Last Course Review
Key Network-Layer Functions

- **forwarding**: move packets from router’s input to appropriate router output
- **routing**: determine route taken by packets from source to dest.
- **Routing algorithms**

**Datagram networks**

- no call setup at network layer
- routers: no state about end-to-end connections
  - no network-level concept of “connection”
- packets forwarded using destination host address
  - packets between some source-dest pair may take different paths

**Longest prefix matching**

<table>
<thead>
<tr>
<th>Prefix Match</th>
<th>Link Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>11001000 00010111 00010110 10100001</td>
<td>1</td>
</tr>
</tbody>
</table>

**Chapter 4: Network Layer**

4.1 Introduction
4.2 Virtual circuit and datagram networks
4.3 What’s inside a router
4.4 IP: Internet Protocol
  - Datagram format
  - IPv4 addressing
  - ICMP
  - IPv6
4.5 Routing algorithms
  - Link state
  - Distance Vector
  - Hierarchical routing
4.6 Routing in the Internet
  - RIP
  - OSPF
  - BGP
4.7 Broadcast and multicast routing

**Interplay between routing and forwarding**

**Datagram or VC network: why?**

**Internet**
- data exchange among computers
  - “elastic” service, no strict timing requirement
- “smart” end systems (computers)
  - can adapt, perform control, error recovery
  - simple inside network, complexity at “edge”
- many link types
  - different characteristics
  - uniform service difficult

**ATM**
- evolved from telephony
  - human conversation
  - strict timing, reliability requirements
  - need for guaranteed service
- “dumb” end systems
  - telephones
  - complexity inside network
Chapter 4: Network Layer

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What’s inside a router
- 4.4 IP: Internet Protocol
  - Datagram format
  - IPv4 addressing
  - ICMP
  - IPv6
- 4.5 Routing algorithms
  - Link state
  - Distance Vector
  - Hierarchical routing
- 4.6 Routing in the Internet
  - RIP
  - OSPF
  - BGP
- 4.7 Broadcast and multicast routing

Router Architecture Overview

Two key router functions:
- run routing algorithms/protocol (RIP, OSPF, BGP)
- forwarding datagrams from incoming to outgoing link

Input Port Functions

Physical layer: bit-level reception
Data link layer: e.g., Ethernet see chapter 5

Decentralized switching:
- given datagram dest, lookup output port using forwarding table in input port memory
- goal: complete input port processing at line speed
- queuing: if datagrams arrive faster than forwarding rate into switch fabric

Three types of switching fabrics

Switching Via Memory

First generation routers:
- traditional computers with switching under direct control of CPU
- packet copied to system’s memory
- speed limited by memory bandwidth (2 bus crossings per datagram)

Switching Via a Bus

- datagram from input port memory to output port memory via a shared bus
- bus contention: switching speed limited by bus bandwidth
- 1 Gbps bus, Cisco 1900: sufficient speed for access and enterprise routers (not regional or backbone)
Switching Via An Interconnection Network

- overcome bus bandwidth limitations
- Banyan networks, other interconnection nets initially developed to connect processors in multiprocessor
- Advanced design: fragmenting datagram into fixed length cells, switch cells through the fabric.
- Cisco 12000: switches Gbps through the interconnection network

Output Ports

- Buffering required when datagrams arrive from fabric faster than the transmission rate
- Scheduling discipline chooses among queued datagrams for transmission

Output port queueing

- buffering when arrival rate via switch exceeds output line speed
- queuing (delay) and loss due to output port buffer overflow!

Input Port Queuing

- Fabric slower than input ports combined -> queuing may occur at input queues
- Head-of-the-Line (HOL) blocking: queued datagram at front of queue prevents others in queue from moving forward
- queuing delay and loss due to input buffer overflow!

Chapter 4: Network Layer

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
  - Datagram format
  - IPv4 addressing
  - ICMP
  - IPv6
- 4.5 Routing algorithms
  - Link state
  - Distance Vector
  - Hierarchical routing
- 4.6 Routing in the Internet
  - RIP
  - OSPF
  - BGP
- 4.7 Broadcast and multicast routing

The Internet Network layer

Host, router network layer functions:

- Transport layer: TCP, UDP
- IP protocol
  - addressing conventions
  - datagram format
  - packet handling conventions
- ICMP protocol
  - error reporting
- routing signaling

Network layer

Link layer

physical layer
Chapter 4: Network Layer

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What’s inside a router
- 4.4 IP: Internet Protocol
  - Datagram format
  - IPv4 addressing
  - ICMP
  - IPv6
- 4.5 Routing algorithms
  - Link state
  - Distance Vector
  - Hierarchical routing
- 4.6 Routing in the Internet
  - RIP
  - OSPF
  - BGP
  - 4.7 Broadcast and multicast routing

IP datagram format

- Header length (bytes) "type" of data
- Max number of hops (decimated at each router)
- Time To Live (TTL)
- Protocol
- Source and Destination IP Address
- Options (if any)

E.g. timestamp, record route taken, specify list of routers to visit.

IP Fragmentation and Reassembly

- Network links have MTU (max transfer size) - largest possible link-level frame.
  - Different link types, different MTUs
  - Large IP datagram divided ("fragmented") within net
    - One datagram becomes several datagrams
    - "reassembled" only at final destination
    - IP header bits used to identify, order related fragments

Example

- 4000 byte datagram
- MTU = 1500 bytes

1480 bytes in data field

IP Addressing: introduction

- IP address: 32-bit identifier for host, router interface
- Interface: connection between host/router and physical link
- Interface’s typically have multiple interfaces
- Host may have multiple interfaces
- IP addresses associated with each interface

IP addresses: 223.1.1.1 - 223.1.1.4 (IPv4 addresses)
**Subnets**

- **IP address:**
  - subnet part (high order bits)
  - host part (low order bits)

- **What’s a subnet?**
  - device interfaces with same subnet part of IP address
  - can physically reach each other without intervening router

**Recipe**

- To determine the subnets, detach each interface from its host or router, creating islands of isolated networks. Each isolated network is called a subnet.

**IP addressing: CIDR**

**CIDR: Classless InterDomain Routing**

- subnet portion of address of arbitrary length
- address format: `a.b.c.d/x`, where `x` is # bits in subnet portion of address

**IP addresses: how to get one?**

- **Q:** How does host get IP address?
  - hard-coded by system admin in a file
  - Wintel: control-panel-network->configuration->tcp/ip->properties
  - UNIX: `/etc/rc.config`
  - DHCP: Dynamic Host Configuration Protocol: dynamically get address from as server
  - “plug-and-play” (more in next chapter)

- **Q:** How does network get subnet part of IP addr?
  - gets allocated portion of its provider ISP’s address space
Hierarchical addressing: route aggregation

Hierarchical addressing allows efficient advertisement of routing information:

- Send me anything with addresses beginning 200.23.16.0/20
- Send me anything with addresses beginning 199.31.0.0/16

ISPs-R-Us
Organization 0
Organization 1
Organization 2
Organization 7

Fly-By-Night-ISP
Internet

Hierarchical addressing: more specific routes

ISPs-R-Us has a more specific route to Organization 1

- Send me anything with addresses beginning 200.23.16.0/23
- Send me anything with addresses beginning 199.31.0.0/16

IP addressing: the last word...

Q: How does an ISP get block of addresses?
A: ICANN: Internet Corporation for Assigned Names and Numbers
- allocates addresses
- manages DNS
- assigns domain names, resolves disputes

NAT: Network Address Translation

Implementation: NAT router must:
- outgoing datagrams: replace (source IP address, port #) of every outgoing datagram to (NAT IP address, new port #)
  - ... remote clients/servers will respond using (NAT IP address, new port #) as destination addr.
- remember (in NAT translation table) every (source IP address, port #) to (NAT IP address, new port #) translation pair
- incoming datagrams: replace (NAT IP address, new port #) in dest fields of every incoming datagram with corresponding (source IP address, port #) stored in NAT table
Chapter 4: Network Layer

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What’s inside a router
- 4.4 IP: Internet Protocol
  - Datagram format
  - IPv4 addressing
  - ICMP
  - IPv6
- 4.5 Routing algorithms
  - Link state
  - Distance Vector
  - Hierarchical routing
- 4.6 Routing in the Internet
  - RIP
  - OSPF
  - BGP
- 4.7 Broadcast and multicast routing

ICMP: Internet Control Message Protocol

- used by hosts & routers to communicate network-level information
- error reporting: unreachable host, network, port, protocol
- echo request/reply (used by ping)
- network-layer "above" IP:
  - ICMP msg carried in IP datagram
  - ICMP message: type, code plus first 8 bytes of IP datagram causing error

Traceroute and ICMP

- Source sends series of UDP segments to dest
  - First has TTL=1
  - Second has TTL=2, etc.
  - Unlikely port number
- When nth datagram arrives to nth router:
  - Router discards datagram
  - And sends to source an ICMP message (type 11, code 0)
- Message includes name of router & IP address
- When ICMP message arrives, source calculates RTT
- Traceroute does this 3 times
- Stopping criterion
  - UDP segment eventually arrives at destination host
  - Destination returns ICMP "host unreachable" packet (type 3, code 3)
  - When source gets this ICMP, stops.
**IPv6**

- Initial motivation: 32-bit address space soon to be completely allocated.
- Additional motivation:
  - header format helps speed processing/forwarding
  - header changes to facilitate QoS
- IPv6 datagram format:
  - fixed-length 40 byte header
  - no fragmentation allowed

**IPv6 Header (Cont)**

- Priority: identify priority among datagrams in flow
- Flow Label: identify datagrams in same "flow" (concept of "flow" not well defined)
- Next header: identify upper layer protocol for data

**Other Changes from IPv4**

- Checksum: removed entirely to reduce processing time at each hop
- Options: allowed, but outside of header, indicated by "Next Header" field
- ICMPv6: new version of ICMP
  - additional message types, e.g. "Packet Too Big"
  - multicast group management functions

**Transition From IPv4 To IPv6**

- Not all routers can be upgraded simultaneously
  - no "flag days"
  - How will the network operate with mixed IPv4 and IPv6 routers?
- Tunneling: IPv6 carried as payload in IPv4 datagram among IPv4 routers

**Tunneling**

Logical view: A IPv6  B IPv6  E IPv6  F IPv6

Physical view:

- Flow X Src: A Dest: F
- Flow X Src: A Dest: F
- Flow X Src: A Dest: F
- Flow X Src: A Dest: F
- Flow X Src: A Dest: F
- Flow X Src: A Dest: F

A-to-B: IPv6
B-to-C: IPv6 inside IPv4
E-to-F: IPv6