Authors: CJ. Bernardos   A. Mourad
         UC3M   InterDigital

RAW multidomain extensions

Abstract

   This document describes the multi-domain RAW problem and explores
   and proposes some extensions to enable RAW multi-domain operation.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction and Problem Statement

Wireless operates on a shared medium, and transmissions cannot be
fully deterministic due to uncontrolled interferences, including
self-induced multipath fading. RAW (Reliable and Available Wireless)
is an effort to provide Deterministic Networking on across a path
that include a wireless interface. RAW provides for high reliability
and availability for IP connectivity over a wireless medium. The
wireless medium presents significant challenges to achieve
deterministic properties such as low packet error rate, bounded
consecutive losses, and bounded latency. RAW extends the DetNet
Working Group concepts to provide for high reliability and
availability for an IP network utilizing scheduled wireless segments
and other media, e.g., frequency/time-sharing physical media
resources with stochastic traffic: IEEE Std. 802.15.4 timeslotted
channel hopping (TSCH), 3GPP 5G ultra-reliable low latency
communications (URLLC), IEEE 802.11ax/be, and L-band Digital
Aeronautical Communications System (LDACS), etc. Similar to DetNet,
RAW technologies aim at staying abstract to the radio layers
underneath, addressing the Layer 3 aspects in support of
applications requiring high reliability and availability.

As introduced in [I-D.ietf-raw-architecture], RAW separates the path
computation time scale at which a complex path is recomputed from
the path selection time scale at which the forwarding decision is
taken for one or a few packets. RAW operates at the path selection
time scale. The RAW problem is to decide, amongst the redundant
solutions that are proposed by the Patch Computation Element (PCE),
which one will be used for each packet to provide a Reliable and
Available service while minimizing the waste of constrained
resources. To that effect, RAW defines the Path Selection Engine
(PSE) that is the counter-part of the PCE to perform rapid local
adjustments of the forwarding tables within the diversity that the
PCE has selected for the Track. The PSE enables to exploit the
richer forwarding capabilities with Packet (hybrid) ARQ,
Replication, Elimination and Ordering (PAREO), and scheduled
transmissions at a faster time scale.

There are several use cases [I-D.ietf-raw-use-cases] where reliability and availability are key requirements for wireless heterogeneous networks. A couple of relevant examples are (i) the manufacturing sector, where a plethora of devices are interconnected and generate data that need to be reliably delivered to the control and monitoring agents; and (ii) the residential gaming, with eXtended Reality (XR).

We can refer to domains managed by a single PCE, as "single-domain RAW", where nodes are typically run and managed by a single administration entity. In this scenario, the PSE can make use of "tracks" and paths involving only the nodes belonging to this RAW domain.

There are scenarios where hosts are connected to different RAW domains and they need to communicate to each other with certain reliability and/or availability guarantees, for example in large factories where networks might be organized in domains (per production lines or building/sites), in residential environments where there are different networks (e.g., one at home and one in the garden), or even vehicular scenarios (e.g., hosts connected to different vehicles).

```
 _____
|                        |
|              ( ( o ) )  |
|               *    ^    |
|              *    / \   |
|             *    /   \  |
|            **   ------+--    |
|           *    | RAW |P|    |
| ( ( o ) )*       *      |node |S|    |
|    ^      *( ( o ) )  | 1-1 |E|    +------+
|   / \        ^    *  ------+--    | PCE1 |
|  /   \      / \    **              +------+
| +-----+    /   \    *( ( o ) )  |
| |host1|   ------+--       ^       |
| |     |   | RAW |P|      / \      |
| |     |   |node |S|     /   \     |
| | o   |   | 1-2 |E|    ------+--   |
| +-----+   ------+--   | RAW |P|   |
|            \   /      |node |S|   |
|     RAW     \ /       | 1-3 |E|   |
|    domain 1  v        ------+--   |
|          ( ( o ) )                |
|            ****                   |
|_____*_____****_____|
 _____*_____*****__
|            *              **       |
|           *          ****( ( o ) ) |
|          *       ****      ^        |
|     ( ( o ) )****        / \       |
|        ^       *        /   \      |
|       / \       *      ------+--    |
|      /   \       *     | RAW |P|    |
|    ------+--      *    |node |S|    |
|    | RAW |P|      *    | 2-1 |E|    +------+
|    |node |S|     *  ------+--    | PCE2 |
|    | 2-2 |E|     *               +------+
|    ------+--     *( ( o ) )      |
|               *      ^           |
|           ( ( o ) )  / \         |
|             ^       /   \        |
|            / \     ------+--      |
|           /   \    | RAW |P|      |
|         +-----+    |node |S|      |
|    RAW  |host2|    | 2-3 |E|      |
|   domain 2  +-----+    ------+--      |
|_____|
```
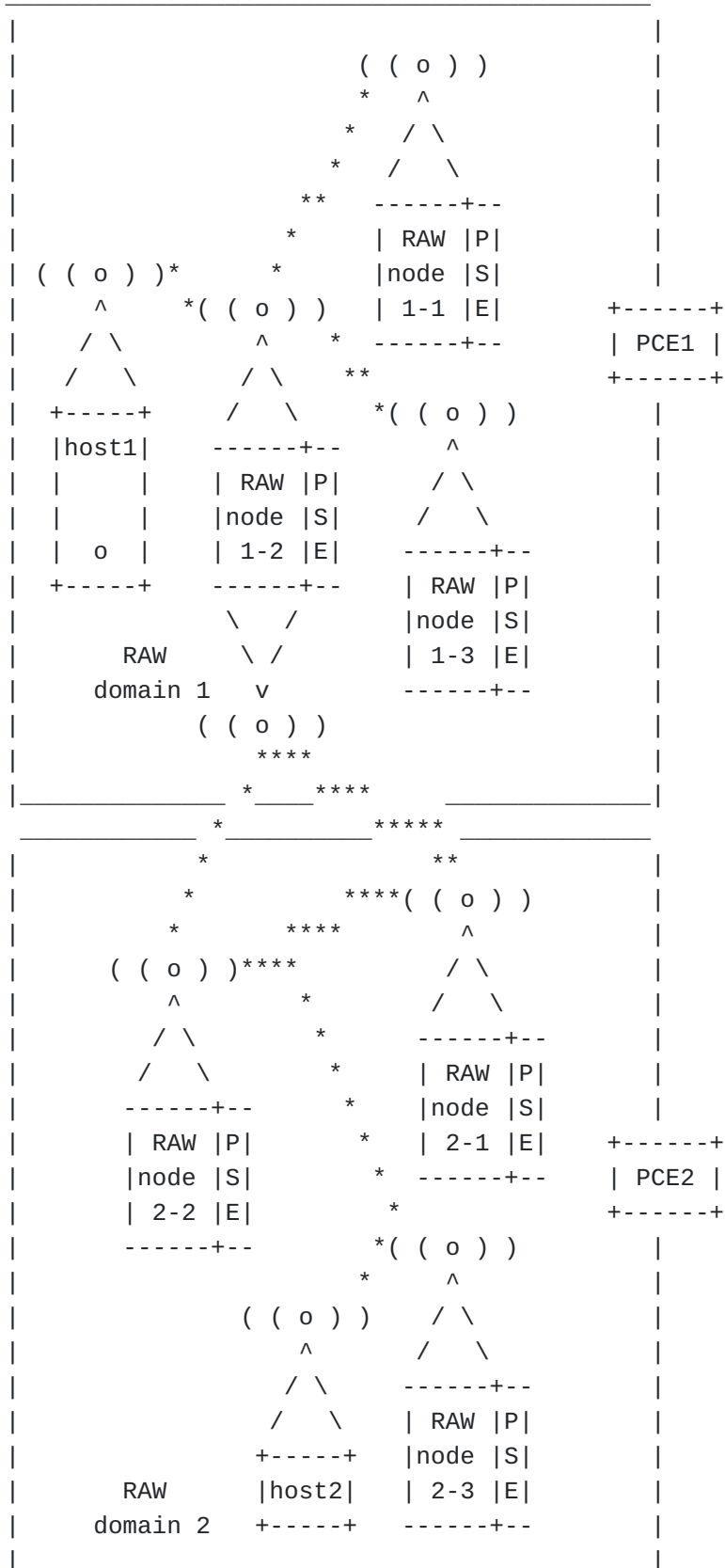
                Figure 1: Exemplary scenario showing multiple RAW domains

[Figure 1](#) shows an example of communication involving two RAW domains. As opposed to a single-domain scenario, where a single PCE may compute all possible "tracks" at longer time scale, and the PSE functionality may perform "subtrack" selection and optimization at a shorter time scale using all information available at the domain, multidomain scenarios pose additional burdens that are not solved yet.

Each RAW domain operates independently of the other domains. While there exist inter-PCE solutions today, allowing one domain's PCE to learn some inter-domain paths, this would not be sufficient, as the PSE of one domain would not have full visibility nor capability to act on the other domains (e.g., there are no multi-domain OAM solutions in place yet), limiting its capability to guarantee any given SLA. Therefore, there is a need to define inter-PSE coordination mechanisms across domains.

There exist today standardized solutions, such as the ones in the context of Path Computation Element (PCE), enabling computing multi-/inter-domain paths. As an example, the Hierarchical PCE (G-PCE) was defined in RFC 6805 [[RFC6805](#)] and is described hereafter. A parent PCE maintains a domain topology map that contains the child domains (seen as vertices in the topology) and their interconnections (links in the topology). The parent PCE has no information about the content of the child domains; that is, the parent PCE does not know about the resource availability within the child domains, nor does it know about the availability of connectivity across each domain because such knowledge would violate the confidentiality requirement and either would require flooding of full information to the parent (scaling issue) or would necessitate some form of aggregation. The parent PCE is used to compute a multi-domain path based on the domain connectivity information. A child PCE may be responsible for single or multiple domains and is used to compute the intra-domain path based on its own domain topology information.

Solutions like the above are not sufficient alone to solve the multi-domain RAW problem, as the PSEs need to have some additional information from the other involved domains to be sensitive/reactive to transient changes, in order to ensure a certain level of reliability and availability in a multi-domain wireless heterogeneous mesh network.

Within a single domain, the RAW framework architecture works, by having the PCE in charge of computing the paths (tracks) and the PSE(s) taking the short time decisions of which sub-tracks to use. Note that the PSE is assumed to be either a distributed functionality (performed by every RAW router of the path, which takes forwarding decisions based on the local and OAM information

that they have), or a centralized functionality played by the entry
(ingress) router in the domain (note that if there are multiple
ingress nodes, then there might be multiple PSEs), which then
performs source routing.

In scenarios with multiple connected RAW domains, running
uncoordinated RAW solutions in each domain is not sufficient. PSEs
would need to have global end-to-end information as well as be
capable of running OAM mechanisms [I-D.ietf-raw-oam-support] to
monitor the quality of the selected paths.

2. Terminology

The following terms used in this document are defined by the IETF:

   PAREO. Packet (hybrid) ARQ, Replication, Elimination and
   Ordering. PAREO is a superset Of DetNet's PREOF that includes
   radio-specific techniques such as short range broadcast, MUMIMO,
   constructive interference and overhearing, which can be leveraged
   separately or combined to increase the reliability.

   PSE. The Path Selection Engine (PSE) is the counter-part of the
   PCE to perform rapid local adjustments of the forwarding tables
   within the diversity that the PCE has selected for the Track. The
   PSE enables to exploit the richer forwarding capabilities with
   PAREO and scheduled transmissions at a faster time scale over the
   smaller domain that is the Track, in either a loose or a strict
   fashion.

3. RAW multi-domain extensions

Here we specify the new mechanisms and signaling extensions to
enable inter-domain RAW connectivity.

```
+-----+-+       +----+         +----+        +-----+-+       +-----+-+
| RAW |P|       |    |         |    |        | RAW |P|       | RAW |P|
|node |S|       |PCE1|         |PCE2|        |node |S|       |node |S|
| 1-2 |E|       |    |         |    |        | 2-1 |E|       | 2-2 |E|
+-----+-+       +----+         +----+        +-----+-+       +-----+-+
    |               |              |              |              |
1.Path compute req  |              |              |              |
(src=node1-2,       |              |              |              |
 dst=node2-3, SLA)  |              |              |              |
     |..........>|              |              |              |
         |        |2.Path compute req          |              |
         |        |(src={node2-1,node2-2},     |              |
         |        | dst=node2-3)               |              |
         |        |..........>|                |              |
         |        |3.Path compute resp         |              |
         |        |({tracks2},{links_quality})|              |
         |        |<..........|                |              |
4.Path compute resp           |              |              |
({{tracks1},{tracks2}},       |              |              |
 PSE={node2-1,node2-2},       |              |              |
 {SLA1,SLA2})       |         |              |              |
     |<..........|            |              |              |
     |5.RAW inter-domain path |              |              |
     |({{tracks1,tracks2}},{SLA1,SLA2})      |              |
     |....................................>|              |
     |.......................................................>|
     |          |           6.RAW inter-domain path ACK        |
     |<...................................|              |
     |<.......................................................|
     |          |              |              |              |
     |7.RAW OAM(flow/track,SLA1)            |              |
 <...>|<...>       |              | 7.RAW OAM(flow/track,SLA1)|
     |          |              |        <..>|<..>    <..>|<..>
     |            8.RAW OAM (flow/track, metrics)           |
     |<...................................>|              |
     |<.......................................................>|
     |          |              |              |              |
```
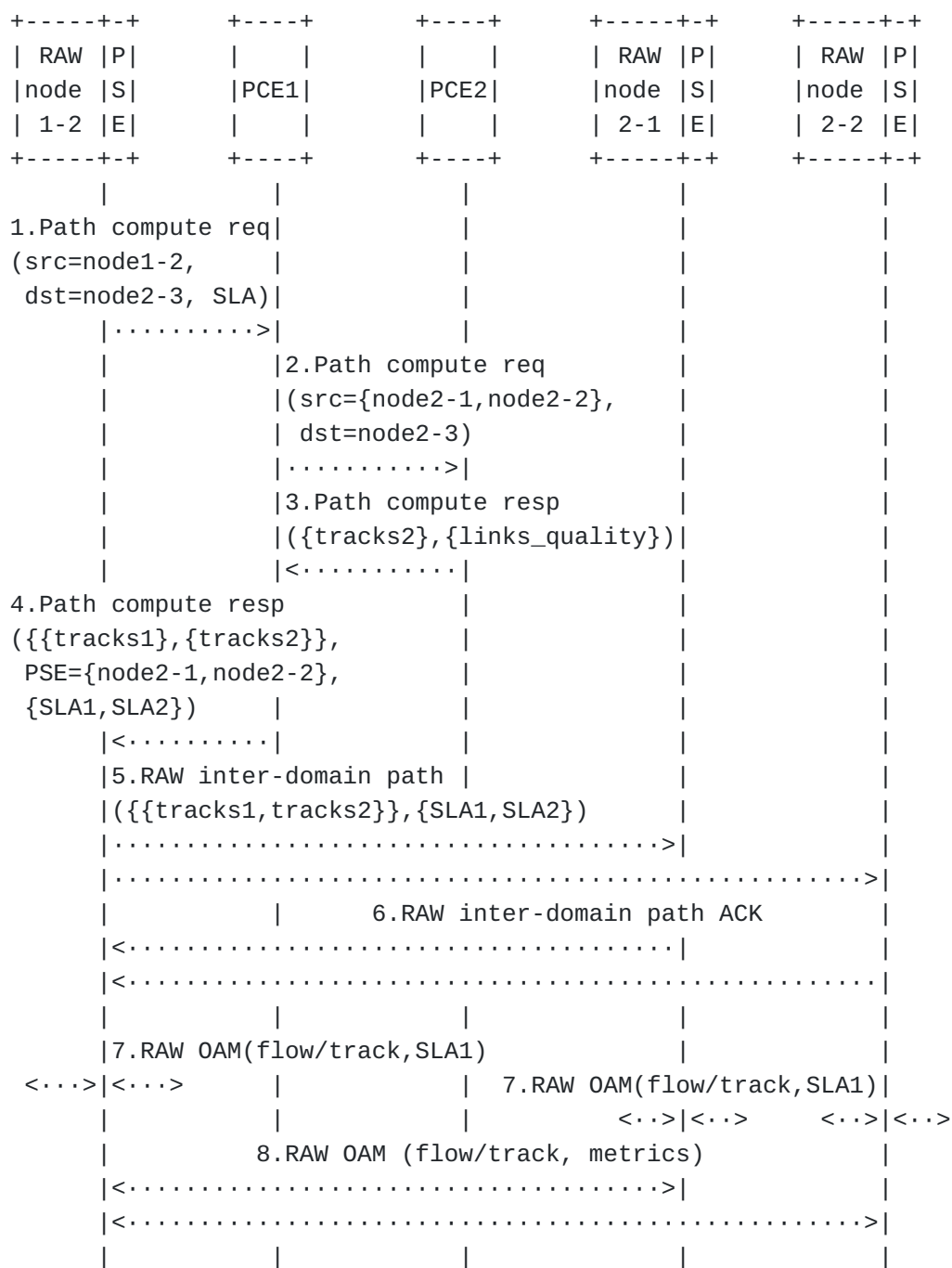
           Figure 2: Multi-domain RAW signaling

   [Figure 2](#) shows a signaling flow diagram, taking as baseline scenario
   the one shown in [Figure 1](#), where host1 (connected to node1-2) wants
   to communicate with host2 (connected to node2-3). An ingress RAW
   node (node1-2) gets a request for connectivity, with a given
   destination RAW node (node2-3) and the desired SLA in terms of
   reliability and availability. The source and/or destination RAW

nodes might be hostss. We next explain each of the steps illustrated
in the figure:

1. The ingress node plays the role of PSE (also referred to as
    PSE@domain1) and requests the computation of the tracks towards
    the destination node2-3 with the intended SLA to the PCE of the
    domain (PCE1).

2. PCE1 knows that the destination is in another domain (domain2)
    and that the PCE of the destination domain is PCE2. PCE1 also
    knows the ingress nodes in domain2 that are connected to
    domain1. How this is done is outside of the scope of this
    document. These nodes (node2-1 and node2-2) play the role of
    PSEs@domain2. PCE1 requests to PCE2 to compute the available
    tracks from PSEs@domain2 to the destination, and the
    characteristics of the links (link_quality) forming these
    tracks. The detail and nature of the information provided by
    PCE2 regarding the links might vary depending on the
    deployment, and is meant to be used by PCE1 and the PSE@domain1
    (node1-2) to compute how to distribute the SLA among the
    domains.

3. PCE2 computes the tracks and responds to PCE1, including also
    the characteristics of the links (link_quality). Examples of
    potential information elements including in the link_quality
    are: available bandwidth, observed reliability, delay, link
    variability/mobility, etc.

4. PCE1 provides to the PSE@domain1 the tracks to reach the
    destination, as well as the split of SLAs among domain1 and
    domain2 (SLA1 and SLA2). An SLA, or a Quality of Service (QoS)
    figure, may include aspects such as, among others: max. delay,
    assured BW, max. Jitter, packet loss ratio, availability ratio,
    etc. PCE1 also provides the PSEs@domain2.

5. The PSE@domain1 sends a message to each PSE@domain2, in order
    to set-up a direct communication channel to provide OAM
    information useful to the PSE@domain1 for computing the
    subtracks to use for the traffic. This message includes the SLA
    that each domain has to monitor and guarantee (SLA1 and SLA2).

6. Each of the PSEs@domain2 acknowledges the message. At this
    point, the communication channel is established and the
    PSE@domain1 can start taking decisions at a forwarding time
    scale regarding which paths (subtracks) to use.

7. All PSEs, at each domain, start performing OAM procedures [I-
    D.ietf-raw-oam-support], which are key to observe if traffic is
    meeting the desired SLAs (SLA1 and SLA2) and adapt the

subtracks and tracks if needed. OAM mechanisms can be applied
in-band (sharing the traffic's fate) or out-of band. Note that
this per-domain distributed OAM is critical to ensure that the
required SLAs (reliability and availability) are met by
reacting on a short time scale at each of the involved domains.

8. PSEs share aggregated and pre-processed information among them
   to facilitate early detection of issues and computation of
   subtracks. If a violation of an SLA is detected, the respective
   PSE would notify the domain PCE and the other PSE, so a
   reaction measure can be taken (e.g., selecting different
   subtracks, taking different PAREO decisions, requesting the
   PCEs to recompute the paths and/or adjust the split of the SLAs
   across the domains).

Note that this example covers the direction host1-to-host2. If there
is traffic in the opposite direction, the process has to be repeated
in the reverse direction, as paths might not be bidirectional.

4. IANA Considerations

TBD.

5. Security Considerations

TBD.

6. Acknowledgments

7. Informative References

[I-D.ietf-raw-architecture] Thubert, P. and G. Z. Papadopoulos,
           "Reliable and Available Wireless Architecture", Work in
           Progress, Internet-Draft, draft-ietf-raw-architecture-03,
           14 January 2022, <https://www.ietf.org/archive/id/draft-
           ietf-raw-architecture-03.txt>.

[I-D.ietf-raw-oam-support] Theoleyre, F., Papadopoulos, G. Z.,
           Mirsky, G., and C. J. Bernardos, "Operations,
           Administration and Maintenance (OAM) features for RAW",
           Work in Progress, Internet-Draft, draft-ietf-raw-oam-
           support-04, 5 March 2022, <https://www.ietf.org/archive/
           id/draft-ietf-raw-oam-support-04.txt>.

[I-D.ietf-raw-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://www.ietf.org/archive/id/draft-
ietf-raw-use-cases-05.txt>.

[RFC6805] King, D., Ed. and A. Farrel, Ed., "The Application of the
Path Computation Element Architecture to the
Determination of a Sequence of Domains in MPLS and
GMPLS", RFC 6805, DOI 10.17487/RFC6805, November 2012,
<https://www.rfc-editor.org/info/rfc6805>.

Authors' Addresses

Carlos J. Bernardos
Universidad Carlos III de Madrid
Av. Universidad, 30
28911 Leganes, Madrid
Spain

Phone: +34 91624 6236
Email: cjbc@it.uc3m.es
URI: http://www.it.uc3m.es/cjbc/

Alain Mourad
InterDigital Europe

Email: Alain.Mourad@InterDigital.com
URI: http://www.InterDigital.com/