COMP9332
Network Routing & Switching

IPv6, Routing Basics

http://www.cse.unsw.edu.au/~cs9332/

This lecture

- IP version 6 (IPv6)
- Introduction to routing
  - Basic mechanics of IP packet delivery

IP version 6 (IPv6)

Motivations for IPv6

- IPv4 addresses are running out
  - Inherent problem of network-host hierarchy
    - » 100% address assignment efficiency is not possible
    - » E.g. Even with CIDR, a network with 600 hosts requires a network with 1024 addresses
  - Proliferation in the number of networks
  - Growth in the number of and type of devices having Internet connectivity

New features of IPv6

- Expanded address space: 128-bit address (c.f. 32-bit for IPv4)
- Address autoconfiguration
- Support for
  - Real-time service
  - IP multicast
  - Mobile IP
  - Security
  - Anycast
  - Note: Most of these services are added onto IPv4 but IPv6 must support them

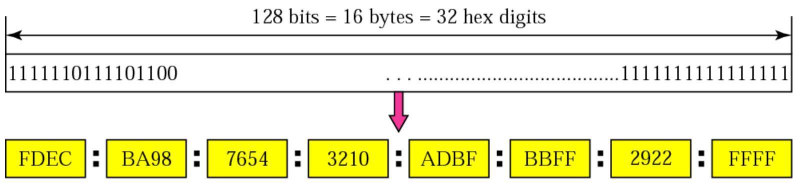
Topics for IPv6

- IPv6 addressing
- Autoconfiguration
- Transition from IPv4 to IPv6

IPv6 addressing

- 128 bits means you can have 2^128 addresses, which is 340,282,366,920,938,463,463,374,607,431,768,211,456
  - This is approximately 3.4 x 10^38
  - Compare with 4 x 10^9 IPv4 addresses, IPv6 has 10^29 times more addresses
- Earth's surface area (land + water) is 500 x 10^14 sq. metres ⇒7 x 10^21 addresses per sq. metre
- One reason why the address space is so large
  - Address assignment can never achieve 100% efficiency
  - but there are other reasons (to be discussed later)

IPv6 address format



Abbreviated address

Unabbreviated

FDEC : BA98 : 0074 : 3210 : 000F : BBFF : 0000 : FFFF

FDEC : BA98 : 74 : 3210 : F : BBFF : 0 : FFFF

Abbreviated

Abbreviated address with consecutive zeros

Abbreviated

FDEC : 0 : 0 : 0 : 0 : BBFF : 0 : FFFF

FDEC : : BBFF : 0 : FFFF

More Abbreviated

The address can also be abbreviated as FDEC:0:0:0:0:BBFF::FFFF

Exercises

- What is the complete IPv6 address for
  - ABBA:CAB:1234::FEED:3:BEEF

Solution and more exercise

- Solution: The complete address for ABBA:CAB:1234::FEED:3:BEEF is ABBA:0CAB:1234:0000:0000:FEED:0003:BEEF
- Exercise: What is the complete address for
  - ABBA::7::FEED?

Solution

- Solution: The address is invalid. Only one group of zeros can be suppressed.

IPv6 address architecture

- All addresses are classless
- Some defined address prefixes (rfc3513)

Prefix	Type
001	Global unicast address
1111 1110 10	Link-Local Unicast Addresses
1111 1111	Multicast Addresses

Host configuration

- A host needs to know its
  - IP address
  - Prefix length
- IPv4 uses BOOTP and DHCP
  - DHCP server maintains a pool of available IP address and gives them out on request
  - DHCP server keeps track which address has been used (stateful configuration)

IPv6 autoconfiguration (1)

- Autoconfiguration is done as follows
  - Step 1: (A bit later)
  - Step 2: Host sends out a router solicitation message
  - Step 3: Router responses with router advertisement which includes
    - » Network prefix
    - » Prefix length

IPv6 autoconfiguration (2)

- Step 3: The host IP address is

Network prefix	Padding (0's)	64-bit interface ID

- The number of zeros in padding is chosen to make it a 128 bit address

IPv6 autoconfiguration (3)

- The 64-bit interface ID is formed from the physical address of the interface
  - The new Ethernet address is 64 bits long and is inserted into the interface ID
  - The old Ethernet address is 48 bits long which consists of company code (24 bits) plus Ethernet extension identifier (24 bits). The interface ID is

Company code	$FFFE_{16}$	Ethernet extension identifier
24 bits	16 bits	24 bits

IPv6 autoconfiguration (4)

- Since Ethernet MAC addresses are unique, autoconfiguration ensures that host IP addresses are also unique
- However, just to make sure, there is step 1
- Step 1:
  - Host form a link local address using prefix 1111 1110 10 + [zeros] + interface ID
  - Host sends out a Neighbour Discovery (part of ICMPv6) using the link local address as the target address
  - If another host on the network has the same link local address, it will reply ⇒ autoconfiguration fails
  - Otherwise, continue onto step 2

IPv6 autoconfiguration (5)

- IPv6 autoconfiguration
  - Does not require a special server
  - Is stateless as routers do not need to keep track of which address is used
- If IPv6 autoconfiguration is used, what will the size of the smallest possible network be?

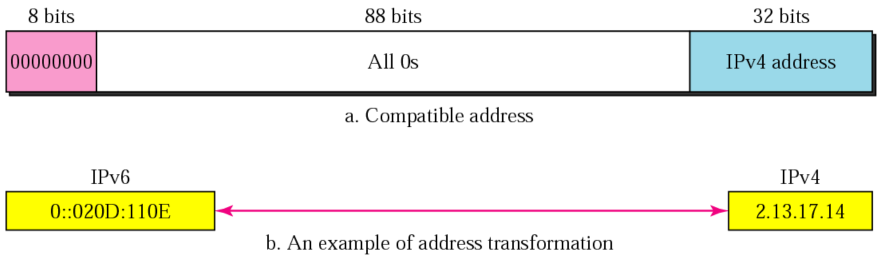
Why 128 bits?

- Allows autoconfiguration
  - Simply "dump" the hardware address in the last part
  - The smallest subnets have 2^64 addresses
- There are still 64 bits for different networks
  - Plenty of flexibility!
- Plenty of addresses
  - This is a reason but not the only one!

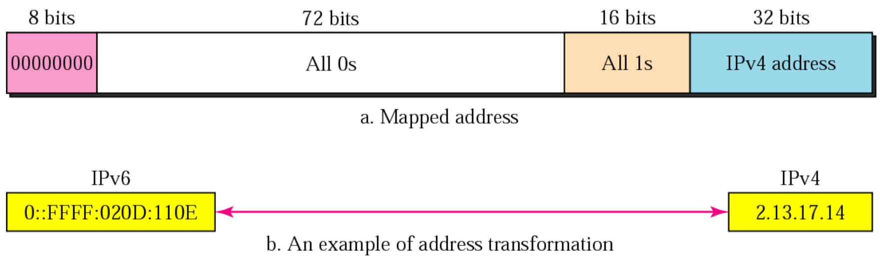
IPv4-friendly IPv6

- Two special IPv6 address formats to help transition from IPv4 to IPv6
  - IPv4 compatible address
  - IPv4 mapped address
- In both formats, IPv4 address is contained within IPv6 address

IPv4 compatible address

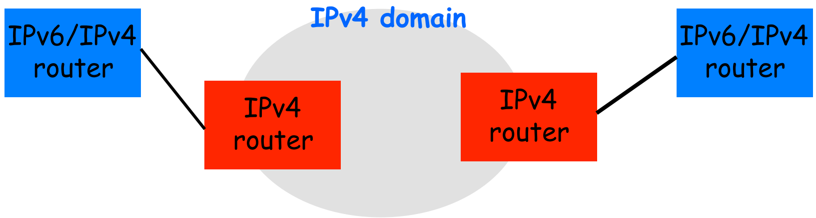


IPv4 mapped address



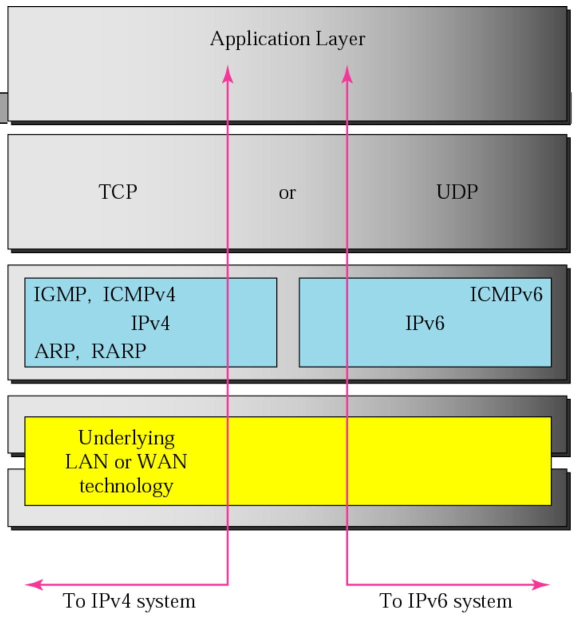
Transition from IPv4 to IPv6

- Transition is progressive
- Problem



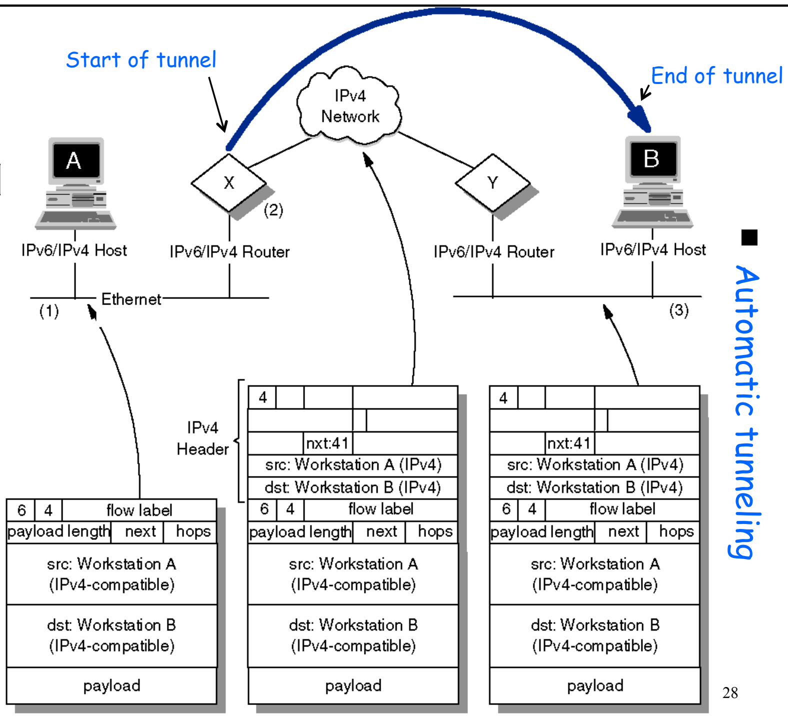
Three transition strategies
- Dual stack
- Tunnelling
- Header translation

Dual stack



Tunneling

- Tunnelling allows an IPv6 packet to transit through one or more IPv4 domains
- In automatic tunnelling (next slide), tunnel endpoints are determined automatically without any explicit configuration
  - Automatic tunneling is triggered when IPv6 addresses are IPv4 compatible

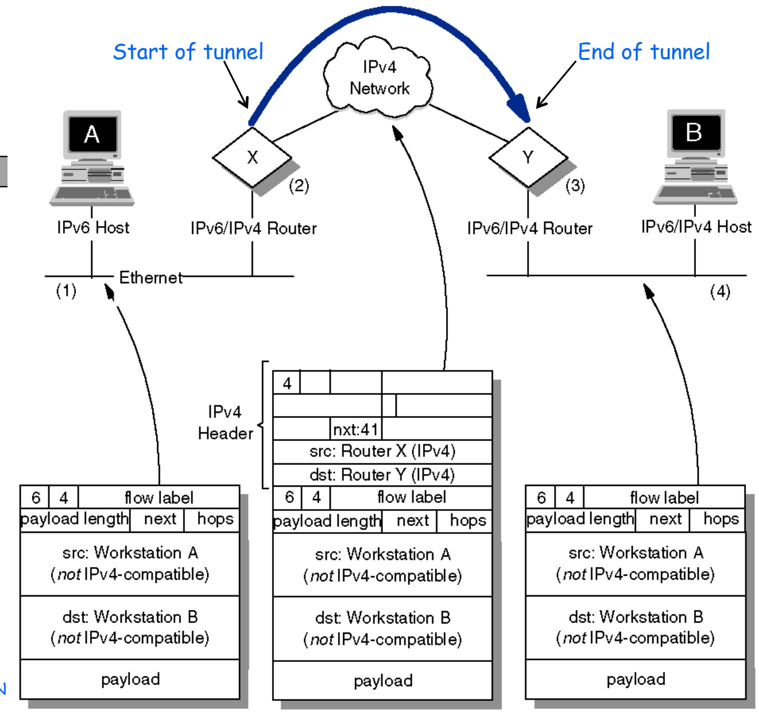


- Automatic tunneling

Tunnelling (2)

- In configured tunneling, tunnel endpoints are configured explicitly
  - by human or by automatic service, eg Tunnel Broker
- Configured tunneling is usually more deterministic
  - easier to debug
  - Recommended for more complex networking environment
- Configured tunnelling (next slide) uses IPv4 mapped addresses

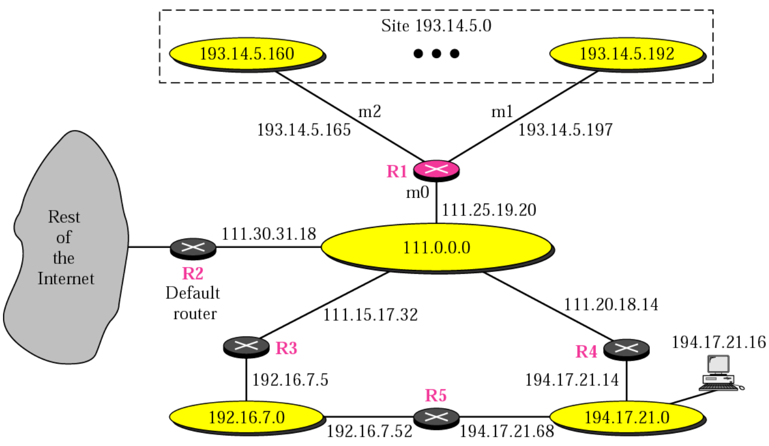
- Configured tunneling



Delivery and routing of IP packets

Key idea: Routing table tells a router how packets are to be delivered

"Our Internet"



Internet and IP addressing

- The Internet is organised into networks
- IP addressing are hierarchical
  - Network id, subnet id, host id
  - Network prefix, host id (CIDR)
- All hosts/router interfaces within a network have the same network prefix
- An IP network is identified by a network number and a network mask

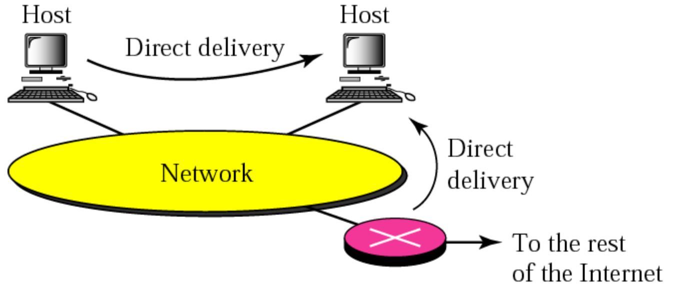
Internet and routing

- Basic function of the Internet
  - To allow any two hosts to talk to each other using IP packets
- Routing enables data packets to find the way through the Internet
- Depending on the locations of the two hosts, the delivery can be
  - Direct, or
  - Indirect

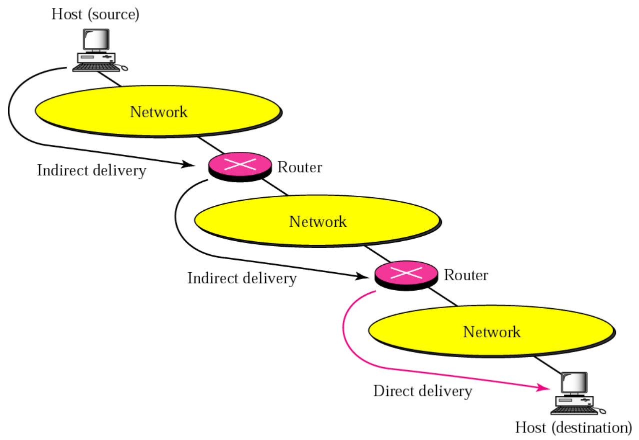
This lecture

- IP version 6 (IPv6)
- Introduction to routing
  - Basic mechanics of IP packet delivery

Direct delivery



Indirect delivery



IP delivery strategies

- IP delivery is primarily network-based
- Host X is to send a packet to host Y
  - Case 1: Hosts X and Y are in the same network
    - » Direct delivery
  - Case 2: They are in different networks
    - » Indirect delivery
    - » The last hop is direct delivery from a router in the destination network to the destination

How do hosts make routing decisions?

- When a host X receives a packet to be delivered to Y
  - Host X checks whether Y is within the same subnet
  - If yes, directly deliver the packet to host Y
  - If no, deliver the packet to the appropriate router
- Two questions
  - How can host X tell whether Y is in the same network?
  - Which is the appropriate router?

Exercise

- Host X with IP address 130.130.10.10 and network masks 255.255.255.128 receives the following two packets:
  - Packet A destined for 130.130.10.56
  - Packet B destined for 130.130.10.156
- Q: Is 255.255.255.128 the subnet mask of 130.130.10.56?
- Determine whether they will be delivered directly or indirectly.

Solution – Method 1

- Host IP address is 130.130.10.10
- Subnet mask is 255.255.255.128
- Network address is 130.130.10.0
- Address range 130.130.10.0 to 130.130.10.127
- Packet A will be delivered directly
- Packet B will be delivered indirectly

Solution – Method 2 (1)

- General setting
  - Given
    - » Host X with IP address IPX and subnet mask MX
      ⇒ network id of X = IPX & MX
    - » Destination host Y with address IPY
  - If network id of Y = network id of X, then X and Y are in the same network; otherwise no
  - Problem: Can't find network id of Y because we don't know the subnet mask for Y
    - » Note: subnet mask for Y can be different from that of X

Solution – Method 2 (2)

- Method of contradiction
  - A statement is either true for false. If assuming that the statement is true leads to contradiction, then the statement must be false.

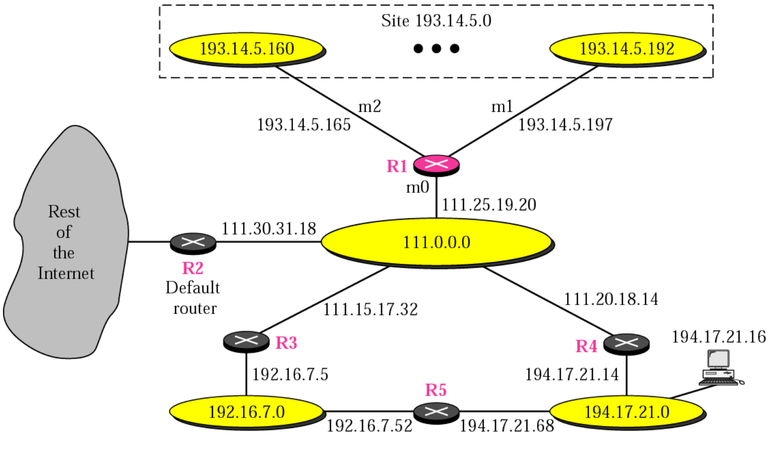
Solution – Method 2 (3)

- Method
  - Assume "Host X and Y are in the same network" is true
  - Since all hosts in a subnet has the same subnet mask, the assumption implies MX is also the subnet mask of Y
    $\Rightarrow$ network id of Y = IPY & MX
  - Since all hosts in a subnet has the same network number, the assumption also implies
    - » IPY & MX = network id of X (Eqn)
    - » If (Eqn) is false $\Rightarrow$ contradiction $\Rightarrow$ assumption false
    - » Otherwise assumption true

Solution – Method 2 (4)

- Host X is in network 130.130.10.0 with mask 255.255.255.128
- Packet A
  - 130.130.10.56 AND 255.255.255.128 = 130.130.10.0 ⇒ direct delivery
- Packet B
  - 130.130.10.156 AND 255.255.255.128 = 130.130.10.128 ⇒ indirect delivery

Our Internet – which router to use for indirect delivery?



Routing table

- In case of indirect delivery, a host looks up a routing table to determine which router to use
  - Most networks have only one router (known as the default router) – it is not necessary to maintain a routing table in this case
- A router also uses a routing table to determine how a packet is to be delivered

How is routing table organized?

- Main issue: size of the routing table must be manageable
  - Cost: A larger routing table needs more memory
  - Performance: It takes longer to search a large routing table
- Different techniques
  - Next-hop based versus route-based
  - Network-based versus host-based
  - Host-specific routing
  - Default routing
- Why different techniques?

Next-hop routing versus route-based routing

Routing table for host A

Destination	Route
Host B	R1, R2, Host B

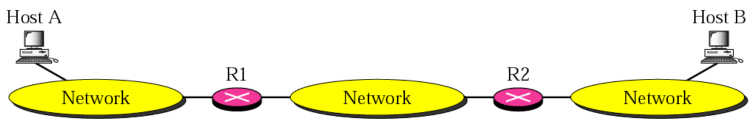
Routing table for R1

Destination	Route
Host B	R2, Host B

Routing table for R2

Destination	Route
Host B	Host B

a. Routing tables based on route



Routing table for host A

Destination	Next Hop
Host B	R1

Routing table for R1

Destination	Next Hop
Host B	R2

Routing table for R2

Destination	Next Hop
Host B	—

b. Routing tables based on next hop

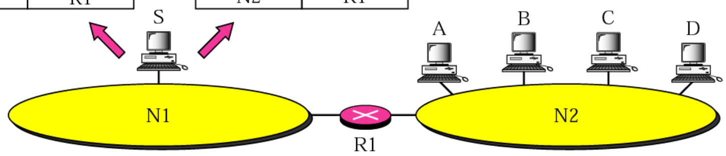
Network-specific versus host-specific

Routing table for host S based on host-specific routing

Destination	Next Hop
A	R1
B	R1
C	R1
D	R1

Routing table for host S based on network-specific routing

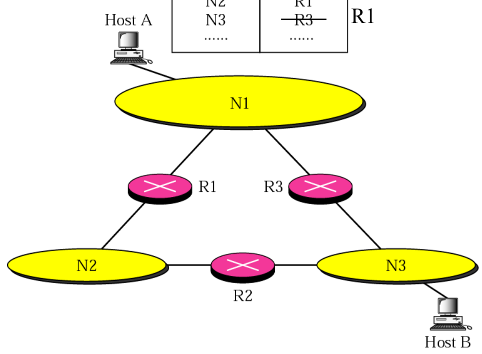
Destination	Next Hop
N2	R1



Host-specific routing

Routing table for host A

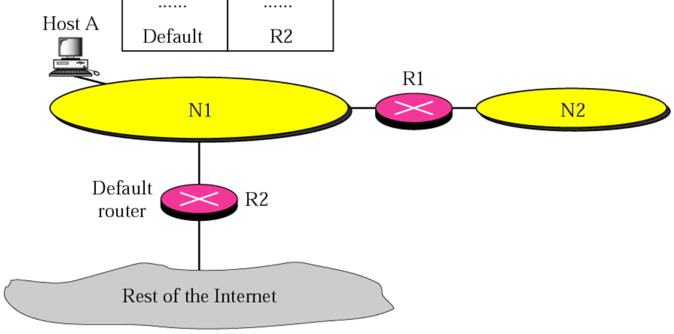
Destination	Next Hop
Host B	R3
N2	R1
N3	~~R3~~ R1
......	......



Default routing

Routing table for host A

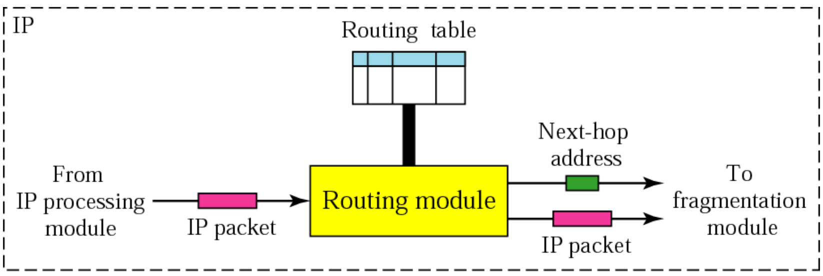
Destination	Next Hop
N2	R1
......	......
Default	R2



Routing table entries

- Routing tables are usually based on next-hop routing
- A routing table may contain these type of entries:
  - Network-specific (the majority)
  - Host-specific
  - Default

Routing module and routing table



Routing table

Mask	Destination address	Next-hop address	Flags	Reference count	Use	Interface
255.0.0.0	124.0.0.0	145.6.7.23	UG	4	20	m2
....	....	....	....	....	....	....
....	....		....	....	....	....

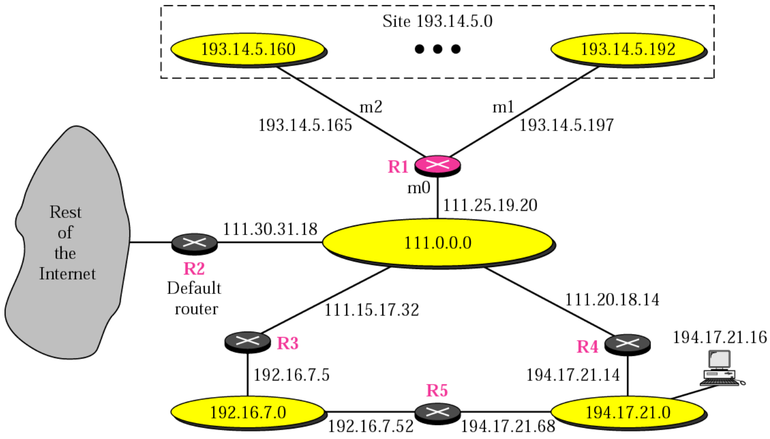
Flags

U The router is up and running.
G The destination is in another network.
H Host-specific address.

Static routing table

- Routing table can be
  - Static
  - Dynamic
- Static routing table are entered manually by the administrator

"Our Internet"



Routing table for R1

Mask	Destination	Next Hop	Interface
255.0.0.0	111.0.0.0	--	m0
255.255.255.224	193.14.5.160	-	m2
255.255.255.224	193.14.5.192	-	m1
255.255.255.255	194.17.21.16	111.20.18.14	m0
255.255.255.0	192.16.7.0	111.15.17.32	m0
255.255.255.0	194.17.21.0	111.20.18.14	m0
0.0.0.0	0.0.0.0	111.30.31.18	m0

Note: The order of the entries is: direct delivery, host-specific, network-specific, and lastly default.

Exercise

- Router R1 receives a packet for the host 192.16.7.14. How will the packet be delivered?
- Given the routing table, the routing module applies the masks row by row until a match is found.
  - A packet for host X
  - A row = (network mask M,network id N)
  - A match means "X & M = N"

Solution

Start matching from the direct delivery part of the routing table (repeated below):

Mask	Destination	Next Hop	Interface
255.0.0.0	111.0.0.0	--	m0
255.255.255.224	193.14.5.160	-	m2
255.255.255.224	193.14.5.192	-	m1

The matching process:

192.16.7.14 & 255.0.0.0 = 192.0.0.0 no match to 111.0.0.0

192.16.7.14 & 255.255.255.224 = 192.16.7.0 no match to 193.14.5.160

192.16.7.14 & 255.255.255.224 = 192.16.7.0 no match to 193.14.5.192

Solution (2)

Since no match has been found, the matching process continues.
The rest of the routing table is repeated below.

Mask	Destination	Next Hop	Interface
255.255.255.255	194.17.21.16	111.20.18.14	m0
255.255.255.0	192.16.7.0	111.15.17.32	m0
255.255.255.0	194.17.21.0	111.20.18.14	m0
0.0.0.0	0.0.0.0	111.30.31.18	m0

The matching process

- Host-specific
  192.16.7.14 & 255.255.255.255 = 192.16.7.14 no match to 194.17.21.16
- Network-specific
  192.16.7.14 & 255.255.255.0 = 192.16.7.0 **match to 192.16.7.0**

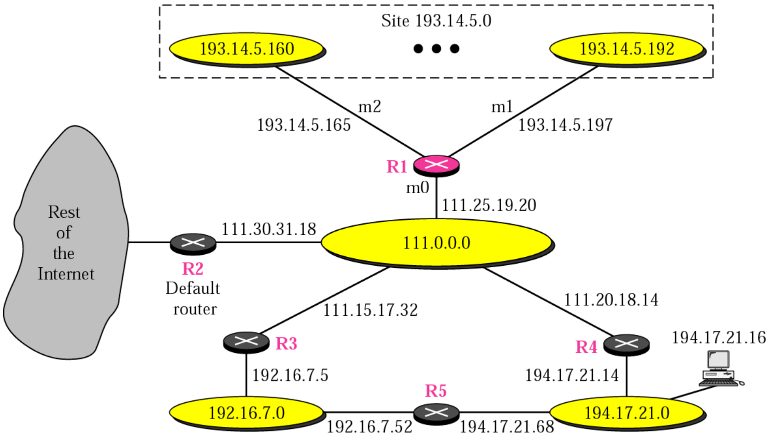
Methods to reduce the size of routing table (1)

- A router needs to know how to route a packet to any host in the Internet
- To reduce the size of routing table, IP addressing is organized so that all hosts within a network have the same network prefix
  - External routers only need to know the network prefix, not individual host addresses in the network

Methods to reduce the size of routing table (2)

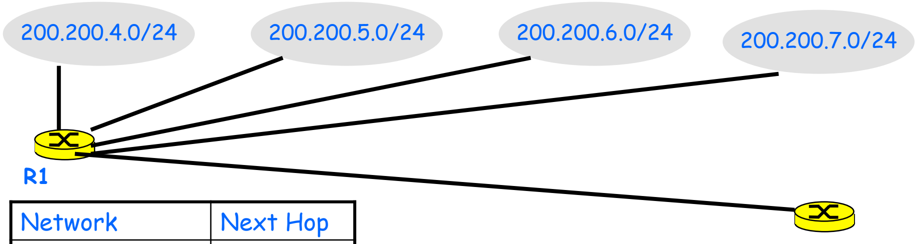
- Routers external to a subnet do not need to know the subnet address
  - Example: R2, R3, R4, R5 in "our Internet" only need to have a routing table entry for 193.14.5.0
  - They don't need to know about 193.14.5.160 etc [Next page]
- Use a default entry to summarise all other routes [Next page]

"Our Internet"



Methods to reduce the size of routing table (3)

With CIDR, routing entries can be aggergated



Network	Next Hop
200.200.4.0/24	
200.200.5.0/24	
200.200.6.0/24	
200.200.7.0/24	

Network	Next Hop
200.200.4.0/22	R1

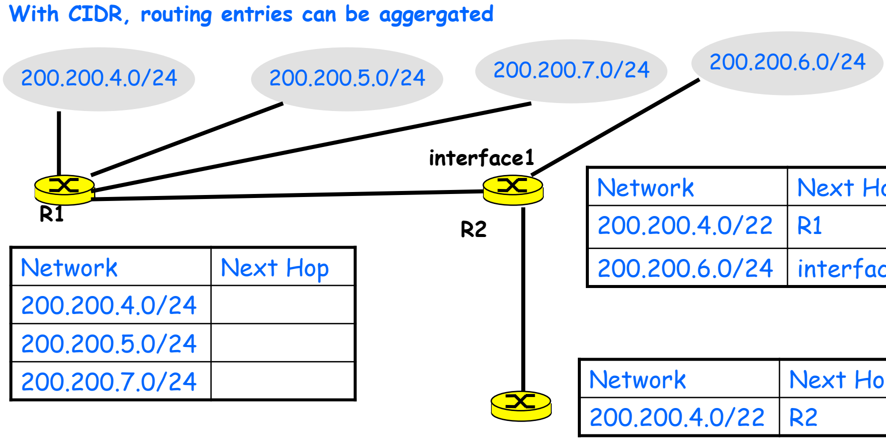
Address aggregation

200.200.4.0/24	1100 1000	1100 1000	0000 0100	0000 0000
200.200.5.0/24	1100 1000	1100 1000	0000 0101	0000 0000
200.200.6.0/24	1100 1000	1100 1000	0000 0110	0000 0000
200.200.7.0/24	1100 1000	1100 1000	0000 0101	0000 0000

Since the first 22 bits are identical, these networks can be aggregated (summarised) as 200.200.4.0/22

Methods to reduce the size of routing table (4)

With CIDR, routing entries can be aggergated



Network	Next Hop
200.200.4.0/24	
200.200.5.0/24	
200.200.7.0/24	

Network	Next Hop
200.200.4.0/22	R1
200.200.6.0/24	interface1

Network	Next Hop
200.200.4.0/22	R2

Address aggregation

200.200.4.0/24	1100 1000	1100 1000	0000 0100	0000 0000
200.200.5.0/24	1100 1000	1100 1000	0000 0101	0000 0000
200.200.7.0/24	1100 1000	1100 1000	0000 0111	0000 0000

Since the first 22 bits are identical, these 3 networks can be aggregated as 200.200.4.0/22

Exercise

- A packet destined for 200.200.6.32 arrives at router R2 which has the following routing table

Network	Next Hop
200.200.4.0/22	R1
200.200.6.0/24	interface1

Q1: Does the address match 200.200.4.0/22?

Q2: Does the address match 200.200.6.0/24?

Solution

- Q1: 200.200.6.32 ?matches? 200.200.4.0/22
  - First 22 bits of 200.200.6.32 ?=? First 22 bits of 200.200.4.0
  - First 22 bits = First 2 bytes + Next 6 bits
  - First 2 bytes certainly match
  - 3rd byte of 200.200.6.32 = 0000 0110
  - 3rd byte of 200.200.4.0 = 0000 0100
  - Yes. A match.

Solution (cont'd)

- Q1: 200.200.6.32 ?matches? 200.200.6.0/24
  - First 24 bits of 200.200.6.32 ?=? First 24 bits of 200.200.6.0
  - Yes.
- Question
  - The IP address matches 2 entries, how should the packet be delivered?

Longest prefix match

- If CIDR address aggregation is used, an IP address may match more than 1 entry in the routing table
- In this case, the match that has the longest prefix length should be chosen
  - "Longest prefix match"
  - E.g. In the example earlier, 200.200.6.0/24 should be chosen instead of 200.200.4.0/22 because the former has a longer prefix length (24) than the latter (22)

References

- Private addresses and NAT
  - IBM Redbook Section 21.4
- IPv6
  - IBM Redbook, Sections 17.3.2, 17.7
- Routing and delivery of IP packet
  - Forouzan Chapter 6