

# CS3611 Computer Networks (Spring 2023)

## Lab 4: Overlay Network and VXLAN

Xiangyuan Xue (521030910387)

- 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.

```

1 import os
2 from mininet.topo import Topo
3 from mininet.net import Mininet
4 from mininet.link import TCLink
5 from mininet.cli import CLI
6 from mininet.log import setLogLevel
7
8
9 class SingleSwitchTopo(Topo):
10     def build(self):
11         h1 = self.addHost('h1', ip='10.0.0.1')
12         s1 = self.addSwitch('s1', ip='10.0.0.3')
13         self.addLink(s1, h1, bw=10, loss=0, delay='5ms')
14
15     def main():
16         os.system('sudo ifconfig enp0s8 192.168.56.101 netmask 255.255.0.0')
17         setLogLevel('info')
18         topology = SingleSwitchTopo()
19         network = Mininet(topology, link=TCLink)
20         network.start()
21         h1 = network.get('h1')
22         h1.cmd('ifconfig h1 eth0 10.0.0.1 netmask 255.0.0.0')
23         h1.cmd('ifconfig s1 eth0 10.0.0.3/8 up')
24         os.system('sudo ifconfig s1 10.0.0.3/8 up')
25         os.system('sudo ovs-vsctl add-br br1')
26         os.system('sudo ovs-vsctl add-port br1 enp0s8')
27         os.system('sudo ifconfig br1 192.168.56.101/24 up')
28         os.system('sudo route add default gw 192.168.56.2')
29         os.system('sudo ovs-vsctl add-port s1 vx1 -- set interface vx1 type=vxlan options:remote_ip=192.168.56.102')
30         CLI(network)
31         network.stop()
32
33 if __name__ == '__main__':
34     main()
35

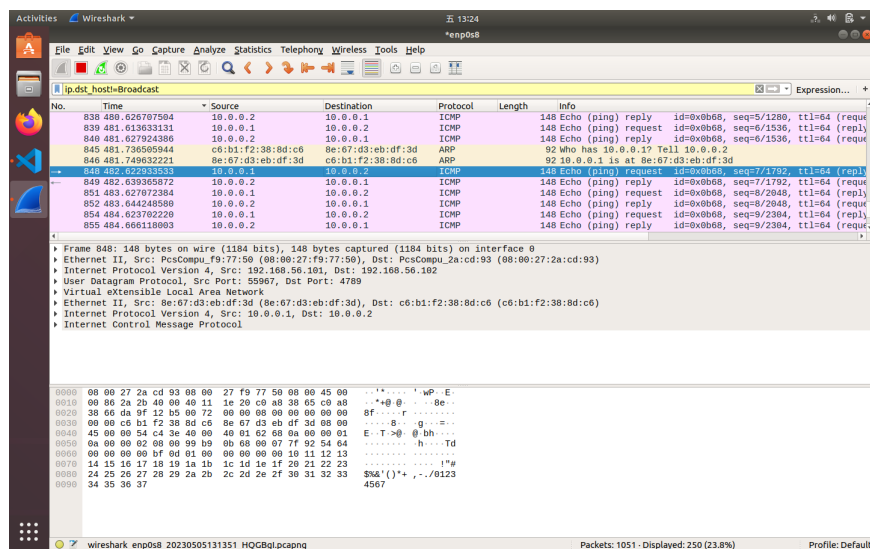
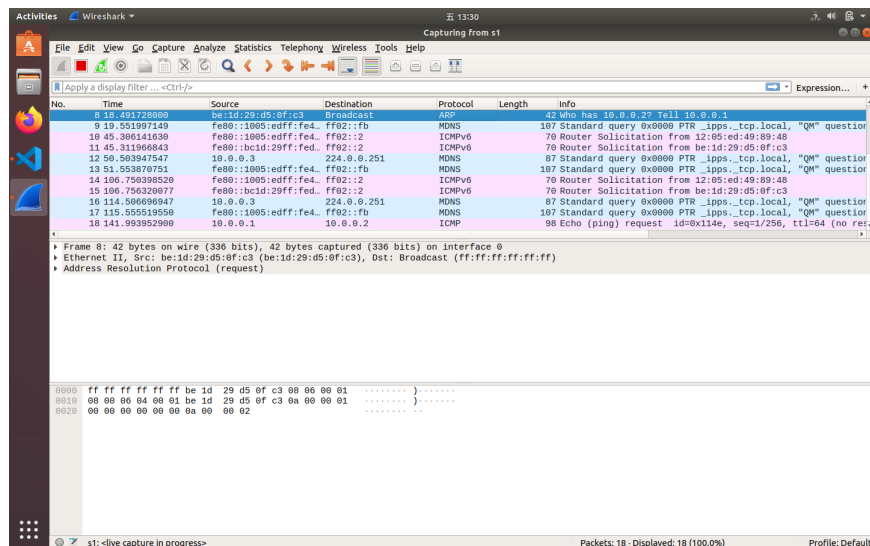
```

```

*** Creating network
*** Adding controller
*** Adding hosts:
h1
*** Adding switches:
s1
*** Adding links:
(10.000bit Sns delay 0.00000% loss) (10.000bit Sns delay 0.00000% loss) (s1, h1)
*** Configuring hosts
h1
*** Starting controller
c0
*** Starting 1 switches
s1 ... (10.000bit Sns delay 0.00000% loss)
mininet> h1 ping 10.0.0.3 -c 3
PING 10.0.0.3 (10.0.0.3): 56(84) bytes of data:
64 bytes from 10.0.0.3: icmp_seq=1 ttl=64 time=25.9 ms
64 bytes from 10.0.0.3: icmp_seq=2 ttl=64 time=11.3 ms
64 bytes from 10.0.0.3: icmp_seq=3 ttl=64 time=19.6 ms
--- 10.0.0.3 ping statistics ---
3 packets transmitted, 3 received, 0% packet loss, time 2007ms
rtt min/avg/max/mdev = 11.382/18.989/25.948/5.965 ms
mininet> s1 ping 10.0.0.4 -c 3
PING 10.0.0.4 (10.0.0.4): 56(84) bytes of data:
64 bytes from 10.0.0.4: icmp_seq=1 ttl=64 time=11.2 ms
64 bytes from 10.0.0.4: icmp_seq=2 ttl=64 time=2.09 ms
64 bytes from 10.0.0.4: icmp_seq=3 ttl=64 time=0.925 ms
--- 10.0.0.4 ping statistics ---
3 packets transmitted, 3 received, 0% packet loss, time 2002ms
rtt min/avg/max/mdev = 0.925/4.753/11.332/4.039 ms
mininet>

```

- 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.

```

1 import os
2 from mininet.topo import Topo
3 from mininet.net import Mininet
4 from mininet.link import TCLink
5 from mininet.cli import CLI
6 from mininet.log import setLogLevel
7
8
9 class SingleSwitchTopo(Topo):
10     def build(self):
11         h1 = self.addHost('h1', ip='10.0.0.1')
12         s1 = self.addSwitch('s1', ip='10.0.0.3')
13         self.addLink(s1, h1, bw=10, loss=0, delay='5ms')
14
15 def main():
16     os.system('sudo ifconfig ens8 192.168.56.101 netmask 255.255.255.0')
17     setLogLevel('info')
18     topology = SingleSwitchTopo()
19     network = Mininet(topology, Link=TCLink)
20     network.start()
21     h1 = network.get('h1')
22     h1.cmd('ifconfig h1-eth0 10.0.0.1 netmask 255.0.0.0')
23     os.system('sudo ifconfig s1 10.0.0.3/8 up')
24     os.system('sudo ovs-vsctl add-br br1')
25     os.system('sudo ovs-vsctl add-port br1 ens8')
26     os.system('sudo ifconfig br1 192.168.56.101/24 up')
27     os.system('sudo route add default gw 192.168.56.2')
28     os.system('sudo ovs-vsctl add-port s1 vx1 -- set interface vx1 type=vxlan options:remote_ip=192.168.56.102')
29     CLI(network)
30     network.stop()
31
32
33 if __name__ == '__main__':
34     main()

```

```

*** Creating network
*** Adding controller
*** Adding hosts:
h1
*** Adding switches:
s1
*** Adding Links:
(10.00Mbit Sns delay 0.00000% loss) (10.00Mbit Sns delay 0.00000% loss) (s1, h1)
*** Configuring hosts
h1
*** Starting controller
c0
*** Starting 1 switches
s1 ... (10.00Mbit Sns delay 0.00000% loss)
*** Starting CLI.
mininet> s1 iperf3 -c 10.0.0.4
Connecting to host 10.0.0.4 port 5201
[ 4] local 10.0.0.3 port 47500 connected to 10.0.0.4 port 5201
[ ID] Interval      Transfer     Bandwidth    Retr  Cwnd
[ 4] 0.00-1.00 sec  57.6 KBytes  758 Kbits/sec  2    1.41 KBytes
[ 4] 1.00-2.00 sec  0.00 Bytes  0.00 bits/sec  0    1.41 KBytes
[ 4] 2.00-3.00 sec  0.00 Bytes  0.00 bits/sec  0    1.41 KBytes
[ 4] 3.00-4.00 sec  0.00 Bytes  0.00 bits/sec  0    1.41 KBytes
[ 4] 4.00-5.00 sec  0.00 Bytes  0.00 bits/sec  0    1.41 KBytes
[ 4] 5.00-6.01 sec  0.00 Bytes  0.00 bits/sec  0    1.41 KBytes
[ 4] 6.01-7.00 sec  0.00 Bytes  0.00 bits/sec  1    1.41 KBytes
[ 4] 7.00-8.00 sec  0.00 Bytes  0.00 bits/sec  0    1.41 KBytes
[ 4] 8.00-9.00 sec  0.00 Bytes  0.00 bits/sec  0    1.41 KBytes
[ 4] 9.00-10.00 sec 0.00 Bytes  0.00 bits/sec  0    1.41 KBytes
[ ID] Interval      Transfer     Bandwidth    Retr  Cwnd
[ 4] 0.00-10.00 sec 57.6 KBytes  75.9 Kbits/sec  3    sender
[ 4] 0.00-10.00 sec 0.00 Bytes  0.00 bits/sec  0    receiver
iperf Done.
mininet>

```

```

1 import os
2 from mininet.topo import Topo
3 from mininet.net import Mininet
4 from mininet.link import TCLink
5 from mininet.cli import CLI
6 from mininet.log import setLogLevel
7
8 class SingleSwitchTopo(Topo):
9     def build(self):
10         h2 = self.addHost('h2', ip='10.0.0.2')
11         s2 = self.addSwitch('s2', ip='10.0.0.4')
12         self.addLink(s2, h2, bw=10, loss=0, delay='5ms')
13
14 def main():
15     os.system('sudo ifconfig ens8 192.168.56.102 netmask 255.255.255.0')
16     setLogLevel('info')
17     topology = SingleSwitchTopo()
18     network = Mininet(topology, link=TCLink)
19     network.start()
20     h2 = network.get('h2')
21     h2.cmd('ifconfig h2-eth0 10.0.0.2 netmask 255.0.0.0')
22     os.system('sudo ifconfig s2 10.0.0.4/8 up')
23     os.system('sudo ovs-vsctl add-br br2')
24     os.system('sudo ovs-vsctl add-port br2 ens8')
25     os.system('sudo ifconfig br2 192.168.56.102/24 up')
26     os.system('sudo route add default gw 192.168.56.2')
27     os.system('sudo ovs-vsctl add-port s2 vx2 -- set interface vx2 type=vxlan options:remote_ip=192.168.56.101')
28     CLI(network)
29     network.stop()
30
31 if __name__ == '__main__':
32     main()

```

```

*** Creating network
*** Adding controller
*** Adding hosts:
h2
*** Adding switches:
s2
*** Adding links:
(10.00Mbit Sns delay 0.00000% loss) (10.00Mbit Sns delay 0.00000% loss) (s2, h2)
*** Configuring hosts
h2
*** Starting controller
C0
*** Starting 1 switches
s2 ... (10.00Mbit Sns delay 0.00000% loss)
*** Starting CLI:
mininet> s2 iiperf3 -s
Server listening on 5201
Accepted connection from 10.0.0.3, port 47486
[ 5] local 10.0.0.4 port 5201 connected to 10.0.0.3 port 47500
[ 5] 0.00-1.00 sec 0.00 Bytes 0.00 bits/sec
[ 5] 1.00-2.00 sec 0.00 Bytes 0.00 bits/sec
[ 5] 2.00-3.00 sec 0.00 Bytes 0.00 bits/sec
[ 5] 3.00-4.00 sec 0.00 Bytes 0.00 bits/sec
[ 5] 4.00-5.00 sec 0.00 Bytes 0.00 bits/sec
[ 5] 5.00-6.01 sec 0.00 Bytes 0.00 bits/sec
[ 5] 6.01-7.00 sec 0.00 Bytes 0.00 bits/sec
[ 5] 7.00-8.01 sec 0.00 Bytes 0.00 bits/sec
[ 5] 8.01-9.00 sec 0.00 Bytes 0.00 bits/sec
[ 5] 9.00-10.00 sec 0.00 Bytes 0.00 bits/sec
[ 5] 10.00-10.00 sec 0.00 Bytes 0.00 bits/sec
[ ID] Interval Transfer Bandwidth
[ 5] 0.00-10.00 sec 0.00 Bytes 0.00 bits/sec
sender receiver
Server listening on 5201

```

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 sharply 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.

```
iiperf3 -c 10.0.0.4 -M 1410
```

The results are shown in the following figure.

```

1 import os
2 from mininet.topo import Topo
3 from mininet.net import Mininet
4 from mininet.link import TCLink
5 from mininet.cli import CLI
6 from mininet.log import setLogLevel
7
8 class SingleSwitchTopo(Topo):
9     def build(self):
10         h1 = self.addHost('h1', ip='10.0.0.1')
11         s1 = self.addSwitch('s1', ip='10.0.0.3')
12         self.addLink(s1, h1, bw=10, loss=0, delay='5ms')
13
14 def main():
15     os.system('sudo ifconfig enp0s8 192.168.56.101 netmask 255.255.255.0')
16     setLogLevel('info')
17     topology = SingleSwitchTopo()
18     network = Mininet(topology, link=TCLink)
19     network.start()
20     h1 = network.get('h1')
21     h1.cmd('ifconfig h1-eth0 10.0.0.1 netmask 255.0.0.0')
22     os.system('sudo ifconfig s1 10.0.0.3/8 up')
23     os.system('sudo ovs-vsctl add-br br1')
24     os.system('sudo ovs-vsctl add-port br1 enp0s8')
25     os.system('sudo ovs-vsctl add-port s1 vxl -- set interface vxl type=vlan options:remote_ip=192.168.56.102')
26     CLI(network)
27     network.stop()
28
29 if __name__ == '__main__':
30     main()

```

```

*** Creating network
*** Adding controller
*** Adding hosts:
h1
*** Adding switches:
s1
*** Adding Links:
(10.00Mbit 5ms delay 0.00000% loss) (10.00Mbit 5ms delay 0.00000% loss) (s1, h1)
*** Configuring hosts
h1
*** Starting controller
c0
*** Starting 1 switches
s1 ... (10.00Mbit 5ms delay 0.00000% loss)
*** Starting CLI:
mininet> s1 iperf3 -c 10.0.0.4 -M 1410
[ 4] local 10.0.0.3 port 60012 connected to 10.0.0.4 port 5201
[ ID] Interval      Transfer     Bandwidth   Retr  Cwnd
[ 4] 0.0-1.00 sec   134 Mbytes  1.13 Gbits/sec   84   401 Kbytes
[ 4] 1.00-2.00 sec   147 Mbytes  1.24 Gbits/sec   23   401 Kbytes
[ 4] 2.00-3.00 sec   139 Mbytes  1.17 Gbits/sec    8   372 Kbytes
[ 4] 3.00-4.00 sec   148 Mbytes  1.24 Gbits/sec   11   531 Kbytes
[ 4] 4.00-5.00 sec   147 Mbytes  1.23 Gbits/sec   12   377 Kbytes
[ 4] 5.00-6.00 sec   116 Mbytes  0.972 Gbits/sec    9   326 Kbytes
[ 4] 6.00-7.00 sec   97.3 Mbytes  0.816 Mbits/sec    9   377 Kbytes
[ 4] 7.00-8.00 sec   104 Mbytes  0.876 Mbits/sec   28   408 Kbytes
[ 4] 8.00-9.00 sec   102 Mbytes  0.858 Mbits/sec   56   294 Kbytes
[ 4] 9.00-10.00 sec  95.7 Mbytes  0.803 Mbits/sec   20   367 Kbytes
[ ID] Interval      Transfer     Bandwidth   Retr  Cwnd
[ 4] 0.00-10.00 sec  1.20 Gbytes  1.03 Gbits/sec   359
[ ID] Interval      Transfer     Bandwidth   Retr  Cwnd
[ 4] 0.00-10.00 sec  1.20 Gbytes  1.03 Gbits/sec
iperf Done.
mininet>

```

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.

```

xxyy@ubuntu:~/Project$ sudo ifconfig enp0s8 mtu 1590
[sudo] password for xxyy:
xxyy@ubuntu:~/Project$

```

```

*** Creating network
*** Adding controller
*** Adding hosts:
h1
*** Adding switches:
s1
*** Adding Links:
(10.00Mbit 5ms delay 0.00000% loss) (10.00Mbit 5ms delay 0.00000% loss) (s1, h1)
*** Configuring hosts
h1
*** Starting controller
c0
*** Starting 1 switches
s1 ... (10.00Mbit 5ms delay 0.00000% loss)
*** Starting CLI:
mininet> s1 iperf3 -c 10.0.0.4
Connecting to host 10.0.0.4, port 5201
[ 4] local 10.0.0.3 port 48796 connected to 10.0.0.4 port 5201
[ ID] Interval      Transfer     Bandwidth   Retr  Cwnd
[ 4] 0.00-1.00 sec   152 Mbytes  1.27 Gbits/sec   88   354 Kbytes
[ 4] 1.00-2.00 sec   164 Mbytes  1.37 Gbits/sec   77   408 Kbytes
[ 4] 2.00-3.00 sec   163 Mbytes  1.27 Gbits/sec   69   438 Kbytes
[ 4] 3.00-4.00 sec   164 Mbytes  1.37 Gbits/sec   41   452 Kbytes
[ 4] 4.00-5.00 sec   160 Mbytes  1.34 Gbits/sec   14   389 Kbytes
[ 4] 5.00-6.00 sec   145 Mbytes  1.22 Gbits/sec   55   382 Kbytes
[ 4] 6.00-7.00 sec   154 Mbytes  1.29 Gbits/sec   42   378 Kbytes
[ 4] 7.00-8.00 sec   125 Mbytes  1.28 Gbits/sec    6   602 Kbytes
[ 4] 8.00-9.00 sec   164 Mbytes  1.37 Gbits/sec   134   487 Kbytes
[ 4] 9.00-10.00 sec  165 Mbytes  1.38 Gbits/sec   10   583 Kbytes
[ ID] Interval      Transfer     Bandwidth   Retr  Cwnd
[ 4] 0.00-10.00 sec  1.35 Gbytes  1.33 Gbits/sec   722
[ ID] Interval      Transfer     Bandwidth   Retr  Cwnd
[ 4] 0.00-10.00 sec  1.35 Gbytes  1.33 Gbits/sec
iperf Done.
mininet>

```

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](https://en.wikipedia.org/wiki/Address_Resolution_Protocol).
- [2] Wikipedia. Internet Control Message Protocol. [https://en.wikipedia.org/wiki/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](https://en.wikipedia.org/wiki/Internet_Protocol_version_4).
- [4] Wikipedia. User Datagram Protocol. [https://en.wikipedia.org/wiki/User\\_Datagram\\_Protocol](https://en.wikipedia.org/wiki/User_Datagram_Protocol).
- [5] Wikipedia. Maximum Transmission Unit. [https://en.wikipedia.org/wiki/Maximum\\_transmission\\_unit](https://en.wikipedia.org/wiki/Maximum_transmission_unit).