ROLL                                                        P. Thubert, Ed.
Internet-Draft                                               J. Pylakutty
Intended status: Standards Track                                     Cisco
Expires: June 10, 2017                                    December 07, 2016


                    Root initiated routing state in RPL
                    draft-ietf-roll-dao-projection-00

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 by DAO messages, this draft projects the routes from
   the root of the DODAG.

Table of Contents

1.  Introduction

   The Routing Protocol for Low Power and Lossy Networks [RFC6550] (LLN)
   (RPL) specification defines a generic Distance Vector protocol that
   is designed for very low energy consumption and adapted to a variety
   of LLNs.  RPL forms Destination Oriented Directed Acyclic Graphs
   (DODAGs) which 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.

   In the non-storing mode (NSM) of operation (MOP), the root also
   computes routes down the DODAG towards the end device and leverages
   source routing to get there, while the default route via the root is
   used for routing upwards within the LLN and to the Internet at large.
   NSM is the dominant MOP because because networks may get arbitrary
   large and in Storing Mode, the amount of memory in nodes close to the
   root may unexpectedly require memory beyond a node's capabilities.

   But as a network gets deep, the size of the source routing header
   that the root must add to all the downward packets may also become an
   issue for far away target devices.  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 has in non-storing mode.

Additionally, RPL storing mode is optimized or 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.  In NSM, all peer-to-peer routes travel all the
way to the root, which adds a source routing header and forwards the
packet down to the destination, resulting in the longest stretch and
overload of the radio bandwidth near the root.  A controller, for
instance collocated with the RPL root, with enough topological
awareness of the connectivity between nodes, would be able to compute
more direct routes, avoiding the vicinity of the root whenever
possible.

The 6TiSCH architecture [I-D.ietf-6tisch-architecture] leverages the
Deterministic Networking Architecture [I-D.finn-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 mode routes within
the RPL domain, along a selected set of nodes and for a selected
duration, thus providing routes from suitable than those obtained

from the distributed operation of RPL in either storing and non-
storing modes.

## 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 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 hops to reach the target(s), presented in the
same order as they would appear in a routing header.

## 3.1.  Via Information

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

```
       0                   1                   2                   3
       0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |   Type = 0x0A | Option Length | Path Sequence | Path Lifetime |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                                                               |
```

```
         +                                                               +
         .                                                               .
         .                      Next-Hop Address                         .
         .                                                               .
         +                                                               +
         |                                                               |
         +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

                   Figure 1: Eliding the RPLInstanceID

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

   Option Length:  Variable, depending on whether or not Parent Address
         is present.

   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.

   Next-Hop Address:  8 or 16 bytes.  IPv6 Address of the next hop
         towards the destination(s) indicated in the target option that
         immediately precede the VIO.  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 is expressed as the 8-bytes suffix
         only, otherwise it is expressed as 16 bytes.

4. Loose Source Routing in Non-storing Mode

   A classical RPL implementation in a very constrained LLN 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 builds a path to a destination down the DODAG by
   concatenating this information.

```
                 ------+---------
                       |          Internet
                       |
                 +-----+
                 |     | Border Router
                 |     |  (RPL Root)
                 +-----+                     ^     |        |
                    |                        | DAO | ACK    |
              o    o   o    o                |     |        | Strict
            o o   o  o   o  o  o o    o       |     |        | Source
           o  o o  o o     o   o   o  o   o   |     |        | Route
           o   o    o  o     o  o    o  o  o  |     |        |
          o  o   o  o   o       o    o o      |     v        v
          o         o            o     o
                      LLN
```

              Figure 2: RPL non-storing operation

   Nodes are not expected to store downward routing state via their
   children, and the routing operates in strict source routing mode as
   detailed in An IPv6 Routing Header for Source Routes with RPL
   [RFC6554]

   This draft proposes an addition whereby the root projects a route
   through an extended DAO to an arbitrary node down the DODAG,
   indicating a child or a direct sequence of children via which a
   certain destination (target) may be reached.  The root is expected to
   use the mechanism optimally and with required parsimony to fit within
   the device resources, but how the root figures the amount of
   resources that are available is out of scope.

```
                 ------+---------
                       |          Internet
```

```
                    |
              +-----+
              |     |  Border Router
              |     |   (RPL Root)
              +-----+                    |    ^         |
              |                          | DAO | ACK    |
         o    o    o    o                |     |        | Loose
       o o   o o   o o o o    o          |  ^           | Source
      o  o o  o o     o   o   o o   o     |  |  DAO      | Route
       o    o    o o     o o    o  o o    | ^            |
      o  o   o  o   o         o    o o    v | DAO        v
       o         o              o    o
                       LLN
```
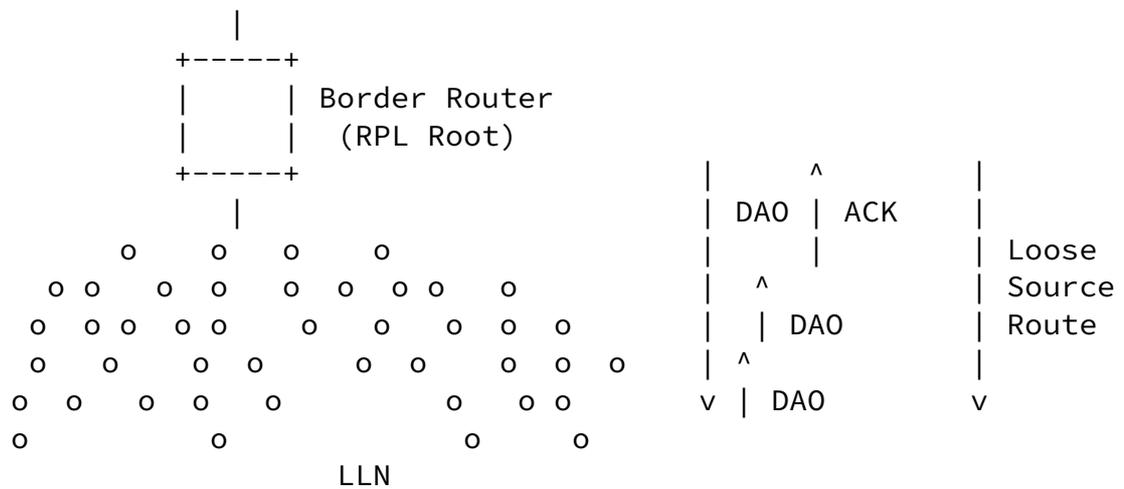
Figure 2: Non-Storing with Projected routes

When a RPL domain operates in non-storing Mode of Operation (NS-MOP),
only the root possesses routing information about the whole network.
A packet that is generated within the domain first reaches the root,
which can then apply a source routing information to reach the
destination.  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.

In NS-MOP, the root, or some associated centralized computation
engine, can thus 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).

Additionally, the DAG root knows the whole DAG topology, so when the
source of a packet is also in the RPL domain, 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 below diagram, if the source is 41
and the destination 52, the common parent is the node 22.

           ------+---------
```

```
                    |               Internet
                    |
              +-----+
              |     |  Border Router
              |     |   (RPL Root)
              +-----+
               | \  \____
              /   \       \
            o 11   o 12     o  13
            /      |      /  \
          o 22      o 23   o 24  o 25
          / \       | \      \
       o 31   o 32  o   o      o 35
       /        /    |    \     |    \
     o 41    o 42   o     o    o 45   o 46
     |        |      |     |     \      |
     o 51    o 52   o 53  o      o 55 o 56
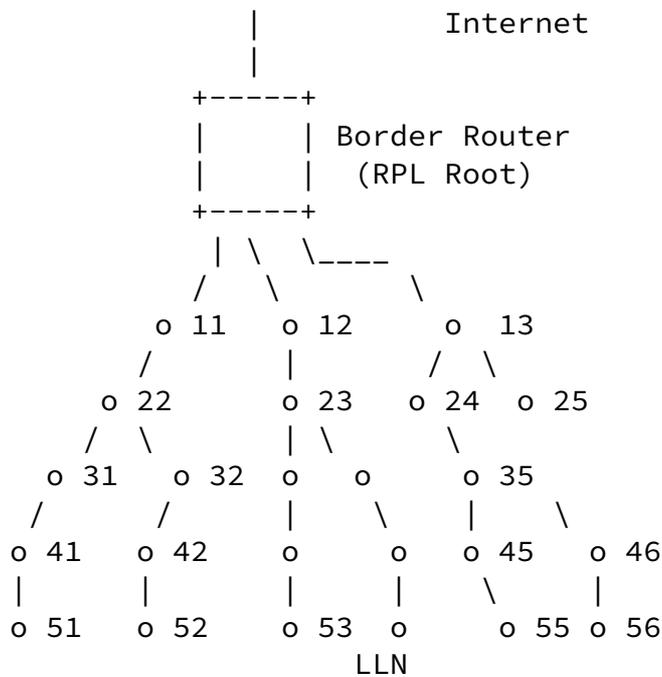                         LLN
```

                Figure 3: Non-Storing with Projected routes

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

   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].  Based on that information, the root
   computes which segment should be routed 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 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 other
   paths.

   This draft introduces a new mode of operation for loose source

routing in the LLN, the Non-Storing with Projected routes MOP.  With
this new MOP, the root sends a unicast DAO message to the last node
of the routing segment that must be installed.  The DAO message
contains the ordered list of hops along the segment as a list of Via
Information options that are preceded by one or more RPL Target
options to which they relate.  Each Via Information option contains a
lifetime for which state is to be maintained.

The root sends the DAO directly to the last node in the segment,
which is expected to be able to route to the targets on its own.

The last node in the segment may have another information to reach
the target(s), such as a connected route or an already installed
projected route.  If it does not have such a route then the node
should lookup the address on the relevant interfaces.  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), and continue the process for those targets
that could be located if any.

For the targets that could be located, last node in the segment
generates a DAO to its loose predecessor in the segment as indicated
in the list of Via Information options.

The node strips the last Via Information option which corresponds to
self, and uses it as source address for the DAO to the predecessor.
The address of the predecessor to be used as destination for the DAO
message is found in the now last Via Information option.  The
predecessor is expected to have a route to the address used as
source, either connected, installed previously as another DAO, or
from other means.

The predecessor is expected to have a route to the address used as
source and that is his successor.  If it does not and cannot locate
the successor, the predecessor node MUST answer to the root with a
negative DAO-ACK indicating the successor that could not be located.
The DAO-ACK contains the list of targets that could not be routed to
(suggested status 11).

If the predecessor can route to the successor node, then it installs
a route to the targets via the successor.  If that route is not
connected then a recursive lookup will take place to reach the
target(s).  From there, the node strips the last Via Information
option and either answers to the root with a positive DAO-ACK that
contains the list of targets that could be routed to, or propagates
the DAO to its own predecessor.

A NULL lifetime in the Via Information option along the segment is used to clean up the state.

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.

5.  Centralized Computation of Optimized Peer-to-Peer Routes

With the initial specifications of RPL [RFC6550], the P2P path from a source to a destination is often stretched, as illustrated in [RFC6550]:

   - 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.

   - 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 DAO 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.

It results that it is often beneficial to enable additional P2P routes, either if the RPL route present a stretch from shortest path, or if the new route is engineered with a different objective.

```
                   ------+---------
                   |              Internet
                   |
              +-----+
              |     | Border Router
              |     |  (RPL Root)
              +-----+
                 X
           ^     v    o     o
         ^ o   o v   o  o  o o    o
        ^   o o  o v     o    o    o   o
        ^     o   o v      o o    o  o  o
       S  o   o  o  D        o   o o
       o         o           o      o
                      LLN
```

                   Figure 4: Routing Stretch

   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.

   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:

      - 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.

      - 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 VIA options, and the flow is
similar to the flow used to obtain loose source routing.

The root sends the DAO with the target option and the Via Option to
the lest router in the path; the last router removes the last Via
Option and passes the DAO to the previous hop.

```
                 ------+---------
                      |              Internet
                      |
                +-----+
                |     |  Border Router
                |     |   (RPL Root)
                +-----+
                    | Projected DAO message to C
           o        |   o     o
         o o    o   |     o  o  o o    o
          o  o o    | o      o   o    o   o
          o   o  V  o       o  o    o  o  o
         S   A   B  C    D          o    o o
          o           o              o      o
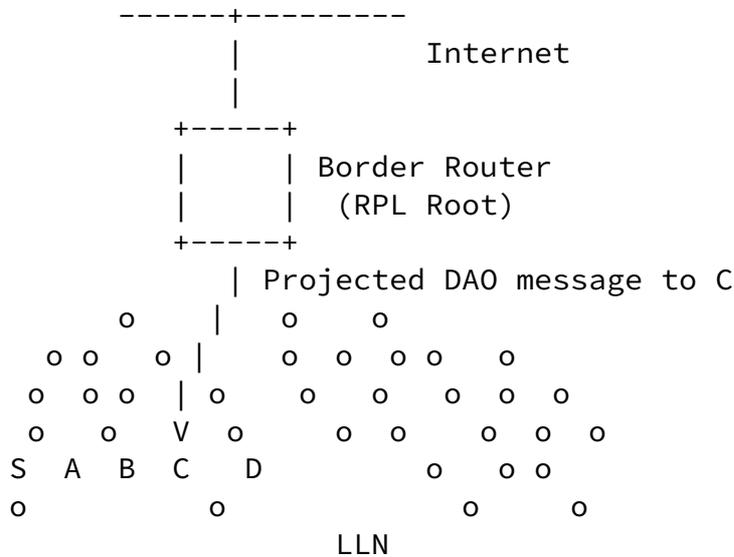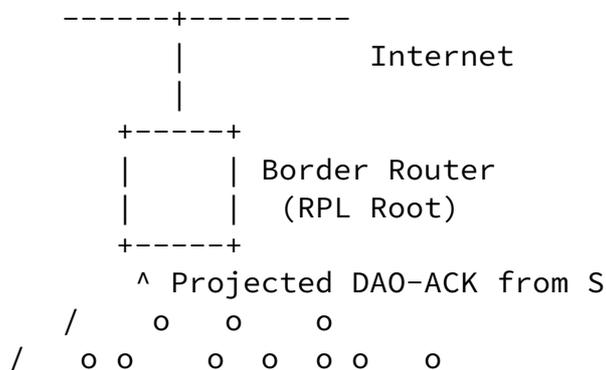                         LLN
```

Figure 5: Projected DAO from root

The process recurses till the destination which sends a DAO-ACK to
the root.  In the example above, for target D, the list of via
options is S, A, B and C.  The projected DAO is sent by the root to

```
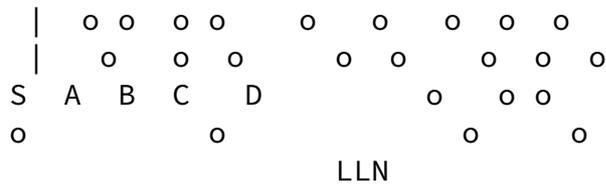                 ------+---------
                      |              Internet
                      |
                +-----+
                |     |  Border Router
                |     |   (RPL Root)
                +-----+
                 ^ Projected DAO-ACK from S
            /    o   o    o
          /    o o    o  o  o o    o
```

```
             |  o o   o o     o    o    o  o   o
             |   o    o  o      o  o   o    o  o  o
            S  A   B  C   D          o   o o
            o        o           o      o
                           LLN
```

                  Figure 6: Projected DAO-ACK to root

   The process recurses till the destination which sends a DAO-ACK to
   the root.  In the example above, for target D, the list of via
   options is S, A, B and C.  The projected DAO is sent by the root to

```
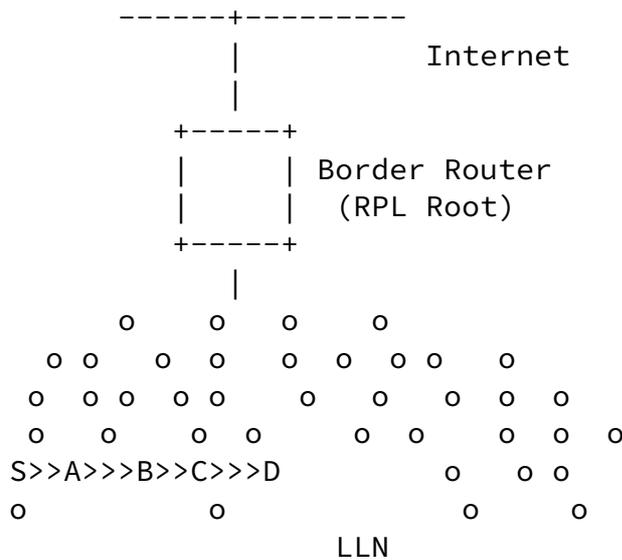                ------+---------
                      |              Internet
                      |
               +-----+
               |     | Border Router
               |     |  (RPL Root)
               +-----+
                  |
            o      o    o     o
          o o   o  o    o  o  o o    o
          o  o o  o o     o    o    o  o  o
          o   o    o  o     o   o    o  o  o
         S>>A>>>B>>C>>>D           o    o o
          o        o               o      o
                           LLN
```

                  Figure 7: Projected Transversal Route

6.  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.

7.  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]

8.  Acknowledgments

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

9.  References

9.1.  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>.

   [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>.

   [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>.

9.2.  Informative References

   [I-D.finn-detnet-architecture]
              Finn, N. and P. Thubert, "Deterministic Networking
              Architecture", draft-finn-detnet-architecture-08 (work in
              progress), August 2016.

   [I-D.ietf-6tisch-architecture]
              Thubert, P., "An Architecture for IPv6 over the TSCH mode
              of IEEE 802.15.4", draft-ietf-6tisch-architecture-10 (work
              in progress), June 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>.

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