    Exploiting Packet Replication and Elimination in Complex Tracks in LLNs
                    draft-papadopoulos-paw-pre-reqs-00

Abstract

   Packet Replication and Elimination mechanism consists in duplicating
   data packets into several paths in the network to increase
   reliability and provide low jitter.  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 this.

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 August 1, 2019.

Copyright Notice

Table of Contents

## 1.  Introduction

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

   The PAW use cases explicitly do not propose or suggest any specific
   solution or design for PAW architecture or protocols.  These are
   subjects of other PAW drafts.  The PAW use cases are not considered
   as concrete requirements by the PAW 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).

In the following Section, we will detail Deterministic Networks PRE
techniques.

## 4. Packet Replication and Elimination principles

In a nutshell, PRE consists in establishing several paths in a
network to provide redundancy and parallel transmissions to bound the
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 depicted in Figure 1, many different paths are possible for
S to reach R.  A simple way to take 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 one level of the path to the upper level.

The PRE may also take advantage of the shared properties of the
medium to compensate for the potential loss that is incurred with
radio transmissions.  For instance, when the source sends to A, B may
listen 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.

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

          source (S)                      (R) (root)

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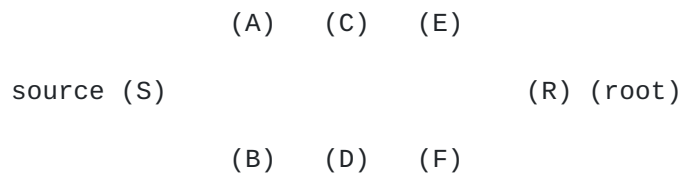
Figure 1: A Typical Ladder Shape with Two Parallel Paths Toward the
Destination

PRE model can be implemented in both centralized and distributed
scheduling approach.  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, a default parent and alternate parent(s) should be
selected to set up complex tracks.

In the following Subsections, detailed description of all required
operations defined by PRE, namely, Alternative Path Selection, Packet
Replication, Packet Elimination and Promiscuous Overhearing, will be
described.

## 4.1.  Packet Replication

The objective of PRE is to provide deterministic networking
properties, with high reliability and bounded latency.  To achieve
this goal, determinism in every hop of the forwarding paths MUST be
guaranteed.  By employing Packet Replication procedure, each node
transmits (i.e., replicates) each data packet to both its Default
Parent (DP) and Alternative Parent (AP).  To do so, each node (i.e.,
source and intermediate nodes) transmits the data packet twice 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.  Thus, the packet
eventually obtains parallel paths to the destination.

```
               ===> (A) => (C) => (E) ===
              //          \\//    \\//          \\
      source (S)          //\\    //\\          (R) (root)
              \\          //  \\ //  \\          //
               ===> (B) => (D) => (F) ===
```
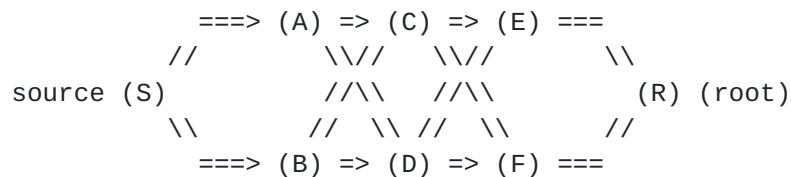
Figure 2: Packet Replication: S transmits twice the same data packet,
to its DP (A) and to its AP (B).

## 4.2.  Packet Elimination

The replication operation increases the traffic load in the network,
due to packet duplications.  Thus, 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 following copies.  Because the first copy
that reaches a node is the one that counts, it is the only copy that
will be forwarded upward.  Then, once a node performed the Packet
Elimination operation, it will proceed with 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, DP and AP eventually have more
chances to receive the data packets.  In Figure 3, 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 correlated 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.

```
                ===> (A) ====> (C) ====> (E) ====
             //      ^ | \\                     \\
      source (S)     | |   \\                     (R) (root)
             \\      | v     \\                  //
                ===> (B) ====> (D) ====> (F) ====
```
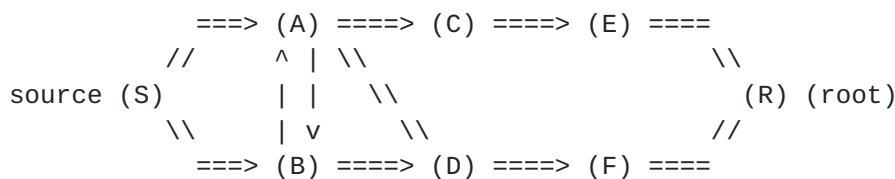
Figure 3: Unicast to DP with Overhearing: by employing Promiscuous
Overhearing, DP, AP and the Sibling nodes have more opportunities to
receive the same data packet.

## 5.  Requirements

## 5.1.  Requirements Related to Alternative Parent Selection

To perform the 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.

Related requirements are:

Req1.1: The routing protocol SHOULD be extended to allow for each
node to select AP(s) in addition to DP.  Thus, packet duplication
(i.e., replication) to multiple parents could be possible.

Req1.2: Considering that the Replication procedure significantly
increases the traffic in a network, when proposing solutions for
Alternative Parent Selection, it 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

To select an Alternative Parent, nodes MUST be aware of their
grandparent node sets.  Thus, it is necessary nodes to propagate such
information to their neighbors.  RPL [RFC6550] defines DODAG
Information Object (DIO) Control Message to allow nodes to propagate
information about themselves to potential children.  In Figure 4, DIO
control message with a DAG Metric Container option is illustrated.
However, RPL [RFC6550], does not indicates how to propagate parent
set related information.

```
 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 4: Example DIO Message with a DAG Metric Container option

Related requirements are:

Req2.1: DIO control messages can include multiple options.  DAG
Metric Container option [RFC6551] is structurally suitable for
transferring parent node set information.  Therefore, to enable PRE,
nodes MUST broadcast their parent node set to their potential
children through the extended DIO control message.  For instance,
"RPL DAG Metric Container (MC) Node State and Attribute (NSA) object
type extension" [I-D.koutsiamanis-roll-nsa-extension] focuses on
extending the DAG Metric Container [RFC6551] by defining new type-
length-value (TLV), entitled Parent Node Set (PNS) which CAN be
carried in the Node State and Attribute (NSA) object.

## 5.3.  Requirements Related to Cell Reservation

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.

Related requirements are:

Req3.1: The destination address filtering is performed at the MAC
layer.  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.  Thus, IEEE std. 802.15.4
implementation SHOULD bypass this filtering either by configuration
forcing to accept such the receiving frame or by using anycast/
multicast address as destination.

Req3.2: The 6top Protocol [I-D.ietf-6tisch-6top-protocol] SHOULD be
extended to possibly allow a cell reservation with two receivers,
i.e., DP and AP.  Considering that each frame may be transmitted
twice in unicast to each parent, then depending the transmission,
either DP will acknowledge the frame or AP will.

Req3.3: Next, to request the overhearing cells, the 6P ADD Request
Format SHOULD be transmitted either twice to each parent, i.e., DP
and AP, or once in multicast to both parents.  This procedure SHOULD
be considered in 6top Protocol [I-D.ietf-6tisch-6top-protocol]
specification.

## 5.4.  Requirements Related to Cells without ACKs

As stated in BCP 210 [RFC8180], each date frame is acknowledged by
the receiving node.  However, by employing promiscuous overhearing
operation, particular attention should be given to who will
acknowledge a transmission, i.e., the DP, and / or one of the AP(s)

Related requirements are:

Req4.1: To avoid the ACK collision, the TSCH Schedule as per BCP 210
[RFC8180], only the destination node of a packet MUST acknowledge the
data packet.

Req4.2: The overhearing node can be configured with the timeslot set
to shared, thus, there will be no acknowledgement from it.  However,
there is the security issue that needs to be considered.  Since, the
overhearing case imply that it is not possible to have per-pair
keying, thus, there MUST be a key that the overhearing node will be
aware of.  Hence, Minimal Security Framework for 6TiSCH
[I-D.ietf-6tisch-architecture] specification should consider such
scenario.

Req4.3: Optionally, to achieve further consistency the overheard
transmission need be acknowledged by both parents, i.e., DP and AP.
To do so, MAC layer operation MUST be extended accordingly.

## 5.5.  Requirements Related to Packet Elimination

By employing packet replication operation, the wireless network
expects to perform the packet elimination operation along a complex
Track to bound the number of the duplicated packets, i.e., the
unnecessary traffic.

Related requirements are:

Req5.1: As per 6TiSCH Architecture [I-D.ietf-6tisch-architecture],
6TiSCH has no position about how the sequence numbers would be tagged
in the packet.  However, it comes with Tagging Packets for Flow
Identification.  More specifically, a wireless network expects that
timeslots corresponding to copies of a same frame along a complex
Track are correlated by configuration and, thus, does not need to
process the sequence numbers.

## 6.  Security Considerations

TODO.

## 7.  IANA Considerations

   This document has no IANA considerations.

## 8.  References

### 8.1.  Informative references

   [I-D.ietf-6tisch-6top-protocol]
             Wang, Q., Vilajosana, X., and T. Watteyne, "6TiSCH
             Operation Sublayer Protocol (6P)", draft-ietf-6tisch-6top-
             protocol-12 (work in progress), June 2018.

   [I-D.ietf-6tisch-architecture]
             Thubert, P., "An Architecture for IPv6 over the TSCH mode
             of IEEE 802.15.4", draft-ietf-6tisch-architecture-19 (work
             in progress), December 2018.

   [I-D.ietf-detnet-architecture]
             Finn, N., Thubert, P., Varga, B., and J. Farkas,
             "Deterministic Networking Architecture", draft-ietf-
             detnet-architecture-10 (work in progress), December 2018.

   [I-D.koutsiamanis-roll-nsa-extension]
             Koutsiamanis, R., Papadopoulos, G., Montavont, N., and P.
             Thubert, "RPL DAG Metric Container Node State and
             Attribute object type extension", draft-koutsiamanis-roll-
             nsa-extension-04 (work in progress), November 2018.

   [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>.

**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


   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