3 Packet Switching and Internetworking
Packet switching

Packet switching is a basic method for resource sharing in communication networks. "Multiplexing in the large".

Switching node (switch, router, bridge)

dynamic resource allocation

Multiplexer

Demultiplexer
Switching and forwarding of packets

- Data stream is broken into packets at or close to the source. Packets are then operated on as units.
- Packets are interpreted at intermediate and end nodes.
- Problem / design space
  - Implications for the application, quality of service (QoS)
  - Parameters for packetization (length, variable / fixed)
  - Routing finding an appropriate path through the network
  - Congestion control
  - Traffic management and planning
- Variants of packet switching
  - Datagrams
  - Virtual Circuits (virtuelle Schaltungen)
3.1.1 Virtual Circuits (Virtuelle Schaltungen)

- Connection-oriented
  - Connection set-up (parameters: source, destination, QoS)
  - Data transfer
  - Connection release

- Route is determined at connection set-up and kept constant while the connection is up

- Question: how can we identify packets belonging to a connection?
Connection identification (e.g. in X.25, ATM)

Connections are unidirectional, but are often created in pairs for bidirectional communication.

VCIs are carried in packet headers to identify a packet on a link.

Forwarding- and VCI mapping table for node D and the marked connection

<table>
<thead>
<tr>
<th>in port</th>
<th>VCI in</th>
<th>out port</th>
<th>VCI out</th>
<th>QoS</th>
<th>Accounting</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>4</td>
<td>19</td>
<td>xxx</td>
<td>xxx</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>14</td>
<td>1</td>
<td>5</td>
<td>yyy</td>
<td>yyy</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.1.2 Datagrams (Datengramme)

- Packets contain source and destination addresses and are routed according to the destination address
- Connection-less service
- No connection-specific state information needs to be kept in intermediate systems (routers)
- Communication is robust against router failures
3.1.3 Forwarding techniques

- **Incremental forwarding**
  - Routers have a forwarding table specifying the "next hop"

- **Source routing**
  - The route a packet should take is explicitly stored in the packet
  - Fully specified route or partial route only (a subset of all nodes, "loose source routing")

<table>
<thead>
<tr>
<th>Dest. Address</th>
<th>Next hop</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>1</td>
</tr>
<tr>
<td>Y</td>
<td>3</td>
</tr>
<tr>
<td>Z</td>
<td>2</td>
</tr>
<tr>
<td>G</td>
<td>2</td>
</tr>
<tr>
<td>H</td>
<td>1</td>
</tr>
</tbody>
</table>
Packet switching and traffic statistics

- Packet switching may provide a "multiplexing gain" if the traffic has certain statistical properties:
- The traffic of n communications with peak load L may therefore be transmitted over a link with capacity < n*L
- This applies for typical telephony traffic, but not necessarily to traffic presently found in the Internet

(b) yields a higher throughput than (a) if traffic statistics are appropriate
Characteristics of data network vs. telephone network traffic

![Graphs showing packet traffic over time for data networks and telephone networks.](image)

- **Long time scale**
- **Short time scale**

Packets/time unit
3.2 Bridges and LAN switches

Bridges and switches serve to create a LAN-based internetwork - i.e. a network of LANs

Why would we want this?

- LANs have a limited geographical dimension (200-500 m with Ethernet)
- Stations connected to the same medium (hub, cable) form a collision domain
- Many stations in the same collision domain, generating a high traffic, cause high collision rates and high medium access delays
- Operation of LANs of different technology (Ethernet, Token Ring, FDDI) in the same environment
Self-learning, transparent bridge

- Solution: Use a device to break a LAN into separate collision domains
- Bridge accepts all frames on LAN 1 and forwards them to LAN 2, using an internal forwarding table (store-and-forward principle)
- Self-learning, transparent bridge
  - Broadcasts to all ports, if location of destination unknown
  - Learns topology by monitoring frame origins
- Forwarding table of the bridge when all hosts are known (in steady state)
IEEE 802.x MAC addresses

- IEEE 802.3 and 4: LSB is transmitted first
- IEEE 802.5: MSB is transmitted first
- Place of transmission of G/I bit is different for 802.3 vs. 802.5
- Canonical format: LSB is first

<table>
<thead>
<tr>
<th>Group / Individual selector bit</th>
</tr>
</thead>
<tbody>
<tr>
<td>OUI1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Global / Local selector bit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LSB</th>
<th>MSB</th>
<th>LSB</th>
<th>MSB</th>
<th>MSB</th>
</tr>
</thead>
<tbody>
<tr>
<td>22 Bit</td>
<td>24 Bit</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- IEEE 802.3 and 4: LSB is transmitted first
- IEEE 802.5: MSB is transmitted first
- Place of transmission of G/I bit is different for 802.3 vs. 802.5
- Canonical format: LSB is first

<table>
<thead>
<tr>
<th>IEEE 802.x MAC addresses</th>
</tr>
</thead>
<tbody>
<tr>
<td>OUI1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Global / Local selector bit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LSB</th>
<th>MSB</th>
<th>LSB</th>
<th>MSB</th>
<th>MSB</th>
</tr>
</thead>
<tbody>
<tr>
<td>22 Bit</td>
<td>24 Bit</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- IEEE 802.3 and 4: LSB is transmitted first
- IEEE 802.5: MSB is transmitted first
- Place of transmission of G/I bit is different for 802.3 vs. 802.5
- Canonical format: LSB is first

<table>
<thead>
<tr>
<th>IEEE 802.x MAC addresses</th>
</tr>
</thead>
<tbody>
<tr>
<td>OUI1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Global / Local selector bit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LSB</th>
<th>MSB</th>
<th>LSB</th>
<th>MSB</th>
<th>MSB</th>
</tr>
</thead>
<tbody>
<tr>
<td>22 Bit</td>
<td>24 Bit</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Problem: Topology with multiple paths

After initialization of the bridges:
- Frame from X to V is forwarded to port 2 by B1
- ... and is forwarded to port 1 by B3 ...
- ... and is forwarded to port 1 by B2 ...
- ... and ...

Diagram:

- Nodes: X, Y, Z, V
- Bridges: B1, B2, B3
- Connections:
  - X to Y
  - Y to B1
  - B1 to B2
  - B2 to Z
  - Z to V
  - V to B3
  - B3 to B2
  - B2 to B1
  - B1 to X
Avoiding loops with a spanning tree

- This topology contains potential loops.
- It cannot be implemented with transparent bridges.
- Bridges use an algorithm to build a spanning tree.
Network represented as a graph
(LAN: vertex; bridge: edge)

- Spanning tree (marked in red): Acyclic, contiguous subgraph of an arbitrary graph
  - contains all vertices
  - edges are deleted such that no cycles exist
  - there are many possible spanning trees (optimization!)
Active bridge ports corresponding to the spanning tree

Note: B5 is not used (appears to be obsolete)
A dynamic protocol may be able to change the spanning tree at runtime
Algorithm and protocol see Peterson Kap. 3.2.2.
Distributed spanning tree algorithm (Radia Perlman)

- Spanning tree is built from a root bridge
- Each bridge assumes initially that it is the root
- Each bridge has a unique ID, $\text{ID}_{\text{min}} < \text{ID} < \text{ID}_{\text{max}}$
- Bridges periodically send configuration messages
  - ID, ID of supposed root bridge, Distance in hops between sender and supposed root bridge
- Each bridge stores "best" configuration received so far
- A configuration is "better" than another if
  - it reports a root bridge with a smaller ID
  - if it reports a root bridge with the same ID, but a smaller distance
  - if root and distance are equal, but the sending bridge has a smaller ID
- Bridge adds 1 to the distance and stores a "better" config
- Only root bridge(s) generate messages; non-root bridges simply add 1 to distance and forward received messages
Restrictions and disadvantages of bridges

- Broadcast frames need to be forwarded in the whole internetwork
- So far no hierarchical spanning tree algorithms has been found
- Scalability to large networks not possible; practical sizes are several ten LANs
- Pragmatic solution: Partitioning the network in several virtual LANs (VLAN)
3.3 Virtual LANs (VLANs)

- A set of hosts to belong to the same VLAN is assigned a VLAN-id.
- Each bridge will forward frames addressed to VLAN $x$ to a LAN, iff (if and only if) at least one host belonging to $x$ is reachable via this LAN.
- Assigning hosts to VLAN is done by a (manual) management operation.
VLAN example

VLAN-id=21
VLAN-id=14
VLAN-id=5

LLC header is extended by a VLAN id
VLAN standard: IEEE 802.1q
3.4 LAN switch

q Similar to a LAN bridge: forwards frames from input port to output port

q Each connected host is connected via a unique medium (cable)

q Depending on performance properties and price tag:
   • Switch may be non-blocking: Simultaneously accepts frames from any or all ports
   • Cut-through switching: A frame is only partially buffered and is forwarded as soon as the destination port is known and available (storage amount is a few bytes, up the destination address)

q Buffering required if several frames are simultaneously directed at the same output port

q Switches are available for 10 / 100 / 1000 Mbit/s Ethernets
3.5 Cell-based switching (ATM)

q ATM: Asynchronous Transfer Mode
q Developed between 1985 and 1995 as a core technology for the Broadband ISDN (B-ISDN)
q Specific interest by the telecom industry (manufacturers / operators), later also by the computer industry (ITU-T; ATM Forum)
q Applications (initial vision, pre-Internet era!)
  • High speed multi-service digital public network
  • Integration of voice and data
  • ATM to the desktop
  • Interconnection of PABXes, interconnection of company multi-service (voice, video, data, ... = multimedia) networks
ATM core characteristics

- Uses short 53-byte cells
- Connection-oriented; connections are set-up manually or by Q.2931 signaling protocol
- Extensive support for QoS: for manually or dynamically created connections
- ATM uses three protocol layers:
  - AAL: service-dependent adaptation
  - ATM: cell-based routing and forwarding
  - PHY: transmission of data
- All three layers implemented in hardware (only signaling is done in software)
ATM protocol layer

q **Physical Layer (PHY):**
   - Digital transmission on fiber / (D)WDM, SONET/SDH and other media

q **ATM Layer (ATM):**
   - Transport of ATM cells (53 bytes), cell switching
   - Virtual Channels, Virtual Channel Connections
   - QoS support
   - Flow and congestion control

q **ATM Adaptation Layer (AAL):**
   - Offers several application-friendly services
   - Segmentation and assembly of cells to frames with a service-specific format
ATM cell format

Format of a cell at the User-Network Interface (UNI)

Format of a cell at the Network-Network Interface (NNI)
Model for the use of virtual paths and virtual channels

Virtual Path Connection
(switching done on VPI, typically only in public network domain)

Virtual Channel Connection
(switching done on VCI, only needed in corporate network domain)
ATM Adaptation Layer 5

User data, max. 64 kByte

Convergence - forming the CS-PDU

CS-PDU

Segmentation of the CS-PDU into ATM cells

CS-PDU Trailer

Pad  Res  Len  CRC-32

PT bit in cell header signals end of frame
The rise and fall of ATM

- Cell switching conceived in the mid-eighties as a packet switching technology for multi-service networks --> ATM
- Pushed as "The Technology for B-ISDN" by ITU-T and ATM-Forum; Goal: "ATM to the desktop"
- First products available early nineties
- Progress hampered by complexity of signalling protocols, slow speed of standardization
- 1993/94: The arrival of the Internet
- Mid-nineties: ATM used for multimedia experiments (e.g. high-quality videoconferencing)
- ATM to the desktop vision abandoned
The rise and fall of ATM (continued)

q 1996-98: IP forwarding recognized as a performance bottleneck in the backbone
   • IP Routers with ATM switching fabric
   • ATM in the backbone: Static and dynamic (IP switching); work on MPLS

q IP forwarding problem solved
   • algorithmic solutions speed up IP forwarding
   • new, high-performance hardware

q MPLS development focuses on traffic engineering

q 2000: MPLS matures; ATM becomes obsolete

q ATM currently only used as a link layer technology in ADSL access technology
3.6 Switching fabrics for packet switching
Types of switch fabrics

- **Space switch**: Arriving cells are forwarded within the switch fabric on different parallel paths (banyan and crossbar switch).

- **Time switch**: Arriving cells are brought on a medium which is shared among all the input ports (bus or ring) and are taken from there to the output ports.

- **Shared memory switch**: Arriving cells are stored in a memory shared by all the input ports and are read out from there to the output ports.
Space switch (Crossbar)

- Very simple basic approach
- Output port has to accept frames from several inputs simultaneously (→ buffering)
- Complexity grows with n^2
- Knockout switch is a practical realization (s. Petersen)

Petersen figure 3.28
Space switch (Banyan)

- Internal blocking is possible, e.g. packets addressed simultaneously to output ports 011 and 010 from inputs B or H would collide.

- Blocking at the output ports is always possible, if packets addressed simultaneously to one output port.
Shared memory switch
The need for buffers

- at the input: collision in the fabric (input buffering)
- at the output: collision at the output port
- in the fabric: collisions at internal points of contention
- problem of head-of-line (HOL) blocking