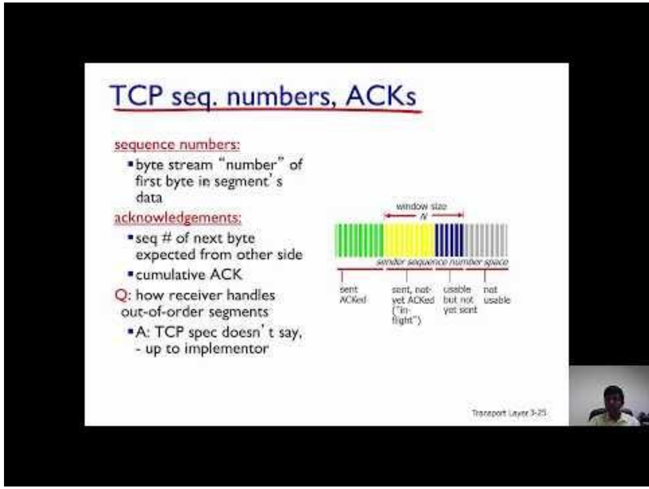
Computer Networks

Transmission Control Protocol

Jianping Pan
Fall 2022



TCP seq. numbers, ACKs

sequence numbers:

- byte stream "number" of first byte in segment's data

acknowledgements:

- seq # of next byte expected from other side
- cumulative ACK

Q: how receiver handles out-of-order segments

- A: TCP spec doesn't say, - up to implementor

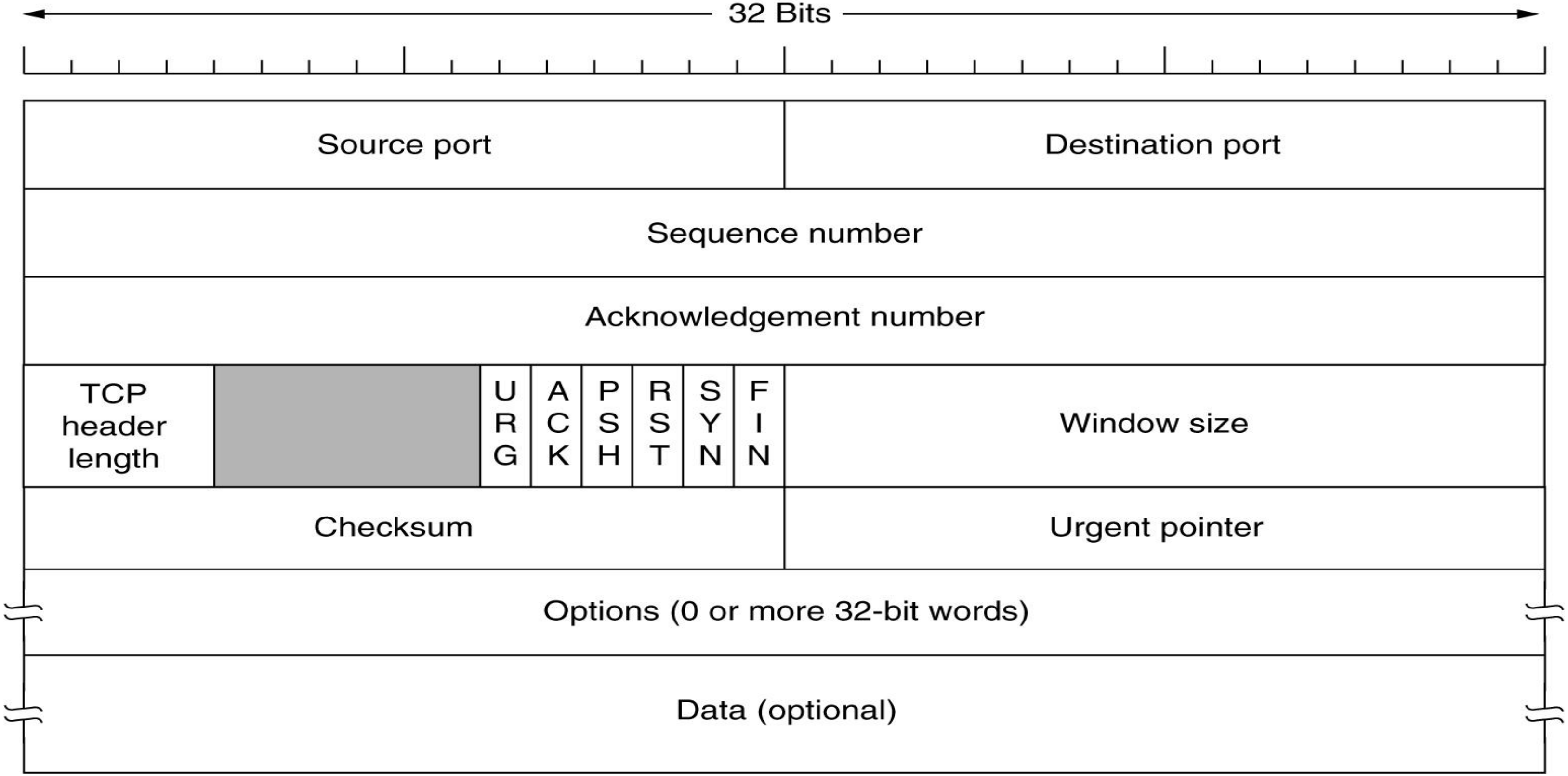
window size
N
sender sequence number space
sent ACKed
sent, not-yet ACKed ("in-flight")
usable but not yet sent
not usable

Transport Layer 3-25

TCP

- Transmission control protocol [RFC793]
- Services provided by TCP
  - connection-oriented, point-to-point, bi-directional
  - reliable, in-sequence, stream-like
- Services provided by IP
  - packets: duplicated, lost, reordered, corrupted
- TCP protocol mechanisms
  - connection management
  - flow, error and congestion control

TCP packet header



32 Bits

Source port

Destination port

Sequence number

Acknowledgement number

TCP header length

URG

ACK

PSH

RST

SYN

FIN

Window size

Checksum

Urgent pointer

Options (0 or more 32-bit words)

Data (optional)

Port number

- TCP port number (16-bit)
  - source, destination port numbers
    - addressing and multiplexing
- Port number allocation (ref: iana.org)
  - well-known port numbers (0~1023, privileged)
    - e.g., 80: http; 443: https
  - registered port numbers (/etc/services)
    - http-alt   8080/tcp
  - dynamically allocated port numbers

bind(), connect()

TCP connection ID

- TCP connections
  - connection: initiator, responder
    - (initiator IP, *initiator port*, responder IP, responder port)
- One connection: one flow in each direction
  - for each flow: source, destination
    - (source IP, source port, destination IP, destination port)
  - 5-tuple (or 4-tuple when protocol ID is implied)
    - (src IP, src port, protocol ID, dst IP, dst port)
- Socket, connection, flow

listen()

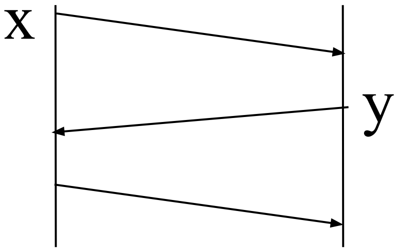
connect(), accept()

send(), recv()

Q: port reuse?

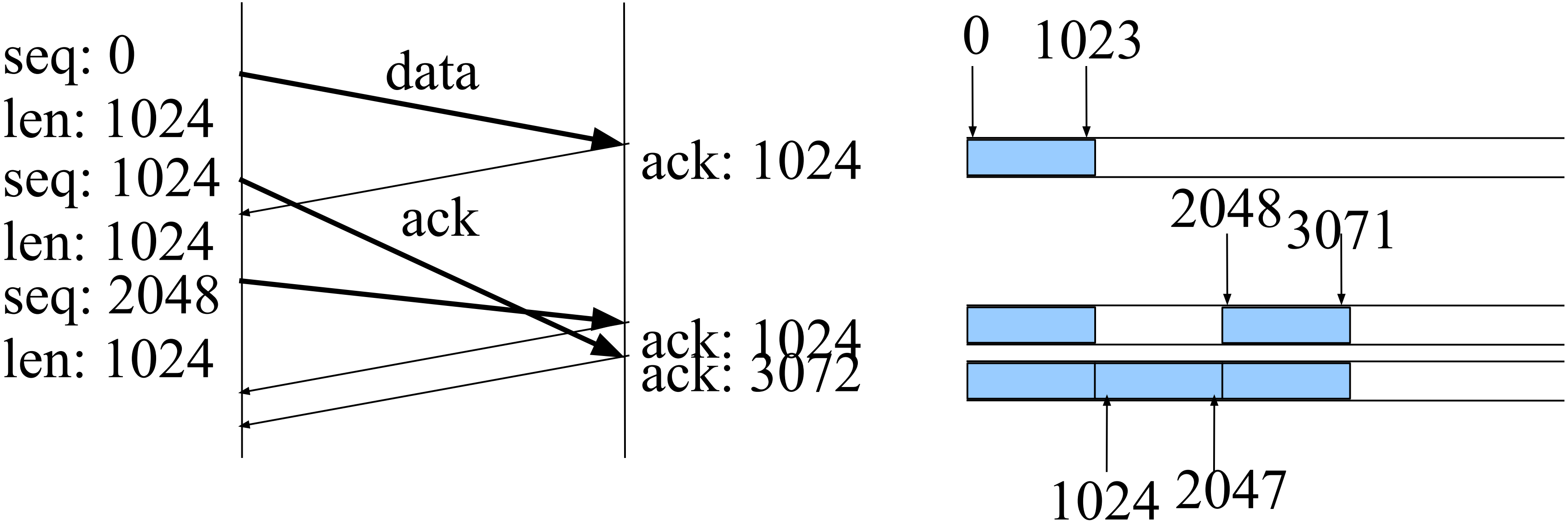
Sequence number

- TCP sequence number (32-bit)
  - byte sequence for the ***first*** byte in payload
    - exception: SYN/FIN sequence number
  - random initial sequence number
    - exchanged during 3-way handshake
  - sequence number rollover
- Acknowledgment number (32-bit)
  - byte sequence for the ***next*** byte to expect



Q: why random ISN?

Sequence vs acknowledgment



Q: byte vs packet sequence?

Header length/Data offset

- Variable-length header due to TCP options
- TCP header length/data offset (4-bit)
  - number of 32-bit words!
  - at least 5 for fixed TCP header fields *
  - maximum: 15
    - i.e., 40 bytes in total for TCP options
- TCP design feature: bit-alignment
  - fields of x-bit length at offset n*x (n: an integer)
  - 32-bit word

* why not start from 0 for 5 (or 20)?

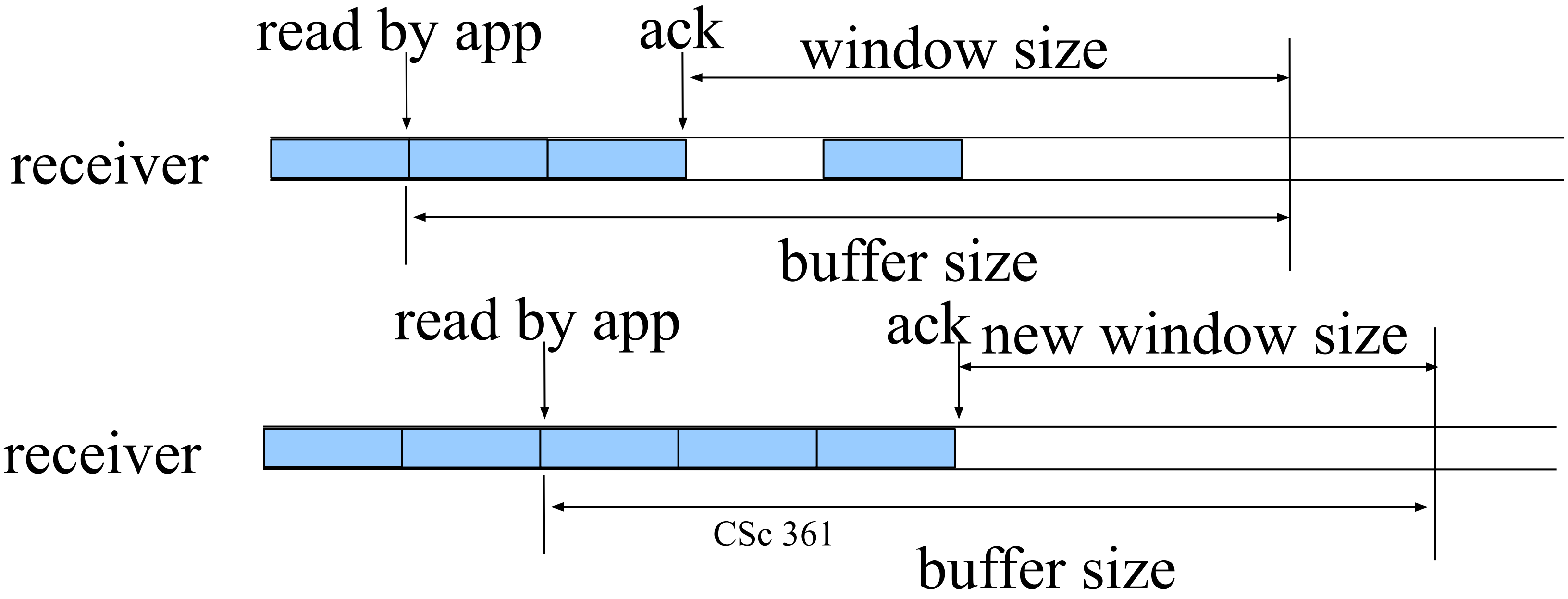
Q: why alignment?

Control flags

- URG: urgent pointer meaningful
- ACK: acknowledgment number meaningful
- PSH: logic message boundary
- RST: connection reset
- SYN: synchronization (conn. establishment)
- FIN: finish (graceful connection release)
  - stay tuned: “TCP connection management”

Window size

- TCP window size (16-bit)
  - stay tuned: “TCP flow control”



Checksum

- TCP checksum (16-bit)
  - “one's complement of one's complement sum”
    - stay tuned: “TCP error control”
  - cover
    - TCP header (including options, if any)
    - TCP payload
    - TCP pseudo header
      - source and destination IP address
      - protocol ID
      - TCP segment size

Q: why TCP pseudo header?

Stone, Jonathan, and Craig Partridge. "When the CRC and TCP checksum disagree." ACM SIGCOMM Computer Communication Review. Vol. 30. No. 4. ACM, 2000.

Urgent pointer

- TCP urgent pointer (16-bit)
  - offset of the LAST byte for urgent data
    - not (LAST+1) per RFC 1122: Host requirements
  - from the current sequence number!
  - for out-of-band (OOB) control information*
    - e.g., interrupt an ongoing file transfer
  - Socket interface
    - send(s, buf, len, MSG_OOB);
  - receiver should process the urgent data immediately

* can be overridden by MSG_OOBINLINE

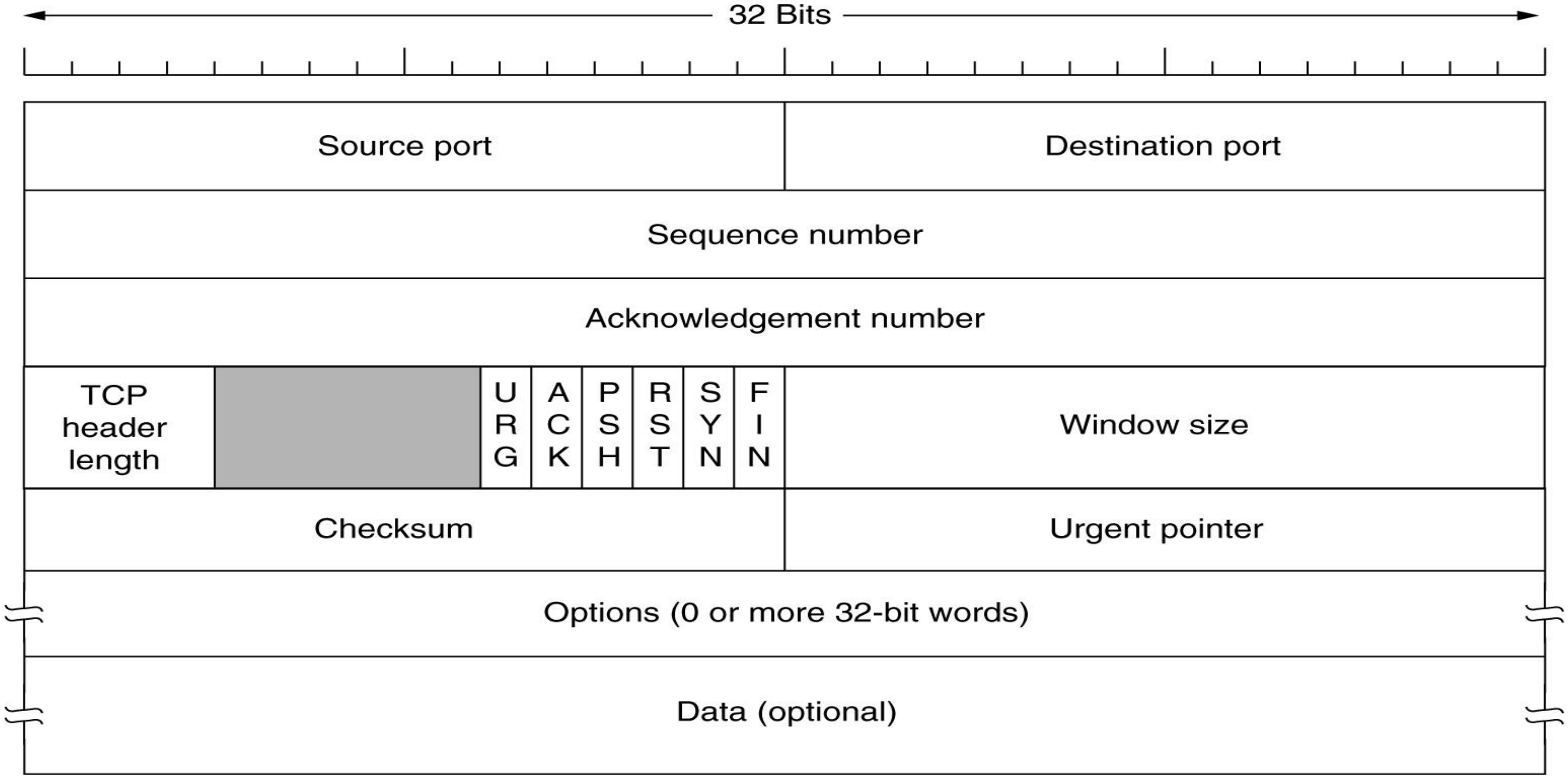
Q: URG vs PSH?

TCP options

- TLV-like options
  - option-kind: 1-byte
  - option-length: 1-byte, for the entire option
  - option-data:variable length
- E.g., Maximum Segment Size (MSS)
  - exchanged during connection establishment
  - default: 536 bytes
- E.g., Selective Acknowledgment (SACK)
  - stay tuned: “TCP congestion control”
- Zero-padding to keep 32-bit alignment

Q: big vs small segment?

TCP packet header (again)



32 Bits

Source port

Destination port

Sequence number

Acknowledgement number

TCP header length

URG

ACK

PSH

RST

SYN

FIN

Window size

Checksum

Urgent pointer

Options (0 or more 32-bit words)

Data (optional)

Q: any redundant fields?

This lecture

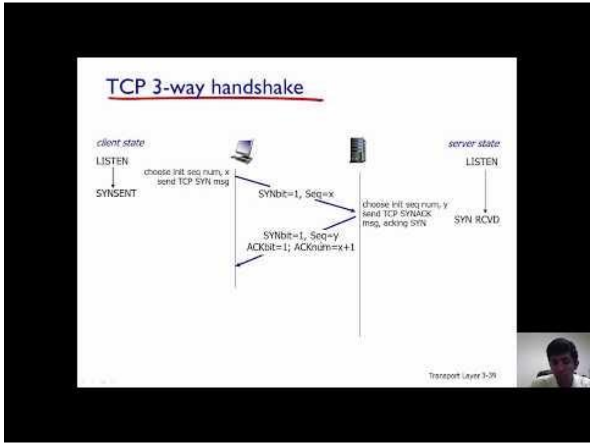
RIP, Sally!

- TCP
  - – services provided by TCP
    - protocol mechanisms to support TCP services
  - – TCP header fields
    - control information exchanged to support TCP
- Explore further
  - – Qs on previous slides
  - – the limitation of TCP and proposed changes
    - http://www.icir.org/floyd/tcp_small.html

“Research - TCP/IP over something - Air Links - Wireless Links. PAN Jianping's annotated pointers.” now at http://www.cs.uvic.ca/~pan/tcpair

Next lecture

- This Friday, Oct 7, 9:30am: M1
- Next week: TCP connection management
  - read KR4: Computer Networking
    - Chapter 3 (all sections required this month)



TCP 3-way handshake

Transport Layer 3-39

Midterm 1 (M1)

- Friday, October 7, 2022, 9:30am Victoria Time (GMT-7)
  - In ECS116 (sorry, the time/space has been determined with CAL already)
  - Let see how it goes, and whether we shall make adjustments for future ones
- Coverage
  - All lectures, tutorials and labs, and their corresponding text materials, slides, assignments, etc
  - From the beginning of the term to Wednesday, October 5, 2022, all inclusive
- Type of questions
  - Concepts questions: describe and compare related concepts
  - Role-playing questions: If I am the computer (network), what shall I do?
  - Play-a-role questions: If I was given the (real) problem to solve, what shall I do?
- Format of the exam
  - Friday 9:30am to 10:20am, Victoria time, ECS116
    - If you have CAL accommodation, the time duration (50 minutes) will be prorated at CAL
- Exams are closed book/notes/Internet, and have to be done by yourself alone