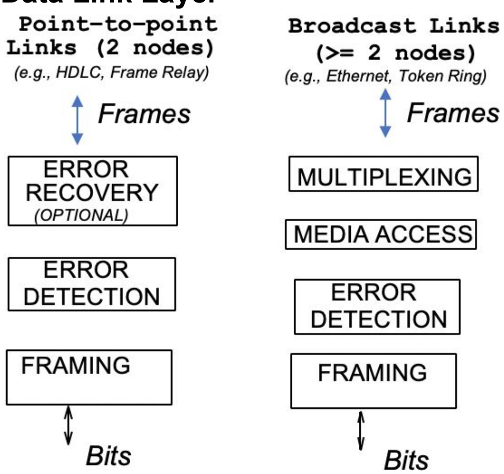
5. Application Layer - i.g Web Browser (HTTP HyperText Transfer Protocol) (FTP File Transfer Protocol) (SMTP Simple Mail Transfer Protocol)
4. Transport Layer -TCP/UDP - Segment/Datagram
3. Network Layer - IP (Every time connected with network, has its own IP address) - Routing Protocols - packet
2. Data Link Layer - Wi-Fi/Ethernet - frame
1. Physical Layer (PHY) - bits

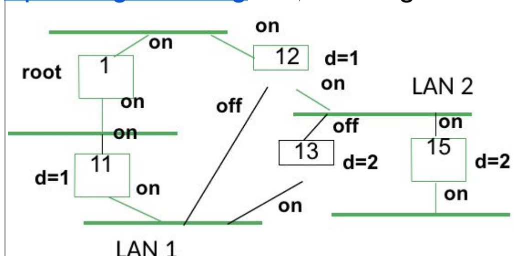
Medium	Speed	Distance Span	Pros	Cons
Twisted Pair	1 Mps -1 G (Cat 1 – Cat 5)	1 – 2 Km	Cheap, easy to install	Low distance
Digital Coax	10-100 Mbps	1- 2 km	broadcast	Hard to install in building
Analog Coax	100-500 Mbps	100 Km	Cable companies Use it now	Expensive amplifiers
Fiber	Terabits	100 km	Security, low noise, BW	No broadcast, Needs digging
Microwave	10-100 Mbps	100 km	Bypass, no right Of way need	Fog outages
Satellite	100-500 Mbps	worldwide	Cost independent of distance	250 msec delay Antenna size
RF/Infrared	1 – 100 Mbps, < 4 Mbps	1 km 3 m	wireless	Obstacles for infrared

.Nyquist limit- sluggishness affects the max signaling rates: Sending signal at rate of 2B signals/sec (max baud rate) without causing intersymbol interference (* ⇒ low bandwidth let to tight encoding (media) *)
3.Shannon theorem -may use diff voltage for diff output - noise affects the max bit rates:
Bit rate = #bit per symbol * baud rate

Data Link Layer

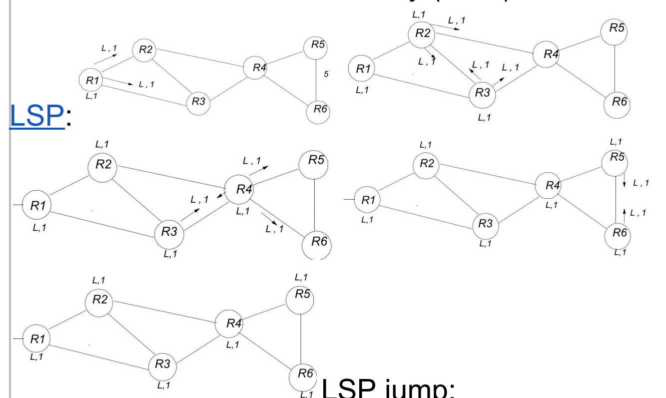


Strict Multiplexing: (Every source gets a bandwidth of B=N(# of possible sources100-1000)) Bad since traffic bursty. I.e. voice call
Statistical Multiplexing: give each user B=x(# of users that currently wish to use the system 1-10).
MAC(Media access control): 1. channel partition. 2. Random access (collisions, ALOHA, CAMA/CD) 3. Taking turns ⇒ efficient, fair, simple, decentralized
Aloha: ack ⇒ 18% max utilization
Slotted Aloha: reduces vulnerable period by half but requires a common clock⇒ 37% max util
CSMA/CD: Carrier Sense(wait until carrier ends) Multiple Access/ Collision Detection(all nodes can detect the collision)
Ethernet: CS and deference, CD, Exponential Backoff. Main Idea: **Slot time** (2T, 51.2us), **Min packet size** (64 bytes, add pad if needed), **Jam**(transmit small # of bits after detect a collision to ensure that others detect the collision), **Collision Detection**(use Manchester with average DC level per bit, collision: increase voltage)
802.11(WLAN): CSMA: RTS(backoff)/CTS
Physical Channels: 12 channels available in US. Only 3 orthogonal channels(1,6,11). Using others causes interference.
*Bridge: -address incompatibility, max packet size incompatibility, bandwidth incompatibility
+Generality, less cost, small control traffic.
Bridge view: All routers are just MAC address.
Router view: IP, MAC, Translation
Multicast and broadcast: certain group vs. all
Spanning Tree Alg: 11,12 designated bridge

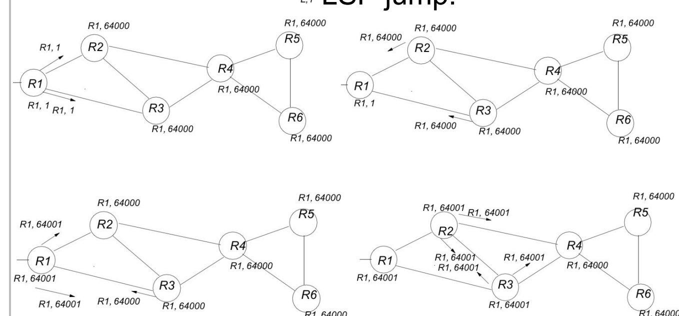


Network Layer
Why router(hierarchical)?: Bridges have to learn all addresses in an extended LAN. (more memory)
Routers only learn addresses within each level of hierarchy + Bridge Spanning Tree inefficient. (increased latency and smaller throughput.)
IP: network address + host
Supernet: CIDR scheme. Assigns new organizations multiple contiguous Class C(1H). Reduce core router table size by aggregating by a common prefix.
CIDR(classless interdomain routing) and NAT(private and global): helped the internet handle exponential growth with a finite 32 bit address space
Distance vector: count to infinity (MST)

LSP:



LSP jump:



Dijkstra's Algorithm
Intra domain routing: distance vector and link state(within AS)
*BGP(path vector): No looping(record path), apply policies. BGP chooses based on policies set by managers, which uses complex functions of the policies specified in every router.
(BGP considers a destination unreachable when all the routes to that destination have a path list that

includes this routers AS number?) -Instability -Scalability -Performance

AS relationship: Customer/Provider(ISP), Multihoming(more than one provider), peer to peer(provider->pro) (customer-> cus).

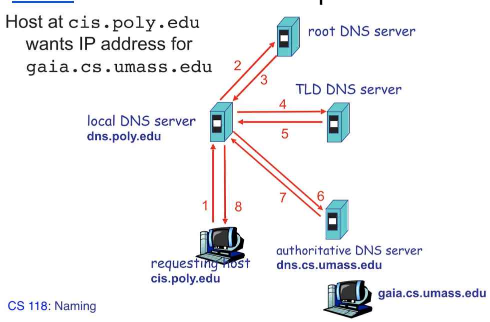
Customer/Provider: 1. Customer needs to be reachable from everyone 2. Customer does not want to provide transit service

Steps: 1. A node learns multiple paths to destination. 2. Stores all of the routes in a routing table. 3. Applies policy to select a single active route 4.... and may advertise the route to its neighbors. 5. Incremental updates unlike distance vector(Announcement and withdrawal)

BGP Attributes: **AS path:** ASs the announcement traversed. **Next-hop:** where the route was heard from. **Origin:** Route came from IGP or EGP. **Local pref:** Statically configured ranking of routes within AS. **Multi Exit Discriminator:** preference for where to exit network. **Community:** opaque data used for tag routes that are to be treated equivalently. Highest local pref, shortest AS path, lowest MED, prefer eBGP over iBGP, lowest IGP cost, router id

Types of BGP Mes: Open: Establish a peering session. Keep Alive: Handshake at regular intervals. Notification: Shuts down a peering session. Update: Announce new routes or withdraw previously announced routes.

*DNS: host name <==> ip address



DNS servers are replicated折叠的. Cache responses to decrease load

*ARP: IP ⇒ MAC 1. Broadcast: “Who has IP address x.x.x.x?” 2. Response: “MAC address yy:yy:yy:yy:yy:yy” 3. Sender caches the result in its ARP table

*DHCP: MAC ⇒ Unique IP (Automates host boot-up process) Bootstrapping problem: Host doesn’t have an IP address yet.(host doesn’t know what source address to use)/Host doesn’t know who to ask for an IP address.(host doesn’t know what destination address to use) ⇒ shout on LAN using well known DHCP multicast address (like ARP, but not broadcast) to discover server who can help + Install DHCP server on the LAN to answer distress calls

Broadcast-based LAN protocol algorithm 1. Host broadcasts “DHCP discover” on LAN (e.g. Ethernet broadcast) 2. DHCP server responds with “DHCP offer” message 3. Host requests IP address: “DHCP request” message 4. DHCP server sends address: “DHCP ack” message w/IP address

*NAT: Challenge: 1.End hosts may not be aware of external IP address 2.NAT’s end hosts are not reachable from the Internet

Transport Layer

Fragmentation: MTU discovery protocol. Send a packet with don’t fragment bit set. Keep decreasing message length until one arrives. Since frag is expensive. Memory and CPU overhead for datagram reconstruction.

Ports: server use well known port #, client use OS-assigned temporary(ephemeral) port

UDP: only multiplexing, UDP Delivery, UDP checksum(data, UDP header, IP pseudoheader). Cheaper in bandwidth and processing.

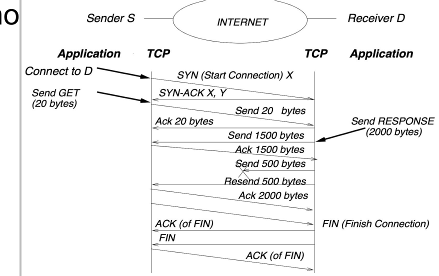
TCP: provide the abstraction of a shared queue(socket).

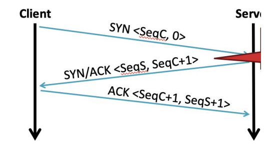
connection-oriented, flow control, congestion control - connection management, Network instead of FIFO⇒ 3 way handshakes

3-Way Handshake: each byte of data in a segment carries a sequence number. ACKs are cumulative.

*1. We wait 2*MSL (maximum segment lifetime of 60 seconds) before completing the close 2. ACK might have been lost and so FIN will be resent.

Large sequence number: TCP works over a network where packets can be delayed for large amounts of time, duplicates can be created by packet looping, and packets can be sent on different routes leading to re-ordering. ⇒ avoid duplicate and wrong order.





Flow Control: matching speed of sender to receiver speed. Adjust the window size over sliding window protocol on info from receiver.

*Can provide dynamic flow control if receiver acks indicate Lower and Upper Window Edge.+ Receiver tells the sender how big their buffer is Called the advertised window. Window may goes to 0: Need to avoid deadlock if window is reduced to 0 and then increase to c > 0. In OSI, receiver keeps sending c. In TCP, sender periodically probes an empty window.

Advertised Window0: receiver starts persist timer

Congestion Control: matching speed of sender to network speed. ⇒ reduce network load

Explicit signal: ICMP, DECBit/ECN, implicit: packet delay/loss.

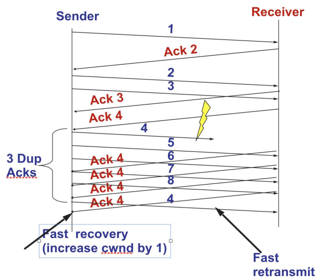
Window-based congestion control: 1. Unified congestion control and flow control mechanism 2. rwin: advertised flow control window from receiver 3. cwnd(use AIMD): congestion control window 4. Estimate of how much outstanding data network can deliver in a round-trip time 5. Sender can only send MIN(rwin,cwnd) at any time

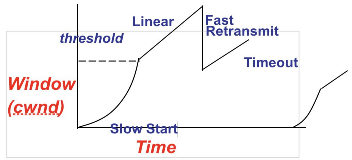
Congestion Avoidance: **AIMD**⇒ Increase sending rate by a constant (e.g. MSS). Decrease sending rate by a linear factor (e.g. divide by 2)

Slow start: ⇒ quickly find the equilibrium sending rate

Fast Retransmit: Timeouts are slow + Use 3 duplicate ACKs to indicate a loss

[Fast recovery]: avoid stalling停转 after loss + If there are still ACKs coming in, then no need for slow start





[DRR]: Multiple queue, separate the UDP and TCP ⇒ Scheduling the queue (round robin)

[Red]: drop a perfectly good packet as an early form of congestion warning if one does not have a congestion bit(ECN bit).

Link State Routing: Why link state routing depends on a primitive flooding protocol to send LSPs through the network instead of using the existing routing table to send LSPs. Existing table may out of date.

Peering: Why ISPs peer with each other though money is exchanged. Help each other to both get fast service.

DNS: Why is the DNS a good idea? No need to memorize all IP address, more human friendly.

BGP: Why BGP uses a path vector instead of distance vector. Avoid loop+apply policy

Fragmentation: Why the Packet ID field in the IP header is mostly useless today? No fragmenting packet anymore. (MTU discovery protocol)

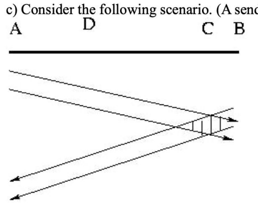
Homework

*c) Nethernet also requires the normal means of detecting collisions (i.e., more than one signal at the same point is detected by an increase in voltage) as well as the new mechanism? Explain with an example why this is needed so that all stations can detect a collision.

d) Suppose we use the mechanism in c) as well as the new Nethernet mechanism to detect collisions. Show using an example that it is still possible for some station to not detect collisions.

e) Use the results of b) and d) to show that Nethernet collisions can result in duplicate packets being received by a receiver.

c) Consider the following scenario. (A sends to B and B sends to A)

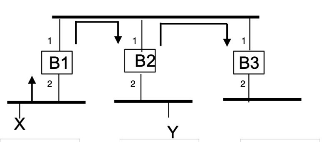


We need the old mechanism because in this example node C would not detect the collision without the old mechanism

d) Using the same diagram , node D will not be able to detect any collision as it is not transmitting and will not wait for 51.2 microsec nor will there an increase in voltage. e) If node D is the receiver (i.e A sends a packet to D in the above example) D will receive the packet. correctly as it does not detect collision: A will detect collision and will retransmit. D will accept the packet again causing duplicates

****

2) When only X updates to all bridges, it shows like the following picture.

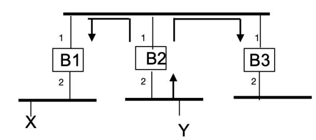


B1	
address	interface
G	2

B2	
address	interface
G	1

B3	
address	interface
G	1

when Y joins



B1	
address	interface
G	2,1

B2	
address	interface
G	1,2

B3	
address	interface
G	1

3) After Y crashed, bridges won't receive updates packets from Y, therefore after time T, they will forget the corresponding interface.

B1	
address	interface
G	2

B2	
address	interface
G	1

B3	
address	interface
G	1

4) The bridge will just forward the packages to all the interface other than the interface the package comes from.

***Ans:** Endnode B does not know the data link address of A so an ARP request is made. A receives it and sends an ARP response back to B with the incorrect data link address of 1's. B sends the data packet to A with all the 1's address is the destination Data Link address. This is sent by the bridge to all stations on the other LAN as well and all stations pick up this frame because of the (incorrect) all 1's address.

-All these nodes will pick up the frame, look at the routing address, and (except for B) will realize that its not for them. Thus, for instance, say C will try and forward the frame to A as part of being "helpful". If C does not know A's data link address, C will also try and ARP to A and the entire process will repeat with every node receiving a copy from C. Since this happens for all nodes C, if N nodes receive a copy the first time, they create N copies after the second round, each of these N copies create a further N copies, and the number of copies grows as N^k, where k is the number of iterations of this process. These kind of storms really take down Ethernets.

-If the bridge is replaced by an IP router the problem gets a bit better because the router isolates the problem to the LAN in which A is attached. Thus we have the same exponential growth but a smaller N, the number of nodes in A's LAN and not the number of nodes in the Extended LAN as before. Since routers isolate storms, that is one of the reasons cited for using routers instead of bridges.

***Ans:**Distance(P,R)=MinimumAcrossallNeighborsN of (Distance(P,N)+Distance(R,N))

And the neighbor N which achieves this minimum is the neighbor we route to, let it be N'

MinMaxPacketSize(P,R) = minimum(MinMaxPacketSize(P,N'), MinMaxPacketSize(R,N'))

Final2017

Media: Why it is still cheaper to run twisted pair to workstations though the cost of wire is fairly cheap: Optics are expensive things like LEDs lasers photodiodes

Physical Layer: If a designer of a physical link finds that there is Intersymbol Interference ISI when he tries to send bits over a link. What can the designer do. Send at a slower rate or use a better quality link that has less capacitance

Addressing: Why Ethernet addresses are 48 bits in length although most LANs have only 1000 stations? To make them globally unique so that stations can use the same Ethernet address wherever they move.

Protocol Specifications: Besides the specification of how the protocol responds to various events and the interfaces to higher and lower layers. What is the other major component of a protocol spec.The message formats bit and byte order in which message formats are to be transmitted.

Bridges versus Routers: Peter Protocol is building an application that needs low latency Peter decides he wants his network to be full of routers although the bridges are slightly faster Explain why? Routers offer shortest path routing which can be less hops that routing along a spanning tree

Spanning Tree Protocol: Although we did not tell you this in class. bridges timeout learned addresses faster after a spanning tree topology change. Why? Because normally the reason to time out an address is because a station physically moves which can take minutes; however a spanning tree change can make a large group of stations change sides wrt a bridge in seconds

Link State Routing: Why a source S sending a link state packet may get a packet with source S and a higher sequence number than S is currently using? may have been sending a large sequence number before crashing and will (by the intelligent flooding rules) cause other routers to send back the old number to [LSP jump]

Distance Vector Routing: Hugh Hopeful suggests stopping the count_up of Distance Vector when the distance reaches the diameter of the network. What is the problem with Hugh's suggestion. Diameter is not well defined if we have node or link failures. A network in the shape of a wheel with a central router connecting every node and where every node is also connected in a ring can have a diameter of two,If the central router fails then the diameter increases to halve the number of nodes.

Congestion Control: Why can the throughput of a network go to zero (congestion collapse) if too much traffic is allowed to enter the network? Because the network can be filled with traffic all of which reaches part of the way to the destination and gets dropped because of other traffic that has a similar property.

Transport: What resources does a transport connection consume at a workstation even when the user of the connection is not sending any data.Bandwidth for sending keep alive messages and memory in connection tables

Que: BGP versus distance vector Explain briefly, the main differences between BGP and Distance vector in terms of a) how routes are chosen b) how routes are considered to be Unreachable. Distance vector always chooses shortest path routes BGP chooses routes based on policies set by managers_ since routers only pass routes that fit their policy to other routers the result is a complex function of the policies specified in every router. Distance vector considers a destination unreachable when the distance goes beyond some limit (e.g., 16 in RIP) BGP considers a destination unreachable when all the routes to that destination have a path list that includes this routers AS number.

Que: Bridging and Learning Hugh Hopeful notices that at very high speeds it is hard for bridges to learn information from the source addresses in every packet. So Hugh suggests that bridges look at source addresses only in multicast packets_ Since routing endnode protocols typically ensure that endnodes send multicast packets (e.g. ARPs_OSI hello) this should ensure, that each bridge periodically hears a multicast packet from each endnode_ Also since multicast traffic is so much less than non multicast traffic_ the processing load on bridges to do learning will be considerably reduced_ Peter Protocol_ who is brought in as a consultant points out that not all endnodes send multicast periodically

1. As usual, bridges will flood unknown destination frames What is one disadvantage of using Hugh's scheme of learning from multicast messages only based on Peter Protocol_s comment If a station X does not send multicast all frames addressed to X will be flooded causing unnecessary traffic

2. All IEEE 802 LANS are supposed to support the SYSID_REQ message. When a station X on a LAN sends a SYSID_REQ message to the broadcast address all stations are supposed to send a SYSID_RESP message back to X. This can be used for instance by a manager to and how many stations there are on a LAN. How can Hugh use the SYSID_REQ message to avoid Peter Protocol's objection Every bridge periodically sends a SYSID_REQ message to the broadcast address on all ports. If station Y sends a SYSIDRESP message back to bridge X that arrives on Port m of bridge X then bridge X learns that Y is reachable through Port m

3. Would the SYSID scheme work well in a large Extended LAN with 8000 stations? The SYSID-RESP from stations like Y that do not send multicast may be lost in the flood of messages caused by 8000 responses many of which the bridge already has information about.

Que: It is also theoretically possible to not limit ourselves to equal cost paths. For example, in the figure above there is a path of cost 5 between R0 and R6 through R5. It seems that we could do better load balancing by having R0 send a small fraction of its packets through R5 as well. However, this kind of load balancing can lead to packet looping unless care is taken. Explain why.
At each hop on the path the packet may be routed along a path longer than the shortest cost.But routing along shortest cost paths is the only way to ensure progress and avoid loops

Que: 1) The algorithm used by a router to reply to a QUERY is trickier than you might think. It is obvious that R5 already knows that R1 and R3 are the best ways for R5 to get to D. However, S may choose to ask R1. How is R1 to know that R3 is also an equally good way to get to D? Assume the use of distance vector routing.
Since R1 is using distance vector it knows the set of all neighbors (including R3 to D). Thus router R1 can easily calculate the set of neighbors that are on the same LAN as S and R1 that have the same cost to D as R1 It then sends this info to S.

2) Suppose S has a cache entry for D that says the best two routers are R1 and R3. Then the link from R1 to R2 crashes. R1 quickly calculates that the

best route to D is through R3 but S may still have an old cache entry. How should R1 react when S sends a packet from D to R1. How can S use this information to update its cache?
R1 can send a REDIRECT to S saying that its sending packets to D through R3. R1 then removes R1 from its cache of equal cost routers to get to D.

Que: **Modifying Transport Protocols to Deal with Load Balancing**: Hugh Hopeful uses a sliding window transport protocol. over a routing protocol and everything works fine. Hugh Hopeful later modifies the routing protocol so that it can do load balancing as shown above. However, he finds that performance actually decreases when he does load balancing.

1)Why does performance go down? Performance goes down because packets are not being buffered out of order and are being dropped.

2)What simple change does Hugh need to make to his transport protocol implementation? He needs to buffer out of order packets at the very least and switch to selective reject at the very best.

Que: **Reverse Path Congestion Control:** Recall that in conges- tion control we had two separate problems. A router had to sense congestion on a link and then send feedback to all sources using the link. In class, we described the DECbit/ECNbit scheme in which the bit is passed to the destination and then back to the source. Here we describe another scheme in which a congestion bit is passed from the router directly back to the source.

In the figure below, assume that S is sending packets to D through R1 and all acks from D return on the same path. The reverse path congestion rule is as follows: if a packet p is received from a link l that is congested in the outbound direction, then a congested bit is set in the routing header of packet p. Thus in the figure, if the link from R1 to R2 gets congested, then the outbound queue at R1 and going to R2 will build up. When any packet from D to S arrives on this link (for example an ack), R1 will set the congested bit and this bit will get to S which then can send at a slower rate.

1) What is one advantage of reverse path congestion control over the DECbit/ECNbit scheme? It provides faster feedback from point of congestion directly back to the source instead of through destination) It also does not require a path to pass bits from the congested router to the destination.

2)The correctness of reverse path congestion control depends on an assump- tion. What is the assumption and why does it not always hold for all routing algorithms? It assumes that the route from S to D is the same as that from D to S. This is not guranteed by distance vector and link state because when there are many equal cost routes from S to Dand vice versa the S to D and D to S calculations can pick different routes

---

Assume that all the following algorithms are implemented in TCP congestion control: slow start, congestions avoidance, fast retransmit, fast recovery, and retransmission upon timeout. If ssthresh equals to cwnd, use the congestion avoidance algorithm in your calculation.

- Initially ssthresh at the sender is set to 4. Assume cwnd and ssthresh are measured in segments, and the transmission time for each segment is negligible. Retransmission timeout (RTO) is initially set to 500ms at the sender and is unchanged during the connection lifetime. The RTT is 100ms for all transmissions.
- The connection starts to transmit data at time t = 0, and the initial sequence number starts from 1. Assume each segment consumes 1 byte. Segment with sequence number 13 is lost once. No other segments are lost. How long does it take, in milliseconds, for the sender to receive the ACK for the segment with the sequence number 20? show your intermediate steps or your diagram.

-What happens if segment 13, 14, 15, and 16 are lost? ⇒ wait timeout to retransmit.

Total 6 RTTs: 600ms