A probabilistic scheme for fast Router Advertisement responses in IPv6
draft-daley-dna-prob-fastra-00.txt

Status of this Memo

By submitting this Internet-Draft, I certify that any applicable patent or other IPR claims of which I am aware have been disclosed, and any of which I become aware will be disclosed, in accordance with RFC 3668.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

This Internet-Draft will expire on August 11, 2005.

Copyright Notice

Copyright (C) The Internet Society (2005). All Rights Reserved.

Abstract

The IPv6 Neighbour Discovery RFC provides for random delays of between 0 to 500 milliseconds, in order to spread responses to solicitations when multiple routers exist on a link. This document describes an alternative scheme for responding to Router Solicitations which estimates the number of routers on the link, and aims to reduce the delay spread for responding Router Advertisements.
# Table of Contents

1. Introduction ............................................... 3
2. Probabilistic Fast Router Advertisement ............. 3
3. Router estimation ........................................... 3
4. Delay calculation ........................................... 4
5. Discussion of heuristic .................................. 5
6. Rate limiting responses .................................. 5
7. Message Format ............................................ 5
8. Configuration ............................................... 6
9. Protocol constants ........................................ 6
10. IANA Considerations ..................................... 7
11. Security Considerations ................................. 7
12. Acknowledgments ......................................... 8
13. References ................................................ 8
13.1 Normative References ................................ 8
13.2 Informative References ................................. 8
A. Authors’ Addresses ....................................... 8

Intellectual Property and Copyright Statements ........ 10
1. Introduction

Responding to Router Solicitations quickly allows a router to advertise up-to-date configuration information to hosts which are either discovering routers or checking their current configuration. The IPv6 Neighbour Discovery RFC [2] requires a random delay of between 0 and 500ms before responding to a Router Solicitation (RS) with a Router Advertisement (RA).

Simultaneous transmissions of Router Advertisements by multiple routers may cause MAC congestion, or even packet loss in some environments. The random delay described in RFC 2461 prevents routers from continually sending RAs at the same time, and therefore causing congestion.

While some delay is needed to prevent continual collision by responding routers, the delays caused by the existing scheme induce a mean 250ms delay before responses are received.

An alternative scheme described in this document allows for a more constrained spread of delays by estimating the number of routers on the link, and randomly choosing a transmission delay within a small range of values.

2. Probabilistic Fast Router Advertisement

Probabilistic Fast Router Advertisement attempts to provide sufficiently fast Router Solicitation responses based only on a rough estimate of the number of routers on a link.

Routers do not coordinate their responses, nor know anything about peers, other than that another (purported) router is advertising.

This allows routers which do not know or trust each other to make choices about the router occupancy on the list with bounded worst case performance, and good usual case performance.

3. Router estimation

On each link a router is likely to receive periodic Router Advertisements from most or all of the other routers on a link. While routers do not use RAs for configuration, they may be used to provide an indication of the number of routers occupying a link.

This router estimate is used by Probabilistic FastRA in order to calculate RS response delays when scheduling Router Advertisement transmission.
In order to provide an accurate estimate, Fast Routers MUST set the Probabilistic Fast Routers Flag in the Router Advertisement header as described in Section 7. Probabilistic Fast Routers SHOULD only count routers advertising this flag in their estimate of fast routers.

Routers inspecting Router Advertisements SHOULD store the link-local source address, advertisement intervals and last transmission time of a (bounded) number of routers on the link. The number of routers so stored provides the basis of the estimate used in scheduling responses, with routers including themselves in the estimate.

Routers always consider themselves to be present, even if they do not receive their own advertisements.

In order to provide robustness against unknown routers and spurious router entries, the router bounds its estimate between MIN_PROB_ROUTERS and MAX_PROB_ROUTERS. The delay so calculated is stored in the variable RouterEst.

This ensures that a Probabilistic FastRA router always acts as if there was at least one other router on the link, even if it has seen no other routers.

Also, in order to facilitate estimation by other routers, a fast router SHOULD advertise periodically, and MAY include an Advertisement Interval Option in its multicast advertisement [3].

4. Delay calculation

When a router receives an RS it processes it in the way described in section 6.2.6 of the Neighbour Discovery RFC [3]. Instead of randomly delaying for 0 to MAX_RA_DELAY_TIME, though, the router may use the algorithm described in this section.

As described above RouterEst is the estimate of the number of routers on a link. RouterEst is in the range: [MIN_PROB_ROUTERS, MAX_PROB_ROUTERS].

Probabilistic FastRA requires routers to delay by choosing one of RouterEst + 1 slots, in the range from [0, RouterEst], where the delay duration is:

slotNumber * ProbFastRASlotTime

Slot choices are based on probabilities, which depend on RouterEst. The first slot (slot 0) is given greater weight, and all other slots equal weight from the remaining probability:
P(slot_0) = 1/RouterEst
P(slot_i) = (RouterEst - 1)/(RouterEst^2), for all i \[1,RouterEst\]

5. Discussion of heuristic

The algorithm presented above provides some random delay, which is tied to the link’s current router occupancy. This limits the effect of collisions, while providing bounded response scheduling delays.

Selection of zero delay is preferred, particularly with low estimates of routers (50% if only one or two routers). This is deliberate, since it allows a fast mean response time, but allows re-ordering upon retransmission if slot collisions and consequent MAC collision or contention occur.

When only one router operates, the response time maximum is 2 slot times (default 40ms) with 25% probability, the expected delay is 0.75 slot times (15ms). Where two routers operate, the expected delay reduces to 0.3125 slot times (6.25ms). With two routers though, there is a 37.5% chance of a slot collision, where both routers choose the same slot.

For larger numbers of routers, there is the likelihood of at least one slot collision. Particularly for odd-numbered estimates of routers, there is a high likelihood that (at least) one of the routers will pick a slot which is not chosen by other routers. These singleton responses may be successfully transmitted in some environments even if slot collisions from other routers’ transmissions produce garbled, lost or delayed RA reception on other computers.

The estimates provided for router numbers are bounded for this algorithm. This may make it inappropriate on links with many advertising routers.

6. Rate limiting responses

It is necessary to limit the number of fast Router Advertisements sent by a router within an interval. It is therefore RECOMMENDED that hosts use a token bucket scheme to allow not allow more than ProbFastRABucket Router advertisements to be outstanding, with transmission tokens being replenished once every ProbFastRATokInterval.

7. Message Format

This document specifies a new flag to be sent in Router
Advertisements in order to identify members of the Fast Router set.

This is described below:

```
<table>
<thead>
<tr>
<th>Type</th>
<th>Code</th>
<th>Checksum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cur Hop Limit</td>
<td>M</td>
<td>O</td>
</tr>
<tr>
<td>Router Lifetime</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reachable Time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Retrans Timer</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Options ...
```

The new field "F" describes the Probabilistic Fast Router Flag. This flag is set by routers participating in Probabilistic FastRA.

8. Configuration

Probabilistic FastRA runs per-interface, and its use SHOULD be configurable on a per-interface basis. All configuration variables listed below SHOULD be configurable on a per-interface basis.

The configuration parameter ProbFastRASlotTime is described in this document. It provides the delay between adjacent scheduling slots for responding router advertisements. This parameter SHOULD default to PROB_FASTRA_SLOT_TIME. The default for this parameter MAY vary based on the class of interface, and the transmission medium.

ProbFastRABucket provides the maximum number of fast Router Advertisements which may be transmitted at any time. It should be a positive integer, if Probabilistic FastRA is in use. It defaults to PROB_FASTRA_BUCKET.

ProbFastRATokInterval number of milliseconds defining the time it takes to increase the number of tokens by one. It defaults to PROB_FASTRA_TOKEN_INTERVAL. For intervals under a second, a router MAY group intervals together and add multiple tokens to the bucket each time the longer interval’s timer fires.

9. Protocol constants

MIN_PROB_ROUTERS: 2
IANA Considerations

A protocol flag (1 bit - suggest 0x04) in the Router Advertisement’s "reserved and flags" octet is required to be allocated to support Probabilistic FastRA identification.

Security Considerations

Probabilistic FastRA does not vouch for the validity of Router Advertisements it receives, nor does it check their security.

Each router up to the fourth (fifth including itself) has an impact on the advertising state of the other routers on the link, but cannot force advertisements at a particular time, nor can they predict which of the slots will be used by another router.

The state required to store router information is thus bounded, and only four other routers need to be stored (information from other routers may be discarded).

Router information received may be spoofed, and recipient routers SHOULD be careful in handling Advertisement Intervals and lifetimes for routers. They SHOULD consider timing out routers based on the default advertised lifetime if no other measures are available (9000 seconds) [2]. Where Advertisement Interval options are used, robustness to packet loss is desirable, and the router should remain in the list for at least two intervals. This may extend the duration of an attack by spurious routers, but cannot cause a more severe delays without transmitting Router Advertisements itself.

The maximum expected duration before an RA is scheduled in this case is bounded at 2.4 slots (around 48ms), even when every other router is bogus. Where this is an acceptable delay, no other action needs to be taken.

Where two real Probabilistic FastRA routers exist on the link, the expected delay comes to around 1.4 slots (about 28 ms), even if all other routers are bogus.
12. Acknowledgments

13. References

13.1 Normative References


13.2 Informative References

Authors’ Addresses

Greg Daley
Centre for Telecommunications and Information Engineering
Department of Electrical and Computer Systems Engineering
Monash University
Clayton, Victoria 3800
Australia

Phone: +61 3 9905 4655
EMail: greg.daley@eng.monash.edu.au

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

Phone: +1 609 734 7599
EMail: sathya@Research.Panasonic.COM
URI:
Appendix A. Implementation status

This has been implemented under RADVD.
Intellectual Property Statement

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

The IETF has been notified of intellectual property rights claimed in regard to some or all of the specification contained in this document. For more information consult the online list of claimed rights.

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

Copyright (C) The Internet Society (2005). This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.
Acknowledgment

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