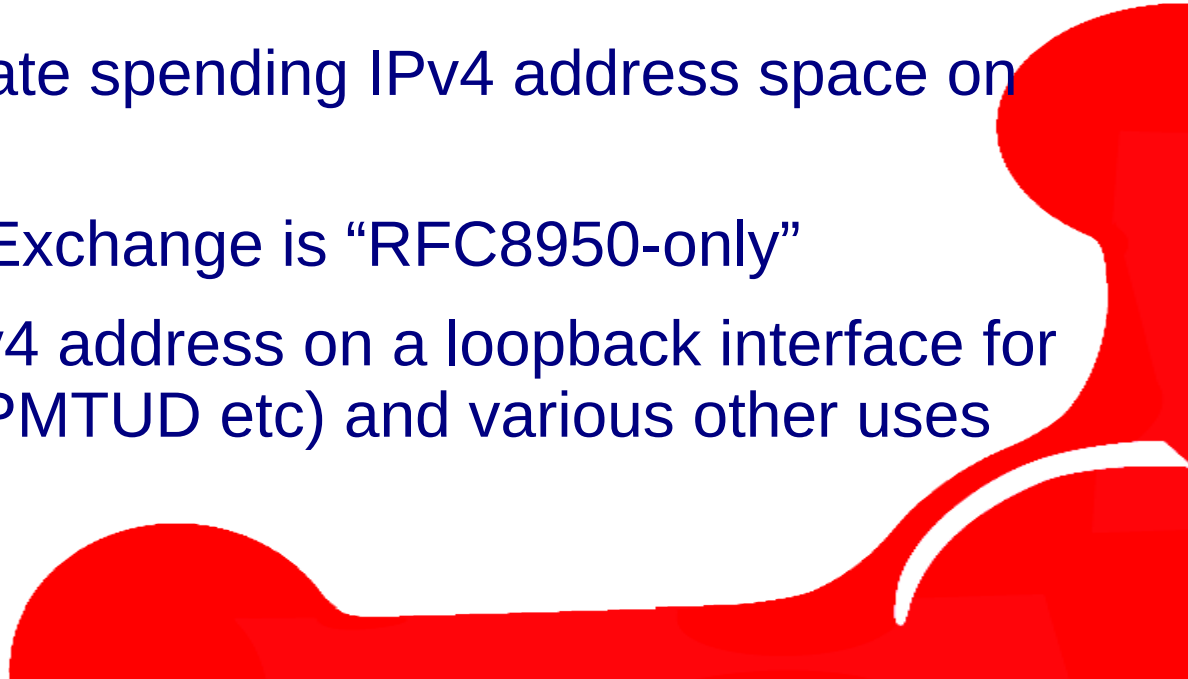
Peering at the Speed of RFC8950

Aleksi Suhonen

June 2025

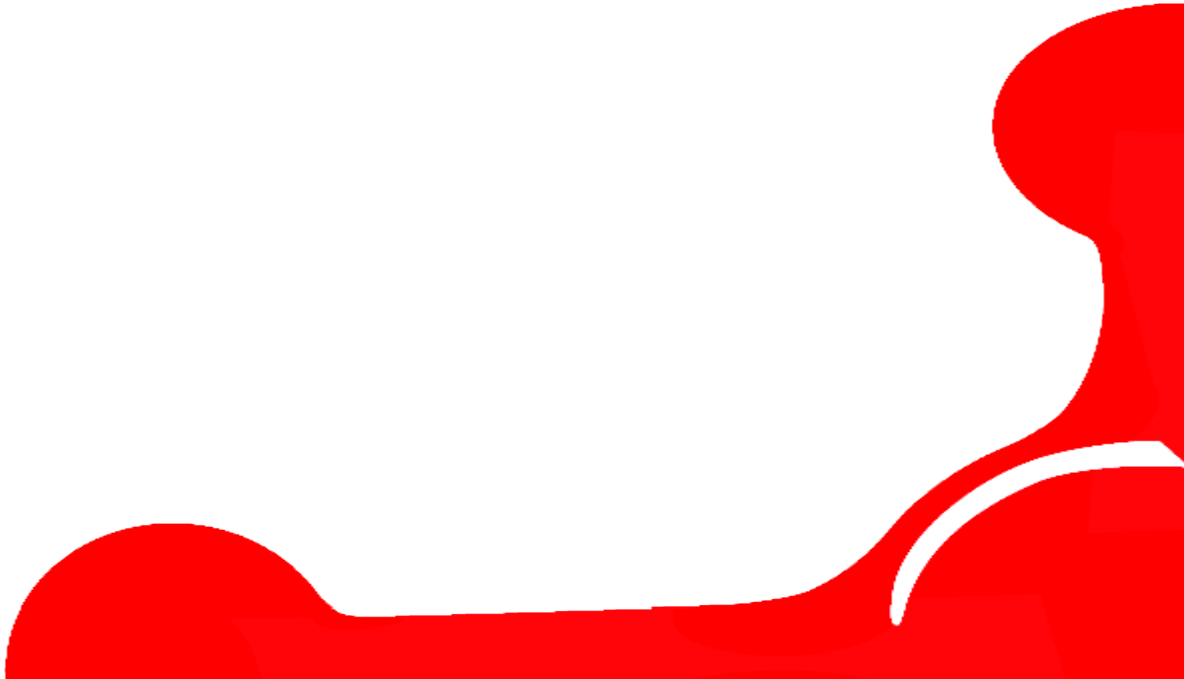
What's This All About?

- RFC8950 specifies announcing IPv4 routes with IPv6 next-hop addresses using BGP
- RFC8950 is the future!
- Allows networks to eliminate spending IPv4 address space on link networks
- The new Turku Regional Exchange is “RFC8950-only”
- Routers still need **one** IPv4 address on a loopback interface for ICMP errors (traceroute, PMTUD etc) and various other uses



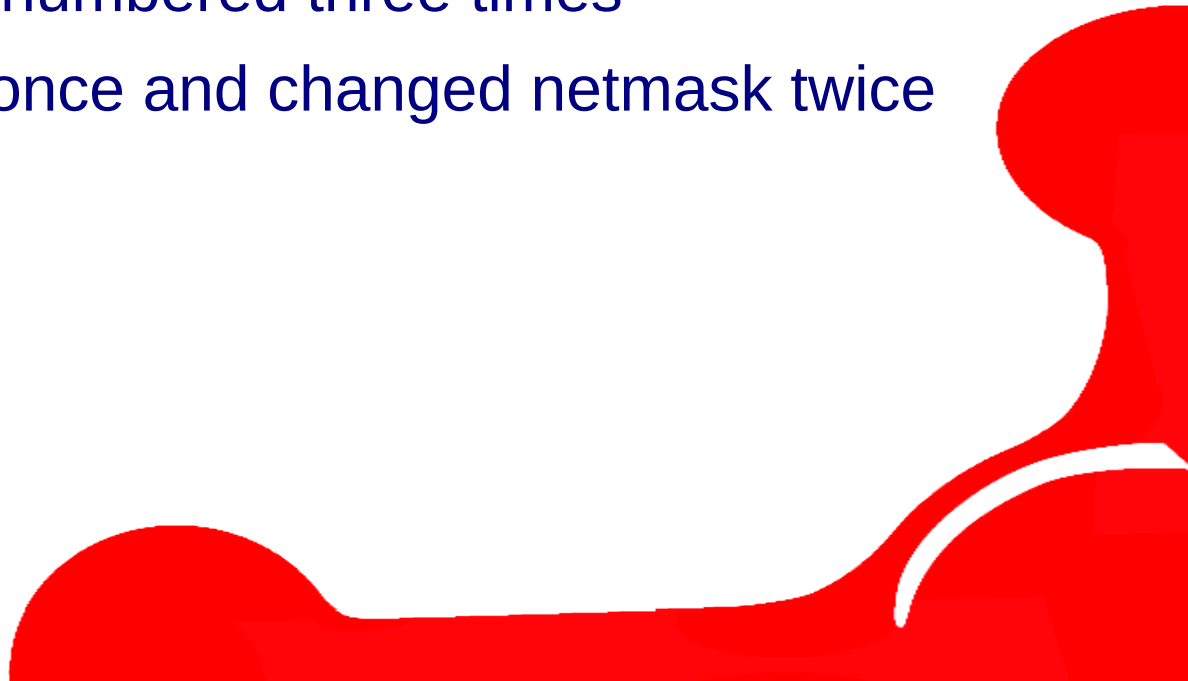
But why?

- We already had a /26 of IPv4 addresses earmarked for turku
- But renumbering into a bigger network later would be painful



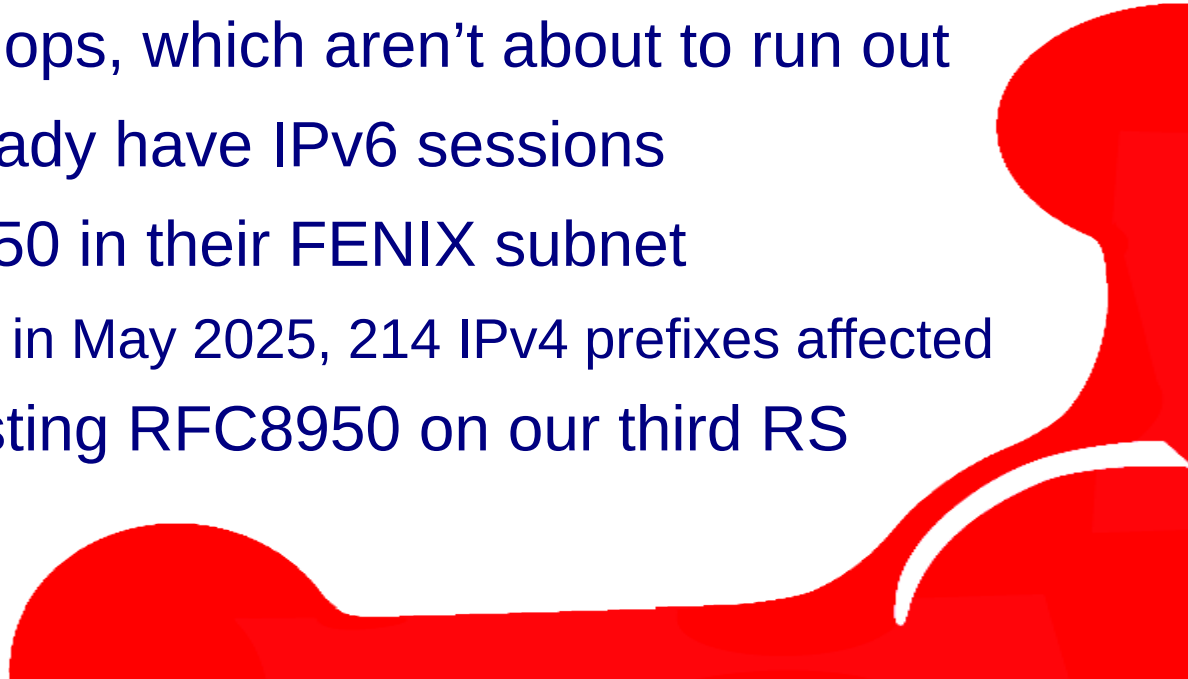
Renumbering an IXP Is Painful

- France-IX renumbering took almost two years 2019-2021
  - Thanks to Radu-Adrian Feurdean for dates
- LINX and AMS-IX have renumbered three times
- DE-CIX has renumbered once and changed netmask twice



Alternative: One Final "Renumbering" Event

- When the peering subnet becomes full, think about whether you really want to go through the pain of renumbering again and again...
- ... or switch to IPv6 next-hops, which aren't about to run out
- All members typically already have IPv6 sessions
- NIX (cz) is testing RFC8950 in their FENIX subnet
  - 12 members, 24 sessions in May 2025, 214 IPv4 prefixes affected
- TREX Tampere is also testing RFC8950 on our third RS



What About New IXPs?

- If we start with an IPv4 subnet, it will be difficult to get members to give it up later
- So what if take a chance, and never have to renumber at all?
- We are testing this IPv6-only approach at TREX Turku
  - A new IXP, not a lot of connected members yet
  - Second site is not yet up and running

RFC8950-ixp Working Group

- Euro-IX started a working group to map out the challenges and to work on best practices for adopting RFC8950 at IXPs
  - Chairmen: André Grüneberg (BCIX) and Aleksi Suhonen (TREX)
  - Members from DE-CIX, LINX, NIX.CZ, nic.cz, NetDEF, ...
- https://github.com/euro-ix/rfc8950-ixp
  - Pull requests accepted...
- There’s also a mailing list and a mattermost chat channel

RFC8950 Challenges

- Most of our members use platforms that have pretty mature RFC8950 support already (like Juniper and Nokia)
- Some new members use platforms with incomplete support:
  - Mikrotik kernel has supported it for a long while,
  - but their BGP implementation doesnt...
  - Huawei has said they support it, but hasn’t provided config docs
- IBGP next-hop-self even more important now
- ICMP Unreachables unpredictable on some platforms

Junos Config Example 1/2

```
[edit interfaces irb unit 7]
description EXCHANGE_Turku;
family inet {
    filter {
        input spoof-protect-trex;
    }
}
family inet6 {
    filter {
        input 6spoof-protect-trex;
    }
    address 2001:7f8:1d:7::72f8:1/64;
}
```

- The family inet section has to exist for the router to process IPv4 traffic
- But it doesn’t need an address.
- It can have an unrelated address, that will show up in traceroute tho...

Junos Config Example 2/2

```
[edit protocols bgp group peering-rfc8950]
type external;
export [ 6export_peering 4export_peering deny-all ];
import [ 4hygiene 6hygiene deny-tier1s 4import_peering 6import_peering deny-all ];
family inet {
    unicast {
        prefix-limit maximum 9999;
        extended-nexthop;
    }
}
family inet6 {
    unicast {
        prefix-limit maximum 999;
    }
}
```

Junos Show Route Example

```
axu@betty> show route 193.163.5.0 terse

inet.0: 975756 destinations, 1499988 routes (507404 active, 0 holddown, 728982 hidden)
+ = Active Route, - = Last Active, * = Both

A V Destination        P Prf   Metric 1   Metric 2  Next hop        AS path
* V 193.163.5.0/24     B 170        300             >2001:7f8:1d:7:0:3:289b:2 207003 I
  V                    B 170        300             >2001:7f8:1d:7:0:3:289b:1 207003 I
  V                    B 170        100             >195.140.192.47  207003 I
  V                    B 170         60           0 >195.140.192.13  6667 207003 I

axu@betty>
```