RAW   F. Theoleyre
Internet-Draft   CNRS
Intended status: Standards Track   G. Papadopoulos
Expires: January 11, 2021   IMT Atlantique
                                                               G. Mirsky
                                                               ZTE Corp.
                                                           July 10, 2020

      Operations, Administration and Maintenance (OAM) features for RAW
                    draft-theoleyre-raw-oam-support-03

Abstract

   Some critical applications may use a wireless infrastructure.
   However, wireless networks exhibit a bandwidth of several orders of
   magnitude lower than wired networks. Besides, wireless transmissions
   are lossy by nature; the probability that a packet cannot be decoded
   correctly by the receiver may be quite high. In these conditions,
   guaranteeing the network infrastructure works properly is
   particularly challenging, since we need to address some issues
   specific to wireless networks. This document lists the requirements
   of the Operation, Administration, and Maintenance (OAM) features
   recommended to construct a predictable communication infrastructure
   on top of a collection of wireless segments. This document describes
   the benefits, problems, and trade-offs for using OAM in wireless
   networks to achieve Service Level Objectives (SLO).

Table of Contents

1. Introduction

   Reliable and Available Wireless (RAW) is an effort that extends
   DetNet to approach end-to-end deterministic performances over a
   network that includes scheduled wireless segments. In wired
   networks, many approaches try to enable Quality of Service (QoS) by
   implementing traffic differentiation so that routers handle each type
   of packets differently. However, this differentiated treatment was
   expensive for most applications.

   Deterministic Networking (DetNet) [RFC8655] has proposed to provide a
   bounded end-to-end latency on top of the network infrastructure,
   comprising both Layer 2 bridged and Layer 3 routed segments. Their
   work encompasses the data plane, OAM, time synchronization,
   management, control, and security aspects.

   However, wireless networks create specific challenges. First of all,
   radio bandwidth is significantly lower than for wired networks. In
   these conditions, the volume of signaling messages has to be very
   limited. Even worse, wireless links are lossy: a layer 2
   transmission may or may not be decoded correctly by the receiver,
   depending on a broad set of parameters. Thus, providing high
   reliability through wireless segments is particularly challenging.

   Wired networks rely on the concept of _links_. All the devices
   attached to a link receive any transmission. The concept of a link
   in wireless networks is somewhat different from what many are used to
   in wireline networks. A receiver may or may not receive a
   transmission, depending on the presence of a colliding transmission,
   the radio channel's quality, and the external interference. Besides,
   a wireless transmission is broadcast by nature: any _neighboring_
   device may be able to decode it. The document includes detailed
   information on what the implications for the OAM features are.

   Last but not least, radio links present volatile characteristics. If
   the wireless networks use an unlicensed band, packet losses are not
   anymore temporally and spatially independent. Typically, links may
   exhibit a very bursty characteristic, where several consecutive
   packets may be dropped. Thus, providing availability and reliability
   on top of the wireless infrastructure requires specific Layer 3
   mechanisms to counteract these bursty losses.

   Operations, Administration, and Maintenance (OAM) Tools are of
   primary importance for IP networks [RFC7276]. It defines a toolset
   for fault detection, isolation, and performance measurement.

   The primary purpose of this document is to detail the specific
   requirements of the OAM features recommended to construct a

predictable communication infrastructure on top of a collection of wireless segments. This document describes the benefits, problems, and trade-offs for using OAM in wireless networks to provide availability and predictability.

In this document, the term OAM will be used according to its definition specified in [RFC6291]. We expect to implement an OAM framework in RAW networks to maintain a real-time view of the network infrastructure, and its ability to respect the Service Level Objectives (SLO), such as delay and reliability, assigned to each data flow.

1.1. Terminology

o OAM entity: a data flow to be controlled;

o Maintenance End Point (MEP): OAM devices crossed when entering/ exiting the network. In RAW, it corresponds mostly to the source or destination of a data flow. OAM message can be exchanges between two MEPs;

o Maintenance Intermediate endPoint (MIP): OAM devices along the flow; OAM messages can be exchanged between a MEP and a MIP;

o Defect: a temporary change in the network (e.g., a radio link which is broken due to a mobile obstacle);

o Fault: a definite change which may affect the network performance, e.g., a node runs out of energy.

1.2. Acronyms

OAM Operations, Administration, and Maintenance

DetNet Deterministic Networking

SLO Service Level Objective

QoS Quality of Service

SNMP Simple Network Management Protocol

SDN Software-Defined Network

1.3. Requirements Language

   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.

2. Role of OAM in RAW

   RAW networks expect to make the communications reliable and
   predictable on top of a wireless network infrastructure. Most
   critical applications will define an SLO to be required for the data
   flows it generates. RAW considers network plane protocol elements
   such as OAM to improve the RAW operation at the service and the
   forwarding sub-layers.

   To respect strict guarantees, RAW relies on an orchestrator able to
   monitor and maintain the network. Typically, a Software-Defined
   Network (SDN) controller is in charge of scheduling the transmissions
   in the deployed network, based on the radio link characteristics, SLO
   of the flows, the number of packets to forward. Thus, resources have
   to be provisioned a priori to handle any defect. OAM represents the
   core of the pre-provisioning process and maintains the network
   operational by updating the schedule dynamically.

   Fault-tolerance also assumes that multiple paths have to be
   provisioned so that an end-to-end circuit keeps on existing whatever
   the conditions. The Packet Replication and Elimination Function
   ([PREF-draft]) on a node is typically controlled by a central
   controller/orchestrator. OAM mechanisms can be used to monitor that
   PREOF is working correctly on a node and within the domain.

   To be energy-efficient, reserving some dedicated out-of-band
   resources for OAM seems idealistic, and only in-band solutions are
   considered here.

   RAW supports both proactive and on-demand troubleshooting.

   The specific characteristics of RAW are discussed below.

2.1. Link concept and quality

   In wireless networks, a _link_ does not exist. A common convention
   is to define a wireless link as a pair of devices that have a non-
   null probability of transmitting and decoding a packet. Similarly,
   we designate as *neighbor* any device which as a link with a specific
   transmitter.

Each wireless link is associated with a link quality, often measured
as the Packet Delivery Ratio (PDR), i.e., the probability that the
receiver can decode the packet correctly. It is worth noting that
this link quality depends on many criteria, such as the level of
external interference, the presence of concurrent transmissions, or
the radio channel state. This link quality is even time-variant.

2.2. Broadcast Transmissions

In modern switching networks, the unicast transmission is delivered
uniquely to the destination. Wireless networks are much closer to
the ancient shared access wireless networks. Unicast transmission is
similar to a broadcast one and can be received by any neighbor.

However, contrary to wired networks, we cannot be sure that a packet
is received by *all* the devices attached to the network. It depends
on the radio channel state between the transmitter(s) and the
receiver(s). In particular, concurrent transmissions may be possible
or not, depending on the radio conditions.

2.3. Complex Layer 2 Forwarding

Multiple neighbors may receive a transmission. Thus, anycast layer-2
forwarding helps to maximize the reliability by assigning multiple
receivers to a single transmission. That way, the packet is lost
only if none of the receivers decode it. Practically, it has been
proven that different neighbors may exhibit very different radio
conditions, and that reception independency may hold for some of them
[anycast-property].

3. Operation

OAM features will enable RAW with robust operation both for
forwarding and routing purposes.

3.1. Information Collection

Several solutions (e.g., Simple Network Management Protocol (SNMP),
YANG-based data models) are already in charge of collecting the
statistics. That way, we can encapsulate these statistics in
specific monitoring packets, to send them to the controller.

3.2. Continuity Check

We need to verify that two endpoints are connected. In other words,
there exists "one" way to deliver the packets between two endpoints A
and B. The solution may not here defer from those of detnet.

3.3. Connectivity Verification

   Additionally, to the Continuity Check, we have to verify the
   connectivity. This verification considers additional constraints,
   i.e., the absence of misconnection.

   In particular, the resources have to be reserved by a given flow, and
   no packets from other flows steal the corresponding resources.
   Similarly, the destination does not receive packets from different
   flows through its interface.

   Because of radio transmissions' broadcast nature, several receivers
   may be active at the same time to enable anycast Layer 2 forwarding.
   Thus, the connectivity verification must test any combination. We
   also consider priority-based mechanisms for anycast forwarding, i.e.,
   all the receivers have different probabilities of forwarding a
   packet. To verify a delay SLO for a given flow, we must also
   consider all the possible combinations, leading to a probability
   distribution function for end-to-end transmissions. If this
   verification is implemented naively, the number of combinations to
   test may be exponential and too costly for wireless networks with low
   bandwidth.

   It is worth noting that the control and data packets may not follow
   the same path. The connectivity verification has to be conducted in-
   band without impacting the data traffic. Test packets MUST share the
   fate with the monitored data traffic without introducing congestion
   in normal network conditions.

3.4. Route Tracing

   ICMP tools are comprehensive tools for diagnostic. They help to
   identify a subset of the list of routers in the route. To ensure
   predictable performance, resources are reserved per flow in RAW.
   Thus, we need to define route tracing tools able to track the route
   for a specific flow.

   Wireless networks are meshed by nature: we have many redundant radio
   links. These meshed networks are both an asset and a drawback: while
   several paths exist between two endpoints, and we should choose the
   most efficient one(s), concerning specifically the reliability, and
   the delay.

   Thus, multipath routing can be considered to make the network fault-
   tolerant. Even better, we can exploit the broadcast nature of
   wireless networks to exploit meshed multipath routing: we may have
   multiple Maintenance Intermediate Endpoints (MIE) for each hop in the
   path. In that way, each Maintenance Intermediate Endpoint has

several possible next hops in the forwarding plane. Thus, all the
possible paths between two maintenance endpoints should be retrieved,
which may quickly become untractable if we apply a naive approach.

3.5. Fault Verification/detection

RAW expects to operate fault-tolerant networks. Thus, we need
mechanisms able to detect faults, before they impact the network
performance.

Wired networks tend to present stable performances. On the contrary,
wireless networks are time-variant. We must consequently make a
distinction between _normal_ evolutions and malfunction.

The network has to detect when a fault occurred, i.e., the network
has deviated from its expected behavior. While the network must
report an alarm, the cause may not be identified precisely. For
instance, the end-to-end reliability has decreased significantly, or
a buffer overflow occurs.

3.6. Fault Isolation/identification

The network has isolated and identified the cause of the fault.
While detnet already expects to identify malfunctions, some problems
are specific to wireless networks. We must consequently collect
metrics and implement algorithms tailored for wireless networking.
For instance, the quality of a specific link has decreased, requiring
more retransmissions, or the level of external interference has
locally increased.

4. Administration

The network has to expose a collection of metrics to support an
operator making proper decisions, including:

o Packet losses: the time-window average and maximum values of the
   number of packet losses have to be measured. Many critical
   applications stop to work if a few consecutive packets are
   dropped;

o Received Signal Strength Indicator (RSSI) is a very common metric
   in wireless to denote the link quality. The radio chipset is in
   charge of translating a received signal strength into a normalized
   quality indicator;

o Delay: the time elapsed between a packet generation / enqueuing
   and its reception by the next hop;

o Buffer occupancy: the number of packets present in the buffer, for each of the existing flows.

These metrics should be collected:

o per virtual circuit to measure the end-to-end performance for a given flow. Each of the paths has to be isolated in multipath routing strategies;

o per radio channel to measure, e.g., the level of external interference, and to be able to apply counter-measures (e.g., blacklisting).

o per device to detect misbehaving node, when it relays the packets of several flows.

4.1. Collection of metrics

We have to minimize the number of statistics / measurements to exchange:

o energy efficiency: low-power devices have to limit the volume of monitoring information since every bit consumes energy.

o bandwidth: wireless networks exhibit a bandwidth significantly lower than wired, best-effort networks.

o per-packet cost: it is often more expensive to send several packets instead of combining them in a single link-layer frame.

Thus, localized and centralized mechanisms have to be combined together, and additional control packets have to be triggered only after a fault detection.

4.2. Worst-case metrics

RAW aims to enable real-time communications on top of a heterogeneous architecture. Wireless networks are known to be lossy, and RAW has to implement strategies to improve reliability on top of unreliable links. Hybrid Automatic Repeat reQuest (ARQ) has typically to enable retransmissions based on the end-to-end reliability and latency requirements.

To make correct decisions, the controller needs to know the distribution of packet losses for each flow, and each hop of the paths. In other words, the average end-to-end statistics are not enough. They must allow the controller to predict the worst-case.

4.3. Energy efficiency constraint

   RAW targets also low-power wireless networks, where energy represents
   a key constraint. Thus, we have to take care of power and bandwidth
   consumption. The following techniques aim to reduce the cost of such
   maintenance:

      on-path collection: some control information is inserted in the
      data packets if they do not fragment the packet (i.e., the MTU is
      not exceeded). Information Elements represent a standardized way
      to handle such information;

      flags/fields: we have to set-up flags in the packets to monitor to
      be able to monitor the forwarding process accurately. A sequence
      number field may help to detect packet losses. Similarly, path
      inference tools such as [ipath] insert additional information in
      the headers to identify the path followed by a packet a
      posteriori.

5. Maintenance

   RAW needs to implement a self-healing and self-optimization approach.
   The network must continuously retrieve the state of the network, to
   judge about the relevance of a reconfiguration, quantifying:

      the cost of the sub-optimality: resources may not be used
      optimally (e.g., a better path exists);

      the reconfiguration cost: the controller needs to trigger some
      reconfigurations. For this transient period, resources may be
      twice reserved, and control packets have to be transmitted.

   Thus, reconfiguration may only be triggered if the gain is
   significant.

5.1. Replication / Elimination

   When multiple paths are reserved between two maintenance endpoints,
   they may decide to replicate the packets to introduce redundancy, and
   thus to alleviate transmission errors and collisions. For instance,
   in Figure 1, the source node S is transmitting the packet to both
   parents, nodes A and B. Each maintenance endpoint will decide to
   trigger the replication/elimination process when a set of metrics
   passes through a threshold value.

```
                   ===> (A) => (C) => (E) ===
                 //        \\//   \\//        \\
         source (S)        //\\   //\\          (R) (root)
                 \\       // \\ // \\      //
                   ===> (B) => (D) => (F) ===
```
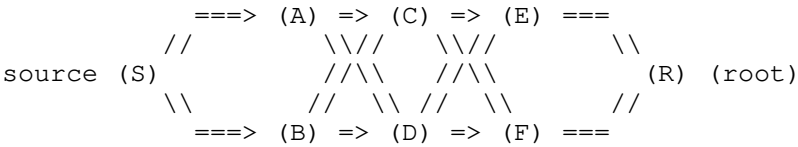
   Figure 1: Packet Replication: S transmits twice the same data packet,
                    to its DP (A) and to its AP (B).

5.2. Dynamic Resource Reservation

   Wireless networks exhibit time-variant characteristics. Thus, the
   network has to provide additional resources along the path to fit the
   worst-case performance. This time-variant characteristics make the
   resource reservation very challenging: over-reaction waste radio and
   energy resources. Inversely, under-reaction jeopardize the network
   operations, and some SLO may be violated.

5.3. Reliable Reconfiguration

   Wireless networks are known to be lossy. Thus, commands may be
   received or not by the node to reconfigure. Unfortunately,
   inconsistent states may create critical misconfigurations, where
   packets may be lost along a path because it has not been properly
   configured.

   We have to propose mechanisms to guarantee that the network state is
   always consistent, even if some control packets are lost. Timeouts
   and retransmissions are not sufficient since the reconfiguration
   duration would be, in that case, unbounded.

6. IANA Considerations

   This document has no actionable requirements for IANA. This section
   can be removed before the publication.

7. Security Considerations

   This section will be expanded in future versions of the draft.

8. Acknowledgments

   TBD

9. Informative References

[anycast-property]
           Teles Hermeto, R., Gallais, A., and F. Theoleyre, "Is
           Link-Layer Anycast Scheduling Relevant for IEEE
           802.15.4-TSCH Networks?", 2019,
           <https://doi.org/10.1109/LCNSymposium47956.2019.9000679>.

[ipath]   Gao, Y., Dong, W., Chen, C., Bu, J., Wu, W., and X. Liu,
           "iPath: path inference in wireless sensor networks.",
           2016, <https://doi.org/10.1109/TNET.2014.2371459>.

[PREF-draft]
           Thubert, P., Eckert, T., Brodard, Z., and H. Jiang, "BIER-
           TE extensions for Packet Replication and Elimination
           Function (PREF) and OAM", 2018,
           <https://tools.ietf.org/html/draft-thubert-bier-
           replication-elimination>.

[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>.

[RFC6291] Andersson, L., van Helvoort, H., Bonica, R., Romascanu,
           D., and S. Mansfield, "Guidelines for the Use of the "OAM"
           Acronym in the IETF", BCP 161, RFC 6291,
           DOI 10.17487/RFC6291, June 2011,
           <https://www.rfc-editor.org/info/rfc6291>.

[RFC7276] Mizrahi, T., Sprecher, N., Bellagamba, E., and Y.
           Weingarten, "An Overview of Operations, Administration,
           and Maintenance (OAM) Tools", RFC 7276,
           DOI 10.17487/RFC7276, June 2014,
           <https://www.rfc-editor.org/info/rfc7276>.

[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>.

[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>.

Authors' Addresses

   Fabrice Theoleyre
   CNRS
   Building B
   300 boulevard Sebastien Brant - CS 10413
   Illkirch - Strasbourg 67400
   FRANCE

   Phone: +33 368 85 45 33
   Email: theoleyre@unistra.fr
   URI: http://www.theoleyre.eu

   Georgios Z. Papadopoulos
   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

   Greg Mirsky
   ZTE Corp.

   Email: gregimirsky@gmail.com