Cell Switching (ATM)

- Connection-oriented packet-switched network
- Used in both WAN and LAN settings
- Signaling (connection setup) Protocol: Q.2931
- Specified by ATM forum
- Packets are called cells
  - 5-byte header + 48-byte payload
- Commonly transmitted over SONET
  - other physical layers possible

Variable vs Fixed-Length Packets

- No Optimal Length
  - if small: high header-to-data overhead
  - if large: low utilization for small messages
- Fixed-Length Easier to Switch in Hardware
  - simpler
  - enables parallelism

Big vs Small Packets

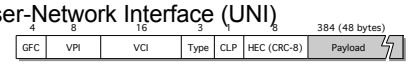
- Small Improves Queue behavior
  - finer-grained preemption point for scheduling link
    - maximum packet = 4KB
    - link speed = 100Mbps
    - transmission time = 4096 x 8/100 = 327.68us
    - high priority packet may sit in the queue 327.68us
    - in contrast, 53 x 8/100 = 4.24us for ATM
  - near cut-through behavior
    - two 4KB packets arrive at same time
    - link idle for 327.68us while both arrive
    - at end of 327.68us, still have 8KB to transmit
    - in contrast, can transmit first cell after 4.24us
    - at end of 327.68us, just over 4KB left in queue

Big vs Small (cont)

- Small Improves Latency (for voice)
  - voice digitally encoded at 64KBps (8-bit samples at 8KHz)
  - need full cell's worth of samples before sending cell
  - example: 1000-byte cells implies 125ms per cell (too long)
  - smaller latency implies no need for echo cancellers
- ATM Compromise: 48 bytes = (32+64)/2

Cell Format

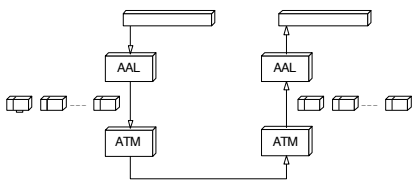
- User-Network Interface (UNI)



  - host-to-switch format
  - GFC: Generic Flow Control (still being defined)
  - VCI: Virtual Circuit Identifier
  - VPI: Virtual Path Identifier
  - Type: management, congestion control, AAL5 (later)
  - CLPL Cell Loss Priority
  - HEC: Header Error Check (CRC-8)
- Network-Network Interface (NNI)
  - switch-to-switch format
  - GFC becomes part of VPI field

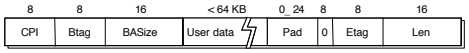
Segmentation and Reassembly

- ATM Adaptation Layer (AAL)
  - AAL 1 and 2 designed for applications that need guaranteed rate (e.g., voice, video)
  - AAL 3/4 designed for packet data
  - AAL 5 is an alternative standard for packet data



AAL 3/4

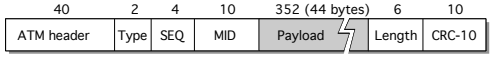
- Convergence Sublayer Protocol Data Unit (CS-PDU)



  - CPI: commerce part indicator (version field)
  - Btag/Etag:beginning and ending tag
  - BAsize: hint on amount of buffer space to allocate
  - Length: size of whole PDU

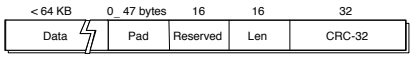
Cell Format

  - Type
    - BOM: beginning of message
    - COM: continuation of message
    - EOM end of message
  - SEQ: sequence of number
  - MID: message id
  - Length: number of bytes of PDU in this cell



AAL5

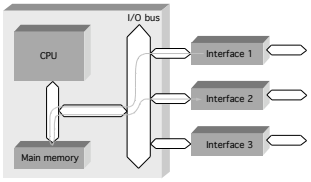
- CS-PDU Format



  - pad so trailer always falls at end of ATM cell
  - Length: size of PDU (data only)
  - CRC-32 (detects missing or misordered cells)
- Cell Format
  - end-of-PDU bit in Type field of ATM header

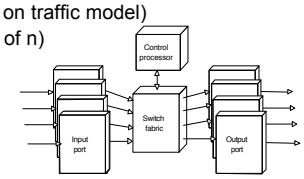
Workstation-Based

- Aggregate bandwidth
  - 1/2 of the I/O bus bandwidth
  - capacity shared among all hosts connected to switch
  - example: 1Gbps bus can support 5 x 100Mbps ports (in theory)
- Packets-per-second
  - must be able to switch small packets
  - 300,000 packets-per-second is achievable
  - e.g., 64-byte packets implies 155Mbps



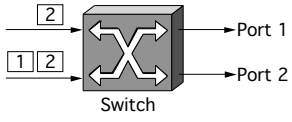
Switching Hardware

- Design Goals
  - throughput (depends on traffic model)
  - scalability (a function of n)
- Ports
  - circuit management (e.g., map VCIs, route datagrams)
  - buffering (input and/or output)
- Fabric
  - as simple as possible
  - sometimes do buffering (internal)

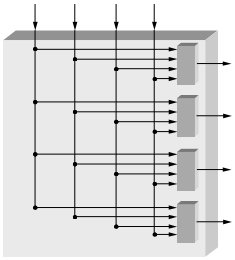


Buffering

- Wherever contention is possible
  - input port (contend for fabric)
  - internal (contend for output port)
  - output port (contend for link)
- Head-of-Line Blocking
  - input buffering

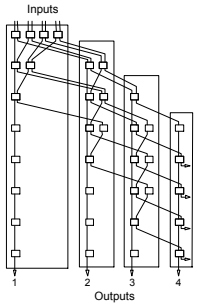


Crossbar Switches



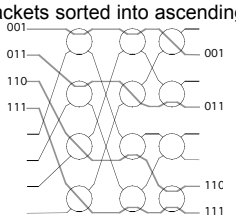
Knockout Switch

- Example crossbar
- Concentrator
  - select l of n packets
- Complexity: n^2



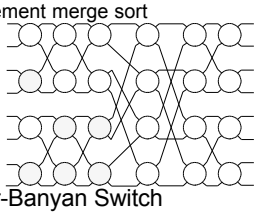
Self-Routing Fabrics

- Banyan Network
  - constructed from simple 2 x 2 switching elements
  - self-routing header attached to each packet
  - elements arranged to route based on this header
  - no collisions if input packets sorted into ascending order
  - complexity: n $\log_2$ n



Self-Routing Fabrics (cont)

- Batcher Network
  - switching elements sort two numbers
    - some elements sort into ascending (clear)
    - some elements sort into descending (shaded)
  - elements arranged to implement merge sort
  - complexity: n log22 n

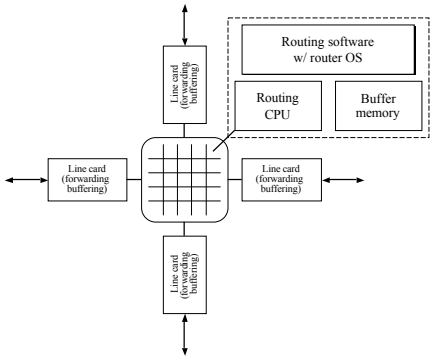


- Common Design: Batcher-Banyan Switch

High-Speed IP Router

- Switch (possibly ATM)
- Line Cards + Forwarding Engines
  - link interface
  - router lookup (input)
  - common IP path (input)
  - packet queue (output)
- Network Processor
  - routing protocol(s)
  - exceptional cases

High-Speed Router



Alternative Design