IPv6 address architecture on point-to-point link

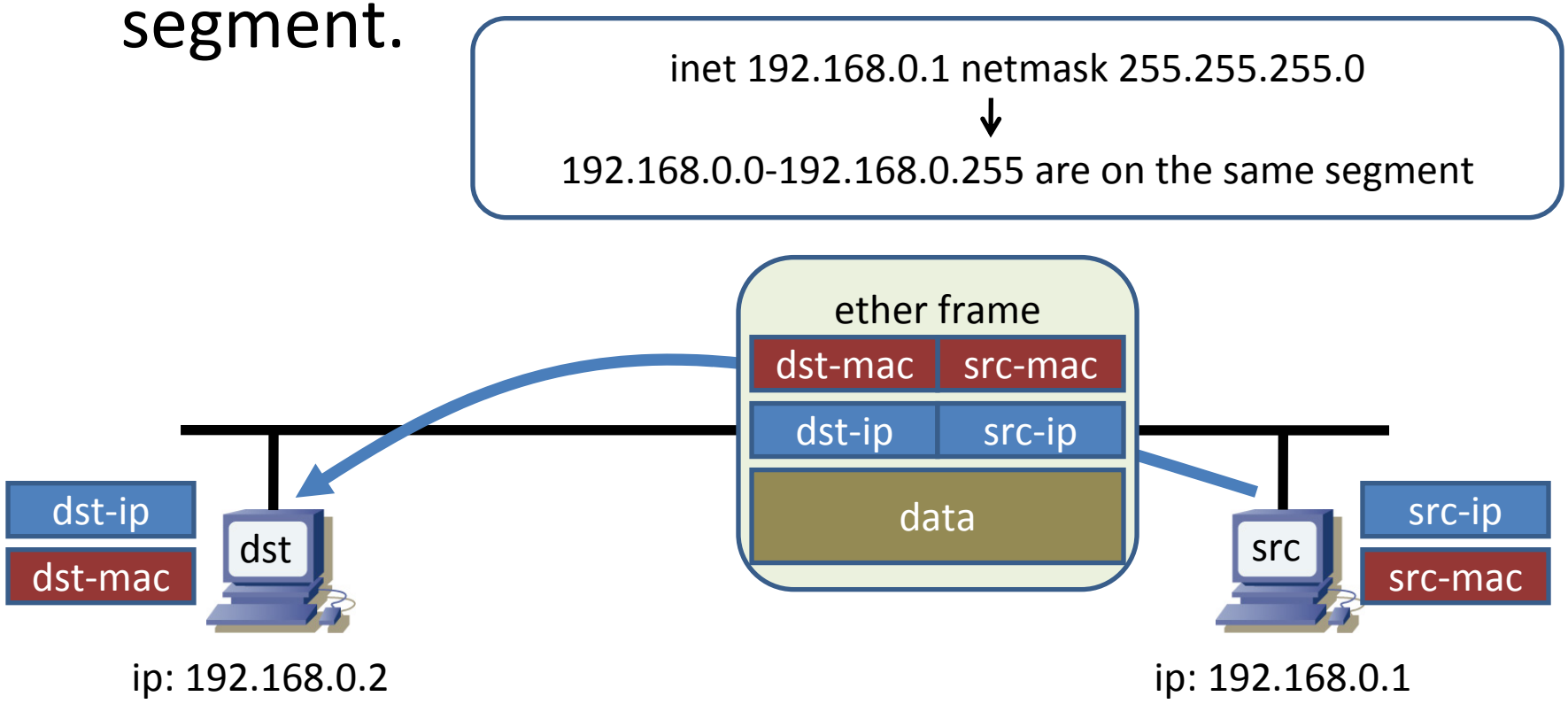
Matsuzaki ‘maz’ Yoshinobu
<maz@iij.ad.jp>

IPv4 and IPv6

- similar in routing, forwarding and so on
- The basis are almost same, but…
  - IP Header is a little bit different
  - more bits in the address field
    - IPv4 32bit -> IPv6 128bit

IPv4 packet forwarding

- Packets are delivered directly on the same segment.



arp (Address Resolution Protocol)

- MAC address is needed on Ethernet.
  - We know destination IPv4 address, then we need to know the layer 2(MAC) address from the IPv4 address.
- ARP
  - RFC826

```
arp who-has 192.168.0.2 tell 192.168.0.1
 0x0000:  ffff ffff ffff 0019 bb27 37e0 0806 0001
 0x0010:  0800 0604 0001 0019 bb27 37e0 c0a8 0001
 0x0020:  0000 0000 0000 c0a8 0002
arp reply 192.168.0.2 is-at 00:16:17:61:64:86
 0x0000:  0019 bb27 37e0 0016 1761 6486 0806 0001
 0x0010:  0800 0604 0002 0016 1761 6486 c0a8 0002
 0x0020:  0019 bb27 37e0 c0a8 0001 0000 0000 0000
 0x0030:  0000 0000 0000 0000 0000 0000
```

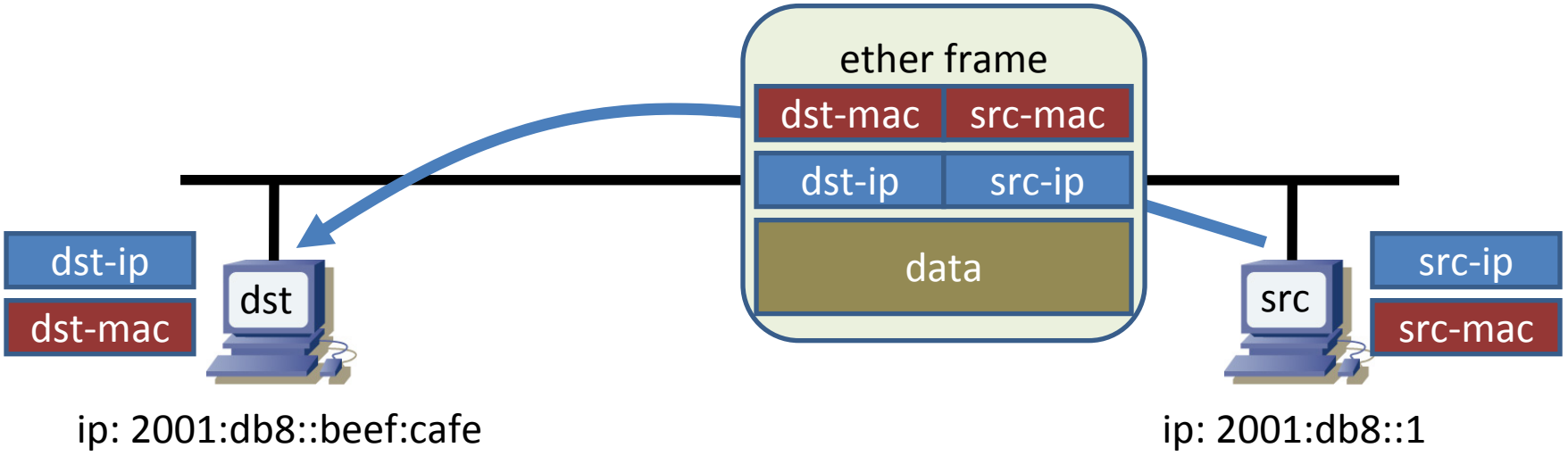
IPv6 packet forwarding

- Packets are delivered directly on the same segment.

inet6 2001:db8::1 prefixlen 64

↓

2001:db8:: - 2001:db8::ffff:ffff:ffff:ffff are on the same segment



ndp (Neighbor Discovery Protocol)

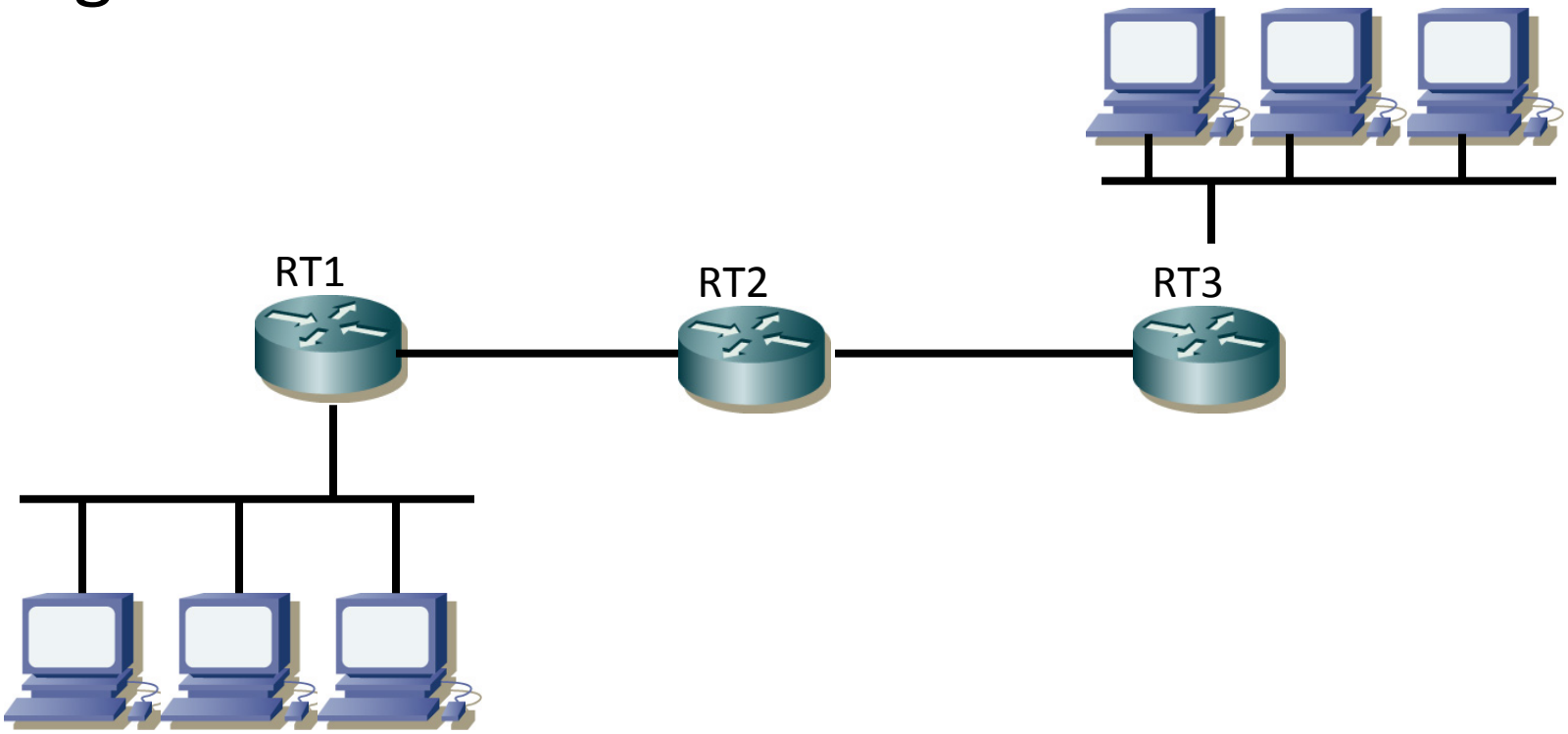
- MAC address is needed on Ethernet.
  - We know destination IPv6 address, then we need to know the layer 2(MAC) address from the IPv6 address.
- ndp
  - RFC4861
    - it uses ICMPv6 to solicit MAC address of destination host.

ndp - solicitation and advertisement

```
IP6 2001:db8::1 > ff02::1:ffef:cafe
 ICMP6, neighbor solicitation, who has 2001:db8::beef:cafe
 source link-address option: 00:19:bb:27:37:e0
        0x0000:  3333 ffef cafe 0019 bb27 37e0 86dd 6000
        0x0010:  0000 0020 3aff 2001 0db8 0000 0000 0000
        0x0020:  0000 0000 0001 ff02 0000 0000 0000 0000
        0x0030:  0001 ffef cafe 8700 9a90 0000 0000 2001
        0x0040:  0db8 0000 0000 0000 0000 beef cafe 0101
        0x0050:  0019 bb27 37e0
IP6 2001:db8::beef:cafe > 2001:db8::1
 ICMP6, neighbor advertisement, tgt is 2001:db8::beef:cafe
 destination link-address option: 00:16:17:61:64:86
        0x0000:  0019 bb27 37e0 0016 1761 6486 86dd 6000
        0x0010:  0000 0020 3aff 2001 0db8 0000 0000 0000
        0x0020:  0000 beef cafe 2001 0db8 0000 0000 0000
        0x0030:  0000 0000 0001 8800 c1fd 6000 0000 2001
        0x0040:  0db8 0000 0000 0000 0000 beef cafe 0201
        0x0050:  0016 1761 6486
```

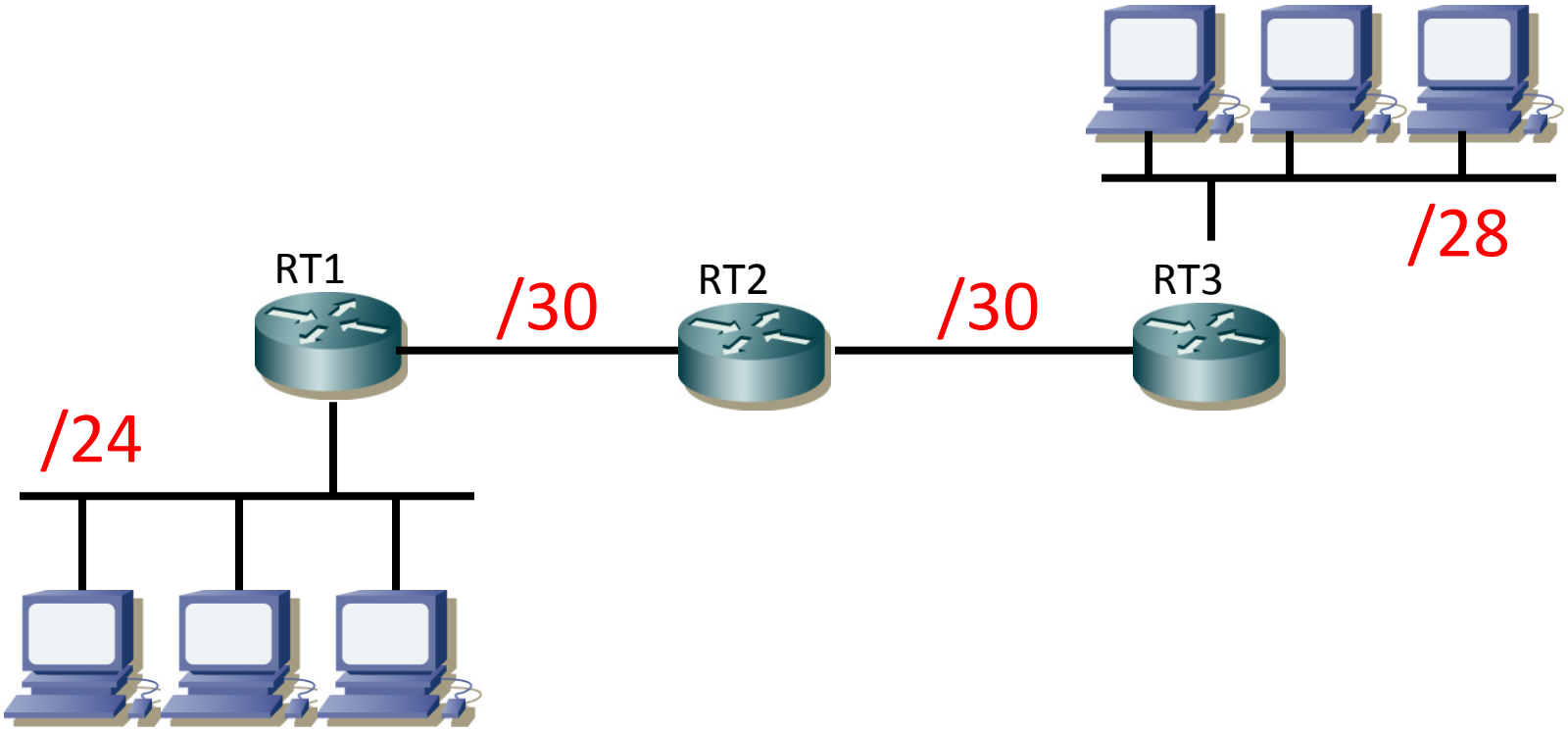
network architecture

- inter-router connections, router-hosts segments and so on.



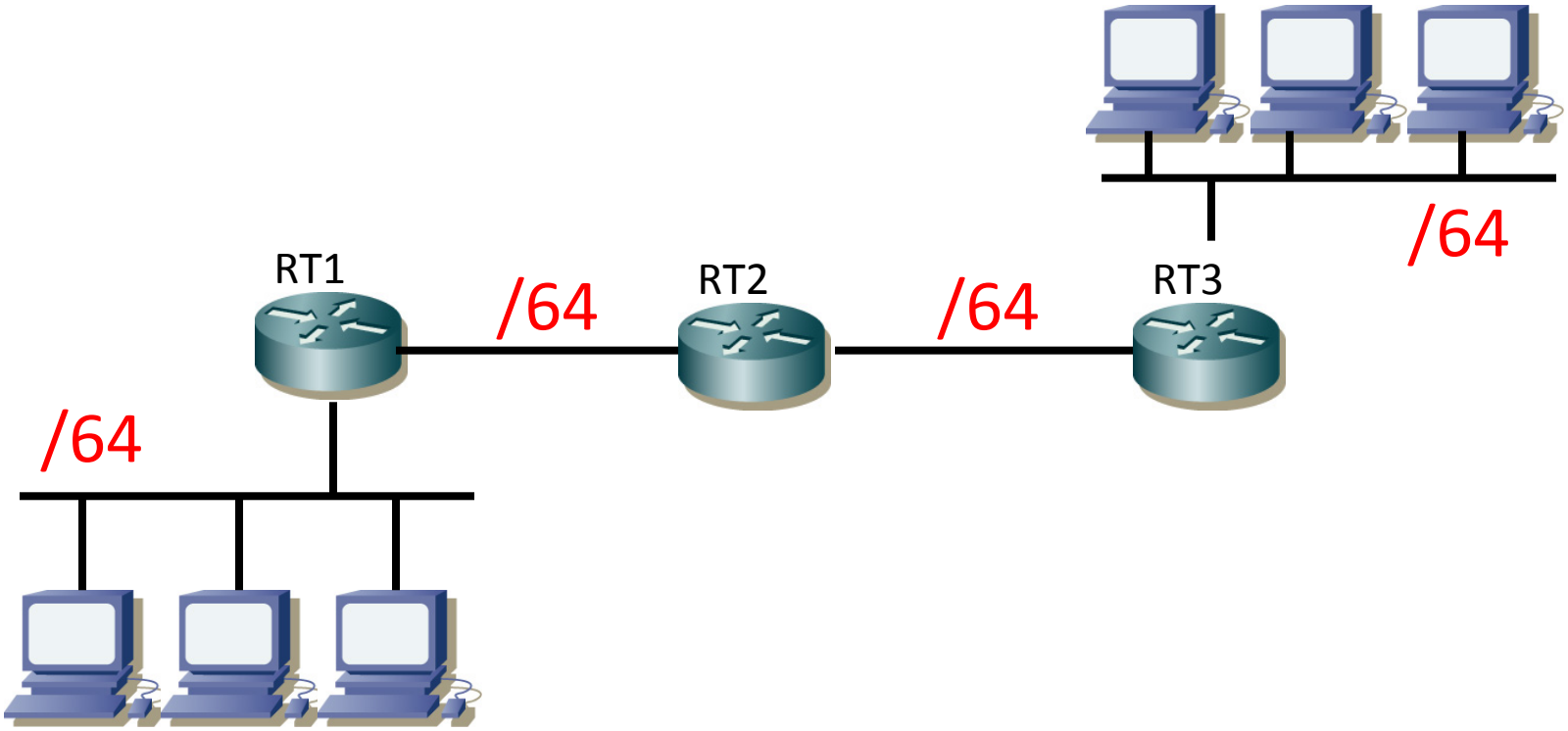
IPv4 address architecture

- We assign optimal net blocks based on needs.



IPv6 address architecture

- /64 is used everywhere.



point-to-point link

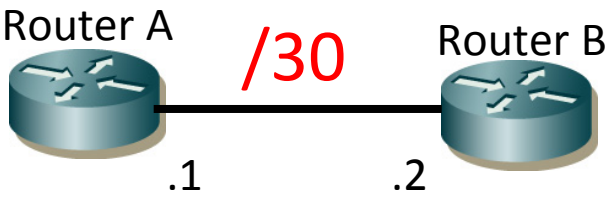
- It mainly used on an inter-router link.
  - POS link, Serial link and the like
    - useful for troubleshoot
  - Actually a tunnel link is also point-to-point link.
- Routers just throw packets to an opposite router via the link.
  - A layer2 address resolution like arp is not required.

point-to-point link and addressing

- We configured an address of the opposite router on each link in ancient days.
  - remote-address or dest_address
  - There are a few routers that still support remote-address.
- Now we use /30s or /64s for a point-to-point link as if there is a segment on the link.
  - We don’t care if the link is Ethernet or POS.

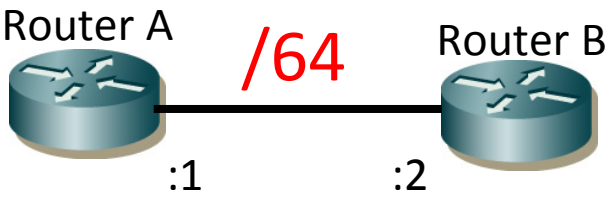
/30 for inter-router link

- 10.0.0.0/30
  - 10.0.0.0 <- network address
  - 10.0.0.1 <- Router A
  - 10.0.0.2 <- Router B
  - 10.0.0.3 <- broadcast address



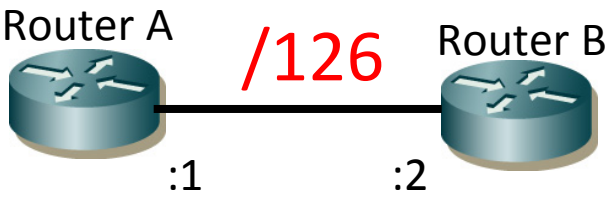
/64 for inter-router link

- 2001:db8::/64
  - 2001:db8::0 <- Subnet Router-anycast address
  - 2001:db8::1 <- Router A
  - 2001:db8::2 <- Router B
  - 2001:db8::3-2001:db8::ffff:ffff:ffff:ffff <- unused



even if you assign /126 for the link

- 2001:db8::/64
  - 2001:db8::0 <- Subnet Router-anycast address
  - 2001:db8::1 <- Router A
  - 2001:db8::2 <- Router B
  - 2001:db8::3 <- unused

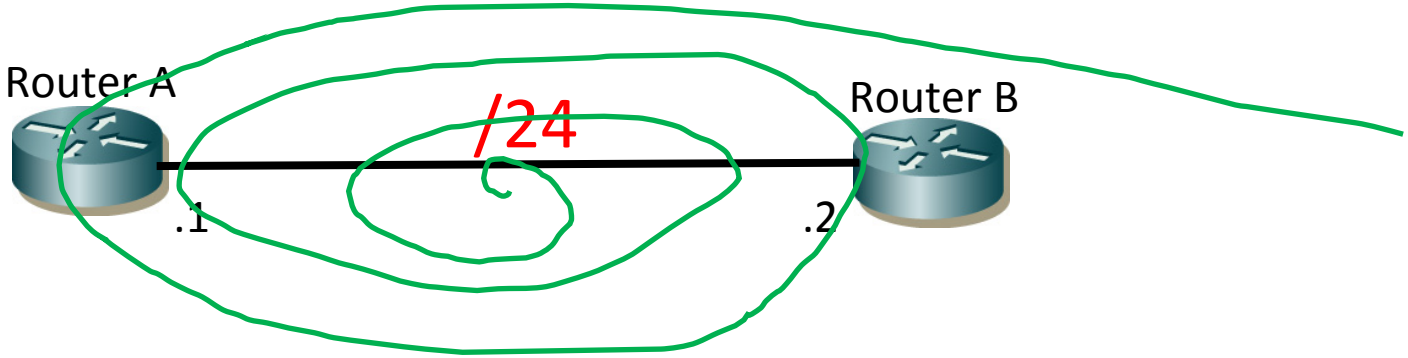


directed unused-address packets

- If a packet is coming to the unused-address on the segment….
- Ethernet
  - Layer2 address resolution is performed by arp or ndp. And icmp host-unreach message is replied if no one use the address.
- point-to-point link
  - The packet will be thrown to the link….

point-to-point link and unused-address

- One assigns 10.0.0.0/24 for the link.
  - 10.0.0.1 for Router A and 10.0.0.2 for Router B
- In this case, what is happened if the destination of a packet is 10.0.0.13?



ping pong

- A packet destined for unused-address on the point-to-point link could be loop on the link.
  - In IPv4 case, there is no vacant address, because people use /30s or /31s for these links.
  - In IPv6 case , there are vacant addresses on the link.
- Of course this issue was noticed early, and has been discussed.
  - Actually there is a solution in a RFC.

RFC4443 – ICMPv6

3. ICMPv6 Error Messages

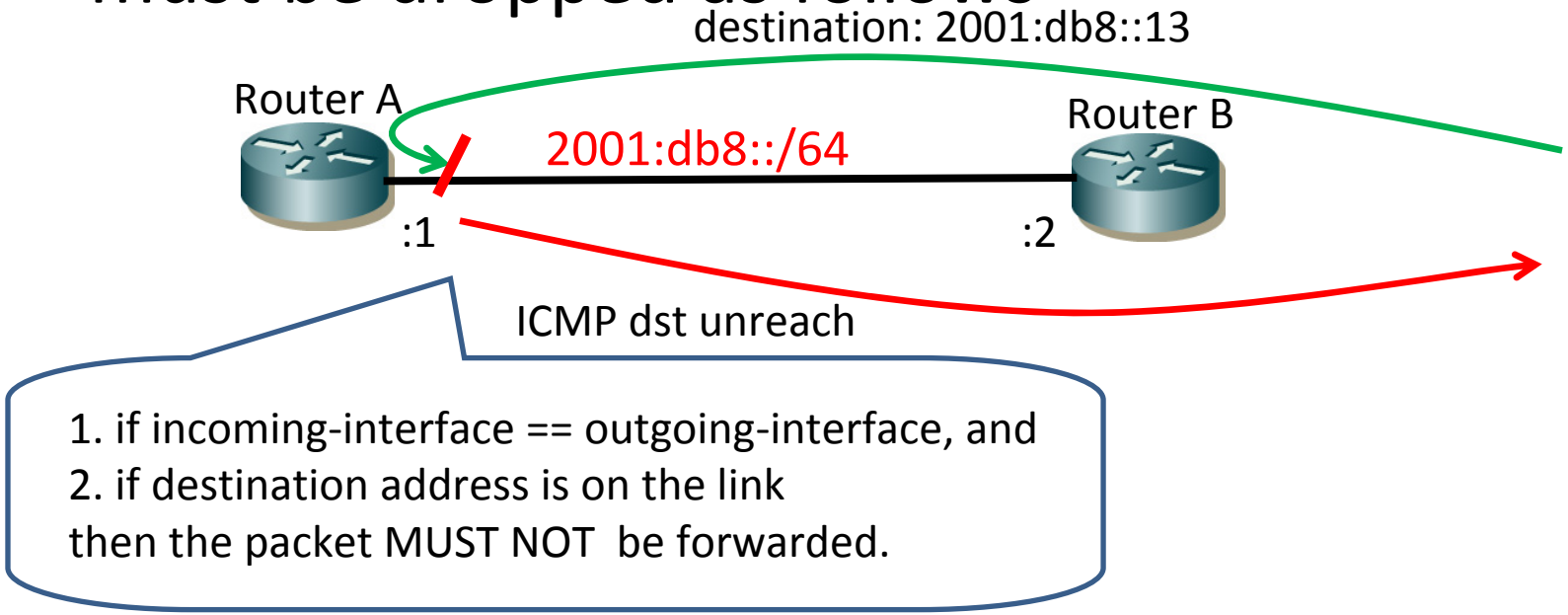
3.1. Destination Unreachable Message

<snip>

One specific case in which a Destination Unreachable message is sent with a code 3 is in response to a packet received by a router from a point-to-point link, destined to an address within a subnet assigned to that same link (other than one of the receiving router's own addresses). In such a case, the packet MUST NOT be forwarded back onto the arrival link.

RFC says

- When you use 2001:db8::/64 for a point-to-point link, a packet destined for 2001:db8::13 must be dropped as follows



awareness is important

- Specs are okay, but implementations are….
  - common matter
- Sometimes ‘special cases’ might be forgotten, so we have to check these before use it.
  - We might forget,
  - And also vendors might forget.

other ideas for this issue

- Operators need workarounds.
  - insurance saves us ☺
  - In case of emergency, we really need workarounds.

1. link-local addressing
2. messy packet-filter
3. /127 for the link

1. link-local for inter-router link

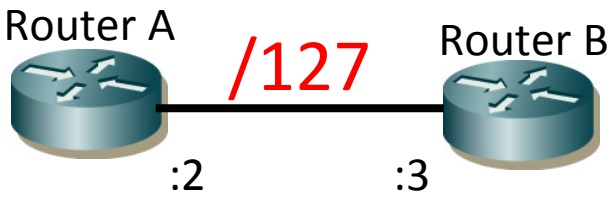
- In IPv6 case, a router does not require global address for a inter-router link.
  - Only loopback interface needs global address
  - Neighboring routers can exchange routing information by link-local address
- Issues of ‘no global address on interfaces’
  - ping from remote site
  - eBGP session (nexthop rewrite is needed)
  - traceroute become useless a bit

2. messy packet-filter

- What we call ‘infrastructure ACL’
  - allocate address block for its infrastructure first, then put a filter that denies packets from outside to these addresses.
  - IPv6 has more bits, so we can allocate infrastructure address easily
- Issues
  - how to maintain these filters
  - inter-AS connections

3. /127 for the link

- 2001:db8::2/127
  - 2001:db8::2 <- Router A
  - 2001:db8::3 <- Router B



- No vacant address on the link ☺
- This was discussed before

/127 and RFC

- ‘use of /127 considered harmful’ [RFC3627]
  - The first address is reserved for subnet router anycast address
    - It might cause problems, though almost of all routers do not support the subnet router anycast address at this moment.
    - We can write a RFC that do not use the subnet router anycast address with /127 addressing, but this brings another ‘special case’
  - Issue of ‘longer prefix other than /64’

Summary

- ‘Specs’ is not ‘Implementation’
  - Check before use
    - please let vendors know when you find a problem
  - caution needed for ‘special cases’
    - ping-pong on point-to-point links
    - tunnels are also ‘point-to-point’
- Even a small difference between IPv4 and IPv6 might cause some issues for your operation.
  - prejudice or bias

END