**Last Course Review**

**Principles of Congestion Control**

*Congestion:*
- informally: "too many sources sending too much data too fast for network to handle"
- different from flow control
- manifestations:
  - lost packets (buffer overflow at routers)
  - long delays (queueing in router buffers)
- a top-10 problem!

**Approaches towards congestion control**

Two broad approaches towards congestion control:

**End-end congestion control:**
- no explicit feedback from network
- congestion inferred from end-system observed loss, delay
- approach taken by TCP

**Network-assisted congestion control:**
- routers provide feedback to end systems
- single bit indicating congestion (SNA, DECbit, TCP/IP ECN, ATM)
- explicit rate sender should send at

---

**TCP Congestion Control**

- end-end control (no network assistance)
- sender limits transmission:
  \[ \text{LastByteSent} - \text{LastByteAcknowledged} \leq \text{CongWin} \]
- Roughly,
  \[ \text{rate} = \frac{\text{CongWin}}{\text{RTT}} \text{ Bytes/sec} \]
- CongWin is dynamic, function of perceived network congestion

**TCP AIMD**

- multiplicative decrease: cut CongWin in half after loss event
- additive increase: increase CongWin by 1 MSS every RTT in the absence of loss events: probing

---

**TCP Slow Start (more)**

- When connection begins, increase rate exponentially until first loss event:
  - double CongWin every RTT
  - done by incrementing CongWin for every ACK received
- **Summary:** initial rate is slow but ramps up exponentially fast

---

**TCP Fairness**

**Fairness goal:** if K TCP sessions share same bottleneck link of bandwidth R, each should have average rate of R/K
Fairness (more)

- Multimedia apps often do not use TCP
  - do not want rate throttled by congestion control
- Instead use UDP:
  - pump audio/video at constant rate, tolerate packet loss
- Research area: TCP friendly

Chapter 4: Network Layer

Chapter goals:
- understand principles behind network layer services:
  - routing (path selection)
  - dealing with scale
  - how a router works
  - advanced topics: IPv6, mobility
- instantiation and implementation in the Internet

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

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

Interplay between routing and forwarding
Connection setup

- 3rd important function in some network architectures:
  - ATM, frame relay, X.25
- Before datagrams flow, two hosts and intervening routers establish virtual connection
  - Routers get involved
- Network and transport layer connection service:
  - Network: between two hosts
  - Transport: between two processes

Network service model

**Q:** What service model for "channel" transporting datagrams from sender to receiver?

**Example services for individual datagrams:**
- Guaranteed delivery
- Guaranteed delivery with less than 40 msec delay

**Example services for a flow of datagrams:**
- In-order datagram delivery
- Guaranteed minimum bandwidth to flow
- Restrictions on changes in inter-packet spacing

Network layer service models:

<table>
<thead>
<tr>
<th>Network Architecture</th>
<th>Service Model</th>
<th>Guarantees?</th>
<th>Congestion feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internet</td>
<td>best effort</td>
<td>none</td>
<td>no (inferred via loss)</td>
</tr>
<tr>
<td>ATM</td>
<td>CBR</td>
<td>constant rate</td>
<td>yes</td>
</tr>
<tr>
<td>ATM</td>
<td>VBR</td>
<td>guaranteed rate</td>
<td>yes</td>
</tr>
<tr>
<td>ATM</td>
<td>ABR</td>
<td>guaranteed minimum</td>
<td>no</td>
</tr>
<tr>
<td>ATM</td>
<td>UBR</td>
<td>none</td>
<td>no</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
  - RREP
  - OSPF
  - BGP
- 4.7 Broadcast and multicast routing

Network layer connection and connection-less service

- Datagram network provides network-layer connectionless service
- VC network provides network-layer connection service
- Analogous to the transport-layer services, but:
  - Service: host-to-host
  - No choice: network provides one or the other
  - Implementation: in the core

Virtual circuits

- "source-to-dest path behaves much like telephone circuit"
  - performance-wise
  - network actions along source-to-dest path
- call setup, teardown for each call before data can flow
- each packet carries VC identifier (not destination host address)
- every router on source-dest path maintains "state" for each passing connection
- link, router resources (bandwidth, buffers) may be allocated to VC
**VC implementation**

A VC consists of:
1. Path from source to destination
2. VC numbers, one number for each link along path
3. Entries in forwarding tables in routers along path

- Packet belonging to VC carries a VC number.
- VC number must be changed on each link.
  - New VC number comes from forwarding table

---

**Virtual circuits: signaling protocols**

- used to setup, maintain, teardown VC
- used in ATM, frame-relay, X.25
- not used in today’s Internet

---

**Forwarding table**

<table>
<thead>
<tr>
<th>Destination Address Range</th>
<th>Link Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>11001000 00010111 00010000 00000000 through</td>
<td>0</td>
</tr>
<tr>
<td>11001000 00010111 00011001 11111111 through</td>
<td>1</td>
</tr>
<tr>
<td>11001000 00010111 00011110 00000000 through</td>
<td>2</td>
</tr>
<tr>
<td>11001000 00010111 00011111 11111111 otherwise</td>
<td>3</td>
</tr>
</tbody>
</table>

- 4 billion possible entries

---

**Longest prefix matching**

<table>
<thead>
<tr>
<th>Prefix Match</th>
<th>Link Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>11001000 00010111 00011000 10101001</td>
<td>1</td>
</tr>
<tr>
<td>11001000 00010111 00011000 00011111 otherwise</td>
<td>2</td>
</tr>
<tr>
<td>11001000 00010111 00011111 11111111 otherwise</td>
<td>3</td>
</tr>
</tbody>
</table>

- Examples
  - DA: 11001000 00010111 00010110 10100001 Which interface?
  - DA: 11001000 00010111 00011000 10101010 Which interface?
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
  - IP 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
- 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
- Queueing (delay) and loss due to output port buffer overflow!

Input Port Queuing

- Fabric slower than input ports combined → queueing may occur at input queue
- Head-of-the-Line (HOL) blocking: queued datagram at front of queue prevents others in queue from moving forward
- Queueing 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

Routing protocols
• path selection
• RIP, OSPF, BGP

IP protocol
• addressing conventions
• packet handling conventions

ICMP protocol
• error reporting
• router signaling

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

Introduction 1-37

Introduction 1-38