University of California at Berkeley
College of Engineering
Department of Electrical Engineering and Computer Sciences

EE122
MIDTERM EXAMINATION
Thursday, 11 October 2012
Scott Shenker

INSTRUCTIONS—READ THEM NOW! This examination is CLOSED BOOK/CLOSED NOTES. You will not require a calculator, iPhone®, laptop computer, or other calculation aid. Please put them away right now! You MAY use one 8.5"x11" double-sided crib sheet, packed with notes, formulas, and diagrams (but no smaller than 8pt font). All work should be done on the attached pages, and there are four blank pages at the end for you to use as scratch paper. Don't be scared off by the number of questions; many of them are simple warm-up exercises, and you should have plenty of time to finish them all.

In general, if something is unclear, write down your assumptions as part of your answer. If your assumptions are reasonable, we will endeavor to grade the question based on them. If necessary, of course, you may raise your hand, and a TA or the instructor will come to you. Please try not to disturb the students taking the examination around you.

Please write your SID on each page!

Name – Please Print! ______________________ (Signature) ______________________

SID | | | | | | | | |

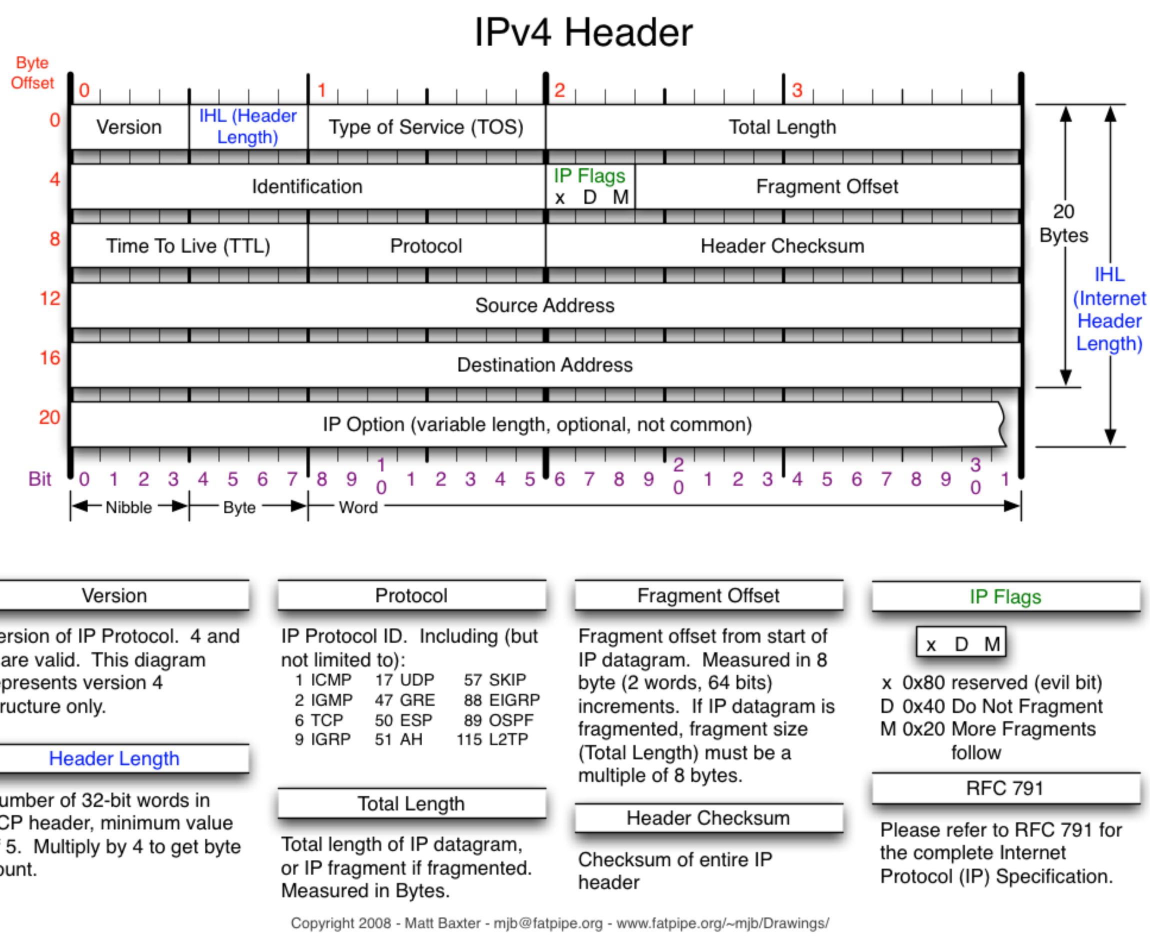
Discussion Section (Day/Time) ______________________

Question		Possible Points	Points Obtained
1	True or False	20	
2	General Multiple-Choice	20	
3	TCP Basics	10	
4	Sequence of Messages	15	
5	The Real World	6	
6	Timer Values	8	
7	Addressing	14	
8	Learning Switch	12	
9	DNS	4	
10	Sliding Window	10	
	Bonus Questions	0.00001	
Total		**119.0001**	

Fun Facts to Remember

The IP header, without options, is 20 bytes
The TCP header, without options, is 20 bytes
The UDP header is 8 bytes.

1 TB = 10^{12} bytes; 1 GB is 10^{9} bytes; 1 MB is 10^{6} bytes
1Gbps is 10^{9} bits/sec; 1Mbps is 10^{6} bits/sec
1 msec is 10^{-3} seconds



1. True or False [20 points]

Pick one and only one answer for each statement (indicate by circling)

1.	T	F	Poisoned reverse is used to ensure that packets flooded on a spanning tree do not loop infinitely.
2.	T	F	The fact that fragmentation reassembly is performed at end-hosts is a good example of the end-to-end-principle.
3.	T	F	Layering dictates that functionality should not be implemented in the network, even if doing so leads to a significant performance improvement.
4.	T	F	Fate sharing suggests that state should be stored on the entities that depend on that state.
5.	T	F	A reliable transport protocol must send packets in sequential order
6.	T	F	Distance Vector involves global computation and local state, whereas Link State Routing involves local computation and global state.
7.	T	F	If your browser opens up two TCP connections to the same server, the server can tell them apart because their arriving packets use different source transport ports.
8.	T	F	On a fast (100Gbps) cross-country link, propagation delay is usually the most significant component of the end-to-end delay.
9.	T	F	The size of an IPv4 packet header is always 20 bytes.
10.	T	F	The size of an IPv4 packet header is usually 20 bytes.
11.	T	F	Poisoned Reverse completely solves the Count-to-Infinity problem.
12.	T	F	Routers only implement the physical, data-link and network layers
13.	T	F	Hosts only implement the transport and application layers
14.	T	F	DNS typically uses UDP rather than TCP
15.	T	F	Everyone agrees on a single interpretation of the End-to-End Principle
16.	T	F	A car carrying a 1TB disk from Berkeley to Stanford in an hour carries 1TB faster than a 10Gbps link between Berkeley and Stanford.
17.	T	F	Each TCP packet has a TCP header that either includes the sequence number or the ACK number, but never includes both.
18.	T	F	The UDP packet header has a sequence number field but no ACK field.
19.	T	F	Packet switching uses bandwidth more efficiently than circuit switching
20.	T	F	Circuit switching requires routers to remember each connection

2. General Multiple-Choice [20 points]

Pick one and only one answer for each question (indicate by circling)

1. The appropriate layer for implementing reliable transfer is:
   a. The physical layer
   b. The data link layer
   c. The network layer
   d. **The transport layer**
   e. The application layer

2. A router running distance-vector routing knows:
   a. Nothing
   b. The entire network's topology
   c. **The distance to destinations through its immediate neighbors**

3. A router running link-state routing knows:
   a. Nothing
   b. **The entire network's topology**
   c. Only the distance to destinations through its immediate neighbors

4. Which of these fields does an IP router need to modify before sending a packet to the next hop?
   a. Source Address
   b. Destination Address
   c. Protocol
   d. **TTL**
   e. None of the above

5. Which of these fields in the IP header protect against corruption of the data payload?
   a. TTL
   b. Fragmentation flags
   c. Protocol
   d. Checksum
   e. **None of the above**

6. Which of these fields does the UDP packet header have?
   a. Source address
   b. Destination address
   c. **Checksum**
   d. Protocol
   e. None of the above

7. Which of these fields is present in IPv4, but not in IPv6?
    a. Source Address
    b. Destination Address
    c. **Fragmentation flags**
    d. Protocol/Next Header
    e. None of the above

8. Which of these fields is present in IPv6, but not in IPv4?
    a. **Flow Identifier**
    b. Destination Address
    c. Destination port
    d. Protocol/Next Header
    e. Checksum

9. Which one of the following apply to MAC addresses?
    a. May be assigned dynamically
    b. Used in the physical layer
    c. **Used in the data link layer**
    d. Used in the network layer
    e. Hierarchical

10. Which one of the following apply to IP addresses?
    a. **May be assigned dynamically**
    b. Burned into the device
    c. Used in the physical layer
    d. Used in the data link layer
    e. Always globally unique

11. Consider an IP packet (without options) with total length 1500 bytes. The packet is split into fragments by a network that can only handle IP packets of up to 500 bytes. Which set of IP packet lengths could describe the set of fragments?
    a. 500, 500, 500
    b. 400, 400, 400, 300
    c. **420, 420, 420, 300**
    d. 500, 500, 500, 80
    e. None of the above

12. Assuming circuit switching, if the network has allocated exactly enough bandwidth to handle the application’s peak bandwidth P, and the application has average bandwidth A, the level of utilization of that application’s circuit is given by (where / denotes division):
    a. **A/P**
    b. P/A
    c. (P - A)/A
    d. (P - A)/P

13. Once this test is over, you go to a take-out only In-N-Out, our course sponsor. While waiting you notice that on average a new customer goes in to the store every 5 minutes, and a customer receives their order and goes out on average 3 minutes after arriving. On average, how many customers are inside the In-N-Out:
    a. 2 customers
    b. 5 customers
    c. **0.6 customers**
    d. There is insufficient information provided
    e. The In-N-Out has customers?

14. Consider the case of a researcher at Berkeley trying to transfer a 6TB file to Stanford. Recall that 1TB = 1000GB, and GB means giga-*bytes*. If they send this over a 10Gbps link (recall that Gbps is gigabits per second), roughly how long will it take for the file to arrive (to the nearest minute)?
    a. 6 minutes
    b. 8 minutes
    c. 60 minutes
    d. **80 minutes**
    e. 90 minutes

15. If a laptop is connected to Ethernet and sends a DNS request, what is the sequence of packet headers on this request packet as it leaves the laptop (**starting from the outermost header**). (We are using the term header loosely when it comes to DNS.)
    a. **Ethernet, IP, UDP, DNS**
    b. DNS, UDP, IP, Ethernet
    c. IP, TCP
    d. Ethernet, TCP, IP, DNS
    e. None of the above

16. All hosts in the networks have been operational for several hours when the DHCP server goes down. What happens to the hosts that have obtained service from the DHCP server?
    a. The hosts will not be able to communicate with any other hosts.
    b. **The hosts will continue to communicate normally for a period of time.**
    c. The hosts will only be able to communicate with hosts outside their own network.
    d. The hosts will only be able to communicate with other hosts by IP address, not by hostname.

17. Someone invents a new compression technique that reduces data to a quarter of its original size. As chief Internet architect, you suggest deploying this technology in all:
    a. **End hosts**
    b. Routers
    c. Switches
    d. NAT boxes

18. If the round-trip time for packets going from Berkeley to MIT is roughly 100msec (i.e., 0.1sec), and the bandwidth available to a connection is roughly 1Mbps (remember this is 10^6 bits per second), then how large of a window (in bits) is necessary to "fill the pipe"?
    a. 100 bits (10^2)
    b. 1000 bits (10^3)
    c. 10000 bits (10^4)
    d. **100000 bits (10^5)**
    e. 1000000 bits (10^6)

19. Using the same parameters as above, what is the end-to-end throughput that can be achieved if the source uses a window size that is double the size needed to fill the pipe?
    a. .1 Mbps
    b. .5 Mbps
    c. **1 Mbps**
    d. 2 Mbps
    e. 10 Mbps

20. Again using the parameters above, what is the end-to-end throughput that can be achieved if the source uses a window size that is half the size needed to fill the pipe?
    a. .1 Mbps
    b. **.5 Mbps**
    c. 1 Mbps
    d. 2 Mbps
    e. 10 Mbps

3. TCP Basics [10 points]

Pick one and only one answer (indicate by circling)

A TCP connection has been established between hosts A and B. A receives a packet from B that has a 52 byte TCP payload and the following field values shown below:

Sequence: 1001
Acknowledgement: 5001
Window size: 4000

Assume that B has not sent any acknowledgments beyond 5001.

For each of the packets below, indicate whether they are possible valid **responses** from A:

1) Sequence: 5001
Acknowledgement: 1053 — **Valid** — Invalid
Window size: 200
Payload: 2000

2) Sequence: 1053
Acknowledgement: 5001 — Valid — **Invalid**
Window size: 3000
Payload: 2000

3) Sequence: 8001
Acknowledgement: 1053 — **Valid** — Invalid
Window size: 200
Payload: 1000

4) Sequence: 6001
Acknowledgement: 1053 — Valid — **Invalid**
Window size: 200
Payload: 4000

5) Sequence: 6231
Acknowledgement: 1053 — **Valid** — Invalid
Window size: 2000
Payload: 1000

Validity test:

- $\text{Sequence}_{\text{response}}$ >= 5001
- ($\text{Sequence}_{\text{response}}$ + $\text{Payload}_{\text{response}}$) <= ($\text{Acknowledgement}_{\text{first}}$ + $\text{Window size}_{\text{first}}$) = (5001 + 4000) = 9001

4. Sequence of Messages [15 points]

Consider a laptop trying to access a file via HTTP from a server. The laptop shows up on the network without an IP address, and without knowing the IP address of the server. Below please list the following messages that appear on Laptop's network (in the order that they would occur), by listing the letter associated with the message. All messages are used once except (i). Assume that the HTTP request and response are each single packets.

a. DHCP: DHCP ACK
b. DHCP: DHCP Offer
c. DNS: DNS Response
d. Laptop: DHCP Discovery
e. Laptop: DHCP Request (aka "DHCP Accept")
f. Laptop: DNS Request
g. Laptop: HTTP Request to Server
h. Laptop: TCP SYN to Server
i. Laptop: TCP ACK to Server **(used three times!)**
j. Laptop: TCP FIN to Server
k. Server: HTTP Response
l. Server: TCP SYNACK to Laptop
m. Server: TCP FINACK to Laptop

Messages in order:

1. ______**d**
2. ______**b**
3. ______**e**
4. ______**a**
5. ______**f**
6. ______**c**
7. ______**h**
8. ______**l**
9. ______**i**
10. ______**g**
11. ______**k**
12. ______**i**
13. ______**j**
14. ______**m**
15. ______**i**

5. The Real World (according to EE122) [6 points]

Pick one and only one answer for each question (indicate by circling)

1. Your grandfather calls you because, in his words, "the Internet is broken." You ask him for details, and he explains that he cannot access any web sites, including google.com and npr.org. Which of the following is most likely the cause of the problem?
   a. **Your grandfather's DHCP server is no longer working, so he hasn't been assigned an IP address.**
   b. Your grandfather's DHCP server is no longer working, so he hasn't been assigned an Ethernet address.
   c. Your grandfather is behind a NAT box, which means that he cannot access anything outside of his local network.
   d. The "com" top level domain server is down.
   e. Google is down.

2. The next day your grandfather calls you back to thank you for fixing his computer, but he now says that he can reach all sites except your home page www.yourfavoritegrandchild.com. You know the site is up (because it is served off of your home PC, and you know your PC is connected to the Internet). Which of the following is the most likely cause of this new problem?
   a. Your grandfather's DHCP server is no longer working, so he hasn't been assigned an IP address.
   b. Your grandfather's DHCP server is no longer working, so he hasn't been assigned an Ethernet address.
   c. Your grandfather is behind a NAT box, which means that he cannot access anything outside of his local network.
   d. The "com" top level domain server is down.
   e. **The authoritative DNS server for yourfavoritegrandchild.com is down.**

3. The following day your grandfather calls you back to thank you for fixing his computer again, but he now says that he can't reach google.com or bing.com, but he can reach cs.berkeley.edu and npr.org. Which of the following is the most likely cause of this latest problem?
   a. Your grandfather's DHCP server is no longer working, so he hasn't been assigned an IP address.
   b. Your grandfather's DHCP server is no longer working, so he hasn't been assigned an Ethernet address.
   c. Your grandfather is behind a NAT box, which means that he cannot access anything outside of his local network.
   d. **The "com" top level domain server is down.**
   e. All authoritative DNS servers are down.

6. Timer Values [8 points]

Pick one and only one answer for each question (indicate by circling)

1. An offer message from a DHCP server includes a lease time (along with other fields like an IP address, subnet mask, DNS server address, and so on). Why is a lease time necessary?
   a. The lease field is necessary for scalability: without a lease time, a DHCP server would store one entry for every host that ever used the network, and eventually this state would become infinite.
   b. The lease time helps network operators bill hosts correctly for the amount of time that they were using the network.
   c. The lease time allows the DHCP server to reclaim IP addresses, even if the client never explicitly releases the address (e.g., if the client crashes).
   d. The lease time ensures that addresses assigned to hosts remain private, and aren't sent outside of the private network where they are being used.

2. The IP packet header includes a time-to-live field that is decremented by each router along the path. Why is the time-to-live field necessary?
   a. The TTL field is decremented at each router along the path, and routers drop packets with a TTL of 0, so the TTL field prevents packets from looping indefinitely if they are stuck in a forwarding loop.
   b. The TTL field allows hosts to determine the appropriate window size.
   c. The TTL field is necessary for scalability. Routers store each packet for the amount of time given by the TTL field; without the TTL field, routers would need to store packets forever, which would consume infinite space.
   d. The TTL field is not necessary, which is why it was removed in IPv6.

3. NAT has a timer associated with each map entry. Why is this necessary?
   a. The timer field is included in each packet and decremented at each router along the path. Routers drop packets with a timer value of 0, so the timer field prevents packets sent by the NAT box from looping indefinitely.
   b. NAT stands for network address timer; the timer is fundamental to a NAT's functionality.
   c. Eventually, the NAT box needs to reclaim port numbers to use for new data transfers. Without a timer, the NAT box would eventually use up all ports and not be able to allow new connections.
   d. When a NAT box assigns an IP address to a client and the client never responds, the timer allows the NAT box to eventually reclaim the address.

4. DNS responses have a TTL field. Why is this necessary?
    a. The TTL field is decremented at each DNS server that the response passes through on its way to the client, and servers drop responses with a TTL of 0, so the TTL field prevents responses from looping indefinitely.
    b. The TTL field allows DNS servers to prevent cache poisoning.
    c. The TTL field is necessary for scalability: if DNS servers could never time out entries, over time they would accumulate infinite state.
    d. **The TTL field causes DNS servers to delete entries after some time, so that if the host moves and the underlying address changes, the server will eventually get the correct address.**

7. Addressing [14 points]

Pick one and only one answer for each question (indicate by circling)

1. What is the network portion of the address 4.5.6.7 with mask 255.192.0.0?
    a. 0000001100
    b. 00000100000
    c. 0000010011
    d. 00010100101
    e. **0000010000**

```
                4    .    5    .    6    .    7
Mask        11111111.11000000.00000000.00000000
Address     00000100.00unimportant bits omitted
```

2. What is the host portion of the address 4.5.6.7 with mask 255.255.255.0?
    a. 1100110010
    b. **00000111**
    c. 00001010
    d. 111000000
    e. 00001000

```
                4    .    5    .    6    .    7
Mask        11111111.11111111.11111111.00000000
Address     unimportant  bits  omitted 00000111
```

3. What is the address range covered by the prefix 128.8.16.0/20?
    a. 128.8.16.0 - 128.8.16.255
    b. 128.8.16.0 - 128.8.20.255
    c. 128.8.16.0 - 128.8.24.255
    d. **128.8.16.0 - 128.8.31.255**

```
              128   .    8    .   16    .    0
Address     10000000.00001000.00010000.00000000
Mask        11111111.11111111.11110000.00000000
```

Conveniently, this means the highest possible address is

```
Address'    10000000.00001000.00011111.11111111
```

which is 128.8.31.255.

For the next four questions, consider the following routing table:

128.8.16.0/20	Port 1	10000000.00001000.00010000.00000000	128.8.16.0	128.8.31.255
128.8.24.0/21	Port 2	10000000.00001000.00011000.00000000	128.8.24.0	128.8.31.255
128.8.128.0/24	Port 3	10000000.00001000.10000000.00000000	128.8.128.0	128.8.128.255
128.8.128.0/28	Port 4	10000000.00001000.10000000.00000000	128.8.128.0	128.8.128.15
Default	Port 5			

For each of the following addresses, indicate which port the packet is sent out:

4. 128.8.128.252

Port 1 Port 2 **Port 3** Port 4 Port 5

This is within the range corresponding to Port 3, but none other.

5. 128.8.128.5

Port 1 Port 2 Port 3 **Port 4** Port 5

This is within the range of Ports 3 and 4, but the range corresponding to Port 4 is more specific (/28).

6. 128.8.25.223

Port 1 **Port 2** Port 3 Port 4 Port 5

This is within the range corresponding to Port 2.

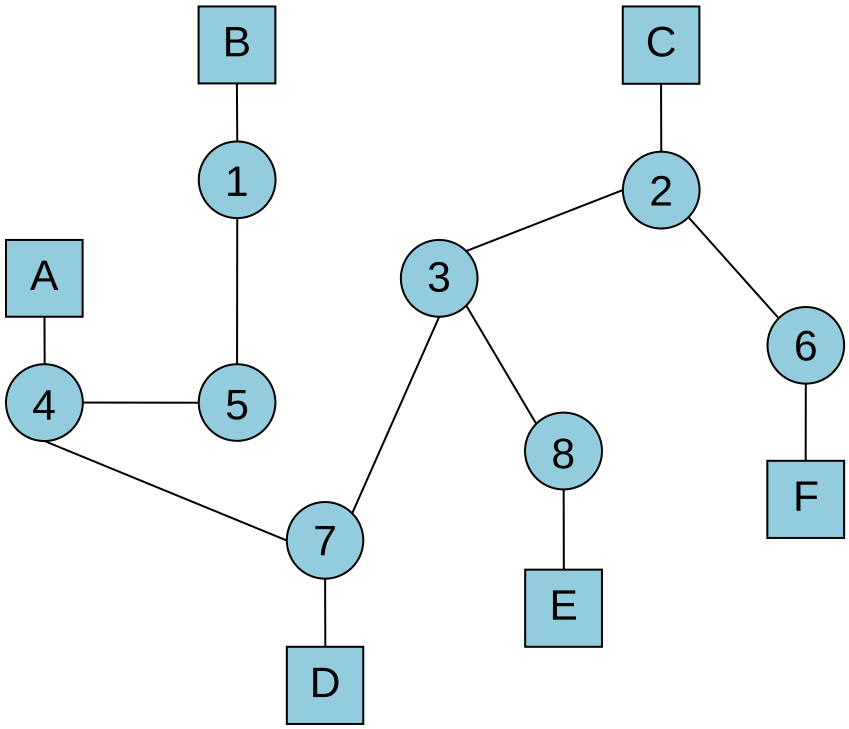
7. 128.8.128.10

Port 1 Port 2 Port 3 **Port 4** Port 5

Same as question 5.

8. Learning Switch [12 points]

In the following graph, the switches are represented by circles and the hosts are represented by squares. The switches in the following network initially do not know how to reach any other node.

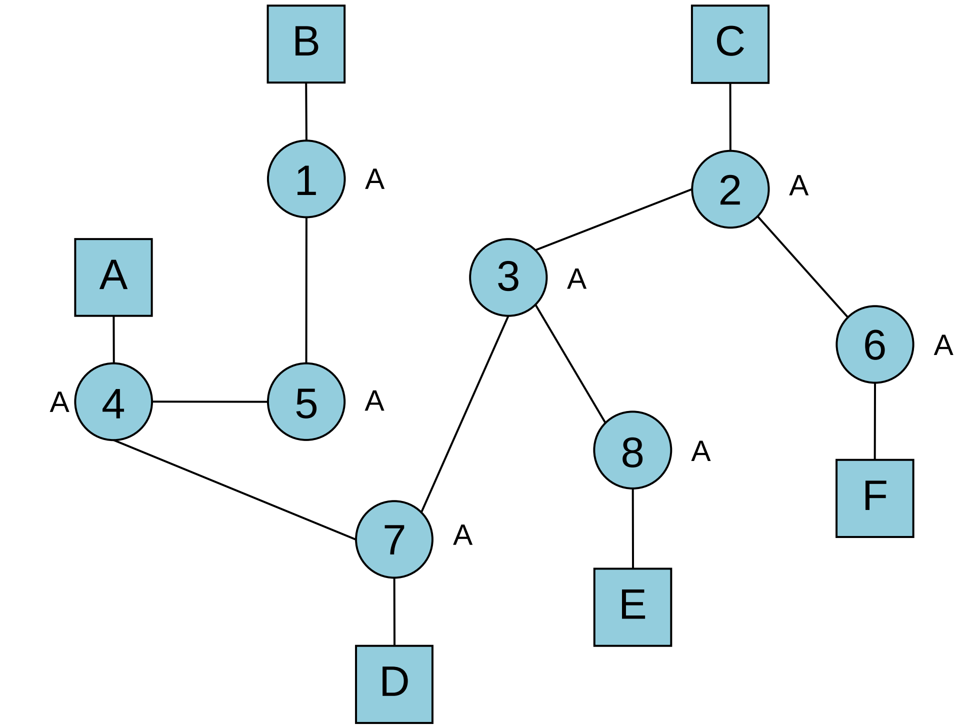


Which switches does the packet visit (indicate by circling) if the following packet exchanges occur? For each question, assume that the previous packet exchanges have occurred.

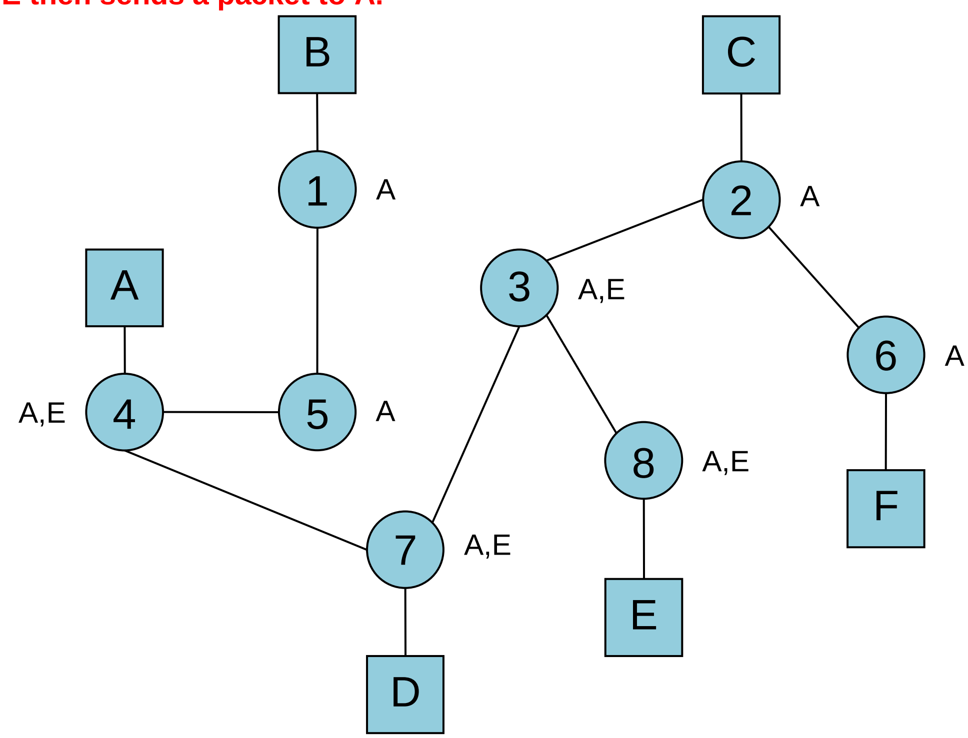
1. A sends a packet to F. **1 2 3 4 5 6 7 8**
2. E then sends a packet to A. 1 2 **3 4** 5 6 **7 8**
3. C then sends a packet to E. 1 **2 3** 4 5 **6** 7 **8**
4. D then sends a packet to C. **1 2 3 4 5** 6 **7** 8

Assume no packet exchanges other than the ones specified in the questions.

After A sends a packet to F, the switches know the links for the following destinations[1]:

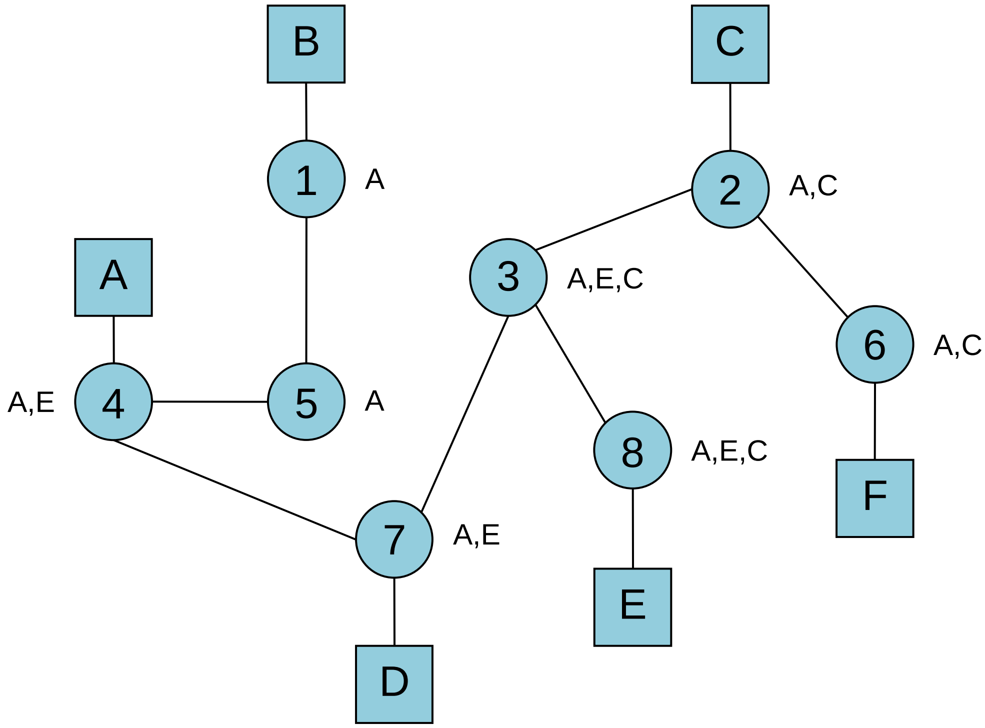


After E then sends a packet to A:



[1] We cheat by listing the known hosts next to the switch, but not the specific link.

After C then sends a packet to E:



9. DNS [4 points]

For the questions below, pick one and only one answer for each step

Consider a scenario with a Host (call it A) in domain berkeley.edu, its local DNS server (call it B), and the rest of the DNS infrastructure. Thus, there are the following entities that might be involved in this scenario (and some that aren't, which were accidentally included in the list):

a. Host A
b. Local DNS server B
c. Root
d. TLD server for .com
e. TLD server for .edu
f. TLD server for .org
g. Authoritative server for foo.com
h. Authoritative server for berkeley.edu

A asks B to resolve hostname mine.foo.com. Assume there are no cached entries relevant to this request. Write down the steps taken to resolve foo.com *and* respond to A. In each step identify who B contacts. Identify who is being contacted by one of the letters above (a-h):

Step 1: B contacts _____**c** **Root**

Step 2: B contacts _____**d** **TLD server for .com**

Step 3: B contacts _____**g** **Authoritative server for foo.com**

Step 4: B contacts _____**a** **Host A**

10. Sliding Window [10 points]

Pick one and only one answer for each question (indicate by circling)

Consider a sender and receiver exchanging data using a sliding window flow control. We use the notation Di to refer to the i'th data packet, and Ai to refer to the ACK packet acknowledging receipt of all data packets up to and including packet (i-1). Assume that the transport protocol has a fixed window size of 5, and will time out packets after 1 second and retransmit them (no fast retransmits). Assume further that the packet transmission time and RTT time are much smaller than 1 second.

We will give you a series of ACK packets received by the sender, and you need to tell us which of the several scenarios could be responsible for this series of observed ACK packets.

Scenarios:

a. D3 was dropped by the network
b. The ACKs sent in response to packets D1 and D3 were dropped by the network
c. A3 was duplicated by the network
d. D3 was duplicated by the network
e. D2 and the ACK sent in response to D4 was dropped by the network

Each of the following independent questions shows the first five ACKs received by the sender. For each one, pick which one of the scenarios was most likely to produce this result:

1. A2 A3 A3 A4 A5 pick one of: a b **c** d e

2. A2 A2 A2 A2 A7 pick one of: a b c d **e**

3. A2 A3 A3 A3 A3 pick one of: **a** b c d e

4. A2 A3 A4 A4 A5 pick one of: a b c **d** e

5. A3 A5 A6 A7 A8 pick one of: a **b** c d e

N.B. The following table shows the answers are sufficient. It does not prove they are the only possible correct cases.

X0 = duplicated packet
~~X0~~ = packet dropped

	Sender generates	Receiver gets	Receiver generates	Sender gets	Scenario
1	D1	D1	A2	A2	A3 was duplicated by the network
	D2	D2	A3	A3	
				A3	
	D3	D3	A4	A4	
	D4	D4	A5	A5	
2	D1	D1	A2	A2	D2 and the ACK sent in response to D4 was dropped by the network
	D2	~~D2~~			
	D3	D3	A2	A2	
	D4	D4	A2	~~A2~~	
	D5	D5	A2	A2	
	D6	D6	A2	A2	
	D2	D2	A7	A7	
3	D1	D1	A2	A2	D3 was dropped by the network
	D2	D2	A3	A3	
	D3	~~D3~~			
	D4	D4	A3	A3	
	D5	D5	A3	A3	
	D6	D6	A3	A3	
4	D1	D1	A2	A2	D3 was duplicated by the network
	D2	D2	A3	A3	
	D3	D3	A4	A4	
		D3	A4	A4	
	D4	D4	A5	A5	
5	D1	D1	A2	~~A2~~	The ACKs sent in response to packets D1 and D3 were dropped by the network
	D2	D2	A3	A3	
	D3	D3	A4	~~A4~~	
	D4	D4	A5	A5	
	D5	D5	A6	A6	
	D6	D6	A7	A7	
	D7	D7	A8	A8	

Alternative reasoning:

- Question 5 is the only one where A2 was dropped, hence this must be scenario b.

- Only question 1 has a single duplication of A3, hence this must be scenario c.
  (We should also check that no other scenario matches question 1, because the question never claimed that scenarios/questions were a bijection.)

- Question 2 implies that, by the time packet 2 was received, they receiver had already got packet 6. This means packet 2 was either delayed or dropped; the only scenario that matches this is e.

- Question 4 could be:
  - receiver gets D1, D2, D3, D(!= 4), D4
    or
  - duplication of A4

  The only matching scenario is d.

- We can check that scenario a, but no other, matches 3.

Bonus Questions [0.00001 points]

What is the port number for the Quote-of-the-Day service? **17**

Who is the unsung hero of the Internet? **Kay Ousterhout/David Clark**

What letter caused the first demonstration of Internet technology to fail? **G**

In what year did Larry Roberts connect two computers by phone? **1965**