NSWI184 – Řízení počítačových sítí

Přednáška dvanáctá

Ondřej Zajíček, Kateřina Kubecová

2025-01-07

Host network

- It should just work.™

Host network

- It should just work.™

What should work?

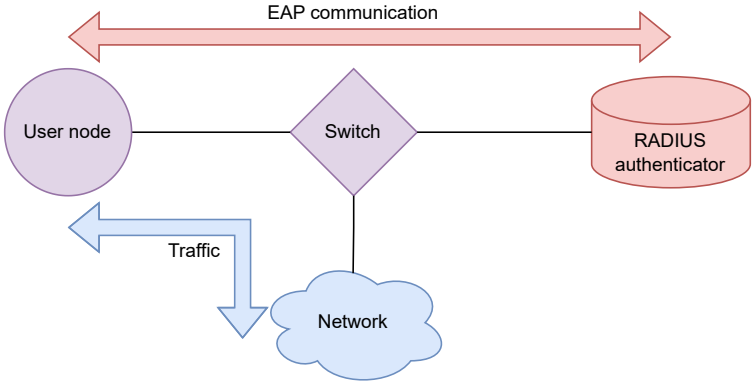
- Media access
- Address assignment
- Forwarding / routing
- DNS

Media access control: Wifi Protected Access

- Framework: WPA3 (802.11w), some older devices need WPA2 (802.11i)
- Authentication by BIP (Broadcast Identity Protocol)
- Encryption by CCMP (Counter Mode Cipher Block Chaining Message Authentication Code Protocol)
- Often used with a pre-shared key configuration
- *WEP, TKIP: deprecated, vulnerable to lots of attacks*
- Advanced uses:
  - Authenticate by EAP
  - Generate keys for BIP / CCMP from the EAP context

Media access control on wire: 802.1x

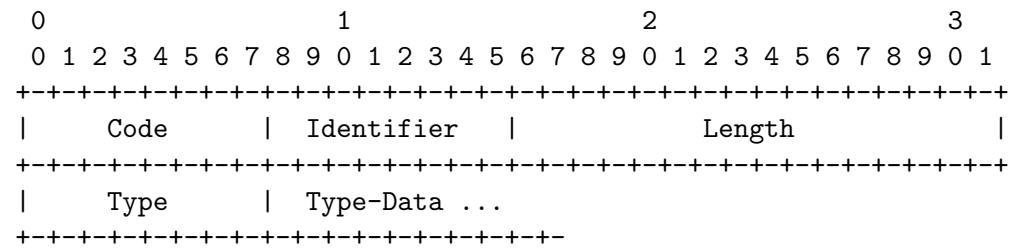
- Switch: Drop everything but EAP frames
- EtherType: 0x888e
- Wait until the device authenticates.
- Unblock the port.



EAP: Extensible Authentication Protocol

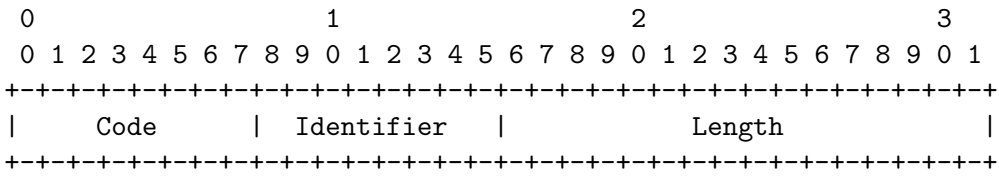
- Lower Layer: Ethernet, wifi, whatever, including PPP
- Nodes may forward EAP packets
- Upper layer: Lots of different methods
- Just a transportation layer
- RFC 3784

EAP Request, Response



- Code: 1 for Request (sent by the Authenticator), 2 for Response (sent by the Supplicant)
- Identifier: Unique for every request, response copies it back.
- Type: the actual authentication method, IDs assigned by IANA

EAP Success, Failure

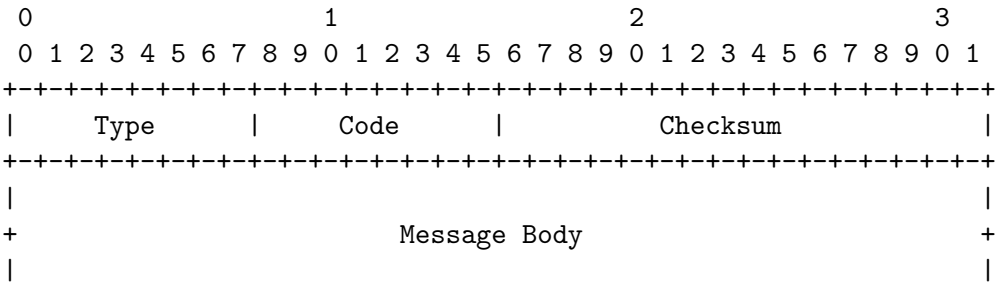


- Code: 3 for Success, 4 for Failure
- Identifier: which Response caused this result
- Type: the actual authentication method, IDs assigned by IANA
- Must be discarded unless allowed explicitly by the state machine

RADIUS: Remote Authentication Dial In User Service

- The switch (NAS) translates between EAP communication and RADIUS packets
- Available also for PPP, not dependent on EAP
- NAS has a trusted connection to its local RADIUS server (via UDP)
- Requests may be proxied to remote server $\rightarrow$ this is how Eduroam works
- RFC 2865, 3579

ICMP: Internet Control Message Protocol



- Short messages informing about the network status
- ICMPv6: RFC 4443, Next Header = 58
- ICMPv4: RFC 792, Protocol = 1

ICMP Error Messages

- Destination unreachable (with reasons)
- Packet too big
- Time Exceeded (hop count)
- Parameter problem: weird header, unknown protocol, unknown option

---

- All ICMPv6 Error Messages contain as much as possible of the original packet. (Only header + 64 bytes by default for ICMPv4.)
- Dropped "Packet too big" message = PMTUD fail

Ping and Pong: ICMP Echo

- Echo Request (Type=128), Echo Response (Type=129)
- Has an identifier (16b) to identify, and a sequence number (16b)
- May contain arbitrary data; the data must be sent back unmodified

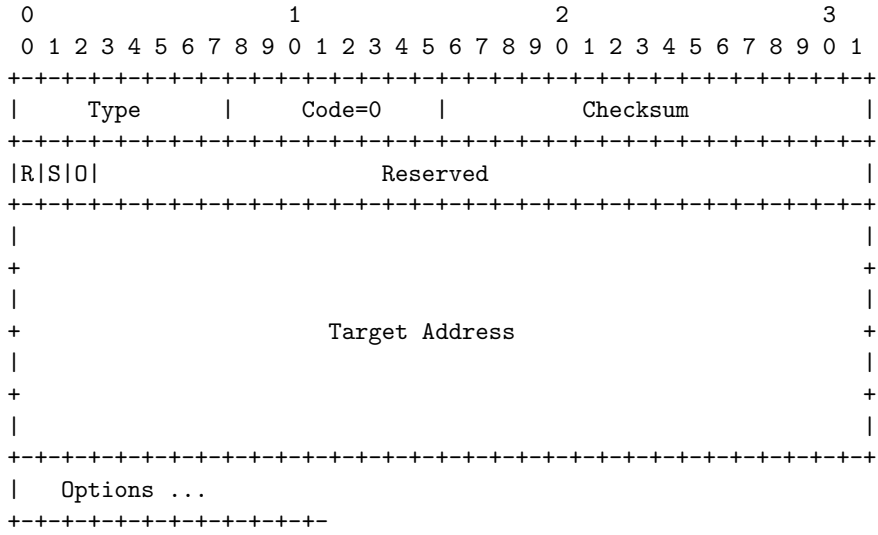
---

- RFC 4884: Extended ICMP
- MPLS Label Stack (RFC 4950)
- Interface information (RFC 5837)
- Node information (draft-ietf-intarea-extended-icmp-nodeid)

Link-Local Address Assignment

- Combine *some interface-specific value* with `fe80::/64` (reserved is /10 tho)
- For Ethernet: EUI-64
- Cryptographic: RFC 3972
- Privacy: RFC 4941 (random value)
- Stable privacy: RFC 7217
  - `CryptoHash(Prefix, Net_Iface, Network_ID, DAD_Counter, secret_key)`
  - Take 64 bits from bottom
- Perform Duplicate Address Detection
- Address is `tentative` until DAD succeeds

Neighbor Solicitation + Advertisement



Neighbor Solicitation + Advertisement

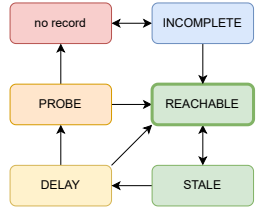
- RFC 4861
- Solicitation (Type=135): Is there anybody out there with this address?
  - Multicast for lookups, unicast for reachability check
  - Multicast address: "Solicited Node Address"
    (`ff02::1:ff<24-bit address suffix>`)
  - Received Solicitation where Target is me → reply with Advertisement.
- Advertisement (Type=136): I am here with this address.
  - Unicast as a response to Solicitation (S=1)
  - Multicast as unsolicited (S=0)
  - Override flag: The link-layer address has changed (O=1)
  - Router flag: The node is actually a router (R=1)
- Used for Neighbor Address Resolution and Unreachability Detection:
  - send Solicitation, wait for Advertisement
  - populate Neighbor Cache by IP + MAC

Duplicate Address Detection

- RFC 4862
- Send a bunch of solicitations with zero source address
- Received Solicitation and/or Advertisement $\rightarrow$ do not use this address.
- Failed DAD for Link-Local based on MAC address $\rightarrow$ do not use this interface.
  "... and trying to recover from it by configuring another IP address will not result in a usable network. In fact, it probably makes things worse by creating problems that are harder to diagnose ... "
- Enhanced DAD: add also a random-value Nonce to distinguish looped-back Solicitations (RFC 7527)
- Optimistic DAD: assign the address right away as Deprecated (RFC 4429)

Neighbor Cache

- Incomplete: sent Solicitation, no Advertisement received
- Reachable: OK
- Stale: timed out (or looks outdated) but nobody cares yet
- Delay: sent a packet using a Stale record, not acked yet
- Probe: trying to re-check, Solicitation sent

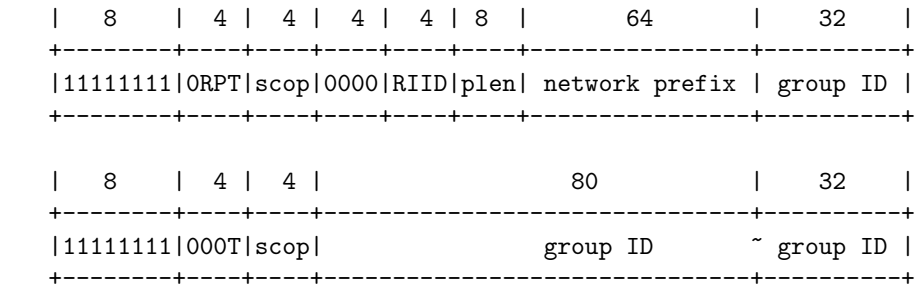


SEcure Neighbor Discovery

- An attempt to make neighbor discovery secure
- RFC 3971, 3972, 6494, 6495
- You have a /64 prefix assigned, with a certificate → calculate the lower 64 bits
- Include CGA option (data used to calculate the 64 bits) and RSA signature of the packet
- No implementation in Linux kernel, found in FreeBSD, Cisco and Juniper
- No data on actual deployment; uses SHA1 and RSA, not easily extensible

There are several more extensions to ICMPv6 with little or no use.

Multicast Address



- R: <network prefix>::RIID is the PIM-SM rendez-vous point (Protocol Independent Multicast Sparse Mode)
- P: the network prefix is the owner of this group
- T: dynamically assigned group address

For Ethernet: mapped to address 33:33:<last 32 bits of group ID>

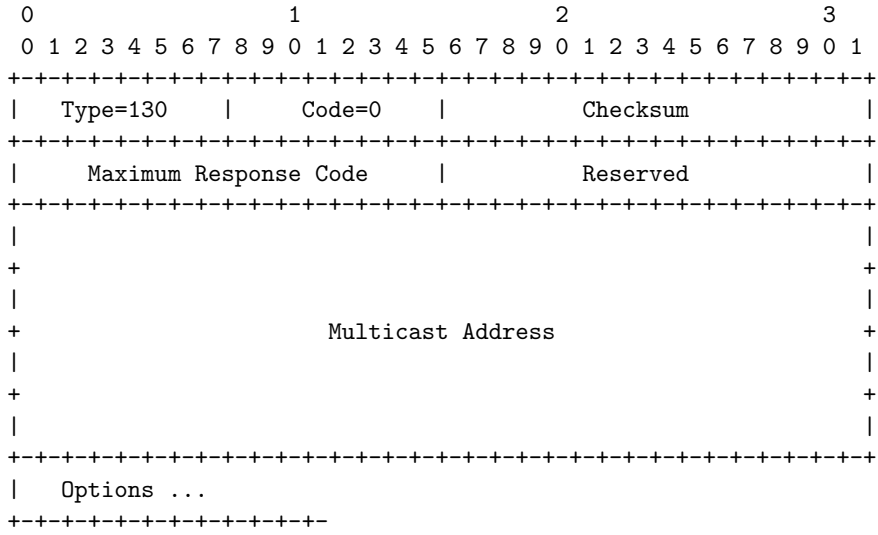
Multicast Address

Scope	
0	Reserved
1	Interface-Local scope
2	Link-Local scope
3	Realm-Local scope
4	Admin-Local scope
5	Site-Local scope
8	Organization-Local scope
E	Global scope
F	Reserved

Well-known addresses	
`ff0x::1`	Nodes
`ff0x::2`	Routers
`ff0x::5`	OSPF
`ff0x::6`	OSPF DR
`ff0x::9`	RIP
`ff0x::12`	VRRP
`ff0x::16`	MLDv2 nodes
`ff0x::fb`	mDNS
`ff0x::1:2`	DHCP relays and servers
`ff0x::1:3`	DHCP servers only
`ff0x::1:6`	Babel
`ff02::1:ffXX:XXXX`	Nodes with this suffix

. . . and many more

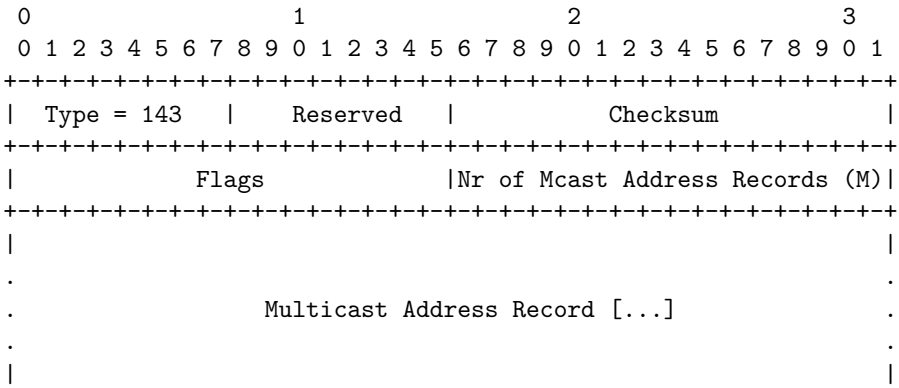
Multicast Listener Query



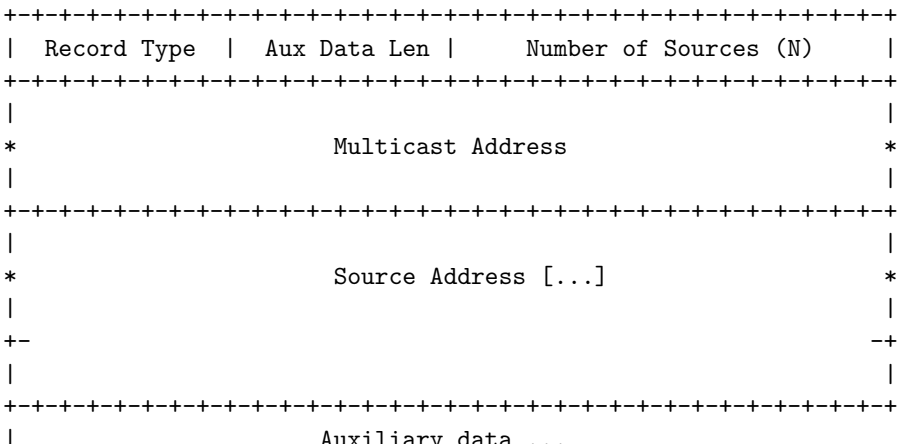
Multicast Listener Query

- Is there anybody listening to this group?
- Anybody listening should answer by Report
- May also include Source addresses for Source Specific Reporting
- Maximum Response Code is Maximum Response Delay in milliseconds
  - ≤ 32767: Directly the value
  - `1|e(3b)|m(12b)`: $(\mathrm{m} + 4096) \cdot 2^{3+\mathrm{e}}$
- Querier: The router with the lowest IPv6 address sends periodical queries

Multicast Listener Report



Multicast Listener Report Address Record

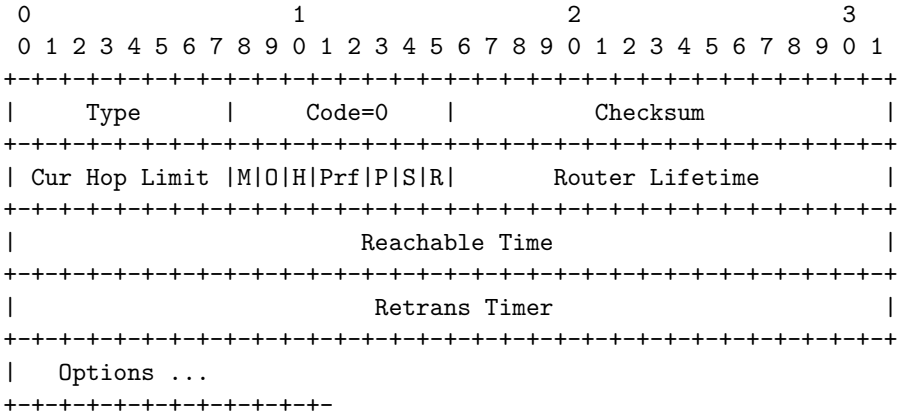


Multicast Listener Report

- Listening status info and changes
- Send as a response to Query, or when something changes
- Retransmitted several times, just to be sure
- Record Type:
  - Mode is Include (1) / Exclude (2): I do (not) listen for this address
  - Mode change to Include (3) / Exclude (4)
  - Allow (5) / Block (6) these sources

... more in RFC 9777

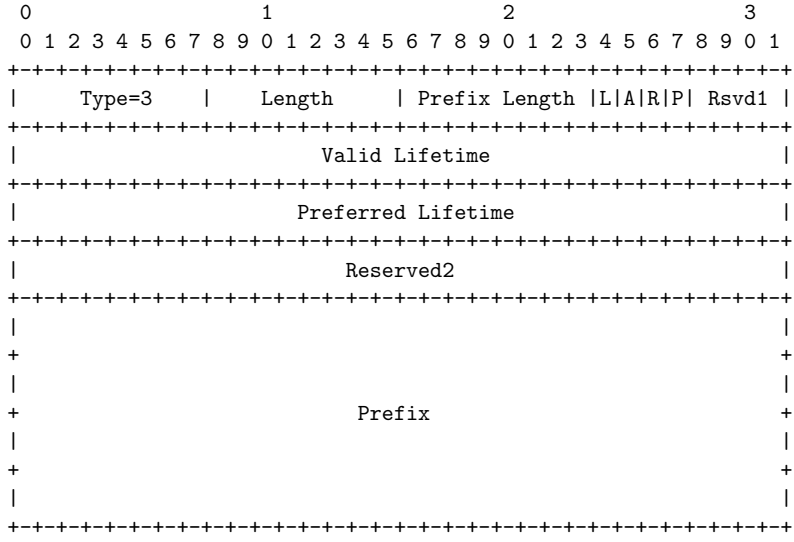
Router Solicitation + Advertisement



Router Solicitation + Advertisement

- Solicitation (Type=133): I want a router right now!
- Advertisement (Type=134): I'm a router.
  - M: Managed, use DHCPv6 to get an address
  - O: Other configuration available in DHCPv6 (e.g. DNS)
  - Prf: Preference (to choose between multiple RA's)
  - P: This is a proxied RA, don't proxy more
  - S: This is a SNAC router
  - Router Lifetime (seconds): For how long is this going to be valid (zero = don't use)
    RFC 7772: Minimum feasible value is 45–90 minutes due to battery saving.
  - Reachable Time: For how long the Neighbor Cache record is valid
  - Retrans Timer: How often do we send Advertisements
- Received Solicitation on a router $\rightarrow$ send an Advertisement.

ICMPv6 Prefix Option



StateLess Address AutoConfiguration using Prefix Option

- L: Addresses inside this prefix are on-link (use Neighbor Discovery for them)
- A: You may assign your address from this prefix
- P: We prefer if you asked for prefix delegation via DHCPv6 instead
- MTU option: Use this MTU on this link
- Lifetime (s):
  - Valid = for how long is this prefix available
  - Preferred = for how long you may use the auto-assigned address
- After assigning your address from the prefix, perform DAD.
- New advertisement $\rightarrow$ renew the lifetimes
- Preferred is over $\rightarrow$ mark as deprecated
- Valid is over $\rightarrow$ stop using the address at all

Notable ICMPv6 Options

- RDNSS, DNSSL: information about DNS resolvers and local domain (RFC 8106)
- Encrypted DNS: provider of DNS over TLS / HTTPS / QUIC
- Flags expansion: 48 more bits for router advertisements
- Captive Portal URL
- NAT64 prefix

Miscellaneous notes

- ICMPv6 Redirect message: there is a better router for this exact address *there*
- Stub Network Auto Configuration
  - Connect weird link types together
  - SNAC Router: a smart proxy filtering BUM
  - Still in a draft phase at the IETF SNAC WG: `https://datatracker.ietf.org/wg/snac/about/`
- Mobile IPv6: There is quite a lot of RFCs on address mobility with apparently little or no use.
- Homenet, KIRA, ...: Various zeroconf approaches.

DHCP: Stateful address autoconfiguration

- UDP port 546 (server to client), 547 (client to server)

- Information Request
  - triggered by O-bit in RA
  - multicast to *DHCP Servers and Relays*
- Simple address request
  - Client sends multicast Solicit indicating Expedited assignment
  - Server replies with Reply assigning the resources
  - The resources must be renewed periodically

Stateful Address AutoConfiguration: DHCP

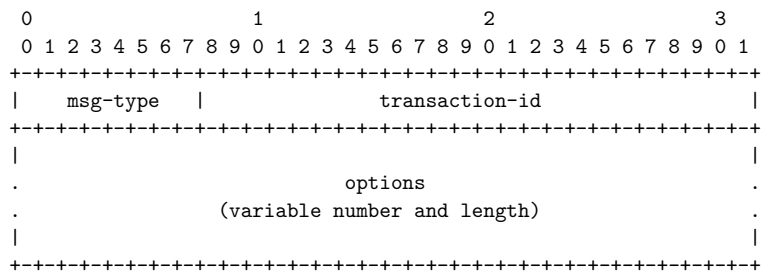
- Client sends Solicit message to find available servers
- Servers reply with Advertisement
- Client chooses one and sends Request
- Server replies with Reply assigning the resources
- The resources must be renewed periodically
- If renew is unavailable, client should rebind to another server

Stateful Address AutoConfiguration: DHCP

- Confirm: Client has an outdated resource and wants to check its validity
- Decline: Client refuses to use the assigned resources
- Release: Client has stopped using the assigned resources
- Reconfigure: Server has changed configuration, clients should renew / rebind / inforequest

The server replies by Reply.

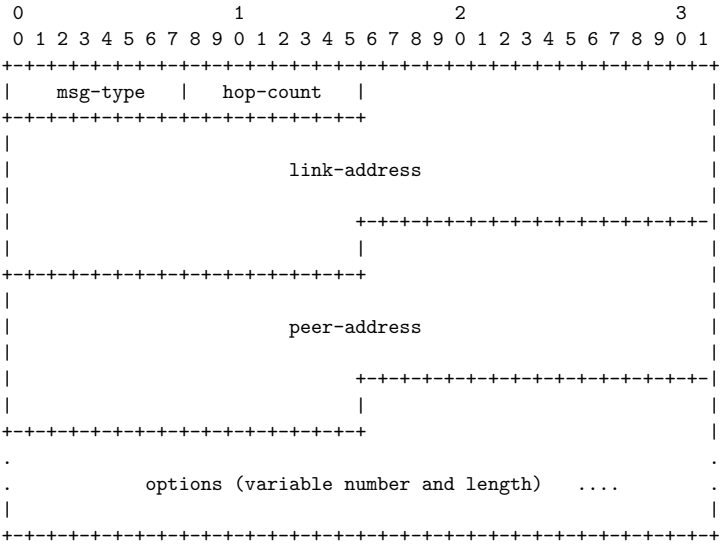
Stateful Address AutoConfiguration: DHCP



DHCP Notable Options

- Client ID, Server ID
- IA_NA, IA_TA, IA_PD:
  - Non-temporary and temporary addresses, prefix delegation
  - lifetimes similar to Router Advertisement
  - Client sends with hints
  - Server confirms or assigns the resources
- Rapid Commit: The Client wants the server to issue the resources immediately
- Option Request: The Client wants the server to include these options
- Preference (of the server among others)
- Server Unicast: The Client may contact the server at this address by unicast
- Status Code: success, failure, no address, no binding, not on link, no prefixes, use multicast

Stateful Address AutoConfiguration: DHCP Relay



QED