DNA Working Group
Internet-Draft
Expires: December 27, 2006

J. Kempf
DoCoMo Communications Labs USA
S. Narayanan
Panasonic
E. Nordmark
Sun Microsystems
B. Pentland, Ed.
Monash University CTIE
JH. Choi
Samsung AIT
June 25, 2006

Detecting Network Attachment in IPv6 Networks (DNAv6)
draft-ietf-dna-protocol-01.txt

Status of this Memo

Copyright Notice

Abstract

Efficient detection of network attachment in IPv6 needs the following

two components: a method for the host to query routers on the link to identify the link (Link Identification) and a method for the routers on the link to consistently respond to such a query with minimal delay (Fast RA). Solving the link identification based strictly on RFC 2461 is difficult because of the flexibility offered to routers in terms of prefixes advertised in a router advertisement (RA) message. Similarly, the random delay in responding to router solicitation messages imposed by RFC 2461 makes to it difficult to receive an RA quickly. In this memo, an integrated solution is presented. Monitoring of prefixes by both hosts and routers is used to achieve link identification while router advertisements are sent rapidly in a deterministic order by combining router solicitation source addresses with hash-based router tokens.

Table of Contents

1. Introduction 5

2. Terms and Abbreviations 5

3. Overview 5
3.1 Link Identification 5
3.2 Fast Router Advertisement 7

4. Message Formats 8
4.1 Router Advertisement 8
4.2 Prefix Information Option LinkID Bit 9
4.3 Landmark Option 10
4.4 Learned Prefix Option 12

5. DNA Operation 13
5.1 DNA Router Operation 13
5.1.1 Data Structures 14
5.1.2 Router Configuration Variables 15
5.1.3 Bootstrapping DNA Data Structures 16
5.1.4 Processing Router Advertisements 16
5.1.5 Processing Router Solicitations 17
5.1.6 Complete Router Advertisements 18
5.1.7 LinkID 18
5.1.8 Scheduling Fast Router Advertisements 19
5.1.9 Scheduling Unsolicited Router Advertisements 20
5.1.10 Removing a Prefix from an Interface 20
5.1.11 Prefix Reassignment 21
5.2 DNA Host Operation 21
5.2.1 Data Structures 21
5.2.2 Host Configuration Variables 22
5.2.3 Selection of a Landmark Prefix 22
5.2.4 Sending Router Solicitations 23
5.2.5 Processing Router Advertisements 23
5.2.6 DNA and Address Configuration 25

6. Backward Compatibility 29
6.1 Non-DNA Host with DNA Routers 29
6.2 DNA Host with Non-DNA Routers 29

7. Security Considerations 29
7.1 Attacks on the Token Bucket 29
7.2 Attacks on DNA Hosts 29

8. IANA Considerations 30

9. Acknowledgments 30

10. References 31
10.1 Normative References 31
10.2 Informative References 31

Authors' Addresses 32

A. How the Goals are Met? 33

Intellectual Property and Copyright Statements 35

1. Introduction

The proposed scheme in this memo is built upon the following solutions catalogued in [16]: Complete RA is used for the link identification, and Hash-based Fast RA is used to achieve fast response to RS messages. Aspects of prefix-based LinkID and Requested Landmark are included to allow for a decrease in the packet sizes associated with Complete RA.

The rest of the document refers to this approach by the term "DNAv6".

2. Terms and Abbreviations

There is an existing DNA terminology draft [13]. This draft does not introduce any new terminology not already used by existing drafts.

The term "link" is used as defined in RFC 2460 [2]. NOTE: this is completely different from the term "link" as used by IEEE 802, etc.

3. Overview

The DNA protocol presented in this document tries to achieve the following objectives:

- Eliminate the delays introduced by RFC 2461 in discovering the configuration.

- Make it possible for the hosts to accurately detect the identity of their current link from a single RA.

DNAv6 assumes that the host's wireless link interface software and hardware is capable of delivering a 'link up' event notification when layer 2 on the host is configured and sufficiently stable for IP traffic. This event notification acts as a hint to the layer 3 DNA procedures to check whether or not the host is attached to the same link as before. DNAv6 also assumes that an interface on the host is never connected to two links at the same time. In the case that the layer 2 technology is capable of having multiple attachments (for instance, multiple layer 2 associations or connections) at the same time, DNAv6 requires the individual layer-2 associations to be represented as separate (virtual interfaces) to layer 3 and DNAv6 in particular.

3.1 Link Identification

DNAv6 identifies a link by the set of prefixes that are assigned to the link, which is quite natural and doesn't require introducing any new form of identifier. However, this choice implies that the

protocol needs to be robust against changes in the set of prefixes assigned to a link, including the case when a link is renumbered and the prefix is later reassigned to a different link. The protocol handles this during graceful renumbering (when the valid lifetime of the prefix is allowed to decrease to zero before it is removed and perhaps reassigned to a different link), it describes how to remove and reassign prefixes earlier than this without any incorrect behaviour, and will also recover in case where a prefix is reassigned without following the draft recommendations.

DNAv6 is based on using a Router Solicitation/Router Advertisement exchange to both verify whether the host has changed link, and if it has, provide the host with the configuration information for the new link. The base method for detecting link change involves getting routers to listen to all of the prefixes that are being advertised by other routers on the link. They can then respond to solicitations with complete prefix information. This information consists of the prefixes a router would advertise itself as per RFC 2461, and also, the prefixes learned from other routers on the link that are not being advertised by itself. These learned prefixes are included in a new Learned Prefix Option in the Router Advertisement.

A host receiving one of these "Complete RAs" - so marked by a flag - then knows all of the prefixes in use on a link, and by inference all those that are not. By comparing this with previously received prefixes the host can correctly decide whether it is connected to the same link as previously, or whether this Router Advertisement is from a new link. Unlike CPL [15], the host does not have to wait for multiple advertisements before making a decision.

Though frequently all routers on a link will advertise the same set of prefixes and thus experience no cost in making the RAs complete, there is potential for the RAs to be large when there are many prefixes advertised. Two mechanisms are defined that allow certain RAs to be reduced in size.

One uses a technique called a "landmark", where the host chooses one of the prefixes as a landmark prefix, and then includes this in the Router Solicitation message in the form of a question "am I still connected to the link which has this prefix?". The landmark is carried in a new option, called the Landmark Option.

In the case when the host is still attached to the same link, which might occur when the host has changed from using one layer 2 access point to another, but the access points are on the same link, the Router Advertisement(s) it receives will contain a "yes, that prefix is on this link" answer, and no other information. Thus, such RA messages are quite small.

In the case when the landmark prefix is unknown to the responding router, the host will receive a "No" answer to its landmark question, and also the information it needs to configure itself for the new link. The routers try to include as much information as possible in such messages, so that the host can be informed of all the prefixes assigned to the new link as soon as possible.

A second mechanism for reducing packet sizes applies to unsolicited Router Advertisements. By selecting one prefix on the link to be the "link identifier", and making sure that it is included in every advertisement, it is possible to omit some prefixes. Such advertisements will not inform a host of all of the prefixes at once, but in general these unsolicited advertisements will not be the first advertisement received on a link. Inclusion of the link identifier prefix simply ensures that there is overlap between the sets of prefixes advertised by each router on a link and that hosts will thus not incorrectly interpret one of these incomplete RAs as an indication of movement.

The Router Advertisement messages are, in general, larger than the solicitations, and with multiple routers on the link there will be multiple advertisements sent for each solicitation. This amplification can be used by an attacker to cause a Denial of Service attack. Such attacks are limited by applying a rate limit on the unicast Router Advertisements sent directly in response to each solicitation, and using multicast RAs when the rate limit is exceeded.

In order for the routers be able to both respond to the landmark questions and send the complete RAs, the routers need to track the prefixes that other routers advertise on the link. This process is initialized when a router is enabled, by sending a Router Solicitation and collecting the resulting RAs, and then multicasting a few RAs more rapidly as already suggested in RFC 2461. This process ensures with high probability that all the routers have the same notion of the set of prefixes assigned to the link.

3.2 Fast Router Advertisement

According to RFC 2461 a solicited Router Advertisement should have a random delay between 0 and 500 milliseconds, to avoid the advertisements from all the routers colliding on the link causing congestion and higher probability of packet loss. In addition, RFC 2461 suggests that the RAs be multicast, and multicast RAs are rate limited to one message every 3 seconds. This implies that the response to a RS might be delayed up to 3.5 seconds.

DNAv6 avoids this delay by using a different mechanism to ensure that

two routers will not respond at exactly the same time while allowing one of the routers on the link to respond immediately. Since the hosts might be likely to use the first responding router as the first choice from their default router list, the mechanism also ensures that the same router doesn't respond first to the RSs from different hosts.

The mechanism is based on the routers on the link determining (from the same RAs that are used in section Section 3.1 to determine all the prefixes assigned to the link), the link-local addresses of all the other routers on the link. With this loosely consistent list, each router can independently compute some function of the (link-local) source address of the RS and each of the routers' link-local addresses. The results of that function are then compared to create a ranking, and the ranking determines the delay each router will use when responding to the RS. The router which is ranked as #0 will respond with a zero delay.

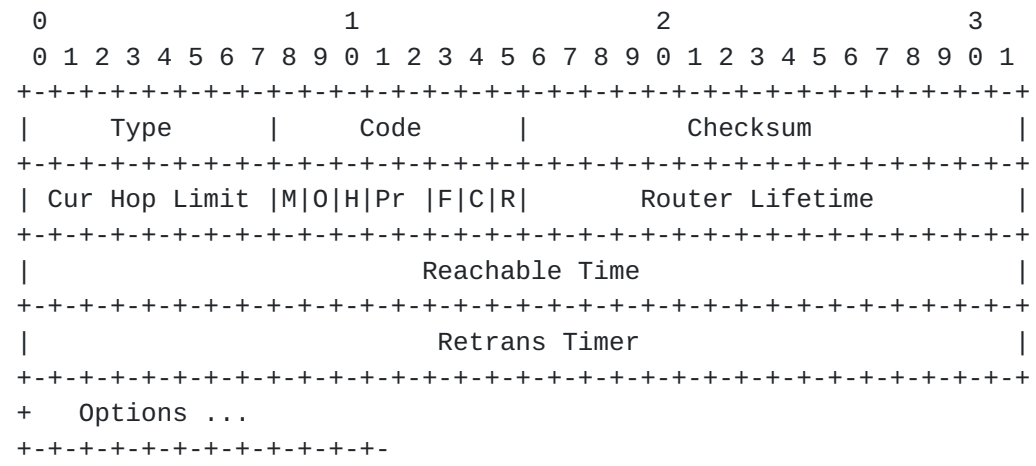
If the routers become out-of-sync with respect to their learned router lists, two or more routers may respond with the same delay, but over time the routers will converge on their lists of learned routers on the link.

4. Message Formats

This memo defines two new flags for inclusion in the router advertisement message and two new options.

4.1 Router Advertisement

DNAv6 modifies the format of the Router Advertisement message by defining a bit to indicate that the router sending the message is participating in the DNAv6 protocol as well as a flag to indicate the completeness of the set of prefixes included in the Router Advertisement. The new message format is as follows:



FastRA (F)

The FastRA (F) bit indicates that the router sending the RA is participating in the DNAv6 protocol. Other routers should include this router in calculating response delay tokens.

Complete (C)

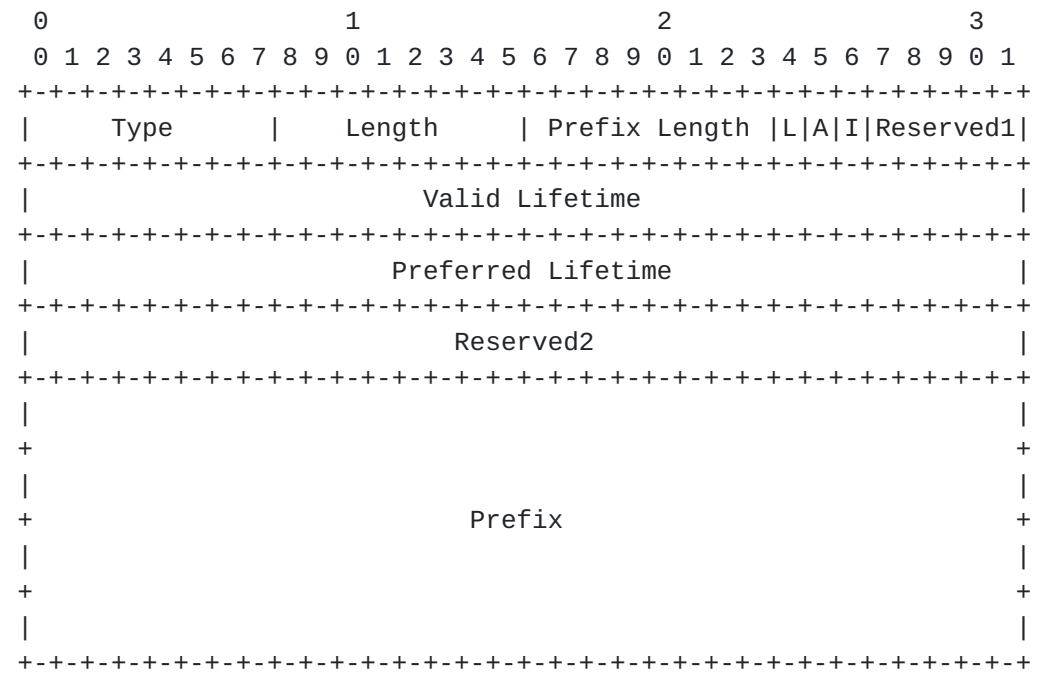
The Complete (C) bit indicates that the Router Advertisement contains PIOs for all prefixes explicitly configured on the sending router, and, if other routers on the link are advertising additional prefixes, a Learned Prefix Option containing all additional prefixes that the router has heard from other routers on the link.

Reserved (R)

The reserved field is reduced from 3 bits to 1 bit.

4.2 Prefix Information Option LinkID Bit

DNAv6 modifies the format of the Prefix Information Option by defining a bit to indicate that the enclosed prefix is currently being used as the Link Identifier. The new message format is as follows:



LinkID (I)

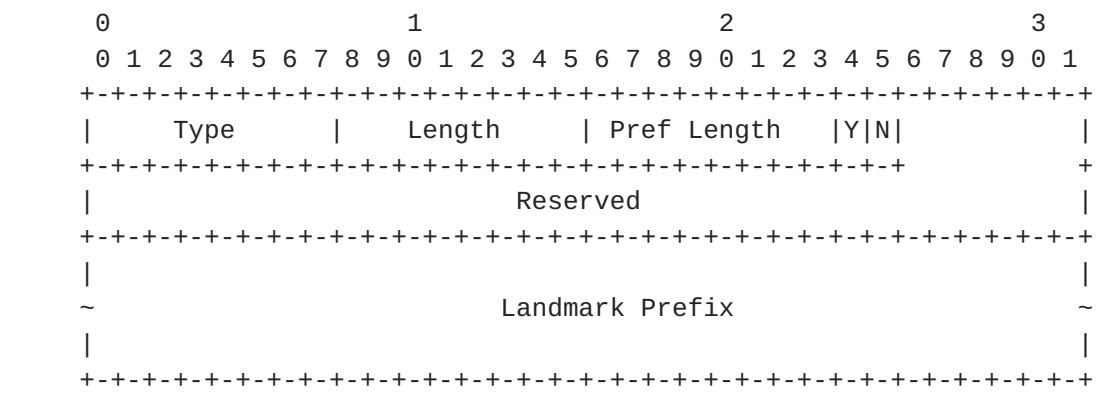
The LinkID (I) bit indicates that the prefix in the Prefix field of this option is currently being used as the Link Identfier (LinkID).

Reserved1

The Reserved1 field is reduced from 6 bits to 5 bits.

4.3 Landmark Option

The Landmark Option is used by hosts in a Router Solicitation message to ask the routers on a link if the specified prefix is being advertised by some router on the link. It is used by routers in a Router Advertisement to reply to a corresponding question in a Router Solicitation, indicating whether the prefix referred to is being advertised by any router on the link.



Type

TBA

Length

8 bit unsigned integer indicating the length of the option in units of 8 octets. Set to 2 or 3.

Pref Length

An 8 bit unsigned integer representing the number of bits in the prefix to be used for matching.

Yes (Y)

The Yes (Y) bit, when included in a Landmark Option in a Router Advertisement, indicates that the prefix referred to in the Prefix field of this option is being advertised by one or more routers on the current link. In a Landmark Option in a Router Solicitation, this bit MUST be set to zero and ignored by receivers.

No (N)

The No (N) bit, when included in a Landmark Option in a Router Advertisement, indicates that the prefix referred to in the Prefix field of this option is not being advertised by any router on the current link. In a Landmark Option in a Router Solicitation, this bit MUST be set to zero and ignored by receivers.

Reserved

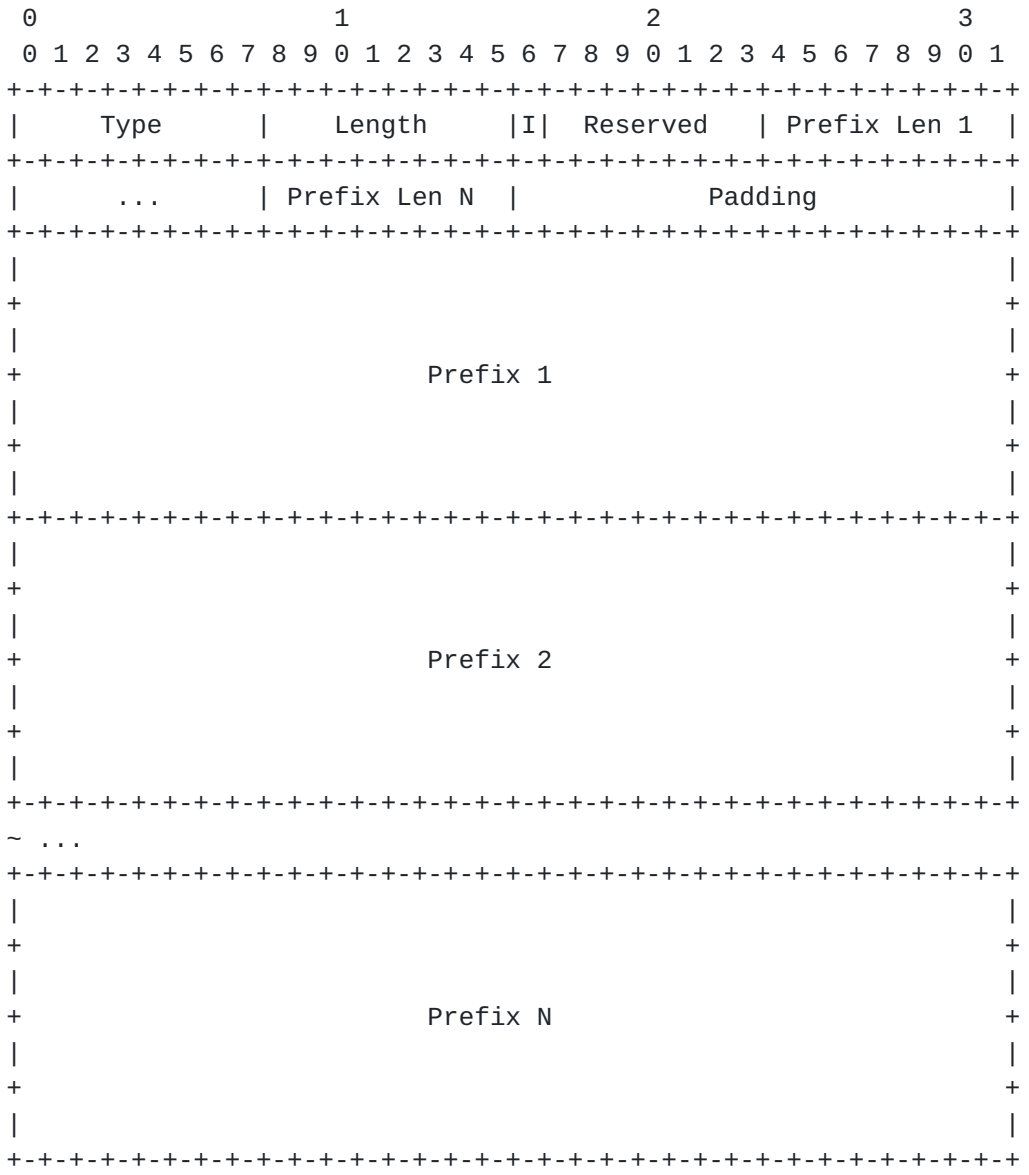
A 38 bit unused field. It MUST be initialised to zero by the sender, and ignored by the receiver.

Prefix

A prefix being used by the host currently for a global IPv6 address, padded at the right with zeros. If the prefix length is less than 65 bits, only 64 bits need be included, otherwise 128 bits are included.

4.4 Learned Prefix Option

The Learned Prefix Option (LPO) is used by a router to indicate prefixes that are being advertised in PIOs by other routers on the link, but not by itself.



Type

TBA

Length

8 bit unsigned integer indicating the length of the option in units of 8 octets.

I

LinkID (I) flag. When set indicates that the first prefix in this option is the LinkID for this link.

Prefix Len

One or more fields (N) each consisting of an 8-bit unsigned integer representing the prefix lengths of the following prefixes. The Prefix Len fields are ordered the same as the Prefix fields so that the first Prefix Len field represents the prefix length of the prefix contained in the first prefix field, and so on.

Padding

Zero padding sufficient to align the following prefix field on an 8-octet boundary.

Prefix

One or more fields (N) each containing a 128-bit address representing a prefix that has been heard on the link but is not explicitly configured on this router.

Description

This option MUST only be included in a Router Advertisement. This option contains prefixes that are being advertised on the link but are not explicitly configured on the sending router. The router MUST NOT include any prefixes with a zero valid lifetime in the LPO.

5. DNA Operation

5.1 DNA Router Operation

Routers MUST collect information about the other routers that are advertising on the link.

5.1.1 Data Structures

The routers maintain a set of conceptual data structures for each interface to track the prefixes advertised by other routers on the link, and also the set of DNA routers (the routers that will quickly respond to RSs) on the link.

For each interface, routers maintain a list of all prefixes learned from other routers on the link but not explicitly configured on the router's own interface. The list will be referred to in this document as "DNARouterPrefixList". Prefixes are learned by their reception within Prefix Information Options [3] in Router Advertisements. Prefixes in Learned Prefix Options (see Section 4.4) MUST NOT update the contents of DNARouterPrefixList. For each prefix the router MUST store sufficient information to identify the prefix and to know when to remove the prefix entry from the list. This may be achieved by storing the following information:

1. Prefix

2. Prefix length

3. Prefix valid lifetime

4. Expiry time

The expiry time for entries in DNARouterPrefixList is 1.5 hours (three times the maximum value of the Router Advertisement interval) after the last received Router Advertisement affecting the entry, or the scheduled expiry of the prefix valid lifetime, whichever is earlier.

For each interface, routers also maintain a list of the other routers advertising on the link. The list will be referred to in this memo as "DNARouterList". For each router from which a Router Advertisement is received with the FastRA flag set, the following information MUST be stored:

1. Link-local source address of advertising router

2. Token equal to the first 64 bits of an SHA-1 hash of the above address

3. Expiry time

Each router MUST include itself in the DNARouterList and generate a token for itself as described above based on the link-local address used in its RA messages.

The expiry time for entries in DNARouterList is 1.5 hours after the last received Router Advertisement affecting the entry.

5.1.2 Router Configuration Variables

A DNAv6 router MUST allow for the following conceptual variables to be configured by the system management. Default values are set to ease configuration load.

UnicastRAInterval

The interval corresponding to the maximum average rate of Router Solicitations that the router is prepared to service with unicast responses. This is the interval at which the token bucket controlling the unicast responses is replenished.

Default: 50 milliseconds

MaxUnicastRABurst

The maximum size burst of Router Solicitations that the router is prepared to service with unicast responses. This is the maximum number of tokens allowed in the token bucket controlling the unicast responses.

Default: 20

RASeparation

The separation between responses from different routers on the same link to a single Router Solicitation.

Default: 20 milliseconds

MulticastRADelay

The delay to be introduced when scheduling a multicast RA in response to a RS message when the token bucket is empty.

Default: 3000 milliseconds

FastRAThreshold

The maximum number of fast responses that a host should receive when soliciting for Router Advertisements.

Default: 3

5.1.3 Bootstrapping DNA Data Structures

When an interface on a router first starts up, it SHOULD transmit up to MAX_RTR_SOLICITATIONS Router Solicitations separated by RTR_SOLICITATION_INTERVAL [3] in order to quickly learn of the other routers and prefixes active on the link.

Upon startup, a router interface SHOULD also send a few unsolicited Router Advertisements as recommended in Section 6.2.4 of RFC 2461 [3], in order to inform others routers on the link of its presence.

During the bootstrap period ((MAX_RTR_SOLICITATIONS - 1) * RTR_SOLICITATION_INTERVAL + RetransTimer [3]), a router interface both sends unsolicited Router Advertisements and responds to Router Solicitations, but with a few restrictions on the message content. Router Advertisements MUST NOT include any DNA specific options except that the FastRA flag MUST be set. The FastRA flag is set so that other routers will know to include this router in their timing calculations for fast RA transmission. Other DNA options are omitted because the router may have incomplete information during bootstrap.

During the bootstrap period, the Complete flag in Router Advertisements MUST NOT be set.

During the bootstrap period, the timing of Router Advertisement transmission is as specified in RFC 2461.

5.1.4 Processing Router Advertisements

When a router receives a Router Advertisement, it first validates the RA as per the rules in RFC 2461, and then performs the actions specified in RFC 2461. In addition, each valid Router Advertisement is processed as follows:

If the FastRA flag is set in the RA, the router checks if there is an entry in its DNARouterList. Thus it looks up the source address of the RA in that list and, if not found, a new entry is added to DNARouterList, including the source address and a token equal to the first 64 bits of an SHA-1 hash of the source address. The entry's expiry time is updated.

Regardless of the state of the FastRA flag, each PIO in the RA is examined. If the prefix is not in the router's DNARouterPrefixList and not in the router's AdvPrefixList [3], it is added to the DNARouterPrefixList, and its expiry time is set.

5.1.5 Processing Router Solicitations

The usual response to a Router Solicitation SHOULD be a unicast RA. However, to keep control of the rate of unicast RAs sent, a token bucket is used. The token bucket is filled at one token every UnicastRAInterval. A maximum of MaxUnicastRABurst tokens are stored.

When a Router Solicitation is received, the router checks if it is possible to send a unicast response. A unicast response requires that the following conditions to be met:

- A unicast send token is available.
- The source address of the Router Solicitation is NOT the unspecified address (::).

If a unicast response is possible and the Router Solicitation contains a Landmark Option whose prefix is included in DNARouterPrefixList or AdvPrefixList, the router SHOULD send an abbreviated Router Advertisement.

This abbreviated advertisement includes only the Landmark Option, with the "Y" flag set, plus the base RA header and any SEND options as appropriate. The FastRA flag MUST be set. The Complete flag MUST NOT be set. This is the one exception where the LinkID MAY be omitted as the Y flag implies that link change has not occured and that the previously received LinkID is still current.

If there is NO Landmark Option in the received Router Solicitation or it contains a Landmark Option whose prefix is NOT included in DNARouterPrefixList or AdvPrefixList or a unicast response is not possible, then the router SHOULD generate a Complete RA as specified in Section 5.1.6. The Router Advertisement MUST include the LinkID, as described in Section 5.1.7.

If a unicast response is possible, then a token is removed and the Router Advertisement is scheduled for transmission as specified in Section 5.1.8.

If a unicast response is not possible and there is no multicast RA already scheduled for transmission in the next MulticastRADelay the RA MUST be sent to the link-scoped all-nodes multicast address at the current time plus MulticastRADelay.

If a unicast response is not possible but there is a multicast RA already scheduled for transmission in the next MulticastRADelay, then the Router Solicitation MUST be silently discarded.

5.1.6 Complete Router Advertisements

A CompleteRA is formed as follows:

Starting with a Router Advertisement with all fixed options (MTU, Advertisement Interval, flags, etc.), the FastRA flag is set. As many Prefix Information Options for explicitly configured prefixes as will fit are added to the Router Advertisement. If there is sufficient room, a Learned Prefix Option as defined in Section 4.4 containing as many of the learned prefixes as will fit is added.

It may not be possible to include all of the prefixes in use on the link due to MTU or administrative limitations. If all Prefix Information Options and a Learned Prefix Option containing all of the learned prefixes were included in the RA, then the Complete flag in the Router Advertisement header is set.

If it is not possible to generate a Complete RA but the Router Solicitation that this Router Advertisement is in response to, if any, includes a Landmark Option containing a prefix that is not in the router's DNARouterPrefixList and not in the router's AdvPrefixList then the router SHOULD include a Landmark Option with the "N" flag set. If there are known to be prefixes that are not included in the Router Advertisement, then the Complete flag MUST NOT be set.

Note that although it may not be possible to fit all of the prefixes into an RA, the LinkID MUST be included.

5.1.7 LinkID

One of the prefixes in use on a link is chosen to be the LinkID.

The LinkID is the numerically smallest prefix stored in either of DNARouterPrefixList or AdvPrefixList whose lifetime is greater than 1.5 hours. For comparing prefixes, they are padded to the right with zeros to make them 128 bit unsigned integers.

The prefix may be included in the RA in either a PIO or LPO as appropriate. If the prefix is included in a PIO, then the "I" flag in that PIO MUST be set. If the prefix is included in an LPO, then the prefix MUST be placed in the first prefix field in the LPO, and the LPO "I" flag MUST be set.

5.1.7.1 Changing the LinkID

When either a new prefix is added to a link that is numerically smaller than all those previously advertised or the lifetime of the

prefix that is currently being used as the LinkID falls below 1.5 hours, a new LinkID is determined. In order to ensure that there is overlap between consecutive RAs on the link, the old LinkID must continue to be advertised for some time alongside the new LinkID.

For the purposes of propagating information, it is assumed that after three advertisements of a change, all routers have been made aware of the change.

If the instant that a router sends its first unsolicited advertisement is time T, then by T + 1 hour at least three such advertisements will have been made and all routers can be assumed to have received it. Thus by time T + 1.5 hours, all routers on the link should have also sent at least one advertisement with the new LinkID.

1.5 hours after first sending an advertisement with a new LinkID it is safe to consider the old LinkID gone and omit the corresponding prefix from RAs if desired.

Following a change of LinkID, the old LinkID MUST be included in RAs for the following 1.5 hours.

5.1.7.1.1 Non-Prefix LinkIDs

Although this memo only discusses LinkIDs that are prefixes, a future specification or ammendment may describe a mechanism to select a LinkID that is not a prefix.

Information from the Learned Prefix Option is only stored in DNAHostPrefixList, and is only used for DNA purposes. Because a length field is used, it is possible to carry any variable length identifier less than or equal to 128 bits in an LPO and store it in DNAHostPrefixList (Section 5.2.1).

Following a change of LinkID, the old LinkID MUST be included in RAs in an LPO for the following 1.5 hours.

Future specifications MUST NOT treat the information in an LPO as prefixes such as they would the prefixes found in a Prefix Information Option. Future specifications MUST NOT assume that the entries in a host's DNAHostPrefixList are actaul prefixes in use on the link.

5.1.8 Scheduling Fast Router Advertisements

RAs may need to be delayed to avoid collisions in the case that there is more than one router on a link. The delay is calculated by

determining a ranking for the router for the received RS, and multiplying that by RASeparation.

A Host Token is needed from the RS to calculate the router's ranking. The first 64 bits of an SHA-1 hash of the source address of the RS MUST be used as the RS host token.

A router's ranking is determined by taking the XOR of the RS Host Token and each of the stored Router Tokens. The results of these XOR operations are sorted lowest to highest. The router corresponding to the first entry in the sorted list is ranked zero, the second, one, and so on.

> Note: it is not necessary for a router to actually sort the whole list. Each router just needs to determine its own position in the sorted list.

If Rank < FastRAThreshold, then the RA MUST be scheduled for transmission in Rank * RASeparation milliseconds. When the router is ranked as zero, the resulting delay is zero, thus the RA SHOULD be sent immediately.

If Rank >= FastRAThreshold, then the RA MUST be replaced with a Complete RA, if it is not one already, and scheduled for multicast transmission as in RFC 2461.

5.1.9 Scheduling Unsolicited Router Advertisements

Unsolicited router advertisements MUST be scheduled as per RFC 2461.

The "F" flag in the RA header MUST be set.

They MAY be Complete RAs or MAY include only a subset of the configured prefixes, but MUST include the LinkID.

This ensures that there will be overlap in the sets of prefixes contained in consecutive RAs on a link from DNA routers, and thus an absence of that overlap can be used to infer link change.

5.1.10 Removing a Prefix from an Interface

When a prefix is to stop being advertised in a PIO in RAs by an interface before the expiry of the prefix's valid lifetime, then the router should treat it as though it has just learned a prefix that is not explicitly configured on it. After sending the last RA containing the prefix in a PIO, the router MUST add the prefix to the DNARouterPrefixList and set it to expire in 1.5 hours or at the expiry of the last advertised valid lifetime, whichever is earlier.

This ensures that to hosts there will be overlap in the prefixes in the RAs they see and prevent them from incorrectly interpreting changed prefixes as movement.

5.1.10.1 Early Removal of the LinkID Prefix

If the LinkID prefix is to be withdrawn early from a link, that is before the expiry of its previously advertised valid lifetime, it MUST be advertised for at least 1.5 hours with a valid lifetime of less than 1.5 hours. This ensures that all of the other routers are notified to begin the process of changing the LinkID as well, and hosts will always see overlap between the prefixes in consecutive RAs and thus not mistake an RA for an indication of link change.

5.1.11 Prefix Reassignment

A prefix whose lifetime has expired after counting down in real time for at least 1.5 hours may be reassigned to another link immediately after expiry. If a prefix is withdrawn from a link without counting down to the expiry of its valid lifetime, it SHOULD NOT be reassigned to another link for at least 1.5 hours or until the original expiry time, whichever is earlier. This gives sufficient time for other routers that have learned the prefix to expire it, and for hosts that have seen advertisements from those routers to expire the prefix as well.

Earlier reassignment may result in hosts that move from between the old and new links failing to detect the movement.

5.2 DNA Host Operation

Hosts collect information about the prefixes available on the link to which they are connected to facilitate change detection.

5.2.1 Data Structures

Hosts MUST maintain a list of prefixes advertised on the link. This is separate from the RFC 2461 "Prefix List" and will be referred to here as the "DNAHostPrefixList". All prefixes SHOULD be stored, however an upper bound MUST be placed on the number stored to prevent overflow. For each prefix stored the host MUST store the following information:

1. Prefix

2. Prefix length

3. Expiry time

If a host is not able to store this information for every prefix, there is a risk that the host will incorrectly decide that it has moved to a new link, when it receives advertisements from a non-DNA router.

Prefix entries in the DNAHostPrefixList expire and MUST be removed 1.5 hours after they are last seen in a received Router Advertisement (in either a PIO or LPO) or at the expiry of the valid lifetime of the prefix, whichever is earlier.

Hosts MUST also maintain a list of all LinkIDs seen on the current Link. This list will be referred to as "DNAHostLinkIDList". This list is identical in structure to DNAHostPrefixList but contains LinkIDs instead of prefixes.

At this time LinkIDs are also prefixes but in future may be able to be identifiers other than prefixes. A list is stored rather than a single entry to allow for changes in the LinkID used on a link.

Entries are expired from DNAHostLinkIDList in the same way as DNAHostPrefixList.

Hosts SHOULD also maintain a "Landmark Prefix" as described in Section 5.2.3.

5.2.2 Host Configuration Variables

Hosts MUST make use of the following conceptual variables and they SHOULD be configurable:

DNASameLinkDADFlag

Boolean value indicating whether or not a host should re-run DAD when DNA indicates that link change has not occurred.

Default: False

5.2.3 Selection of a Landmark Prefix

For each interface, hosts SHOULD choose a prefix to use as a Landmark Prefix in Router Solicitations. The following rules are used in selecting the landmark prefix:

The prefix MUST have a non-zero valid lifetime. If the valid lifetime of a previously selected Landmark Prefix expires, a new

Landmark Prefix MUST be selected.

The prefix MUST be one of those that the hosts has used to assign a non-link-local address to itself

The prefix SHOULD be chosen as the one with the longest preferred lifetime, but it is not necessary to switch to different prefix if the preferred lifetime of the current landmark prefix changes.

5.2.4 Sending Router Solicitations

Upon the occurrence of a Layer 2 link-up event notification, hosts SHOULD send a Router Solicitation. Hosts SHOULD apply rate limiting and/or hysteresis to this behaviour as appropriate to the link technology subject to the reliability of the hints.

Hosts SHOULD include a Landmark Option (LO) in the RS message with the landmark prefix chosen based on the rules in Section 5.2.3.

Hosts SHOULD include a tentative source link layer address option (TSLLAO) in the RS message [7]. The router solicitation message is sent to the All_Routers_Multicast address and the source address MUST be the link local address of the host.

The host MUST consider its link local address to be in the "Optimistic" state for duplicate address detection [6] until either the returned RA confirms that the host has not switched to a new link or, if an link change has occurred, the host has performed optimistic duplicate address detection for the address.

5.2.5 Processing Router Advertisements

When the host receives a Router Advertisement, the host checks for the conditions and derives the associated conclusions given below:

If the RA contains a Landmark Option with the "Y" flag set that matches the Landmark Option in the last transmitted Router Solicitation, then that indicates that no link change has occurred and current configuration can be assumed to still be current.

If the RA includes any prefixes in either a PIO or LPO that matches a prefix in DNAHostPrefixList then the host can conclude that no link change has occurred and current configuration can be assumed to still be current.

If the RA includes a LinkID that matches an entry in DNAHostLinkIDList, then the host can conclude that no link change

has occurred and the current configuration can be assumed to still be current.

If the RA is a Complete RA, as indicated by the "Complete" flag in the RA header, and there are no prefixes included in it in either a PIO or LPO that are also in the hosts DNAHostPrefixList, then the host can conclude that it has changed link and SHOULD initiate re-configuration using the information in the received Router Advertisement.

If the RA is not a CompleteRA, but includes a LinkID that is not in DNAHostLinkIDList and no prefixes that match entries in DNAHostPrefixList, then the host can conclude that it has changed link and SHOULD initiate re-configuration using the information in the received Router Advertisement.

If the received RA is not complete, contains no prefixes that are stored in DNAHostPrefixList, does not contain a Landmark Option that matches a corresponding option in the most recent RS and contains no LinkID, then the host SHOULD use CPL logic to decide whether or not to reconfigure as described in [15].

5.2.5.1 Maintaining the DNAHostPrefixList

If a Router Advertisement does not indicate a link change, the host updates its DNAHostPrefixList, adding any new prefixes if necessary.

If the Router Advertisement has the C flag set, then the host SHOULD make the DNAHostPrefixList match the contents of the advertisement. Any new prefixes are added and any prefixes in the list that are absent in the advertisement are removed. Expiry times on prefixes are updated if the prefix was contained in a PIO, but not if it was contained in an LPO.

If the Router Advertisement does not have the C flag set, then the host SHOULD add any new prefixes and update expiry times as above, but SHOULD NOT remove any entries from DNAHostPrefixList.

When initiating reconfiguration due to link change, the host MUST remove all prefixes in the DNAHostPrefixList and repopulate it with the prefixes in the Prefix Information Options and Learned Prefix Option, if any, in the RA.

5.2.5.2 Router Reachability Detection and Default Router Selection

The receipt of a unicast RA from a router in response to a multicast RS indicates that the host has bi-directional reachability with the

routers that responded. Such reachability is necessary for the host to use a router as a default router, in order to have packets routed off the host's current link. If the destination address of the received RA is a unicast address, the host knows the router heard its RS, and therefore that the host has reachability with the router.

Prior to sending a DNA RS in response to an indication of link change, the host SHOULD set all Neighbor Cache entries for routers on its Default Router List to STALE. When the host receives an RA in reply to the RS, the host SHOULD mark that router's Neighbor Cache Entry [3] as REACHABLE, or add a Neighbor Cache Entry in the REACHABLE state if one does not currently exist.

The host SHOULD also update its Default Router List in the following fashion. If any of the routers returning RAs are already on the default router list, the host SHOULD use the information in the RA to update the Default Route List entry with the new information. The host SHOULD add entries to the Default Router List for any routers returning RAs that are not on the list. The host SHOULD confine selection of a router from the Default Router List to those routers whose Neighbor Cache entries are in the REACHABLE state. Note that the Default Router List SHOULD be updated as described here regardless of whether the RA indicates that the host has changed to a new IP link, since changes in router reachability are possible on some link types even if the host remains on the same IP link.

Note that this procedure does not prevent a MN from sending packets to its current default router while the RA solicitation is in progress and if reachability with the current default router is unchanged, there should be no change in default router after the RA solicitation completes. If the current default router is still reachable, it will forward the packets.

5.2.6 DNA and Address Configuration

When a host moves to a new point of attachment, a potential exists for a change in the validity of its unicast and multicast addresses on that network interface. In this section, host processing for address configuration is specified. The section considers both statelessly and statefully configured addresses.

5.2.6.1 Duplicate Address Detection

A DNA host MUST support optimistic Duplicate Address Detection [6] for autoconfiguring unicast link local addresses. If a DNA host uses address autoconfiguration [8] for global unicast addresses, the DNA host MUST support optimistic Duplicate Address Detection for autoconfiguring global unicast addresses.

5.2.6.2 DNA and the Address Autoconfiguration State Machine

When a link level event occurs on a network interface indicating that the host has moved from one point of attachment to another, it is possible that a change in the reachability of the addresses associated with that interface may occur. Upon detection of such a link event and prior to sending the RS initiating a DNA exchange, a DNA host MUST change the state of addresses associated with the interface in the following way (address state designations follow RFC 2461):

- Addresses in the "Preferred" state are moved to the "Optimistic" state, but the host defers sending out an NS to initiate Duplicate Address Detection.

- Addresses in the "Optimistic" state remain in the "Optimistic" state, but the host defers sending out an NS to initiate Duplicate Address Detection.

- Addresses in the "Deprecated" state remain in the "Deprecated" state.

- No addresses should be in the "Tentative" state, since this state is unnecessary for nodes that support optimistic Duplicate Address Detection.

A host MUST keep track of which "Preferred" addresses are moved to the "Optimistic" state, so it is possible to know which addresses were in the "Preferred" state and which were in the "Optimistic" state prior to the change in point of attachment.

In order to perform the DNA transaction, the DNA host SHOULD select one of the unicast link local addresses that was in the "Preferred" state prior to switching to "Optimistic" and utilize that as the source address on the DNA RS. If the host had no "Preferred" unicast link local address but did have an address in the "Optimistic" state, it MUST utilize such an address as the source address. If the host currently has no unicast link local addresses, it MUST construct one and put it into the "Optimistic" state and note this address as having been in the "Optimistic" state previously, but defer sending the NS to confirm. Note that the presence of a duplicate unicast link local address on the link will not interfere with the ability of the link to route a unicast DNA RA from the router back to the host nor will it result in corruption of the router's neighbor cache, because the TSLLA option is included in the RS and is utilized by the router on the RA frame without changing the neighbor cache.

When the host receives unicast or multicast RAs from the router, if

the host determines from the received RAs that it has moved to a new link, the host MUST immediately move all unicast global addresses to the "Deprecated" state and configure new addresses using the subnet prefixes obtained from the RA. For all unicast link local addresses, the host MUST initiate NS signaling for optimistic Duplicate Address Detection to confirm the uniqueness of the unicast link local addresses on the new link.

If the host determines from the received RAs that it has not moved to a new link (i.e. the link has not changed) and the previous state of an address was "Optimistic", then the host MUST send an NS to confirm that the address is unique on the link. This is required because optimistic Duplicate Address Detection may not have completed on the previous point of attachment, so the host may not have confirmed address uniqueness. If the previous state of an address was "Preferred", whether or not the host initiates optimistic Duplicate Address Detection depends on the configurable DNASameLinkDADFlag flag. A host MUST forgo sending an NS to confirm uniqueness if the value of the DNASameLinkDAD flag is False. If, however, the DNASameLinkDAD flag is True, the host MUST perform optimistic duplicate address detection on its unicast link local and unicast global addresses to determine address uniqueness.

5.2.6.3 DNA and Statefully Configured Addresses

The DHCPv6 specification [9] requires hosts to send a DHCPv6 CONFIRM message when a change in point of attachment is detected. Since the DNA protocol provides the same level of movement detection as the DHCPv6 CONFIRM, it is RECOMMENDED that DNA hosts not utilize the DHCPv6 CONFIRM message when a DNA RA is received, to avoid excessive signaling. If, however, a non-DNA RA is received, the host SHOULD use the DHCPv6 CONFIRM message as described in RFC 3315 [9] rather than wait for additional RAs to perform CPL, since this will reduce the amount of time required for the host to confirm whether or not it has moved to a new link. If the CONFIRM message validates the addresses, the host can continue to use them.

When a DNA RA is received and the received RA indicates that the host has not moved to a new link, the host SHOULD apply the same rules to interpreting the 'M' flag in the received RA and any subsequently received RAs as in Section 5.5.3 of RFC 2461 [3]. That is, if an RA is received with the 'M' flag set, then the 'M' flag value is copied into the ManagedFlag, and if the ManagedFlag changes from False to True the host should run DHCPv6, but if the ManagedFlag changes from True to False, the host should continue to run DHCPv6. If, however, the value of the ManagedFlag remains the same both before and after the change in point of attachment on the same link has been confirmed, it is NOT RECOMMENDED that the host run DHCPv6 to obtain

new addresses, since the old addresses will continue to be valid.

If the DNA RA indicates that the host has moved to a new link or the DHCPv6 CONFIRM indicates that the addresses are invalid, the host MUST move its old addresses to the "Deprecated" state and MUST run DHCPv6 to obtain new addresses. Normally, the DHCPv6 operation is 4-message exchange, however, this exchange allows for redundancy (multiple DHCPv6 servers) without wasting addresses, as addresses are only provisionally assigned to a host until the host chooses and requests one of the provisionally assigned addresses. If the DNA host supports the Rapid Commit Option [9], the host SHOULD use the Rapid Commit Option in order to shorten the exchange from 4 messages to 2 messages.

5.2.6.4 Packet Delivery During DNA

The specification of packet delivery before, during, and immediately after DNA when a change in point of attachment occurs is out of scope for this document. The details of how packets are delivered depends on the mobility management protocols (if any) available to the host's stack.

5.2.6.5 Multicast Address Configuration

If the returning RAs indicate that the host has not moved to a new link, no further action is required for multicast addresses to which the host has subscribed using MLD Report [10]. In particular, the host MUST NOT perform MLD signaling for any multicast addresses unless such signaling was not performed prior to movement to the new point of attachment. For example, if an address is put into the "Optimistic" state prior to movement but the MLD Report for the Solicited_Node_Multicast_Address is not sent prior to movement to a new point of attachment, the host MUST send the MLD Report on the new point of attachment prior to performing optimistic Duplicate Address Detection. The host SHOULD use the procedure described below for sending an MLD Report.

If, on the other hand, the DNA RA indicates that the host has moved to a new link, the host MUST issue a new MLD Report to the router for subscribed multicast addresses. MLD signaling for the Solicited_Node_Multicast_Addresses [8] MUST be sent prior to performing signaling for optimistic DAD.

To avoid lengthy delays in address reconfiguration, it is RECOMMENDED that the host send the MLD Report for newly configured addresses immediately, as soon as the addresses have been constructed, rather than waiting for a random backoff.

Hosts MUST defer MLD signaling until after the results of DNA have confirmed whether or not a link change has occurred.

6. Backward Compatibility

6.1 Non-DNA Host with DNA Routers

The RS message sent by non-DNA hosts will not contain any of the new options defined by this document. The host will receive a Complete RA in response to the solicitation message and process it as per RFC 2461. This means that it will drop the unrecognised Learned Prefix option, but process the included PIOs and non-DNA flags normally.

6.2 DNA Host with Non-DNA Routers

The routers will behave based in the recommendations of RFC 2461 [3] and ignore the new options defined in this memo. Hosts will receive RA message without the FastRA flag in the RA header set and will fallback on CPL for link identification. Obviously, the objective of receiving fast response for RS message can not be achieved.

This case can occur on a link with no DNA routers or on a link with a mix of the two. In the latter, usually a response from the DNA router(s) will be received first and CPL will just be used with the non-DNA Router Advertisement to confirm that no movement has taken place since the previous DNA advertisement.

7. Security Considerations

7.1 Attacks on the Token Bucket

A host on the link could easily drain the token bucket(s) of the router(s) on the link by continuously sending RS messages on the link. For example, if a host sends one RS message every UnicastRAInterval, and send a additional RS every third UnicastRAInterval, the token bucket in the router(s) on the link will drain within MaxUnicastRABurst * UnicastRAInterval * 3 time-units. For the recommended values of UnicastRAInterval and MaxUnicastRABurst, this value is 3000 milliseconds. It is not clear whether arrival of such RS messages can be recognized by the router as a DoS attack. This attack can also be mitigated by aggregating responses. Since only one aggregation is possible in this interval due to MIN_DELAY_BETWEEN_RAS restriction, the routers may not be able protect the tokens in the bucket.

7.2 Attacks on DNA Hosts

RFC 3756 outlines a collection of threats involving rouge routers.

Since DNAv6 requires a host to obtain trustworthy responses from routers, such threats are relevant to DNAv6. In order to counter such threats, DNAv6 hosts SHOULD support RFC 3971 (SEND) secure router discovery.

8. IANA Considerations

This memo defines two new Neighbor Discovery [3] options, which must be assigned Option Type values within the option numbering space for Neighbor Discovery messages:

1. The Landmark option, described in Section 4.3; and

2. The Learned Prefix option, described in Section 4.4.

9. Acknowledgments

The design presented in this document grew out of discussions among the members of the DNA design team (JinHyeock Choi, Tero Kauppinen, James Kempf, Sathya Narayanan, Erik Nordmark and Brett Pentland). The spirited debates on the design, and the advantages and dis-advantages of various DNA solutions helped the creation of this document.

Thanks to Syam Madanapalli who co-authored draft-jinchoi-dna-protocol2 from which this draft draws ideas, as well as providing feedback on draft-pentland-dna-protocol from which most of the text for this draft comes.

Thanks to Greg Daley for much feedback on draft-pentland-dna-protocol and for helping to work out how to merge the two drafts into this one.

Thanks to Jari Arkko, Jim Bound, Tero Kauppinen, Syam Madanapalli, Mohan Parthasarathy, Subba Reddy, and Christian Vogt for their review of this document.

Thanks to Gabriel Montenegro for his review of draft-pentland-dna-protocol.

Thanks also to other members of the DNA working group for their comments that helped shape this work.

10. References

10.1 Normative References

[1] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[2] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[3] Narten, T., Nordmark, E., and W. Simpson, "Neighbor Discovery for IP Version 6 (IPv6)", RFC 2461, December 1998.

[4] Johnson, D., Perkins, C., and J. Arkko, "Mobility Support in IPv6", RFC 3775, June 2004.

[5] Arkko, J., Kempf, J., Sommerfeld, B., Zill, B., and P. Nikander, "SEcure Neighbor Discovery (SEND)", draft-ietf-send-ndopt-06 (work in progress), July 2004.

[6] Moore, N., "Optimistic Duplicate Address Detection for IPv6", draft-ietf-ipv6-optimistic-dad-05 (work in progress), February 2005.

[7] Daley, G., "Tentative Source Link-Layer Address Options for IPv6 Neighbour Discovery", draft-daley-ipv6-tsllao-00 (work in progress), June 2004.

10.2 Informative References

[8] Thomson, S. and T. Narten, "IPv6 Stateless Address Autoconfiguration", RFC2462 2462, December 1998.

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

[10] Vida, R. and L. Costa, "Multicast Listener Discovery Version 2 (MLDv2) for IPv6", RFC 3810, June 2004.

[11] Choi, J., "Detecting Network Attachment in IPv6 Goals", draft-ietf-dna-goals-04 (work in progress), December 2004.

[12] Narayanan, S., Daley, G., and N. Montavont, "Detecting Network Attachment in IPv6 - Best Current Practices", draft-narayanan-dna-bcp-00 (work in progress), June 2004.

[13] Yamamoto, S., "Detecting Network Attachment Terminology", draft-yamamoto-dna-term-00 (work in progress), February 2004.

[14] Manner, J. and M. Kojo, "Mobility Related Terminology", draft-ietf-seamoby-mobility-terminology-06 (work in progress), February 2004.

[15] Choi, J. and E. Nordmark, "DNA with unmodified routers: Prefix list based approach", draft-ietf-dna-cpl-00 (work in progress), April 2005.

[16] Pentland, B., "An Overview of Approaches to Detecting Network Attachment in IPv6", draft-dnadt-dna-discussion-00 (work in progress), February 2005.

Authors' Addresses

James Kempf
DoCoMo Communications Labs USA
USA

Phone:
Email: kempf@docomolabs-usa.com

Sathya Narayanan
Panasonic Digital Networking Lab
Two Research Way, 3rd Floor
Princeton, NJ 08536
USA

Phone: 609 734 7599
Email: sathya@Research.Panasonic.COM
URI:

Erik Nordmark
Sun Microsystems, Inc.
17 Network Circle
Mountain View, CA
USA

Phone: +1 650 786 2921
Email: erik.nordmark@sun.com

Brett Pentland (editor)
Centre for Telecommunications and Information Engineering
Department of Electrical and Computer Systems Engineering
Monash University
Clayton, Victoria 3800
Australia

Phone: +61 3 9905 5245
Email: brett.pentland@eng.monash.edu.au

JinHyeock Choi
Samsung Advanced Institute of Technology
PO Box 111
Suwon 440-600
Korea

Phone: +82-31-280-8194
Email: jinchoe@samsung.com

Appendix A. How the Goals are Met?

The DNA goals document [11] contains a list of goals identified by G1 to G10. This is also enumerated in the solutions discussion document [16] generated by the DNA design team. This section discusses how the proposed scheme addresses each of these goals.

G1 The Complete RA contains the complete list of prefixes advertised on the link allowing the host to determine whether link change has occurred and to re-configure if necessary.

G2 Under normal circumstances the host will receive a RA response within round-trip time and some processing time on the router. If the first RA message is lost, if another router is on the link, a second RA should arrive within a slot time and so on.

G3 Non movement scenarios will be correctly identified because the landmark will be confirmed by the router(s) on the link or the Complete RA will have prefixes that have already been seen, indicating non-movement.

G4 A single RS/RA message exchange is initiated in response to a hint that link change may have occurred.

G5 The existing RS/RA signalling is used along with unsolicited RA messages. Some new options and flags are proposed.

G6 Only link scope signaling is used.

G7 SEND can be used to protect the RS and RA messages exchanged.

G8 If SEND is not deployed, then a rogue device could cause a host to think its configuration is invalid by sending an RA that answers the RS question incorrectly. A similar effect is already possible, however, by a rogue device sending an RA with valid prefixes with zero lifetimes.

G9 The CPL logic allows a graceful fallback position for dealing with non-DNA routers and non DNA hosts will still receive the benefit of receiving an RA response from its current router faster than RFC 2461.

G10 This technique is carried out on an interface by interface basis. A host with multiple interfaces can get information about changes to configuration on each interface, but would need a higher level process to decide how the information from the various interfaces relates to each other.

Intellectual Property Statement

Disclaimer of Validity

Copyright Statement

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.