More LAN

YONSHIK CHOI, Ph.D.

Illinois Institute of Technology
Department of Computer Science
CS 455 Rice Campus

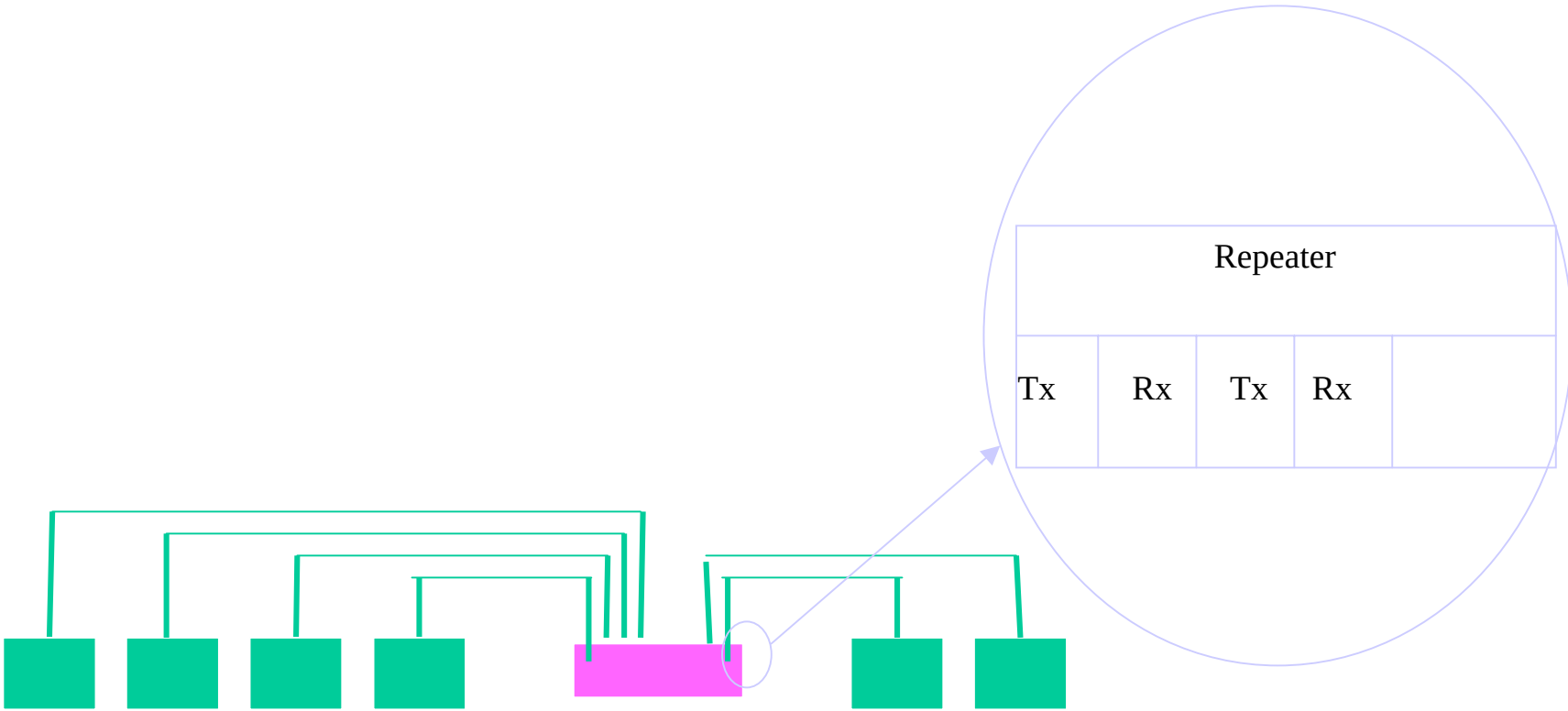
Hub

- Recently using hub with star topology is more common rather than using thick/thin coaxial cable using tab.
- Hub has multiple ports with Tx and Rx and connects to multiple DTEs.
- As physically there is a separate transmit and receive pair of wires and hub repeats/retransmits the incoming/outgoing signal.

Hub

- It emulates broadcast mode of transmission used with coaxial cable and allows collisions to be detected by each attached DTE in the normal way.
- Only one transmission can be in progress at anytime.

Hub



Ethernet Switching

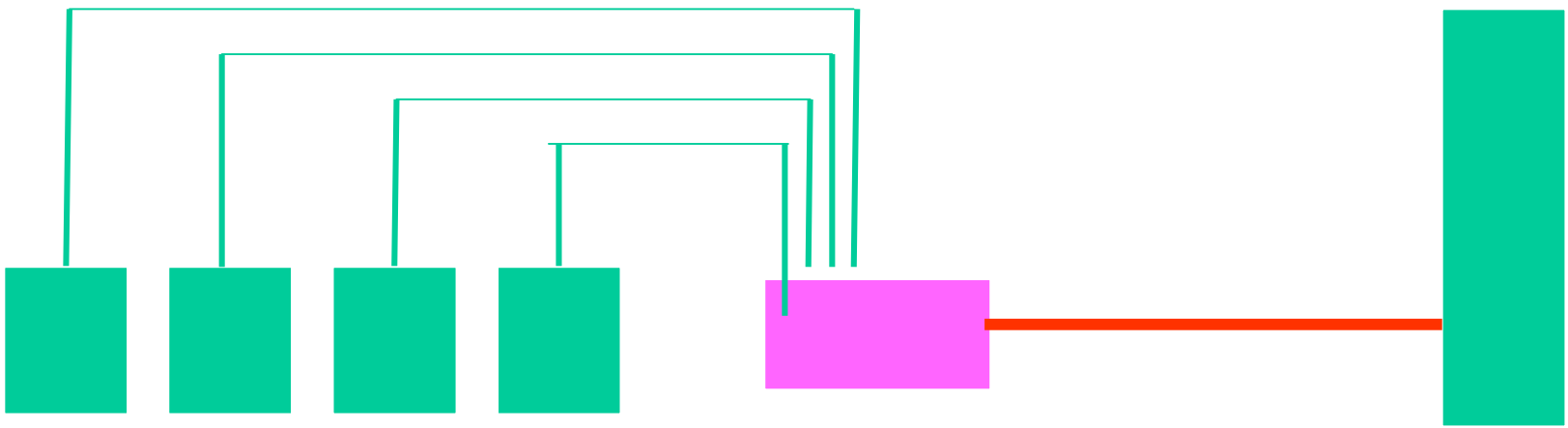
- Motivation:
  - As the applications of LAN have grown, the throughput of LAN has also grown.
  - More bandwidth required applications have developed.
  - LAN should improve its throughput and number of transactions greatly.
  - To meet these requirements, higher speed LANs have been developed: Ethernet switching, Fast Ethernet, 802.12 (Gigabit Ethernet).

Ethernet Switching

- By increasing the complexity of the repeater electronics of hub, the hub can operate in a non-broadcast mode.
- A hub can have a routing table with source/destination address on the MAC header and sending only to the destined output port, it doesn’t bother the other DTEs. It is principle of Ethernet Switching.
- Each link can work with 10 Mbps.

Ethernet Switching

- Collision occurs only when a received frame requires a destination port that is already receiving a frame from another port. To overcome this, hub to server can have higher rate.
- Fast Ethernet: If allowable distance is limited, we can achieve higher rate (100 Mbps).



Beyond LAN

Bridges

- Need to expend beyond the confines of a single LAN $\Rightarrow$ to provide interconnection to other LANs and to WANs.
- Two approaches: bridges and routers
- A Bridge is simpler than a router: provides a means of interconnecting similar LANs
- It is designed for use between LANs that use identical protocols for physical and data link layers.

Bridges

- More sophisticated Bridge are capable of mapping from one MAC format to another (e.g. Ethernet and token ring LAN)
- Then, why not use one large LAN? Other than using bridge..
  - Reliability: fault on a network may disable communication for all nodes
  - Performance: In general, performance declines with an increase in number of devices or length of the wire. A number of smaller LANs will often improve performance.

Bridges

- Security: Desirable to keep different level of user traffic (accounting, personnel, etc.)
- Geography: If two buildings are separated by a highway, then it may be easier to use a microwave bridge link than wire

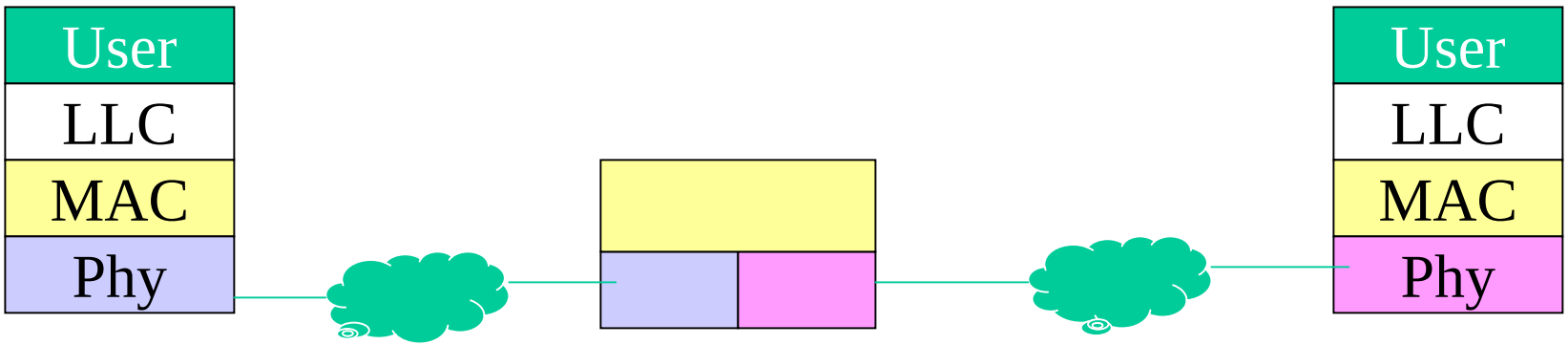
Functions of a Bridge

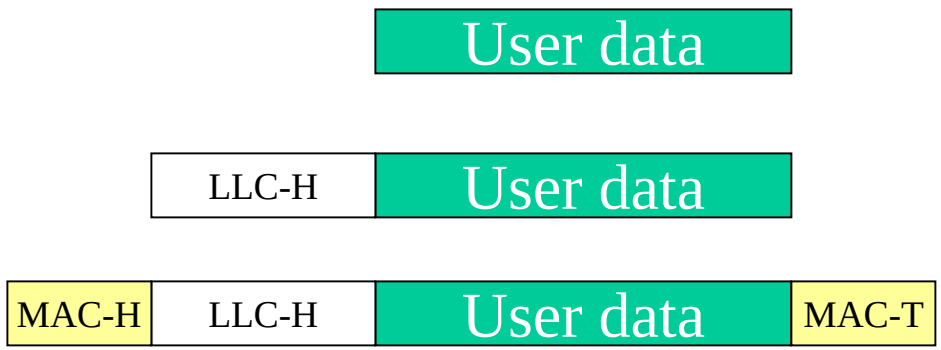
- Read all frames transmitted on LAN A and accept those addressed to any station on B.
- Using the MAC protocol for B, retransmit each frame on B.
- Do the same for B-to-A traffic.
- The bridge makes no modification to the content or format of the frames it receives, nor does it encapsulate them with an additional header. Each frame to be transferred is simply copied from one LAN and repeated with exactly the same bit pattern as the other LAN.

Functions of a Bridge

- The bridge should contain enough buffer space to meet peak demands.
- The bridge must contain addressing and routing intelligence. The bridge should know which addresses are on each network. There may be more than two LANs interconnected by a number of bridges, then a frame may have to be routed.
- A bridge may connect more than two LANs.

Bridge Protocol Architecture





Router

- Interconnects a variety of LANs and WANs
- Essential functions that a router must perform:
  - Provide a link between networks.
  - Provide for the routing and delivery of data between processes on end systems attached to different networks.
  - Provide these functions in such a way as not to require modifications of the networking architecture of any of the attached subnetworks.

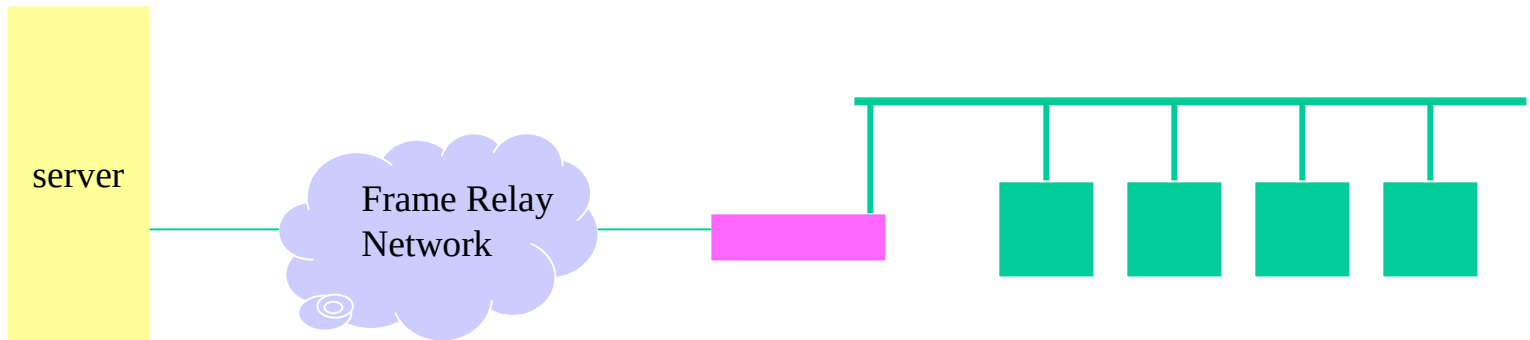
Router

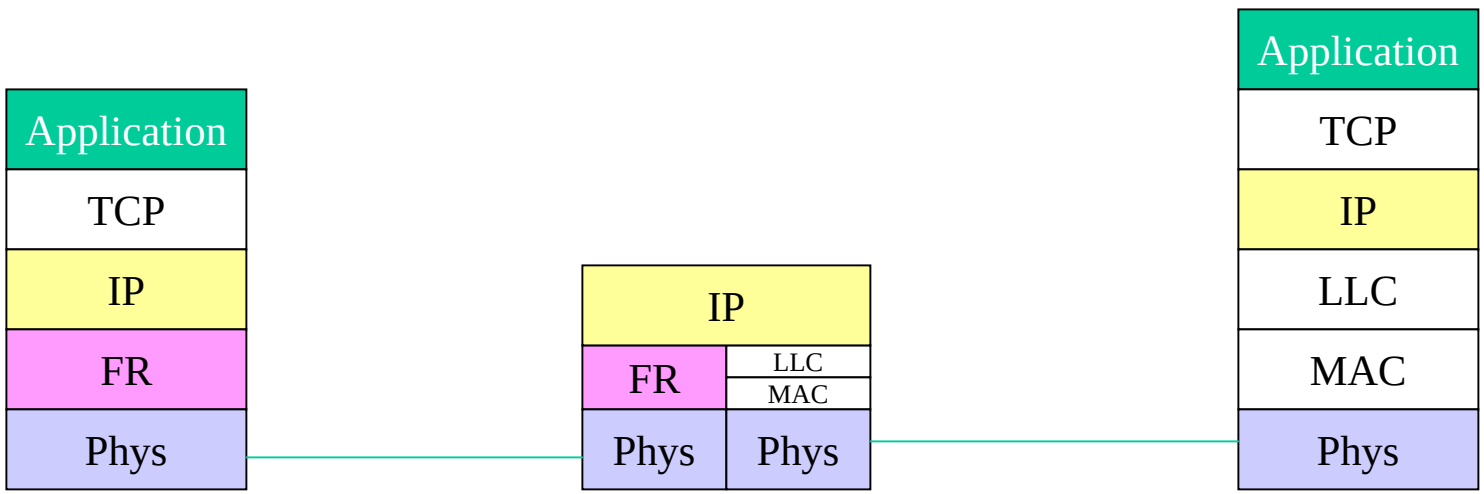
- A router must accommodate a number of differences among networks:
  - Addressing schemes: Networks may use different addresses and router should work.
  - Maximum packet size: segmentation of a packet for a network accepts smaller packet size.

Router

- Interfaces: the hardware & software interfaces to various networks differ. Independence of these differences.
- Reliability: the operation of the router should not depend on an assumption of network reliability. Some network service provide a reliable end-to-end virtual circuit.

Router Protocol Architecture





WAN

- Circuit Switched Network
  - Telephone network using SS7.
- Packet Switched Network
  - Datagram
  - Virtual Circuit

IP

- Source address: Internetwork address of sending IP entity.
- Destination address: Internetwork address of destination IP entity.
- Protocol: Recipient protocol entity (an IP user, such as TCP).
- Type of service indicators: Used to specify the treatment of the data unit in its transmission through component networks.

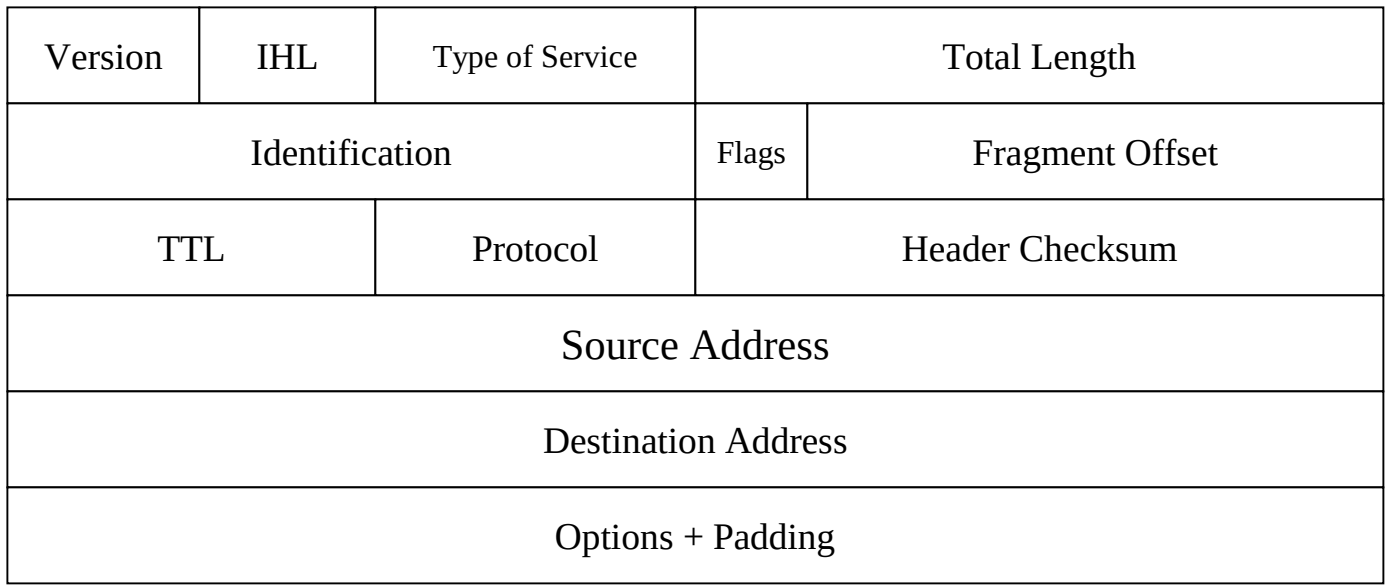
IP

- Identification: Used in combination with the source and destination addresses and user protocol to identify the data unit uniquely. This parameter is needed for reassembly and error reporting.
- Don’t fragment identifier: Indicates whether IP can fragment data to accomplish delivery.
- Time to live: Measured in seconds.

IP

- Data length: Length of data being transmitted.
- Option data: Options requested by the IP user.
- Data: User data to be transmitted.

IPv4 header



IHL: Internet Header Length
ToS: Specifies reliability, precedence, delay, and throughput parameters
TL: total datagram length, in octets
I: sequence number
F: only two bits are currently defined.
FO: indicates where in the original datagram this fragment belongs
TTL: specifies how long, in seconds, a datagram is allowed to remain in the Internet
P: indicates the next higher level protocol that is to receive the data field at the destination

IPv6

- Expanded address space: uses 128 bit address instead of 32 bit of IPv4.
- Increase of address space by a factor of 2^{96}
- Improved option mechanism: IPv6 options are placed in separate optional headers that are located between the IPv6 header and the transport-layer header. Most of these optional headers are not examined or processed by any router on the packet's path. This simplifies and speeds router processing of IPv6 packets compared to IPv4 datagrams.

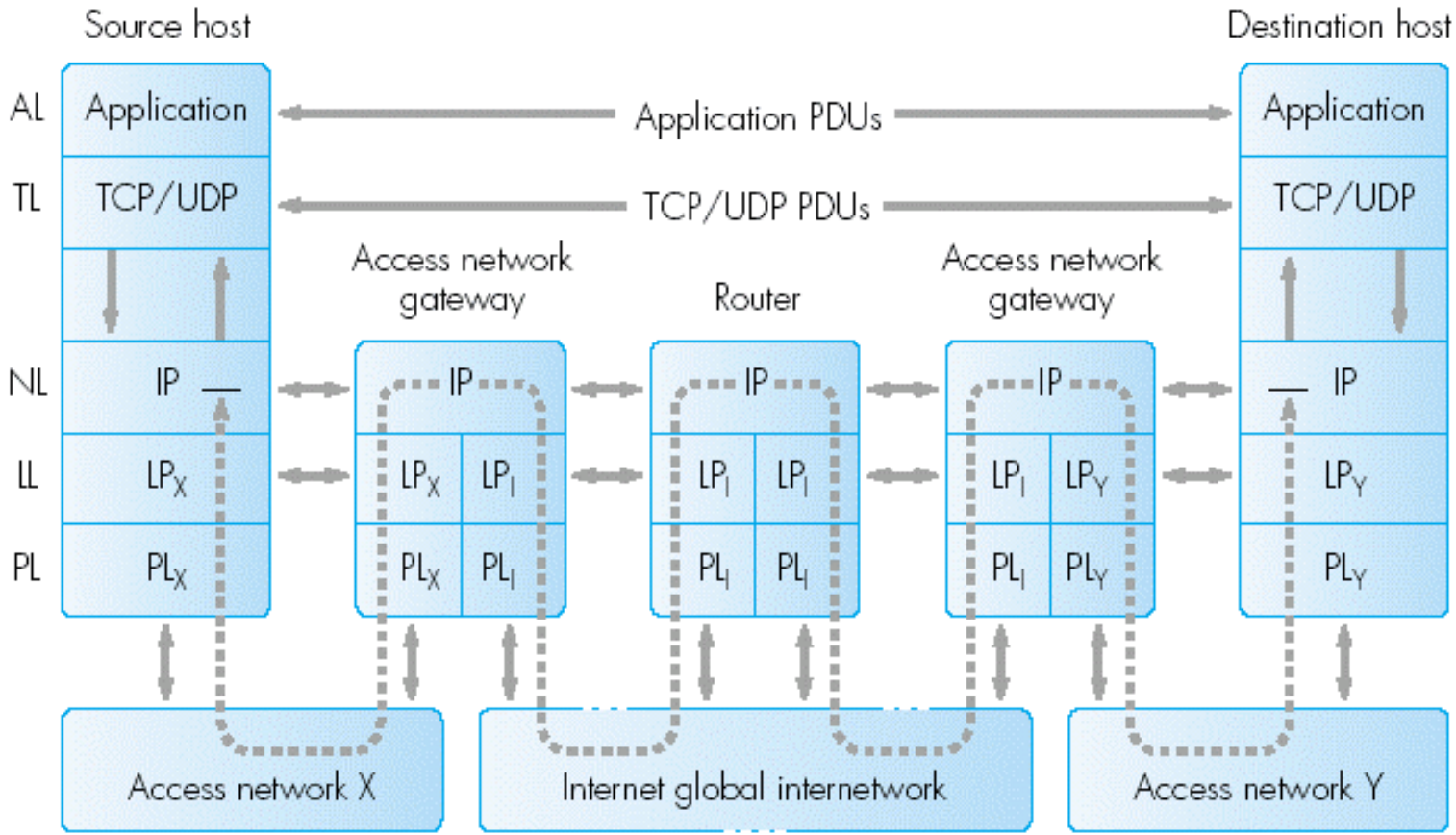
IPv6

- Address autoconfiguration:This capability provides for dynamic assignment of IPv6 addresses.
- Increased addressing flexibility: IPv6 includes the concept of an anycast address, for which a packet is delivered to just one of a set of nodes. The scalability of multicast routing is improved by adding a scope field to multicast addresses.

IPv6

- Support for resource allocation: Instead of the type-of-service field in IPv4, IPv6 enables the labeling of packets belong to a particular traffic flow for which the sender requests special handling. This aids in the support of specialized traffic such as real-time video.

Internet Protocol Diagram



= logical communications path of protocol data units (PDUs)

= actual path

TCP/UDP = transmission control protocol/user datagram protocol
IP = Internet protocol LP = link protocol PL = physical layer

Internet Protocol Diagram

- Address Resolution Protocol (ARP)
  - Used by IP in hosts attached to a broadcast LAN
  - Determines the MAC address of a host or gateway given its IP address

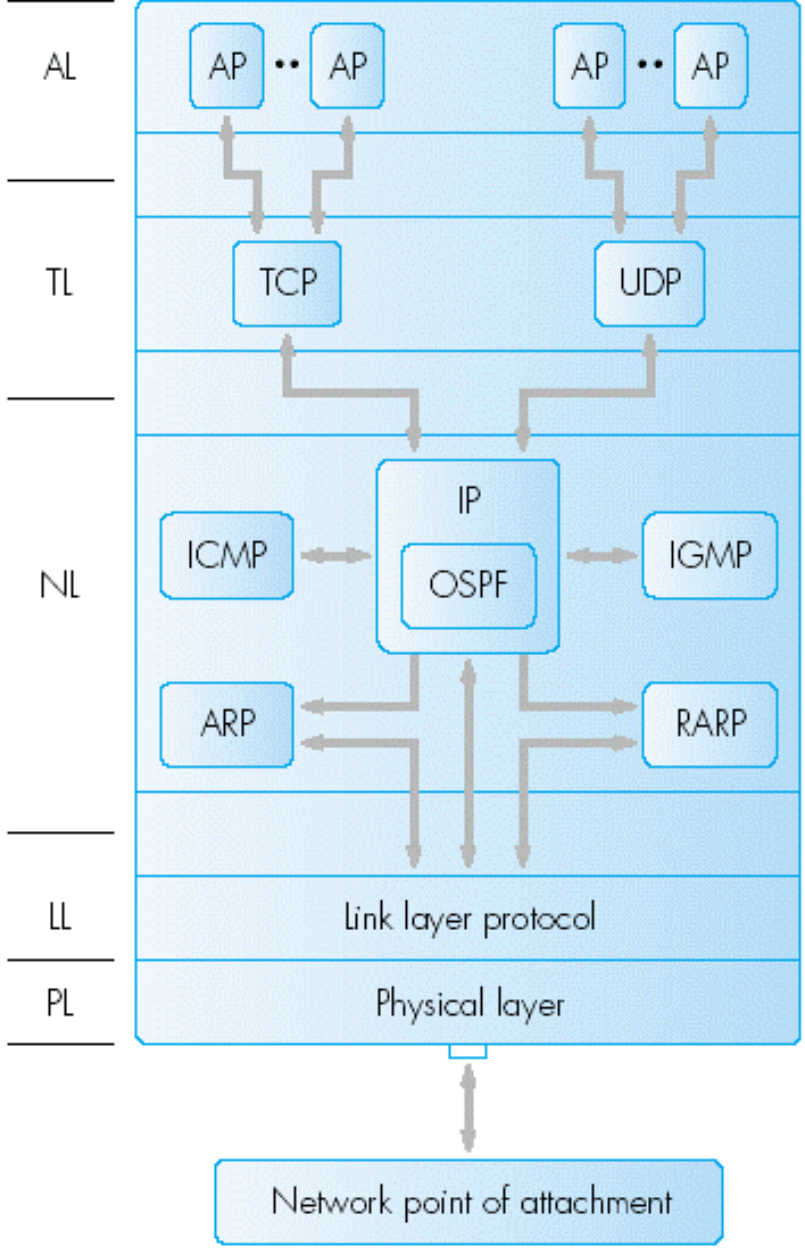
- Reverse ARP (RARP)
  - Reverse function of ARP

- Open Shortest Path First (OSPF)
  - One of the routing protocols
  - Used in the routers to build the routing table

Internet Protocol Diagram

- Internet Control Message Protocol (ICMP)
  - Used by the IP in a host or gateway to exchange control messages (e.g., error information) with the IP in another host or gateway

- Internet Group Management Protocol (IGMP)
  - Used by a host to multicast a datagram to other hosts in the same multicast group



AP = application protocol/process
IP = Internet protocol
ARP = address resolution protocol
RARP = reverse ARP
ICMP = Internet control message protocol
IGMP = Internet group message protocol
OSPF = open shortest path first

Fragmentation and Reassembly (FAR)

- Operation
  - If the size of the packet is larger than the Maximum Transmission Unit (MTU) in an intermediate network (or the destination access network), the IP in the intermediate router (or the destination gateway) divides the packet into smaller fragments.
  - The IP in the destination host reassembles the fragments.

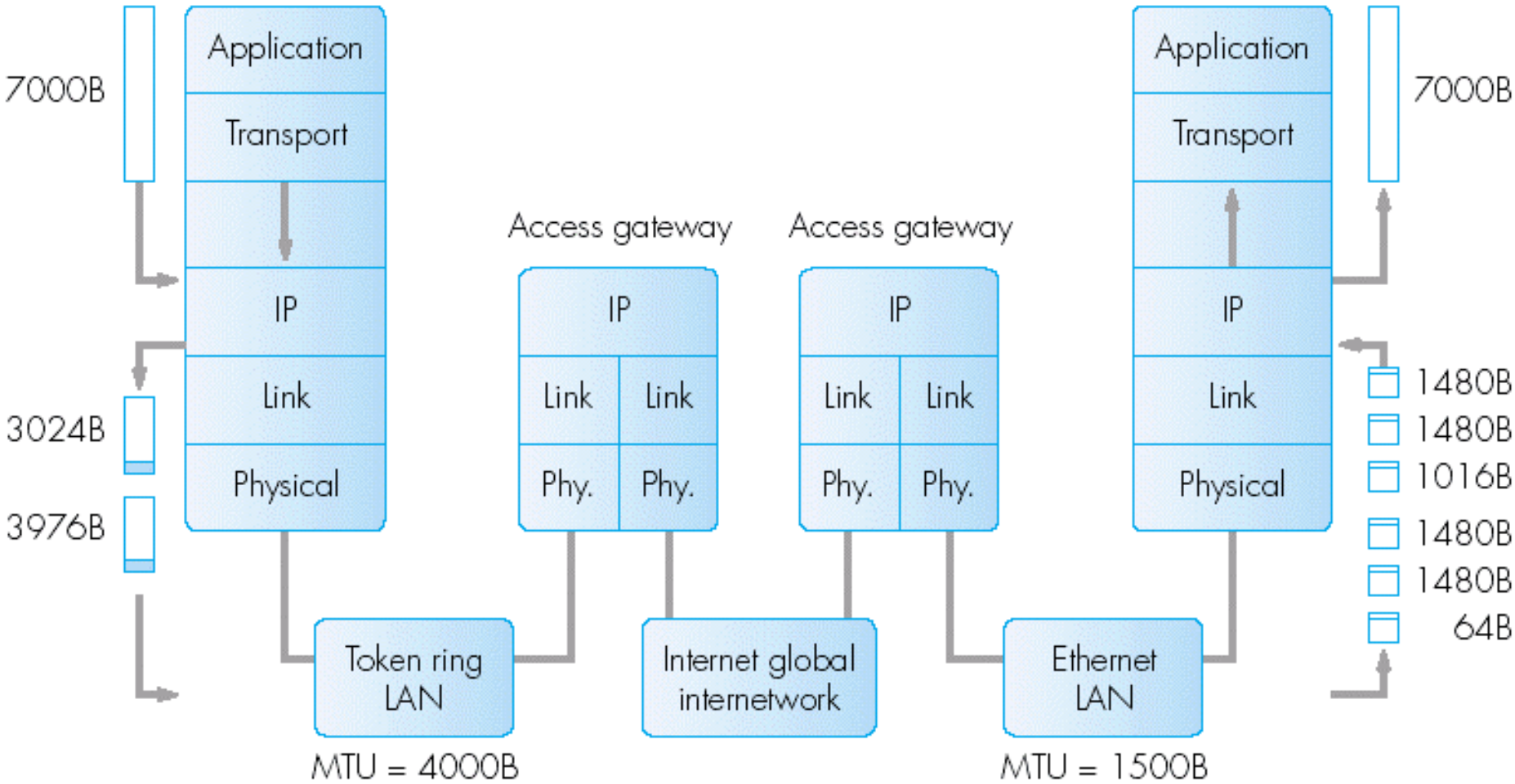
- Datagram fields used
  - Identification
  - Total length
  - Fragmentation offset
  - More fragments

Fragmentation and Reassembly (FAR)

- Drawbacks
  - Source TCP will retransmit a block if ACK not received within maximum time limit
  - If one of the fragments is delayed or discarded, TCP will retransmit entire block

- Alternatives
  - TCP limits maximum block size
  - Source IP determines MTU for the path prior to sending packet

Fragmentation and Reassembly



Note: All values shown are the amounts of user data in each packet/frame in bytes

Fragmentation and Reassembly

(b) *Token ring LAN*:

	(*i*)	(*ii*)
Identification	20	20
Total length	7000	7000
Fragment offset	0	497
(User data)	3976	3024
M-bit	1	0

(c) *Ethernet LAN*:

	(*i*)	(*ii*)	(*iii*)	(*iv*)	(*v*)	(*vi*)
Identification	20	20	20	20	20	20
Total length	7000	7000	7000	7000	7000	7000
Fragment offset	0	185	370	497	682	867
(User data)	1480	1480	1016	1480	1480	64
M-bit	1	1	1	1	1	0

ARP & RARP

- ARP used by the IP in hosts attached to a broadcast LAN to determine the MAC address of another host gateway port given the IP address (RFC 826)
- RARP performs the reverse operation (RFC 903)

- ARP operation: e.g.
  - Each host has two addresses (IP and MAC) stored in configuration file of the host on the hard disk
  - Each ARP has an ARP cache: routing table with IP/MAC address pairs
  - Host A sends datagram to host B
    * If translation not in cache, ARP broadcast request msg
    * Host B sends ARP reply

ARP & RARP

- Host A sends datagram to a host on a different LAN
  * ARP broadcast request msg
  * Gateway ARP returns its own pair: proxy ARP

• RARP operation: e.g.
  - Diskless hosts have only the MAC address
  - IP address obtained initially from the server with RARP request and reply messges

Traffic Management of TCP

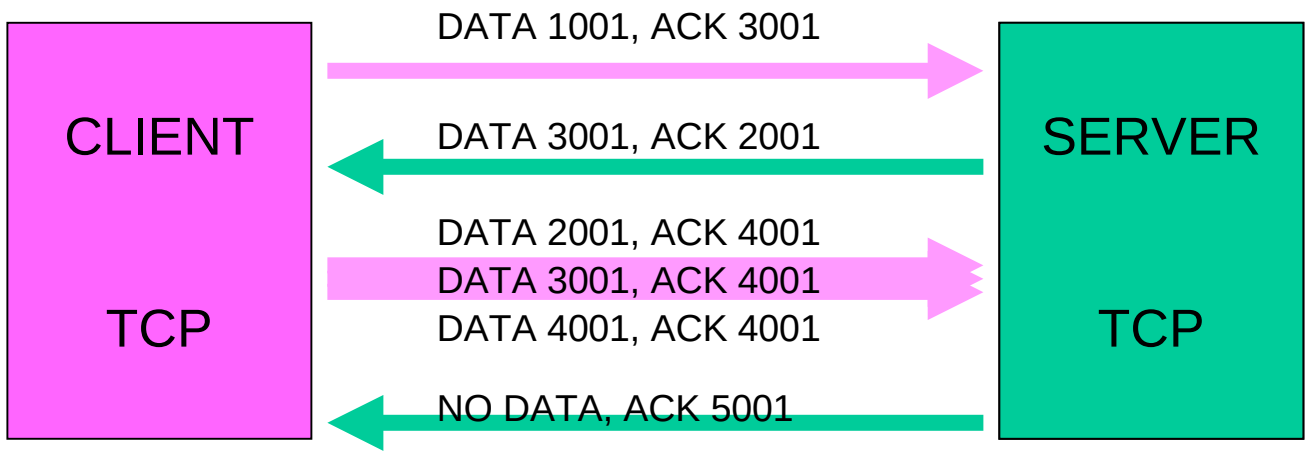
Data Transfer

- Data transfer begins after completion of the three-way handshake.
- Each segment's TCP header includes an ACK field which identifies the sequence number of the next byte expected from the partner.

Data Transfer

- For example (next slide);
  - The first segment sent by the client contains bytes 1001 to 2000. Its ACK field announces that 3001 is the sequence number of the next byte expected from the server.
  - The ACK field from the server indicates that bytes 1001 to 2000 have been received in perfect condition, so the sequence number of the next byte expected from the client is 2001.

Data Transfer



Three Way Handshake

To establish a TCP connection:

1. The requesting end (*client*) sends a SYN segment specifying the port number of the *server* that the client wants to connect to, and the client's *initial sequence number* (ISN*). This is segment 1.

Three Way Handshake

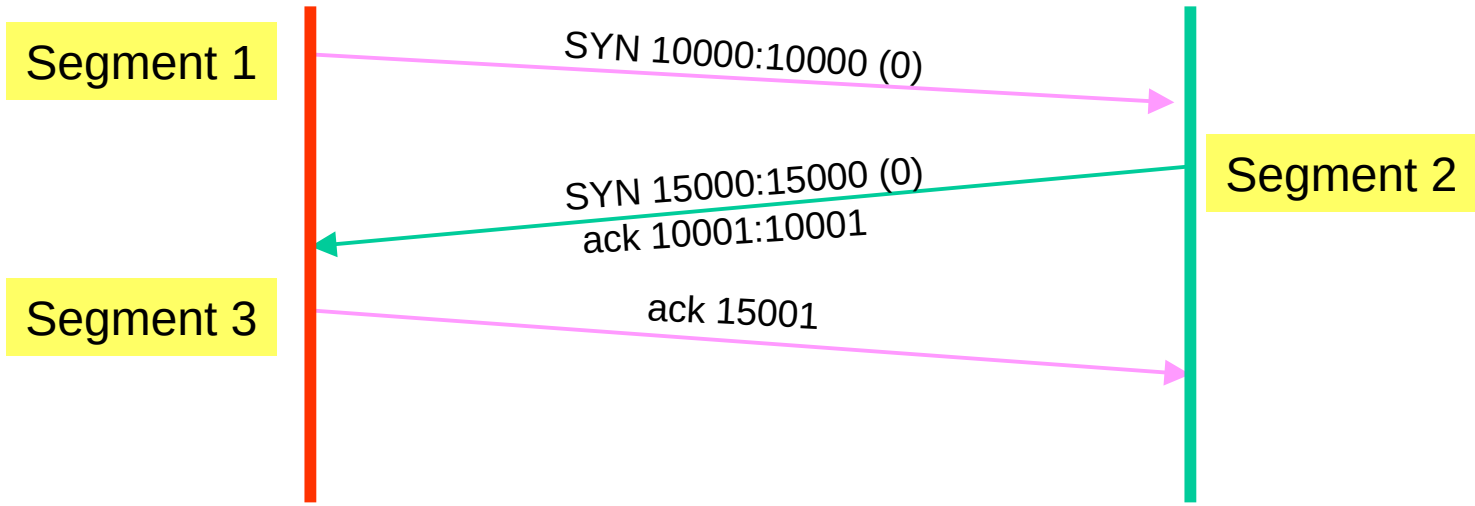
2. The server responds with its own SYN segment containing the server's initial sequence number (segment 2). The server also acknowledges the client's SYN by ACKing the client's ISN plus one. A SYN consumes one sequence number.
3. The client must acknowledge this SYN from the server by ACKing the server's ISN plus one (segment 3).

Three Way Handshake

These three segments complete the connection establishment.

* ISN is chosen by the system.

Three Way Handshake



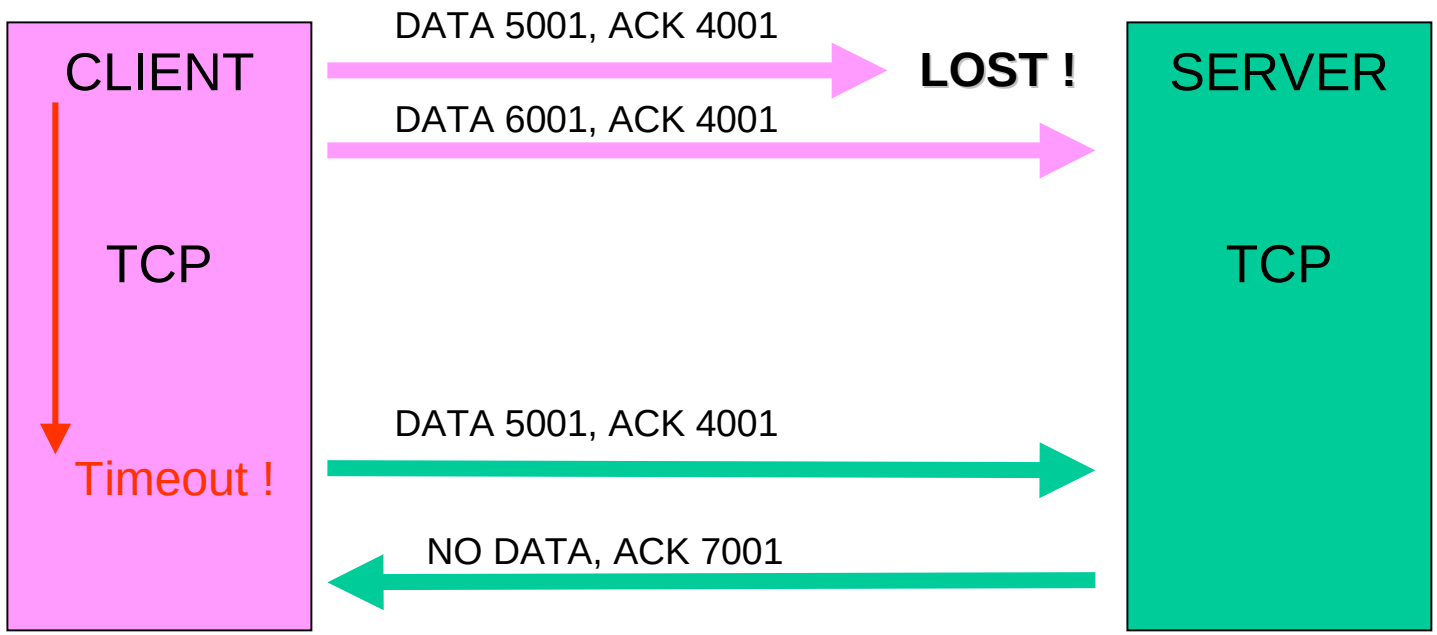
Data Loss and Retransmission

- If the first segment is lost like above diagram, the lost segment should be retransmitted after a timeout period.
- TCP uses implicit ACK so there is no negative ACK for the lost segment.
- The server TCP never sends back for the lost segment.

Data Loss and Retransmission

- Client TCP timeouts. Retransmit the lost segment.
- When the server TCP receives the lost segment, it acknowledges both at a time (see ACK 7001 sent by server TCP).

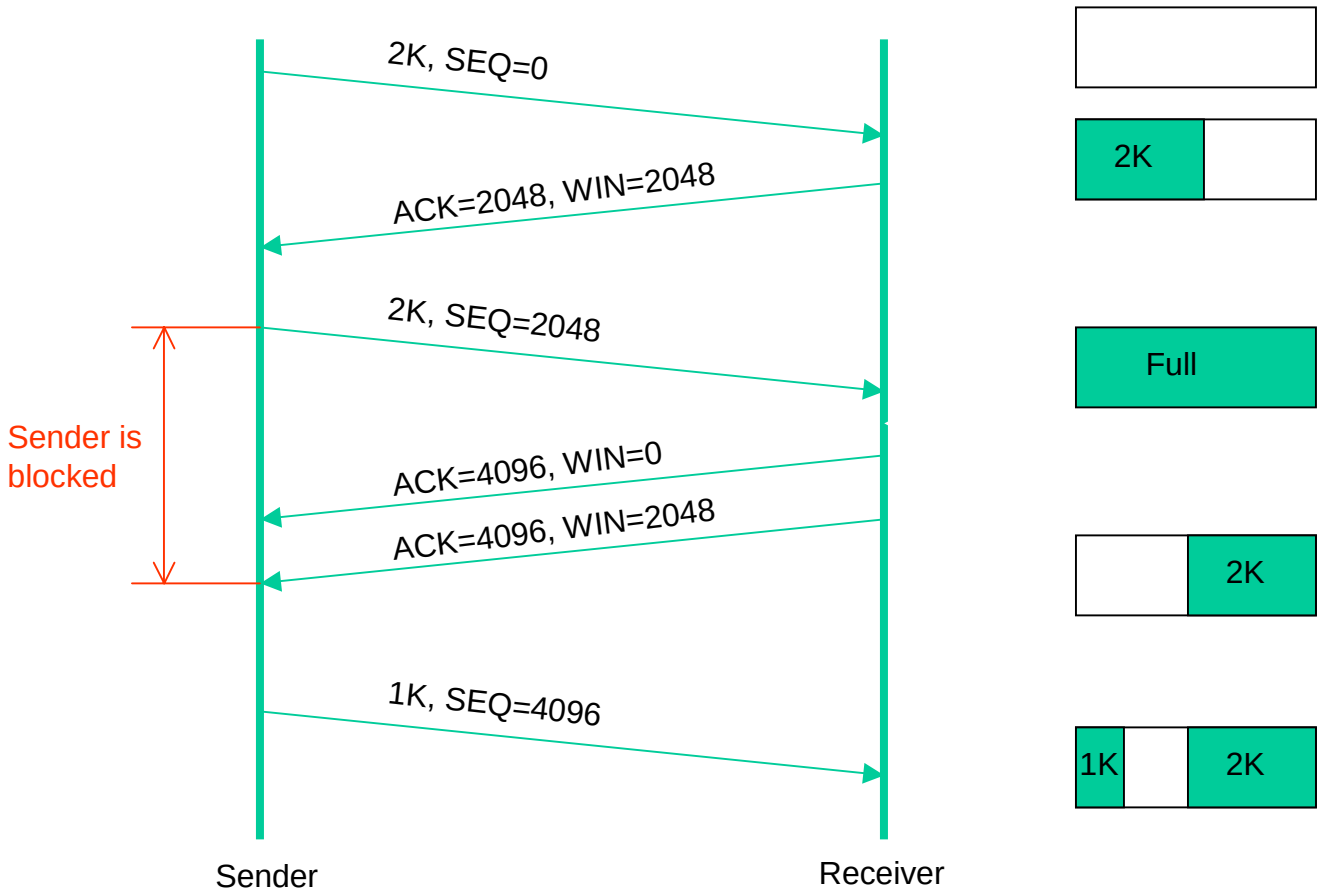
Data Loss and Retransmission



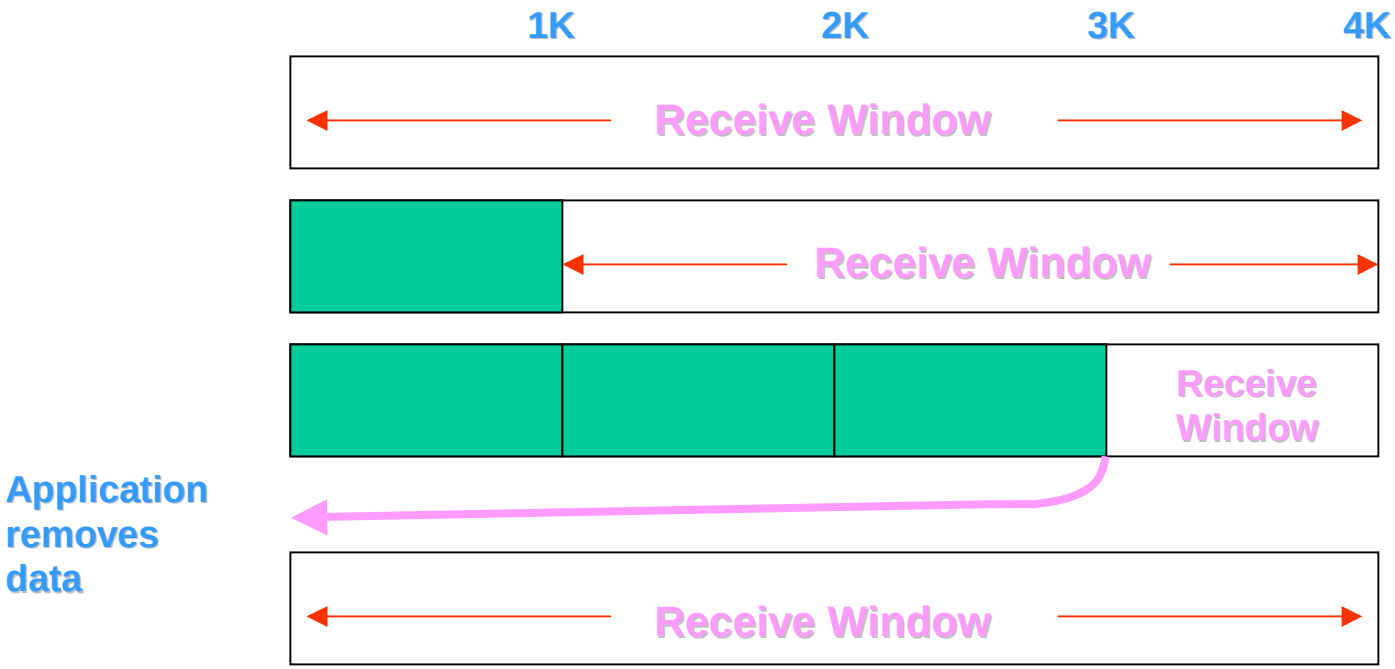
Flow Control

- The TCP data receiver is in charge of its incoming flow of data.
- The receiving TCP decides how much data it is willing to accept, and the sending TCP must stay within this limit.

Flow Control

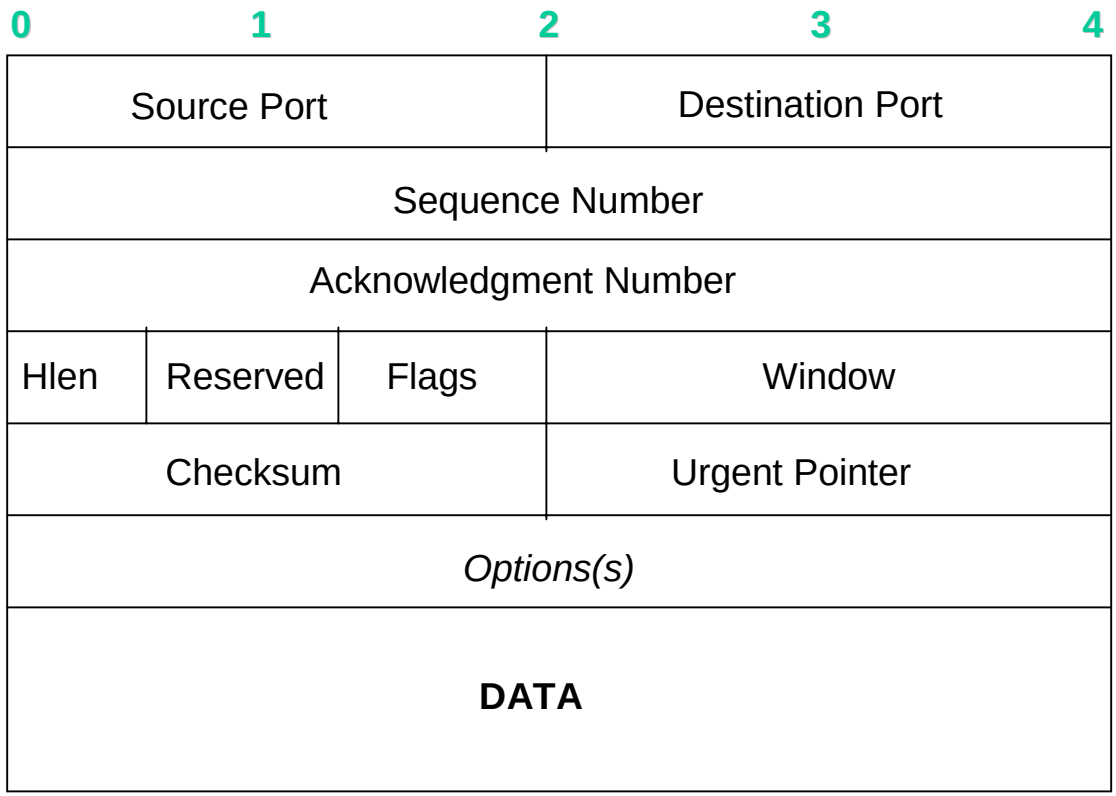


Sliding Window



- The diagram above is an example for the receive window.
- The send window shrinks and extends by acknowledges.

TCP Header



URG: 1 if urgent data is included

ACK: 1 for all but the initial SYN segment

PSH: Indicates that data should be delivered promptly

RST: Indicates an error; also used to abort a session

SYN: Set to 1 during connection setup

FIN: Set to 1 during graceful close

TCP Congestion Control

- When the load offered to any NW is more than it can handle, congestion builds up.

Case 1: A receiver buffer is small.

If a sender does not send more traffic than receiver's availability, no buffer overflow.

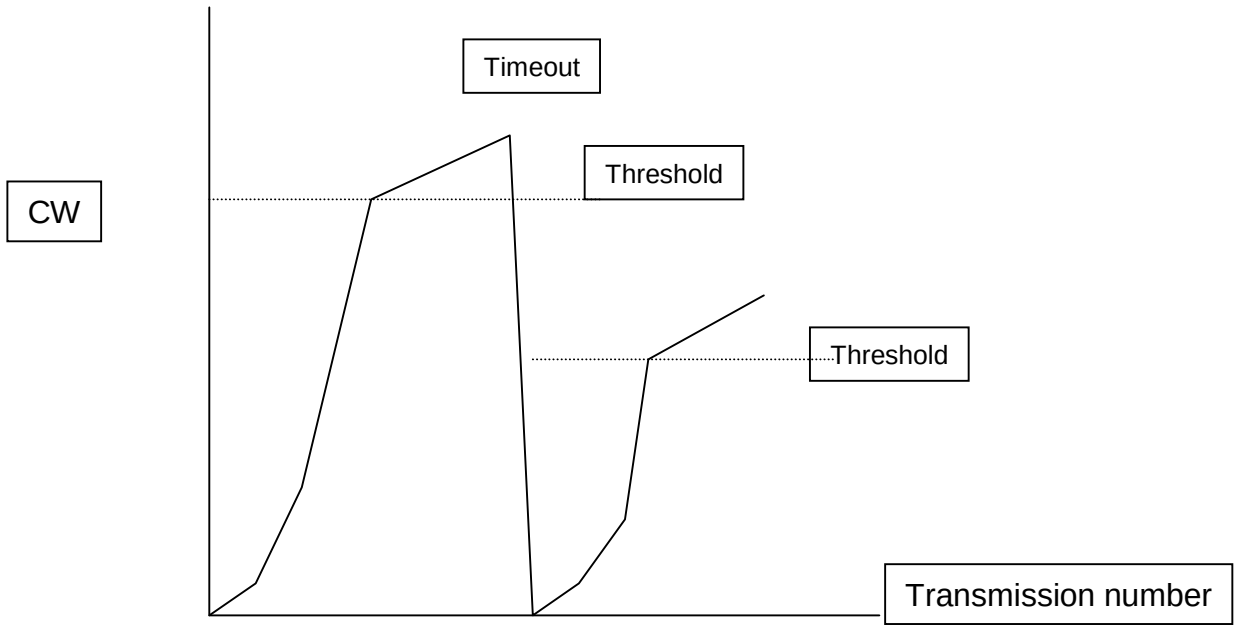
Case 2: Network has internal congestion.

Requires advanced control $\Rightarrow$ Congestion window

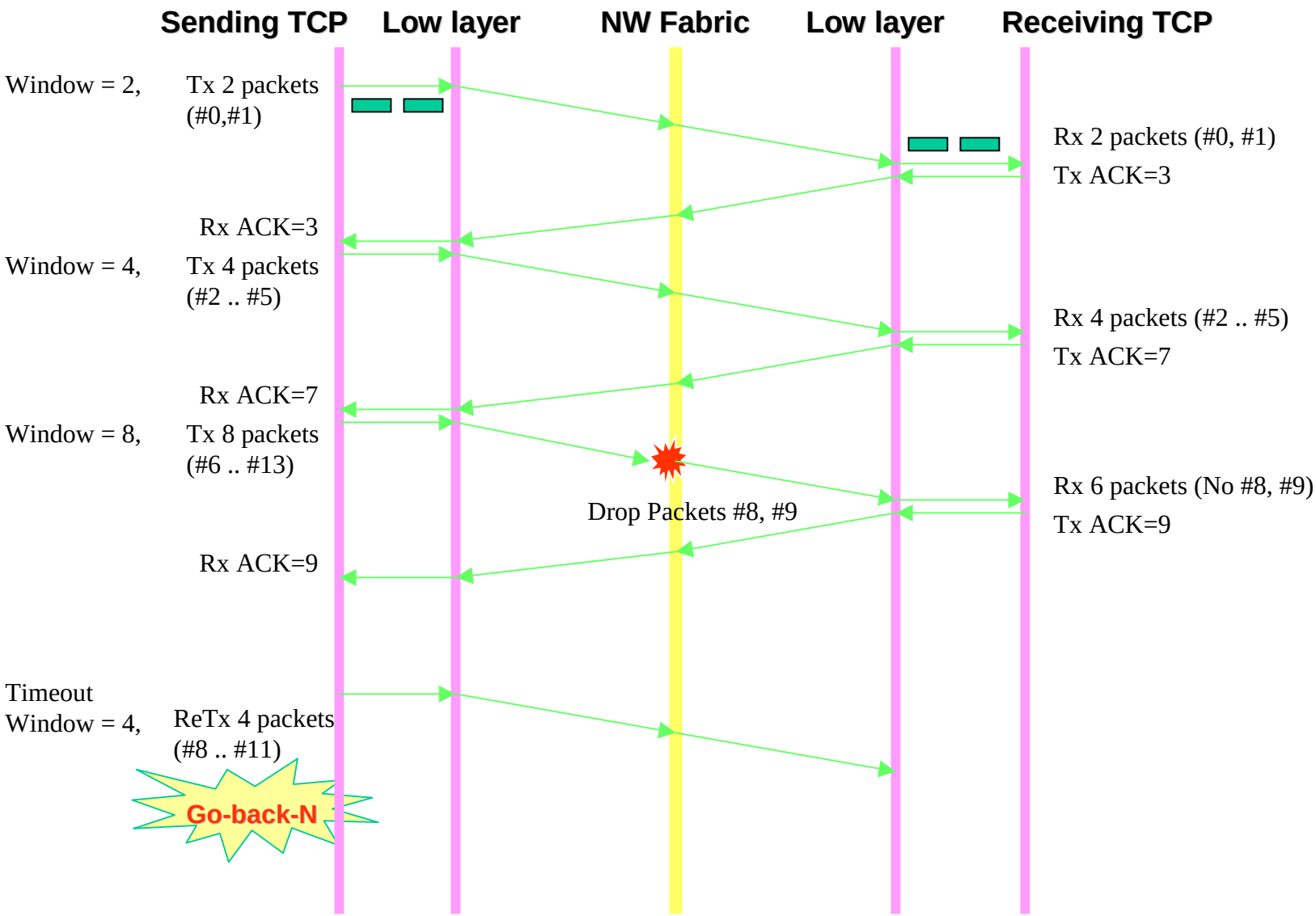
TCP Congestion Control

- The Internet solution : sender maintains two window.
- Slow start : sets a threshold (usually 64KB).
  When timeouts, congestion window reduces half and grows exponentially up to threshold limit.
  Congestion window grows linearly after it hits threshold (normally by one segment)

TCP Congestion Control

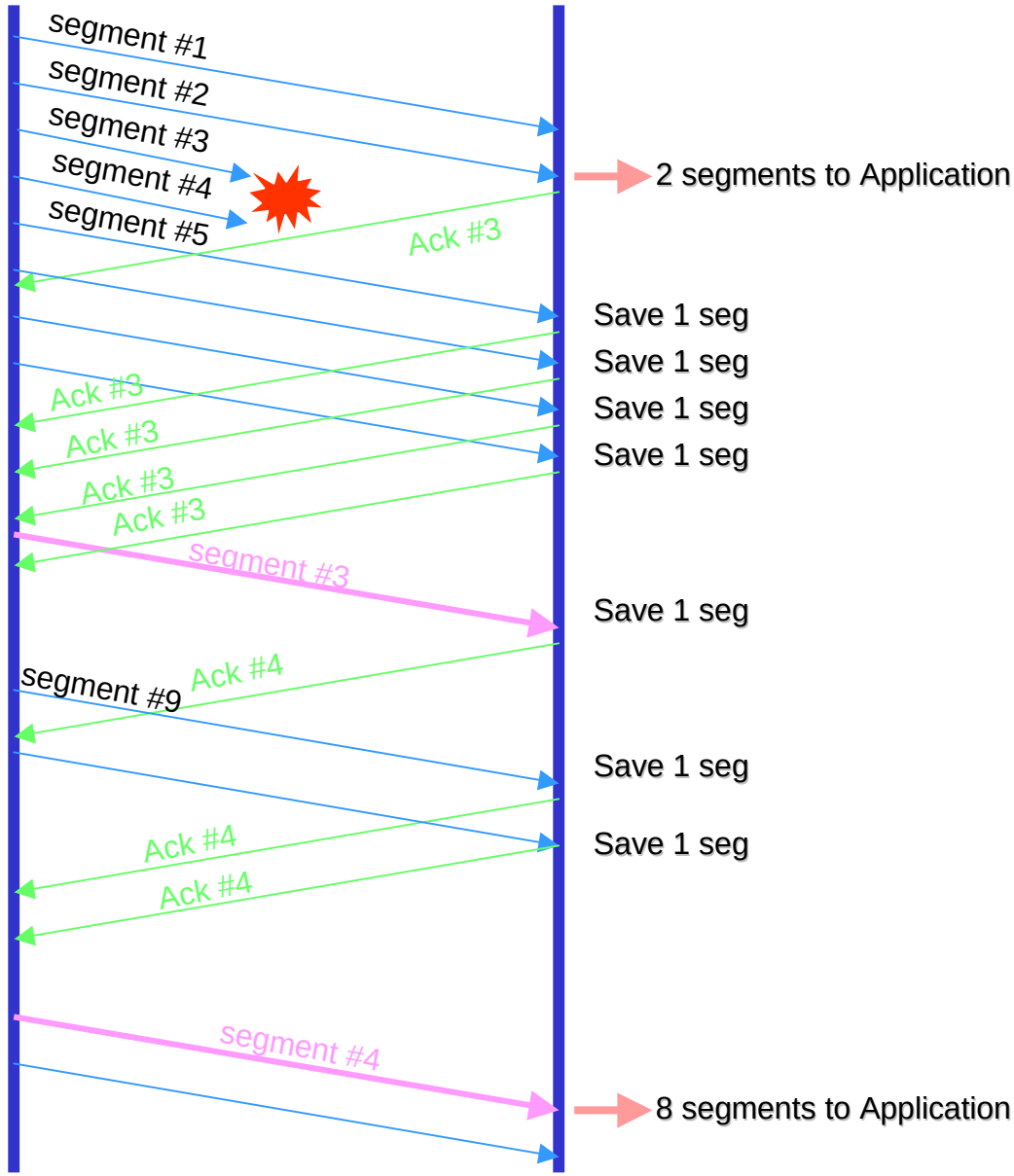


TCP Packet Loss & Retransmission



Fast Recovery

When 3 same ACKs arrive, send TCP consider a segment is lost and decides to retransmit.



TCP Timer Management

Updated RTT = αRTT + (1 - α)M

where, RTT : round-trip time between src and dest

α : smoothing factor that determines how much weight is given to the

old value. Typically α = 7/8.

M : old smoothed round trip time

TCP Timer Management

$D = \alpha D + (1 - \alpha) |RTT - M|$

where, D is smoothed deviation

Timeout = RTT + 4*D