

ROLL
Internet-Draft
Intended status: Standards Track
Expires: September 13, 2017

S. Anamalamudi
Huaiyin Institute of Technology
M. Zhang
AR. Sangi
Huawei Technologies
C. Perkins
Futurewei
S.V.R.Anand
Indian Institute of Science
March 12, 2017

Asymmetric AODV-P2P-RPL in Low-Power and Lossy Networks (LLNs)
draft-ietf-roll-aodv-rpl-01

Abstract

Route discovery for symmetric and asymmetric Point-to-Point (P2P) traffic flows is a desirable feature in Low power and Lossy Networks (LLNs). For that purpose, this document specifies a reactive P2P route discovery mechanism for hop-by-hop routing (storing mode) based on Ad Hoc On-demand Distance Vector Routing (AODV) based RPL protocol. Two separate Instances are used to construct directional paths in case some of the links between source and target node are asymmetric.

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 September 13, 2017.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the 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

1. Introduction	2
2. Terminology	3
3. Overview of AODV-RPL	5
4. AODV-RPL Mode of Operation (MoP)	5
5. RREQ Message	8
6. RREP Message	9
7. Gratuitous RREP	11
8. Operation of Trickle Timer	12
9. IANA Considerations	12
9.1. New Mode of Operation: AODV-RPL	12
9.2. AODV-RPL Options: RREQ and RREP	12
10. Security Considerations	12
11. References	12
11.1. Normative References	13
11.2. Informative References	14
Appendix A. ETX/RSSI Values to select S bit	14
Authors' Addresses	15

1. Introduction

RPL[RFC6550], the IPv6 distance vector routing protocol for Low-power and Lossy Networks (LLNs), is designed to support multiple traffic flows through a root-based Destination-Oriented Directed Acyclic Graph (DODAG). For traffic flows between routers within the DODAG (i.e., Point-to-Point (P2P) traffic), this means that data packets either have to traverse the root in non-storing mode (source routing), or traverse a common ancestor in storing mode (hop-by-hop routing). Such P2P traffic is thereby likely to flow along sub-optimal routes and may suffer severe traffic congestion near the DAG root [RFC6997], [RFC6998].

To discover optimal paths for P2P traffic flows in RPL, P2P-RPL [RFC6997] specifies a temporary DODAG where the source acts as temporary root. The source initiates "P2P Route Discovery mode (P2P-RDO)" with an address vector for both non-storing mode (H=0) and storing mode (H=1). Subsequently, each intermediate router adds its

IP address and multicasts the P2P-RDO message, until the message reaches the target node (TargNode). TargNode sends the "Discovery Reply" option. P2P-RPL is efficient for source routing, but much less efficient for hop-by-hop routing due to the extra address vector overhead. In fact, when the P2P-RDO message is being multicast from the source hop-by-hop, receiving nodes are able to determine a next hop towards the source in symmetric links. When TargNode subsequently replies to the source along the established forward route, receiving nodes can determine the next hop towards TargNode. In other words, it is efficient to use only routing tables for P2P-RDO message instead of "Address vector" for hop-by-hop routes (H=1) in symmetric links.

RPL and P2P-RPL both specify the use of a single DODAG in networks of symmetric links. But, application-specific routing requirements that are defined in IETF ROLL Working Group [RFC5548], [RFC5673], [RFC5826] and [RFC5867] may need routing metrics and constraints enabling use of asymmetric bidirectional links. For this purpose, [I-D.thubert-roll-asymlink] describes bidirectional asymmetric links for RPL [RFC6550] with Paired DODAGs, for which the DAG root (DODAGID) is common for two Instances. This can satisfy application-specific routing requirements for bidirectional asymmetric links in base RPL [RFC6550]. P2P-RPL for Paired DODAGs, on the other hand, requires two DAG roots: one for the source and another for the target node due to temporary DODAG formation. For networks composed of bidirectional asymmetric links (see Section 4), AODV-RPL specifies P2P route discovery, utilizing RPL with a new MoP. AODV-RPL makes use of two multicast messages to discover possibly asymmetric routes. AODV-RPL eliminates the need for address vector control overhead, significantly reducing the control packet size which is important for Constrained LLN networks. Both discovered routes meet the application specific metrics and constraints that are defined in the Objective Function for each Instance [RFC6552].

2. Terminology

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 [RFC2119]. Additionally, this document uses the following terms:

AODV

Ad Hoc On-demand Distance Vector Routing[RFC3561].

AODV-Instance

Either the RREQ-Instance or RREP-Instance

Bi-directional Asymmetric Link

A link that can be used in both directions but with different link characteristics (see [I-D.thubert-roll-asymlink]).

DODAG RREQ-Instance (or simply RREQ-Instance)

AODV Instance built using the RREQ option; used for control transmission from OrigNode to TargNode, thus enabling data transmission from TargNode to OrigNode.

DODAG RREP-Instance (or simply RREP-Instance)

AODV Instance built using the RREP option; used for control transmission from TargNode to OrigNode thus enabling data transmission from OrigNode to TargNode.

downstream

Routing along the direction from OrigNode to TargNode.

hop-by-hop routing

Routing when each node stores routing information about the next hop.

OrigNode

The IPv6 router (Originating Node) initiating the AODV-RPL route discovery to obtain a route to TargNode.

Paired DODAGs

Two DODAGs for a single application.

P2P

Point-to-Point -- in other words, not constrained to traverse a common ancestor.

RREQ message

An AODV-RPL MoP DIO message containing the RREQ option. The InstanceID in DIO object of RREQ option MUST be always an odd number.

RREP message

An AODV-RPL MoP DIO message containing the RREP option. The InstanceID in DIO object of RREP option MUST be always an even number (InstanceID of RREQ-Instance+1).

source routing

The mechanism by which the source supplies the complete route towards the target node along with each data packet. [RFC6997].

TargNode

The IPv6 router (Target Node) for which OrigNode requires a route and initiates Route Discovery within the LLN network.

upstream

Routing along the direction from TargNode to OrigNode.

3. Overview of AODV-RPL

With AODV-RPL, routes from OrigNode to TargNode within the LLN network established are "on-demand". In other words, the route discovery mechanism in AODV-RPL is invoked reactively when OrigNode has data for delivery to the TargNode but existing routes do not satisfy the application's requirements. The routes discovered by AODV-RPL are point-to-point; in other words the routes are not constrained to traverse a common ancestor. Unlike base RPL [RFC6550] and P2P-RPL [RFC6997], AODV-RPL can enable asymmetric communication paths in networks with bidirectional asymmetric links. For this purpose, AODV-RPL enables discovery of two routes: namely, one from OrigNode to TargNode, and another from TargNode to OrigNode. When possible, AODV-RPL also enables symmetric routing along Paired DODAGs (see Section 4).

4. AODV-RPL Mode of Operation (MoP)

In AODV-RPL, route discovery is initiated by forming a temporary DAG rooted at the OrigNode. Paired DODAGs (Instances) are constructed according to a new AODV-RPL Mode of Operation (MoP) during route formation between the OrigNode and TargNode. The RREQ-Instance is formed by route control messages from OrigNode to TargNode whereas the RREP-Instance is formed by route control messages from TargNode to OrigNode (as shown in Figure 2). Intermediate routers join the Paired DODAGs based on the rank as calculated from the DIO message. Henceforth in this document, the RREQ-Instance message means the AODV-RPL DIO message from OrigNode to TargNode, containing the RREQ option. Similarly, the RREP-Instance means the AODV-RPL DIO message from TargNode to OrigNode, containing the RREP option. Subsequently, the RREQ-Instance is used for data transmission from TargNode to OrigNode and RREP-Instance is used for Data transmission from OrigNode to TargNode.

The AODV-RPL Mode of Operation defines a new bit, the Symmetric bit ('S'), which is added to the base DIO message as illustrated in Figure 1. OrigNode sets the the 'S' bit to 1 in the RREQ-Instance message when initiating route discovery.

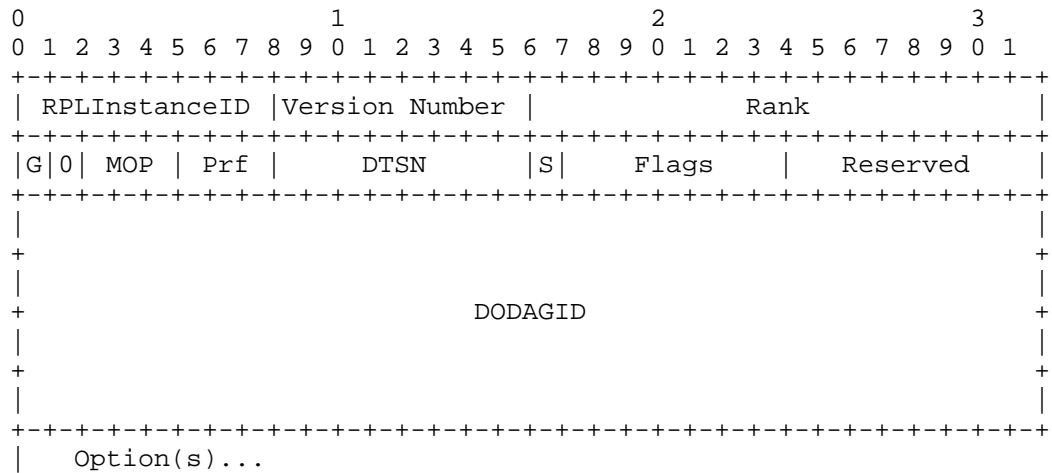


Figure 1: DIO modification to support asymmetric route discovery

A device originating a AODV-RPL message supplies the following information in the DIO header of the message:

'S' bit

Symmetric bit in the DIO base object

MOP

MOP operation in the DIO object MUST be set to "5(TBD1)" for AODV-RPL DIO messages

RPLInstanceID

RPLInstanceID in the DIO object MUST be the InstanceID of AODV-Instance(RREQ-Instance). The InstanceID for RREQ-Instance MUST be always an odd number.

DODAGID

For RREQ-Instance :

DODAGID in the DIO object MUST be the IPv6 address of the device that initiates the RREQ-Instance.

For RREP-Instance

DODAGID in the DIO object MUST be the IPv6 address of the device that initiates the RREP-Instance.

Rank

Rank in the DIO object MUST be the the rank of the AODV-Instance (RREQ-Instance).

Metric Container Options

AODV-Instance(RREQ-Instance) messages MAY carry one or more Metric Container options to indicate the relevant routing metrics.

The 'S' bit is set to mean that the route is symmetric. If the RREQ-Instance arrives over an interface that is known to be symmetric, and the 'S' bit is set to 1, then it remains set at 1, as illustrated in Figure 2.

In this figure:

S := OrigNode; R := Intermediate nodes; D := TargNode

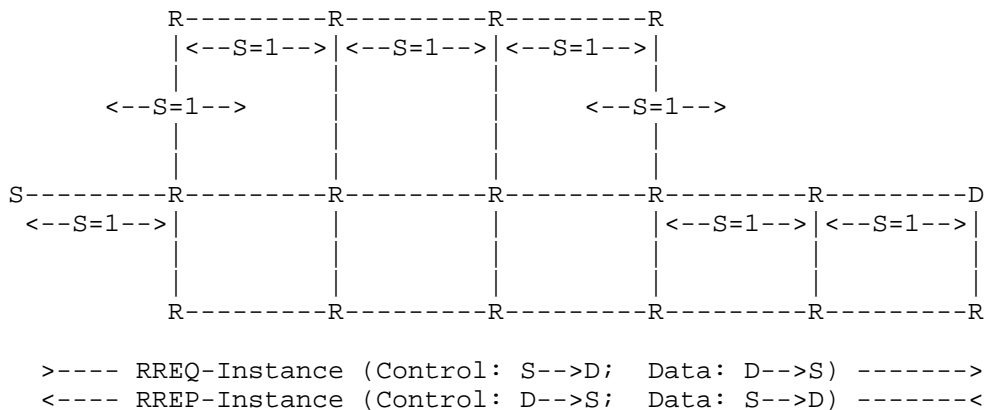


Figure 2: AODV-RPL with Symmetric Paired Instances

If the RREQ-Instance arrives over an interface that is not known to be symmetric, or is known to be asymmetric, the 'S' bit is set to be 0. Moreover, if the 'S' bit arrives already set to be '0', it is set to be '0' on retransmission (Figure 3). Based on the 'S' bit received in RREQ-Instance, the TargNode decides whether or not the route is symmetric before transmitting the RREP-Instance message upstream towards the OrigNode. The metric used to determine symmetry (i.e., set the "S" bit to be "1" (Symmetric) or "0" (asymmetric)) is implementation specific. We used ETX/RSSI to verify the feasibility of the protocol operations in this draft, as discussed in Appendix A.

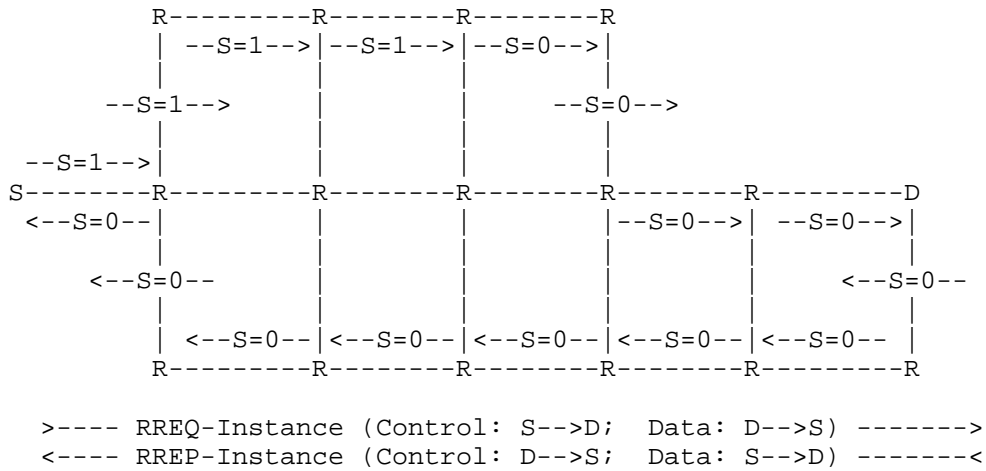


Figure 3: AODV-RPL with Asymmetric Paired Instances

5. RREQ Message

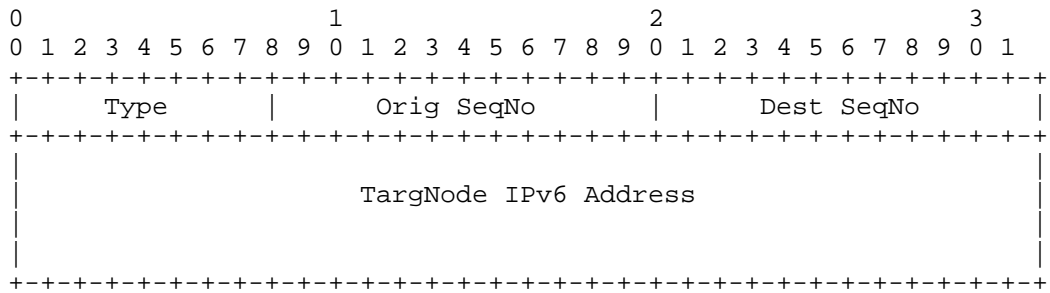


Figure 4: DIO RREQ option format for AODV-RPL MoP

OrigNode supplies the following information in the RREQ option of the RREQ-Instance message:

Type

The type of the RREQ option(see Section 9.2).

Orig SeqNo

Sequence Number of OrigNode.

Dest SeqNo

If nonzero, the last known Sequence Number for TargNode for which a route is desired.

TargNode IPv6 Address

IPv6 address of the TargNode that receives RREQ-Instance message. This address MUST be in the RREQ option (see Figure 4) of AODV-RPL.

In order to establish the upstream route from TargNode to OrigNode, OrigNode multicasts the RREQ-Instance message (see Figure 4) to its one-hop neighbours. In order to enable intermediate nodes R_i to associate a future RREP message to an incoming RREQ message, the InstanceID of RREQ-Instance MUST assign an odd number.

Each intermediate node R_i computes the rank for RREQ-Instance and creates a routing table entry for the upstream route towards the source if the routing metrics/constraints are satisfied. For this purpose R_i must use the asymmetric link metric measured in the upstream direction, from R_i to its upstream neighbor that multicasted the RREQ-Instance message.

When an intermediate node R_i receives a RREQ message in storing mode, it MUST store the OrigNode's InstanceID (RREQ-Instance) along with the other routing information needed to establish the route back to the OrigNode. This will enable R_i to determine that a future RREP message (containing a paired InstanceID for the TargNode) must be transmitted back to the OrigNode's IP address.

If the paths to and from TargNode are not known, the intermediate node multicasts the RREQ-Instance message with updated rank to its next-hop neighbors until the message reaches TargNode (Figure 2). Based on the 'S' bit in the received RREQ message, the TargNode will decide whether to unicast or multicast the RREP message back to OrigNode.

As described in Section 7, in certain circumstances R_i MAY unicast a Gratuitous RREP towards OrigNode, thereby helping to minimize multicast overhead during the Route Discovery process.

6. RREP Message

The TargNode supplies the following information in the RREP message:

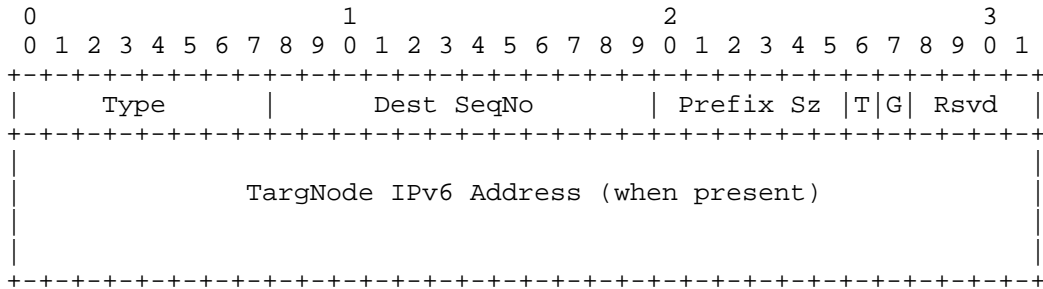


Figure 5: DIO RREP option format for AODV-RPL MoP

Type

The type of the RREP option (see Section 9.2)

Dest SeqNo

The Sequence Number for the TargNode for which a route is established.

Prefix Sz

The size of the prefix which the route to the TargNode is available. This allows routing to other nodes on the same subnet as the TargNode.

'T' bit

'T' is set to true to indicate that the TargNode IPv6 Address field is present

'G' bit

(see Section 7)

TargNode IPv6 Address (when present)

IPv6 address of the TargNode that receives RREP-Instance message.

In order to reduce the need for the TargNode IPv6 Address to be included with the RREP message, the InstanceID of the RREP-Instance is paired, whenever possible, with the InstanceID from the RREQ message, which is always an odd number. The pairing is accomplished by adding one to the InstanceID from the RREQ message and using that, whenever possible, as the InstanceID for the RREP message. If this is not possible (for instance because the incremented InstanceID is

still a valid InstanceID for another route to the TargNode from an earlier Route Discovery operation), then the 'T' bit is set and an alternative even number is chosen for the InstanceID of RREP from TargNode.

The OrigNode IP address for RREQ-Instance is available as the DODAGID in the DIO base message (see Figure 1). When TargNode receives a RREQ message with the 'S' bit set to 1 (as illustrated in Figure 2), it unicasts the RREP message with the 'S' bit set to 1. In this case, route control messages and application data between OrigNode and TargNode for both RREQ-Instance and RREP-Instance are transmitted along symmetric links. When 'T' bit set is to "1" in the RREP-Instance, then TargNode IPv6 Address is transmitted in RREP option. Otherwise, the TargNode IPv6 Address is elided in RREP option.

When (as illustrated in Figure 3) the TargNode receives RREQ message with the 'S' bit set to 0, it also multicasts the RREP message with the 'S' bit set to 0. Intermediate nodes create a routing table entry for the path towards the TargNode while processing the RREP message to OrigNode. Once OrigNode receives the RREP message, it starts transmitting the application data to TargNode along the path as discovered through RREP messages. Similarly, application data from TargNode to OrigNode is transmitted through the path that is discovered from RREQ message.

7. Gratuitous RREP

Under some circumstances, an Intermediate Node that receives a RREQ message MAY transmit a "Gratuitous" RREP message back to OrigNode instead of continuing to multicast the RREQ message towards TargNode. For these circumstances, the 'G' bit of the RREP option is provided to distinguish the Gratuitous RREP sent by the Intermediate node from the RREP sent by TargNode.

When an Intermediate node R receives a RREQ message and has recent information about the cost of an upstream route from TargNode to R, then R MAY unicast the Gratuitous RREP (GRREP) message to OrigNode. R determines whether its information is sufficiently recent by comparing the value it has stored for the Sequence Number of TargNode against the DestSeqno in the incoming RREQ message. R also must have information about the metric information of the upstream route from TargNode. The GRREP message MUST have PrefixSz == 0 and the 'G' bit set to 1. R SHOULD also unicast the RREQ message to TargNode, to make sure that TargNode will have a route to OrigNode.

8. Operation of Trickle Timer

The trickle timer operation to control RREQ-Instance/RREP-Instance multicast is similar to that in P2P-RPL [RFC6997].

9. IANA Considerations

9.1. New Mode of Operation: AODV-RPL

IANA is required to assign a new Mode of Operation, named "AODV-RPL" for Point-to-Point(P2P) hop-by-hop routing under the RPL registry. The value of TBD1 is assigned from the "Mode of Operation" space [RFC6550].

Value	Description	Reference
TBD1 (5)	AODV-RPL	This document

Figure 6: Mode of Operation

9.2. AODV-RPL Options: RREQ and RREP

Two entries are required for new AODV-RPL options "RREQ-Instance" and "RREP-Instance", with values of TBD2 (0x0A) and TBD3 (0x0B) from the "RPL Control Message Options" space [RFC6550].

Value	Meaning	Reference
TBD2 (0x0A)	RREQ Option	This document
TBD3 (0x0B)	RREP Option	This document

Figure 7: AODV-RPL Options

10. Security Considerations

This document does not introduce additional security issues compared to base RPL. For general RPL security considerations, see [RFC6550].

11. References

11.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<http://www.rfc-editor.org/info/rfc2119>>.
- [RFC3561] Perkins, C., Belding-Royer, E., and S. Das, "Ad hoc On-Demand Distance Vector (AODV) Routing", RFC 3561, DOI 10.17487/RFC3561, July 2003, <<http://www.rfc-editor.org/info/rfc3561>>.
- [RFC5548] Dohler, M., Ed., Watteyne, T., Ed., Winter, T., Ed., and D. Barthel, Ed., "Routing Requirements for Urban Low-Power and Lossy Networks", RFC 5548, DOI 10.17487/RFC5548, May 2009, <<http://www.rfc-editor.org/info/rfc5548>>.
- [RFC5673] Pister, K., Ed., Thubert, P., Ed., Dwars, S., and T. Phinney, "Industrial Routing Requirements in Low-Power and Lossy Networks", RFC 5673, DOI 10.17487/RFC5673, October 2009, <<http://www.rfc-editor.org/info/rfc5673>>.
- [RFC5826] Brandt, A., Buron, J., and G. Porcu, "Home Automation Routing Requirements in Low-Power and Lossy Networks", RFC 5826, DOI 10.17487/RFC5826, April 2010, <<http://www.rfc-editor.org/info/rfc5826>>.
- [RFC5867] Martocci, J., Ed., De Mil, P., Riou, N., and W. Vermeylen, "Building Automation Routing Requirements in Low-Power and Lossy Networks", RFC 5867, DOI 10.17487/RFC5867, June 2010, <<http://www.rfc-editor.org/info/rfc5867>>.
- [RFC6550] Winter, T., Ed., Thubert, P., Ed., Brandt, A., Hui, J., Kelsey, R., Levis, P., Pister, K., Struik, R., Vasseur, JP., and R. Alexander, "RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks", RFC 6550, DOI 10.17487/RFC6550, March 2012, <<http://www.rfc-editor.org/info/rfc6550>>.
- [RFC6552] Thubert, P., Ed., "Objective Function Zero for the Routing Protocol for Low-Power and Lossy Networks (RPL)", RFC 6552, DOI 10.17487/RFC6552, March 2012, <<http://www.rfc-editor.org/info/rfc6552>>.

[RFC6997] Goyal, M., Ed., Baccelli, E., Philipp, M., Brandt, A., and J. Martocci, "Reactive Discovery of Point-to-Point Routes in Low-Power and Lossy Networks", RFC 6997, DOI 10.17487/RFC6997, August 2013, <<http://www.rfc-editor.org/info/rfc6997>>.

[RFC6998] Goyal, M., Ed., Baccelli, E., Brandt, A., and J. Martocci, "A Mechanism to Measure the Routing Metrics along a Point-to-Point Route in a Low-Power and Lossy Network", RFC 6998, DOI 10.17487/RFC6998, August 2013, <<http://www.rfc-editor.org/info/rfc6998>>.

11.2. Informative References

[I-D.thubert-roll-asymlink]
 Thubert, P., "RPL adaptation for asymmetrical links", draft-thubert-roll-asymlink-02 (work in progress), December 2011.

Appendix A. ETX/RSSI Values to select S bit

We have tested the combination of "RSSI(downstream)" and "ETX (upstream)" to decide whether the link is symmetric or asymmetric at the intermediate nodes. The example of how the ETX and RSSI values are used in conjunction is explained below:

Source----->NodeA----->NodeB----->Destination

Figure 8: Communication link from Source to Destination

RSSI at NodeA for NodeB	Expected ETX at NodeA for nodeB->nodeA
> -15	150
-25 to -15	192
-35 to -25	226
-45 to -35	662
-55 to -45	993

Table 1: Selection of 'S' bit based on Expected ETX value

We tested the operations in this specification by making the following experiment, using the above parameters. In our experiment, a communication link is considered as symmetric if the ETX value of NodeA->NodeB and NodeB->NodeA (See Figure.8) are, say, within 1:3 ratio. This ratio should be taken as a notional metric for deciding link symmetric/asymmetric nature, and precise definition of the ratio

is beyond the scope of the draft. In general, NodeA can only know the ETX value in the direction of NodeA -> NodeB but it has no direct way of knowing the value of ETX from NodeB->NodeA. Using physical testbed experiments and realistic wireless channel propagation models, one can come up with a relationship between RSSI and ETX that can be represented as an expression or a mapping table. Such a relationship in turn can be used to estimate ETX value at nodeA for link NodeB-->NodeA from the received RSSI from NodeB. Whenever nodeA determines that the link towards the nodeB is bi-directional asymmetric then the "S" bit is set to "S=0". Later on, link from NodeA to Destination is asymmetric with "S" bit remains to "0".

Authors' Addresses

Satish Anamalamudi
Huaiyin Institute of Technology
No.89 North Beijing Road, Qinghe District
Huaian 223001
China

Email: satishnaidu80@gmail.com

Mingui Zhang
Huawei Technologies
No. 156 Beiqing Rd. Haidian District
Beijing 100095
China

Email: zhangmingui@huawei.com

Abdur Rashid Sangi
Huawei Technologies
No.156 Beiqing Rd. Haidian District
Beijing 100095
P.R. China

Email: sangi_bahrian@yahoo.com

Charles E. Perkins
Futurewei
2330 Central Expressway
Santa Clara 95050
Unites States

Email: charliep@computer.org

S.V.R Anand
Indian Institute of Science
Bangalore 560012
India

Email: anand@ece.iisc.ernet.in

ROLL
Internet-Draft
Intended status: Standards Track
Expires: September 11, 2017

P. Thubert, Ed.
J. Pylakutty
Cisco
March 10, 2017

Root initiated routing state in RPL
draft-ietf-roll-dao-projection-01

Abstract

This document proposes a protocol extension to RPL that enables to install a limited amount of centrally-computed routes in a RPL graph, enabling loose source routing down a non-storing mode DODAG, or transversal routes inside the DODAG. As opposed to the classical route injection in RPL that are injected by the end devices, this draft enables the root of the DODAG to projects the routes that are needed on the nodes where they should be installed.

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 September 11, 2017.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the 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

1. Introduction	2
2. Terminology	3
3. New RPL Control Message Options	3
3.1. Via Information Option	4
4. Projected DAO	5
4.1. Non-storing Mode Projected DAO	6
4.2. Storing-Mode Projected DAO	8
5. Applications	10
5.1. Loose Source Routing in Non-storing Mode	10
5.2. Transversal Routes in storing and non-storing modes	11
6. RPL Instances	13
7. Security Considerations	14
8. IANA Considerations	14
9. Acknowledgments	14
10. References	14
10.1. Normative References	14
10.2. Informative References	15
Appendix A. Examples	15
A.1. Using storing mode P-DAO in non-storing mode MOP	16
A.2. Projecting a storing-mode transversal route	17
Authors' Addresses	18

1. Introduction

The Routing Protocol for Low Power and Lossy Networks (LLN)(RPL) [RFC6550] is a generic Distance Vector protocol that is well suited for application in a variety of low energy Internet of Things (IoT) networks. RPL forms Destination Oriented Directed Acyclic Graphs (DODAGs) in which the root often acts as the Border Router to connect the RPL domain to the Internet. The root is responsible to select the RPL Instance that is used to forward a packet coming from the Internet into the RPL domain and set the related RPL information in the packets.

The 6TiSCH architecture [I-D.ietf-6tisch-architecture] leverages RPL for its routing operation and considers the Deterministic Networking Architecture [I-D.ietf-detnet-architecture] as one possible model whereby the device resources and capabilities are exposed to an external controller which installs routing states into the network based on some objective functions that reside in that external entity.

Based on heuristics of usage, path length, and knowledge of device capacity and available resources such as battery levels and reservable buffers, a Path Computation Element ([PCE]) with a global visibility on the system could install additional P2P routes that are more optimized for the current needs as expressed by the objective function.

This draft enables a RPL root, with optionally the assistance of a PCE, to install and maintain additional storing and non-storing mode routes within the RPL domain, along a selected set of nodes and for a selected duration, thus providing routes more suitable than those obtained with the distributed operation of RPL. Those routes may be installed in either storing and non-storing modes RPL instances, resulting in potentially hybrid situations where the mode of the projected routes is different from that of the other routes in the instance.

2. Terminology

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 [RFC2119].

The Terminology used in this document is consistent with and incorporates that described in 'Terminology in Low power And Lossy Networks' [RFC7102] and [RFC6550].

3. New RPL Control Message Options

Section 6.7 of [RFC6550] specifies Control Message Options (CMO) to be placed in RPL messages such as the Destination Advertisement Object (DAO) message. The RPL Target Option and the Transit Information Option (TIO) are such options; the former indicates a node to be reached and the latter specifies a parent that can be used to reach that node. Options may be factorized; one or more contiguous TIOs apply to the one or more contiguous Target options that immediately precede the TIOs in the RPL message.

This specification introduces a new Control Message Option, the Via Information option (VIO). Like the TIO, the VIO MUST be preceded by one or more RPL Target options to which it applies. Unlike the TIO, the VIO are not factorized: multiple contiguous Via options indicate an ordered sequence of routers to reach the target(s), presented in the order of the packet stream, source to destination, and in which a routing state must be installed.

The Via Information option MUST contain at least one Via Address.

3.1. Via Information Option

The Via Information option MAY be present in DAO messages, and its format is as follows:

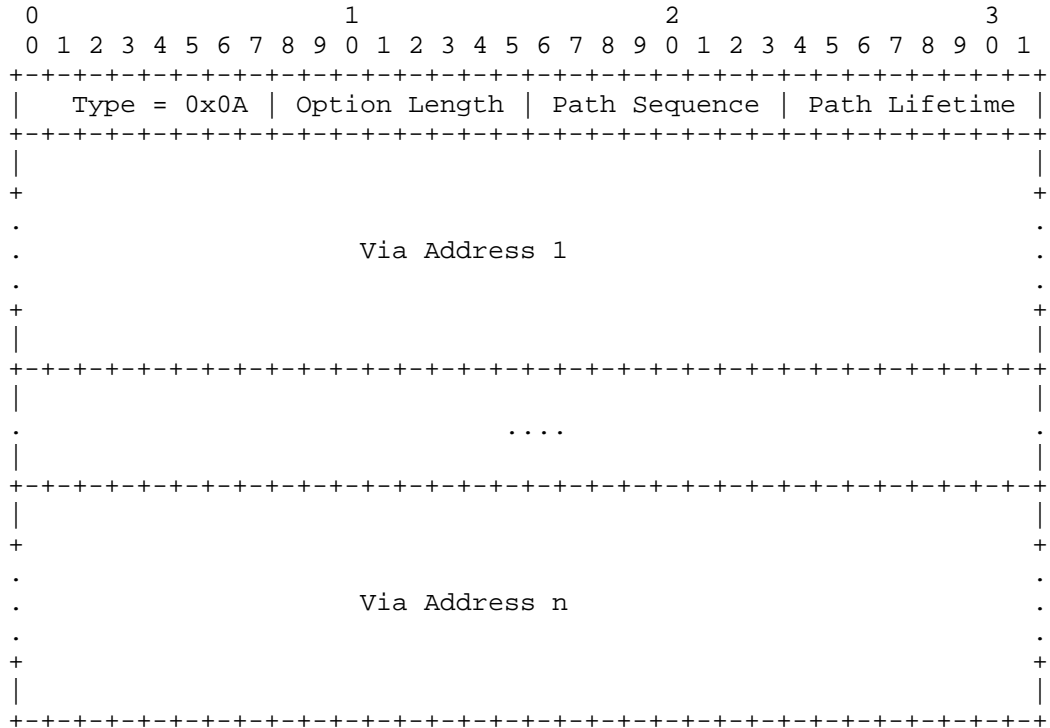


Figure 1: Via Information option format

Option Type: 0x0A (to be confirmed by IANA)

Option Length: In bytes; variable, depending on the number of Via Addresses.

Path Sequence: 8-bit unsigned integer. When a RPL Target option is issued by the root of the DODAG (i.e. in a DAO message), that root sets the Path Sequence and increments the Path Sequence each time it issues a RPL Target option with updated information. The indicated sequence deprecates any state for a given Target that was learned from a previous sequence and adds to any state that was learned for that sequence.

Path Lifetime: 8-bit unsigned integer. The length of time in Lifetime Units (obtained from the Configuration option) that the prefix is valid for route determination. The period starts when a new Path Sequence is seen. A value of all one bits (0xFF) represents infinity. A value of all zero bits (0x00) indicates a loss of reachability. A DAO message that contains a Via Information option with a Path Lifetime of 0x00 for a Target is referred as a No-Path (for that Target) in this document.

Via Address: 16 bytes. IPv6 Address of the next hop towards the destination(s) indicated in the target option that immediately precede the VIO. TBD: See how the /64 prefix can be elided if it is the same as that of (all of) the target(s). In that case, the Next-Hop Address could be expressed as the 8-bytes suffix only, otherwise it is expressed as 16 bytes, at least in storing mode.

4. Projected DAO

This draft adds a capability to RPL whereby the root projects a route through an extended DAO message called a Projected-DAO (P-DAO) to an arbitrary router down the DODAG, indicating a next hop or a sequence of routers via which a certain destination indicated in the Target Information option may be reached.

A P-DAO message MUST contain at least a Target Information option and at least one VIA Information option following it.

Like a classical DAO message, a P-DAO is processed only if it is "new" per section 9.2.2. "Generation of DAO Messages" of the RPL specification [RFC6550]; this is determined using the Path Sequence information from the VIO as opposed to a TIO. Also, a Path Lifetime of 0 in a VIO indicates that a route is to be removed.

There are two kinds of P-DAO, the storing mode and the non-storing mode ones.

The non-storing mode P-DAO discussed in section Section 4.1 has a single VIO with one or more Via Addresses in it, the list of Via Addresses indicating the source-routed path to the target to be installed in the router that receives the message, which replies to the root directly with a DAO-ACK message.

The storing mode P-DAO discussed in section Section 4.2 has at least two Via Information options with one Via Address each, for the ingress and the egress of the path, and more if there are intermediate routers. The Via Addresses indicate the routers in

which the routing state to the target have to be installed via the next Via Address in the sequence of VIO. In normal operations, the P-DAO is propagated along the chain of Via Routers from the egress router of the path till the ingress one, which confirms the installation to the root with a DAO-ACK message. Note that the root may be the ingress and it may be the egress of the path, that it can also be neither but it cannot be both.

The root is expected to use these mechanisms optimally and with required parsimony to limit the state installed in the devices to fit within their resources, but how the root figures the amount of resources that is available in each device is out of scope for this document.

In particular, the draft expects that the root has enough information about the capability for each node to store a number of routes, which can be discovered for instance using a Network Management System (NMS) and/or the RPL routing extensions specified in Routing for Path Calculation in LLNs [RFC6551].

A route that is installed by a P-DAO is not necessarily installed along the DODAG, though how the root and the optional PCE obtain the additional topological information to compute other routes is out of scope for this document

4.1. Non-storing Mode Projected DAO

As illustrated in Figure 2, the non-storing mode P-DAO enables the root to install a source-routed path towards a target in any particular router; with this path information the router can add a source routed header reflecting the path to any packet for which the current destination either is the said target or can be reached via the target, for instance a loose source routed packet for which the next loose hop is the target, or a packet for which the router has a routing state to the final destination via the target.

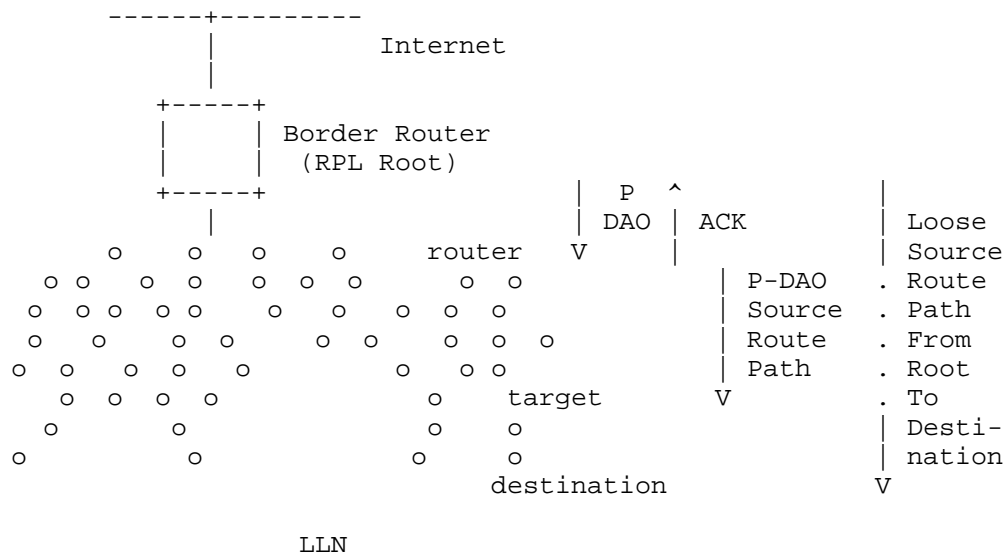


Figure 2: Projecting a non-storing route

A router that receives a non-storing P-DAO installs a source routed path towards each of the consecutive targets via a source route path indicated in the following VIO.

When forwarding a packet to a destination for which the router determines that routing happens via the target, the router inserts the source routing header in the packet to reach the target.

In order to do so, the router encapsulates the packet with an IP in IP header and a non-storing mode source routing header (SRH) [RFC6554].

In the uncompressed form the source of the packet would be self, the destination would be the first Via Address in the VIO, and the SRH would contain the list of the remaining Via Addresses and then the target.

In practice, the router will normally use the IPv6 over Low-Power Wireless Personal Area Network (6LoWPAN) Paging Dispatch [RFC8025] to compress the RPL artifacts as indicated in the 6LoWPAN Routing Header [I-D.ietf-roll-routing-dispatch] specification. In that case, the router indicates self as encapsulator in an IP-in-IP 6LoRH Header, and places the list of Via Addresses in the order of the VIO and then the target in the SRH 6LoRH Header.

4.2. Storing-Mode Projected DAO

As illustrated in Figure 3, the storing mode P-DAO enables the root to install a routing state towards a target in the routers along a segment between an ingress and an egress router; this enables the routers to forward along that segment any packet for which the next loose hop is the said target, for instance a loose source routed packet for which the next loose hop is the target, or a packet for which the router has a routing state to the final destination via the target.

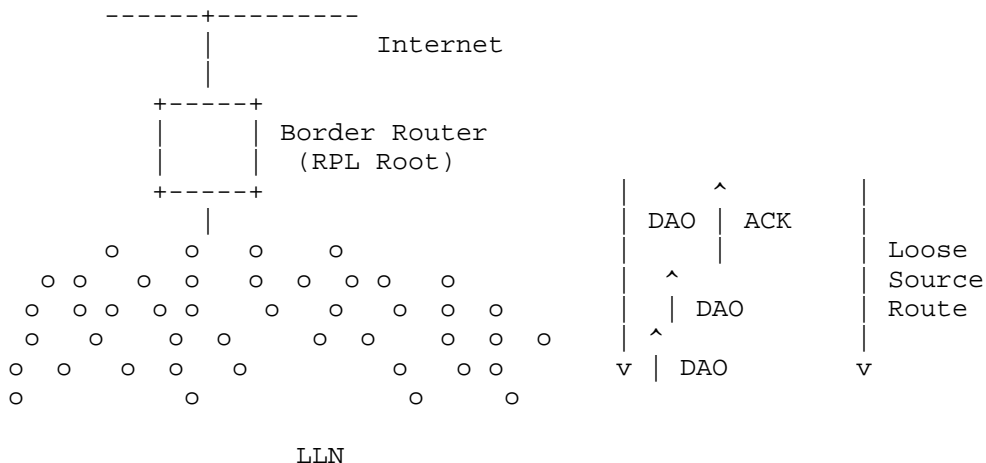


Figure 3: Projecting a route

Based on available topological, usage and capabilities node information, the root or an associated PCE computes which segment should be optimized and which relevant state should be installed in which nodes. The algorithm is out of scope but it is envisaged that the root could compute the ratio between the optimal path (existing path not traversing the root, and the current path), the application service level agreement (SLA) for specific flows that could benefit from shorter paths, the energy wasted in the network, local congestion on various links that would benefit from having flows routed along alternate paths.

In order to install the relevant routing state along the segment between an ingress and an egress routers, the root sends a unicast P-DAO message to the egress router of the routing segment that must be installed. The P-DAO message contains the ordered list of hops along the segment as a direct sequence of Via Information options that are preceded by one or more RPL Target options to which they

relate. Each Via Information option contains a Path Lifetime for which the state is to be maintained.

The root sends the P-DAO directly to the egress node of the segment, which in that P-DAO, the destination IP address matches the Via Address in the last VIO. This is how the egress recognizes its role. In a similar fashion, the ingress node recognizes its role as it matches Via Address in the first VIO.

The egress node of the segment is the only node in the path that does not install a route in response to the P-DAO; it is expected to be already able to route to the target(s) on its own. It may either be the target, or may have some existing information to reach the target(s), such as a connected route or an already installed projected route. If one of the targets cannot be located, the node MUST answer to the root with a negative DAO-ACK listing the target(s) that could not be located (suggested status 10 to be confirmed by IANA).

If the egress node can reach all the targets, then it forwards the P-DAO with unchanged content to its loose predecessor in the segment as indicated in the list of Via Information options, and recursively the message is propagated unchanged along the sequence of routers indicated in the P-DAO, but in the reverse order, from egress to ingress.

The address of the predecessor to be used as destination of the propagated DAO message is found in the Via Information option that precedes the one that contains the address of the propagating node, which is used as source of the packet.

Upon receiving a propagated DAO, an intermediate router as well as the ingress router install a route towards the DAO target(s) via its successor in the P-DAO; the router locates the VIO that contains its address, and uses as next hop the address found in the Via Address field in the following VIO. The router MAY install additional routes towards the addresses that are located in VIOs that are after the next one, if any, but in case of a conflict or a lack of resource, a route to a target installed by the root has precedence.

The process recurses till the P-DAO is propagated to ingress router of the segment, which answers with a DAO-ACK to the root.

Also, the path indicated in a P-DAO may be loose, in which case the reachability to the next hop has to be asserted. Each router along the path indicated in a P-DAO is expected to be able to reach its successor, either with a connected route (direct neighbor), or by routing, for instance following a route installed previously by a DAO

or a P-DAO message. If that route is not connected then a recursive lookup may take place at packet forwarding time to find the next hop to reach the target(s). If it does not and cannot reach the next router in the P-DAO, the router MUST answer to the root with a negative DAO-ACK indicating the successor that is unreachable (suggested status 11 to be confirmed by IANA).

A Path Lifetime of 0 in a Via Information option is used to clean up the state. The P-DAO is forwarded as described above, but the DAO is interpreted as a No-Path DAO and results in cleaning up existing state as opposed to refreshing an existing one or installing a new one.

5. Applications

5.1. Loose Source Routing in Non-storing Mode

A RPL implementation operating in a very constrained LLN typically uses the non-storing mode of operation whereby a RPL node indicates a parent-child relationship to the root, using a Destination Advertisement Object (DAO) that is unicast from the node directly to the root, and the root typically builds a source routed path to a destination down the DODAG by recursively concatenating this information.

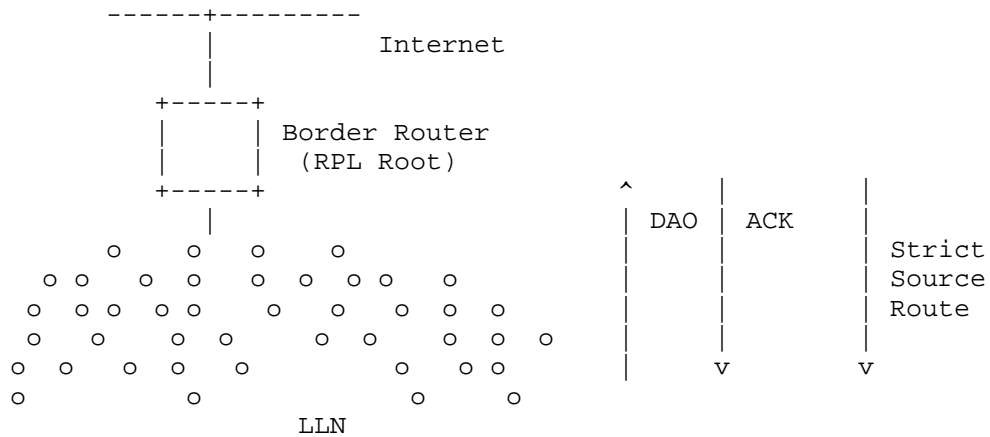


Figure 4: RPL non-storing mode of operation

Based on the parent-children relationships expressed in the non-storing DAO messages, the root possesses topological information about the whole network, though this information is limited to the structure of the DODAG for which it is the destination. A packet that is generated within the domain will always reach the root, which

can then apply a source routing information to reach the destination if the destination is also in the DODAG. Similarly, a packet coming from the outside of the domain for a destination that is expected to be in a RPL domain reaches the root.

It results that the root, or then some associated centralized computation engine such as a PCE, can determine the amount of packets that reach a destination in the RPL domain, and thus the amount of energy and bandwidth that is wasted for transmission, between itself and the destination, as well as the risk of fragmentation, any potential delays because of a paths longer than necessary (shorter paths exist that would not traverse the root).

As a network gets deep, the size of the source routing header that the root must add to all the downward packets becomes an issue for nodes that are many hops away. In some use cases, a RPL network forms long lines and a limited amount of well-targeted routing state would allow to make the source routing operation loose as opposed to strict, and save packet size. Limiting the packet size is directly beneficial to the energy budget, but, mostly, it reduces the chances of frame loss and/or packet fragmentation, which is highly detrimental to the LLN operation. Because the capability to store a routing state in every node is limited, the decision of which route is installed where can only be optimized with a global knowledge of the system, a knowledge that the root or an associated PCE may possess by means that are outside of the scope of this specification.

This specification enables to store source-routed or storing mode state in intermediate routers, which enables to limit the excursion of the source route headers in deep networks. Once a P-DAO exchange has taken place for a given target, if the root operates in non storing mode, then it may elide the sequence of routers that is installed in the network from its source route headers to destination that are reachable via that target, and the source route headers effectively become loose.

5.2. Transversal Routes in storing and non-storing modes

RPL is optimized for Point-to-Multipoint (P2MP), root to leaves and Multipoint-to-Point (MP2P) leaves to root operations, whereby routes are always installed along the RPL DODAG. Transversal Peer to Peer (P2P) routes in a RPL network will generally suffer from some stretch since routing between 2 peers always happens via a common parent, as illustrated in Figure 5:

- o in non-storing mode, all packets routed within the DODAG flow all the way up to the root of the DODAG. If the destination is in the same DODAG, the root must encapsulate the packet to place a

Routing Header that has the strict source route information down the DODAG to the destination. This will be the case even if the destination is relatively close to the source and the root is relatively far off.

- o In storing mode, unless the destination is a child of the source, the packets will follow the default route up the DODAG as well. If the destination is in the same DODAG, they will eventually reach a common parent that has a route to the destination; at worse, the common parent may also be the root. From that common parent, the packet will follow a path down the DODAG that is optimized for the Objective Function that was used to build the DODAG.

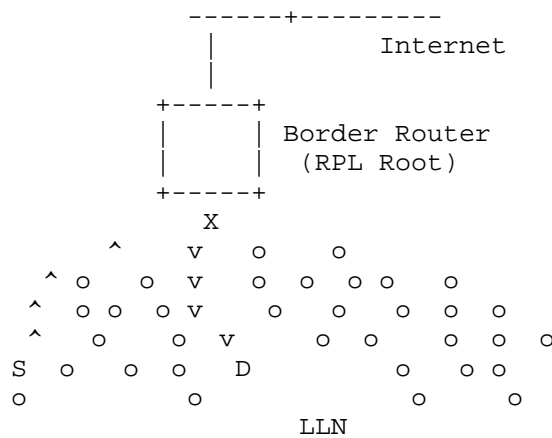


Figure 5: Routing Stretch between S and D via common parent X

It results that it is often beneficial to enable transversal P2P routes, either if the RPL route presents a stretch from shortest path, or if the new route is engineered with a different objective. For that reason, earlier work at the IETF introduced the Reactive Discovery of Point-to-Point Routes in Low Power and Lossy Networks [RFC6997], which specifies a distributed method for establishing optimized P2P routes. This draft proposes an alternate based on a centralized route computation.

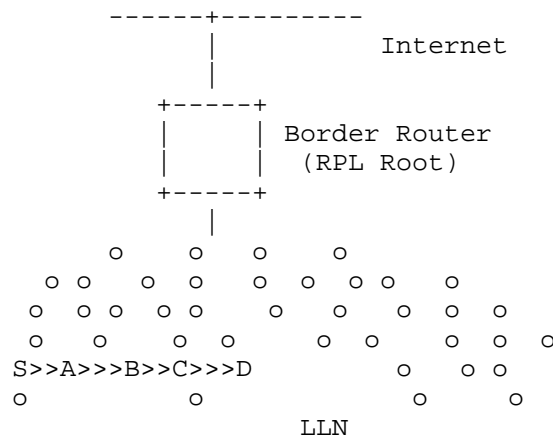


Figure 6: Projected Transversal Route

This specification enables to store source-routed or storing mode state in intermediate routers, which enables to limit the stretch of a P2P route and maintain the characteristics within a given SLA. An example of service using this mechanism could be a control loop that would be installed in a network that uses classical RPL for asynchronous data collection. In that case, the P2P path may be installed in a different RPL Instance, with a different objective function.

6. RPL Instances

It must be noted that RPL has a concept of instance but does not have a concept of an administrative distance, which exists in certain proprietary implementations to sort out conflicts between multiple sources. This draft conforms the instance model as follows:

- o if the PCE needs to influence a particular instance to add better routes in conformance with the routing objectives in that instance, it may do so. When the PCE modifies an existing instance then the added routes must not create a loop in that instance. This is achieved by always preferring a route obtained from the PCE over a route that is learned via RPL.
- o If the PCE installs a more specific (Traffic Engineering) route between a particular pair of nodes then it should use a Local Instance from the ingress node of that path. Only packets associated with that instance will be routed along that path.

In all cases, the path is indicated by a new Via Information option, and the flow is similar to the flow used to obtain loose source routing.

7. Security Considerations

This draft uses messages that are already present in [RFC6550] with optional secured versions. The same secured versions may be used with this draft, and whatever security is deployed for a given network also applies to the flows in this draft.

8. IANA Considerations

This document updates the IANA registry for the Mode of Operation (MOP)

4: Non-Storing with Projected routes [this]

This document updates IANA registry for the RPL Control Message Options

0x0A: Via descriptor [this]

9. Acknowledgments

The authors wish to acknowledge JP Vasseur and Patrick Wetterwald for their contributions to the ideas developed here.

10. References

10.1. Normative References

- [I-D.ietf-roll-routing-dispatch]
Thubert, P., Bormann, C., Toutain, L., and R. Cragie,
"6LoWPAN Routing Header", draft-ietf-roll-routing-
dispatch-05 (work in progress), October 2016.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
Requirement Levels", BCP 14, RFC 2119,
DOI 10.17487/RFC2119, March 1997,
<<http://www.rfc-editor.org/info/rfc2119>>.
- [RFC6550] Winter, T., Ed., Thubert, P., Ed., Brandt, A., Hui, J.,
Kelsey, R., Levis, P., Pister, K., Struik, R., Vasseur,
JP., and R. Alexander, "RPL: IPv6 Routing Protocol for
Low-Power and Lossy Networks", RFC 6550,
DOI 10.17487/RFC6550, March 2012,
<<http://www.rfc-editor.org/info/rfc6550>>.

- [RFC6551] Vasseur, JP., Ed., Kim, M., Ed., Pister, K., Dejean, N., and D. Barthel, "Routing Metrics Used for Path Calculation in Low-Power and Lossy Networks", RFC 6551, DOI 10.17487/RFC6551, March 2012, <<http://www.rfc-editor.org/info/rfc6551>>.
- [RFC6554] Hui, J., Vasseur, JP., Culler, D., and V. Manral, "An IPv6 Routing Header for Source Routes with the Routing Protocol for Low-Power and Lossy Networks (RPL)", RFC 6554, DOI 10.17487/RFC6554, March 2012, <<http://www.rfc-editor.org/info/rfc6554>>.
- [RFC8025] Thubert, P., Ed. and R. Cragie, "IPv6 over Low-Power Wireless Personal Area Network (6LoWPAN) Paging Dispatch", RFC 8025, DOI 10.17487/RFC8025, November 2016, <<http://www.rfc-editor.org/info/rfc8025>>.

10.2. Informative References

- [I-D.ietf-6tisch-architecture]
Thubert, P., "An Architecture for IPv6 over the TSCH mode of IEEE 802.15.4", draft-ietf-6tisch-architecture-11 (work in progress), January 2017.
- [I-D.ietf-detnet-architecture]
Finn, N. and P. Thubert, "Deterministic Networking Architecture", draft-ietf-detnet-architecture-00 (work in progress), September 2016.
- [PCE] IETF, "Path Computation Element", <<https://datatracker.ietf.org/doc/charter-ietf-pce/>>.
- [RFC6997] Goyal, M., Ed., Baccelli, E., Philipp, M., Brandt, A., and J. Martocci, "Reactive Discovery of Point-to-Point Routes in Low-Power and Lossy Networks", RFC 6997, DOI 10.17487/RFC6997, August 2013, <<http://www.rfc-editor.org/info/rfc6997>>.
- [RFC7102] Vasseur, JP., "Terms Used in Routing for Low-Power and Lossy Networks", RFC 7102, DOI 10.17487/RFC7102, January 2014, <<http://www.rfc-editor.org/info/rfc7102>>.

Appendix A. Examples

A.1. Using storing mode P-DAO in non-storing mode MOP

In non-storing mode, the DAG root maintains the knowledge of the whole DODAG topology, so when both the source and the destination of a packet are in the DODAG, the root can determine the common parent that would have been used in storing mode, and thus the list of nodes in the path between the common parent and the destination. For instance in the diagram shown in Figure 7, if the source is node 41 and the destination is node 52, then the common parent is node 22.

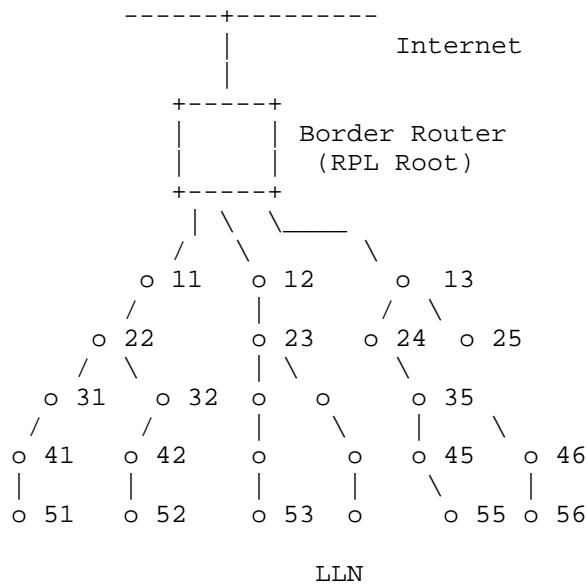


Figure 7: Example DODAG forming a logical tree topology

With this draft, the root can install a storing mode routing states along a segment that is either from itself to the destination, or from one or more common parents for a particular source/destination pair towards that destination (in this particular example, this would be the segment made of nodes 22, 32, 42).

In the example below, say that there is a lot of traffic to nodes 55 and 56 and the root decides to reduce the size of routing headers to those destinations. The root can first send a DAO to node 45 indicating target 55 and a Via segment (35, 45), as well as another DAO to node 46 indicating target 56 and a Via segment (35, 46). This will save one entry in the routing header on both sides. The root may then send a DAO to node 35 indicating targets 55 and 56 a Via segment (13, 24, 35) to fully optimize that path.

Alternatively, the root may send a DAO to node 45 indicating target 55 and a Via segment (13, 24, 35, 45) and then a DAO to node 46 indicating target 56 and a Via segment (13, 24, 35, 46), indicating the same DAO Sequence.

A.2. Projecting a storing-mode transversal route

In this example, say that a PCE determines that a path must be installed between node S and node D via routers A, B and C, in order to serve the needs of a particular application.

The root sends a P-DAO with a target option indicating the destination D and a sequence Via Information option, one for S, which is the ingress router of the segment, one for A and then for B, which are an intermediate routers, and one for C, which is the egress router.

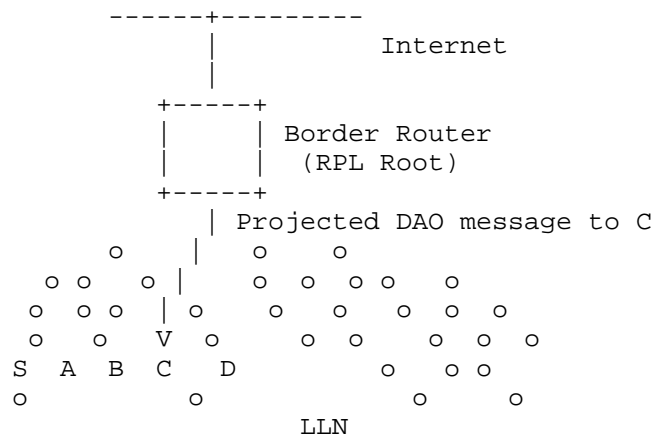


Figure 8: Projected DAO from root

Upon reception of the P-DAO, C validates that it can reach D, e.g. using IPv6 Neighbor Discovery, and if so, propagates the P-DAO unchanged to B.

B checks that it can reach C and of so, installs a route towards D via C. Then it propagates the P-DAO to A.

The process recurses till the P-DAO reaches S, the ingress of the segment, which installs a route to D via A and sends a DAO-ACK to the root.

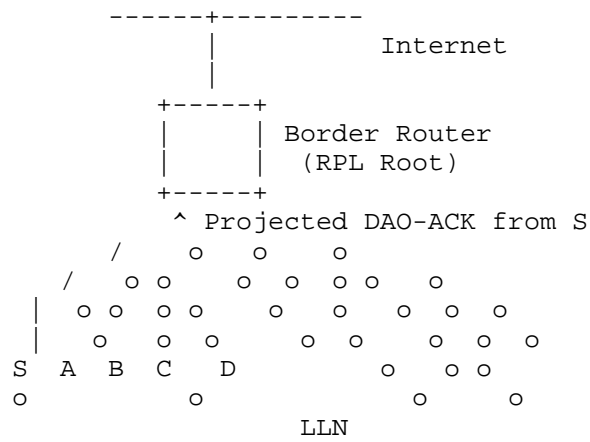


Figure 9: Projected DAO-ACK to root

As a result, a transversal route is installed that does not need to follow the DODAG structure.

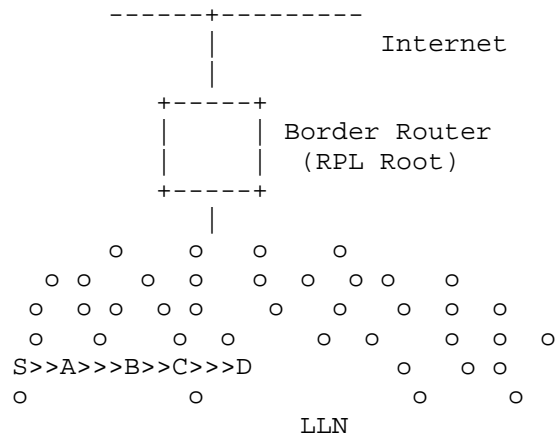


Figure 10: Projected Transversal Route

Authors' Addresses

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

Phone: +33 4 97 23 26 34
Email: pthubert@cisco.com

James Pylakutty
Cisco Systems
Cessna Business Park
Kadubeesanahalli
Marathalli ORR
Bangalore, Karnataka 560087
INDIA

Phone: +91 80 4426 4140
Email: mundenma@cisco.com

ROLL
Internet-Draft
Intended status: Standards Track
Expires: August 18, 2017

R. Jadhav, Ed.
R. Sahoo
Z. Cao
Huawei Tech
February 14, 2017

No-Path DAO modifications
draft-jadhav-roll-efficient-npdao-00

Abstract

This document describes the problems associated with the use of No-Path DAO messaging in RPL and a signaling improvement to improve route invalidation efficiency.

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 August 18, 2017.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the 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

1.	Introduction	2
1.1.	Requirements Language and Terminology	3
1.2.	Current No-Path DAO messaging	3
1.3.	Cases when No-Path DAO may be used	4
1.4.	Why No-Path DAO is important?	5
2.	Problems with current No-Path DAO messaging	5
2.1.	Lost NP-DAO due to link break to the previous parent	5
2.2.	Invalidate routes to dependent nodes of the switching node	5
2.3.	Route downtime caused by asynchronous operation of NPDAO and DAO	6
3.	Requirements for the No-Path DAO Optimization	6
3.1.	Req#1: Tolerant to the link failures to the previous parents	6
3.2.	Req#2: Dependent nodes route invalidation on parent switching	6
3.3.	Req#3: No impact on traffic while NP-DAO operation in progress	7
4.	Existing Solution	7
4.1.	NP-DAO can be generated by the parent node who detects link failure to the child	7
4.2.	NP-DAO can be generated once the link is restored to the previous parent	8
5.	Proposed changes to NPDAO signaling	8
5.1.	Change in NPDAO semantics	8
5.2.	DAO message format changes	9
5.2.1.	Path Sequence number in the reverse NPDAO	11
5.3.	Example messaging	11
5.4.	Other considerations	12
5.4.1.	Dependent Nodes invalidation	12
6.	Acknowledgements	12
7.	IANA Considerations	12
8.	Security Considerations	13
9.	References	13
9.1.	Normative References	13
9.2.	Informative References	13
Appendix A.	Additional Stuff	14
Authors' Addresses	14

1. Introduction

RPL [RFC6550] specifies a proactive distance-vector based routing scheme. The specification has an optional messaging in the form of DAO messages using which the 6LBR can learn route towards any of the nodes. In storing mode, DAO messages would result in routing entries

been created on all intermediate hops from the node's parent all the way towards the 6LBR.

RPL also allows use of No-Path DAO (NPDAO) messaging to invalidate a routing path and thus releasing of any resources utilized on that path. A No-Path DAO is a DAO message with route lifetime of zero, signaling route invalidation for the given target. This document studies the problems associated with the current use of No-Path DAO messaging, which creates route inefficiency and inconsistency. This document also discusses the requirements for an optimized No-Path DAO messaging scheme.

1.1. Requirements Language and Terminology

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].

The document only caters to the RPL's storing mode of operation (MOP). The non-storing mode does not require use of NPDAO for route invalidation since routing entries are not maintained on 6LRs in case of non-storing MOP.

Common Ancestor node: 6LR node which is the first common node on the old and new path for the child node.

Current parent: Parent 6LR node before switching to the new path.

New parent: Parent 6LR node after switching to the new path.

NPDAO: No-Path DAO. A DAO message which has target with lifetime 0.

Reverse NPDAO: A No-Path DAO message which traverses downstream in the network.

Regular DAO: A DAO message with non-zero lifetime.

This document also uses terminology described in [RFC6550] and [RFC6775].

1.2. Current No-Path DAO messaging

RPL introduced No-Path DAO messaging in the storing mode so that the node switching its current parent can inform its parents and ancestors to invalidate the existing route. Subsequently parents or ancestors would release any resources (such as the routing entry) it maintains on behalf of that child node. The No-Path DAO message

always traverses the RPL tree in upward direction, originating at the target node itself.

For the rest of this document consider the following topology:

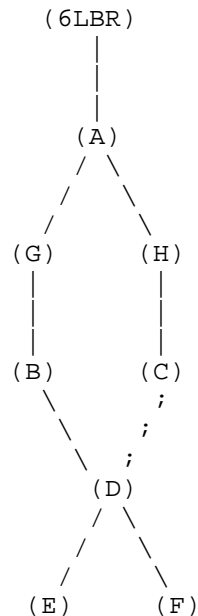


Figure 1: Sample topology

Node (D) is connected via preferred parent (B). (D) has an alternate path via (C) towards the BR. Node (A) is the common ancestor for (D) for paths through (B)-(G) and (C)-(H). When (D) switches from (B) to (C), [RFC6550] suggests sending No-Path DAO to (B) and regular DAO to (C).

1.3. Cases when No-Path DAO may be used

There are following cases in which a node switches its parent and may employ No-Path DAO messaging:

Case I: Current parent becomes unavailable because of transient or permanent link or parent node failure.

Case II: The node finds a better parent node i.e. the metrics of another parent is better than its current parent.

Case III: The node switches to a new parent whom it "thinks" has a better metric but does not in reality.

The usual steps of operation when the node switches the parent is that the node sends a No-Path DAO message via its current parent to invalidate its current route and subsequently it tries to establish a new routing path by sending a new DAO via its new parent.

1.4. Why No-Path DAO is important?

Nodes in LLNs may be resource constrained. There is limited memory available and routing entry records are the one of the primary elements occupying dynamic memory in the nodes. Route invalidation helps 6LR nodes to decide which entries could be discarded to better achieve resource utilization in case of contention. Thus it becomes necessary to have efficient route invalidation mechanism. Also note that a single parent switch may result in a "sub-tree" switching from one parent to another. Thus the route invalidation needs to be done on behalf of the sub-tree and not the switching node alone. In the above example, when Node (D) switches parent, the route invalidation needs to be done for (D), (E) and (F). Thus without efficient route invalidation, a 6LR may have to hold a lot of unwanted route entries.

2. Problems with current No-Path DAO messaging

There are following problems with the usage of current NP-DAO messaging

2.1. Lost NP-DAO due to link break to the previous parent

When the node switches its parent, the NPDAO is to be sent via its previous parent and a regular DAO via its new parent. In cases where the node switches its parent because of transient or permanent parent link/node failure then the NPDAO message is bound to fail. RPL assumes communication link with the previous parent for No-Path DAO messaging.

RPL mentions use of route lifetime to remove unwanted routes in case the routes could not be refreshed. But route lifetimes in case of LLNs could be substantially high and thus the route entries would be stuck for long.

2.2. Invalidate routes to dependent nodes of the switching node

No-path DAO is sent by the node who has switched the parent but it does not work for the dependent child nodes below it. The specification does not specify how route invalidation will work for sub-children, resulting in stale routing entries on behalf of the sub-children on the previous route. The only way for 6LR to invalidate the route entries for dependent nodes would be to use route lifetime expiry which could be substantially high for LLNs. In the example

topology, when Node (D) switches its parent, Node (D) generates an NPDAO on its behalf. Post switching, Node (D) transmits a DIO with incremented DTSN so that child nodes, node (E) and (F), generate DAOs to trigger route update on the new path for themselves. There is no NPDAO generated by these child nodes through the previous path resulting in stale entries on nodes (B) and (G) for nodes (E) and (F).

2.3. Route downtime caused by asynchronous operation of NPDAO and DAO

A switching node may generate both an NPDAO and DAO via two different paths at almost the same time. There is a possibility that an NPDAO generated may invalidate the previous route and the regular DAO sent via the new path gets lost on the way. This may result in route downtime thus impacting downward traffic for the switching node. In the example topology, consider Node (D) switches from parent (B) to (C) because the metrics of the path via (C) are better. Note that the previous path via (B) may still be available (albeit at relatively bad metrics). An NPDAO sent from previous route may invalidate the existing route whereas there is no way to determine whether the new DAO has successfully updated the route entries on the new path.

An implementation technique to avoid this problem is to further delay the route invalidation by a fixed time interval after receiving an NPDAO, considering the time taken for the new path to be established. Coming up with such a time interval is tricky since the new route may also not be available and it may subsequently require more parent switches to establish a new path.

3. Requirements for the No-Path DAO Optimization

We identify the following requirements for the NP-DAO optimization.

3.1. Req#1: Tolerant to the link failures to the previous parents

When the switching node send the NP-DAO message to the previous parent, it is normal that the link to the previous parent is prone to failure. Therefore, it is required that the NP-DAO message MUST be tolerant to the link failure during the switching.

3.2. Req#2: Dependent nodes route invalidation on parent switching

While switching the parent node and sending NP-DAO message, it is required that the routing entries to the dependent nodes of the switching node will be updated accordingly on the previous parents and other relevant upstream nodes.

3.3. Req#3: No impact on traffic while NP-DAO operation in progress

While sending the NP-DAO and DAO messages, it is possible that the NP-DAO successfully invalidates the previous path, while the newly sent DAO gets lost (new path not set up successfully). This will result into downstream unreachability to the current switching node. Therefore, it is desirable that the NP-DAO is synchronized with the DAO to avoid the risk of routing downtime.

4. Existing Solution

4.1. NP-DAO can be generated by the parent node who detects link failure to the child

RPL states mechanisms which could be utilized to clear DAO states in a sub-DODAG. [RFC6550] Section 11.2.2.3 states "With DAO inconsistency loop recovery, a packet can be used to recursively explore and clean up the obsolete DAO states along a sub-DODAG".

Thus in the sample topology in Figure 1, when Node (B) detects link failure to (D), (B) has an option of generating an NP-DAO on behalf of Node (D) and its sub-children, (E) and (F).

This section explains why generation of an NP-DAO in such cases may not function as desired. Primarily the DAO state information in the form of Path Sequence plays a major role here. Every target is associated with a Path Sequence number which relates to the latest state of the target. [RFC6550] Section 7.1 explains the semantics of Path Sequence number. The target node increments the Path Sequence number every time it generates a new DAO. The router nodes en-route utilize this Path Sequence number to decide the freshness of target information. If a non-target node has to generate an NP-DAO then it could use following two possibilities with Path Sequence number:

Let the Path Sequence number of old regular DAO that flowed through (B) be x . The subsequent regular DAO generated by Node (D) will have sequence number $x+1$.

i. Node (B) uses the previous Path Sequence number from the regular DAO i.e. NP-DAO(pathseq= x)

ii. Node (B) increments the Path Sequence number i.e. NP-DAO(pathseq= $x+1$)

In case i, the NP-DAO(pathseq= x) will be dropped by all the intermediate nodes since the semantics of Path Sequence number dictates that any DAO with an older Path Sequence number be dropped.

In case ii, there is a risk that the NP-DAO(pathseq=x+1) traverses up the DODAG and invalidates all the routes till the root and then the regular DAO(pathseq=x+1) from the target traverses upwards. In this case the regular DAO(pathseq=x+1) will be dropped from common ancestor node to the root. This will result in route downtime.

Another problem with this scheme is its dependence on the upstream neighbor to detect that the downstream neighbor is unavailable. There are two possibilities by which such a detection might be put to work:

i. There is P2P traffic from the previous sub-DODAG to any of nodes in the sub-tree which has switched the path. In the above example, lets consider that Node (G) has P2P traffic for either of nodes (D), (E), or (F). In this case, Node (B) will detect forwarding error while forwarding the packets from Node (B) to (D). But dependence on P2P traffic may not be an optimal way to solve this problem considering the reactive approach of the scheme. The P2P traffic pattern might be sparse and thus such a detection might kick-in too late.

ii. The other case is where Node (B) explicitly employs some mechanism to probe directly attached downstream child nodes. Such kind of schemes are seldom used.

4.2. NP-DAO can be generated once the link is restored to the previous parent

This scheme solves a specific scenario of transient links. The child node can detect that the connection to previous parent is restored and then transmit an NP-DAO to the previous parent to invalidate the route. This scheme is stateful, thus requires more memory and solves a specific scenario.

5. Proposed changes to NPDAO signaling

5.1. Change in NPDAO semantics

As described in Section 1.2, currently the NPDAO originates at the node switching the parent and traverses upstream towards the root. In order to solve the problems as mentioned in Section 2, the draft proposes to change the way NPDAO originates and traverses the network. The new NPDAO proposed does not originate at the node but instead originates at a common ancestor node between the new and old path. The trigger for the common ancestor node to generate this NPDAO is the change in the next hop for the node on reception of an update message in the form of regular DAO for the target.

In the Figure 1, when node D decides to switch the path from B to C, it sends a regular DAO to node C with reachability information containing target as address of D and a incremented path sequence number. Node C will update the routing table based on the reachability information in DAO and in turn generate another DAO with the same reachability information and forward it to H. Node H also follows the same procedure as Node C and forwards it to node A. When node A receives the regular DAO, it finds that it already has a routing table entry on behalf of the target address of node D. It finds however that the next hop information for reaching node D has changed i.e. the node D has decided to change the paths. In this case, Node A which is the common ancestor node for node D along the two paths (previous and new), may generate an NPDAO which traverses downwards in the network. The document in the subsequent section will explain the message format changes to handle this downward flow of NPDAO.

5.2. DAO message format changes

Every RPL message is divided into base message fields and additional Options. The base fields apply to the message as a whole and options are appended to add message/use-case specific attributes. As an example, a DAO message may be attributed by one or more "RPL Target" options which specifies the reachability information for the given targets. Similarly, a Transit Information option may be associated with a set of RPL Target options.

The draft proposes a change in DAO message to contain "Invalidate previous route" (I) bit. This I-bit which is carried in regular DAO message, signals the common ancestor node to generate a downstream NPDAO on behalf of the target node. The I-bit is carried in the transit container option which augments the reachability information for a given set of RPL Target(s).

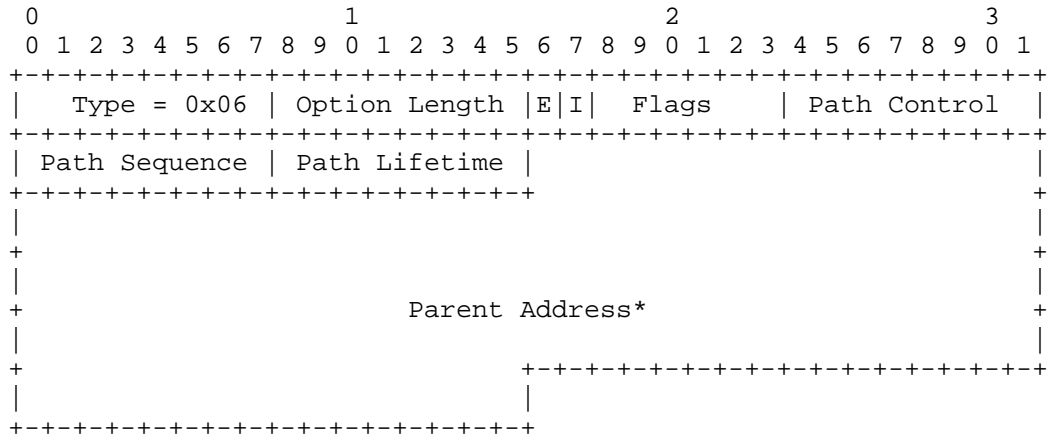


Figure 2: Updated Transit Information Option (New I flag added)

I (Invalidate previous route) bit: 1 bit flag. The 'I' flag is set by the target node to indicate that it wishes to invalidate the previous route by a common ancestor node between the two paths.

The NPDAO thus generated by the common ancestor node needs to traverse downstream. An additional flag called as "Reverse NPDAO" (R) is added in the base DAO object to signal this change.

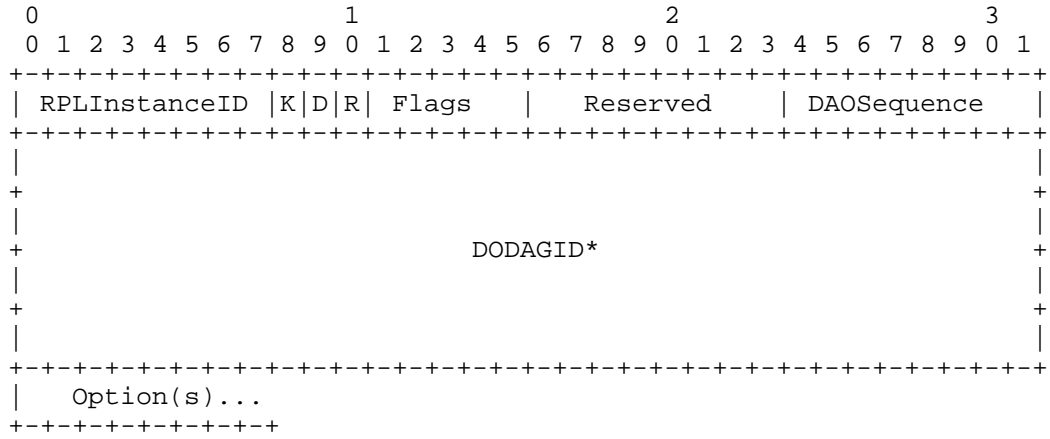


Figure 3: Updated DAO base object (New R flag added)

R (Reverse DAO) bit: 1 bit flag. The 'R' flag is used to signal that the DAO traverses downwards.

5.2.1. Path Sequence number in the reverse NPDAO

Every DAO message may contain a Path Sequence in the transit information option to identify the freshness of the DAO message. The Path Sequence in the downward NPDAO generated by common ancestor should use the same Path Sequence number present in the regular DAO message.

5.3. Example messaging

In Figure 1, node (D) switches its parent from (B) to (C). The sequence of actions is as follows:

1. Node D switches its parent from node B to node C
2. D sends a regular DAO($\text{tgt}=\text{D}$, $\text{pathseq}=\text{x}+1$, $\text{I_flag}=1$) in the updated path to C
3. C checks for routing entry on behalf of D, since it cannot find an entry on behalf of D it creates a new routing entry and forwards the reachability information of the target D to H in a DAO.
4. Similar to C, node H checks for routing entry on behalf of D, cannot find an entry and hence creates a new routing entry and forwards the reachability information of the target D to H in a DAO.
5. A receives the DAO, and checks for routing entry on behalf of D. It finds a routing entry but checks that the next hop for target D is now changed. Node A checks the I_flag and generates downstream NPDAO($\text{tgt}=\text{D}$, $\text{pathseq}=\text{x}+1$, $\text{R_flag}=1$) to previous next hop for target D which is G. Subsequently, A updates the routing entry and forwards the reachability information of target D upstream DAO($\text{tgt}=\text{D}$, $\text{pathseq}=\text{x}+1$, $\text{I_flag}=\text{x}$) (the I_flag carries no significance henceforth).
6. Node G receives the downstream NPDAO and invalidates routing entry of target D and then checks the reverse (R) flag and forwards the (un)reachability information downstream to B.
7. Similarly, B processes the downstream NPDAO by invalidating the routing entry of target D and then checks the reverse (R) flag and forwards the (un)reachability information downstream to D.
8. D ignores the downstream NPDAO since the target is itself.

5.4. Other considerations

5.4.1. Dependent Nodes invalidation

Current RPL [RFC6550] does not provide a mechanism for route invalidation for dependent nodes.

This section describes approaches for invalidating routes of dependent nodes if the implementation chooses to solve this problem. The common ancestor node realizes that the paths for dependent nodes have changed (based on next hop change) when it receives a regular DAO on behalf of the dependent nodes. Thus dependent nodes route invalidation can be handled in the same way as the switching node. Note that there is no way that dependent nodes can set the `I_flag` in the DAO message selectively since they are unaware that their parent/grand parent node is switching paths. There are two ways to handle dependent node route invalidation:

1. One way to resolve is that the common ancestor does not depend upon the `I_flag` to generate the reverse NPDAO. The only factor it makes the decision will be based on `next_hop` change for an existing target to generate the NPDAO. Thus when the switching nodes and all the below dependent nodes advertise a regular DAO, the common ancestor node will detect a change in next hop and generate NPDAO for the same target as in the regular DAO.
2. Another way is that the nodes always set the `I_flag` whenever they send regular DAO. Thus common ancestor will first check whether `I_flag` is set and then check whether the `next_hop` has changed and subsequently trigger NPDAO if required.

This document recommends the approach in point 2. The advantage with `I_flag` is that the generation of downstream NPDAO is still controlled by the target node and thus is still in control of its own routing state.

6. Acknowledgements

We would like to thank Cenk Gundogan, Simon Duquennoy and Pascal Thubert for their review and insightful comments.

7. IANA Considerations

IANA is requested to allocate bit 11 in the DAO base object defined in RPL [RFC6550] section 6.4 for reverse 'R' NPDAO flag.

IANA is requested to allocate bit 18 in the Transit Information Option defined in RPL [RFC6550] section 6.7.8 for Invalidate route 'I' flag.

8. Security Considerations

TBA

9. References

9.1. Normative References

- [CONTIKI] Thingsquare, "Contiki: The Open Source OS for IoT", 2012, <<http://www.contiki-os.org>>.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<http://www.rfc-editor.org/info/rfc2119>>.

9.2. Informative References

- [RFC3552] Rescorla, E. and B. Korver, "Guidelines for Writing RFC Text on Security Considerations", BCP 72, RFC 3552, DOI 10.17487/RFC3552, July 2003, <<http://www.rfc-editor.org/info/rfc3552>>.
- [RFC5191] Forsberg, D., Ohba, Y., Ed., Patil, B., Tschofenig, H., and A. Yegin, "Protocol for Carrying Authentication for Network Access (PANA)", RFC 5191, DOI 10.17487/RFC5191, May 2008, <<http://www.rfc-editor.org/info/rfc5191>>.
- [RFC6345] Duffy, P., Chakrabarti, S., Cragie, R., Ohba, Y., Ed., and A. Yegin, "Protocol for Carrying Authentication for Network Access (PANA) Relay Element", RFC 6345, DOI 10.17487/RFC6345, August 2011, <<http://www.rfc-editor.org/info/rfc6345>>.
- [RFC6550] Winter, T., Ed., Thubert, P., Ed., Brandt, A., Hui, J., Kelsey, R., Levis, P., Pister, K., Struik, R., Vasseur, JP., and R. Alexander, "RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks", RFC 6550, DOI 10.17487/RFC6550, March 2012, <<http://www.rfc-editor.org/info/rfc6550>>.

[RFC6775] Shelby, Z., Ed., Chakrabarti, S., Nordmark, E., and C. Bormann, "Neighbor Discovery Optimization for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs)", RFC 6775, DOI 10.17487/RFC6775, November 2012, <<http://www.rfc-editor.org/info/rfc6775>>.

Appendix A. Additional Stuff

This becomes an Appendix.

Authors' Addresses

Rahul Arvind Jadhav (editor)
Huawei Tech
Kundalahalli Village, Whitefield,
Bangalore, Karnataka 560037
India

Phone: +91-080-49160700
Email: rahul.ietf@gmail.com

Rabi Narayan Sahoo
Huawei Tech
Kundalahalli Village, Whitefield,
Bangalore, Karnataka 560037
India

Phone: +91-080-49160700
Email: rabinarayans@huawei.com

Zhen Cao
Huawei Tech
W Chang'an Ave
Beijing 560037
China

Email: zhencao.ietf@gmail.com

roll
Internet-Draft
Intended Status: Standards Track
Expires: August 10, 2017

M.Qasem
A.Al-Dubai
I.Romdhani
B.Ghaleb
Edinburgh Napier University
February 6, 2017

Load Balancing Objective Function in RPL
draft-qasem-roll-rpl-load-balancing-00

Abstract

This document proposes an extended Objective Function(OF) that balances the number of children nodes of the parent nodes to avoid the overloading problem and ensure node lifetime maximization in the IPv6 Routing Protocol for Low-Power and Lossy Networks (RPL). The standard OFs are used to build a Destination Oriented Directed Acyclic Graph (DODAG) where the bottleneck nodes may suffer from unbalanced traffic load. As a result, a part of the network may be disconnected as the energy of the overloaded preferred parent node will drain much faster than other candidate parents. Thus, a new RPL metric has been introduced to balance the traffic load over the network. The potential extra overhead has been mitigated using a new utilization technique. Finally, the proposed OF amends the DODAG Information Object (DIO) message format.

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and BCP 79.

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/lid-abstracts.html>

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

Copyright and License Notice

Copyright (c) 2017 IETF Trust and the persons identified as the 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

1	Introduction	3
1.1	Terminology	3
2	DODAG construction in a nutshell	4
3	Load balancing in RPL	4
4	The proposed objective function	6
4.1	Balancing the load traffic	6
4.2	A new utilization technique for the amended DIO	6
4.3	The amended DIO message	7
5	Security Considerations	7
6	IANA Considerations	7
7	References	8
7.1	Normative References	8
7.2	Informative References	9
	Appendix A.	9
	Authors' Addresses	9

1 Introduction

IPv6 Routing Protocol for LLNs (RPL) [RFC6550] defined two OFs to optimize the path selection towards the root node, namely, the OF zero (OF0) [RFC6552], and the Minimum Rank with Hysteresis OF (MRHOF)[RFC6719]. The Destination Oriented Directed Acyclic Graph (DODAG) construction is built by the RPL OF, that specify how nodes select the preferred parent node by translating one or more metrics into the rank value.

The used OF calculates the rank based on some routing metrics [RFC 6551] such as hop-count, delay, energy, and so forth. The parent node in RPL can serve more than one child if it is chosen by them as preferred parent. Consequently, the overloaded preferred parents will become fragile nodes as their energy risks to drain much quicker than other nodes.

Having conducted simulation experiments and rigours analysis, it is concluded that the current OFs lead to build a topology that suffers from an unbalanced load traffic in bottleneck nodes especially for the first hop nodes (i.e., from the root). Consequently, this problem has a crucial impact on the lifetime of these types of nodes. The battery depletion of that overloaded parent node may affect the network reliability negatively.

This challenging problem is still an open issue. In an attempt to overcome this problem, this draft proposes a new OF to mitigate the overusing of the bottleneck node to prolong its battery lifetime.

This draft proposes an extended Objective Function(OF) that balances the number of children nodes for the overloaded nodes to ensure node lifetime maximization in RPL and can be summarized as follows. First, a new RPL metric has been used to balance the load traffic among the bottleneck nodes. Second, the DODAG Information Object (DIO) message has been amended by injecting the IP address of the chosen parent before broadcasting it. Third, a new utilization technique has been proposed for the amended DIO message to avoid increasing the overhead of the handshaking and acknowledgment processes. Simulation experiments have been conducted to validate the extended OF performance as detailed in Appendix A.

1.1 Terminology

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].

2 DODAG construction in a nutshell

RPL is a proactive distance vector routing protocol designed for LLNs [RFC6550], it constructs a DODAG using a certain OF that suits the application requirements. Essentially, RPL relies on a DODAG Information Object (DIO) control message to build the DODAG.

Thus, the starting point begins when the root node broadcasts the DIO message to the downstream neighbor nodes. As soon as the closest node receives the message, it can decide whether to join this DODAG or not based on the calculated rank according to the equations (1) and (2) [RFC6719].

$$\text{Rank}(N) = \text{Rank}(PN) + \text{RankIncrease} \quad (1)$$

$$\text{RankIncrease} = \text{Step} * \text{MinHopRankIncrease} \quad (2)$$

Where Step represents a scalar value and MinHopRankIncrease represents the minimum RPL parameter. If the node decides to join, then it adds the DIO sender to the candidate parent list. Next, the preferred parent, i.e. the next hop to the root, will be chosen based on the rank from this list to receive all traffic from the child node. Then, it computes its own rank with a monotonical increase according to the selected OF.

After that, the node propagates its own DIO with all updated information to all its neighbors including the preferred parent. [RFC 6551] defined the number of node metrics/constraints (e.g. hop count and energy) and the link metrics/constraints (e.g. ETX and throughput) that might be used in the OFs [RFC6552][RFC6719].

3 Load balancing in RPL

RPL is designed with several robust features such as exiguous delay, quick configuration, loop-free topology, and self-healing. However, the load imbalance is considered as a significant weakness in this protocol. More specifically, RPL is dealing with non-uniform distribution in large-scale LLNs, which may lead to unequal data traffic distribution. Consequently, the energy of the overloaded nodes will be drained much faster than other nodes. Furthermore, this problem has more harmful impacts if the overloaded node is a bottleneck node (i.e. with the first hop to the root) as shown in Figure 1 for node A and B.

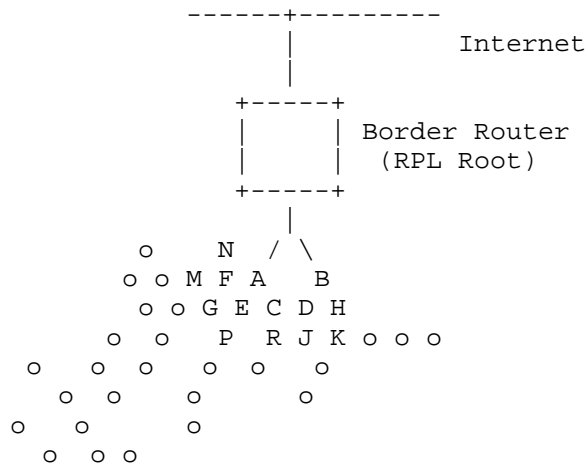


Figure 1: Bottleneck nodes in RPL

Figure 2 depicts the selection of the preferred parent for those nodes are within the first hop from nodes A or B. Clearly, node A has more children as it is surrounded by the nodes (N,M,F,G,E,P). Despite the fact that A has more children, it dominates the shared nodes (C,D,R,J) that are also located within the shared area of node B (i.e., within the transmission range of A and B). That unbalanced parent selection approach in RPL left node B only with two children, while node A has ten children.

Parent nodes	Children nodes	Shared nodes between A and B
A	N,M,F,G,E,P,C,D,R,J	C,D,R,J
B	H,K	C,D,R,J

Figure 2: The selection of the preferred parents

It is notable that the connection of all nodes through A is fragile as it is the only link to the root with an overloaded bottleneck node, thus, disconnecting part of the network if node A dies. In particular, this serious problem occurs in RPL due to omitting the number of children in existing parent selection technique.

To this end, the node sticks with the current preferred parent and influences its rank, even if this parent deteriorates with more load (i.e. being a parent for more children). The only conceivable scenarios to change the current parent to another candidate parent are as follow: first, if the current parent dies due to battery depletion. The second possibility, when the lossy percentage becomes higher than before, so no acknowledgement message can be heard from the preferred parent for a certain period of time.

4 The proposed objective function

The proposed OF leverages the lifetime of the entire network. The load balanced OF (LB-OF) balances the data traffic by taking into account the number of children for each candidate parent.

4.1 Balancing the load traffic

As aforementioned, being a preferred parent for more children means more overhead and unbalanced load, thus leads to drain its own energy much faster than other candidate parents. To solve this problem, a new metric has been proposed. The children set created in section 4.2 provides each preferred parent with the number of children it has. Based on that, the number of children in the rank calculation in formula (1) is considered.

Specifically, the parent with the least number of children will be elected as preferred parent. To this end, the balance has been achieved by declining the number of children of the overloaded bottleneck node. As a result, the majority of children (i.e., the shared nodes between A and B) will choose another preferred parent according to the lower rank, and surely has less number of children.

4.2 A new utilization technique for the amended DIO

Generally, in the upward routes the root initiates the DODAG construction by sending the first DIO message. Once other nodes receive this DIO, they select the sender as a preferred parent, and then they start calculating their own ranks based on the assigned OF. After that, each node broadcasts its own DIO message (i.e. the updated DIO that contains the new calculated rank value) to all neighbors including the chosen preferred parent which sent the original DIO message. In the standard OFs, the preferred parent ignores the DIOs that come from its child based on the rank.

In this stage, the aim is to allow each parent to count its number of children to avoid later possible overloading situations. However, that

is not possible in the upward routes (i.e., while maintaining the DODAG through DIOs), as the only control message that can be acknowledged by the destination is the Destination Advertisement Object (DAO) message in the downward routes to recognize the number of children for each parent.

Alternatively, a handshaking mechanism between parent and children can be set up to count the number of children for each parent, but this also brings an extra overhead for the entire network and subsequently increases the power consumption massively. To overcome this problem, LB-OF using a new technique is proposed as detailed below. In LB-OF algorithm, the received DIO from the child node is counted by the preferred parent node using a special buffer (set) created for this purpose. As detailed in section 4.3, the amended DIO contains the IP address of the chosen preferred parent. Thus, for each received DIO, the node matches its own IP address with the preferred parent IP address which is inserted in the DIO message, then increments the number of children set by one for this node if there is a matching.

Hence, this technique evades increasing the overhead that can be caused by the handshaking process (which could be used for the entire network). In addition, the coming DIOs from the children nodes has been utilized to allow each preferred parent to distinguish the number of its children during the DODAG construction stage.

4.3 The amended DIO message

Typically, the DIO carries the RPL InstanceID, DODAG identifier, version number, Rank and the OF that has been used to calculate the rank, in addition to others identifiers [RFC6550]. The DIO has been amended by injecting the chosen preferred parent IP address into the propagated DIO as shown in Figure 3.

5 Security Considerations

Since the routing metrics/constraints are carried within RPL message, the security routing mechanisms defined in [RFC6550] apply here.

6 IANA Considerations

The IANA defined in [RFC6551] apply here according to [RFC5226].

7.2 Informative References

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.

[Contiki] Contik O.S and cooja simulator <http://www.contiki-os.org/>

Appendix A.

The protocol has been simulated with Cooja simulator based on Instant Contiki 2.7 operating system [Contiki]. Collected results corroborate the superiority of our OF over the existing ones in terms of lifetime, power consumption and packet delivery ratio.

Authors' Addresses

Mamoun Qasem (editor)
Edinburgh Napier University
10 Colinton Road, EH10 5DT, Edinburgh, UK

Phone: +44 131 455 2772
Email: m.qasem@napier.ac.uk

Ahmed Al-Dubai
Edinburgh Napier University
10 Colinton Road, EH10 5DT, Edinburgh, UK

Phone: +44 131 455 2796
Email: a.al-dubai@napier.ac.uk

Imed Romdhani
Edinburgh Napier University
10 Colinton Road, EH10 5DT, Edinburgh, UK

Phone: +44 131 455 2726
Email: i.romdhani@napier.ac.uk

Baraq Ghaleb
Edinburgh Napier University

10 Colinton Road, EH10 5DT, Edinburgh, UK

Email: b.ghaleb@napier.ac.uk