

Turning IPv6 on by default

draft-roy-v6ops-v6onbydefault-01.txt

Sébastien Roy

Alain Durand

Jim Paugh

Disclaimer

- What you are going to see is the result of actual experiments made in the context of real deployment.
- We do not claim that the issues we raised are universal. There may be deployment scenarios where those issues are moot.
- The name of the culprits have been changed to protect the innocent.

Where do we come from?

- We come from a world where breaking existing technology (IPv4) when introducing new technology (IPv6) is not something we like.

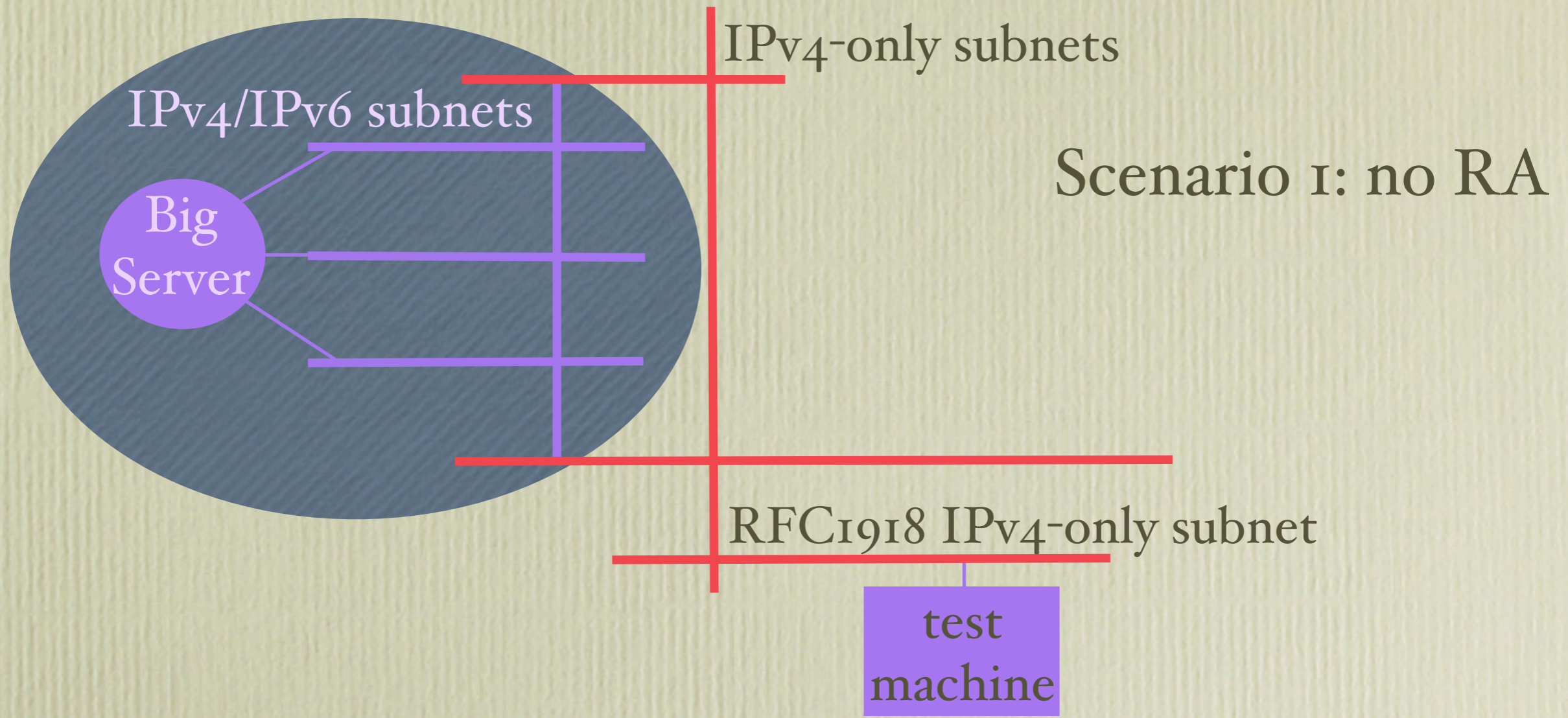
Introduction

- Partial IPv6 deployment in our network
 - 6to4 based, no relay router, no Isatap
 - some subnets are v6 enable, some not
 - some v4 subnets use private addresses
 - some access to the enterprise network are through IPv4-only VPN
 - IPv6 addresses are published in the internal DNS.
- (Big) Internal server answering v4 & v6
 - 17 interfaces! => 17 AAAA & 17 A records
- “client nodes” are v4-only or dual stack

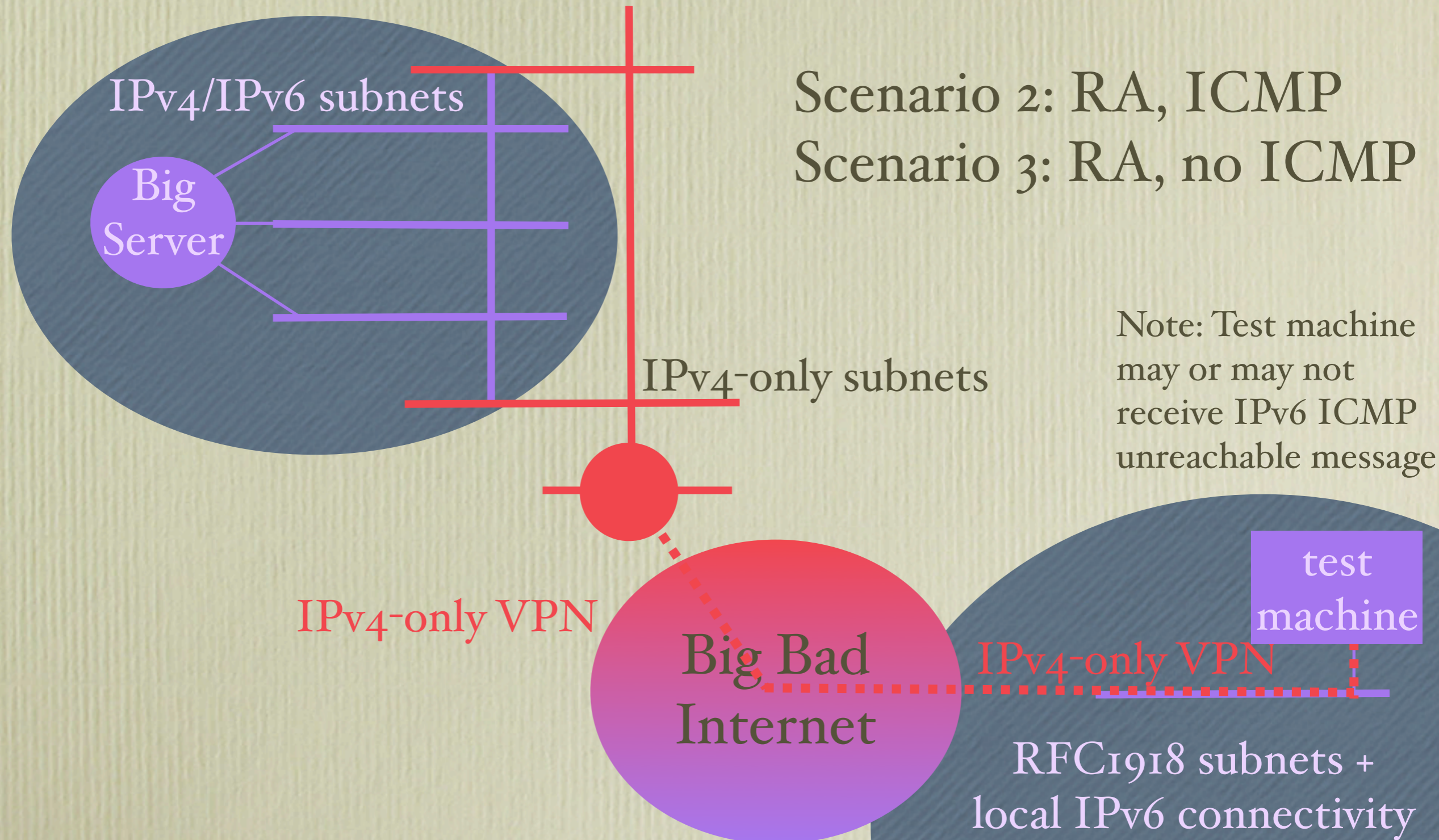
Methodology

- We took 3 well-known implementations with IPv6 turned ON.
- We put them on several network with different “interesting” IPv6 characteristics and RFC1918 IPv4 addresses.
- On each implementation, we tried: “telnet BIGserver”.
- We measured the time it took to fall back from one address to the other.
- Total fallback time was 17 times this number.

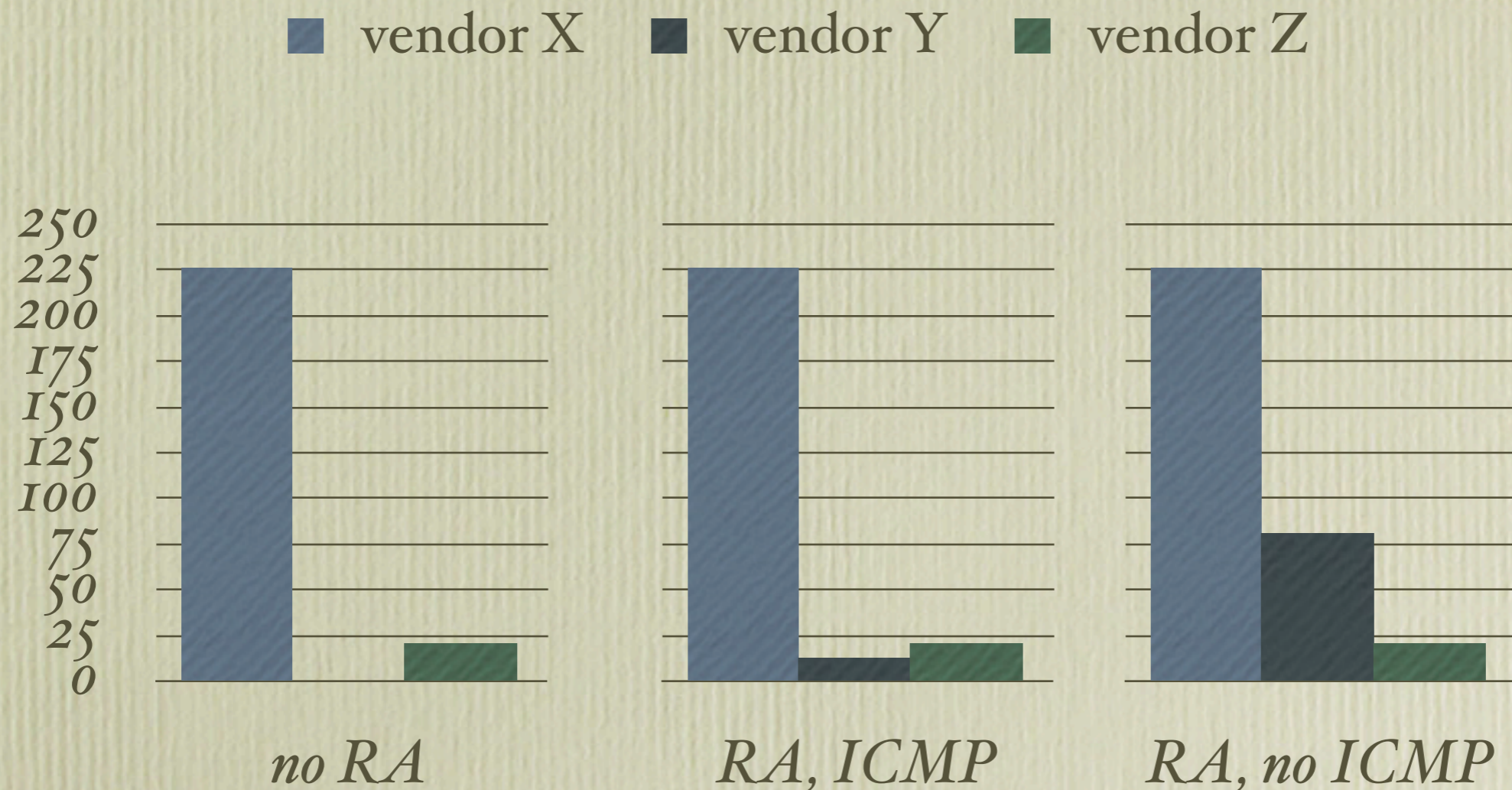
IPv6 isolated node



Split IPv6 connectivity

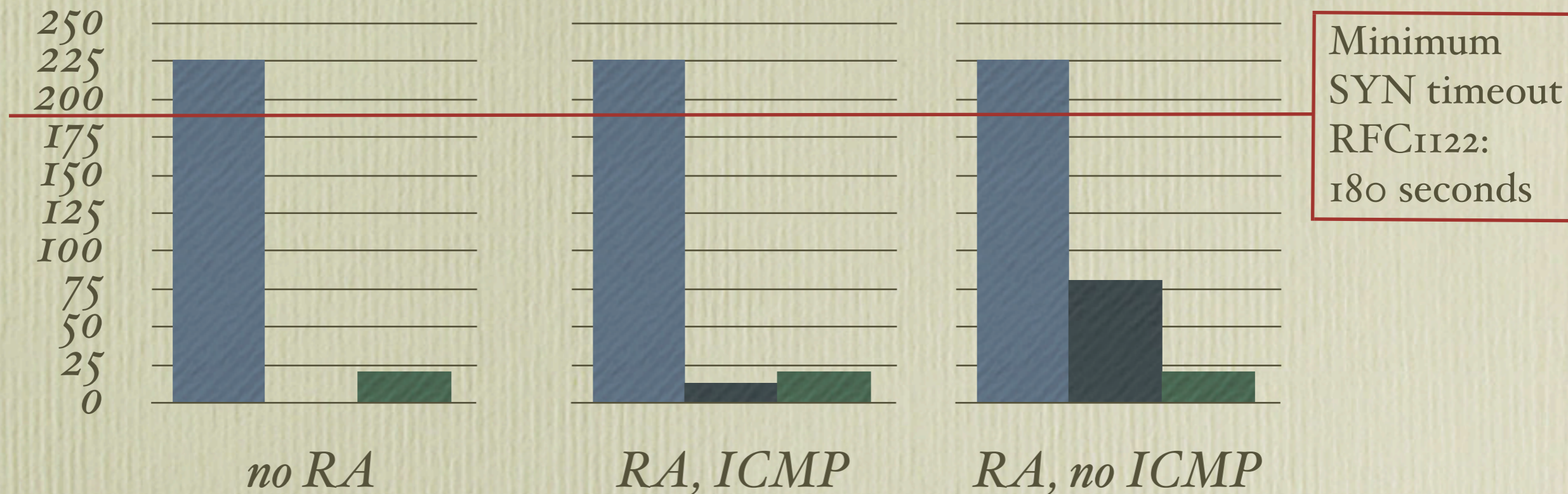


Measured timeouts



Time to fall back from on address to the next

Timeouts origin



ND on-link assumption
default address selection

RFC1122: TCP MUST
ignore ICMP while in
SYN state

RFC1122: TCP initial
timeout

Issues with specs /1

- RFC2461 on Neighbor Discovery (5.2)
 - If the Default Router List is empty, the sender assumes that the destination is on-link.

Issues with specs /2

- RFC3484 Default Address Selection (6)
 - Rule 1: Avoid unusable destinations

If DB is known to be unreachable or if Source(DB) is undefined, then prefer DA. Similarly, if DA is known to be unreachable or if Source(DA) is undefined, then prefer DB.

Issues with specs /3

- RFC1122 on TCP (4.2.3.5)
 - However, the values of R_1 and R_2 may be different for SYN and data segments. In particular, R_2 for a SYN segment **MUST** be set large enough to provide retransmission of the segment for at least 3 minutes. The application can close the connection (i.e., give up on the open attempt) sooner, of course.

DISCUSSION:

Some Internet paths have significant setup times, and the number of such paths is likely to increase in the future.

(our) Conclusions /I

- The “no RA” case should be “zero delay”
Something in the IPv6 specs needs to be fixed:
 - ND: do not make the “on-link” assumption” or
 - Addr Select: add rule 2.5
“avoid link-local source address when destination is non link local”
Note: this solution does not help if literal addresses are used.

(our) Conclusions /2

- In case of incomplete v6 connectivity, there are transport level issues:
 - e.g. TCP-SYN assumption
 - Remark: not all vendors respect RFC1122
- Be extra careful before putting AAAA in your DNS....

(our) Conclusions /3

- Use v6-aware VPN, if not:
 - TCP timeout delays
 - **security issues:** the v6 packet does not go where you believe it is!
 - VPN should intercept/sink IPv6 packets
- We have identified a number of other less critical issues (deployment BCP), please read the draft.

Additional materials

Destination Address Ordering

ND “on-link” assumption makes this irrelevant

SRC6= LL, DST6=GL vs SRC4=SL, DST4=GL

Here we prefer v6 over v4

- Rule 1: Avoid unusable destinations.
- Rule 2: Prefer matching scope.
- Rule 2.5: Avoid LL SRC when DST is not LL
- Rule 3: Avoid deprecated addresses.
- Rule 4: Prefer home addresses.
- Rule 5: Prefer matching label.
- Rule 6: Prefer higher precedence.
- Rule 7: Prefer native transport.
- Rule 8: Prefer smaller scope.
- Rule 9: Use longest matching prefix.
- Rule 10: Otherwise, leave the order unchanged.