RAW   G. Papadopoulos, Ed.
Internet-Draft   R. Koutsiamanis
Intended status: Standards Track   N. Montavont
Expires: December 22, 2019   IMT Atlantique
                                                                 P. Thubert
                                                                      Cisco
                                                              June 20, 2019

Exploiting Packet Replication and Elimination in Complex Tracks in LLNs
                  draft-papadopoulos-raw-pareo-reqs-00

Abstract

   The Packet Replication and Elimination (PRE) mechanism duplicates
   data packets into several paths in the network to increase
   reliability and provide low jitter. PRE may be used to complement
   layer-2 Automatic Repeat reQuest (ARQ) and receiver-end Ordering to
   form the PAREO functions. Over a wireless medium, this technique can
   take advantage of communication Overhearing, when parallel
   transmissions over two adjacent paths are scheduled. This document
   presents the concept and details the required changes to the current
   specifications that will be necessary to enable the PAREO functions.

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 December 22, 2019.

Table of Contents

1. Introduction

   This draft describes industrial use cases which require deterministic
   flows over wireless multi-hop paths.

   The RAW use cases explicitly do not propose any specific solution or
   design for the RAW architecture or protocols. These are the subjects
   of other RAW drafts. The RAW use cases are not considered to be
   concrete requirements by the RAW Working Group.

   The industrial use cases covered in this draft are professional
   audio, wireless for industrial applications and amusement parks.

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. Tracks

3.1. Tracks Overview

   The 6TiSCH architecture introduces the concept of Tracks in 6TiSCH
   Architecture [I-D.ietf-6tisch-architecture]. A simple track is
   composed of a sequence of cells (a combination of a transmitter, a
   receiver and a given channel offset) to ensure the transmission of a
   single packet from a source node to a destination node across a
   multihop path.

3.2. Complex Tracks

   A Complex Track is designed as a directed acyclic graph from a source
   node towards a destination node to support multi-path forwarding, as
   introduced in 6TiSCH Architecture [I-D.ietf-6tisch-architecture]. By
   employing DetNet [I-D.ietf-detnet-architecture] Packet Replication
   and Elimination (PRE) functions, several paths may be computed, and
   these paths may be more or less independent. For example, a complex
   Track may branch off and rejoin over non-congruent paths (branches).

   Some more details for Deterministic Network PRE techniques are
   presented in the following Section.

4. Packet Replication and Elimination principles

   In a nutshell, PRE establishes several paths in a network to provide
   redundancy and parallel transmissions to bound the end-to-end delay
   to traverse the network. Optionally, promiscuous listening between
   paths is possible, such that the nodes on one path may overhear
   transmissions along the other path. Considering the scenario shown
   in Figure 1, many different paths are possible for S to reach R. A
   simple way to benefit from this topology could be to use the two
   independent paths via nodes A, C, E and via B, D, F. But more
   complex paths are possible by interleaving transmissions from the
   lower level of the path to the upper level.

   PRE may also take advantage of the shared properties of the wireless
   medium to compensate for the potential loss that is incurred with
   radio transmissions. For instance, when the source sends to A, B may
   listen also and get a second chance to receive the frame without an
   additional transmission. Note that B would not have to listen if it

already received that particular frame at an earlier timeslot in a
dedicated transmission towards B.

```
                    (A)   (C)   (E)

        source (S)                      (R) (root)

                    (B)   (D)   (F)
```
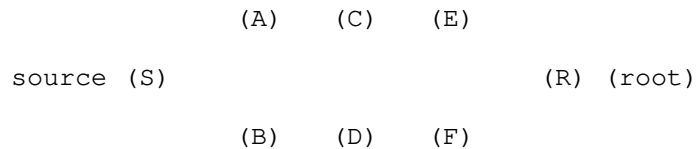
Figure 1: A Typical Ladder Shape with Two Parallel Paths Toward the
Destination

The PRE model can be implemented in both centralized and distributed
scheduling approaches. In the centralized approach, a Path
Computation Element (PCE) scheduler calculates the routes and
schedules the communication among the nodes along a circuit such as a
Label switched path. In the distributed approach, each node selects
its route to the destination, typically using a source routing
header. In both cases, at each node in the paths, a default parent
and alternative parent(s) should be selected to set up complex
tracks.

In the following Subsections, all the required operations defined by
PRE, namely, Alternative Path Selection, Packet Replication, Packet
Elimination and Promiscuous Overhearing, are described.

4.1. Packet Replication

The objective of PRE is to provide deterministic networking
properties: high reliability and bounded latency. To achieve this
goal, determinism in every hop of the forwarding paths MUST be
guaranteed. By employing a Packet Replication procedure, each node
forwards a copy of each data packet to multiple parents: its Default
Parent (DP) and multiple Alternative Parents (APs). To do so, each
node (i.e., source and intermediate node) transmits the data packet
multiple times in unicast to each parent. For instance, in Figure 2,
the source node S is transmitting the packet to both parents, nodes A
and B, in two different timeslots within the same TSCH slotframe. An
example TSCH schedule is shown in Figure 3. Thus, the packet
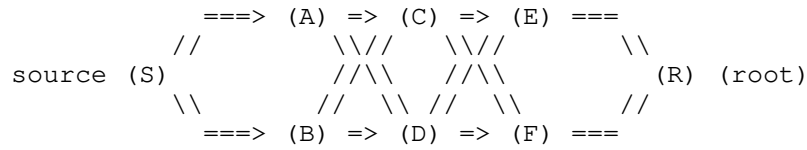eventually obtains parallel paths to the destination.

```
                      ===> (A) => (C) => (E) ===
                     //        \\//   \\//        \\
          source (S)          //\\   //\\          (R) (root)
                     \\        // \\ // \\        //
                      ===> (B) => (D) => (F) ===
```

Figure 2: Packet Replication: S transmits twice the same data packet,
                    to its DP (A) and to its AP (B).

```
                                  Timeslot
+---------++------+------+------+------+------+------+------+
| Channel || 0    | 1    | 2    | 3    | 4    | 5    | 6    |
+---------++======+======+======+======+======+======+======+
| 0       || S->A | S->B | B->C | B->D | C->F | E->R | F->R |
+---------++------+------+------+------+------+------+------+
| 1       ||      | A->C | A->D | C->E | D->E | D->F |      |
+---------++------+------+------+------+------+------+------+
```

Figure 3: Packet Replication: Sample TSCH schedule

4.2. Packet Elimination

   The replication operation increases the traffic load in the network,
   due to packet duplications. Thus, a Packet Elimination operation
   SHOULD be applied at each RPL DODAG level to reduce the unnecessary
   traffic. To this aim, once a node receives the first copy of a data
   packet, it discards the subsequent copies. Because the first copy
   that reaches a node is the one that matters, it is the only copy that
   will be forwarded upward. Then, once a node performs the Packet
   Elimination operation, it will proceed with the Packet Replication
   operation to forward the packet toward the RPL DODAG Root.

4.3. Promiscuous Overhearing

   Considering that the wireless medium is broadcast by nature, any
   neighbor of a transmitter may overhear a transmission. By employing
   the Promiscuous Overhearing operation, a DP and some AP(s) eventually
   have more chances to receive the data packets. In Figure 4, when
   node A is transmitting to its DP (node C), the AP (node D) and its
   sibling (node B) may decode this data packet as well. As a result,
   by employing corellated paths, a node may have multiple opportunities
   to receive a given data packet. This feature not only enhances the
   end-to-end reliability but also it reduces the end-to-end delay and
   increases energy efficiency.

```
                ===> (A) ====> (C) ====> (E) ====
             //       ^  | \\                      \\
   source (S)         |  |   \\                       (R) (root)
             \\       |  v    \\                     //
                ===> (B) ====> (D) ====> (F) ====
```
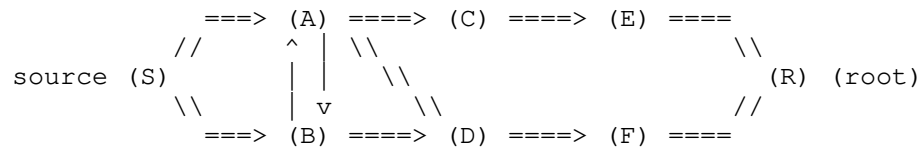
Figure 4: Unicast to DP with Overhearing: by employing Promiscuous Overhearing, DP, AP and the sibling nodes have more opportunities to receive the same data packet.

5. Requirements

5.1. Requirements Related to Alternative Parent Selection

To perform the Packet Replication procedure, it is necessary to define the Alternative Parent(s) and, consequently, the path to the destination node, for each node in the wireless network. An AP can be selected in many different ways, and is dependent on the implementation.

The requirements are:

Req1.1: The routing protocol SHOULD be extended to allow for each node to select AP(s) in addition to the DP. This enables packet replication to multiple parents.

Req1.2: Considering that the Packet Replication procedure significantly increases the traffic in a network, when proposing solutions for Alternative Parent Selection, they should be efficient enough to mitigate the potential uncontrolled packet duplications.

Req1.3: The topology SHOULD be defined when proposing solutions for Alternative Parent Selection. For instance, the ladder topology should be defined explicitly e.g., number of parallel paths, density.

5.2. Requirements Related to Propagated Information

For Alternative Parent(s) selection, nodes MAY need additional information about the network topology. This draft does not prescribe the information required for AP selcetion or how it is to be propagated to the nodes that need to select AP(s). TODO: To be discussed.

The requirement is:

   Req2.1: Nodes MUST have a way of receiving the required information
           for efficient Alternative Parent Selection.

   As an example, it is possible to use and extend the RPL [RFC6550]
   DODAG Information Object (DIO) Control Message to allow nodes to
   propagate information about themselves to potential children. For
   instance, "RPL DAG Metric Container (MC) Node State and Attribute
   (NSA) object type extension" [I-D.ietf-roll-nsa-extension] focuses on
   extending the DAG Metric Container [RFC6551] by defining a new type-
   length-value (TLV), entitled Parent Set (PS) which can be carried in
   the Node State and Attribute (NSA) object.

5.3. Requirements Related to Promiscuous Overhearing

   As stated previously, to further increase the network reliability and
   to achieve deterministic packet deliveries at the destination node,
   Promiscuous Overhearing can be considered.

   As it is described in BCP 210 [RFC8180], in TSCH mode, the data
   frames are transmitted in unicast mode and are acknowledged by the
   receiving neighbor. To perform the promiscuous overhearing
   procedure, there SHOULD be an option for the transmitted frames,
   i.e., in unicast, to be overheard by the potential neighborhood node.

   Destination address filtering is performed at the Medium Access
   Control (MAC) layer. For example, according to IEEE std. 802.15.4
   [IEEE802154-2015], a node receiving a packet with a destination
   address different than its own and different to 0xFF discards the
   packet. A change is needed to be able to receive packets whose
   destination address is neither multicast nor the overhearing node's
   MAC address.

   The requirements are:

   Req3.1: The MAC implementation MUST be able to disable MAC address
           filtering to accept the overheard frame.

   Req3.2: The 6top Protocol [RFC8480] specification MUST be extended
           to indicate disabling MAC filtering in a receiving cell. This
           can be achieved by reserving a bit in the 6P CellOptions Bitmap
           (Section 6.2.6 [RFC8480]) for this purpose.

   Req3.3: The overhearing node can be configured with the timeslot set
           to shared reception, thus, there will be no acknowledgement
           from it. However, there is the security issue that needs to be
           considered. Since the overhearing case implies that it is not
           possible to have per-pair keying, there MUST be a key that the
           overhearing node will be aware of. Hence, the Minimal Security

Framework for 6TiSCH [I-D.ietf-6tisch-architecture]
specification should consider such a scenario.

5.4. Requirements Related to Packet Elimination

By employing Packet Replication, the wireless network is expected to
also perform Packet Elimination to restrict the number of the
duplicated packets, i.e., the unnecessary traffic. As per the 6TiSCH
Architecture [I-D.ietf-6tisch-architecture], 6TiSCH has no position
about how the sequence numbers would be tagged in the packet.

The requirement is:

Req4.1: To perform Packet Elimination the packet copies MUST contain
         a sequence number which allows identifying the copies.

6. Security Considerations

TODO.

7. IANA Considerations

This document has no IANA considerations.

8. References

8.1. Informative references

[I-D.ietf-6tisch-architecture]
          Thubert, P., "An Architecture for IPv6 over the TSCH mode
          of IEEE 802.15.4", draft-ietf-6tisch-architecture-20 (work
          in progress), March 2019.

[I-D.ietf-detnet-architecture]
          Finn, N., Thubert, P., Varga, B., and J. Farkas,
          "Deterministic Networking Architecture", draft-ietf-
          detnet-architecture-13 (work in progress), May 2019.

[I-D.ietf-roll-nsa-extension]
          Koutsiamanis, R., Papadopoulos, G., Montavont, N., and P.
          Thubert, "RPL DAG Metric Container Node State and
          Attribute object type extension", draft-ietf-roll-nsa-
          extension-01 (work in progress), March 2019.

[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>.

   [RFC8180] Vilajosana, X., Ed., Pister, K., and T. Watteyne, "Minimal
              IPv6 over the TSCH Mode of IEEE 802.15.4e (6TiSCH)
              Configuration", BCP 210, RFC 8180, DOI 10.17487/RFC8180,
              May 2017, <https://www.rfc-editor.org/info/rfc8180>.

   [RFC8480] Wang, Q., Ed., Vilajosana, X., and T. Watteyne, "6TiSCH
              Operation Sublayer (6top) Protocol (6P)", RFC 8480,
              DOI 10.17487/RFC8480, November 2018,
              <https://www.rfc-editor.org/info/rfc8480>.

8.2. Other Informative References

   [IEEE802154-2015]
              IEEE standard for Information Technology, "IEEE standard
              for Information Technology, "IEEE Std 802.15.4-2015
              Standard for Low-Rate Wireless Personal Area Networks
              (WPANs)", December 2015".

Authors' Addresses

   Georgios Papadopoulos (editor)
   IMT Atlantique
   Office B00 - 102A
   2 Rue de la Chataigneraie
   Cesson-Sevigne - Rennes 35510
   FRANCE

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

   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

   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

   Pascal Thubert
   Cisco Systems, Inc
   Building D
   45 Allee des Ormes - BP1200
   MOUGINS - Sophia Antipolis 06254
   FRANCE

   Phone: +33 497 23 26 34
   Email: pthubert@cisco.com