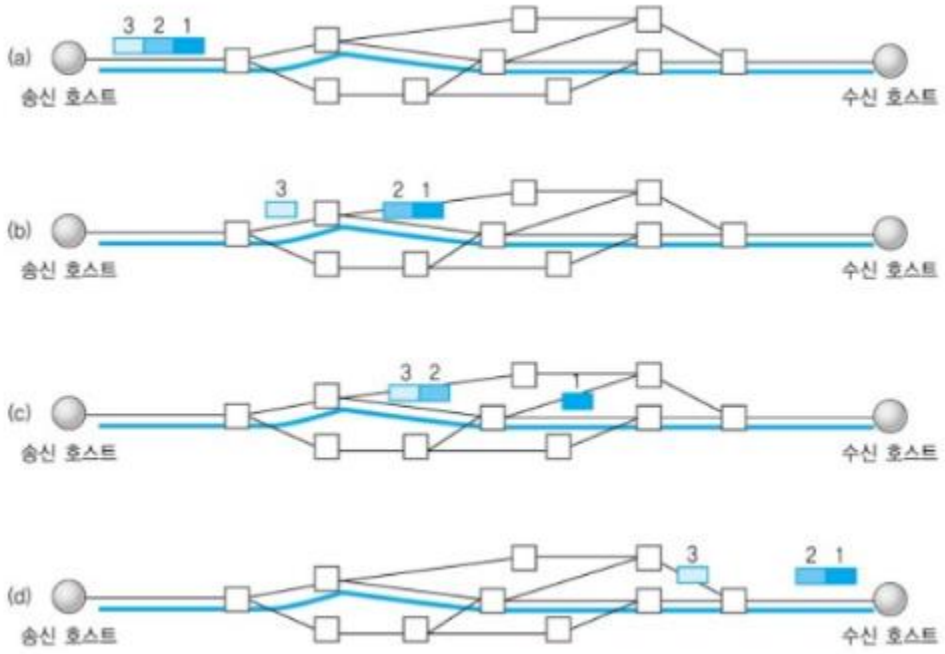
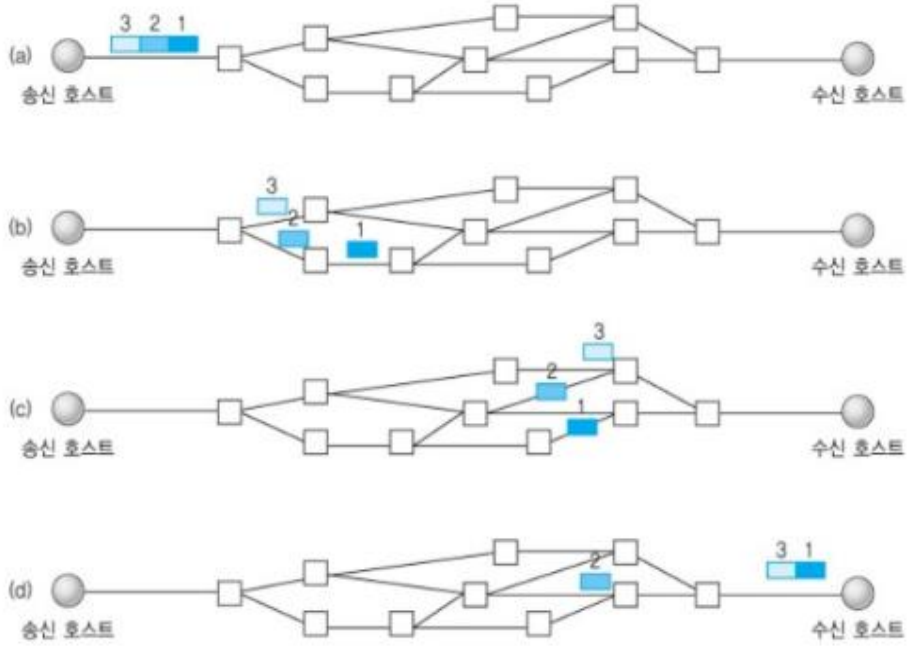
2.1 Introduction (Transport layer)

1. TCP



sophisticated, reliable byte-stream protocol . 연결형

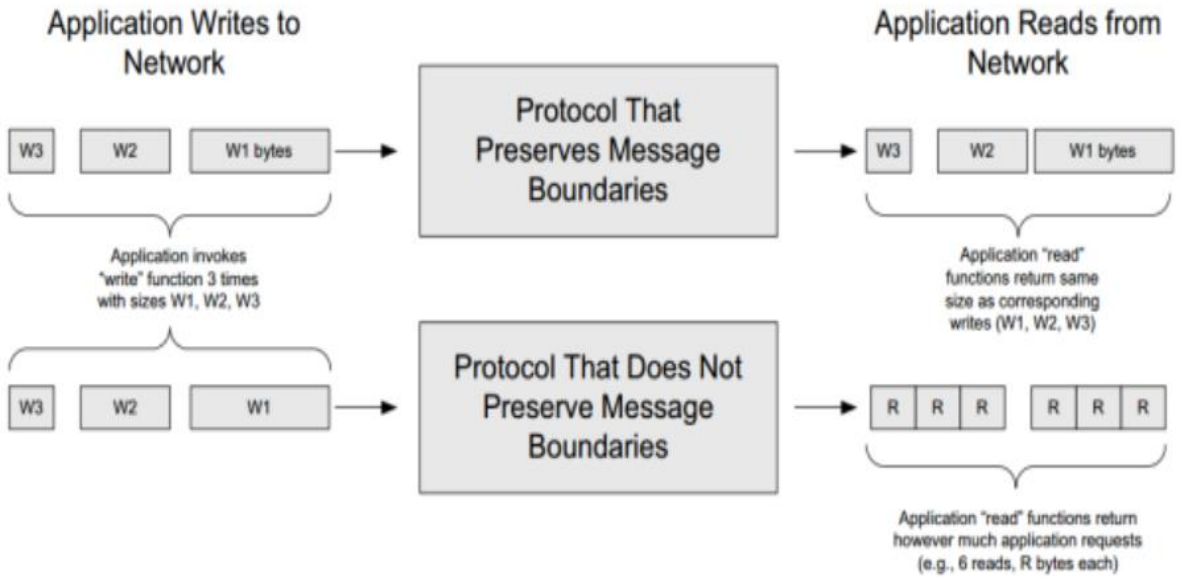
2. UDP



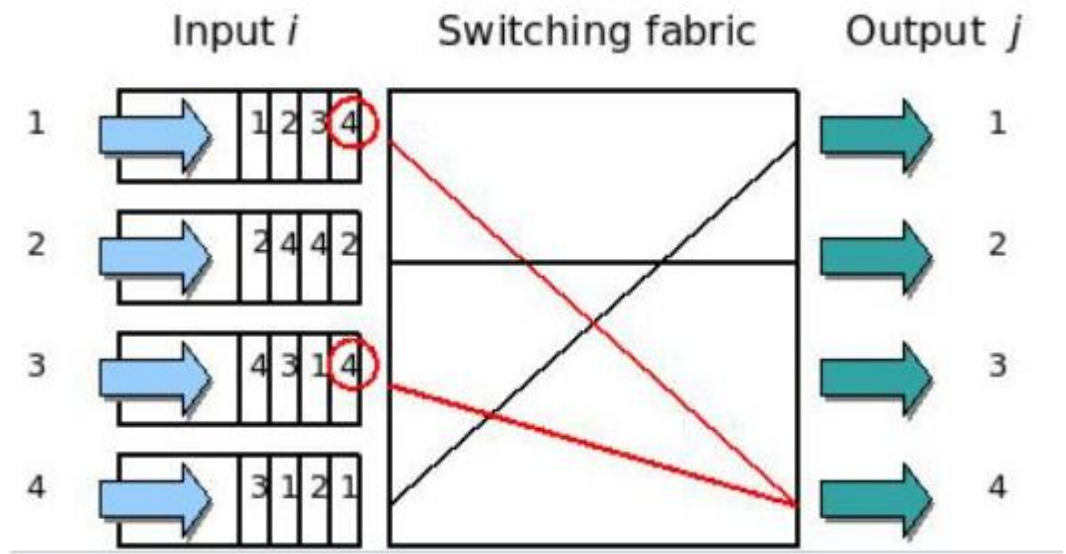
simple, unreliable datagram, speed, 비연결형

2.1 Introduction (Transport layer)

3. SCTP

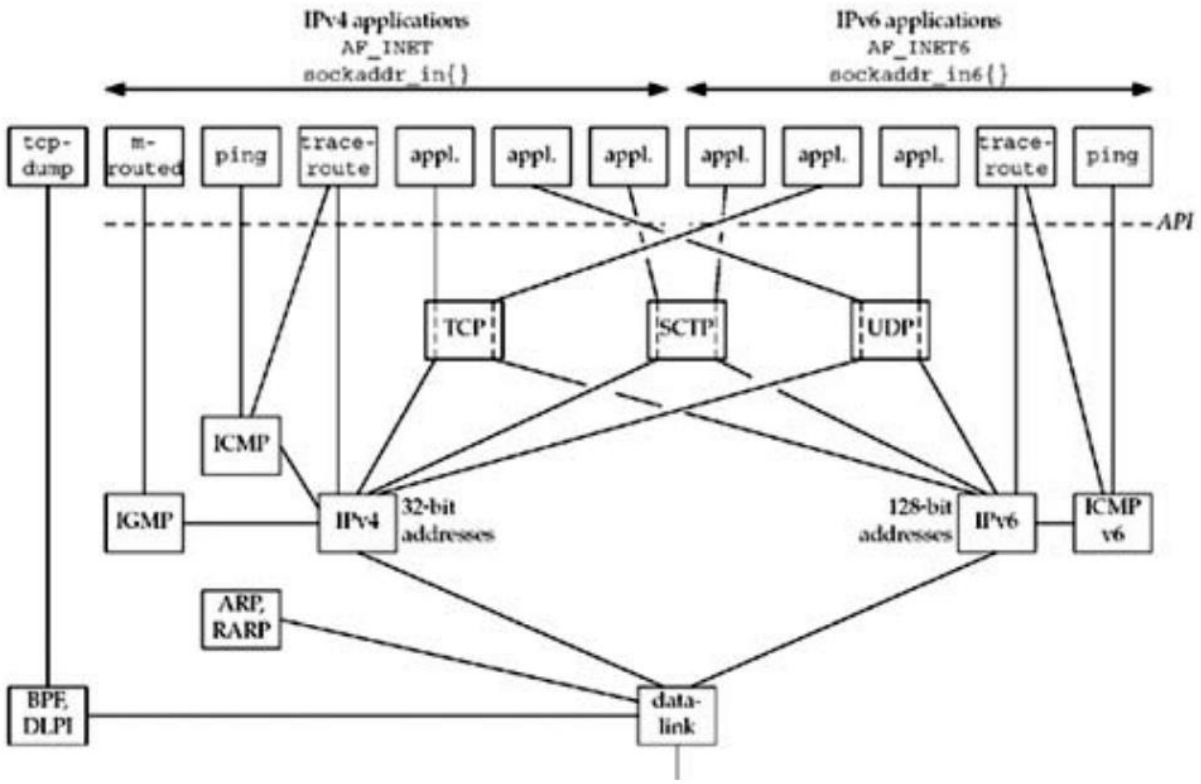


Message Boundaries



Head-Of-Line blocking

2.2 The Big Picture



	IPv4	IPv6
Address Family	AF_INET	AF_INET6
Structure	sockaddr_in	sockaddr_in6
Address	32bits	64bits

2.2 Protocol

-IPv4 : Internet Protocol version 4. It uses 32-bit addresses.
IPv4 provides packet delivery service for TCP, UDP, SCTP, ICMP, and IGMP.

-IPv6 : Internet Protocol version 6. It uses 128-bit addresses.
IPv4 provides packet delivery service for TCP, UDP, SCTP, and ICMPv6.

-TCP : TCP is a connection-oriented protocol that provides a reliable, full-duplex byte stream to its users.
It can use either IPv4 or IPv6.

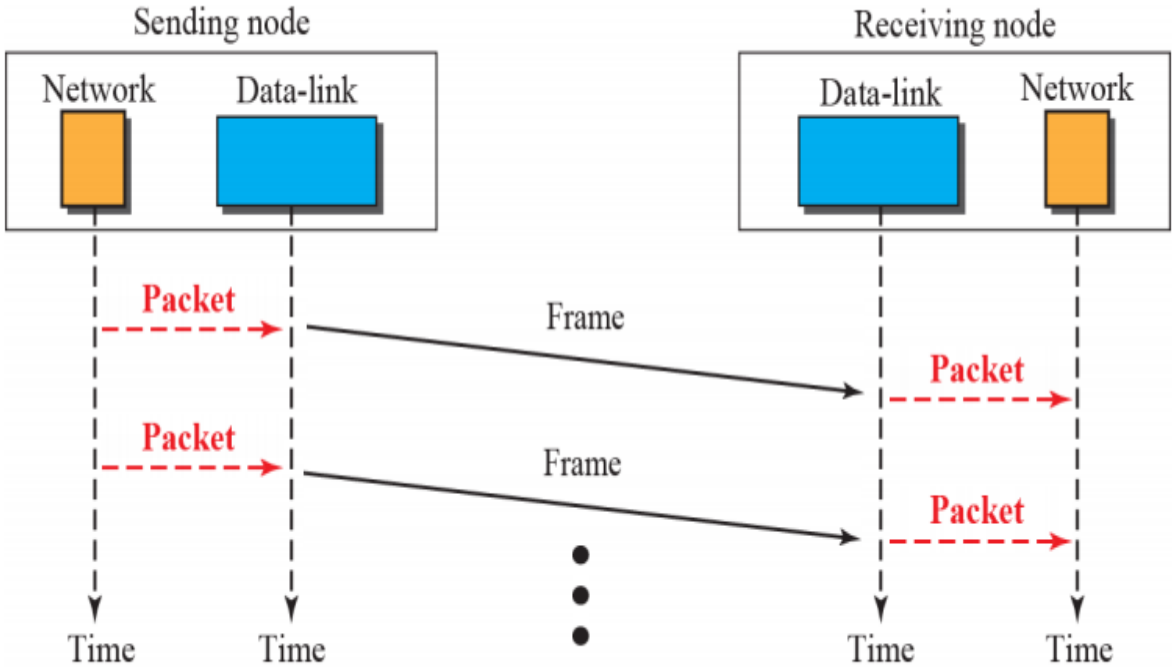
-UDP : UDP is a connectionless protocol that provides unreliable. It can use either IPv4 or IPv6.

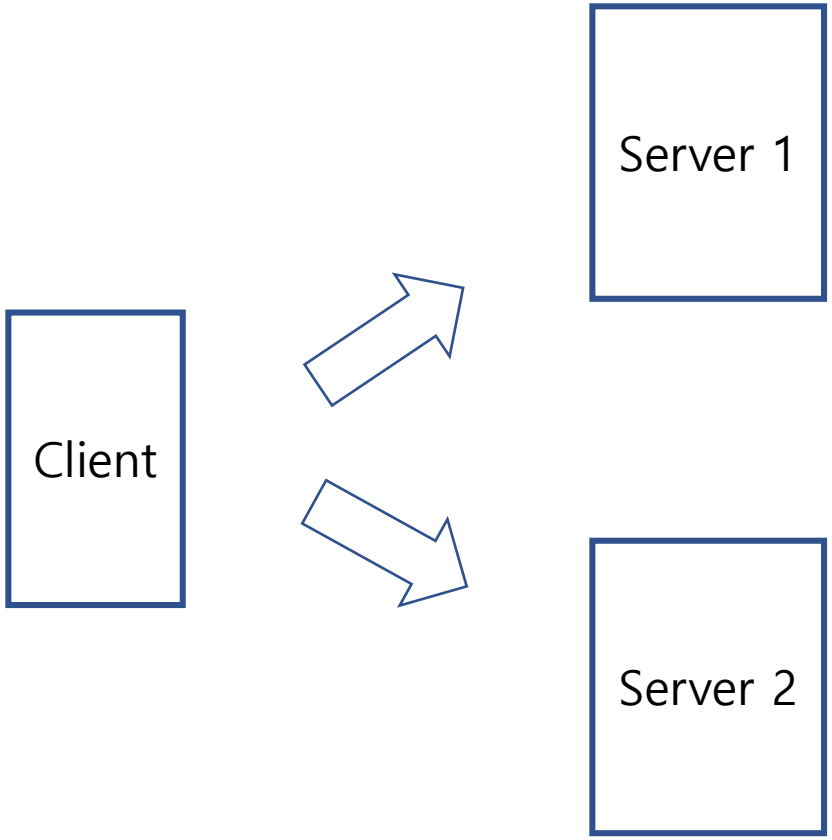
-SCTP : SCTP is a connection-oriented protocol that provides a reliable full-duplex association.
It is multihomed. It can use either IPv4 or IPv6.

-ICMP : ICMP handles error and control information between routers and hosts

Each Internet protocol is defined by one or more documents called a Request for Comments (RFC)

2.3 User Datagram Protocol (UDP)

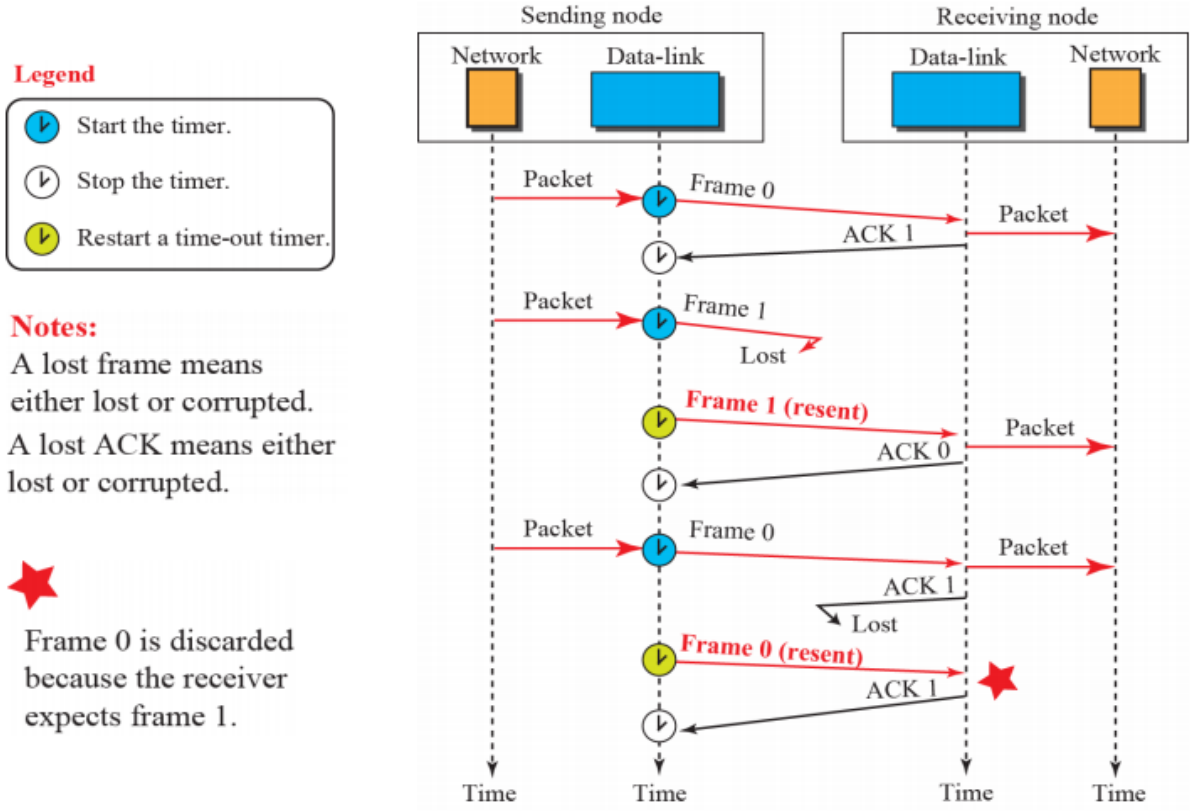




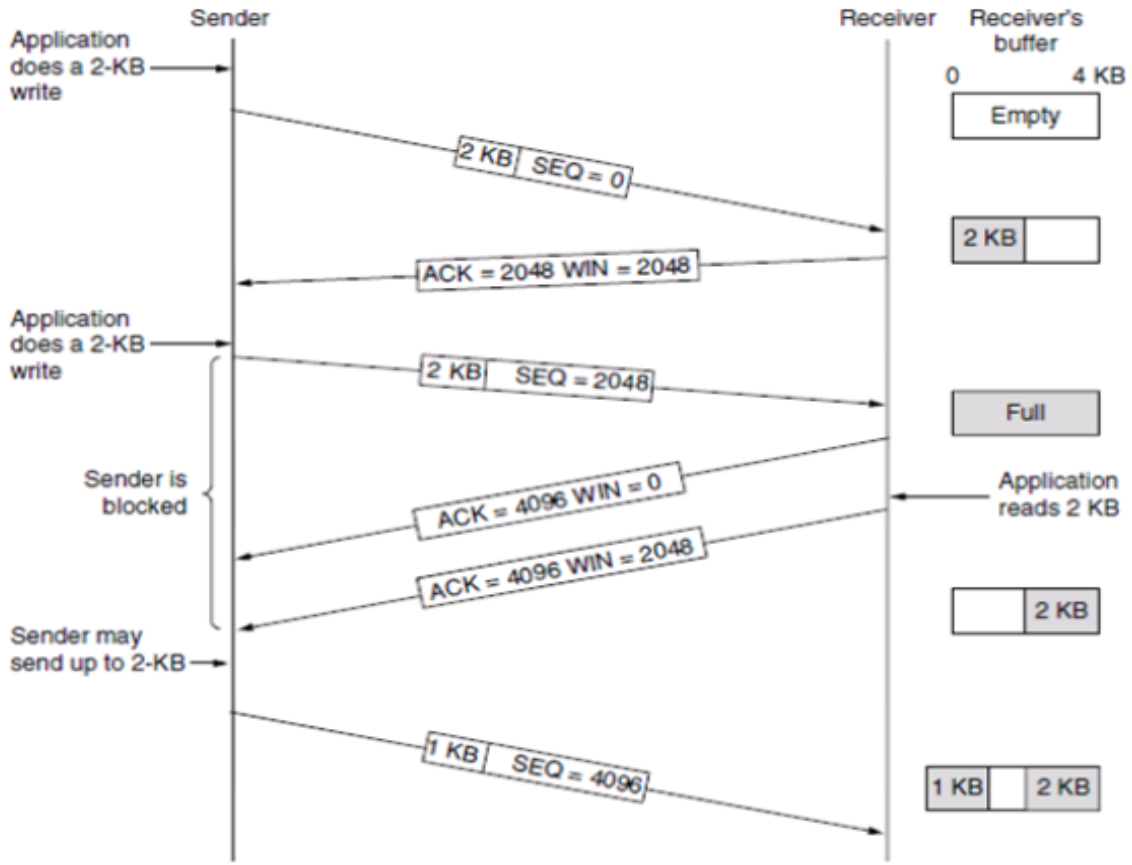
UDP is a simple transport-layer protocol.
It is lack of reliability.

UDP client can create a socket and send a datagram to a given server and then immediately send another datagram on the same socket to a different server.

2.4 Transmission Control Protocol (TCP)

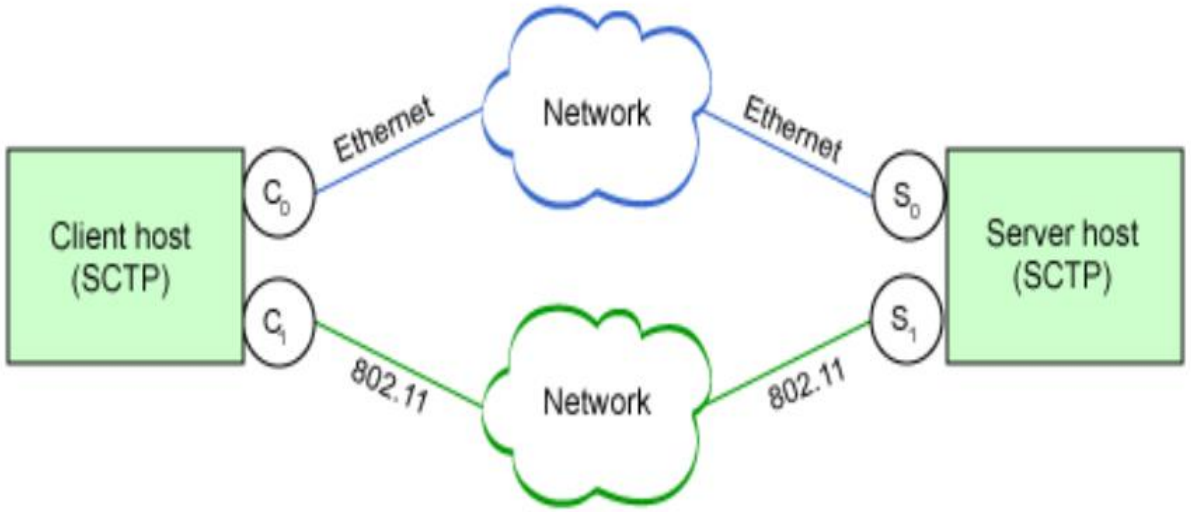


TCP provides reliability.
If TCP receives duplicate data from its peer, it can detect that the data has been duplicated, and discard the duplicate data.



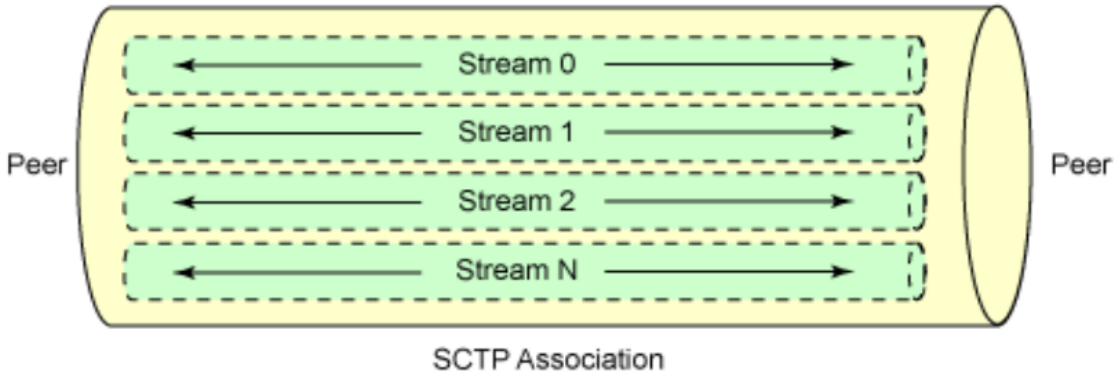
TCP provides flow control.

2.5 Stream Control Transmission Protocol (SCTP)



Multi-Homing

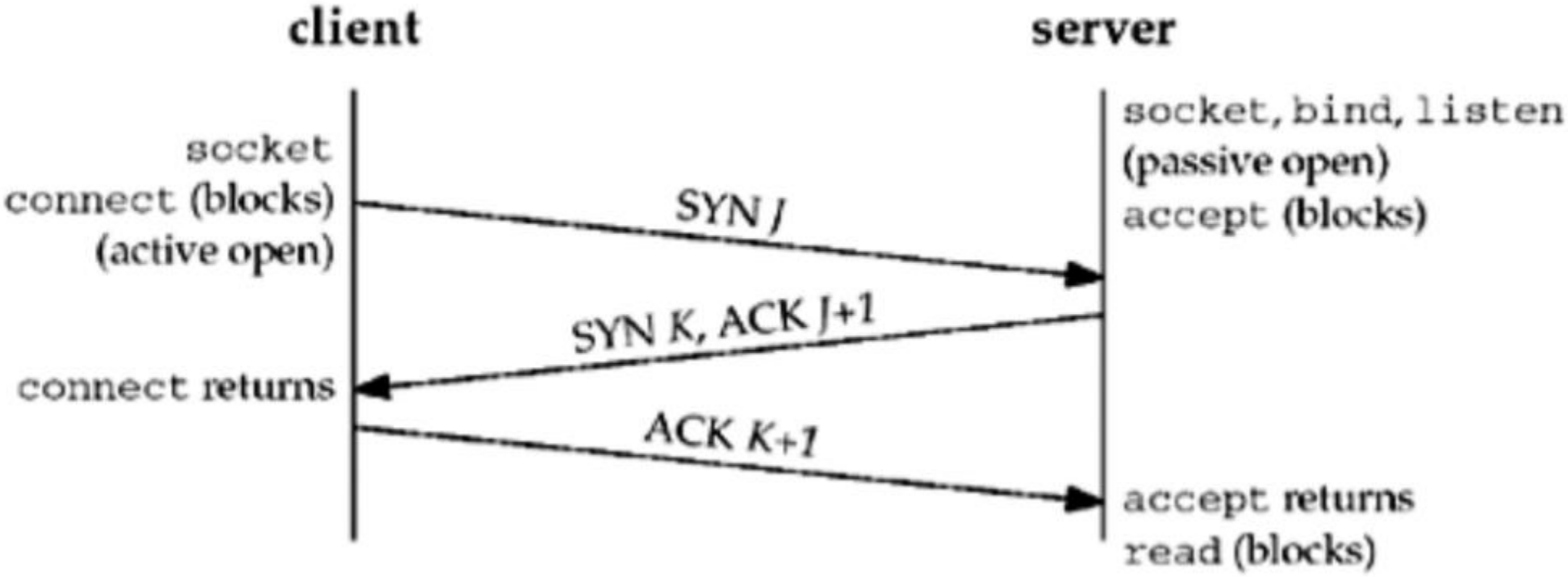
This feature can provide increased robustness against network failure. An endpoint can have multiple redundant network connections



Multi-Streaming

A lost message in one of these streams does not block delivery of messages in any of the other streams.

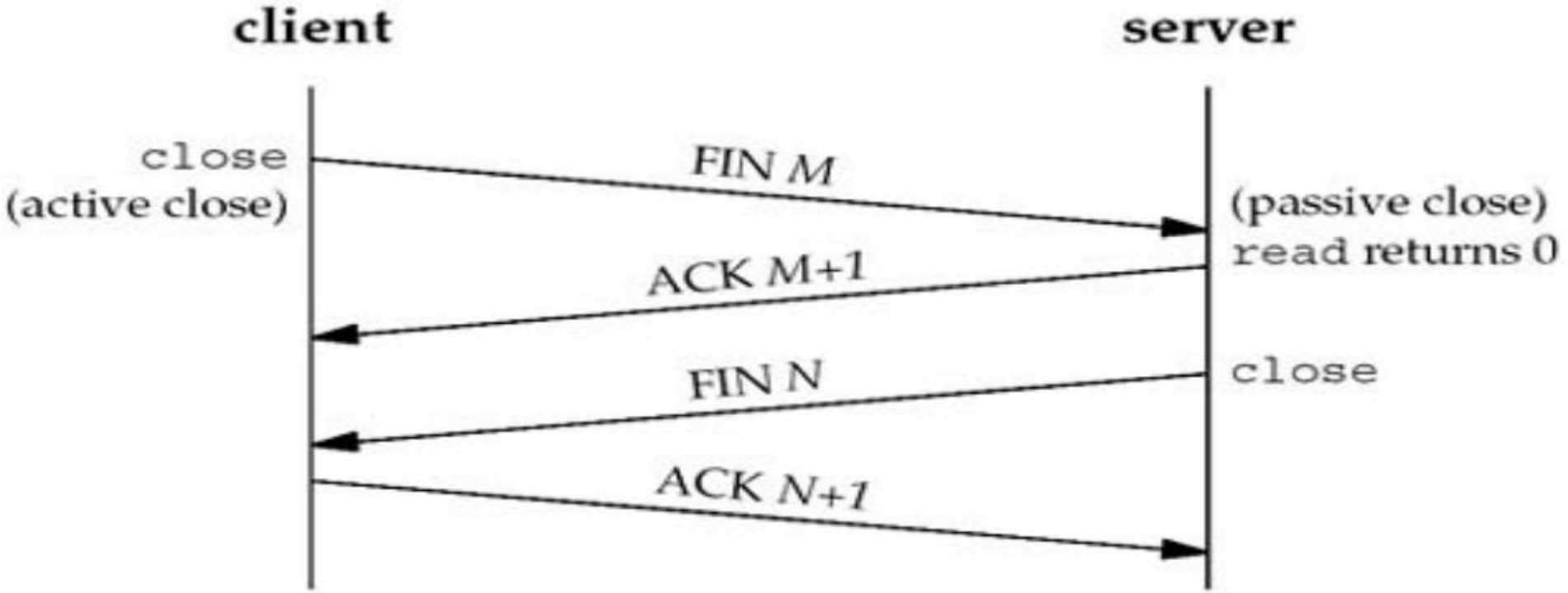
2.6 TCP Connection Establishment and Termination



Three-Way Handshake

The minimum number of packets required for this exchange is three; hence, this is called TCP's three-way handshake

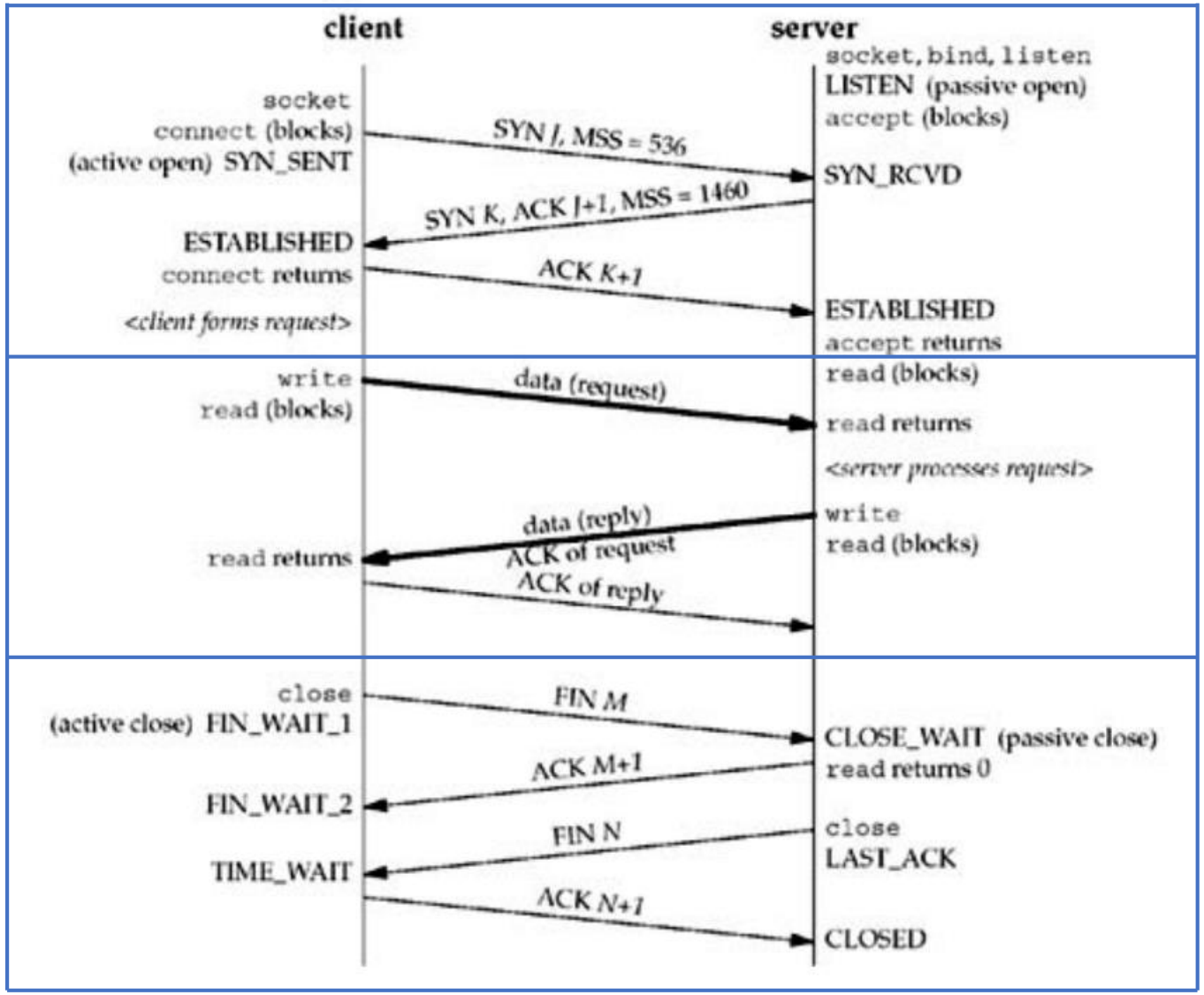
2.6 TCP Connection Establishment and Termination



TCP Connection Termination

it takes four to terminate a connection.

2.6 TCP Connection Establishment and Termination

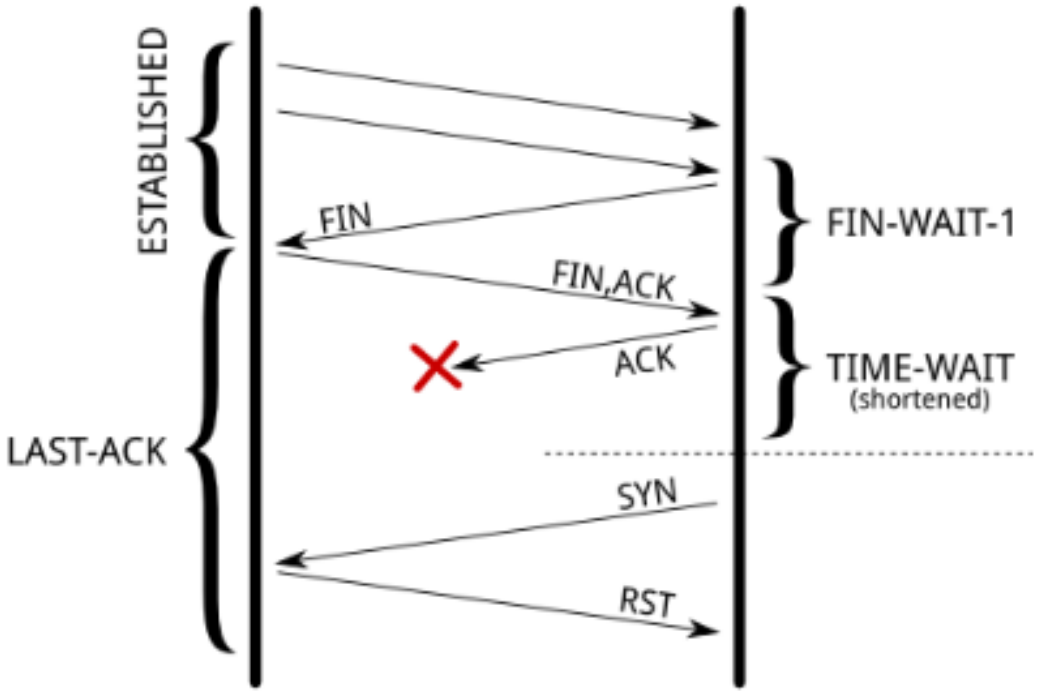


Three-Way Handshake

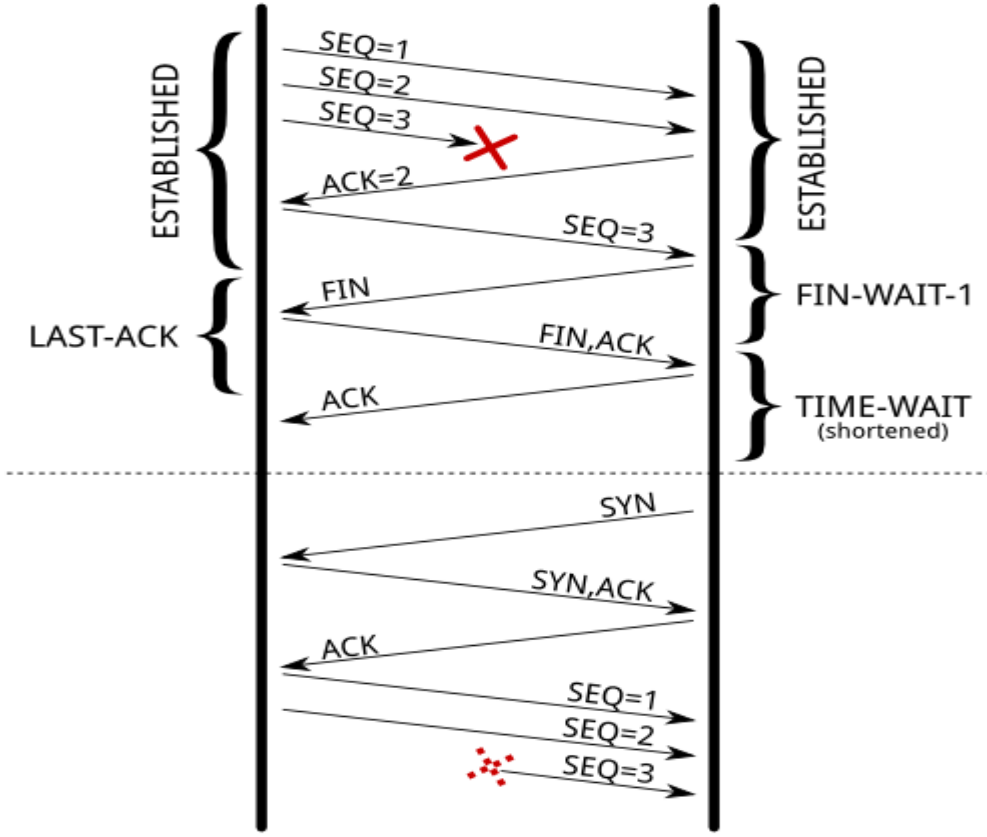
Data transfer

Connection Termination

2.7 TIME_WAIT State

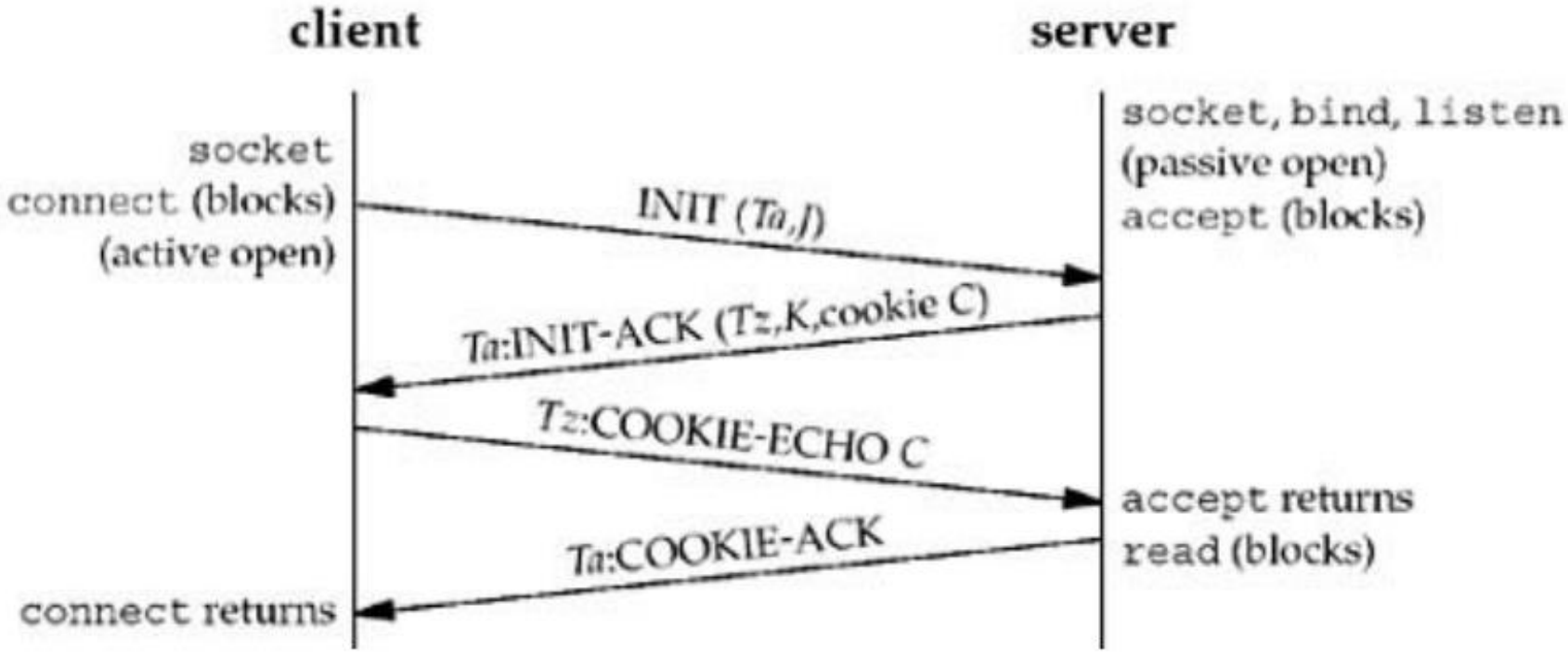


1. To implement TCP's full-duplex connection termination reliably



2. To allow old duplicate segments to expire in the network

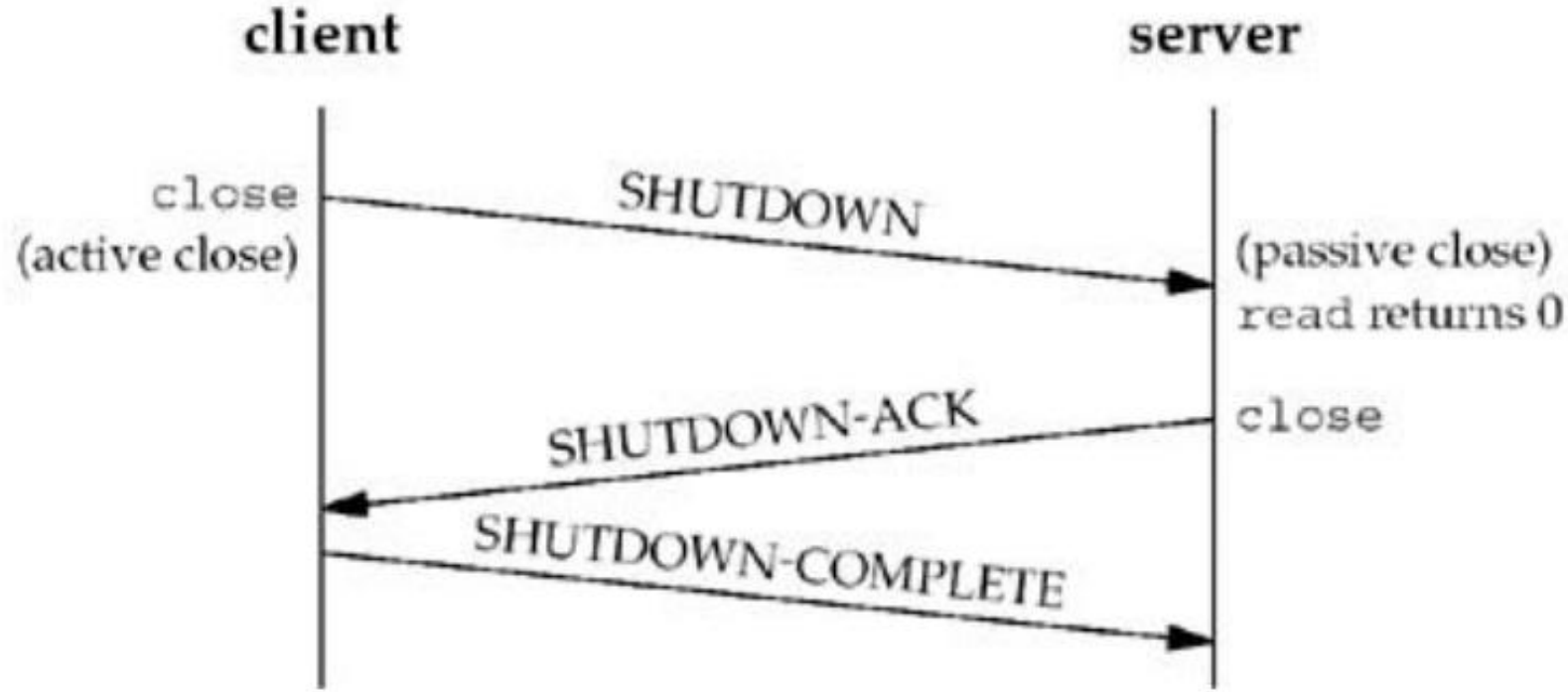
2.8 SCTP Association Establishment and Termination



SCTP Four-Way Handshake

The minimum number of packets required for this exchange is four; hence, this process is called SCTP's four-way handshake

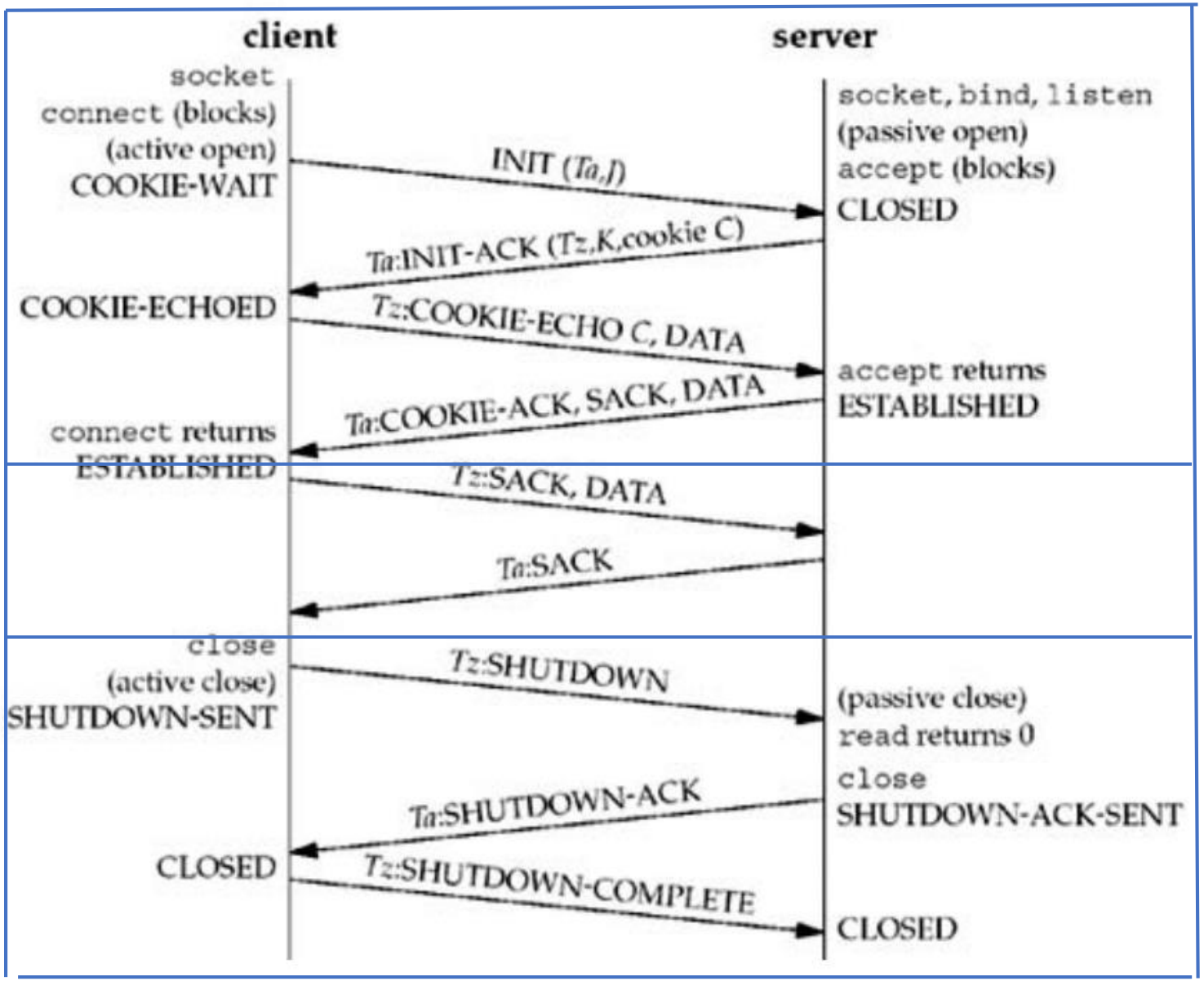
2.8 SCTP Association Establishment and Termination



Association Termination

Unlike TCP, SCTP does not permit a "half-closed" association.
SCTP instead places verification tag values in TIME_WAIT.

2.8 SCTP Association Establishment and Termination



Four-Way Handshake

Data transfer

Association Termination

2.9 Port Numbers

All three transport layers use 16-bit integer port numbers to differentiate between these processes.

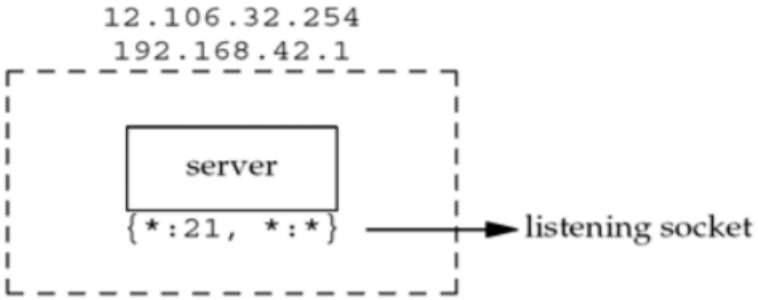
1. The well-known ports: 0 through 1023. These port numbers are controlled and assigned by the IANA.

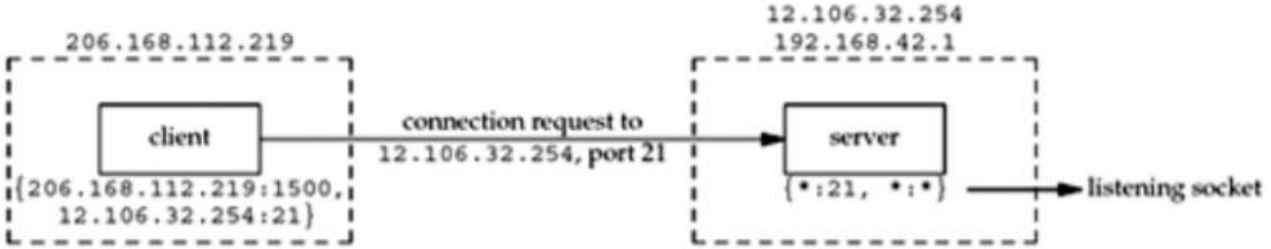
2. The registered ports: 1024 through 49151. These are not controlled by the IANA, but the IANA registers and lists the uses of these ports as a convenience to the community.

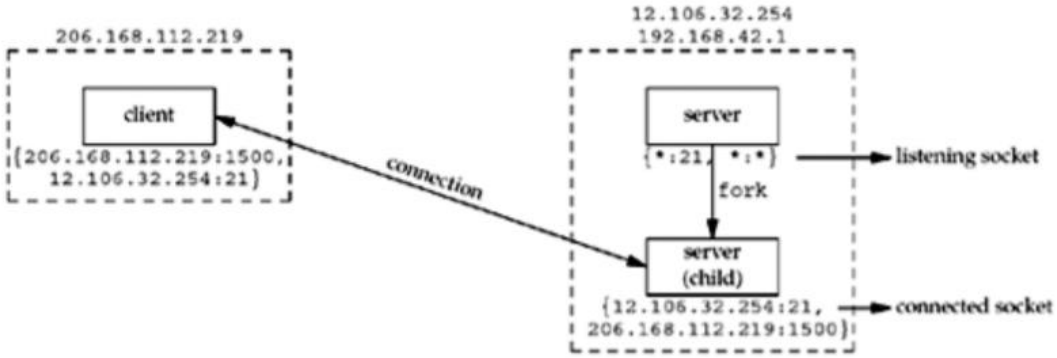
3. The dynamic or private ports: 49152 through 65535. The IANA says nothing about these ports. These are what we call ephemeral ports

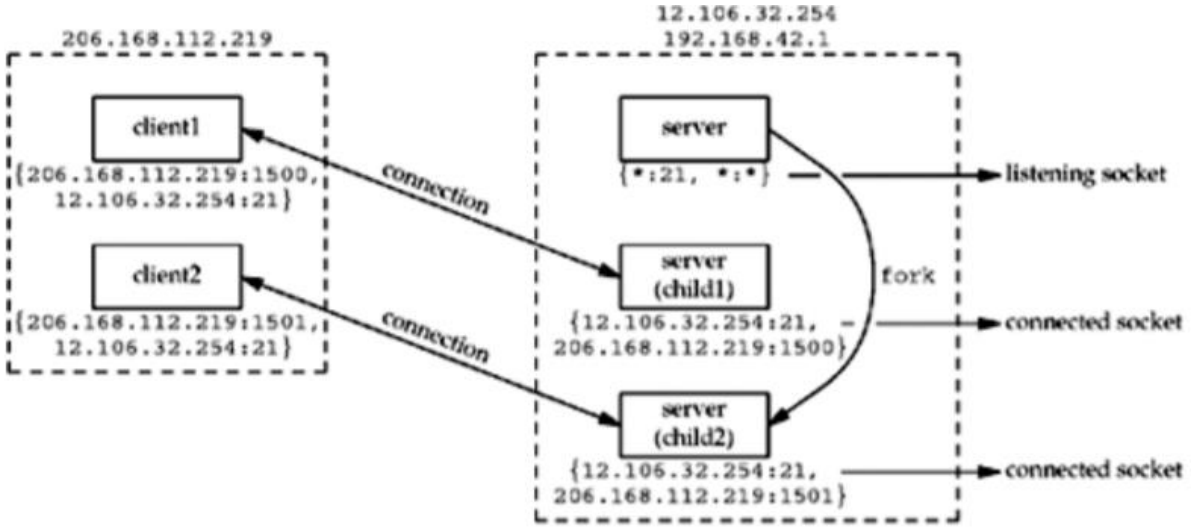
Socket Pair={local IP address, local port, foreign IP address, foreign port.}

2.10 TCP Port Numbers and Concurrent Servers









With a concurrent server,
where the main server loop spawns a child to handle each new connection

2.11 Buffer Sizes and Limitations

-MTU : maximum transmission unit ex) Ethernet is 1500bytes

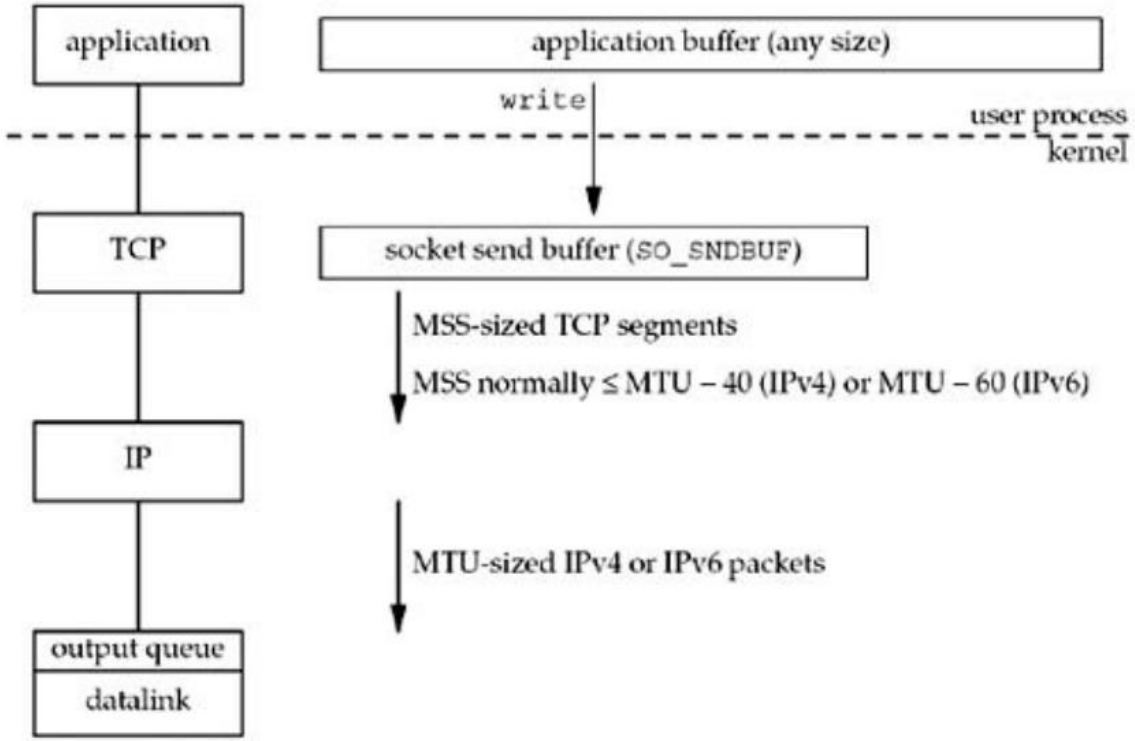
-MSS : maximum segment size ex) MSS=MTU(1500)-IPv4 header(20)-TCP header(20)=1460bytes

If the size of the datagram exceeds the link MTU, Fragmentation is performed by both IPv4 and IPv6.
The fragments are not normally reassembled until they reach the final destination.

If the "don't fragment" (DF) bit is set in the IPv4 header, it specifies that this datagram must not be fragmented,

2.11 Buffer Sizes and Limitations

-TCP Output

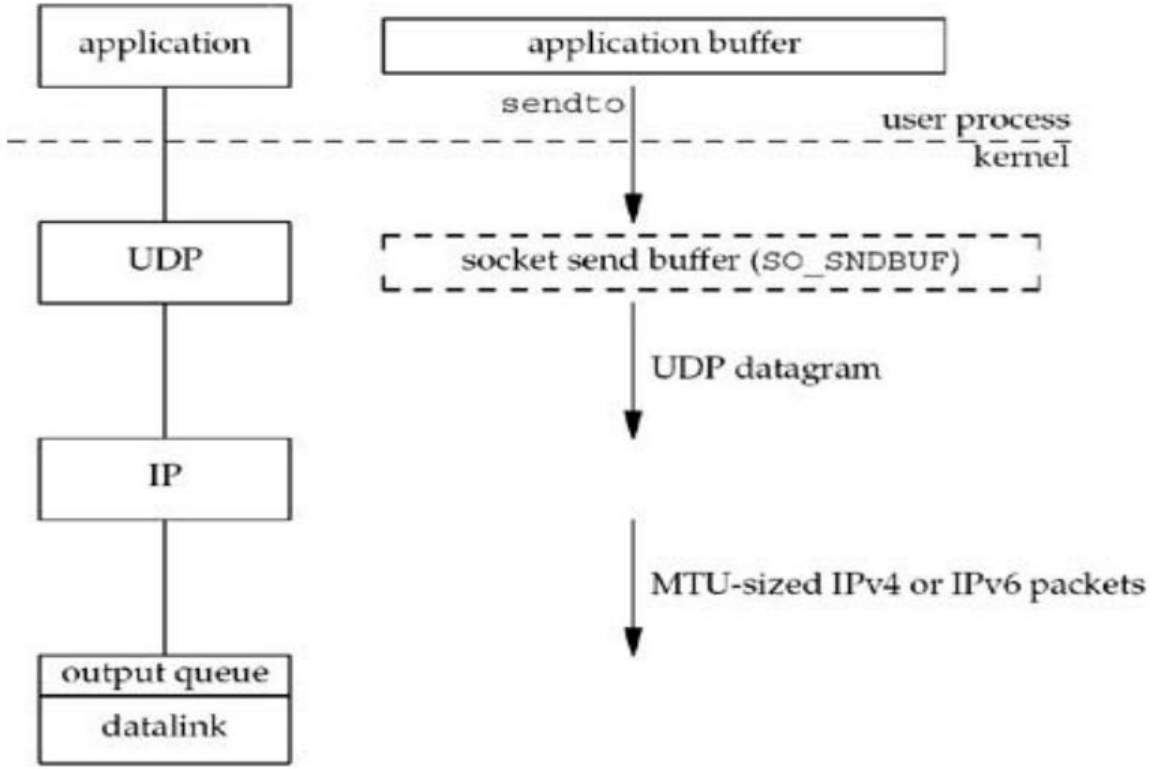


The kernel copies all the data from the application buffer into the socket send buffer.

TCP must keep a copy of our data until it is acknowledged by the peer.

2.11 Buffer Sizes and Limitations

-UDP Output



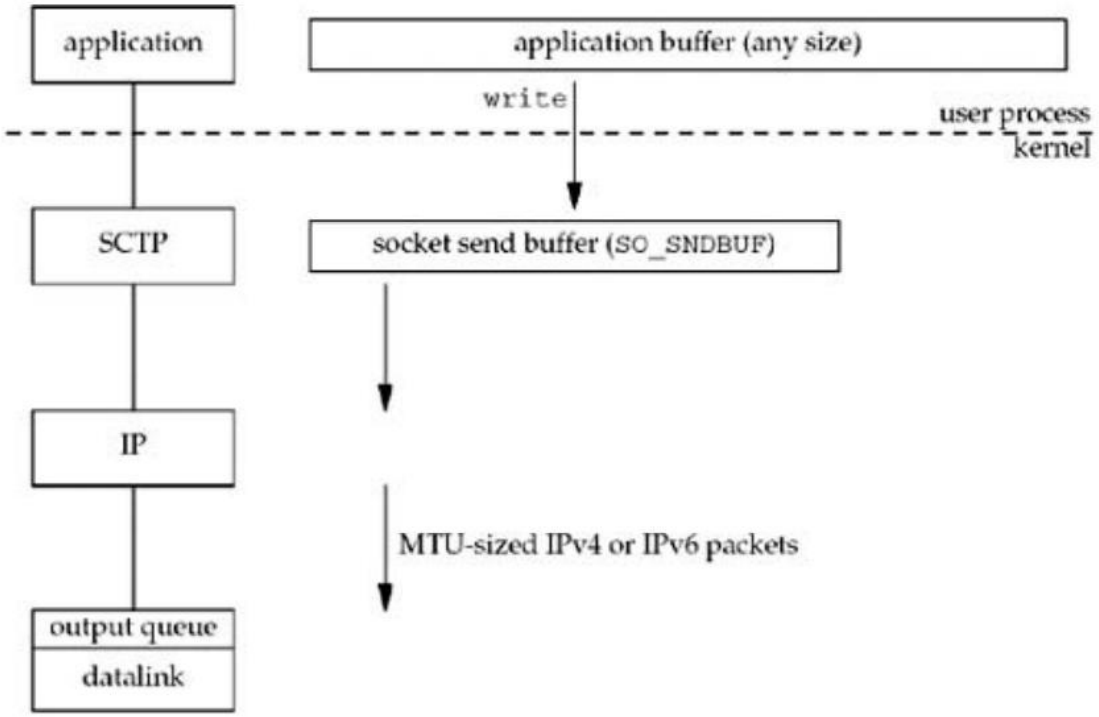
UDP does not have socket send buffer because it is unreliable.

UDP simply prepends its 8-byte header and passes the datagram to IP

There is a much higher probability of fragmentation than with TCP, because TCP breaks the application data into MSS-sized chunks, something that has no counterpart in UDP

2.11 Buffer Sizes and Limitations

-SCTP Output



The kernel copies all the data from the application buffer into the socket send buffer.