Alternatives to DHCPv6 Failover

This document provides alternatives to Dynamic Host Configuration Protocol Version 6 (DHCPv6) failover, as no DHCPv6 failover implementation is presently available.

One of the main reasons for DHCPv4 failover was to avoid having to split the address space between two DHCPv4 servers. This was problematic because IPv4 address space was (and now is even more so) a scarce resource and few could obtain twice the address space they required. DHCPv4 failover eliminated the need to split the address space, though it does require a small increase in the address space in order to allow each failover partner some pool of addresses to use when unable to communicate with its partner. IPv6 generally does not have these limitations.

There are also some other important differences between DHCPv4 and DHCPv6 or, more properly, between IPv4 and IPv6. In particular IPv6 provides for graceful address (or prefix) transition. IPv4 had a general concept of one address per interface. IPv6 does not have this concept - an interface is expected to have multiple addresses and each has a lifetime. Addresses transition between states (see RFC 4862):

- Tentative to preferred (this occurs after testing for address conflicts). An address in the preferred state can be used for communication.
- Preferred to deprecated (this occurs when the preferred lifetime has expired). An address in the deprecated state may continue to be used for existing communications but should not be used for new ones.
- Deprecated to invalid (when the valid lifetime has expired the address is removed and is no longer usable).

Basic DHCPv6 Operation Overview

A DHCPv6 client sends a Solicit message and waits for one or more Advertise messages. The client selects one of the Advertise messages, and sends a Request message. The server selected by the client then responds with a Reply message with the lease or leases that the client has been assigned. This is very similar to the DHCPv4 DHCPDISCOVER/DHCPOFFER and DHCPREQUEST/DHCPACK exchange used with DHCPv4.

Two Servers - Split Address Space

One technique for redundancy in DHCPv6 is for the two servers to split the address space. For example, if devices are being assigned addresses from a /64, one server could be given the **64-bit-prefix**:0:0:0:0/65 from which to allocate addresses and the other server the **64-bit-prefix**:8000:0:0:0/65 from which to allocate addresses.

When a client gets a lease, it will get it from one of the servers and will usually be able to renew the address from that same server. Only if that server is down during the renewal period (50 percent to 85 percent of the lease time; see the “Longer Lease Times” section below) will the client likely end up having to obtain a new address.

But if the client gets a lease from a different server, it will transition to that new address gracefully per the IPv6 rules. (Note: Some clients, such as cable modems, do not necessarily follow the IPv6 address transition rules and may switch to the new address immediately.)

There are two possibilities for configuring Cisco Network Registrar DHCPv6 servers for split address spaces:

Configure /64 prefixes with a /65 range (recommended)

Configure /65 prefixes (not recommended)

Configuring the /64 prefixes with a /65 range is the recommended approach. The reason for this is that, if a DHCPv6 client issues a Confirm message and the /64 prefixes are configured, either server will respond with success to the Confirm message if the client has an address in that /64 prefix. If instead the /65 prefixes are configured, one of the servers will respond with NOT ON LINK, which is not desirable (as this will cause the client to solicit a new address).

An example of prefixes configured for address assignment is shown below (note that many unset/default attributes have been elided for clarity):

On Server 1:

```
nrcmd> prefix server1
100 Ok
server1:
    address = 2001:db8::/64
    dhcp-type = [default=dhcp]
    range = 2001:db8::/65
```

On Server 2:

```
nrcmd> prefix server2
100 Ok
server2:
    address = 2001:db8::/64
    dhcp-type = [default=dhcp]
    range = 2001:db8:0:0:8000::/65
```

Note: The above applies to prefix delegation as well. For example, if you wanted to delegate prefixes from a /48 prefix, you would configure a prefix delegation prefix on both servers with the /48 prefix but one server with range of **48-bit-prefix**:0:0:0:0:0/49 and the other with the range **48-bit-prefix**:8000:0:0:0:0/49.

An example of prefixes configured for prefix delegation is shown below (note that many unset/default attributes have been elided):

On Server 1:

```
nrcmd> prefix server1-pd
100 Ok
server1-pd:
    address = 2001:db8::/32
    dhcp-type = prefix-delegation
    range = 2001:db8::/33
```

On Server 2:

```
nrcmd> prefix server2-pd
100 Ok
server2-pd:
    address = 2001:db8::/32
    dhcp-type = prefix-delegation
    range = 2001:db8:8000::/33
```

Two Servers - Interface-Identifier Addresses

If only doing address assignment (not prefix delegation), it may be possible to use the client's IPv6 link-local address interface identifier to generate the address (see RFC 4291, Appendix A, on EUI-64 formatted interface identifiers). Both servers would then always generate the same address and thus the client could use either server just as easily.

To enable this, a prefix's allocation algorithm attribute must be configured to use only interface identifiers.

An example of a prefix configured for interface-identifier-based address assignment is shown below (note that many unset/default attributes have been elided):

```
nrcmd> prefix sample1
100 Ok
sample1:
    address = 2001:db8::/64
    allocation-algorithms = interface-identifier
    dhcp-type = [default=dhcp]
```

Two Servers - Using Client Reservations

Another technique, which might be particularly interesting to customers doing prefix delegation to clients, is to use external address or prefix delegation assignments. This is done using the client reservation feature introduced in Cisco Network Registrar 7.2. Client reservations provides for assignment of reserved address assignments or prefix delegations. Client reservations are similar to lease reservations, but are not configured explicitly in the server as reservations. Instead, client reservations use information stored in Cisco Network Registrar client entries, LDAP client entries, or obtained from an external source. If an external source is used, an extension needs to be written to obtain the reserved addresses or prefixes and set the client reservation feature's reserved-ip6addresses or reserved-prefixes environment dictionary data items.

Multiple Techniques

Cisco Network Registrar's DHCPv6 server can actually use different techniques for different prefixes at the same time. This may be useful to optimize the assignment technique based on the client types or prefix type (stateful address versus prefix delegation). The techniques used are controlled by the prefix configuration, especially the allocation-algorithm attribute.

Longer Lease Times

One reason DHCPv4 lease times are configured to be relatively short is that a service provider has a limited number of addresses. However, the DHCPv6 addresses are not in short supply. Thus, the preferred and valid lifetimes for DHCPv6 leases can be set much longer.

Clients have between 50 percent of the preferred lifetime to 85 percent of the preferred lifetime in which to renew the address from the original server. Thus, for a 30-day lease, this means that between 15 days and 25.5 days the client will periodically try to renew its lease, and only after 10.5 days will it attempt to contact any available server through a Rebind message - which may then trigger the client to solicit a new lease.

The important point here though is that with longer lease times an operator has more time in which to restore a downed server before existing clients will attempt to contact another DHCPv6 server (10.5 days with 30-day preferred lease times). However, new clients will need to get leases from the other servers during this time.

Preferring One Server over Another

The DHCPv6 protocol was designed with a means to control which server a client chooses to use. This is done through the DHCPv6 preference option (see RFC 3315, sections 22.8 and 17.1.2).

Thus, one server can be configured with a high preference or the highest preference (255). Using a preference of 255 has the advantage that a client need not wait for additional server responses after receiving an advertisement with a preference of 255 (as normally a client waits for a short period to receive responses from all servers). The other server can be configured with a lower preference and thus will be less preferred.

When both servers respond to a client's Solicit message with an Advertise message, the client will prefer the higher preference server (or immediately use the 255 preference server).

Note that this preference only applies to the Solicit/Advertise messages phase. And, for Cisco Network Registrar, the preference is specified by adding an instance of the preference option on an appropriate policy. For example to set the default server wide preference:

```
nrcmd> policy system_default_policy setv6option preference 255
```

Issues with Independent Servers

Of course, there are some issues with using multiple independent servers:

- Domain Name System (DNS) updates will occur correctly and multiple addresses can be entered under a device's name. This generally isn't a problem, but may slow down connections to that device if one of those addresses is not in use by the device. This can happen if Server A provides the original lease, the device is shut down, Server A goes down, and the device is powered up and obtains a new address from Server B - the DNS will now contain both addresses, but only Server B's address is in use by the client.
- DNS updates may be prematurely removed if an external source assigns the address or delegated prefix or if interface-identifier-based address generation is used. This can happen in the just-mentioned situation, as when Server A returns, it may find that the lease has expired and therefore remove the DNS mapping for the client.
- Lease query requests for Cable Source Address verify or lease recovery by a relay agent (a cable modem termination system, or CMTS) after reboot can be more complex and troublesome. When one of the DHCPv6 servers is down, it cannot respond to lease query requests and thus the relay agent is unable to obtain information. And when both servers are up, one server may respond that the lease is leased to the client but the other indicates it is not.
- Lease information can be lost if a server's disk is lost or the lease database is irretrievably corrupted. DHCPv4 failover has been used to provide a "backup" of the lease data. For DHCPv6, this can result in the risk of duplicate address assignment if the disk or database is lost on one server. However, as Cisco Network Registrar generates addresses using a random number generator (that uses up to 62 bits of randomness), the probability is fairly small that a duplicate address will be generated.

These are key reasons why a DHCPv6 failover solution will have significant value.

DHCPv6 Failover Protocol

As of this writing, the Internet Engineering Task Force (IETF) Dynamic Host Configuration Working Group (DHC WG) has not begun work on a DHCPv6 failover protocol. However, it is anticipated that many of the concepts from the DHCPv4 failover protocol are applicable and usable for DHCPv6.

References

- RFC 3315 - Dynamic Host Configuration Protocol for IPv6 (DHCPv6)
- RFC 3633 - IPv6 Prefix Options for Dynamic Host Configuration Protocol (DHCP) version 6
- RFC 4291 - IP Version 6 Addressing Architecture
- RFC 4862 - IPv6 Stateless Address Autoconfiguration
- draft-ietf-dhc-failover-12.txt (expired) - DHCP Failover Protocol

Americas Headquarters Cisco Systems, Inc. San Jose, CA	Asia Pacific Headquarters Cisco Systems (USA) Pte. Ltd. Singapore	Europe Headquarters Cisco Systems International BV Amsterdam, The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**

Printed in USA C11-682438-00 09/11