Advanced Unix System Administration

Lecture 13
October 29, 2008

Steven Luo
<sluo+decal@OCF.Berkeley.EDU>

IP: the Internet Protocol

- IPv6 con't
  - IPv4 and IPv6 can coexist, but v6 is not backwards-compatible
  - Transition mechanisms:
    - IPv6-in-IPv4 tunnels
    - Mapping the IPv4 space into the IPv6 space
    - Protocol translation mechanisms
  - Transition is going slowly, but Catch-22 involved: no one will deploy until commonly used, no one will use until commonly deployed

The Transport Layer

- User Datagram Protocol (UDP)
  - Extremely simple, but provides basically no features
  - Unreliable, message-oriented, stateless
  - Each message is packaged into a single datagram and sent over the network
  - No guarantees on delivery, order of arrival
  - UDP packet: source port (16 bits), destination port (16 bits), length (16 bits), checksum (16 bits), data

The Transport Layer

- Ports
  - In TCP and UDP, each socket is assigned a port number to identify the traffic to it
  - Kernel examines destination port to decide which process to give data to; source port used to determine destination of replies
  - By convention:
    - 1-1023 are “Well Known Ports” for services, 1024-49151 are “Registered Ports” for services,
    - 49152-65535 are dynamic ports used as source ports

The Transport Layer

- Transmission Control Protocol (TCP)
  - Much more elaborate and featureful than UDP
  - Reliable, stream-oriented, connection-oriented
  - Applications send streams of data which TCP packages into packets and sends over the network
  - Correct and in-order delivery is guaranteed even on unreliable networks

The Transport Layer

- The TCP packet
  - Source port (16 bits), destination port (16 bits)
  - Sequence number (32 bits)
  - Acknowledgment number (32 bits)
  - Data offset (4 bits), gives size of header in 32-bit words, reserved field (4 bits)
  - TCP flags (8 bits): CWR, ECE, URG, ACK, PSH, RST, SYN, FIN
  - Window size (16 bits), gives number of bytes sender is willing to receive before ACK

The Transport Layer

- The TCP packet con't
  - Checksum (16 bits)
  - Urgent pointer (16 bits)
  - Options, padded to an integral multiple of 32 bits
  - Data
- TCP connections
  - Three phases of connections: establishment, data transfer, teardown

The Transport Layer

- TCP connections con't
  - Establishment (3-way handshake):
    - Client sends packet with SYN set to server
    - Server replies with SYN/ACK
    - Client sends ACK
    - Unexpected/unwanted connections rejected with RST
  - Data transfer
    - The sequence number of the packets with SYN set give initial sequence numbers (ISNs)
    - Each byte of data in the stream is given a sequence number, starting with ISN+1

The Transport Layer

- TCP connections con't
  - Data transfer con't
    - Receipt of each packet is acknowledged with an ACK with ack number set to the last byte in sequence received + 1
      - Selective packet acknowledgment is available as an option
    - Packets not acknowledged will be retransmitted; duplicates will be dropped silently
    - Number of bytes a sender will send before waiting for ACK is controlled by the window size

The Transport Layer

- TCP connections con't
  - Data transfer con't
    - TCP implementations use data such as retransmissions, ACK rates, and the like to adjust to conditions (via changing the window, slowing transmission rate, etc.)
  - Teardown
    - FIN is sent to announce that one has no more data to send
    - That half of the connection is closed when the ACK reply is received