Traceroute problems on Guest OS of VMware

1. Case 1

1.1. TEST Environments

1. Host OS : Windows XP SP3

1.1. ipconfig -a

```
Windows IP Configuration
        Host Name . . . . . . . . . . . . : PRUTTA
        Primary Dns Suffix  . . . . . . . :
        Node Type . . . . . . . . . . . . : Unknown
        IP Routing Enabled. . . . . . . . : No
        WINS Proxy Enabled. . . . . . . . : No
        DNS Suffix Search List. . . . . . : 

Ethernet adapter VMware Network Adapter VMnet8:
        Connection-specific DNS Suffix  . :
        Description . . . . . . . . . . . : VMware Virtual Ethernet Adapter for VMnet8
        Physical Address. . . . . . . . . : 00-50-56-C0-00-08
        Dhcp Enabled. . . . . . . . . . . : No
        IP Address. . . . . . . . . . . . : 192.168.18.1
        Subnet Mask . . . . . . . . . . . : 255.255.255.0
        Default Gateway . . . . . . . . . :

Ethernet adapter VMware Network Adapter VMnet1:
```

```
        Connection-specific DNS Suffix  . :
        Description . . . . . . . . . . . : VMware Virtual Ethernet Adapter for VMnet1
        Physical Address. . . . . . . . . : 00-50-56-C0-00-01
        Dhcp Enabled. . . . . . . . . . . : No
        IP Address. . . . . . . . . . . . : 192.168.235.1
        Subnet Mask . . . . . . . . . . . : 255.255.255.0
        Default Gateway . . . . . . . . . :

Ethernet adapter ワイヤレス ネットワーク接続:
        Media State . . . . . . . . . . . : Media disconnected
        Description . . . . . . . . . . . : Intel(R) Wireless WiFi Link 4965AGN
        Physical Address. . . . . . . . . : 00-1F-3B-C6-83-C1

Ethernet adapter ローカル エリア接続:
        Connection-specific DNS Suffix  . :
        Description . . . . . . . . . . . : Intel(R) 82566MM Gigabit Network Connection
        Physical Address. . . . . . . . . : 00-1D-72-92-C4-E9
        Dhcp Enabled. . . . . . . . . . . : Yes
        Autoconfiguration Enabled . . . . : Yes
        IP Address. . . . . . . . . . . . : 192.168.0.14
        Subnet Mask . . . . . . . . . . . : 255.255.255.0
        Default Gateway . . . . . . . . . : 192.168.0.1
        DHCP Server . . . . . . . . . . . : 192.168.0.1
        DNS Servers . . . . . . . . . . . : 122.1.115.75
                                            122.1.115.76
                                            202.249.37.87
        Lease Obtained. . . . . . . . . . : 2009 年 4 月 7 日  9:53:13
        Lease Expires . . . . . . . . . . : 2009 年 4 月 8 日  5:53:13
```

2. Guest OS : Fedora 8 x86_64 on **VMware Player 2.5.2**

2.1. Network Setting : NAT

2.2. Uname –a and ifconfig eth0,

```
[root@localhost ~]# uname -a
Linux localhost.localdomain 2.6.26.8-57.fc8 #1 SMP Thu Dec 18 18:59:49 EST 2008 x86_64 x86_64 x86_64 GNU/Linux
```

```
[root@localhost ~]# ifconfig eth0
eth0      Link encap:Ethernet  HWaddr 00:0C:29:EB:29:BC
          inet addr:192.168.18.129  Bcast:192.168.18.255  Mask:255.255.255.0
          inet6 addr: fe80::20c:29ff:feeb:29bc/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:2902 errors:0 dropped:0 overruns:0 frame:0
          TX packets:1787 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:3428054 (3.2 MiB)  TX bytes:123609 (120.7 KiB)
          Interrupt:17 Base address:0x1080
```

1.2. TEST1: traceroute www.google.com from Host and Guest

1.2.1. Traceroute from Guest

```
[root@localhost ~]# traceroute www.google.com
traceroute to www.google.com (66.249.89.147), 30 hops max, 60 byte packets
 1  192.168.18.2 (192.168.18.2)  0.163 ms  0.107 ms  0.101 ms
 2  jp-in-f147.google.com (66.249.89.147)  5.255 ms  5.052 ms  4.937 ms
```

Traceroute shows only first and last hops. I think that host OS have changed TTL to 128 when host (vmnet-natd?) forwards the UDP packets received from guest (see also 1.2.3)

1.2.2. Traceroute from Host

```
Tracing route to www.l.google.com [66.249.89.147]
over a maximum of 30 hops:
  1    <1 ms    <1 ms    <1 ms  192.168.0.1
```

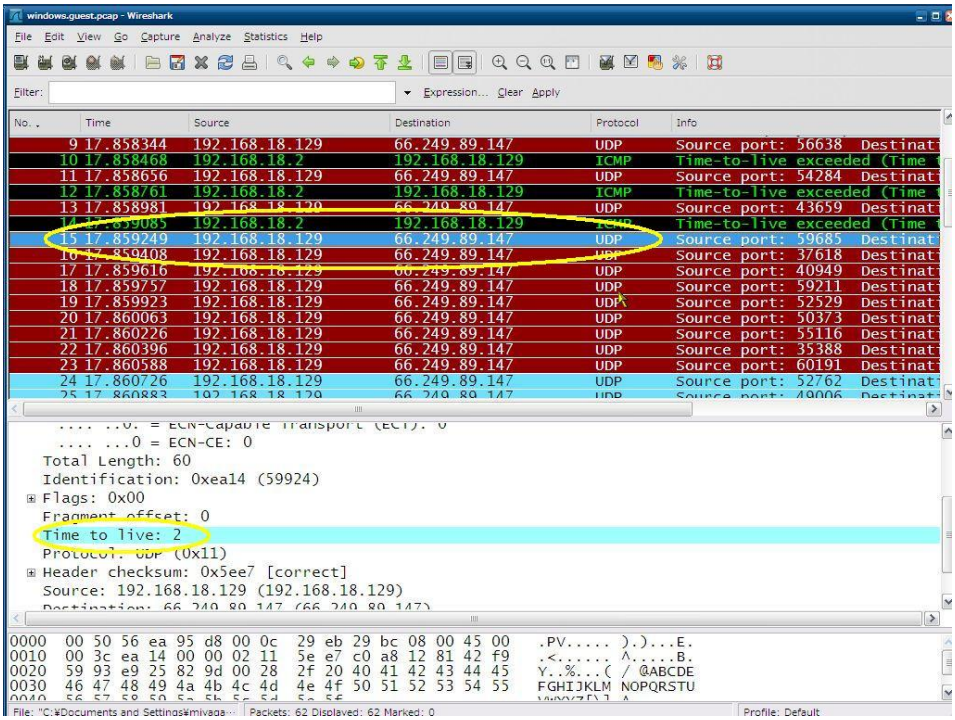
2	1 ms	1 ms	1 ms	118.23.99.139
3	2 ms	2 ms	2 ms	118.23.99.158
4	4 ms	3 ms	6 ms	221.184.12.233
5	2 ms	1 ms	2 ms	60.37.11.41
6	2 ms	2 ms	2 ms	210.254.188.141
7	2 ms	2 ms	2 ms	210.254.188.146
8	2 ms	21 ms	2 ms	122.28.104.90
9	2 ms	2 ms	2 ms	210.163.230.234
10	3 ms	2 ms	2 ms	xe-1-1.a17.tokyjp01.jp.ra.gin.ntt.net [61.213.169.66]
11	3 ms	3 ms	3 ms	xe-1-3.a17.tokyjp01.jp.ra.gin.ntt.net [61.213.169.174]
12	3 ms	3 ms	3 ms	209.85.241.94
13	4 ms	14 ms	3 ms	72.14.236.126
14	4 ms	3 ms	3 ms	jp-in-f147.google.com [66.249.89.147]

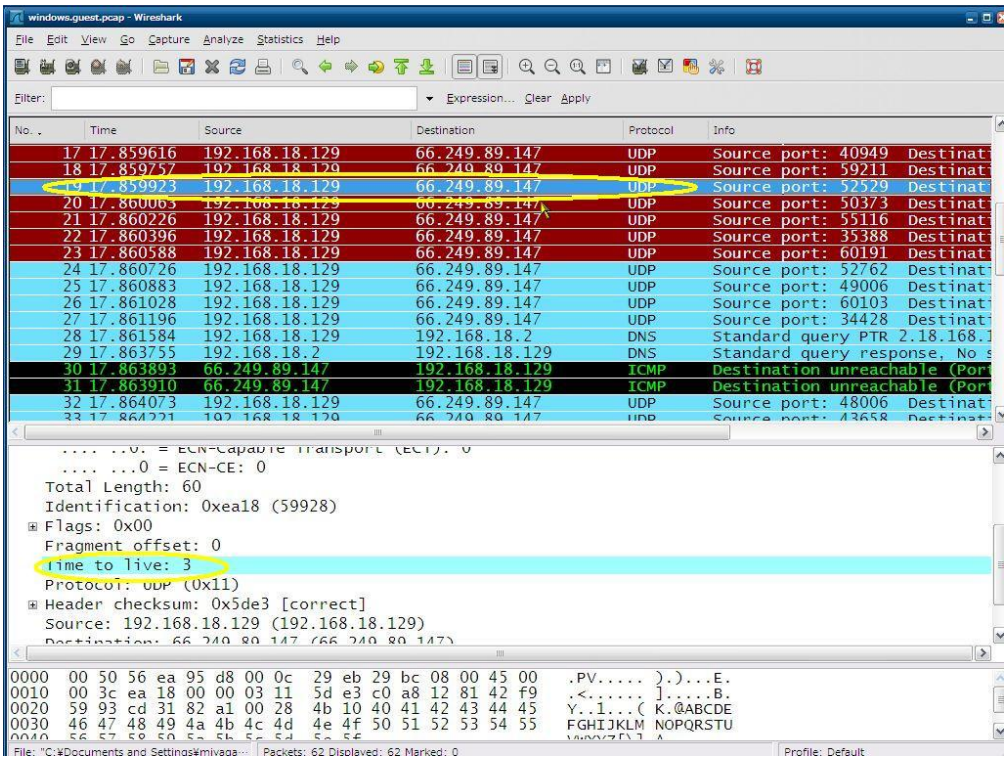
Trace complete.

Tracert.exe@host shows every hops correctly.

1.2.3. Captured packets of Traceroute flow with Wireshark

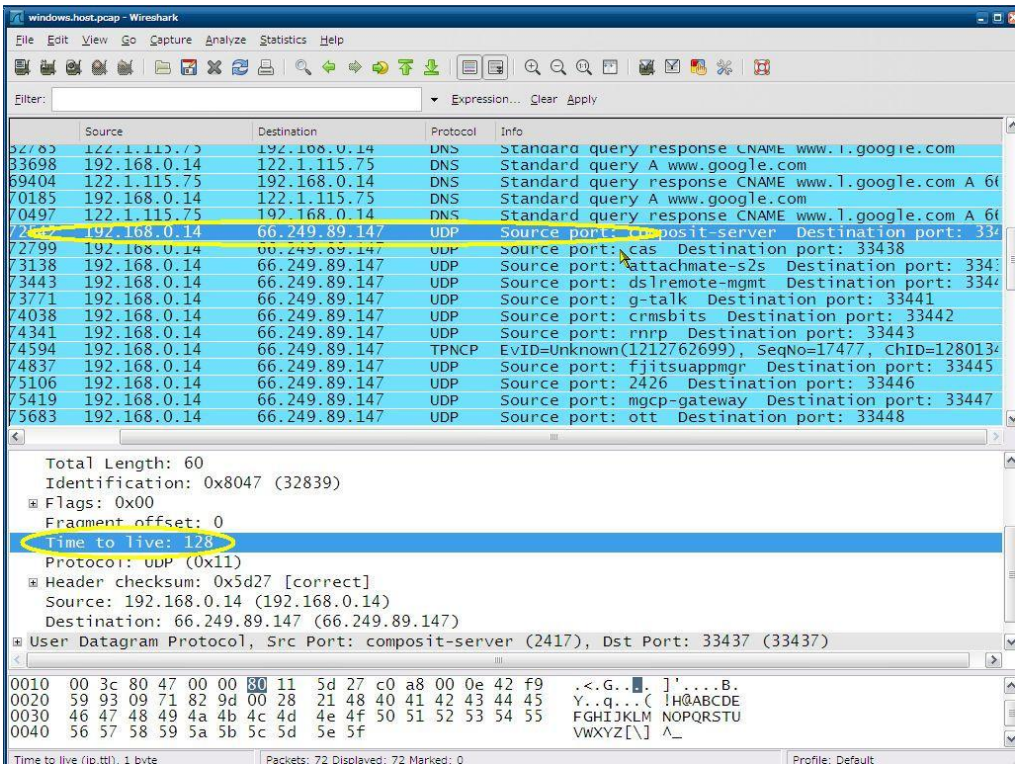
1.2.3.1. @Guest





Traceroute was sending UDP packets with TTL = 1,2,3….

1.2.3.2. @Host



Host OS was sending traceroute UDP packets received from Guest OS with TTL = 128!

1.3. TEST2: ping www.google.com from Guest

No	TEST	Result
1.	ping www.google.com . with no option	[root@localhost ~]# ping www.google.com PING www.l.google.com (66.249.89.104) 56(84) bytes of data.
2.	Set option " –t 1". All packets are sent with TTL == 1	[root@localhost ~]# ping **-t 1** www.google.com PING www.l.google.com (66.249.89.99) 56(84) bytes of data. From 192.168.18.2 icmp_seq=1 Time to live exceeded
3.	Set option " –t 2". All packets are sent with TTL == 1	[root@localhost ~]# ping **-t 2** www.google.com PING www.l.google.com (66.249.89.147) 56(84) bytes of data. **64 bytes from jp-in-f147.google.com (66.249.89.147): icmp_seq=1 ttl=128 time=4.75 ms**

Result of No3 is strange. This means host OS sent traceroute packet from Guest OS with large TTL.

2. Case 2

2.1. TEST Environments

3. Host OS : MacOSX 10.5.6

3.1. ipconfig -a

```
lo0: flags=8049<UP,LOOPBACK,RUNNING,MULTICAST> mtu 16384
        inet6 fe80::1%lo0 prefixlen 64 scopeid 0x1
        inet 127.0.0.1 netmask 0xff000000
        inet6 ::1 prefixlen 128
gif0: flags=8010<POINTOPOINT,MULTICAST> mtu 1280
stf0: flags=0<> mtu 1280
en0: flags=8863<UP,BROADCAST,SMART,RUNNING,SIMPLEX,MULTICAST> mtu 1500
        inet6 fe80::223:dfff:fe80:bd20%en0 prefixlen 64 scopeid 0x4
        inet 192.168.0.11 netmask 0xffffff00 broadcast 192.168.0.255
        ether 00:23:df:80:bd:20
        media: autoselect (1000baseT <full-duplex,flow-control>) status: active
        supported media: none autoselect 10baseT/UTP <half-duplex> 10baseT/UTP <full-duplex>
10baseT/UTP <full-duplex,flow-control> 10baseT/UTP <full-duplex,hw-loopback> 100baseTX <half-duplex>
100baseTX <full-duplex> 100baseTX <full-duplex,flow-control> 100baseTX <full-duplex,hw-loopback>
1000baseT <full-duplex> 1000baseT <full-duplex,flow-control> 1000baseT <full-duplex,hw-loopback>
en1: flags=8823<UP,BROADCAST,SMART,SIMPLEX,MULTICAST> mtu 1500
        ether 00:23:6c:80:f5:a3
        media: autoselect (<unknown type>) status: inactive
        supported media: autoselect
vmnet8: flags=8863<UP,BROADCAST,SMART,RUNNING,SIMPLEX,MULTICAST> mtu 1500
        inet 172.16.48.1 netmask 0xffffff00 broadcast 172.16.48.255
        ether 00:50:56:c0:00:08
vmnet1: flags=8863<UP,BROADCAST,SMART,RUNNING,SIMPLEX,MULTICAST> mtu 1500
        inet 172.16.121.1 netmask 0xffffff00 broadcast 172.16.121.255
        ether 00:50:56:c0:00:01
```

4. Guest OS : Fedora 8 x86_64 on **VMware Fusion 2.0.3**

4.1. Same of case 1

4.2. Network Setting : NAT

4.3. uname –a and ifconfig eth1,

```
[root@localhost ~]# uname -a
Linux localhost.localdomain 2.6.26.8-57.fc8 #1 SMP Thu Dec 18 18:59:49 EST 2008 x86_64 x86_64 x86_64 GNU/Linux
[root@localhost ~]# ifconfig eth1
eth1      Link encap:Ethernet  HWaddr 00:0C:29:30:5F:67
          inet addr:172.16.48.129  Bcast:172.16.48.255  Mask:255.255.255.0
          inet6 addr: fe80::20c:29ff:fe30:5f67/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:2799 errors:0 dropped:0 overruns:0 frame:0
          TX packets:1670 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:3412212 (3.2 MiB)  TX bytes:104036 (101.5 KiB)
          Interrupt:17 Base address:0x1080
```

2.2. TEST1: traceroute www.google.com from Host and Guest

2.2.1. Traceroute from Guest

```
traceroute to www.google.com (66.249.89.99), 30 hops max, 60 byte packets
 1  172.16.48.2 (172.16.48.2)  0.347 ms  0.171 ms  0.239 ms
 2  * * *
 3  * * *
 4  * * *
 5  * * *
 6  * * *
 7  * * *
 8  * * *
 9  * * *
10  * * *
11  * * *
12  * * *
13  * * *
14  * * *
15  * * *
16  * * *
17  * * *
18  * * *
19  * * *
20  * * *
21  * * *
22  * * *
23  * * *
24  * * *
25  * * *
26  * * *
27  * * *
28  * * *
29  * * *
30  * * *
```

No time exceed ICMP packets were received at Guest OS. Therefore, traceroute wait timer was expired. On the other hand, in the result of captured packets at Host OS (see also 2.2.3.1), Host OS received time exceed ICMP packets from every hops. But, these packets are not forwarded to Guest OS. Why??

2.2.2. Traceroute from Host

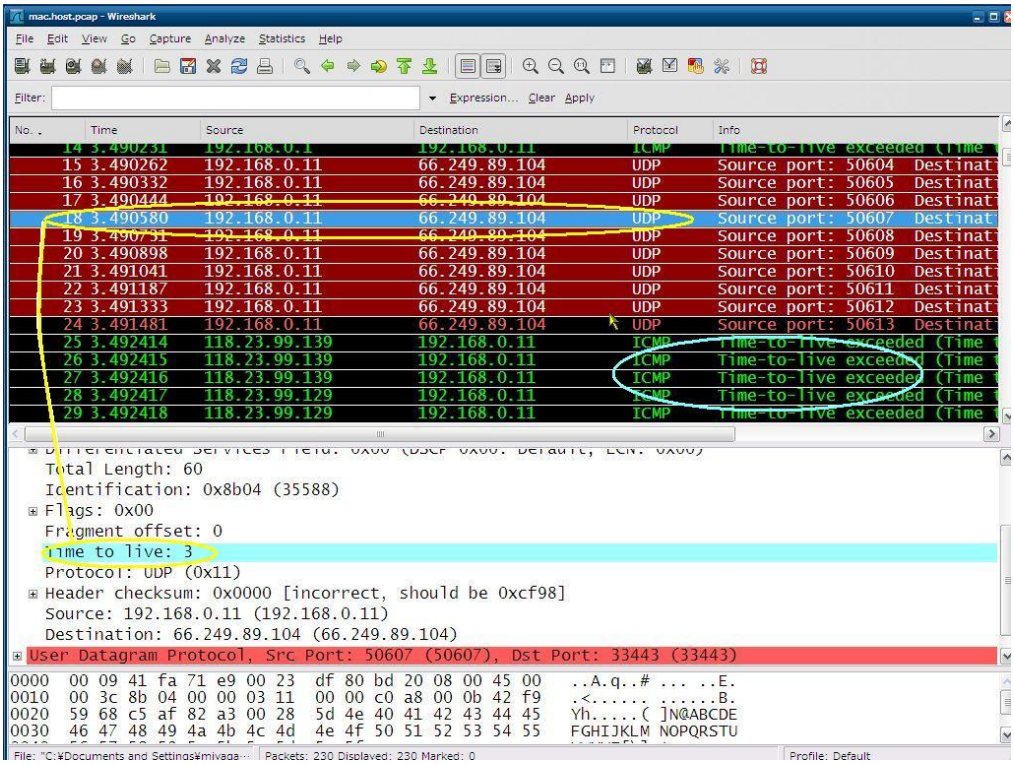
```
mbu05:~user$ traceroute www.google.com
1  192.168.0.1 (192.168.0.1)  0.662 ms  0.196 ms  0.252 ms
 2  118.23.99.139 (118.23.99.139)  1.862 ms  2.373 ms  2.015 ms
```

```
 3  118.23.99.158 (118.23.99.158)  2.314 ms  2.158 ms  2.110 ms
 4  221.184.12.233 (221.184.12.233)  5.582 ms  3.785 ms  3.698 ms
 5  60.37.11.41 (60.37.11.41)  2.300 ms  1.833 ms  2.043 ms
 6  210.254.188.141 (210.254.188.141)  1.787 ms  2.719 ms  2.595 ms
 7  210.254.188.146 (210.254.188.146)  2.996 ms  2.764 ms  3.321 ms
 8  122.28.104.106 (122.28.104.106)  2.774 ms  3.007 ms  2.870 ms
 9  210.163.230.238 (210.163.230.238)  3.131 ms  3.380 ms  3.235 ms
10  xe-2-1.a17.tokyjp01.jp.ra.gin.ntt.net (61.213.169.70)  3.129 ms  2.597 ms  3.043 ms
11  xe-1-3.a17.tokyjp01.jp.ra.gin.ntt.net (61.213.169.174)  3.756 ms  3.495 ms  3.059 ms
12  209.85.241.94 (209.85.241.94)  3.632 ms  3.923 ms  3.537 ms
13  72.14.236.126 (72.14.236.126)  4.171 ms  15.918 ms  4.253 ms
14  jp-in-f147.google.com (66.249.89.147)  4.272 ms  4.039 ms  4.371 ms
```

Traceroute from host OS is no problem.

2.2.3. Captured packets of Traceroute flow with Wireshark

2.2.3.1. @Host



- TTL fields of UDP packets from Gest OS traceroute (yellow) are not changed unlike windows host case.
- Host OS received time exceed ICMP packets (blue circle), but did not forward them to guest

2.3. TEST2: ping www.google.com from Guest

No	TEST	Result
1.	ping www.google.com . with no option	[root@localhost ~]# ping www.google.com PING www.l.google.com (66.249.89.99) 56(84) bytes of data. 64 bytes from jp-in-f99.google.com (66.249.89.99): icmp_seq=1 ttl=128 time=6.07 ms
2.	Set option " –t 1". All packets are sent with TTL == 1	[root@localhost ~]# ping **-t 1** www.google.com PING www.l.google.com (66.249.89.99) 56(84) bytes of data. From 172.16.48.2 icmp_seq=1 Time to live exceeded
3.	Set option " –t 2". All packets	[root@localhost ~]# ping **-t 2** www.google.com

	are sent with TTL == 1	PING www.l.google.com (66.249.89.99) 56(84) bytes of data. --- www.l.google.com ping statistics --- 3 packets transmitted, 0 received, **100% packet loss**, time 2000ms

3. Questions

1. Why does **Windows host** change the TTL field of packets received from its Guest OS?
2. Why doesn't MacOS host and natd forward time exceed ICMP packets to Guest?
3. How to fix theses problem