THE HONG KONG POLYTECHNIC UNIVERSITY

Department of Computing

This is an open-book examination.

(COMP5311)

Internet Infrastructure and Protocols

14 December 2009 3.5 hours

[Answer at most 7 questions in section A and both questions in section B.]

Section A: Please answer AT MOST SEVEN questions in this section [8 marks each, making up a total of 56 marks out of 100]. You should always attach a succinct explanation to your answer.

---

1. (IP subnet mask) An IP network is configured with a subnet mask of `255.255.254.0`. A host and four routers (R1-R4) are connected to the subnet. The host is configured with an IP address of `158.132.11.10`; and R1, R2, R3, and R4 are configured with `158.132.10.20`, `158.132.11.30`, `158.132.12.40`, and `158.132.13.50`, respectively. Which router(s) could be the host's default router(s) and why?

2. (IP forwarding) Consider an IP network in Figure 1 in which two IP subnets (subnets 1 and 2) have been configured. In each host's routing table, there are two main routing entries: one to its own subnet and the other one to its default router.

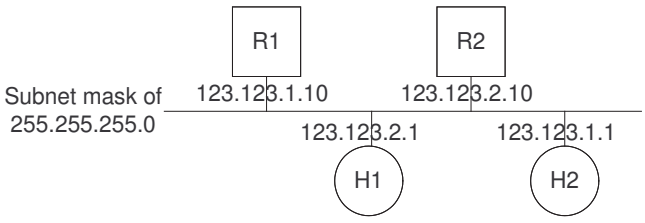


Figure 1: A network of two IP hosts and two IP routers.

   (a) (3 marks) If H1 sends an IP packet to H2, how is the packet forwarded to the destination?

   (b) (3 marks) If H1 sends an IP broadcast packet to its own IP subnet using `123.123.2.255`, will the packet be received by the IP layer of H2? Explain your answer.

   (c) (2 marks) In (a), is there an ICMP redirect message sent out by a router? Explain your answer.

3. (ARP) Consider that there are four hosts A, B, C, and D on a network. Initially, their ARP caches are all empty. Consider the following sequence of IP packet transmissions: (1) $A \rightarrow C$, (2) $D \rightarrow B$, (3) $C \rightarrow A$, (4) $B \rightarrow D$, and (5) A broadcasts locally. At the end of each successful packet transmission, what is the ARP cache content (i.e., which IP host's IP-MAC address binding) in each host? Assume that none of the cache items, if any, expire. Explain your answers.

4. (ICMP error messages) Answer the following questions concerning ICMP error messages which report to the sender of an offensive packet on various delivery errors, such as destination/port unreachable, packet too big, time exceeded, etc.

   (a) (2 marks) If the offensive IP packet is a non-tunneled packet without IP option, and it carries a TCP segment, could the sender determine whether the offensive IP packet contains a TCP RESET segment? Explain your answer.

   (b) (3 marks) If the offensive IP packet is a non-tunneled packet with IP option, and it carries a TCP segment, could the sender determine the port numbers of the TCP packet encapsulated by the offensive IP packet? Explain your answer.

   (c) (3 marks) If the offensive IP packet is a tunneled packet (i.e., IP in IP), is it possible for the tunnel encapsulator to determine the source IP address of the original IP packet encapsulated in the IP tunnel? Explain your answer.

5. (Traceroute ping) Figure 2 shows the first ping packet for traceroute purpose. Figure 3 shows the reply packet elicited by the ping packet. Fill in the missing fields (a)-(h) in Figure 3. Note that some fields in the reply packet are still missing.

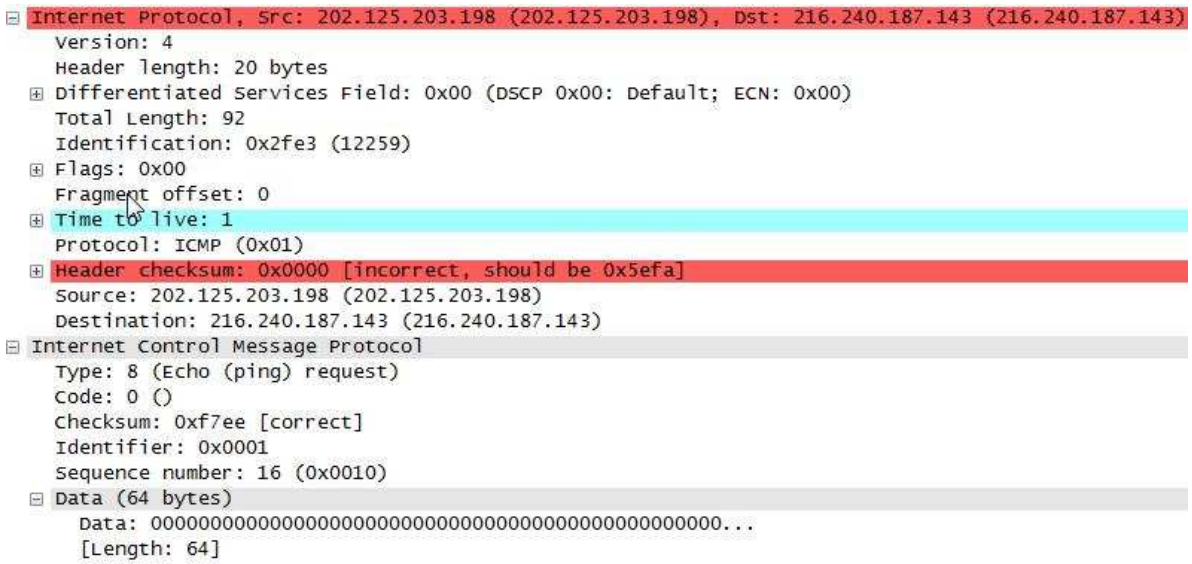


Figure 2: The first ping packet.

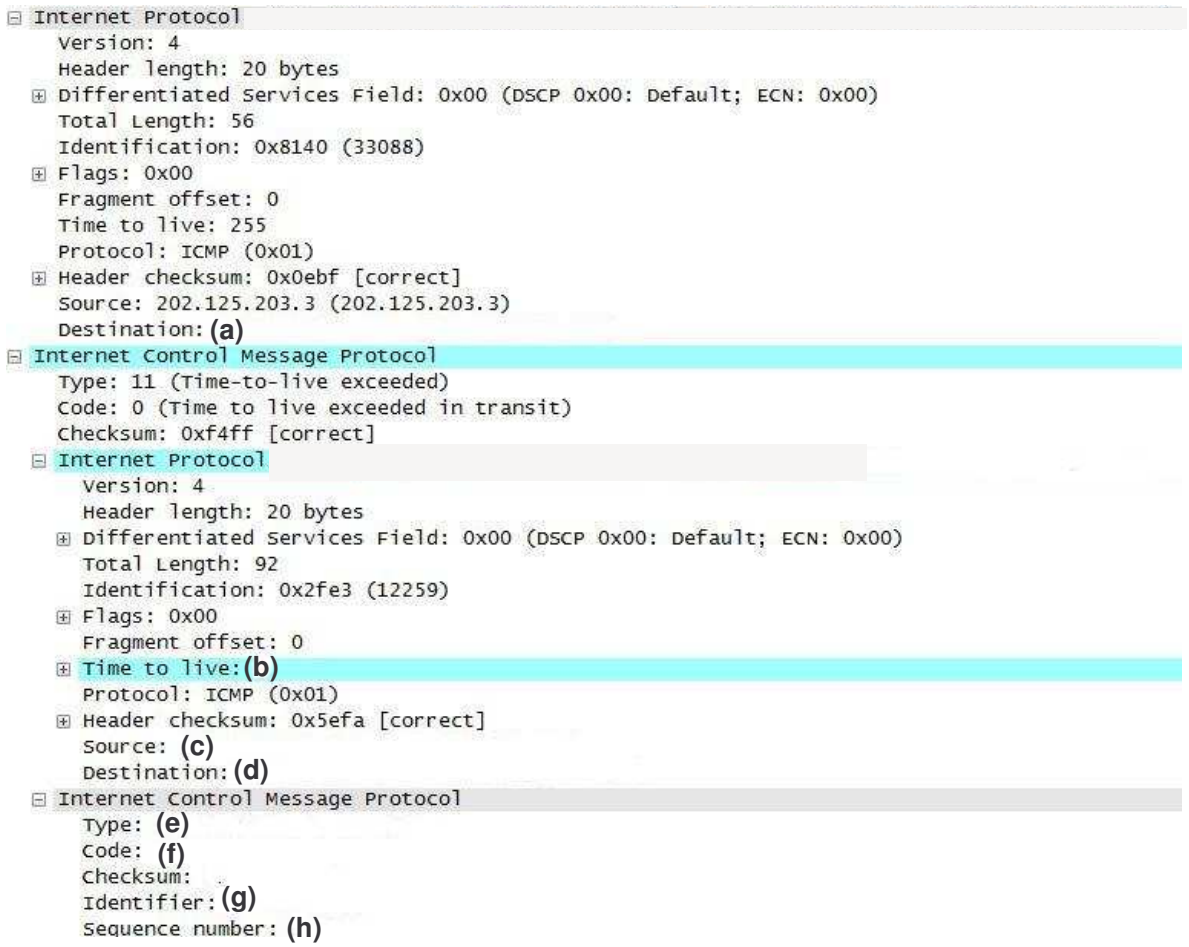


Figure 3: The reply to the first ping packet.

6. (IP tunnels) RFC 792 says that if a host reassembling a fragmented datagram cannot complete the reassembly due to missing fragments within its time limit, it discards the datagram and may send a time exceeded error message. Consider the network in Figure 4 in which there are two nested IP tunnels. Who will send the ICMP error message and whom is the message sent to for each scenario below? Explain your answers.

   (a) (2 marks) A packet is only fragmented in A, and one of the fragments is lost.

   (b) (3 marks) A packet is fragmented in A and then one of the fragments is fragmented again in B, and one of the latter fragments is lost.

   (c) (3 marks) A packet is fragmented in A and then one of the fragments is fragmented again in C, and one of the latter fragments is lost.

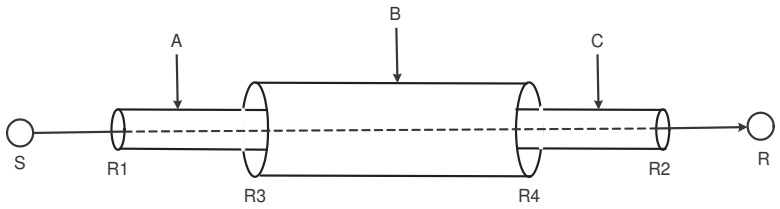


Figure 4: Two nested IP tunnels.

7. (TCP transmissions) Figure 5 shows a Wireshark trace of a TCP connection. Assume that both sides use an initial sequence number (SN) of 0.

   (a) (2 marks) What is the SN of the TCP segment 10?

   (b) (2 marks) What is the acknowledgment number (AN) of the TCP segment 10?

   (c) (2 marks) What is the SN of the TCP segment 9?

   (d) (2 marks) What is the AN of the TCP segment 9?

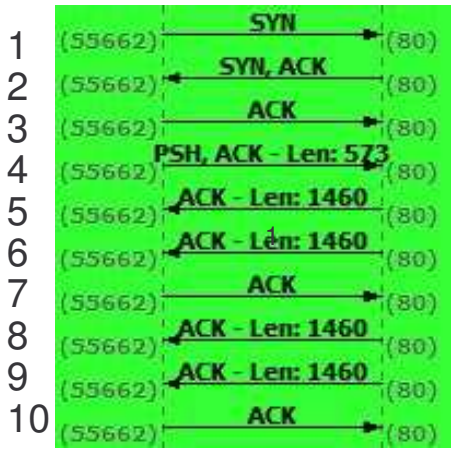


Figure 5: TCP packet transmissions.

8. (TCP congestion control) Referring to Figure 6 for TCP packet transmissions in a TCP Reno connection, a number of full-sized TCP data segments were sent from a sender to a receiver, and the segments are numbered starting from 0. Each ACK acknowledges only one TCP data segment. The square symbol refers to a TCP data segment transmission, whereas a small dot refers to an ACK transmission. The TCP data segment and its ACK are drawn on the same line for easy reference (an example is shown for segment 10). TCP Reno implements the fast retransmission and fast recovery algorithms that we discussed in class. The figure shows that only segment 13

was lost, and it was fast retransmitted after receiving a third duplicate ACK. Note that there were a number of duplicate ACKs received. The send window was always given by the `cwnd`, and the sender always had data to send.

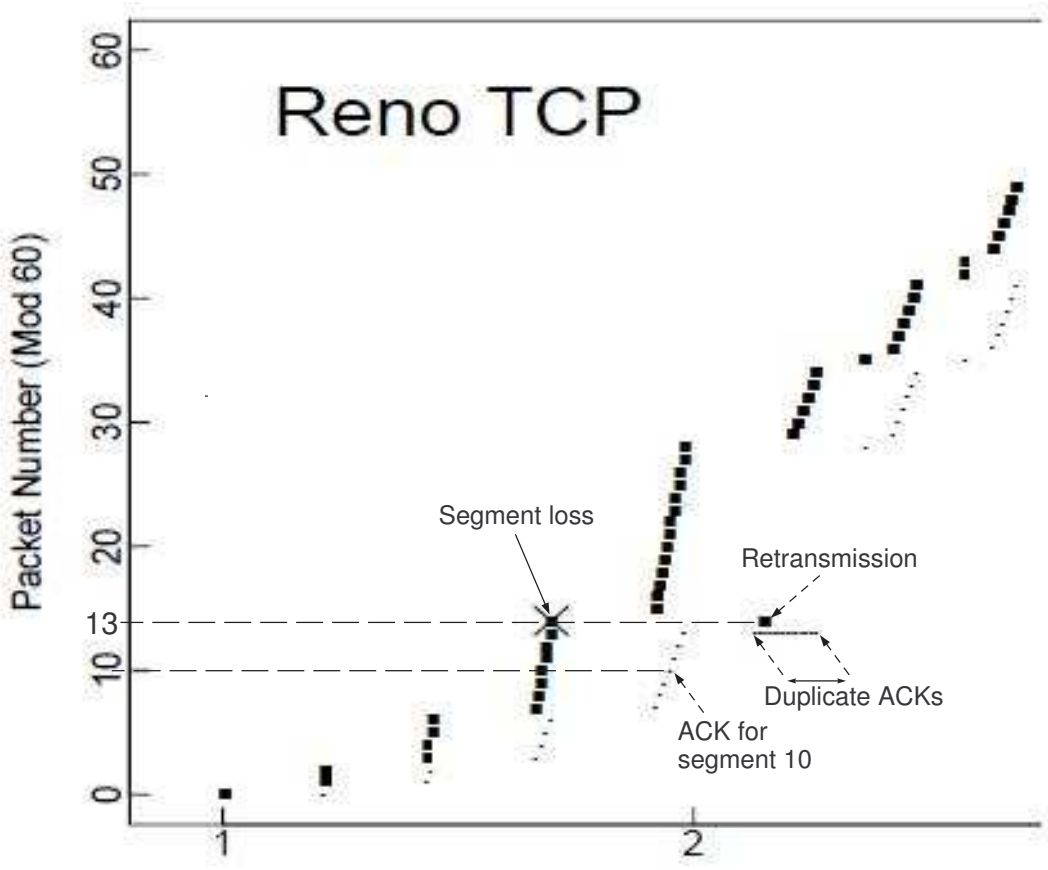


Figure 6: Packet transmission sequence of a TCP Reno connection.

Answer the following questions about the trace in Figure 6.

(a) (2 marks) How many ACKs were received for segment 13? Explain your answer.

(b) (2 marks) What was the size of the `cwnd` (in terms of the number of data segments) immediately after starting the fast recovery phase? Explain your answer.

(c) (2 marks) At what time (in terms of an event) was the fast recovery phase ended? Explain your answer.

(d) (2 marks) At the time that the fast recovery phase was ended, what was the size of the `cwnd`? Explain your answer.

9. (Distance vector) Consider that three RIP routers R1, R1, and R3 are connected on layer two and they exchange their distance vectors for a certain prefix. Initially, they announce (R1, 4), (R2, 6), and (R3, 8) to one another using IP broadcast or multicast. Recall that a distance of 16 is designated as "unreachable" in RIP. Assume that split horizon with poisonous reverse is used.

   (a) (3 marks) When the routing protocol converges and the network topology does not change, what are the distance vectors announced on the network? Explain your answer.

   (b) (3 marks) After receiving a better route from R1, how do R2 and R3 further announce this route?

   (c) (2 marks) After some time, R1 fails and stops sending distance vectors, what are the distance vectors announced by R2 and R3 after the routing protocol converges? The network topology remains unchanged.

10. (Link state) Recall from the slides on link-state routing protocol that the backbone database for the example OSPF network with areas is:

```
                **FROM**

          |RT|RT|RT|RT|RT|RT|RT
          |3 |4 |5 |6 |7 |10|11|
       -------------------------
       RT3|  |  |  |6 |  |  |  |
       RT4|  |  |8 |  |  |  |  |
       RT5|  |8 |  |6 |6 |  |  |
       RT6|8 |  |7 |  |  |5 |  |
       RT7|  |  |6 |  |  |  |  |
   *  RT10|  |  |  |7 |  |  |2 |
   *  RT11|  |  |  |  |  |3 |  |
   T    N1|4 |4 |  |  |  |  |  |
   O    N2|4 |4 |  |  |  |  |  |
   *    N3|1 |1 |  |  |  |  |  |
   *    N4|2 |3 |  |  |  |  |  |
        Ia|  |  |  |  |  |5 |  |
        Ib|  |  |  |7 |  |  |  |
        N6|  |  |  |  |1 |1 |3 |
        N7|  |  |  |  |5 |5 |7 |
        N8|  |  |  |  |4 |3 |2 |
 N9-N11,H1|  |  |  |  |  |  |1 |
       N12|  |  |8 |  |2 |  |  |
       N13|  |  |8 |  |  |  |  |
       N14|  |  |8 |  |  |  |  |
       N15|  |  |  |  |9 |  |  |
```

   Draw a spanning tree rooted at RT6 for only internal networks (i.e., N1-N4, N6-N8, and N9-N11) which is the result of running the OSPF protocol.

Section B: Please answer both questions in this section. Each question carries 22 marks.

---

11. This question concerns two mechanisms for the IPv4 to IPv6 transitions.

    (a) (IPv6/IPv4 protocol translation) A protocol translation approach has been proposed for IPv6-only nodes to communicate with IPv4-only nodes, as depicted in Figure 7.

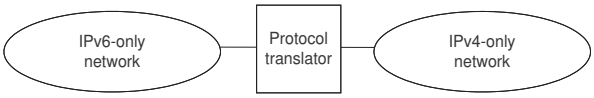
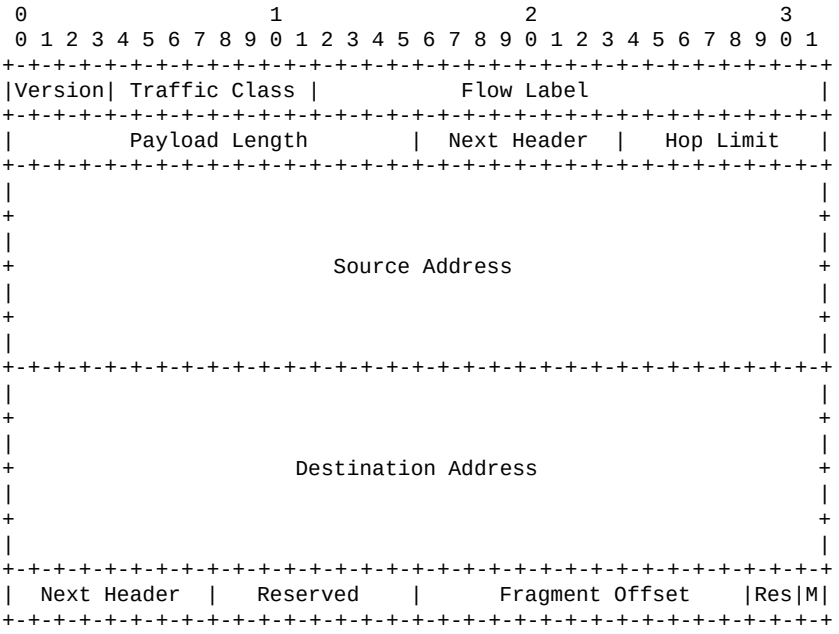


Figure 7: Protocol translation between IPv4 and IPv6.

The idea behind this approach is for the translator to translate an IPv6 packet that is destined to an IPv4-only node into a corresponding IPv4 packet transparently. There are quite a number of issues to consider, for example, address translation. To make it simple, here we only concentrate on the case that the translator receives such an IPv6 packet with a fragment extension header, and there are no other extension headers. An IPv6 header and a fragment extension header are shown below. The ID, Fragment Offset and M bit in the fragment are similar to those in the IPv4 packet. The Next Header contains a protocol ID (e.g., 6 for TCP) to indicate the IPv6 payload type.



```
|                         Identification                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
+                                                               +
+                         IPv6 payload                          +
```

Describe how the translator fills in the following fields in the corresponding IPv4 packet (the Total Length field has already been filled in for you):

i. Total Length: Payload length value from the IPv6 header minus 8 for the fragment header, and plus the size of the IPv4 header.

ii. (3 marks) Identification

iii. (3 marks) Flags

iv. (2 marks) Fragment offset

v. (2 marks) Protocol

(b) (IPv6 tunneling) This question concerns IPv6 tunneling which creates a "virtual link" between two IPv6 nodes, as depicted in Figure 8.

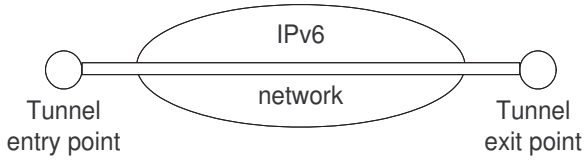


Figure 8: An IPv6 tunnel.

- The payloads of the tunnel packet can be IPv6 packets or IPv4 packets.
- A tunnel entry point, servicing like a source of IPv6 packets, must support fragmentation of tunnel IPv6 packets.
- A tunnel intermediate node that forwards a tunnel packet must not fragment a packet undergoing forwarding.
- Each tunnel is characterized by a tunnel MTU which is the path MTU between the tunnel endpoints minus the size of the tunnel header.

It is very important to note that IPv6 requires that every link in the IPv6 network must have an MTU of at least 1280 bytes. Answer the following questions.

i. (6 marks) When an original IPv6 packet enters the IPv6 tunnel, and if the original packet size exceeds the tunnel MTU,

If (the original packet size $>$ 1280 bytes),
the tunnel entry point discards the packet and sends an ICMPv6 "Packet Too Big" message to the source address with the recommended MTU size equal to

_____

Else

_____

ii. (3 marks) When an original IPv4 packet enters the IPv6 tunnel, and if the original packet size exceeds the tunnel MTU,

If (the Don't Fragment bit is set in the original IPv4 packet),
the tunnel entry point discards the packet and returns an ICMP "packet too big" message with the recommended MTU = tunnel MTU.

Else

_____

iii. (3 marks) In (i), assuming that the sending host receives the ICMPv6 "Packet Too Big" message and sends an IPv6 packet using the recommended MTU, is it possible that this IPv6 packet will be fragmented by the tunnel entry point? Explain your answer.

12. Figure 9 shows TCP segment and ACK transmissions for a TCP Reno connection. Please refer to question 8 on how to read this graph. Unlike question 8, there were two TCP segment losses (segments 14 and 28). The lost segment 14 was recovered by a fast retransmission. During the fast recovery, five more segments were sent (segments 29-33). The fast recovery phase was ended when receiving an ACK for segment 27. In the following we focus on the subsequent TCP segment (re)transmissions and the sender's `cwnd`. The send window was always given by the `cwnd`, and the sender always had data to send.

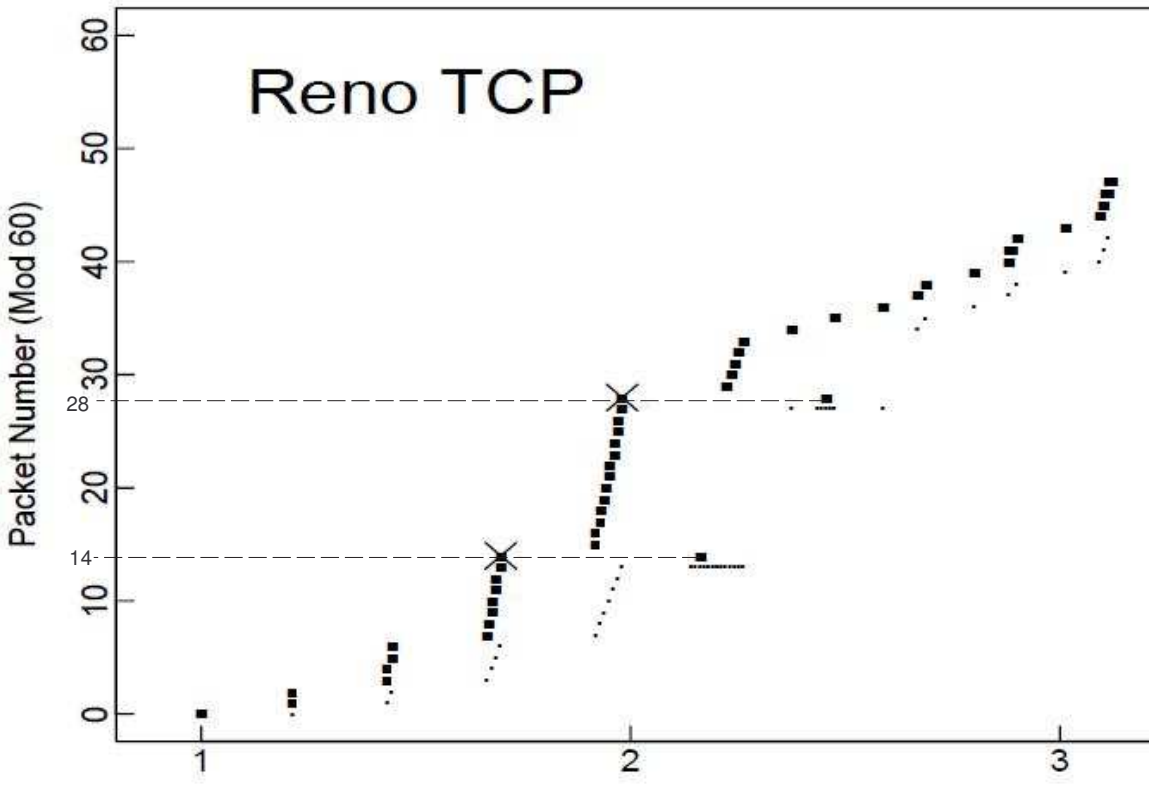


Figure 9: TCP segment and ACK transmissions of another TCP Reno connection.

(a) (3 marks) What event was responsible for the first ACK for segment 27?

(b) (3 marks) When the first ACK for segment 27 was received, the sender was able to send only segment 34. Based on this observation, what was the sender's `cwnd` at the time of sending segment 34?

(c) (3 marks) Segment 28 was fast retransmitted upon receiving the third duplicate ACK for segment 27. During the fast recovery phase, how many new segments were sent out and what were they, if any?

(d) (3 marks) What event was responsible for the last duplicate ACK for segment 27?

(e) (3 marks) When was the second fast recovery phase ended (in terms of event)?

(f) (3 marks) What was the sender's `cwnd` when the second fast recovery phase was ended?

(g) (4 marks) In later part, the figure shows that a new ACK could allow transmitting only a new segment (instead of two). Give an explanation for this observation.

— End of the Examination Paper —