Networker's handbook
part 2

Matěj Grégr
igregr@fit.vutbr.cz

Agenda

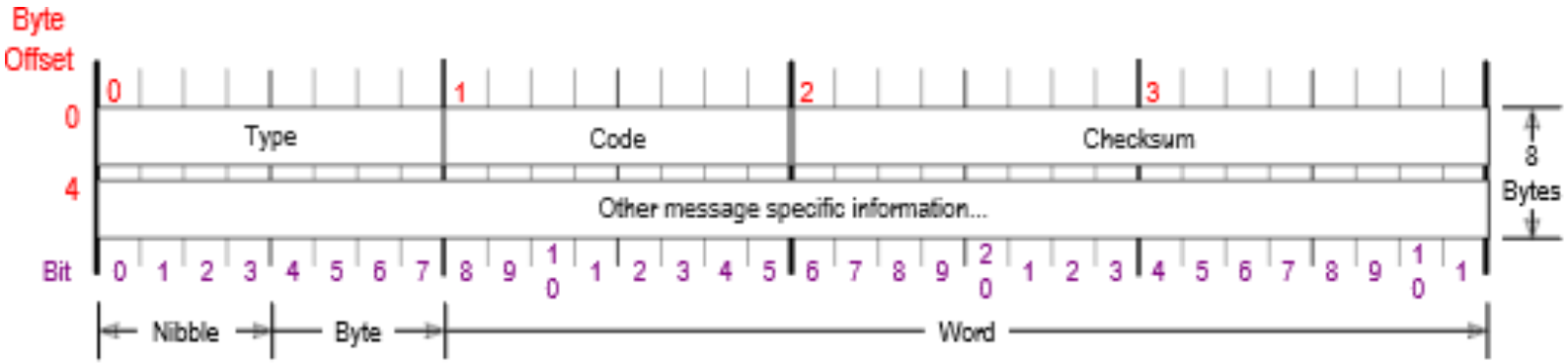
- ICMP, ICMPv6
- DHCP, SLAAC, DHCPv6
- NAT
- DNS
- Troubleshooting

ICMP, ICMPv6

ICMP

- RFC 792
- **Internet Control Message Protocol (ICMP)**
  - Service protocol for IPv4 functionality
  - Employed by hosts and routers to diagnose network and to announce errors during packet transfer
- Messages
  - Encapsulated into IPv4 packet, Protocol number 1
  - Two message kinds – error notifications and service queries
  - May contain the part of original packet that caused generation of this ICMP message
- Applications
  - Ping
  - Traceroute

ICMP header



ICMP Message Types

Type Code/Name

- 0 Echo Reply
- 3 Destination Unreachable
  - 0 Net Unreachable
  - 1 Host Unreachable
  - 2 Protocol Unreachable
  - 3 Port Unreachable
  - 4 Fragmentation Required, and DF set
  - 5 Source Route Failed
  - 6 Destination Network Unknown
  - 7 Destination Host Unknown
  - 8 Source Host Isolated
  - 9 Network Administratively Prohibited
  - 10 Host Administratively Prohibited
  - 11 Network Unreachable for TOS
- 3 Destination Unreachable (continued)
  - 12 Host Unreachable for TOS
  - 13 Communication Administratively Prohibited
- 4 Source Quench
- 5 Redirect
  - 0 Redirect Datagram for the Network
  - 1 Redirect Datagram for the Host
  - 2 Redirect Datagram for the TOS & Network
  - 3 Redirect Datagram for the TOS & Host
- 8 Echo
- 9 Router Advertisement
- 10 Router Selection
- 11 Time Exceeded
  - 0 TTL Exceeded
  - 1 Fragment Reassembly Time Exceeded
- 12 Parameter Problem
  - 0 Pointer Problem
  - 1 Missing Required Operand
  - 2 Bad Length
- 13 Timestamp
- 14 Timestamp Reply
- 15 Information Request
- 16 Information Reply
- 17 Address Mask Request
- 18 Address Mask Reply
- 30 Traceroute

Checksum

Checksum of ICMP header.

RFC 792

Please refer to RFC 792 for the Internet Control Message protocol (ICMP) specification.

ICMPv6

- *ICMPv6 is more important than ICMP for IPv4!!!*
- Supports all messages as ICMPv4
  - Destination Unreachable, Packet Too Big, Time Exceeded, Parameter Problem, Echo/Echo Reply, Redirect
  - Add new messages
    - Router Solicitation, Advertisement (plug-an-play configuration)
    - Neighbor Solicitation, Advertisement (ARP replacement)
- Added features are used for:
  - Finding routers
  - IPv6 plug-an-play configuration
  - IPv6 address to MAC address translation
  - Duplicate IPv6 address detection

ICMPv6 header

ICMPv6 packet

Bit offset	0–7	8–15	16–31
0	Type	Code	Checksum
32	Message body		

- Checksum is computed from IPv6 pseudo-header

ICMPv6 pseudo-header

Bit offset	0 - 7	8–15	16–23	24–31
0	Source address			
32				
64				
96				
128	Destination address			
160				
192				
224				
256	ICMPv6 length			
288	Zeros			Next header

DHCP

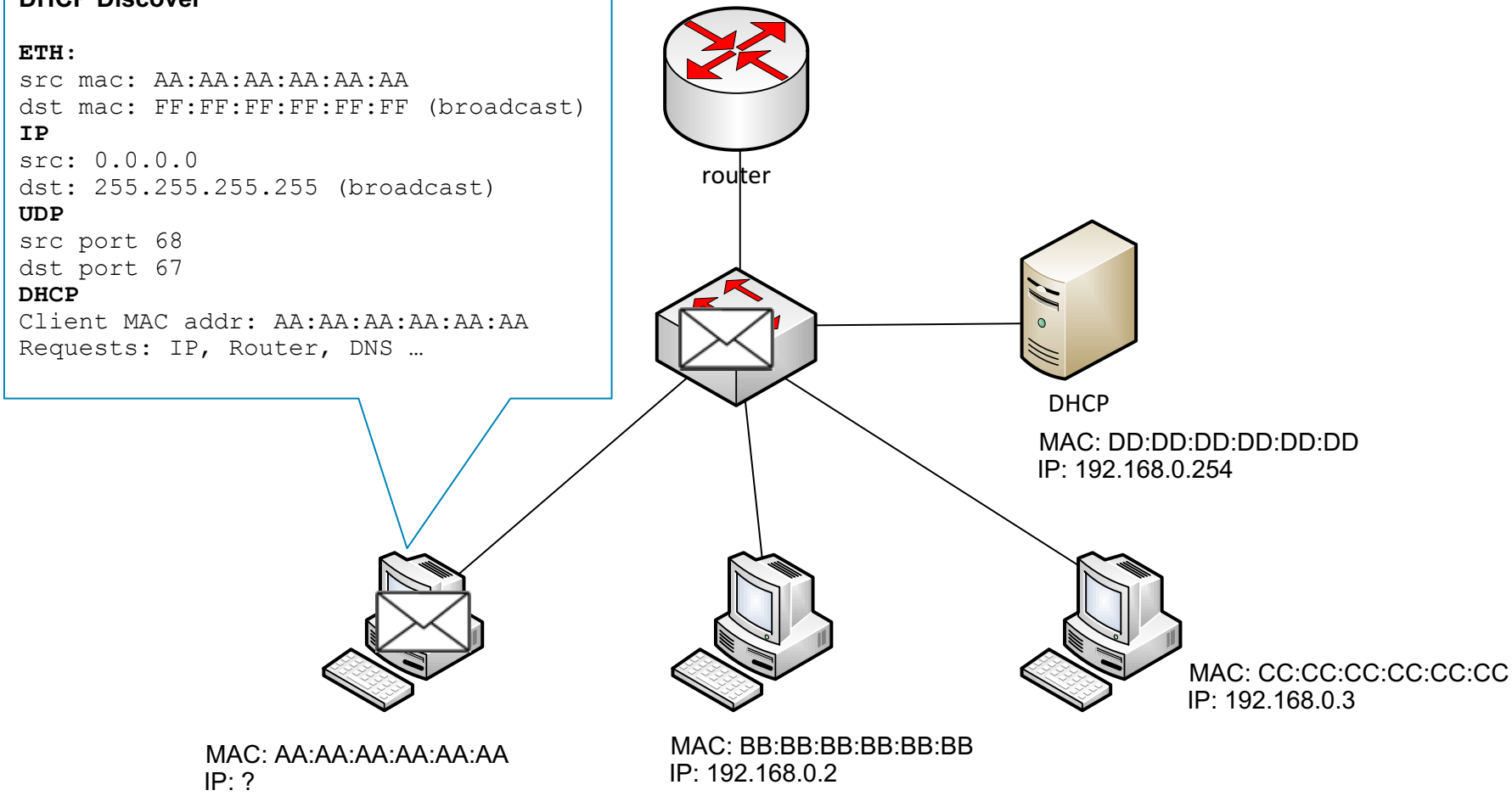
Dynamic Host Configuration Protocol

- Client, Server
- Assigns IP address and other configuration parameters
- UDP ports 67 (client listens) and 68 (server listens)
- Parameters are valid for a given period of time (lease time) then they need to be renegotiated
- **Messages**
  - Most used messages: Discover, Offer Request, Ack
  - Other messages: Decline, Negative Acknowledgment,Release, Inform

DHCP – example

DHCP Discover

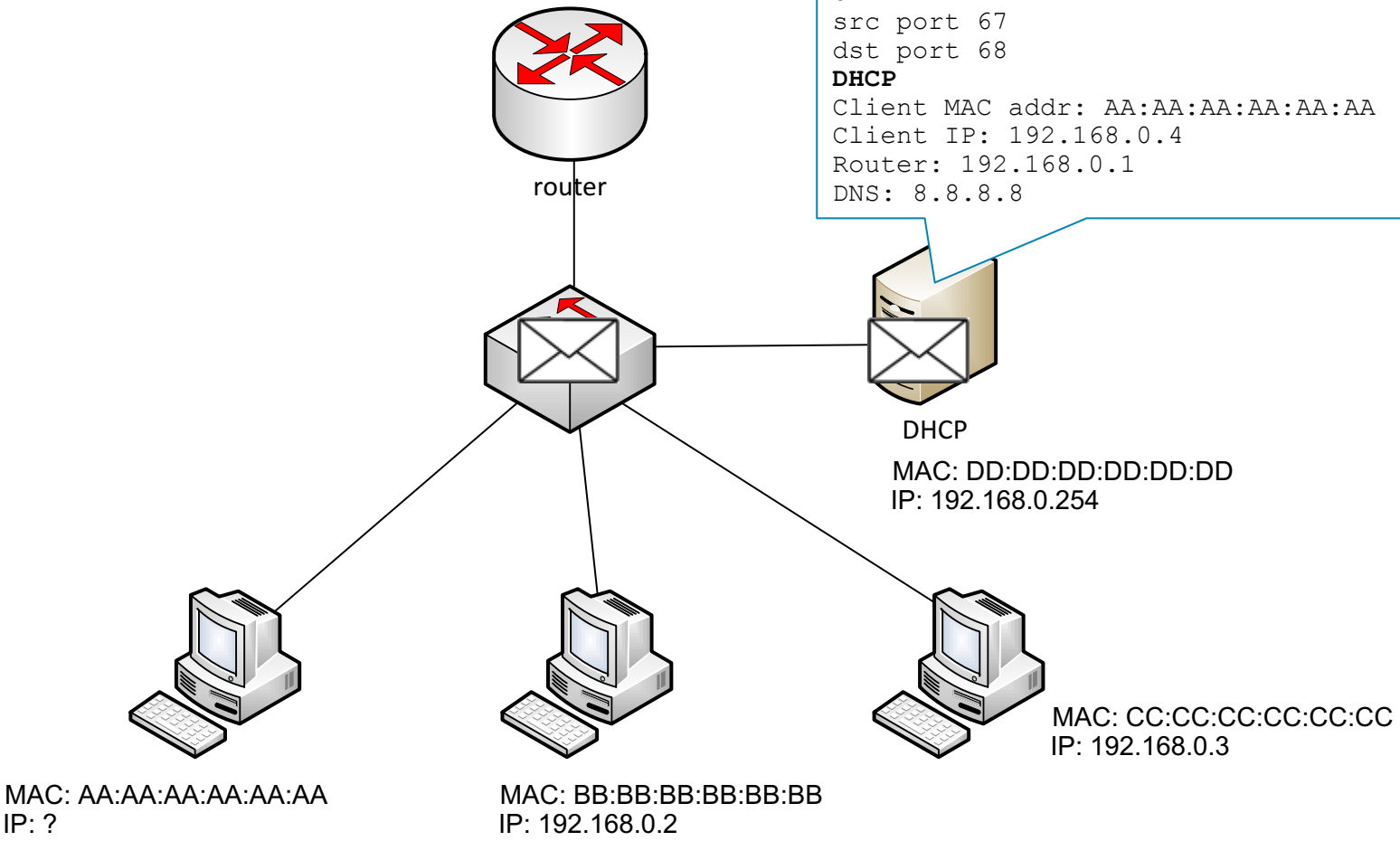
ETH:
src mac: AA:AA:AA:AA:AA:AA
dst mac: FF:FF:FF:FF:FF:FF (broadcast)
IP
src: 0.0.0.0
dst: 255.255.255.255 (broadcast)
UDP
src port 68
dst port 67
DHCP
Client MAC addr: AA:AA:AA:AA:AA:AA
Requests: IP, Router, DNS …



DHCP – example

DHCP Offer

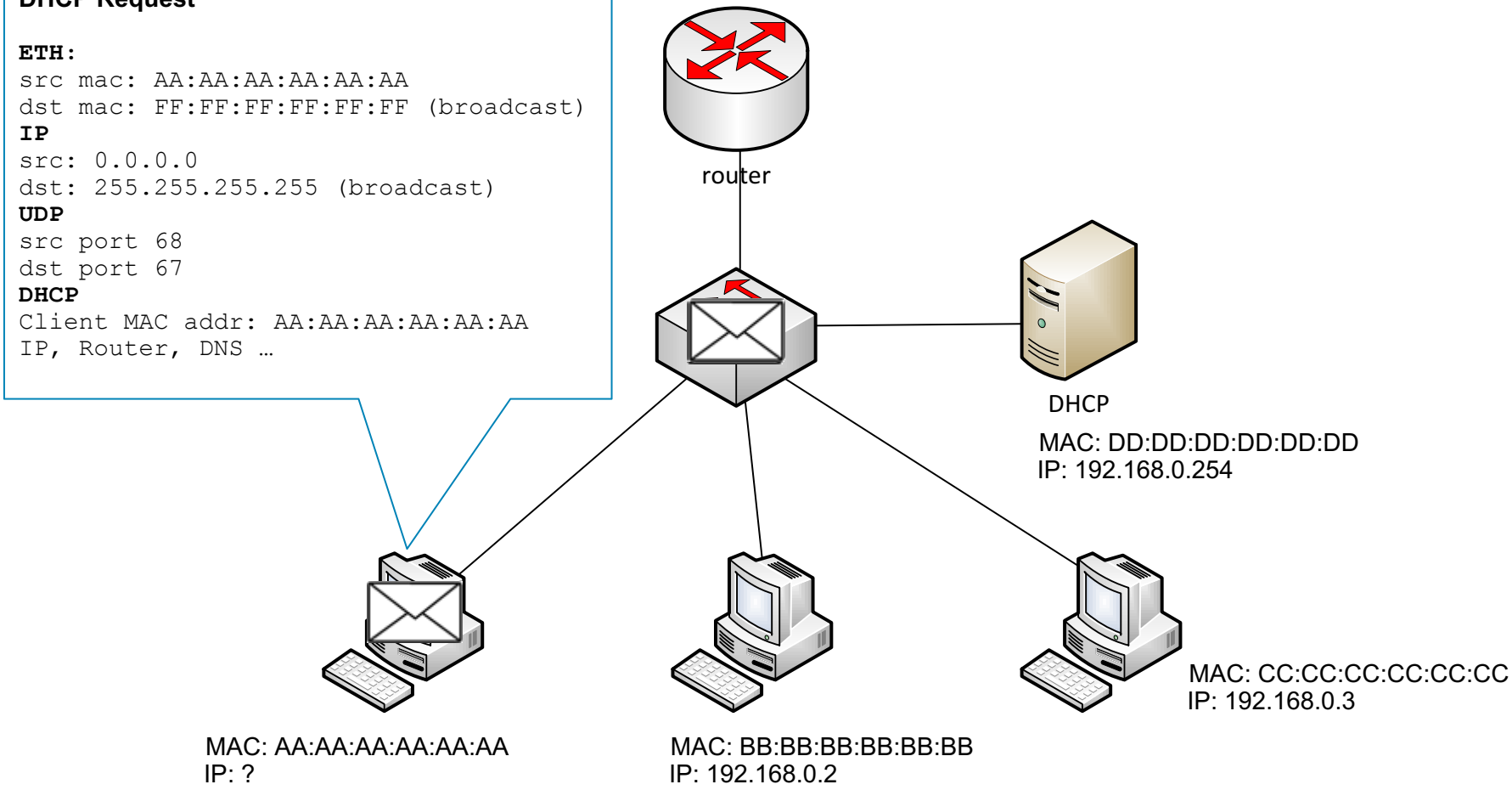
ETH:
src mac: DD:DD:DD:DD:DD:DD
dst mac: AA:AA:AA:AA:AA:AA
IP
src: 192.168.0.254
dst: 192.168.0.4
UDP
src port 67
dst port 68
DHCP
Client MAC addr: AA:AA:AA:AA:AA:AA
Client IP: 192.168.0.4
Router: 192.168.0.1
DNS: 8.8.8.8



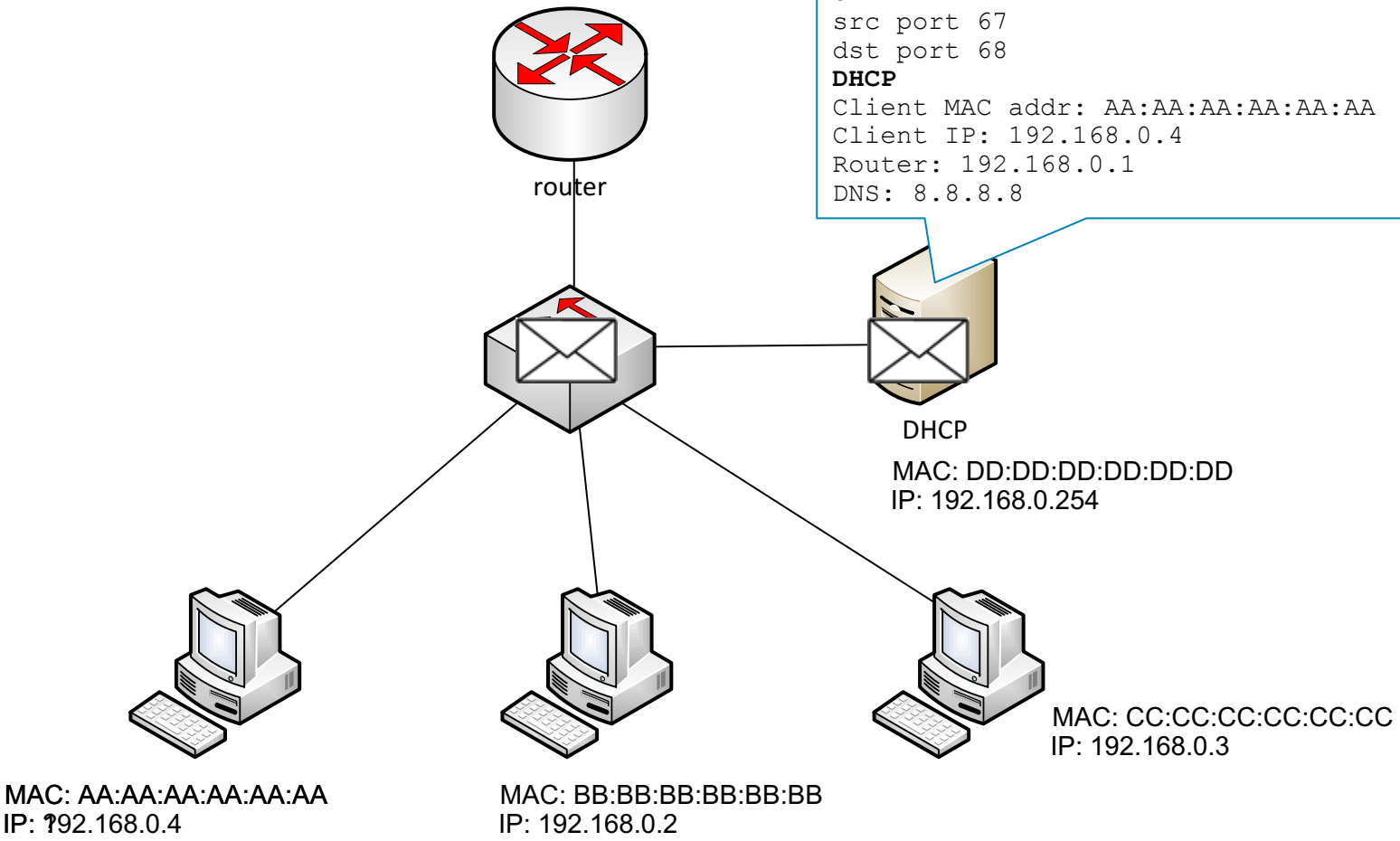
DHCP – example

DHCP Request

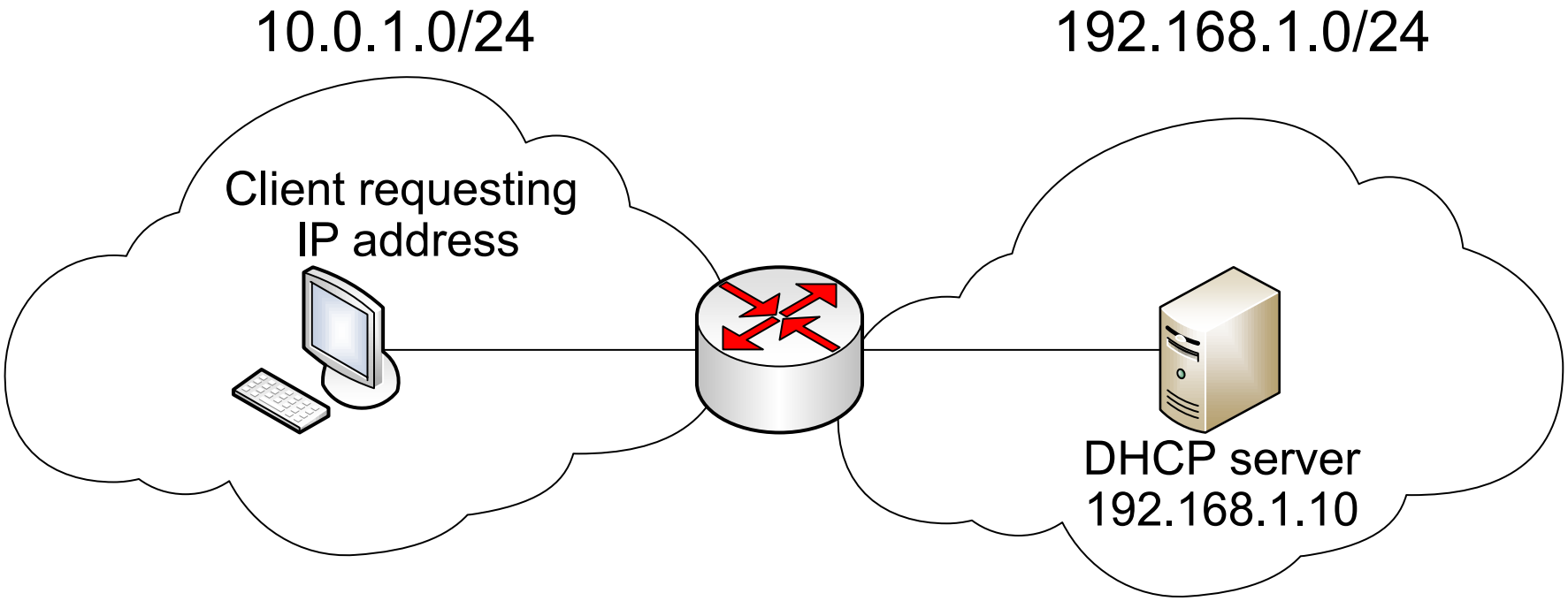
ETH:
src mac: AA:AA:AA:AA:AA:AA
dst mac: FF:FF:FF:FF:FF:FF (broadcast)
IP
src: 0.0.0.0
dst: 255.255.255.255 (broadcast)
UDP
src port 68
dst port 67
DHCP
Client MAC addr: AA:AA:AA:AA:AA:AA
IP, Router, DNS …



DHCP – example



DHCP in different subnet?



SLAAC

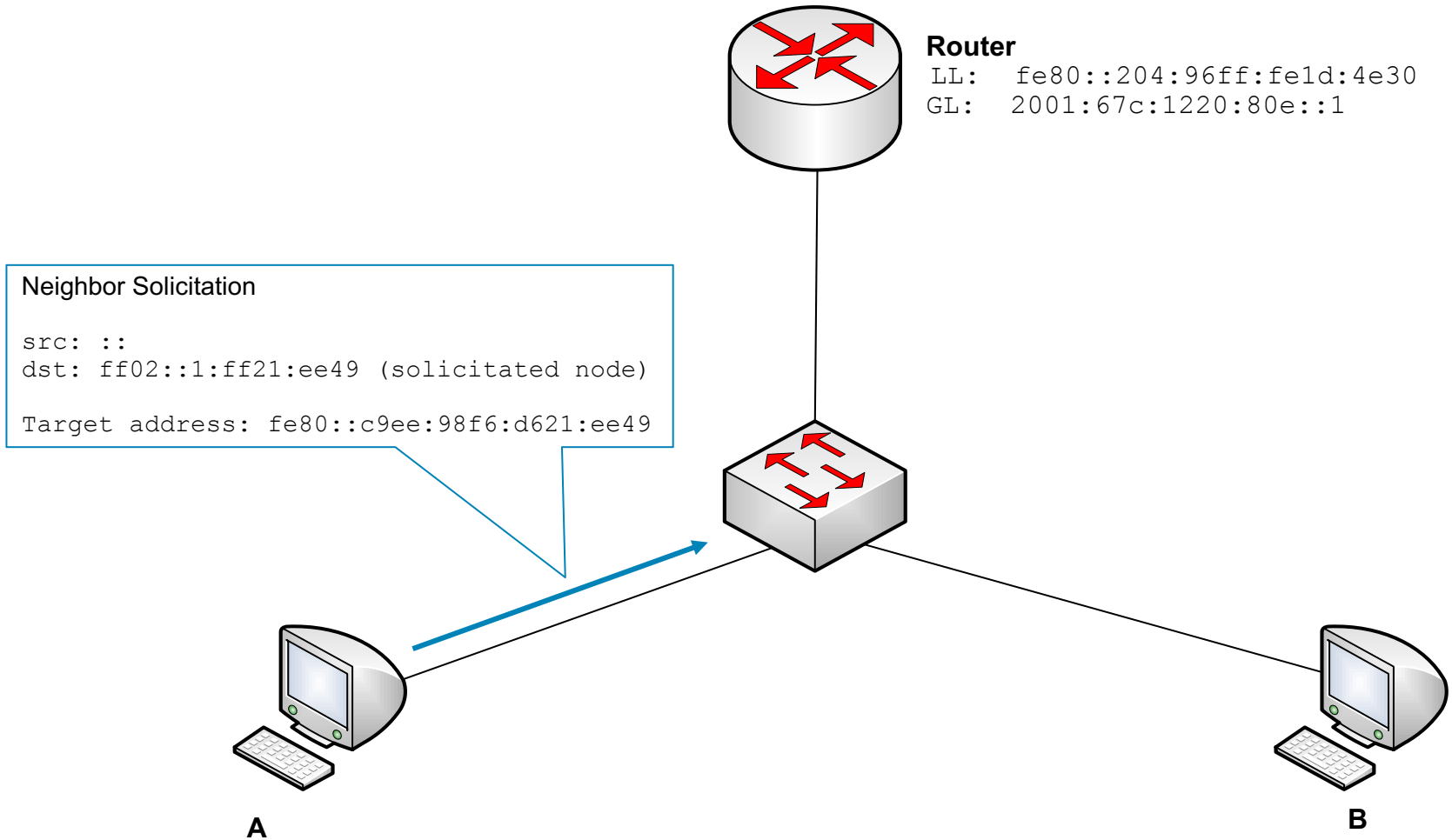
SLAAC

- Stateless address autoconfiguration in IPv6
  - Plug&play configuration for IPv6 nodes
- A router sends network information to all the nodes on the local link (**Router Advertisement messages, RA**)
- A host can autoconfigure itself by appending its IPv6 interface identifier (64-bit format) to the local link prefix (64 bits)
  - 64 bits? → RFC 7421
  - How to choose IID? → RFC 7217, RFC 4941, RFC 4862, ...
  - DAD is necessary!

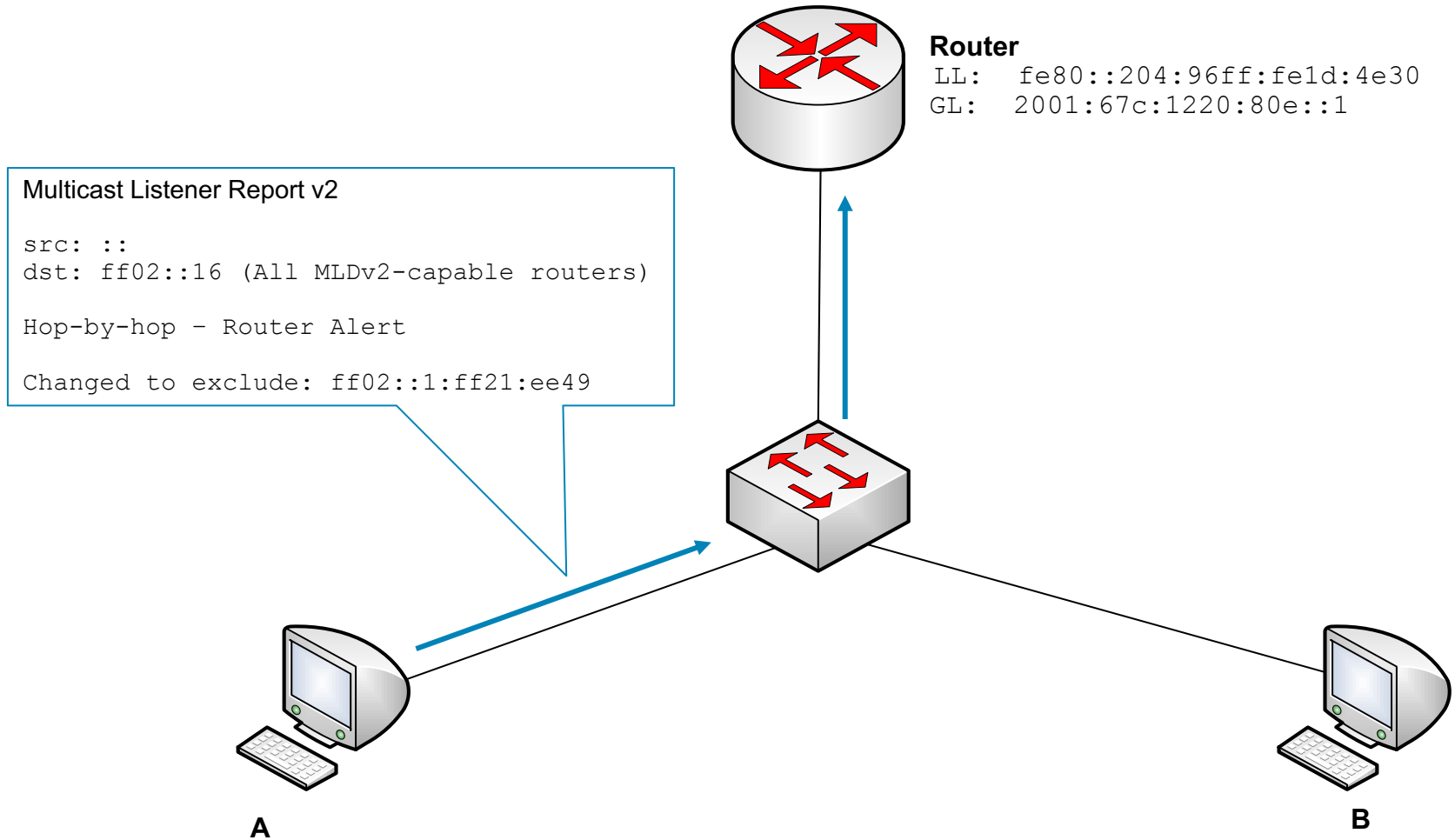
Router Advertisement message

```
▸ Ethernet II, Src: ExtremeN_1d:4e:30 (00:04:96:1d:4e:30), Dst: IPv6mcast_01 (33:33:00:00:00:01)
▸ Internet Protocol Version 6, Src: fe80::204:96ff:fe1d:4e30, Dst: ff02::1
▾ Internet Control Message Protocol v6
    Type: Router Advertisement (134)
    Code: 0
    Checksum: 0x99bf [correct]
    [Checksum Status: Good]
    Cur hop limit: 64
  ▾ Flags: 0x80
      1... .... = Managed address configuration: Set
      .0.. .... = Other configuration: Not set
      ..0. .... = Home Agent: Not set
      ...0 0... = Prf (Default Router Preference): Medium (0)
      .... .0.. = Proxy: Not set
      .... ..0. = Reserved: 0
    Router lifetime (s): 1800
    Reachable time (ms): 30000
    Retrans timer (ms): 1000
  ▾ ICMPv6 Option (MTU : 1500)
      Type: MTU (5)
      Length: 1 (8 bytes)
      Reserved
      MTU: 1500
  ▾ ICMPv6 Option (Source link-layer address : 00:04:96:1d:4e:30)
      Type: Source link-layer address (1)
      Length: 1 (8 bytes)
      Link-layer address: ExtremeN_1d:4e:30 (00:04:96:1d:4e:30)
  ▾ ICMPv6 Option (Prefix information : 2001:67c:1220:80c::/64)
      Type: Prefix information (3)
      Length: 4 (32 bytes)
      Prefix Length: 64
    ▾ Flag: 0xc0
        1... .... = On-link flag(L): Set
        .1.. .... = Autonomous address-configuration flag(A): Set
        ..0. .... = Router address flag(R): Not set
        ...0 0000 = Reserved: 0
      Valid Lifetime: 2592000
      Preferred Lifetime: 604800
      Reserved
      Prefix: 2001:67c:1220:80c::
```

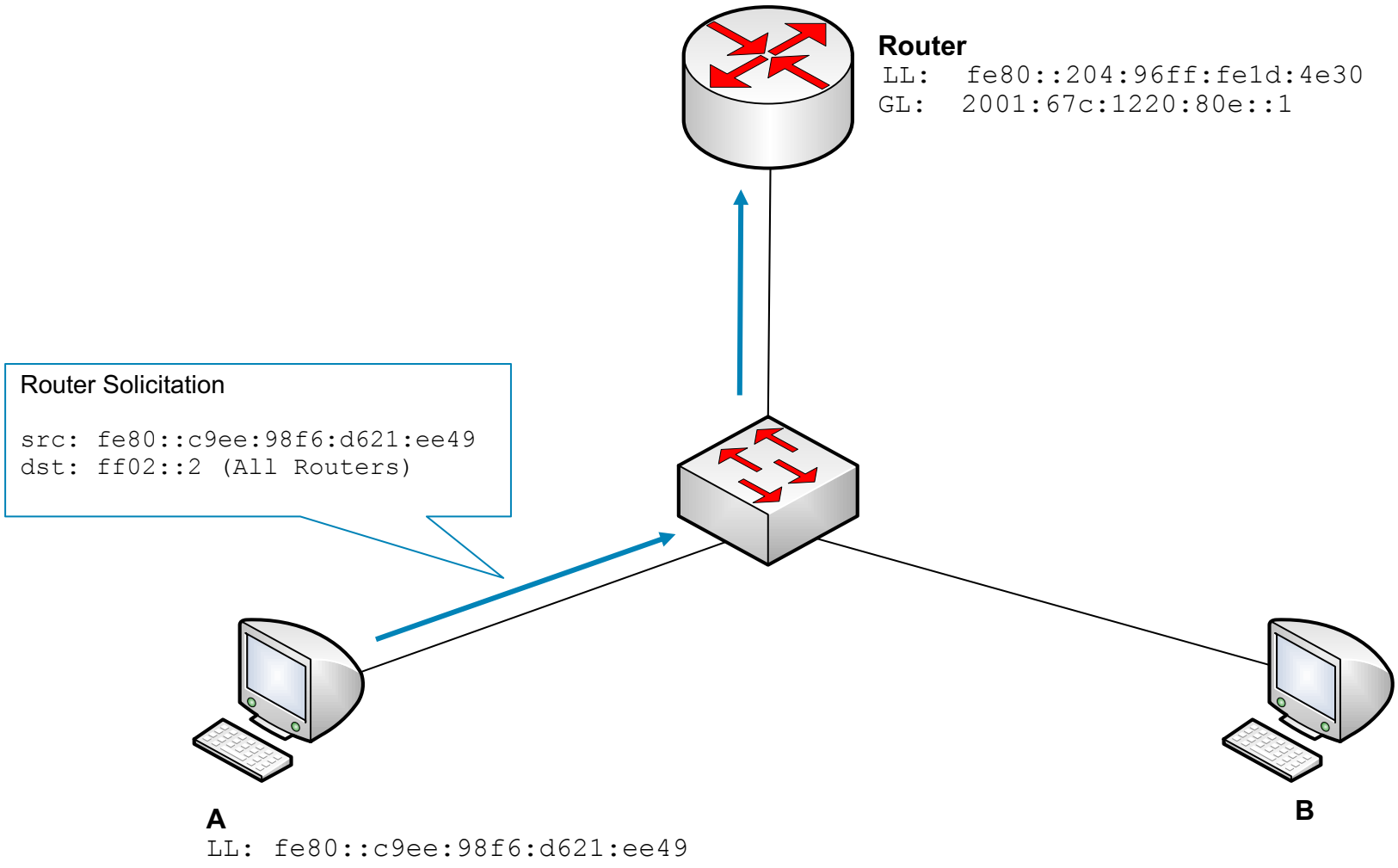
Link local address + DAD



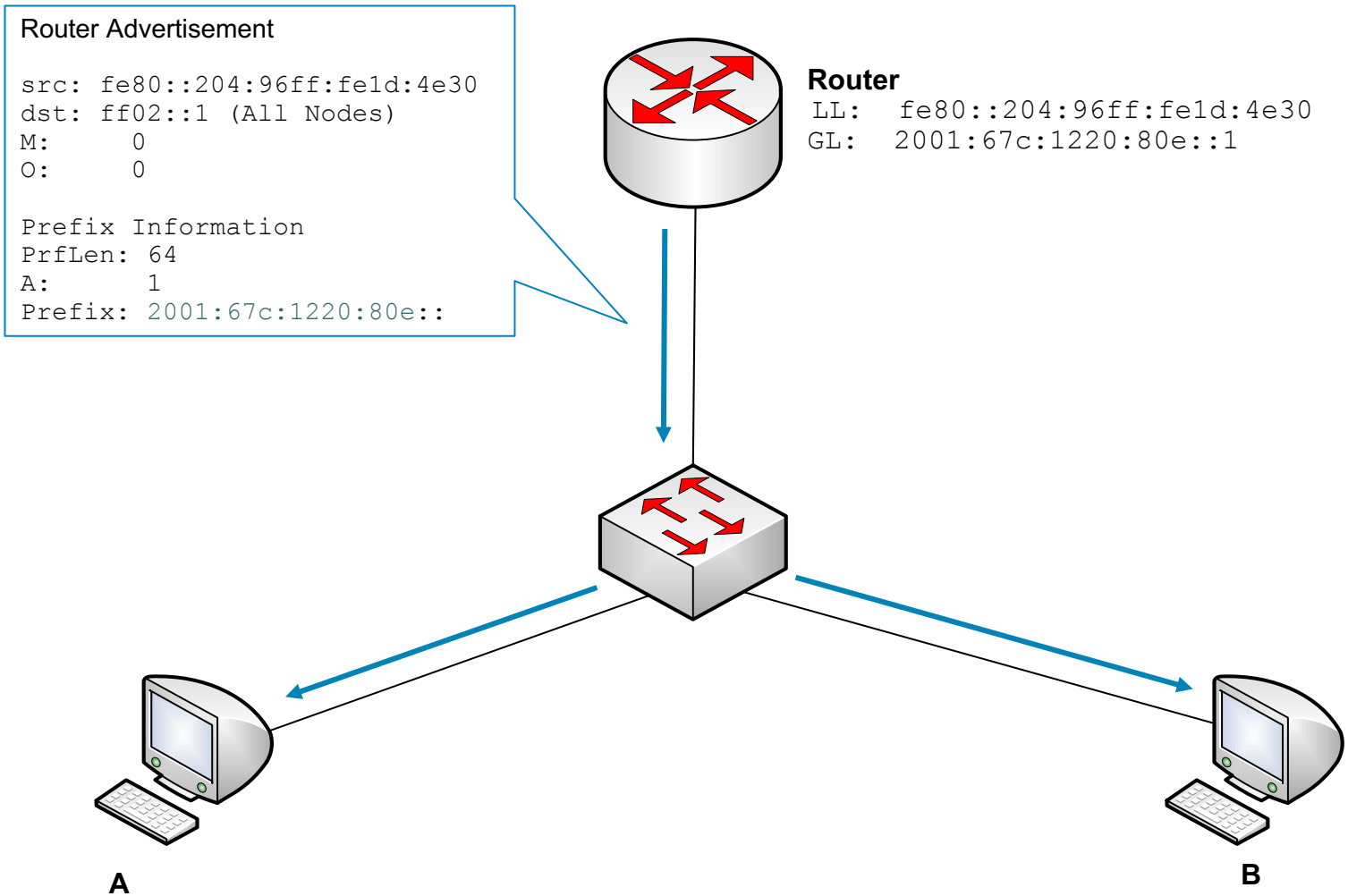
MLD Report for LL address



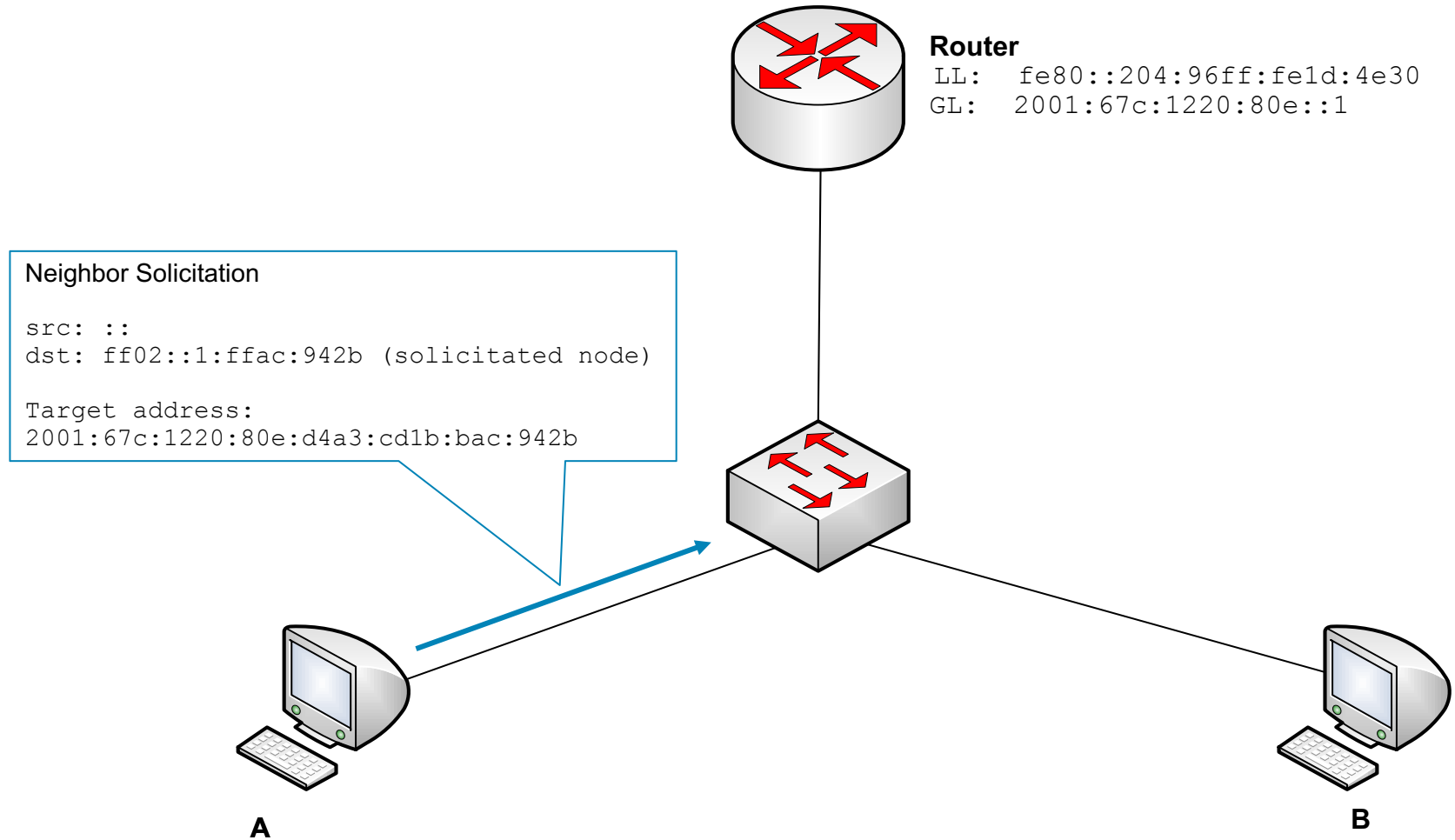
Global address



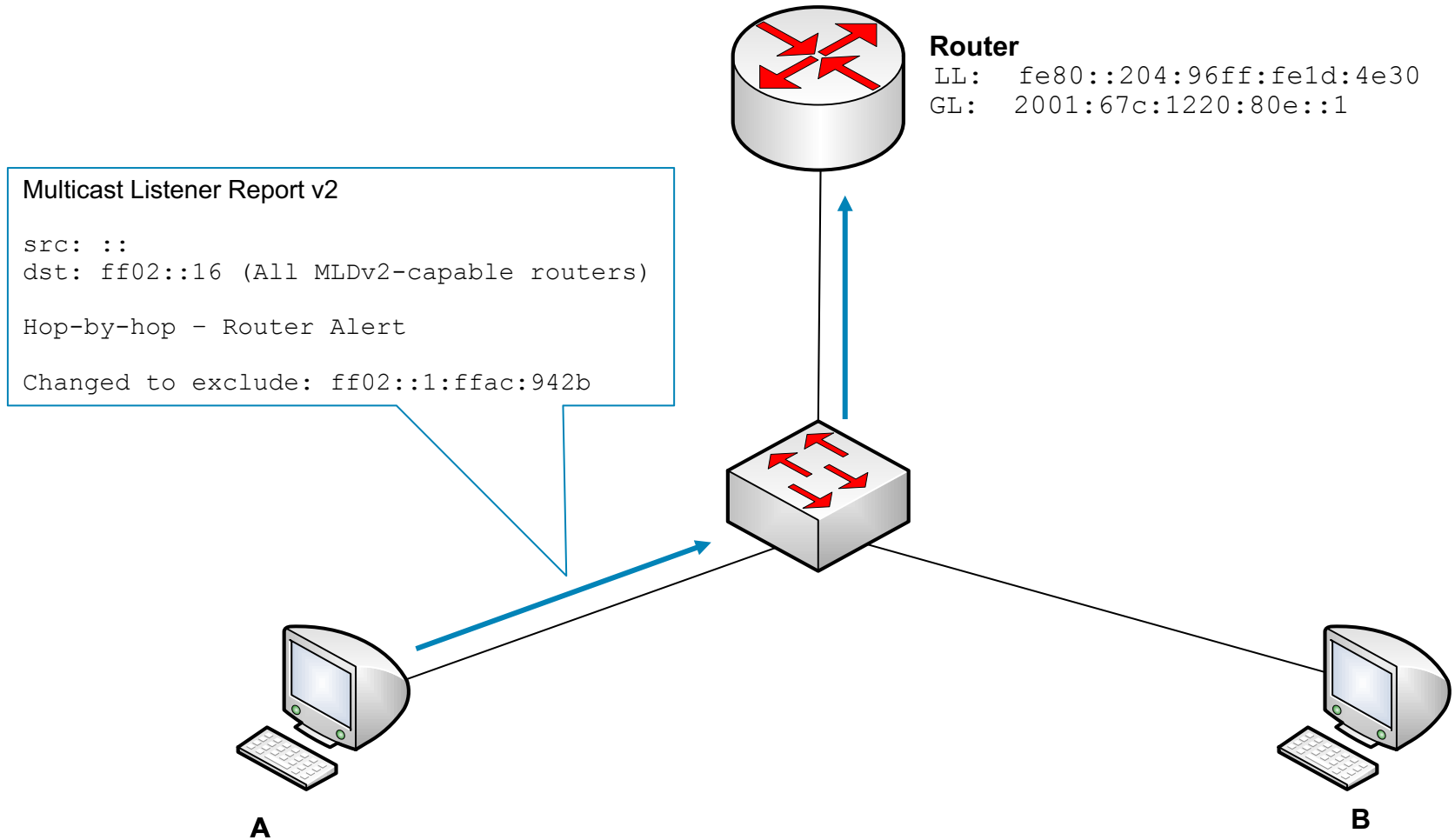
Global address (RFC 7217 only)



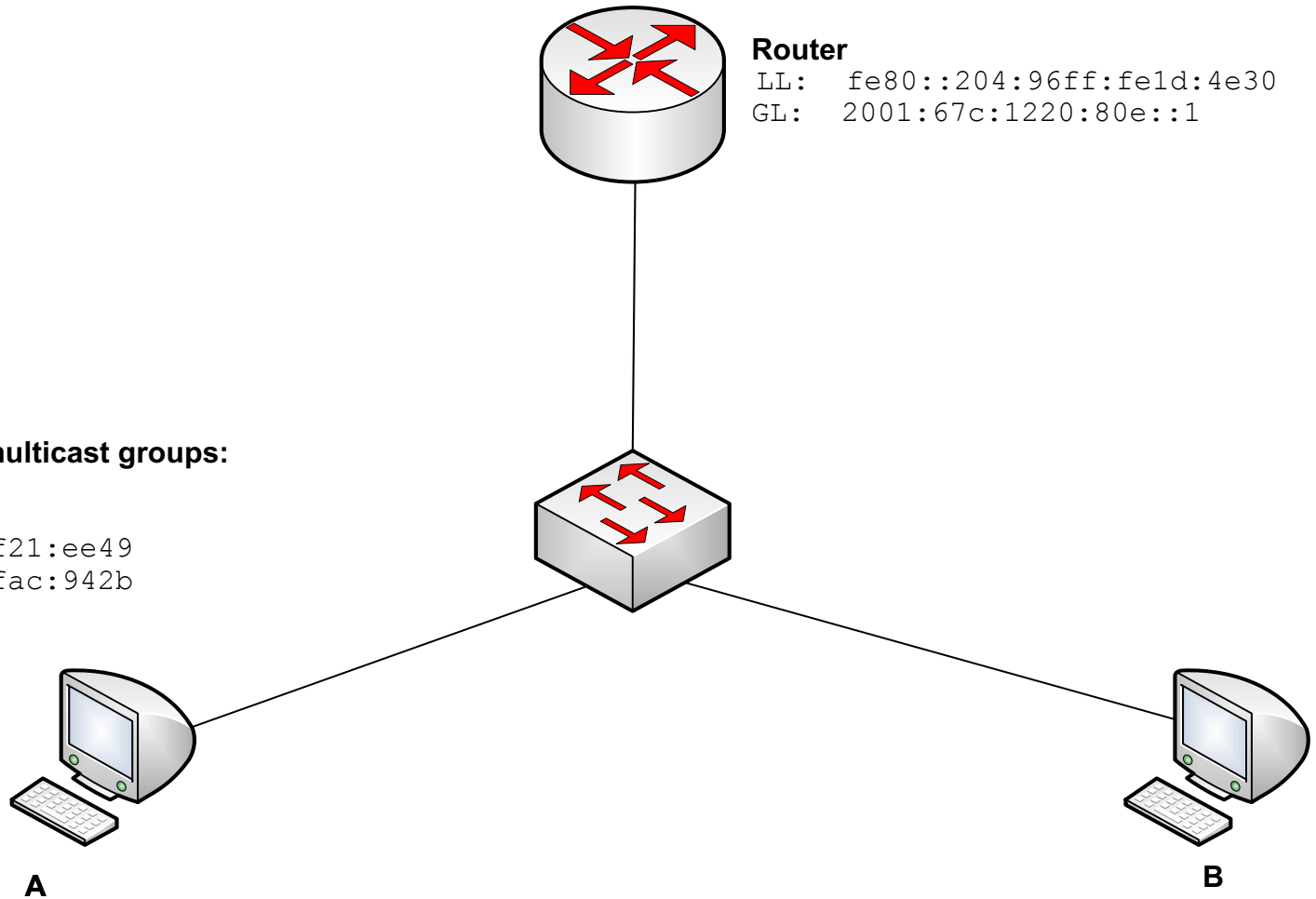
DAD for global address



MLD Report for global address



Final state


Router
LL: fe80::204:96ff:fe1d:4e30
GL: 2001:67c:1220:80e::1
A's joined multicast groups:
ff01::1
ff02::1
ff02::1:ff21:ee49
ff02::1:ffac:942b
A
LL: fe80::c9ee:98f6:d621:ee49
GL: 2001:67c:1220:80e:d4a3:cd1b:bac:942b
B

DHCPv6

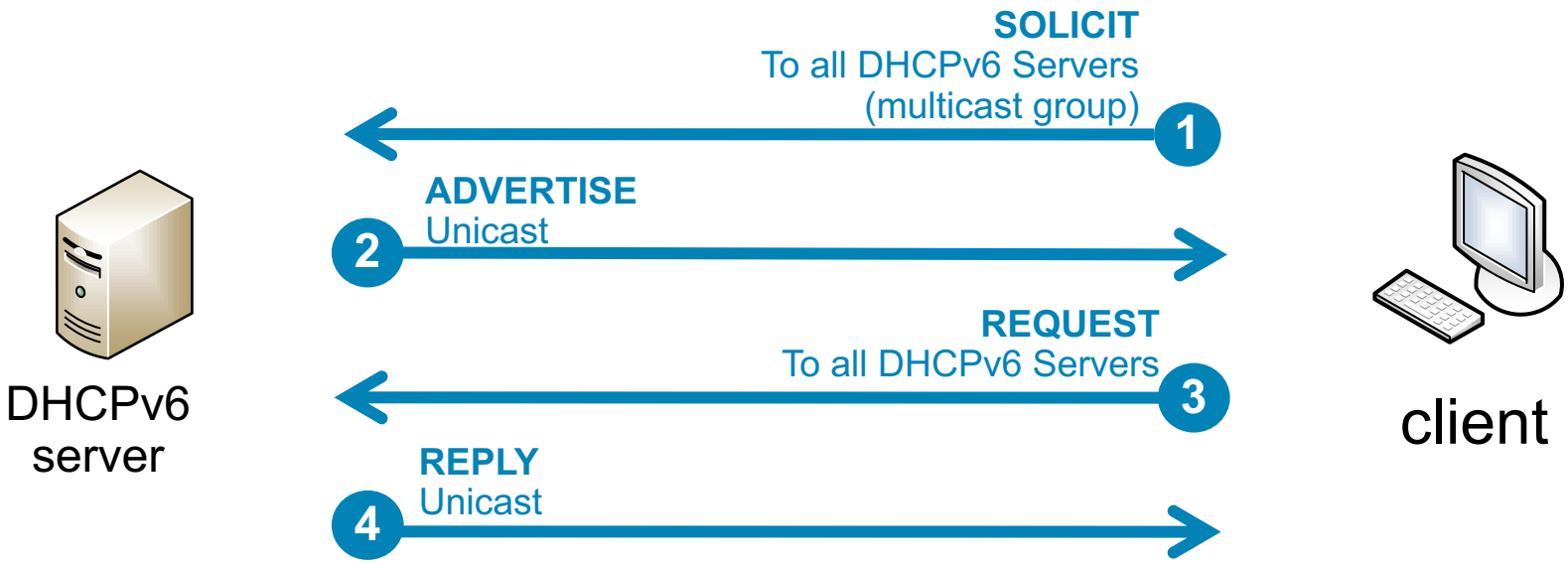
DHCPv6

- Stateless and stateful
- Tied with Route Advertisement – flags M/O
  - How should be these flags interpreted?
- Different client's identifier (DUID) compare to DHCPv4
- It is **not** mandatory – there could be devices that **does not** support DHCPv6 (e.g., Android)

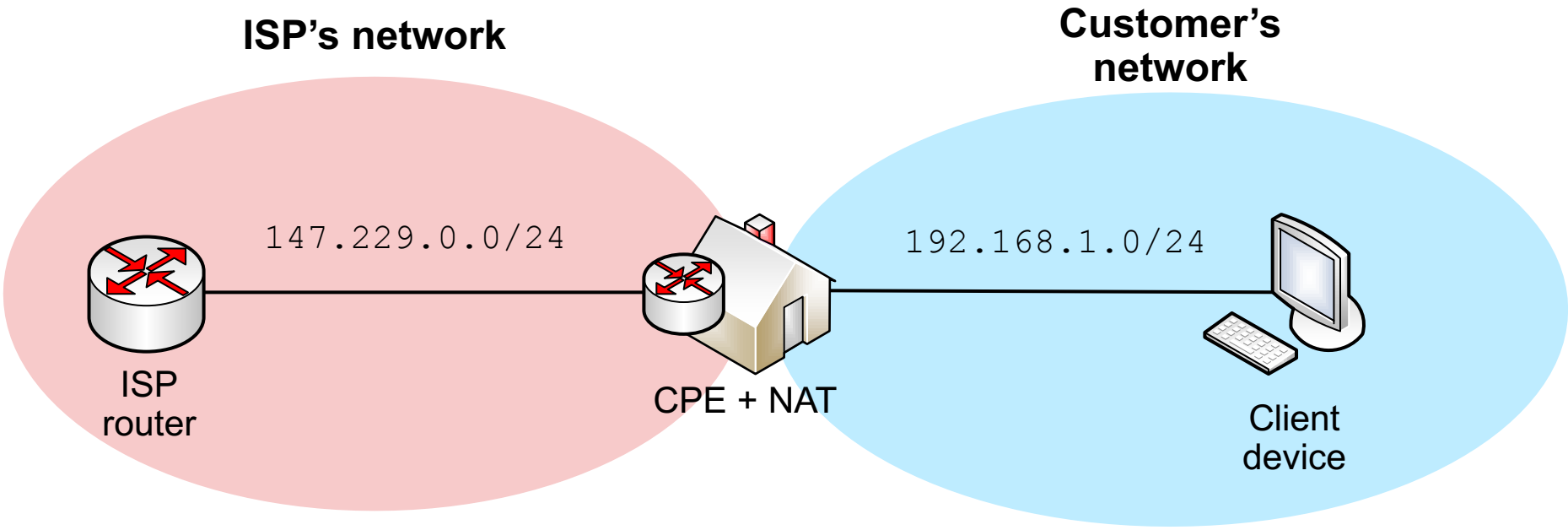
- **Stateless DHCPv6**
  - O flag in RA message set to 1
  - Information-request, Reply
  - Basically only for DNS name server + DNS search list

Stateful DHCPv6

- Managed flag in RA message set to 1
- Provide configuration parameters (IPv6 address, DNS, …), but **except** default gateway and prefix length
  - Where does a client find this information?

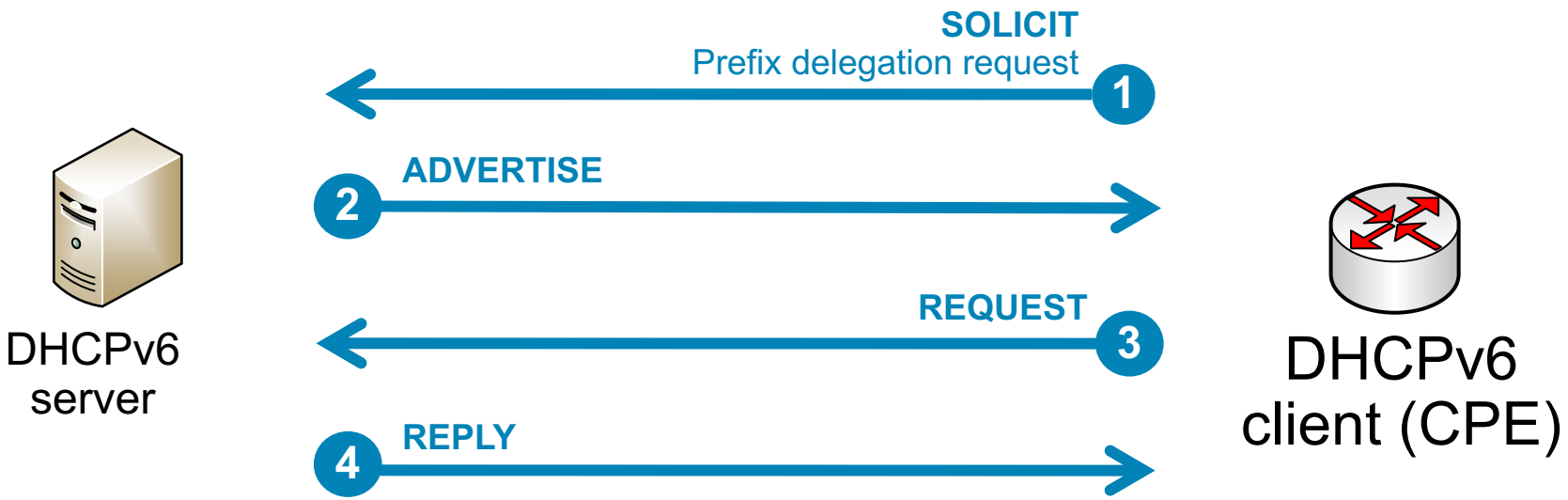


IPv4 world

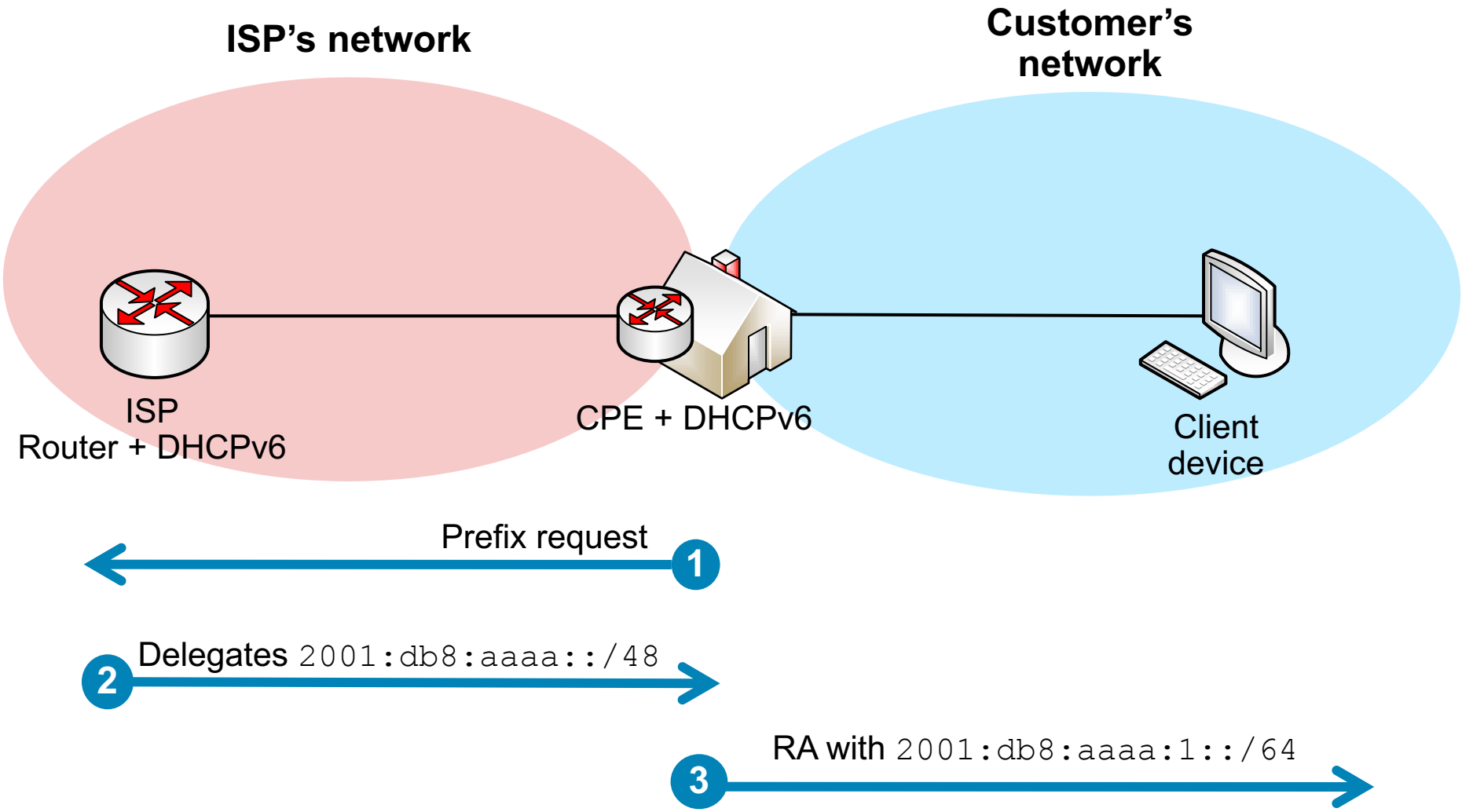

ISP's network
Customer's network
147.229.0.0/24
192.168.1.0/24
ISP router
CPE + NAT
Client device

DHCPv6 Prefix delegation

- Restores end-to-end reachability in IPv6
- DHCPv6 client requests IPv6 prefix that can use for addressing internal network
  - DHCPv6-PD does not address end clients – there must be another DHCPv6 server or SLAAC that uses delegated prefix



DHCPv6 Prefix delegation



DHCPv6 PD – questions

- What happens if DHCPv6 server does not run on ISP router?
- Does CPE requires global IPv6 connectivity on WAN interface?
- What should be the prefix size?
- Considering RA between ISP and CPE – should there be RA with M or O flag?

NAT

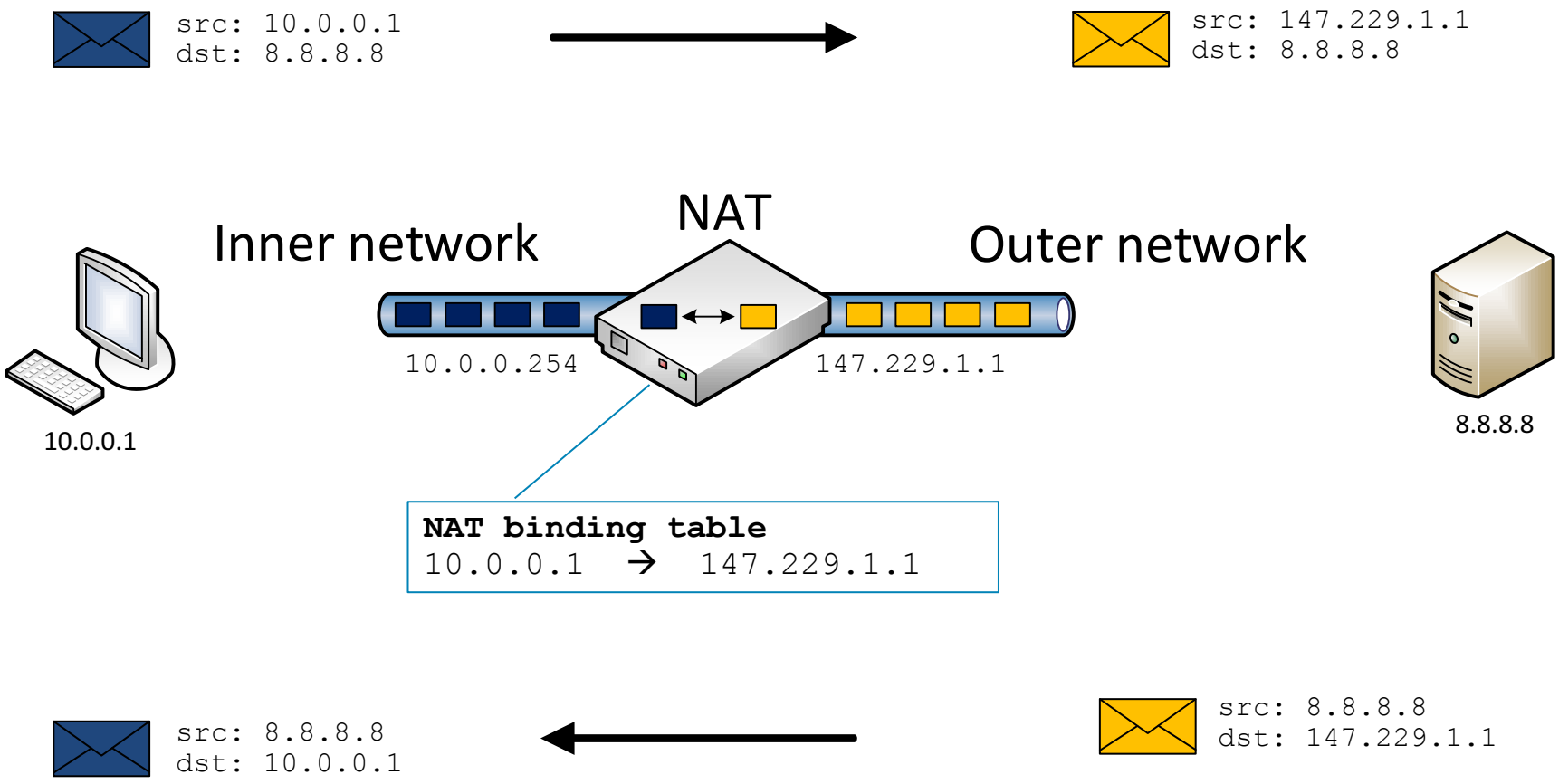
Network Address Translation

- There is no need for own address range, just one public address from ISP
- No renumbering problem
  - *When we change ISP, we do not need to change local addresses*
- Cheap multihoming – There is no need for own ASN
- Inside private addresses are not directly accessible from Internet
- NAT modifies IP, can modified TCP/UDP headers
- Destination NAT, port forwarding
- There is no standard – RFC 1631 is only Informational

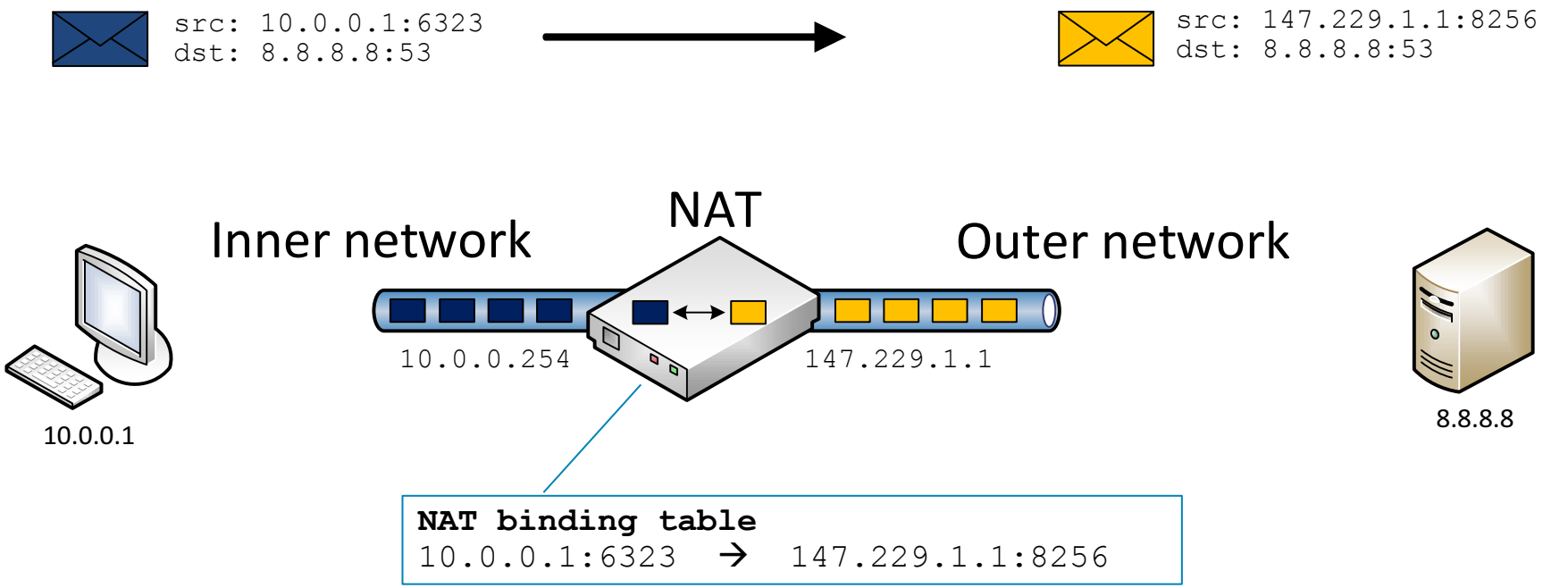
NAT

- Different variants are described in RFC 3489
  - One-to-One = Full Cone = Static
  - Restricted cone, Port Restricted cone, Symetric
- Real implementations use own algorithms (NAT behavior is not standardized)
- CGNAT
  - NAPT for ISP (NAT444)

NAT behavior



NAPT behavior



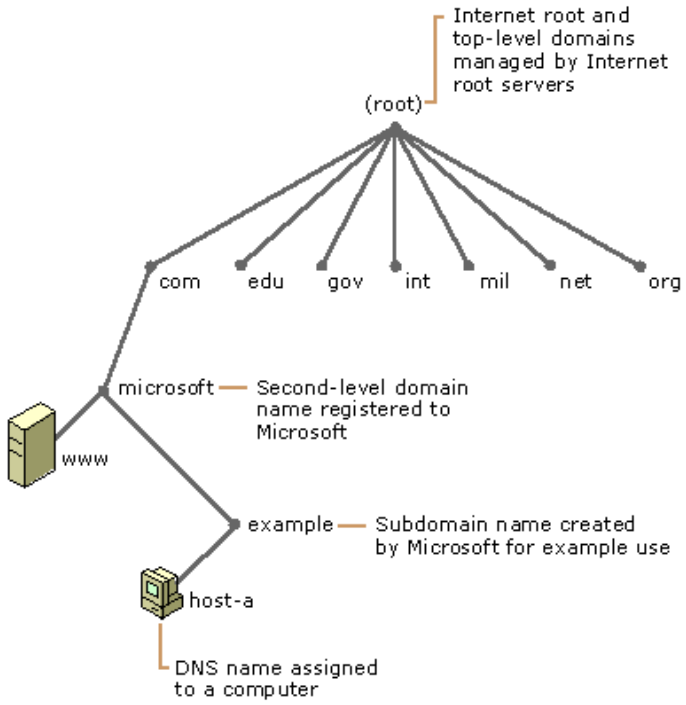
DNS

DNS

- **DNS service**
  - Maps domain name onto IP address and vice-versa
  - Database of names and IP addresses (and other records)
  - Database on special dedicated server – **DNS server**
- **DNS architecture**
  - Domain namespace
  - Distributed database spread over many servers
  - DNS protocol defines message format and resolution of queries

Domain Namespace

- Hierarchical structure
  - Tree with special root node
  - Domain = subtree
  - Domain name = path in the tree
  - Tree leaves = **fully-qualified names (FQDN)** of devices



Server types

- Whole namespace is divided across multiple DNS servers
  - Each one contains own zone file

- **Primary**
  - a.k.a **authoritative server** with zone file
  - Only one for each domain
- **Secondary**
  - Backup server with copy of zone in case of failure of primary server
- **Caching only**
  - Only relays queries and answer using its own cache of previously relayed answers
  - Provides non-authoritative answers

Root Servers

- Geoff Huston: The Root of the DNS
- http://root-servers.org/map/

DNS Query Resolution

- **Resolver**
  - Program obtaining answer from DNS server
  - Dispatches queries from clients using DNS protocol RFC 1035
- **Resolution**
  - = looking up for the answer in DNS
  - Tree-like hierarchy of DNS namespace
  - Recursive query – made by a client to a DNS server – server solves the resolutin and return an answer
  - Iterative query – DNS server to return the best answer it can give based on its cache or zone data

DNS – basic tools

- dig/nslookup

- Things to try:
  - Find the A, AAAA, MX record for a domain
  - Run own resolver and validates DNSSEC on your laptop
  - Try manual validation to learn about the process:
    - http://backreference.org/2010/11/17/dnssec-verification-with-dig/