Dynamic Host Configuration (DHC)
Internet-Draft
Intended status: BCP
Expires: May 3, 2012

J. Brzozowski
Comcast Cable Communications
J. Tremblay
Videotron Ltd.
J. Chen
Time Warner Cable
T. Mrugalski
ISC

October 31, 2011

DHCPv6 Redundancy Deployment Considerations
draft-ietf-dhc-dhcpv6-redundancy-consider-02

Abstract

This document documents some deployment considerations for those who wishing to use DHCPv6 to support their deployment of IPv6. Specifically, providing semi-redundant DHCPv6 services is discussed in this document.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on May 3, 2012.

Copyright Notice

Table of Contents

1. Introduction

To support the deployment of IPv6 redundancy and high availability are required for many if not all components. This document provides information specific to the proposed near term approach for deploying semi-redundant DHCPv6 services in advance of DHCPv6 server implementations that support a standards based failover or redundancy protocol.

2. Scope and Assumptions

This document specifies an interim architecture to provide a semi-redundant DHCPv6 solution before the availability of vendor or standard based solutions. The proposed architecture may be used in wide range of networks, two notable deployment models are discussed: service provider and enterprise network environments. The described architecture leverages only existing and implemented DHCPv6 standards. This document does not address a standards based solution for DHCPv6 redundancy. In the absence of a standards based DHCPv6 redundancy protocol and implementation, some analogies are loosely drawn with the DHCPv4 failover protocol for reference. Specific discussions related to DHCPv4 failover and redundancy is out of scope for this document. Reader interested in initial work being done in DHCPv6 failover is recommended to read [I-D.mrugalski-dhc-dhcpv6-failover-requirements].

Although DHCPv6 redundancy may be useful in a wide range of scenarios, they may be generalized for illustration purposes in the two aforementioned. The following assumptions were made with regards to the existing DHCPv6 infrastructure, regardless of the model used:

1. At least two DHCPv6 servers are used to service to the same clients, but the number of servers is not restricted.

2. Existing DHCPv6 servers will not directly communicate or interact with one another in the assignment of IPv6 addresses and configuration information to requesting clients.

3. DHCPv6 clients are instructed to run stateful DHCPv6 to request at least one IPv6 address. Configuration information and other options like a delegated IPv6 prefix may be also requested.

4. Clients requesting IPv6 addresses, prefixes, and or options care of DHCPv6 must recognize and honor the DHCPv6 preference option. Furthermore, the requesting clients must process DHCPv6 ADVERTISE messages per [RFC3315] when the preference option is present.

5. DHCPv6 server failure does not imply failure of any other network service or protocol, e.g. TFTP servers. Redundancy of any additional services configured by means of DHCPv6 are outside of scope of this document. For example, a single DHCPv6 server may configure multiple TFTP servers, with preference for each TFTP server, as specified in [RFC5970].

While techniques described in this document provide some aspects of redundancy, it should be noted that complete redundancy will not be available until DHCPv6 protocol is standardized. Initial work toward that goal is described in [I-D.mrugalski-dhc-dhcpv6-failover-requirements].

2.1. Service provider model

The service provider model represents cases, where end-user devices may be configured directly, without any intermediate devices (like home routers used in service provider model). DHCPv6 clients include cable modems, customer gateways or home routers, and end-user devices. In some cases hosts may be configured directly using the service provider DHCPv6 infrastructure or via intermediate router, that is in turn being configured by the provider DHCPv6 infrastructure. The service provider DHCPv6 infrastructure may be semi-redundant in either case. Cable modems, customer gateways or home routers, and end-user devices are commonly referred to as CPE (Customer Premises Equipment). The following additional assumptions were made, besides the ones made in Section 2:

1. The service provider edge routers and access routers (CMTS for cable or DSLAM/BRAS for DSL for example) are IPv6 enabled when required.

2. CPE devices are instructed to perform stateful DHCPv6 to request at least one IPv6 address, delegated prefix, and or configuration information. CPE devices may also be instructed to leverage stateless DHCPv6 [RFC3736] to acquire configuration information only. This assumes that IPv6 address and prefix information has been acquired using other means.

3. The primary application of this BCP is for native IPv6 services. Use and applicability to transition mechanisms is out of scope for this document.

4. CPE devices must implement a stateful DHCPv6 client [RFC3315], support for DHCPv6 prefix delegation [RFC3633] or stateless DHCPv6 [RFC3736] may also be implemented.

2.2. Enterprise model

The enterprise model represents cases where end-user devices are most often configured directly without any intermediate devices (like home routers used in service provider model). However enterprise IPv6 environments quite often use or require that DHCPv6 relay agents are in place to support the use of DHCPv6 for the acquisition of IPv6 addresses and or configuration information. The assumptions here extend those that are defined in the beginning of Section 2:

1. DHCPv6 clients are hosts and are considered end nodes. Examples of such clients include computers, laptops, and possibily mobile devices.

2. DHCPv6 clients generally do not require the assignment of an IPv6 prefix delegation and as such do not support DHCPv6 prefix delegation [RFC3633].

3. Protocol requirements

The following sections outline the requirements that must be satisfied by DHCPv6 clients, relays, and servers to ensure the desired behavior is provided using pre-existing DHCPv6 server implementations as is. The objective is to provide a semi-redundant DHCPv6 service to support the deployment of IPv6 where DHCPv6 is required for the assignment of IPv6 addresses, prefixes, and or configuration information.

3.1. DHCPv6 Servers

This interim architecture requires DHCPv6 servers that are [RFC3315] compliant and support the necessary options required to support this solution. Essential to the the use of the interim architecture is support for stateful DHCPv6 and the DHCPv6 preference option both which are specified in [RFC3315]. For deployment scenarios where IPv6 prefix delegation is employed DHCPv6 servers must support DHCPv6 prefix delegation as defined by [RFC3633]. Further, where stateless DHCPv6 is used support for [RFC3736] is required by DHCPv6 servers.

3.2. DHCPv6 Relays

There are no specific requirements regarding relays. However, it is implied that DHCPv6 relay agents must be [RFC3315] compliant and must support the ability to relay DHCPv6 messages to more than one destination minimally.

3.3. DHCPv6 Clients

DHCPv6 clients are required to be compliant to [RFC3315] and support the necessary options required to support this solution depending on the mode of operations and desired behavior. Where prefix delegation is required DHCPv6 clients will be required to support DHCPv6 prefix delegation as defined in [RFC3633]. Clients used with this semi-redundant DHCPv6 deployment model must support the acquistion of at least one IPv6 address and configuration information using stateful DHCPv6 as specified by [RFC3315]. The use of stateless DHCPv6 which is also specified in [RFC3315] may also be supported. DHCPv6 client must recognize and adhere to the processing of the advertised DHCPv6 preference options sent by the DHCPv6 servers.

4. Deployment models

At the time of this writing a standards-based DHCPv6 redundancy protocol and implementations are not available. As a result DHCPv6 server implementations will be used as-is to provide best effort, semi-redundant DHCPv6 services. Behavior of the DHCPv6 services will in part be governed by the configuration used by each of the servers. Additionally, various aspects of the DHCPv6 protocol [RFC3315] will be leveraged to yield the desired behavior. No inter-server or inter-process communications will be used to coordinate DHCPv6 events and or activities. DHCP services for both IPv4 and IPv6 may operate simultaneously on the same physical server(s) or may operate on different ones.

4.1. Split Prefixes

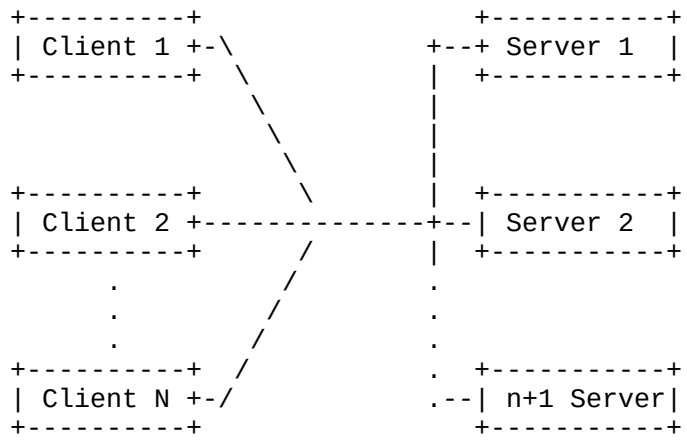
In the split prefixes model, each DHCPv6 server is configured with a unique, non-overlapping range derived from the /64 prefix deployed for use within an IPv6 network. Distribution between two servers, for example, would require that an allocated /64 be split in two /65 ranges. 2001:db8:1:0001:0000::/65 and 2001:db8:1:0001:8000::/65 would be assigned to each DHCPv6 server for allocation to clients derived from 2001:db8:1:0001::/64 prefix.

Each DHCP server allocates IPv6 addresses from the corresponding ranges per device class. Each DHCPv6 server will be simultaneously active and operational. Address allocation is governed largely through the use of the DHCPv6 preference option, so server with higher preference value is always prefered. Additional proprietary mechanisms can be leveraged to further enforce the favoring of one DHCP server over another. Example of such scenario is presented in Figure 1.

It is important to note that over time, it is possible that bindings may be disproportionally distributed amongst DHCPv6 servers and not any one server will be authoritative for all bindings.

Per [RFC3315], a DHCPv6 ADVERTISE messages with a preference option of 255 is an indicator to a DHCPv6 client to immediately begin a client-initiated message exchange by transmitting a REQUEST message. Alternatively, a DHCPv6 ADVERTISE messages with a preference option of any value lesser than 255 or absent preference option is an indicator to the client that it must wait for subsequent ADVERTISE messages before proceeding, as defined in Section 17.1.2 of [RFC3315]. Additionally, in the event of a DHCPv6 server failure it is desirable for a server other than the server that originally responded to be able to rebind the client. It is not critical, that the DHCPv6 server be able to rebind the client in this scenario, however, this is generally desirable behavior. Given the proposed architecture, the remaining active DHCPv6 server will have a different range configured making it technically incorrect for the same to rebind the client in its current state. Ultimately, when rebinding fails the client will acquire a new binding from the configured range unique to an active server. Furthermore, shorter T1, T2, valid, and preferred lifetimes can be used to reduce the possibility that a client or some other element on the network will experience a disruption in service or access to relevant binding data. The values used for T2, preferred and valid lifetime can be adjusted or configured to minimize service disruption. Ideally T2, preferred and valid lifetimes that are equal or near equal can be used to trigger a DHCPv6 client to reacquire IPv6 address, prefix, and or configuration information almost immediately after rebinding fails. It is important to note that shorter values will most certainly create additional load and processing for the DHCPv6 server, which must be considered.

Using a split prefix configuration model dynamic updates to DNS can be coordinated to ensure that the DNS is properly updated with current binding information. Challenges arise with regards to the update of PTR for IPv6 addresses since the DNS may need to be overwritten in a failure condition. The use of a split prefixes enables the differentiation of bindings and binding timing to determine which represents the current state. This becomes particularly important when DHCPv6 Leasequery [RFC5007] and/or DHCPv6 Bulk Leasequery [RFC5460] are leveraged to determine lease or binding state. An additional benefit is that the use of separate ranges per DHCPv6 server makes failure conditions more obvious and detectable.



Server 1
========
Prefix=2001:db8:1:0:0::/64
Range=2001:db8:1:0:0::/65
Preference=255

Server 2
========
Prefix=2001:db8:1:0:0::/64
Range=2001:db8:1:0:8000::/65
Preference=0

Server n+1
==========
Prefix, range, and preference would
vary based on range definition

Split prefixes approach.

Figure 1

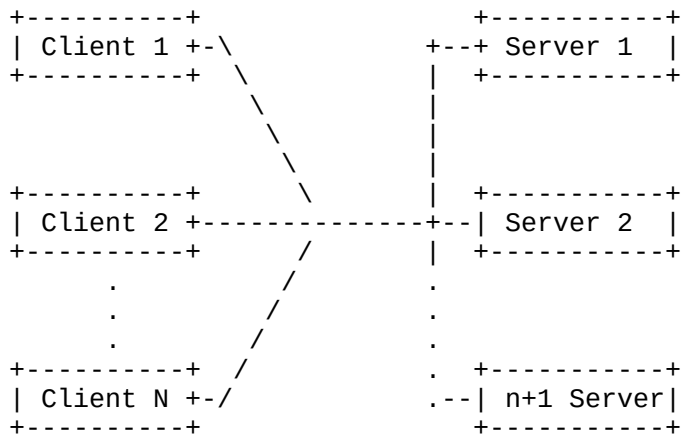
4.2. Multiple Unique Prefixes

In the multiple prefix model, each DHCPv6 server is configured with a unique, non-overlapping prefix. A /64 range equal to the prefix is configured on each server. For example, the range 2001:db8:1:0000::/64 would be assigned to a single DHCPv6 server for allocation to clients equal to its parent prefix 2001:db8:1:0000::/64. Subsequently the second DHCPv6 server could use 2001:db8:1:0001:::/64 as range and prefix. This would be repeated for each active DHCP server. Example of this scenario is presented in Figure 2.

This approach uses a unique prefix and ultimately range per DHCPv6 server with corresponding prefixes configured for use in the network. The corresponding network infrastructure must in turn be configured to use multiple prefixes on the inteface(s) facing the DHCPv6 client. The configuration is similar on all the servers, but a different prefix and a different preference is used per DHCPv6 server.

This approach would drastically increase the rate of consumption of IPv6 prefixes and would also yield operational and management challenges related to the underlying network since a significantly higher number of prefixes would need to be configured and routed. This approach also does not provide a clean migration path to the desired solution leveraging a standards-based DHCPv6 redundancy or failover protocol, which of course has yet to be specified.

The use of multiple unique prefixes provides benefits similar to those referred to in Section 4.1 related to dynamic updates to DNS. The use of multiple unique prefixes enables the differentiation of bindings and binding timing to determine which represents the current state. This becomes particularly important when DHCPv6 Leasequery [RFC5007] and/or DHCPv6 Bulk Leasequery [RFC5460] are leveraged to determine lease or binding state. The use of separate prefixes and ranges per DHCPv6 server makes failure conditions more obvious and detectable.




```
Server 1
========
Prefix=2001:db8:1:0000::/64
Range=2001:db8:1:0000::/64
Preference=255

Server 2
========
Prefix=2001:db8:1:1000::/64
Range=2001:db8:1:1000::/64
Preference=0

Server 3
========
Prefix=2001:db8:1:2000::/64
Range=2001:db8:1:2000::/64
Preference=(>0 and <255)
```

Multiple unique prefix approach.

Figure 2

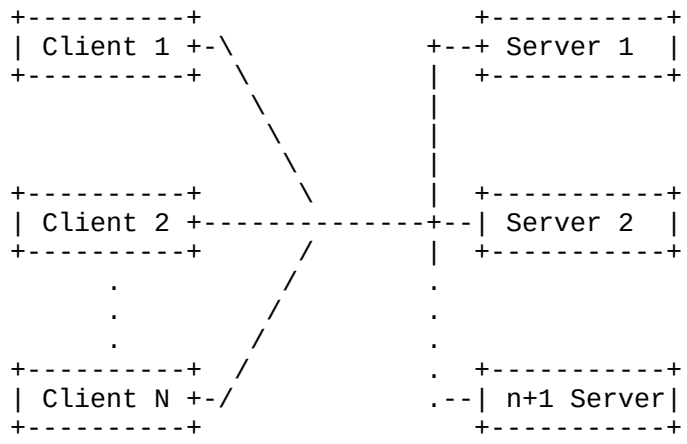
4.3. Identical Prefixes

In the identical prefix model, each DHCPv6 server is configured with the same overlapping prefix and range deployed for use within an IPv6 network. Distribution between two or more servers, for example, would require that the same /64 prefix and range be configured on all DHCP servers. For example, the range 2001:db8:1:0001:0000::/64 would be assigned to all DHCPv6 server for allocation to clients derived from 2001:db8:1:0001::/64 prefix. This would be repeated for each active DHCP server. Example of such scenario is presented in

Figure 3.

This approach uses the same prefix, length, and range definition across multiple DHCPv6 servers. All other configuration remaining the same the only other attribute of configuration option configured differently per DHCPv6 server would be DHCPv6 preference. This approach conceivably eases the migration of DHCPv6 services to fully support a standards based redundancy or failover protocol. Similar to the split prefix architecture described above this approach does not place any additional addressing requirements on network infrastructure.

The use of identical prefixes provides no benefit or advantage related to dynamic DNS updates, support of DHCPv6 Leasequery [RFC5007] or DHCPv6 Bulk Leasequery [RFC5460]. In this case all DHCP servers will use the same prefix and range configurations making it less obvious that a failure condition or event has occurred.




```
Server 1
========
Prefix=2001:db8:1:0000::/64
Range=2001:db8:1:0000::/64
Preference=255

Server 2
========
Prefix=2001:db8:1:0000::/64
Range=2001:db8:1:0000::/64
Preference=0

Server 3
========
Prefix=2001:db8:1:0000::/64
Range=2001:db8:1:0000::/64
Preference=(>0 and <255)
```

Identical prefix approach.

Figure 3

5. Challenges and Issues

The lack of interaction between DHCPv6 servers introduces a number of challenges related to the operations of the same in a production environment. The following areas are of particular concern:

- o In indentical prefixes scenario, both servers must follow the same address allocation procedure, i.e. they both must use the same algorithm and the same policy to determine which address is going

to be assigned to a specific client. Otherwise there is a distinct chance that each server will assign the same address to two different clients.

- Interactions with DNS server(s) to support the dynamic update of the same address when one or more DHCPv6 servers have become unavailable. This specifically becomes a challenge when or if nodes that were initially granted a lease:

  1. Attempt to renew or rebind the lease originally granted, or

  2. Attempt to obtain a new lease

  DHCID Resource Record, defined in [RFC4701], allows identification of the current owner for specific DNS data that can be used during DNS Update procedure [RFC2136]. [RFC4704] specifies how DHCPv6 servers and/or client may perform updates. [RFC4703] provides a way how to solve conflicts between clients. Although it deals with most cases, it is still possible to leave abandoned RR records. Consider following scenario. There are two independent servers. Server A assigns a lease to a client and updates DNS with AAAA record for assigned address and name. When the client renews, server A is not available and server B assigns a different lease. DNS is again updated (now two AAAA RRs are in the DNS for the client). Anyone trying to use the DNS information doesn't know which of the two leases is active. And, if server A never recovers, its information may never be removed.

- Interactions with DHCPv6 servers to facilitate the acquisition of IPv6 lease data care of the DHCPv6 Leasequery [RFC5007] or DHCPv6 Bulk Leasequery [RFC5460] protocols when one or more DHCPv6 servers have become unavailable and have granted leases to DHCPv6 clients. If IPv6 lease data is required and the granting server is unavailable it will not be possible to obtain any information about leases granted until one of the following has taken place.

  1. The granting DHCPv6 server becomes available with all lease information restored

  2. The client has renewed or rebound its lease against a different DHCPv6 server

  It is important to note that with DHCPv6 until such time that a redundancy or failover protocol is available binding updates and synchronization will not occur between DHCPv6 servers.

6. IANA Considerations

IANA is not requested to assign any numbers at this time.

7. Security Considerations

Security considerations specific to the operation of the DHCPv6 protocol are created through the use of this interim architecture for DHCPv6 redundancy beyond what has been cited for Dynamic Host Configuration Protocol for IPv6 (DHCPv6) [RFC3315]. There are considerations related to DNS, specifically the dynamic updating of DNS, when such models are employed. Potential opportunities are created to overwrite valid DNS resource records when provisions have been made accommodate some of the models cited in this document. In some cases this is desirable to ensure that DNS remains up to date when using one or more of these models, however, abuse of the same could result in undesirable behavior.

8. Acknowledgements

Many thanks to Bernie Volz, Kim Kinnear, Ralph Droms, David Hankins and Chuck Anderson for their input and review.

This work has been partially supported by Department of Computer Communications (a division of Gdansk University of Technology) and the Polish Ministry of Science and Higher Education under the European Regional Development Fund, Grant No. POIG.01.01.02-00-045/09-00 (Future Internet Engineering Project).

9. References

9.1. Normative References

[RFC2136] Vixie, P., Thomson, S., Rekhter, Y., and J. Bound, "Dynamic Updates in the Domain Name System (DNS UPDATE)", RFC 2136, April 1997.

[RFC3315] Droms, R., Bound, J., Volz, B., Lemon, T., Perkins, C., and M. Carney, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3315, July 2003.

[RFC3633] Troan, O. and R. Droms, "IPv6 Prefix Options for Dynamic Host Configuration Protocol (DHCP) version 6", RFC 3633, December 2003.

[RFC3736] Droms, R., "Stateless Dynamic Host Configuration Protocol (DHCP) Service for IPv6", RFC 3736, April 2004.

[RFC4701] Stapp, M., Lemon, T., and A. Gustafsson, "A DNS Resource Record (RR) for Encoding Dynamic Host Configuration Protocol (DHCP) Information (DHCID RR)", RFC 4701, October 2006.

[RFC4703] Stapp, M. and B. Volz, "Resolution of Fully Qualified Domain Name (FQDN) Conflicts among Dynamic Host Configuration Protocol (DHCP) Clients", RFC 4703, October 2006.

[RFC4704] Volz, B., "The Dynamic Host Configuration Protocol for IPv6 (DHCPv6) Client Fully Qualified Domain Name (FQDN) Option", RFC 4704, October 2006.

[RFC5007] Brzozowski, J., Kinnear, K., Volz, B., and S. Zeng, "DHCPv6 Leasequery", RFC 5007, September 2007.

[RFC5460] Stapp, M., "DHCPv6 Bulk Leasequery", RFC 5460, February 2009.

[RFC5970] Huth, T., Freimann, J., Zimmer, V., and D. Thaler, "DHCPv6 Options for Network Boot", RFC 5970, September 2010.

9.2. Informative References

[I-D.mrugalski-dhc-dhcpv6-failover-requirements] Mrugalski, T. and K. Kinnear, "DHCPv6 Failover Requirements", draft-mrugalski-dhc-dhcpv6-failover-requirements-00 (work in progress), June 2011.

Authors' Addresses

John Jason Brzozowski
Comcast Cable Communications
1306 Goshen Parkway
West Chester, PA 19380
USA

Phone: +1-609-377-6594
Email: john_brzozowski@cable.comcast.com

Jean-Francois Tremblay
Videotron Ltd.
612 Saint-Jacques
Montreal, Quebec H3C 4M8
Canada

Email: jf@jftremblay.com

Jack Chen
Time Warner Cable
13820 Sunrise Valley Drive
Herndon, VA 20171
USA

Email: jack.chen@twcable.com

Tomasz Mrugalski
Internet Systems Consortium, Inc.
950 Charter St.
Redwood City, CA 94063
USA

Phone: +1 650 423 1345
Email: tomasz.mrugalski@gmail.com