

CS339 计算机网络

Lab 7: Overlay Network and VXLAN

目录

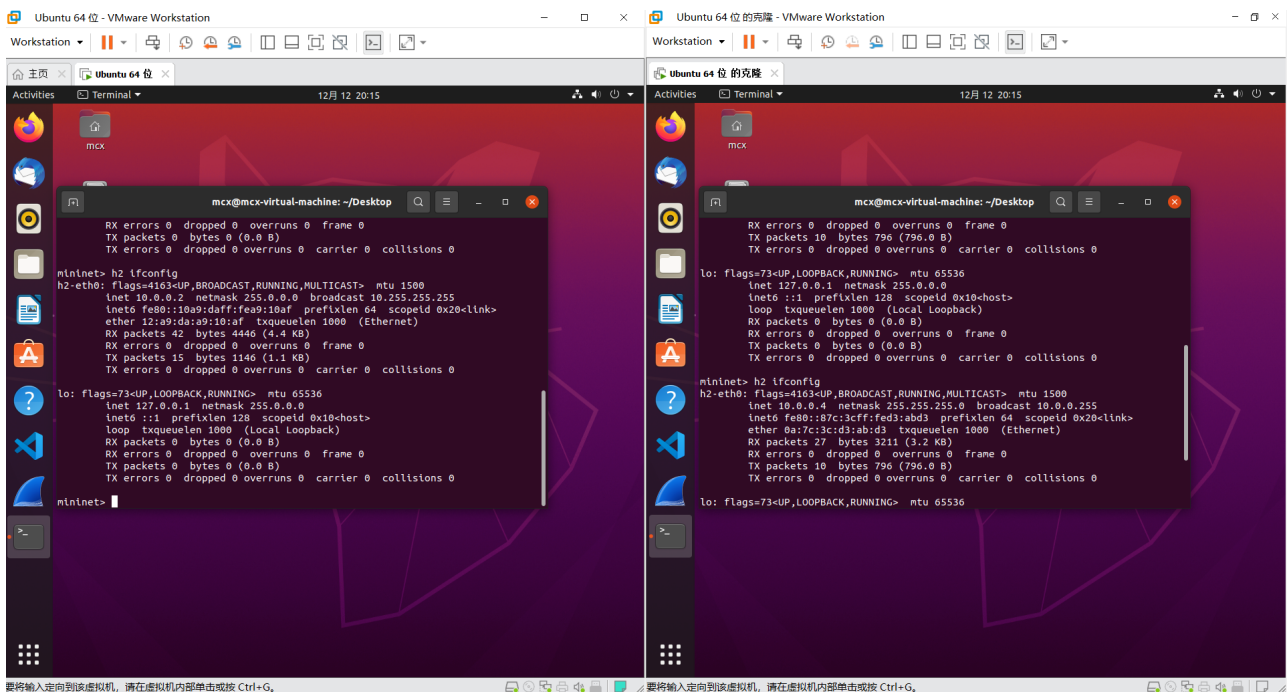
- 1 实验准备
- 2 Problem1
- 3 Problem2
- 4 Problem3

1 实验准备

创建两台虚拟机，分别运行 mininet 默认网络结构。

为VM2上的Host设置ip地址。VM1上Host的ip地址保持默认。

```
1 # VM2
2 mininet> h1 ifconfig h1-eth0 10.0.0.3 netmask 255.255.255.0
3 mininet> h2 ifconfig h2-eth0 10.0.0.4 netmask 255.255.255.0
```



尝试使用VM1上的h1 ping VM2上的h1，结果如下图所示，发现目标地址无法到达。

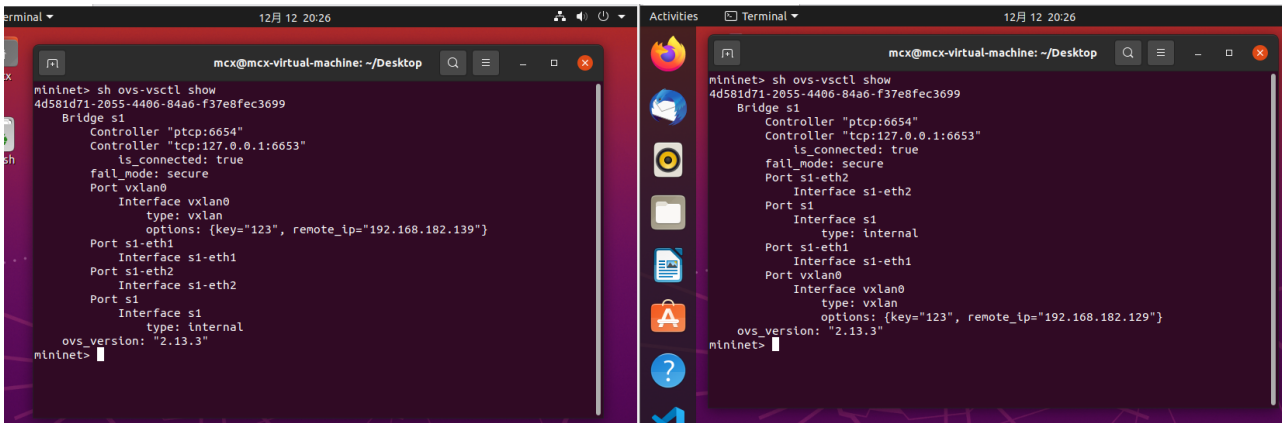
```
1 # VM1
2 h1 ping 10.0.0.3
```

```
mininet> h1 ping 10.0.0.3
PING 10.0.0.3 (10.0.0.3) 56(84) bytes of data.
From 10.0.0.1 icmp_seq=1 Destination Host Unreachable
From 10.0.0.1 icmp_seq=2 Destination Host Unreachable
From 10.0.0.1 icmp_seq=3 Destination Host Unreachable
From 10.0.0.1 icmp_seq=4 Destination Host Unreachable
From 10.0.0.1 icmp_seq=5 Destination Host Unreachable
From 10.0.0.1 icmp_seq=6 Destination Host Unreachable
From 10.0.0.1 icmp_seq=7 Destination Host Unreachable
From 10.0.0.1 icmp_seq=8 Destination Host Unreachable
From 10.0.0.1 icmp_seq=9 Destination Host Unreachable
^C
--- 10.0.0.3 ping statistics ---
12 packets transmitted, 0 received, 100% packet loss, time 11267ms
```

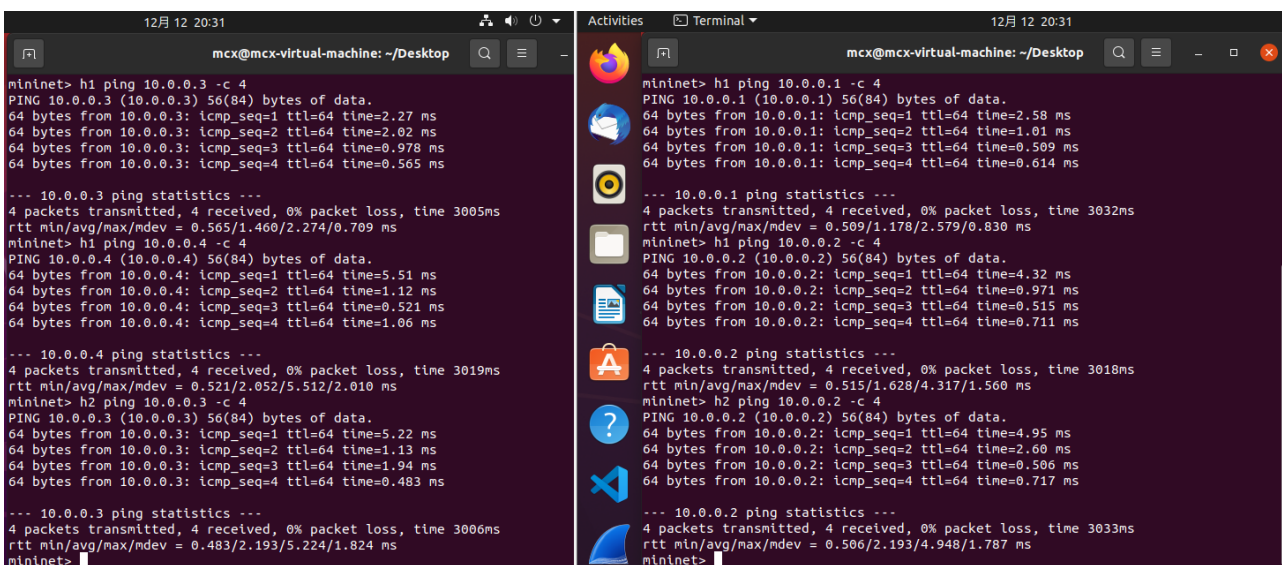
在VM1与VM2之间建立VXLAN。

```
1 # VM1
2 ovs-vsctl add-port s1 vxlan0 -- set interface vxlan0 type=vxlan
  options:remote_ip=192.168.182.139 option:key=123
3 # VM2
4 sh ovs-vsctl add-port s1 vxlan0 -- set interface vxlan0 type=vxlan
  options:remote_ip=192.168.182.129 option:key=123
```

使用 `ovs-vsctl show` 查看桥接状况，可以看出VXLAN端口已经建立。



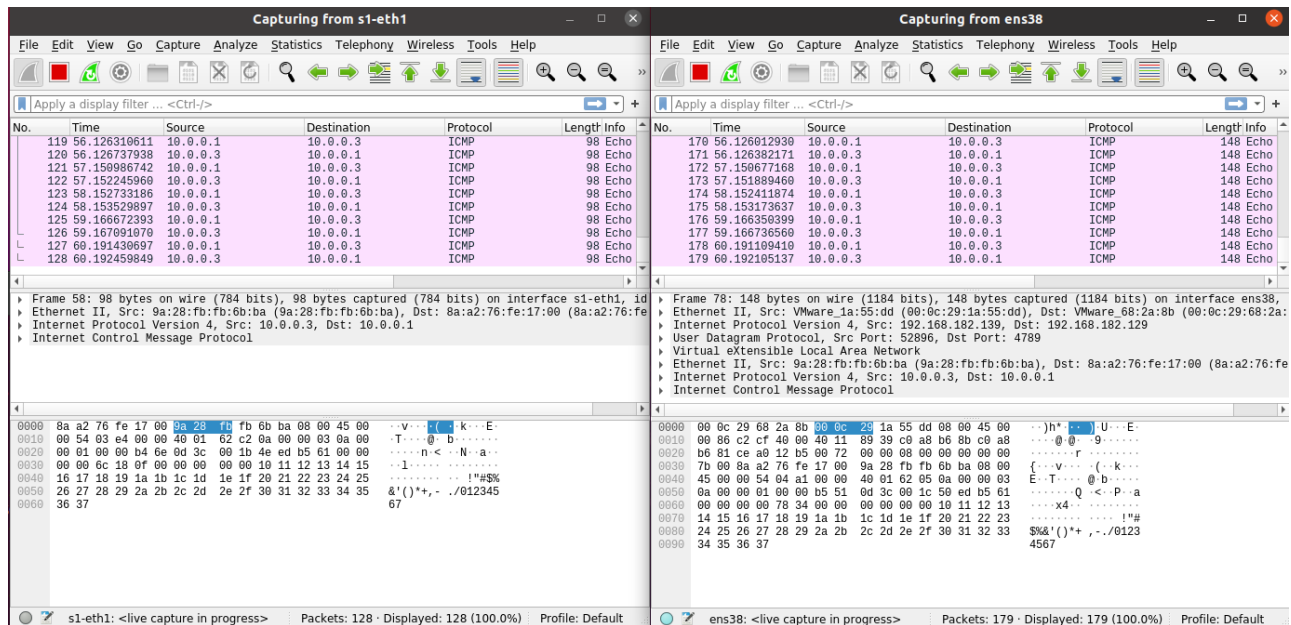
再次尝试使用VM1上的Host与VM2上的Host互相ping，结果如下图。两台主机间已经可以通信。



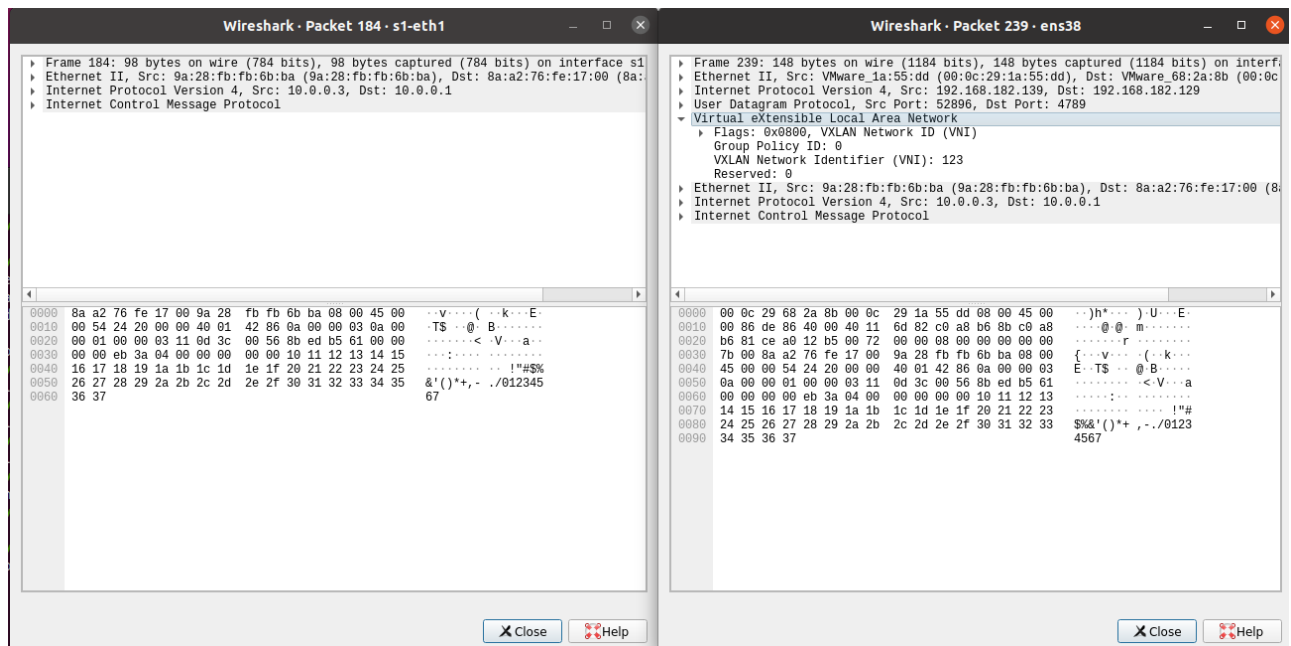
2 Problem1

使用VM1上的h1 ping VM2上的h1，同时启动两个wireshark，分别监听 s1-eth0 与 ens38，其中 s1-eth0 是 h1 与 s1 连接的接口，ens38 是VM1的网卡接口。

如下图所示，s1-eth0 接口上每一个分组都对应着 ens38 接口上的一个分组，对比两个接口上的分组信息，可以看出两个接口上分组的长度有差异，ens38 上的分组长度比 s1-eth0 长50字节。



选取对应的一组分组查看详情，可以看到 s1-eth0 的分组被完整包装在 ens38 对应分组的VXLAN字段中，而整个VXLAN数据报文又基于UDP发送到VM2。VXLAN、UDP、IP等协议的报文头长度总和为50字节，因此 ens38 上的分组长度比 s1-eth0 长50字节。



3 Problem2

使用 iperf3 测试VM1与VM2之间的连接速度。

```
mcx@mcx-virtual-machine:~/Desktop$ iperf3 -c 192.168.182.139
Connecting to host 192.168.182.139, port 5201
[ 5] local 192.168.182.129 port 35450 connected to 192.168.182.139 port 5201
[ ID] Interval      Transfer    Bitrate    Retr  Cwnd
[ 5] 0.00-1.00    sec    131 MBytes  1.10 Gbits/sec    0   197 KBytes
[ 5] 1.00-2.00    sec    158 MBytes  1.33 Gbits/sec    0   197 KBytes
[ 5] 2.00-3.00    sec    183 MBytes  1.53 Gbits/sec    0   264 KBytes
[ 5] 3.00-4.00    sec    179 MBytes  1.50 Gbits/sec    0   206 KBytes
[ 5] 4.00-5.00    sec    157 MBytes  1.32 Gbits/sec    0   216 KBytes
[ 5] 5.00-6.00    sec    161 MBytes  1.35 Gbits/sec    0   198 KBytes
[ 5] 6.00-7.00    sec    164 MBytes  1.38 Gbits/sec    0   212 KBytes
[ 5] 7.00-8.00    sec    189 MBytes  1.58 Gbits/sec    0   223 KBytes
[ 5] 8.00-9.00    sec    202 MBytes  1.70 Gbits/sec    0   211 KBytes
[ 5] 9.00-10.00   sec    170 MBytes  1.43 Gbits/sec    0   263 KBytes
-----
[ ID] Interval      Transfer    Bitrate    Retr
[ 5] 0.00-10.00   sec    1.65 GBytes  1.42 Gbits/sec    0
[ 5] 0.00-10.00   sec    1.65 GBytes  1.42 Gbits/sec    0
sender
receiver
iperf Done.
```

使用 iperf3 测试VM1上h1、h2与VM2上h1之间的连接速度。

```
1 | h1 iperf3 -c 10.0.0.3 -M 1400
2 | h2 iperf3 -c 10.0.0.3 -M 1400
```

由于VXLAN协议会额外添加一系列报文头，如果使用默认的MTU会导致通过链路层实际发送的数据分组长度大于链路层MTU限制，导致数据分组被链路层丢弃。所以需要使用 -M 参数指定MTU，这里我们设置为1400。

```
mininet> h1 iperf3 -c 10.0.0.3 -M 1400
Connecting to host 10.0.0.3, port 5201
[ 5] local 10.0.0.1 port 44986 connected to 10.0.0.3 port 5201
[ ID] Interval      Transfer    Bitrate    Retr  Cwnd
[ 5] 0.00-1.00    sec    99.0 MBytes  830 Mbits/sec    9   244 KBytes
[ 5] 1.00-2.00    sec    140 MBytes  1.17 Gbits/sec    0   218 KBytes
[ 5] 2.00-3.00    sec    151 MBytes  1.27 Gbits/sec    0   225 KBytes
[ 5] 3.00-4.00    sec    157 MBytes  1.32 Gbits/sec    1   220 KBytes
[ 5] 4.00-5.00    sec    153 MBytes  1.29 Gbits/sec    0   229 KBytes
[ 5] 5.00-6.00    sec    153 MBytes  1.28 Gbits/sec    0   241 KBytes
[ 5] 6.00-7.00    sec    157 MBytes  1.32 Gbits/sec    0   232 KBytes
[ 5] 7.00-8.00    sec    156 MBytes  1.31 Gbits/sec    0   220 KBytes
[ 5] 8.00-9.00    sec    157 MBytes  1.32 Gbits/sec    0   228 KBytes
[ 5] 9.00-10.00   sec    156 MBytes  1.31 Gbits/sec    0   233 KBytes
-----
[ ID] Interval      Transfer    Bitrate    Retr
[ 5] 0.00-10.00   sec    1.44 GBytes  1.24 Gbits/sec   10
[ 5] 0.00-10.00   sec    1.44 GBytes  1.24 Gbits/sec   10
sender
receiver
iperf Done.
```

```
mininet> h2 iperf3 -c 10.0.0.3 -M 1400
Connecting to host 10.0.0.3, port 5201
[ 5] local 10.0.0.2 port 52178 connected to 10.0.0.3 port 5201
[ ID] Interval      Transfer    Bitrate    Retr  Cwnd
[ 5] 0.00-1.00    sec    60.6 MBytes  508 Mbits/sec    0   130 KBytes
[ 5] 1.00-2.00    sec    57.2 MBytes  480 Mbits/sec    0   130 KBytes
[ 5] 2.00-3.00    sec    141 MBytes  1.18 Gbits/sec    2   230 KBytes
[ 5] 3.00-4.00    sec    163 MBytes  1.37 Gbits/sec    0   217 KBytes
[ 5] 4.00-5.00    sec    163 MBytes  1.37 Gbits/sec    0   272 KBytes
[ 5] 5.00-6.00    sec    160 MBytes  1.34 Gbits/sec    0   236 KBytes
[ 5] 6.00-7.00    sec    158 MBytes  1.32 Gbits/sec    0   229 KBytes
[ 5] 7.00-8.00    sec    155 MBytes  1.30 Gbits/sec    0   232 KBytes
[ 5] 8.00-9.00    sec    163 MBytes  1.37 Gbits/sec    0   228 KBytes
[ 5] 9.00-10.00   sec    150 MBytes  1.26 Gbits/sec    0   245 KBytes
-----
[ ID] Interval      Transfer    Bitrate    Retr
[ 5] 0.00-10.00   sec    1.34 GBytes  1.15 Gbits/sec    2
[ 5] 0.00-10.00   sec    1.34 GBytes  1.15 Gbits/sec    2
sender
receiver
iperf Done.
```

总结测试数据如下表所示。

iperf3测试结果

测试对象	连接速度(Gbits/s)
VM1 & VM2	1.42
h1(VM1) & h1(VM2)	1.24
h1(VM1) & h2(VM2)	1.15

从测试结果可以看出使用VXLAN协议的连接速度略低于VM1与VM2的连接速度，由于VXLAN协议进行报文封装时需要添加额外报文头，额外报文头占用了一部分传输带宽，所以会造成传输速度上的性能损失。

4 Problem3

使用 ping 测试VM1与VM2之间的连接延迟。

```
mcx@mcx-virtual-machine:~/Desktop$ ping 192.168.182.139 -c 10
PING 192.168.182.139 (192.168.182.139) 56(84) bytes of data.
64 bytes from 192.168.182.139: icmp_seq=1 ttl=64 time=0.499 ms
64 bytes from 192.168.182.139: icmp_seq=2 ttl=64 time=0.443 ms
64 bytes from 192.168.182.139: icmp_seq=3 ttl=64 time=0.378 ms
64 bytes from 192.168.182.139: icmp_seq=4 ttl=64 time=0.556 ms
64 bytes from 192.168.182.139: icmp_seq=5 ttl=64 time=2.06 ms
64 bytes from 192.168.182.139: icmp_seq=6 ttl=64 time=0.781 ms
64 bytes from 192.168.182.139: icmp_seq=7 ttl=64 time=0.429 ms
64 bytes from 192.168.182.139: icmp_seq=8 ttl=64 time=0.361 ms
64 bytes from 192.168.182.139: icmp_seq=9 ttl=64 time=0.435 ms
64 bytes from 192.168.182.139: icmp_seq=10 ttl=64 time=0.849 ms
--- 192.168.182.139 ping statistics ---
10 packets transmitted, 10 received, 0% packet loss, time 9168ms
rtt min/avg/max/mdev = 0.361/0.679/2.060/0.486 ms
```

使用 ping 测试VM1上h1、h2与VM2上h1之间的连接延时。

```
1 | h1 ping 10.0.0.3 -c 10
2 | h2 ping 10.0.0.3 -c 10
```

```
mininet> h1 ping 10.0.0.3 -c 10
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=4.77 ms
64 bytes from 10.0.0.3: icmp_seq=2 ttl=64 time=2.15 ms
64 bytes from 10.0.0.3: icmp_seq=3 ttl=64 time=0.875 ms
64 bytes from 10.0.0.3: icmp_seq=4 ttl=64 time=0.525 ms
64 bytes from 10.0.0.3: icmp_seq=5 ttl=64 time=1.26 ms
64 bytes from 10.0.0.3: icmp_seq=6 ttl=64 time=0.453 ms
64 bytes from 10.0.0.3: icmp_seq=7 ttl=64 time=0.494 ms
64 bytes from 10.0.0.3: icmp_seq=8 ttl=64 time=0.487 ms
64 bytes from 10.0.0.3: icmp_seq=9 ttl=64 time=0.486 ms
64 bytes from 10.0.0.3: icmp_seq=10 ttl=64 time=2.29 ms
--- 10.0.0.3 ping statistics ---
10 packets transmitted, 10 received, 0% packet loss, time 9127ms
rtt min/avg/max/mdev = 0.453/1.378/4.767/1.308 ms
```

```
mininet> h2 ping 10.0.0.3 -c 10
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=2.33 ms
64 bytes from 10.0.0.3: icmp_seq=2 ttl=64 time=1.60 ms
64 bytes from 10.0.0.3: icmp_seq=3 ttl=64 time=1.49 ms
64 bytes from 10.0.0.3: icmp_seq=4 ttl=64 time=0.590 ms
64 bytes from 10.0.0.3: icmp_seq=5 ttl=64 time=0.574 ms
64 bytes from 10.0.0.3: icmp_seq=6 ttl=64 time=0.430 ms
64 bytes from 10.0.0.3: icmp_seq=7 ttl=64 time=1.91 ms
64 bytes from 10.0.0.3: icmp_seq=8 ttl=64 time=0.833 ms
64 bytes from 10.0.0.3: icmp_seq=9 ttl=64 time=1.59 ms
64 bytes from 10.0.0.3: icmp_seq=10 ttl=64 time=0.464 ms
--- 10.0.0.3 ping statistics ---
10 packets transmitted, 10 received, 0% packet loss, time 9076ms
rtt min/avg/max/mdev = 0.430/1.181/2.326/0.648 ms
```

总结测试数据如下表所示。

ping测试结果

测试对象	连接延迟(ms)
VM1 & VM2	0.679
h1(VM1) & h1(VM2)	1.378
h1(VM1) & h2(VM2)	1.181

从测试结果可以看出使用VXLAN协议的连接延迟略高于VM1与VM2直接连接，由于VXLAN协议通信的数据包需要在两端交换机经历封装、拆封过程，所以会增加一定的连接延迟，增加的连接延迟在0.6ms左右，对于现实网络连接中数十至数百毫秒的延迟而言，VXLAN协议带来的额外延迟基本可以忽略不计。