Packet Switching

Key Idea: Transmit data in packets (short bundles) with headers (control)
- Large messages are spread over many packets which move through the network using store-and-forward

ADVANTAGES
- Share link usage ➔ Greater link efficiency
- Links can transmit at different rates ➔ Rate conversion
- Under heavy load, calls can still be accepted but with greater delay
- Can prioritize packet transmission

DISADVANTAGES
- Packet delay at each node (Processing, Queueing, Transmission)
- A packet can acquire jitter while passing through the network
  - Causes: Different route, packet sizes, processing times at each node
- Packets may contain metadata ➔ Higher overhead than circuit switching (e.g., telephone system)
- More processing required at each node in the network

Packet Switching Methods

- Datagram (Connectionless)
  - Each packet carries a header with full destination address
  - Each packet is treated independent of other packets
  - Each network node chooses the next hop for each packet

- Virtual Circuit Switching (Connection-Oriented)
  - Each packet header contains a virtual circuit identifier (VCI)
  - Each node routes packets based on the VCI field
    - Per switch (or port) translation table translates incoming VCI to outgoing VCI and port number
  - A preplanned route is established before sending packets
    - Identifier-to-port mapping requires a call setup phase
  - Faster routing

- Source Routing: Full path stored in packet
**Datagram Forwarding**

- Each switch moves a packet closer to the destination by using its forwarding (routing) table and the destination in the packet header.

**Virtual Circuit Switching (1)**

- The incoming packet header gets (perhaps) a new VCI before being forwarded out a port.

**Virtual Circuit Switching (2)**

- Creating a new connection
  - Assign a new VCI for the connection on EACH link that the connection will traverse
  - The VCI must be unique to the new connection
- Approaches to establishing connection state
  - Permanent virtual circuit (PVC): Done by network administrator
  - Switched virtual circuit (SVC)
    - Signalling: Host sends message to network asking for a connection
    - Done dynamically

**Call Setup (Signalling) (1)**

- Source Setup Switch 1 Setup Destination
  - Call Proceeding
  - Connect
  - Connect ACK
### Signalling (2)
- Host sends connection setup request to its switch
  - Includes destination address and initial VCI
- Switch
  - Has to know how to forward setup request
  - Records request (including incoming VCI)
  - Picks a free VCI for the outgoing VCI
  - Forwards request to next switch
- Other switches along the path repeat above
- Destination
  - Picks free VCI for outgoing VCI
  - Acknowledges setup request back upstream
- Acknowledgement follows reverse path

### Virtual Circuit Switching
- At least 1 RTT delay before data can be sent
- Data packets have a small VCI in place of a full destination address
  - Smaller header overhead than datagram approach
- If a switch or link fails, connections will need to be torn down (free storage) and reestablished
- Switches (their control processors) need a routing algorithm to determine which outgoing port to use during connection setup

### Source Routing
- Store the entire sequence of output ports in the packet header
  - Sender must know the network topology
  - Doesn’t scale
  - Header processing
    - Label rotation
    - Label stripping
    - Label pointer
  - Variable length packets

### Asynchronous Transfer Mode
- **Goal:** Combine flexibility of the Internet with QoS guarantees of telephone network
  - Potential to create a unified infrastructure for carrying voice, video, and data
  - Basically cell relay (as opposed to frame relay)
- **ATM Layer**
  - Connection-oriented (Q.2931 Signalling), in-sequence, unreliable, QoS assured cell transport
- Small fixed-length, packet switched ATM cells
  - Simplified processing → High data rates
  - Statistical multiplexing → Reduced blocking
- Minimal error/flow control → High data rates
- Virtual Circuits
**ATM Cell (Fixed-Size Packet)**

<table>
<thead>
<tr>
<th>5 bytes</th>
<th>48 bytes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Header</strong></td>
<td><strong>Data (Payload)</strong></td>
</tr>
</tbody>
</table>

- **GFC(4/0):** Generic Flow Control (Not Used!)
- **VCI(16):** Virtual Channel Identifier (Routing Field)
- **VPI(12/16):** Virtual Path Identifier (Routing Field)
- **PT(3):** Payload Type
  - 0**: User Information
  - 0*1: Explicit Forward Congestion Indication (EFCI)
  - 0*1: End of AAL5 frame
- **CLP(1):** Cell Loss Priority (Discard Priority)
  - 0: High priority; 1: Subject to discard (e.g., contract violation)
- **HEC(8):** Header Error Control (CRC-8)
  - Correct single bit error in header (most optical link errors)
  - Detect multiple errors in header

**UNI:** User-Network Interface  
**NNI:** Network-Network Interface

---

**ATM Cell Fields**

- **GFC(4/0):** Generic Flow Control (Not Used!)
- **VCI(16):** Virtual Channel Identifier (Routing Field)
- **VPI(12/16):** Virtual Path Identifier (Routing Field)
- **PT(3):** Payload Type
  - 0**: User Information
    - 0*1: Explicit Forward Congestion Indication (EFCI)
    - 0*1: End of AAL5 frame
  - 1**: Maintenance or Management Information (e.g., inband control)
- **CLP(1):** Cell Loss Priority (Discard Priority)
  - 0: High priority; 1: Subject to discard (e.g., contract violation)
- **HEC(8):** Header Error Control (CRC-8)
  - Correct single bit error in header (most optical link errors)
  - Detect multiple errors in header

---

**Small Fixed Sized Cells**

- **Benefits**
  - Reduce line delay for high-priority cells
  - Simpler buffer hardware
    - Easier dynamic storage allocation
  - Simpler line scheduling
    - Controls bandwidth and delay characteristics of a virtual circuit
  - Easier to build large parallel switches
    - Input buffers, switch fabric, output buffers
    - Simultaneous partial packet transfer
  - Controls bandwidth and delay characteristics of a virtual circuit
  - Easier to build large parallel switches
    - Input buffers, switch fabric, output buffers
    - Simultaneous partial packet transfer
- **Problems with large packets**
  - Must be broken down into cells at the source and reassembled at the destination
  - Fragmentation from last cell of a packet

---

**Why 48-Byte Payload?**

- **U.S. wanted small packetization delay**
  - Geographically spread out → Speed-of-light delay is significant
  - Cell size of 64 bytes to get 8 ms packetization delay
    - Audio uses 8000 samples per sec → 1 byte every 125 usec
    - Tuned to time-division multiplex onto a SONET channel
- **48-byte payload was a compromise**
### Line Delay and Multiplexing

- **Reduce line delay for high-priority cells**
  - Maximum queueing delay at 100 Mbps link interface
    - \(53 \times 8 \text{ bits} / 100 \text{ Mbps} = 4.24 \text{ usec}\)
  - 5300-byte cells would result in 424 usec
- **Statistical Multiplexing**
  - **The Problem**: Given \(n\) input links operating at rate \(r\), determine the proper output link rate \(R\).
  - **Worst-Case**: \(R = nr\)
  - **The Bet**: Low probability that aggregate input rate will exceed output link rate
  - Switches store cells \(\Rightarrow\) Trade off delay for slower output line speed
  - **Statistical Multiplexing Gain**: \(nr/R\)

### Statistical Multiplexing

- **Input Link Rate** = 2 cells/sec
- **Output Link Rate** = 4 cells/sec
- **Statistical Multiplexing Gain** = \(\frac{4 \times 2}{4} = 2\)

- **Hope to avoid**: Sudden maximum input burst \(\Rightarrow\) Buffer overflow \(\Rightarrow\) Switch congestion

### Segmentation and Reassembly

- **Adaptation Layers**
  - **AAL1** and **AAL2**: Support applications (e.g. voice) that require guaranteed bit rate service
  - **AAL3/4**: Connection-oriented and Connectionless service
  - **AAL5**: More efficient AAL3, connection-oriented

### AAL3/4 Frame (1)

- **Convergence Sublayer**
  - **Higher Layer PDU**
  - **CPCS-PDU**
  - **CPCS-HEC**
  - **COM**
  - **SAR-RI**
  - **SAR-T**
  - **SAR-PDU**

- **Protocol Data Unit**
  - **Segmentation and Reassembly**
  - **ATM Cell**
  - **ATM Payload**
AAL3/4 Frame (2)

- **CPCS**: Common Part Convergence Sublayer
- **CPCS Header and Trailer Fields**
  - CPI (Common Part Indicator): Version of CS-PDU format
  - Etag: Matches Etag to detect merging of 2 cells
  - BASize: Buffer Allocation Size for reassembly
  - AL = 0
  - Etag: Matches Etag
  - Length: Real length of PDU
- Each cell has SAR header and trailer
  - 2 bytes each

AAL3/4 (3)

- **Type**
  - BOM(10): Beginning of Message
  - COM(00): Continuation of Message
  - EOM(01): End of Message
  - SSM(11): Single-Segment Message
- **SEQ**: Sequence number to detect loss/reordering
- **MID**: Multiplexing identifier (multiple PDUs)
- **Length**: Number of bytes in PDU
- **CRC-10**: To detect errors in 48-byte payload

AAL5 FRAME (1)

<table>
<thead>
<tr>
<th>0-47</th>
<th>1</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>Bytes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - 64KB Payload</td>
<td>Pad</td>
<td>UU</td>
<td>Length</td>
<td>CRC</td>
<td></td>
</tr>
</tbody>
</table>

- Most fields of AAL3/4 frame can be omitted, if:
  - ATM level has low error rates and
  - Higher layer takes care of connection management

AAL5 Frame (2)

- **Goal**: Reduce overhead of AAL3/4 cell payload format
- **Type**
  - PayloadType: 1 bit in each ATM header (PayloadType) of an AAL5 cell used to signal last cell
- **Trailer**
  - UU: User-to-User field for transmitting information to receiver
  - Length: Length of data bytes (excluding padding)
  - CRC: 32-bit CRC over entire EPDU
ATM Connection

Virtual Channel Connection (VCC)
- Setup between endpoints
- Variable-rate, full-duplex flow of fixed-length cells
- Channel Uses: (1) User Data; (2) Control Signalling; (3) Network Management and Routing

Virtual Path Connection (VPC)
- A group of VCCs with the same endpoints
- Amortizes high cost of connection setup over all connections with same VPI
- Control: (1) Calculate route; (2) Allocate capacity; (3) Store connection state

VP/VC CHARACTERISTICS
- Quality of Service: Cell delay (max), delay variation, loss ratio, rate (minimum)
- Switched and Permanent Virtual Connections
  - SVCs (on-demand connection) require signalling for setup and teardown
  - PVCs (long duration) are setup by network management
- Cell Sequence Integrity
  - Cell sequence within a VCC is preserved
- Traffic Parameter Negotiation/Usage Monitoring
  - User: Provides traffic descriptor
    - Rate (average, peak), burstiness, peak duration
  - Network: Monitors users’ traffic for compliance
    - Network can drop non-conforming cells

What Layer Is ATM?
- ATM layer is not a data link layer
  - Not single-hop protocol (ATM layer is multihop)
  - But we can run IP over ATM
- ATM layer could be a network layer
  - Has end-to-end switching (forwarding) and routing
  - Has global addressing (NSAP: Network Service Access Point or E.164)
- ATM layer is not a transport layer
  - There is no guarantee of 100% reliable delivery of cells
  - There are no acknowledgements
- Cells sent on a virtual circuit (VC) will be ordered
  - No guarantee about cells on different virtual circuits

Service Categories
- Constant Bit Rate (CBR) (uncompressed video and audio)
- Real-Time Variable Rate (rt-VBR) (compressed video)
- Non-Real-Time Variable Rate (nrt-VBR) (transaction and monitoring systems)
- Unspecified Bit Rate (UBR) (file transfer, email)
- Available Bit Rate (ABR)
**ATM in the LAN**

- **Advantage of ATM over Ethernet as a LAN**
  - No distance limitation for ATM links
  - Early introduction of high bandwidth (622 Mbps)

- **Typical LAN configuration**
  - ATM backbone connected to Ethernet switches
  - Now, Gigabit Ethernet has overtaken ATM in the LAN

- **Problem with ATM in LAN**
  - Doesn’t look like a traditional LAN
    - ATM is not a shared medium
    - Many protocols (e.g., Address Resolution Protocol) depend on broadcast and multicast
  - Solution 1: Redesign protocols (ATMARP)
  - Solution 2: Make ATM look like shared-media (LANE)

**LAN Emulation (LANE)**

- Each ATM device has an ATM address AND a MAC address

![LAN Emulation Diagram](image)

**LANE Servers**

- LAN emulation configuration server (LECS)
  - Boot: LAN parameters (e.g., type, MTU) over known VCI
- LAN emulation server (LES)
  - Hosts register their MAC and ATM addresses
- Broadcast and unknown server (BUS): Point-to-multipoint VC

![LANE Server Diagram](image)

**Generic ATM Switch**

- VCXT (VC Translation Table)
  - Input Buffers
  - Port Mappers
  - Switch Fabric
  - Output Buffers
LAN Switches (Bridges)
- A node placed between multiple LANs
  - Forward frames from one or more LANs to other LANs
  - Buffers frames \( \rightarrow \) At edge of collision domains
  - Aggregate capacity is greater than single LAN

Learning Bridge
- Builds forwarding table by looking at all source addresses
- Entries timeout and are discarded

Spanning Tree Algorithm (1)
- Avoids forwarding loops by avoiding some ports
- Why loops can occur
  - Provide redundancy
  - The network is a composite of multiple administrative domains
- A spanning tree is an *acyclic* cover of the network nodes in a complete network graph
  - Spans all network nodes and contains no loops
  - Nodes: Bridges and networks
  - Links: Between bridges and networks

Spanning Tree Algorithm (2)
- Each bridge
  - Computes shortest path to root bridge
- Each LAN
  - Has a designated bridge that is closest to root bridge
  - Break ties by choosing lowest bridge ID
- Configuration Message
  - \((\text{Sndr}, \text{distance}, \text{Root})\)
  - Initial message from all bridges: \((\text{Sndr}, 0, \text{Sndr})\)

Spanning Tree Algorithm (3)
- Better configuration message \((\text{Sndr}, d', \text{Root}'\))
  - Smaller root ID or
    - \((\text{Root}', d') < (\text{Root}^*, d^*)\)
  - Same root ID but smaller distance or
    - \((\text{Root}' - \text{Root}^*) \cdot d' < (\text{Root}^* - d^*)\)
  - Root ID and distance are the same as the current one, but the sending bridge has a smaller ID than my ID
- Increment \(d\) and forward better configuration messages
Limitations of Bridges

- Spanning tree algorithm scales linearly
  » Diameter could be linear in number of nodes, not logarithmic as in a hierarchy
- Bridges forward ALL broadcast traffic
- Bridges can interconnect only limited network types (e.g., Ethernet to Ethernet or 802.5 ring)
- VLAN (Virtual LAN) can overcome some scalability
  » e.g., Insert VLAN tag after Ethernet header

![Diagram of VLANs](image)