

SAVI  
Internet-Draft  
Intended status: Standards Track  
Expires: December 6, 2012

P. Thubert, Ed.  
Cisco Systems  
June 4, 2012

Throttling RAs on constrained interfaces  
draft-thubert-savi-ra-throttler-01

## Abstract

In a large switched topology with either many routers or routers that transmit a high rate of multicast advertisements per router, as suggested to support mobile nodes, the cost of distributing the many resulting multicast messages to certain classes of devices might be prohibitive. This is the case of a device that runs on batteries, or a device that is reachable over a wireless interface. For this device, it can be beneficial to filter a certain amount of multicast messages such as the Router Advertisement while preserving the functionality expected in the IPv6 Neighbor Discovery Protocol.

## Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC 2119](#) [[RFC2119](#)].

## Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

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

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

This Internet-Draft will expire on December 6, 2012.

## Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the

Internet-Draft

ra-throttler

June 2012

document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

<a href="#">1.</a>	Introduction . . . . .	<a href="#">3</a>
<a href="#">2.</a>	Terminology . . . . .	<a href="#">4</a>
<a href="#">3.</a>	Problem statement . . . . .	<a href="#">5</a>
<a href="#">3.1.</a>	Routers behavior . . . . .	<a href="#">5</a>
<a href="#">3.2.</a>	Wireless Mobility domain . . . . .	<a href="#">5</a>
<a href="#">4.</a>	Operation . . . . .	<a href="#">6</a>
<a href="#">4.1.</a>	Throttling scope and period . . . . .	<a href="#">7</a>
<a href="#">4.2.</a>	Pending Hosts List . . . . .	<a href="#">7</a>
<a href="#">4.3.</a>	Advertising Routers List . . . . .	<a href="#">8</a>
<a href="#">4.4.</a>	RA with an Advertisement Interval Option . . . . .	<a href="#">8</a>
<a href="#">4.5.</a>	Final RA . . . . .	<a href="#">9</a>
<a href="#">4.6.</a>	Throttling Policy . . . . .	<a href="#">10</a>
<a href="#">5.</a>	Manageability . . . . .	<a href="#">11</a>
<a href="#">6.</a>	IANA Considerations . . . . .	<a href="#">11</a>
<a href="#">7.</a>	Security Considerations . . . . .	<a href="#">11</a>
<a href="#">8.</a>	Acknowledgements . . . . .	<a href="#">11</a>
<a href="#">9.</a>	References . . . . .	<a href="#">11</a>
<a href="#">9.1.</a>	Normative References . . . . .	<a href="#">11</a>
<a href="#">9.2.</a>	Informative References . . . . .	<a href="#">11</a>
	Author's Address . . . . .	<a href="#">12</a>

## 1. Introduction

The protection of the network is not necessarily a security function such as the defense against a specific attack. It might also have to do with other activities that include the control of multicast storms, as provided to some extent by Multicast Listener Discovery (MLD) Snooping [[RFC4541](#)], or of the overuse of the network that might degrade the service for all users as provided for instance by Call Admission Control [[RFC5865](#)].

In particular, the wireless edge of a large Layer 2 topology will require some special protections to conserve the limited bandwidth that is available over the radio medium, such protections certainly involving the reduction of multicast operations.

MLD Snooping helps control the impact of multicast messages but:

It does not apply to the all-nodes link-scoped multicast address (FF02::1) as defined in the IP Version 6 Addressing Architecture [[RFC4291](#)].

MLD snooping is generally not implemented for link scope multicast messages anyway.

If MLD snooping runs in instance as opposed to the Access Point (AP) and if there is at least one listener associated to the AP then the AP will still get the multicast and transmit it to all the devices that are associated to the AP.

MLD snooping has the granularity of a group as opposed to a binding table that has the granularity of a target - a host.

This document focusses on the protection against an excessive bandwidth consumption by multicast IPv6 Neighbor Discovery (ND) [[RFC4861](#)] Protocol (NDP) Router Advertisement (RA) messages over the wireless edge of a switched network.

RA messages are link-scoped messages that provide the recipient node with link information such as availability and characteristics of routers that are present on the link and a list of prefixes that are usable for IPv6 NDP Stateless Address Autoconfiguration (SLAAC) [[RFC4862](#)].

RA messages coming from different routers may carry different information, including information about the router itself, but also different prefix lists and other information such as the period of transmission [[RFC6275](#)].

Thubert

Expires December 6, 2012

[Page 3]

---

Internet-Draft

ra-throttler

June 2012

In a number of cases, the fact that a node misses an RA does not impact the node operation in a notable fashion, either because the information is fully redundant with information that the node already has (e.g. multiple RAs of a same content from a same router in rapid sequence) or because the information is not critical to the node (e.g. yet another router that is not preferable as default gateway). In other cases, the loss of an RA is eventually recovered, but the node will not operate optimally in the meantime and such a loss should be avoided.

This document studies situations when an Ethernet Switch, an IEEE 802.11 Access Point, or a Control And Provisioning of Wireless Access Points (CAPWAP) Access Controller (AC) [[RFC5415](#)] can, with no notable effect, make the decision not to copy an RA message onto a port or a set of ports, typically one or more ports that connect to an IEEE 802.11 wireless domain, the consequences of doing so and the eventual recovery. In the remainder of this document, the term throttling will refer to the decision not to copy a message over such ports, and the layer 2 device in charge of making that decision will simply be referred to as switch, whether it is a classical Ethernet switch or any one of the sorts of devices listed above.

## [2.](#) Terminology

The draft uses the following terminology:

**Switch:** A layer 2 device that distributes packets over one or more ports. This broad definition includes but is not limited to

Ethernet and IEEE802.3 switches, CAPWAP Access Controllers and IEEE 802.11 Access Points.

**Throttling:** The decision not to copy a given multicast message onto a given port or a set of ports after the determination that the RA would be redundant for most hosts across the port(s). A multicast packet that is throttled over a given port might still be copied for unicast delivery to selected hosts on a that port if it is determined that they will benefit from receiving the RA. The packet might still be passed on to other ports such as trunk ports, for further switching along a VLAN for instance.

**Throttled port:** A port on the switch where throttling is active. The port might be an access port that is directly connected to a host, but it might also be a multipoint port, for instance if it is connected to another switch such as an Access Point.

Thubert

Expires December 6, 2012

[Page 4]

---

Internet-Draft

ra-throttler

June 2012

### [3.](#) Problem statement

#### [3.1.](#) Routers behavior

Assuming that the solicitor's source address is not the unspecified address, a router may choose to respond to an ND Router solicitation (RS) with a unicast RA message directly to the soliciting host's address. But it is common that the router aggregates multiple requests and sends a single multicast response to the all-nodes group. This RA is received by all the nodes on link, though a host that did not issue an RS is probably not very interested in receiving the solicited RA response message. Yet, in a wireless environment, a host will usually issue an RS each time it reassociates, which can be quite often if the host is as mobile as a smartphone.

A traditional (wireline) router will typically not rate limit its RA emissions based on consistent RA messages received from other routers, though such a behavior is required in the Routing Protocol for Low Power and Lossy Networks (RPL) [[I-D.ietf-roll-rpl](#)] that is specifically designed for constrained environments such as wireless mesh networks. As a result, the number of RAs on a switched topology increase roughly linearly with the number of deployed routers.

As it goes, the whole RA operation denotes an implicit expectation that the cost for a multicast operation is not substantially different from that of a unicast transmission and that the cost is roughly similar on all segments of the link. Sadly, this is not true in the case of a composite network with a switched Ethernet backbone and an IEEE 802.11 Extended Service Set (ESS) wireless edge. The situation is even worse if the edge is a mesh network, e.g. an IEEE 802.11S or a Low Power Lossy Network as described in [\[I-D.phinney-roll-rpl-industrial-applicability\]](#) and [\[I-D.thubert-lowpan-backbone-router\]](#). In any case, a rate of RAs that might appear acceptable on the backbone can rapidly become excessive on the wireless edge.

### [3.2.](#) Wireless Mobility domain

A number of (layer 3) Network-Based Localized Mobility Management (NetLMM) techniques have been deployed that enable IP mobility transparently to the host, that is without requiring the active participation of the host in any mobility-related signaling. This is achieved by hiding its mobility to the host and more specifically by presenting the host with a consistent link appearance as it roams at layer 3, in particular through tailored RA messages. An example of such NetLMM solutions would be the adaption of Proxy Mobile IPv6 (PMIPv6) [\[RFC5213\]](#) within a proprietary mobility framework.

Thubert

Expires December 6, 2012

[Page 5]

---

Internet-Draft

ra-throttler

June 2012

As a result, within a same radio environment (say an IEEE 802.11 Service Set Identification or SSID), some of the associated nodes may be local nodes and some other nodes may be roaming devices that are virtually part of some other link or VLAN in a remote location. If all the associated nodes received a local RA announcing an local IPv6 prefix, roaming devices would detect their movement, form new addresses and defeat the mobility functionality to the point that the entire mobility domain would appear as one flat single IPv6 link.

To avoid that problem, a dedicated RA is unicast to each of the associated devices as opposed to sent once as a layer 2 broadcast to all devices in a single shot. A very common method consists in rewriting the link layer multicast address in a frame that carries the layer 3 multicast message onto a layer 2 unicast address. This isolates which L3 multicast packet gets to which host, and more

importantly which multicast packet will not get to a given host for which it is not destined.

When a multicast packet is converted into multiple unicast frames by a switch such as an Access Point or an Access Controller, a single packet that is sent to the all-nodes group can consume a large amount of bandwidth that is roughly a factor of the number of associated devices, and disrupt sensitive applications such as voice over IEEE 802.11. The NDP assumption that a multicast does not cost more than a unicast is severely broken. It results that the ND Protocol is not really suited for the wireless medium, and that some tailoring is required in instance to reduce the impact of the multicast messages, in particular RA messages.

#### [4.](#) Operation

The NDP messages Router Advertisements are scoped to a link. They are sent on a given IPv6 link (e.g. a Virtual Local Area Network) and should be delivered only to IPv6 nodes that reside on that link. An RA message can be transmitted over the medium either as unicast response or as a multicast message that is sent to the link-scoped all-nodes multicast address (FF02::1) as defined in [[RFC4291](#)], which all IPv6 nodes on the link listen to.

If a Router Advertisement is sent to a unicast destination address, instance MUST forward the packet to the destination device. But as opposed to other ND Protocol operations such as the Duplicate Address Detection (DAD) that occurs only when a node obtains or forms a new address, multicast RAs are sent periodically and might be quite frequent for the duration of the network activity. In that case, instance MAY drop the multicast RA if it is redundant. The question becomes to determine whether a multicast RA is redundant.

A switch might connect ports of different natures. Some ports may need throttling of the RA messages, and some node. It is expected that some mechanism is in place to determine which ports require throttling, for instance a configured policy or an automatic discovery.

##### [4.1.](#) Throttling scope and period

The scope of a throttling activity is a link (a VLAN). Within that scope, some ports on instance are determined to be throttled, while others are not. A throttling period is associated to that scope. A policy dictates how many and under which conditions multicast RAs are throttled. The policy is based on counters that count RAs per router and counters that aggregate the numbers to the throttling scope (the VLAN). The counters are reset at each throttling period.

NDP does not mandate that routers on a link expose the same prefixes. It is possible that a router advertises a prefix that none of the other routers does for instance. Or a router might advertise a better preference for a given destination [[RFC4291](#)]. It is this important that the throttling mechanism does not starve any given router. instance SHOULD attempt to distribute fairly the amount of RAs per source router, and to serve at least once any given router on the link (VLAN) within a given period of time.

#### [4.2.](#) Pending Hosts List

A multicast RA that is the response to an RS is probably redundant for all nodes that did not solicit the RA in the first place. But it is certainly useful to nodes that issued an RS over a throttled port since the last multicast RA happened. instance needs to keep track of all those hosts as discovered through their RS messages, in an abstract list referred to as the Pending Hosts List (PHL). There is one PHL per link (that is, typically, per VLAN).

A host is anchored to a port on instance, a link layer address, and eventually one or more VLAN identifier(s) depending on the deployment. An IPv6 Link Local Address [[RFC4291](#)] might be available to qualify the host. A PHL entry SHOULD contain all the anchor parameter and MAY indicate additional information such as the host Link Local Address.

instance SHOULD add a host to the PHL when it receives an RS from that host over a throttled port, and upon a layer 2 trigger that indicates that the port has flapped, typically an association or a reassociation event in an Access Point or an Accesss Controller. instance MAY remove a host from the PHL when a RA is forwarded to the host, either as a unicast, a multicast, or a unicast copy of a

throttled multicast, and SHOULD remove the entry after a number or



RAs are forwarded, depending on the policy that applies to the host.

#### 4.3. Advertising Routers List

The primary cause of RA redundancies is a router that sends multiple identical RAs in a short sequence, for instance as stimulated by hosts joining the link. instance identifies such redundancies by keeping track of all the routers as discovered through RA messages, and eventually of the content of those RA messages, in an abstract list referred to as the Advertising Routers List (ARL). There is one ARL per link (VLAN).

A router is anchored to a port on instance, a link layer address, and eventually one or more VLAN identifier(s) depending on the deployment. The router Link Local Address is found as the source address of the RA. A ARL entry SHOULD contain all the anchor parameters, the router Link Local Address, and a number of counters that indicate the router activity over the last period and MAY contain additional information from the RA such as a prefix list or the router preferences [[RFC4191](#)]. The entry MUST also contain counters that are necessary for the throttling operation, typically the number of multicast RAs that were copied and the number that were throttled during the current throttling period.

instance SHOULD add a router to the ARL when it receives an RA from that router on any port that belong to that link (VLAN). instance SHOULD remove the router from the ARL when the throttle period elapses. instance MAY maintain a list of routers that were part of the ARL for the previous period in an alternate list to keep additional history and improve runtime performances.

#### 4.4. RA with an Advertisement Interval Option

The Mobility Support in IPv6 (MIPv6) [[RFC4191](#)] [section 7.3](#) introduces the Advertisement Interval Option (AIO), used in RA messages to advertise the interval at which the advertising router sends unsolicited multicast Router Advertisements. When this option is present, a switch SHOULD NOT interfere with a routers attempt to live up to its claim that at least one RA message will be posted every advertisement interval.

There is more than one way for instance to comply with this requirement, as controlled by a policy that applies to the throttling operation:

instance MAY never copy RAs from a given router that carry the AIO over throttled ports.

Or instance MAY copy all RAs from a given router that carry the AIO over throttled ports.

Alternatively, instance MAY monitor the timing of RA emissions from a given router and refrain from throttling at least one RA per advertisement interval from that router. It might then happen that the router arms its timer on a message that instance throttles later. In that case, the next RA that is not throttled can be separated by substantially more time than one advertisement interval though less than 2 intervals. This should not impact the MIPv6 operation that does not take action until no RA is received within two and a half advertisement intervals.

It can be noted that the advertisement interval that is used to support mobility can be very short and load the radio medium quite dramatically, depending on the available bandwidth on that medium. The policy in place SHOULD probably make it so that RAs with too short intervals are not copied on throttled ports unless no other option is available. If mobile devices are expected on the wireless link, then it might be preferred to block all routers advertising AIO but one or two that would preferably use an acceptable interval.

#### 4.5. Final RA

[Section 6.2.5](#) of the Neighbor Discovery specification [[RFC4861](#)] describe the router operation when it ceases to advertise on a given interface. In particular, the router needs to transmit one or more final (multicast) RA messages on the interface with a Router Lifetime field of zero.

This information is critical to any host that utilizes the router either as default gateway or more preferred gateway for a given destination prefix since filtering out a final RA might leave such host without connectivity till the host discovers that the router is gone. A switch SHOULD NOT take actions that would prevent such a host from receiving at least one final RA that indicates that a given router ceases to be available as a IPv6 gateway on the link (VLAN) where throttling applies.

There is more than one way for instance to comply with that requirement, as controlled by a policy that applies to the throttling operation:

instance MAY never throttle an RA with a Router Lifetime field set to zero.

Alternatively, instance MAY throttle further multicast final RAs arrive immediately after a first final RA from a same router.

It can be noted that the advertisement interval that is used to support mobility can be very short and load the system quite dramatically. The policy in place should probably make it so that RAs with short intervals are not copied on throttled ports unless no other option is available.

#### [4.6.](#) Throttling Policy

An implementation SHOULD allow to configure a policy whereby the RA throttling operation is based on the history of received RAs during the current throttling period.

Suggested policy parameters per link (VLAN) include:

**throttle-period:** This is the duration of the throttling period. A suggestion is to keep this value under the highest `MaxRtrAdvInterval` used in the network. `MaxRtrAdvInterval` is defined in [[RFC4861](#)] with a default of 600 seconds. The policy that provides that parameter MAY apply to the link (VLAN) or instance.

**max-through:** This is a maximum number of RAs that may pass before for all routers during a throttling period. `rAdvInterval` is defined in [[RFC4861](#)] with a default of 600 seconds. The policy that provides that parameter MAY apply to the link (VLAN) or instance. A suggested default is 1.

**at-least:** This is the minimum guaranteed number of RAs that pass before the first RA is throttled for a given router. The policy that provides that parameter MAY apply to an individual router, a port, the link (VLAN) or instance. A suggested default is 1. This parameter takes precedence over the `max-through` parameter that is defined at the link (VLAN) level so as not to starve any router.

**at-most:** This is a maximum number of RAs that may pass before for a given router during a throttling period. The policy that provides that parameter MAY apply to the router, the port, the link (VLAN) or instance. A suggested default is 1.

interval-option: This parameter indicates the behaviour upon RAs with the IA0 as discussed in [Section 4.4](#). The policy that provides that parameter MAY apply to the router, the port, the link (VLAN) or instance. A suggested default is never to copy RAs with IA0 on a throttled port.

Thubert

Expires December 6, 2012

[Page 10]

---

Internet-Draft

ra-throttler

June 2012

## [5.](#) Manageability

An implementation SHOULD allow the administrator to define one or more throttling policies and to apply them on the relevant targets (routers, ports, links and switch). The implementation should count the number of RAs that passed and RAs that are throttled per target.

## [6.](#) IANA Considerations

This specification does not require IANA action.

## [7.](#) Security Considerations

This specification is not found to introduce new security threat.

## [8.](#) Acknowledgements

The author wishes to thank Eric Levy-Abegnoli for his kind mentorship all along this project.

## [9.](#) References

### [9.1.](#) Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing

Architecture", [RFC 4291](#), February 2006.

[RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", [RFC 4861](#), September 2007.

[RFC4862] Thomson, S., Narten, T., and T. Jinmei, "IPv6 Stateless Address Autoconfiguration", [RFC 4862](#), September 2007.

## [9.2.](#) Informative References

[I-D.ietf-roll-rpl]

Brandt, A., Vasseur, J., Hui, J., Pister, K., Thubert, P.,  
Levis, P., Struik, R., Kelsey, R., Clausen, T., and T.  
Winter, "RPL: IPv6 Routing Protocol for Low power and  
Lossy Networks", [draft-ietf-roll-rpl-19](#) (work in

---

Thubert Expires December 6, 2012 [Page 11]

---

Internet-Draft ra-throttler June 2012

progress), March 2011.

[I-D.phinney-roll-rpl-industrial-applicability]

Phinney, T., Thubert, P., and R. Assimiti, "RPL  
applicability in industrial networks",  
[draft-phinney-roll-rpl-industrial-applicability-00](#) (work  
in progress), October 2011.

[I-D.thubert-lowpan-backbone-router]

Thubert, P., "LoWPAN Backbone Router",  
[draft-thubert-lowpan-backbone-router-00](#) (work in  
progress), November 2007.

[RFC4191] Draves, R. and D. Thaler, "Default Router Preferences and  
More-Specific Routes", [RFC 4191](#), November 2005.

[RFC4541] Christensen, M., Kimball, K., and F. Solensky,  
"Considerations for Internet Group Management Protocol  
(IGMP) and Multicast Listener Discovery (MLD) Snooping  
Switches", [RFC 4541](#), May 2006.

[RFC5213] Gundavelli, S., Leung, K., Devarapalli, V., Chowdhury, K.,  
and B. Patil, "Proxy Mobile IPv6", [RFC 5213](#), August 2008.

[RFC5415] Calhoun, P., Montemurro, M., and D. Stanley, "Control And

Provisioning of Wireless Access Points (CAPWAP) Protocol Specification", [RFC 5415](#), March 2009.

[RFC5865] Baker, F., Polk, J., and M. Dolly, "A Differentiated Services Code Point (DSCP) for Capacity-Admitted Traffic", [RFC 5865](#), May 2010.

[RFC6275] Perkins, C., Johnson, D., and J. Arkko, "Mobility Support in IPv6", [RFC 6275](#), July 2011.

Thubert

Expires December 6, 2012

[Page 12]

---

Internet-Draft

ra-throttler

June 2012

#### Author's Address

Pascal Thubert (editor)  
Cisco Systems  
Village d'Entreprises Green Side  
400, Avenue de Roumanille  
Batiment T3  
Biot - Sophia Antipolis 06410  
FRANCE

Phone: +33 497 23 26 34  
Email: [pthubert@cisco.com](mailto:pthubert@cisco.com)

