1. **Retransmission of lost data can be done at the link, transport, and application layers. What are the pros and cons of doing it at each layer?**

Answer: Here we analyze them separately.

Link layer:

Advantages:It improves the performance because the receiver can wait a new packet other than this old packet. And at the same time, link layer retransmissions avoid having to retransmit over the entire path.

Disadvantages: It doesn't estimate the need of the higher layers to retransmit, and even cause bad result to upper layers. In this layer, retransmission is complex and overhead.

Transport layer:

Advantages:It provide end-to-end retransmission, so it doesn't need to rely on the link layer to achieve correctness. Another good result is that many applications can use the same transport protocol implemented in a library or in the operating system.

Disadvantages: Obviously, the efficiency of sending packets is low. Another point is that we can not guarantee the reliability of application layer.

Application layer:

Advantages:The application is the most suitable position to know what piece of data needs to be retransmitted and what is the suitable timeline in this period.

Disadvantages: It push every application has to retransmit, which means there is a waste in using resources.

2. **Compare Implicit versus Explicit congestion signals. What are the advantages and disadvantages of each?**

Answer:

For implicit congestion signals:
Advantages: 1. An implicit congestion signal does not require route support because it based on in the internet. 2. An implicit congestion signal require transport protocols adapt them anytime, which means it can not cause a lack of congestion.
Disadvantages : 1. Congestion is detected when the sender detects that some packet was loss. Packet loss only can be detected through timeout. It means that this method is a little incorrect.

For explicit congestion signals:
Advantages: 1. An explicit congestion signal requires lower latency to interpret than implicit congestion signals. 2. An explicit congestion signal can be used for congestion avoidance. Routers can detect signal congestion when their queues

grow long, but before they have to drop packets.
Disadvantages : 1. Explicit signals requires the assistance of routers and therefore, it is NOT useful practically because present day Internet routers do not flip bits in packets

3. What are the main functions of the transport layer? Describe briefly.
Answer:

1) Service point addressing: Transport layer header includes service point addressing which is port address. This layer gets the message to the correct process on the computer unlike network layer, which gets each packet to the correct computer.
2) Segmentation and reassembling: A message is divided into segments; each segment contains sequence number, which enables this layer in reassembling the message. Message is reassembled correctly upon arrival at the destination and replaces packets which were lost in transmission.
3) Flow control: In this layer, flow control is performed end to end.
4) Error control: Error control is performed end to end in this layer to ensure that the complete message arrives at the receiving transport layer without any error. Error correction is done through retransmission.

4. How does the transport layer perform multiplexing and demultiplexing?

Answer:
For demultiplexing: Each transport-layer segment has a field that contains information that is used to determine the process to which the segment's data is to be delivered. At the receiving end, the transport layer can then examine this field to determine the receiving process, and then direct the segment to that process. This job of delivering the data in a transport-layer segment to the correct application process is called demultiplexing.

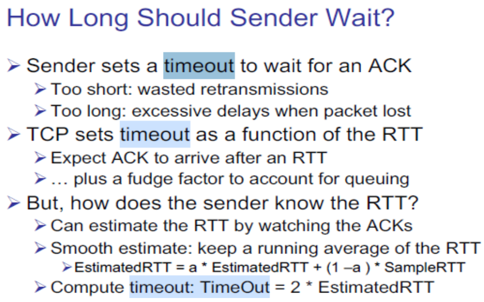
For multiplexing: The job of gathering data at the source host from different application processes, enveloping the data with header information to create segments, and passing the segments to the network layer is called multiplexing.

5. Why does TCP wait for three duplicate acknowledgments before retransmitting a packet? What do the triple duplicate acks represent?

Answer:
Because TCP does not know whether a duplicate ACK is caused by a lost segment or just a reordering of segments, it waits for a small number of duplicate ACKs to be received. It is assumed that if there is just a reordering of the segments, there will be only one or two duplicate ACKs before the reordered segment is processed, which will then generate a new ACK. If three or more duplicate ACKs

are received in a row, it is a strong indication that a segment has been lost. TCP then performs a retransmission of what appears to be the missing segment, without waiting for a retransmission timer to expire.

6. How does TCP set its timeout value?



Answer: Firstly, the sender set an initial timeout value wo wait for an ACK. Secondly, TCP sets a timer at the sender and if the data is not acknowledged before the timer expires, the sender resend the data. We can get the duration of Round Trip Time (RTT). Finally, the timeout value is about 2 times of RTT.

7. TCP congestion avoidance is done via AIMD. Explain.

Answer:
TCP congestion avoidance is based on AIMD, which stands for Additive Increase Multiplicative Decrease. This means that TCP increases the congestion window for every window of packets acknowledged and cuts the congestion window in half for every packet that is lost (or inferred lost due to receiving triple duplicate acks). The idea of AIMD is to probe the network congestion limits by increasing the congestion window and then back off when loss occurs. This cycle of probe and backoff is the main principle behind TCP congestion avoidance.

8. What is goal of network fairness? Is TCP fair? If so, explain what resources TCP allocates in a fair manner.

Answer:
Fairness in computer networks deals with the distribution of network resources among applications; i.e., fairness is achieved when network resources are distributed in a fair way. Investigating fairness in computer network aims at two goals. The first goal is to improve the behavior of networking architectures by adding the valuable concept of distributing resources fairly, which should be considered both for existing and for new scenarios. The second goal is to enable new (fair) applications that are currently not implemented in existing networks for various reasons.

Certainly, TCP is fair.

TCP fairness requires that a new protocol receive no larger share of the network than a comparable TCP flow. This is important as TCP is the dominant transport protocol on the Internet, and if new protocols acquire unfair capacity they tend to cause problems such as congestion collapse

9. What is the throughput of TCP? The throughput is the average rate that packets are successfully decoded at the receiver. Note that the rate that packets are sent by the sender is an upper bound on the actual throughput and since it is easily computable, we use it to estimate the throughput.

Answer:
The sending rate of a flow is computed as follows. A flow with a window size of MSS packets has a throughput of (MSS /RTT) bit/second. The RTT here is the total round trip time (in seconds), consisting of all possible delays, such as those due to processing, propagation, queuing and transmission delays.

But when we consider the loss in transformation, we should consider it.

analysis of the congestion-control behavior of TCP has shown t
steady state, TCP's throughput is approximately

$$Rate = \left(\frac{1.2 \times MSS}{RTT \times \sqrt{\rho}}\right)$$

So the result is $throughput = (\frac{1.2 \times MSS}{RTT \times \sqrt{\rho}})$

10. We said that the goal of the transport layer (layer 4) was end-to-end reliability. Recall that layer 2 also has reliability, but it is hop-by-hop reliability. Why do we have reliability at both layers? Are they both necessary?

Answer:

1. The reliability at two layers is different. For Layer2, it is hop-by-hop reliability. As for Layer4, the reliability more focus on end-to-end.
2. At the Layer4, TCP has the advantage of running on top of IP, which allows the routing of packets over multiple hops to reach a final destination,
3. AT the layer2, the protocols only focus on single hops which are related to themselves.
4. Only we both guarantee the reliability on two layers, the internet service is the best.

In a word, they are both necessary.