CS3611 Computer Networks (Spring 2023)
Lab 4: Overlay Network and VXLAN

Xiangyuan Xue (521030910387)

1. For Step 1, we add a default router for the bridge. On VM 1, we run the following command.

   ```
   sudo route add default gw 192.168.56.2
   ```

   On VM 2, we run the following command.

   ```
   sudo route add default gw 192.168.56.2
   ```

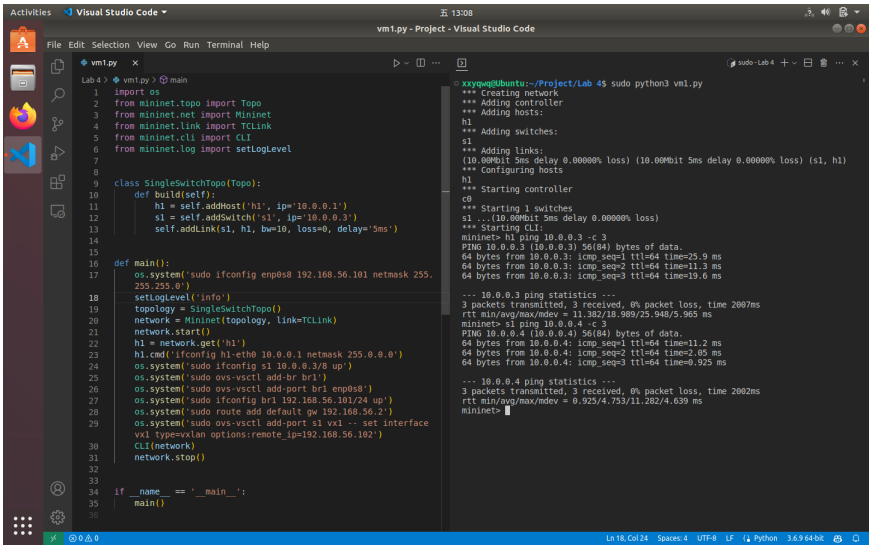
   For Step 2, we build the VXLAN tunnel and set the remote IP address. On VM 1, we run the following command.

   ```
   sudo ovs-vsctl add-port s1 vx1 -- set interface vx1 type=vxlan options
       :remote_ip=192.168.56.102
   ```

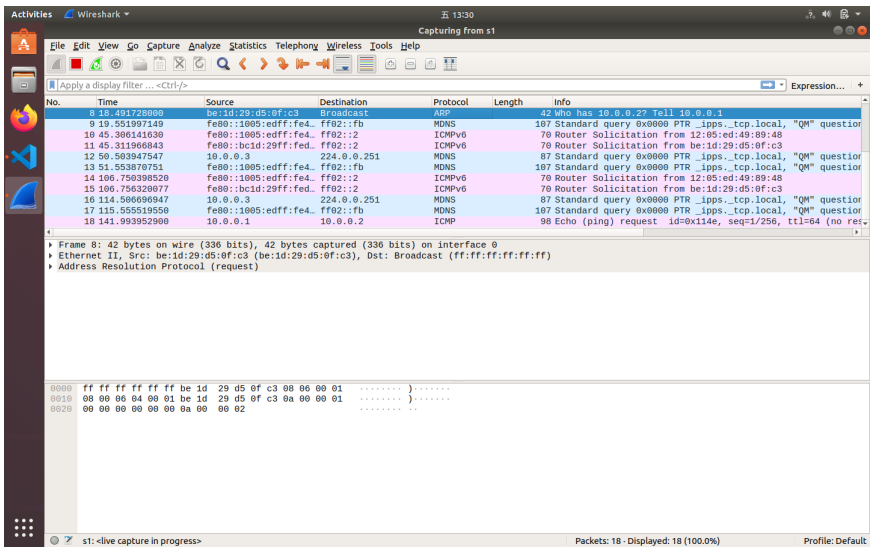
   On VM 2, we run the following command.

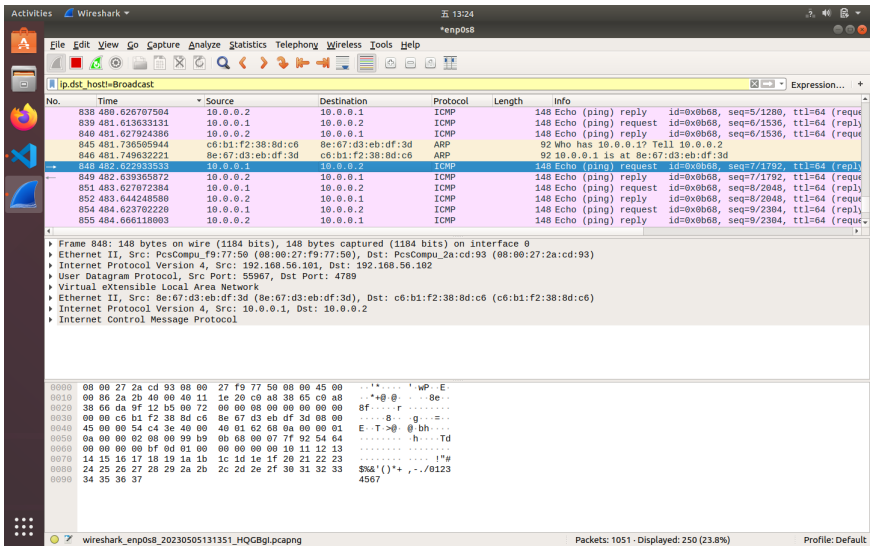
   ```
   sudo ovs-vsctl add-port s2 vx2 -- set interface vx2 type=vxlan options
       :remote_ip=192.168.56.101
   ```

   Then we can ping 10.0.0.2/10.0.0.4 from 10.0.0.1/10.0.0.3 as the figure below.



2. Use Wireshark to monitor the interfaces s1 and enp0s8 on VM 1. The results are shown in the following figures.





Explicitly, Address Resolution Protocol (ARP) and Internet Control Message Protocol (ICMP) are used in this procedure. In addition, Internet Protocol version 4 (IPv4) is implicitly and User Datagram Protocol (UDP) is encapsulated inside the VXLAN packet. Detailed descriptions are as follows.

- ARP: *ARP is a communication protocol used for discovering the link layer address, such as a MAC address, associated with a given internet layer address, typically an IPv4 address. This mapping is a critical function in the Internet protocol suite. ARP was defined in 1982 by RFC 826, which is Internet Standard STD 37.*[1]
- ICMP: *ICMP is a supporting protocol in the Internet protocol suite. It is used by network devices, including routers, to send error messages and operational information indicating success or failure when communicating with another IP address, for example, an error is indicated when a requested service is not available or that a host or router could not be reached.*[2]

- IPv4: *IPv4 is the fourth version of the Internet Protocol (IP). It is one of the core protocols of standards-based internetworking methods in the Internet and other packet-switched networks. IPv4 was the first version deployed for production on SATNET in 1982 and on the ARPANET in January 1983.*[3]
- UDP: *UDP is one of the core communication protocols of the Internet protocol suite used to send messages (transported as datagrams in packets) to other hosts on an Internet Protocol (IP) network. Within an IP network, UDP does not require prior communication to set up communication channels or data paths.*[4]

The workflow of VXLAN is specified as follows: The original packet is passed to the VXLAN Tunnel Endpoints (VTEP). The VTEP adds the VXLAN header and the UDP header to the original packet and sends the packet to the destination VTEP. After receiving the packet, the destination VTEP removes the UDP header and sends the original packet to the destination host according to the VXLAN Network Identifier (VNI) in the VXLAN header.

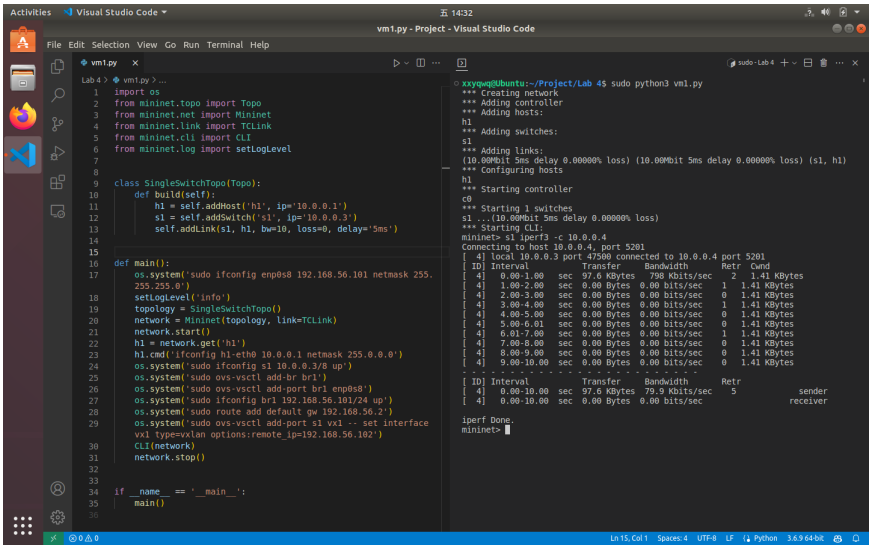
3. Start iperf3 server on VM 2. We run the following command on s2.

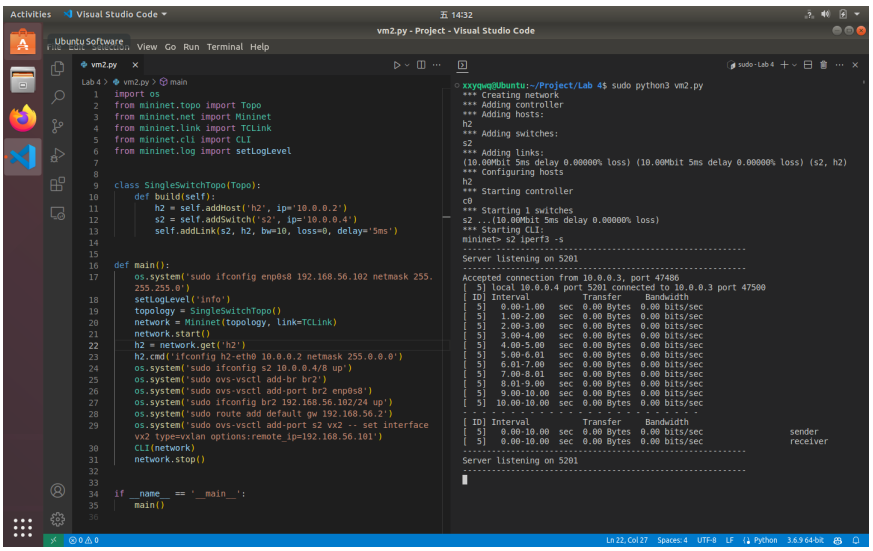
```
iperf3 -s
```

Start iperf3 client on VM 1. We run the following command on s1.

```
iperf3 -c 10.0.0.4
```

The results are shown in the following figures.





We can see that the bandwidth between 192.168.56.101 and 192.168.56.102 is approximately 798Kbits/sec at the beginning, which is extremely low. Later the bandwidth drops to zero and all the packets are lost. This problem arises from the unreasonable Maximum Transmission Unit (MTU) setting. Run the following command on VM 1.

```
ifconfig br1
```

We can see that the MTU size of the bridge is 1500Bytes. However, the packet is encapsulated in the VXLAN header and the UDP header, which enlarges the packet size. Since the packet size exceeds the MTU limitation, it will be split into several packets. However, the VTEP does not support the packet fragmentation, so the bandwidth drops shaply and other packets are lost.[5]

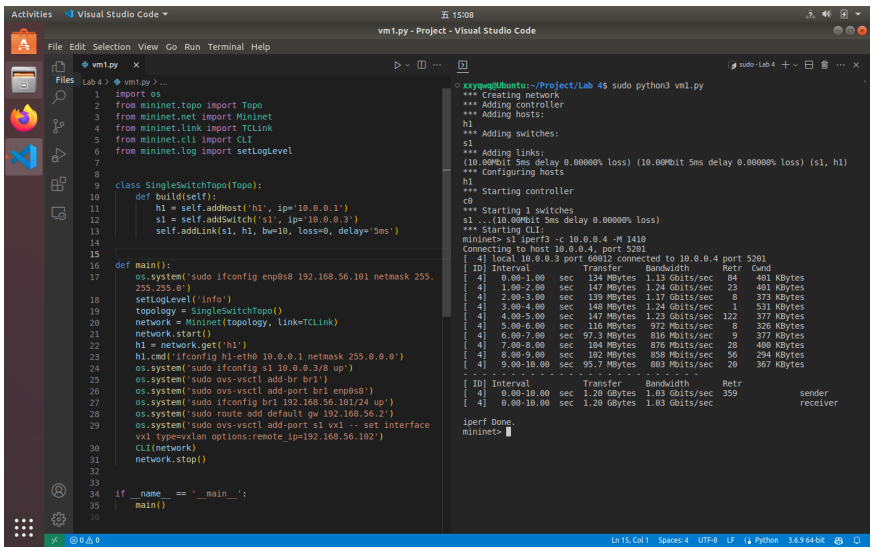
4. (1) Now we try to determine the reasonable MTU size. The MTU size of the bridge is 1500Bytes. The VXLAN header requires 50Bytes. The UDP header requires 20Bytes, in which 12Bytes are used for checksum. Another 20Bytes are required for the IPv4 header. Therefore, the reasonable MTU size is specified as

$$1500\text{Bytes} - 50\text{Bytes} - 20\text{Bytes} - 20\text{Bytes} = 1410\text{Bytes}$$

With the reasonable MTU size, we run the following command on s1.

```
iperf3 -c 10.0.0.4 -M 1410
```

The results are shown in the following figure.



We can see that the bandwidth between 192.168.56.101 and 192.168.56.102 becomes 1.03Gbits/sec, which is a great improvement compared with the previous results.

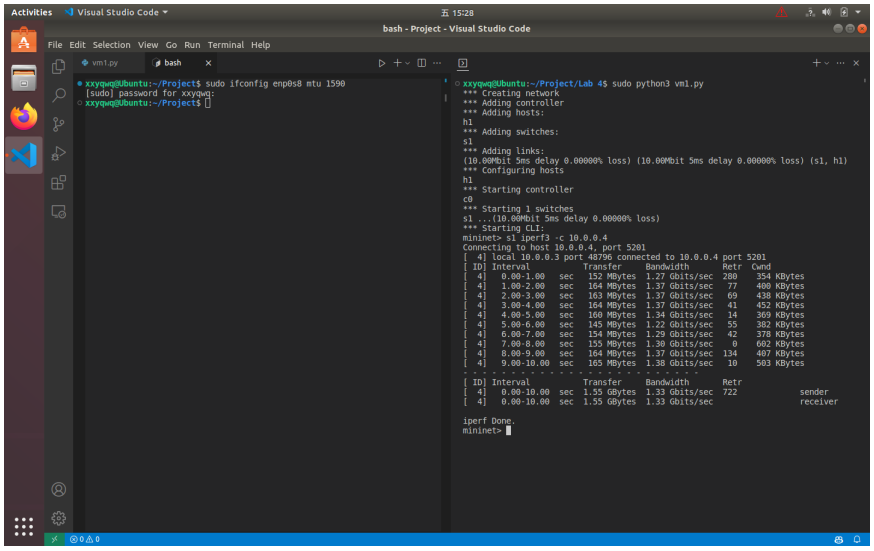
(2) Similarly, for the two enp0s8 interfaces, the reasonable MTU size is specified as

$$1500\text{Bytes} + 50\text{Bytes} + 20\text{Bytes} + 20\text{Bytes} = 1590\text{Bytes}$$

To modify the MTU size, we run the following command on both VM 1 and VM 2.

```
sudo ifconfig enp0s8 mtu 1590
```

Then we test the bandwidth again. The results are shown in the following figure.



We can see that the bandwidth between 192.168.56.101 and 192.168.56.102 becomes 1.33Gbits/sec, which is a also great improvement.

References

[1] Wikipedia. Address Resolution Protocol. https://en.wikipedia.org/wiki/Address_Resolution_Protocol.

[2] Wikipedia. Internet Control Message Protocol. https://en.wikipedia.org/wiki/Internet_Control_Message_Protocol.

[3] Wikipedia. Internet Protocol version 4. https://en.wikipedia.org/wiki/Internet_Protocol_version_4.

[4] Wikipedia. User Datagram Protocol. https://en.wikipedia.org/wiki/User_Datagram_Protocol.

[5] Wikipedia. Maximum Transmission Unit. https://en.wikipedia.org/wiki/Maximum_transmission_unit.