DNA Working Group B. Pentland
Internet-Draft Monash University CTIE
Expires: August 15, 2005 February 14, 2005

An Overview of Approaches to Detecting Network Attachment in IPv6
draft-dnadt-dna-discussion-00.txt

Status of this Memo

Copyright Notice

Abstract

This document is a discussion of potential solutions to the problem of rapidly and reliably detecting attachment to a network and determining when host configuration change is required.

Table of Contents

1. Introduction

This document represents, an overview of the discussion of the DNA design team on potential solutions to the problem of rapid and reliable network attachment detection. The design team was comprised of JinHyeock Choi, Tero Kauppinen, James Kempf, Sathya Narayanan, Erik Nordmark, and Brett Pentland.

While we were unable to settle on a single solution, a number of different techniques were discussed at length, and this document provides a summary of them, including their advantages and disadvantages.

In general, it was felt that fast attachment detection will require some kind of hint from layer two. The link layer event notifications draft [13], provides a discussion of layer two information available from a number of link types. Though it talks about "link-up" and "link-down", only link-up is considered here, for its utility in triggering a router solicitation from a host.

In considering DNA, we have discussed how to get to the point where a host knows whether or not its IP configuration is likely to be valid, and has enough information to be able to start reconfiguration if necessary. At the end of "DNA" it either knows that it is connected to the same link as its current configuration implies, or it knows that it's connected to a new link, its IP configuration is invalid, as it has at least one prefix that it can use for Stateless Address Auto-configuration, if appropriate. It might also just use this as a trigger to start DHCP and/or any higher layer authorization/ authentication as required.

2. Terminology

The following terms are used throughout the document

Link - As defined in RFC 2461 [2].

Landmark - Some attribute that may be associated with a specific link such as a prefix or global router address.

Link Identifier
- A consistent identifier used in some way by all routers on a link to allow hosts to distinguish the link from other links.

3. Objectives for a Solution to the Problem of DNA

Because detecting network attachment will frequently happen on a network that a host has not seen before, information about that network will be required in order to configure addresses, etc. for future communication. The usual mechanism for obtaining this information is the Router Advertisement. A solicitation will be required in order to take advantage of hints from lower layers and speed up the detection process. It was felt that by crafting the information in the solicitation and advertisement, a single RS/RA exchange should be sufficient for DNA.

In order to detect attachment rapidly, RA response should be as fast as possible. To this end, a DNA protocol ideally should not include any timer-induced delays for the first RA in response to an RS, though if a technique is sufficiently superior in other areas, some delay may be acceptable.

4. Assumptions

- All of the routers on a link can be expected to receive packets sent to the "all-routers" multicast address (ff02::2) with some packet-loss probability < 1.

- Hosts with interfaces that can connect to more than one link concurrently are able to distinguish packets from the separate links and thus abstract the link connections as separate virtual interfaces.

5. The Problem

Given the above objectives, the problem to be solved can then be broken into two parts:

1. Crafting the information in the RS/RA exchange to allow accurate determination of whether a link change has occurred, necessitating a change to existing configuration, and including the necessary prefixes, etc. to allow those configuration changes to be made if needed.

2. Ensuring that an RA is received as rapidly as possible in response to solicitation.

6. Determining Whether Link Change Has Occurred

There are a number of ways that the RFC 2461 RS/RA exchange could be modified to allow unambiguous determination of whether configuration

change is required with a single exchange. These may be broadly divided into two classes:

1. Techniques that add information to the RA to allow hosts to make a correct decision.

2. Techniques that add information to the RS to ask a question of routers on the link, getting back an answer in the RA.

6.1 Techniques that add information to the RA.

6.1.1 Explicit Link Identifier.

The routers on a link, by some mechanism, agree on a single identifier for the link that is different from any corresponding identifier used on another link that a host is likely to transition to directly. The identifier would then be included in router advertisements to allow hosts to immediately recognise whether this link is the same as the one they believe themselves to be connected to.

Link Identifier techniques have an associated cost of needing a (secure) mechanism for the routers to come to agreement on the value of the link ID and also a means to change it if necessary without confusing the hosts on the link.

6.1.1.1 Random Link Identifiers.

There are a number of options for the identifier. It could be simply a random number carried in a new advertisement option. This is quite simple, independent of changes to prefixes or routers on a link, but takes some extra bytes per RA and has a non-zero probability of there being multiple links using the same identifier.

6.1.1.2 Prefix Based Link Identifiers.

A global prefix is another possible candidate for use as a Link Identifier. This would have the advantage of being unique, requires no additional RA bytes if a router is already advertising the prefix and just needs to add a flag to indicate that it is also a link identifier. If it is not possible to find a global prefix that all of the routers on the link are announcing, then some routers will need to include a prefix that is only a link identifier, thus increasing the size of the RA. There also needs to be a mechanism to change the Link Identifier if the chosen prefix ceases to be used on the link.

Another way to generate an identifier from prefixes is to collect all of the prefixes active on the link and take some kind of hash over the ordered list of them, or alternatively, just one of them. This would generate an identifier like in Section 6.1.1.1, but with a guarantee of uniqueness. As with other prefix based link identifiers, the prefixes need to be monitored to ensure that they are current. A mechanism is needed to be able to change the identifier in response to changes in the prefixes. Using a hash based on a single prefix would be less vulnerable to changes than one based on all of the prefixes.

6.1.2 Complete Router Advertisement.

Routers on a link listen for other routers' advertisements and include a complete list of all prefixes in use on the link in RAs they send. The RA would carry a flag to indicate that the RA did indeed include a complete list.

Care should be taken to differentiate prefixes that are learnt from those that were originally configured on a router. The prefixes that are only learnt would need to be included in special PIOs so that hosts only use them for identification purposes and not as regular PIOs. The special PIOs could have their own code so that they are unrecognised by hosts that don't implement a new DNA specification. Another alternative would be to use conventional PIOs but with the L, A and R flags not set, and with a new D-flag (DNA) to indicate that the prefix is reflected for DNA purposes. PIOs with the L, A and R flags all cleared have no effect on non-DNA hosts.

This technique has the advantages of forming an implicit identifier, is quite simple, requires no changes to solicitations, and makes it easy for a host to generate a complete prefix list for a link, allowing it to deal easily with an RA from a non-DNA router. The main cost is the size of the RAs. If routers on a link have matching sets of prefixes, then this is no cost but if there are differences then some of the RAs will be larger. If the sets are disjoint, then all of the solicited RAs will be larger and there is no implicit upper bound on the increase.

6.2 Techniques That Ask a Question in the RS.

6.2.1 Requested Landmark.

Routers on a link listen for other routers' advertisements and collect the prefixes so that they know all of the active prefixes on the link. Hosts, when soliciting, select a prefix that they have seen previously and include it in the solicitation. Routers responding to the solicitation can then included a yes or no flag (as

distinct from no flag at all) that says whether or not the prefix is in use on this link.

This technique is again quite simple. The cost is an increase in the size of the RS (with no corresponding increase in the RA) but the increase is known, fairly modest, and fixed. There is, in general, a 1:N ratio of RSs to RAs, where N is the number of routers on the link. The result of this is that increases to the RS size are less costly than increases to the RA size.

In certain other techniques it is possible to reduce the number of RAs by using multicast to answer multiple solicitations at once. This can only be achieved if delays are added which is something to be avoided if possible for the first response to an RS. Thus the number of RAs is always >= the number of RSs and hence increases to the RA size are still more costly than increases to the RS size.

6.2.2 Priority Landmark.

Hosts include the address of their default router in solicitations. A fast RA mechanism is used that guarantees that if that router is on the link then it will respond first. If the first response is not from the default router then the host can assume that it has moved to a new link, possibly after checking that the included PIOs support this.

This technique has the advantage of confirming bi-directional reachability with the default router when movement has not taken place. The cost is an increase in the RS size and a dependence on a mechanism to ensure that the requested router is always the first to respond.

6.2.3 Hybrid Landmark.

Routers on a link include at least one PIO in unsolicited advertisements that includes an R-bit, and monitor the R-bit PIOs of other routers on the link. This gives out addresses that can be used as landmarks for the link. Hosts pick a landmark that they have seen most recently and include it in solicitations. Routers responding to this solicitation include flags to indicate whether or not this landmark has been seen on the link. The fast RA mechanism can be designed to allow the router with the requested landmark to respond first.

This again has the advantage of confirming bi-directional reachability with the default router when movement has not taken place. It also allows any router to respond and give a definitive answer to the link-change question. The main cost is the increased

RS size.

This, the two preceding landmark schemes, and Complete RA all require routers to place timers on the gathered prefixes to be sure that old information can be discarded if prefixes are moved to different links.

7. Fast Responses to Solicitation

Again there are a number of ways that standard procedures could be modified to allow a router advertisement to be received quickly following solicitation.

7.1 Fast Router Discovery.

draft-jinchoi-dna-frd-00.txt

Access Points cache the most recent RA(s) and forward it(them) to a host upon detection of its association.

This is very simple and potentially very fast and places no requirements on hosts or routers but is link-specific and raises some security concerns since it is, by definition, a "man in the middle".

Where there are multiple routers on a link it needs to be determined which RA(s) will be forwarded, and how to time out old cache entries.

7.2 Simple Fast RA.

draft-mkhalil-ipv6-fastra-05.txt

Select one router on each link to be allowed to respond immediately to solicitations.

Again very simple, but requires a mechanism to select the fast router, introduces a single point of failure, and may result unbalanced loading of routers.

7.3 Deterministic Fast RA.

draft-daley-dna-det-fastra-01.txt

Routers on a link negotiate amongst themselves an ordering for responding to solicitations. Responses are made in order at fixed intervals starting from zero delay for the first router.

This removes the single point of failure problem and means that losing an RA doesn't slow down the RS/RA exchange much (unless there

is only one router). The costs include the necessity for the routers to engage in negotiation to select the ordering and fact that that ordering may result in unbalanced loading of the routers.

It would be fairly simple to alter the behaviour to reorder the responses based on some function of the source of the RS to spread load evenly.

7.4 Negotiation-free Deterministic Fast RA.

Routers on a link listen for advertisements from other routers and form tokens for them from the source addresses. Hosts include a tentative source link layer address option (TSLLAO) [11] in solicitations. When an RS is received by a router, some function of the TSLLAO is XORed with each of the router tokens to create a ranking. One or more of the routers then respond in order with fixed delays starting from zero.

The advantage here is that routers just need to listen to RAs to determine an ordering that will vary from solicitation to solicitation, it doesn't have a single point of failure and will result in multiple (if there are multiple routers) RAs quickly. The main cost is the need for the RSs from hosts to include a TSLLAO, but this will be necessary for any technique where sending unicast RAs is required, unless a separate NS/NA exchange is done between the RS and RA.

If multicast RAs are to be used, then TSLLAOs are not necessary for the transmission of the RA to the host without an NS/NA exchange. The cost of including a TSLLAO might be removed by determining a base ordering for the routers based on the tokens, and then perturbing that ordering using a function of the time that the RS is received (for example, shifting the ordering by the minute of the reception time, modulo the number of routers). The cost of this is the necessity to synchronize the router's clocks and a small period of ambiguity around the time when the part of the timestamp used changes (e.g. when the clock ticks over from one minute to the next).

7.5 Probabilistic Fast RA.

Routers on a link listen for advertisements from other probabilistic fast RA routers (as defined by a flag) and count the number of them. Responses to solicitations are scheduled into one of count+1 slots spaced, say, 20 ms apart starting from zero. A slight bias towards the zero slot can be done to improve average response. Maximum and minimum values for the number of slots can be set to limit the effects of unknown or spurious routers. Details of this technique can be found in [10] including claimed intellectual property rights.

Again, routers just have to listen to other routers on the link to get the information they need to determine the sending delay. The trust requirements are even lower, having no need for security associations between routers. This is because they are just counting routers, storing minimal information about them, and abnormally high counts are easily ignored.

An upper bound is placed on introduced delay, and average delays are quite low, albeit non-zero. Hosts will often, but not always receive an immediate response.

8. Dealing with Legacy Routers

It is likely that hosts will encounter links that have routers that don't have any enhancements to support DNA. It is important that they are still able make correct decisions quickly about whether link change has occurred. By maintaining a list of all of the prefixes in use on a link, they can then use any prefix in an RA to make a decision. One way to maintain such a list is described in [8].

An unsolicited RA might indicate an added prefix or router, rather than movement, but can be used as a trigger to send an RS to test the link. There is a small chance of an erroneous decision, even after an RS, if a prefix or router is added to a link. An implementation may choose to delay making configuration changes until further confirmation if the cost of an incorrect decision is high. It may wait for further RAs or even re-solicit to achieve that confirmation.

The Complete Router Advertisement technique described in Section 6.1.2 integrates well with this because a single RA can provide all of the prefixes in use on a link, simplifying the process of gathering them.

9. Putting Things Together

For a complete solution, a fast RA technique needs to be mated up with a technique for using the RS/RA exchange for identification. The two parts are largely but not completely independent. For example, deterministic fast RA defines a "router to router" message that can be reused to negotiate a link identifier.

To evaluate solutions, the way they meet the goals laid out in the DNA goals document [9] should be considered. Quoting from the goals document:

G1 DNA schemes should detect the identity of the currently attached link to ascertain the validity of the existing IP configuration. They should recognize and determine whether a link change has

occurred and initiate the process of acquiring a new configuration if necessary.

G2 DNA schemes should detect the identity of an attached link with minimal latency lest there should be service disruption.

G3 In the case where a host has not changed a link, DNA schemes should not falsely assume a link change and an IP configuration change should not occur.

G4 DNA schemes should not cause undue signaling on a link.

G5 DNA schemes should make use of existing signaling mechanisms where available.

G6 DNA schemes should make use of signaling within the link (particularly link-local scope messages), since communication off-link may not be achievable in the case of a link change.

G7 DNA schemes should be compatible with security schemes such as Secure Neighbour Discovery [3].

G8 DNA schemes should not introduce new security vulnerabilities. The node supporting DNA schemes should not expose itself or others on a link to additional man-in-the-middle, identity revealing, or denial of service attacks.

G9 The nodes, such as routers or hosts, supporting DNA schemes should work appropriately with unmodified nodes, such as routers or hosts, which do not support DNA schemes.

G10 Hosts, especially in wireless environments, may perceive routers reachable on different links. DNA schemes should take into consideration the case where a host is attached to more than one link at the same time.

9.1 Requested Landmark with Negotiation-free Deterministic Fast RA

(How routers collect the information needed to determine RS response order)

- Routers send periodic advertisements including a flag that indicates that they are DNA routers.
  - An interval option (rfc3775) should be included in multicast RAs to facilitate detection of lost routers
  - If more than one SA is in use then a PIO with the R-bit (rfc3775) should be included.
- Routers listen to other routers' advertisements and use them to

maintain a list of all active prefixes on the link.

  - Upper bound needed on list length to prevent memory overflow.
  - routers collect the source addresses of routers they have received RAs from.
    - a token equal to a hash of the IID (could be the full SA, but presumably the IID is where the variation is) of the collected addresses is stored for each router, associated with a timer related to the interval option in the received RA.
      - the IIDs are also stored
      - R-bit PIOs should be monitored to detect the use of multiple SAs by a single router - only the lowest IID and its token should be stored.
    - an upper bound is needed on the number of tokens that will be stored (to prevent overflow).
      - a fallback position is needed in the case of a full list.

(How hosts request information from the link)

- Hosts solicit including a TSLLAO (for unicast responses), an 8 bit counter that is incremented for each RS sent, and an option including the most recently received prefix.

(How routers respond to solicitation)

- Unicast is used by routers for responding to solicitations, subject to rate control by a token bucket scheme.
  - Exhaustion of the token bucket results in the use of multicast, subject to the controls of rfc2461.
- Routers examine the TSLLAO of incoming RSs, XOR the IID contained with each of the stored tokens (including its own) and compare them to calculate a ranking for themselves.
  - The top ranked router responds immediately.
  - Some number of the others follow at fixed intervals.

(How hosts interpret RAs received)

- The response RAs contain one of two flags indicating whether or not the requested prefix is active on this link, and a copy of the counter in the received RS.
- Hosts look at the flags in the received RA to decide on a course of action.
  - Yes flag set: no action required - maybe NS/NA exchange with current default router at leisure.
  - No flag set: clear NC, and use one of the prefixes in the RA to form a new address - test with optiDAD, etc.
  - Both set: not allowed - treat as though none set
  - None set: invoke CPL logic
    - CPL logic: (going from memory, may need correcting - more aggressive approach to new prefixes)
      - hosts try to form a complete list of all prefixes available

on a link.
- send (possibly multiple) RSs at suitable intervals
- RAs received in this time considered to be part of prefix list building
- RAs with all prefixes disjoint from current prefix list assumed to be from a new link (maybe test with NS/NA to current default router with short timeout)
  - Clear NC, form new address, etc., restart CPL building.

9.1.1 How the goals are met

G1 The answer to the landmark question gives a positive indication of whether link change has occurred, and the RA will contain the information required to reconfigure if necessary.

G2 Under normal circumstances, a host that sends an RS should get an RA back from a router in one round trip time plus a small processing delay. If that RA is lost another should arrive after a small delay if there is more than on router on the link.

G3 Non-movement situations are correctly detected.

G4 A single RS/RA exchange is initiated in response to a hint that a link change may have occurred. Routers build the state they need to respond to RSs simply by listening to the unsolicited RAs of other routers.

G5 The RS/RA mechanism is all that is required. A new option is defined for the RS and a pair of new flags is required in RAs.

G6 Only link-scope signalling is used.

G7 SeND can be used to protect RSs with a specified source address and will protect the new option against tampering. It will also protect the new flags in the RA against tampering.

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 hosts and routers.

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.

9.2 CompleteRA with Probabilistic Fast RA

(How routers gather all the routers and prefixes on a link.)

- Routers include a "D" flag (DNA) in RAs to indicate that they will participate in probabilistic fast RA.
- Routers listen to other routers' advertisements and use them to maintain a list of all active prefixes on the link.
  - They also keep a count of the number of DNA routers on the link.
  - An upper bound needed on list lengths to prevent memory overflow.

(How routers decide when to respond to an RS)

- Upon reception of an RS, an RA is scheduled into one of count+1 time slots starting from zero with, say, 20 ms spacing.
  - count is set to the number of probabilistic routers heard with configurable upper and lower bounds
  - multicast RAs may be used (subject to the rate limiting restrictions of RFC 2461) and if they are, solicitations that would result in the scheduling of an RA after an already scheduled RA may be ignored

(How routers advertise CompleteRA)

- CompleteRA is the RA that contains the complete set of all prefixes on the link, including a flag to indicate that the list is indeed complete.
  - PIOs seen on the link but not originating from the sending router could use a new type code (as distinct from a flag which would be ignored by non-dna hosts).
  - Routers advertise the CompleteRA upon receiving an RS as indicated above.
  - If too many prefixes are in use to fit in an RA then the complete flag cannot be set and CPL may be relied on with conventional logic.

(How hosts check for link change with CompleteRA.)

- A host receiving an RA compares the prefixes in the RA to their own list of current prefixes.
  - if there is overlap between the prefix lists (they should match exactly) then nothing needs to be done - maybe NS/NA exchange with current default router at leisure.

  - if they are disjoint, then it is a new link and the NC can be cleared and new addresses formed, etc.

(How hosts generate the Complete Prefix List with a single CompleteRA)

- Upon receiving a CompleteRA, hosts can generate the Complete Prefix List without further action.

(How to interoperate with non-supporting links)

- The Complete Prefix List logic is simpler:
  - similar to above
  - when building the CPL, if an RA is received with the complete flag set, then those prefixes constitute the CPL and the host can go straight to the state where list is considered built.
  - In the built state, if a new prefix is received that has a disjoint prefix set, then a new link is implied.
    - reconfiguration should be commenced, possibly after an attempted NS/NA exchange with default router with a short timeout if the cost for an incorrect decision is high (could just be a new router and prefix on the link).

9.2.1 How the goals are met

G1 The complete set of prefixes in the RA gives a positive indication of whether link change has occurred, and contains the information required to reconfigure if necessary.

G2 The router advertisement is usually transmitted to the host in one round trip time plus a processing delay. Sometimes there will be a slot delay if no routers schedule for slot zero, adding 20 ms to the delay.

G3 Non-movement situations are correctly detected.

G4 A single RS/RA exchange is initiated in response to a hint that a link change may have occurred. Routers build the state they need to respond to RSs simply by listening to the unsolicited RAs of other routers. If a complete RA is received without solicitation, then no solicitation is required; the RA contains enough information.

G5 The RS/RA mechanism is all that is required. A new option is defined for the RA to carry learned (but unused) prefixes and new flags are required to indicate completeness and participation in probabilistic fast RA.

G6 Only link-scope signalling is used.

G7 SeND can be used to protect the RAs. The new option can be protected against tampering but not necessarily that they are authorized to be included in the RA. Since they are only used for link identification, this is no different to the flag protection in the previous section. It will also protect the new flags in the RA against tampering.

G8 If SeND is not deployed, then a rogue device could cause a host to think its configuration is invalid by sending an RA with bogus prefixes. 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 hosts and routers.

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.

9.3 Prefix-based LinkID with Fast Router Discovery

-(How to choose a common Prefix LinkID on a link)
- Routers listen to other routers' advertisements and use them to maintain a list of all active prefixes on the link.
  - Upper bound needed on list length to prevent memory overflow.
- The routers choose the smallest prefix as the Link Identifier, i.e. Prefix LinkID.

(How to advertise the Prefix LinkID in an RA)
- The routers advertise the prefix in every RA with PIO including a new "I" (identification) bit to indicate that the prefix is the Link Identifier, i.e. Prefix LinkID.
  - If the prefix is not explicitly configured on the sending router, the L, A and R flags should be set off, so that the PIO would have no effect on hosts other than link identification.

(How hosts use Prefix LinkID)
- Hosts keep the Prefix LinkID of the currently attached link.

(How to quickly forward Prefix LinkID to hosts with FRD)
- Access Points on the network cache an RA with the Prefix LinkID.

- When an Access Point detects (through layer two means) that a host has arrived on the link, it immediately forward it a copy of the cached RA.

(How a host checks for link change with Prefix LinkID)

- The host receiving the RA compares the Prefix LinkID in the RA to its currently stored one.
  - If they are the same, the host remains at the same link and no further DNA action is required.
  - If they differ, the host assumes a link change and immediately initiates a new IP configuration.

(How to interoperate with non-supporting links)

- Prefix LinkID scheme allows a host to detect a link change properly when it moves FROM and TO the link supporting the scheme. Backup mechanism such as CPL is needed only when a host moves between non-supporting links.

9.3.1 How the goals are met

G1 The reception of the LinkID gives a host a positive indication of whether link change has occurred, and the RA will contain the information required to reconfigure if necessary.

G2 The router advertisement is transmitted to the host as soon as the AP detects that it has associated and is able to receive packets on the link.

G3 Non-movement situations are correctly detected.

G4 Only a single RA is required.

G5 Only RAs are required. A new flag is added to PIOs to indicate that a prefix is in fact the Link ID as well.

G6 Only link-scope signalling is used.

G7 SeND can be used to protect RAs and show authorization for a set of prefixes. For routers with the prefix used as the LinkID explicitly configured, SeND may not show authorization. In this case there will be no evidence that the LinkID is valid. Hosts should only accept RAs that contain another authorized prefix. It will also protect the new flag in the RA against tampering.

G8 If SeND is not deployed, then a rogue device could cause a host to think its configuration is invalid by sending an RA with a bogus Link ID. 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 hosts and routers.

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.

10. On the Wire Costs

The number of bytes sent onto the wire (air) is highly dependent on the number of routers on a link and the way in which prefixes are distributed across them. In the very simplest case where there is only one router and it only has a single prefix to advertise, the variation in costs is quite small. Considering only unicast RA examples, the RS/RA would take 160 octets for CompleteRA, or for LinkID where the link identifier is the prefix being advertised. Using the hybrid landmark scheme would take 184 octets.

As the topology gets more complex and there are more routers and/or prefixes the number of octets in the exchange increases dramatically. In general, however, the growth is fairly consistent across the combinations of techniques. The exception is combinations including CompleteRA. The re-advertising of prefixes makes the size of its exchanges grow much more quickly if there are non-matching sets of prefixes on routers. For example, a medium case where there are two routers each with one prefix (but not the same one), the prefix-based requested landmark scheme takes 280 octets for the exchange. Complete RA takes 328 octets. In a much worse case of four routers, each with two prefixes and none matching, the two exchanges are 616 and 1240 octets respectively.

The worst case performance of Complete RA can be improved substantially by defining a new RA option to carry all of the re-advertised 64-bit prefixes at once. This reduces the above case exchange to 824 octets, but it is still unbounded. It needs to be considered how likely such topologies are.

The actual sizes of RAs will depend on which options are needed but an example of the sizes is shown in the following table. In this case "typical" options counted are Maximum Transmission Unit (MTU) and Link Layer Address Option (LLAO).

Technique	RS size	RA size
Random LinkID	56 octets (basic + 8 for TSLLAO)	40 + 48 + p*32 octets (basic + 8 (LLAO) + 8 (MTU) + 16 (LinkID) + PIOs)
Prefix LinkID	56 octets (basic + 8 for TSLLAO)	40 + 48 + p*32 octets as above _OR_ 40 + 32 + p*32 octets if one of the prefixes is the link ID
CompleteRA	56 octets (basic + 8 for TSLLAO)	40 + 32 + P*32 octets (basic + 8 (LLAO) + 8 (MTU) + all PIOs)
CompleteRA with re-advertised prefix option	56 octets (basic + 8 for TSLLAO)	40 + 32 + p*32 + 8 + p2*32 (basic + 8 (LLAO) + 8 (MTU) + own PIOs + opt header learned prefixes
Requested Landmark	72 octets (basic + TSLLAO + 16 octet landmark option)	40 + 32 + p*32 octets (basic + 8 (LLAO) + 8 (MTU) + PIOs
Priority Landmark	80 octets (basic + TSLLAO + 24 octet landmark option)	40 + 32 + p*32 octets as above
Hybrid Landmark	80 octets (basic + TSLLAO + 24 octet landmark option)	40 + 32 + p*32 octets as above

p = number of prefixes router advertises
P = total number of prefixes on link
p2 = number of prefixes re-advertised in case of CompleteRA

Note that unicast RAs have been assumed here necessitating the TSLLAO in the RS if an immediate RA is desired.

Note that RA size assumes that flags can be placed in existing RA flag fields - if an option is required the RA will be 8 octets larger.

Note also that the CompleteRA and LinkID techniques could have value even without an RS at all.

As mentioned above, the number of routers on a link and the distribution of prefixes has a large effect on the number and size of packets sent onto the link. Some examples are shown below.

Technique	1 Router 1 prefix	2 Router 1 prefix	2 Router 1 prefix each	2 Router 2 prefix disjoint	4 Router 2 prefix disjoint
Random LinkID	56+120 =176	56+2*120 =296	56+2*120 =296	56+2*152 =360	56+4*152 =664
Prefix LinkID	56+104 =160	56+2*104 =264	56+104+ 120=280	56+136+ 152=344	56+136+ 3*152 =648
CompleteRA	56+104 =160	56+2*104 =264	56+2*136 =328	56+2*200 =456	56+4*296 =1240
CompleteRA with re-advertised prefix option	56+104 =160	56+2*104 =264	56+2*120 =296	56+2*160 =376	56+4*192 =824
Requested Landmark	72+104 =176	72+2*104 =280	72+2*104 =280	72+2*136 =344	72+4*136 =616
Priority Landmark	80+104 =184	80+2*104 =288	80+2*104 =288	80+2*136 =352	80+4*136 =624
Hybrid Landmark	80+104 =184	80+2*104 =288	80+2*104 =288	80+2*136 =352	80+4*136 =624

Note this table assumes that for each RS there will be N RAs, where N is the number of routers on the link. It may be possible to multicast any delayed RAs and if a group of RSs arrive very close together, to have one RA answer multiple RSs.
If the first RA is not delayed, then #RAs is always >= #RSs and in general, #RAs = N * #RSs.

11. IANA Considerations

No new options or messages are defined in this document.

12. Security Considerations

All of the techniques described in this document are modifications to router discovery. SeND [12] has be design for the express purpose of securing neighbour and router discovery exchanges. It follows then that there are two cases to consider: networks with and without SeND.

SeND can be used to show that a router is authorised to advertise particular prefixes. This can be used equally well by routers checking the prefixes of other routers as it can by hosts checking the same. In the case of link identifiers SeND may not be able to show that a linkid is correct for a given router but it can protect the packet against tampering.

There may be some performance issues introduced by SeND. The first time a host comes to a link there may need to be a packet exchange to get certificates chains. This may be mitigated by allowing certificates to cover larger prefixes, e.g. for a site/organization.

Where SeND is not deployed there are many attacks against neighbour/ router discovery already possible and it is just necessary to investigate whether proposed DNA techniques make the network or hosts any more vulnerable than they already are.

The main difference between the threat to RFC2461 devices and the threat to devices implementing techniques discussed in this document comes from the desire to speed things up. The goal is to have a single RA able to give enough information to decide if a link change has occurred, and if so, reconfigure addressing, etc., to allow packet exchanges to begin on the new link. A result of this is that if a single carefully crafted packet can cause a host to make the decision that it has changed links, it can then cause that host enter a state where logically all of its existing configuration is invalid. If a host has in fact moved to a new link, then that configuration is invalid. It be prudent, however, to move the configuration to a placeholder in case it is possible to recognise to false advertisement and restore the old configuration.

12.1 General Threats

1 A bogus router advertises a landmark or identifier that convinces a host that it has moved when it in fact not.

2 A bogus router advertises a landmark or identifier that convinces a host that it has not moved when it in fact has.

3 A mischievous host may, depending on the mechanisms available in the fast RA scheme employed, be able to cause a flood of RAs to be sent onto the link. Even unicast RAs can cause disruption to all nodes on certain link types, such as those employing CSMA/CA like 802.11b. It is probably worth designing in mechanisms to limit the effect of this even when SeND is not employed because of the potential for a multiplicative effect where there are more than one router on the link: 1 RS -> N RAs.

13. Acknowledgments

Thanks to all members of the design team - JinHyeock Choi, Tero Kauppinen, James Kempf, Sathya Narayanan, and Erik Nordmark - upon whose discussion the text of this document is based, and for their help in shaping the content.

Thanks also to Greg Daley for some very useful insight.

14. References

14.1 Normative References

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

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

14.2 Informative References

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

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

[5] Choi, J., "Fast Router Discovery with RA Caching", draft-jinchoi-dna-frd-00 (work in progress), July 2004.

[6] Kempf, J., Khalil, M. and B. Pentland, "IPv6 Fast Router Advertisement", draft-mkhalil-ipv6-fastra-05 (work in progress), July 2004.

[7] Daley, G., "Deterministic Fast Router Advertisement Configuration", draft-daley-dna-det-fastra-01 (work in progress), October 2004.

[8] Choi, J., "DNA with unmodified routers: Prefix list based approach", draft-jinchoi-dna-cpl-01 (work in progress), October 2004.

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

[10] Daley, G., Narayanan, S. and G. Perkins, "A probabilistic scheme for fast Router Advertisement responses in IPv6", draft-daley-dna-prob-fastra-00 (work in progress), February 2005.

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

[12] 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.

[13] Yegin, A., "Link-layer Event Notifications for Detecting Network Attachments", draft-ietf-dna-link-information-00 (work in progress), September 2004.

Author's Address

Brett Pentland
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

Intellectual Property Statement

Disclaimer of Validity

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Copyright Statement

Acknowledgment

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