Tunneling: topology and examples

D. Berardi, F. Callegati, A. Melis

In una rete Ethernet tradizionale, uno switch raccoglie e connette tutti i dispositivi all'interno di una rete locale (LAN). Quando i dispositivi sono collegati a uno switch, questi dispositivi possono comunicare tra loro all'interno della stessa rete fisica. Lo switch può essere utilizzato per connettere dispositivi all'interno della stessa rete IP

Topologia

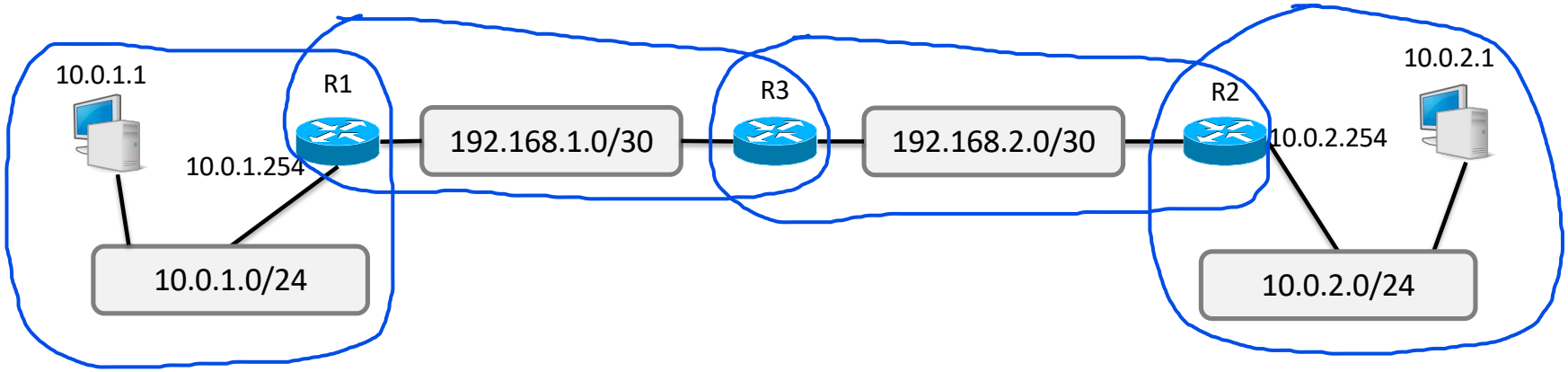


Tabelle di routing

R3

Destination	Gateway	Genmask	Flags	Metric	Ref	Use	Iface
10.0.1.0	192.168.1.1	255.255.255.0	UG	0	0	0	veth13
10.0.2.0	192.168.2.1	255.255.255.0	UG	0	0	0	veth23
192.168.1.0	0.0.0.0	255.255.255.252	U	0	0	0	veth13
192.168.2.0	0.0.0.0	255.255.255.252	U	0	0	0	veth23

R1

Destination	Gateway	Genmask	Flags	Metric	Ref	Use	Iface
10.0.1.0	0.0.0.0	255.255.255.0	U	0	0	0	eth-H1
10.0.2.0	192.168.1.2	255.255.255.0	UG	0	0	0	veth1
192.168.1.0	0.0.0.0	255.255.255.252	U	0	0	0	veth1
192.168.2.0	192.168.1.2	255.255.255.252	UG	0	0	0	veth1

R2

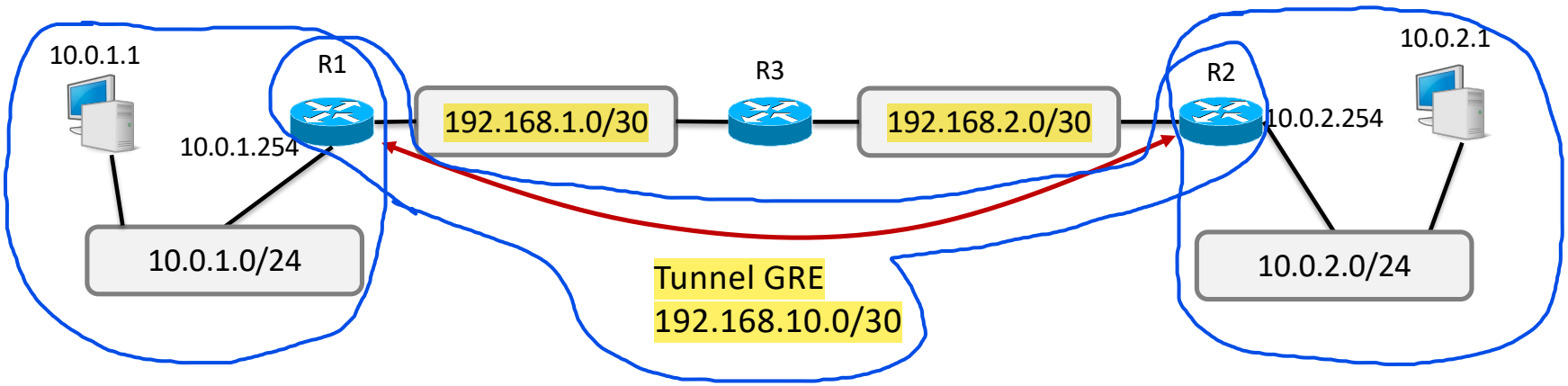
Destination	Gateway	Genmask	Flags	Metric	Ref	Use	Iface
10.0.1.0	192.168.2.2	255.255.255.0	UG	0	0	0	veth2
10.0.2.0	0.0.0.0	255.255.255.0	U	0	0	0	eth-H2
192.168.1.0	192.168.2.2	255.255.255.252	UG	0	0	0	veth2
192.168.2.0	0.0.0.0	255.255.255.252	U	0	0	0	veth2

H1 = (10.0.1.1)

Raggiungibilità

```
H1> ping -c 3 10.0.2.1

PING 10.0.2.1 (10.0.2.1) 56(84) bytes of data.
64 bytes from 10.0.2.1: icmp_seq=1 ttl=61 time=2.88 ms
64 bytes from 10.0.2.1: icmp_seq=2 ttl=61 time=0.097 ms
64 bytes from 10.0.2.1: icmp_seq=3 ttl=61 time=0.101 ms

--- 10.0.2.1 ping statistics ---
3 packets transmitted, 3 received, 0% packet loss, time 2018ms
rtt min/avg/max/mdev = 0.097/1.026/2.880/1.310 ms

H1> traceroute -n 10.0.2.1

traceroute to 10.0.2.1 (10.0.2.1), 30 hops max, 60 byte packets
 1  10.0.1.254  0.087 ms  0.011 ms  0.007 ms (Da 10.0.1.1 vado a R1)
 2  192.168.1.2  0.039 ms  0.013 ms  0.010 ms (Da R1 ad R3)
 3  192.168.2.1  0.019 ms  0.012 ms  0.011 ms (Da R3 a R2)
 4  10.0.2.1  0.021 ms  0.016 ms  0.017 ms (Da R2 a 10.0.2.1)
```

GRE tunneling



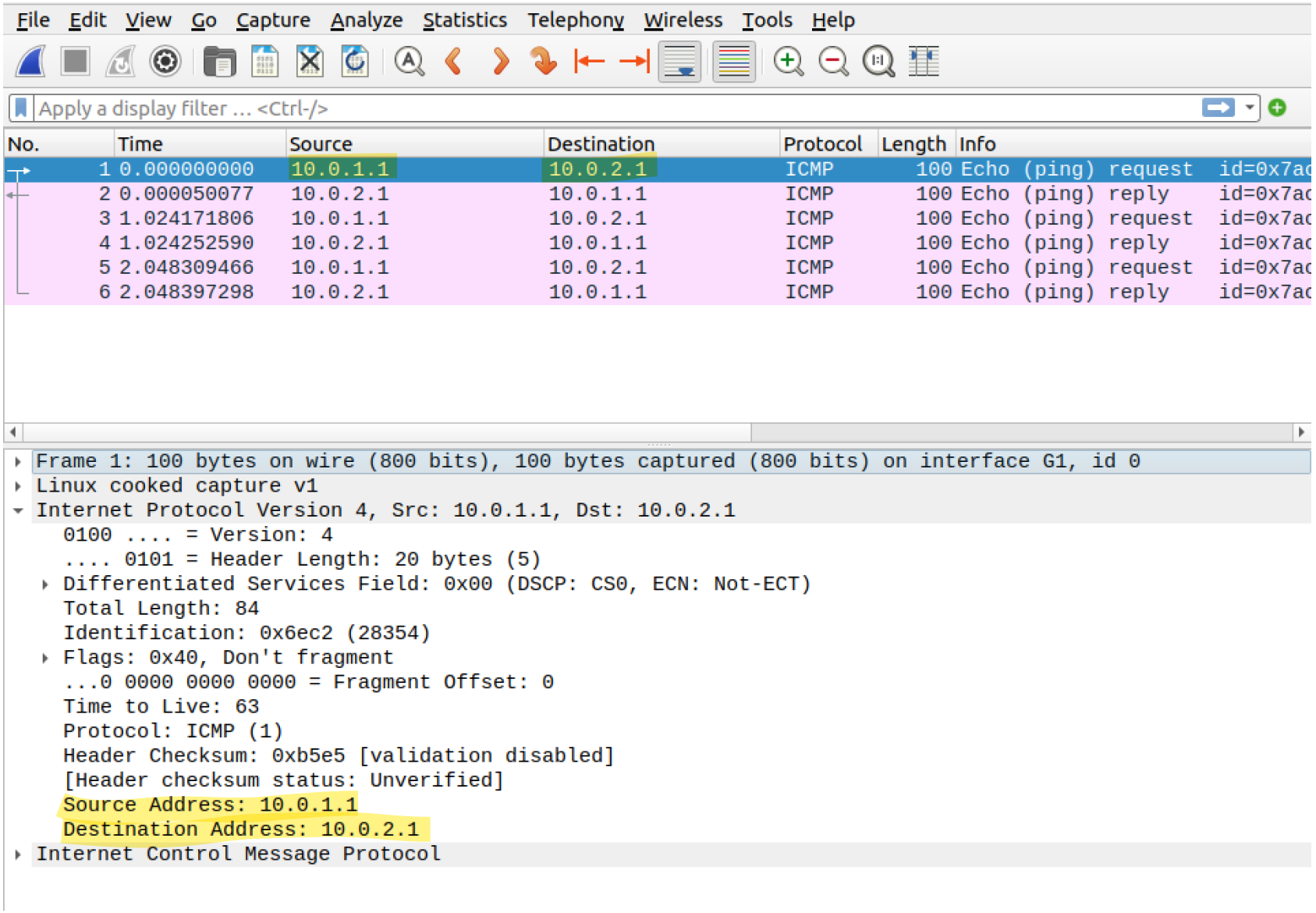
H1 = (10.0.1.1)

Raggiungibilità con GRE

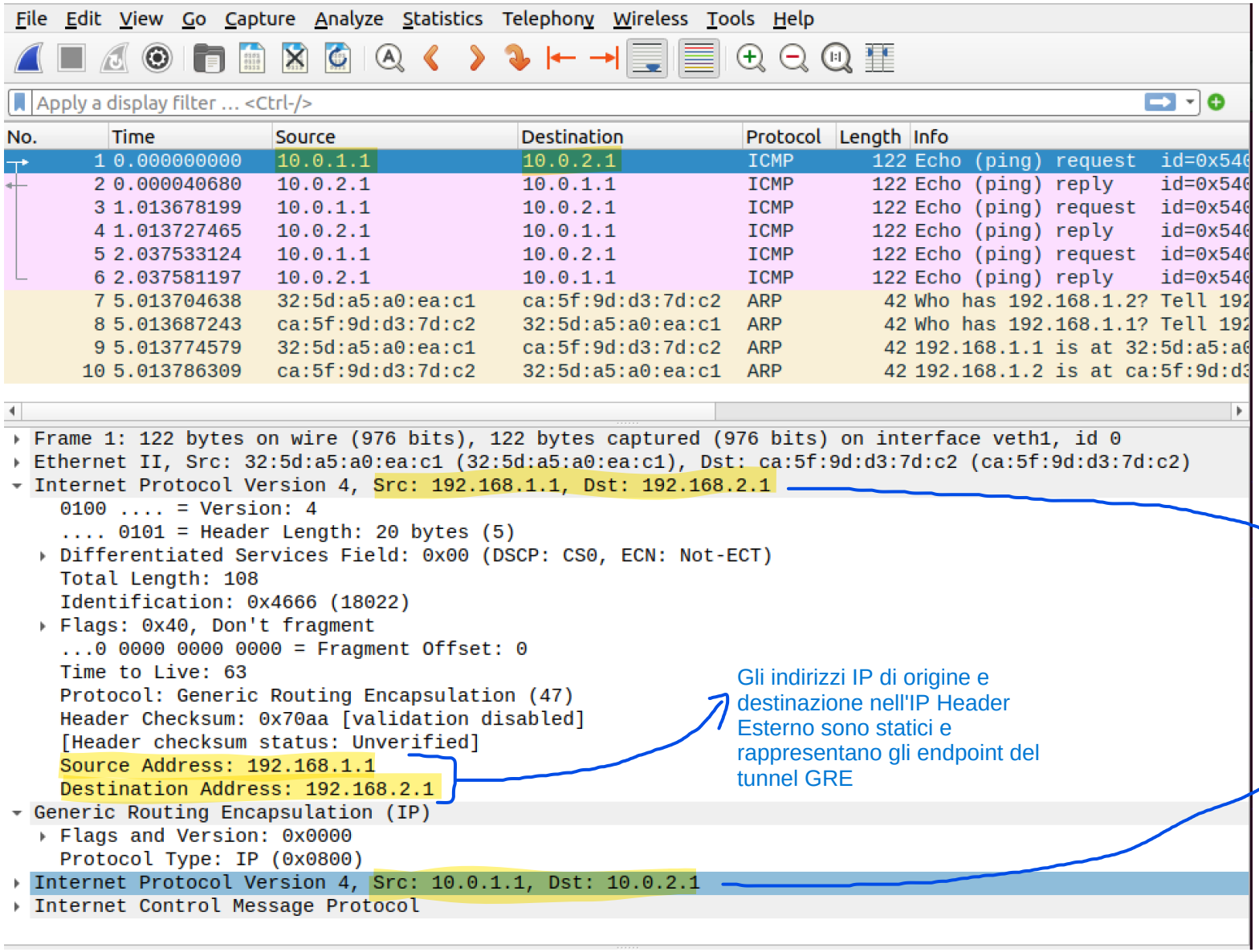
```
H1> ping -c 3 10.0.2.1
PING 10.0.2.1 (10.0.2.1) 56(84) bytes of data.
64 bytes from 10.0.2.1: icmp_seq=1 ttl=62 time=0.153 ms
64 bytes from 10.0.2.1: icmp_seq=2 ttl=62 time=0.150 ms
64 bytes from 10.0.2.1: icmp_seq=3 ttl=62 time=0.141 ms

--- 10.0.2.1 ping statistics ---
3 packets transmitted, 3 received, 0% packet loss, time 2025ms
rtt min/avg/max/mdev = 0.141/0.148/0.153/0.005 ms

H1> traceroute -n 10.0.2.1
traceroute to 10.0.2.1 (10.0.2.1), 30 hops max, 60 byte packets
 1  10.0.1.254  0.066 ms  0.011 ms  0.006 ms (Da 10.0.1.1 a R1)
 2  192.168.10.2  0.073 ms  0.023 ms  0.017 ms (Da R1 a R2, tramite GRE)
 3  10.0.2.1  0.028 ms  0.020 ms  0.021 ms (Da R2 a 10.0.2.1)
```

significa che rispetto a prima, senza GRE, ho fatto un salto in meno

SENZA GRE


File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help
Apply a display filter … <Ctrl-/>
No. Time Source Destination Protocol Length Info
1 0.000000000 10.0.1.1 10.0.2.1 ICMP 100 Echo (ping) request id=0x7ac
2 0.000050077 10.0.2.1 10.0.1.1 ICMP 100 Echo (ping) reply id=0x7ac
3 1.024171806 10.0.1.1 10.0.2.1 ICMP 100 Echo (ping) request id=0x7ac
4 1.024252590 10.0.2.1 10.0.1.1 ICMP 100 Echo (ping) reply id=0x7ac
5 2.048309466 10.0.1.1 10.0.2.1 ICMP 100 Echo (ping) request id=0x7ac
6 2.048397298 10.0.2.1 10.0.1.1 ICMP 100 Echo (ping) reply id=0x7ac
Frame 1: 100 bytes on wire (800 bits), 100 bytes captured (800 bits) on interface G1, id 0
Linux cooked capture v1
Internet Protocol Version 4, Src: 10.0.1.1, Dst: 10.0.2.1
0100 .... = Version: 4
.... 0101 = Header Length: 20 bytes (5)
Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)
Total Length: 84
Identification: 0x6ec2 (28354)
Flags: 0x40, Don't fragment
...0 0000 0000 0000 = Fragment Offset: 0
Time to Live: 63
Protocol: ICMP (1)
Header Checksum: 0xb5e5 [validation disabled]
[Header checksum status: Unverified]
Source Address: 10.0.1.1
Destination Address: 10.0.2.1
Internet Control Message Protocol

CON GRE



IP Header Esterno: Contiene gli indirizzi IP origine e destinazione dei router che stanno gestendo il tunnel GRE. Questi indirizzi sono utilizzati per il trasporto del pacchetto attraverso la rete pubblica (o il percorso di rete che ospita il tunnel).

Gli indirizzi IP di origine e destinazione nell'IP Header Esterno sono statici e rappresentano gli endpoint del tunnel GRE

IP Header Interno: Contiene gli indirizzi IP origine e destinazione dei dispositivi che comunicano. Questi indirizzi sono quelli originali del pacchetto prima che fosse incapsulato nel tunnel GRE.

Le due rotte aggiunte o modificate nella tabella di routing di R1 dopo la configurazione del tunnel GRE sono utilizzate per indirizzare il traffico attraverso il tunnel, e il software GRE si occupa dell'incapsulamento (quando il pacchetto viene inviato nel tunnel) e del decapsulamento (quando il pacchetto arriva all'endpoint del tunnel). Queste operazioni permettono di trasportare pacchetti attraverso una rete pubblica, simulando una rete privata tra i router, e mantenendo l'indirizzamento originale per i pacchetti.

R1 dopo il GRE

Questo indica che, dopo l'implementazione del tunnel GRE, il traffico verso 10.0.2.0 viene instradato attraverso un router remoto (192.168.10.2), il quale gestisce il tunnel GRE.

```
R1> route -n
Kernel IP routing table
Destination     Gateway         Genmask         Flags Metric Ref    Use Iface
10.0.1.0        0.0.0.0         255.255.255.0   U     0      0        0 eth-H1
10.0.2.0        192.168.10.2    255.255.255.0   UG    0      0        0 G1
192.168.1.0     0.0.0.0         255.255.255.252 U     0      0        0 veth1
192.168.2.0     192.168.1.2     255.255.255.252 UG    0      0        0 veth1
192.168.10.0    0.0.0.0         255.255.255.252 U     0      0        0 G1
```

R1 prima del GRE

```
R1> route -n
Kernel IP routing table
Destination     Gateway         Genmask         Flags Metric Ref    Use Iface
10.0.1.0        0.0.0.0         255.255.255.0   U     0      0        0 eth-H1
10.0.2.0        192.168.1.2     255.255.255.0   UG    0      0        0 veth1
192.168.1.0     0.0.0.0         255.255.255.252 U     0      0        0 veth1
192.168.2.0     192.168.1.2     255.255.255.252 UG    0      0        0 veth1
```

Le due rotte aggiunte o modificate nella tabella di routing di R1 dopo la configurazione del tunnel GRE sono utilizzate per indirizzare il traffico attraverso il tunnel, e il software GRE si occupa dell'incapsulamento (quando il pacchetto viene inviato nel tunnel) e del decapsulamento (quando il pacchetto arriva all'endpoint del tunnel). Queste operazioni permettono di trasportare pacchetti attraverso una rete pubblica, simulando una rete privata tra i router, e mantenendo l'indirizzamento originale per i pacchetti.

R2 dopo il GRE

Questo indica che, dopo l'implementazione del tunnel GRE, il traffico verso 10.0.1.0 viene instradato attraverso un router remoto (192.168.10.1), il quale gestisce il tunnel GRE.

```
R2> route -n
Kernel IP routing table
Destination     Gateway         Genmask         Flags Metric Ref    Use Iface
10.0.1.0        192.168.10.1    255.255.255.0   UG    0      0        0 G1
10.0.2.0        0.0.0.0         255.255.255.0   U     0      0        0 eth-H2
192.168.1.0     192.168.2.2     255.255.255.252 UG    0      0        0 veth2
192.168.2.0     0.0.0.0         255.255.255.252 U     0      0        0 veth2
192.168.10.0    0.0.0.0         255.255.255.252 U     0      0        0 G1
```

R2 Prima del GRE

```
R2> route -n
Kernel IP routing table
Destination     Gateway         Genmask         Flags Metric Ref    Use Iface
10.0.1.0        192.168.2.2     255.255.255.0   UG    0      0        0 veth2
10.0.2.0        0.0.0.0         255.255.255.0   U     0      0        0 eth-H2
192.168.1.0     192.168.2.2     255.255.255.252 UG    0      0        0 veth2
192.168.2.0     0.0.0.0         255.255.255.252 U     0      0        0 veth2
```

R3 prima e dopo (Non cambia nulla)

Prima del GRE

```
R3> route -n
Kernel IP routing table
Destination     Gateway         Genmask         Flags Metric Ref    Use Iface
10.0.1.0        192.168.1.1     255.255.255.0   UG    0      0        0 veth13
10.0.2.0        192.168.2.1     255.255.255.0   UG    0      0        0 veth23
192.168.1.0     0.0.0.0         255.255.255.252 U     0      0        0 veth13
192.168.2.0     0.0.0.0         255.255.255.252 U     0      0        0 veth23
```

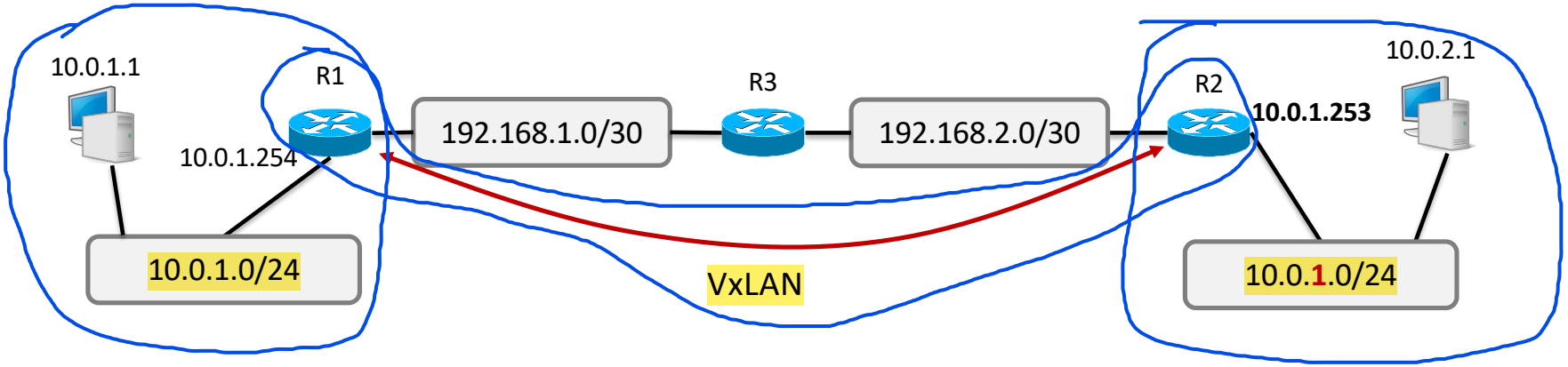
Dopo il GRE

```
R3> route -n
Kernel IP routing table
Destination     Gateway         Genmask         Flags Metric Ref    Use Iface
10.0.1.0        192.168.1.1     255.255.255.0   UG    0      0        0 veth13
10.0.2.0        192.168.2.1     255.255.255.0   UG    0      0        0 veth23
192.168.1.0     0.0.0.0         255.255.255.252 U     0      0        0 veth13
192.168.2.0     0.0.0.0         255.255.255.252 U     0      0        0 veth23
```

Far parte della stessa rete IP significa che due o più dispositivi condividono la stessa sottorete IP, consentendo loro di comunicare direttamente senza la necessità di un router.

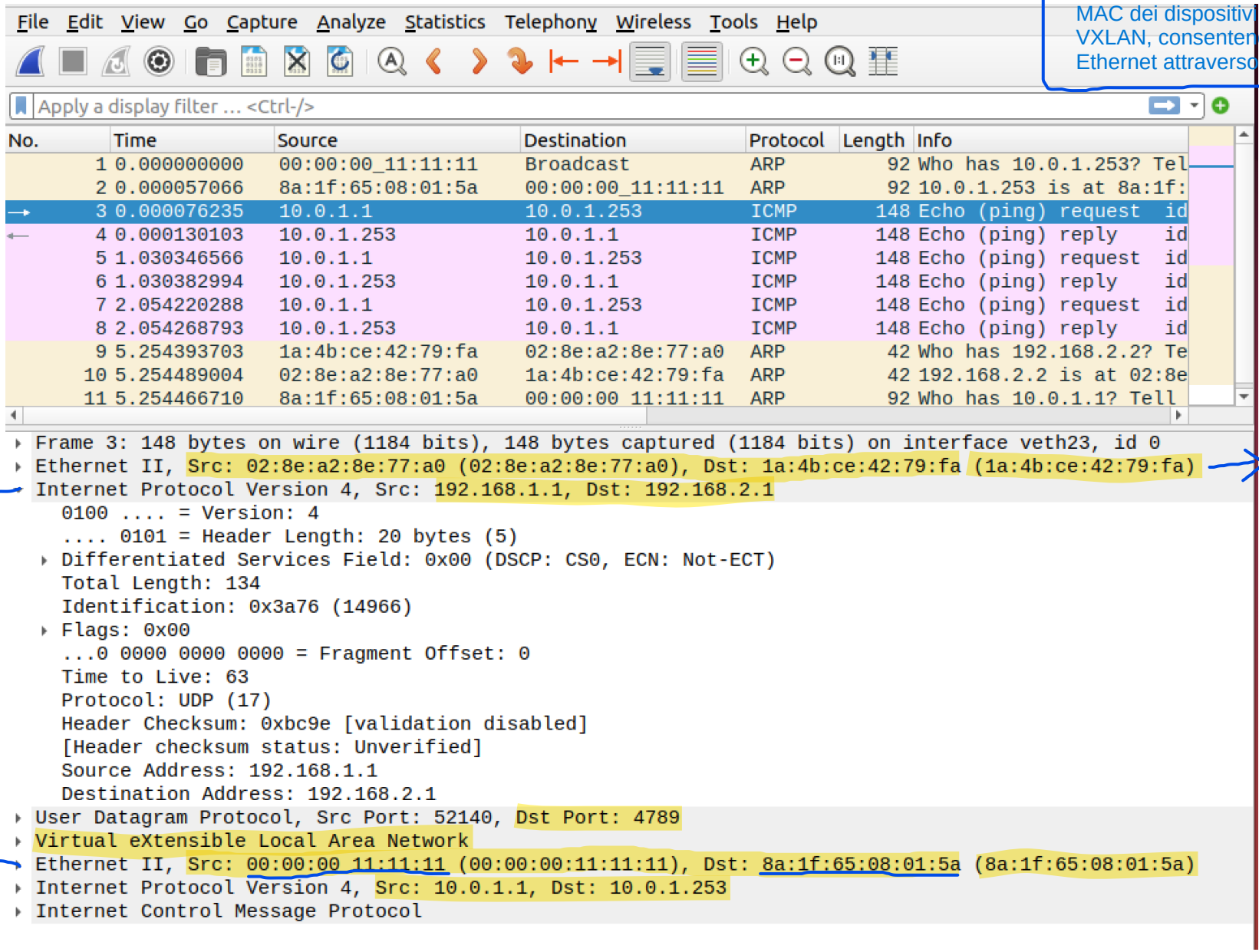
VxLAN tunneling

In una configurazione VXLAN, i router possono essere virtualmente considerati come appartenenti alla stessa rete locale (Layer 2), ma non alla stessa rete IP (Layer 3).



- Indirizzo IP di sorgente: Indica l'indirizzo IP del VTEP di origine, ossia il dispositivo che incapsula il traffico Ethernet in pacchetti VXLAN per il trasporto attraverso la rete IP sottostante.
- Indirizzo IP di destinazione: Indica l'indirizzo IP del VTEP di destinazione, ossia il dispositivo che riceve i pacchetti VXLAN, li decapsula e inoltra il traffico Ethernet alla destinazione finale.

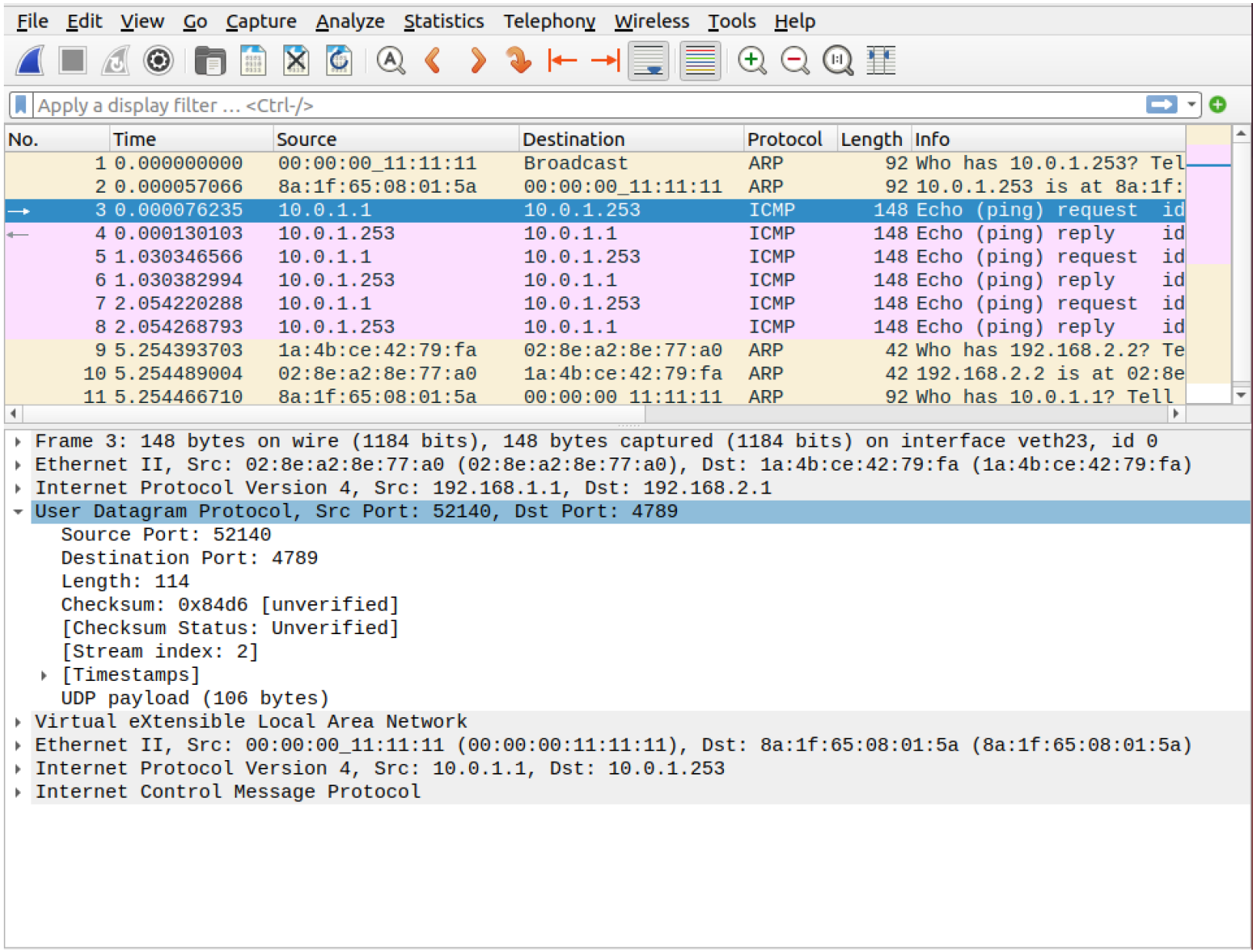
- Inner Ethernet Header: Rappresenta gli indirizzi MAC dei dispositivi finali all'interno della rete locale, permettendo la comunicazione tra dispositivi sulla stessa VLAN.
- Outer Ethernet Header: Rappresenta gli indirizzi MAC dei dispositivi di rete che gestiscono il tunnel VXLAN, consentendo il trasporto del traffico Ethernet attraverso una rete IP.

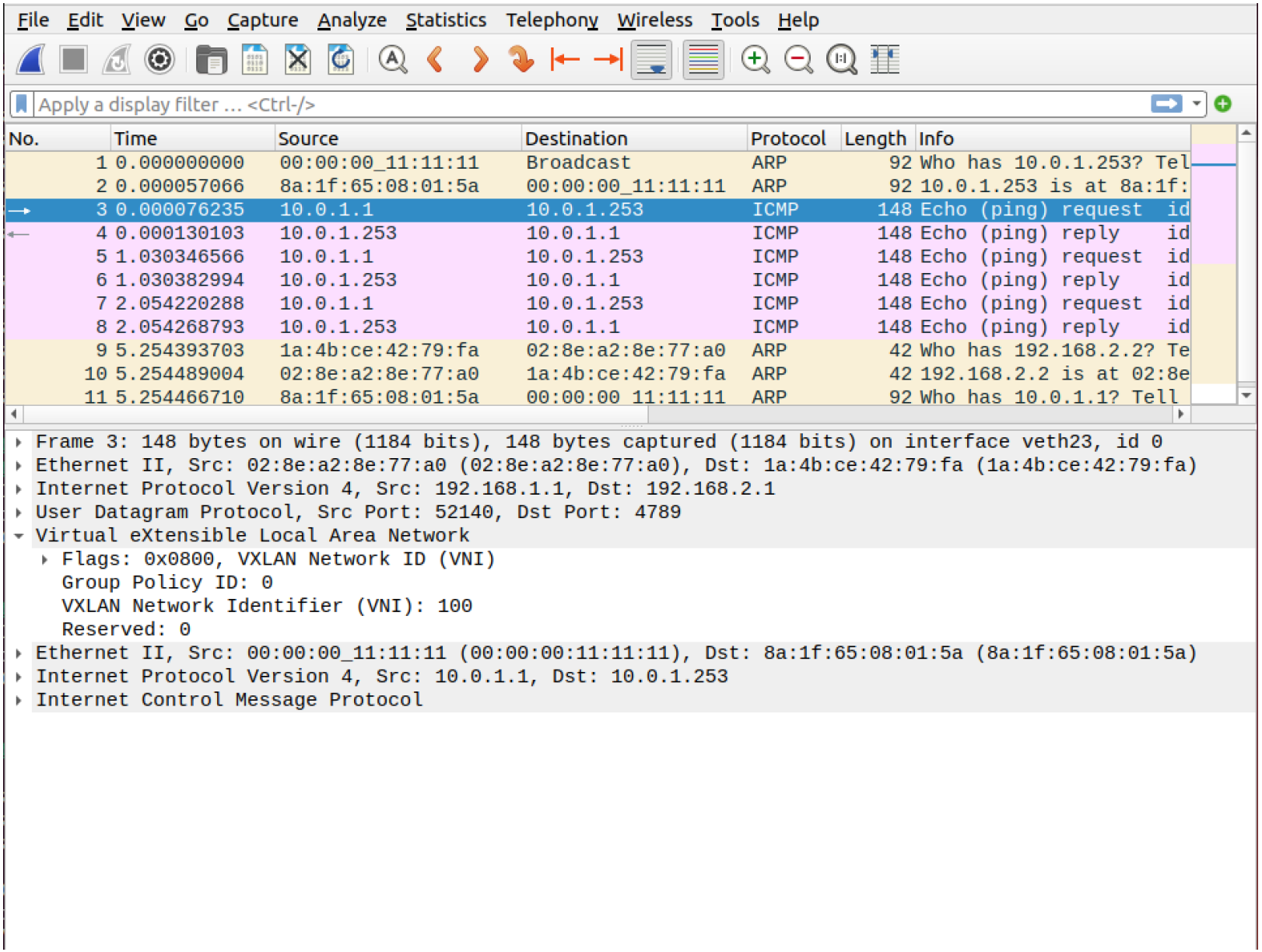


Contiene gli indirizzi MAC di origine e destinazione dei dispositivi di rete che gestiscono il tunnel VXLAN (i VTEP, ovvero i punti finali del tunnel):

- Destinazione MAC: Indirizzo MAC del VTEP di destinazione, ossia il dispositivo che riceve il pacchetto incapsulato.
- Origine MAC: Indirizzo MAC del VTEP di origine, ossia il dispositivo che invia il pacchetto incapsulato.

Pacchetto Ethernet Originale dove Sorgente e Destinazione sono gli Indirizzi MAC originali


File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help
Apply a display filter ... <Ctrl-/>
No. Time Source Destination Protocol Length Info
1 0.000000000 00:00:00_11:11:11 Broadcast ARP 92 Who has 10.0.1.253? Tel
2 0.000057066 8a:1f:65:08:01:5a 00:00:00_11:11:11 ARP 92 10.0.1.253 is at 8a:1f:
3 0.000076235 10.0.1.1 10.0.1.253 ICMP 148 Echo (ping) request id
4 0.000130103 10.0.1.253 10.0.1.1 ICMP 148 Echo (ping) reply id
5 1.030346566 10.0.1.1 10.0.1.253 ICMP 148 Echo (ping) request id
6 1.030382994 10.0.1.253 10.0.1.1 ICMP 148 Echo (ping) reply id
7 2.054220288 10.0.1.1 10.0.1.253 ICMP 148 Echo (ping) request id
8 2.054268793 10.0.1.253 10.0.1.1 ICMP 148 Echo (ping) reply id
9 5.254393703 1a:4b:ce:42:79:fa 02:8e:a2:8e:77:a0 ARP 42 Who has 192.168.2.2? Te
10 5.254489004 02:8e:a2:8e:77:a0 1a:4b:ce:42:79:fa ARP 42 192.168.2.2 is at 02:8e
11 5.254466710 8a:1f:65:08:01:5a 00:00:00 11:11:11 ARP 92 Who has 10.0.1.1? Tell
Frame 3: 148 bytes on wire (1184 bits), 148 bytes captured (1184 bits) on interface veth23, id 0
Ethernet II, Src: 02:8e:a2:8e:77:a0 (02:8e:a2:8e:77:a0), Dst: 1a:4b:ce:42:79:fa (1a:4b:ce:42:79:fa)
Internet Protocol Version 4, Src: 192.168.1.1, Dst: 192.168.2.1
User Datagram Protocol, Src Port: 52140, Dst Port: 4789
Source Port: 52140
Destination Port: 4789
Length: 114
Checksum: 0x84d6 [unverified]
[Checksum Status: Unverified]
[Stream index: 2]
[Timestamps]
UDP payload (106 bytes)
Virtual eXtensible Local Area Network
Ethernet II, Src: 00:00:00_11:11:11 (00:00:00:11:11:11), Dst: 8a:1f:65:08:01:5a (8a:1f:65:08:01:5a)
Internet Protocol Version 4, Src: 10.0.1.1, Dst: 10.0.1.253
Internet Control Message Protocol


File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help
Apply a display filter ... <Ctrl-/>
No. Time Source Destination Protocol Length Info
1 0.000000000 00:00:00_11:11:11 Broadcast ARP 92 Who has 10.0.1.253? Tel
2 0.000057066 8a:1f:65:08:01:5a 00:00:00_11:11:11 ARP 92 10.0.1.253 is at 8a:1f:
3 0.000076235 10.0.1.1 10.0.1.253 ICMP 148 Echo (ping) request id
4 0.000130103 10.0.1.253 10.0.1.1 ICMP 148 Echo (ping) reply id
5 1.030346566 10.0.1.1 10.0.1.253 ICMP 148 Echo (ping) request id
6 1.030382994 10.0.1.253 10.0.1.1 ICMP 148 Echo (ping) reply id
7 2.054220288 10.0.1.1 10.0.1.253 ICMP 148 Echo (ping) request id
8 2.054268793 10.0.1.253 10.0.1.1 ICMP 148 Echo (ping) reply id
9 5.254393703 1a:4b:ce:42:79:fa 02:8e:a2:8e:77:a0 ARP 42 Who has 192.168.2.2? Te
10 5.254489004 02:8e:a2:8e:77:a0 1a:4b:ce:42:79:fa ARP 42 192.168.2.2 is at 02:8e
11 5.254466710 8a:1f:65:08:01:5a 00:00:00 11:11:11 ARP 92 Who has 10.0.1.1? Tell
Frame 3: 148 bytes on wire (1184 bits), 148 bytes captured (1184 bits) on interface veth23, id 0
Ethernet II, Src: 02:8e:a2:8e:77:a0 (02:8e:a2:8e:77:a0), Dst: 1a:4b:ce:42:79:fa (1a:4b:ce:42:79:fa)
Internet Protocol Version 4, Src: 192.168.1.1, Dst: 192.168.2.1
User Datagram Protocol, Src Port: 52140, Dst Port: 4789
Virtual eXtensible Local Area Network
Flags: 0x0800, VXLAN Network ID (VNI)
Group Policy ID: 0
VXLAN Network Identifier (VNI): 100
Reserved: 0
Ethernet II, Src: 00:00:00_11:11:11 (00:00:00:11:11:11), Dst: 8a:1f:65:08:01:5a (8a:1f:65:08:01:5a)
Internet Protocol Version 4, Src: 10.0.1.1, Dst: 10.0.1.253
Internet Control Message Protocol