Network technology

November 12, 1998

Topics

• Overview
• Telephone system
• Ethernet
• ATM
Course Theme

Abstraction is good, but don’t forget reality!

Earlier courses to date emphasize abstraction
  • Abstract data types
  • Asymptotic analysis

These abstractions have limits
  • Especially in the presence of bugs
  • Need to understand underlying implementations

Useful outcomes
  • Become more effective programmers
    – Able to find and eliminate bugs efficiently
    – Able to tune program performance
  • Prepare for later “systems” classes
    – Compilers, Operating Systems, Networks, Computer Architecture
“Harsh Realities” of Computer Science

- Int’s are not integers; float’s are not reals
  - Must understand characteristics of finite numeric representations
- You’ve got to know assembly
  - Basis for understanding what really happens when execute program
- Memory matters
  - Memory referencing bugs especially difficult
    » Violates programming language abstraction
  - Significant performance issues
    » E.g., cache effects
- There’s more to performance than asymptotic complexity
  - Constant factors also matter
- Computers do more than execute programs
  - Get data in and out
  - Communicate with each other via networks
Computer system

Processor and L1 cache

Interrupt controller

Keyboard controller

Keyboard

Mouse

Serial port controller

Modem

Parallel port controller

Network adaptor

Video adaptor

Display

Network

Local/IO Bus

Memory

IDE disk controller

SCSI controller

SCSI bus

Disk

cdrom

Disk

Disk

Modem

Keyboard controller

Serial port controller
Simple example

Starting Point: Want to send bits between 2 computers
- FIFO queue on each end
- Can send both ways ("full duplex")
- Name for standard group of bits sent: "packet"
- Packet format and rules for communicating them ("protocol")

Simple request/response protocol and packet format:

<table>
<thead>
<tr>
<th>header</th>
<th>payload</th>
</tr>
</thead>
<tbody>
<tr>
<td>0/1</td>
<td>data/address</td>
</tr>
</tbody>
</table>

0: please send the data word at “address”
1: here is the data word you asked for.
Questions about simple example

What if more than 2 computers want to communicate?
  • Need computer address field in packet?

How do multiple machines share the interconnect?
  • multiple paths? arbitration? congestion control?

What if a packet is garbled in transit?
  • Add error detection field in packet?

What if a packet is lost?
  • More elaborate protocols to detect loss?

What if multiple processes/machine?
  • one queue per process? separate field in packet to identify process?

Warning: You are entering a buzzword-rich environment!!!
Generic network

protocol stack

adaptor/interface card

link

link

link

kernel code

Interconnect

s/w interface

h/w interface

s/w interface

h/w interface

s/w interface

h/w interface

s/w interface

h/w interface
Protocols

A protocol defines the format of packets and the rules for communicating them across the network.

Different protocols provide different levels of service:

- simple error correction (ethernet)
- uniform name space, unreliable best-effort datagrams (host-host) (IP)
- reliable byte streams (TCP)
- unreliable best-effort datagrams (process-process) (UDP)
- multimedia data retrieval (HTTP)
Protocols provide specialized services by building on services provided by other protocols.

Unreliable best effort datagram delivery (process-process)

Unreliable best effort datagram delivery (host-host)

Reliable byte stream delivery (process-process)

Physical connection
Protocol stacks

Host A

- physical
- data link
- network
- transport
- application

Repeaters/Bridges/Routers

Host B

- physical
- data link
- network
- transport
- application

Telnet, FTP, HTTP, email

TCP/UDP

IP

CSMA/CD

10Base-T

Reliable, efficient end user service

Routing

Flow control

Congestion

Framing, Error recovery, media access

Xmit raw bits

Xmit raw bits

Framing, Error recovery, media access

Routing

Flow control

Congestion

Reliable, efficient end user service
Transmission media

**twisted pair:** (1-2 Mb/s at 1 km)
- 2 insulated copper wires
- Braided outer conductor
- Plastic cover

**coaxial cable:** (1-2 Gb/s at 1 km)
- Stiff copper wire
- Insulator
- Plastic cover

**fiber:** (100-200 Gb/s at 1 km)
- Light source
- Silica

**station wagon full of mag tapes hurtling down the highway:**
- (15 Gb/s at 1 hour)
- 7 GBytes/tape
- 1000 tapes/station wagon (50x50x50cm)
- 7,000 GBytes total
- 7,000 GBytes/3600 minutes = 15 Gb/s

$5/tape reused 10 times -> $500 tape cost
$200 for shipping -> 10 cents /GByte
Shared vs switched media

**Shared media (e.g., Ethernet)**

**Switched media (e.g., ATM)**

```
<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
</table>
```

```
<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
</table>
```

```
<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
</table>
```

```
<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
</table>
```

```
<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
</table>
```

```
<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
</table>
```

```
<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
</table>
```

```
<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
</table>
```

```
<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
</table>
```

```
<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
</table>
```
Network performance measures
## Example performance measures

<table>
<thead>
<tr>
<th></th>
<th>SAN</th>
<th>LAN</th>
<th>WAN</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interconnect Example</strong></td>
<td>CM-5</td>
<td>Ethernet</td>
<td>ATM</td>
</tr>
<tr>
<td><strong>Bisection BW</strong></td>
<td>N x 5MB/s</td>
<td>1.125 MB/s</td>
<td>N x 10 MB/s</td>
</tr>
<tr>
<td><strong>Int./Link BW</strong></td>
<td>20 MB/s</td>
<td>1.125 MB/s</td>
<td>10 MB/s</td>
</tr>
<tr>
<td><strong>Latency</strong></td>
<td>5 μsec</td>
<td>15 μsec</td>
<td>50 to 10,000 μs</td>
</tr>
<tr>
<td><strong>HW Overhead to/from</strong></td>
<td>0.5/0.5 μs</td>
<td>6/6 μs</td>
<td>6/6 μs</td>
</tr>
<tr>
<td><strong>SW Overhead to/from</strong></td>
<td>1.6/12.4 μs</td>
<td>200/241 μs</td>
<td>207/360 μs</td>
</tr>
</tbody>
</table>

(TCP/IP on LAN/WAN)
Importance of Overhead (+ Latency)

Ethernet / SS10: 9 Mb/s BW, 900 μsecs ovhd
ATM Synoptics: 78 Mb/s BW, 1,250 μsecs ovhd.

NFS trace over 1 week: 95% msgs < 200 bytes

- Link bandwidth is as misleading as MIPS
Basic network types

System area network (SAN)
- same room (meters)
- 300 MB/s Cray T3E

Local area network (LAN)
- same bldg or campus (kilometers)
- 10 Mb/s Ethernet
- 100 Mb/s Fast Ethernet
- 100 Mb/s FDDI
- 150 Mb/s OC-3 ATM
- 622 Mb/s OC-12 ATM

Metropolitan area network (MAN)
- same city (10’s of kilometers)
- 800 Mb/s Gigabit Nectar

Wide area network (WAN)
- nationwide or worldwide (1000’s of kilometers)
- telephone system
- 1.544 Mb/s T1 carrier
- 44.736 Mb/s T3 carrier
AT&T Telephone Hierarchy

10 regional offices (fully interconnected)

67 sectional offices

230 primary offices

1,300 toll offices

19,000 end offices

200 million telephones

local loops

200 million telephones

local loops
Computer-to-computer calls

V.34 modem \rightarrow codec \rightarrow codec \rightarrow V.34 modem

home computer \rightarrow codec \rightarrow local office \rightarrow toll office \rightarrow local office \rightarrow home computer

digital (short cable or bus) 33 MB/s

28.8 Kb/s analog local loop

1.544 Mb/s (T1 carrier)

digital

digital

digital (short cable or bus) 33 MB/s
Modulating digital signals

- Binary signaling
- Sine wave carrier (1kHz-2kHz)
- Amplitude modulation
- Phase modulation

- 00: no shift
- 01: 1/4 shift left
- 10: 1/2 shift left
- 11: 3/4 shift left

(shifts are relative to previous wave)
Quadrature amplitude modulation (QAM)

Modern modems use a combination of amplitude and phase modulation to encode multiple bits per “symbol”, i.e. amplitude/phase pair.

3 bits/symbol QAM modulation (8 symbols)

phase angle is 1/4

4 bits/symbol QAM modulation (16 symbols)

1/8
Conventional Modems

MOdulate: convert from digital to analog
DEModulate: convert from analog to digital

<table>
<thead>
<tr>
<th>modem standards:</th>
</tr>
</thead>
<tbody>
<tr>
<td>type</td>
</tr>
<tr>
<td>v.32</td>
</tr>
<tr>
<td>v.32.bis</td>
</tr>
<tr>
<td>v.34</td>
</tr>
</tbody>
</table>

Theoretical limit for modulated signals is approx 35 Kb/s:

Shannon's law: max bits/s = H log2(1 +S/N), where H is bdwdth and S/N is signal to noise ratio. For phone network, H~3,600 and S/N is 30 dB. Thus max rate is ~35 Kb/s.
T1 carrier (1.544 Mb/s)

Digital part of phone system based on the T1 carrier:

193 bit frame (125 us, 8000 samples/s, 8 bits/sample/channel)

- Channel 1
- Channel 2
- Channel 3
- Channel 24

Bit 1 is a framing code

8 data bits per channel

Each channel has a data rate of 8000 samples/s * 8 bits/channel = 64 Kb/s
56KB “Modems”

- Asymmetric: home to SP uses conventional v.34 modem
- SP has digital connection into phone system
  - Channel sending 8000 samples/second, up to 8-bits/sample
- DAC encodes each sample with 92 or 128 voltage levels
  - Not enough precision on analog side to handle finer resolution
- Receiver converts samples back to digital values
  - Must match frequency & phase of senders DAC
  - Establish using “training” signals from sender

Key: no analog conversion at ISP
Ethernet

History
- 1976- proposed by Metcalfe and Boggs at Xerox PARC
- 1978 - standardized by Xerox, Intel, DEC

Bandwidth
- 10 Mbits/sec (old) , 100 Mbits/sec (new)

Key features
- broadcast over shared bus (the ether)
  - no centralized bus arbiter
  - each adapter sees all bits
- each adapter has a unique (for all time!) 48-bit address
- variable length frames (packets) (64 - 1518 bytes)
Ethernet cabling

<table>
<thead>
<tr>
<th>Name</th>
<th>Cable</th>
<th>Max Segment</th>
<th>Nodes/Segment</th>
<th>Advantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>10Base5</td>
<td>thick coax</td>
<td>500 m</td>
<td>100</td>
<td>good for backbones</td>
</tr>
<tr>
<td>10base2</td>
<td>thin coax</td>
<td>200 m</td>
<td>30</td>
<td>cheapest</td>
</tr>
<tr>
<td>10Base-T</td>
<td>twisted pair</td>
<td>100 m</td>
<td>1024</td>
<td>easy maintenance</td>
</tr>
<tr>
<td>10Base-F</td>
<td>fiber</td>
<td>2000 m</td>
<td>1024</td>
<td>best between bldgs</td>
</tr>
</tbody>
</table>
Repeaters directly transfer their inputs to their outputs.
Bridges maintain a cache of hosts on their input segments.

Selectively transfer packets from their inputs to their outputs.
### Ethernet packet (frame) format

64 - 1518 bytes

<table>
<thead>
<tr>
<th>Preamble</th>
<th>Dest addr</th>
<th>Src addr</th>
<th>Frame type</th>
<th>Payload</th>
<th>CRC</th>
<th>Postamble</th>
</tr>
</thead>
<tbody>
<tr>
<td>64 bits</td>
<td>48 bits</td>
<td>48 bits</td>
<td>16 bits</td>
<td>368-12000 bits</td>
<td>32 bits</td>
<td>8 bits</td>
</tr>
</tbody>
</table>

visible from the host

*Preamble*: 101010101 (synch)

*dest and src addr*: unique ethernet addresses

*payload*: data

*CRC*: cyclic redundancy check (error detection/correction)
Ethernet receiving algorithm

Ethernet adaptor receives all frames.

Accepts:
- frames addressed to its own address
- frames addressed to broadcast address (all 1’s).
- frames addressed to multicast address (1xxx...), if it has been instructed to listen to that address
- all frames, if it has placed in promiscuous mode

Passes to the host only those packets it accepts.
Ethernet sending algorithm (CSMA/CD)

Problem: how to share one wire without centralized control.

Ethernet solution: Carrier Sense Multiple Access with Collision Detection (CSMA/CD):

1. Adaptor has frame to send and line is idle:
   • then send frame immediately

2. When adaptor has frame to send and line is busy:
   • wait for line to become idle, then send frame immediately.

3. If “collision” (simultaneous sends) occurs during transmission:
   • send at least 1024 bits
   • send “jam signal” to notify receivers
   • wait some period of time (binary exponential backoff)
   • retry
Binary exponential backoff algorithm:

• after 1st collision, wait 0 or 1 slots, at random.
• after 2nd collision, wait 0, 1, 2, 3 slots at random.
• etc up to 1023 slots.
• after 16 collisions, exception.
Why the 64 byte minimum packet size?

Assume propagation delay from A to B is $\tau$.

A sends to B at time 0

packet almost at B at time $\tau - \epsilon$

B sends at time $\tau$: collision

Noise burst gets back to A at time $2\tau$

Conclusions: Senders must take more than $2\tau$ seconds to send their packets.

For ethernet, $\tau$ is specified by standard (2500 m cable w/ 4 repeaters) to be 51.2 usecs, which at 10 Mb/s is 512 bit times, or 64 bytes.

Rough estimate: propagation through copper is about 20 cm/ns. With a 2500 m cable, $\tau$ is 12.5 us and $2\tau$ is 25 us.

As speeds increase there are two possibilities:
1. increase packet sizes
2. decrease maximum cable length

Neither is particularly appealing.
Ethernet pros and cons

Pros:
- simple
- robust
- cheap ($50/adapter in 1998)

Cons:
- no quality of service guarantees
  - OK for data
  - not OK for real-time bit streams like video or voice
- fixed bit rate
  - not keeping up with faster processors
  - workstation can produce data at 10-50 MBytes/sec
- prone to congestion
  - processors getting faster
  - bridged Ethernets can help some
Asynchronous transfer mode (ATM)

History
- 1988- proposed by international ATM forum
- telecommunications and computer vendors

Goal:
- mechanism for integrated transport of bit streams with different performance and reliability requirements (quality of service)
  - video: 1.5 Mbits/sec, latency and variance sensitive, some bit loss OK
  - voice: 64 Kbits/sec, latency and variance sensitive, some bit loss OK
  - data: high data rates, latency and variance insensitive, bit loss not OK
ATM overview (cont)

Bandwidths

- OC-1: 51.84 Mbits/sec
- OC-3: 155.52 Mbits/sec (current LAN rate)
- OC-12: 622.08 Mbits/sec (current LAN rate)
- OC-24: 1244.16 (Gigabit network)

Key features:

- virtual connections (VC’s)
  - allow bandwidth reservation
- fixed cell (frame) size of 53 bytes
  - simplifies high-speed switching
- small cell size
  - allows fine-grained allocation of network bandwidth
# ATM cell format

53 bytes (fixed)

<table>
<thead>
<tr>
<th>Generic flow ctl</th>
<th>VCI/ VPI</th>
<th>Payload type</th>
<th>Priority</th>
<th>Header checksum</th>
<th>Payload</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 bits</td>
<td>24 bits</td>
<td>2 bits</td>
<td>2 bits</td>
<td>8 bits</td>
<td>48 bytes</td>
</tr>
</tbody>
</table>

- **VCI**: virtual connection (channel, circuit) identifier
- **VPI**: virtual path identifier
- **payload**: data
ATM cell routing

Input port routing table

Input ATM cell

vci/vpi

Output ATM cell

vci/vpi'

Switch

Routing table

Input link

output port

output link

output port
ATM pros and cons

Pros:

- bandwidth can be reserved (connections)
- scalable aggregate bandwidth (wide range of supported bit rates)
- support for network traffic with different quality of service requirements (small, fixed, easily multiplexed cells)
- potential for high speed switching (small fixed-size cells)

Cons:

- maximum user bandwidth still limited by link bandwidth
- connections make broadcast and multicast more difficult
- quality of service is still a research issue