CS 168 Spring 2024 Networking | Introduction to the Internet: Architecture and Protocols | Sylvia Ratnasamy & Rob Shakir | Discussion 12: Host

1 Multiple Choice

1. Which of the following is true of datacenters?

   (a) Many applications that run in datacenters have extreme performance requirements.

   (b) The hardware cost of the network is more expensive than that of the compute.

   (c) CPU cycles are especially important since a cloud provider may rent them to customers for additional profit.

   (d) Datacenter operators have full administrative control over their datacenters.

   Solution: (a), (c), (d)

2. Why would one use kernel bypass?

   (a) To customize networking behavior

   (b) To avoid making changes to kernel code

   (c) For performance optimization

   (d) To save money on hardware

   Solution: (a), (b), (c)

3. T/F: Stateless offloads are easier to implement since all information needed is contained in each individual packet.

   (a) True (b) False

   Solution: True

4. T/F: NIC offloads can be easily changed.

   (a) True (b) False

   Solution: False

5. T/F: Segmentation offload is a stateless offload.

   (a) True (b) False

   Solution: True

6. Which of the following are true of congestion control signals?

(a) ECN contains more information than loss.
(b) ECN contains more information than delay.
(c) Delay can be difficult to measure at high precision.
(d) Receive rate (bandwidth) contains more information than ECN.

Solution: (a), (c)

7. Which of the following are true of QoS?

(a) It only matters when a link is at 100% utilization.
(b) Only one flow can be in each class.
(c) It is used to prioritize some types of traffic.
(d) The queue can enforce the policy with weighted fair queueing.

Solution: (a), (c), (d)

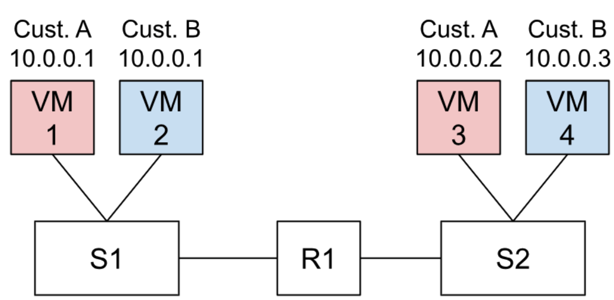
2 Network Virtualization Offload

Suppose you are a cloud provider implementing network virtualization. You want to allow a single server to host multiple VMs with potentially the same private IP addresses. In order for the servers hosting VMs to know which VM to send a packet to if there are, the packets must have some sort of identifier other than the destination IP. Accordingly, you want to add a virtual network header to each packet to distinguish it from other customer's virtual networks.

1. Why would you want to allow this addressing overlap?
Solution: For better scalability (not limited by the number of private IPs in how many VMs you can host) and so that tenants do not have to worry about others' address spaces.

2. Where could you implement this functionality?
Solution: This can be done in the hypervisor/software switch or the NIC.

3. Assume the setup below. What rules should your SDN controller program into S1 in order for the below network to be routable? Fill in the table with the necessary match-action rules.



Assume the packets in the underlay network look like the one below. The virtual network header has one field, the ID, which has the name of the customer in it.

IP	VNET ID	IP	Payload

Header Field	Value	Action
Virtual Network ID		

Solution:

Header Field	Value	Action
Virtual Network ID	A	Forward to VM 1
Virtual Network ID	B	Forward to VM 2

4. What other functionality might you put in this MA table?

   (a) Bandwidth limiting

   (b) HTTP caching

   (c) Access control

   (d) Accounting

   Solution: (a), (c), (d)

3 RDMA

Assume your application is about to use RDMA to transfer some data. In this question, we will walk through the steps to complete the process. Assume that we will use the RDMA Send operation discussed in lecture.

1. First, what benefits come from using RDMA?

   (a) Reduce complexity

   (b) Better performance

   (c) Save CPU cycles

   (d) Faster design iteration

   Solution: (b), (c)

2. Which sides must register buffers before the transfer?
   Solution: both

3. How does the sender application start the transfer?
   Solution: It places a pointer to the memory to send in the the send queue

4. Does the receiver have to do anything for the transfer to work? If so, what?
   Solution: It must place a pointer to the buffer in the receive queue

5. Which component of each host creates a completion queue entry?
   Solution: The NIC

6. How does the application know that the transfer is complete?
   Solution: It checks the completion queue

4 Timing Wheel

Assume that there are two traffic classes at a host, A and B. Class A gets a rate of 1 Gbps while class B gets a rate of 0.4 Gbps. Assume the following about the recent packets:

- The last packet for Class A was sent at time t = $100\mu s$
- The last packet for Class B was sent at time t = $100\mu s$

Assume all packets are size 1500B.

Assume the following about the timing wheel used to shape this traffic.

- All the queues are currently empty
- A traffic class is allowed to have at most one packet in a given queue
- The horizon is 4s
- The current `now` pointer references the 3rd queue
- The current time is $110\mu s$
- The `now` pointer last moved at $107\mu s$

1. If we want to support a maximum rate of 1.5Gbps for any traffic class, what should the time granularity (amount of time covered by each queue) be? Assume this value for the granularity for the rest of the problem.
   Solution: Max packets per second = $m = 1.5 \cdot 10^9 * (8 * 1500)$
   Min time between packets = $\frac{1}{m}$ = 8μs
2. What is the minimum rate supported by the configuration from the last question?
   Solution: The minimum rate is achieved by queueing one packet per horizon, so the min rate = $(1500 * 8)/(4s) = 3Kbps$
3. How many queues are there?
   Solution: $horizon/granularity = 4s/8\mu s = 0.5 \cdot 10^6$
4. Assume a packet is ready for transmission from a socket in class A. What timestamp will it be enqueued with? Which queue is this?
   Solution: Timestamp = $100\mu s + \frac{pkt\_size}{rate} = 100\mu\ s + \frac{1500*8}{1Gbps} = 100\mu s + 12\mu s = 112\mu s$.
   It will be in enqueued in queue 3.
5. Assume a packet is ready for transmission from a socket in class B. What timestamp will it be enqueued with? Which queue is this?
   Solution: Timestamp = $100\mu s + \frac{pkt\_size}{rate} = 100\mu s + \frac{1500*8}{0.4Gbps} = 100\mu s + 30\mu s = 130\mu s$.
   It will be in enqueued in queue 5.
6. Assume the current time is now t = $112\mu s$ and a packet from a socket in class A has a new packet ready for transmission. Since t is $> 107\mu s$, you can assume all packets in the current queue have already been sent. Which packet will be dequeued next?
   Solution: The new packet from A will be enqueued into queue 5 with timestamp $124\mu s$. Even though its timestamp is sooner, it was enqueued after the packet from B and packets are dequeued in FIFO order, so B's packet will be sent first.

7. When will the `now` pointer move next?
Solution: The last shift was at time $107\mu s$, so it will move at $115\mu s$