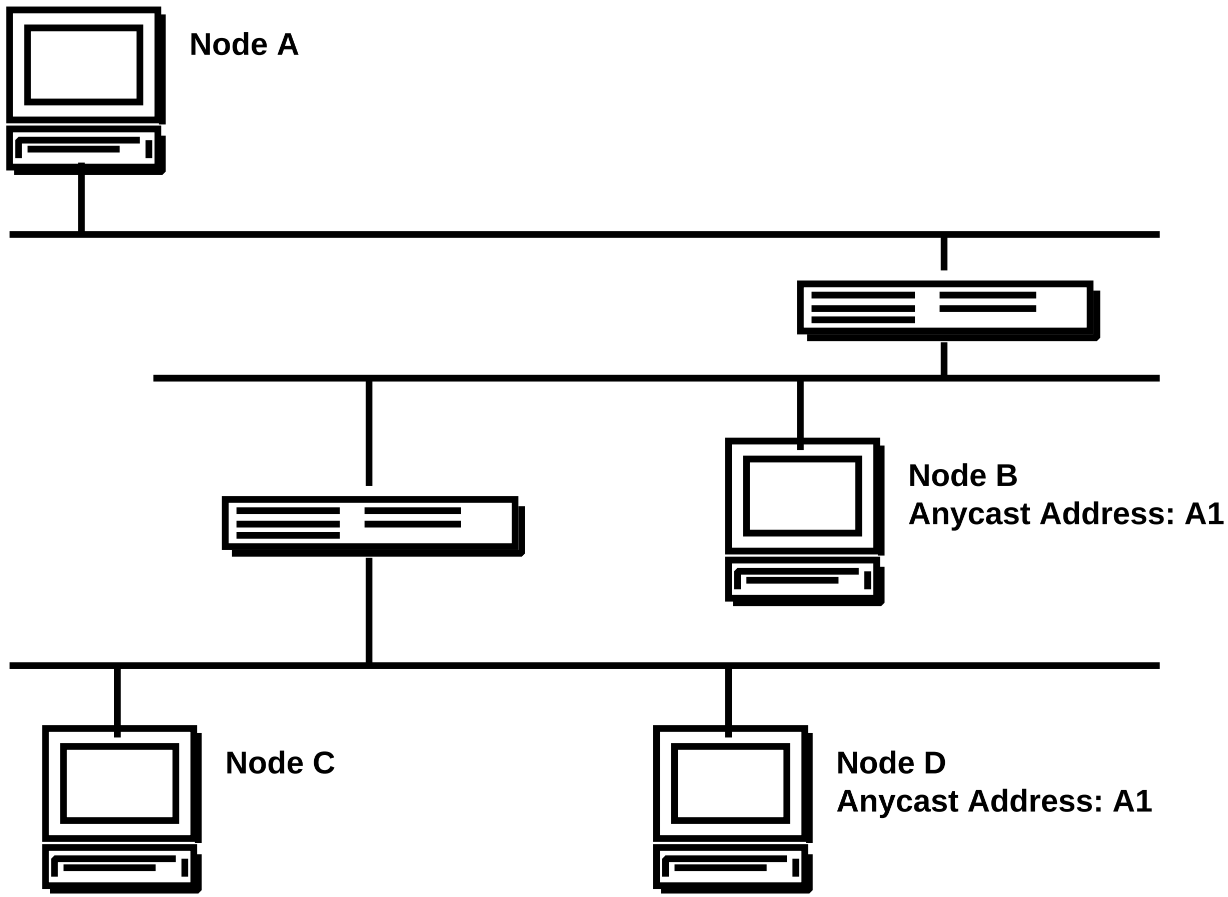
- There is no broadcast address in IPv6
- Use multicast address instead
- Special multicast addresses are defined
  - All-node multicast address
  - All-router multicast address
- Some protocols have its own multicast address
  - Datalink-layer address resolution
  - OSPF
  - RIP
  - PIM
  - DHCP
  - etc

Anycast address

- Represents a nearest interface in the sense of routing
- Address format is same as that of unicast
- What's for?
  - Service discovery like a DNS server discovery
- Need more study for using anycast addresses

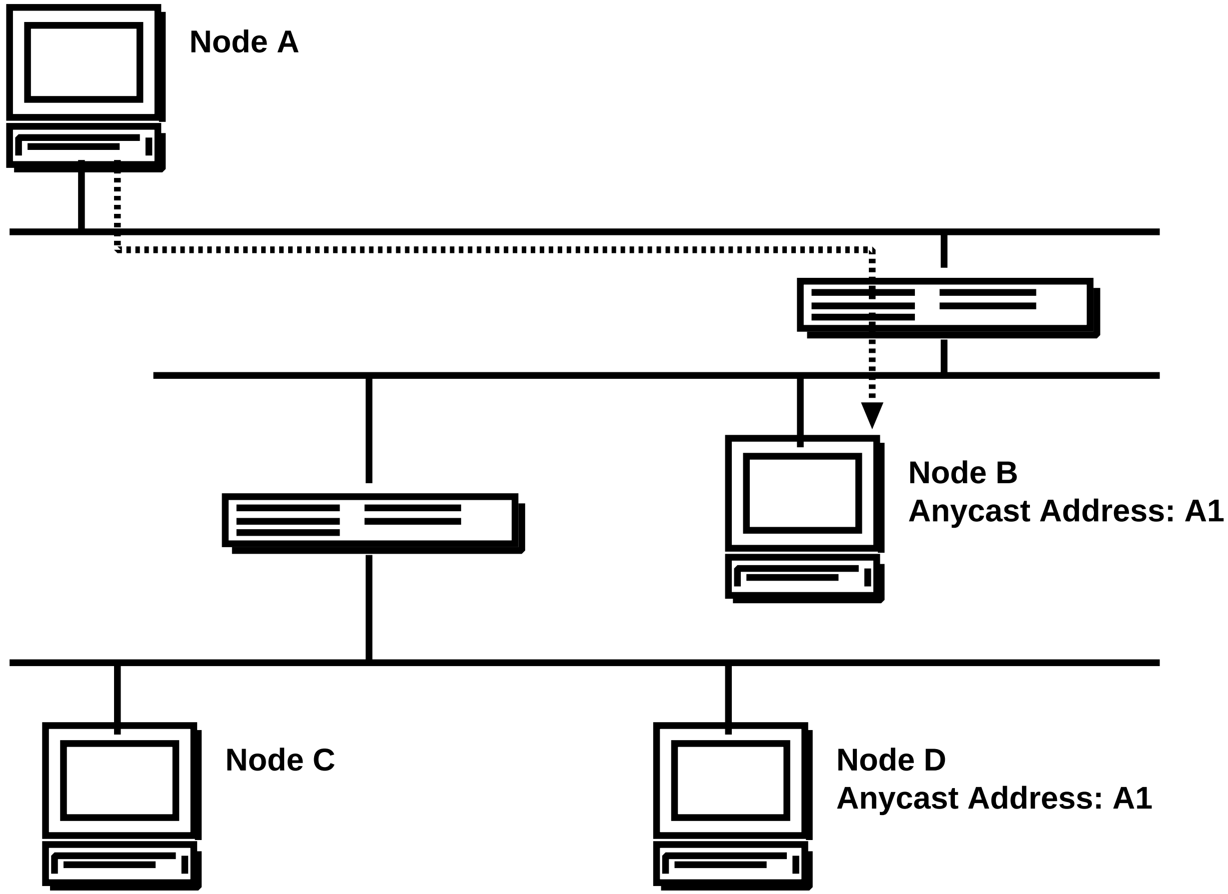
Anycast address

- Many nodes have a same anycast address



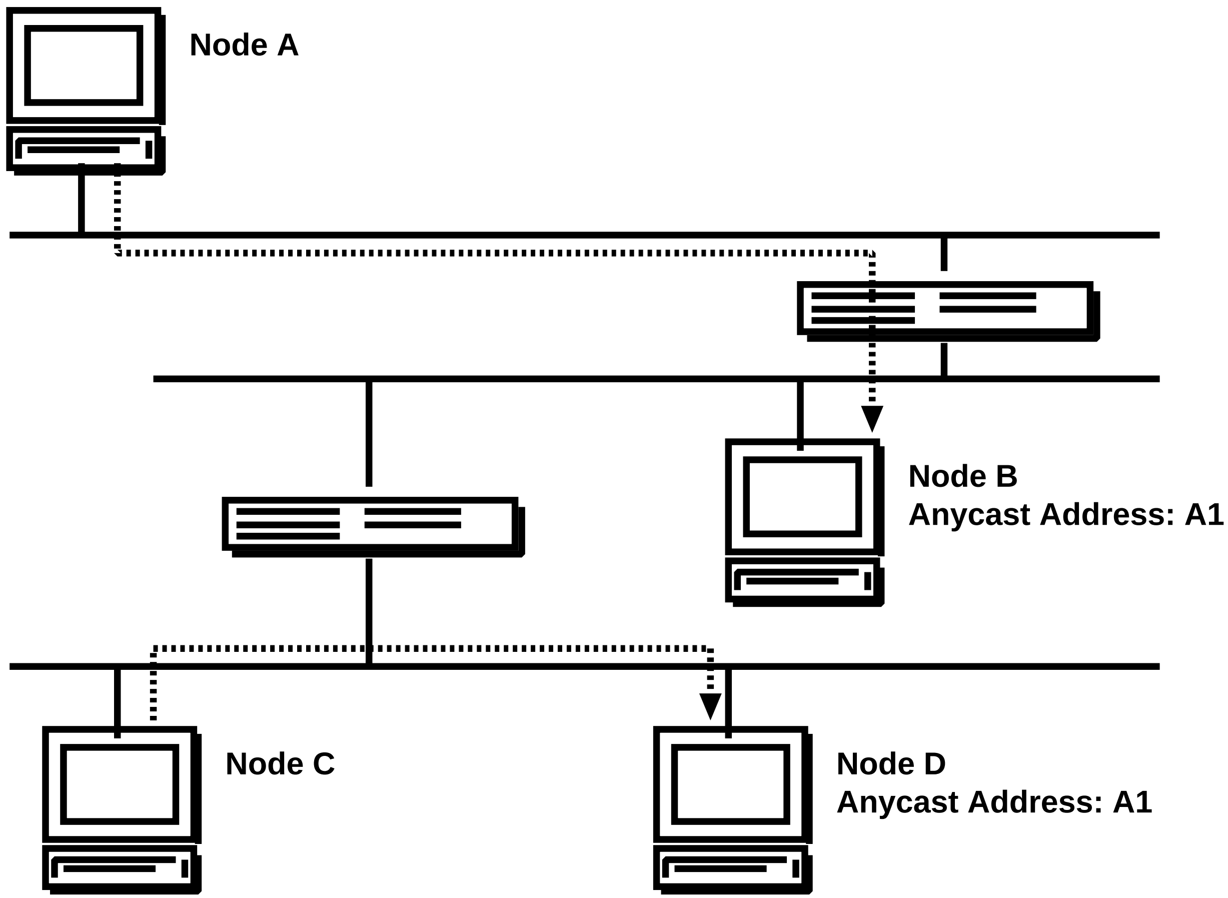
Anycast address

- Many nodes have a same anycast address
- Packets are sent to the nearest node



Anycast address

- Many nodes have a same anycast address
- Packets are sent to the nearest node



Text representation of addresses

- x:x:x:x:x:x:x:x
  - Where ’x’s are the hex values of 16-bit
  - Separated by colons(:)
- Example
  - fe80:0000:0000:0000:0203:47ff:fe3d:02bd
- Leading 0 can be ommited
  - fe80:0000:0000:0000:0203:47ff:fe3d:02bd
  - fe80:0:0:0:203:47ff:fe3d:2bd
- 0 can be compressed, but only once
  - fe80:0:0:0:203:47ff:fe3d:2bd
  - fe80::203:47ff:fe3d:2bd
- Specify prefix length using slash
  - fe80::203:47ff:fe3d:2bd/64

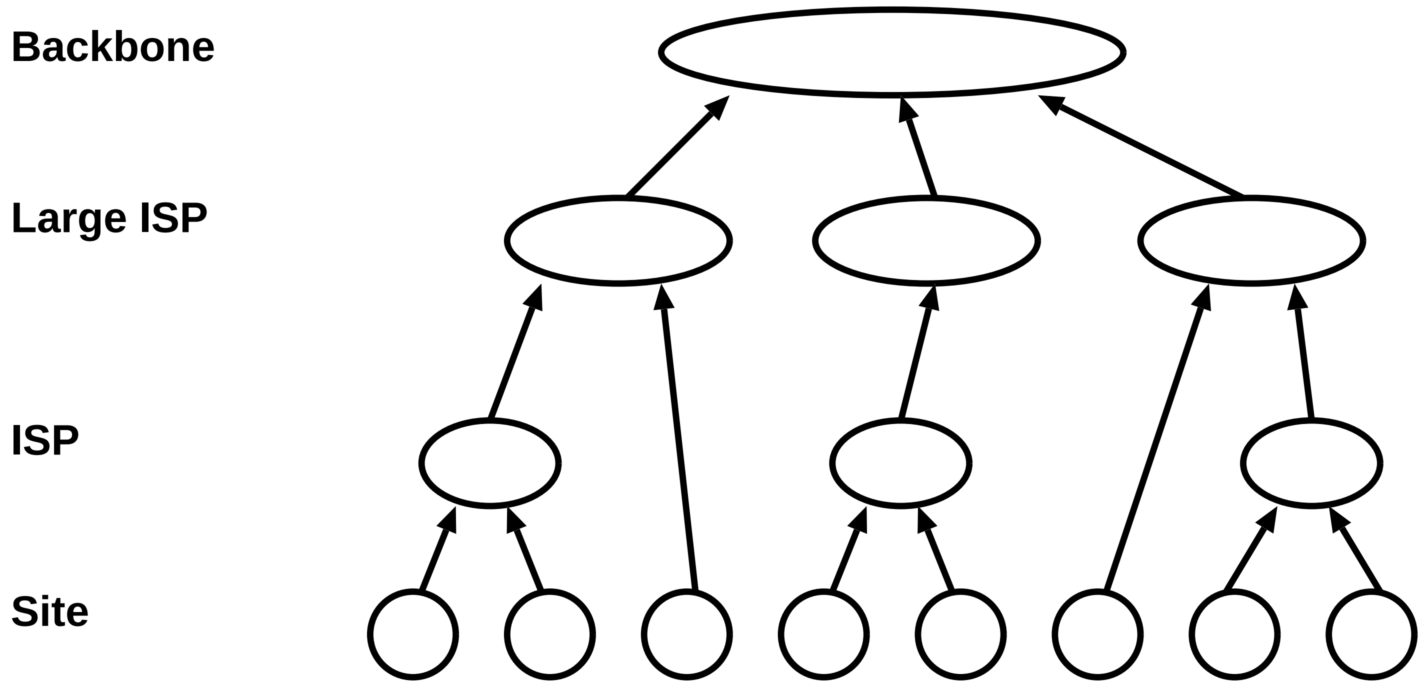
Address blocks

- The high-order bits represents address blocks

Unicast	Global	0000000000 ------------ 0000000000	
		0010000000 ------------ 0000000000	2000::/3
		0100000000 ------------ 0000000000	
	Link-local	1111111010 ------------ 0000000000	fe80::/10
	Site-local	1111111011 ------------ 0000000000	fec0::/10
Multicast		1111111100 ------------ 0000000000	ff00::/8
		1111111111 ------------ 1111111111	

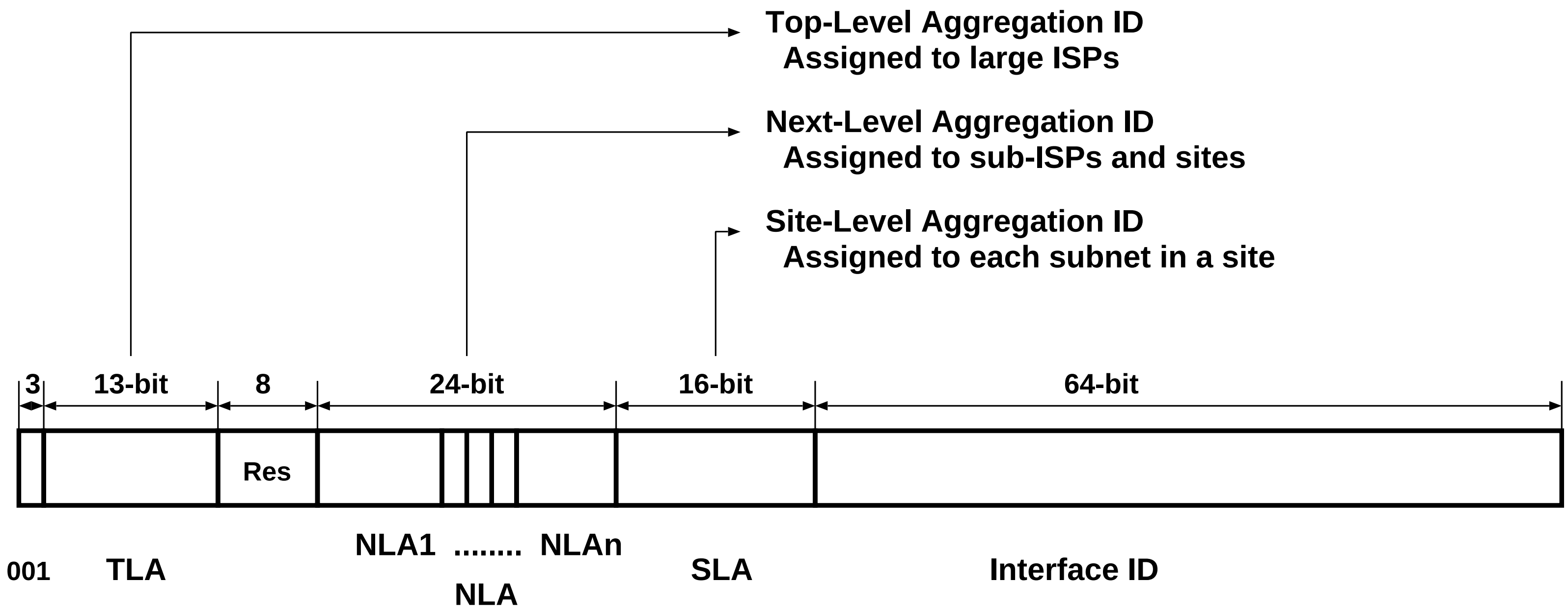
Aggregatable addressing architecture

- Hierarchical address allocation
- Aggregate routing information
  - Manages only downstream ISPs/Sites' routes



Aggregatable addressing architecture

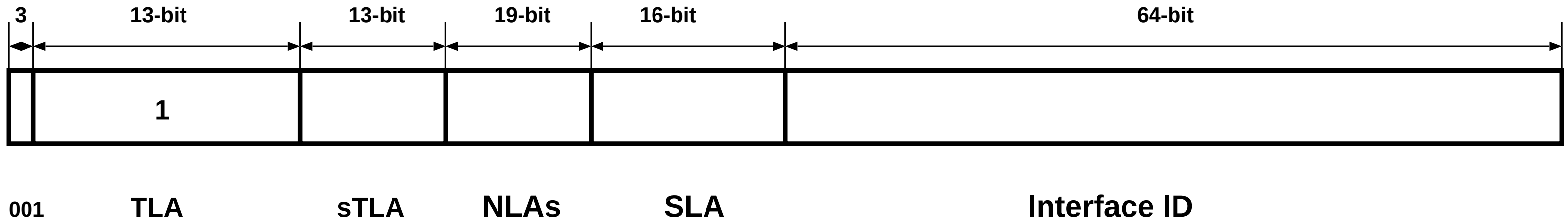
- Aggregate routes in each level



- The backbone only manages routes for TLAs
- A large ISP assigned TLA ID only manages routes fot its NLA1s
- And so on...

Current Status

- We are now in the initial stage
- Using one TLA ID (2001::/16)
- The TLA ID 1 (2001::/16) has 13-bit Sub-TLA (sTLA)
  - A large ISP has a sTLA ID



2001:0200::/29 - 2001:03f8::/29 APNIC
2001:0400::/29 - 2001:05f8::/29 ARIN
2001:0600::/29 - 2001:07f8::/29 RIPE NCC

Address allocation policy

- LIR can get /32 space from RIR
  - LIR...large ISPs
  - RIR...APNIC, ARIN, RIPE
- A large ISP can get a huge space for their customers by default
  - Potentially, 65536 customers
- Current allocation status can be found
  - http://www.ripe.net/cgi-bin/ipv6allocs
- A site will have /48 address space from ISP
  - 65536 subnets with /64 prefix

Why do we use IPv6?

IPv6 Addresses

Link-layer address resolution

Auto-configuration mechanism

DNS

Transision mechanisms

Deployment status

Recent event report

Neighbour Discovery Protocol (NDP)

- ARP (Address Resolution Protocol) for IPv4
- Do not use broadcasting
  - Use multicasting
  - Lightweight than ARP
- NDP is designed as ICMP
  - Datalink independent
- New features
  - Duplicate Address Detection
  - Neighbour Unreachability Detection
- Integrated functions
  - Redirection

How does NDP work?

- A special multicast address
  - Related to a node's IPv6 address
  - All nodes must join to its special multicast address
- Querier sends Neighbour Solicitation (NS) to that special multicast address
- A target node replys by Neighbour Advertisement (NA)
- NA includes a datalink address

Solicited node multicast address

- A special multicast address
- Calculated from node's interface ID
- Interface ID creation (Ethernet)

Ethernet MAC address (48-bit)

`00:03:47:3d:02:bd`

Interface ID (64-bit)

Solicited node multicast address

- A special multicast address
- Calculated from node's interface ID
- Interface ID creation (Ethernet)

Ethernet MAC address (48-bit)

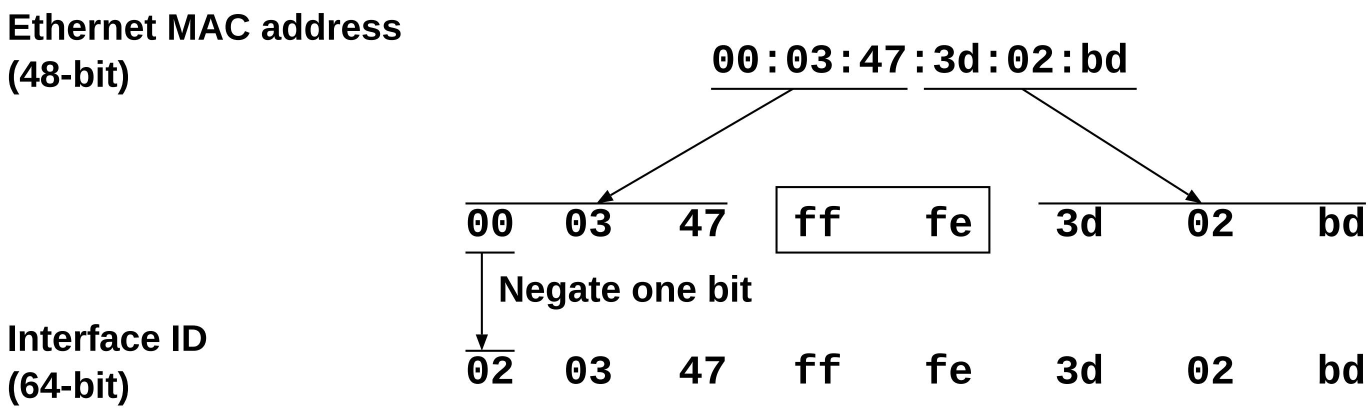
00:03:47:3d:02:bd

00 03 47 ff fe 3d 02 bd

Interface ID (64-bit)

Solicited node multicast address

- A special multicast address
- Calculated from node's interface ID
- Interface ID creation (Ethernet)



Solicited node multicast address

- A special multicast address
- Calculated from node's interface ID
- Interface ID creation (Ethernet)

Ethernet MAC address (48-bit)

00:03:47:3d:02:bd

00 03 47 ff fe 3d 02 bd

Negate one bit

Interface ID (64-bit)

02 03 47 ff fe 3d 02 bd

- Solicited node multicast address calculation

Interface ID

02 03 47 ff fe 3d 02 bd

Solicited node multicast address

- A special multicast address
- Calculated from node's interface ID
- Interface ID creation (Ethernet)

Ethernet MAC address (48-bit)

`00:03:47:3d:02:bd`

`00 03 47 ff fe 3d 02 bd`

Negate one bit

Interface ID (64-bit)

`02 03 47 ff fe 3d 02 bd`

- Solicited node multicast address calculation

Link-local multicast prefix

`ff02::/16`

Interface ID

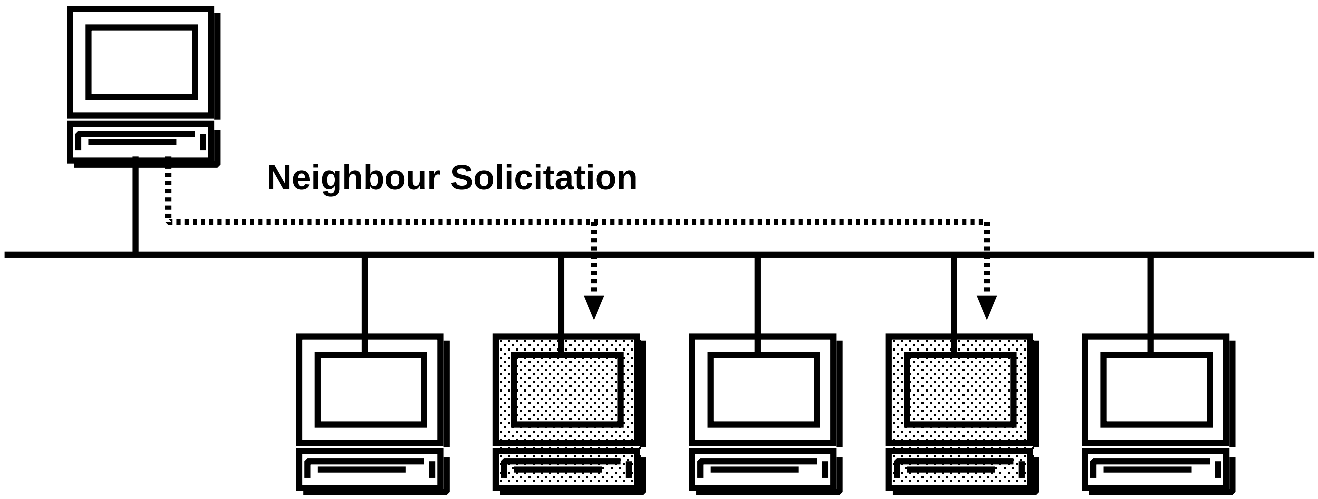
`02 03 47 ff fe 3d 02 bd`

Lower 24-bit

`ff 02 00 00 00 00 00 00 00 00 00 01 ff 3d 02 bd`

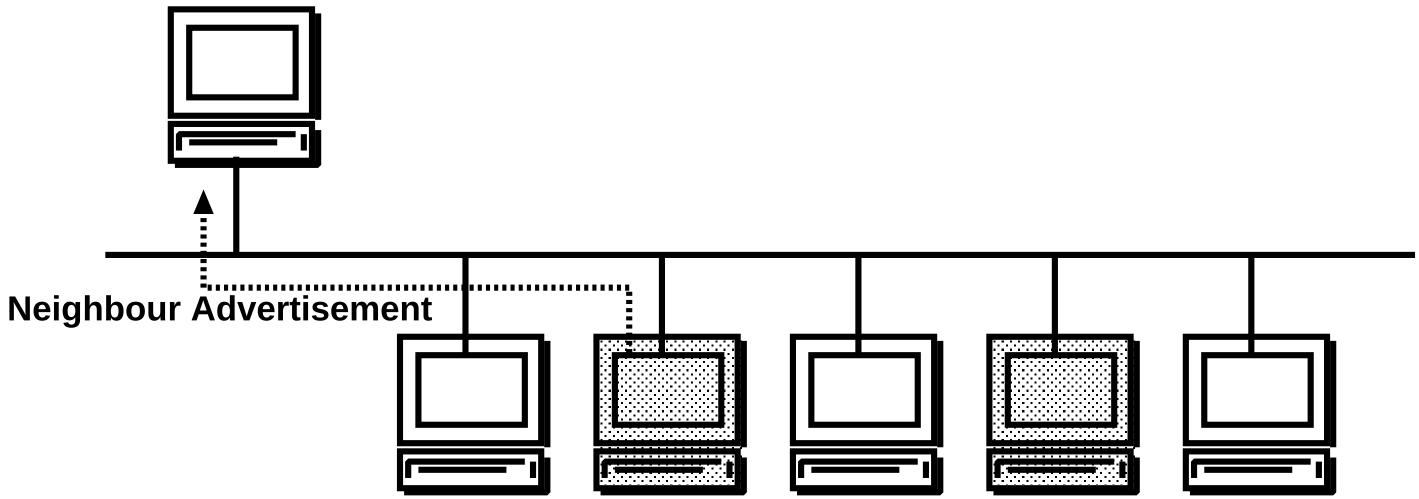
NS/NA transmission

- Solicited node multicast address
  - Represents a set of nodes including a target node
  - Lower 24-bits are the same
- Address resolution request is sent to this address



NS/NA transmission

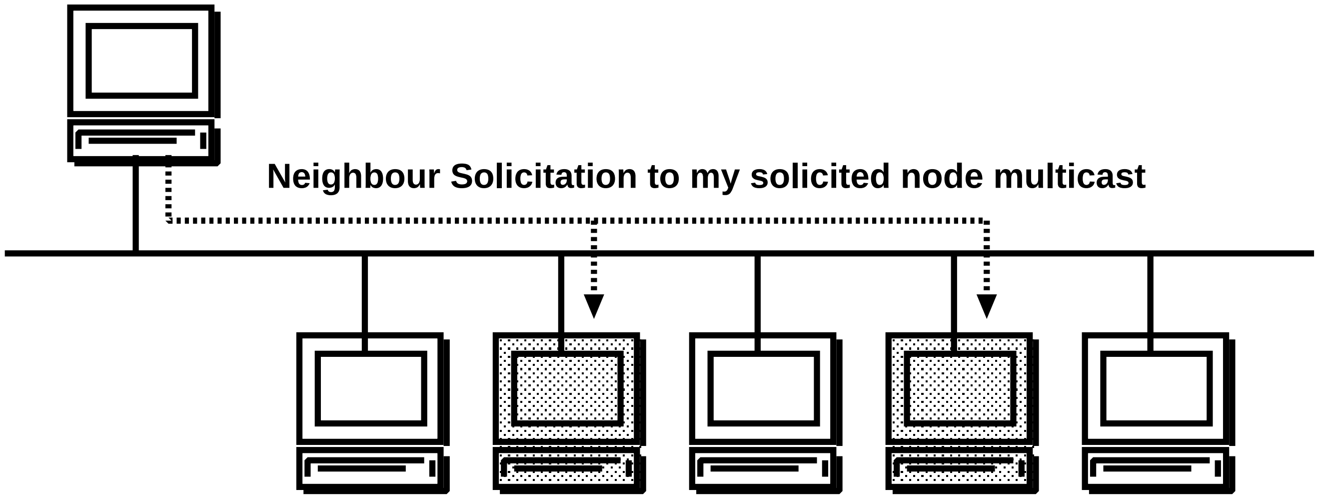
- Solicited node multicast address
  - Represents a set of nodes including a target node
  - Lower 24-bits are the same
- Address resolution request is sent to this address



- In most cases, solicited node multicast address includes only the target node
  - It is rare to have same lower 24-bit address
  - Address resolution is done between only two nodes

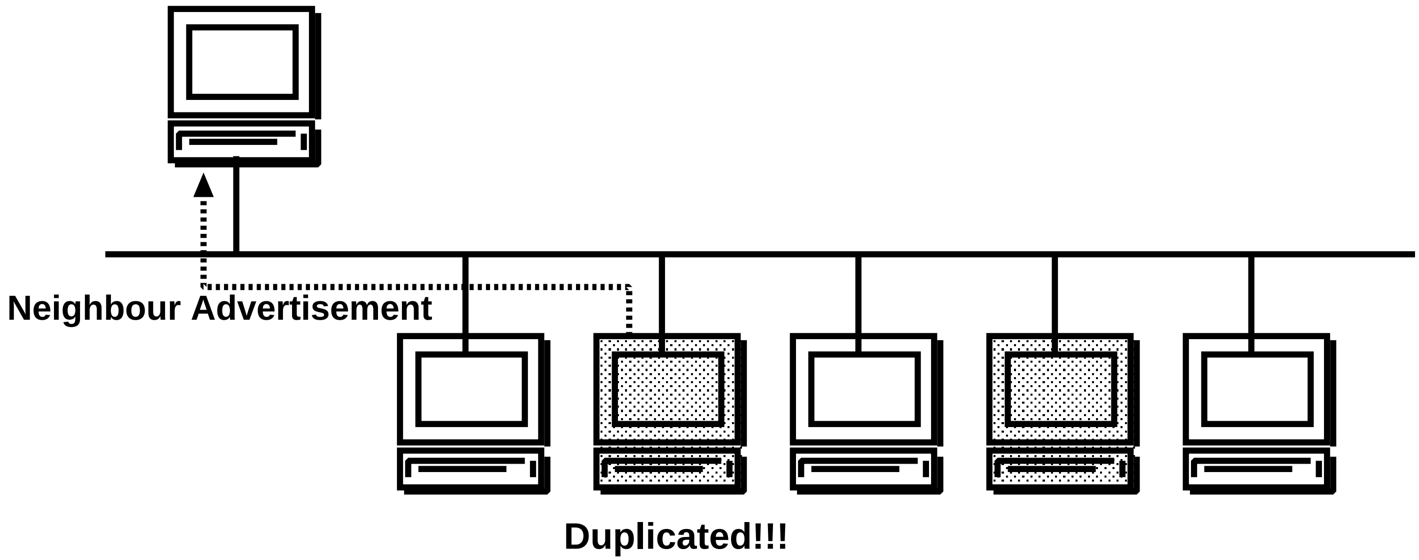
Duplicate address detection (DAD)

- Send NS to MY solicited node multicast address
- No answer will come if no duplication



Duplicate address detection (DAD)

- Try to resolve my IPv6 address
- Send NS to "MY" solicited node multicast address
- No answer will come if no duplication



Neighbour Unreachability Detection (NUD)

- Datalink addresses are cached
  - Expire in a short time (default 30sec)
  - ARP has 20min expiration time, too long
- Probe nodes using NS when expired
  - The cache can be used
  - No additional wait for resolution
- If the node stays, NA will come
- If the node disappers, NA will not come
  - Datalink address cache is removed
- Fast detection of node reachability

Why do we use IPv6?

IPv6 Addresses

Link-layer address resolution

Auto-configuration mechanism

DNS

Transision mechanisms

Deployment status

Recent event report

Why is auto-configuration important?

- IPv6 has a huge address space
  - It is nightmare to manage them by hand
- Many small devices will appear
  - They may not have a console
  - Should be plag-and-play

IPv6 auto-configuration

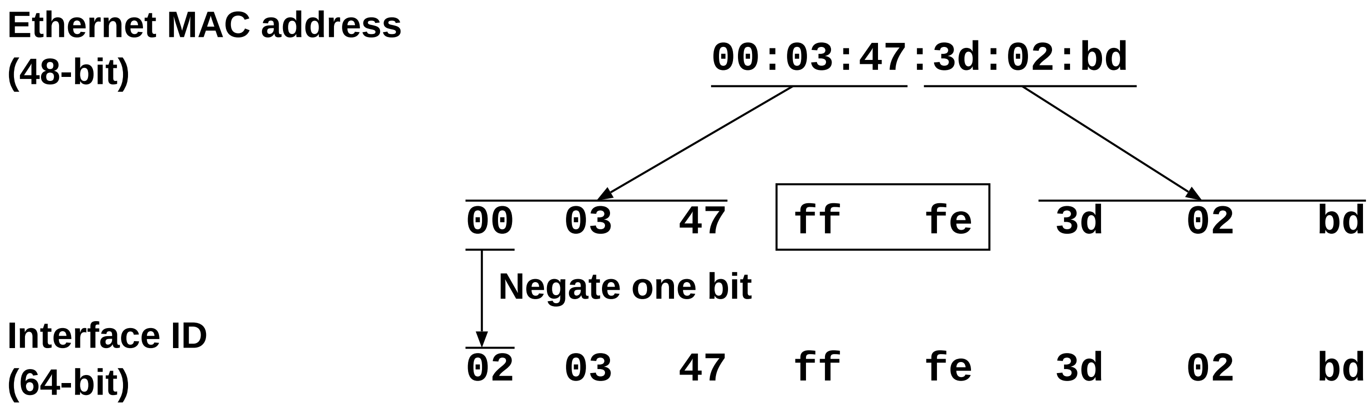
- Host configuration
  - Address auto-configuration
  - Defualt router discovery
- Edge-router configuration
  - Prefix Delegation

Stateless address auto-configuration

- Auto-configuration steps
  - Create interface ID
  - Assign a link-local address
  - Receive prefix information from routers
  - Assign global address(es)
- No need for a central server like DHCP
- Defacto stadard for IPv6 address auto-configuration

Create interface ID

- Interface ID is calculated from MAC address
- No additional information
- Calculation methods are defined by RFC for each datalink
- Example (Ethernet)



Link-local address creation

- Concatinate link-local prefix and interface ID
  - Link-local prefix fe80::/64
  - interface ID is calculated from the MAC address
- Example

Link-local prefix	**Interface ID**
`fe80::/10`	**`02 03 47 ff fe 3d 02 bd`**
↓	↓
`fe 80 00 00 00 00 00 00`	**`02 03 47 ff fe 3d 02 bd`**

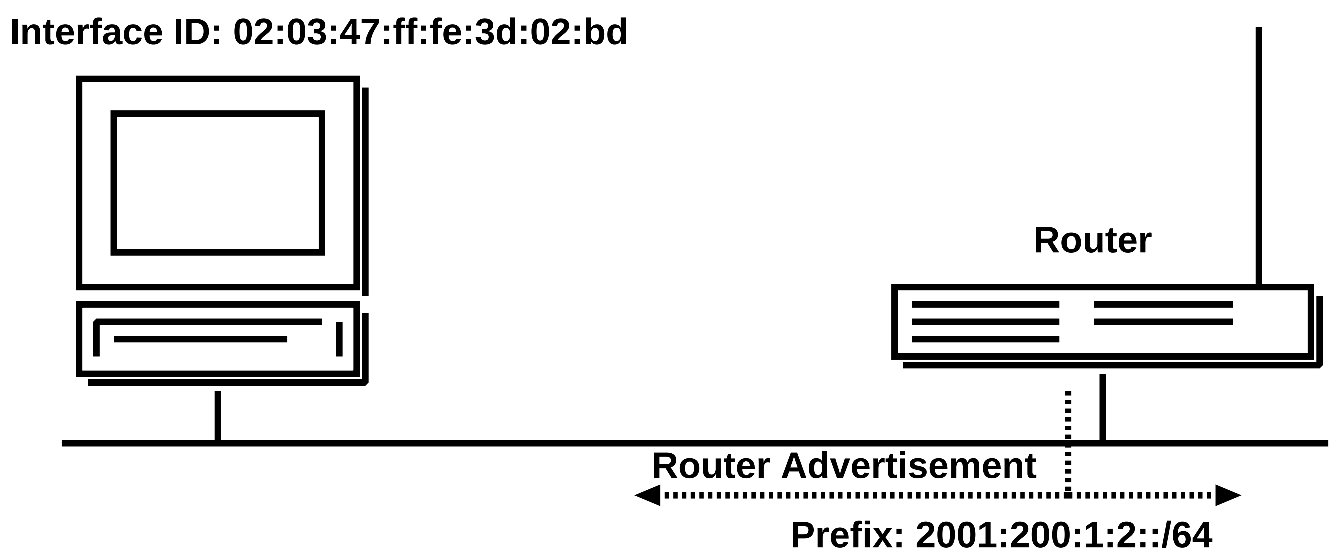
- With link-local addresses, we can communicate other nodes on the same link

Receive prefix information

- Router advertisement (RA)
  - Multicasted periodically from routers to all nodes connected to the same link
  - Routers use link-local addresses to communicate with nodes
- RA includes link information
  - Global/Site-local prefixes
  - MTU size, etc
- Nodes receive prefix information and create global/site-local addresses

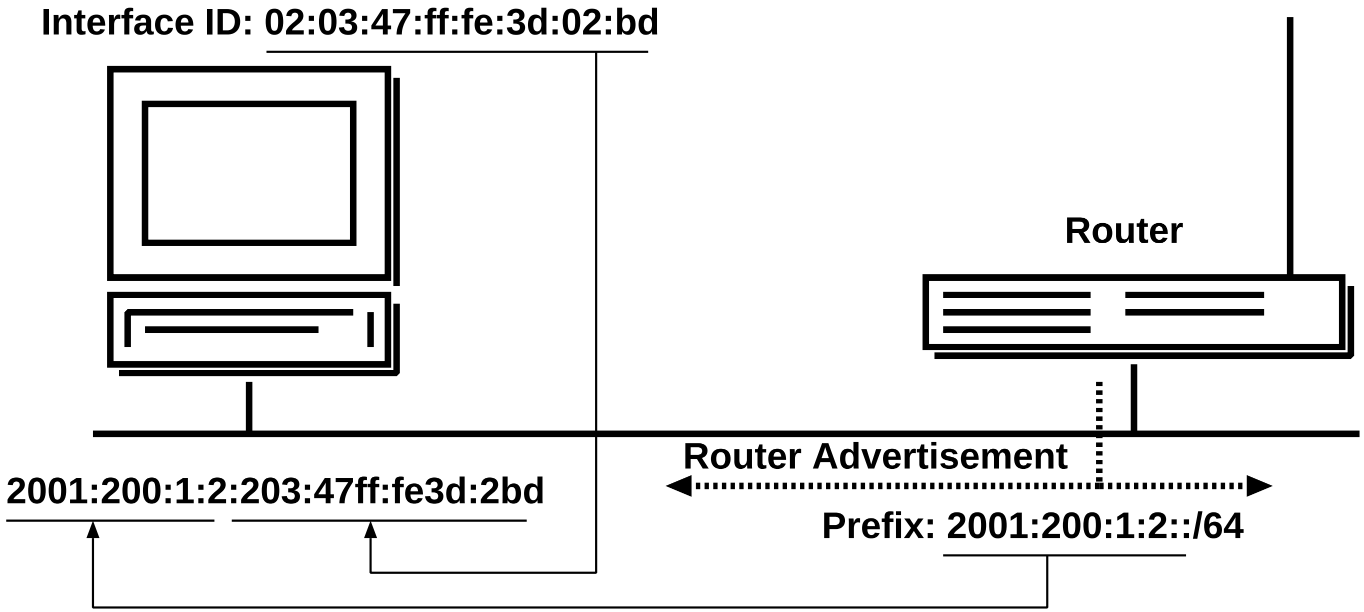
Global/Site-local address creation

- Extract prefix information from RA
- Concatinate global/site-local prefix and interface ID



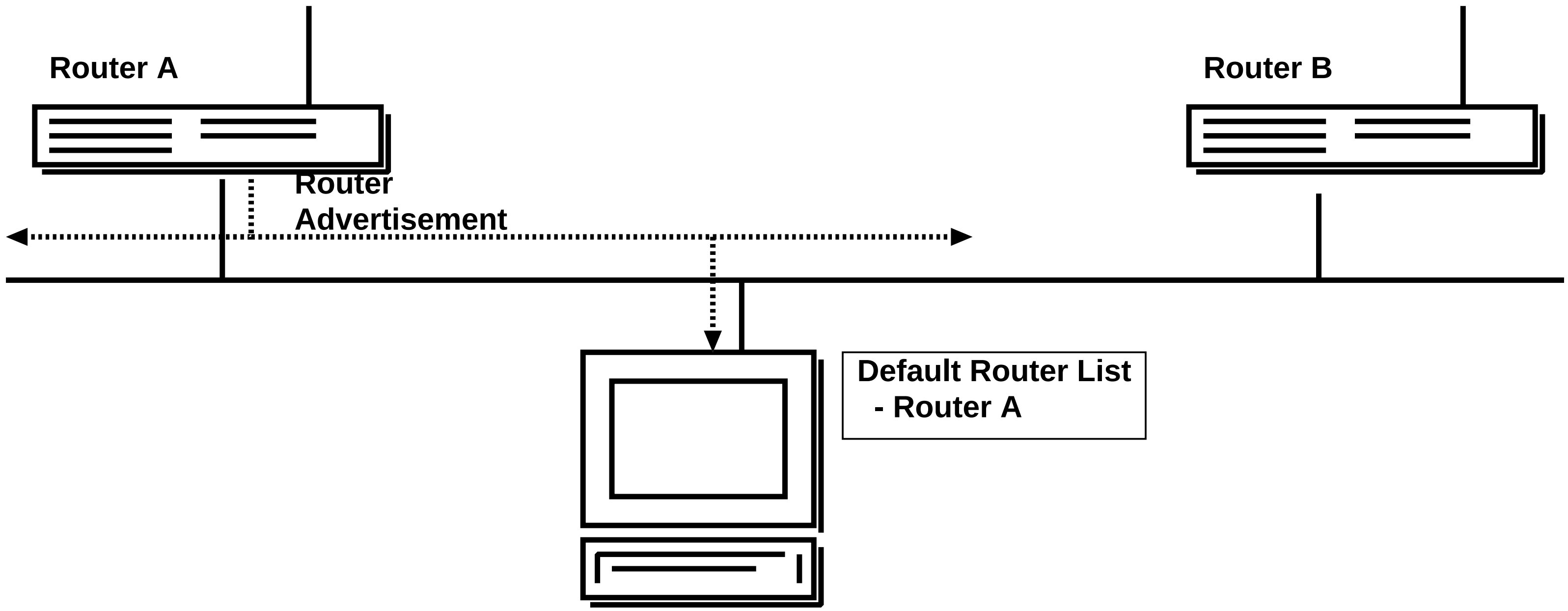
Global/Site-local address creation

- Extract prefix information from RA
- Concatinate global/site-local prefix and interface ID



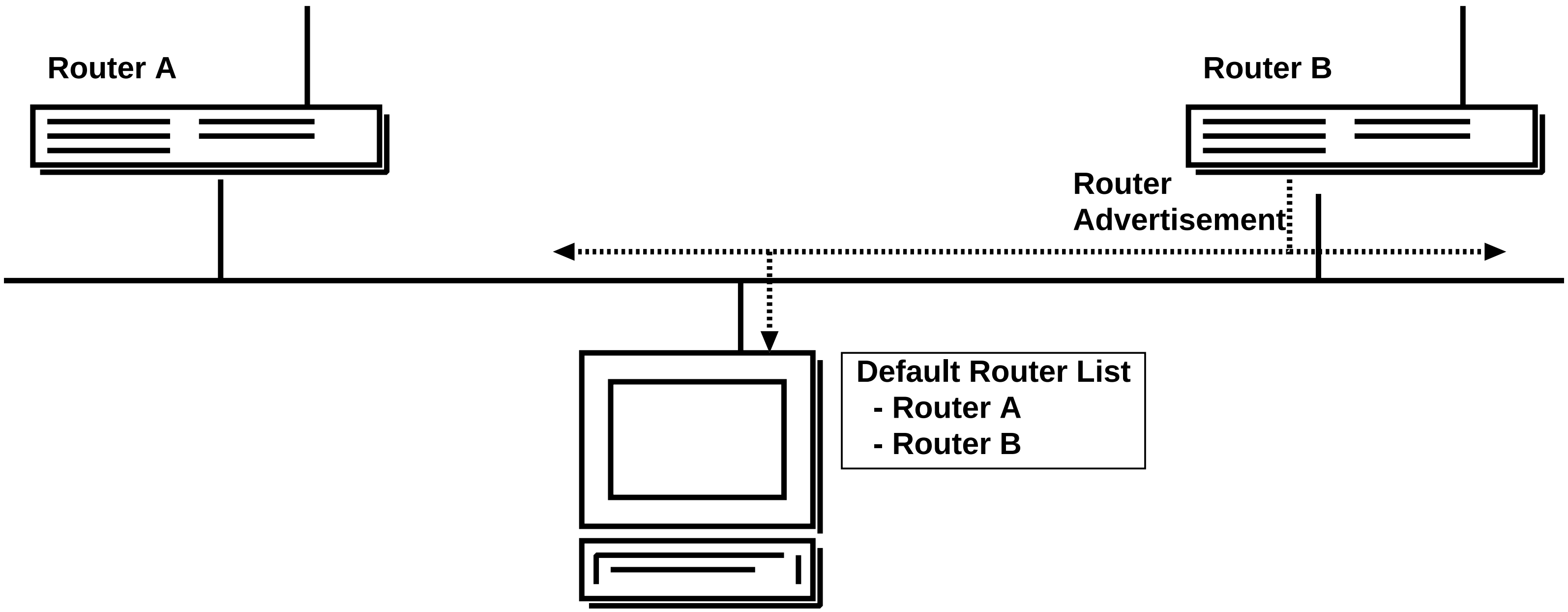
Default router discovery

- Routers send RA periodically
- Those routers are the candidates of the default router
- A host selects one router from the default router list



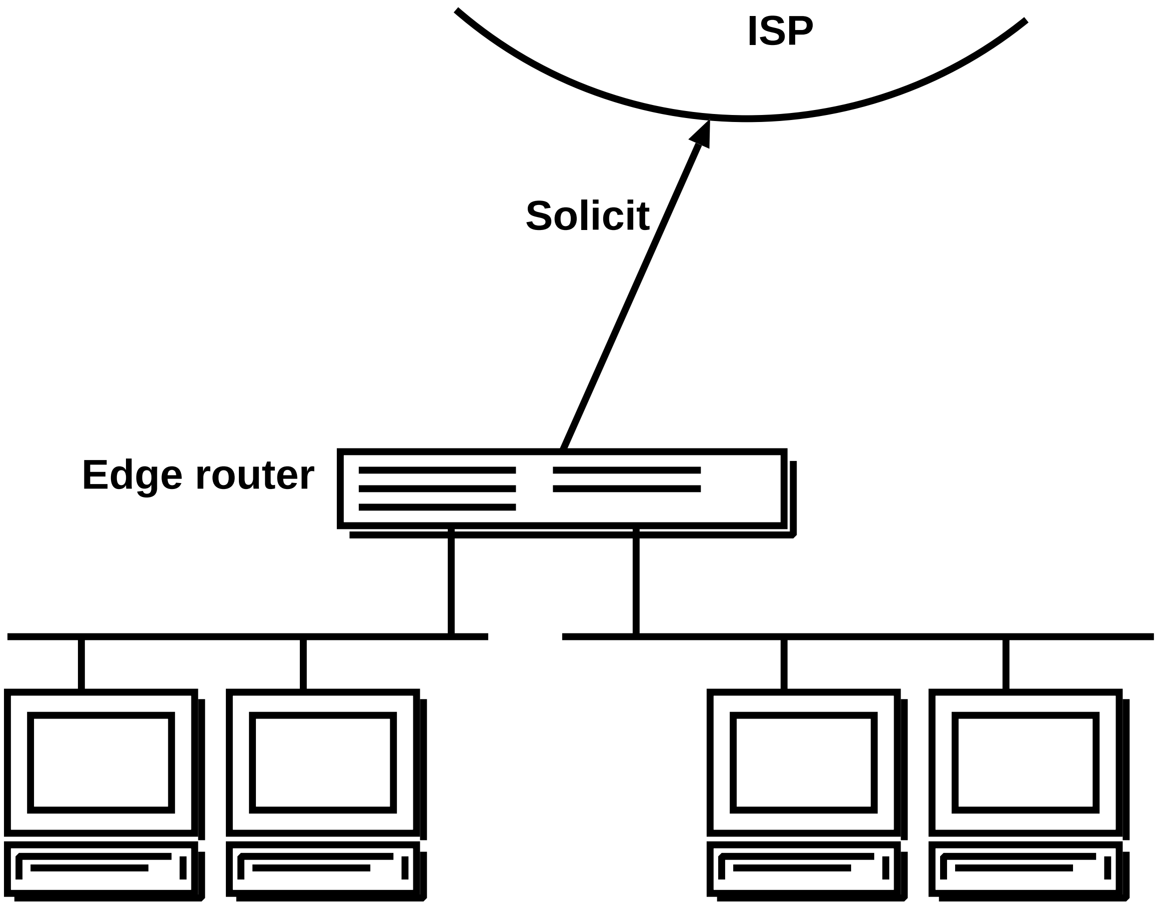
Default router discovery

- Routers send RA periodically
- Those routers are the candidates of the default router
- A host selects one router from the default router list



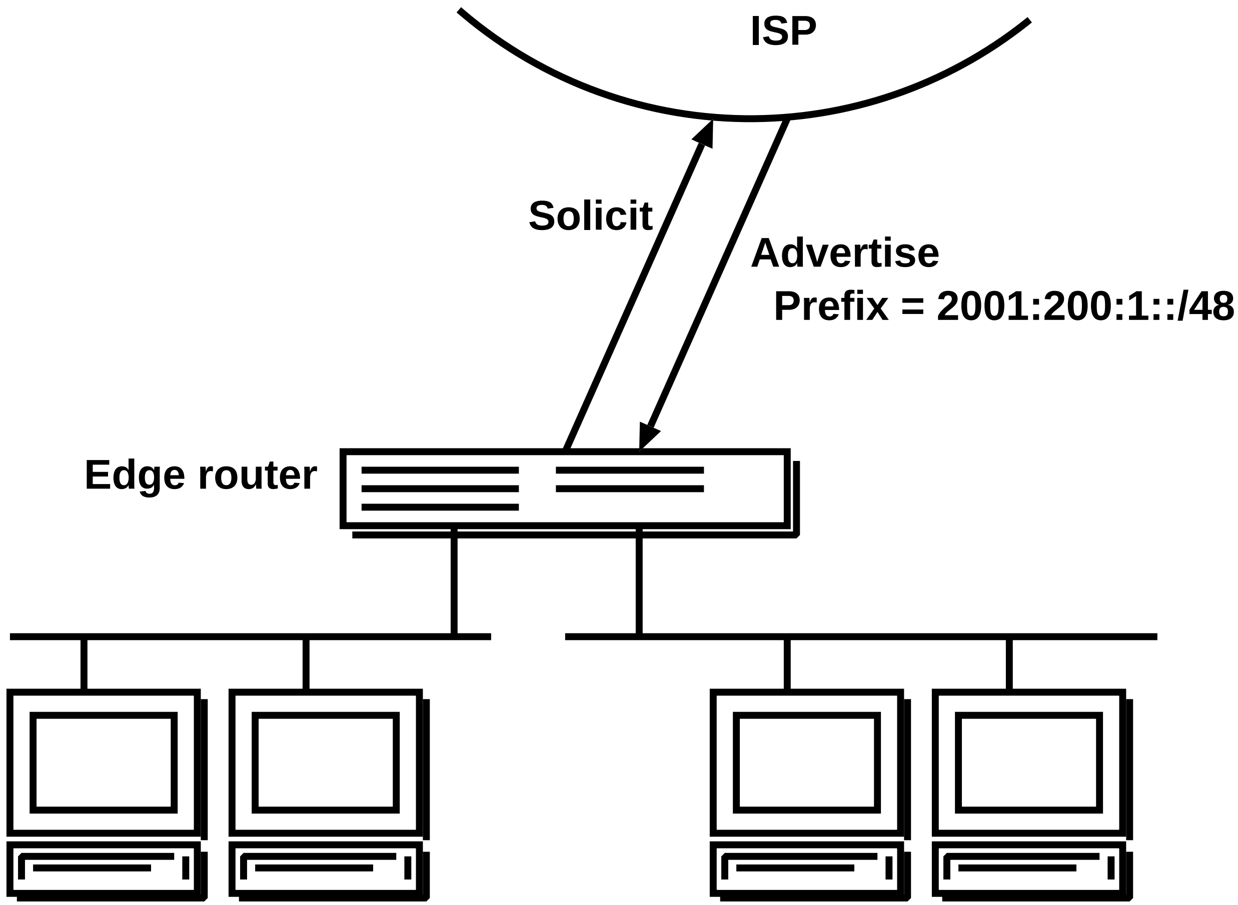
Prefix Delegation

- Provide prefix to an edge router
- No need to configure site prefixes by hand



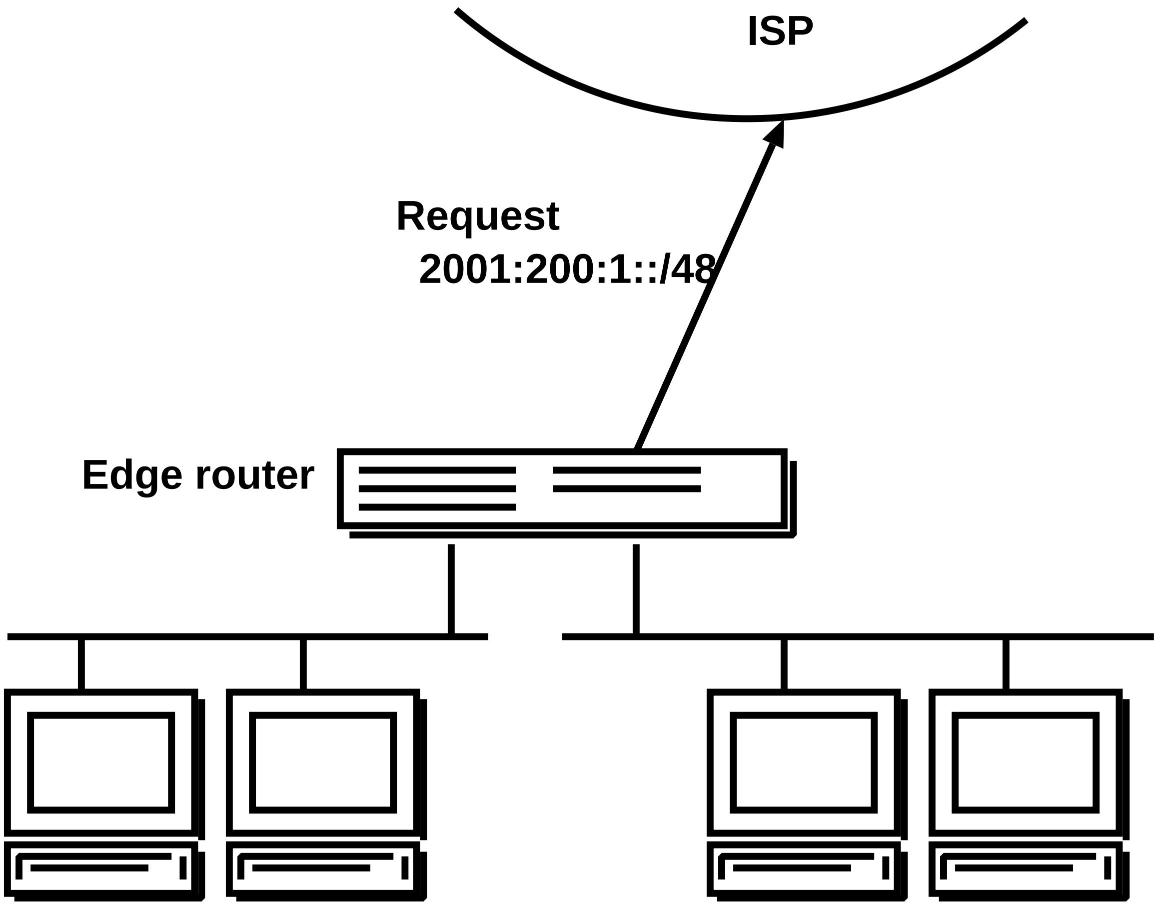
Prefix Delegation

- Provide prefix to an edge router
- No need to configure site prefixes by hand



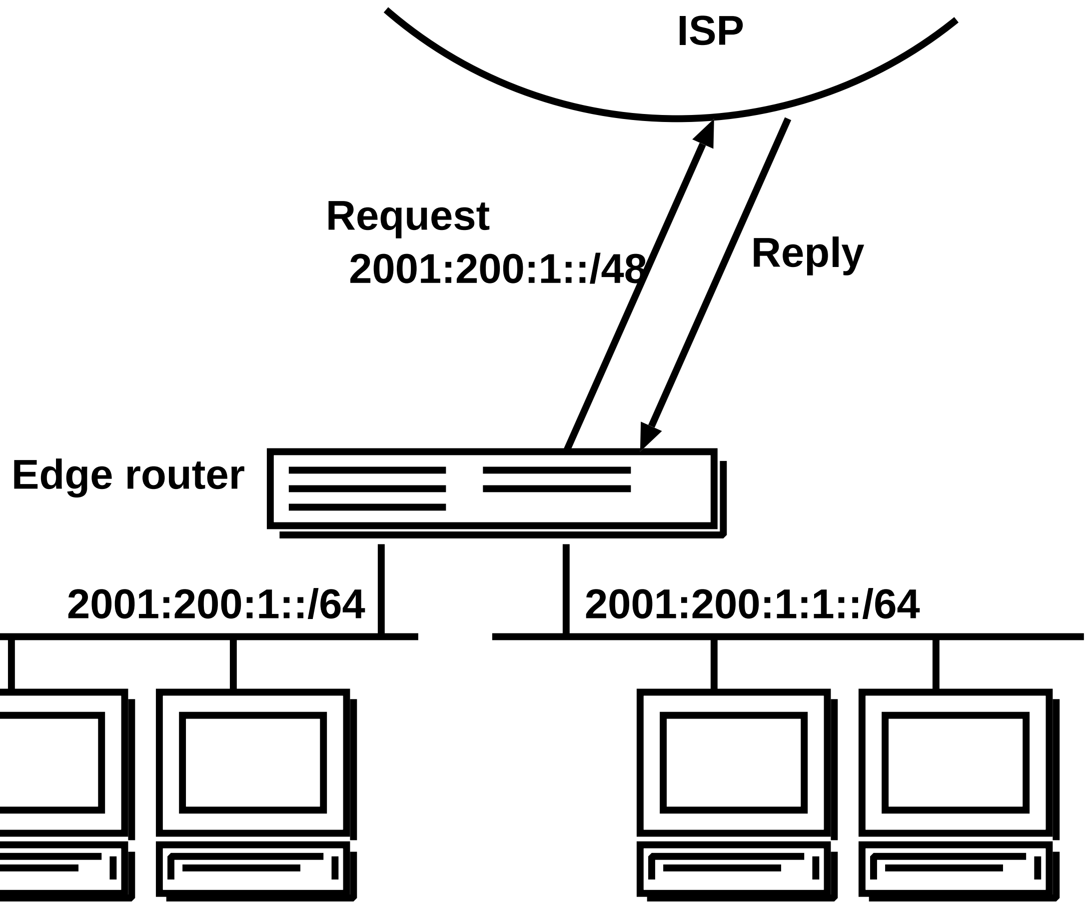
Prefix Delegation

- Provide prefix to an edge router
- No need to configure site prefixes by hand



Prefix Delegation

- Provide prefix to an edge router
- No need to configure site prefixes by hand



Why do we use IPv6?

IPv6 Addresses

Link-layer address resolution

Auto-configuration mechanism

DNS

Transision mechanisms

Deployment status

Recent event report

Accessing IPv6 services

- IPv6 nodes can be specified by hostnames as we can in IPv4
- Users are not aware of which protocol they are using
  - ping6 www.iij.ad.jp
  - You use IPv6 if your PC is connected to IPv6 cloud
  - You use IPv4 if your PC is not connected to IPv6
- Textual representtion can be used, of course
  - ping6 2001:240::80
  - Problem with using URL
    - ':' is used to specify a port number
    - http://www.iij.ad.jp:8080/
    - http://[2001:240::80]:8080/

DNS records

- AAAA record for IPv6 forward lookup

```
$ORIGIN iij.ad.jp.
www  IN  AAAA  2001:240::80
www  IN  A     202.232.2.10
```

- PTR record for reverse lookup

```
$ORIGIN 0.0.0.0.0.0.0.0.0.4.2.0.1.0.0.2.IP6.ARPA.
0.8.0.0.0.0.0.0.0.0.0.0.0  IN  PTR  www.iij.ad.jp.
$ORIGIN 2.232.202.IN-ADDR.ARPA.
10                         IN  PTR  www.iij.ad.jp.
```

- Other resource records are same as IPv4

DNS transport

- DNS query and answer can be on IPv4/IPv6
- Some resolver don't support IPv6 transport yet
  - DNS query/answer are done by IPv4
  - Such a node must be a dual stack node
  - But, users can use IPv6 applications
- Root DNS
  - Currently, root DNS servers are not IPv6 ready
  - DNS servers must be a dual stack node
  - A client can be an IPv6 only node

Why do we use IPv6?

IPv6 Addresses

Link-layer address resolution

Auto-configuration mechanism

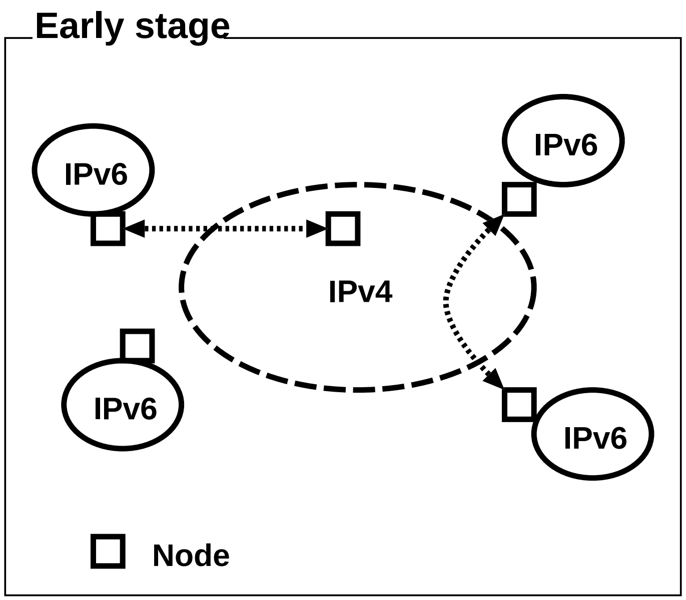
Transision mechanisms

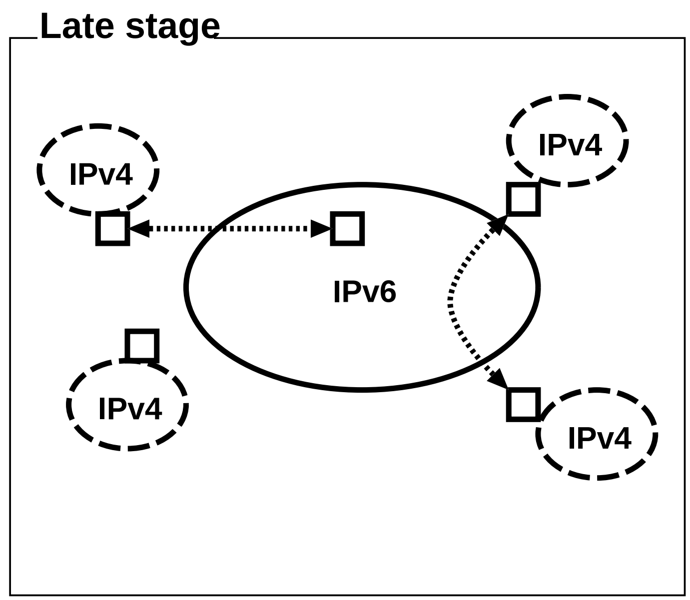
Deployment status

Recent event report

Transision stages

- Early stage
  - IPv4 network is wider than IPv6 network
  - There are many IPv6 islands
- Late stage
  - IPv4 networks are isolated



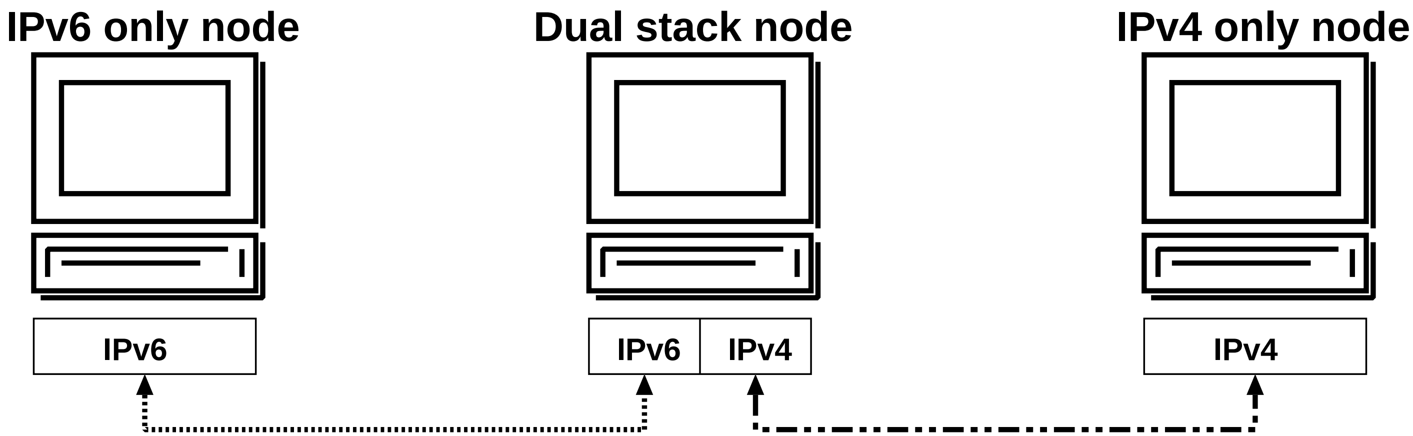


Transision mechanism types

- Dual stack node
  - Support both IPv4 and IPv6
- Tunneling
  - Encapsulate IPv6 packet in IPv4 packet (for early stage)
  - Encapsulate IPv4 packet in IPv6 packet (for late stage)
- Translator
  - Translate IPv6 packet to IPv4, and vice versa

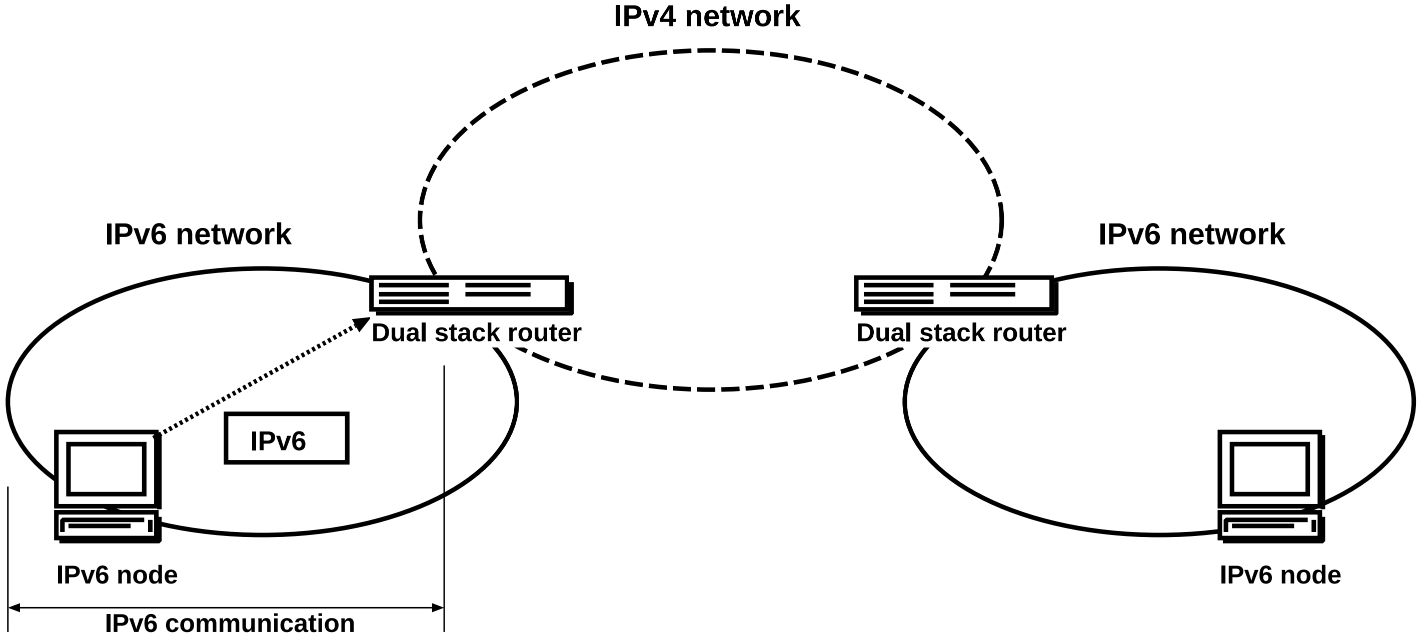
Dual stack node

- Dual stack node has both IPv4 and IPv6 address
- Use IPv4 address when communicating with IPv4 node
- Use IPv6 address when communicating with IPv6 node



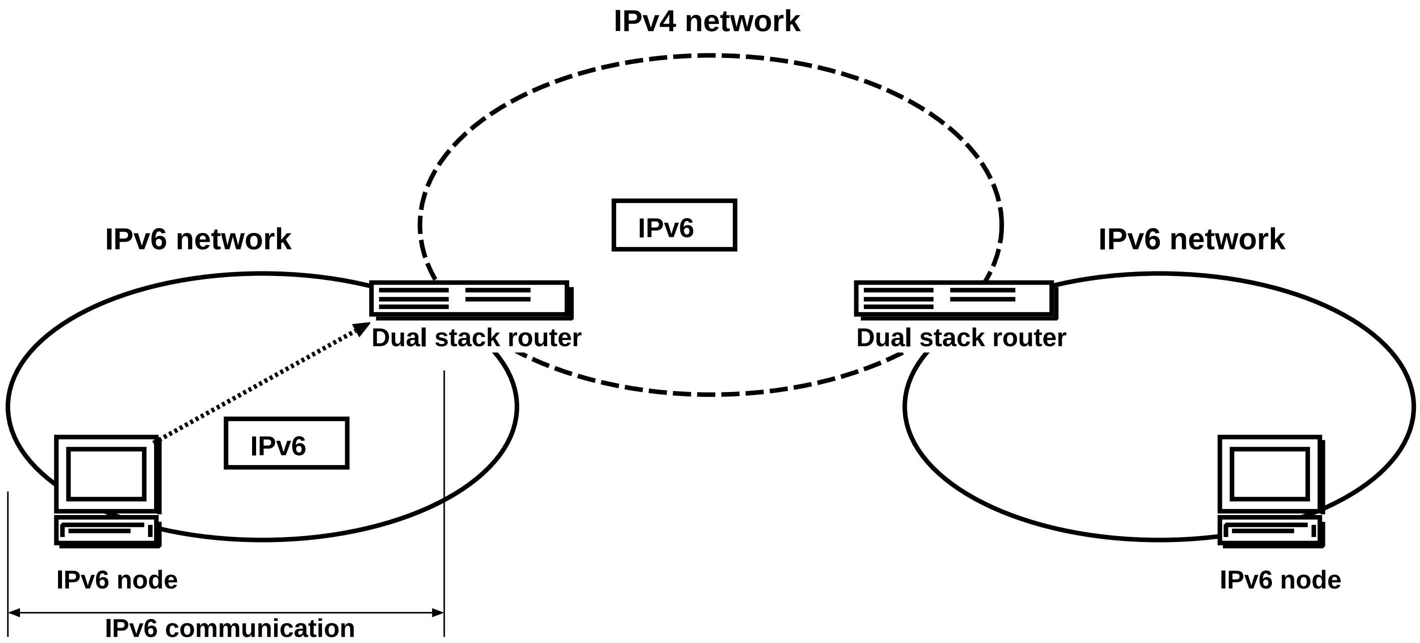
Tunneling

- IP in IP encapsulating
- Use IPv4(IPv6) as a datalink layer of IPv6(IPv4)
- Connect isolated IPv6(IPv4) networks/hosts over IPv4(IPv6) network
- Bordar routers must be a dual stack node



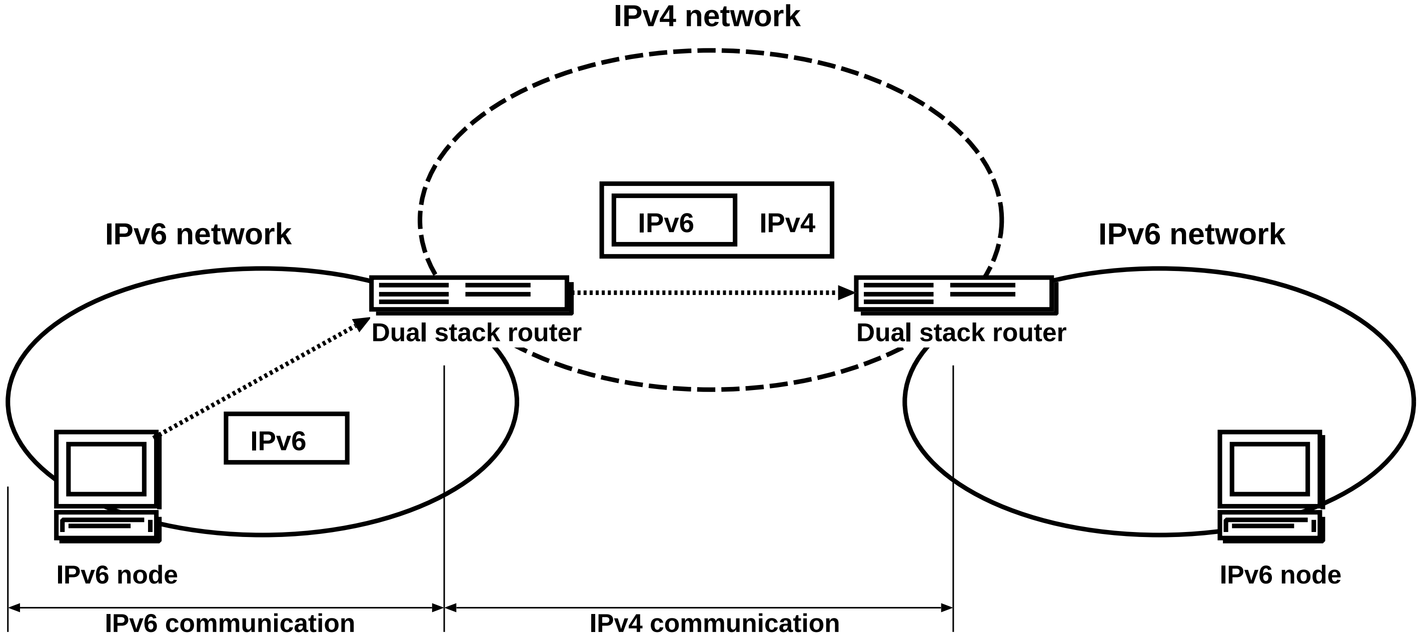
Tunneling

- IP in IP encapsulating
- Use IPv4(IPv6) as a datalink layer of IPv6(IPv4)
- Connect isolated IPv6(IPv4) networks/hosts over IPv4(IPv6) network
- Bordar routers must be a dual stack node



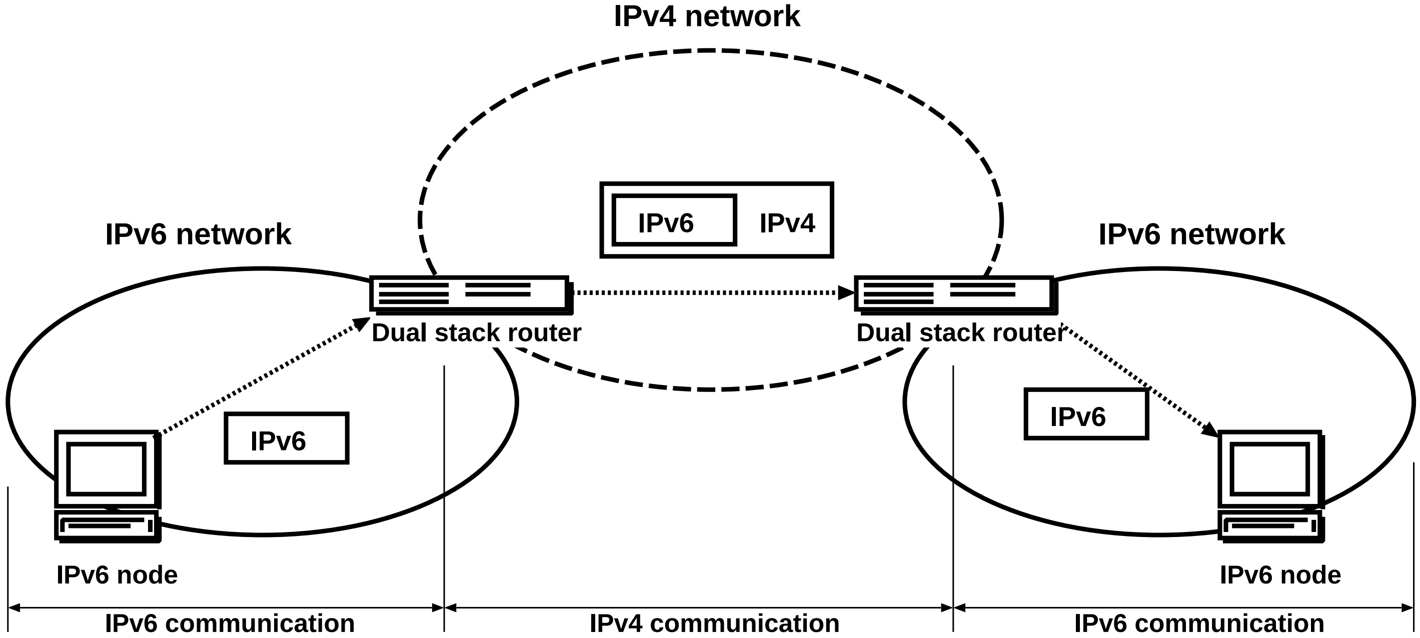
Tunneling

- IP in IP encapsulating
- Use IPv4(IPv6) as a datalink layer of IPv6(IPv4)
- Connect isolated IPv6(IPv4) networks/hosts over IPv4(IPv6) network
- Bordar routers must be a dual stack node



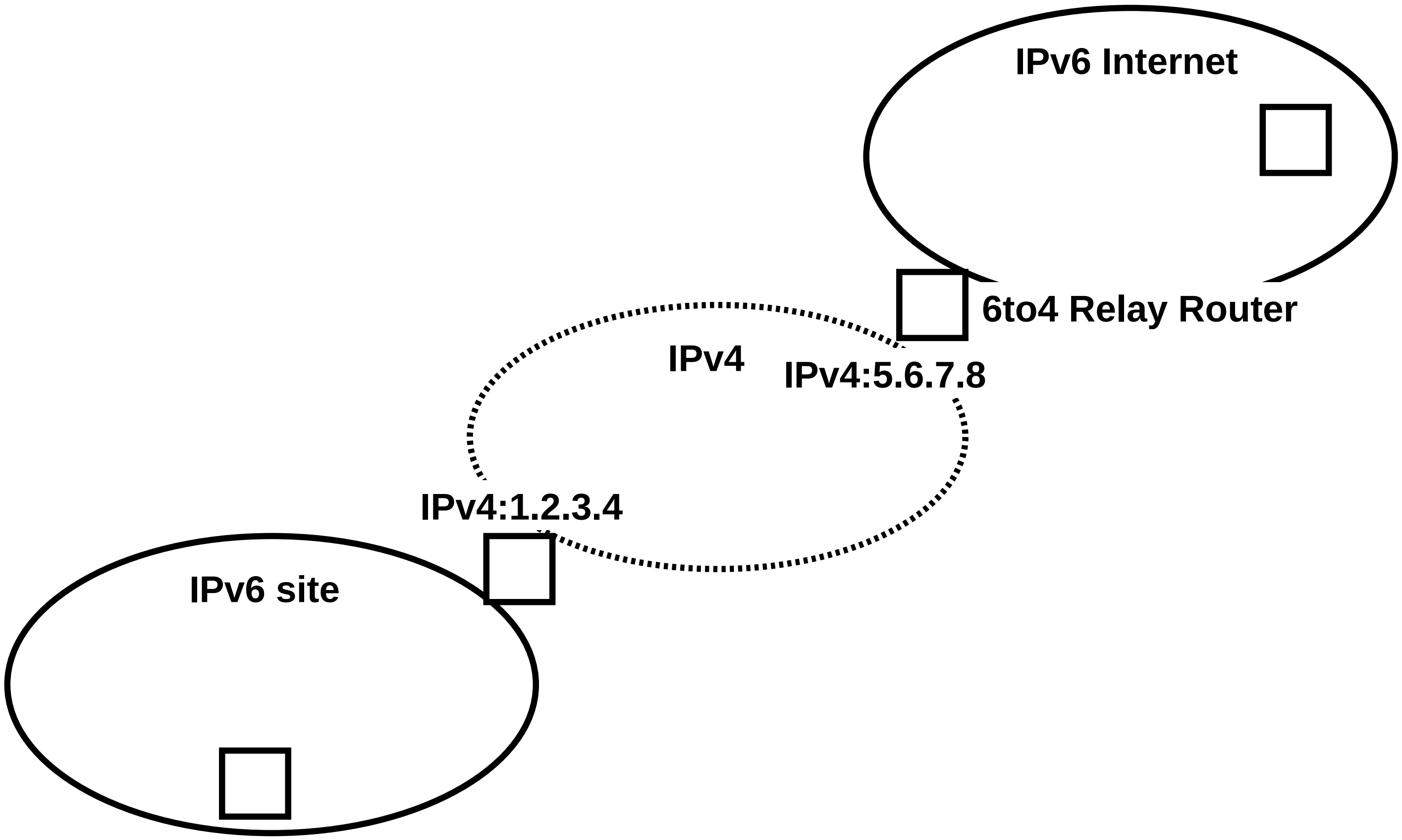
Tunneling

- IP in IP encapsulating
- Use IPv4(IPv6) as a datalink layer of IPv6(IPv4)
- Connect isolated IPv6(IPv4) networks/hosts over IPv4(IPv6) network
- Bordar routers must be a dual stack node



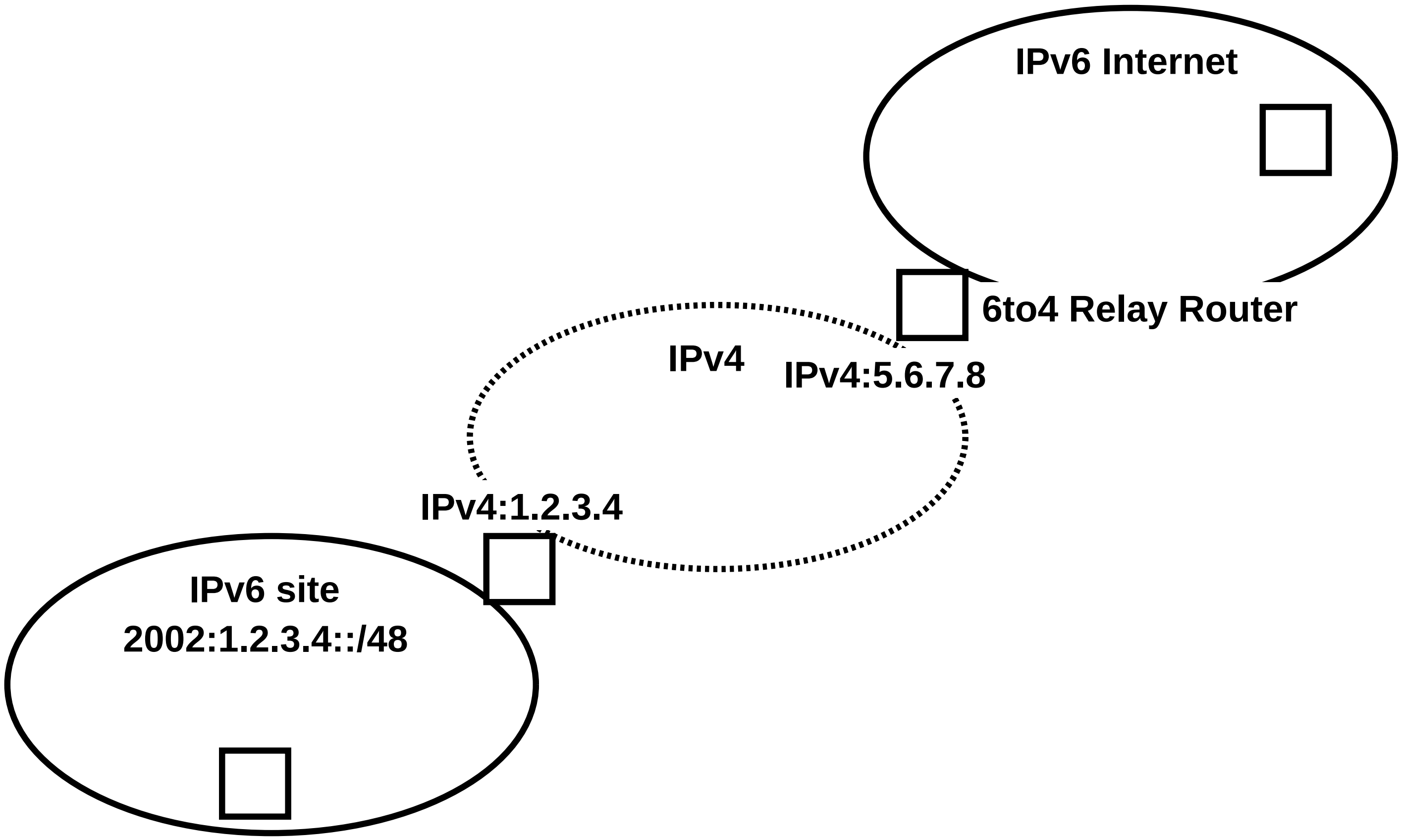
6to4 automatic tunneling

- Use other TLA ID (2) for tunneling
- Embed IPv4 address in IPv6 prefix
- A user can get /48 address space over tunnel



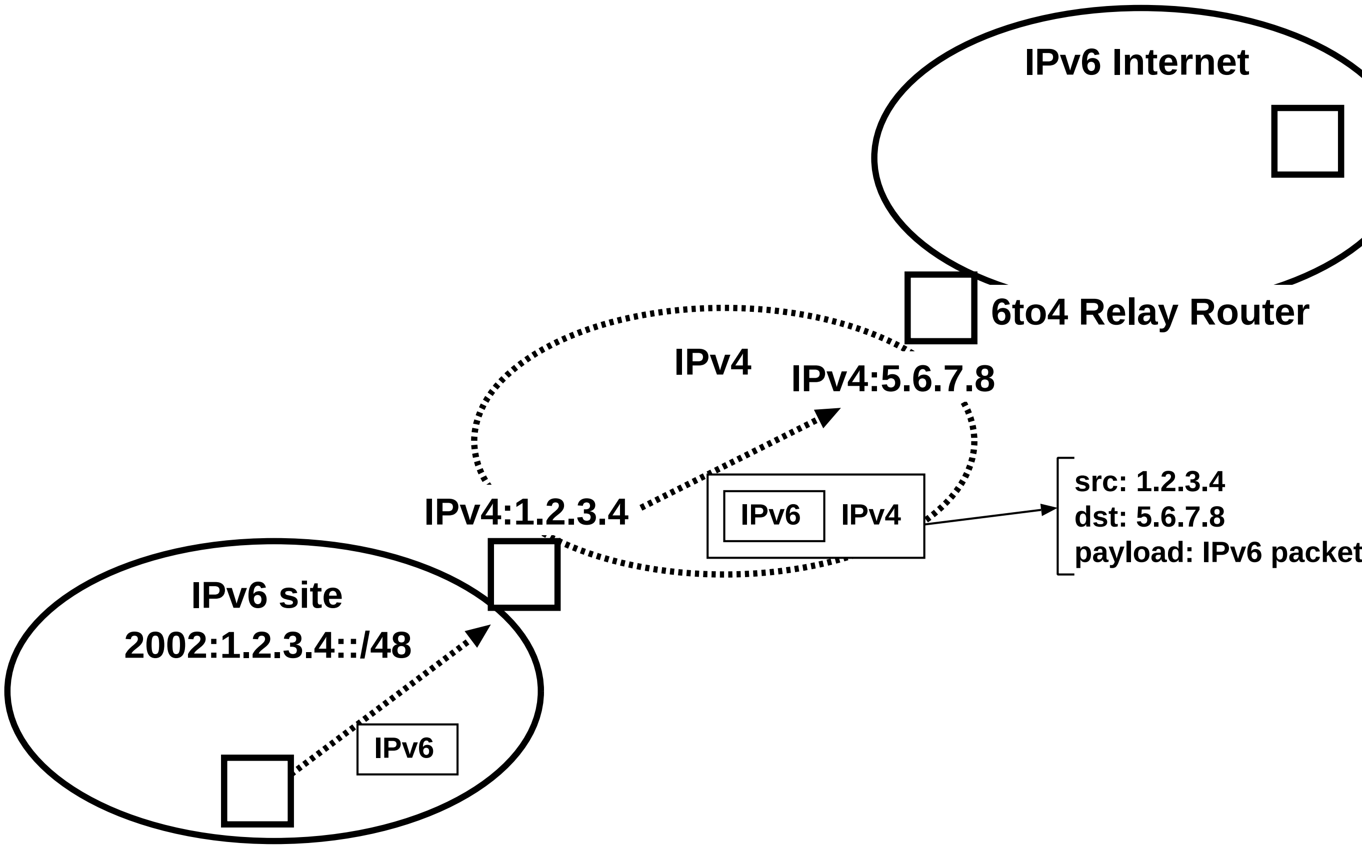
6to4 automatic tunneling

- Use other TLA ID (2) for tunneling
- Embed IPv4 address in IPv6 prefix
- A user can get /48 address space over tunnel



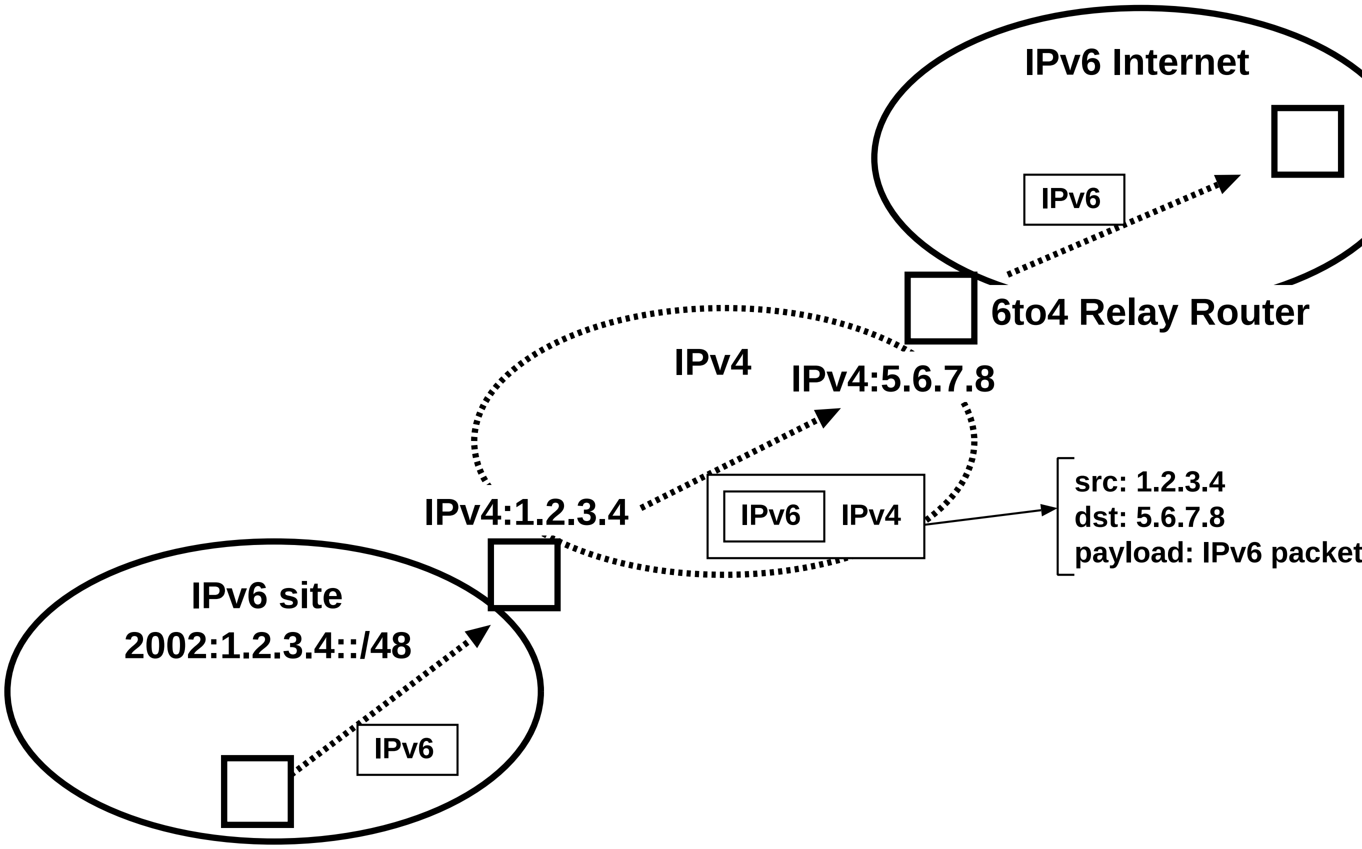
6to4 automatic tunneling

- Use other TLA ID (2) for tunneling
- Embed IPv4 address in IPv6 prefix
- A user can get /48 address space over tunnel



6to4 automatic tunneling

- Use other TLA ID (2) for tunneling
- Embed IPv4 address in IPv6 prefix
- A user can get /48 address space over tunnel



6to4 automatic tunneling

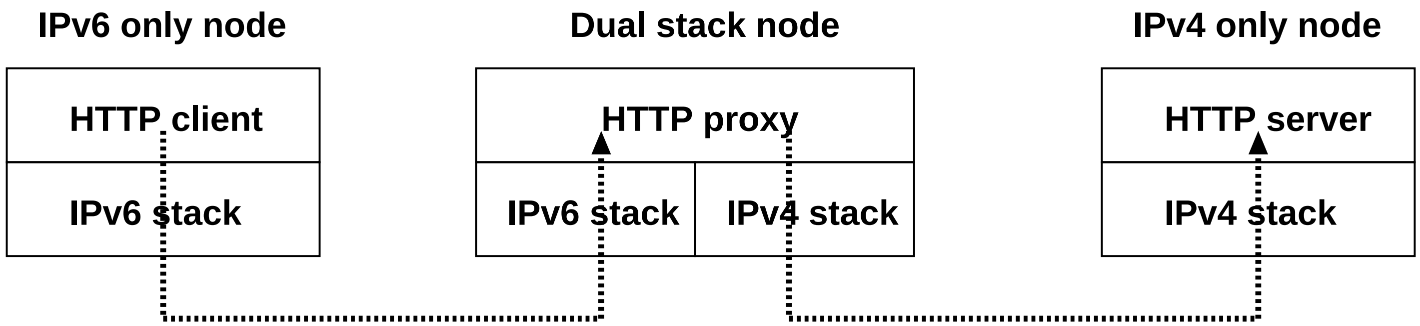
- Requirement
  - A user must have one (static) IPv4 global address
  - A user must know 6to4 relay router’s IPv4 address
- RFC3068 defines a special address for 6to4 relay router
- 6to4 relay router’s IP address may be provided statically from 6to4 service provider
- Public 6to4 relay routers
  - http://www.kfu.com/~nsayer/6to4/

Translator

- IPv4 never disappear
  - IPv6 and IPv4 will co-exist
- We must provide the way for them to communicate with each other
- Translator mechanisms
  - Application level gateway
    - Proxy (HTTP, FTP, and so on)
  - NAT-PT

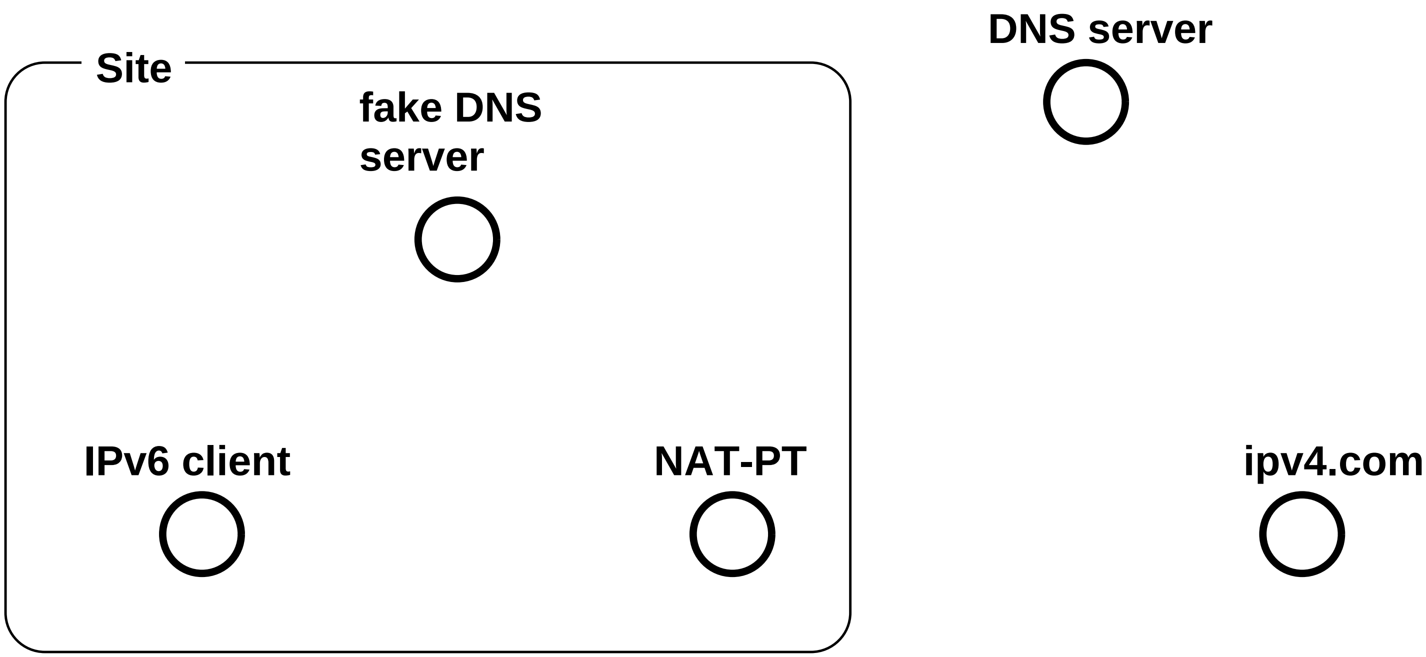
Application level gateway

- A kind of a proxy
- Proxy must be a dual stack node
- Proxy receives requests on its IPv6 interface from IPv6 client
- Proxy sends requests to IPv4 server using its IPv4 interface
- Example



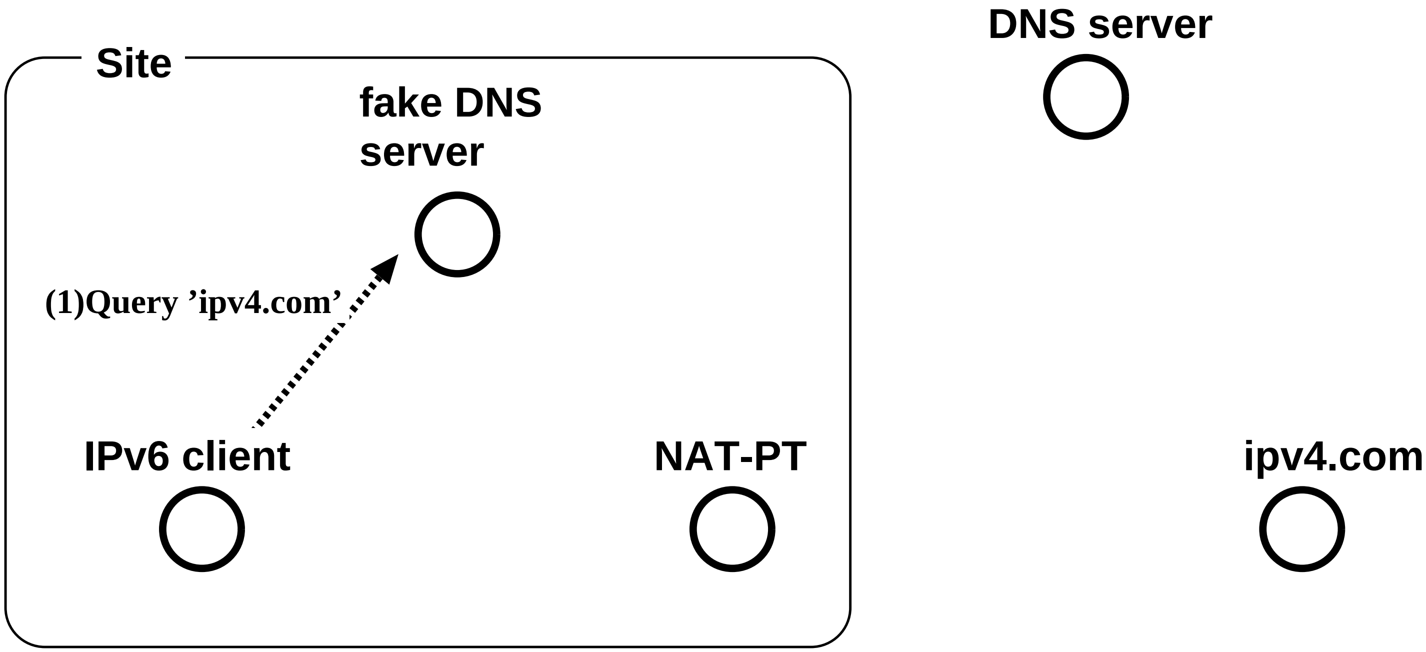
NAT-PT

- Map IPv4 addresses to special IPv6 addresses using a fake DNS server
- Provide transparent connection to IPv6 nodes
- IPv6 nodes communicates with IPv4 node as if it is IPv6 node



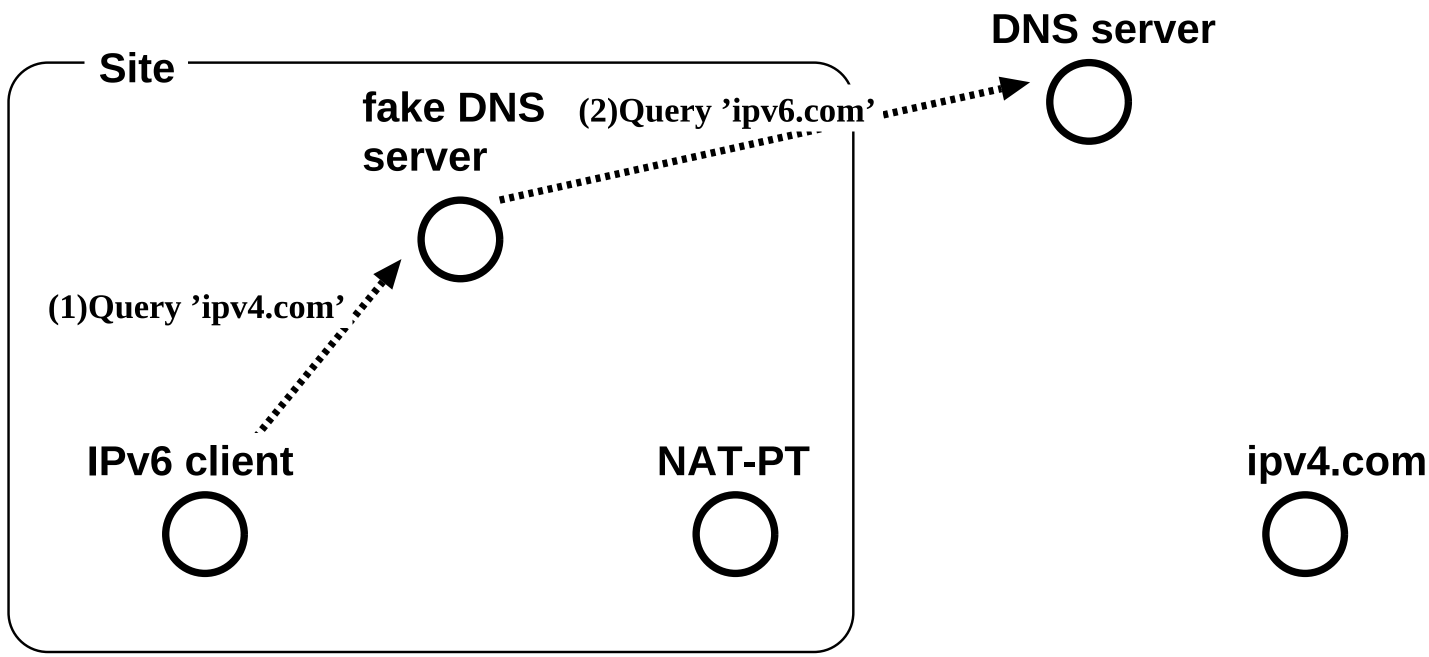
NAT-PT

- Map IPv4 addresses to special IPv6 addresses using a fake DNS server
- Provide transparent connection to IPv6 nodes
- IPv6 nodes communicates with IPv4 node as if it is IPv6 node



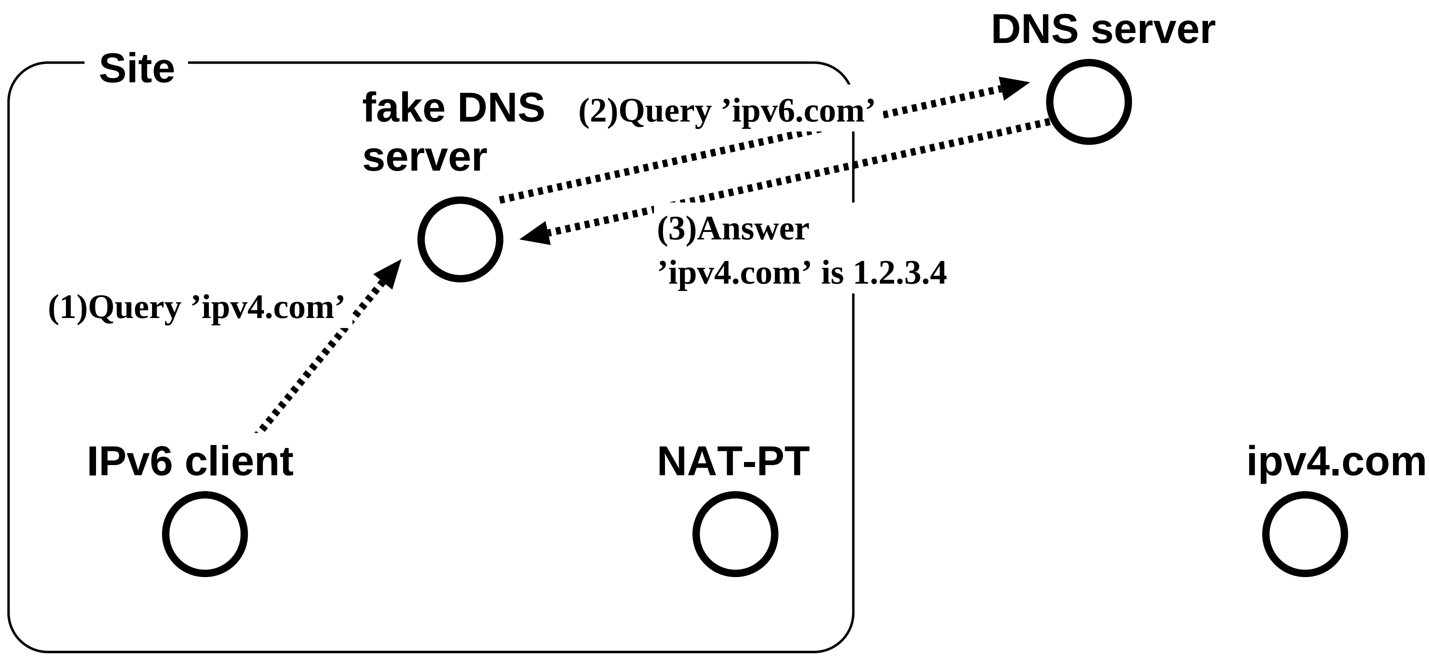
NAT-PT

- Map IPv4 addresses to special IPv6 addresses using a fake DNS server
- Provide transparent connection to IPv6 nodes
- IPv6 nodes communicates with IPv4 node as if it is IPv6 node



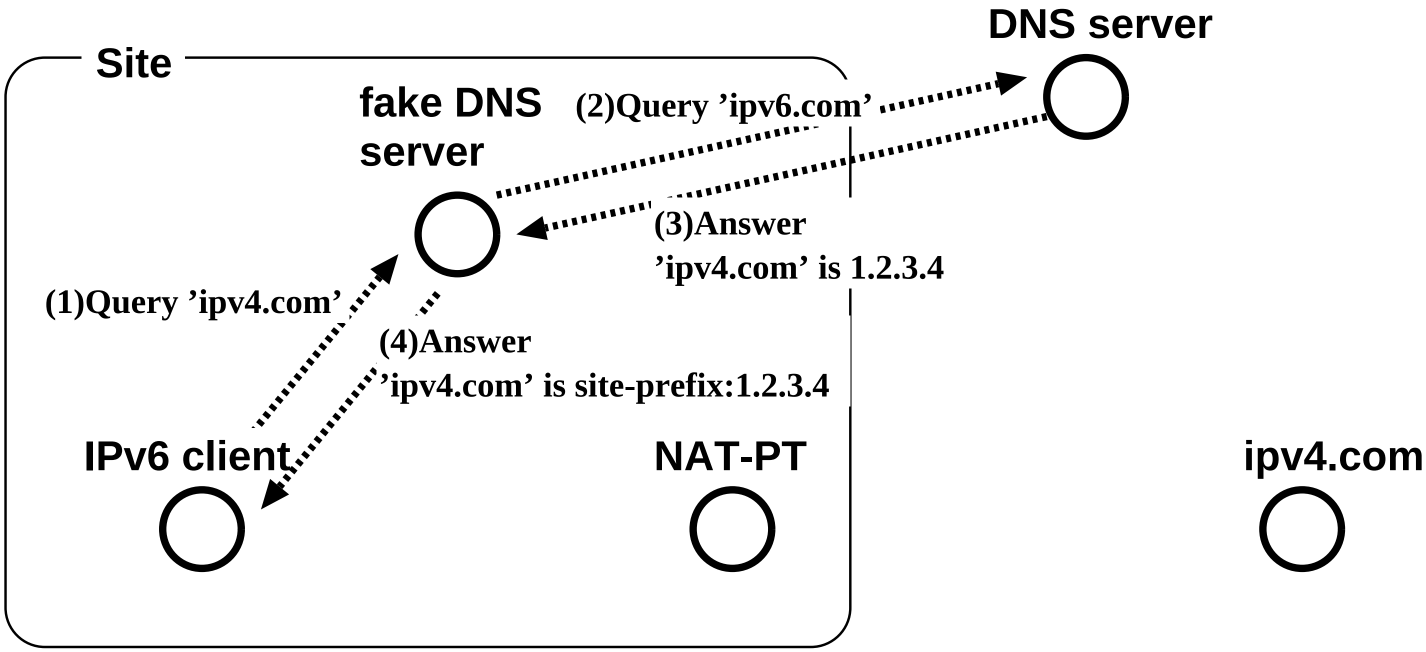
NAT-PT

- Map IPv4 addresses to special IPv6 addresses using a fake DNS server
- Provide transparent connection to IPv6 nodes
- IPv6 nodes communicates with IPv4 node as if it is IPv6 node



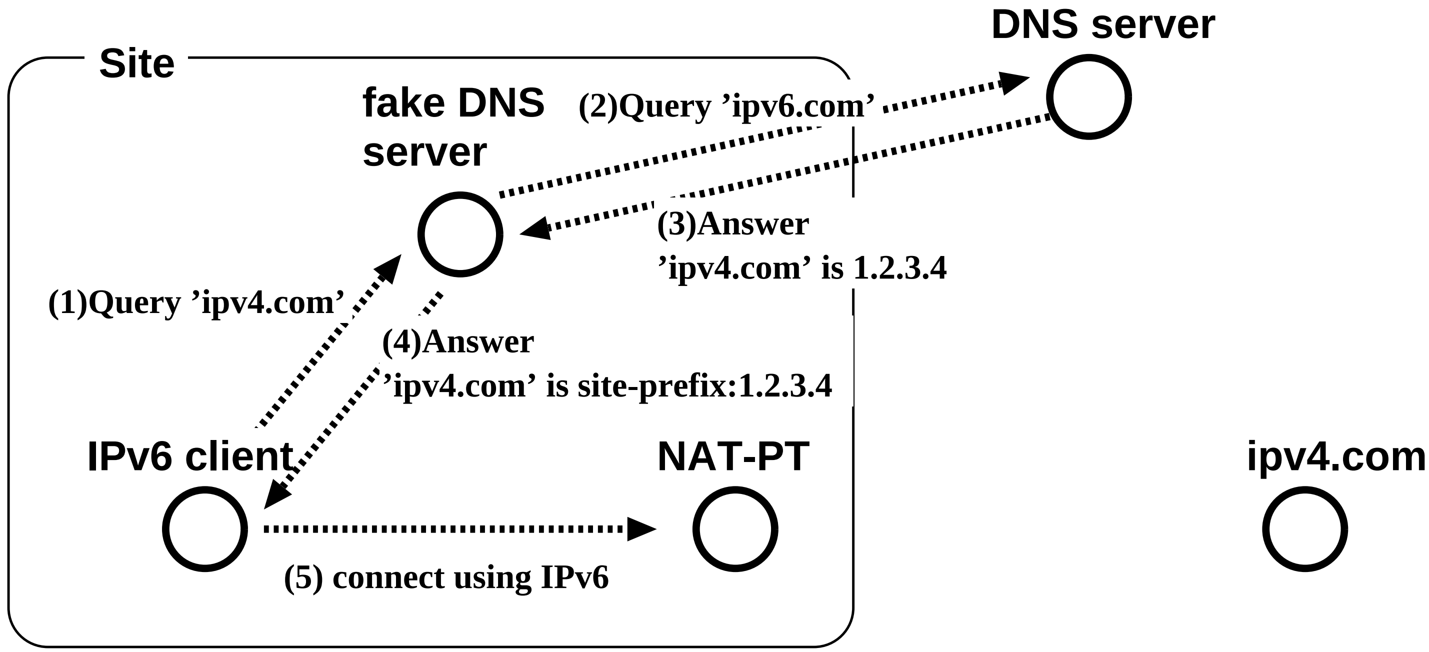
NAT-PT

- Map IPv4 addresses to special IPv6 addresses using a fake DNS server
- Provide transparent connection to IPv6 nodes
- IPv6 nodes communicates with IPv4 node as if it is IPv6 node



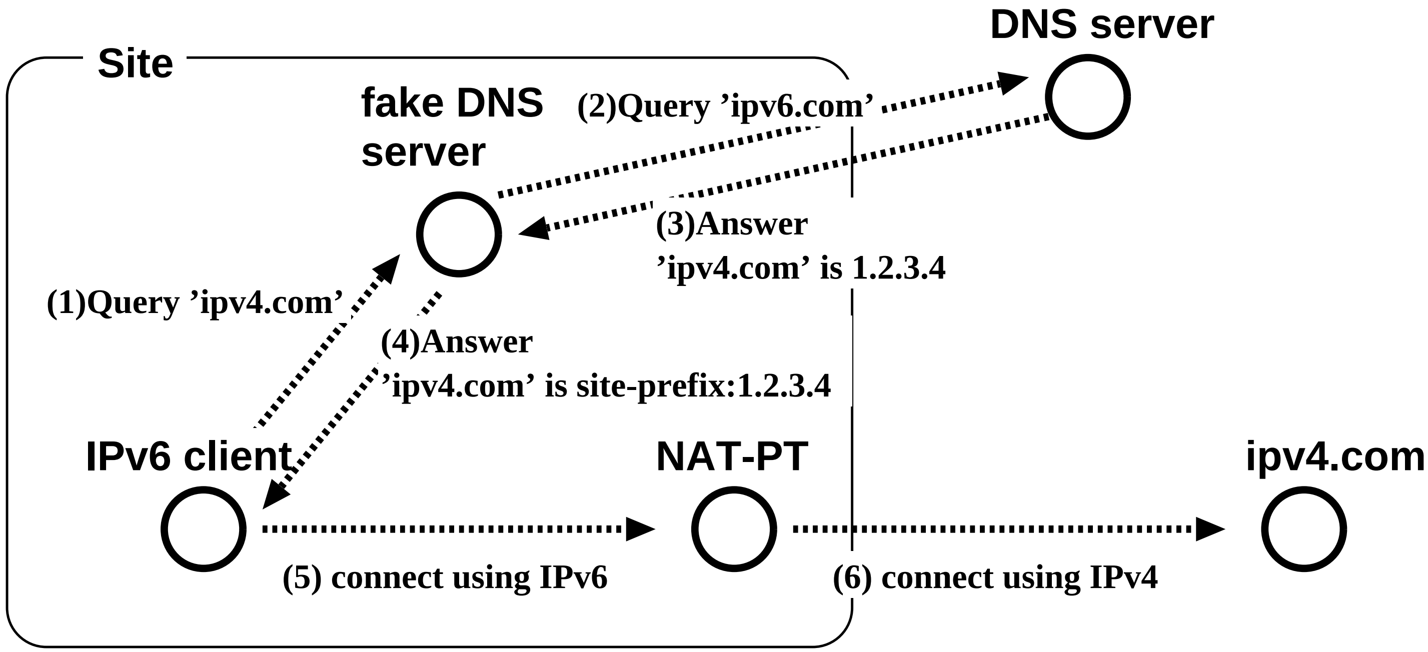
NAT-PT

- Map IPv4 addresses to special IPv6 addresses using a fake DNS server
- Provide transparent connection to IPv6 nodes
- IPv6 nodes communicates with IPv4 node as if it is IPv6 node



NAT-PT

- Map IPv4 addresses to special IPv6 addresses using a fake DNS server
- Provide transparent connection to IPv6 nodes
- IPv6 nodes communicates with IPv4 node as if it is IPv6 node



Problems of translator

- Have same problems which NAT has
- Break end-to-end security
- Hard to translate if the protocol itself utilizes address information (e.g. FTP, VoIP)
    - We need a special gateway per protocol

Why do we use IPv6?

IPv6 Addresses

Link-layer address resolution

Auto-configuration mechanism

DNS

Transision mechanisms

Deployment status

Recent event report

Deployment areas

- Network products
  - Routers, Switches
- User end products
  - Operating Systems
- ISP
  - Consumer/Prosumer ISP services
- Software

Network products

- Many vendors are shipping IPv6 enabled boxes
  - Cisco Systems
  - Hitachi
  - Juniper Networks
  - Nortel Networks
  - 6Wind
  - IIJ
  - YAMAHA
  - NEC
  - Fujitsu
  - 3Com
  - many other...

User end products

- Many Operating Systems support IPv6
  - UNIX
    - NetBSD, FreeBSD, OpenBSD, BSD/OS
    - Linux
    - Solaris
    - HP-UX
    - IRIX
    - AIX
    - etc
  - Windows
    - Windows XP
    - Windows 2000 (additional patches needed)
    - Windows CE.NET
  - Macintosh
    - MacOS X 10.2
  - Embeded OS
    - VxWorks
    - TRON

ISP

- In Japan, many ISPs provide IPv6 services
- Commercial service
  - NTT Communications
  - IIJ
  - Japan Telecom
  - PoweredCom
  - IMASY
- Experimental service
  - JENS
  - KDDI
  - AboveNet
  - Chita Media Network
  - Nifty
  - KMN
  - Miako net

Software

- Many software supports IPv6
  - Network programs bundled with BSD/Linux
  - Sendmail/Postfix
  - Apache
  - Mozilla/Internet Explorer
  - BIND

Why do we use IPv6?
IPv6 Addresses
Link-layer address resolution
Auto-configuration mechanism
DNS
Transision mechanisms
Deployment status

Recent event report

IPv6 ShowCase (N+I 2002, July 2002)

- IPv6 town image is presented
- 3 zones
  - ISP/Datacenter zone
  - Home zone
  - Mobile zone
- Over 30 companies/organizations participated

ISP/Datacenter zone

- ISP services
  - Connectibily
  - Prefix Delegation
- Router/Switch products
  - Many vendor supports IPv6
- Radius products

ISP/Datacenter zone

- Routers and Switches

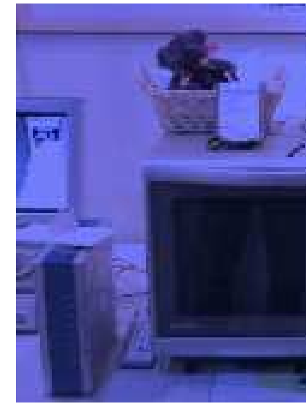
Home zone

- Home appliances
  - Digital camera
  - Microoven
  - Refrigerator
- VoD software
- P2P application
- Live camera

Home zone

- Home appliances





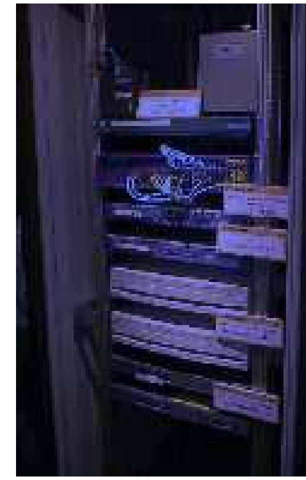
- Game console / P2P application


Mobile zone

- Mobile IPv6
  - Mobile Video/Music player
  - Mobile conference tools
- Network mobility
  - Internet car
- Many small devices IPv6/Mobile IPv6 enabled
  - PDA
  - Handheld PC
  - Note PC

Mobile zone

- Mobile nodes and home agents



- Internet car

Many IPv6 related products

- One chip IPv6 processer
- IPv6 network management tools
- Radius servers
- IP phone over IPv6
- Cipher chip for IP security
- Embeded OSes which support IPv6

Summary

- IPv6 is not a next generation protocol
- IPv6 is a current protocol
- It is not too early to start IPv6
  - IPv6 has many advantages
  - Hardware/Software are ready
  - Network infrastructure is ready
- Not to be late!