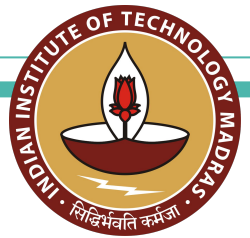
CS3205 : Introduction to Computer Networks

Assignment 2: TCP Congestion Control

Chigullapally Sriharsha

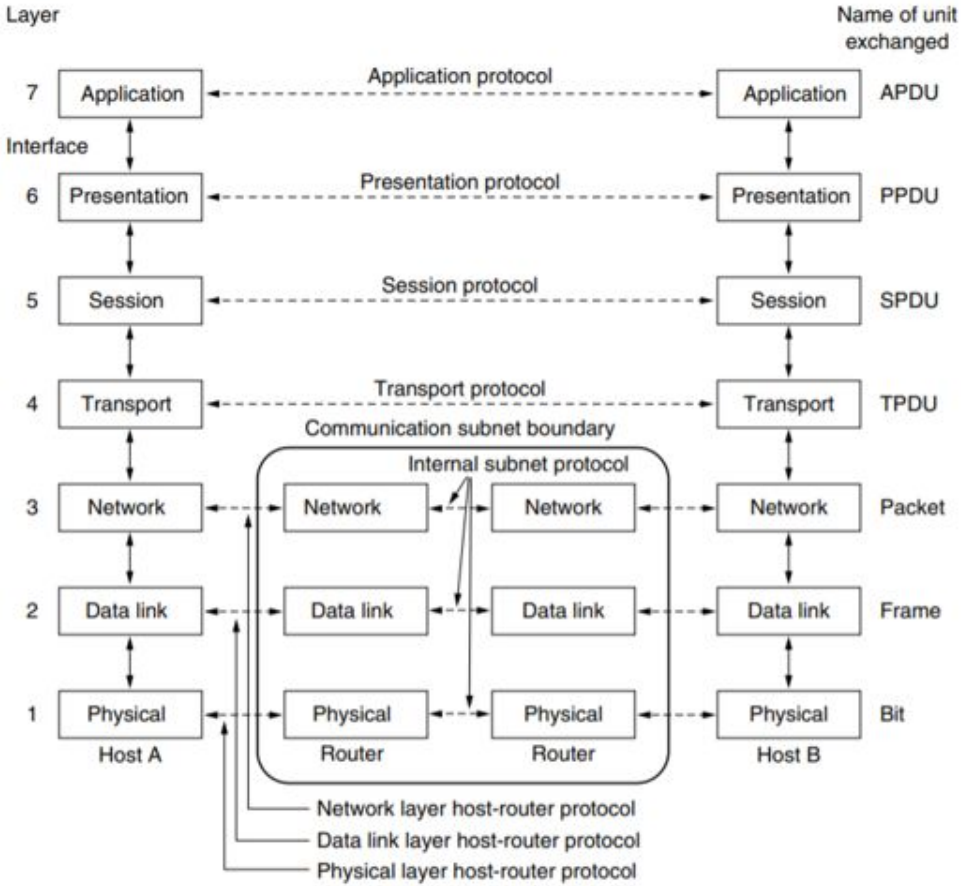


Outline

Transmission Control Protocol (TCP)

- Reference models
- Sliding Window ARQ
- Transport Layer
- TCP
  - TCP Segment header
  - TCP Connections
  - Timers
  - Error Control
  - Flow Control
  - Congestion Control

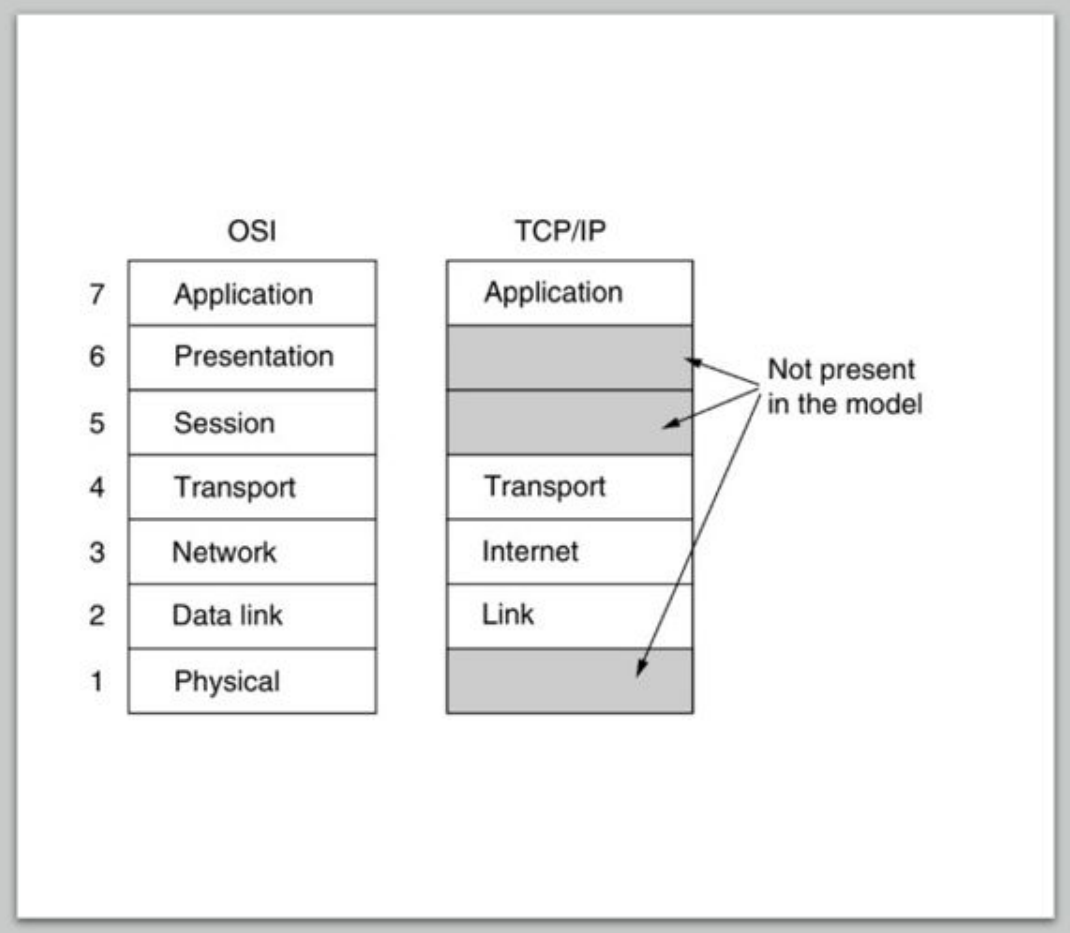
The OSI reference model



Source: Andrew S. Tanenbaum and David J. Wetherall. 2010. Computer Networks (5th. ed.). Prentice Hall Press, USA.

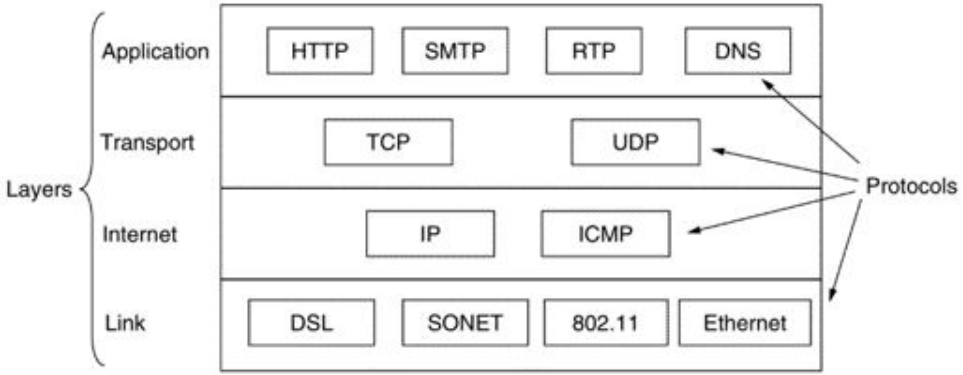
The TCP/IP Protocol Suite

- developed by US Defense Advanced Research Project Agency (DARPA)
- for ARPANET packet switched network
- used by the global Internet
- protocol suite comprises a large collection of standardized protocols
- Concise version of OSI model



Source: Andrew S. Tanenbaum and David J. Wetherall. 2010. Computer Networks (5th. ed.). Prentice Hall Press, USA.

TCP / IP Model

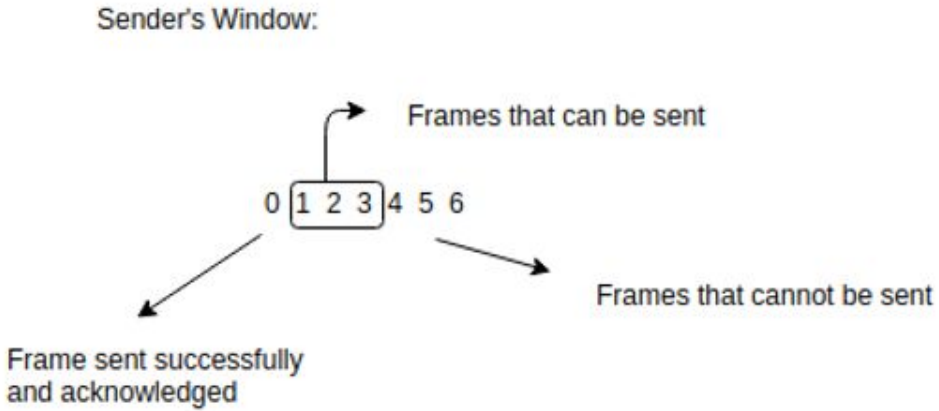


Source: Andrew S. Tanenbaum and David J. Wetherall. 2010. Computer Networks (5th. ed.). Prentice Hall Press, USA.

TCP/IP model

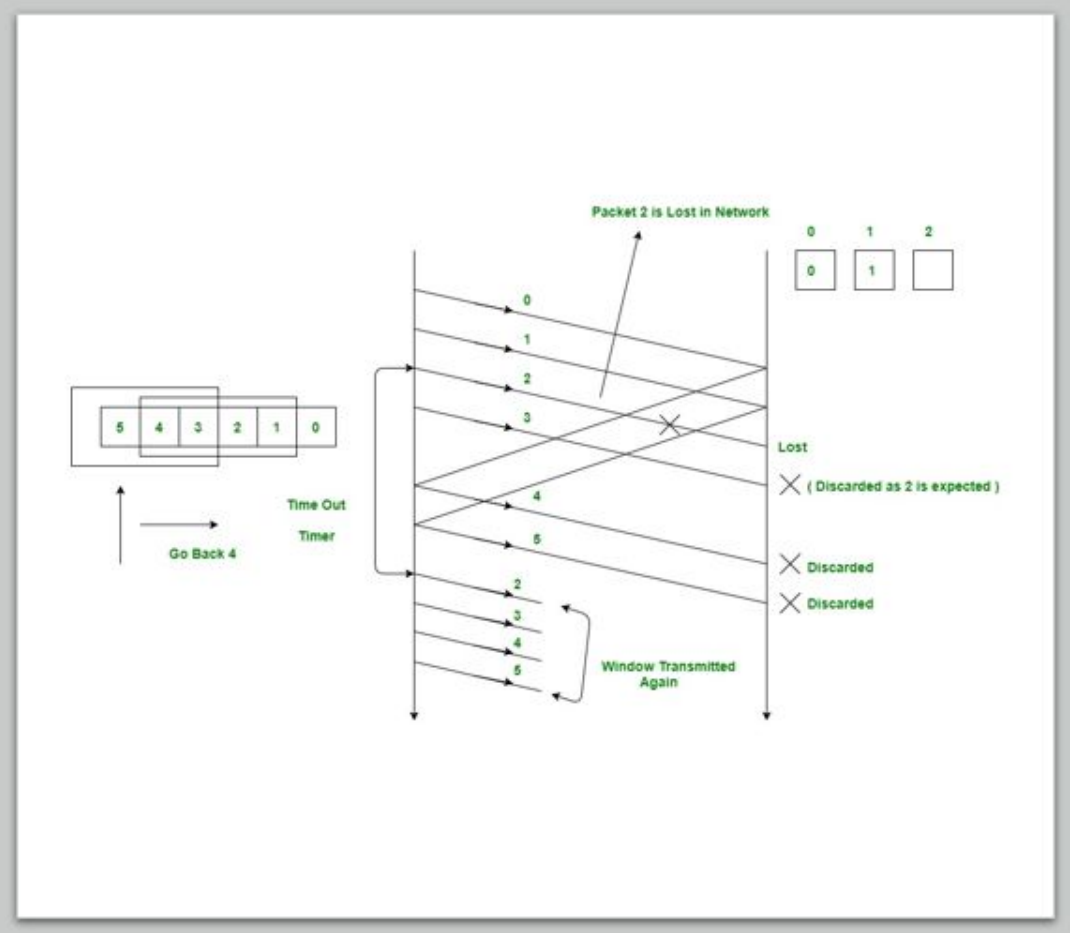
- Link Layer:
  - framing,
  - error control and flow control at link-level,
  - physical addressing.
- Internet Layer:
  - Routing,
  - Logical addressing,
  - Congestion control,
  - Fragmentation.
- Transport Layer:
  - End-to-end connectivity,
  - error control and flow control between two end-points,
  - segmentation,
  - congestion control,
  - multiplexing,
  - service point addressing.
- Application Layer:
  - Provides Applications with access to network services.
  - Applications produce the data that is transferred over the network.

Sliding Window ARQ



Go back N Sliding Window Protocol

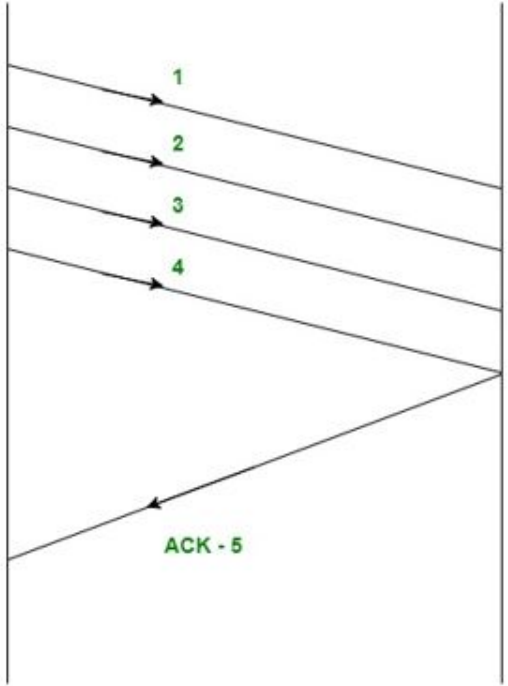
- Size of sender window = N
- Size of receiver window = 1
- The minimum number of sequence numbers required is N+1
- If the sent frame is found to be suspected then all the frames are re-transmitted from the lost packet to the last packet transmitted.
- Out of order frames are discarded.
- Cumulative acknowledgement is used



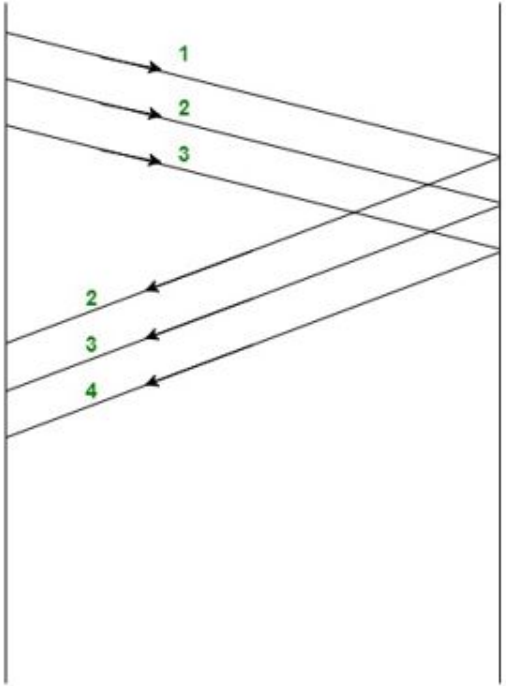
Source: https://www.geeksforgeeks.org/sliding-window-protocol-set-2-receiver-side/

Acknowledgements

- Cumulative Ack
- Independent Ack

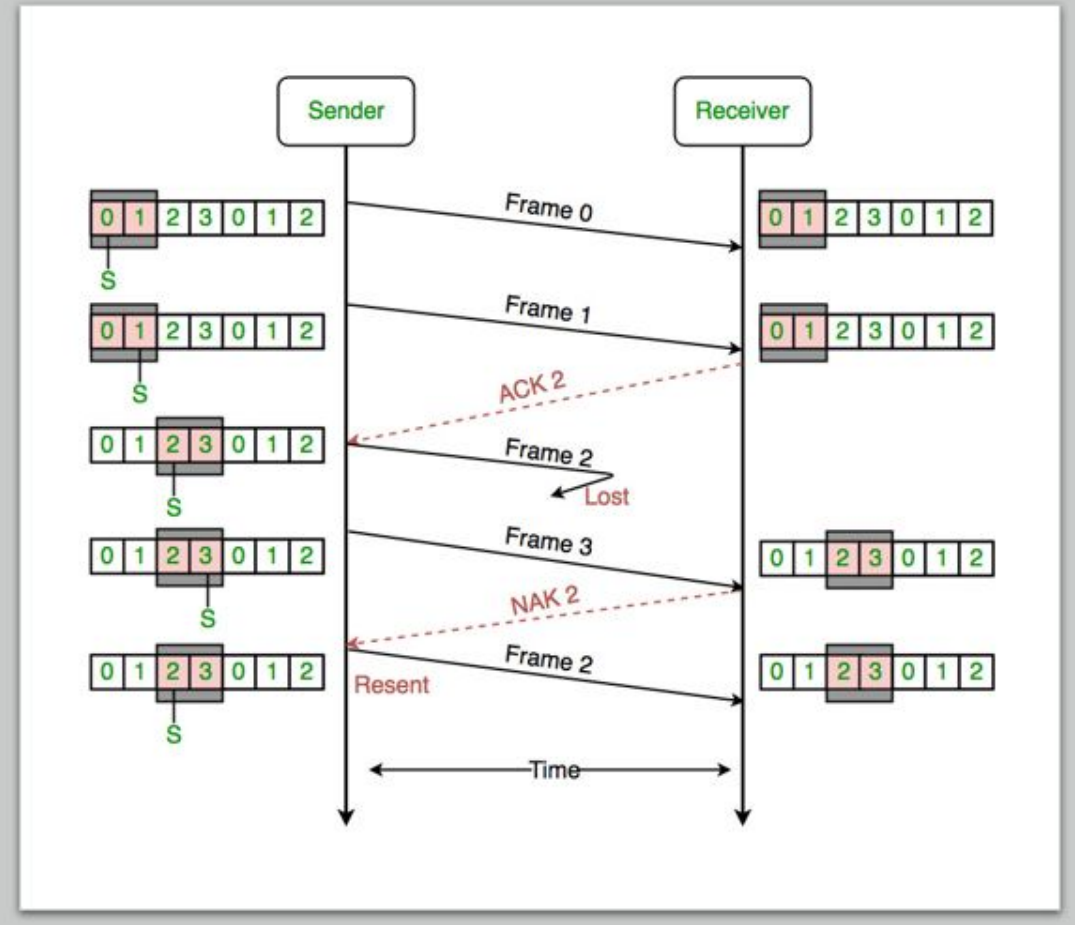


Cummulative



INDEPENDENT

Source: https://www.geeksforgeeks.org/sliding-window-protocol-set-2-receiver-side/



Selective Repeat Sliding Window Protocol

- Here, both sender and receiver have same window size
- If window size of both sender and receiver is W, then minimum number of sequence numbers required is 2W
- Only those frames are re-transmitted which are found suspected
- Out-of-Order frames are accepted, if they are in receiver's window
- Individual acknowledgment is used
- If receiver receives a corrupt packet, it immediately sends a negative acknowl edgement and hence only the selective packet is retransmitted

Source: https://www.geeksforgeeks.org/sliding-window-protocol-set-3-selective-repeat/

Transport Layer

- End-to-end connectivity,
- error control and flow control between two end-points,
- segmentation,
- congestion control,
- multiplexing,
- service point addressing.

TCP

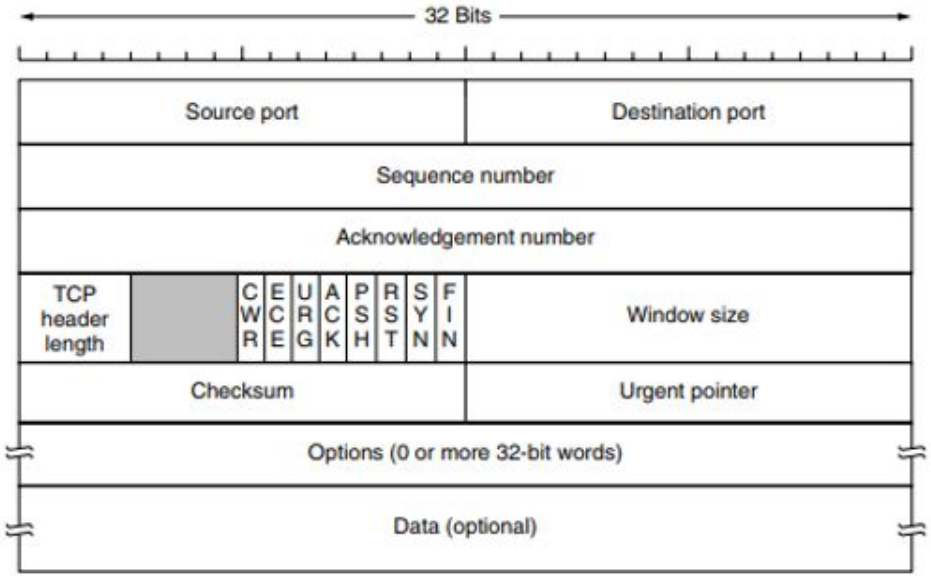
The various services provided by the TCP (Transmission Control Protocol) to the application layer are as follows:

- Process-to-Process Communication: Port numbers are 16 bit long that help identify which process is sending or receiving data on a host.
- Stream oriented: This means that the data is sent and received as a stream of bytes (unlike UDP or IP that divides the bits into datagrams or packets).
- Full duplex service: This means that the communication can take place in both directions at the same time.
- Connection oriented service: It defines 3 different phases: Connection establishment, Data transfer, Connection termination. This is a logical connection, not physical.

TCP

- Reliability: TCP is reliable as it uses checksum for error detection, attempts to recover lost or corrupted packets by re-transmission, acknowledgement policy and timers. It uses features like byte number and sequence number and acknowledgement number so as to ensure reliability. Also, it uses congestion control mechanisms.
- Multiplexing: TCP does multiplexing and demultiplexing at the sender and receiver ends respectively as a number of logical connections can be established between port numbers over a physical connection.
- Sequence number and Acknowledgement number: All the data bytes that are to be transmitted are numbered and the beginning of this numbering is arbitrary.

TCP Segment Header



- The header of a TCP segment can range from 20-60 bytes.
- 40 bytes are for options.
- If there are no options, header is of 20 bytes else it can be of utmost 60 bytes.

Source: Andrew S. Tanenbaum and David J. Wetherall. 2010. Computer Networks (5th. ed.). Prentice Hall Press, USA.

TCP Segment Header

- Source Port Address – 16 bit field that holds the port address of the application that is sending the data segment.
- Destination Port Address – 16 bit field that holds the port address of the application in the host that is receiving the data segment.
- Sequence Number – 32 bit field that holds the sequence number, i.e., the byte number of the first byte that is sent in that particular segment. It is used to reassemble the message at the receiving end if the segments are received out of order.
- Acknowledgement Number – 32 bit field that holds the acknowledgement number, i.e., the byte number that the receiver expects to receive next. It is an acknowledgment for the previous bytes being received successfully.
- Header Length (HLEN) – This is a 4 bit field that indicates the length of the TCP header by number of 4-byte words in the header, i.e, if the header is of 20 bytes(min length of TCP header), then this field will hold 5 (because 5 x 4 = 20) and the maximum length: 60 bytes, then it'll hold the value 15(because 15 x 4 = 60). Hence, the value of this field is always between 5 and 15.

TCP Segment Header

- Control flags – These are 8 1-bit control bits that control connection establishment, connection termination, connection abortion, flow control, mode of transfer etc. Their function is:

  CWR and ECE are used to signal congestion when ECN (Explicit Congestion Notification) is used.

  - ECE is set to signal an ECN-Echo to a TCP sender to tell it to slow down when the TCP receiver gets a congestion indication from the network.
  - CWR is set to signal Congestion Window Reduced from the TCP sender to the TCP receiver so that it knows the sender has slowed down and can stop sending the ECN-Echo
  - URG: Urgent pointer is valid
  - ACK: Acknowledgement number is valid(used in case of cumulative acknowledgement)
  - PSH: Request for push
  - RST: Reset the connection
  - SYN: Synchronize sequence numbers
  - FIN: Terminate the connection

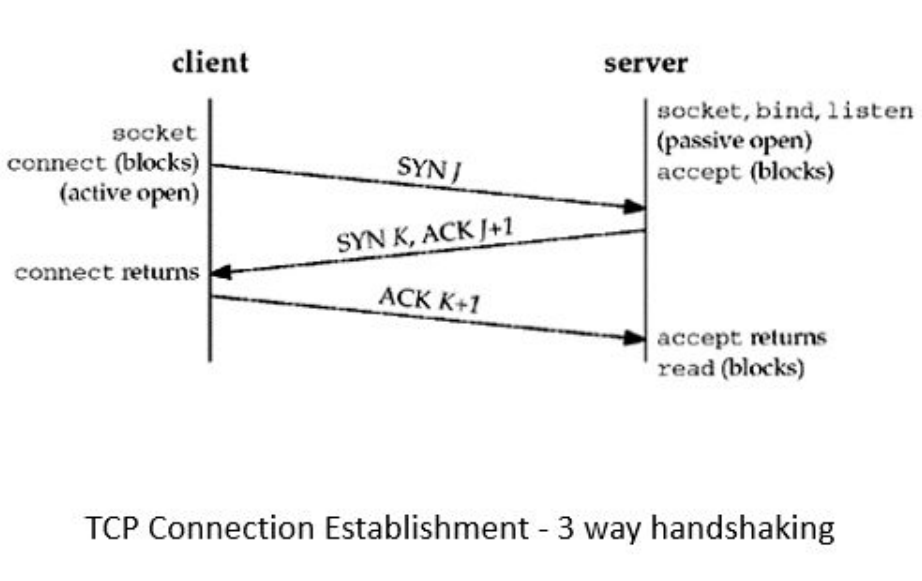
TCP Segment Header

Window size – It contains the size of the receiving window of the sender. It advertises how much data (in bytes) the sender can receive without acknowledgement. Thus, window size is used for Flow Control.

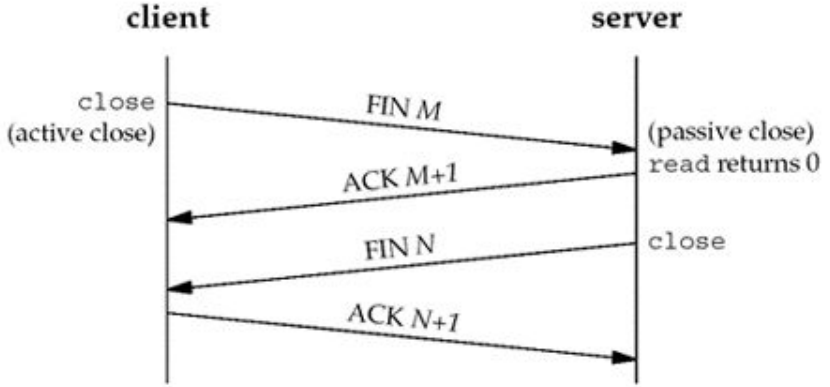
Checksum – This field holds the checksum for error control. It is mandatory in TCP as opposed to UDP.

Urgent pointer – This field (valid only if the URG control flag is set) indicates how much data in the current segment counting from the first data byte is urgent. Urgent pointer added to the sequence number indicates the end of urgent data byte

TCP Connections



TCP Connection Establishment - 3 way handshaking



TCP Connection Termination - 4 way handshaking

Source: poincare.matf.bg.ac.rs/~ivana/courses/tos/sistemi_knjige/pomocno/unp/UNP/0131411551_ch02lev1sec6.html.

Timers

Retransmission Timer – To retransmit lost segments, TCP uses retransmission timeout (RTO). Jacobson's Algorithm, Karn's modification are used for computing the value of time out timer dynamically.

Persistent Timer – To deal with a zero-window-size deadlock situation. Sender sends a probe(1B dummy data) every 60 seconds until the receiver responds with a non-zero window size

Keep Alive Timer – To prevent a long idle connection between two TCPs. If a client opens a TCP connection to a server transfers some data and crashes, the connection remains open forever. So a keepalive timer is used. Usually, it is set to 2 hours. After 2 hours a probe is sent. If there is no response after 10 probes, each of which is 75 s apart, it assumes that the client is down and terminates the connection.

Time Wait Timer – Used during TCP connection termination. The timer starts after sending the last Ack for 2nd FIN and closing the connection. It is used to receive any delayed segments form the server. It is set to twice the lifetime of a segment.

Error Control

Error control in TCP is mainly done through use of three simple techniques :

1. Checksum - If the segment is corrupted, then that segment is discarded by the destination TCP and is considered as lost.
2. Acknowledgement
3. Retransmission – Segments are retransmitted only during two events: when the sender receives three duplicate acknowledgements (ACK) or when a retransmission timer expires.
    1. Retransmission after RTO : An RTO timer for all sent but not acknowledged segments. When the timer runs out of time, the earliest segment is retransmitted. In TCP, RTO value is dynamic in nature and it is updated using round trip time (RTT) of segments. RTT is the time duration needed for a segment to reach receiver and an acknowledgement to be received to the sender. This method works well when RTO is small.
    2. Retransmission after Three duplicate ACK segments : Sometimes one segment is lost and the receiver receives so many out-of-order segments that they cannot be saved. In order to solve this situation, three duplicate acknowledgement method is used and missing segment is retransmitted immediately instead of retransmitting already delivered segment.

Flow Control

- TCP uses sliding window protocol for flow control.
- The window size is variable, and is controlled by the receiver.
- It is similar to GBN as no NAKs are used
- It is also similar to Selective repeat as out-of-order segments are not discarded.
- TCP uses cumulative acknowledgment.

Congestion

What is **congestion**?

A state occurring in network layer when the message traffic is so heavy that it slows down network response time.

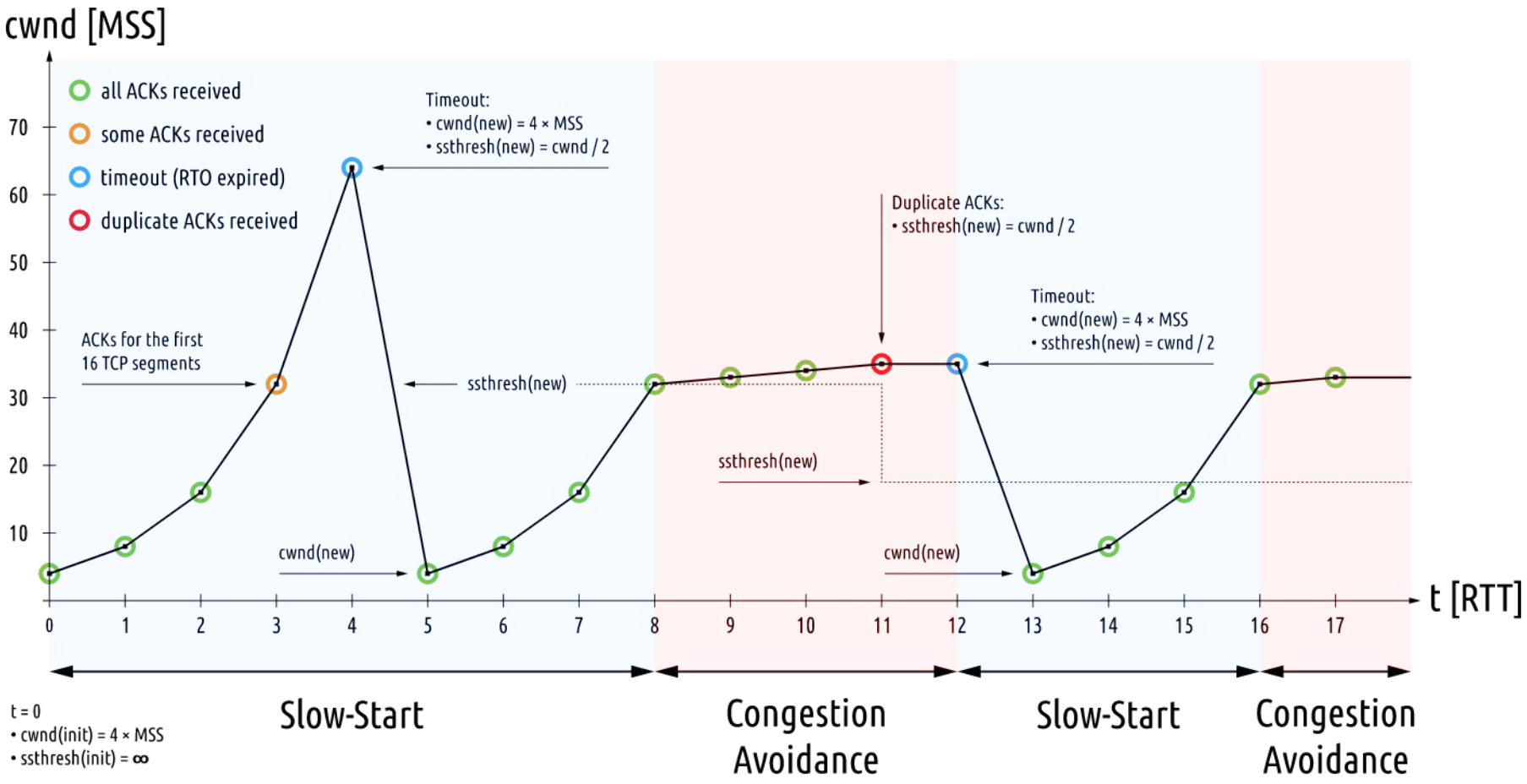
Effects of Congestion

- As delay increases, performance decreases.
- If delay increases, retransmission occurs, making situation worse.

Congestion Control

Additive Increase Multiplicative Decrease (AIMD) algorithm is used for congestion control. It has three phases

- Slow Start Phase: starts slowly – increment of congestion window is exponential till threshold is reached
  - Each time an ACK is received by the sender, the congestion window is increased by 1 segment. So, CWND approximately doubles on every RTT
- Congestion Avoidance Phase: After reaching the threshold, the congestion window is incremented by 1 for every RTT
- Congestion Detection Phase: multiplicative decrement – If congestion occurs, sender goes back to Slow start phase or Congestion avoidance phase. Congestion is detected through retransmission
  - Case 1 : Retransmission due to Timeout – congestion possibility is high. (Fast retransmission)
    (a) ssthresh = cwnd / 2
    (b) set cwnd = Initial value
    (c) start with slow start phase again.
  - Case 2 : Retransmission due to 3 duplicate Acks – congestion possibility is less. (Fast recovery)
    (a) cwnd = ssthresh = cwnd / 2
    (b) start with congestion avoidance phase



Source: https://witestlab.poly.edu/blog/tcp-congestion-control-basics/

Assignment 2 - Due Date: March 26, 2021, 11:59 PM

Parameter combinations for the technical report:

$K_i \in \{1,4\}$

$K_m \in \{1,1.5\}$

$K_n \in \{0.5,1\}$

$K_f \in \{0.1,0.3\}$

$P_s \in \{0.01,0.0001\}$