v4-via-v6 routing

Juliusz Chroboczek
IRIF
Université Paris-Cité

(Joint work
with Toke Høiland-Jørgensen
and Warren Kumari)

17 March 2025

v4-via-v6 routing

V4-via-v6 routing is a technique for forwarding IPv4 through routers with no IPv4 address.

- encourages IPv6 adoption
  (deploy IPv6, get IPv4 for free);
- simplifies administration
  (no IPv4 addresses inside the network);
- some minor difficulties with ICMP (solved).

History

- May 2009: RFC 5549 for BGP;
- Summer 2020: implemented in Babel (joint work with Théophile Bastian);
- May 2022: RFC 9229 (Experimental), Babel protocol extension.

Deployed in production since May 2022:

- enabled by default in babeld;
- explicitly enabled in BIRD.

Nobody complained!

Next-hop forwarding

Forwarding IPv4 packets:

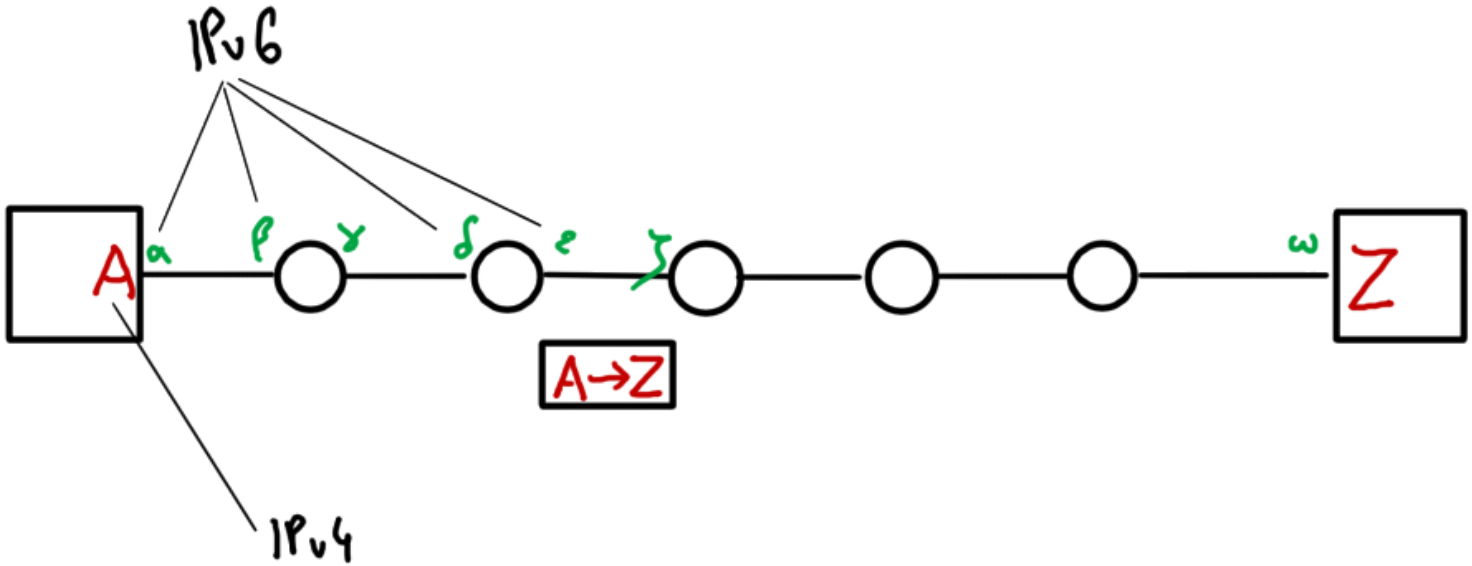
- map IPv4 destination ⟶ IPv4 next hop;
- map IPv4 next hop ⟶ MAC address (ARP);
- forward the packet to the MAC address.

Forwarding IPv6 packets:

- map IPv6 destination ⟶ IPv6 next hop;
- map IPv6 next hop ⟶ MAC address (IPv6 ND);
- forward the packet to the MAC address.

The next-hop IP never appears on the wire.
The next-hop IP is an internal implementation detail.

v4-via-v6 forwarding



V4-via-v6 forwarding:

- map IPv4 destination ⟶ IPv6 next hop;
- map IPv6 next hop ⟶ MAC address (IPv6 ND);
- forward the packet to the MAC address.

Routing table maps IPv4 addresses to IPv6 next-hops.

v4-via-v6 forwarding: implementations

Routing table maps IPv4 addresses to IPv6 next-hops:

`203.0.113.0/24` ⟶ `fe80::1` `via eth1`

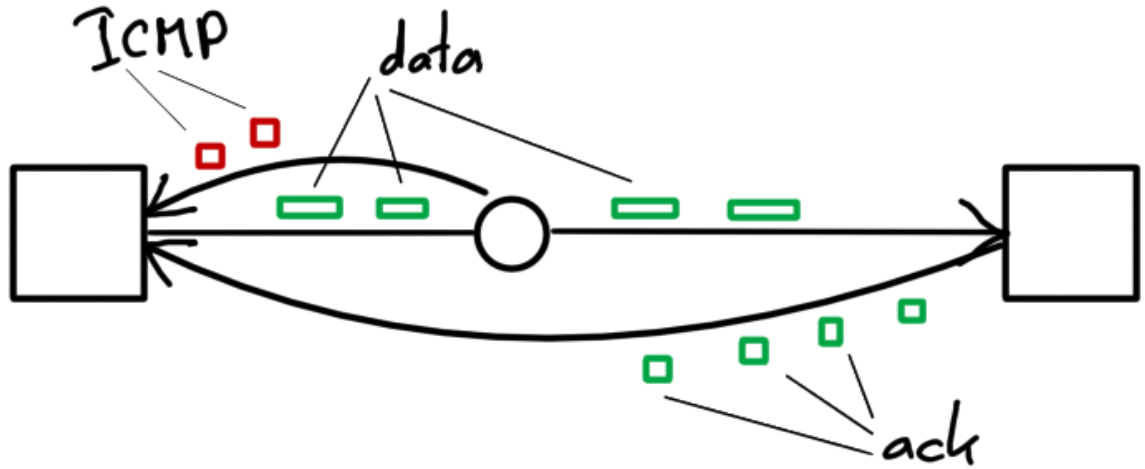
This is implemented on:

- Linux (since version 5.2.0);
- Arista EOS;
- Mikrotik;
- probably others.

ICMP issues

Modern IPv4 relies on ICMP from intermediate routers:

- PMTU discovery relies on *ICMP fragmentation required*.



With v4-via-v6 routing, the ICMP originator does not necessarily have an IPv4 address assigned.

ICMP solutions

Intermediate routers need to orginate ICMPv4 packets. The ICMPv4 packet's source address doesn't really matter.

Solution:

- borrow an IPv4 address from another interface, if available;
- as a last resort, use 192.0.0.8, the "IPv4 dummy address" (RFC 7600).

Fallback procedure implemented in Linux since version 5.13.0 (Toke Høiland-Jørgensen). Other systems unknown.

Makes traceroute confusing, should be combined with draft-ietf-intarea-extended-icmp-nodeid.

Results

Routing table

```
$ ip route show 10.0.0.2
10.0.0.2 via inet6 fe80::216:3eff:fe00:1
    dev lxcbr0 proto babel onlink
```

Traceroute

```
$ traceroute 10.0.0.2
traceroute to 10.0.0.2 (10.0.0.2) [...]
  1 192.0.0.8 (192.0.0.8) [...] 0.014 ms
  2 192.0.0.8 (192.0.0.8) [...] 0.042 ms
  3 192.0.0.8 (192.0.0.8) [...] 0.030 ms
  4 10.0.0.2 (10.0.0.2) [...] 0.039 ms
```

PMTUD

```
19:58:47.402871 IP 192.168.0.27.60046 > 10.0.0.2.22:
   Flags [.], seq 33:1481 [...] length 1448
19:58:47.402874 IP 192.168.0.27.60046 > 10.0.0.2.22:
   Flags [P.], seq 1481:1537 [...] length 56
19:58:47.402906 IP 192.0.0.8 > 192.168.0.27:
   ICMP 10.0.0.2 unreachable - need to frag (mtu 1420),
   length 556
19:58:47.402919 IP 10.0.0.2.22 > 192.168.0.27.60046:
   Flags [.], ack 33 [...] sack 1 1481:1537] [...]
19:58:47.402934 IP 192.168.0.27.60046 > 10.0.0.2.22:
   Flags [.], seq 33:1401 [...] length 1368
```

Please adopt?

V4-via-v6 routing:

- encourages deployment of IPv6
  (deploy IPv6, get IPv4 for free);
- simplifies administration
  (no IPv4 addresses on intermediate routers);
- deployed in production since 2022.

Described in RFC 9229:

- Babel-specific;
- Experimental.

draft-chroboczek-intarea-v4-via-v6:

- routing-protocol agnostic;
- aims for Standards Track.

Please adopt?