CS641 Final Examination, Nov 15 2016 (9.30am to 12.30am) [40 Marks]

Your Name: ____________________________ **Your Roll Number:** _____________

Code of Honour: I hereby declare that I will not cheat (or have not cheated) in this examination. ________________ (signature)

Q1	Q2	Q3	Q4	Q5	Total

General Instructions:

- Handwritten notes (only) are allowed.
- State any assumptions clearly. Draw neatly. Please be crisp and to the point in your answers.
- **No marks for answers without explanation**

Overview: Question 1 [17 Marks]

These are short-answer questions. Please answer in 2-4 lines and not in paragraphs. ***Explain your reasoning behind the answer.*** *Yes/no answers will not fetch you marks.*

1. Suppose there are N active peers in a Gnutella network, and each peer is connected to every other peer through a TCP connection. Suppose, each TCP connection passes through a total of M 'IP' routers. How many nodes and edges are there in the corresponding overlay network?

   An overlay doesn't involve IP routers, so
   No. of nodes = N
   No. of edges = N(N-1)/2

2. Specify a cluster selection strategy in CDNs which is optimal for the client (as opposed to Local DNS server).

   Using real-time measurements via TCP handshake where client is directed to a different cluster each time.

3. Name a scenario in which HTTP 2.0 does worse than HTTP1.1.
   There could be multiple answers here. One is as under:
   HTTP 2.0 using a single TCP connection as opposed to HTTP1.1 which can use parallel connections. Under high loss rate, the single connection will suffer while HTTP1.1 will do better.

4. Consider two multimedia systems, one employing dynamic adaptive streaming (resolution of video can vary over the duration of video) and the other traditional streaming (resolution of video fixed at beginning and does not change over time). Assume both use HTTP byte control headers to periodically generate URLs to request portions of the videos. If an organization is employing a proxy to cache these urls, which system will provide better bandwidth savings for the organization. Why?

   Traditional streaming will work better for caching since dynamic variation of content results in overall less match in the URLs

5. Give one advantage of multi-process architecture over multi-threaded server architecture.

   Multi process does not require synchronization while multi-threaded which operates over shared memory.

6. MPLS forwarding employs exact-match. Further MPLS label need not be globally unique across all routers or switches. At a router that swaps an incoming label with an outgoing label, which of these two labels can the router choose locally? Justify.

   The router has to choose the incoming label locally so as to keep it distinct. Only then it can do exact match. If they were same, they will refer to the same location in memory.

7. Two hosts both behind NAT cannot communicate with each other. However use of a third-party always-on server can facilitate this communication. Explain how along with the steps involved.

The two hosts let us call them A and B. Whenever A and B come online, they have to contact the always on third party and maintain a connection with it. If A wants to contact B, it can contact the third party which will tunnel its packets to B. This is possible since it has a connection to B initiated by B itself.

8. Consider a SDN based network where firewall functionality has to be implemented at a gateway router/switch. In the process, it is desired to minimize the traffic unnecessarily sent to the controller. [2 Marks]

   a) If we wish to have a rule where an external host cannot send traffic to an internal host unless the internal host contacts it first. What initial rules should be configured in the switch before any internal or external traffic arrives? For each rule, clearly indicate both the pattern and the action.

      pattern: input any packet from external-interface, action: drop
      pattern: input any packet from internal-interface, action: sendto-controller

      (Note: We should allow packets from outside of a connection originating from inside. But this rule will be set up later by the controller once it gets a packet from inside)

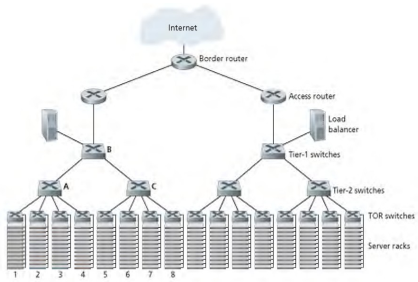
   b) Now suppose an internal host with IP address 10.2.3.4 sends a packet to an external host with IP address 15.16.7.8, what new rule(s) should the controller install in the switch?

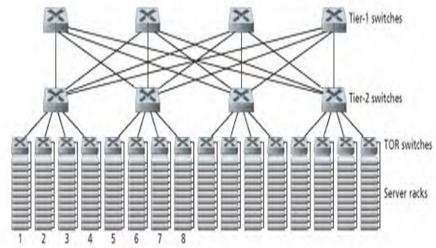
      pattern: internal interface, src-ip 10.2.3.4, dst-ip 15.16.7.8 ; action: forward
      pattern: external interface, src-ip: 15.16.7.8, dst-ip 10.2.3.4; action: forward

      (Assuming that this is done after NATing)

9. In class the following two architectures were discussed (a thumb-nail figure has been provided to jog your memory). Each rack has 10 hosts. Suppose there are 10 flows between rack 1 and 9, 10 flows between rack 2 and10 so on, totaling 80 flows. Like

before all links in the network are 10Gbps except the ones between the host and the TOR switch which are 1Gbps. [2 Marks]





(fig 1) (fig 2)

a) What is the over-subscription ratio in figure 1? Justify.

There are 80 flows that have to traverse the top level link (access router to border router) and each will get 10Gbps/80 = 125Mbps
over subscription is 8:1

b) What is the over-subscription ratio in figure 2? Justify.

There are sufficient paths, where the first two tier-1 switches will spread the traffic to all four tier-2 switches. And these top tier1 switches in turn will push the traffic to the last two tier-1 switches. There are basically 8 independent path for each rack. So, each flow will still get 1Gbps leading to an over subscription of 1:1

10. Specify the processing path taken by a packet generated by an app in a VM until it is sent out of the host's NIC. How many times is the packet normally copied in the process?

Guest app → guest OS → vnic → vswitch → host OS → NIC
three copies; one from guest app to guest OS; guest OS to host OS; host OS to NIC

11. TCP Newreno unlike reno handles multiple packet losses per window. Under what circumstances can timeouts still happen in Newreno.

There can be more than one answer. One solution is
If all the partial acks are lost it has to recover only via timeout

12. Suppose that two TCP connections A and B traverse a network consisting only of RED gateways. They both traverse through the same congested gateway, and do not encounter congestion in any of the other gateways on their path. Assume that both flows have the same RTT and packet size. [2 Marks]

a) What is the relation between the throughput achieved by connection A and B?

The congested RED gateway introduces the same loss rate to both the flows. Since RTT & pkt size are same for both the flows they experience same throughput ie. ThrA = ThrB

b) Suppose the two connections traverse the same congested gateway, but also traverse some other unshared congested gateways. What is the relation between the throughput achieved by connection A and B?

No relation exists since the loss rate suffered by A would be unrelated to that suffered by B.

13. Consider three switches S1, S2 and S3. S1 needs to serve three groups Red, Blue and Green. S2 needs to server two groups Red and Blue. S3 needs to server two groups Red and Green. It is intended to hook up the three switches so as to form three VLANs corresponding to the three groups Red, Blue and Green. In the answers below, specify clearly the number of ports needed on each switch to facilitate this interconnection and what color will be assigned to them? [2 Marks]

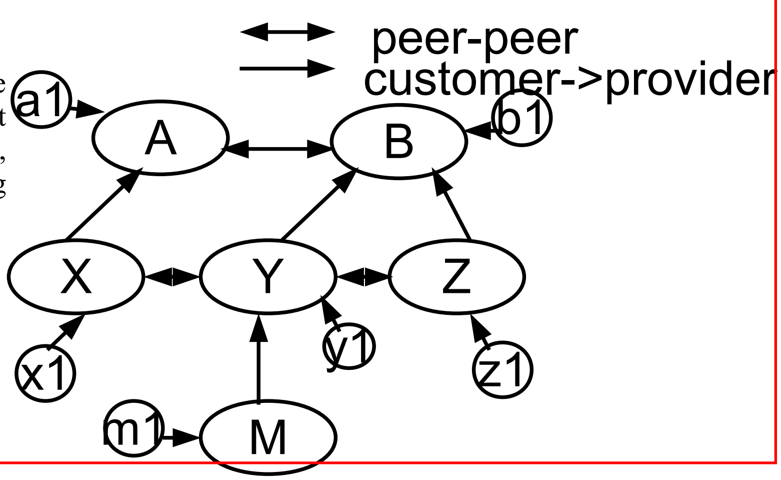
a) How would you interconnect the switches if no VLAN trunking can be employed?

There could be more than one answer.
(R:Red, B: Blue, G: Green)
S1: need 4 ports (R to S2, R to S3, B to S2 and G to S3)
S2: 2 ports (R to S1, B to S1)
S3: 2 ports (R to S1, G to S1)
Note in above S1 interconnects S2 and S3 to serve R
(S1 has 3 and S2 has 3 and S3 has 2 is also valid, if S2 interconnects S1 and S3 to serve R.)
(S1 has 3 and S2 has 2 and S3 has 3 is also valid, if S3 interconnects S1 and S1 to serve R.)

b) How would you interconnect the switches if VLAN trunking can be employed

S1: 2 ports (one to S2 marked R/B and one to S3 marked R/G)
S2: one port to S1marked R/B
S3: one port to S1 marked R/G

BGP: Question 2: [5 Marks]
The network below depicts the relationship between different autonomous systems (A,B, X, Y, Z, M) that employ BGP routing protocol.



Consider the following paths between pairs of end-hosts. For each path, specify whether the path is valid or not (note that a valid path may still not be selected by BGP for other reasons like AS hop count etc). Explain the reasoning behind your answers.

1. x1 to b1: x1-X-A-B-b1

Ans: Valid
Reason: A & B are in peering relation but have agreed to serve each other' customers

2. x1 to b1: x1-X-Y-B-b1

Ans: Not valid
Reason: Y wouldn't carry X' traffic since b1 is not it' customers

3. m1 to z1: m1-M-Y-Z-z1

Ans: Valid
Reason: Both peers y & Z will carry each other' traffic

4. x1 to z1: x1-X-Y-Z-z1

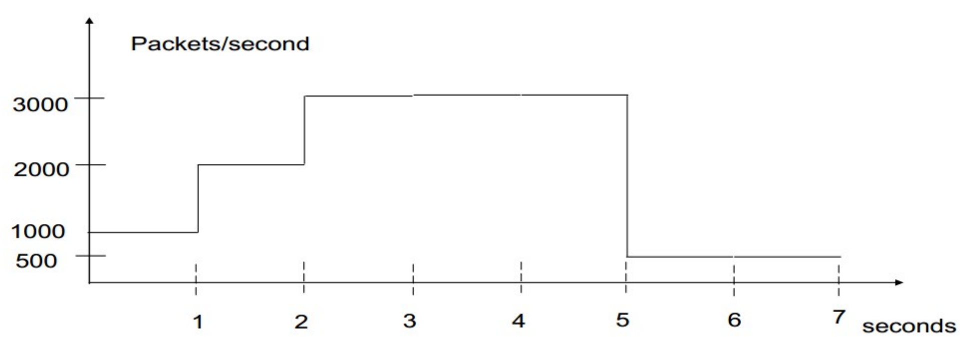
Ans: Not valid
Reason: Y will not carry X's traffic not destined to it's customers

5. x1 to z1: x1-X-A-B-Z-z1

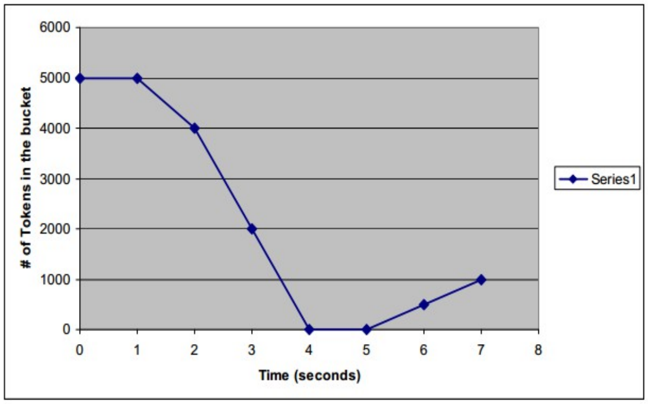
Ans: Valid
Reason: A & B (peers) have agreed to serve each other's customers

Token Bucket: Question 3: [4 Marks]

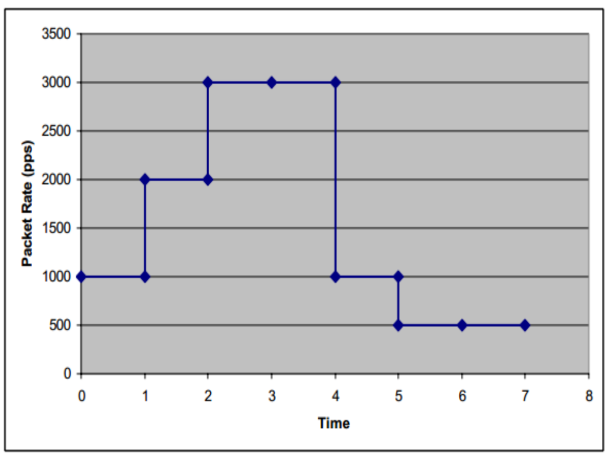
A router uses the token bucket algorithm to regulate flows. Suppose an incoming flow declares the following parameters for the token bucket algorithm: Bucket depth b = 5000 packets. Average rate r = 1000 packets per second. The actual arrival rate of the flow in terms of packets as a function of time is shown in the graph below. Assume initially the bucket is full and the arrival packet is dropped if there is no token in the bucket.



1. Draw the number of tokens in the bucket as a function of time.



2. Draw the packet rate sent out by the router as a function of time.

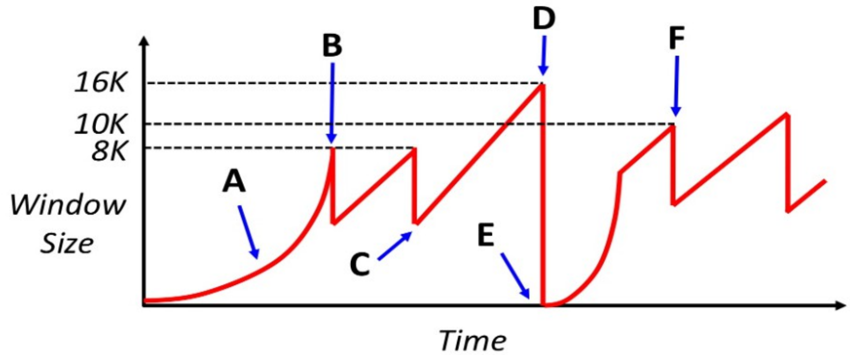


3. Are there lost packets? If so, how many?

Yes, from 4 to 5 seconds, # of packet loss is (3000-1000)*(5-4) = 2000 packets.

TCP: Question 4: [7 Marks]

Consider the following graph of TCP throughput (NOT DRAWN TO SCALE), where the y-axis describes the TCP window size of the sender. **State any assumptions made.**



1. Name the event at B that causes the sender to decrease its window.

   Triple duplicate acks

2. Does the event at B positively mean that the network discarded a packet? Why or why not?

   No, the packet could be reordered

3. Name the event at D which occurs that causes the sender to decrease its window.

   Time Out

4. Does the event at D necessitate that the network discarded a packet? Why or why not?
   No. Congestion in either direction could cause RTT > RTO (retrans. timeout).

5. Assume that the network has an MSS of 1000 bytes and the round triptime between sender and receiver of 100 milliseconds. Assume at time 0 the sender starts sending data. How much time has progressed by point B?

You don't have to consider TCP handshake time here. Further answer depends on if you interpret B as before or after you received the triple duplicate ACK, so both answers are accepted.

start (1, 2, 4, (8) MSS) = 3 or 4 RTT = 300 or 400 ms

6. How much time has progressed between points C and D?

   4 MSS to 16 MSS = 12 periods of RTT = 1.2 s

7. How much time has progressed between points E and F?

slow start to 8K, followed by AI to 10. If you interpret F as before or after you received the triple duplicate ACK, so both answers are accepted.
Either 5 or 6 RTT = 500 or 600 ms.

8. Give one explanation for why point D is higher than point B?

The cross-traffic by other concurrent senders across same route can vary over time and give rise to different values of cwnd.

Scheduling: Question 5: [7 Marks]
Consider a system of four queues being serviced according toWFQ scheduling policy. The weights given to the four queues A, B, C and D are 3, 2 , 4 and 1 respectively. They are being serviced at a total rate of 10Mbps. The table below gives a list of different input traffic rates in Mbps at the four input queue. Fill in the resultant output rates for each of the four queues. Justify your logic in a few lines.

Input Rates (Mbps)				Output Rates (Mbps)			
A	B	C	D	A	B	C	D
2	2	1	1	2	2	1	1
6	6	1	2	4.5	3	1	1.5
8	0	0	8	7.5	0	0	2.5
5	1	3	5	4.5	1	3	1.5

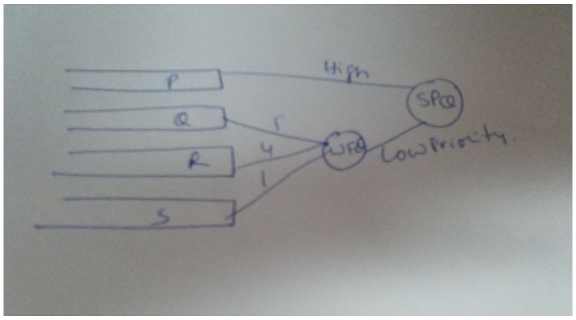
The logic is that you give the minimum of 3,2,4 and 1 to the four queues. If there were a surplus, given that some queue does not require so much, the surplus is again distribute in the ratio of weights.

If TDM (time division multiplexing) is used instead, where each queue obtains a fixed share of the bandwidth based on the same weights as WFQ, under which circumstances TDM and WFQ are equivalent?

When all input rates are all smaller than their fair share or all larger than their fair share.

Apart from WFQ and TDM, consider another queuing discipline called strict priority queuing (SPQ). In SPQ, the queue with higher priority is always scheduled ahead of the queue with lower priority. Now consider the following scenario:

There are four classes of traffic P, Q, R and S. Traffic class P has higher priority than traffic class Q, R and S. Class Q should get 5 times the amount of bandwidth of class S. Class R should get 4 times the amount of bandwidth of class S. Given these requirements, assign each traffic class into one queue and draw a diagram to describe how you would schedule the traffic using what type of scheduling algorithms? Assume 10Mbps outgoing link rate. Also fill in the table below based on the above specifications. Justify the values.



Input Rates				Output Rates			
P	Q	R	S	P	Q	R	S
2	6	2	1	2	5	2	1
1	6	10	6	1	4.5	3.6	0.9

P gets whatever input rate it has. After you subtract this input rate from 10Mbps, the remaining bandwidth is distributed via WFQ using the same logic as before.

***** Phew! Done. Have a great winter break!****