INSTRUCTIONS:

- There are 13 pages (numbered at the bottom). Make sure you have all of them.
- Please write your name on this cover and put your initials at the top of each page in this booklet except the last.
- If you find a question ambiguous, be sure to write down any assumptions you make.
- Be clear and concise. Limit your answers to the space provided.
- This exam has a total of 83 points.

<table>
<thead>
<tr>
<th>Question</th>
<th>Points</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Total:</td>
<td>83</td>
<td></td>
</tr>
</tbody>
</table>
A Miscellaneous

1. In the following, keep your answers brief and to the point (i.e. 1-2 sentences).

(a) (1 point) What are the two types of “connections” used in Click?

Solution: Push and Pull

(b) (2 points) Can a module have both connections? Either give an example or explain why not?

Solution: Queue has both

(c) (3 points) For each of the following tasks, which 4D plane should take the most responsibility for it? Pick one of \{Decision, Dissemination, Discovery, Data\} plane for each task.

Identify how much RAM a router has?

Determining the fraction of traffic between a source and destination pair should be sent through a particular path?

On a per-packet basis in real-time, determine which output interface a packet should be sent through?

Solution: (a) Discovery plane, (b) Decision plane, (c) Data plane

(d) (2 points) Circle either OpenFlow or Active Networks for the system that accomplishes each goal better.

Active Network or OpenFlow: Easily supports new data plane behaviors
Active Network or OpenFlow: Easy to implement on existing switch hardware

Solution: Active Networks, Active Networks, OpenFlow
(e) (2 points) T/F questions on SDN, NFV and middleboxes.

[True, False] To maximize network utility, CoMb HyperApp partitions a single flow processing across multiple middlebox boxes.

[True, False] NFV depends on SDN forwarding the appropriate traffic through the middlebox appliance. As a result, NFV without SDN has no value.

**Solution:** F, F

(f) (4 points) Harry Bovik was watching streaming video on his phone in a room full of people, when he realized the wireless connection might be the reason for the bad video quality. So he decided to implement a new link-layer protocol that would automatically retransmit each link-layer segment. The new link-layer protocol would retransmit a segment for at most five times, and he found this significantly improves the performance. He is concerned that his design violates the end-to-end argument since it tries to provide reliability on a hop instead of on an end-to-end basis. Briefly explain why this is against or not against end-to-end argument?

**Solution:** Does not. This is purely for performance benefit.
TCP Performance

2. Harry Bovik builds the topology below and he is responsible for setting congestion control algorithms and queue length. Machine A is connected with Machine B via a router R. The link capacity of the link between A and R and the link between R and B is $BW$, and the round-trip time between A and B is $RTT$. Machine A and B uses standard TCP congestion control. You can make following assumptions: (i) no packet loss other than packet dropped by the router, (ii) the receiving buffer on machine A and B is infinite, (iii) the router implements drop-tail policy, (iv) A sends a long flow (infinitely amount of bytes) to B

(a) (4 points) If the router has zero queue (i.e., packets are dropped if the router doesn’t have enough capacity), what average bandwidth will TCP achieve in the steady state (i.e., AIMD)?

Solution: $0.75BW$

(b) (3 points) Harry changes the AIMD algorithm on machine A so that the congestion control window will increase by 2 segments, instead of 1, if no packet loss happens in one RTT. It is still halved if packet loss happens. What average bandwidth will TCP achieve in the steady state (i.e., AIMD)?

Solution: $0.75BW$
(c) (3 points) Then, Harry keeps the change of part (b) and further changes the AIMD algorithm on machine A so that the congestion control window will be multiplied by 0.75, instead of 0.5, if packet loss happens in one RTT. What average bandwidth will TCP achieve in the steady state (i.e., AIMD)?

**Solution:** \(0.875BW\)

(d) (4 points) Now, Harry keeps the changes of part (b) and (c), and tries to configure the queue length. How long the queue length should be in order to achieve 100% link utility? You can assume that the round-trip time is always RTT.

**Solution:** \(W \times 0.75 > RTT \times BW, W = RTT \times BW + Qsize, so Qsize = (1/0.75 - 1)RTT \times BW = 0.33RTT \times BW\)
C  BGP

3. Answer the following questions about BGP

(a) (4 points) Consider the AS graph below and answer the following questions about BGP routing policies.

The vertices are individual ASes and edges are links between them. The arrows represent customer-provider relationships where the customer points to its provider. An edge without arrows represents a link between peers.

The below list are some BGP routes to IP P (a prefix in AS 1) *RECEIVED* (not necessarily used) by particular ASes. Circle the events that could NEVER happen.

A. P (4,2,1) received by AS3
B. P (5,4,3,1) received by AS2
C. P (3,2,1) received by AS4
D. P (4,3,2,1) received by AS5

Solution: (c)(d)

Note that the question asked which routes could never be received by each AS, not which routes it would prefer not to use. In general, following Gao and Rexford’s rules will make a route possible (though not necessarily preferred) if and only if it is “valley-free”—that is, it goes up the hierarchy, may plateau, then comes back down. It may never plateau while going up or down and may not go up once it starts going down.

(b) (3 points) Suppose you are the network operator of AS X and decide what BGP advertisement to announce. You think that AS Y drops too many packets and there is an alternative route to destination AS Z. Using only BGP, is it possible for you to implement a policy stating that “traffic outbound from my AS should not cross Y?” Why or why not?
Solution: Yes, Prefer paths that don’t contain Y. Of course, if the only path to a destination contains Y, it can not reach that destination without going through Y.
During a high-profile incident in 2008, Pakistan Telecom (AS 17557) managed to block many YouTube servers (AS 36561) from most of the Internet for two hours. On that day, Pakistan Telecom falsely claimed that it “owned” YouTube servers in the range 208.65.153.0/24. Meanwhile, Google, which owned YouTube, advertised that it owned the range of address in 208.65.152.0/22 (which included YouTube and other Google servers). There is nothing to prevent two different ASes from both claiming to own the same IP address, although one of the routers must be mistaken. Within just two minutes after Pakistan Telecom announced the bad route, all the tier-1 ISPs used its route and users could not access YouTube.

(c) (3 points) Why did *EVERY* ISP use the bad route even though these IPs were also advertised by YouTube AS?

**Solution:** Because Pakistan Telecom advertised a more specific IP range.

(d) (3 points) Pakistan Telecom is connected to the AS of a US-based ISP PCCW (AS 3491). Assume PCCW is the only AS connecting with Pakistan Telecom and PCCW is the provider of Pakistan Telecom. Once Pakistan Telecom announced the bad route, what should have PCCW done to prevent the bad route from being used by other ASes?

**Solution:** It should drop the announcement. It could be because it requires manual intervention, so it takes time.

(e) (3 points) Unfortunately, PCCW failed to implement the above fix. Without changing the existing BGP or cooperation from any other group (such as PCCW), how can Google reclaim the hijacked /24 IP prefix?

**Solution:** It could (as it did) advertise more specific (e.g., /25) IP prefixes.
D  Fairness

4. Consider the topology below. Machine A is connected to B and C through a router R. The capacity of each link is shown above the line.

(a) (3 points) Initially, machine B sends a long-lived flow to machine A, there is no traffic between A and C. Router R implements fair-queuing. In terms of max-min fairness, what is the rate of this flow on the link from R to A?

Solution: 8Mbps

(b) (4 points) Now, machine C also sends a long-lived flow to machine A, while the traffic from B to A still exists. Router R implements fair-queuing. In terms of max-min fairness, what are the flow rates for C-to-A flow and B-to-A flow on the link from R to A?

Solution: C-to-A: 3Mbps, B-to-A: 7Mbps

(c) (3 points) If the router R uses CSFQ, what will be the fair share rate $\alpha$ on the link from R to A? What is the drop rate for each flow?

Solution: $\alpha = 7Mbps$. Drop rate of C-to-A is 0, and B-to-A is 1/8.
E  Router Design

5. Please briefly answer following questions on router design.

(a) (3 points) Describe what is a virtual output queue (VOQ)

Solution: Maintain per output buffer at input

(b) (2 points) Give one advantage that VOQs have over input queuing.

Solution: Solve head-of-line blocking

(c) (2 points) Give one advantage that VOQs have over output queuing.

Solution: doesn’t need high speedup

(d) (3 points) What task does iSLIP perform in a VoQ system?

Solution: It determines which VoQ should send to which output port in each time slot. It ensures that this is done fairly and efficiently.
F  Data Centers

6. Answer following questions about the design of PortLand.

(a) (4 points) Data center networks like PortLand use ethernet switches instead of IP routers. Please give two advantages that design choice provides.

Solution: Plug and play, and seamless VM migration.

(b) (4 points) However, the flat addressing schemes used in Ethernet completely destroy the scalability of routing. How do schemes like Portland get around this?

Solution: Although MAC addresses are traditionally viewed as having a flat namespace, in Portland, Pseudo MAC (PMAC) addresses are specially created for each individual vm which provide very specific location information to keep routing scalable.
7. This question is based on an simplified example from 4D. The figure below shows an example of three ASes. They peer with each other via three external BGP sessions (one eBGP session along each of the links shown in the figure) maintained by five border routers R1-R5. Assume that AS1 is a major transit network, and it announces a route to destination $d$ in its eBGP session with AS2. If AS1’s policy is to not provide AS3 with transit service for $d$, it does not announce $d$ in its eBGP sessions with AS3. However, if AS3 wishes to be unscrupulous (e.g., use AS1 for transit service without paying), it can assume AS1 does know a way to $d$.

(a) (4 points) How can AS3 send a packet to $d$ within the existing BGP protocol?

Solution: Send it through the border router R1.

(b) (5 points) To ensure AS1 doesn’t provide transit service traffic from AS3 to $d$, network operators of AS1 decide to replace one of its border routers (R1, R4) with Ethane-enabled router (the router that communicate with the Ethane Controller). Which router will you replace with an Ethane-enabled router? Briefly explain how the Ethane controller enforces the policy?

Solution: R1. The Ethane Controller then checks the source and destination of the packet, and decides to drop the packet (and subsequent ones) if the source is in AS3 and the destination is in $d$. 

G SDN in Action
Anonymous Feedback

8. (2 points) Tear this sheet off to **receive two points**. We’d love it if you handed it in either at the end of the exam or, if time is lacking, to the course secretary.

(a) Please list one thing you’d like to see improved in this class in the current or a future version.

(b) Please list one good thing you’d like to make sure continues in the current or future versions of the class.