IPv6 Neighbor Discovery DeepDive

Jen Linkova aka Furry
furry13@gmail.com
Nov 2015

Most Important Slide of This Talk

IPv6 is

NOT

"like IPv4 but longer addresses"

What Can Be Improved in IPv6?

- Just one control protocol?
  - IPv4: ARP for L2, ICMP for Internet
- Less chatty protocol?
  - ARP using broadcast
- do more than just map IP <> L2 addresses?
  - ARP does not confirm reachability

Protocol Choice

- ICMPv6 is already here as a control protocol
- No reason to use non-IP protocol
- Flexible
  - new message types can be created
- Like in ARP, 2 messages are needed:
  - **Neighbor Solicitation**, "Who has this IPv6 address"?
  - **Neighbor Advertisement**, "I have this IPv6 address"

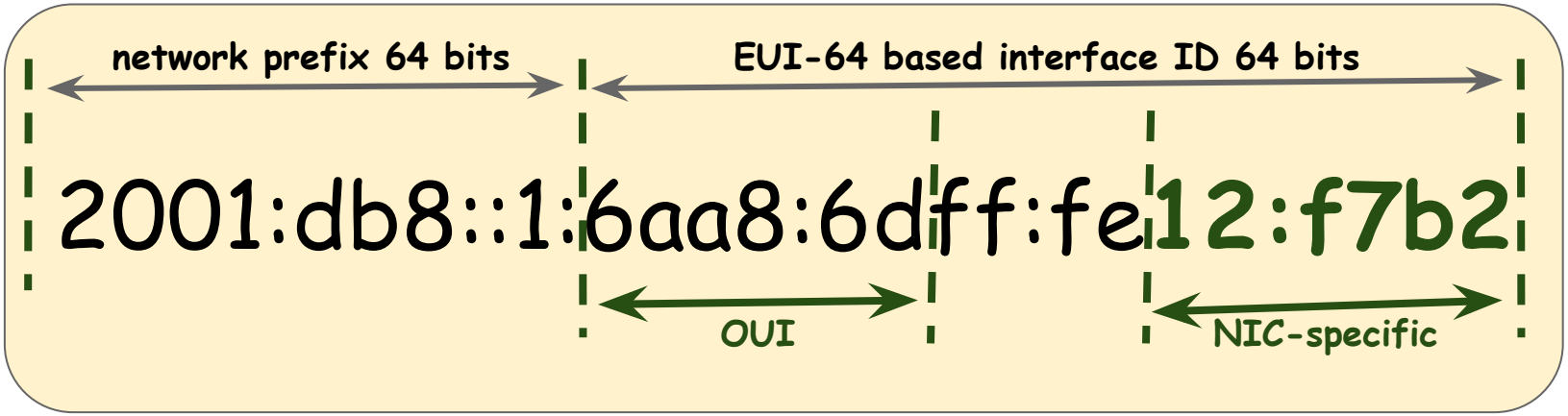
ARP Is Too Chatty

- Broadcast
  - creates noise
  - consumes network resources
  - Kills the battery on mobile devices
- NS message should be received by?
  - all hosts? => broadcast 
  - or hosts which might have that address? 

MULTICAST

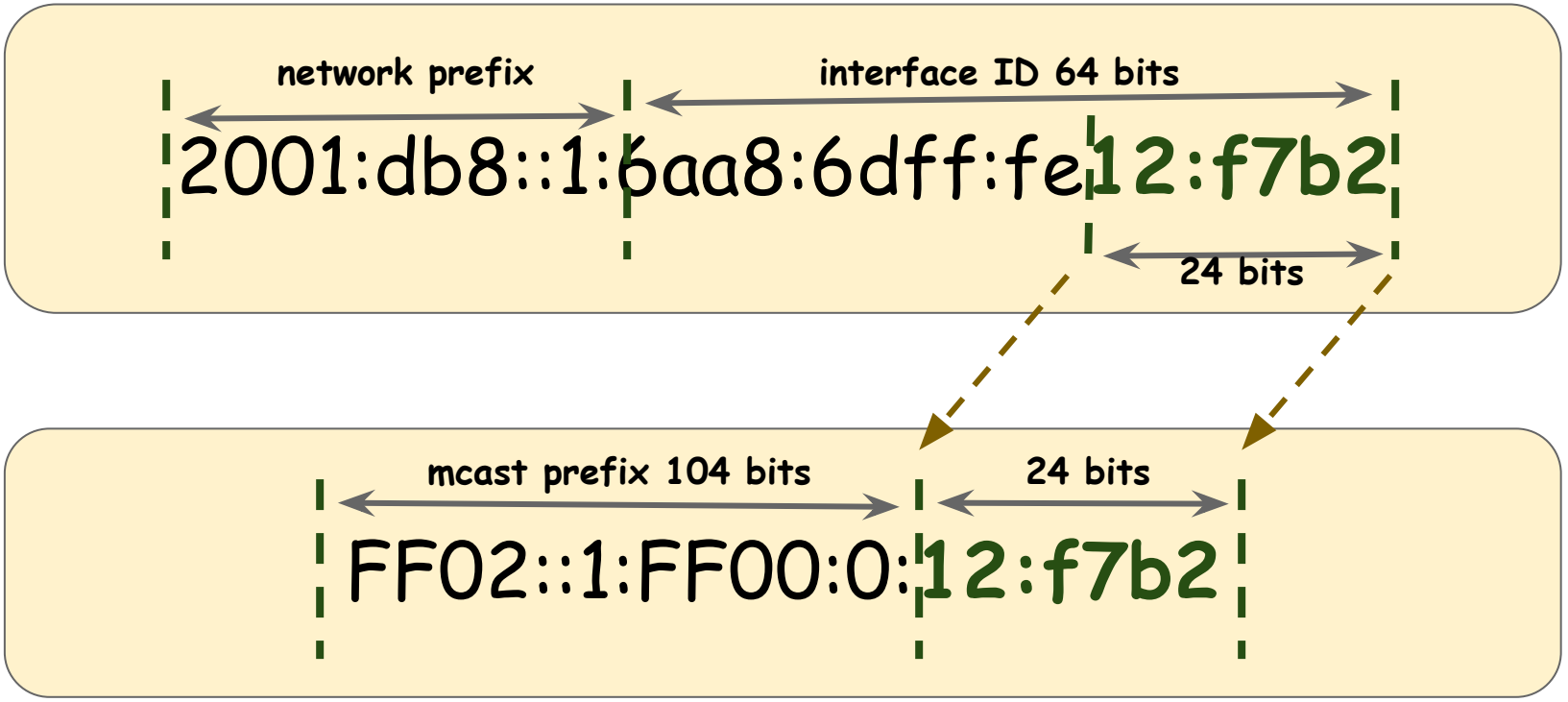
Multicast Group Address

- What about: one IPv6 address - one mcast group (ff02::$interface_id)?
  - Too many mcast groups
  - Local resolution to L2 address - only 32 bits of L3 addresses going to L2 mcast address
  - EUI-64: highest 24 bits are OUI, then fffe
  - Manually assigned interface_id: no highest bits set



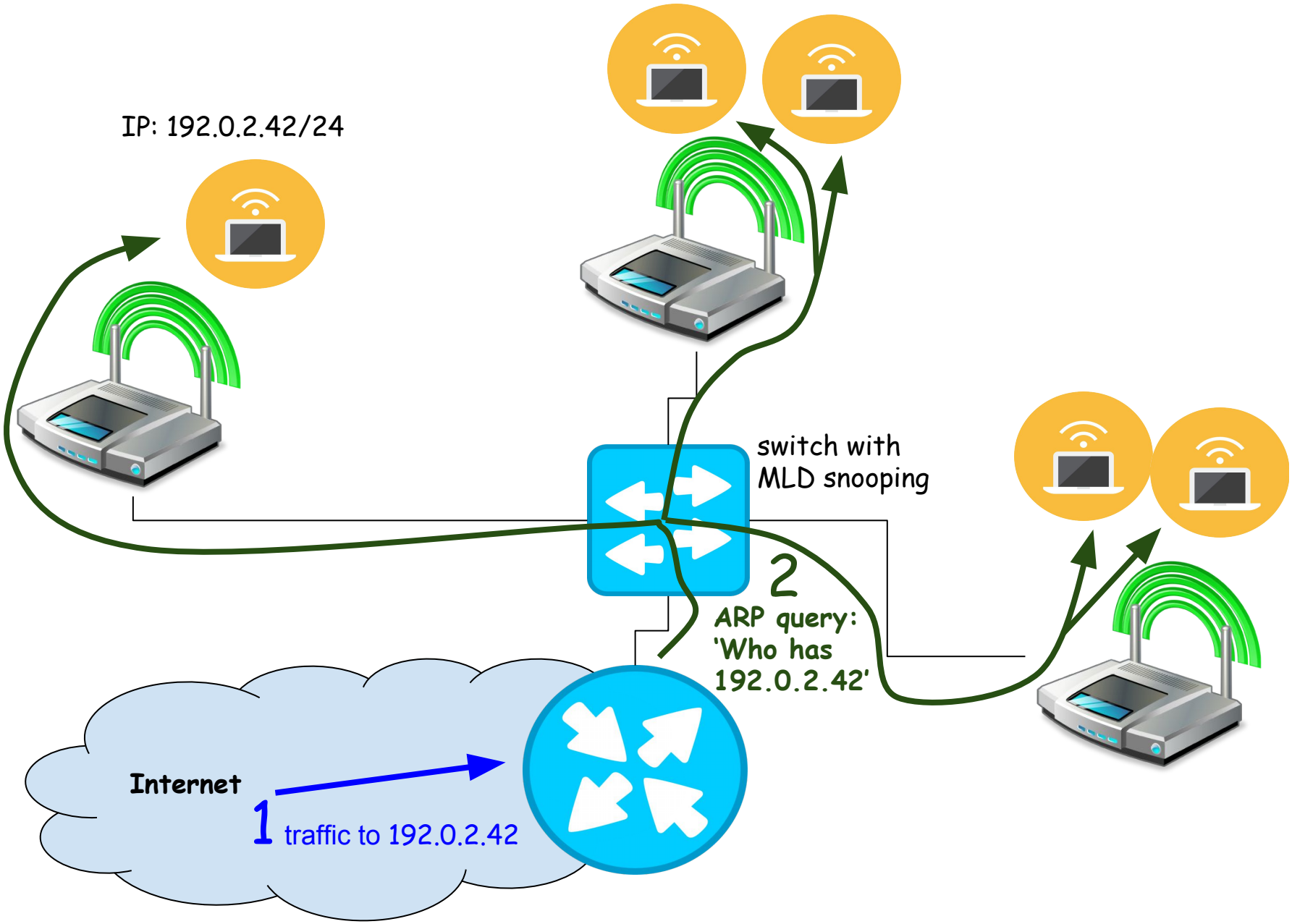
Solicited Node MCast Address

- Lower 24 bits of IPv6 address
- Solicited-node multicast address format:
  - Globally-assigned prefix FF02::1:FF00:0:/104
  - low-order 24 bits of a node address

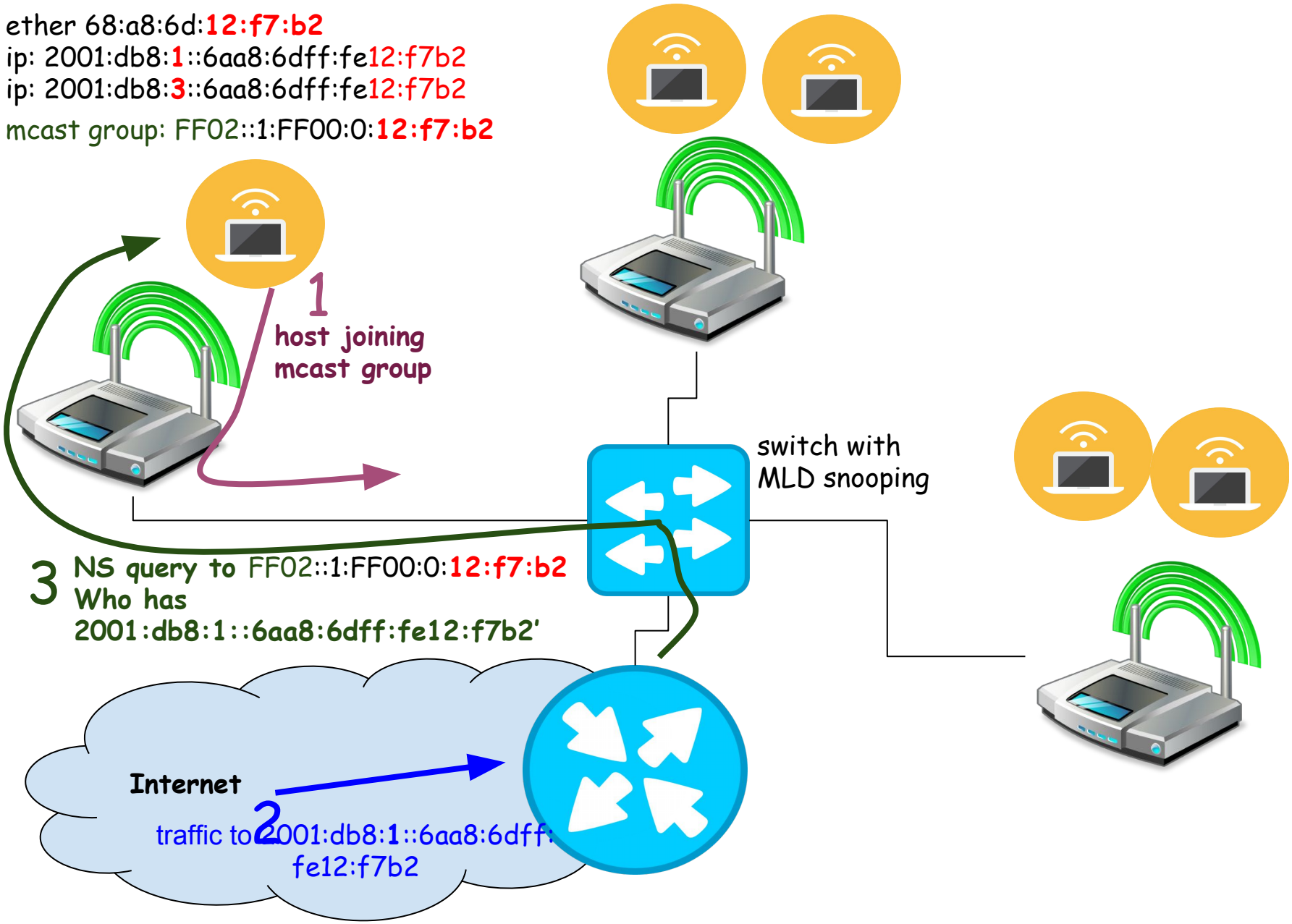


solicited node multicast address

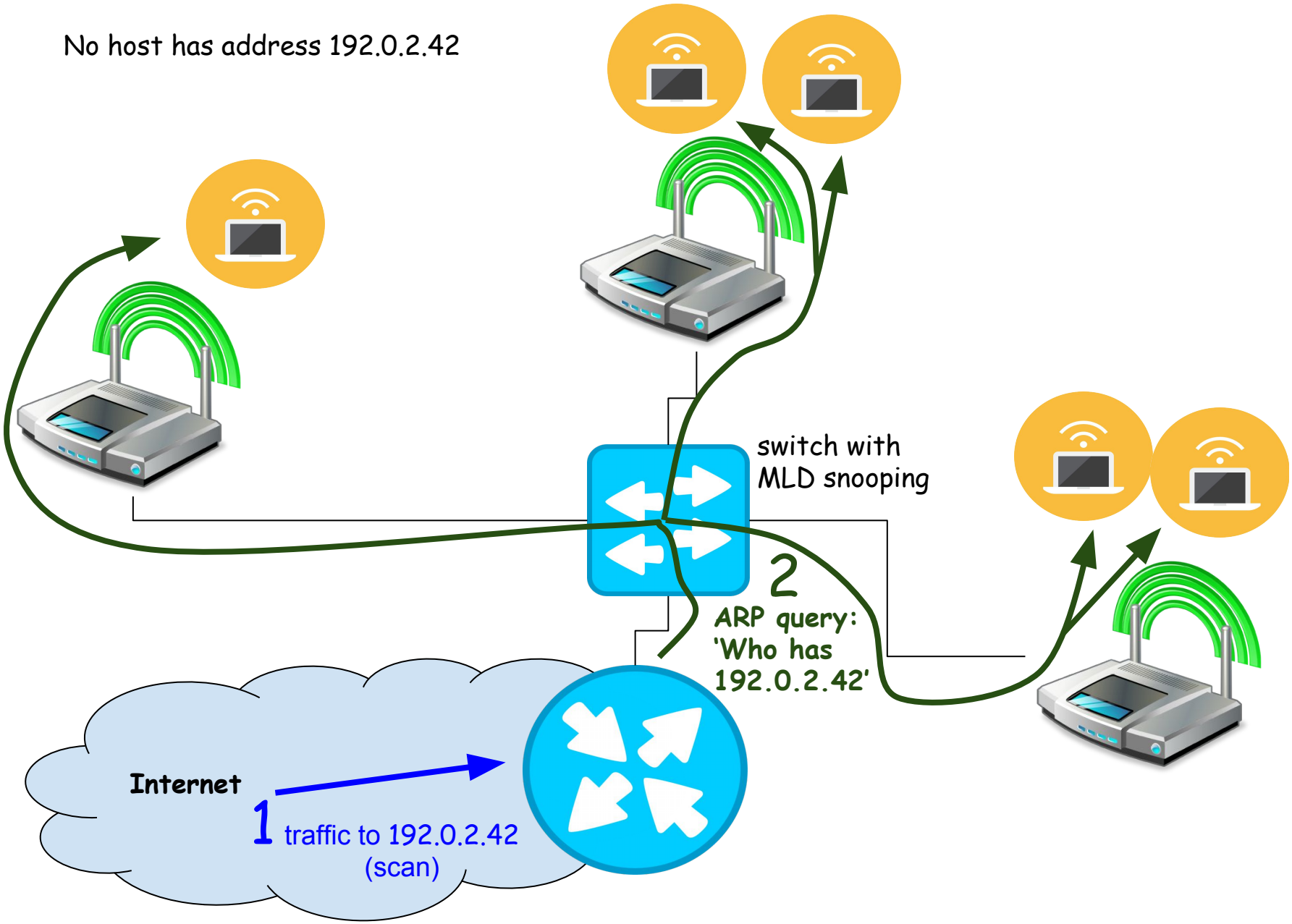
IPv4/ARP: WiFi



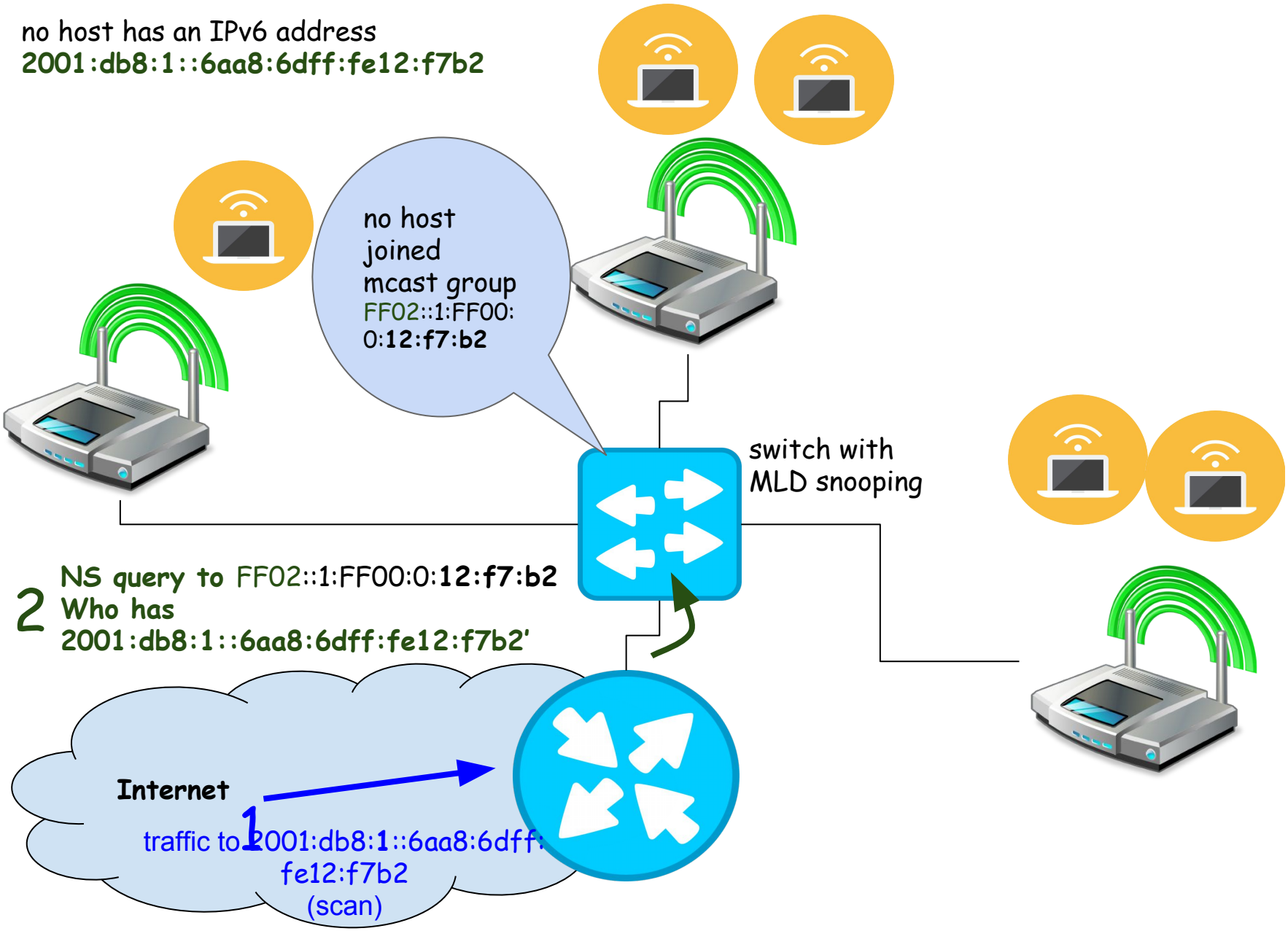
Multicast, ND and MLD: WiFi



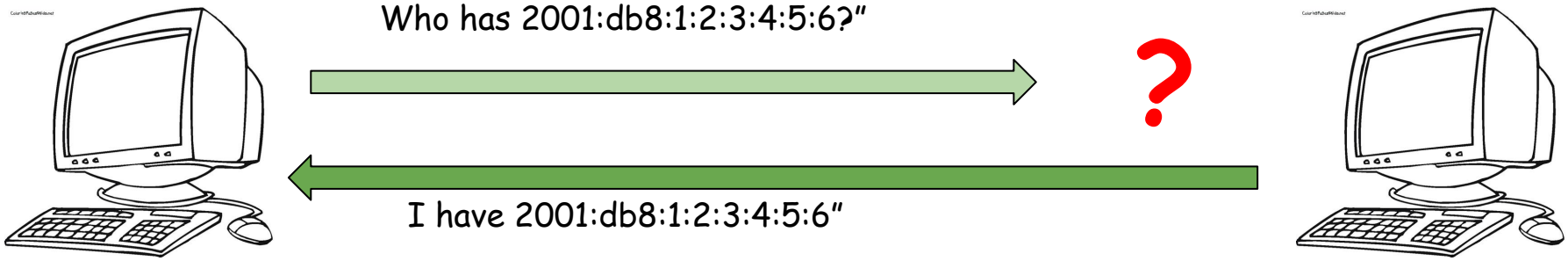
Non-Existent Address & WiFi: V4



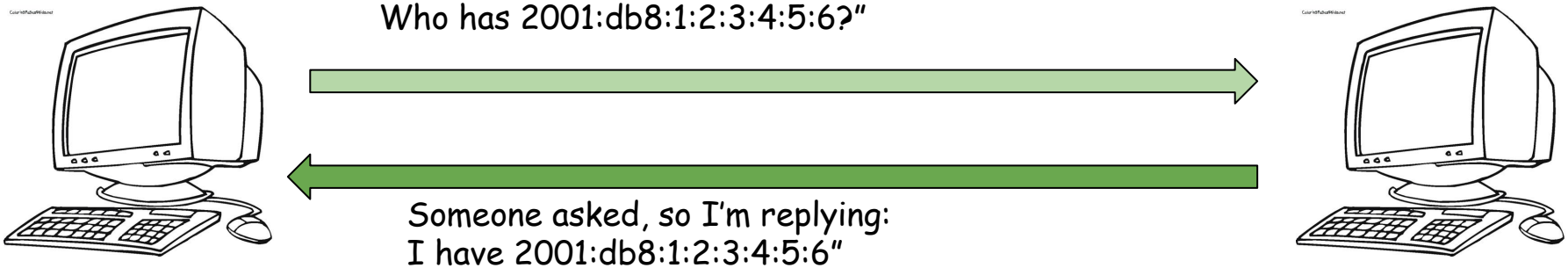
Non-existent Address and WiFi: V6



Improve ND Robustness (1)

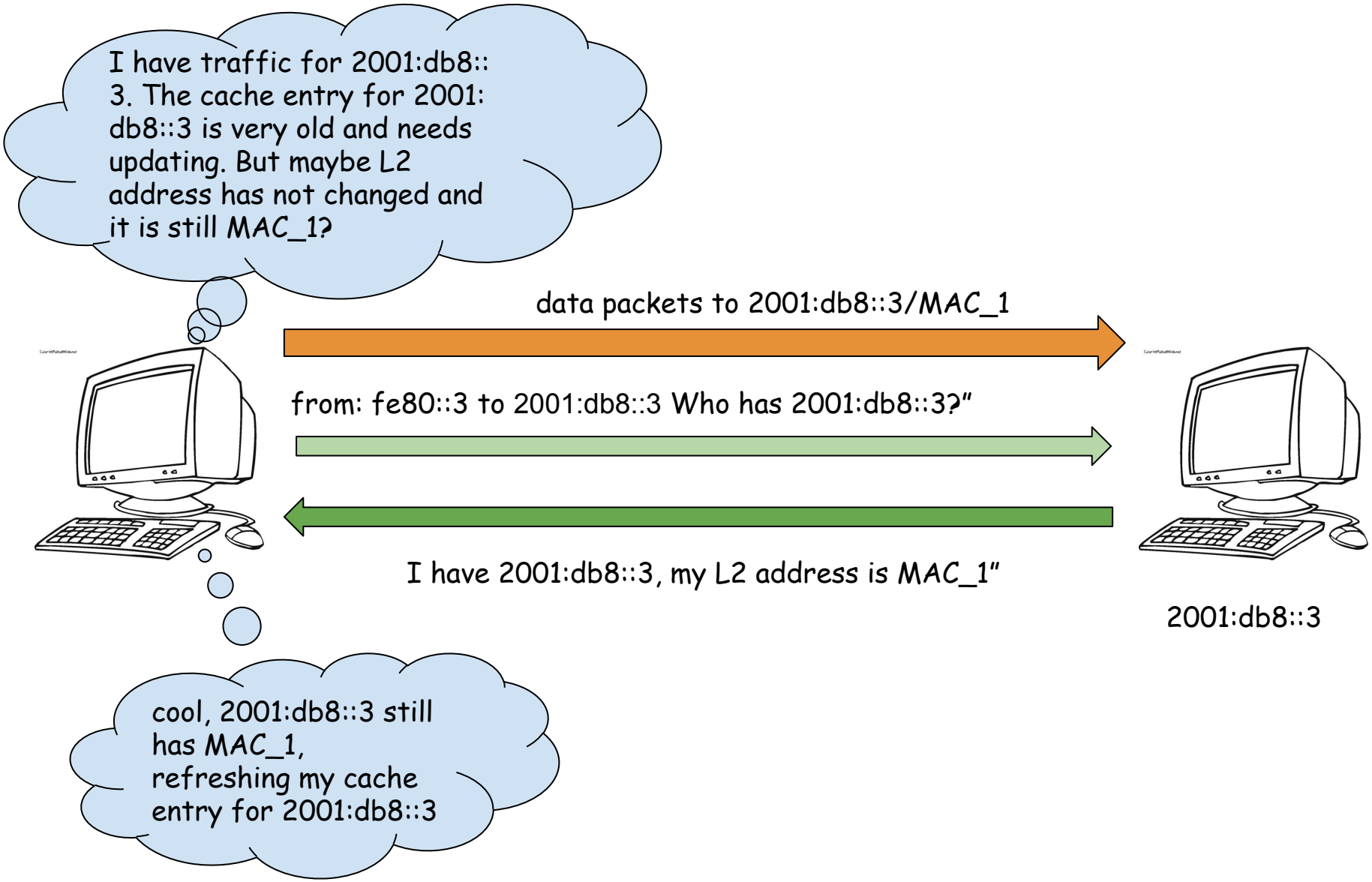


one-way communication possible

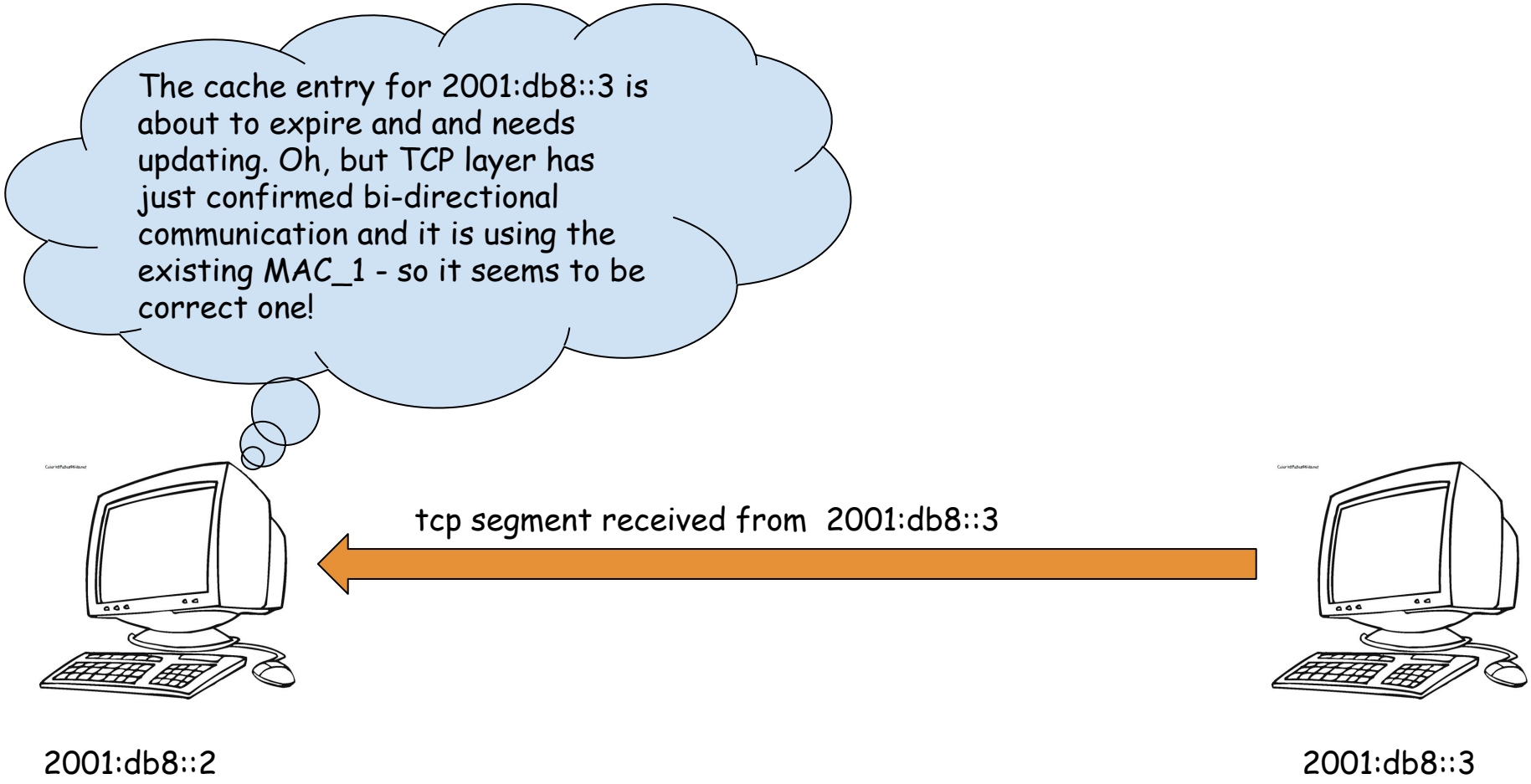


most likely communication is bi-directional

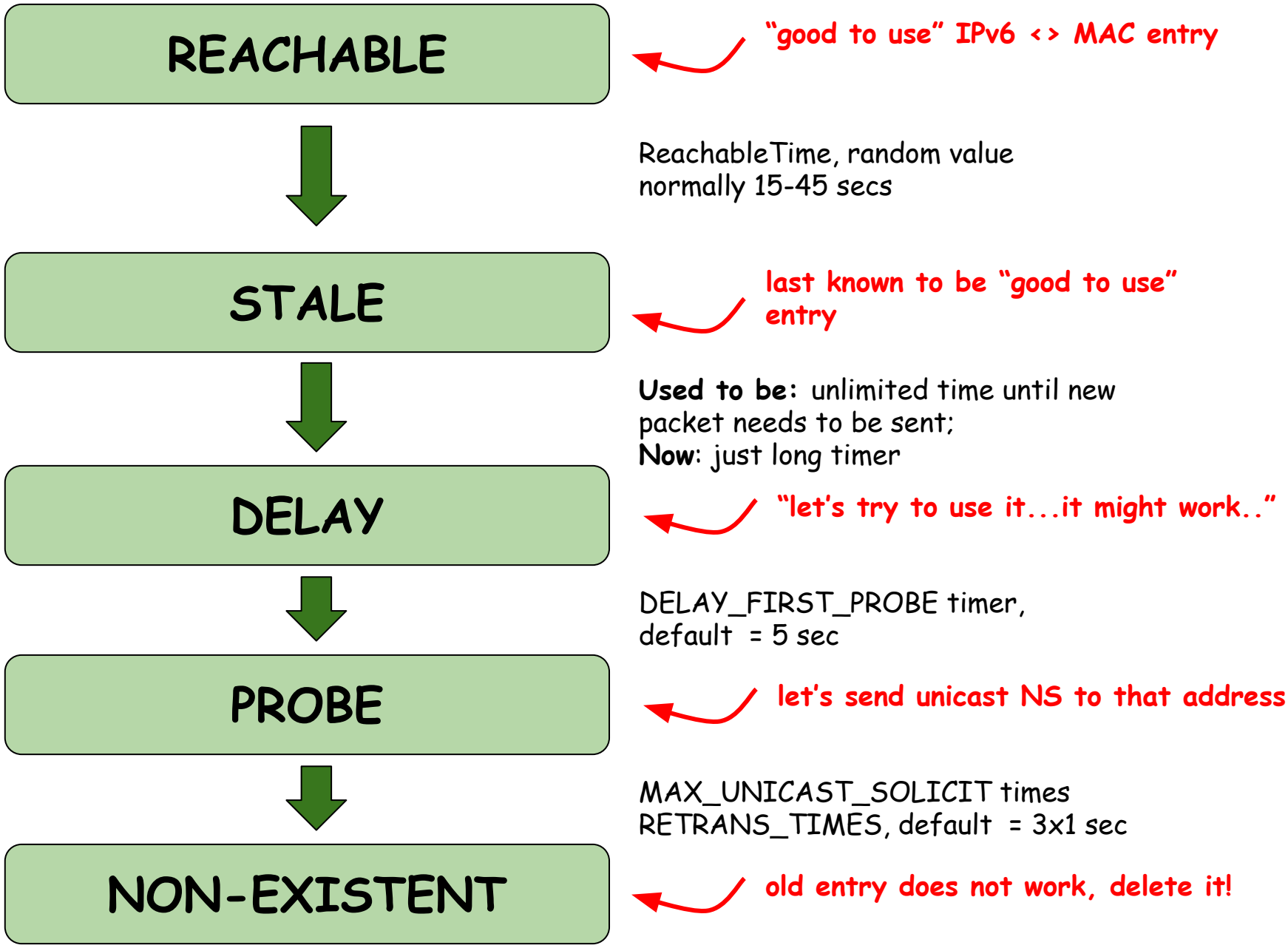
Improve ND Robustness (2)

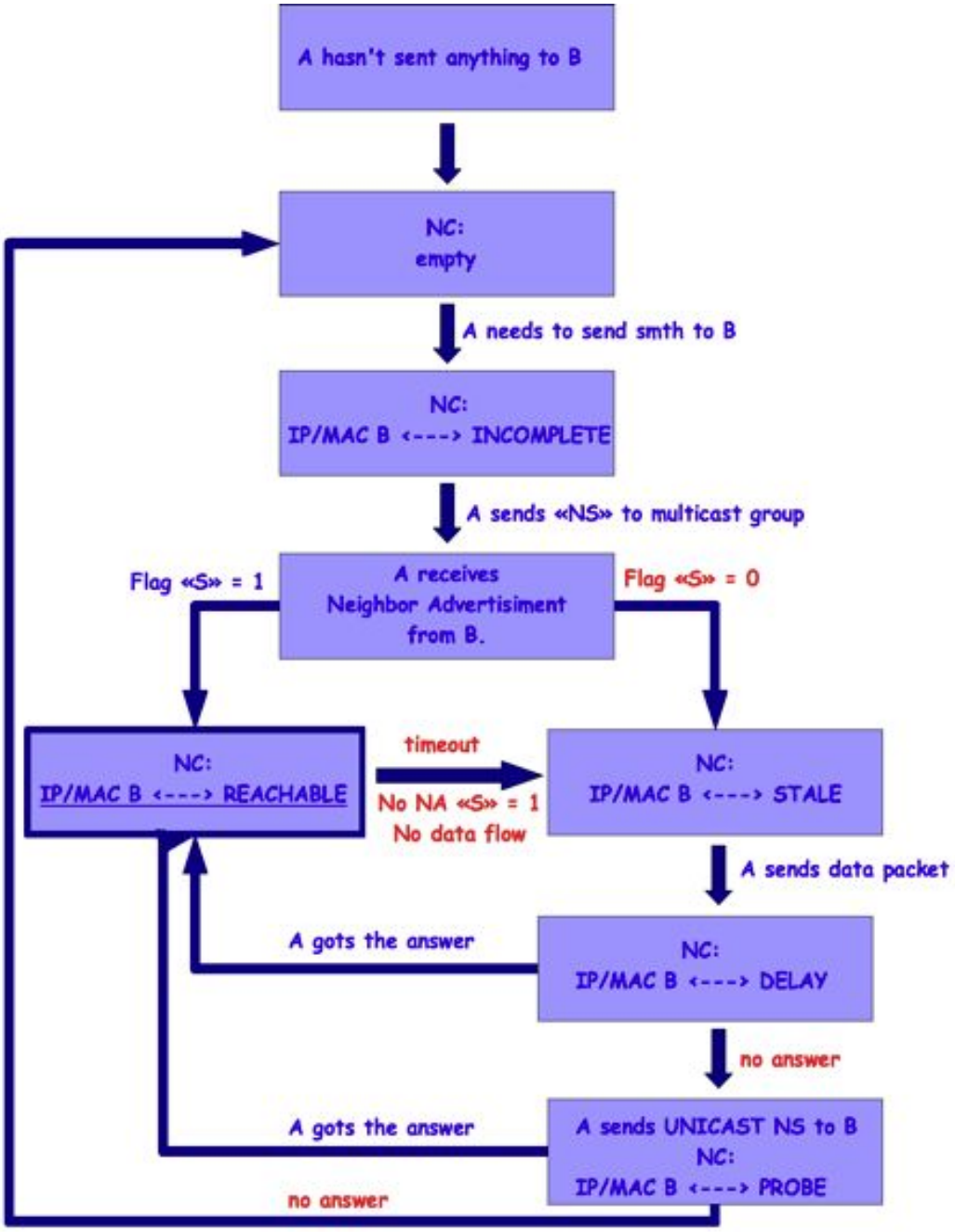


Upper-Layer Integration



Neighbor Caches Entry Lifecycle




A hasn't sent anything to B
NC:
empty
A needs to send smth to B
NC:
IP/MAC B <---> INCOMPLETE
A sends «NS» to multicast group
A receives
Neighbor Advertisiment
from B.
Flag «S» = 1
Flag «S» = 0
NC:
IP/MAC B <---> REACHABLE
timeout
No NA «S» = 1
No data flow
NC:
IP/MAC B <---> STALE
A sends data packet
A gots the answer
NC:
IP/MAC B <---> DELAY
no answer
A gots the answer
A sends UNICAST NS to B
NC:
IP/MAC B <---> PROBE
no answer

IPv6 Autoconfiguration

How to Configure a Host

- Manual configuration
- Controlled by external system (DHCP)
  - stateful or stateless
- **S**tateLess **A**ddress **A**uto-**C**onfiguration (SLAAC)
  - network prefix
    - routers know prefixes configured
    - well-known link-local fe80::/10
  - host can generate interface id:
    - EUI-64
    - random
    - other options

Interface ID: Avoiding Duplicates

interface id generated by a host might not be unique

before using an address - check for duplicates

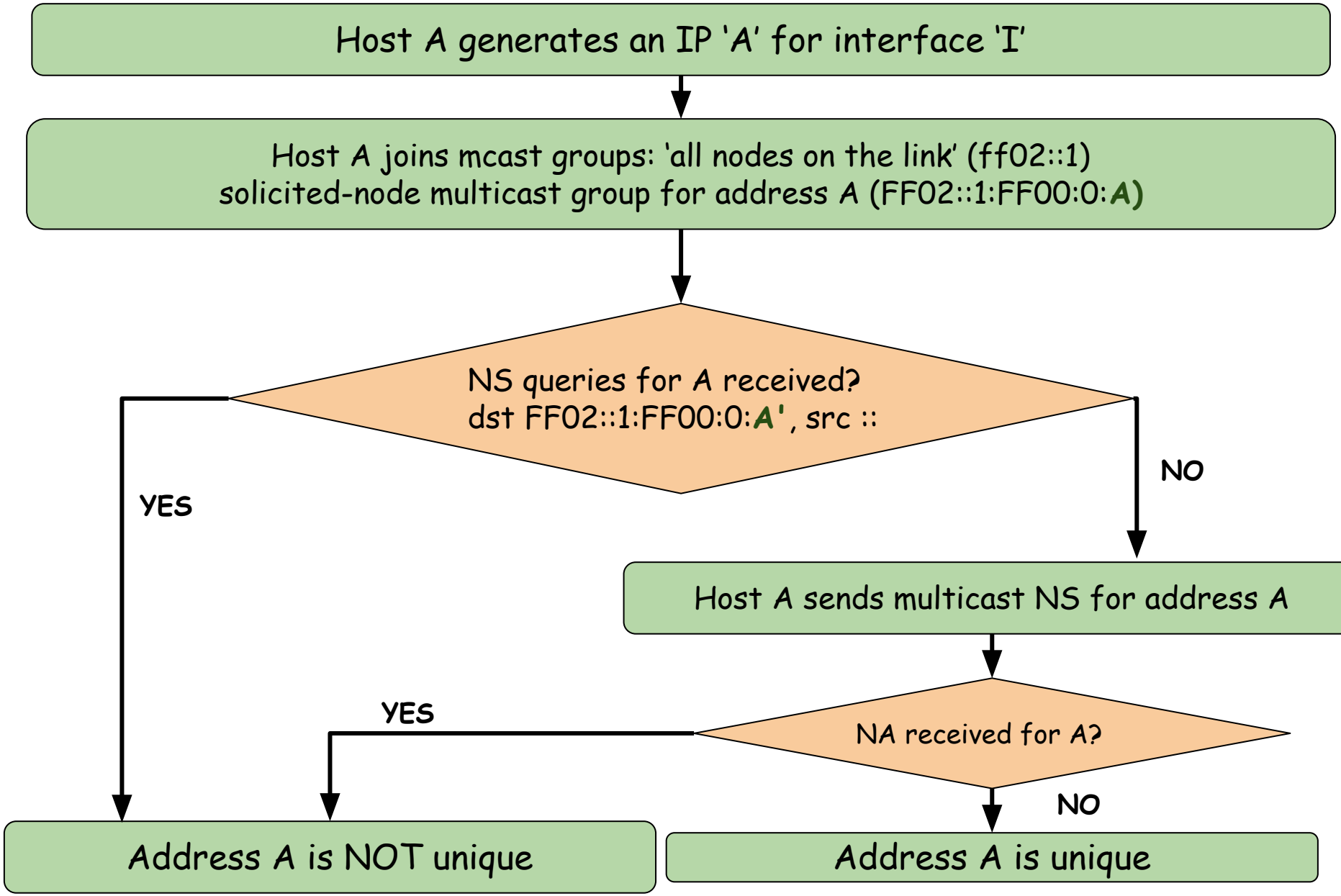
???HOW???

host needs to know if there is another owner for that IP

Neighbor Discovery locates owners of IP addresses

Duplicate Address Detection

Duplicate Address Detection (DAD)



Duplicate Address Detection (contd)

- DAD **MUST** be performed for all addresses (ex. anycast)
- Heavily relies on multicast so false negatives might happen:
  - on busy wifi networks
  - if multicast is broken

After DAD failure manual actions are required.

Prefix and Router Discovery

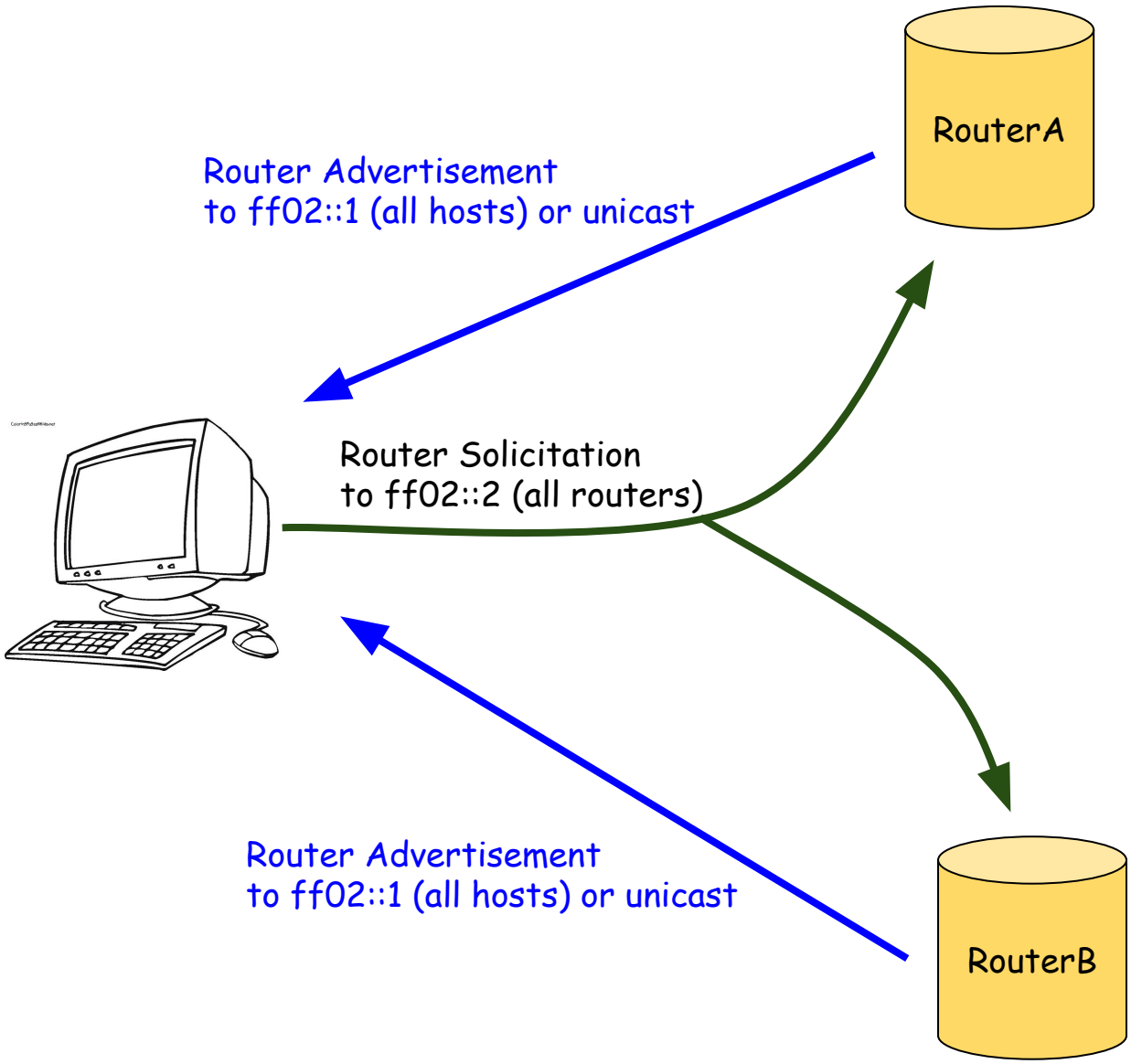
Critical Info to Discover

- **Network prefix (to complete the address configuration)**
- **Default router(s)**
- Routes to some destinations
- DNS servers and search list

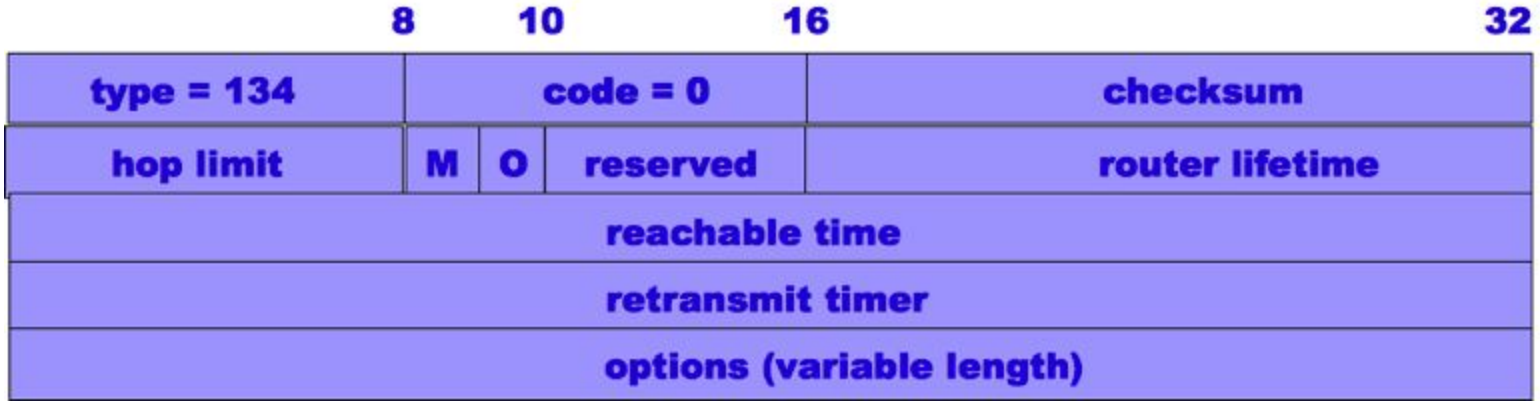
Routers have most of this info

Router are neighbors => ND can be used!

Router Solicitation and Advertisement



Router Advertisement Message



Src IP = link-local, Dst IP = the source IP of the RS query or FF02::1

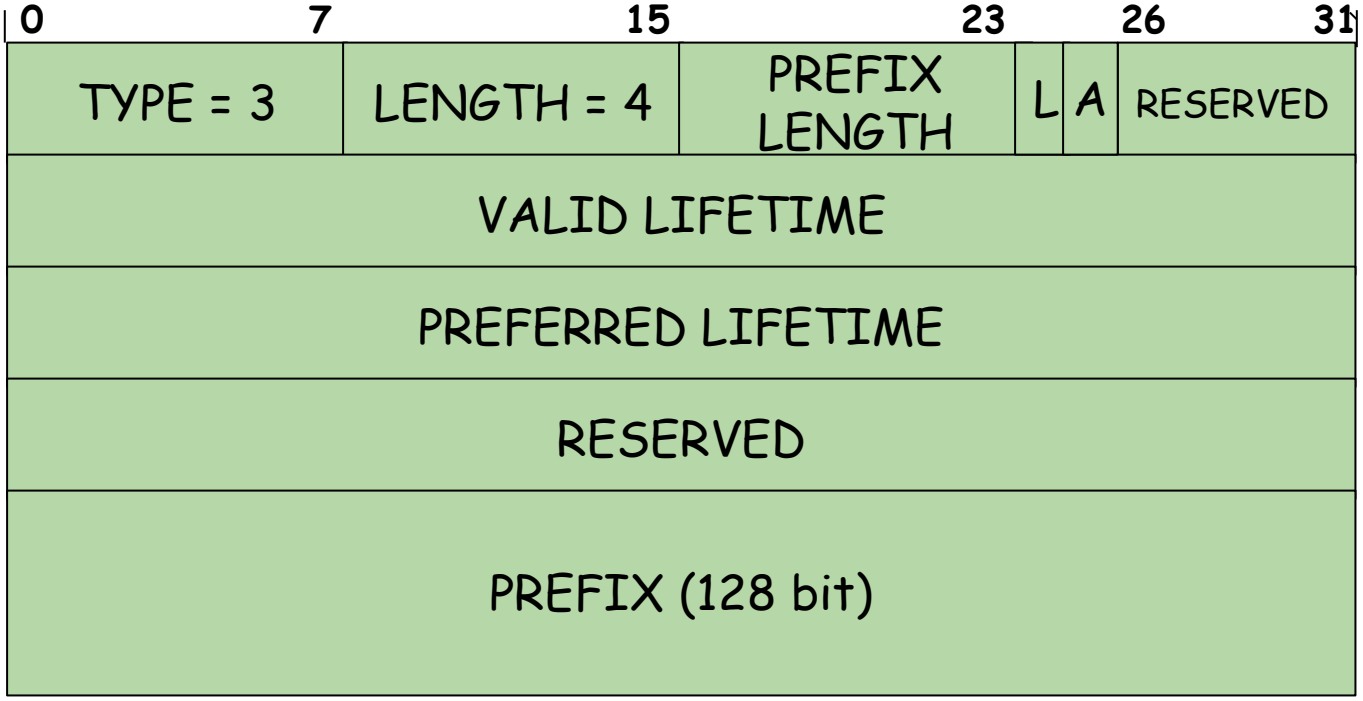
- **M,O flags**: indicate that addresses (M) or other configuration info (O) is available via DHCPv6
- **Router lifetime** (in seconds) – the lifetime associated with the default router (0 - the router isn't default router, shouldn't appear on the default router list)
- **Reachable time** (millisecs) – how long the neighbor is reachable after receiving a reachability confirmation (NC record goes from Reachable -> Stale then)
- **Retransmit timer** (millisecs) – the interval between retransmitted NS messages

RAs: What Can Possibly Go Wrong #1

- Solicited RAs: in response to RS but rate-limited to 1/ MIN_DELAY_BETWEEN_RAS sec (default: 3 secs)
- Large-scale wifi network:
  - every 3 secs
  - kills the battery on mobile devices
- Solution: to send solicited RAs unicast

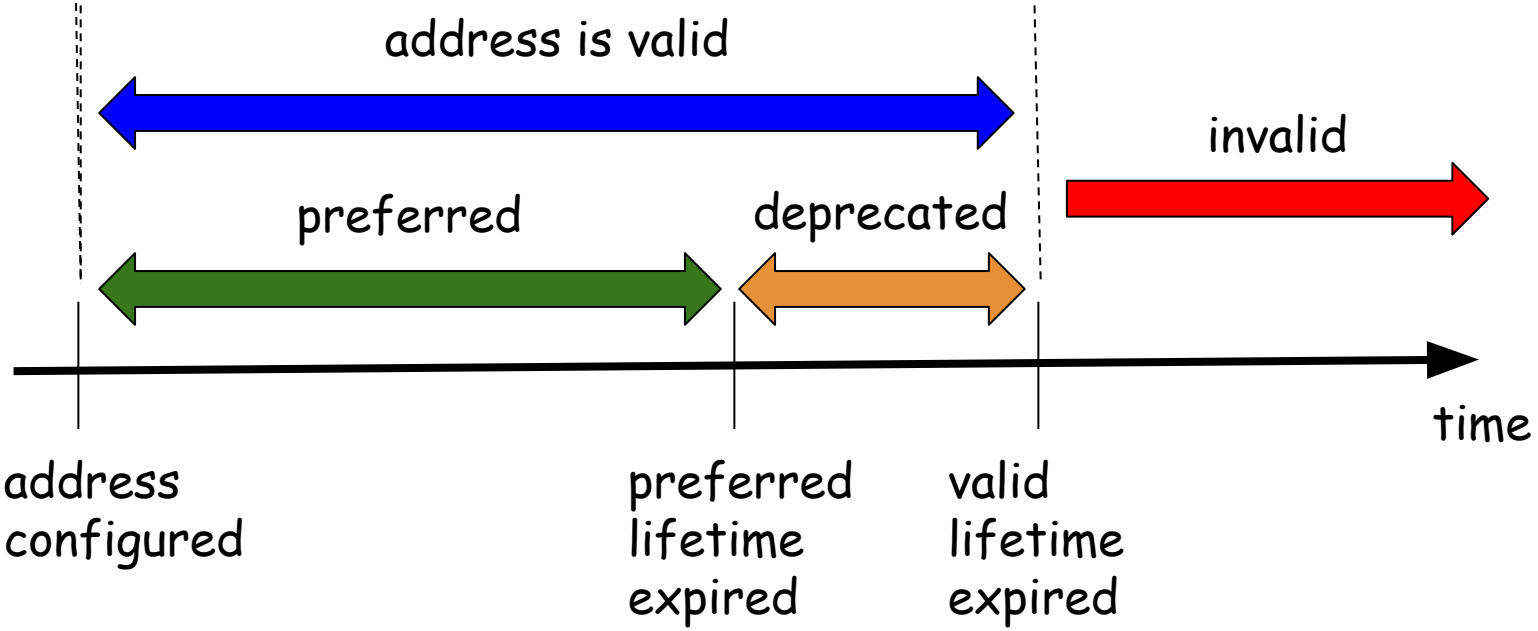
see draft-ietf-v6ops-reducing-ra-energy-consumption

Prefix Information Option (PIO)



- L - prefix is on-link
- A - prefix can be used for SLAAC

It's All About Timers

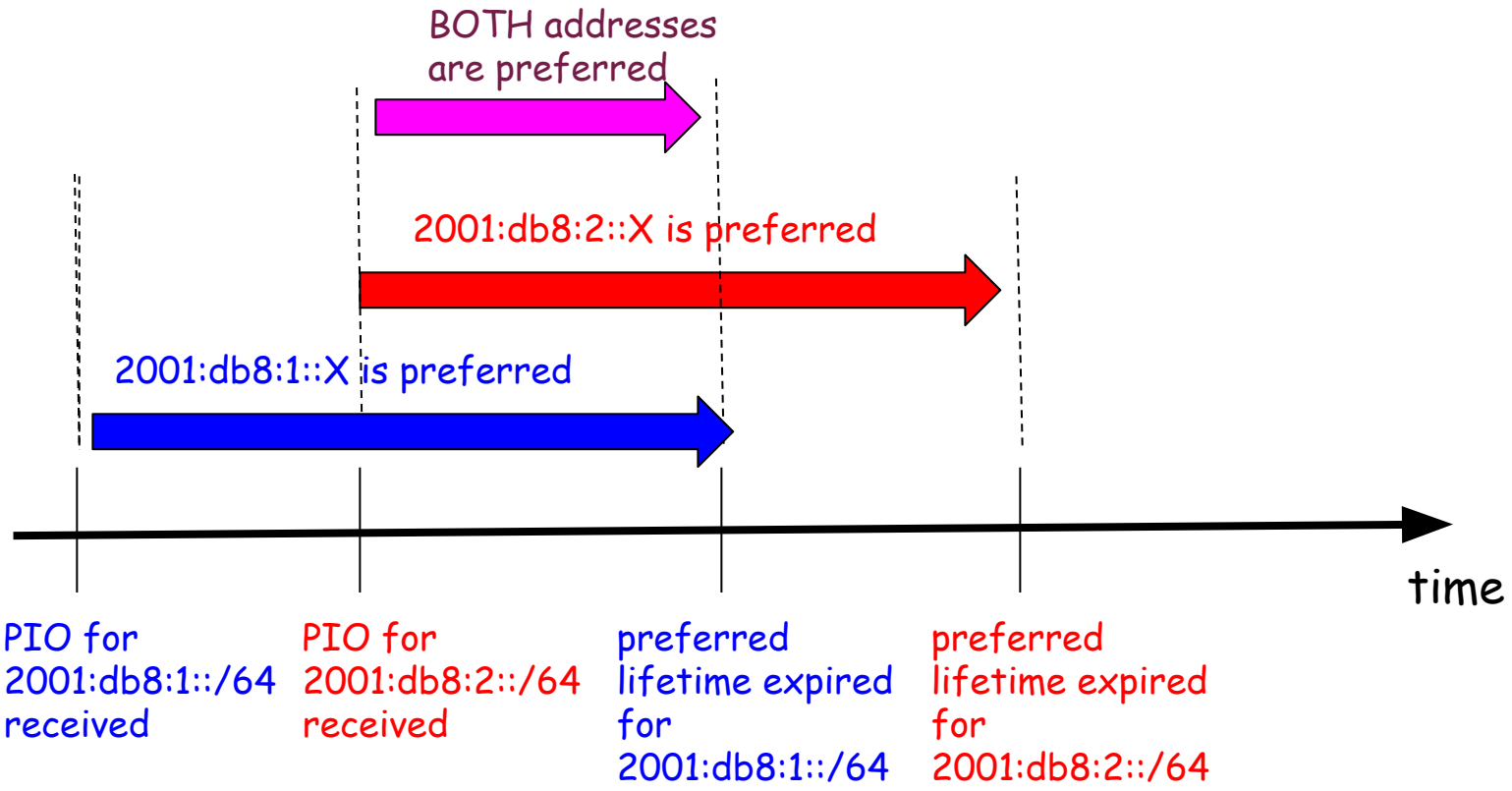


- Deprecated address
  - SHOULD be used for existing communications
  - SHOULD NOT be used for new ones
- Preferred lifetime <= Valid lifetime
- Can not set valid lifetime < 2hrs

Default values:

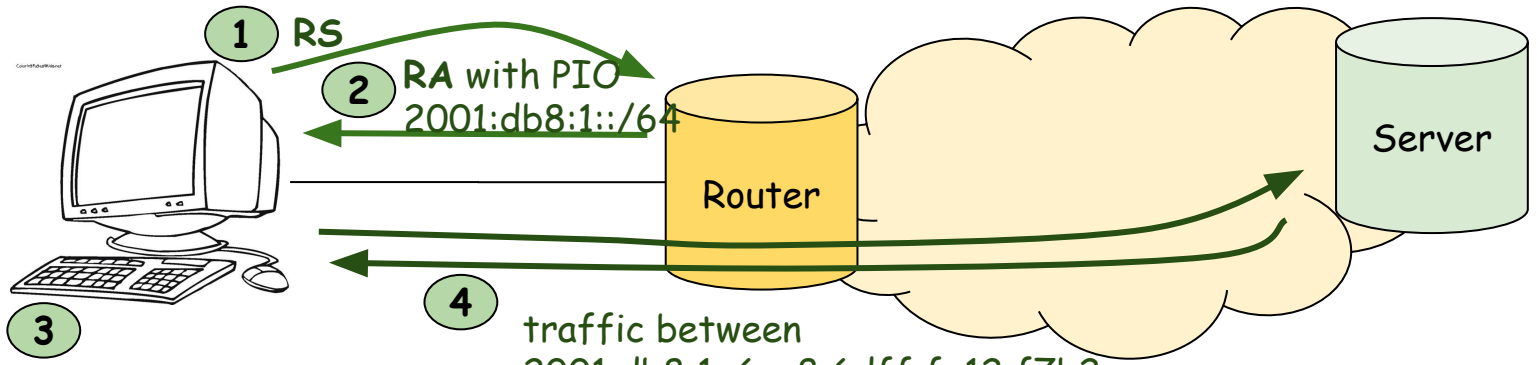
- preferred lifetime - 7d
- valid lifetime - 30d

Multiple PIOs

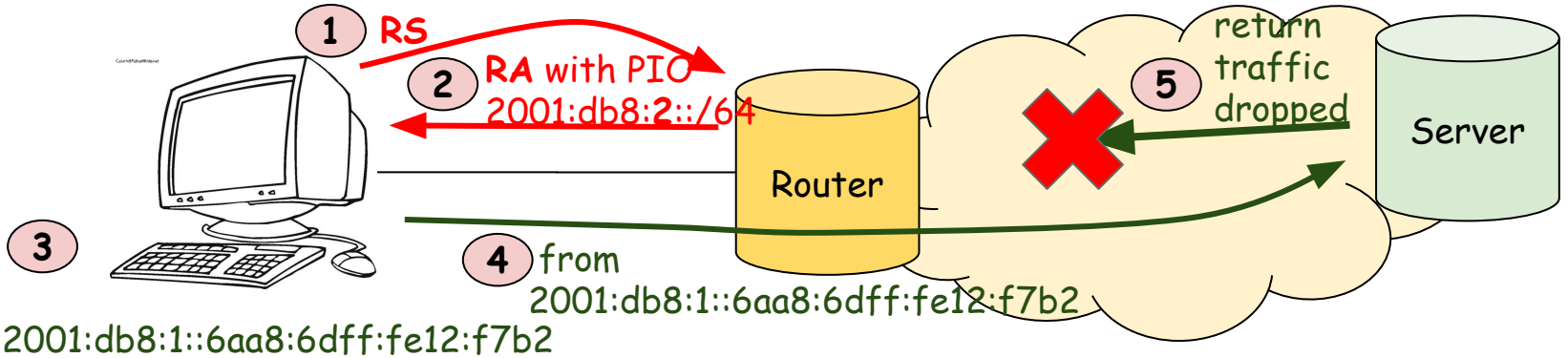


- **Removing a prefix from router configuration DOES NOT mean hosts stop using it**
- To renumber a new RA with PIO with preferred lifetime = 0 needs to be sent

How NOT TO Renumber



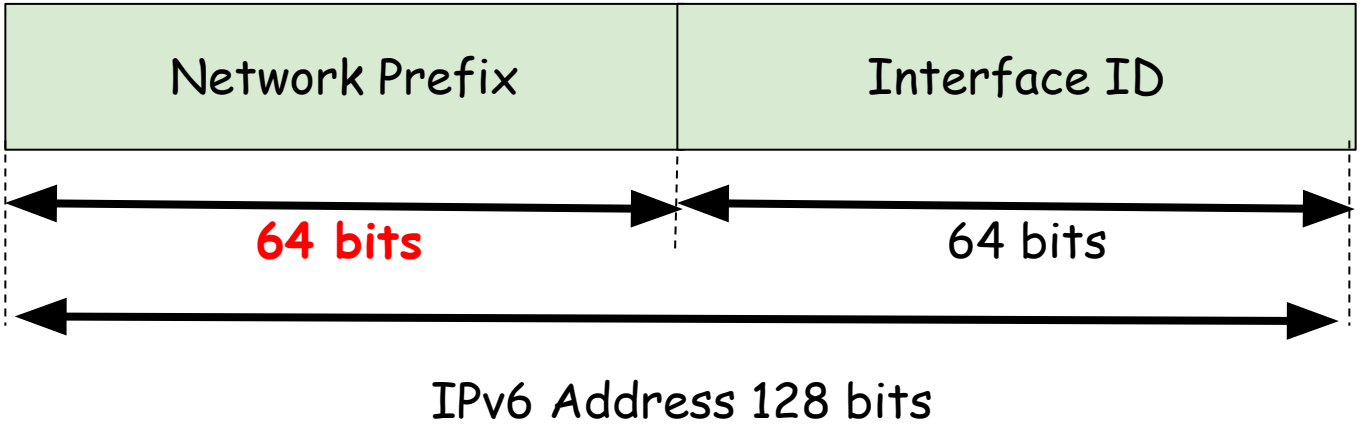

```
# show | compare
[edit protocols router-advertisement interface ae1]

+  prefix 2001:db8:2::/64;

-  prefix 2001:db8:1::/64;
```



Second Most Important Slide

Never
Ever
Use
Prefixlength
between 64 and 127 bytes

RAs, PIOs and Prefix Length



SLAAC **DOES NOT** work with PIOs if prefix length !=64

What the prefix length field is for then?

The Host, the Link, and the Subnet

What Does "Subnet Mask" Mean

- IPv4: IP address + netmask => on-link prefix
- IPv6: On-link prefix & address assignment are separated!
  - Link-local – always on-link
  - Any other addresses - if explicitly told so

See "IPv6 Subnet Model: The Relationship between Links and Subnet Prefixes", RFC5942

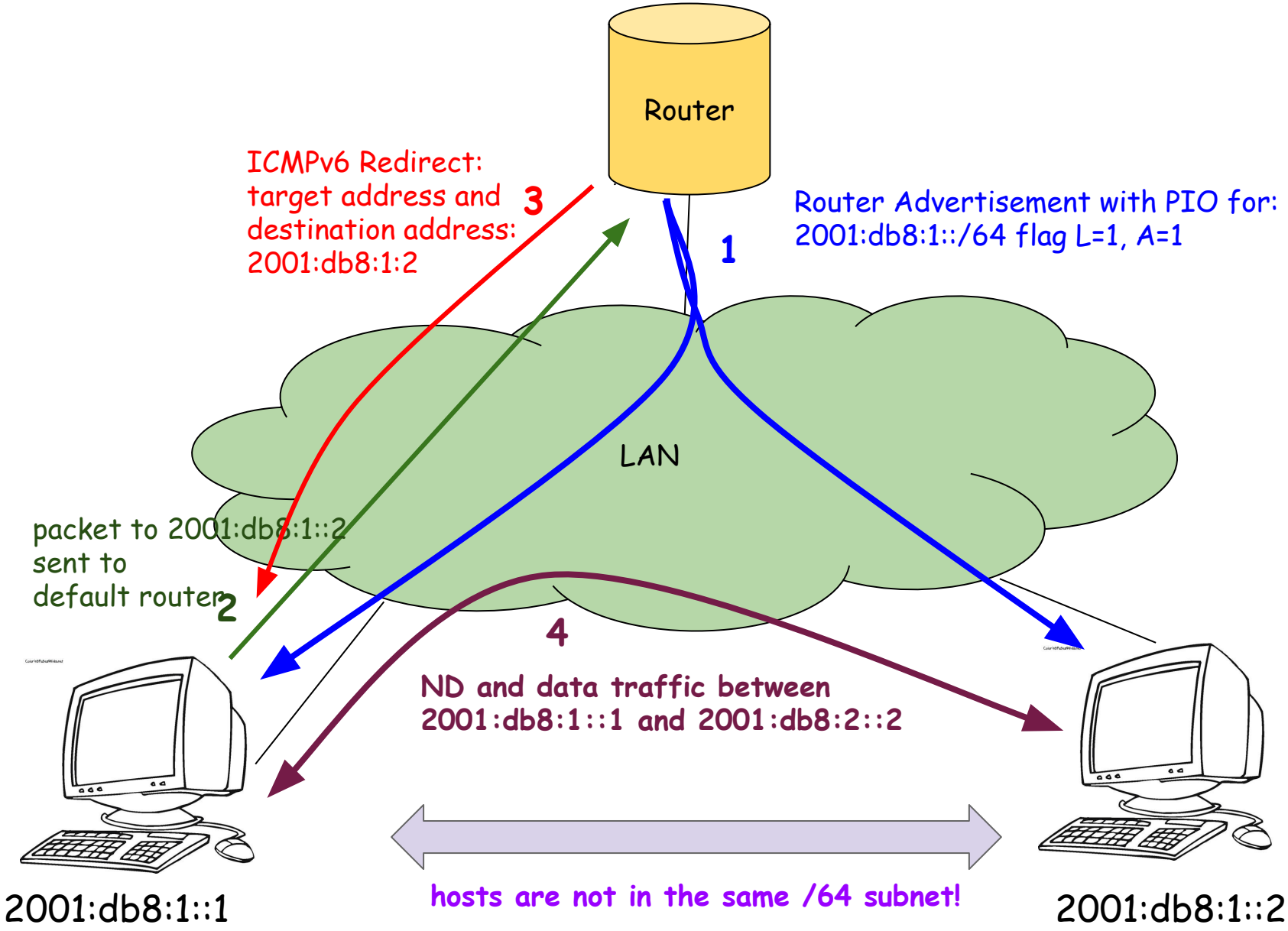
On-link Prefix List

Prefix is on-link if

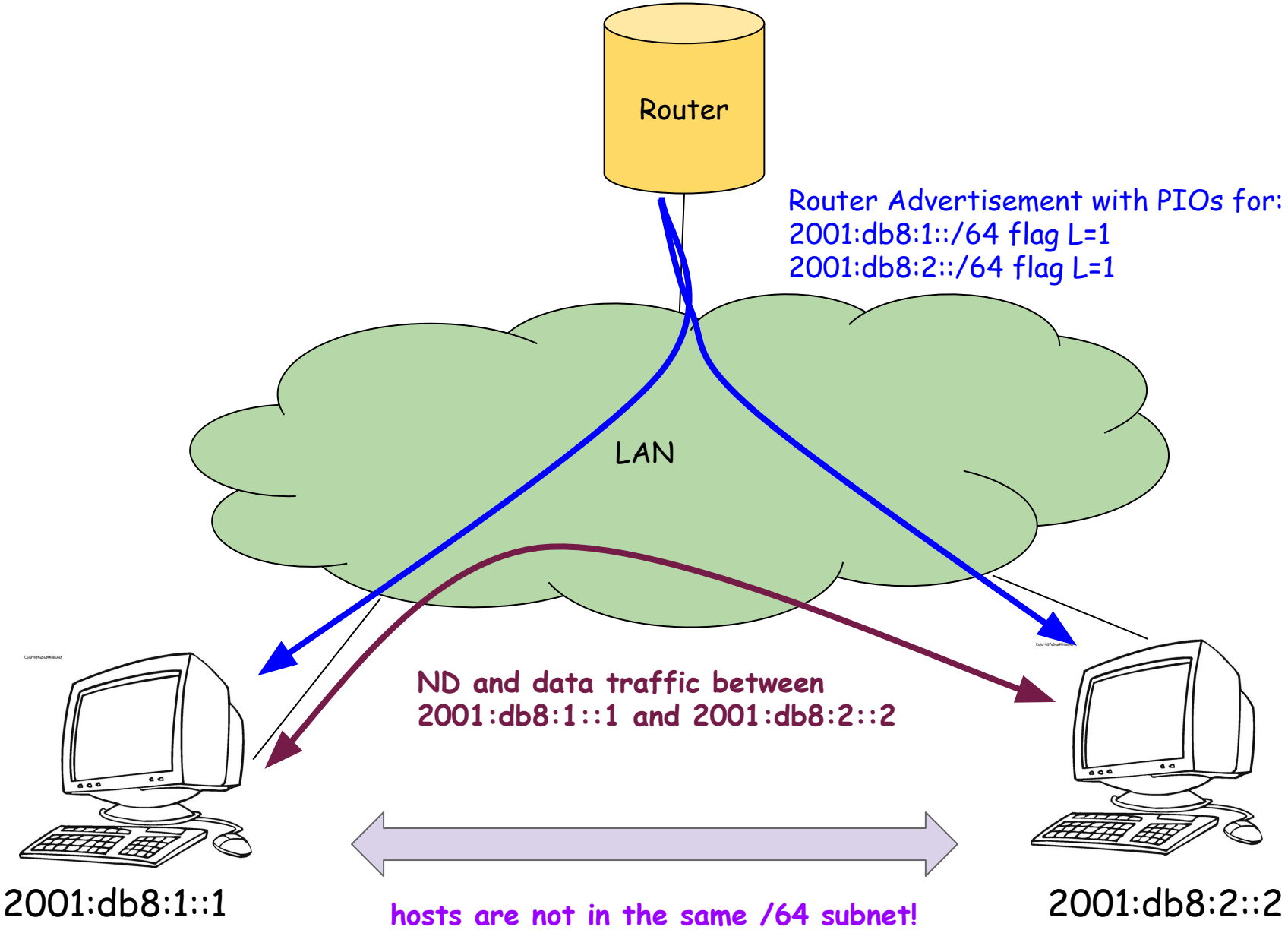
- PIO with L flag received
- ICMPv6 Redirect received from a router
- a Neighbor Advertisement message is received **for** the (target) address
- any Neighbor Discovery message is received **from** the address
- it is link-local fe80::/10

Everything else is OFF-LINK

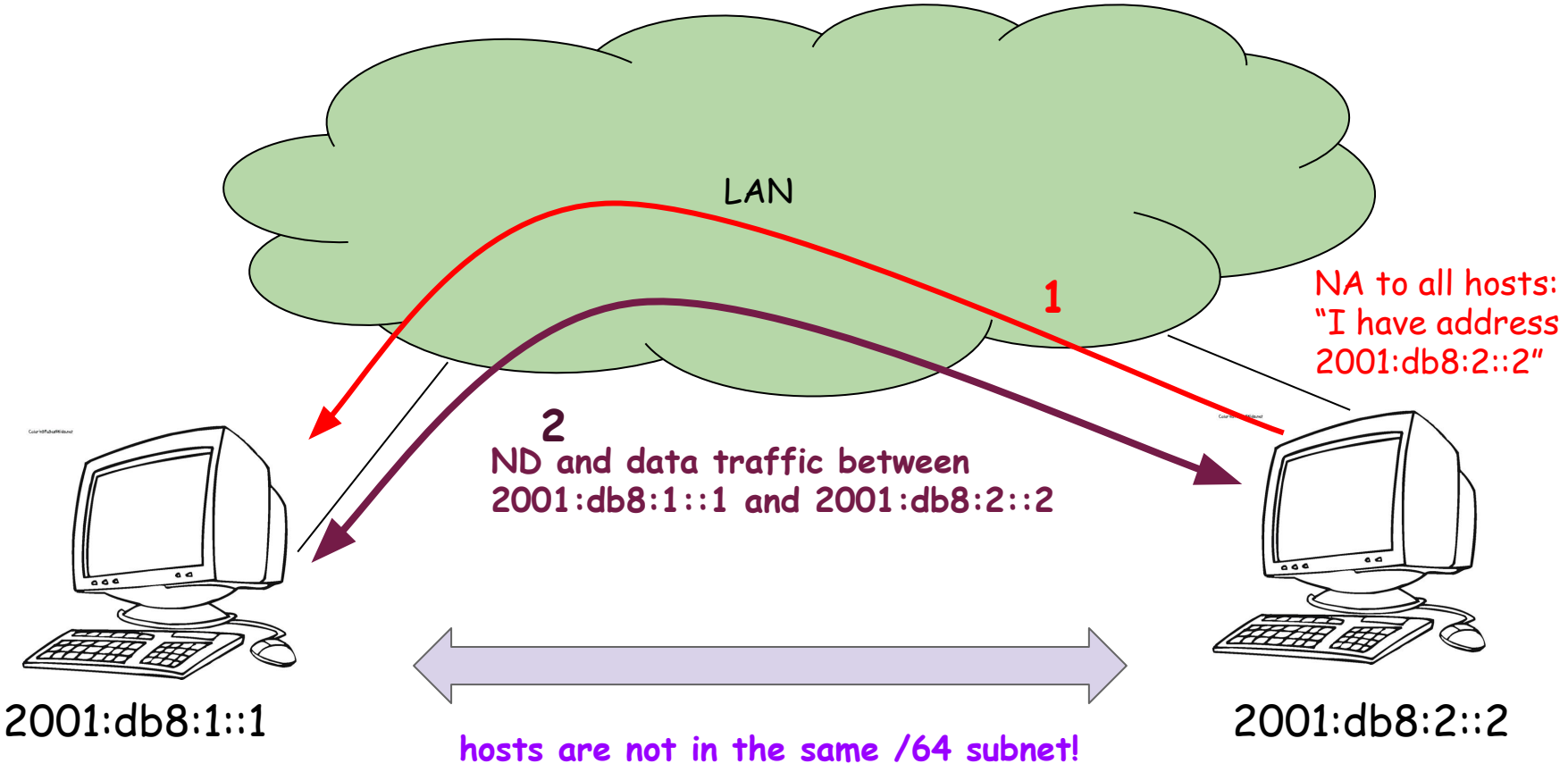
On-link Prefixes: Examples



On-link Prefix: Examples



On-link Prefix: Examples



Other RA Options

- Route Information Option
  - more specific routes
- MTU
- DNS
  - Recursive DNS Server
  - Search List

Questions?