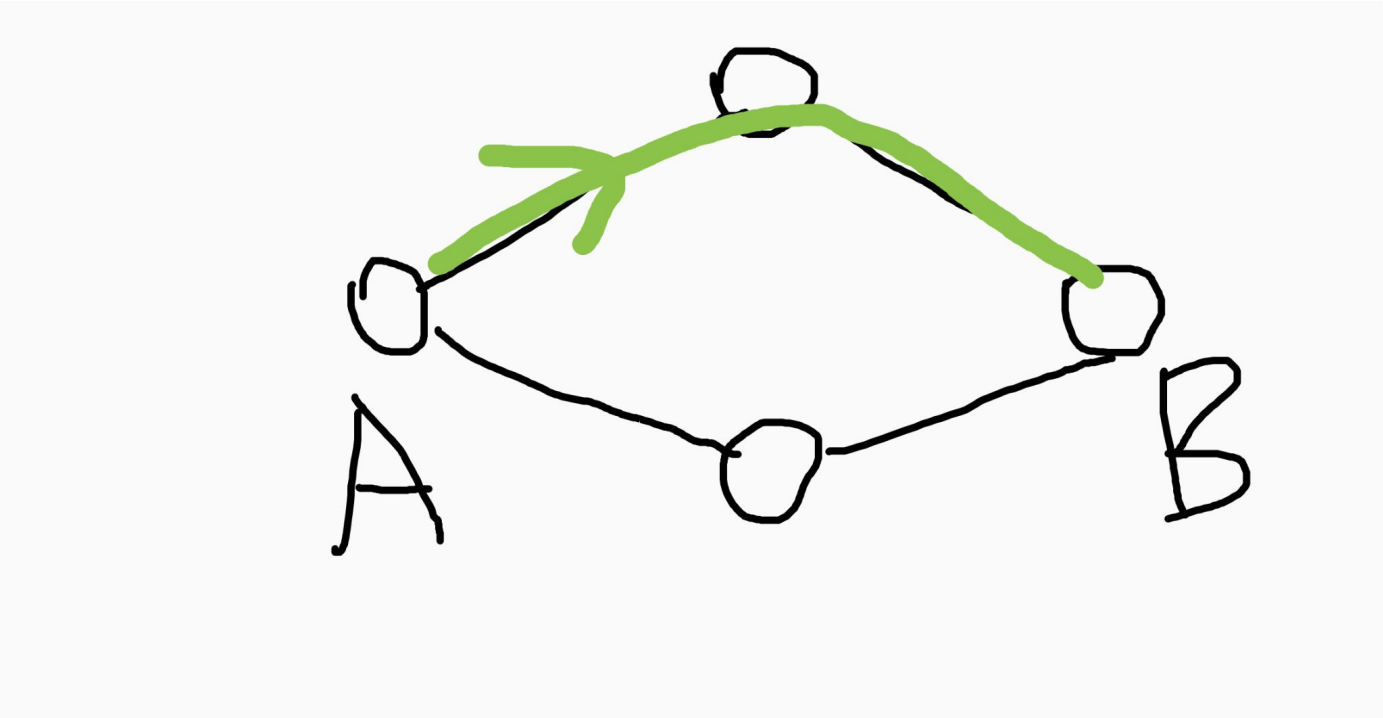
QRATOR LABS

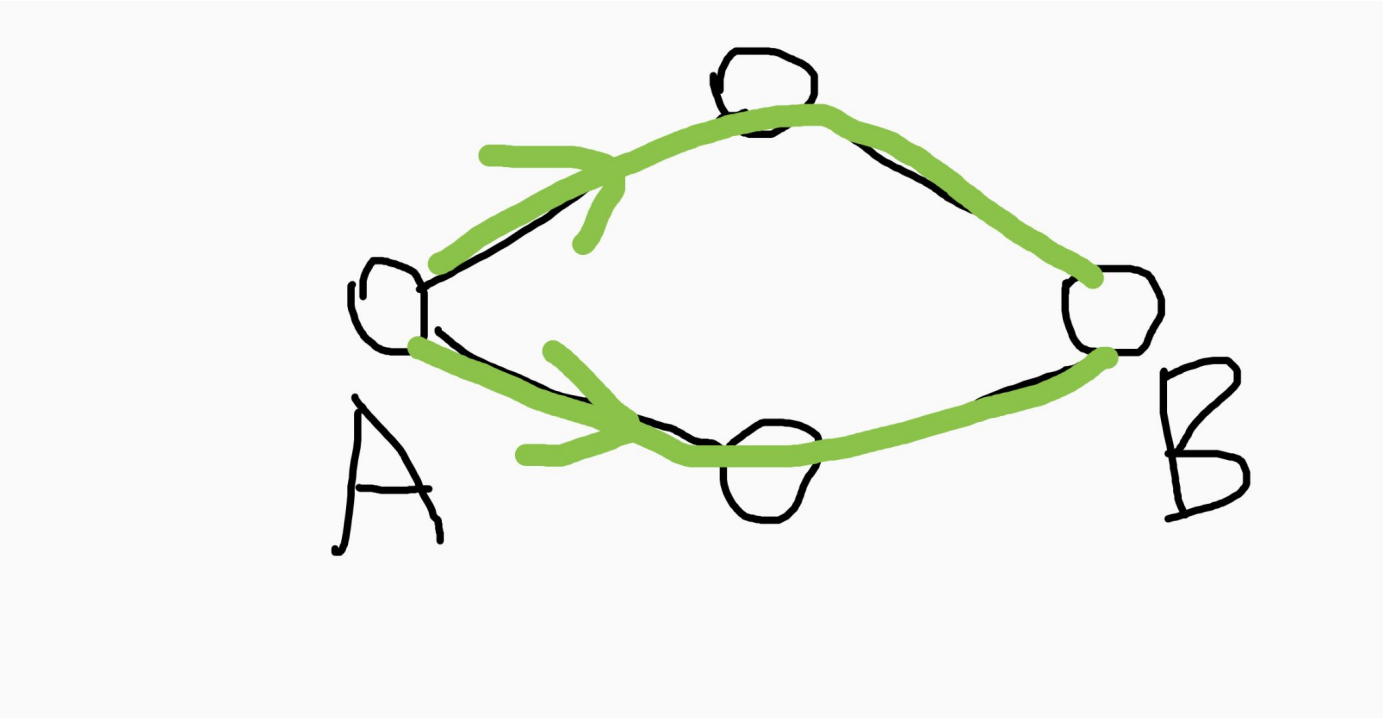
Multipath diagnostics

Alexander Zubkov
green@qrator.net
2026-01-22 CSNOG 2026

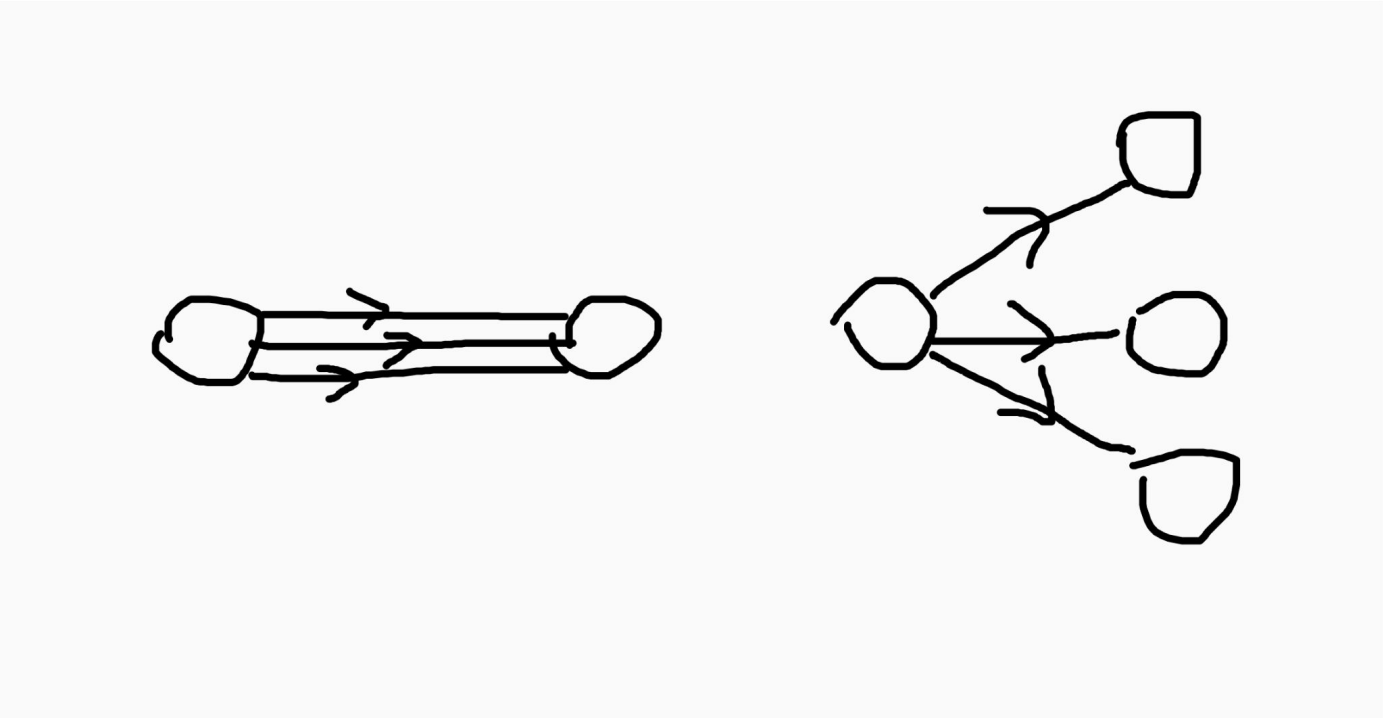
Single path routing



Multipath routing



Multipath routing



Multipath routing

- simultaneous best routes from A to B
- L2 (link aggregation) / L3 (equal routes)
- hash
  - L3: source / destination IP
  - L4: + protocol / next header, port
  - IPv6 flow id
  - inner fields (GRE, IP-in-IP, …)

Multipath problems

- path broken
- packet loss
- balancing broken

Diagnostics tools

- end-to-end test
  - ping, telnet, curl
- explore paths
  - mtr, trippy
- tricky cases
  - hping3
  - tcpdump, wireshark

mtr

- better traceroute
- stats
- interactive / report
- ICMP (default), UDP, TCP
- source / destination IP / port
  - source port UDP only
- AS lookup

mtr example (ICMP)

```
qrator@ovh01:~$ mtr --report-wide cesnet01.ring.nlnog.net
Start: 2026-01-04T18:17:30+0000
HOST: ovh01.ring.nlnog.net         Loss%   Snt   Last   Avg  Best  Wrst StDev
  1.|-- 2001:41d0:a:64ff:ff:ff:ff:fe  0.0%    10    1.2   1.1   1.0   1.2   0.1
  2.|-- 2001:41d0:0:50::1:8a68        0.0%    10    1.1   1.2   1.1   1.4   0.1
  3.|-- fdff:f000:8::1                0.0%    10    1.1   1.1   0.8   1.8   0.3
  4.|-- fdff:f003:10::3a              0.0%    10    3.0   3.0   2.2   3.7   0.5
  5.|-- 2001:41d0:20a:600::29         0.0%    10    1.0   1.1   0.9   1.2   0.1
  6.|-- be103.fra-fr5-sbb1-nc5.de.eu  0.0%    10    9.3   9.3   9.2   9.5   0.1
  7.|-- 2001:41d0::2725               0.0%    10   18.4  18.4  18.2  19.2   0.3
  8.|-- nix2-100ge.ipv6.cesnet.cz     0.0%    10   18.7  20.3  18.7  33.7   4.7
  9.|-- 2001:718:0:600:0:135:dc01:11  0.0%    10   19.0  19.0  18.8  19.3   0.1
 10.|-- 2001:718:0:600:0:135:dc01:11  0.0%    10   18.7  18.8  18.6  19.0   0.1
 11.|-- cesnet01.ring.nlnog.net       0.0%    10   18.4  18.4  18.4  18.5   0.0
```

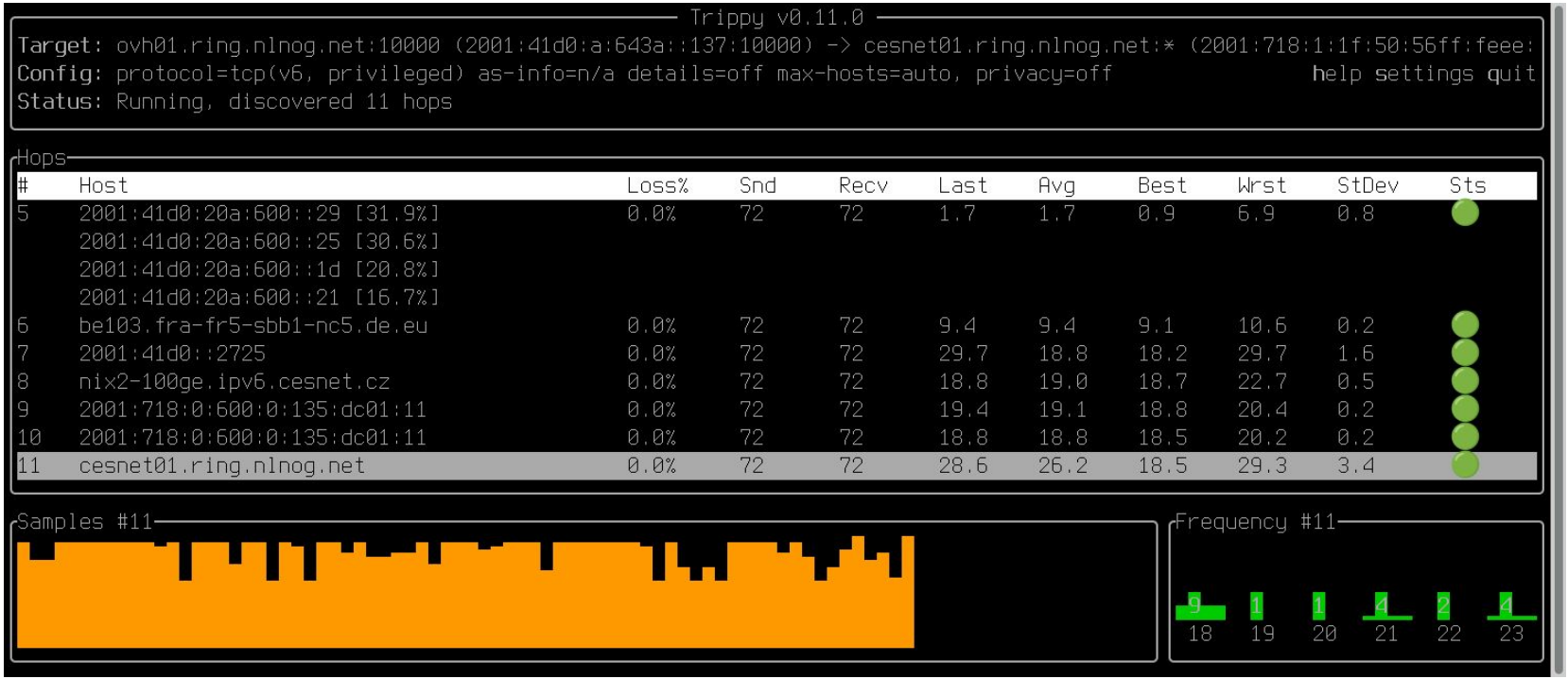
mtr example (TCP)

```
qrator@ovh01:~$ mtr --tcp --report-wide cesnet01.ring.nlnog.net
Start: 2026-01-04T18:17:54+0000
HOST: ovh01.ring.nlnog.net         Loss%   Snt   Last   Avg  Best  Wrst StDev
  1.|-- 2001:41d0:a:64ff:ff:ff:ff:fe  0.0%    10    1.2   1.1   1.0   1.3   0.1
  2.|-- 2001:41d0:0:50::1:8a6a        0.0%    10    1.4   1.4   1.2   1.5   0.1
        2001:41d0:0:50::1:8a68
  3.|-- fdff:f000:8::3                0.0%    10    1.2   1.1   0.9   1.2   0.1
        fdff:f000:8::
        fdff:f000:8::1
        fdff:f000:8::2
  4.|-- fdff:f003:10::2c              0.0%    10    2.5   3.0   2.3   4.1   0.6
        fdff:f003:10::28
        fdff:f003:10::22
        fdff:f003:10::3a
        fdff:f003:10::52
        fdff:f003:10::24
        fdff:f003:10::50
        fdff:f003:10::30
  5.|-- 2001:41d0:20a:600::29         0.0%    10    6.9   2.8   1.2   7.9   2.4
        2001:41d0:20a:600::21
        2001:41d0:20a:600::25
  6.|-- be103.fra-fr5-sbb1-nc5.de.eu  0.0%    10    9.4   9.5   9.2  10.2   0.3
  7.|-- 2001:41d0::2725               0.0%    10   21.4  19.1  18.3  21.7   1.3
  8.|-- nix2-100ge.ipv6.cesnet.cz     0.0%    10   19.1  19.0  18.8  19.2   0.1
  9.|-- 2001:718:0:600:0:135:dc01:11  0.0%    10   19.0  19.1  18.9  19.7   0.2
 10.|-- 2001:718:0:600:0:135:dc01:11  0.0%    10   18.8  18.8  18.6  19.1   0.1
 11.|-- cesnet01.ring.nlnog.net       0.0%    10   18.6  18.5  18.5  18.7   0.0
```

trippy

- mtr with TUI
- source TCP port
  - but not both ports
- multipath strategy

trippy TUI


Trippy v0.11.0
Target: ovh01.ring.nlnog.net:10000 (2001:41d0:a:643a::137:10000) -> cesnet01.ring.nlnog.net:* (2001:718:1:1f:50:56ff:feee:
Config: protocol=tcp(v6, privileged) as-info=n/a details=off max-hosts=auto, privacy=off
help settings quit
Status: Running, discovered 11 hops
Hops
Host Loss% Snd Recv Last Avg Best Wrst StDev Sts
5 2001:41d0:20a:600::29 [31.9%] 0.0% 72 72 1.7 1.7 0.9 6.9 0.8
2001:41d0:20a:600::25 [30.6%]
2001:41d0:20a:600::1d [20.8%]
2001:41d0:20a:600::21 [16.7%]
6 be103.fra-fr5-sbb1-nc5.de.eu 0.0% 72 72 9.4 9.4 9.1 10.6 0.2
7 2001:41d0::2725 0.0% 72 72 29.7 18.8 18.2 29.7 1.6
8 nix2-100ge.ipv6.cesnet.cz 0.0% 72 72 18.8 19.0 18.7 22.7 0.5
9 2001:718:0:600:0:135:dc01:11 0.0% 72 72 19.4 19.1 18.8 20.4 0.2
10 2001:718:0:600:0:135:dc01:11 0.0% 72 72 18.8 18.8 18.5 20.2 0.2
11 cesnet01.ring.nlnog.net 0.0% 72 72 28.6 26.2 18.5 29.3 3.4
Samples #11
Frequency #11
9 1 1 4 2 4
18 19 20 21 22 23

Diagnostics with “traceroute”

- rate limiting
- asymmetrical routing
  - trace from both sides
- each hop use its own path
- try ICMP, TCP, UDP
- many paths
  - try different IPs, ports

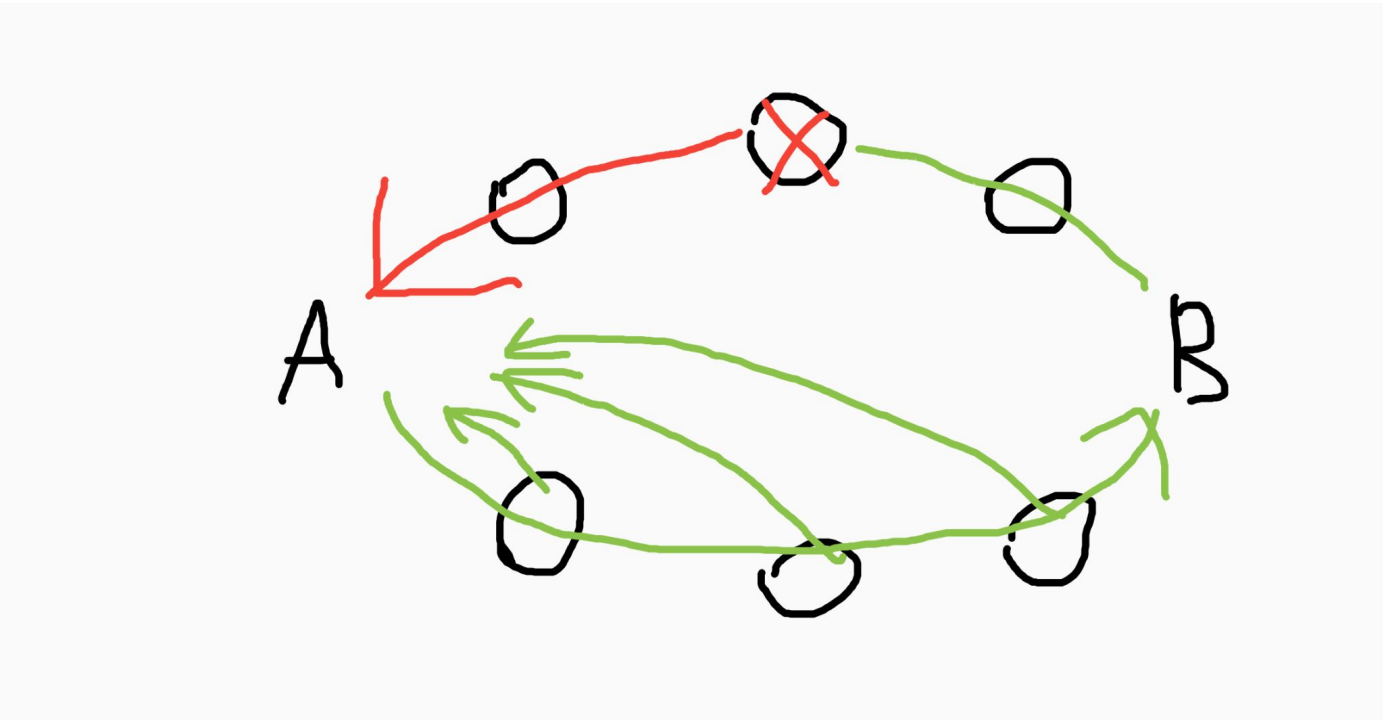
Asymmetry example: mtr a -> b

```
[root@a]# mtr -nr -c100 2001:db8::b
Start: 2026-01-18T14:13:22+0100
HOST: a                          Loss%   Snt   Last   Avg  Best  Wrst StDev
  1.|-- 3fff::1:a1:2              0.0%   100    0.3   0.3   0.2   0.4   0.0
  2.|-- 3fff::1:12:2              0.0%   100    0.5   0.5   0.3   0.7   0.1
  3.|-- 3fff::1:23:2              0.0%   100    0.7   0.6   0.4   0.9   0.1
  4.|-- 3fff::1:34:2              0.0%   100    1.0   0.8   0.6   1.1   0.1
  5.|-- 3fff::1:45:2              0.0%   100    1.0   1.0   0.8   1.3   0.1
  6.|-- 3fff::1:56:2              0.0%   100    1.1   1.2   1.0   1.6   0.1
  7.|-- 3fff::1:67:2              0.0%   100    1.5   1.4   1.0   1.7   0.1
  8.|-- 3fff::1:78:2              0.0%   100    1.8   1.6   1.2   1.8   0.1
  9.|-- 3fff::1:89:2              0.0%   100    1.9   1.7   1.5   2.0   0.1
 10.|-- 2001:db8::b               44.0%  100    2.0   2.1   1.8   2.4   0.1
```

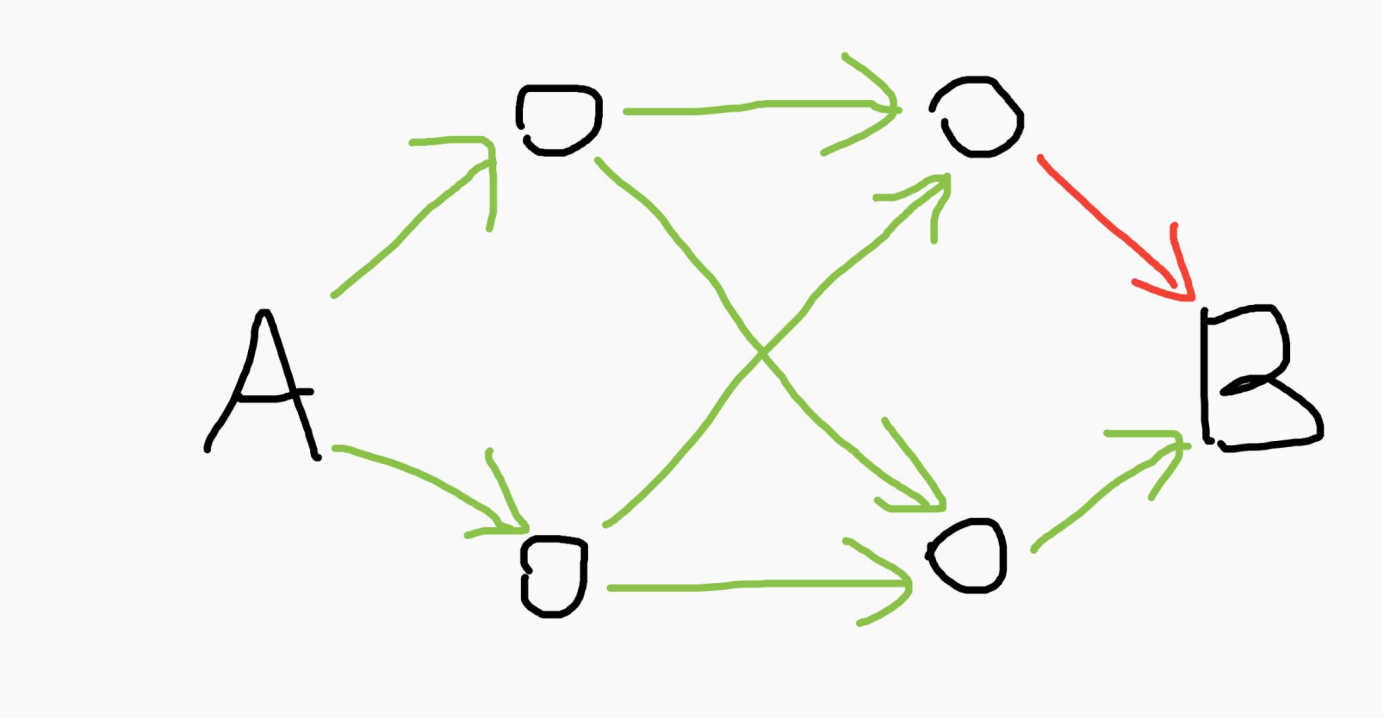
Asymmetry example: mtr b -> a

```
[root@b]# mtr -nr -c100 2001:db8::a
Start: 2026-01-18T14:16:15+0100
HOST: b                          Loss%   Snt   Last   Avg  Best  Wrst StDev
  1.|-- 3fff::2:9b:1              0.0%   100    0.2   0.3   0.2   0.4   0.0
  2.|-- 3fff::2:89:1              0.0%   100    0.4   0.4   0.3   0.7   0.1
  3.|-- 3fff::2:78:1              0.0%   100    0.6   0.6   0.5   0.8   0.1
  4.|-- 3fff::2:67:1             51.0%   100    0.6   0.8   0.6   0.9   0.1
  5.|-- 3fff::2:56:1             39.0%   100    0.7   1.0   0.7   1.3   0.1
  6.|-- 3fff::2:45:1             43.0%   100    1.2   1.2   0.9   1.4   0.1
  7.|-- 3fff::2:34:1             56.0%   100    1.3   1.4   1.1   1.6   0.1
  8.|-- 3fff::2:23:1             53.0%   100    1.3   1.5   1.1   1.9   0.1
  9.|-- 3fff::2:12:1             59.0%   100    1.5   1.7   1.2   2.1   0.2
 10.|-- 2001:db8::a              45.0%   100    1.7   2.1   1.4   2.3   0.1
```

Asymmetry example



Multipath example

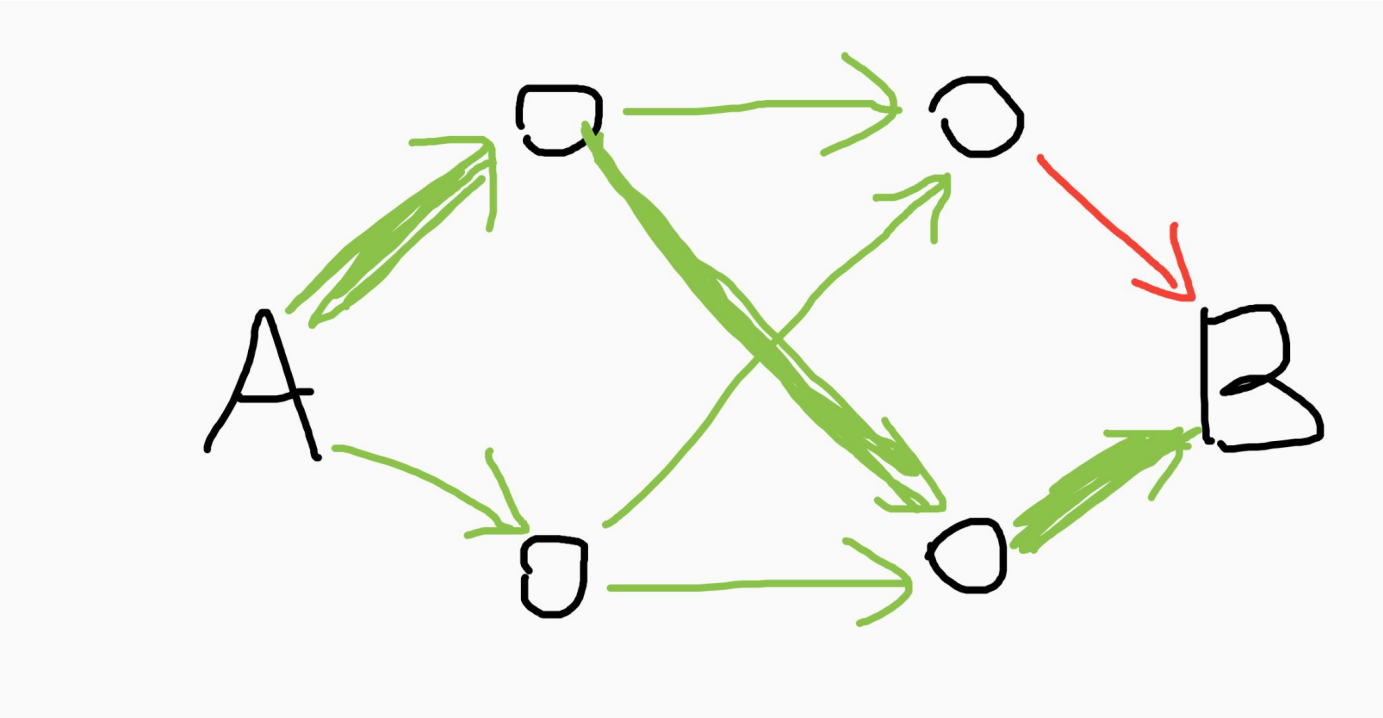


Multipath example: trippy ICMP

```
[root@a]# trip -m pretty --icmp -C 100 2001:db8::b
```

Hop	IPs	Addrs	Loss%	Snt	Recv	Last	Avg	Best	Wrst	StdDev
1	3fff::a:1:2	3fff::a:1:2	0.0	100	100	0.2	0.3	0.2	0.4	0.0
2	3fff::ab:12:b	3fff::ab:12:b	0.0	100	100	0.3	0.4	0.3	0.6	0.1
3	2001:db8::b	2001:db8::b	0.0	100	100	0.6	0.5	0.3	0.7	0.1

Multipath example: ICMP

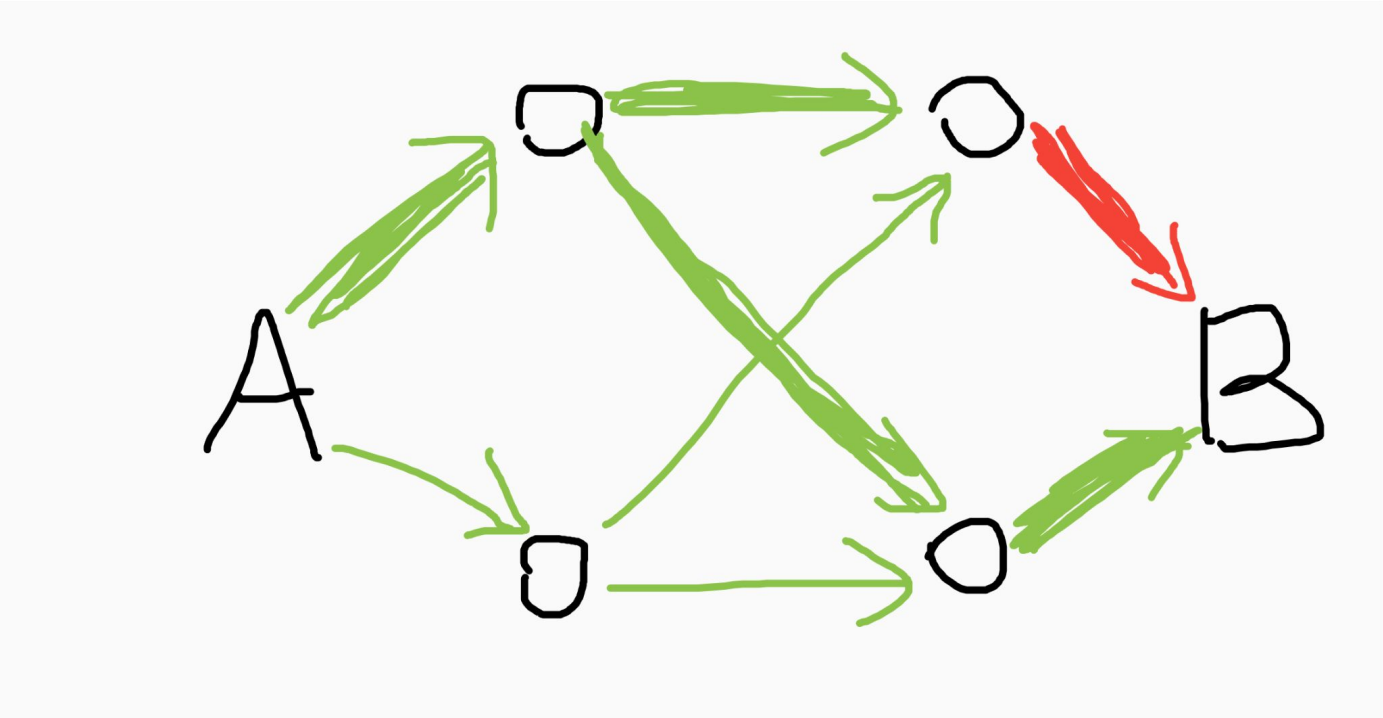


Multipath example: trippy UDP

```
[root@a]# trip -m pretty --udp -C 100 2001:db8::b
```

Hop	IPs	Addrs	Loss%	Snt	Recv	Last	Avg	Best	Wrst	StdDev
1	3fff::a:2:2	3fff::a:2:2	0.0	100	100	0.2	0.3	0.2	0.9	0.1
2	3fff::ab:21:b 3fff::ab:22:b	3fff::ab:21:b 3fff::ab:22:b	0.0	100	100	0.4	0.4	0.2	0.6	0.1
3	2001:db8::b	2001:db8::b	25.0	100	75	0.5	0.6	0.4	0.8	0.1

Multipath example: UDP



Multipath example: trippy TCP

```
[root@a]# trip -m pretty --tcp -C 100 2001:db8::b
```

Hop	IPs	Addrs	Loss%	Snt	Recv	Last	Avg	Best	Wrst	StdDev
1	3fff::a:2:2 3fff::a:1:2	3fff::a:2:2 3fff::a:1:2	0.0	100	100	0.3	0.3	0.2	0.5	0.0
2	3fff::ab:12:b 3fff::ab:21:b 3fff::ab:11:b 3fff::ab:22:b	3fff::ab:12:b 3fff::ab:21:b 3fff::ab:11:b 3fff::ab:22:b	0.0	100	100	0.5	0.5	0.4	0.6	0.1
3	2001:db8::b	2001:db8::b	30.0	100	70	10.2	10.2	10.2	10.4	0.0

Multipath example: mtr TCP

```
[root@a]# mtr -nr -c 100 --tcp 2001:db8::b
Start: 2026-01-18T17:16:13+0100
HOST: a                          Loss%   Snt   Last   Avg  Best  Wrst StDev
  1.|-- 3fff::a:1:2               0.0%   100    0.3   0.3   0.2   0.5   0.0
        3fff::a:2:2
  2.|-- 3fff::ab:21:b             0.0%   100    0.6   0.5   0.4   0.8   0.1
        3fff::ab:12:b
        3fff::ab:11:b
        3fff::ab:22:b
  3.|-- 2001:db8::b               0.0%   100  3065. 545.8   0.5 5150. 1016.3
```

Multipath example: trippy UDP fixed port, good

```
[root@a]# trip -m pretty --udp -S 10000 -P 80 -R dublin -C 100 2001:db8::b
```

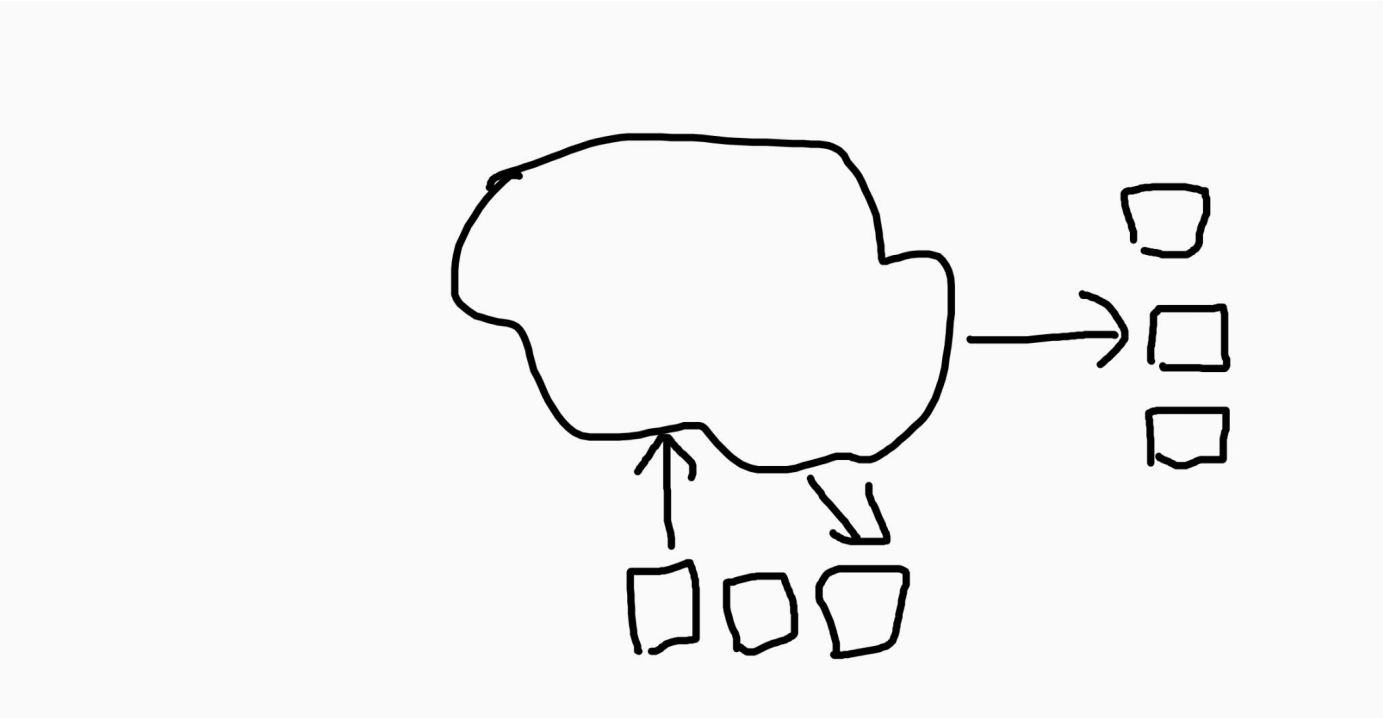
Hop	IPs	Addrs	Loss%	Snt	Recv	Last	Avg	Best	Wrst	StdDev
1	3fff::a:2:2	3fff::a:2:2	0.0	100	100	0.3	0.3	0.2	0.5	0.0
2	3fff::ab:22:b	3fff::ab:22:b	0.0	100	100	0.5	0.5	0.4	0.6	0.1
3	2001:db8::b	2001:db8::b	0.0	100	100	0.7	0.6	0.4	0.8	0.1

Multipath example: trippy UDP fixed port, bad

```
[root@a]# trip -m pretty --udp -S 10001 -P 80 -R dublin -C 100 2001:db8::b
```

Hop	IPs	Addrs	Loss%	Snt	Recv	Last	Avg	Best	Wrst	StdDev
1	3fff::a:2:2	3fff::a:2:2	0.0	100	100	0.2	0.3	0.2	0.5	0.0
2	3fff::ab:21:b	3fff::ab:21:b	0.0	100	100	0.3	0.4	0.3	0.6	0.1
3	2001:db8::b	2001:db8::b	50.0	100	50	0.4	0.7	0.4	0.8	0.1
4	2001:db8::b	2001:db8::b	83.3	6	1	0.6	0.6	0.6	0.6	0.0

Traceroute with anycast source



hping3

- highly customizable “ping”
  - TTL
  - protocol
  - port
  - TCP flags
- no IPv6
- pseudo traceroute
  - stops at first missing answer
- custom payload

tcpdump: IPv4

- trace 78.128.211.50 (cesnet01) -> 37.187.111.58 (ovh01)
- 37.187.111.58 = 0x4e80d332
- offset = 4 (icmp header) + 16 (dst addr in ip header) = 20
- https://en.wikipedia.org/wiki/IPv4
- https://en.wikipedia.org/wiki/Internet_Control_Message_Protocol
- pcap filter:

(dst host 78.128.211.50 and src host 37.187.111.58)

or

(dst host 78.128.211.50 and icmp and icmp[20:4] = 0x4e80d332)

tcpdump: IPv6

- trace 2001:718:1:1f:50:56ff:feee:50 (cesnet01) -> 2001:41d0:a:643a::137 (ovh01)
- 2001:41d0:a:643a::137 = 0x200141d0 0x000a643a 0x00000000 0x00000137
- offset = 40 (ip header) + 4 (icmp header) + 4 (icmp time exceeded) +
  24 (dst addr in ip header) = 72
- https://en.wikipedia.org/wiki/IPv6_packet
- https://en.wikipedia.org/wiki/Internet_Control_Message_Protocol
- pcap filter:

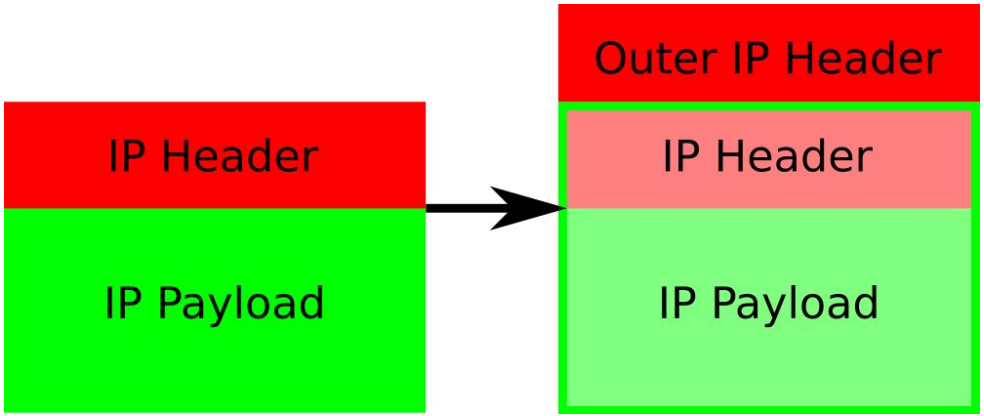
(dst host 2001:718:1:1f:50:56ff:feee:50 and src host 2001:41d0:a:643a::137)
or
(dst host 2001:718:1:1f:50:56ff:feee:50 and icmp6 and
ip6[72:4] = 0x200141d0 and ip6[76:4] = 0x000a643a and
ip6[80:4] = 0x00000000 and ip6[84:4] = 0x00000137)

Broken IPIP case

- TCP connection in IPIP tunnel to anycast address fails
- Tcpdump shows splitted TCP connection
- Usual diagnostics shows stable path
- Extract packets with wireshark
- Use hping3 to reproduce

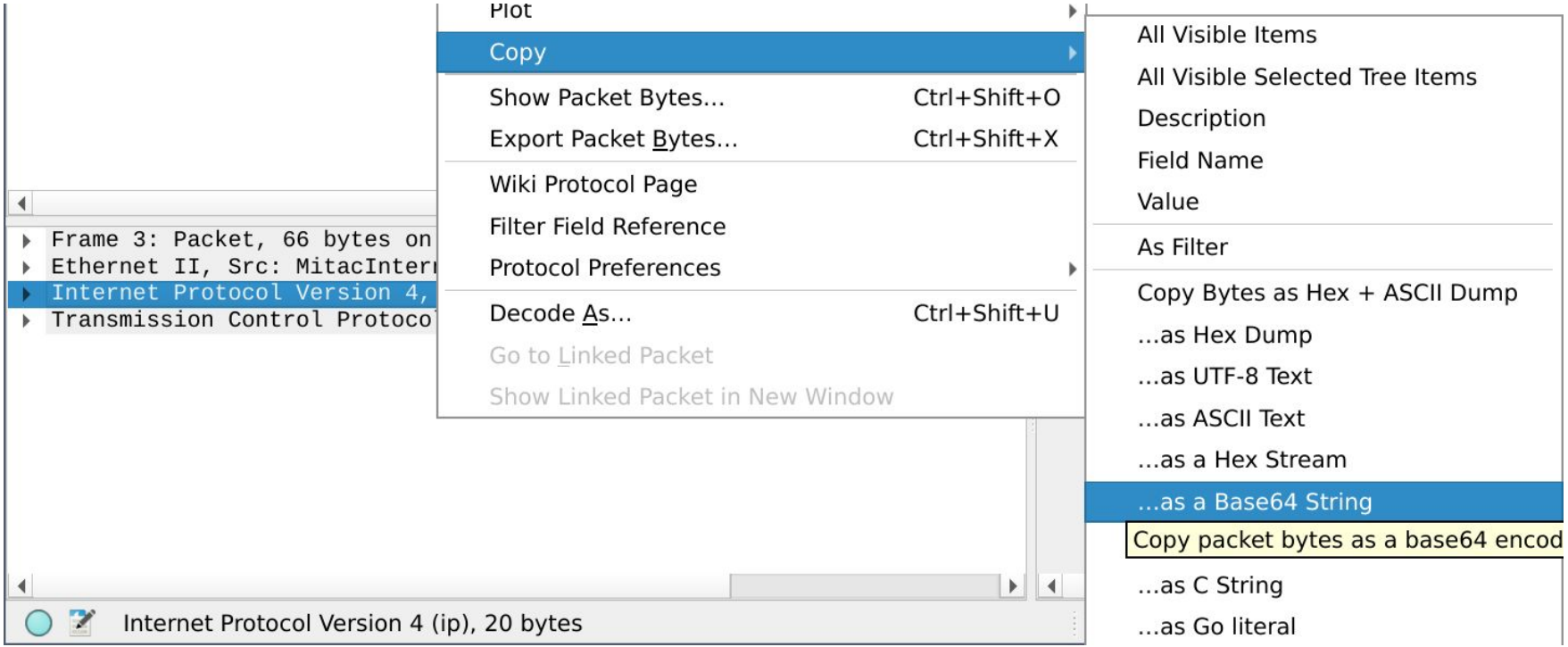
Tunneling

- IPIP, GRE
- IPIP:
  - IPv4 packet is encapsulated into another IPv4 packet
  - IP proto = 4



By Cybjit - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=17065586

Broken IPIP: wireshark



Broken IPIP: hping

```
qrator@ovh01:~$ echo "<base64>" | base64 -d >> tcp.1
...
qrator@ovh01:~$ echo "<base64>" | base64 -d >> tcp.2
...
qrator@ovh01:~$ hping3 --traceroute --rawip --ipproto 4 --data 52 --file tcp.1 <dst_ip>
hop=1 TTL 0 during transit from ip=37.187.111.253 name=UNKNOWN
hop=2 TTL 0 during transit from ip=10.17.137.104 name=UNKNOWN
hop=3 TTL 0 during transit from ip=10.73.2.146 name=UNKNOWN
hop=4 TTL 0 during transit from ip=172.20.16.0 name=UNKNOWN
hop=5 TTL 0 during transit from ip=37.59.16.31 name=UNKNOWN
hop=6 TTL 0 during transit from ip=213.186.32.215 name=be102.par-th2-sbb1-nc5.fr.eu
...
qrator@ovh01:~$ hping3 -T -0 -H 4 -d 52 -E tcp.2 <dst_ip>
hop=1 TTL 0 during transit from ip=37.187.111.253 name=UNKNOWN
hop=2 TTL 0 during transit from ip=10.17.137.104 name=UNKNOWN
hop=3 TTL 0 during transit from ip=10.73.2.146 name=UNKNOWN
hop=4 TTL 0 during transit from ip=172.20.16.0 name=UNKNOWN
hop=5 TTL 0 during transit from ip=37.59.16.29 name=UNKNOWN
hop=6 TTL 0 during transit from ip=91.121.215.177 name=be102.par-gsw-sbb1-nc5.fr.eu
...
```

QRATOR LABS

Thank you

Alexander Zubkov
green@qrator.net