...g Systems

...tion Networks

Goddard
...cse.unl.edu

.../~goddard/Courses/CSCE855

CSCE 455/855

Communication for Distributed Systems

- Communication is central
  - » process management: messages synchronize, coordinate
  - » file mgmt: messages access and transmit files and directory information
  - » device mgmt: messages carry data, access devices
  - » memory mgmt: messages carry data
- Network Layer
  - » facilities to send and receive messages to addressed locations
  - » routing: messages are forwarded
    - ❖ the Internet

Communication Networks

- Open System Interconnect (OSI) 7-layer model
  - » Physical
  - » Data Link
  - » Network
  - » Transport
  - » Session
  - » Presentation
  - » Application

OSI 7-layer model

- Physical
  - » transmit bits and bytes
  - » LANs
- Data Link
  - » translate signals (bits/bytes) into *frames*
  - » checksums, source and destination
- Network
  - » translate frames into packets
  - » packet routing
  - » datagrams

OSI 7-layer model

- Transport
  - » reliable end-to-end byte streams
  - » packet re-sending, packet ordering
  - » virtual circuits
- Session
  - » high-level naming, bi-directional streams
  - » managing more than one communication session
- Presentation
  - » translation between protocols
  - » heterogeneous systems
- Application
  - » user applications

OSI 7-layer model (cont.)

- Higher-level facilities
  - » Application
  - » Presentation
  - » Session
- Basic network communication services
  - » transport
  - » network
- Physical medium and LANs
  - » data-link
  - » physical

OSI 7-layer model (cont.)

- General tradeoff
  - » quite flexible
    - supports a wide range of applications
    - makes communication as transparent as user needs it (i.e. user can choose level as needed)
  - » each layer adds overhead
    - reduce number of layers
    - simplify layers
    - improve implementations
- Specific protocols for different problems

Network and Transport Layers

- Network layer does not provide reliability
  - » packets may be lost
    - no means to detect errors
    - higher levels must provide for detection of errors and re-sending packets
  - » packets may be received out-of-order
    - large messages broken down into fixed packet size
    - packets reconstructed to make message at recipient

Network and Transport Layers (cont.)

- Transport layer
  - » transparent transfer of data
  - » reliability provided
    - ❖ re-sending lost messages
    - ❖ packet ordering
- Network layer provides services for transport layer
  - » connectionless (IP - datagram service)
  - » connection-oriented (X.25 - virtual circuit service)

Connectionless and Connection-Oriented Services

- Connectionless services
  - » datagrams: single message sent from point-to-point
  - » no relationship established between packets
  - » advantages:
    - ❖ protocol is simple
    - ❖ data delivery is fast
  - » disadvantages:
    - ❖ no error handling, ordering of packets
    - ❖ each packet self-identifying; leads to long headers
    - ❖ packets may arrive out of order

Connectionless and Connection-Oriented Services (cont.)

- Connection-oriented service
  - » virtual circuit: data path between endpoints
  - » communication can have a state
    - send a reply every 5 messages
  - » three phases
    - establish connection
    - transfer data
    - release connection
  - » advantage: reliable communication
  - » disadvantage:
    - protocol complexity makes communication slower
    - error handling, flow control add overhead

Connection-Oriented Services

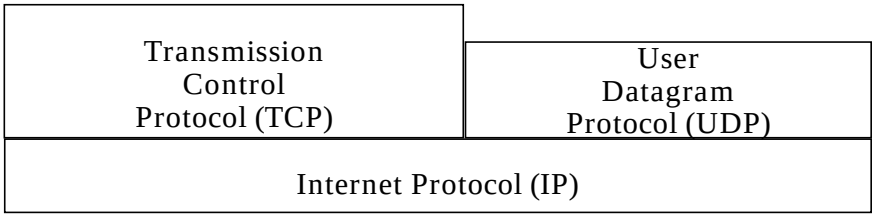
- Overhead from “virtual circuits”
  - » connection establishment, release
  - » reliability
    - detecting lost messages (time-outs, etc)
    - re-sending lost messages
    - message acknowledgment
    - packet ordering
    - ordering algorithm
    - sending additional order information
- Network layer
  - » virtual circuits do not guarantee reliability
  - » connectionless services (datagrams) dominate

Connection-Oriented Services (cont.)

- Transport layer
  - » virtual circuits guarantee reliability
    - TCP
  - » some connectionless services have reliability services
    - IP: guaranteed packet delivery with TCP over IP

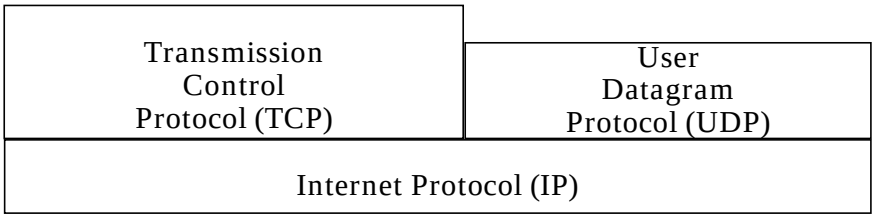
Internet Protocols

- Internet Protocol (IP)
  - » connection-less
  - » network routing
  - » datagram construction



Internet Protocols

- Transmission Control Protocol (TCP)
  - » connection-oriented; establish a (logical) virtual circuit
  - » positive acks, time-out
  - » sequence numbers
  - » connection procedures
  - » state information is kept
- User Datagram Protocol (UDP)
  - » no acks, messages may arrive out-of-order
  - » essentially IP with some minor additions



Asynchronous Transfer Mode (ATM)

- Connection Oriented
- Virtual circuits
- Fixed-size blocks (cells)
- Connection is established and all cells follow the same route over a switched network

ATM Rationale

- Voice transmissions require steady bandwidth
  - » bandwidth needs are low, but need to be consistent
- Data (esp. real-time) is bursty
  - » high rates needed when transmitting data, no bandwidth otherwise
- Want networks to handle both
  - » solution: small packets that can be rapidly switched

ATM Characteristics

- Fixed size blocks sent over virtual circuits
  - » routing info stored in switches
- A packet-switching network
  - » meaning packet transmissions can be interleaved
- Packets broken into very small cells
- Allowed to drop cells
  - » usually results in re-transmission of entire packet

ATM (lowest three) Layers

- Physical
  - » same functionality as OSI Layer 1
- ATM
  - » OSI Layer 2 and part of OSI Layer 3
- Adaptation
  - » OSI Layer 4 but without reliable end-to-end service

ATM
Physical Layer

- Designed to use optical technology
- Essentially digital switch technology
  - » star topology with switch as central node
  - » each machine has dedicated connection to switch
  - » multiple communication paths can be open simultaneously
- Switching networks...
  - » allow scaling to large networks

ATM
ATM Layer

- Connection-oriented cell routing
  - » connection set up only if sufficient resources are available
- Cell structure
  - » 48 bytes of data
  - » 5 header fields (53 bytes total)
  - » Virtual Path Identifier (VPI)
  - » Virtual Circuit Identifier (VCI)

ATM
ATM Layer (cont.)

- Virtual channel (VC)
  - » unidirectional association between source and destination
  - » refers to specific channels inside the virtual path
  - » allocated dynamically at connection setup
- Virtual paths (VP)
  - » collection of VCs
  - » (semi-)permanent connection between pairs of endpoints

ATM Adaptation Layer

- Essentially chops packets into cells ... then re-assembles them
- Cells can be dropped
  - » adaptation layer not reliable
- Need higher layers for transport connections
  - » use ATM cells to carry TCP/IP packets
  - » TCP/IP will take care of reliability
  - » means entire packet will need to be re-sent

ATM Switching

- VC and VP together provide routing information
  - » VPI: refers to virtual path on the physical link
  - » VCI: refers to specific VC inside VP
- General routing strategy
  - » VPI field used by routing tables to determine next destination
  - » VPI field modified at each hop
  - » if virtual path used by more than one cell
    - ❖ use VCI field to determine destination
    - ❖ VCI field also changed at each hop

ATM Switching (cont.)

- Cells needing the same output line
  - » must choose whether to buffer or not
    - standard allows to just drop a cell
    - don't want to do that often
  - » buffering at the input port
    - pick one cell to forward, hold others
    - long input queues may result
    - this blocks cells wanting to go to other output ports
    - ...known as *head-of-line blocking*
  - » buffering at the output port
    - queue located at output port
    - removes head-of-line blocking
    - can also have a pool of input and output buffers

Local Area Networks

- Three dominant topologies
  - » star (digital switch, ATM)
  - » ring
  - » bus
- Ethernet
  - » multi-access bus technology
  - » messages broadcast to all nodes
  - » all nodes listen to bus
    - receives only messages addressed to the node
  - » bus contention: single communication line

Ethernet

- Implements physical and data link layers
  - » multi-access bus
  - » transmits data link frames
- Access Method: Carrier sense multiple access with collision detection (CSMA/CD)
  - » CSMA to reduce the chance of collisions
  - » CD to detect collisions (and retransmit with back off)

Ethernet CSMA/CD

- Carrier Sense Multiple Access (CSMA)
  - » carrier sense: listen for clear bus
    - ❖ if busy, wait for clear carrier
    - ❖ if clear, send message (transmit a packet)
  - » listen to bus while transmitting for CD
- Collision Detection (CD)
  - » sender compares outgoing message to received message
    - ❖ if mismatch occurs, assume collision has occurred
  - » if collision has occurred
    - ❖ each sender waits a period of time (back off)
    - ❖ then re-send packet

Ethernet Collision Detection

- Assume a collision occurs
  - » all nodes back off 2 sec and re-transmits
  - » what will happen?
- Back off intervals
  - » nodes detection collision back off a random time interval
  - » what if another collision occurs
    - ❖ may want to back off a longer time period
  - » binary exponential backoff
    - ❖ i^{th} collision back off between 0 and 2^i-1 interval

Ethernet vs. ATM

- Ethernet
  - » keep traffic fairly sparse to avoid collisions
    - ❖ gateways to divide network into smaller units
  - » limit transmission time
    - ❖ keep packet size small
    - ❖ keep length of network small
    - ❖ increase transmission speed
- ATM
  - » packets can be transmitted in parallel
  - » “collisions” handled by ATM buffers
    - ❖ buffers are a finite size
    - ❖ cells can be dropped - re-transmit packet
  - » scaling to larger networks
    - ❖ use larger switches
    - ❖ switching networks
    - ❖ network speed a factor, but importance reduced by parallel transfer in ATM
    - ❖ busy machine can be a bottleneck