Source Address Selection In Multi-Router Multi-Prefix Scenario

IETF123 Hackathon Report

David Lamparter, Jen Linkova

Recap: RFC6724 Rule 5.5

New flow: selecting a source for the given destination

RFC6724 Source Address Selection Rule 5.5:

> "Prefer addresses in a prefix advertised by the next-hop"

Select the destination -> next-hop -> source address

Rule 5.5 Example

1. Sending a packet to

   2001:db8:aaa6:bbb::1

2. Select **R1** as next-hop

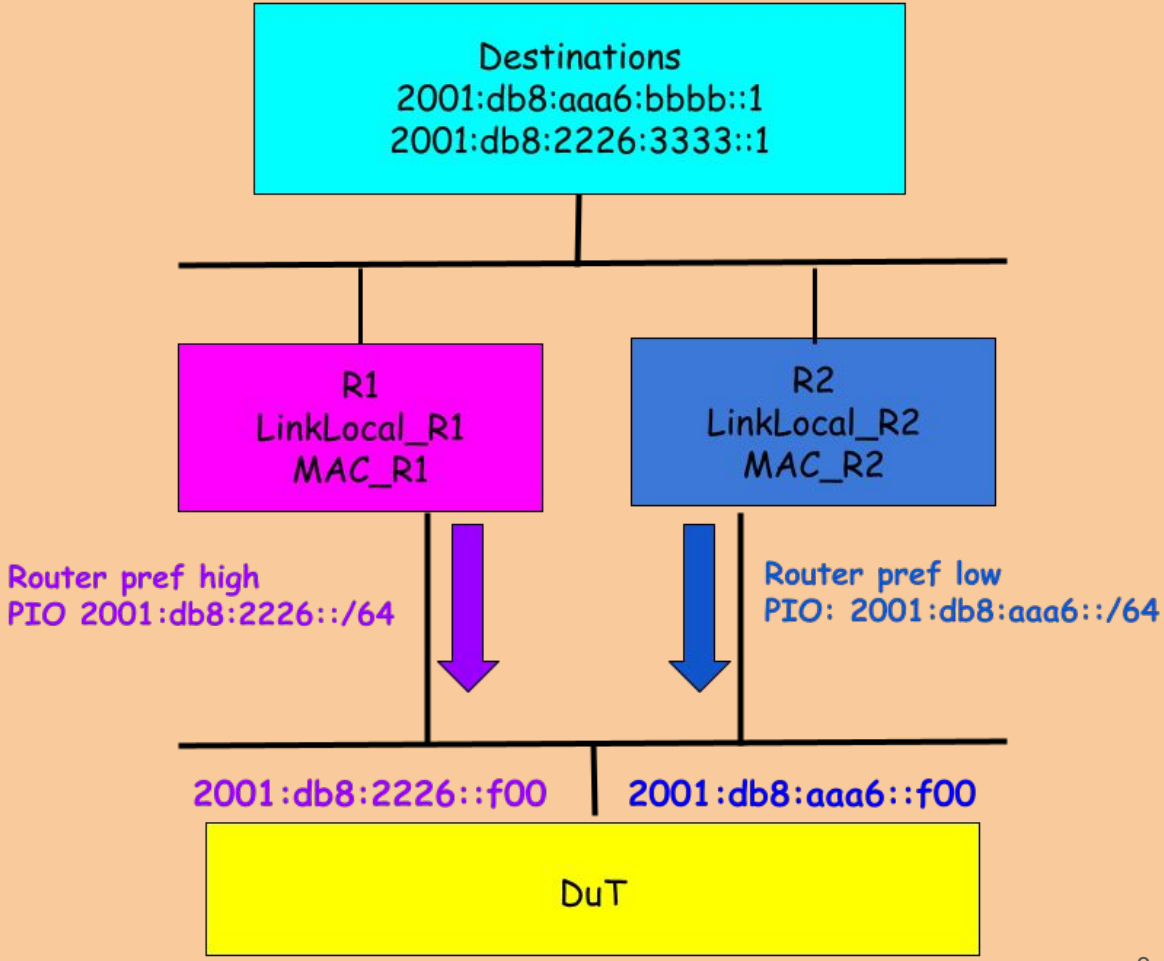
3. Select source IPv6

   2001:db8:2226::f00

w/o Rule 5.5: Rule 8 chooses the

longest prefix match:

2001:db8:aaa6::f00



Recap: RFC8028

Routing packets for existing flows when multiple routers are present

RFC8028: "A host SHOULD select default routers for each prefix it is assigned an address in" selecting the correct nexthop for the given source address"

The only implementation found uses Destination-Source routing

- Source IPv6 address is pre-defined (hardcoded)
- Network topology has changed

RFC8028 Example

1. Host sends a packet from

   2001:db8:aaa6::f00

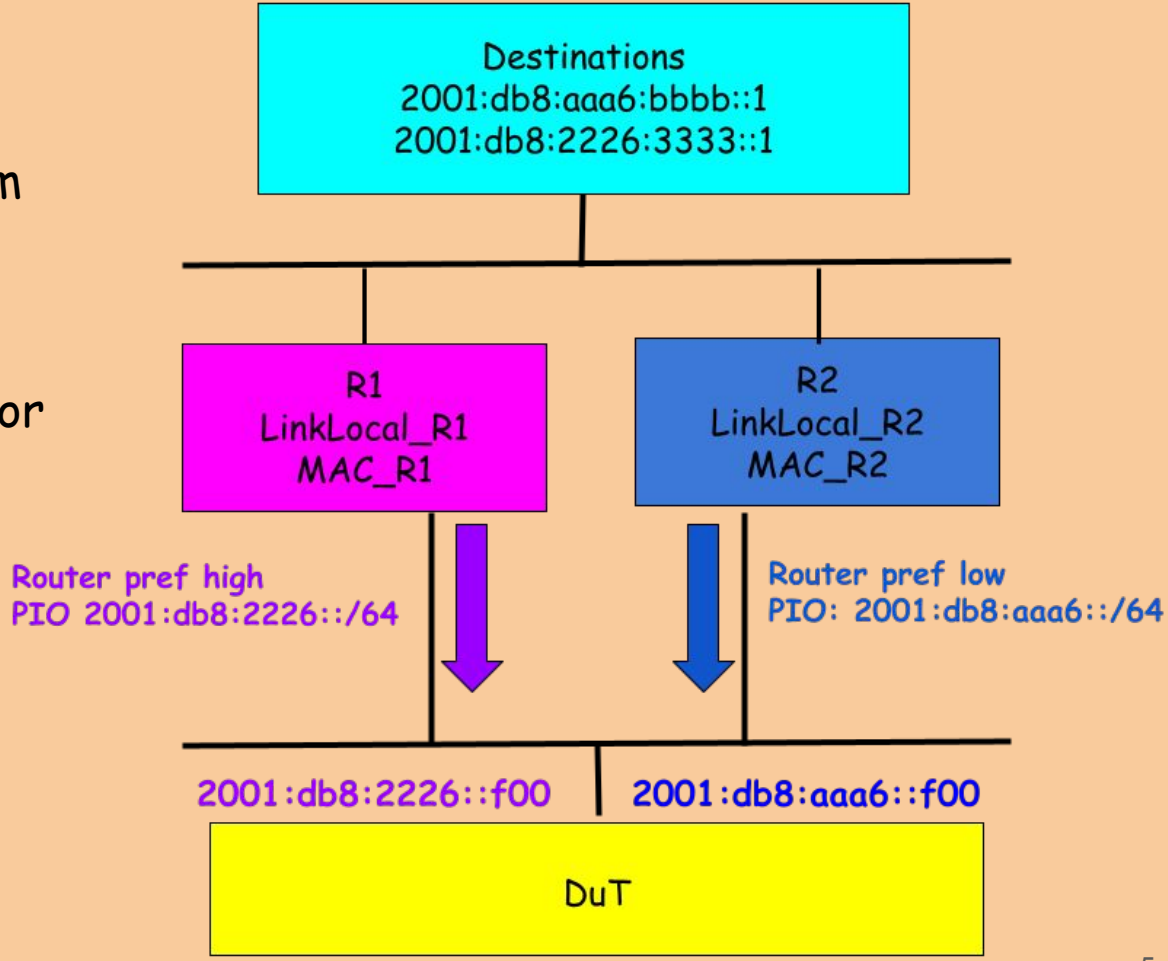
2. The default router list for

   2001:db8:aaa6::/64

   contains **R2**

3. Select **R2** as next-hop

*w/o RFC8028 support: **R1** is chosen*



Motivation for Hackathon

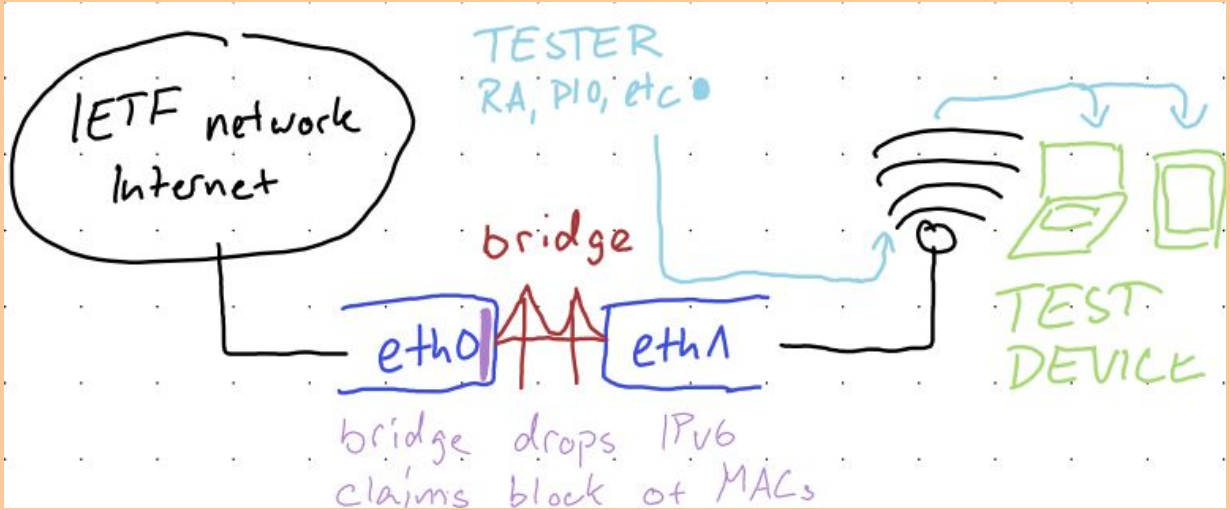
Test multiple routers and multiple prefixes on a single link:

- Source address selection (RFC6724 Rule 5.5 specifically)
- RFC8028 support
- Failover cases when a router disappears or re-appears

Real-life applicability:

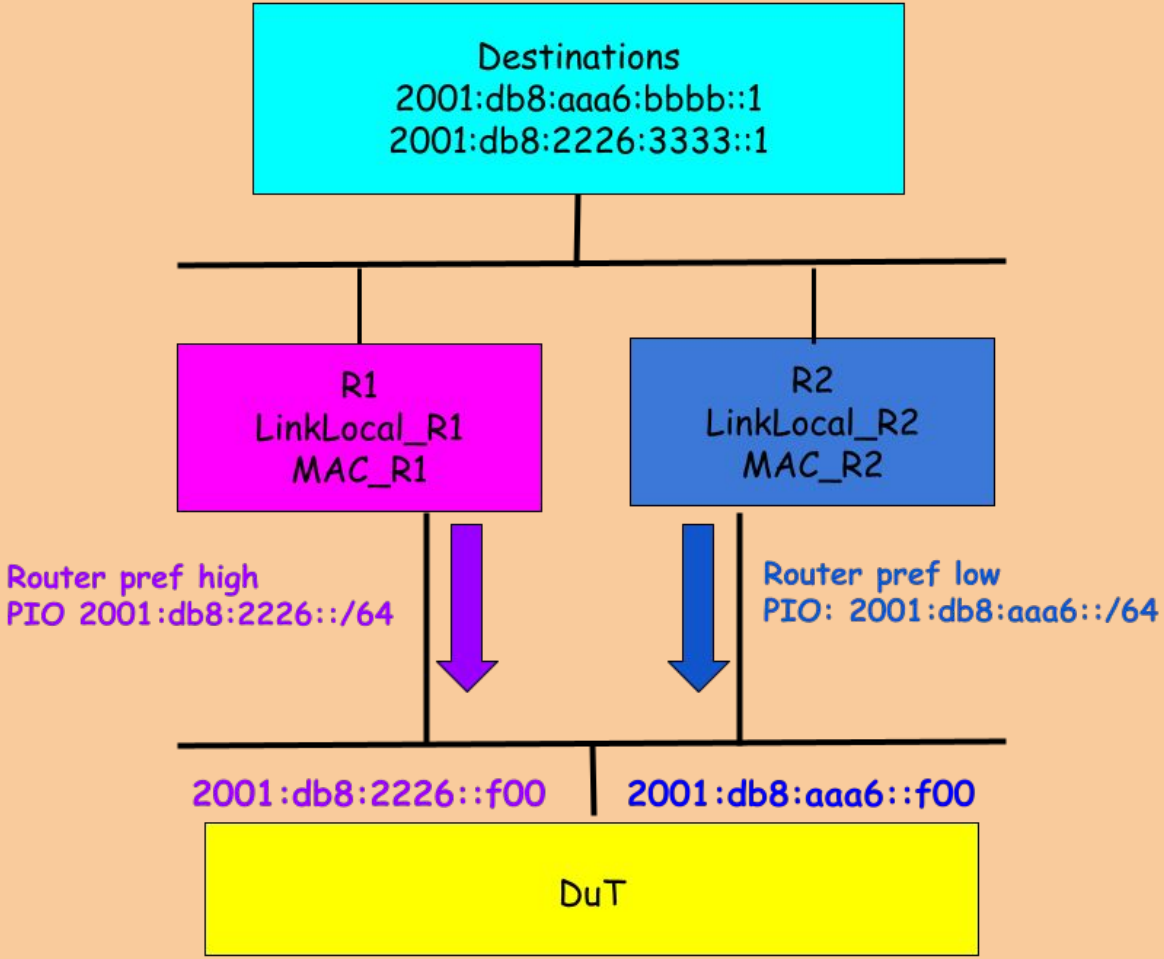
- Multihoming
- Flash renumbering
- Applicability of draft-link-6man-gulla

Testbed Physical Topology



- Python asyncio web server
- Clients connect to the test WiFi, reach the server over IPv4
- The server
  - Instructs the client to open connections (JavaScript)
  - Sends RAs
- Code available at Github: https://github.com/eqvinox/rule5p5-tests

Logical Topology



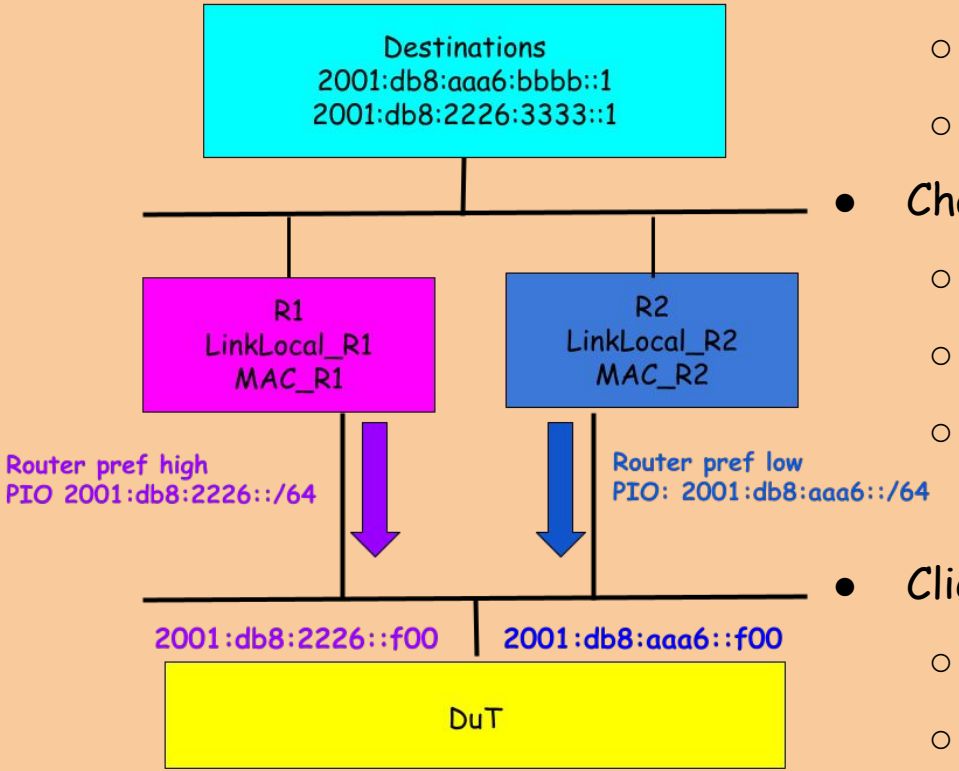
Test Cases

- RFC6724 Rule 5.5 when two routers (two PIOs) are present
  - Selecting the correct source address for new flows
- RFC8028 behaviour
  - selecting the correct nexthop for the given source address
- Reaction to topology changes:
  - PIO deprecation
  - RIO changes
  - Router disappears (and then comes back)
    - Explicit signal: RA with router lifetime = 0
    - Implicit signal: NUD (router stops responding to NSes)
      - Emulates flash renumbering with draft-link-6man-gulla
    - Destination can be reachable via ::/0 or via RIO

Devices Under Test

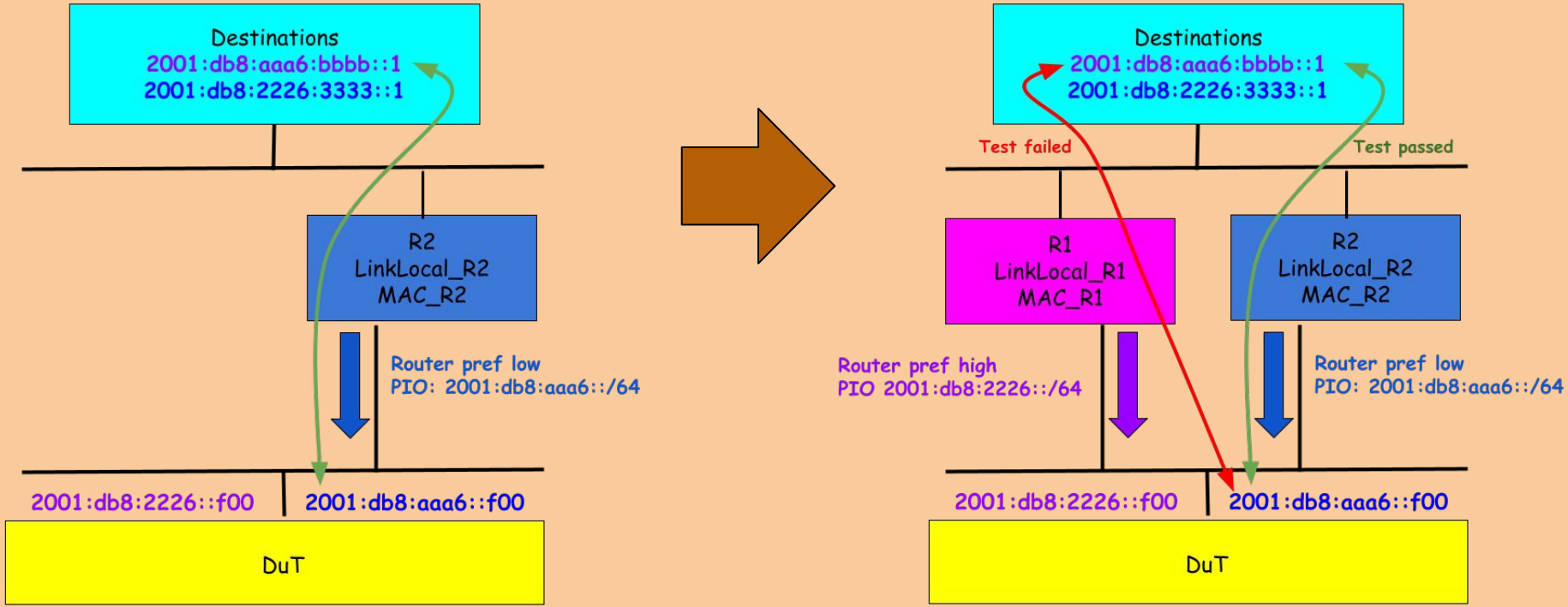
- MacOS 15.5
- Windows 10
- Android 16
- ChromeOS ver 16328.25.0 (Browser version 139.0.7258.43)
- Linux with RAs processed by kernel
  - kernel 6.12
  - Kernel 6.14 + Rule 5.5 patch
- Linux with NetworkManager 1.46.0-1u2.2

RFC6724 Rule 5.5 Tests

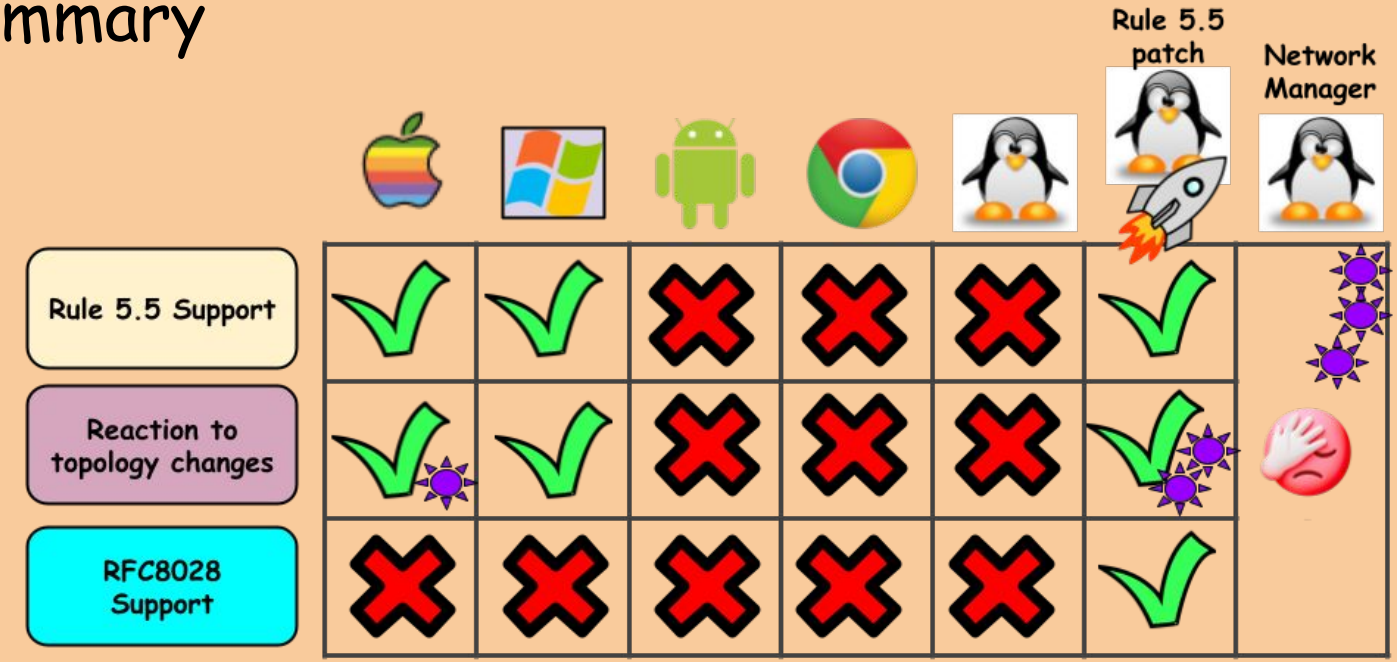


- Client connects to **2001:db8:aaa6:bbbb:1**
  - Pass: 2001:db8:2226::/64 is used
  - Fail: **2001:db8:aaa6::/64** is used (Rule 8)
- Change the topology (3 tests):
  - Deprecate R1
  - Deprecate 2001:db8:2226::/64
  - R1 stops responding to NSes (draft-link-6man-gulla)
- Client connects to **2001:db8:2226:3333::1**
  - Pass: **2001:db8:aaa6::/64** is used
  - Fail: 2001:db8:2226::/64 is used (Rule 8)

RFC8028: Selecting Next-hop for a Source Address



Results Summary



A bug discovered when a router is deprecated (see the next slides)

More work needed to process PIO deprecation

NM performs ECMP for multiple routers

Results

- Both MacOS and Windows support Rule 5.5
- No RFC8028 support yet
- Linux-based systems lack RFC6724 Rule 5.5 support
  - David's patch mostly works
    - adds Rule 5.5 and RFC8028 support to Linux
    - Not dealing with PIO deprecation yet
  - But there is ~~a catch~~ NetworkManager…

NetworkManager does NetworkManager Things © David

NM creates a default route with multiple nexthops, does ECMP

Even if routers have different Default Router Preferences

Therefore can not behave correctly in multi-router multi-prefix environment

Issue with MacOS 15.5

Two routers are present (**R1**, **R2**), **R1** is selected as default gateway

default	***fe80::7d12:ca25:4e5a:e476%en0***	*UGcg*	*en0*

When R1 is deprecated (router lifetime = 0), R2 appears in the routing table

That route is scoped to the interface

default	***fe80::7ba1:d555:9fe5:c92%en0***	*UGcIg*	*en0*

That default is not used until an RA received from R2 and the 'I' flag is cleared

default	***fe80::7ba1:d555:9fe5:c92%en0***	*UGcg*	*en0*

The issue has been reported

Conclusion

Rule 5.5 support is already here – it's just not evenly distributed

Linux needs to catch up

NetworkManager (and things like that) are going to complicate things

Next Steps

Test CLAT traffic

Failures are expected though

Other suggestions for testing?