CS-552/452 **Name:** ____________________
Spring 2023
Exam 2
4/28/2023
Time Limit: Take home (10:00 AM – 11:59 PM) B-Number: ____________________

This exam contains 13 pages (including this cover page) and 4 questions.Total of points is 85. Please answer the questions concisely, convincingly, and to the point.

Grade Table (for teacher use only)

Question	Points	Score
1	15	
2	27	
3	16	
4	27	
Total:	85	

1. (15 points) **Q1: Scalable Databases – NoSQL**
   (1) What does NoSQL stand for? (3 points)

   (2) What is the CAP Theorem? (3 points)

   (3) Please list one scenario demanding high availability? (3 points)

   (4) Please list one scenario demanding high consistency? (3 points)

   (5) What do you think are the fundamental reasons that NoSQL has better scalability than Relational (SQL) databases (3 points)?

2. (27 points) **Q2: Object Storage – Mapping**
For object storage, mapping data/objects to underlying devices is important — we would like to evenly distribute them among available physical devices. One solution is to use hash functions to calculate the mapping values, based on which we can distribute data/objects (e.g., 1:1 mapping). Suppose our mapping hash function is:

$hash\_function = file\_key \ \% \ number\_of\_device$

"%" is the modulo operation. For example, if the file key is 10001, and the device number is 8, the hash value would be 10001 % 8 = 1.

Suppose we have 10 objects with the following file keys (the mapping values for 8, 7, and 6 devices are also provided, so you do not need to calculate from the above hash functions). Given this information, please answer the following questions.

Filename	file keys	mapping values (% 8)	mapping values (% 7)	mapping values (% 6)
object1	10001	1	5	5
object2	10002	2	6	0
object3	10003	3	0	1
object4	10004	4	1	2
object5	10005	5	2	3
object6	10006	6	3	4
object7	10007	7	4	5
object8	10008	0	5	0
object9	10009	1	6	1
object10	10010	2	0	2

(1) Suppose we have 8 devices as shown in Figure 1, how do you distribute the above 10 objects based on their mapping values? Please draw it in Figure 1. (4 points)

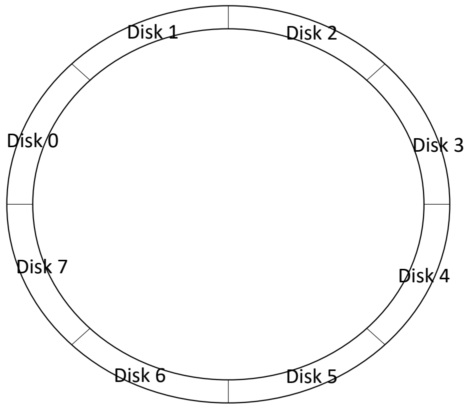

Figure 1: 8 devices.

(2) Now we remove one device, and only have 7 devices left as shown in Figure 2. How will the 10 objects be re-distributed? Please draw it in Figure 2 as well. (4 points)

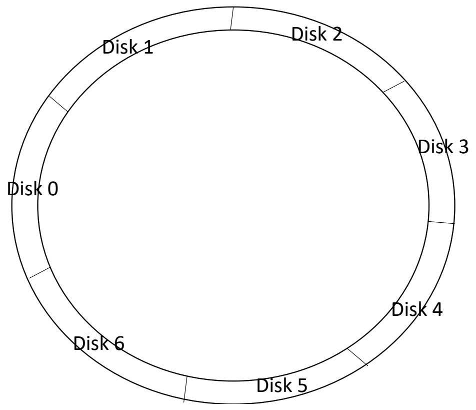

Figure 2: 7 devices.

(3) What are the main problem(s) from this (above) hash function based mapping, when we need to delete (or add) devices? (4 points)

(4) *Consistent hashing algorithm* is used to solve the above problem(s) caused by simple hash functions for object distribution/mapping among available devices. Please come up with a "consistent hashing" solution for distributing the above 10 objects among 8 devices, and demonstrate how it solves the problem(s) caused by the simple hash functions. For example, when we need to delete (or add) one device, what would happen for your "consistent hashing" solution? (5 points)

(5) Please state any issues with your "consistent hashing" solution, and briefly describe the possible solutions. If you think there are no issues, please also say so. (5 points)

(6) “Consistent hashing” does not consider the heterogeneity of the underlying devices. For example, some devices are more powerful (e.g., with more space), while some devices are not. Any solutions to consider the device heterogeneity for distributing objects? (5 points)

3. (16 points) **Q3: Object Storage – Reliability**
Data Replication is a straightforward-yet-powerful technique in large-scale, distributed systems such as object storage. For example, if one node containing one data replica fails, we can always get the data back from other replicas. However, it brings complexity for managing these replicas.

(1) Do you see any complexity for data reads? Actually, are there any benefits for reads brought by data replication? (4 points)

(2) For each update to the data, definitely we have to update all its replicas (say 3 replicas). To ensure strong data consistent — clients always see the same value of the data from different replicas — how do you design the update protocol? (4 points)

(3) What is the major drawback of the (above update protocol that ensures strong data consistency? Do you have some ideas to address this (hint: we may not need data consistency right away. Instead, we can make data consistent *eventually*, after a while.)? (4 points)

(4) Under some scenarios, we can replace "data replication" with "erasure coding" to achieve the same goals such as data durability (i.e., data is always correctly stored) and fault tolerance. Please list the pros and cons of these two techniques. Can you come up with *one* representative scenario that we prefer "erasure coding" instead of "data replication"? (4 points)

4. (27 points) **Q4: Software Defined Network**

   (1) Alice implements a new network algorithm running on the controller side as shown in Figure 4. Suppose there are no rules in all OpenFlow switches, initially. Please describe the main steps how Alice's algorithm works in such an OpenFlow based SDN network setup. (5 points)

   (2) Communication between the SDN controller and OpenFlow switches is expensive. Which types of flows, short flows (i.e., containing fewer packets) or long flows (i.e., containing more packets), suffer more from such communication cost, and why? What can we do to reduce this overhead. (5 points)

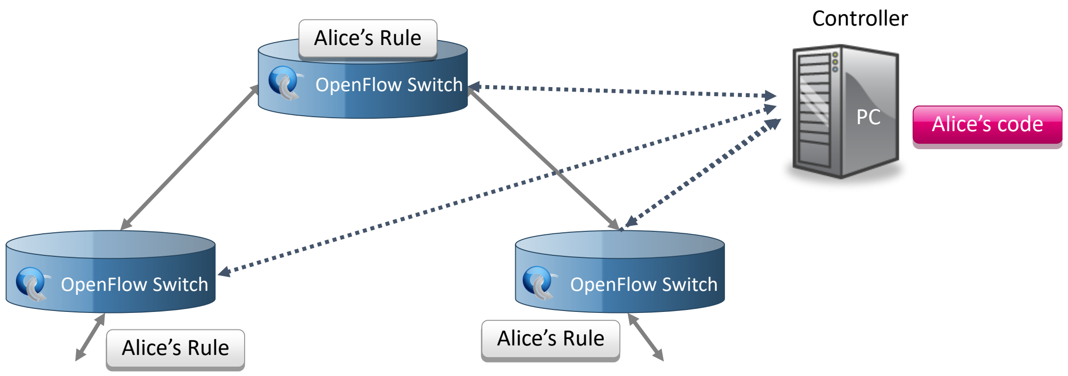

Figure 3: An SDN network setup.

In OpenStack, network services (i.e., neutron) leverages Open vSwitch (OVS) to create virtual networks for different tenants, as shown in Figure 4.

(3) To ensure isolation among VMs (e.g., VM1 and VM1') belonging to different tenants within a single physical host, what isolation technique should OVS use? Which network service of OpenStack (L2 or L3 agent) will be involved? Note that, L2 agent is for switching services, while L3 agent is for routing services. (4 points)

(4) How to ensure that VM1 can communicate with VM2 (both belonging to Tenant 1)? Which network services of OpenStack (L2 or L3 agent) will be involved? (4 points)

(5) What functions/services do we need to enable Tenant 1 and Tenant 2 to talk to each other? Which network services of OpenStack (L2 or L3 agent) will be involved? (4 points)

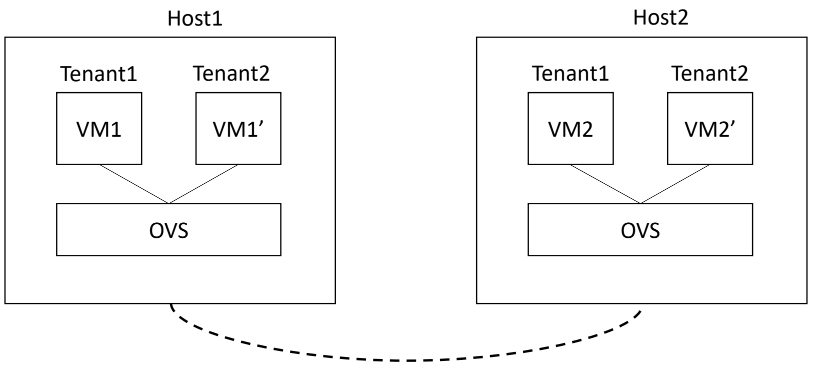

Figure 4: Dual I/O stacks under virtualization.

(6) As we can see from Google Cloud Platform (and also from OpenStack), Firewall rules are associated with VMs (e.g., blocking most of the ports). We can realize such rules within OVS. Please realize the following Firewall rules using ovs flow table entries: Blocking all ports except for port 80 and 22, for VM1 with IP being 192.168.1.1. Please state all your assumptions. (5 points)