

ROLL
Internet-Draft
Intended status: Standards Track
Expires: 24 September 2022

P. Thubert, Ed.
Cisco Systems
R.A. Jadhav
Huawei Tech
M. Richardson
Sandelman
23 March 2022

Root initiated routing state in RPL
draft-ietf-roll-dao-projection-25

Abstract

THIS RFC extends RFC 6550, RFC 6553, and RFC 8138 to enable a RPL Root to install and maintain Projected Routes within its DODAG, along a selected set of nodes that may or may not include self, for a chosen duration. This potentially enables routes that are more optimized or resilient than those obtained with the classical distributed operation of RPL, either in terms of the size of a Routing Header or in terms of path length, which impacts both the latency and the packet delivery ratio.

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 <https://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 24 September 2022.

Copyright Notice

Copyright (c) 2022 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 (<https://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 Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction	4
2. Terminology	4
2.1. Requirements Language	4
2.2. References	5
2.3. Glossary	5
2.4. Domain Terms	5
2.4.1. Projected Route	6
2.4.2. Projected DAO	6
2.4.3. Path	6
2.4.4. Routing Stretch	6
2.4.5. Track	7
3. Context and Goal	9
3.1. RPL Applicability	10
3.2. RPL Routing Modes	11
3.3. Requirements	12
3.3.1. Loose Source Routing	12
3.3.2. East-West Routes	13
3.4. On Tracks	15
3.4.1. Building Tracks With RPL	15
3.4.2. Tracks and RPL Instances	16
3.5. Serial Track Signaling	16
3.5.1. Using Storing Mode Segments	18
3.5.2. Using Non-Storing Mode joining Tracks	24
3.6. Complex Tracks	31
3.7. Scope and Expectations	33
3.7.1. External Dependencies	33
3.7.2. Positioning vs. Related IETF Standards	33
4. Extending existing RFCs	35
4.1. Extending RFC 6550	35
4.1.1. Projected DAO	36
4.1.2. Projected DAO-ACK	38
4.1.3. Via Information Option	39
4.1.4. Sibling Information Option	39
4.1.5. P-DAO Request	39
4.1.6. Amending the RPI	40
4.1.7. Additional Flag in the RPL DODAG Configuration Option	40
4.2. Extending RFC 6553	41
4.3. Extending RFC 8138	42
5. New RPL Control Messages and Options	43

5.1.	New P-DAO Request Control Message	43
5.2.	New PDR-ACK Control Message	45
5.3.	Via Information Options	46
5.4.	Sibling Information Option	49
6.	Root Initiated Routing State	51
6.1.	RPL Network Setup	51
6.2.	Requesting a Track	52
6.3.	Identifying a Track	53
6.4.	Installing a Track	54
6.4.1.	Signaling a Projected Route	55
6.4.2.	Installing a Track Segment with a Storing Mode P-Route	56
6.4.3.	Installing a Track Leg with a Non-Storing Mode P-Route	58
6.5.	Tearing Down a P-Route	60
6.6.	Maintaining a Track	60
6.6.1.	Maintaining a Track Segment	61
6.6.2.	Maintaining a Track Leg	61
6.7.	Encapsulating and Forwarding Along a Track	62
6.8.	Compression of the RPL Artifacts	64
7.	Lesser Constrained Variations	66
7.1.	Storing Mode Main DODAG	66
7.2.	A Track as a Full DODAG	68
8.	Profiles	69
9.	Backwards Compatibility	71
10.	Security Considerations	72
11.	IANA Considerations	72
11.1.	RPL DODAG Configuration Option Flag	72
11.2.	Elective 6LoWPAN Routing Header Type	73
11.3.	Critical 6LoWPAN Routing Header Type	73
11.4.	Subregistry For The RPL Option Flags	73
11.5.	RPL Control Codes	74
11.6.	RPL Control Message Options	74
11.7.	SubRegistry for the Projected DAO Request Flags	75
11.8.	SubRegistry for the PDR-ACK Flags	75
11.9.	Subregistry for the PDR-ACK Acceptance Status Values	76
11.10.	Subregistry for the PDR-ACK Rejection Status Values	76
11.11.	SubRegistry for the Via Information Options Flags	77
11.12.	SubRegistry for the Sibling Information Option Flags	77
11.13.	Destination Advertisement Object Flag	77
11.14.	Destination Advertisement Object Acknowledgment Flag	78
11.15.	New ICMPv6 Error Code	78
11.16.	RPL Rejection Status values	78
12.	Acknowledgments	79
13.	Normative References	79
14.	Informative References	81
	Authors' Addresses	83

1. Introduction

RPL, the "Routing Protocol for Low Power and Lossy Networks" [RPL] (LLNs), is an anisotropic Distance Vector protocol that is well-suited for application in a variety of low energy Internet of Things (IoT) networks where stretched P2P paths are acceptable vs. the signaling and state overhead involved in maintaining shortest paths across.

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 IP backbone and routes along that graph up, towards the Root, and down towards the nodes.

With this specification, an abstract routing function called a Path Computation Element [PCE] (e.g., located in a central controller or collocated with the Root) interacts with the RPL Root to compute Peer to Peer (P2P) paths within a pre-existing RPL Main DODAG. The topological information that is passed to the PCE is derived from the DODAG that is already available at the Root in RPL Non-Storing Mode. This specification introduces protocol extensions that enrich the topological information that is available at the Root and passed to the PCE.

Based on usage, path length, and knowledge of available resources such as battery levels and reservable buffers in the nodes, the PCE with a global visibility on the system can optimize the computed routes for the application needs, including the capability to provide path redundancy. This specification also introduces protocol extensions that enable the Root to translates the computed paths into RPL and install them as Projected Routes (aka P-Routes) inside the DODAG on behalf of a PCE.

2. Terminology

2.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in THIS RFC are to be interpreted as described in BCP 14 [RFC2119][RFC8174] when, and only when, they appear in all capitals, as shown here.

In addition, the terms "Extends" and "Amends" are used as per [I-D.kuehlewind-update-tag] section 3.

2.2. References

In THIS RFC, readers will encounter terms and concepts that are discussed in the "Routing Protocol for Low Power and Lossy Networks" [RPL], the "6TiSCH Architecture" [RFC9030], the "Deterministic Networking Architecture" [RFC8655], the "Reliable and Available Wireless (RAW) Architecture" [RAW-ARCHI], and "Terminology in Low power And Lossy Networks" [RFC7102].

2.3. Glossary

THIS RFC often uses the following acronyms:

CMO: Control Message Option
DAO: destination Advertisement Object
DAG: Directed Acyclic Graph
DODAG: destination-Oriented Directed Acyclic Graph; A DAG with only one vertex (i.e., node) that has no outgoing edge (i.e., link)
GUA: IPv6 Global Unicast Address
LLN: Low-Power and Lossy Network
MOP: RPL Mode of Operation
P-DAO: Projected DAO
P-Route: Projected Route
PDR: P-DAO Request
RAN: RPL-Aware Node (either a RPL router or a RPL-Aware Leaf)
RAL: RPL-Aware Leaf
RH: Routing Header
RPI: RPL Packet Information
RTO: RPL Target Option
RUL: RPL-Unaware Leaf
SIO: RPL Sibling Information Option
ULA: IPv6 Unique Local Address
NSM-VIO: A Source-Routed Via Information Option, used in Non-Storing Mode P-DAO messages.
SLO: Service Level Objective
TIO: RPL Transit Information Option
SM-VIO: A strict Via Information Option, used in Storing Mode P-DAO messages.
VIO: A Via Information Option; it can be a SM-VIO or an NSM-VIO.

2.4. Domain Terms

This specification uses the following terminology:

2.4.1. Projected Route

A RPL P-Route is a RPL route that is computed remotely by a PCE, and installed and maintained by a RPL Root on behalf of the PCE. It is installed as a state that signals that destinations (aka Targets) are reachable along a sequence of nodes.

2.4.2. Projected DAO

A DAO message used to install a P-Route.

2.4.3. Path

Quoting section 1.1.3 of [INT-ARCHI]:

At a given moment, all the IP datagrams from a particular source host to a particular destination host will typically traverse the same sequence of gateways. We use the term "path" for this sequence. Note that a path is uni-directional; it is not unusual to have different paths in the two directions between a given host pair.

Section 2 of [I-D.irtf-panrg-path-properties] points to a longer, more modern definition of path, which begins as follows:

A sequence of adjacent path elements over which a packet can be transmitted, starting and ending with a node. A path is unidirectional. Paths are time-dependent, i.e., the sequence of path elements over which packets are sent from one node to another may change. A path is defined between two nodes.

It follows that the general acceptance of a path is a linear sequence of nodes, as opposed to a multi-dimensional graph. In the context of this document, a path is observed by following one copy of a packet that is injected in a Track and possibly replicated within.

2.4.4. Routing Stretch

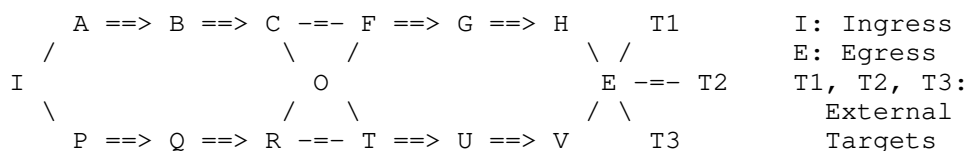
RPL is anisotropic, meaning that it is directional, or more exactly polar. RPL does not behave the same way "down" with multicast DIO messages that form the DODAG and "up" with unicast DAO messages that follow the DODAG. This is in contrast with traditional IGPs that operate the same in all directions and are thus called isotropic.

The term Routing Stretch denotes the length of a path, as compared with a shortest path, which can be an abstract concept in RPL when the metrics are statistical and dynamic, and the concept of short varies with the Objective Function.

The RPL DODAG optimizes the P2MP (from Root) and MP2P (to Root) paths, but the P2P (node to node) traffic has to follow the same DODAG. Following the DODAG, the RPL datapath passes via a common parent in Storing Mode and via the Root in Non-Storing Mode. This typically involves more hops and more latency than the minimum possible for a direct P2P path that an isotropic protocol would compute. We refer to this elongated path as stretched.

2.4.5. Track

A networking graph that can be followed to transport packets with equivalent treatment; as opposed to the definition of a path above, a Track is not necessarily linear. It may contain multiple paths that may fork and rejoin, and may enable the RAW Packet ARQ, Replication, Elimination, and Overhearing (PAREO) operations.



I ==> A ==> B ==> C : a segment to targets F and O

I --> F --> E : a leg to targets T1, T2, T3

I, A, B, C, F, G, H, E : a path to T1, T2, T3

Figure 1: A Track and its Components

This specification builds Tracks that are DODAGs oriented towards a Track Ingress, and the forward direction for packets (aka East-West) is from the Track Ingress to one of the possibly multiple Track Egress Nodes, which is also down the DODAG.

The Track may be strictly connected, meaning that the vertices are adjacent, or loosely connected, meaning that the vertices are connected using Segments that are associated to the same Track.

2.4.5.1. TrackID

A RPL Local InstanceID that identifies a Track using the namespace owned by the Track Ingress. The TrackID is associated with the IPv6 Address of the Track Ingress that is used as DODAGID, and together they form a unique identification of the Track (see the definition of DODAGID in section 2 of [RPL]).

2.4.5.2. Namespace

The term namespace is used to refer to the scope of the TrackID. The TrackID is locally significant within its namespace. The namespace is identified by the DODAGID for the Track. The tuple (DODAGID, TrackID) is globally unique.

2.4.5.3. Serial Track

A Track that has only one path.

2.4.5.4. Stand-Alone

A single P-DAO that fully defines a Track, e.g., a Serial Track installed with a single Storing Mode Via Information option (SM-VIO).

2.4.5.5. Stitching

This specification using the term stitching to indicate that a track is piped to another one, meaning that traffic out of the first is injected in the other.

2.4.5.6. Leg

An end-to-end East-West serial path. A leg can be a serial Track by itself or a subTrack of a complex Track with the same Ingress and Egress Nodes. With this specification, a Leg is installed by the Root of the main DODAG using a Non-Storing Mode P-DAO message, and it is expressed as a loose sequence of nodes that are joined by Track Segments.

As the Non-Storing Mode Via Information option (NSM-VIO) can only signal sequences of nodes, it takes one Non-Storing Mode P-DAO message per Leg to signal the structure of a complex Track.

Each NSM-VIO for the same TrackId but a different Segment ID signals a different leg that the Track Ingress adds to the topology.

2.4.5.7. subTrack

A Track within a Track, formed by a non-empty collection of Legs of the Track.

2.4.5.8. Segment

A serial path formed by a strict sequence of nodes, along which a P-Route is installed. With this specification, a Segment is typically installed by the Root of the main DODAG using Storing Mode P-DAO messages. A Segment is used as the topological edge of a Track joining the loose steps along the Legs that form the structure of a complex Track. The same segment may be leveraged by more than one Leg where the Legs overlap.

Since this specification builds only DODAGs, all Segments are oriented from Ingress (East) to Egress (West), as opposed to the general Track model in the RAW Architecture [RAW-ARCHI], which allows North/South Segments that can be bidirectional as well.

2.4.5.8.1. Section of a Segment

A continuous subset of a segment that may be replaced while the segment remains. for instance, in segment $A \Rightarrow B \Rightarrow C \Rightarrow D \Rightarrow E \Rightarrow F$, say that the link C to D might be misbehaving. The section $B \Rightarrow C \Rightarrow D \Rightarrow E$ in the segment may be replaced by $B \Rightarrow C' \Rightarrow D' \Rightarrow E$ to route around the problem. The segment becomes $A \Rightarrow B \Rightarrow C' \Rightarrow D' \Rightarrow E \Rightarrow F$.

2.4.5.8.2. Segment Routing and SRH

The terms Segment Routing and SRH refer to using source-routing to hop over segments. In a Non-Storing mode RPL domain, the SRH is typically a RPL Source Route Header (the IPv6 RH of type 3) as defined in [RFC6554].

If the network is a 6LoWPAN Network, the expectation is that the SRH is compressed and encoded as a 6LoWPAN Routing Header (6LoRH), as specified in section 5 of [RFC8138].

On the other hand, if the RPL Network is less constrained and operated in Storing Mode, as discussed in Section 7.1, the Segment Routing operation and the SRH could be as specified in [RFC8754]. This specification applies equally to both forms of source routing and SRH.

3. Context and Goal

3.1. RPL Applicability

RPL is optimized for situations where the power is scarce, the bandwidth constrained and the transmissions unreliable. This matches the use case of an IoT LLN where RPL is typically used today, but also situations of high relative mobility between the nodes in the network (aka swarming), e.g., within a variable set of vehicles with a similar global motion, or a toon of drones.

To reach this goal, RPL is primarily designed to minimize the control plane activity, that is the relative amount of routing protocol exchanges vs. data traffic, and the amount of state that is maintained in each node. RPL does not need converge, and provides connectivity to most nodes most of the time.

RPL may form multiple topologies called instances. Instances can be created to enforce various optimizations through objective functions, or to reach out through different Root Nodes. The concept of objective function allows to adapt the activity of the routing protocol to the use case, e.g., type, speed, and quality of the LLN links.

RPL instances operate as ships passing in the night, unbeknownst of one another. The RPL Root is responsible to select the RPL Instance that is used to forward a packet coming from the Backbone into the RPL domain and set the related RPL information in the packets. 6TiSCH leverages RPL for its distributed routing operations.

To reduce the routing exchanges, RPL leverages an anisotropic Distance Vector approach, which does not need a global knowledge of the topology, and only optimizes the routes to and from the RPL Root, allowing P2P paths to be stretched. Although RPL installs its routes proactively, it only maintains them lazily, in reaction to actual traffic, or as a slow background activity.

This is simple and efficient in situations where the traffic is mostly directed from or to a central node, such as the control traffic between routers and a controller of a Software Defined Networking (SDN) infrastructure or an Autonomic Control Plane (ACP).

But stretch in P2P routing is counter-productive to both reliability and latency as it introduces additional delay and chances of loss. As a result, [RPL] is not a good fit for the use cases listed in the RAW use cases document [USE-CASES], which demand high availability and reliability, and as a consequence require both short and diverse paths.

3.2. RPL Routing Modes

RPL first forms a default route in each node towards the a Root, and those routes together coalesce as a Directed Acyclic Graph upwards. RPL then constructs routes to destinations signaled as Targets in the reverse direction, down the same DODAG. So do so, a RPL Instance can be operated either in RPL Storing or Non-Storing Mode of Operation (MOP). The default route towards the Root is maintained aggressively and may change while a packet progresses without causing loops, so the packet will still reach the Root.

In Non-Storing Mode, each node advertises itself as a Target directly to the Root, indicating the parents that may be used to reach self. Recursively, the Root builds and maintains an image of the whole DODAG in memory, and leverages that abstraction to compute source route paths for the packets to their destinations down the DODAG. When a node changes its point(s) of attachment to the DODAG, it takes single unicast packet to the Root along the default route to update it, and the connectivity is restored immediately; this mode is preferable for use cases where internet connectivity is dominant, or when, like here, the Root controls the network activity in the nodes.

In Storing Mode, the routing information percolates upwards, and each node maintains the routes to the subDAG of its descendants down the DODAG. The maintenance is lazy, either reactive upon traffic or as a slow background process. Packets flow via the common parent and the routing stretch is reduced vs. Non-Storing, for a better P2P connectivity. On the other hand, a new route takes a longer time to propagate to the Root, time for the Distance-Vector protocol to operate hop-by-hop, and the Internet connectivity is restored more slowly upon movement.

Either way, the RPL routes are injected by the Target nodes, in a distributed fashion. To complement RPL and eliminate routing stretch, this specification introduces an hybrid mode that combines Storing and Non-Storing operations to build and project routes onto the nodes where they should be installed. This specification uses the term Projected Route (P-Route) to refer to those routes.

A P-Route may be installed in either Storing and Non-Storing Mode, potentially resulting in hybrid situations where the Mode of the P-Route is different from that of the RPL Main DODAG. P-Routes can be used as stand-alone segments to reduce the size of the source routing headers with loose source routing operations down the main RPL DODAG. P-Routes can also be combined with other P-Routes to form a more complex forwarding graph called a Track.

3.3. Requirements

3.3.1. Loose Source Routing

A RPL implementation operating in a very constrained LLN typically uses the Non-Storing Mode of Operation as represented in Figure 2. In that mode, 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.

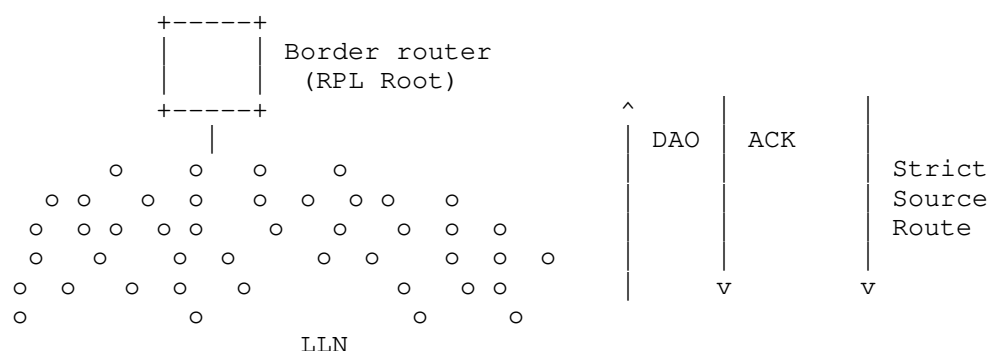


Figure 2: RPL Non-Storing Mode of operation

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 wireless bandwidth near the Root is the gating factor for all transmissions towards or within the domain, and that the Root is a single point of failure for all connectivity to nodes within its domain.

The RPL Root must add a source routing header to all downward packets. As a network grows, the size of the source routing header augments with the depth of the nodes. In some use cases, a RPL network forms long lines along physical structures such as streets for lighting. Limiting the packet size is directly beneficial to the energy budget, but, mostly, it reduces the chances of frame loss and packet fragmentation, which are highly detrimental to the LLN operation. A limited amount of well-targeted routing state would

allow the source routing operation to be loose as opposed to strict, and save packet size. 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.

Being on path for all packets in Non-Storing mode, the Root may determine the number of P2P packets in its RPL domain per source and destination, the latency incurred, and the amount of energy and bandwidth that is consumed to reach the self and then down, including a possible fragmentation when encapsulating larger packets. Enabling a shorter path that would not traverse the Root for select P2P source/destinations may improve the latency, lower the consumption of constrained resources, free bandwidth at the bottleneck near the Root, improve the delivery ratio and reduce the latency for those P2P flows with a global benefit for all flows of reducing the load at the Root.

This requirement is to store a routing state associated with the Main DODAG in selected RPL routers, to limit the excursion of the source route headers in deep networks. The Root may elide the sequence of routers that is installed in the network from its source route header, which becomes loose while it is strict in [RPL].

3.3.2. East-West Routes

[RPL] optimizes Point-to-Multipoint (P2MP) routes from the Root, Multipoint-to-Point (MP2P) routes to the DODAG Root, and Internet access when the Root also serves as Border Router. All routes are installed North-South (aka up/down) along the RPL DODAG. Peer to Peer (P2P) East-West routes in a RPL network will generally suffer from some elongated (stretched) path versus a direct (optimized) path, since routing between two nodes always happens via a common parent, as illustrated in Figure 3:

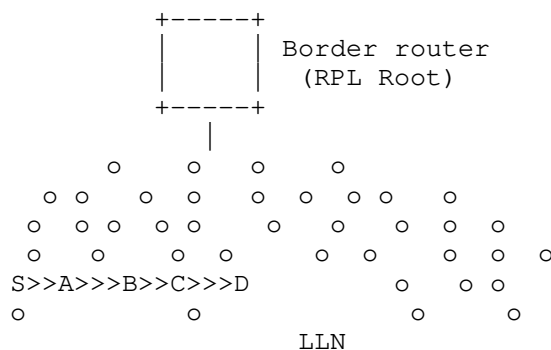


Figure 4: More direct East-West Route between S and D

The requirement is to install additional routes in the RPL routers, to reduce the stretch of some P2P routes and maintain the characteristics within a given SLO, e.g., in terms of latency and/or reliability.

3.4. On Tracks

3.4.1. Building Tracks With RPL

The concept of a Track was introduced in the "6TiSCH Architecture" [RFC9030], as a collection of potential paths that leverage redundant forwarding solutions along the way. This can be a DODAG or a more complex structure that is only partially acyclic (e.g., per packet).

With this specification, a Track is shaped as a DODAG, and following the directed edges leads to a Track Ingress. Storing Mode P-DAO messages follow the direction of the edges to set up routes for traffic that flows the other way, towards the Track Egress(es). If there is a single Track Egress, then the Track is reversible to form another DODAG by reversing the direction of each edge. A node at the Ingress of more than one Segment in a Track may use one or more of these Segments to forward a packet inside the Track.

A RPL Track is a collection of (one or more) parallel loose source routed sequences of nodes ordered from Ingress to Egress, each forming a Track Leg. The nodes that are directly connected, reachable via existing Tracks as illustrated in Section 3.5.2.3 or joined with strict Segments of other nodes as shown in Section 3.5.1.3. The Legs are expressed in RPL Non-Storing Mode and require an encapsulation to add a Source Route Header, whereas the Segments are expressed in RPL Storing Mode.

A Serial Track comprises provides only one path between Ingress and Egress. It comprises at most one Leg. A Stand-Alone Segment implicitly defines a Serial Track from its Ingress to Egress.

A complex Track forms a graph that provides a collection of potential paths to provide redundancy for the packets, either as a collection of Legs that may be parallel or cross at certain points, or as a more generic DODAG.

3.4.2. Tracks and RPL Instances

Section 5.1. of [RPL] describes the RPL Instance and its encoding. There can be up to 128 global RPL Instances, for which there can be one or more DODAGs, and there can be 64 local RPL Instances, with a namespace that is indexed by a DODAGID, where the DODAGID is a Unique Local Address (ULA) or a Global Unicast Address (GUA) of the Root of the DODAG. Bit 0 (most significant) is set to 1 to signal a Local RPLInstanceID, as shown in Figure 5. By extension, this specification expresses the value of the RPLInstanceID as a single integer between 128 and 191, representing both the Local RPLInstanceID in 0..63 and Bit 0 set.

```

0 1 2 3 4 5 6 7
+---+---+---+---+---+---+
|1|D|   ID   |  Local RPLInstanceID in 0..63
+---+---+---+---+---+---+

```

Figure 5: Local RPLInstanceID Encoding

A Track is normally associated with a Local RPL Instance which RPLInstanceID is used as the TrackID, more in Section 6.3. A Track Leg may also be used as a subTrack that extends the RPL main DODAG. In that case, the TrackID is set to the global RPLInstanceID of the main DODAG, which suffices to identify the routing topology. As opposed to local RPL instances, the Track Ingress that encapsulates the packets over a subtrack is not Root, and that the source address of the encapsulated packet is not used to determine the Track.

3.5. Serial Track Signaling

This specification enables to set up a P-Route along either a Track Leg or a Segment. A P-Route is installed and maintained by the Root of the main DODAG using an extended RPL DAO message called a Projected DAO (P-DAO), and a Track is composed of the combination of one or more P-Routes.

A P-DAO message for a Track signals the TrackID in the RPLInstanceID field. In the case of a local RPL Instance, the address of the Track Ingress is used as source to encapsulate packets along the Track. The Track is signaled in the DODAGID field of the Projected DAO Base Object, see Figure 8.

This specification introduces the Via Information Option (VIO) to signal a sequence of hops in a Leg or a Segment in the P-DAO messages, either in Storing Mode (SM-VIO) or Non-Storing Mode (NSM-VIO). One P-DAO messages contains a single VIO, associated to one or more RPL Target Options that signal the destination IPv6 addresses that can reached along the Track, more in Section 5.3.

Before diving deeper into Track Legs and Segments signaling and operation, this section provides examples of what how route projection works through variations of a simple example. This simple example illustrates the case of host routes, though RPL Targets can be prefixes.

Say we want to build a Serial Track from node A to E in Figure 6, so A can route packets to E's neighbors F and G along A, B, C, D and E as opposed to via the Root:

```

A ==> B ==> C ==> D ==> E < /==> F
                             \==> G

```

Figure 6: Reference Track

Conventionally we use ==> to represent a strict hop and --> for a loose hop. We use "-to-", such as in C==>D==>E-to-F to represent coma-separated Targets, e.g., F is a Target for Segment C==>D==>E. In this example, A is Track Ingress, E is Track Egress. C is a stitching point. F and G are "external" Targets for the Track, and become reachable from A via the Track A(ingress) to E (Egress and implicit Target in Non-Storing Mode) leading to F and G (explicit Targets).

Figure 5 depicts the format of the RPLInstanceID encoding for a local RPLInstanceID .

In a general manner the desired outcome is as follows:

- * Targets are E, F, and G
- * P-DAO 1 signals C==>D==>E

- * P-DAO 2 signals $A \Rightarrow B \Rightarrow C$

- * P-DAO 3 signals F and G via the $A \dashrightarrow E$ Track

P-DAO 3 may be omitted if P-DAO 1 and 2 signal F and G as Targets.

Loose sequences of hops must be expressed in Non-Storing Mode, so P-DAO 3 contains a NSM-VIO. With this specification, the DODAGID to be used by the Ingress as source address is signaled if needed in the DAO base object, the via list starts at the first loose hop and matches the source route header, and the Egress of a Non-Storing Mode P-DAO is an implicit Target that is not listed in the RTO.

3.5.1. Using Storing Mode Segments

$A \Rightarrow B \Rightarrow C$ and $C \Rightarrow D \Rightarrow E$ are segments of a same Track. Note that the Storing Mode signaling imposes strict continuity in a segment, since the P-DAO is passed hop by hop, as a classical DAO is, along the reverse datapath that it signals. One benefit of strict routing is that loops are avoided along the Track.

3.5.1.1. Stitched Segments

In this formulation:

- * P-DAO 1 signals $C \Rightarrow D \Rightarrow E$ -to-F,G

- * P-DAO 2 signals $A \Rightarrow B \Rightarrow C$ -to-F,G

Storing Mode P-DAO 1 is sent to E and when it is successfully acknowledged, Storing Mode P-DAO 2 is sent to C, as follows:

Field	P-DAO 1 to E	P-DAO 2 to C
Mode	Storing	Storing
Track Ingress	A	A
(DODAGID, TrackID)	(A, 129)	(A, 129)
SegmentID	1	2
VIO	C, D, E	A, B, C
Targets	F, G	F, G

Table 1: P-DAO Messages

As a result the RIBs are set as follows:

Node	destination	Origin	Next Hop(s)	TrackID
E	F, G	P-DAO 1	Neighbor	(A, 129)
D	E	P-DAO 1	Neighbor	(A, 129)
"	F, G	P-DAO 1	E	(A, 129)
C	D	P-DAO 1	Neighbor	(A, 129)
"	F, G	P-DAO 1	D	(A, 129)
B	C	P-DAO 2	Neighbor	(A, 129)
"	F, G	P-DAO 2	C	(A, 129)
A	B	P-DAO 2	Neighbor	(A, 129)
"	F, G	P-DAO 2	B	(A, 129)

Table 2: RIB setting

Packets originated by A to F or G do not require an encapsulation as the RPI can be placed in the native header chain. For packets that it routes, A must encapsulate to add the RPI that signals the trackID; the outer headers of the packets that are forwarded along the Track have the following settings:

Header	IPv6 Source Addr.	IPv6 Dest. Addr.	TrackID in RPI
Outer	A	F or G	(A, 129)
Inner	X != A	F or G	N/A

Table 3: Packet Header Settings

As an example, say that A has a packet for F. Using the RIB above:

- * From P-DAO 2: A forwards to B and B forwards to C.
- * From P-DAO 1: C forwards to D and D forwards to E.
- * From Neighbor Cache Entry: E delivers the packet to F.

3.5.1.2. External routes

In this example, we consider F and G as destinations that are external to the Track as a DODAG, as discussed in section 4.1.1. of [RFC9008]. We then apply the directives for encapsulating in that case, more in Section 6.7.

In this formulation, we set up the Track Leg explicitly, which creates less routing state in intermediate hops at the expense of larger packets to accommodate source routing:

- * P-DAO 1 signals C==>D==>E-to-E
- * P-DAO 2 signals A==>B==>C-to-E
- * P-DAO 3 signals F and G via the A-->E-to-F,G Track

Storing Mode P-DAO 1 and 2, and Non-Storing Mode P-DAO 3, are sent to E, C and A, respectively, as follows:

	P-DAO 1 to E	P-DAO 2 to C	P-DAO 3 to A
Mode	Storing	Storing	Non-Storing
Track Ingress	A	A	A
(DODAGID, TrackID)	(A, 129)	(A, 129)	(A, 129)
SegmentID	1	2	3
VIO	C, D, E	A, B, C	E
Targets	E	E	F, G

Table 4: P-DAO Messages

Note in the above that E is not an implicit Target in Storing mode, so it must be added in the RTO.

As a result the RIBs are set as follows:

Node	destination	Origin	Next Hop(s)	TrackID
E	F, G	P-DAO 1	Neighbor	(A, 129)
D	E	P-DAO 1	Neighbor	(A, 129)
C	D	P-DAO 1	Neighbor	(A, 129)
"	E	P-DAO 1	D	(A, 129)
B	C	P-DAO 2	Neighbor	(A, 129)
"	E	P-DAO 2	C	(A, 129)
A	B	P-DAO 2	Neighbor	(A, 129)
"	E	P-DAO 2	B	(A, 129)
"	F, G	P-DAO 3	E	(A, 129)

Table 5: RIB setting

Packets from A to E do not require an encapsulation. The outer headers of the packets that are forwarded along the Track have the following settings:

Header	IPv6 Source Addr.	IPv6 Dest. Addr.	TrackID in RPI
Outer	A	E	(A, 129)
Inner	X	E (X != A), F or G	N/A

Table 6: Packet Header Settings

As an example, say that A has a packet for F. Using the RIB above:

- * From P-DAO 3: A encapsulates the packet the Track signaled by P-DAO 3, with the outer header above. Now the packet destination is E.
- * From P-DAO 2: A forwards to B and B forwards to C.
- * From P-DAO 1: C forwards to D and D forwards to E; E decapsulates the packet.
- * From Neighbor Cache Entry: E delivers packets to F or G.

3.5.1.3. Segment Routing

In this formulation leverages Track Legs to combine Segments and form a Graph. The packets are source routed from a Segment to the next to adapt the path. As such, this can be seen as a form of Segment Routing [RFC8402]:

- * P-DAO 1 signals C==>D==>E-to-E
- * P-DAO 2 signals A==>B-to-B,C
- * P-DAO 3 signals F and G via the A-->C-->E-to-F,G Track

Storing Mode P-DAO 1 and 2, and Non-Storing Mode P-DAO 3, are sent to E, B and A, respectively, as follows:

	P-DAO 1 to E	P-DAO 2 to B	P-DAO 3 to A
Mode	Storing	Storing	Non-Storing
Track Ingress	A	A	A
(DODAGID, TrackID)	(A, 129)	(A, 129)	(A, 129)
SegmentID	1	2	3
VIO	C, D, E	A, B	C, E
Targets	E	C	F, G

Table 7: P-DAO Messages

Note in the above that the Segment can terminate at the loose hop as used in the example of P-DAO 1 or at the previous hop as done with P-DAO 2. Both methods are possible on any Segment joined by a loose Track Leg. P-DAO 1 generates more signaling since E is the Segment Egress when D could be, but has the benefit that it validates that the connectivity between D and E still exists.

As a result the RIBs are set as follows:

Node	destination	Origin	Next Hop(s)	TrackID
E	F, G	P-DAO 1	Neighbor	(A, 129)
D	E	P-DAO 1	Neighbor	(A, 129)
C	D	P-DAO 1	Neighbor	(A, 129)
"	E	P-DAO 1	D	(A, 129)
B	C	P-DAO 2	Neighbor	(A, 129)
A	B	P-DAO 2	Neighbor	(A, 129)
"	C	P-DAO 2	B	(A, 129)
"	E, F, G	P-DAO 3	C, E	(A, 129)

Table 8: RIB setting

Packets originated at A to E do not require an encapsulation, but carry a SRH via C. The outer headers of the packets that are forwarded along the Track have the following settings:

Header	IPv6 Source Addr.	IPv6 Dest. Addr.	TrackID in RPI
Outer	A	C till C then E	(A, 129)
Inner	X	E (X != A), F or G	N/A

Table 9: Packet Header Settings

As an example, say that A has a packet for F. Using the RIB above:

- * From P-DAO 3: A encapsulates the packet the Track signaled by P-DAO 3, with the outer header above. Now the destination in the IPv6 Header is C, and a SRH signals the final destination is E.
- * From P-DAO 2: A forwards to B and B forwards to C.
- * From P-DAO 3: C processes the SRH and sets the destination in the IPv6 Header to E.
- * From P-DAO 1: C forwards to D and D forwards to E; E decapsulates the packet.
- * From the Neighbor Cache Entry: E delivers packets to F or G.

3.5.2. Using Non-Storing Mode joining Tracks

In this formulation:

- * P-DAO 1 signals C==>D==>E-to-F,G
- * P-DAO 2 signals A==>B==>C-to-E,F,G

A==>B==>C and C==>D==>E are Tracks expressed as Non-Storing P-DAOs.

3.5.2.1. Stitched Tracks

Non-Storing Mode P-DAO 1 and 2 are sent to C and A respectively, as follows:

	P-DAO 1 to C	P-DAO 2 to A
Mode	Non-Storing	Non-Storing
Track Ingress	C	A
(DODAGID, TrackID)	(C, 131)	(A, 131)
SegmentID	1	1
VIO	D, E	B, C
Targets	F, G	E, F, G

Table 10: P-DAO Messages

As a result the RIBs are set as follows:

Node	destination	Origin	Next Hop(s)	TrackID
E	F, G	ND	Neighbor	Any
D	E	ND	Neighbor	Any
C	D	ND	Neighbor	Any
"	E, F, G	P-DAO 1	D, E	(C, 131)
B	C	ND	Neighbor	Any
A	B	ND	Neighbor	Any
"	C, E, F, G	P-DAO 2	B, C	(A, 131)

Table 11: RIB setting

Packets originated at A to E, F and G do not require an encapsulation, though it is preferred that A encapsulates and C decapsulates. Either way, they carry a SRH via B and C, and C needs to encapsulate to E, F, or G to add an SRH via D and E. The encapsulating headers of packets that are forwarded along the Track between C and E have the following settings:

Header	IPv6 Source Addr.	IPv6 Dest. Addr.	TrackID in RPI
Outer	C	D till D then E	(C, 131)
Inner	X	E, F, or G	N/A

Table 12: Packet Header Settings between C and E

As an example, say that A has a packet for F. Using the RIB above:

- * From P-DAO 2: A encapsulates the packet with destination of F in the Track signaled by P-DAO 2. The outer header has source A, destination B, an SRH that indicates C as the next loose hop, and a RPI indicating a TrackId of 131 from A's namespace, which is distinct from TrackId of 131 from C's.
- * From the SRH: Packets forwarded by B have source A, destination C, a consumed SRH, and a RPI indicating a TrackId of 131 from A's namespace. C decapsulates.
- * From P-DAO 1: C encapsulates the packet with destination of F in the Track signaled by P-DAO 1. The outer header has source C, destination D, an SRH that indicates E as the next loose hop, and a RPI indicating a TrackId of 131 from C's namespace. E decapsulates.

3.5.2.2. External routes

In this formulation:

- * P-DAO 1 signals C==>D==>E-to-E
- * P-DAO 2 signals A==>B==>C-to-C,E
- * P-DAO 3 signals F and G via the A-->E-to-F,G Track

Non-Storing Mode P-DAO 1 is sent to C and Non-Storing Mode P-DAO 2 and 3 are sent A, as follows:

	P-DAO 1 to C	P-DAO 2 to A	P-DAO 3 to A
Mode	Non-Storing	Non-Storing	Non-Storing
Track Ingress	C	A	A
(DODAGID, TrackID)	(C, 131)	(A, 129)	(A, 141)
SegmentID	1	1	1
VIO	D, E	B, C	E
Targets	E	E	F, G

Table 13: P-DAO Messages

As a result the RIBs are set as follows:

Node	destination	Origin	Next Hop(s)	TrackID
E	F, G	ND	Neighbor	Any
D	E	ND	Neighbor	Any
C	D	ND	Neighbor	Any
"	E	P-DAO 1	D, E	(C, 131)
B	C	ND	Neighbor	Any
A	B	ND	Neighbor	Any
"	C, E	P-DAO 2	B, C	(A, 129)
"	F, G	P-DAO 3	E	(A, 141)

Table 14: RIB setting

The encapsulating headers of packets that are forwarded along the Track between C and E have the following settings:

Header	IPv6 Source Addr.	IPv6 Dest. Addr.	TrackID in RPI
Outer	C	D till D then E	(C, 131)
Middle	A	E	(A, 141)
Inner	X	E, F or G	N/A

Table 15: Packet Header Settings

As an example, say that A has a packet for F. Using the RIB above:

- * From P-DAO 3: A encapsulates the packet with destination of F in the Track signaled by P-DAO 3. The outer header has source A, destination E, and a RPI indicating a TrackId of 141 from A's namespace. This recurses with:
- * From P-DAO 2: A encapsulates the packet with destination of E in the Track signaled by P-DAO 2. The outer header has source A, destination B, an SRH that indicates C as the next loose hop, and a RPI indicating a TrackId of 129 from A's namespace.
- * From the SRH: Packets forwarded by B have source A, destination C, a consumed SRH, and a RPI indicating a TrackId of 129 from A's namespace. C decapsulates.
- * From P-DAO 1: C encapsulates the packet with destination of E in the Track signaled by P-DAO 1. The outer header has source C, destination D, an SRH that indicates E as the next loose hop, and a RPI indicating a TrackId of 131 from C's namespace. E decapsulates.

3.5.2.3. Segment Routing

In this formulation:

- * P-DAO 1 signals C==>D==>E-to-E
- * P-DAO 2 signals A==>B-to-C
- * P-DAO 3 signals F and G via the A-->C-->E-to-F,G Track

Non-Storing Mode P-DAO 1 is sent to C and Non-Storing Mode P-DAO 2 and 3 are sent A, as follows:

	P-DAO 1 to C	P-DAO 2 to A	P-DAO 3 to A
Mode	Non-Storing	Non-Storing	Non-Storing
Track Ingress	C	A	A
(DODAGID, TrackID)	(C, 131)	(A, 129)	(A, 141)
SegmentID	1	1	1
VIO	D, E	B	C, E
Targets		C	F, G

Table 16: P-DAO Messages

As a result the RIBs are set as follows:

Node	destination	Origin	Next Hop(s)	TrackID
E	F, G	ND	Neighbor	Any
D	E	ND	Neighbor	Any
C	D	ND	Neighbor	Any
"	E	P-DAO 1	D, E	(C, 131)
B	C	ND	Neighbor	Any
A	B	ND	Neighbor	Any
"	C	P-DAO 2	B, C	(A, 129)
"	E, F, G	P-DAO 3	C, E	(A, 141)

Table 17: RIB setting

The encapsulating headers of packets that are forwarded along the Track between A and B have the following settings:

Header	IPv6 Source Addr.	IPv6 Dest. Addr.	TrackID in RPI
Outer	A	B till D then E	(A, 129)
Middle	A	C	(A, 141)
Inner	X	E, F or G	N/A

Table 18: Packet Header Settings

The encapsulating headers of packets that are forwarded along the Track between B and C have the following settings:

Header	IPv6 Source Addr.	IPv6 Dest. Addr.	TrackID in RPI
Outer	A	C	(A, 141)
Inner	X	E, F or G	N/A

Table 19: Packet Header Settings

The encapsulating headers of packets that are forwarded along the Track between C and E have the following settings:

Header	IPv6 Source Addr.	IPv6 Dest. Addr.	TrackID in RPI
Outer	C	D till D then E	(C, 131)
Middle	A	E	(A, 141)
Inner	X	E, F or G	N/A

Table 20: Packet Header Settings

As an example, say that A has a packet for F. Using the RIB above:

- * From P-DAO 3: A encapsulates the packet with destination of F in the Track signaled by P-DAO 3. The outer header has source A, destination C, an SRH that indicates E as the next loose hop, and a RPI indicating a TrackId of 141 from A's namespace. This recurses with:

- * From P-DAO 2: A encapsulates the packet with destination of C in the Track signaled by P-DAO 2. The outer header has source A, destination B, and a RPI indicating a TrackId of 129 from A's namespace. B decapsulates forwards to C based on a sibling connected route.
- * From the SRH: C consumes the SRH and makes the destination E.
- * From P-DAO 1: C encapsulates the packet with destination of E in the Track signaled by P-DAO 1. The outer header has source C, destination D, an SRH that indicates E as the next loose hop, and a RPI indicating a TrackId of 131 from C's namespace. E decapsulates.

3.6. Complex Tracks

To increase the reliability of the P2P transmission, this specification enables to build a collection of Legs between the same Ingress and Egress Nodes and combine them with the same TrackID, as shown in Figure 7. Legs may cross at the edges of loose hops or remain parallel.

The Segments that join the loose hops of a Leg are installed with the same TrackID as the Leg. But each individual Leg and Segment has its own P-RouteID which allows it to be managed separately. When Legs cross within respective Segment, the next loose hop (the current destination of the packet) indicates which Leg is being followed and a Segment that can reach that next loose hop is selected.

CPF

CPF

CPF

CPF

Southbound API

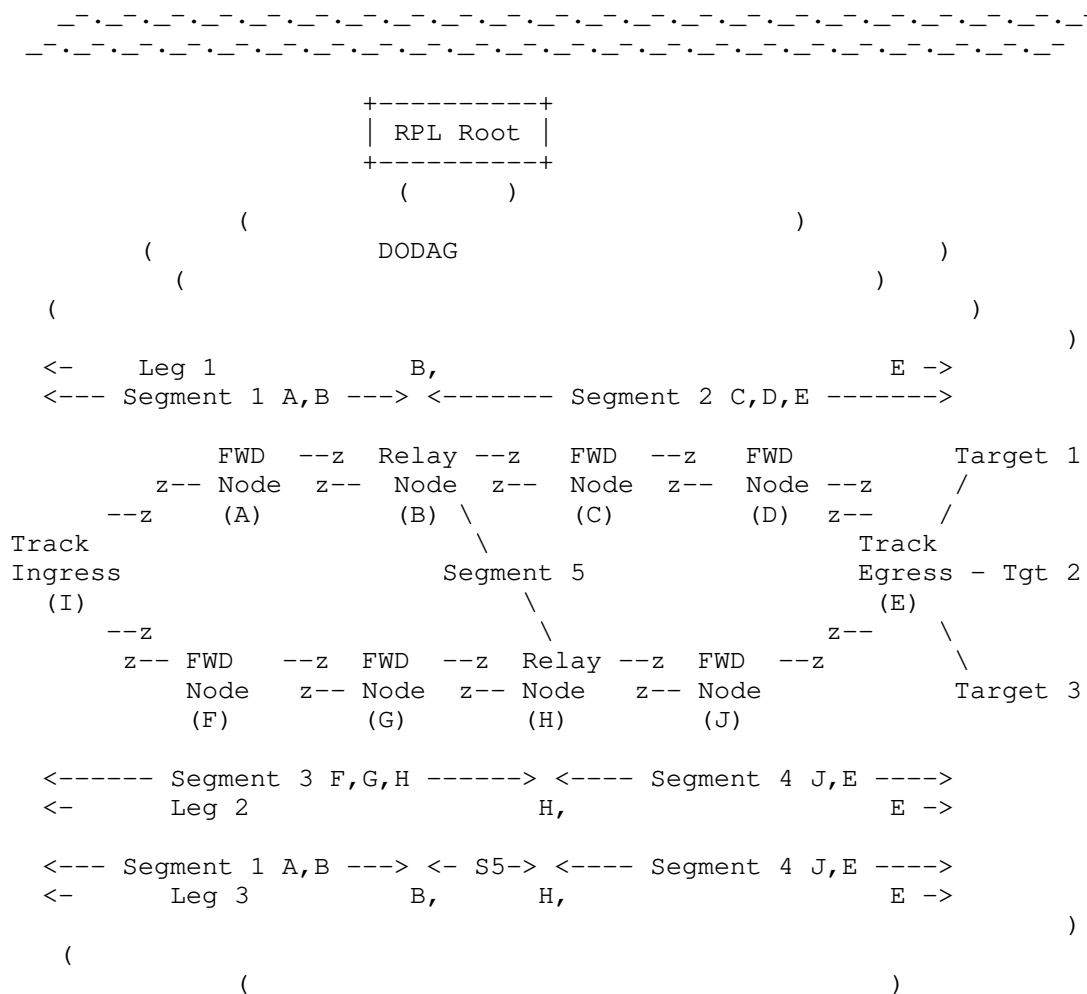


Figure 7: Segments and Tracks

Note that while this specification enables to build both Segments inside a Leg (aka East-West), such as Segment 2 above which is within Leg 1, and Inter-Leg Segments (aka North-South), such as Segment 2 above which joins Leg 1 and Leg 2, it does not signal to the Ingress which Inter-Leg Segments are available, so the use of North-South Segments and associated PAREO functions is currently limited. The only possibility available at this time is to define overlapping Legs

as illustrated in Figure 7, with Leg 3 that is congruent with Leg 1 till node B and congruent with Leg 2 from node H on, abstracting Segment 5 as an East-West Segment.

3.7. Scope and Expectations

3.7.1. External Dependencies

This specification expects that the RPL Main DODAG is operated in RPL Non-Storing Mode to sustain the exchanges with the Root. Based on its comprehensive knowledge of the parent-child relationship, the Root can form an abstracted view of the whole DODAG topology. THIS RFC adds the capability for nodes to advertise additional sibling information to complement the topological awareness of the Root to be passed on to the PCE, and enable the PCE to build more / better paths that traverse those siblings.

P-Routes require resources such as routing table space in the routers and bandwidth on the links; the amount of state that is installed in each node must be computed to fit within the node's memory, and the amount of rerouted traffic must fit within the capabilities of the transmission links. The methods used to learn the node capabilities and the resources that are available in the devices and in the network are out of scope for THIS RFC. The method to capture and report the LLN link capacity and reliability statistics are also out of scope. They may be fetched from the nodes through network management functions or other forms of telemetry such as OAM.

3.7.2. Positioning vs. Related IETF Standards

3.7.2.1. Extending 6TiSCH

The "6TiSCH Architecture" [RFC9030] leverages a centralized model that is similar to that of "Deterministic Networking Architecture" [RFC8655], whereby the device resources and capabilities are exposed to an external controller which installs routing states into the network based on its own objective functions that reside in that external entity.

3.7.2.2. Mapping to DetNet

DetNet Forwarding Nodes only understand the simple 1-to-1 forwarding sublayer transport operation along a segment whereas the more sophisticated Relay nodes can also provide service sublayer functions such as Replication and Elimination.

One possible mapping between DetNet and this specification is to signal the Relay Nodes as the hops of a Leg and the forwarding Nodes as the hops in a Segment that join the Relay nodes as illustrated in Figure 7.

3.7.2.3. Leveraging PCE

With DetNet and 6TiSCH, the component of the controller that is responsible of computing routes is a PCE. The PCE computes its routes based on its own objective functions such as described in [RFC4655], and typically controls the routes using the PCE Protocol (PCEP) by [RFC5440]. While this specification expects a PCE and while PCEP might effectively be used between the Root and the PCE, the control protocol between the PCE and the Root is out of scope.

This specification also expects a single PCE with a full view of the network. Distributing the PCE function for a large network is out of scope. This specification uses the RPL Root as a proxy to the PCE. The PCE may be collocated with the Root, or may reside in an external Controller. In that case, the protocol between the Root and the PCE is out of scope and abstracted by / mapped to RPL inside the DODAG; one possibility is for the Root to transmit the RPL DAOs with the SIOs that detail the parent/child and sibling information.

The algorithm to compute the paths and the protocol used by the PCE and the metrics and link statistics involved in the computation are also out of scope. The effectiveness of the route computation by the PCE depends on the quality of the metrics that are reported from the RPL network. Which metrics are used and how they are reported is out of scope, but the expectation is that they are mostly of long-term, statistical nature, and provide visibility on link throughput, latency, stability and availability over relatively long periods.

3.7.2.4. Providing for RAW

The RAW Architecture [RAW-ARCHI] extends the definition of Track, as being composed of East-West directional segments and North-South bidirectional segments, to enable additional path diversity, using Packet ARQ, Replication, Elimination, and Overhearing (PAREO) functions over the available paths, to provide a dynamic balance between the reliability and availability requirements of the flows and the need to conserve energy and spectrum. This specification prepares for RAW by setting up the Tracks, but only forms DODAGs, which are composed of aggregated end-to-end loose source routed Legs, joined by strict routed Segments, all oriented East-West.

The RAW Architecture defines a dataplane extension of the PCE called the Path Selection Engine (PSE), that adapts the use of the path redundancy within a Track to defeat the diverse causes of packet loss. The PSE controls the forwarding operation of the packets within a Track. This specification can use but does not impose a PSE and does not provide the policies that would select which packets are routed through which path within a Track, IOW, how the PSE may use the path redundancy within the Track. By default, the use of the available redundancy is limited to simple load balancing, and all the segments are East-West unidirectional only.

A Track may be set up to reduce the load around the Root, or to enable urgent traffic to flow more directly. This specification does not provide the policies that would decide which flows are routed through which Track. In a Non-Storing Mode RPL Instance, the Main DODAG provides a default route via the Root, and the Tracks provide more specific routes to the Track Targets.

4. Extending existing RFCs

This section explains which changes are extensions to existing specifications, and which changes are amendments to existing specification. It is expected that extensions to existing specifications do not cause existing code on legacy 6LRs to malfunction, as the extensions will simply be ignored. New code is required for an extension. Those 6LRs will be unable to participate in the new mechanisms, but may also cause projected DAOs to be impossible to install. Amendments to existing specifications are situations where there are semantic changes required to existing code, and which may require new unit tests to confirm that legacy operations will continue unaffected.

4.1. Extending RFC 6550

This specification Extends RPL [RPL] to enable the Root to install East-West routes inside a Main DODAG that is operated as Non-Storing Mode. The Root issues a Projected DAO (P-DAO) message (see Section 4.1.1) to the Track Ingress; the P-DAO message contains a new Via Information Option (VIO) that installs a strict or a loose sequence of hops to form respectively a Track Segment or a Track Leg.

The new P-DAO Request (PDR) is a new message detailed in Section 5.1. As per [RPL] section 6, if a node receives this message and it does not understand this new Code, then discards the message. When the root initiates to a node that it has not communicated with before, and to which it does not know if this specification has been implemented (by means such as capabilities), then the root SHOULD request a PDR-ACK.

A P-DAO Request (PDR) message enables a Track Ingress to request the Track from the Root. The resulting Track is also a DODAG for which the Track Ingress is the Root, the owner the address that serves as DODAGID and authoritative for the associated namespace from which the TrackID is selected. In the context of this specification, the installed route appears as a more specific route to the Track Targets, and the Track Ingress routes the packets towards the Targets via the Track using the longest match as usual.

To ensure that the PDR and P-DAO messages can flow at most times, it is RECOMMENDED that the nodes involved in a Track maintain multiple parents in the Main DODAG, advertise them all to the Root, and use them in turn to retry similar packets. It is also RECOMMENDED that the Root uses diverse source route paths to retry similar messages to the nodes in the Track.

4.1.1. Projected DAO

Section 6 of [RPL] introduces the RPL Control Message Options (CMO), including the RPL Target Option (RTO) and Transit Information Option (TIO), which can be placed in RPL messages such as the destination Advertisement Object (DAO). A DAO message signals routing information to one or more Targets indicated in RTOs, providing one hop information at a time in the TIO.

THIS RFC Amends the specification of the DAO to create the P-DAO message. This Amended DAO is signaled with a new "Projected DAO" (P) flag, see Figure 8.

A Projected DAO (P-DAO) is a special DAO message generated by the Root to install a P-Route formed of multiple hops in its DODAG. This provides a RPL-based method to install the Tracks as expected by the 6TiSCH Architecture [RFC9030] as a collection of multiple P-Routes.

The Root MUST source the P-DAO message with its address that serves as DODAGID for the main DODAG. The receiver MUST NOT accept a P-DAO message that is not sent by the Root of its DODAG and MUST ignore such message silently.

The 'P' flag is encoded in bit position 2 (to be confirmed by IANA) of the Flags field in the DAO Base Object. The Root MUST set it to 1 in a Projected DAO message. Otherwise it MUST be set to 0. It is set to 0 in Legacy implementations as specified respectively in Sections 20.11 and 6.4 of [RPL].

The P-DAO is control plane signaling and should not be stuck behind high traffic levels. The expectation is that the P-DAO message is sent as high QoS level, above that of data traffic, typically with the Network Control precedence.

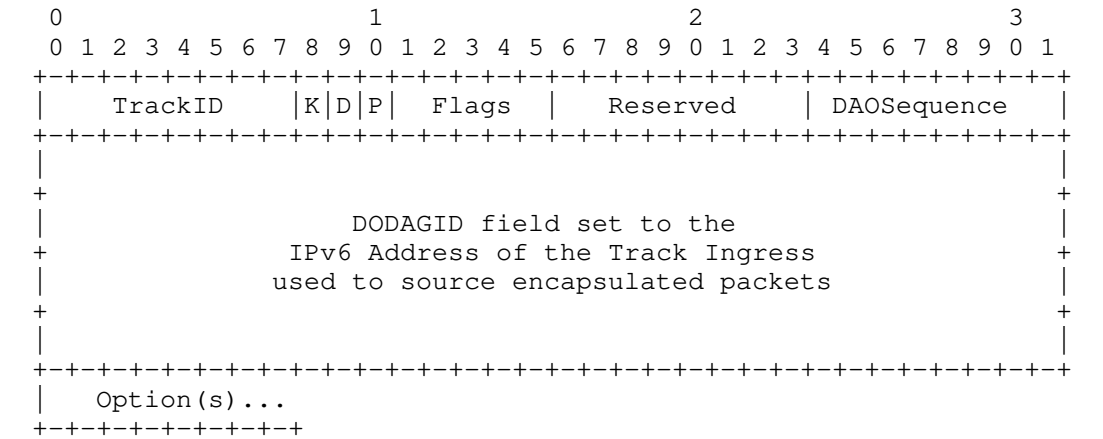


Figure 8: Projected DAO Base Object

New fields:

TrackID: The local or global RPLInstanceID of the DODAG that serves as Track, more in Section 6.3

P: 1-bit flag (position to be confirmed by IANA).

The 'P' flag is set to 1 by the Root to signal a Projected DAO, and it is set to 0 otherwise.

The D flag is set to one to signal that the DODAGID field is present. It may be set to zero if and only if the destination address of the P-DAO-ACK message is set to the IPv6 address that serves as DODAGID and it MUST be set to one otherwise, meaning that the DODAGID field MUST then be present.

In RPL Non-Storing Mode, the TIO and RTO are combined in a DAO message to inform the DODAG Root of all the edges in the DODAG, which are formed by the directed parent-child relationships. The DAO message signals to the Root that a given parent can be used to reach a given child. The P-DAO message generalizes the DAO to signal to the Track Ingress that a Track for which it is Root can be used to reach children and siblings of the Track Egress. In both cases, options may be factorized and multiple RTOs may be present to signal a collection of children that can be reached through the parent or the Track, respectively.

4.1.2. Projected DAO-ACK

THIS RFC also Amends the DAO-ACK message. The new P flag signals the projected form.

The format of the P-DAO-ACK message is thus as illustrated in Figure 9:

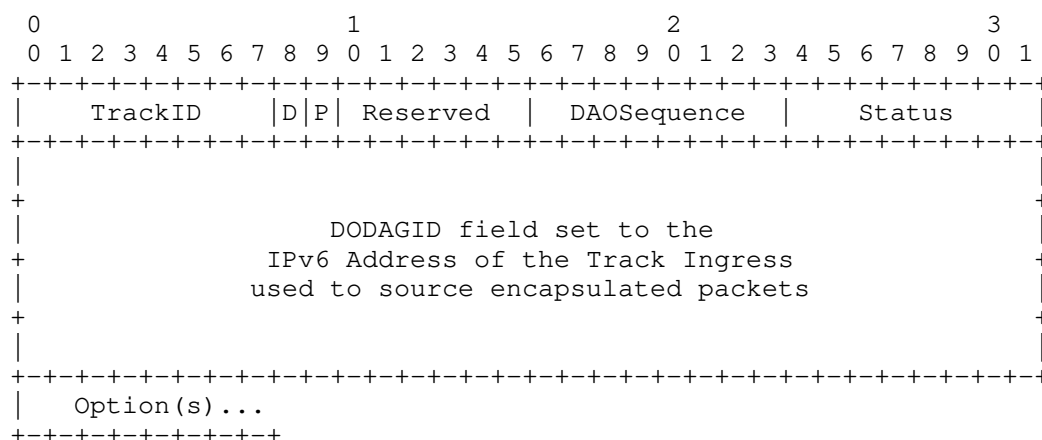


Figure 9: Projected DAO-ACK Base Object

New fields:

TrackID: The local or global RPLInstanceID of the DODAG that serves as Track, more in Section 6.3

P: 1-bit flag (position to be confirmed by IANA).

The 'P' flag is set to 1 by the Root to signal a Projected DAO, and it is set to 0 otherwise.

The D flag is set to one to signal that the DODAGID field is present. It may be set to zero if and only if the source address of the P-DAO-ACK message is set to the IPv6 address that serves as DODAGID and it MUST be set to one otherwise, meaning that the DODAGID field MUST then be present.

4.1.3. Via Information Option

THIS RFC Extends the CMO to create new objects called the Via Information Options (VIO). The VIOs are the multihop alternative to the TIO, more in Section 5.3. One VIO is the stateful Storing Mode VIO (SM-VIO); an SM-VIO installs a strict hop-by-hop P-Route called a Track Segment. The other is the Non-Storing Mode VIO (NSM-VIO); the NSM-VIO installs a loose source-routed P-Route called a Track Leg at the Track Ingress, which uses that state to encapsulate a packet IPv6_in_IPv6 with a new Routing Header (RH) to the Track Egress, more in Section 6.7.

A P-DAO contains one or more RTOs to indicate the Target (destinations) that can be reached via the P-Route, followed by exactly one VIO that signals the sequence of nodes to be followed, more in Section 6. There are two modes of operation for the P-Routes, the Storing Mode and the Non-Storing Mode, see Section 6.4.2 and Section 6.4.3 respectively for more.

4.1.4. Sibling Information Option

This specification Extends the CMO to create the Sibling Information Option (SIO). The SIO is used by a RPL Aware Node (RAN) to advertise a selection of its candidate neighbors as siblings to the Root, more in Section 5.4. The SIO is placed in DAO messages that are sent directly to the Root of the main DODAG.

4.1.5. P-DAO Request

The set of RPL Control Messages is Extended to include the P-DAO Request (PDR) and P-DAO Request Acknowledgement (PDR-ACK). These two new RPL Control Messages enable an RPL-Aware Node to request the establishment of a Track between itself as the Track Ingress Node and a Track Egress. The node makes its request by sending a new P-DAO Request (PDR) Message to the Root. The Root confirms with a new PDR-ACK message back to the requester RAN, see Section 5.1 for more.

4.1.6. Amending the RPI

Sending a Packet within a RPL Local Instance requires the presence of the abstract RPL Packet Information (RPI) described in section 11.2. of [RPL] in the outer IPv6 Header chain (see [RFC9008]). The RPI carries a local RPLInstanceID which, in association with either the source or the destination address in the IPv6 Header, indicates the RPL Instance that the packet follows.

This specification Amends [RPL] to create a new flag that signals that a packet is forwarded along a P-Route.

Projected-Route 'P': 1-bit flag. It is set to 1 in the RPI that is added in the encapsulation when a packet is sent over a Track. It is set to 0 when a packet is forwarded along the main Track, including when the packet follows a Segment that joins loose hops of the Main DODAG. The flag is not mutable en-route.

The encoding of the 'P' flag in native format is shown in Section 4.2 while the compressed format is indicated in Section 4.3.

4.1.7. Additional Flag in the RPL DODAG Configuration Option

The DODAG Configuration Option is defined in Section 6.7.6 of [RPL]. Its purpose is extended to distribute configuration information affecting the construction and maintenance of the DODAG, as well as operational parameters for RPL on the DODAG, through the DODAG. This Option was originally designed with 4 bit positions reserved for future use as Flags.

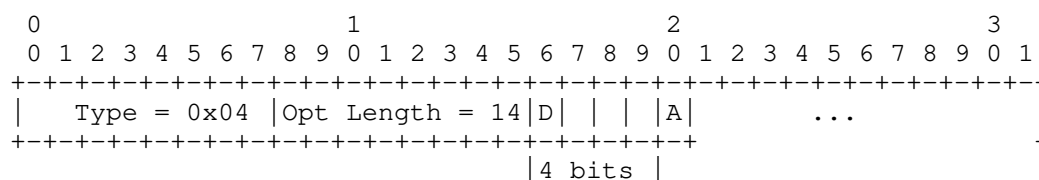


Figure 10: DODAG Configuration Option (Partial View)

This specification Amends the specification to define a new flag "Projected Routes Support" (D). The 'D' flag is encoded in bit position 0 of the reserved Flags in the DODAG Configuration Option (this is the most significant bit) (to be confirmed by IANA but there's little choice). It is set to 0 in legacy implementations as specified respectively in Sections 20.14 and 6.7.6 of [RPL].

The 'D' flag is set to 1 to indicate that this specification is enabled in the network and that the Root will install the requested Tracks when feasible upon a PDR message.

Section 4.1.2. of [RFC9008] updates [RPL] to indicate that the definition of the Flags applies to Mode of Operation values from zero (0) to six (6) only. For a MOP value of 7, the implementation MUST consider that the Root accepts PDR messages and will install Projected Routes.

The RPL DODAG Configuration option is typically placed in a DODAG Information Object (DIO) message. The DIO message propagates down the DODAG to form and then maintain its structure. The DODAG Configuration option is copied unmodified from parents to children.

[RPL] states that:

```
| Nodes other than the DODAG root MUST NOT modify this information
| when propagating the DODAG Configuration option.
```

Therefore, a legacy parent propagates the 'D' flag as set by the root, and when the 'D' flag is set to 1, it is transparently flooded to all the nodes in the DODAG.

4.2. Extending RFC 6553

"The RPL Option for Carrying RPL Information in Data-Plane Datagrams" [RFC6553] describes the RPL Option for use among RPL routers to include the abstract RPL Packet Information (RPI) described in section 11.2. of [RPL] in data packets.

The RPL Option is commonly referred to as the RPI though the RPI is really the abstract information that is transported in the RPL Option. [RFC9008] updated the Option Type from 0x63 to 0x23.

This specification Amends the RPL Option to encode the 'P' flag as follows:

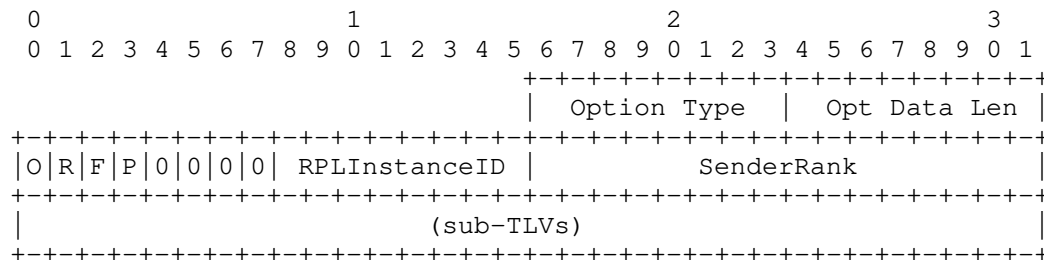


Figure 11: Amended RPL Option Format

Option Type: 0x23 or 0x63, see [RFC9008]

Opt Data Len: See [RFC6553]

'O', 'R' and 'F' flags: See [RFC6553]. Those flags MUST be set to 0 by the sender and ignored by the receiver if the 'P' flag is set.

Projected-Route 'P': 1-bit flag as defined in Section 4.1.6.

RPLInstanceID: See [RFC6553]. Indicates the TrackId if the 'P' flag is set, as discussed in Section 4.1.1.

SenderRank: See [RFC6553]. This field MUST be set to 0 by the sender and ignored by the receiver if the 'P' flag is set.

4.3. Extending RFC 8138

The 6LoWPAN Routing Header [RFC8138] specification introduces a new IPv6 over Low-Power Wireless Personal Area Network (6LoWPAN) [RFC6282] dispatch type for use in 6LoWPAN route-over topologies, which initially covers the needs of RPL data packet compression.

Section 4 of [RFC8138] presents the generic formats of the 6LoWPAN Routing Header (6LoRH) with two forms, one Elective that can be ignored and skipped when the router does not understand it, and one Critical which causes the packet to be dropped when the router cannot process it. The 'E' Flag in the 6LoRH indicates its form. In order to skip the Elective 6LoRHs, their format imposes a fixed expression of the size, whereas the size of a Critical 6LoRH may be signaled in variable forms to enable additional optimizations.

When the [RFC8138] compression is used, the Root of the Main DODAG that sets up the Track also constructs the compressed routing header (SRH-6LoRH) on behalf of the Track Ingress, which saves the complexities of optimizing the SRH-6LoRH encoding in constrained code. The SRH-6LoRH is signaled in the NSM-VIO, in a fashion that it is ready to be placed as is in the packet encapsulation by the Track Ingress.

Section 6.3 of [RFC8138] presents the formats of the 6LoWPAN Routing Header of type 5 (RPI-6LoRH) that compresses the RPI for normal RPL operation. The format of the RPI-6LoRH is not suited for P-Routes since the O,R,F flags are not used and the Rank is unknown and ignored.

This specification extends [RFC8138] to introduce a new 6LoRH, the P-RPI-6LoRH that can be used in either Elective or Critical 6LoRH form, see Table 22 and Table 23 respectively. The new 6LoRH MUST be used as a Critical 6LoRH, unless an SRH-6LoRH is present and controls the routing decision, in which case it MAY be used in Elective form.

The P-RPI-6LoRH is designed to compress the RPI along RPL P-Routes. Its format is as follows:

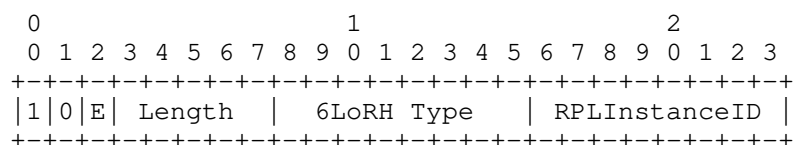


Figure 12: P-RPI-6LoRH Format

Type: IANA is requested to define the same value of the type for both Elective and Critical forms. A type of 8 is suggested.

Elective 'E': See [RFC8138]. The 'E' flag is set to 1 to indicate an Elective 6LoRH, meaning that it can be ignored when forwarding.

RPLInstanceID : In the context of this specification, the RPLInstanceID field signals the TrackID, see Section 3.4 and Section 6.3 .

Section 6.8 details how a a Track Ingress leverages the P-RPI-6LoRH Header as part of the encapsulation of a packet to place it into a Track.

5. New RPL Control Messages and Options

5.1. New P-DAO Request Control Message

The P-DAO Request (PDR) message is sent by a Node in the Main DODAG to the Root. It is a request to establish or refresh a Track where this node is Track Ingress, and signals whether an acknowledgment called PDR-ACK is requested or not. A positive PDR-ACK indicates that the Track was built and that the Roots commits to maintain the Track for the negotiated lifetime.

The main Root MAY indicate to the Track Ingress that the Track was terminated before its time and to do so, it MUST uses an asynchronous PDR-ACK with an negative status. A status of "Transient Failure" (see Section 11.10) is an indication that the PDR may be retried after a reasonable time that depends on the deployment. Other

negative status values indicate a permanent error; the tentative must be abandoned until a corrective action is taken at the application layer or through network management.

The source IPv6 address of the PDR signals the Track Ingress to-be of the requested Track, and the TrackID is indicated in the message itself. At least one RPL Target Option MUST be present in the message. If more than one RPL Target Option is present, the Root will provide a Track that reaches the first listed Target and a subset of the other Targets; the details of the subset selection are out of scope. The RTO signals the Track Egress, more in Section 6.2.

The RPL Control Code for the PDR is 0x09, to be confirmed by IANA. The format of PDR Base Object is as follows:

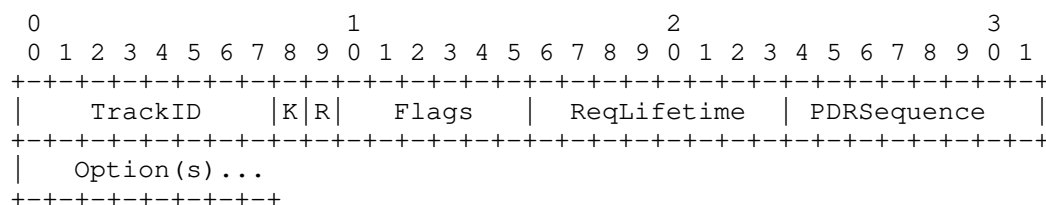


Figure 13: New P-DAO Request Format

TrackID: 8-bit field. In the context of this specification, the TrackID field signals the RPLInstanceID of the DODAG formed by the Track, see Section 3.4 and Section 6.3. To allocate a new Track, the Ingress Node must provide a value that is not in use at this time.

K: The 'K' flag is set to indicate that the recipient is expected to send a PDR-ACK back.

R: The 'R' flag is set to request a Complex Track for redundancy.

Flags: Reserved. The Flags field MUST be initialized to zero by the sender and MUST be ignored by the receiver

ReqLifetime: 8-bit unsigned integer. The requested lifetime for the Track expressed in Lifetime Units (obtained from the DODAG Configuration option).

A PDR with a fresher PDRSequence refreshes the lifetime, and a PDRLifetime of 0 indicates that the Track should be destroyed, e.g., when the application that requested the Track terminates.

PDRSequence: 8-bit wrapping sequence number, obeying the operation

in section 7.2 of [RPL]. The PDRSequence is used to correlate a PDR-ACK message with the PDR message that triggered it. It is incremented at each PDR message and echoed in the PDR-ACK by the Root.

5.2. New PDR-ACK Control Message

The new PDR-ACK is sent as a response to a PDR message with the 'K' flag set. The RPL Control Code for the PDR-ACK is 0x0A, to be confirmed by IANA. Its format is as follows:

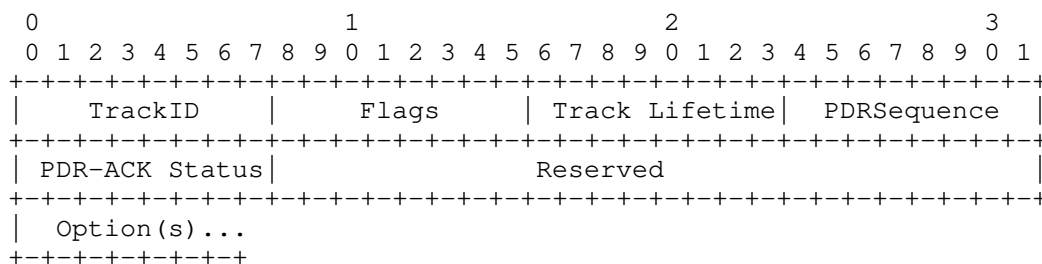


Figure 14: New PDR-ACK Control Message Format

TrackID: Set to the TrackID indicated in the TrackID field of the PDR messages that this replies to.

Flags: Reserved. The Flags field MUST initialized to zero by the sender and MUST be ignored by the receiver

Track Lifetime: Indicates that remaining Lifetime for the Track, expressed in Lifetime Units; the value of zero (0x00) indicates that the Track was destroyed or not created.

PDRSequence: 8-bit wrapping sequence number. It is incremented at each PDR message and echoed in the PDR-ACK.

PDR-ACK Status: 8-bit field indicating the completion. The PDR-ACK Status is substructured as indicated in Figure 15:

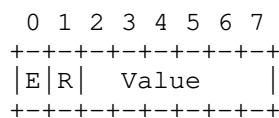


Figure 15: PDR-ACK status Format

E: 1-bit flag. Set to indicate a rejection. When not set, the

value of 0 indicates Success/Unqualified Acceptance and other values indicate "not an outright rejection".

R: 1-bit flag. Reserved, MUST be set to 0 by the sender and ignored by the receiver.

Status Value: 6-bit unsigned integer. Values depending on the setting of the 'E' flag, see Table 28 and Table 29.

Reserved: The Reserved field MUST be initialized to zero by the sender and MUST be ignored by the receiver

5.3. Via Information Options

A VIO signals the ordered list of IPv6 Via Addresses that constitutes the hops of either a Leg (using Non-Storing Mode) a Segment (using storing mode) of a Track. A Storing Mode P-DAO contains one Storing Mode VIO (SM-VIO) whereas a Non-Storing Mode P-DAO contains one Non-Storing Mode VIO (NSM-VIO)

The duration of the validity of a VIO is indicated in a Segment Lifetime field. A P-DAO message that contains a VIO with a Segment Lifetime of zero is referred as a No-Path P-DAO.

The VIO contains one or more SRH-6LoRH header(s), each formed of a SRH-6LoRH head and a collection of compressed Via Addresses, except in the case of a Non-Storing Mode No-Path P-DAO where the SRH-6LoRH header is not present.

In the case of a SM-VIO, or if [RFC8138] is not used in the data packets, then the Root MUST use only one SRH-6LoRH per Via Information Option, and the compression is the same for all the addresses, as shown in Figure 16, for simplicity.

In case of an NSM-VIO and if [RFC8138] is in use in the Main DODAG, the Root SHOULD optimize the size of the NSM-VIO if using different SRH-6LoRH Types make the VIO globally shorter; this means that more than one SRH-6LoRH may be present.

The format of the Via Information Options is as follows:

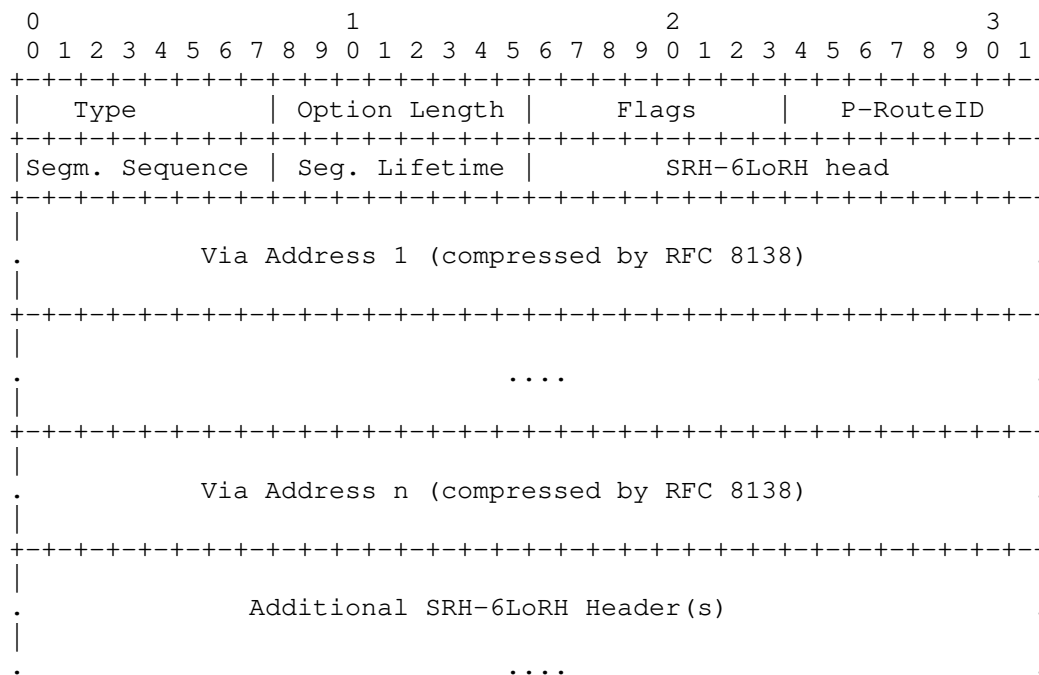


Figure 16: VIO format (uncompressed form)

Option Type: 0x0E for SM-VIO, 0x0F for NSM-VIO (to be confirmed by IANA), see Table 26

Option Length: 8-bit unsigned integer, representing the length in octets of the option, not including the Option Type and Length fields, see section 6.7.1. of [RPL]; the Option Length is variable, depending on the number of Via Addresses and the compression applied.

P-RouteID: 8-bit field that identifies a component of a Track or the Main DODAG as indicated by the TrackID field. The value of 0 is used to signal a Serial Track, i.e., made of a single segment/Leg. In an SM-VIO, the P-RouteID indicates an actual Segment. In an NSM-VIO, it indicates a Leg, that is a serial subTrack that is added to the overall topology of the Track.

Segment Sequence: 8-bit unsigned integer. The Segment Sequence obeys the operation in section 7.2 of [RPL] and the lollipop starts at 255.

When the Root of the DODAG needs to refresh or update a Segment in a Track, it increments the Segment Sequence individually for that Segment.

The Segment information indicated in the VIO deprecates any state for the Segment indicated by the P-RouteID within the indicated Track and sets up the new information.

A VIO with a Segment Sequence that is not as fresh as the current one is ignored.

A VIO for a given DODAGID with the same (TrackID, P-RouteID, Segment Sequence) indicates a retry; it MUST NOT change the Segment and MUST be propagated or answered as the first copy.

Segment Lifetime: 8-bit unsigned integer. The length of time in Lifetime Units (obtained from the Configuration option) that the Segment is usable.

The period starts when a new Segment Sequence is seen. The value of 255 (0xFF) represents infinity. The value of zero (0x00) indicates a loss of reachability.

SRH-6LoRH head: The first 2 bytes of the (first) SRH-6LoRH as shown in Figure 6 of [RFC8138]. As an example, a 6LoRH Type of 4 means that the VIA Addresses are provided in full with no compression.

Via Address: An IPv6 ULA or GUA of a node along the Segment. The VIO contains one or more IPv6 Via Addresses listed in the datapath order from Ingress to Egress. The list is expressed in a compressed form as signaled by the preceding SRH-6LoRH header.

In a Storing Mode P-DAO that updates or removes a section of an already existing Segment, the list in the SM-VIO may represent only the section of the Segment that is being updated; at the extreme, the SM-VIO updates only one node, in which case it contains only one IPv6 address. In all other cases, the list in the VIO MUST be complete.

In the case of an SM-VIO, the list indicates a sequential (strict) path through direct neighbors, the complete list starts at Ingress and ends at Egress, and the nodes listed in the VIO, including the Egress, MAY be considered as implicit Targets.

In the case of an NSM-VIO, the complete list can be loose and excludes the Ingress node, starting at the first loose hop and ending at a Track Egress; the Track Egress MUST be considered as an implicit Target, so it MUST NOT be signaled in a RPL Target Option.

5.4. Sibling Information Option

The Sibling Information Option (SIO) provides indication on siblings that could be used by the Root to form P-Routes. One or more SIO(s) may be placed in the DAO messages that are sent to the Root in Non-Storing Mode.

To advertise a neighbor node, the router MUST have an active Address Registration from that sibling using [RFC8505], for an address (ULA or GUA) that serves as identifier for the node. If this router also registers an address to that sibling, and the link has similar properties in both directions, only the router with the lowest Interface ID in its registered address needs report the SIO, with the B flag set, and the Root will assume symmetry.

The SIO carries a flag (B) that is set when similar performances can be expected both directions, so the routing can consider that the information provided for one direction is valid for both. If the SIO is effectively received from both sides then the B flag MUST be ignored. The policy that describes the performance criteria, and how they are asserted is out of scope. In the absence of an external protocol to assert the link quality, the flag SHOULD NOT be set.

The format of the SIO is as follows:

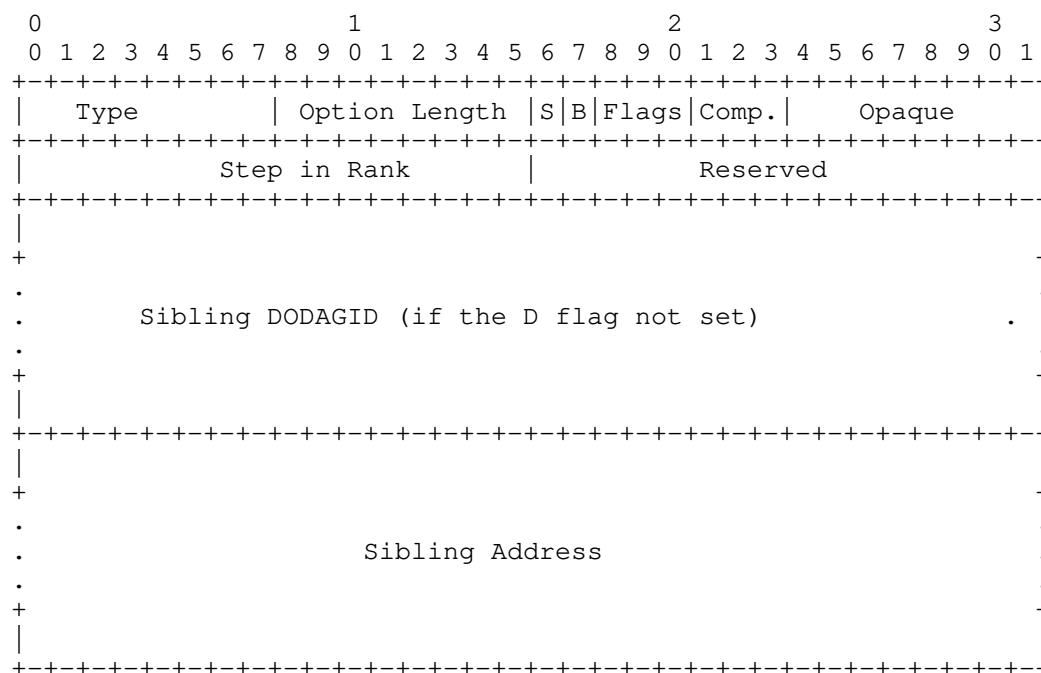


Figure 17: Sibling Information Option Format

Option Type: 0x10 for SIO (to be confirmed by IANA), see =Table 26

Option Length: 8-bit unsigned integer, representing the length in octets of the option, not including the Option Type and Length fields, see section 6.7.1. of [RPL].

Reserved for Flags: MUST be set to zero by the sender and MUST be ignored by the receiver.

B: 1-bit flag that is set to indicate that the connectivity to the sibling is bidirectional and roughly symmetrical. In that case, only one of the siblings may report the SIO for the hop. If 'B' is not set then the SIO only indicates connectivity from the sibling to this node, and does not provide information on the hop from this node to the sibling.

S: 1-bit flag that is set to indicate that sibling belongs to the same DODAG. When not set, the Sibling DODAGID is indicated.

Flags: Reserved. The Flags field MUST be initialized to zero by the sender and MUST be ignored by the receiver

Opaque: MAY be used to carry information that the node and the Root understand, e.g., a particular representation of the Link properties such as a proprietary Link Quality Information for packets received from the sibling. An industrial Alliance that uses RPL for a particular use / environment MAY redefine the use of this field to fit its needs.

Compression Type: 3-bit unsigned integer. This is the SRH-6LoRH Type as defined in figure 7 in section 5.1 of [RFC8138] that corresponds to the compression used for the Sibling Address and its DODAGID if resent. The Compression reference is the Root of the Main DODAG.

Step in Rank: 16-bit unsigned integer. This is the Step in Rank [RPL] as computed by the Objective Function between this node and the sibling, that reflects the abstract Rank increment that would be computed by the OF if the sibling was the preferred parent.

Reserved: The Reserved field MUST be initialized to zero by the sender and MUST be ignored by the receiver

Sibling DODAGID: 2 to 16 bytes, the DODAGID of the sibling in a [RFC8138] compressed form as indicated by the Compression Type field. This field is present if and only if the D flag is not set.

Sibling Address: 2 to 16 bytes, an IPv6 Address of the sibling, with a scope that MUST be make it reachable from the Root, e.g., it cannot be a Link Local Address. The IPv6 address is encoded in the [RFC8138] compressed form indicated by the Compression Type field.

An SIO MAY be immediately followed by a DAG Metric Container. In that case the DAG Metric Container provides additional metrics for the hop from the Sibling to this node.

6. Root Initiated Routing State

6.1. RPL Network Setup

To avoid the need of Path MTU Discovery, 6LoWPAN links are normally defined with a MTU of 1280 (see section 4 of [6LoWPAN]). Injecting packets in a Track typically involves an IP-in-IP encapsulation and additional IPv6 Extension Headers. This may cause a fragmentation if the resulting packets exceeds the MTU that is defined for the RPL domain.

Though fragmentation is possible in a 6LoWPAN LLN, e.g., using [6LoWPAN], [RFC8930], and/or [RFC8931], it is RECOMMENDED to allow an MTU that is larger than 1280 in the main DODAG and allows for the additional headers while exposing only 1280 to the 6LoWPAN Nodes.

6.2. Requesting a Track

This specification introduces the PDR message, used by an LLN node to request the formation of a new Track for which this node is Ingress. Note that the namespace for the TrackID is owned by the Ingress node, and in the absence of a PDR, there must be some procedure for the Root to assign TrackIDs in that namespace while avoiding collisions, more in Section 6.3.

The PDR signals the desired TrackID and the duration for which the Track should be established. Upon a PDR, the Root MAY install the Track as requested, in which case it answers with a PDR-ACK indicating the granted Track Lifetime. All the Segments MUST be of a same mode, either Storing or Non-Storing. All the Segments MUST be created with the same TrackID and the same DODAGID signaled in the P-DAO.

The Root designs the Track as it sees best, and updates / changes the Segments overtime to serve the Track as needed. Note that there is no protocol element to notify to the requesting Track Ingress when changes happen deeper down the Track, so they are transparent to the Track Ingress. If the main Root cannot maintain an expected service level, then it needs to tear down the Track completely. The Segment Lifetime in the P-DAO messages does not need to be aligned to the Requested Lifetime in the PDR, or between P-DAO messages for different Segments. The Root may use shorter lifetimes for the Segments and renew them faster than the Track is, or longer lifetimes in which case it will need to tear down the Segments if the Track is not renewed.

When the Track Lifetime that was returned in the PDR-ACK is close to elapse - vs. the trip time from the node to the Root, the requesting node SHOULD resend a PDR using the TrackID in the PDR-ACK to extend the lifetime of the Track, else the Track will time out and the Root will tear down the whole structure.

If the Track fails and cannot be restored, the Root notifies the requesting node asynchronously with a PDR-ACK with a Track Lifetime of 0, indicating that the Track has failed, and a PDR-ACK Status indicating the reason of the fault.

6.3. Identifying a Track

RPL defines the concept of an Instance to signal an individual routing topology, and multiple topologies can coexist in the same network. The RPLInstanceID is tagged in the RPI of every packet to signal which topology the packet actually follows.

This draft leverages the RPL Instance model as follows:

- * The Root MAY use P-DAO messages to add better routes in the main (Global) RPL Instance in conformance with the routing objectives in that Instance.

To achieve this, the Root MAY install a Segment along a path down the main Non-Storing Mode DODAG. This enables a loose source routing and reduces the size of the Routing Header, see Section 3.3.1. The Root MAY also install a Track Leg across the Main DODAG to complement the routing topology.

When adding a P-Route to the RPL Main DODAG, the Root MUST set the RPLInstanceID field of the P-DAO Base Object (see section 6.4.1. of [RPL]) to the RPLInstanceID of the Main DODAG, and MUST NOT use the DODAGID field. A P-Route provides a longer match to the Target Address than the default route via the Root, so it is preferred.

- * The Root MAY also use P-DAO messages to install a Track as an independent routing topology (say, Traffic Engineered) to achieve particular routing characteristics from an Ingress to an Egress Endpoints. To achieve this, the Root MUST set up a local RPL Instance (see section 5 of [RPL]), and the Local RPLInstanceID serves as TrackID. The TrackID MUST be unique for the IPv6 ULA or GUA of the Track Ingress that serves as DODAGID for the Track.

This way, a Track is uniquely identified by the tuple (DODAGID, TrackID) where the TrackID is always represented with the D flag set to 0 (see also section 5.1. of [RPL]), indicating when used in an RPI that the source address of the IPv6 packet signals the DODAGID.

The P-DAO Base Object MUST indicate the tuple (DODAGID, TrackID) that identifies the Track as shown in Figure 8, and the P-RouteID that identifies the P-Route MUST be signaled in the VIO as shown in Figure 16.

The Track Ingress is the Root of the DODAG ID formed by the local RPL Instance. It owns the namespace of its TrackIDs, so it can pick any unused value to request a new Track with a PDR. In a

particular deployment where PDR are not used, a portion of the namespace can be administratively delegated to the main Root, meaning that the main Root is authoritative for assigning the TrackIDs for the Tracks it creates.

With this specification, the Root is aware of all the active Tracks, so it can also pick any unused value to form Tracks without a PDR. To avoid a collision of the Root and the Track Ingress picking the same value at the same time, it is RECOMMENDED that the Track Ingress starts allocating the ID value of the Local RPLInstanceID (see section 5.1. of [RPL]) used as TrackIDs with the value 0 incrementing, while the Root starts with 63 decrementing.

6.4. Installing a Track

A Serial Track can be installed by a single P-Route that signals the sequence of consecutive nodes, either in Storing Mode as a single-Segment Track, or in Non-Storing Mode as a single-Leg Track. A single-Leg Track can be installed as a loose Non-Storing Mode P-Route, in which case the next loose entry must recursively be reached over a Serial Track.

A Complex Track can be installed as a collection of P-Routes with the same DODAGID and Track ID. The Ingress of a Non-Storing Mode P-Route is the owner and Root of the DODAGID. The Ingress of a Storing Mode P-Route must be either the owner of the DODAGID, or a hop of a Leg of the same Track. In the latter case, the Targets of the P-Route must include the next hop of the Leg if there is one, to ensure forwarding continuity. In the case of a Complex Track, each Segment is maintained independently and asynchronously by the Root, with its own lifetime that may be shorter, the same, or longer than that of the Track.

A route along a Track for which the TrackID is not the RPLInstanceID of the Main DODAG MUST be installed with a higher precedence than the routes along the Main DODAG, meaning that:

- * Longest match MUST be the prime comparison for routing.
- * In case of equal length match, the route along the Track MUST be preferred vs. the one along the Main DODAG.
- * There SHOULD NOT be 2 different Tracks leading to the same Target from same Ingress node, unless there's a policy for selecting which packets use which Track; such policy is out of scope.

- * A packet that was routed along a Track MUST NOT be routed along the main DODAG again; if the destination is not reachable as a neighbor by the node where the packet exits the Track then the packet MUST be dropped.

6.4.1. Signaling a Projected Route

This draft adds a capability whereby the Root of a main RPL DODAG installs a Track as a collection of P-Routes, using a Projected-DAO (P-DAO) message for each individual Track Leg or Segment. The P-DAO signals a collection of Targets in the RPL Target Option(s) (RTO). Those Targets can be reached via a sequence of routers indicated in a VIO.

Like a classical DAO message, a P-DAO causes a change of state only if it is "new" per section 9.2.2. "Generation of DAO Messages" of the RPL specification [RPL]; this is determined using the Segment Sequence information from the VIO as opposed to the Path Sequence from a TIO. Also, a Segment Lifetime of 0 in a VIO indicates that the P-Route associated to the Segment is to be removed. There are two Modes of operation for the P-Routes, the Storing and the Non-Storing Modes.

A P-DAO message MUST be sent from the address of the Root that serves as DODAGID for the Main DODAG. It MUST contain either exactly one sequence of one or more RTOs followed one VIO, or any number of sequences of one or more RTOs followed by one or more TIOs. The former is the normal expression for this specification, where as the latter corresponds to the variation for lesser constrained environments described in Section 7.2.

A P-DAO that creates or updates a Track Leg MUST be sent to a GUA or a ULA of the Ingress of the Leg; it must contain the full list of hops in the Leg unless the Leg is being removed. A P-DAO that creates a new Track Segment MUST be sent to a GUA or a ULA of the Segment Egress and MUST signal the full list of hops in Segment; a P-DAO that updates (including deletes) a section of a Segment MUST be sent to the first node after the modified Segment and signal the full list of hops in the section starting at the node that immediately precedes the modified section.

In Non-Storing Mode, as discussed in Section 6.4.3, the Root sends the P-DAO to the Track Ingress where the source-routing state is applied, whereas in Storing Mode, the P-DAO is sent to the last node on the installed path and forwarded in the reverse direction, installing a Storing Mode state at each hop, as discussed in Section 6.4.2. In both cases the Track Ingress is the owner of the Track, and it generates the P-DAO-ACK when the installation is successful.

If the 'K' Flag is present in the P-DAO, the P-DAO must be acknowledged using a DAO-ACK that is sent back to the address of the Root from which the P-DAO was received. In most cases, the first node of the Leg, Segment, or updated section of the Segment is the node that sends the acknowledgment. The exception to the rule is when an intermediate node in a Segment fails to forward a Storing Mode P-DAO to the previous node in the SM-VIO.

In a No-Path Non-Storing Mode P-DAO, the SRH-6LoRH MUST NOT be present in the NSM-VIO; the state in the Ingress is erased regardless. In all other cases, a VIO MUST contain at least one Via Address, and a Via Address MUST NOT be present more than once, which would create a loop.

A node that processes a VIO MAY verify whether one of these conditions happen, and when so, it MUST ignore the P-DAO and reject it with a RPL Rejection Status of "Error in VIO" in the DAO-ACK, see Section 11.16.

Other errors than those discussed explicitly that prevent the installing the route are acknowledged with a RPL Rejection Status of "Unqualified Rejection" in the DAO-ACK.

6.4.2. Installing a Track Segment with a Storing Mode P-Route

As illustrated in Figure 18, a Storing Mode P-DAO installs a route along the Segment signaled by the SM-VIO towards the Targets indicated in the Target Options. The Segment is to be included in a DODAG indicated by the P-DAO Base Object, that may be the one formed by the RPL Main DODAG, or a Track associated with a local RPL Instance.

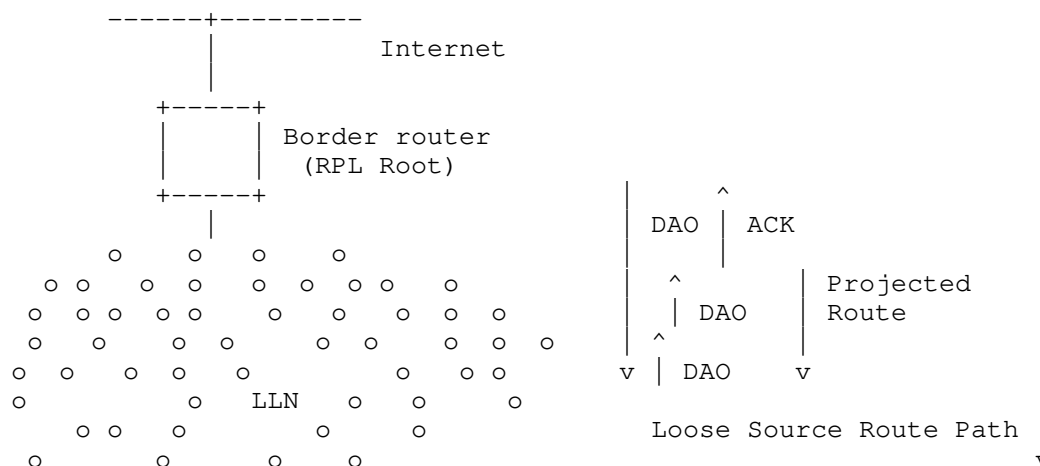


Figure 18: Projecting a route

In order to install the relevant routing state along the Segment , the Root sends a unicast P-DAO message to the Track Egress router of the routing Segment that is being installed. The P-DAO message contains a SM-VIO with the strict sequence of Via Addresses. The SM-VIO follows one or more RTOs indicating the Targets to which the Track leads. The SM-VIO contains a Segment Lifetime for which the state is to be maintained.

The Root sends the P-DAO directly to the Egress node of the Segment. In that P-DAO, the destination IP address matches the last Via Address in the SM-VIO. This is how the Egress recognizes its role. In a similar fashion, the Segment Ingress node recognizes its role as it matches first Via Address in the SM-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) based on its existing tables. If one of the Targets is not known, the node MUST answer to the Root with a DAO-ACK listing the unreachable Target(s) in an RTO and a rejection status of "Unreachable Target".

If the Egress node can reach all the Targets, then it forwards the P-DAO with unchanged content to its 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 Address the precedes the one that contain the address of the propagating node, which is used as source of the message.

Upon receiving a propagated DAO, all except the Egress router MUST install a route towards the DAO Target(s) via their successor in the SM-VIO. A router that cannot store the routes to all the Targets in a P-DAO MUST reject the P-DAO by sending a DAO-ACK to the Root with a Rejection Status of "Out of Resources" as opposed to forwarding the DAO to its predecessor in the list. The router MAY install additional routes towards the VIA Addresses that are the SM-VIO after self, if any, but in case of a conflict or a lack of resource, the route(s) to the Target(s) are the ones that must be installed in priority.

If a router cannot reach its predecessor in the SM-VIO, the router MUST send the DAO-ACK to the Root with a Rejection Status of "Predecessor Unreachable".

The process continues till the P-DAO is propagated to Ingress router of the Segment, which answers with a DAO-ACK to the Root. The Root always expects a DAO-ACK, either from the Track Ingress with a positive status or from any node along the segment with a negative status. If the DAO-ACK is not received, the Root may retry the DAO with the same TID, or tear down the route.

6.4.3. Installing a Track Leg with a Non-Storing Mode P-Route

As illustrated in Figure 19, a Non-Storing Mode P-DAO installs a source-routed path within the Track indicated by the P-DAO Base Object, towards the Targets indicated in the Target Options. The source-routed path requires a Source-Routing header which implies an IP-in-IP encapsulation to add the SRH to an existing packet. It is sent to the Track Ingress which creates a tunnel associated with the Track, and connected routes over the tunnel to the Targets in the RTO. The tunnel encapsulation MUST incorporate a routing header via the list addresses listed in the VIO in the same order. The content of the NSM-VIO starting at the first SRH-6LoRH header MUST be used verbatim by the Track Ingress when it encapsulates a packet to forward it over the Track.

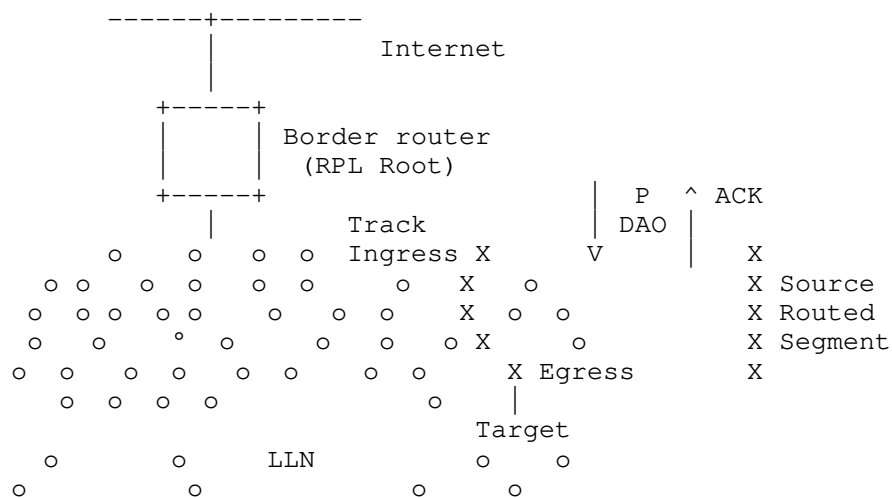


Figure 19: Projecting a Non-Storing Route

The next entry in the source-routed path must be either a neighbor of the previous entry, or reachable as a Target via another P-Route, either Storing or Non-Storing, which implies that the nested P-Route has to be installed before the loose sequence is, and that P-Routes must be installed from the last to the first along the datapath. For instance, a Segment of a Track must be installed before the Leg(s) of the same Track that use it, and stitched Segments must be installed in order from the last that reaches to the Targets to the first.

If the next entry in the loose sequence is reachable over a Storing Mode P-Route, it MUST be the Target of a Segment and the Ingress of a next segment, both already setup; the segments are associated with the same Track, which avoids the need of an additional encapsulation. For instance, in Section 3.5.1.3, Segments A==>B-to-C and C==>D==>E-to-F must be installed with Storing Mode P-DAO messages 1 and 2 before the Track A-->C-->E-to-F that joins them can be installed with Non-Storing Mode P-DAO 3.

Conversely, if it is reachable over a Non-Storing Mode P-Route, the next loose source-routed hop of the inner Track is a Target of a previously installed Track and the Ingress of a next Track, which requires a de- and a re-encapsulation when switching the outer Tracks that join the loose hops. This is exemplified in Section 3.5.2.3 where Non-Storing Mode P-DAO 1 and 2 install strict Tracks that Non-Storing Mode P-DAO 3 joins as a super Track. In such a case, packets are subject to double IP-in-IP encapsulation.

6.5. Tearing Down a P-Route

A P-DAO with a lifetime of 0 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. To tear down a Track, the Root must tear down all the Track Segments and Legs that compose it one by one.

Since the state about a Leg of a Track is located only on the Ingress Node, the Root cleans up the Leg by sending an NSM-VIO to the Ingress indicating the TrackID and the P-RouteID of the Leg being removed, a Segment Lifetime of 0 and a newer Segment Sequence. The SRH-6LoRH with the Via Addresses in the NSM-VIO are not needed; it SHOULD not be placed in the message and MUST be ignored by the receiver. Upon that NSM-VIO, the Ingress node removes all state for that Track if any, and replies positively anyway.

The Root cleans up a section of a Segment by sending an SM-VIO to the last node of the Segment, with the TrackID and the P-RouteID of the Segment being updated, a Segment Lifetime of zero (0) and a newer Segment Sequence. The Via Addresses in the SM-VIO indicates the section of the Segment being modified, from the first to the last node that is impacted. This can be the whole Segment if it is totally removed, or a sequence of one or more nodes that have been bypassed by a Segment update.

The No-Path P-DAO is forwarded normally along the reverse list, even if the intermediate node does not find a Segment state to clean up. This results in cleaning up the existing Segment state if any, as opposed to refreshing an existing one or installing a new one.

6.6. Maintaining a Track

Repathing a Track Segment or Leg may cause jitter and packet misordering. For critical flows that require timely and/or in-order delivery, it might be necessary to deploy the PAREO functions [RAW-ARCHI] over a highly redundant Track. This specification allows to use more than one Leg for a Track, and 1+N packet redundancy.

This section provides the steps to ensure that no packet is lost due to the operation itself. This is ensured by installing the new section from its last node to the first, so when an intermediate node installs a route along the new section, all the downstream nodes in the section have already installed their own. The disabled section is removed when the packets in-flight are forwarded along the new section as well.

6.6.1. Maintaining a Track Segment

To modify a section of a Segment between a first node and a second, downstream node (which can be the Ingress and Egress), while conserving those nodes in the Segment, the Root sends an SM-VIO to the second node indicating the sequence of nodes in the new section of the Segment. The SM-VIO indicates the TrackID and the P-RouteID of the Segment being updated, and a newer Segment Sequence. The P-DAO is propagated from the second to the first node and on the way, it updates the state on the nodes that are common to the old and the new section of the Segment and creates a state in the new nodes.

When the state is updated in an intermediate node, that node might still receive packets that were in flight from the Ingress to self over the old section of the Segment. Since the remainder of the Segment is already updated, the packets are forwarded along the new version of the Segment from that node on.

After a reasonable time to enable the deprecated sections to empty, the Root tears down the remaining section(s) of the old segments are torn down as described in Section 6.5.

6.6.2. Maintaining a Track Leg

This specification allows the Root to add Legs to a Track by sending a Non-Storing Mode P-DAO to the Ingress associated to the same TrackID, and a new Segment ID. If the Leg is loose, then the Segments that join the hops must be created first. It makes sense to add a new Leg before removing one that is becoming excessively lossy, and switch to the new Leg before removing the old. Dropping a Track before the new one is installed would reroute the traffic via the root; this may augment the latency beyond acceptable thresholds, and load the network near the root. This may also cause loops in the case of stitched Tracks; the packets that cannot be injected in the second Track may be routed back at reinjected at the Ingress of the first.

It is also possible to update a Track Leg by sending a Non-Storing Mode P-DAO to the Ingress with the same Segment ID, an incremented Segment Sequence, and the new complete list of hops in the NSM-VIO. Updating a live Leg means changing one or more of the intermediate loose hops, and involves laying out new Segments from and to the new loose hops before the NSM-VIO for the new Leg is issued.

Packets that are in flight over the old version of the Track Leg still follow the old source route path over the old Segments. After a reasonable time to enable the deprecated Segments to empty, the Root tears down those Segments as described in Section 6.5.

6.7. Encapsulating and Forwarding Along a Track

When injecting a packet in a Track, the Ingress router must encapsulate the packet using IP-in-IP to add the Source Routing Header with the final destination set to the Track Egress.

All properties of a Track operations are inherited from the main RPL Instance that is used to install the Track. For instance, the use of compression per [RFC8138] is determined by whether it is used in the RPL Main DODAG, e.g., by setting the "T" flag [RFC9035] in the RPL configuration option.

The Track Ingress that places a packet in a Track encapsulates it with an IP-in-IP header, a Routing Header, and an IPv6 Hop-by-Hop Option Header that contains the RPL Packet Information (RPI) as follows:

- * In the uncompressed form the source of the packet is the address that this router uses as DODAGID for the Track, the destination is the first Via Address in the NSM-VIO, and the RH is a Source Routing Header (SRH) [RFC6554] that contains the list of the remaining Via Addresses terminating by the Track Egress.
- * The preferred alternate in a network where 6LoWPAN Header Compression [RFC6282] is used is to leverage "IPv6 over Low-Power Wireless Personal Area Network (6LoWPAN) Paging Dispatch" [RFC8025] to compress the RPL artifacts as indicated in [RFC8138].

In that case, the source routed header is the exact copy of the (chain of) SRH-6LoRH found in the NSM-VIO, also terminating by the Track Egress. The RPI-6LoRH is appended next, followed by an IP-in-IP 6LoRH Header that indicates the Ingress router in the Encapsulator Address field, see as a similar case Figure 20 of [RFC9035].

To signal the Track in the packet, this specification leverages the RPL Forwarding model follows:

- * In the data packets, the Track DODAGID and the TrackID MUST be respectively signaled as the IPv6 Source Address and the RPLInstanceID field of the RPI that MUST be placed in the outer chain of IPv6 Headers.

The RPI carries a local RPLInstanceID called the TrackID, which, in association with the DODAGID, indicates the Track along which the packet is forwarded.

The D flag in the RPLInstanceID MUST be set to 0 to indicate that the source address in the IPv6 header is set to the DODAGID, more in Section 6.3.

- * This draft conforms to the principles of [RFC9008] with regards to packet forwarding and encapsulation along a Track, as follows:
 - With this draft, the Track is a RPL DODAG. From the perspective of that DODAG, the Track Ingress is the Root, the Track Egress is a RPL-Aware 6LR, and neighbors of the Track Egress that can be reached via the Track, but are external to it, are external destinations and treated as RPL-Unaware Leaves (RULs). The encapsulation rules in [RFC9008] apply.
 - If the Track Ingress is the originator of the packet and the Track Egress is the destination of the packet, there is no need for an encapsulation.
 - So the Track Ingress must encapsulate the traffic that it did not originate, and add an RPI.

A packet that is being routed over the RPL Instance associated to a first Non-Storing Mode Track MAY be placed (encapsulated) in a second Track to cover one loose hop of the first Track as discussed in more details Section 3.5.2.3. On the other hand, a Storing Mode Track must be strict and a packet that it placed in a Storing Mode Track MUST follow that Track till the Track Egress.

The forwarding of a packet along a track will fail if the Track continuity is broken, e.g.:

- * In the case of a strict path along a Segment, if the next strict hop is not reachable, the packet is dropped.
- * In the case of a loose source-routed path, when the loose next hop is not a neighbor, there must be a Segment of the same Track to that loose next hop. When that is the case the packet is forwarded to the next hop along that segment, or a common neighbor with the loose next hop, on which case the packet is forwarded to that neighbor, or another Track to the loose next hop for which this node or a neighbor is Ingress; in the last case, another encapsulation takes place and the process possibly recurses; otherwise the packet is dropped.

- * When a Track Egress extracts a packet from a Track (decapsulates the packet), the destination of the inner packet must be either this node or a direct neighbor, or a Target of another Segment of the same Track for which this node is Ingress, otherwise the packet MUST be dropped.

In case of a failure forwarding a packet along a Segment, e.g., the next hop is unreachable, the node that discovers the fault MUST send an ICMPv6 Error message [RFC4443] to the Root, with a new Code "Error in P-Route" (See Section 11.15). The Root can then repair by updating the broken Segment and/or Tracks, and in the case of a broken Segment, remove the leftover sections of the segment using SM-VIOs with a lifetime of 0 indicating the section of one or more nodes being removed (See Section 6.6).

In case of a permanent forwarding error along a Source Route path, the node that fails to forward SHOULD send an ICMP error with a code "Error in Source Routing Header" back to the source of the packet, as described in section 11.2.2.3. of [RPL]. Upon this message, the encapsulating node SHOULD stop using the source route path for a reasonable period of time which duration depends on the deployment, and it SHOULD send an ICMP message with a Code "Error in P-Route" to the Root. Failure to follow these steps may result in packet loss and wasted resources along the source route path that is broken.

Either way, the ICMP message MUST be throttled in case of consecutive occurrences. It MUST be sourced at the ULA or a GUA that is used in this Track for the source node, so the Root can establish where the error happened.

The portion of the invoking packet that is sent back in the ICMP message SHOULD record at least up to the RH if one is present, and this hop of the RH SHOULD be consumed by this node so that the destination in the IPv6 header is the next hop that this node could not reach. If a 6LoWPAN Routing Header (6LoRH) [RFC8138] is used to carry the IPv6 routing information in the outer header then that whole 6LoRH information SHOULD be present in the ICMP message.

6.8. Compression of the RPL Artifacts

When using [RFC8138] in the Main DODAG operated in Non-Storing Mode in a 6LoWPAN LLN, a typical packet that circulates in the Main DODAG is formatted as shown in Figure 20, representing the case where :

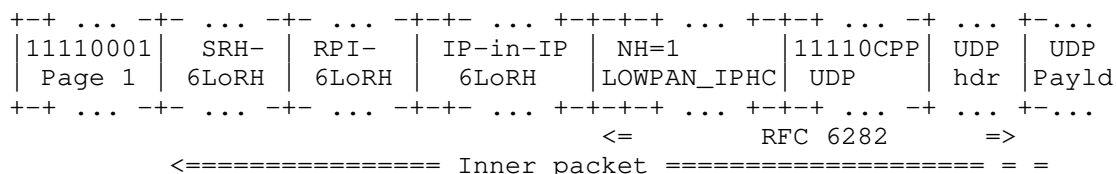


Figure 20: A Packet as Forwarded along the Main DODAG

Since there is no page switch between the encapsulated packet and the encapsulation, the first octet of the compressed packet that acts as page selector is actually removed at encapsulation, so the inner packet used in the descriptions below start with the SRH-6LoRH, and is verbatim the packet represented in Figure 20 from the second octet on.

When encapsulating that inner packet to place it in the Track, the first header that the Ingress appends at the head of the inner packet is an IP-in-IP 6LoRH Header; in that header, the encapsulator address, which maps to the IPv6 source address in the uncompressed form, contains a GUA or ULA IPv6 address of the Ingress node that serves as DODAG ID for the Track, expressed in the compressed form and using the DODAGID of the Main DODAG as compression reference. If the address is compressed to 2 bytes, the resulting value for the Length field shown in Figure 21 is 3, meaning that the SRH-6LoRH as a whole is 5-octets long.

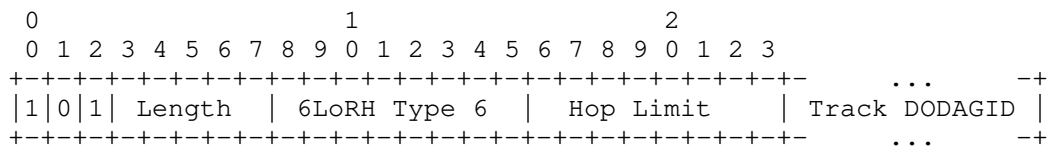


Figure 21: The IP-in-IP 6LoRH Header

At the head of the resulting sequence of bytes, the track Ingress then adds the RPI that carries the TrackID as RPIinstanceID as a P-RPI-6LoRH Header, as illustrated in Figure 12, using the TrackID as RPIinstanceID. Combined with the IP-in-IP 6LoRH Header, this allows to identify the Track without ambiguity.

The SRH-6LoRH is then added at the head of the resulting sequence of bytes as a verbatim copy of the content of the SR-VIO that signaled the selected Track Leg.

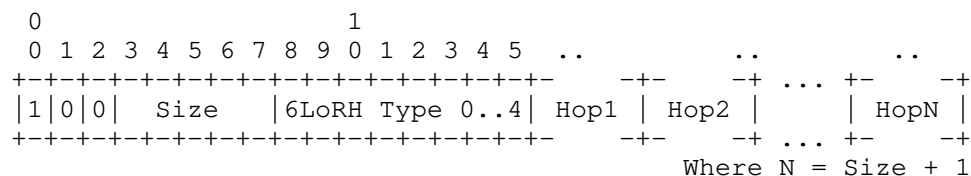
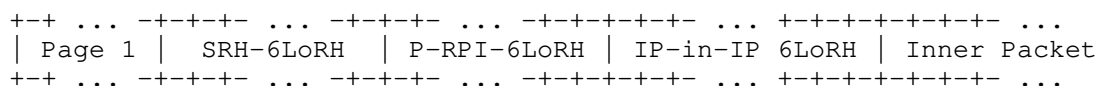


Figure 22: The SRH 6LoRH Header

The format of the resulting encapsulated packet in [RFC8138] compressed form is illustrated in Figure 23:



Signals : Loose Hops : TrackID : Track DODAGID :

Figure 23: A Packet as Forwarded along a Track

7. Lesser Constrained Variations

7.1. Storing Mode Main DODAG

This specification expects that the Main DODAG is operated in Non-Storing Mode. The reasons for that limitation are mostly related to LLN operations, power and spectrum conservation:

- * In Non-Storing Mode The Root already possesses the DODAG topology, so the additional topological information is reduced to the siblings.
- * The downwards routes are updated with unicast messages to the Root, which ensures that the Root can reach back to the LLN nodes after a repair faster than in the case of Storing Mode. Also the Root can control the use of the path diversity in the DODAG to reach to the LLN nodes. For both reasons, Non-Storing Mode provides better capabilities for the Root to maintain the P-Routes.
- * When the Main DODAG is operated in Non-Storing Mode, P-Routes enable loose Source Routing, which is only an advantage in that mode. Storing Mode does not use Source Routing Headers, and does not derive the same benefits from this capability.

On the other hand, since RPL is a Layer-3 routing protocol, its applicability extends beyond LLNs to a generic IP network. RPL requires fewer resources than alternative IGPs like OSPF, ISIS,

EIGRP, BABEL or RIP at the expense of a route stretch vs. the shortest path routes to a destination that those protocols compute. P-Routes add the capability to install shortest and/or constrained routes to special destinations such as discussed in section A.9.4. of the ANIMA ACP [RFC8994].

In a powered and wired network, when enough memory to store the needed routes is available, the RPL Storing Mode proposes a better trade-off than the Non-Storing, as it reduces the route stretch and lowers the load on the Root. In that case, the control path between the Root and the LLN nodes is highly available compared to LLNs, and the nodes can be reached to maintain the P-Routes at most times.

This section specifies the additions that are needed to support Projected Routes when the Main DODAG is operated in Storing Mode. As long as the RPI can be processed adequately by the dataplane, the changes to this specification are limited to the DAO message. The Track structure, routes and forwarding operations remain the same. Since there is no capability negotiation, the expectation is that all the nodes in the network support this specification in the same fashion, or are configured the same way through management.

In Storing Mode, the Root misses the Child to Parent relationship that forms the Main DODAG, as well as the sibling information. To provide that knowledge the nodes in the network MUST send additional DAO messages that are unicast to the Root as Non-Storing DAO messages are.

In the DAO message, the originating router advertises a set of neighbor nodes using Sibling Information Options (SIO)s, regardless of the relative position in the DODAG of the advertised node vs. this router.

The DAO message MUST be formed as follows:

- * The originating router is identified by the source address of the DAO. That address MUST be the one that this router registers to neighbor routers so the Root can correlate the DAOs from those routers when they advertise this router as their neighbor. The DAO contains one or more sequences of one Transit Information Option and one or more Sibling Information Options. There is no RPL Target Option so the Root is not confused into adding a Storing Mode route to the Target.

- * The TIO is formed as in Storing Mode, and the Parent Address is not present. The Path Sequence and Path Lifetime fields are aligned with the values used in the Address Registration of the node(s) advertised in the SIO, as explained in Section 9.1. of [RFC9010]. Having similar values in all nodes allows to factorise the TIO for multiple SIOs as done with [RPL].
- * The TIO is followed by one or more SIOs that provide an address (ULA or GUA) of the advertised neighbor node.

But the RPL routing information headers may not be supported on all type of routed network infrastructures, especially not in high-speed routers. When the RPI is not supported in the dataplane, there cannot be local RPL Instances and RPL can only operate as a single topology (the Main DODAG). The RPL Instance is that of the Main DODAG and the Ingress node that encapsulates is not the Root. The routes along the Tracks are alternate routes to those available along the Main DODAG. They MAY conflict with routes to children and MUST take precedence in the routing table. The Targets MUST be adjacent to the Track Egress to avoid loops that may form if the packet is reinjected in the Main DODAG.

7.2. A Track as a Full DODAG

This specification builds parallel or crossing Track Legs as opposed to a more complex DODAG with interconnections at any place desirable. The reason for that limitation is related to constrained node operations, and capability to store large amount of topological information and compute complex paths:

- * With this specification, the node in the LLN has no topological awareness, and does not need to maintain dynamic information about the link quality and availability.
- * The Root has a complete topological information and statistical metrics that allow it or its PCE to perform a global optimization of all Tracks in its DODAG. Based on that information, the Root computes the Track Leg and predigest the source route paths.
- * The node merely selects one of the proposed paths and applies the associated pre-computed routing header in the encapsulation. This alleviates both the complexity of computing a path and the compressed form of the routing header.

The RAW Architecture [RAW-ARCHI] actually expects the PSE at the Track Ingress to react to changes in the forwarding conditions along the Track, and reroute packets to maintain the required degree of reliability. To achieve this, the PSE need the full richness of a DODAG to form any path that could make meet the Service Level Objective (SLO).

This section specifies the additions that are needed to turn the Track into a full DODAG and enable the main Root to provide the necessary topological information to the Track Ingress. The expectation is that the metrics that the PSE uses are of an order other than that of the PCE, because of the difference of time scale between routing and forwarding, more in [RAW-ARCHI]. It follows that the PSE will learn the metrics it needs from an alternate source, e.g., OAM frames.

To pass the topological information to the Ingress, the Root uses a P-DAO messages that contains sequences of Target and Transit Information options that collectively represent the Track, expressed in the same fashion as in classical Non-Storing Mode. The difference is that the Root is the source as opposed to the destination, and can report information on many Targets, possibly the full Track, with one P-DAO.

Note that the Path Sequence and Lifetime in the TIO are selected by the Root, and that the Target/Transit information tuples in the P-DAO are not those received by the Root in the DAO messages about the said Targets. The Track may follow sibling routes and does not need to be congruent with the Main DODAG.

8. Profiles

THIS RFC provides a set of tools that may or may not be needed by an implementation depending on the type of application it serves. This sections described profiles that can be implemented separately and can be used to discriminate what an implementation can and cannot do. This section describes profiles that enable to implement only a portion of this specification to meet a particular use case.

Profiles 0 to 2 operate in the Main RPL Instance and do not require the support of local RPL Instances or the indication of the RPL Instance in the data plane. Profile 3 and above leverage Local RPL Instances to build arbitrary Tracks Rooted at the Track Ingress and using its namespace for TrackID.

Profiles 0 and 1 are REQUIRED by all implementations that may be used in LLNs; Profiles 1 leverages Storing Mode to reduce the size of the Source Route Header in the most common LLN deployments. Profile 2 is

RECOMMENDED in high speed / wired environment to enable traffic Engineering and network automation. All the other profile / environment combinations are OPTIONAL.

Profile 0 Profile 0 is the Legacy support of [RPL] Non-Storing Mode, with default routing Northwards (up) and strict source routing Southwards (down the main DODAG). It provides the minimal common functionality that must be implemented as a prerequisite to all the Track-supporting profiles. The other Profiles extend Profile 0 with selected capabilities that this specification introduces on top.

Profile 1 (Storing Mode P-Route Segments along the Main DODAG) Profile 1 does not create new paths; compared to Profile 0, it combines Storing and Non-Storing Modes to balance the size of the Routing Header in the packet and the amount of state in the intermediate routers in a Non-Storing Mode RPL DODAG.

Profile 2 (Non-Storing Mode P-Route Segments along the Main DODAG) Profile 2 extends Profile 0 with Strict Source-Routing Non-Storing Mode P-Routes along the Main DODAG, which is the same as Profile 1 but using NSM VIOs as opposed to SM VIOs. Profile 2 provides the same capability to compress the SRH in packets down the Main DODAG as Profile 1, but it requires an encapsulation, in order to insert an additional SRH between the loose source routing hops. In that case, the Tracks MUST be installed as subTracks of the Main DODAG, the main RPL Instance MUST be used as TrackID, and the Ingress node that encapsulates is not the Root as it does not own the DODAGID.

Profile 3 In order to form the best path possible, those Profiles require the support of Sibling Information Option to inform the Root of additional possible hops. Profile 3 extends Profile 1 with additional Storing Mode P-Routes that install segments that do not follow the Main DODAG. If the Segment Ingress (in the SM-VIO) is the same as the IPv6 Address of the Track Ingress (in the projected DAO base Object), the P-DAO creates an implicit Track between the Segment Ingress and the Segment Egress.

Profile 4 Profile 4 extends Profile 2 with Strict Source-Routing Non-Storing Mode P-Routes to form East-West Tracks that are inside the Main DODAG but do not necessarily follow it. A Track is formed as one or more strict source source routed paths between the Root that is the Track Ingress, and the Track Egress that is the last node.

Profile 5 Profile 5 Combines Profile 4 with Profile 1 and enables to

loose source routing between the Ingress and the Egress of the Track. As in Profile 1, Storing Mode P-Routes connect the dots in the loose source route.

Profile 6 Profile 6 Combines Profile 4 with Profile 2 and also enables to loose source routing between the Ingress and the Egress of the Track.

Profile 7 Profile 7 implements profile 5 in a Main DODAG that is operated in Storing Mode as presented in Section 7.1. As in Profile 1 and 2, the TrackID is the RPLInstanceID of the Main DODAG. Longest match rules decide whether a packet is sent along the Main DODAG or rerouted in a track.

Profile 8 Profile 8 is offered in preparation of the RAW work, and for use cases where an arbitrary node in the network can afford the same code complexity as the RPL Root in a traditional deployment. It offers a full DODAG visibility to the Track Ingress as specified in Section 7.2 in a Non-Storing Mode Main DODAG.

Profile 9 Profile 9 combines profiles 7 and 8, operating the Track as a full DODAG within a Storing Mode Main DODAG, using only the Main DODAG RPLInstanceID as TrackID.

9. Backwards Compatibility

This specification can operate in a mixed network where some nodes support it and some do not. There are restrictions, though. All nodes that need to process a P-DAO MUST support this specification. As discussed in Section 3.7.1, how the root knows whether the nodes capabilities and whether it support this specification is out of scope.

This specification defines the 'D' flag in the RPL DODAG Configuration Option (see Section 4.1.7) to signal that the RPL nodes can request the creation of Tracks. The requester may not know whether the Track can effectively be constructed, and whether enough nodes along the preferred paths support this specification. Therefore it makes sense to only set the 'D' flags in DIO when the conditions of success are in place, in particular when all the nodes that could be on path of tracks are upgraded.

10. Security Considerations

It is worth noting that with [RPL], every node in the LLN is RPL-aware and can inject any RPL-based attack in the network. This draft uses messages that are already present in RPL [RPL] 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.

The LLN nodes depend on the 6LBR and the RPL participants for their operation. A trust model is necessary to ensure that the right devices are acting in these roles, so as to avoid threats such as black-holing, (see [RFC7416] section 7). This trust model could be at a minimum based on a Layer-2 Secure joining and the Link-Layer security. This is a generic 6LoWPAN requirement, see Req5.1 in Appendix B.5 of [RFC8505].

In a general manner, the Security Considerations in [RPL], and [RFC7416] apply to this specification as well. The Link-Layer security is needed in particular to prevent Denial-Of-Service attacks whereby a rogue router creates a high churn in the RPL network by constantly injected forged P-DAO messages and using up all the available storage in the attacked routers.

With this specification, only the Root may generate P-DAO messages. PDR messages may only be sent to the Root. This specification expects that the communication with the Root is authenticated but does enforce which method is used.

Additionally, the trust model could include a role validation (e.g., using a role-based authorization) to ensure that the node that claims to be a RPL Root is entitled to do so. That trust should propagate from Egress to Ingress in the case of a Storing Mode P-DAO.

This specification suggests some validation of the VIO to prevent basic loops by avoiding that a node appears twice. But that is only a minimal protection. Arguably, an attacker that can inject P-DAOs can reroute any traffic and deplete critical resources such as spectrum and battery in the LLN rapidly.

11. IANA Considerations

11.1. RPL DODAG Configuration Option Flag

IANA is requested to assign a flag from the "DODAG Configuration Option Flags for MOP 0..6" [RFC9010] registry as follows:

Bit Number	Capability Description	Reference
0 (suggested)	Projected Routes Support (D)	THIS RFC

Table 21: New DODAG Configuration Option Flag

IANA is requested to add [THIS RFC] as a reference for MOP 7 in the RPL Mode of Operation registry.

11.2. Elective 6LoWPAN Routing Header Type

THIS RFC updates the IANA registry titled "Elective 6LoWPAN Routing Header Type" that was created for [RFC8138] and assigns the following value:

Value	Description	Reference
8 (Suggested)	P-RPI-6LoRH	THIS RFC

Table 22: New Elective 6LoWPAN Routing Header Type

11.3. Critical 6LoWPAN Routing Header Type

THIS RFC updates the IANA registry titled "Critical 6LoWPAN Routing Header Type" that was created for [RFC8138] and assigns the following value:

Value	Description	Reference
8 (Suggested)	P-RPI-6LoRH	THIS RFC

Table 23: New Critical 6LoWPAN Routing Header Type

11.4. Subregistry For The RPL Option Flags

IANA is required to create a subregistry for the 8-bit RPL Option Flags field, as detailed in Figure 11, under the "Routing Protocol for Low Power and Lossy Networks (RPL)" registry. The bits are indexed from 0 (leftmost) to 7. Each bit is Tracked with the following qualities:

- * Bit number (counting from bit 0 as the most significant bit)
- * Indication When Set
- * Reference

Registration procedure is "Standards Action" [RFC8126]. The initial allocation is as indicated in Table 27:

Bit number	Indication When Set	Reference
0	Down 'O'	[RFC6553]
1	Rank-Error (R)	[RFC6553]
2	Forwarding-Error (F)	[RFC6553]
3 (Suggested)	Projected-Route (P)	THIS RFC

Table 24: Initial PDR Flags

11.5. RPL Control Codes

THIS RFC extends the IANA Subregistry created by RFC 6550 for RPL Control Codes as indicated in Table 25:

Code	Description	Reference
0x09 (Suggested)	Projected DAO Request (PDR)	THIS RFC
0x0A (Suggested)	PDR-ACK	THIS RFC

Table 25: New RPL Control Codes

11.6. RPL Control Message Options

THIS RFC extends the IANA Subregistry created by RFC 6550 for RPL Control Message Options as indicated in Table 26:

Value	Meaning	Reference
0x0E (Suggested)	Stateful VIO (SM-VIO)	THIS RFC
0x0F (Suggested)	Source-Routed VIO (NSM-VIO)	THIS RFC
0x10 (Suggested)	Sibling Information option	THIS RFC

Table 26: RPL Control Message Options

11.7. SubRegistry for the Projected DAO Request Flags

IANA is required to create a registry for the 8-bit Projected DAO Request (PDR) Flags field. Each bit is Tracked with the following qualities:

- * Bit number (counting from bit 0 as the most significant bit)
- * Capability description
- * Reference

Registration procedure is "Standards Action" [RFC8126]. The initial allocation is as indicated in Table 27:

Bit number	Capability description	Reference
0	PDR-ACK request (K)	THIS RFC
1	Requested path should be redundant (R)	THIS RFC

Table 27: Initial PDR Flags

11.8. SubRegistry for the PDR-ACK Flags

IANA is required to create an subregistry for the 8-bit PDR-ACK Flags field. Each bit is Tracked with the following qualities:

- * Bit number (counting from bit 0 as the most significant bit)
- * Capability description
- * Reference

Registration procedure is "Standards Action" [RFC8126]. No bit is currently defined for the PDR-ACK Flags.

11.9. Subregistry for the PDR-ACK Acceptance Status Values

IANA is requested to create a Subregistry for the PDR-ACK Acceptance Status values.

- * Possible values are 6-bit unsigned integers (0..63).
- * Registration procedure is "Standards Action" [RFC8126].
- * Initial allocation is as indicated in Table 28:

Value	Meaning	Reference
0	Unqualified Acceptance	THIS RFC

Table 28: Acceptance values of the PDR-ACK Status

11.10. Subregistry for the PDR-ACK Rejection Status Values

IANA is requested to create a Subregistry for the PDR-ACK Rejection Status values.

- * Possible values are 6-bit unsigned integers (0..63).
- * Registration procedure is "Standards Action" [RFC8126].
- * Initial allocation is as indicated in Table 29:

Value	Meaning	Reference
0	Unqualified Rejection	THIS RFC
1	Transient Failure	THIS RFC

Table 29: Rejection values of the PDR-ACK Status

11.11. SubRegistry for the Via Information Options Flags

IANA is requested to create a Subregistry for the 5-bit Via Information Options (Via Information Option) Flags field. Each bit is Tracked with the following qualities:

- * Bit number (counting from bit 0 as the most significant bit)
- * Capability description
- * Reference

Registration procedure is "Standards Action" [RFC8126]. No bit is currently defined for the Via Information Options (Via Information Option) Flags.

11.12. SubRegistry for the Sibling Information Option Flags

IANA is required to create a registry for the 5-bit Sibling Information Option (SIO) Flags field. Each bit is Tracked with the following qualities:

- * Bit number (counting from bit 0 as the most significant bit)
- * Capability description
- * Reference

Registration procedure is "Standards Action" [RFC8126]. The initial allocation is as indicated in Table 30:

Bit number	Capability description	Reference
0 (Suggested)	"S" flag: Sibling in same DODAG as Self	THIS RFC

Table 30: Initial SIO Flags

11.13. Destination Advertisement Object Flag

THIS RFC modifies the "Destination Advertisement Object (DAO) Flags" registry initially created in Section 20.11 of [RPL] .

Section 4.1.1 also defines one new entry in the Registry as follows:

Bit Number	Capability Description	Reference
2 (Suggested)	Projected DAO (P)	THIS RFC

Table 31: New Destination Advertisement Object (DAO) Flag

11.14. Destination Advertisement Object Acknowledgment Flag

THIS RFC modifies the "Destination Advertisement Object (DAO) Acknowledgment Flags" registry initially created in Section 20.12 of [RPL] .

Section 4.1.2 also defines one new entry in the Registry as follows:

Bit Number	Capability Description	Reference
1 (Suggested)	Projected DAO-ACK (P)	THIS RFC

Table 32: New Destination Advertisement Object Acknowledgment Flag

11.15. New ICMPv6 Error Code

In some cases RPL will return an ICMPv6 error message when a message cannot be forwarded along a P-Route.

IANA has defined an ICMPv6 "Code" Fields Registry for ICMPv6 Message Types. ICMPv6 Message Type 1 describes "destination Unreachable" codes. This specification requires that a new code is allocated from the ICMPv6 Code Fields Registry for ICMPv6 Message Type 1, for "Error in P-Route", with a suggested code value of 8, to be confirmed by IANA.

11.16. RPL Rejection Status values

This specification updates the Subregistry for the "RPL Rejection Status" values under the RPL registry, as follows:

Value	Meaning	Reference
2 (Suggested)	Out of Resources	THIS RFC
3 (Suggested)	Error in VIO	THIS RFC
4 (Suggested)	Predecessor Unreachable	THIS RFC
5 (Suggested)	Unreachable Target	THIS RFC
6..63	Unassigned	

Table 33: Rejection values of the RPL Status

12. Acknowledgments

The authors wish to acknowledge JP Vasseur, Remy Liubing, James Pylakutty, and Patrick Wetterwald for their contributions to the ideas developed here. Many thanks to Dominique Barthel and SVR Anand for their global contribution to 6TiSCH, RAW and this RFC, as well as text suggestions that were incorporated. Also special thanks Li Zhao and Toerless Eckert for their in-depth reviews, with many excellent suggestions that improved the readability and well as the content of the specification. Many thanks to Remous-Aris Koutsiamanis for his review during WGLC.

13. Normative References

[INT-ARCHI]

Braden, R., Ed., "Requirements for Internet Hosts - Communication Layers", STD 3, RFC 1122, DOI 10.17487/RFC1122, October 1989, <<https://www.rfc-editor.org/info/rfc1122>>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.

[RFC4443] Conta, A., Deering, S., and M. Gupta, Ed., "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", STD 89, RFC 4443, DOI 10.17487/RFC4443, March 2006, <<https://www.rfc-editor.org/info/rfc4443>>.

- [RFC4655] Farrel, A., Vasseur, J.-P., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, DOI 10.17487/RFC4655, August 2006, <<https://www.rfc-editor.org/info/rfc4655>>.
- [RFC6282] Hui, J., Ed. and P. Thubert, "Compression Format for IPv6 Datagrams over IEEE 802.15.4-Based Networks", RFC 6282, DOI 10.17487/RFC6282, September 2011, <<https://www.rfc-editor.org/info/rfc6282>>.
- [RPL] 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, <<https://www.rfc-editor.org/info/rfc6550>>.
- [RFC6553] Hui, J. and JP. Vasseur, "The Routing Protocol for Low-Power and Lossy Networks (RPL) Option for Carrying RPL Information in Data-Plane Datagrams", RFC 6553, DOI 10.17487/RFC6553, March 2012, <<https://www.rfc-editor.org/info/rfc6553>>.
- [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, <<https://www.rfc-editor.org/info/rfc6554>>.
- [RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <<https://www.rfc-editor.org/info/rfc8126>>.
- [RFC8138] Thubert, P., Ed., Bormann, C., Toutain, L., and R. Cragie, "IPv6 over Low-Power Wireless Personal Area Network (6LoWPAN) Routing Header", RFC 8138, DOI 10.17487/RFC8138, April 2017, <<https://www.rfc-editor.org/info/rfc8138>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.
- [RFC8754] Filsfils, C., Ed., Dukes, D., Ed., Previdi, S., Leddy, J., Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header (SRH)", RFC 8754, DOI 10.17487/RFC8754, March 2020, <<https://www.rfc-editor.org/info/rfc8754>>.

- [RFC9008] Robles, M.I., Richardson, M., and P. Thubert, "Using RPI Option Type, Routing Header for Source Routes, and IPv6-in-IPv6 Encapsulation in the RPL Data Plane", RFC 9008, DOI 10.17487/RFC9008, April 2021, <<https://www.rfc-editor.org/info/rfc9008>>.

14. Informative References

- [6LoWPAN] Montenegro, G., Kushalnagar, N., Hui, J., and D. Culler, "Transmission of IPv6 Packets over IEEE 802.15.4 Networks", RFC 4944, DOI 10.17487/RFC4944, September 2007, <<https://www.rfc-editor.org/info/rfc4944>>.
- [RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, DOI 10.17487/RFC5440, March 2009, <<https://www.rfc-editor.org/info/rfc5440>>.
- [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, <<https://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, <<https://www.rfc-editor.org/info/rfc7102>>.
- [RFC7416] Tsao, T., Alexander, R., Dohler, M., Daza, V., Lozano, A., and M. Richardson, Ed., "A Security Threat Analysis for the Routing Protocol for Low-Power and Lossy Networks (RPLs)", RFC 7416, DOI 10.17487/RFC7416, January 2015, <<https://www.rfc-editor.org/info/rfc7416>>.
- [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, <<https://www.rfc-editor.org/info/rfc8025>>.
- [RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402, July 2018, <<https://www.rfc-editor.org/info/rfc8402>>.

- [RFC8505] Thubert, P., Ed., Nordmark, E., Chakrabarti, S., and C. Perkins, "Registration Extensions for IPv6 over Low-Power Wireless Personal Area Network (6LoWPAN) Neighbor Discovery", RFC 8505, DOI 10.17487/RFC8505, November 2018, <<https://www.rfc-editor.org/info/rfc8505>>.
- [RFC8655] Finn, N., Thubert, P., Varga, B., and J. Farkas, "Deterministic Networking Architecture", RFC 8655, DOI 10.17487/RFC8655, October 2019, <<https://www.rfc-editor.org/info/rfc8655>>.
- [RFC8930] Watteyne, T., Ed., Thubert, P., Ed., and C. Bormann, "On Forwarding 6LoWPAN Fragments over a Multi-Hop IPv6 Network", RFC 8930, DOI 10.17487/RFC8930, November 2020, <<https://www.rfc-editor.org/info/rfc8930>>.
- [RFC8931] Thubert, P., Ed., "IPv6 over Low-Power Wireless Personal Area Network (6LoWPAN) Selective Fragment Recovery", RFC 8931, DOI 10.17487/RFC8931, November 2020, <<https://www.rfc-editor.org/info/rfc8931>>.
- [RFC8994] Eckert, T., Ed., Behringer, M., Ed., and S. Bjarnason, "An Autonomic Control Plane (ACP)", RFC 8994, DOI 10.17487/RFC8994, May 2021, <<https://www.rfc-editor.org/info/rfc8994>>.
- [RFC9010] Thubert, P., Ed. and M. Richardson, "Routing for RPL (Routing Protocol for Low-Power and Lossy Networks) Leaves", RFC 9010, DOI 10.17487/RFC9010, April 2021, <<https://www.rfc-editor.org/info/rfc9010>>.
- [RFC9030] Thubert, P., Ed., "An Architecture for IPv6 over the Time-Slotted Channel Hopping Mode of IEEE 802.15.4 (6TiSCH)", RFC 9030, DOI 10.17487/RFC9030, May 2021, <<https://www.rfc-editor.org/info/rfc9030>>.
- [RFC9035] Thubert, P., Ed. and L. Zhao, "A Routing Protocol for Low-Power and Lossy Networks (RPL) Destination-Oriented Directed Acyclic Graph (DODAG) Configuration Option for the 6LoWPAN Routing Header", RFC 9035, DOI 10.17487/RFC9035, April 2021, <<https://www.rfc-editor.org/info/rfc9035>>.

[RAW-ARCHI]

Thubert, P. and G. Z. Papadopoulos, "Reliable and Available Wireless Architecture", Work in Progress, Internet-Draft, draft-ietf-raw-architecture-04, 4 March 2022, <<https://datatracker.ietf.org/doc/html/draft-ietf-raw-architecture-04>>.

[USE-CASES]

Bernardos, C. J., Papadopoulos, G. Z., Thubert, P., and F. Theoleyre, "RAW use-cases", Work in Progress, Internet-Draft, draft-ietf-raw-use-cases-05, 23 February 2022, <<https://datatracker.ietf.org/doc/html/draft-ietf-raw-use-cases-05>>.

[I-D.kuehlewind-update-tag]

Kuehlewind, M. and S. Krishnan, "Definition of new tags for relations between RFCs", Work in Progress, Internet-Draft, draft-kuehlewind-update-tag-04, 12 July 2021, <<https://datatracker.ietf.org/doc/html/draft-kuehlewind-update-tag-04>>.

[I-D.irtf-panrg-path-properties]

Enghardt, T. and C. Krähenbühl, "A Vocabulary of Path Properties", Work in Progress, Internet-Draft, draft-irtf-panrg-path-properties-05, 7 March 2022, <<https://datatracker.ietf.org/doc/html/draft-irtf-panrg-path-properties-05>>.

[PCE]

IETF, "Path Computation Element", <<https://dataTracker.ietf.org/doc/charter-ietf-pce/>>.

Authors' Addresses

Pascal Thubert (editor)
Cisco Systems, Inc
Building D
45 Allee des Ormes - BP1200
06254 Mougins - Sophia Antipolis
France
Phone: +33 497 23 26 34
Email: pthubert@cisco.com

Rahul Arvind Jadhav
Huawei Tech
Kundalahalli Village, Whitefield,
Bangalore 560037
Karnataka
India
Phone: +91-080-49160700
Email: rahul.ietf@gmail.com

Michael C. Richardson
Sandelman Software Works
Email: mcr+ietf@sandelman.ca
URI: <http://www.sandelman.ca/>

ROLL
Internet-Draft
Intended status: Standards Track
Expires: 13 May 2022

K. Iwanicki
University of Warsaw
9 November 2021

RNFD: Fast border router crash detection in RPL
draft-iwanicki-roll-rnfd-01

Abstract

By and large, a correct operation of a RPL network requires border routers to be up. In many applications, it is beneficial for the nodes to detect a crash of a border router as soon as possible to trigger fallback actions. This document describes RNFD, an extension to RPL that expedites border router failure detection, even by an order of magnitude, by having nodes collaboratively monitor the status of a given border router. The extension introduces an additional state at each node, a new type of RPL Control Message Options for synchronizing this state among different nodes, and the coordination algorithm itself.

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 <https://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 13 May 2022.

Copyright Notice

Copyright (c) 2021 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 (<https://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 Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction	2
1.1. Effects of LBR Crashes	3
1.2. Design Principles	4
2. Terminology	5
3. Overview	6
3.1. Protocol State Machine	7
3.2. Counters and Communication	8
4. The RNFD Option	9
4.1. General CFRC Requirements	9
4.2. Format of the Option	10
5. RPL Router Behavior	12
5.1. Joining a DODAG Version and Changing the RNFD Role . . .	12
5.2. Detecting and Verifying Problems with the DODAG Root . .	13
5.3. Disseminating Observations and Reaching Agreement . . .	15
5.4. DODAG Root's Behavior	16
5.5. Activating and Deactivating the Protocol on Demand . . .	17
5.6. Processing CFRCs of Incompatible Lengths	17
5.7. Summary of RNFD's Interactions with RPL	18
5.8. Summary of RNFD's Constants	19
6. Manageability Considerations	19
6.1. Role Assignment and CFRC Size Adjustment	20
6.2. Virtual DODAG Roots	20
7. Security Considerations	21
8. IANA Considerations	22
9. Acknowledgements	22
10. References	22
10.1. Normative References	22
10.2. Informative References	23
Author's Address	24

1. Introduction

RPL is an IPv6 routing protocol for low-power and lossy networks (LLNs) [RFC6550]. Such networks are usually constrained in device energy and channel capacity. They are formed largely of nodes that offer little processing power and memory, and links that are of variable qualities and support low data rates. Therefore, the main challenge that a routing protocol for LLNs has to address is minimizing resource consumption without sacrificing reaction time to network changes.

One of the main design principles adopted in RPL to minimize node resource consumption is delegating much of the responsibility for routing to LLN border routers (LBRs). A network is organized into destination-oriented directed acyclic graphs (DODAGs), each corresponding to an LBR and having all its paths terminate at the LBR. To this end, every node is dynamically assigned a rank representing its distance, measured in some metric, to a given LBR, with the LBR having the minimal rank, which reflects its role as the DODAG root. The ranks allow each non-LBR node to select from among its neighbors (i.e., nodes to which the node has links) those ones that are closer to the LBR than the node itself: the node's parents in the graph. The resulting DODAG paths, consisting of the node-parent links, are utilized for routing packets upward: to the LBR and outside the LLN. They are also used by nodes to periodically report their connectivity upward to the LBR, which allows in turn for directing packets downward, from the LBR to these nodes, for instance, by means of source routing [RFC6554]. All in all, not only do LBRs participate in routing but also drive the process of DODAG construction and maintenance underlying the protocol.

To play this central role, LBRs are expected to be more capable than regular LLN nodes. They are assumed not to be constrained in computing power, memory, and energy, which often entails a more involved hardware-software architecture and tethered power supply. This, however, also makes them more prone to failures, especially since in large deployments it is often difficult to ensure a backup power supply for every LBR.

1.1. Effects of LBR Crashes

When an LBR crashes, the nodes in its DODAG lose the ability to communicate with other Internet hosts. In addition, a significant fraction of DODAG paths interconnecting the nodes become invalid, as they pass through the LBR. The others also degenerate as a result of DODAG repair attempts, which are bound to fail. In effect, routing inside the DODAG also becomes largely impossible. Consequently, it is desirable that an LBR crash be detected by the nodes fast, so that they can leave the broken DODAG and join another one or trigger additional application- or deployment-dependent fallback mechanisms, thereby minimizing the negative impact of the disconnection.

Since all DODAG paths lead to the corresponding LBR, detecting its crash by a node entails dropping all parents and adopting an infinite rank, which reflects the node's inability to reach the LBR. Depending on the deployment settings, the node can then remain in such a state, join a different DODAG, or even become itself the root of a floating DODAG. In any case, however, achieving this state for all nodes is slow, can generate heavy traffic, and is difficult to implement correctly [Iwanicki16] [Paszowska19] [Ciolkosz19].

To start with, tearing down all DODAG paths requires each of the LBR's neighbors to detect that its link with the LBR is no longer up. Otherwise, any of the neighbors unaware of this fact can keep advertising a finite rank and can thus be other nodes' parent or ancestor in the DODAG: such nodes will incorrectly believe they have a valid path to the LBR. Detecting a crash of a link by a node normally happens when the node has observed sufficiently many forwarding failures over the link. Therefore, considering the low-data-rate applications of LLNs, the period from the crash to the moment of eliminating from the DODAG the last link to the LBR may be long. Subsequently learning by all nodes that none of their links can form any path leading to the LBR also adds latency, partly due to parent changes that the nodes independently perform in attempts to repair their broken paths locally. Since a non-LBR node has only local knowledge of the network, potentially inconsistent with that of other nodes, such parent changes often produce paths containing loops, which have to be broken before all nodes can conclude that no path to the LBR exists globally. Even with RPL's dedicated loop detection mechanisms [RFC6553], this also requires traffic, and hence time. Finally, switching a parent or discovering a loop can also generate cascaded bursts of control traffic, owing to the adaptive Trickle algorithm for exchanging DODAG information [RFC6202]. Overall, the behavior of the network when handling an LBR crash is highly suboptimal, thereby not being in line with RPL's goal of minimizing resource consumption and reaction latencies.

1.2. Design Principles

To address this issue, this document proposes an extension to RPL, dubbed Root Node Failure Detector (RNFD). To minimize the time and traffic required to handle an LBR crash, the RNFD algorithm adopts the following design principles, derived directly from the previous observations:

1. Explicitly coordinating LBR monitoring between nodes instead of relying only on the emergent behavior resulting from their independent operation.

2. Avoiding probing all links to the dead LBR so as to reduce the tail latency when eliminating these links from the DODAG.
3. Exploiting concurrency by prompting proactive checking for a possible LBR crash when some nodes suspect such a failure may have taken place, which aims to further reduce the critical path.
4. Minimizing changes to RPL's existing algorithms by operating in parallel and largely independently (in the background), and introducing few additional assumptions.

While these principles do improve RPL's performance under a wide range of LBR crashes, their probabilistic nature precludes hard guarantees for all possible corner cases. In particular, in some scenarios, RNFD's operation may result in false negatives, but these situations are peculiar and will eventually be handled by RPL's own aforementioned mechanisms. Likewise, in some scenarios, notably involving highly unstable links, false positives may occur, but they can be alleviated as well. In any case, the principles also guarantee that RNFD can be deactivated at any time, if needed, in which case RPL's operation is unaffected.

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 BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

The Terminology used in this document is consistent with and incorporates that described in "Terms Used in Routing for Low-Power and Lossy Networks (LLNs)" [RFC7102], "RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks" [RFC6550], and "The Routing Protocol for Low-Power and Lossy Networks (RPL) Option for Carrying RPL Information in Data-Plane Datagrams" [RFC6553]. Other terms in use in LLNs can be found in "Terminology for Constrained-Node Networks" [RFC7228].

In particular, the following acronyms appear in the document:

DIO DODAG Information Object (a RPL message)

DIS DODAG Information Solicitation (a RPL message)

DODAG Destination-Oriented Directed Acyclic Graph

LLN Low-power and Lossy Network

LBR LLN Border Router

In addition, the document introduces the following concepts:

Sentinel One of the two roles that a node can play in RNFD. For a given DODAG Version, a Sentinel node is the DODAG root's neighbor that monitors the DODAG root's status. There are normally multiple Sentinels for a DODAG root. However, being the DODAG root's neighbor need not imply being Sentinel.

Acceptor The other of the two roles that a node can play in RNFD. For a given DODAG Version, an Acceptor node is a node that is not Sentinel.

Locally Observed DODAG Root's State (LORS) A node's local knowledge of the DODAG root's status, specifying in particular whether the DODAG root is up.

Conflict-Free Replicated Counter (CFRC) Conceptually represents a dynamic set whose cardinality can be estimated. It defines a partial order on its values and supports element addition and union. The union operation is order- and duplicate-insensitive, that is, idempotent, commutative, and associative.

3. Overview

As mentioned previously, LBRs are DODAG roots in RPL, and hence a crash of an LBR is global in that it affects all nodes in the corresponding DODAG. Therefore, each node running RNFD for a given DODAG explicitly tracks the DODAG root's current condition, which is referred to as Locally Observed DODAG Root's State (LORS), and synchronizes its local knowledge with other nodes.

Since monitoring the condition of the DODAG root is performed by tracking the status of its links (i.e., whether they are up or down), it must be done by the root's neighbors; other nodes must accept their observations. Consequently, depending on their roles, non-root nodes are divided in RNFD into two disjoint groups: Sentinels and Acceptors. A Sentinel node is the DODAG root's neighbor that monitors its link with the root. The DODAG root thus normally has multiple Sentinels but being its neighbor need not imply being Sentinel. An Acceptor node is in turn a node that is not Sentinel. Acceptors thus mainly collect and propagate Sentinels' observations. More information on Sentinel selection can be found in Section 6.1.

3.1. Protocol State Machine

The possible values of LORS and transitions between them are depicted in Figure 1. States "UP" and "GLOBALLY DOWN" can be attained by both Sentinels and Acceptors; states "SUSPECTED DOWN" and "LOCALLY DOWN"--by Sentinels only.

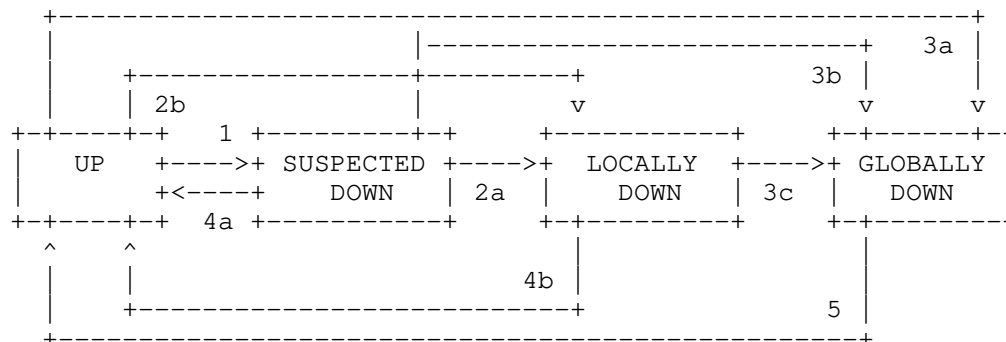


Figure 1: RNFD States and Transitions

To begin with, when any node joins a DODAG Version, the DODAG root must appear alive, so the node initializes RNFD with its LORS equal to "UP". For a properly working DODAG root, the node remains in state "UP".

However, when a node--acting as Sentinel--starts suspecting that the root may have crashed, it changes its LORS to "SUSPECTED DOWN" (transition 1 in Figure 1). The transition from "UP" to "SUSPECTED DOWN" can happen based on the node's observations at either the data plane, for instance, link-layer triggers about missing hop-by-hop acknowledgments for packets forwarded over the node's link to the root, or the control plane, for example, a significant growth in the number of Sentinels already suspecting the root to be dead. In state "SUSPECTED DOWN", the Sentinel node may verify its suspicion and/or inform other nodes about the suspicion. When this has been done, it changes its LORS to "LOCALLY DOWN" (transition 2a). In some cases, the verification need not be performed and, as an optimization, a direct transition from "UP" to "LOCALLY DOWN" (transition 2b) can be done instead.

If sufficiently many Sentinels have their LORS equal to "LOCALLY DOWN", all nodes--Sentinels and Acceptors--consent globally that the DODAG root must have crashed and set their LORS to "GLOBALLY DOWN", irrespective of the previous value (transitions 3a, 3b, and 3c). State "GLOBALLY DOWN" is terminal in that the only transition any node can perform from this to another state (transition 5) takes

place when the node joins a new DODAG version. When a node is in state "GLOBALLY DOWN", RNFD forces RPL to maintain an infinite rank and no parent, thereby preventing routing packets upward in the DODAG. In other words, this state represents a situation in which all non-root nodes agree that the current DODAG version is unusable, and hence, to recover, the root has to give a proof of being alive by initiating a new DODAG Version.

In contrast, if a node--either Sentinel or Acceptor--is in state "UP", RNFD does not influence RPL's packet forwarding: a node can route packets upward if it has a parent. The same is true for a Sentinel node in states "SUSPECTED DOWN" and "LOCALLY DOWN". Finally, while in any of the two states, a Sentinel node may observe some activity of the DODAG root, and hence decide that its suspicion is a mistake. In such a case, it returns to state "UP" (transitions 4a and 4b).

3.2. Counters and Communication

To enable arriving at a global conclusion that the DODAG root has crashed (i.e., transiting to state "GLOBALLY DOWN"), all nodes count locally and synchronize among each other the number of Sentinels considering the root to be dead (i.e., those in state "LOCALLY DOWN"). This process employs structures referred to as conflict-free replicated counters (CFRCs). They are stored and modified independently by each node and are disseminated throughout the network in options added to RPL link-local control messages: DODAG Information Objects (DIOs) and DODAG Information Solicitations (DISs). Upon reception of such an option from its neighbor, a node merges the received counter with its local one, thereby obtaining a new content for its local counter.

The merging operation is idempotent, commutative, and associative. Moreover, all possible counter values are partially ordered. This enables ensuring eventual consistency of the counters across all nodes, irrespective of the particular sequence of merges, shape of the DODAG, or general network topology.

Each node in RNFD maintains two CFRCs for a DODAG:

- * PositiveCFRC, counting Sentinels that have considered or still consider the root node as alive in the current DODAG Version,
- * NegativeCFRC, counting Sentinels that have considered or still consider the root node as dead in the current DODAG Version.

PositiveCFRC is always greater than or equal to the NegativeCFRC in terms of the partial order defined for the counters. The difference between the value of PositiveCFRC and the value of NegativeCFRC is thus nonnegative and estimates the number of Sentinels that still consider the DODAG root node as alive.

4. The RNFD Option

RNFD state synchronization between nodes takes place through the RNFD Option. It is a new type of RPL Control Message Options that is carried in link-local RPL control messages, notably DIOs and DISSs. Its main task is allowing the receivers to merge their two CFRCs with the sender's CFRCs.

4.1. General CFRC Requirements

CFRCs in RNFD MUST support the following operations:

value(c) Returns a non-negative integer value corresponding to the number of nodes counted by a given CFRC, c.

zero() Returns a CFRC that counts no nodes, that is, has its value equal to 0.

self() Returns a CFRC that counts only the node executing the operation.

infinity() Returns a CFRC that counts all possible nodes and represents a special value, infinity.

merge(c1, c2) Returns a CFRC that is a union of c1 and c2 (i.e., counts all nodes that are counted by either c1, c2, or both c1 and c2).

compare(c1, c2) Returns the result of comparing c1 to c2.

saturated(c) Returns TRUE if a given CFRC, c, is saturated (i.e., no more new nodes should be counted by it), or FALSE otherwise.

The partial ordering of CFRCs implies that the result of compare(c1, c2) can be either:

- * smaller, if c1 is ordered before c2 (i.e., c2 counts all nodes that c1 counts and at least one node that c1 does not count);
- * greater, if c1 is ordered after c2 (i.e., c1 counts all nodes that c2 counts and at least one node that c2 does not count);

```
*   equal, if c1 and c2 are the same (i.e., they count the same
    nodes);
```

* incomparable, otherwise.

In particular, `zero()` is smaller than all other values and `infinity()` is greater than any other value.

The properties of merging in turn can be formalized as follows for any c_1 , c_2 , and c_3 :

```
* idempotence: c1 = merge(c1, c1);
```

```
* commutativity: merge(c1, c2) = merge(c2, c1);
```

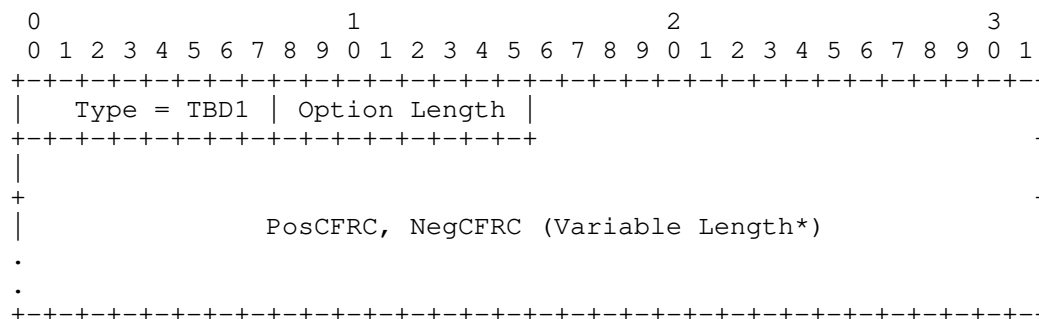
```
*  associativity: merge(c1, merge(c2, c3)) = merge(merge(c1, c2),
c3).
```

In particular, `merge(c, zero())` always equals `c` while `merge(c, infinity())` always equals `infinity()`.

There are many algorithmic structures that can provide the aforementioned properties of CFRC. Although in principle RNFD does not rely on any specific one, the option adopts so-called linear counting [Whang90].

4.2. Format of the Option

The format of the RNFD Option conforms to the generic format of RPL Control Message Options:



The '*' denotes that the fields have equal lengths of at least one octet each. In effect, together they occupy at least two octets.

Figure 2: Format of the RNFD Option

Option Type TBD1

Option Length 8-bit unsigned integer. Denotes the length of the option in octets excluding the Option Type and Option Length fields. Its value MUST be even. A value of 0 denotes that RNFD is disabled in the current DODAG Version.

PosCFRC, NegCFRC Two variable-length, octet-aligned bit arrays carrying the sender's PositiveCFRC and NegativeCFRC, respectively.

The length of the arrays constituting the PosCFRC and NegCFRC fields is the same and is derived from Option Length as follows. The value of Option Length is divided by 2 to obtain the number of octets each of the two arrays occupies. The resulting number of octets is multiplied by 8 which yields an upper bound on the number of bits in each array. As the actual bit length of each of the arrays, the largest prime number less than the upper bound is assumed. For example, if the value of Option Length is 16, then each array occupies 8 octets, and its actual bit length is 61, as this is the largest prime number less than 64.

Furthermore, for any bit equal to 1 in the NegCFRC, the bit with the same index MUST be equal to 1 also in the PosCFRC. Any unused bits (i.e., the bits beyond the actual bit length of each of the arrays) MUST be equal to 0. Finally, if PosCFRC has all its bits equal to 1, then NegCFRC MUST also have all its bits equal to 1.

The CFRC operations are defined for such bit arrays of a given length as follows:

value(c) Returns the smallest integer value not less than $-LT \cdot \ln(L0/LT)$, where $\ln()$ is the natural logarithm function, $L0$ is the number of bits equal to 0 in the array corresponding to c and LT is the bit length of the array.

zero() Returns an array with all bits equal to 0.

self() Returns an array with a single bit, selected uniformly at random, equal to 1.

infinity() Returns an array with all bits equal to 1.

merge(c1, c2) Returns a bit array that constitutes a bitwise OR of $c1$ and $c2$, that is, a bit in the resulting array is equal to 0 only if the same bit is equal to 0 in both $c1$ and $c2$.

compare(c1, c2) Returns:

- * equal if each bit of c1 is equal to the corresponding bit of c2;
- * less if c1 and c2 are not equal and, for each bit equal to 1 in c1, the corresponding bit in c2 is also equal to 1;
- * greater if c1 and c2 are not equal and, for each bit equal to 1 in c2, the corresponding bit in c1 is also equal to 1;
- * incomparable, otherwise.

saturated(c) Returns TRUE, if more than 63% of the bits in c are equal to 1, or FALSE, otherwise.

5. RPL Router Behavior

Although RNFD operates largely independently of RPL, it does need interact with RPL and the overall protocol stack. These interactions are described next and can be realized, for instance, by means of event triggers.

5.1. Joining a DODAG Version and Changing the RNFD Role

Whenever RPL running at a node joins a DODAG Version, RNFD--if active--MUST assume for the node the role of Acceptor. Accordingly, it MUST set its LORS to "UP" and its PositiveCFRC and NegativeCFRC to zero().

The role MAY then change between Acceptor and Sentinel at any time. However, while a switch from Sentinel to Acceptor has no preconditions, for a switch from Acceptor to Sentinel to be possible, all of the following conditions MUST hold:

1. LORS is "UP";
2. saturated(PositiveCFRC) is FALSE;
3. a neighbor entry for the DODAG root is present in RPL's DODAG parent set;
4. the neighbor is considered reachable via its link-local IPv6 address.

A role change also **REQUIRES** appropriate updates to LORS and CFRCs, so that the node is properly accounted for. More specifically, when changing its role from Acceptor to Sentinel, the node **MUST** add itself to its PositiveCFRC as follows. It **MUST** generate a new CFRC value, `selfc = self()`, and **MUST** replace its PositiveCFRC, denoted `oldpc`, with `newpc = merge(oldpc, selfc)`. In contrast, the effects of a switch from Sentinel to Acceptor vary depending on the node's value of LORS before the switch:

- * for "GLOBALLY DOWN", the node **MUST NOT** modify its LORS, PositiveCFRC, and NegativeCFRC;
- * for "LOCALLY DOWN", the node **MUST** set its LORS to "UP" but **MUST NOT** modify its PositiveCFRC and NegativeCFRC;
- * for "UP" and "SUSPECTED DOWN", the node **MUST** set its LORS to "UP", **MUST NOT** modify its PositiveCFRC, but **MUST** add itself to NegativeCFRC, that is, replace its NegativeCFRC, denoted `oldnc`, with `newnc = merge(oldnc, selfc)`, where `selfc` is the counter generated with `self()` when the node last added itself to its PositiveCFRC.

5.2. Detecting and Verifying Problems with the DODAG Root

Only nodes that are Sentinels take active part in detecting crashes of the DODAG Root; Acceptors just disseminate their observations, reflected in the CFRCs.

The DODAG root monitoring SHOULD be based on both internal inputs, notably the values of CFRCs and LORS, and external inputs, such as triggers from RPL and other protocols. External input monitoring SHOULD be performed preferably in a reactive fashion, also independently of RPL, and at both data plane and control plane. In particular, it is RECOMMENDED that RNFD be directly notified of events relevant to the routing adjacency maintenance mechanisms on which RPL relies, such as Layer 2 triggers [RFC5184] or the Neighbor Unreachability Detection [RFC4861] mechanism. Only events concerning the DODAG root need be monitored to this end. For example, RNFD can conclude that there may be problems with the DODAG root if it observes a lack of multiple consecutive L2 acknowledgments for packets transmitted by the node via the link to the DODAG root. Internally, in turn, it is RECOMMENDED that RNFD take action whenever there is a change to its local CFRCs, so that a node can have a chance to participate in detecting potential problems even when normally it would not exchange packets over the link with the DODAG root during some period. In particular, RNFD SHOULD conclude that there may be problems with the DODAG root, when the fraction $\text{value}(\text{NegativeCFRC}) / \text{value}(\text{PositiveCFRC})$ has grown by at least `RNFD_SUSPICION_GROWTH_THRESHOLD` since the node last set its LORS to "UP".

Whenever having its LORS set to "UP" RNFD concludes--based on either external or internal inputs--that there may be problems with the link with the DODAG root, it MUST set its LORS to either "SUSPECTED DOWN" or, as an optimization, to "LOCALLY DOWN".

The "SUSPECTED DOWN" value of LORS is temporary: its aim is to give RNFD an additional opportunity to verify whether the link with the DODAG root is indeed down. Depending on the outcome of such verification, RNFD MUST set its LORS to either "UP", if the link has been confirmed not to be down, or "LOCALLY DOWN", otherwise. The verification can be performed, for example, by transmitting RPL DIS or ICMPv6 Echo Request messages to the DODAG root's link-local IPv6 address and expecting replies confirming that the root is up and reachable through the link. Care SHOULD be taken not to overload the DODAG root with traffic due to simultaneous probes, for instance, random backoffs can be employed to this end. It is RECOMMENDED that the "SUSPECTED DOWN" value of LORS is attained and verification takes place if RNFD's conclusion on the state of the DODAG root is based only on indirect observations, for example, the aforementioned growth of the CFRC values. In contrast, for direct observations, such as missing L2 acknowledgments, the verification MAY be skipped, with the node's LORS effectively changing from "UP" directly to "LOCALLY DOWN".

For consistency with RPL, when detecting potential problems with the DODAG root, RNFD also MUST make use of RPL's independent knowledge. More specifically, a node MUST switch its LORS from "UP" or "SUSPECTED DOWN" directly to "LOCALLY DOWN" if a neighbor entry for the DODAG root is removed from RPL's DODAG parent set or the neighbor ceases to be considered reachable via its link-local IPv6 address.

Finally, while having its LORS already equal to "LOCALLY DOWN", a node may make an observation confirming that its link with the DODAG root is actually up. In such a case, it SHOULD set its LORS back to "UP" but MUST NOT do this before the previous conditions 2-4 necessary for a node to change its role from Acceptor to Sentinel all hold.

To appropriately account for the node's observations on the state of the DODAG root, the aforementioned LORS transitions are accompanied by changes to the node's local CFRCs as follows. Changes between "UP" and "SUSPECTED DOWN" do not affect any of the two CFRCs. During a switch from "UP" or "SUSPECTED DOWN" to "LOCALLY DOWN", in turn, the node MUST add itself to its NegativeCFRC, as explained previously. By symmetry, a transition from "LOCALLY DOWN" to "UP" REQUIRES the node to add itself to its PositiveCFRC, again, as explained previously.

5.3. Disseminating Observations and Reaching Agreement

Nodes disseminate their observations by exchanging CFRCs in the RNFD Options embedded in link-local RPL control messages, notably DIOs and DISs. When processing such a received option, a node--acting as Sentinel or Acceptor--MUST update its PositiveCFRC and NegativeCFRC to respectively $\text{newpc} = \text{merge}(\text{oldpc}, \text{recvpc})$ and $\text{newnc} = \text{merge}(\text{oldnc}, \text{recvnc})$, where oldpc and oldnc are the values of the node's PositiveCFRC and NegativeCFRC before the update, while recvpc and recvnc are the received values of option fields PosCFRC and NegCFRC, respectively.

In effect, the node's value of $\text{fraction}(\text{value}(\text{NegativeCFRC})/\text{value}(\text{PositiveCFRC}))$ may change. If the fraction reaches at least `RNFD_CONSENSUS_THRESHOLD` (with $\text{value}(\text{PositiveCFRC})$ being greater than zero), then the node consents on the DODAG root being down. Accordingly, it MUST change its LORS to "GLOBALLY DOWN" and set its PositiveCFRC and NegativeCFRC both to infinity().

The "GLOBALLY DOWN" value of LORS is terminal: the node MUST NOT change it and MUST NOT modify its CFRCs until it joins a new DODAG Version. With this value of LORS, RNFD at the node MUST also prevent RPL from having any DODAG parent and advertising any Rank other than `INFINITE_RANK`.

Since the RNFD Option is embedded, among others, in RPL DIO control messages, updates to a node's CFRCs may affect the sending schedule of these messages, which is driven by the DIO Trickle timer [RFC6206]. It is RECOMMENDED to use for RNFD a dedicated Trickle timer, different from RPL's DIO Trickle timer. In such a setting, whenever RNFD's timer fires and no DIO message containing the RNFD Option has been sent to the link-local all-RPL-nodes multicast IPv6 address since the previous firing, the node sends a DIO message containing the RNFD Option to the address. In contrast, in the absence of a dedicated Trickle timer for RNFD, an implementation SHOULD ensure that the RNFD Option is present in multicast DIO messages sufficiently often to quickly propagate changes to the node's CFRCs. In either case, a node MUST reset its Trickle timer when it changes its LORS to "GLOBALLY DOWN", so that information about the detected crash of the DODAG root is disseminated in the DODAG fast. Likewise, a node SHOULD reset its Trickle timer when any of its local CFRCs changes significantly.

5.4. DODAG Root's Behavior

The DODAG root node MUST assume the role of Acceptor in RNFD and MUST NOT ever switch this role. It MUST also monitor its LORS and local CFRCs, so that it can react to various events.

To start with, the DODAG root MUST generate a new DODAG Version, thereby restarting the protocol, if it changes its LORS to "GLOBALLY DOWN", which may happen when the root has restarted after a crash or the nodes have falsely detected its crash. It MAY also generate a new DODAG Version if $\text{fraction value(NegativeCFRC)/value(PositiveCFRC)}$ approaches `RNFD_CONSENSUS_THRESHOLD`, so as to avoid potential interruptions to routing.

Furthermore, the DODAG root SHOULD either generate a new DODAG Version or increase the bit length of its CFRCs if `saturated(PositiveCFRC)` becomes TRUE. This is a self-regulation mechanism that helps adjust the CFRCs to a potentially large number of Sentinels (see Section 6.1).

In general, issuing a new DODAG Version effectively restarts RNFD. The DODAG root MAY thus perform this operation also in other situations.

5.5. Activating and Deactivating the Protocol on Demand

RNFD CAN be activated and deactivated on demand, once per DODAG Version. The particular policies for activating and deactivating the protocol are outside the scope of this document. However, the activation and deactivation SHOULD be done at the DODAG root node; other nodes MUST comply.

More specifically, when a non-root node joins a DODAG Version, RNFD at the node is initially inactive. The node MUST NOT activate the protocol unless it receives for this DODAG Version a valid RNFD Option containing some CFRCs, that is, having its Option Length field positive. In particular, if the option accompanies the message that causes the node to join the DODAG Version, the protocol SHOULD be active from the moment of the joining. RNFD then remains active at the node until it is explicitly deactivated or the node joins a new DODAG Version. An explicit deactivation MUST take place when the node receives an RNFD Option for the DODAG Version with no CFRCs, that is, having its Option Length field equal to zero. When explicitly deactivated, RNFD MUST NOT be reactivated unless the node joins a new DODAG Version. In particular, when the first RNFD Option received by the node has its Option Length field equal to zero, the protocol MUST remain deactivated for the entire time the node belongs to the current DODAG Version.

When RNFD at a node is initially inactive for a DODAG Version, the node MUST NOT attach any RNFD Option to the messages it sends (in particular, because it may not know the desired CFRC length--see Section 5.6). When the protocol has been explicitly deactivated, the node MAY also decide not to attach the option to its outgoing messages. However, it is RECOMMENDED that it sends sufficiently many messages with the option to the link-local all-RPL-nodes multicast IPv6 address to allow its neighbors to learn that RNFD has been deactivated in the current DODAG version. In particular, it MAY reset its Trickle timer to this end but also MAY use some reactive mechanisms, for example, replying with a unicast DIO or DIS containing the RNFD Option with no CFRCs to a message from a neighbor that contains the option with some CFRCs, as such a neighbor appears not to have learned about the deactivation of RNFD.

5.6. Processing CFRCs of Incompatible Lengths

The merge() and compare() operations on CFRCs require both arguments to be compatible, that is, to have the same bit length. However, the processing rules for the RNFD Option (see Section 4.2) do not necessitate this. This fact is made use of not only in the mechanisms for activating and deactivating the protocol (see Section 5.5), but also in mechanisms for dynamic adjustments of

CFRCs, which aim to enable deployment-specific policies (see Section 6.1). A node thus MUST be prepared to receive the RNFD Option with fields PosCFRC and NegCFRC of a different bit length than the node's own PositiveCFRC and NegativeCFRC. Assuming that it has RNFD active and that fields PosCFRC and NegCFRC in the option have a positive length, the node MUST react as follows.

If the bit length of fields PosCFRC and NegCFRC is the same as that of the node's local PositiveCFRC and NegativeCFRC, then the node MUST perform the merges, as detailed previously (see Section 5.3).

If the bit length of fields PosCFRC and NegCFRC is smaller than that of the node's local PositiveCFRC and NegativeCFRC, then the node MUST ignore the option and MAY reset its Trickle timer.

If the bit length of fields PosCFRC and NegCFRC is greater than that of the node's local PositiveCFRC and NegativeCFRC, then the node MUST extend the bit length of its local CFRCs to be equal to that in the option and set the CFRCs as follows:

- * If the node's LORS is "GLOBALLY DOWN", then both its local CFRCs MUST be set to infinity().
- * Otherwise, they both MUST be set to zero(), and the node MUST account for itself in so initialized CFRCs. More specifically, if the node is Sentinel, then it MUST add itself to its PositiveCFRC, as detailed previously. In addition, if its LORS is "LOCALLY DOWN", then it MUST also add itself to its NegativeCFRC, again, as explained previously. Finally, the node MUST perform merges of its local CFRCs and the ones received in the option (see Section 5.3) and MAY reset its Trickle timer.

In contrast, if the node is unable to extend its local CFRCs, for example, because it lacks resources, then it MUST stop participating in RNFD, that is, until it joins a new DODAG Version, it MUST NOT send the RNFD Option and MUST ignore this option in received messages.

5.7. Summary of RNFD's Interactions with RPL

In summary, RNFD interacts with RPL in the following manner:

- * While having its LORS equal to "GLOBALLY DOWN", RNFD prevents RPL from routing packets and advertising upward routes in the corresponding DODAG (see Section 5.3).
- * In some scenarios, RNFD triggers RPL to issue a new DODAG Version (see Section 5.4).

- * Depending on the implementation, RNFD may cause RPL's DIO Trickle timer resets (see Section 5.3, Section 5.5, and Section 5.6).
- * RNFD monitors events relevant to routing adjacency maintenance as well as those affecting RPL's DODAG parent set (see Section 5.1 and Section 5.2).
- * Using RNFD entails embedding the RNFD Option into link-local RPL control messages (see Section 4.2).

5.8. Summary of RNFD's Constants

The following is a summary of RNFD's constants:

RNFD_SUSPICION_GROWTH_THRESHOLD A threshold concerning the value of fraction value(NegativeCFRC)/value(PositiveCFRC). If the value at a Sentinel node grows at least by this threshold since the time the node's LORS was last set to "UP", then the node's LORS is set to "SUSPECTED DOWN" or "LOCALLY DOWN", which implies that the node suspects or assumes a crash of the DODAG root (see Section 5.2). The default value of the threshold is 0.12. The higher the value the longer the detection period but the lower risk of increased traffic due suspicion verification.

RNFD_CONSENSUS_THRESHOLD A threshold concerning the value of fraction value(NegativeCFRC)/value(PositiveCFRC). If the value at a Sentinel or Acceptor node reaches the threshold, then the node's LORS is set to "GLOBALLY DOWN", which implies that consensus has been reached on the DODAG root node being down (see Section 5.3). The default value of the threshold is 0.51. The higher the value the longer the detection period but the lower the risk of false positives.

The means of configuring the constants at individual nodes are outside the scope of this document.

6. Manageability Considerations

RNFD is largely self-managed, with the exception of protocol activation and deactivation, as well as node role assignment and the related CFRC size adjustment, for which only the aforementioned mechanisms are defined, so as to enable adopting deployment-specific policies. This section outlines some of the possible policies.

6.1. Role Assignment and CFRC Size Adjustment

One approach to node role and CFRC size selection is to manually designate specific nodes as Sentinels in RNFD, assuming that they will have chances to satisfy the necessary conditions for attaining this role (see Section 5.1), and fixing the CFRC bit length to accommodate these nodes.

Another approach is to automate the selection process: in principle, any node satisfying the necessary conditions for becoming Sentinel (see Section 5.1) can attain this role. However, in networks where the DODAG root node has many neighbors, this approach may lead to saturated(PositiveCFRC) quickly becoming TRUE, which--without additional measures--may degrade RNFD's performance. This issue can be handled with a probabilistic solution: if PositiveCFRC becomes saturated with little or no increase in NegativeCFRC, then a new DODAG Version can be issued and a node satisfying the necessary conditions can become Sentinel in this version only with probability 1/2. This process can be continued with the probability being halved in each new DODAG Version until PositiveCFRC is no longer quickly saturated. Another solution is to increase, potentially multiple times the bit length of the CFRCs by the DODAG root if PositiveCFRC becomes saturated with little or no growth in NegativeCFRC, which does not require issuing a new DODAG Version but lengthens the RNFD Option. In this way, again, a sufficient bit length can be dynamically discovered or the root can conclude that a given bit length is excessive for (some) nodes and resort to the previous solution. Increasing the bit length can be done, for instance, by doubling it, respecting the condition that it has to be a prime number (see Section 4.2).

In either of the solutions, Sentinel nodes SHOULD preferably be stable themselves and have stable links to the DODAG root. Otherwise, they may often exhibit LORS transitions between "UP" and "LOCALLY DOWN" or switches between Acceptor and Sentinel roles, which gradually saturates CFRCs. Although as a mitigation the number of such transitions and switches per node MAY be limited, having Sentinels stable SHOULD be preferred.

6.2. Virtual DODAG Roots

RPL allows a DODAG to have a so-called virtual root, that is, a collection of nodes coordinating to act as a single root of the DODAG. The details of the coordination process are left open in the specification [RFC6550] but, from RNFD's perspective, two possible realizations are worth consideration:

- * Just a single (primary) node of the nodes comprising the virtual root acts as the actual root of the DODAG. Only when this node fails, does another (backup) node take over. As a result, at any time, at most one of the nodes comprising the virtual root is the actual root.
- * More than one of the nodes comprising the virtual root act as actual roots of the DODAG, all advertising the same Rank in the DODAG. When some of the nodes fail, the other nodes may or may not react in any specific way. In other words, at any time, more than one node can be the actual root.

In the first realization, RNFD's operation is largely unaffected. The necessary conditions for a node to become Sentinel (Section 5.1) guarantee that only the current primary root node is monitored by the protocol. This SHOULD be taken into account in the policies for node role assignment, CFRC size selection, and, possibly, the setting of the two thresholds (Section 5.8). Moreover, when a new primary has been elected, to avoid polluting CFRCs with observations on the previous primary, it is RECOMMENDED to issue a new DODAG Version, especially if the new primary has different neighbors compared to the old one.

In the second realization, the fact that the virtual root consists of multiple nodes is transparent to RNFD. Therefore, employing RNFD in such a setting can be beneficial only if the nodes comprising the virtual root may suffer from correlated crashes, for instance, due to global power outages.

7. Security Considerations

RNFD is an extension to RPL and is thus both vulnerable to and benefits from the security issues and solutions described in [RFC6550] and [RFC7416]. Its specification in this document does not introduce new traffic patterns or new messages, for which specific mitigation techniques would be required beyond what can already be adopted for RPL.

In particular, RNFD depends on information exchanged in the RNFD Option. If the contents of this option were compromised, then failure misdetection may occur. One possibility is that the DODAG root may be falsely detected as crashed, which would result in an inability of the nodes to route packets, at least until a new DODAG Version is issued by the root. Another possibility is that a crash of the DODAG root may not be detected by RNFD, in which case RPL would have to rely on its own mechanisms. Moreover, compromising the contents of the RNFD Option may also lead to increased traffic due to DIO Trickle timer resets. Consequently, RNFD deployments are RECOMMENDED to use RPL security mechanisms if there is a risk that control information might be modified or spoofed.

In this context, RNFD's two features are worth highlighting. First, unless a DODAG root's all neighbors are compromised, a false positive can always be detected by the root based on its local CFRCs. If the frequency of such false positives becomes problematic, RNFD can be disabled altogether, for instance, until the problem has been diagnosed. This procedure can be largely automated at LBRs. Second, some types of false negatives can also be detected this way. Those that pass undetected, in turn, are likely not to have major negative consequences on RPL apart from the lack of improvement to its performance upon a DODAG root's crash, at least if RPL's other components are not attacked as well.

8. IANA Considerations

To represent the RNFD Option, IANA is requested to allocate the value TBD1 from the "RPL Control Message Options" sub-registry (<https://www.iana.org/assignments/rpl/rpl.xhtml#control-message-options>) of the "Routing Protocol for Low Power and Lossy Networks (RPL)" registry.

9. Acknowledgements

The authors would like to acknowledge Piotr Ciolkosz and Agnieszka Paszkowska. Agnieszka contributed to deeper understanding and formally proving various aspects of RPL's behavior upon an LBR crash. Piotr in turn developed a prototype implementation of RNFD dedicated for RPL to verify earlier performance claims.

TODO More likely to follow.

10. References

10.1. 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, <<https://www.rfc-editor.org/info/rfc2119>>.
- [RFC6206] Levis, P., Clausen, T., Hui, J., Gnawali, O., and J. Ko, "The Trickle Algorithm", RFC 6206, DOI 10.17487/RFC6206, March 2011, <<https://www.rfc-editor.org/info/rfc6206>>.
- [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, <<https://www.rfc-editor.org/info/rfc6550>>.
- [RFC6553] Hui, J. and JP. Vasseur, "The Routing Protocol for Low-Power and Lossy Networks (RPL) Option for Carrying RPL Information in Data-Plane Datagrams", RFC 6553, DOI 10.17487/RFC6553, March 2012, <<https://www.rfc-editor.org/info/rfc6553>>.
- [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, <<https://www.rfc-editor.org/info/rfc6554>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.

10.2. Informative References

- [Ciolkoszl9] Ciolkosz, P., "Integration of the RNFD Algorithm for Border Router Failure Detection with the RPL Standard for Routing IPv6 Packets", Master's Thesis, University of Warsaw, 2019.
- [Iwanicki16] Iwanicki, K., "RNFD: Routing-layer detection of DODAG (root) node failures in low-power wireless networks", In IPSN 2016: Proceedings of the 15th ACM/IEEE International Conference on Information Processing in Sensor Networks, IEEE, pp. 1--12, DOI 10.1109/IPSN.2016.7460720, 2016, <<https://doi.org/10.1109/IPSN.2016.7460720>>.

- [Paszkowska19] Paszkowska, A. and K. Iwanicki, "Failure Handling in RPL Implementations: An Experimental Qualitative Study", In Mission-Oriented Sensor Networks and Systems: Art and Science (Habib M. Ammari ed.), Springer International Publishing, pp. 49--95, DOI 10.1007/978-3-319-91146-5_3, 2019, <https://doi.org/10.1007/978-3-319-91146-5_3>.
- [RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861, DOI 10.17487/RFC4861, September 2007, <<https://www.rfc-editor.org/info/rfc4861>>.
- [RFC5184] Teraoka, F., Gogo, K., Mitsuya, K., Shibui, R., and K. Mitani, "Unified Layer 2 (L2) Abstractions for Layer 3 (L3)-Driven Fast Handover", RFC 5184, DOI 10.17487/RFC5184, May 2008, <<https://www.rfc-editor.org/info/rfc5184>>.
- [RFC6202] Loreto, S., Saint-Andre, P., Salsano, S., and G. Wilkins, "Known Issues and Best Practices for the Use of Long Polling and Streaming in Bidirectional HTTP", RFC 6202, DOI 10.17487/RFC6202, April 2011, <<https://www.rfc-editor.org/info/rfc6202>>.
- [RFC7102] Vasseur, JP., "Terms Used in Routing for Low-Power and Lossy Networks", RFC 7102, DOI 10.17487/RFC7102, January 2014, <<https://www.rfc-editor.org/info/rfc7102>>.
- [RFC7228] Bormann, C., Ersue, M., and A. Keranen, "Terminology for Constrained-Node Networks", RFC 7228, DOI 10.17487/RFC7228, May 2014, <<https://www.rfc-editor.org/info/rfc7228>>.
- [RFC7416] Tsao, T., Alexander, R., Dohler, M., Daza, V., Lozano, A., and M. Richardson, Ed., "A Security Threat Analysis for the Routing Protocol for Low-Power and Lossy Networks (RPLs)", RFC 7416, DOI 10.17487/RFC7416, January 2015, <<https://www.rfc-editor.org/info/rfc7416>>.
- [Whang90] Whang, K.-Y., Vander-Zanden, B.T., and H.M. Taylor, "A Linear-time Probabilistic Counting Algorithm for Database Applications", In ACM Transactions on Database Systems, DOI 10.1145/78922.78925, 1990, <<https://doi.org/10.1145/78922.78925>>.

Author's Address

Konrad Iwanicki
University of Warsaw
Banacha 2
02-097 Warszawa
Poland

Phone: +48 22 55 44 428
Email: iwanicki@mimuw.edu.pl