Today’s Lecture

- Switched Networks
  - Switching Concepts
  - Ethernet Switches
    - Learning bridge
    - Spanning tree
    - Multicast
  - Asynchronous Transfer Mode (ATM) Switching
    - Protocol overview
    - Virtual Circuit and Virtual Path
    - ATM Quality of Service (QoS)

- Switched SDN (Software Defined Network) (Sumit)
Intro to Switching

- Build a large network by interconnecting a number of switches
- Easily add new hosts
- Switching Techniques
  - Datagram or connectionless (Ethernet) \(\rightarrow\) SDN
    - Unique address
    - No need to setup connection
  - Virtual circuit or connection-oriented (ATM)
    - Set up connection and maintain connection state
  - Source routing
    - Source specify the whole or partial route to the destination
Ethernet Hub

- Hub is just a repeater
  - Receive signal from one port and broadcast to all other ports
- Extends max distance between nodes, but collisions are propagated
  - Individual segment collision domains become one large collision domain
- Cannot interconnect different LAN technologies, e.g. 10BaseT & 100BaseT
Bridges/LAN switches

- Link layer device
  - stores and forwards frames
  - examines frame header and **selectively** forwards frame based on MAC dest address
  - when frame is to be forwarded on segment, uses the corresponding MAC to access segment (e.g. CSMA/CD for Ethernet)
Bridges/LAN switches (Cont.)

- Interconnect multiple LANs, possibly even support different IEEE 802.x types, e.g. 802.3 and 802.5, 802.11, but NOT 802.x with ATM
Ethernet Hubs vs. Ethernet Switches

• An **Ethernet switch** is a packet switch for Ethernet frames
  - Buffering of frames prevents collisions
  - Each port is isolated and builds its own collision domain
    - Break subnet into LAN segments
  - Host can directly connect to switch, no collision, full duplex

• An **Ethernet Hub** does not perform buffering:
  - Collisions occur if two frames arrive at the same time.
A Switched Enterprise Network
Forwarding

- Which port to forward a frame?
  - Use forwarding database/table
    < MAC address, port, Time-to-Live (TTL)>

- How to build the forwarding table???
  - A routing problem

### Forwarding Table

<table>
<thead>
<tr>
<th>Destination</th>
<th>Port</th>
<th>Time-to-Live (TTL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>C</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>D</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>E</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>F</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>G</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>H</td>
<td>0</td>
<td>5</td>
</tr>
</tbody>
</table>
Transparent Bridges

Three parts to transparent bridges:

(1) Learning of Addresses
(2) Forwarding of Frames
(3) Spanning Tree Algorithm
Self Learning (Learning Bridges)

- Forwarding tables entries are set automatically with a simple heuristic:
  - The source address field of a frame that arrives on a port tells which host is reachable from this port.
  - When a frame received, switch “learns” location of sender
  - records sender/location pair in forwarding table with TTL = MAX_TTL
  - TTL reset to MAX_TTL every time a frame with the same source addr is received to refresh the existing table entry
  - Entry removed when TTL counts down to 0
Frame Forwarding/Filtering

**When switch receives a frame:**

index forwarding table using MAC dest address

\[
\begin{align*}
\text{if} & \quad \text{entry found for destination} \\
& \quad \text{then}\{ \\
& \quad \quad \text{if} \quad \text{dest on the same port from which frame arrived} \\
& \quad \quad \quad \text{then} \quad \text{drop the frame (filtering)} \\
& \quad \quad \quad \text{else} \quad \text{forward the frame on port indicated} \\
& \quad \}\}
\end{align*}
\]

else flood

Forward on all but the port on which the frame arrived
Example

• Consider the following packets: \((\text{Src}=A, \text{Dest}=F), (\text{Src}=C, \text{Dest}=A), (\text{Src}=E, \text{Dest}=C)\)

• What have the bridges learned?
Danger of Loops

- Consider the two LANs that are connected by two bridges.
- Assume *host n* is transmitting a frame F with unknown destination.

**What is happening?**

- Bridges A and B flood the frame to LAN 2.
- Bridge B sees F on LAN 2 (with unknown destination), and copies the frame back to LAN 1.
- Bridge A does the same.
- The copying continues.
Spanning Trees / Transparent Bridges

- A solution is to prevent loops in the topology.
- IEEE 802.1d has an algorithm that organizes the bridges as spanning tree in a dynamic environment. Note: Trees don’t have loops.
- Bridges that run 802.1d are called transparent bridges.
Spanning Tree Protocol (STP)

- Each bridge has a unique ID (MAC addr + priority level)
- Select the bridge with the smallest ID as the root of the spanning tree, called “root bridge”
  - All the ports on the root bridge are active (forwards the frames)
- Each bridge determines the minimum-cost path from itself to the root and nodes which of its port is on the path (root port)
  - **Link cost**: the cost of traversing a single network segment (link)
  - **Path cost**: the sum of the costs of the segments (links) on the path
    - an administrator can configure the cost of traversing a particular segment (link)
    - E.g. set the cost for every segment to 1, the path cost is a count of the number of bridges along the path.
  - **Root path cost**: the cost of the minimum-cost path from this bridge to the root
  - **Root port**: the port connecting to the minimum-cost path on this bridge
  - Breaking ties: When multiple paths from a bridge are min-cost paths, choose the path using the neighbor bridge with the lower bridge ID.
  - If the multiple ports connects this bridge and the neighbor bridge on the root path, choose the port with the lowest port ID as the root port.
Spanning Tree Protocol (Cont.)

- Select a single “designated bridge” and its designated port on each LAN segment
  - **Designated bridge**: the bridge on that LAN segment with the minimum-cost path to the root. Only designated bridge allowed to forward frames to and from this LAN segment.
    - If two or more bridges have the same root path cost, choose the one with the lowest bridge ID
  - **Designated port**: the port connecting the designated bridge to this LAN segment
    - If the designated bridges has two or more ports attached to this LAN, choose the port with the lowest port ID
- Any port that is not a root port or a designated port is blocked.
Spanning Tree Protocol (Cont.)

- Bridges exchange messages to configure the bridge (Configuration Bridge Protocol Data Unit, CBPDUs) to cut the loop and build the tree.
  - Source addr: port MAC addr, Dest. addr: STP multicast address
  - \(<\text{sending bridge ID, root bridge ID, root path cost}>\)

- At the beginning, each bridge considers itself to be the root, sends CBPDU identifying itself as root
- Upon receiving a CBPDU, check if the new path is better
  - if better, update its STP record, forward the message after updating the root path cost in the message
  - After stabilization, only the root bridge generates new CBPDUs regularly, others stops generate CBPDUs once learning it is not a root

- From a non-root port, receives a CBPDU indicating it is not the designated bridge for that segment, goes to blocking state
  - BPDU is still received in blocking state.
Spanning Tree Protocol Example

1. B3 receives (B2,0,B2)
2. Since 2<3, B3 accepts B2 as root
3. B3 adds 1 to distance advertised by B2 and sends (B2,1,B3) to B5
4. Meanwhile B2 accepts B1 as root because it has lower ID and sends (B1,1,B2) to B3
5. B5 accepts B1 as root and sends (B1,1,B5) towards B3
6. B3 accepts B1 as root and notes that both B2 and B5 are closer to the root than it is.
7. B3 stops forwarding messages on both its interfaces
Virtual LAN

- Group the stations in a broadcast domain, regardless of their physical location.
- A VLAN ID (VID) in the frame
- A frame is not forwarded/broadcasted from one VLAN to another VLAN
- Each VLAN establishes its own spanning tree
- Assign a port to one or multiple or all VLANs (static or dynamic)
ATM Introduction

• 1990’s standards for high-speed (155Mbps to 622 Mbps and higher) Broadband Integrated Service Digital Network (BISDN) architecture
• Goal: integrated, end-end transport of carry voice, video, data
• meeting timing/quality of service (QoS) requirements of voice, video (versus Internet best-effort model)
• “next generation” telephony: technical roots in telephone world
• packet-switching (fixed length packets, called “cells”) using virtual circuits
ATM Vision

The Ultimate Integrated Services Network

- ATM network moves cells (fixed length packets) with low delay and low delay variation at high speeds
- Devices at ends translate (e.g., segment and reassemble) between cells and original traffic
ATM System Architecture

Diagram illustrating the ATM System Architecture with layers and components.
The ATM Cell

- Small Size (low delay, but high overhead)
  - 5 Byte Header
  - 48 Byte Payload
- Fixed Size (easy switch implementation, but padding overhead)
- Header contains virtual circuit information
- Payload can be voice, video or other data types
Packetization Delay Advantage of Small Cells

Percent Overhead and Packetization Delay for 64 Kbps Voice

![Graph showing the relationship between overhead and payload bytes]

- **% Overhead**
- **Delay (ms)**

Payload (Bytes) vs. % Overhead and Delay (ms)
ATM Adaptation Layer (AAL)

- Only at edge of ATM network (end system)
- Roughly analogous to Internet transport layer
- Provides mapping of applications (IP or native ATM applications) to ATM service of the same type
- Segments/Reassembles into 48 Payloads
- Hands 48 Byte Payloads To ATM Layer

### AAL Types

1. Circuit Emulation
   - Constant Bit Rate (CBR)
2. Low Bit Rate Voice (Real Time)
   - Variable Bit Rate (VBR)
3/4. Time Invariant Data
5. "Simple" Data
ATM Layer

- Adds/Removes Header To 48 Byte Payload
  - Header Contains Connection Identifier, multiplexes 53 Byte cells into virtual connections,
- ATM’s “Network” layer
  - Transport cells across ATM network (analogous to IP network layer, but very different strategy and services than IP network layer)
  - Signaling, cell switching, routing
**ATM Interfaces**

UNI User Network Interface
NNI Network Node Interface
B-ICI BISDN Inter-Carrier Interface
ATM DXI Data eXchange Interface
FUNI ATM Frame Based UNI Interface
ATM UNI Cell

CLP = Cell Loss Priority

5 Byte Header

48 Byte Payload

CLP = Cell Loss Priority
ATM NNI Cell

- **Virtual Path Identifier**
- **Virtual Channel Identifier**
- **Payload Type Identifier**
- **CLP**
- **Header Error Check**
- **Payload (48 bytes)**

**CLP = Cell Loss Priority**

- **5 Byte Header**
- **48 Byte Payload**
• Bundles of Virtual Channels are switched via Virtual Paths
  – Better scalability (i.e. more capable of growing to large numbers of circuits)
ATM VCs

• Advantages of ATM VC approach:
  – QoS performance guarantee for connection mapped to VC (bandwidth, delay, delay jitter)

• Drawbacks of ATM VC approach:
  – Inefficient support of datagram traffic
  – one PVC between each source/dest pair) does not scale (N*2 connections needed)
  – SVC introduces call setup latency, processing overhead for short lived connections
Switched Virtual Circuits

- Switch and terminal exchange signalling messages using the predefined signalling channel, VPI/VCI = 0/5
Permanent Virtual Circuits

- Long setup time (especially with human intervention) means that connections are left active for long periods of time e.g., days, weeks
- VPI/VCI tables setup in terminals and switches
Call Control Signalling

Call control protocol is used to establish, maintain, and clear virtual channel connections between a user and network.
A wants to communicate with B

- Setup message
  - Call reference
  - Called party address
  - Calling party address
  - Traffic characteristics
  - Quality of service

- Call proceeding message
  - Call reference
  - VPI/VCI
Setting Up a Call - 2

- Internal network processing
  - Resource availability checking
  - Virtual channel or path routing
  - Function of the Network Node Interface (NNI)
Setting Up a Call - 3

- Setup message
  - Call reference
  - Called party address
  - Calling party address
  - Traffic characteristics
  - Quality of service
  - VPI/VCI

Call Proceeding

Call Proceeding

Call Proceeding

Call Proceeding

Call Proceeding

Called user deciding to accept call

Call reference

Call Proceeding
Setting Up a Call - 4

- Connect message
  - Call reference
  - Indicates call acceptance
- Connect Acknowledge
  - Call reference
• Calling party informed that call is available for user information exchange
Bandwidth Negotiation

UNI  NNI  UNI

Setup (20 Mb/s)  Setup (15 Mb/s)  Setup (10 Mb/s)

Connect (10 Mb/s)  Connect (10 Mb/s)  Connect (10 Mb/s)
NNI

- Supports $2^{12}$ Virtual Paths
- Supports virtual connection routing
  - Distribution of topology information
  - Distribution of resource availability information
- Public version being standardized by ITU TS
- Private version specified by ATM Forum Technical Working Group
ATM Service Categories

- **Constant Bit Rate (CBR)**
  - Continuous flow of data with tight bounds on delay and delay variation

- **Real-Time Variable Bit Rate (rt-VBR)**
  - Variable bandwidth with tight bounds on delay and delay variation

- **Non-Real-Time Variable Bit Rate (nrt-VBR)**
  - Variable bandwidth with tight bound on cell loss

- **Available Bit Rate (ABR)**
  - Guarantee minimum
  - Flow control on source with tight bound on cell loss

- **Unspecified Bit Rate (UBR)**
  - No guarantees (i.e., best effort delivery)

<table>
<thead>
<tr>
<th>Service Model</th>
<th>Guarantees?</th>
<th>Congestion feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bandwidth</td>
<td>Loss</td>
</tr>
<tr>
<td>CBR</td>
<td>Constant rate</td>
<td>yes</td>
</tr>
<tr>
<td>VBR</td>
<td>Guaranteed rate</td>
<td>yes</td>
</tr>
<tr>
<td>ABR</td>
<td>Guaranteed minimum rate</td>
<td>no</td>
</tr>
<tr>
<td>UBR</td>
<td>none</td>
<td>no</td>
</tr>
</tbody>
</table>
• 48 Bytes of Data per Cell
• Uses a payload type identifier (PTI) bit in the ATM header to Indicate Last Cell
• Only One PDU at a Time on a Virtual Connection
• CRC-32: Per PDU CRC for error checking
ATM Addressing

- Public networks
  - E.164 numbers (telephone numbers)
  - Up to 15 digits
- Private networks
  - 20 byte address
  - Format modeled after OSI NSAP (Network Service Access Point)
  - Mechanisms for administration exist
  - Hierarchical structure will facilitate virtual connection routing in large ATM networks
  - MAC address will be encapsulated within NSAP
ATM End System Address (AESA)

**AESA Format**

- **Network-Supplied**
  - Data Country Code (DCC)
    - 39: DCC
  - International Code Designator (ICD)
    - 47: IDC
  - E.164 Private Address
    - 45: E.164 Number

- **End System-Supplied**
  - End System ID
  - SEL

**Private UNI**

- Selector (Not used by Network for Routing)

**Higher Order Domain- Specific Part (HO-DSP)-routing field**
**IP over ATM**

- **Classic IP**
  - Layer 3 “networks”
  - connect LANs
  - MAC (802.3) and IP addresses

- **IP over ATM**
  - Replace LAN segments with ATM network
  - ATM addresses, IP addresses
IP-over-ATM (Cont)

- Packet journey in IP-over-ATM network
  - at Source Host (IP-over-ATM router):
    - IP layer maps between IP and ATM dest address
      - IP packet into ATM AAL5 PDUs
      - from IP addresses to ATM addresses just like IP addresses to 802.3 MAC addresses (ARP)
    - passes datagram to AAL5
    - AAL5 encapsulates data, segments cells, passes to ATM layer
      - ATM network: moves cell along VC to destination
    - at Destination Host (IP-over-ATM router):
      - AAL5 reassembles cells into original datagram
      - if CRC OK, packet is passed to IP
Ethernet Switching vs. Virtual Circuit Switching

- No connection setup (connection less)
- Packet carries dest. addr.
- Switching based on globally unique MAC address
- A host does not know whether the network is capable of delivering the packet when it sends the packet
- Each packet is forwarded independently and may be out of order
- A switch and link failure might not have any serious effect if it is possible to find an alternate route

- Establish connection state before sending any data (connection oriented)
  - Setup latency, processing overhead, scalability (capability to grow to a large network)
- Packet/cell carries VCI
- Switching based on incoming port + VCI (unique per port)
  - VCI changed at the output port
- Negotiate the QoS parameters and allocate resources (buffer, bandwidth) to VC
  - If not enough resource, reject the connection request
  - QoS performance guaranteed for connection (bandwidth, delay, delay jitter)
- Each cell is routed along the established connection in order
- If a switch or a link fail, tear down the old connection and establish a new connection

Many ATM ideas adopted in IP networks called MPLS
Today’s Homework

• Peterson & Davie, Chap 3, 4th ed
  -3.1
  -3.5
  -3.7
  -3.8
  -3.13
  -3.26

• Download and browse IEEE 802.3 and ATM UNI4.0 spec and relate contents to today’s lecture
Consider a 1 Mbps 802.11b wireless LAN access point operating in the infrastructure mode. There are 4 active users A, B, C and X all within radio range of each other and the access point. Station X ends its transmission cycle at time $T_0$ on the channel as shown, while stations A, B and C each generate a packet to transmit at time $T_A<T_B<T_C<T_0$ as shown. The packet lengths to be transmitted by A, B and C are 250B, 500B and 1500B respectively, inclusive of all headers. Assume that transmitters use the RTS/CTS mode only for packet sizes larger than 1000B. Using standard values of SIFS (10 msec), DIFS (20 msec), RTS (352 msec), CTS & ACK (304 msec), initial collision window [0,31], CW slot size (10 msec), show an example realization of the 1 Mbps 802.11b channel timeline starting from $T_0$ until packets from A, B and C are successfully transmitted and acknowledged.
Ethernet users U1, U2, U3 and U4 transmit packets P1, P2, P3, P4 of length 5T, 12T, 3T, 8T respectively, where T (51.2 ms) is the end-to-end propagation delay. Assume for simplicity that packet arrivals and transmissions at terminals occur at virtual slot boundaries t=T, 2T, 3T, 4T, ..., nT. Also, assume that the shortest Ethernet packet (after collision) is T sec, and that backoff algorithms delay by kT units of time, where k={0,1}, {0,1,2,3}, {0,1,2,3,4,5,6,7},.... as determined by the number of collisions. Now suppose U1, U2, U3 and U4 have newly arriving packets in their transmit buffers at t1=t3=1 and t2=t4=4 respectively as shown in the table below. Specify the sequence of transmissions and retransmissions on the channel from slot #1 to the time when all four packets P1, P2, P3, P4 have been transmitted successfully assuming no other traffic on the LAN. Provide brief explanations for each channel event and identify any random values used for backoff. Use the following tabular format to give your results, and explain important entries with numbered footnotes.

<table>
<thead>
<tr>
<th>Time slot #</th>
<th>Station U1</th>
<th>U2</th>
<th>U3</th>
<th>U4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CS/packet arrives</td>
<td>Idle</td>
<td>CS/packet arrives</td>
<td>Idle</td>
</tr>
<tr>
<td>2</td>
<td>...</td>
<td>Idle</td>
<td>...</td>
<td>Idle</td>
</tr>
<tr>
<td>3</td>
<td>...</td>
<td>Idle</td>
<td>...</td>
<td>Idle</td>
</tr>
<tr>
<td>4</td>
<td>...</td>
<td>CS/packet arrives</td>
<td>...</td>
<td>CS/packet arrives</td>
</tr>
<tr>
<td>5</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>6</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>