EECS 122, Lecture 25

Today's Topics:

Introduction to ATM

Kevin Fall, kfall@cs.berkeley.edu

Asynchronous Transfer Mode (ATM) Concepts

- Motivations:
  - flexibility of the Internet
  - per-connection QoS facilities of telephone network
- Basic architecture:
  - virtual circuits
  - small fixed-size packets (cells)
  - statistical multiplexing
  - traffic types and QoS support

STM and Packet Switching

- Issues with Synchronous Transfer Mode:
  - unused time slots wasted because receiver knows owner of data only by timing
  - because of fixed cyclical schedules in TDM, can only get bandwidth in multiples of the available schedule (e.g. n * 64kb/s)
- Packet Switching:
  - by placing headers on packets, can provide statistical multiplexing (better utilization) at the added cost of additional buffering

Packet Switched Networks

- The datagram method (like Internet)
  - every packet has complete destination (and possibly source) information; wastes bandwidth if addresses are long
  - independent forwarding; reordering allowed
- The virtual circuit (VC) method
  - every packet has small identifier which is mapped at each hop to connection info (smaller overhead)
  - requires connection setup, followed by data

Virtual Circuit Identifiers (VCIs)

- How to know which identifier (VCI) should be placed in a packet?
  - unique global selection is difficult in large networks
  - instead, use a locally-unique identifier
  - requires translation (swapping) at each intermediate switch
- Each switch must maintain per-connection state to facilitate routing and label swapping

Implications of VC Switching

- switches maintain tables
  - switch failure kills connections
- data follows the same route after set-up
  - re-ordering tricky or not allowed
  - duplication unlikely
- data/signaling separation
- no guarantee of low loss
- required set-up takes an RTT

Some Particulars

- Types of VC:s
  - Permanent (pre set-up) VCs (PVCs)
  - Switched (set up on demand) VCs (SVCs)
  - pt-to-pt (uni/bi directional), multipoint (uni)
- Virtual Path (VP) and Circuit (VC)
  - each contain an ID, often combined
  - some switches provide for VC aggregation by switching VPs rather than VCs
- Can be used for tag/L3/MPLS approaches

ATM Signaling

- Separation of signaling
  - UNI - node and network; NNI - internal to the network
  - UNI 4.0 (latest, 136 pps) includes recv-inititated joins
- UNI 3.0/3.1 based on Q.2931, in turn based on Q.931 (used for N-ISDN)
- One-pass method of circuit set-up
  - routing and admit control based on dest address, QoS

Cell Switching

- ATM uses fixed-sized "cells"
  - simpler buffer hardware
  - simpler line scheduling (recall WFQ)
  - large parallel switches easier to build
- Actually uses 5 byte header and 48 byte payload:
  - 48 is average of 32 and 64; 9.4% overhead
  - what you get when designed by committee

Buffer Hardware

- Usually need to pre-allocate buffer areas prior to receiving packets:
  - if variable-sized packets are used, can vastly under-utilize allocated space
  - if variable-sized buffers are used, more complexity in managing multiple free lists
- With single fixed size, only 1 list necessary; minimizes memory fragmentation

Line Scheduling

- If only fixed-sized packets are ever used, not hard to provide ratios of bandwidth allocated to various circuits
- With variable-sized packets, things are more complicated:
  - need to account for larger packet sizes
  - recall bit-by-bit round-robin and how it is approximated with fair queueing

Large Parallel Switches

- For large switches, generally need parallel processing to perform switching function
- If packets always the same size, can predict amount of time spent in each hardware resource/pipeling stage
- Variable-sized packets can degrade parallelism

Problems with Fixed Sizes

- Source wishing to send a message larger than fixed size must divide it up (segmentation)
- Receiver must be ready to reconstruct original message (reassembly)
- If message size is < than fixed size, excess bandwidth is wasted; similar problem for last cell of message

The Overhead/Delay Tradeoff

- A smaller fixed-size cell provides for a smaller packetization delay (time to wait before sending a message)
- A larger cell provides a better overhead rate
- With ATM 5 bytes of 53 are overhead, so can never be more than 90.57% efficient

Integrated Services

- ATM provides the ability to specify a Quality of Service (QoS) associated with virtual circuits (a traffic descriptor)
- During setup, an admission control procedure is invoked to allocate resources
- If successful, traffic policing is used to verify user does not violate agreement

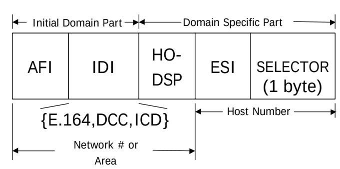
ATM Addressing

- ATM is envisioned to be use both as a public and private network
- Public network addresses are worldwide-unique "phone numbers" (E.164)
- For private networks, addressing is based on ISO standard (called NSAP) addresses

ATM Addressing

- Network Service Access Point (NSAP) addresses, originally from ISO:
  - variable-length (between 7 and 20 bytes)
  - hopelessly cumbersome set of acronyms…
  - and to make it more confusing, there is a way to encode E.164 addresses in NSAP format

The NSAP Format…



- IDI can be E.164 "phone number"
- DCC Format: IDI is data country code, country identifiers from ISO 3166. Administered by ISO national members
- ICD: Intl Code Designator (from BSI)

NSAP Acronym Decode

- NSAP, network service access point
- AFI, authority & format identifier
- IDI, initial domain identifier
- ICD, international code designator (BSI)
- DCC, data country code (ISO)
- IDP, initial domain part (AFI + IDI)
- DSP, domain specific part (end sys addr)
- HO-DSP, high-order bits of DSP (prefix)
- ESI, end system identifier (like host part)
- SEL, selector (like link type field)

Service Provided by ATM

- connection-oriented (VCs)
  - point-to-point (unidirectional, bidirectional)
  - point-to-multipoint (unidirectional only)
- in-sequence (no re-ordering)
- unreliable (no retransmission)
- QoS on a per-VC basis

ATM Cell Header Format

8 bits across

GFC[UNI]/VPI	VPI	
VPI	VCI	
VCI		
VCI	PT	CLP
HEC		

- VPI/VCI: virtual path/circuit ID
- PT: payload type, CLP: cell loss priority
- HEC: header error check (CRC8)
- GFC: generic flow control (mostly ignored)

ATM Cell Payload Type Field

- First PT bit 1 means management cell
- First PT bit 0 means user cell:
  - next bit is explicit forward congestion indication (EFCI)
  - 3rd bit is "user signaling" bit, used primarily as EOF indicator for AAL5 (a sort of layer violation)

ATM Adaptation Layers (AAL)

- AAL1: constant bit-rate CBR service
- AAL3/4: VBR service, checksums on each cell, detects loss/corruption; 4 bytes of 48 byte payload used for overhead [see book, p 183]
- AAL5: VBR service, checksums on each frame, detects loss/corruption; simpler header, less overhead than AAL3/4
- AAL5 is most popular…

AAL5

- Using a layer violation, use one bit in the ATM header PT field to mark end of PDU
- No per-cell CRCs, all frame meta-data appears at *end* of AAL5-layer frame
- Provides up to 64KB data in frame:

CRC	LEN	CPI	UU	PAD	DATA
4 bytes	2	1	1	0-47	<64KB

AAL5 Frame Fields

- Pad: used to fill-out frame to be an even multiple of 48 bytes. Allows for frame meta-data to always appear at end of entire AAL5 frame
- UU: user-to-user info, treated as opaque data by ATM, unspecified [in 1996]
- CPI: a type field, current value of zero
- Length: length of overall PDU
- CRC: CRC over entire payload

IP and ATM

- In the late 80's and early 90's there was hope among ATM providers that TCP/IP could be abandoned in favor of a completely ATM-based network
- They lost.
- ATM is use today primarily in wide-area support of IP-based networking
- So, how to use IP with ATM?

IP over ATM

- ATM characteristics:
  - connection-oriented
  - fixed small cells
  - non-broadcast
  - QoS support
  - several address formats
- IP characteristics:
  - connectionless
  - variable-sized datagrams
  - ARP requires broadcast
  - best-effort service
  - single address format

IP Encapsulation on AAL5

- need to place IP packet in AAL5 frame, and translate IP address to ATM address
- so, just place IP packet in payload part:
  - well, what if non-IP traffic is also there?
  - use 802-compatible LLC/SNAP header
  - uses up 8 bytes, with 2 for protocol selection
- how to translate the address?
  - can't broadcast on ATM

Address Mapping for ATM

- Create ARP Server, used to hold IP-to-ATM address mappings for the LAN
- ATM hosts must contact server to ascertain destination ATM address (during exchange, server learns requester's mappings too)
- To avoid scaling problems with large ATM network, divide up ATM net into *logical IP subnets* (LISs); broadcast only within LIS [called Classical IP on ATM]

NHRP (Next-Hop Routing Protocol)

- Problem is that we need to go thru an IP router even on same ATM net
- NHRP solves this by providing mappings for all reachable ATM hosts
- NHRP servers self-coordinate (a bit like DNS servers)
- See RFC 2332 (Apr 1998)

Multicast

- How to do receiver-based join given source-based multipoint VCs?
- MARS - multicast address resolution servers
- MARS servers map IP multicast packets to all interested receivers
- Can also map IP multicast address to set of ATM addresses for source based multipoint VCs

LAN Emulation (LANE)

- A way of using ATM net essentially as a layer 2 bridge; (spanning tree protocol still used for loop avoidance)
- Broadcasting provided by a broadcast server
- Hides QoS support, so good for migration to ATM, but doesn t really take advantage of it
- Only Ethernet and Token Ring for now…

Virtual LANs (VLANs)

- Administrator can set up multiple different emulated LANs
- Removes hard association between physical location and attached LAN
- Possible to apply different policy to different VLANs

ATM Routing (P-NNI)

- ATM standard hierarchical routing protocol (>2 levels of hierarchy)
- Link state approach, but uses source routing set up at edge nodes
- Switch controllers form peer groups, elect a peer group leader to enter in the next-higher-level group
- Can annotate route info with QoS constraints

PNNI QoS Routing

- ATM QoS types:
  - CBR (constant bit-rate w/timing)
  - VBR (variable bit-rate w/timing)
  - ABR (available bit-rate, no timing)
  - UBR (unspecified bit rate)
- Admission control at each node; on failure resorts to crankback to find an acceptable path