ROLL
Internet-Draft
Intended status: Standards Track
Expires: April 22, 2019

C. Ji, Ed.
Alexander TEI of Thessaloniki
R. Koutsiamanis
G. Papadopoulos
IMT Atlantique
D. Dujovne
Universidad Diego Portales
N. Montavont
IMT Atlantique

October 19, 2018

Traffic-aware Objective Function
draft-ji-roll-traffic-aware-objective-function-03

Abstract

This document proposes a remaining throughput metric for parent and DODAG selection. This metric represents the amount of remaining traffic handling capacity that the node has. This document also proposes an Objective Function (OF) which uses the proposed metric for parent and DODAG selection to balance the amount of traffic between nodes and DODAGs.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

RPL [RFC6550] is an IPv6 Routing protocol for LLNs. It uses Objective Functions (OF) to construct the Destination Oriented Directed Acyclic Graph (DODAG) containing the nodes of the network. The existing OFs defined are OF Zero (OF0) [RFC6552] and Minimum Rank with Hysteresis OF (MRHOF) [RFC6719]. These OFs specify how nodes in a DODAG select their preferred parent using different metrics.

The metrics can be separated into two different types, link metrics (e.g. ETX) and node metrics (e.g. energy). Experimental results [I-D.qasem-roll-rpl-load-balancing] conclude that using the current OFs leads to an unbalanced network within which some nodes are overloaded. Here, a node is overloaded in the sense that it forwards many more packets than it otherwise would if the network were balanced. This problem has consequences for the lifetime of the network because overloaded nodes drain quicker than others, a problem which becomes even more significant when the overloaded nodes are near the DODAG root [I-D.qasem-roll-rpl-load-balancing].

Similarly, one DODAG might be overloaded in the same sense compared to another DODAG, and this will lead to the same consequences for the whole DODAG as for a specific node.

This problem is still an open issue. This draft proposes a new way of parent and DODAG selection as an attempt towards a solution. This draft proposes a new OF that considers the remaining throughput as a representation of the remaining traffic handling capacity each node possesses and which uses this information to balance the amount of traffic between nodes and DODAGs.

In brief, each node tracks its remaining throughput and appends this information as a DAG Metric Container option to DIO messages it sends. When the DIO message is received by child nodes or potential child nodes, the remaining throughput information is stored and used to influence the result when RPL parent or DODAG selection is performed.

2. Terminology

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

3. DODAG construction in RPL

RPL uses OFs to construct a DODAG. OFs define the way the nodes select their preferred parent and DODAG and how they compute the new rank. A node's rank is always larger than its parent's rank because the calculation of rank is based on an increment to the parent's rank. This increment differs for each OF but all OFs include the MinHopRankIncrease, which is the minimum increase in rank between a node and a node's parent and a step. Different OFs use different metrics or constraints to select the preferred parent and DODAG and to define the step, depending on application requirements. Nodes obtain these values from DODAG Information Object (DIO) control messages sent by their neighbor nodes.

The construction of a DODAG starts when the root node sends DIO messages to its neighbors. After receiving the DIO, these neighbor nodes select the root as their preferred parent if they wish to join the DODAG. To announce that they joined the DODAG as its child node, they send a Destination Advertisement Object (DAO) to their preferred parent - the DODAG root. After joining the DODAG, these nodes send their own DIO messages with the new computed rank to their neighbors. This procedure repeats for every node which joins the DODAG.

4. Load distribution problem in RPL

According to the experiments conducted using existing OFs RPL faces a load distribution problem in large LLNs. With RPL using existing OFs, such as MRHOF, an unbalanced network is formed with some nodes

overloaded and other nodes at rest. This problem is severe for network performance because overloaded nodes will use up their available energy faster than other nodes. This is exacerbated for nodes near the root (within 1 hop distance) or nodes which are the only parent candidate for other nodes. Additionally, when the overloaded node shuts down, a big part of the network will become disconnected and will have to be transferred to another parent or DODAG. There is a high probability that the children nodes will also select the same new node as their parent or the same DODAG, leading to another overloaded node/DODAG. Also, when a node has selected its parent, it will change only when the parent node is not reachable (due to battery depletion or packet losses).

The existing OFs usually use a single metric to compare parent candidates, for example, as described in [RFC6719] the default metric used in MRHOF is ETX [RFC6551], which represents the number of transmissions a node expects to make to a destination to successfully deliver one packet. The result from using a single metric is that nodes prefer to select the same node as their parent, which according to [I-D.qasem-roll-rpl-load-balancing] leads to an unbalanced network with overloaded nodes (node load is indicated by a node's child count). But the child count does not accurately indicate the load because among the child nodes some may have higher traffic load and others may have lower.

The network traffic can be quantified by tracking the packets a node generates/sends/receives and the amount of energy it consumes. Energy consumption is strongly correlated to the amount of network traffic handled by a node since the energy consumption for the operation of the radio is the primary energy consumer in typical nodes. However, directly measuring the remaining throughput is both more accurate and also works when nodes have atypical energy consumption profiles (e.g. increased node processing or high energy consumption sensors).

Calculating the remaining throughput then requires knowledge of the total throughput supported by a node and subtraction of the used throughput.

4.1. Parent selection problem

```
           T:4p/s               |         T:4p/s
           U:4p/s (R)           |         U:4p/s (R)
                   ^            |                 ^
                   |            |                 |
            +--------------+    |          +------+-------+
            |              |    |          |              |
    T:2p/s  +      T:2p/s  +    |  T:2p/s  +      T:2p/s  +
    U:3p/s (A)     U:1p/s (B)   |  U:2p/s (A)     U:2p/s (B)
            ^              ^    |          ^              ^
            |              |    |          |              |
      +-----+------+       |    |    +-----+       +------+
      |     |      |       |    |    |     |       |      |
      +     +      +       +    |    +     +       +      +
     (C1)  (C2)   (C3)    (D1)  |   (C1)  (C2)    (C3)   (D1)
   U:1p/s U:1p/s U:1p/s  U:1p/s | U:1p/s U:1p/s  U:1p/s U:1p/s
                                |
             Unbalanced         |            Balanced
```

Figure 1: Use of Remaining Throughput with nodes with the same requirements

As a first simple example, an unbalanced network with nodes which all use the same throughput ("U:") is shown in Figure 1. Nodes A and B have the same total throughput ("T:"), but node A is overloaded due to trying to handle more than its ability while node B has a spare throughput of 2-1=1p/s. Its transformation into a balanced network is shown on the right and it involves a node (C3) switching parents from A to B so that the capacity of its parent is no longer exceeded.

```
            T:6p/s               |          T:6p/s
            U:6p/s (R)           |          U:6p/s (R)
                    ^            |                  ^
                    |            |                  |
             +------+-------+    |           +--------------+
             |              |    |           |              |
     T:3p/s  +      T:3p/s  +    |   T:3p/s  +      T:3p/s  +
     U:2p/s (A)     U:4p/s (B)   |   U:3p/s (A)     U:3p/s (B)
             ^              ^    |           ^              ^
             |              |    |           |              |
       +-----+       +------+    |     +-----+------+       |
       |     |       |      |    |     |     |      |       |
       +     +       +      +    |     +     +      +       +
      (C1)  (C2)    (D1)   (D2)  |    (C1)  (C2)   (D1)    (D2)
    U:1p/s U:1p/s  U:1p/s U:3p/s |  U:1p/s U:1p/s U:1p/s  U:3p/s
                                 |
                Unbalanced       |            Balanced
```

Figure 2: Use of Remaining Throughput with nodes with different requirements

As a second simple example, an unbalanced network with nodes which have different throughput ("U:") is shown in Figure 2. In this case, node B is overloaded and node D1 should move to parent A, which as a space throughput of 1p/s. Its transformation into a balanced equivalent network is shown on the right.

4.2. DODAG selection problem

The purpose of the following example is to show the problem of DODAG selection, and not to focus on selecting the best parent.

```
.... DODAG 1 ... ... DODAG 2 .... | .... DODAG 1 ... ... DODAG 2 ....
.               .               . | .               .               .
.  T:4p/s       .  T:4p/s       . | .  T:4p/s       .  T:4p/s       .
.  U:4p/s (R1)  .  U:3p/s (R2)  . | .  U:5p/s (R1)  .  U:3p/s (R2)  .
.          ^    .          ^    . | .          ^    .          ^    .
.          |    .          |    . | .          |    .          |    .
.    +-----+    .    +-----+    . | .    +-----+    .    +-----+    .
.    |     |    .    |     |    . | .    |     |    .    |     |    .
.    +     +    .    +     +    . | .    +     +    .    +     +    .
.   (A1)  (B1)  .   (A2)  (B2)  . | .   (A1)  (B1)  .   (A2)  (B2)  .
. U:3p/s U:1p/s . U:2p/s U:1p/s . | . U:3p/s U:1p/s . U:2p/s U:1p/s .
................ ................ | ................ ................
                                  |            ^
                                  |            |
                ^                 |            +----+
                |                 |                 |
                +                 |                 +
               (C)                |                (C)
             U:1p/s               |              U:1p/s
                                  |
            Joining               |        Joined - Unbalanced
```

Figure 3: DODAG selection example leading to unbalanced traffic with RT metric

In the example in Figure 3, there are two DODAGs (DODAG 1 and DODAG 2) that belong to the same RPL Instance and a node (C) that must select a DODAG. Node C has to pick from the information provided by its two reachable neighbors: B1 and A2. On the left, node C is shown before selecting the preferred DODAG, while on the right it is shown after the DODAG selection.

Node C might choose B1 in DODAG 1 to be its preferred parent since the traffic information of the two DODAGs is not available. However, at the root node (R1), it can be observed that the total network traffic is higher in DODAG 1 and that after node C joins it, the traffic handling capacity of the root R1 has been exceeded.

```
   .... DODAG 1 ... ... DODAG 2 .... | .... DODAG 1 ... ... DODAG 2 ....
   .               .               . | .               .               .
   .  T:4p/s       .  T:4p/s       . | .  T:4p/s       .  T:4p/s       .
   .  U:4p/s (R1)  .  U:3p/s (R2)  . | .  U:4p/s (R1)  .  U:4p/s (R2)  .
   .          ^    .          ^    . | .          ^    .          ^    .
   .          |    .          |    . | .          |    .          |    .
   .    +-----+    .    +-----+    . | .    +-----+    .    +-----+    .
   .    |     |    .    |     |    . | .    |     |    .    |     |    .
   .    +     +    .    +     +    . | .    +     +    .    +     +    .
   .   (A1)  (B1)  .   (A2)  (B2)  . | .   (A1)  (B1)  .   (A2)  (B2)  .
   . U:3p/s U:1p/s . U:2p/s U:1p/s . | . U:3p/s U:1p/s . U:2p/s U:1p/s .
   ................ ................ | ................ ................
                                     |                      ^
                                     |                      |
                   ^                 |                 +----+
                   |                 |                 |
                   +                 |                 +
                  (C)                |                (C)
                U:1p/s               |              U:1p/s
                                     |
               Joining               |          Joined - Balanced
```

Figure 4: DODAG selection example leading to balanced traffic with RT metric

If the traffic handling capacity information is available, then node C could make a more efficient decision by using DODAG 2 and selecting node A2 as the preferred parent, as shown in Figure 4. Such a selection is based on the traffic of the entire DODAG and would not lead to exceeding the traffic handling capacity of the root R2 since this root had spare capacity.

5. TAOF description

In this specification, a metric is proposed to be used in the parent and DODAG selection mechanism, the Remaining Throughput (RT) which represents the number of packets each node can transmit (send or forward) during a certain time period. The period used, named THROUGHPUT_PERIOD, is a parameter common to the whole RPL instance. This parameter CAN be pre-configured on all the nodes. The period used SHOULD coincide with a sliding window of the same size used to calculate the packets transferred during this period to facilitate the calculation of the remaining possible packet transmissions. Therefore, whenever the RT value is reported it will refer to the previous THROUGHPUT_PERIOD period of time. This information is added in DIO messages and is broadcast to every neighbor.

At first, each node MUST identify from their neighbor set which nodes are acceptable to be selected as a parent. For this purpose, the metric ETX is used as a filter to filter out parent candidates with low link quality with a preference for nodes with link quality below a given threshold. The ETX threshold SHOULD be different depending on application requirements. The suggested value for the relevant threshold MAX_PATH_COST from MRHOF [RFC6719] is 32768, which means the specific path has expected transmission counts greater than 256.

When the ETX value is used as a filter, nodes with bad link quality will not be included in the parent set. This ensures that undue retransmissions caused by bad links will be avoided. After all the filtering is done, if any, the node chooses the parent candidate or DODAG with the highest remaining throughput.

For the purpose of DODAG specifically, the A field in the Routing Metric/Constraint Flag field object [RFC6551] SHOULD be set to 1, indicating that the value reported is a maximum. Furthermore, when a node is calculating the value of RT to broadcast in a DIO, the value reported SHOULD be the minimum of two values: its parent RT and the node's own calculated remaining throughput. Thus, the value broadcasted will be the available remaining throughput in the whole path from the node to the DODAG root.

This proposal is expected to increase the frequency of parent changes because the remaining throughput is more likely to be different between DIO messages, even for DIO messages from the same node. There are multiple ways to minimize the frequency of unnecessary parent changes:

a. Use the remaining throughput in combination with another metric (e.g. child count, hop counts).

b. Use a threshold when comparing the remaining throughput, similar to the approach in MRHOF [RFC6719]. Switch parents when the difference of remaining throughput between the original parent and the alternative parent is above a threshold. This threshold depends on different factors (e.g. network size, average traffic load) and SHOULD be defined differently for each use case.

6. DIO Metric Container Type extension

```
     0                   1
     0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |              RT               |    RT: Remaining Throughput
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5: DAG metric container type format.

A DIO message carries fields as described in RFC6550 [RFC6550] and the available options for the DAG metric container are described in RFC6551 [RFC6551]. In this specification, a metric container option is proposed and the detailed format is shown in Figure 5. The information carried is the RT, represented as a 2 byte unsigned integer and the unit is packets per THROUGHPUT_PERIOD time.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| RPLInstanceID |Version Number |             Rank              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|G|0| MOP | Prf |     DTSN      |     Flags     |   Reserved    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
+                                                               +
|                                                               |
+                            DODAGID                            +
|                                                               |
+                                                               +
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| DAGMC Type (2)| DAGMC Length  |                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+                               |
|                                                               |
//                   DAG Metric Container data                 //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 6: Example DIO Message with a DAG Metric Container option

The structure of the DIO Control Message when a DAG Metric Container option is included is shown in Figure 6. The DAG Metric Container option type (DAGMC Type in Figure 6) has the value 0x02 as per the IANA registry for the RPL Control Message Options and is defined in [RFC6550]. The DAG Metric Container option length (DAGMC Length in Figure 6) expresses the DAG Metric Container length in bytes. DAG Metric Container data holds the actual data and is shown further expanded in Figure 7.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Routing-MC-Type|Res Flags|P|C|O|R| A   |  Prec | Length (bytes)|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             RT                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 7: DAG Metric Container (MC) data with Remaining Throughput (RT) object body

An example DAG Metric Container containing the proposed Metric Container object is shown in Figure 7. The explicit definition of the fields is:

Routing-MC-Type: TBD1. The type of the proposed DAGMC extension. To be assigned by IANA.

Remaining Throughput (RT): The remaining throughput, represented as a 2-byte unsigned integer in units of packets per THROUGHPUT_PERIOD time.

7. Enrollment

The RT metric SHOULD be used not only for ongoing parent selection but also especially within enrollment, i.e. the process a node follows to join a 6TiSCH network. In accordance to [I-D.ietf-6tisch-enrollment-enhanced-beacon], the value in RT SHOULD be used to affect the enrollment process so that a new node will be able to directly select a DODAG which will be able to cover its traffic needs and to spread the traffic load between different DODAGs. More specifically, the pan priority field described in [I-D.ietf-6tisch-enrollment-enhanced-beacon] can be derived from the RT value. For this purpose, the RT value SHOULD be used with the maximum value aggregation mode (A field in the Routing Metric/ Constraint Flag field object [RFC6551] set to 1), to report the maximum remaining throughput in the whole path to the DODAG root. The pan priority field is an unsigned 8-bit integer with lower values signifying higher priority while the RT value is a 16-bit unsigned integer with higher values signifying more remaining throughput. To convert the RT value to a pan priority the following formula should be used:

pan priority = 16 - FLOOR(LOG2(RT + 1))

where LOG2 is the logarithm function with a base of 2. The use of the LOG function allows having higher accuracy in the low values of

the remaining throughput, where small value differences are significant, and lower accuracy in the high values of the remaining throughput, where small differences are less significant. The addition of 1 to the RT allows converting RT=0.

8. Security Considerations

The structure of the DIO control message is extended, within the pre-defined DIO options. Therefore, the security mechanisms defined in RPL [RFC6550] apply to this proposed extension.

9. IANA Considerations

This proposal requests the allocation of a new value TBD1 for the metric type "RT" in the Routing-MC-Type field in the DAG MC from IANA.

Additionally, an Objective Code Point (OCP) with value TBD2 for TAOF needs to be assigned in the Objective Code Point Registry as described in Section 20.5 of [RFC6550].

10. Informative references

[I-D.ietf-6tisch-enrollment-enhanced-beacon]
Dujovne, D. and M. Richardson, "IEEE802.15.4 Informational Element encapsulation of 6tisch Join and Enrollment Information", draft-ietf-6tisch-enrollment-enhanced-beacon-00 (work in progress), July 2018.

[I-D.qasem-roll-rpl-load-balancing]
Qasem, M., Al-Dubai, A., Romdhani, I., Ghaleb, B., Hou, J., and R. Jadhav, "Load Balancing Objective Function in RPL", draft-qasem-roll-rpl-load-balancing-02 (work in progress), October 2017.

[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>.

[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>.

[RFC6551] Vasseur, JP., Ed., Kim, M., Ed., Pister, K., Dejean, N., and D. Barthel, "Routing Metrics Used for Path Calculation in Low-Power and Lossy Networks", RFC 6551, DOI 10.17487/RFC6551, March 2012, <https://www.rfc-editor.org/info/rfc6551>.

[RFC6552] Thubert, P., Ed., "Objective Function Zero for the Routing Protocol for Low-Power and Lossy Networks (RPL)", RFC 6552, DOI 10.17487/RFC6552, March 2012, <https://www.rfc-editor.org/info/rfc6552>.

[RFC6719] Gnawali, O. and P. Levis, "The Minimum Rank with Hysteresis Objective Function", RFC 6719, DOI 10.17487/RFC6719, September 2012, <https://www.rfc-editor.org/info/rfc6719>.

Authors' Addresses

Chenyang Ji (editor)
Alexander TEI of Thessaloniki
Department of Informatics
Thessaloniki 57400
GREECE

Email: jichenyang920@gmail.com

Remous-Aris Koutsiamanis
IMT Atlantique
Office B00 - 126A
2 Rue de la Chataigneraie
Cesson-Sevigne - Rennes 35510
FRANCE

Phone: +33 299 12 70 49
Email: aris@ariskou.com

Georgios Papadopoulos
IMT Atlantique
Office B00 - 114A
2 Rue de la Chataigneraie
Cesson-Sevigne - Rennes 35510
FRANCE

Phone: +33 299 12 70 04
Email: georgios.papadopoulos@imt-atlantique.fr

Diego Dujovne
Universidad Diego Portales
Escuela de Informatica y Telecomunicaciones, Av. Ejercito 441
Santiago, Region Metropolitana
Chile

Phone: +56 (2) 676-8121
Email: diego.dujovne@mail.udp.cl

Nicolas Montavont
IMT Atlantique
Office B00 - 106A
2 Rue de la Chataigneraie
Cesson-Sevigne - Rennes 35510
FRANCE

Phone: +33 299 12 70 23
Email: nicolas.montavont@imt-atlantique.fr