Department of Electrical Engineering and Computer Science
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

6.033 Computer Systems Engineering: Spring 2019

Quiz I

There are 11 questions and 14 pages in this quiz booklet. Most questions have multiple parts. Answer each question according to the instructions given. You have two hours to answer the questions.

- The questions are organized (roughly) by topic. They are not ordered by difficulty nor by the number of points they are worth.
- **If you find a question ambiguous, write down any assumptions you make.** Be neat and legible.
- Some students will be taking a make-up exam at a later date. Do not discuss this quiz with anyone who has not already taken it.
- You may use the back of the pages for scratch work, but do not write anything on the back that you want graded. We will not look at the backs of the pages.
- Write your name in the space below. Write your initials at the bottom of each page.

This is an open-book, open-notes, open-laptop quiz, but you may NOT use your laptop, or any other device, for communication with any other entity (person or machine).

Turn all network devices, including your phone, off.

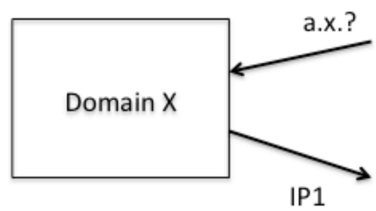
Name:

Email: **@mit.edu**

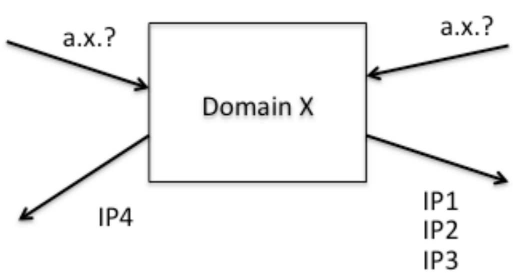
Please use only your MIT email ID.

DNS

1. [10 points] The following diagram shows an abstract view of a simple DNS query. A server that is authoritative for the domain x is asked about the name "a.x." and returns the IP address of that machine:



Now consider the following situation, in which two different queries are asking about the same name at the same time but receiving different replies. All of the IP addresses (IP1, IP2, IP3, IP4) returned do in fact correspond to different machines that are all prepared to deliver the same service(s) associated with the name a.x.



A. What is the most likely explanation of this phenomenon? Circle the BEST answer:

a) There is a race condition in the name server

b) The different requests are coming from different RON nodes on the same local area network

c) The name server is operated by a CDN

d) The name server's cache has timed out

e) Hackers have taken control of the name server and corrupted it

Initials:

A user is having a puzzling problem with one of their applications, and your troubleshooting has narrowed it down: for some reason this name server resolved a.x. to IP5, which is not a server that supports the service(s) expected.

You are equipped to watch DNS traffic going to and from this server. Joe suggests having the user's application retry the action that caused resolution of a.x. and watching the traffic. Assume you follow Joe's advice.

B. What, if anything, are you sure to learn? Circle the BEST answer:

a) How the name server currently resolves the name a.x. for that user
b) Nothing, because of potential caching between the user and the name server.
c) Whether the name server is still offering IP5 as one of the possible results
d) Nothing, because the resolution would start at the DNS root.

Unix

2. [10 points] As a result of a fork, there are two processes running on a machine: the parent and the child A. Immediately after returning from the fork() call, the parent forks again, creating child B. Neither child process has been scheduled yet (i.e., they have not yet had an opportunity to execute anything after the return from fork()). We are asking about an instant when the two children are fully created and completely ready to run, but before either has had a chance to run.

Circle True or False for the each of the following statements:

a) **True / False** If virtual address *a* maps to physical address *p* in process A, then virtual address *a* maps to physical address *p* in process B
b) **True / False** Processes A and B have identical file descriptors.
c) **True / False** Child A knows the pid of Child B.
d) **True / False** Child B knows the pid of Child A.

Initials:

Threads and Eraser

3. [12 points] Consider the following function designed to transfer the contents of bank account ***a*** into account ***b***. Each account has a text *name*, a numeric *amount* and a *lock*. Assume that some other part of the system has ensured that all variables are initialized and have sensible values, so you don't need to be concerned with handling those kinds of errors. You also don't need to be concerned about any case where ***a*** and ***b*** are the exact same account.

```
transfer1(account a, account b) {
  if (a.name < b.name)
    lock = a.lock
  else
    lock = b.lock

  acquire(lock)
  b.amount = b.amount + a.amount
  a.amount = 0
  release(lock)
}
```

The `transfer1` code is run in a multi-threaded environment with Eraser, with sufficient test cases (possibly with different arguments) and scheduling control to ensure that all "interesting" paths and combinations are covered.

A. Circle True or False for each of the following statements:

a) **True / False** Deadlock can occur.

b) **True / False** A data race can occur.

c) **True / False** Eraser will raise a warning.

Initials:

Next, consider this alternative function for the same purpose:

```
transfer2(account a, account b) {
  acquire(a.lock)
  acquire(b.lock)
  b.amount = b.amount + a.amount
  a.amount = 0
  release(a.lock)
  release(b.lock)
}
```

Now the `transfer2` code is run in a multi-threaded environment with Eraser, again using sufficient test cases (possibly with different arguments) and scheduling control to ensure that all "interesting" paths and combinations are covered.

B. Circle True or False for each of the following statements:

a) **True / False** Deadlock can occur.

b) **True / False** A data race can occur.

c) **True / False** Eraser will raise a warning.

Initials:

Virtual Memory

4. [15 points] Katrina's operating system (kOS) uses a two-level page-based virtual addressing system. Each virtual address is 32 bits. All page table entries are 4 bytes. A single page is 4096 bytes. Assume the size of each first-level table is the same as the size of each second-level table.

   Each virtual address consists of m bits for page level 1, n bits for page level 2, and p bits of offset within a page. All bits of the address are used ($m + n + p = 32$).

   Write a SINGLE CLEAR answer for each of the following questions about kOS:

   A. What is the value of m?

   B. What is the value of n?

   C. What is the value of p?

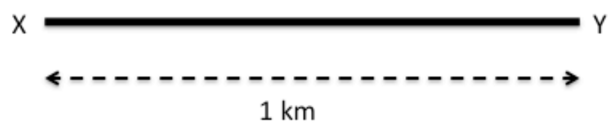
   D. Using this system, what is the maximum quantity of memory that could possibly be occupied by page tables? Give your answer as an expression using m, n, and p.

   Now assume that initially no page tables are allocated. Then, only a single virtual address translation takes place.

   E. After that single virtual address translation, what is the quantity of memory occupied by page tables? Give your answer as an expression using m, n, and p.

Initials:

Ethernet



5. [8 points] In the original 3 Mb/s Ethernet that we read about, the maximum length of an Ether segment is 1 km. Assume that the time for a signal to go from one end (X) to the opposite end (Y) is at most 8 microseconds. The round-trip time (RTT) is 16 microseconds.

A. Why is the RTT of the segment important for determining whether collisions occur? Circle the BEST answer:

a) Every packet fills up an RTT-length slot

b) Two stations at opposite ends might both start transmitting at roughly the same time

c) The speed of light is the same in coax as in a vacuum

d) Both A and B

e) Both B and C

f) None of the above

B. Circle True or False for each of the following statements:

a) **True / False** Ethernet receivers send acknowledgement (ACK) packets upon receipt of a packet that passes a CRC integrity check.

b) **True / False** Ethernet receivers discard duplicate packets.

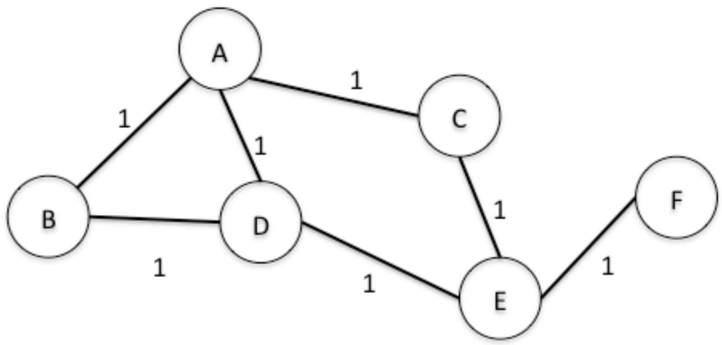
c) **True / False** Each Ethernet transceiver makes local decisions about packet sending and receiving, to avoid a single central point of failure.

d) **True / False** In the worst case, a single failed Ethernet transceiver can cause the entire segment to be unusable.

Initials:

Routing

6. [10 points] Consider the following network topology and link costs:



First assume that the network is using Link-State Routing.

A. How many total advertisements will be sent? Circle the BEST answer:
   a) 14
   b) 21
   c) 30
   d) 36
   e) 84

Now assume that the network is using Distance-Vector Routing.
Node A sends advertisements at timesteps 1, 11, 21, ..., 10N+1.
Node B sends advertisements at timesteps 2, 12, 22, ... , 10N+2.
Likewise for the remaining nodes:
Node C sends at 3, 13…10N+3; Node D at 4, 14…10N+4; Node E at 5, 15…10N+5; and Node F at 6, 16…10N+6.

Assume that a single timestep is sufficient time for an advertisement to arrive at its destination and be processed, so that in one time step A's advertisements have reached B, C, and D. Further assume that no advertisements are dropped, and that no "sophisticated" mechanisms like Split Horizon are in use.

After a point when the network has stabilized, E learns that the EF link has failed at timestep 26.

Initials:

B. At timestep 46, what does E believe is the distance of its minimum path to F? Write your SINGLE answer CLEARLY in the box below:

C. Circle True or False for each of the following statements about routing protocols:

a) **True / False** In Distance-Vector routing, each node sends advertisements to all nodes in the network through a process known as flooding.

b) **True / False** Link-State Routing is appropriate for a network the size of MIT's network.

c) **True / False** In Distance-Vector Routing, INFINITY is set to be a very large integer, typically the max integer value (~ 2^32).

d) **True / False** In both Distance-Vector Routing AND Link-State Routing, nodes learn about their neighbors via the HELLO protocol.

RON

7. [3 points] Ben is excited by the possibility of using a RON but didn't read the paper very carefully. Responding to his statement below, circle the ONE BEST correction to help him learn:

"When there's a RON node at both MIT and Stanford, all the traffic between the two schools will become more resilient against network failures."

a) Only applications that are configured to use RON can benefit

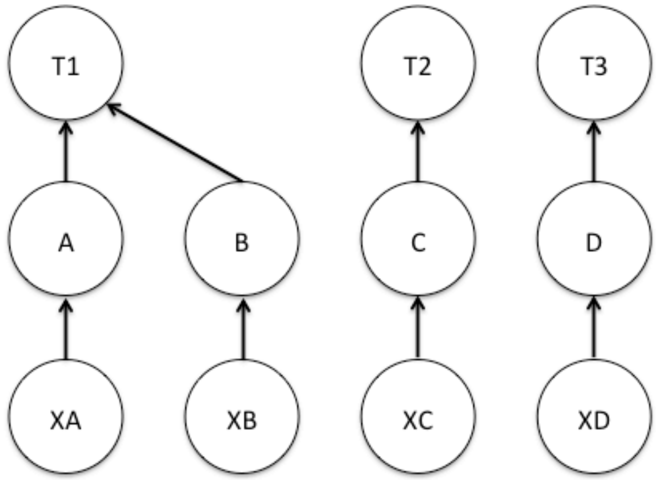
b) Only applications that can access Internet-2 can benefit

c) Only applications that understand CIDR address ranges can benefit

d) No traffic benefits unless all the network devices on the path between MIT and Stanford include RON functionality

BGP

8. [10 points] Consider the following AS graph. (A → T1 means that A purchases transit from T1; equivalently, T1 provides transit for A)



XA wants to send traffic to XC.

A. Circle ALL of the relationships that, if added to the graph, would make this possible:
   a) T1 and T2 form a peering relationship
   b) B and C form a peering relationship
   c) A and D form a peering relationship; D and C form a peering relationship
   d) A and C form a peering relationship
   e) None of the above

Now assume that B and C have formed a peering relationship, in addition to the relationships in the graph above. XA still wants to send traffic to XC.

B. What is the minimum number of RON nodes that, if added to the graph, would make this possible? Circle the BEST answer:
   a) No RON nodes needed; XA can already send traffic to XC
   b) 2 RON nodes, at XA and XC
   c) 2 RON nodes, at XA and B
   d) 3 RON nodes, at XA, B, and XC
   e) 4 RON nodes, at XA, B, C, and XC

Initials:

DCTCP and MapReduce

9. [12 points] We have a large-scale web application using the partition/aggregate design pattern with many workers and real-time requirements. Assume there is NO background traffic (no other applications on the network) and the responses from workers are all small (1-2 packets each).

   A. Which of these problems is likely to be a concern in our network? Circle ALL correct answers:

      a) Incast

      b) Queue buildup

      c) Buffer pressure

      d) None of these problems is likely

   Max has a bright idea of changing the aggregator structure for the partition/aggregate pattern. Instead of having one aggregator to aggregate the answer of 1000 workers, Max proposes using 10 aggregators each aggregating the answer from 100 workers, and then 1 aggregator to aggregate the answer from the 10 aggregators.

   B. For our application, what is the most significant potential problem with this approach?

      Circle the BEST answer:

      a) It increases the number of machines required

      b) It roughly doubles the time required for aggregation

      c) It creates longer queues

      d) It intensifies incast

   C. The MapReduce paper doesn't mention incast as an issue. Why not?

      Circle the BEST answer:

      a) Actually it does, but MapReduce talks about "stragglers" instead of "incast."

      b) The MapReduce paper doesn't describe many details of the reshuffling that happens between stages.

      c) Incast problems only started after the MapReduce paper had already been published.

      d) MapReduce operates in rigidly-scheduled stages, so it's impossible to have multiple messages arrive at a single switch at the same time.

      e) Google used exotic custom switches, described in the paper, to avoid these kinds of networking problems.

Initials:

D. TCP with RED/ECN has some similarities to DCTCP. What distinguishes DCTCP from TCP with RED/ECN?

Circle ALL correct statements on how DCTCP is DIFFERENT:

a) DCTCP implements a central service to actively optimize the queue length of each switch

b) DCTCP can be implemented in commercially-available switches

c) DCTCP automatically determines the queue-length threshold K

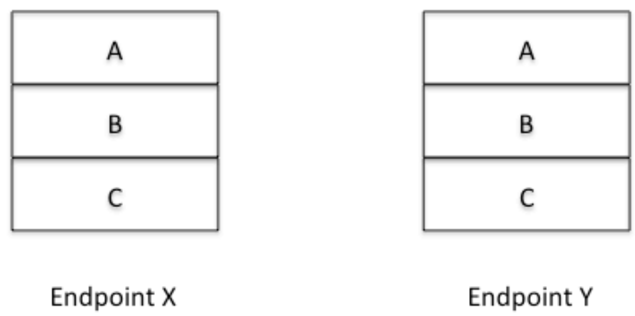
d) DCTCP does not adjust the size of the congestion window.

e) DCTCP tries to minimize queue occupancy, but at the cost of lower throughput than TCP with RED/ECN.

f) None of the above statements distinguishes DCTCP from TCP with RED/ECN.

Initials:

End-to-End



10. [9 points] The layer diagram above shows two communicating endpoints on a local-area network (LAN). Protocol A is a TCP-like reliable transport protocol. Protocol B is an IP-like network protocol. Protocol C is an Ethernet-like protocol for hardwired LANs. Assume that programs at endpoints X and Y are sending data to each other using protocol A.

   A. By some technical means, we eavesdrop on the LAN traffic. We don't alter or interfere with the traffic in any way; but we can see every packet that's transmitted, in exactly the form it takes on the LAN. If we consider the very first header of any such packet, which protocol defines its format and meaning?

   Circle the BEST answer:

   a) A
   b) B
   c) C
   d) None of the above

   B. Assume that we replace protocol C with protocol C', where there are more frequent packet losses. We make no changes in the communicating applications. What change do we need to make in protocol A?

   Circle the BEST answer:

   a) Cumulative ACKs
   b) Ordered byte sequence
   c) Additive increase
   d) Multiplicative decrease
   e) All of the above
   f) No change needed

Initials:

C. Now we replace protocol B with protocol B', where B' guarantees delivery of every packet. Again, we make no changes in the communicating applications. Which protocol layer(s) can we now remove?

Circle the BEST answer:

a) A

b) B'

c) C

d) None, because protocol A is the interface used by the programs

e) None, because protocol A may be providing other services like congestion control, in addition to reliable delivery

f) Both d and e

Cheap Point

11. [1 point] Consider this to be our market research: We want to know which papers you particularly liked or disliked.

Please pick papers that we read this term, and describe them in a way that we can recognize.

Feel free to comment further about why you like or don't like a paper, but first be sure you've completed all the previous questions.

A. What was your favorite paper?

B. What was your least favorite paper?

End of Quiz 1

Please double-check that you have written your name and email ID on the first page and initialed all of the pages

Initials: