Network Working Group
Internet-Draft
Intended status: Informational
Expires: December 22, 2011

W. Dec
Cisco Systems

June 20, 2011

IPv6 Router Solicitation Driven Access Considered Harmful
draft-dec-6man-rs-access-harmful-00

Abstract

This document presents issues regarding the reliance of IPv6 Router Solicitation messages for creating or initializing router state necessary to enable IPv6 users' connectivity, particularly in situations where such users have bridged ethernet connectivity with the router. A number of alternative solution approaches are also presented and discussed.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD 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 22, 2011.

Copyright Notice

Table of Contents

1. Introduction

Recent proposals for including subscriber line identifiers alongside host sourced Router Solicitation (RS) messages ([I-D.ietf-6man-lineid]) in an environment where the host has no direct link layer adjacency with the router (eg when using Ethernet bridging), have highlighted the intent of using these RS messages on the receiving router as a trigger for specific functions & processes. Without the execution of these processes, such as host or line authorization, the host will not receive Router Advertisements (RAs) that allow the establishment of full IPv6 connectivity. Similar RS triggered processes, although without line identifiers, are proposed in specifications concerning WiFi access and appear to share the same pitfalls.

In analyzing the impact of these proposals it is useful to refer to the basics of the IPv6 Neighbour Discovery protocol as defined in [RFC4861], which defines the Router Solicitation (RS) message type. This message is intended to be used by hosts to request routers to generate Router Advertisements sooner than at their next scheduled time. The Router Solicitation mechanism is intended to be used in a very specific set of cases, or not at all, and a regular IPv6 network can work fully without any RS message ever being sent. In general, as per Section 6.3.7 of [RFC4861], Router Solicitations may be sent by a host after any of the following events:

- The interface is initialized at system startup time.
- The interface is reinitialized after a temporary interface failure or after being temporarily disabled by system management.
- The system changes from being a router to being a host, by having its IP forwarding capability turned off by system management.
- The host attaches to a link for the first time.

o The host re-attaches to a link after being detached for some time.

Notably in the above a host is at no stage required to periodically send RS messages, nor to send RS messages after a period of not receiving any RAs.

Furthermore [RFC4861] states that once a host "receives a valid Router Advertisement with a non-zero Router Lifetime, the host MUST desist from sending additional solicitations on that interface, until the next time one of the above events occurs." This effectively signifies that following the reception of any given RA message, sent by any device, a host will not issue RS messages until it is

reattached or re-initialized.

The following text from [RFC4861] also illustrates another aspect relating to the rule governing a host's ceasing of RS sending.

"If a host sends MAX_RTR_SOLICITATIONS solicitations, and receives no Router Advertisements after having waited MAX_RTR_SOLICITATION_DELAY seconds after sending the last solicitation, the host concludes that there are no routers on the link"

Experimental evidence conducted on a number of IPv6 implementations confirms that the above behaviour is indeed currently the norm, with specific implementations differing in terms of the default timers (eg MAX_RTR_SOLICITATION_DELAY) used. One implementation has been found to send RS messages at evenly spaced 4 second intervals for up to 12 seconds after the link event. Another implementation has been found to exponentially increase the sending interval for successive messages and stopping RS sending after 90 seconds.

The RS sending mechanism was thus clearly not designed nor is implemented to be periodic, nor reliable, nor expected to be sent by a host that has timed out or received an RA. Any mechanism that presupposes any of these RS sending characteristics, or requires them to work reliably, requires a thorough review.

2. Problem Overview

The main intent of the [I-D.ietf-6man-lineid] proposal is to convey

from an Ethernet bridging Access-Node to an upstream IPv6 router, the subscriber-line-id information indicating the origin of downstream host sourced RS messages. All this is envisaged to be done by tunneling such RS messages using IPinIP tunneling between the Access-Node and the Router, with the access node inserting the subscriber-line-id for each tunnelled RS. The reception by the router of such RS messages with the subscriber-line-id is expected to be the trigger for authorizing and allowing the subscriber's connectivity to the network. It is crucial to note that only after successful authorization will the router send RA messages that contain IPv6 Prefix Information Option (PIO) that allow the host to configure a global IPv6 address. A direct example of this usage goal can be found in Section 6.5 and Appendix A of [TR-177].

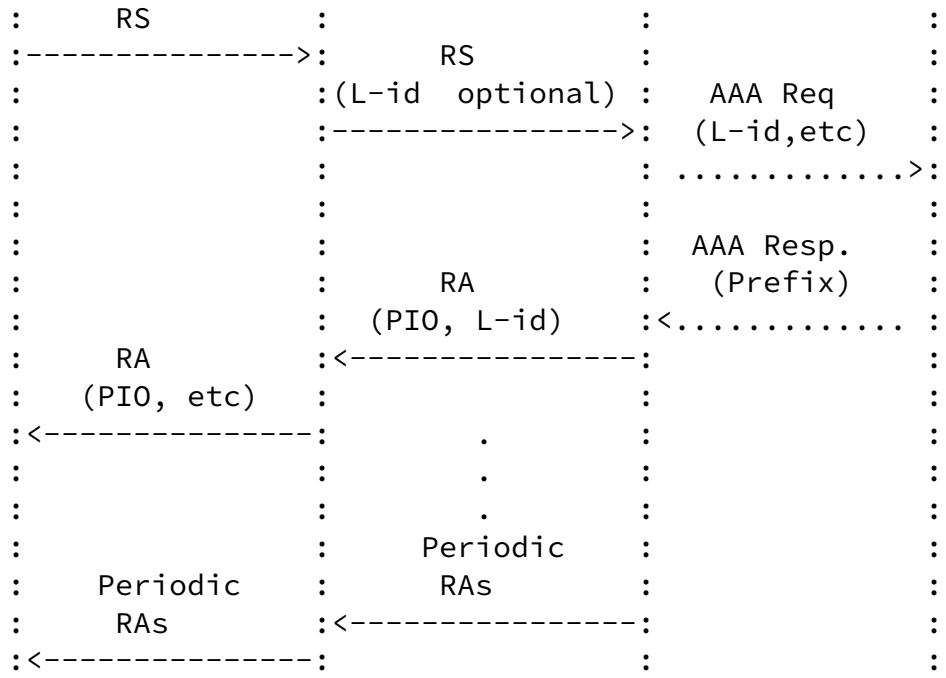
In generic terms, the principle of such mechanism is shown in Figure 1, and the goal is to create a dynamic user driven IPv6 access system that is in conductive to:

a. Triggering by means of subscriber sourced ND (RS) messages, processes on the IP edge router which serve to provide and setup hosts/subscribers with IPv6 connectivity.

b. Deriving from the received messages host identifiers and/or information regarding where the host is connected to in the Layer 2 network (eg based on MAC address and/or subscriber line id) and using that information in performing access and/or address authorization prior to granting connectivity.

c. Being used in an environment where the host/subscriber has no directly link layer adjacency with the router, but rather indirect connectivity (eg via a bridged Ethernet RG/CPE, and/or a bridging DSLAM).

d. Being used in an environment where IPv6 hosts implement *only* [RFC4861] as the control protocol, and without any further host changes or client protocols (eg DHCPv6)



A number of deployment contexts that seek to realize such a system will result in the end user having no IPv6 connectivity, and being left without any automated means of recovery it, all very detrimental to the success of the IPv6 deployment.

One such deployment context is the residential broadband N:1 VLAN environment, as described by [I-D.ietf-6man-lineid]. This features hosts indirectly connected to the edge router over a bridged Layer 2 VLAN set-up (aka an N:1 VLAN). End subscriber hosts connect to Ethernet bridging devices, such as an RG/modem and an Access-Node/ DSLAM, which provide indirect link connectivity for the host with the router. From each end hosts perspective, its local LAN link state is as presented by the RG/modem's LAN interface, eg Ethernet or WiFi. This state is decoupled from the RG/modem's uplink interface state, or that of the DSLAM links, or that of the IP edge router interface(s). Hence, each host's interface is expected to be "up" even when no DSL WAN link synchronisation has been established, or when the WAN link is being established following a modem reboot (an event lasting 2 minutes or more is not uncommon), etc. Given this, and in consideration of the RS sending characteristics described in

Section 1 , it is near certain that following a bridge/modem reload, or a DSLAM reload, any and all RS messages sent by hosts will never arrive at the intended IP edge router within the time hosts send RS messages. Since the reception of such RS messages by the edge router is required to trigger the announcement of RAs containing the chosen user address prefix option (PIO) towards the hosts, the host will be left without any addressing information and thus no IPv6 connectivity. The only recourse a user has is manual intervention on the host's interface.

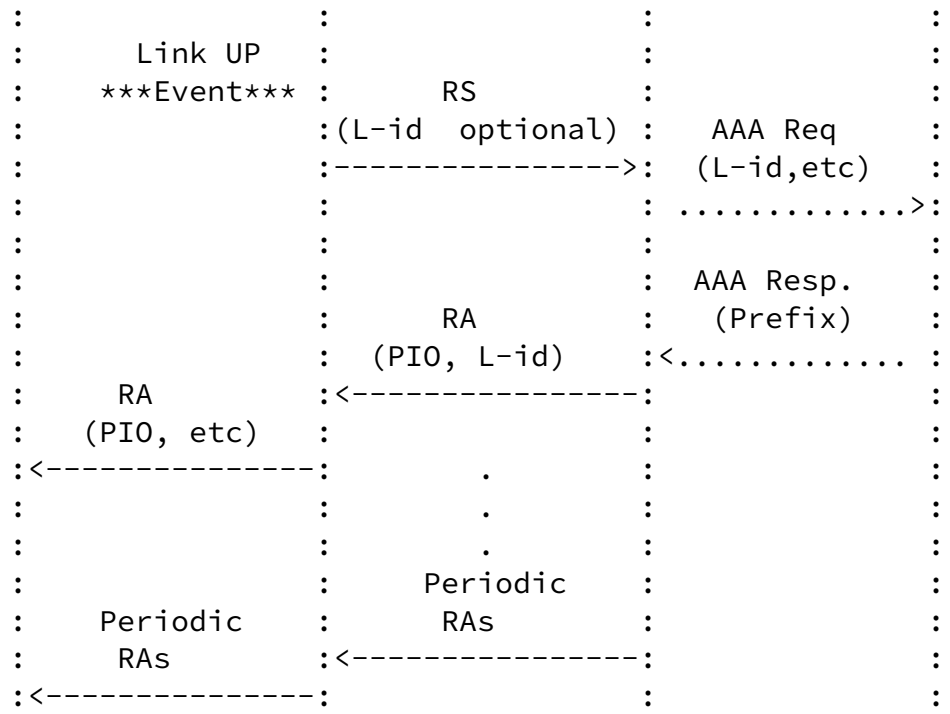
Note: The example of DSL is used above, but the case applies to other media, eg cable modems, that exhibit similar "modem reload" events. Moreover, the same problem appears to apply to each deployment that seeks to realize the mentioned goals and features hosts that have no direct link layer adjacency with a router, eg IEEE 802.11 WiFi architectures.

It's significant to note that the [I-D.ietf-6man-lineid] mechanism implicitly assumes behaviour which by itself will result in the system failing non-deterministically. As expemplieifed by its usage described in [TR-177], "empty RAs" (ie Router-Advertisement messages that contain no addressing/prefix information) are to be multicast to all subscribers and hosts from the router, in parallel to any specific RAs containing prefix information and the line option. Again, following the cited rules of [RFC4861], should a subscriber host receive such an empty RA prior to issuing an RS, that host will never send an RS and thus never trigger the authorization process necessary to get global IPv6 addressing & connectivity.

2.1. RS Sending Proxy

An update to the [I-D.ietf-6man-lineid] draft proposal has somewhat recognized the critical flaw described in Section 2. It also

attempted a remedy in the form of introducing an Access-Node feature, as described in Section 5.3 of [I-D.ietf-6man-lineid]. This feature, consists in the Access Node issuing RS messages towards the Router driven by subscriber link activation (and only activation) state (ie when the link is "brought up"). The term "proxy-RS sender" rather aptly describes the feature, as denoted in Figure 2 below.



[I-D.ietf-6man-lineid] indicates that a finite number of RS messages are to be sent and that sending should stop after the Access Node receives an RA with a matching subscriber line information option back from the edge router. This remedy, in the context of the overall solution, is not only insufficient, but introduces further problems, consisting of:

1. Unreliability of RS messaging: There is no assurance that the RS messages sent by the proxy will reach the edge router. Eg it is not uncommon for spanning tree protocol events take place on the Ethernet segments, or other similar events, which result in loss of connectivity with the edge router ranging from a couple of seconds to a couple of minutes - this is often the case during access-node activation. Any RS messages sent by the RS-proxy, on behalf of bridged subscribers connected to this access node, would be lost and all the relevant subscribers left without IPv6 connectivity.

2. Lack of subscriber host identifiers: In many of today's broadband

deployments end host identifiers are required for the purpose of authorization besides intermediate identifiers such as subscriber line-id. For example, it is quite common to identify and authorize devices like WiFi smart phones or TV set-top-boxes by their unique MAC address. With the RS-proxy mechanism, these identifiers are not be available, and effectively do not meet goal b) of the system

3. No ability to clean up state/recover: Each "active" subscriber link is intended to induce IPv6 subscriber state in the router. Short of manual intervention by the operator there is no mechanism on the router to remove such state should a link ever become "inactive". In other words, there is no equivalent of a "link down" message, nor does the ND protocol provide for such extensibility, and the router and operator are likely to be burdened with a large amount of stale state, besides inefficient use of resources.

4. In ability to recover from node failures: Given that an RS-proxy eventually stops sending RSes, should the edge router loose for any reason any or all of the RS induced state, including the route to the subscriber, the system will fall into a state of unrecoverable connectivity loss for end users, even as they continue to have a valid IPv6 address. Basically, a host that received a previous RA from an Edge Router will following rfc4862 NOT send an further RS messages, while a router without the necessary state will NOT forward traffic to the subscriber. Similarly, neither will the RS-proxy send RS messages as long as the line is still "active".

Given the above issues, while the introduction of the RS-sending-proxy was intended to fix a critical flaw with the original proposal, if not only left the issue in place, but it introduced further issues undermining its overall purpose and compromising the usability and scalability of the system.

3. Discussion of possible solutions

Its readily apparent that any solution based on proxy functionality that is driven by link state changes cannot meet all of the system goals as presented in Section 2 (eg goals a, b and c), while satisfying the constraint of no changes to end hosts (goal e) and within the context of a bridged/indirect-link host-router set-up (goal d). At best compromises to the goals or combinations of solutions need to be adopted. The solutions below indicate such compromises:

3.1. Modifying RFC4861

One possible, solution, that would solve a handful of issues, would be to modify [RFC4861] in such a way as to give the protocol a semblance of reliability and persistence. For example, it could be stipulated that host RS sending behaviour needs to be periodic and continue irrespective of RA messages being received. Router behaviour would need to be modified to detect periods of RS inactivity. All this would be a substantial change to the original protocol specification, and would naturally require changes to any existing IPv6 ND implementations to be useful, falling short of goal e). Besides this, it would also significantly increase the RS processing load on any router.

3.2. Modifying RS-proxy and router behaviour

Modifying the RS-proxy mechanism to issue periodic RS messages driven subscriber link state, or doing so whenever no RA is received for a given subscriber line over a certain period of time could be seen as a possible solution to some, but not all, of the problems identified. In essence this modification transforms RS/RA messaging into link-state notification messages. Unfortunately it also introduces several other flaws, besides not meeting the Section 2 goals a), b) and possibly c):

- Unknown timers: For the mechanism to function, the behaviour of both the RS-proxy and the edge router need to be modified in terms of RS processing and RA sending, around a timer driven state machine, where both the Access-Node and Router share the timers. Defining for this purpose a new timer negotiation protocol appears a major ND or IPinIP protocol change, while relying on "well known" timers (ie hard set) is highly inflexibility not conductive to automated, reliable and inter operable deployments.

- Increased load on AAA system: Following the intent of the system, for each RS message for which no authorization state exists on the edge router, authorization from an AAA server is to be requested. With RS messages being periodic, this will place additional burden on any AAA infrastructure, besides being analogous to issuing AAA requests for each link keepalive received.

- Subscriber management: One of the main premises of an architecture that features a Layer 2 Access Node and an upstream aggregating IP Edge Router is the notion of subscriber management on the IP Edge Router. Operators deploying this architecture seek to use the IP Edge Router as the node on which subscriber related configuration

and control is applied - hence the desire to perform dynamic subscriber authorization at/by the router. Introducing into this

architecture a mechanism where periodic RS messages sent by a proxy could lead to similarly periodic denial of authorizations at the edge router, eg for subscriber lines that are not authorized to use the service, with the only way of disabling such RS sending is by maintaining on the Access-Node subscriber configuration information, is counter to the premise of the architecture itself.

- ND customization: One of the design goals for using the IPinIP tunneling mechanism was to avoid changes to the ND protocol or implementations. Unfortunately, the processing of custom tunneled RS messages as well as generation of custom tunneled RA messages, in effect requires a highly customized ND implementation, the likes of which diverges from typically ND implementations.

Given the above, modifying the RS-proxy mechanism to be periodic would not only require a fairly major extension to the proposal, including the definition of timers covering message sending periodicity discovery and/or negotiation, but also result in more issues to the overall system. Above all, such a modification would in the end only mimic a link-state signalling/keepalive protocol, without actually resolving all of the identified problems, and without actually being one.

3.3. Ethernet Connectivity Fault Monitoring

A core issue in the a system driven by host sourced RS, is the end hosts inability to detect when an indirect link has failed, translating into the hosts inability to re-send RS messages. On links such as PPP, which offer link state keepalives, the issue does not come up, but neither does the need of driving router authorization events via RS messages due to the link layer negotiation stage of PPP. Over Ethernet, a link state keepalive mechanisms could fill in part of that gap. The closest equivalent can be found in Ethernet Connectivity Fault Monitoring that is a component of the IEEE 802.1ag Ethernet OAM specification [802.1ag]. The implementation of such extensions on hosts and routers would allow the regular [RFC4861] RA sending rules to respond appropriately to connectivity or device failures. Unfortunately, there is no known end host implementation of 802.1ag today, which translates that this

solution does not meet goal e) (no end host modifications). Nevertheless, it appears like a valid approach, whose realization however does not appear to be within the IETF's specification direct sphere of influence.

3.4. Access-Node based DHCPv6 Proxy Client

An alternative solution to some of the problems identified in relation to periodic RA sending, would be to define an RS/

RA-DHCPv6-proxy function, whose role would be to transform host sourced RS messages into DHCPv6 Solicit/etc messages towards the edge router. The access-node would thus be a multi DUID DHCPv6 client as seen by the rest of the operator's network. Regular mechanisms of DHCPv6 relaying by the edge router and prefix delegation would be used to assign /64 prefixes for each subscriber line. The RS/ RA-DHCPv6 proxy would also be responsible for announcing the DHCPv6 derived prefixes in regular RA messages to downstream hosts. An additional bonus of this solution is the fact that the existing DHCPv6 specification allows for the subscriber line-id to be included in the DHCPv6 messages [RFC3315], [RFC6221]. Hence, no additional RS subscriber line id or IPinIP tunnel header extensions would be required, effectively obviating all of the [I-D.ietf-6man-lineid] protocol extension requirements. Similarly, none of the upstream devices, would appear to be affected in supporting this solution.

Though this solution solves the problem of error recovery, state deletion and timer discovery/negotiation, besides removing the need to define any protocol extensions to convey line-id information, in its RS triggered form it remains prone to the critical flaws described in Section 2. Hence, a more reliable version of this solution would see the DHCPv6 proxy client be invoked by line-state changes. Unfortunately, this variant again does not meet goals a), b) and possibly c). Nevertheless, with these usability caveats clearly recognized, it appears that this solution is still superior to what is currently found in [I-D.ietf-6man-lineid], and does not require protocol extensions.

3.5. DHCPv6 client on end hosts

A solution that would see most of the goals realized, without the need to define any new protocol extensions, would be to rely on

DHCPv6 [rfc3315] client functionality in the end host. DHCPv6 was designed to offer the degree of reliability sought for, as well as periodic retransmissions of messages, along with client identifiers. The compromise in this solution would be that it does not appear to fit goal e), at least when looked from a universal current host implementation perspective, namely that some end hosts would be required to implement a DHCPv6 client.

Given the relation of the problem being addressed to the bridged connectivity model, a non technical variant of this solution at the service level is to stipulate in the user's terms and conditions it is supported only with DHCPv6 clients. This approach has been effectively assumed by the Cablelabs specifications for bridged media connectivity [MULPI], as well as put into practice by several Ethernet FTTx network operators.

3.6. ANCP

The Access Node Control Protocol (ANCP) [I-D.ietf-ancp-protocol] defines a suite of mechanisms for conveying information pertaining to the state of a subscriber access line between a Layer 2 access node physically terminating the subscriber access line and a separate Layer 3 router. One of the key capabilities of the protocol is that to signal line state changes from the access-node to the router, as well as to apply dynamic configuration on access-lines retrieved from the router. In the combination, these two capabilities offer another alternative solution, at least in so far as a line-state driven mechanism can provide.

The basic premise of the solution would see the Access-Node use existing ANCP "Port-UP" or "Port-Down" messages, which also convey line-id, to signal line state changes to the edge router. These could be considered as the trigger events to drive the edge router to send to the Access-Node either "Line Configuration" messages with IPv6 parameters, or define a new "Raw data" message type which would ferry a raw RA to be sent on the access-line.

As with any of the other Access-Node line state driven solutions, meeting goals a) and b) would not be possible. Despite that, ANCP offers a robust and reliable (TCP based) line-state communication mechanism between an Access-Node and Edge Router, which does not need

re-inventing.

3.7. Other

The solution proposed by [I-D.ietf-6man-lineid], consisting in adding a subscriber-line-id parameter as part of an IPinIP encapsulation header, can be realized practically by various other tunneling protocols. Specifically, L2TPv3 already defines AVPs for subscriber-line-id information. As with other solutions that rely only on tunneling host sourced RAs, this will be prone to host connectivity impediments.

4. Conclusions

Due to the inherent design and implementation characteristics of the ND protocol, mechanisms that gate IPv6 user connectivity based on the reception of an RS message are likely to lead to serious IPv6 connectivity failures for end users, and leave both users and operators with no automated means of recovering from the situation. The issues are particularly severe in cases when the end users do not have a direct link adjacency to the router, as is often the case in bridged Ethernet or WiFi based broadband access networks. Moreover,

such a mechanism appears not to meet the expected more general usage goals as presented in Section 2. As such, the definition and deployment of such mechanisms is considered to be harmful to the success of IPv6 usage, and thus should be discouraged in favour of alternative solutions.

Two alternative solutions presented in Sections 3.4 and 3.6, can comprehensively meet the majority of the Section 2 goals. The solution presented in Section 3.5, which has proven to meet the requirements of many operators, indicating the imposed host constraints might not be universally applicable, remains a valid approach which requires no protocol extensions.

Solution variants seek to redress the lack of direct link state adjacency by using an intermediate link state driven messaging proxy function incur a shortcoming. This consist in their inability to be able to provide the to the authorization system information such as the end host MAC address. Thus, any such solution carries usage

constraints, that should be clarified.

The solution variant proposed by [I-D.ietf-6man-lineid] introduces itself numerous issues of reliability and deployability, whose resolution is not trivial without major ND protocol extensions, if not other protocol work. Alternatives, as presented in Section 3.4, 3.6 and 3.7 all offer more robust and deployable mechanisms that in most cases leverage already defined protocols and mechanisms hence appear to offer a much more viable solution path.

5. IANA Considerations

This document does not raise any IANA considerations.

6. Security Considerations

The security of the solutions outlined needs to be evaluated in specific solution documents.

7. Contributors and Acknowledgements

The author would like to thank Erik Nordmark, Ole Troan, and Sean Cavanaugh for reviewing this document.

8. References

8.1. Normative References

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

8.2. Informative References

[I-D.ietf-6man-lineid] Krishnan, S., Kavanagh, A., Varga, B., Ooghe, S., and E. Nordmark, "The Line Identification Destination Option", draft-ietf-6man-lineid-01 (work in progress), March 2011.

[I-D.ietf-ancp-protocol]
Wadhwa, S., Moisand, J., Haag, T., Voigt, N., and T. Taylor, "Protocol for Access Node Control Mechanism in Broadband Networks", draft-ietf-ancp-protocol-17 (work in progress), April 2011.

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

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

[RFC5851] Ooghe, S., Voigt, N., Platnic, M., Haag, T., and S. Wadhwa, "Framework and Requirements for an Access Node Control Mechanism in Broadband Multi-Service Networks", RFC 5851, May 2010.

[RFC6221] Miles, D., Ooghe, S., Dec, W., Krishnan, S., and A. Kavanagh, "Lightweight DHCPv6 Relay Agent", RFC 6221, May 2011.

[TR-177] - Broadband Forum,<http://www.broadband-forum.org/technical/download/TR-177.pdf>

[IEEE802.1ag] - IEEE,<http://www.ieee802.org/1/pages/802.1ag.html>

Author's Address

Wojciech Dec
Cisco Systems

Haarlerbergweg 13-19
1101 CH Amsterdam
The Netherlands

Email: wdec@cisco.com