CERN DHCPv6 implementation

WLCG Workshop 2017 - Manchester

21st of June 2017

edoardo.martelli@cern.ch

Why DHCPv6 and not SLAAC

CERN DHCP[v6] servers offer a lease only to registered MAC addresses, because of:

- device tracking for
    - user's own security
    - user traceability
    - user support

- DNS and Firewall automation

- static address assignment (optional)

DHCPv6 client-server exchange

1. The **client** sends a **Solicit** message to the All_DHCP_Relay_Agents_and_Servers multicast address, requesting the assignment of addresses and other configuration information

2. The **server** responds with a **Reply** message that contains the confirmed addresses and configuration. Each address assigned to the client has associated preferred and valid lifetime

...

3a. Periodically, the **client** sends a **Renew** message to the server to request an extension of the lifetimes of an address

3b. The **server** sends a **Reply** message to the client with the new lifetimes, allowing the client to continue to use the address without interruption

[RFC3315]

DHCPv6 Options

- IPv6 address, OPTION_IAADDR *[RFC3315]*

- DHCPv6 timers:
  - preferred-lifetime (time a valid address is preferred)
  - valid-lifetime (time an address remains in valid state)
  [RFC2462]

- DNS servers, OPTION_DNS_SERVERS *[RFC3646]*

Other options exist

DHCPv6 client-server exchange

tcpdump -i eth0 -vvv -n ip6 and port 547

3:52:22.644967 IP6 (hlim 60, next-header UDP (17) payload length: 106) 2001:db8:305:10::2.dhcpv6-server > 2001:db8:1000::9.dhcpv6-server: [udp sum ok] dhcp6 relay-fwd (linkaddr=2001:db8:221:5::1 peeraddr=fe80::16:3eff:fe01:b3a1 (relay-message (**dhcp6 solicit** (xid=d3c81a (client-ID hwaddr/time type 1 time 550227784 02163e01b3a1) (option-request DNS DNS-name) (elapsed-time 0) (IA_NA IAID:1040298913 T1:3600 T2:5400))) (interface-ID 01252707...))

13:52:22.645383 IP6 (hlim 64, next-header UDP (17) payload length: 187) 2001:db8:1000::9.dhcpv6-server > 2001:db8:305:10::2.dhcpv6-server: [udp sum ok] **dhcp6 relay-reply** (linkaddr=2001:db8:221:5::1 peeraddr=fe80::16:3eff:fe01:b3a1 (interface-ID 01252707...) (relay-message (dhcp6 advertise (xid=d3c81a (IA_NA IAID:1040298913 T1:0 T2:0 (**IA_ADDR 2001:db8:d0:9::100:13e pltime:604800 vltime:604800**)[|dhcp6ext]) (client-ID hwaddr/time type 1 time 550227784 02163e01b3a1) (server-ID hwaddr/time type 1 time 455785913 b4b52f67dfbc) (**DNS 2001:db8:1100::5 2001:db8:1200::5**) (DNS-name))))

Not existing DHCPv6 Options

Compared to DHCP for IPv4, DHCPv6 **cannot** send these Options

- **Default gateway** (critical!)

- **Netmask length** (some clients may need to know it)

They must be communicated by other means

DHCPv6 relies on RAs

RAs needed to
- communicate default-gateway address
- prefix netmask length

CERN doesn't want SLAAC to happen, so routers are configured to send RAs that:
- tell clients to not use SLAAC, but DHCPv6 (M=1)
- tell clients the link local address of the default gateway
- tell clients the netmask of the prefix, but to not use it for SLAAC (A=0)

RAs flags

M: 1-bit "**Managed address configuration**" flag. When set, indicates that addresses are available via DHCPv6. If the M flag is set, the O flag is redundant and can be ignored because DHCPv6 will return all available configuration information

O: 1-bit "**Other configuration**" flag. When set, indicates that other configuration information is available via DHCPv6. Examples of such information are DNS-related information or information on other servers within the network

L: 1-bit "**On-link**" flag, defined in the PIO (Prefix Information Option). When set, indicates that this prefix can be used for for determining whether another address shares the same link (on-link determination)

A: 1-bit "**Autonomous**" flag, defined in the PIO (Prefix Information Option). When set, it indicates that the prefix can be used for stateless address configuration.

[RFC4861, draft-dhcpv6-slaac-problem]

RAs configuration examples

Brocade NetIron

```
int ve 2
 ipv6 nd managed-config-flag
 ipv6 nd other-config-flag
 ipv6 nd prefix-advertisement 2001:db8:1:2::/64 2592000 604800 onlink
```

HPE Aruba ProVision

```
vlan 2
 ipv6 nd ra managed-config-flag
 ipv6 nd ra other-config-flag
 ipv6 nd ra prefix default no-advertise
 ipv6 nd ra prefix 2001:db8:1:3::/64 2592000 604000 no-autoconfig
 ipv6 nd reachable-time 15000
```

RA packet

```
# tcpdump -i eth0 -vvv -n ip6 host ff02::1

17:08:11.734729 IP6 (hlim 255, next-header ICMPv6 (58) payload length: 112)
fe80::215:60ff:feed:ce00 > ff02::1: [icmp6 sum ok] ICMP6, router advertisement,
length 112
    hop limit 64, Flags [managed, other stateful], pref medium, router lifetime
1800s, reachable time 0s, retrans time 0s
      source link-address option (1), length 8 (1): 00:15:60:ed:ce:00
        0x0000:  0015 60ed ce00
      prefix info option (3), length 32 (4): 2001:db8:21:7a::/64, Flags [onlink],
valid time 2592000s, pref. time 604000s
        0x0000:  4080 0027 8d00 0009 3760 0000 0000 fd01
        0x0010:  1458 0201 007a 0000 0000 0000 0000
      prefix info option (3), length 32 (4): 2001:db8:22:229::/64, Flags [onlink],
valid time 2592000s, pref. time 604000s
        0x0000:  4080 0027 8d00 0009 3760 0000 0000 2001
        0x0010:  1458 0202 0229 0000 0000 0000 0000
      dnssl option (31), length 24 (3):  lifetime 1200s, domain(s): db8.org.
        0x0000:  0000 0000 04b0 0463 6572 6e02 6368 0000
        0x0010:  0000 0000 0000
```

RAs limitations

In multi-homed networks and shared media, it’s not possible to assign different gateways to given clients

In networks sharing different IPv6 prefixes, it’s not possible to not assign certain addresses to given clients

All available prefixes are exposed to the clients

MAC address identification

CERN client devices are identified by their MAC address

When using DHCPv6 relays, the source MAC address of the supplicant is lost: the multicast request is transformed in an unicast packet sourced by the router, to allow the DHCPv6 Lease to come back to the correct router

The clients can put their own MAC (Link Layer) address in the **DUID field** of the DHCPv6 Solicit

DUID types

1. DUID-LLT: Link-layer address plus time

2. DUID-EN: Vendor-assigned unique ID based on Enterprise Number

3. DUID-LL: Link-layer address

4. DUID-UUID: UUID-Based DUID

[RFC3315, RFC6355]

Example of Solicit with DUID-LLT

tcpdump -i eth0 -vvv -n ip6 and port 547

3:52:22.644967 IP6 (hlim 60, next-header UDP (17) payload length: 106)
2001:db8:305:10::2.dhcpv6-server > 2001:db8:1000::9.dhcpv6-server: [udp sum ok] dhcp6
relay-fwd (linkaddr=2001:db8:221:5::1 peeraddr=fe80::16:3eff:fe01:b3a1 (relay-message
(**dhcp6 solicit** (xid=d3c81a (**client-ID hwaddr/time type 1 time 550227784 02163e01b3a1**)
(option-request DNS DNS-name) (elapsed-time 0) (IA_NA IAID:1040298913 T1:3600 T2:5400)))
(interface-ID 01252707...))

Problems with requiring DUID-LL[T]

- Not all the DHCPv6 client implementation allows the users to set a preferred DUID

- DHCPv6 clients don’t have to use DUID-LL[T] by default

- DHCPv6 clients may not put in the DUID-LL[T] the MAC address of the interface they send the request from

RFC 6939 to the rescue

RFC6939: “Client Link-Layer Address Option in DHCPv6”

It specifies the format and mechanism that is to be used for encoding the client link-layer address in DHCPv6 Relay-Forward messages by defining a new DHCPv6 Client Link-Layer Address option.

Must be implemented by the router relaying the DHCPv6 requests

Very few implementations so far. Ask your router vendor to implement it

RFC 6939 configuration examples

Brocade NetIron from 6.0:

```
interface ve 1
  ipv6 dhcp-relay destination 2001:DB8::9
  ipv6 dhcp-relay destination 2001:DB8::A
  ipv6 dhcp-relay include-options client-mac-address
```

HPE ProVision from K15.16.0013:

```
dhcpv6-relay
dhcpv6-relay option 79
```

ISC DHCPv6 server configuration example

```
#
# CERN dhcpv6 C4.3 configuration file
#
dhcpv6-lease-file-name "/etc/dhcpd/dhcpd6.leases";
pid-file-name "/etc/dhcpd/dhcpdv6.pid";
# By default not auth, but auth for PB services
not authoritative;
deny bootp;
ignore unknown-clients;
# Services configuration
option dhcp6.domain-search "cern.ch";
option dhcp6.sntp-servers 2001:db8:1040::69,2001:db8:1140::69;
# Default DHCP timers
min-lease-time 86400;
max-lease-time 86400;
default-lease-time 86400;
preferred-lifetime 86400;
# T1 and T2 timers
option dhcp-renewal-time 32400;
option dhcp-rebinding-time 48600;
option client-class-information code  97 = string;
log-facility local1;
omapi-port 7912;
omapi-key dhcpuser;
```

ISC DHCPv6 server configuration example

```
# DNS updates for dyndns6.cern.ch and 2001:db8:202::/48
ddns-update-style standard;
deny client-updates;
ddns-domainname "dyndns6.cern.ch";
ddns-ttl 60;
# Update AAAA and PTR records also for static leases
update-static-leases on;
# Allow DDNS updates without DHCID check
update-conflict-detection false;
# Update with every renewal/assignment
update-optimization false;
one-lease-per-client true;
# Supply the defined hostname to the client
use-host-decl-names true;
# Dynamic DNS updates information
key "dhcp6-ddns" {
    algorithm hmac-md5;
    secret "abcdzyx";
};

zone dyndns6.cern.ch. {
    primary 11.1.1.7;
    key "dhcp6-ddns";
}

zone 2.0.2.0.8.b.d.0.1.0.0.2.ip6.arpa. {
    primary 11.1.1.7;
    key "dhcp6-ddns";
}
```

ISC DHCPv6 server configuration example

```
option space MSFT;
option MSFT.release-on-shutdown code 2 = unsigned integer 32;

if (option vendor-class-identifier = "MSFT 5.0") {
    vendor-option-space MSFT;
    option MSFT.release-on-shutdown 1;
}

key "dhcpuser" {
        algorithm HMAC-MD5.SIG-ALG.REG.INT;
        secret abcdefxyz;
};

option space LINUX;
option LINUX.pxelinux-magic code 208 = string;
option LINUX.pxelinux-configfile code 209 = text;
option LINUX.pxelinux-pathprefix code 210 = text;
option LINUX.pxelinux-reboottime code 211 = unsigned integer 32;
option LINUX.discovery-control code 6 = unsigned integer 8;
option LINUX.boot-servers code 8 = array of { unsigned integer 16, unsigned integer 8, ip-address };
option client-architecture code 93 = unsigned integer 16;


# This section enables blocking at the pool level. The access-list is applied in
# the pools after all other allow/deny statements making it the most authoritative
class "blocked-list" {
        match hardware;
        log ( debug, concat("BLOCKED: ", hardware));
}

subclass "blocked-list" 01:00:00:11:22:33:44;
subclass "blocked-list" 01:00:02:11:22:33:55;
```

ISC DHCPv6 server configuration example

```
subnet6 2001:db8:1::/64 {
    option dhcp6.name-servers 2001:db8:1000::5,2001:db8:1100::5;
}

subnet6 2001:db8:2::/64 {
    option dhcp6.name-servers 2001:db8:1000::5,2001:db8:1100::5;
}

shared-network "Shared-S104" {
    authoritative;
    subnet6 2001:db8:3::/64 {
        option dhcp6.name-servers 2001:db8:1000::5,2001:db8:1100::5;
    }
    subnet6 fd01:db8:4::/64 { }
    subnet6 2001:db8:5::/64 {
        pool6 {
            deny unknown-clients;
            deny members of "blocked-list";
            option dhcp6.name-servers 2001:db8:1000::5,2001:db8:1100::5;
            range6 2001:db8:5::101:0/112;
        }
    }
}

group { # Static Clients
  host HOST1 { hardware ethernet 00:11:22:33:44:55; fixed-address6 2001:db8:1::100:42; option host-name "host1"; }
  host HOST1-RFC6939 { host-identifier v6relopt 1 dhcp6.client-linklayer-addr 00:01:00:11:22:33:44:55; fixed-address6
2001:db8:5::100:42; option host-name "host1"; }
}
```

Limitations of using DHCPv6

Android doesn’t provide any DHCPv6 client yet

Simple IOT devices may not have a DHCPv6 client implemented

Future developments

DHCPv6bis IETF working group:
https://www.ietf.org/mailman/listinfo/dhcpv6bis

https://tools.ietf.org/html/draft-ietf-dhc-rfc3315bis-08
This document updates the text from RFC3315, the original DHCPv6 specification, and incorporates **prefix delegation** (RFC3633), **stateless DHCPv6** (RFC3736), an option to specify **an upper bound for how long a client should wait before refreshing information** (RFC4242), a mechanism for **throttling DHCPv6 clients when DHCPv6 service is not available** (RFC7083), and clarifies the interactions between modes of operation (RFC7550)

ISC DHCPv6 servers

ISC develops two DHCP server software:
- DHCP – most used, but almost in maintenance mode
- KEA – new, active development

https://ripe74.ripe.net/wp-content/uploads/presentations/140-kea-ripe74-final.pdf

Questions?

edoardo.martelli@cern.ch