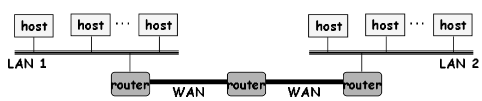
CS 640: Introduction to Computer Networks

Aditya Akella

Lecture 7 -
IP: Addressing and Forwarding

What is an Internetwork?

- Multiple incompatible LANs can be physically connected by specialized computers called *routers*
- The connected networks are called an *internetworks*
  - The *Internet* can be viewed as an internetwork of internetworks

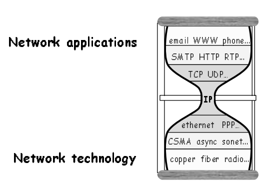


LAN 1 and LAN 2 might be completely different, totally incompatible LANs (e.g., Ethernet and ATM)

2

Internet Protocol (IP)

- Hour Glass Model
  - Create abstraction layer that hides underlying technology from network application software
  - Make as minimal as possible
  - Allows range of current & future technologies
  - Can support many different types of applications



3

Designing an Internetwork

- How do I designate a distant host?
  - Addressing
- How do I send information to a distant host?
  - Underlying service model
    - What gets sent?
    - How fast will it go?
    - What happens if it doesn't get there?
  - Routing/Forwarding: What path is it sent on?
- Challenges
  - Heterogeneity
    - Assembly from variety of different networks
  - Scalability
    - Ensure ability to grow to worldwide scale

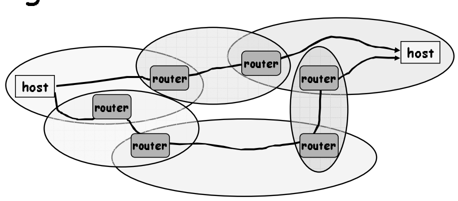
4

The Road Ahead

- Methods for packet forwarding
- Traditional IP addressing
- CIDR IP addressing
- Forwarding examples

5

Logical Structure of Internet



- Ad hoc interconnection of internetworks, owned by different organizations called ISPs.
  - No particular topology
  - Vastly different router & link capacities
- Send packets from source to destination by hopping through networks
  - Router forms bridge from one network to another
  - Different packets may take different routes

6

Approaches to Forwarding Packets

Forwarding: which path to send a packet on? Choices arise both at Layer 2 and Layer 3, but we will discuss in the context of Layer 3.

1. Table of global addresses – "packet switching"
   - Routers keep next hop for destination
   - Packets carry destination address
   - Very common
2. Source routing
   - Packet carries path
3. Table of virtual circuits – "virtual circuit switching"
   - Connection routed through network to setup state
   - Packets forwarded using connection state

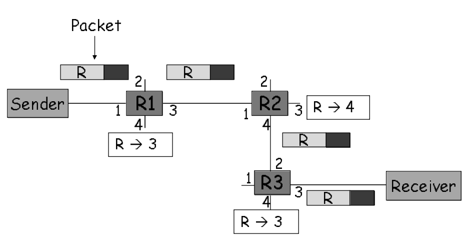
7

Global Addresses

- Each packet has destination address
- Each router has forwarding table of (destination → next hop)
  - Routing table is static – does not change with flows (cf. VCs)
- Distributed routing algorithm for calculating forwarding tables
  - Next class

8

Global Address Example



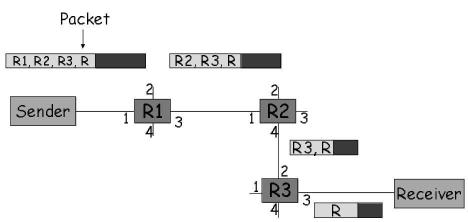
9

Source Routing

- List entire path in packet
- Router processing
  - Strip first step from packet
  - Examine next step in directions
  - Forward to next step

10

Source Routing Example



11

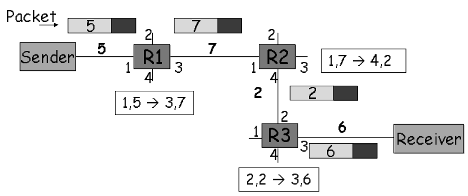
Simplified Virtual Circuits

- Connection-oriented packet-switching
  - Telephone: connection-oriented circuit-switching
- Connection setup phase
  - Use other means to route setup request
  - Each router allocates flow ID on local link
    - Local significance
    - Currently unused on link
  - Set up connection state
- Each packet carries connection ID
  - Sent from source with 1st hop connection ID
- Router processing
  - Lookup flow ID – simple table lookup
  - Replace flow ID with outgoing flow ID
  - Forward to output port

12

Virtual Circuits Example

- Network picks a path
- Assigns VC numbers for flow on each link
- Populates forwarding table



13

Source Routing

- Advantages
  - Switches can be very simple and fast
- Disadvantages
  - Variable (unbounded) header size
  - Sources must know or discover topology
    - Must also deal with failures
- Typical uses
  - Ad-hoc wireless networks
  - Loose source routing in overlays

14

Virtual Circuits

- Advantages
  - Source knows route exists and receiver willing to receive
  - Efficient lookup (simple table lookup)
  - More flexible (different path for each flow)
  - Can reserve bandwidth at connection setup
  - Easier for hardware implementations
- Disadvantages
  - Still need to route connection setup request
  - More complex failure recovery – must recreate connection state
- Typical use → fast router/switch implementations
  - ATM – combined with fix sized cells
  - MPLS – tag switching for IP networks

15

Global Addresses

- Advantages
  - No per connection state (per source/flow)
  - Aggregation → also helps
    - Scalability
- Disadvantages
  - Routers must know routes even for inactive destinations
    - Potentially large tables
  - All packets to destination take same route
    - Little flexibility
  - Need routing protocol to fill table
    - Complex distributed process

16

Comparison

	Source Routing	Global Addresses	Virtual Circuits
Header Size	Worst	OK – Large address	Best
Router Table Size	None	Number of hosts (prefixes)	Number of circuits
Forward Overhead	Best	Prefix matching	Pretty Good
Setup Overhead	None	None	Connection Setup
Error Recovery	Tell all hosts	Tell all routers	Tell all routers and Tear down circuits and re-route

17

Router Table Size

- One entry for every host on the Internet?
  - 300M entries, doubling every 18 months
- One entry for every LAN?
  - Every host on LAN shares prefix
  - Still too many and growing quickly
- One entry for every organization? Better…
  - Every host in organization shares prefix
  - Requires careful address allocation

18

Addressing in IP: Considerations

- Hierarchical vs. flat
  - Wisconsin / Madison / UW-Campus / Aditya
    vs.
    Aditya:123-45-6789
  - Ethernet addresses are flat
- What information would routers need to route to Ethernet addresses?
  - Hierarchical structure crucial for designing scalable binding from interface name to route
  - Route to a general area, then to a specific location
- What type of Hierarchy?
  - How many levels?
  - Same hierarchy depth for everyone?
- Address broken in segments of increasing specificity
  - Uniform for everybody: needs centralized management
  - Non-uniform: more flexible, needs careful decentralized management

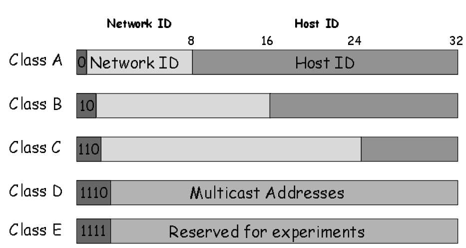
19

IP Addresses

- Fixed length: 32 bits
- Total IP address size: 4 billion
- Initial class-ful structure (1981)
  - Class A: 128 networks, 16M hosts
  - Class B: 16K networks, 64K hosts
  - Class C: 2M networks, 256 hosts

20

IP Address Classes
(Some are Obsolete)



21

Original IP Route Lookup

- Address would specify prefix for forwarding table
  - Simple lookup
- www.cmu.edu address 128.2.11.43
  - Class B address – class + network is 128.2
  - Lookup 128.2 in forwarding table
  - Prefix – part of address that really matters for routing
- Forwarding table contains
  - List of class+network entries
  - A few fixed prefix lengths (8/16/24)
- Large tables
  - 2 Million class C networks

22

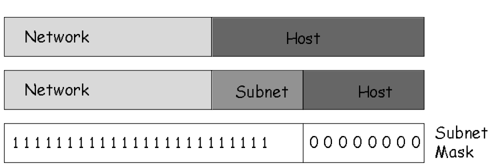
Subnet Addressing: RFC917 (1984)

- Original goal: network part would uniquely identify a single physical network
- Inefficient address space usage
  - Class A & B networks too big
    - Also, very few LANs have close to 64K hosts
    - Easy for networks to (claim to) outgrow class-C
  - Each physical network must have one network number
- Routing table size is too high
- Need simple way to reduce the number of network numbers assigned
  - Subnetting: Split up single network address ranges
  - Fizes routing table size problem, partially

23

Subnetting

- Add another "floating" layer to hierarchy
- Variable length subnet masks
  - Could subnet a class B into several chunks



24

Subnetting Example

- Assume an organization was assigned address 150.100 (class B)
- Assume < 100 hosts per subnet (department)
- How many host bits do we need?
  - Seven
- What is the network mask?
  - 11111111 11111111 11111111 10000000
  - 255.255.255.128

25

Forwarding Example

- Host configured with IP adress and subnet mask
- Subnet number = IP (AND) Mask
- (Subnet number, subnet mask) → Outgoing I/F

```
D = destination IP address
For each forwarding table entry (SN, SM → OI)
  D1 = SM & D
  if (D1 == SN)
    Deliver on OI
  Else
    Forward to default router
```

26

Inefficient Address Usage

- Address space depletion
  - In danger of running out of classes A and B
  - Why?
    - Class C too small for most domains
    - Very few class A – very careful about giving them out
    - Class B poses greatest problem
  - Class B sparsely populated
    - But people refuse to give it back

27

Classless Inter-Domain Routing (CIDR) – RFC1338

- Allows arbitrary split between network & host part of address
  - Do not use classes to determine network ID
  - Use common part of address as network number
  - Allows handing out arbitrary sized chunks of address space
  - E.g., addresses 192.4.16 - 192.4.31 have the first 20 bits in common. Thus, we use these 20 bits as the network number → 192.4.16/20
- Enables more efficient usage of address space (and router tables)
  - Use single entry for range in forwarding tables
  - Combine forwarding entries when possible

28

CIDR Example

- Network is allocated 8 contiguous chunks of 256-host addresses 200.10.0.0 to 200.10.7.255
  - Allocation uses 3 bits of class C space
  - Remaining 20 bits are network number, written as 201.10.0.0/21
- Replaces 8 class C routing entries with 1 combined entry
  - Routing protocols carry prefix with destination network address

29

IP Addresses: How to Get One?

Network (network portion):

- Get allocated portion of ISP's address space:

ISP's block	11001000	00010111	00010000	00000000	200.23.16.0/20
Organization 0	11001000	00010111	00010000	00000000	200.23.16.0/23
Organization 1	11001000	00010111	00010010	00000000	200.23.18.0/23
Organization 2	11001000	00010111	00010100	00000000	200.23.20.0/23
...		…..		….	….
Organization 7	11001000	00010111	00011110	00000000	200.23.30.0/23

30

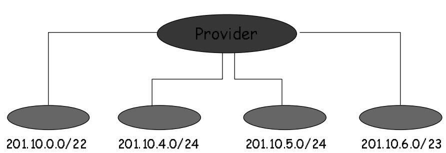
IP Addresses: How to Get One?

- How does an ISP get block of addresses?
  - From **Regional Internet Registries** (RIRs)
    - ARIN (North America, Southern Africa), APNIC (Asia-Pacific), RIPE (Europe, Northern Africa), LACNIC (South America)
- How about a single host?
  - Hard-coded by system admin in a file
  - DHCP: Dynamic Host Configuration Protocol: dynamically get address: "plug-and-play"
    - Host broadcasts "DHCP discover" msg
    - DHCP server responds with "DHCP offer" msg
    - Host requests IP address: "DHCP request" msg
    - DHCP server sends address: "DHCP ack" msg

31

Back to CIDR

Provider is given 201.10.0.0/21



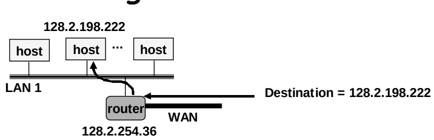
CIDR implications:

Longest prefix match

Route aggregation

32

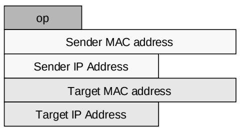
Finding a Local Machine



- Routing Gets Packet to Correct Local Network
  - Based on IP address
  - Router sees that destination address is of local machine
- Still Need to Get Packet to Host
  - Using link-layer protocol
  - Need to know hardware address
- Same Issue for Any Local Communication
  - Find local machine, given its IP address

33

Address Resolution Protocol (ARP)

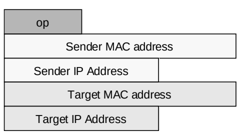


- op: Operation
  - 1: request
  - 2: reply
- Sender
  - Host sending ARP message
- Target
  - Intended receiver of message

  - Diagrammed for Ethernet (6-byte MAC addresses)
- Low-Level Protocol
  - Operates only within local network
  - Determines mapping from IP address to hardware (MAC) address
  - Mapping determined dynamically
    - No need to statically configure tables
    - Only requirement is that each host know its own IP address

34

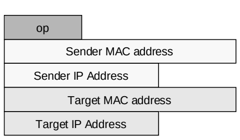
ARP Request



- op: Operation
  - 1: request
- Sender
  - Host that wants to determine MAC address of another machine
- Target
  - Other machine
- Requestor
  - Fills in own IP and MAC address as "sender"
    - Why include its MAC address?
- Mapping
  - Fills desired host IP address in target IP address
- Sending
  - Send to MAC address `ff:ff:ff:ff:ff:ff`
    - Ethernet broadcast

35

ARP Reply



- op: Operation
  - 2: reply
- Sender
  - Host with desired IP address
- Target
  - Original requestor
- Responder becomes "sender"
  - Fill in own IP and MAC address
  - Set requestor as target
  - Send to requestor's MAC address

36

Host Routing Table Example

Destination	Gateway	Genmask	Iface
128.2.209.100	0.0.0.0	255.255.255.255	eth0
128.2.0.0	0.0.0.0	255.255.0.0	eth0
127.0.0.0	0.0.0.0	255.0.0.0	lo
0.0.0.0	128.2.254.36	0.0.0.0	eth0

- Host 128.2.209.100 when plugged into CS ethernet
- Dest 128.2.209.100 → routing to same machine
- Dest 128.2.0.0 → other hosts on same ethernet
- Dest 127.0.0.0 → special loopback address
- Dest 0.0.0.0 → default route to rest of Internet
  - Main CS router: gigrouter.net.cs.cmu.edu (128.2.254.36)

37

The Next Lecture

- IP packet structure
  - Fragmentaiton
- Routers and route lookup

38