CS 925

Lecture 10

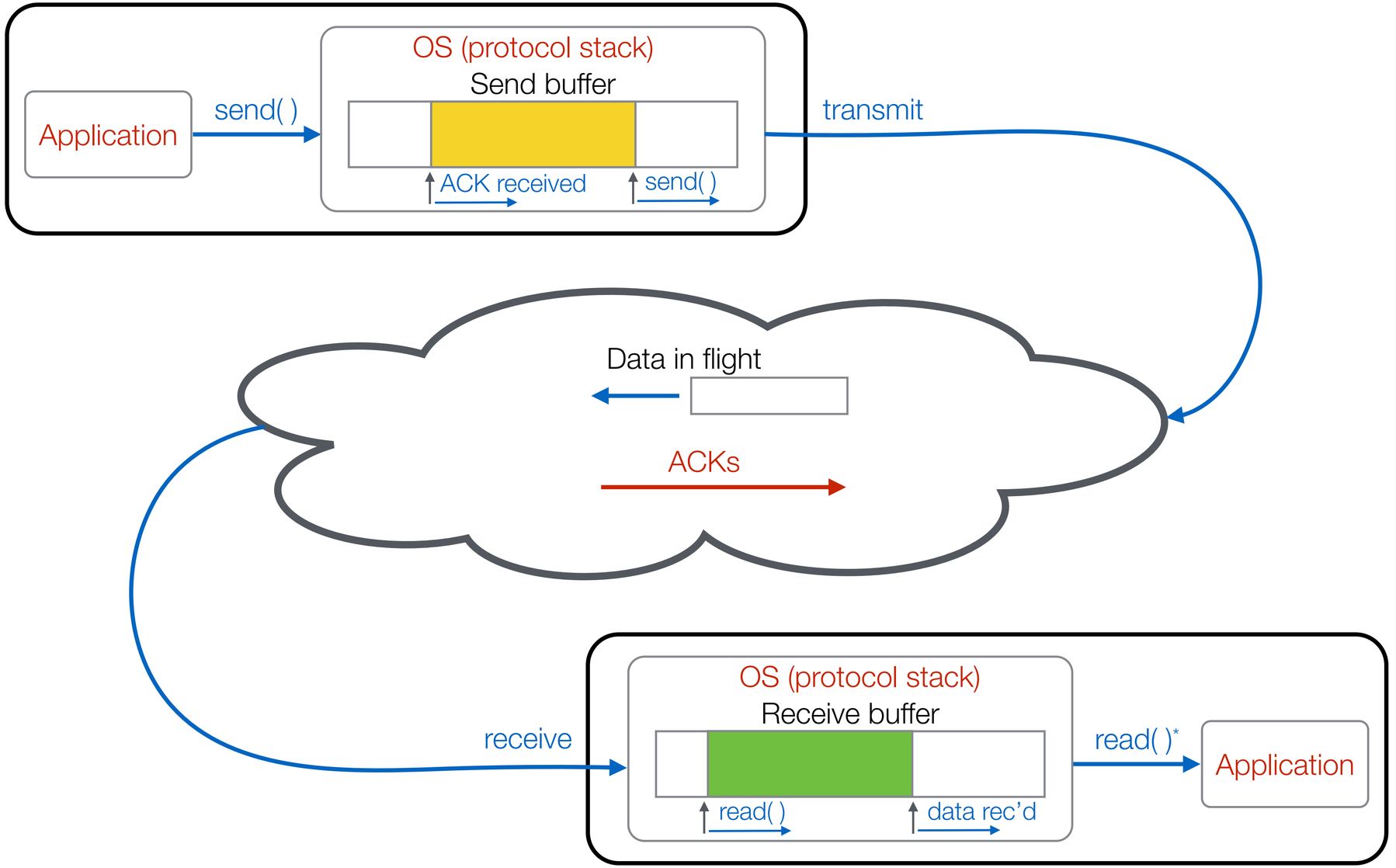
TCP Congestion Control

Thursday, February 22, 2024

TCP Congestion Control

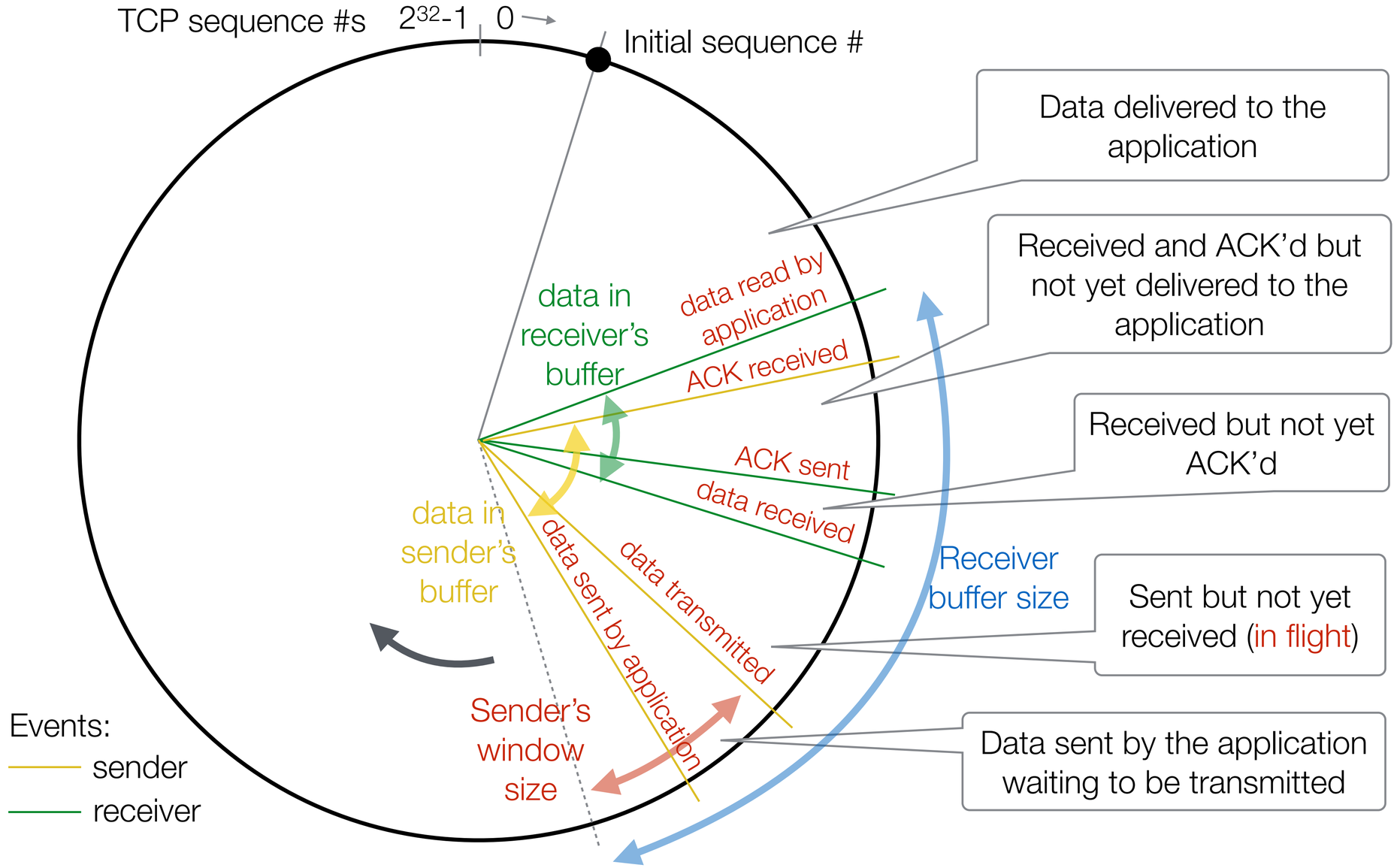
- Receiver congestion (flow control)
  - **Window Size field** - explicitly reported by the receiver
  - TCP Window Scale Option
- Network congestion
  - **Retransmission timeout**
  - **Transmission window**
  - Set based on observed RTT and on detected packet loss

TCP buffering and data flow

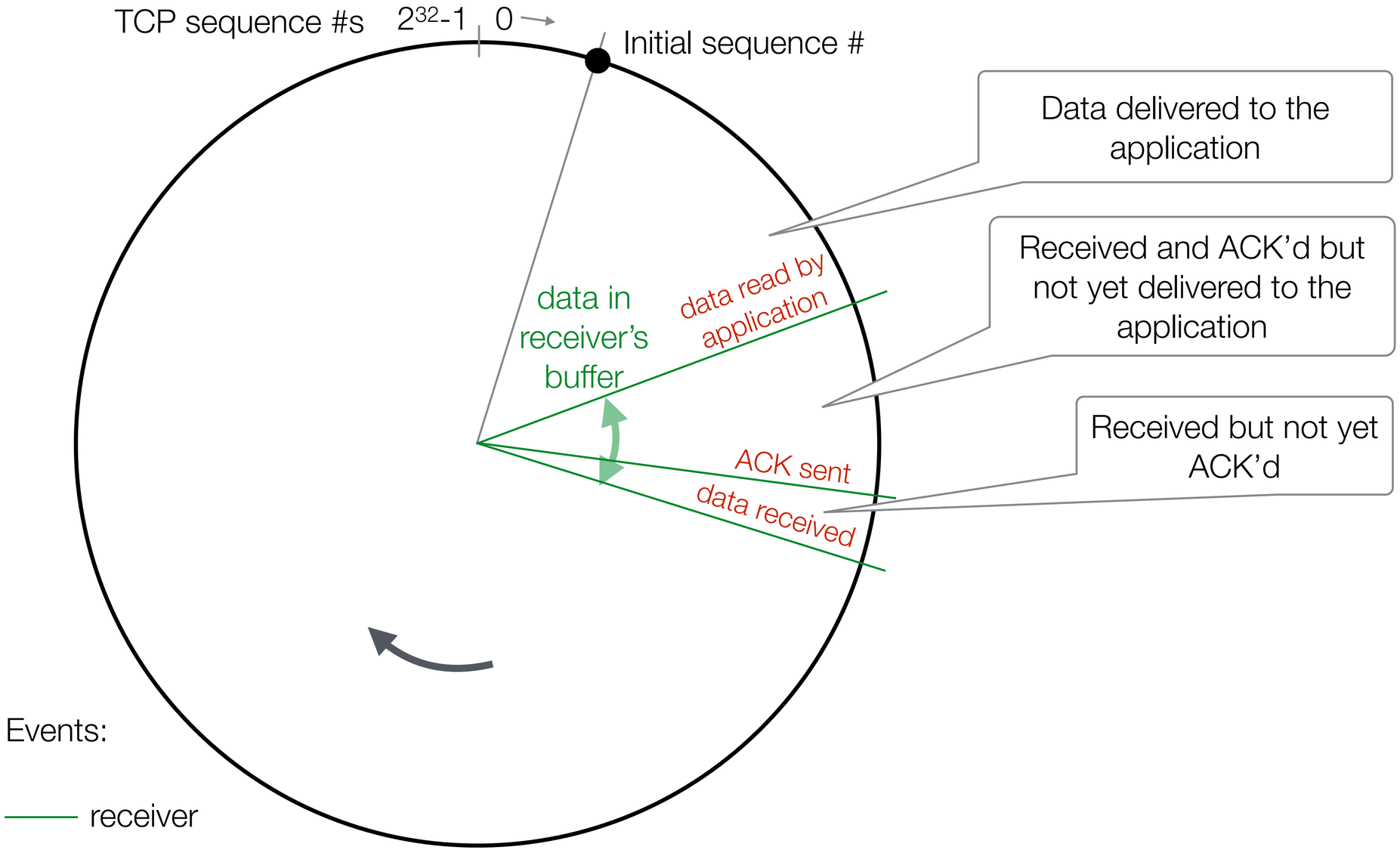


(*) many APIs call the read() operation "receive" (eg: recv()), read is used here to avoid confusion with receiving data on an interface

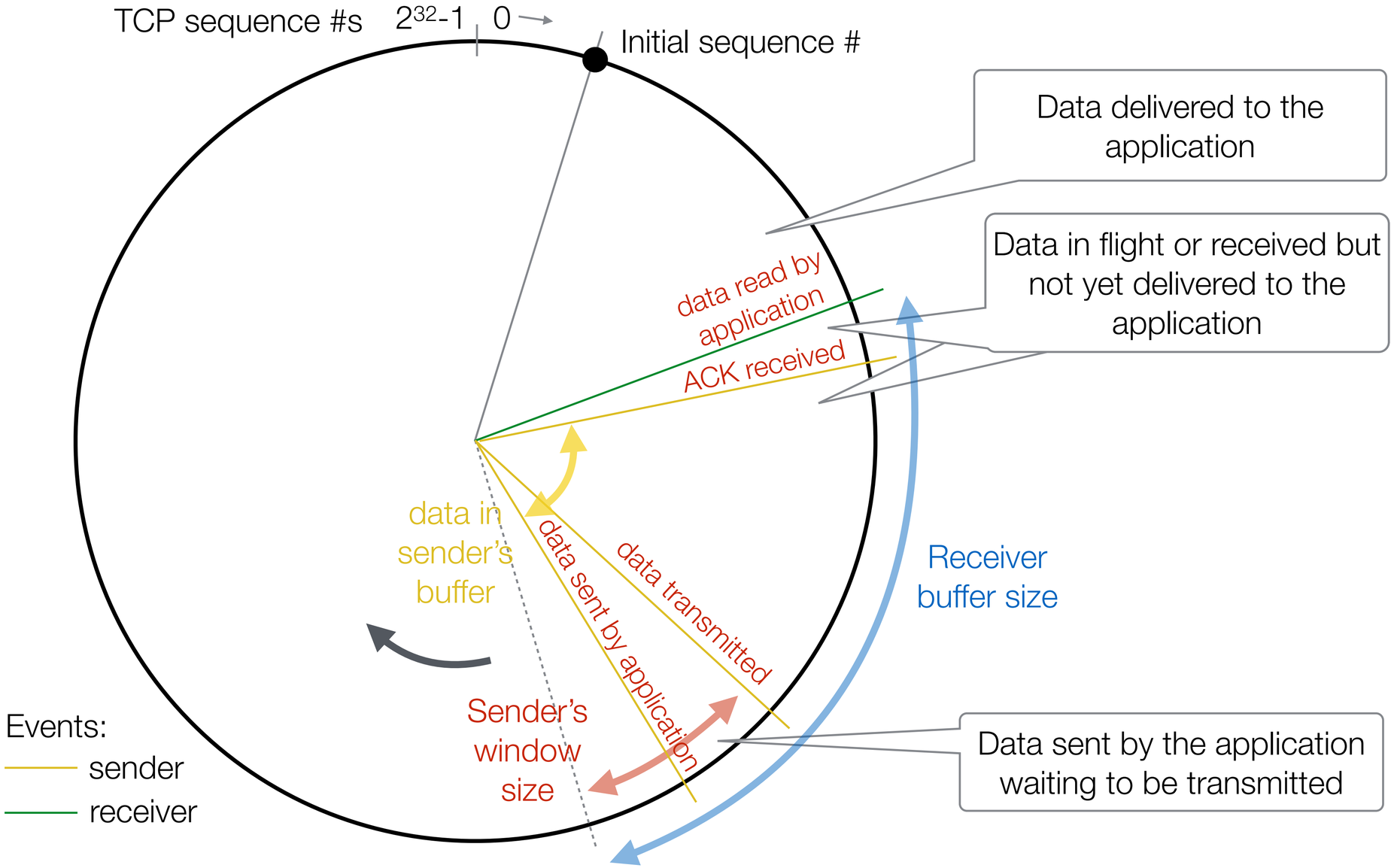
TCP Sliding Window



TCP Sliding Window

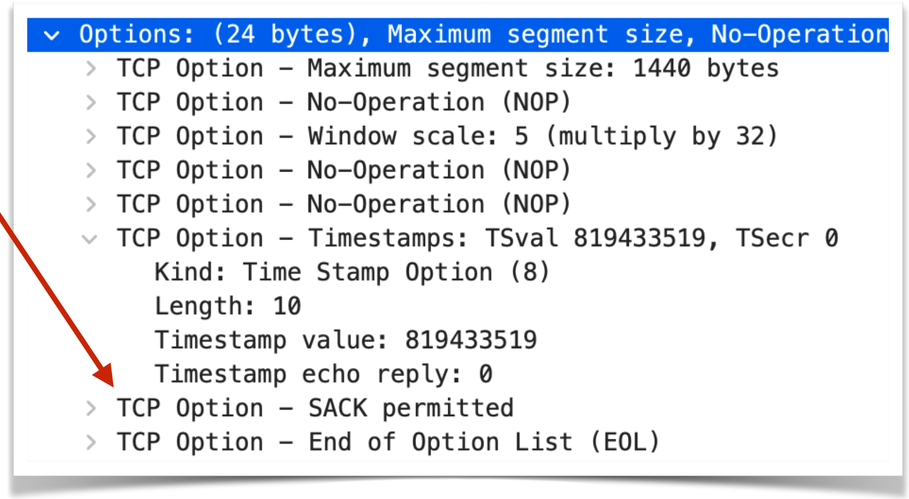


TCP Sliding Window



TCP SACK

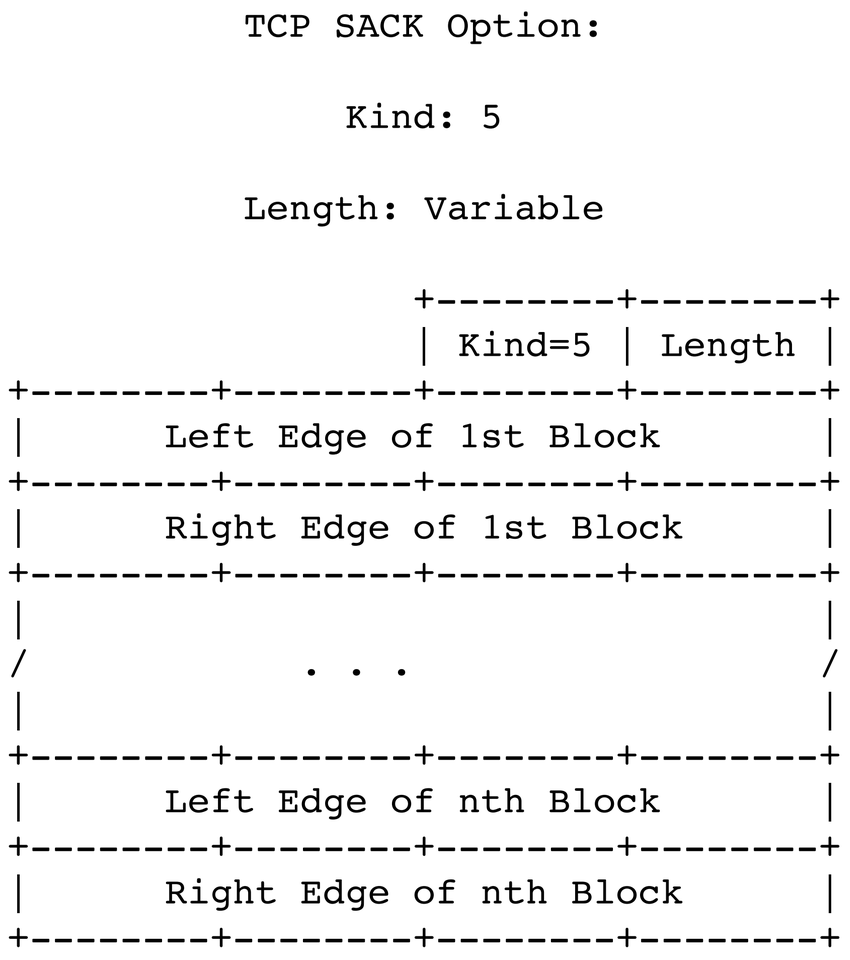
- TCP Selective Acknowledgment (SACK)
  - negotiated during connection open
  - additional info about received data on top of TCP's cumulative acknowledgment (ACK# field)
  - a TCP option negotiated when a connection is opened



Wireshark decode of a TCP SYN packet

TCP SACK

- TCP Selective Acknowledgment (SACK)
  - negotiated during connection open
  - additional info about received data on top of TCP's cumulative acknowledgment (ACK# field)
  - a TCP option negotiated when a connection is opened



Source: RFC2018

Network Congestion Control

- ACK self-clocking
- Retransmission timer management
- Additive Increase Multiplicative Decrease (AIMD)
- Slow start mechanism
- ...

Retransmission Timeout

RFC 6298

Initialization:

RTO ← 1 sec

After the first measurement:

SRTT ← R

RTTVAR ← R/2

RTO ← SRTT + max (G, K * RTTVAR)

After subsequent measurements:

RTTVAR ← (1 - beta) * RTTVAR + beta * |SRTT - R'|

SRTT ← (1 - alpha) * SRTT + alpha * R'

RTO ← SRTT + max (G, K * RTTVAR)

Where:

R - first RTT measurement
R' - subsequent RTT measurement
RTTVAR - RTT variance
SRTT - smoothed RTT estimate
RTO - retransmission timeout
G - clock granularity

Recommended values:

alpha=1/8, beta=1/4, K=4

Exponential Back-off

RTO after a timeout:

RTO ← q * RTO

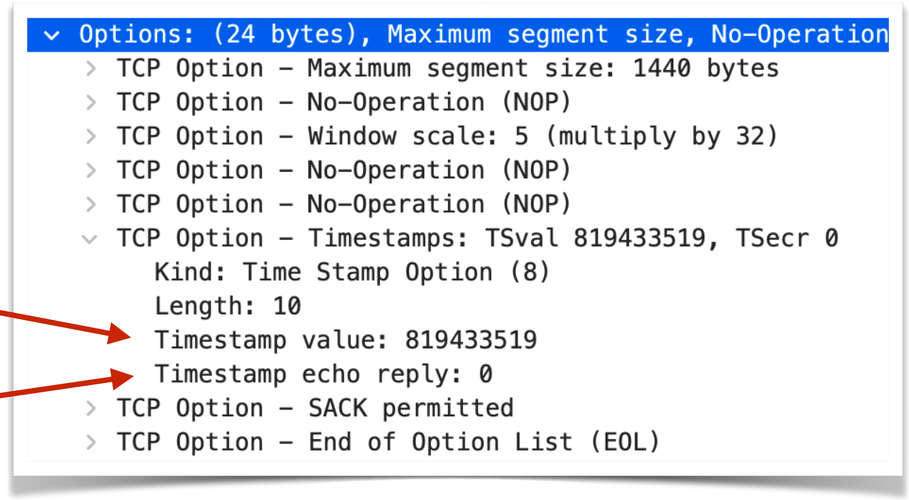
Recommended value: q = 2

This a congestion control mechanism since retransmissions are delayed after packet loss detected. The delay is increasing exponentially with more packet losses.

TCP Timestamp

- Question:
  - ACK for what?

- RTTM - RTT Management
  - TCP option, two 4-byte values
  - *TS value* (TSval) - current “timer” value
  - *TS echo reply value* (TSecr) - most recently received TSval (only if it acknowledges new data)



Wireshark decode of a TCP SYN packet