CS244
Advanced Topics in Computer Networks
Midterm Exam – Tuesday, May 19

OPEN BOOK, OPEN NOTES

Your Name: ______________________________

SUNet ID: ________________@stanford.edu

In accordance with both the letter and the spirit of the Stanford Honor Code, I neither received nor provided any assistance on this exam

Signature: ______________________________

1	/25
2	/25
3	/25
4	/25
Total	/100

- The exam has 4 questions totaling 100 points. You have 80 minutes to complete them.
- Some questions may be harder than others.

- All questions require you to justify your answer to receive full credit, even multiple choice questions for which you circle the correct answer(s).
- Keep your answers concise. We will deduct points for a correct answer that also includes incorrect or irrelevant information.

Assumptions (write any other assumptions you make in the space below)

1. Assume the speed of propagation along a cable or optical fiber is 2×10^8m/s.
2. $\left(\frac{a-1}{a}\right)^a \to 1/e \approx 37\%$, as integer $a \to \infty$.

Question 1: [25 points] Buffer Sizing

a) Explain why a router needs a packet buffer.

A network with a single router carries a single TCP Reno flow between two end hosts. The source host is connected to the router by an Ethernet link running at 1Gb/s. The destination host is connected via a link at rate 100Mb/s. Both links are 10km long. The source host is capable of sending at full line rate and delivers a long file, as fast as it is allowed to by TCP, to the destination. All packets are 1,000 bits long.

b) What is the minimum round-trip time for a 1,000bit packet to be sent from the source to the destination and back again? You should include both the propagation and packetization (*aka* serialization) delays in your answer.

c) How large does the packet buffer need to be in the router in order to maintain full 100% sustained utilization of the bottleneck link?

d) If the router in fact contains 10 times as much buffering as your answer to part (d), will the sustained utilization still be 100%? Explain your answer.

e) List three negative consequences of using a packet buffer 10x larger than needed.

f) We learned in class that if AIMD reduces the window size by a factor k when it detects a packet drop (i.e. $W \rightarrow \frac{W}{k}$), then the network can sustain 100% throughput for a single flow if $B \geq aC$, where $k = 1 + \frac{a}{2T}$ and $2T$ is the minimum RTT. Prove that this is true.

g) What are the consequences of the result in (f) on the necessary buffer size for 100% throughput?

Question 2: [25 points] Switching

a) Explain why an output queued (OQ) switch minimizes the average delay of packets passing through it.

b) Explain why packets arriving to an input queued (IQ) switch, with a single FIFO queue at each input, generally have a larger delay through the switch than for an OQ switch.

c) Consider an IQ packet switch with 100 inputs and 100 outputs all operating at rate 1Mb/s. The switch has a single FIFO queue at each input to hold packets. If the input links are continuously busy with fixed-length packets, and if an arriving packet throws a fair 100-sided die to decide which output to go to, write down the approximate throughput of the switch?

d) In class, we learned about Parallel Iterative Matching (PIM), an algorithm used to schedule “cells” (fixed length packets) in an input-queued switch with virtual output queues (VOQs). Consider a 2×2 switch (i.e. a switch with 2 inputs and outputs) running the PIM algorithm **with only one iteration** to schedule packets across a crossbar. If all the VOQs are full all the time:

   i. Explain why the size (or cardinality) of the bipartite match is always either 1 or 2.

   ii. Prove that the average match size is 1.5.

iii. [**Hard**] Consider now an $N \times N$ switch under the same conditions (PIM with one iteration; and the VOQs are always non-empty). Write down an expression for the average throughput of the switch; *i.e.* the average size of the match each time the PIM algorithm is run, as a function of N and line-rate R.

iv. Prove that the throughput converges to approximately $0.63 \times R$ for large N.

e) [**Hard**] It is known that if an IQ switch with VOQs uses a maximum weight matching (MWM) algorithm (where the weight might be the current occupancy of the VOQ), then it leads to higher throughput than if it uses a maximum size (or cardinality) matching instead. But a maximum size match maximizes (by definition) the instantaneous number of cells transferred. Explain why the MWM might lead to higher throughput in the long-run.

Question 3: [25 points] Internet Architecture for Mobility

a) Imagine an Internet connection between host A and server B using the standard Ethernet/IP/TCP stack, where A is being served by a router R1. What is the primary purpose each layer of Ethernet/IP/TCP is serving?

b) Assume that there is an ongoing TCP flow between host A and server B. If A were to move to new network served by router R2, what part of the stack fails in its operation and what is the main reason for it?

c) Between layering and end-to-end principles, which principle was violated in the design choice responsible for lack of connection-oriented mobility and why?

d) Do you think the particular design choice you mentioned in part b is essential for the primary purposes of each layer to be served correctly? Explain briefly.

e) Suggest a simple solution that fixes the mobility problem and still preserves the layering and end-to-end principles. In keeping with the end-to-end principle, you are allowed full freedom in the design of the end points but not in the network components (routers).

Question 4. An End to the End-to-End? [25 points]:

For this question, we will consider the End-to-End Argument presented by Saltzer, Reed, and Clark in the 1981 paper “End-to-End Arguments in System Design”:

> *The function in question can completely and correctly be implemented only with the knowledge and help of the application standing at the end points of the communication system. Therefore, providing that questioned function as a feature of the communication system itself is not possible. (Sometimes an*

incomplete version of the function provided by the communication system may be useful as a performance enhancement.)

Below are the systems from a few papers we have read so far this quarter. For each system, please answer the following and justify:

- Is the system a departure from the end-to-end argument? In what ways?
- If the system departs from the end-to-end argument: What benefits does it receive by doing so? Are there potential problems or weaknesses?
- If the system does not depart from the end-to-end argument: What benefits does the system receive by adhering to the end-to-end argument, and can you outline an alternative architecture that would be more efficient (but does not adhere as faithfully to the end-to-end argument)?

Please be clear about what the "communication system" is in your answer, and **please make specific references** to the papers or to the design of the system in question.

b) TCP congestion control (Van Jacobson & Karels)

b) Software Defined Networking, e.g. P4 or OpenFlow

c) TLS

d) Espresso (Yap et al)