               Deterministic Networking Problem Statement

Abstract

   This paper documents the needs in various industries to establish
   multi-hop paths for characterized flows with deterministic
   properties.

Status of This Memo

   This document is not an Internet Standards Track specification; it is
   published for informational purposes.

   This document is a product of the Internet Engineering Task Force
   (IETF).  It represents the consensus of the IETF community.  It has
   received public review and has been approved for publication by the
   Internet Engineering Steering Group (IESG).  Not all documents
   approved by the IESG are candidates for any level of Internet
   Standard; see Section 2 of RFC 7841.

   Information about the current status of this document, any errata,
   and how to provide feedback on it may be obtained at
   https://www.rfc-editor.org/info/rfc8557.

Table of Contents

1.  Introduction

   "Deterministic Networking Use Cases" [RFC8578] illustrates that
   beyond the classical case of Industrial Automation and Control
   Systems (IACSs) there are in fact multiple industries with strong,
   and relatively similar, needs for deterministic network services with
   latency guarantees and ultra-low packet loss.

   The generalization of the needs for more deterministic networks has
   led to the IEEE 802.1 Audio Video Bridging (AVB) Task Group becoming
   the Time-Sensitive Networking (TSN) [IEEE-802.1TSNTG] Task Group
   (TG), with a much-expanded constituency from the industrial and
   vehicular markets.

   Along with this expansion, the networks considered here are becoming
   larger and structured, requiring deterministic forwarding beyond the
   LAN boundaries.  For instance, an IACS segregates the network along
   the broad lines of the Purdue Enterprise Reference Architecture
   (PERA) [ISA95], typically using deterministic LANs for Purdue level 2
   control systems, whereas public infrastructures such as electricity
   automation require deterministic properties over the wide area.
   Implementers have come to realize that the convergence of IT and
   Operation Technology (OT) networks requires Layer 3, as well as
   Layer 2, capabilities.

   While the initial user base has focused almost entirely on Ethernet
   physical media and Ethernet-based bridging protocols from several
   Standards Development Organizations (SDOs), the need for Layer 3, as
   expressed above, must not be confined to Ethernet and Ethernet-like
   media.  While such media must be encompassed by any useful
   Deterministic Networking (DetNet) architecture, cooperation between
   the IETF and other SDOs must not be limited to the IEEE or the

IEEE 802 organizations.  Furthermore, while both completed and
ongoing work in other SDOs, and in IEEE 802 in particular, provides
an obvious starting point for a DetNet architecture, we must not
assume that these other SDOs' work confines the space in which the
DetNet architecture progresses.

The properties of deterministic networks will have specific
requirements for the use of routed networks to support these
applications, and a new model must be proposed to integrate this
determinism in IT implementations.  The proposed model should enable
a fully scheduled operation orchestrated by a central controller and
may support a more distributed operation with (probably lesser)
capabilities.  At any rate, the model should not compromise the
ability of a network to keep carrying the sorts of traffic that is
already carried today in conjunction with new, more deterministic
flows.  Note: "Deterministic Networking Architecture" [DetNet-Arch]
was produced by the DetNet Working Group to describe that model.

At the time of this writing, it is expected that

o  once the abstract model is agreed upon, the IETF will specify
   (1) the signaling elements to be used to establish a path and
   (2) the tagging elements to be used to identify the flows that are
   to be forwarded along that path

o  the IETF will specify the necessary protocols or protocol
   additions, based on relevant IETF technologies, to implement the
   selected model

A desirable outcome of the work is the ability to establish a
multi-hop path over the IP or MPLS network for a particular flow with
given timing and precise throughput requirements and to carry this
particular flow along the multi-hop path with such characteristics as
low latency and ultra-low jitter, reordering and/or replication and
elimination of packets over non-congruent paths for a higher delivery
ratio, and/or zero congestion loss, regardless of the amount of other
flows in the network.

Depending on the network capabilities and the current state, requests
to establish a path by an end node or a network management entity may
be granted or rejected, an existing path may be moved or removed, and
DetNet flows exceeding their contract may face packet
declassification and drop.

2.  On Deterministic Networking

   The Internet is not the only digital network that has grown
   dramatically over the last 30-40 years.  Video and audio
   entertainment, as well as control systems for machinery,
   manufacturing processes, and vehicles, are also ubiquitous and are
   now based almost entirely on digital technologies.  Over the past
   10 years, engineers in these fields have come to realize that
   significant advantages in both cost and the ability to accelerate
   growth can be obtained by basing all of these disparate digital
   technologies on packet networks.

   The goals of Deterministic Networking are to (1) enable the migration
   of applications with critical timing and reliability issues that
   currently use special-purpose fieldbus technologies (High-Definition
   Multimedia Interface (HDMI), Controller Area Network (CAN bus),
   PROFIBUS [PROFIBUS], etc. ... even RS-232!) to packet technologies in
   general and to IP in particular and (2) support both these new
   applications and existing packet network applications over the same
   physical network.  In other words, a deterministic network is
   backwards compatible with (capable of transporting) statistically
   multiplexed traffic while preserving the properties of the accepted
   deterministic flows.

   [RFC8578] indicates that applications in multiple fields need some or
   all of a suite of features that includes:

   1.  Time synchronization of all host and network nodes (routers
       and/or bridges), accurate to something between 10 nanoseconds and
       10 microseconds, depending on the application.

   2.  Support for deterministic packet flows that:

       *  Can be unicast or multicast.

       *  Need absolute guarantees of minimum and maximum latency
          end to end across the network; sometimes a tight jitter is
          required as well.

       *  Need a packet loss ratio beyond the classical range for a
          particular medium, in the range of $10^{-9}$ to $10^{-12}$ or better
          on Ethernet and on the order of $10^{-5}$ in wireless sensor mesh
          networks.

       *  Can, in total, absorb more than half of the network's
          available bandwidth (that is, massive over-provisioning is
          ruled out as a solution).

   *  Cannot suffer throttling, congestion feedback, or any other
      network-imposed transmission delay, although the flows can be
      meaningfully characterized by either (1) a fixed, repeating
      transmission schedule or (2) a maximum bandwidth and packet
      size.

3. Multiple methods for scheduling, shaping, limiting, and otherwise
   controlling the transmission of critical packets at each hop
   through the network data plane.

4. Robust defenses against misbehaving hosts, routers, or bridges,
   in both the data plane and the control plane, with guarantees
   that a critical flow within its guaranteed resources cannot be
   affected by other flows, whatever the pressures on the network.
   For more on the specific threats against DetNet, see
   "Deterministic Networking (DetNet) Security Considerations"
   [DetNet-Security].

5. One or more methods for reserving resources in bridges and
   routers to carry these flows.

Time-synchronization techniques need not be addressed by an IETF
working group; there are a number of standards available for this
purpose, including IEEE 1588 [IEEE-1588], IEEE 802.1AS [IEEE-8021AS],
and more.

The needs related to multicast, latency, loss ratio, and throttling
avoidance exist because the algorithms employed by the applications
demand it.  They are not simply the transliteration of fieldbus needs
to a packet-based fieldbus simulation; they also reflect fundamental
mathematics of the control of a physical system.

With classical forwarding of latency-sensitive and loss-sensitive
packets across a network, interactions among different critical flows
introduce fundamental uncertainties in delivery schedules.  The
details of the queuing, shaping, and scheduling algorithms employed
by each bridge or router to control the output sequence on a given
port affect the detailed makeup of the output stream, e.g., how
finely a given flow's packets are mixed among those of other flows.

This, in turn, has a strong effect on the buffer requirements, and
hence the latency guarantees deliverable, by the next bridge or
router along the path.  For this reason, the IEEE 802.1 TSN TG has
defined a new set of queuing, shaping, and scheduling algorithms that
enable each bridge or router to compute the exact number of buffers
to be allocated for each flow or class of flows.

Networking protocols commonly need robustness.  Note that robustness
plays a particularly important part in real-time control networks,
where expensive equipment, and even lives, can be lost due to
misbehaving equipment.

Reserving resources before packet transmission is the one fundamental
shift in the behavior of network applications that is impossible to
avoid.  In the first place, a network cannot deliver finite latency
and practically zero packet loss to an arbitrarily high offered load.
Secondly, achieving practically zero packet loss for unthrottled
(though bandwidth-limited) flows means that bridges and routers have
to dedicate buffer resources to specific flows or classes of flows.
The requirements of each reservation have to be translated into the
parameters that control each host's, bridge's, and router's queuing,
shaping, and scheduling functions and delivered to the hosts,
bridges, and routers.

## 3.  Problem Statement

### 3.1.  Supported Topologies

In some use cases, the end point that runs the application is
involved in the Deterministic Networking operation -- for instance,
by controlling certain aspects of its throughput, such as rate or
precise time of emission.  In such a case, the deterministic path is
end to end from application host to application host.

On the other end, the deterministic portion of a path may be a tunnel
between an ingress point and an egress router.  In any case, routers
and switches in between should not need to be aware of whether the
path is end to end or a tunnel.

While it is clear that DetNet does not aim to set up deterministic
paths over the global Internet, there is still a lack of clarity
regarding the limits of a domain where a deterministic path can be
set up.  These limits may depend on the technology that is used to
set the path up, whether it is centralized or distributed.

### 3.2.  Flow Characterization

Deterministic forwarding can only apply to flows with such
well-defined characteristics as periodicity and burstiness.  Before a
path can be established to serve them, the expression of those
characteristics, and how the network can serve them (for instance, in
shaping and forwarding operations), must be specified.

3.3.  Centralized Path Computation and Installation

   A centralized routing model, such as that provided with a Path
   Computation Element (PCE) (see [RFC4655]), enables global and
   per-flow optimizations.  This type of model is attractive, but a
   number of issues remain to be solved -- in particular:

   o  whether and how the path computation can be installed by

      *  an end device or

      *  a network management entity

      and

   o  how the path is set up -- either

      *  by installing state at each hop with a direct interaction
         between the forwarding device and the PCE or

      *  along a path by injecting a source-routed request at one end of
         the path, following classical Traffic Engineering (TE) models

   To enable a centralized model, DetNet should produce a description of
   the high-level interaction and data models to:

   o  report the topology and device capabilities to the central
      controller

   o  establish a direct interface between the centralized PCE and each
      device under its control in order to enable vertical signaling

   o  request a path setup for a new flow with particular
      characteristics over the service interface and control it through
      its life cycle

   o  provide support for life-cycle management for a path
      (instantiate/modify/update/delete)

   o  provide support for adaptability to cope with such various events
      as loss of a link

   o  expose the status of the path to the end devices (User-Network
      Interfaces (UNIs))

o  provide additional reliability through redundancy, particularly
   with Packet Replication, Elimination, and Ordering Functions
   (PREOF), where redundant paths may deliver packets out of order
   and PREOF may need to correct the ordering

o  indicate the flows and packet sequences in-band with the flows.
   This is needed for flows that require PREOF in order to isolate
   duplicates and reorder packets at the end of the sequence

## 3.4.  Distributed Path Setup

Whether a distributed alternative without a PCE can be valuable could
be studied as well.  Such an alternative could, for instance, build
upon Resource Reservation Protocol - TE (RSVP-TE) flows [RFC3209].
But the focus of the work should be to deliver the centralized
approach first.

To enable functionality similar to that of RSVP-TE, the following
steps would take place:

1.  Neighbors and their capabilities would be discovered and exposed
    to compute a path that would fit the DetNet constraints --
    typically those of latency, time precision, and resource
    availability.

2.  A constrained path would be calculated with an improved version
    of Constrained Shortest Path First (CSPF) that is aware of
    DetNet.

3.  The path may be installed using a control protocol such as
    RSVP-TE, extended to enable flow identification and install new
    per-hop behavior such as Packet Replication, Elimination, and
    Ordering, and to reserve physical resources for the flow.  In
    that case, traffic flows could be transported through an MPLS-TE
    tunnel, using the reserved resources for this flow at each hop.

## 3.5.  Duplicated Data Format

In some cases, the duplication and elimination of packets over
non-congruent paths are required to achieve a sufficiently high
delivery ratio to meet application needs.  In these cases, a small
number of packet formats and supporting protocols are required
(preferably just one of each) to serialize the packets of a DetNet
stream at one point in the network, replicate them at one or more
points in the network, and discard duplicates at one or more other
points in the network, including perhaps the destination host.  Using
an existing solution would be preferable to inventing a new one.

4.  Security Considerations

   Security in the context of Deterministic Networking has an added
   dimension; the time of delivery of a packet can be just as important
   as the contents of the packet itself.  A man-in-the-middle attack,
   for example, can impose and then systematically adjust additional
   delays into a link, and thus disrupt or subvert a real-time
   application without having to crack any encryption methods employed.
   See [RFC7384] for an exploration of this issue in a related context.

   Typical control networks today rely on complete physical isolation to
   prevent rogue access to network resources.  DetNet enables the
   virtualization of those networks over a converged IT/OT
   infrastructure.  Doing so, DetNet introduces an additional risk of
   flows interacting and interfering with one another as they share
   physical resources such as Ethernet trunks and the radio spectrum.
   The requirement is that there is no possible data leak from and into
   a deterministic flow.  Stated more generally, there is no possible
   influence whatsoever from the outside on a deterministic flow.  The
   expectation is that physical resources are effectively associated
   with a given flow at a given point in time.  In that model, the
   time-sharing of physical resources becomes transparent to the
   individual flows, as these flows have no clue regarding whether or
   not the resources are used by other flows at other times.

   The overall security of a deterministic system must cover:

   o  the protection of the signaling protocol

   o  the authentication and authorization of the controlling nodes,
      including plug-and-play participating end systems

   o  the identification and shaping of the flows

   o  the isolation of flows from leakage and other influences from any
      activity sharing physical resources

   The specific threats against DetNet are further discussed in
   [DetNet-Security].

5.  IANA Considerations

   This document has no IANA actions.

6.  Informative References

   [DetNet-Arch]
             Finn, N., Thubert, P., Varga, B., and J. Farkas,
             "Deterministic Networking Architecture", Work in
             Progress, draft-ietf-detnet-architecture-13, May 2019.

   [DetNet-Security]
             Mizrahi, T., Grossman, E., Ed., Hacker, A., Das, S.,
             Dowdell, J., Austad, H., Stanton, K., and N. Finn,
             "Deterministic Networking (DetNet) Security
             Considerations", Work in Progress,
             draft-ietf-detnet-security-04, March 2019.

   [IEEE-1588]
             IEEE, "IEEE Standard for a Precision Clock Synchronization
             Protocol for Networked Measurement and Control Systems",
             IEEE Standard 1588-2008, <https://standards.ieee.org/
             findstds/standard/1588-2008.html>.

   [IEEE-802.1TSNTG]
             IEEE Standards Association, "IEEE 802.1 Time-Sensitive
             Networking Task Group",
             <http://www.ieee802.org/1/pages/avbridges.html>.

   [IEEE-8021AS]
             IEEE, "IEEE Standard for Local and Metropolitan Area
             Networks - Timing and Synchronization for Time-Sensitive
             Applications in Bridged Local Area Networks",
             IEEE 802.1AS-2011,
             <http://www.ieee802.org/1/pages/802.1as.html>.

   [ISA95]   ANSI/ISA, "Enterprise-Control System Integration - Part 1:
             Models and Terminology", <https://www.isa.org/isa95/>.

   [PROFIBUS] IEC, "PROFIBUS Standard - DP Specification (IEC 61158
             Type 3)", <https://www.profibus.com/>.

   [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V.,
             and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP
             Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001,
             <https://www.rfc-editor.org/info/rfc3209>.

   [RFC4655] Farrel, A., Vasseur, J.-P., and J. Ash, "A Path
             Computation Element (PCE)-Based Architecture", RFC 4655,
             DOI 10.17487/RFC4655, August 2006,
             <https://www.rfc-editor.org/info/rfc4655>.

   [RFC7384]   Mizrahi, T., "Security Requirements of Time Protocols in
               Packet Switched Networks", RFC 7384, DOI 10.17487/RFC7384,
               October 2014, <https://www.rfc-editor.org/info/rfc7384>.

   [RFC8578]   Grossman, E., Ed., "Deterministic Networking Use Cases",
               RFC 8578, DOI 10.17487/RFC8578, May 2019,
               <https://www.rfc-editor.org/info/rfc8578>.

Acknowledgments

Authors' Addresses

   Norman Finn
   Huawei Technologies Co. Ltd
   3755 Avocado Blvd.
   PMB 436
   La Mesa, California  91941
   United States of America

   Phone: +1 925 980 6430
   Email: norman.finn@mail01.huawei.com


   Pascal Thubert
   Cisco Systems, Inc.
   Building D, 45 Allee des Ormes - BP1200
   Mougins - Sophia Antipolis  06254
   France

   Phone: +33 4 97 23 26 34
   Email: pthubert@cisco.com